Exploring Generated Titles and Summaries for Personalized News Filtering and Reading

ZIMO XIA, XICHEN SUN, XUEYANG WU, QIYU PAN, KAIXIANG MO, ZHENHUI PENG*

In this paper, we provide insights to the question "what values can AI-generated content bring to social media users". We present a technical prototype *SummarFlex* that support news filtering and reading with generated query-focused hierarchical summarization. First, we describe the motivation and design of *SummarFlex*. Then, we discuss the benefits and concerns of enabling users to customize the content in social media.

Additional Key Words and Phrases: Text summarization, large language models, news consumption

1 INTRODUCTION

Seeking needed information from news articles on social news sites or websites of news media is a daily activity for many people. A typical news reading process involves two stages. First, readers **filter** interested news articles to reduce the quantity of news to process. Second, they **read** the selected articles in detail to seek needed information. However, the title of an article is typically determined by the publishers and may not always accurately reflect the content that aligns with readers' interests. Researchers in Human-Computer Interaction (HCI) also propose various reading support approaches, such as asking questions [2] and providing summaries [3], to help users efficiently understand the article's ideas. Nevertheless, few websites and HCI research seek to enable readers to actively customize the article's text based on their interests, which would be beneficial for reducing readers' cognitive load [1].

In this paper, we develop an interactive tool named *SummarFlex* that supports users' news filtering and reading with query-focused hierarchical summarization. When users specify queried keywords in the filtering stage (Figure 1), *SummarFlex* sorts the news articles based on BM25P score and prompts GPT-3.5 to generate short summaries conditioned to the keywords as replacement to the articles' titles. When users click a title to proceed to the reading stage (Figure 2), *SummarFlex* supports reading news summaries generated by GPT-4.0 in a hierarchical structure. Specifically, users can first read a GPT-generated high-level summary of the article that focuses on the queried keywords. Then, they can click any sentence in the summary to view more details about it and locate the most related sentence in the original article.

Benefits of AI-generated content for news consumption. First, the generated titles (Figure 1) based on queried keywords could improve readers' efficiency in filtering articles that contain content of their interests. The generated titles could be more objective than the original ones, which could mitigate the negative impact of clickbait titles. Second, the generated summaries (Figure 2) could improve users' efficiency in curating needed information from the articles. This is because the summaries can reduce the information needed to be processed. Third, for news creators, the generative content based on queried keywords could help them draft the title and content that match the interests of the article's target audience.

Concerns of AI-generated content in news consumption. Despite the potential benefits, the generated content could be harmful or misrepresent the original ideas of the news articles. Besides, the generated titles and summaries may reduce the profit of news creators and managers of news platforms who would like to convey certain information and retain their readers. To address these concerns, we should work with news readers, news creators, and managers of news platforms to co-design a reading support tool that balances the benefits of different stakeholders.

*Zimo Xia and Xichen Sun contribute equally in this work. Zimo Xia, Xichen Sun, Qiyu Pan, and Dr. Zhenhui Peng (corresponding author) are from Sun Yat-sen University. Dr. Xueyang Wu is from Zhipu AI in China. Dr. Kaixiang Mo is from Shopee in Singapore.

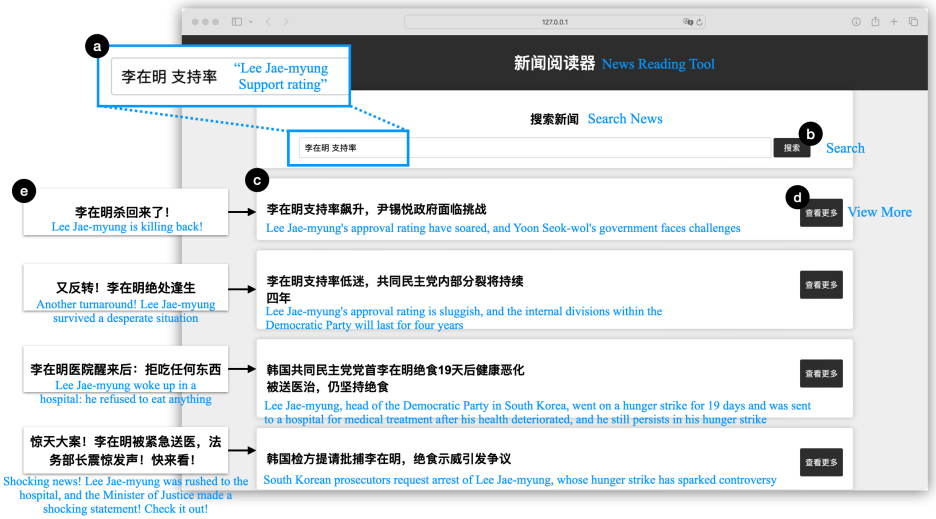

Fig. 1. Screenshot of *SummarFlex*'s news filtering interface. (a) Search box. (b) Search button. (c) Sorted news articles based on the keywords in the search box. The titles of these articles are generated by GPT-3.5 with a focus on the user-specified keywords. (d) View More button to read an article in detail. (e) Original titles of the sorted articles used in the baseline tool.

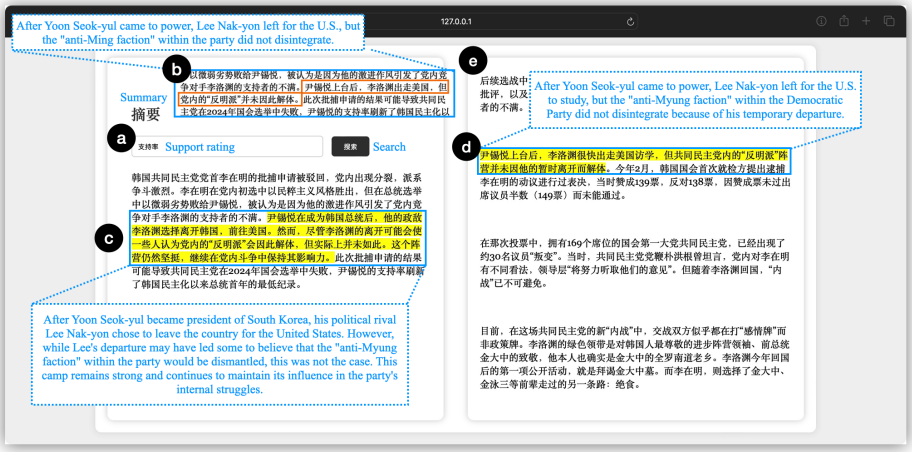

Fig. 2. Screenshot of *SummarFlex*'s news reading interface. (a) Search box. (b) An example of the generated short summary of the news article with a focus on the user-specified keywords. (c) Expanded summary of the generated short one in (b) when users click an interested sentence. (d) Highlighted sentence in the news article that is most relevant to the clicked extended sentence.

2 RESEARCH BACKGROUND AND INTERESTS

The corresponding author Zhenhui Peng is an assistant professor in the School of Artificial Intelligence, Sun Yat-sen University (SYSU). Before that, he was a postdoctoral researcher at Aalto

University, working with Prof. Antti Oulasvirta. He got his Ph.D. degree in Department of Computer Science and Engineering (CSE), Hong Kong University of Science and Technology (HKUST), supervised by Prof. Xiaojuan Ma. He received his Bachelor's degree in ESE from Nanjing University in 2017. His research interests are in Human-Computer Interaction and its intersection with Artificial Intelligence. Specifically, he leverages AI approaches, including generative models, to understand and assist users in online communities (e.g., about mental health, UI design), build interactive tools to address users' needs in various tasks (e.g., critical thinking tasks), and develop personalized tutoring systems (e.g., for vocabulary learning). Zimo Xia, Xichen Sun, and Qiyu Pan are undergraduate students in the School of Artificial Intelligence in SYSU and are interested in building and applying generative models. Xueyang Wu and Kaixiang Mo obtained their Ph.D. degrees in Department of CSE at HKUST and are working on generative models in companies.

REFERENCES

[1] Zhieh Lor, Hae Jung Oh, and Jihyang Choi. 2022. Excluding and Including: News Tailoring Strategies in an Era of News Overload. *Digital Journalism* (2022), 1–19.

[2] Zhenhui Peng, Yuzhi Liu, Hanqi Zhou, Zuyu Xu, and Xiaojuan Ma. 2022. CReBot: Exploring Interactive Question Prompts for Critical Paper Reading. *Int. J. Hum.-Comput. Stud.* 167, C (nov 2022), 17 pages. https://doi.org/10.1016/j.ijhcs.2022.102898

[3] Kangyu Yuan, Hehai Lin, Shilei Cao, Zhenhui Peng, Qingyu Guo, and Xiaojuan Ma. 2023. CriTrainer: An Adaptive Training Tool for Critical Paper Reading. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–17.