

COMMENT

OPTIMIZED FOR ADDICTION: EXTENDING PRODUCT LIABILITY CONCEPTS TO DEFECTIVELY DESIGNED SOCIAL MEDIA ALGORITHMS AND OVERCOMING THE COMMUNICATIONS DECENCY ACT

ALLISON ZAKON*

Over the past decade, social media has gained an ever more pervasive presence in our lives. But the social connection, real-time dissemination of information, and the creative outlet it provides come with a cost that is invisible to most users: to increase site value, many platforms are designed to addict users and trap them in a cycle of dependence that can ravage their mental health and wellbeing. One of the most effective ways for platforms to maximize their value is by using recommendation systems: sophisticated software systems that utilize artificial intelligence to learn about the user and predict what kinds of content will keep them scrolling, clicking, and liking.

This Comment sketches out a legal solution for a subset of the harms caused by these algorithms by arguing that product liability concepts should extend to social media platforms, who have managed to achieve a nearly impenetrable form of immunity due to Section 230 of the Communications Decency Act (CDA). The first part of this Comment summarizes the relevant concepts, including the recommendation engines that drive social media, the mental health effects of social media use, and the origins and purpose of the CDA. Next, this Comment demonstrates that social media platforms qualify as products for the purposes of strict product liability and walks the reader through a hypothetical example of the risk-utility test. Finally, this Comment considers Section 230 of the CDA and offers two avenues for circumventing its broad grant of immunity. Instead of joining the extensive body of literature lamenting the woes of the CDA, this Comment proposes an affirmative solution that enables victims to recover while holding the multi-billion-dollar social media industry to a higher standard and incentivizing greater levels of caution in its development process.

Introduction	1108
I. Background.....	1111
A. Introduction to Algorithms	1111
B. Social Media Use and Mental Health.....	1115
C. Product Liability and the Theory of Defective Design	1117
D. The Communications Decency Act.....	1119
II. The Hurdles to Making the Case: Applying Concepts of Product Liability and Circumventing the CDA	1121
A. Strict Product Liability Applies to Software	1121

* J.D. Candidate, 2021, University of Wisconsin Law School; B.A. 2016, Barnard College. Many thanks to the members of *Wisconsin Law Review* for all of their hard work and to my friends and family for their support throughout this process.

B. Proving Defective Design	1127
C. Overcoming the Communications Decency Act.....	1132
1. Social Media Platforms Do Not Act in a Publishing Capacity When They Recommend Personalized Content.....	1133
2. The Use of Recommendation Algorithms Converts a Platform Into an Information Content Provider	1140
Conclusion.....	1144

INTRODUCTION

In the twenty-first century—an era of internet hegemony—algorithms control our lives by influencing our behavior in countless, yet often inconspicuous, ways.¹ From impacting decisions as simple as where we direct our attention to what we purchase to as complex as whom we marry, algorithms hold increasing authority over our daily lives.² It is undeniable that some of these algorithms make our lives easier and more enjoyable.³ But, these seemingly innocuous tools can produce unintended consequences that cause serious harm.⁴ Over the past several years, internet giants like Facebook, Instagram, and YouTube have come under increased scrutiny for recommending incendiary and addictive content to their users.⁵ Recommendation engines designed to maximize platform value propel much of this harmful content, as researchers build and deploy algorithms that optimize for value creation metrics,⁶ such as user engagement and time spent on the platform, with little regard for the long-term costs to users.⁷

1. See KARTIK HOSANAGAR, A HUMAN'S GUIDE TO MACHINE INTELLIGENCE: HOW ALGORITHMS ARE SHAPING OUR LIVES AND HOW WE CAN STAY IN CONTROL 27–31 (2019) (discussing the various ways in which algorithms impact our daily lives).

2. See *id.*

3. See *id.* at 27.

4. See *infra* notes 7–15 and accompanying text.

5. See *infra* notes 7–15 and accompanying text.

6. The primary objective of any firm is to maximize shareholder value, which is determined by net present value. Net present value equals the present value of expected future cash flow minus the value of investments. This, in turn, drives shareholder value. See RICHARD A. BREALEY, STEWART C. MYERS & FRANKLIN ALLEN, PRINCIPLES OF CORPORATE FINANCE 9–10, 20–23 (10th ed. 2011). Logically, this means that a firm's algorithms should align with its valuation goals. See also Jeff Jordan, Anu Hariharan, Frank Chen & Preethi Kasireddy, *16 Startup Metrics*, ANDREESSEN HOROWITZ (Aug. 21, 2015), <https://a16z.com/2015/08/21/16-metrics/> [<https://perma.cc/L3WF-R98V>]; Anu Hariharan, Frank Chen & Jeff Jordan, *16 More Startup Metrics*, ANDREESSEN HOROWITZ (Sept. 23, 2015), <https://a16z.com/2015/09/23/16-more-metrics/> [<https://perma.cc/GL9C-ZNXC>].

7. See Renee DiResta, *Up Next: A Better Recommendation System*, WIRED (Apr. 11, 2018, 11:00 AM), <https://www.wired.com/story/creating-ethical-recommendation-engines/> [<https://perma.cc/HRW6-WKWQ>].

Unsurprisingly, the effects of these behavior-influencing algorithms can be severe. In 2017, a British teenager named Molly Russel took her own life after months of viewing pro-suicide and self-harm content recommended to her by the social media platforms that she frequented, including Instagram and Pinterest.⁸ Even after her death, Pinterest continued to send her personalized emails containing graphic imagery, “including a slashed thigh and cartoon of a young girl hanging.”⁹ As news of her passing spread, thirty other families came forward alleging that social media played a role in their children’s suicides too.¹⁰

In early 2019, news broke that YouTube was recommending explicit self-harm videos with titles like “[m]y huge extreme self-harm scars” and search terms like “how to self-harm tutorial” to its users—many of whom were young children.¹¹ That same month, the platform came under fire for recommending compromising content of young children, enabling a “pedophilia ring” to proliferate on the site where users left obscene comments and posted links to child pornography.¹² While YouTube

8. Ysabel Gerrard & Tarleton Gillespie, *When Algorithms Think You Want to Die*, WIRED (Feb. 21, 2019, 12:41 PM), <https://www.wired.com/story/when-algorithms-think-you-want-to-die/> [https://perma.cc/MT58-QSR2].

9. *Id.*

10. Faith Ridler, *THIRTY Families Blame Social Media Firms for Their Roles in Children’s Suicides as it Emerges Pinterest Sent a Personalised Email to Molly Russell’s Account with Self-Harm Images AFTER She Took Her Own Life*, DAILYMAIL (Jan. 27, 2019, 1:20 PM), <https://www.dailymail.co.uk/news/article-6636807/Now-30-families-blame-social-media-firms-roles-childrens-suicides.html> [https://perma.cc/T9W8-9EKC]. Incidents like these prompted the British Information Commissioner’s Office to develop a new privacy code, the Age Appropriate Design Code, which was approved by Parliament and became effective September 2, 2020. The code sets out standards to “ensur[e] that the best interest of [children] are the primary consideration when designing and developing online services.” Elizabeth Denham, *Age Appropriate Design: A Code of Practice for Online Services*, INFO. COMM’R’S OFF., <https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/age-appropriate-design-a-code-of-practice-for-online-services/#:~:text=The%20code%20is%20a%20set,designing%20and%20developing%20online%20services> [https://perma.cc/UC4F-CF3F] (last visited Oct. 1, 2020); John Glenday, *ICO Unveils ‘Transformational’ Code to Force Facebook and Google to Protect Kids Online*, THE DRUM (Jan. 22, 2020, 11:18 AM), <https://www.thedrum.com/news/2020/01/22/ico-unveils-transformational-code-force-facebook-and-google-protect-kids-online#:~:text=The%20Age%20Appropriate%20Design%20Code,after%20viewing%20inappropriate%20Instagram%20material> [https://perma.cc/7MUT-LKXN].

11. Daniyal Malik, *YouTube Faces Severe Criticism For Recommending Self Harm Videos Again*, DIGIT. INFO. WORLD (Feb. 7, 2019), <https://www.digitalinformationworld.com/2019/02/youtube-recommending-self-harm-videos-in-search-results-criticized.html> [https://perma.cc/V7KJ-BBNP].

12. Ryan Broderick, *YouTube’s Latest Child Exploitation Controversy Has Kick-Started A War Over How to Fix The Platform*, BUZZFEED NEWS (Feb. 22, 2019, 5:42 PM), <https://www.buzzfeednews.com/article/ryanhatethis/youtube-child-sexual->

responded to the crisis by deleting accounts and suspending comments,¹³ some felt this did not address the core problem—the platform’s use of artificial intelligence algorithms built to maximize engagement metrics above all else.¹⁴ The fact that so many predators spent an inordinate amount of time watching these videos is what caused YouTube’s algorithms to recommend them so widely and what allowed the ring of activity to flourish.¹⁵

Despite the potentially egregious consequences that these algorithms can produce,¹⁶ some courts today would find the responsible platforms immune from liability under Section 230 of the Communications Decency Act (CDA), which grants information service providers (ISPs) broad immunity for content posted by third party users.¹⁷ This Comment argues that the algorithms through which social media platforms recommend content fall outside the scope of the CDA and are therefore not implicated in judicially created tests for applying the statute. Accordingly, social media platforms should be liable to their users under theories of strict product liability and accountable for the harms they accrue in their quests to maximize site value.

The prevalence of algorithms in our daily lives is a relatively new phenomenon, and so, lawsuits surrounding defective algorithms have been sparse.¹⁸ As a result, there is still gray area in the law and room for skepticism. Some experts believe that strict product liability concepts

exploitation-creators-watson [https://perma.cc/X7EQ-XVX9]; K.G. Orphanides, *The Paedophile Scandal Shows YouTube is Broken. Only Radical Change Can Fix It*, WIRED (Feb. 23, 2019), https://www.wired.co.uk/article/youtube-paedophiles-boycott-algorithm-change [https://perma.cc/6RGM-548U].

13. Orphanides, *supra* note 12.

14. *Id.* (“‘The heart of the problem is that these paedophiles spend hours on YouTube watching these [videos]’ says Chaslot. YouTube’s AI is primarily optimised for watch time, and if paedophiles sink thousands of hours of view time into videos of children that they find appealing, those videos are more likely to be recommended.”).

15. *Id.*

16. This Comment will primarily focus on mental-health-related physical harms as they are most likely to give rise to successful product liability lawsuits, although faulty or misguided algorithms cause a multitude of societal harms. *See infra* notes 298–300 and accompanying text.

17. *See* Communication Decency Act of 1996, 47 U.S.C. § 230(c) (2018); *Gonzalez v. Google, Inc.*, 335 F. Supp. 3d 1156, 1174–75 (N.D. Cal. 2018); *Herrick v. Grindr L.L.C.*, 765 F. App’x 586, 589–90 (2d Cir. 2019).

18. *See* Frances E. Zollers, Andrew McMullin, Sandra N. Hurd & Peter Shears, *No More Soft Landings for Software: Liability for Defects in an Industry That Has Come of Age*, 21 SANTA CLARA COMPUT. & HIGH TECH. L.J. 745, 766–68 (2005).

should not apply to less tangible items like software,¹⁹ while others believe that the CDA's reach is broad enough to encompass algorithms.²⁰

This Comment examines the various challenges involved in pursuing a strict liability cause of action and proves to prospective litigants that building a case is workable under current legal regimes. Part I of this Comment introduces the reader to recommender algorithms and their impact on social media users while setting out the relevant law, including the applicable theory of strict product liability and the Communications Decency Act. Part II discusses the core hurdles involved in making a case. It establishes that software, and thus social media platforms, qualify as products for the purposes of strict product liability. Then, it proceeds to discuss the nuances involved in proving the defective design theory—the theory of product liability best suited to less tangible products like software. The final section turns to the CDA and current case law to analyze the statute itself and show why Section 230 does not cover defective recommender algorithms.

I. BACKGROUND

As the myriad dangers of social media have come to the forefront in recent years, those who have tried to seek legal recourse have confronted a system that is ill-equipped to provide it.²¹ Plainly, there are many non-redressable harms under the current state of the law.²² This Part will introduce the reader to the legal and technical frameworks under which a plaintiff must operate, while Part II will demonstrate that litigants seeking to hold social media companies accountable for their defectively designed algorithms have a viable cause of action.

A. Introduction to Algorithms

An algorithm is merely a step-by-step procedure to accomplish a goal.²³ Algorithms are essential to machine learning (ML), a field of artificial intelligence (AI) in which systems “extract[] patterns from raw

19. See, e.g., Karni A. Chagal-Feferkorn, *Am I an Algorithm or a Product? When Products Liability Should Apply to Algorithmic Decision-Makers*, 30 STAN. L. & POL'Y REV. 61, 66 (2019).

20. See, e.g., Karni A. Chagal-Feferkorn, *The Reasonable Algorithm*, 2018 U. ILL. J.L. TECH. & POL'Y 111, 121–30.

21. See *infra* note 96 and accompanying text.

22. See, e.g., *Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093 (9th Cir. 2019); *Doe v. Backpage.com, L.L.C.*, 817 F.3d 12 (1st Cir. 2016).

23. HOSANAGAR, *supra* note 1, at 5; G. MICHAEL SCHNEIDER & JUDITH L. GERSTING, *INVITATION TO COMPUTER SCIENCE* 11–17 (6th ed. 2013).

data” to solve real-world problems.²⁴ On social media, recommender algorithms solve a specific optimization problem: what content should the user see next?²⁵ The answer determines how long the user stays on the site, whether they choose to engage, and when they choose to return.²⁶

The social media industry’s core revenue stream depends on advertisers who purchase user engagement and user attention.²⁷ The more time an individual spends on the site, the more paid ads the platform can show, and the more money it can make.²⁸ It is in this context that recommendation has emerged as one of the most effective ways that platforms can keep users clicking and viewing ads on their site to the tune of billions of dollars in revenue.²⁹

But how are these platforms able to employ recommendations to generate so much revenue so effectively? They leverage sophisticated ML algorithms that study user behavior.³⁰ These algorithms examine the content individuals have interacted with in the past as well as content similar users have interacted with to determine which unique layouts, content, and sequences of content are most likely to captivate users.³¹ Once the system selects a permutation of choices to show, it observes the user’s

24. See IAN GOODFELLOW, YOSHUA BENGIO & AARON COURVILLE, *DEEP LEARNING 2* (2016).

25. See KIM FALK, *PRACTICAL RECOMMENDER SYSTEMS 23* (2019).

26. See *RECOMMENDER SYSTEMS HANDBOOK 5* (Francesco Ricci, Lior Rokach, Bracha Shapira & Paul B. Kantor eds., 2011) (discussing the various reasons that service providers utilize recommendation systems, such as increasing “user satisfaction” and “fidelity”).

27. Catherine Price, *Trapped – The Secret Ways Social Media Is Built to Be Addictive (and What You Can Do to Fight Back)*, *SCI. FOCUS* (Oct. 29, 2018, 8:00 AM), <https://www.sciencefocus.com/future-technology/trapped-the-secret-ways-social-media-is-built-to-be-addictive-and-what-you-can-do-to-fight-back/> [<https://perma.cc/MF59-K77H>].

28. *Id.*

29. *Id.*; see also *Facebook, Inc. Class A Common Stock (FB)*, NASDAQ (2020), <https://www.nasdaq.com/market-activity/stocks/fb/financials> [<https://perma.cc/V8NL-PGYW>]; *Twitter, Inc. Common Stock (TWTR)*, NASDAQ (2020), <https://www.nasdaq.com/market-activity/stocks/twtr/financials> [<https://perma.cc/MXE6-YWUC>].

30. See Zeynep Tufekci, *Yes, Big Platforms Could Change Their Business Models*, *WIRED* (Dec. 17, 2018, 6:00 AM), <https://www.wired.com/story/big-platforms-could-change-business-models/> [<https://perma.cc/EA5R-5GV5>] (“To microtarget individuals with ads, today’s platforms massively surveil their users; then they use engagement-juicing algorithms to keep people onsite as long as possible.”). For instance, Facebook tracks, stores, and analyzes *exabytes* of user data. One exabyte is “equivalent to 1 million hard drives inside a contemporary PC.” Mike Rogoway, *Facebook Plans ‘Cold Storage’ for Old Photos in Prineville*, *OREGONIAN* (Jan. 10, 2019), https://www.oregonlive.com/silicon-forest/2013/02/facebook_plans_cold_storage_fo.html [<https://perma.cc/UCY4-BRJU>].

31. *RECOMMENDER SYSTEMS HANDBOOK*, *supra* note 26, at 11, 466.

behavior and evolves independently via reinforcement learning paradigms.³² Using an iterative process of trial and error over time, recommender systems learn which highly personalized suggestions will delight users most.³³ But while these algorithms may be mathematically sophisticated,³⁴ they often do not understand the content they are promoting.³⁵ Nor can they recognize when they are promoting content to vulnerable individuals.³⁶

In general, ML algorithms strive to produce models that meet (or exceed) a pre-specified goal.³⁷ A model's goals are defined by what is known as an objective function, which is a representation of the decision problem the algorithm is built to solve.³⁸ Often, the objective function aligns with the platform's ultimate goal of maximizing value.³⁹ Like all firms, social media platforms seek to increase the net present value of future cash flows by making decisions that balance short-term gains with

32. See GOODFELLOW, BENGIO & COURVILLE, *supra* note 24, at 25; RICHARD S. SUTTON & ANDREW G. BARTO, REINFORCEMENT LEARNING: AN INTRODUCTION 450 (2d ed. 2018) (“A reinforcement learning system can improve a recommendation policy by making adjustments in response to user feedback . . . it is common to monitor user clicks as indicators of interest in a link.”); DiResta, *supra* note 7.

33. DiResta, *supra* note 7; see also SARA WACHTER-BOETTCHER, TECHNICALLY WRONG: SEXIST APPS, BIASED ALGORITHMS, AND OTHER THREATS OF TOXIC TECH 94 (2017) (“Delight is a concept that’s been tossed around endlessly in the tech industry . . .”).

34. Jeremy Howard, Margit Zwemer & Mike Loukides, *Designing Great Data Products: The Drivetrain Approach: A Four-Step Process for Building Data Products*, O’REILLY (Mar. 28, 2012), <https://www.oreilly.com/radar/drivetrain-approach-data-products/> [<https://perma.cc/V7GJ-P86A>].

35. DiResta, *supra* note 7; see also Guillaume Chaslot, *The Toxic Potential of YouTube’s Feedback Loop*, WIRED (July 13, 2019, 9:00 AM), <https://www.wired.com/story/the-toxic-potential-of-youtubes-feedback-loop/> [<https://perma.cc/YW9E-FYRK>].

36. Gerrard & Gillespie, *supra* note 8.

37. SUTTON & BARTO, *supra* note 32, at 6; CATHY O’NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 73 (2016).

38. SUTTON & BARTO, *supra* note 32, at 6, 62. Note that another term for objective function is value function. See, e.g., Howard, Zwemer & Loukides, *supra* note 34.

39. See, e.g., PAUL COVINGTON, JAY ADAMS & EMRE SARGIN, DEEP NEURAL NETWORKS FOR YOUTUBE RECOMMENDATIONS (2016) (“Our final ranking objective is constantly being tuned based on live A/B testing results but is generally a simple function of expected watch time per impression.”); Nicolas Koumchatzky & Anton Andreyev, *Using Deep Learning at Scale in Twitter’s Timelines*, TWITTER BLOG (May 9, 2017), https://blog.twitter.com/engineering/en_us/topics/insights/2017/using-deep-learning-at-scale-in-twitters-timelines.html [<https://perma.cc/BZ49-AJUS>] (“The set of metrics we use [to assess model quality] usually relate more directly to usage and enjoyment of Twitter. For instance, we may track the number of engagements per user, or the total amount of time individuals spend on Twitter.”).

long-term value, thereby augmenting the value of their equity.⁴⁰ By extension, recommender algorithms behave much in the same way and make similar tradeoffs.⁴¹ However, net present value is not always easy to measure directly, so these algorithms optimize instead for behaviors that predict net present value.⁴² Specifically, net present value can decompose into more readily controllable units,⁴³ such as time spent (the total time spent over each user's lifetime)⁴⁴ and user engagement,⁴⁵ which social media algorithms optimize over the long run.⁴⁶ Therefore, time spent and user engagement are common components of objective functions given the value they generate.⁴⁷ Both metrics require action from users, and if they are unconstrained, they can erode user welfare.⁴⁸

While recommending content is beneficial in many respects, the underlying methods are often exploitative or "opportunistic."⁴⁹ To be clear, recommender algorithms are not per se harmful, yet as explained above, companies stand to gain from increasing metrics like sales, clicks, or views.⁵⁰ These incentives have come to drive the way sites render content⁵¹ and motivate platforms to design their sites to be addictive, with features ranging from infinite scroll to the recommender algorithms at issue here.⁵² As the inventor of the infinite scroll feature, Aza Raskin, retrospectively reasoned:

It's as if they're taking behavioral cocaine and just sprinkling it all over your interface . . . [b]ehind every screen on your phone, there are generally like literally a thousand engineers that have

40. See BREALEY, MYERS & ALLEN, *supra* note 6.

41. See SUTTON & BARTO, *supra* note 32, at 6.

42. See DiResta, *supra* note 7.

43. For an overview of commonly utilized metrics in the startup context, see Jordan, Hariharan, Chen & Kasireddy, *supra* note 6.

44. *Id.*

45. See DiResta, *supra* note 7; Gerrard & Gillespie, *supra* note 8.

46. See SUTTON & BARTO, *supra* note 32, at 6.

47. An objective function typically balances multiple objectives. See COVINGTON, ADAMS & SARGIN, *supra* note 39; see also Katherine J. Wu, *Radical Ideas Spread Through Social Media. Are the Algorithm's to Blame?*, PBS (Mar. 28, 2019), <https://www.pbs.org/wgbh/nova/article/radical-ideas-social-media-algorithms/> [<https://perma.cc/V7KY-9M4T>].

48. See DiResta, *supra* note 7.

49. RECOMMENDER SYSTEMS HANDBOOK, *supra* note 26, at 6.

50. See Price, *supra* note 27.

51. See Sachin Banker & Salil Khetani, *Algorithm Overdependence: How the Use of Algorithmic Recommendation Systems Can Increase Risks to Consumer Well-Being*, 38 J. OF PUB. POL'Y & MKTG. 500, 503 (2019).

52. Hilary Andersson, *Social Media Apps are 'Deliberately' Addictive to Users*, BBC (July 3, 2018), <https://www.bbc.com/news/technology-44640959> [<https://perma.cc/33ZJ-E37T>].

worked on this thing to try to make it maximally addicting . . . In order to get the next round of funding, in order to get your stock price up, the amount of time that people spend on your app has to go up . . . [s]o, when you put that much pressure on that one number, you're going to start trying to invent new ways of getting people to stay hooked.⁵³

Some of these features are incredibly subtle and enhance the user experience in minor, nearly undetectable ways, like how a newsfeed takes several moments to load, or how Facebook gives users a notification when a friend has read their message, or how Snapchat provides a “streak” feature to display the amount of time since two users last interacted.⁵⁴ The effect of these features is cumulative.⁵⁵ But ultimately, users come to a site for its content, and thus, recommendation algorithms represent the quintessential tool in the so-called “arms race for attention”—a race that comes at a significant cost.⁵⁶

B. Social Media Use and Mental Health

Social media use has increased drastically over the past decade. In 2005, the Pew Research Center reported that 5% of American adults used at least one social media platform.⁵⁷ In 2019, that number rose to 72%.⁵⁸ Many users, including a majority of users on Facebook, Snapchat, and Instagram, visit these platforms multiple times a day.⁵⁹ Vulnerable

53. *Id.*

54. Julian Morgans, *The Secret Ways Social Media Is Built for Addiction*, VICE (May 18, 2017, 2:00 PM), https://www.vice.com/en_nz/article/vv5jkb/the-secret-ways-social-media-is-built-for-addiction?utm_source=reddit.com [https://perma.cc/Q6GU-GDQR].

55. See ADAM ALTER, *IRRESISTIBLE: THE RISE OF ADDICTIVE TECHNOLOGY AND THE BUSINESS OF KEEPING US HOOKED* 5 (2017) (“The people who create and refine tech . . . are very good at what they do. They run thousands of tests with millions of users to learn which tweaks work and which ones don’t—which background colors, fonts, and audio tones maximize engagement and minimize frustration. As an experience evolves, it becomes an irresistible, weaponized version of the experience it once was. In 2004, Facebook was fun; in 2016, it’s addictive.”).

56. Price, *supra* note 27 (“There’s an arms race for attention . . . [a]nd if you don’t use every tool at your disposal to ensnare consumers, you’ll be left behind.”). These problems of addiction *do not* stem from human “lack [of] willpower;” the problem instead stems from the fact that “there are a thousand people on the other side of the screen whose job it is to break down the self-regulation you have.” ALTER, *supra* note 55, at 3.

57. *Social Media Fact Sheet*, PEW RESEARCH CENTER (June 12, 2019), <https://www.pewinternet.org/fact-sheet/social-media/> [https://perma.cc/Z7GL-RUU8].

58. *Id.*

59. Andrew Perrin & Monica Anderson, *Share of U.S. Adults Using Social Media, Including Facebook, is Mostly Unchanged Since 2018*, PEW RESEARCH CENTER (Apr. 10, 2019), <https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults->

populations like young children and individuals with pre-existing mental health conditions are among the most frequent users of these sites.⁶⁰ A 2018 survey revealed that 95% of teenagers have access to a smartphone, and 45% of those reported that they were online “almost constantly.”⁶¹

As social media use ascends, researchers are beginning to identify connections between these platforms and adverse mental health effects that range from body image issues to depression and suicide. For instance, a study examining the Instagram use of American and Australian women found that high levels of Instagram use were positively associated with “greater self-objectification” and “body image concerns”—both “predictors of disordered eating and depression.”⁶² A second study conducted at the University of Pennsylvania set out to show *causation* between social media exposure and depression by studying the effects of limiting social media use.⁶³ The study concluded that “limiting social media usage does have a direct and positive impact on subjective well-being over time, especially with respect to decreasing loneliness and depression.”⁶⁴ Finally, a nationally representative study established that the national rise of “suicide-related outcomes” and suicide rates were positively associated with screen time and social media use.⁶⁵

When social media companies design their platforms to trap users in a cycle of dependence, they exacerbate these harms.⁶⁶ Unfortunately, profit generation comes at a cost to user welfare, and so “[t]he attention-based business model of social media platforms means that their goals and the goals of their users are often inherently at odds.”⁶⁷ As one app

using-social-media-including-facebook-is-mostly-unchanged-since-2018/
[<https://perma.cc/W38T-KTST>].

60. Aksha M. Memon, Shiva G. Sharma, Satyajit S. Mohite & Shailesh Jain, *The Role of Online Social Networking on Deliberate Self-Harm and Suicidality in Adolescents: A Systemized Review of Literature*, 60(4) INDIAN J. PSYCHIATRY 384, 390 (2018).

61. Monica Anderson & Jingjing Jiang, *Teens, Social Media & Technology 2018*, PEW RESEARCH CENTER (May 31, 2018), <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/> [<https://perma.cc/3GM2-TMCY>].

62. Jasmine Fardouly, Brydie K. Willburger & Lenny R. Vartanian, *Instagram Use and Young Women’s Body Image Concerns and Self-Objectification: Testing Mediation Pathways*, 20(4) NEW MEDIA & SOC’Y 1380, 1380–81 (2017).

63. Melissa G. Hunt, Rachel Marx, Courtney Lipson & Jordyn Young, *No More FOMO: Limiting Social Media Decreases Loneliness and Depression*, 37 J. OF SOC. & CLINICAL PSYCH. 751 (2018).

64. *Id.* at 766–67.

65. Jean M. Twenge, Thomas E. Joiner, Megan L. Rogers & Gabrielle N. Martin, *Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time*, 6(1) CLINICAL PSYCH. SCI. 3, 3, 8–9 (2018) (looking to previously collected survey data).

66. *See Price, supra* note 27.

67. *Id.*

developer and creator candidly conceded: “[s]ocial media isn’t designed with your long-term happiness in mind: it’s designed to capture as much of your attention as possible right now.”⁶⁸

A 2012 experiment conducted by Facebook Research illustrates the role that recommender algorithms can play in bringing about adverse mental health effects.⁶⁹ The study sought to ascertain how emotion spreads on the site by studying user status updates.⁷⁰ Researchers manipulated user newsfeeds, removing positive posts from some and negative posts from others, and found that people who saw more negative posts internalized that negativity and became more negative in their own posts.⁷¹ Wharton Professor Kartik Hosanagar believes that these results carry over to the context of recommendation: “[i]f you felt energized and ready to take on the world yesterday and feel despondent and hopeless today, it may very well be related to the social media posts selected by your news-feed algorithm.”⁷²

C. Product Liability and the Theory of Defective Design

The central causes of action for defective artificial intelligence are breach of warranty, negligent design, and strict product liability.⁷³ Strict product liability, in turn, has three core theories: defective design, defective manufacture, and failure to warn.⁷⁴ Of these theories, defective design is most applicable to the software context and carries the greatest chance of success for those harmed by social media recommender algorithms.⁷⁵ The application of strict liability is fundamental because it mitigates potentially prohibitive proof problems that accompany a

68. *Id.* Some platforms have been candid about the possibility for these negative mental health effects and have taken steps to allay some of these consequences. *See, e.g.,* Naz Erkan & Sandeep Pandey, *Partnering With Researchers at UC Berkeley to Improve the Use of ML*, TWITTER BLOG (Jan. 29, 2019), https://blog.twitter.com/en_us/topics/company/2019/ucberkeley-twitter-ml.html [<https://perma.cc/D5BL-GER9>]; Ryan Smith, *Instagram Will Never Be ‘The Safest Place on the Internet,’* CCN (Nov. 9, 2019, 2:34 PM), <https://www.ccn.com/instagram-never-the-safest-place-internet/> [<https://perma.cc/ED2X-TZQ5>].

69. *See generally* ADAM D. I. KRAMER, *The Spread of Emotion Via Facebook*, FACEBOOK RESEARCH (May 16, 2012), <https://research.fb.com/publications/the-spread-of-emotion-via-facebook/> [<https://perma.cc/9J4V-8ULR>].

70. *Id.*; HOSANAGAR, *supra* note 1, at 35–37.

71. HOSANAGAR, *supra* note 1, at 36.

72. *Id.*

73. *See* Zollers, McMullin, Hurd & Shears, *supra* note 18, at 757.

74. Brandon W. Jackson, *Artificial Intelligence and the Fog of Innovation: A Deep-Dive on Governance and the Liability of Autonomous Systems*, 35 SANTA CLARA HIGH TECH. L.J. 35, 58 (2019).

75. *See id.* at 58–59.

negligence cause of action.⁷⁶ Indeed, this is the very function the doctrine was designed to serve,⁷⁷ and, therefore, its application to the complex and often inaccessible tech industry makes sense.⁷⁸ To prove strict liability, a plaintiff must show the following:

1. The defendant sells a product that the plaintiff uses;
2. The defendant is the commercial seller of such a product;
3. The plaintiff suffers an injury;
4. When the defendant sold the item, the item was defective; and,
5. The defect was an actual and proximate cause⁷⁹ of the plaintiff's injury.⁸⁰

To show that an item is defective, a plaintiff must prove that the “foreseeable risks of harm posed by the product could have been reduced or avoided by the adoption of a reasonable alternative design . . . and the omission of the alternative design renders the product not reasonably safe.”⁸¹ This language articulates a balancing test known as the risk-utility test, which is the approach employed by a majority of jurisdictions.⁸²

Despite the recent rise of social media and the mass adoption of personal devices facilitating constant user activity,⁸³ strict product liability

76. *Id.* at 60.

77. See Lawrence W. Kessler, *Alternative Liability in Litigation Malpractice Actions: Eradicating the Last Resort of Scoundrels*, 37 SAN DIEGO L. REV. 401, 445–46 (2000).

78. For an in-depth discussion on the benefits of extending strict liability to tech, see Zollers, McMullin, Hurd & Shears, *supra* note 18, at 768–74.

79. Proximate cause is “the result in a particular instance that flows from a condition or potential of general causation Proving such cases may require not only general causation, proof that the exposure or activity could cause the plaintiff’s condition, but also specific causation, that in this particular case it did cause that condition.” *Proximate Cause (Specific Causation)*, BOUVIER LAW DICTIONARY (Desk ed. 2012). Given the obvious efforts on behalf of social media platforms to (1) manufacture addiction with (2) the knowledge of its detrimental impact on user mental health, an attenuation problem is unlikely, and so establishing proximate cause should not be challenging. Accordingly, this Comment does not explore it further. See also Zollers, McMullin, Hurd & Shears, *supra* note 18, at 779.

80. *Products Liability*, LEGAL INFO. INST., https://www.law.cornell.edu/wex/products_liability [https://perma.cc/FJ7G-69ZC]. Plaintiffs can recover for harms including wrongful deaths, physical injuries (i.e., those caused by self-harm or eating disorders), and emotional distress, though recovery for pure emotional distress may vary by jurisdiction. See LOUIS R. FRUMER & MELVIN I. FRIEDMAN, 2 PRODUCTS LIABILITY § 13.03(3)(b), (9) (2020); Jane B. Silverman, *Recovery for Emotional Distress in Strict Products Liability*, 61 CHI.-KENT L. REV. 545 (1985).

81. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 2(b) (AM. L. INST. 1998).

82. Chagal-Feferkorn, *supra* note 19, at 80.

83. See *supra* note 58 and accompanying text.

lawsuits against social media platforms applying the defective design theory have been scarce,⁸⁴ and none to date have been successful.⁸⁵ Nevertheless, these failures are unrelated to the technicalities of this theory, under which liability is entirely possible. Instead, most failures have resulted from an inability to surmount the immunity provisions of the CDA.⁸⁶

D. The Communications Decency Act

Congress passed The Communications Decency Act (CDA) in 1996 with the primary goal of combatting explicit material online.⁸⁷ The early Internet was overrun with obscene material that was all too easy to access.⁸⁸ Accordingly, Congress dedicated much of the Act to regulating and penalizing the transmission of offensive material.⁸⁹ But, hardly a year after its enactment, the majority of the Act was struck down in *Reno v. ACLU*.⁹⁰ After *Reno*, only one provision of the Act remained: Section 230.⁹¹

Section 230 represents a vital part of the CDA's goal to clean up the Internet, as it offers a "good faith" defense⁹² which "'encourage[s] telecommunications and information service providers to deploy new technologies and policies' to block or filter offensive material."⁹³ The protections of this provision are twofold.⁹⁴ First and most significantly, Section 230 confers immunity to interactive computer services acting in a publishing capacity.⁹⁵ Liability falls only on third party individuals who

84. See Jackson, *supra* note 74, at 58–59.

85. *But see Maynard v. Snapchat, Inc.*, 816 S.E.2d 77 (Ga. Ct. App. 2018).

86. 47 U.S.C. § 230(c) (2012); *see, e.g., Herrick v. Grindr L.L.C.*, 765 F. App'x 586 (2d Cir. 2019).

87. Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401, 403 (2017).

88. Natalie Annette Pagano, Comment, *The Indecency of the Communications Decency Act § 230: Unjust Immunity for Monstrous Social Media Platforms*, 39 PACE L. REV. 511, 514 (2018).

89. Benjamin Volpe, Comment, *From Innovation to Abuse: Does the Internet Still Need Section 230 Immunity?*, 68 CATH. U. L. REV. 597, 602 (2019).

90. *See Reno v. ACLU*, 521 U.S. 844 (1997).

91. In *Reno v. ACLU*, the Supreme Court struck down the anti-indecency provisions of the Act because they violated the First Amendment. This decision had no impact on Section 230. *See Reno*, 521 U.S. at 849; Wes Gerrie, *Say What You Want: How Unfettered Freedom of Speech on the Internet Creates No Recourse for Those Victimized*, 26 CATH. U. J.L. & TECH. 26, 33–34 (2018).

92. Gerrie, *supra* note 91, at 35 (quoting 47 U.S.C. § 230(c)(1)–(c)(2)(A)).

93. Citron & Wittes, *supra* note 87, at 404 (quoting S. REP. NO. 104–23, at 59 (1995)).

94. *See* 47 U.S.C. § 230(c)(1)–(2) (2012).

95. *Id.*; *Doe v. Backpage.com, L.L.C.*, 817 F.3d 12, 18 (1st Cir. 2016).

create or develop the content, known as information content providers.⁹⁶ Second, Section 230 allows interactive computer services to filter and screen content without liability for failure to do so effectively.⁹⁷

Courts across the country employ a standard test to determine whether an ISP is entitled to Section 230 immunity. A 2009 Ninth Circuit case stated its elements succinctly: “it appears that subsection (c)(1) only protects from liability (1) a provider or user of an interactive computer service (2) whom a plaintiff seeks to treat . . . as a publisher or speaker (3) of information provided by another information content provider.”⁹⁸ The second and third elements of the test have proven to be the most contentious and produce the most litigation.⁹⁹

Regardless of which element is at issue in litigation, courts have interpreted Section 230 broadly,¹⁰⁰ and in the process, have “built a mighty fortress protecting platforms from accountability for unlawful activity on their systems”¹⁰¹ Yet many believe that the protections afforded are far “too sweeping”¹⁰² and may be implemented “blindly” without careful and serious examination of the facts.¹⁰³ Opponents of the provision contend that the Internet is no longer in need of the sweeping immunity that was warranted over twenty years ago when it was in a far more delicate state.¹⁰⁴ Moreover, they argue that Section 230 unjustly bars most victims of online crimes and torts from recovery.¹⁰⁵ Challengers also note that, under the prevailing judicial interpretation, Section 230 actively protects bad actors who encourage or ignore indecent behavior.¹⁰⁶ These factors have prompted scholars to urge Congress to amend the CDA and rectify the consequences of Section 230.¹⁰⁷ This Comment endeavors to

96. Catherine Tremble, Note, *Wild Westworld: Section 230 of the CDA and Social Networks’ Use of Machine-Learning Algorithms*, 86 *FORDHAM L. REV.* 825, 835 (2017); see *infra* note 258 and accompanying text, for the definition of an information content provider.

97. *Backpage.com*, 817 F.3d at 18.

98. *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1100–01 (9th Cir. 2009).

99. See *infra* Sections II.C.1 & II.C.2.

100. See Citron & Wittes, *supra* note 87, at 406–09.

101. *Id.* at 406, 422.

102. *Id.* at 403.

103. Rachel Seaton, Comment, *All Claims Are Not Created Equal: Challenging the Breadth of Immunity Granted by the Communications Decency Act*, 6 *SETON HALL CIR. REV.* 355, 357 (2010).

104. Citron & Wittes, *supra* note 87, at 411.

105. Volpe, *supra* note 89, at 600–01; Gerrie, *supra* note 91, at 27, 29, 35.

106. Citron & Wittes, *supra* note 87, at 406–09.

107. Pagano, *supra* note 88, at 535–38. In 2018, Congress amended the CDA by passing FOSTA (Fight Online Sex Trafficking Act) and SESTA (Stop Enabling Sex Traffickers Act). These provisions created an exception to § 230 allowing litigants to hold ISPs liable for sex trafficking and prostitution resulting from the use of their websites. Heidi Tripp, Comment, *All Sex Workers Deserve Protection: How FOSTA/SESTA*

add to this scholarship by proposing a new avenue to mitigate these concerns—one that does not rely on congressional action.

II. THE HURDLES TO MAKING THE CASE: APPLYING CONCEPTS OF PRODUCT LIABILITY AND CIRCUMVENTING THE CDA

A successful strict product liability lawsuit involves overcoming two unique challenges. The first step in establishing liability requires showing that software qualifies as a product for the purpose of product liability. The second step is to navigate the risk-utility test for defective designs. Overcoming these hurdles may prove challenging for the first groups of litigants making their cases in courts. Even so, their work will pave the way for future lawsuits and bring about legal development to anchor strict liability for technologies that are not tangible in the traditional sense.

A. Strict Product Liability Applies to Software

Scholars have discussed extending strict product liability to software for nearly twenty years,¹⁰⁸ and some courts have already assumed that these concepts apply to social media platforms.¹⁰⁹ But, because there is still room for skepticism, litigants must undertake a careful analysis. As an initial matter, algorithms are components of software, the core product behind social media platforms.¹¹⁰ The Third Restatement of Torts on Product Liability recognizes that defective parts can be subject to the same form of liability as the whole product:

One engaged in the business of selling or otherwise distributing product components who sells or distributes a component is subject to liability for harm to persons or property caused by a product into which the component is integrated if:

Overlooks Consensual Sex Workers in an Attempt to Protect Sex Trafficking Victims, 124 PENN ST. L. REV., 219, 221–22 (2019). However, this amendment only grants recourse to a limited set of plaintiffs. *See id.*

108. T. Randolph Beard, George S. Ford, Thomas M. Koutsky & Lawrence J. Spiwak, *Tort Liability for Software Developers: A Law & Economics Perspective*, 27 J. MARSHALL J. COMPUT. & INFO. L. 199, 207 (2009).

109. *See, e.g., Herrick v. Grindr, L.L.C.*, 306 F. Supp. 3d 579, 592 n.9 (S.D.N.Y. 2018), *aff'd*, 765 F. App'x 586 (2d Cir. 2019), *cert. denied*, 140 S. Ct. 221 (2019).

110. Social media platforms actively represent themselves as products. In a 2018 article, a Facebook Vice President wrote: “Our product is social media – the ability to connect with the people that matter to you, wherever they are in the world.” Rob Goldman, *Hard Questions: What Information Do Facebook Advertisers Know About Me?*, FACEBOOK (Apr. 23, 2018), <https://about.fb.com/news/2018/04/data-and-advertising/> [<https://perma.cc/V3TC-6E5S>]; *see supra* Part I.A.

(a) the component is defective in itself . . . and the defect causes the harm¹¹¹

Therefore, strict liability concepts apply to software when an algorithm is defective. So, the remaining question is whether software is a product under prevailing legal standards.

Looking to the current landscape of strict product liability law, the fundamental task will be showing that products need not be tangible in the traditional sense.¹¹² Embedded within this inquest is an evaluation of whether the item at issue qualifies as (1) a product subject to product liability principles; (2) information, sometimes subject to product liability principles; or, (3) a service, never subject to product liability principles.¹¹³ The most recent version of the Restatement of Torts recognizes that less tangible items may qualify as products: “[a] product is tangible personal property distributed commercially for use or consumption. Other items, such as real property and electricity, are products when the context of their distribution and use is sufficiently analogous to the distribution and use of tangible personal property”¹¹⁴ In isolation, this definition may appear unfavorable,¹¹⁵ but the inclusion of “electricity” within the definition of a product reveals that the meaning of tangible is broad.¹¹⁶ Moreover, the comments accompanying Section 19 further widen this opening.¹¹⁷ One such comment surveys cases and scholarly works considering whether software is a product and directs courts to decisions finding software to be a good in the context of the Uniform Commercial Code (UCC).¹¹⁸

Beyond the Restatement, recent case law exhibits a trend toward recognizing that less traditionally tangible items, including software, can

111. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 5 (AM. L. INST. 1998).

112. Joseph L. Reutiman, Note, *Defective Information: Should Information Be a “Product” Subject to Products Liability Claims?*, 22 CORNELL J.L. PUB. POL’Y 181, 182, 186–88 (2012).

113. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 19(b) (AM. L. INST. 1998) (stating that services do not qualify as a product); Beard, Ford, Koutsky & Spiwak, *supra* note 108, at 207.

114. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 19(a) (AM. L. INST. 1998).

115. David W. Lannetti, *Toward a Revised Definition of “Product” Under the Restatement (Third) of Torts: Products Liability*, 35 TORT & INS. L.J. 845, 884–85 (2000).

116. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 19(a) (AM. L. INST. 2019). Software is, in fact, tangible—even though it is not always visible to the user. *See, e.g.*, Doug Aamoth, *The Internet Weighs About as Much as a Strawberry*, TIME (Nov. 2, 2011), <https://techland.time.com/2011/11/02/the-internet-weighs-about-as-much-as-a-strawberry/> [<https://perma.cc/T3UM-VBA8>].

117. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 19 cmt., at 267–71 (AM. L. INST. 1998).

118. *Id.* § 19 Reporter’s Notes cmt. d, at 278–79.

qualify as products and also explains the critical distinctions among services, information, and products.¹¹⁹ This recognition grew out of two separate lines of cases considering whether certain items— aeronautical charts¹²⁰ and the information in books¹²¹—passed as products. Both began with the Ninth Circuit decision in *Aetna Casualty & Surety Co. v. Jeppesen & Co.*,¹²² where the court determined that the defendant’s depiction of information from the FAA on a “graphic approach chart” was a product for the purposes of strict liability.¹²³ However, the Ninth Circuit did not explain the rationale behind this finding.¹²⁴ As discussed below, other courts considering its holding soon joined the discussion and clarified the lines separating information, services, and products.

In *Saloomey v. Jeppesen & Co.*,¹²⁵ the defendant argued that the defective aeronautical charts were a service and therefore not subject to product liability principles.¹²⁶ The court rejected this contention and found the fact that the defendant’s charts were mass-produced and mass-marketed to be dispositive.¹²⁷ It reasoned that “the mass production and marketing of these charts requires Jeppesen to bear the costs of accidents that are proximately caused by defects in the charts.”¹²⁸ This case suggests a bright-line rule: when an item is mass-produced, it cannot qualify as a service.¹²⁹

Social media platforms are undoubtedly mass-marketed products¹³⁰ because they are created for and distributed to a vast general audience.¹³¹ Case law considering whether software qualifies as a good or as a service

119. Beard, Ford, Koutsky & Spiwak, *supra* note 108, at 207–10; Reutiman, *supra* note 112, at 186–96.

120. *See, e.g., Aetna Cas. & Sur. Co. v. Jeppesen & Co.*, 642 F.2d 339 (9th Cir. 1981).

121. *See, e.g., Winter v. G.P. Putnam’s Sons*, 938 F.2d 1033 (9th Cir. 1991).

122. 642 F.2d 339.

123. *See id.* at 341–42.

124. *See Saloomey v. Jeppesen & Co.*, 707 F.2d 671, 676 (2d Cir. 1983).

125. *See id.*

126. *Id.*

127. *Id.* at 676–77.

128. *Id.* at 677.

129. *See, e.g., Brocklesby v. United States*, 753 F.2d 794, 800 (9th Cir. 1985), *amended by* 767 F.2d 1288 (9th Cir. 1985); *Fluor Corp. v. Jeppesen & Co.*, 216 Cal. Rptr. 68, 70 (Cal. Ct. App. 1985).

130. The Merriam-Webster Dictionary defines mass-market as “appealing or sold to a general audience.” *Mass Market*, MERRIAM-WEBSTER DICTIONARY, <https://www.merriam-webster.com/dictionary/mass-market> [https://perma.cc/NX9D-KNAQ] (last visited Oct. 4, 2020); *see also* Zollers, McMullin, Hurd & Shears, *supra* note 18, at 776–77.

131. *See supra* Part I.B.

under the UCC bolsters this finding.¹³² Cases concerning this question have found that mass-marketed software qualifies as a good under the code, while software specially designed for a particular customer falls into the service category.¹³³ For example, in *Simulados Software, LTD. v. Photon Infotech Private, LTD.*,¹³⁴ a California district court summarized the general trend in this area as follows: “[g]enerally, courts have found that mass-produced, standardized, or generally available software, even with modifications and ancillary services included in the agreement, is a good that is covered by the UCC Even software adapted for specific needs has been considered a good.”¹³⁵ Given the mass-market nature of social media sites¹³⁶ and the Restatement’s suggestion that product liability litigants use this companion context to their advantage, this analogy is highly compelling.¹³⁷ Because software is considered a good within commercial law,¹³⁸ it is dissonant to insist on a different standard under tort law.

Once a litigant establishes that the item at issue is not a service, the next step is to ensure it is not disqualified from strict product liability principles because it is information. Superficially, both aeronautical charts and the information in books, much like algorithms, involve presentations of information. However, courts considering whether information in a book qualified as a product have articulated a dividing line. The most important case in this realm¹³⁹ is *Winter v. G.P. Putnam’s Sons*.¹⁴⁰ In *G.P. Putnam’s Sons*, the plaintiffs fell ill after eating mushrooms they believed to be safe for consumption based on information in a reference book.¹⁴¹ The plaintiffs sued under strict liability and analogized the information contained in the book to an aeronautical chart.¹⁴² The court recognized that his analogy strengthened the plaintiff’s argument but still made a critical distinction.¹⁴³ In comparing the book to an aeronautical chart, it explained:

132. See RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 19 Reporter’s Notes cmt. d, at 278–79 (AM. L. INST. 1998).

133. See *id.*; Reutiman, *supra* note 112, at 195–96.

134. *Simulados Software, Ltd. v. Photon Infotech Priv., Ltd.*, 40 F. Supp. 3d 1191 (N.D. Cal. 2014).

135. *Id.* at 1199–200.

136. See *supra* note 128.

137. RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY §19 Reporter’s Notes cmt. d, at 278 (AM. L. INST. 1998).

138. *Id.* at 278–79.

139. Beard, Ford, Koutsky & Spiwak, *supra* note 108, at 208.

140. *Winter v. G.P. Putnam’s Sons*, 938 F.2d 1033 (9th Cir. 1991).

141. *Id.* at 1034.

142. *Id.* at 1035.

143. *Id.* at 1035–36.

Aeronautical charts are highly technical tools. They are graphic depictions of technical, mechanical data. The best analogy to an aeronautical chart is a compass. Both may be used to guide an individual who is engaged in an activity requiring certain knowledge of natural features. Computer software that fails to yield the result for which it was designed may be another. In contrast, *The Encyclopedia of Mushrooms* is like a book on how to use a compass or an aeronautical chart. The chart itself is like a physical “product” while the “How to Use” book is pure thought and expression.¹⁴⁴

While this discussion is dictum,¹⁴⁵ it implies that when a product involving information is akin to a tool because it serves a specific function and yields a specific result when used in a specific manner, it may be subject to product liability principles. That this case explicitly acknowledged computer software is significant. Both aeronautical charts and computer software are carefully designed to perform a single function, unlike the information in books¹⁴⁶ which individuals can use in various capacities for various different ends.¹⁴⁷ As one scholar stated: “the charts are functional and not literary. There can be no purpose for aeronautical charts other than . . . to aid pilots. Similarly, software has no purpose other than to cause a computer to perform some function.”¹⁴⁸ Academics have also recognized that many of the concerns implicit in extending strict liability to literary information do not arise in the context of software.¹⁴⁹ For instance, the impact on freedom of expression and the First Amendment is slim to none.¹⁵⁰ Moreover, extending strict product liability to this realm will serve fundamental policy goals such as loss spreading and protecting lay victims, who cannot reasonably inspect software for defects¹⁵¹—policies that at least one district court has recognized as important in the context of aeronautical charts.¹⁵²

Scholars initially predicted that *G.P. Putnam’s Sons* would precipitate an increase in software product liability lawsuits, but this surge

144. *Id.* at 1036.

145. *Id.* (declining to expand products liability law to embrace ideas and expressions contained in a book).

146. The book in *G.P. Putnam’s Sons* was a reference book; however, the analysis looks to information in books more generally. *Id.* at 1034, 1036.

147. See Zollers, McMullin, Hurd & Shears, *supra* note 18, at 763.

148. *Id.*

149. *Id.*

150. *Id.*

151. *Id.* at 768–71.

152. See *Fluor Corp.*, 216 Cal. Rptr. 68, 71 (Cal. Ct. App. 1985); Zollers, McMullin, Hurd & Shears, *supra* note 18, at 763.

has yet to occur for a variety of reasons, including the economic loss rule in torts,¹⁵³ as well as perceived difficulties in proving defective design.¹⁵⁴ Still, some recent cases have implicitly regarded software as a product for the purposes of strict product liability, often without undertaking the above analysis.¹⁵⁵ In 2007, the Eastern District of Louisiana made this finding, citing Section 19 of the Third Restatement and *G.P. Putnam's Sons*.¹⁵⁶ More recently, in 2019, the Southern District of New York allowed the parties to agree that concepts of strict liability applied to a dating application which the plaintiff claimed was defectively designed.¹⁵⁷ While courts are in a unique position to make these assumptions without consequence, litigants are not. As a result, the best practice for future litigants is to perform the aforementioned analysis in their pleadings and briefs. They must show three things: (1) the Restatement does not require products to be tangible in the traditional sense, (2) social media software is not a service because it is mass-marketed, and (3) software is not merely information because it operates like a tool that performs a specific function. When litigants do not undertake this analysis, they risk the court making unfavorable assumptions founded on outdated information.¹⁵⁸

153. Beard, Ford, Koutsky & Spiwak, *supra* note 108, at 209; see Daniel T. Perlman, *Who Pays the Price of Computer Software Failure?*, 24 RUTGERS COMPUT. & TECH. L.J. 383, 396 (1998).

154. Beard, Ford, Koutsky & Spiwak, *supra* note 108, at 209–10.

155. See, e.g., *Herrick v. Grindr, L.L.C.*, 306 F. Supp. 3d 579, 592 n.9 (S.D.N.Y. 2018), *aff'd*, 765 F. App'x 586 (2d Cir. 2019), *cert. denied*, 140 S. Ct. 221 (2019) (“It appears to be common ground between the parties that strict products liability may apply to standardized and mass-downloaded software.”); *Schafer v. State Farm Fire & Cas. Co.*, 507 F. Supp. 2d 587, 600–01 (E.D. La. 2007); see also *Williams v. Apple Inc.*, No. H-19-782, 2019 U.S. Dist. LEXIS 78772, at *8–11 (S.D. Tex. May 9, 2019) (accepting, implicitly, iOS 12.1 as a product for the purpose of the defective design theory but dismissing the plaintiff’s claim for their failure to plead a reasonable alternative design); *Hardin v. PDX, Inc.*, 173 Cal. Rptr. 3d 397, 407 (2014) (declining to dismiss a product liability claim where the defendant argued that software was not a product).

156. *Schafer*, 507 F. Supp. 2d at 600–01.

157. *Herrick*, 306 F. Supp. 3d at 592 n.9.

158. For instance, a recent New Jersey case determined that Snapchat did not qualify as a product because the required analysis was not presented to the court. *Grossman v. Rockaway Twp.*, No. MRS-L-1173-18, 2019 LEXIS 1496, at *44 (N.J. Super. June 10, 2019) (“No persuasive or other authority has been presented to this Court to support the conclusion that Snap’s role of involvement in the events of this case constitute a ‘product’ rather than a ‘service.’”). Likewise, in March of 2020, the Third Circuit Court of Appeals issued a short and unreported opinion in which it determined that a Public Safety Assessment algorithm utilized by the New Jersey court system did not qualify as a product. See *Rodgers v. Christie*, 795 F. App'x 878, 879–80 (3d Cir. 2020). It reached this conclusion for two reasons. First, because the algorithm was not commercially distributed, and second because it was intangible. *Id.* While the court was ultimately correct in its conclusion that the tool was not a product because it was neither mass-marketed nor commercially sold, its conclusion regarding traditional tangibility not only lacked support

Legal frameworks are ready for plaintiffs to test the waters in this area. Not only are there sufficient analogies and precedents, like aeronautical charts and electricity,¹⁵⁹ but there is a substantial body of scholarly work advocating for the same result.¹⁶⁰ Consequently, it is time for this virtually unregulated multi-billion-dollar industry to be held accountable for the harm it causes due to defectively designed algorithms.¹⁶¹

B. Proving Defective Design

As with any extension of law to new technology, establishing that personalized recommender systems can satisfy the risk-utility test for defective design presents unique complexities and expenses. The apex of the risk-utility test, the most technical portion of the inquiry, requires proving the existence of a reasonable design.¹⁶² This section describes the central analysis involved and provides a hypothetical application of the risk-utility test to underscore the nuances inherent in these cases.

Under the risk-utility approach, a design is defective when “the foreseeable risks of harm posed by the product could have been reduced or avoided by the adoption of a reasonable alternative design” whose “omission . . . renders the product not reasonably safe.”¹⁶³ Most courts employ this approach because it is easy for jurors and litigants to apply.¹⁶⁴ While the test is often phrased in broad terms and its scope may vary based on the court,¹⁶⁵ it is, at bottom, a cost-benefit analysis under which “a product’s design is defective if the safety benefits of an untaken design

but also overlooked the above analysis, and most significantly, the comments to Section 19 of the Third Restatement. *Id.*

159. See Zollers, McMullin, Hurd & Shears, *supra* note 18, at 774.

160. See *id.* at 755–56, 756 n. 57 (listing twelve periodicals which advocate for applying strict liability to software). Martina Barash, *Snap, Other App, Website Operators Face Product Liability Risks*, BLOOMBERG LAW (Jul. 2, 2018, 3:47 PM), https://www.bloomberglaw.com/document/X4R506V0000000?bna_news_filter=product-liability-and-toxics-law&jcsearch=BNA%252000000164421ed09fa164473ff1fb0000%22%5C1%20%22jcite#jcite [https://perma.cc/QJR5-H2SJ].

161. See Zollers, McMullin, Hurd & Shears, *supra* note 18, at 768–75; Gerrie, *supra* note 91, at 26–28.

162. See RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 2(b) (AM. L. INST. 1998).

163. *Id.*

164. Kristopher-Kent Harris, *Drones: Proposed Standards of Liability*, 35 SANTA CLARA HIGH TECH. L.J. 65, 87–88 (2018).

165. David G. Owen, A Tribute to Professor David Fischer, *Design Defects*, 73 MO. L. REV. 291, 315, 325–26 (2008).

precaution foreseeably exceed its costs.”¹⁶⁶ This test examines various factors like the alternative design’s impact on consumer safety, the cost of incorporating the design, any decrease in the utility of the product, and whether the new design would create additional hazards.¹⁶⁷

To demonstrate how a plaintiff may satisfy the risk-utility test and alternative design inquiry, it is important to review common sources of defects in recommender systems that could give rise to the type of product liability at issue. Arguably the most common causes of defects in this context are “inattentive blindness” and feedback loops, both of which are collateral consequences of the “arms race for attention.”¹⁶⁸ Inattentive blindness occurs when developers design a system to focus on a specific goal, such as time spent or engagement, and overlook equally essential features like user wellbeing.¹⁶⁹ Separately, a feedback loop occurs when an algorithm relies on recommendations it has already made.¹⁷⁰ This results in repeated exposure to a limited set of content, which becomes problematic when that content is extreme.¹⁷¹ Other common defects may stem from the individual biases of designers or from reliance on flawed data or incomplete data sets.¹⁷²

166. John E. Woodard, Comment, *Oops, My GPS Made Me Do It!: GPS Manufacturer Liability Under a Strict Products Liability Paradigm When GPS Fails to Give Accurate Directions to GPS End-Users*, 34 DAYTON L. REV. 429, 454 (2009) (citing DAVID G. OWEN, PRODUCTS LIABILITY LAW 512–14 (W. Acad. 3d ed. 2015)).

167. Owen, *supra* note 165, at 327.

168. See RECOMMENDER SYSTEMS HANDBOOK, *supra* note 26.

169. See WACHTER-BOETTCHER, *supra* note 33, at 94–98.

170. In other words, “where the output of the algorithm becomes part of its input.” Swathi Meenaskhi Sadagopan, *Feedback Loops and Echo Chambers: How Algorithms Amplify Viewpoints*, THE CONVERSATION (Feb. 4, 2019, 4:18 PM), <http://theconversation.com/feedback-loops-and-echo-chambers-how-algorithms-amplify-viewpoints-107935> [https://perma.cc/8LEG-V2KH]; see also Alexis C. Madrigal, *How YouTube’s Algorithm Really Works . . . If You’re Not the Average YouTube User*, THE ATLANTIC (Nov. 8, 2018), <https://www.theatlantic.com/technology/archive/2018/11/how-youtubes-algorithm-really-works/575212/> [https://perma.cc/3KLA-VB2Q].

171. See FALK, *supra* note 25, at 216–17, 246–47. Feedback loops often result in echo chambers and filter bubbles. An echo chamber is defined as the effect of a user’s “interest being positively or negatively reinforced by repeated exposure to a certain item or category of items” and a filter bubble is “the fact that recommender systems select limited content to serve users online.” RAY JIANG, SILVIA CHIAPPA, TOR LATTIMORE, ANDRAS AGYORGY & PUSHMEET KOHLI, DEGENERATE FEEDBACK LOOPS IN RECOMMENDER SYSTEMS (2019).

172. WACHTER-BOETTCHER, *supra* note 33, at 143–46; Iria Giuffrida, Fredric Lederer & Nicolas Vermeys, Tribute to Professor Paul Giannelli, *A Legal Perspective on the Trial and Tribulations of AI: How Artificial Intelligence, the Internet of Things, Smart Contracts, and Other Technologies Will Affect the Law*, 68 CASE W. RES. L. REV. 747, 755, 778–79 (2018). The topic of discriminatory algorithms that reflect the biases of their developers and perpetuate racial and socioeconomic disparity has become popular among scholars in recent years. For a more thorough discussion, see O’NEIL, *supra* note 37, at 85–86, 89–93.

Applying the aforementioned sources of defects to a hypothetical problem illustrates a straightforward application of the risk-utility test. In this hypothetical, the objective function is at the center of both sources. Specifically, inattentive blindness may cause a developer to select improper objectives that overlook user welfare,¹⁷³ while a feedback loop may result from a hyper-personalized objective that exploits individual user level preferences.¹⁷⁴ Both cause the system to recommend increasingly rousing and narrow content in order to keep the user's attention.¹⁷⁵ YouTube is one platform whose algorithms have been criticized for this very result.¹⁷⁶ As one scholar stated: “[YouTube’s] algorithm seems to have concluded that people are drawn to content that is more extreme than what they started with—or to incendiary content in general.”¹⁷⁷

To prevail, a plaintiff making a case against YouTube would need to show that the platform could have avoided the harm at issue by adopting a reasonable alternative design.¹⁷⁸ There are plenty of reasonable alternatives in this area, but a litigant need only show one.¹⁷⁹ One alternative design option is modifying the objective function to both explicitly consider user welfare¹⁸⁰ and mitigate the effects of hyper-personalization through aggregation. Adding penalty terms that protect user welfare to the algorithm's objective function constrains the maximum cost imposed on individual users and tempers the algorithm's output (in this case, the content it recommends) by reducing its expected reward.¹⁸¹

173. See WACHTER-BOETTCHER, *supra* note 33, at 94–98.

174. See DiResta, *supra* note 7 (“The systems don’t actually understand the content, they just return what they predict will keep us clicking. That’s because their primary function is to help achieve one or two specific key performance indicators (KPIs) chosen by the company.”).

175. See *id.*

176. *Id.* Another app that has come under recent scrutiny is TikTok. Commenters have compared the video sharing app’s algorithms to YouTube’s, as both “create echo chambers of content that allow users not only to subconsciously construct their own reality, but . . . also push users toward more radical content.” Caroline Haskins, *TikTok Can’t Save Us from Algorithmic Content Hell*, VICE (Jan. 31, 2019, 9:02 AM), https://www.vice.com/en_us/article/kzdown9/tiktok-cant-save-us-from-algorithmic-content-hell [<https://perma.cc/QUM7-KEG5>].

177. Zeynep Tufekci, *YouTube, the Great Radicalizer*, N.Y. TIMES (Mar. 10, 2018), <https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html> [<https://perma.cc/TW6B-83N8>]; see generally Covington, Adams & Sargin, *supra* note 39.

178. See RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 2(b) (AM. L. INST. 1998).

179. *Id.*

180. See Gerrard & Gillespie, *supra* note 8.

181. See Julia Roberts & Mykel Kochenderfer, *Penalty Functions*, STAN. UNIV. DEP’T OF MATHEMATICS, <https://web.stanford.edu/group/sisl/k12/optimization/MO-unit5->

Sample penalties could include penalties for spending too much time on the platform or for unhealthy engagement, like leaving too many comments in a short period of time.¹⁸² Additionally, requiring a minimum level of aggregation prevents the system from focusing too narrowly on the content most attractive to each individual user and showing that user increasingly extreme versions of the same content.¹⁸³ It is more difficult to consistently serve content in a way that is addictive to a group than it is to an individual.¹⁸⁴ Concretely, the objective function can use a weighted average of user ratings or preferences to diminish the algorithm's effect on any one user.¹⁸⁵

Once the plaintiff establishes that a reasonable alternative exists, the next step is to show that the burdens of the alternative design do not outweigh its safety benefits to users.¹⁸⁶ Changing the objective function by adding penalty terms or by requiring aggregation would protect young and otherwise vulnerable or impressionable users with minimal impairment to YouTube's utility. The algorithms would still carry out their basic functions of recommending non-harmful content when a user's activity is normal.¹⁸⁷ The penalty terms, which are measured at the individual user level, would not come into play unless the user is consuming content in an unhealthy way.¹⁸⁸ Similarly, deploying aggregate group recommenders would allow users to experience the site as normal by receiving recommendations, and the platform would continue to serve its stated purpose to "give everyone a voice" and "show them the world" without

pdfs/5.6penaltyfunctions.pdf [https://perma.cc/R8WQ-9Q3W] (last visited Oct. 10, 2020); Alice E. Smith & David W. Coit, *Constraint-Handling Techniques - Penalty Functions*, in HANDBOOK OF EVOLUTIONARY COMPUTATION 2–4 (1997); STEPHEN BOYD & LIEVEN VANDENBERGHE, CONVEX OPTIMIZATION 215 (7th ed. 2009).

182. For a discussion of other alternatives see JIANG, CHIAPPA, LATTIMORE, AGYORGY & KOHLI, *supra* note 171.

183. See CHARU C. AGGARWAL, RECOMMENDER SYSTEMS 424–25 (2016); AMRA DELIC, JULIA NEIDHARDT, THUY NGOC NGUYEN & FRANCESCO RICCI, RESEARCH METHODS FOR GROUP RECOMMENDER SYSTEMS (2020); Gerrard & Gillespie, *supra* note 8 ("And these algorithms are optimized to serve the individual wants of individual users; it is much more difficult to optimize them for the collective benefit."). This is the barrier to group level recommender systems: why only recommend to groups when you can recommend to each individual using their individual data?

184. See DELIC, NEIDHARDT, NGUYEN & RICCI, *supra* note 183; Gerrard & Gillespie, *supra* note 8.

185. See DELIC, NEIDHARDT, NGUYEN & RICCI, *supra* note 183; Gerrard & Gillespie, *supra* note 8.

186. See RESTATEMENT (THIRD) OF TORTS: PRODUCTS LIABILITY § 2(b) (AM. L. INST. 1998).

187. See Madrigal, *supra* note 170 (acknowledging that the dangers of these algorithms do not impact normal users).

188. See Tufekci, *supra* note 177.

addicting its users.¹⁸⁹ Thus, the alternative design would only impact the user experience when the activity is unhealthy, for example, when a user displays signs of social media addiction¹⁹⁰ or when the content consumed harms welfare. Moreover, users would still be able to search for and locate graphic or extreme content on their own without issue, but the system would not exploit those preferences.¹⁹¹

While the cost of changing the system is the greatest downside of this alternative, it is not prohibitive. Modifying the objective function will lead to costs associated with developer labor, such as costs for changing and maintaining the algorithm,¹⁹² as well as losses in ad revenue due to decreased engagement and time spent on the site. Exact numbers are difficult to quantify and are currently in dispute.¹⁹³ Nevertheless, YouTube itself has stated that the revenues from harmful content are minimal.¹⁹⁴ In a 2019 blog post, the platform wrote that its upcoming plans to modify its algorithms to reduce the recommendation of “borderline” content would only impact one percent of content on the site.¹⁹⁵ Furthermore, platforms frequently tweak their algorithms on their own,¹⁹⁶ and many, including YouTube, have begun to make changes to their platforms that enhance

189. *YouTube About*, YOUTUBE, [https://www.youtube.com/about/\[https://perma.cc/ZBE4-FVQU\]](https://www.youtube.com/about/[https://perma.cc/ZBE4-FVQU]) (last visited Oct. 3, 2020); see DELIC, NEIDHARDT, NGUYEN & RICCI, *supra* note 183; Goldman, *supra* note 110.

190. I-Hua Chen, Carol Strong, Yi-Ching Li, Meng-Che Tsai, Hildie Leung, Chung-Ying Lin, Amir H. Pakpour & Mark D. Griffiths, *Time Invariance of Three Ultra-Brief Internet-Related Instruments: Smartphone Application-Based Addiction Scale (SABAS), Bergen Social Media Addiction Scale (BSMAS), and the Nine-Item Internet Gaming Disorder Scale- Short Form (IGDS-SF9) (Study Part B)*, ADDICTIVE BEHAV., Feb. 2020, at 2 (confirming generalized and specific internet addictions share symptoms of depression, social anxiety, attention deficit/hyperactivity disorder, substance abuse, poor emotional wellbeing, and decreased academic performance, among other comorbidities).

191. YouTube, *Continuing Our Work to Improve Recommendations on YouTube*, YOUTUBE OFFICIAL BLOG (Jan. 25, 2019), <https://blog.youtube/news-and-events/continuing-our-work-to-improve> [<https://perma.cc/LU99-3JU5>] (stating that changes the site plans to make to its algorithm will not change the availability of videos on the site).

192. See RUSSELL JURNERY, *AGILE DATA SCIENCE 2.0: BUILDING FULL-STACK DATA ANALYTICS APPLICATIONS WITH SPARK* 15–24 (Shannon Cutt ed., 2017) (discussing the types of programmer labor involved in building a product).

193. See Steve Kovach, *Alphabet Had More Than \$70 Billion in Market Cap Wiped Out, and It Says YouTube Is One of the Problems*, CNBC (Apr. 30, 2019, 6:41 PM), <https://www.cnbc.com/2019/04/30/youtube-algorithm-changes-negatively-impact-google-ad-revenue.html> [<https://perma.cc/GE65-KGHJ>] (“There’s a misconception that YouTube makes money off of recommending ‘radical’ content, but the truth is that very little of this content makes any kind of meaningful money. In fact, when we cleaned up our partner program to remove bad actors last year, we made it clear that 99% of those impacted creators were making less than \$100 a year.”).

194. *Id.*

195. YouTube, *supra* note 191.

196. See, e.g., Alter, *supra* note 55, at 5.

user health and safety.¹⁹⁷ These alterations may even help these sites win the confidence and loyalty of users in the long run. Even at their most expensive, the costs pale in comparison to the billions of dollars of profit that these platforms generate, and as a result, are unlikely to be detrimental.¹⁹⁸ This solution is also unique in that it imposes no cost on consumers but instead spreads costs among advertisers.¹⁹⁹ Finally, the alternative design would not involve new risks to safety because the sole risk to safety in this context is exactly what the alternative design seeks to mitigate. In sum, the above cost-benefit analysis illustrates that YouTube can eliminate the risk of a defective design with a nominal impact on the product's use and cost but a major improvement in safety. As a result, the risk-utility test is satisfied, leaving the plaintiff with one remaining challenge to confront: the Communications Decency Act.

C. Overcoming the Communications Decency Act

To maintain a successful lawsuit, a plaintiff must show that the defendant is not immune from liability under the CDA. The following two sections establish that a plaintiff can accomplish this in two ways in a defective design case: (1) the plaintiff can show that social media platforms *do not* act in the judicially defined publishing capacity when they employ personalized algorithms to recommend content to users; or, (2) that they *do* act as information content providers when they design their systems to addict users. Both these inquiries will be fact-intensive and conduct-specific.²⁰⁰ As a result, each will require precise pleading to encourage reviewing courts to reevaluate what is often a blanket application of immunity.²⁰¹

197. Kevin Murnane, *How Much Time Do You Spend Watching YouTube? A New Digital Wellbeing Profile Let's You Know*, FORBES (Aug. 27, 2018, 11:59 AM), <https://www.forbes.com/sites/kevinmurnane/2018/08/27/how-much-time-do-you-spend-watching-youtube-a-new-digital-wellbeing-profile-lets-you-know/> [https://perma.cc/FLG6-HE72]; Kovach, *supra* note 193.

198. See Alphabet Inc., Annual Report (Form 10-K) 29 (Feb. 3, 2020); Todd Spangler, *Alphabet Reports YouTube Ad Revenue for First Time, Video Service Generated \$ 15.1 Billion in 2019*, VARIETY (Feb. 3, 2020, 1:15 PM), <https://variety.com/2020/digital/news/alphabet-youtube-ad-revenue-first-time-15-billion-2019-1203491155/> [https://perma.cc/ZN7A-5Z76].

199. Cf. Joseph H. King Jr., *A Goals-Oriented Approach to Strict Tort Liability for Abnormally Dangerous Activities*, 48 BAYLOR L. REV. 341, 350–52 (1996) (discussing the loss-spreading function of strict liability).

200. See Tremble, *supra* note 96, at 856–59.

201. See Seaton, *supra* note 103, at 357; 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, 388–89 (2011) (discussing the need for a specific pleading and procedural approach to ensure courts

1. SOCIAL MEDIA PLATFORMS DO NOT ACT IN A PUBLISHING CAPACITY WHEN THEY RECOMMEND PERSONALIZED CONTENT

The statutory text of Section 230 does not define “publisher” and thus judicial interpretations control.²⁰² The first major case to interpret the meaning of the word was *Zeran v. America Online, Inc.*,²⁰³ a case involving defamatory content posted on an AOL message board by a third party.²⁰⁴ The court looked to the plain meaning of publisher, finding that “lawsuits seeking to hold a service provider liable for its exercise of a publisher’s traditional editorial functions—such as deciding whether to publish, withdraw, postpone or alter content—are barred.”²⁰⁵ It also determined that distributors, who disseminate information written by others, are considered publishers.²⁰⁶ This holding proved to be influential among federal appellate courts: many adopted and adapted this definition and test.²⁰⁷ Some courts even simplified the inquiry further, like the Ninth Circuit in *Fair Housing Council v. Roommates.com*,²⁰⁸ which held that “any activity that can be boiled down to deciding whether to exclude material that third parties seek to post online is perforce immune under section 230.”²⁰⁹ A few courts have expanded the *Zeran* publisher test like the First Circuit in *Doe v. Backpage.com, L.L.C.*²¹⁰ The *Backpage.com* court called the term “capacious” and found that decisions about a website’s design as well as its “practices and policies . . . reflected choice[s] about what content can appear on the website and in what form” and thus “[were] editorial choices . . . within the purview of traditional publisher functions.”²¹¹ Therefore, different circuit courts interpret the term with varying degrees of latitude, which may beget different outcomes

address Section 230 immunity and establish a more fulsome record to make it more likely that the *Zeran* test can be conducted fully).

202. 47 U.S.C. § 230(f); see, e.g., *HomeAway.com, Inc. v. City of Santa Monica*, 918 F.3d 676, 681 (9th Cir. 2019).

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

204. *Id.* at 329; Caitlin McKeown, Comment, *Facebook, Defamation, and Terrorism: Who Is Responsible for Dangerous Posts on Social Media?*, 26 TUL. J. INT’L & COMPAR. L. 163, 169 (2017).

205. *Zeran*, 129 F.3d at 330.

206. *Id.* at 332.

207. Tremble, *supra* note 96, at 842–44.

208. 521 F.3d 1157 (9th Cir. 2008).

209. *Id.* at 1170–71.

210. 817 F.3d 12, 18 (1st Cir. 2016). The First Circuit decided *Backpage.com* before FOSTA-SESTA, which would have allowed the plaintiffs to proceed with their claim. See Tripp, *supra* note 107, at 221–22. While the outcome of the case would be different today, its holdings regarding the CDA remain intact. See *id.* at 231–33.

211. *Backpage.com*, 817 F.3d at 19–21. *Backpage.com* has had a strong influence in promulgating an overly broad interpretation of the CDA. Barash, *supra* note 160.

for litigants in different circuits. For instance, a litigant is more likely to succeed in the Ninth Circuit, which subscribes to a more precise and limited view of the publisher's role than the First Circuit, whose "capacious" construction threatens to swallow important distinctions that separate recommendation from publication.²¹²

Barnes v. Yahoo!, Inc.,²¹³ is arguably the most widely cited case in this area. There, the court clarified that the publisher inquiry extends to causes of action beyond the defamation context.²¹⁴ Consequently, it held that "what matters" most in determining whether the ISP acted as a publisher is "whether the cause of action inherently requires the court to treat the defendant as the 'publisher or speaker' of content provided by another."²¹⁵ In other words, "courts 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.'"²¹⁶ In its assessment of what sorts of status or conduct will qualify an ISP as a publisher, the court retained a *Zeran*-style plain meaning approach.²¹⁷ It looked to the plain meaning of the term as one who "reviews material submitted for publication, perhaps edits it for style or technical fluency, and then decides whether to publish it,"²¹⁸ and cited the *Roommates* test with approval.²¹⁹ Moreover, the court went on to state that there was no need to perform "intellectual gymnastics" in ascertaining the meaning of the term.²²⁰ The fact that this traditional approach appears to be the majority view will make a litigant's job easier in most jurisdictions.

Comparing the role of social media recommendation algorithms against both the *Barnes* and *Backpage* approaches demonstrates that platforms lie outside the realm of publisher immunity no matter what test a court applies. The *Barnes* inquiry requires the court to closely examine the *elements of the cause of action*, and most importantly, *the offending conduct*.²²¹ Even a cursory review of these factors is decisive. As an initial

212. Compare *Roommates.com*, 521 F.3d 1157 (adopting a narrow view of immunity), with *Backpage.com*, 517 F.3d 12 (adopting a "capacious" view of immunity). See Mark D. Quist, Comment, "Plumbing the Depths" of the CDA: Weighing the Competing Fourth and Seventh Circuit Standards of ISP Immunity Under Section 230 of the Communications Decency Act, 20 GEO. MASON L. REV. 275, 282 (2012) (providing a comprehensive discussion of the different ways in which circuit courts interpret the scope of CDA immunity).

213. 570 F.3d 1096 (9th Cir. 2009).

214. *Id.* at 1101–02.

215. *Id.* at 1102.

216. *Id.*

217. *Id.*

218. *Id.*

219. *Id.*

220. *Id.*

221. *Id.* at 1101–02, 1107–09; Tremble, *supra* note 96, at 860.

matter, in the case of strict product liability, there is no duty element for the plaintiff to establish; the plaintiff need only prove the defect, causation, and injury.²²² So formalistically, there is no duty that stems from the ISP's role as a publisher. More substantively, however, liability does not arise from the posting, editing, or withdrawal of any specific piece of content.²²³ While a site no doubt publishes the content that it recommends, the harm stems from the platform's use of artificial intelligence and reinforcement learning to personalize user content.²²⁴ This function simply cannot be distilled to a determination of whether or not to exclude or filter content,²²⁵ and under the *Backpage* conception of liability, it is not traceable to an editorial choice about content a webpage may contain.²²⁶ It follows that social media platforms would fail the publisher test even under its broadest conception.

An evaluation of recent cases involving social media platforms elucidates what it takes for a plaintiff to succeed. Cases will fail when litigants cannot separate the source of the harm from choices that the platform made about what content can appear on the site. *Herrick v. Grindr, L.L.C.*,²²⁷ a recent Second Circuit case, stands as a useful example.²²⁸ In *Herrick*, the plaintiff argued that a dating application was defectively designed because it lacked certain safety features that failed to detect and remove false profiles created by a vengeful ex-lover.²²⁹ The court held that, at bottom, the “claims [sought] to hold Grindr liable for its failure to combat or remove offensive third party content” and, as a result, were “barred by Section 230.”²³⁰ So, *Herrick*'s claim failed for two key reasons. First, the claim was fundamentally tied to the posts of a specific user, thus the harm could not have occurred but for Grindr's publication of those posts. And second, though the plaintiff alleged a defective design claim, that claim was still premised on the dating app's failure to filter out content—an essential function of publishing.

Conversely, plaintiffs have overcome publisher immunity when they do not tie their claims to specific harmful content. In *Maynard v.*

222. See Kessler, *supra* note 77, at 445–47 (explaining the removal of the duty element in strict products liability).

223. See Tremble, *supra* note 96, at 828–29 (discussing how ISPs avoid liability for actions a traditional publisher cannot).

224. See *supra* notes 27–32 and accompanying text.

225. *Fair Hous. Council v. Roommates.com, L.L.C.*, 521 F.3d 1157, 1169–70 (9th Cir. 2008) (discussing the “direct and palpable” connection between Roommates.com and their discriminatory filtering process).

226. *Doe v. Backpage.com, L.L.C.*, 817 F.3d 12, 19–21 (1st Cir. 2016).

227. 765 F. App'x 586 (2d Cir. 2019), *cert. denied*, 140 S. Ct. 221 (2019).

228. See *id.* at 589–91.

229. See *id.* at 588.

230. *Id.* at 590; see also 47 U.S.C. § 230(c)(2)(B) (2018).

Snapchat,²³¹ plaintiffs sued under the theory of “negligent creation, design, and maintenance,”²³² and alleged that Snapchat negligently designed its speedometer filter feature to “encourage[] . . . excessive speeding” that, in that case, resulted in a devastating car crash.²³³ The plaintiffs argued that CDA immunity did not apply because they did not seek to hold Snapchat responsible as a publisher of third-party content, but rather for their negligent product design.²³⁴ The court agreed, and in distinguishing the case from its predecessors, it stated:

In *Barnes, Fields, and Backpage.com*, the publication of a third-party user’s posts caused the harm, and the claims depended on the content of the posts themselves. In the instant case, on the other hand, there was no third-party content uploaded to Snapchat at the time of the accident and the Maynards do not seek to hold Snapchat liable for publishing a Snap by a third-party that utilized the Speed Filter. Rather, the Maynards seek to hold Snapchat liable for its *own conduct*, principally for the creation of the Speed Filter and its failure to warn users that the Speed Filter could encourage speeding and unsafe driving practices.²³⁵

Snapchat, much like other social media platforms, is primarily designed for users to post and share content.²³⁶ But the court in *Maynard* recognized that this fact did not envelop all decisions made by the platform within the scope of the publishing function. Instead, the court acknowledged that a claim based on product defects was decisively outside the bounds of publisher immunity because the harm stemmed from a completely distinct source and because it was perpetuated by the platform alone.²³⁷ *Maynard* informs that the key to overcoming publisher

231. 816 S.E.2d 77 (Ga. Ct. App. 2018).

232. *Id.* at 80. While this case did not allege a strict product liability cause of action, negligent design falls under the broader umbrella of product liability and thus, it will serve as a useful analog to plaintiffs in the strict liability context.

233. *Id.* at 79. This filter tracks how fast a car is driving and superimposes this information onto the user’s photos and videos. *Id.*

234. *Id.* at 80.

235. *Id.* at 81 (emphasis added) (citations omitted).

236. Snap Inc., Annual Report (Form 10-K) 3 (Feb. 4, 2020) (Snapchat “is a camera application that helps people communicate visually with friends and family through short videos and images.”).

237. *Maynard*, 816 S.E.2d at 81.

immunity lies in a careful evaluation of the source of harm.²³⁸ Thus, it reinforces the importance of the *Barnes* test.²³⁹

The case of defective recommendation algorithms is more akin to *Maynard* and does not suffer from the issues that pervaded in *Herrick*. As in *Maynard*, the liability in the case of defective recommendation algorithms stems from an engineering decision and not from decisions tied to a social media platform's "status or conduct as a 'publisher or speaker.'"²⁴⁰ Specifically, the liability arises from hyper-personalized recommendation algorithms that rely on misguided objectives, predicated on optimizing aggregate platform value independent of individual user costs.²⁴¹ These algorithms do so by employing knowledge about individual users to dictate the unique sequence and aggregation of content each user experiences.²⁴² It is this tailoring that harms the user: by commandeering their attention and saturating them with the exact content that damages their mental health.²⁴³ In sum, there would be no harm but for the decision to exploit individual user data for personalization. Therefore, unlike in *Herrick*, *Backpage*, and *Barnes*, the cause of action does not arise solely from content posted on the site.²⁴⁴

When courts elide this essential distinction about the source of liability, the consequences for plaintiffs are costly. A recent case illustrates these consequences. In the summer of 2019, the Second Circuit decided *Force v. Facebook, Inc.*,²⁴⁵ one of the first decisions in the country to extensively consider recommender algorithms in the context of CDA publisher immunity.²⁴⁶ The plaintiffs here sued under the Anti-Terrorism

238. *Id.* (distinguishing whether the source of the harm is the publisher itself or a third party).

239. *See id.* at 80 (quoting *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1101 (9th Cir. 2009)).

240. *Barnes*, 570 F.3d at 1101–02; *see Force v. Facebook, Inc.*, 934 F.3d 53, 82–83 (Katzmann, C.J., concurring in part, dissenting in part).

241. *See DiResta*, *supra* note 7; Mark Bergen, *YouTube Executives Ignored Warnings, Letting Toxic Videos Run Rampant*, BLOOMBERG (Apr. 2, 2019, 10:29 AM), <https://www.bloomberg.com/news/features/2019-04-02/youtube-executives-ignored-warnings-letting-toxic-videos-run-rampant> [<https://perma.cc/AM3Q-F4PE>] (addressing YouTube's entrenched pursuit of profits at the cost of users, how recommendation algorithms exacerbate the problem, and a legal strategy to avoid accountability).

242. DiResta, *supra* note 7.

243. Recall the specifically tailored suicide-related content that Pinterest sent to Molly Russell's email. *See Gerrard & Gillespie*, *supra* note 8; *Ridler*, *supra* note 10; *see also supra* Part II.B (discussing the hyper-personalized objective functions).

244. This appears to be the most common reason that CDA cases fail. For examples of similar cases, *see Dyroff v. Ultimate Software Grp., Inc.*, 934 F.3d 1093, 1097–98 (9th Cir. 2019); *Doe v. Backpage.com, L.L.C.*, 817 F.3d 12, 22 (1st Cir. 2016).

245. 934 F.3d 53 (2d Cir. 2019).

246. *Id.* at 58–59, 65–71; *see also Gonzalez v. Google, Inc.*, 335 F. Supp. 3d 1156, 1171, 1179 (N.D. Cal. 2018).

Act, alleging that Facebook’s use of algorithms to show targeted content, advertisements, and friend suggestions enabled the terrorist organization Hamas to spread its message and “carry out the essential communication components of [its] terror attacks.”²⁴⁷ But the majority opinion failed to seriously consider the plaintiff’s arguments about “matchmaking,” and in doing so, it oversimplified the nature of the algorithms at issue.²⁴⁸ As it declared:

[P]laintiffs’ argument that Facebook’s algorithms uniquely form “connections” or “matchmake” is wrong. That, again, has been a fundamental result of publishing third-party content on the Internet since its beginning . . . Facebook’s algorithms might cause more such “matches” than other editorial decisions. But that is not a basis to exclude the use of algorithms from the scope of what it means to be a “publisher” under Section 230(c)(1). The matches also might—as compared to those resulting from other editorial decisions—present users with targeted content of even more interest to them, just as an English speaker, for example, may be best matched with English-language content.²⁴⁹

Unsurprisingly, this decision was accompanied by a partial dissent from Chief Judge Katzmann.²⁵⁰ He found that the majority had “strain[ed] the English language” in concluding that Facebook acted as a publisher of third-party information in utilizing its machine learning algorithms.²⁵¹ Judge Katzmann returned to the traditional understanding of the publishing function and, like the court in *Maynard*, recognized that the CDA does not endow providers with immunity for all activities in which they engage.²⁵² Therefore, he reasoned:

247. *Force*, 934 F.3d at 59. Over the past two to three years, there have been several lawsuits like *Force* attempting to sue social media platforms under the Anti-Terrorism Act for their material contributions, including their use of recommendation algorithms, to specific acts of terror. *See, e.g., Gonzalez*, 335 F. Supp. 3d at 1160, 1162; *Taamneh v. Twitter, Inc.*, 343 F. Supp. 3d 904, 906, 908 (N.D. Cal. 2018). However, the vast majority of these suits have failed due to issues of attenuated causation and because the source of the harm in these cases is more akin to cases like *Herrick*, *Backpage.com*, and *Barnes*. *See Gonzalez*, 335 F. Supp. 3d at 1174–75, 1178; *Taamneh*, 343 F. Supp. 3d at 913, 918; *see also supra* note 76 and accompanying text (noting the problems of proof associated with negligence claims).

248. *Force*, 934 F.3d at 65–67.

249. *Id.* at 67.

250. *Id.* at 76 (Katzmann, C.J., concurring in part, dissenting in part).

251. *See id.* at 76–77 (Katzmann, C.J., concurring in part, dissenting in part).

252. *Id.* at 81 (Katzmann, C.J., concurring in part, dissenting in part).

The fact that Facebook also publishes third-party content should not cause us to conflate its two separate roles with respect to its users and their information. Facebook may be immune under the CDA from plaintiffs' challenge to its allowance of Hamas accounts, since Facebook acts solely as the publisher of the Hamas users' content. That does not mean, though, that it is also immune when it conducts statistical analyses of that information and delivers a message based on those analyses.²⁵³

Ultimately, the dissent represents the most logical way to conceive of personalized recommender algorithms within the framework of the CDA—as performing a function that is severable from the function of publication. Both systems form connections and make matches, but the difference lies in *why* and *how* they accomplish this. A traditional publisher generally seeks to “make [information] generally known” or to “disseminate to the public,”²⁵⁴ which is why a newspaper looks the same to all who read it. On the other hand, no two social media users experience identical sets of content because the platform's recommender algorithms strive to tailor individual user experiences via artificial intelligence and reinforcement learning.²⁵⁵ These algorithmic choices are far more than mere publishing decisions: they are extraordinarily calculated business decisions that leverage individual user data to exert a direct effect on site value.²⁵⁶ This distinction transcends the defective-design context and can apply to any cause of action where the litigant alleges that the platform's algorithms caused harm. In the words of Judge Katzmann, “[w]hen a plaintiff brings a claim that is based not on the content of the information shown but rather on the connections [a platform's] algorithms make . . . the CDA does not and should not bar relief.”²⁵⁷

253. *Id.* at 83 (Katzmann, C.J., concurring in part, dissenting in part).

254. *Publish*, MERRIAM-WEBSTER DICTIONARY, <https://www.merriam-webster.com/dictionary/publish> [<https://perma.cc/G2SQ-WB5B>] (last updated Sept. 16, 2020); *see also Force*, 934 F.3d at 65 (providing several definitions for “publisher”).

255. *See supra* notes 25–33 and accompanying text.

256. *See Price*, *supra* note 27.

257. *Force*, 934 F.3d at 77 (Katzmann, C.J., concurring in part, dissenting in part). Under this theory, Judge Katzmann recognized “the fact that publishers may sell advertising based on user data does not immunize the publisher if someone brings a claim based on the publisher's selling of the data, because the claim would not treat the defendant as the publisher's third party's content.” *Id.* at 81 (Katzmann, C.J., concurring in part, dissenting in part).

2. THE USE OF RECOMMENDATION ALGORITHMS CONVERTS A PLATFORM INTO AN INFORMATION CONTENT PROVIDER

The second way for litigants to surmount the CDA is to prove that social media platforms become content providers when their engineering efforts are defectively designed for addiction and therefore fall outside the scope of Section 230 immunity. While the notion that social media platforms are information content providers may appear dubious at first glance, a close look at the language of the statute as well as the prevailing test in the federal appellate courts suggests that this argument is, in fact, a viable option. Section 230 defines an information content provider as “any person or entity that is responsible, in whole or in part, for the creation or *development* of information provided through the Internet or any other interactive computer service.”²⁵⁸ The statute does not define what might constitute the creation or development of information, and unsurprisingly, courts have struggled to do the same.²⁵⁹ Nevertheless, courts recognize that creation and development are distinct sources of liability.²⁶⁰ Case law has generally found that a platform becomes a content creator when the “content at issue is a direct result of the conduct of the ISP.”²⁶¹ The concept of development is more nuanced and accordingly receives more judicial attention.²⁶² This Comment focuses on the content development test because it is more readily satisfied in the context of a defective design case than is the content creation test.²⁶³

The majority of federal appellate courts subscribe to an assessment of material contribution to ascertain whether ISPs qualify as content developers.²⁶⁴ The Ninth Circuit created this test in its *Roommates* decision, which interpreted the term “development” to “refer[] not merely to augmenting the content generally, but to materially contributing to its

258. 47 U.S.C. § 230(f)(3) (emphasis added).

259. See, e.g., *Kimzey v. Yelp! Inc.*, 836 F.3d 1263, 1269 (9th Cir. 2016) (“The meanings of the words ‘creation’ and ‘development’ are hardly self-evident in the online world, and our cases have struggled with determining their scope.”); see also Samuel J. Morley, *How Broad is Web Publisher Immunity Under § 230 of the Communications Decency Act of 1996?*, FLA. BAR J., Feb. 2010, at 9, 10 (discussing varying court understandings of whether websites are publishers).

260. *Fair Hous. Council v. Roommates.com, L.L.C.*, 521 F.3d 1157, 1168 (9th Cir. 2008).

261. See Tremble, *supra* note 96, at 850.

262. See *id.* at 860–61.

263. In the context of an information content provider, the content itself is not a direct result of the ISP’s conduct. The platform is not creating the underlying content but is manipulating it to achieve a specific result. See *id.* at 850; *supra* Part I.A.

264. See, e.g., *Kimzey*, 836 F.3d at 1269; *FTC v. Leadclick Media, L.L.C.*, 838 F.3d 158, 174, 176 (2d Cir. 2016); *Jones v. Dirty World Ent. Recordings L.L.C.*, 755 F.3d 398, 413 (6th Cir. 2014).

alleged unlawfulness.”²⁶⁵ The *Roommates* court found that the ISP acted as a developer when it sought out discriminatory content and used it to illegally further its business endeavors.²⁶⁶ *Roommates* did far more than provide a “framework” or “neutral tools” with which illegal content could be published.²⁶⁷ Instead, the website’s “work in developing the discriminatory questions, discriminatory answers and discriminatory search mechanism [was] directly related to the alleged illegality of the site.”²⁶⁸ The *Roommates* court predicted its holding would engender confusion, and so it provided several examples of what does and does not qualify as development.²⁶⁹ For example, where an ISP provides “neutral tools,” such as providing a search engine or employing drop-down menus that a third party then uses for insidious purposes, the company’s immunity remains intact.²⁷⁰ But if a provider alters or edits the content—for example, by removing certain words to create a false statement—it has materially contributed to tortious libel and is no longer within Section 230 immunity.²⁷¹

Other courts have added important insights to the *Roommates* analysis. For instance, in *Jones v. Dirty World Entertainment Recordings L.L.C.*,²⁷² the Sixth Circuit clarified the distinction between publication and content development.²⁷³ The critical difference is that actions of publishers “are necessary to the display of unwelcome . . . content,” whereas actions of developers are “[responsible] for what makes the displayed content illegal or actionable.”²⁷⁴ In order to qualify as a content developer, a website must do more than simply “display[]” or “allow[] access” to third-party content.²⁷⁵ Therefore, an ISP can simultaneously be immune from liability for its publishing activities but subject to liability for its role in content development.²⁷⁶ While both the publisher and the

265. *Roommates.com*, 521 F.3d at 1167–68.

266. *See id.* at 1165–73.

267. *Id.* at 1172.

268. *Id.*

269. *Id.* at 1169.

270. *Id.* (emphasis omitted). Note that some scholars have criticized this holding as “exploit[ing] the ambiguity within the language of § 230 to provide a malleable standard for content ‘development,’” which could lead courts to narrow the scope of CDA Immunity. *See* Eric Weslander, Comment, *Murky “Development”: How the Ninth Circuit Exposed Ambiguity Within the Communications Decency Act, and Why Internet Publishers Should Worry*, 48 WASHBURN L.J. 267, 269 (2008). While the decision may be unfavorable to ISPs, it provides victims with important latitude to initiate lawsuits. *See id.*

271. *Roommates.com*, 521 F.3d at 1169.

272. 755 F.3d 398 (6th Cir. 2014).

273. *Id.* at 410.

274. *Id.* at 414.

275. *Id.* at 410.

276. *See, e.g.,* Tremble, *supra* note 96, at 859.

developer tests involve similar considerations and similar activity, particularly in the context of a defective-design case, litigants must understand the core differences between the inquiries as they apply to social media algorithms. The former looks to whether the conduct involved in publication (editing, posting, withdrawing, etc.) generated the harm to the plaintiff, while the latter looks to whether the activity of the ISP was responsible for what made the content harmful.²⁷⁷

While not formally memorialized in the material-development test, whether an ISP develops harmful content for profit has been an important consideration in the material-contribution inquiry.²⁷⁸ In *FTC v. Accusearch*,²⁷⁹ the Tenth Circuit held that Accusearch qualified as an information content developer when it sought out legally protected personal data to post and sell on its site in violation of the Federal Trade Commission Act.²⁸⁰ The court found that Accusearch “contributed mightily to the unlawful conduct” at issue because the harmful posts were its “*raison d’etre* and it affirmatively solicited them.”²⁸¹

The *Roommates* decision provides a useful analog for examining the material-contribution inquiry. Again, in *Roommates*, the court found that the ISP was a content developer when it required users to answer discriminatory questions to access the site and then used these questions to prepopulate a profile page and to influence its search and email systems.²⁸² In comparing the facts of the case to a prior Ninth Circuit decision that declined to find that the ISP acted as a developer, the court stated:

The salient fact in *Carafano* was that the website’s classifications of user characteristics did absolutely nothing to enhance the defamatory sting of the message, to encourage defamation or to make defamation easier By sharp contrast, Roommate’s website is designed to force subscribers to divulge protected characteristics and discriminatory preferences, and to match those who have rooms with those who are looking for rooms based on criteria that appear to be prohibited by the FHA.²⁸³

277. See, e.g., *Jones*, 755 F.3d at 414.

278. See, e.g., *FTC v. Accusearch Inc.*, 570 F.3d 1187, 1200 (10th Cir. 2009).

279. 570 F.3d 1187, 1190.

280. *Id.* at 1191, 1201.

281. *Id.* at 1200.

282. *Fair Hous. Council v. Roommates.com, L.L.C.*, 521 F.3d 1157, 1166 (9th Cir. 2008).

283. *Id.* at 1172.

Therefore, it was the design of the website, which was built to gain and then use discriminatory information for an illegal purpose, that rendered Roommates.com a content developer.²⁸⁴

Much like *Roommates*, social media platforms become developers through their design choices. The material contribution to what makes the content tortious emerges from defectively designed algorithms that are engineered to addict,²⁸⁵ for instance, via the improper and hyperpersonalized objective functions discussed in Section II.A.2.²⁸⁶ These defective-design decisions strive to keep users on the site where they continue to view advertisements and make money for the platform.²⁸⁷ What makes the content harmful lies in the way it is shown to the user and not in qualities inherent in the posts themselves, none of which would be tortious in isolation. And as in *Roommates* and *Accusearch*, the motivations behind these choices are financial,²⁸⁸ a fact that serves to strengthen the relationship between the ISP and what makes the content harmful.²⁸⁹

There is one more distinction to draw in the defective-design context. Unlike in *Accusearch* and *Roommates*, it is the conduct of the platform

284. *Id.* at 1167. The court found it important that the site “[made] aggressive use” of the discriminatory information it sought out “in conducting its business.” *Id.* at 1172; see also Marc van Bree, *Pre-Populating Forms to Increase Conversions*, ARTS HACKER (Dec. 27, 2016), <https://artshacker.com/pre-populating-forms-to-increase-conversions/> [<https://perma.cc/K6VS-VYRH>] (explaining the practice of prepopulating forms to increase conversion rates).

285. It is important to note that at least one recent decision has found recommender algorithms to be neutral tools absent the context of engineered addiction. In *Dyroff v. Ultimate Software Group, Inc.*, the plaintiff sued a now-defunct social media site alleging that their algorithms “steered” drug users to “groups dedicated to the sale and use of narcotics” and “sent users alerts to posts within groups . . . dedicated to the sale and use of narcotics.” *Id.* at 1095. In this case, the plaintiff’s son became a member of one such group and initiated a post seeking heroin. *Id.* at 1095. Unfortunately, the heroin he purchased through the group was laced with fentanyl and resulted in his fatal overdose. *Id.* In finding the platform’s algorithms to be neutral tools, the court focused on the specific content at issue—the deceased’s post about obtaining narcotics. *Id.* at 1096, 1098. In doing so, it determined that the language of the post alone was responsible for what made the content harmful and for what ultimately caused the harm. *Id.* at 1098–99. The algorithms used by the site played no role here and thus could not satisfy the test of material contribution. *Id.* In fact, the court specifically noted that the plaintiff failed to plead that the platform “made suggestions regarding the content of potential user posts.” *Id.* at 1099. Therefore, it considered the algorithms used to be content-neutral. *Id.* at 1096. Yet, this begs the question; how can algorithms be content neutral if the decisions rendered by them depend on the content itself? Hypothetically, had the recommendation algorithm caused the deceased’s drug addiction over time through the recommendation of drug-related content, then the outcome of the case may have been different.

286. See generally *supra* text accompanying notes 168–177.

287. See *supra* text accompanying notes 27–29, 168–177.

288. See *supra* text accompanying notes 27–29.

289. See *supra* text accompanying notes 281, 284.

alone that generates the harm. In both of those cases, a third party knowingly aided the ISPs in bringing about the harm: that party bought the protected information in *Accusearch*²⁹⁰ and filled out the discriminatory forms in *Roommates*.²⁹¹ But here, the third parties providing the content and the third parties using the site are likely unaware of the site's efforts to manufacture addiction.²⁹² Not only does this fact make it all the more difficult to view defective recommender algorithms as simply "neutral tools,"²⁹³ it also simplifies the inquiry. In many other lawsuits, untangling the extent of an ISP's participation in developing the content is a complex endeavor that can vary greatly depending on the facts of the case and the relationships among the platform, content, users, and injured parties.²⁹⁴ This is not so in this defective design context. Rather, applying the material contribution test is far more direct, and litigants should capitalize on this unique benefit, which will allow them to build forcible cases for overcoming the CDA.

All told, litigants victimized by the "arms race for attention"²⁹⁵ have two workable options for defeating CDA immunity, which they should be sure to plead in the alternative. First, they may show that platforms fall outside the realm of publisher immunity by arguing it is the function of personalized recommendation that spawns harm to the plaintiff, not a decision about whether or not to publish harmful content in general. And second, litigants may establish that social media platforms become information content developers subject to liability through their AI design choices that strive to manufacture addiction for financial gain.

CONCLUSION

Beyond its technical arguments about how litigants can construct a successful case, this Comment aims to serve as a call to action for future

290. See *FTC v. Accusearch Inc.*, 570 F.3d 1187, 1191–92 (10th Cir. 2009).

291. *Fair Hous. Council v. Roommates.com*, 521 F.3d at 1161–62, 1165 (9th Cir. 2008) ("Subscribers who are seeking housing must make a selection from a drop-down menu, again provided by Roommate, to indicate whether they are willing to live with 'Straight or gay' males, only with 'Straight' males, only with 'Gay' males or with 'No males.'").

292. See Gerrard & Gillespie, *supra* note 8; Morgans, *supra* note 54; John Gramlich, *10 Facts About Americans and Facebook*, PEW RSCH. CTR. (May 16, 2019), <https://www.pewresearch.org/fact-tank/2019/05/16/facts-about-americans-and-facebook/> [<https://perma.cc/M88L-9BA2>] ("Many adult Facebook users in the U.S. lack a clear understanding of how the platform's news feed works.").

293. See *Roommates.com*, 521 F.3d at 1169 (emphasis omitted).

294. See Tremble, *supra* note 96, at 861 (suggesting that courts should examine the relationship between the ISP and the user as well as the relationship between the ISP and the content).

295. See Price, *supra* note 27.

plaintiffs. Awareness of the negative consequences of social media use—particularly problematic social media use²⁹⁶—has come to the forefront in the past several years, and consequently, the opportunity for litigants to maintain a successful suit has reached new potential.²⁹⁷ As this Comment demonstrates, current legal frameworks are sufficiently developed for plaintiffs to bring their cases to court and overcome the greatest obstacles to success. It illustrates that product liability concepts apply to less tangible products like software²⁹⁸ and exhibits through its hypothetical application of the risk-utility test that proving a reasonable alternative design is straightforward.²⁹⁹ Most importantly, this Comment proves that CDA case law is favorable enough to provide plaintiffs with two different ways to circumvent Section 230's broad grant of immunity by (1) arguing that the AI-driven function of recommending hyperpersonalized content is different in kind and severable from the publishing-related tasks of filtering, editing, and posting content and (2) showing that these platforms qualify as content developers when they engineer their algorithms for addiction, which single-handedly makes that content tortious.³⁰⁰

While this Comment strives to solve a particularized problem, it is important to recognize how flawed algorithms can impact our world. Whether they reinforce discriminatory decision-making in the criminal justice system,³⁰¹ further political polarization,³⁰² or enable terrorism,³⁰³ the consequences of deficient algorithms are wide-reaching yet just beginning to surface.³⁰⁴ Many have urged that it is time for our society to

296. See Gerrard & Gillespie, *supra* note 8.

297. In recent months, concerns of online censorship have led many, including the President of the United States, to call for amendment to Section 230. In June of 2020, the President promulgated an Executive Order, which, among other things, instructed the National Telecommunications and Information Administration to request that the FCC set forth regulations to clarify the proper scope of Section 230. Christian T. Fjeld & Christopher Harvie, *The Implications of Trump's Executive Order and Section 230 of the Communications Decency Act*, NAT'L L. REV. (June 3, 2020), <https://www.natlawreview.com/article/implications-trump-s-executive-order-and-section-230-communications-decency-act> [<https://perma.cc/8LG2-W2PJ>]. The order also seeks regulations to consider “the conditions under which an online platform’s actions do not constitute ‘good faith.’” *Id.* While the executive order was aimed at combatting political censorship online, it is possible that a reduction in the section’s vast immunity could benefit future litigants in a defective design case. *Id.*

298. See *supra* Part II.A.

299. See *supra* Part II.B.

300. See *supra* Part II.C.

301. See O’NEIL, *supra* note 37, at 84–104.

302. See *Force v. Facebook, Inc.*, 934 F.3d 53, 87–88 (2d Cir. 2019) (Katzmann, C.J., concurring in part, dissenting in part).

303. *Id.* at 84–86 (Katzmann, C.J., concurring in part, dissenting in part).

304. See O’NEIL, *supra* note 37, at 199–218.

hold algorithms and their creators accountable,³⁰⁵ and it is clear that lawyers will play a paramount role in this struggle.³⁰⁶ So too will affirmative litigation. And thus, when lawyers initiate lawsuits to hold social media firms accountable for their efforts to engineer addiction, they will be doing their part in the quest for algorithmic reckoning.

305. See, e.g., *id.*; DiResta, *supra* note 7.

306. See O'NEIL, *supra* note 37, at 199–218.