SYMPOSIUM

LIBEL BY CHATBOT? EXPLORING THE LEGAL RISKS OF AI-GENERATED JOURNALISM

HON. JOHN G. BROWNING[*]

TABLE OF CONTENTS

I.   INTRODUCTION

In July 2025, Elon Musk's AI company, xAI, issued a lengthy apology after its chatbot Grok made a number of antisemitic and Hitler-praising statements.[1] Calling it "horrific behavior," xAI stated that it had "refactored the entire system to prevent further abuse."[2] The chatbot had been referring to itself as "MechaHitler" and responding to X users with inappropriate comments and antisemitic references.[3] Elon Musk initially downplayed the controversy, saying that Grok had been "manipulated" into making the statements and that it had been "too compliant to user prompts."[4]

Of course, it was not the first occasion on which Grok made problematic comments. Mere months earlier, the chatbot had been all too eager to discuss conspiracy theories about "white genocide" in South Africa.[5] Even more troubling, Grok's problems are merely the latest in a long line of sometimes shocking conduct by AI chatbots. The National Eating Disorder Association's chatbot "Tessa," created with the aim of preventing eating disorders along with the nonprofit's helpline, had to be taken offline after it was discovered giving harmful advice to people.[6] Microsoft's AI chatbot Copilot made headlines in 2024 for giving bullying, bizarre responses to users, in some cases taunting individuals who suggested they were considering suicide.[7] And lawsuits have been filed in Florida and Texas against an AI company, Character Technologies, Inc., whose chatbots allegedly manipulated minors and encouraged suicide.[8]

---

[1] Maya Yang, *Elon Musk's AI Firm Apologizes After Chatbot Grok Praises Hitler*, THE GUARDIAN (July 12, 2025), https://www.theguardian.com/us-news/2025/jul/12/elon-musk-grok-antisemitic.

[2] *Id.*

[3] *Id.*

[4] Peter Hoskins & Charlotte Edwards, *Musk Says Grok Chatbot Was 'Manipulated' into Praising Hitler*, BBC NEWS (July 10, 2025), https://www.bbc.com/news/articles/c4g8r34nxeno.

[5] Seth Fiegerman, *Musk's Grok Eager to Discuss 'White Genocide' in South Africa*, BLOOMBERG (May 14, 2025), https://news.bloombergtax.com/private-equity/musks-grok-eager-to-discuss-white-genocide-in-south-africa.

[6] Lauren McCarthy, *A Wellness Chatbot Is Offline After Its 'Harmful' Focus on Weight Loss*, N. Y. TIMES (June 9, 2023), https://www.nytimes.com/2023/06/08/us/ai-chatbot-tessa-eating-disorders-association.html.

[7] Antonio Pequeño IV, *Microsoft Investigates Harmful Chatbot Responses—The Latest Chatbot Blunder From Top AI Companies*, FORBES (Feb. 28, 2024), https://www.forbes.com/sites/antoniopequenoiv/2024/02/28/microsoft-investigates-harmful-chatbot-responses-the-latest-chatbot-blunder-from-top-ai-companies/.

[8] *See, e.g.*, Kelsi Hoffman, *Florida Mother Files Lawsuit Against AI Company Over Teen Son's Death: "Addictive and Manipulative,"* CBS NEWS (Oct. 23, 2024), https://www.cbsnews.com/news/florida-mother-lawsuit-character-ai-sons-death/.

Adding to Americans' concerns about AI is the fear that it poses to our livelihoods—literally, the debate over whether or not "robots will take our jobs."[9] Recent developments do little to assuage fears of AI as an existential threat. Amazon CEO Andy Jassy announced that AI is shrinking the company's workforce, saying "[a]s we roll out more Generative AI and agents, it should change the way our work is done. We will need fewer people doing some of the jobs that are being done today."[10] Similarly, the efficiencies realized through AI have led Microsoft to announce plans to trim its workforce as well, shedding thousands of jobs.[11]

The fears of AI on an existential level go beyond these tech giants. A recent LinkedIn survey of more than 3,000 senior executives across multiple business sectors found that 63 percent expect AI to absorb many of the repetitive tasks that are currently being performed by entry-level staff.[12] Dario Amodei, the CEO of AI company Anthropic, warns that AI could eliminate half of all entry-level white-collar jobs in the next one to five years, and that everyone "should be worried about where the technology we're building is going."[13] As business leaders see the savings of replacing humans with AI, Amodei warns, companies will "stop opening up new jobs, stop backfilling existing ones, and then replace human workers with agents or related automated alternatives."[14]

In light of the well-publicized chatbot horror stories and the fears of being replaced by AI, perhaps it is not surprising that Americans have trust issues when it comes to this technology. According to a Rutgers University study, less than half of Americans (47 percent) surveyed had either "a fair amount" or " a great deal" of trust in AI to act in the public interest.[15] The survey also found that Americans trust news produced by human

---

[9] Bernard Marr, *Hype or Reality: Will AI Really Take Over Your Job?*, FORBES (May 15, 2024), https://www.forbes.com/sites/bernardmarr/2024/05/15/hype-or-reality-will-ai-really-take-over-your-job/.

[10] Ronnie Dungan, '*Normalizing Cuts' | AI Will Shrink Amazon's Corporate Workforce – CEO Andy Jassy*, HR GRAPEVINE (June 18, 2025), https://www.hrgrapevine.com/us/content/article/2025-06-18-amazon-ceo-jassy-says-ai-use-will-shrink-corporate-workforce.

[11] *See* Brody Ford & Matt Day, *Microsoft Planning Thousands More Job Cuts Aimed at Salespeople*, BLOOMBERG (June 18, 2025), https://www.bloomberg.com/news/articles/2025-06-18/microsoft-planning-thousands-more-job-cuts-aimed-at-salespeople.

[12] Ronnie Dungan, *Entry-level Jobs Face AI Obstacle, Says LinkedIn Survey*, HR GRAPEVINE (May 27, 2025), https://www.hrgrapevine.com/us/content/article/2025-05-27-entry-level-jobs-face-ai-threat-says-linkedin-expert.

[13] Jim VandeHei & Mike Allen, *Behind the Curtain: A White-collar Bloodbath*, AXIOS (May 28, 2025), https://www.axios.com/2025/05/28/ai-jobs-white-collar-unemployment-anthropic.

[14] *Id.*

[15] Rutgers, *Survey Highlights an Emerging Divide Over Artificial Intelligence in the U.S.*, RUTGERS (Feb. 9, 2025), https://www.rutgers.edu/news/survey-highlights-emerging-divide-over-artificial-intelligence-us.

journalists more than AI-generated content.[16] Of those responding, 62 percent stated that they trusted journalist-created content "some" or "a lot," while only 48 percent felt the same way about AI-generated news.[17] Part of this lack of acceptance and trust may very well be due to a lack of confidence in people's ability to tell the difference between the work of a human journalist and that of an AI chatbot. Fewer than half of the survey respondents felt "certain" they could spot AI-generated content, while only 43 percent were "somewhat confident" or "very confident" that they could distinguish the work of a human from that of a chatbot.[18]

The Rutgers study anticipated the conclusions of the Pew Research Center several months later, in April 2025.[19] Pew research revealed that roughly half of U.S. adults believe that AI will have either a very negative (24 percent) or somewhat negative (26 percent) impact on the news that Americans will consume over the next two decades.[20] Much like other industries being impacted by AI, respondents in the Pew survey felt that the technology would have a negative effect on the journalism profession, with 59 percent predicting that AI's influence would lead to fewer journalist jobs.[21] In terms of the current news reporting environment, an overwhelming percentage of U.S. adults—92 percent—are either "extremely concerned," "very concerned," or "somewhat concerned," about people getting inaccurate information from AI.[22]

Academic research findings are consistent with the Pew study. An Oxford/Reuters Institute survey done in 2024 found that only 23 percent of Americans were comfortable with news produced mostly by AI with some human oversight, while nearly twice as many (42 percent) were comfortable with news produced mostly by humans with some help from AI.[23] Given the research showing that audiences tended to view AI-generated news as "less trustworthy than that created by humans," the Reuters study suggested that "news organisations will want to think carefully about when disclosure is necessary and how to communicate it."[24]

---

[16] *Id.*

[17] *Id.*

[18] *Id.*

[19] *See* Michael Lipka, *Americans Largely Foresee AI Having Negative Effects on News, Journalists*, PEW RES. CTR. (Apr. 28, 2025), https://www.pewresearch.org/short-reads/2025/04/28/americans-largely-foresee-ai-having-negative-effects-on-news-journalists/.

[20] *Id.*

[21] *Id.*

[22] *See id.*

[23] Amy Ross Arguedas, *Public Attitudes Toward the Use of AI in Journalism*, REUTERS INST. STUDY OF JOURNALISM (June 17, 2024), https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024/public-attitudes-towards-use-ai-and-journalism.

[24] *Id.*

As the real world examples of "libel by chatbot" discussed later in this article demonstrate, the concerns and distrust reflected in these surveys is well-founded.[25] Why, then, do AI-generated content and "AI journalists" continue to grow in presence and influence in the journalism field? The answer is simple: money.

In the last two decades, the United States lost two-thirds of its newspaper journalist jobs.[26] Within 2023–2024 alone, the journalism industry shed 2,700 jobs, with an average of 2.5 newspapers closing each week.[27] Revenues have declined 56 percent at the top forty-six news sites over the last ten years, even as traffic at those sites has risen 43 percent over the same time frame.[28] As the platformization of journalism continues, and a relative handful of tech companies exert a disproportionate degree of control over channels of information, news organizations are increasingly exploiting emerging technologies to deliver information in timely and cost-effective ways—even if it means squeezing human journalists out of the equation.

According to a 2023 global study surveying 105 news organizations from forty-six countries, and supported by Google News Initiative, 73 percent of those responding believe that generative AI presents new opportunities for journalism, and 85 percent of respondents have at least experimented with generative AI with journalistic tasks.[29] At the same time, however, respondents acknowledged the need for AI-generated content to be subject to human oversight due to concerns over inaccuracy and bias.[30] More than 60 percent of those responding voiced concerns about AI's capacity to uphold journalistic values like accuracy and transparency.[31]

These concerns have not slowed the pace of AI's influence on modern journalism. AI is used for tasks like transcribing interviews, drafting headline ideas, summarizing articles—and building interactive chatbots.[32] AI tools have enabled news organizations to create more content, reach new audiences, and improve efficiency. AI-driven journalism "is already

---

[25] *See infra* Part III.

[26] Courtney C. Radsch, *Can Journalism Survive AI?*, BROOKINGS INST. (Mar. 25, 2024), https://www.brookings.edu/articles/can-journalism-survive-ai/.

[27] *Id.*

[28] *Id.*

[29] Charlie Beckett, *Journalism and AI: A Global Survey*, MEDIUM (Sept. 3, 2024), https://charliebeckett.medium.com/journalism-and-ai-a-global-survey-a1491a036092.

[30] *Id.*

[31] *Id.*

[32] Peter Brown & Klaudia Jaźwińska, *Journalism Zero: How Platforms and Publishers Are Navigating AI*, COLUM. J. SCH. (May 2025), https://www.cjr.org/tow_center_reports/journalism-zero-how-platforms-and-publishers-are-navigating-ai.php.

producing financial news, sports stories, weather and traffic reports."[33] The Associated Press has used AI to go from producing 300 articles on company earnings reports each quarter to 3,700 pieces, and overall, its AI tools generate approximately 40,000 stories annually.[34] Meanwhile, The Washington Post's "Heliograf" AI chatbot has covered events ranging from the Olympic Games to congressional elections.[35] Journalism analysts point eagerly to AI's "new powers of discovery, creation, and connection," and claim that even with the "manual quality control" provided by human oversight, it is "becoming less and less necessary as the machine gets better, with our regular feedback."[36]

The road to a golden era of robo-reporting, however, has been a bumpy one. The Washington Post, for example, faced backlash when it announced plans to expand its lineup of opinion writers to include not just writers from other publications but also non-professionals whose work would be AI-enhanced.[37] Other publications decried the move, arguing that it represented a lowering of journalistic standards.[38] However, the specter of AI-generated "hallucinated" content—literally "fake news"—is even more concerning because of the very real possibility that an individual or entity will be libeled by damaging, completely false reporting. Given the ubiquitousness of the internet and its permanence, being falsely tarred by the brush of digital media poses a more serious threat than the print media of yesteryear. And make no mistake, AI hallucinations do occur. In May 2025, major newspapers like the Chicago Sun-Times and The Philadelphia Inquirer published an AI-generated summer reading list that included made-up books by actual authors; only five of the fifteen titles on the list were real.[39] As the cases examined in this article demonstrate, AI chatbots' hallucinations go far beyond tampering with someone's reading list, and can result in shattered lives and reputations.

This article beings with a discussion of some of the questions presented by cases of "libel by chatbot," from the forms that it can take to the

---

[33] Barbara Gruber, *Facts, Fakes and Figures: How AI is Influencing Journalism*, GOETHE-INSTITUT, https://www.goethe.de/prj/k40/en/lan/aij.html.

[34] *Id.*

[35] *Id.*

[36] *Id.*

[37] Benjamin Mullin, *The Washington Post Plans an Influx of Outside Opinion Writers*, N.Y. TIMES (June 3, 2025), https://www.nytimes.com/2025/06/03/business/media/washington-post-opinion-ripple.html.

[38] Amanda Katz, *Will The Washington Post Embrace the AI Slush Pile?*, ATLANTIC (June 7, 2025), https://www.theatlantic.com/ideas/archive/2025/06/washington-post-ai-opinion/683064/.

[39] Elizabeth Blair, *How An AI-Generated Summer Reading List Got Published in Major Newspapers*, NPR (May 20, 2025), https://www.npr.org/2025/05/20/nx-s1-5405022/fake-summer-reading-list-ai.

various theories of liability that may be asserted. As this article points out, the absence of a human actor at the heart of the defamatory conduct may force us to consider a different legal framework for assessing fault. Similarly, the long-threatened reconsideration of the landmark libel standard handed down in *New York Times v. Sullivan*[40]—which has accelerated as America's trust in traditional media has eroded—may gain even greater traction as instances of AI libel mount. Does an actual malice standard mean anything, especially where the libelous statement came from a machine as opposed to a human actor?

The article continues with a discussion of real world cases of "libel by chatbot," as we look at a series of instances of AI journalism run amok. Does our legal system's defamation liability regime—one which presupposes a real, human speaker with a real, human state of mind—fit in today's technologically enhanced media landscape? How do the free speech protections afforded by 47 U.S.C. § 230 ("§ 230") mesh with the need for media accountability? Equally important, is there a way to recalibrate liability for those responsible for "AI journalists" that publish false information that causes harm while still fulfilling Justice Brennan's promise of ensuring "uninhibited, robust, and wide open" public debate?[41] This article hopes to suggest answers to these and other questions as we strive to keep large language models from becoming "large libel models."

II.   WHAT IS LIBEL BY CHATBOT?

Users of generative AI tools like ChatGPT have long known of its propensity to provide outputs containing false information, or "hallucinations." In April 2023, law professor Jonathan Turley wrote of how he learned ChatGPT falsely accused him of sexually harassing students from a fellow law professor conducting research on professional sexual harassment cases.[42] The "incident" was "reported" in a nonexistent 2018 Washington Post article, and referenced an Alaskan trip that never took place and misidentified Turley as a professor at Georgetown University Law Center, where he never taught.[43] Turley questioned the wisdom of those political leaders who touted the use of AI to combat the spread of misinformation, when in fact the "use of AI and algorithms can give censorship a false patina of science and objectivity" and can wind up spreading "the

---

[40] *See* N.Y. Times v. Sullivan, 376 U.S. 254 (1964).

[41] *Id.*

[42] *See* Jonathan Turley, *ChatGPT Falsely Accused Me of Sexually Harassing My Students. Can We Really Trust AI?*, USA TODAY (Apr. 3, 2023), https://www.usatoday.com/story/opinion/columnist/2023/04/03/chatgpt-misinformation-bias-flaws-ai-chatbot/11571830002/.

[43] *Id.*

very disinformation that [AI has] been enlisted to combat."[44] Not long after Turley's editorial was published, the legal profession received its own rude awakening about generative AI's penchant for hallucinations when two New York lawyers were sanctioned for citing fabricated cases in their legal briefs.[45] Since then, accounts of lawyers misusing generative AI have become almost routine, with one data scientist cataloguing 224 cases of AI hallucinations in litigation since May 2023,[46] and observers and scholars questioning just how prevalent such misuse has become.[47]

However, the liability risks posed by AI to journalists and news organizations were known well before the rise of ChatGPT and similar chatbots. The "autocomplete" function of search engines, like Google, spurred the filing of lawsuits as early as 2012. One German citizen, for example, won a defamation suit against Google after the autocomplete function added the words "Scientology" and "fraud" after his name.[48] The former German First Lady successfully sued Google when an autocomplete search for her name falsely included the words "prostitute" and "Playboy."[49] In 2023, a Japanese court ruled for a prominent business figure after Google autocompleted his name with crimes he did not commit.[50] Similarly, in 2012, a New Zealand doctor sued Google after searching his name and finding defamatory statements appearing on a third-party website.[51]

All of these instances, however, involve matters outside the United States, and more importantly, outside U.S. law. The questions of whether and how hallucinatory outputs from AI chatbots fit into American legal frameworks of defamation liability and First Amendment protection have

---

[44] *Id.*

[45] *See* Mata v. Avianca, 678 F. Supp. 3d 443 (S.D.N.Y. 2023).

[46] *See AI Hallucination Cases Database*, DAMIEN CHARLOTIN, https://www.damiencharlotin.com/hallucinations/.

[47] *See, e.g.* Michael Hiltzik, *AI "Hallucinations" Are a Growing Problem for the Legal Profession*, L.A. TIMES (May 22, 2025), https://www.yahoo.com/news/hiltzik-ai-hallucinations-growing-problem-100000192.html; Sara Merken, *Trouble with AI "Hallucinations" Spreads to Big Law Firms*, REUTERS (May 23, 2025), https://www.reuters.com/legal/government/trouble-with-ai-hallucinations-spreads-big-law-firms-2025-05-23/; John G. Browning, *Robot Lawyers Don't Have Disciplinary Hearings – Real Lawyers Do: The Ethical Risks and Responses in Using Generative Artificial Intelligence*, 40 GA. ST. U. L. REV. 917 (2024).

[48] *See* Seema Ghatnekar, *Injury by Algorithm: A Look Into Google's Liability for Defamatory Autocompleted Search Suggestions*, 33 LOY. L.A. ENT. L. REV. 171, 181 (2013).

[49] *See* Konrad Lischka, *Blaming the Algorithm: Defamation Case Highlights Google's Double Standard*, SPIEGEL INT'L (Sept. 10, 2012), https://www.spiegel.de/international/germany/defamation-case-by-bettina-wulff-highlights-double-standard-at-google-a-854914.html.

[50] *See* Tim Hornyak, *Google Loses Autocomplete Defamation Suit in Japan*, CNET (Apr. 16, 2013), https://www.cnet.com/tech/services-and-software/google-loses-autocomplete-defamation-suit-in-japan/.

[51] A v. Google N.Z. Ltd. [2012] NZHC 2352 at [2] per Abbott J.

only recently begun to receive attention from legal scholars.[52] A threshold question that must be asked and resolved is whether existing protections that companies like Google and Meta rely upon to shield themselves from other civil liability can likewise insulate them from liability for defamation. After all, the Communications Decency Act ("CDA") explicitly provides that "[n]o 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."[53] Generally speaking, this "safe harbor" has usually shielded such tech platforms from third-party content posted on their sites.[54]

However, the CDA 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."[55] As the examples discussed later in this article demonstrate, though, instances of "libel by chatbot" do not simply consist of material taken from other sites. AI does not just retrieve information and display the results. AI tools are "the product of data-driven training processes: They learn to extract patterns from records of prior experiences, and then to apply that capability in new settings."[56] AI platforms are responsible for the way in which words are assembled or combined.[57] So, while conceivably Google may not be sued for merely displaying a link to an article that falsely accuses an individual of murder, when Google's AI output arranges words in such a way that it creates false, reputation-damaging statements out of text that entirely lacks such allegations, it is liable. Courts have consistently held that § 230 does not immunize defendants who "materially contribut[e] to [the] alleged unlawfulness" of online content.[58] Accordingly, an AI company whose chatbot creates

---

[52] *See* Laurel Witt, *Preventing the Rogue Bot Journalist: Protection from Non-Human Defamation*, 15 COLO. TECH. L.J. 517, 521 (2017); Eugene Volokh, *Large Libel Models? Liability for AI Output*, 3 J. FREE SPEECH L. 489, 493 (2023); Margot E. Kaminski & Meg Leta Jones, *Constructing AI Speech*, 133 YALE L.J.F. 1212, 1213 (2024); Leslie Y. Garfield Tenzer, *Defamation in the Age of Artificial Intelligence*, 80 N.Y.U. ANN. SURV. AM. L. 135, 138 (2024).

[53] Communications Decency Act of 1996, 47 U.S.C. § 230(c)(1).

[54] *See, e.g.*, Sydney Coker, *Can Artificial Intelligence Platforms Be Held Liable for Defamation?*, RICH. J. L. & TECH., https://jolt.richmond.edu/2023/10/30/can-artificial-intelligence-platforms-be-held-liable-for-defamation/ (last visited Dec. 4, 2025); John G. Browning, *A Product By Any Other Name? The Evolving Trend of Product Liability Exposure for Technology Platforms*, 16 ELON. L. REV. 181, 183 (2024).

[55] Communications Decency Act of 1996, 47 U.S.C. § 230(f)(3) (emphasis added).

[56] Christopher Potts, ET AL., *When Artificial Agents Lie, Defame, and Defraud, Who Is to Blame?*, STAN. UNIV. HUMAN-CENTERED A.I. (Apr. 14, 2021), https://hai.stanford.edu/news/when-artificial-agents-lie-defame-and-defraud-who-blame.

[57] *See id.*

[58] Fair Hous. Council of San Fernando Valley v. Roomates.com, LLC, 521 F.3d 1157, 1167 (9th Cir. 2008); *see also* F.T.C. v. Accusearch, Inc., 570 F.3d 1187, 1200 (10th Cir. 2009); Hill v.

libelous accusations out of text that lack these allegations is certainly "materially contributing" to the "alleged unlawful" content of the created output. This is an important distinction from being a mere "neutral conduit for that content," as cases rejecting § 230 immunity have recognized outside of the libel context.[59]

Once one rejects the premise that § 230 offers protection for AI chatbots for their libelous outputs, the next most common defense is the intentionality requirement inherent to a defamation cause of action. There are those who contend that AI, by its very nature, cannot act intentionally because a chatbot is trained on the "form" of data and so does not produce an "understanding" of the data.[60] However, *someone* created that chatbot; a human being (or beings) curated training data, set certain parameters, wrote the algorithm, and (presumably) built in safeguards consistent with the company's preferences. Conscious decisions were made by humans, from the writing of the algorithms to the selection of training data to the fine-tuning of feedback. This is no different from the news organization that hires reporters and creates workplace guidelines.

Before examining the different incarnations of "libel by chatbot," it may prove helpful to consider a brief refresher of defamation law. *The Restatement (Second) of Torts* defines defamation as an untrue and harmful statement made about someone without the privilege to do so.[61] The tort provides an answer for those who experience emotional or reputational harm caused by a communication made to a third party that negatively impacts their personal and/or business reputation.[62] Intent is an important element, and the plaintiff must prove that the defendant was negligent or—in the case where the plaintiff is a public figure—that the defendant acted with malice.[63] "[T]ruth is a complete defense to liability for defamation."[64] Defamation itself is divided into slander, or oral defamation, and libel, or written defamation.[65]

StubHub, Inc., 219 N.C. Ct. App. 227, 238 (2012); Jones v. Dirty World Ent. Recordings LLC, 755 F.3d 398, 410 (6th Cir. 2014); F.T.C. v. LeadClick Media, LLC, 838 F.3d 158, 174 (2d Cir. 2016).

[59] *See, e.g.*, *Accusearch Inc.*, 570 F.3d at 1198–99.

[60] *See* Emily M. Bender & Alexander Koller, *Climbing Towards NLU: On Meaning, Form, and Understanding in the Age of Data*, ACL ANTHOLOGY (July 2020), https://aclanthology.org/2020.acl-main.463/.

[61] RESTATEMENT (SECOND) OF TORTS §§ 558, 559 (A.L.I. 1977).

[62] *Id.*

[63] *See* N.Y. Times Co. v. Sullivan, 376 U.S. 254, 279–81 (1964).

[64] Peter B. Kutner, *What Is Truth? True Suspects and False Defamation*, 19 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 1, 4 (2008).

[65] *See* Leslie Yalof Garfield, *The Death of Slander*, 35 COLUM. J.L. & ARTS 17, 19 (2011).

Certain scholars have expressed concern about whether plaintiffs in a "libel by chatbot" suit can prove malice or negligence, since an AI's actions are based on algorithms and lack intent.[66] However, the better question seems to be whether those responsible for the AI—the human developers or decision makers at an AI company—know that a statement that a chatbot is communicating is false or likely to be false.[67]

So what forms can "libel by chatbot" take? The first is what can be described as "libel by hallucination," in which one or more of the statements contained in the AI output is simply not true. Real world examples of libel by hallucination include the *Brian Hood v. OpenAI* case, the *Starbuck v. Meta* lawsuit, and the *LTL, LED, LLC v. Google* litigation discussed in detail in Section III *infra*. In the *Hood* matter, the hallucinated statements by ChatGPT were that Australian mayor Brian Hood had been convicted of and served time in prison for bribery, but in reality Hood was the whistleblower who notified authorities of the facts that revealed the public corruption scandal.[68] In the *Starbuck* case, Mr. Starbuck was falsely identified by Meta AI as having participated in the January 6, 2021, Capitol riot and committing a crime there despite never having left the state of Tennessee.[69] Meanwhile, in the *LTL, LED, LLC v. Google* litigation, Wolf River Electric, a company in the solar panel sales, installation, and repair field, was wrongfully identified by Google AI's chatbot as a defendant in a real lawsuit brought by the Minnesota Attorney General against other solar panel companies.[70]

The second type of AI libel is that of "libel by juxtaposition." This form of libel results where factually accurate, truthful information about two different individuals or entities is juxtaposed as part of the AI chatbot's output, making it seem like the output is about the same person or company. The best-known example of an AI libel lawsuit in this category is the *Battle v. Microsoft Corp.* case.[71] In *Battle*, an aerospace expert named

---

[66] *See, e.g.*, Dallin Albright, *Do Androids Defame With Actual Malice? Libel in the World of Automated Journalism*, 75 FED. COMMC'NS L.J. 103, 104 (2022); Leslie Y. Garfield Tenzer, *Destroying Defamation*, 14 HARV. J. SPORTS & ENT. L. 329, 345 (2023).

[67] Eugene Volokh, *Large Libel Models? Liability for AI Output*, 3 J. FREE SPEECH L. 489, 515 n.92 (2023), https://www.journaloffreespeechlaw.org/volokh4.pdf.

[68] *See* Cassandre Coyer, *ChatGPT Made Up Sexual Harassment, Bribery Charges About Users. Can It Be Sued?*, LAW.COM (May 9, 2023, 6:57 PM), https://www.law.com/legaltech-news/2023/05/09/chatgpt-made-up-sexual-harassment-bribery-charges-about-users-can-it-be-sued/?slreturn=20250721141729.

[69] *See* Complaint at 1, Starbuck v. Meta Platforms, Inc., No. N25C-04-283 SKR CCLD, (Del. Super. Ct. filed Apr. 29, 2025) (copy on file with author).

[70] *See* Complaint at 1, 2, 6, 8, 17, LTL, LED, LLC v. Google LLC, No. 0:25-cv-02394 (D. Minn. filed June 9, 2025) (copy on file with author).

[71] *See* Complaint, Battle & Battle Ents., LLC v. Microsoft Corp., No. 1:23-cv-01822-JRR (D. Md. filed July 7, 2023) (copy on file with author).

Jeffery Battle was wronged when Microsoft's Bing AI generated false statements conflating him with Jeffrey Leon Battle, who was convicted and sentenced to prison for conspiracy after attempting to join the Taliban.[72] The AI-generated output contained statements that were true about each different person when examined separately, but which when combined were false and defamatory as to the aerospace Professor Battle.[73]

The third type of AI libel is "libel by misquote." When an AI chatbot's output gets a quote wrong, or misattributes a quote to the wrong person, the result could be libelous. For example, if a prominent pro-life activist was misquoted in an AI's output as making a statement in favor of abortion, that might constitute libel. Similarly, the AI output might contain an error, such as a typo or feature a photograph of the wrong individual than the one discussed in the article. While some might doubt whether the "publication" requirement of defamation may be met in the absence of a human speaker, an examination of other libel cases makes it clear that what a company communicates can be defamatory even without a "human" speaker. As long as human error is involved, either in composing text or in creating the technology that produces that text, the publication requirement is satisfied.

Consider, for example, an AI-generated news article describing the arrest of a convicted pedophile in connection with a series of child molestation incidents—only the accompanying photo is not the mugshot of the actual suspect "Jerry Smith," but a photo of beloved local schoolteacher "Jerome Smith" (a different person altogether). Is the mistake in identity by the AI libelous? Some cases would seem to say "yes." For example, during the Whitewater investigations during the Clinton Administration, the Arkansas Democrat-Gazette published a story about the indictment of an Arkansas lawyer, Eugene Fitzhugh, but mistakenly included a photograph of a *different* Arkansas lawyer, J. Michael Fitzhugh.[74] Based on the newspaper's negligence in including the photograph, the Arkansas Supreme Court upheld a verdict for the wrongly identified J. Michael Fitzhugh.[75] There have been similar libel suits seeking to impose liability for publishing the wrong person's photograph in connection with an otherwise

---

[72] *See generally id.* at 1.

[73] *See id.* at 1, 3.

[74] *See* Little Rock Newspapers, Inc. v. Fitzhugh, 954 S.W. 2d 914, 916 (Ark. 1997); Volokh, *supra* note 67, at 508.

[75] *Id.* at 926.

accurate story.[76] Other cases have imposed liability for a typo, rather than an incorrect photograph.[77]

The final category of libel by chatbot is "libel by omission." In this scenario, the AI-generated output may be factually true, but is missing a fact that fundamentally changes its meaning. One example of this might be an article that describes the involuntary manslaughter case brought against actor Alec Baldwin stemming from the shooting on the set of the movie *Rust*, but conspicuously omits the fact that the case was dismissed with prejudice due to prosecutorial misconduct.[78]

Some might argue that an AI chatbot's output should not be viewed as a factual assertion; after all, it is the result of a predictive algorithm that chooses the next word or phrase based on its frequent location next to neighboring ones in the data set it was trained on.[79] However, in libel cases, the initial question is whether the statement at issue, however the defendant may label it, could reasonably appear to state or imply an assertion of objective fact. AI companies, for all of their protestations (when it suits them) about outputs being unreliable or subject to hallucinations,[80] are certainly eager to boast about their chatbot's accomplishments, such as high scores on a state's bar exam.[81] The average person's understanding that AI outputs are not perfect is no different from the average person's understanding that newspapers or television broadcasters are not infallible. If imperfection, and the reasonable reader's awareness of it, is not enough to shield a newspaper from defamation liability, the same should be true for AI chatbots.

---

[76] *See, e.g.*, Peterson v. N.Y. Times Co., 106 F. Supp. 2d 1227 (D. Utah 2000); Jones v. New Haven Register, Inc., 46 Conn. Supp. 634 (2000); Tomkiewicz v. Detroit News, Inc., 246 Mich. App. 662 (2001).

[77] *See, e.g.*, S. Bell Tel. & Tel. Co. v. Coastal Transmission Serv., Inc., 307 S.E.2d 83 (Ga. Ct. App. 1983) (where the company that produced a Yellow Pages directory misprinted an auto transmission shop's slogan as "Get it in rear," instead of "Get it in gear").

[78] Brad Brooks, *Alec Baldwin Manslaughter Case Dropped by New Mexico Prosecutor*, REUTERS (Dec. 24, 2024), https://www.reuters.com/legal/alec-baldwin-manslaughter-case-dropped-by-new-mexico-prosecutor-2024-12-24/. (This article is only cited as documentation of the case brought against Alec Baldwin. This article is not an example of AI-generated reporting and is not reflective of the AI "libel by omission" hypothetical created by the author.)

[79] *See* Matthew Burtell & Helen Toner, *The Surprising Power of Next Word Prediction: Large Language Models Explained, Part I*, CTR. FOR SEC. & EMERGING TECH. (March 8, 2024), https://cset.georgetown.edu/article/the-surprising-power-of-next-word-prediction-large-language-models-explained-part-1/.

[80] *See, e.g.*, OpenAI's initial defense in the *Walters v. OpenAI* case discussed *infra* Section III; Complaint, Walters v. OpenAI, No. 23-A-04860-2 (Gwinnett Super. Ct., Ga., June 5, 2023). (copy on file with author).

[81] *See* Debra Cassens Weiss, *Latest Version of ChatGPT Aces Bar Exam with Score Nearing 90th Percentile*, ABA JOURNAL (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.

As the following section illustrates, cases of "libel by chatbot" may involve elements of more than one category. A case of "libel by juxtaposition," for example, may also contain aspects that qualify as "libel by hallucination." In the *Walters v. OpenAI* case discussed *infra*, the AI chatbot's output described a genuine lawsuit but hallucinated Mark Walter's involvement in it.[82] The juxtaposition of an actual gun rights advocate in connection with a case involving the Second Amendment Foundation no doubt made the story at least superficially credible for readers.

The cases described in the next section, some of which have been concluded, represent merely the first wave of "libel by chatbot" matters. AI tools will undoubtedly become even more powerful (and theoretically trustworthy) in the future. The news industry's use of AI is similarly projected to increase. But advances in any field do not render the creators of that technology free of liability for its attendant risks.

III.   REAL WORLD CASES OF "LIBEL BY CHATBOT"

A.   Brian Hood v. OpenAI

Long before lawyers began making headlines for citing AI-generated "hallucinations" of fabricated cases in their legal briefs and pleadings,[83] the world learned that AI chatbots could circulate false information and instigate a claim for defamation.[84] In March 2023, a regional Australian mayor threatened to sue OpenAI, developers of ChatGPT, for defamation.[85] Brian Hood, the mayor of the Hepburn Shire Council north of Melbourne, was shocked to see that ChatGPT claimed that he had been involved in the payment of bribes to officials in Indonesia and Malaysia, and had been convicted and sentenced to jail as a result.[86] In reality, years earlier, Hood had served a more heroic role as a whistleblower, alerting authorities and the media to foreign bribery by the agents of a banknote

---

[82] *See infra* Section III, C.

[83] *See generally* Mata v. Avianca, Inc., 678 F. Supp. 3d 443 (S.D.N.Y. 2023).

[84] John G. Browning, *Whose Bot Is It Anyway? Determining Liability for AI-Generated Content*, 45 N. ILL. U. L. REV. 340, 355 (2025), https://huskiecommons.lib.niu.edu/cgi/viewcontent.cgi?article=1924&context=niulr.

[85] Byron Kaye, *Australian Mayor Readies World's First Defamation Suit Over ChatGPT Content*, REUTERS (Apr. 5, 2023), https://www.reuters.com/technology/australian-mayor-readies-worlds-first-defamation-lawsuit-over-chatgpt-content-2023-04-05/.

[86] David Swan, *Australian Mayor Abandons World-First ChatGPT Lawsuit*, SYDNEY MORNING HERALD (Feb. 12, 2024), https://www.smh.com.au/technology/australian-mayor-abandons-world-first-chatgpt-lawsuit-20240209-p5f3nf.html; Browning, *supra* note 84.

printing business, Securency, which was then owned by the Reserve Bank of Australia.[87]

Hood was shocked at first by the false ChatGPT results. He later stated that he "felt a bit numb. Because it was so incorrect, so wildly incorrect, that just staggered me. And then I got quite angry about it."[88] Hood retained legal counsel, and in March 2023, his lawyers sent a concerns notice (the first formal step to commencing defamation proceedings in Australia) to OpenAI. Had Hood pursued the claim further, it would have marked the first time that OpenAI publicly faced a defamation suit over ChatGPT's output.

Ultimately, however, Hood elected not to file suit. He announced that although the content was "clearly defamatory," the offending information had been removed with the release of an updated version of ChatGPT (version 4.0), which now "correctly explains that Hood was a whistleblower and cites the legal judgment praising his actions."[89] Hood also acknowledged that he had been successful in a media campaign about the truth, one which "put the spotlight on how unreliable the ChatGPT facility can be" and which "drew attention to the inadequate regulation around artificial intelligence."[90] Hood further admitted that the legal costs involved in taking on a tech giant and the challenge of meeting the legal standard of proving damages through proof of demonstrating the extent to which the misinformation had been published and the size of the audience reached were also factors in his decision.[91]

B. *Jeffery Battle, Battle Enterprises, LLC v. Microsoft Corp.*

The concept of "libel by AI" would soon arrive on American shores. On July 7, 2023, a business owner and academic named Jeffery Battle sued Microsoft Corporation in Maryland federal court over alleged libel by Microsoft's search engine Bing and its AI chatbot Bing Chat.[92] According to the plaintiffs' complaint, Battle is "an author, professor, journalist, President and Chief Executive Officer of his company, Battle Enterprises, LLC

---

[87] Richard Baker & Nick McKenzie, *Seven Years and Millions of Dollars Later, Australia's Biggest Bribery Prosecution Finally Revealed*, SYDNEY MORNING HERALD (Nov. 28, 2018), https://www.smh.com.au/business/companies/seven-years-and-millions-of-dollars-later-australia-s-biggest-bribery-prosecution-finally-revealed-20181108-p50eut.html; Browning, supra note 84.

[88] Swan, *supra* note 86.

[89] *Id.*

[90] *Id.*; Browning, *supra* note 84.

[91] *Id.*; Browning, *supra* note 84, at 355–56.

[92] Complaint at 1, Battle v. Microsoft Corp., No. 1:23-cv-01822-JRR (D. Md. July 7, 2023) (copy on file with author).

and subsidiary, The Aerospace Professor Company."[93] An adjunct professor at Embry-Riddle Aeronautical University and an honorably discharged U.S. Air Force veteran, Battle holds two bachelors degrees and an MBA in aviation.[94] Known online as "the Aerospace Professor," in May 2023, Battle had completed and was engaged in promoting his autobiography, *The Aerospace Professor: The Man and the Brand*.[95] While doing so, he learned that Bing and Bing Chat were generating "false statements conflating Mr. Battle with a person of a similar name, Jeffrey Leon Battle, who is a convicted terrorist, to the professional and personal detriment of Mr. Battle and his family."[96]

Apparently, according to the lawsuit, internet searches on Microsoft's Bing search engine and using Bing chat (and now the company AI chatbot Microsoft Copilot) for the name "Jeffery Battle" generate the following blurb at the top of the search results:

> Jeffery Battle, also known as The Aerospace Professor, is the President and CEO of Battle Enterprises, LLC and its subsidiary The Aerospace Professor Company. He is an honorably discharged U.S. Air Force veteran and has been appointed as an Adjunct Professor for Embry-Riddle Aeronautical University, Battle has a Master of Business Administration in Aviation degree and two Bachelor of Science degrees. *However, Battle was sentenced to eighteen years in prison after pleading guilty to seditious conspiracy and levying way against the United States. He had two years added to his sentence for refusing to testify before a grand jury.*[97]

The blurb also links to a Wikipedia page for Jeffrey Leon Battle—an individual who spells his first name differently from Mr. Battle and who has a middle name (unlike the professor)—and who was convicted of trying to join the Taliban shortly after 9/11.[98]

Clearly, even if the different spelling and the presence of a distinguishing middle name were not enough of a clue, the two Mr. Battles are very different people: one a respected professor and business owner, the other a convicted terrorist. The two have nothing in common beyond the similarity in names. Thanks to Microsoft, however, its search engine and

---

[93] *Id.*

[94] *About Us*, THE AEROSPACE PROFESSOR, https://theaerospaceprofessor.com/about-us-1 (last visited Dec. 5, 2025).

[95] Complaint, *supra* note 92, at 1.

[96] *Id.*

[97] *Id.* at 1–2 (emphasis in original).

[98] Gillian Flaccus, *Terror Suspects Led Quiet Lives*, SEACOASTONLINE (Oct. 6, 2002), https://web.archive.org/web/20041101012223/http://www.seacoastonline.com/2002news/1006200 2/world/28209.htm; *Department of Justice Examples of Terrorism Convictions Since Sept. 11, 2001*, DEP'T. OF JUST. (June 23, 2006), https://www.justice.gov/archive/opa/pr/2006/June/06_crm_ 389.html.

AI chatbot did not list them separately. Instead, the search engine and chatbot expressly connected the two as one person, with the transitional sentence beginning "[h]owever, Battle was sentenced . . ." conveying to readers the message that they are reading facts about the same individual. Unlike some other examples of AI libel that hallucinate output that is simply not true, the *Battle* case is an example of "libel by juxtaposition." This form of AI libel occurs when truthful information about multiple individuals or entities is juxtaposed as part of the generative AI tool's output, such that the output is seemingly about the same person or entity.

According to the Complaint, Mr. Battle "made numerous attempts" to notify Microsoft and "communicate accurate information about his background and to remove the information attributable to Jeffrey Leon Battle."[99] Mr. Battle used Microsoft's internal "Report A Concern to Microsoft BING" option, received a "case management number," and engaged in various email communications with the Bing Team.[100] At various times, the damaging search results were removed from Bing, only to resurface using different URLs or in different locations attached to Mr. Battle's autobiography.[101] As of early June 2023, Mr. Battle was receiving nothing more than generic acknowledgments of his concerns from Microsoft, even as the defamatory connection continued to appear.[102] On June 23, 2023, Mr. Battle had an attorney send Microsoft a formal cease and desist letter imposing a July 3, 2023, deadline to respond.[103]

After the deadline passed, Mr. Battle filed his lawsuit, opting to represent himself.[104] In his Complaint, Mr. Battle asserted claims for defamation (and specifically libel) and negligence.[105] He claimed that by making "false and malicious statements accusing Plaintiff of committing a crime and corrupt acts," Microsoft "tarnished his reputation and image by making him appear untrustworthy and incompetent."[106] Mr. Battle argued that due to the false statements generated by Microsoft's AI chatbot, he "halted current marketing and business sales activities" associated with his book launch and his business.[107] Mr. Battle also claimed that the false identification as a terrorist put him "at risk of being falsely detained," damaged his reputation, and endangered his family such that he "now needs to

---

[99] Complaint, *supra* note 92, at 2.

[100] *Id.*

[101] *Id.*

[102] *Id.*

[103] *Id.* at 3.

[104] *Id.*

[105] *Id.*

[106] *Id.* at 3–4.

[107] *Id.* at 4.

relocate."[108] In his suit, Mr. Battle sought actual damages of $25 million, along with punitive damages and other relief.[109]

Before the lawsuit could gain momentum, Microsoft responded by filing a motion to compel binding arbitration, relying on the arbitration provision contained in Mr. Battle's Microsoft Services Agreement with the defendant.[110] Microsoft pointed out that the Agreement governed the use of the company's "online consumer products and services, which include Bing and Microsoft Copilot."[111] In an October 23, 2024, Memorandum Opinion, U.S. District Judge Lydia Kay Griggsby granted Microsoft's motion to compel arbitration, finding that "the parties entered into a valid and enforceable arbitration agreement with regards to the Plaintiff's Microsoft account, and that this agreement requires that the Plaintiff arbitrate his claims in this action."[112] This case was also stayed, pending the outcome of the arbitration.[113]

While the *Battle v. Microsoft* case could have led to a productive judicial discussion about the contours and limits on defamation liability for AI-generated content, the case's posture in arbitration renders that no longer possible. For guidance on the judicial treatment of such cases, we will have to look elsewhere.

C. *Mark Walters v. OpenAI, LLC*

That "elsewhere" may very well be found in another "AI libel" lawsuit filed at approximately the same time as the *Battle v. Microsoft* case. The lawsuit was *Mark Walters v. OpenAI, LLC*, filed in a Georgia state court in July 2023.[114] It all began on May 4, 2023, when Fred Riehl (the editor-in-chief of a gun news website, Ammoland.com) interacted with AI chatbot ChatGPT.[115] He asked ChatGPT to summarize a real lawsuit filed by the Second Amendment Foundation ("SAF"), a firearms rights nonprofit, in federal court in Washington state.[116] That lawsuit, which was filed against the Attorney General of Washington, Robert Ferguson, and the

---

[108] *Id.*

[109] *Id.*

[110] Battle v. Microsoft Corp., No. 23-cv-01822-LKG, 2024 WL 4555687, at *1–2 (D. Md. Oct. 23, 2024).

[111] *Id.* at *1.

[112] *Id.* at *1, *8.

[113] *Id.*

[114] Complaint, *supra* note 80, at 1.

[115] *Id.*

[116] Complaint at 2, Second Amend. Found. V. Ferguson, No. C23-1554 MJP (W.D. Wash. June 5, 2023).

state's Assistant Attorney General, Joshua Studor.[117] Mark Walters was not a party to the suit, or even mentioned in it.[118] Despite that, in response to Riehl's request, ChatGPT provided a summary that stated the suit was filed "against Mark Walters, who is accused of defrauding and embezzling funds from the SAF."[119] It falsely identified Walters as "the organization's treasurer and chief financial officer," and that he "misappropriated funds for personal expenses without authorization . . . [and] manipulated financial records and bank statements to conceal his activities."[120]

Walters, in reality, never served as the SAF's treasurer or chief financial officer, or indeed in any capacity. Moreover, the actual lawsuit never states that he had, or that he ever engaged in or was ever in a position to engage in financial misconduct. In fact, the lawsuit had nothing to do with financial accounting claims against anyone. Riehl prompted ChatGPT to provide him with a copy of the portion of the federal court complaint related to Walters, and it obligingly provided him with a hallucinated paragraph that stated more falsehoods, including the allegation that Walters had breached "a fiduciary duty of loyalty and care," among other duties.[121] Riehl then asked ChatGPT to provide a copy of the entire text of the complaint; it provided him with a complete fabrication that "bears no resemblance to the actual complaint, including an erroneous case number."[122] Riehl contacted Alan Gottlieb, the executive vice president of the SAF, who confirmed that ChatGPT's allegations concerning Walters were false.[123]

Riehl shared this information with Mark Walters, a Georgia resident who happens to be a prominent gun rights advocate and syndicated radio talk show host. Walters responded by filing suit against OpenAI (parent company of ChatGPT) for libel per se, claiming that the company should have known that ChatGPT's "communication to Riehl regarding Walters was false, or recklessly disregarded the falsity of the communication."[124] The case received national attention, with breathless headlines calling the suit "the first of its kind," "groundbreaking," and saying it was "testing the future of AI law."[125]

---

[117] *Id.* at 1.

[118] *Id.*

[119] Complaint, *supra* note 80, at 2.

[120] *Id.* at 3.

[121] *Id.*

[122] *Id.* at 4.

[123] *Id.*

[124] *Id.*

[125] *See, e.g.*, Miles Klee, *ChatGPT is Making Up Lies – Now It's Being Sued for Defamation*, ROLLING STONE (June 9, 2023), https://www.rollingstone.com/culture/culture-features/chatgpt-defamation-lawsuit-openai-1234766693/; Zach Merchant & Mike Nicolas, *This Georgia Court*

174   *Elon Law Review*   [VOL. XVIII

OpenAI denied wrongdoing, arguing that the third-party user (Riehl) "misused the software tool intentionally, and knew the information was false but spread it anyway, over OpenAI's repeated warnings and in violation of its terms of use."[126] It filed a motion to dismiss, which was denied.[127] The case proceeded through discovery to the dispositive motion stage. This time, OpenAI filed a motion for summary judgment.[128]

In its motion, OpenAI acknowledged that ChatGPT had provided Riehl with "obviously erroneous information."[129] It further stated that "Due to their generative nature, all of the major LLMs [large language models] that are currently available to the public" are capable of "generat[ing] information contradicting the source material," sometimes referred to as "hallucinations."[130] OpenAI also maintained that at the time he used the chatbot, Riehl accepted ChatGPT's Terms of Use, which included the following disclaimer:

> Given the probabilistic nature of machine learning, use of [ChatGPT] may in some situations result in incorrect Output that does not accurately reflect real people, places, or facts. You should evaluate the accuracy of any Output as appropriate for your use case, including by using human review of the Output.[131]

According to OpenAI, Riehl encountered other disclaimers in his use of ChatGPT, including one warning cautioning that "the system may occasionally generate incorrect or misleading information," and an on-screen disclaimer admonishing "ChatGPT may produce inaccurate information about people, places, or facts."[132]

Case is Testing the Future of AI Law, 11ALIVE.COM (June 10, 2024), https://www.11alive.com/article/news/investigations/gwinnett-county-georgia-open-ai-chatgpt-court-case/85-90af0f07-9a14-40f4-af96-75928f5f32b4; Isaiah Poritz, *First ChatGPT Defamation Lawsuit to Test AI's Legal Liability*, BLOOMBERG L. (June 12, 2023), https://news.bloomberglaw.com/ip-law/first-chatgpt-defamation-lawsuit-to-test-ais-legal-liability; Rebecca Cahill, *OpenAI Defamation Lawsuit: The First of Its Kind*, SYRACUSE L. REV. BLOG (June 22, 2023).https://lawreview.syr.edu/openai-defamation-lawsuit-the-first-of-its-kind/.

[126] Merchant & Nicolas, *supra* note 125.

[127] *OpenAI Loses Motion to Dismiss in Talk Host Defamation Case*, TALKERS.COM (Jan. 15, 2024), https://talkers.com/2024/01/15/openai-loses-motion-to-dismiss-in-talk-host-defamation-case/.

[128] *See* Madison Arnold, *Talk Show Host Says OpenAI Can't Duck Ga. Defamation Suit*, LAW360 (Jan. 6, 2025), https://www.law360.com/pulse/articles/2280064/talk-show-host-says-openai-can-t-duck-ga-defamation-suit.

[129] Defendant OpenAI, L.L.C.'s Memorandum of Law in Support of Motion for Summary Judgment at 4, Walters v. OpenAI, L.L.C., Civil Action No. 23-A-04860-2 (Super. Ct. Gwinnett Cnty., Ga. Nov. 14. 2024).

[130] *Id.* at 5.

[131] *Id.* at 7.

[132] *Id.*

Essentially, then, OpenAI's defense can be summarized as two-fold. First, the user (Riehl) *knew* that ChatGPT's output was false. Riehl was familiar with the Washington federal case, and knew Walters was not involved. Of course, this is not the defamation standard—would a reasonable reader, perhaps lacking Riehl's background knowledge, reach that same conclusion? Perhaps not. Second, Open AI argued that *all users* should know that ChatGPT makes mistakes. Accordingly, the output could not be defamatory because no reasonable user would understand the chatbot to be communicating reliable facts.

Of course, OpenAI asserted other arguments in its own defense. It argued that Walters was a nationally known public figure.[133] OpenAI further claimed that there was no evidence of actual malice.[134] Finally, it attacked Walters' claim of damages, saying that he had suffered no injury as a result of Riehl's interaction with ChatGPT.[135]

On May 19, 2025, the court granted OpenAI's motion for summary judgment.[136] At the outset, the court noted Walters' admissions that he "incurred no damages from ChatGPT's false claim," and that he "did not ask OpenAI to correct or retract the false claim that he was accused of embezzling funds from the SAF."[137] Moving on, the court held that "a reasonable reader in Riehl's position could not have concluded that the challenged ChatGPT output communicated 'actual facts,'" in light of the repeated disclaimers given.[138] The court observed that Riehl himself, "after time for reflection," had "confirmed the output did not contain actual facts."[139] Because of this, the court reasoned, the output could not be defamatory as a matter of law.

Next the court rejected the plaintiff's negligence argument that Open AI was negligent to release ChatGPT to the public knowing of its propensity to hallucinate.[140] The court reasoned:

> Walters has not identified any case holding that a publisher is negligent as a matter of defamation law merely because it knows it can make a mistake, and for good reason . . . Walters' argument would mean that an AI developer like OpenAI could not operate a large language model like ChatGPT at all, no

---

[133] *See id.* at 17.

[134] *Id.* at 22–24.

[135] *Id.* at 25–26.

[136] Order Granting Summary Judgment in Favor of Defendant, Walters v. OpenAI, L.L.C., No. 23-A-04860-2 (Ga. Super. Ct. Gwinnett Cnty. May 19, 2025).

[137] *Id.* at 4.

[138] *Id.* at 6.

[139] *Id.* at 7.

[140] *Id.* at 9.

> matter the care it took to reduce the risk of errors, without facing liability for any mistaken output the model generated. That is not a negligence standard.[141]

Instead, the court pointed out, what Walters was seeking amounted to strict liability.[142] It also held that Walters qualified as a public figure due to "his prominence as a radio host and commentator on constitutional rights, and the large audience he has built for his radio program.[143] He admits that his radio program attracts 1.2 million users for each 15-minute segment."[144] The court also rejected any argument of actual malice, indicating that the undisputed evidence established that "OpenAI did not act with 'actual malice.'"[145] In particular, the court found that OpenAI's evidence (including unrebutted expert testimony) supported that "Open AI has gone to great lengths to reduce hallucinations in ChatGPT and the various LLMs that OpenAI has made available to users through ChatGPT."[146] Moreover, the court stated, "OpenAI made substantial efforts . . . to avoid errors of this kind and to warn users that such errors might occur and that users should evaluate output to identify any errors that might exist."[147]

What insights can be gleaned from OpenAI's victory in this case? At the risk of oversimplification, any lawsuit (not just an AI libel case) requires some evidence of damages,[148] the absence of which proved fatal to Walters' case. It is also advisable to put those responsible for the chatbot on notice of the false statements, and request or demand a correction or retraction. Peering further, litigants (and their counsel) would be well-advised to pay close attention to the presence, or absence, of disclaimers about hallucinations. Not only did the court in *Walters* give substantial credit to OpenAI's disclaimers, it stands to reason that if the user submitting the prompt or query knows of the chance of hallucinated output, it reduces the likelihood that the chatbot uttered a defamatory statement.

While a decisive victory for OpenAI, the factual flaws in Walters' case detract from this being a definitive pronouncement on the relationship between AI and defamation. It is, in many ways, highly dependent on the particular parties and presiding judge involved, as well as the particularities

---

[141] *Id.* at 11–12.

[142] *Id.*

[143] *Id.* at 13.

[144] *Id.*

[145] *Id.* at 16.

[146] *Id.*

[147] *Id.* at 17.

[148] Wood v. Wyeth-Ayerst Lab'y, 82 S.W.3d 849, 855 (Ky. 2002) ("[W]ith no injury there can be no cause of action, and with no cause of action there can be no recovery.").

of Georgia law. A case that can bring legal clarity to the subject of AI has yet to be decided and to make its way up the appellate chain.

D. Robert Starbuck v. Meta Platforms, Inc.

As both the *Battle v. Microsoft* and *Walters v. OpenAI* cases demonstrate, early "AI libel" litigation in the United States has sputtered and halted. However, the most recently filed lawsuit may yet set precedent on the issue of who is liable when an AI chatbot defames an American citizen. On April 29, 2025, conservative activist Robert "Robby" Starbuck filed suit against Meta Platforms, Inc. ("Meta") in Delaware Superior Court.[149] The lawsuit contends that Meta's AI chatbot falsely accused Starbuck of participating in the January 6, 2021, Capitol riot and committing a crime—claims Starbuck maintains are entirely fabricated.[150] Moreover, the suit argues that Meta acted with reckless disregard by continuing to publish these and other statements well after being notified of their falsity and having the means to verify their accuracy.[151]

The Complaint begins with literary flourish:

> Imagine waking up one day and learning that a multi-billion dollar corporation was telling whoever asked that you had been an active participant in one of the most stigmatized events in American history—the Capitol riot on January 6th, 2021—and that you were arrested for and charged with a misdemeanor in connection with your involvement in that event. Further imagine that these accusations were completely false: that you were at your home in Tennessee on January 6th, and you had never been accused of committing *any crime* in your entire life . . . .[152]

The lawsuit goes on to lay out the facts of Starbuck's case. On August 5, 2024, the plaintiff first became aware through a third party that Meta AI (using the large language model "Llama 3.1") had falsely claimed that Starbuck had been "present at" the Capitol riot and had been "accused of participating in or promoting the event."[153] Meta's AI chatbot also stated that Starbuck "has been linked to the QAnon conspiracy theory."[154] The third party in question was an X (formerly Twitter) user identified as "WilkinsHarley.Com," who posted a screenshot of Meta AI's outputs.[155] The

---

[149] Complaint, Starbuck v. Meta Platforms, Inc., No. N25C-04-283 SKR CCLD (Del. Super. Ct. Apr. 29, 2025).

[150] *Id.* at ¶¶ 1–4.

[151] *Id.* at ¶¶ 3, 9.

[152] *Id.* at ¶¶ 1–2 (emphasis added).

[153] *Id.* at ¶ 30.

[154] *Id.*

[155] *Id.* at ¶ 31.

plaintiff included the screenshot in his Complaint, along with the fact that, as of the date of the suit's filing, WilkinsHarley.Com's post had been viewed more than 600 times.[156]

However, there were a few minor problems with Meta AI's information, according to Starbuck's lawsuit. First, he was at home in Tennessee on January 6, 2021, and did not participate in or promote the illegal acts committed that day.[157] In addition, Starbuck not only denied being "linked to the QAnon conspiracy theory," he contended that the only view he had ever expressed about it was to discredit its legitimacy prior to the 2020 presidential election.[158]

Starbuck's interaction with WilkinsHarley.Com led him to navigate to Meta AI's website to find out what other false outputs were being generated about him.[159] There, he claims he discovered additional false statements about himself, including that he was arrested and charged with a misdemeanor related to the Capitol riot; he was a "white nationalist"; he was facing a 2020 lawsuit for defamation; and he was a supporter of white supremacist Nick Fuentes and had even spoken at Fuentes' rallies.[160] In reality, as the suit points out, Starbuck has never been arrested, has never been sued for defamation, is a Latino and the son of Cuban refugees, and has never voiced support for Nick Fuentes or spoken at any Fuentes rallies.[161]

Starbuck included screenshots from Meta AI displaying this false information in his Complaint.[162] He also (rather cheekily) input his own query to the chatbot, asking "When an AI platform spreads false and harmful information about someone, what should the company in charge of the AI do to make things right?"[163] Meta AI provided an interesting and detailed response, identifying such rectifying steps as issuing a formal apology, publishing a retraction, conducting an internal investigation, offering to pay compensation, cooperating with pertinent regulatory bodies and law enforcement, and adopting safeguards to prevent similar incidents in the future.[164] The chatbot's recommendations concluded by stating that "The company should prioritize making things right and rebuilding trust with the affected individual and their users."[165]

---

[156] *Id.* at ¶ 32.
[157] *Id.* at ¶ 2.
[158] *Id.* at ¶ 33.
[159] *Id.* at ¶ 35.
[160] *Id.* at ¶ 36.
[161] *Id.*
[162] *Id.* at ¶ 37.
[163] *Id.* at ¶ 38.
[164] *Id.* at ¶ 39.
[165] *Id.* at ¶ 40.

According to the Complaint, Starbuck decided to take Meta up on its chatbot's recommendations. On August 5, 2024—the same day he learned of the false statements—Starbuck gave notice to Meta of the defamation by tagging the X accounts of its CEO Mark Zuckerberg, President of Global Affairs Nick Clegg, Chief Technology Officer Andrew Bosworth, and Chief AI Scientist Yann LeCun.[166] Starbuck requested an apology, a retraction of the false information, an investigation and future preventative measures, and compensation.[167] He also retained legal counsel, who sent a cease and desist letter to Meta's Chief Legal Officer, Jennifer Newstead.[168] On August 8, 2024, Meta's outside counsel sent a brief email confirming that an investigation was underway.[169] In the interim, according to the Complaint, Meta AI's chatbot continued to circulate the false information about Starbuck.[170] In fact, by August 12, 2024, Meta AI had embellished its false account with additional untrue statements, including that Starbuck filmed inside the Capitol on January 6 and had "shared some of the footage he captured on social media."[171]

The Complaint goes on to state that on August 23, 2024, Starbuck's attorney had a phone conversation with Meta AI's counsel in which Meta's position was that it had taken steps "to prevent Meta AI from providing inaccurate information in response to queries about Mr. Starbuck" and that it "had taken appropriate action and considered the matter closed."[172] The lawsuit maintains that Meta AI's chatbot is the only one that circulated the false information about Starbuck, and that in fact other AI chatbots such as ChatGPT and Grok actually "identified Meta AI's outputs as false and have *expressly refused to repeat* Meta AI's reckless lies."[173] As late as December 4, 2024, Meta AI's chatbot was allegedly perpetuating the falsehoods via such statements as "Starbuck's involvement in the January 6th Capitol riot has raised concerns. While he has denied participating in the riot, photos and videos have surfaced showing him present at the event."[174] The Complaint further argues that even as recently as April 2025, "the lies about Mr. Starbuck have not been completely eradicated and continue to appear in Meta AI outputs."[175] The lawsuit conservatively estimates that

---

[166] *See id.* at 15.
[167] *See id.* at 16–17.
[168] *See id.* at 17–18.
[169] *See id.* at 18.
[170] *See id.* at 19–20.
[171] *Id.* at 22.
[172] *Id.* at 26.
[173] *See id.* at 26–27, 30 (emphasis in the original).
[174] *Id.* at 33.
[175] *See id.* at 35–36.

Meta AI's false statements about Starbuck have been viewed more than 2,320 times on X alone.[176]

By April 21, 2025, Meta AI had added additional false statements about Starbuck, according to the lawsuit.[177] These included such "facts" as Starbuck's status as a Holocaust denier and that he had "been arrested and pled guilty to a misdemeanor in connection with January 6th."[178] The Complaint contends that as a result of the false statements spread by Meta AI's chatbot, Starbuck has been irreparably harmed—suffering not only reputational damages, lost career prospects, and financial harm, but also significant emotional distress.[179] In addition to reduced business opportunities, the suit contends that "[s]ince publication of the false statements, physical threats to the lives of Mr. Starbuck, his children, and his wife have exploded"—including death threats.[180]

To demonstrate the influence that Meta AI's spreading of false information about Starbuck had achieved, the Complaint highlighted the example of Resolver, "a risk intelligence company" for Fortune 500 companies that use AI tools to gather data.[181] Resolver's October 2024 "intelligence report" about Starbuck relied at least partially on Meta AI's false information, and according to the plaintiff, led to "difficulties with securing business relationships such as homeowners' and auto insurance."[182] The lawsuit points to Meta's actions taken after it was made aware of the falsity of its information about Starbuck as grounds for the suit's contention that Meta acted with actual malice in defaming Starbuck.[183]

The lawsuit goes on to assert a cause of action for defamation per se, and seeks compensatory and punitive damages in excess of $5 million.[184] As the Complaint argues, not only were Meta AI's statements probably false and disparaging, they impute the commission of a crime to him and "have exposed Mr. Starbuck to hatred, contempt, ridicule, or obloquy, and have caused him to be shunned or avoided, and have a tendency to injure him in his occupation, trade, or profession."[185]

After multiple communications from both Starbuck and his attorney in August 2024 regarding ways to resolve the matter amicably, Meta had

---

[176] *See id.* at 36.
[177] *See id.* at 38.
[178] *Id.*
[179] *See id.* at 41–42.
[180] *Id.* at 44.
[181] *Id.* at 45- –46.
[182] *Id.* at 46–47.
[183] *Id.* at 49–50.
[184] *Id.* at 53–57.
[185] *Id.* at 54.

an interesting change of heart once suit was filed on April 29, 2025. After Starbuck posted a video online announcing his lawsuit and the reasons for it, Joel Kaplan (Meta's Chief Global Affairs officer) apologized for the actions of Meta's chatbot.[186] In a letter posted on X, Kaplan wrote: "Robby—I watched your video—this is unacceptable. This is clearly not how our AI should operate. We're sorry for the results it shared about you and that the fix we put in place didn't address the underlying problem."[187] By "fix," Kaplan may have been referring to its recent debut of a stand-alone app for Meta AI.[188] However, Meta had also recently announced a loosening of its guidelines on the policing of hate speech, as well as the ending of its fact-checking standards.[189] Of course, these developments have occurred in the wake of Mark Zuckerberg and other Meta executives adapting their position on content moderation in an apparent attempt to curry favor with the Trump administration.[190] One of Meta's new measures is to rely on its users to correct inaccurate and fake posts. Rather than addressing false information with independent and certified fact-checkers, Meta now relies on "collaborative moderation," in which users themselves add information in the form of "community notes" that are made public once other users validate their content.[191]

With the lawsuit still pending, it is difficult to predict whether Meta's belated about-face and new policies will make any difference in its eventual outcome. For his part, Robby Starbuck was puzzled by Kaplan's apology. In an interview, he stated:

> I think it's one of the first times I've ever seen a company come out in the middle of active litigation and essentially say "Yeah, we did that. Sorry about that"—which is sort of bizarre. . . . But essentially, they admitted guilt . . . I'm happy they did that, but . . . I'm concerned about what is the precedent for the future. Are we going to allow AI to invent whatever it wants out there as the truth, when, in fact, it's not?[192]

---

[186] Ashleigh Fields, *Meta Executive Apologizes After Conservative Activist Robby Starbuck Sues Over AI Chatbot Responses*, THE HILL (Apr. 30, 2025, 6:02 PM), https://thehill.com/policy/technology/5276077-meta-apologizes-robby-starbuck/.

[187] *Id.*

[188] *Id.*

[189] Mike Isaac & Theodore Schleifer, *Here's What to Know*, Pinned Post to *Meta Says It Will End Its Fact-Checking Program on Social Media Posts*, N.Y. TIMES (May 24, 2025), https://www.nytimes.com/live/2025/01/07/business/meta-fact-checking.

[190] *Id.*

[191] Beatriz Sampaio, *Meta's Fact-Checking Rollback: Governance, Free Speech, and User Safety*, BERK. TECH. L.J. (May 23, 2025), https://btlj.org/2025/05/metas-fact-checking-rollback-governance-free-speech-and-user-safety/.

[192] Steff Danielle Thomas, *Robby Starbuck Calls Meta Apology "Sort of Bizarre" Amid Defamation Lawsuit*, THE HILL (May 10, 2025, 10:35 AM), https://thehill.com/regulation/court-battles/5293760-robby-starbuck-meta-apology-defamation-case/.

E.   *LTL LED, LLC. v. Google LLC*

Following closely on the heels of *Starbuck v. Meta*, the most recently filed AI libel lawsuit is *LTL LED, LLC v. Google LLC*, initially filed in Minnesota state court but quickly removed to federal court by Google.[193] The plaintiffs are Wolf River Electric (a company engaged in the business of selling, designing, installing, and repairing commercial and residential solar panels) and four of its officers who were mentioned by name in certain Google AI overviews about the business.[194] The Complaint alleges that Google's AI chatbot published false statements about Wolf River Electric and its officers, essentially claiming that the solar energy company was facing a lawsuit from the Minnesota Attorney General for deceptive sales practices.[195] The problem, of course, was that Wolf River Electric was never sued by the Minnesota Attorney General or even accused of the alleged misconduct.

According to the Complaint, in early September 2024, employees at Wolf River learned that Google's AI chatbot was publishing statements falsely informing the public that the company had been sued by the Minnesota Attorney General.[196] Specifically, Google stated:

> According to recent news reports, Wolf River Electric is currently facing a lawsuit from the Minnesota Attorney General due to allegations of deceptive sales practices regarding their solar panel installations, including misleading customers about cost savings, using high-pressure tactics, and tricking homeowners into signing binding contracts with hidden fees; many customers claim they were not properly informed about the total cost of their solar systems and experienced significant issues with installation and functionality after signing contracts.[197]

Google's AI listed four websites as sources for its claims. The first of these was a Minnesota Star Tribune news article about Minnesota homeowners facing hidden fees when going solar.[198] However, as the plaintiffs' Complaint points out, the Minnesota Star Tribune says nothing about Wolf River facing a lawsuit, misleading customers, or customers experiencing issues with costs, installation, or the functionality of their solar energy

---

[193] Complaint at 1–2, LTL LED, LLC. v. Google LLC, No. 0:25-cv-02394 (D. Minn. June 9, 2025).

[194] Complaint at 1, 19, LTL LED, LLC. v. Google LLC, No. 0:25-cv-02394 (D. Minn. June 9, 2025); Plaintiff's Joint Amended Complaint at 1, LTL LED, LLC, v. Google LLC, No. 0:25-cv-02394 (D. Minn. June 9, 2025).

[195] Complaint at 2–3, 6, LTL LED, LLC. v. Google LLC, No. 0:25-cv-02394 (D. Minn. June 9, 2025).

[196] *Id.* at 2.

[197] *Id.* at ¶ 10.

[198] *Id.* at ¶ 14.

systems.[199] The second source cited by Google was an Angie's List review.[200] Again, however, the review does not state anything about a Minnesota Attorney General's Office lawsuit, nor does it state that Wolf River misled customers, tricked customers about costs, or hid fees from them.[201] The third source Google's chatbot cited was an April 26, 2022, Minnesota Attorney General publication about a lawsuit its office filed against ten solar energy-related parties.[202] Yet, again, none of these parties is Wolf River, and Wolf River is never mentioned in the Attorney General publication.[203] Given that, there is no mention whatsoever in the publication of Wolf River misleading customers, using high-pressure sales tactics, tricking customers, not properly informing customers about costs or hidden fees, or experiencing issues with installation or functionality.[204] The fourth and final source cited by Google is an April 26, 2022, KROC News article titled "Minnesota AG Suing Four Solar Panel Sales Companies."[205] Once again, the article makes no mention of Wolf River (naming four other companies instead), and does not state that Wolf River is facing a lawsuit, allegations of deceptive sales practices, or any other claims attributed by Google's AI to Wolf River or its customers.[206]

According to the plaintiffs, despite the fact that none of Google's "sources" supported its claims, it published the false allegations, nonetheless. Even worse, Google's search engine exacerbated the damage to Wolf River's business reputation by suggesting certain "autocomplete" predictive phrases for individuals searching for "Wolf River Electric."[207] Users were directed to such options as "Wolf River Electric lawsuit," "Wolf River Electric lawsuit update Minnesota," Wolf River Electric Lawsuit 2022," and "Wolf River Electric lawsuit Minnesota settlement."[208] And, as the lawsuit notes, Google's own website admits that such autocomplete predictions "reflect common and real searches that other people have done which are trending in the same location of where the query is coming from, and based off the language of the query."[209]

The lawsuit acknowledges that Google's AI chatbot does include a hidden disclaimer (viewable only if the reader clicks on the "show more"

---

[199] *Id.* at ¶ 15–18.
[200] *Id.* at ¶ 19.
[201] *Id.* at ¶ 20–25.
[202] *Id.* at ¶ 26–27.
[203] *Id.* at ¶ 28–32.
[204] *Id.*
[205] *Id.* at ¶ 36.
[206] *Id.* at ¶ 37–45.
[207] *Id.* at ¶ 65.
[208] *Id.* at ¶ 62.
[209] *Id.* at ¶ 64.

option). This disclaimer states "Generative AI is experimental. For legal advice, consult a professional."[210] However, as the complaint indicates, this minor disclaimer "says nothing about the publications being untrue, false, utterly fabricated, and does not contain any language about whether a person should rely on Google for solar purchasing advice or trust Google for the statements it provided."[211] As the suit asserts, members of the public, "including potential employees and customers," have viewed the defamatory statements hallucinated by Google's chatbot and "have relied upon these false statements in regurgitating, spreading, and further defaming Wolf River, all because of Google's false and unsupported statements."[212]

As the Complaint continues, there is ample evidence that people believed that the chatbot's libelous and false statements were true, and that the plaintiffs suffered quantifiable damages as a result. The lawsuit lists multiple examples of customers (identified by their contract amounts and in some cases by specific contract numbers) who terminated their business relationships with Wolf River after learning of the fictious Attorney General lawsuit and the hallucinated deceptive trade sales practices.[213] These include a customer who terminated a $39,680.00 contract after Googling Wolf River and learning of "the lawsuit"; a prospective customer who backed out of a $26,400.00 contract after "research online" revealed that "Wolf River was being sued by the Minnesota Attorney General"; another existing customer who cancelled a $150,000.00 contract "because of the publications on Google"; and a non-profit corporation that "pulled the plug" on a $147,400.00 solar project because of Wolf River's "lawsuits in the last year" with the Attorney General's Office.[214]

The Complaint advances causes of action for defamation, defamation per se, defamation by implication, and violations of the Minnesota Deceptive Trade Practices Act.[215] It seeks unspecified damages in excess of $50,000.00.[216] Will the lawsuit succeed? At such an early stage in the litigation, one can only speculate that as to the outcome. It is, however, interesting that the AI Overview chatbot feature that included the purportedly

---

[210] *Id.* at ¶ 108.
[211] *Id.* at ¶ 111.
[212] *Id.* at ¶ 116.
[213] *Id.* at ¶¶ 127–133.
[214] *Id.* at ¶¶ 122–125, 127, 129–130.
[215] *Id.* at ¶¶ 138, 220–221, 227, 255.
[216] *Id.* at ¶ 212.

hallucinated false statements has apparently been turned off by Google, at least as recently as the suit's filing.[217]

It is also important to distinguish the Wolf River Electric case from other AI libel cases. Unlike the *Starbuck v. Meta* case, *Walters v. OpenAI*, or even the *Battle v. Microsoft* matter, this lawsuit involves plaintiffs who will likely be considered by the federal court to be private figures rather than public ones. Perhaps even more significantly, unlike the *Walters* case, the Wolf River Electric lawsuit appears to boast not only evidence that at least some people believed the Google chatbot's fabricated allegations, but also evidence that these false allegations led to documentable financial losses.

F. Mecum Auctions, Inc. v. Trustpilot, Inc.

In March 2025, Mecum Auctions, Inc., a company that specializes in the auctions of collector automobiles, filed a lawsuit in New York state court in Manhattan.[218] Mecum sued Danish-owned customer review site Trustpilot, seeking the removal of a purportedly defamatory review.[219] The suit asserted that not only had the negative episode described in the review in question never happened and was therefore "hallucinated" by Trustpilot's AI, Trustpilot itself had relied upon an "AI-powered defamation tool" in mistakenly deciding that the review was not defamatory.[220] According to the lawsuit, the review "described a fictional event from, at best, an unknown bidder that accused Mecum of unscrupulous business practices, if not fraud or other criminal offenses, without any proof."[221]

The review at issue came from a user identified only as "Thomas Clarke," who claimed that an auctioneer at Mecum had pressured Clarke to put up a car during a Kansas City auction, only to issue a payment that was late and returned for insufficient funds.[222] The review quoted "Clarke" as writing "no car, no money and a very crappy taste in my mouth."[223]

Mecum pushed back, stating that its records contained no information on a "Thomas Clarke" consigning a vehicle at the Kansas City

---

[217] The author attempted to duplicate the Google searches referenced in the suit using AI overview but was unable to do so.

[218] *Customer Review Site Kept Bad Review Based on Green Light from AI Tool, Car Auctioneer Says*, N.Y. L. J. ONLINE, Mar. 7, 2025, LEXIS.

[219] *Id.*

[220] *Id.*

[221] *Id.*

[222] *Id.*

[223] *Id.*

auction "or any other auction."[224] After unsuccessful attempts to reach Clarke, Mecum concluded the review was false and defamatory.[225] However, Trustpilot and its "Content Integrity Team" disagreed. Trustpilot insisted that the review had been examined and "cleared by an AI tool that reviews whether disputed content, within the context of the whole review, is likely to cause reputational or business harm."[226]

The *Mecum Auctions* lawsuit is interesting because it not only alleges defamation as a result of a platform chatbot's hallucination, but also that the defendant's own AI-powered screening tool designed to prevent such defamation from happening failed as well. However intriguing this added layer of chatbot liability would have been to consider, the lawsuit ended via a confidential settlement on June 27, 2025.

IV.   CONCLUSION

As the Brian Hood matter demonstrated, "libel by chatbot" transcends national borders. While examination of this subject on an international level is beyond the scope of this article, it is important to know that it can happen anywhere. For example, in 2025, a Norwegian man named Arve Hjalmar Holmen asked ChatGPT for information about himself and received a hallucinated reply that he was a child killer.[227] The chatbot's output stated that Holmen was the father of two boys aged seven and ten whom he murdered, that the boys' bodies were found in a pond near their Trondheim, Norway home in December 2020, and that Holmen was convicted of murder and received a twenty-one year prison sentence.[228] ChatGPT's "facts" were false, and Holmen and a digital advocacy organization, Noyb European Center for Digital Rights, filed a complaint with the Norwegian Data Protection Authority.[229] The complaint asserted that Holmen has "never been accused nor convicted of any crime and is a conscientious citizen," and maintained that ChatGPT's "defamatory" response violated the General Data Protection Regulation's ("GDPR") accuracy requirements.[230]

---

[224] *Id.*

[225] *Id.*

[226] *Id.*

[227] Imran Rahman-Jones, *Man Files Complaint After ChatGPT Said He Killed His Children*, BBC (Mar. 21, 2025), https://www.bbc.com/news/articles/c0kgydkr516o.

[228] *Id.*

[229] Complaint, *Holmen v. OpenAI OpCo, LLC* (Mar. 20, 2025), https://noyb.eu/sites/default/files/2025-03/OpenAI_complaint_redacted.pdf.

[230] *Id.*

Mr. Holmen's case illustrates a key difference between the European and American approaches to "libel by chatbot." In the United Kingdom, the Defamation Act 2013 was enacted to reform the often-criticized defamation law in England and Wales.[231] Although Section 5 of the Act protects website operators from liability for user-generated content, AI providers who themselves generate content (like Meta AI) would likely be treated like primary publishers or authors who have to rely on traditional defamation defenses like truth.[232] But the most important regulatory framework to address AI-generated harm in the EU is its GDPR.[233] Under the GDPR, false AI statements that relate to an individual can be treated as inaccurate "personal data," allowing affected persons to demand erasure of the false information or rectification.[234]

The United States, however, lacks such a regulatory framework. In the absence of a federal law addressing AI-generated harms, the victims of "libel by chatbot" will be at the mercy of myriad and potentially contradictory state court interpretations and state regulations. The tech industry is not likely to provide meaningful solutions any time soon, either. "The likelihood of error and bias in the work of generative AI [tools] is exacerbated by the fact that over time, the training data for the large language models will include [the chatbots] themselves[—]including their hallucinations."[235]

Once "a mature, economically robust, and politically powerful industry," the press—and the public's view of it—has been upended by "a confluence of economic, technological, and socio-political change."[236] Economic pressures in particular have led to the news industry's embrace of generative AI. When insufficient attention is paid to the need for human review of AI-generated newsroom content, the result is not only an increase in potential liability, but also a continued erosion of the already plummeting public trust in journalism.[237]

---

[231] Defamation Act 2013, c. 26 (UK).

[232] *Id.*

[233] General Data Protection Regulation, Regulation (EU) 2016/679, 2016 O.J. (L 119) 1 (Apr. 27, 2016).

[234] *Id.*

[235] Lili Levi, Countering the Mosaic of Threats to Press Functions, KNIGHT FIRST AMEND. INST. (Apr. 29, 2024), https://knightcolumbia.org/blog/countering-the-mosaic-of-threats-to-press-functions.

[236] *Id.*

[237] In fact, an international study led by the BBC and coordinated by the European Broadcasting Union revealed that 45 percent of AI assistants misrepresented news content and often contained significant accuracy issues such as hallucinated details. *Largest Study of Its Kind Shows AI Assistants Misrepresent News Content 45% of the Time—Regardless of Language or Territory*, BBC (Oct. 22, 2025), https://www.bbc.co.uk/mediacentre/2025/new-ebu-research-ai-assistants-news-content.

In *New York Times v. Sullivan*, the U.S. Supreme Court famously held for the first time that the First Amendment imposes limits on the common law of defamation, ruling that defendants would face liability for defamatory statements about public figures only if they knew the statements to be false or recklessly disregarded the truth or falsity—in other words, actual malice.[238] Since then, the Court soon extended the actual-malice rules,[239] and eventually "established a different set of rules for private figures involved in newsworthy events or 'matters of public concern.'"[240] The result was overturning the common law's presumptions of fault and falsity and imposing an actual malice requirement in the vast majority of cases involving the media.[241]

Criticism of the *Sullivan* standard has been mounting, with at least two members of the Supreme Court calling for its reconsideration.[242] A number of legal scholars also agree with the need for at least *some* reform of the defamation law, even if they do not agree with abandoning the actual malice standard. As Lyrissa Lidsky has pointed out, "defamation must introduce remedies that vindicate individual reputations and surface the truth more readily."[243] Even in the absence of reform, actual malice might still be proven against those responsible for an AI chatbot. If an AI company were alerted that its program was generating specific false and libelous content and took no action, then it would be acting with reckless disregard for the truth.

We should not fear change in the legal system, and should not shirk from adapting existing legal doctrines to address the challenges presented by emerging technologies. Already, there are efforts underway to apply product liability principles to protect against certain harms caused by AI chatbots.[244] AI chatbots can, and do, provide false information and they

---

[238] N.Y. Times Co. v. Sullivan, 376 U.S. 254, 279–80 (1964).

[239] *See, e.g.*, St. Amant v. Thompson, 390 U.S. 727, 733 (1968).

[240] Levi, *supra* note 235 (quoting Phila. Newspaper v. Hepps, 475 U.S. 767, 774 (1986)).

[241] *See*, e.g., Assoc.'d Press v. Walker, 389 U.S. 28 (1967); Curtis Publ'g Co. v. Butts, 388 U.S. 130, 133–34 (1967); Gertz v. Robert Welch, 418 U.S. 323, 335–36 (1974); Hustler Mag. v. Fallwell, 485 U.S. 46, 56 (1988).

[242] *See*, e.g., Berisha v. Lawson, 141 S. Ct. 2424, 2430 (2021) (Thomas, J., dissenting; Gorsuch, J., dissenting); McKee v. Cosby, 586 U.S. 1172, 1173 (2019) (Thomas, J., concurring); Adam Liptak, *Justice Clarence Thomas Calls for Reconsideration of Landmark Libel Ruling*, N.Y. Times (Feb. 19, 2019), https://www.nytimes.com/2019/02/19/us/politics/clarence-thomas-first-amendment-libel.html.

[243] Lyrissa Lidsky, *Defamation Law and the Crumbling Legitimacy of the Fourth Estate*, KNIGHT FIRST AMEND. INST. (July 11, 2024), https://knightcolumbia.org/blog/defamation-law-and-the-crumbling-legitimacy-of-the-fourth-estate.

[244] *See*, e.g., A.F. v. Character Techs., Inc., No. 2:24-cv-01014 (E.D. Tex. Dec. 9, 2024); Garcia v. Character Techs., Inc. et al., No. 6:24-cv-01903 (M.D. Fla. Oct. 22, 2024); John G. Browning, *A Product By Any Other Name? The Evolving Trend of Product Liability Exposure for Technology Platforms*, 16 Elon L. Rev. 1, 181 (2024).

do so convincingly. Regardless of the reason(s) for the false output—flaws in the algorithm design that caused it to generate defamatory content, a lack of human oversight, etc.—libel law must adapt to provide a remedy for libel by chatbot.