



March 2020

Why Am I Seeing This?

How Video and E-Commerce Platforms Use Recommendation Systems to Shape User Experiences

Spandana Singh

Acknowledgments

In addition to the many stakeholders across civil society and industry who have taken the time to talk to us about ad targeting and delivery over the past few months, we would particularly like to thank Dr. Nathalie Maréchal from Ranking Digital Rights for her help in closely reviewing and providing feedback on this report. We would also like to thank Craig Newmark Philanthropies for its generous support of our work in this area.

About the Author(s)

Spandana Singh is a policy analyst in New America's Open Technology Institute.

About New America

We are dedicated to renewing America by continuing the quest to realize our nation's highest ideals, honestly confronting the challenges caused by rapid technological and social change, and seizing the opportunities those changes create.

About Open Technology Institute

OTI works at the intersection of technology and policy to ensure that every community has equitable access to digital technology and its benefits. We promote universal access to communications technologies that are both open and secure, using a multidisciplinary approach that brings together advocates, researchers, organizers, and innovators.

Contents

Introduction	6
An Overview of Algorithmic Recommendation Systems	8
Case Study: YouTube	10
A Technical Overview of YouTube’s Recommendation System	11
Controversies Related to YouTube’s Recommendation System	17
User Controls Related to YouTube’s Recommendation System	21
Case Study: Amazon	22
A Technical Overview of Amazon’s Recommendation System	23
Controversies Related to Amazon’s Recommendation System	28
User Controls Related to Amazon’s Recommendation System	31
Case Study: Netflix	32
A Technical Overview of Netflix’s Recommendation System	33
Controversies Related to Netflix’s Recommendation System	39
User Controls Related to Netflix’s Recommendation System	41

Contents Cont'd

Promoting Fairness, Accountability, and Transparency Around Algorithmic Recommendation Practices	42
Recommendations for Internet Platforms	42
Recommendations for Policymakers	45

Introduction

Today, personalized recommendation systems, particularly those that are based on machine learning (and the choice architecture decisions associated with them),¹ have come to govern many internet platforms, including social media, e-commerce, and media streaming platforms. Recommendation systems are algorithmic tools that internet platforms use to identify and recommend content, products, and services that may be of interest to their users. These systems are responsible for recommending a range of content including friends, posts, ads, news articles, trending topics, items to purchase, jobs, and more. In doing so, these systems are able to influence user interests, opinions, and behaviors as well as their social group formation.²

Many internet platforms assert that these systems enhance users' experiences through personalized and relevant recommendations. However, it is important to note that in deploying these systems, internet platforms also seek to retain user attention on their services. This translates to significant financial benefits for the companies, as they can then target these users with advertisements and recommend further content to consume or items to purchase.³ In addition, definitions of relevance vary across platforms and are largely based on what a platform believes a user is interested in through its data collection and inference practices.

Widely used by internet platforms today, recommendation systems have a significant amount of influence over how users engage with—and are influenced by—the online sphere.⁴ For example, recommender systems have the power to influence product purchases. They can also determine what content—such as which news articles—a user sees. This power has raised concerns around the use of algorithmic recommendation systems to intentionally or unintentionally create echo chambers in which users have a homogenized experience and engage with only certain viewpoints, or only with popular or trending topics.⁵

In addition, researchers have found that these recommender systems create a number of concerning outcomes. Notably, these include reinforcing societal biases and augmenting harmful perspectives, such as those of extremists and conspiracy theorists. Internet platforms that deploy these recommendation systems do not currently provide meaningful transparency and accountability around how these systems are created, how they operate, and how they make decisions.⁶ This makes it very difficult to analyze and combat the problematic recommendations that come from these systems.⁷ Because of this, critics have called recommendation systems “the biggest threat to societal cohesion on the internet” and a major contributor to offline threats.⁸

Further, recommender systems now also influence the operations of internet platforms themselves. For example, platforms such as Amazon and Netflix produce films and television shows based on behavioral data collected on their users through these systems. As a result, these recommender systems are not only influencing what existing content users see and engage with online, but they are also shaping the database of options that users have to choose from.⁹ To the extent that these productions are based on popularity signals, this could create a feedback loop that narrows the choices available to users.

This report is the final report in a series of four reports that explore how major technology companies rely on automated tools to shape the content we see and engage with online, and how internet platforms, policymakers, and researchers can promote greater fairness, accountability, and transparency around these algorithmic decision-making practices. This report focuses on the use of automated tools to provide recommendations to users. It relies on case studies on three internet platforms—YouTube, Amazon, and Netflix—to highlight the different ways algorithmic tools can be deployed by technology companies to enable recommendations. These case studies will also highlight the challenges associated with these practices.

Editorial disclosure: This report discusses policies by Google (YouTube), which is a funder of work at New America but did not contribute funds directly to the research or writing of this report. New America is guided by the principles of full transparency, independence, and accessibility in all its activities and partnerships. New America does not engage in research or educational activities directed or influenced in any way by financial supporters. View our full list of donors at www.newamerica.org/our-funding.

An Overview of Algorithmic Recommendation Systems

There are three main types of algorithms that can be used in recommendation systems:

1. Content-based Recommender Systems: Content-based recommender systems operate by suggesting items to a user that are similar in attributes to items that a user has previously demonstrated interest in.¹⁰ Content-based recommender systems evaluate the attributes of an item, such as its metadata (e.g. tags or text).¹¹ Although different recommendation systems can vary in their exact composition, most content-based recommender systems operate by creating a profile of a user that outlines their interests and preferences, such as alternative rock music or drug store makeup. The system will then compare this user profile to existing items in its database in order to identify a match.¹² A user's profile is based on explicit and implicit feedback that a user provides on recommendations as a way to indicate their preferences and interests further.¹³ For example, a user may explicitly indicate preferences by rating a particular item, or implicitly by clicking on certain suggested items and not others. The algorithm uses this information to categorize these preferences and refine recommendations going forward. Unlike other systems discussed below, content-based recommender systems do not typically account for the preferences or actions of other users when making recommendations. Rather, their recommendations are primarily based on a given user's past interactions.¹⁴

2. Collaborative-filtering Recommender Systems: Collaborative-filtering recommender systems operate by suggesting items to a user based on the interests and behaviors of other users who are identified as having similar preferences or tastes.¹⁵ The process is considered collaborative because these recommender systems make automated predictions (known as filtering) about a user's interests and preferences based on data about other user's who have similar preferences and interests.¹⁶ Collaborative-filtering recommender systems make these determinations by mining user behavior, such as purchase records and user ratings,¹⁷ and through techniques like pattern recognition.¹⁸

Collaborative-filtering recommender systems are considered particularly useful because they allow comparison and ranking of completely different types of items. This is because these algorithms do not need to know about the attributes of an item. Rather, they only need to know which items are bought together.¹⁹ In addition, because collaborative-filtering recommender systems make recommendations based on a vast dataset of other user behaviors and preferences, researchers have found that this approach yields more accurate results compared to other techniques. For example, keyword searching offers a more narrow assessment of datasets.²⁰ Keyword searching is typically deployed

during behavioral targeting, which is when advertisers use information on a user's browsing history and behavior to customize the ad targeting and delivery process.²¹

In addition, some researchers consider collaborative-filtering recommender systems to be able to deliver more accurate results for users with both mainstream and niche interests when compared to the other types of recommender systems outlined in this report. This is because a programmer can control how many users in the database should be considered as part of the data set for calculating recommendations; as a result, the programmer can optimize the algorithm to balance recommending popular and niche results.²² According to researchers, collaborative-filtering recommender systems are modeled after the ways individuals solicit feedback and recommendations from their social circle offline.²³ Such recommender systems seek to automate this process based on the notion that if two similar individuals like an item or piece of content, there are also likely several other items that they would both find interesting.²⁴

Collaborative-filtering recommender systems rely on two primary types of algorithms: user-based and item-based collaborative-filtering algorithms. Both types rely on users' ratings on items. In the user-based category of algorithms, the algorithm scores an item a user has not rated by combining the ratings of similar users.²⁵ In the item-based approach, the algorithm matches a user's purchased and rated items with similar items, and then combines these similar items into a list of recommended products for the initial user.²⁶

3. Knowledge-based Recommender Systems: Knowledge-based recommender systems make suggestions based on the attributes of a user and an item. These systems typically rely on data-mining methods and advanced natural language processing (NLP) to identify and evaluate an item's attributes (e.g. price or technical specifications such as HD or BluRay functionality). The system then identifies similarities between Item A's attributes and User A's preferences (e.g. preference for high-end equipment), and makes recommendations based on its findings. For example, such a system could identify attribute similarities between a job seeker's resume and a job description. This recommender system does not typically consider a user's past behaviors, and it is therefore considered most effective when engaging with a new user or a new item.²⁷

Most recommendation systems deploy a hybrid of these three filter categories.²⁸ Traditionally, both content-based and collaborative-filtering recommender systems rely on explicit input from a user. For example, algorithms can collect user ratings, likes, or reactions to an item or piece of content. These systems can also use implicit user input, which is drawn from data on user activities and behaviors.²⁹ These can include a user's clicks, search queries, and purchase history, as well as a user adding an item to their cart, completing a purchase, and reading an entire article.³⁰ Using both of these data types, these systems can be refined so that they deliver more personalized recommendations.

Case Study: YouTube



Source: PixieMe / Shutterstock.com

YouTube is an American video-sharing company founded in 2005 by Chad Hurley, Steve Chen, and Jawed Karim.³¹ In 2006, YouTube was acquired by Google for \$1.65 billion, and has since operated as a subsidiary of the company.³² YouTube is the world's largest online video source,³³ with approximately 2 billion users worldwide.³⁴ The company currently ranks second for global internet engagement on Alexa rankings.³⁵ According to a Pew Research Center study, 94 percent of Americans between the ages of 18 and 24 use YouTube, a higher percentage than for any other online platform.³⁶

Today, individuals turn to YouTube to access a range of content, including music videos, instructional videos, and the news. The company operates a vast database of videos, and has been referred to as a library of content.³⁷ YouTube utilizes an algorithmic recommendation system to generate personalized video recommendations to its users.³⁸ According to YouTube, although many users visit the platform to search for something specific, the company has expanded its recommendation system in order to also engage those who did not come to the platform with a specific idea of what they wanted to watch.³⁹ YouTube's videos also often appear in Google search results.⁴⁰ YouTube seeks to maximize the time that users spend on the platform as it enables the company to deliver more ads to users. Given that the recommendation system is designed to infer user interests and behaviors, and subsequently suggest content that may be of interest

to a user, the system is part and parcel of the company's revenue generation model.

YouTube's recommendation system determines what content should appear on a user's home page and in the user's "Up Next" sidebar, which appears next to videos that a user is currently watching. The Up Next feature autoplays recommended content unless a user turns the autoplay off.⁴¹ Today, YouTube's recommendation system is responsible for generating over 70 percent of viewing time on the platform.⁴² This has a significant impact on its users. According to a Pew Research Center study, 81 percent of YouTube users say that they at least occasionally watch recommended videos, including 15 percent who say they watch recommended videos regularly.⁴³

YouTube is one of the largest video repositories on the internet, and many users incorrectly equate the site's popularity with the credibility of its recommendation system. However, despite the fact that YouTube's recommendation system is responsible for shaping how billions of individuals engage with content on the service, and influencing how they see the world, the company has provided relatively little transparency around how this system works.⁴⁴ According to YouTube, user recommendations and search results are influenced by factors such as the videos a user has liked, the playlists a user has created,⁴⁵ and a user's watch history and activity on YouTube, Google, and Chrome. Some researchers have suggested that the system also considers data points such as a user's account preferences⁴⁶ and the keywords they search for.⁴⁷ The company has not, however, offered comprehensive disclosures outlining the key factors its recommendation system considers.⁴⁸ This lack of transparency is concerning, as the company's recommendation system has been found to suggest controversial and harmful videos, including those that promote extremist propaganda, conspiracy theories, and misinformation. Further, YouTube provides users with only a limited set of controls over how they would like their platform experience to be shaped by such algorithmic decision-making. Without insight into how YouTube's recommendation systems work, it is difficult to understand why these suggestions are made, and how to develop targeted interventions to prevent them.

A Technical Overview of YouTube's Recommendation System

According to a 2016 paper authored by three researchers at Google, in order for YouTube's recommendation system to deliver personalized recommendations, it has to be able to process YouTube's expansive user base and collection of videos.⁴⁹ In addition, given that over 500 hours of new content are uploaded to YouTube every minute,⁵⁰ the recommendation system also needs to be responsive enough that it can rapidly integrate these new videos as well as any new user behaviors and patterns into its suggestions.⁵¹

According to YouTube, it makes minor changes to the recommendation system every year.⁵² But, the company has provided little transparency around how its recommendation system is structured and makes decisions, and how the system has changed over time. However, numerous researchers and journalists have attempted to document the system's various iterations and evolutions.

Prior to March 2012, the recommendation algorithm was designed to maximize user views by recommending videos that the system calculated users were likely to click on. However, many creators figured out how to influence this recommendation system and gain more views on their videos. In addition, the prioritization of user views by the algorithm meant that creators had a greater incentive to produce clickbait content⁵³ that garnered a large number of clicks, such as content with sensational titles, compared to content that a user would actually want to fully watch.⁵⁴

In response to these concerns, YouTube altered the recommendation algorithm so that it placed more weight on a user's watch time rather than a video's views.⁵⁵ The platform defines watch time as how much time a user spends viewing content on the platform. YouTube asserts that this change encouraged creators to produce "higher quality content" that users would watch fully, rather than content users would click on and then abandon.⁵⁶ This would in turn increase the likelihood that users would be satisfied with the service, view more videos and advertisements, and generate more revenue.⁵⁷

The introduction of the watch time metric also influenced how the company displays videos in search results, runs ads, and pays video creators on the platform.⁵⁸ After introducing this new model, the company also changed their rules so that all creators—rather than a vetted group—could run ads on their content and accrue revenue from them.⁵⁹ A few weeks after the company introduced these changes to its recommendation system, YouTube reported that the number of views on the platform was decreasing.⁶⁰ The overall watch time, however, was increasing and it grew 50 percent a year for three consecutive years.⁶¹ However, critics both inside and outside the company argued that this metric also rewarded offensive, harmful, and often fringe content that garnered high watch times.⁶²

In 2015, Google's artificial intelligence division, Google Brain, began reconstructing YouTube's recommendation system around neural networks. A neural network is a series of algorithms that aims to identify relationships by finding patterns in a dataset by mimicking how animal brains operate.⁶³ Prior to Google Brain, YouTube had implemented machine learning tools in their recommendation system through a Google-produced system known as Sibyl.⁶⁴ However, the new algorithm introduced by Google Brain brought in a range of new functionalities. For example, Google Brain used supervised learning systems, which operate using inputs (training data sets that are pre-labeled by

humans) and predetermined output data.⁶⁵ These supervised learning techniques enabled the system to identify adjacent relationships between videos, and generalize these findings in ways that humans could not. Before the use of such supervised learning techniques, if a user watched an episode from a particular beauty vlogger, the recommendation system suggested videos with high degrees of similarity. However, by identifying adjacent relationships, the Google Brain model was able to suggest other vloggers who were comparable, but not exactly identical.

In addition, the Google Brain algorithm was able to identify important patterns in consumption. For example, the algorithm noted and optimized a relationship between a user's device and their watch times, recommending shorter videos for mobile users and longer videos for YouTube TV users.⁶⁶ The Brain algorithm also enabled YouTube to incorporate insights on a user's behavior into its recommendations at a faster rate, thus making it easier for the company to identify trending topics and offer updated recommendations.⁶⁷ Following the introduction of the Google Brain model, 70 percent of the time that users spend on the site consuming content has been driven by the recommendation system.⁶⁸

This recommendation system was created by combining two deep-learning neural networks: one for candidate generation and another for ranking. Candidate generation is the first stage of the recommendation process. During this stage, the recommendation system is given a query and produces a group of relevant videos as candidates. The ranking network comes second, and is responsible for ranking these candidates in order. The candidate generation network uses information on a user's behaviors and history on the platform to identify a small group of a couple hundred videos from YouTube's larger corpus that are considered broadly of interest to the user. The candidate generation network relies on collaborative filtering to produce these personalized results. The ranking network is then tasked with delivering a select number of best recommendations to each user. It does this by assigning each video a score based on information about the video (e.g. length) and information about the user (e.g. whether they watch long videos or short videos). The videos that are assigned the highest scores are then ranked and displayed to the user. During the ranking stage, the model has greater access to information on a specific video and a user's relationship to the video than it does during the candidate generation stage, as the model only needs to consider a small group of a couple hundred videos. In simplified terms, the ranking function can be thought of as expected watch time per impression. Researchers from Google have asserted that this formula promotes more "relevant" videos to users compared to ones that emphasize click-through rate, as click-through rate functions often result in the promotion of clickbait.⁶⁹

This two-stage model enables YouTube to make personalized video recommendations from a large database of content.⁷⁰ In addition, these deep neural networks create an average of a user's search history and watch history in

order to make recommendations. It also considers additional data points, including a user's geographic region, device, gender, logged-in status, and age.⁷¹ YouTube's recommendation system has undergone a number of changes over the past few years. It is unclear how much of Google Brain's changes are still in use.

YouTube evaluates and refines this recommendation system using a range of offline metrics, such as precision, ranking loss, and recall.⁷² The company also runs A/B tests⁷³ during live experiments. During these experiments, researchers can measure minor changes in click-through rate, watch time, and other user engagement metrics.⁷⁴ This method of testing is considered the gold standard for evaluating the effectiveness of a recommendation system, compared to offline testing, which has a number of problems.⁷⁵ However the company has not provided any insight into the results of such tests and how they contribute to the company's assessment of how effective its recommendation system is. The examples that the recommendation system model has been trained on consist of more than just videos that the system recommended. It features videos from all YouTube watches, including embedded videos from other sites, so as to ensure the recommendation system can surface new content (especially content that may not be widely viewed yet, but is still considered of interest to a user).⁷⁶ The company noted that it has been able to identify other mechanisms by which users are discovering new content, and integrate it into the model, although it did not specify what these mechanisms are.⁷⁷

This new two-stage algorithm also created new problems. It often forced users into specific content niches by consistently recommending content that was similar in nature to videos the user had previously watched. As a result, users got bored. This prompted researchers at Google Brain to explore whether they could maintain user engagement by guiding them to content in other sectors of the platform, rather than just in existing interest buckets. These questions led the company to test a new algorithm, which incorporated a type of artificial intelligence known as reinforcement learning.

Reinforcement learning is used to train machine-learning models to make a certain sequence of decisions using a trial and error process that features rewards and penalties.⁷⁸ The team called this new algorithm Reinforce. Its primary goal was to predict which video recommendations would broaden the range of subjects that a user would watch content on, pushing users to consume more content and maximizing user engagement over time. A YouTube spokesperson also said that Reinforce was intended to improve the accuracy of recommendations on the platform by mitigating the recommendation system's bias toward popular content.⁷⁹ Google considered the introduction of Reinforce a massive success. Sitewide views across YouTube increased by almost 1 percent, a staggering amount given the platform's size. This gain translated into millions more hours of watch time and a significant bump in the company's ad revenue.⁸⁰

As YouTube's algorithm has evolved, the company has shared that the fundamental components of this model remain intact today.⁸¹

According to a YouTube spokeswoman, in late 2016, the company adopted social responsibility as a core value for the company.⁸² During this time, the recommendation system was altered so it considered inputs such as how many times a video was shared, liked, and disliked.⁸³ These changes were introduced amidst growing pressure on internet platforms to be more proactive in their efforts to combat harmful content, such as extremist propaganda, disinformation and misinformation, and content unsafe for children.⁸⁴ Recently, the company has provided more detail around this concept of responsibility, outlining that it consists of the four Rs of responsibility:⁸⁵

- Removing content that violates the platform's Community Guidelines as quickly as possible
- Raising up authoritative information sources, especially during breaking news moments
- Reducing the spread of content that comes close to violating, but does not violate the platform's Community Guidelines (known as borderline content)
- Rewarding trusted creators

Also in 2016, a YouTube spokesperson stated that the recommendation system had changed significantly, and was no longer geared to optimize for watch time. Rather, the system began to emphasize satisfaction to ensure users were happy with the content they were viewing,⁸⁶ and so that the recommendation system would suggest clickbait videos less often.⁸⁷ This new metric aimed to balance watch time with factors such as likes, dislikes, shares, and satisfaction surveys that the company prompts users to fill out after they finish certain videos on the platform.⁸⁸ According to the company, it receives millions of survey responses every week.⁸⁹

Further, between 2017 and 2019, the company introduced two new internal metrics for evaluating how videos are performing on the site. The first metric monitors the total amount of time users are spending on YouTube, including by posting and reading the comments. The second metric, known as "quality watch time," aims to identify content that goes beyond just retaining a user's attention; the company has not explained what this involves. In calculating these two new metrics, YouTube aimed to reward content considered generally acceptable to YouTube users and advertisers and push back on criticisms that it uses its algorithmic recommendation system to capture user attention and make the platform more addictive.⁹⁰ A recent Pew Research Center study indicates,

however, that the algorithmic system still seeks to reel users in and get them to consume more content. In addition, the study found that the longer a user spends watching videos on the platform, the more the system recommends longer and more popular content. In the initial stages of the Pew study, the recommendation engine suggested videos that were on average nine minutes and 31 seconds in length. During the final stages of the study, the recommended videos were on average 15 minutes in length.⁹¹

Since 2017, YouTube has introduced changes to its recommendation algorithm designed to promote videos from sources that the company considers to be authoritative,⁹² such as top and local news channels.⁹³ YouTube determines which sources it considers to be authoritative based on Google News's assessments of publishers and whether they abide by Google News's content policies and are producing reliable content. These sources can also include organizations such as public health institutions.⁹⁴ A publisher's status as an authoritative source is not dependent on how many subscribers its YouTube channel has.⁹⁵ If a publisher is considered authoritative, and it has a YouTube channel, then YouTube will promote this publisher's related content when a user searches for information or news related queries.⁹⁶ This is intended to deter the spread of misinformation and conspiracy theories on the platform.⁹⁷

In June 2019, the company also began promoting authoritative sources in cases in which a user has consumed multiple videos that are close to violating the platform's Community Guidelines,⁹⁸ such as conspiracy theory videos,⁹⁹ as well as queries related to election news.¹⁰⁰ YouTube also promotes authoritative sources in its top news and breaking news shelves for more recent events.¹⁰¹ The company shared that it could expand this to other categories of content, such as entertainment. However, this comes with trade-offs: It is challenging to define authoritative sources across more subjective verticals, as these determinations are based on personal preference and taste.¹⁰²

In January 2019, YouTube also altered its recommendation algorithm to reduce suggestions of "borderline" videos, such as harmful content and misinformation.¹⁰³ According to YouTube, this change resulted in a 50 percent drop in watchtime for this type of content.¹⁰⁴ However, this data has not been verified by independent researchers.¹⁰⁵ The company implemented these changes using machine learning, as well as human evaluators and experts across the United States. These evaluators are responsible for providing input on the quality of the videos they review.¹⁰⁶ This data is then used to train the machine-learning recommendation-generation systems.¹⁰⁷ According to YouTube, the human evaluators themselves are trained using precise guidelines.¹⁰⁸

The company stated that it would roll out these efforts to other countries to minimize recommendations of harmful content as its systems become more accurate in implementing this rule in the United States.¹⁰⁹ According to a

company blog post, this change was anticipated to impact less than 1 percent of the videos on YouTube and would only affect recommendations of these borderline videos, not the availability of this content on the platform as a whole. This means that users who search for or subscribe to channels that post such content would still be able to view these videos.¹¹⁰ According to YouTube, the company took this approach in order to adequately safeguard free expression on the platform.¹¹¹

YouTube collects both explicit and implicit feedback from its users. Explicit feedback is collected through the thumbs up and thumbs down features in the product, as well as product surveys¹¹² that the company runs to find out if a user enjoyed a video that was recommended to them.¹¹³ Some of the implicit data points that the company collects and uses include user activity on YouTube, Google, and Chrome,¹¹⁴ and user watch history.¹¹⁵ YouTube relies on this implicit feedback to inform recommendations, as well as to train the recommendation models.¹¹⁶ In such training processes, the fact that a user has finished a video is a positive signal.¹¹⁷

As described above, YouTube provides very little transparency and accountability around how its recommendation system is structured, how it operates, and how it makes decisions.¹¹⁸ Research has suggested that promoting awareness of the use of algorithmic tools and enabling users to control their own experiences on a platform are fundamental steps in building trust with users. This lack of transparency from YouTube therefore limits the agency users have over their own experiences.¹¹⁹

Controversies Related to YouTube's Recommendation System

In addition, the lack of transparency and accountability from YouTube is also concerning considering the level of controversy and backlash around this system's recommendations. It has also made evaluating these criticisms difficult, as most studies are operating with insufficient data.¹²⁰ This makes it hard to draw reliable conclusions on how YouTube's recommendation system shapes user perceptions and behaviors. In addition, a lack of transparency around the company's training datasets also makes it challenging to run assessments to ensure the system is providing the same level of utility to its diversity of users (e.g. users of different genders and ethnicities),¹²¹ a common problem associated with defining fairness in recommender systems.¹²² This is important as such assessments can identify potential instances of bias within recommender systems.

Over the past decade, the company has come under particular criticism for enabling its recommendation engine to suggest content to users containing misleading or false information and conspiracy theories. For example, after a fire

broke out in Paris's Notre-Dame cathedral in April 2019, a number of conspiracy theories began circulating on the platform, claiming that the fire was an act of terrorism, and promoting Islamophobic rhetoric. One vlogger on YouTube also claimed that the French government had started the fire as a covert operation, and that French President Emmanuel Macron could not be trusted. The video was viewed almost 50,000 times overnight,¹²³ and this expanded to over 100,000 views soon after.¹²⁴ Shortly after the video was posted, it was also monetized through the use of advertisements.¹²⁵ YouTube's recommendation system continued to recommend the video despite changes implemented by the company in January 2019, which aimed to limit algorithmic promotion of conspiracy theory content and instead promote content from authoritative sources.¹²⁶ Although YouTube has stated that conspiracy theory videos make up less than 1 percent of all content on the platform, this is still a staggering amount of content, and the problem is compounded whenever the recommendation algorithm promotes this content.

YouTube has also introduced algorithmically-recommended information panels with links to third-party sources that offer to fact-check video content. These panels appear next to content featuring common conspiracies, as well as content posted by some foreign state-run media outlets.¹²⁷ However, journalists found that the panels the company appended to livestreams of the Notre-Dame fire rerouted users to information about the 9/11 terrorist attacks in the United States instead.¹²⁸

Similarly, in 2017, academics and other commenters spotlighted the role of the company's search and recommendation algorithms in promoting disinformation during the 2016 U.S. presidential election. Based on these findings, prominent sociologist and technology critic Zeynep Tufekci termed these algorithms "misinformation engines."¹²⁹

In addition, YouTube has faced particular criticism for creating a "rabbit hole" effect, in which the algorithm delivers personalized recommendations that prompt users to consume harmful or radical content¹³⁰ that they did not originally seek out.¹³¹ In 2019, Mozilla began publishing anecdotes of how anonymous users encountered such rabbit holes on the platform under a project known as YouTube Regrets. The project aimed to push YouTube to let independent researchers study their algorithmic decision-making systems.¹³² In September 2019, YouTube representatives met with the Mozilla Foundation to discuss the issues raised in the campaign.¹³³ In addition, a YouTube spokesman told CNET that while the company welcomed research on these issues, it had not seen the videos, screenshots, or data that Mozilla was using and was therefore unable to review Mozilla's claims.¹³⁴

The Media Manipulation Initiative at Data & Society Research Institute spearheaded a research project exploring the far-right in the United States and

Germany, through which they examined YouTube's recommendation system. Their research uncovered that the system concerningly combined communities associated with Fox News and GOP accounts with communities associated with conspiracy theory channels, such as those belonging to far-right commentator Alex Jones. Similarly, the researchers found that the recommendation system categorized communities associated with the religious right together with communities associated with the international right-wing. The researchers outlined that these categorizations could create a rabbit hole effect, because if a user is consuming content produced by conservative groups on the platform, they are only a few clicks away from receiving recommendations for content produced by far-right extremist groups.¹³⁵ In addition, the researchers raised concerns that this grouping also generates a filter bubble on the platform, in which users are prevented from accessing information that may challenge their perspectives or broaden their horizons because of a system's predictions on what they will like.¹³⁶

Several researchers, including former YouTube employee-turned-critic Guillaume Chaslot, have argued that it is in the company's business interest to promote such polarizing and fringe videos and channels, as they drive engagement and greater watch times.¹³⁷ Further, critics such as Chaslot have suggested that the recommendation system is biased toward promoting divisive, sensational, and conspiratorial content,¹³⁸ perhaps because the system has learned that such content is engaging.¹³⁹ Given the vast number of users who consume recommended content, this raises significant concerns about the platform serving as a radicalization pipeline.¹⁴⁰

YouTube executives have contested these notions, claiming that the company considers more than just watch time when making content recommendations, and that because advertisers do not want their content to appear alongside such harmful content, there is no financial interest in promoting these videos.¹⁴¹ The company also stated that after a user consumes a video, the recommendation system does not account for whether the content of the video was less or more extreme, and it therefore would not seek to necessarily recommend similar videos. Rather, recommendations in such instances rely on the context and user behavior associated with the consumption of the initial video.¹⁴² YouTube also stated that after reviewing internal testing data, they found that on average, users who watch one extreme video are subsequently recommended more moderate content¹⁴³ suggesting that the rabbit hole effect toward more radical content is not inevitable.¹⁴⁴ However, because the company provides little transparency around the factors that its recommendation system actually does consider, and because the company has not shared any of the data it collected or evaluated during these tests, it is difficult to corroborate these statements.¹⁴⁵

It is important to note that some researchers, such as Penn State Political Scientists Kevin Munger and Joseph Phillips, have also pushed back against the notion that the company's recommendation system is a central component of

online radicalization.¹⁴⁶ They contend that prior studies have not been able to determine that the algorithm has had a noticeable effect on radicalization, and that instead this narrative has been highlighted by policymakers and the media because it offers simple policy prescriptions.¹⁴⁷ These researchers instead suggest that radicalization online is similar to radicalization offline, in that it relies on providing an individual with new, radicalizing information at scale, and that the supply of such content (and the ease with which producers can create content on YouTube) caters to this demand.¹⁴⁸ Other researchers, such as Data & Society's Becca Lewis, also point to the role of other algorithms, like the company's search algorithm, as well as less technical factors, such as online social-networking interactions between creators and audiences, as more significant factors in promoting radicalization. As a result, these researchers suggest that focusing solely on the algorithmic recommendation component of the radicalization process provides a limited view of the overall problem and hinders potential solutions.¹⁴⁹

YouTube has also faced significant backlash over how its recommendation system—which makes content recommendations both on YouTube's main platform as well as on YouTube Kids—interfaces with children. YouTube Kids is a separate video app, curated by humans and algorithms, that features age-tailored content.¹⁵⁰ However, children's videos are some of the most watched categories of content on YouTube's main platform as well. As a result, producing and reproducing popular children's content has emerged as a lucrative business on the service, as it enables creators to reap the benefits of advertising dollars.¹⁵² According to a Pew Research Center Study,¹⁵³ children's videos constituted the majority of the 10-most recommended posts on YouTube,¹⁵⁴ and 80 percent of parents said they occasionally let their children watch content on YouTube.¹⁵⁵

In 2017, the company was hit with the "ElsaGate scandal," in which its recommendation system recommended seemingly child-friendly content featuring characters such as Elsa from Disney's *Frozen*, but that actually contained inappropriate themes related to topics like violence, sex, drugs, and alcohol.¹⁵⁶ In addition, researchers, journalists, and YouTube creators have found that the recommendation engine was suggesting ordinary videos of children that were rampant with sexualized comments as well as comments suggesting timestamps in which children were in sexualized positions.¹⁵⁷ These videos were often recommended by YouTube's system after a user searched for videos of adult women, such as using the term "bikini haul," raising concerns about how the system was making links about searches for adult women and children.¹⁵⁸ In response, the company said they would implement changes such as closing down the comments section of such posts and more rigorously removing posts that were found to violate its Community Guidelines.¹⁵⁹ However, YouTube has said little publicly about how the company's recommendation system would be altered to prevent the promotion of such content going forward.¹⁶⁰

Although YouTube has introduced some technical and policy changes to combat the spread of misinformation, conspiracy theories, and egregious content, numerous reports have circulated, often with employee input,¹⁶¹ claiming that YouTube executives repeatedly ignored warnings and suggestions to alter the company's recommendation system in a more significant manner.¹⁶² This raises concerns that the company is placing profits over ensuring the company's use of automated tools is responsible, transparent, and accountable.

User Controls Related to YouTube's Recommendation System

As highlighted, YouTube does not provide significant transparency around how its recommendation system operates, thus limiting the agency users have over their personal YouTube experience. The company does, however, offer its users a limited set of controls over how this system shapes their platform experience.

As previously mentioned, YouTube users have the ability to turn off the AutoPlay feature.¹⁶³ In June 2019, the company announced that they were expanding the controls users have over the homepage and Up Next recommendations.¹⁶⁴ These changes made it easier for signed-in users to view recommendations on both of these areas of the platform.¹⁶⁵ The changes also let users mark certain channels so that they do not appear in their recommendations. However, if a user subscribes to the channel, searches for the channel, or visits the channels page, they will still see its content. In addition, if the channel appears in the Trending tab, the user will still see its content.¹⁶⁶

YouTube's expansion of controls also enables users to learn why a video may have been suggested to them, particularly on the homepage.¹⁶⁷ Users can also remove specific videos from their watch history and specific queries from their search history to prevent these data points from being considered in recommendations. They can also pause their watch and search history, or clear them altogether.¹⁶⁸ Further, users can remove videos, channels, sections, and playlists from their homepage, and indicate that they are not interested in this content or do not want to be recommended content based on these factors.¹⁶⁹ They can also remove liked videos from their playlists or edit and delete playlists to further control their recommendations.¹⁷⁰

If a user wants to revert to having all the information YouTube has collected on their behaviors and interests used for personalizing their recommendations, they can clear their "not interested" and "don't recommend channel" feedback through a tool in their My Activity tab.¹⁷¹

Case Study: Amazon



Source: Puckpao / Shutterstock.com

Amazon is an American technology company that offers services such as e-commerce, cloud computing, and streaming. The company was founded in 1994 by Jeff Bezos¹⁷² and is considered the largest online marketplace in the world in terms of revenue and market capitalization, and the largest internet company in the world in terms of revenue.¹⁷³ The company currently ranks fourteenth for global internet engagement on Alexa rankings.¹⁷⁴ In 2019, Amazon accounted for 37.7 percent of all U.S. e-commerce sales.¹⁷⁵

Amazon deploys recommendation systems across a number of its services, including its e-commerce platform and its streaming platform. Given the large scale influence the introduction of Amazon's recommendation system has had on online commerce, this analysis will focus on Amazon.com, the company's e-commerce platform. Amazon generates revenue from its e-commerce platform through product sales, targeted advertising, subscriptions (e.g. Amazon Prime), and the fees sellers pay in order to be able to sell on the platform. The longer a user spends on the platform, the more products they will see and potentially buy and the more ads they will see. Given that the company's recommendation system is designed to understand and predict user interest and behaviors, and make recommendations based on these insights, it is an integral tool for driving user purchases and increasing and maintaining user attention and engagement on the platform. Therefore, the recommendation system is an important contributor to revenue generation on the platform. In addition, some researchers

and commentators have argued that Amazon uses its recommendation system to promote its own brands over others, thereby further increasing its profits, maintaining its market dominance, and raising antitrust concerns.¹⁷⁶

Amazon's e-commerce recommendation engine is powerful, and it engages with users at every stage of their journey on the website. It can therefore influence everything from what products a user sees to which items they eventually buy. Despite the extensive influence this system has over user behaviors and purchasing decisions, Amazon has provided little transparency around how these algorithmic systems are designed and how they operate. This is especially concerning given that numerous researchers and journalists have highlighted how the platform's recommendation engine has suggested products that are misleading, false, or conspiracy-theory based. This lack of transparency therefore makes it challenging to understand why these recommendations are made, and how to prevent them going forward. In addition, the platform offers users a limited set of controls around whether and how they would like their platform experience to be shaped by such algorithmic decision-making processes.

A Technical Overview of Amazon's Recommendation System

Amazon first deployed a recommendation system across its e-commerce platform almost two decades ago.¹⁷⁷ Before deploying this system, Amazon made product recommendations to users based on human curation and best-seller lists. However, according to Amazon, this approach was found to be inherently biased and did not sufficiently provide recommendations to users with niche interests.¹⁷⁸ The company later developed and deployed an algorithmic system that matches a user's purchased and rated items with similar items, and then combines these similar items into a list of recommended products for a user.¹⁷⁹ This unique approach to making recommendations became known as "item-based collaborative filtering" or "item-to-item collaborative filtering."¹⁸⁰

According to a 2001 paper authored by three Amazon employees, the company opted to use the item-to-item collaborative-filtering model rather than a traditional collaborative-filtering model, or models such as content-based models. This is because its algorithm's online computation grows at a rate that is not connected to the growth in the number of customers and items in the product catalog. As a result, this model is able to generate real-time recommendations, scale to large data sets, and produce recommendations that are more likely to be of interest to users. This model is also less computationally expensive than the other models previously outlined.¹⁸¹ When a model is computationally expensive it requires a considerable amount of resources to complete. These resources include the overall run time, processing power, and memory usage required to complete a function. In addition, researchers and journalists have found that the item-to-item collaborative-filtering model helped the company recommend

niche items to shoppers in a compelling manner, thus increasing their potential to gain revenue on slow-moving inventory.¹⁸² Further, given the vast size of Amazon's product catalog, the use of this algorithmic approach also helped the company address the issues of which recommendations to present and in which order (known in data science as the "learning to rank" problem), and how to ensure that there are a diversity of products in each recommendation set.¹⁸³

The introduction of the item-to-item collaborative-filtering model has served as a significant targeted marketing tool and has transformed the e-commerce space tremendously.¹⁸⁴ It has also enabled the company to generate a personalized shopping experience for each user in a novel manner. Amazon asserts that providing a personalized experience will enhance users' overall experience with the platform. However, in doing so, the company also strives to increase the time a user spends on the platform, as well as the company's average order value (the average amount spent every time a user places an order on the platform), and overall revenue the company generates from each user.¹⁸⁵ The company also sells advertisements, and therefore users who spend a longer time on the platform will see and potentially click on more ads, thus driving revenue through ad impressions (views) and clicks.

Amazon's recommendation system relies on numerous explicit and implicit data points. Users are able to provide explicit feedback data to the platform by rating items from one to five stars, through both public and private ratings. The private ratings are not shared with other Amazon customers, nor do they impact the average customer review for an item. These private ratings are used to refine the recommendations that the user receives.¹⁸⁶ Although Amazon publicly notes how it uses the private ratings, the company has not indicated whether and how its recommendation systems rely upon the public ratings and feedback that users provide on items they have purchased and sellers they have purchased from.¹⁸⁷ Each action that a user takes on the platform provides implicit feedback data to the company, which it uses to refine its recommendations in real time.¹⁸⁸ These data points include a user's browsing and purchase histories, which items a user added to their virtual shopping cart, which items a user rated and liked, and what items similar users have browsed and purchased.¹⁸⁹ For new users, there is less implicit and explicit data available to the company. As the user continues using the Amazon platform, however, the company is able to collect a vast amount of additional data, which can then be used to refine recommendations.¹⁹⁰

Today, Amazon uses recommendation algorithms to offer different categories of recommendations to users at a range of junctures on its e-commerce platform. The company, however, provides little transparency around these recommendation system use cases. This limits the understanding and agency that users have over these tools. Several engineers, journalists, and researchers, however, have identified some of the categories of product recommendations

that the company's recommendation system generates. These categories are broken down below:

1. **Recommended For You:** When a user visits the Amazon.com webpage and¹⁹¹ logs in, they will see a tab on the main toolbar that is tied to their account (e.g. Gabby's Amazon.com). Once they click on this they are presented with a range of product recommendations across multiple categories.¹⁹² For example, this can include recommendations specific to Amazon Books, apparel, or electronics.
2. **Frequently Bought Together:** In order to increase average order value, the company makes recommendations on items that have been purchased together in the past. These product recommendations aim to convince customers to purchase an additional item (known as up-selling) and/or purchase a different product or service (known as cross-selling) by providing suggestions based on items a user has added to their shopping cart or below items that they are currently exploring on the website.¹⁹³
3. **Similar Items:** The company also makes recommendations for products similar to ones that a user has viewed recently. These recommendations are based on a user's browser history and the recommended items typically vary in terms of shape, size, and brand.¹⁹⁴
4. **General Browsing History:** The company will make product recommendations based on a user's browsing history in case they want to access and purchase something they have previously demonstrated an interest in.¹⁹⁵
5. **Items Recently Viewed:** The company also makes recommendations based on the items that a user recently viewed. Amazon has the same goal in making these recommendations as with the Similar Items recommendations, in that they want to suggest products based on a user's recent browsing history.¹⁹⁶
6. **Related To/Based On Items You Viewed:** These categories of recommendations suggest products that are similar to items that a user recently viewed. For example, if a user searched for hangers on the platform, these recommendations would suggest hangers of different shapes, sizes, brands, etc. These recommendations are also based on a user's recent browsing history.
7. **Customers Who Bought This Item Also Bought:** This category of recommendations suggests items that have been purchased collectively by users in the past. This category of recommendations is similar to the

Frequently Bought Together section and it similarly aims to increase average order value through up-selling and cross-selling.¹⁹⁷

8. Recommended Items Other Customers Often Buy Again: This category of recommendations suggests items that similar users often purchase multiple times.

9. New Version of This Item: This category of recommendations is based on the assumption that users like to upgrade items, such as electronics, that they purchased. As a result, this category of recommendations informs users when a new edition of an item they purchased is available.

¹⁹⁸

10. Recommended For You Based on a Previous Purchase/Inspired By Your Purchases/Inspired By Your Shopping Trends: These categories of recommendations make product suggestions to a user based on a recent purchase they made. After a user makes a purchase on the platform, they are directed to an order details page. On this page, the user will receive further recommendations for items that can be paired with the initial order. For example, if a user purchases an iPad on Amazon, they might receive recommendations for iPad covers on the subsequent order details page. These recommendations also appear on the homepage. This category of recommendations aims to encourage users to make a second purchase by offering a relevant cross-sell offer.¹⁹⁹

11. Best-Selling in Different Categories: This category of recommendations features top-selling items across the different categories of products on the platform. It is based on the notion that an item that has been widely purchased by other users is validated as worthwhile. In addition, these recommendations aim to help users identify popular products and make purchases from new categories of products that they have not made purchases from before. This produces a range of up-sell and cross-sell opportunities for the platform.²⁰⁰

12. Popular in Brands You May Like: This category recommends items that are popular and are sold by brands or sellers that a user may be interested in.

13. Off-site Email Recommendations: Amazon also sends users product recommendations via email. Users can opt out of receiving these marketing emails, and can also select certain categories of items (e.g. beauty, books, Amazon Echo) for which they would like to receive marketing emails. These emails vary in their content and focus, but

contain similar categories of recommendations as those available on the platform.²⁰¹

In the early 2000s, Amazon's product recommendation system relied on both automated and human decision-making. Initially, the recommendations put forth on the website relied more on automated decision-making, whereas the email recommendations users received were generated by humans. These recommendations were produced by Amazon employees who were tasked with using software tools to target users based on their purchasing and browsing history. These employees were also assigned a certain product and were responsible for identifying similar items to recommend along with the original item.²⁰²

In the early 2010s, journalists noted that the company relied on a range of metrics when constructing and deploying its email recommendations. These included email open rates, click rates, opt out rates, as well as revenue-prioritizing metrics. The company's use of revenue-prioritizing metrics and related marketing and targeting techniques meant that if the recommendation system determined a user should receive a recommendation email for best-selling shoes and best-selling books, the company would only send them the email with the highest average revenue-per-mail. In this way, the company aimed to avoid spamming users and therefore maximize purchases. According to Sucharita Mulpuru, a Forrester analyst, the conversion rate and efficiency of Amazon's recommendation emails were high, and considerably more effective than the on-site recommendations a user receives.²⁰³ Mulpuru shared that based on the performance of other e-commerce sites, one could estimate that Amazon's on-site recommendation conversion is approximately 60 percent, suggesting a much more staggering rate of success for email recommendations.²⁰⁴ The company also offers users the option to receive marketing newsletters by traditional mail.

The introduction of a robust recommendation system has radically transformed Amazon as a business. Between 2011 and 2012, the company integrated its recommendation system across all stages of the purchasing process, beginning with the product discovery stage, when a user first identifies an item of interest, and ending with the checkout stage. In the second fiscal quarter of 2012, the company reported a 29 percent increase in sales (approximately \$12.83 billion). This was a significant increase from the \$9.9 billion in sales that the company reported during the same quarter the previous year.²⁰⁵ According to the 2013 data released by McKinsey & Company, recommendation systems drive 35 percent of purchases at Amazon.²⁰⁶

Controversies Related to Amazon's Recommendation System

As outlined, recommendation algorithms are used widely on the Amazon platform. They have the power to significantly influence what products a user sees, and can subsequently drive purchases. The company, however, has provided little transparency around the data it uses to train these recommendation systems and how they contribute to the company's determinations on how effective its recommendation system is. Therefore, although the recommendation system has impacted the company's success as a business, it is difficult to know how this system impacts user behaviors and whether its recommendations are equally as refined for all of its users, across demographics. This therefore also prevents researchers from identifying and understanding any sources of bias. In addition, the results of these automated decision-making processes are not always positive. For example, over the past few years, journalists and researchers have identified numerous instances of the platform's recommendation engine-making suggestions for products that are misleading, false, or conspiracy-theory based.

For example, in March 2019 *WIRED* reported finding that Amazon recommended a number of anti-vaccine books in the best-selling categories of various health sections on Amazon Books, including the Epidemiology, Emergency Pediatrics, History of Medicine, and Chemistry categories. Many of these books were marked as #1 Best Sellers in their respective categories.²⁰⁷ In addition, the outlet also found that in the Oncology category, one of the books marked as a Best Seller made recommendations on cancer treatment that are contrary to the consensus of medical experts, such as consuming juice instead of undergoing chemotherapy. Further, when *WIRED* journalists searched for "cancer," a misinformation-filled book titled *The Truth About Cancer* was one of the first product recommendations that appeared. The outlet found that the book had 1,684 reviews and had a 96 percent five-star rating.

Also in March 2019, NBC News reported that the company's recommendation engine was recommending *QAnon: An invitation to the great awakening* in its Hot New Releases section of Amazon Books. The book outlined a number of the common beliefs held by the QAnon conspiracy theory movement,²⁰⁸ including that Democrats murder and eat children and that the U.S. government manufactured both AIDS and the movie *Monsters Inc.*²⁰⁹ The book rapidly climbed to the top 75 books sold on Amazon that month. The company did not respond to questions about how its algorithmic recommendation system had been used to make these product suggestions, and they did not clarify whether the recommendation of the book in the Amazon Books section meant that the book had also been recommended through its other recommendation mechanisms on the platform.²¹⁰ This raised significant concerns regarding the lack of oversight over how product recommendations are made, and the subsequent consequences that could result from recommending a harmful and misleading product such as this book.²¹¹ Further, cases like these have raised

concerns around popularity-based recommendations on the platform, as researchers and academics have suggested that popularity is a weak metric for quality, and it homogenizes results.²¹²

In 2015, Amazon introduced the “Amazon Choice” label to indicate highly-recommended items across the platform. According to the Amazon website, these recommendations are based on highly-rated, well-priced products that can be shipped immediately. However, there is little transparency around how these signals are weighted and how the final recommendations are made, and the company has not shared any information around the performance of its recommendation system and how and whether it is able to deliver personalized recommendations to users.²¹³

The lack of transparency around how the company uses algorithmic decision-making is also concerning, as there have been numerous instances indicating that the company’s recommendation algorithms can be taken advantage of. The rise of the QAnon book in Amazon’s top books sold list, as described above, is an example of the ease with which the company’s algorithms can be gamed. Although the QAnon movement has a relatively small base, through coordinated purchasing behaviors, the group was able to generate a spike in book sales and therefore foster its rise in Amazon’s rankings. The group also coordinated subsequent reviews of the book, giving it five stars and prompting the platform’s algorithms to continue recommending it. As a result of this, the book, which peddles fringe ideas, was able to gain significant viewership. The book’s sales peak may not have lasted long, and the boost in viewership may not have resulted in a significant number of subsequent purchases, but it did succeed in making the book and its ideas seem mainstream.²¹⁴

Such coordinated inauthentic activities extend far beyond small conspiracy theory-based groups who are seeking to amplify their ideas. Because of the vast influence that Amazon has over users’ consumption habits, sellers have a strong incentive to appear on the first page of product search results and recommendations. In addition, according to *WIRED*, Amazon reviews appear to significantly influence the company’s ranking and recommendation algorithms²¹⁵ (although the company has not confirmed this²¹⁶). Further, the company has stated that it accounts for profitability when ranking products and making recommendations on its service. This means that items that could earn more revenue would be ranked higher, and therefore there is an incentive to engage in additional black hat operations in order to secure a place in these rankings.²¹⁷

As competition on the platform has risen significantly, it has resulted in the growth of a vast black market industry²¹⁸ where sellers can purchase “black hat” services that can manipulate the Amazon platform to help the sellers gain an advantage on their competitors. The services that they advertise include helping sellers appear at the top of a product search result page, promoting a seller’s products on the platform such as by removing negative reviews, and taking

advantage of loopholes on the Amazon platform to raise a product's overall sales ranking.

For example, for the book *The Truth About Cancer* discussed above, Reviewmeta.com, a site that seeks to help customers evaluate whether online reviews are legitimate, has suggested that over 1,000 of the existing reviews were suspicious based on the time, language, and reviewer behavior.²¹⁹

As rules-violating sellers have gained more of an advantage of the platform, traditionally rule-abiding sellers have come under increased competitive pressure and often opt to use these services as well. Typically, Amazon users rely on reviews as an indication of an item's quality and reliability. However, the rise of this black hat industry has made it easy for sellers to purchase services that result in the production of persuasive positive or negative reviews.²²⁰ Amazon has stated that it deploys a team of investigators, automated technology, and machine learning technology to prevent and detect inauthentic reviews at scale, and to enforce its policies against actors who violate their policies.²²¹ However, many outlets and researchers have found that although these black hat activities are a violation of Amazon's terms of service, the platform's enforcement of these rules has been weak,²²² perhaps because of the sheer size of the company's product catalog.²²³ In addition, there is little transparency around the scope and scale of these enforcement efforts. The platform has also made no indication that it will intervene to change these recommendations, such as by flagging items as misleading, or downranking their presentation in its overall recommendation and rankings. In addition, the company has in the past removed items such as two books that contained misleading claims on how to fight autism for violating its content guidelines,²²⁴ however it is unclear how widely and consistently these rules are enforced. The company has removed some misleading content from its streaming service, after facing pressure from the media and legislators in the United States,²²⁵ but, as highlighted, fewer consequential actions have been taken on its e-commerce platform.

Amazon's Frequently Bought Together recommendation category has also raised some concerns. A 2017 investigation by a team at Channel 4 News in the United Kingdom found that when a user searched for a common chemical that was used in certain food products, the Amazon recommender system would suggest other items the user could buy, which collectively could be used to make black powder, a chemical explosive. The system also recommended other items, such as ball bearings, which could be used as shrapnel in homemade explosives.²²⁶ In response, Amazon said it was reviewing its website to ensure that all products were being "presented in an appropriate manner."²²⁷ Although merely purchasing these items is not illegal, this recommendation raised concerns around whether users could easily access and purchase items to cause large scale harm and destruction with relative ease and whether Amazon's algorithmic systems were enabling or even encouraging these actions. In addition, in the

United Kingdom, there have been some successful prosecutions against individuals who have bought items that can be combined to make a bomb.²²⁸

Finally, Amazon's use of recommendation algorithms has also raised significant concerns around privacy of user data. In 2019, the company began testing a program that relied on machine learning tools and its broader recommendation engine to identify brands and products that a user may be interested in purchasing, based on data the company had collected on the user. Based on these inferences, Amazon sent users free samples of products to test, and this enabled the company to get new brands or brands of interest directly in front of users. The program raised a number of concerns among users, however, who were seeing first hand and offline manifestations of the company's collection of their personal data. Later in 2019, the company announced it will discontinue the program in 2020.²²⁹

User Controls Related to Amazon's Recommendation System

As discussed, Amazon has faced significant controversy over its recommendation system, issues that are compounded by the fact that the company does not provide meaningful transparency around how its recommendation system is structured and how it makes decisions. The company does offer its users some limited controls over how this system impacts their individual experiences on the platform, however.

For example, when a user is logged into their Amazon account, they have access to a page titled Recommendations. The page explains that when making recommendations, the company examines the items a user has purchased, the items a user has proactively indicated they own, and items they have rated. The page also explains that the company makes recommendations by comparing a user's activity with the activity of other users. Further, the page states that a user's recommendations will constantly change, based on whether they purchase or rate a new item, and whether the interests of similar users change. The company also asks that users indicate what items they are interested in by adding products to their Wish List or Shopping Cart in order to improve the personalization of their recommendations.²³⁰

In addition to this explanatory page, users can access a page titled Improve Your Recommendations. This page gives users a list of all of the items that they have purchased, the videos they have watched on Amazon Prime Video, the items they have marked as "I own it," the items they have rated, the items they have marked as "not interested," and the items that they have marked as gifts. On this page, users have the option to select "I prefer not to use these for recommendations" for any item, rate each item from one to five stars, and mark items as a gift.²³¹ Further, if a user is unsure why they have been recommended a particular product, they are able to see an explanation of why the item was recommended to them, such as because of a previous item rating, wish list addition, and so on.

Case Study: Netflix



Source: *sitthiphong / Shutterstock.com*

Netflix is an American digital content streaming and production company that was founded in 1997 by Reed Hastings and Marc Randolph. The company initially offered a mail-based DVD rental subscription service, but it introduced digital video streaming services in 2007.²³³ In 2016, the company separated its DVD service onto a new platform, known as DVD.com: A Netflix Company.²³⁴ The company then transitioned its main platform to subscription-based digital video streaming. The company is regarded as the world's largest subscription streaming service,²³⁵ with most recent estimates suggesting that the company now has 167.1 million subscribers around the world, with over 100 million of these users residing outside the United States.²³⁶ The company currently ranks twenty-first for global internet engagement on Alexa rankings.²³⁷ According to Nielsen's Total Audience Report released in February 2020,²³⁸ Netflix accounted for 31 percent of streaming time in U.S. homes capable of over-the-top streaming.²³⁹

Netflix's recommendation system is an important contributor to its revenue generation model, driving approximately 80 percent of hours of content streamed on the platform.²⁴⁰ Unlike YouTube and Amazon, the platform does not deliver targeted advertisements to its users. Rather, the company relies on subscriptions to both its digital video streaming service and DVD-delivery service to generate revenue. As competition between video streaming platforms

has heated up, the company has also invested heavily in producing original content. The platform's financial success relies on attracting and maintaining user attention, and preventing users from leaving Netflix in favor of competitors. Its recommendation engine is vital to achieving this, and it is therefore key to the company's business model.

Netflix's recommendation system holds a significant amount of influence over how the platform operates and how users engage with the service. Despite this, the platform has offered limited transparency around how this system is designed and how it operates. This is concerning given that the company's recommendation systems have raised concerns related to biased and discriminatory outcomes. In addition, Netflix offers users only a limited set of controls over how algorithmic decision-making shapes their platform experience.

A Technical Overview of Netflix's Recommendation System

Netflix operates its own proprietary recommendation engine that delivers personalized title recommendations to users. When the company was a mail-based DVD rental company, this recommendation system was known as Cinematch, and it was designed to help users fill their order queues with titles that they wanted to receive in the mail over the coming days, weeks, and months. Under the DVD-rental business model, each title took a few days to be delivered and users could only order a limited number of titles at a time. Netflix's recommendation system aimed to predict the number of stars that a user would rate a title, on a scale of one to five, after they had watched it. Once a user returned a title, they provided their actual rating. This served as the primary source of feedback from the user that the company then used to retrain and optimize its recommendation algorithms.²⁴¹

In 2006, the company launched the Netflix Prize challenge in order to improve predictions of user enjoyment based on individuals' movie preferences.²⁴² The company made a portion of its film database public for the competition, challenging participants to develop a collaborative filtering algorithm that could improve Netflix's own results by 10 percent.²⁴³ The winning team, BellKor's Pragmatic Chaos,²⁴⁴ was able to improve the results by 10.06 percent,²⁴⁵ and they were awarded a grand prize of \$1 million.²⁴⁶ Netflix still uses the algorithm the team developed to help predict ratings.²⁴⁷

However, as Netflix shifted toward a video streaming-focused service, it began transforming its recommendation system to align with new parameters and goals. When the company expanded into more customer segments and geographies, it realized a growing need to deliver and recommend personalized and interesting titles to different categories of users.²⁴⁸ In 2014, the company invested \$150 million (approximately 3 percent of its revenue at the time) in creating a team of 300 employees dedicated to improving the company's

recommendation engine.²⁴⁹ When Netflix introduced its services globally in 2016, it developed a single global recommender system that shares and relies on data from across each of its countries of operation. The company hoped that doing so would improve its recommendations for users in smaller markets without negatively impacting recommendations for users in larger markets.²⁵⁰

As a video streaming-focused company, Netflix had to alter its recommendation engine to provide users with multiple real-time suggestions on what to watch in the moment. Unlike subscribers to Netflix's DVD rental service, streaming subscribers can sample as many titles as they want before settling on one. In addition, these users are able to view numerous titles in one sitting or in quick succession.²⁵¹ Consumer research has suggested that the average Netflix user loses interest in the service after approximately 60 to 90 seconds of browsing and considering potential titles. Once a user loses interest, it is likely that they will switch over to another streaming service.²⁵² As a result, user retention is dependent on the company's recommendation engine being able to provide real-time, personalized recommendations.²⁵³ Unlike the recommendation algorithms used before 2006, the recommendation algorithms that the company now uses can be optimized in real time using explicit data as well as granular implicit data.

²⁵⁴

Netflix collects explicit feedback data from users by enabling them to provide a thumbs up or thumbs down on titles.²⁵⁵ This is similar to the explicit data the company collected during its DVD rental period, albeit at a larger scale. The company also collects implicit data, which consists of behavioral user data that is collected in real-time as a user navigates and engages with content on the service. Implicit data can include what titles a user has watched, the time of day and day of the week that a user is watching, what devices a user is watching on, and how long they are watching certain titles.²⁵⁶ The company also collects data on where in each row of recommendations the selected title appeared, and what titles were recommended to a user but not selected.²⁵⁷ It also considers data from other users, such as the combined ratings for a title made by all users who are similar to a given user.²⁵⁸ Netflix has stated that the implicit data it collects does not include demographic data such as race, age, and gender,²⁵⁹ although presumably user patterns and behaviors could enable the company to infer this information. Because implicit data provides more granular insights into a user's behaviors and preferences, the company considers it to be more useful than explicit data for constructing and retraining recommendation algorithms.²⁶⁰

Netflix's recommendation engine is at play at all stages of a user's journey on the platform. When a user creates a Netflix account or adds a new profile to their account, they are first prompted to select a few titles that they like. These titles are used to jump start their recommendations. This process, however, is optional. If a user elects to forego this process, then the initial recommendations the system delivers will be for a diverse and popular set of titles.²⁶¹ Once a user

begins watching titles on the platform, the system will use the explicit and implicit data it collects on the user and their viewing practices to adjust its recommendations. These data points will supersede any initial preference indications that a user previously provided, suggesting that revealed preferences can carry greater weight in the recommendations process than a user's initial-stated preferences. As a user watches more content over time, the recommendation engine will begin to weigh titles that the user consumed more recently heavier than titles that were consumed in the past.²⁶²

Today, Netflix's recommendation engine is a machine learning-driven collection of algorithms that serve different purposes and collectively create the Netflix streaming experience.²⁶³ This recommendation system offers multiple categories of recommendations. Most of these categories of recommendations reside on the homepage, which is the page that a user first sees when they log into their profile on any device. The homepage is the main hub for a user's personalized recommendations and it is where two of every three hours of content streamed on the platform are discovered.²⁶⁴

Below is a breakdown of the different categories of recommendation algorithms that comprise Netflix's recommendation system:

- 1. Personalized Video Ranker (PVR):** The PVR algorithm operates on the Netflix homepage and presents users with the entire catalog of titles available on the platform for the region in which they live, as well as certain categories of titles filtered by specific genre-based themes in a personalized manner. The Netflix homepage is designed in a matrix-like layout. In this matrix, each entry is a recommended title and each row of recommended videos are grouped together and labeled based on an overarching theme (e.g. Award-Winning Documentaries, Soapy TV Dramas, etc.).²⁶⁵ On average, each user will find approximately 40 rows on their homepage, with up to 75 videos per row. However, these figures may vary depending on the functionality of the user's device. Because the homepage experience is personalized to each user, the titles in each row, the order of the titles in each row, and the rows themselves vary per person. However, because the PVR algorithm is applied widely across the platform, it is designed so that it can provide broader and more generalized content recommendations and rankings. This limits the extent of the personalization that the algorithm can subsequently generate. As a result, this algorithm works best when it is producing recommendations based on a combination of personalized signals and general popularity signals, which are used to generate the recommendations in the Popular row.²⁶⁶
- 2. Top N Video Ranker:** The Top N video-ranker algorithm is used to produce recommendations for titles that appear in the Top Picks row on

the homepage. This algorithm is designed to identify a limited number of personalized recommendations from the entire Netflix catalog based on titles that are ranked highly. Whereas the PVR algorithm is assessed and optimized using metrics and algorithms that focus on the ranking that is produced for the entire catalog, the Top N video-ranker algorithm relies on metrics and algorithms that focus on the top percentiles of the catalog ranking. Like the PVR algorithm, the Top N video-ranker algorithm combines personalization and popularity metrics when making its recommendations. In addition, like the PVR algorithm, it is also able to identify and integrate user viewing trends it has collected over different time periods, ranging from a day to a year.²⁶⁷

3. Continue Watching Video Ranker (CWR): The CWR is a video ranking algorithm that is used to order the titles that appear in the Continue Watching row on the homepage. Whereas the majority of the video rankers that Netflix deploys order unviewed titles based on implicit data on user preferences and interests, the CWR sorts and ranks titles that a user has recently watched based on its calculations of whether the user will continue watching the title, or rewatch it, and whether a user abandoned a title because they found it uninteresting. The CWR algorithm considers signals such as the time that has elapsed since a user viewed the title, when a user stopped watching the title (e.g. mid-title, at the beginning of the title, or at the end of the title), whether the user has viewed different titles since abandoning the title, what device the user was watching the title on, etc.²⁶⁸

4. Video-Video Similarity (Sims): The Sims algorithm is an algorithm used to generate the Because You Watched (BYW) row, which features recommendations that are generated based on a user's consumption of one particular title. The algorithm evaluates every single title in the Netflix catalogue in order to identify titles that are similar to a title a user has recently watched. It then ranks and presents these similar titles in the BYW row. Although the Sims algorithm generates groupings that are not personalized to a specific user (i.e. everyone will be recommended the same predetermined list of similar titles if they watch a certain title), which BYW rows eventually appear on a user's homepage (e.g. BYW recommendations for Title A or Title B that a user has watched), as well as which titles within a predetermined BYW list appear, is personalized. This personalization is based on estimates of which titles a user would be interested in, and what they have already watched.²⁶⁹

Each of the platform's video ranking algorithms relies on different mathematical and statistical signals and data as input. They also require different model training scenarios that are constructed based on the specific goal of each algorithm.²⁷⁰ Whether or not the recommendation system is able to suggest titles

that a user is likely to watch strongly influences user retention on the platform. As a result, this is a key metric for indicating quality and effectiveness of the recommender system.²⁷¹ However, there is little transparency around the results of these tests, and around the training data the company uses to structure its recommendation system. This makes it difficult to understand how and if the company's recommendation system caters to the needs of different categories of users (e.g. users in different regions, users of different genders, etc.). It also makes it difficult to identify instances of bias.

In February 2020, Netflix added another layer of personalized recommendations to the homepage of its users, known as the Top 10 list. The list is updated daily and the positioning of the row on a user's homepage is dependent on how relevant Netflix believes the titles in the list are to each user. When a user clicks on Movies and TV Shows tabs, they are also able to see lists outlining the top 10 movies and top 10 TV shows in their country at that time. Titles that appear in Top 10 lists that also appear in other recommendation rows on a user's homepage are marked with a badge that reads Top 10. This recommendation feature was initially piloted in the United Kingdom and Mexico, and was introduced worldwide in February 2020.²⁷²

In each row on a user's homepage, there are three distinct types of personalization: the choice of which rows are displayed in which order (e.g. Because You Watched, Continue Watching, Romantic Comedies, etc.), which titles appear in each row, and the ranking or order that these titles appear within each row. Titles that are the most highly recommended will appear at the top of the homepage, and will be ordered from left to right in each row. However, for users who have set their language preference to Arabic or Hebrew, this is reversed, and the most highly recommended titles are ordered from right to left, since this is the direction in which those languages are read.²⁷³ The company's recommendation algorithm calculates a percent match score for each user that is displayed next to each title. This score is different for each user and it is a prediction of how likely a user is to like a given title.²⁷⁴

In addition to these four-core recommendation algorithms, the platform also deploys other algorithms, which work in conjunction with the recommendation system. These algorithms include:

1. **Evidence Selection:** Evidence Selection algorithms assess all information that is available about a title and determine which characteristics to feature in the description of the title. This choice is based on the algorithm's prediction of what the company believes a certain user will find most helpful when they are considering a recommendation. Evidence is the information that a user can see on the top left of their home page, such as the synopsis of a title, thumbnail images of a title, and other relevant information such as the cast members and awards. For example, evidence selection algorithms determine

whether a title being recommended to a user should be labeled as an Oscar-winning film or a film that is similar to another film that the user recently watched. Evidence selection algorithms are also responsible for determining which of the image thumbnails should be displayed to a user.

²⁷⁵ In 2017, Netflix began personalizing the artwork or thumbnails that each user can see when they browse different titles. For example, for the TV show *Stranger Things*, there are nine different artwork options that a user could potentially see while browsing. Which thumbnail a user eventually sees is based on their browsing history and whether they have demonstrated an interest in comedy, horror, or suspense titles in the past.
²⁷⁶

2. **Search:** As previously noted, Netflix’s recommendation system accounts for 80 percent of the hours of content streamed on the platform. Its search functionality, on the other hand, accounts for the remaining 20 percent. Using the search functionality, users can search the entire Netflix catalog available in their country for a particular title, actor, genre, and so on. When a user inputs a query, the results are based on the top results related to the actions of other users who have entered the same or similar queries.
²⁷⁷ The search functionality relies on its own set of algorithms. If an item is not in Netflix’s catalog, the search algorithms recommend alternative content based on the search query. This can be challenging given that many search queries are incomplete phrases or terms that consist of only a few letters. The search functionality relies on three different algorithms:²⁷⁸

1. The first algorithm aims to identify titles that match a search query (e.g. delivering Fantasia when a user searches “fan”).
2. The second algorithm predicts a user’s interest in a specific concept when they enter a partial search query (e.g. identifying and suggesting the concept “fantasy” when a user searches for “fan”).
3. The third algorithm provides title recommendations for a specific concept (e.g. providing specific title recommendations for fantasy movies).

According to Todd Yellin, Netflix’s former vice president of innovation, now vice president of product, Netflix’s recommendation engine relies on three distinct categories of information:²⁷⁹

1. **Information on Netflix Members:** As previously mentioned, Netflix collects granular implicit behavioral data on its users including their viewing history, when they are watching, and on what devices.

2. Information on Content and Titles: Netflix employs numerous in-house and freelance individuals to watch each Netflix title and assign tags to them. These tags are granular in nature and can include information such as whether a title has an ensemble cast, is set in space, and stars a strong female lead. The company deploys the same tags globally and many of these tags are visible to users when they are navigating through various titles on the platform. However, a smaller subset of tags are used across different regions, languages, and cultures in order to enhance localization (e.g. the English language tag “gritty drama” may not translate well into other languages, and as a result a more localized tag will be used).²⁸⁰

3. Information Produced by Machine Learning Algorithms: Netflix’s machine learning algorithms combine the information the platform has collected on its users and the information produced during the tagging process to identify important elements and patterns and assign them relevant weights. Based on this data, the algorithms generate thousands of “taste communities.” Taste communities are categories of users who are interested in similar titles.²⁸¹ According to Netflix, examples of these communities include users who watched *House of Cards* and also *It’s Always Sunny in Philadelphia*, or *Making a Murderer* and the *John Mulaney: The Comeback Kid* comedy special.²⁸² Each user will fit into numerous taste communities and which taste communities a user is assigned to impact the various recommendations they ultimately receive on the platform.

As outlined, Netflix’s recommendation engine involves a complex interplay of various algorithms and signals. Although the company does not sell ads, the ability to retain user attention on the platform still significantly influences the company’s bottom line as it means users will consume more titles on the platform, and not be consuming content on other services. According to 2016 estimates, the combination of such automated personalization and recommendation tools saved the company over \$1 billion per year, as they enabled the platform to reduce the monthly number of customers that stopped using the service (known as customer churn) by a few percentage points. This has become increasingly important as competing streaming platforms have emerged in the United States and around the world.²⁸³

Controversies Related to Netflix’s Recommendation System

Netflix’s recommendation systems, however, have surfaced some concerns regarding biased and discriminatory outcomes. For example, in 2017, the company introduced a new algorithm to personalize the artwork or thumbnail images a user sees for each title. The company introduced this algorithm after its own research found that these visual elements were the biggest influencing

factor when a viewer was deciding which title to watch, comprising 82 percent of their focus. The artwork that a user sees changes based on the user's tastes and viewing history. After the algorithm was introduced, however, numerous African-American users in the United States found that the thumbnail images they were seeing were racially and ethnically driven, and were often misrepresentative of the actual cast of the movie. For example, one user saw a thumbnail for the romantic comedy film *Love Actually* featuring British black actor Chiwetel Ejiofor, who plays a minor role in the film, and Keira Knightley. Other users saw a thumbnail featuring the main stars of the film: Hugh Grant, Emma Thompson, and Colin Firth.²⁸⁴ Users reported similar outcomes for titles such as the murder mystery series *The Good Cop*.²⁸⁵ These results caused outrage among many Netflix users, who were concerned that Netflix was tailoring its recommendations to them based on their race and ethnicity. Many also expressed concern that in altering the artwork to include black characters, the platform was falsely suggesting plot lines and actor prominence in titles in order to attain user engagement and attention.²⁸⁶ The company responded by stating "we don't ask members for their race, gender or ethnicity so we cannot use this information to personalize their individual Netflix experience. The only information we use is a member's viewing history."²⁸⁷ Nonetheless, these incidents provide a further example of how algorithmic tools can make inferences on a user's demographic characteristics using patterns identified in data points such as viewing history. This can result in biased and discriminatory outcomes. In this case, the outcomes did not have significant, harmful offline effects. However, they do demonstrate a need for the company to provide greater transparency and accountability around how these algorithms are trained and how they engage in decision-making. This could help researchers identify when Netflix's recommendation system is making concerning inferences (especially in situations like the above where the system seemingly made race-based inferences although the company does not collect racial data). Given that the company is used widely around the world by users who may be part of marginalized communities, it is important that the company invests in ensuring these individuals are not being misrepresented on the platform.

In addition, although Netflix states it does not collect information such as race, gender, and ethnicity, information on user's viewing patterns is also a highly sensitive category of data. This has been demonstrated in debates over U.S. surveillance law. Section 215 of the Foreign Intelligence Surveillance Act, (which has more recently received attention as the surveillance law under which the government has collected phone records), was previously known as the "library records provision." Section 215 authorizes the government to collect business records, including financial records and library records. The provision thus can require librarians to provide information on customer reading and computer records. Numerous librarians have opposed the provision,²⁸⁸ explaining that information on viewing patterns is highly sensitive, and requiring librarians to hand over this data amounts to a serious privacy violation.²⁸⁹

Given the extent to which recommendation algorithms are used on the Netflix platform, Netflix should provide greater transparency around these tools' impact on user's Netflix experiences. Netflix has made some positive strides in this regard.

Many of the company's executives have publicly spoken and written about the company's recommendation engine in varying levels of technicality.²⁹⁰ In addition, in its online help center, the platform published a page that provides a high level overview of how the company's recommendation system works. This page outlines the different factors the system considers when producing recommendations; how the initial jump start recommendation process works when a user creates a new profile; how recommendations are defined by row, rank and title; and how the company improves its recommendation systems. Although this page does not provide a granular and technical overview of the company's recommendation systems, it does present this information in a user-friendly and digestible manner. This is a positive first step toward building user trust and agency, and in providing transparency and accountability around how the company uses automated decision-making in its recommendations process.

User Controls Related to Netflix's Recommendation System

Netflix presents an interesting case study in that the company discloses information on how its recommendation system works, thus offering users a sense of agency through a limited set of controls that aim to help promote awareness of its algorithms. For example, in the My Profile section, each user has access to pages that outline all of the titles a user has previously rated and all of the titles they have watched and that are stored in their watch history. Through these pages, users can change or remove their rating for a title and remove a title from their watch history. In addition, users can also access their recent device streaming activity, which outlines which devices the user profile has recently watched Netflix on, where, and on what dates and times. However, Netflix does not offer users the ability to view the various characteristics and signals its recommendation system uses to make video recommendations or view explanations of why a particular title was recommended to them. Users also do not have the option to control or opt out of having certain factors considered by Netflix's recommendation system. In addition, Netflix does not offer users the ability to opt out of receiving title suggestions from the recommendation system altogether.

This is an interesting comparison with companies such as YouTube and Amazon, which offer less transparency and accountability around how their recommendation systems work, but offer users a greater set of controls over their personal platform experiences.

Promoting Fairness, Accountability, and Transparency Around Algorithmic Recommendation Practices

As explained in this report, the use of algorithmic systems to curate and recommend content gives internet platforms significant power to shape the perspectives and behaviors of their users. The use of recommendation systems has also significantly transformed, and contributed to the success of, companies' business models because these systems help platforms retain user attention, thus enabling them to target users with advertisements and/or further recommendations. However, the use of these algorithmic recommendation systems have also sparked a number of controversies. Despite this, internet platforms do not offer a significant amount of transparency and accountability around how these systems are structured, how they operate, and how they engage in decision-making. This makes it difficult to understand how these systems impact users, their worldviews, and their behaviors. Going forward, internet platforms and policymakers should consider the following set of recommendations in order to promote greater fairness, accountability, and transparency around algorithmic decision-making.²⁹¹ This report does not offer recommendations for researchers, as there is currently far too little corporate data available to researchers related to algorithmic recommendation systems. Once internet platforms begin implementing the recommendations outlined below, we will be able to suggest more tangible recommendations for researchers.

Recommendations for Internet Platforms

- 1. Disclose to users the situations in which the platform uses an algorithmically-curated recommendation system and provide comprehensive and meaningful explanations to users around how their recommendation systems work.** These explanations should include information on the explicit and implicit data points the system considers (especially sensitive data points such as demographic information). These explanations should also include the various signals (e.g. when a user stopped watching a video, how popular an item is) and factors (e.g. user location, on what device a user is visiting the platform) that a recommendation system considers and weighs in order to generate its recommendations. If these signals or factors change, the company should publicly disclose this and explain why these changes have been made. These disclosures should also include an overview of instances in which the company will manually intervene in algorithmic decision-making processes to change outcomes (e.g. to downrank or hide

recommendations that include misinformation). The disclosure and these explanations should be publicly available, easily accessible on the company's website, and written in a manner that is easily comprehensible by the average user.

2. **Explain to users why a recommendation was made to them.** Users should be able to access information on why a particular video, item, etc. was recommended to them. This explanation should at a minimum include information on the different signals and user characteristics the recommendation system considered to make the recommendation. It should also include an easy link to relevant user controls (per recommendation seven below) that could let the user change their recommendation preferences.
3. **Disclose granular data around how the company trains its algorithmic recommendation systems.** At a minimum, this should include information on the categories of users that a company's training data sets are trained on (e.g. which demographic groups).
4. **Enable independent researchers to conduct audits to review and verify relevant internal models and data.** In particular, companies should permit pre-vetted researchers to review and verify its training models, the results of tests the company runs to evaluate how effective their recommendation system is, any statistics the company has publicly released related to the impact of algorithmic changes on the operations of the recommendation system, and data related to controversial categories of recommendations such as extremist propaganda, conspiracy theories, and misinformation.
5. **Hire independent auditors to conduct regular periodic audits of recommendation algorithms in order to identify potentially harmful outcomes and take steps to address findings of audits, including mitigating discrimination and bias.** These audits should specifically evaluate how algorithmic recommendation systems can inappropriately influence or manipulate user perspectives and behaviors, promote concerning topics of information, and cause discrimination. Internet platforms should conduct these audits proactively, as well as in response to concerns surfaced by community partners, civil society organizations, researchers, activists, etc. Companies should take affirmative steps to address any problematic findings from the audits, including using the results of these audits to refine, retrain, and improve their recommendation systems and make them more fair, accountable, and transparent. Companies should also work to reduce instances of discrimination and bias that result from the use of their algorithmic recommendation systems. These audits should be conducted by an

external third party, and companies should make summaries publicly available.

6. Share granular data related to how the company tests its recommendation systems and how it determines how effective the company's systems are. At a minimum, this should include information on how well these systems predict the preferences of different demographic groups. In addition, this data should be continuously updated to indicate how various algorithmic changes have impacted the company's metrics and conclusions related to the overall effectiveness of the company's recommendation system.

7. Improve user controls so that users can easily manage whether and how their data is collected and inferred, how this data is used, and how it influences the recommendations that they see. These user controls should be easy to access and understand. They should be available to all logged in users of a service. In addition, these controls should be accompanied with an explanation of how using these controls will impact a user's overall platform experience. At a minimum, these user controls should include the ability to:

1. Select and change the factors (e.g. demographic information, browsing history, purchase history, ratings, interests) that a recommendation system may consider when generating recommendations for them. These settings should include the ability to completely opt out from having any of these factors considered. It should also include the ability to completely clear a user's watch, browsing, and purchase history. These controls are integral for protecting user privacy.
2. Exclude certain videos, titles, channels, sellers, or items from factoring into their recommendations.
3. Choose whether recommendations are influenced by a user's activity on partner or related products and websites. This should include the option to opt out entirely from having such data considered.
4. Opt out of the autoplay feature on video and streaming-based services. Ideally, users should have to opt into receiving autoplay recommendations on any platform.
5. Decide whether they want to receive algorithmically-curated recommendations at all. Ideally, users should have to opt into receiving such recommendations on any platform. At a minimum,

users should have access to controls that enable them to fully opt out of the recommendation process. Users should have easy to use controls that let them opt out of all practices at once.

8. **Share the platform's Terms of Service Community Guidelines related to topics such as content and purchases, and how they are enforced.** These guidelines should be easily accessible and comprehensible to the average user. They should clearly explain what kinds of content and behaviors are and are not permissible on the platform. They should also explain how the company will enforce these policies, and what the consequences for violating these policies are. If the company changes these Terms of Service or enforcement policies, they should announce these changes and explain why these changes have been made.
9. **Publish a transparency report outlining the scope and scale of Terms of Service enforcement actions in all of the regions in which it operates.** These transparency reports should provide granular and meaningful data around how the company has enforced its Terms of Service. In addition, this transparency report should be published at regular intervals (e.g. annually, quarterly, etc.). All of the data in the transparency report should be available in a structured data format (e.g. comma separated values), rather than or in addition to a flat PDF file. This is helpful to researchers who want to make use of the report data, as it simplifies the data extraction process and makes reports more accessible.
10. **Explain how the company uses human evaluators to review and train its algorithmic and machine-learning models.** These explanations should include an overview of the role of the evaluators and a publicly available copy of the guidelines these evaluators use.

Recommendations for Policymakers

The recommendations for policymakers in this report are focused on U.S. policymakers. This is both because the platforms discussed are U.S. companies and also because the First Amendment of the U.S. Constitution imposes unique constraints on the extent to which U.S. policymakers can regulate how companies decide which content to permit on their platforms.

In order to help protect privacy and prevent harmful outcomes as a result of algorithmic decision-making in recommendation systems, U.S. policymakers should:

Enact rules to require greater transparency from online platforms regarding their use of algorithmic recommendation systems. The U.S. government is limited in the extent to which it can direct platforms how to decide what content to permit on their sites. However, Congress could improve accountability mechanisms by requiring greater transparency around the use of algorithmic recommendation systems.

Notes

- 1 Renee DiResta, "Up Next: A Better Recommendation System," *WIRED*, April 11, 2018, <https://www.wired.com/story/creating-ethical-recommendation-engines/>
- 2 Renee DiResta, "How Amazon's Algorithms Curated a Dystopian Bookstore," *WIRED*, March 5, 2019, <https://www.wired.com/story/amazon-and-the-spread-of-health-misinformation/>
- 3 Spandana Singh, *Special Delivery: How Internet Platforms Use Artificial Intelligence to Target and Deliver Ads*, February 18, 2020, <https://www.newamerica.org/oti/reports/special-delivery/>
- 4 Zeynep Tufekci, "How Recommendation Algorithms Run the World," *WIRED*, April 22, 2019, <https://www.wired.com/story/how-recommendation-algorithms-run-the-world/>
- 5 Azadeh Nematzadeh et al., *How Algorithmic Popularity Bias Hinders Or Promotes Quality*, July 14, 2017, <https://arxiv.org/pdf/1707.00574.pdf>
- 6 Sinha and Swearingen, *The Role*.
- 7 Ryan Bigge, "Better Personalized Recommendations Through Transparency and Content Design," *Medium* (blog), entry posted February 6, 2019, <https://medium.com/@biggeidea/better-personalized-recommendations-through-transparency-and-content-design-f5b2af16f8b>
- 8 Diresta, "Up Next".
- 9 Nematzadeh et al., *How Algorithmic*.
- 10 Ido Guy et al., "Social Media Recommendation based on People and Tags," *SIGIR '10: Proceedings Of The 33rd International ACM SIGIR Conference on Research and Development In Information Retrieval*, July 2010, <https://www.research.ibm.com/haifa/dept/imt/papers/guySIGIR10.pdf>
- 11 Michael D. Ekstrand et al., "All The Cool Kids, How Do They Fit In? Popularity and Demographic Biases in Recommender Evaluation and Effectiveness," *Proceedings of Machine Learning Research: Conference on Fairness, Accountability, and Transparency* 81 (2018):, <http://proceedings.mlr.press/v81/ekstrand18b/ekstrand18b.pdf>
- 12 Michael J. Pazzani and Daniel Billsus, "Content-Based Recommendation Systems," in *The Adaptive Web: Methods and Strategies of Web Personalization* (2007), https://link.springer.com/chapter/10.1007/978-3-540-72079-9_10
- 13 Pazzani and Billsus, "Content-Based Recommendation".
- 14 Pavel Kordík, "Recommender Systems Explained," *Recombee* (blog), entry posted July 12, 2016, <https://medium.com/recombee-blog/recommender-systems-explained-d98e8221f468>
- 15 Guy et al., "Social Media".
- 16 Remus Titiriga, "Social Transparency through Recommendation Engines and its Challenges: Looking Beyond Privacy," *Informatica Economică* 15, no. 4 (2011): https://pdfs.semanticscholar.org/ad3/e37bfc81e82a8ce3e43d02c4a267eb1ee568.pdf?_ga=2.137990704.1671841003.1571852964-1797958481.1570491365
- 17 Ekstrand et al., "All The Cool".
- 18 Titiriga, "Social Transparency".
- 19 Charu K. Aggarwal, *Recommender Systems* (Springer International Publishing, 2016).
- 20 John Riedl and Joseph Konstan, *Word of Mouse: The Marketing Power of Collaborative Filtering* (Warner Books, 2002), https://books.google.com/books/about/Word_of_Mouse.html?id=-NcvmAEACAAJ

- 21 Titiriga, "Social Transparency".
- 22 Kordík, "Recommender Systems," *Recombee* (blog).
- 23 Rashmi Sinha and Kirsten Swearingen, *The Role of Transparency in Recommender Systems*, 2002, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.8.8787&rep=rep1&type=pdf>
- 24 Titiriga, "Social Transparency".
- 25 Kordík, "Recommender Systems," *Recombee* (blog).
- 26 Michael Martinez, "Amazon: Everything You Wanted To Know About Its Algorithm and Innovation," IEEE Computer Society, <https://www.computer.org/publications/tech-news/trends/amazon-all-the-research-you-need-about-its-algorithm-and-innovation>
- 27 Aggarwal, *Recommender Systems*.
- 28 Ekstrand et al., "All The Cool".
- 29 Aggarwal, *Recommender Systems*.
- 30 Guy et al., "Social Media".
- 31 Christopher McFadden, "YouTube: Its History and Impact on the Internet," *Interesting Engineering*, October 4, 2019, <https://interestingengineering.com/youtube-its-history-and-impact-on-the-internet>
- 32 Michael Arrington, "Google Has Acquired YouTube," *TechCrunch*, October 9, 2006, <https://techcrunch.com/2006/10/09/google-has-acquired-youtube/>
- 33 Joan E. Solsman, "Mozilla Is Sharing YouTube Horror Stories To Prod Google For More Transparency," *CNET*, October 15, 2019, <https://www.cnet.com/news/mozilla-is-sharing-youtube-horror-stories-to-prod-google-for-more-transparency/>
- 34 Maryam Mohsin, "10 Youtube Stats Every Marketer Should Know in 2020 [Infographic]," Oberlo, last modified November 11, 2019, <https://www.oberlo.com/blog/youtube-statistics>
- 35 "Youtube.com Competitive Analysis, Marketing Mix and Traffic," Alexa Internet, <https://www.alexa.com/siteinfo/youtube.com>
- 36 Kevin Roose, "The Making of a YouTube Radical," *New York Times*, June 8, 2019, <https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html?mtrref=mail.google.com&mtrref=www.nytimes.com&assetType=REGIWALL&mtrref=undefined&assetType=REGIWALL>
- 37 Ben Popken, "As Algorithms Take Over, YouTube's Recommendations Highlight A Human Problem," *NBC News*, April 19, 2018, <https://www.nbcnews.com/tech/social-media/algorithms-take-over-youtube-s-recommendations-highlight-human-problem-n867596>
- 38 Adrienne LaFrance, "The Algorithm That Makes Preschoolers Obsessed With YouTube," *The Atlantic*, July 25, 2017, <https://www.theatlantic.com/technology/archive/2017/07/what-youtube-reveals-about-the-toddler-mind/534765/>
- 39 Casey Newton, "How YouTube Perfected The Feed," *The Verge*, August 30, 2017, <https://www.theverge.com/2017/8/30/16222850/youtube-google-brain-algorithm-video-recommendation-personalized-feed>
- 40 Popken, "As Algorithms".
- 41 Matt Elliott, "How To Turn Off YouTube's New Autoplay Feature," *CNET*, March 20, 2015, <https://www.cnet.com/how-to/how-to-turn-off-youtubes-new-autoplay-feature/>
- 42 Jack Nicas, "How YouTube Drives People to the Internet's Darkest Corners," *Wall Street Journal*, February 7, 2018, [\[newamerica.org/oti/reports/why-am-i-seeing-this/\]\(http://newamerica.org/oti/reports/why-am-i-seeing-this/\)](https://www.wsj.com/articles/how-
</div>
<div data-bbox=)

youtube-drives-viewers-to-the-internets-darkest-corners-1518020478

43 Aaron Smith, Skye Toor, and Patrick Van Kessel, "Many Turn to YouTube for Children's Content, News, How-To Lessons," Pew Research Center, last modified November 7, 2018, <https://www.pewresearch.org/internet/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons/>

44 "The RDR Corporate Accountability Index: Transparency and Accountability Standards for Targeted Advertising and Algorithmic Systems — Pilot Study and Lessons Learned," Ranking Digital Rights, March 2020, [rankingdigitalrights/pilot-report-2020](https://rankingdigitalrights.org/pilot-report-2020)

45 "Manage Your Recommendations and Search Results," YouTube Help, <https://support.google.com/youtube/answer/6342839?hl=en&co=GENIE.Platform=Desktop>

46 Jonas Kaiser and Adrian Rauchfleisch, "Unite the Right? How YouTube's Recommendation Algorithm Connects The U.S. Far-Right," *D&S Media Manipulation: Dispatches from the Field* (blog), entry posted April 11, 2018, <https://medium.com/@MediaManipulation/unite-the-right-how-youtubes-recommendation-algorithm-connects-the-u-s-far-right-9f1387ccfabd>

47 Caroline O'Donovan et al., "We Followed YouTube's Recommendation Algorithm Down The Rabbit Hole," *BuzzFeed News*, January 24, 2019, <https://www.buzzfeednews.com/article/carolineodonovan/down-youtubes-recommendation-rabbithole>

48 "Manage Your," YouTube Help, "The RDR Corporate Accountability Index: Transparency and Accountability Standards for Targeted Advertising and Algorithmic Systems — Pilot Study and Lessons Learned," Ranking Digital Rights, March 2020, [rankingdigitalrights/pilot-report-2020](https://rankingdigitalrights.org/pilot-report-2020)

49 Paul Covington, Jay Adams, and Emre Sargin, "Deep Neural Networks for YouTube Recommendations," *Proceedings of the 10th ACM Conference on Recommender Systems, ACM, New York, NY, USA*, 2016, <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45530.pdf>

50 Roose, "The Making".

51 Covington, Adams, and Sargin, "Deep Neural".

52 Roose, "The Making".

53 Mark Bergen and Lucas Shaw, "To Answer Critics, YouTube Tries A New Metric: Responsibility," *The Star*, April 15, 2019, <https://www.thestar.com.my/tech/tech-news/2019/04/15/to-answer-critics-youtube-tries-a-new-metric-responsibility/>

54 Roose, "The Making".

55 Roose, "The Making".

56 Bergen and Shaw, "To Answer."

57 Roose, "The Making".

58 Bergen and Shaw, "To Answer."

59 Roose, "The Making".

60 Michael Learmonth, "YouTube's Video Views Are Falling -- By Design," *AdAge*, May 14, 2012, <https://adage.com/article/digital/youtube-s-video-views-falling-design/234735>

61 Newton, "How YouTube".

62 Bergen and Shaw, "To Answer."

63 "Neural Network," DeepAI, <https://deepai.org/machine-learning-glossary-and-terms/neural-network>

64 Alex Woodie, "Inside Sibyl, Google's Massively Parallel Machine Learning Platform," *Datanami*, last

modified July 17, 2014, <https://www.datanami.com/2014/07/17/inside-sibyl-googles-massively-parallel-machine-learning-platform/>

65 Margaret Rouse and Matthew Haughn, "Supervised Learning," Search Enterprise AI, <https://searchenterpriseai.techtarget.com/definition/supervised-learning>

66 Newton, "How YouTube".

67 Newton, "How YouTube".

68 Newton, "How YouTube".

69 Covington, Adams, and Sargin, "Deep Neural".

70 Covington, Adams, and Sargin, "Deep Neural".

71 Covington, Adams, and Sargin, "Deep Neural".

72 In this context, precision can be understood as how useful search results are. Recall can be understood as how complete search results are. Ranking loss functions winnow down the list of potential recommendations.

73 A/B testing compares two versions of a variable by testing a user's response to variable A against variable B, and establishing which of the two variables is more effective.

74 Covington, Adams, and Sargin, "Deep Neural".

75 Ekstrand et al., "All The Cool".

76 Alexis C. Madrigal, "How YouTube's Algorithm Really Works," *The Atlantic*, November 8, 2018, <https://www.theatlantic.com/technology/archive/2018/11/how-youtubes-algorithm-really-works/575212/>

77 Covington, Adams, and Sargin, "Deep Neural".

78 Błażej Osiński and Konrad Budek, "What Is Reinforcement Learning? The Complete Guide," Deep Sense AI, last modified July 5, 2018, [https://](https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/)

deepsense.ai/what-is-reinforcement-learning-the-complete-guide/

79 Roose, "The Making".

80 Roose, "The Making".

81 "[YouTube Recommendations] Ask us anything! YouTube Team will be here Friday February 8th.," YouTube Help, last modified February 8, 2019, <https://support.google.com/youtube/thread/1456096?hl=en>

82 Mark Bergen, "YouTube Executives Ignored Warnings, Letting Toxic Videos Run Rampant," *Bloomberg*, April 2, 2019, <https://www.bloomberg.com/news/features/2019-04-02/youtube-executives-ignored-warnings-letting-toxic-videos-run-rampant>

83 Bergen, "YouTube Executives".

84 YouTube, "The Four Rs of Responsibility, Part 1: Removing Harmful Content," *Official YouTube Blog*, entry posted September 3, 2019, <https://youtube.googleblog.com/2019/09/the-four-rs-of-responsibility-remove.html>

85 YouTube, "Susan Wojcicki: Preserving Openness Through Responsibility," *Official YouTube Blog*, entry posted August 27, 2019, <https://youtube-creators.googleblog.com/2019/08/preserving-openness-through-responsibility.html>

86 Paul Lewis, "'Fiction is Outperforming Reality': How YouTube's Algorithm Distorts Truth," *The Guardian*, February 2, 2018, <https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth>

87 YouTube, "Continuing Our Work To Improve Recommendations On YouTube," *Official YouTube Blog*, entry posted January 25, 2019, <https://youtube.googleblog.com/2019/01/continuing-our-work-to-improve.html>

88 Popken, "As Algorithms".

- 89 Bergen, "YouTube Executives".
- 90 Bergen and Shaw, "To Answer."
- 91 Smith, Toor, and Kessel, "Many Turn," Pew Research Center.
- 92 YouTube, "The Four Rs of Responsibility, Part 2: Raising Authoritative Content and Reducing Borderline Content and Harmful Misinformation," *Official YouTube Blog*, entry posted December 3, 2019, <https://youtube.googleblog.com/2019/12/the-four-rs-of-responsibility-raise-and-reduce.html>
- 93 YouTube, "Building a Better News Experience On YouTube, Together," *Official YouTube Blog*, entry posted July 9, 2018, <https://youtube.googleblog.com/2018/07/building-better-news-experience-on.html>
- 94 Conversation with representatives from YouTube on March 2, 2020
- 95 Conversation with representatives from YouTube on March 2, 2020
- 96 Conversation with representatives from YouTube on March 2, 2020
- 97 Roose, "The Making".
- 98 Roose, "The Making".
- 99 YouTube, "Our Ongoing Work To Tackle Hate," *Official YouTube Blog*, entry posted June 5, 2019, <https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>
- 100 YouTube, "How YouTube Supports Elections," *Official YouTube Blog*, entry posted February 3, 2020, <https://youtube.googleblog.com/2020/02/how-youtube-supports-elections.html>
- 101 Conversation with representatives from YouTube on March 2, 2020
- 102 Kevin Roose, "YouTube's Product Chief On Online Radicalization, Algorithmic Rabbit Holes," *SF Gate*, April 6, 2019, <https://www.sfgate.com/business/article/YouTube-s-product-chief-on-online-13746004.php>
- 103 YouTube, "Continuing Our Work," *Official YouTube Blog*.
- 104 YouTube, "Our Ongoing," *Official YouTube Blog*.
- 105 Solsman, "Mozilla Is Sharing".
- 106 YouTube, "Continuing Our Work," *Official YouTube Blog*.
- 107 YouTube, "Continuing Our Work," *Official YouTube Blog*.
- 108 "External Evaluators and Recommendations," YouTube Help, <https://support.google.com/youtube/answer/9230586>
- 109 YouTube, "Continuing Our Work," *Official YouTube Blog*.
- 110 YouTube, "Continuing Our Work," *Official YouTube Blog*.
- 111 YouTube, "Continuing Our Work," *Official YouTube Blog*.
- 112 Covington, Adams, and Sargin, "Deep Neural".
- 113 Newton, "How YouTube".
- 114 "Manage Your," YouTube Help.
- 115 Kaiser and Rauchfleisch, "Unite the Right?," *D&S Media Manipulation: Dispatches from the Field* (blog).
- 116 Covington, Adams, and Sargin, "Deep Neural".
- 117 Covington, Adams, and Sargin, "Deep Neural".
- 118 "The RDR Corporate Accountability Index: Transparency and Accountability Standards for

Targeted Advertising and Algorithmic Systems — Pilot Study and Lessons Learned,” Ranking Digital Rights, March 2020, rankingdigitalrights.org/pilot-report-2020

119 Jaron Harambam, Natali Helberger, and Joris van Hoboken, "Democratizing Algorithmic News Recommenders: How To Materialize Voice In A Technologically Saturated Media Ecosystem," *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*, October 2018, https://www.ivor.nl/publicaties/download/Harambam-et-al_-Democratizing-algorithmic-news-recommenders-2018.pdf

120 Chris Stokel-Walker, "YouTube's Deradicalization Argument Is Really a Fight About Transparency," *FFWD* (blog), entry posted December 29, 2019, <https://ffwd.medium.com/youtubes-deradicalization-argument-is-really-a-fight-about-transparency-fe27af2f3963>

121 Ekstrand et al., "All The Cool".

122 Ekstrand et al., "All The Cool".

123 Jesselyn Cook, "YouTube And Google Algorithms Promoted Notre Dame Conspiracy Theories," *The Huffington Post*, April 17, 2019, https://www.huffpost.com/entry/notre-dame-youtube-google-conspiracy-theories_n_5cb77c2be4b096f7d2db8183

124 Cook, "YouTube And Google".

125 Cook, "YouTube And Google".

126 Cook, "YouTube And Google".

127 O'Donovan et al., "We Followed".

128 Cook, "YouTube And Google".

129 Lewis, "'Fiction is Outperforming".

130 Roose, "The Making".

131 Roose, "YouTube's Product"

132 "YouTube Regrets," Mozilla Foundation, <https://foundation.mozilla.org/en/campaigns/youtube-regrets/>

133 Email conversation with representative from the Mozilla Foundation

134 Solsman, "Mozilla Is Sharing".

135 Kaiser and Rauchfleisch, "Unite the Right?," *D&S Media Manipulation: Dispatches from the Field* (blog).

136 Pariser, *The Filter Bubble: What the Internet is Hiding From You*.

137 Cook, "YouTube And Google".

138 Lewis, "'Fiction is Outperforming".

139 Lewis, "'Fiction is Outperforming".

140 O'Donovan et al., "We Followed".

141 Roose, "YouTube's Product"

142 Roose, "YouTube's Product" Kevin Roose, "YouTube's Product Chief on Online Radicalization and Algorithmic Rabbit Holes," *New York Times*, March 29, 2019, <https://www.nytimes.com/2019/03/29/technology/youtube-online-extremism.html?rref=collection%2Fbyline%2Fkevin-roose&action=click&contentCollection=undefined&on=stream&module=inline&version=latest&contentPlacement=1&pgtype=collection>

143 Roose, "The Making".

144 Roose, "YouTube's Product"

145 Roose, "The Making".

146 Charlie Warzel, "Big Tech Was Designed to Be Toxic," *New York Times*, April 3, 2019, <https://www.nytimes.com/2019/04/03/opinion/facebook-youtube-disinformation.html>

- 147 Paris Martineau, "Maybe It's Not YouTube's Algorithm That Radicalizes People," *WIRED*, October 23, 2019, <https://www.wired.com/story/not-youtubes-algorithm-radicalizes-people/#>
- 148 Martineau, "Maybe It's",
- 149 Becca Lewis, "All of YouTube, Not Just the Algorithm, is a Far-Right Propaganda Machine," *FFWD* (blog), entry posted January 8, 2020, <https://ffwd.medium.com/all-of-youtube-not-just-the-algorithm-is-a-far-right-propaganda-machine-29b07b12430>
- 150 LaFrance, "The Algorithm".
- 151 Conversation with representatives from YouTube on March 2, 2020.
- 152 LaFrance, "The Algorithm".
- 153 Smith, Toor, and Kessel, "Many Turn," Pew Research Center.
- 154 Madrigal, "How YouTube's".
- 155 Madrigal, "How YouTube's".
- 156 Russell Brandom, "Inside Elsagate, The Conspiracy-Fueled War on Creepy YouTube Kids Videos," *The Verge*, December 8, 2017, <https://www.theverge.com/2017/12/8/16751206/elsagate-youtube-kids-creepy-conspiracy-theory>
- 157 Natasha Lomas, "YouTube Under Fire For Recommending Videos Of Kids With Inappropriate Comments," *TechCrunch*, February 18, 2019, <https://techcrunch.com/2019/02/18/youtube-under-fire-for-recommending-videos-of-kids-with-inappropriate-comments/>
- 158 Julia Alexander, "YouTube Still Can't Stop Child Predators In Its Comments," *The Verge*, February 19, 2019, <https://www.theverge.com/2019/2/19/18229938/youtube-child-exploitation-recommendation-algorithm-predators>
- 159 YouTube, "5 Ways We're Toughening Our Approach To Protect Families On YouTube and YouTube Kids," *Official YouTube Blog*, entry posted November 22, 2017, <https://youtube.googleblog.com/2017/11/5-ways-were-toughening-our-approach-to.html>
- 160 Alexander, "YouTube Still".
- 161 Bergen, "YouTube Executives".
- 162 Bergen, "YouTube Executives".
- 163 Elliott, "How To Turn".
- 164 YouTube, "Giving You More Control Over Your Homepage And Up Next Videos," *Official YouTube Blog*, entry posted June 26, 2019, <https://youtube.googleblog.com/2019/06/giving-you-more-control-over-homepage.html>
- 165 YouTube, "Giving You More," *Official YouTube Blog*.
- 166 YouTube, "Giving You More," *Official YouTube Blog*.
- 167 YouTube, "Giving You More," *Official YouTube Blog*.
- 168 "Manage Your," YouTube Help.
- 169 "Manage Your," YouTube Help.
- 170 "Manage Your," YouTube Help.
- 171 "Manage Your," YouTube Help.
- 172 Lydia De Pillis and Ivory Sherman, "Amazon's Extraordinary 25-Year Evolution," *CNN Business*, October 4, 2018, <https://www.cnn.com/interactive/2018/10/business/amazon-history-timeline/index.html>
- 173 "Fortune Global 500," *Fortune*, <https://fortune.com/global500/search/>

- 174 "Amazon.com Competitive Analysis, Marketing Mix and Traffic," Alexa Internet, <https://www.alexa.com/siteinfo/amazon.com>
- 175 Andrew Lipsman, "US Ecommerce 2019," eMarketer, last modified June 27, 2019, <https://www.emarketer.com/content/us-ecommerce-2019>
- 176 Julie Creswell, "How Amazon Steers Shoppers to Its Own Products," *The New York Times*, June 23, 2018, <https://www.nytimes.com/2018/06/23/business/amazon-the-brand-buster.html>
- 177 Martinez, "Amazon: Everything," IEEE Computer Society.
- 178 Stephanie Condon, "Amazon Shares How It Leverages AI Throughout The Business," *ZDNet*, June 5, 2019, <https://www.zdnet.com/article/amazon-shares-how-it-leverages-ai-throughout-the-business/>
- 179 "The Amazon Recommendations Secret to Selling More Online," Rejoinder, <http://rejoinder.com/resources/amazon-recommendations-secret-selling-online/>
- 180 Martinez, "Amazon: Everything," IEEE Computer Society.
- 181 Greg Linden, Brent Smith, and Jeremy York, *Amazon.com Recommendations: Item-to-Item Collaborative Filtering* (e IEEE Computer Society, 2003), <http://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf>
- 182 Shabana Arora, "Recommendation Engines: How Amazon and Netflix Are Winning the Personalization Battle," Martech Advisors, last modified June 28, 2016, <https://www.martechadvisor.com/articles/customer-experience-2/recommendation-engines-how-amazon-and-netflix-are-winning-the-personalization-battle/>
- 183 Arora, "Recommendation Engines," Martech Advisors.
- 184 Martinez, "Amazon: Everything," IEEE Computer Society.
- 185 "The Amazon," Rejoinder.
- 186 Amazon, "Improve Your Recommendations," Help & Customer Service, https://www.amazon.com/gp/help/customer/display.html/ref=hp_left_v4_sib?ie=UTF8&nodeId=GWXKS5URFHJUVUDN
- 187 Amazon, "Improve Your," Help & Customer Service.
- 188 Linden, Smith, and York, *Amazon.com Recommendations*.
- 189 JP Mangalindan, "Amazon's Recommendation Secret," *Fortune Magazine*, July 30, 2012, <https://fortune.com/2012/07/30/amazons-recommendation-secret/>
- 190 Linden, Smith, and York, *Amazon.com Recommendations*.
- 191 "The Amazon," Rejoinder.
- 192 This is a customized page that is available to each logged in user.
- 193 "The Amazon," Rejoinder.
- 194 "The Amazon," Rejoinder.
- 195 "The Amazon," Rejoinder.
- 196 "The Amazon," Rejoinder.
- 197 "The Amazon," Rejoinder.
- 198 "The Amazon," Rejoinder.
- 199 "The Amazon," Rejoinder.
- 200 "The Amazon," Rejoinder.
- 201 "The Amazon," Rejoinder.
- 202 Mangalindan, "Amazon's Recommendation".

- 203 Mangalindan, "Amazon's Recommendation".
- 204 Mangalindan, "Amazon's Recommendation".
- 205 Mangalindan, "Amazon's Recommendation".
- 206 Ian MacKenzie, Chris Meyer, and Steve Noble, "How Retailers Can Keep Up With Consumers," McKinsey & Company, <https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers#>
- 207 DiResta, "How Amazon's".
- 208 QAnon is a far-right conspiracy theory that suggests the "deep state" is secretly plotting against current U.S. President Donald Trump. The conspiracy theory first surfaced in October 2017 on 4chan.
- 209 Ben Collins, "On Amazon, a Qanon Conspiracy Book Climbs The Charts — With An Algorithmic Push," *NBC News*, March 4, 2019, <https://www.nbcnews.com/tech/tech-news/amazon-qanon-conspiracy-book-climbs-charts-algorithmic-push-n979181>
- 210 Collins, "On Amazon".
- 211 Collins, "On Amazon".
- 212 Nematzadeh et al., *How Algorithmic*.
- 213 Mangalindan, "Amazon's Recommendation".
- 214 Collins, "On Amazon".
- 215 DiResta, "How Amazon's".
- 216 DiResta, "How Amazon's".
- 217 Sidney Fussell, "Algorithms Are People," *The Atlantic*, September 18, 2019, <https://www.theatlantic.com/technology/archive/2019/09/is-amazons-search-algorithm-biased-its-hard-to-prove/598264/>
- 218 Elizabeth Dwoskin and Craig Timberg, "How Merchants Use Facebook To Flood Amazon With Fake Reviews," *The Washington Post*, April 23, 2018, https://www.washingtonpost.com/business/economy/how-merchants-secretly-use-facebook-to-flood-amazon-with-fake-reviews/2018/04/23/5dad1e30-4392-11e8-8569-26fda6b404c7_story.html?utm_term=.c3927b4ed0ca
- 219 DiResta, "How Amazon's".
- 220 Dwoskin and Timberg, "How Merchants".
- 221 <https://www.buzzfeednews.com/article/leticiamiranda/amazon-marketplace-sellers-black-hat-scams-search-rankings>
- 222 DiResta, "How Amazon's".
- 223 DiResta, "How Amazon's".
- 224 Brandy Zadrozny, "Amazon Removes Books Promoting Autism Cures And Vaccine Misinformation," *NBC News*, March 12, 2019, <https://www.nbcnews.com/tech/internet/amazon-removes-books-promoting-autism-cures-vaccine-misinformation-n982576>
- 225 DiResta, "How Amazon's".
- 226 April Glaser, "Amazon Is Suggesting 'Frequently Bought Together' Items That Can Make a Bomb," *Slate*, September 20, 2017, <https://slate.com/technology/2017/09/amazons-algorithm-is-suggesting-items-frequently-bought-together-that-can-make-a-bomb.html> , Siobhan Kennedy, "Potentially Deadly Bomb Ingredients Are 'Frequently Bought Together' On Amazon," *Channel 4 News*, September 18, 2017, <https://www.channel4.com/news/potentially-deadly-bomb-ingredients-on-amazon>
- 227 Glaser, "Amazon Is Suggesting".
- 228 Glaser, "Amazon Is Suggesting".
- 229 Annie Palmer, "Amazon Kills Program That Sent Shoppers Free Items Based On Prior Purchases," *CNB*

C, November 26, 2019, [<https://www.cnn.com/2019/11/26/amazon-axes-program-that-sent-shoppers-free-items-based-on-purchases.html>]

230 Amazon, "Recommendations," Help & Customer Service, https://www.amazon.com/gp/help/customer/display.html/ref=hp_left_v4_sib?ie=UTF8&nodeId=GE4KRSZ4KAZZB4BV

231 Amazon, "Improve Your," Help & Customer Service.

232 Amazon, "Improve Your," Help & Customer Service.

233 "About Netflix," Netflix Media Center, <https://media.netflix.com/en/about-netflix>

234 Ingrid Lunden, "Netflix Sharpens Focus on DVDs With DVD.com, But Don't Cry Qwikster. (It's Staying)," TechCrunch, March 30, 2012, <https://techcrunch.com/2012/03/30/netflix-sharpens-focus-on-dvds-with-dvd-com-but-dont-cry-qwikster-its-staying/>

235 Dan Moskowitz, "Who Are Netflix's Main Competitors?," Investopedia, last modified October 28, 2019, <https://www.investopedia.com/articles/markets/051215/who-are-netflixs-main-competitors-nflx.asp>

236 Mike Snider, "Netflix Adds 8.8 Million New Subscribers, 'The Witcher' Tracks As Most-Viewed New Series," *USA Today*, January 21, 2020, <https://www.usatoday.com/story/tech/2020/01/21/netflix-adds-8-8-million-new-subscribers-now-has-100-m-outside-us/4490819002/>

237 "Netflix.com Competitive Analysis, Marketing Mix and Traffic," Alexa Internet, <http://www.alexa.com/siteinfo/netflix.com>

238 Nielsen, *Nielsen Total Audience Report*, February 2020, <https://www.nielsen.com/us/en/client-learning/tv/nielsen-total-audience-report-february-2020/>

239 Over-the-top streaming services are streaming services that are available to users directly via the internet, rather than through outlets such as cable. Examples of other such streaming services are YouTube and Hulu.

240 Carlos A. Gomez-Uribe and Neil Hunt, "The Netflix Recommender System: Algorithms, Business Value, and Innovation," *ACM Transactions on Management Information Systems* 6, no. 4 (December 2015): <https://dl.acm.org/doi/pdf/10.1145/2843948?download=true>

241 Paul Sawers, "Remember Netflix's \$1m Algorithm Contest? Well, Here's Why It Didn't Use The Winning Entry.," *The Next Web*, April 13, 2012, <https://thenextweb.com/media/2012/04/13/remember-netflixs-1m-algorithm-contest-well-heres-why-it-didnt-use-the-winning-entry/>, Gomez-Uribe and Hunt, "The Netflix".

242 Sawers, "Remember Netflix's".

243 Titiriga, "Social Transparency".

244 Titiriga, "Social Transparency".

245 Andreas Toscher, Michael Jahrer, and Robert M. Bell, *The BigChaos Solution to the Netflix Grand Prize*, September 5, 2009, https://www.netflixprize.com/assets/GrandPrize2009_BPC_BigChaos.pdf

246 Sawers, "Remember Netflix's".

247 Gomez-Uribe and Hunt, "The Netflix".

248 Roettgers, "Netflix Spends".

249 Janko Roettgers, "Netflix Spends \$150 Million On Content Recommendations Every Year," *Gigaom*, October 9, 2014, <https://gigaom.com/2014/10/09/netflix-spends-150-million-on-content-recommendations-every-year/>

250 Gomez-Uribe and Hunt, "The Netflix".

- 251 Sawers, "Remember Netflix's".
- 252 Gomez-Uribe and Hunt, "The Netflix".
- 253 Gomez-Uribe and Hunt, "The Netflix".
- 254 Sawers, "Remember Netflix's".
- 255 "Netflix Ratings & Recommendations," Netflix Help Center, <https://help.netflix.com/en/node/9898>
- 256 "How Netflix's Recommendations System Works," Netflix Help Center, <https://help.netflix.com/en/node/100639>
- 257 Gomez-Uribe and Hunt, "The Netflix".
- 258 "Netflix Ratings," Netflix Help Center.
- 259 "How Netflix's," Netflix Help Center.
- 260 Libby Plummer, "This Is How Netflix's Top-Secret Recommendation System Works," *WIRED*, August 22, 2017, <https://www.wired.co.uk/article/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like>
- 261 "How Netflix's," Netflix Help Center.
- 262 "How Netflix's," Netflix Help Center.
- 263 Gomez-Uribe and Hunt, "The Netflix".
- 264 Gomez-Uribe and Hunt, "The Netflix".
- 265 Gomez-Uribe and Hunt, "The Netflix".
- 266 Gomez-Uribe and Hunt, "The Netflix".
- 267 Gomez-Uribe and Hunt, "The Netflix".
- 268 Gomez-Uribe and Hunt, "The Netflix".
- 269 Gomez-Uribe and Hunt, "The Netflix".
- 270 Gomez-Uribe and Hunt, "The Netflix".
- 271 Ekstrand et al., "All The Cool".
- 272 Netflix, "Now - For The First Time - You Can See What's Popular On Netflix," Netflix Media Center, last modified February 24, 2020, <https://media.netflix.com/en/company-blog/see-whats-popular-on-netflix>
- 273 "How Netflix's," Netflix Help Center.
- 274 "Netflix Ratings," Netflix Help Center.
- 275 Gomez-Uribe and Hunt, "The Netflix".
- 276 Gina Barton, "Why Your Netflix Thumbnails Don't Look Like Mine," *Vox*, November 21, 2018, <https://www.vox.com/2018/11/21/18106394/why-your-netflix-thumbnail-coverart-changes>
- 277 "How Netflix's," Netflix Help Center.
- 278 Gomez-Uribe and Hunt, "The Netflix".
- 279 Plummer, "This Is How Netflix's".
- 280 Plummer, "This Is How Netflix's".
- 281 Plummer, "This Is How Netflix's".
- 282 Nicole Nguyen, "Netflix Wants To Change The Way You Chill," *BuzzFeed News*, December 13, 2018, <https://www.buzzfeednews.com/article/nicolenguyen/netflix-recommendation-algorithm-explained-binge-watching>
- 283 Gomez-Uribe and Hunt, "The Netflix".
- 284 Nosheen Iqbal, "Film Fans See Red Over Netflix 'Targeted' Posters For Black Viewers," *The Guardian*, October 20, 2018, <https://www.theguardian.com/media/2018/oct/20/netflix-film-black-viewers-personalised-marketing-target>
- 285 Iqbal, "Film Fans".
- 286 Iqbal, "Film Fans".
- 287 Iqbal, "Film Fans".

288 Andrea Peterson, "Librarians Won't Stay Quiet About Government Surveillance," *The Washington Post*, October 3, 2014, <https://www.washingtonpost.com/news/the-switch/wp/2014/10/03/librarians-wont-stay-quiet-about-government-surveillance/>

289 April Glaser, "Long Before Snowden, Librarians Were Anti-Surveillance Heroes," *Slate*, June 3, 2015, <https://slate.com/technology/2015/06/usa-freedom-act-before-snowden-librarians-were-the-anti-surveillance-heroes.html>

290 Gomez-Urbe and Hunt, "The Netflix", Plummer, "This Is How Netflix's".

291 Ranking Digital Rights, an affiliate program at New America, has released a set of draft indicators which seek to measure corporate disclosures related to algorithmic systems. Our research concluded recommendations that are in line with many of their research-based indicators."RDR Corporate Accountability Index: Draft Indicators," Ranking Digital Rights, last modified October 2019, https://rankingdigitalrights.org/wp-content/uploads/2019/10/RDR-Index-Draft-Indicators_-Targeted-advertising-algorithms.pdf



This report carries a Creative Commons Attribution 4.0 International license, which permits re-use of New America content when proper attribution is provided. This means you are free to share and adapt New America's work, or include our content in derivative works, under the following conditions:

- **Attribution.** You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

For the full legal code of this Creative Commons license, please visit **creativecommons.org**.

If you have any questions about citing or reusing New America content, please visit **www.newamerica.org**.

All photos in this report are supplied by, and licensed to, **shutterstock.com** unless otherwise stated. Photos from federal government sources are used under section 105 of the Copyright Act.