WILD WESTWORLD: SECTION 230 OF THE CDA AND SOCIAL NETWORKS' USE OF MACHINE-LEARNING ALGORITHMS

*Catherine Tremble**

On August 10, 2016, a complaint filed in the Eastern District of New York accused Facebook of aiding the execution of terrorist attacks. The complaint depicted user-generated posts and groups promoting and directing the perpetration of terrorist attacks. Under § 230 of the Communications Decency Act, interactive service providers (ISPs), such as Facebook, cannot be held liable for user-generated content where the ISP did not create or develop the content at issue. However, this complaint stands out because it seeks to hold Facebook liable not only for the content of third parties but also for the effect its personalized machine-learning algorithms—or "services"—have had on the ability of terrorists to execute attacks. In alleging that Facebook's actual services, as well as its publication of content, allow terrorists to more effectively execute attacks, the complaint seeks to negate the applicability of § 230 immunity.

This Note argues that Facebook's services—specifically the personalization of content through machine-learning algorithms—constitute the "development" of content and as such do not qualify for § 230 immunity. This Note analyzes the evolution of § 230 jurisprudence to help inform the development of a revised framework. This framework is guided by congressional and public policy goals and creates brighter lines for technological immunity. It tailors immunity to account for user data mined by ISPs and the pervasive effect that the use of that data has on users—two issues that courts have yet to confront. This Note concludes that under the revised framework, machine-learning algorithms' content organization—made effective through the collection of individualized data—make ISPs codevelopers of content and thus bar them from immunity.

* J.D. Candidate, 2018, Fordham University School of Law; B.A., 2013, Williams College. This Note would not have been possible without the guidance and insight of Professor Joel Reidenberg, the hard work of the *Fordham Law Review* members, and the love and support of my family and friends.

INTRODUCTION

To say that every online interaction you have is curated to influence the way you live and how you feel is hyperbolic. However, to say that one website has the influence and reach to determine the political landscape of the United States[1] or that that same site may be responsible for your foul mood on Monday evening[2] is not as alarmist as it first sounds. You might be incredulous, but both voter turnout and behavioral studies conducted on unsuspecting subjects in real time on the Facebook platform have demonstrated these hypotheticals are capable of becoming reality.[3]

Facebook attained this influence on users' behavior and mood through the use of machine-learning algorithms.[4] The impact of machine-learning algorithms has not been concretely quantified,[5] but studies highlight one definite trend: human decision-making is intensely susceptible to their influence.[6] Machine-learning algorithms' subtle but pervasive influence alters human behavior and, to a certain extent, the human experience.[7] The increasing ubiquity of machine learning has not gone unnoticed. Even technologists are hesitant about the loss of human autonomy that comes with the implementation of machine learning.[8] Although the effects of this

1. *See* Adam Rogers, *Google's Search Algorithm Could Steal the Presidency*, WIRED (Aug. 6, 2015, 1:24 PM), https://www.wired.com/2015/08/googles-search-algorithm-steal-presidency/ [https://perma.cc/MM2Z-4ALM] (describing a study in which Facebook exposed sixty-one million people to a message encouraging them to vote during the 2010 congressional elections, which generated 340,000 extra voters). This influence has raised concerns regarding the ability to create a type of "digital gerrymandering" by only targeting those users Facebook predicts are most likely to vote for certain candidates. *Id.*; *see* Zoe Corbyn, *Facebook Experiment Boosts US Voter Turnout*, NATURE (Sept. 12, 2012), http://www.nature.com/news/facebook-experiment-boosts-us-voter-turnout-1.11401 [https://perma.cc/7KKJ-FPFG].

2. It is not just a case of the Mondays. Facebook conducted a study, unbeknownst to users, which demonstrated that users' moods and behaviors were directly influenced by the tone of their newsfeeds. Positive newsfeeds yielded positive posts, and negative newsfeeds yielded negative posts. *See* Adam D.I. Kramer et al., *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*, 111 PNAS 8788, 8789 (2014).

3. *See id.*; *see also* Corbyn, *supra* note 1.

4. *See* Will Oremus, *Who Controls Your Facebook Feed*, SLATE (Jan. 3, 2016, 8:02 PM), http://www.slate.com/articles/technology/cover_story/2016/01/how_facebook_s_news_feed_algorithm_works.html [https://perma.cc/3URB-NFW2].

5. *See* Jason Tanz, *Soon We Won't Program Computers. We'll Train Them Like Dogs*, WIRED (May 17, 2016, 6:50 AM), https://www.wired.com/2016/05/the-end-of-code/ [https://perma.cc/FM5G-EH48] (noting that these algorithms do not abide by traditional coding rules, but instead operate more like a human brain, which makes them more opaque and harder to understand).

6. *See* Corbyn, *supra* note 1; *see also* Kramer et al., *supra* note 2, at 8789; *infra* Part I.C.2.

7. For further discussion on the actual neurological effects technological dependency has on human brains and behaviors, see generally NICHOLAS CARR, THE GLASS CAGE: AUTOMATION AND US (2014).

8. *See* Tanz, *supra* note 5 ("'Instead of being masters of our creations, we have learned to bargain with them, cajoling and guiding them in the general direction of our goals . . . .' This can all be pretty frightening." (quoting Danny Hillis, *The Enlightenment Is Dead, Long Live the Entanglement*, PUBPUB (Aug. 3, 2016), https://www.pubpub.org/pub/enlightenment-to-entanglement [https://perma.cc/VRZ2-HWE7])); *see also* CARR, *supra* note 7, at 2 ("[A]utomation also has deeper, hidden effects. . . . It can narrow our perspectives and limit our choices.").

technology are constantly being revealed, the application of knowledge of behavioral influence to legal theories of liability remains untested in the courts for a host of reasons detailed below.[9]

Though Facebook and other interactive service providers (ISPs)[10] are commonly used as trading posts for illegal goods and services,[11] they are rarely held liable for aiding unlawful enterprises. If Facebook were an offline publication or service, it would face intermediary liability. This theory of liability, traditionally assigned to publishers, allows a victim to hold the third-party disseminator of information liable for a content provider's misconduct.[12]

Facebook, however, has avoided intermediary liability for three main reasons: First, because most algorithms are proprietary technology, people are unaware of both the potential influence and harm to which they are exposed.[13] Second, the use of machine-learning technology and the studies demonstrating its mood- and behavior-altering influence only reach back as far as 2007 and 2012, respectively, giving legal experts little time to assess

9. One of the first complaints to address the issue of liability for algorithms deployed on a social network is being adjudicated in the courts as of the date of publication of this Note. *See* Complaint at 16, Cohen v. Facebook, Inc., 1:16-cv-04453 (E.D.N.Y. Aug. 10, 2016) [hereinafter *Cohen* Complaint]. This complaint and a related consolidated complaint were dismissed for lack of standing and due to the court's finding that Facebook was entitled to complete immunity under § 230 of the Communications Decency Act. Cohen v. Facebook, Inc., Nos. 16-CV-4453 (NGG) (LB), 16-CV-5158 (NGG) (LB), 2017 WL 2192621, at *15 (E.D.N.Y. May 18, 2017). The only cases that consider machine-learning algorithms to date are search engine defamation lawsuits, in which § 230 immunity has applied across the board. *See* Michael L. Smith, Note, *Search Engine Liability for Autocomplete Defamation: Combating the Power of Suggestion*, 2013 U. ILL. J.L. TECH. & POL'Y 313, 314 (discussing former First Lady of Germany Bettina Wulff's attempt to sue Google in Germany for autocompletes that suggested she was an "escort"). These search engine cases are distinguishable from social network cases on numerous grounds. The business goals, services offered, relationships to users, and freedom of the user to control his or her data differ in kind between social networks and search engines; as such, they do not provide dispositive precedent. Those factors are all relevant in determining whether a company's actions are deserving of § 230 immunity.

10. The term "interactive service providers" (ISPs) refers to online platforms that provide interactive services.

11. *See, e.g.*, Felicia Bolton & WMCActionNews5.com Staff, *Child Porn Posted to Facebook, Those Responsible Face Felony Charges*, WMCACTIONNEWS5.COM (May 30, 2015, 3:14 PM), http://www.wmcactionnews5.com/story/29193793/child-porn-posted-to-facebook-those-responsible-face-felony-charges [https://perma.cc/A4WY-JN65] (tracking the investigation of two men who posted photos of underage students on Facebook in violation of criminal laws); *see also* Alex Heath, *People Are Already Selling Drugs, Animals, and Adult Services on Facebook's New Craigslist Competitor*, BUS. INSIDER (Oct. 3, 2016, 5:01 PM), http://www.businessinsider.com/facebook-marketplace-drugs-animals-adult-services-2016-10 [https://perma.cc/9HAQ-Z7X7] (noting the wide variety of illegal and bizarre postings that have popped up on Facebook's new marketplace product).

12. *See* K.A. Taipale, *Secondary Liability on the Internet: Towards a Performative Standard for Constitutive Responsibility* (Ctr. for Advanced Stud., Working Paper No. 04-2003, 2003), https://ssrn.com/abstract=712101 [https://perma.cc/9XRQ-JAFG].

13. *See, e.g.*, *The Age of the Algorithm*, 99% INVISIBLE (Sept. 5, 2017), http://www.99percentinvisible.org/episode/the-age-of-the-algorithm/ [https://perma.cc/C9QS-8ZCF] (noting that "[c]omputer algorithms now shape our world in profound and mostly invisible ways"); *see also* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 8 (2015).

appropriate legal theories to apply in cases that arise.[14] Third, and most important, ISPs enjoy broad immunity under § 230 of the Communications Decency Act (CDA), which bars most claims against online intermediaries in the early pleading stages.[15] These three factors converge to create a new "Wild West" of internet regulation,[16] a realm in which an ISP will almost never face civil or criminal liability for the content it hosts, even when its curated content may increase the impetus to engage in illegal activities.

Section 230 of the CDA provides that where an ISP hosts user-generated content, none of it—however illegal—implicates the provider in the way a traditional publisher would be implicated. Instead, liability lies with the "information content provider,"[17] a term that encompasses anyone who "create[s] or develop[s]"[18] the content at issue. It is possible for an ISP to have protection as the provider of an interactive service and to lose that protection if the ISP created or developed the offending content.[19] Judges, guided by the congressional findings and policies provided in the statute,[20] have read the immunity provision broadly to protect against issues ranging from traditional defamation liability[21] to failure-to-warn claims for connecting users to sexual abusers.[22] In most cases the legal claim at issue is not addressed; such is the strength of § 230 immunity.[23] Recent years have chipped away at this expansive interpretation, however, and judges have begun to distill the importance of noting the ISP's actual involvement in the content and conduct at issue.

By drawing on several courts' analyses of § 230 immunity,[24] this Note examines the complex and novel legal situation that arises when a social media company is allegedly liable for the effective radicalization of its users, not via third-party content, but through its own individualized and influential content curation.[25] This Note also examines the claim that the

14. For an example of one of the first articles proposing a new type of liability regime specific to online intermediaries, see generally Marcelo Thompson, *Beyond Gatekeeping: The Normative Responsibility of Internet Intermediaries*, 18 VAND. J. ENT. & TECH. L. 783 (2016).

15. *See* 47 U.S.C. § 230 (2012).

16. *See* 142 CONG. REC. S1646 (daily ed. Mar. 7, 1996) (statement of Rep. Exon) ("Before the passage of the Communications Decency Act, the Internet had been described as the Wild West.").

17. 47 U.S.C. § 230(f)(3) (defining the term "information content provider" as "any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service").

18. *Id.*

19. *See, e.g.*, Chi. Lawyers' Comm. for Civil Rights Under Law, Inc. v. Craigslist, Inc., 519 F.3d 666, 670 (7th Cir. 2008) (noting that the immunity is not without limitation).

20. 47 U.S.C § 230(b)(1)–(5).

21. *See, e.g.*, Carafano v. Metrosplash.com, Inc., 339 F.3d 1119, 1125 (9th Cir. 2003).

22. *See, e.g.*, Doe v. MySpace, Inc., 528 F.3d 413, 419 (5th Cir. 2008); *see also* Doe II v. MySpace Inc., 96 Cal. Rptr. 3d 148, 151 (Ct. App. 2009).

23. *See, e.g.*, *Carafano*, 339 F.3d at 1125.

24. *See infra* Part II.C–D.

25. *See* Thompson, *supra* note 14, at 797 ("Intermediaries are the designers of the heart valves through which the lifeblood of our information environment flows. Actions they take or refrain from taking can fundamentally alter medium and message, structure and content of information we impart and receive. In other words, intermediaries can transform the very constitution of the environments we inhabit and the lives we live therein." (footnote omitted)).

algorithmically driven connective services[26] effectively aid offline radicalization by creating connections among radicalized users and promoting their inclusion in extremist events and groups.

This Note ultimately seeks to determine whether machine-learning algorithms providing individualized content continue to qualify as services envisioned for protection under § 230;[27] or whether courts should recognize the rapidly growing distinction between traditional media and new individualized and predictive media to create liability schemes that best comport with the public interest and align with legislative purpose.[28] Ultimately, it concludes that § 230 immunity should not be interpreted to protect machine-learning algorithms for the sake of convenience. Instead, § 230 must be interpreted to give meaning to its language and to bolster the legislative intent to withhold immunity from creators and codevelopers of content where their technology aids individuals and groups in their extremist agendas on social media. Such an interpretation may necessarily find that immunity cannot, in fact, be granted where machine-learning technology is employed in conjunction with user-generated content.

Part I begins with a look at the cases that inspired the CDA and an examination of the congressional findings[29] and policy goals[30] undergirding this law. Part I then outlines the development of the algorithm and its influence on Facebook's growth as a social media platform. Next, Part II explores the judicial evolution of § 230 doctrine, highlighting the important factors courts analyze when addressing a platform's eligibility for immunity.

Part III proposes an updated framework for analyzing § 230 immunity and applies this framework to determine whether social media platforms that use machine-learning algorithms should be granted immunity. This reenvisioned framework is applied to the facts in *Cohen v. Facebook, Inc.*,[31] where plaintiffs sought to implicate Facebook for aiding terrorist connections, meetings, and event attendance through its ability to curate engaging content for each user's predicted interests through negligence claims and claims under the statute prohibiting material support to designated foreign terrorist

---

26. The term "connective services" refers to services that recommend groups, friends, or activities to the user without any prompting.

27. 47 U.S.C. § 230(f)(4) (2012).

28. *See, e.g.*, Thompson, *supra* note 14, at 802 (differentiating online intermediaries by the amount of control they have over content and noting that online-intermediary immunity does not create a regime that is grounded in normative concerns of justice). Thompson advocates for an approach that asks what justice requires of intermediaries. Specifically, he notes that state action should occur "whenever social forms of harm leave us without a meaningful range of options based on which to author our lives" through actions that have "the 'forward-looking aspect' of 'diminishing our prospects', of 'adversely affecting our possibilities.'" *Id.* (quoting JOSEPH RAZ, THE MORALITY OF FREEDOM 108 (1986)).

29. *See* 47 U.S.C. § 230(a) (stating Congress's findings that the internet: (1) has huge educational and informational value, (2) has a diversity of viewpoints and cultures that stimulate intellectual curiosity, and (3) requires very little interference from government to function well).

30. *See* 47 U.S.C. § 230(b); *see also id.* § 230(a).

31. Nos. 16-CV-4453 (NGG) (LB), 16-CV-5158 (NGG) (LB), 2017 WL 2192621 (E.D.N.Y. May 18, 2017).

organizations (FTOs).[32] This Part focuses on answering two questions. The first is whether algorithmic-based services—which include Facebook's "conduct" of connecting people based on their interests—are properly categorized as the "conduct" of a publisher and as such qualify for immunity. The second is whether the effects of algorithmic influence on users, in combination with the ISP's intent to engage users as effectively as possible, constitutes codevelopment of the content shown to users. Finally, Part III discusses the policy rationales behind withholding immunity in cases where algorithmically manipulated content or the conduct of the ISP is at issue. This Part assesses the competing concerns of the administrability of online-intermediary liability with the importance of transparency concerning algorithms with behavioral influence.

I. SPURRING INTERNET GROWTH: THE CDA'S CREATION AND INTERNET INNOVATIONS THAT FLOURISHED AFTER ITS PASSAGE

After outlining the historical events and political motivations that formed the basis for the Communications Decency Act (CDA), this Part outlines Facebook's evolving technology and corresponding growth in the post-CDA world. Part I.A provides the legislative background courts use to justify their decisions. Part I.B then briefly explains how Facebook's machine-learning algorithms function. Finally, Part I.C traces the evolution of Facebook from a startup to a Fortune 500 company that arrests the attention of its users with ever-increasing ingenuity and precision. Facebook's growth reveals how and why machine-learning algorithms are developed and their intended and actual effects on users, which will serve as important considerations when applying § 230 in Part III.

A. Cleaning Up the Internet: The Passage of the CDA

This Part discusses the two cases that spurred the creation of the CDA. These cases created a liability scheme that could have stunted the growth of the internet writ large. But Congress's actions to correct the untenable liability scheme created by the application of common law principles to ISPs allowed the internet to flourish instead of flounder. This section discusses the drafting attempts to promote constitutionally protected speech online while limiting the harmful—and sometimes illegal—content available online.

1. Untenable Online Liability: The Tension Created by *Stratton* and *Cubby*

Section 230 was written to remedy the asymmetrical application of intermediary liability to ISPs. In the fledgling period of the internet, any company that allowed users to freely post comments on its online platform

32. *See* 18 U.S.C. § 2339B (2012). *See generally Cohen* Complaint, *supra* note 9; *see also* Complaint, Force v. Facebook, No. 1:16-cv-05158 (S.D.N.Y. July 10, 2016) [hereinafter *Force* Complaint].

exposed itself to massive liability.[33] Under intermediary liability, ISPs were liable for any content on their sites, such as defamatory comments by users, if the site exercised any type of editorial power over any comments—even deletion.[34]

For example, in *Stratton Oakmont, Inc. v. Prodigy Services Co.*,[35] Prodigy, the site host, was found liable as the "publisher" of statements posted by a third-party user of their online bulletin Money Talk.[36] The court held that because Prodigy used an automated screening tool for offensive language and removed some offending content, it acted as a "publisher" and was liable for the defamatory statements posted.[37] The holding extended common law publisher liability (a specific type of intermediary liability) to online services.

This ruling created an incentive for ISPs to forgo monitoring entirely to escape publisher liability.[38] This reality was realized in *Cubby, Inc. v. CompuServe, Inc.*,[39] where a company escaped liability for a user's defamatory statements because it acted merely as a "distributor" of the information.[40] Because CompuServe had not altered the post in question and did not remove or alter content generally, it was not considered a publisher.[41] Instead, the court compared CompuServe to a bookstore, a common example of a distributor.[42] To establish intermediary liability of a distributor under common law, the plaintiff must prove that the defendant had actual knowledge of defamation.[43] Thus, *Stratton* controlled and the ISP was liable where an ISP attempted to remove offensive content; but where the ISP did

33. *See generally* Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710, at *6 (N.Y. Sup. Ct. May 24, 1995).

34. *Id.*; *see also* David S. Ardia, *Free Speech Savior or Shield for Scoundrels: An Empirical Study of Intermediary Immunity Under Section 230 of the Communications Decency Act*, 43 LOY. L.A. L. REV. 373, 392 n.74 (2010) (noting that contributory liability, the form of intermediary liability commonly at issue in CDA cases, "applies when a party 'with knowledge of the infringing activity, induces, causes or materially contributes to the infringing conduct of another'" (quoting Gershwin Publ'g v. Columbia Artists Mgmt., 443 F.2d 1159, 1162 (2d Cir. 1971))). Ardia also notes that, "[i]n the criminal context, aiding-and-abetting and conspiracy may also create secondary liability for intermediaries. Under secondary liability doctrines, intermediaries generally do not take on an affirmative duty to act or to prevent tortious or illegal conduct, but only a duty not to facilitate known wrongdoing." *Id.*

35. No. 31063/94, 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995).

36. *Id.* at *6.

37. *Id.*

38. *See* Cubby, Inc. v. CompuServe, Inc., 776 F. Supp. 135, 139 (S.D.N.Y. 1991) (holding that the online platform was not liable for libel as a *distributor* where it neither knew nor had reason to know of the libelous posting).

39. 776 F. Supp. 135 (S.D.N.Y. 1991).

40. *Id.*

41. *Id.*

42. *Id.*

43. *Id.* ("Ordinarily, 'one who repeats or otherwise republishes defamatory matter is subject to liability as if he had originally published it. . . .' With respect to entities such as news vendors, book stores, and libraries, however, 'New York courts have long held that vendors and distributors of defamatory publications are not liable if they neither know nor have reason to know of the defamation.'" (first quoting Cianci v. New Times Publ'g Co., 639 F.2d 54, 61 (2d Cir. 1980); then quoting Lerman v. Chuckleberry Publ'g, Inc., 521 F. Supp. 228, 235 (S.D.N.Y. 1981))).

not attempt to remove content, there was no liability under *CompuServe*.[44] The common law created a regime where ISPs were better served by leaving offensive or defamatory information online because removing it created liability.

Congress recognized that these cases created untenable liability for newly emerging internet businesses and the perverse incentive to refrain from removing or editing inappropriate content.[45] To prevent the shadow of liability from chilling commerce and to incentivize filtering inappropriate content, Congress passed § 230 of the CDA in 1996 as part of a greater attempt to purify the internet.[46] Section 230(c) absolved online platforms of liability in which they were treated as the "publisher or speaker" of the information, in an effort to encourage the filtration and removal of inappropriate and illegal content.[47]

2. The Congressional Solution: Immunity for "Good Samaritans"

The CDA began as an attempt to purify the internet of its amassed catalog of indecent content.[48] Senator J. James Exon introduced the Act to curtail the growing access to pornography that the internet abetted.[49] Through § 223 of the CDA, Exon proposed to extend common-carrier indecency laws to

44. *See id.*

45. *See infra* Part I.A.2.a.

46. *See* Communications Decency Act of 1996, Pub. L. No. 104-104, § 509, 110 Stat. 133, 137 (codified at 47 U.S.C. § 230 (2012)) (stating that the first policy goal was "to promote the continued development of the Internet"); *see also* 47 U.S.C. § 230(b)(2) (stating that the second policy goal was "to preserve the vibrant and competitive free market that presently exists for the Internet").

47. *See* 47 U.S.C. § 230.

48. This characterization of the internet was supported by a study published in the *Georgetown Law Review* that proclaimed that the internet was comprised of 83.5 percent pornography. *See* Marty Rimm, *Marketing Pornography on the Information Superhighway: A Survey of 917,410 Images, Descriptions, Short Stories, and Animations Downloaded 8.5 Million Times by Consumers in over 2000 Cities in Forty Countries, Provinces, and Territories*, 83 GEO. L.J. 1849, 1867 (1995). The study was subsequently deemed scientifically flawed. *See* Robert Cannon, *The Legislative History of Senator Exon's Communications Decency Act: Regulating Barbarians on the Information Superhighway*, 49 FED. COMM. L.J. 51, 55 n.16 (1996). However, the characterization persisted and heavily influenced Senate votes in favor of the CDA as they did not want to appear "pro-pornography." *See id.* at 64.

49. *See* 141 CONG. REC. S1953 (daily ed. Feb. 1, 1995) (statement of Rep. Exon) (noting that the internet could be kept from becoming a "red light district" by extending "the standards of decency which have protected telephone users to new telecommunications devices").

ISPs.[50] It passed in the Senate by a large margin[51] but was met with pushback in the House,[52] as it would allow the FCC to determine the meaning of decency on the internet, criminalize vast swaths of speech,[53] and was unlikely to pass constitutional muster.[54] In keeping with the spirit of protecting society from indecency while simultaneously limiting the government regulation of the internet, Representatives Jim Cox and Ron Wyden created § 230.[55]

Cox and Wyden used § 230 to reverse the liability regime established in *Stratton*, which was the major obstacle to promoting private filtration of indecent content.[56] They sought to establish the freedom for private actors to determine, without civil liability, what constituted indecency without government interference.[57] ISPs supported this approach as a way to avoid

50. *See* Cannon, *supra* note 48, at 57 ("The CDA . . . extends the antiharassment, indecency, and antiobscenity restrictions currently placed on telephone calls to 'telecommunications devices' and 'interactive computer services.'"). This legislation made it illegal to "knowingly send to or display in a manner available to 'a person under 18 years of age, any comment, request, suggestion, proposal, image, or other communication that, in context, depicts or describes, in terms patently offensive as measured by contemporary community standards, sexual or excretory activities or organs, regardless of whether the user of such service placed the call or initiated the communication.'" *Id.* at 58 (quoting 47 U.S.C. § 223(d)(1)(B)).

51. *See* Roll Call Vote 104th Congress—1st Session, U.S. SENATE, https://www.senate.gov/legislative/LIS/roll_call_lists/roll_call_vote_cfm.cfm?congress=104&session=1&vote=00263 [https://perma.cc/RV2R-MP5P] (last visited Oct. 16, 2017) (showing Sen. Exon's proposal passing by a vote of 84-16).

52. *See* Cannon, *supra* note 48, at 66 (characterizing the House reception of the CDA as "frigid"). The House had recently taken an affirmative step toward integrating with the internet, and deregulation was also a top priority. *See* Cannon, *supra* note 48, at 53; Ann Reilly Dowd, *The Net's Surprising Swing to the Right*, FORTUNE (July 10, 1995), http://archive.fortune.com/magazines/fortune/fortune_archive/1995/07/10/204243/index.htm [https://perma.cc/PG9G-VDLN] (noting a successful initiative to post all congressional legislation records online).

53. *See* 141 CONG. REC. H8470 (daily ed. Aug. 4, 1995) (statement of Rep. Wyden) (noting that Senator Exon's Act would take the task of defining indecent communications away from communities, resting it in the hands of the government).

54. *See id.* at H8472 (statement of Rep. Goodlate) (finding that the Cox-Wyden amendment likely "doesn't violate free speech or the right of adults to communicate with each other").

55. *See id.* (statement of Rep. Wyden) (stating that the law should provide relief from indecency on the web "without Federal regulation" being involved).

56. H.R. Rep. No. 104-458, 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."); *see also* David Lukmire, *Can the Courts Tame the Communications Decency Act?: The Reverberations of* Zeran v. America Online, 66 N.Y.U. ANN. SURV. AM. L. 371, 380 (2010).

57. Contrary to what is stated as the primary reasoning for the statute in many cases, the driving force behind the Cox-Wyden amendment was more likely the result of Congress's desire to keep the internet free of regulation than its desire to protect robust free speech. *See* Lukmire, *supra* note 56, at 380–81 ("Representatives Cox and Wyden envisioned that their amendment would discourage bureaucratic oversight and thereby encourage the robust growth of the Internet, largely by avoiding the unappetizing regulatory implications of Senator Exon's proposal. Their prevailing aim was not to create a liability shield . . . .").

the chilling effects of *Stratton*'s intermediary liability.[58] When the U.S. Supreme Court struck down the criminalization of indecency in § 223 as overbroad and violative of the First Amendment,[59] § 230 remained.

Section 230(c), titled "Protection for 'Good Samaritan' blocking and screening of offensive material," is the immunity provision.[60] It is broken into two sections: (1) The "Treatment of Publisher or Speaker" and (2) "Civil Liability."[61] Under § 230(c)(1), "[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."[62] And, under § 230(c)(2), no ISP will be liable for good-faith restriction of material that may be objectionable, regardless of whether that material is constitutionally protected.[63]

Section 230(c)(1) and (2), taken together, provide that claims seeking to hold an ISP responsible for conduct that would have implicated an offline publisher are barred.[64] Instead, liability only lies with the "information content provider." This term, as defined in § 230(f), encompasses anyone who creates or develops the information at issue.[65] Under this statute, it is possible for an ISP to fall under § 230(c) protection as a provider of an "interactive computer *service*," as well as outside of that same protection as a "content provider" or developer of content.[66]

The procedural application of § 230(c) is unclear.[67] A majority of courts consider § 230 immunity prior to discovery, and of those courts, a majority never address whether this the proper procedure.[68] Some courts, however, treat § 230 as an affirmative defense, which allows plaintiffs to plead claims without having to address the immunity provision.[69] This procedural choice

58. Robert T. Langdon, Note, *The Communications Decency Act § 230: Make Sense? Or Nonsense?—A Private Person's Inability to Recover If Defamed in Cyberspace*, 73 ST. JOHN'S L. REV. 829, 844 (1999).

59. *See* Reno v. ACLU, 521 U.S. 844, 885 (1997) (discussing the unconstitutionality of § 223(a) and (d) and striking them down, with the exception of the portions regarding child pornography).

60. 47 U.S.C. § 230(c) (2012).

61. *Id.*

62. *Id.* § 230(c)(1).

63. *Id.* § 230(c)(2).

64. *See id.* § 230(c)(1)–(2).

65. *See id.* § 230(f)(3) (defining the term "information content provider" as "any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service").

66. *See, e.g.*, Chi. Lawyers' Comm. for Civil Rights Under Law, Inc. v. Craigslist, Inc., 519 F.3d 666, 670 (7th Cir. 2008).

67. *See* Ardia, *supra* note 34, at 482–83 (finding there is not consensus in the case law "as to whether section 230 can be raised in a motion to dismiss or otherwise addressed before the parties have had a chance to engage in discovery").

68. *See id.* at 483. Ardia notes that, of the 51 percent of cases that consider § 230 prior to full discovery, roughly 85 percent do not raise the issue of the proper timing of this defense. *Id.* Additionally, courts only refused to address the issue of § 230 before discovery in 7.6 percent of cases. *Id.*

69. *See, e.g.*, Barnes v. Yahoo!, Inc., 570 F.3d 1096, 1100 (9th Cir. 2009); *see also* Doe v. GTE Corp., 347 F.3d 655, 657 (7th Cir. 2003) (holding that § 230 is an affirmative defense

is important as it may determine whether the court permits the case to be dismissed in the initial pleading stages on the basis of immunity, whether the defendant must answer the complaint, and whether there can be discovery.[70] Certain courts have allowed limited discovery before addressing a motion to dismiss,[71] but fact patterns are often simple and require little, if any, discovery.[72]

The procedural use of § 230 may be pivotal in any case that implicates the use of machine-learning algorithms in a complaint. Without access to discovery it is almost impossible to determine whether algorithms had an effect on specific terrorists' ability to organize successful attacks because information about user accounts is likely proprietary.[73] This is where the divide from many previous cases is most pronounced: the nonobvious operation of algorithms and the limited access to user data make content development difficult to assess and, further down the road, make culpability difficult to assign.[74]

B. The Machine-Learning Algorithm Explained

Before discussing the use of machine-learning algorithms and their effects on business, it is helpful to discuss how they differ from conventional programming, and the importance of that difference. In conventional programming, programmers create discrete sets of step-by-step instructions to be followed by a computer.[75] This type of programming was likely employed in the initial construction of Facebook's newsfeed.[76] For instance, a program might be written to instruct a newsfeed to rank and display recent posts from the people a user communicated with most.[77] This stands in stark

---

and that because affirmative defenses cannot justify a dismissal under Fed. R. Civ. P. 12(b)(6), the plaintiffs are not required to "plead around" the § 230 defense).

70. *See* Ardia, *supra* note 34, at 482 (noting that "some courts have refused to permit the plaintiff to engage in discovery until the court addressed whether section 230 preempted the claims at issue in the case" because of cost considerations).

71. *See, e.g.*, Ben Ezra, Weinstein & Co. v. Am. Online, Inc., 206 F.3d 980, 983–84 (10th Cir. 2000) (granting the plaintiff the opportunity to conduct limited discovery on issues relating to whether the defendant qualified for § 230 immunity).

72. *See* Eric Taubel, Note, *The ICS Three-Step: A Procedural Alternative for Section 230 of the Communications Decency Act and Derivative Liability in the Online Setting*, 12 MINN. J.L. SCI. & TECH. 365, 369 (2011) (noting that the evolution of technology has caused the inquiry of § 230 to become an increasingly fact-intensive analysis).

73. *See* PASQUALE, *supra* note 13, at 3–4 (noting the intentional information imbalance where ISPs seek to aggregate all information about users, but divulge none of it to maintain power).

74. *See id.* Pasquale goes on to note that users are left "in the dark" about the "[s]ecret algorithmic rules for organizing information, and wars against those who would defeat" or try to learn those rules. *Id.* at 66. He also points out that users are not technology companies' customers; they are the product sold to advertisers. *Id.*

75. *See* Tanz, *supra* note 5.

76. *See* Oremus, *supra* note 4 (describing the first newsfeed engine as a "crude algorithm" based on what engineers believed people might like).

77. This is a hypothetical. Facebook does not share the details of their newsfeed rankings. *See* Caleb Garling, *Tricking Facebook's Algorithm*, ATLANTIC (Aug. 8, 2014), http://www.theatlantic.com/technology/archive/2014/08/tricking-facebooks-algorithm/375801/ [https://perma.cc/8ZTS-4UGM].

contrast to machine-learning algorithms, which are defined by their self-teaching abilities.[78]

A machine-learning algorithm ingests information and creates inferences in order to categorize and act based on overarching goals defined by programmers.[79] In traditional programming, mechanisms operate as the result of concrete rules; as such, problems are solved by correcting the programmers' previously written rules to yield a different output.[80] By contrast, if the output of a machine-learning algorithm is unsatisfactory, the program needs more exposure to trial and error; it will self-teach to achieve its goal.[81]

The difficulty with machine-learning programs is that the engineers are no longer fully in control. These programs' operations are not digestible lines of code but rather an indecipherable web of what has been learned, over time, to ascertain categorizations.[82] There are not rigid rules designating a path to the desired destination and so there are no concrete judgments to be corrected.[83] Outsourcing decision-making processes to technology that has no discernable logic may have unknowable consequences on human perceptions of self and the world at large.[84]

C. Facebook's Evolution: Developing the Science of Engagement

Facebook engineers work to find the most relevant material for users to consume.[85] In this effort, their techniques have evolved from rudimentary guesses to science-based methods and behavioral testing.[86] The mission is to create a behavioral-response cycle to a research process that Pavlov would envy.[87] In the struggle to discern "relevancy score[s]" for posters, the development team has expressed a philosophy that centers on the importance

78. Tanz, *supra* note 5.
79. *Id.*
80. *Id.*
81. *Id.* An excellent hypothetical is excerpted here:

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

Id.
82. *Id.*
83. *Id.*
84. *Id.* ("And as these black boxes assume responsibility for more and more of our daily digital tasks, they are not only going to change our relationship to technology—they are going to change how we think about ourselves, our world, and our place within it.").
85. *See* Oremus, *supra* note 4.
86. *See id.*; *see also supra* note 2 and accompanying text; *infra* Part I.C.2.
87. *See* Bianca Bosker, *The Binge Breaker*, ATLANTIC (Nov. 2016), http://www.theatlantic.com/magazine/archive/2016/11/the-binge-breaker/501122/ [https://perma.cc/BW5T-W25U] (arguing that "technology has become better at controlling us").

of aggregating the most meaningful experience for each individual: the one that will most influence time spent on the site.[88]

This Part briefly outlines the history of Facebook's platform—specifically looking at how and why its algorithms developed—and then explores the effects those algorithms have on users. This analysis aids the later consideration of machine-learning algorithms' relationship to profitability and their influence on human decision-making, which will be important when assessing whether content is developed and thus ineligible for immunity under § 230.

1. The Evolution from "Likes" to "Relevancy Scores"

Facebook began as a directory-type network in 2004.[89] The newsfeed—introduced in 2006 to show users their friends' updates and photos—was first ordered based on the guesswork of engineers.[90] In 2007, the "like" button was introduced to begin to eliminate the guesswork and to facilitate understanding of engagement on the site.[91] This innovation was significant as it marked the first time Facebook employed an algorithm that was tied to user action.[92] The "like" button was accompanied by a feature that allowed users to delete posts that they did not want to see.[93] Both additions allowed Facebook to begin to understand user behavior on an individualized basis and predict what specific users wanted to see. Users were now equipped to unwittingly improve their engagement with the site by telling Facebook precisely what they found interesting.[94]

In 2009, an update to the algorithm prioritized posts not only based on user's previous interests but also on a post's ability to gain "likes."[95] However, in 2013, Facebook realized that users "like" things that they don't necessary engage with so that newsfeeds were filling with posts that were more reactionary than engaging.[96] Continuing to use such a crude metric would undoubtedly stilt engagement.[97] To combat this, Facebook utilized

88. *See* Oremus, *supra* note 4.

89. *Facebook Newsfeed Algorithm History*, WALLAROO (Aug. 17, 2017), http://wallaroomedia.com/facebook-newsfeed-algorithm-change-history/ [https://perma.cc/Z8QN-DC74].

90. *See* Oremus, *supra* note 4.

91. *See Facebook Newsfeed Algorithm History*, *supra* note 89.

92. *Id.*

93. *Id.*

94. *See* Oremus, *supra* note 4 ("That users didn't realize they were doing this was perhaps the most ingenious part. . . . Facebook's news feed algorithm was one of the first to surreptitiously enlist users in personalizing their experience—and influencing everyone else's.").

95. *See id.*

96. *Id.* ("Publishers, advertisers, hoaxsters, and even individual users began to glean the elements that viral posts tended to have in common—the features that seemed to trigger reflexive likes . . . . Social-media consultants sprung up to advise people on how to game Facebook's algorithm: the right words to use, the right time to post, the right blend of words and pictures. 'LIKE THIS,' a feel-good post would implore, and people would do it, even if they didn't really care that much about the post.").

97. *See id.*

metrics like time spent on the page, the relative amount of time users spent on some posts over others, and whether someone had "liked" the post before or after engaging with it.[98] This structure even logs user behaviors that are not direct interactions with the site in order to learn what constitutes more engaging material for individual users.[99]

As the data flowed in, Facebook was able to better predict and distill its users' content desires. The rise of machine-learning algorithms, driven by historic behavioral data, has made it apparent that companies can "use code to understand our most intimate ties."[100] Facebook's founder and CEO Mark Zuckerburg has suggested that there might be a "fundamental mathematical law underlying human relationships that governs the balance of who and what we all care about."[101]

The intense focus on the development of technology that better understands the mathematics of human relationships is driven primarily by the byproduct: money. For many online businesses, revenue comes primarily from advertising.[102] Advertisers pay more for ad space on websites where users spend more time, exposing the user to the ad for a longer period. This pricing structure has created an "attention economy," in which companies are incentivized to hold users' attention to increase ad revenue.[103] This business model has enabled a "race to the bottom of the brain stem."[104] Facebook is no exception to this rule; their science is driven and funded by a need to increase revenues.[105]

When user-generated data become the basis for revenue, those data become necessary for the continuing success of the business.[106] Under a § 230 analysis, this information becomes relevant to the assessment of the level of development that the content has undergone in service of the ISP's ends. Courts have considered whether a defendant has a vested interest in the collection of information when determining whether the actions of the ISP fall outside the boundaries of publisher actions.[107]

---

98. *See id.*
99. *See id.*
100. Tanz, *supra* note 5.
101. *Id.*
102. *See* Lev Grossman, *The Great Ad-Blocker Battle*, TIME (Oct. 8, 2015) http://time.com/4065962/our-attention-is-just-a-pawn-in-the-great-game-of-silicon-valley/ [https://web.archive.org/web/20151012050137/http://time.com:80/4065962/our-attention-is-just-a-pawn-in-the-great-game-of-silicon-valley/]; *see also supra* note 74.
103. Bosker, *supra* note 87 (comparing the tech industry to "Big Tobacco before the link between cigarettes and cancer was established" because the industry is "keen to give customers more of what they want, yet simultaneously inflicting collateral damage on their lives").
104. *Id.*
105. *See* Jane Wakefield, *What Is Facebook Doing with My Data?*, BBC NEWS (Nov. 10, 2015), http://www.bbc.com/news/magazine-34776191 [https://perma.cc/8TX8-WTB7] ("Advertising revenue is Facebook's biggest source of income, jumping 45 percent this year.").
106. *See* PASQUALE, *supra* note 13, at 66; *supra* note 74 and accompanying text.
107. *See* Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 521 F.3d 1157, 1172 (9th Cir. 2008) (en banc) ("Roommate both elicits the allegedly illegal content and makes aggressive use of it in conducting its business."); *see also id.* at 1166 ("Roommate makes answering the discriminatory questions a condition of doing business.").

2. Facebook's Demonstrable Effects on Behavior

Facebook's interactive experience represents much more than the bulletin board in *Stratton*.[108] Facebook's use of algorithms has tangible effects on its bottom line and intangible effects on its users. The same algorithms that are focused on increasing time spent on the site can influence users in ways that go beyond creating bad time-management habits. In two separate studies, Facebook demonstrated its actual power to control users. One study (the "mood study") shows how, by adjusting the tone of messages on users' newsfeeds, Facebook can make a certain type of response more likely.[109] The second study (the "voting study") shows that, by feeding people messages related to voting, Facebook enticed 340,000 additional voters to go to the polls.[110]

The mood study was conducted in January 2012.[111] For one week, newsfeeds of almost 700,000 people were adjusted to show happier-than-average posts or sadder-than-average posts in an effort to see if the messages sparked any type of response.[112] They did.[113] Users began posting messages corresponding to the mood they had been primed with.[114] This study was controversial because it was conducted without the express consent of the people involved.[115]

The mood study's express purpose was to make Facebook users' content "as relevant and engaging as possible."[116] The engineers were testing whether overexposure to negativity on users' newsfeeds would lead to avoidance of the platform.[117] Presumably, Facebook now knows the extent of the influence of its algorithms and has used that knowledge to manage users' feeds.

The voting study went further. This study demonstrated the link between online and offline behaviors.[118] In the study, Facebook pushed generic

---

108. *See* Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710, at *6 (N.Y. Sup. Ct. May 24, 1995).

109. *See* Kramer et al., *supra* note 2, at 8788 (showing "that emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness" and without direct interaction between people).

110. *See* Corbyn, *supra* note 1.

111. Kramer et al., *supra* note 2, at 8789.

112. *Id.* at 8788–89.

113. *Id.* at 8789.

114. *Id.*

115. Robinson Meyer, *Everything We Know About Facebook's Secret Mood Manipulation Experiment*, ATLANTIC (June 28, 2014), http://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/ [https://perma.cc/V945-75Q6] (noting that there was no legal recourse as language in the terms of service suggested this possibility).

116. *Id.*

117. Gail Sullivan, *Facebook Responds to Criticism of Its Experiment on Users*, WASH. POST (June 30, 2014), https://www.washingtonpost.com/news/morning-mix/wp/2014/06/30/facebook-responds-to-criticism-of-study-that-manipulated-users-news-feeds/ [https://perma.cc/G7ZP-2VJY].

118. Rogers, *supra* note 1 ("In their paper, Epstein and Robertson equate digital gerrymandering to what a political operative might call GOTV—Get Out the Vote, the mobilization of activated supporters. It's a standard campaign move when your base agrees

messages to sixty-one million users extolling them to vote.[119] For certain users, the message was augmented with notifications that other people in their network had already voted.[120] Facebook calculated that by the end of the 2010 congressional election, it had influenced an additional 340,000 people to go to their polling places.[121]

Although no court has yet been faced with evaluating the specific level of influence created by Facebook's algorithms, when assessing whether the defendant is a codeveloper of content the court will look to the extent to which users are restricted when generating and reviewing their own content and the content of others.[122] To the extent that user content or actions can be directly linked to what users have seen, data collected by Facebook to display the consumed content plausibly played a role in users' overall decision-making.[123]

II. THE EVOLUTION OF § 230 IMMUNITY

This Part outlines the judicial interpretation of § 230. Part II.A begins with a discussion of *Zeran v. America Online, Inc.*,[124] the first court of appeals case interpreting § 230. Using § 230's policy goals as a guide, *Zeran* established the broad construction of the statute and created the three-prong test used to determine whether § 230 immunity applies.[125] Next, Part II.B traces *Zeran's* impact on later courts' broad construction of immunity and notes logical difficulties that occur when the statute is stretched beyond its limit. Part II.C then outlines the backlash to broad immunity sparked by *Fair Housing Council of San Fernando Valley v. Roommate.com* ("*Roommates*")[126] and discusses the impact of that case on the analysis of ISP's as developers, rather than creators, of content. Part II.D concludes with an overview of the effects of *Roommates* on § 230 doctrine.

Since its debut in the Fourth Circuit in 1997,[127] § 230 immunity has been a powerful tool for ISP defendants. Although legislative history suggests that immunity was intended to apply to third-party publishers accused of

---

with your positions but isn't highly motivated—because they feel disenfranchised, let's say, or have problems getting to polling places.").

119. *Id.*

120. *See* Corbyn, *supra* note 1.

121. *Id.*

122. *See* Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 521 F.3d 1157, 1167 (9th Cir. 2008) (en banc).

123. *See* Tom Simonite, *What Facebook Knows*, MIT TECH. REV. (June 13, 2012), https://www.technologyreview.com/s/428150/what-facebook-knows/ [https://perma.cc/UW2R-3CFK] (discussing a study that found that "our close friends strongly sway which information we share, but overall their impact is dwarfed by the collective influence of numerous more distant contacts—what sociologists call 'weak ties'").

124. 129 F.3d 327 (4th Cir. 1997).

125. *See* 47 U.S.C. § 230(a) (2012); *see also Zeran*, 129 F.3d at 330 (holding that the policies and findings of Congress require a broad reading of immunity).

126. 521 F.3d 1157 (9th Cir. 2008) (en banc).

127. *See Zeran*, 129 F.3d at 327.

defamation, as in *Stratton*,[128] the Fourth Circuit interpreted immunity to bar all tort-based liability.[129] In the spirit of broad immunity, subsequent courts have construed "creation or development" in the statute narrowly, meaning an ISP's conduct is likely to be characterized as "publisher" conduct and receive immunity as opposed to conduct being considered codevelopment of content.[130] However, rebellions against this broad judicial construction have been cropping up since 2004.[131] Courts have intermittently broadened what it means to "develop" information and narrowed what can be considered "traditional publishing activities," which creates caveats in this once-immutable doctrine.

A. Establishing Expansive Immunity: Zeran

Zeran marks the first instance of § 230's judicial interpretation.[132] The *Zeran* court handed down pivotal holdings that influenced the scope of immunity moving forward. The first was the court's determination that the word "publisher" in the statute would take on its plain meaning, as opposed to the legal meaning in defamation law.[133] A website qualified as a "publisher" when it undertook any type of publication-related activity, such as posting another's content or making small grammatical edits.[134] The court also noted that, regardless of the provision's origins in defamation law, § 230 provided general immunity to all theories of liability.[135] These two interpretations, which were by no means required,[136] set the tone for courts to apply a broad construction of § 230 and expand immunity in other ways.

In *Zeran*, an anonymous user of an AOL bulletin board, purporting to be Kenneth Zeran, offered T-shirts for sale with slogans praising the Oklahoma City bombings.[137] The user also included Zeran's personal contact information.[138] As a result, Zeran received death threats and hate mail for the ensuing month.[139] Zeran contacted AOL on numerous occasions and

128. Stratton Oakmont, Inc. v. Prodigy Servs. Co., No. 31063/94, 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995).

129. *Zeran*, 129 F.3d at 330 (noting that immunity applies "to any cause of action" that seeks to impose liability on ISPs for information originating with users).

130. *See, e.g.*, Blumenthal v. Drudge, 992 F. Supp. 44, 50 (D.D.C. 1998) (holding that AOL was not the creator or developer of defamatory content, despite the editorial control they had over editorialist Matt Drudge's gossip website and the licensing agreement they had with Drudge to produce content).

131. *See* Barrett v. Rosenthal, 9 Cal. Rptr. 3d 142, 167 (Ct. App. 2004) (holding that "section 230 does not restrict distributor liability under the common law"), *rev'd*, 146 P.3d 510 (Cal. 2006).

132. *Zeran*, 129 F.3d at 327.

133. *See* Lukmire, *supra* note 56, at 397. ("After *Zeran*, courts expanded the reach of section 230 immunity beyond defamation law not by construing the term 'publisher' in subsection (c)(1) as a term of art, but instead using it as it appears in ordinary parlance.").

134. *See id.* at 390–91.

135. *See Zeran*, 129 F.3d. at 330.

136. *See* Lukmire, *supra* note 56, at 384.

137. *See Zeran*, 129 F.3d at 329.

138. *Id.*

139. *Id.*

asked the company to remove the posts and post a redaction, but AOL's attempted removal was ineffective, and it refused to post a redaction.[140]

Zeran then sued AOL and claimed that AOL's behavior was negligent under a "distributor" liability scheme[141] and that, as such, AOL was ineligible to claim § 230 immunity, which applied only to *publishers* of information.[142] The Fourth Circuit explained that a narrow interpretation of "publisher" was out of line with Congress's stated policy goals and concluded that a broad construction was warranted.[143] Accordingly, the court had properly found that distributor liability was a subsection of publisher liability.[144]

Even though the Fourth Circuit was only presented with the issue of whether § 230 granted immunity for negligence as well as defamation, the court indicated that § 230 could protect defendants from any cause of action that treated it as the "publisher or speaker" of information provided by a user.[145] This maneuver—untethering § 230 from defamation (and possibly from torts altogether)[146]—influenced almost every ensuing judicial consideration.[147] This holding created a no-holds-barred immunity under which some courts claim there are only two areas of law exempt from § 230's protection: intellectual property law and federal criminal law.[148]

B. The Norm of Expansion of Immunity: Zeran*'s Legacy*

Zeran's influence on § 230 jurisprudence manifests in two ways. Subsequent courts have used it to support a strong presumption in favor of immunity and applied the test from *Zeran* to evaluate eligibility for

---

140. *Id.*

141. Publishers can be held liable without proof of specific knowledge of the defamatory statement's existence in the work, but to find a distributor's conduct tortious, the minimum requirement is actual knowledge of the defamatory statements. *See id.* at 330–31.

142. *Id.*; *see also* 47 U.S.C. § 230(c)(1) (2012) ("No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.").

143. *Zeran*, 129 F.3d. at 330 (stating that the purpose of the act was to "preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, *unfettered by Federal or State regulation*" (quoting 47 U.S.C. § 230(b)(2)).

144. *Id.* at 332 (finding that distributor and publisher were separate tiers of liability under the same cause of action—"publication" (citing W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 113 (5th ed. 1984))).

145. *Id.* at 330.

146. *Id.* ("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 that service." (emphasis added)).

147. *See* Lukmire, *supra* note 56, at 384 (noting that *Zeran* has been cited in almost every decision regarding § 230 immunity and arguing that this initial interpretation was too expansive and set dangerous precedent with its generous construction).

148. *See* Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 521 F.3d 1157, 1187 (9th Cir. 2008) (en banc) (McKeown, J., dissenting) (noting that these are the only exemptions explicit in the statute under § 230(e)). There have been courts that refuse to let states prosecute purveyors of child pornography and hold that ISP prosecution is prohibited under § 230. *See, e.g.*, Backpage.com, LLC v. McKenna, 881 F. Supp. 2d 1262, 1275 (W.D. Wash. 2012) (granting a preliminary injunction against enforcement of a state anti-child-pornography law where the plaintiffs speech would be improperly restricted and their argument would likely succeed on the merits due to the supremacy of the federal statute).

immunity. As such, courts continue to presume that immunity applies as new and different technology comes before them, while using a test originally created in 1996 for online bulletin boards.

The Fourth Circuit's interpretation of § 230 encouraged further expansion. Courts have stretched *Zeran*'s broad constructions—the definition of "traditional"-publisher conduct and the scope of protection—to their logical extremes. Types of ISPs considered publishers (as opposed to "content providers") for § 230 purposes include email listservs,[149] dating websites,[150] social networking sites,[151] and gossip sites.[152] Similarly, causes of action barred from adjudication by § 230 include defamation,[153] negligence,[154] Fair Housing Act (FHA) claims,[155] and even state criminal charges.[156] Furthermore, in the spirit of expanding the immunity to its broadest point, post-*Zeran* courts have explicitly narrowed the definition of creation or development of content, resulting in few, if any, actions that fall outside the scope of publisher conduct and into the role of codeveloper of information.[157] As such, the judiciary has contributed to the iron-clad immunity ISPs enjoy today in an analytical and normative manner.

First, even without the addition of courts' narrow construction of development, the *Zeran* construction creates almost impenetrable immunity. The question of whether the cause of action permits the consideration of § 230 immunity is almost always answered in favor of the defendant.[158] Courts grant immunity for claims in which the defendant took on the role of a "traditional" publisher when editing the content, but the definition of "traditional" publisher is not derived from the statute or common law.[159] It

149. *See, e.g.*, Batzel v. Smith, 333 F.3d 1018, 1030–31 (9th Cir. 2003).

150. *See* Carafano v. Metrosplash.com, Inc., 339 F.3d 1119, 1121, 1125 (9th Cir. 2003).

151. *See* Doe v. MySpace, Inc., 474 F. Supp. 2d 843, 846, 849–50 (W.D. Tex. 2007), *aff'd*, 528 F.3d 413 (5th Cir. 2008).

152. *See* Blumenthal v. Drudge, 992 F. Supp. 44, 52–53 (D.D.C. 1998).

153. *See* Jones v. Dirty World Entm't Recordings LLC, 755 F.3d 398, 407 (6th Cir. 2014).

154. Zeran v. Am. Online, Inc., 129 F.3d 327, 330 (4th Cir. 1997).

155. *See, e.g.*, Chi. Lawyers' Comm. for Civil Rights Under Law, Inc. v. Craigslist, Inc., 519 F.3d 666, 670 (7th Cir. 2008).

156. Backpage.com, LLC v. McKenna, 881 F. Supp. 2d 1262, 1275 (W.D. Wash. 2012) (holding that an injunction against enforcement of a criminal statute prosecuting pages that hosted ads depicting child pornography was proper where the immunity provision and First Amendment rights both indicated that the Plaintiff would likely prevail when criminal charges were brought).

157. *See* Recent Case, FTC v. Accusearch, Inc., *No. 06-CV-105, 2007 WL 4356786 (D. Wyo. Sept. 28, 2007)*, 121 HARV. L. REV. 2246, 2246 n.3 (2008) (noting that the courts have utilized three "broad interpretative policy levers" to control immunity: the first two, as noted in *Zeran*, are the broad construction of "internet service providers" and the allowance of any cause of action, and the third is the judicial license to determine what encompasses the definition of creation and development narrowly).

158. *See* Ardia, *supra* note 34, at 477.

159. *See supra* note 133 (noting that "publisher" is a term of art); *see* also *Zeran*, 129 F.3d at 330 (noting that the "exercise of a publisher's traditional editorial functions" include "deciding whether to publish, withdraw, postpone or alter content"); Blumenthal v. Drudge, 992 F. Supp. 44, 49–53 (D.D.C. 1998) (finding that defendant was not a content provider despite its exercising editorial control over a gossip column's content and having a licensing agreement employing the author).

has been crafted by the courts.[160] Ultimately, an expansive interpretation of the claims covered by § 230 (in which harmful third-party content is not the basis of the legal claim), combined with an expansive definition of traditional-publisher activities defined by the court stretches immunity to its broadest point. Courts have even found that immunity applies where the claim brought is unrelated to the specific content posted by a third party.[161]

This expansive interpretation was employed in *Doe v. MySpace, Inc.*,[162] where the court found that a publisher could not be held liable for any action in which a third party played a part, thus separating the doctrine from the realm of content publication altogether.[163] In *Doe*, a mother sued on behalf of her underage daughter who had been sexually assaulted by a man she connected with on MySpace.[164] She asserted that she was not suing about the content on the platform but rather for MySpace's failure to warn that this type of assault could happen as a result of joining the website.[165] The court held that even where the plaintiff alleged harm on grounds other than the posting of content, § 230 "precludes courts from entertaining claims that would place a [website] in a publisher's role."[166] This construction absolves ISPs of liability even where the content the third party posted is not related to the harm alleged. Instead, the ability to allege a failure-to-warn claim—implicating ISPs acting as a business rather than a publisher—is destroyed.[167]

In addition to continuing to apply *Zeran*'s broad interpretation of publisher duties, courts have narrowed the definition of creation or development, which results in a defendant-friendly immunity.[168] Courts' narrow interpretation of

160. *See supra* note 133.

161. Doe v. MySpace, Inc., 474 F. Supp. 2d 843, 847 (W.D. Tex. 2007), *aff'd*, 528 F.3d 413 (5th Cir. 2008).

162. 474 F. Supp. 2d 843 (W.D. Tex. 2007), *aff'd*, 528 F.3d 413 (5th Cir. 2008).

163. *See* Lukmire, *supra* note 56, at 398 (noting that the immunity "even extended . . . to cover MySpace's *own* failure to act" without respect to the content exchanged on the website). *But see* Doe v. Internet Brands, Inc., 824 F.3d 846, 850 (9th Cir. 2016) (holding that immunity did not apply to a website that models used to advertise their freelance services where information on the site was used by two men to lure models to Florida to rape them). In *Internet Brands*, the court allowed the failure to warn claim to proceed because the duty to warn did not derive from the company's actions as a publisher (editing or removing content) but from its failure as a service to warn of attendant threats of the service unrelated to posted content. *Id.* at 851. This opinion continues the trend of using precision when defining publisher conduct as compared to ISP conduct. The court in *Internet Brands*, however, further distinguished *MySpace* by the fact that Internet Brands had actual knowledge of the scheme. *Id.* at 853.

164. *MySpace*, 474 F. Supp. 2d at 848.

165. *Id.*

166. *Id.* at 847 (quoting Dimeo v. Max, 433 F. Supp. 2d 523, 528 (E.D. Pa. 2006)); *see also* Lukmire, *supra* note 56, at 398 (discussing how even though this ruling brought the statute "to its breaking point, the court accused the plaintiff of disingenuous 'artful pleading' for accusing MySpace of negligence").

167. *See* Lukmire, *supra* note 56, at 398.

168. *See, e.g.*, Batzel v. Smith, 333 F.3d 1018, 1031 (9th Cir. 2003) (holding that "development" of content "means something more substantial than merely editing . . . and selecting material"); *see also* Ben Ezra, Weinstein & Co., Inc. v. Am. Online, Inc., 206 F.3d 980, 985–86 (10th Cir. 2000); Blumenthal v. Drudge, 992 F. Supp. 44, 52–53 (D.D.C. 1998); Paul Ehrlich, Note, *Communications Decency Act § 230*, 17 BERKELEY TECH. L.J. 401, 406–11 (2002) (analyzing the expansive approach courts have taken regarding immunity).

development reinforces the broad interpretation of what a "traditional" publisher does: the more broadly the court construes a publisher's actions, the smaller the window becomes for a publisher to develop content. In addition to granting immunity for actions that are "traditionally" those of a publisher,[169] courts have expanded the definition of publisher duties to include employing content providers.[170] This appears to conflate the duties of publishers with general duties of ISPs.

In *Blumenthal v. Drudge*,[171] AOL was sued for defamatory remarks originating with a column it published to its readers.[172] The allegedly defamatory remarks were authored by Matt Drudge, a paid writer of a gossip column for AOL over which AOL had editorial control.[173] The court held that § 230 applied because the story was written by Drudge without substantive edits made by AOL.[174] This case demonstrated that, despite soliciting and paying for the content under the specific pretense of gossip, the ISP could still escape liability. This holding did not address that § 230 was enacted to encourage "Good Samaritan" blocking[175] by shielding from liability those websites that attempted to edit lewd, defamatory, or obscene material.[176] This purpose becomes obstructed where ISPs are not responsible for content that they have solicited via licensing agreements with columnists.

Finally, as apparent in *Blumenthal*, neither the tone nor the purpose of a website influences the analysis of development under § 230 immunity. Even those who encourage gossip,[177] derogatory claims,[178] and facilitate child prostitution[179] are not liable where the courts consider the invitation to criminal or tortious acts within the bounds of the publishers' role. Thus, by broadening the role of the traditional publisher to include a broad range of actions, courts have created a smaller pool of activity that falls outside the role of a traditional publisher. By making that window smaller, courts have all but guaranteed immunity, regardless of ISP conduct.

---

169. *See supra* note 159.
170. *Blumenthal*, 992 F. Supp. at 50.
171. 992 F. Supp. 44 (D.D.C. 1998).
172. *Id.* at 46.
173. *Id.* at 47, 50.
174. *Id.* at 50.
175. 47 U.S.C. § 230(c) (2012).
176. *See* Chi. Lawyers' Comm. for Civil Rights Under Law, Inc. v. Craigslist, Inc., 519 F.3d 666, 670 (7th Cir. 2008) (noting that the title is "hardly an apt description if its principal effect is to induce ISPs to do nothing about the distribution of indecent and offensive materials via their services"); *see also* Part I.A.2.
177. *Blumenthal*, 992 F. Supp. at 51.
178. *See* Jones v. Dirty World Entm't Recordings LLC, 755 F.3d 398, 407–08 (6th Cir. 2014); *see also* Glob. Royalties, Ltd. v. Xcentric Ventures, LLC, 544 F. Supp. 2d 929, 933 (D. Ariz. 2008) (noting that while it was "obvious that a website entitled Ripoff Report encourages the publication of defamatory content," ultimately § 230 provides immunity where Ripoff Report did not create or develop the offending content).
179. *See* Backpage.com, LLC v. McKenna, 881 F. Supp. 2d 1262, 1275 (W.D. Wash. 2012); Doe v. Bates, No. 5:05-CV-91DF-CMC, 2006 WL 3813758, at *5 (E.D. Tex. Dec. 27, 2006).

Other than the three levels of broad construction,[180] the most explicit remnant of *Zeran* is the formal analysis courts apply to determine the applicability of § 230. This test requires the consideration of three questions: (1) whether the defendant qualifies as a provider of an "interactive computer service,"[181] (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."[182] Though most early decisions accepted the *Zeran* construction, dissension was apparent as early as 2004.[183] As the internet began to flourish and the largest companies became technology based, the attitude of "internet exceptionalism" began to shift: the broadest possible construction was no longer the only option considered.[184]

C. Curtailing Broad Immunity: Roommates *Assesses Content at Issue and Redefines "Development"*

Although broad construction of § 230 immunity dominated early opinions, judicial opposition to sweeping immunity was apparent within the first ten years of its application.[185] Judges demonstrated discontent with the position of legal privilege that internet businesses gained over every other business.[186] The fact that companies with an online presence could advance illegal conduct with virtually no threat of ensuing civil litigation grates against certain judges' sense of justice.[187] In a departure from a period of total

180. The applicability of § 230 to any cause of action, the broad construction of the duties of a traditional publisher, and the narrow construction of "development" are the three interpretations that have expanded the applicability of § 230 immunity. *See supra* Part II.A–B.

181. This inquiry requires very little of the court. It merely asks whether the defendant operates a web-based service.

182. James D. Shanahan, *Rethinking the Communications Decency Act: Eliminating Statutory Protections of Discriminatory Housing Advertisements on the Internet*, 60 FED. COMM. L.J. 135, 139 (2007); *see, e.g.*, Doe v. Internet Brands, Inc., 824 F.3d 846, 850 (9th Cir. 2016) ("Separated into its elements, subsection (c)(1) precludes liability for '(1) a provider or user of an interactive computer service (2) whom a plaintiff seeks to treat, under a state law cause of action, as a publisher or speaker (3) of information provided by another information content provider.'" (quoting Barnes v. Yahoo!, Inc., 570 F.3d 1096, 1100–01 (9th Cir. 2009))).

183. *See* Barrett v. Rosenthal, 9 Cal. Rptr. 3d 142, 167 (Ct. App. 2004) (holding that "section 230 does not restrict distributor liability under the common law"), *rev'd*, 146 P.3d 510 (Cal. 2006) (reversing in favor of the *Zeran* analysis in part because of the practical implications of forum shopping).

184. *See* Eric Goldman, *The Third Wave of Internet Exceptionalism*, TECH. & MARKETING L. BLOG (Mar. 11, 2009), http://blog.ericgoldman.org/archives/2009/03/the_third_wave.htm [https://perma.cc/HUR3-CZNJ].

185. *See* Chi. Lawyers' Comm. for Civil Rights Under Law, Inc. v. Craigslist, Inc., 519 F.3d 666, 671 (7th Cir. 2008) (critiquing the drafting of the statute but ultimately holding that the most sensible construction provides immunity for Craigslist from the Fair Housing Act claims); *see also supra* note 183.

186. *See supra* note 185; *see also* Joey Ou, Note, *The Overexpansion of the Communications Decency Act Safe Harbor*, 35 HASTINGS COMM. & ENT. L.J. 455, 470 (2013).

187. *See supra* notes 185–86; *see also* Christopher Zara, *The Most Important Law in Tech Has a Problem*, WIRED (Jan. 3, 2017, 12:00 AM), https://www.wired.com/2017/01/the-most-important-law-in-tech-has-a-problem/ [https://perma.cc/HP63-KGPZ] (noting "a disquieting number of courtroom losses for Section 230" in the past year).

deference to § 230's immunity, some courts have found that, depending on the claim brought, a website's structure may bar it from immunity.[188] But courts struggle to delineate precisely where this disqualification occurs. At what point does an ISP's involvement with the content move it from the role of publisher to the role of "developer" under this more attentive regime?

The first case to put substantive limits on what constitutes the publisher's role and analyze what qualifies as codevelopment of information is *Roommates*.[189] Part 1 describes the holding in that case and then dissects the analytical considerations that characterized the decision. It then discusses the logical difficulties that are presented by the final statement of the rule. Next, Part 2 parses the opinion's analysis of the statute and the facts, while specifically limiting references to "underlying illegality" to avoid the jurisprudential difficulties it creates. Finally, Part 3 addresses the difficulties of considering illegality in § 230 analysis and why an abandonment of the "underlying illegality" standard will enable courts to approach the issue of immunity more fairly.

1. *Roommates*: The Holding

Roommate.com, LLC operated a site, Roommates.com,[190] which allowed users to find prospective roommates and shared living spaces.[191] When users set up profiles, they were required to fill out profiles with prepopulated drop-down menus and selection boxes as well as "additional comments" boxes.[192] Answers to questions regarding the residence location, description, and details, as well as to questions about the renter's sex, sexual orientation, and familial status were required.[193] Other questions regarding preferences, as well as an "additional comments" box, were optional.[194] Users could then filter options and opt in to receive emails through Roommates.com using these criteria.[195]

The Fair Housing Council of San Fernando Valley (FHC) brought suit alleging that Roommates.com violated the FHA. The FHA specifically prohibits housing discrimination on the basis of "race, color, religion, sex, handicap, familial status, or national origin."[196] The FHC alleged that the Roommates.com platform violated the FHA first by posting illegal questions on its website; second, by forcing users to complete them and by using the information provided to send emails with profiles that reflect selected

188. For a discussion of the limitation of immunity, see *infra* Part II.C.2, 4.

189. Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 521 F.3d 1157, 1161 (9th Cir. 2008) (en banc).

190. When referring to the 2008 Ninth Circuit en banc decision, this Note uses "*Roommates*" as a short form. This Note uses "Roommates.com" when referring to the ISP and its services.

191. *Roommates*, 521 F.3d at 1161.

192. *Id.* at 1165.

193. *Id.* at 1161.

194. *Id.* at 1181 (McKeown, J., dissenting).

195. *Id.*

196. 42 U.S.C. § 3604(c) (2012) (making it unlawful to "make, print, or publish" discriminatory statements or advertisements related to the sale or rental of a dwelling).

preferences; and third, by posting the information in additional comments boxes even where the comments were discriminatory.[197] Roommates.com claimed that § 230 shielded it from liability for FHA violations where third parties were the ones using these categories to find housing. The question before the court was directed to the third prong of the *Zeran* analysis: whether Roommates.com was "responsible, in whole or in part, for the creation or development of [the] information."[198] The court held:

> By requiring subscribers to provide the information as a condition of accessing its service, and by providing a limited set of pre-populated answers, Roommate becomes much more than a passive transmitter of information provided by others; it becomes the developer, at least in part, of that information. And section 230 provides immunity only if the interactive computer service does not "creat[e] or develop[]" the information "in whole or in part."[199]

Under this ruling, Roommates.com lost immunity for the questions it posed, for the resulting profiles and emails it assembled using the information gathered, and for the search engine it provided that limited results based on that information.[200] However, Roommates.com did retain immunity for the "additional comments" boxes.[201]

The holding above, which stripped Roommates.com of most immunity, was subsequently clouded by other parts of the opinion. The court stated whether an ISP had "developed" content is contingent on whether the *unlawful* conduct is being furthered by the site's actions or design.[202] Under this definition, a court would be required to make a judgment on the merits of the claim before deciding whether immunity exists to bar liability.[203]

This created a legal regime in which unlawfulness is relitigated over multiple proceedings. For example, in 2008, the Ninth Circuit stripped Roommates.com of immunity because its actions contributed to the underlying FHA violation. In 2012, however, when the case reappeared in the Ninth Circuit on the merits of the FHA claim, the court determined that

197. Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 489 F.3d 921, 926 (9th Cir. 2007), *modified on reh'g en banc*, 521 F.3d 1157 (9th Cir. 2008).

198. *Id.* (alteration in original).

199. *Roommates*, 521 F.3d at 1166 (alterations in original) (quoting 47 U.S.C. § 230(f)(3)).

200. *Id.*

201. *See id.* at 1173–74 (holding that because Roommates.com was not responsible "for the development" of content within the box because it did "not provide any specific guidance as to what the essay should contain, nor [did] it urge subscribers to input discriminatory preferences"). This holding mimicked the Seventh Circuit's holding in *Craigslist*. *See* Chicago Lawyers' Commission for Civil Rights Under Law, Inc. v. Craigslist, Inc., 519 F.3d 666 (7th Cir. 2008).

202. *See generally* Jeffrey R. Doty, Note, *Inducement or Solicitation? Competing Interpretations of the "Underlying Illegality" Test in the Wake of* Roommates.com, 6 WASH. J.L. TECH. & ARTS 125 (2010); *see also Roommates*, 521 F.3d at 1168.

203. *Roommates*, 521 F.3d at 1182 (McKeown, J., dissenting); *see also* Eric Goldman, *Roommates.com Isn't Dealing in Illegal Content, Even Though the Ninth Circuit Denied Section 230 Immunity Because It Was*, TECH. & MARKETING L. BLOG (Feb. 6, 2012), http://blog.ericgoldman.org/archives/2012/02/roommatescom_is.htm [https://perma.cc/AES6-FEM8].

Roommates.com's actions did not violate the FHA.[204] Under a test that considers underlying illegality, immunity should have been granted in 2008 because in 2012, when the issue was litigated, the court found no underlying illegality.[205]

2. *Roommates* Analysis: Assessing "Development," and the ISP Collection and Use of User-Generated Data

Roommates served as a useful catalyst for other courts to reexamine the third prong of the *Zeran* analysis—the question of creation or codevelopment of content. While the first issue that *Roommates* discusses is whether the claim treats the defendant as a publisher, the court promptly qualifies that the question is only dispositive where the defendant is not also found to be the creator or developer of the content at issue.[206] As such, the *Roommates* analysis hinges directly on the third prong of the *Zeran* analysis.

In its analysis of the third prong, the Ninth Circuit first identified the specific content at issue for each cause of action and discerned the relationship of that content to the ISP's conduct. This analysis created a novel bright line between conduct as "creation" and conduct as "development."[207] Importantly, if the content at issue is a direct result of the conduct of the ISP, then the content is considered "created" by the ISP and immunity does not apply. If the ISP is alleged to have "developed" the information, the court must consider whether the ISP's control in soliciting and utilizing user data constituted development.[208]

a. ISP as Creator of Content

The Ninth Circuit began its immunity analysis with the simplest issue: whether immunity applies to the questions and prepopulated answers created by Roommates.com where that content allegedly indicates intent to discriminate in violation of the FHA.[209] The court reasoned that since Roommates.com created the questions and the selectable answers and since these questions were alleged to be a direct violation of the FHA (choosing housing based on designated attributes), Roommates.com created the content at issue.[210] Section 230 immunity was denied as it could not protect content that Roommates.com had actually created. While this reading comports with the statute, in order to hold that immunity did not apply, the court had to narrow an earlier Ninth Circuit ruling on § 230 immunity.

---

204. Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 666 F.3d 1216, 1223 (9th Cir. 2012).

205. *Id.*; *see also* Goldman, *supra* note 203.

206. *Roommates*, 521 F.3d at 1162 (majority opinion).

207. *Id.* at 1167–69, 1172.

208. *Id.* at 1172.

209. *See id.* at 1164 ("Councils allege that requiring subscribers to disclose their sex, family status and sexual orientation 'indicates' an intent to discriminate against them, and thus runs afoul of both the FHA and state law.").

210. *Id.*

That earlier decision, *Carafano v. Metrosplash.com, Inc.*,[211] involved a fake dating profile that was created using the images of the plaintiff.[212] The profile also contained sexually suggestive language and made her personal contact information and address available.[213] The disclosure of this type of personal information was not required for the service.[214] Carafano sued for invasion of privacy, misappropriation of the right of publicity, defamation, and negligence.[215] The court had held that the ISP was not liable under the CDA because "no profile has any content until a user actively creates it."[216]

The court in *Roommates* expressly narrowed this holding, which otherwise would have included the Roommates.com profiles.[217] The court noted that even though both cases involved questions and prepopulated answers, the harm in *Carafano* had not derived from the questions; as such, the questions were not the content at issue in that case. Instead, the harm was entirely the product of the third party's posting of private contact information, leading to privacy invasions, and the posting of false answers, leading to defamation.[218] Nothing on the Metrosplash platform encouraged or solicited personal information from individuals or information that could defame, because the user was supposedly creating his or her own profile.[219] The claim against Metrosplash was that it failed to review every profile for potentially defamatory content or personal information.[220] This claim had nothing to do with the actual profile created. The content at issue was defamatory material. Nothing alleged related to how the development of a profile would lead to the defamation or how the questions solicited an invasion of privacy. By contrast, Roommates.com asked and provided limited answers for the questions that formed the basis of the intent to discriminate required to state a claim under the FHA.

b. ISP as Codeveloper

In *Roommates*, the FHC alleged that Roommates.com developed content that violated the FHA: specifically, the completed profiles and emails that

211. 339 F.3d 1119 (9th Cir. 2003).
212. *Id.* at 1121.
213. *Id.*
214. *Id.*
215. *Id.* at 1122.
216. *Id.* at 1124.
217. *See* Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 521 F.3d 1157, 1171 (9th Cir. 2008) (en banc) (calling the holding in *Carafano* "unduly broad").
218. In this analysis, the court also stressed that the conduct contributed to the illegality of the content, whereas in *Carafano*, it did not. However, the questions posed by Roommates.com did not violate the FHA. *See infra* Part II.C.3.
219. Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 489 F.3d 921, 928 (9th Cir. 2007), *modified on reh'g en banc*, 521 F.3d 1157 (9th Cir. 2008) ("The website sought information about the individual posting the information, not about unwitting third parties. Nothing in the questions the dating service asked suggested, encouraged or solicited posting the profile of another person . . . .").
220. *Roommates*, 521 F.3d at 1171–72 ("With respect to the defamatory content, the website operator was merely a passive conduit and thus could not be held liable for failing to detect and remove it.").

enabled users to search for roommates using sensitive characteristics.[221] To assess this claim, the Ninth Circuit redefined "development" and used its definition to analyze the relationship between the ISP and the user-generated content.[222] Ultimately, the court found that the restricted options available to users and the importance of the collected data to the profitability of the ISP were two factors that differentiated the survey-enabled services (search and email functions) from the additional comment box services that retained immunity.[223]

The Ninth Circuit faced an uphill battle to redefine ISP conduct as "development" given the acceptance of most ISP conduct as "traditional publisher" conduct.[224] However, using close textual analysis, the court demonstrated that there was necessarily room for its interpretation of the word "development."[225] This allowed the court to move away from the proimmunity "traditional publisher" norm.[226] While it initially conceded that the broadest definition of development would eviscerate immunity by encompassing almost all functions of websites and search engines, it demonstrated that the dissent's narrow definition effectively ignored the word "development" and only permitted immunity to be withheld from content-creator ISPs.[227] The court ultimately likened development to the solicitation and use of information.[228]

The court established that an ISP could be a developer of content and, as such, analyzed Roommates.com's solicitation and use of content despite the fact that the content was user generated.[229] Under the solicitation analysis, the court found that where the type of information collected was dictated so strictly by the manner in which it was collected, it was deemed to be codeveloped by the ISP.[230] This occurred where the information was both required and limited to prepopulated options. The required disclosure of information and the limitation of survey answers are the most easily discernable differences between the survey-based services, which lost immunity, and the "additional comments" box, which did not.[231] Because Roommates.com required answers from prepopulated selections as conditions of using the service, Roommates.com was deemed a

221. *Id.* at 1167.
222. *Id.* at 1166–68.
223. *Id.* at 1172.
224. *See supra* Part II.B.
225. *Roommates*, 521 F.3d at 1167.
226. For a discussion of the evolution of § 230 doctrine, see *supra* Part II.A–B.
227. *Roommates*, 521 F.3d at 1167 ("[R]eading the exception for co-developers as applying only to content that originates entirely with the website—as the dissent would seem to suggest—ignores the words 'development . . . in part' in the statutory passage 'creation *or development* in whole *or in part*.' We believe that both the immunity for passive conduits and the exception for co-developers must be given their proper scope . . . ." (citation omitted) (quoting 47 U.S.C. § 230(f)(3) (2012))).
228. *Id.* at 1166 (using the terms "development" and "solicitation" synonymously by saying "solicit (a.k.a. 'develop')").
229. *Id.* at 1169.
230. *Id.* at 1172.
231. *See supra* note 201 (describing the upholding of immunity for the "additional comments" box).

codeveloper.[232] The ISP's limitation of "user control" discussed in the statute[233] could be a reason for the court's hostility toward the format that Roommates.com presented to users.[234]

The specific organization of the information, and the inability for users to control that information as a contingency of using the service, caused the court to find that Roommates.com did not qualify for immunity.[235] Some commenters have proposed guidelines to assess whether restriction goes so far as to limit the user to the role of codeveloper of the information.[236] Such guidelines consider the number of choices available to the user and whether the user has the ability to abstain from the collaboration altogether.[237]

When analyzing the ISP's use of the user-generated content, the court also considered the extent to which the collection of the required content enabled the ISP to meet business objectives.[238] The court undertook this analysis when assessing the survey-based services offered. It noted that Roommates.com's email- and platform-filtration capabilities were enabled by the manner in which the information was collected.[239] This filtration also furthered the alleged conduct by allowing the potential for discriminatory conduct.[240] It distinguishes "merely" publishing the profiles to other users from Roommates.com's channeling of profiles based on users' preferences and the exclusion of profiles that do not fit those preferences.[241] The court

232. *Roommates*, 521 F.3d at 1167 ("If Roommate has no immunity for asking the discriminatory questions, as we concluded above, . . . it can certainly have no immunity for using the answers to the unlawful questions to limit who has access to housing." (citation omitted)).

233. 47 U.S.C. § 230(b)(3) (2012) (stating that it is the policy of the United States to "encourage the development of technologies which *maximize user control* over what information is received by individuals, families, and schools who use the Internet and other interactive computer services." (emphasis added)).

234. Hattie Harman, Note, *Drop-Down Lists and the Communications Decency Act: A Creation Conundrum*, 43 IND. L. REV. 143, 163–64 (2009) (noting that where there is free will, one can encourage a result without incurring liability, but where someone puts another in danger, and where there is "no real choice," a duty to rescue is created); *see also* Carafano v. Metrosplash.com, Inc., 339 F.3d 1119, 1124 (9th Cir. 2003) (stating that where the third party "willingly provide[d] the essential published content," the service provider retains immunity).

235. *Roommates*, 521 F.3d at 1164 ("Roommate created the questions and choice of answers, and designed its website registration process around them. Therefore, Roommate is undoubtedly the 'information content provider' as to the questions and can claim no immunity for posting them on its website, or for forcing subscribers to answer them as a condition of using its services.").

236. *See* Harman, *supra* note 234, at 171–74 (providing a fuller discussion of the proposed indicia).

237. *Id.*

238. *See supra* note 107. Some scholars suggest that assessing the commercial value of the content collected is also important to this analysis as it drives the objectives of the company and its use of data. *See* Harman, *supra* note 234, at 172–73 (2009) (noting that, when assessing the level of development, the court can look to whether there was an incentive for the particular information collected because "it is logical to conclude the site has a stake in the outcome of the user's selection and therefore encourages it").

239. *Roommates*, 521 F.3d at 1169.

240. *Id.*

241. *Id.* at 1167 ("Roommate is not entitled to CDA immunity for the operation of its search system, which filters listings, or of its email notification system, which directs emails to subscribers according to discriminatory criteria.").

noted Roommates.com's decision to design "its search system so it would steer users based on the preferences and personal characteristics that Roommate itself forces subscribers to disclose" barred it from immunity.[242]

The court differentiates between "neutral" search tools, which can receive immunity, and nonneutral tools, such as those utilized by Roommates.com.[243] Because Roommates.com's survey-based services allowed the platform to limit available options for certain users and override "user-defined" filter options, the platform did not qualify as a neutral tool.[244] A nonneutral tool is one that aligns results of a search to limit listings available based on preferences that the program knows about the user.[245] In this instance, required information from a user determined the basis of the information that Roommates.com would show to a user, thus affecting the very discrimination against roommates that the FHA sought to eliminate. This limitation, based on the system's knowledge of the user, destroyed immunity for Roommates.com as it was seen as interference with content that went beyond that of a publisher. In contrast, search tools that have no underlying information about the user have no potential to silently effect the information a user receives from a publisher.

3. The Underlying Illegality Test for Assessing "Development"

In response to the question whether Roommates.com could lose immunity due to the answers Roommates.com solicited in its survey, the court "interpret[ed] the term 'development' as referring not merely to augmenting the content generally, but to materially contributing to its alleged unlawfulness."[246] This language about "unlawfulness" caused the most apparent rift with the dissent and created a difficult standard to apply.[247] Differentiating the augmentation of content generally from the augmentation of content that contributes to illegality is not useful where general content development will not require immunity due to the absence of underlying illegality.[248]

The dissent argues that assessing illegality tarnishes the court's impartiality by addressing the underlying claim in the same breath as the

242. *Id.*; *see also* Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 489 F.3d 921, 929 (9th Cir. 2007) ("By categorizing, channeling and limiting the distribution of users' profiles, Roommate provides an additional layer of information that it is 'responsible' at least 'in part' for creating or developing."), *modified on reh'g en banc*, 521 F.3d 1157 (9th Cir. 2008).

243. *Roommates*, 521 F.3d at 1169.

244. *See id.*

245. *See id.*

246. *Id.* at 1167–68 (clarifying the point by adding that "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").

247. *See id.* at 1166 ("Unlawful questions solicit (a.k.a. 'develop') unlawful answers.").

248. *See id.* at 1182–83 (McKeown, J., dissenting) ("Immunity has meaning only when there is something to be immune *from*, whether a disease or the violation of a law. It would be nonsense to claim to be immune only from the innocuous.").

available immunity.[249] It argues that this interpretation strains the textual bounds of the statute.[250] Finally, the ultimate result of this case demonstrates the awkward legal ramifications of characterization of actions as "illegal" before the claim is adjudicated. This Part addresses the two main concerns raised by the dissent and the uncomfortable outcomes that arise out of this standard.

The primary concern is the bias introduced by discussion of substantive liability in relation to the immunity analysis. The dissent argues the issue of immunity should not be biased by a judge's cursory review of the pleadings, as the two issues are analytically distinct.[251] Judge M. Margaret McKeown states that the holding is flawed as it considers the analysis "built on substantive liability."[252] She states that there must be procedural respect given to adjudication of the merits of the claims "[i]nstead of foreshadowing a ruling on the FHA."[253] The dissent's concerns are somewhat validated as the majority occasionally abandons the modifier "alleged" when referring to the "discriminatory" actions of Roommates.com.[254]

The majority addresses the dissent in a footnote, and claims that the definition "does not depend on finding substantive liability" but "merely requires analyzing the context in which a claim is brought."[255] The court agrees that "finding that a defendant is not immune is quite distinct from finding liability" but continues in the body of the decision to note Roommates.com's discriminatory conduct.[256] In this footnote, it states its test without the use of the "underlying illegality" claim, saying: "the CDA provides no immunity to Roommate's actions in soliciting and developing the content of its website."[257]

The second critique of this formulation of the word development is that it goes beyond the meaning of the statute. The dissent notes that there is no mention of "unlawfulness" in the statute.[258] The statute contemplates culpability when it provides for "Good Samaritans," and yet never writes a good- or bad-faith requirement into the statute. Similarly, the dissent notes that there is no definition of the word "development" that connotes any type of negativity or unlawfulness.[259]

---

249. *Id.* at 1182 ("The majority's definition of 'development' epitomizes its consistent collapse of substantive liability with the issue of immunity.").

250. *Id.*

251. *Id.* ("Whether Roommate is entitled to immunity for publishing and sorting profiles is wholly distinct from whether Roommate may be liable for violations of the FHA.").

252. *Id.* at 1182–83 ("[T]he majority's immunity analysis is built on substantive liability: to the majority, CDA immunity depends on whether a webhost materially contributed to the unlawfulness of the information.").

253. *Id.* at 1178.

254. *Id.* at 1172 (majority opinion).

255. *Id.* at 1171 n.30.

256. *Id.*

257. *Id.*

258. *See id.* at 1182 (McKeown, J., dissenting) ("Where in the statute does Congress say anything about unlawfulness?").

259. *Id.* (noting that "[t]his definition . . . springs forth untethered to anything in the statute").

Finally, outcomes produced by the underlying illegality language are potentially contradictory. As noted earlier, in 2012, when *Roommates* returned to the Ninth Circuit on the issue of the FHA violation, the court ruled that there was no FHA violation.[260] The 2012 ruling expressly indicated that the First Amendment protected this type of roommate selection within a dwelling.[261] The court, however, did not acknowledge this when it discussed the 2008 ruling in the 2012 decision.[262]

Scholars who have considered this decision have characterized this language as imprecise and unhelpful.[263] So, while *Roommates* was pivotal in giving new meaning to the term "development" and providing tools for future courts to assess the ISP-content relationship, it created problems as well.

D. The Aftereffects of Roommates

Roommates had immediate and lasting reverberations. A spike in CDA legal scholarship followed the controversial opinion,[264] as did many cases approaching the statute in a more nuanced manner.[265] This section discusses how courts approached § 230 cases after *Roommates*. Specifically, it breaks down the two key shifts that evolved out of the decision in *Roommates*: the redirection of the definition of "development" and the scrutiny of what publisher conduct includes.

1. Culpability Refocused: Conduct as a Means of Doing Business

Outside of the Ninth Circuit, other circuit courts began to wrestle with the meaning of "development" more intently. When unpacking *Roommates*'s underlying illegality language in *FTC v. Accusearch*,[266] the Tenth Circuit analyzed the plain meaning of the words "develop" and "responsible" in the statute.[267] The court found that to develop something was to make the content "usable," "visible," or "active."[268] It held that an ISP is "responsible"

260. *See generally* Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 666 F.3d 1216 (9th Cir. 2012).

261. *Id.*

262. *Id.* at 1219.

263. *See* Doty, *supra* note 202, at 131 ("As might be expected, decisions following *Roommates.com* have not applied the underlying illegality test consistently. Instead, the case law seems to reflect two different approaches to defining culpable behavior: one based on 'solicitation' and the other on 'inducement.'").

264. JOEL R. REIDENBERG ET AL., FORDHAM CTR. ON LAW & INFO. POLICY, SECTION 230 OF THE COMMUNICATIONS DECENCY ACT: A SURVEY OF THE LEGAL LITERATURE AND REFORM PROPOSALS 9 (2012), http://www.fordham.edu/download/downloads/id/1825/clip_section_230_of_the_communications_decency_act_report.pdf [https://perma.cc/N4NP-N73K] (noting that in the late 2000's there was a spike of literature that rose up in defense of broad immunity and that since 2007—the year the first *Roommates* decision came out, before the en banc hearing—there has been a renewed stream of scholarly articles).

265. *See, e.g.*, FTC v. Accusearch, Inc., 570 F.3d 1187, 1198–99 (10th Cir. 2009) (spending paragraphs discussing the meaning of both "development" and "responsible" in the statute).

266. 570 F.3d 1187 (10th Cir. 2009).

267. *Id.* at 1197.

268. *Id.* at 1198.

where it is "more than a neutral conduit for that content" and has "specifically encourage[d] development of what is offensive about the content."[269] This standard hinges on the ISP's awareness of its conduct and moves away from the awareness of specific illegal content. The court evoked a type of moral accountability for the development of the offending content, evocative of proximate causation, but it did not implicate a criminal "illegality"-type standard.[270] In so doing, the Tenth Circuit created a sort of "culpability-toward-development" standard, rather than a "culpability-toward-illegality" standard.

The assessment of the culpability-toward-development standard for ISPs was also taken up in *NPS LLC v. StubHub, Inc*.[271] In this case, StubHub lost immunity because, while not intending to aid illegal scalping, it enabled it by providing technology that allowed users to mask the location of the seats they were selling.[272] This masking technology prevented sports franchises from being able to pinpoint the ticket holders who were selling their seats.[273] StubHub was found to have developed the illegal information not because their conduct of masking seats was illegal but because it allowed the information to be processed in a manner that enabled the content at issue (tickets being sold in violation of antiscalping laws) to be sold.[274] This material contribution to the content at issue was enough to convince the court that StubHub relinquished immunity.[275]

More recently, the holdings in *Roommates* and *Accusearch* were recast by the Seventh Circuit in *Huon v. Denton*.[276] The explanatory parentheticals were devoid of any consideration of the underlying claim as unlawful, illegal, or immoral. Rather, the standards evoked when citing those cases speak to the requiring or solicitation of potentially neutral information in the conduct of business.[277] While some state courts still invoke the illegality standard,

269. *Id.* at 1199.

270. *See id.* at 1198 ("In this context—responsibility for harm—the word *responsible* ordinarily has a normative connotation. . . . [O]ne definition of *responsible* [i]s '[m]orally accountable for one's actions.'" (first, second, and third alternations in original) (quoting *Responsible*, OXFORD ENGLISH DICTIONARY 742 (2d ed. 1998))).

271. No. 06-4874-BLS1, 2009 WL 995483 (Mass. Super. Ct. Jan. 26, 2009).

272. *Id.* at *13.

273. *Id.* at *8.

274. *Id.* at *12.

275. *Id.*

276. 841 F.3d 733 (7th Cir. 2016).

277. *See id.* at 742. The decision characterizes *Roommates* as "concluding that a website was not a 'passive transmitter of information provided by others' but instead helped develop the information by 'requiring subscribers to provide the information as a condition of accessing its service, and by providing a limited set of pre-populated answers.'" *Id.* (quoting Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 521 F.3d 1157, 1166–67 (9th Cir. 2008) (en banc)). The decision characterizes *Accusearch* as "concluding that a website developed the information by 'solicit[ing] requests' for the information and then 'pa[ying] researchers to obtain it.'" *Id.* (quoting FTC v. Accusearch, Inc., 570 F.3d 1187, 1199–1200 (10th Cir. 2009)).

the standard is tempered by language that considers the design of the website and demands made of the user.[278]

2. ISP Conduct in Relation to Content at Issue: What Would a Traditional Publisher Do?

The first post-*Roommates* decision was also from the Ninth Circuit and dealt primarily with distinguishing the duties of publishers from those of ISPs in § 230 cases. Months after *Roommates* was decided, the Ninth Circuit took the opportunity to engage with the second prong of the *Zeran* analysis—whether the cause of action treats the ISP as a publisher. In *Barnes v. Yahoo!, Inc.*,[279] the plaintiff brought a traditional defamation claim regarding third-party posted commentary but also brought a claim for promissory estoppel.[280] While the first claim was barred, the second claim was not. This was because the second claim did not have to do with the content directly, but rather the fact that Yahoo! had breached a "contract" by promising to remove the content and failing to do so in a reasonable amount of time.[281] This court distinguished the claims by assessing the ISP's duties toward those parties for each claim. Where the ISP, independent of publication, had made a legally enforceable agreement, the ISP would not be excused from that agreement merely because it related tangentially to the duties of publication.[282]

The concurring opinion in *Accusearch* also highlighted some distinctions about the responsibility for the conduct toward the content at issue rather than the responsibility for the content itself.[283] While this analysis was not part of the holding, it demonstrates a departure from the blanket immunity granted by earlier courts and a willingness to use the window *Roommates* and *Barnes* opened.[284] This line drawing of ISPs' actions as distinct from those of a publisher continues to appear in recent opinions in the Second Circuit and the

278. *See, e.g.*, People v. Bollaert, 203 Cal. Rptr. 3d 814, 833 (Ct. App. 2016) (noting that the website was "designed to force" the divulgence of information in order to participate).

279. 570 F.3d 1096 (9th Cir. 2009).

280. *Id.* at 1107 ("[Contract law] generates a legal duty distinct from the conduct at hand, be it the conduct of a publisher, of a doctor, or of an overzealous uncle."); *see also* Julia M. MacAllister, Note, *The Doxing Dilemma: Seeking a Remedy for the Malicious Publication of Personal Information*, 85 FORDHAM L. REV. 2451, 2458 (2017).

281. *See id.*

282. *See id.* at 1101–02 ("[C]ourts must ask whether the duty that the plaintiff alleges the defendant violated derives from the defendant's status or conduct as a 'publisher or speaker.' If it does, section 230(c)(1) precludes liability.").

283. *See* FTC v. Accusearch, 570 F.3d 1187, 1204–05 (10th Cir. 2009) (Tymkovich, J., concurring) ("[T]he FTC sought and ultimately held Accusearch liable for its *conduct* rather than for the *content* of the information it was offering on the Abika.com website. Section 230 only immunizes publishers or speakers for the *content* of the information from other providers that they make public.").

284. It was also the main reason the Judge William Downes provided in the district court decision affirmed by the Tenth Circuit. *See* FTC v. Accusearch, Inc., No. 06-CV-105-D, 2007 WL 4356786, at *6–7 (D. Wyo. Sept. 28, 2007), *aff'd*, 570 F.3d 1187 (10th Cir. 2009) (discussing whether the conduct at issue sought to treat the ISP as a publisher and holding that regardless of that fact, the solicitation and resale of third party information fell within the bounds of creation or development).

Ninth Circuit.[285] Courts have crafted effective congressional intent arguments to support narrowing of immunity by asserting that holding ISPs liable for their nonpublisher actions does "not discourage the core policy of section 230(c), 'Good Samaritan' filtering of third party content."[286]

In the most robust example of pushback against blanket immunity, a court held that even where the act of publication is involved, if the user's content is assembled to create a new impression on the viewer and is mobilized for profit, the conduct of publication becomes "development." This case, *Fraley v. Facebook, Inc.*,[287] established that a sponsored post, which consisted of a user's image and the fact that they had "liked" a product on Facebook, constituted development because the ad generated by third parties changed the meaning of the user content.[288] Facebook argued that the "like" and the profile picture were both content provided by users and so were not developed by the ISP; but the impression that the action and the image ultimately created allowed Facebook to monetize them in a way that constituted a new use.[289] This is significant because, even though the action of publication was involved, the fact that Facebook—through its automated assemblage of the information—created a piece of content that was regarded as "new" demonstrated the fullest expression of the pendulum's swing in the other direction. The new line for actions that differentiate publishers from content providers can account for the effect that the assembled content has on the viewer.

III. A NEW FRONTIER: SECTION 230 DOCTRINE REENVISIONED FOR THE CHALLENGES OF MACHINE-LEARNING ALGORITHMS ON SOCIAL MEDIA

This Part proposes a revised framework to analyze immunity, applies that framework to a case that implicates Facebook's algorithmically based services, and addresses the concerns associated with the outcomes. Part III.A outlines a reenvisioned test for § 230 immunity that incorporates the elements recently deployed by circuit courts as well as elements suggested by scholars. This test is then applied in Part III.B to a case recently decided in district

285. *See* FTC v. Leadclick Media, LLC, 838 F.3d 158, 176–77 (2d Cir. 2016) (withholding immunity and holding Leadclick "accountable for its *own* deceptive acts or practices—for directly participating . . . by providing edits to affiliate webpages, for purchasing media" and because the liability did not "derive[] from its status as a publisher or speaker"); Doe v. Internet Brands, Inc., 824 F.3d 846, 852 (9th Cir. 2016) (holding that there was no immunity where the content was not at issue, but rather the ISP's liability derived from its conduct in failing to warn).

286. *Internet Brands*, 824 F.3d at 852.

287. 830 F. Supp. 2d 785 (N.D. Cal. 2011). The court noted that Facebook "ignores the nature of Plaintiffs' allegations, which accuse Defendant not of publishing tortious content, but rather of creating and developing commercial content that violates their statutory right of publicity" and held that immunity could not apply as a result of that development. *Id.* at 801. The court was not persuaded that this statutory claim even treated Facebook as a publisher because the grouping of content "transformed the character of Plaintiffs' words, photographs, and actions into a commercial endorsement to which they did not consent." *Id.* at 802.

288. *Id.* at 792.

289. *See id.* at 799 (discussing the profit Facebook earned from selling sponsored posts).

court, which alleges that Facebook committed both tortious and statutory violations by aiding a foreign terrorist organization in the commission of terrorist attacks. The framework addresses the structural ambiguities that machine-learning algorithms may present to the court and suggests a solution. Finally, Part III.C discusses the rationale for this analysis and addresses concerns that may arise from its application.

A. The Zeran *Framework 2.0*

The *Zeran* framework has undergone major analytical augmentation[290] that has not been formally recognized in judicial opinions. While the first prong has become an ancillary question, the second and third prongs have evolved beyond the simple inquiries that the Fourth Circuit initially considered. As such, this Note proposes an updated framework to address the complexity that algorithmically based services pose. While the first and second questions of the *Zeran* framework require little, if any, formal change, the third prong requires a more thorough approach to "development" that includes consideration of the ISP-user and ISP-content relationships.[291]

The first question of the proposed revision remains the same: whether the defendant-ISP qualifies as an ISP under the statute. Social media networks, as providers of interactive services, will win on this issue of immunity. The second question—whether the claims treat the ISP as a publisher or speaker of the content at issue—requires courts to consider the distinct conduct alleged by each claim. Where the cause of action springs from publisher conduct (harm through editing, disseminating, or deleting), the claim is barred. But, where the claim does not, it must be considered outside the realm of immunity.[292]

The third question requires the most in-depth reformation. When considering whether the content was wholly provided by another content provider, the court must first determine what content forms the root of the action brought. Where that content is created or developed by the ISP, immunity cannot apply. Determination of creatorship is relatively simple, as demonstrated in *Roommates*,[293] but determination of development requires further inquiry.

To determine whether the content at issue is codeveloped by the ISP, the court must inquire into the ISP-user relationship and the ISP-content relationship. The consideration of development fashioned in this way is devoid of any hint of "illegality," and thus avoids the pitfalls of such an analysis.[294] Consideration of the ISP-user relationship and the ISP-content relationship requires the complaint to identify the offending content and allows the court to assess the claim with more precision while sidelining the question of "illegality."

---

290. *See supra* Part II.C–D.

291. ISP-to-user and ISP-to-content relationships will hereinafter be shortened to ISP-user and ISP-content relationships.

292. *See* Barnes v. Yahoo!, Inc., 570 F.3d 1096, 1101–02 (9th Cir. 2009).

293. *See supra* Part II.C.2.a.

294. *See supra* Part II.C.3.

The ISP-user relationship assesses the amount of user control relative to the ISP control exercised over the content at issue. ISP control is evaluated by assessing whether use of the service is contingent upon obtaining specific information and what degree of choice a user has when submitting that content.[295] The ISP-content relationship assesses an ISP's relationship to the gathered content—specifically to its organization and use.[296] Analysis of this relationship looks to whether the ISP has a vested interest in the ingestion of the content and whether content organization is specifically designed to aid the business objectives of the ISP.[297] When combined, where content is required as a contingency of service and the ISP requires that the user divulge the information within specific parameters required by a business, there is most likely "development" of content to the extent that the ISP had sufficient control in the input to alter the end content. This combination is evocative of the culpability-toward-development standard used in *Accusearch*.[298] This standard ultimately views development of content as purposeful, but not necessarily malicious or illegal. Where the contribution to such content makes the content more effective, the publisher can be seen as reaching beyond publication and into the realm of codevelopment.[299]

B. The Application of the Framework to Cohen

There have been two cases consolidated for consideration in the Eastern District of New York that hinge on the issue of § 230 immunity. The first seeks to implicate Facebook for negligently aiding terrorist connections, meetings, and event attendance through its ability to curate engaging content for each person's predicted interests.[300] The second claims a statutory violation of the material-support statute for aiding and abetting a terrorist organization.[301]

The factual background of both cases is similar. Both actions are brought against Facebook because of its alleged role in the incitement of terrorist attacks carried out by Hamas, a group formally recognized as an FTO by the United States.[302] The plaintiffs were either injured in the attacks or are family members of the deceased. The attacks range from bombings to stabbings that have occurred in Israel.[303] Hamas has formally taken responsibility for some of the attacks.[304]

In *Cohen v. Facebook, Inc.*[305]—which is also the name of the consolidated case—the *Cohen* complaint alleges that Facebook acted negligently under

295. *See supra* Part II.C.2.
296. *See supra* Part II.C.2.
297. *See supra* Part II.C.2.
298. *See supra* Part II.C.2.
299. *See, e.g.*, Fraley v. Facebook, Inc., 830 F. Supp. 2d 785, 801 (N.D. Cal. 2011).
300. *See generally Cohen* Complaint, *supra* note 9.
301. *See generally Force* Complaint, *supra* note 32.
302. 8 U.S.C. § 1189(a) (2012).
303. *See, e.g.*, *Cohen* Complaint, *supra* note 9; *Force* Complaint, *supra* note 32.
304. *See Force* Complaint, *supra* note 32, at 26.
305. Nos. 16-CV-4453 (NGG) (LB), 16-CV-5158 (NGG) (LB), 2017 WL 2192621 (E.D.N.Y. May 18, 2017).

Israeli statutory tort law and as a result alleges a breach of statutory duty claim and vicarious liability claim under Israeli law.[306] The complaint also alleges prima facie tort, intentional infliction of emotional distress, and it seeks declaratory judgment for civil conspiracy.[307] It states that § 230 should not apply.[308] The *Force* complaint alleges that Facebook "knowingly provided material support and resources to Hamas" in contravention of the material-support statute.[309]

Each claim's success depends on the content at issue. Some of the content at issue is the actual profiles, groups, and events that terrorists have created.[310] Claims deriving from this content will likely be barred by § 230 where Facebook merely passively allows the content's existence. A claim that requires Facebook to remove content can be dispatched under step two of *Zeran*.[311] But, where the complaint alleges that the conduct was outside the realm of publisher or that the content was codeveloped by the company, Facebook could lose immunity. This strategy is outlined in the framework below.

Under the revised framework, there are two separate points at which Facebook could lose immunity. The first appears under step two of *Zeran*: the argument to be considered is that Facebook's conduct as a social networking service, and not publisher, has aided the effectiveness of the content-at-issue. As such, Facebook could lose immunity where the claim doesn't treat it as publisher. The second arises under the codevelopment exception for publisher immunity. The argument here is that Facebook intentionally creates an engaging and influential space in the "race to the bottom of the brain stem"[312] that not only amplifies users' messages but changes their potency. This argument shows that the medium is a new message where the augmentation of user-generated content's effectiveness actually encourages both the duplication of violent messages and the enactment of violent ends.

1. Facebook: A Traditional Publisher?

Under step two of *Zeran*, where Facebook merely allows for content to be created by third parties online, it is entitled to immunity. But to what extent do Facebook's algorithmically driven services—such as the exposure of users with common interests to each other or exposure to events or groups with nefarious motivations—go beyond those of a traditional publisher?

306. *Cohen* Complaint, *supra* note 9. To avoid the discussion of conflict of laws, for the purposes of this Note § 230 could apply to all claims.

307. *Cohen* Complaint, *supra* note 9, at 20–33.

308. *Id.* at 30–32.

309. *Force* Complaint, *supra* note 32, at 2–3.

310. *See, e.g.*, *Cohen* Complaint, *supra* note 9, at 9–15.

311. *See, e.g.*, Klayman v. Zuckerberg, 753 F.3d 1354, 1359 (D.C. Cir. 2014) (granting immunity to Facebook for a claim that sought to hold them liable for failure to remove a user's page about the "Third Palestinian Intifada").

312. Bosker, *supra* note 87.

In *Cohen*, Cohen's complaint seeks to deny immunity to Facebook because its conduct—via its services—has materially contributed to the effectiveness of dissemination and networking of a known FTO.[313] The plaintiffs allege that the system's ability to bucket and categorize people based on their interests makes Facebook more than a mere publisher.[314] The complaint makes clear that the content at issue is the connective tissue of the website, specifically the algorithms:

> Facebook's algorithms suggest friends based on such factors as friends of friends, group membership, geographic location, event attendance, language, etc. Thus, Facebook purposefully provides users who have expressed an interest in the "Knife Intifada" or stabbing Jews, by joining groups or attending events with those themes, with friend suggestions of other like-minded people, and thereby helps to build large groups of people sharing and cultivating similar bigoted hatreds and murderous inclination.[315]

These connective services are not those of a traditional publisher. It is not "the very essence of publishing"[316] to assemble personalized suggestions based on historical knowledge of a user. However, given past courts' reliance on underlying policies of § 230, specifically the utilitarian aspect of relieving ISPs of the duty to monitor,[317] the court may deem these services within the scope of publisher conduct. An alternate interpretation would require the intensive monitoring originally deemed impractical.

To overcome the purposivist argument that the statute is meant to relieve the ISP of a publisher's duty to monitor, the court could assess the level of technological advancement a site utilizes in providing these services. To the extent that the algorithms are bucketing users and organizations as "interests: Knife Intifada," the site may reasonably have the capability to monitor potential violence. But, as machine-learning algorithms do not operate with such transparency, this capability is likely to be grouped in the publisher camp, and, as such, to be protected by § 230.

The second prong of the *Zeran* analysis may require discussion of the operation of the algorithms that drive the services. The court must either conduct limited discovery to demonstrate whether the services are the type that congressional intent would indicate fall into the scope of § 230 immunity, or whether new, nonpublisher technology is easily subject to

---

313. *Cohen* Complaint, *supra* note 9, at 21 ("By providing *Services* that facilitate, encourage, and broker the connections between and among Palestinian terrorist organizations and hundreds of thousands of individuals and groups who have indicated through their Facebook pages that they sympathize with the terrorist-cause, defendant Facebook has performed acts which a reasonably prudent person would not have committed under the same circumstances . . . ." (emphasis added)).

314. *Id.* at 20 ("Facebook's active involvement in making connections between its users, as outlined above, renders it as far more than a neutral and passive bulletin board for information provided by others. Its active role in making connections between terrorist inciters and terrorists, leading to murder, requires that it be held accountable for its actions.").

315. *Id.* at 17–18.

316. *Klayman*, 753 F.3d at 1359 (noting "the very essence of publishing is making the decision whether to print or retract a given piece of content").

317. *See supra* Part I.A.2.

monitoring. Assuming that the algorithmically driven connective services cannot effectively be monitored due to their opacity, Facebook will likely be considered a publisher with respect to these claims.

2. Assessing Codevelopment

There are two distinct claims that can be made under the development prong of the *Zeran* analysis.[318] The first considers the content at issue to be the terrorists' expanded network and claims Facebook is a codeveloper of these relationships. The second considers the content at issue to be the organization of content that a user engages with; the claim being that this organization saps the user of enough control as to constitute codevelopment of the content that user views and is inspired by.

In assessing the codevelopment prong, the first step is to assess the ISP-user relationship to determine whether the actions taken constitute development. In this analysis, the central considerations are (1) whether the collection of information is required as a contingency of use and (2) whether the ISP limits the parameters of the expression of that information.[319]

The assessment of the ISP-user relationship where the content at issue is the introduction of terrorists to new recruits likely cuts against Facebook. Under the first ISP-user issue, the data collected about a user's interests is required by Facebook, and under the second question the types of engagement Facebook tracks are limited and controlled wholly by the company. This type of limitation of expression combined with the users' required participation, however, is not dispositive of the issue.

The second major consideration is the ISP's use of the user's content: the ISP-content relationship. Where the content at issue is terrorist relationships, Facebook cannot fairly be considered the "creator" or "developer" of a relationship because while Facebook suggests connecting users, the fact that users have ultimate control over the acceptance of the suggestions from Facebook is key. Facebook neither forced the users together nor initiated contact between them, and therefore cannot be held liable here. Thus, the platform likely will not lose immunity as a codeveloper for connective services with minimal suggestive influence.

However, where the content at issue is the suggestive organization of content, which incorporates the power that the medium-as-message has over the user, the issue becomes more difficult to determine. The first consideration of the ISP-user relationship cuts against Facebook where the collection of behavioral data and network data[320] cannot be controlled by the user. Granting Facebook the license to use this data is a required condition of using the platform.[321] Therefore, the information used to assemble the

318. As this is a new framework, the claims in the pleadings are not plead to fit within these considerations and as such, this Note utilizes the basic fact pattern to propose the way claims can be brought in similar instances.

319. *See supra* Part III.A.

320. *See, e.g.*, *supra* note 315 and accompanying text.

321. *See Data Policy*, FACEBOOK (Sept. 29, 2016), https://www.facebook.com/full_data_use_policy [https://perma.cc/SB5V-YANM].

content to be as engaging as possible is content in its own right that is gleaned from the user without user-defined controls. The resulting display of third-party content is similarly created and displayed using engagement-focused organization rather than user filters.

Second, when considering whether there are adequate options for expression of the information required by the ISP, one must look to the information being required from the user and what control the user has over the manner in which collection occurs. This does not fit neatly into a prepopulated select-a-box or additional-comments type of analysis as in *Roommates*. The issue giving rise to liability in *Roommates* was that the prepopulated questions had suggested answers and limited the possible answers to those suggestions.[322] When assessing Facebook's content control, this can be likened to the top posts, suggested posts, recommended events, or groups posts that are on the page. This consideration is not dispositive against Facebook because, while Facebook can suggest friends, events and posts, the user also has the open search box to use. However, in *Roommates*, it was not the presence of the opportunity to select all categories from a drop-down menu that made the profiles lose immunity; rather, it was Roommates.com's limitation of the options that lost immunity. This consideration cuts slightly against Facebook where a newsfeed, minimally controllable by users, appears to be curated to engage the user through subtle psychological forces, rather than overt user action.[323]

The second consideration—the ISP-content relationship—also cuts against Facebook. While users own the content and information on Facebook and are free to post what they wish, Facebook gleans information from that content.[324] User content and the data gleaned from user content have materially different functions and are not analogous. The former is used for communication between users while the latter is used by Facebook to conduct its business of creating engaged users and selling their attention to advertisers.[325] The processing that goes into each users' newsfeed is the result of behavioral analysis conducted on the user and the users' network. This aggregated information is then applied to the personalized information of users to create the highest engagement rate.

Finally, as to culpability and the material contribution to development, Cohen's complaint makes plain that Facebook is aware of its active role in creating engagement.[326] This fact is magnified by the studies Facebook has

322. *See supra* Part II.C.1.

323. *See supra* note 94 and accompanying text.

324. *See supra* Part I.C.1.

325. *See supra* notes 102–07 and accompanying text; *see, e.g.*, Fair Hous. Council of San Fernando Valley v. Roommate.com, LLC, 521 F.3d 1157, 1172 (9th Cir. 2008) (en banc) ("Roommate both elicits the allegedly illegal content and makes aggressive use of it in conducting its business.").

326. *Cohen* Complaint, *supra* note 9, at 17 ("Facebook's computers implement algorithms which utilize the data it collects to suggest friends, groups, products, services and local events, and target ads that will be 'as relevant and interesting' as possible to each individual user. This enables Facebook to customize its Services to the specific likes and interests of each of its users.").

conducted. Facebook is aware of the mechanisms that allow FTOs to network and incite effectively and categorize them before they are deployed on profiles. By processing this information, the complaint alleges that Facebook "has the technical ability to monitor the material that appears on its websites."[327] The complaint alleges that Facebook's acknowledgement that sophisticated algorithms are behind the categorization and matching makes it more than a mere passive conduit with respect to users' information.[328] Indeed, Facebook uses the data "actively [to] direct those who are most interested in carrying out Terror Attacks to the incitement that will lead them to their goal, and . . . introduce[] the inciters to people who are interested in murder."[329] It is the algorithmic introduction of engaging content, and not the content itself, that is at issue here.

Further, the civil conspiracy claim alleges that "[a]t all relevant times, defendant Facebook knew that those users posting incitement to commit Terror Attacks were utilizing defendant's Services in order to direct their murderous Incitement to the widest possible audience of individuals most likely to act upon it."[330] The complaint brings the allegation one step further by implicating Facebook's knowledge of its own ability to incite action among people by stating that Facebook knew the information disseminated "likely would result in Terrorist Attacks targeting plaintiffs, causing the harm to plaintiffs set forth herein."[331]

Where Facebook merely posts information to the page of the user and the pages of other users, there is not codevelopment. But where the platform has gleaned new information from user content and user behavior, in order to create a site architecture that affects both mood and behavior, influence can be likened to the limitation of profiles emailed to users in *Roommates*[332] or the creation of new content found in *Fraley*.[333] The platform uses knowledge it has gathered from users to better understand their needs and to run a more effective business.

C. The Rationale Behind Denying Immunity Where the Content at Issue Is Subject to "Development" or Where the ISP's Conduct, Rather than the Content, Is at Issue

As technology becomes more sophisticated, the approach to § 230 immunity must become more nuanced.[334] Determining a set of principles to guide the application of § 230 is critically important to society and global business operations.[335] Blanket immunity may have been satisfactory when

327. *Id.* at 2.
328. *Id.*
329. *Id.*
330. *Id.* at 29.
331. *Id.* In this way, the complaint contends that there is no protection for this speech at all, because "fighting words" are not considered "information" under the statute. *Id.* at 31.
332. *See supra* Part I.C.2.
333. *See supra* notes 287–89 and accompanying text.
334. *See* Taubel, *supra* note 72, at 388–89.
335. *See* Bosker, *supra* note 87.

the primary concern was defamation,[336] but claims are no longer confined to traditional tort liability for traditional publishers' actions. Instead, they are rooted in the consequences that new manifestations of technology have on the physical world. Section 230's purpose was to protect ISPs from the liability posed by creating the space for speech, but liability is no longer derived solely from amplification of third-party conduct. ISPs once provided a soapbox and a bullhorn but now craft elaborate set designs of user content; as such, they are worthy of accreditation where applicable. Courts must understand and analyze the precise allegations, and the relationships between the content and the users, before assigning labels to content providers and ISPs.

While courts' traditional broad construction of § 230 immunity can act as a shield to discovery in close cases, a framework that considers the appropriate ISP-user and ISP-content relationships should be applied where the technology in question may have a hand in development of information or where the conduct of that technology is the source of the alleged harm. The importance of a judicial shift toward more nuanced evaluation is paramount because it opens the doors to liability. This shift may spur consideration of the ethical considerations that have long gone unaddressed in the development of algorithmic technology.[337] Thus, any bar to giving more power to these considerations is not one that should be casually granted.

The ability to appropriately assign liability based on algorithmic configuration will determine the course of computer-science education[338] and will determine the transparency we require of companies' online behavioral analyses. If judges cannot articulate a clear manner in which to approach issues of algorithmic site architecture, then society will need to fill in the gaps with additional legislation. But that is not the necessary conclusion. In fact, it is one that should be avoided. Punting this issue back to the legislature will leave the determination of normative liability in the hands of the most powerful and influential people in the public sphere—the ISPs and their lobbyists.[339] It will also detract from the flexibility that the judiciary provides in close cases.

While innovation is good for business, there should be some check in place whereby ISPs can be held liable for the services, and not the content, that they provide. A judicial solution will afford the greatest flexibility to determine, in highly fact-specific instances of algorithmic development,

336. *See supra* Part I.A.1.

337. *See* Bosker, *supra* note 87 (noting that there is a lack of understanding and consensus about the ethics of rapidly evolving technology).

338. *See id.* (noting that some coders wish to create a type of Hippocratic Oath to hold computer scientists accountable for perceived psychological tampering).

339. *See* Tom Porter, *Congress Versus Internet: Lawmakers Want to Stop Russian Interference and Sex Traffickers on the Web*, NEWSWEEK (Sept. 20, 2017, 2:49 PM) http://www.newsweek.com/senators-battle-sillicon-valley-sex-trafficking-russia-interference-668371 [https://perma.cc/GQ37-VUKW] (highlighting the deployment of technology companies' "substantial lobbying resources," which they use to combat a proposed revision of § 230 immunity).

whether development has taken away user control with respect to the content at issue.

CONCLUSION

While § 230 has aided the development and growth of the internet, that purpose has to a large extent been served.[340] Internet companies are some of the most powerful[341] and profitable[342] in existence. Broad civil and criminal immunity should no longer blindly be granted where new technology drastically changes the analytical foundations of preceding case law. The internet has outgrown § 230 immunity as a broadly applied doctrine. The narrowing of immunity as more statutory claims are brought is evidence of that.[343] Defamation and other tort claims that rely on simple fact patterns to assign vicarious liability to "publishers" should, due to free speech interests and justiciability concerns, continue to be dismissed pursuant to § 230. But, algorithmic technology that springs forth distinct from user-generated content and is powerful enough to influence human behavior should be given due consideration in a revised framework, instead of obtaining customary immunity.[344]

Facebook is aware that the way in which it disseminates material has a real effect on users. It is also aware that such dissemination affects the brain differently than the way a "traditional" publisher's dissemination affects the brain. Because of these considerations, some of Facebook's services do not fall within the bounds of immunity. Some of its actions, though akin to publication (services that disseminate user content to the masses), are conducted with knowledge and intent to have a substantive effect on users' mental state furthered by the intentional development of data. Where allegations incorporate the effects of social, emotional, and behavioral contagion powered by data gleaned from users' interactions into their claim, a platform like Facebook will fall outside of § 230's immunity.

A solution that compels discussion and disclosure of at least some of the technology being deployed forces accountability for the powerful mechanisms at play. This will catalyze responsible decision-making in

340. *See* Zara, *supra* note 187 ("What began as a provision to promote the growth of an emerging technology is now a legal tool to protect the business interests of the powerful."); *see also* Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401, 409–10 (2017).

341. *See* Chenda Ngak, *SOPA and PIPA Internet Blackout Aftermath, Staggering Numbers*, CBS NEWS (Dec. 19, 2012, 4:48 PM), http://www.cbsnews.com/news/sopa-and-pipa-internet-blackout-aftermath-staggering-numbers/ [https://perma.cc/39WS-NRYP] (reporting that in a one-day protest, websites banned together to fight the passage of SOPA and PIPA, mobilizing 4.5 million people to sign a petition and an additional 350,000 to email representatives).

342. Erin Griffith, *Here Are the 51 Technology and Telecommunications Companies of the Fortune 500*, FORTUNE (June 7, 2016, 6:00 AM), http://fortune.com/2016/06/07/fortune-500-technology-companies/ [https://perma.cc/95A7-MMHW] (noting just over one-tenth of Fortune 500 companies are technology companies).

343. *See* Zara, *supra* note 187 (discussing the recent trend of judges chipping away at § 230 immunity); *see also supra* notes 183–84 and accompanying text.

344. *See supra* Part I.C (discussing the types of data, unrelated to content, that Facebook gathers to inform personalized newsfeeds).

technology sectors moving forward. No longer will the mindless activation of base human impulses be the driving factor and main concern of the ISP. The ISP will shoulder the burden of potential liability for the services it provides that have broken the boundaries of the publisher role. As a result, perhaps more responsible technologists will mean that fewer victims of terrorist incitement will have to turn the other cheek.