Enhancing Viewpoint Diversity with AI: A Debate-Driven Smart News Platform

*Irene Minhee Lee, *Juri Lee, *Hyon Hee Kim

*Dept. of Information & Statistics, Dongduk Women's University

Cover Note

This writing sample is derived from my research project "Enhancing Viewpoint Diversity with AI: A Debate-Driven Smart News Platform." I served as the primary researcher responsible for system design, prompt engineering, experimental evaluation, and manuscript writing. The paper has been submitted and is currently under revision for publication in the Journal of the Korean Society of Multimedia.

Abstract

Online news platforms widely adopt personalized news recommender systems to help users efficiently consume relevant content, but long-term personalization can amplify information bias and opinion homogeneity, resulting in filter bubbles and echo chambers. In this paper, we propose Be Kinder, Be Smarter (BKBS), a smart news platform that enhances viewpoint diversity by integrating a generative AI-based debate chatbot with debate-informed news recommendation. To evaluate BKBS, we compare it with a cosine-similarity baseline using a Naver News dataset collected in 2025 across two policy-related topics. Experimental results from an offline comparison show that the proposed debate-based recommendation improves content diversity, as measured by Intra-List Diversity (ILD), and source diversity, as indicated by a lower Community Gini Index (CGI). Overall, these results suggest that BKBS can broaden the range of viewpoints presented to users.

Keywords: Recommender System, Smart News Platform, Generative AI, Chatbot, Filter Bubble

1. Introduction

Online news platforms increasingly rely on personalized recommendation systems to help users navigate vast amounts of information. While such systems improve efficiency by tailoring content to users' interests and reading behaviors, they also risk reinforcing informational bias and ideological homogeneity. Over time, this personalization can lead users to encounter only like-minded views, amplifying social polarization through filter bubble and echo chamber effects. Prior research has shown that algorithmic filtering can negatively affect users' decision-making processes and even the quality of democratic discourse.

Several platforms have attempted to mitigate these issues. For example, AllSides provides news articles spanning left, center, and right political perspectives, and The Guardian's "Burst Your Bubble" feature exposes readers to opposing viewpoints. However, these approaches rely heavily on users' willingness to actively explore diverse content and do not sufficiently engage them in deeper understanding or dialogue. Existing work incorporating large language models (LLMs) into news systems also faces limitations, including hallucinations, high computational cost, and difficulty handling long contextual information.

To address these challenges, this study proposes Be Kinder, Be Smarter (BKBS), a smart news platform that integrates generative AI to enhance viewpoint diversity. Unlike traditional systems that merely juxtapose articles from different perspectives, BKBS uses an interactive debate chatbot to present counterarguments and facilitate multi-perspective discussion. It also generates news recommendations informed by the content of these debates.

The key contributions of this study are as follows:

1. We propose a novel semantic-context-based recommendation framework using GPT to enhance content diversity beyond simple similarity measures.
2. We experimentally demonstrate that GPT-based recommendations produce greater news diversity than traditional approaches.
3. We develop an interactive "news debate chatbot" that engages users in argumentation to help reduce filter bubble effects.
4. We present an integrated AI-driven architecture that connects news recommendation, keyword extraction, summarization, image generation, and debate-based interaction.

The remainder of the paper reviews related work, introduces the BKBS system design and implementation, presents our experimental setup and diversity evaluation, and concludes with contributions and future directions.

2. Related Works

Research on news recommendation systems has evolved from early content-based and collaborative filtering to deep learning–based approaches that capture complex user–item interactions. While these models significantly improved accuracy, they also intensified personalization, raising concerns about reduced informational diversity and the amplification of filter bubble and echo chamber effects. As a result, recent research increasingly emphasizes diversity, fairness, and the social implications of recommendation algorithms.

Traditional news recommendation methods rely heavily on historical behavior and similarity metrics. To address their limitations, emerging studies incorporate large language models (LLMs) into recommendation pipelines through zero-shot prompting or conversational recommendation interfaces. These approaches enhance semantic understanding but often remain static, offering limited user control. Building on this trend, BKBS proposes a prompt-driven, user-influenced recommendation structure that allows individuals to directly specify perspectives and argumentative focus.

Efforts to mitigate filter bubbles have explored balancing accuracy with diversity, fairness, or serendipity through multi-objective optimization or reinforcement learning. These methods broaden exposure but do not fully capture users' real-time cognitive engagement. Recent work on user-controllable systems and multi-agent LLM frameworks expands viewpoint exposure but still focuses mainly on passive presentation. In contrast, BKBS introduces a debate-oriented mechanism that actively engages users in counter-perspective reasoning, moving beyond exposure toward cognitively meaningful interaction.

3. System Architecture

3.1 System Overview

The BKBS platform is designed as an interactive smart news system that integrates generative AI to support article understanding, debate-based viewpoint exploration, and personalized news recommendations. The overall architecture (Fig. 1) consists of five core modules—content understanding, image generation, debate chatbot, debate-based recommendation, and a user-facing interface—connected through a Flask-based backend API that processes all user inputs in conversational form.

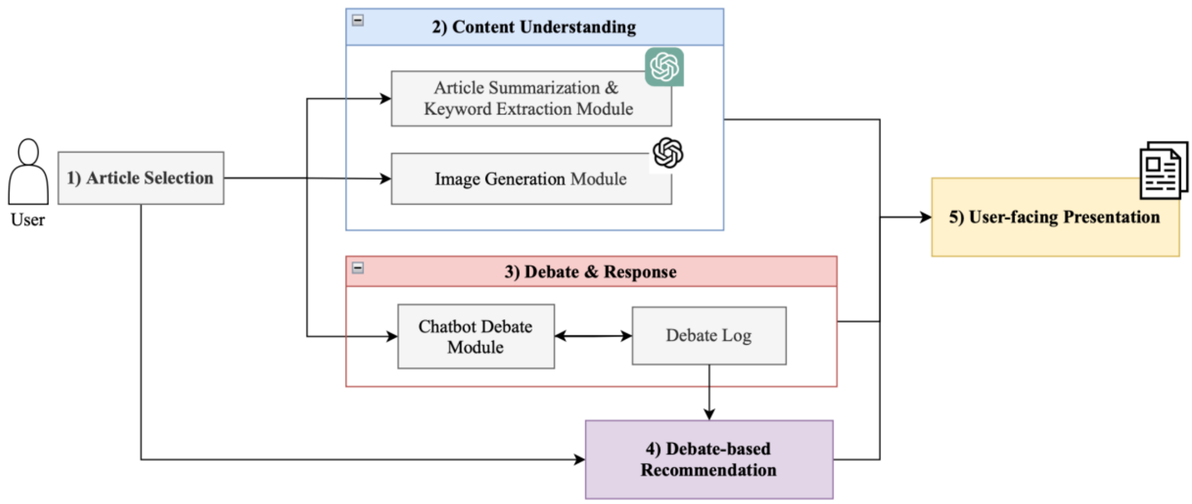

Fig. 1. BKBS System Overview

The system workflow follows five main stages in a recurring interaction loop.

1. Article Selection: The user selects a news article of interest.
2. Content Understanding: The selected article is summarized, key terms are extracted, and a four-panel comic-style illustration is generated.
3. Debate & Response: The user expresses an opinion, and the chatbot generates logical counterarguments based on the article and the ongoing discussion.
4. Debate-based Recommendation: Using the debate log as contextual input, the system recommends ten related news articles that can deepen or broaden the discussion.
5. User-facing Presentation: The final recommendations and interactive debate interface are presented through the web UI, enabling continuous exploration and sequential debates.

3.2 Content Understanding

This stage leverages GPT-3.5-Turbo for zero-shot summarization and keyword extraction, and DALL·E 3 for visual storytelling. Summarization condenses the article into a single sentence, which becomes the basis for generating a four-panel comic-style illustration. Five key technical terms are also extracted and defined to support user comprehension.

We implemented the modules using zero-shot prompting templates for summarization, keyword extraction, image generation, debate generation, and contextual recommendation, designed to ensure semantic consistency and safety. All outputs follow the ethical and safety guidelines outlined in Section 3.6.

3.3 Chatbot Debate Module

The debate chatbot uses GPT-3.5-Turbo with prompts designed to generate counterarguments to the user's statements. The system maintains conversation history and the full article text as contextual input, ensuring coherent dialogue across turns. User messages are processed to produce concise, logically structured rebuttals.

3.4 Debate-based Recommendation Module

GPT-4o-mini is prompted to generate a list of ten relevant articles that align with the ongoing debate. The model receives the article text, debate log, and a pool of candidate articles, and excludes the currently discussed article from the output. Recommendations include a title, article ID, link, and justification. This module enables dynamic viewpoint expansion by connecting argument-driven discussion with curated content suggestions.

3.5 User Interface

Fig. 2 shows the user interface implemented using Flask and HTML templates, which consolidate article details, summaries, extracted keywords, visualizations, chatbot debate, and recommendations into a single coherent layout. Users can explore articles, engage in interactive debates, and request debate-based recommendations directly from the interface. Fig. 3 shows the debate section with recommended articles generated from the conversation context.

Fig. 2. User interface of the BKBS article detail page

Fig. 3. Generative AI debate chatbot interface

3.6 LLM/DALL·E Ethical and Safety Considerations

Given the use of LLMs and generative image models, BKBS incorporates multi-layered policies to mitigate risks related to hallucination, bias, harmful content, and copyright issues. User data is anonymized before being processed through GPT APIs, ensuring compliance with privacy regulations and preventing identifiable information from being used for model training. Summaries and images avoid reproducing exact copyrighted content, and the image generation prompts explicitly prohibit real people, logos, or trademarks. All outputs are meant to serve as assistive interpretations rather than authoritative facts, aligning with responsible AI principles.

4. Experimental Design

This section describes the experimental design used to evaluate whether the proposed BKBS system enhances viewpoint diversity and helps mitigate filter bubble effects during news recommendation. The experiment was conducted as an offline comparison between the baseline content-based recommender and the debate-based BKBS system, as illustrated in Fig. 4.

A dataset of 100 news articles was collected from Naver News on September 29, 2025, covering two socially relevant topics—U.S. immigration policy and the proposed 4.5-day workweek—each represented by 50 articles. Metadata such as titles, full texts, publishers, publication dates, and URLs were included, and duplicates or erroneous entries were removed to ensure reliable diversity evaluation.

Two participants were involved to generate debate contexts for controlled offline comparison. For each topic, they selected an article and received recommendations from both systems under identical dataset conditions. The baseline model generated the top 10 articles using vector similarity between texts, whereas BKBS recommended 10 articles by incorporating the debate context generated from the participant's opinion and the AI chatbot's counterargument. In both systems, the originally selected article was excluded from the recommendation pool.

Recommendation quality was evaluated using an offline evaluation loop. To assess the system's potential to broaden informational exposure, we measured diversity from two perspectives: content diversity using ILD (Intra-List Diversity)@K and source diversity using CGI (Community Gini Index)@K. These metrics quantify whether the recommended results avoid topical or publisher concentration and instead present a wider range of viewpoints. The goal of the experiment was not to generalize user behavior statistically but to analyze how the structural differences between the baseline model and BKBS influence diversity under controlled, reproducible conditions.

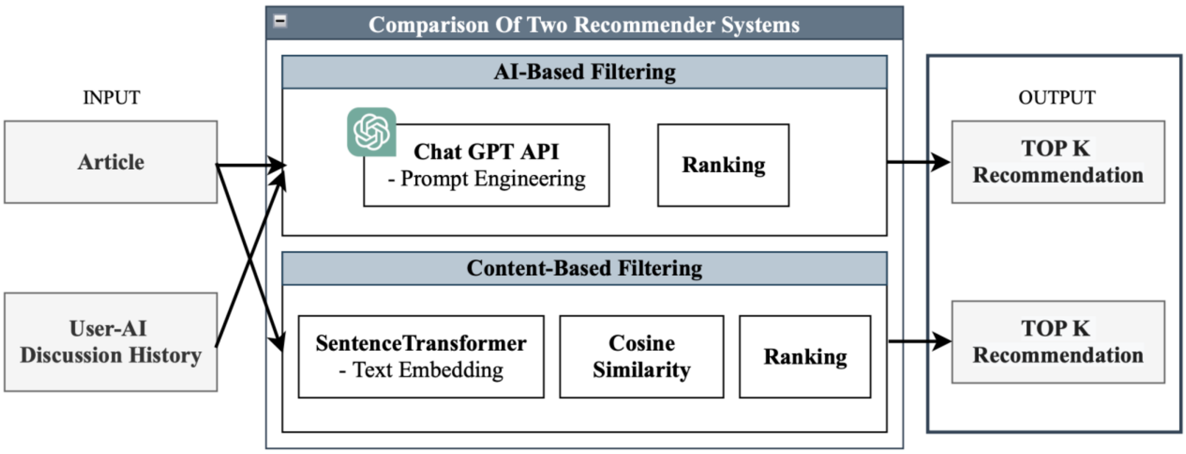

Fig. 4. Overall experimental procedure for comparing the baseline content-based recommender system and the proposed debate-based recommender (BKBS). Both systems generate top-K news recommendations under identical dataset conditions.

5. Discussion

This section evaluates whether the proposed BKBS system enhances recommendation diversity compared to the traditional similarity-based baseline. Diversity was assessed from two perspectives: content diversity, measured using ILD (Intra-List Diversity)@K, and source diversity, measured using CGI (Community Gini Index)@K. These metrics were used to determine whether the recommended articles broadened users' exposure to diverse viewpoints or, conversely, were overly concentrated in specific topics or publishers.

5.1 User Debate Context

To generate debate-based recommendations, two participants interacted with the BKBS system by expressing their opinions on selected news articles. Both participants engaged in debate sessions for two topics—U.S. immigration policy and Korea's proposed 4.5-day workweek (contrasted with Silicon Valley's "996" work culture). These brief debates provided contextual input that guided the BKBS recommendation engine. The subsequent sections (5.2 and 5.3) present the diversity analysis, comparing BKBS with the baseline model.

5.2 Content Diversity Based on ILD

To assess semantic diversity among recommended articles, we computed ILD@K (Intra-List Diversity), which measures the average dissimilarity between all article pairs within the top-K results. The definition of ILD@K is presented below:

$$ILD@K = \frac{2}{K(K-1)} \sum_{i<j} (1 - \mathrm{sim}(i, j))$$

Higher ILD values indicate that the recommended items cover a broader range of topics and perspectives. Text embeddings were generated using Sentence-Transformer models, and cosine similarity served as the basis for pairwise comparisons. Table 6 compares ILD scores for the baseline similarity-based recommender and the BKBS debate-based system across two topics. BKBS consistently produced higher ILD values than the baseline. In the table, P1 refers to Participant 1, and P2 refers to Participant 2.

Topic	Method	ILD
U.S. Immigration Policy	Similarity-based	0.177
	BKBS (AI-based)	P1: 0.276 P2: 0.280
4.5-Day Workweek	Similarity-based	0.113
	BKBS (AI-based)	P1: 0.137 P2: 0.148

Table 6. Intra-List Diversity (ILD) Comparison Between Methods

- U.S. immigration policy: The baseline achieved an ILD of 0.177, while BKBS scored 0.276 (P1) and 0.280 (P2), indicating a clear increase in semantic breadth.

- 4.5-day workweek: Although baseline diversity was lower (0.113), BKBS still showed moderate improvements (0.137 for P1 and 0.148 for P2).

Overall, BKBS broadened content diversity across both topics, though the magnitude of improvement varied depending on the subject domain.

5.3 Source Diversity Based on CGI

Source diversity was evaluated using CGI@K (Community Gini Index), which measures how evenly recommended articles are distributed across different news publishers. The definition of CGI@K is presented below:

$$CGI@K = 1 - \frac{2\sum_{i=1}^{k-1} Si}{kSk} - \frac{1}{k}$$

Lower CGI values represent more balanced exposure to diverse sources, while higher values indicate concentration within a few outlets. Table 7 shows the CGI results for each topic.

Topic	Method	CGI
U.S. Immigration Policy	Similarity-based	0.736
	BKBS (AI-based)	P1: 0.672 P2: 0.672
4.5-Day Workweek	Similarity-based	0.615
	BKBS (AI-based)	P1: 0.654 P2: 0.739

Table 7. Community Gini Index (CGI) Comparison Between Methods

- U.S. immigration policy: The baseline exhibited the highest source concentration (0.736). BKBS reduced this value to 0.672 for both participants, suggesting a meaningful decrease in publisher bias.

- 4.5-day workweek: The baseline already displayed a relatively low concentration (0.615). For this topic, BKBS produced higher CGI values for P1 (0.654) and P2 (0.739), indicating greater reliance on a narrower set of sources. This may reflect how participants' debate inputs aligned more strongly with certain publishers in this specific context.

These findings show that the effect of BKBS on diversity is topic-dependent: while content diversity consistently increased, source diversity improvements varied according to the distributional characteristics of each news domain.

6. Conclusion

This study introduced BKBS (Be Kinder, Be Smarter), a smart news platform designed to mitigate filter bubbles and echo chambers by enhancing diversity in news recommendations. Unlike traditional systems that simply list articles from different perspectives, BKBS integrates a generative debate chatbot and a debate-aware recommendation mechanism. Through interactive counterargument generation, article summarization, keyword extraction, and visual abstraction, the platform supports deeper user understanding while expanding exposure to diverse viewpoints.

To explore its effectiveness, an offline comparison experiment was conducted using a dataset of 100 policy-related articles collected from Naver News in 2025. Results showed that BKBS consistently increased content diversity (ILD) compared to similarity-based recommendations, particularly for the topic of U.S. immigration policy. Source diversity (CGI) also improved for this topic, though the 4.5-day workweek dataset exhibited mixed results, suggesting that debate context and source distribution can interact to create unintended concentration in certain cases.

Overall, BKBS presents a practical approach to promoting viewpoint exploration through generative AI–driven interactive debate and contextual recommendation. By integrating recommendation, comprehension support, debate, and re-recommendation into a unified loop, the system embodies design principles that encourage critical thinking and balanced information consumption—elements often overlooked in accuracy-focused recommender systems. However, the findings are exploratory due to the limited dataset size and small number of participants. Future work will involve larger and more diverse user studies and additional topic domains to validate and extend these results.

References

[1] S. Raza, "Bias Reduction News Recommendation System," *Digital 2024*, Vol. 4, No. 1, pp. 92-103, 2023.

[2] S. Raza and C. Ding, "News Recommender System: A Review of Recent Progress, Challenges, and Opportunities," *arXiv Preprint,* arXiv:2009.04964, 2020.

[3] Eli. Pariser, *The Filter Bubble : What the Internet Is Hiding from You*, Penguin, New York, 2011.

[4] C. Sunstein, *Republic.Com*, Princeton University Press, New Jersey, 2001.

[5] N. Helberger, "On the Democratic Role of News Recommenders," *Digital Journalism*, 7(8), pp. 993-1012, 2019.

[6] E. Bozdag and J. van den Hoven, "Breaking the Filter Bubble: Democracy and Design," *Ethics and Information Technology*, Vol. 17, No. 4, pp. 249-265, 2015.

[7] AllSides | Balanced News via Media Bias Ratings for an Unbiased News Perspective, https://allsides.com (accessed November 19, 2025).

[8] Burst Your Bubble | The Guardian, https://www.theguardian.com/us-news/series/burst-your-bubble (accessed November. 19, 2025).

[9] R. Wang, V. Liesaputra, and Z. Huang, “A Survey on LLM-Based News Recommender Systems,” *arXiv Preprint*, arXiv:2502.09797, 2025.

[10] W. IJntema, F. Goossen, F. Frasincar, and F. Hogenboom, “Ontology-Based News Recommendation,” *Proceedings of the 2010 Extending Database Technology / International Conference on Database Theory Workshops*, Article 16, pp. 1-6, 2010.

[11] M. Kompan and M. Bieliková, “Content-Based News Recommendation,” *E-Commerce and Web Technologies,* Vol. 61, pp. 61-72, 2010.

[12] O. Phelan, K. McCarthy, M. Bennett, and B. Smyth, “Terms of a Feather: Content-Based News Recommendation and Discovery Using Twitter,” *Lecture Notes in Computer Science,* Vol. 6611, pp. 448-459, 2011.

[13] R. Devooght and H. Bersini, “Collaborative Filtering with Recurrent Neural Networks,” *arXiv Preprint*, arXiv:1608.07400, 2016.

[14] A. S. Das, M. Datar, A. Garg, and S. Rajaram, “Google News Personalization: Scalable Online Collaborative Filtering,” *Proceedings of the 16th International World Wide Web Conference*, pp. 271-280, 2007.

[15] B. Marlin and R. S. Zemel, “The Multiple Multiplicative Factor Model for Collaborative Filtering,” *Proceedings of the Twenty-First International Conference on Machine Learning*, pp. 576-583, 2004.

[16] G. De Francisci Morales, A. Gionis, and C. Lucchese, “From Chatter to Headlines: Harnessing the Real-Time Web for Personalized News Recommendation,” *Proceedings of the 5th ACM International Conference on Web Search and Data Mining*, pp. 153-162, 2012.

[17] L. Li, D. Wang, T. Li, D. Knox, and B. Padmanabhan, “SCENE: A Scalable Two-Stage Personalized News Recommendation System,” *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 125-134, 2011.

[18] J. Liu, P. Dolan, and E. R. Pedersen, “Personalized News Recommendation Based on Click Behavior,” *Proceedings of the International Conference on Intelligent User Interfaces,* pp. 31-40, 2010.

[19] X. Wang, L. Yu, K. Ren, G. Taor, W. Zhang, Y. Yu, et al., “Dynamic Attention Deep Model for Article Recommendation by Learning Human Editors’ Demonstration,” *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Vol. Part F129685, pp. 2051-2059, 2017.

[20] G. Zheng, F. Zhang, Z. Zheng, Y. Xiang, N. J. Yuan, X. Xie, et al., “DRN: A Deep Reinforcement Learning Framework for News Recommendation,” *Proceedings of the World Wide Web Conference*, pp. 167–176, 2018.

[21] Z. Zhao, W. Fan, J. Li, Y. Liu, X. Mei, Y. Wang, et al., "Recommender Systems in the Era of Large Language Models (LLMs)," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 36, No. 11, pp. 6889-6907, 2024.

[22] A. Zhang, Y. Chen, L. Sheng, X. Wang, and T. S. Chua, "On Generative Agents in Recommendation," *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Vol. 1, pp. 1807-1817, 2023.

[23] J. Li, W. Zhang, T. Wang, G. Xiong, A. Lu, and G. Medioni, "GPT4Rec: A Generative Framework for Personalized Recommendation and User Interests Interpretation," *Proceedings of eCom'23: ACM SIGIR Workshop on eCommerce*, Vol. 3589, 2023.

[24] Y. Zhang, H. DING, Z. Shui, Y. Ma, J. Zou, A. Deoras, et al., "Language Models as Recommender Systems: Evaluations and Limitations," *Proceedings of the ICBINB Workshop at NeurIPS 2021,* Poster, 2021.

[25] Y. Gao, T. Sheng, Y. Xiang, Y. Xiong, H. Wang, and J. Zhang, "Chat-REC: Towards Interactive and Explainable LLMs-Augmented Recommender System," *arXiv Preprint,* arXiv:2303.14524, 2023.

[26] M. Haim, A. Graefe, and H. B. Brosius, "Burst of the Filter Bubble?: Effects of Personalization on the Diversity of Google News," *Digital Journalism*, Vol. 6, No. 3, pp. 330-343, 2018.

[27] S. Raza and C. Ding, "A Regularized Model to Trade-off between Accuracy and Diversity in a News Recommender System," *Proceedings of 2020 IEEE International Conference on Big Data*, pp. 551-560, 2020.

[28] A. Chakraborty, S. Ghosh, N. Ganguly, and K. P. Gummadi, "Optimizing the Recency-Relevance-Diversity Trade-Offs in Non-Personalized News Recommendations," *Information Retrieval Journal*, Vol. 22, No. 5, pp. 447-475, 2019.

[29] A. Maksai, F. Garcin, and B. Faltings, "Predicting Online Performance of News Recommender Systems through Richer Evaluation Metrics," *Proceedings of the 9th ACM Conference on Recommender Systems*, pp. 179–186, 2015.

[30] P. Banerjee, W. Chen, and L. V. S. Lakshmanan, "Mitigating Filter Bubbles Under a Competitive Diffusion Model," *Proceedings of ACM on Management of Data*, Vol. 1, No. 2, Article 175, pp. 1-26, 2023.

[31] W. Wang, F. Feng, L. Nie, and T.-S. Chua, "User-Controllable Recommendation Against Filter Bubbles," *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1251–1261, 2022.

[32] Y. Zhang, J. Sun, L. Feng, C. Yao, M. Fan, L. Zhang, et al., "See Widely, Think Wisely: Toward Designing a Generative Multi-Agent System to Burst Filter Bubbles," *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, Article 484, pp. 1-24, 2024.

[33] OpenAI, GPT-3.5 Turbo, 2023, https://platform.openai.com/.

[34] OpenAI, DALL·E (version 3), 2024, https://openai.com/index/dall-e-3/.

[35] OpenAI, GPT-4o mini, 2024, https://platform.openai.com/docs/models/gpt-4o-mini.

[36] Y. Du, S. Ranwez, N. Sutton-Charani, and V. Ranwez, "Is Diversity Optimization Always Suitable? Toward a Better Understanding of Diversity within Recommendation Approaches," *Information Processing & Management*, Vol. 58, No. 6, p. 102721, 2021.

[37] M. Tang, X. Huang, and J. Sang, "Mitigating Filter Bubble from the Perspective of Community Detection: A Universal Framework," *arXiv Preprint*, arXiv:2508.11239, 2025.

[38] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks," *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, pp. 3982-3992, 2019.