

# IDEA – Informfully Dataset with Enhanced Attributes

Lucien Heitz<sup>1,2,†</sup>, Nicolas Mattis<sup>3</sup>, Oana Inel<sup>1</sup> and Wouter van Atteveldt<sup>3</sup>

<sup>1</sup>Department of Informatics, University of Zurich, Zurich, Switzerland

<sup>2</sup>Digital Society Initiative, University of Zurich, Zurich, Switzerland

<sup>3</sup>Department of Communication Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

## Abstract

In this paper, we present the Informfully Dataset with Enhanced Attributes (IDEA) for news article recommendations. The dataset consists of an open-source collection of user profiles, news articles with a high topic and outlet diversity, item recommendations, and rich user-item interactions from a field study on behavioral changes in news consumption. The records include both quantitative data from real-time session tracking as well as self-reported data from user surveys on satisfaction with news, knowledge acquisition, and personal background information. This paper outlines the data collection procedure and potential use cases of the dataset for designing normative recommender systems. It provides the documentation of all data collections together with insights into the data quality.

## Keywords

choice architecture, news dataset news recommender design, political learning, topic personalization

## 1. Introduction and Background

User experiments are vital for the evaluation and understanding of the societal impact of content curation in the news domain [1, 2]. However, despite voices in the community demanding more field studies [3, 4], empirical research remains rare [5], especially in the normative domain [6]. A complicating factor in this is that there are only a handful of datasets publicly available to the research community that share results from user experiments with news [7, 8]. Table 1 provides an overview of the most prominent resources available. All datasets provide user-item interactions, mostly in combination with the corresponding news articles. Additionally, some of the existing data collections provide information on users [9, 8, 10] and article recommendations [10, 11]. Unfortunately, many datasets have sparsity issues [12], a lack of meaningful content diversity [11, 13], or omit descriptions of users’ backgrounds [14, 15, 11]. Furthermore, while previous studies found that the image as well as the visualization/display style of a news article are crucial for predicting user engagement [16, 17], such information is not present in any of the existing datasets. To counteract and alleviate these weaknesses, we present the Informfully Dataset with Enhanced Attributes (IDEA).<sup>1</sup>

---

*NORMALize 2024: The Second Workshop on the Normative Design and Evaluation of Recommender Systems, October 18, 2024, co-located with the ACM Conference on Recommender Systems 2024 (RecSys 2024), Bari, Italy*

\*Corresponding author.

✉ heitz@ifi.uzh.ch (L. Heitz); n.m.mattis@vu.nl (N. Mattis); inel@ifi.uzh.ch (O. Inel); wouter.van.attedt@vu.nl (W. v. Atteveldt)

🆔 0000-0001-7987-8446 (L. Heitz); 0000-0001-5110-6612 (N. Mattis); 0000-0003-4691-6586 (O. Inel); 0000-0003-1237-538X (W. v. Atteveldt)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

<sup>1</sup>GitHub repository of the dataset: <https://github.com/Informfully/Datasets>

**Table 1**

Overview of the attributes available in open-source datasets in the news domain: user data on session activity, survey data for background information on users, text and image of news articles, user-article interactions, like/dislike ratings, recommendation lists, and reference screenshots for article presentation.

Dataset	User Data	Survey Data	Article Text	Article Images	Interactions	Article Ratings	Recom. List	Presentation
Adressa [14]	×	×	✓	×	✓	×	×	×
Globo [15]	×	×	×	×	✓	×	×	×
MIND [11]	×	×	✓(URL)	×	✓	×	✓	×
NeMig [9]	✓	✓	✓	×	✓	×	×	×
NPR [8]	✓	×	✓	×	✓	×	×	×
Plista [10]	✓	×	✓	×	✓	×	✓	×
IDEA	✓	✓	✓(URL)	✓(URL)	✓	✓	✓	✓

The Informfully Dataset with Enhanced Attributes includes news articles (text and images via URLs, with high topic and source diversity), interactions (including enhanced attributes, e.g., reading progress, like/dislike ratings for each article, and session data), recommendation lists, and visual references for the item presentation. Further enhancements include timestamps for all datapoints and the entire navigation history for session reconstruction. By providing information on users, items (news articles), user-item interactions, and the item presentation, our dataset enables researchers to build different content-based and collaborative filtering recommender systems (RS). Complementary to the in-app user data (i.e., interaction and page view/session data), we provide self-reported survey measures for all users.

The inclusion of these datapoints is motivated by the *normative thinking* at the heart of the NORMalize Workshop [18]. In doing so, our dataset enables switching from interaction-driven thinking (i.e., an assessment of users based solely on their interaction data, cf. [19]) to normative-driven thinking (i.e., an assessment of users based on survey data to check if the normative goals behind the recommender design have been fulfilled).

Looking not only at the impact of algorithms on user engagement *within* RS, but also the impact on user attitudes and behavior *outside* of RS is critical, as the *relevant values* [18] promoted through *normative* curation strategies might not be visible in engagement alone. After all, the goals of normative RS are not primarily a change in user interactions, but a change towards a given normative goal that results from these interactions (e.g., knowledge acquisition, critical thinking, or becoming aware of important societal issues). In news recommendations, these normative goals can foster an increase of political participation, active deliberation, and provide a voice to underrepresented minorities [2, 20].

By including measures such as users’ political preferences, attitudes towards algorithms, diversity needs, and satisfaction with the news, IDEA enables researchers to explore news engagement in relation to its antecedent effects. Since IDEA’s enhanced set of attributes goes beyond simple user-item interactions, it opens new ways of operationalizing norms and values in the context of RS. It helps to assess the interplay of algorithmic recommendation and individual-level preferences in determining normatively meaningful news engagement and content curation strategies in the news domain.

## 2. Dataset Creation

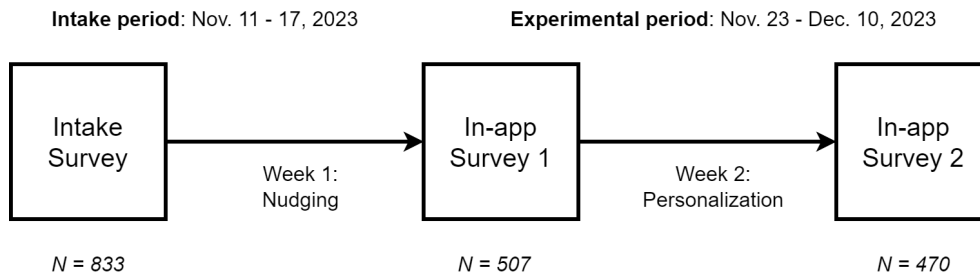
We conducted a pre-registered field experiment in the United Kingdom in November and December of 2023.<sup>2</sup> The dataset was created using the *Informfully Platform* [21, 22, 23], allowing us to track participants in real-time and record all their app interactions over the course of a two-week long field study that required participants to use the app on a daily basis.<sup>3</sup>

### 2.1. Experimental Setting

The experiment for creating IDEA is based on a two-week long user study in which users were exposed to different nudging and personalization conditions: To complete the study, participants needed to 1) express interest and fill in an online Qualtrics recruitment survey, 2) download and install the Informfully app on their phone, 3) use Informfully during two consecutive experimental weeks that differed in their experimental manipulations (see Section 2.3 for the details on the experimental treatment), and 4) complete two in-app surveys.

The recruitment survey measured a number of relevant control variables that can affect user behavior irrespective of the experimental manipulations.<sup>4</sup> The in-app surveys measured recall, subjective knowledge, and user satisfaction—all of which could have been influenced by the experimental manipulations themselves. During the experiment, we asked participants to use the app on a daily basis and to spend at least five minutes reading news each day. To maximize external validity, we asked participants to use the app in the exact same way as they would use a commercial product. Thus, participants were free to choose what to read, when to read, and—as long as the minimum level of engagement was achieved—how long to read.

Within the app, we showed users news from six different news outlets (incl. The BBC, The Guardian, The Independent, i news, Sky News, and the Evening Standard) that were automatically scraped and recommended.<sup>5</sup> We aimed to provide a diverse range of content from news outlets that differ in their target audience as well as the style and focus of news coverage.



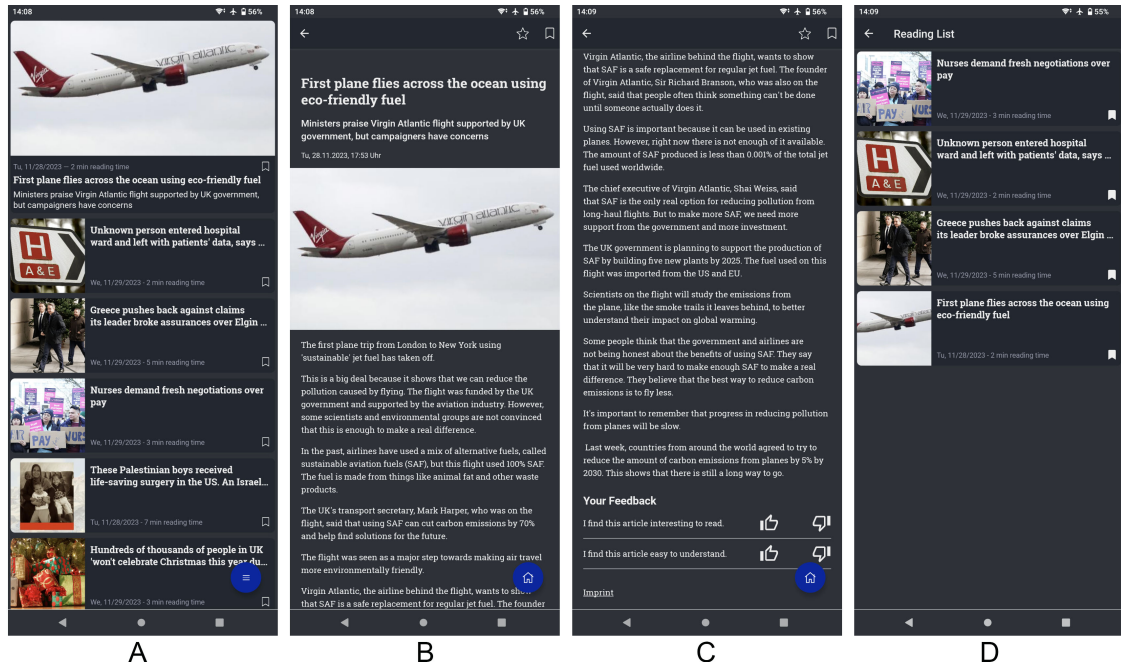
**Figure 1:** Timeline of the experimental procedure and participant involvement.

<sup>2</sup>We used a moving time window for recruitment, as the onboarding was done in multiple waves. While the effective starting date varies between subjects, the overall duration of the experiment remains the same for everyone.

<sup>3</sup>GitHub repository of the platform: <https://github.com/Informfully/Platform>

<sup>4</sup>We chose to measure a broad range of attitudes, preferences, and behaviors that had been either theoretically or empirically linked to news engagement. For example, we measured users' interest in different news topics, political interest and ideology, and attitudes toward algorithmic and journalistic news curation.

<sup>5</sup>GitHub repository of the scrapers: <https://github.com/Informfully/Scrapers>



**Figure 2:** Screenshots of the user interface of the news app. (A) shows the home screen with the 26 articles. (B) is the article view where users read the recommended item. (C) includes the item rating at the bottom of the article. (D) provides an overview of the bookmark collection.

Respondents were shown 26 articles per day, covering 13 different news topics (two articles per topic).<sup>6</sup> During the experiment, users could see the expected reading time for the articles inside the experimental environment, which is, on average, four to five minutes.<sup>7</sup> As such, we predominantly recommended articles that were suitable for brief reading sessions.

The Informfully Platform was used to automatically track and collect interaction data. It is an all-in-one research platform for content distribution. The infrastructure of this platform includes an app for content delivery, giving researchers complete control over when and what items are shown to the experimental participants. Informfully has been evaluated and assessed in previous studies (for details, please see [21, 22]), with its design and item visualization refined by information retrieval insights [24, 25]. By using Informfully, we were able to control the presentation of all news items and make it consistent across platforms (i.e., Android and iOS devices). Figure 2 presents screenshots of how the news feed and items were displayed to participants inside the app. A built-in tutorial enabled people to familiarize themselves with the news apps and its features (e.g., setting bookmarks for creating a reading list or rating articles). The home feed in Figure 2A is automatically updated each day, allowing participants to always have access to the most recent news articles.

<sup>6</sup>The articles presented to participants present only a small subset of the total number of articles that were scraped for the experiment. The dataset also includes articles that were not recommended to participants. These articles can be leveraged to, e.g., calculate outlet-specific text complexity scores and understand the various viewpoints or stances news outlets communicate for topics of interest, among others.

<sup>7</sup>Calculations are based on the average reading speed for native adult speakers of English.

## 2.2. Participants

Participants were recruited online via a third-party marketing agency. They were paid £3 for completing the intake survey and another £15 for finishing the complete study. If they finished the study, they were also eligible for one of two £100 prizes from a draw. Once participants had expressed interest in the study and completed an initial intake survey, they were given login details for Informfully, to then download and use for the remainder of the experiment. In total,  $N = 593$  users filled in the intake survey and used the app at least one time (for details, please see Figure 1). Overall, the dataset includes 199 male, 387 female, and 7 non-binary participants, who were on average 37 ( $M = 37.01$ ;  $SD = 11.73$ ) years old.<sup>8</sup>

Participants were rather highly educated, having on average obtained a college degree, were fairly interested in news ( $M = 5.03$ ;  $SD = 2.17$ ), diversity ( $M = 5.97$ ;  $SD = 0.76$ ), and had high political interest ( $M = 5.19$ ;  $SD = 1.34$ ). Furthermore, their attitude towards both algorithmic ( $M = 3.80$ ;  $SD = 1.30$ ) and journalistic ( $M = 3.55$ ;  $SD = 1.35$ ) preference was similarly close to neutral. On average, respondents were active on 11 days ( $M = 10.61$ ;  $SD = 4.02$ ) and opened 6 unique articles per day ( $M = 5.98$ ;  $SD = 4.30$ ). While user engagement varied considerably between participants at times, we opted to include as many respondents as possible in the dataset. This gives researchers the freedom to decide on their own whom to include and at what cut-off points they want to filter out participants.

## 2.3. User Groups

Table 2 presents an overview of the four different user feeds/groups of the experiments. These feeds differed in terms of item placement and text complexity for one of the articles. Group A received an original environmental news article (Env. OG) in the first position of their feed and the most popular news items from the previous day in position five. Group B received a feed identical to Group A, with the difference that the environmental article in the top position is rewritten (Env. RW) to be more accessible. Group C has the feed of Group A, and Group D has the feed of Group B, with the environmental article and popular article switching places.

During the first week of the experiment, the random articles in positions 2-4 and 6-26 were the exact same across all groups. To ensure topic diversity within the recommendation list, each of these 24 random positions was populated using two articles from each of the twelve available topics.<sup>9</sup> In the second week, we introduced explicit and implicit conditions for topic personalization for a subgroup of users. Table 3 provides an overview of the different conditions. *Condition 1* means participants had the same recommender algorithm as in the first week. *Condition 2 and 3* were exposed to personalization.

Overall, this created  $3 \times 4$  different strategies for constructing news feeds. Implicit preferences are based on the log files, where we calculated which article topics participants spent the most time on during the first week. To determine explicit user preferences, we ask participants after the first week to select their most liked topic in the in-app survey.

<sup>8</sup>Since our sampling included an element of self-selection, our final sample is not representative of the UK.

<sup>9</sup>We manually mapped the topics of each outlet to a unified topic list in order to have a consistent naming convention. The topics present in the unified topic list are: business, crime, entertainment & arts, football, health, life & style, politics, science, sport, technology, UK news, and world news. Environment was an additional topic, but it never appeared in a random position (it was limited to position 1 or 5 of the feed).

**Table 2**

Overview of the different news feeds for each user group in the **first week**. Overall, each feed included 26 news items. The article positions indicates at what place of the feed a given item was placed (position 1 being the top-most article and position 26 being the article on the bottom).

Position	Group A	Group B	Group C	Group D
1	Env. OG	Env. RW	Popular Article	Popular Article
2-4	Random Articles	Random Articles	Random Articles	Random Articles
5	Popular Article	Popular Article	Env. OG	Env. RW
6-8	Random Articles	Random Articles	Random Articles	Random Articles
9-26	Random Articles	Random Articles	Random Articles	Random Articles

**Table 3**

Overview of the different modifications for the news feeds for each user group in the **second week**. Within each group, participants were randomly assigned to one of three conditions. Articles on position 1 and 5 remain fixed to what they received in the first week, with positions 6-26 subject to condition-specific changes for implicit and explicit topic preferences.

Position	Condition 1	Condition 2	Condition 3
1	Fixed	Fixed	Fixed
2-4	Random Articles	Implicit Preference	Explicit Preference
5	Fixed	Fixed	Fixed
6-8	Random Articles	Implicit	Explicit
9-26	Random Articles	9x Implicit, 9x Random	9x Explicit, 9x Random

If topic personalization was present, we populated positions 2-4 and 6-8 according to their preferences.<sup>10</sup> The remaining positions, 9-26, consisted of nine preference-based articles (three articles per topic preference) and nine filler articles (one random article for each non-preferred topic to ensure sufficient topic diversity across the news feed). Random articles were picked from a shared pool, with three articles for each topic.<sup>11</sup>

### 3. Documentation and Analysis

The Informfully Dataset with Enhanced Attributes features 593 users together with 10,954 news articles and a total of 34,890 user-article interactions. Overall, the dataset consists of nine document collections that provide detailed tracking and interaction data across two weeks. All data is exported from the Informfully back end together with user background information from Qualtrics. Section 3.1 provides the description of each of these collections and Section 3.2 offers insight into the dataset quality.

<sup>10</sup>The group allocation outlined here is for users that completed the onboarding and in-app surveys on time. A subset of users, however, delayed activating the app and/or completing the survey. This is reflected in the data records, as their group allocation schedule varies between the first and second week.

<sup>11</sup>More details on the curation process are listed in the online codebook.



### 3.1. Document Collections

The dataset includes the collection of news articles (Articles) retrieved from six different news outlets, reading list and favorites (Bookmarks, Favorites), all user-item interactions (Interactions), article ratings (Ratings), the list of article recommendations (Recommendation), in-app user surveys (Surveys), users and their survey responses (Users), and the session navigation history (Views).<sup>12</sup> The collections contain the following information:

**Articles:** Collection that holds all articles that were retrieved from six different news outlets (i.e., The BBC, The Guardian, The Independent, i news, Sky News, and the Evening Standard) and displayed to users during the study. For each article, the collection holds the title, lead, and accompanying metadata, such as the publication outlet, author, and image URL.<sup>13</sup> (Total size: 10,954 entries.)

**Bookmarks & Favorites:** Holds users' bookmarks in the reading list and their favorites in the archive. (Total size: 2,479 bookmark entries and 3,115 favorite entries.)

**Interactions:** Records of each time a user has selected and opened an article. The collection stores both the item and user ID, together with a timestamp, the reading time, and the maximum scroll percentage of an article. (Total size: 34,890 entries.)

**Ratings:** This collection records each instance where a user agreed (thumbs up) or disagreed (thumbs down) with a statement below an article. During the study, respondents could indicate (dis)agreement with two statements, namely "I find this article interesting" and "I find this article easy to read." (Total size: 28,382 entries.)

**Recommendations:** Contains all article recommendations that were made for any given user over the course of the experiment. Holds the article and user ID together with a timestamp and list position for each recommendation. Includes 26 recommendations for each user per day.<sup>14</sup> (Total size: 207,220 entries.)

**Surveys:** Collection that stores the weekly in-app survey items. Among others, this collection holds the wording of and response to each survey item that respondents were shown in the questionnaires.<sup>15</sup> (Total size: 43,078 entries.)

**Users:** Holds a range of self-reported measures from the Qualtrics intake survey and context variables such as respondents' experimental conditions in the experiment's first and second week. Data on participants' backgrounds consists of replies to questions on: 1) internal political efficacy, 2) political interest and position, 3) news interest and consumption habits, 4) environmental news interests, and 5) attitude towards algorithmic content curation and diversity. (Total size: 593 users, 14 different self-reported measures.)

**Views:** Record of the navigation throughout the entire app. Each page/mask of the app (see Figure 2A-D) has a unique ID (e.g., home screen or article view). This collection tracks and timestamps the transition from one page to another, allowing to reconstruct all user sessions in their entirety. (Total size: 84,747 entries.)

<sup>12</sup>For an in-depth technical documentation, please see the online documentation: <https://informfully.readthedocs.io/en/latest/database.html>; please see the codebook for a non-technical explanation of all attributes and reliability analyses for all survey scales: <https://github.com/Informfully/Datasets/blob/main/IDEA/Codebook.pdf>

<sup>13</sup>For legal reasons, the article text and image are only shared via URL references in the dataset.

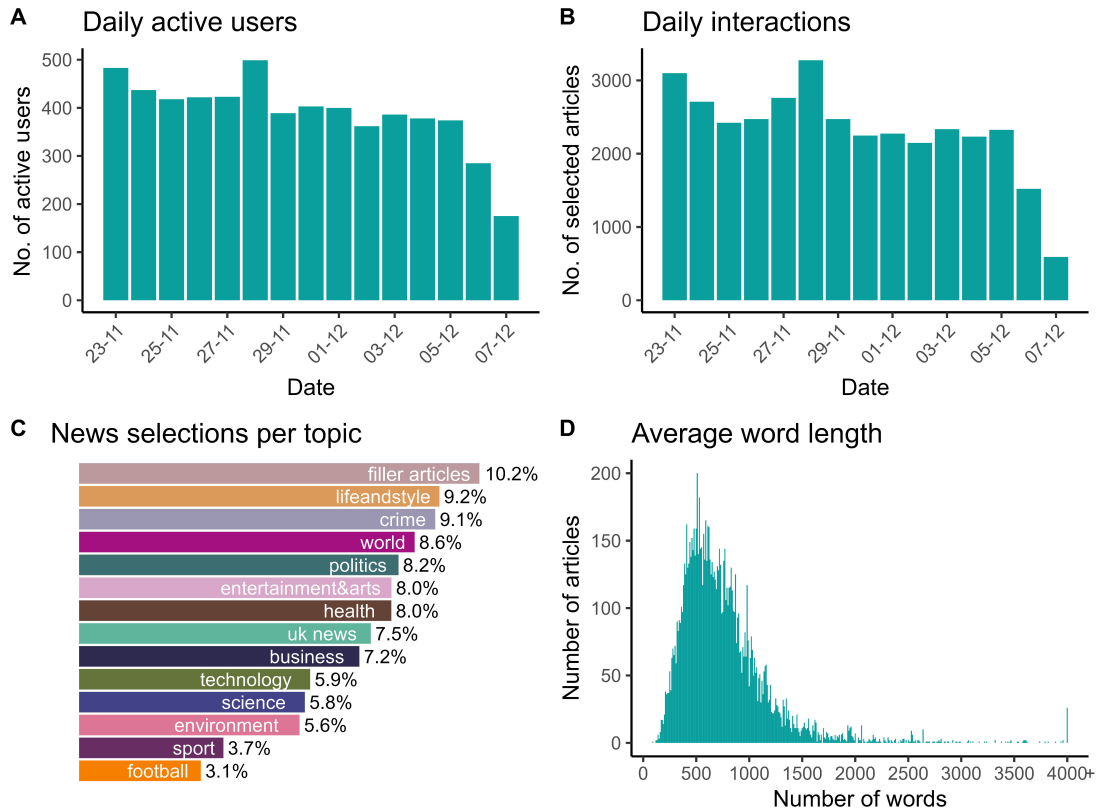
<sup>14</sup>Recommendations were accessible for 24h. They were removed upon inserting the new daily batch.

<sup>15</sup>The experiment included two in-app surveys that were shown to respondents after the first and second week.

### 3.2. Data Analysis

Figure 3 presents an overview of the quantitative data on daily active users (Figure 3A), daily user-item interactions (Figure 3B), distribution of news topics among the opened and read articles (Figure 3C), and article length (Figure 3D).<sup>16</sup> We see almost constant daily active users counts around 400 ( $M = 388.93$ ,  $SD = 77.62$ ) and more than two thousand ( $M = 2326.00$ ,  $SD = 632.02$ ) daily interactions with news articles. When looking at the recommendation lists and the user interactions, the dataset has a sparsity of 83% and an item Gini coefficient of 0.36. The spike in active users and daily interaction on day six of the experiment coincided with us sending out reminders for the upcoming survey at the end of the first week.

We can also see that users' news engagement was quite varied, spanning a broad selection of topics.<sup>17</sup> As such, the data leaves room for analyzing user behavior across different groups, over time, and in combination with changing news supply. Further analyses of the news content itself (e.g., in the form of annotations such as valence and viewpoints) may provide additional insights into the overall selection patterns as well as their relation to self-reported measures.



**Figure 3:** Distributions of daily active users (A), daily interactions (B), engagement with different news topics (C), and average article length (D) .

<sup>16</sup>Please note only articles with more than 500 words have been recommended to users.

<sup>17</sup>"Filler articles" is a catch-all topic for randomly selected articles that were included in the news feed. These fillers were present if there were too few articles to meet the quotas of the curation strategy.



## 4. Discussion and Limitations

The Informfully Dataset with Enhanced Attributes (IDEA) is a first controlled attempt at providing a resource that documents the entire recommendation procedure and can be used to inform the design and development of normative-aware news recommender systems. In the following, we highlight the applicability of the dataset and point out its limitations. But first, we want to reiterate that the sampling procedure for study participants involved an element of self-selection. The final sample was not representative of the UK and it is likely to feature participants who are more interested in news than what is to be expected of the average citizen. IDEA presents the interaction profile of a specific part of the UK population at a specific point in time. Therefore, the generalizability of deriving engagement dynamics for users is limited.

One limitation of IDEA is that participants' news engagement might not always be completely genuine, as they needed to fulfill certain criteria/daily quotas to be eligible for remuneration. Using a realistic news app and providing a broad range of content has likely alleviated this issue. However, compared to other datasets, such as the MIND [11], IDEA ultimately comes from a controlled experiment. It presents only a partial insight into the news consumption habits of individuals and does not track engagement with external news resources that participants might have used parallel to participating in the experiment.

IDEA is composed of more than 10,000 high-quality news articles, covering six sources from the United Kingdom, namely, The BBC, The Guardian, The Independent, i news, Sky News, and the Evening Standard. Due to our controlled study design, participants were exposed to only a fraction of the news articles included in our dataset. Nevertheless, the diverse and rich selection of the outlets allows for designing news recommender systems incorporating several dimensions of diversity, such as source diversity as well as more normative aspects focusing on exposure to diverse topical or political viewpoints.

Finally, leveraging IDEA to develop normative RS might necessitate additional analysis steps. For example, if researchers want to examine normative aspects such as readers' engagement with opposing viewpoints or minority voices, these dimensions must first be extracted and annotated from the body of the news articles. Nonetheless, by combining granular behavioral and self-reported data, IDEA provides ample room to examine how person- and context-specific characteristics co-determine user engagement. These insights could eventually inform and be translated into normative RS designs. While collecting these rich characteristics in a controlled study come at the expense of the dataset size, to the best of our knowledge, no other news dataset (see Table 1) provides such an extensive list of features.

## 5. Conclusion

News engagement is a complex phenomenon and reducing it to clicks alone is reductive [26]. IDEA supports complex user engagement analyses by recording users' news reading behavior together with the associated reading time, scroll percentage, information on articles' likes/dislikes, bookmarks, favorites, and references of how articles were presented. More importantly, as opposed to the majority of existing datasets, these analyses can be correlated with rich user attitudes and perceptions, such as political interest and orientation, diversity values, and

preferences for algorithmic curation and journalism, among others.

With its combination of behavioral and self-reported data, IDEA allows researchers to explore the drivers and dynamics of news engagement across a diverse range of topics and news outlets within an externally valid field experiment. It covers the effects of different recommendation algorithms of four groups across three conditions per group (random vs. based on explicit or implicit user preferences) for a total of twelve experimental conditions. Thus, despite a rather simplistic underlying recommendation logic, we hope the Informfully Dataset with Enhanced Attributes can become a useful resource for researchers across various disciplines, ranging from computer science and information retrieval to communication science and journalism.

## Acknowledgments

This work was partially funded by the Digital Society Initiative (DSI) of the University of Zurich under a grant of the DSI Excellence Program, the Graduate Campus (GRC) of the University of Zurich under a Travel Grant, GRC grant no. 2023\_Q1\_TG\_095), as well as by the Dutch Research Council (NWO), NWO grant no. 406.DI.19.073; project lead: Prof. Wouter van Atteveldt.

## References

- [1] D. Jannach, M. Zanker, M. Ge, M. Gröning, Recommender systems in computer science and information systems—a landscape of research, in: *E-Commerce and Web Technologies: 13th International Conference, EC-Web 2012, Vienna, Austria, September 4-5, 2012. Proceedings 13*, Springer, 2012, pp. 76–87.
- [2] N. Helberger, On the democratic role of news recommenders, in: *Algorithms, Automation, and News*, Routledge, 2021, pp. 14–33.
- [3] A. Bernstein, C. De Vreese, N. Helberger, W. Schulz, K. Zweig, L. Heitz, S. Tolmeijer, et al., Diversity in news recommendation, *Dagstuhl Manifestos 9* (2021) 43–61.
- [4] H. Sargeant, E. Pirkova, M. C. Kettemann, M. Wisniak, M. Scheinin, E. Bevensee, K. Pentney, L. Woods, L. Heitz, B. Kostic, et al., Spotlight on artificial intelligence and freedom of expression: A policy manual, *Organization for Security and Co-operation in Europe* (2022).
- [5] D. Jannach, C. Bauer, Escaping the mcnamara fallacy: Towards more impactful recommender systems research, *Ai Magazine* 41 (2020) 79–95.
- [6] L. Heitz, O. Inel, S. Vrijenhoek, Recommendations for the recommenders: Reflections on prioritizing diversity in the recsys challenge, in: *Proceedings of the Recommender Systems Challenge 2024*, 2024.
- [7] C. Treuillier, S. Castagnos, E. Dufraisse, A. Brun, Being diverse is not enough: Rethinking diversity evaluation to meet challenges of news recommender systems, in: *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, 2022, pp. 222–233.
- [8] J. P. Lucas, J. F. G. da Silva, L. F. de Figueiredo, Npr: a news portal recommendations dataset, in: *Proceedings of the First Workshop on the Normative Design and Evaluation of Recommender Systems*, 2023.

- [9] A. Iana, M. Alam, A. Grote, N. Nikolajevic, K. Ludwig, P. Müller, C. Weinhardt, H. Paulheim, Nemig—a bilingual news collection and knowledge graph about migration, arXiv preprint arXiv:2309.00550 (2023).
- [10] B. Kille, F. Hopfgartner, T. Brodt, T. Heintz, The plista dataset, in: Proceedings of the 2013 international news recommender systems workshop and challenge, 2013, pp. 16–23.
- [11] F. Wu, Y. Qiao, J.-H. Chen, C. Wu, T. Qi, J. Lian, D. Liu, X. Xie, J. Gao, W. Wu, et al., Mind: A large-scale dataset for news recommendation, in: Proceedings of the 58th annual meeting of the association for computational linguistics, 2020, pp. 3597–3606.
- [12] M. Singh, Scalability and sparsity issues in recommender datasets: a survey, Knowledge and Information Systems 62 (2020) 1–43.
- [13] S. Vrijenhoek, Do you mind? reflections on the mind dataset for research on diversity in news recommendations, in: International Workshop on Algorithmic Bias in Search and Recommendation, Springer, 2023, pp. 147–154.
- [14] J. A. Gulla, L. Zhang, P. Liu, Ö. Özgöbek, X. Su, The adressa dataset for news recommendation, in: Proceedings of the international conference on web intelligence, 2017, pp. 1042–1048.
- [15] G. de Souza Pereira Moreira, F. Ferreira, A. M. Da Cunha, News session-based recommendations using deep neural networks, in: Proceedings of the 3rd workshop on deep learning for recommender systems, 2018, pp. 15–23.
- [16] J. Beel, H. Dixon, The ‘unreasonable’ effectiveness of graphical user interfaces for recommender systems, in: Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, 2021, pp. 22–28.
- [17] L. Heitz, Classification of normative recommender systems, in: Proceedings of the First Workshop on the Normative Design and Evaluation of Recommender Systems, 2023.
- [18] S. Vrijenhoek, L. Michiels, J. Kruse, A. Starke, N. Tintarev, J. Viader Guerrero, Normalize: The first workshop on normative design and evaluation of recommender systems, in: Proceedings of the 17th ACM Conference on Recommender Systems, 2023, pp. 1252–1254.
- [19] M. Kaminskas, D. Bridge, Diversity, serendipity, novelty, and coverage: a survey and empirical analysis of beyond-accuracy objectives in recommender systems, ACM Transactions on Interactive Intelligent Systems (TiIS) 7 (2016) 1–42.
- [20] N. Mattis, P. Masur, J. Möller, W. Van Atteveldt, Nudging towards news diversity: A theoretical framework for facilitating diverse news consumption through recommender design, New Media & Society (2022) 26. doi:0.1177/14614448221104413.
- [21] L. Heitz, J. A. Lischka, A. Birrer, B. Paudel, S. Tolmeijer, L. Laugwitz, A. Bernstein, Benefits of diverse news recommendations for democracy: A user study, Digital Journalism 10 (2022) 1710–1730.
- [22] L. Heitz, J. A. Lischka, R. Abdullah, L. Laugwitz, H. Meyer, A. Bernstein, Deliberative diversity for news recommendations: Operationalization and experimental user study, in: Proceedings of the 17th ACM Conference on Recommender Systems, 2023, pp. 813–819.
- [23] L. Heitz, J. A. Croci, M. Sachdeva, A. Bernstein, Informfully - research platform for reproducible user studies, in: Proceedings of the 18th ACM Conference on Recommender Systems, 2024.
- [24] L. Rossetto, M. Baumgartner, R. Gasser, L. Heitz, R. Wang, A. Bernstein, Exploring graph-querying approaches in lifegraph, in: Proceedings of the 4th Annual on Lifelog Search

Challenge, Springer, 2021, pp. 7–10.

- [25] L. Rossetto, M. Baumgartner, N. Ashena, F. Ruosch, R. Pernisch, L. Heitz, A. Bernstein, Videograph—towards using knowledge graphs for interactive video retrieval, in: International Conference on Multimedia Modeling, Springer, 2021, pp. 417–422.
- [26] T. Groot Kormelink, I. Costera Meijer, A user perspective on time spent: Temporal experiences of everyday news use, *Journalism Studies* 21 (2020) 271–286.