Fordham Law Review

Volume 89 | Issue 2 Article 9

2020

Programmed Defamation: Applying § 230

Michael R. Bartels
Fordham University School of Law

NOTES

PROGRAMMED DEFAMATION: APPLYING § 230 OF THE COMMUNICATIONS DECENCY ACT TO RECOMMENDATION SYSTEMS

*Michael R. Bartels**

Section 230 of the Communications Decency Act was originally intended to promote online innovation for the good faith moderation of interactive computer services. Since Congress enacted the statute, innovators, flourishing under statutory immunity, have been able to master the technological frontier, with most web traffic now consumed through highly curated and specialized feeds resembling a personal newspaper. The resulting free market of information is stronger than at any point in human history.

The new technological regime, however, has created another problem. Most of the content shared on these platforms originates from third parties, often anonymous or judgment-proof. Section 230, meanwhile, protects social media sites from liability as the publishers or speakers of such information. Although the websites develop the content by displaying it to others, they are generally not considered responsible for what comes from a third party. In this regime, defamation spreads quickly and easily, and plaintiffs have little legal recourse for relief.

This Note contends that social media sites' increasingly frequent use of recommendation algorithms to sort content for users should not be afforded § 230 immunity against liability. The creation of a recommendation feed, this Note argues, should be treated as a second creation point for which the interactive computer service should be held entirely responsible. This Note also contends that such a result may be achieved through judicial application of § 230 to the recommendation feeds as their own products, rather than amending the statute. Accordingly, this Note further contends that the

* J.D. Candidate, 2021, Fordham University School of Law; B.A., 2018, Providence College. I would like to thank Professor Sepehr Shahshahani for his support and commitment to my Note. I would also like to thank my parents, Robert A. Bartels and Ruth A. Bartels, and my brother, Daniel Bartels, for their love and support. I would also like to thank Professor Joseph Cammarano, whose advice and guidance at Providence College cultivated my early interest in social media and its interaction with the wider world. Lastly, I would like to thank Claire Abrahamson and the rest of the *Fordham Law Review* for their helpful comments and thorough editing.

allowance of liability against such systems could drive technological innovation toward addressing the dangers of online defamation and assign responsibility for the content on one's platform.

INTRODUCTION 653
I. SECTION 230 AND THE RECOMMENDATION FEED EXPLAINED 656
A. *Section 230* 656
1. Online Defamation Before § 230 656
2. Congress Writes an Exception 657
3. The Text of § 230 658
B. *Judicial Interpretation of § 230* 658
1. The Foundational Case: *Zeran* 658
2. Material Contribution as a Limit to Immunity 660
3. Other Cases Limiting § 230 662
4. Further Expansions of § 230 Immunity 664
C. *The Rise of the Recommendation Feed* 668
D. *Problems Created by § 230's Expansion* 672
II. POSSIBLE APPLICATIONS OF § 230 TO A RECOMMENDATION SYSTEM 676
A. *Parameters of the Discussion: Facebook as a Case Study* 676
B. *Arguments for Limiting § 230 Immunity* 677
1. Applying *Roommates.com* and *Accusearch* Directly to Facebook's Current Practices 677
2. Removing Protections by a Second Creation Point 679
3. Facebook Should Be Responsible for Its ICPs' Actions 682
C. *Arguments for Extending Immunity* 684
1. Applying *Roommates.com* to Individual User Posts 685
2. The ICP's Ultimate Responsibility for Offensive Content 686
3. Further Limits on § 230 Immunity Will Offend Speech Concerns 686
III. A PROPOSED REFINEMENT OF § 230 IMMUNITY 688
A. *Foundations of the Proposal* 688
B. *A Proposed Interpretation: Algorithm as ICP* 690
C. *The Broader Impact of This Proposal* 691
D. *Limits of This Proposal* 692
CONCLUSION 693

INTRODUCTION

Uzoma Igbonwa never learned the identities behind the fake profiles who defamed him.[1] A native Nigerian, Igbonwa found himself the target of two unknown Facebook users operating under aliases in May of 2016.[2] The users employed a Facebook page, operated by a third-party user with an alias, to spread falsehoods about Igbonwa; he alleged that moderators designed the forum to harm his reputation.[3] The posters alleged crimes and unflattering details about the plaintiff's personal life, accusing him of money laundering and beating his wife.[4] According to Igbonwa, by the time the District Court for the Northern District of California addressed his case more than two years later, unknown people still continued to humiliate him; the accusations were apparently so severe that they prompted threats to his life.[5] Despite his hardship, Igbonwa could not confront the accusations, as he could not find the people who posted the lies.[6] Unable to pursue claims against the unknown tortfeasors, Igbonwa joined Facebook and Mark Zuckerberg as defendants in his civil suit and advanced his case as a pro se plaintiff.[7]

Despite Igonbwa's efforts to save his reputation and clear his name, Facebook never had to address his case.[8] Because Igbonwa never alleged that Facebook created the content in question, § 230 of the Communications Decency Act of 1996[9] (CDA) barred all of his claims against the site and its creator.[10]

Section 230 is a federal statute that has been interpreted to give many online platforms broad protections against liability for their content, so long as another person posted that content.[11] In most other forms of media, such as newspapers, a publisher who repeats a third-party-sourced statement, even when attributing the statement to the original source, risks triggering liability for its veracity.[12] Publishers may even be held liable if they expressly deny the truth of a third-party libel but still reprint the statement.[13] This principle

1. Igbonwa v. Facebook, Inc., No. 18-cv-02027, 2018 U.S. Dist. LEXIS 173769, at *3 (N.D. Cal. Oct. 9, 2018).
2. *Id.* at *2.
3. *Id.* at *2–3.
4. *Id.*
5. *Id.* at *4.
6. *Id.*
7. *Id.* at *1.
8. *See generally id.*
9. Pub. L. No. 104-104, tit. V, 110 Stat. 56, 133–43 (codified as amended in scattered sections of 18 and 47 U.S.C.).
10. *Igbonwa*, 2018 U.S. Dist. LEXIS 173769, at *14.
11. ELLEN P. GOODMAN & RYAN WHITTINGTON, GERMAN MARSHALL FUND, SECTION 230 OF THE COMMUNICATIONS DECENCY ACT AND THE FUTURE OF ONLINE SPEECH 2 (2019), https://www.gmfus.org/publications/section-230-communications-decency-act-and-future-online-speech [https://perma.cc/9F5Y-QLDS].
12. RESTATEMENT (SECOND) OF TORTS § 578 (AM. L. INST. 1977).
13. *See* Benjamin C. Zipursky, *Online Defamation, Legal Concepts, and the Good Samaritan*, 51 VAL. U. L. REV. 1, 4 (2016); *see also* Cianci v. New Times Publ'g Co., 639 F.2d 54, 60–61 (2d Cir. 1980).

is known as the "republication rule."[14] The same rules apply to television and radio, which face even steeper penalties for their own defamation because of the high volume of information they process and distribute.[15]

Modern social media sites produce more data in seconds than a television network could possibly dream of distributing;[16] under the traditional rules for libel and defamation, such a site would trigger enormous liability for content that others post through their services.[17] Section 230, however, lifts the impossible burden of traditional tort law from the shoulders of many websites by crafting a large exception.[18] Under § 230, any "interactive computer service" (ICS) cannot be held liable as the "publisher or speaker" of content originating from a third-party "information content provider" (ICP).[19] An ICS can be defined extremely broadly[20] but is generally understood to encompass all websites,[21] including such diverse services as Facebook,[22] Google,[23] AOL,[24] and Grindr.[25] ICPs include "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."[26] Section 230 broadly prevents most websites from incurring liability for problematic third-party content by dictating that they shall not be recognized as its original publishers or speakers.[27] Such protections also

14. Zipursky, *supra* note 13, at 4.

15. Anthony Ciolli, *Chilling Effects: The Communications Decency Act and the Online Marketplace of Ideas*, 63 U. MIA. L. REV. 137, 143–44 (2008).

16. *See* Bernard Marr, *How Much Data Do We Create Every Day?: The Mind-Blowing Stats Everyone Should Read*, FORBES (May 21, 2018), https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#161f1c1860ba [https://perma.cc/H9RA-9C9G].

17. GOODMAN & WHITTINGTON, *supra* note 11, at 2.

18. *See* 47 U.S.C. § 230(c).

19. *See id.*

20. *Id.* § 230(f)(2) ("The term 'interactive computer service' means 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 or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions.").

21. Universal Commc'n Sys. v. Lycos, Inc., 478 F.3d 413, 419 (1st Cir. 2007) ("A web site, such as the Raging Bull site, 'enables computer access by multiple users to a computer server,' namely, the server that hosts the web site. Therefore, web site operators, such as Lycos, are providers of interactive computer services within the meaning of Section 230." (quoting 47 U.S.C. § 230(f)(2))).

22. Force v. Facebook, Inc., 934 F.3d 53, 64 (2d Cir. 2019).

23. Bennett v. Google, Inc., No. 16-cv-02283, 2017 WL 2692607, at *2 (D.D.C. June 21, 2017) ("[I]t has long been held across jurisdictions that Google qualifies as an interactive computer service under Section 230(c)(1).").

24. Zeran v. Am. Online, Inc. (*Zeran II*), 129 F.3d 327, 330 n.2 (4th Cir. 1997) ("The parties do not dispute that AOL falls within the CDA's 'interactive computer service' definition . . . ." (quoting 47 U.S.C. § 230(e)(3))).

25. Herrick v. Grindr LLC, 765 F. App'x 586, 590 (2d Cir. 2019) ("[W]e see no error in the district court's conclusion that Grindr is an ICS.").

26. 47 U.S.C. § 230(f)(3).

27. *Id.* § 230(c).

apply where the site takes pains to moderate content posted by others in a way usually considered editorial.[28]

Within these broad protections, courts have also recognized that sourcing tortious content to a third party does not always absolve a third-party host of responsibility.[29] Where websites materially contribute to third-party content,[30] actively solicit it,[31] or act in a way that cannot be classified as publishing or speaking,[32] circuit courts have found them responsible for content even though a third party still contributed it. While these cases may have been easier to explore when § 230 was enacted, modern advancements have spurred ICS interaction with user content far beyond the passive bulletin boards of the mid-1990s.[33] Perhaps the most advanced are modern social media sites, many of which use a recommendation system operated by a machine-learning algorithm to produce personalized compilations of user posts.[34]

The content at issue under § 230 is mind-boggling in scope. On Facebook alone, 510,000 comments are posted and 293,000 statuses are updated every minute.[35] Under current interpretations of § 230, every asserted fact, thought, musing, and proposition contained in this rapidly proliferating content exists under a modified defamation liability regime by virtue of its being hosted by an ICS.[36]

This Note explores whether the product of a machine-learning algorithm, such as that of social media giant, Facebook, should render an ICS its own ICP, therefore leaving this massive flood of content largely unprotected by § 230. This Note argues that recommendation systems should be held responsible as the ICPs of the additional layers of development they afford to defamatory content, such as inclusion in a recommendation feed. Part I summarizes the background of § 230 and its subsequent interpretation by the courts. It also describes how modern recommendation algorithms, widely employed by websites claiming § 230 protections, have changed from the

28. *Id*. § 230(c)(2) ("No provider or user of an interactive computer service shall be held liable on account of—(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected . . . .").

29. *See infra* Part I.

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

31. FTC v. Accusearch, Inc., 570 F.3d 1187, 1198–200 (10th Cir. 2009).

32. Oberdorf v. Amazon.com Inc., 930 F.3d 136, 153 (3d Cir. 2019) ("[T]o the extent that Oberdorf's negligence and strict liability claims rely on Amazon's role as an actor in the sales process, they are not barred by the CDA. However, to the extent that Oberdorf is alleging that Amazon failed to provide or to edit adequate warnings regarding the use of the dog collar, we conclude that that activity falls within the publisher's editorial function. . . . [T]hese failure to warn claims are barred by the CDA.").

33. *See infra* Part I.B.

34. *How News Feed Works*, FACEBOOK HELP CTR., https://www.facebook.com/help/1155510281178725 [https://perma.cc/EX3D-92VT] (last visited Oct. 3, 2020).

35. Marr, *supra* note 16.

36. *See* Zipursky, *supra* note 13, at 4.

simple online forums which prompted § 230's inclusion in the CDA. Part II describes the various arguments for and against continuing § 230 protections of websites that employ recommendation algorithms in constructing original feeds of third-party content for each individual user. Part III proposes that a recommendation system, using the example of the Facebook News Feed, can be ultimately responsible for its own content and, therefore, not subject to § 230 protections.

I. SECTION 230 AND THE RECOMMENDATION FEED EXPLAINED

This part discusses the origins of § 230 and the developments that have created the legal gray area in which recommendation systems reside. Part I.A discusses how § 230 was enacted. Part I.B elaborates on the ways in which subsequent court interpretations have expanded and limited its protections. Part I.C then provides context for the expanding framework of recommendation systems that could soon be influenced by those same interpretations.

A. Section 230

This section discusses defamation law's application online before § 230's passage, as well as the statute's enactment. While traditional defamation law punished republication of a libel,[37] § 230 gave some protections against that rule to online services that otherwise risked liability by innovatively moderating their sites.[38]

1. Online Defamation Before § 230

Section 230 began as a reaction to what contemporaries saw as the stifling effect of tort liability and, specifically, defamation liability, for third-party user content hosted on the internet. Before § 230, online platforms faced a dilemma in which they could only escape liability for third-party content posted on their sites by completely abstaining from moderation or curation of the content on their sites.[39] For businesses, this meant that no innovation to mold an online space was possible without massive risk.[40] For the general public, this meant that highly offensive content would permeate most of the rapidly developing internet.[41]

The tension was best exemplified by two cases that were decided before § 230's passing, reflecting opposite aspects of the same issue.[42] The first case, *Cubby, Inc. v. CompuServe, Inc.*,[43] protected a website that did not moderate its content from legal accountability for content it hosted.

37. *See* Cianci v. New Times Publ'g Co., 639 F.2d 54, 60–61 (2d Cir. 1980); *see also* Zipursky, *supra* note 13, at 4.
38. *See infra* Parts I.A.2–3.
39. GOODMAN & WHITTINGTON, *supra* note 11, at 4.
40. Ciolli, *supra* note 15, at 147–48.
41. *Id.* at 148.
42. GOODMAN & WHITTINGTON, *supra* note 11, at 4.
43. 776 F. Supp. 135 (S.D.N.Y. 1991).

CompuServe did not moderate its sites, which the Southern District of New York claimed protected the site from liability as a publisher.[44] In the second case, *Stratton Oakmont v. Prodigy Services Co.*,[45] the New York Supreme Court found a website liable for defamatory statements posted on a bulletin board it hosted. Because the site took pains to moderate content on the platform, the New York court reasoned that it had exercised editorial discretion, triggering liability for the content its fora hosted after curation.[46]

These two cases applied traditional notions of publishing to nontraditional platforms and created a lopsided incentive to turn a blind eye to destructive activity in the process.[47] Sites that took no action to protect their platforms from undesirable content completely insulated themselves from liability.[48] Conversely, a site that made a good faith effort to curate its content—even if its efforts could not have stopped tortious conduct—rendered itself a "publisher" vulnerable to suit.[49] The result, contemporaries feared, would have disincentivized curation of harmful content by the providers in the best position to stop it if adopted by other jurisdictions.[50]

2. Congress Writes an Exception

Congress enacted § 230 shortly after *Stratton* to eliminate that disincentive.[51] While Congress recognized that ICSs needed protection from liability to confidently police their content, it also recognized that many sites lacked the technology or resources to effectively do so.[52] Section 230's primary purpose was to curtail the negative effects of *Stratton* and other potential cases like it.[53] The provision's title, "Protection for 'Good Samaritan' blocking and screening of offensive material," says volumes about its purpose, which was to provide "Good Samaritan" protections to any

44. *See id*. at 135.

45. No. 31063/94, 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995), *superseded by statute*, Communications Decency Act of 1996, Pub. L. No. 104-104, § 509, 110 Stat. 56, 137–39, *as recognized in* Shiamili v. Real Est. Grp. of N.Y., Inc., 952 N.E.2d 1011 (N.Y. 2011).

46. *See id*. at *4.

47. *See* GOODMAN & WHITTINGTON, *supra* note 11, at 4.

48. *See Cubby*, 776 F. Supp. at 135.

49. *See Stratton*, 1995 WL 323710, at *4.

50. *See* 47 U.S.C. § 230(b)(3)–(5); *see also* 141 CONG. REC. 22,044–45 (1995) (statement of Rep. Charles Cox).

51. 47 U.S.C. § 230(c)(2)(A).

52. *See* Olivier Sylvain, *Discriminatory Designs on User Data*, KNIGHT FIRST AMEND. INST. (Apr. 1, 2018), https://knightcolumbia.org/content/discriminatory-designs-user-data [https://perma.cc/FWL5-59M2]. The *Stratton* court also recognized the impossibility of monitoring every third-party post with traditional editorial scrutiny but decided that it was irrelevant to applying defamation law. *See Stratton*, 1995 WL 323710, at *4 ("That such control is not complete . . . does not minimize or eviscerate the simple fact that PRODIGY has uniquely arrogated to itself the role of determining what is proper for its members to post and read on its bulletin boards.").

53. S. REP. NO. 104-230, at 194 (1996) (Conf. Rep.) ("One of the specific purposes of this section is to overrule *Stratton-Oakmont v. Prodigy* and any other similar decisions which have treated such providers and users as publishers or speakers of content that is not their own because they have restricted access to objectionable material.").

ICS that exercised modest discretion to restrict objectionable content.[54] Rather than punish the ICS with liability for attempting to create a family friendly environment, legislators wanted to bolster its efforts with rewards for taking on the risk of content moderation to help users in distress.[55]

3. The Text of § 230

Section 230 protects, among other things, "any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected."[56] While the statute eliminates incentives against content moderation, it imposes no obligation to moderate and, in fact, insulates an ICS from liability for failure to do so.[57]

Section 230's protections, importantly, only apply to an ICS when a third party provides its content.[58] An ICP provides content when it is "responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service."[59] If an ICS is responsible for its own content, then it could lose § 230 protections for its actions because it would then be responsible for the information connected to users by its service.[60]

B. Judicial Interpretation of § 230

This section discusses the various significant court cases interpreting § 230 and how they have expanded, or restrained, its broad protections. The discussion begins with the first case to expansively interpret § 230's protections for an ICS and contrasts it with the first case to articulate § 230's limitations. This section then identifies other key cases from circuit courts that have further developed these limitations in ways potentially applicable to a recommendation system.

1. The Foundational Case: *Zeran*

Ever since § 230's enactment, courts have interpreted the statute broadly, providing nearly unlimited immunity for ICSs when third parties do harm through their conduits. These broad interpretations began with *Zeran v.*

54. *See* 47 U.S.C. § 230(c); *see also* S. REP. NO. 104-230, at 194 ("This section provides 'Good Samaritan' protections from civil liability for providers or users of an interactive computer service for actions to restrict or to enable restriction of access to objectionable online material."); *see also Luke* 10:25–37 (King James) (originating the term "Good Samaritan" as one who risks harm by extending help to another in distress). The only obligation that § 230 imposes on an ICS is to notify users of commercially available parental controls, and even this provision only applies in a way the ICS deems appropriate. *See* 47 U.S.C. § 230(d).

55. *See* S. REP. NO. 104-230, at 194.

56. 47 U.S.C. § 230(c)(2)(A).

57. *See id.* § 230.

58. *Id.* § 230(c)(1).

59. *Id.* § 230(f)(3).

60. *See id.*; *see also supra* Part I.B.2.

America Online, Inc.[61] In this case, a third party had advertised plaintiff Kenneth Zeran as a sales contact for offensive shirts referring to the 1995 Oklahoma City bombings, including his first name and phone number in the post.[62] The advertisements were posted on a bulletin board feature of a website operated by AOL, similar to the Prodigy Services board in *Stratton*.[63] Because of the defamatory posting, Zeran received several calls to his home, including some threatening his life.[64] A radio station even discovered the posting and, believing it to be true, broadcast it to listeners, encouraging them to call Zeran's home.[65] Zeran contacted website representatives about the issue; they repeatedly said that the posting account would be taken down.[66] Although the initial post was eventually removed, it continued to reappear across AOL's site, generating a steady stream of harassment for Zeran.[67] At the height of the reaction, he received harassing phone calls every two minutes.[68] Zeran eventually responded by bringing a claim against AOL as a distributor, instead of a publisher, of the information, reasoning that § 230 only insulated "publisher[s] or speaker[s]" from liability for third-party content.[69]

The Fourth Circuit rejected Zeran's reasoning and granted ICS AOL immunity under § 230 from a negligence claim.[70] According to the Fourth Circuit, the statute barred all claims, not just defamation claims, stemming from liability for third-party content.[71] Like many future expansions of § 230 protections, the preference for a broad interpretation stemmed from concerns that tort law would stifle internet speech due to government interference.[72]

---

61. 129 F.3d 327 (4th Cir. 1997).

62. *Id.* at 329.

63. *Compare* Zeran v. Am. Online, Inc. (*Zeran I*), 958 F. Supp. 1124, 1127 (E.D. Va. 1997), *aff'd*, 129 F.3d 327 (4th Cir. 1997), *with* Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710, at *4 (N.Y. Sup. Ct. May 24, 1995), *superseded by statute*, Communications Decency Act of 1996, Pub. L. No. 104-104, § 509, 110 Stat. 56, 137–39, *as recognized in* Shiamili v. Real Est. Grp. of N.Y., Inc., 952 N.E.2d 1011 (N.Y. 2011).

64. *Zeran II*, 129 F.3d at 329.

65. *Id.*

66. *Id.*

67. *Zeran I*, 958 F. Supp. at 1127–28.

68. *Id.*

69. *Zeran II*, 129 F.3d at 331 (quoting 47 U.S.C. § 230(c)(1)).

70. *Id.* at 328 ("[Zeran] also contends that § 230 does not apply here because his claims arise from AOL's alleged negligence prior to the CDA's enactment. Section 230, however, plainly immunizes computer service providers like AOL from liability for information that originates with third parties.").

71. *Id.* at 330 ("By its plain language, § 230 creates a federal immunity to any cause of action that would make service providers liable for information originating with a third-party user of the service."). The court did, however, consider Zeran's negligence claim to be "indistinguishable from a garden variety defamation action." *Id.* at 332.

72. *See id.* at 330 ("The purpose of this statutory immunity is not difficult to discern. Congress recognized the threat that tort-based lawsuits pose to freedom of speech in the new and burgeoning Internet medium. The imposition of tort liability on service providers for the communications of others represented, for Congress, simply another form of intrusive government regulation of speech. Section 230 was enacted, in part, to maintain the robust nature of Internet communication and, accordingly, to keep government interference in the medium to a minimum.").

In interpretations after *Zeran*, courts have applied a three-part test to determine whether an ICS is protected under § 230, drawing from the language of the statute. The test asks three questions of the ICS at issue: "(1) whether the defendant qualifies as a provider of an 'interactive computer service,' (2) whether the asserted claims treat the defendant as a publisher or speaker of the information, and (3) whether the content was wholly provided by another 'information content provider.'"[73]

2. Material Contribution as a Limit to Immunity

The first case to provide an exception to the nearly unlimited expansions of immunity by *Zeran* was *Fair Housing Council of San Fernando Valley v. Roommates.com, LLC*.[74] The defendant in the case, Roommates.com, matched users with potential roommates based on personal descriptions and preferences expressed through a questionnaire.[75] The Fair Housing Councils of San Diego and San Fernando Valley brought suit against the matching service on the grounds that, along with other state laws, the questionnaire violated the Fair Housing Act[76] (FHA) by requiring users to disclose characteristics such as sex, family status, or sexual orientation.[77] By requesting this information, plaintiffs argued, Roommates.com expressed discriminatory intent and discriminated in fact against users by funneling them toward or away from profiles based on the controversial characteristics.[78]

The Ninth Circuit ultimately held that an ICS could act in a way that constituted codevelopment of third-party content and thus surrender § 230 protections.[79] The court reasoned that because Roommates.com had required the requested information of its users as a condition of participating in the matchmaking service, it was responsible in part for the content, even though the content consisted of third-party inputs.[80] Roommates.com's search engine similarly fell beyond § 230 protections; because the site steered users toward choices based on the search engine's unlawful criteria, it "force[d]

73. Catherine Tremble, Note, *Wild Westworld: Section 230 of the CDA and Social Networks' Use of Machine-Learning Algorithms*, 86 FORDHAM L. REV. 825, 847 (2017) (footnote omitted); *see, e.g.*, Doe v. Internet Brands, Inc., 824 F.3d 846, 850 (9th Cir. 2016); *see also* Force v. Facebook, Inc., 934 F.3d 53, 64 (2d Cir. 2019); Herrick v. Grindr LLC, 765 F. App'x 586, 589 (2d Cir. 2019); FTC v. Accusearch, Inc., 570 F.3d 1187, 1196 (10th Cir. 2009); Barnes v. Yahoo!, Inc., 570 F.3d 1096, 1100–01 (9th Cir. 2009).

74. 521 F.3d 1157 (9th Cir. 2008) (en banc).

75. *Id.* at 1164.

76. Title VIII of the Civil Rights Act of 1968, Pub. L. No. 90-284, 82 Stat. 81 (codified as amended in scattered sections of the U.S.C.).

77. *Roommates.com*, 521 F.3d at 1164.

78. *Id.* at 1165.

79. *See id.* at 1166.

80. *Id.* ("When a business enterprise extracts such information from potential customers as a condition of accepting them as clients, it is no stretch to say that the enterprise is responsible, at least in part, for developing that information. For the dissent to claim that the information in such circumstances is 'created solely by' the customer, and that the business has not helped in the least to develop it strains both credulity and English." (citation omitted) (quoting *id*. at 1181–82 (McKeown, J., dissenting))).

users to participate in its discriminatory process" and was "designed to achieve illegal ends."[81]

The holding in *Roommates.com* had extremely narrow application. The codevelopment activity only included prepopulating an answer field and designing a search engine to achieve an illegal goal.[82] The codevelopment in that case only created liability when the site made direct, material contributions to the specific illegal aspects of the content at hand.[83] The Roommates.com search engine triggered liability for illegal content because it was designed to classify search results by protected characteristics; another search engine that did not limit search results based on unlawful characteristics would not trigger the same liability as Roommates.com's did.[84] As the Ninth Circuit understood it, Google exemplified this unbiased standard because its search engine sorted results independently of any legally restricted preference.[85]

The Ninth Circuit here established an important limit on § 230's expanding application in *Roomates.com* by establishing a concrete ceiling as to how far the definition of "publisher" could be expanded.[86] This limit became stricter in *Roomates.com*'s aftermath, diminishing the importance of the third prong of the *Zeran* analysis: whether an ICS served as its own ICP.[87] Specifically, *Roommates.com* established that an ICS could still, in some situations, be held liable under the third prong of the analysis if it created, or codeveloped, the specific illegal aspects of third-party content.[88] The court likened development to solicitation and held that "unlawful questions solicit (a.k.a. 'develop') unlawful answers."[89] In other words, an ICS materially contributes to illegal content when it conditions participation in the ICS on its provision.[90] This was why Roommates.com was not responsible for content submitted into a blank text box but was responsible for a prepopulated answer field.[91]

The Ninth Circuit outlined additional criteria on search engines in its decision:

> Roommate designed its search system so it would steer users based on the preferences and personal characteristics that Roommate itself forces

81. *Id.* at 1167.
82. *See id.* at 1166–67.
83. *See id.* at 1167–68 ("We believe that both the immunity for passive conduits and the exception for co-developers must be given their proper scope and, to that end, we interpret the term 'development' as referring not merely to augmenting the content generally, but to materially contributing to its alleged unlawfulness. In other words, a website helps to develop unlawful content, and thus falls within the exception to section 230, if it contributes materially to the alleged illegality of the conduct.").
84. *Id.* at 1167.
85. *Id.*
86. *See id.* at 1166.
87. Tremble, *supra* note 73, at 845–46.
88. *See Roommates.com*, 521 F.3d at 1167–68.
89. *See id.* at 1166.
90. *See id.*
91. *See id.* at 1174 n.38.

> subscribers to disclose. . . . Councils charge that limiting the information a subscriber can access based on that subscriber's protected status violates the Fair Housing Act and state housing discrimination laws. . . . If such screening is prohibited when practiced in person or by telephone, we see no reason why Congress would have wanted to make it lawful to profit from it online.[92]

After *Roommates*, an ICS may be held liable not only for soliciting illegal content but also for designing a search engine to structure illegal results.[93]

3. Other Cases Limiting § 230

After *Roommates.com*, another major case, *FTC v. Accusearch Inc.*,[94] arose in the Tenth Circuit and tackled the development prong of the *Zeran* analysis. Defendant Accusearch operated a website, Abika.com, which allowed users to search databases of information mostly provided by third-party ICPs.[95] The site styled itself as an intermediary that connected paying customers with third-party researchers' findings.[96] One of these databases involved personal telephone records.[97] Because of restrictions imposed by the Telecommunications Act of 1996,[98] anyone who obtained the records had presumably done so illegally.[99] The Federal Trade Commission (FTC) accordingly brought an unfair practice claim against Accusearch for disclosing telephone records.[100]

The Tenth Circuit held that Accusearch developed content when it "specifically encourage[d] development of what is offensive about the content" by making it more usable, visible, or active.[101] Resting its decision on the third prong of the *Zeran* analysis,[102] the court investigated the etymology of the word "development," the action that renders an entity an ICP of content, and found its "core meaning" to be the "drawing out" of information to expand usefulness.[103] The court asked whether the content was developed through Abika.com and who was responsible for that development.[104] Exposing information to public view, the court found, qualified as "development" of it.[105] Taken in full context, the court held that

92. *Id.* at 1166.
93. *See id.*
94. 570 F.3d 1187 (10th Cir. 2009).
95. *Id.* at 1191.
96. *Id.*
97. *Id.*
98. Pub. L. No. 104-104, 110 Stat. 56 (codified as amended in scattered sections of 15, 18, and 47 of the U.S.C.).
99. *Accusearch*, 570 F.3d at 1192.
100. *Id.* at 1190.
101. *Id.* at 1198–99.
102. *Id.* at 1197.
103. *Id.* at 1198.
104. *Id.*
105. *Id.*

a website collecting and publicly displaying legally problematic content *because* of its legally problematic nature becomes the ICP of that content.[106]

In *Accusearch*, the public display of the telephone records constituted development under the Tenth Circuit definition.[107] The site had knowingly solicited the phone information from researchers obtaining the data illegally and constructed a portion of Abika.com to develop the illegal data as it came in.[108] The court therefore used a slightly broader theory than that of *Roommates.com* and ruled that the site was an ICP.[109]

To distinguish exactly where Accusearch went wrong in developing its content, the court compared Abika.com with the more recent actions of a familiar defendant: AOL.[110] In *Ben Ezra, Weinstein, & Co. v. American Online, Inc.*,[111] AOL, also the defendant in *Zeran*, returned triumphant in a § 230 defense against claims of defamation and negligence.[112] The plaintiff alleged that AOL had published faulty stock prices and share volumes associated with Ben Ezra on three separate occasions.[113]

AOL developed the information, but it did not seek out the faulty information; it even requested corrections from the third-party vendor providing the content.[114] For this reason, while both Accusearch and AOL developed the content on their sites, Accusearch was responsible for its offensiveness, whereas AOL in *Ben Ezra* was not.[115]

The combined effect of *Accusearch* and *Roommates.com* has triggered further liability for other defendants in the following years. In *FTC v. LeadClick Media, LLC*,[116] the FTC sued an affiliate marketing network for unfair trade practices.[117] The network, LeadClick, had been connecting affiliate, LeadSpa, with fake news sites to advertise LeadSpa products.[118] LeadClick had even suggested alterations to the fake news content that would meet requests on the part of both LeadClick and LeadSpa.[119] LeadClick attempted a § 230 defense by alleging that its affiliates published the deceptive statements as independent ICPs.[120] However, the court found that because LeadClick helped to develop the fake news sites with which it

106. *Id.* at 1198–99 ("[O]ne is not 'responsible' for the development of offensive content if one's conduct was neutral with respect to the offensiveness of the content (as would be the case with the typical Internet bulletin board).").
107. *Id.* at 1198.
108. *Id.*
109. *Id.* at 1200.
110. *Id.* at 1199.
111. 206 F.3d 980 (10th Cir. 2000).
112. *See id.* at 983.
113. *Id.*
114. *Id.* at 985.
115. *Accusearch*, 570 F.3d at 1199.
116. 838 F.3d 158 (2d Cir. 2016).
117. *See id.* at 167.
118. *Id.* at 171.
119. *Id.* at 167.
120. *Id.* at 175.

worked, it was responsible for the deceptiveness of the content it published and was not merely a neutral assistant or conduit.[121]

4. Further Expansions of § 230 Immunity

Even as courts have established boundaries to § 230 defenses, they have also continued its broad application. In *Doe No. 1 v. Backpage.com, LLC*,[122] the First Circuit evaluated whether Backpage.com, a website that featured advertisements for prostitutes, unlawfully promoted and profited from sex trafficking.[123] The appellants were three victims of human trafficking who sued pseudonymously, stating that Backpage.com hosted online prostitution advertisements featuring them.[124] The appellants alleged that the advertisements, which their traffickers posted on Backpage.com, helped customers find and rape them over a combined 1900 times.[125] They therefore claimed both that Backpage.com encouraged sex trafficking through its policies and that it illegally profited off of its proliferation.[126]

The First Circuit held for Backpage.com and stated that § 230 barred appellants' claims.[127] The actions of Backpage.com, the court held, constituted traditional publisher functions that barred claims arising from their exercise.[128] Even though the court acknowledged that Backpage.com's rules facilitated illegal conduct, the prevailing prohibition of claims based on publishing functions occupied the court's attention.[129] The court did not, however, address the factual question of whether Backpage.com actually encouraged sex trafficking through its website or whether it was responsible for the development of the alleged sex traffickers' advertisements.[130]

In another case, *Herrick v. Grindr LLC*,[131] the Second Circuit applied § 230 to Grindr, an application that "matches users based on their interests and location."[132] Appellant Matthew Herrick faced a harassment campaign in which his ex-boyfriend posted a fake profile of Herrick on Grindr and directed Grindr users to go to his home and workplace.[133] Herrick then sued

121. *Id.* at 176.
122. 817 F.3d 12 (1st Cir. 2016).
123. *Id.* at 16.
124. *Id.*
125. *Id.* at 17.
126. *Id.* at 16–17. The site was alleged to encourage sex trafficking by scrubbing metadata from photos, not requiring the provision of a phone number or email verification, and employing a content filter for prohibited terms that could be easily circumnavigated. *Id*. In the victims' cases, the advertisements posted also encouraged anonymous payments, contained altered phone numbers, and used coded language signaling the availability of underage girls. *Id*. at 17.
127. *Id.* at 24.
128. *Id.* at 22.
129. *See id.* ("We hold that claims that a website facilitates illegal conduct through its posting rules necessarily treat the website as a publisher or speaker of content provided by third parties and, thus, are precluded by section 230(c)(1).").
130. *Compare id.*, *with* FTC v. Accusearch, Inc., 570 F.3d 1187 (10th Cir. 2009).
131. 765 F. App'x. 586 (2d Cir. 2019).
132. *Id.* at 588.
133. *Id.*

Grindr under several causes of action[134] on the theory that Grindr's website lacked features to prevent impersonation or other dangerous misconduct such as that which he experienced.[135] Grindr moved to dismiss all but the copyright infringement claims based on § 230 protections.[136]

The Second Circuit sided with Grindr.[137] The court reasoned that Grindr was an ICS that collected and displayed information provided exclusively by other ICPs, namely its users.[138] Grindr also successfully defeated the appellant's claims on the grounds that it could not be held liable for failing to remove offensive content, i.e., exercising its ordinary editorial functions.[139] The case again did not discuss whether Grindr was responsible for developing the content at issue[140] but explained that it could not be held responsible either way, as it only provided neutral assistance "in the form of tools and functionality available equally to bad actors and the app's intended users."[141]

The Sixth Circuit presented one of the broadest examples of § 230 protections in *Jones v. Dirty World Entertainment Recordings LLC*.[142] The Sixth Circuit characterized the website in question, www.TheDirty.com, as "a user-generated tabloid primarily targeting nonpublic figures."[143] The plaintiff in that case was the target of several defamatory posts and brought an action in federal court against the site and its owners for defamation, libel per se, false light, and intentional infliction of emotional distress.[144]

The Dirty generated very little of its own content, instead relying primarily on user-generated content for the site.[145] The moderators' curation of the site only entailed deletion to ensure "nudity, obscenity, threats of violence, profanity, and racial slurs are removed" in a method not dissimilar to that in *Prodigy Services*.[146] Moderators would occasionally mark a post with their own brief comments, distinguished from user-generated content with boldface.[147]

134. These included "negligence, deceptive business practices and false advertising, intentional and negligent infliction of emotional distress, failure to warn, and negligent misrepresentation." *Id.* An amended complaint also added causes of action for "products liability, negligent design, promissory estoppel, fraud, and copyright infringement." *Id.* at 589.
135. *Id.* at 588.
136. *Id.* at 589.
137. *Id.* at 591.
138. *Id.* at 589–90.
139. *Id.* at 591.
140. *Compare id.*, *with* FTC v. Accusearch, Inc., 570 F.3d 1187 (10th Cir. 2009).
141. *Herrick*, 765 F. App'x. at 591.
142. 755 F.3d 398 (6th Cir. 2014).
143. *Id.* at 401.
144. *Id.* at 401–02.
145. *Id.* at 402.
146. *Compare id.* at 403, *with* Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710, at *5 (N.Y. Sup. Ct. May 24, 1995), *superseded by statute*, Communications Decency Act of 1996, Pub. L. No. 104-104, § 509, 110 Stat. 56, 137–39, *as recognized in* Shiamili v. Real Est. Grp. of N.Y., Inc., 952 N.E.2d 1011 (N.Y. 2011).
147. *Jones*, 755 F.3d at 403.

The Sixth Circuit, after realizing that The Dirty had developed the defamatory content at issue, then asked whether it was responsible for the development (the same two questions posed in *Accusearch*).[148] Because the site was neutral as to the content it asked parties to submit, never asked users to submit defamatory information explicitly, and never suggested that it only published defamation, the Sixth Circuit held that it was not responsible for the defamation in user posts.[149] While the court found The Dirty responsible for its own comments, the comments themselves contributed nothing new to the defamation that third parties had already volunteered.[150]

Most recently, the Second Circuit upheld § 230 protections for Facebook in *Force v. Facebook, Inc*.[151] That case began with five attacks by Hamas on Americans in Israel between 2014 and 2016, resulting in the deaths of four American citizens.[152] The appellants, which included one survivor and representatives of the estates of the others, brought multiple claims against the site,[153] alleging that Facebook aided Hamas in the attacks when its algorithm helped Hamas spread incitement to potential terrorists and when the site failed to enforce official guidelines prohibiting the use of Facebook by designated foreign terrorist organizations.[154] Facebook, the appellants argued, had developed the terrorists' online network.[155]

The appellants attempted several new arguments to demonstrate how Facebook's role exceeded that of a publisher. They argued that Facebook's "matchmaking" role between users and content did not constitute a form of publishing[156] and that the Facebook algorithm's automation of traditional editorial functions removes it from the traditional scope of publishing.[157] They also argued that Facebook had become its own ICP by "developing" Hamas propaganda and making it more visible to users through the news feed.[158]

The Second Circuit disagreed with the appellants' analysis.[159] Matchmaking, as long as it was performed through tools neutral to the alleged offensiveness of the content, replicated "an essential result of publishing" in the eyes of the court,[160] although the Second Circuit acknowledged that the matchmaking was possibly made more efficient through Facebook's

148. *See id.* at 409.
149. *Id.* at 415–16.
150. *Id.* at 416–17.
151. 934 F.3d 53, 63 (2d Cir. 2019).
152. *Id.* at 57–58.
153. *Id.* at 61 ("Plaintiffs claimed that . . . Facebook was civilly liable for aiding and abetting Hamas's acts of international terrorism; conspiring with Hamas in furtherance of acts of international terrorism; providing material support to terrorists; and providing material support to a designated foreign terrorist organization.").
154. *See id.* at 59.
155. *Id.* at 61.
156. *Id.* at 65.
157. *Id.* at 67.
158. *Id.*
159. *Id.* at 71.
160. *Id.* at 66.

system.[161] The use of an algorithm, the court explained, was irrelevant so long as a third party provided "the essential published content."[162] As for the argument that Facebook was its own ICP, the court imposed its own interpretation of the "material contribution test" implemented in *LeadClick* and *Roommates.com*.[163] Unlike the Tenth Circuit in *Accusearch*, the Second Circuit cited a D.C. Circuit case[164] to assert that a website did not "cross the line into content development" by making information available for public display.[165] Even if it developed content by displaying it, use of a neutral algorithm would protect it from liability: "Merely arranging and displaying others' content to users of Facebook through such algorithms—even if the content is not actively sought by those users—is not enough to hold Facebook responsible as the 'develop[er]' or 'creat[or]' of that content."[166]

As the Second Circuit case indicates, recommendation systems, such as those employed by Facebook, are generally insulated from liability by § 230.[167] There is, however, some discrepancy within the case law interpreting § 230 as to the third prong of the *Zeran* analysis[168] and what line renders an ICS its own ICP of offensive content. More specifically, courts disagree on what constitutes "development." In *Accusearch*, the Tenth Circuit held that websites develop content by making it more available than it otherwise would be. The question there is whether a site is responsible for that development in some way, which includes indirect measures, such as solicitation,[169] as well as direct measures, such as creating the choices for users to select.[170] The view favoring broader § 230 application, recently held by the Second Circuit, requires direct, material contribution to the illegal aspects of the content,[171] similar to the direct micromanagement of fake news sites in *LeadClick*.[172]

While courts have begun applying these standards to AOL bulletin boards and roommate matchmaking sites, the nature of the internet entities raising § 230 defenses has changed dramatically. Whether development stems only from direct contribution or comes from public display and creates a question of responsibility, the answers to these questions require a broader understanding of the recommendation algorithms ultimately handling the content at issue. Aside from a surface-level aesthetic resemblance to the

161. *Id.* at 67.
162. *Id.*
163. *Id.* at 68 ("[A] defendant will not be considered to have developed third-party content unless the defendant directly and 'materially' contributed to what made the content itself 'unlawful.'").
164. Marshall's Locksmith Serv. Inc. v. Google, LLC, 925 F.3d 1263, 1266 (D.C. Cir. 2019).
165. *Force*, 934 F.3d at 69.
166. *Id.* at 70.
167. *See id.*
168. *See supra* note 71.
169. FTC v. Accusearch, Inc., 570 F.3d 1187, 1198–99 (10th Cir. 2009).
170. *See* Fair Hous. Council v. Roommates.com, 521 F.3d 1157, 1166, 1174 n.38 (9th Cir. 2008) (en banc).
171. *Force*, 934 F.3d at 68.
172. FTC v. LeadClick Media, LLC, 838 F.3d 158, 176 (2d Cir. 2016).

Prodigy Services bulletin board, recommendation feeds of the 2010s employ vastly different strategies that may have greater relevance for § 230.

C. The Rise of the Recommendation Feed

Today's algorithms are far from entirely passive conduits and can make extremely personal inferences about users' lives. Indeed, service providers of all kinds, even beyond the internet, have been able to exploit this abundance of data in various ways as an emerging market.[173] Retailers such as Target have been able to leverage this data to build deep insights into their customers' lives. For instance, a recommendation algorithm even as far back as 2012 could predict a customer's pregnancy based on shopping habits.[174]

One group of online services that has grown around these insights is social media providers, which can be defined as the subcategory of ICSs whose services allow for the creation and distribution of third-party content to network socially.[175] While the phrase describes a broad range of sites, a growing subcategory of these, such as Facebook, employs a machine-learning algorithm[176] that bases its content displays on how likely a user is to engage with a given post in that display, among other factors.[177] The algorithmic sorting most often manifests itself in a displayed compilation of relevant information.[178] Examples of these recommendation systems include the Facebook News Feed[179] and Twitter feeds.[180] Unlike the chat room services that were protected in cases like *Stratton*, which functioned similarly to online bulletin boards, these website features compose an independent compilation of resources based on your likelihood to engage with them.[181]

What distinguishes a modern social media algorithm from the Prodigy Services bulletin board, among other things, is the more advanced and

---

173. *See* JOHN GANTZ & DAVID REISEL, INT'L DATA CORP., THE DIGITAL UNIVERSE IN 2020: BIG DATA, BIGGER DIGITAL SHADOWS, AND BIGGEST GROWTH IN THE FAR EAST 16 (2012), https://www.speicherguide.de/download/dokus/IDC-Digital-Universe-Studie-iView-11.12.pdf [https://perma.cc/DX59-3GLV] ("Big Data is going to be a big boon for the IT industry. Web sites that gather significant data need to find ways to monetize this asset. . . . Further, companies that deliver the most creative and meaningful ways to display the results of Big Data analytics will be coveted and sought after.").

174. *See* Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES MAG. (Feb 16, 2012), https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?searchResultPosition=3 [https://perma.cc/7CZX-CYX2].

175. *Social Media*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/social%20media [https://perma.cc/75TR-8PJG] (last visited Oct. 3, 2020).

176. *See* Tremble, *supra* note 73, at 838–40.

177. *See* AJ Agrawal, *What Do Social Media Algorithms Mean for You?*, FORBES (Apr. 20, 2016), https://www.forbes.com/sites/ajagrawal/2016/04/20/what-do-social-media-algorithms-mean-for-you/#17720dcfa515 [https://perma.cc/HLC4-7E3N].

178. *See id.*

179. *How News Feed Works*, *supra* note 34.

180. *See Twitter Ads Targeting*, TWITTER BUS. https://business.twitter.com/en/advertising/targeting.html [https://perma.cc/3MGP-R9VL] (last visited Oct. 3, 2020).

181. *See* Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710, at *4 (N.Y. Sup. Ct. May 24, 1995), *superseded by statute*, Communications Decency Act of 1996, Pub. L. No. 104-104, § 509, 110 Stat. 56, 137–39, *as recognized in* Shiamili v. Real Est. Grp. of N.Y., Inc., 952 N.E.2d 1011 (N.Y. 2011); *How News Feed Works*, *supra* note 34.

complex programming of the ICS's interactive service.[182] Consequently, it is necessary to explore, on a basic level, how current advances differ from the message boards § 230 originally contemplated.

The Prodigy Services model provides context for services legislators might have expected to protect, as well as the actions to which they may have intended § 230 to apply.[183] Prodigy Services operated several bulletin boards where almost two million users communicated through posts.[184] The site also contracted "Board Leaders" who also posted on the bulletin boards.[185] The Board Leaders enforced community guidelines (which the plaintiffs labeled as standards of editorial control) via an "emergency delete function" by which the leaders could remove problematic content.[186] The bulletin board also employed a rudimentary program that automatically screened postings for offensive language.[187] Aside from the posts of board leaders, Prodigy Services made no positive contributions to its users' posts; its only "editorial" actions were blocking offensive content and deleting posts that violated user guidelines.[188]

Modern social media differs from the bulletin boards of Prodigy's day in several ways but perhaps most prominently through the implementation of algorithmically generated content.[189] Instead of users posting material to a neutral bulletin board, an algorithm sorts the massive volume of content available to display and creates a custom feed based on what each individual user is most likely to engage with.[190] The generated feed is unique for every user, based on their own personal interactions with the site.[191]

All such feeds, such as YouTube's "recommended" tabs, the Facebook News Feed, or Twitter's timeline, so long as they are managed by a computer program, fall into a broad category called recommendation systems.[192] These systems can themselves be sorted into three categories;[193]

182. *See supra* note 176 and accompanying text.
183. *See* GOODMAN & WHITTINGTON, *supra* note 11, at 4.
184. *Stratton*, 1995 WL 323710, at *3.
185. *Id.*
186. *Id.* at *3, *5–6.
187. *Id.* at *5.
188. *See id.*
189. *See supra* note 176 and accompanying text.
190. *See How News Feed Works*, *supra* note 34.
191. Will Oremus, *Who Really Controls What You See in Your Facebook Feed—and Why They Keep Changing It*, SLATE (Jan. 3, 2016, 8:02 PM), https://www.slate.com/articles/technology/cover_story/2016/01/how_facebook_s_news_feed_algorithm_works.html [https://perma.cc/3FRM-22LK].
192. *See generally* Maxim Naumov et al., Deep Learning Recommendation Model for Personalization and Recommendation Systems (May 31, 2019) (unpublished manuscript), https://arxiv.org/pdf/1906.00091.pdf [https://perma.cc/LNB7-2P5H].
193. Gabriel Chartrand et al., *Deep Learning: A Primer for Radiologists*, 37 RADIOGRAPHICS 2113, 2114–15 (2017). Chartrand et al. list four categories, but the fourth category, deep learning, is a subset of the third one, representation learning. For the purposes of this Note, the distinction between representation learning and deep learning is immaterial. As applied to a recommendation system, both refer to a mostly or fully automated process that sorts content from third-party ICPs into categories of its own making.

understanding the basis of these categories is necessary to understand how § 230 applies to them.

The first category, generic rules-based artificial intelligence (AI), consists of a stable, unchanging computer program that turns an input into an output.[194] An example of such a program would be an online bulletin board or even a screening device for posts, such as the programs used by the defendants in *Stratton*.[195] The court in *Stratton* held that such a program would normally constitute the exercise of editorial discretion, triggering liability as a publisher in a defamation suit,[196] but the context and legislative history of § 230 show that the law was meant to eliminate that liability.[197] Such a program would not, at least in theory, itself violate the third prong of *Zeran* by any court's standard because, while possibly "developing" the content by generating the display,[198] it does not materially contribute to the illegal aspect of a defamatory post.[199] Unless the rules-based AI were somehow programmed to itself contain offensive elements (as was the case in *Roommates.com*),[200] its programmer would not be responsible for the offensiveness of third-party content it displayed.

The second category, classic machine learning, is a subset of AI that changes over time to most efficiently perform a task.[201] In a recommendation system, this usually entails a human identifying trends in data and feeding those data to an AI.[202] When users engage with that data, for example, by "liking" a social media post, the machine-learning AI can classify them concerning their responses to the data, altering an output to fit their past preferences in future iterations of the same process.[203] Apple News exemplifies this system on a basic level.[204] Editors handpick and categorize stories and feed them into the recommendation system connected to an application on user devices.[205] The recommendation system is a classic

194. *See id.* at 2114.

195. *Compare* Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710, at *5 (N.Y. Sup. Ct. May 24, 1995), *superseded by statute*, Communications Decency Act of 1996, Pub. L. No. 104-104, § 509, 110 Stat. 56, 137–39, *as recognized in* Shiamili v. Real Est. Grp. of N.Y., Inc., 952 N.E.2d 1011 (N.Y. 2011) (explaining how Prodigy's screening program filtered user posts for explicit appearances of objectionable words), *with* Chartrand et al., *supra* note 193, at 2115 (describing a rules-based AI as entirely composed of an explicit computer program).

196. *Stratton*, 1995 WL 323710, at *4 ("By actively utilizing technology and manpower to delete notes from its computer bulletin boards on the basis of offensiveness and 'bad taste', for example, PRODIGY is clearly making decisions as to content . . . and such decisions constitute editorial control.").

197. *See* S. REP. NO. 104-230, at 194 (1996) (Conf. Rep.).

198. FTC v. Accusearch, Inc., 570 F.3d 1187, 1198 (10th Cir. 2009).

199. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1166–67 (9th Cir. 2008) (en banc).

200. *See id.*

201. *See* Chartrand et al., *supra* note 193, at 2115.

202. *Id.*

203. *See* Naumov et al., *supra* note 192.

204. *See generally* APPLE NEWS+, https://www.apple.com/apple-news/ [https://perma.cc/TK88-ZA6J] (last visited Oct. 3, 2020).

205. *Id.*

machine-learning algorithm;[206] as users sort through stories and expose preferences for one category over another, the algorithm "learns" about those preferred categories and prioritizes them in future recommendations.[207]

The third category, representation learning, operates as a subset of classic machine learning in which the AI self-categorizes raw user data to best obtain a predetermined goal without independent human assistance.[208] While third parties can still input and categorize data, the algorithm not only sorts its own data based on user preferences but also divines those user preferences based on its own interpretation.[209]

In *Force*, the appellants alleged, and the Second Circuit agreed, that the Facebook algorithm amounted to an automation of decision-making by the platform.[210] As the dissent noted, the categories by which Facebook sorts content are created and modified by the algorithm to at least some degree to best achieve its intended goals.[211] Facebook does in fact generate its advertising categories on its own, although it can still manually tweak those categories when they promote offensive content.[212] As a recommendation system that creates its own categories as gleaned from user input to best construct its News Feed, Facebook's algorithm is a form of representation learning.[213]

Specifically, Facebook's programmers want to employ deep learning regarding most user-contributed content.[214] A subset of representation learning, deep learning identifies simple features of third-party content as features within the more complex categories that the algorithm creates.[215] Commonly used in image classification or facial recognition programs, deep learning works by "picking out edges first, then circles, then faces."[216] In

206. *Id.* ("Apple News only uses on-device intelligence to recommend stories . . . .").

207. *Id.* ("As you read, Apple News gets a better read on your interests, then suggests relevant stories.").

208. *See* Chartrand et al., *supra* note 193, at 2114.

209. *See id.*

210. Force v. Facebook, Inc., 934 F.3d 53, 68 (2d Cir. 2019).

211. *Id.* at 85 (discussing how Facebook's algorithm created the category "Hitler did nothing wrong," among others, to maximize content engagement, and automatically modified the anti-Semitic category to include Second Amendment supporters); *see also* Julia Angwin et al., *Facebook Enabled Advertisers to Reach 'Jew Haters*,*'* PROPUBLICA (Sept. 14, 2017, 4:00 PM), https://www.propublica.org/article/facebook-enabled-advertisers-to-reach-jew-haters [https://perma.cc/HG2N-J2PV].

212. *See* Angwin et al., *supra* note 211.

213. *See* Chartrand et al., *supra* note 193, at 2114.

214. *See generally* Naumov et al., *supra* note 192. Facebook has recently implemented new features to promote reliable news sources over known sources of fake news. The changes do not constitute a change to the central Facebook News Feed but rather the creation of an alternative "Facebook News" feature that exists separately from the recommendation system that this Note discusses and is, therefore, outside the scope of this Note. *See* Gerry Smith & Sarah Frier, *Facebook Launches News Section to Compensate Publishers*, BLOOMBERG (Oct. 25, 2019, 6:00 AM), https://www.bloomberg.com/news/articles/2019-10-25/facebook-creates-news-section-to-compensate-restive-publishers [https://perma.cc/57AG-9S6A].

215. *See* Chartrand et al., *supra* note 193, at 2115.

216. Gideon Lewis-Kraus, *The Great A.I. Awakening*, N.Y. TIMES MAG. (Dec. 14, 2016), https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html [https://perma.cc/B62Y-BRK5].

one of the earlier examples of true deep learning, a Google AI, after viewing millions of YouTube still frames, had taught itself what cats are and how to identify them.[217] The AI had never been previously educated on what a cat is.[218] Since its initial AI achievement, Google also managed to apply deep learning to Google Translate, better improving recommended translations to suit what their users are most likely searching for.[219]

D. Problems Created by § 230's Expansion

Section 230 has been hailed by supporters as "the twenty-six words that created the internet."[220] The statute aimed, in its own words, to promote the development of the internet and protect free market innovation from the stifling effects of defamation liability, as seen in *Stratton*.[221] In many respects, these policy goals have been achieved. Internet usage has climbed from 52 percent of the U.S. population in 2000 to about 90 percent in 2019.[222] Racial gaps in pure internet access have also narrowed significantly, though they persist in subtler forms, such as smartphone dependence.[223] While only about 1 percent of U.S. adults had home broadband internet access in 2000, almost three-quarters do today and an additional one in five accesses home internet exclusively through a smartphone.[224] Social media sites, consisting only of nascent bulletin boards when § 230 was first passed, produce and collect massive amounts of data, which has opened up new economic opportunities both in the United States and abroad.[225] Facebook is now arguably the second largest news provider in the United States; in the United Kingdom, Facebook, Google, and Twitter all fall into the top ten.[226] The shift to online news consumption has effectively broken the brand loyalty bubbles of traditional news sources, the ICPs of the pre-internet world.[227] With the collapse of those bubbles, internet users have accessed an even

217. *Id.*; *see also* Quoc V. Le et al., *Building High-Level Features Using Large Scale Unsupervised Learning*, *in* PROCEEDINGS OF THE 29TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING 507, 507 (2012) ("The focus of this work is to build *high-level*, class specific feature detectors from *unlabeled* images.").

218. *See* Lewis-Kraus, *supra* note 216.

219. *Id.*

220. *See generally* JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET (2019).

221. 47 U.S.C. § 230(b).

222. *Internet/Broadband Fact Sheet*, PEW RSCH. CTR. (June 12, 2019), https://www.pewresearch.org/internet/fact-sheet/internet-broadband [https://perma.cc/HVU2-583Y].

223. *Id.*

224. *Id.*

225. *See generally* GANTZ & REISEL, *supra* note 173.

226. GEORGE J. STIGLER CTR. FOR THE STUDY OF THE ECON. & THE STATE, STIGLER COMMITTEE ON DIGITAL PLATFORMS: FINAL REPORT 9 (2019) [hereinafter STIGLER REPORT], https://www.chicagobooth.edu/research/stigler/news-and-media/committee-on-digital-platforms-final-report [https://perma.cc/Q75Z-PTGZ].

227. *Id.* at 156 ("Recent studies of audience news consumption behavior have indicated that news users increasingly rely on multiple news media and seem to shop for the best news across outlets online. As a consequence, they follow the news on multiple media platforms. It has been well-documented that the Internet has reduced loyalty to any single outlet, in particular for technological reasons.").

greater free market of information than ever before, empowering formerly underrepresented and concealed voices in a way that exceeds even the diversity of thought that emerged before FCC regulations ensured corporate dominance of radio broadcasts in the late 1920s.[228]

For defamation law, the transformation has posed a unique problem. Under traditional defamation law, the medium in which a defamatory statement was published had significant bearing on how the publisher experienced liability for it.[229] A printed work, for instance, triggered heavier damages than everyday speech because courts viewed print's permanence as more severely damaging than a fleeting slanderous remark.[230] Television and radio, which disseminated the formerly less harmful spoken word to millions, triggered new debates about damages, ultimately being treated as involving more serious libel because of their destructive potential.[231]

Online defamation defies many assumptions made by traditional defamation law. For instance, the creator and disseminator of tortious content (other than the ICS) might be a fake or anonymous profile who proves nearly impossible to unmask.[232] In Igbonwa's case, the plaintiff could not pursue defamation claims against the "faceless men" who defamed him for this very reason.[233] Where an ICS is shielded from liability, but also shields the third-party ICPs through display of an alias, a unique problem arises of how to unmask the tortfeasor. A court can order a content provider to disclose an anonymous user's information,[234] but this process is often more complex than it seems. Where a user employs an alias, the user's identity would have to be determined via an IP address, an internet user's unique signature based on connection point.[235] Obtaining the IP address through a court could prove costly, time-consuming, and possibly fruitless; if the tortfeasor connected through a public router used by many people or through another person's private router, the IP address narrows the list of possible culprits but cannot identify the tortfeasor.[236] More determined parties can even employ a virtual private network (VPN), which can render them nearly impossible to find even by the ICS, let alone the court system.[237]

---

228. *See generally* ELENA RAZLOGOVA, THE LISTENER'S VOICE: EARLY RADIO AND THE AMERICAN PUBLIC (2011). The Federal Communications Commission's regulation of radio encouraged a shift from niche stations serving niche audiences to larger corporate stations serving broader ones, the reverse of what recommendation systems encourage.

229. *See* Ciolli, *supra* note 15, at 141–43.

230. *Id.*

231. *Id.* at 143–44.

232. *See generally* Ronen Perry & Tal Z. Zarski, *Who Should Be Liable for Online Anonymous Defamation?*, 82 U. CHI. L. REV. DIALOGUE 162, 165 (2015).

233. *See* Igbonwa v. Facebook, Inc., No. 18-cv-02027, 2018 U.S. Dist. LEXIS 173769, at *18–22 (N.D. Cal. Oct. 9, 2018).

234. Perry & Zarski, *supra* note 232, at 165.

235. *Id.* at 166.

236. *Id.*

237. *Compare* Daiyuu Nobori & Yasushi Shinjo, *VPN Gate*, *in* PROCEEDINGS OF THE 11TH USENIX SYMPOSIUM ON NETWORKED SYSTEMS DESIGN AND IMPLEMENTATION 229, 235 (2014) ("Without a packets logging function, criminals could abuse VPN Gate to hide their client IP addresses. When a criminal uses a VPN server, the owner can pass packets logs to a

Victims facing such obstacles may never be able to discover their tortfeasors. Even in successful cases, the financial burden for courts and plaintiffs dramatically increases the cost of making the victim whole.[238]

Beyond the individual level, the entities that most benefit from § 230 have produced significant negative externalities as the spread of unverifiable information and "fake news" generally have become a serious issue on a national and even global scale.[239] Even when news sources are not intentionally producing fake news, the fragmentation of news sites' audiences by the proliferation of social media aggregators has decreased the incentive to produce costly, well-researched reporting when the product can be copied in minutes, rendering any profit motive pointless.[240]

While the conversation about unverifiable information online has centered on fake news, and specifically the 2016 presidential election,[241] the problem of rapidly spreading false information about private individuals has received less attention, despite being palpable as early as the first case to involve a § 230 defense.[242] Information shared on social media is typically detached from clear association with any site besides the recommendation feed of choice and may generally be seen as more trustworthy if shared by a person's friend, rather than by a reliable source.[243] The resulting potential for the spread of misinformation has been compared by prominent scholars to that of medieval Europe; rather than trust verified information from objective experts, consumers trust information shared by friends that reflects their own subjective beliefs.[244] In short, the ability to produce information about anyone, by anyone, is rapidly escalating just as the economic incentive to produce thoroughly researched and trustworthy information about anyone, by anyone, is disappearing. The resulting civilizational crisis in faith has led many to seriously contemplate life in a "post-truth society."[245]

---

law enforcement agency. The 'VPN Gate Anti-Abuse Policy' on the web site clearly states that each VPN Gate server records packet logs in order to prevent such abuse."), *with No-Log VPN Service*, NORDVPN, https://nordvpn.com/features/strict-no-logs-policy [https://perma.cc/LSQ6-G82D] (last visited Oct. 3, 2020) ("We do not store connection timestamps, session information, used bandwidth, traffic data, IP addresses, or other data.").

238. *See* Perry & Zarski, *supra* note 232, at 166.

239. *See generally* STIGLER REPORT, *supra* note 226.

240. *Id*. at 156–57.

241. *See generally* Andrea Butler, *Protecting the Democratic Role of the Press: A Legal Solution to Fake News*, 96 WASH. U. L. REV. 419 (2018).

242. *See Zeran I*, 958 F. Supp. 1124, 1127 (E.D. Va. 1997), *aff'd*, 129 F.3d 327 (4th Cir. 1997).

243. *See generally* Oren Livio and Jonathan Cohen, *'Fool Me Once, Shame on You': Direct Personal Experience and Media Trust*, 19 JOURNALISM 684 (2018).

244. STIGLER REPORT, *supra* note 226, at 165–66 ("Lack of transparency and use of native advertising are said by consumers to make them less trusting of the media. . . . Just as they did in the Middle Ages, audiences trust information that is familiar and/or comes from friends." (footnotes omitted)).

245. *See* S. I. Strong, *Alternative Facts and the Post-truth Society: Meeting the Challenge*, 165 U. PA. L. REV. ONLINE 137, 137–38 (2017); *see, e.g.*, Shanto Iyengar & Douglas S. Massey, *Scientific Communication in a Post-truth Society*, 116 PNAS 151 (2019); Molly Worthen, *The Evangelical Roots of Our Post-truth Society*, N.Y. TIMES (Apr. 13, 2017),

Section 230 aimed to spur innovation in a free market of internet communication,[246] but in the process, it helped drive the growth of its corresponding downsides,[247] while also raising the bar for everyday people to find protection against a wave of misinformation about themselves.[248] Even when plaintiffs overcome the hurdle of § 230 and successfully bring suit against a third-party ICP, judgment-proof defendants who lack the personal resources to compensate a victim's loss often provide their own systemic bar to recovery. [249]

With data production from social media use increasing exponentially in recent years, § 230's limitations on liability for ICSs generally, and social media sites specifically, will likely only produce more victims who find defamation law virtually useless to aid them.[250] The organizations that have empowered this change and are arguably most responsible for the crisis are also immunized from liability by the same statute that allowed them to develop.[251]

The current situation has led to growing pushback against broad interpretations of § 230.[252] Because the statute has been construed so broadly, proposals to address its shortcomings tend to focus on amending the statute rather than salvaging it.[253] Both Republicans[254] and Democrats[255] have proposed that § 230 should be amended to rein in its broad application by the courts. Many of these reforms, however, embrace irreconcilable ideas that are unlikely to find bipartisan appeal beyond the recognition that § 230 is somehow flawed.[256] Some legislators, recognizing the growing prevalence

---

https://www.nytimes.com/2017/04/13/opinion/sunday/the-evangelical-roots-of-our-post-truth-society.html [https://perma.cc/GD5B-8UY5].

246. 47 U.S.C. § 230(b).

247. *See generally* STIGLER REPORT, *supra* note 226, at 156–57.

248. *See, e.g.*, Perry & Zarski, *supra* note 232, at 165–66.

249. *See* Salina Tariq, *Revenge: Free of "Charge?,"* 17 SMU SCI. & TECH. L. REV. 227, 228 (2014).

250. ANTONIO GARCÍA MARTÍNEZ, CHAOS MONKEYS: OBSCENE FORTUNE AND RANDOM FAILURE IN SILICON VALLEY 506–07 (2016). García Martínez, a former Facebook product manager, who has also been an adviser to Twitter, discusses what is essentially § 230 immunity as the reason for expanding misinformation at every social level, although he cites the Digital Millennium Copyright Act, a source of related immunity in copyright, as the cause.

251. *Id.* at 507.

252. *See* Sylvain, *supra* note 52; *see also* Force v. Facebook, Inc., 934 F.3d 53, 76 (2d Cir. 2019) (Katzmann, J., dissenting); Butler, *supra* note 241.

253. *See, e.g.*, Butler, *supra* note 241 (proposing an amendment to § 230 to expand defamation liability); *see also* GOODMAN & WHITTINGTON, *supra* note 11, at 8.

254. Ending Support for Internet Censorship Act, S. 1914, 116th Cong. (2019) (proposing that § 230 protections be suspended for social media companies "unless the social media company obtains certification from the Federal Trade Commission that it does not moderate information on its platform in a manner that is biased against a political party, candidate, or viewpoint").

255. GOODMAN & WHITTINGTON, *supra* note 11, at 1 (discussing how House of Representatives Speaker Nancy Pelosi and Representative Adam Schiff have both called for reforms to § 230 as a grant of power without responsibility).

256. *Id.* at 6 ("While it is true that Section 230 is increasingly a bipartisan issue, proposals to reform it are often attempting to achieve irreconcilable goals. Many recent ones are designed to encourage firms to adopt greater responsibility for policing their platforms, either

of online defamation, want to expand platforms' incentives to moderate and censor their users as a countermeasure.[257] Other legislators, fearful of the growing dominance of ICSs over news and platforming decisions, want to cut back protections in a way that limits their moderation capabilities.[258] President Donald Trump has adopted a goal of pushing the latter view through executive order, but such a move could well be seen as an attempt by the executive branch to overturn long-standing judicial precedent.[259] This difficult reality calls into question whether § 230 may be somehow salvaged by the courts themselves without changing the text of the statute or existing legal precedent. Could courts instead apply current interpretations of § 230 in a way that more accurately reflects the role recommendation systems behind the massive social media aggregators of our time handle the third-party content that they then distribute to the public?

II. POSSIBLE APPLICATIONS OF § 230 TO A RECOMMENDATION SYSTEM

This part discusses the potential arguments for and against the application of § 230 immunity to recommendation systems directly. Using Facebook as a case study, Part II.B explores potential arguments for why a recommendation system such as the Facebook News Feed should face liability as the ICP of its own material. Part II.C poses potential problems with those arguments, as well as corresponding arguments against limiting § 230's expansive reach.

A. Parameters of the Discussion: Facebook as a Case Study

Section 230's protections touch a broad spectrum of internet services, from advertisers[260] to search engines.[261] Even limiting the scope of a discussion about possible § 230 reforms to recommendation systems, such systems come in various levels of complexity[262] and apply to different platforms with different designs.[263] Part II of this Note focuses on how § 230 interpretations would apply to the representation-learning algorithmic recommendation systems, using the Facebook News Feed as its anecdotal focus.

---

through the creation of additional exceptions to Section 230's immunity, or through the establishment of preconditions for immunity. Other proposals aim to limit the amount of moderation platforms can employ. There are also important differences in whether the proposals seek to shape platform content ex ante with new content-based Section 230 exceptions or to condition immunity after the fact based on platform reasonableness or processes.").

257. *Id.* at 6.

258. *Id.*

259. Maggie Haberman & Kate Cronger, *Trump Prepares Order to Limit Social Media Companies' Protections*, N.Y. TIMES (June 2, 2020), https://www.nytimes.com/2020/05/28/us/politics/trump-executive-order-social-media.html [https://perma.cc/BY9B-P7DR].

260. FTC v. LeadClick Media, LLC, 838 F.3d 158, 176 (2d Cir. 2016).

261. Marshall's Locksmith Serv. Inc. v. Google, LLC, 925 F.3d 1263, 1266 (D.C. Cir. 2019).

262. *See supra* notes 190–213 and accompanying text.

263. *See supra* notes 175–78 and accompanying text.

This Note selects Facebook as its recommendation system of choice to analyze for several reasons. First, Facebook has found itself to be the lightning rod for regulators' attention to the recommendation systems at issue.[264] Second, its 2.7 billion users comprise more than a third of the global population.[265] Third, as the already massive proliferation of data accessible through social media is only projected to increase,[266] a fully automated recommendation system like that used in Facebook's News Feed will likely be not only ideal but necessary to operate a functioning social media site in the 2020s and beyond.[267]

B. Arguments for Limiting § 230 Immunity

Given the many problems created by § 230 and the limited probability of meaningful amendment, this Note considers interpretations of § 230 that can salvage the statute as it stands. Most of these involve applying § 230 with a more comprehensive understanding of how a recommendation system differs from an online bulletin board.

1. Applying *Roommates.com* and *Accusearch* Directly to Facebook's Current Practices

The Facebook algorithm might have a more complex result than that of Roommate.com's search engine but operates on the same basic principles: the recommendation system filters through a content pool based on the categorization of individual elements within that pool and presents the user with a result.[268] The main difference between the two stems from the source of the categories. On Roommate.com, users categorized themselves based on categories preselected by the website's human operators.[269] On Facebook, users are categorized by the recommendation system itself in an entirely automated process.[270] Based on a user's interactions with posts the algorithm recognizes as involving "football," for example, Facebook may categorize that user as a "football fan" and feed the user more content involving football.[271]

---

264. Marcy Gordan, *From Toast of Town to Toxic: Facebook CEO on Outs with Dems*, ASSOCIATED PRESS (Nov. 4, 2019), https://apnews.com/article/f9cbd0627e4a4a90a778d5b819967cda [https://perma.cc/4RUT-UDCJ]; *see also* Adam Satariano, *Facebook Dodged a Bullet from the F.T.C. It Faces Many More.*, N.Y. TIMES (July 13, 2019), https://www.nytimes.com/2019/07/13/technology/facebook-privacy-investigations.html?searchResultPosition=9 [https://perma.cc/EP8J-PNGM].

265. Mike Isaac, *Facebook's Mark Zuckerberg Says He'll Shift Focus to Users' Privacy*, N.Y. TIMES (Mar. 6, 2019), https://www.nytimes.com/2019/03/06/technology/mark-zuckerberg-facebook-privacy.html?searchResultPosition=1 [https://perma.cc/JE37-SPX5].

266. GANTZ & REISEL, *supra* note 173, at 1 ("From now until 2020, the digital universe will about double every two years.").

267. *See* Sylvain, *supra* note 52.

268. *See* Oremus, *supra* note 191.

269. Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1164 (9th Cir. 2008) (en banc).

270. *How News Feed Works*, *supra* note 34.

271. *See id.*

Because Facebook's categories are automatically generated to meet the goals of its algorithm, problematic categories have appeared on its site in the past, generating bad publicity for the company.[272] Facebook has even ended up in the same troublesome situation as Roommate.com by targeting advertisements based on race, though it has since made changes to avoid violating antidiscrimination laws like the FHA.[273] The site currently employs a comprehensive nondiscrimination policy for advertisers.[274] Even with § 230 protections intact, the social media giant understood that its protections would not extend if their recommendation system materially contributed to third-party violations of the FHA by designing a different News Feed for people of different races.[275]

This understanding could possibly revive some of the protections of defamation law when the same logic is applied to defamation. If Facebook theoretically targeted pieces of content that are identifiably defamatory to users most likely to engage with defamation, then, by the holding of *Accusearch*—that a website is responsible for its content if it specifically encourages what generates liability—it would be responsible for the defamation; it "developed" content by sending it to a third party and was responsible for doing so because its algorithm sought out defamatory material.[276] Courts could even distinguish such a situation from that in *Jones*; in that case, the moderators reposted user submissions regardless of their defamatory nature.[277] Unlike the moderators of The Dirty, such a situation would mean Facebook's algorithm objectively knew the content was defamatory and chose it for that purpose.[278]

Such a hypothetical would make an interesting thought experiment if it were not already somewhat true. Facebook's Ad Manager tool allows any customer to create and market an advertisement to selected audiences, limiting and expanding the scope of targeted recipients by various traits.[279] One such category of interests, "Slander," is described as "people who have expressed an interest in or like pages related to Slander."[280] Therefore, Facebook's current algorithm might unambiguously target third-party

272. Angwin et al., *supra* note 211.

273. *See Roommates.com*, 521 F.3d at 1157; Sapna Maheshwari & Mike Isaac, *Facebook Will Stop Some Ads from Targeting Users by Race*, N.Y. TIMES (Nov. 11, 2016), https://www.nytimes.com/2016/11/12/business/media/facebook-will-stop-some-ads-from-targeting-users-by-race.html?searchResultPosition=1 [https://perma.cc/S5YP-FAWZ].

274. *Advertising Policies: Prohibited Content*, FACEBOOK, https://www.facebook.com/policies/ads/prohibited_content/discriminatory_practices [https://perma.cc/T5GJ-2ZES] (last visited Oct. 3, 2020).

275. *See* Maheshwari & Isaac, *supra* note 273.

276. *See* FTC v. Accusearch, Inc., 570 F.3d 1187, 1198–2000 (10th Cir. 2009).

277. Jones v. Dirty World Enter. Recordings LLC, 755 F.3d 398, 403 (6th Cir. 2014).

278. *Id.*

279. *Business Help Center*: *About Ad Creation in Ads Manager*, FACEBOOK FOR BUS., https://www.facebook.com/business/help/282701548912119?id=649869995454285 [https://perma.cc/ZQ24-SKEW] (last visited Oct. 3, 2020).

280. *Id.* Upon creating a Facebook page and a corresponding ad, if one clicks the "Edit" button next to the "Detailed Targeting" category and types "slander" into the resulting text box, hovering over the category reveals the description.

content to potential recipients because of their expressed interest in slander, fulfilling the hypothetical this Note outlines.[281] In addition to this category, Facebook also gives advertisers the option to allow "targeting expansion" in which Facebook can automatically expand content reach to new categories not selected by the advertisers.[282] One of these categories, presumably, would be slander.

With regard to content labeled with the "slander" category or something similar ("libel" or even "defamation"), the argument for liability would be a relatively straightforward one: because Facebook developed content by making it publicly available and was responsible for its development by seeking out defamatory content to display, it became an ICP of that specific piece of content even though it did not create it, just as Accusearch became an ICP of its phone records even though it did not itself obtain them.[283]

2. Removing Protections by a Second Creation Point

While Judge Robert Katzmann's lone dissent in *Force* focused on Facebook's potential to influence actions beyond the scope of internet speech (which the Third Circuit has already recognized as a valid exception to § 230 immunity),[284] he touches on a detail distinguishing the Facebook algorithm from the Prodigy message boards: there is a second point of creation between a third-party posting a Facebook status and a user's reception of that content.[285] Under the third prong of the *Zeran* analysis, an ICS that creates or provides its own content cannot receive protection under § 230 from liability for that content.[286] By creating the recommendation feed and customizing content for individual users, Facebook adds another level of development to user posts, for which it is almost entirely responsible.[287] Facebook would therefore be the ICP of that additional level of development.

Suits against Facebook and other sites failing to overcome the § 230 defense generally focus on assailing an individual third-party contribution, or a collection of disaggregated third-party contributions, as themselves impetuses for liability for the ICS.[288] Such cases typically fail because, if Facebook is assumed to be a neutral collector and displayer of information, it cannot be said to be responsible for the content at issue.[289]

---

281. *Id.*

282. *Business Help Center: About Targeting Expansion*, FACEBOOK FOR BUS., https://www.facebook.com/business/help/128066880933676?id=176276233019487 [https://perma.cc/GR6J-GRGH] (last visited Oct. 3. 2020).

283. *See* FTC v. Accusearch, Inc., 570 F.3d 1187, 1198–2000 (10th Cir. 2009).

284. Oberdorf v. Amazon.com Inc., 930 F.3d 136, 153 (3d Cir. 2019).

285. Force v. Facebook, Inc., 934 F.3d 53, 76 (2d Cir. 2019) (Katzmann, J., dissenting).

286. *See supra* Part I.B.2.

287. *See supra* Part I.B.3 (identifying development as a universal action performed by every ICS, with the material question being whether the ICS was responsible for that development).

288. *See, e.g.*, *Force*, 934 F.3d at 66–67.

289. *Id.*

Facebook's recommendation system is anything but a neutral conduit. Through the design of its programmers and the evolution of its representation-learning algorithm, it can send different content to different people based on its own categories and target audiences.[290] Facebook can even influence the behavior of its users, who are its ICPs, through its design.[291] The questions then arise of whether Facebook develops the content processed through that design and whether it is responsible for that development.[292]

The divide between circuit interpretations of development here becomes important. The Tenth Circuit held in *Accusearch* that all websites "develop" content by hosting it for public display.[293] By this definition, the only question remaining is whether the site is responsible for the development of the offensive content.[294] Other courts, such as the Second Circuit, have held that the display of information alone does not constitute development.[295] Instead, the Second Circuit holds that development only occurs when an intermediary is not neutral with respect to the offensive trait of the content at issue.[296] In other words, development only occurs when an intermediary distinguishes between offensive content and all other content.[297]

All of these questions, however, become irrelevant if the News Feed itself is thought of as its own product. Facebook, understandably, closely guards its algorithm as its property.[298] The News Feed is arguably a product of Facebook; it cannot be produced by any other source.[299] The Facebook News Feed would be empty if not for the provision of third-party content, but third parties themselves likewise cannot produce it.[300]

In copyright law, compilations of third-party content are treated as the intellectual property of their compilers so long as they contain a modicum of originality. As outlined in *Feist Publications, Inc. v. Rural Telephone Service Co.*,[301] an original selection of facts is eligible for copyright, though the copyright would then be limited to the arrangement and not the facts copyrighted.[302] In a world where Facebook does not qualify as an ICS, the compilation of content to best match user interests would result in millions,

290. *How News Feed Works*, *supra* note 34; *see also* Michal Lavi, *Evil Nudges*, 21 VAND. J. ENT. & TECH. L. 4, 18–20 (2018) (discussing the various ways in which the design of social networks can influence the dissemination of information).
291. Lavi, *supra* note 290, at 20.
292. *See* FTC v. Accusearch, Inc., 570 F.3d 1187, 1198–99 (10th Cir. 2009).
293. *Id.*
294. *See id.* at 1199.
295. Force v. Facebook, Inc., 934 F.3d 53, 70 (2d Cir. 2019).
296. *See id.* at 69.
297. *Id.*
298. Oremus, *supra* note 191.
299. *See id.*
300. *See id.*
301. 499 U.S. 340 (1991).
302. *Id*. at 350–51.

if not billions, of personalized, copyrightable original publications,[303] analogous to the creation of a personalized physical newspaper.[304]

This informs the § 230 analysis because it clarifies a legal precedent that Facebook's content emerges from two points of creation or development.[305] The first development point occurs when the Facebook user, presumably by free choice,[306] contributes content to the site voluntarily by posting or sharing the content of another.[307] The second point of development, classified by some as a publishing function[308] and others as development of equal import,[309] occurs when Facebook displays the content on a user's feed. If a website displays all content in a neutral fashion, the circuits agree that there should be no responsibility for that content.[310] The question is whether, at the second development point, a website is responsible for not just developing content but developing the illegal aspect of that content.[311] If Facebook developed a post not categorized as defamation (or slander),[312] which still was, in fact, defamation, should it then be absolved of responsibility for that development?

A deep-learning algorithm sorts content by categories, but those categories are not magically distilled from thin air; they are the more complex interpretations of simpler traits.[313] In recognizing a face, for instance, an image-categorizing, deep-learning AI might first categorize edges, then categorize shapes, including circles, based on combinations of edges, and then understand combinations of shapes to be a face.[314] Even if Facebook barred all categories of defamation, slander, libel, or other related words from being used as categories of representation, the system would still categorize content based on the foundational elements that it formerly categorized as defamation because those elements still exist.[315] In the image categorization example from Google, if an image had whiskers, pointy ears, fur, four legs, and two eyes, that creature might be categorized as a "cat."[316] If Google intervened with an order barring all categorization of image search results as

303. *See generally id.; How News Feed Works*, *supra* note 34.

304. Cass R. Sunstein, #REPUBLIC: DIVIDED DEMOCRACY IN THE AGE OF SOCIAL MEDIA 1–3 (2017) (calling Facebook, as well as Twitter and Google, a close approximation to the "Daily Me" personal newspaper predicted in 1995 by technology specialist Nicholas Negroponte); *see also* STIGLER REPORT, *supra* note 226, at 9 (calling Facebook the second largest news provider in the United States).

305. *See* FTC v. Accusearch, Inc., 570 F.3d 1187, 1198–99 (10th Cir. 2009).

306. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1166, 1174 n.38 (9th Cir. 2008) (en banc). *But see generally* Lavi, *supra* note 290 (arguing that design patterns of a website can substantially erode a user's free will in content creation).

307. *How News Feed Works*, *supra* note 34.

308. *See* Force v. Facebook, Inc., 934 F.3d 53, 68 (2d Cir. 2019).

309. *See Accusearch*, 570 F.3d at 1198–99.

310. *See supra* Part I.B.

311. *Accusearch*, 570 F.3d at 1198–99.

312. *See supra* Part II.B.1.

313. Chartrand et al., *supra* note 193, at 2115.

314. Lewis-Kraus, *supra* note 216.

315. *See id.*; *see also* Chartrand et al., *supra* note 193, at 2115.

316. Lewis-Kraus, *supra* note 216.

"cat," the images would still be classified as commonly containing similar whiskers, pointy ears, fur, legs, and eyes.[317]

What this means is that even if Facebook were to take the surface-level measure of banning a category, its site would still adapt to prioritize defamation, just in another form; the existence of a presumably autogenerated slander category further evidences this.[318] Facebook has, in its past, considered legal ramifications in structuring its product.[319] If the problem of intensive defamation spreading by algorithm persists past an easy fix, not only could § 230 immunity not apply but it may provide an impetus for systematic change to better control the site's flow of information.

3. Facebook Should Be Responsible for Its ICPs' Actions

The theory behind § 230 protections for Facebook is a simple one: it should not be held liable for information contributed to its site by a third party.[320] The text of § 230 embraces this rationale in no uncertain terms.[321] Courts have nonetheless found that § 230 liability ought not apply when a site is "responsible" for the development, even by a third party; this theory has already defeated the third prong of the *Zeran* analysis in past cases.[322]

In a situation where users freely decide of their own volition what to post or advertise, solicitation is not usually an issue.[323] Such an interpretation squares with the text of § 230, which applies explicitly to content for which an ICS has no responsibility.[324] This may not be the case with many modern websites, where the ICS and its users are not isolated actors interacting tangentially with each other, but two parties in a relationship that affects both involved.[325] Through both "dark patterns" obscuring user choice[326] and "evil nudges" pushing users in directions that the recommendation system desires,[327] user actions on Facebook may not entirely be the product of their own intentions. The recommendation algorithm, in part, is "learning" from user "preferences" that might have originated from the design of the website,[328] or even the algorithm itself,[329] rather than the user. This is simultaneously analogous to the limitation of user choice in

317. *See supra* notes 305–08.
318. *See infra* Part II.B.1.
319. GARCÍA MARTÍNEZ, *supra* note 250, at 507. García Martínez discusses how he was "constantly sprinting half-blind through a minefield of potential legal problems . . . . and trying to find some legal rubric that would forgive (or at least defensibly excuse) our next depredation with user data." *Id.* at 317; *see also* Isaac, *supra* note 265; Satariano, *supra* note 264.
320. *See* Sylvain, *supra* note 52.
321. *See* 47 U.S.C. § 230(c).
322. *See supra* Part I.B.
323. *See supra* Part I.B.
324. *See* 47 U.S.C. § 230(c).
325. Lavi, *supra* note 290, at 18–20.
326. STIGLER REPORT, *supra* note 226, at 210.
327. Lavi, *supra* note 290, at 18–20.
328. *Id.*
329. *Id.*

Roommates.com[330] and the solicitation in *Accusearch*;[331] if users' actions are mostly shaped by the design of a site, they are not acting in full independence, and Facebook, not the user, is the ICP of that content.

The idea that free will of consumers can be manipulated by small nudges is by no means a novel one or one confined to the internet. Exposing a focus group to a lesson on habit formation, for instance, can dramatically increase time spent exercising.[332] The chime of a smartphone notification likewise serves as a nudge to check emails, even if the user wants to focus on working undisturbed.[333] The nudges are not always obvious connections to that which they nudge; environmental factors such as geographic location can greatly "nudge" certain individuals to commit crimes, for example.[334]

Like any other business, a social media site can use nudges and dark patterns to trigger desired outcomes.[335] Dark patterns in this setting refer to website structures that either distract users from doing something or manipulate them into acting in some way.[336] A common, if simple, example of this might be where a website prominently places an "accept" button for a newsletter and subtly places a "decline" option. Another dark pattern may appear to require acceptance of the newsletter to continue browsing a site, even if that is not the case.[337]

With the extreme personal knowledge that modern recommendation systems have of their users, more sophisticated dark patterns can not only target users with advertisements but also with nudges to create or share desired content.[338] An extremely basic example of this would be Facebook "reacts."[339] At least one former executive of the site has remarked that because engagement with the content a user posts on Facebook triggers reward centers in the brain, bad actors can manipulate user choices by exploiting patterns that lead to more engagement.[340]

Whether or not Facebook intentionally prompts its users to spread defamation in order to increase user engagement, it may not have to. Falsehoods spread through online social networks "significantly farther

330. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1166, 1174 n.38 (9th Cir. 2008) (en banc).

331. *See* FTC v. Accusearch, Inc., 570 F.3d 1187, 1198–99 (10th Cir. 2009).

332. Duhigg, *supra* note 174.

333. *Id.*

334. JOHN F. PFAFF, SENTENCING LAW AND POLICY 216–17 (Robert C. Clark et al. eds., 2016).

335. *See* Lavi, *supra* note 290, at 18–20.

336. STIGLER REPORT, *supra* note 226, at 237.

337. *Id.* at 238.

338. *See* Amy B. Wang, *Former Facebook VP Says Social Media Is Destroying Society with 'Dopamine-Driven Feedback Loops,'* WASH. POST (Dec. 12, 2017), https://www.washingtonpost.com/news/the-switch/wp/2017/12/12/former-facebook-vp-says-social-media-is-destroying-society-with-dopamine-driven-feedback-loops [https://perma.cc/R9NV-7DNT].

339. *See id.* (reporting on Facebook's former vice president for user growth's remarks that Facebook employs "dopamine-driven feedback loops" to exploit users); *see also* Duhigg, *supra* note 174.

340. *See* Wang, *supra* note 338.

faster, deeper, and more broadly than the truth."[341] This sharing bias for false information, presumably because of its novelty,[342] persists even when a social network's design does not intentionally encourage it.[343] A social media post becomes "novel" when it contrasts with previously existing information; false rumors, which are usually significantly more "novel" than true rumors, may be more likely to be shared by users when the corpus of information they have been exposed to contradicts the false rumor.[344]

Fact-checking and other efforts to promote "true information," which have little to no effect on their own,[345] try to remedy the prevalence of false rumors by promoting truthful, verified content for users to share and engage with.[346] While novelty may not be the only factor that causes false rumors to spread faster than truthful content, novel information is more likely to be shared, and falsehoods are more likely to be novel.[347] An ICS that did not interfere at all would still see the spread of false rumors faster than truthful ones.[348] By promoting verified content, be it news or statements about individuals, the ICS expands the corpus of truthful information seen by the average reader, even if the reforms only apply to a small minority of users.[349] Novelty, however, has a correlative relationship with the existing corpus of contrasting information the user has seen.[350] Efforts to dominate the Facebook News Feed with verified information may then produce a perverse effect by expanding the novelty of and therefore user engagement with defamation.[351]

C. Arguments for Extending Immunity

Despite the many problems created by § 230's expansion, refusing its protections to recommendation systems would face some opposition. This section discusses alternative views of the recommendation system, the legal issues with questioning user agency as an ICP, and the potential broader policy issues of refusing § 230 immunity to recommendation systems.

---

341. Lavi, *supra* note 290, at 20; *see also* Soroush Vosoughi et al., *The Spread of True and False News Online*, 359 SCIENCE 1146, 1146 (2018) (finding that false tweets spread faster than true tweets).

342. Vosoughi et al., *supra* note 341, at 1149.

343. *Id.* at 1150 ("The greater likelihood of people to retweet falsity more than the truth is what drives the spread of false news, despite network and individual factors that favor the truth.").

344. *Id.* at 1149–50.

345. STIGLER REPORT, *supra* note 226, at 166.

346. *See, e.g.*, Smith & Frier, *supra* note 214.

347. Vosoughi et al., *supra* note 341, at 1150.

348. *Id.*

349. Smith & Frier, *supra* note 214 ("[I]t's better to have a separate tab, which will definitely draw a smaller audience, maybe 10% to 20% of the main feed, but that will still be significant.").

350. Vosoughi et al., *supra* note 341, at 1150 ("The last two [novelty] metrics measure differences between probability distributions representing the topical content of the incoming tweet and the corpus of previous tweets to which users were exposed.").

351. *See supra* notes 337–42 and accompanying text.

1. Applying *Roommates.com* to Individual User Posts

No matter how the mechanics and constructs of a recommendation system are defined and analyzed, one concrete fact weighs solidly against the incursion of liability: if the individual fragments of a recommendation system's resulting display that were manipulated, distributed, and compiled by the site originate with a third-party post.[352] In *Roommates.com*, potential liability for the operation of a recommendation algorithm came from a series of prepopulated answers originating from a drop-down field.[353] Such broad choice presented to a user would not ordinarily create liability for an ICS.[354]

While a drop-down menu of illegal categories is not usually presented to a user, the use of a recommendation system can complicate the matter because an algorithm or human expert can sort users and posts into categories that they have not chosen.[355] Dark patterns and similar phenomena may make it harder for a user to control what their news feed contains, but the user can still ultimately control their Facebook News Feed.[356] Users can hide content they dislike and request to see less of it in the future.

The recommendation algorithm largely sorts information based on user activity within the site, which the user agrees to provide to Facebook when it uses the site.[357] Beyond some basic information, such as a telephone number and name, Facebook requires nothing of a user before they start using Facebook and voluntarily surrendering their preferences and desires by navigating the site.[358] Regarding the "slander" category of advertising, the situation is not that Facebook seeks out slander to send unwitting users. Users demonstrate their willingness to engage with slander by making choices in what content to engage with. In response to the demonstrated user preference, Facebook's automated recommendation system responds to the user's expressed desires.[359] The user, at least in part, is responsible for the defamatory content in the user's own News Feed.

The News Feed might also be considered a choice of presentation, rather than an individual compilation; this was the argument embraced by *Force*.[360] The Facebook algorithm, proprietary or not, is an algorithm choosing how to arrange individual pieces of content in the Facebook News Feed. The Facebook user can also request changes in an individual News Feed, which the algorithm then responds to by changing content in the future.[361]

---

352. *How News Feed Works*, *supra* note 34.
353. Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1166 (9th Cir. 2008) (en banc).
354. *Id.* at 1174 n.38.
355. *How News Feed Works*, *supra* note 34.
356. *Id.*
357. *Id.*
358. Force v. Facebook, Inc., 934 F.3d 53, 70 (2d Cir. 2019).
359. *How News Feed Works*, *supra* note 34.
360. *Force*, 934 F.3d at 70.
361. *How News Feed Works*, *supra* note 34.

2. The ICP's Ultimate Responsibility for Offensive Content

Dark patterns and nudges aside, the Facebook user exercises ultimate control over what is posted on the site. All pieces of content begin with what is essentially a blank box; users can contribute defamatory statements to the site and may even be encouraged to by the site's design but by no means are they required to.[362] Courts that have rejected "nudge" arguments have generally done so on these grounds.[363]

The Sixth Circuit's protection of The Dirty provides the best example of this counter to nudge arguments.[364] So long as websites do not develop user content *because* of its problematic nature and do not contribute anything problematic to the content at issue, they act neutrally regarding their hosted content.[365] The Dirty posted defamatory content, but the content originated almost entirely from third parties and that which did not, the Sixth Circuit held, was not defamatory.[366] While, The Dirty did, however, republish defamatory information from other ICPs, some have called § 230 protections for such behavior a full repudiation of tort law's applicability to republication.[367] As courts have already built the framework of § 230 around eliminating the republication rule, third-party liability for defamatory content of users becomes nearly impossible unless the content at issue originated explicitly from the ICS that hosted it.[368]

3. Further Limits on § 230 Immunity Will Offend Speech Concerns

Part of the purpose of § 230 was to preserve a free and open marketplace of ideas and to spur online innovation. This intent resided so saliently in the minds of legislators that it found its own place in the statute.[369] These same concerns were also what drove the adoption of an extremely expansive reading of the statute in *Zeran*.[370] These interpretations have furthermore driven the development of an information economy built around § 230 immunity.[371] A court's attempt to reverse this trend in a meaningful way, be it through algorithm liability or applying § 230 to "evil nudges," might prove impractical to the extent that the modern economy has embraced its

362. *See id.*

363. *See, e.g.*, GW Equity LLC v. Xcentric Ventures LLC, No. 07-CV-976, 2009 WL 62173, at *5 (N.D. Tex. Jan. 9, 2009); Whitney Info. Network, Inc. v. Xcentric Ventures, LLC, No. 04-CV-47-FtM-34, 2008 WL 450095, at *11 (M.D. Fla. Feb. 15, 2008); *see also* Jones v. Dirty World Enter. Recordings LLC, 755 F.3d 398, 403 (6th Cir. 2014).

364. *Jones*, 755 F.3d at 403.

365. *See supra* Part I.B.3.

366. *Jones*, 755 F.3d at 403.

367. *See* Zipursky, *supra* note 13, at 14–16.

368. *Id.*

369. 47 U.S.C. § 230(b).

370. *Zeran II*, 129 F.3d 327, 330 (4th Cir. 1997).

371. STIGLER REPORT, *supra* note 226, at 156–57.

immunity.[372] Modern means of communication may have advanced so far that a correction would only come at a massive cost to society.[373]

There are also concerns that, under the existing interpretations of § 230 as limited by cases like *Roommates.com* and *Accusearch*, ICSs are dissuaded from the exact sort of Good Samaritan moderation of their content that § 230 was intended to encourage.[374] Further limitations on § 230 immunity would inevitably trigger greater pains by ICSs to avoid that liability, driving § 230 even further from accomplishing one of its original goals.[375] Every online recommendation system that would lose liability under the new regime could then be prevented from implementing the good faith moderation § 230 and its backers sought to encourage.[376] Recall that technology has advanced, and more data is being created and processed than ever before in human history.[377] Section 230, in part, aims to promote the development of technologies to simultaneously encourage and control the massive flow of information.[378] For two decades, these advancements have occurred under a § 230 immunity regime. A sudden change in that regime could render those advancements legally impractical and economically irreplaceable but in a far more advanced cyberspace. Online providers might strive to look more like CompuServe, the defendant in *Cubby* (the companion case to *Stratton*).[379] In that case, CompuServe avoided liability by passing the buck of moderation duties to a third party.[380]

A similar initiative, if it emerged today, could devastate the prospects of recovery for plaintiffs in defamation suits involving a recommendation system. If a *Cubby*-style case developed today, it might not result in a curbing of online defamation but rather its proliferation, as tortfeasors run wild in unregulated online spaces.[381] Such tortfeasors might still frequently operate under an aliases or otherwise render themselves impossible to discover.[382] When found, they would still be more frequently judgment-proof[383] than the parties who actually benefited, via website traffic, from the defamatory content.[384] In other words, the problems a proponent of liability

372. GARCÍA MARTÍNEZ, *supra* note 250, at 508–10 (giving the opinion that modern media has adapted so far in the direction of "new Media Medievalism" that navigation of the new normal, rather than a reversion, is the most practical path of technological development).

373. *See id*.

374. *See* Tariq, *supra* note 249, at 237 ("Under this analysis, most websites, including 'MyEx.com,' clearly foresee that most information on their site will be non-consensual in nature, enjoy CDA immunity as long as all they do is provide space for vengeful posts, such as offensive texts, images and videos, and do not actively participate in forming the content themselves.").

375. GOODMAN & WHITTINGTON, *supra* note 11, at 9.

376. *See supra* notes 371–73 and accompanying text.

377. *See generally* GANTZ & REISEL, *supra* note 173.

378. 47 U.S.C. § 230(c).

379. *See* Cubby, Inc. v. CompuServe, Inc., 776 F. Supp. 135, 137 (S.D.N.Y. 1991).

380. *See id.*

381. *See* GOODMAN & WHITTINGTON, *supra* note 11, at 9.

382. *See generally* Perry & Zarski, *supra* note 232.

383. *See generally id.*

384. *See* Tariq, *supra* note 249, at 237.

for recommendation systems would hope to solve might end up worsened, not improved, as ICSs adapt to the updated § 230.

III. A PROPOSED REFINEMENT OF § 230 IMMUNITY

This part discusses a refinement of § 230 interpretations that could resolve some legal and policy problems caused by the statute without amending it and while also limiting negative consequences. After discussing the foundations of the proposal, it explains the legal viability of adopting this view in future interpretations of the statute, before analyzing its broader impact on society at large.

A. Foundations of the Proposal

When legislators drafted and passed § 230, they aimed to protect incentives for innovation in navigating the unexplored frontier of the internet.[385] They saw the potential that tort law, as it existed, could stifle technological innovation in the way of establishing a free and unprecedentedly broad market of information for the republic and by extension block that potential achievement through the courts.[386] Since § 230's passage, the internet titans of our time have brought about some version of the future those representatives sought. The technologies they created can not only help humans navigate cyberspace but do so by themselves faster than any human could hope to.[387] The result has been a technological revolution, opening not only metaphorical free markets of information but literal markets of economic opportunity.[388]

In the midst of this innovation, however, new technologies have surpassed what legislators or judges in 1996 addressed in drafting § 230. Section 230 arose from a fear that legal precedent would stifle innovation and leave the internet otherwise completely vulnerable to an onslaught of objectionable material.[389] The objectionable material in question came primarily from third-party posters, who had no affiliation with the websites they engaged with. Section 230 responded by restricting liability for those third-party ICPs attributable to the computer services that hosted them.[390]

The measure made sense when the line between content provider and content host was as clear as who posted what on a bulletin board. As online services broadened and grew more complex, however, the lines began to blur, and courts have stepped in to refine the emerging gray area between original ICS product and third-party contribution.[391] The emerging rule became that where an ICS was responsible for the development of content through its

385. *See supra* Part I.A.2.
386. *See supra* Part I.A.2.
387. *See supra* Part I.C.
388. *See supra* Part I.D.
389. *See supra* Part I.A.2.; *see also* Sylvain, *supra* note 52.
390. *See supra* Part I.A.3.
391. *See supra* Parts I.B.2–3.

service, be it by encouragement or direct contribution, it could act as the provider of its own information content.[392]

The two conceptualizations of development, exemplified by the Tenth and Second Circuits, are relevant only for structuring the third prong of the *Zeran* analysis. One side poses that a website develops all content it displays, regardless of what other actions it takes; it becomes responsible for that content when it specifically encourages illegal or tortious aspects.[393] The other side essentially dodges the question of development, holding that a website cannot develop content until it materially contributes to the same problematic aspects.[394] Of course, a website can develop content under the second conception in completely acceptable ways; it only loses its protections from liability as a developer when it develops illegality.[395] The second standard's phrasing creates more confusion than the first but, understood properly, they should accomplish the same result. The relevant question is whether, in the development of content, an ICS contributed to the aspect of the content at issue that creates liability.

Even with this clarification, a narrower gray zone remains: what constitutes contribution? What § 230's drafters conceived of as a wilderness of information yet untamed has become, in many cases, a highly tailored and curated space.[396] Access to the wider internet also increasingly flows through such spaces, giving the curators more leverage over where and how that information spreads.[397] Far from needing encouragement to develop content moderation systems, an ICS can now exercise extensive control over what types of information its users will see.[398] The curation systems § 230 incentivized are now so far advanced that ICS users have little to no control over where their contributed content goes.[399] Commentators have even called into question whether the users exert full control over their content even at the creation stage.[400]

A deep-learning recommendation system falls into the new gray area. Not only does it develop content by displaying it, but it takes the further step of deciding to whom it will be displayed.[401] While the users create fragments of content, the algorithm assembles those fragments into unique compilations based on users' prior tastes.[402] Those compilations, in Facebook's case, the News Feed, can then run against established legal norms about

392. *See supra* Part I.B.2.
393. *See supra* notes 291–92 and accompanying text.
394. *See supra* notes 293–96 and accompanying text.
395. *See supra* notes 293–95 and accompanying text.
396. *See supra* Part II.B.3.
397. Adrienne LaFrance, *Facebook Is Eating the Internet*, ATLANTIC (Apr. 29, 2015), https://www.theatlantic.com/technology/archive/2015/04/facebook-is-eating-the-internet/391766/ [https://perma.cc/NZ85-J863].
398. *See supra* Parts I.C, II.B.3.
399. *See supra* Part I.C.
400. *See supra* Part II.B.3.
401. *See supra* Part I.C.
402. *See supra* Part I.C.; *see also How News Feed Works*, *supra* note 34.

defamation.[403] Where defamation is shared, the relevant question, again, is: who is responsible?[404]

B. A Proposed Interpretation: Algorithm as ICP

A recommendation system has multiple components. These involve third-party content, the recommendation algorithm itself, and a display feed of recommendations, like a news feed.[405] Each of these must be independently created and inserted into the system for it to work. Thus, they should be analyzed as separate products. Third-party content, unless solicited or prompted in some way by an ICS,[406] stems from a third party. The algorithm itself is developed and operated by the ICS. The resulting feed is a blend of the efforts of the two; the third-party content is sorted into a compilation by the ICS itself. To call this an equal effort, however, would not accurately reflect the process. With the minor exception of advertising, the algorithm, and not the user, controls exactly to whom content is displayed (unless the user bypasses the algorithm entirely by posting it on someone's profile).[407] The recommendation feed, however, can pinpoint exact subsets of individuals most receptive to specific types of third-party compilations.[408] A user post's destination is entirely at the mercy of the algorithm's recommendation feed.[409] Consequently, the recommendation system is almost entirely responsible when its recommendation feeds favor the inclusion of defamatory content, and interpretations of § 230 should be modified to reflect this reality.

The Ninth Circuit has already established that a search engine can trigger liability for its owner if it acts in a way that would be illegal had a person done it.[410] The individual pieces of content in a social media site are created by users, but most of their reach will come from the site's recommendation algorithm, not user effort.[411] The modern process of posting content online less resembles speaking than convincing an omnipotent machine to speak on your behalf. Where an algorithm picks and chooses which content to display based on its tortious or illegal effect, § 230 should not be construed to protect its ICS owner.

Section 230 incentivized the development of tools to encourage the good faith moderation of an undeveloped cyberspace frontier.[412] If applied to Facebook's News Feed, § 230 would protect the most efficient catalysts for defamation ever created. Even ignoring the manifold problems of online defamation, such as unidentifiable defendants, ICSs cause almost all the

403. *See supra* Parts II.B.1–2.
404. *See supra* notes 291–95 and accompanying text.
405. *See supra* Part I.C.
406. *See generally* FTC v. Accusearch, Inc., 570 F.3d 1187 (10th Cir. 2009).
407. *See supra* Part II.B.2.
408. *See supra* Part I.C.
409. *See supra* Part I.C.
410. *See supra* Part I.C.
411. *See supra* Part II.B.2.
412. *See supra* Part I.A.2.

development an online libel will receive.[413] Because of this, an ICS that spreads defamation should be considered the ICP of that content, especially where it uses it to create defamatory feeds of information.[414]

C. The Broader Impact of This Proposal

Policy considerations also favor the treatment of recommendation systems as ICPs. While § 230 promoted the development of a new free market of information, the system that has since emerged expanded the reach of all information, true or false.[415] Because the new internet favors false but novel information over verifiable but familiar information, recommendation systems that have sprung up under § 230's protection have not only undermined the economic incentive for truthful reporting but also opened the door for widespread misinformation about anyone, anytime, anywhere.[416] Much has been discussed about the growing national and global problem of "fake news" and its impact on the republic, but equally important are the problems of individual rumors, defamation, and the impact on those targeted by them.[417] The problems may even intertwine when the trustworthiness of news is affected more by trust in the third-party poster than in the news outlet that published the original story.[418]

Limiting § 230's protections for recommendation systems could force a change of course in the developing misinformation epidemic. Faced with legal consequences for actively promoting misinformation, social media developers will be forced to consider a more careful course of development. Instead of "move fast and break things,"[419] new technological innovations will have to avoid actively promoting defamatory material.[420] Because removing defamatory categories of information and users might not actually change the spread of defamation,[421] the required shift in technology might be significant and potentially costly.[422] To justify such a shift economically, one need only review the logic behind § 230's original creation: technological innovation to solve a problem will better progress if the law allows it.[423] As of now, data is extremely profitable, and multiple decades of development have gone into the current course of exploiting that data.[424] Limiting § 230's protections where recommendation systems are fundamentally responsible for defamation could raise the cost of continuing

413. *See supra* Part II.B.2.
414. *See supra* Part II.B.2.
415. *See supra* Part I.D.
416. *See supra* notes 240–44 and accompanying text.
417. *See supra* notes 240–44 and accompanying text.
418. *See* STIGLER REPORT, *supra* note 226, at 165–66.
419. GARCÍA MARTÍNEZ, *supra* note 248, at 257.
420. GOODMAN & WHITTINGTON, *supra* note 11, at 9.
421. *See supra* notes 312–14 and accompanying text.
422. *See supra* notes 376–78 and accompanying text.
423. *See* 47 U.S.C. § 230(b)(3)–(5).
424. *See generally* GANTZ & REISEL, *supra* note 173, at 16.

on the current course of development enough to incentivize the creation of a more trustworthy alternative.[425]

Holding against some recommendation systems might lead to concerns about tortfeasors running wild in a newly unregulated space. After all, if the current technologies for curation and moderation developed with the assumption of § 230 immunity, removing those protections would also set back efforts at content moderation.[426] Such fears, however, ignore the reality that those allegedly protective systems help, not hinder, online defamation's spread.[427] Removing protections for these wolves in sheep's clothing would not exacerbate the power vacuum of inadequate moderation even in the short term; if anything, its immediate effect would be to curtail it. Without the ability to continue a system that promotes defamation, the absolute worst-case scenario would be reversion to a wholly unregulated space, which, though not working against tortfeasors, would also avoid supporting them.[428] The immediate effect of this would be to reduce the spread of defamation by removing an algorithmic incentive to include it in select news feeds.

D. Limits of This Proposal

Although this Note's proposed treatment of recommendation systems has broad implications for the world of technology, its reach has limits. The Sixth Circuit demonstrated those limits in its holding for The Dirty. Where an ICS is responsible for the spread of defamation, it can, and should, be liable for that spread.[429] Sometimes, however, an ICS develops defamatory content in a neutral way—that is, without regard to its potential as a tort.[430] The Dirty, which engaged in good faith moderation for extreme obscenity but otherwise posted whatever content users submitted, acted neutrally towards that content's defamation.[431] The site never encouraged actual defamation (though it encouraged gossip and "dirt") or added defamation beyond what the third-party user initially provided.[432]

This Note proposes applying § 230 to recommendation systems because of their active role in targeting defamatory statements by inserting them into larger, intentionally defamatory recommendation feeds. Where an ICS merely develops the content at issue without favoring it *because of its defamation*, it should still qualify for § 230 protections, even where it otherwise exercises some editorial discretion. To say otherwise would defeat

425. *See* GOODMAN & WHITTINGTON, *supra* note 11, at 9.
426. *See supra* Part II.C.3.
427. *See supra* notes 312–16 and accompanying text.
428. *See* GOODMAN & WHITTINGTON, *supra* note 11, at 9.
429. *See supra* Part III.B.
430. *See, e.g.*, Herrick v. Grindr LLC, 765 F. App'x 586, 591 (2d Cir. 2019).
431. Jones v. Dirty World Enter. Recordings LLC, 755 F.3d 398, 401–02 (6th Cir. 2014).
432. *Id.* at 416–17.

the original purpose of § 230 and would likely require amending the statute to properly bring about the change.[433]

CONCLUSION

Since its passage, § 230 has done much good and accomplished part of its original goal of encouraging innovation for ICSs.[434] It has also, however, contributed to many issues of growing importance that define our time.[435] By recognizing the true nature of recommendation systems and treating them as the independent products they are, courts have the power to repair some of these issues by restoring a legal incentive for modern social media to better observe the traditional republication rule.[436]

In the case that opened this Note, plaintiff Igbonwa focused on finding the "faceless men" who ruined his life.[437] While a third party posted the content, it was Facebook who took that content, identified the users most likely to appreciate it, and created News Feed with its own proprietary algorithm to maximize spread.[438] While Igbonwa's case is a small one, it should elicit great concern. Anyone can encounter the faceless men. Anyone might now wake up one morning to find their reputation ruined, their career destroyed, and their social circle alienated by a malevolent or even careless actor, a possibility created by the modern recommendation feed, which does most of the work.[439] To help combat this modern crisis and preserve the effectiveness of defamation law, courts should recognize the responsibility for development those systems bear, deny them § 230's protections, and, where necessary, hold them accountable.

433. *See supra* Part I.A.3. *But cf.* Zipursky, *supra* note 13, at 5 (arguing that Congress's modifications of the republication rule through § 230 have been greatly exaggerated by courts' interpretations).

434. *See supra* Part I.D.

435. *See supra* Part I.D.

436. *See supra* Part III.C.

437. Igbonwa v. Facebook, Inc., No. 18-cv-02027, 2018 U.S. Dist. LEXIS 173769, at *4 (N.D. Cal. Oct. 9, 2018).

438. *See supra* Part II.B.2.

439. *See supra* Part I.D.