ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

Research on AI-Driven Information Cocoons in New Media and Their Breaking Strategies

Xiangyu Jiang[1], *Yanxin Lin*[2], *Xin Hong*[2], *Kun Hu*[2] *and Yijia Tang*[2,*]

[1] School of General Education, Yango University, Fuzhou 350015, Fujian, China
[2] School of Metaverse and New Media, Yango University, Fuzhou, 350015, Fujian, China

SUMMARY: *Given that artificial intelligence (AI) technology has now spread widely in the field of new media, personalised recommendation algorithms have reshaped the pattern of information distribution; at the same time, they have also given rise to or worsened the problem of "information cocoons". AI-driven information cocoons are chosen as the research objects, and in combination with theoretical analysis and case studies, their basic concepts and attributes have been defined, formation mechanisms under the three dimensions of technical logic, commercial logic and user psychology have been examined, and their adverse effects on individual cognition and social ecology have been investigated. Based on the above, three types of breaking strategies at the levels of government regulation, industry self-discipline and individual empowerment have been proposed to provide a theoretical basis and practical path for building an open and diverse new media information ecosystem.*

KEYWORDS: *Artificial Intelligence; New Media; Information Cocoons; Personalized Recommendation; Collaborative Governance*

1 Introduction

With the rapid development of AI technology, Big Data and machine learning are widely used to build algorithms that guide the new media environment today. AI-driven personalised recommendation systems are used to collect data on how users behave, match the relevant information, and spread it more efficiently to improve the user experience. Technology supports cognition but also presents the problem of excessive dependence. Users are increasingly surrounded by homogeneous content promoted through algorithms and have gradually fallen into a "information cocoon". A large-scale empirical study conducted by a research group at Tsinghua University found that more than 57% of active users had experienced some degree of reduction in information entropy after one year of online interaction; that is to say, the diversity of their information exposure had decreased. A typical problem in the governance of an intelligent society is now the phenomenon of "deep information cocoons".

The formation of information cocoons is not a new phenomenon; however, the introduction of AI technology has given them new characteristics such as automation, invisibility and self-reinforcement, and thus poses a dual threat to individual cognitive upgrading and the construction of social consensus. At present, most academic research on information cocoons has focused on describing phenomena and discussing single-dimensional countermeasures, and systematic theoretical studies for the specific case of AI-driven

*paf7200@163.com
https://doi.org/10.65102/is20261293 Published: 20 June, 2026

information cocoons are still lacking. Therefore, in-depth research has been carried out on the formation mechanism of AI-driven information cocoons in new media and the exploration of multi-subject collaborative breaking strategies to provide strong theoretical support and offer urgent practical guidance for regulating algorithm application, maintaining cyberspace order, and promoting harmonious social development. Based on theoretical research and a number of typical cases, this paper establishes a research framework of "mechanism analysis - hazard assessment - strategy construction" to provide systematic solutions for the problem of information cocoons in the era of artificial intelligence.

The present discussion of information cocoons should be viewed in light of the wider issue of personalisation and public communication. Pariser proposed that the filter bubble is a personalised digital space in which algorithmic optimisation may exclude new knowledge from users [1]. Sunstein believes that a lack of contact with all the public will make it more difficult for them to find common ground [2]. These reasons show that AI-enabled recommendations are not just technical optimisation devices, but also modes of communication capable of capturing public attention.

Based on empirical studies, algorithmic ranking and individual choice work together to reduce exposure diversity. A relatively large-scale study on the exposure of Facebook news showed that both platform ranking and users' own click habits influenced how widely cross-cutting political information spread [3]. This result directly relates to new media platforms, and the reasons for the recommendation outcome are not known to be algorithms. Ranking systems for machines and objectives of e-commerce platforms are also motivated by user behaviour.

Therefore, this revision keeps the original research object but strengthens its research design. The paper adds a literature review, a measurable index system, three mechanism figures, four analysis tables and six formulas. A modified system will be established to promote fact-based conversations about the formation process of AI-driven information cocoons, define indicators of severe cocoons, and foster an inclusive governance mode to address their adverse effects.

2 Core Concepts and Features of AI-driven Information Cocoons in New Media

2.1 Definitions of Core Concepts

The concept of information cocoons was first put forward by the American scholar Sunstein, who referred to a phenomenon where, due to their preferences for selective exposure in the process of obtaining information, individuals either voluntarily or involuntarily surround themselves with a circle of information that aligns with their own wishes and form cognitive closure. With the emergence of AI-driven new media, the concept of an "information cocoon" has expanded to describe a systematic communication crisis where AI technology, through personalised recommendation algorithms, continually provides users with uniform information based on the construction of user profiles, behaviour data mining and precise content matching; thus, users' cognitive boundaries have hardened, and their information horizons have been narrowed.

Compared with information cocoons in the traditional media era, the main difference of AI-driven information cocoons is that the driving subject has changed from active individual choice to passive algorithmic shaping. The formation of traditional information cocoons is driven by people's own interests, preferences and information-filtering behavior; on the other hand, artificial intelligence technology uses a closed-loop process of "data collection - model training - precise push - feedback reinforcement" to make the formation of information cocoons more subtle and inevitable. Users often unknowingly enter a state of cognitive solidification.

2.2 Core Features

First, the invisibility of automated generation. AI algorithms will build user profiles and recommend relevant content automatically through the processing of background data without manual operation. Users are unable to recognise that the information has been filtered. For example, by continuously watching a certain kind of content, short video platform users will find that their recommended feeds are gradually dominated by such content. However, people are likely to believe that "the platform knows me" and therefore fail to realise that algorithms also limit the scope of information. It is difficult to identify information cocoons in time because they are invisible; thus, they are more likely to solidify.

Second, a self-reinforcing closed-loop system. AI-driven information dissemination is a positive feedback loop of "user behaviour - data collection - algorithm optimisation - precise push". The clicks, likes, and other behaviours users have exhibited for a particular type of information are used by the algorithm as an optimisation basis, and then more similar content is pushed out. Conversely, if users neglect diverse information, algorithms will reduce the push of such content; thus, it forms a closed-loop effect of "the more it is used, the more specialised it becomes". The human-AI adaptive dynamics model put forward by Tsinghua University suggests that this closed-loop interaction will cause the system to gradually shift from a state of diversity to one of deep information cocoons.

Thirdly, amplification of the effect of group polarization. AI algorithms may enhance individual cognition, but at the same time, they can also promote group division by "circled recommendation"; by identifying users with similar opinions in the same information circle, algorithms gather such users to reinforce the spread of specific views within that circle and thus lead to the phenomenon of group polarization. For example, when discussing public issues in a particular area of society online, many algorithms have separated supporters from opponents into different information spaces, making communication between the two difficult or even fostering animosity.

2.3 Literature Review and Research Gap

The three groups of the existing research are as follows. The first group has been studying filter bubbles and echo chambers in the context of political communication. Flaxman, Goel and Rao have found that online news use can increase both the number of available sources and ideological separation among highly engaged users simultaneously [4]. Therefore, the Internet will not be inherently open, and platforms and user behaviours together determine how much diversity in access there is.

The second group is based on algorithms for personalisation. Haim, Graefe and Brosius investigated personalisation in Google News and discovered that although there will be some reduction in diversity, not all of it is eliminated; rather, the proportion of visible information changes [5]. Geschke, Lorenz and Holtz put forward a triple-filter-bubble system that combines individual cognition, social networks and algorithmic selection [6]. Based on the above research, information cocoons are better viewed as multiple-factor systems rather than a single technological result.

The third group covers echo chambers and recommender systems. Cinelli and his colleagues have studied how far echo chambers have developed on various social media sites and determined that these sites' structural features contribute to polarisation and content separation [7]. Nguyen and others have studied recommender systems and content diversity to find that recommendations can extend a user's experience of culture [8]. Recent systematic reviews have also shown that filter bubbles in recommender systems are related to exposure bias, popularity bias and over-specialised ranking [9]. Research on recommendation bias and debiasing has also shown that ranking models may reproduce historical exposure patterns without correction mechanisms [10].

Therefore, there is a lack of relevant studies. Many studies have investigated either the psychology of users or recommendation algorithms, but fewer have combined algorithmic logic, platform monetization and user cognition in a governance-oriented way. Therefore, this paper will link the explanation of mechanisms to operating indicators and multiple governance models.

Table 1: Comparison of Main Research Directions on Information Cocoons.

Perspective	Main focus	Typical contribution	Limitation
Filter bubble	Personalized filtering of information	Explains invisible narrowing of exposure	Often underestimates user agency
Echo chamber	Social reinforcement among similar users	Explains group polarization and identity closure	May overlook algorithmic ranking
Recommender systems	User-content matching and ranking	Provides measurable technical mechanism	May prioritize accuracy over diversity
This paper	AI-driven cocoon in new media	Integrates algorithm, platform and cognition	Requires future platform data validation

Table 1 shows the theoretical basis of this paper. AI-driven information cocoons are not the same as general selective exposure. These arise when algorithmic ranking, commercial interests and the behaviour of users interact within a personalised media environment.

3 Formation Mechanism of AI-Driven Information Cocoons in New Media.

The formation of AI-driven information cocoons in new media is not the result of a single factor, but rather the product of the interplay and synergy among technical logic, commercial logic, and user psychology. Together, the three form a triple-driving mechanism for the construction of the "technology-business-psychology" system for information cocoons.

The three-factor model is as follows: Figure 1. The figure is added here to show the path of the formation more clearly and is not merely textually described.

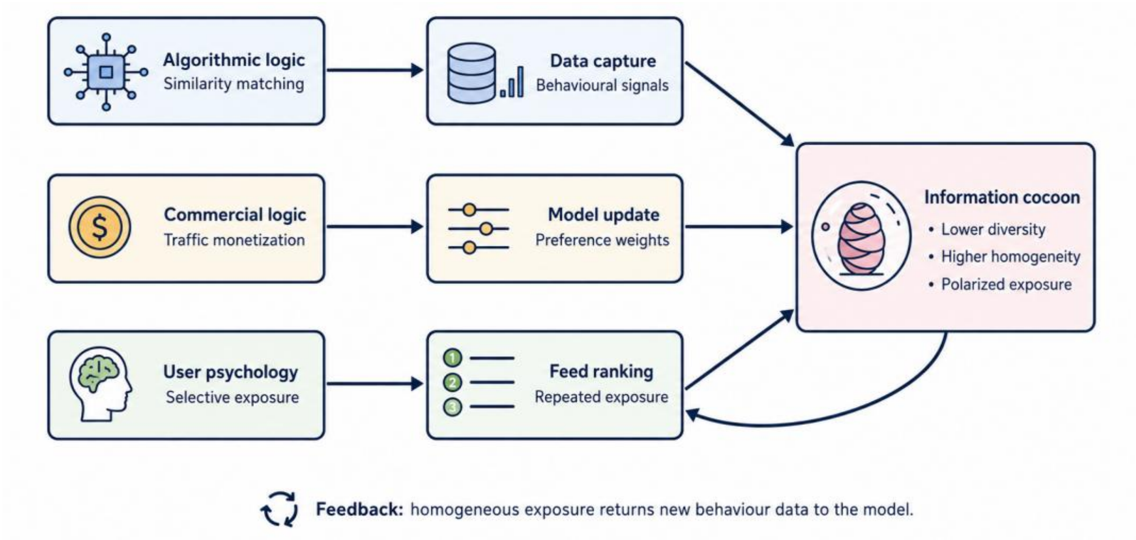

Figure 1: Formation Mechanism of AI-driven Information Cocoons

Figure 1 shows that information cocoons appear in a closed loop: behavioural data are collected, preference weights are updated, homogeneous content is ranked higher, and repeated exposure generates new behavioural signals. As a result, there will be fewer kinds of information and thus less cognitive learning.

3.1 Inherent Deficiencies of Algorithm Recommendation Mechanisms in Technical Logic

The purpose of the core of AI-powered personalised recommendation algorithms is to enhance the accuracy of information matching, but they tend to be uniform in their technical design. The current mainstream recommendation algorithms are collaborative filtering algorithms and content-based recommendation algorithms, and both take "similarity matching" as their basic logic; collaborative filtering algorithms recommend items based on the similarity between users or content, and content-based recommendation algorithms distribute information according to how closely it matches users' past preferences and the characteristics of the content. Therefore, based on the technical logic, these algorithms will tend to promote popular content and fail to show diverse and new materials.

Algorithms are black boxes, and the problem of information cocoons has been further intensified. In order to ensure the commercial competitiveness of the platform, the main parameters and decision-making processes of the platform's algorithms are often not disclosed to the public, forming an "algorithm black box"; thus, regulatory authorities and users are unable to supervise or intervene in the information-screening logic of algorithms, and algorithms continuously strengthen homogeneous recommendations in the pursuit of matching efficiency. An algorithm engineer at an Internet company has admitted that the purpose of algorithms is to increase the user retention rate, and they will not be proactively recommended by the system if users are not interested in the content.

3.2 Profit-Oriented Motivation of Traffic Monetization in Commercial Logic.

The purpose of a new-media platform is to generate income through traffic, so AI algorithms are used to maximise this traffic. Accurately present content that interests users to inspire them and extend their time of use on the platform; this will lead to more advertising and sales opportunities. As profit-oriented enterprises, these platforms tend to strengthen the information-cocoon effect.

On the one hand, the platform can conduct precise marketing by using "user labeling" to classify different groups of people based on their age, gender, interests, etc., and then provide personalised content and advertisements. Label Management restricts users to a particular information circle. For example, an e-commerce platform can recommend related products based on the purchase history of users, and social media can spread relevant information according to a user's interests; thus, some information diversity is sacrificed for the sake of improving the efficiency of commercial conversion.

At the same time, platforms also prefer to distribute entertainment and provide less news. Entertaining content is generally easy to communicate and widely liked by users; thus, it will have a large number of viewers. Therefore, AI algorithms often favour pushing entertaining and superficial content, and serious culture, public affairs, and other valuable but less interesting content are marginalised. Based on the survey data in Wenzhou universities, 92% of students use short-video platforms frequently, but the recommendation ratio for local cultural content is less than 5%; thus, algorithmic preference for entertainment has led to a decline in information diversity.

3.3 Selective Exposure and Preference for Cognitive Comfort Zones in User Psychology

Psychological traits of the users are the internal motivations for forming information cocoons. Cognitive psychology believes that people are naturally inclined to "selective exposure" and will therefore tend to seek out information that is in line with their own beliefs, interests and values, ignoring other kinds of information. This psychological tendency aligns with the drive for the AI algorithm and is thus promoting the formation of information cocoons.

First of all, users' desire for a cognitive comfort zone motivates them to voluntarily withdraw from the world. Different types of information need to be understood and analysed; if all the information were the same, people would not feel stressed. Thus, the user is more likely to select a well-known type of content voluntarily. A mother, Mrs. Li is a professional woman who has long been watching beauty and fashion content, and as a result, her platform's recommendation feed has gradually simplified and formed a cognitive inertia.

Secondly, the desire to belong to a group leads to self-selected exposure behaviour. People need to belong to a group psychologically. By accessing information in line with the group's view, they can feel that they belong to the group. In the era of new media, people have been organising communities based on shared interests and following accounts with similar opinions; at the same time, artificial intelligence algorithms are spreading standardised content within these communities to strengthen group cohesion and are thus more likely to reject different information outside their own.

3.4 Operating Measurement of Information Cocoon Intensity

To increase the verifiability of the analysis, a small set of measurable indicators can be used to describe the formation of information cocoons. Digital media effects are generated by social drivers and algorithmic mechanisms that form a feedback loop, so platform log data and exposure records can be used to evaluate cocoon intensity [11]. The first is the user interest vector shown in Formula (1).

$$U_i = (w_{i1}, w_{i2}, \dots, w_{in}) \tag{1}$$

In Formula (1), U_i denotes the interest vector of user i, and w_{in} denotes the weight assigned to topic or feature n. The matching score between a user and a content item can be represented by cosine similarity, as shown in Formula (2).

$$S(U_i, C_j) = \frac{U_i \cdot C_j}{\|U_i\|\|C_j\|} \tag{2}$$

In Formula (2), C_j is the content vector of item j. If this matching score becomes the dominant ranking objective, familiar content is likely to be recommended repeatedly. Exposure diversity can be measured using Shannon entropy, as shown in Formula (3).

$$D_i = -\sum_{k=1}^{m} p_{ik} \ln(p_{ik}) \tag{3}$$

In Formula (3), p_{ik} is the proportion of content category k in the recommendation list observed by user i. A higher D_i indicates broader topic exposure. Content homogeneity can be measured by the concentration of repeated categories, as shown in Formula (4).

$$H_i = \sum_{k=1}^{m} p_{ik}^2 \tag{4}$$

Formula (4) increases when only a few categories account for the feed. With the demand for fairness and diversity in recommender systems, the cocoon intensity index should incorporate elements of diversity, homogeneity, repetition, and cross-domain exposure [12]. The composite index is as follows: Formula (5).

$$\mathrm{ICI}_i = \alpha\left(1 - \frac{D_i}{D_{\max}}\right) + \beta H_i + \gamma R_i - \delta E_i \tag{5}$$

In Formula (5), ICI_i denotes information cocoon intensity, R_i denotes repeated-content ratio, E_i denotes cross-domain exposure ratio, and α, β, γ and δ are adjustable weights. The governance effect of an intervention can then be evaluated by the relative change in exposure diversity, as shown in Formula (6).

$$\mathrm{GE} = \frac{D_{\text{after}} - D_{\text{before}}}{D_{\text{before}}} \times 100\% \tag{6}$$

Formula (6) is the specific form of governance. If the platform logs of the experimental recommendation feeds or survey-based exposure records are available, then the proposed indicators can be calculated before and after the intervention. Fairness research in recent years has also proposed that recommendation models should be evaluated by indicators other than accuracy and click-through rate [13].

Table 2: Operational Indicators for Assessing Information Cocoon Intensity

Indicator	Measurement	Interpretation
Matching score	Cosine similarity between user and content vectors	High score means strong preference-based ranking
Information diversity	Entropy of exposed content categories	Low entropy indicates narrow exposure
Homogeneity	Sum of squared category proportions	High value indicates repeated categories
Repeated-content ratio	Share of semantically similar items	High ratio means redundancy
Cross-domain exposure	Share of content outside dominant interests	Higher ratio means broader information access

Table 2 transforms the abstract concept of an information cocoon into operational variables. The subjects of this study are news, short videos, cultural communication and education, etc.; however, the basic idea is the same: to expand the scope of dissemination rather than focusing on a single channel.

The Structure of the Feedback Loop is as follows: Figure 2. It locates the place in the recommendation process for adding a diversity constraint.

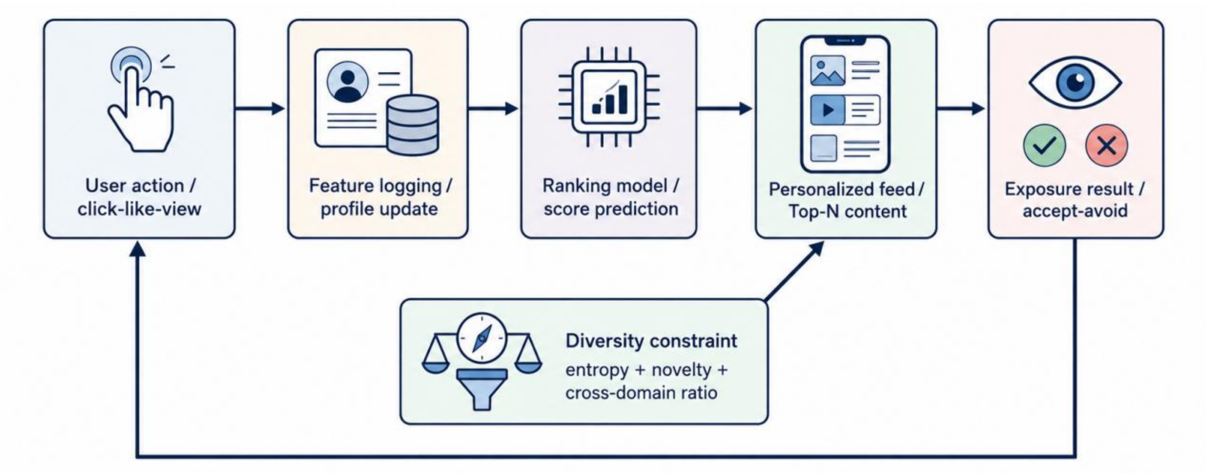

Figure 2: Algorithmic Feedback Loop and Diversity Constraint Module

As shown in Figure 2, the system will not be personalised. A diversity constraint can be added before the final ranking instead. This module will introduce entropy, novelty and cross-domain exposure limits to reduce repetition in the recommendation list.

4 Negative Effects of AI-driven Information Cocoons in New Media

4.1 Cognitive Consolidation and Limited Development at the Individual Level

Information cocoons first have a negative impact on individual cognition by narrowing the scope of thought and restricting thinking. Over time, if a person is only exposed to a single type of information, they will not be exposed to all kinds of information and thus fail to develop a multi-dimensional view and the ability to handle complex problems. For example, users who have been reading extremist views for a long time will gradually form biased thinking and fail to view social phenomena objectively and all-encompassing.

At the same time, information cocoons will reduce the willingness of people to innovate and adapt. Innovation is the result of the combination and adaptation of different knowledge and perspectives, but information cocoons restrict individuals' access to this variety of information, keeping them in a "knowledge closed loop". Furthermore, extended dependence on algorithms for information will reduce people's capacity to seek out and filter information independently; thus, it will be difficult to quickly adapt to, and respond to the complex information needs arising outside this cocoon. According to a survey by Wenzhou University, students who are enclosed in "information cocoons" have a weaker sense of their local culture, and it can be seen that information cocoons inhibit the construction of their knowledge systems.

4.2 Group Polarization and Fragmentation of Social Consensus at the Social Level

AI-driven information cocoons increase group division by using circled recommendations and are thus unable to reach social consensus. The different groups are limited to their own information circles, do not communicate and understand each other effectively, have formed different cognitions on the same public issue, and have even generated hostile emotions. For example, in discussions of public policy on a particular social platform, one frequently sees a bipolarised pattern; that is to say, supporters and opponents merely strengthen their own views within their own circles and thus cannot reach a rational consensus.

Information cocoons may also serve as a place for the propagation of fake news and extremist ideas. Algorithms are homogeneous push mechanisms that spread false information quickly in certain areas, and people living in these areas are likely to believe the false information because they do not receive cross-verification from various sources. At the same time, extreme ideas can be precisely driven by algorithms to reach target groups and continuously reinforce these extreme views, thus posing a risk to social stability. Professor Xiong Wenzhao believes that information cocoons are emerging or growing and hindering social communication and development.

4.3 Loss of Diversity and Discontinuity of Inheritance at the Cultural Level

AI algorithms are more fond of entertaining content, so high-quality content such as serious culture and traditional culture has been neglected and lost. For example, intangible cultural heritage in Wenzhou, such as Ouxiu Embroidery and boxwood carving, has very little exposure in the recommendation feeds of short-video platforms because it lacks entertaining qualities; thus, young people are unable to learn about and recognize them. This will reduce the diversity of culture and may lead to a break in the chain of inheritance for traditional culture.

At the same time, information cocoons also lead to the closure of cultural circles. A shortage of communication among various cultural groups and lack of interaction thus hinder the development of culture. Over time, this will result in a stagnation of cultural development and fail to boost social cultural confidence and soft power.

Recently, several comparative studies on echo-chamber research have used different concepts and operating indicators, and thus have reached different conclusions [14]. Therefore, the scope of impact analysis will exclude individual cognitions, social consensus and cultural dissemination. Audit work on the pathways to radicalisation on YouTube has also shown that recommendation routes can create repeated exposure to increasingly homogenous or extreme content [15].

Table 3: Negative Impacts and Observable Manifestations of AI-driven Information Cocoons

Level	Main impact	Observable manifestation	Potential risk
Individual cognition	Narrowed horizon	Reduced topic categories and repeated viewpoints	Lower judgement quality
Social communication	Group polarization	Weak contact with opposing or neutral views	Fragmented consensus
Cultural circulation	Loss of diversity	Low exposure of serious and local cultural content	Inheritance discontinuity
Information security	Rumour reinforcement	Repeated exposure to unverified claims	Amplified misinformation

As shown in Table 3, the harm of information cocoons is not confined to individual preference. Repeatedly limiting the scope of exposure, recommendation systems will also affect public communication, cultural inheritance and information-security governance.

5 Breaking Strategies for AI-Driven Information Cocoons in New Media

Breaking out of the AI-driven information cocoon in new media is a systematic project that requires the formation of collaborative governance forces among the government, industry and individuals, and the construction of an all-encompassing breaking system at three levels: regulation, technology and cognition.

5.1 Strengthen Regulatory and Supervisory Guidance Systems

Policy will take a top-down regulatory approach to strengthen institutional guarantees for breaking information cocoons by modifying laws and regulations and increasing supervisory work. First, strengthen the legal and regulatory system for algorithmic governance. Clarify the use conditions of platform algorithms, require platforms to increase the openness of algorithms, and regularly disclose the main reasons and settings of algorithm recommendations. At the same time, strengthen the legal provisions and regulations for data security and the protection of personal information, restrict the collection of user data by platforms to what is necessary for their functions, and standardize the operating foundation of algorithms at the source.

Second, reform the regulatory system and strengthen the impact of reforms. Set up a "regulatory sandbox" mechanism to offer enterprises a controlled test-and-learn area for technological iteration under the condition of ensuring compliance with the bottom line, thus promoting the combination of regulation and innovation. Build third-party algorithm evaluation institutions to regularly assess the information diversity of platform algorithms and publish reports; construct a normalised supervision mechanism. In addition, strengthen the collaboration and supervision across departments, unify the regulatory resources of cyberspace administration, radio and television, market supervision, and other relevant departments, form a supervision synergy, and strictly deal with illegal activities such as algorithm abuse and the spread of false information.

Strengthen investment in public cultural resources and optimise the structure of information supply. Promote the balanced distribution of spiritual and cultural resources, encourage mainstream media to introduce AI-based communication innovation, and produce more high-quality and diverse public cultural works. Through policy guidance, support the spread of traditional culture and serious cultural works on new media platforms, increase the exposure of high-quality content, and offset the damage to information diversity caused by commercial algorithms.

5.2 Platforms Optimise Algorithms and Service Models

As the main subject of algorithm application, new media platforms should take the initiative to assume social responsibilities, promote the concept of "technology for good", and break through information cocoons through algorithm optimisation and service innovation. First, optimize algorithm design and introduce diversity recommendation mechanisms. Based on the existing similarity-based recommendations, add algorithm modules such as "random recommendation" and "cross-domain exploration", and incorporate content novelty and diversity into the algorithm evaluation indicators. For example, a platform can put out content that is outside of a user's interests but has a public benefit in the recommended list to broaden the user's knowledge. A pilot experiment on "algorithm transparency" at Wenzhou University found that by optimising algorithms to give more weight to promoting local culture, the variety of information for users was increased significantly, and the cultural cognition of the experimental group rose by 31%.

Second, enhance the transparency of algorithms and protect the right to know and choose for users. Platforms should inform users about the general principles of algorithm-driven recommendations, offer functions to turn off personalised recommendations and fine-tune the options for customisation, and allow users to adjust the diversity of recommended content independently. At the same time, set up a convenient feedback channel for users, provide an easy-to-use feedback entry point such as "reduce recommendations of the same type", and allow users to actively adjust the algorithm's push logic.

Thirdly, build a human-machine collaborative information review and distribution system. Give the role of media workers as "gatekeepers", conduct secondary reviews of content pushed by algorithms, and promptly remove false information and extreme content. In the process of traffic generation, pay more attention to the public value of information, increase the dissemination of content with mainstream values and public affairs information, guide users to be aware of social public issues, and promote information exchange among different groups.

5.3 Users Improve Literacy and Cognition to Break Free from Information Cocoons.

Individuals are the internal subjects for breaking out of the information cocoon, and should proactively improve their information literacy, strengthen their cognition and critical thinking abilities concerning algorithms, and actively venture out of their cognitive comfort zones. Improve Algorithm's cognition ability first. Study algorithm-related knowledge to learn how personalised recommendation algorithms work and what they can and cannot do, thereby reducing over-reliance on them. Rationally view the content promoted by algorithms, know that it is in fact the manifestation of commercial interests of the platform, and not unbiased and comprehensive information delivery.

Secondly, actively expand the channels for obtaining information. Regularly organise the follow lists, proactively follow accounts in various fields and with different perspectives, and increase the opportunities for accessing heterogeneous information. In addition to new media platforms, people can also increase the opportunities for "information serendipity" by reading books, visiting museums, participating in offline community activities, etc., and build a multi-faceted system of information acquisition. At the same time, be able to identify the type of information, have a sense of self-correction that validates whether it is true, and avoid blindly accepting or spreading unknown content.

Finally, take part in public discussions voluntarily and improve one's ability to communicate and understand. Facing a public problem, actively gather opinions from all sectors of society, and based on a comprehensive overview of this issue from all corners, form a fair and objective judgement. Through online and offline exchanges and interactions, break the circle of information, promote the collision and integration of different opinions, and help form a consensus in society.

Diversity of exposure has been put forward as a design principle for recommender systems, and therefore, platforms should build diversity into their technical goals rather than viewing it as an external ethical concept [16]. From the perspective of democratic design, breaking filter bubbles also needs user autonomy, meaningful choices and interface-level support for heterogeneous information access [17].

The cooperative governance model shown in this paper is Figure 3.

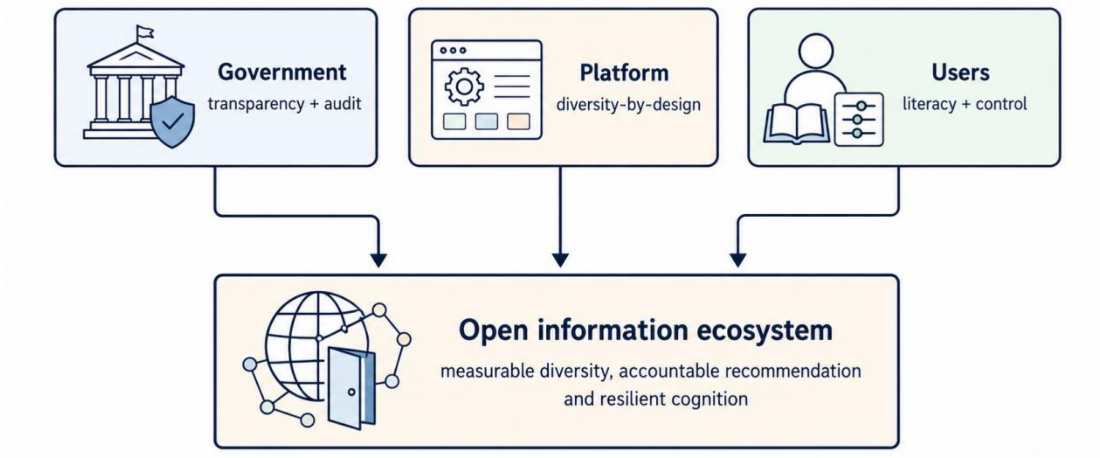

Figure 3: Collaborative Governance Framework for Breaking AI-Driven Information Cocoons

Figure 3 shows that none of the above parties can solve the information-cocoon problem independently. Government regulations introduce requirements for openness and responsibility; Platforms reform recommendation systems and user-controlled interfaces; Users boost media literacy and actively expand their sources of information.

Table 4: Governance Measures and Expected Effects

Governance subject	Main measure	Evaluation indicator	Expected effect
Government	Algorithm disclosure and third-party audit	Transparency report and risk assessment	Higher accountability
Platform	Diversity-by-design ranking	Entropy and cross-domain exposure	Reduced homogeneity
Platform	User control over recommendation settings	Use rate of non-personalized or diversified modes	Greater user autonomy
User	Media literacy and source diversification	Number of active information channels	Stronger critical judgement
Public sector	Supply of high-quality cultural content	Exposure share of public-interest content	Improved cultural diversity

Table 4 links each governance subject to an observable effect. It does not use the style of a slogan and is thus more suitable for future empirical studies.

The Digital Services Act has created a problem of recommender-system transparency in practice. Fabbri holds that the two goals of the transparency requirement for recommender systems are explanations and self-determination [18]. Söderlund and others have also shown that high-reach AI systems need vertical transparency and a system for assessing risks of widespread public attention in the event of a platform failure [19]. Therefore, these studies indicate that legal disclosure, platform audit and user-facing control tools all need to be combined.

At the implementation level, the user control should not be a simple button for turning off personalisation. Reviglio and Fabbri have proposed some transparency and control scenarios for very large online platforms, and interface design can influence how users perceive and modify recommendation systems [20]. Therefore, breaking information cocoons needs to have visible, usable and adjustable recommendation settings rather than just abstract pledges of algorithmic responsibility.

5.4 Summary

AI-driven information cocoons in new media are the result of the joint action of technological progress and commercial profit-seeking. They are characterised by invisibility, a closed-loop nature, and an amplifying effect, thus negatively impacting individual cognition, social consensus, and cultural inheritance. To break out of this dilemma, a collaborative governance system for the government, industry and society needs to be established; the government will provide strong institutional support and regulations; industry will improve the supply of good information through technology; and individuals will raise their own literacy to enhance their cognitive ability.

With the continuous development of AI technology, both the form and the structure of information cocoons will also change; thus, new problems in governance work will emerge. Future research will further explore the mechanism of human-AI interaction and dynamically adjust breaking strategies in combination with the development of new AI technologies, such as generative AI. At the same time, cross-disciplinary studies should also be carried out to integrate theoretical ideas from computer science, sociology, psychology and other areas to provide stronger theoretical support and practical measures for building an open, diverse and healthy new media information ecosystem.

About the Author

Xiangyu Jiang was born in Hunan, China, in 1984. From 2003 to 2007, I studied at Hunan Normal University and graduated from university in 2007. From 2007 to 2011, I was at Guangdong Ocean University. From 2015 to 2018, I was a student at Fujian Normal University and graduated with a master's degree in 2018. I have been working at Yango University since 2011. Five papers have been published, and two of them have been included in EI. The first two are studies on new media art and applications of AI creation.

Yanxin Lin was born in Fuzhou, Fujian, China in 1998. She has graduated from Fujian Normal University for her master's degree. She is now a member of the faculty at the School of Metaverse and New Media, Yango University. She has been doing some research on communication, and some of her specific research topics include the effect of media on society and new media technology.

Xin Hong was born on April 12, 1996, in Fuzhou, Fujian Province, People's Republic of China. China. She has earned a Master's degree from Hanseo University in South Korea and now works at Fuzhou Yango University. The three research areas for the new media content creation, new media marketing and video art are as follows.

Hu Kun is a graduate of Nanjing Forestry University for the master's degree in Broadcasting and Television. He is currently working at the School of Metaverse and New Media in Yango University. The main research direction of him is AIGC Image Communication.

Yijia Tang was born in 1991 and is from Hunan Province, China. From 2009 to 2013, she was a student at China Women's University and graduated with a bachelor's degree in 2013. From 2014 to 2017, she was a student at Capital Normal University and graduated with a master's degree in 2017. From 2020 to 2025, she studied at Cheongju University in South Korea and obtained a doctorate in 2025. She is now working at Yango University. Twelve papers have

been published so far, and the main directions of research are digital media art and the dissemination of cultural content. Email: paf7200@163.com

References

[1] Pariser E. The Filter Bubble: What the Internet Is Hiding from You[M]. New York: Penguin Press, 2011.

[2] Sunstein C R. #Republic: Divided Democracy in the Age of Social Media[M]. Princeton: Princeton University Press, 2017.

[3] Bakshy E, Messing S, Adamic L A. Exposure to ideologically diverse news and opinion on Facebook[J]. Science, 2015, 348(6239): 1130-1132.

[4] Flaxman S, Goel S, Rao J M. Filter bubbles, echo chambers, and online news consumption[J]. Public Opinion Quarterly, 2016, 80(S1): 298-320.

[5] Haim M, Graefe A, Brosius H B. Burst of the filter bubble? Effects of personalization on the diversity of Google News[J]. Digital Journalism, 2018, 6(3): 330-343.

[6] Geschke D, Lorenz J, Holtz P. The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers[J]. British Journal of Social Psychology, 2019, 58(1): 129-149.

[7] Cinelli M, De Francisci Morales G, Galeazzi A, et al. The echo chamber effect on social media[J]. Proceedings of the National Academy of Sciences, 2021, 118(9): e2023301118.

[8] Nguyen T T, Hui P M, Harper F M, et al. Exploring the filter bubble: The effect of using recommender systems on content diversity[C]//Proceedings of the 23rd International Conference on World Wide Web. New York: ACM, 2014: 677-686.

[9] Areeb Q M, Nadeem M, Sohail S S, et al. Filter bubbles in recommender systems: Fact or fallacy-a systematic review[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2023, 13(6): e1512.

[10] Chen J, Dong H, Wang X, et al. Bias and debias in recommender system: A survey[J]. ACM Transactions on Information Systems, 2023, 41(3): 1-39.

[11] Metzler H, Garcia D. Social drivers and algorithmic mechanisms on digital media[J]. Perspectives on Psychological Science, 2024, 19(5): 735-748.

[12] Zhao Y, Wang Y, Liu Y, et al. Fairness and diversity in recommender systems: A survey[J]. ACM Transactions on Intelligent Systems and Technology, 2025, 16(1): 1-28.

[13] Wang Y, Ma W, Zhang M, et al. A survey on the fairness of recommender systems[J]. ACM Transactions on Information Systems, 2023, 41(3): 1-43.

[14] Hartmann D, Wang S M, Pohlmann L, et al. A systematic review of echo chamber research: Comparative analysis of conceptualizations, operationalizations, and varying outcomes[J]. Journal of Computational Social Science, 2025, 8(2): 1-59.

[15] Ribeiro M H, Ottoni R, West R, et al. Auditing radicalization pathways on YouTube[C]//Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. New York: ACM, 2020: 131-141.

[16] Helberger N, Karppinen K, D'Acunto L. Exposure diversity as a design principle for recommender systems[J]. Information, Communication & Society, 2018, 21(2): 191-207.

[17] Bozdag E, van den Hoven J. Breaking the filter bubble: Democracy and design[J]. Ethics and Information Technology, 2015, 17(4): 249-265.

[18] Fabbri M. Self-determination through explanation: An ethical perspective on the implementation of the transparency requirements for recommender systems set by the Digital Services Act of the European Union[C]//Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society. New York: ACM, 2023: 653-661.

[19] Söderlund K, Engström E, Haresamudram K, et al. Regulating high-reach AI: On transparency directions in the Digital Services Act[J]. Internet Policy Review, 2024, 13(1): 1-31.

[20] Reviglio U, Fabbri M. Navigating the Digital Services Act: Scenarios of transparency and user control in VLOPs' recommender systems[C]//CEUR Workshop Proceedings. Aachen: CEUR-WS, 2024, 3898: 1-9.