University of Arkansas at Little Rock Law Review

Volume 46 | Issue 1 Article 2

2023

Cracking The Shield: CDA Section 230, Algorithms, and Product Liability

Kevin Ofchus

Follow this and additional works at: https://lawrepository.ualr.edu/lawreview

 Part of the Torts Commons

Recommended Citation

Kevin Ofchus, *Cracking The Shield: CDA Section 230, Algorithms, and Product Liability*, 46 U. ARK. LITTLE ROCK L. REV. 27 (2023).
Available at: https://lawrepository.ualr.edu/lawreview/vol46/iss1/2

This Article is brought to you for free and open access by Bowen Law Repository: Scholarship & Archives. It has been accepted for inclusion in University of Arkansas at Little Rock Law Review by an authorized editor of Bowen Law Repository: Scholarship & Archives. For more information, please contact mmserfass@ualr.edu.

CRACKING THE SHIELD: CDA SECTION 230, ALGORITHMS, AND PRODUCT LIABILITY

Kevin Ofchus[*]

I. INTRODUCTION: PRACTICAL APPLICATIONS AND COMPLEXITY OF *GONZALEZ V. GOOGLE LLC*

Imagine if you walked outside your office, home, onto your patio, or even stopped your car on the interstate highway, as thousands of cars passed by, and shouted the random phrase: "Buy the EthereumMax crypto stock!"[1] This statement would by itself, absent some added relevant context, not create much interest or liability. If, however, you were on an internet highway such as Facebook (Meta), Twitter, Snapchat, Instagram, or Tik Tok, your financial

---

[*] Assistant District Attorney with the Griffin Judicial District, Georgia. J.D. Mercer University School of Law. B.A. Philosophy and English from Randolph-Macon College. I am indebted to my family for their moral support and my editors for their due diligence. This product is dedicated to Brendan Alexai; destined for great things, one who loves numbers, never short on words, and unremitting in his search for knowledge.

1. For context, see Arushi Geol et al., *Sanctioning a Cryptocurrency Protocol: What does that mean for Web3?*, WORLD ECON. F. (Oct. 17, 2022), https://www.weforum.org/agenda/2022/10/cryptocurrency-regulation-sanctions-web3/. "In August 2022, the Office of Foreign Assets Control (OFAC) of the United States Treasury Department sanctioned a cryptocurrency 'mixer'—a program used to increase the anonymity of crypto transactions—for its alleged use in money laundering. It also blacklisted several Ethereum addresses associated with the protocol." *Id.* (cleaned up). OODA is an acronym for Observe-Orient-Decide-Act. *About OODA* Loop, OODA, https://www.oodaloop.com/about/ (June 20, 2023). Less than sixty days later the Securities Exchange Commission (SEC) boldly asserted the law regarding statements of information exploited over the internet. *SEC Charges Kim Kardashian for Unlawfully Touting Crypto Security*, SEC. EXCH. COMM'N. (Oct. 3, 2022), https://www.sec.gov/news/press-release/2022-183. The SEC's press release stated that federal securities laws were clear regarding culpability for a celebrity or other individuals promoting a crypto asset security. *Id.* Gurbir S. Grewal, Director of the SEC's Division of Enforcement, emphasized the disclosure of bias in the publicity of a security was controlled by law: "[i]nvestors are entitled to know whether the publicity of a security is unbiased . . . ." *Id.* On October 3, 2022, Kim Kardashian agreed to settle SEC charges against her and pay $1.26 million in penalties, disgorgement, and interest. *Id.* The penalty was for SEC violations for touting on social media a crypto asset security offered and sold by EthereumMax without disclosing the payment she received for the promotion. *Id.* This was a violation of the anti-touting provision of federal securities laws. *Id.* On December 2, 2022, Chris Wray, Director of the FBI, expressed concern the Chinese Government had the ability to control Tik Tok's recommendation algorithm, allowing the Chinese Government to manipulate content and, if they so choose, to use it for influence operations. Eric Tucker, *FBI Director Raises National Security Concerns About TikTok*, AP NEWS (Dec. 2, 2022), https://apnews.com/article/technology-china-united-states-national-security-government-and-politics.

interest—for example, compensation for advertising EthereumMax stock—could expose you to liability.[2] The monetization of speech and information carries culpability when one's interest or bias in the speech is present.[3] Because of the elevated potential for deception and confusion in commercial speech, the United States Supreme Court established a factor test to determine whether speech is commercial speech or non-commercial speech, in *Bolger v. Youngs Drug Products Corp*.[4] Speech may be properly characterized as commercial when: (1) it is concededly an advertisement, (2) it refers to a specific product, or (3) it is motivated by an economic interest in selling the product.[5] The U.S. Court of Appeals for the Tenth Circuit, following the precedent set in *Bolger*, stated that a common-sense distinction exists when speech is motivated by an economic interest in selling a product.[6] The "common-sense" distinction suggests that certain speech, namely, commercial advertising, is subject to less protection than pure information-based speech.[7] For the EthereumMax example, the distinction turns on whether the speaker has a vested interest in gaining the economic benefit from promoting EthereumMax stock or if the speaker is disinterested entirely. In the context of investments, vis-à-vis the Securities Act of 1933 and 15 USCA § 77q(a)–(b), the Supreme Court has limited speech when such speech relates to economic interest.[8] Section 230 of the Communications Decency Act (CDA) offers a shield.[9] So, while parties may profit from advertising a third party's phrase—"Buy the EthereumMax crypto stock!"—immunity resides in the parties' role as a publisher.[10]

---

2. 15 U.S.C. § 77q(b) provides that "[i]t shall be unlawful for any person, by the use of any means or instruments of transportation or communication in interstate commerce or by the use of the mails, to publish, give publicity to, or circulate any notice, circular, advertisement, newspaper, article, letter, investment service, or communication which, though not purporting to offer a security for sale, describes such security for a consideration received or to be received, directly or indirectly, from an issuer, underwriter, or dealer, without fully disclosing the receipt, whether past or prospective, of such consideration and the amount thereof."

3. 15 U.S.C. § 77q(a) (falling under the title "[u]se of interstate commerce for purposes of fraud or deceit.").

4. Bolger v. Youngs Drug Products Corp., 463 U.S. 60, 65, 68 (1983) ("Advertisers should not be permitted to immunize false or misleading product information from government regulation simply by including references to public issues.").

5. *Id.* at 66–67.

6. United States v. Wenger, 427 F.3d 840, 846–48 (10th Cir. 2005) (citing *Bolger,* 463 U.S. at 66–67). Speech is commercial if the speech is motivated by an economic interest in selling the product. *Id.*

7. *Wenger*, 427 F.3d at 847.

8. *Bolger*, 463 U.S at 67.

9. Communications Decency Act of 1996 (CDA), 47 U.S.C. § 230(c).

10. *See id.*

The power of Section 230 immunity, as interpreted by U.S. courts, is expansive.[11] While terrorism defined in 18 U.S.C. § 2331 may be at issue for Section 230 immunity in *Gonzalez v. Google LLC* and *Force v. Facebook*,[12] the tempest which creates the liability is much closer to shore.[13] Instead of posting cryptocurrency financial advice on social media, the statement could have easily been something far more dangerous; "Hang Mike Pence!" could have been the statement, instead of the rally to buy stock. The issue would turn to the people who had an interest in hanging the sitting Vice President of the United States, rallying to violent action, and inflicting great harm on others.[14] Further, add in the element of a social media platform algorithm specifically and directly targeting certain viewers.[15] Compounding the direct targeting, what if the responses to the original "Hang Mike Pence!" post were then transferred to a feedback loop by the algorithm exclusively for those within the algorithm's profile?[16] With the intent and basis for creating the algorithm to generate interest and increase advertising revenue.[17] Upon viewing this post, the targeted groups communicated a time and place to amass at the Capital on the day electors were to be submitted for the election of the President of the United States. At the Capital, the targeted groups displayed effigies of

---

11. Force v. Facebook, Inc., 934 F.3d 53, 63 (2d Cir. 2019); *see also* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1168, 1171–72 (9th Cir. 2008) (distinguishing broad and unduly broad application of 47 U.S.C. § 230).

12. *See* Gonzalez v. Google LLC, 2 F.4th 871, 889 (9th Cir. 2021); *Force*, 934 F.3d at 67.

13. Anti-Terrorism Clarification Act of 2018 (ATA), 18 U.S.C. § 23331(1)–(5). Sections 1 –5 define international and domestic terrorism as "violent acts or acts dangerous to human life that are a violation of the criminal laws of the United States or of any State, or that would be a criminal violation if committed within the jurisdiction of the United States or of any State." *Id.* The ATA incudes acts which appear to be "(i) intended to intimidate or coerce a civilian population; (ii) influence the policy of a government by intimidation or coercion; or (iii) to affect the conduct of a government by mass destruction, assassination, or kidnapping." *Id.* (1)(B)(i)–(iii). While the ATA is not addressed specifically here, once the distinction is shown in the facts of a claim, between § 230(f)(2) and (f)(3), the analysis for liability presents itself as abandoning the shield of immunity afforded under § 230(c)(1).

14. Martin Pengelly, *"Hang Mike Pence": Twitter Stops Phrase Trending After Capitol Riot*, GUARDIAN (Jan. 10, 2021), https://www.theguardian.com/us-news/2021/jan/10/hang-mike-pence-twitter-stops-phrase-trending-capitol-breach.

15. Twitter suspended President Donald J. Trump's Twitter access for incitement of violence. *See id.*

16. This leads to algorithmic confounding, loss of utility, and the homogenizing of user behavior. *See* Allison J.B. Chaney et al., *How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility*, ARXIV (Nov. 27, 2018), https://arxiv.org/pdf/1710.11214.pdf (Recommender Systems Conference, Vancouver, BC, Canada, Oct. 2–7, 2018).

17. Allison Zakon, Comment, *Optimized for Addiction: Extending Product Liability Concepts to Defectively Designed Social Media Algorithms and Overcoming the Communications Decency Act*, 2020 WIS. L. REV. 1107, 1113–14 (2020).

a noose and invaded the building seeking to hang Mike Pence.[18] For those committed to the violent action, liability of an algorithm targeting and redistributing certain communication shifts from an unprotected economic interest in commercial speech to incitement of foreseeable harm with a likely and probable consequence which can reasonably be anticipated.[19]

When Section 230 of the CDA (hereinafter § 230) was crafted in 1996, the prism of Internet cyberspace, particularly social media, was beyond the grasp of the legislation.[20] However, there was an emphasis on acknowledging the rapid growth of the Internet.[21] Additionally, there was a focus on promoting the development of technologies to maximize user control over what information was received by individuals, families, and schools.[22] Congress balanced concerns about Internet development with an intent to preserve its vibrant and competitive free market.[23]

Algorithms are critical to the Internet marketplace. Algorithms are an essential product for the market to generate over fifty billion dollars in yearly advertising revenue from Google to Snapchat.[24] Principally, algorithms are products created to exploit, develop, and repackage data; the content of the algorithm uses the raw material of re-packaged data as a marketing product which advertisers pay for on social media.[25] The algorithms—created and produced within a nexus of machine learning, artificial intelligence (AI), and recommender feedback loops[26]—are not disinterested. They are invented, created, and produced to maximize economic interests.[27] Shielded from liability, algorithms have generally operated with impunity.

---

18. *See* Pengelly, *supra* note 14. Mike Pence fled the Capital under guard by the Secret Service as his life was threatened. *See id.*

19. *Id.* Jim Bourg, a Reuters editor in Washington on January 6, 2021, said on Twitter: "I heard at least three different rioters at the Capitol say that they hoped to find Vice-President Mike Pence and execute him by hanging him from a Capitol Hill tree as a traitor. It was a common line being repeated. Many more were just talking about how the VP should be executed." *Id.*

20. *See* 47 U.S.C. § 230. Section 230 does not mention social media, nor obviously, could it have. However, it was explicitly cognizant of harm to and abuse of users of the internet as identified in § 230(b) Policy. *See id* § 230(b).

21. *See id.* § 230(b)(1).

22. *See id.* § 230(b)(3).

23. *See id.* § 230(b)(2).

24. *See Social Network Advertising Revenues in the United States from 2017 to 2021*, STATISTA (Jan. 6, 2023), https://www.statista.com /statistics/271259/advertising-revenue-of-social-networks-in-the-us/; *see also* Kalev Leetaru, *What Does It Mean for Social Media Platforms to "Sell" Our Data?*, FORBES (Dec. 15, 2018, 3:56 PM), https://www.forbes.com/sites/kalevleetaru/2018/12/15/what-does-it-mean-for-social-media-platforms-to-sell-our-data/?sh=763a80452d6c.

25. *See* Michael S. Gal & Nicolas Petit, *Radical Restorative Remedies for Digital Markets*, 36 BERKELY TECH. L.J. 617, 636–40 (2021).

26. *See infra* Section II.B.

27. *See* Gal & Petit, *supra* note 25, at 626.

Absent meaningful legislation on the relationship between algorithms, immunity, liability, and § 230, the Federal Circuit Courts of Appeals have struggled in applying § 230.[28] This Article seeks to analyze the application of product liability to algorithms. Specifically, the examination will focus on § 230 liability within algorithms as defective products,[29] algorithms' functional relationship to content production not as a service provider,[30] and how algorithms act in material contribution to certain conduct.[31] The analysis leads this Article to conclude that algorithms can have a defective design in the line of production.[32] Elements of machine learning, AI, and recommender algorithms play a central role in creating these defective products.[33] Thus, by not addressing the critical issue of algorithms as products in *Gonzalez v. Google LLC*,[34] the Court left open a perilous gateway. The liability of algorithms in machine learning, AI, and recommender algorithms will not abate as the daily tide of market use and misuse increases.

Section II begins with an analysis of § 230(c), the purpose of the statute, and subsections (b), (e), and (f).[35] Section 230(c) is the basis for a string of cases appearing in multiple Circuit Courts of Appeals,[36] forming the genealogy of *Gonzalez v. Google*.[37] The key determining factor in these cases is the delineation of content provision from service provision. Service provision is shielded from liability, while content provision is not.[38] Each case touches upon § 230 in varying degrees of scope by balancing the statute's text and

---

28. Kimzey v. Yelp! Inc., 836 F.3d 1263, 1269 (9th Cir. 2016); *see also* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1169–72 (9th Cir. 2008) (en banc).

29. *See infra* Section III.

30. *See infra* Section II.

31. *See infra* Section IV.

32. *See infra* Section V.

33. *See* Chaney et al., *supra* note 16 (describing the implications and unintended consequences of algorithmic confounding which impacts user behavior and may decree utility).

34. Gonzalez v. Google LLC, 2 F.4th 871, 894–96 (9th Cir. 2021). The court addresses algorithms providing a neutral platform not treating content on the platform YouTube differently, and thus immune under Section 230. *Id.* However, the court does not reach the issue of machine learning or artificial intelligence algorithms as products which may materially contribute to content creation and thus not shielded under CDA Section 230. *See id.*

35. *See infra* Section II.

36. *See infra* Section II.A.1.a–c.

37. Across multiple federal circuit courts and one state supreme court, each case addresses an element of immunity and potential liability relative to algorithms either directly or indirectly. *See, e.g.*, *Gonzalez*, 2 F.4th at 886; Erie Ins. Co. v. Amazon.com, Inc., 925 F.3d 135, 139 (4th Cir. 2019); A.M. v. Omegle.com, LLC, 614 F. Supp. 3d 814, 818 (D. Or. 2022); Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1161 (9th Cir. 2008) (en banc); Lemmon v. Snap, Inc., 870 S.E.2d 739, 743 (Ga. 2022) (offering a line of reasoning for a product liability claim). In response, the most potent counterargument lies in Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1094 (9th Cir. 2019), which is examined *infra* Section II.

38. 47 U.S.C. § 230(c)(1)–(2), (f)(1)–(3).

intent with its common law application regarding content and service.[39] Additionally, the recent holding of *Maynard v. Snapchat*,[40] taking on the issue of algorithms and products liability in March 2022, adds the perspective of state courts.[41] Section II then examines how algorithms exist as part of content-specific products, the algorithm's role in the development of hyper-personalized content, and how algorithms can perpetuate feedback loops.[42] Section III addresses algorithms defined within the domain of product liability, namely, intangible goods, defective conditions, product design elements, the reasonableness of danger, and risk-utility.[43]

Section IV addresses the impact of algorithms on causation and material contribution vis-a-vis substantial assistance and foreseeability.[44] In that Section, intended use and use relative to the liability of the conduct are examined. The analysis of cost benefits and risk-utility considers how algorithms might work to encourage certain conduct that leads to liability inherent in their design and use. Finally, this Article offers a solution establishing a set of factors to consider in a court's ruling or the legislature's amending of § 230 in response to the Supreme Court's decision in *Gonzalez v. Google LLC* and beyond.[45]

II. ESTABLISHING THE DISTINCTION: CONTENT PROVIDERS AND NOT SERVICE PROVIDERS

After addressing the relevant elements of § 230, this Section examines recent decisions giving rise to an evolving algorithm presence as content providers since the CDA was established in 1996. Then, Section II addresses algorithm creation, development, and an algorithm's relationship with content, thereby forming a product.

---

39. *See generally Gonzalez*, 2 F.4th 886 (18 U.S.C. § 2333(d) Justice Against Sponsors of International Terrorism Act (JASTA) not limiting the Communications Decency Act (CDA) § 230; *see Erie Ins. Co.*, 925 F.3d at 144 (holding Amazon not shielded by § 230, however Amazon was not the seller thus not liable to the insurer under Maryland law for the product liability claim); *see Omegle.com*, 614 F. Supp. 3d at 820–22 (citing *Lemmon* by analogy by stating a party may sue and not violate § 230 immunity "if the conduct underlying the claim constitutes a violation of Section 1591 (18 U.S.C. 2421(a)—Fight Online Sex Trafficking Act (FOSTA)—where mens rea is actual knowledge mental state); *see Roommates.com*, 521 F.3d at 1161; *see Lemmon*, 870 S.E.2d at 743 (offering a line of reasoning for a product liability claim).

40. Maynard v. Snapchat, Inc., 870 S.E.2d 739 (Ga. 2022).

41. *Id.*

42. *See infra* Section II.B.

43. *See infra* Section III.A.

44. *See infra* Section IV.

45. *See infra* Section IV–V.

A. Section 230 of the CDA and the Double-Edged Sword: Business is Content, Service is a Proxy

Following Section 223 of the CDA,[46] Congress enacted § 230 with a specific purpose in mind. Section 223, while limited to obscene materials or communications with the intent to abuse, threaten, or harass another person,[47] offers insight into the conflicting issues at the core of § 230. When § 230 was enacted, telecommunications, absent the Internet, were the most expansive communication platform requiring regulation.[48] Among restrictions and judicial remedies, various defenses to liability were enacted.[49] Specifically, corporate liability was shielded by statute unless the conduct was within the scope of employment or agency, and the corporate employer knowing of such conduct, authorized or ratified the conduct, or the conduct was recklessly disregarded by the corporate employer.[50] A defense to corporate liability existed where there had been good faith, reasonable, and effective action to prevent the conduct prohibited under Section 223.[51] This included instituting any technologically feasible regulatory measures.[52] Algorithms, software applications, and the ubiquitous cellular telephone were not considered nor mentioned in the statute.[53] The statute expressly distinguished telecommunications devices from interactive computer services under § 230(e).[54] Section 230(f)(2)–(3), respectively, defines the interactive computer service and the information content provider.[55]

While never addressing algorithms, the United States Court of Appeals for the Fourth Circuit distinguished these two elements in *Zeran v. America Online, Inc.* shortly after the statute was enacted.[56] The court cited § 230(c)(1) in holding that the statute's plain language creates an immunity shielding any cause of action making service providers liable for information originating

---

46. 47 U.S.C. § 223 (targeting obscene or harassing telephone calls).

47. *Id.*

48. VALERIE C. BRANNON & ERIC N. HOLMES, CONG. RSCH. SERV., RL46751, SECTION 230: AN OVERVIEW 2–3 (2021) (discussing the modernization of "existing protections against obscene, lewd, indecent and harassing uses of telephone lines").

49. *See* ,

50. *See id.* § 223(e)(4).

51. *See id.* § 223(e)(5)(A)–(B).

52. *Id.*

53. *See id.* § 223(a)(1)(A)(ii).

54. *See id.* § 230(e), (h)(1)).

55. *See* 47 U.S.C. 230(f)(2)–(3). The interactive computer service applies to any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service that provides access to the internet. *Id.* Under § 230(f)(3), an information content provider means any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the internet or any other internet computer service. *Id.*

56. Zeran v. America Online, Inc., 129 F.3d 327, 330–31 (4th Cir. 1997).

with a third-party user of the service.[57] The holding specifically applied to a computer service provider's role in a "publisher's traditional editorial function," barring liability.[58] The court was silent on content providers under § 230(f)(3).

In 2008, the United States Court of Appeals for the Ninth Circuit addressed algorithms and content in *Fair Housing Council v. Roommates.com*.[59] The Ninth Circuit, while glossing over algorithms, construed § 230 to allow service providers the ability "to edit user-generated content without becoming liable for all defamatory or otherwise unlawful messages they didn't edit or delete."[60] In finding against Roommates, the court determined that "[c]ongress sought to immunize the removal of user-generated content, not the creation of content."[61] Despite being buried in a footnote of the *Roommates* holding,[62] algorithms in discovery became critical to distinguishing conduct as either a service or a content provider. While *Gonzalez v. Google* later addressed algorithms, *Gonzalez* had split holdings on the motion to dismiss.[63] Most cases in this domain do not survive the Rule 12(b)(6) motion.[64]

1. Developing Content

In asserting liability as a content provider and not as a service, it is incumbent upon the party bringing the suit to show that the entity is in *whole or in part* responsible "for the creation or development of information provided through the internet."[65] The claim of immunity arises under the argument that companies using algorithms are not liable because the companies are attenuated from the cause of harm. Section 230 immunity as a defense from liability has been applied to a wide range of claimants, each offering insight regarding the distinction between algorithms within content and service.

---

57. *Id*.

58. *Id.*

59. Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1167 (9th Cir. 2008) (en banc) (describing that when a prospective subscriber applies, Roommates's server presumably checks to ensure all required fields are complete and any credit card information is not fraudulent or erroneous; some algorithm developed by Roommates then decoded the input, transformed it into a profile page, and notified other subscribers of a new applicant or individual offering housing matching their preferences).

60. *Id.* at 1163.

61. *See id.* at 1164–75.

62. *See id.* at 1174 n.33.

63. *See* Gonzalez v. Google LLC, 2 F.4th 871, 910–13 (9th Cir. 2021).

64. David S. Ardia, *Free Speech Savior or Shield for Scoundrels: An Empirical Study of Intermediary Immunity Under Section 230 of the Communications Decency Act*, 43 LOY. L.A. L. REV. 373, 483 (2010) ("In the vast majority (84.8%) of the decisions addressing a section 230 defense prior to full discovery, neither the court nor the parties appear to have raised the issue of the proper timing for a section 230 defense.").

65. 47 U.S.C. § 230(f)(3).

a. *Roommates* and filtering tools

In *Roommates*, the website allowed users to create profiles that required disclosure of personal information during registration.[66] The information was used for offering and finding rooms for people to rent via user profile pages.[67] The information included discriminatory preferences used to filter searches and notify available matches via profile pages.[68] Roommates also included a "blank box" feature where statements could be posted.[69] Filtering tool neutrality acting with content creation was at issue. In *Roommates*, the court found that the use of neutral tools[70] on a website does not count as development under § 230 immunity.[71] If third-party user data is provided in response to service provider requirements as a condition of service and the data is used in a discriminatory fashion, in this case offending the Fair Housing Act, liability shifts to the service provider.[72] The service provider is no longer immune because it becomes a part of the developer of that information.[73]

b. *Force* and neutrality

In 2019, however, the Second Circuit held in *Force v. Facebook, Inc.* that Facebook's use of algorithms was not outside the scope of publishing, and thus immune, "as long as a third party willingly provides the essential published content."[74] While in *Roommates* the Fair Housing Act was at issue, in *Force*, the Anti-Terrorism Act was the focus.[75] In *Force*, the court held that the social media provider did not "develop" content, even though the postings on Facebook were made by officially designated terrorist groups.[76] The court reasoned that Facebook, developing algorithms designed to utilize users' information to match them with other users, was entitled to immunity as a publisher under § 230 despite its promotion of the organization's terrorist activities.[77] The court further held that immunity applied because the service,

66. Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1166 (9th Cir. 2008) (en banc).

67. *Id.* at 1165.

68. *See id.* at 1167.

69. *Id.* at 1173.

70. CLAIMS AGAINST SOCIAL NETWORKS, 4 E-COM. AND INTERNET L. 37.05[6] (2020) (stating a "neutral tool" is a function for data input that is considered by the algorithm, for example, a dropdown menu provided by the service provider for the third party to fill in the content).

71. *See Roommates.com*, 521 F.3d at 1169; *see also* 47 U.S.C. § 220(e)(4).

72. *Roomates.com*, 521 F.3d at 1166.

73. *Id.*

74. Force v. Facebook, Inc., 934 F.3d 53, 67 (2d Cir. 2019).

75. *See* Anti-Terrorism Clarification Act, 18 U.S.C. § 2333.

76. *See Force*, 934 F.3d at 68–69 (addressing material contribution).

77. *Id.* at 65–66.

Facebook in this case, acted as a "neutral intermediary" in making third-party content more visible, available, and usable.[78] Algorithms as neutral intermediaries were distinguished from algorithms that may materially alter the underlying third-party information.[79]

In *Force*, immunity applied to the interactive service provider, regardless of the specific editorial process.[80] However, the court stated that the term development was "undefined" and it broadly construed the term "publisher" as not being considered to have developed third-party content unless the defendant directly and materially contributed to what made the content itself unlawful.[81]

c. *Gonzalez*—contribution and the final straw

In *Gonzalez v. Google*, the Ninth Circuit revisited both *Force* and *Roommates*.[82] Plaintiffs in *Gonzalez*, like *Force*, argued that algorithms played a role in the development of content leading to terrorism by the international terror group ISIS.[83] The plaintiff argued that the internet provider contributed to the development of terrorism and profited from the utilization of the content which furthered ISIS terrorism.[84] The court reiterated that a "website's use of a content-neutral algorithm, without more, does not expose liability for content posted by third parties."[85] The court explicitly stated, however, that "we do not hold that machine-learning algorithms can *never* produce content . . . ."[86]

The court then broadened the scope of immunity utilizing the rationale from its *Dyroff v. Ultimate Software Group, Inc*. decision. *Dyroff* addressed an algorithm's anonymity feature enabling illegal narcotics distribution on a website.[87] The defendant's machine learning algorithm encouraged users to join groups which were discriminating them based on their posts and attributes.[88] The anonymity feature shielded users conducting illegal activity based on the defendant's impetus for the "least amount of inhibition as possible."[89] The plaintiff argued that the anonymity feature's purpose was to facilitate

78. *Id.* at 69–70.
79. *Id.*
80. *Id.* at 69.
81. *Id.* at 68 (citing Kimzey v. Yelp! Inc., 836 F.3d 1263, 1269 (9th Cir. 2016)) (taking action to display content and responsibility for what makes the content itself illegal).
82. *See generally* Gonzalez v. Google LLC, 2 F.4th 871 (9th Cir. 2021).
83. *Id.* at 881–82.
84. *Id.* at 888.
85. *Id.* at 896.
86. *Id.* (cleaned up).
87. Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1094 (9th Cir. 2019).
88. *Id*. at 1095.
89. *Id.*

conduct known to be illegal.[90] The plaintiff's son was directed to a supplier of heroin laced with fentanyl and died the next day.[91] The court held that certain algorithm functions, such as recommendations and notifications, were tools meant to facilitate communication and content of others, not content in and of itself.[92]

Across the appellate circuits, the body of cases focusing on a service provider's use of algorithms shows a trend in evolving case law narrowing the scope of the statute. In *Roommates*, the court targeted neutral and non-neutral tools in content development.[93] In *Force*, the issue turned towards material contribution and how material contribution might make the content itself unlawful.[94] The Second Circuit, in *Force*, analogized *Force* with *Roommates* in the context of algorithms, the development of content, and discrimination of protected classes regarding the specific and actual content developed.[95] While conceding caution against the overly broad analysis of § 230, the court in *Gonzalez* collapses both issues within its analysis and its application of *Dyroff*.[96] In order for *Gonzalez* to be distinguished from *Dyroff*, something more would be needed.[97] What the court meant by *more* was the focus of intense dissent and continues to be the ultimate issue.[98] *Gonzalez*'s holding that online neutral and non-neutral features and functions, including algorithms, as communicating content to others and not content itself, provided the final straw to break the camel's back. A significant legal and factual threshold was asserted regarding algorithms and their role across industry and legal disciplines.[99] The Supreme Court is aware of this assertion.[100] While §

90. *Id*.

91. *Id.*

92. Gonzalez v. Google LLC, 2 F.4th 871, 894 (9th Cir. 2021) (citing *Dyroff*, 934 F.3d at 1093).

93. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1166 (9th Cir. 2008) (en banc) (discussing use of illegal conduct and discriminatory questions as a condition in conducting business).

94. Force v. Facebook, Inc., 934 F.3d 53, 68 (2d Cir. 2019) (citing Kimzey v. Yelp, 836 F.3d 1263 (9th Cir. 2016)) (taking action to display content and responsibility for what makes the content itself illegal).

95. *Id.* at 69 (describing how the 9th Circuit addressed the Fair Housing Act, 42 U.S.C. § 3601 prohibitions on discrimination on the basis of a protected class in activities related to housing—Roommates's websites requiring users to use pre-populated responses to answer inherently discriminatory questions amounted to developing the actionable information for purposes of the plaintiff's discrimination claim).

96. *See Gonzalez*, 2 F.4th at 894–95 (citing Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093 (9th Cir. 2019)).

97. *Id.* at 895.

98. *Id.* at 896–97.

99. *See id.*

100. *See id.* at 925 n.9 (Gould, J., concurring in part and dissenting in part) (citing Malwarebytes, Inc. v. Enigma Software Grp. USA, 946 F.3d 1040, 1053 (9th Cir. 2019), *cert.*

230 states what an information content provider ("ICP"), means, "any person or entity responsible, in whole or in part, for the creation or development of information provided through the internet,"[101] substantial uncertainty remains. The uncertainty with non-neutral tools or machine learning algorithms used on the Internet lies in the creation and development of information.[102] Thus, liability is extinguished even when an algorithm created by the company distributes content it knows or reasonably should know is illegal.[103]

B. Algorithm Creation, Collaboration, and Content Development

In *Gonzalez v. Google*, and undoubtedly in future cases to come before the Supreme Court of the United States, a fact-based analysis of algorithms is required. First, addressing how the algorithm is created and data content is developed should be considered. The examination then turns to how the content-created functions as the product. Algorithms use data for mining, exploitation, hyper-personalized profiling, and discriminatory targeting for a specific purpose.[104] This purpose is the basis of content development and creation.[105] The intent of § 230(f)(3), following the policy provision of section (b), was to address this result.[106] The analysis will show how algorithms function as content developers, sometimes inherently, and thus immunity does not apply.[107]

1. Algorithm Creation

Algorithms are step-by-step mathematical procedures for solving a problem or accomplishing an end.[108] Algorithms are carefully crafted and intentionally built, designed with specific objectives and structured decision-

---

denied, 141 S. Ct. 13 (2020) (statement of Thomas, J., respecting the denial of certiorari)) (arguing that the courts have construed § 230, overly broad, outside of Congressional intent).

101. *See* Chaney et al., *supra* note 16.

102. Catherine Tremble, *Wild Westworld: Section 230 of the CDA and Social Networks' Use of Machine-Learning Algorithms,* 86 FORDHAM L. REV. 825, 837, 854–55 (2017).

103. *See* Chaney et al., *supra* note 16.

104. *See* Leetaru, *supra* note 24; *see also* Gal & Petit, *supra* note 25, at 636–637; Chaney et al.*, supra* note 16.

105. *See* Gal & Petit, *supra* note 25; *see also* Brandon W. Jackson, *Artificial Intelligence and the Fog of Innovation: A Deep-Dive on Governance and the Liability of Autonomous Systems*, 35 SANTA CLARA HIGH TECH. L.J. 35, 42 (2019); *see also* Chris Brummer & Yesha Yadav, *Fintech and the Innovation Trilemma*, 107 GEO. L. J. 235, 267, 271 (2019).

106. 47 U.S.C. § 230(b)(1)–(5), (f)(3).

107. Force v. Facebook, Inc.*,* 934 F.3d 53, 83–85 (2d Cir. 2019) (Katzmann, C.J., concurring in part and dissenting in part); *see also Gonzalez*, 2 F.4th at 914–15 (Berzon, J., concurring).

108. *Algorithm*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/algorithm (last visited Sept. 13, 2023).

making processes.[109] Algorithms, including self-teaching machine learning algorithms take labor, engineering, and a significant level of effort to create.[110] Algorithm source code is unique, and companies vigorously protect their proprietary algorithms.[111] Algorithms are automated computational creations generating procedures for decisional outcomes based on data inputs.[112] The content that drives the algorithms is controlled by data parameters, which in some cases may redirect the algorithm, that serves as the basis for machine learning.[113] Algorithms can be designed to set or redefine their decision-making parameters based on the data input and the decision-making criteria in which they are coded to perform.[114] These are commonly known as "learning algorithms".[115] Thus, an algorithm's knowledge of consumers, the mosaic or tapestry of data communicating time, place, and manner, and even intent of actions, determines its function.[116] Data content is the raw material which feeds the algorithm; without data content, the algorithm is meaningless.[117] Data content—regarding consumer buying habits, interests, dislikes, and even hatred—is the content that the algorithm ingests.[118] The algorithm sees and develops what works, depending on the criteria or bias, and then exploits that relationship.[119]

Algorithms are products designed for a specific purpose on an Internet platform. The purpose of the algorithm, depending on the digital platform, is to exploit its knowledge of consumers to extract profit.[120] Depending on the data set, which in the cases of *Gonzalez* and *Force* consisted of billions of

109. *See* Gal & Petit, *supra* note 25, at 636–37.

110. *Machine Learning*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/machine%20learning (last visited Sept. 13, 2023). "Machine learning" is defined as "the process by which a computer can improve its own performance (as in analyzing image files) by continuously incorporating new data into an existing statistical mode." *Id.*

111. *See* Katherine B. Forrest, *AI and Algorithmic Bias—Seeking Disclosure of the Algorithm*, *in* N.Y. PRAC., COM. LITIG. IN N.Y. STATE COURTS § 79:13 (5th ed. 2022).

112. *See* Gal & Petit, *supra* note 25, at 637–38.

113. *See* Chaney et al., *supra* note 16.

114. *See* Gal & Petit, *supra* note 25, at 637–38.

115. *Id.* at 673–38.

116. *See* Tremble, *supra* note 102, at 839.

117. *See* Gal & Petit, *supra* note 25, at 636–37.

118. IAN COCKBURN ET AL., THE ECONOMICS OF ARTIFICIAL INTELLIGENCE: AN AGENDA 127 (2019). The focus here is regarding deep learning algorithms in the public domain and the data pools that are essential to generating predictions. In some cases, this can be minute-by-minute knowledge in other cases it may be a general pattern of life knowledge. *Id.*

119. *Id.* at 146.

120. *See* STIGLER CTR. FOR THE STUDY OF THE ECON. & THE STATE, STIGLER COMMITTEE ON DIGITAL PLATFORMS FINAL REPORT 58 (2019), https://www.chicagobooth.edu/-/media/research/stigler/pdfs/digital-platforms---committee-report---stigler-center.pdf. Note generally, behavioral economics helps improve our understanding of real consumer choices and suggests that consumer exploitation is common. There are several systematic consumer biases that, when incorporated into economic analysis, affect outcomes and welfare. *Id.*

users, the bias or learning algorithm often directly affects behavior—generating massive profits. The provider takes the content of users and data and applies a specifically created product—the algorithm—to the data, creating a continual loop of content development, the sum of which is its own unique content beyond the individual parts.[121] Defining when exactly the threshold of content creation outside of immunity is reached is a question that must consider how content development unfolds.

2. *Basis for Content: Creation and Collaboration of Non/Neutral Algorithms*

When data is obtained, the courts focus on the importance of how it is obtained.[122] If it is neutral, or untainted, there is greater deference towards immunity from liability associated with its use. However, if the service provider filters or uses a notification system that directs communication of users according to discriminatory criteria, the service provider is no longer neutral and relinquishes immunity.[123] In *Roomates.com*, defendant Roommates argued that it was not responsible for the information on the user profile page because the data required the subscriber's choices and decisions leading to the publication of each profile.[124] Because § 230 states that "information content provider" means anyone responsible, even in part, for the creation and development of information provided through the internet, Roommates's non-neutral, discriminatory practices stripped them of immunity.[125] The court qualified Roommates's involvement as a collaborative effort between Roommates and the subscriber.[126] Thus, the Ninth Circuit held non-neutral collaboration to be content provider development, as envisioned with § 230(f)(3). The holding in *Roommates* set the standard that if collaboration exists and it develops illegal content, then liability may attach.[127]

To determine the performance of neutrality in the algorithm, decision parameters, data being mined and ingested, predictive purposes, and statistical patterns must be identified, and two primary factors need to be

---

121. *See* Force v. Facebook, Inc., 934 F.3d 53, 83 (2d Cir. 2019) (Katzmann, C.J., concurring in part and dissenting in part) (commenting on the cumulative effect of recommending several friends, several groups, or events has an impact greater than the sum of each suggestion; the process envelops the user, immersing her in an entire universe filled with people, ideas, and events she may never have discovered on her own.).

122. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1169 (9th Cir. 2008) (en banc).

123. *Id.*

124. *Id.* at 1167.

125. *Id.*

126. *Id.* at 1167–70.

127. *Id.* at 1165, 1175.

considered.[128] First is the quality of the data input—the essential raw information.[129] In *Roommates*, plaintiffs alleged that this data was based on discriminatory preferences of sex, sexual orientation, and the presence of children, in violation of the Fair Housing Act.[130] Second is the quality of decisional outcomes produced from the raw data filtered by the algorithm.[131] Here, qualitative success or failure is measured by the algorithm's ability to consume the initial data and form a rational basis for decisional outcomes, focusing on patterns and correlations identified from the ingested data.[132] Lacking a better analogy with *Roommates*, discrimination in–discrimination out, is the outcome. Thus, the basis of the data creating the algorithm's decisional outcomes is critical to defining content. How the data set is discriminately filtered matters regarding content creation.

If the data set is, for example, one million accounts and the algorithm parameters instruct it to hunt for common postings showing, discussing, or advocating violence, one parameter targeted may be the phrase "Hang Mike Pence." If the objective is to increase revenue by engaging more consumers targeted with directed advertising, then the outcome of increased engagement is driven by the increase in the objective, advertising revenue. Thus, if the algorithm, by collaborating with data, is driven by decisional outcomes related to advertising revenue correlated to the data, then the algorithm has a particular objective. The objective is non-neutral, because the parameters defining the objective are pre-determined. The algorithm ingests and filters the data, and it targets content based upon user bias and predetermined criteria. The end results are obtaining more advertising revenue.[133] Thus, in the examples, from cryptocurrency to terrorism, content data as property affects algorithmic performance.[134] The algorithm uses data content to seek out words, pictures, and statements that carry specific meanings based on the algorithm's instruction to search out similar correlated patterns of bias, even anger and hate.[135] Absent the data content, the algorithm and its existential collaborative and developmental properties sit on the shelf.

---

128. *See* Gal & Petit, *supra* note 25, at 637–38.

129. *Id.* at 637; *Roommates.com*, 521 F.3d at 1166.

130. *See Roommates.com*, 521 F.3d at 1175 (identifying discriminatory data requirements); *see also* Fair Housing Act, 42 U.S.C. § 3604.

131. Gal & Petit, *supra* note 25, at 635–38.

132. *Id.*

133. *Id.* at 637–38; *see also* JAMES MANYIKA ET AL., BIG DATA: THE NEXT FRONTIER FOR INNOVATION, COMPETITION, AND PRODUCTIVITY 5, 18, 28–29 (2011).

134. *Id.*; *see* Leetaru, *supra* note 24.

135. *See id.* at 636–38; *see also* TIMOTHY HINRICHS ET AL., TRANSFER LEARNING LEVEL DEFINITIONS 1 (Stanford Logic Grp. Tech. Rep. 2007). The algorithmic process occurs across internet networks. Initially, we saw how the SEC applied statutes to mitigate a form of bias. The SEC held Ms. Kardasian accountable for failing to disclose conflicts with her underlying

a. Algorithms learn

As identified earlier, content data is filled with discriminatory factors, invidious or otherwise, often within re-sequenced recommendations called feedback loop algorithms, which "fine-tune their decisional parameters."[136] Algorithms learn. Algorithm "transfer learning" leverages prior knowledge and data to improve learning when presented with new data and information related to the original set of data and information.[137] Algorithm learning is considered the development of information within the scope of § 230(f)(3) and, thus, part of an information content provider offering no protection under the statute's immunity provision.

Additionally, because algorithms use data content to develop, learn, and achieve a specific objective, neutrality becomes problematic. Put more unnervingly, algorithms are self-teaching.[138] The result is that the industry is "raising a generation of algorithms . . . that don't really learn the material, but they do well on the test."[139] In an effort to create "self-supervised learning" algorithms, they began ingesting massive amounts of raw human audio and visual data which shows a closer correspondence to human brain function.[140] Thus, the so-called "neutrality" of the algorithm is controlled by a set of predicates that may inherently be part of the discriminatory self-supervised learning algorithms. When the algorithm itself incorporates the worldview of the one programming the algorithm or the data, bias is inherent.[141]

Neutrality becomes even more opaque if the users on the platform are not even human, but automated bots.[142] Thus, automated bots feed the

---

motivations for expressing her bias. The party profited from the bias while others may have been induced to take risks causing financial injury because of it. *Id.*

136. *See* Gal & Petit, *supra* note 25, at 637–38.

137. Lilyana Mihalkova et al., *Mapping and Revising Markov Logic Networks for Transfer Learning*, DEP'T. OF COMPUT. SC. U. TEX. (July 2007), https://www.cs.utexas.edu/users/ml/papers/mihalkova-aaai07.pdf.

138. *Id.*; *see* Jackson, *supra* note 105, at 38–39; *see also* Brummer & Yadav, *supra* note 105, at 270–71.

139. *See* Anil Ananthaswamy, *Self-Taught AI Shows Similarities to How the Brain Works*, QUANTA MAG. (Aug. 11, 2022), https://www.quantamagazine.org/self-taught-ai-shows-similarities-to-how-the-brain-works-20220811/.

140. *Id.*

141. *See* Jackson, *supra* note 105, at 42; *see* Adam Rogers, *Google's Search Algorithm Could Steal the Presidency*, WIRED (Aug. 6, 2015, 1:24 PM), https://www.wired.com/2015/08/googles-search-algorithm-steal-presidency/. Neutrality is ambiguous. "It's not really possible to have a completely neutral algorithm," says Jonathan Bright, a research fellow at the Oxford Internet Institute on election studies. *Id.*; *see also* Jason Tanz, *Soon We won't Program Computers. We'll Train them like Dogs*, WIRED (May 17, 2016, 6:50 AM), https://www.wired.com/2016/05/the-end-of-code/.

142. *Bot*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/bots (last visited Sept. 13, 2023). Bots are "computer program[s] that perform automatic repetitive tasks." *Id.*

algorithm content with terms or statements correlated to whatever terms generate the most advertising revenue, which is based on the number of times the post is viewed, forwarded, or deleted.[143] This process is then accelerated via algorithm feedback loops.[144] The court in *Roommates*, from this perspective, appears to have anticipated the issue. Roommates's presumed objective was to generate revenue, regardless of the invidious process, and it relied on the subscriber's content collaborating with the algorithm to achieve the result.[145]

3. Basis for Content: Personalized Content Development, Feedback Loops

Algorithms are products created to collaborate with user data on their own in a self-learning environment.[146] Algorithms are taught objectives, given direction, and then provided data on which to feed.[147] Engineers then instruct the algorithm with various decision parameters to "coach" itself based on the data content.[148] More data means more personalized content development and a lot of "coaching." For the algorithm to maximize its objective, for example, to generate interest in a topic to create revenue as argued in *Gonzalez*, *Roommates*, or *Force*, certain actions need to be fine-tuned to achieve maximum utility.[149] The data content is what algorithms collaborate with and develop as defined under § 230(f)(3) and exemplified in *Roommates.*

---

143. *See* Jared Schroeder, *Marketplace Theory in the Age of AI Communicators*, 17 FIRST AMEND. L. REV. 22, 28–32 (2019) (analyzing the use and impact of bots across industries); *see* Brummer &Yadav, *supra* note 105, at 279 (disinformation spread on Twitter and Facebook (Meta) can "impact allocation of capital"); *see also* Zakon, *supra* note 17, at 1113 (algorithm and machine learning impact on social media revenue, advertising, and user engagement).

144. *See* Gal & Petit, *supra* note 25, at 637–38. "[P]arameters used by the algorithm to make new predictions improve over time as the algorithm learns by analyzing the effects of its past predictions." *Id.*

145. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1161 (9th Cir. 2008) (en banc).

146. *See* COCKBURN ET AL., *supra* note 118, at 116, 118, 124; *see also* Tanz, *supra* note 141.

147. *See generally* Tanz, *supra* note 141.

148. *Id.* As stated by Tanz:

> In traditional programming, an engineer writes explicit, step-by-step instructions for the computer to follow. With machine learning, programmers don't encode computers with instructions. They train them. If you want to teach a neural network to recognize a cat, you don't tell it to look for whiskers, ears, fur, and eyes. You simply show it thousands and thousands of photos of cats, and eventually, it works things out. If it keeps misclassifying foxes as cats, you don't rewrite the code. You just keep coaching it.

Id.

149. *Id.*; Gal & Petit, *supra* note 25, at 639.

a. Personalized content development

Tuning or tweaking the self-learning algorithm enables very specific parameters to be included or excluded.[150] These "hyper-parameters" are enabled by the previous experience that the algorithm has with the data, deciding what to include and what to exclude.[151] The content reconstructs itself based on the expertise, knowledge, and direction of hyper-parameters.[152] These very specific parameters may be fine-tuned by the algorithm in real-time via transfer learning data sets.[153] As the algorithm consumes a vast ocean of raw data, content development is refined as directed by the compass of algorithm parameters. Thus, not only are the algorithms non-neutral, but they are also purely based on content development squarely within § 230(f)(3).[154]

Data content, like in *Roommates*, is data that the algorithm can use to recalibrate itself, further developing and building on the new content to achieve the objective, financial or otherwise.[155] This collaborative process is in part what guided the Ninth Circuit in denying § 230 immunity to Roommates.[156] Thus, personalized content development is conducted within the boundary of the Information Content Provider under § 230(f)(3) and not immune as a service provider.[157] The reason for its personalized content development is its value. Advertisers pay billions of dollars annually by using personalized content development to influence choice and decision-making.[158] In 2010, a mass social network study showed the power and influence of social media via content development.[159] An experiment involving sixty-one million users on Facebook resulted in an additional 340,000 voters turning out in the 2010 United States congressional elections.[160] Further, research estimated 280,000 people were "indirectly nudged to the polls" by seeing specific messages targeting

150. *See* Gal & Petit, *supra* note 25, at 638–39.
151. *Id.* at 639.
152. *Id.* at 639.
153. *See id.*; *see also* Mihalkova et al., *supra* note 137.
154. 47 U.S.C. § 230(f)(3) (defining an information content provider a person *or* entity that is responsible in whole or in part for the creation or development of information provided through the Internet or other interactive service).
155. *See* Brummer & Yadav, *supra* note 105, at 271–72, 274 (stating that an algorithm's use of a broad spectrum of data can help shape the course of decisions.); *see also* Zakon, *supra* note 17, at 1113 (stating that there is a machine learning effect on social media revenue, advertising, and user engagement).
156. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1167 (9th Cir. 2008) (en banc).
157. *Id.* at 1162–63, 1169.
158. *See* Zakon, *supra* note 17, at 1112.
159. *See* Zoe Corbyn, *Facebook Experiment Boosts US Voter Turnout*, NATURE (Sept. 12, 2012), http://www.nature.com/news/facebook-experiment-boosts-us-voter-turnout-1.11401.
160. *Id.*

them in their news feeds.[161] In this instance, the specific parameter targeting user content was a message sent to voters stating that "I voted" along with profile pictures of randomly selected Facebook friends who had clicked the "I voted" button.[162] Thus, raw data of who voted and who had friends who voted was distilled to target voters by using the friend feature. By adding raw content data with the friend feature for targeting, researchers determined that voters were impacted.[163]

b. Feedback loops modify data content in real-time

Personalized content development increases intensity with recommender algorithms.[164] Recommenders, as seen in *Roommates*, generally function in two ways. First, they can ask whether the content data, as in our example with EthereumMax crypto stock, is somehow similar to the content that the user liked or disliked.[165] This determination will be based on the mosaic of intimately personal data which the algorithm has previously obtained in a hyper-personalized context.[166] This may include everything from recent data searches to postings, and even geolocation.[167] Second, the algorithm may filter and develop content in a "collaborative filter system."[168] Here, though the algorithm is active, it is essentially idling with a set of predetermined recommendations in wait for data to be ingested. Idling means the search engine algorithm continues to function without any feedback through user actions.[169] For example, when you travel to a new area, the geographic options change the content feeding the algorithm.[170] The learning recommender algorithm

161. *Id.*
162. *Id.*
163. *Id.*
164. *See generally* Chaney et al., *supra* note 16.
165. Renee DiResta, *Up Next: A Better Recommendation System*, WIRED (Apr. 11, 2018, 11:00 AM), https://www.wired.com/story/creating-ethical-recommendation-engines/ ("Algorithms used by Facebook, YouTube, and other platforms keep us clicking. But those systems may promote misinformation, abuse, and polarization.").
166. Gal & Petit, *supra* note 25, at 639, 641; *see* Zakon, *supra* note 17, at 1129, 1138.
167. *See* Brummer & Yadav, *supra* note 105, at 271–72, 274 (stating that the algorithm uses a broad spectrum of personal data). Everything from punctuation in text messages to geolocation and shopping preferences may be exploited. Here, reference is made to the intentional "mosaic" of life created by technology. This resembles a perspective held by the Supreme Court in United States v. Jones, 565 U.S. 400 (2012). The Court held that such ubiquitous monitoring was an unconstitutional invasion of privacy constituting a violation of the Fourth Amendment protection against unreasonable search and seizure. *Id.* at 412. It is worth noting that Justice Scalia's majority opinion included a very broad range of judicial philosophy on the Court. *See e.g.*, *id.* at 404–10.
168. *See* DiResta, *supra* note 165.
169. *Id.*
170. *Id.*

will recalibrate based on a new set of data and collaborate with the data content, develop it for marketable consumption and profit. Here, the feedback loop plays an important role.

Feedback loops are a form of repetition, a re-sequence of recommendations.[171] Based on the mosaic of patterned behavior, the feedback loop will cull similar data from others with similar profiles and re-enforce targeted recommendations.[172] Here, the output of the algorithm becomes part of its input, thus, the algorithm is part of the content and develops the content directly.[173] Under § 230(f)(3), this is a function of information *content* providers. Reengaging the original metaphor, recommender algorithms and feedback loops become the eye of the tempest where similarly recommended and distilled data "feedback" is distilled and sent to users repeatedly.[174] Options become increasingly narrow, and user choices can be restricted to increasingly extreme content.[175]

Having shown how algorithms are created as products, their method of producing content, and their inherent collaboration with user data and content to develop information through the Internet, the elements needed for applying § 230(f)(3) are met.[176] Because the service provider is also functioning as a content provider and a content provider is responsible in whole or in part for the creation or development of the offending content, immunity does not apply.[177] The analysis will now turn to the standard of liability that must be shown regarding the algorithms as defective products.

III. ALGORITHMS AS DEFECTIVELY DESIGNED PRODUCTS

Algorithms, while not tangible in the traditional sense, are products. Algorithms are mathematical calculations on which decisions are based.[178] The

171. Gal & Petit, *supra* note 25, at 637; *see also* Chaney et al., *supra* note 16 (discussing the consequences of feedback loops).

172. *See* Chaney et al., *supra* note 16 (feedback loops increasing homogeneity and decreasing utility).

173. Swathi M. Sadagopan, *Feedback Loops and Echo Chambers: How Algorithms Amplify Viewpoints*, CONVERSATION (Feb. 4, 2019, 4:18 PM), http://theconversation.com/ feedback-loops-and-echo-chambers-how-algorithms-amplify-viewpoints-107935.

174. *See* Chaney et al., *supra* note 16. This leads to algorithmic confounding, loss of utility, and homogenizing user behavior. *See id.*

175. *Id.*

176. Force v. Facebook, Inc., 934 F.3d 53, 83–85 (2d Cir. 2019) (Katzmann, C.J., concurring in part and dissenting in part); *see also* Gonzalez v. Google LLC, 2 F.4th 871, 914–15 (9th Cir. 2021) (Bezon, J., concurring).

177. *Gonzalez*, 2 F.4th at 914 (citing Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1162 (9th Cir. 2008)).

178. *See* Gal & Petit, *supra* note 25, at 636–37; Mihalkova et al., *supra* note 137; Tanz, *supra* note 141.

functional components of the software used on the Internet are algorithms.[179] Absent the algorithm ingesting the raw data, liability is mitigated.

Because of feedback loops and machine learning, the product continually re-creates itself.[180] Recreation occurs based on parameters or "decisions" to improve predictions and performance based on data inputs.[181] The analysis regarding product liability must consider and balance multiple competing factors: (1) intended performance, and (2) the liability of foreseeable and probable risks, namely, conduct, associated with performance.

Piercing immunity by asserting algorithm product liability in § 230(f)(3) can achieve two important objectives: (1) not limiting First Amendment rights, while (2) the marketplace remains accountable for defective design and risk-utility. Post-*Zeran* and leading to *Gonzalez*, accountability has been the fundamental concern. How much risk should the consumer and the public accept and how valuable in utility is this product? Justice Thomas seems to wrestle with this issue in questioning the plausible scope of publication within § 230 when the product collaborates with the content.[182] By not confronting the core issue of algorithms functioning as products creating and developing content, the holding in *Gonzalez* leaves product liability unresolved. Algorithm products, such as artificial intelligence and machine learning algorithms, that are inherently creative and developmental in nature will continue to accelerate absent any reasonable guardrails for liability.

A. Algorithm Product Liability Defined

A claim of product liability must first establish that there is a product, either tangible or intangible.[183] Here, the algorithm is an intangible product. Second, it must be shown that the product has been brought to market by a commercial supplier in the business of selling or distributing the product, including the product into which the component is integrated.[184] Here, the algorithm is brought to market as an indefeasible part of the software or interface used by consumers on the Internet. Third, the product must be defective by manufacturing, design, or information.[185] When applying the third element, it is helpful to view the liability of defective algorithms as a Venn diagram.

---

179. *See* Zakon, *supra* note 17, at 1121–22, 1138; *see generally* MANYIKA ET AL., *supra* note 133, at 27 (discussing software and big data applications across industries).

180. *See* Tanz, *supra* note 141; *see* Gal & Petit, *supra* note 25.

181. *See* Gal & Petit, *supra* note 25, at 636–37.

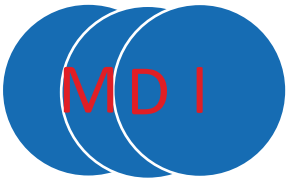
182. *See* Malwarebytes, Inc. v. Enigma Software Grp. USA, 946 F.3d 1040, 1053 (9th Cir. 2019), *cert. denied*, 141 S. Ct. 13 (2020) (statement of Thomas, J., respecting the denial of certiorari).

183. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 19(a) (AM. L. INST. 1998).

184. *See id.* §§ 1, 5.

185. *See id.* § 2.

While defective manufacturing, design, or information may all potentially establish an algorithm product liability claim, defective design is the most applicable due to the existential nature of how algorithms function.



186

Fourth, the product must reach the consumer without substantial change; the algorithm reaches the consumer without substantial change. Finally, causation, the algorithm causes harm to the user.[187] Algorithms are within the domain of products liability found in the Restatement (Third) of Torts §§ 2, 5, 19, and Restatement (Second) of Torts § 402A. [188]

1. Algorithms as the Product Sold

An algorithm is a set of calculations and numbers, devoid of meaning outside of its application as part of a software product, constructed for a specific purpose, the purpose being to ingest raw data, apply the algorithm as constructed to the raw data, and achieve an intended result.[189] The intended result is defined by the decision parameters which construct the algorithm.[190] Further, algorithms can refine their own decision parameters based on the raw data ingested in real-time.[191] Here, the machine learning algorithms facilitate artificial intelligence enabling the host to learn from the data it analyzes, absent explicitly being programmed.[192]

Consequently, algorithms create unique capabilities impacting both descriptive and predictive decision-making.[193] In practical terms, the razor-sharp chain is not exposed to liability with a chainsaw until the chain is put on the

186. The Venn diagram is important because it may provide a framework for future determinations on liability. *See infra* Section III.A.2.

187. RESTATEMENT (SECOND) OF TORTS § 402(a) (AM. LAW. INST. 1965).

188. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. §§ 2, 5, 19 (AM. L. INST. 1998); *see also* RESTATEMENT (SECOND) OF TORTS: PROD. LIAB. § 402(a) (AM. L. INST. 1965).

189. *See* Gal & Petit, *supra* note 25, at 636; *see also* THOMAS H. CORMEN ET AL., INTRODUCTION TO ALGORITHMS 5, 11–13 (3d ed. 2009).

190. *See* ORG. FOR ECON. CO-OPERATION & DEV., DATA-DRIVEN INNOVATION: BIG DATA FOR GROWTH AND WELL-BEING 152–58 (2015).

191. *Id.*

192. *Id.*; *see also* Gal & Petit, *supra* note 25.

193. *See* Gal & Petit, *supra* note 25.

saw. The chainsaw does not know what trees or how many trees will need to be harvested, absent the one controlling the machine. Likewise, in the forest of data that is the Internet, algorithms play a critical role when applied via software. As defined in the Restatement (Third) of Torts, Products Liability § 19, a product is tangible personal property distributed commercially for use or consumption. Other items are products when the context of their distribution and use is sufficiently analogous to the distribution and use of tangible personal property.[194]

When the algorithm is created and embedded in the software, it becomes part of the product. Algorithms are unique products in some cases. As used by Google and Meta (formerly Facebook), they are valuable and protected commodities.[195] Product liability applies when the algorithm is distributed commercially for use or consumption.[196] When considering strict liability within the scope of computer software created with algorithms, courts have considered the treatment of software under the Uniform Commercial Code and product liability law.[197] Under the Uniform Commercial Code, mass-marketed software is considered a good.[198] The Restatement on product liability notes the differences between tangible and intangible products and applies the rules based on the distribution of the product and how sufficiently similar and appropriate its use is compared to tangible personal property.[199] For example, under the Restatement, if one was to load a disk of Microsoft software, that would be a tangible product distributed for commercial use. If, however, instead of buying a Microsoft disk, a person digitally downloads the same data from the Microsoft website, that person would still obtain the same product in an intangible medium. Thus, the Restatement defers to a fact-based analysis.[200]

a. It is all about the intangibles

The Restatement addresses intangible property in product liability in two categories.[201] One category involves intangible harm-causing products such

194. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. §§ 1, 5 (AM. L. INST. 1998).

195. *See generally* IAN C. BALLON, UNIQUE INTELLECTUAL PROP. ISSUES IN SEARCH ENGINE MARKETING, OPTIMIZATION AND RELATED INDEXING, INFORMATION LOCATION TOOLS AND INTERNET AND SOCIAL MEDIA ADVERTISING PRACTICES, E-COMMERCE AND INTERNET LAW PART II INTELLECTUAL PROP. § 9 (2020).

196. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 19 cmt. d (AM. L. INST. 1998).

197. *See id.*

198. *See generally* Advent Sys. Ltd. v. Unisys Corp., 925 F.2d 670 (3d Cir. 1991) (applying Pennsylvania law).

199. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 19 cmt. b, d (AM. L. INST. 1998).

200. *Id.* § 19 cmt. d–f.

201. *Id.* § 19.

as electricity after passing through a meter.[202] The second category consists of information in media, specifically books, maps, and charts.[203] Within a fact-based analysis, algorithms combine determinative attributes of both categories.[204]

Regarding electricity, courts have consistently held that electricity becomes a product only when it passes through the customer's meter and enters the customer's premises.[205] Thus, until the product enters a medium used by the customer in the stream of commerce, it is not a product but a service.[206] For example, as in *Roommates*, the algorithm sitting by itself is void of liability, it is only when the customer engages the discriminatory algorithm that liability attaches.[207] Because the algorithm was infected with discriminatory data requirements on preferences of sex, sexual orientation, and the presence of children, it was a defective product violating the Fair Housing Act; consequently, § 230 immunity did not attach.[208] The intangible algorithm is analogous to electricity because, while the algorithm is a product before use, it has not yet been sold or otherwise distributed.[209] However, once the algorithm is put into the stream of commerce by entering the device, it may create or develop content, thereby losing immunity under § 230.[210] Like the chainsaw waiting to be used on a forest of trees, the algorithm awaits use by the end user's device.

The essential purpose of the algorithm is, as directed, to develop and make decisions based on raw content data collected.[211] This purpose is not realized until and upon the algorithm traveling to or from a user's mobile or fixed media access control ("MAC") address.[212] Each device, computer, server, switch, or cellular phone carries a unique MAC address, like a real property address.[213] Upon entry, like electricity to the real property address, algorithms travel along the Internet like an electrical current travels before entering the destination address. Some courts have reasoned that the electric

202. *Id.* § 19 cmt. d.

203. *Id.*

204. *Id.* Algorithms are defined as products within the context of their use and distribution sufficiently analogous to the use and distribution of tangible personal property as generally defined in the Restatement. *See id.* A product is personal tangible personal property distributed commercially for use or consumption. *See id.*

205. *E.g.*, Smith v. Home Light & Power Co., 695 P.2d 788, 789 (Colo. App. 1984).

206. RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 19 cmt. d (AM. L. INST. 1998).

207. Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1175 (9th Cir. 2008) (en banc).

208. *Id.* (identifying discriminatory data requirements).

209. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 19 cmt. d (AM. L. INST. 1998).

210. *See id.*

211. Gal and Petit, *supra* note 25, at 637.

212. Naeim Abedi et al., *Tracking Spatio-Temporal Movement of Human in Terms of Space Utilization Using Media-Access-Control Dat*a, 51 APPLIED GEO. 72–81 (2014).

213. *See id.* at 73.

power, though a product, is not recoverable in a product liability claim unless it is sold and passes through the customer meter.[214] However, an algorithm defeats this argument because the algorithm is engaged by the customer—a cyber invitee.[215] In this regard, algorithms are more robust than intangible electricity. Algorithms passing through the MAC address, acting as the meter, harvest raw content data.[216] The purpose of harvesting the data is to further develop content that can be monetized.[217] Electricity, however, does not offer engagement—it is binary.[218] Algorithms, alternatively, can be directed to control data content and who receives it.

In the second category of products liability for intangible products, plaintiffs allege product liability because the information in the medium of a book, map, or chart is defective.[219] In those instances, some courts have emphasized that some charts, specifically navigational, are used for their physical characteristics and not the ideas in them.[220] This analogy is similar to an algorithm. The algorithms are used for their numerical computational characteristics, to develop the content as applied to, wittingly or unwittingly in user terms and agreements The algorithm is void of liability, however, until it is engaged by the customer, as in *Roommates*. As analogized in product liability claims with navigational charts, courts have noted that a pilot's total reliance on navigational charts directly links the charts to accidents.[221] Similarly, reliance on

214. Monroe v. Savannah Elec. & Power Co., 471 S.E.2d 854, 857 (Ga.1996).

215. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 19 cmt. d (AM. L. INST. 1998).

216. *See* Leetaru, *supra* note 24.

217. *Id.*

218. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 19 cmt. d (AM. L. INST. 1998) (electricity as a commercially distributed intangible product for use becomes a product only when it enters the home). With electricity, the issue is pre- and post-delivery engagement with the commercially distributed product. *See id.* The commercial product is either delivered or not, thus on or off. With electricity we can see the lights go on and off when we flip the switch. However, algorithms are far more ubiquitous. The commercially distributed algorithm is often more directly controlled, filtered, and targeted towards end users—even absent the consumer's control or knowledge. The specific capability is used to harvest and monetize data. Thus, the algorithm as a commercially used and distributed product is more tangible than electricity because it is on all the time and everywhere, commercially used for harvesting and collecting data. *See id.*

219. *Id.*

220. *Id.*; *see* Zakon, *supra* note 17, at 1124–25 (quoting Winter v. G.P. Putnam's Sons, 938 F.2d 1033 (9th Cir. 1991)).

221. *See* Brocklesby v. United States, 767 F.2d 1288, 1298 (9th Cir.1985); *see also* BALLON, *supra* note 195, § 9.01 ("The ways in which information is accessed and distributed online—and the tools used by companies and individuals to lure people to their sites or divert them away from other locations—implicate intellectual property law questions. . . . Concerns about requiring users to view legal terms and conditions or other information today are addressed by using pop up windows or click-through screens.")

algorithms is more deceptive and may occur without the user ever realizing it exists.[222] This is further explained in Section IV.[223]

2. Defective Condition—The Whole is Greater Than the Sum of the Parts

Having defined an algorithm as an intangible good, liability attaches for the harm caused by a product into which the component is integrated if the component is defective and the defect causes harm, or the seller or distributor of the component substantially participates in the integration of the component into the design of the product.[224]

In *Roommates*, the defective condition occurred when the non-neutral content was added to the algorithm.[225] It was Roommates's addition to the content, plaintiffs alleged, which triggered the discrimination breach of the Fair Housing Act.[226] In *Erie Insurance Company v. Amazon.com*, *Inc.*,[227] the Fourth Circuit provided some clarity on defective condition, but only as it related to immunity under § 230, not regarding the algorithm as a product.[228] In a defective condition claim, where Amazon resold a defective product, the Fourth Circuit held that § 230(c)(1) protected internet intermediaries in the online publication of a third-party's information.[229] Amazon escaped product liability under Maryland law because it was not the seller of the product in a defective condition.[230] The importance of *Erie* is that the court limited immunity as it relates to internet intermediaries while simultaneously distinguishing Amazon's role as a product reseller.[231] The Fourth Circuit specifically held that, had Amazon transferred ownership of property for a price, liability as a seller for a defective product would have attached.[232] The possessory interest of the party was a key factor.[233]

---

222. Matthew Stewart, *The Limitations of Machine Learning*, MEDIUM (July 29. 2019), https://towardsdatascience.com/the-limitations-of-machine-learning-a00e0c3040c6#:~:text=This%20is%20the%20most%20obvious,and%20lack%20of%20good%20data.&text=Many%20machine%20learning%20algorithms%20require,begin%20to%20give%20useful%20results.

223. *See infra* Section IV.

224. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. §§ 1, 5 (AM. L. INST. 1998).

225. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157 (9th Cir. 2008) (en banc).

226. *See id.*

227. 925 F.3d 135 (4th Cir. 2019).

228. *See id.*at 139–40. Section 230(c)(1) is not a blanket shield from immunity. *See id.*

229. *Id.* at 138.

230. *Id.* at 137–38.

231. *Id.* at 141–42.

232. *Erie Ins. Co.*, 925 F.3d at 144.

233. *Id.* at 141 (distinguishing liability for defective products on sellers and manufacturers and imposed on owners of personal property). The court explained that a party who transfers

For an algorithm, when user data is one's personal property that is transferred and exploited for profit on the platform, the logical result would be that a product developed and constructed on unlawful user content exposes it to liability.[234] Artificial intelligence algorithms, for example, absorb and repurpose a multiplicity of raw data, including things such as one's spending habits, grammar usage, hobbies, and shopping preferences.[235] When an algorithm takes data, which is another's property, and repackages it for distribution as a marketable commodity, the algorithm has transferred that property.[236] In the ocean of internet data, mined with perilous icebergs, roadways of driverless cars, and ChatGPT, algorithms act as navigational charts. Ample evidence supports this assertion through the collaborative content valuation of algorithms, estimated at over fifty billion dollars a year in advertising and marketing alone.[237]

The boundaries set by *Roommates* and *Erie* represent two ends of the spectrum for defective product liability, as applied to § 230. In *Roommates*, adding content to the algorithm directly affected the behavior and resulted in liability.[238] In *Erie*, the Fourth Circuit held, "[t]o be sure" Amazon selling its own goods on its website carries the burden and liability of a seller for a defective product.[239] Algorithms that develop and construct data content generating advertising revenue are the intangible goods sold by Google, Meta, and Twitter to willing buyers of the data content developed.[240] The value of the algorithms as a product is their ability to generate revenue.[241] Revenue is existentially reliant upon content development; the basis of § 230(f)(3) as an information content provider is content development.[242]

---

title (having a possessory interest) in a bargained for exchange price is a seller. *Id.* Those who do not take title to the property during distribution are not sellers. *Id.* Those who render services for distribution or sale are not sellers. *Id.*

234. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1167 n.20 (9th Cir. 2008) (en banc) (Roommates's commercially used and developed an algorithm decoding subscriber input and transforming it into a profile page, which plaintiffs alleged violated the Fair Housing Act). Roommates's use of the algorithm directly participated in the development of the alleged illegality. *Id.*

235. *See* Gal & Petit, *supra* note 25, at 637; *see also* Brummer & Yadav, *supra* note 105, at 234–74.

236. *See* Leetaru, *supra* note 24; *see also* Gal & Petit, *supra* note 25, at 638.

237. *See* Leetaru, *supra* note 24.

238. *Roommates.com*, 521 F.3d at 1167.

239. Erie Ins. Co. v. Amazon.com, Inc., 925 F.3d 135, 144 (4th Cir. 2019).

240. *See* Leetaru, *supra* note 24; *see* Tremble, *supra* note 102, at 838–841.

241. *See* Leetaru, *supra* note 24.

242. 47 U.S.C. § 230(f)(3) (The term "information content provider" means any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service."); *see* Leetaru, *supra* note 24; *see* Tremble, *supra* note 102, at 838–841.

Intangible algorithms can be defective overtly, as in *Roommates*, or more subtly, as found in the *Gonzalez* holding where the causal connection to material contribution is less clear.[243] What is known is that algorithms generate and impact viewer frequency, purchasing choices, and target decision-making.[244] Machine learning and recommender algorithms have a verifiable impact on individual decisions.[245] From regulatory holdings by the Securities and Exchange Commission to civil tort litigation, the functional impact of algorithms is palpable.[246] Algorithms "can influence how users perceive the world by filtering access to media, pushing political dialogue towards extremes, or filtering out contrary opinions."[247]

The defective condition is clear when algorithm systems impact crucial decision-making processes. But algorithms have an impact across a wide spectrum of human activity. For example, determining who is marketed to for loan approvals, who is subject to criminal profiling, and who may receive medical intervention for an illness.[248] Justice Thomas's comments in denial of certiorari in *Malwarebytes, Inc. v. Enigma Software Group USA*,[249] seem to have been well founded regarding immunity under § 230 for "content" known to be illegal.[250] The facts revealing the defective condition of the intangible product, the algorithm, are more clearly defined in some cases than in others. Regardless, showing that the defective condition exists is paramount. This is achieved in three ways.

a. Defective condition: algorithm manufacturing, design, inadequate warning

A product's defective condition and resulting liability are generally discovered in one of three ways when, at the time of sale or distribution, it (1) contains a manufacturing defect, (2) is defective in design, or (3) is defective

243. *See Roommates.*com, 521 F.3d at 1167; *See* Gonzalez v. Google LLC, 2 F.4th 871, 894–95 (9th Cir. 2021) (finding that Google's recommender algorithm did not treat terrorist created content differently than other third-party created content and thus immune from liability under § 230).

244. *See* Leetaru, *supra* note 24; *see also* Chaney et al., *supra* note 16; COCKBURN ET AL., *supra* note 118, at 126; STIGLER COMMITTEE ON DIGITAL PLATFORMS FINANCIAL REPORT, *supra* note 120, at 58.

245. *See* Chaney et al., *supra* note 16.

246. P.J. Tobia et al., *Why Kicking Alex Jones Off Social Media is not Legally Censorship*, PBS (Aug. 8, 2018 6:30 PM), https://www.pbs.org/newshour/show/why-kicking-alex-jones-off-social-media-is-not-legally-censorship; *see also* Geol, *supra* note 1.

247. *See* Chaney et al., *supra* note 16 (commenting on the loss of utility).

248. *Id.*

249. 946 F.3d 1040, 1053 (9th Cir. 2019), *cert. denied*, 141 S. Ct. 13 (2020) (statement of Thomas, J., respecting the denial of certiorari).

250. *See* Gonzalez v. Google LLC, 2 F.4th 871, 926 n.9 (9th Cir. 2021) (Gould, J., concurring in part and dissenting in part).

because of inadequate instructions or warnings.[251] As we will see through case law, and as identified below, there may be overlap among the three defective conditions.

(1) Manufacturing defects focus on a single product unit.[252] In *Roommates*, for example, the algorithm would be manufactured defectively based on the discriminatory parameters of the algorithm which plaintiffs alleged violated the Fair Housing Act.[253]

(2) Defective design is far broader, condemning the entire product line.[254] In some commercial markets, this may include bias which may be misleading.[255] For example, loan and credit markets create significant implications for machine learning and artificial intelligence as raw data content may be false, dangerous, and intentionally misleading.[256]

(3) Defective conditions regarding inadequate warnings "arise when the specific product unit conforms to the intended design but the intended design itself, or its sale without adequate instructions or warnings, renders the product not reasonably safe," and, thus, "every unit in the same product line is potentially defective."[257] The speed filter feature in Snapchat's algorithm demonstrates a good example of how there may be overlap between defective design and inadequate warning.[258]

i. Manufacturing defect

Manufacturing defect claims turn on whether the product departs from its originally intended design by the seller.[259] For commercial entities on the Internet, algorithms have a specific intended design with instructions on what decisions to make and when to make them regarding the data with which they

---

251. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 (AM. L. INST. 1998).

252. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 3(a)–(b) (AM. L. INST. 1998).

253. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1165–67, 1175 (9th Cir. 2008) (en banc).

254. OWEN & DAVIS ON PROD. LIAB. § 8:1 (4th ed. 2023).

255. *See* HINRICHS ET AL., *supra* note 135, at 1.

256. *See* Brummer & Yadav, *supra* note 105, at 274 –76. AI and machine-learning algorithms raise dangers regarding the proper workings of algorithms relying on input of clear, correct, and codable data. *Id.* When algorithms are fed data from informational sources (like alternative data) that are ambiguous, falsified, or overly noisy, algorithm output will be tainted by error and thus unreliable. *Id.* Moreover, AI and machine learning automation means that the impact of such misfiring can spread exponentially fast as AI and machine learning algorithms respond automatically to new information. *Id.*

257. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 (AM. L. INST. 1998).

258. Maynard v. Snapchat, Inc., 870 S.E.2d 739, 743–44 (Ga. 2022) (holding that plaintiff adequately alleged that the manufacturer owed the driver a design duty with respect to particular risk of harm, and there is no "blanket exception to a manufacturer's design duty in all cases of intentional or tortious third-party product misuse.").

259. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(a) (AM. L. INST. 1998).

are presented.[260] Data algorithms are presented with the content of a medium—digital communications, written, visual, or otherwise. Data is content acquired by the algorithm which then decides on how to create and develop the information and how to use the content.[261] Development and creation of content is the basis of § 230(f)(3) precluding immunity.[262] To determine if an algorithm is functioning as intended, a party must be able observe the parts of the source code parameters which create the algorithm—in other words, map the source code.[263] Thus, the only way to know if the algorithm is acting as it was designed is to open the product up and examine the parts from which it is constructed.[264] Here, like a defective chain saw or an automobile, the algorithm is like any other product. As identified in Sections II and III, instructions include parameters directing AI, machine learning, feedback loops, and recommendations that lead to causation, discussed in Section IV.[265] A metaphorical interrogatory of the algorithm is required. From a manufacturing defect theory of strict liability, because an algorithm's decisional parameter is instructed to consume certain data, its actions are binary based upon an instruction which may be biased or even unlawful.[266] When the basis of the algorithm's decisions is known, in other words, upon an examination of the products parts as manufactured and designed, a claim can be fairly adjudicated.

A critical issue presents another gray area. If the algorithm is machine learning, making decisions, and redefining itself every minute of every day as it feeds on billions of pieces of data, then the algorithm is creating, developing, and functioning as its designers intended.[267] The algorithm is manufactured to change, adjust, and adapt, like a human brain.[268] This leads to a defective design analysis under the Third Restatement of Torts.[269]

ii. Defective design

Defective design exists when the foreseeable risk of the harm posed by an algorithm could have been reduced or avoided by reasonable alternative

260. *See* Gal & Petit, *supra* note 25 (discussing algorithms); Mihalkova et al., *supra* note 137; Ananthaswamy, *supra* note 139; Tanz *supra* note 141 (discussing exploitation of algorithms).

261. *See* Gal & Petit, *supra* note 25.

262. *See supra* Section II.A.

263. *See* Mihalkova et al., *supra* note 137.

264. *See* Chaney et al., *supra* note 16.

265. *See infra* Section IV.

266. *See* Jackson, *supra* note 105, at 42 (discussing subjective bias in algorithm creation); s*ee* also Chaney et al., *supra* note 16; *see, e.g.*, Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1165–67, 1175 (9th Cir. 2008) (en banc).

267. *See* Ananthaswamy, *supra* note 139.

268. *Id.*

269. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(b) (AM. L. INST. 1998).

design and omitting the alternative design renders the product not reasonably safe.[270] Because internet algorithms are using data to develop content as intended in § 230(f)(3),[271] the analysis must consider not merely the data, but its development of data.[272] The balance of First Amendment freedoms is critical. Indeed, people may say what they want, regardless of how offensive it may be.[273] The question for internet technology companies is whether the algorithms as constructed create a foreseeable risk in their development, construction, and exploitation of information.[274]

Interaction models with empirical data show the foreseeable consequences of feedback loop algorithms.[275] Algorithms, taking user data and content, ingesting it by machine learning or into a feedback loop, create recommendation effects on choice and decision making.[276] While the empirical data shows there are foreseeable negative risks, foreseeability may not be specific to the facts in any set of events that may come before the law.[277]

In *Roommates*, the foreseeable consequences of the algorithm were clear. But there are important differences in each case regarding causation.[278] In *Maynard v. Snapchat*, the Georgia Supreme Court made an important holding regarding whether a manufacturer owes a decisional design duty regarding a particular risk of harm posed by a product.[279] *Maynard* addressed product liability associated with the "speed filter" feature within Snapchat's mobile phone application.[280] The social media application recorded real-life speed on a photo or video that users could then share with other Snapchat users.[281] While using the social media application at over one hundred miles per hour in her vehicle, a user was involved in a car accident that severely injured the plaintiff.[282] In *Maynard*, the question turned on whether the risk

---

270. *See id.*

271. 47 U.S.C. § 230(f)(3).

272. *See* Leetaru, *supra* note 24 (describing the use and development of the monetization of data).

273. *See* Schroeder, *supra* note 143, at 24–25.

274. *See e.g.*, Maynard v. Snapchat, Inc., 870 S.E.2d 739, 743–44, 747 (Ga. 2022) (social media providers may breach duty to exercise reasonable care in design).

275. *See* Chaney et al., *supra* note 16 (discussing consequences of the feedback loop); DiResta, *supra* note 165; *see* Ananthaswamy, *supra* note 139.

276. *See* Chaney et al., *supra* note 16 (discussing causality in recommendation systems).

277. *See id.* (discussing consequences of the feedback loop).

278. *See Maynard*, 870 S.E.2d at 743–44, 747 (the speed filter, designed for the application, created a foreseeable risk to the user and manufacturer had duty of reasonable care in selecting alternative designs); *see also* A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814, 819 n.2 (D. Or. 2022) (describing how the anonymity feature was reasonably related to causation).

279. *See Maynard*, 870 S.E.2d at 743–44.

280. *Id.* at 743.

281. *Id.*

282. *Id.*

posed by the "speed filter" was reasonably foreseeable.[283] The court held that social media providers, as manufacturers, may breach their duty to exercise reasonable care in design when the manufacturer fails to adopt a reasonable, safer design that would have reduced the foreseeable risk of harm presented by the product.[284]

Reasonable design in algorithms requires more than just managing the content, which is then developed, it requires regulatory or corporate oversight regarding risk-limiting safeguards on the AI, machine learning, and feedback loop algorithms that carry foreseeable risk.[285] *Maynard*, like *Roommates*, is an example of a foreseeable risk that creates the necessity of adopting a reasonable alternative design.[286] In both cases, there is some interaction by the user collaborating substantively with the social media product.[287] In *Roommates*, the service provider created the drop-down boxes for customer data, subsequently manipulating and using the data, which the plaintiffs alleged violated the Fair Housing Act.[288] In *Maynard*, the "speed filter" software application was created by the service provider and used by customers.[289] In these cases, the service provider's design created a foreseeable risk of harm; in this way, algorithms may have a defective design that creates a foreseeable risk before use. This has posed a problematic issue in judicial decisions from *Roommates* to *Gonzalez*.[290]

In May 2021, the Ninth Circuit held that "[t]he duty to design a reasonably safe product is independent of Snap's [the social media service provider's] role in monitoring or publishing third-party content."[291] *Lemmon v. Snap, Inc.* was another social media speed filter accident case.[292] Agreeing specifically on the issue with the Georgia Supreme Court, the Ninth Circuit reasoned that while publishing content was "a but-for cause of just about everything," it did not mean the plaintiff's claim sought to hold the defendant

283. *Id.* at 746.

284. *Id.* at 747. Additionally, the policy considerations are significant and were not ignored in the holding. *Id.* at 754–56. Notwithstanding policy concerns, some products carry higher consequential value than others. *Id.*

285. *See* Chaney et al., *supra* note 16 (discussing causation and consequences of feedback loop algorithms); *see also* Schroeder, *supra* note 143, at 33, 54 (discussing algorithm use and control over the marketplace of ideas on the Internet).

286. *Maynard*, 870 S.E.2d at 745–46; *see* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1175 (9th Cir. 2008) (en banc).

287. *Maynard*, 870 S.E.2d at 745–46; *see Roomates.com*, 521 F.3d at 1175.

288. *Roommates.com*, 521 F.3d at 1165, 1169.

289. *Maynard*, 313 Ga. at 544–45.

290. *See, e.g.*, *Roomates.com*, 521 F.3d at 1175; Gonzalez v. Google LLC, 2 F.4th 871, 923–26 (9th Cir. 2021) (Gould, J., concurring in part and dissenting in part).

291. Lemmon v. Snap, Inc., 995 F.3d 1085, 1093 (9th Cir. 2021). The court emphasized *Maynard*, holding immunity unavailable. *Id.*

292. *Id.* at 1088–90.

responsible as a "publisher or speaker."[293] The Ninth Circuit held that, because the plaintiff's claim did not seek to hold the media service provider responsible as a publisher, but only liable for its own conduct, § 230 immunity did not apply to the creation of the speed filter.[294] The duty to design a reasonably safe product is independent of the social media service provider's role in monitoring or publishing third-party content.[295]

This analysis was applied to internet service providers more directly in *A.M. v. Omegle.com*.[296] Omegle, a free online chat room, paired the plaintiff with a man in his late thirties who then forced her to send pornographic images and videos of herself to him, perform for him and his friends, and recruit additional minors for abuse.[297] The predator threatened the juvenile with the release of the videos and pictures.[298] The court summarized the reasoning from *Lemmon*, that what mattered in *Lemmon*, for resolving the issue of § 230 liability, was the interplay between the speed filter and the reward system of the application.[299] Reward functions were thus shown to have a causal nexus within recommendation algorithm systems.[300] Likewise, in *Omegle*, the design of the chatroom was allegedly defective, and the plaintiffs alleged that the defective design led to the interaction between the eleven-year-old girl and a sexual predator.[301]

In *Omegle*, § 230 immunity for third-party content communication was at issue.[302] The court, holding that § 230 immunity did not apply, pointed out that the Plaintiff's contention was that the product's design connected individuals that should not be connected (minors and adult sexual predators), and that the design did so before any content was exchanged between them.[303] The alleged defective design was evidenced by the website's user anonymity and the absence of age restrictions.[304] The complaint alleged that the design defect created the predictable consequence of attracting both unsuspecting children and predatory adults, facilitating and encouraging dangerous behavior and harm to children using the product.[305]

---

293. *Id.* at 1093 (citation omitted).
294. *Id.*
295. *Id.*
296. *See* A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814, 819 (D. Or. 2022).
297. *Id.* at 817.
298. *Id.*
299. *Id.* at 819; *see Lemmon*, 995 F.3d at 1091–93.
300. *See* Chaney et al., *supra* note 16 (providing an example that social media such as Twitter or Facebook will have a "like" or "friend" feature regarding posts).
301. *Omegle.com*, 614 F. Supp. 3d. at 817.
302. *Id.* at 820–21.
303. *Id.*
304. *Id.* at 819–820 n.2.
305. *Id.*

iii. Inadequate warning

Before addressing reasonable alternative design, it is worth briefly addressing product defects because of inadequate instructions or warnings.[306] When the foreseeable risk of harm posed by a product could have been reduced by a reasonable instruction or warning, and the omission of instruction or warning renders the product unsafe, it may be considered defective.[307] The standard is reasonableness.[308] Relating back to the chainsaw analogy, warning stickers, directions, and standards for use abound. With algorithms, the user is unaware of how the algorithm may use data to learn, revise, and conduct transfer learning.[309] Algorithm transfer learning is a form of machine learning where, instead of starting off with a blank slate, the machine has already learned a required task.[310] The algorithm then creates, develops, and builds on the data presented at an extremely fast rate. Providing a warning regarding algorithm content development continually building upon itself may seem unreasonable. Yet, many social media service providers regularly hold users accountable for various infractions regarding the content of the data posted and exploited by the algorithms.[311]

b. Defective design and a reasonable alternative design

Reasonableness within the context of defective design product liability is distinct from negligence actions, which focus on the conduct of parties toward one another.[312] Product liability parties are generally using products created to generate an economic or alternative benefit.[313] Strict liability negligence turns on whether the seller failed to use reasonable care.[314] The question

306. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(d), (i) (AM. L. INST. 1998).
307. *Id.*
308. *Id.*
309. *See* Mihalkova et al., *supra* note 137.
310. *See* HINRICHS ET AL., *supra* note 135, at 1. The Abstract provides that "Transfer Learning is a generalization of machine learning where instead of starting with no information when given the target task, the machine has already been able to learn in a source task." *Id.* "The degree to which the information gained by learning from the source is useful for the target is dependent on the relationship between the source and target tasks." *Id.*
311. Ariana Tobin & Jeremy B. Merrill, *Besieged Facebook Says New Ad Limits Aren't Response to Lawsuits*, PROPUBLICA (Aug. 23, 2018, 12:48 PM) https://www.propublica.org/article/facebook-says-new-ad-limits-arent-response-to-lawsuits#:~:text=Facebook's%20move%20to%20eliminate%205%2C000,discrimination%2C%20the%20company%20said%20Wednesday.
312. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. f (AM. L. INST. 1998).
313. *Id.*
314. RONALD W. EADES, MASTERING PRODUCTS LIABILITY 34 (2008).

is simply whether the nature of the product, as created, is safe?[315] In addition to the seller's reasonable care, the issue is the product itself.[316]

A manufacturing defect is claimed when the product fails to meet the manufacturing design specifications.[317] A defective design is claimed when the product has met the specifications but the specifications themselves are what create unreasonable risk.[318] As addressed in Section II, algorithm parameters, recommender algorithms, transfer learning, and machine learning algorithms would be specifications that may create unreasonable risk. [319]

In *Roommates*, *Lemmon*, *Maynard*, and *Omegle*, the common thread is that social media providers can be liable for an existing defective design that creates an unreasonable risk before use.[320] In *Omegle*, the alleged defective design could have been remedied simply by refraining from altering or changing content posted by users.[321] The foreseeable risk of harm within the Omegle's product, the risk that anonymous adults and children interacting within cyberspace, could have easily been avoided by creating several reasonable safeguards and warnings.[322] Likewise, in *Lemmon* and *Maynard*, the social media service provider had a duty to adopt a reasonable and safer design knowing that a foreseeable risk of harm existed in the common use of the product.[323] Reasonable alternative designs for an algorithm can be as simple as changing a set of parameters.[324] This may include reasonable limitations on the algorithm application depending on user time, place, and manner of use. The data and content fed to the algorithm may be limited. The reasonable alternative design for algorithms can thus be narrowed or broadened and even tailored to data for specific purposes. Reasonable alternative designs would need to consider the collaboration of data, decision parameters, and designated machine-learning objectives balanced against the utility of use below.

Internet algorithms, mining data as designed, establish connections between individuals before any content exchange.[325] The algorithm design

315. *Id.*
316. *Id.*
317. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(b) (AM. L. INST. 1998).
318. *Id.* § 2 cmt. d–f.
319. *See* Chaney et al.*, supra* note 16; s*ee also* Gal & Petit, *supra* note 25.
320. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1175 (9th Cir. 2008) (en banc); Maynard v. Snapchat, Inc., 870 S.E.2d 739, 748 (Ga. 2022); Lemmon v. Snap, Inc., 995 F.3d 1085, 1093 (9th Cir. 2021); A.M. v. Omegle.com, LLC, 614 F. Supp. 3d 814, 819 (D. Or. 2022).
321. *See Omegle.*com, 614 F. Supp. 3d. at 819.
322. For an example of a warning on Google, see GOOGLE, *Social Media Warning*, CHROME WEB STORE, https://chrome.google.com/webstore/detail/social-media-warning/jldmmfhhjgopdfogeihbopebhhogelfc (last visited Sept. 13, 2023).
323. *Maynard*, 870 S.E.2d at 747.
324. *See* Chaney et al.*, supra* note 16.
325. *See* Leetaru, *supra* note 24.

presupposes the content and while the design relies on the content, it is still independent of it. The basis for a defective design, and thus any reasonable alternative design, will be balanced with the recommendations and machine learning parameters.[326] Once the algorithm collaborating and developing content within § 230(f)(2), is demonstrated to be a defective product, it must be shown as unreasonably dangerous.[327] The reasonableness of the alternative design would be a cost-benefit analysis either in consumer expectation or risk utility.[328]

c. Reasonableness of danger reaching the consumer

In *Roommates*, § 230 did not shield the social media service provider from liability for defective design claims arising out of its product.[329] In *Lemmon*, *Maynard*, and *Omegle*, the courts held that the social media service provider's owed a design duty to provide products that are not unreasonably dangerous.[330] In general, liability turns on collaboration between users.[331] The *Lemmon* court explained that liability arises in the interplay between the speed filter and reward algorithm system.[332] In *Omegle*, the court explained that liability arises in the interplay between the website and anonymous users.[333] In either case, the defendant's owed a design duty to mitigate the reasonably foreseeable risk in using the products as designed.[334]

The issue that the Court must address is whether a reasonably foreseeable risk within an algorithm itself exists; or, is the algorithm's liability attached to the data that it then uses, thus potentially attenuating foreseeability?[335] Citing the previous certification on this question from the Eleventh Circuit, *Maynard* reiterates that use of a product is not a predicate to liability.[336]

---

326. *See* Chaney et al.*, supra* note 16; Gal & Petit, *supra* note 25.

327. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. d–f (AM. L. INST. 1998).

328. *See id.*

329. Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1175 (9th Cir. 2008) (en banc).

330. *See, e.g.*, Maynard v. Snapchat, Inc., 870 S.E.2d 739, 751 (Ga. 2022); Lemmon v. Snap, Inc., 995 F.3d 1085, 1094 (9th Cir. 2021); A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814, 819–21 (D. Or. 2022).

331. *See, e.g.*, *Maynard*, 870 S.E.2d at 751; *Lemmon*, 995 F.3d at 1094; *Omegle.com*, 614 F. Supp. 3d. at 819–21.

332. *Lemmon*, 995 F.3d at 1094.

333. *Omegle.com*, 614 F. Supp. 3d. at 819–821.

334. *Id.*

335. *See* Gonzalez v. Google LLC, 2 F.4th 871, 913 (9th Cir. 2021) (noting whether social media companies should continue to enjoy immunity for published third-party content they publish and whether algorithms used by entities ought to be regulated).

336. *Maynard*, 870 S.E.2d 739, 748–49 (citing Jones v. NordicTrack, Inc.*,* 550 S.E.2d 101 (Ga. 2001)). In *Jones v. NordicTrack, Inc.*, the Eleventh Circuit Court of Appeals asked the Georgia Supreme Court to certify the question as to whether a product needed to be in use at

The fundamental element of a design defect case is whether the manufacturer breaches its duty to reduce the foreseeable risk of harm presented by the product in failing to adopt a reasonable alternative design.[337] Regardless of whether foreseeability resides within the algorithm or is tethered to data content, the function of an algorithm inherently turns to content development as envisioned in § 230(f)(3) and is thus not immune from liability.

The question is not about negligent use by the defendant but about the product itself. By way of analogy, like an algorithm culling specific data, throwing some to the side, and leaving others to remain, a chainsaw does not cause an injury until it is used. Absent facts showing the chainsaw was built unsafely, generally negligent manner of use will lead to injury. The chainsaw must be directed in a specific manner to create injury. Thus, these products reasonably lead to foreseeable risk and carry multiple safeguards and warnings in multiple languages because the seller and the buyer know the risk of using chainsaws.

i. Risk utility of algorithms

In the chainsaw product liability analogy, the risk and the utility of the product are known. The creator and the user are on equal footing. Equal knowledge is conveyed, and warnings provide adequate notice. To cull the forest, a dangerous product must be used. A chainsaw cannot be safe. However, the danger of the chainsaw is what makes it useful as a product. The public expects the chainsaw to function and operate a certain way to do a certain job. Outside of these parameters or boundaries lies a gray area which generally, if not used as intended, bars product liability. Consumer expectation, also known as the consumer expectation test for product liability, offers one approach to determining liability with algorithms.[338] Algorithms are not chainsaws, however, and the chainsaw doesn't recreate itself into an ax, a tree shredder, or a forest fire on its own based on its use like an algorithm does absorbing harmful data or content. Algorithms are complex products. Use of these products and user expectation of use requires knowledge. When the algorithm is redefining itself, learning from content data provided in real-time, and recreating and re-prioritizing itself, consumer expectations and knowledge of the product are unrealistic.[339]

---

the time of injury for a manufacturer to be held liable for a defective design. 550 S.E.2d 101, 102 (Ga. 2001). The Court held that use of the product was not determinative of liability. *Id.*

337. *Jones*, 550 S.E.2d at 103.

338. *See* CORMEN ET AL., *supra* note 189, at 36.

339. *See* Chaney et al., *supra* note 16 (discussing fine-tuning algorithms, machine learning, bias, and causation relative to algorithms); *see also* Gal & Petit, *supra* note 25, at 639–40 (same).

In determining a reasonable alternative design, the most appropriate test is balancing the reasonableness of risk-utility.[340] Assuming the algorithm meets the manufacturer's design, the analysis turns to the design specification creating unreasonable risk.[341] When judging the defectiveness of product design, the risks created by the product are balanced against the utility created by the product.[342] Here, the cost-benefit analysis is clear. First, would a reasonable alternative design, at a reasonable cost, reduce the foreseeable risk of harm posed by the algorithm?[343] Second, does the omission of the alternative design by the seller render the algorithm unreasonably safe?[344] Finally, does the use of a reasonable alternative design, balanced with the risk-utility render the algorithm worthless?[345] Would the proposed alternative design defeat the purpose for which the algorithm is designed to achieve? The standard for product design assessment compared to alternative design is that of a reasonable person. [346]

First, algorithms are ubiquitous in the 21st-century marketplace. The economic value of an algorithm is quantifiable and can be high. In advertising alone, algorithms may create early revenue streams in the amount of tens of billions of dollars.[347] Algorithms may assist in determining crucial factors in life and death decisions and choices of who receives treatment, and why.[348] Because the reasonable value of the algorithms can be high in some cases, the risk-utility balance may be clear. From the reasonable person standard, the value and risk-utility of the algorithm are case specific. For example, a person would likely find a proposed management algorithm evaluating pancreatic cystic lesions more reasonable than an algorithm that enables the communication and dissemination of videos on how to kill other humans.[349] Reasonableness has its boundaries.

As courts have shown in *Roommates*, *Lemmon*, *Maynard*, and *Omegle*, there are limits on reasonableness. In these cases, the cost of changing the algorithm was held to be reasonable even if the product had significantly less

340. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(b) (AM. LAW. INST. 1998).
341. *See id.* § 2 cmt. d.
342. *Id.*
343. *See id.*
344. *See id.*
345. *See id.* § 2 cmt. d–f.
346. RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. d (citing RESTATEMENT (SECOND) OF TORTS § 283 cmt. c (AM. L. INST. 1965)).
347. *See* Leetaru, *supra* note 24.
348. *See* Joshua Sharfstein, *How Health Care Algorithms and AI Can Help and Harm*, JOHN HOPKINS BLOOMBERG SCH. PUB. HEALTH (May 2, 2023), https://publichealth.jhu.edu/2023/how-health-care-algorithms-and-ai-can-help-and-harm.
349. Sherry J. Lim et al., *Preoperative Evaluation of Pancreatic Cystic Lesions: Cost-Benefit Analysis and Proposed Management Algorithm*, 138 SURGERY 672 (2005); s*ee* Force v. Facebook, Inc., 934 F.3d 53, 58–59 (2d Cir. 2019).

value or was even made useless.[350] In *Omegle*, the main purpose was generating contact between anonymous users, also serving as the primary way causation was attached to the harm.[351] But for the anonymity feature, it can be reasonably inferred that child sex predators would not use the product. The same principle is held up in *Roommates*, *Lemmon*, and *Maynard*.

ii. Risk utility—machines arise

The question the Court left open in wake of *Gonzalez v. Google* is what cost-utility standard should be applied. Google will argue that § 230 provides immunity for re-publication and its limitation is cost prohibitive. Balancing regulation and the resulting impact on marketing revenue would not, however, absent some statutory regulation, address the fundamental issue of reasonableness.[352] The standard is whether there is a reasonable alternative design that exists that would have reasonably mitigated foreseeable harm due to the use of the product irrespective of whether the omission of an alternative design made the product unreasonably safe.[353] *Gonzalez* would need to show that the product itself is what is egregious, not the speech or re-publication under § 230. The product used in a manner that is unreasonable and which creates foreseeable risk is at issue. Algorithms may ingest data and learn like humans, but they are not accountable for liability under the law.[354] At least not yet.

The first consideration in balancing risk-utility is the reasonableness standard applied to those creating the algorithms. The foreseeable risk of the algorithm design is what product manufacturers, like Google, must address. The costs for reducing the foreseeable harm could be minimal and may even prove to be profitable. For example, creating vetting procedures for how data is used and classified, and building self-regulation algorithms whose sole purpose is to identify, assess, and report risk would be low-cost fixes to reduce foreseeable harm.[355] Another option under reasonable alternative design is to have a machine learning algorithm police itself. When algorithms are introduced into the marketplace, a reasonable alternative design is assessable, the costs are part of reasonable operational labor expenses, and these reasonable

350. *See generally* Maynard v. Snapchat, Inc., 870 S.E.2d 739 (Ga. 2022); *see also* A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814 (D. Or. 2022).

351. *See Omegle.com*, 614 F. Supp.3d. at 819–820.

352. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. d–f (AM. L. INST. 1998).

353. *Id.*

354. *See* Chaney et al., *supra* note 16.

355. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. a (AM. L. INST. 1998) (products are not generally defective because they are dangerous, e.g., chainsaws). Risk utility balancing assesses advantages and disadvantages of the product design. *See id.* Trade-offs are considered which balance the costs of injury with the potential added costs to the product. *See id.*

alternatives reduce the foreseeable risk of harm posed by algorithms that function in similarity to the human mind, yet unaccountable for their decisions.[356]

Second, in some algorithms, omitting a reasonable alternative design renders the product itself unsafe.[357] Transfer learning, machine learning, or AI algorithms, as designed, may act independently of direct human instruction.[358] Conduct, choices, and decisions made on information the machine itself acquires and analyzes, without human monitoring or oversight, may lead to probable and foreseeable unreasonably unsafe consequences.[359] The moment of algorithm technical singularity, that is the ability of an algorithm to reach human thinking, is at the very least reaching a liability threshold.[360] The manufacturer who built the product carries liability.[361] Because algorithms may make highly consequential decisions in circumstances that may not be anticipated by, let alone directly addressed by, the machine algorithms creators, accountability by the creator should stay intact.[362]

Conversely, in balancing risk-utility and liability, developers and manufacturers of autonomous algorithms, machine learning, or AI systems may be discouraged from taking a product to market lacking in effective safeguards.[363] This would not be an unreasonable risk-utility balance conforming to common business practice. While safeguards may reign in the speed of innovation becoming available to consumers, product liability uncertainty would maximize utility and minimize risk before introducing potentially harmful algorithms.[364]

An algorithm can have a reasonable alternative design, at a reasonable cost, and reduce foreseeable risk. A reasonable alternative design can be achieved by creating better algorithmic oversight, more efficient management of content, and more efficient management of content development within the algorithm itself.[365] This would not render the product worthless, and it would

356. *See* Zakon, *supra* note 17, at 1128; *see also* David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence*, 89 WASH. L. REV. 117, 135–36, 138, 145 (2014); Nick Bostrom, *When Machines Outsmart Humans*, 35 FUTURES 759, 763 (2000).

357. *See, e.g.*, Chaney et al., *supra* note 16; Gal & Petit *supra* note 25; Brummer & Yadav, *supra* note 105, at 274–75; *cf.* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. d–f (AM. L. INST. 1998).

358. *See* Vladeck, *supra* note 356, at 121–22, 148; Bostrom, *supra* note 356, at 763. Artificial intelligence uses the term "singularity" or "technical singularity" to describe the moment in time when machines exceed human intelligence—becoming fully sentient. Consequently, a cascade of complex philosophical and legal questions will arise and society as well as the Court will have to wrestle with them. *Id.*

359. *Id.*

360. *Id.*

361. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. d (AM. L. INST. 1998).

362. *See* Vladek, *supra* note 356, at 121.

363. *See* Jackson, *supra* note 105, at 60.

364. *Id.* at 60.

365. *See* Chaney et al., *supra* note 16; *see* Gal & Petit, *supra* note 25.

not carry an unreasonable burden in cost. The innovation required may likely increase product development and safety. Failing to create proper oversight, omitting the alternative design with parameters and rules on how data is managed, exploited, developed, and sold to consumers may render an algorithm reasonably unsafe.

d. Reaching the consumer without substantial change: a paradox

An algorithm reaching consumers without substantial change may seem self-evident. This would be a gross oversimplification. The social media or search engine internet provider deploys algorithms to mine and exploit the content provided to them.[366] User content data and algorithm recommendations are learned from the original content data.[367] Algorithm decision parameters and choices, including many discriminatory biases of the algorithm then become part of the input for a newly constructed algorithm.[368] Thus, algorithms, such as AI, machine learning, and recommendation algorithms, develop and re-produce content based on the continual re-development of user data. This conduct is outside of § 230 immunity. The conduct, development, and redevelopment of data is a confounding problem.[369] Once unleashed, the misapplication, or abuse, of the algorithm may be the basis of the defective design as it left the "hands of the defendant".[370] Algorithm value, the unchanging attribute, is its adaptability and resiliency to continually change on its own, simultaneously presenting a foreseeable and probable risk for liability.[371] In some algorithms, change and development are inherent. This seems to create a paradox. There is no substantial change when the existential attribute of the product is the ability to change itself; without this attribute, the product has less value. The court's position on this conduct was specifically left open in *Gonzalez v. Google*.[372] The Ninth Circuit explicitly stated that machine learning algorithms may produce content within the meaning of § 230,[373] thus, they would not qualify for immunity.

---

366. For discussion on algorithms learning, deciding, choosing, bias, parameters, and decreasing utility, see, for example, Chaney et al., *supra* note 16; Mihalkova, et al., *supra* note 137; Ananthaswamy, *supra* note 139; Tanz, *supra* note 141.

367. *See* Chaney et al., *supra* note 16; *see also* Gal & Petit, *supra* note 25.

368. *See* Sadagopan, *supra* note 173.

369. *See* Chaney et al., *supra* note 16. While linguistically similar, this should not be confused with the detailed analysis of confounding algorithms. *See id.*

370. *See* EADES, *supra* note 314, at 34.

371. *See* Brummer & Yadav, *supra* note 105, at 274–76.

372. Gonzalez v. Google LLC, 2 F.4th 871, 895–96 (9th Cir. 2021).

373. *Id.* at 896 ("[W]e do not hold that machine-learning algorithms can never produce content within the meaning of Section 230.").

In Sections II and III, algorithms as products were defined and examined showing how they create and develop content within the scope of § 230(f)(3) and outside of immunity. Since *Zeran*, two lines of cases have developed along with advances in algorithms. *Roommates*, *Lemmon*, *Maynard*, and *Omegle* show how elements for product liability can be established regarding Internet service providers acting as Internet content providers.[374] The uncertainty remains in the parallel track with *Force*, *Gonzalez*, and *Dyroff.* In this second line of cases, the key evolving element is causation and material contribution.[375] Section IV addresses how the intended use of algorithms might differ from material contribution to what makes the conduct unlawful.

IV. CAUSATION AND MATERIAL CONTRIBUTION—ALGORITHMS USE IN DEFECTIVE DESIGN

Arguably, the most challenging element for successfully showing an algorithm product liability claim is causation. Intuitively, this makes sense. One might argue that algorithms are only a bunch of numbers strewn together. The courts have held that merely showing the algorithm developing content within § 230(f)(3) is not sufficient.[376] The intended use of the algorithm must be attached to the conduct for which it is used.[377] Thus, the development of content must be shown to materially contribute to the alleged illegality of the conduct.[378] Within the domain of an Internet service or content provider, material contribution does not mean simply displaying illegal conduct.[379] For example, in *Lemmon* and *Maynard*, simply posting and showing the car wreck did not materially contribute to the conduct.[380] The application itself enabled and materially contributed to the foreseeable risk.[381] The injury was the likely and

374. *See generally, e.g.*, Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157 (9th Cir. 2008) (en banc); Maynard v. Snapchat, Inc., 870 S.E.2d 739 (Ga. 2022); Lemmon v. Snap, Inc., 995 F.3d 1085 (9th Cir. 2021); A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814 (D. Or. 2022).

375. *Gonzalez*, 2 F.4th at 893, 917, 923; *see* Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1099 (9th Cir. 2019); *see also* Force v. Facebook, Inc*.,* 934 F.3d 53, 68–69 (2d Cir. 2019).

376. *See Gonzalez*, 2 F.4th at 892–93 (stating a website is not transformed into a content creator or developer by virtue of supplying "neutral tools" that deliver content in response to user inputs).

377. *Id.*

378. Jones v. Dirty World Ent. Recordings LLC, 755 F.3d 398, 410–11 (6th Cir. 2014) (citing Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1167–68 (9th Cir. 2008)).

379. *Id.*

380. Lemmon v. Snap, Inc., 995 F.3d 1085, 1093 (9th Cir. 2021) (citing Maynard v. Snapchat, Inc., 870 S.E.2d 739 (Ga. 2022)). The court has since upheld *Maynard*. *Id.*

381. *Id.*

probable consequence of using the product.[382] In *Omegle*, liability attached to the product's random pairing anonymity feature.[383]

The anonymity feature materially contributed to the illegality of the conduct.[384] Randomly pairing adults with children in an anonymous forum, as in *Omegle*, provided an environment for children to be sexually exploited.[385] Absent the anonymity feature, or in the alternative, a product warning that adult sexual predators could be on the website with children, the illegal conduct was not reasonably foreseeable.[386] The court found that the website's function of randomly matching children with adults caused the danger.[387] Thus, anonymity was the risk that made the reasonably foreseeable conduct harmful. Without anonymity, sexual predators would likely be prevented from engaging in this illegal conduct on the website.

Courts have held that algorithm content must materially contribute to the alleged unlawfulness—not simply augment content or data more generally.[388] At least one state has taken action to address algorithmic decision-making bias as material contribution.[389] Effective January 1, 2023, the City of New York's novel shield cracking statute regulates the use of artificial intelligence in hiring and promotion decisions.[390] The law specifically targets any "computational process derived from machine learning . . . or recommendation used to substantially assist or replace discretionary decision making . . . ."[391]

---

382. *Id.*

383. A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814, 820 (D. Or. 2022) (discussing similarity with the *Lemmon* rationale in that, Omegle could have satisfied its alleged obligation by designing its product differently—for example, by designing a product so that it did not match minors and adults. Omegle would not have to alter the content posted by its users—it would only have to change its design and warnings).

384. *Id.* at 821–22 (discussing the applicability of the Fight Online Sex Trafficking Act (FOSTA), 18 U.S.C. § 2421A, to the conduct).

385. *Id.* at 820 n.2 (random pairing function of adults and children and the service's accessibility to both adults and children work in tandem; plaintiff's claims have nothing to do with information provided by a user).

386. *Id.* at 819–820 n.2.

387. *Id.*

388. Jones v. Dirty World Ent. Recordings LLC, 755 F.3d 398, 410–11 (6th Cir. 2014) (citing Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1167 –68 (9th Cir. 2008) (en banc)) (discussing material contribution); *see also* Gonzalez v. Google LLC, 2 F.4th 871, 892–95 (9th Cir. 2021).

389. *See* Forrest, *supra* note 111 (referencing N.Y.C. ADMIN. CODE §§ 20-870–20-874).

390. *Id.*

391. *See id.* The tool must be subjected to a bias audit no more than one year prior to its use, and the employer must publish a summary of the most recent audit results on the employer's website. *Id.* Further, the employer must notify candidates or employees of its use of the tool and the metrics the tool will use to assess them, allow them to request an accommodation or alternative selection process, and, upon written request, provide them with information about the type and source of data collected and the employer's data retention policy, if that information is not already on the employer's website. *Id.*

Additional claims have been initiated and bills have been proposed in the New York State Senate regarding criminal profiling and regulation of social media content.[392] The intent of the legislation prohibits the use of algorithms or other automated systems which prioritize content by methods outside of objective parameters, such as a date and time stamp.[393] Considering the jurisprudential history leading to *Gonzalez v. Google* and the developing posture of state legislatures, limitations on § 230 immunity are imminent. The issue behind pernicious algorithms' intended use and use within potentially unlawful conduct now turns on material contribution, substantial assistance, and encouragement.[394]

A. Material Contribution Rule

Material contribution, in its simplest form, is the development of information beginning with the collaboration of the algorithm to the content data making the content illegal or actionable.[395] For algorithms within the domain of § 230, material contribution continues with the development of content fed to the algorithm.[396] This is not general content augmentation. The algorithm must materially contribute to the content's alleged illegality.[397] Material contribution to alleged illegality is more than merely displaying unlawful content; it means being responsible for what makes the content allegedly unlawful.[398] The court in *Jones* went to great lengths in citing *Roommates* to provide examples of material contribution.[399] Material contribution cannot be passive, but, it can be developmental.[400] Development may include some functions a website operator may conduct with respect to content originating from a third party.[401] Ratification or adoption of content, however, does not make one a creator or developer of content within § 230(f)(3).[402] In addition, "neutral

392. *Id.*
393. *Id.*
394. *See* Gonzalez v. Google LLC, 2 F.4th 871, 913–18 (9th Cir. 2021) (Katzmann, C.J., concurring in part and dissenting in part) (discussing Facebook's friend and content suggestion algorithms under the Communications Decency Act).
395. *See id.* at 892–93 (majority opinion); *see also* DATA-DRIVEN INNOVATION: BIG DATA FOR GROWTH AND WELL-BEING, *supra* note 190, at 152–58; Leetaru, *supra* note 24; Gal & Petit, *supra* note 25.
396. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1167–68 (9th Cir. 2008).
397. *Id.*
398. *See Gonzalez*, 2 F.4th at 891–92, 917, 923.
399. Jones v. Dirty World Ent. Recordings LLC, 755 F.3d 398, 410–11 (6th Cir. 2014).
400. *Id.* at 410–12.
401. *Id.*
402. *Id.*

tools" to carry out what may be unlawful or illicit is not development and protected under § 230.[403]

In order for material contribution to arise, choices, and in the case of algorithms, decisional parameters, are necessarily required.[404] Because Roommates required information about specific protected characteristics and engineered its search and email systems recommending or filtering to limit access to housing listings based on those protected characteristics, the court held that the website materially contributed to the alleged illegality of hiding certain listings.[405] When *Roommates* engineered its search and email systems limiting listings based upon certain protected characteristics it materially contributed to the alleged illegal discrimination.[406] Thus, engineering limitations based upon characteristics, bias, or certain decisional parameters are conduct giving rise to material contribution.[407] The manifestation of material contribution begs an important question that is unanswered in the wake of *Gonzales*; what if an AI or machine learning algorithm decided on its own to discriminate?[408]

In *Gonzalez*, the Ninth Circuit analyzed Google's algorithm recommending content of the ISIS terrorist organization to its users.[409] On this issue, the Ninth Circuit cited its analysis in *Dyroff v. Ultimate Software Group, Inc.*[410] In *Dyroff*, the action was brought against Ultimate Software Group who provided the platform for the "Experience Project," a social media website messaging board.[411] The social media platform allowed users to anonymously communicate in a "blank box" approach.[412]

A person using the website could join groups on his or her own, while at the same time the website also recommended groups for users to join.[413] One of the site's key product features enabled anonymity of the user within the

403. *Id.*; *see also* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1169 (9th Cir. 2008) (en banc).

404. *See* Chaney et al., *supra* note 16; *See also* Vladeck, *supra* note 356, at 121–22; Bostrom, *supra* note 356, at 763.

405. *See Jones*, 755 F.3d at 411.

406. *Id.* at 411–12.

407. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1172 (9th Cir. 2008) (en banc).

408. *See* Chaney et al., *supra* note 16; *see also* Matthew Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. L.J. TECH. 354, 364–65, 367 (2016).

409. Gonzalez v. Google LLC, 2 F.4th 871, 892–95 (9th Cir. 2021).

410. *Id.* at 894 (citing Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093 (9th Cir. 2019)).

411. Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1094 (9th Cir. 2019).

412. *Id.*; *see also* BALLON, E-COMMERCE & INTERNET LAW: CLAIMS AGAINST SOCIAL NETWORKS § 37.05[6] (2d ed. 2020). A "blank box" approach allows communication on any topic without limiting or promoting the type of experiences shared *Id.*

413. *Dyroff*, 934 F.3d at 1095.

groups.[414] Site revenue accrued from both advertising and the sale of tokens that users obtained to communicate with their groups.[415] Partly due to the user's anonymity, the site was used to facilitate illegal drug sales.[416]

Plaintiff's son suffered from opioid addiction and posted on the site in a heroin-related chat group inquiring where he could obtain heroin in Jacksonville, Florida.[417] The website then sent him a notification when a companion user, an Orlando drug dealer, posted in the same chat group.[418] Greer, the plaintiff's son, then met in-person with the narcotics dealer and bought heroin.[419] The heroin contained fentanyl, killing Greer.[420] The illegal distributor admitted to selling heroin laced with fentanyl on the Experience Project site.[421] The court held Ultimate Software did not create the content on the Experience Project, in whole or in part, because its functions, including notifications, and recommendations, were content-neutral tools used to facilitate communication.[422]

Machine learning, recommendation, and feedback loop algorithms, however, function as anything but content neutral.[423] Algorithms learn from the content, make recommendations based on content and facilitate communication based on the algorithm's decision-making parameters.[424] These algorithmic attributes could constitute material contribution to unlawful conduct. But for the algorithm's level of subjective decision-making prioritization, facilitated by the development of the content data, Greer would not have been introduced to the drug dealer who gave him the fatal dose of fentanyl. Likewise, in *Gonzalez*, the argument that the algorithms do not treat ISIS content differently than any other third-party content belies the inherent material contribution existential to machine learning, AI, and recommendation algorithms.[425] An algorithm's inherent purpose is its material contribution in deciding and choosing which data materially contributes to achieving the decisional parameters assigned.[426] Material contribution occurs when the algorithm decides

414. *Id.*
415. *Id.*
416. *Id.*
417. *Id.*
418. *Id.*
419. *Dyroff*, 934 F.3d at 1095.
420. *Id.*
421. *Id.*
422. *Id.* at 1096 (citing Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1167–69 (9th Cir. 2008)).
423. *See* Chaney et al., *supra* note 16. Digital content development and its exploitation via decisional parameters, financially driven or otherwise, carry non-neutrality. *See* Tanz, *supra* note 141.
424. Chaney et al., *supra* note 16.
425. *See* Gonzalez v. Google LLC, 2 F.4th 871, 896–97 (9th Cir. 2021).
426. *See* Tremble, *supra* note 102, at 864; Gal & Petit, *supra* note 25, at 636–39.

what data to exploit, what content to develop, and why.[427] The algorithm's material contribution includes a bias designed to develop data content to generate a determined objective, financial or otherwise.[428]

In May 2023, the Supreme Court of the United States consolidated *Gonzalez v. Google* and *Twitter v. Taamneh*, focusing the issue on whether the platforms aided and abetted terrorism.[429] This was a narrow and limited holding.[430] The Court's holding, including its characterizations of the social-media platforms and algorithms, was case and issue specific, "[o]ther cases presenting different allegations and different records may lead to different conclusions."[431]

In *Gonzalez* and *Twitter*, arguments focused on determining liability for those aiding and abetting terrorism under 18 U.S.C. § 2333 by "knowingly providing substantial assistance," or conspiring with a person who commits an act of international terrorism.[432] In addressing civil aiding and abetting and conspiracy liability for terrorism under the Justice Against Sponsors of Terrorism Act (JASTA), the Court distinguished certain algorithms.[433] In *Gonzalez* and *Twitter*, the Court held that the recommender algorithm was outside of civil liability because the algorithm: (1) was part of the infrastructure, (2) appeared to be "agnostic to the nature of the content," and (3) appeared to be in "passive non-feasance not active in abetting the injury."[434] The Court, by narrowing the focus of the ruling as case specific to JASTA, explicitly left the door open and balked on the issue of product liability in the context of AI, machine learning, or other recommender algorithms.[435] Product liability was not presented, and these specific recommendation algorithms were shown to be immune under § 230, even if they participated in injury.[436]

The Court noted that its holding must consider the wrongdoer's actions against the slippery slope of holding a communication provider liable for any wrongdoing merely for knowing that the wrongdoers were using its services and failing to stop them.[437] While this would take aiding and abetting "far

427. *See Gonzalez*, 2 F.4th at 896–97.
428. *See* Twitter, Inc. v. Taamneh, 598 U.S. 471, 478 (2023).
429. *See id.*
430. *See id.* at 500–02 (distinguishing between liability for intentional decisions to promote content related to injury from other cases where a lesser showing of scienter might be sufficient).
431. *Id.* at 507 (Jackson, J., concurring).
432. 18 U.S.C. § 2333(d)(2); *see Taamneh*, 598 U.S. at 478.
433. *Id.* at 499–500.
434. *Id.* at 500–02.
435. *See id.* at 502–04.
436. *See id.* at 500–02
437. *Id.* at 503.

beyond its essential culpability moorings," the decision does not address the core issue of algorithm product liability.[438]

B. Substantial Assistance

Material contribution, alone, is not sufficient to assign liability.[439] There must be a higher threshold of substantial assistance for an algorithm to be considered a material contribution.[440] The threshold of substantial assistance is achieved when the unlawful conduct is determined to be a cause resulting from the algorithm; the burden then shifts further towards showing responsibility for what makes displayed content illegal or actionable.[441]

But this analysis fails to address the crucible of the digital marketplace. That is the point of Justice Thomas's lengthy discussion in denying certiorari in *Malwarebytes*.[442] Social media and the Internet generally, which § 230 regulates, is a digital marketplace.[443] The digital marketplace is a vast set of locations, a repository, providing digital premises with addresses, and possessory interests like real property.[444] Algorithms, learning from data, like human beings, functionally and purposefully invite users, marketplace consumers, onto the cyber premises and receive a financial benefit from the invitation.[445] As marketplace consumer invitees, a reasonable standard of care to protect invitees against known and observable defects may arise.[446] While this theory

438. *Taamneh*, 598 U.S. at 503.

439. *See id.* at 502 ("There may be . . . situations where the provider of routine services [on a social media platform] does so in an unusual way or provides such dangerous wares that selling those goods to a terrorist group could constitute aiding and abetting a foreseeable terror attacks.").

440. *Id.* at 494–95, 501–05.

441. *See e.g.*, Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1099–1100 (9th Cir. 2019). The court focuses on the distinction of responsibility with alleged unlawfulness of content. *See id.* at 1099 (citing Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1175 (9th Cir. 2008); Kimzey v. Yelp! Inc., 836 F.3d 1263, 1269 (9th Cir. 2016)).

442. Malwarebytes, Inc. v. Enigma Software Grp. USA, 946 F.3d 1040, 1053 (9th Cir. 2019), *cert. denied*, 141 S. Ct. 13 (2020) (statement of Thomas, J., respecting the denial of certiorari) (discussing content it knows to be illegal).

443. *See* Schroeder, *supra* note 143, at 22 (covering content creation, creators, and virtual communities); *see also* Leetaru, *supra* note 24 (discussing algorithms as proprietary property with interests including intellectual property); Gal & Petit, *supra* note 25, at 645–46.

444. *See* Gal & Petit, *supra* note 25, at 645–46; *see also* Vladeck, *supra* note 356, at 121–22.

445. *See* Tremble, *supra* note 102, at 827, 838; *see also* Gal & Petit, *supra* note 25, at 645-46; Schroeder, *supra* note 143.

446. *See* Abedi et al., *supra* note 212. The distinction, here, is the comparison between property and services which carry liability in a tangible space or forum—like an airplane, ship, or a business with a property address—and cyber space where algorithms manifest their conduct and control. *See* RESTATEMENT (SECOND) OF TORTS § 332 (AM. L. INST. 1965). The

of liability seems to creep into negligence, an algorithm here may be dangerous chattel when used as intended as defined in Section 392 of the Restatement Second of Torts, the basis of liability is still firmly grounded in the product on the digital premises.[447] Thus, substantial assistance of algorithms regarding causation within the marketplace becomes clearer. Defective algorithms pose a foreseeable risk on digital premises.[448]

In the digital marketplace in which *Gonzalez v. Google* is at issue, websites and social media Internet providers are premises in more ways than one.[449] Digital mediums have MAC addresses; like real property, they engage and invite entrants to use their business, and the algorithms work within the cyber storefront recruiting users.[450] Recruitment occurs based on the content developed by the algorithms aggregating user data and influencing engagement.[451] This is substantial assistance in what may be unlawful, or at the least foreseeably negligent, conduct.[452] Substantial assistance by the product and

---

Restatement considers real tangible property; here, liability accrues in cyber space with digital addresses and the algorithm is chattel property. *See id.* § 392.

447. The negligence argument may be applicable in this instance. *See* RESTATEMENT (SECOND) OF TORTS § 392 (AM. L. INST. 1965). Ownership of chattel is immaterial. Algorithms, functioning as chattel property outside of ownership, are used for business purposes. *See id.* Under § 392,

> One who supplies to another, directly or through a third person, a chattel to be used for the supplier's business purposes is subject to liability to those for whose use the chattel is supplied, or to those whom he should expect to be endangered by its probable use, for physical harm caused by the use of the chattel in the manner for which and by persons for whose use the chattel is supplied: (a) if the supplier fails to exercise reasonable care to make the chattel safe for the use for which it is supplied, or (b) if he fails to exercise reasonable care to discover its dangerous condition or character, and to inform those whom he should expect to use it.

Id.

448. *See* Scherer, *supra* note 408 , at 357, 359, 365–66; *see also* Chaney et al., *supra* note 16.

449. *See* Abedi et al., *supra* note 212 (tracking and utilization of Media Access Control Data); *see also* Schroeder, *supra* note 143, at 24–26 (discussing how cyber environments are different than traditional, physical, and societally constructed environments and have "substantially shifted the marketplace").

450. *See* Tremble, *supra* note 102, at 837–39 (developing the science of engagement); *see also* Leetaru, *supra* note 24; Abedi et al., *supra* note 212; Schroeder, *supra* note 143, at 24–26.

451. *See* Tremble, *supra* note 102, at 837–39 (discussing algorithms aggregating data, influencing users time spent on websites in cyber space, and how algorithms effect users).

452. *See* Tremble, *supra* note 102, at 838–39 (algorithms using behavioral science data to "prioritize posts not only based on user's previous interests but also on a post's ability to gain (likes)," increases engagement and use of the product); *e.g.*, Maynard v. Snapchat, Inc., 870 S.E.2d 739, 747–50 (Ga. 2022) (speed filter defective design encouraging engagement as a foreseeable risk).

substantial assistance for what makes the conduct unlawful are not mutually exclusive in some algorithms.[453]

Algorithms are components of software.[454] Component sellers may be liable when the components themselves are defective or when component providers substantially participate in the integration of components into the design of the other products.[455] Establishing algorithms as substantially assisting the software for the development of content on the Internet, algorithms must then be shown to substantially assist unlawful conduct.[456] In *Jones*, *Gonzalez*, and *Dyroff*, the threshold for substantial assistance is in making the displayed content allegedly unlawful.[457] The allegedly unlawful conduct must occur. If the plaintiffs in *Roommates*, *Dyroff*, *Maynard*, *Lemmon*, and the juvenile victim in *Omegle*, never enter the digital premises they are invited to enter, they experience no harm.

The algorithms are designed, through feedback loops and machine learning, to invite and exploit these individuals like customers entering their digital premises.[458] In each case, the algorithm's conduct would not have occurred but-for their specifically designed parameters. Similar conduct, while not as clearly defined, occurred in *Gonzalez* and *Force*.[459] The confusion is conflating algorithms as communication, thus subject to immunity under § 230 when the conduct creating liability is that of intangible property acting in agency of a principle.[460]

In *Gonzalez* and *Twitter*, the Court distinguished algorithms as part of an infrastructure, in other words agents acting on behalf of the principle, which could be liable.[461] In concurrence, Justice Jackson explicitly stated that

---

453. *See* Tremble, *supra* note 102, at 838–39.

454. *See e.g.*, Gal & Petit, *supra* note 25, at 636–37; Mihalkova et al., *supra* note 137; Tanz, *supra* note 141.

455. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 5 (AM. L. INST. 1998).

456. Gonzalez v. Google LLC, 2 F.4th 871, 893, 917, 923 (9th Cir. 2021) (there must be something more than "neutral tools" that deliver content in response to user inputs); *see* Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1099 (9th Cir. 2019); Force v. Facebook, Inc., 934 F.3d 53, 68–69 (2d Cir. 2019).

457. *See Gonzalez*, 2 F.4th at 894; *see* Jones v. Dirty World Ent. Recordings LLC, 755 F.3d 398, 410–11 (6th Cir. 2014) (citing Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1167–68 (9th Cir. 2008)).

458. *See* Chaney et al., *supra* note 16 ("This suggests that the 'tyranny of majority' and niche 'echo chamber' effects may both be manifestations of the same problem: over-exploitation of recommendation models.").

459. *See Force*, 934 F.3d at 77 (Katzmann, C.J., concurring in part and dissenting in part); *see also Gonzalez*, 2 F.4th at 914–15 (Berzon, J., concurring).

460. *See* Vladeck, *supra* note 356, at 141, 145, 150.

461. *See* Twitter, Inc. v. Taamneh, 598 U.S. 471, 500 (2023) ("Viewed properly, defendants' 'recommendation' algorithms are merely part of that infrastructure [a social media platform]. 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."); *But see id.* at 507 (Jackson, J., concurring) (there may be situations where the social media

liability may attach in certain circumstances just not in the cause of action as presented under JASTA.[462]

1. Foreseeability

But-for an algorithm's substantial assistance facilitating the illegal content, the actionable tort does not arise. The intent of the algorithm's conduct may create a foreseeable risk of harm. The purpose and actions of the algorithm functioning on digital premises are to produce an economic benefit.[463] The value of the economic benefit created by the use of the algorithm is balanced with the ability to mitigate the risk of harm.[464] When algorithms are deployed, their conduct is determined by the parameters and choices which direct their conduct. These parameters may be determined by human designation, machine learning, or artificial intelligence.[465] Based on the line of cases across multiple circuit courts, a theory of liability takes root in a sliding scale of the level of foreseeable risk.[466] On one end of the spectrum, there are *Roommates*, *Maynard*, *Lemmon*, and *Omegle* where there is overtly cognizable conduct with foreseeable risk. The "proximate-cause inquiry asks whether a prudent [manufacturer] would foresee an appreciable risk that, because of an unreasonable design decision, some harm would happen according to ordinary and usual experience."[467] In each instance, the products were functioning as intended relative to an alleged defective design.[468] The design was created by one who reasonably knew or should have known that it would have created the risk which led to harm.[469]

---

platform provides services in an unusual or dangerous way that it leads to foreseeable risk of harm).

462. *Id.* at 507 (Jackson, J., concurring).

463. *See* Leetaru, *supra* note 24 (discussing algorithms mining and selling data); *see also* Tremble, *supra* note 102, at 837–39 (developing the science of engagement).

464. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2 cmt. a–d (AM. L. INST. 1998); *see also supra* Section III.

465. *See* Vladeck, *supra* note 356, at 120–21; *see also* Gal & Petit, *supra* note 25, at 636–37; Mihalkova et al., *supra* note 137.

466. *See e.g.*, Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1165–67, 1175 (9th Cir. 2008) (en banc) (providing discriminatory parameters under the Fair Housing Act); Maynard v. Snapchat, Inc., 870 S.E.2d 739, 747 (Ga. 2022) (speed filter designed on the application created a foreseeable risk to the user); Lemmon v. Snap, Inc., 995 F.3d 1085, 1092 (9th Cir. 2021); A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814, 820 (D. Or. 2022) (the anonymity parameters were reasonably related to causation).

467. *Maynard*, 870 S.E.2d at 747.

468. *Id.* at 748–49; *Omegle.com*, 614 F. Supp. 3d. at 820; *Lemmon*, 995 F.3d at 1092.

469. *See Maynard*, 870 S.E.2d at 748–49; *see* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. § 2(b) cmt. f (AM. L. INST. 1998); *see also* OWEN & DAVIS ON PROD. LIAB. § 10:6 (4th ed. 2023).

In *Gonzalez*, *Force*, and *Dyroff*, algorithm collaboration with data content and causation with the internet provider is not as clear. This distinction is at the heart of *Gonzalez*. The Ninth Circuit held in *Gonzalez* that algorithms do not treat ISIS-created content differently than any other third-party content and are immune under § 230.[470] However, the algorithm's inherent purpose and design as a product are to exploit bias and treat third-party content differently.[471] This is non-neutral conduct. Algorithms have been shown to make decisions and choices regarding information based on a set of criteria that carry inherent bias.[472] *Dyroff*, like *Omegle*, sought to work around this issue by creating anonymity.[473] In *Dyroff*, no duty was applied because the Ninth Circuit found the features of the website amounted to content "neutral" functions not creating the risk of harm, and these content functions were used regardless of in which groups the user participated.[474] The court's reasoning would be accurate absent the inherent bias defect designs, but for which, the conduct would not have occurred.[475] Notably, the text of § 230 does not allow for misfeasance or unlawful conduct for some users just because all users are able to perpetuate such conduct, yet do not.[476]

While the holdings were different, in *Dyroff*, like *Omegle*, the foreseeable risk not only remained intact, but, because of the anonymity feature, a reasonable and probable inference would be that risk of injury likely increased.[477] Developing content in this manner would lead to a reasonable circumstantial evidentiary inference supporting the conclusion that the defect was a contributing cause of the harm.[478] The anonymity feature and the algorithms collaborating with it were a determinative component creating a probable and foreseeable consequence, the unlawful conduct.[479]

---

470. Gonzalez v. Google, LLC, 2 F.4th 871, 895 (9th Cir. 2020).

471. *See* Jackson, *supra* note 105, at 42, *see* COCKBURN ET AL., *supra* note 118, at 126–27; *see also* Rogers, *supra* note 141.

472. *See* Rogers *supra* note 141*; see also* Sadagopan, *supra* note 173.

473. Dyroff v. Ultimate Software Grp., Inc., 934 F.3d 1093, 1100 (9th Cir. 2019).

474. *Id.* at 1100–01.

475. *See* RESTATEMENT (THIRD) OF TORTS: PROD. LIAB. §§ 2–3 (AM. L. INST. 1998). When circumstantial evidence supports the conclusion that a defect was a contributing cause of the harm and that the defect existed at the time of sale, it is unnecessary to identify the specific nature of the defect and meet the requisites of Section 2. *Id.* It is important to emphasize the difference between a general inference of defect under Section 3 and claims of defect brought directly under Sections 1 and 2. Section 3 claims are limited to situations in which a product fails to perform its manifestly intended function, thus supporting the conclusion that a defect of some kind is the most probable explanation. *Id.*

476. *See Gonzalez,* 2 F.4th at 926 n.9.

477. *Dyroff*, F.3d at 1100 ("[O]nline privacy is a ubiquitous public concern for both users and technology companies").

478. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1175 (9th Cir. 2008) (en banc).

479. *See* A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814, 820 (D. Or. 2022).

While *Dyroff* and *Gonzalez* are a different genus from *Omegle* and *Roommates*, collectively, these cases offer guidance on establishing elements for an equitable ruling vis-a-vis § 230. The greater the foreseeability of conduct, harmful or otherwise, the more likely substantial assistance will be found.[480] Algorithms are intended to effectuate an outcome, the more substantial assistance the better.[481] Thus, if the algorithm is trying to resolve pancreatic cancer, the risk-utility favors the product.[482] However, if the algorithm is directed towards anonymous drug dealing, sexual abuse, or terrorism, the risk-utility evaporates.[483] Generally, this distinction is understood by humans, but algorithms not as much.[484] At least not yet.

C. Encouragement

The policy rationale of the Communications Decency Act is to "encourage the development of technologies which maximize user control."[485] Congress could not have foreseen the power, adaptability, and influence algorithms have across the economy or within the law.[486] Yet the intent of § 230(b)(3) remains clear—to maximize user control.[487] Algorithms are designed and produced to supplant, manipulate, and undermine user control.[488] Evidence provides a reasonable argument demonstrating some algorithms, as designed, work independently of user control.[489] This conduct flies in the face of the policy undergirding 47 U.S.C. § 230(b)(3) which encourages an increase in user control of information, *not* decreasing user control.[490]

---

480. *See* Gal & Petit, *supra* note 25, at 636–38; COCKBURN ET AL., *supra* note 118, at 140. In answering this issue, digital data and technology companies will pivot to the argument that algorithms are created within multiple layers of computer-generated algorithms making it difficult to determine particular "decisions" based on data inputs. Because the algorithm re-creates itself repeatedly, assigning liability with the product would fall back to the content creating the algorithm. The content would be argued as attenuated from any harm because that is how the algorithm re-creates and develops content. While this may obfuscate liability attenuation does not escape proximate cause. There is someone behind the curtain, even if there are thousands of curtains to pull.

481. *See* Gal & Petit, *supra* note 25, at 636–38; COCKBURN ET AL., *supra* note 118, at 140.

482. *See* Gal & Petit, *supra* note 25, at 636–38; COCKBURN ET AL., *supra* note 118, at 140.

483. *See* Gal & Petit, *supra* note 25, at 636–38; COCKBURN ET AL., *supra* note 118, at 140.

484. *See* Gal & Petit, *supra* note 25, at 636–38; COCKBURN ET AL., *supra* note 118, at 140.

485. 47 U.S.C. § 230(b)(3).

486. Gonzalez v. Google, LLC, 2 F.4th 871, 897, 913 (9th Cir. 2020) (inviting Congress to address the issue directly).

487. 47 U.S.C. § 230(b)(3).

488. *See* Chaney et al., *supra* note 16; *see also* Scherer, *supra* note 408, at 363, 366, 373 (addressing Artificial Intelligence foreseeability and causation).

489. *See* Bostrom, *supra* note 356, at 763; *see also* Chaney et al., *supra* note 16; Vladeck, *supra* note 356, at 135–36, 138, 145; Gal & Petit, *supra* note 25.

490. 47 U.S.C. § 230(b)(3).

Statutorily, the policy shows no textual intent to benefit corporate control of consumers.

Encouragement is more directly related to conduct which drives economic utility. In *Lemmon* and *Maynard*, the plaintiffs argued that the conduct creating liability was encouraged, however, the courts' holdings turned on whether the injury was a reasonably foreseeable and proximately caused by the alleged defective design.[491] In *Lemmon*, the court held "the CDA does not shield Snap from liability for the predictable consequences" of a defectively designed product that allegedly "encourages dangerous behavior."[492] Thus, the standard for a prudent product creator is one who "foresee(s) an appreciable risk" and that "as a result of the design decision" some harm happens.[493] While in *Roommates* encouragement was less proximate,[494] in *Omegle*, the anonymity of the feature as a design defect created the predictable consequence of attracting both unsuspecting children and predatory adults to the social media internet platform.[495] The court found this facilitated and encouraged dangerous behavior and harm to children using the product.[496] Encouragement can thus be both explicit and implicit depending on the level of engagement and the users. However, the product facilitating the unlawful content, the algorithm, contributes to making this distinction. As in *Dyroff*, one could reasonably argue algorithmic misfeasance or even non-feasance.[497] Either the defendant put the plaintiff in a worse position by elevating risk or the defendant should have helped the plaintiff but failed to do so.[498] The Ninth Circuit found that the algorithms were content-neutral and did not create a duty because no risk of harm was created.[499] This assumes important factors within algorithms do not exist, such as decisional parameters that guide and control how algorithms are monetized and applied to data.[500]

Algorithms may be designed, created, and trained, to think, learn, choose, and decide like humans.[501] When doing so, algorithms are interested parties, servants to a master, inviting witting or unwitting users onto their cyberspace premises. Some are granted entry and engaged, while others are not.

---

491. *See* Maynard v. Snapchat, Inc., 870 S.E.2d 739, 747–48 (Ga. 2022).
492. *See* Lemmon v. Snapchat, Inc.*,* 995 F.3d 1085, 1094 (9th Cir. 2021).
493. *Maynard*, 870 S.E.2d at 747–48; *see Lemmon*, 995 F.3d at 1091.
494. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1175 (9th Cir. 2008) (en banc).
495. A.M. v. Omegle.com, LLC, 614 F. Supp. 3d. 814, 819–20 (D. Or. 2022).
496. *Id.*
497. Dyroff v. Ultimate Software Grp., Inc.*,* 934 F.3d. 1093, 1100-01 (9th Cir. 2019).
498. *Id.* at 1101.
499. *Id.*
500. *See* Bostrom, *supra* note 356, at 763; Vladeck, *supra* note 356, at 135–36, 138, 145; Gal & Petit, *supra* note 25; Leetaru, *supra* note 24.
501. *See* Vladeck, *supra* note 356, at 120–21, 126; *see also* Scherer, *supra* note 408, at 364–65, 367.

Their ambitions are driven by their master and a duty of reasonable standard of care may attach.

In *Gonzalez*, the Ninth Circuit left the door open on encouragement.[502] While finding that Google's algorithm was content-neutral, the court explicitly stated that it was not holding that machine learning algorithms could never produce content within the meaning of § 230.[503] Inviting clarity as to the threshold required for material contribution, the Ninth Circuit presented the issue to the Supreme Court within the substantially limited application of the Justice Against Sponsors of Terrorism Act (JASTA).[504] As shown earlier, the use of algorithms is now ubiquitous and, absent direct engagement with algorithms product liability, the Supreme Court's later holding in the case leaves the door open for further litigation across industries.[505] The Court now appears ready to address that some algorithms, as designed, may develop data content not only substantially assisting and materially contributing but in encouraging, either implied or explicit, alleged unlawful conduct.[506]

V. CONCLUSION

Gonzalez v. Google ripened an issue that Congress has failed to address. While Congress could not have enacted § 230 anticipating every possible outcome, Congress anticipated injury and harm related to the Internet as it developed.[507] Twenty-five years later, the industry has failed to effectively self-regulate, Congress has attempted to address injury and harmful acts over the Internet via the Anti-Terrorism Act and the Fight Online Sex Trafficking Act, yet algorithms, like the human mind, continue to develop and adapt. Algorithms are the nerve center for artificial intelligence, machine learning, robotics, and the advent of modern technology. The intersection of these products and a multitude of applications has created an inflection point with the law. The product acts and conducts itself much like humans, the more it does the more effective it may be. Like humans, however, not all algorithms are the same. Some are dangerous.

Algorithms are products and thus carry product liability. Algorithms functioning to develop data content, as identified in § 230(f)(3) do not fall

502. Gonzalez v. Google LLC, 2 F.4th 871, 892 (9th Cir. 2019) (describing how encouragement materially contributes to alleged unlawfulness).

503. *Id.* at 896.

504. *See* 18 U.S.C. § 2333(d)(2).

505. *See Gonzalez*, 2 F.4th at 892, 896.

506. *See* Twitter, Inc. v. Taamneh, 598 U.S. 471, 493 (2023) (distinguishing material contribution for criminal and tort liability within the framework of aiding and abetting and "consciously and culpably" participating in the alleged conduct). However, the Court failed to recognize that machine learning algorithms that utilize artificial intelligence which, as shown *supra* Section IV, may make "conscious and culpable" decisions based on data.

507. 47 U.S.C. § 230(b)–(f).

within the immunity provided by § 230. Algorithm content development may in certain circumstances contribute to what makes the content unlawful. Thus, just because all users can perpetuate such conduct yet do not, this does not mean algorithms that materially contribute to harm should be shielded.

In cracking the § 230 shield of immunity, the presumption may need to be that the algorithm, *depending on its use* as pleaded in the complaint, inherently withstands the initial motion to dismiss under Federal Rule of Civil Procedure 12(b)(6).[508] Using some algorithms, depending on the threshold of risk, may create a rebuttable presumption favoring discovery. Consequently, when the expressed purpose for creating algorithms is to develop content and algorithm existence requires data, the totality of the facts in the pleading must be evaluated. Thus, depending on the species of algorithm, the decision parameters, and the data ingested, some algorithms such as AI or machine learning carry more risk than others.[509]

If a high-risk algorithm, as identified in machine learning, AI, or recommender feedback loops, are in use, a reasonable rebuttable presumption should need to be overcome. Specifically, time, manner, and place within the material contribution, substantial assistance, and encouragement should be weighed against reasonable alternative design. Mitigating foreseeable risk is the focus. This will require a detailed analysis of algorithms to see if and how they are deployed or utilized.

Because parties are generally liable for the risks which make their conduct negligent to begin with, the analysis will require an examination of decision parameters, protocols, and internal operating procedures for the entities employing certain high-risk algorithms. These determinations create foreseeable risks. When a probable likelihood of conduct as a consequence of the algorithms or third parties relying on them, wittingly or unwittingly, is established, a reasonable and prudent standard would be applicable.

A sliding scale of rebuttable presumption might balance strict product liability and algorithm autonomy, such as AI and machine learning, against less liability for neutral non-biased algorithms. Considerations might include, readily discoverable coding design and protocols, identifying foreseeable risk within the content and development, risk-utility, warning and guidance with regular postings and updates, and preventive parameters to offset foreseeable risks. Until effective guardrails are implemented, the tempest will not abate.

---

508. *See Taamneh*, 598 U.S. at 507 (Jackson, J., concurring) (stating explicitly that the Court's "characterizations of the social media platforms and algorithms" is extremely limited to those allegations set forth in *Taamneh* and *Gonzalez* as applied to 18 U.S.C. § 2333(d)(2)); FED. R. CIV. P. 12(b)(6) (failure to state a claim upon which relief can be granted).

509. *Id.*