The Asian Thinker
A Quarterly Bilingual Peer-Reviewed Journal for Social Sciences and Humanities
Year-8 Volume: II, April-June, 2026 Issue-30 ISSN: 2582-1296 (Online)
Website: www.theasianthinker.com **Email**: asianthinkerjournal@gmail.com

13. Feedback and Feed Forward Study of Role of AI-Based Recommendation Feedback in Social Media Platforms

Aniket Kasaudhan
Student, Journalism and Mass Communication
IIMT College of Management, Gr. Noida

Abstract

Artificial Intelligence has fundamentally changed the way social media platforms operate. No longer are these platforms simply spaces for people to share posts and connect with friends — they are now intelligent, adaptive systems that learn from every action a user takes and continuously adjust what content is shown to them. At the heart of this transformation lies the AI-based recommendation system, a technology that analyses user behaviour, preferences, and interaction history to suggest content that keeps users engaged for as long as possible.

This study examines the role of AI-based recommendation feedback in shaping user interaction on major social media platforms including Instagram, YouTube, Facebook, Twitter (now X), TikTok, and LinkedIn. It investigates how algorithms collect and process user data, how feedback loops are created and reinforced, and what the consequences of this process are — both for individual users and for society at large. The study also discusses the ethical concerns surrounding algorithmic curation, including issues of filter bubbles, misinformation amplification, mental health impacts, and user autonomy.

Drawing on existing academic literature, platform documentation, and publicly available research data, this paper presents a comprehensive analytical overview of AI recommendation systems in social media. The study argues that while these systems offer genuine benefits in terms of content discovery and personalisation, they also carry significant risks that are currently insufficiently addressed by either platform operators or regulators. The paper concludes with a set of recommendations for more responsible, transparent, and human-centred AI recommendation design.

Keywords: Artificial Intelligence, Recommendation System, Social Media Algorithm, User Interaction, Filter Bubble

Introduction

In less than two decades, social media has gone from being a simple communication tool to one of the most powerful forces shaping human attention, opinion, and behaviour. Platforms like Facebook, YouTube, Instagram, TikTok, and Twitter command billions of active users globally,

The Asian Thinker
A Quarterly Bilingual Peer-Reviewed Journal for Social Sciences and Humanities
Year-8 Volume: II, April-June, 2026 Issue-30 ISSN: 2582-1296 (Online)
Website: www.theasianthinker.com **Email**: asianthinkerjournal@gmail.com

and the time people spend on these platforms is measured not in minutes but in hours each day. This extraordinary hold over human attention did not happen by accident — it was engineered, deliberately and sophisticatedly, using Artificial Intelligence.

At the centre of this engineering effort is the recommendation system. A recommendation system is an algorithmic technology that looks at what a user has done on a platform — what they watched, liked, shared, searched for, paused on, or scrolled past — and uses that information to predict what content they are likely to engage with next. It then surfaces that content, creating a personalised feed that is unique to every individual user.

The implications of this technology are vast. For users, it means an experience that feels intuitive and relevant — the platform seems to 'know' what they want. For platforms, it means dramatically higher engagement metrics: longer sessions, more content consumed, more advertisements seen, and greater revenue. But for society, it raises a set of deeply important and troubling questions about the nature of information consumption, democratic discourse, mental health, and the manipulation of human psychology at scale.

This study explores these questions systematically. It begins by explaining what AI-based recommendation systems are and how they work at a technical level, before examining the ways in which user interaction feeds back into and shapes the algorithm — a process known as the feedback loop. It then analyses the specific recommendation systems used by major platforms, the psychological effects on users, the social consequences such as filter bubbles and misinformation, and the ethical and regulatory debates that have emerged in response.

The study is designed to be accessible to anyone — regardless of their technical or academic background — while maintaining the rigour expected of academic research. It draws on established communication theory, computer science concepts, published research studies, and documented platform policies to construct a thorough and honest analysis of one of the most consequential technologies of our time.

1.1 Scope and Significance of the Study

Social media recommendation algorithms now influence not only what entertainment people consume but also what news they read, what political opinions they encounter, what products they buy, and what beliefs they form about the world. Understanding how these systems work is no longer a niche academic interest — it is a pressing civic and democratic necessity.

This study is particularly relevant to students and practitioners of journalism and mass communication, who work at the intersection of content creation, distribution, and audience engagement. As the gatekeeping function traditionally performed by editors and journalists is increasingly supplemented or replaced by algorithmic curation, media professionals need to understand the logic of recommendation systems to understand the information environment in which they operate.

2. Understanding AI-Based Recommendation Systems

2.1 What is a Recommendation System?

A recommendation system — sometimes called a recommender system — is a type of artificial intelligence application designed to predict what a particular user will find interesting or useful, and to present that content to them proactively. The concept is not new: libraries have always recommended books, and friends have always suggested films. What makes the AI-based version revolutionary is its scale, speed, and sophistication.

A modern social media recommendation system can analyse millions of pieces of content and match them to millions of users in real time, using dozens of different signals simultaneously. It does this not by following a set of fixed rules written by human programmers, but by learning patterns from data — which is the defining characteristic of machine learning, the branch of AI that underpins these systems.

2.2 The Three Core Approaches

There are three fundamental approaches to building recommendation systems, and most modern platforms use a combination of all three:

2.2.1 Collaborative Filtering

Collaborative filtering works on the principle that people who have agreed in the past tend to agree in the future. If User A and User B have both liked a similar set of videos, and User A then watches and enjoys a new video, the system assumes that User B might enjoy it too, and recommends it to them. This approach does not require the system to understand anything about the actual content of the video — it only needs to understand patterns of user behaviour.

The limitation of collaborative filtering is the 'cold start' problem: when a new user joins a platform, the system has no behavioural data for them, and therefore cannot make accurate collaborative recommendations. Similarly, when a piece of new content is posted, it has not yet been interacted with by enough users for collaborative patterns to form.

2.2.2 Content-Based Filtering

Content-based filtering takes a different approach: instead of looking at patterns of user behaviour, it analyses the characteristics of the content itself. A content-based system might analyse the genre, duration, language, and keywords of a video, and then recommend other videos with similar characteristics to a user who has previously engaged with content of that type.

The Asian Thinker
A Quarterly Bilingual Peer-Reviewed Journal for Social Sciences and Humanities
Year-8 Volume: II, April-June, 2026 Issue-30 ISSN: 2582-1296 (Online)
Website: www.theasianthinker.com **Email**: asianthinkerjournal@gmail.com

This approach handles the cold start problem better for content — a new video can be recommended as soon as its characteristics are analysed — but it has a different limitation: it tends to create a narrow, repetitive experience in which users are shown content very similar to what they have already seen, limiting exposure to new or unexpected content.

2.2.3 Hybrid Systems

Modern platforms invariably use hybrid systems that combine collaborative and content-based filtering with additional layers of machine learning. These hybrid systems can incorporate contextual signals (time of day, device being used, location), social signals (who the user follows and what their contacts are engaging with), and real-time behavioural signals (how long the user paused on a piece of content, whether they replayed it, whether they shared it).

Approach	Core Logic	Key Limitation
Collaborative Filtering	Similar users = similar interests	Cold start for new users
Content-Based Filtering	Similar content = relevant to user	Creates narrow, repetitive feed
Hybrid System	Combination of both + contextual data	Complexity and opacity
Deep Learning Models	Neural networks find hidden patterns	Requires massive data, hard to explain

Table 2.1: Comparison of Core Recommendation Approaches

3. How Algorithms Work: The Technical Framework

3.1 Data Collection: The Foundation of the Algorithm

Every recommendation algorithm begins with data. Without data about what users do on the platform, the algorithm has nothing to learn from and nothing to base its predictions on. Social media platforms are extraordinarily effective data collection machines, capturing a vast range of user signals with every interaction.

These signals can be divided into two broad categories: explicit signals and implicit signals. Explicit signals are actions that users take consciously and deliberately — pressing the like button, leaving a comment, sharing a post, subscribing to a channel, or rating content. Implicit signals are actions that are recorded without the user necessarily intending to provide feedback — how long they watched a video before scrolling away, whether they opened a link, how often they return to a specific creator's content, or whether their eyes paused on a particular advertisement.

Modern platforms have become highly sophisticated in collecting and interpreting implicit signals, because these signals are far more abundant and, arguably, more honest than explicit ones. A user might like a post out of social obligation, but the fact that they watched a particular type of video for twenty minutes in the middle of the night reveals something deeper about their genuine interests.

3.2 The Machine Learning Process

Once data is collected, it is fed into machine learning models that are trained to predict user engagement. The most common objective function — the goal that the model is trained to optimise — is engagement: the algorithm learns to predict which content a given user is most likely to interact with (watch, like, share, comment on), and it surfaces that content preferentially.

This training process involves the algorithm making predictions, comparing those predictions to what actually happened (did the user engage with the recommended content or not?), and adjusting its internal parameters to make better predictions in future. Over millions of iterations across billions of user interactions, the model becomes extraordinarily good at predicting what will hold a user's attention.

Deep learning models, and particularly neural networks, have significantly advanced the capability of recommendation systems in recent years. These models can identify complex, non-obvious patterns in data — for example, detecting that users who watch a particular type of cooking video at night are also likely to be interested in a specific genre of travel content — that would be impossible to identify through manual rule-writing.

3.3 The Engagement Optimisation Problem

The central technical choice that shapes everything else about a recommendation system is the choice of what to optimise for. Most commercial social media platforms optimise for engagement — the quantity and depth of user interaction with the platform. This makes sense from a business perspective: engagement translates directly into advertising revenue. But it creates a problem that is now widely documented in academic literature and technology journalism.

Optimising for engagement does not always mean optimising for content that users find genuinely valuable, informative, or enriching. Research has consistently shown that content that provokes strong emotional reactions — outrage, fear, anxiety, moral indignation — tends to generate more engagement than calm, factual, or nuanced content. An algorithm optimising for engagement will therefore systematically surface emotionally provocative content, not because it was designed to do harm, but because it was designed to maximise interaction, and emotionally provocative content generates interaction.

3.4 Ranking and Scoring Systems

The practical output of the recommendation algorithm is a ranked list of content items. For every user at every moment, the system assigns a predicted engagement score to every piece of potentially relevant content, and then ranks that content from highest to lowest predicted score. The top-ranked items are what the user sees first in their feed or in their 'For You' or 'Recommended' sections.

These scores are typically composite scores that weight multiple signals. For example, YouTube's recommendation system is known to weight watch time heavily — a video that users tend to watch all the way through receives a higher score than one that users abandon halfway. Instagram weights recency, relationship closeness, and content type. TikTok, perhaps the most sophisticated recommendation system currently in widespread use, places extraordinary weight on completion rate — whether users watched a short video from start to finish — making it particularly effective at identifying and surfacing content that holds attention. 4. User Interaction and Feedback Loops

4.1 What is a Feedback Loop?

A feedback loop, in the context of AI recommendation systems, refers to the self-reinforcing cycle that is created when user behaviour shapes the algorithm, and the algorithm in turn shapes user behaviour. This is one of the most important and consequential aspects of how social media recommendation systems operate, and it is essential to understanding both their power and their risks.

The basic logic of the feedback loop works as follows. A user interacts with a piece of content — say, they watch a video about a particular political topic all the way through. The algorithm notes this signal and infers that the user is interested in that topic. It then recommends more content on the same topic. The user watches this content too, reinforcing the signal. The algorithm recommends even more content on the topic, and shows the user less and less content on other subjects. Over time, the user's feed becomes increasingly dominated by content related to that one topic — not because the user consciously chose this, but because the feedback loop amplified an initial interaction into a content monoculture.

4.2 Types of Feedback Signals

Signal Type	Example Actions	Algorithm Interpretation
Strong Positive	Share, Save, Comment, Subscribe	High interest — prioritise similar content strongly
Moderate Positive	Like, Full watch, Replay	Interest confirmed — increase similar content
Weak Positive	Click, Partial watch (>50%)	Potential interest — test with more similar content
Neutral / Ambiguous	Impression (seen but not acted on)	Uncertain — may reduce frequency

Weak Negative	Partial watch (<30%), Skip	Low interest — reduce similar content slightly
Strong Negative	Hide post, 'Not interested', Mute	Strong disinterest — significantly reduce similar content
Explicit Report	Report as inappropriate, Block	Content violation flag — remove and flag creator

4.3 The Positive Feedback Loop: Personalisation and Satisfaction

Not all feedback loops are harmful. The fundamental premise of personalisation — that users should see content relevant to their genuine interests — is a reasonable and user-serving goal. When a person who loves cooking discovers a new chef's channel through a recommendation, or when a student interested in science stumbles upon an educational creator they had never heard of, the recommendation system has performed its intended function. It has connected a user with content they value, and in doing so has served both the user and the content creator.

Research by Spotify and Netflix — two platforms with highly sophisticated recommendation systems — has documented that recommendation-driven content discovery leads to higher user satisfaction and greater willingness to engage with unfamiliar content than browsing or searching alone. In this sense, well-designed recommendation feedback can expand rather than narrow a user's content universe.

4.4 The Negative Feedback Loop: Radicalisation and Reinforcement

The darker consequence of feedback loops is their potential to reinforce and amplify harmful content consumption patterns. A user who watches one conspiracy theory video may find their feed progressively populated with more extreme versions of similar content, not because they sought this out, but because each engagement reinforced the algorithm's inference about their interests.

This process — sometimes called algorithmic radicalisation — has been studied in the context of political extremism, health misinformation, and body image disorders. The algorithm does not intend to radicalise; it intends to engage. But when engaging content happens to be extreme or harmful, the feedback loop amplifies it. Guilluy (2019) and Ribeiro et al. (2020) both documented evidence of YouTube's recommendation system creating pathways from mainstream political content to increasingly extremist material, though the extent and mechanism of this process remains debated in academic literature.

5. Platform-wise Analysis of Recommendation Systems

5.1 YouTube: Watch Time and the Recommendation Rabbit Hole

YouTube is the world's largest video-sharing platform and operates one of the most influential recommendation systems in existence. With over 2.7 billion logged-in users monthly and over 500 hours of video uploaded every minute, YouTube's recommendation system performs the essential function of connecting users with relevant content from an inconceivably large library.

YouTube's algorithm has evolved significantly since the platform was founded in 2005. Until 2012, the system optimised primarily for clicks — any click on a recommended video counted as a success. This created a strong incentive for 'clickbait' content with sensational thumbnails and misleading titles. In 2012, YouTube shifted its optimisation objective to watch time, reasoning that if users were watching a video rather than clicking away, the recommendation had been genuinely successful. This reduced some of the worst clickbait content but created new problems: it rewarded content that was good at holding attention over long periods, which, as discussed earlier, often means emotionally stimulating or anxiety-provoking content.

Since 2016, YouTube has incorporated a 'satisfaction' signal into its algorithm — using survey data from users about whether they were satisfied with a recommendation — in an attempt to align engagement metrics with genuine user satisfaction. This is an important methodological step, though critics argue that the incentive structure of the platform continues to favour engagement over quality.

5.2 Instagram and Facebook: Social Graph and Interest Signals

Meta's platforms — Instagram and Facebook — use recommendation systems that blend social graph signals (who you follow and who follows you) with interest signals (what content you engage with). For many years, both platforms showed users content primarily from accounts they followed, in reverse chronological order. The shift to algorithmic feeds — which Meta introduced to Facebook around 2009 and to Instagram around 2016 — was enormously controversial among creators and users who felt they had lost control of what they saw.

Facebook's News Feed algorithm, known internally as EdgeRank in its original form, scores content based on three factors: affinity (how close the relationship between the user and the content creator is), weight (how much engagement the post has received), and time decay (how recent the post is). Later iterations incorporated hundreds of additional signals and machine learning layers. Instagram's algorithm similarly considers relationship closeness, interest matching, and recency, but places particularly high weight on Reels — short-form videos — reflecting the platform's strategic competition with TikTok.

5.3 TikTok: The Interest Graph Revolution

TikTok represents perhaps the most significant innovation in social media recommendation in recent years. Unlike platforms that build primarily on the social graph — showing users content from people they know or follow — TikTok is built around what it calls the interest graph: a model of what content a user finds interesting, derived almost entirely from implicit behavioural signals.

The Asian Thinker
A Quarterly Bilingual Peer-Reviewed Journal for Social Sciences and Humanities
Year-8 Volume: II, April-June, 2026 Issue-30 ISSN: 2582-1296 (Online)
Website: www.theasianthinker.com **Email**: asianthinkerjournal@gmail.com

A new TikTok user does not need to follow anyone to receive highly personalised recommendations. Within hours of joining, and after watching only a handful of videos, TikTok's algorithm can build a sufficiently accurate interest model to surface content that feels uncannily relevant. This is achieved through the platform's emphasis on completion rate — whether users watch short videos from beginning to end — which provides a clean, high-frequency signal about what content genuinely holds their attention.

TikTok's recommendation system has been credited with enabling content creators without any existing following to achieve viral reach overnight, and with providing users with a more diverse discovery experience than social-graph-based platforms. It has also attracted intense scrutiny from regulators in the United States, India, and Europe on concerns related to data privacy and national security, given its ownership by Chinese company ByteDance.

5.4 Twitter/X and LinkedIn: Real-Time and Professional Contexts

Twitter (rebranded as X in 2023) has historically emphasised real-time content and chronological feeds, but has increasingly incorporated algorithmic recommendations. Its 'For You' tab, which became the default feed after Elon Musk's acquisition of the platform in 2022, uses engagement signals and interest modelling to surface content from outside a user's following. This shift was controversial among users who valued the predictability and transparency of the chronological feed.

LinkedIn's recommendation system operates in a professional context, optimising for career-relevant content and professional network development. It uses a combination of professional graph signals (connections, industries, job titles) and content engagement signals to surface posts, articles, and job opportunities. LinkedIn's system is distinctive in that it explicitly tries to balance virality with quality — it applies a 'human review' layer on top of its algorithmic recommendations to prevent the spread of low-quality or inappropriate content.

6. The Role of Data in Personalisation

6.1 What Data is Collected?

The richness of personalisation on social media platforms is a direct function of the richness of the data collected about users. This data collection is pervasive, continuous, and extends far beyond what most users are aware of. Understanding what data is collected, and how it is used, is essential to understanding both the power and the potential harms of AI recommendation systems.

The primary categories of data collected by social media platforms include: behavioural data (every action taken on the platform), profile data (demographic information, stated interests, location), device data (type of device, operating system, browser, IP address), contextual data (time of day, location if permitted, network type), cross-platform data (data shared by third-party websites and apps through tracking pixels and APIs), and inferred data (characteristics and interests that the platform infers from behavioural patterns, even if the user never explicitly provided this information).

6.2 Inferred vs. Declared Data

A particularly important distinction is between declared data — information a user consciously provides — and inferred data — information the algorithm deduces from behaviour. Platforms have become extraordinarily adept at inferring highly sensitive personal characteristics from relatively innocuous-seeming behavioural data.

Research published by Kosinski, Stillwell, and Graepel (2013) in the Proceedings of the National Academy of Sciences demonstrated that Facebook likes alone could accurately predict users' sexual orientation, political affiliation, religious views, intelligence, and personality traits with startling accuracy. Users who had not disclosed any of this information were nonetheless profiled for it based on their engagement patterns. This finding illustrates a fundamental asymmetry in AI recommendation systems: the platform knows far more about the user than the user knows about what the platform knows.

6.3 The Data Economy of Social Media

Platform	Primary Data Types Used	Est. Data Points per User	Third-Party Data?
Facebook / Meta	Social graph, engagement, location, off-platform activity	~52,000	Yes (extensive)
Google / YouTube	Search history, watch history, location, device data	~70,000	Yes (across Google services)
TikTok	Watch time, replay, device signals, clipboard access (historical)	~30,000+	Limited
Instagram	Engagement, social graph, hashtags, story interactions	~35,000	Yes (Meta network)
Twitter / X	Follow graph, engagement, search queries, location	~20,000	Moderate
LinkedIn	Professional profile, connections, job activity, content engagement	~18,000	Limited

Table 6.1: Comparative Data Collection Across Major Social Media Platforms (Estimated)

6.4 Consent, Transparency, and Data Rights

The collection of this data is technically consented to through terms of service agreements, but research consistently shows that users do not read these agreements and have little genuine understanding of what they are consenting to. A 2019 study by Obar and Oeldorf-Hirsch found that users spent an average of 73 seconds 'reading' a terms of service document before clicking accept — far too short to comprehend documents that routinely run to 10,000 words or more.

This gap between formal consent and meaningful informed consent is at the heart of the debate around data rights in the digital age. The European Union's General Data Protection Regulation (GDPR), implemented in 2018, represents the most ambitious attempt to date to give users genuine rights over their data, including the right to access it, correct it, and request its deletion. India's Digital Personal Data Protection Act (2023) represents a similar legislative effort in the Indian context.

7. Regulatory Landscape and Policy Debates

7.1 The Global Regulatory Response

The recognition that AI-based recommendation systems carry significant risks — for individual users, for democratic discourse, and for social cohesion — has prompted regulatory responses across multiple jurisdictions. These regulatory efforts reflect different values and priorities and have produced a varied and still-evolving global policy landscape.

Regulation / Policy	Jurisdiction	Key Requirements	Status (2025)
General Data Protection Regulation (GDPR)	European Union	Data minimisation, consent, right to explanation, data portability	In force (2018)
Digital Services Act (DSA)	European Union	Algorithmic transparency, non-personalised feed option, risk assessments	In force (2024)
Digital Personal Data Protection Act	India	Consent-based data processing, data localisation provisions, user rights	In force (2023)
AI Act	European Union	Risk-based AI classification, transparency for high-risk AI systems	In force (2024)
Kids Online Safety Act (KOSA)	United States	Platform duties of care for minors, restrictions on harmful algorithmic recommendations	Under debate
Information Technology Rules	India	Grievance mechanisms, content takedown requirements, traceability	In force (amended 2023)

7.2 The Indian Regulatory Context

India presents a particularly complex and important regulatory context for social media recommendation systems. With over 800 million internet users — making it the world's largest social media market by number of users — the Indian policy environment around social media has significant global implications.

India's Information Technology Act (2000) and its subsequent amendments have established a framework for platform liability and content regulation. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, imposed new obligations on significant social media intermediaries (those with more than five million users in India), including requirements for grievance redressal mechanisms, content traceability (the ability to identify the origin of viral messages), and local offices with designated officers accountable to Indian law.

The Digital Personal Data Protection Act (DPDPA) of 2023 establishes a consent-based framework for personal data processing in India, with significant implications for the data collection practices that underpin recommendation systems. The Act introduces the concept of 'Data Principals' (users) and 'Data Fiduciaries' (platforms), establishing rights and obligations that echo GDPR principles while reflecting Indian priorities.

7.3 The Self-Regulation Debate

A central tension in the regulatory debate is between those who argue that government regulation is necessary to address the harms created by AI recommendation systems, and those who argue that platforms should be trusted to self-regulate. Proponents of self-regulation argue that government regulators lack the technical expertise to effectively oversee rapidly evolving AI systems, and that heavy-handed regulation risks stifling innovation. Critics of self-regulation argue that platforms have a fundamental conflict of interest — engagement-maximising recommendation systems are enormously profitable, and platforms cannot be trusted to voluntarily constrain systems that generate their revenue.

The evidence from the past decade tends to support the critics: platforms have repeatedly demonstrated willingness to prioritise engagement over user safety when left to self-regulate, and significant platform behaviour changes have generally followed only external pressure from regulators, advertisers, or major public controversies.

8. The Future of AI Recommendations in Social Media

8.1 Generative AI and Recommendation Systems

The rapid advancement of generative AI — AI systems capable of producing new content such as text, images, audio, and video — is beginning to intersect with recommendation systems in ways that have profound implications for the future of social media. Generative AI makes it possible to create content at scale and at very low cost, raising the prospect of social media feeds increasingly populated by AI-generated content rather than human-created content.

When AI systems both generate content and determine which content is recommended to which users, the entire information environment of social media becomes AI-mediated in a new and more fundamental sense. The potential for large-scale manipulation of public opinion through strategically generated and algorithmically targeted content is a significant and growing concern. The 2024 global election cycle — with major elections in India, the United States, the European Union, and the United Kingdom — provided early evidence of the challenges that AI-generated political content poses for electoral integrity.

8.2 Towards Responsible Recommendation Design

A growing body of research and practitioner work is exploring what responsible, human-centred recommendation design might look like. Several promising directions have been identified. Value-sensitive design approaches attempt to incorporate human values — beyond mere engagement — directly into the objective functions of recommendation systems, optimising for user wellbeing, information diversity, or civic engagement alongside or instead of pure engagement metrics.

Controllability — giving users greater ability to understand, adjust, and override their recommendation algorithms — is another important direction. Research has found that users who are given more control over their algorithmic feeds report higher satisfaction, greater trust in the platform, and lower feelings of manipulation. Some platforms have begun offering 'algorithmic transparency' features, though these remain limited and superficial in most cases.

9. Conclusion

This study has traced the arc of AI-based recommendation feedback in social media platforms — from the technical foundations of machine learning and collaborative filtering, through the complex dynamics of user interaction and feedback loops, to the profound social consequences of engagement-optimised algorithmic curation. The picture that emerges is one of a technology of extraordinary power and sophistication that is currently deployed in ways that serve commercial interests more reliably than human interests.

This is not because the engineers who build these systems are malicious, or because the platforms that deploy them are indifferent to user welfare. It is because the structural incentives of the attention economy — in which user attention is the product, engagement is the metric, and advertising is the revenue source — create a systematic pressure to optimise for what is engaging over what is good. An algorithm optimised purely for engagement will, predictably, surface content that provokes emotion over content that informs, content that confirms over content that challenges, and content that entertains over content that enriches.

The consequences of this structural misalignment are now well-documented: filter bubbles that narrow information exposure, misinformation that spreads faster than truth, mental health impacts particularly on young people, erosion of epistemic autonomy, and risks to democratic discourse. These are not isolated harms — they are systemic effects of a technology designed and deployed in a particular way.

The good news is that there is nothing inevitable about these outcomes. Recommendation systems can be designed differently — to optimise for a richer set of objectives that include user wellbeing, information quality, and content diversity alongside engagement. Regulators can establish frameworks that require transparency, accountability, and meaningful user control. Educators can build the media and algorithmic literacy that enables citizens to navigate these environments more consciously. And users, armed with greater understanding of how these systems work, can make more deliberate choices about their own digital information diets.

10. References

1. Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. Science, 348(6239), 1130–1132.
2. Bucher, T. (2018). If...Then: Algorithmic Power and Politics. Oxford University Press.
3. Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for YouTube recommendations. Proceedings of the 10th ACM Conference on Recommender Systems (RecSys 2016), 191–198.
4. Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7(2), 117–140.
5. Floridi, L. (2014). The Fourth Revolution: How the Infosphere is Reshaping Human Reality. Oxford University Press.
6. Haugen, F. (2021). The Facebook Papers [Whistleblower disclosure to the US Securities and Exchange Commission and the US Congress]. United States Congress.
7. Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences, 110(15), 5802–5805.
8. Lasswell, H. D. (1948). The structure and function of communication in society. In L. Bryson (Ed.), The Communication of Ideas (pp. 37–51). Harper.
9. McQuail, D. (2010). McQuail's Mass Communication Theory (6th ed.). Sage Publications.

10. Ministry of Electronics and Information Technology (MeitY). (2023). Digital Personal Data Protection Act. Government of India.
11. Obar, J. A., & Oeldorf-Hirsch, A. (2020). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. Information, Communication & Society, 23(1), 128–147.
12. Pariser, E. (2011). The Filter Bubble: What the Internet is Hiding from You. Penguin Press.
13. Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A. F., & Meira Jr., W. (2020). Auditing radicalization pathways on YouTube. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT*), 131–141.
14. Skinner, B. F. (1938). The Behavior of Organisms: An Experimental Analysis. Appleton-Century-Crofts.
15. Telecom Regulatory Authority of India (TRAI). (2022). Recommendations on Regulatory Framework for Over-the-Top (OTT) Communication Services. TRAI.