SYNC: A Crowdsourcing Platform for News Co-editing

Jin-An Lin
National Yang Ming Chiao Tung
University
Hsinchu, Taiwan
lls0308abc.cs07g@nctu.edu.tw

Feng-Yi Hsu
National Yang Ming Chiao Tung
University
Hsinchu, Taiwan
goeetw123.cs08g@nctu.edu.tw

Hsin-Yu Yao
National Yang Ming Chiao Tung
University
Hsinchu, Taiwan
christyyao.aa06g@nctu.edu.tw

Shang-Hsun Lu
National Yang Ming Chiao Tung
University
Hsinchu, Taiwan
daniellu620.aa06g@nctu.edu.tw

Tsai-Yu Kuo
National Yang Ming Chiao Tung
University
Hsinchu, Taiwan
kuotsaiyu.06g@g2.nctu.edu.tw

Chieh-Kai Lin
National Yang Ming Chiao Tung
University
Hsinchu, Taiwan
axk51013@gapp.nthu.edu.tw

Yung-Ju Chang
National Yang Ming Chiao Tung
University
Hsinchu, Taiwan
armuro@cs.nctu.edu.tw

ABSTRACT

We present SYNC, a crowdsourcing platform that allows news audiences to read news curated, aggregated, organized, and edited by the crowd, and to participate in news aggregation and editing at anytime. Through crowdsourcing, SYNC brings together news information from diverse sources and from contributors with different perspectives, enabling news audiences to obtain a more complete context of specific news events, thereby synchronizing their knowledge of the event. SYNC employs "blocks" and a timeline to help with editors structure and organize news information, and a news material panel to facilitate finding news material to help with aggregation and editing. Our user evaluation showed that participants were positive about the usefulness and societal impacts of SYNC, and found the user interface easy to follow. They also indicated improvements to make to better support news aggregation and editing.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing**.

KEYWORDS

crowdsourcing, news, co-editing

1 INTRODUCTION

Internet users have a vast array of platforms on which they can encounter news, either through active information-seeking behaviors or by serendipity [11]. On these platforms, users' exposure to online news is largely determined by algorithms, many of which leverage users' behavioral history to provide personalized content [2]. However, news-content personalization has been critiqued as potentially creating "filter bubbles" [9] or "echo chambers" [3] that narrow users' exposure to news from diverse perspectives and sources [10, 11], especially to directly opposing viewpoints. Meanwhile, these echo chambers increase the news audience's exposure to biased, homogeneous, and incomplete pictures of news events. A number of solutions to diversifying viewpoints and sources have been proposed, including helping news audiences understand the relationship between different positions via visualization [4]. For example, Vera et al. [6] suggest that providing the valence and the magnitude of the information source position is useful for encouraging exposure to diverse information. StarryThoughts, which positioned the stance of opinions onto a 2-D space, helped news audiences understand and discover the distribution of opinions. [5]. Other researchers also argue that content recommender systems should help users be exposed to more diverse aspects and contents [4, 11].

Another approach is crowdsourcing, bringing people with diverse backgrounds to curate and aggregate news, facts, and stances from different information sources. Unlike the previous approaches, this approach focuses on aggregating news written from different perspectives and providing different aspects of a news event, not biased toward any particular news source from a specific point of view. Wikitribune, which leverages the framework of Wikipedia,

aims to enable crowdsourcing news curation and editing. Nevertheless, it was reported that currently its articles are mainly contributed by recruited editors [8]. In addition, the features of Wikitribune were not designed to particularly support news aggregation or understanding of the temporal context of the news event. Reddit [1] is a social news aggregation and forum platform allowing users to share images, news or text content. A recent study regarding news aggregation on Reddit showed that aggregating news via crowdsourcing is effective [1]. The study inspired us to design a news aggregation system which improve the user experience of reading and editing. It also let us know that news aggregation on the Internet is a feasible solution for us to solve the pain point. More details will be introduced in the next paragraph.

In this paper, we present SYNC, a platform that leverages the power of the crowd to collaboratively curate, aggregate, organize, and edit news information. Its name was inspired by our motivation to synchronize news audiences' knowledge and understanding of news events. This is underpinned by our observation that news media in our country are politically and ideologically fragmented, that misinformation and biased news are also ubiquitous, and that the majority of the news audience is mostly "unsynchronized" with the news information they receive. We consider that we lack a single, accepted site or service on which people can view information from different perspectives and viewpoints. The objective of developing SYNC, therefore, is to curate a more synchronized form of news, marked by clearer contexts and a wider array of viewpoints and sources, via a combination of crowdsourcing. We aim to allow people who possess different levels and types of knowledge and read different news articles about the same event that are written from a wide range of perspectives to be able to collaboratively curate, aggregate, and consolidate such information in a unified, non-segregated manner and distribute it via a range of different platforms, feeds, and posts.

We have developed the core functionality of SYNC to support news material finding, news editing, and news reading. Our user evaluation showed that participants had positive feedback on the societal impacts that SYNC can provide and on the usefulness of its features for supporting them in aggregating and editing news information. Later we will introduce the design process, the core functionality, and implementation of SYNC.

2 SYNC: CROWDSOURCING NEWS CO-EDITING

2.1 Design Process

We employed a user-centered design process, with a primary objective of designing a platform on which 1) the news audience can view news event information from diverse and organized perspectives and sources; and 2) news editors can efficiently identify and use the materials relevant to the present news for editing. With this objective in mind, our design process focused on the features needed by the news audience and editors.

We initially created a paper prototype of SYNC, inspired by the user interface of existing news platforms and co-editing services, respectively. In the first user evaluation, we invited five participants

---

who had no news-editing experience but who paid close attention to public affairs to join our prototype evaluation. In this phase, we sought to understand how the news audience would perceive and accept the concept of news aggregation using crowdsourcing. We walked them through the prototype to explore their perceptions of and attitudes toward the concept and the interaction flow of the system. We also asked them how they expected that they or other people would use the system if it was available online. Through this process, we validated the design concept and learned users' needs for the features that should be included in the system, such as facilitating finding news material, reading about specific topics under a news event, and convenient citation of sources, etc.

Based on the feedback, we built a hi-fi prototype and conducted another formative usability study. We focused on the usability of the user interface of editing, and meanwhile also explored if other needs for news-editing existed that had not been uncovered in the previous study. We recruited five participants who had written news, articles and had aggregated and organized news information to join the formative study. Based on these studies, we have summarized the core features that SYNC should include, which have also been implemented in the system. We introduce the core features of SYNC and implementation below.

There are two modes in SYNC, audience mode and editor mode; readers visit the website in audience mode and can switch to editor mode when they want to provide information or contribute to the news articles.

2.2 Main Component of SYNC

2.2.1 Block-Based Editor. We learned from the study that our participants, when aggregating and organizing news, would create "blocks" of text and links, with a summary text that describes the topic and ideas involved in the block. This is because different paragraphs serve different roles to contextualize an event, such as background, different topics or storylines, and so on. Because a news event typically lasts for several days or longer, during which various topics derived from events may further lead to their own series of follow-up news articles, participants would collect and organize news information according to the specific topics. According to them, separating these blocks allows them to create a clear structure to describe different aspects and storylines about the event, and meanwhile increase its readability and enable the news audience to focus on specific topics instead of needing to read the entire article. To support these practices, we determined to leverage the concept of *block* in the editor mode to help editors structure aggregated news information. Specifically, rather than editing a single article, editors create a news event and create blocks within that event. For each block, editors provide meta-data about the block and then add news material. As shown in Fig. 1a, each block has its own editor interface, in which editors can enter all information, including news title, news event time, news content, hyperlink, and editor information, which are modifiable by anyone on the platform.

The block design can be viewed from both the editor and audience perspectives, but in the audience mode, blocks are not editable. Another advantage of the block-design is that is allows us to place
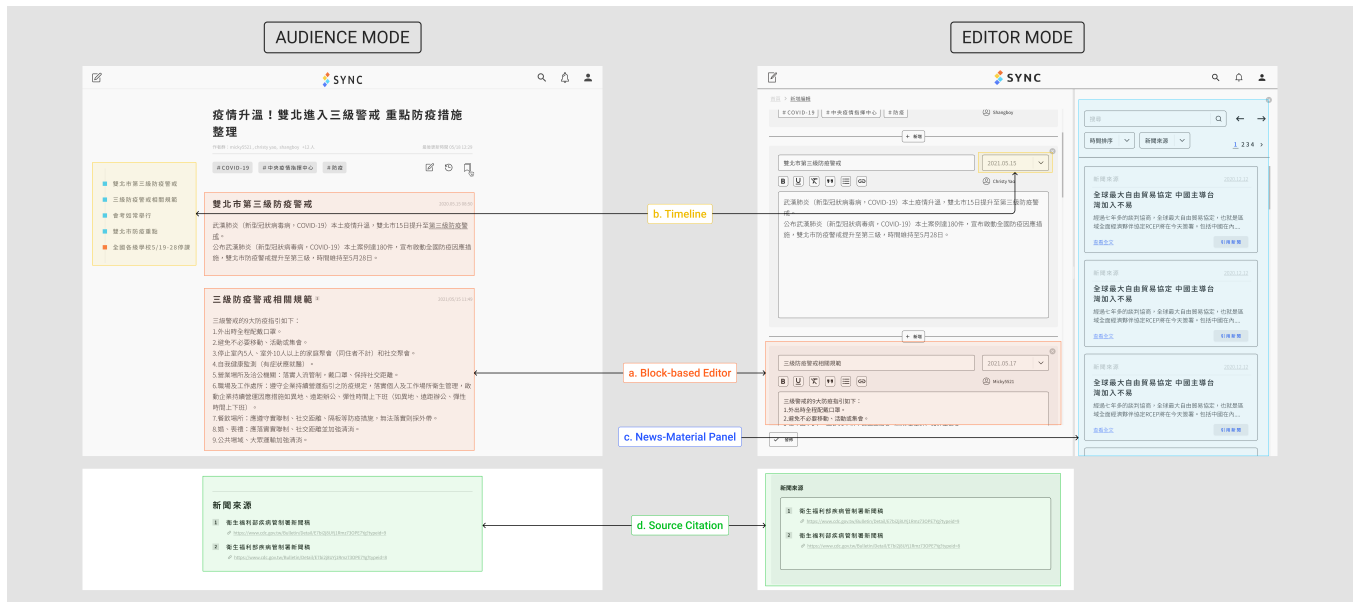
Figure 1: a. Block-Based Editor: Provide the context of news events corresponding to the news storylines. b. Timeline: Present the timestamps of the news event. c. News-Material Panel: Search for news from different platform sources. d. Source Citation: Import and use the news content from the system to automatically generate sources in the News-Material Panel.

anchors into blocks, which serves as a shortcut that enables the audience to directly read specific blocks that contain information only related to specific topics about the event. Below we describe the feature Timeline that leverages these shortcuts to facilitate targeted editing and reading.

2.2.2 Timeline. As shown in Fig. 1b, the Timeline is composed of blocks, which are presented in order of time, according to the timestamp that the editors specify in those blocks. Currently, it is available in the audience mode. The Timeline is designed to give the audience an overview of how the news event is organized in the article. It also reveals the crowd's perceived storyline of the news event. As mentioned earlier, the Timeline also serves as a navigator, on which each item maps to a specific block and creates a shortcut that directly leads the audience to that block. Thus, it helps the news audience quickly find and go to the topics they are interested in to know more about the news event.

2.2.3 News-Material Panel. We learned that participants often switch between material searching, curating and organizing, and choosing information to aggregate [1]. One pain point is the frequent switching between these tasks since they take place in separate application windows. To facilitate these tasks and reduce the switching cost, SYNC includes a News-Material Panel that enables editors to search for news articles in a sidebar (as shown in Fig. 1c). In addition, as we also learned that news aggregation involves many copy-and-paste actions, SYNC includes a text-import feature that allows the editor to directly import the text from the selected news article. When importing and using the text from the material, the system automatically generate a source citation, as shown in Fig 1d. Editors can also manually add citations using the panel.

2.3 Implementation

On the main user interface, the Block-Based Editor is on the left-hand side; it is a WYSIWYG (What You See Is What You Get) editor. The News-Material Panel is on the right-hand side. This side-by-side layout is intended to enable the system's users to easily switch between news editing and managing news material. The News-Material Panel allows users to search for related news by keywords and rank the results by semantic similarities [7]. The high-level architecture of SYNC consists of a web-service interface, a back-end server, a number of news crawlers, a natural language understanding module, and back-end databases. The crawlers, developed in Python, grab the latest news from 15 news websites every 60 minutes. We use MongoDB [2], a NoSQL [3] database, to store news and log data. We use the Express.js [4] web framework and Node.js [5] to develop and implement the back-end API server. For front-end development, we use the Vue.js [6] framework. The SYNC web service is now deployed on Docker [7] in our server.

3 USER EVALUATION

We conducted a user study to explore how the news audience would perceive the features of SYNC for supporting their aggregation and organization of news information. We recruited six participants, half of whom had experience of editing articles on Wikipedia or organizing news information on their personal page on social media. We asked them to aggregate news and browse edited news using

---

[2]MongoDB: https://www.mongodb.com/

[3]NoSQL: https://en.wikipedia.org/wiki/NoSQL

[4]Express.js: https://expressjs.com/

[5]Node.js: https://nodejs.org/en/

[6]Vue.js: https://vuejs.org/

[7]Docker: https://www.docker.com/

SYNC. In the aggregating news part, participants were asked to aggregate news that they did not know about and were allowed to use any other applications to support their task, if necessary. From this, we also observed for what current practices SYNC did not provide enough support.

Overall, participants were positive about the societal impact that SYNC creates. Several participants anticipated that SYNC will facilitate more citizen participation in public affairs. Others mentioned that SYNC could save their time finding relevant news articles to understand a news event because they are all organized in one place, as P6 commented, *"It's good to see a platform to provide aggregations of news; there is so much news media in our country. Our time is precious, it's hard to spend too much time on a news event. I consider that this platform will be a time-saver for citizens who care about current affairs.".*

The News-Material Panel also helped participants accelerate finding news material when editing. However, some participants initially were not clear about the functionality of the News-Material Panel and thought it was for searching within the edited news article. One participant preferred to search for news articles on search engines and attributed it to personal habits. However, this suggests that the News-Material Panel should offer more practical benefits for news aggregators, such as proactively providing recommended news articles based on the content, the ideological position, and the perspective of the edited article, or filtering news based on their quality. Finally, all participants considered that the Timeline was helpful. They mentioned that it helped them quickly clarify the temporal context of the news event. It also provided them with a shortcut to a paragraph about a specific news update; as P2 said, *"I think the timeline can help me when I want to track the latest event in this topic."*

4 FUTURE WORK AND CONCLUSION

We introduce SYNC, a system aimed to provide the news audience with a place to see aggregated news, hoping to help the news audience synchronize news information from diverse sources organized by people with different perspectives. We currently have implemented the core functionality of SYNC. To make the system more useful, we plan to implement features including version comparison and conflict resolution, a personal contribution page, and a more flexible timeline to sort blocks not only based on temporal characteristics but also topics. We also need to establish policies and rules for news material use and for editing. Once a critical mass of people are participating in SYNC, we will evaluate whether it can help the news audience be more aware of the complete context of news events and of different perspectives. Even more compelling to us is the possibility of leveraging the co-edited data on the platform to train generative models for automatically aggregating and editing news text. We envision a human-AI co-editing ecology on SYNC and are looking forward to observing how human collaborate with AI to curate and aggregate news on the platform.

ACKNOWLEDGMENTS

REFERENCES

[1] John J Robinson Alex Leavitt. 2017. Upvote My News: The Practices of Peer Information Aggregation for Breaking News on reddit.com. (2017).
[2] Abhinandan S Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. 2007. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*. 271–280.
[3] Alex France. 2017. Cass R. Sunstein:# Republic: Divided Democracy in the Age of Social Media. *Ethical Theory and Moral Practice* 20, 5 (2017), 1091–1093.
[4] Mingkun Gao, Kyo Jin Do, and Wai-Tat Fu. 2018. Burst Your Bubble! An Intelligent System for Improving Awareness of Diverse Social Opinions. (2018).
[5] Hyunwoo Kim, Haesoo Kim, Kyungje Jo, and Juho Kim. 2021. StarryThoughts: Facilitating Diverse Opinion Exploration on Social Issues. (2021).
[6] Q. Vera Liao and Wai-Tat Fu. 2014. Can You Hear Me Now? Mitigating the Echo Chamber Effect by Source Position Indicators. *Computer Supported Cooperative Work (CSCW '14)* (2014).
[7] Bhaskar Mitra, Nick Craswell, et al. 2018. *An introduction to neural information retrieval.* Now Foundations and Trends.
[8] Sheila O'Riordan, Gaye Kiely, Bill Emerson, and Joseph Feller. 2019. Do you have a source for that? Understanding the Challenges of Collaborative Evidence-based Journalism. *International Symposium on Open Collaboration(OpenSym '19)* (2019).
[9] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you.* Penguin UK.
[10] F Pasquale. 2015. The black box society: Harvard University Press. (2015).
[11] Urbano Reviglio. 2017. Serendipity by Design? How to Turn from Diversity Exposure to Diversity Experience to Face Filter Bubbles in Social Media. (2017).