



MY AI HELPER SERIES

The Dark Side of Algorithms

An Investigative Report on YouTube's Recommendation Engine

Investigative Report

February 2026

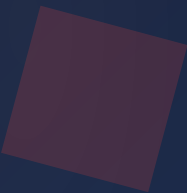


Table of Contents

Executive Summary

1. Introduction: The Algorithm That Shapes Our Reality

2. How YouTube's Algorithm Works

2.1 The Two-Stage Architecture

2.2 The Signals That Drive Recommendations

2.3 Evolution: From Clicks to Satisfaction

3. Criticism and Controversies

3.1 The Mozilla Foundation Study

3.2 Radicalization Concerns

3.3 Children's Content and Safety

4. Academic Research Findings

4.1 The University of Pennsylvania Study

4.2 The Princeton/UC Davis Study

4.3 Misinformation Filter Bubbles

5. Content Moderation and Policy

6. Regulatory Response: The EU Digital Services Act

7. Impact on the Creator Economy

8. Conclusion

References

Executive Summary

YouTube's recommendation algorithm stands as one of the most powerful and influential artificial intelligence systems ever deployed. Processing over **80 billion signals daily** from more than **2 billion active users**, this system determines what billions of people watch, learn, and believe across the globe.

This investigative report examines the mechanics, impacts, and controversies surrounding YouTube's recommendation engine. Our analysis draws exclusively from verified, reputable sources including peer-reviewed academic research, official Google documentation, and investigations by established organizations.

Key Findings:

- YouTube's algorithm uses a sophisticated two-stage deep neural network architecture that has evolved significantly from optimizing for clicks to prioritizing "valued watch time" and user satisfaction.
- A landmark Mozilla Foundation study found that **71% of "regretted" videos** came from algorithmic recommendations rather than user searches, with recommended videos being **40% more likely** to be regretted.
- Academic research presents conflicting evidence on radicalization: while some studies suggest the algorithm can create filter bubbles, others indicate user preferences play a larger role than algorithmic influence.
- Children's safety remains a significant concern, with research demonstrating that YouTube's algorithm can rapidly expose young users to violent, sexual, or extremist content.
- The European Union's Digital Services Act represents the most significant regulatory attempt to mandate algorithmic transparency and accountability for platforms like YouTube.

The evidence suggests that while YouTube has made genuine efforts to address harmful recommendations, fundamental tensions remain between engagement-based business models

and user wellbeing. The opacity of these systems continues to limit meaningful public oversight and accountability.

1. Introduction: The Algorithm That Shapes Our Reality

YouTube is the world's second-most visited website, hosting over **800 million videos** and serving more than **2 billion logged-in users monthly** ^[1]. But beneath this vast library of content lies something far more consequential: an artificial intelligence system that quietly shapes the information diet of nearly a quarter of the world's population.

The platform's recommendation algorithm has been described as potentially "one of the most powerful radicalizing instruments of the 21st century"^[2]. This characterization, whether accurate or hyperbolic, reflects genuine concerns about the societal impact of a system that determines what billions of people watch, often without their explicit awareness of how those choices are being made.

Unlike traditional media with human editors making conscious decisions, YouTube's recommendations are determined by deep neural networks operating at a scale that defies human comprehension. The system processes over **80 billion pieces of information**—called "signals"—every single day^[3]. These signals include watch history, search queries, likes, dislikes, and countless behavioral patterns that most users never realize they're generating.

This report provides a comprehensive, evidence-based examination of YouTube's recommendation algorithm. We analyze how it works, what the research says about its impacts, and what measures are being taken—or should be taken—to ensure these powerful systems serve the public interest.

2. How YouTube's Algorithm Works

Understanding YouTube's recommendation system requires examining both its technical architecture and its evolutionary history. The system has transformed dramatically since its inception, reflecting both technological advances and growing awareness of the responsibilities that come with such influence.

2.1 The Two-Stage Architecture

At its core, YouTube employs a sophisticated two-stage recommendation process powered by deep neural networks^[4]. This architecture is necessary because the platform must select from hundreds of millions of videos to present just a handful of recommendations to each user—an extraordinary computational challenge performed billions of times daily.

Stage One: Candidate Generation

The first neural network, called the candidate generation network, performs what engineers call "broad recall." It takes a user's activity history—watch patterns, search queries, demographic information—and reduces the universe of possible videos from hundreds of millions to just a few hundred candidates^[4].

This network uses several key techniques:

- **Embedding retrieval:** Both users and videos are mapped into a shared mathematical space where similarity can be calculated. Users with similar viewing patterns are clustered together, enabling collaborative filtering.
- **Collaborative filtering:** The system identifies users with similar histories and recommends videos that similar users have enjoyed.
- **Metadata features:** For new users without extensive history (the "cold start" problem) or for newly uploaded videos, the system relies on titles, descriptions, tags, and categories.
- **Popularity signals:** Trending or viral content receives boosts, particularly for users without established preference profiles.

Stage Two: Ranking

The second neural network, the ranking model, performs fine-grained scoring of the candidate videos. This network uses a much richer set of features to assign each video a score predicting how likely the user is to

engage with it^[4].

The ranking model considers:

- Predicted watch time
- Expected user satisfaction (based on survey data and behavioral patterns)
- Video freshness and novelty
- Thumbnail quality and appeal
- Diversity considerations (preventing repetitive recommendations)

Crucially, YouTube's ranking system uses multi-objective optimization—no single metric dominates. The system balances engagement metrics with satisfaction signals, attempting to maximize long-term user value rather than immediate clicks.

2.2 The Signals That Drive Recommendations

YouTube's system relies on two categories of signals: explicit feedback that users consciously provide, and implicit behavioral signals that users generate without necessarily realizing it^[3].

Table 1 Primary Signals Used by YouTube's Recommendation System

Signal Category	Specific Signals	Impact on Recommendations
Explicit Feedback	"Not Interested" clicks	Immediately deprioritizes similar content
	"Don't Recommend Channel"	Removes channel from recommendation pool
	Satisfaction surveys (1-5 stars)	Directly measures "valued watch time"
	Likes/Dislikes	Informs future content predictions

Implicit Signals	Watch time and completion rates	Primary engagement metric
	Skip and abandon patterns	Negative satisfaction indicator
	Click-through rate	Initial interest measure
	Search history	Explicit intent signal
	Context (device, time, location)	Situational relevance

A critical insight from Google's published research is that not all watch time is equal. YouTube has shifted from optimizing for raw watch time to what they call "valued watch time"—time spent watching content that users actually find satisfying, as measured through surveys and behavioral patterns^[5].

2.3 Evolution: From Clicks to Satisfaction

YouTube's algorithm has undergone a dramatic transformation over the past decade, reflecting both technological capabilities and growing awareness of the system's societal impact^[6]:

2008-2011: The Popularity Era

In its earliest form, YouTube's recommendation system was essentially a "Trending" page—videos were ranked primarily by view count. This created a frustrating experience where users interested in niche topics would see only mainstream popular content.

2012: The Watch Time Shift

YouTube announced a major change: optimizing for watch time rather than clicks. The logic was that longer viewing indicated higher-quality content. However, this created unintended consequences as creators began padding videos with filler content or using clickbait thumbnails to drive initial clicks.

2015-Present: The Satisfaction Revolution

Around 2015, YouTube implemented a more sophisticated AI-driven system that attempts to measure actual user satisfaction, not just engagement. This system uses machine learning models trained on survey responses to predict which videos users will find genuinely valuable^[6].

According to Todd Beaupre from YouTube's Growth and Discovery Team, recommendations aren't calculated in advance: "It's at that moment that the recommendation engine is engaged... Given that this viewer is coming to YouTube on this type of device, at this time of day, with this video history, what are the best videos that we can show them right now?"^[6]

2019-Present: Responsibility Measures

Following widespread criticism, YouTube implemented significant changes to reduce recommendations of borderline content—material that doesn't violate Community Guidelines but is potentially harmful or misleading. The company claims these changes reduced consumption of borderline content by **70%** ^[7].

3. Criticism and Controversies

Despite YouTube's efforts to improve its recommendation system, the platform has faced sustained criticism from researchers, journalists, and civil society organizations. These criticisms center on three primary concerns: the amplification of harmful content, potential radicalization pathways, and inadequate protection of children.

3.1 The Mozilla Foundation Study

In 2021, the Mozilla Foundation published what remains the largest crowdsourced investigation into YouTube's recommendation algorithm^[8]. The study, conducted in partnership with researchers from the University of Exeter, analyzed **3,362 videos** that volunteers reported as "regrettable"—content they wished they hadn't watched.

Key Finding: Recommendations Drive Regrets

71% of all regret reports came from videos recommended by YouTube's algorithm, not videos users searched for. Recommended videos were **40% more likely** to be regretted than searched videos.

The study revealed troubling patterns in the content being recommended:

- **Misinformation (20%):** Conspiracy theories and political disinformation
- **COVID-19 fearmongering (12%):** Health misinformation during a global pandemic
- **Violent/graphic content (14%):** Disturbingly violent videos
- **Hate speech (12%):** Content targeting specific groups

Perhaps most concerning, **9% of the regrettable videos**—which had accumulated **160 million views**—were later removed by YouTube for violating the platform's own Community Guidelines^[8]. This suggests the algorithm was actively promoting content that even YouTube deemed unacceptable.

The study also revealed significant geographic disparities. Users in non-English-speaking countries experienced regrettable content at a rate **60% higher** than English-speaking users, with Brazil, Germany, and France showing the highest incidence rates. This raises serious questions about YouTube's content moderation capabilities across different languages and regions.

"When it's actively suggesting that people watch content that violates YouTube's policies, the algorithm seems to be working at odds with the platform's stated aims, their own community guidelines, and the goal of making the platform a safe place for people."

— Brandi Geurkink, Mozilla Senior Manager of Advocacy^[9]

3.2 Radicalization Concerns

The question of whether YouTube's algorithm actively radicalizes users has been the subject of intense debate and conflicting research findings. The "rabbit hole" theory suggests that the algorithm progressively exposes users to increasingly extreme content, creating pathways to radicalization.

A 2023 study from Princeton University and UC Davis put this theory to empirical test^[2]. The researchers created "sock puppet" accounts—automated accounts that simulate user behavior—across the political spectrum and tracked what recommendations they received.

The study's findings were nuanced:

- The researchers did **not** find meaningful increases in ideological extremity in video recommendations as users went deeper into their viewing sessions.
- However, they observed a "growing proportion" of recommendations from "problematic" channels (Alt-right, Conspiracy, Socialist, QAnon) for users categorized as "right" and "very-right."
- On average, **36.1%** of sock puppets across the political spectrum received recommendations from problematic channels. This rose to **40%** for far-right sock puppets.
- For "very-right" sock puppets, the chances of encountering far-right recommendations increased by **37%** the longer they engaged with YouTube.

The researchers concluded that their audit "suggests the presence of right-wing bias in YouTube's algorithm"^[2].

However, other research has challenged the radicalization narrative. A 2024 study from the University of Pennsylvania's Computational Social Science Lab found that users' own political interests and preferences play the primary role in what they choose to watch^[10]. The researchers created bots that either followed recommendations or ignored them, finding that bots following recommendations actually consumed *less* partisan content than real users.

"On average, relying exclusively on the recommender results in less partisan consumption... This gap corresponds to an intrinsic preference of users for such content relative to what the algorithm recommends."

— Homa Hosseinmardi, University of Pennsylvania^[10]

These conflicting findings highlight the complexity of studying algorithmic influence. The algorithm's effects may vary dramatically based on user behavior, content type, and political context.

3.3 Children's Content and Safety

Perhaps no area of YouTube's algorithm has generated more concern than its impact on children. Despite YouTube's claims that its main platform "has never been for kids under 13," research consistently shows that children constitute a significant portion of the user base^[11].

A 2023 investigation by the Tech Transparency Project created simulated accounts mimicking 9-year-old and 14-year-old boys interested in video games^[12]. The findings were deeply concerning:

Key Finding: Gun Content Exposure

Accounts that clicked on YouTube's recommendations received **382 firearms-related videos** in a single month—approximately 12 per day. Accounts that ignored recommendations received only 34 such videos.

The recommended content included graphic videos about school shootings, tactical gun training, and instructions on making firearms fully automatic. One video featured an elementary school-age girl wielding a handgun; another showed a shooter using a .50 caliber gun to fire on a dummy head filled with lifelike blood and brains^[12].

Research published in the journal *Pediatrics* examined YouTube Kids specifically, finding that the platform's algorithm can create increasingly restricted content "diets" for young children, potentially building profiles of children as digital consumers from a very young age^[13].

The study noted additional concerns:

- Children copying inappropriate behaviors seen on screen
- Impact on communication and language development
- Effects on attention spans and ability to concentrate

- Association with increased emotional and behavioral problems

4. Academic Research Findings

Beyond the headline-grabbing controversies, academic researchers have conducted rigorous studies to understand how YouTube's algorithm actually functions in practice. These studies provide crucial evidence for understanding both the capabilities and limitations of the platform's recommendation system.

4.1 The University of Pennsylvania Study

Published in 2024, the University of Pennsylvania Computational Social Science Lab study represents one of the most sophisticated attempts to disentangle user behavior from algorithmic influence^[10].

The researchers created bots trained on the actual YouTube watch histories of **87,988 real users** collected from October 2021 to December 2022. These bots were then assigned to different experimental conditions:

- **Control bots:** Continued watching videos from the user's actual history
- **"Up Next" bots:** Always selected the first video from sidebar recommendations
- **Random sidebar bots:** Randomly selected from the top 30 sidebar recommendations
- **Homepage bots:** Randomly selected from the top 15 homepage recommendations

The key finding challenged the radicalization narrative: counterfactual bots following recommendations consumed *less* partisan content than the corresponding real users. The researchers interpreted this as evidence that users have "intrinsic preferences" for partisan content that exceeds what the algorithm would recommend on its own^[10].

This study's methodology addressed a key limitation of earlier research: the inability to separate what the algorithm recommends from what users choose to watch. By creating counterfactual scenarios where bots followed different rules, the researchers could isolate the algorithm's true influence.

4.2 The Princeton/UC Davis Study

The 2023 study by researchers from Princeton University and UC Davis took a different approach, creating sock puppet accounts across the political spectrum and measuring the ideological slant of recommendations^[2].

The researchers compiled a list of **4,150 problematic YouTube channels** known to promote extremist content. They then tracked whether videos from these channels appeared in recommendation trails for sock puppets with different political profiles.

The study's methodology was notable for its scale and comprehensiveness. As researcher Manoel Ribeiro of EPFL commented: "Their findings indicate that, indeed, following the 'algorithmic rabbit hole' increases the recommendation of extreme content. This paper stands out because the authors conducted a high quality, comprehensive audit on an unprecedented scale"^[2].

However, the study also had important limitations. The researchers acknowledged that using simulated users rather than real ones means "it is unclear to what extent their experiment translates to the experience of someone using YouTube." Real users have agency—they don't blindly follow recommendations as sock puppets do.

4.3 Misinformation Filter Bubbles

A 2022 study published in the Proceedings of the ACM on Human-Computer Interaction examined how quickly users can enter "misinformation filter bubbles" and what it takes to escape them^[14].

The researchers used sock puppet methodology to have agents watch misinformation-promoting content, then attempt to "burst the bubble" by watching debunking content. They recorded

17,405 unique videos, manually annotating 2,914 for misinformation content.

Key findings included:

- Filter bubbles don't always form, but when they do, they can be burst by watching debunking content
- The effect varies significantly by topic
- Watching debunking videos after misinformation videos caused a "sudden decrease" in misinformation recommendations

- The researchers did not observe significant improvements in the overall quantity of recommended misinformation compared to earlier studies

A 2024 study by the Institute for Strategic Dialogue (ISD) expanded this research by creating profiles tailored to specific ages and genders, monitoring what content was recommended when showing initial interests toward mainstream categories like gaming, lifestyle content, and news^[15].

The ISD study found that regardless of profile age or interests, misinformation and problematic content eventually surfaced through YouTube's algorithm. The researchers highlighted "shocking" findings regarding violent and sexually explicit content recommended to profiles set up for 14-year-old boys and girls in the context of video games.

"We need more data transparency, more transparency from the platforms—give researchers and journalists and academics access to the data around the algorithm so we can actually understand how they work."

—Aoife Gallagher, ISD Senior Analyst^[15]

5. Content Moderation and Policy

YouTube's approach to content moderation has evolved significantly in response to criticism and controversies. The platform now employs multiple layers of safeguards intended to prevent harmful recommendations while maintaining user engagement.

Borderline Content Demotion

YouTube classifies "borderline content" as material that doesn't violate Community Guidelines but is potentially harmful or misleading—conspiracy theories, sensational misinformation, or content that comes close to violating policies. The company states that videos classified as borderline are demoted in recommendations^[5].

This classification relies on:

- **Topic detection algorithms:** Flagging sensitive or high-risk themes
- **Quality scoring models:** Assessing production quality, accuracy, and trustworthiness
- **Human evaluation:** Trained evaluators assess content against publicly available rating guidelines
- **Expert consultation:** Medical doctors for health content, for example

Elevating Authoritative Sources

For topics like health, science, and breaking news, YouTube boosts content from vetted publishers and subject-matter experts. This includes source whitelisting, context-specific ranking overrides, and policy-driven adjustments during high-stakes events like elections or public health crises^[5].

Recent Policy Changes

In December 2024, YouTube quietly rolled back some content moderation policies, raising the threshold for how much potentially prohibited content a video can contain before removal from one-quarter to one-half of a video's content, provided the video is deemed "in the public interest"^[16].

This change, which was not publicly disclosed until reported by the New York Times, reflects a broader shift in how tech platforms approach content moderation amid political pressures. According to Boston University professor Michelle Amazeen, the change "opens the door to messy interpretations of what's in the public interest, and at whose expense"^[16].

6. Regulatory Response: The EU Digital Services Act

The European Union's Digital Services Act (DSA), which entered into force in November 2022 and became applicable across the EU in February 2024, represents the most significant regulatory attempt to mandate algorithmic transparency and accountability^[17].

The DSA imposes specific obligations on "Very Large Online Platforms" (VLOPs)—those with over 45 million EU users—including YouTube, Facebook, Instagram, TikTok, and others^[18].

Key DSA Requirements for Algorithmic Systems:

- **Transparency on recommender systems:** Platforms must clearly explain how their content moderation and algorithmic recommender systems work in their terms of service
- **Alternative recommender options:** Users must be offered at least one alternative feed not based on profiling
- **Advertising transparency:** Clear information about why users were targeted with ads
- **Annual risk assessments:** VLOPs must formally assess how their products, including algorithmic systems, may exacerbate risks to society

- **Independent audits:** Requirements for external oversight of platform practices

The DSA also bans targeting advertisements to children and profiling individuals based on "sensitive" traits like religious affiliation or sexual orientation. It introduces limits on "dark patterns"—design practices that deceive and manipulate users^[17].

As VU Amsterdam expert Sarah Eskens explains: "The law does not determine which content is banned, but forces platforms to be transparent about how they moderate it. If platforms use AI for moderation, they must be transparent about how it works"^[19].

Violations can result in fines of up to 6% of global annual revenue—potentially billions of euros for platforms like YouTube.

7. Impact on the Creator Economy

YouTube's algorithm doesn't just affect viewers—it fundamentally shapes the livelihoods of millions of content creators. The platform's recommendation system determines which creators thrive, which struggle, and which never get discovered at all.

According to industry analysis, YouTube creators can be categorized into tiers based on average views per video^[20]:

Table 2 YouTube Creator Tiers and Characteristics

Tier	Average Views	Algorithm Relationship
Micro Creator	Under 50,000	Highly volatile; algorithm discovery critical
Mid Creator	50,000 - 300,000	Growing algorithmic consistency
Macro Creator	Over 300,000	Algorithm favors; broader audience reach

As one industry executive noted: "Macrocreators drive more efficient reach... YouTube is more likely to recommend their content. Their audience is broader, meaning that for a U.S.-based brand, they might also be paying for some level of international viewership"^[20].

This creates what creators call the "algorithm lottery"—where success depends not just on content quality but on whether the algorithm decides to promote a video. The resulting "YouTube Vibecession" has creators in panic over falling viewership as algorithm changes can devastate established channels overnight^[20].

The fundamental tension is clear: creators must optimize for an algorithm that prioritizes engagement metrics, even when that optimization may conflict with creating genuinely valuable content. This has led to widespread creator burnout and a homogenization of content as creators chase whatever the algorithm currently favors.

8. Conclusion

YouTube's recommendation algorithm represents both a remarkable technological achievement and a profound societal challenge. The system's ability to process billions of signals and deliver personalized recommendations to billions of users is genuinely impressive. Yet the evidence reviewed in this report raises serious questions about whether this power is being wielded responsibly.

What the Evidence Shows:

The research presents a complex picture. YouTube has made genuine efforts to address harmful recommendations—shifting from click optimization to satisfaction metrics, implementing borderline content demotion, and elevating authoritative sources. The company claims these efforts have reduced consumption of borderline content by 70%.

However, independent research tells a more nuanced story. The Mozilla Foundation's investigation found that 71% of regretted videos came from algorithmic recommendations, with 9% of these videos later removed for Community Guidelines violations. Studies have documented how the algorithm can expose children to violent content, promote misinformation, and potentially create filter bubbles.

The academic research on radicalization remains contested. Some studies suggest the algorithm can amplify extreme content, particularly for users with existing right-wing preferences. Others find that user choice plays a larger role than algorithmic influence. What is clear is that the relationship between algorithmic recommendations and user behavior is complex and context-dependent.

Unanswered Questions:

Fundamental gaps in our understanding remain. YouTube does not release data on what percentage of views come from recommendations versus searches. The company does not allow independent researchers to audit

its algorithm. The "violative view rate"—the percentage of views from content that violates guidelines—is disclosed (0.16-0.18%) but cannot be independently verified.

The Path Forward:

The European Union's Digital Services Act represents a promising step toward meaningful algorithmic accountability. By mandating transparency, requiring alternative recommender options, and imposing significant penalties for non-compliance, the DSA creates incentives for platforms to prioritize user wellbeing alongside engagement.

However, regulation alone cannot solve these challenges. What is needed is a fundamental shift in how platforms approach recommendation systems—moving from optimization for engagement metrics to genuine optimization for user flourishing. This requires not just policy changes but changes in corporate culture, business models, and the metrics by which success is measured.

As this report demonstrates, the "dark side" of YouTube's algorithm is not a simple story of evil corporations or malicious AI. It is a story of complex systems with unintended consequences, of business incentives misaligned with user interests, and of the challenges of governing technologies that evolve faster than our understanding of them.

The question is not whether YouTube's algorithm is "good" or "bad." The question is whether we, as a society, are willing to accept a system that shapes what billions of people know and believe based on opaque calculations optimized for engagement rather than human flourishing. The evidence suggests we should not.

References

- [1] Shaped.ai. (2025). How YouTube's Algorithm Works: A Guide to Recommendations. *Shaped Blog*. Retrieved from <https://www.shaped.ai/blog/how-youtubes-algorithm-works>
- [2] Tech Policy Press. (2023). New Study Suggests Right-Wing Bias in YouTube Recommendation Algorithm. Retrieved from <https://techpolicy.press/new-study-suggests-rightwing-bias-in-youtube-recommendation-algorithm>
- [3] Google Support. (n.d.). How YouTube recommendations work. *Google Support*. Retrieved from <https://support.google.com/youtube/answer/16089387>
- [4] Covington, P., Adams, J., & Sargin, E. (2016). Deep Neural Networks for YouTube Recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems*, 191-198. <https://research.google.com/pubs/archive/45530.pdf>

- [5] Goodrow, C. (2021). On YouTube's recommendation system. *YouTube Official Blog*. Retrieved from <https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/>
- [6] Hootsuite. (2025). How the YouTube algorithm works in 2025. *Hootsuite Blog*. Retrieved from <https://blog.hootsuite.com/youtube-algorithm/>
- [7] Agorapulse. (2024). YouTube Algorithm in 2024: Everything You Need to Know. Retrieved from <https://www.agorapulse.com/blog/youtube/youtube-algorithm-in-2024-everything-you-need-to-know/>
- [8] Mozilla Foundation. (2021). YouTube Regrets: A crowdsourced investigation into YouTube's recommendation algorithm. Retrieved from <https://www.mozillafoundation.org/en/youtube/findings/>
- [9] NBC News. (2021). YouTube's recommendations still push harmful videos, crowdsourced study finds. Retrieved from <https://www.nbcnews.com/tech/tech-news/youtubes-recommendations-still-push-harmful-videos-crowdsourced-study-rcna1355>
- [10] University of Pennsylvania Annenberg School. (2024). The YouTube Algorithm Isn't Radicalizing People. Retrieved from <https://www.asc.upenn.edu/news-events/news/youtube-algorithm-isnt-radicalizing-people>
- [11] Wired. (2019). YouTube Has Kid Troubles Because Kids Are a Core Audience. Retrieved from <https://www.wired.com/story/youtube-kid-troubles-kids-core-audience/>
- [12] PBS NewsHour. (2023). YouTube algorithm sends gun videos to kids as young as 9, study finds. Retrieved from <https://www.pbs.org/newshour/economy/youtube-algorithm-sends-gun-videos-to-kids-as-young-as-9-study-finds>
- [13] Radesky, J., et al. (2020). Algorithmic Content Recommendations on a Video-Sharing Platform Used by Children. *Pediatrics*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11137630/>
- [14] Srba, I., et al. (2022). Auditing YouTube's Recommendation Algorithm for Misinformation Filter Bubbles. *Proceedings of the ACM on Human-Computer Interaction*. <https://arxiv.org/abs/2210.10085>
- [15] Institute for Strategic Dialogue. (2024). ISD study reveals how YouTube's algorithm pushes problematic content despite user interest. Retrieved from <https://www.isdglobal.org/isd-in-the-news/isd-study-reveals-how-youtubes-algorithm-pushes-problematic-content-despite-user-interest/>
- [16] Boston University. (2025). YouTube Is the Latest Media Platform to Loosen Content Moderation. *BU Today*. Retrieved from <https://www.bu.edu/articles/2025/youtube-content-moderation-changes/>
- [17] AlgorithmWatch. (2023). A guide to the Digital Services Act, the EU's new law to rein in Big Tech. Retrieved from <https://algorithmwatch.org/en/dsa-explained/>
- [18] New America. (2021). Regulating Platform Algorithms: Approaches for EU and U.S. Policymakers. Retrieved from <http://newamerica.org/oti/briefs/regulating-platform-algorithms/>
- [19] VU Amsterdam. (2025). The Digital Services Act: how the EU is tackling Big Tech. Retrieved from <https://vu.nl/en/stories/the-digital-services-act-how-the-eu-is-tackling-big-tech>

- [20] Digiday. (2026). Not all creators are the same: How the creator economy breaks down by business model. Retrieved from <https://digiday.com/media/not-all-creators-are-the-same-how-the-creator-economy-breaks-down-by-business-model/>