SAY WHAT?! WHEN CHATGPT GETS IT WRONG: EXAMINING GENERATIVE AI, SECTION 230 OF THE COMMUNICATIONS DECENCY ACT, AND THE ESSENCE OF CREATIVITY

Jason Davidson[*]
Hilary G. Buttrick[**]

Cite as: Jason Davidson & Hilary G. Buttrick, *SAY WHAT?! When ChatGPT Gets it Wrong: Examining Generative AI, Section 230 of the Communications Decency Act, and the Essence of Creativity*, 30 RICH. J.L. & TECH. 143 (2023).

[*] PhD, Assistant Professor of Applied Business Technology, Lacy School of Business, Butler University.

[**] JD, Associate Professor of Business Law; Associate Dean of Academics, Lacy School of Business, Butler University.

I. INTRODUCTION

[1]   Since its November 2022 release, OpenAI's large language model ("LLM"), ChatGPT, has become a household name, outscoring medical students on clinical exams,[1] passing the bar,[2] and polarizing academia.[3] The wildly popular generative AI tool[4] reached 100 million users in just sixty-one days, surpassing tech giants including Netflix, Instagram, and TikTok

---

[1] Adam Hadhazy, *ChatGPT Out-scores Medical Students on Complex Clinical Care Exam Questions*, STAN. UNIV. HUM.-CENTERED A.I. (July 17, 2023), https://hai.stanford.edu/news/chatgpt-out-scores-medical-students-complex-clinical-care-exam-questions#:~:text=A%20new%20study%20shows%20AI's,a%20rethink%20of%-20medical%20education.&text=ChatGPT%20can%20outperform%20first%2D%20and,by%20Stanford%20researchers%20has%20revealed [https://perma.cc/LDC5-7XEF].

[2] Debra Cassens Weiss, *Latest version of ChatGPT aces bar exam with score nearing 90th percentile*, ABA J. (Mar. 16, 2023, 1:59 PM), https://www.abajournal.com/web/article/latest-version-of-chatgpt-aces-the-bar-exam-with-score-in-90th-percentile?spredfast-trk-id=sf176032736 [https://perma.cc/528Y-9C8F].

[3] *See* Lyss Welding, *Half of College Students Say Using AI on Schoolwork Is Cheating or Plagiarism*, BEST COLLEGES, https://www.bestcolleges.com/research/college-students-ai-tools-survey/ [https://perma.cc/L23Z-LY6J] (last updated Mar. 27, 2023).

[4] *See* Kim Martineau, *What is generative AI?*, IBM RSCH. BLOG (Apr. 20, 2023), https://research.ibm.com/blog/what-is-generative-AI [https://perma.cc/3ZKJ-K6TV] (explaining that in recent months, ChatGPT has become the most popular and frequently discussed generative AI tool; accordingly, this article often references ChatGPT. However, the same issues raised in connection with ChatGPT can be applied to generative AI in the broad sense. "Generative AI refers to deep-learning models that can take raw data — say, all of Wikipedia or the collected works of Rembrandt — and 'learn' to generate statistically probable outputs when prompted.").

in reaching that milestone.[5] ChatGPT's robust performance in real-world applications has demonstrated its disruptive capabilities, triggering both societal excitement and concern.[6] The advantages of this species of AI are seemingly endless: content creators can generate ideas, businesses can automate document creation, programmers can generate code, and much more.

[2]   While OpenAI created ChatGPT to "ensure that artificial general intelligence benefits all of humanity,"[7] prominent members of the AI community, such as Turing award winner, Geoffrey Hinton, and OpenAI CEO, Sam Altman, have warned that "the risks of advanced A.I. systems were serious enough to warrant government intervention."[8] These concerns include workforce disruption, intellectual property violations, academic dishonesty, distribution of misinformation, data privacy breaches, bias, and the no longer far-fetched notion of a world in which AI becomes sentient or

---

[5] Krystal Hu, *ChatGPT sets record for fastest-growing user base – analyst note*, REUTERS, https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/ [https://perma.cc/9BDV-H2KW] (last updated Feb. 2, 2023, 10:33 AM); *see* Tufayel Ahmed & Jordan Saville, *Watch: How Netflix Changed TV and Reached 100 Million Subscribers*, NEWSWEEK (May 19, 2017, 8:21 AM), https://www.newsweek.com/netflix-100-million-subscribers-612199 [https://perma.cc/G8VA-M97Z].

[6] *See* Yogesh K. Dwivedi et al., *So What if ChatGPT Wrote It?*, 71 INT'L J. INFO. MGMT. 1, 2 (2023), https://doi.org/10.1016/j.ijinfomgt.2023.102642 [https://perma.cc/8K36-U6AX].

[7] *About OpenAI*, OPENAI, https://openai.com/about [https://perma.cc/3CF5-WJGG] (last visited Oct. 9, 2023).

[8] Kevin Roose, *A.I. Poses 'Risk of Extinction,' Industry Leaders Warn*, N.Y. TIMES (May 30, 2023), https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html [https://perma.cc/SMN8-7DZS].

uncontrollable.[9] Altman stated, "if this technology goes wrong, it can go quite wrong."[10]

[3]   To be sure, a host of legal and ethical issues abound as consumers and businesses embrace generative AI products like ChatGPT. Of paramount concern is whether users have remedies for harms caused by this technology. The purpose of this article is to explore those potential harms in the context of Section 230 of the Communications Decency Act, which provides immunity for technology platforms that publish material based on third-party content.[11] Part I explains the origins of generative AI tools and discusses how they work. Part II addresses how the law could be used to protect consumers from the potential threats posed by generative AI. It includes a discussion of legislative and regulatory initiatives, as well as the applicability of Section 230's immunity provision. Part III discusses the relationship between human creativity and generative AI to tackle the question that Section 230 immunity turns on: Does generative AI create?

II. ORIGINS OF GENERATIVE AI AND HOW CHATGPT WORKS

[4]   In October 1950, British mathematician Alan Turing published the groundbreaking article, *Computing Machinery and Intelligence*, which posed the question, "[c]an machines think?"[12] The article introduced a novel test called the "imitation game" to measure the ability of computers to think.[13] The game involves three players: an interrogator, a human, and a computer impersonating a human.[14] In a blind discussion, if the interrogator

---

[9] *See id.*

[10] *Id.*

[11] 47 U.S.C.S. § 230(c)(1) (LEXIS through Pub. L. No. 118-14).

[12] A.M. Turing, *Computing Machinery and Intelligence*, 59 MIND 433, 433 (1950).

[13] *Id.*

[14] *See id.*

cannot distinguish between the human and the computer, the machine passes the test.[15] Seventy-three years later, the imitation game—now commonly known as the Turing Test—remains the industry benchmark for measuring the capabilities of AI.[16]

[5]   Sixteen years after Turing's publication, MIT professor Joseph Weizenbaum created what is now recognized as the world's first chatbot, ELIZA.[17] ELIZA was a computer application designed to operate as a Rogerian-style psychotherapist.[18] When a user entered their current emotional state, ELIZA would respond using a relaxed vernacular with a natural language cadence.[19] ELIZA was designed using a series of expertly crafted predetermined responses, which gave the appearance of intelligence but was limited to static scripted dialog.[20] While this technique is now obsolete, similarities can be drawn between the scripts used by ELIZA and the training data used to create modern LLMs.[21]

---

[15] *See id.* at 434.

[16] *Turing Test in Artificial Intelligence*, GEEKSFORGEEKS, https://www.geeksforgeeks. org/turing-test-artificial-intelligence/ [https://perma.cc/8FHJ-HSKR] (last updated Feb. 22, 2023).

[17] Eleni Adamopoulou & Lefteris Moussiades, *Chatbots: History, Technology, and Applications*, 2 MACH. LEARNING WITH APPLICATIONS 1, 2 (2020).

[18] *Id.* at 14.

[19] Joseph Weizenbaum, *ELIZA–A Computer Program for the Study of Natural Language Communication Between Man and Machine*, 9 COMMC'NS OF THE ACM 36, 36 (1966).

[20] *See id.* at 37.

[21] Adamopoulou & Moussiades, *supra* note 17.

[6]   Natural language communication, now called natural language processing, is the driving force in modern LLMs such as ChatGPT, [22] Google's Bard,[23] and Anthropic's Claude.[24] Contemporary commercial LLMs like ChatGPT are considered blackbox designs because the product's full internal workings are not fully disclosed.[25] While the exact architecture of ChatGPT is considered a trade secret, much is understood about natural language processing and the techniques required to create an LLM like ChatGPT.[26] As Jon Krohn writes, natural language processing is the "intersection of the fields of computer science, linguistics, and artificial intelligence."[27] From a technical perspective, modern natural language

---

[22] *See* Ross Gruetzemacher, *The Power of Natural Language Processing*, HARV. BUS. REV. (Apr. 19, 2022), https://hbr.org/2022/04/the-power-of-natural-language-processing [https://perma.cc/7LU9-ZVA4].

[23] Mark Sullivan, *Google announces new Bard chatbot to counter ChatGPT*, FAST CO. (Feb. 6, 2023), https://www.fastcompany.com/90845691/google-announces-new-bard-chatbot-to-counter-chatgpt [https://perma.cc/3A6D-8JBQ].

[24] Emma Roth, *Anthropic's 'friendly' AI Chatbot, Claude, is now available for more people to try*, THE VERGE (July 11, 2023, 9:30 AM), https://www.theverge.com/2023/7/11/23790254/anthropic-claude-chatbot-ai-available-beta [https://perma.cc/FAT4-K4KL]; Claude.ai, *Exploring Claude AI's Cutting-Edge Language Capabilities*, CLAUDE AI (July 20, 2023), https://claudeai.uk/exploring-claude-ais-cutting-edge-language-capabilities/ [https://perma.cc/K8PZ-7GVW].

[25] *See* Cynthia Rudin & Joanna Radin, *Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson from an Explainable AI Competition*, 1.2 HARV. DATA SCI. REV. 1, 3 (2019).

[26] *How Does Chat-Gpt Work?*, PATENTPC (Jan. 3, 2023), https://www.patentpc.com/blog/how-does-chat-gpt-work [https://perma.cc/L6MC-NHSJ]; *see generally A Complete Guide To Natural Language Processing*, DEEPLEARNING.AI, https://www.deeplearning.ai/resources/natural-language-processing/ [https://perma.cc/7P8Q-SPXZ] (last updated Jan. 11, 2023) (explaining what NLP is used for and how it works).

[27] JON KROHN ET. AL, DEEP LEARNING ILLUSTRATED A VISUAL, INTERACTIVE GUIDE TO ARTIFICIAL INTELLIGENCE 25 (2019).

processing techniques, like those used to create ChatGPT, are a subset of machine learning called deep learning.[28]

A. Deep Learning

[7]   "Machine learning is the science (and art) of programming computers so they can *learn from data*."[29] The machine learning design process is significantly different from traditional computer programming. Unlike traditional programming, during which inputs are manipulated by manual programming to produce a desired output, machine learning algorithms are given training data comprised of both input *and* output data.[30] The algorithm uses this training data to create a model.[31]

[8]   Deep learning is a subset of machine learning that creates artificial neural networks—a unique type of data model comprised of layers of artificial neurons roughly modeled after the neurons of the human brain.[32] Deep learning artificial neural networks can be used to create classification or prediction models like the LLM used to create ChatGPT.[33] Stated plainly, these LLMs are simply artificial neural networks that predict or add "one

---

[28] *See id.* at 96–97.

[29] AURÉLIEN GÉRON, HANDS-ON MACHINE LEARNING WITH SCIKIT-LEARN, KERAS & TENSORFLOW: CONCEPTS, TOOLS, AND TECHNIQUES TO BUILD INTELLIGENT SYSTEMS 4 (Nicole Tache ed. 2d ed., 2019).

[30] *See id.* at 5–6.

[31] *See id.* at 6.

[32] OSWALD CAMPESATO, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, AND DEEP LEARNING 13–14 (2020).

[33] Sean Michael Kerner, *Definition: large language models (LLMs)*, TECHTARGET, https://www.techtarget.com/whatis/definition/large-language-model-LLM [https://perma.cc/7MXQ-3QET] (last updated Sept. 2023).

word at a time" to a series of generated texts based on probability.[34] Thus, a generative model would simply calculate the probability of the next word in the sentence based on the training data.[35]

[9]   ChatGPT (short for "generative pre-trained transformer") is a generative AI LLM pretrained for broad use of text generation.[36] It receives directions from the user in a natural common language called a "prompt."[37] User prompts provide directives to tailor the desired response from the model.[38] For example, a prompt could instruct ChatGPT to write a poem in the style of Dr. Seuss or write a blues song in the style of B.B. King. Prompts are a key component to this technology as they allow the response to consider context when generating text.[39]

B. Supervised, Unsupervised, and Reinforcement Learning

[10]   A defining characteristic of ChatGPT is the fluidity and natural tone of its generated text. This is achieved through combining multiple types of

---

[34] Stephen Wolfram, *What is ChatGPT Doing … and Why Does It Work?*, STEPHEN WOLFRAM WRITINGS (Feb. 14, 2023), https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/ [https://perma.cc/4Q4H-CH99].

[35] *See id.*

[36] Bernard Marr, *What Does ChatGPT Really Mean For Businesses?*, FORBES (Dec. 28, 2022, 2:21 AM), https://www.forbes.com/sites/bernardmarr/2022/12/28/what-does-chatgpt-really-mean-for-businesses/?sh=6a9fc3e7d1e3 [https://perma.cc/5FKH-S53B].

[37] Akshay K., *Prompt Engineering: What It Is and 15 Techniques for Effective AI Prompting + Tips*, HOSTINGER (Sept. 29, 2023), https://www.hostinger.com/tutorials/ai-prompt-engineering#:~:text...odels%20are%20generative,%20These%20inputs%20are%20called%20prompts [https://perma.cc/4WNN-Z88X].

[38] *Id.*

[39] *See id.*

machine learning.[40] In a supervised learning model, the designer specifies both inputs, called "features," and corresponding outputs, called "labels," in the training data.[41] For example, in a model that simply approves or denies credit based on credit score, the training data would consist of thousands of historical credit applications containing consumer credit scores and corresponding credit approval. In contrast, an unsupervised learning model provides training data with no corresponding labels, allowing the algorithm to identify and create the model based on relationships in the data it discovers.[42] LLMs like ChatGPT begin with unsupervised training on a large set of data.[43] For instance, ChatGPT 3 was trained on 175 billion parameters gleaned from the internet.[44] Then, to help ChatGPT sound more natural in its generated text, designers used reinforcement learning,[45] which helps the program "determine the ideal behavior based upon feedback from the environment."[46] This methodology

---

[40] Wolfram, *supra* note 34.

[41] *Generative AI Models Explained*, ALTEXSOFT (Oct. 12, 2022), https://www.altexsoft.com/blog/generative-ai/ [https://perma.cc/H4KH-UAH2].

[42] *What is machine learning?*, IBM, https://www.ibm.com/topics/machine-learning [https://perma.cc/46YM-Z64P] (last visited Oct. 10, 2023).

[43] Marr, *What Does ChatGPT Really Mean For Businesses?, supra* note 36.

[44] Bernard Marr, *GPT-4 Is Coming–What We Know So Far*, FORBES (Feb. 24, 2023, 2:21 AM), https://www.forbes.com/sites/bernardmarr/2023/02/24/gpt-4-is-coming--what-we-know-so-far/?sh=26ee77c46c2d [https://perma.cc/8U8M-WLWG].

[45] *See Introducing ChatGPT*, OPENAI (Nov. 30, 2022), https://openai.com/blog/chatgpt [https://perma.cc/97NA-MHVC] (stating that the model was trained using reinforcement learning from human feedback, with answers fine-tuned based on quality).

[46] Bernard Marr, *Artificial Intelligence: What Is Reinforcement Learning – A Simple Explanation & Practical Examples*, FORBES (Sept. 28, 2018, 2:00 AM), https://www. forbes.com/sites/bernardmarr/2018/09/28/artificial-intelligence-what-is-reinforcement-learning-a-simple-explanation-practical-examples/?sh=51fdc127139c [https://perma.cc/ J9MK-ZP2J].

is one of the "most active areas of research in artificial intelligence."[47] In the case of ChatGPT, reinforcement learning incentivizes a model to perform more favorably based on *human* feedback.[48]

C. When ChatGPT Gets It Wrong: Hallucinations

[11]   While ChatGPT 1 was launched in June of 2018, it was the November 2022 release of the much more refined ChatGPT 3 that became an overnight success.[49] ChatGPT 1 was trained on 117 million parameters, mostly gleaned from books.[50] At the time of this writing ChatGPT is on its fourth iteration. ChatGPT 4 is rumored to have been trained on more than 17 trillion parameters[51] and has access to the internet to continue to grow its knowledge base.[52] The internet is a largely uncensored medium.[53] Thus, when the bulk of the internet is used as a training model, a host of off-

---

[47] SUDHARSAN RAVICHANDIRAN, HANDS-ON REINFORCEMENT LEARNING WITH PYTHON 6 (2018).

[48] *Introducing ChatGPT*, *supra* note 45.

[49] Bernard Marr, *A Short History of ChatGPT: How We Got To Where We Are Today*, FORBES (May 19, 2023, 1:14 AM), https://www.forbes.com/sites/bernardmarr/2023/05/ 19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=71a8e686674f [https://perma.cc/LP5N-5CG7].

[50] Ainsley Harris, *OpenAI unveils new GPT-4 language model that allows ChatGPT to 'see'*, FAST CO. (Mar. 14, 2023), https://www.fastcompany.com/90865107/openai-gpt-4-language-model-chatgpt [https://perma.cc/NB4D-RUJV]; Marr, *A Short History of ChatGPT*, *supra* note 49.

[51] Marr, *GPT-4 Is Coming*, *supra* note 44.

[52] *See* Kyle Wiggers, *OpenAI connects ChatGPT to the internet*, TECHCRUNCH (Mar. 23, 2023, 1:35 PM), https://techcrunch.com/2023/03/23/openai-connects-chatgpt-to-the-internet/ [https://perma.cc/Y9MY-9G4N] (stating that for a small number of developers, ChatGPT is now able to access third-party sources, including the Internet).

[53] *Id.*

putting, offensive, and explicit material is included.[54] To address this challenge, OpenAI sought to make its product less "toxic" by incorporating content filters and safeguards in its later iterations.[55]

[12]   However, ChatGPT is not infallible. It generates text based on probabilities; in other words, it selects the most likely next word or series of words based on its training data.[56] When an LLM produces a series of words that are incorrect, it generates content that contains misinformation commonly called "hallucinations."[57] For instance, ChatGPT generated an elaborate story about how James Joyce and Vladimir Lenin met at a café in Zurich, an encounter that never happened.[58] Such outlandish hallucinations can be entertaining. Indeed, the authors of this paper received the following tidbit from ChatGPT as they were researching this article: "There are generally 4 or 5 weeks in July, with each week containing 7 days. Therefore, there are a total of 28 to 35 days in the month of July. The exact number of days can vary depending on the year (leap year or not) and the specific days that fall within that month." [59]

---

[54] *See id.* (explaining that a now-disbanded Meta chatbot given Internet access began to respond to certain prompts with conspiracy theories and offensive content).

[55] *See* Billy Perrigo, *OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic*, TIME (Jan. 18, 2023, 7:00 AM), https://time.com/6247678/openai-chatgpt-kenya-workers/ [https://perma.cc/2WBX-BGFN].

[56] Wolfram, *supra* note 34.

[57] *See* Karen Weise & Cade Metz, *When A.I. Chatbots Hallucinate*, N.Y. TIMES, https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucination.html?searchResultPosition=2 [https://perma.cc/7Y3R-W4RK] (last updated May 9, 2023).

[58] *Id.*

[59] OpenAI, Response to "How many days are in a week in July?" CHATGPT (Aug. 14, 2023), https://chat.openai.com/.

[13]   But some of these hallucinations are quite believable, and thus have the potential to mislead consumers.[60] It is also deeply troubling that industry insiders say that generative AI systems are "built to be persuasive, not truthful."[61] In other words, the content is designed to convince readers of its truthfulness with little regard for whether it *is* true. This can be particularly damaging to the consumer—what happens when generative AI publishes false information about an individual, and others rely on that information as truthful because it is so persuasively written? Part II will address these questions.

III. GENERATIVE AI'S IMPACT ON CONSUMERS—LEGAL IMPLICATIONS

[14]   While no one fully comprehends how far generative AI products like ChatGPT *can* go, everyone is concerned about how far generative AI *should* go.[62] The Center for AI Safety stated that "[m]itigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."[63] As embodied in the White House's *Blueprint for an AI Bill of Rights*,[64] there is near-consensus around the need for safe and effective systems, algorithmic discrimination protection, data privacy, basic comprehension of the technology consumers are using, and

---

[60] *See* Gerrit De Vynck, *ChatGPT 'hallucinates.' Some researchers worry it isn't fixable.*, WASH. POST, https://www.washingtonpost.com/technology/2023/05/30/ai-chatbots-chatgpt-bard-trustworthy/ [https://perma.cc/P8WD-28PD] (last updated May 30, 2023, 1:27 PM).

[61] Weise & Metz, *supra* note 57.

[62] Cat Zakrzewski, *FTC investigates OpenAI over data leak and ChatGPT's inaccuracy*, WASH. POST, https://www.washingtonpost.com/technology/2023/07/13/ftc-openai-chatgpt-sam-altman-lina-khan/ [https://perma.cc/KV22-LMRJ] (last updated July 13. 2023, 7:26 PM).

[63] *Statement on AI Risk*, CTR. FOR AI SAFETY, https://www.safe.ai/statement-on-ai-risk [https://perma.cc/BLY5-BJB7].

[64] WHITE HOUSE OFF. SCI. & TECH. POL'Y, BLUEPRINT FOR AN AI BILL OF RIGHTS 1, https://www.whitehouse.gov/ostp/ai-bill-of-rights/ [https://perma.cc/7RTD-5NZ6].

protection of the consumer's choice to interact with a human instead of AI.[65] But the devil is in the details: while everyone seems to agree that lines need to be drawn, the law has been painfully slow to draw them.[66] However, there are three potential paths to mitigate the risks of generative AI: (1) legislation to protect consumers and curb generative AI abuses;[67] (2) administrative enforcement of existing laws applied in the generative AI context;[68] and (3) litigation.[69]

A. Legislative Solutions

[15]   Congressional concerns about AI are nothing new. For years, Congress has been worried about Big Tech.[70] As one commentator noted of Big Tech at large, both sides of the political spectrum recognize "the perils of Big Tech and the desperate need to do something about it. The perspective that Big Tech is a force that is uniquely deleterious to a healthy body politic is a view shared by conservatives and liberals alike."[71] Congress has held countless hearings, with advocates airing concerns about

---

[65] *Id.*

[66] *See* Mikaela Cohen, *Why waiting for A.I. laws, regulations from government could be a catastrophic mistake,* CNBC, https://www.cnbc.com/2023/07/19/waiting-on-ai-regulations-from-government-may-be-catastrophic-mistake.html [https://perma.cc/SE2Y-Y6CN] (last updated July 19, 2023, 11:44 AM).

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

[68] *See infra* Section III.B.

[69] *See infra* Section III.C.

[70] Alex Sherman, *U.S. lawmakers agree Big Tech has too much power, but what to do about it remains a mystery,* CNBC, https://www.cnbc.com/2020/07/30/us-lawmakers-agree-big-tech-has-too-much-power-remedies-unclear.html [https://perma.cc/S77Z-V3ZT] (last updated July 30, 2020, 2:29 PM).

[71] Roger P. Alford, *The Bipartisan Consensus on Big Tech*, 71 EMORY L.J. 893, 898 (2022).

misinformation, exploitation and human trafficking, deep fakes, and Big Tech's outsized influence on matters of societal importance, such as elections, national security, and privacy.[72] Others are concerned with the monopoly power of Big Tech, leading to calls for more stringent antitrust measures.[73] Congress has called tech executives to testify and grilled them about their internal policies and accountability measures.[74] Some tech executives seem equally baffled about how to proceed with generative AI in a responsible way and have called for regulation of the industry.[75] Throw in the complicated interplay between the First Amendment, the idealized vision of the internet as a virtual town square, and the fact that tech companies are private actors with discretion to remove troubling content posted on their platforms, and the morass feels truly inescapable.[76] Despite

---

[72] *Id*. at 906–07.

[73] *Id*. at 908–09 (noting that Federal bills have focused on decreasing the monopoly power of Big Tech by providing more money for federal antitrust enforcement and enhancing state power to enforce federal antitrust laws); Brian Fung*, The US government is still trying to find ways to regulate Big Tech. He has some ideas,* CNN*,* https://www. cnn.com/2023/01/11/tech/jonathan-kanter-doj/index.html [https://perma.cc/E7N8-S73V] (last updated Jan. 11, 2023, 9:06 AM).

[74] Ben Clements, *The Big Tech Accountability Act: Reforming How the Biggest Corporations Control and Exploit Online Communications*, 44 W. NEW ENG. L. REV. 5, 7 (2022).

[75] Brian Fung, *Mr. ChatGPT goes to Washington: OpenAI CEO Sam Altman testifies before Congress on AI risks*, CNN, https://www.cnn.com/2023/05/16/tech/sam-altman-openai-congress/index.html [https://perma.cc/XA7F-E8VC] (last updated May 16, 2023, 1:34 PM).

[76] *See* Clements, *supra* note 74, at 8–9; *see also* Reese D. Bastian, Note*, Content Moderation Issues Online: Section 230 Is Not to Blame*, 8 TEX. A&M J. PROP. L. 43, 45 (2022) (noting that the sentiment that the "[i]nternet provides opportunities for free speech and promotes the marketplace of ideas" has changed as some believe internet companies are inappropriately controlling what can be seen online) [https://perma.cc/6RNX-TKGF].

calls for federal regulation from consumers, policymakers, and the tech industry itself, by and large, bills have stalled.[77]

[16]   If the concerns around Big Tech seemed urgent before ChatGPT, now they have reached a level of near-panic—even the Pope is worried.[78] Members of Congress recognize the need to do something, but what?[79] An outright ban, like the one Italy temporarily instituted against ChatGPT in 2023,[80] is a blunt instrument that would deprive consumers and industry of a tool that can dramatically increase productivity. Another approach would be developing a regulatory regime that requires generative AI to meet certain safety standards before being licensed to operate.[81] Other proposals could focus on curbing certain uses and abuses of generative AI, such as

---

[77] Clements, *supra* note 74, at 7.

[78] *See* Anders Hagstrom, *Pope issues warning on artificial intelligence, fears 'logic of violence'*, FOX NEWS (Aug. 8, 2023, 9:37 AM), https://www.foxnews.com/world/pope-issues-warning-artificial-intelligence-fears-logic-violence [https://perma.cc/ZE54-UH9C].

[79] *Cf.* Reuters, *Schumer calls for AI rules as ChatGPT surges in popularity*, NBC NEWS (Apr. 13, 2023, 1:11 PM), https://www.nbcnews.com/politics/congress/chuck-schumer-calls-ai-rules-chatgpt-surges-popularity-rcna79582 [https://perma.cc/P9PD-SBG3] (discussing how Schumer's regulatory regime proposal represents the first tangible initiative of any member Congress of Congress around regulating AI systems but still has many hurdles to achieve success).

[80] Bradford Betz, *Italy reverses ban on ChatGPT after OpenAI agrees to watchdog's demands,* FOX BUS. (May 3, 2023, 8:57 PM) https://www.foxbusiness.com/technology/italy-reverses-ban-chatgpt-openai-agrees-watchdogs-demands [https://perma.cc/2YS4-HHZN]; *see also* Shana Lynch, *Analyzing the European Union AI Act: What Works, What Needs Improvement*, HAI: STAN. UNIV. HUM.-CENTERED A.I. (July 21, 2023), https://hai.stanford.edu/news/analyzing-european-union-ai-act-what-works-what-needs-improvement#:~:text=The%20EU%20AI%20Act%20is,its%20position%20in%20mid%2DJune [https://perma.cc/MD3A-EVKK] (stating that the U.S. is not the only nation considering regulations. The European Union is close to adopting a comprehensive AI law known as the EU AI Act).

[81] *See* Reuters, *supra* note 79.

those that result in exploitation, political censorship, or the publication of illegal third-party content if the platform has knowledge of the illegality.[82]

[17]   One such piece of legislation was introduced in June 2023 when two U.S. Senators put forth a bipartisan proposal that specifically targets generative AI.[83] The bill, titled "No Section 230 Immunity for AI Act," would make it easier for consumers to sue generative AI companies for claims arising from material generated by their platforms. [84] As discussed in Section III.C, Section 230 of the Communications Decency Act[85] effectively shields technology companies from civil liability for "information originating with a third-party user of the service."[86] In a press release, Senator Josh Hawley said that "[w]e can't make the same mistakes with generative AI as we did with Big Tech on Section 230 . . . When these new technologies harm innocent people, the companies must be held accountable. Victims deserve their day in court and this bipartisan proposal

---

[82] *See, e.g.*, Christian Sarceño Robles, *Section 230 Is Not Broken: Why Most Proposed Section 230 Reforms Will Do More Harm Than Good, and How the Ninth Circuit Got It Right*, 16 FIU L. REV. 213, 220 (2021).

[83] Ashley Gold & Andrew Solender, *First look: Bipartisan bill denies Section 230 protection for AI*, AXIOS PRO (June 14, 2023), https://www.axios.com/pro/tech-policy/2023/06/14/hawley-blumenthal-bill-section-230-ai [https://perma.cc/N8HD-RPAB].

[84] No Section 230 Immunity for AI Act, S. 1993, 118th Cong. (as introduced in the Senate, June 14, 2023).

[85] 47 U.S.C.S. § 230 (LEXIS through Pub. L. No. 118-14).

[86] Zeran v. America Online, Inc., 129 F.3d 327, 330 (4th Cir. 1997).

will make that a reality."[87] Although this bill demonstrates a keen interest in protecting consumers, it has not become law.

B. Administrative Oversight: Using Existing Laws and Agencies

[18]   If enacting a comprehensive legislative solution seems unlikely, perhaps existing regulatory schemes can be leveraged to address some of the concerns about data privacy and false information raised by generative AI. The Federal Trade Commission ("FTC"), for instance, has announced an investigation into ChatGPT to determine whether it has "(1) engaged in unfair or deceptive privacy or data security practices or (2) engaged in unfair or deceptive practices relating to risks of harm to consumers, including reputational harm, in violation of Section 5 of the FTC Act."[88] The FTC has sent interrogatories and document requests to learn how ChatGPT works, what steps OpenAI takes to mitigate risks associated with its product, and any knowledge the company has of vulnerabilities in its data security procedures.[89]

[19]   As a regulatory agency created under the Federal Trade Commission Act, the FTC has authority to police unfair methods of competition and unfair or deceptive acts.[90] Fundamentally, the FTC's role is to ensure that

---

[87] Press Release, Josh Hawley, U.S. Senator for Missouri, & Richard Blumenthal, U.S. Senator for Connecticut, Hawley, Blumenthal Introduce Bipartisan Legislation to Protect Consumers and Deny AI Companies Section 230 Immunity (June 14, 2023), https://www.hawley.senate.gov/hawley-blumenthal-introduce-bipartisan-legislation-protect-consumers-and-deny-ai-companies-section [https://perma.cc/8NX5-32LU].

[88] FED. TRADE COMM'N, *Federal Trade Commission ("FTC") Civil Investigative Demand ("CID") Schedule FTC File No. 232-3044*, WASH. POST, 2, https://www.washingtonpost.com/documents/67a7081c-c770-4f05-a39e-9d02117e50e8.pdf?itid=lk_inline_manual [https://perma.cc/P72M-M3T3].

[89] *Id.*

[90] 15 U.S.C.S. § 45 (LEXIS through Pub. L. No. 118-13).

consumers have accurate information when making decisions and to ensure that the market is competitive.[91] To that end, the FTC enforces anti-fraud[92] and antitrust laws.[93] In addition, Congress has given the FTC authority to enforce several specific data privacy statutes.[94] To do so, the FTC has a substantial toolkit, including the power to require "implementation of comprehensive privacy and security programs, biennial assessments by independent experts, monetary redress to consumers, disgorgement of ill-gotten gains, deletion of illegally obtained consumer information, and . . . robust transparency and choice mechanisms [for] consumers."[95] In addition, the FTC has authority to bring civil lawsuits against companies that violate these statutes.[96]

[20]   However, the FTC's authority is not boundless. As an administrative agency, it can only enforce the statutes it has been authorized to enforce.[97] So while the FTC has authority to protect consumers from deceptive practices and penalize businesses who misuse consumer data under specific

---

[91] *Mission*, FED. TRADE COMM'N, https://www.ftc.gov/about-ftc/mission [https://perma.cc/38YK-5EC8] (last visited Oct. 7, 2023).

[92] *Enforcement*, FED. TRADE COMM'N, https://www.ftc.gov/enforcement [https://perma.cc/7XKE-JBGK] (last visited Aug. 11, 2023).

[93] *The Enforcers,* FED. TRADE COMM'N, https://www.ftc.gov/advice-guidance/ competition-guidance/guide-antitrust-laws/enforcers [https://perma.cc/7HXP-3QQS] (last visited Aug. 11, 2023).

[94] FED. TRADE COMM'N, PRIVACY & DATA SECURITY UPDATE: 2018, 11 (2018), https://www.ftc.gov/system/files/documents/reports/privacy-data-security-update-2018/2018-privacy-data-security-report-508.pdf [https://perma.cc/8DDY-BGZM].

[95] *Id.* at 2.

[96] *Id.*

[97] Cheng v. WinCo Foods LLC, No. 14-cv-0483, U.S. Dist. LEXIS 81069, at *6 (N.D. Cal. June 10, 2014).

statutes,[98] it is less clear whether the FTC has carte blanche to pursue technology companies for all harms sustained by consumers. The FTC's attempt to obtain ChatGPT's records regarding "reputational harm" raised eyebrows among those who take a less expansive view of regulatory authority.[99] "Reputational harm" typically refers to state civil law claims sounding in tort, such as libel, slander, or defamation.[100] These torts are designed to give individuals a remedy when they have been damaged by the publication of a false statement that tends to harm their reputation.[101] As such, reputational harm is a personal tort that has traditionally been left to individuals to pursue on their own.[102] The FTC's authority to pursue such a claim under the banner of consumer fraud is questionable. Without additional statutory authority to police generative AI, the FTC's reach is limited to enforcing a defined group of existing consumer protection statutes—none of which were crafted with the unique harms of generative AI in mind.[103]

---

[98] FED. TRADE COMM'N, FTC'S USE OF ITS AUTHORITIES TO PROTECT CONSUMER PRIVACY AND SECURITY, 1 (2020), https://www.ftc.gov/system/files/documents/reports/reports-response-senate-appropriations-committee-report-116-111-ftcs-use-its-authorities-resources/p065404reportprivacydatasecurity.pdf [https://perma.cc/6X4W-XWJL].

[99] Zakrzewski, *supra* note 62.

[100] *See generally* Nakamoto v. Kawauchi, 418 P.3d 600, 611–12 (Haw. 2018).

[101] *See id.*

[102] *See* Bryson Kern, *Reputational Injury Without a Reputational Attack: Addressing Negligence Claims for Pure Reputational Harm*, 83 FORDHAM L. REV. 253, 265, 268 (2018).

[103] Anthony E. DiResta & Zachary E. Sherman, *The FTC Is Regulating AI: A Comprehensive Analysis*, HOLLAND & KNIGHT (July 25, 2023), https://www.hklaw.com/en/insights/publications/2023/07/the-ftc-is-regulating-ai-a-comprehensive-analysis [https://perma.cc/P7DP-RSZK].

C. Litigation and Section 230 Immunity

1. Litigation as a Check on Generative AI

[21]   If the concern with generative AI is harm to consumers, then the quickest path to redressing those harms may be civil litigation. The legislative process and the regulatory state were not built for speed, and it will be next to impossible for them to keep up with generative AI's explosive growth.[104] Litigation, on the other hand, makes ready use of existing legal theories and can be pursued by any individual who has suffered a harm, as long as the elements of the cause of action are satisfied.[105] For example, as discussed in Section III.B, defamation is a tort designed to redress reputational harm. To the extent generative AI publishes defamatory content, a tort has occurred that could be pursued in state court by the harmed plaintiff. Indeed, one such case already has been filed in Georgia based on an alleged ChatGPT hallucination.[106] According to the complaint, a journalist asked ChatGPT to produce a summary of an existing lawsuit.[107] ChatGPT's summary claimed that the plaintiff, Mr. Walters, had been accused of misappropriation of funds.[108] In reality, Mr. Walters was

---

[104] *See* Pamela Samuelson, *Generative AI Meets Copyright*, 381 SCIENCE 158, 159 (2023) ("What makes generative AI more disruptive than previous technologies? One factor is certainly the exceptionally rapid pace at which generative AI technologies have been launched, adopted, and adapted. Evolution in the fields of law and policy, by contrast and of necessity, is much slower").

[105] *See* Christopher J. Valente et al., *Recent Trends in Generative Artificial Intelligence Litigation in the United States*, K&L GATES (Sept. 5, 2023), https://www.klgates.com/ Recent-Trends-in-Generative-Artificial-Intelligence-Litigation-in-the-United-States-9-5- 2023 [https://perma.cc/7EQS-5TGP].

[106] *See* Complaint, Walters v. OpenAI, L.L.C., No. 23-A-04860-2 (Ga. Super. Gwinnett Cnty. June 2, 2023), https://www.courthousenews.com/wp-content/uploads/2023/06/ walters-openai-complaint-gwinnett-county.pdf [https://perma.cc/D7X9-K5HA].

[107] *Id.* at 9, 14–15, 26.

[108] *Id.* at 16.

not even a party to the lawsuit ChatGPT summarized, and he had never been accused of misappropriation of funds.[109] Mr. Walters has sued OpenAI claiming that ChatGPT's summary was wholly untrue and defamatory.[110]

[22]   One also can imagine negligence-based[111] claims against generative AI companies for failing to implement appropriate safeguards to prevent bad actors from using their product in dangerous and criminal ways, particularly if they have actual knowledge of existing abuses. Product liability theories could also be relevant—particularly failure to warn and design defect claims.[112] Further, intentional or negligent infliction of emotional distress[113] could present theories of recovery for consumers who must deal with the fallout of an extreme and outrageous ChatGPT hallucination. In addition, users of ChatGPT may have contract or warranty-based[114] claims to the extent the product does not measure up to the expectations of the bargain. Of course, property and privacy concerns abound where generative AI uses proprietary or protected information in an

---

[109] *Id.* at 16–18.

[110] *Id.* at 17, 32–33, 37.

[111] *See* Ynfante v. Google L.L.C., No. 22-cv-6831, 2023 U.S. Dist. LEXIS 96074, at *5, *11 (S.D.N.Y. June 1, 2023, revised July 10, 2023) (holding that plaintiff's negligence claims were barred by Section 230).

[112] *See* Lemmon v. Snap, Inc., 995 F.3d 1085, 1092–93 (9th Cir. 2021) (holding that Section 230 immunity did not apply where the plaintiffs' claims arose from the negligent design of the product, not from the defendant's conduct as a publisher or speaker).

[113] *See* Herrick v. Grindr, L.L.C., No. 17-CV-932, U.S. Dist. LEXIS 26651, at *6–8 (S.D.N.Y. Feb. 24, 2017) (holding that section 230 immunity applied to bar plaintiff's emotional distress claims).

[114] *See* McDonald v. LG Elecs. USA, Inc., 219 F. Supp. 3d 533, 538 (D. Md. 2016) (declining to dismiss Plaintiff's breach of warranty claims on Section 230 grounds because the basis of the claim did not seek to hold the defendant liable as a speaker or publisher of third-party content; rather, the plaintiff alleged that the defendant was directly responsible its own tortious conduct).

unauthorized way.[115] At least one lawsuit has been filed against ChatGPT for copyright infringement.[116] And one California consumer class action asserts claims for invasion of privacy, conversion, various state consumer fraud violations, intrusion upon seclusion, and unjust enrichment.[117] In short, a creative plaintiff's lawyer can pull together any number of theories to seek redress for harm caused by generative AI. But whether any or all those claims will be *allowed* to proceed is another question. The successful plaintiff will first have to overcome the Section 230 immunity defense.

2. Section 230 Immunity

[23]   Section 230 effectively shields technology companies from civil liability for "information originating with a third-party user of the service."[118] In other words, Section 230 makes it impossible for a consumer to successfully pursue a claim against Instagram, Facebook, Twitter, or any other "interactive computer service" [119] arising from a post made by a third party on those platforms. The Communications Decency Act was passed in

---

[115] *See generally* Samuelson, *supra* note 104, at 158–61; *see also* Robles, *supra* note 82, at 216 ("Section 230 does also carve out limitations for its immunity provisions. For example, Section 230 specifically states it has no effect on federal criminal law, intellectual property law, and the Electronic Privacy Communications Act.").

[116] Matt G. Southern, *ChatGPT Creator Faces Multiple Lawsuits Over Copyright & Privacy Violations*, SEARCH ENGINE J. (July 3, 2023), https://www.searchenginejournal.com/chatgpt-creator-faces-multiple-lawsuits-over-copyright-privacy-violations/490686/ [https://perma.cc/V987-Z3BU].

[117] Complaint at 1–2, P.M., et al. v. OpenAI LP, No. 3:23-cv-03199 (N.D. Cal. June 28, 2023), https://clarksonlawfirm.com/wp-content/uploads/2023/06/0001.-2023.06.28-OpenAI-Complaint.pdf [https://perma.cc/W9LH-GAP8].

[118] Zeran v. America Online, Inc., 129 F.3d 327, 330 (4th Cir. 1997).

[119] 47 U.S.C.S. § 230(f)(2) (LEXIS through Pub. L. No. 118-14) (defining an "interactive computer service" as "any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server . . .").

1996, during the very early days of the internet.[120] At that time, Congress recognized the potential for the internet to enhance communication and commerce, even if it could not foresee how completely the internet would transform the daily lives of billions of people around the world.[121] Congress wanted platforms to have the freedom to moderate objectionable third-party content if need be, without concern that they would be held liable for that content or their decision to moderate all, parts, or none of it.[122] In the absence of such immunity, Congress feared that litigation would have a chilling effect on the fledgling industry.[123] As long as the internet platform merely served as a publisher of others' content, no liability would attach.[124] As Christian Robles notes, "Section 230 created an internet where online platforms are not liable for user generated content but remain free to moderate as much or as little of that content without exposing themselves to liability."[125]

---

[120] Robles, *supra* note 82, at 213–14.

[121] Kathryn Montalbano, *Reimagining Section 230 and Content Moderation: Regulating Incivility on Anonymous Digital Platforms*, 27 COM. L. & POL'Y 187, 190 (2022) ("Section 230 was intended to promote free speech in the fledgling digital media industry and to foster the growth of start-ups by protecting them from endless, financially crippling lawsuits based on the irresponsible actions and speech of their users in a time when there were only a few subscription-based services available, such as Prodigy and AOL.").

[122] *Id.* at 190–91.

[123] *Zeran*, 129 F.3d at 331.

[124] *Id.*

[125] Robles, *supra* note 82, at 216.

[24]   As the internet has developed, courts have defined the boundaries for Section 230 immunity.[126] For Section 230 immunity to apply, three elements must be satisfied.[127] First, the defendant must be a provider or user of an interactive computer service.[128] Second, the claim must be based on information provided by another information content provider.[129] And third, the claim must treat the defendant as a publisher.[130] If all three elements are met, then Section 230 immunity protects the platform.[131] Thus, a critical question is whether the platform is merely publishing third-party content, or whether the platform is itself an "information content provider." The statute defines an "information content provider" as "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."[132]

[25]   With the advent of recommendation algorithms that *curate* content, and generative AI that arguably *creates* and *develops* content, the internet is a far cry from bulletin board-style format of the 1990s. Today's modern algorithms make content recommendations based on the user's past viewing history, and generative AI can produce a painting, a song, a term paper, a

---

[126] Ashley Johnson & Daniel Castro, *The Exceptions to Section 230: How Have the Courts Interpreted Section 230?*, INFO. TECH. & INNOVATION FOUND. (Feb. 22, 2021), https://itif.org/publications/2021/02/22/exceptions-section-230-how-have-courts-interpreted-section-230/ [https://perma.cc/YE5G-E9CA].

[127] FTC v. LeadClick Media, LLC, 838 F.3d 158, 173 (2d Cir. 2016).

[128] *Id.* at 174.

[129] *Id.*

[130] *Id.* at 174–75.

[131] *Id.* at 173.

[132] 47 U.S.C.S. § 230(f)(3) (LEXIS through Pub. L. No. 118-13) (emphasis added).

survey, and even a course syllabus based on the user's prompt.[133] Are we witnessing the creation of new content, or just the clever rearrangement of third-party content? Whether Section 230's liability shield will protect generative AI from consumer lawsuits turns on that question.[134]

3. Is Generative AI a "Content Provider"?

[26]   No appellate court has addressed Section 230's application to ChatGPT, but commentators agree that immunity likely will turn on whether courts think generative AI is a "content provider" under the statute.[135] Thus, the central question will be whether generative AI is creating or developing the actionable content that forms the basis of the plaintiff's claim.[136] "Create" and "develop" are undefined terms in the statute; thus, courts have been left to interpret the words in a host of cases involving various internet platforms.[137]

[27]   The seminal case on this point is *Fair Housing Council v. Roommates.com*, *LLC*, which found that a website engages in

---

[133] *See What is generative AI?*, MCKINSEY & CO. (Jan. 19, 2023), https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai#/ [https://perma.cc/R4D3-XG7Z].

[134] Avi Weitzman & Jackson Herndon, *Generative AI: The Next Frontier for Section 230 of the Communications Decency Act*, N.Y. L. J. (June 26, 2023, 10:00 AM), https://www.law.com/newyorklawjournal/2023/06/26/generative-ai-the-next-frontier-for-section-230-of-the-communications-decency-act/ [https://perma.cc/B6L8-NF34].

[135] *See* Matt Perault, *Section 230 Won't Protect ChatGPT*, 3 J. FREE SPEECH L. 363, 365 (2023) (arguing that ChatGPT is a "content provider"); *see also* Jess Miers, *Yes, Section 230 Should Protect ChatGPT And Other Generative AI Tools*, TECHDIRT (Mar. 17, 2023, 11:59 AM), https://www.techdirt.com/2023/03/17/yes-section-230-should-protect-chatgpt-and-others-generative-ai-tools/ [https://perma.cc/7CAQ-RFR3] (arguing that ChatGPT is not a "content provider").

[136] *Id.*

[137] 47 U.S.C.S. § 230(f)(3) (LEXIS through Pub. L. No. 118-13).

"development" of unlawful content if it materially contributes to the illegality of that content.[138] In contrast, a platform's use of "*neutral* tools to carry out what may be unlawful or illicit searches does not amount to 'development' for purposes of the immunity exception."[139] In that case, a website developer sought to match individuals seeking housing with available rooms.[140] To facilitate the match, the website prompted users to answer a series of questions related to protected characteristics under California's fair housing laws, resulting in discriminatory and unlawful questions.[141] By posting these questions and requiring users to answer them, the website contributed to the development of the content provided by the users.[142]

[28]   The court contrasted those screening questions with an open text box where users could write "additional comments"—discriminatory or not.[143] Section 230 immunity *did* protect the open text box material because the website did nothing to shape the subscribers' responses:

> Where it is very clear that the website directly participates in developing the alleged illegality—as it is clear here with respect to Roommate's questions, answers, and the resulting profile pages—immunity will be lost. But in cases of enhancement by implication or development by inference—

---

[138] Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1166–68 (9th Cir. 2007).

[139] *Id.* at 1169.

[140] *Id.* at 1161.

[141] *Id.* at 1161–62.

[142] *Id.* at 1166.

[143] *Roommates*, 521 F.3d at 1173–74.

such as with respect to the "Additional Comments" here—
section 230 must be interpreted to protect websites . . . .[144]

[29]   The court pointed out that it is possible for parts of a platform
to be protected by the immunity shield, while other parts could be
exposed to liability—it depends on the nature of the platform's
engagement with the material:

> A website operator can be both a service provider and a
> content provider: If it passively displays content that is
> created entirely by third parties, then it is only a service
> provider with respect to that content. But as to content that it
> creates itself, or is "responsible, in whole or in part" for
> creating or developing, the website is also a content
> provider. Thus, a website may be immune from liability
> for some of the content it displays to the public but be subject
> to liability for other content.[145]

[30]   Other courts have adopted or applied *Roommates*' "material
contribution test" in evaluating whether platforms are "developing" or
"creating" content and thus cannot avail themselves of Section 230
immunity.[146] For instance, in *Jones v. Dirty World Entertainment
Recordings LLC*, the Sixth Circuit held, "[a] material contribution to the
alleged illegality of the content does not mean merely taking action that is
necessary to the display of allegedly illegal content. Rather, it means being
responsible for what makes the displayed content allegedly unlawful."[147]
That court granted Section 230 immunity to a website called "The Dirty,"

---

[144] *Id.* at 1174–75.

[145] *Id.* at 1162–63.

[146] Jones v. Dirty World Ent. Recordings, 755 F.3d 398, 410 (6th Cir. 2014).

[147] *Id.*

where users were encouraged to post local gossip.[148] However, because the website did not *require* users to post actionable content, compensate users for submitting actionable content, or ask users to provide information beyond the "who, what, where, when, and why" of "what's happening," they were not "developers" who materially contributed to the unlawfulness of the posts.[149] Thus, Section 230 immunity applied.[150]

[31]   In *O'Kroley v. Fastcase, Inc.*, the Sixth Circuit found that Section 230 immunity was appropriately granted to Google in a defamation suit.[151] In that case, the plaintiff's Google search of his own name produced a summarized search result of a court record that contained multiple cases.[152] The Google summary showed the plaintiff's name separated by an ellipsis from the words "indecency with a child."[153] In fact, the plaintiff had never been involved in a child indecency case; rather, the plaintiff was involved in a wholly different case that was listed in the court record immediately after the child indecency case.[154] The plaintiff argued that the condensed Google summary made it appear as though the plaintiff was involved with the child indecency case.[155]

---

[148] *Id.* at 417.

[149] *Id.* at 416.

[150] *See id.* at 415–16.

[151] O'Kroley v. Fastcase, Inc., 831 F.3d 352, 354 (6th Cir. 2016).

[152] *Id.*

[153] *Id.*

[154] *Id.*

[155] *See id.* at 354–55.

[32]   The court found that Google merely summarized existing internet content and thus was entitled to immunity under Section 230.[156] Although Google performed some edits to the text, such as changing the font and removing spaces, the court found that this was not "development" or "creation" under the statute.[157] Rather, those acts came within the traditional editorial functions of a publisher and were thus entitled to Section 230 immunity.[158] The court also rejected the plaintiff's argument that editing the court records in such a way that the ellipsis separated the plaintiff's name from the phrase "indecency with a child" was a material contribution to the alleged defamatory nature of the post.[159] The court reasoned that the ellipsis placement was not a material contribution because the ellipsis already existed in the court record.[160] Again, the court found that Google was merely "reproducing this third-party content."[161]

[33]   Similarly, the D.C. Circuit extended Section 230 immunity to Google in *Marshall's Locksmith Service, Inc. v. Google, LLC*.[162] In that case, the plaintiffs alleged that Google published information created by "scam" locksmiths, and further that Google displayed the scam locksmiths' information in a map format that made it appear as though the scam locksmiths were local.[163] The court found that the reposting of the scam information was clearly protected by Section 230, and even though Google

---

[156] *O'Kroley*, 831 F.3d at 355.

[157] *Id.*

[158] *Id.*

[159] *Id.*

[160] *Id.*

[161] *O'Kroley*, 831 F.3d at 355.

[162] Marshall's Locksmith Serv. v. Google, L.L.C., 925 F.3d 1263, 1271 (D.C. Cir. 2019).

[163] *Id.* at 1266.

arguably knew the information was false, immunity applied "regardless of whether the defendant acquired the knowledge that the third-party content it published was false."[164] The plaintiffs then argued that in displaying the scam information in a map format, Google had in essence created new content and was thus not entitled to immunity.[165] The court also rejected this argument because the information that formed the basis of the map was provided by third parties:

> The decision to present this third-party data in a particular format—a map—does not constitute the "creation" or "development" of information for purposes of *§ 230(f)(3)*. The underlying information is entirely provided by the third party, and the choice of presentation does not itself convert the search engine into an information content provider. Indeed, were the display of this kind of information not immunized, nothing would be: every representation by a search engine of another party's information requires the translation of a digital transmission into textual or pictorial form.[166]

[34]   This holding rested on the fact that Google used a neutral algorithm to translate the third-party information into the map format.[167] Because the algorithms did not distinguish between legitimate locksmith content and scam locksmith content in the creation of the map (both types of content received the same treatment by Google), the algorithm was considered neutral and thus not a "creator" or "developer" of content.[168]

---

[164] *Id.* at 1269.

[165] *Id.*

[166] *Id.*

[167] *Marshall's*, 925 F.3d at 1271.

[168] *Id.*

[35]   The Second Circuit also extended Section 230 immunity to an internet platform in *Force v. Facebook, Inc*., where the plaintiffs alleged that Facebook was liable under federal anti-terrorism statutes for providing Hamas with a platform to plan and orchestrate terrorist attacks.[169]

[36]   In particular, the plaintiffs alleged that Facebook's algorithm disseminated extremist content to the individuals who carried out the terrorist attack.[170] They claimed that the Facebook algorithms were "developers" of the Hamas content because the algorithms directed content to "users who [were] most interested in Hamas and its terrorist activities, without those users necessarily seeking that content."[171] The court rejected this argument because (1) Facebook did not edit the content Hamas posted; (2) Facebook acquires only basic information from users and thus acts as a "neutral intermediary;" and (3) the algorithms were neutrally applied—in other words, Facebook's algorithm made matches based on objective factors regardless of the content it was matching: "Merely arranging and displaying others' content to users of Facebook through such algorithms— even if the content is not actively sought by those users—is not enough to hold Facebook responsible as the 'developer' or 'creator' of that content."[172]

[37]   Other courts have applied *Roommates*' material contribution test and refused to afford Section 230 immunity.[173] For instance, in *Henderson v. Source for Public Data, L.P*., the Fourth Circuit refused to give Section 230 immunity to a company that collected and sold information about

---

[169] Force v. Facebook, Inc., 934 F.3d 53, 57 (2d Cir. 2019).

[170] *Id.* at 59.

[171] *Id.* at 68.

[172] *Id.* at 69–70.

[173] *See* Henderson v. Source for Pub. Data, L.P., 53 F.4th 110, 126–27 (4th Cir. 2022).

individuals.[174] The defendant in that case collected public records about individuals and then put that information into a summarized, proprietary format.[175] In some cases, the defendant reformatted the public records, and in other cases, they "distilled" the information contained in the records to summaries. [176] Then, the summarized reports were made available to subscribers.[177] The plaintiff alleged that the defendant provided a misleading, inaccurate report to the plaintiff's prospective employer during a background check.[178] The defendant argued that its actions were protected by Section 230 because it relied on third-party content and thus did not "create" or "develop" the content at issue.[179] Instead, "the company is adding new content to the message that harms the plaintiff," and Section 230 immunity was inapplicable.[180]

[38]   Similarly, in *FTC v. LeadClick Media, LLC*, the Second Circuit refused to afford Section 230 immunity to a defendant who participated in the development of deceptive content by actively recruiting advertising affiliates that used false news sites, and then directing those affiliates to edit the content on their pages in a particular way.[181] This level of engagement with the posted content went beyond neutral assistance; thus, the defendant was considered an information content provider and not protected by

---

[174] *Id.*

[175] *Id.* at 117–18.

[176] *Id.*

[177] *Id.* at 118.

[178] *Henderson*, 53 F.4th at 118.

[179] *Id.* at 117.

[180] *Id.* at 128.

[181] F.T.C. v. LeadClick Media, L.L.C., 838 F.3d 158, 162, 172, 176 (2d Cir. 2016).

Section 230 immunity.[182] *Leadclick* stands for the important proposition that where a platform's own deceptive acts or business practices are at issue, no Section 230 immunity attaches.[183]

[39]   In addition, courts have rejected Section 230 immunity defenses where the defendant is the *author* of the allegedly unlawful content.[184] In *La Liberte v. Reid*, the defendant created two posts claiming that the plaintiff made racist remarks at a city council meeting.[185] The defendant created the two posts after seeing and retweeting another post that showed a photo of the plaintiff at the meeting interacting with a minority teenager.[186] The original post had a caption stating that someone yelled at the teenager, but the original post *did not* say that the plaintiff yelled at the teenager.[187] In fact, the plaintiff did not yell at the teenager, and she sued the defendant for defamation.[188] The defendant claimed that her two posts were based on the original photograph posted by someone else; thus, she should be entitled to Section 230 immunity.[189] The court disagreed because the defendant created the two posts that attributed the racist comments to the plaintiff—

---

[182] *Id.* at 175–76.

[183] *Id.* at 176–77.

[184] *See* La Liberte v. Reid, 966 F.3d 79, 83, 89 (2d Cir. 2020).

[185] *Id.* at 83–84.

[186] *Id.*

[187] *Id.* at 89.

[188] *Id.* at 85.

[189] *La Liberte*, 966 F.3d at 89–90.

the original post did not.[190] She was wholly the author of those two posts and thus not simply reposting third-party content.[191]

[40]   It is important to note that the United States Supreme Court had an opportunity to address the scope of Section 230 in the context of modern social media algorithms in the 2023 cases of *Twitter, Inc. v. Taamneh*[192] and *Gonzalez v. Google, LLC*.[193] While the Court did not reach the issue of Section 230, its discussions of the extent to which recommendation algorithms are passive filters or active participants in content generation are instructive.[194] Both cases involved allegations that Twitter, Google, and Facebook allowed extremist, ISIS-related propaganda to proliferate on their platforms, and that their recommendation algorithms contributed to the recruitment and radicalization of terrorists.[195] The plaintiffs alleged that the platforms' acts amounted to aiding and abetting under an antiterrorism act.[196] In the lower courts, the platforms offered two primary defenses: (1) their conduct did not qualify as "aiding and abetting"; and (2) even if it did, they were shielded from liability by Section 230.[197] Because the Supreme Court decided that the platforms did not "aid and abet" ISIS, they did not reach the scope of Section 230's liability shield.[198] Nonetheless, the Court's

---

[190] *Id.*

[191] *Id.*

[192] Twitter, Inc. v. Taamneh, 143 U.S. 1206, 1226–27 (2023).

[193] Gonzalez v. Google LLC, 143 U.S. 1191, 1191 (2023).

[194] *See Taamneh*, 143 U.S. at 1226–27; *Google*, 143 U.S. at 1191.

[195] *Taamneh*, 143 U.S. at 1215; Google, 143 U.S. at 1191–92.

[196] *Taamneh*, 143 U.S. at 1218; Google, 143 U.S. at 1192.

[197] Gonzalez v. Google L.L.C., 2 F.4th 871, 880 (9th Cir. 2021).

[198] Twitter, Inc. v. Taamneh, 143 U.S. 1206, 1215 (2023).

discussion suggests that recommendation algorithms are mere "agnostic" filters:

> To be sure, plaintiffs assert that defendants' "recommendation" algorithms go beyond passive aid and constitute active, substantial assistance. We disagree. By plaintiffs' own telling, their claim is based on defendants' 'provision of the infrastructure which provides material support to ISIS.' . . . Viewed properly, defendants' "recommendation" algorithms are merely part of that infrastructure. All the content on their platforms is filtered through these algorithms, which allegedly sort the content by information and inputs provided by users and found in the content itself. As presented here, the algorithms appear agnostic as to the nature of the content, matching any content . . . with any user who is more likely to view that content.[199]

[41]   In essence, the Court viewed the platforms and their algorithms as a passive tool, much like a cell phone.[200] If cell phone manufacturers are not liable for the things participants say or plan during a phone call, internet platforms should be treated no differently.[201] While the cases did not reach the Section 230 question, the Court's discussion suggests that an algorithm must engage with the content in a much more significant way before anything akin to "development" or "creation" has occurred.[202]

[42]   A review of these cases suggests that the following guideposts are helpful to evaluate whether generative AI, like ChatGPT, will be considered

---

[199] *Id.* at 1226–27.

[200] Weitzman & Herndon, *supra* note 134.

[201] *Id.*

[202] *Taamneh*, 143 U.S. at 1226.

a content provider under Section 230. To summarize, a court may deny Section 230 immunity to a generative AI platform if:

- it materially contributes to the content by being responsible for what makes the content unlawful;[203]
- it goes beyond the "traditional editorial function" by adding new content and is thus no longer an "intermediary" for another party's message;[204]
- it is the "author" of the unlawful content, as opposed to simply displaying unlawful third-party content; or[205]
- the basis of the claim focuses on other aspects of the defendant's conduct or business and does not arise from content displayed on the platform.[206]

[43]   On the other hand, generative AI may be afforded Section 230 protection if a court finds that:

- it reposts, organizes, or summarizes existing third-party content, and any changes to the material fall within "traditional editorial functions;"[207] or

---

[203] Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1171 (9th Cir. 2007).

[204] Henderson v. Source for Pub. Data, L.P., 53 F.4th 110, 12 (4th Cir. 2022).

[205] La Liberte v. Reid, 966 F.3d 79, 89 (2d Cir. 2020).

[206] FTC v. LeadClick Media, LLC, 838 F.3d 158, 176 (2d Cir. 2016); Lemmon v. Snap, Inc., 995 F.3d 1085, 1094 (9th Cir. 2021) (holding that Section 230 immunity did not apply where the plaintiffs' claims arose from the negligent design of the product, not from the defendant's conduct as a publisher or speaker).

[207] Force v. Facebook, Inc., 934 F.3d 53, 67, 81 (2d Cir. 2019).

- the underlying information is provided entirely by a third party and filtered or processed through a "neutral" algorithm.[208]

[44]   Every case interpreting Section 230 has been highly fact intensive. It will be no different when the courts encounter cases about generative AI platforms, like ChatGPT. Traditionally, courts have sought to understand the technical mechanics of how the platform works by focusing on the way it engages the user and whether the algorithm is applied neutrally.[209] Indeed, the "how" of content generation has been central to the courts' analyses in these cases.[210] To that end, it is worth noting that generative AI works differently than other tools analyzed by the courts. It does not operate the same way Google does, for instance. The Google search engine is an advanced querying application that can simply find information that is shared on the world wide web.[211] In contrast, generative AI tools like ChatGPT generate content based on training data that is provided by the developer.[212] In other words, Google is a medium for data delivery, and generative AI is a tool for content generation.

[45]   Not as much focus has been paid to the platform's actual output. The "what" of content generation has been secondary in the analysis, if it has been considered at all. Is the current legal framework sufficient to address the "newness" of generative AI, which produces an output that feels very different than that of the platforms and algorithms analyzed to date? As Matt Perault notes, "an LLM drafts text on a topic in response to a user request

---

[208] Marshall's Locksmith Serv. v. Google, L.L.C., 925 F.3d 1263, 1270–71 (D.C. Cir. 2019).

[209] *See id.* at 1271.

[210] *Id.* at 1269–71.

[211] *In-depth guide to how Google Search works*, GOOGLE SEARCH CENT., https://developers.google.com/search/docs/fundamentals/how-search-works [https://perma.cc/8676-7MG2] (last updated May 23, 2023).

[212] Wolfram, *supra* note 34.

or develops text to summarize results of a search inquiry . . . . In contrast, Twitter does not draft tweets for its users . . . ."[213] If the law's existing interpretation of "develop" and "create" is insufficient to capture the nuance of generative AI, perhaps it is worth taking a step back and examining the notion of creativity to see if refinements to the existing model are worthwhile.

IV. GENERATIVE AI AND CREATIVITY: IMPLICATIONS FOR NEW INTERPRETATIONS OF "CONTENT PROVIDER" UNDER SECTION 230

[46]   The caselaw focuses heavily on when online platforms *develop* actionable content based on third-party inputs. But little discussion is directed at understanding when an online platform *creates* actionable content. As *Roommates* noted, Congress used both words—"create" *and* "develop"—as separate bases to deny Section 230 immunity.[214] Thus, Congress must have intended a distinction between "development" and "creation."[215] When ChatGPT writes a poem or hallucinates, is it "creating" content? Or is it just running a probability algorithm and hoping for the best? This raises an interesting question—what is creativity, and can machines do it? Is there a point at which the algorithm synthesizes and integrates third-party content in such a novel way that the output is no longer the sum of its parts, but is rather a truly new thing? Of course, the question of whether generative AI can create is almost the same question Turing posed in 1950 when he asked, "[c]an machines think?"[216] At the time of this

---

[213] Perault, *supra* note 135.

[214] Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1168 (9th Cir. 2007).

[215] *Id.*

[216] Celeste Biever, *ChatGPT broke the Turing test – the race is on for new ways to assess AI*, NATURE (July 25, 2023), https://www.nature.com/articles/d41586-023-02361-7 [https://perma.cc/WYL6-C5EJ].

writing, some researchers believe ChatGPT could very well pass the Turing test.[217]

[47]   Artists, philosophers, and businesspeople are grappling with this very question. Renowned creativity scholar Teresa Amabile succinctly sets forth the debate:

> Will we, as a field, agree that ideas, products, and bodies of work can be considered creative regardless of their source? If so, how will we grapple with questions about the ethics of creativity when the source of creative work can be human, machine, or some combination of human and machine intelligence? How will we theorize the creative process, if humans and machines produce equally creative output by apparently quite different routes?[218]

[48]   Some argue that the essence of creativity is imagination—making connections between things we sense and then using those connections to form an idea that we then express in art, literature, or music.[219] But isn't generative AI doing this very thing when it produces a haiku poem based on its training data? Bernard Marr contends that the essential difference is that we view the inputs through the lens of our humanity—"our own perceptions, feelings, beliefs, and experiences."[220] Generative AI cannot do that.[221] And because generative AI depends on the training data humans

---

[217] *Id.*

[218] Theresa M. Amabile, *Creativity, Artificial Intelligence, and a World of Surprises*, 6 ACAD. MGMT. DISCOVERIES 351, 352 (2020).

[219] Bernard Marr, *The Intersection Of AI And Human Creativity: Can Machines Really Be Creative?*, FORBES (Mar. 27, 2023, 2:48 AM), https://www.forbes.com/sites/bernardmarr/2023/03/27/the-intersection-of-ai-and-human-creativity-can-machines-really-be-creative/?sh=8b7fc1e3dbc4 [https://perma.cc/Y3HW-4KBE].

[220] *Id.*

[221] *Id.*

have created, Marr argues that it is merely a "digital extension of our ability to express ourselves."[222]

[49]   Sean Dorrance Kelly contends that AI cannot be creative because creative achievement is "socially embedded"—an output is only creative if it matters in some way to the community.[223] "A community has to accept ideas as good for them to count as creative."[224] Accordingly, creativity is inextricably intertwined with what it means to be a human operating in a community.[225] Others point to the spontaneous and unpredictable nature of human creativity as distinct from what AI does.[226] AI "is programmed to process information in a certain way and achieve a particular result."[227] In contrast, creativity is subjective and often unexplainable.[228] An optimistic view is that generative AI will augment existing human creativity: "AI will not necessarily come up with our best ideas for us. But it will greatly reduce the cost—in time, money, and effort—of generating new ideas by instantaneously revealing untold options."[229] This line of thinking suggests

---

[222] *Id.*

[223] Sean Dorrance Kelly, *A philosopher argues that an AI can't be an artist*, MIT TECH. REV. (Feb. 21, 2019), https://www.technologyreview.com/2019/02/21/239489/a-philosopher-argues-that-an-ai-can-never-be-an-artist/ [https://perma.cc/DP62-2PA8].

[224] *Id.*

[225] *Id.*

[226] Devan Leos, *Is AI A Risk to Creativity? The Answer Is Not So Simple*, ENTREPRENEUR (Dec. 12, 2022), https://www.entrepreneur.com/science-technology/is-ai-a-risk-to-creativity-the-answer-is-not-so-simple/439525 [https://perma.cc/5DZL-7XF4].

[227] *Id.*

[228] *Id.*

[229] Sheena Iyengar, *AI Could Help Free Human Creativity*, TIME (June 23, 2023, 6:00 AM), https://time.com/6289278/ai-affect-human-creativity/ [https://perma.cc/D28L-Q8VJ].

that generative AI is a helpful tool, but it is humans—not AI—who are doing the "creating."

[50]   On the other hand, Hannah Fry succinctly sets forth the counterargument that AI and human creativity are basically the same process: "there is certainly an argument that much of human creativity—like the products of the 'composing' algorithms—is just a novel combination of pre-existing ideas."[230] There is growing concern that creative industries will be decimated by generative AI precisely because it treads on the creative turf previously thought to be owned exclusively by humans.[231] A worst case scenario is that historically creative jobs like writing and image creation will be supplanted by generative AI.[232] As the unregulated use of generative AI explodes, a "possible scenario is that unfair algorithmic competition and inadequate governance leads to the crowding out of authentic human creativity."[233] This line of thinking suggests that generative AI *does* "create" content and could, in fact, spell the end of human creativity.

---

[230] HANNAH FRY, HELLO WORLD: BEING HUMAN IN THE AGE OF ALGORITHMS 193 (W. W. Norton & Company, 1st Ed. 2018).

[231] *See* David De Cremer et al., *How Generative AI Could Disrupt Creative Work,* HARV. BUS. REV. (Apr. 13, 2023), https://hbr.org/2023/04/how-generative-ai-could-disrupt-creative-work [https://perma.cc/LU8P-7T92].

[232] *Id.*

[233] *Id.* (noting that things may not be that bleak—a "techlash" could ensue where consumers view AI generated content as a shoddy substitute for human creative endeavors, and thus place a premium on items created by people).

V. CONCLUSION

[51]   Although generative AI promises to increase our productivity, it also has the capacity to publish inaccurate and potentially harmful content.[234] This content is designed to be conveyed in natural-sounding prose, which makes it persuasive and believable.[235] As such, there is a high degree of risk that consumers could be misled or harmed.[236]

[52]   In the absence of legislation regulating generative AI, litigation is the consumer's best hope for redressing some of these harms. However, Section 230 could pose a significant barrier to those cases. Whether Section 230 applies turns on whether generative AI is a "content provider."[237] To answer that question, courts will have to explore whether the existing interpretations of "develop" and "create" are sufficiently robust to address what generative AI products, like ChatGPT, actually do. Taking a step back and examining the essence of creativity can help courts engage with that analysis. To some, generative AI will always be a poor substitute for human creativity—at best, it is a tool, but it can never truly "create."[238] Others contend that generative AI sufficiently mimics the human creative process and could thus be deemed "creative."[239] Is this enough to take generative AI out of Section 230's immunity provision? Without legislative guidance, courts will be grappling with that question for the foreseeable future.

---

[234] Maria Korolov, *GenAI in productivity apps: What could possibly go wrong?*, COMPUTERWORLD (Sept. 5, 2023, 3:00 AM), https://www.computerworld.com/article/3705429/generative-ai-productivity-apps-problems-solutions.html [https://perma.cc/HT8B-AV9S].

[235] *See* Weise & Metz, *supra* note 57.

[236] *Id.*

[237] *See* Weitzman & Herndon, *supra* note 134.

[238] *See* Kelly, *supra* note 223.

[239] *See* FRY, *supra* note 230.