

THE REVOLUTION WILL BE DIGITIZED: GENERATIVE AI, SYNTHETIC MEDIA, AND THE MEDIUM OF DISRUPTION

JON M. GARON*

OpenAI's text generation program ChatGPT and the text-to-image generators Stable Diffusion and Dall-E have broken records for early public adoption, capital investment, and a technological shift potentially more far-reaching than even the Internet itself. The broad category of generative AI has the potential to disrupt industry, art, and culture, both if done poorly and if done well. Despite significant problems with accuracy and deep concerns about the social and legal consequences of the premature adoption of these technologies, global multinational enterprises are moving these projects out of the test labs and into everyday use. This article provides a comprehensive, but introductory overview of the development of generative AI, the training methods used to produce artificially generated content, the industry opportunities for generative AI, and the legal considerations that enterprises adopting these technologies should consider involving intellectual property. After discussing the development and implementation of the technology, the article emphasizes the key concerns regarding copyright, trademark, and trade secret. The article also identifies areas in which the growth of generative AI will require new federal legislation to retain the balance of creativity and commercial development within intellectual property laws.

* Professor of Law, Nova Southeastern University Shepard Broad College of Law. An earlier version of this paper, "A Practical Introduction to Generative AI, Synthetic Media, and the Messages Found in the Latest Medium," was presented for the American Bar Association Business Law Section Spring Meeting and the Nova Southeastern University Learning and Educational Center. Special thanks to Cheryl Booth, Cheryl Burtzel, Nikola Datzov, Dan Garon, Stacy Blumberg Garon, Edward A. Morse, and Becka Rich, Susan Stephan, and Steven Teppler for their many helpful comments, insights, and reflections.

I. INTRODUCTION	141
II. THE EVOLUTION OF AI FROM FICTION TO REALITY	146
III. A NONTECHNICAL INTRODUCTION TO THE TECHNOLOGIES	153
IV. THE OPPORTUNITIES OF GENERATIVE AI.....	164
A. Image Generation.....	166
B. Text Generation.....	170
C. Audio, Including Music, Voice, and Sound.....	174
D. Code	179
E. Product Design and Drug Design	182
V. INTELLECTUAL PROPERTY ISSUES EMERGING FROM AI-GENERATED CONTENT	184
A. Copyright in the Dataset	184
B. Copyright by the End User and the Commercial Licensee....	194
C. Trade Secrets and Nonpublic Information.....	197
D. Trademarks	203
E. Rights of Publicity in the Data Set and the Output.....	211
F. Where Congress May Wish to Tread.....	215
VI. CONCLUSION	222

I. INTRODUCTION

Story—sacred and profane—is perhaps the main cohering force in human life. A society is composed of fractious people with different personalities, goals, and agendas. What connects us beyond our kinship ties? Story. . . . Story is the counterforce to social disorder, the tendency of things to fall apart. Story is the center without which the rest cannot hold.

— Jonathan Gottschall, *The Storytelling Animal: How Stories Make Us Human*

Mythology, in other words, is psychology misread as biography, history, and cosmology.

— Joseph Campbell, *The Hero with a Thousand Faces*

The question was simple: how does one explain to a 9-year-old what the James Webb Space Telescope has discovered? The answer should have been simple as well. But Google Bard got it wrong, and the embarrassment cost the company \$100 billion in market value.¹ That was the indelible lesson from the Google preview of Google Bard, the generative artificial intelligence (AI) program designed to use natural language to glean facts from across the Internet and provide interactive chat with users on whatever subjects might cross their minds.

Bard incorrectly stated that the James Webb telescope took the “very first pictures of a planet outside of our own solar system.”² Researchers at the Very Large Telescope in Chile are recognized for

¹ Jeran Wittenstein, *A Factual Error by Bard AI Chatbot Just Cost Google \$100 Billion*, TIME (Feb. 9, 2023, 12:36 PM), <https://time.com/6254226/alphabet-google-bard-100-billion-ai-error/> [<https://perma.cc/RV6Z-UH97>] (“Google parent Alphabet’s shares tumbled 7.7% on Wednesday after concerns surfaced about the competency of Bard, the ChatGPT rival it unveiled on Feb. 6. The selloff continued on Thursday The rout has erased about \$170 billion in market value.”).

² Sundar Pichai, *A Message From Our CEO: An Important Next Step on Our AI Journey*, GOOGLE (Feb. 6, 2023), <https://www.blog.google/technology/ai/bard-google-ai-search-updates/> [<https://perma.cc/852Y-SAV3>]; see Tristan Bove, *A Robot’s \$100 Billion Error: Alphabet Shares Tank After Its ChatGPT Rival Makes a Mistake in Its Very First Ad*, FORTUNE (Feb. 8, 2023, 3:42 PM), <https://fortune.com/2023/02/08/google-bard-ai-mistake-ad-stock-price-market-cap/> [<https://perma.cc/XK3R-2ZS5>].

having achieved this scientific first.³ The answer shook a stock market that had been betting on Alphabet having a technological answer ready to compete with OpenAI and the incorporation of ChatGPT technology into Bing and Office.

The stock market had been giddy with the prospect of Microsoft, Alphabet, Meta, and Nvidia (maker of the graphical processing units providing most of the hardware for next generation AI⁴) and the hundreds of companies trying to participate in the cultural phenomenon unleashed by the public launch of OpenAI's ChatGPT.⁵ "ChatGPT's wide range of uses, status as the fastest-growing consumer app in history, and potential to disrupt internet searches has pushed rivals to throw out the slow and cautious strategy that has dominated AI research for years."⁶

Using a variety of different training methods (which are discussed in section III), new and emerging generative AI systems can produce data that mimics the content of human creativity.⁷ These neural

³ See *Very Large Telescope*, EUR. S. OBSERVATORY, <https://www.eso.org/public/teles-instr/paranal-observatory/vlt/> [https://perma.cc/37F7-K4SG] ("The VLT has stimulated a new age of discoveries, with several notable scientific firsts, including the first image of an extrasolar planet (eso0428), tracking individual stars moving around the supermassive black hole at the centre of the Milky Way (eso0846), and observing the afterglow of the furthest known Gamma-Ray Burst."); see Bove, *supra* note 2 ("But the very first image of an exoplanet was captured by the Very Large Telescope, a ground-based array in Chile, in 2004 and confirmed as an exoplanet in 2005, according to NASA—long before James Webb's 2021 launch. James Webb, however, is being used to identify and catalog exoplanets.").

⁴ See Kif Leswing, *Meet the \$10,000 Nvidia Chip Powering the Race for A.I.*, CNBC (Feb. 23, 2023, 5:10 PM), <https://www.cnbc.com/2023/02/23/nvidias-a100-is-the-10000-chip-powering-the-race-for-ai-.html> [https://perma.cc/3RV5-EQ4S] ("Nvidia takes 95% of the market for graphics processors that can be used for machine learning, according to New Street Research. . . . AMD and Intel have competing graphics processors, and big cloud companies like Google and Amazon are developing and deploying their own chips specially designed for AI workloads.").

⁵ See Bern Elliott, *Why is ChatGPT Making Waves in the AI Market?*, GARTNER (Dec. 8, 2022), <https://www.gartner.com/en/newsroom/press-releases/2022-12-08-why-is-chatgpt-making-waves-in-the-ai-market> [https://perma.cc/SYA3-DS2H] ("Artificial intelligence (AI) research and deployment company OpenAI recently announced the official launch of ChatGPT, a new model for conversational AI. According to OpenAI, the dialogue provided by this platform makes it possible for ChatGPT to 'answer follow-up questions, admit its mistakes, challenge incorrect premises and reject inappropriate requests.'").

⁶ Bove, *supra* note 2.

⁷ See Rob Toews, *The Next Generation of Large Language Models*, FORBES (Feb. 7, 2023, 11:00 AM), <https://www.forbes.com/sites/robtoews/2023/02/07/the-next-generation-of-large-language-models/?sh=a6a71bd18dbc> [https://perma.cc/AQ2B-WMY9].

networks can generate content in the form of text, voice, pictures, videos, software, physical and molecular designs, audiovisual works combining these features, and more.⁸

As a story-telling medium, “ChatGPT has been used to create thousands of books that have been published on Amazon. Interestingly enough, the genres and topics covered by these AI-generated books are quite wide, ranging from self-help to cookbooks to romance novels.”⁹ Generative AI is very different from the “extractive AI” that drives search engines, mapping data, auto correct, and similar services. Generative AI systems combine and synthesize the known information to provide an answer of their own. The best can evaluate the information, combining the extractive functions with the generative ones. Trained on the data with which they are provided, they develop their own responses to the questions presented.

Generative AI is just one of many AI implementations, but the focus of this paper is limited to the potential for large networks to create new works as a result of their synthesis of the information they are provided. In some developing AI models, the generative AI service can be integrated with an “extractive” or “abstractive” process that produces a shortened or summarized version of texts or documents.¹⁰ “Extractive AI is all about taking existing information and using it to answer specific questions or generate new content, while generative AI is all about creating new information from scratch.”¹¹

Generative AI is fundamentally different from other forms of AI because it reflects a new form of automation in which the creative

⁸ *Id.*; see also Lee, *infra* note 87; Wang et al., *infra* note 93, at 588 (“It is now possible using GANs to generate photorealistic object images such as birds and faces, generate indoor or outdoor scenes, translate images from a source domain to the target domain, generate high-definition images from low-definition images, and so on.”).

⁹ ODSC Team, *AI Generated Books are Flooding Amazon Kindle Store*, ODSC (Feb. 23, 2023), <https://opendatascience.com/ai-generated-books-are-flooding-amazon-kindle-store/> [<https://perma.cc/Y4ED-26WV>]; see also Reuters, *These Authors Are Using ChatGPT to Write Books and Sell Them on Amazon*, N.Y. POST (Feb. 23, 2023, 6:56 PM), <https://nypost.com/2023/02/21/chatgpt-launches-boom-in-ai-written-e-books-on-amazon/> [<https://perma.cc/5DTF-7R9E>] (“There were over 200 e-books in Amazon’s Kindle store as of mid-February listing ChatGPT as an author or co-author, . . . [a]nd the number is rising daily. There is even a new sub-genre on Amazon: Books about using ChatGPT, written entirely by ChatGPT.”).

¹⁰ Stefano Ferilli et al., *A Similarity-Based Abstract Argumentation Approach to Extractive Text Summarization in AI*IA 2017: ADVANCES IN A.I.* 87 (2017).

¹¹ *Extractive AI vs Generative AI*, NOWIGENCE, <https://nowigence.com/blog/extractive-ai-vs-generative-ai/> [<https://perma.cc/QW22-D2GZ>] (“It’s like the difference between copy-pasting a Wikipedia article and writing a completely original one.”).

output upon which all societies depend may be generated by narrow AI systems that do not actually understand what they are producing.¹² The power of a generative AI to compete with humans to tell stories, make music, and create art will trigger significant upheavals across a multitude of industries.

At this stage, most of the commercial attention for AI innovation remains steered toward business rather than creative industries.¹³ AI systems are excellent at automating routine tasks, while ensuring that rote procedures are followed consistently time after time. The systems identify patterns in data that help identify customer interests and behaviors, increasing productivity and profitability. The systems also have the potential to augment the tasks that humans find tedious, increasing workforce satisfaction.¹⁴

Even though these new systems remain error prone, companies are jostling to be first with practical applications. In March 2023, Salesforce “introduced Einstein GPT, an A.I. system that it hopes will help salespeople, marketers, and customer service agents do their jobs more efficiently, as well as detailed plans to integrate the chatbots into the Slack messaging system.”¹⁵ Since then, Microsoft has aggressively

¹² See generally Oladeji M. Tiamiyu, *The Impending Battle for the Soul of ODR: Evolving Technologies and Ethical Factors Influencing the Field*, 23 CARDOZO J. CONFLICT RESOL. 75, 88 (2022) (“The role of AI in art helps to provide a framework for the possibilities of AI in ODR.”); Molly K. Land & Jay D. Aronson, *Human Rights and Technology: New Challenges for Justice and Accountability*, 16 ANN. REV. L. & SOC. SCI. 223 (2020); Gia Jung, *Do Androids Dream of Copyright?: Examining AI Copyright Ownership*, 35 BERKELEY TECH. L.J. 1151 (2020); Harry Surden, *Artificial Intelligence and Law: An Overview*, 35 GA. ST. U.L. REV. 1305 (2019); Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54 (2019); Frank Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, 87 GEO. WASH. L. REV. 1 (2019); Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 889 (2018).

¹³ Saarthak Bakshi, *Why Artificial Intelligence is Taking the World by Storm*, BUS. WORLD (June 29, 2019), <https://www.businessworld.in/article/Why-Artificial-Intelligence-Is-Taking-The-World-By-Storm/29-06-2019-172451/> [<https://perma.cc/DSJ6-A4B5>].

¹⁴ *Id.* (“Artificial Intelligence helps in directing and locating the entrepreneur to their potential customers, . . . maintain the several business activities like sales, the flow of cash, inventory and various other records . . . , assists in better problem-solving, . . . involves no errors and faults in its functions . . . [and] actually reduces the amount of pain that humans have to take . . .”).

¹⁵ Chris Morris, *Salesforce is Joining the ChatGPT Frenzy by Launching an OpenAI-powered Tool on Slack*, FORTUNE (Mar. 7, 2023, 11:02 AM), <https://fortune.com/2023/03/07/salesforce-artificial-intelligence-slack-einstein/> [<https://perma.cc/KV3F-KVNQ>].

expanded its use of OpenAI tools in Bing, Word, Excel, and coding in Github's Co-Pilot.¹⁶ Expedia has created a travel-agent-styled chat service to allow users to have conversations about their upcoming vacations.¹⁷ Cheggmate, a college assistance platform, has added ChatGPT to its service.¹⁸ These are just a few of the many examples.

Salesforce CEO Marc Benioff explained some of the uses for generative AI in Slack, the business messaging app, describing the impact of generative AI as transformative for commerce. "The world is experiencing one of the most profound technological shifts with the rise of real-time technologies and generative AI. . . . This comes at a pivotal moment as every company is focused on connecting with their customers in more intelligent, automated, and personalized ways."¹⁹

Salesforce is not alone. "Every industry that requires humans to create original work—from social media to gaming, advertising to architecture, coding to graphic design, product design to law, marketing to sales—is up for reinvention."²⁰ And yet, the problems have barely been identified and certainly not addressed.

Industry leaders likely hope that the rapid innovations mean that many of the problems will be resolved by the new technology. While the rapid iteration will resolve some of the challenges discussed below, it will only exacerbate others. The anticipated technical improvements may not help with the legal liabilities or the social upheaval. Moreover, in the long run, the potential for generative AI to create a disruptive force to labor markets, cultural institutions, and social norms will go well beyond the sales and marketing concerns being discussed in boardrooms today.

This article addresses the requirements to fulfill corporate accountability obligations with respect to laws and regulations around patent, copyright, trademark, trade secret, and publicity rights issues. A companion article will focus on products liability, privacy,

¹⁶ Bernard Marr, *10 Amazing Real-World Examples of How Companies Are Using ChatGPT in 2023*, FORBES (May 30, 2023, 3:03 AM), <https://www.forbes.com/sites/bernardmarr/2023/05/30/10-amazing-real-world-examples-of-how-companies-are-using-chatgpt-in-2023/> [https://perma.cc/XS95-7JEA].

¹⁷ *See id.*

¹⁸ *See id.*

¹⁹ Morris, *supra* note 15.

²⁰ Sonya Huang, Pat Grady & GPT-3, *Generative AI: A Creative New World*, SEQUOIA (Sept. 19, 2022), <https://www.sequoiacap.com/article/generative-ai-a-creative-new-world/> [https://perma.cc/8A7C-GPVT].

employment, and nondiscrimination.²¹ The article also provides a primer for those in business who are not familiar with the AI technologies and the fields impacted by the advances in these technologies.

II. THE EVOLUTION OF AI FROM FICTION TO REALITY

Although automata with clockwork mechanics had been used since the Middle Ages,²² the first “robot” appeared in modern literature with the 1920 play, *R.U.R.: Rossum’s Universal Robots*, written by Karel Čapek.²³ The term robot was a Czech word for forced labor, and the play depicted an early conception of the “singularity,” which is the point at which the intelligent, self-aware machines come to dominate humanity rather than serve as its labor force.²⁴ Perhaps even better known is the 1927 film, *Metropolis*, by Fritz Lang, that explored similar themes.²⁵ Even before the depiction of self-aware robots in film, “elaborate mechanical inventions” were creating music and stories.²⁶ “John Clark’s The Eureka (1845) . . . could produce Latin verses with a pull of a lever. . . . Because of the strict rules of Latin hexameter, this wooden machine was capable of flawlessly randomizing words and arranging them in the plausible output.”²⁷

The fictional development of robots was combined with massive artificial intelligence and popularized by Isaac Asimov, who wrote more than three dozen books involving self-aware robots, as well as a dozen stories featuring “Multivac,” a massive government-run answer

²¹ See Jon M. Garon, *An AI’s Picture Paints a Thousand Lies: Designating Responsibility for Visual Libel*, 3 UCLA J. FREE SPEECH L. 425 (2023).

²² See Jessica Riskin, *Machines in the Garden*, 1 REPUBLICS OF LETTERS 24 (2010), <https://arcade.stanford.edu/rofl/machines-garden> [<https://perma.cc/RG7U-G3JR>].

²³ *R.U.R.*, ENCYCLOPEDIA BRITANNICA, <https://www.britannica.com/topic/RUR> [<https://perma.cc/TCT7-KCFR>].

²⁴ *Id.*

²⁵ See *Metropolis*, IMDB, <https://www.imdb.com/title/tt0017136/> [<https://perma.cc/WS5V-K7P8>].

²⁶ Vladimir Todorović, *Wandering Machines, Narrativity in Generative Art*, 11 CITAR J. 50 (2019), <https://journals.ucp.pt/index.php/jsta/article/view/7334/7114> (listing these examples: JACQUES VAUCONSON’S FLUTE PLAYER (1730s), JAQUET DROZ’S AUTOMATA (1768–1774), JOSEPH FABER’S TALKING MACHINE EUPHONIA (1845) & JOHN CLARK’S THE EUREKA (1845)).

²⁷ *Id.* at 50–51.

machine that could understand natural language queries and provide thoughtful answers.²⁸

The fictional depictions often frame the popular reference to the emergence of these technologies. John McCarthy is attributed with coining the academic term “artificial intelligence” during a 1956 Dartmouth College conference dedicated to capturing human-like thought processes by computers.²⁹

Although work has been underway for decades to better understand ways to mirror the processes of human thinking and decision-making, the past five years have seen an explosion of AI adoption. “Artificial intelligence is driving important developments in technology, from controlling autonomous vehicles, to developing medical diagnoses, to combating climate change.”³⁰ Each success generates new research. “The total number of AI publications grew from 162,444 in 2010 to 334,497 in 2021, according to the 2022 AI Index Report. In 2021, a PubMed search of papers with the keyword ‘deep learning’ returned 14,685 citations, up from 107 papers in 2010.”³¹

The rapid growth of AI includes experiments in self-driving vehicles,³² healthcare,³³ facial recognition to unlock a cell phone,³⁴ and uses in the production of music,³⁵ digital images,³⁶ deepfake videos, and

²⁸ See *Multivac*, WIKIPEDIA, <https://en.wikipedia.org/wiki/Multivac> [<https://perma.cc/6U2W-7K7Y>] (identifying the Asimov stories *Franchise* and *The Last Question* as framing the exploration and expansion of the technology).

²⁹ See HERBERT L. ROITBLAT, *ALGORITHMS ARE NOT ENOUGH: CREATING GENERAL ARTIFICIAL INTELLIGENCE 1* (2020).

³⁰ Rose Acoraci Zeck, *Analysis: Patents Forecast Widespread Reach of AI Tech in 2023*, BLOOMBERG L. (Nov. 13, 2022, 9:01 PM), <https://news.bloomberglaw.com/bloomberg-law-analysis/analysis-patents-forecast-widespread-reach-of-ai-tech-in-2023> [<https://perma.cc/7L7D-DNJ2>].

³¹ Gil Press, *Recent Anecdotal Evidence of Practical AI*, FORBES (May 30, 2022, 9:00 AM), <https://www.forbes.com/sites/gilpress/2022/05/30/recent-anecdotal-evidence-of-practical-ai/?sh=3290d5121aee> [<https://perma.cc/JVD4-8VH6>].

³² Claire D, *Top 13 Examples of Artificial Intelligence in Daily Life*, DIGITALOLOGY (Aug. 19, 2020), <https://blog.digitalogy.co/best-examples-of-artificial-intelligence-in-everyday-life/> [<https://perma.cc/WF89-MLRA>].

³³ West & Allen, *infra* note 68 (discussing Merantix, a German company that analyzes Computer Tomography (CT) images using deep learning to identify tumors).

³⁴ See Claire D, *supra* note 32.

³⁵ Computer Music, *A Short History of AI in Music Production*, MUSICRADAR (June 14, 2022), <https://www.musicradar.com/news/the-history-of-ai-in-music-production> [<https://perma.cc/4UQ7-YM2L>].

³⁶ Brian Jarvis, *How Artificial Intelligence Has Changed Photo Editing*, DIGIT. PHOTOGRAPHY SCH., <https://digital-photography-school.com/artificial-intelligence->

much more.³⁷ AI even plays a substantial role in national defense. Through its Project Maven, the American military is deploying AI “to sift through the massive troves of data and video captured by surveillance and then alert human analysts of patterns or when there is abnormal or suspicious activity.”³⁸

The expanding range of applications mirrors the increase in innovation and investment.

The global AI market was valued at nearly \$59.7 billion in 2021, and is estimated to reach \$422.4 billion by 2028. . . .

The number of patents issued by the US Patent & Trademark Office for AI technologies has surged over the past five years. During this timeframe, the number of AI-related patents issued has increased from 3,267 in

changed-photo-editing/ [https://perma.cc/QHH6-7G5D] (“One of the hottest topics in the imaging community is the power of AI to take low-quality images and upscale them into high-resolution files.” Discussing that AI also automates masking, sky replacement, color adjustment and other image attributes.)

³⁷ See Claire D, *supra* note 32; see also Thomas H. Davenport & Rajeev Ronanki, *Artificial Intelligence for the Real World*, HARV. BUS. REV. (Jan.–Feb. 2018), <https://hbr.org/2018/01/artificial-intelligence-for-the-real-world> [https://perma.cc/YG7N-XAQB].

Of the 152 projects [James Wheaton and Andrew Nguyen] studied, the most common type was the automation of digital and physical tasks—typically back-office administrative and financial activities—using robotic process automation technologies. RPA is more advanced than earlier business-process automation tools, because the “robots” (that is, code on a server) act like a human inputting and consuming information from multiple IT systems. Tasks include:

- transferring data from e-mail and call center systems into systems of record—for example, updating customer files with address changes or service additions;
- replacing lost credit or ATM cards, reaching into multiple systems to update records and handle customer communications;
- reconciling failures to charge for services across billing systems by extracting information from multiple document types; and
- “reading” legal and contractual documents to extract provisions using natural language processing.

³⁸ West & Allen, *infra* note 68 (quoting Christian Davenport, *Future Wars May Depend as Much on Algorithms as on Ammunition, Report Says*, WASH. POST (Dec. 3, 2017)).

2017 to 18,753 in 2021. As of Nov. 2, [2022] the USPTO has issued 15,992 AI-related patents³⁹

Although patents for AI inventions are growing at a rapid pace, this does not mean that patents generated by AI are having the same success. “Under the legal systems of the UK, US, China, Germany and many others, however, only a human can be recognised as an inventor.”⁴⁰ The premise is being tested globally by Stephen Thaler, the creator of DABUS, which stands for “Device for Autonomous Bootstrapping of Unified Sentience.”⁴¹ DABUS “has so far conceived of two inventions: a ‘fractal container’, designed to allow for coupling of multiple containers, improve grip, and improve heat transfer; and a ‘neural flame’, a flashing light designed to emit a uniquely-identifiable light signal that attracts enhanced human attention.”⁴²

The argument that an AI device could obtain patent protection was dismissed as inconsistent with the plain language of the statute and the meaning of the word “individual” in the Patent Act.⁴³ “Here, there is no ambiguity: The Patent Act requires that inventors must be natural persons; that is, human beings.”⁴⁴

Although litigation and applications continue around the globe, only South Africa has allowed the registration of an AI-authored invention, and South Africa does not review the merits of a patent

³⁹ Zeck, *supra* note 30.

⁴⁰ Tom Dines, *A Patent Predicament: Who Owns an AI-generated Invention?*, FIN. TIMES (Oct. 6, 2019), <https://www.ft.com/content/84677ec8-be73-11e9-9381-78bab8a70848> [<https://perma.cc/5R6R-R27M>].

⁴¹ See Sam Tobin, *UK Supreme Court Hears Landmark Patent Case Over AI ‘Inventor’*, REUTERS (Mar. 2, 2023, 9:16 AM), <https://www.reuters.com/technology/uk-supreme-court-hears-landmark-patent-case-over-ai-inventor-2023-03-02/> [<https://perma.cc/HN8W-WN43>]; Blake Brittain, *U.S. Scientist Hits Another Dead End in Patent Case Over AI ‘Inventor’*, REUTERS (Oct. 20, 2022, 3:58 PM), <https://www.reuters.com/legal/litigation/us-scientist-hits-another-dead-end-patent-case-over-ai-inventor-2022-10-20/> [<https://perma.cc/3H9P-PZTW>].

⁴² Sanjaya Mendis, Rachel Wasserman & Chuck Rothman, *AI and Patent Law: Can AI Be an ‘Inventor’?*, LEXOLOGY (May 31, 2022), <https://www.lexology.com/library/detail.aspx?g=648f10e3-c937-4604-bf89-c21e66b4d669> [<https://perma.cc/8SJ4-NHBN>].

⁴³ See *Thaler v. Hirshfeld*, 558 F. Supp. 3d 238, 246 (E.D. Va. 2021), *aff’d sub nom. Thaler v. Vidal*, 43 F.4th 1207 (Fed. Cir. 2022); see also *Mohamad v. Palestinian Auth.*, 566 U.S. 449, 453–54 (2012) (“The ordinary meaning of the word [individual], fortified by its statutory context,” referred to a “natural person”).

⁴⁴ *Thaler v. Vidal*, 43 F.4th 1207, 1210 (Fed. Cir. 2022).

application at the time of registration, only its procedural requirements.⁴⁵

The same outcome has occurred in the U.S. for copyrights of works created by non-human parties.⁴⁶ “To qualify for copyright protection, a work must be original to the author.”⁴⁷ In addition, copyright only subsists if the work “possesses at least some minimal degree of creativity.”⁴⁸

Among the recent attempts to claim non-human copyright ownership was the claim brought by PETA on behalf of Naruto, a seven-year-old crested macaque from the island of Sulawesi, Indonesia.⁴⁹ Naruto “borrowed” a camera from wildlife photographer David Slater and took a series of exceptional selfies with the camera. Slater eventually recovered the camera and published the images without Naruto’s permission. But since Naruto had no right to make a copyrightable image,⁵⁰ those works fell into the public domain and were free for Slater or anyone to publish.⁵¹

In cases of alleged divine authorship, the courts distinguish between the uncopyrightable text authored by the non-humans from the various contributions provided by their human editors, including textual questions posed as well as the order, selection, and arrangement of the text provided.⁵² The cases involving divine authorship also point to important distinctions between the various aspects of copyright protection. A recipe may not be subject to copyright protection because

⁴⁵ See Kingsley Egbonu, *The Latest News on the DABUS Patent Case*, IP STARS (Feb. 28, 2023), <https://www.ipstars.com/NewsAndAnalysis/The-latest-news-on-the-DABUS-patent-case/Index/7366> [<https://perma.cc/9R6A-ZVE6>].

⁴⁶ But this is not without controversy. See, e.g., Nina I. Brown, *Artificial Authors: A Case for Copyright in Computer-Generated Works*, 20 COLUM. SCI. & TECH. L. REV. 1, 20 (2018) (“The courts have not yet been confronted with determining whether a computer can meet that definition, or whether an author must be human. Importantly, the Constitution does not define authors as human.”); Arthur R. Miller, *Copyright Protection for Computer Programs, Databases, and Computer-Generated Works: Is Anything New Since CONTU?*, 106 HARV. L. REV. 977 (1993); *Urantia Found. v. Maaherra*, 114 F.3d 955, 958 (9th Cir. 1997).

⁴⁷ *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co., Inc.*, 499 U.S. 340, 345 (1991).

⁴⁸ *Id.*

⁴⁹ *Naruto v. Slater*, 888 F.3d 418, 420 (9th Cir. 2018) (“[A]ll animals, since they are not human—lack[] statutory standing under the Copyright Act.”).

⁵⁰ *Id.*

⁵¹ *Id.*

⁵² See *Urantia Found. v. Maaherra*, 114 F.3d 955, 959 (9th Cir. 1997); *Oliver v. Saint Germain Found.*, 41 F. Supp. 296 (S.D. Cal. 1941).

it is a set of instructions.⁵³ “In contrast, a recipe that creatively explains or depicts how or why to perform a particular activity may be copyrightable [including] any photographs or illustrations that are owned by the applicant.”⁵⁴

The Copyright Office has used this distinction to extend copyright protection to Kristina Kashtanova, author of *Zarya of the Dawn*, for her contribution to her graphic novel, while withholding copyright protection for the images generated through Midjourney.⁵⁵ In the reevaluation later of February 21, 2023, the Copyright Office explained the limitation on the registration for the graphic novel:

We conclude that Ms. Kashtanova is the author of the Work’s text as well as the selection, coordination, and arrangement of the Work’s written and visual elements. That authorship is protected by copyright. However, . . . the images in the Work that were generated by the Midjourney technology are not the product of human authorship.⁵⁶

Relying on the history of copyright cases discerning the modicum of creativity required by authors, on February 1, 2022, the Copyright Review Board denied Steven Thaler’s attempt to register his two-dimensional artwork entitled “A Recent Entrance to Paradise” as lacking the requisite originality required under copyright caselaw.⁵⁷ The Copyright Review Board explained that “copyright law only protects ‘the fruits of intellectual labor’ that ‘are founded in the creative powers of the [human] mind.’”⁵⁸ The Copyright Board quoted a long history of

⁵³ See 17 U.S.C. § 102(b) (2023); U.S. COPYRIGHT OFF., WORKS NOT PROTECTED BY COPYRIGHT (CIRCULAR 33) (Mar. 2021) (“A recipe is a statement of the ingredients and procedure required for making a dish of food. A mere listing of ingredients or contents, or a simple set of directions, is uncopyrightable.”).

⁵⁴ U.S. COPYRIGHT OFF., *supra* note 53.

⁵⁵ See Franklin Graves, *U.S. Copyright Office Backtracks on Registration of Partially AI-Generated Work*, IP WATCHDOG (Nov. 1, 2022), <https://ipwatchdog.com/2022/11/01/us-copyright-office-backtracks-registration-partially-ai-generated-work/id=152451/> [https://perma.cc/L4FR-S7XF].

⁵⁶ Decision on Copyright of *Zarya of the Dawn* (Registration #VAu001480196), Letter from U.S. Copyright Office to Van Lindberg (Feb. 21, 2023), <https://www.copyright.gov/docs/zarya-of-the-dawn.pdf>.

⁵⁷ See Second Request for Reconsideration for Refusal to Register *A Recent Entrance to Paradise* (Correspondence ID 1-3ZPC6C3; SR # 1-7100387071), Letter from U.S. Copyright Office to Ryan Abbott (Feb. 14, 2022), <https://www.copyright.gov/rulings-filings/review-board/docs/a-recent-entrance-to-paradise.pdf>.

⁵⁸ *Id.* at 3 (quoting Trade-Mark Cases, 100 U.S. 82, 94 (1879)).

cases using human pronouns to reinforce the longstanding interpretation that authorship is an act of human creativity and this creativity is the essential minimal requirement for copyright protection.⁵⁹

Applying these cases to generative AI, the production of code, images, videos, or text would not be protected by copyright, but if the questions used to generate the text were of sufficient detail to merit copyright protection, then those may be subject to copyright protection, a question left intentionally unanswered by the Copyright Review Board.⁶⁰ More practically, selections, editing, and arrangements of AI-generated content would also be protected by thin copyrights extending only to the human additions to those works. One can speculate that if this field grows, the copyright registration process will increasingly require applicants to disclaim all content that was AI-generated as a step in the registration process.

In March 2023, the Copyright Office made its next effort to add clarity to the parameters that distinguish copyrightable works from content that does not have the modicum of human creativity or is otherwise outside the scope of copyright protection under §102(b). In publishing registration guidance, the Copyright Office reiterated the longstanding requirement of human authorship.⁶¹ Nonetheless, just as it recognized the human contribution of Kristina Kashtanova, the Office recognized that “a human may select or arrange AI-generated material in a sufficiently creative way that ‘the resulting work as a whole constitutes an original work of authorship.’ Or an artist may modify material originally generated by AI technology to such a degree that the modifications meet the standard for copyright protection.”⁶²

Despite the lack of patent and copyright protection for AI-generated content and inventions, AI has already become deeply embedded in products most people use daily from route predictions in mapping apps to automated customer service. Almost every author has used AI in other forms, such as auto-correction of text, color correction

⁵⁹ *Id.* at 4 (quoting *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53, 56–59 (1884) (explaining that copyright extends “the exclusive right of a man to the production of his own genius or intellect . . .”)).

⁶⁰ *See id.* at 3 n.3 (“Because Thaler has not raised this as a basis for registration, the Board does not need to determine under what circumstances human involvement in the creation of machine-generated works would meet the statutory criteria for copyright protection.”).

⁶¹ Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, 88 Fed. Reg. 16190 (Mar. 16, 2023) (citing *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53, 56 (1884)).

⁶² *Id.* (quoting U.S. COPYRIGHT OFFICE, *Compendium of U.S. Copyright Office Practices* § 507.1 (3d ed. 2021)).

of images, or auto-tune corrections for pitch. As such, the Office directive that “AI-generated content that is more than *de minimis* should be explicitly excluded from the application[,]”⁶³ may lead to more confusion than clarification. The Office made the situation worse because it fails to define AI or its *de minimis* standard while at the same time requiring that applicants and registrants alike “correct the public record by submitting a supplementary registration.”⁶⁴ At the moment, it is unlikely that the Office wants a reminder that Word’s auto-suggestion feature may have supplied word choices to a writer, since presumably the writer’s choice to accept those word choices remains authorial in operation. But as the auto-suggestion features grow from words to phrases to more, the Office’s registration disclaimers will undoubtedly need updating.

In many fields, such as visual works, voice and audio output, and design suggestions, the output of generative AI is already exceeding that which was once within the realm of only science fiction. The success suggests that perhaps AI does not need intellectual property protection to enhance these services.

At the same time, questions continue to be raised about the quality of the output by generative AI systems in factual contexts, the ability of these systems to adequately address legal and ethical requirements for the use of their output, and legal concerns that the process of training these systems and the content generated does so in a manner that respects the intellectual property rights of third parties. These questions, more than the questions of authorship and inventorship, will define the growth of generative AI in commercial enterprises.

III. A NONTECHNICAL INTRODUCTION TO THE TECHNOLOGIES

For purposes of understanding the scope of the technologies at the heart of AI, it is helpful to have some construction of the meaning of “intelligence.”⁶⁵ Two often-used quotes help capture the broad outline of the construct. First, intelligence can be considered as “an

⁶³ *Id.*

⁶⁴ *Id.*

⁶⁵ See Stephan De Spiegeleire, Matthijs Maas & Tim Sweijs, *What is Artificial Intelligence? Artificial Intelligence and the Future of Defense*, HAGUE CENTRE FOR STRATEGIC STUD. 25, 26 (2017), <http://www.jstor.com/stable/resrep12564.7>.

agent's ability to achieve goals in a wide range of environments."⁶⁶ A second construction is slightly more directive: "Intelligence is the computational part of the ability to achieve goals in the world."⁶⁷

Under these definitions, the purpose of intelligence is to be able both to understand the external world and apply that understanding to operations. Those operations can range from answering questions to driving a plane or automobile in a real-world environment. An automaton that can move in a pre-determined manner does not meet either of these criteria. The automaton does not understand the inputs to select appropriate choices and it is undertaking its operations following pre-defined behaviors. "Artificial intelligence algorithms are designed to make decisions, often using real-time data. They are unlike passive machines that are capable only of mechanical or predetermined responses."⁶⁸

In 2005, Ray Kurzweil wrote that "[a]rtificial intelligence permeates our economy."⁶⁹ The AI pervasive at the beginning of this century is described as "narrow" because the focus is on a particular task.⁷⁰ Narrow AI has made headlines and provided the underpinning for many common tools and services. "In 1997, Garry Kasparov, head in hands, lost a chess match to IBM's Deep Blue. Almost exactly 20 years later, Go champion Ke Jie was defeated by the AI company DeepMind's AlphaGo Master."⁷¹ In contrast to these narrow forms of AI, general or strong AI is a machine or system that demonstrates "the full range of human intelligence"⁷² or meets "the full range of human performance across any task."⁷³ Just as narrow AI can exceed humans in its accomplishment of certain goals, the transition to general AI may

⁶⁶ *Id.* (quoting Shane Legg & Marcus Hutter, *A Collection of Definitions of Intelligence*, ARXIV (June 25, 2007), <http://arxiv.org/abs/0706.3639> [<https://perma.cc/UHD6-RVRX>]).

⁶⁷ *Id.* (quoting John McCarthy, *What is Artificial Intelligence: Basic Questions*, STAN. UNIV. FORMAL REASONING GRP. (Nov. 12, 2007), <http://www-formal.stanford.edu/jmc/whatisai/node1.html> [<https://perma.cc/L7KN-LCFB>]).

⁶⁸ Darrell M. West & John R. Allen, *How Artificial Intelligence is Transforming the World*, BROOKINGS (Apr. 24, 2018), <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/> [<https://perma.cc/7F3Q-TGK7>].

⁶⁹ Ray Kurzweil, *Long Live AI*, FORBES (Aug. 15, 2005), <https://www.forbes.com/forbes/2005/0815/030.html?sh=1991ab7d7e8f> [<https://perma.cc/X4DH-AMLY>].

⁷⁰ *Id.*; see also Spiegeleire et al., *supra* note 65 (describing the task-specific AI as "narrow" or "weak" in contrast to general AI or strong AI).

⁷¹ Henry Shevlin et al., *The Limits of Machine Intelligence*, EMBO REPS. (Sept. 18, 2019), <https://doi.org/10.15252/embr.201949177>.

⁷² Kurzweil, *supra* note 69.

⁷³ Spiegeleire et al., *supra* note 65.

bring with it the ability of the machine intelligence to exceed human capacity, and for this stage of development, many researchers use the label “artificial superintelligence.”⁷⁴

In its attempt to develop regulations surrounding the adoption of AI within the European Union (EU), the European Commission created “AI Watch, the European Commission knowledge service to monitor the development, uptake and impact of Artificial Intelligence (AI) for Europe.”⁷⁵ AI Watch published a comprehensive review to help better define the range of technologies involved with AI. That study provided a more operational understanding of AI:

The study of the definitions found in literature leads us to identify four characteristics that are commonly mentioned in AI: i) perception of the environment and real-world complexity, ii) information processing: collecting and interpreting inputs, iii) decision making, including reasoning, learning and taking actions; and iv) achievement of pre-defined goals.⁷⁶

Using these four criteria, AI Watch then proposes this definition for AI:

Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.⁷⁷

The AI Watch definition highlights another of the important taxonomies involved with AI. Different technologies approach their

⁷⁴ *Id.*

⁷⁵ Sofia Samoili et al., *AI Watch. Defining Artificial Intelligence*, EUR 30117 EN, JRC TECH. REPS. 1 (2020).

⁷⁶ *Id.* at 4.

⁷⁷ *Id.*

problem solving using very different strategies. AI is merely a catchphrase for a range of various technologies involving the analysis of informational inputs that result in the system being able to achieve a stated goal.⁷⁸ Much like the range of strategies for training humans and animals, there are a range of techniques to address how the data is provided to the computer systems and how that data is processed.

Traditional computers follow the logic of decision trees provided by the system programmers. Jacquard knitting machines reproduce complex weaving patterns based on the information programmed into the machines, and early mainframe computers with punch card inputs used similar techniques to analyze and sort the data that was coded or structured into those machines. Improvements in the input devices have led to massive increases in the data available to the machine from which to glean patterns.

Since the 1950s, these systems have been known as “machine learning systems.”⁷⁹ The rules of a game such as checkers or blackjack were trained into the computer, and by running millions of simulations, the computer could learn to predict the best next choice in each turn of the game.⁸⁰ The rules and options were programmed into the computer using labeled datasets. Then algorithms or weighted math formulae would be used to iteratively improve the performance of the computer’s

⁷⁸ See Nikola Datzov, *The Role of Patent (In)Eligibility in Promoting Artificial Intelligence Innovation*, 92 UMKC L. REV. 2 (2023) (explaining that although “AI” is nearly ubiquitous, it has no agreed upon definition); see also Tabrez Y. Ebrahim, *Data-Centric Technologies: Patent and Copyright Doctrinal Disruptions*, 43 NOVA L. REV. 287, 291 (2019).

⁷⁹ See *What Is Machine Learning?*, IBM, <https://www.ibm.com/topics/machine-learning> [<https://perma.cc/FY7D-9RMG>] (“Arthur Samuel, is credited for coining the term, ‘machine learning’ [in the 1950s]”); *The IBM 700 Series: Computing Comes to Business*, IBM, <https://www.ibm.com/ibm/history/ibm100/us/en/icons/ibm700series/impacts/> [<https://perma.cc/4K2K-F6X7>] (“On February 24, 1956, Arthur Samuel’s Checkers program, which was developed for play on the IBM 701, was demonstrated to the public on television. In 1962, self-proclaimed checkers master Robert Nealey played the game on an IBM 7094 computer.”); see also *What Is Machine Learning (ML)?*, BERKELEY SCH. OF INFO. (June 26, 2020), <https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/> [<https://perma.cc/F7RS-QYH3>].

⁸⁰ See Julianna Delua, *Supervised vs. Unsupervised Learning: What’s the Difference?*, IBM (Mar. 12, 2021), <https://www.ibm.com/blog/supervised-vs-unsupervised-learning/> [<https://perma.cc/622L-K7NJ>] (“In supervised learning, the algorithm “learns” from the training dataset by iteratively making predictions on the data and adjusting for the correct answer.”).

selection of options for its choices.⁸¹ This process was later labeled “supervised learning”⁸² as alternative systems were developed called “unsupervised learning” that eliminated the need for labeling or quantifying the datasets. “Unsupervised learning uses machine learning algorithms to analyze and cluster unlabeled data sets. These algorithms discover hidden patterns in data without the need for human intervention.”⁸³ “Through machine learning, practitioners develop artificial intelligence through models that can ‘learn’ from data patterns without human direction. The unmanageably huge volume and complexity of data (unmanageable by humans, anyway) being generated has increased the potential of machine learning, as well as the need for it.”⁸⁴

⁸¹ See, e.g., Harry Surden, *Artificial Intelligence and Law: An Overview*, 35 GA. ST. U. L. REV. 1305 (2019).

⁸² See Delua, *supra* note 80.

Supervised learning can be separated into two types of problems when data mining: classification and regression:

- Classification problems use an algorithm to accurately assign test data into specific categories, such as separating apples from oranges. Or, in the real world, supervised learning algorithms can be used to classify spam in a separate folder from your inbox. Linear classifiers, support vector machines, decision trees and random forest are all common types of classification algorithms.
- Regression is another type of supervised learning method that uses an algorithm to understand the relationship between dependent and independent variables. Regression models are helpful for predicting numerical values based on different data points, such as sales revenue projections for a given business. Some popular regression algorithms are linear regression, logistic regression and polynomial regression.

⁸³ *Id.*

Unsupervised learning models are used for three main tasks: clustering, association and dimensionality reduction: **Clustering** is a data mining technique for grouping unlabeled data based on their similarities or differences. . . . **Association** . . . uses different rules to find relationships between variables in a given dataset. . . . **Dimensionality reduction** is a learning technique used when the number of features (or dimensions) in a given dataset is too high. It reduces the number of data inputs to a manageable size while also preserving the data integrity.

⁸⁴ *What Is Generative AI?*, MCKINSEY & CO. (Jan. 19, 2023), <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai#/> [<https://perma.cc/C83D-3SDB>].

A subset of the machine learning techniques is described as neural networks. “Neural networks, or artificial neural networks (ANNs), are comprised of node layers, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, connects to another and has an associated weight and threshold.”⁸⁵ A basic neural network has only three node layers. If the network has more, then it earns the classification as a “deep learning” network.⁸⁶ Changes to the weights of each layer of the node inform and influence the other layers, quickly generating highly complex decision trees.

ChatGPT employs a large language model (LLM), a “deep learning algorithm that can recognize, summarize, translate, predict, and generate text and other forms of content based on knowledge gained from massive datasets.”⁸⁷ In fact, GPT, or “Generative (G) Pre-trained (P) Transformer (T),”⁸⁸ utilizes such an LLM system. The LLM relies on a neural network in which changes to the weights of decisions in each node carry forward to influence the weighting on decisions in subsequent nodes.⁸⁹ The weighting shifts may be very subtle, and the scale of the neural network can be vast, leading the artificial intelligence system to mimic a sizable fraction of intelligent thought.⁹⁰

Called a transformer model, the most sophisticated of the neural networks iteratively adjusts the weighting among the nodes. “A transformer model is a neural network that learns context and thus meaning by tracking relationships in sequential data like the words in this sentence.”⁹¹

Recently, the focus on AI has shifted from the manner in which the systems learn data to the ways in which the learning can be put to use. In a shift away from moving a chess piece or sorting spam email

⁸⁵ *What is Machine Learning?*, *supra* note 79.

⁸⁶ *Id.*

⁸⁷ Angie Lee, *What Are Large Language Models Used For?*, NVIDIA (Jan. 26, 2023), <https://blogs.nvidia.com/blog/2023/01/26/what-are-large-language-models-used-for/> [<https://perma.cc/S2EQ-64CW>].

⁸⁸ Kevin Pocock, *What Does ChatGPT Stand For?*, PCGUIDE (Feb. 28, 2023), <https://www.pcguides.com/apps/what-does-chatgpt-stand-for/> [<https://perma.cc/V6ZP-EE97>].

⁸⁹ Lee, *supra* note 87.

⁹⁰ Rick Merritt, *Beyond Words: Large Language Models Expand AI's Horizon*, NVIDIA (Oct. 10, 2022), <https://blogs.nvidia.com/blog/2022/10/10/llms-ai-horizon/> [<https://perma.cc/X39C-Y865>].

⁹¹ Rick Merritt, *What Is a Transformer Model?*, NVIDIA (Mar. 25, 2022), <https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model/> [<https://perma.cc/6SW2-5HDM>].

from priority email, the focus has been to provide new content, in the form of music, art, videos, and text.

Generative AI refers to artificial intelligence that can generate novel content, rather than simply analyzing or acting on existing data. Generative AI models produce text and images: blog posts, program code, poetry, and artwork. The software uses complex machine learning models to predict the next word based on previous word sequences, or the next image based on words describing previous images. In the shorter term, we see generative AI used to create marketing content, generate code, and in conversational applications such as chatbots.⁹²

LLM is just one of the key training models. Another is the generative adversarial network (GAN), which is used more heavily to train systems to discriminate as regarding the meaning and context of a visual image, again using the deep neural networks to iteratively learn how to map the visual world interpreted through the network.⁹³

One subset of generative AI began making waves five years earlier. In late 2017 and 2018, FakeApp was released to enable users to swap the facial images of one person on a video of someone else.⁹⁴ FakeApp substituted the facial texture or visual features of a person's face onto that of another video, creating a tool to map famous actors' images onto pornographic videos.⁹⁵ Similar faked videos were also used to create simulated revenge porn to embarrass, harass, and extort women.⁹⁶ "As of 2019, 96 percent of deepfakes on the internet were

⁹² Danica Lo, *AI Is Having a Moment—Here's How Businesses Can Lean In*, FAST CO. (Dec. 18, 2022), <https://www.fastcompany.com/90826178/generative-ai> [<https://perma.cc/3KA3-F2KU>] (interviewing Babson College Professor Thomas Davenport and Nitin Mittal, head of U.S. artificial intelligence growth at Deloitte).

⁹³ See Kunfeng Wang et al., *Generative Adversarial Networks: Introduction and Outlook*, 4 IEEE/CAA J. AUTOMATICA SINICA 588 (Oct. 2017).

⁹⁴ See Edvinas Meskys et al., *Regulating Deep Fakes: Legal and Ethical Considerations*, 15 J. INTELL. PROP. L. & PRAC. 24 (2019).

⁹⁵ *Id.*

⁹⁶ See Tatum Hunter, *AI Porn is Easy to Make Now. For Women, That's a Nightmare*, WASHINGTON POST (Feb. 13, 2023), <https://www.washingtonpost.com/technology/2023/02/13/ai-porn-deepfakes-women-consent/> [<https://perma.cc/7G8S-NRFT>] ("[A] person made esports star 'QTCinderella's likeness in computer-generated porn. 'For every person saying it's not a big deal, you don't know how it feels to see a picture of yourself doing things you've never done being sent to your family,' QTCinderella said in a live-streamed video.'").

pornography, according to an analysis by AI firm DeepTrace Technologies, and virtually all pornographic deepfakes depicted women.”⁹⁷

Deepfakes are a subset of generative AI created using GAN model in which two networks are trained together.⁹⁸ Then the “generator tries to generate images that would fool the discriminator into believing they were real.”⁹⁹ While deepfakes demonstrate the early potential for manipulating video imagery, the misuse of pornographic videos and the potential to create photorealistic, defamatory content,¹⁰⁰ or disinformation for state-sponsored espionage¹⁰¹ highlights the need for management and moderation in the deployment of these tools.

Despite the foreshadowing of potential problems with generative AI, the growth of generative AI has been exponential,¹⁰² at

⁹⁷ *Id.*

⁹⁸ Meskys, *supra* note 94, at 26; *see also* Ian Goodfellow et al., *Generative Adversarial Networks*, 63 COMM’NS ACM 139 (Nov. 2020).

⁹⁹ Meskys, *supra* note 94, at 26.

¹⁰⁰ *See* Garon, *supra* note 21, at 432 (“[A] photograph or video asserts that the images are exactly as they appeared. This assertion of accuracy makes a photorealistic image appear self-validating as to its truthfulness.”); *see also* Michael Atleson, *Chatbots, Deepfakes, and Voice Clones: AI Deception for Sale*, FTC BLOG (Mar. 20, 2023), <https://www.ftc.gov/business-guidance/blog/2023/03/chatbots-deepfakes-voice-clones-ai-deception-sale> [<https://perma.cc/99DA-EJMD>]; Regina Rini & Leah Cohen, *Deepfakes, Deep Harms*, 22 J. ETHICS & SOC. PHIL. 143, 148 (2022); Cass R. Sunstein, *Falsehoods and the First Amendment*, 33 HARV. J.L. & TECH. 387, 405 (2020).

¹⁰¹ *See* Jon M. Garon, *When AI Goes to War: Corporate Accountability for Virtual Mass Disinformation, Algorithmic Atrocities, and Synthetic Propaganda*, 49 N. KY. L. REV. 181, 192 (2022) (“The expansion and reliance on artificial intelligence technologies are among the greatest advances in the twenty-first century and greatest threats for misuse.”); Rachel E. VanLandingham, *Jailing the Twitter Bird: Social Media, Material Support to Terrorism, and Muzzling the Modern Press*, 39 CARDOZO L. REV. 1, 13 (2017) (“The link between social media platforms and terrorism competes with privacy concerns as one of the most discussed and most concerning, dynamics emanating from modern society’s explosive utilization of these communication technologies.”).

¹⁰² *See* U.S. CHAMBER OF COMMERCE, *U.S. Chamber’s AI Commission Report Highlights the Promise of AI While Calling for a Risk-Based, Regulatory Framework*, (Mar. 9, 2023), <https://www.uschamber.com/technology/u-s-chambers-ai-commission-report-highlights-the-promise-of-ai-while-calling-for-a-risk-based-regulatory-framework> [<https://perma.cc/44JG-W7PA>] (“The development of AI and the introduction of AI-based systems are growing exponentially. Over the next 10 to 20 years, virtually every business and government agency will use AI. This will have a profound impact on society, the economy, and national security.”).

least as measured by the growth of U.S. venture capital.¹⁰³ Pitchbook reported that “funding of generative AI was up 27% year over year in 2022, hitting \$1.4 billion,” an increase from a mere \$200 million in 2020.¹⁰⁴ “Generative AI was a breakthrough. Rather than simply *perceive* and *classify* a photo of a cat, machine learning is now able to *create* an image or text description of a cat on demand.”¹⁰⁵

Generative AI hit the public in a manner seldom seen with new technologies. “[T]he November launch of [ChatGPT,] the conversational robot from start-up OpenAI marked a turning point in its perception by the general public and investors. ‘Every so often we have platforms that come along and result in an explosion of new companies. We saw this with the internet and mobile, and AI could be the next platform,’ said Shernaz Daver of California-based Khosla Ventures.”¹⁰⁶ ChatGPT hit 1 million users in a mere five days¹⁰⁷ and logged over 100 million unique users two months after its beta launch.¹⁰⁸ For a comparison, Twitter required two years to achieve its first million users and even Instagram required 2.5 months.¹⁰⁹

In addition to the OpenAI service, there is also Stability AI’s Stable Diffusion, an open-source generative AI system. But ChatGPT runs on massive processing from Microsoft Azure and Stable Diffusion runs on Amazon’s cloud system.¹¹⁰ These are driving a resurgence in

¹⁰³ See Sara Lebow, *Generative AI Funding Exploded Over the Past 2 Years*, INSIDER INTELLIGENCE (Jan. 20, 2023), <https://www.insiderintelligence.com/content/generative-ai-funding-exploded-over-past-2-years> [https://perma.cc/WTA7-45QE].

¹⁰⁴ *Id.*

¹⁰⁵ *What Is Generative AI?*, *supra* note 84.

¹⁰⁶ AFP, *ChatGPT Sparks AI ‘Gold Rush’ in Silicon Valley*, THE NEWS (Feb. 19, 2023), <https://www.thenews.com.pk/latest/1042216-chatgpt-sparks-ai-gold-rush-in-silicon-valley> [https://perma.cc/YRC2-4XTE].

¹⁰⁷ Arooj Ahmed, *Chat GPT Achieved One Million Users in Record Time - Revolutionizing Time-Saving in Various Fields*, DIGIT. INFO. WORLD (Jan. 27, 2023), <https://www.digitalinformationworld.com/2023/01/chat-gpt-achieved-one-million-users-in.html> [https://perma.cc/TE3F-9K5H].

¹⁰⁸ See Dan Milmo, *ChatGPT Reaches 100 Million Users Two Months After Launch*, THE GUARDIAN (Feb. 2, 2023), <https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app> [https://perma.cc/SVZ9-2E3M].

¹⁰⁹ Ahmed, *supra* note 107 (citing Statista).

¹¹⁰ See Kenrick Cai & Alex Konrad, *Six Things You Didn’t Know About ChatGPT, Stable Diffusion and the Future of Generative AI*, FORBES (Feb. 2, 2023), www.forbes.com/sites/kenrickcai/2023/02/02/things-you-didnt-know-chatgpt-stable-diffusion-generative-ai/ [https://perma.cc/QLF7-W2FD].

demand for highly specialized processors, particularly those designed by Nvidia as well as those by AMD.¹¹¹

The commitment to specialized graphics processing units (GPUs) in data centers provides another measure of the growth of generative AI. “[A]ccording to research firm Gartner, the share of specialized chips such as GPUs that are used in data centers is expected to rise to more than 15% by 2026 from less than 3% in 2020.”¹¹² “ChatGPT, Silicon Valley’s latest app sensation, has investors rushing to find the next big thing in generative AI, the technology that some hail as the beginning of a new era in big tech.”¹¹³

Even a brief introduction should likely include a competitive technology that has been providing natural language responses to difficult questions since 2009.¹¹⁴ Wolfram|Alpha, created by Stephen Wolfram, has been providing just such a service.¹¹⁵ As he describes it, Wolfram Language and Wolfram|Alpha have built a computational language to give AI “what we can think of as computational knowledge superpowers, that leverage the non-human-like power of structured computation and structured knowledge.”¹¹⁶ Wolfram|Alpha falls into the category of extractive AI, tools that can search the content available to it and retrieve responses based on an AI-weighted evaluation of possible choices.

The inherent limitations of deep neural networks are that they cannot really discern, ignore, and focus—at least not yet. The power of Wolfram|Alpha’s computational language model is that it focuses on

¹¹¹ See Chavi Mehta, *Nvidia Results Show Its Growing Lead in AI Chip Race*, REUTERS (Feb. 23, 2023), <https://www.reuters.com/technology/nvidia-results-show-its-growing-lead-ai-chip-race-2023-02-23/> [<https://perma.cc/4VGQ-EL87>] (“The company got its start in the graphics chip business for PCs by helping video games look more realistic, and then rode the cryptocurrency wave as its chips were used for mining. Now, the next push comes from generative AI.”).

¹¹² *Id.* (Nvidia “controls about 80% of the market for graphic processing units (GPUs). . . . Advanced Micro Devices . . . is the second-biggest player in the GPU industry, with a market share of roughly 20%.”).

¹¹³ AFP, *supra* note 106.

¹¹⁴ See Stephen Wolfram, *Wolfram Alpha as the Way to Bring Computational Knowledge Superpowers to ChatGPT*, STEPHEN WOLFRAM WRITINGS (Jan. 9, 2023), <https://writings.stephenwolfram.com/2023/01/wolframalpha-as-the-way-to-bring-computational-knowledge-superpowers-to-chatgpt/> [<https://perma.cc/G2UE-HNQF>]; Bobbie Johnson, *Where Does Wolfram Alpha Get Its Information?*, THE GUARDIAN (May 20, 2009), <https://www.theguardian.com/technology/2009/may/21/1> [<https://perma.cc/QB4R-JXJB>].

¹¹⁵ See Wolfram, *supra* note 114.

¹¹⁶ *Id.*

accuracy rather than scope. When the processes are combined, however, each improves the other.

Inside Wolfram|Alpha, everything is being turned into computational language, and into precise Wolfram Language code, that at some level has to be “perfect” to be reliably useful. But the crucial point is that ChatGPT doesn’t have to generate this. It can produce its usual natural language, and then Wolfram|Alpha can use its natural language understanding capabilities to translate that natural language into precise Wolfram Language.¹¹⁷

Increasingly, large data providers are recognizing that they must follow this model to combine extractive AI tools as “guardrails” to the generative AI output.¹¹⁸ For example, both of the two largest legal research services, RELX’s Lexis and Thomson Reuters’ Westlaw, each have entered the AI arms race, as have their parent companies more generally. Lexis boasts that it began implementing natural language extractive AI in 2017 and in 2023 launched services to augment extractive AI with generative AI.¹¹⁹ Westlaw has been offering similar tools. “Westlaw implemented an AI technique – natural language processing – in the 1990s, which made it possible for users to search for content with normal search queries instead of complex Boolean search strings.”¹²⁰ The race between these two legal and business giants heated up when Thomson Reuters spent \$650 million to acquire Casetext, which has been innovating through its early access to ChatGPT-4.¹²¹ The investments by both RELX and Thomson Reuters on integration of

¹¹⁷ *Id.*

¹¹⁸ *Id.*

¹¹⁹ See Jake Nelson, *Combining Extractive and Generative AI for New Possibilities*, LEXISNEXIS LEGAL INSIGHTS (June 6, 2023), <https://www.lexisnexis.com/community/insights/legal/b/thought-leadership/posts/combining-extractive-and-generative-ai-for-new-possibilities> [<https://perma.cc/SV5H-52B8>].

¹²⁰ *Not All Legal AI is Created Equal*, THOMSON REUTERS, <https://legal.thomsonreuters.com/en/insights/white-papers/5-things-to-consider-when-evaluating-legal-ai-solutions> [<https://perma.cc/HBT7-HLA3>].

¹²¹ See Kyle Wiggers, *Thomson Reuters Buys Casetext, an AI Legal Tech Startup, for \$650M in Cash*, TECHCRUNCH (June 26, 2023), <https://techcrunch.com/2023/06/26/thomson-reuters-buys-casetext-an-ai-legal-tech-startup-for-650m-in-cash/> [<https://perma.cc/26W5-8U37>]; *Thomson Reuters to Acquire legal AI Firm Casetext for \$650 Million*, REUTERS (June 27, 2023), <https://www.reuters.com/markets/deals/thomson-reuters-acquire-legal-tech-provider-casetext-650-mln-2023-06-27/> [<https://perma.cc/TTT9-4AAC>].

extractive AI to be combined with generative AI will mark a substantial next step in the quality of AI services.

The descriptions in this technical introduction are not particularly technical nor likely to be considered accurate to a data scientist. They are approximations that reflect a non-scientific understanding of the technologies as they are developing. They also reflect the popular descriptions at the time of this writing. As computational language models are used to assist LLMs and the wide array of other training structures are added to the mix, the potential for these systems will only increase. This gives rise to both opportunities and limitations. The next sections discuss both.

IV. THE OPPORTUNITIES OF GENERATIVE AI

Generative AI remains in its infancy and integration with extractive AI is still nascent, but the explosive interest by investors, major tech multinationals, and the general public suggests that the potential for this technology will be more than mere hype. “Since its discrete release in late November, ChatGPT has become one of the fastest growing apps ever and pushed Microsoft and Google to rush out projects that had until now stayed carefully guarded over fears that the technology wasn’t yet ready for the public.”¹²²

NFX, “a venture firm exclusively focused on pre-seed & seed stage startups,”¹²³ has provided extensive coverage of generative AI since October 2022, primarily by its managing partner, James Currier.¹²⁴ Currier introduced generative AI in an introductory article: “Some have called it “Generative AI,” but AI is only half of the equation. AI models are the enabling base layers of the stack. The top layers will be thousands of applications. Generative Tech is about what will actually touch us – what you can do with AI as a partner.”¹²⁵

¹²² AFP, *supra* note 106.

¹²³ NFX, *See What Others Do Not*, NFX, <https://www.nfx.com/> (last visited July 8, 2023); James Currier, SILICON VALLEY OPEN DOORS, <https://www.svod.org/james-currier/> (last visited July 8, 2023).

¹²⁴ *See, e.g.*, James Currier, *Generative Tech Begins*, NFX (Oct. 2022), <https://www.nfx.com/post/generative-tech#Generative-Tech-will-have-unusual-market-dynamics-because-it%E2%80%99s-already-consensus> [<https://perma.cc/Q52Q-3JUF>]; *see also* James Currier, NFX <https://www.nfx.com/team/james-currier> [<https://perma.cc/QH94-9MKA>] (“The biggest change to the Internet since crypto just happened. A whole new world of applications opened up to Founders in the last 6 months.”).

¹²⁵ Currier, *Generative Tech Begins*, *supra* note 124.

These opportunities are not without risks. In addition, open-source projects and projects “in the wild” means that generative AI experimentation will not be limited to a handful of heavily financed multinational corporations whose operations are entangled with regulators across industries and across the globe.¹²⁶ Meta—which admittedly is one of the heavily financed multinationals—created “an A.I. technology, called LLaMA, that can power online chatbots. But instead of keeping the technology to itself, Meta released the system’s underlying computer code into the wild.”¹²⁷ This was not a true open-source release as much as a failure to manage and restrict the use of the code, but it will likely have a similar effect. Access to public domain and open access code is creating additional models. “After the unexpected leak of Meta’s LLaMA, researchers cleverly enhanced it with ChatGPT outputs, creating chatbots Alpaca and Vicuna. These new bots perform nearly as well as GPT-3.5 and cost less to train—Alpaca took just 3 hours and \$600.”¹²⁸ The more experimental generative AI environment has produced multiple projects to render pornography, convert image of individuals to nude photographs, and even to train a “GPT-4chan” LLM model weighting the analysis to be “fine-tuned on 4chan’s politically incorrect board and generated offensive content.”¹²⁹

¹²⁶ See Khari Johnson, *Open Source Llama Upsets the AI Horse Race*, WIRED (July 26, 2023), www.wired.com/story/metasp-open-source-llama-upsets-the-ai-horse-race/ [<https://perma.cc/7UUV-3PY5>] (“Meta’s system is available for free to developers, startups, and others interested in creating custom variations of the model. By supplying a cheaper option, Meta’s Llama 2 makes it easier for small companies or lone coders to create new products and services, potentially accelerating the current AI boom.”); Erin Griffith & Cade Metz, *‘Let 1,000 Flowers Bloom’: A.I. Funding Frenzy Escalates*, FORBES (Mar. 14, 2023), <https://www.nytimes.com/2023/03/14/technology/ai-funding-boom.html> [<https://perma.cc/57YA-W44R>]. See generally Prathamesh Ingle, *Top VC Firms Investing in Artificial Intelligence (AI) Companies*, MARKETPOST (July 18, 2023), <https://www.marktechpost.com/2023/07/18/top-vc-firms-investing-in-artificial-intelligence-ai-companies/> [<https://perma.cc/7458-RWTX>].

¹²⁷ Cade Metz & Mike Isaac, *In Battle Over A.I., Meta Decides to Give Away Its Crown Jewels*, N.Y. TIMES (May 18, 2023), <https://www.nytimes.com/2023/05/18/technology/ai-meta-open-source.html> [<https://perma.cc/C7X2-LSG6>].

¹²⁸ Andre Ferretti, *Navigating the Open-Source AI Landscape: Data, Funding, and Safety*, EFFECTIVE ALTRUISM F. (Apr. 12, 2023), <https://forum.effectivealtruism.org/posts/N25EARxvbxYJa5pbB/navigating-the-open-source-ai-landscape-data-funding-and> [<https://perma.cc/J5SA-WUCK>].

¹²⁹ *Id.* (“Although the model is no longer available on Hugging Face, its weights and dataset can still be found online.”).

While generative AI can potentially be incorporated into almost any human activity, there are some key areas where the technology has already been put to use. “[T]he introduction of APIs [application program interfaces] and open-source tools are allowing entrepreneurs to build upon existing foundation models and narrow in on specific verticals like gaming, graphic design, copywriting, media, and entertainment.”¹³⁰

A study on collaboration among people using AI as a catalyst found that using technology as an augmentation to human creative collaborations proved very promising.

AI may play important roles in influencing social dynamics during creativity by: 1) implicitly seeding common ground at the start of collaboration, 2) acting as a psycho-logical safety net in creative risk-taking, 3) providing a force for group progress, 4) mitigating interpersonal stalling and friction, and 5) altering users’ collaborative and creative roles.¹³¹

Generative AI text, music, video, gaming, design, and software have the power to reshape the world around us and in so doing to also reshape and mediate the ways in which we interact with each other. These areas are ripe for transformational change as AI augmentation and AI creation redefines how media is conceived, developed, and executed.

A. Image Generation

Image generation stands out as one of the key functions of generative AI. An image generator can create a novel image that is functionally indistinguishable from real photographs.¹³² Image

¹³⁰ Clarke Pennington, *Generative AI: The New Frontier for VC Investment*, FORBES (Jan. 17, 2023), www.forbes.com/sites/columbiabusinessschool/2023/01/17/generative-ai-the-new-frontier-for-vc-investment/ [https://perma.cc/6VR2-CG2U].

¹³¹ Minhyang (Mia) Suh et al., *AI as Social Glue: Uncovering the Roles of Deep Generative AI During Social Music Composition*, CHI CONF. ON HUM. FACTORS IN COMPUTING SYS. (May 8, 2021), <https://dl.acm.org/doi/fullHtml/10.1145/3411764.3445219> [https://perma.cc/S9TF-7JSF].

¹³² See DANIEL ZHANG ET AL., THE AI INDEX 2022 ANNUAL REPORT 54 (Stanford Inst. Hum.-Centered AI, 5th ed. Mar. 2022).

generators can also be used to create wholly original works of art.¹³³ “With AI art generators, you can type in a prompt as detailed or vague as you’d like and have the image you were thinking of pop up on your screen instantly. These tools can help with branding, social media content creation, vision boards, and more.”¹³⁴ The most impactful art generation program is DALL-E 2 by OpenAI (the same company that released ChatGPT).

Creating cute pictures of a corgi in a house made of sushi isn’t exactly a profitable business case, at least not for large enterprises,” . . . said Peter van der Putten, director of the AI Lab at software firm Pegasystems Inc. . . . And yet, he said, “generative AI startups are popping up left and right, in areas such as marketing, support, service and other content creation.”¹³⁵

Generated Photos,¹³⁶ an online service, boasts that it can provide “2,677,079 low-res generated photos” to users for free,¹³⁷ with the promise that “[g]enerated photos are created from scratch by AI systems. All images can be used for any purpose without worrying about copyrights, distribution rights, infringement claims, or royalties.”¹³⁸ As discussed in the next section, however, these claims promising no copyright violations and no “infringement” claims (that presumably covers defamation and violations of the rights of publicity) may be overly optimistic.

AI-based image generation has already gone mainstream. Adobe’s highly popular photo editing software, Photoshop, has added an AI integration.¹³⁹ Google has added a “magic eraser” function to its

¹³³ Sabrina Ortiz, *The Best AI Art Generators: DALL-E 2 and Other Fun Alternatives to Try*, ZDNET (Feb. 27, 2023), <https://www.zdnet.com/article/best-ai-art-generator/> [<https://perma.cc/V5T5-X5GP>].

¹³⁴ *Id.* In addition to DALL-E, other services include Dream by WOMBO, Craiyon, Midjourney, and MyHeritage’s AI Time Machine.

¹³⁵ Angus Loten, *Generative AI Startups Attract Business Customers, Investor Funding*, WALL ST. J. (Oct. 25, 2022), <https://www.wsj.com/articles/generative-ai-startups-attract-business-customers-investor-funding-11666736176> [<https://perma.cc/WMS9-28NZ>].

¹³⁶ Generated Photos, <https://generated.photos/> (last visited Mar. 1, 2023).

¹³⁷ Generated Photos, *Pricing*, <https://generated.photos/pricing> (last visited Mar. 1, 2023).

¹³⁸ Generated Photos, *supra* note 136.

¹³⁹ Brian X. Chen, *How to Use A.I. to Edit and Generate Stunning Photos*, N.Y. TIMES (June 2, 2023), <https://www.nytimes.com/2023/06/02/technology/ai-photo-editing.html> [<https://perma.cc/4GQV-3VXN>].

photo apps that is becoming increasingly powerful.¹⁴⁰ Lambda Labs has created one of the first image mixers, that enable a user to upload 3-5 source images to generate a new image based on those works.¹⁴¹ These highly popular tools are complemented by dozens of apps that provide some level of AI-based photo enhancements.

Image generation can move beyond photorealistic images of people, places, and things. “[C]ompanies like NVIDIA have already used image generators to create virtual worlds for gaming.”¹⁴² This has the potential to create ever-expanding videogame and metaverse environments in the manner often depicted in films.¹⁴³ Image generation can also be used in fashion to virtualize different clothing patterns and different applications of color and design on those clothing patterns.¹⁴⁴ For example, “Stitch Fix, the online service that uses recommendation algorithms to personalize apparel, says it has experimented with DALL-E 2 to visualize its products based on specific characteristics like color, fabric and style.”¹⁴⁵ “Rowan Curran, an AI and data science analyst at IT research firm Forrester Research Inc., said he expects generative AI

¹⁴⁰ Dina Berrada & Esteban Kozak, *Magic Eraser Plus More Google Photos Features Coming to Google One*, GOOGLE (Feb. 23, 2023), <https://blog.google/products/photos/magic-eraser-android-ios-google-one/> [<https://perma.cc/XQ5Q-W9NZ>].

¹⁴¹ Damir Yalalov, *Lambda Labs Announced an AI Image Mixer That Can Combine Up to Five Images*, METAVERSE POST (Jan. 20, 2023, 4:59 AM), <https://mpost.io/lambda-labs-announced-an-ai-image-mixer-that-can-combine-up-to-five-images/> [<https://perma.cc/CZ7Z-RE5W>].

¹⁴² ZHANG, *supra* note 132, at 54.

¹⁴³ See Chris E. Hayner, *21 Virtual Reality Movies, Ranked From Awful to Amazing*, GAMESPOT (May 3, 2021, 4:03 PM), <https://www.gamespot.com/gallery/21-virtual-reality-movies-ranked-from-awful-to-amazing/2900-1874/> [<https://perma.cc/YLE4-Y372>] (listing the Matrix as the best such film. Others include Ready Player One, Tron, Spy Kids 3-D, Total Recall, and many others.); Ben Wodecki, *AI in Video Games: From Development to Design*, AI BUS. (Oct. 15, 2021), <https://aibusiness.com/verticals/ai-in-video-games-from-development-to-design> [<https://perma.cc/KAC9-K6FC>] (discussing the various roles AI can take in game design, development, and game play); Devindra Hardawar, *How ‘Microsoft Flight Simulator’ became a ‘living game’ with Azure AI*, ENGADGET (Sept. 25, 2020), <https://www.engadget.com/microsoft-flight-simulator-azure-ai-machine-learning-193545436.html> [<https://perma.cc/5G57-BEUN>] (“Microsoft’s Azure . . . was able to push 2.5 petabytes worth of Bing Maps satellite photo data through Azure machine learning to construct the virtual world of *Flight Simulator*. You could say it’s really the cloud that brings the game to life. Azure also helps to model real-time weather.”).

¹⁴⁴ See ZHANG, *supra* note 132, at 54.

¹⁴⁵ Kyle Wiggers, *Businesses Including Stitch Fix Are Already Experimenting with DALL-E 2*, TECHCRUNCH (Aug. 9, 2022, 3:06 PM), <https://techcrunch.com/2022/08/09/businesses-including-stitch-fix-are-already-experimenting-with-dall-e-2/> [<https://perma.cc/JJ6E-JCH2>].

to impact enterprises from top to bottom. . . . “[T]he use of text-to-image generators for inspiration, ideation and prototyping for marketers, designers and developers has just barely begun.”¹⁴⁶

The visual works need not be limited to still images. Meta has already announced “Make-A-Video, a new AI system that lets people turn text prompts into brief, high-quality video clips.”¹⁴⁷ The output looks like GIFs, very short animation of generated images, but as the proof of concept, they illustrate the next step in text-to-video generation. Other vendors have already surpassed these efforts, including Synthesia, Pictory, Elai.io, and a growing list.¹⁴⁸

The same expansion of skill sets is also improving the scope of deepfakes, the ability to revise and reconceptualize other people’s video with the editor’s content. “Startups like Hour One, Synthesia, Uneeq and D-ID see more prosaic applications for the technology: putting infinite numbers of shiny, happy people in personalized online ads, video tutorials and presentations.”¹⁴⁹ The technology can also be used for clever skits, for industrial videos, or more problematically, for pornography and political disinformation. “A report last month by misinformation watchdog NewsGuard warned of the dangers of GPT on its own, saying it gives peddlers of political misinformation, authoritarian information operations, and health hoaxes the equivalent of ‘an army of skilled writers spreading false narratives.’”¹⁵⁰

Advocates, however, see these synthesized media videos as extensions of the individuals being promoted by enterprises. “As the technology gets faster and cheaper, Hour One . . . hopes to put avatars into real-time video calls, giving users their own “super communicators,” enhanced “extensions” of themselves.”¹⁵¹ Natalie Monbiot, head of strategy for Hour One, explains that through social media “[w]e already do that every day. . . . And this is almost just like an animated version of you that can actually do a lot more than a nice photo. It can actually do work on your behalf.”¹⁵²

¹⁴⁶ Loten, *supra* note 135.

¹⁴⁷ *Introducing Make-A-Video: An AI System That Generates Videos from Text*, META AI (Sept. 29, 2022), <https://ai.facebook.com/blog/generative-ai-text-to-video/> [<https://perma.cc/D97J-C7YE>].

¹⁴⁸ See Alex McFarland, 8 “Best” AI Video Generators (March 2023), BEST OF (Mar. 5, 2023), <https://www.unite.ai/best-ai-video-generators/>.

¹⁴⁹ See Alex Pasternack, *GPT-Powered Deepfakes Are a ‘Powder Keg,’* FAST CO. (Feb. 22, 2023), <https://www.fastcompany.com/90853542/deepfakes-getting-smarter-thanks-to-gpt> [<https://perma.cc/8EVQ-9CZW>].

¹⁵⁰ *Id.*

¹⁵¹ *Id.*

¹⁵² *Id.* (quoting Natalie Monbiot).

B. Text Generation

While generative art and synthetic media were among the first of the generative AI content to become popular, it was the release of ChatGPT that transformed the growth of generative AI into a popular inflection point.¹⁵³ As the Chronicle of Higher Ed recently reported, “[i]t’s hard to believe that ChatGPT appeared on the scene just three months ago, promising to transform how we write. The chatbot, easy to use and trained on vast amounts of digital text, is now pervasive.”¹⁵⁴ The OpenAI “chatbot exploded into the mainstream almost overnight.”¹⁵⁵ Of course, like any overnight success, the effort was based on decades of more systematic research and perseverance.

OpenAI’s ChatGPT and Google’s Bard are perhaps the two systems garnering the most headlines, but they are not alone. Others recommended by Tech.Co include ChatSonic, OpenAI Playground, YouChat, Perplexity AI, Character AI, and Jasper Chat.¹⁵⁶ Some of these services are closer to search engines, helping find answers on the Internet and condensing them as text.¹⁵⁷

Despite some of the early giddiness surrounding the introduction of ChatGPT, none of the text generators are really ready to be generating new novels based on a simple prompt or input.¹⁵⁸ On the other hand, they are often quite capable of delivering a cleverly numbered list of product or service attributes¹⁵⁹ or drafting a coherent five-paragraph

¹⁵³ Will Douglas Heaven, *ChatGPT is Everywhere. Here’s Where It Came from*, MIT TECH. REV. (Feb. 8, 2023), <https://www.technologyreview.com/2023/02/08/1068068/chatgpt-is-everywhere-heres-where-it-came-from/> [<https://perma.cc/8JEP-QPDP>].

¹⁵⁴ Beth McMurtrie, *ChatGPT is Everywhere*, THE CHRON. OF HIGHER EDUC. (Mar. 6, 2023), https://www.chronicle.com/article/chatgpt-is-already-upending-campus-practices-colleges-are-rushing-to-respond?cid=gen_sign_in [<https://perma.cc/CLA8-HZJH>].

¹⁵⁵ Heaven, *supra* note 153.

¹⁵⁶ Aaron Drapkin, *Best ChatGPT AI Alternatives You Need to Try in 2023*, TECH.CO (Feb. 20, 2023), <https://tech.co/news/best-chatgpt-alternatives>.

¹⁵⁷ *See id.* (“YouChat is an AI chatbot and search assistant. It will provide you with similar answers to that of ChatGPT, but will also link you to webpages that relate to whatever query you choose to type in.”).

¹⁵⁸ *See* Amira Loutfi, *12 AI Tools to Write Your Novel for You*, METASTELLAR (Oct. 18, 2021), <https://www.metastellar.com/nonfiction/on-writing/12-artificial-intelligence-tools-to-write-your-novel-for-you/> [<https://perma.cc/9Y75-L7LS>].

¹⁵⁹ *See* Thomas H. Davenport & Nitin Mittal, *How Generative AI is Changing Creative Work*, HARV. BUS. REV. (Nov. 14, 2022), <https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work> [<https://perma.cc/M6MM-S4PH>] (discussing a well

essay.¹⁶⁰ Of course, the systems do not compose in the manner of a human author. Instead, they predict likely appropriate text based on how that text has been used in its training data.

The software uses complex machine learning models to predict the next word based on previous word sequences, or the next image based on words describing previous images. LLMs began at Google Brain in 2017, where they were initially used for translation of words while preserving context. Since then, large language and text-to-image models have proliferated at leading tech firms including Google (BERT and LaMDA), Facebook (OPT-175B, BlenderBot), and OpenAI, a nonprofit in which Microsoft is the dominant investor (GPT-3 for text, DALL-E2 for images, and Whisper for speech). Online communities such as Midjourney (which helped win the art competition), and open-source providers like HuggingFace, have also created generative models.

These models have largely been confined to major tech companies because training them requires massive amounts of data and computing power. GPT-3, for example, was initially trained on 45 terabytes of data and employs 175 billion parameters or coefficients to make its predictions; a single training run for GPT-3 cost \$12 million. Wu Dao 2.0, a Chinese model, has 1.75 trillion parameters. Most companies don't have the data center capabilities or cloud computing budgets to train their own models of this type from scratch.¹⁶¹

The consequence of learning to write based on massive but historical training data is discussed in the following sections. But the

composed list of attributes for generative AI text. The human authors noted GPT-3 “came up with ideas that we didn't think of. The last point about personalized content [being provided by generative AI], for example, is not one we would have considered.”); Ian Bogost, *ChatGPT is Dumber Than You Think*, THE ATLANTIC (Dec. 7, 2022), <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-openai-artificial-intelligence-writing-ethics/672386/> [<https://perma.cc/DZW7-6MNE>] (“In short, it wrote a basic, high-school-style five-paragraph essay. . . . Treat it like a toy, not a tool.”).

¹⁶⁰ See McMurtrie, *supra* note 154 (“One thing ChatGPT does particularly well, he believes, is help students learn the ‘basic building blocks’ of effective academic writing.”).

¹⁶¹ See Davenport & Mittal, *supra* note 159.

difference between structuring an argument and anticipating the next phrases to come is a very important distinction. The current systems utilize only the prompts in front of them to build their arguments. Additional training and modeling will be needed to be able to provide the generative AI an outline of a paper and have it make the connections that generate the flow from one building-block of the paper to the next.

Instead, “[t]he first true ‘killer application’ for generative text, in terms of commercial adoption, has proven to be copywriting: that is, AI-generated website copy, social media posts, blog posts and other marketing-related written content.”¹⁶² While this may not be as dynamic as writing poems, novels, or legal briefs, it represents a very large and somewhat tedious market, attributes that make it perfect for AI support.¹⁶³

As these tools continue to improve, they will expand into more business transactions and more components of those transactions. Contracts will not just be assembled by smart systems; they will likely be drafted from information gleaned directly by the AI about the parties to that transaction.¹⁶⁴

“[M]ake no mistake . . . AI-powered text generation will create many orders of magnitude more value than will AI-powered image generation in the years ahead. Machines’ ability to generate language—to write and speak—will prove to be far more transformative than their ability to generate visual content.”¹⁶⁵ The marketing is shifting to personalized email generation, customer follow-up regarding products and services, and may soon include “real-time coaching and feedback

¹⁶² Rob Toews, *The Biggest Opportunity in Generative AI is Language, Not Images*, FORBES (Nov. 6, 2022, 8:00 PM), <https://www.forbes.com/sites/robtoews/2022/11/06/the-biggest-opportunity-in-generative-ai-is-language-not-images/?sh=6674bcae789d> [https://perma.cc/2LM4-MWRT].

¹⁶³ *See id.*

¹⁶⁴ *See* Beverly Rich, *How AI is Changing Contracts*, HARV. BUS. REV. (Feb. 12, 2018), <https://hbr.org/2018/02/how-ai-is-changing-contracts> [https://perma.cc/LWT7-BWUX].

AI software, however, can easily extract data and clarify the content of contracts. (It could quickly pull and organize the renewal dates and renegotiation terms from any number of contracts.) It can let companies review contracts more rapidly, organize and locate large amounts of contract data more easily, decrease the potential for contract disputes (and antagonistic contract negotiations), and increase the volume of contracts it is able to negotiate and execute.

¹⁶⁵ Toews, *supra* note 162.

to human sales agents on calls; summarize sales discussions and suggest next steps.”¹⁶⁶

These incremental improvements, along with the emerging skills of personalization will lead to AI services that can talk with people in a genuinely companionable manner. This will go beyond marketing to socialization—or socialization with marketing hidden in its center.

Eventually—and perhaps quite quickly—these systems will be able to generate longer blocks of text, building the connection that goes from simple marketing to complex narrative. At least one system, Dramatron, claims to have already cracked this code.¹⁶⁷ “Fundamental to Dramatron is the so-called ‘prompt chaining’. The method chains prompts at lower hierarchy levels with prompts including output from higher hierarchy levels.”¹⁶⁸ Rather than following the traditional LLM model of stream of consciousness text flow, Dramatron uses the common Hollywood strategy of starting with a log line and fleshing out the story from the preexisting architecture of the story.¹⁶⁹ “After a human enters the log line, Dramatron generates a list of characters, the plot, and locations along with their descriptions. All these parameters should be consistent with the log line. Finally, the system generates the dialogs for each scene.”¹⁷⁰ But even here, this is not—at least yet—a substitute for AI authored works. “Dramatron is a co-writing system that has only been used in collaboration with human writers and was not conceived or evaluated to be used autonomously.”¹⁷¹

Enterprises seeking to license generative AI systems will be able to manage the training data and control the output for factual errors, at least to some degree. One possible corporate arms race may well be the speed and depth companies seek to control, train, and exploit proprietary data sets for the best answers and most correct insights from their generative text systems. For fields such as law, medicine, business, and finance, accuracy and depth are the key metrics. Highly accurate,

¹⁶⁶ *Id.*

¹⁶⁷ See Matthias Bastian, *Deepmind’s “Dramatron” Can Write Film and Theater Scripts*, THE DECODER (Dec. 10, 2022), <https://the-decoder.com/deepminds-dramatron-can-write-film-and-theater-scripts/> [<https://perma.cc/4YAS-7YEY>].

¹⁶⁸ *Id.*

¹⁶⁹ *Id.*

¹⁷⁰ *Id.*

¹⁷¹ *About Dramatron, DRAMATRON*, <https://deepmind.github.io/dramatron/details.html> (last visited Mar. 9, 2023). See generally Piotr Mirowski, Kory W. Mathewson, Jaylen Pittman & Richard Evans, *Co-Writing Screenplays and Theatre Scripts with Language Models: An Evaluation by Industry Professionals*, ARXIV (Sept. 29, 2022), <https://arxiv.org/abs/2209.14958> [<https://perma.cc/VU5Y-78R3>].

proprietary text engines in these fields will transform these industries and the enterprises controlling these systems will become the leaders in each field.

C. Audio, Including Music, Voice, and Sound

Friedrich Nietzsche explained simply, “[w]ithout music, life would be a mistake.”¹⁷² “For thousands of years, music has been a form of human expression. Music production has always been an innate pursuit of human beings, and a product of creativity and collaboration.”¹⁷³ Yet, despite the highly personal nature of music, algorithmic composition and digitized music are among the earlier areas of generative success, with early experiments dating back to the 1950s.¹⁷⁴ Those early experiments, however, were human labor intensive and extremely limited in scope. But those days are behind us.¹⁷⁵

Much like the visual and text training models, music AI is built upon deep learning neural networks. “If you’re feeding it different types of music, it will be able to differentiate things like tempo, chords, length, the relationship between notes, and use all of those details to generate new melodies.”¹⁷⁶ Music, however, is more inherently structured, so “[s]ome AI frameworks are even trained to use musical theory to guide the end result.”¹⁷⁷

One particularly interesting example is “The Lost Tapes of the 27 Club,” an AI memorial to some of the many music artists who died early, in the height of their careers, remembered through an AI generated song in the artist’s voice and musical style, produced by Over

¹⁷² FRIEDRICH NIETZSCHE, TWILIGHT OF THE IDOLS, OR, HOW TO PHILOSOPHIZE WITH THE HAMMER 9 (1894).

¹⁷³ Weiming Liu, *Literature Survey of Multi-track Music Generation Model Based on Generative Confrontation Network in Intelligent Composition*, 79 J. SUPERCOMPUTING 6560 (2023).

¹⁷⁴ See Nick Collins, *The Analysis of Generative Music Programs*, 13 ORGANISED SOUND, 237 (2008) (“Composers have spent more than fifty years devising computer programs for the semi-automated production of music.”).

¹⁷⁵ See Eric Sunray, *Sounds of Science: Copyright Infringement in AI Music Generator Outputs*, 29 CATH. U.J.L. & TECH. 185, 191 (2021) (“Applying AI technology to music is nothing new. Automatic music generation can be traced back to the earliest days of AI research. Further, AI already aides the modern music industry by performing functions ranging from talent discovery to playlist curation.”).

¹⁷⁶ *Id.*

¹⁷⁷ *Id.*

the Bridge.¹⁷⁸ “Over the Bridge’s The Lost Tapes of the 27 Club project fed an AI songwriting algorithm with existing songs to generate pretty convincing new compositions and lyrics based on the inputs [for] Amy Winehouse, Jimi Hendrix, and Kurt Cobain.”¹⁷⁹

As explained in one research paper on generative AI music, music composition problem solving is a form of design thinking.¹⁸⁰ “Automatic composition is essentially a note sequence prediction problem, and recurrent neural networks are good at solving this type of problem.”¹⁸¹ The music problem-solving or design thinking strategy will eventually also serve to inform the text generation as well, since the recurrent neural networks utilize both the input of the network at each layer and also the output from the previous moment.¹⁸² As a result, in principle, the recurrent neural network can “map from the entire history of previous inputs to each output.”¹⁸³ This enables the music to be much more coherent and allows for the intersection between melody, composition, and the various components of discrete instrument orchestrations.

Analogizing this to the text problem, future iterations of the large language model learning systems for text may be able to “start at the end,” and train the model to fill in the narrative that ends with a story’s resolution and recursively predict the steps leading up to that story.¹⁸⁴ Such approaches in music composition and production are

¹⁷⁸ *AI in Songwriting: 4 Practical Applications*, THE MIX (July 1, 2022), <https://themix.musixmatch.com/post/ai-in-songwriting-4-practical-applications> [<https://perma.cc/W8XE-2JVP>].

¹⁷⁹ *Id.*

¹⁸⁰ Luo Luo, *Practical Exploration on the Construction of Theoretical Courses of Composition Technology in the Age of Artificial Intelligence*, 2022 MOBILE INFO. SYS. (Aug. 2022); see also Rikke Friis Dam & Teo Yu Siang, *What is Design Thinking and Why is It So Popular?*, INTERACTION DESIGN FOUND. (June 27, 2022), <https://www.interaction-design.org/literature/article/what-is-design-thinking-and-why-is-it-so-popular> (last visited Mar. 13, 2023) (“Design thinking is an iterative process in which you seek to understand your users, challenge assumptions, redefine problems and create innovative solutions which you can prototype and test. The overall goal is to identify alternative strategies and solutions that are not instantly apparent with your initial level of understanding.”).

¹⁸¹ *Id.*

¹⁸² *Id.*

¹⁸³ *Id.*

¹⁸⁴ *But see* Mirowski et al., *supra* note 171, at 3 (“[W]e use log lines to start the hierarchical story generation process because it contains plot elements in the answers to questions: Who? What? When and Where? How? Why?”).

leading to a “golden age” in AI music.¹⁸⁵ The uses go beyond the creation of new works. Artists are using the analytical tools to identify what makes compositions popular or impactful, unlocking some of the science behind the magic of music.¹⁸⁶ AI can also help A&R (artists and repertoire—the music industry’s talent scouts) discover the next generation of talent “by combing through music and trying to identify the next breakout star.”¹⁸⁷

The selection of which artists to promote; the decision of which tracks to feature; the choices of how to shape beat, melody, and orchestration; and even the entire creative process for making music is being influenced and occasionally reproduced using AI. The future of music is rapidly moving into the hands of the machine.

The use of generative AI has also seen great success, made famous by one of the most iconic voices on stage and screen—that of James Earl Jones.¹⁸⁸ “During the creation of the *Obi-Wan Kenobi* TV series, James Earl Jones signed off on allowing Disney to replicate his vocal performance as Darth Vader in future projects using an AI voice-modeling tool called Respeecher.”¹⁸⁹

Generative AI can reproduce speech from other speakers, such as the example where a stand-in delivers the line and then recreate that

¹⁸⁵ Bernard Marr, *How Artificial Intelligence (AI) is Helping Musicians Unlock Their Creativity*, FORBES (May 14, 2021, 12:20 AM), <https://www.forbes.com/sites/bernardmarr/2021/05/14/how-artificial-intelligence-ai-is-helping-musicians-unlock-their-creativity/?sh=29ab2f667004> [https://perma.cc/37QS-BUFZ] (“AIVA uses AI and deep learning algorithms to help mainstream users compose their own soundtrack music and scores. It’s the perfect tool for content creators on YouTube, Twitch, Tik Tok, and Instagram who need a steady supply of music but don’t have an endless budget for royalties.”).

¹⁸⁶ *Id.* (“Grammy-nominated producer Alex da Kid used IBM Watson to analyze five years’ of hit songs, as well as cultural data from films, social media, and online articles to figure out a theme for an AI-generated song that fans would enjoy. The final song, ‘Not Easy,’ reached number four on the iTunes Hot Tracks chart within 48 hours after its release.”).

¹⁸⁷ *Id.* (“Warner Music Group acquired a tech start-up last year that uses an algorithm to review social, streaming, and touring data to find promising talent.”).

¹⁸⁸ See Anthony Breznican, *Darth Vader’s Voice Emanated from War-Torn Ukraine*, VANITY FAIR (Sept. 23, 2022), <https://www.vanityfair.com/hollywood/2022/09/darth-vaders-voice-emanated-from-war-torn-ukraine> [https://perma.cc/B3HZ-NR9X].

¹⁸⁹ Benj Edwards, *Darth Vader’s Voice Will Be AI-generated from Now On*, ARS TECHNICA (Sept. 26, 2022, 11:21 AM), <https://arstechnica.com/information-technology/2022/09/james-earl-jones-signed-darth-vader-voice-rights-to-disney-for-ai-use/> [https://perma.cc/W4HS-MVTM] (“Jones, who is 91, has voiced the iconic Star Wars villain for 45 years, starting with *Star Wars: Episode IV—A New Hope* in 1977 and concluding with a brief line of dialog in 2019’s *The Rise of Skywalker*.”); see Breznican, *supra* note 188.

performance using the preferred actor's voice characteristics, or it can be generated from written text.¹⁹⁰ Engineers for Vall-E, the latest generation of voice software, claim that they can train the voice modeling software off a mere three seconds of a speech sample and use it to largely replicate that speaker.¹⁹¹

Improvements in voice software will enable generative AI characters to come alive in video games and metaverse experiences, to improve customer service experience when speaking to chatbots, and to add creative flexibility in film and television production. But the prevalence, accuracy, and access to these systems are already being exploited for impersonation scams. "Technology is making it easier and cheaper for bad actors to mimic voices, convincing people, often the elderly, that their loved ones are in distress."¹⁹² Federal Trade Commission data from 2022 tallies "over 36,000 reports of people being swindled by those pretending to be friends and family."¹⁹³ The ability to fake one's true voice will only make distinguishing fact from fraud harder.

Another social challenge for the use of voice remodeling software is the ethical considerations when call centers use the software to synthesize away the accents of their call centers.¹⁹⁴ A new entrant in the call center industry, Sanas, hopes to ease away the challenges accents place on the customer experience. "Using data about the sounds of different accents and how they correspond to each other, Sanas's AI engine can transform a speaker's accent into what passes for another one – and right now, the focus is on making non-Americans sound like white Americans."¹⁹⁵ In advocating for the use of the technology, Sanas points to the effort demanded by call center workers in India and the

¹⁹⁰ See Benj Edwards, *Microsoft's New AI Can Simulate Anyone's Voice with 3 Seconds of Audio*, ARS TECHNICA (Jan. 9, 2023, 5:15 PM), <https://arstechnica.com/information-technology/2023/01/microsofts-new-ai-can-simulate-anyones-voice-with-3-seconds-of-audio/> [<https://perma.cc/SDR4-UC9Z>].

¹⁹¹ *Id.* ("While using VALL-E to generate those results, the researchers only fed the three-second 'Speaker Prompt' sample and a text string (what they wanted the voice to say) into VALL-E.")

¹⁹² Pranshu Verma, *They Thought Loved Ones Were Calling for Help. It Was an AI Scam*, WASH. POST (Mar. 5, 2023, 6:00 AM), www.washingtonpost.com/technology/2023/03/05/ai-voice-scam/ [<https://perma.cc/8WM7-VE23>].

¹⁹³ *Id.*

¹⁹⁴ See Wilfred Chan, *The AI Startup Erasing Call Center Worker Accents: Is It Fighting Bias – or Perpetuating It?*, THE GUARDIAN (Aug. 24, 2022, 1:00 PM), <https://www.theguardian.com/technology/2022/aug/23/voice-accent-technology-call-center-white-american> [<https://perma.cc/88N9-4EYJ>].

¹⁹⁵ *Id.*

Philippines to reduce regional accents and drop British pronunciation of words as well as learn common American terms.¹⁹⁶ “They have to mimic the culture as well as neutralize their own culture”¹⁹⁷

The danger, Aneesh said, was that artificially neutralizing accents represented a kind of “indifference to difference,” which diminishes the humanity of the person on the other end of the phone. “It allows us to avoid social reality, which is that you are two human beings on the same planet, that you have obligations to each other. It’s pointing to a lonelier future.”¹⁹⁸

Other critics of accent neutralization point out that “[a]ccents can connect consumers to places, products and history in culturally nuanced ways that, at times, general ‘neutral-accented standard American English’ cannot. . . . Multicultural audiences are substantial across the country, and ignoring their presence does a brand no good.”¹⁹⁹ What is best for marketing may nonetheless have a corrosive effect on the broader society and a long-term reinforcement of normalcy driven by algorithm rather than society.

Despite these ethical concerns, however, the expansion of conversational AI driven by generative AI chatbots and the reengineered call centers those chatbots are being used to replace will continue to be a growing component of the customer experience as the costs continue to drop and the quality continues to rise. “Rapid tech advancements and lower barriers have made AI increasingly popular. Among successful AI-powered technologies such as big data and analytics, conversational voice AI is showing considerable promise.”²⁰⁰

¹⁹⁶ *See id.*

¹⁹⁷ *Id.* (quoting A Aneesh, a sociologist and the incoming director of the University of Oregon’s School of Global Studies and Languages).

¹⁹⁸ *Id.*

¹⁹⁹ Antonio Francisco Lewis, *The Ethical Challenge of AI Accent Neutralization vs. Sonic Diversity*, ADWEEK, <https://www.adweek.com/partner-articles/the-ethical-challenge-of-ai-accent-neutralization-vs-sonic-diversity/> [https://perma.cc/B352-S5ZK] (“The Linguistic Society of America highlights a truth that is increasingly important to acknowledge: Everyone who speaks English does so with an accent.”).

²⁰⁰ Sourabh Gupta, *Why Growing Organizations Should Consider Voice AI*, FORBES (June 22, 2022, 6:30 AM), <https://www.forbes.com/sites/forbestechcouncil/2022/06/22/why-growing-organizations-should-consider-voice-ai/?sh=408310837df2> [https://perma.cc/M7H4-DPLY].

In addition to customer service, the improvements in voice AI will likely reignite issues over control of audiobook rights for authors, add voice to most publications, and blur the creative lines among blogs and podcasts. The increased presence of digital voices from chatbots and from people will undoubtedly be growing as fast as the imagination will allow.

D. Code

Another area in which generative AI has exceptional potential is that of software. Generative AI has been incorporated into dozens of tools that “can analyze patterns in existing code and generate new lines of code that are optimized for readability, efficiency, and error-free execution.”²⁰¹ In addition to writing new code, the AI systems are also good as code editors and reviewers, “identifying potential issues and helping teams maintain high-quality codebases.”²⁰² Given the potential to both write new code and improve the efficiency of existing code, it is no surprise that the potential for using AI has gained traction among software developers. “Use cases for software developers are also exploding—as of September, over 1.2 million developers had used GitHub Copilot’s technical preview. ChatGPT has also proven surprisingly adept at coding applications—from generating full code from text prompts (albeit often with many bugs) to bug-fixing code.”²⁰³ Microsoft launched GitHub Copilot in June 2022 and had over 400,000 subscribers in the first month of operation.²⁰⁴ The following January, “Microsoft Chief Executive Satya Nadella said that more than 1 million people had used Copilot to date.”²⁰⁵

As with other areas of AI interaction, the debate about AI in software is not focused on whether it will continue to be an increasingly significant aspect of the field, but the extent to which it will enhance the human coding profession or supplant it. Codex (yet another product

²⁰¹ *Top 17 Generative AI-based Programming Tools (for Developers)*, TOWARDS AI (Mar. 7, 2023), <https://pub.towardsai.net/top-17-generative-ai-based-programming-tools-for-developers-652977ac47b4> [<https://perma.cc/RVW6-JCHE>].

²⁰² *Id.*

²⁰³ *Id.*

²⁰⁴ *Microsoft Attracting Users to Its Code-writing, Generative AI Software*, ECON. TIMES (Jan. 23, 2023, 10:30 AM), https://economictimes.indiatimes.com/tech/technology/microsoft-attracting-users-to-its-code-writing-generative-ai-software/articleshow/97298152.cms?utm_source=contentofinterest&utm_medium=ext&utm_campaign=cppst [<https://perma.cc/SC5D-7R6A>].

²⁰⁵ *Id.*

from OpenAI) and similar tools are very good at solving complex coding problems.²⁰⁶ One software expert working with Codex even described its ability as “spooky to watch.”²⁰⁷ That same expert, however, does not believe that the tool is a direct threat to employment. Instead, Codex was viewed “as a tool that will end up boosting human productivity. It may even help a whole new generation of people learn the art of computers, by showing them how to write simple pieces of code, almost like a personal tutor.”²⁰⁸

Generative AI has an increasingly important place in new software code generation, in documenting the functions of the code that is created, in revising code for greater efficiency, in reviewing code for vulnerabilities and errors, and in other more technical situations.²⁰⁹ “Generative AI tools for code completion are becoming increasingly popular among developers, as they reduce the time and effort required to write complex code.”²¹⁰ According to GitHub’s own research, “[b]etween 60–75% of users reported they feel more fulfilled with their job, feel less frustrated when coding, and are able to focus on more satisfying work when using GitHub Copilot.”²¹¹ The self-reported study also identified that developers using GitHub Copilot finished a task in 55 percent less time than those developers completing the same task without the virtual assistant.²¹²

²⁰⁶ See Cade Metz, *A.I. Can Now Write Its Own Computer Code. That’s Good News for Humans*, N.Y. TIMES (Sept. 10, 2021), <https://www.nytimes.com/2021/09/09/technology/codex-artificial-intelligence-coding.html> [<https://perma.cc/7SZL-8CTN>].

²⁰⁷ *Id.* (quoting Tom Smith, a seasoned programmer who oversees an A.I. start-up called Gado Images).

²⁰⁸ *Id.* (“‘This is a tool that can make a coder’s life a lot easier,’ Mr. Smith said.”).

²⁰⁹ See *Generative AI for Code: What You Need to Know in 2023*, TURINTECH (Feb. 6, 2023), <https://www.turintech.ai/generative-ai-for-code-what-you-need-to-know-in-2023/> [<https://perma.cc/QQ9W-LXGZ>] (“During the process of writing code, developers are able to use AI-based tools to have suggestions for the code they write. These tools use natural language processing to analyse code, suggest relevant code snippets or keywords, and complete partial code with the full, syntactically correct language.”).

²¹⁰ *Id.*

²¹¹ Eirini Kalliamvakou, *Research: Quantifying GitHub Copilot’s Impact on Developer Productivity and Happiness*, GITHUB BLOG (Sept. 7, 2022), <https://github.blog/2022-09-07-research-quantifying-github-copilots-impact-on-developer-productivity-and-happiness/> [<https://perma.cc/Q3HZ-5LA5>].

²¹² *Id.* (“[The experiment] timed how long it took them to write an HTTP server in JavaScript. . . . [D]evelopers using GitHub Copilot took on average 1 hour and 11 minutes to complete the task, while the developers who didn’t use GitHub Copilot took on average 2 hours and 41 minutes.”).

The code being developed by the generative AI is helpful to the human developers, but like the other fields being discussed, as of this writing, the output remains a work in progress. “In the current generative AI landscape, code produced by generative tools is often far from ideal Code created by generative AI tools can even be faulty.”²¹³ As with other areas, the ability of generative AI tools to seem accurate creates an additional danger that its human users do not review it carefully enough. As a result, “StackOverflow temporarily banned answers generated by ChatGPT to be posted on the site, since it could flood the site with answers that may seem true at the surface level, but are incorrect.”²¹⁴

Despite the concerns, adoption is moving ahead at an accelerating pace. Take the example of CarMax. CarMax CIO Shamim Mohammad explains that his team continues to explore GitHub Copilot, estimating that the tool “could potentially generate up to 40% of their code.”²¹⁵ In an interview, he explained the benefits and the concerns. “This is evolving quickly [Y]ou have to make sure there’s no copyright infringement, fake content or malware embedded if you’re using it to create software. You can’t just plug that code in without oversight.”²¹⁶

Google is also moving forward with more autonomous code strategies. It has two complementary products. Alphabet’s Deepmind has launched Alphacode, which is intended to assist developers in the same manner as Microsoft’s GitHub Copilot.²¹⁷ Separately, Google has launched Project Pitchfork, as part of the Google AI Developer Assistance Program, with the goal to provide more autonomous code production.²¹⁸ “Pitchfork will be able to learn, write and rewrite code from scratch itself, learn different programming styles, and generate new code based on them, according to an internal project description.”²¹⁹

The predictions for the impact on code development range, but “Github CEO Thomas Dohmke expects up to 80 percent of code to be

²¹³ *Generative AI for Code: What You Need to Know in 2023*, *supra* note 209.

²¹⁴ *Id.*

²¹⁵ Robert Mitchell, *Enterprise IT Moves Forward—Cautiously—with Generative AI*, CIO (Mar. 7, 2023), <https://www.cio.com/article/463750/enterprise-it-moves-forward-cautiously-with-generative-ai.html> [<https://perma.cc/F4XS-FTSA>].

²¹⁶ *Id.*

²¹⁷ Matthias Bastian, *Google Looks into AI for Code, but Not Everyone Thinks It’s a Good Idea*, THE DECODER (Nov. 24, 2022), <https://the-decoder.com/google-looks-into-auto-acting-ai-for-code/> [<https://perma.cc/2RP7-JGZD>].

²¹⁸ *Id.*

²¹⁹ *Id.*

written by AI systems in the next five years.”²²⁰ Dohmke suggests that through GitHub, its AI service has already “written 40 percent of the code.”²²¹ Undoubtedly then, in the area of code generation, review, and documentation, AI has already transformed the landscape and will likely do so again as increasingly autonomous systems refine the relationship or even the conception of software authorship.

E. Product Design and Drug Design

Generative AI has also found a home in the development of virtual and rapid prototyping systems that allow the AI to provide a range of iterative and complex design recommendations across industries. Users provide parameters such as materials, size, weight, and operational constraints. Based on these parameters, the AI can generate dozens, hundreds, or thousands of iterations of designs and then select those that maximize the desired parameter, whether that is weight, strength, or manufacturing steps.²²² “Digital simulation and analysis are now so fast that designs can be evaluated in seconds—or even less.”²²³

Generative AI has been used to help designers with the aesthetic look of products, user interface designs, marketing effectiveness, lighter weight components, and much more.²²⁴ The AI can use the simulations to generate additional simulations. “Using artificial-intelligence techniques, these new generative design systems can explore a much larger universe of possible solutions, comparing the results of thousands of simulations to close in on a design that delivers the most favorable combination of attributes.”²²⁵

Although generative design services have already been proven to reduce the cost of production for some items or to achieve redesign goals such as to reduce the weight of an existing product, there remains

²²⁰ Matthias Bastian, *Github CEO Thinks AI Will Write Majority of Code in Just Five Years*, THE DECODER (Oct. 15, 2022), <https://the-decoder.com/github-ceo-thinks-ai-will-write-majority-of-code-in-just-five-years/> [<https://perma.cc/QXK6-HSVU>].

²²¹ *Id.*

²²² Graeme Fulton, *Generative UI Design: Einstein, Galileo, and the AI Design Process*, PROTOTYPR (Feb. 16, 2023), <https://prototypypr.io/post/generative-ai-design> [<https://perma.cc/MMW7-CF9G>].

²²³ Mickael Brossard et al., *How Generative Design Could Reshape the Future of Product Development*, MCKINSEY & CO. (Feb. 5, 2020), <https://www.mckinsey.com/capabilities/operations/our-insights/how-generative-design-could-reshape-the-future-of-product-development> [<https://perma.cc/MB7C-66PZ>].

²²⁴ Fulton, *supra* note 222.

²²⁵ Brossard et al., *supra* note 223.

reluctance to throw open the design process to these new techniques. “One of the first barriers is likely to be stakeholder acceptance of the resulting parts and products: generative algorithms produce designs that may be radically different from their human-designed predecessors. Some observers even find them ‘alien’ or disturbing.”²²⁶ The second common challenge is the danger that AI disturbs the institutional culture among its designers and relies on know-how that comes from the AI system rather than the gained wisdom of the engineers.²²⁷ Even if there is an openness for adoption, the third challenge remains that AI strategies need to be integrated into the broader design, procurement, and construction strategies of organizations.²²⁸

One particular area for potential success already stands out. This is in the field of pharmaceutical research, where the constraints and costs of design and engineering are often prohibitive. Research has shown that generative AI may play a role in creating “de novo drug design.” “De novo drug design (DNDD) refers to the design of novel chemical entities that fit a set of constraints using computational growth algorithms.”²²⁹

The capacity of generative AI to develop new medical treatments as well as medical devices and devices of all kinds demonstrates the powerful potential for generative design products to become increasingly ubiquitous in the planning and development of products. Using a wide range of AI techniques, including generative AI, new molecules can be designed and then tested through simulation to identify where the bioactivities meet the goals for the potential drug or treatment.²³⁰ “In the De novo drug design process, the deep learning models acts as autoencoder to generate an appropriate format for new chemical entities (NCE’s). Therefore, an embedment of autoencoder with multilayer perceptron classifier is also a value-added technique in the generation of NCE’s with predefined physicochemical properties.”²³¹ These tools can radically reduce the time needed to prototype new medicine and increase the likelihood of effectiveness.

²²⁶ *Id.*

²²⁷ *See id.*

²²⁸ *See id.*

²²⁹ Varnavas D. Mouchlis et al., *Advances in De Novo Drug Design: From Conventional to Machine Learning Methods*, 22 INT’L J. MOLECULAR SCI. 1, 2 (2021), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7915729/>.

²³⁰ *See id.* at 9–10 (“A shape-based generative approach for de novo drug design was developed using [a convolutional neural network] to generate novel molecules from a seed compound, its three-dimensional shape, and its pharmacophoric features.”).

²³¹ Suresh Dara et al., *Machine Learning in Drug Discovery: A Review*, 55 A.I. REV. 1947, 1974 (2022) (internal citations omitted).

V. INTELLECTUAL PROPERTY ISSUES EMERGING FROM AI-GENERATED CONTENT

While the uses to which generative AI can be put are essentially unlimited, the focus of the ChatGPT revolution emphasizes the role that generative AI plays in text, image, audiovisual, and related user-interface designs. In other words, the buzz centers on the potential transformation to popular culture; the AI tools that collectively may reshape the Internet and emerging metaverse as an entirely new medium.²³² There are clear lessons from the development of other communications-based technologies, so this section highlights the concerns that should be raised when adopting these new technologies as strategies to work with existing staff creators or in replacing staff creators. These include general considerations regarding intellectual property rights, speech-based torts such as defamation, and obligations regarding the use of new technologies in the workplace or educational environments.

A. Copyright in the Dataset

Under § 106 of the Copyright Act, copyright owners have the exclusive rights to “reproduce the copyrighted work in copies” and “to prepare derivative works based upon the copyrighted work.”²³³ “Copyright infringement is established when the owner of a valid copyright demonstrates unauthorized copying.”²³⁴

In January 2023, Getty Images brought suit against Stability AI for copyright infringement in the High Court of Justice in England for doing just that.²³⁵ Getty followed that claim with a similar lawsuit in the U.S. District Court for the District of Delaware.²³⁶ While this is the most high-profile of the early lawsuits, it is certainly not alone. Stability AI, along with DeviantArt and Midjourney, have been sued by visual artists who claim their works have been copied, transformed, and exploited by

²³² See generally MARSHALL McLuhan & QUENTIN FIORE, *THE MEDIUM IS THE MESSAGE: AN INVENTORY OF EFFECTS* (1967); Jon M. Garon, *Towards a Conceptual Framework of Entertainment Law for the Twenty-First Century*, 102 J. PAT. & TRADEMARK OFF. SOC'Y 203 (2022).

²³³ 17 U.S.C. § 106 (2022).

²³⁴ *Repp v. Webber*, 132 F.3d 882, 889 (2d Cir. 1997).

²³⁵ See *Getty Images Statement*, GETTY IMAGES (Jan. 17, 2023), <https://newsroom.gettyimages.com/en/getty-images/getty-images-statement> [<https://perma.cc/23KH-ZWAH>] (announcing litigation against Stability AI).

²³⁶ See Complaint, *Getty Images (US), Inc. v. Stability AI, Inc.*, 1:23-cv-00135 (D. Del. Feb. 2, 2023).

the search engines.²³⁷ A group of literary authors has begun a class action suit against ChatGPT.²³⁸ Plaintiffs representing open source software authors have sued Microsoft and OpenAI over the GitHub Copilot software writing tool, particularly over concerns that Copilot used open source software to generate AI software that does not have the open source licensing requirements.²³⁹ Comedian Sarah Silverman is lead plaintiff for yet another lawsuit involving the training of OpenAI as well as Meta’s LLaMA.²⁴⁰

As alleged by Getty, “Stability AI has copied at least 12 million copyrighted images from Getty Images’ websites, along with associated text and metadata, in order to train its Stable Diffusion model.”²⁴¹ These 12 million images were gleaned from among the “hundreds of millions” of images owned by Getty.²⁴² The lawsuit itself is limited to 7,216 works-in-suit that are each registered with the U.S. Copyright Office, a requirement for bringing suit in the U.S. to enforce the rights to the work.²⁴³ In addition to the 7,216 works-in-suit, Getty also claims copyright for the detailed title and caption that accompany many of the works.²⁴⁴ Although copyright law does not permit the registration of titles, slogans, or short phrases,²⁴⁵ the illustrations provided by Getty are rather substantial in length.²⁴⁶ The sea lion image described in the

²³⁷ See Complaint at 3, *Andersen v. Stability AI Ltd.*, 3:23-cv-00201 (N.D. Cal. Jan. 13, 2023) (“Stable Diffusion uses the Training Images to produce seemingly new images through a mathematical software process. These ‘new’ images are based entirely on the Training Images and are derivative works of the particular images Stable Diffusion draws from when assembling a given output.”).

²³⁸ See Complaint, *Tremblay v. OpenAI, Inc.*, 3:23-cv-03223 (N.D. Cal. June 28, 2023).

²³⁹ See Complaint, *Doe v. GitHub, Inc.*, 3:22-cv-06823 (N.D. Cal. Nov. 3, 2022).

²⁴⁰ See Complaint, *Silverman v. OpenAI, Inc.*, 3:23-cv-03416 (N.D. Cal. July 7, 2023).

²⁴¹ Complaint at ¶ 8, *Getty Images (US), Inc. v. Stability AI, Inc.*, 1:23-cv-00135 (D. Del. Feb. 2, 2023).

²⁴² *Id.* ¶ 18; see also *id.* ¶ 21 (“Getty Images has more than 500,000 contributors (80,000 of which are exclusive to Getty Images), over 300 premium content partners, more than 115 staff photographers, videographers, and other content experts . . .”).

²⁴³ *Id.* ¶ 24; see 17 U.S.C. § 411 (2022) (requiring preregistration or registration of work as a condition of suit).

²⁴⁴ Complaint, *Getty Images (US), Inc.*, 1:23-cv-00135 ¶ 25.

²⁴⁵ See 17 U.S.C. § 102(b) (2022); U.S. COPYRIGHT OFF., *supra* note 53 (“Words and short phrases, such as names, titles, and slogans, are uncopyrightable because they contain an insufficient amount of authorship.”).

²⁴⁶ Complaint, *Getty Images (US), Inc.*, 1:23-cv-00135 ¶ 25.

For example, for the image below, the accompanying title reads, “Malnourished Sea Lions Continued To Be Rescued Off California Shores” and the accompanying caption reads: “A sick and

complaint includes a caption that is 90 words in length, an amount that is likely well beyond the minimum authorial creativity necessary for copyright to protect the writing.²⁴⁷

“There are two main components of this *prima facie* case of infringement: ‘a plaintiff must first show that his work was actually copied and then must show that the copying amounts to an improper or unlawful appropriation.’”²⁴⁸ The claim by Getty seems to meet the *prima facie* requirements. First, it alleges valid registrations for the works-in-suit as well as millions more that it has not included in the litigation. Second, presuming the accuracy of the images shown in the complaint, the Stable Diffusion illustrations include the actual watermarks and trademarks of Getty in some of their reproductions. Whether or not the mark is itself copyrightable does not matter. It provides substantial direct evidence of copying.²⁴⁹

As a result, unless Stability AI can establish that the massive copying of the Getty images constitutes fair use, Stability AI will be liable for copyright infringement.

The copyright fair use test originated as a common law constraint upon the exclusive rights conveyed to copyright holders and was codified as part of the 1976 Copyright Act. The statute provides:

Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use),

malnourished sea lion pup sits in an enclosure at the Marine Mammal Center on March 18, 2015 in Sausalito, California. For the third winter in a row, hundreds of sick and starving California sea lions are washing up on California shores, with over 1,800 found and treated at rehabilitation centers throughout the state since the beginning of the year. The Marine Mammal Center is currently caring for 224 of the emaciated pups.”

²⁴⁷ See, e.g., *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1990) (requiring merely a modicum of creativity for copyright protection).

²⁴⁸ *Castle Rock Ent., Inc. v. Carol Publ'g Grp., Inc.*, 150 F.3d 132, 137 (2d Cir. 1998) (quoting *Laureyssens v. Idea Grp., Inc.*, 964 F.2d 131, 139–40 (2d Cir. 1992)) (internal emendations, quotation marks, and citations omitted).

²⁴⁹ The use of non-copyrightable elements of a work to highlight the duplication process is sometimes called “striking similarity.” See *Three Boys Music Corp. v. Bolton*, 212 F.3d 477, 485 (9th Cir. 2000) (“[I]n the absence of any proof of access, a copyright plaintiff can still make out a case of infringement by showing that the songs were ‘strikingly similar.’”).

scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.²⁵⁰

“From the infancy of copyright protection, some opportunity for fair use of copyrighted materials has been thought necessary to fulfill copyright’s very purpose, ‘[t]o promote the Progress of Science and useful Arts’”²⁵¹ Yet despite the importance of fair use, since its inception, the emphasis has been on the importance that “one must not put manacles upon science.”²⁵² Justice Joseph Story, who first articulated the fair use doctrine had explained that “[e]very book in literature, science and art, borrows, and must necessarily borrow, and use much which was well known and used before.”²⁵³

In applying the fair use test, the preamble and all four prongs must be considered and weighed holistically.²⁵⁴ Here, Getty alleges that Stability AI has raised in excess of \$1 billion in equity investment, so that even if it is not charging the general public for the use of the image generation, the company itself is undertaking a for-profit, highly commercial enterprise. Secondly, the nature of the copyrighted works

²⁵⁰ 17 U.S.C. § 107 (2022).

²⁵¹ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 575 (1994) (quoting U.S. CONST. art. I, § 8, cl. 8).

²⁵² *Id.* (quoting *Carey v. Kearsley*, 170 Eng. Rep. 679, 681 (K.B. 1803)).

²⁵³ *Id.* (quoting *Emerson v. Davies*, 8 F. Cas. 615, 619 (C.C.D. Mass. 1845) (No. 4,436)).

²⁵⁴ *Id.*

are highly creative photographs rather than fact-based works with merely thin copyright protection.²⁵⁵

The first factor, however, has been given an independent role. Under *Campbell v. Acuff-Rose Music, Inc.*, a separate inquiry must be made regarding the extent to which “the allegedly infringing work ‘merely supersedes’ the original work ‘or instead adds something new, with a further purpose or different character, altering the first with new . . . meaning [] or message,’ in other words ‘whether and to what extent the new work is “transformative.””²⁵⁶

In *Andy Warhol Foundation for the Visual Arts, Inc. v. Goldsmith*,²⁵⁷ the Supreme Court revisited the meaning of the first fair use factor since it adopted *Campbell*. Like the dispute involving Getty and Stable Diffusion, the conflict between the estate of Andy Warhol and photographer Lynn Goldsmith involves the misuse of photographs. At the heart of the transformative use debate is the extent to which widespread copying for a purpose other than to supplant the original should constitute fair use.

The Court explained that “the first fair use factor . . . focuses on whether an allegedly infringing use has a further purpose or different character, which is a matter of degree, and the degree of difference must be weighed against other considerations, like commercialism.”²⁵⁸ Going further, the court reinforced the limited role that fair use plays when a substantially similar, copied work becomes a market substitute for the work that was copied. “[T]he first factor relates to the problem of substitution—copyright’s *bête noire*. The use of an original work to achieve a purpose that is the same as, or highly similar to, that of the original work is more likely to substitute for, or ‘supplan[t],’ the work.”²⁵⁹

The Supreme Court went even further, pointing out that that §106 grants to the original copyright holder the right to transform, recast, or adapt the work. The Court notes that “[t]o be sure, this right is ‘[s]ubject to’ fair use. The two are not mutually exclusive. But an overbroad concept of transformative use, one that includes any further

²⁵⁵ See *Blanch v. Koons*, 467 F.3d 244, 257 (2d Cir. 2006) (“As we recently explained, although ‘the creative nature of artistic images typically weighs in favor of the copyright holder,’ ‘the second factor may be of limited usefulness where the creative work of art is being used for a transformative purpose.’”) (quoting *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 612 (2d Cir. 2006)).

²⁵⁶ *Castle Rock Ent., Inc. v. Carol Pub. Grp., Inc.*, 150 F.3d 132, 142 (2d Cir. 1998) (quoting *Campbell*, 510 U.S. at 579).

²⁵⁷ See *Andy Warhol Found. Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508 (2023).

²⁵⁸ *Id.* at 525.

²⁵⁹ *Id.* at 528.

purpose, or any different character, would narrow the copyright owner's exclusive right to create derivative works."²⁶⁰

The *Goldsmith* opinion reverses a line of cases that had largely equated technological innovation with transformative use. Although the *Goldsmith* court never mentions generative AI, the reemphasis on market substitution will certainly have important implications for the use of generative AI to usurp the marketplace for human-authored works in those situations where the output is substantially similar to that of the plaintiff's source work.

The Getty allegations, if proven correct at trial, illustrate this position. The other three fair use factors appear to strongly favor Getty, at least based on the allegations contained in the complaint. Factor two looks to the nature of the copyrighted work. A work that is highly creative in nature, fictional, like a poem or novel, or a work of art is generally entitled to broad copyright protection, while a work that is largely descriptive or factual in nature generally receives only thin copyright protection.²⁶¹ Photographs, in general, are typically understood to receive broad copyright protection,²⁶² and many of the images in the Getty catalog are highly creative in form, content, or subject matter.

Factor three addresses the amount of the work taken both qualitatively and quantitatively. Here each image was copied completely, indexed, and reproduced as part of the training process. Stability AI "has copied entire works. Though this conclusion does not preclude a finding of fair use, it militates against such a finding."²⁶³

The final factor is the potential and actual impact on the market for the plaintiff's copyrighted works. "Fair use, when properly applied, is limited to copying by others which does not materially impair the marketability of the work which is copied."²⁶⁴ "It requires courts to consider not only the extent of market harm caused by the particular actions of the alleged infringer, but also 'whether unrestricted and widespread conduct of the sort engaged in by the defendant . . . would result in a substantially adverse impact on the potential market' for the original."²⁶⁵

²⁶⁰ *Id.* at 529 (quoting 17 U.S.C. § 107 (2022)).

²⁶¹ *See* *Stewart v. Abend*, 495 U.S. 207, 237 (1990) ("In general, fair use is more likely to be found in factual works than in fictional works.").

²⁶² *See* *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53 (1884).

²⁶³ *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 926 (2d Cir. 1994).

²⁶⁴ *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 566–67 (1985).

²⁶⁵ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 (quoting *MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT*, § 13.05(A)(4), at 13-102.61).

Getty alleges that the purpose of releasing Stable Diffusion into the marketplace is to compete directly for sales with Getty.²⁶⁶ Undoubtedly, Stability AI will respond that its images are not sold but rather provided free to the public without charge. It has not, thus far, detailed how the investors in Stability AI will receive a return on their investment, but it is certainly plausible that Stability AI's revenue will come from licensing aspects of its generative AI for other commercial uses or provide services that do not include charging the public for the images generated by Stable Diffusion. A similar argument was made and rejected in *A&M Records v. Napster, Inc.*²⁶⁷ The lack of direct sales has not proven sufficient to allow wholesale copying of a company's copyrighted library in order to give those works away for free.

While the *Getty* case against the present output of Stable Diffusion should be relatively straight forward for the plaintiff, improvements to the Stable Diffusion training model should enable the company to prohibit works with actual watermarks or artist signatures from appearing in the finished results. The generative AI can be trained to identify such indicia of ownership and remove such artifacts from the images it creates. Without those elements of striking similarity, it may be much harder to demonstrate that any particular image created by a generative AI was copied directly from a particular image among the billions included in the archival database.

The copyright claim for the next generation of generative AI, therefore, will turn less on the striking similarity between plaintiff's works and defendant's works and more on the fair use for the intermediate copying of the images ingested to train the model.

Copyright law has developed the term "intermediate copying" to describe potential infringers that copy but do not reproduce works

²⁶⁶ Complaint ¶¶ 9, 52, *Getty Images (US), Inc. v. Stability AI, Inc.*, 1:23-cv-00135 (D. Del. Feb. 2, 2023).

²⁶⁷ *A&M Recs., Inc. v. Napster, Inc.*, 239 F.3d 1004, 1017 (9th Cir. 2001), *as amended* (Apr. 3, 2001), *aff'd*, 284 F.3d 1091 (9th Cir. 2002) ("Having digital downloads available for free on the Napster system necessarily harms the copyright holders' attempts to charge for the same downloads."); *see also* *L.A. Times v. Free Republic*, 54 U.S.P.Q.2d 1453, 2000 WL 565200 (C.D. Cal. Apr. 4, 2000).

for the public.²⁶⁸ In *Sega Enterprises, Ltd. v. Accolade, Inc.*,²⁶⁹ the Ninth Circuit addressed the question regarding whether the copying during an intermediary step of creating computer software could violate the copyright even if the copyrighted work did not ultimately appear in the final product.²⁷⁰ The court found that it did. “[I]ntermediate copying of computer object code may infringe the exclusive rights granted to the copyright owner in section 106 of the Copyright Act regardless of whether the end product of the copying also infringes those rights.”²⁷¹

In *Sega*, the court treated the intermediate copying of the protected computer code for the purpose of extracting the non-copyrightable aspects of the functional computer code to be fair use. In the case of the database training, Stability AI and other generative AI training companies will assert that they are not copying for the purposes of reproduction but only for the purpose of labeling, deconstruction, and training.

There may be an important distinction between the training databases for software, natural language, and images. In the case of software, *Sega* and similar cases recognize that the code incorporates unprotected functions that cannot be masked by the overlay of copyrighted works.²⁷² In the case of natural language processing, it is likely that the software deconstructs longer texts into short phrases to help predict the words and phrases that tend to follow a series of earlier words and phrases. Although the entirety of a work is being copied, the use of the natural language generative AI should result in the AI “speaking.” While an AI that reproduced or recited the works of authors

²⁶⁸ *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1518 (9th Cir. 1992), *as amended* (Jan. 6, 1993) (“On its face, that language unambiguously encompasses and proscribes “intermediate copying.”); *see* *Walker v. U. Books, Inc.*, 602 F.2d 859, 864 (9th Cir. 1979) (“That an infringing copy may be produced in a medium different than that of the protected work is not, in itself a bar to recovery, nor does the fact that the blueprints themselves were never sold for profit eliminate the possibility of an award of statutory damages for infringement under the Act.”).

²⁶⁹ *Sega Enters. Ltd.*, 977 F.2d 1510.

²⁷⁰ *Id.* at 1519.

²⁷¹ *Id.*

²⁷² *See id.* at 1527–28 (“[W]here disassembly is the only way to gain access to the ideas and functional elements embodied in a copyrighted computer program and where there is a legitimate reason for seeking such access, disassembly is a fair use of the copyrighted work, as a matter of law.”); *see also* *Sony Comput. Ent., Inc. v. Connectix Corp.*, 203 F.3d 596, 603 (9th Cir. 2000) (“[C]opying nonetheless could be protected as a fair use if it was ‘necessary’ to gain access to the functional elements of the software itself. . . . [T]his idea/expression distinction poses ‘unique problems’ because computer programs are ‘in essence, utilitarian articles—articles that accomplish tasks.’”).

still under copyright would undoubtedly be copyright infringement, such an AI that selected only short words or phrases from popular literature would not be infringing the copyright by its output and could arguably be acting within the scope of fair use by deconstructing the text for training purposes.

In the context of natural language training, the process for developing the massive training databases and the legal test of fair use began with the Google Book Project, wherein Google promised to benevolently scan all the world's books.²⁷³ "It was the most ambitious library project of our time—a plan to scan all of the world's books and make them available to the public online."²⁷⁴ The project eventually scanned at least 30 million books,²⁷⁵ which is an impressive corpus from which to begin a natural language training model, a use that was never explicitly part of the litigation. The works that were in the public domain were never at issue, since Google and all others are free to reproduce those works.

For the works protected by copyright, the Second Circuit asked two very narrow questions, "whether Google's search and snippet views functions satisfy the first fair use factor with respect to Plaintiffs' rights in their books" and "whether these functions might infringe upon Plaintiffs' derivative rights"²⁷⁶ In answer to these questions, the Second Circuit found that making a text available to the public through its search function was a highly transformative use.²⁷⁷ The court also emphasized that to the extent that facts were then published from the database, those facts, such as the year Franklin D. Roosevelt contracted polio, are outside of copyright protection.²⁷⁸

²⁷³ See Tim Wu, *What Ever Happened to Google Books?*, THE NEW YORKER (Sept. 11, 2015), <https://www.newyorker.com/business/currency/what-ever-happened-to-google-books> [<https://perma.cc/S HQ9-U6WG>].

²⁷⁴ *Id.*

²⁷⁵ *Id.*

²⁷⁶ *Authors Guild v. Google, Inc.*, 804 F.3d 202, 216 (2d Cir. 2015); see also *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014).

²⁷⁷ *Authors Guild v. Google*, 804 F.3d at 216; *Authors Guild, Inc. v. HathiTrust*, 755 F.3d at 98.

²⁷⁸ *Authors Guild v. Google*, 804 F.3d at 224.

By entering "Roosevelt polio" in a Google Books search, the student would be taken to (among numerous sites) a snippet from page 31 of Richard Thayer Goldberg's *The Making of Franklin D. Roosevelt* (1981), telling that the polio attack occurred in 1921. This would satisfy the searcher's need for the book, eliminating any need to purchase it or acquire it from a library. But what the searcher derived from the snippet was a historical fact. Author Goldberg's

The *Goldsmith* Court highlighted the continued importance of including the focus of the justification for the copying as part of the first prong of fair use. The Court explained “the first factor also relates to the justification for the use. In a broad sense, a use that has a distinct purpose is justified because it furthers the goal of copyright, namely, to promote the progress of science and the arts, without diminishing the incentive to create.”²⁷⁹ The Court quoted *Authors Guild v. Google* to emphasize the point.

The more the appropriator is using the copied material for new, transformative purposes, the more it serves copyright’s goal of enriching public knowledge and the less likely it is that the appropriation will serve as a substitute for the original or its plausible derivatives, shrinking the protected market opportunities of the copyrighted work.²⁸⁰

As a practical matter, the potential for copyright infringement liability for natural language search will likely depend on the extent to which the texts are substantially reproduced by the generative AI rather than on the widespread copying that went into the training. Nonetheless, the *Google* decision was unable to take into account the billions of dollars that Google was going to earn from its natural language voice assistant. In the current economic climate, the values being ascribed to generative AI will mandate that the courts recognize the significant value of the copyrighted works that are being used without authorization to earn profits by parties copying without authorization.

Another way to ask the question, as highlighted by *Goldsmith*, is whether this is a transformative use or a derivative rights use. If there is a ready market for licensing source material to AI models, then usurpation of that market interferes with the derivative rights of the copyright holder. If there is no such market, then the use cannot be a derivative right and the use must therefore be transformative.²⁸¹

More broadly, copyright and fair use are not really equipped to regulate such sweeping wealth transfers, so it may again be time for Congress to consider mechanisms for the blanket licensing of

copyright does not extend to the facts communicated by his book. It protects only the author’s manner of expression.

²⁷⁹ Andy Warhol Found. Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 531 (2023).

²⁸⁰ *Id.* at 531 (quoting *Authors Guild v. Google*, 804 F.3d at 214).

²⁸¹ See, e.g., *Ringgold v. Black Ent. Tel., Inc.*, 126 F.3d 70 (2d Cir. 1997) (discussing the role of an existing market in diminishing claims for fair use).

copyrighted works to reflect the value they are giving to the tech industry and the public that relies on these services.²⁸²

The question for image generation, however, is somewhat different. There is less functionality in a deconstructed image than in software or natural language text. There remains the factual depiction that a picture of a cat contains a cat, but each cat depicted retains its unique artistic expression. More importantly, the production of new works of imagery are often quite recognizable as being reproductions or derivative works from pre-existing works. So unlike software and natural language content, the imagery databases will remain the most susceptible to claims of copyright infringement.

The copyright concerns about implications of training the generative AI using unauthorized content only exist to the extent the publisher has done so. Among the services becoming available are generative AI engines that can be trained using the licensee's own content. If a large database holder such as Adobe, Getty, Bloom, Thomson Reuters, RELX Group (formerly Reed Elsevier), or another large content holder were to start using its proprietary content on its AI, there would be far less likelihood that the resulting work would violate copyright, but ultimately this result would depend on the particular licenses under which the various rights were gathered by that publisher. Artists and authors who provide content into these large data management companies may also have concerns about license provisions adopted before the development of the new AI tools.²⁸³

B. Copyright by the End User and the Commercial Licensee

Although there is likely recourse for Getty against Stability AI for its current and rather blatant publications of Getty images, the

²⁸² See Wu, *supra* note 273 (“A federal judge sided with the critics and threw out the 2008 settlement, adding that aspects of the copyright issue would be more appropriately decided by the legislature. ‘Sounds like a job for Congress,’ James Grimmelman, a law professor [now at Cornell Law School] . . .”). See generally Jonathan Band, *The Google Library Project: Both Sides of the Story*, PLAGIARY (2006), <https://quod.lib.umich.edu/p/plag/5240451.0001.002/--google-library-project-both-sides-of-the-story?rgn=main;view=fulltext> [https://perma.cc/RV2E-RUPK].

²⁸³ See Ron Miller, *Adobe Promises Artists Will be Compensated Fairly with New Generative AI Product, but is Fuzzy on Details*, TECHCRUNCH (Mar. 21, 2023), <https://techcrunch.com/2023/03/21/adobe-promises-artists-will-be-compensated-fairly-with-new-generative-ai-product-but-is-fuzzy-on-details/> [https://perma.cc/3HD2-FQ8J]; see also *New York Times Co. v. Tasini*, 533 U.S. 483 (2001) (interpreting text of non-employee authors contracts only transferred first publication rights).

question remains whether the hosts of embedded generative AI services obtained from Stability AI or OpenAI will be liable for copyright infringement or whether the users of those services, in turn, will be liable in the event that the AI creates a work that is substantially similar to that of a registered copyright holder.

The potential infringement scenario will be fairly common. For example, assuming that web design companies can survive the automation services, such a company builds a website for a florist. In this hypothetical, Acme Florist contracts with Zenith Web Design, LLC to redesign its website to feature images of Seattle where the florist is based and flower arrangements for each of the types of flowers sold by Acme. Zenith uses a free generative AI tool to iterate two dozen potential website layouts and selects four to send to the client. Zenith also generates 100 photographs to populate the website in the manner requested. Acme rejects a few of the images because they do not look like the products Acme sells and authorizes the use of the remainder. None of the images are based on the actual bouquets and arrangements made and sold by Acme.

Among the 100 photographs published to the Acme Florist website, one of those pictures looks strikingly similar to an image of flowers taken by commercial photographer Gloria Ramirez, which she published as part of a gallery photography show. The work, entitled *The Bouquet*, was registered with the copyright office. *The Bouquet*, as exhibited, was a signed, single copy.²⁸⁴

In this scenario, Ramirez will have a cause of action against both Acme Florist and Zenith Web Design for the violation of her exclusive rights in her copyrighted image and for failing to provide attribution to the work as protected under §106A.²⁸⁵ Copyright law does not require that a party's *mens rea* or intent is relevant to the liability, but instead provides that in the measure of statutory damages, the damages can be reduced.²⁸⁶ As a practical matter, if a copyright infringer responds

²⁸⁴ A photograph will meet the definition as a “work of visual art” when “(2) a still photographic image produced for exhibition purposes only, existing in a single copy that is signed by the author, or in a limited edition of 200 copies or fewer that are signed and consecutively numbered by the author. 17 U.S.C. § 101; *see also* 17 U.S.C. § 106A.

²⁸⁵ *See* 17 U.S.C. § 501(a) (“Anyone who violates any of the exclusive rights of the copyright owner as provided by sections 106 through 122 or of the author as provided in section 106A(a), . . . is an infringer of the copyright or right of the author, as the case may be.”).

²⁸⁶ 17 U.S.C. § 504(c)(2) (“In a case where the infringer sustains the burden of proving, and the court finds, that such infringer was not aware and had no reason to believe that

expeditiously to a takedown request, then the copyright owner is unlikely to sue for what could be very modest damages.

There are also the notice and takedown systems under § 512 of the Copyright Act,²⁸⁷ but these would not apply in this situation. An Internet service provider (ISP) can limit its liability for the content published on its site by third parties.²⁸⁸ In this hypothetical, however, the content is being selected and published at the direction of the website owner rather than by third parties. If instead, customers of the florist were able to publish their pictures, and a customer published a photograph protected by the copyright of a third person, then the florist could avoid liability by adhering to the safe harbor requirements of § 512.²⁸⁹

As between Acme Florist and Zenith Web Design, the liability will likely depend on the contractual representations and warranties as well as the limits on liability placed in the written agreements. Savvy web design companies will likely expressly waive any obligation to determine the copyright ownership of the works provided to their clients and obligate the clients to conduct such due diligence. These contracts may also include limitations on liability to the contracting party and indemnification and defense provisions to require that the client pay any fees resulting from the publication of unauthorized third-party works.

The copyright litigation scenario is not significantly different than the present situation for small businesses who rely on third-party vendors to populate their websites with imagery. The only significant difference is the extent to which the parties inaccurately believe that simply because the image was created by a generative AI, that image must be in the public domain and free to use. As the Getty Image litigation highlights, there will be continuing concerns that particular works created by the generative AI are actually near reproductions of works that were placed into the training database and will therefore still be subject to copyright infringement actions when used as output.

his or her acts constituted an infringement of copyright, the court in its discretion may reduce the award of statutory damages to a sum of not less than \$200.”).

²⁸⁷ See 17 U.S.C. § 512(c) (2022) (providing that an internet service provider will not be liable for monetary or equitable relief for copyright infringement resulting from the posting or storage of content at the direction of a user of the service provider’s network or system).

²⁸⁸ See *id.*

²⁸⁹ See 17 U.S.C. § 512.

C. Trade Secrets and Nonpublic Information

While much of the focus on generative AI has focused on the copyright aspects of the databases and the uses, there are also business concerns regarding the impact the service has on trade secret protection for enterprises.²⁹⁰ Generative AI has consequences for the discovery of existing trade secrets and the maintenance of trade secrets that exist in a company's current enterprise.

A trade secret is confidential, commercially valuable information that provides a company with a competitive advantage, such as customer lists, methods of production, marketing strategies, pricing information, and chemical formulae. (Well-known examples of trade secrets include the formula for Coca-Cola, the recipe for Kentucky Fried Chicken, and the algorithm used by Google's search engine.) To succeed in the global marketplace, U.S. firms depend upon their trade secrets, which increasingly are becoming their most valuable intangible assets.²⁹¹

The Restatement of Torts § 757 provides that “[a] trade secret may consist of 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.”²⁹² To be a trade secret under the state and federal protection, the information must be generally kept secret from the public, have economic value because of its secrecy, and be protected through the

²⁹⁰ David Levine, *Generative Artificial Intelligence and Trade Secrecy*, 3 J. FREE SPEECH L. 559, 564 (2023) (“Generative AI raises some significant questions for the core trade secret misappropriation doctrine that has governed trade secret law for decades.”); Jennifer Bennett et al., *Trade Secrets and Generative AI: Protective Measures in an Evolving Technological Landscape*, JD SUPRA (June 12, 2023), <https://www.jdsupra.com/legalnews/trade-secrets-and-generative-ai-1027321/> [<https://perma.cc/D6ZS-EYUH>].

²⁹¹ BRIAN YEH, CONG. RSCH. SERV., R43714, PROTECTION OF TRADE SECRETS: OVERVIEW OF CURRENT LAW AND LEGISLATION Summary (Apr. 22, 2016), <https://sgp.fas.org/crs/secrecy/R43714.pdf>.

²⁹² RESTATEMENT (FIRST) OF TORTS § 757 cmt. b (AM. L. INST. 1939). See generally *E.I. duPont deNemours & Co. v. Christopher*, 431 F.2d 1012 (5th Cir. 1970).

reasonable efforts of the party asserting the trade secret.²⁹³ State law has long recognized trade secret protection and in 2016, the state law protection was augmented by the Defend Trade Secrets Act (DTSA),²⁹⁴ a federal statute that largely duplicates and expands upon the widely adopted Uniform Trade Secrets Act. The DTSA provides a private right of action for “[a]n owner of a trade secret that is misappropriated . . . if the trade secret is related to a product or service used in, or intended for use in, interstate or foreign commerce.”²⁹⁵

As a general matter, a natural language and design system that answers users’ questions without restriction and iteratively generates new explanations and designs is likely to regularly “reinvent” or “rediscover” information treated as trade secrets on a regular basis. As soon as the information is created and divulged by the generative AI, that information is available to the public and the essential component of secrecy will be lost. Out of concerns that proprietary information will be leaked, “organizations including banks like Bank of America, Citi, Deutsche Bank, Goldman Sachs, Wells Fargo, and JPMorgan, as well as Walmart and telecom giant Verizon have also restricted their staff from accessing ChatGPT.”²⁹⁶

While this is no different than the threat that exists from competitors and from academics who strive to find and publish new knowledge, the scope of the generative AI content has the potential to be staggering. Still, this concern is a natural byproduct of increased general knowledge in the world. Indeed, new trade secrets are as likely to be found through improvements in the know-how as are lost because the older, protected ways of doing business are no longer secret.

Trade secrets are also lost, however, when the information is made public by the party holding the trade secret or when it fails to

²⁹³ See UNIF. TRADE SECRETS ACT § 1(4) (NAT’L CONF. OF COMM’RS OF UNIF. STATE L. 1985).

“Trade secret” means information, including a formula, pattern, compilation, program, device, method, technique, or process, that: (i) derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable by proper means by, other persons who can obtain economic value from its disclosure or use, and (ii) is the subject of efforts that are reasonable under the circumstances to maintain its secrecy.

²⁹⁴ 18 U.S.C. §§ 1831–39 (2016).

²⁹⁵ 18 U.S.C. § 1836(b)(1).

²⁹⁶ Ivan Mehta, *Apple Reportedly Limits Internal Use of AI-powered Tools Like ChatGPT and GitHub Copilot*, TECHCRUNCH (May 19, 2023, 9:55 AM), <https://techcrunch.com/2023/05/19/apple-reportedly-limits-internal-use-of-ai-powered-tools-like-chatgpt-and-github-copilot/> [https://perma.cc/NKE7-KTCM].

reasonably protect the trade secret information from disclosure. Here the exuberance for generative AI may come to plague large enterprises that seek to manage their trade secret portfolios.

Whenever an employee, or other individual with access to a trade secret, discloses that information without having taken steps to ensure the information will be treated as confidential by the recipient, that information may lose its trade secret status. While courts generally look for patterns rather than mere instances of inadvertent disclosure, the public disclosure of a trade secret generally ends its secrecy. As a result, if an employee uses proprietary information in a natural language query, that disclosure has the effect of uploading the information regarding the trade secret into the AI's database for future training and usage.

Contractual limitations, therefore, are essential, at a minimum, to create legal protection for the information being provided to the generative search engine.²⁹⁷ The particular language of each service varies considerably and there is likely little or no ability of the enterprise to police the information being uploaded into the natural language database. Worse, the practice of sharing trade secret information with the query engine renders the trade secret public, even if the AI does not republish that information in answers to third parties.

For employers, this requires quick and aggressive training to remind staff and vendors that external generative AI tools cannot be used for sensitive data. For example, a public company that is considering a merger or acquisition of another company might well be prohibited from disclosing such information under the Securities Exchange Act and the associated requirements for specified disclosures in advance of any acquisition. The employees should treat the natural language generator like an outside analyst and provide to it only that information that one could provide to such analysts. While the information that could potentially be gleaned from the massive database available to a generative AI would be useful, such inquiries could also constitute disclosure or even announcements and should be avoided.

Software is another field that relies heavily on trade secret protection. Although copyright protection is available for software its effect is limited. The functional nature of code, the rapid pace of

²⁹⁷ See UAB “Planner5D” v. Facebook, Inc., No. 19-CV-03132-WHO, 2020 WL 4260733, at *7 (N.D. Cal. July 24, 2020) (“Planner 5D’s combined effect of structural and legal barriers ‘consistently walls off both the location and content of the trade secrets at issue here, and under its Terms of Service, separately prohibits circumventing these protections via crawling, scraping, or otherwise accessing its data files.’”).

revision, registration requirements that are prerequisite to litigation, and similar practical considerations make copyright a poor fit for practical protection from unfair competition.²⁹⁸ Instead, companies often rely on trade secret protection to stop direct competitors from unfairly taking advantage of key software components.²⁹⁹

Proprietary software will often have the requisite economic value because of its secrecy.³⁰⁰ Therefore, to obtain trade secret protection, the owners of the software simply must protect the source code through the reasonable efforts of the party asserting the trade secret.³⁰¹ “Companies that own proprietary software typically safeguard their trade secret rights in source code by maintaining the source code in confidence, and releasing it only selectively and under strict restrictions against copying, reverse engineering, transfer and disclosure.”³⁰² Source code—which is the human authored software written in computer languages such as Java, C++, or others—does not need to be distributed.³⁰³ On the other hand, the object code is the machine-readable version of the software which is often distributed as part of the software business.³⁰⁴ Only that object code that is either used only internally or distributed under a highly protective contractual agreement to a narrow pool of end users will likely afford trade secret protection.³⁰⁵ When distributing it outside the enterprise, the licensed source code and object code should be “subject to restrictions against copying, reverse engineering, transfer and disclosure.”³⁰⁶

The popularity of GitHub’s Copilot AI tools has raised additional concerns about software trade secret protection. Apple is

²⁹⁸ See *Rivendell Forest Prods., Ltd. v. Georgia-Pac. Corp.*, 28 F.3d 1042, 1044 (10th Cir. 1994); *Altavion, Inc. v. Konica Minolta Sys. Lab’y, Inc.*, 171 Cal. Rptr. 3d 714, 738 (Cal. Ct. App. 2014) (“algorithms and source code . . . information is unquestionably protectable by trade secret law.”); Elizabeth A. Rowe & Nyja Prior, *Procuring Algorithmic Transparency*, 74 ALA. L. REV. 303, 337 (2022) (“[Trade secrets] protect data, software, and algorithms in these technologies. In general, trade secret rights cover operability and functionality of devices, and algorithmic models are often within one of these categories.”).

²⁹⁹ See *Rivendell Forest Prods.*, 28 F.3d at 1044.

³⁰⁰ See *Altavion, Inc.*, 171 Cal. Rptr. 3d at 738–39.

³⁰¹ See, e.g., UNIF. TRADE SECRETS ACT § 1(4) (NAT’L CONF. OF COMM’RS OF UNIF. STATE L. 1985).

³⁰² Lonnie Finkel, *How to Protect Your Company’s Software Assets*, FINKEL L. GRP., <https://finkellawgroup.com/protect-company-software-assets/> [https://perma.cc/DZ3K-YAN8].

³⁰³ *Id.*

³⁰⁴ *Id.*

³⁰⁵ *Id.*

³⁰⁶ *Id.*

among the most high-profile of companies that has placed restrictions on use of Copilot out of fear of exposing its proprietary software.³⁰⁷ Samsung has done the same.³⁰⁸ Amazon is restricting its engineers to using internal AI tools.³⁰⁹ By uploading software to Copilot for review, software engineers are exporting the software to a common data set and not taking reasonable steps to protect the trade secrets represented by the software. Potentially, that software could make it into third-party products. But even if it doesn't, the failure to maintain confidentiality could jeopardize the trade secret protection against other claims and misuse.

Even more than challenges in the field of software, the intersection of trade secret protection and content design creates a unique challenge. Interactive sites often rely on nonpublic data to help inform the software as to the public information that should be presented.³¹⁰ Potentially, this nonpublic information can be protected by trade secret if it otherwise meets the definitional test. Then, assuming that the hidden data are trade secrets, third-party efforts to acquire such information in violation of the Terms of Service should constitute trade secret theft.³¹¹

Nonpublic information also includes all customer, client, employee, student, and patient data subject to one or more privacy regimes, including protected health records or academic records. A college coach using a natural language search algorithm should not provide it—or the public—private information regarding the student-athlete information despite the potential to use the algorithm to model player performance.³¹² Such concerns are not merely hypothetical. “For example, in February 2023 an AI-based chatbot was prohibited by the

³⁰⁷ Aaron Tilley & Miles Kruppa, *Apple Restricts Employee Use of ChatGPT, Joining Other Companies Wary of Leaks*, WALL ST. J. (May 18, 2023, 7:35 PM ET), <https://www.wsj.com/articles/apple-restricts-use-of-chatgpt-joining-other-companies-wary-of-leaks-d44d7d34> [<https://perma.cc/M7J7-MKFH>].

³⁰⁸ Mehta, *supra* note 111.

³⁰⁹ Tilley & Kruppa, *supra* note 307.

³¹⁰ See generally *Understanding Browser Tracking*, GCFGLOBAL, <https://edu.gcfglobal.org/en/internetsafety/understanding-browser-tracking/1/> [<https://perma.cc/VMY6-AD6Z>].

³¹¹ See *UAB “Planner5D” v. Facebook, Inc.*, No. 19-CV-03132-WHO, 2020 WL 4260733, at *8 (alleging by the plaintiff that Princeton used data scraping techniques to acquire the data protected as trade secrets. Plaintiff’s allegations that it took steps to hide “the locations and contents of its data files,” defendant “had to design and deploy hacking software to obtain this information,” and the Terms of Service prohibited access, the *prima facie* case for trade secret theft was established.).

³¹² See Family Educational Rights and Privacy Act (FERPA), 20 U.S.C. § 1232g, 34 C.F.R. pt. 99 (2023).

Italian data protection authority from further processing of personal data of Italian users.”³¹³ The Italian authority had to step in “after the authority found that the service put children and vulnerable people at risk, did not comply with requirements to provide users with certain information, and lacked a valid legal basis for its processing of personal data.”³¹⁴

Even if there is a general statement of confidentiality in the contract between the AI service provider and the user, that confidentiality provision may fall short of the requirements for specified records of students,³¹⁵ health care patients,³¹⁶ minors,³¹⁷ data subjects in the EU,³¹⁸ or customers in states such as California.³¹⁹ Each of the various regulations differs in form and content,³²⁰ so that a simple provision in a clickwrap terms of service agreement is unlikely to satisfy statutory requirements or regulators. Instead, enterprises need clear policies on what may and may not be shared with natural language systems, and these policies need to be enforced with trainings and testing.

Given the potential fines under privacy regulations such as the European General Data Protection Regulation and California Privacy Rights Act, the obligations to provide data minimization strategies, and other obligations to provide data subjects copies of information relating to the use of personal information, companies should be very wary of mixing protected personal information into broader arrays of training data.³²¹

³¹³ Giles Pratt et al., *Generative AI: Five Things for Lawyers to Consider*, FRESHFIELDS BRUCKHAUS DERINGER LLP (Feb. 21, 2023), <https://technologyquotient.freshfields.com/post/102i82i/generative-ai-five-things-for-lawyers-to-consider> [<https://perma.cc/Q2J7-6NSJ>].

³¹⁴ *Id.*

³¹⁵ *See id.*

³¹⁶ *See* Health Insurance Portability and Accountability Act of 1996 (HIPAA), 42 U.S.C. § 1320d-9 (2023).

³¹⁷ Children’s Online Privacy Protection Act of 1998 (COPPA), 15 U.S.C. 6501–6505, 16 C.F.R. pt. 312 (2023).

³¹⁸ *See* 2016 O.J. (L 119) 1.

³¹⁹ *See* California Privacy Rights Act of 2020 (CPRA), Cal. Civ. Code §§ 1798.100–1798.199.100 (West 2023).

³²⁰ *See generally* JON M. GARON, A SHORT AND HAPPY GUIDE TO PRIVACY AND CYBERSECURITY (2020).

³²¹ *See* Pratt et al., *supra* note 313.

D. Trademarks

Trademark protection is also an essential part of any business enterprise, and the early stages of generative AI services suggest that small but important steps should be taken by trademark holders as well as by the database services regarding the exploitation of trademarks.

As an early experiment with DALL-E 2, “Heinz had DALL-E 2 generate a series of images of ketchup bottles using natural language terms like ‘ketchup,’ ‘ketchup art,’ ‘fuzzy ketchup,’ ‘ketchup in space’ and ‘ketchup renaissance.’³²² The company invited fans to send their own prompts, which Heinz curated and shared across its social channels.”³²³

In most of the “ketchup” designs, DALL-E 2 utilized the Heinz trade dress of its bottle and iconic label shape.³²⁴ This is a victory for the branding power of Heinz ketchup (and perhaps a nod to the power of alternative spellings that ignore the traditional catsup), but if those images are used for purposes other than promoting Heinz products, then the company may have new challenges policing its trademark.

Text-to-image tools like DALL-E 2 are “trained” by crawling the Web and “learning” hundreds of millions of text and image associations. Earlier this year, a TechCrunch report noted that DALL-E 2 maker OpenAI “filtered out” pornography and duplicates from that learning process. But the tool can produce images that include logos, trademarked characters, and other intellectual property—such as, SpongeBob shopping at Best Buy, Homer Simpson in *Psycho*, ancient Rome Spider-Man, Santa shopping on Amazon, or an “angry mob” of Ronald McDonalds protesting working conditions, in the style of Caravaggio.³²⁵

Cartoon characters pose a particular challenge. Many of the training models have intentionally avoided presentations of real individuals, but few have policed their training libraries for the ever-popular characters from Marvel, DC, Star Wars, and the rest of the

³²² Wiggers, *supra* note 145.

³²³ *Id.*

³²⁴ *Id.*

³²⁵ Rob Walker, *What Will AI Do to Branding?*, FAST CO. (Oct. 14, 2022), <https://www.fastcompany.com/90794674/ai-dalle2-branding-copyright-infringement> [<https://perma.cc/35YW-EZUC>].

Disney or Looney Tunes libraries.³²⁶ These images are appearing in many of the early experiments.³²⁷

For famous marks, which are the marks most likely to be broadly incorporated into the training data, there is the potential risk that the appearance of the mark in association with mash-ups posted by third parties or generally incorporated into unrelated images of products and environments could lead to the dilution of those trademarks.³²⁸

For example, in the context of the litigation between Getty and Stability AI, Getty has alleged that the copying and distortion of its trademarks and watermarks by Stable Diffusion violates its trademark rights.³²⁹ Getty alleges that much of the Stable Diffusion content is of “lower quality and at times ranges from the bizarre to the grotesque. Stability AI’s incorporation of Getty Images’ marks into low quality, unappealing, or offensive images dilutes those marks in further violation of federal and state trademark laws.”³³⁰

A more common problem may arise if free images and natural language searches result in the adoption of marks that are not properly and traditionally cleared for their potential likelihood of confusion. Among small businesses, for example, the promise of royalty free images from generative AI might suggest that those images are cleared from any claims of trademark rights. But this more mundane form of trademark infringement could readily result from the overexuberance to which generative AI output is being put.

If a business were to ask a natural language AI for suggested product names, then unless the AI had been specifically designed to exclude from its answers all the product names already in use for that particular field of use and related fields of use, the answers would likely tend to rely heavily on the marks in use.³³¹ Simply put, the same neural

³²⁶ Mikael Thalen, *Artists Fed Up with AI-Image Generators Use Mickey Mouse to Goad Copyright Lawsuits*, DAILY DOT (Dec. 19, 2022, 6:28 AM), <https://www.dailydot.com/debug/ai-art-protest-disney-characters-mickey-mouse/> [<https://perma.cc/M9JS-88KS>].

³²⁷ *Id.*

³²⁸ See Lanham Trademark Act of 1946 § 43(c), 15 U.S.C. § 1125 (“The owner of a famous mark shall be entitled, . . . to an injunction against another person’s commercial use in commerce of a mark or trade name, if such use begins after the mark has become famous and causes dilution of the distinctive quality of the mark, and to obtain such other relief as is provided in this subsection.”).

³²⁹ Complaint ¶ 11, *Getty Images (US), Inc. v. Stability AI, Inc.*, 1:23-cv-00135 (D. Del. Feb. 2, 2023).

³³⁰ *Id.*

³³³ Nancy Friedman, *Can You Use ChatGPT to Create Brand Names?*, MEDIUM (Jan. 24, 2023), <https://wordworking.medium.com/can-you-use-chatgpt-to-create-brand-names-976855ea1622> [<https://perma.cc/HPF2-T7UA>].

network that makes a natural language AI know what to say next is very likely to provide to a user the most likely infringing marks rather than marks that are not in use. Similarly, as the Heinz example highlighted, certain image generation requests will invariably emphasize the existing stylized marks and trade dress that are common in popular culture and therefore likely to result in marks that will cause a likelihood of confusion if adopted.³³²

A related concern may exist for trademark holders who are the target of deliberate competitive campaigns to blur, tarnish, or free-ride on those marks. In the context of trademarks, it is reasonable to expect that the AI relies heavily on the public Internet for brand information. In such cases, a sophisticated attack could use bots and mass disinformation campaigns to cybersquat on the training data in order to boost a competitors product using the target's marks.³³³ In existing click fraud, 36 percent of display ads and 11 percent of search ads are fraudulently identified as viewed by target audience members when in fact they are "clicked" by automated bots.³³⁴ In the likely variation on this fraud, unauthorized third parties will hijack the brands to associate with the competitors goods and promote strategies to encourage the AI to make inaccurate recommendations. Trademark holders will be required to invest heavily in countermeasures to take down or disrupt the campaigns that will themselves be made simple to mount through the use of generative content.

Beyond the risks of relying on generative AI for improperly cleared marks and generative trademark dilution, there may also be a more existential threat to trademarks and brands more generally. Analysts are beginning to ask whether AI will replace marks in the public consciousness as the source identifier for goods and services.³³⁵

³³² See, e.g., *In re Majestic Distilling Co.*, 315 F.3d 1311, 1316 (Fed. Cir. 2003) ("the . . . mistaken belief that [a good] is manufactured or sponsored by the same entity [as another good] . . . is precisely the mistake that §2(d) of the Lanham Act seeks to prevent.").

³³³ See, e.g., Ben Elgin, Michael Riley, David Kocieniewski & Joshua Brustein, *How Much of Our Audience is Fake?*, BLOOMBERG (Sept. 24, 2015), <https://www.bloomberg.com/features/2015-click-fraud/?leadSource=uverify%20wall> [<https://perma.cc/UQW4-3NEG>].

³³⁴ PPC Protect, *The Global PPC Click Fraud Report 2020–21*, SEARCH ENGINE J. (Jan. 19, 2021), <https://www.searchenginejournal.com/the-global-ppc-click-fraud-report-2020-21/391493/> [<https://perma.cc/PJ95-SBVY>] ("Only 13% of accounts monitored had little to no fraudulent or invalid activity.").

³³⁵ See generally Janna Anderson & Lee Rainie, *As AI Spreads, Experts Predict the Best and Worst Changes in Digital Life by 2035*, PEW RSCH. CTR. (June 21, 2023), <https://www.pewresearch.org/internet/2023/06/21/as-ai-spreads-experts-predict-the-best-and-worst-changes-in-digital-life-by-2035/> [<https://perma.cc/YD32-ERP2>].

In a report written to an Emerging Issues Committee at the International Trademark Association,³³⁶ its authors suggest “AI may assume responsibility for setting brand or default purchase preferences, which in turn makes it increasingly difficult for companies to target customers at the critical point of sale.”³³⁷ The report also notes “the risk that AI may eliminate brand preferences altogether, for example, by allowing customers to set default preferences entirely unrelated to brand.”³³⁸ At the same time, the report recognizes that generative AI tools have the potential to help police the marketplace for unauthorized uses of marks and previously unknown examples of new marks that could engender a likelihood of confusion.³³⁹

More generally, however, there are increasing concerns among brand owners and their legal counsel that AI-generated product recommendations in online platforms and predictive tools that anticipate consumer needs are disintermediating the role of trademarks and brand management in the marketing relationship.³⁴⁰ Trademarks “allow cognitively limited humans to simplify decisions by ignoring the context in which trademarks appear.”³⁴¹ AI does not necessarily need this simplification, or it may use a very different set of parameters by which to simplify its contextualization of the environment. As a result, “AI’s advantage lies in its ability to sort through context, [reducing] the importance of maintaining stable trademark meanings”³⁴² As a result, “trademarks would be left with less to do.”³⁴³

While this remains a long-term or theoretical concern, at least as demonstrated thus far, there is little evidence to suggest trademarks and brands are diminishing in value. To the contrary, the core attribute of generative AI is an increase in information. At present, the total amount of human knowledge doubles every twelve months and in the near

³³⁶ See RENEE KEEN ET AL., ARTIFICIAL INTELLIGENCE (AI) AND THE FUTURE OF BRANDS: HOW WILL AI IMPACT PRODUCT SELECTION AND THE ROLE OF TRADEMARKS FOR CONSUMERS? (2019), <https://www.inta.org/wp-content/uploads/public-files/advocacy/committee-reports/AI-and-the-Future-of-Brands-Report-2019-010-18.pdf>.

³³⁷ *Id.* at 2.

³³⁸ *Id.* at 2.

³³⁹ *See id.* at 3, 8.

³⁴⁰ Michael Grynberg, *AI and the “Death of Trademark”*, 108 KY. L.J. 199, 200 (2019–2020).

³⁴¹ *Id.* at 205.

³⁴² *Id.*

³⁴³ *Id.*

future is expected every twelve hours.³⁴⁴ Strategies to simplify the information storm through tools such as brands and trademarks will be more important than ever. At the same time, however, brands will be unable to compete directly with AI for the delivery of factual information. Instead, the evolution of branding will likely further accelerate its drive to provide emotionally and psychographically relevant information to consumers to overcome objective decision-making and to promote brands based on emotional responsiveness.³⁴⁵

Psychographic targeting uses psychological variables and personality traits to predict and shape consumer behavior³⁴⁶ (as well as electoral voting support).³⁴⁷ To develop the psychographic determinants for individuals, researchers and marketers often use self-identification surveys, but these prove unreliable.³⁴⁸ So instead, researchers and marketers tap the power of generative AI to use natural language markers from the conversations among individuals and increasingly

³⁴⁴ Duli Pllana, *Expanding Entire Volume of Knowledge Influences on Incrementing Individual Knowledge*, 19 GLOBAL J. HUM.-SOCIAL SCI. (H) 33, 33 (2019). *See also* Scott Sorokin, *Thriving in a World of “Knowledge Half-Life,”* CIO (Apr. 5, 2019), <https://www.cio.com/article/219940/thriving-in-a-world-of-knowledge-half-life.html> [<https://perma.cc/PL6U-WYTD>].

In 1982, futurist and inventor R. Buckminster Fuller estimated that up until 1900, human knowledge doubled approximately every century, but by 1945 it was doubling every 25 years. And by 1982, it was doubling every 12-13 months. In retrospect, this may sound a little quaint since experts now estimate that by 2020, human knowledge will double every 12 hours.

³⁴⁵ *See* Bhumika Dutta, *A Guide to Psychographics for Marketing*, ANALYTIC STEPS (Nov. 29, 2021), <https://www.analyticssteps.com/blogs/guide-psychographics-marketing> [<https://perma.cc/8EMU-8Y8Y>] (“People understood that merely demographics are insufficient for any effective marketing techniques after the development of psychographic segmentation. . . . This field extends beyond demographics such as age, gender, race, geography, and so on to analyze customer behavior over purchasing.”). *See generally* *What is Psychographics?*, CB INSIGHTS (May 6, 2020), <https://www.cbinsights.com/research/what-is-psychographics/> [<https://perma.cc/6Y3K-X33C>] (“Psychographics seeks to understand the cognitive factors that drive consumer behaviors. This includes emotional responses and motivations; moral, ethical, and political values; and inherent attitudes, biases, and prejudices.”).

³⁴⁶ Hui Liu et al., *Personality or Value: A Comparative Study of Psychographic Segmentation Based on an Online Review Enhanced Recommender System*, 9 APPLIED SCI. 2, 2 (2019).

³⁴⁷ *See* *What is Psychographics?*, *supra* note 345. A discussion of the use of generative AI and psychographic modeling is beyond the scope of this article.

³⁴⁸ Philip S. Brenner & John DeLamater, *Lies, Damned Lies, and Survey Self-Reports? Identity as a Cause of Measurement Bias*, 79 SOC. PSYCH. Q. 333, 333 (2016).

with the chatbots and generative AI services.³⁴⁹ Psychographic research “demonstrates the utility of relying on natural language markers of abstract psychological phenomena, including values and personality, and present significant opportunities to better predict and understand their connection to consumers’ behaviors and thoughts in a broader sense.”³⁵⁰ In this brave new world of psychographic targeting, trademarks will become more prominent than ever, but will likely represent a very different brand strategy than those of the twentieth century.

A final trademark consideration is the potential to see a significant change in the role for Lanham Act § 43(a) false advertising claims.³⁵¹ The federal cause of action provides:

(1) Any person who, . . . uses in commerce any word, term, name, symbol, or device, or any combination thereof, or any false designation of origin, false or misleading description of fact, or false or misleading representation of fact, which—

(A) is likely to cause confusion, or to cause mistake, or to deceive as to the affiliation, connection, or association of such person with another person, or as to the origin, sponsorship, or approval of his or her goods, services, or commercial activities by another person, or

(B) in commercial advertising or promotion, misrepresents the nature, characteristics, qualities, or geographic origin of his or her or another person’s goods, services, or commercial activities,

shall be liable in a civil action by any person who believes that he or she is or is likely to be damaged by such act.³⁵²

As noted earlier, the first iterations of natural language chats have proven to be highly inconsistent in their accuracy.³⁵³ For open environment natural language systems that rely on the information generally available on the Internet, misinformation will continue to

³⁴⁹ See Liu et al., *supra* note 346, at 6 (“Research indicates cases in which natural language data have provided a clearer picture of people’s cognitive and behavioral processes than data collected from a traditional and widely used self-report survey.”).

³⁵⁰ *Id.* at 6.

³⁵¹ See Grynberg, *supra* note 340, at 212–13.

³⁵² 15 U.S.C. § 1125.

³⁵³ See *supra* Section IV.2.

inform the responses from generative AI chatbots. In one potential situation, a response system has “learned” that the most popular automobile on the market today is the Pacer by American Motors Corporation.³⁵⁴ Since the AMC Pacer was thankfully put to rest in 1980, such information will not affect any prospective car buyers. But if instead, the AI had been coached to promote a particular brand in the marketplace using factually inaccurate information, then such information could meet the standard for false marketing under section 43(a).³⁵⁵

Like many of the challenges to generative AI, the question remains who is responsible for this misinformation. At a minimum, complainants will need to establish that they have a “commercial interest in reputation or sales” harmed by the chatbot.³⁵⁶

Certainly, if a competitor could be proven to have falsely trained the AI to promote its brand with factually inaccurate information, that would give rise to liability. But in the absence of such proof, could an AI host be obligated to “remove” such information? Indeed, what does the removal look like, particularly if the site providing access to the chatbot is merely a licensee from the source of the training data?

As between the host website and the provider of the natural language service, Communications Decency Act § 230 might provide an answer.³⁵⁷ Section 230(c) of the CDA provides: “No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”³⁵⁸ Unlike copyright and patent laws, CDA § 230 does not use the term individual,³⁵⁹ and because it is written in the negative, the law potentially sweeps much more broadly to include content that is created by generative AI.

³⁵⁴ See *AMC Pacer*, WIKIPEDIA, https://en.wikipedia.org/wiki/AMC_Pacer [<https://perma.cc/3C43-U96Z>] (“The AMC Pacer is a two-door compact car produced in the United States by American Motors Corporation (AMC) from the 1975 through the 1980 model years.”).

³⁵⁵ See Quentin J. Ullrich, *Is This Video Real? The Principal Mischief of Deepfakes and How the Lanham Act Can Address It*, 55 COLUM. J.L. & SOC. PROBS. 1, 16 (2021).

³⁵⁶ *Lexmark Int’l, Inc. v. Static Control Components, Inc.*, 572 U.S. 118, 131–34 (2014).

³⁵⁷ See 47 U.S.C. § 230(c).

³⁵⁸ 47 U.S.C. § 230(c).

³⁵⁹ *Id.* at § 230(f)(3) (“Information content provider” means “any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the internet or any other interactive computer service.”).

At the same time, however, it cannot provide immunity for the generative AI itself.³⁶⁰ When Simplicity AI or OpenAI create and train their models, they are not using an interactive service, putting them outside the framework of the safe harbor.³⁶¹ In addition, Professor Matt Perault writes that given the definition of interactive service provider, “courts will likely find that ChatGPT and other [systems] are information content providers.”³⁶²

In contrast, the service creating and generating the generative AI is likely an interactive computer service. The content is coming from third parties. Further, the interactive computer service is not substantially modifying or controlling the content. In this situation, the hosting service will likely be able to shield itself from the direct liability for the content provided by the natural language generator. Of course, if the website engages more directly with the content in some manner, then it may lose the protection of the safe harbor as a result.

Separately, to the extent the claim is based on violations of the Lanham Act, CDA § 230 should not serve as a bar to the cause of action. Subsection (e) ensures several legal domains remain unaffected by § 230(c). Most relevant here, the court in *Hepp v. Facebook*,³⁶³ made clear § 230 has “[n]o effect on intellectual property.”³⁶⁴ In *Universal Communication Systems, Inc. v. Lycos, Inc.*,³⁶⁵ the plaintiff alleged that the defendant message board providers allowed posts contained “false, misleading” content about the company’s financial status and position.³⁶⁶ The plaintiff brought claims for both state and federal trademark violations and the court refused to dismiss the claims based on § 230.³⁶⁷ Contrast this with a narrower interpretation in *Perfect 10, Inc. v. CCBill LLC*,³⁶⁸ where the court found that § 230 barred the state intellectual property claims and construed the limitation of § 230(e)(2) to federal claims.³⁶⁹ While such a limitation may have made some public policy sense in the context of an earlier understanding of

³⁶⁰ Matt Perault, *Section 230 Won’t Protect ChatGPT*, LAWFARE BLOG (Feb. 22, 2023, 1:11 PM), <https://www.lawfareblog.com/section-230-wont-protect-chatgpt> [<https://perma.cc/2279-5JL2>].

³⁶¹ *Id.*

³⁶² *Id.*

³⁶³ *Hepp v. Facebook*, 14 F.4th 204, 209 (3d Cir. 2021).

³⁶⁴ *Id.* at 209 (quoting 47 U.S.C. § 230(e)(2)); *see also* *Universal Commc’n Sys., Inc. v. Lycos, Inc.*, 478 F.3d 413, 415–416 (1st Cir. 2007).

³⁶⁵ *Universal Commc’n Sys.*, 478 F.3d at 415–16.

³⁶⁶ *See id.*

³⁶⁷ *Id.* at 422–25.

³⁶⁸ *Perfect 10, Inc. v. CCBill LLC*, 488 F.3d 1102, 1118–19 (9th Cir. 2007).

³⁶⁹ *Id.* at 1118–20.

potential Internet liability, the Third Circuit recognition of the state right of publicity likely reflects the current concerns over Internet immunity and the loss of confidence that courts have found the right balance between § 230 and other laws.³⁷⁰

E. Rights of Publicity in the Data Set and the Output

As noted in *Hepp v. Facebook*, state rights of publicity claims are also intellectual property claims.³⁷¹ The term “right of publicity” was first used by Judge Jerome Frank in 1953 in *Haelan Laboratories, Inc. v. Topps Chewing Gum, Inc.*³⁷² As the field has evolved in the context of online influencer culture and collegiate athlete brand endorsements, today, it is very common to refer to publicity rights as “name, image, and likeness (NIL) licensing.”³⁷³

The right of publicity is an intellectual property right of recent origin which has been defined as the inherent right of every human being to control the commercial use of his or her identity. The right of publicity is a creature of state law and its violation gives rise to a cause of action for the commercial tort of unfair competition.³⁷⁴

³⁷⁰ *Hepp v. Facebook*, 14 F.4th 204, 209 (3d Cir. 2021).

³⁷¹ *Haelan Labs., Inc. v. Topps Chewing Gum, Inc.*, 202 F.2d 866, 868 (2d Cir. 1953).

³⁷² *Id.*; see Melville Nimmer, *The Right of Publicity*, 19 L. & CONTEMP. PROBS. 203, 204 (1954).

³⁷³ See, e.g., Ted Tatos, *The NCAA Goes After College Athletes’ NIL Money—Here are the Antitrust Implications for Workers and Consumers*, PROMARKET (May 20, 2022), <https://www.promarket.org/2022/05/20/ncaa-goes-after-college-athletes-nil-antitrust/> [<https://perma.cc/5M5D-PZJU>] (“[T]he NCAA has signaled its intent to renew its efforts to suppress competition – this time in the nascent, but rapidly growing market for college athlete Name, Image, and Likeness (NIL) licensing.”); Nathan Rubbelke, *How the NCAA’s Name, Image and Likeness Licensing Policies are Creating a New Type of Company*, ST. LOUIS BUS. J. (June 9, 2022), <https://www.bizjournals.com/stlouis/news/2022/06/09/ncaa-nil-policies-creating-new-type-of-company.html> [<https://perma.cc/488K-NZ4B>] (“So-called “NIL collectives” are a new phenomenon, and type of business entity, that pool together money from a college’s boosters and supporters to fund name, image and likeness deals with student-athletes.”).

³⁷⁴ *ETW Corp. v. Jireh Publ’g., Inc.*, 332 F.3d 915, 928–29 (6th Cir. 2003).

For purposes of taxonomy, the right of publicity shares certain attributes with privacy rights from which it was originally derived,³⁷⁵ with trademark law³⁷⁶ because it requires a commercial exploitation of the name, image, or likeness, and even with copyright because of its communicative impact.³⁷⁷

Each of these attributes helps describe the potential for liability that can arise from the use of a person's recognizable name, image, or likeness in the output of a generative AI response, image, or video. In addition, any potential litigation should take into account the ability of an individual to control whether a person can object to having their information collected into the database³⁷⁸ as distinct from whether the output of a person's name, image, or likeness requires permission.

Thus far, claims regarding the inclusion of a person's identity have not proven actionable under state publicity rights law. For example, in *Verde v. Confi-Chek, Inc.*,³⁷⁹ a plaintiff sued the operators of the website peoplefinders.com, claiming that the Illinois Right to Publicity Act³⁸⁰ was violated as a consequence of being included in the peoplesfinder.com database.³⁸¹ The district court refused to find Article

³⁷⁵ See Samuel Warren & Louis Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 193 (1890).

³⁷⁶ See Stacey L. Dogan & Mark A. Lemley, *What the Right of Publicity Can Learn from Trademark Law*, 58 STAN. L. REV. 1161, 1178 (2006). *But see* Daniel Gervais & Martin L. Holmes, *Fame, Property & Identity: The Purpose and Scope of the Right of Publicity*, 25 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 181, 199 (2014) (“Rationales that are used to justify trademark law should not be offered to justify the right of publicity simply because the two rights seem similar. . . . [T]rademark law reduces search costs for consumers and incentivizes quality through accountability; the right of publicity, meanwhile, allows an individual to profit from and exercise control over the commercial use of his name and likeness” But, of course, this analysis ignores all anti-dilution and false endorsement provisions of the Lanham Act).

³⁷⁷ See *Comedy III Prods., Inc. v. Gary Saderup, Inc.*, 21 P.3d 797, 808 (Cal. 2001) (incorporating the transformative use test from copyright fair use into California's right of publicity laws).

³⁷⁸ See Madeline Lamo & Jeff Landis, *Spike in Right of Publicity Cases Against Online Providers*, ZWILLGENBLOG (Dec. 2, 2021), <https://www.zwillgen.com/privacy/spike-right-publicity-cases-online-providers/> [<https://perma.cc/TLM5-EJ26>].

³⁷⁹ *Verde v. Confi-Chek, Inc.*, No. 21 C 50092, 2021 WL 4264674, at *1 (N.D. Ill. Sept. 20, 2021).

³⁸⁰ 765 ILL. COMP. STAT. 1075/30(a), 40(a) (2023) (providing that “[a] person may not use an individual's identity for commercial purposes during the individual's lifetime without having obtained previous written consent. . . .” A person who violates section 30 may be liable for the greater of “(1) actual damages, profits derived from the unauthorized use, or both; or (2) \$1,000.”).

³⁸¹ *Verde*, 2021 WL 4264674, at *1–2.

III standing because the inclusion in the database did not, in and of itself, constitute a harm under the statute.³⁸² However, in more recent motions to dismiss, other Illinois district courts are finding that the use of a person's information in the database used to help sell subscriptions is sufficient to provide a cause of action.³⁸³

In *Hoffower v. Seamless Contacts, Inc.*, for example, the court found that the plaintiff “has sufficiently alleged unconsented commercial use of her identity.”³⁸⁴ Specifically, she alleges that her identity (like those of other class members) is used to advertise Seamless's subscription service.³⁸⁵ “To the extent the product or service advertised has to be separate from the plaintiff's identity itself, that's satisfied here.”³⁸⁶

In much the same way, many of the free generative AI systems provide a limited number of free searches and then the consumer is offered or encouraged to purchase increased access. Under this free-to-try model, the *Hoffower* explanation of commercial use would mean that the information is now being used to promote the paid service.

In the more typical case, the commercial activity requires the association of the individual's identity with some product or service, as opposed to a news story or other communicative work. To the extent that advertising agencies and commercial websites take advantage of the generative AI content to produce materials that could be construed as marketing materials, those uses will require licenses from any persons who are identifiable. For example, in *Hepp*, the complaint was that without her consent, a photograph of the plaintiff standing “in a convenience store, smiling in the center of the frame's foreground . . . [was] used in online advertisements for erectile dysfunction and dating websites.”³⁸⁷ If that image commonly found on the Internet was closely reproduced in an AI generation, it would still have the same basis for liability. Moreover, as noted earlier § 230 may no longer bar claims

³⁸² *Id.* at *5 (quoting *TransUnion LLC v. Ramirez*, 141 S. Ct. 2190, 2209 (2021) (“[P]laintiff fails to allege a concrete injury because she does not allege defendant disclosed any of her information to any third party. ‘[T]he retention of information lawfully obtained, without further disclosure, traditionally has not provided the basis for a lawsuit in American courts.’”).

³⁸³ *Hoffower v. Seamless Contacts, Inc.*, No. 22 C 2079, 2022 WL 17128949, at *2 (N.D. Ill. Nov. 22, 2022); *Siegel v. ZoomInfo Techs., LLC*, No. 21 C 2032, 2021WL 4306148, at *4 (N.D. Ill. Sept. 22, 2021).

³⁸⁴ *Hoffower*, 2022 WL 17128949, at *2.

³⁸⁵ *Id.* at *1.

³⁸⁶ *Id.* at *2.

³⁸⁷ *Hepp v. Facebook*, 14 F.4th 204, 206 (3d Cir. 2021).

under state rights of publicity statutes³⁸⁸ and it would not bar federal causes of action for false endorsement under § 43(a) of the Lanham Act.³⁸⁹

Identity rights are not limited to photographic representations of an individual. Likenesses are generally included in the panoply of rights granted to each individual, and likenesses should reasonably be understood to include avatars.³⁹⁰ Thus far, DALL-E 2 and Stable Diffusion have addressed the rights of publicity and privacy concerns by obscuring and distorting the images of human characters. But this strategy is not universal, structural, nor comprehensive. The choice for any particular generative AI to provide facial information can be made by the operator and changed as market conditions dictate. Generated Photos,³⁹¹ Fotor,³⁹² and other services offer very lifelike photos generated completely by AI.³⁹³ If those images are trained on a database of preexisting photographs, then a person bearing a striking likeness to one of the images might have a valid claim that the person's image is being commercially exploited without permission.

In addition, since avatars are widely generated, there may be nothing that restricts the generative AI from learning the most popular of avatars and basing its output on these. To the extent that these derivative avatars rely too heavily on the source material and they are used for commercial exploitation, then they will invade the rights of publicity for the owners of those avatars and create legal liability as a consequence.

A recently filed lawsuit in Illinois will explore the extent to which the development of avatars based on an individual's likeness requires compliance with state laws protecting biometric information such as the Illinois Biometric Information Privacy Act (BIPA).³⁹⁴ In *Flora v. Prisma Labs*,³⁹⁵ the plaintiffs allege that Prisma's "magic

³⁸⁸ *Id.* at 226.

³⁸⁹ 15 U.S.C. § 1125(a).

³⁹⁰ See Dhruva Krishna, *Dawn of the Dead: Virtual Avatars & New York's Right of Publicity*, HARV. J. SPORTS & ENT. L. (Sept. 29, 2021), <https://journals.law.harvard.edu/jsel/2021/09/dawn-of-the-dead-virtual-avatars-new-yorks-right-of-publicity/> [<https://perma.cc/9JC4-BEGJ>].

³⁹¹ See generally Generated Photos, *supra* note 136.

³⁹² See generally FOTOR, <https://www.fotor.com/features/ai-face-generator/> [<https://perma.cc/U6EA-A67Z>].

³⁹³ *Id.*

³⁹⁴ See generally Biometric Information Privacy Act ("BIPA"), 740 ILL. COMP. STAT. 14/1 (2008).

³⁹⁵ See Complaint for Damages and Equitable Relief at 22 ¶¶ 85–87, *Flora v. Prisma Labs, Inc.*, No. 5:23-cv-00680 (N.D. Cal. Feb. 15, 2023).

avatar” feature enables the company to collect facial geometry of any person’s face, which is used to train the neural network that improves the avatar generation tool. The app does not restrict the use of magic avatar function to one’s own image, so that individuals who did not consent to the terms of service may still be included in the dataset.³⁹⁶ The statute makes clear that a “scan of . . . face geometry” constitutes a “biometric identifier.”³⁹⁷ Biometric identifiers, however, only become “biometric information” when it is used to identify an individual. Still, the statute provides that no one may collect either biometric identifiers or biometric information without informed written consent.³⁹⁸

The lawsuit challenges Prisma’s notice language regarding the personal information collected, and of course, that language is irrelevant to any members of the class who had their biometric identifiers uploaded by friends and family without their own consent. At the same time, the case may also look more closely into the State’s ability to define biometric information in so broad a manner as to interfere with speaker’s use of images and other biometric identifiers more generally.

BIPA could also prove significant in the field of generative AI music and voice duplication. A voiceprint is also a biometric identifier. BIPA provides a state statutory basis for objection to the use of one’s voice without prior consent. BIPA actions will bolster more claims that are often more difficult to establish regarding unfair competition or publicity rights claims for sound-alike infringements.³⁹⁹

F. Where Congress May Wish to Tread

There are a few additional areas where intellectual property law likely needs congressional intervention to improve the fit between intellectual property protections and innovation policy. Rather than

³⁹⁶ *See id.* at 22 ¶¶ 89–91.

³⁹⁷ BIPA, 740 ILL. COMP. STAT. 14/10 (2008).

³⁹⁸ *Id.* at 14/15(b).

³⁹⁹ *See, e.g.,* Midler v. Ford Motor Co., 849 F.2d 460, 463–64 (9th Cir. 1988) (applying common law right of publicity claim to appropriation of singer’s voice by voice-impersonator while statutory claim did not); *Waits v. Frito-Lay, Inc.*, 978 F.2d 1093, 1106–07 (9th Cir. 1992) (noting impersonation of a performer’s voice may give rise to a claim for false endorsement). *But see* *Sinatra v. Goodyear Tire & Rubber Co.*, 435 F.2d 711, 713–14 (9th Cir. 1970) (rejecting claim of false endorsement when singer’s performance was attributed to a single song).

starting with *Goldsmith*⁴⁰⁰ or other recent Supreme Court cases,⁴⁰¹ the lessons can best be learned from the Supreme Court in *White-Smith Music Publishing Co. v. Apollo Co.* from 1908.⁴⁰² In the late nineteenth century, inventors created the phonogram, telephone, motion picture, and player piano. Commercial radio and movie theaters soon followed. The Supreme Court rejected the claim to copyright for protection of a player piano roll, holding such a novelty could not be within the meaning of the law.⁴⁰³ “I cannot convince myself that these perforated strips of paper are copies of sheet music within the meaning of the copyright law. They are not made to be addressed to the eye as sheet music, but they form part of a machine.”⁴⁰⁴

Fortunately, the Court’s decision was not the last word. In 1909, Congress enacted the most comprehensive copyright revision it has ever undertaken.⁴⁰⁵ The new law corrected the arbitrary limit placed on copyright by the Court, extending copyright to player piano rolls that could only be read through a machine or device.⁴⁰⁶ Concerned about the economic monopoly that the change in the law could foster, Congress also introduced the statutory royalty for musical compositions that continues to define and provide balance between the creators of copyrighted works, those that exploit those works, and the consumer public. In doing so, Congress recognized the constitutional mandate that the intellectual property monopoly is intended “[t]o promote the Progress of Science and Useful Arts”⁴⁰⁷

Generative AI and synthetic media have the potential to further progress. Generative AI has the same potential to reshape the creative industries as did the invention of the phonogram, telephone, motion picture, player piano, and commercial radio. Perhaps generative AI is

⁴⁰⁰ See *Andy Warhol Found. Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 532 (2023) (“[T]he first fair use factor considers whether the use of a copyrighted work has a further purpose or different character, which is a matter of degree, and the degree of difference must be balanced against the commercial nature of the use.”).

⁴⁰¹ See, e.g., *Jack Daniel’s Props., Inc. v. VIP Prods. LLC*, 599 U.S. 140, 163 (2023) (“*Rogers v. Grimaldi*, 875 F.2d 994 (1989) does not apply when the challenged use of a mark is as a mark. On dilution, . . . the noncommercial exclusion does not shield parody or other commentary when its use of a mark is similarly source-identifying.”).

⁴⁰² *White-Smith Music Pub. Co. v. Apollo Co.*, 209 U.S. 1 (1908).

⁴⁰³ *Id.* at 18.

⁴⁰⁴ *Id.* at 12 (quoting *Kennedy v. McTammany*, 33 F. 584, 584 (C.C.D. Mass. 1888)).

⁴⁰⁵ 17 U.S.C. § 1(e) (1909) (repealed 1976) (providing, in the original law, that a musical composition included “any system of notation or any form of record in which the thought of an author may be recorded and from which it may be read or reproduced”).

⁴⁰⁶ *Id.*

⁴⁰⁷ U.S. CONST. art. I, § 8, cl. 8.

only as significant as one of those industries or perhaps it will be as transformative as all of these technologies combined. Only time will tell. It also has the potential to be as disruptive as each or as all of them.

To better integrate generative AI into the law, a few additional areas need to be clarified. For example, in the area of “Digital Protection Measures,”⁴⁰⁸ under the statute, to “circumvent a technological measure” means “to descramble a scrambled work, to decrypt an encrypted work, or otherwise to avoid, bypass, remove, deactivate, or impair a technological measure, without the authority of the copyright owner.”⁴⁰⁹ The question is whether a “do not copy” flag is a sufficient digital protection measure to meet this statutory standard. The related question is whether the protections provided by § 1201 are appropriate to stop scraping of content from public websites.

Protections afforded by § 1201 would enable authors and artists greater control over the extent to which their works were incorporated into training sets. Reliance on such flags might prove to be a much better tool for both copyright owners and AI developers. Having a technical system that allows the copyright owner to opt in or out on a work-by-work basis may prove to enable the growth of responsible AI without harming those authors and artists who are most concerned about the use of their works in the training sets. In addition, it may not be sufficient to rely on only those sites that utilize flags to prohibit copying and web crawling. Vast amounts of copyrighted work are reproduced without authorization at repository sites. In the class action lawsuit filed against OpenAI and Meta, the plaintiffs alleged “that OpenAI’s ChatGPT and Meta’s LLaMA were trained on illegally-acquired datasets containing their works, which they say were acquired from ‘shadow library’ websites like Bibliotik, Library Genesis, Z-Library, and others, noting the books are ‘available in bulk via torrent systems.’”⁴¹⁰ The problem of unauthorized copying is greatly exacerbated by generative AI systems that circumvent anti-copying flags and more aggressive technologies, if the LLM developers exploit pirated works for their training set.

⁴⁰⁸ 17 U.S.C. § 1201. Section 1201(a)(1) provides in pertinent part that “[n]o person shall circumvent a technological measure that effectively controls access to a work protected under this title.”

⁴⁰⁹ See 17 U.S.C. § 1201(a)(3)(A).

⁴¹⁰ Wes Davis, *Sarah Silverman is Suing OpenAI and Meta for Copyright Infringement*, THE VERGE (July 9, 2023, 2:14 PM), <https://www.theverge.com/2023/7/9/23788741/sarah-silverman-openai-meta-chatgpt-llama-copyright-infringement-chatbots-artificial-intelligence-ai> [<https://perma.cc/6PGJ-38KD>] (citing Complaint, *Silverman v. OpenAI*, No. 3:23CV-03416 (N.D. Cal. July 7, 2023)).

Updates to § 1201 need to take into account both the need for improvement in the implementation and enforcement of the services that use the flags and also to increase consequences for companies that avoid § 1201 by copying already-pirated works.

Second, the concern over pirated works has only grown since the Digital Millennium Copyright Act crafted the secondary liability compromise reflected in § 512.⁴¹¹ As the Copyright Office has highlighted, § 512 was enacted with a dual purpose: “(1) to enable copyright owners to effectively address the infringement of their works online and (2) to facilitate the development of internet-based platforms by clarifying the obligations and limiting the liability of OSPs [online service providers] with respect to infringement committed by third-party users of their systems.”⁴¹² The notice-and-takedown regime of § 512 has become increasingly ineffective in the decades following its adoption as the scale of user-created content has grown exponentially, the technologies used to identify and monetize content have evolved, and the operational challenges to copyright enforcement have increased.⁴¹³ The more general need for § 512 reform is beyond the scope of this article, but it is important to note that a great deal of the content available to LLMs is online because of third party use rather than through the exploitation of the copyright holder. The broader need to reform § 512 should recognize the impact that the § 512 balance facilitating third party use of copyrighted works now works to divest authors and creators over their control of their works as training content. Since this content is presently free for the LLM systems to exploit, there is no need to develop a commercial market to obtain these works.

Third, as illustrated by the congressional balancing provided in both § 512 and § 1201, the traditional understanding of copyright law was primarily that of a commercial legal system. Copyright law serves as a form of national collective bargaining between creators and media and software industries which attempts to also take into account the best interest of the public. As many of the limitations on copyright highlight,⁴¹⁴ copyright law had been drafted and largely applied in the

⁴¹¹ Digital Millennium Copyright Act of 1998, Pub. L. No. 105-304, 112 Stat. 2860 (1998) (codified as amended at 17 U.S.C. §§ 512(a)–(d)).

⁴¹² U.S. COPYRIGHT OFF., SECTION 512 OF TITLE 17, A REPORT OF THE REGISTER OF COPYRIGHTS 8 (2020) (citing S. REP. NO. 105-190, at 40 (1998); H.R. REP. NO. 105-551, pt. 2, at 49–50 (1998)).

⁴¹³ *See id.* at 137.

⁴¹⁴ *See* 17 U.S.C. § 1001 (Audio Home Recording Act); 17 U.S.C. § 110(1) (limitation on copyright for classroom instruction); 17 U.S.C. § 110(3) (performances in religious settings); 17 U.S.C. § 110(7) (music stores).

context of commercial litigation rather than enforcement against individual infringers. The Home Audio Recording Act, for example, attempted to create a balance, providing revenue to the music industry for digital audio taping while exempting home audio recording from liability.⁴¹⁵

The massive proliferation of social media and user-generated content highlights the importance that copyright law should primarily be enforced against commercial interests. As § 512 attempts to provide, the balance of interests is benefitted if consumer works are subject to removal from the Internet but not to statutory damages that can quickly rise to millions of dollars. To this end, Congress should consider eliminating or modifying the availability of statutory damages against non-commercial infringements of an author's works.

With the growth of social media, peer-to-peer systems, and the Web 2.0 marketplace, copyright law has become more relevant to individuals, potentially as the subject of lawsuits involving statutory damages amounting to hundreds of thousands of dollars.⁴¹⁶ Particularly at an inflection point where the future is hard to predict, it is important that the public is not unduly restricted in their experimentation. Given the potential for an individual to create as much harm as a commercial enterprise, the appropriate balance should not include total immunity. Instead, the availability of injunctive relief, of actual damages, and perhaps the right to receive statutory damages only when certain thresholds have been met would return copyright to its commercial roots. No one wants to sue a fifth grader for the website she created as a class project. A revision of the law should make such activities fall outside the scope of statutory damages.

Fourth and finally, the great interest in generative AI should trigger another look at the extent to which the United States has an international obligation and economic imperative to address the right of attribution under copyright law and *sui generis* publicity rights under new federal law. Both attribution and identity are growing in

⁴¹⁵ Audio Home Recording Act, Pub. L. No. 102-563, 106 Stat. 4237 (1992) (codified as amended at 17 U.S.C. §§ 1001–1010).

⁴¹⁶ *Cf.* Sony BMG Music Ent. v. Tenenbaum, 719 F.3d 67, 71 (1st Cir. 2013) (“Tenenbaum argues that the award of \$675,000 violates due process because it is not tied to the actual injury that he caused, which he estimates to be no more than \$450, or the cost of 30 albums at \$15 each. But this argument asks us to disregard the deterrent effect of statutory damages, the inherent difficulty in proving damages in a copyright suit.”); *Cap. Recs., Inc. v. Thomas-Rasset*, 692 F.3d 899, 902 (8th Cir. 2012) (“[T]he recording companies are entitled to the remedies they seek: damages of \$222,000 and a broadened injunction that forbids Thomas–Rasset to make available sound recordings for distribution.”).

commercial value and the lack of federal protection leads to free-riding and highly inconsistent application of state law.

The right of attribution is fundamental to international copyright pursuant to Article 6^{bis} of the Berne Convention.⁴¹⁷ Since the U.S. joined the Berne Convention in 1988, it has been a treaty obligation to afford this protection.⁴¹⁸ The U.S. skirted this requirement in its admission to Berne by relying upon the Lanham Act's unfair competition protection for the use of an individual's name or likeness in a manner that triggers a false endorsement.⁴¹⁹ This protection, however, has been largely vitiated by the Supreme Court's interpretation of the intersection between copyright and trademark law.⁴²⁰ In *Dastar Corp. v. Twentieth Century Fox Film Corp.*, the Court rejected the approach taken to explain adherence to Berne, instead stating that copyright law preempted trademark law in a conflict over attribution, and that the congressional choice to exclude attribution in most instances limited both the Court's interpretation of copyright and of trademark.⁴²¹

Despite the Court action, attribution has become more important than ever. As leading copyright scholar Jane Ginsburg has explained, the U.S. has a strong interest in recognizing attribution rights.

Perpetuating [the] omission not only allows a source of international embarrassment to continue to fester; it also belittles our own creators. Copyright not only protects the economic interests in a work of authorship, it also secures (or should secure) the dignitary interests that for many authors precede monetary gain. Without established and enforceable attribution rights, U.S.

⁴¹⁷ Berne Convention for the Protection of Literary and Artistic Works art. 6bis, July 24, 1971, 25 U.S.T. 1341, 828 U.N.T.S. 221. See generally, Jon M. Garon, *Commercializing the Digital Canvas: Renewing Rights of Attribution for Artists, Authors, and Performers*, 1 TEX. A&M L. REV. 837, 872 (2014).

⁴¹⁸ See Jane C. Ginsburg, Keynote Address, *The Most Moral of Rights: The Right to be Recognized as the Author of One's Work*, 8 GEO. MASON J. INT'L COM. L. 44, 46 (2016) ("The United States, I believe, is the only country, including among common law countries, not to include attribution rights in its copyright law. Whether or not there is any effective international sanction for non-implementation of this particular Berne Convention obligation, it is not a good thing to be an international scofflaw.").

⁴¹⁹ See Garon, *supra* note 417, at 838 ("At the time, rights of attribution and integrity were seen as quasi-trademark rights, and artists were protected under the Lanham Act.").

⁴²⁰ See *Dastar Corp. v. Twentieth Century Fox Film Corp.*, 539 U.S. 23, 33 (2003).

⁴²¹ *Id.*

copyright neither meets international norms nor fulfills the aspirations of the constitutional Copyright Clause.⁴²²

Attribution for the author of copyrighted works, however, is no longer enough. As courts have recognized, there is also an economic and ethical need to protect individuals who are commercially exploited as the subject matter of works created without authorization.⁴²³ Image generators often use celebrity versions of copyright-protected characters, so that when a prompt requests the image generator to create a picture of “The Avengers,” it will not only reproduce the Marvel-owned cartoon characters, but may also select the faces of Robert Downey Jr. as Iron Man; Scarlett Johansson as Natasha Romanoff; Chris Evans as Captain America, Steve Rogers; Mark Ruffalo as The Hulk; Chris Hemsworth as Thor; Jeremy Renner as Hawkeye, Clint Barton; Brie Larson as Captain Marvel, Carol Danvers; and Samuel L. Jackson as Nick Fury.⁴²⁴ The popular actors are so closely associated with their iconic characters that the LLMs are trained to equate the individual with the character. Much like Bela Lugosi’s Count Dracula and Boris Karloff’s Frankenstein’s Monster, the LLMs can be trained to reproduce particular images in response to the text prompts in a highly consistent manner.⁴²⁵

⁴²² Ginsburg, *supra* note 418, at 81.

⁴²³ See, e.g., *Hart v. Elec. Arts, Inc.*, 717 F.3d 141, 170 (3d Cir. 2013) (holding that student athletes had a right of publicity in avatars when depicted in video games and that such depictions were not sufficiently transformative so as to overcome such rights); *In re NCAA Student-Athlete Name & Likeness Licensing Litig.* (“Keller”), 724 F.3d 1268, 1284 (9th Cir. 2013); see also *Nat’l Collegiate Athletic Ass’n v. Alston*, 141 S. Ct. 2141, 2141 (June 21, 2021) (finding NCAA limitations on income violated antitrust regulations); *O’ Bannon v. Nat’l Collegiate Athletic Ass’n*, 802 F.3d 1049, 1068 (9th Cir. 2015).

⁴²⁴ See *Avengers: Endgame*, IMDB (2019) <https://m.imdb.com/title/tt4154796/fullcredits/cast> [<https://perma.cc/P4WF-63SL>]; Snigdha Nalini Orey, *Ever Wondered How Avengers Would Look if It Was Based in India? These AI Images Are the Answer*, SCOOP WHOOP (June 1, 2023), <https://www.scoopwhoop.com/in-other-news/ai-generated-images-avengers-would-look-if-it-was-based-in-india/> [<https://perma.cc/DC4L-BKSX>].

⁴²⁵ Although the LLM does not store particular images or text in its weighting algorithm, the heavy training of certainly well-known works may mean that the LLM is, in fact, a derivative work of those training sources. For example, if the response to the text prompt for “Dorothy” consistently returns an image of Judy Garland’s picture in pigtails taken as the original publicity shots for *The Wizard of Oz*, then MGM’s photo might be “in” the LLM weighting scheme, not as a visual reproduction but as a functional derivative work of the photo. The actor would also be captured in the LLM but would have no copyright claim.

At present, the legal status of such uses will depend on the vagaries of state laws and state courts interpreting the common law by analogy. When their publicity rights are exploited without their permission, they have legal recourse, but it varies from state to state, adding to complexity and confusion. Standardizing this protection by creating a federal publicity rights statute would greatly benefit all stakeholders. The process of drafting the legislation would help ensure that commercial stakeholders and representatives of personal interests all have the opportunity to provide input into this process.

VI. CONCLUSION

Generative AI has captured the imagination of the public, motivated significant investments of capital, moved securities markets, and lit a fire within industry. As the irrational exuberance of the pre-release ChatGPT bubble deflates, it will likely be replaced with a much more serious and systematic revision to the global economy and every sector that relies on text, data, imagery, or design for its economic wellbeing.

The promises seem magical: The power to automate routine tasks; the ability to create AI-driven but authentic-feeling interactions in marketing and customer service; the power to obtain free art, music, and narrative that can be incorporated into commercial displays; the ability to improve software while reducing its price and speeding its production; the ability to explore new designs almost instantly; and the potential to invent new chemicals, molecules, and medicinal cures. While the list is not infinite, the potential may be limitless.

For companies hoping to jump on the bandwagon, however, there is a significant gulf between the promise and the practice. Enterprises that fail to adhere to common sense practices of fact checking, testing new systems before deploying them, and respecting the intellectual property of the underlying rights holders will find themselves on the wrong side of the AI revolution. Perhaps more importantly, the developers of these tools and the enterprises that adopt them have both a legal and ethical responsibility to understand the broader consequences implicated by adopting these tools. One important part of that process requires a careful reconsideration of the intellectual property rights being impacted by generative AI across industries.

If the change is managed well, then the opportunities for creativity, efficiency, and expanded resources will support the new economy that emerges from the transformation. It is incumbent upon

everyone involved in the generative AI explosion that transparency, diligence, and responsibility are core values of the deployment.

Vive la révolution!