



Regulating Algorithmic Content Distribution and Moderation by Online Platforms

By Josh Withrow and Shoshana Weissmann

Allowing too much regulation around algorithmic content curation invites government control over the free flow of information.

Executive Summary

The last decade has seen an explosion in attempts to regulate how online platforms—particularly social media sites—sort, deliver, and moderate content. This paper categorizes the major legislative approaches to regulating platform algorithms and explains why many of these efforts endanger the free internet and raise serious First Amendment concerns. It accomplishes this by briefly outlining how algorithmic content feeds work, acknowledging some of their potential negative impacts, questioning commonly held assumptions about how algorithms affect society, and explaining why a majority of the legislative mandates under consideration are likely to cause more harm than good.

Table of Contents

Executive Summary	1
Introduction	2
What Algorithms Are and How They Work	2
Algorithmic Content Delivery and Curation	2
Too Much of a Good (or Bad) Thing	3
Algorithmic Content Moderation and Its Limitations	4
Legislative Proposals to Control Algorithms	5
Regulating Platform Design	5
Imposing Liability	5
Algorithmic Transparency Mandates	5
Legal Barriers to Algorithmic Moderation	6
The Free-Speech Risks of Algorithmic Liability	7
Algorithmic Transparency and Its Limits	8
Takeaways for Lawmakers	8
Conclusion	9
About the Authors	9

The sources included in this paper were verified and active at the time of publication.

Introduction

Algorithms shape nearly every aspect of our digital experience. They determine the content and advertisements we see, as well as the way information is prioritized in our feeds. In recent years, the word “algorithm” has taken on a negative connotation, feeding fears that these automated systems negatively influence public discourse, youth, and even democracy itself. This has sparked concerns about how online platforms personalize, serve, and remove content. It has also led to questions about whether algorithms themselves should be regulated.

This paper focuses on the algorithms online platforms use to deliver, curate, and moderate content (i.e., not the machine-learning systems behind large language models or generative artificial intelligence tools, which would warrant separate examination). It explains how these algorithms work, clarifies misconceptions, and demonstrates how one-size-fits-all regulation could undermine individual user experiences and free expression.

What Algorithms Are and How They Work

Although there are many nuanced variations on how to define an algorithm in the context of computing, the National Institute of Standards and Technology provides the following baseline definition: “[a] clearly specified mathematical process for computation; a set of rules that, if followed, will give a prescribed result.”¹ For online platforms, algorithms serve two core purposes: helping users find relevant information and protecting users from unwanted or harmful content.

Social media, e-commerce, and web advertisers rely on algorithms to recommend content and products based on user behavior and preferences.² For example, search engines deploy “web crawlers” to index the internet and rank pages according to complex criteria such as relevance, credibility, and user engagement.³ Content-moderation algorithms identify and suppress harmful material before it reaches users. These systems improve efficiency and safety but also raise questions about transparency and bias.

Algorithmic Content Delivery and Curation

Many early online platforms displayed posts in simple reverse-chronological order, meaning the most recent posts appeared at the top of the feed. Today, few services rely solely on that model; instead, modern algorithms curate content based on relevance and engagement to better reflect users’ interests.

Each online platform applies its own version of relevance using engagement data and user behavior to decide what appears in a feed. For example, X (formerly Twitter) offers two main feeds: “Following,” which shows posts



For online platforms, algorithms serve two core purposes: helping users find relevant information and protecting users from unwanted or harmful content.

1. Kristian Lum and Rumman Chowdhury, “What is an ‘algorithm’? It depends whom you ask,” MIT Technology Review, Feb. 26, 2021. <https://www.technologyreview.com/2021/02/26/1020007/what-is-an-algorithm>; National Institute of Standards and Technology, “Algorithm,” U.S. Department of Commerce, last accessed Nov. 11, 2025. <https://csrc.nist.gov/glossary/term/Algorithm>.
2. Arvind Narayanan, “Understanding Social Media Recommendation Algorithms,” Knight First Amendment Institute, March 9, 2023. <https://knightcolumbia.org/content/understanding-social-media-recommendation-algorithms>.
3. Sergey Brin and Lawrence Page, “The Anatomy of a Large-Scale Hypertextual Web Search Engine,” *Computer Networks and ISDN Systems* 30 (1998), pp. 107-117. <https://snap.stanford.edu/class/cs224w-readings/Brin98Anatomy.pdf>.

from accounts a user follows, and “For You,” which adds algorithmic recommendations.⁴ The algorithm for “For You” includes a variety of factors when determining what posts to display (also called “ranking” posts), such as a user’s likelihood of engaging with a post, whether a user has seen the tweet before, and how much time the user has spent viewing similar posts.⁵ Pinterest personalizes recommendations in a similar way, based on a user’s boards, interests, and searches.⁶

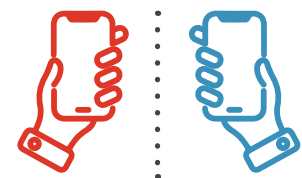
Search engines use similar logic. Google’s algorithm, for example, orders and formats posts based on the type of content searched. Searches about current events give more priority to recency, whereas searches about cooking prioritize recipes, regardless of recency.⁷

This tailoring means that many algorithms include some level of personalization. This may be stronger on platforms like Pinterest, where users go to curate information on specific topics, and weaker on platforms like X, where users seek out recent or popular information, but it has real utility. Without it, platforms like Spotify would not be able to suggest new artists, Yelp would not be able to highlight local services based on your search history or location, and AllTrails would not be able to match users to the highest-rated hiking trails for their location. Much of a given online platform’s success is tied to how well it can deliver this sort of personalized information.

Too Much of a Good (or Bad) Thing

Despite their benefits, algorithms are often cited as contributing to polarization, misinformation, and addiction. During the 2016 elections, critics argued that social media algorithms deepened ideological divides by showing users content that reinforced their existing views.⁸ Because these systems are designed to maximize engagement, they tend to serve up more of the content users already prefer, which critics argue amplifies confirmation biases and fuels the spread of misinformation.⁹ Opponents contend that sensational content draws higher engagement, regardless of whether or not it is true, and that engagement-driven design can foster addictive user behavior.¹⁰

However, recent research suggests that the problem is more complex and goes beyond social media algorithms. One study found that users who switched from algorithmic to reverse-chronological feeds became more—not less—likely to



During the 2016 elections, critics argued that social media algorithms deepened ideological divides by showing users content that reinforced their existing views.

4. X Help Center, “About your For you timeline on X,” X Corp., last accessed Sept. 21, 2025. <https://help.x.com/en/using-x/x-timeline>.
5. Rochi Zalani, “How the Twitter (Now X) Algorithm Works in 2025 (+ 11 Ways You Can Succeed on the Platform),” Buffer Marketing Library, Feb. 7, 2024. <https://buffer.com/resources/twitter-timeline-algorithm>.
6. Pinterest, “Tune Your Home Feed,” Pinterest Help Center, last accessed Sept. 21, 2025. <https://help.pinterest.com/en/article/tune-your-home-feed>; Brian Johnson, “Taste Graph part 1: Assigning interests to Pins,” Pinterest Engineering Blog, Sept. 21, 2017. <https://medium.com/pinterest-engineering/taste-graph-part-1-assigning-interests-to-pins-9158b4c25906>.
7. “How Google Search Determines Ranking Results,” Google, last accessed Sept. 21, 2025. https://www.google.com/intl/en_us/search/howsearchworks/how-search-works/ranking-results.
8. Paul Barrett et al., “How Tech Platforms Fuel U.S. Political Polarization and What Government Can Do About It,” Brookings Institution Commentary, Sept. 27, 2021. <https://www.brookings.edu/articles/how-tech-platforms-fuel-u-s-political-polarization-and-what-government-can-do-about-it/>; “How Algorithms And Filter Bubbles Decide What We See On Social Media,” BBC Bitesize, last accessed Sept. 21, 2025. <https://www.bbc.co.uk/bitesize/articles/zd9tt39>.
9. “Algorithms,” Queens University Library, last accessed Sept. 21, 2025. <https://library.queens.edu/misinformation-on-social-media/algorithms>; Sanja Kelly et al., *Manipulating Social Media to Undermine Democracy* (Freedom House, 2017). https://www.americanbar.org/groups/public_interest/election_law/american-democracy/our-work/democracy-database/manipulating-social-media-undermine-democracy/; Chico Q. Camargo “YouTube’s algorithms might radicalise people — but the real problem is we’ve no idea how they work,” The Conversation, Jan. 21, 2020. <https://theconversation.com/youtubes-algorithms-might-radicalise-people-but-the-real-problem-is-weve-no-idea-how-they-work-129955>.
10. Kelsey Hansen, “Our Social Media Addiction,” *Harvard Business Review Magazine* (November-December 2022). <https://hbr.org/2022/11/our-social-media-addiction>.

click on partisan news sources.¹¹ Other studies have suggested that traditional media—like television and newspapers—may have a greater role in increasing partisan polarization and spreading misinformation than social media.¹² In fact, elite-driven and mass-media-led disinformation campaigns—such as those surrounding mail-in voting fraud—have had a far stronger impact than online sharing alone.¹³

Moreover, an experiment that studied users who abstained from social media during the lead-up to the 2020 election found that these platforms did not have a strong overall influence on polarization; although, in some cases, deactivating Facebook or Instagram actually correlated with increased issue-level polarization.¹⁴ In other words, conspiracy theories, confirmation bias, and polarization are long-standing features of human behavior. Blaming algorithms for these issues in isolation obscures deeper cultural and psychological drivers.

Algorithmic Content Moderation and Its Limitations

Algorithms play a crucial role in moderating online content by identifying and removing harmful or illegal material at scale. Platforms such as X and Snapchat, for example, moderate hundreds of millions of posts every year, relying on automated systems to detect hate speech, child exploitation, and other violations.¹⁵ These systems use algorithms to reduce human exposure to disturbing content by removing posts before they become public.¹⁶

Yet moderation at scale is inherently difficult; decisions about where to draw the line on what content to remove are subjective, and algorithms struggle with nuance, satire, and context.¹⁷ For instance, a system designed to block neo-Nazi content may inadvertently suppress historical documentaries containing swastikas.¹⁸ Likewise, algorithms designed to target content referencing sensitive topics like self-harm, substance abuse, and eating disorders could remove self-help resources and support groups on the same topics.¹⁹ Even with a 99.9 percent accuracy rate, a platform handling 350 million daily uploads would still make hundreds of thousands of mistakes each day.²⁰



Moderation at scale is inherently difficult; decisions about where to draw the line on what content to remove are subjective, and algorithms struggle with nuance, satire, and context.

This is why having a human element involved in content moderation remains essential—to both identify harmful content that evades the algorithms and

11. Andrew M. Guess et al., “How do social media feed algorithms affect attitudes and behavior in an election campaign?,” *Science* 381:6656 (July 27, 2023), pp. 398–404. <https://doi.org/10.1126/science.abp9364>.
12. Daniel Muise et al., “Quantifying Partisan News Diets in Web and TV Audiences,” *Science Advances* 8:eabn0083 (2022). <https://www.science.org/doi/10.1126/sciadv.abn0083>; “Tackling Conspiracy Theories,” Ipsos, December 2021. https://www.ipsos.com/sites/default/files/ct/publication/documents/2021-12/ipsos_Tackling_Conspiracy_Theories.pdf.
13. Yochai Benkler et al., “Mail-In Voter Fraud: Anatomy of a Disinformation Campaign,” Berkman Klein Center, Oct. 1, 2020. <https://cyber.harvard.edu/publication/2020/Mail-in-Voter-Fraud-Disinformation-2020>.
14. Hunt Allcott et al., “The effects of Facebook and Instagram on the 2020 election: A deactivation experiment,” *Proceedings of the National Academy of Sciences of the United States of America* 121:21 (May 13, 2024). <https://www.pnas.org/doi/10.1073/pnas.2321584121>.
15. “Global Transparency Report H2 2024,” X Corp., 2025. <https://transparency.x.com/en/reports/global-reports/2025-transparency-report>; “Transparency Report,” Snap Inc., June 20, 2025. <https://values.snap.com/privacy/transparency>.
16. Ruth Spence et al., “The psychological impacts of content moderation on content moderators: a qualitative study,” *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 17:4 (2023). <https://cyberpsychology.eu/article/view/33166>.
17. Mike Masnick, “Masnick’s Impossibility Theorem: Content Moderation At Scale Is Impossible To Do Well,” *Techdirt*, Nov. 20, 2019. <https://www.techdirt.com/2019/11/20/masnicks-impossibility-theorem-content-moderation-scale-is-impossible-to-do-well>.
18. Jim Waterston, “YouTube blocks history teachers uploading archive videos of Hitler,” *The Guardian*, June 6, 2019. <https://www.theguardian.com/technology/2019/jun/06/youtube-blocks-history-teachers-uploading-archive-videos-of-hitler>.
19. Ashken Kazaryan, “I Built This Algo Brick by Brick,” *The Future of Free Speech*, Sept. 10, 2025. <https://futurefreespeech.org/wp-content/uploads/2025/09/I-Built-This-Algo-Brick-by-Brick-September-2025-The-Future-of-Free-Speech.pdf>.
20. Masnick. <https://www.techdirt.com/2019/11/20/masnicks-impossibility-theorem-content-moderation-scale-is-impossible-to-do-well>.

to hopefully reduce the amount of legitimate speech that is filtered out.²¹ Still, automated moderation has the virtue of being proactive in reducing not only platform users' but also human moderators' exposure to much of the most horrific content posted on the internet.²² The challenge lies not in whether to use algorithms but in how to balance efficiency with fairness and transparency.

Legislative Proposals to Control Algorithms

Lawmakers across the country—at both the state and federal levels—have proposed legislation to regulate content-related algorithms. These efforts generally fall into one or more of three basic approaches: direct regulation of platform design, imposing liability for algorithmic outputs, or mandating algorithmic transparency.

Regulating Platform Design

The “Filter Bubble Transparency Act,” introduced by a bipartisan group of senators in 2021, exemplifies problematic platform-design regulation.²³ This bill would have required even small platforms such as TruthSocial and AllTrails, as well as many news websites, to provide an alternative feed that excluded user-specific data unless explicitly provided for that purpose. Although well-intentioned, this bill would have forced many services to present a version of their product that was useless by default. A location-based service unable to recommend nearby trails or a news feed stripped of relevant content would be less useful, not safer.



Imposing Liability

One of the most recent examples of legislation designed to impose liability on platforms for what their algorithms recommend is California's SB 771.²⁴ This legislation would have established a private right of action for any occasion in which a social media platform's algorithm delivered any content deemed to violate California's civil rights laws. The covered platforms would have legally been presumed to be “knowingly” disseminating such content for the purposes of enforcement. Free speech advocates pointed out that this would force platforms to take down large quantities of constitutionally protected speech, and although the bill passed the state legislature, it was subsequently vetoed by Governor Gavin Newsom.²⁵



Algorithmic Transparency Mandates

Algorithmic transparency is frequently cited as a lighter-touch approach to regulating online platforms—in theory to simply provide consumers, lawmakers, and researchers with better insight into how users' online data



21. “Content Moderation in a New Era for AI and Automation,” Oversight Board, September 2024. <https://www.oversightboard.com/news/content-moderation-in-a-new-era-for-ai-and-automation>.
22. Julia Shaw, “Content moderators pay a psychological toll to keep social media clean. We should be helping them” BBC Science Focus, Nov. 2, 2022. <https://www.sciencefocus.com/news/content-moderators-pay-a-psychological-toll-to-keep-social-media-clean-we-should-be-helping-them>.
23. “S.2024 - Filter Bubble Transparency Act,” U.S. Senate, 117th Congress. <https://www.congress.gov/bill/117th-congress/senate-bill/2024/text>.
24. “SB-771 Personal rights: liability: social media platforms,” California Legislative Information, 2025. https://leginfo.ca.gov/faces/billStatusClient.xhtml?bill_id=202520260SB771.
25. Sophia Mandt, “California Is Advancing a Bill To Punish Social Media Companies for Not Suppressing Speech,” Reason, July 24, 2025. <https://reason.com/2025/07/24/california-is-advancing-a-bill-to-punish-social-media-companies-for-not-suppressing-speech>; Veto notice, Office of Governor Gavin Newsom, Oct. 13, 2025. <https://www.gov.ca.gov/wp-content/uploads/2025/10/SB-771-Veto.pdf>.

is processed and how it affects the content they are presented with. For example, a piece of legislation being advanced at the time of writing this paper in Massachusetts, S. 51, is meant to require third-party audits of how online services' algorithms target and deliver content to minors.²⁶ However, such bills pose an implicit (or, in the case of S. 51, explicit) threat that the findings of such audits might be used for regulation that could easily result in a highly limiting effect on the speech that platforms are willing to host.²⁷

Legal Barriers to Algorithmic Moderation

Two core legal protections shield online platforms from excessive liability related to algorithmic content moderation: the First Amendment and Section 230 of the Communications Decency Act.²⁸ The First Amendment safeguards free speech and editorial discretion, while Section 230 ensures that platforms are not held liable for third-party content or moderation decisions.²⁹ Together, they allow platforms to host vast amounts of speech without crippling legal risk.

Cases such as *Gonzalez v. Google* and *Twitter v. Taamneh* have tested whether platforms could be held liable for user content that incited violence.³⁰ In both cases, lower courts ruled that the platforms could potentially be held liable for hosting content that inspired terrorist attacks or recruitment; however, the U.S. Supreme Court vacated both decisions.³¹ The more recent *TikTok v. Anderson* introduced doubt about whether algorithmic recommendations are fully covered by Section 230 when a content suggestion leads someone (in this case, a child) to engage in a dangerous activity that harms them.³² However, thus far, the U.S. Supreme Court has refused to directly address arguments for limiting Section 230's protection of content algorithms.³³

Meanwhile, at the state level, outraged by what they viewed as partisan bias against conservative speech in content-takedown decisions by many major social media platforms during the COVID-19 pandemic and the 2020 elections, some conservative lawmakers have introduced numerous bills to disable or eliminate Section 230 protections for platforms whose editorial decisions are not "viewpoint neutral." Texas and Florida were the first states to enact variants of this approach, using the threat of liability to pressure companies to change their content moderation practices.³⁴ The Texas law tried to mandate that social media sites maintain "viewpoint neutrality" in their content-removal decisions, while the Florida law held that posts by politicians and political candidates could not be taken down even if they violated a site's terms of service.



The more recent *TikTok v. Anderson* introduced doubt about whether algorithmic recommendations are fully covered by Section 230 when a content suggestion leads someone (in this case, a child) to engage in a dangerous activity that harms them.

26. John C. Velis, "An Act Relative to Social Media, Algorithm Accountability, and Transparency," Massachusetts Legislature. <https://malegislature.gov/Bills/194/SD1313>.

27. Eric Goldman, "The Constitutionality of Mandating Editorial Transparency," *UC Law Journal* 73:5 (2022), pp. 1203-1232. https://uclawjournal.org/wp-content/uploads/2.-Goldman_Final.pdf.

28. See, e.g., Jeff Kossseff, *The Twenty-Six Words that Created the Internet* (Cornell University Press, 2019).

29. Eric Goldman, "Why Section 230 is Better than the First Amendment," *Notre Dame Law Review Reflection* 95:1 (2019). https://scholarship.law.nd.edu/ndlr_online/vol95/iss1/3.

30. *Gonzalez v. Google*, No. 21-1333, May 18, 2023. <https://supreme.justia.com/cases/federal/us/598/21-1333>; *Twitter, Inc. v. Taamneh*, No. 21-1496, May 18, 2023. <https://supreme.justia.com/cases/federal/us/598/21-1496>.

31. Eric N. Holmes, "Liability for Algorithmic Recommendations," Congress.gov, Oct. 12, 2023. <https://www.congress.gov/crs-product/R47753>.

32. *Anderson v. TikTok Inc.*, U.S. Court of Appeals for the Third Circuit, No. 22-3061, Aug. 27, 2024. <https://law.justia.com/cases/federal/appellate-courts/ca3/22-3061/22-3061-2024-08-27.html>.

33. Kazaryan. <https://futurefreespeech.org/wp-content/uploads/2025/09/I-Built-This-Algo-Brick-by-Brick-September-2025-The-Future-of-Free-Speech.pdf>.

34. "HB 20: Relating to censorship of or certain other interference with digital expression, including expression on social media platforms or through electronic mail message," Texas Legislature Online, 2021. <https://capitol.texas.gov/billlookup/History.aspx?LegSess=872&Bill=HB20>; "SB 702: Social Media Platforms," The Florida Senate, 2021. <https://www.flsenate.gov/Session/Bill/2021/7072>.

Both of these laws were vacated by the U.S. Supreme Court and remanded for further review. In both cases, Justice Elena Kagan, writing for the majority, explained:

When the platforms [Facebook and YouTube] use their Standards and Guidelines to decide which third-party content those feeds will display, or how the display will be ordered and organized, they are making expressive choices. And because that is true, they receive First Amendment protection.³⁵

These rulings reinforce that algorithmic decisions are ultimately editorial decisions that parallel those made by publishers when deciding what content or products to feature.³⁶ With certain limited exceptions (such as already-illegal or pornographic content), courts have generally held that the First Amendment, which protects free speech, applies without regard to changes in technology.³⁷

The Free-Speech Risks of Algorithmic Liability

Despite these constitutional and statutory protections largely being upheld by the courts, lawmakers continue to advance proposals that would make online platforms liable for algorithmic choices. For example, California’s 2022 Age-Appropriate Design Code Act (AADC) requires platforms to disable by default any features deemed “materially detrimental to the child’s physical health, mental health, or well-being.”³⁸ A number of similar state and federal bills followed, including the federal Kids Online Safety Act (KOSA).³⁹ In addition, Senator Amy Klobuchar (D-Minn.) introduced the Health Misinformation Act of 2021, which would have allowed online platforms hosting user content to be held liable for “health misinformation” if an algorithm that is not considered “neutral” promotes the content.⁴⁰ Under the bill, the Secretary of Health and Human Services would define what constitutes “health misinformation.”⁴¹

Such measures risk turning government officials into arbiters of truth, undermining both free expression and user choice. Most misinformation is constitutionally protected speech, and existing laws already provide remedies for defamation or fraud.⁴² Courts are likely to strike down these bills as unconstitutional restraints on editorial decisions, as they have with California’s AADC, which has already been enjoined and is widely expected to be overturned.⁴³



Despite these constitutional and statutory protections largely being upheld by the courts, lawmakers continue to advance proposals that would make online platforms liable for algorithmic choices.

-
35. *Moody v. NetChoice, LLC*, Supreme Court of the United States, No. 22-277, Oct. 3, 2023. https://www.supremecourt.gov/opinions/23pdf/22-277_d18f.pdf.
36. Corbin Barthold and Berin Szóka, “No, Florida Can’t Regulate Online Speech,” *Lawfare*, March 12, 2021. <https://www.lawfaremedia.org/article/no-florida-cant-regulate-online-speech>.
37. *District of Columbia v. Heller*, Supreme Court of the United States, June 26, 2008. <https://www.law.cornell.edu/supct/pdf/08-1448P.ZO>.
38. “The California Age Appropriate Design Code Act,” California Legislature, AB 2273, 2022. https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=202120220AB2273.
39. “S. 1748: Kids Online Safety Act,” 119th Congress (2025-2026). <https://www.congress.gov/bill/119th-congress/senate-bill/1748/text>.
40. Canyon Brimhall et al., “R Street Institute Concerns About the Kids Online Safety Act,” R Street Institute, July 27, 2023. <https://www.rstreet.org/commentary/r-street-institute-concerns-about-the-kids-online-safety-act>; U.S. Senate, “S.2448: Health Misinformation Act of 2021,” 117th Congress (2021-2022). <https://www.congress.gov/bill/117th-congress/senate-bill/2448/text>.
41. Ibid.; Amy Klobuchar (@amyklobuchar), “When were you lying, sir?” Asked my colleague Tina Smith of RFK Jr., ‘When you told this committee that you were not anti-vax? Or when you told Americans that there’s no safe and effective vaccine?’ Kennedy replied, ‘Both things are true.’ That’s a double scream: 🤯🤯,” Sept. 4, 2025, 3:29 PM. [Tweet] <https://x.com/amyklobuchar/status/1963685913520681131>; Amy Klobuchar (@amyklobuchar), “Elizabeth’s look of incredulity when listening to RFK Jr explain his decimation of the CDC and all other havoc he’s wrought pretty much says it all,” Sept. 4, 2025. [Threads] <https://www.threads.com/@amyklobuchar/post/DOMSGZkbb4f/elizabeths-look-of-incredulity-when-listening-to-rfk-jr-explain-his-decimation-o>.
42. Valerie C. Brannon, “False Speech and the First Amendment: Constitutional Limits on Regulating Misinformation,” *Congress.gov*, Aug. 1, 2022. <https://www.congress.gov/crs-product/IF12180>; “Fact Sheet: Regulation of False, Misleading, or Intimidating Speech about Elections,” Georgetown Law Institute for Constitutional Advocacy and Protection, last accessed Oct. 10, 2025. <https://www.law.georgetown.edu/icap/wp-content/uploads/sites/32/2024/08/Fact-Sheet-False-Misleading-and-Intimidating-Election-Information.pdf>.
43. Eric Goldman, “California’s Age Appropriate Design Code (AADC) Is Completely Unconstitutional (Multiple Ways)—*NetChoice v. Bonta*,” *Technology & Marketing Law Blog*, March 17, 2025. <https://blog.ericgoldman.org/archives/2025/03/californias-age-appropriate-design-code-aadc-is-completely-unconstitutional-multiple-ways-netchoice-v-bonta.htm>.

Algorithmic Transparency and Its Limits

Given the constitutional difficulties of directly regulating platforms' editorial and design choices, a more indirect legislative approach has been to require online platforms to disclose how their content-delivery and moderation algorithms operate. The concept behind this approach is that regulators should be able to audit a site's algorithms to determine whether decisions to display or remove content are being made fairly according to the site's own terms of service. However, as previously discussed, content moderation is subjective, and violations often depend on interpretation, making such laws vulnerable to First Amendment challenge.⁴⁴

A less-intrusive approach to algorithmic transparency might require online platforms to grant controlled access to their algorithms and the data that informs them so independent experts can assess them more thoroughly. The Platform Accountability and Transparency Act (PATA) is perhaps the most thoughtful such proposal at the federal level.⁴⁵ There may indeed be some benefit to allowing academic or research access to this type of data, but it also presents serious practical and legal obstacles.⁴⁶

For example, any forced disclosure rules would have to take into account data privacy and security concerns, limiting the scope of access and recognizing that some proprietary data may be inappropriate to share. Regulators would also need clear safeguards to prevent misuse of data, such as that seen in the Cambridge Analytica scandal, in which an academic research partner repurposed Facebook data for political targeting.⁴⁷ Such a proposal would also need to define who is legally liable for the release of platforms' data should it be mishandled or breached by external researchers.

Takeaways for Lawmakers

Policymakers aiming to address concerns about algorithms should do so with care. The following takeaways outline guiding principles for lawmakers seeking to balance constitutional limits, market dynamics, and user choice.

1. Regulating Algorithms Means Regulating Speech. Any attempt to regulate online platforms and the algorithms they use to curate content must consider the platform owner's right to editorial discretion and the potential effects that the regulation could have more broadly. For instance, any laws that place undue burdens on free speech remain highly vulnerable to constitutional challenges. Laws regulating platform design, especially with respect to how algorithms sort or recommend content, inherently affect platform owners' freedom of speech and association.

2. Let Markets and Innovation Work. Attempts to define best practices through prescriptive regulation of algorithmic content curation could disrupt innovation and competition. For example, overly intrusive content moderation during



A less-intrusive approach to algorithmic transparency might require online platforms to grant controlled access to their algorithms and the data that informs them so independent experts can assess them more thoroughly.

44. Eric Goldman, "The Constitutionality of Mandating Editorial Transparency," *UC Law Journal* 73:5 (2022), pp. 1203-1232. https://uclawjournal.org/wp-content/uploads/2.-Goldman_Final.pdf.

45. "S.1876 - Platform Accountability and Transparency Act," 118th Congress (2023-2024). <https://www.congress.gov/bill/118th-congress/senate-bill/1876>.

46. Daphne Keller, "Platform Transparency and the First Amendment," *Journal of Free Speech Law* 4:1 (2023), pp. 1-86. <https://www.journaloffreespeechlaw.org/keller2.pdf>.

47. "FTC Imposes \$5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook," Federal Trade Commission, July 24, 2019. <https://www.ftc.gov/news-events/news/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions-facebook>.

the COVID-19 pandemic and 2020 elections created a market for alternative platforms such as Truth Social and Rumble, while changes to content policies by Twitter (now X) caused many users to migrate to Mastodon, BlueSky, and Threads.⁴⁸

Another way that markets are adjusting organically is through decentralization, which gives platform users more control over the algorithms shaping their experience. In recent years, more open, protocol-based platforms such as BlueSky have begun to allow “middleware,” where users can add custom software that builds an experience that moderates content to their personal tastes.⁴⁹ Other platforms, such as Facebook and X, have introduced community-driven, fact-checking systems, which have shown promise in checking the spread of misinformation.⁵⁰

3. Empower Users, Not Bureaucrats. More market-driven approaches give greater agency to individual consumers. Lawmakers might consider ways to further empower people to have better experiences with social media and the internet. For example, rather than removing choice by imposing outright bans on certain algorithmic features for minors, some states have taken the approach of incorporating digital literacy into school curriculums.⁵¹

Conclusion

Social media sites and other content-hosting internet platforms are still reaching maturity, and society continues to learn how these platforms shape public discourse. Lawmakers should proceed with humility, recognizing that markets tend to self-correct, though sometimes more slowly than we might prefer. Allowing too much governmental regulation around algorithmic content curation invites government control over the free flow of information. Many of the problems policymakers hope to solve through regulation are more complex than the algorithms themselves, and poorly designed proposals could make the internet less open and useful without achieving the intended goals.



Lawmakers should proceed with humility, recognizing that markets tend to self-correct, though sometimes more slowly than we might prefer.

About the Authors

Josh Withrow is a resident fellow in Technology and Innovation Policy at the R Street Institute.

Shoshana Weissmann is a former R Street digital media director and policy fellow and is currently the digital director at the Abundance Institute.

48. Galen Stocking et al., “2. Alternative social media sites frequently identify as free speech advocates,” Pew Research Center, Oct. 6, 2022. <https://www.pewresearch.org/journalism/2022/10/06/alternative-social-media-sites-frequently-identify-as-free-speech-advocates>. Jennifer Gregory, “The X-odus: How User Departures Are Reshaping Social Media Dynamics,” Website Planet, Nov. 4, 2025. <https://www.websiteplanet.com/blog/people-and-companies-leaving-x>.
49. Luke Hogg and Renee DiResta, “Enabling a Thriving Middleware Market,” Tech Policy Press, April 11, 2025. <https://www.techpolicy.press/enabling-a-thriving-middleware-market>.
50. Spence Purnell, “Meta’s Bet on Decentralized Moderation May Be Working,” R Street Institute, June 18, 2025. <https://www.rstreet.org/commentary/metas-bet-on-decentralized-moderation-could-be-working>.
51. Anna Merod, “California joins small, growing number of states requiring K-12 media literacy,” K-12Dive, Nov. 15, 2023. <https://www.k12dive.com/news/california-media-literacy-k12-law/699911>.