Where Do News Recommendations Appear? A Systematic Interface Mapping of 75 European News Websites

Hanne Vandenbroucke[1,*], Annelien Smets[1]

[1]imec-SMIT, Vrije Universiteit Brussel, 1050 Brussels, Belgium

Abstract

News recommender systems (NRSs) are commonly evaluated as isolated rankings or lists of recommended articles, yet in practice they are embedded within website interfaces containing many elements competing for users' attention. This paper reports a descriptive interface mapping of the desktop websites of the three most-used national news outlets in 25 European countries. For each outlet we captured the homepage, one category page, and a standard article page, yielding a corpus of 225 webpages collected between December 2025 and February 2026. Every recommendation module on these pages was annotated for its position, format, number of items, label, and inferred curation logic. Homepages combine recency-based, popularity-based, and editorially structured content; category pages contain comparatively few explicit recommendation modules and remain organized largely around chronological streams; article pages are the most recommendation-dense, combining related-content modules with popularity, recency, editorial, and occasionally personalized selections. We argue that NRSs' performance is shaped not only by which articles are ranked, but also by the interface positions under which recommendations are encountered, including module position, visibility, presentation format, and competition from surrounding content. Through annotated screenshots, this resource provides an empirical basis for constructing more realistic experimental interfaces, and evaluating NRSs in ways that account for exposure rather than ranking quality alone.

Keywords

News Recommender Systems, Website Design, Interface analysis, European news websites, Resource paper

News recommender systems (NRSs) are commonly studied as ranked lists that select and order articles according to the particular algorithmical logic for which they are designed [1, 2]. In real-world news environments, however, recommendations are embedded within webpages containing editorially prioritized stories, chronological streams, popularity lists, related-content modules, advertisements, video blocks, and navigation tools. News pages often contain several recommendation modules at once, each potentially serving different user objectives [2], such as staying up to date, exploring popular stories, or following a topic.

NRS performance should therefore not be understood only in terms of item accuracy or relevance, but also in relation to how the recommendation is presented: where the module is positioned, how it is labeled, which objective it serves, and which other modules compete for attention. Clicks, attention time, and other implicit engagement signals remain valuable indicators of user preference, but may overlook how attention is distributed unevenly across a page [3]. An item recommended at the top of a homepage may receive more exposure than an item below a long article, or one placed last in a horizontal carousel next to a competing recommendation list. The design of the surrounding interface therefore shapes the extent to which users even encounter a recommendation [4, 5]. Furthermore, how users respond to recommendations also depends on whether the system's recommendation purpose aligns with the user's current objective. Together, these factors influence observed user engagement and, consequently, how NRS performance is observed and evaluated.

This short resource paper presents a systematic interface analysis of the 75 most-used national news websites across 25 European countries. For each outlet, we mapped recommendation modules on the homepage, one category page, and one article page, resulting in a corpus of annotated screenshots covering 225 webpages. Each module was coded for its format, page position, item count, label, and

IntRS'26: Joint Workshop on Interfaces and Human Decision Making for Recommender Systems, September 28, 2026, Minneapolis, USA

*Corresponding author.
hanne.vandenbroucke@vub.be (H. Vandenbroucke); annelien.smets@vub.be (A. Smets)
01234-5678-9012 (H. Vandenbroucke); 0003-4771-7159 (A. Smets)

inferred curation logic. The resource is intended to support more realistic experimental interfaces that reflect contemporary multi-list practices. The main contribution is to provide empirical evidence for extending NRS design and evaluation beyond isolated ranked lists toward the multi-module environments in which news recommendations are actually encountered. The mapping documents how different modules may coexist on the same page while serving distinct user objectives and recommendation goals. It also offers a descriptive snapshot of how European publishers currently combine and present recommendation modules, including the emerging integration of generative-AI features. As such, the resource provides a starting point for studying how exposure is distributed across modules and individual items, how modules interact or compete, and how changes to one module may influence engagement with others.

1. Background and Related Work

NRS evaluation often focuses on whether the ranked items are accurate or relevant, using offline metrics or implicit engagement signals such as clicks and attention time as indicators of recommendation quality [6]. This approach has two limitations: first, it overlooks that, in practice, recommendations do not occur in isolation; and second, it does not account for the bias in user interactions resulting from the position of the recommendations.

1. In practice, recommendations are not encountered as isolated ranked lists. Rather, recommendations are distributed across the broader website infrastructure, appearing in lists on homepages, category pages, article pages, newsletters, and push notifications. While one-dimensional recommendation lists long dominated news websites, Loepp [7] points out that contemporary websites increasingly display multiple collections of recommendations with different purposes, such as popularity or recency [7, 2]. Such multi-list interfaces, where several NRS modules coexist on the same page, further complicate the evaluation of NRSs [7, 5, 8, 3, 4], since users choose not only between items within one list, but also between lists and carousels. A news page may combine most-read rankings, editorial selections, and related-story recommendations, yet most NRS studies still evaluate one ranking or one module at a time. Prior work on NRSs has suggested that interface design and positioning matter for users' engagement with recommendations [9, 10]. Therefore, interface-level design decisions, such as visibility, ordering, and placement, strongly shape whether recommendations are noticed, appreciated, and acted upon [11, 7, 12, 13, 4].

2. Exposure and attention are unevenly distributed: position bias and visibily Academic work on NRS evaluation often operationalizes 'position' as the ranking of an item within a list of recommendations. However, in multi-list news interfaces, visibility is also determined at the page level [12, 13, 4, 5]. A recommended article may appear above or below the first scroll, in a sidebar, within or below an article, or in a carousel that requires horizontal scrolling. Visibility is further influenced by the size of the module, its visual weight, the surrounding editorial content, advertisements, and competing recommendation lists [7, 13]. Thus, low interaction with a recommendation module may indicate poor ranking quality, but it may also reflect limited exposure due to weak visual prominence or competition from other elements.

Research on position bias has shown that users are more likely to notice and select prominently displayed items, independently of their relevance [14, 11]. This complicates offline NRS evaluation: clicks and implicit engagement signals used to model user preferences might be skewed as exposure is unevenly distributed across the interface. Models trained on such data may therefore reproduce earlier ranking and visibility effects. Recently, news-specific models such as DebiasGAN [15] and DebiasedRec [16] address this problem by separating inferred user preference from presentation effects when predicting clicks. Implicit engagement signals therefore reflect the combined effects of article relevance, ranking policy, module design, and page-level exposure. Yet there is limited systematic evidence documenting the interface environments in which recommendations on operational news websites are placed.

This short paper addresses that gap through a corpus of annotated screenshots mapping recommendation modules across homepages, category pages, and article pages. Here, we use *"recommendation module"* to refer to an interface element suggesting a collection of articles for further consumption, regardless of whether its contents are selected through algorithmic curation (e.g., personalization, popularity, recency, content similarity) or editorial curation.

2. Methodology

2.1. Data Collection

We conducted an interface analysis of desktop websites of the three most-used national news outlets from 25 European countries. The sample was drawn from the Digital News Report 2025, which reports the online news brands with the highest weekly reach across 25 European countries [17]. For each country we selected the three most-visited national news brands, yielding a sample of 75 websites. The sample includes public service media, privately owned commercial outlets, and digital-native publishers with nationwide reach and an accessible desktop website. Our sample does not include local or specialized news outlets, an overview of the news websites in our sample is included in Appendix **??**. For each of the 75 websites, we documented the homepage, one category page, and a standard article page, resulting in a corpus of 225 webpages captured between December 2025 and February 2026. Our data-collection procedure was informed by the walkthrough method, which involves systematically moving through a digital interface and documenting its features, functions, textual cues, and pathways of use [18]. For our walkthrough, the websites were accessed in desktop version on a laptop using Google Chrome. Cookie requests were accepted to allow the available interface features and dynamically loaded elements to appear. When a website was not available in English, Google Chrome's automatic translation function was used to translate the visible interface text into English. For every website, we first opened the homepage and scrolled through the entire page, taking consecutive screenshots of each viewpoint. Depending on screen size, the viewport and amount of scrolling may vary slightly. We then opened the category page corresponding most closely to abroad, international, foreign, or world news and repeated this procedure. Selecting the same broad category across outlets increases comparability between websites. Finally, we opened a recently published article from the selected international news category and captured the complete article page using screenshots. We excluded article formats whose layouts were structurally exceptional, such as live blogs, interactive specials, or video-only pages, because these were unlikely to provide a comparable representation of a standard article page. The presence of such article page formats was nevertheless recorded in the coding (see Subsection 2.2). This procedure produces a static record of each site at a single point in time. The resulting mapping should therefore be read as a descriptive snapshot of interface elements as they appeared during data collection, not as a stable or permanent characterization of these outlets' designs.

2.2. Screenshot Annotation

The screenshots were examined to identify all positions at which news articles were recommended or suggested as possible next reading choices. Our conceptualization of recommendation modules includes, for example, editorial selections, most-read lists, latest-news lists, related-article modules, in-text recommendations, and visibly personalized modules. Standard navigation menus and search functions were not coded as recommendation modules. When multiple recommendation modules appeared on the same page, each module was treated as a separate unit of analysis. Every module was annotated according to four primary dimensions:

1. **Format or structure:** the visual organization of the module, such as a vertical list, horizontal carousel, grid, card-based block, or in-text link;
2. **Position:** the location of the module on the page and its position relative to other recommendation modules, for example, in the main column, sidebar, within the article body, or below the article;

3. **Number of items:** the number of articles visibly included in the module;
4. **Label or heading:** the title associated with the module, where present, such as *"Latest"*, *"Most read"*, *"Related articles"*, or *"Recommended for you"*.

Alongside these categories, we include an open-coding field for notable features (such as unique article formats and presentations, generative AI tools, etc.) to be recorded for later analysis. The annotation scheme was developed iteratively during an initial examination of the corpus and was refined when recurring formats or ambiguous cases were encountered. All screenshots of the systematic analysis are available in this repository[1].

2.3. Reverse Engineering the Curation Logic

To interpret the recommendation purposes of identified modules, we draw on Bucher's reverse-engineering approach [19, 20], which examines observable outputs to make cautious analytical inferences about their underlying curation logic. Our analysis does not attempt to identify the actual technical implementation of NRSs, which cannot be determined reliably from the interface alone. Instead, we coded the inferred curation logic suggested by module labels, visible ordering, and the relationship between the recommended items and the surrounding page. The coding distinguished between:

- **Recency**, such as modules labeled *"Latest"* or visibly ordered by publication time;
- **Popularity**, such as *"Most read"*, *"Most viewed"*, *"Most commented on"* or *"Most shared"*;
- **Content-based**, such as *"Related articles"* or *"More on this topic"*;
- **Editorial curation**, such as *"Editor's recommender"* or *"Selected by editors"*;
- **Personalized curation**, where the interface explicitly signals suggestions tailored to individual users, for example, through labels such as *"(Recommended) For you"*;
- **Region/Local**, for article selection based on regions the user actively selects or derived from location;
- **Unclear**, when insufficient information is provided to support a more specific interpretation, such as *"Recommended stories"*.

2.4. Data Analysis

The annotated corpus was analyzed using two complementary descriptive approaches. First, for each news outlet, we counted how often each recommendation module format and inferred curation logic occurred across the three page types. We also counted the number of articles displayed within each module type. These descriptive counts were used to describe common patterns and typical ranges, not to generalize statistically to all European news websites.

Second, we reviewed the observations recorded in the open-coding field which capture uncommon, hybrid, or emerging design elements that appeared only in a few cases, but could nevertheless be relevant when studying recommendation modules within news website interfaces.

The mapping, including both coding and analysis, was conducted by one researcher. The purpose of this resource paper is to provide a corpus of screenshots of the recommendation modules across 225 webpages rather than a definitive classification of European news website designs. To support consistency, the researcher applied the same annotation framework across the corpus, revisited ambiguous cases after the coding scheme had been refined, and maintained notes on coding decisions.

[1]Full dataset, including codes, will be shared upon publication of this paper. To ensure author anonymity, we cannot include this in the current manuscript; instead, the anonymous repository can be accessed via this link

3. Results

3.1. Most Common Recommendation Modules

3.1.1. Homepages

Homepages most commonly combined recency-based and popularity-based modules. Recency-based lists usually appeared at the top of the page in two different formats. The first was a static vertical sidebar list, usually containing four to six items and occasionally allowing users to scroll through as many as 15. The second was a narrow horizontal row or hero banner containing three to five static headlines or 10–25 carousel items. Popularity modules were most often displayed as vertical, static lists showing the top 3, 5, or 10 items one below the other, either on the side or within the main content feed. The most common number of items was 5, although some *"Most read"* or *"Most commented on"* lists were internally scrollable. Notably, *"Most viewed videos"* lists often appeared in a horizontal, scrollable carousel with 10 videos. Seven of the 75 homepages contained no explicit recommendation module. In addition, 18 homepages featured modules with labels as *"Recommended stories"*, *"Can't miss"*, or *"Most interesting"*, where it was unclear whether the content would be curated editorially or algorithmically. It might be possible this reflects a combination of both curation logics, where users for whom there is insufficient historical data are shown an editorial selection, while for logged-in, frequent readers, this module is personalized. Furthermore, 8 out of the 75 homepages had a dedicated section with local content, tailored to the location of the user. This can be seen as a personalized module. Our findings show that a fully personalized recommendations on the homepage are rare. Instead, the homepage is compiled using a combination of editorial focus, topic sections and various recommendation lists based on topicality and popularity. Recommendation modules, whether editorial curation or (non-)personalized algorithmic recommendations, compete with one another, as well as with various other established content, for readers' attention.

3.1.2. Category Pages: Low-Density Recommendation Environments

Category pages contain the fewest recommendation modules: 36 out of 75 pages contain none. Where modules were present, popularity lists or recency-based recommendations appeared most often, both in the form of a vertical list in the right-hand column containing 3 to 10 items. For both the *"Most read"* and *"Latest"* lists, it was not always explicitly stated whether these are the most-read or most recent articles within the specific topic category of the page selected (in this case: international news), or overall. We find that the common interface design of the category page, of the websites in our sample, consists of a chronological thematic stream which functions as an ordered selection. As the user has already selected a topic, recommendations might be viewed as limited to only ranking within a limited set of candidates. We propose extending beyond recommendations based on predicted accuracy, supporting users in their search for specific content within the selected theme, and aiming to broaden the user's exploration within that defined category. Although our findings suggest that recommendation modules are currently limited, we argue that category pages can also be useful environments for NRSs and personalization.

3.1.3. Article Pages: Multi-List Environments

Article pages were the most recommendation-dense page type and display the widest variety of recommendation logics and formats. The most common format was a single content-related recommendation embedded within the article text, typically presented as an in-text link or compact preview. Longer articles often included several such standalone in-text recommendations. Many websites also displayed a separate related-content module below the article, usually containing between 3 and 6 items. In some cases, both in-text and below-article related recommendations appeared on the same page. Besides content-related article recommendations, popularity- and recency-based modules remain common. These were most often presented as vertical lists in a sidebar and contain between 3 and 20 items. In addition, many websites placed larger recommendation sections below the article, typically containing between 4 and 20 items and labeled with generic headings such as *"Read more"* or *"Read also"*. Because these labels did not communicate why this specific selection is presented to the user, the underlying curation logic often cannot be inferred from the interface alone. The recommended items in these sections frequently cover a mix of topics and are sometimes interspersed with advertisements. Our findings show that the simultaneous presence of several recommendation modules on one article page was common. Article pages should consequently be understood as multi-list environments that combine in-text related links, recommendations under the article, popularity rankings, recency lists, editorial selections, and advertising.

3.2. Emergence of Generative AI

Beyond conventional recommendation modules, the mapping identified an emerging set of generative AI features that extend recommendations beyond ranked lists suggesting articles for further read, to also provide users with tools to tailor how they consume articles. We most commonly observed bullet-point article summaries and text-to-speech features on article pages. These tools do not themselves recommend which article to read next, but they shape how users access and consume (recommended) content, by lowering reading effort and offering alternative modes of engagement. Furthermore, they point to a further level of personalization beyond item selection: adapting how a articles is presented, as a short summary, an audio version, or a visual explanation, based on each individual user's format preference. This raises questions about how reliably such preferences can be inferred, whether users should retain explicit control over these adaptations, and how well AI-generated representations preserve the meaning, nuance, and editorial integrity of the original reporting. We see this as relevant to this resource paper, because these features increasingly occupy interface space and attention alongside traditional recommendation modules. Generative AI tools are becoming part of the page-level competition for attention, and should be taken into consideration when studying personalization of news websites, which we will highlight in Section 4.1.

In addition, some websites in our sample have implemented conversational agents to assist users in their search for specific content. In these examples, such conversational agent can be considered an extension of the search bar, aimed at supporting users in formulating their information needs, or guiding their search process by proposing follow-up questions. The prompts used to interact with a conversational AI could provide interesting insights into how users formulate their needs in search for specific news content; We do not view this as a separate innovation, since the data from conversational search could then be used to further personalize the news offering. Unlike traditional, static recommendation lists, generative AI is shifting the way articles are recommended: from a passive presentation of items to an interactive form that suggests articles an individual might be interested in, presented in a tailored format, in which the user provides explicit input themselves. This introduces a new level of complexity to NRS design and evaluation. Although such conversational modules are still rare in our dataset, there presence points to an emerging trend and indicates that news website interfaces are personalizing their entire offering by combining recommendation, retrieval, summarization, and explanation within a single interaction loop.

4. Discussion

4.1. Multi-List Recommendation Interfaces and Partial Exposure

Our mapping shows that recommendation modules on news websites rarely appear as standalone lists. This is particularly evident on homepages and article pages, where multiple modules are usually displayed side by side and compete for attention, which is in line with Loepp [7]'s observation that multi-list interfaces have become the default for news. Vertical lists, presented in a sidebar, were the most common format for recency and popularity-based modules, and in most such cases all items in those lists were visible once it entered the view. However, we also observed some vertically scrollable blocks and horizontally scrollable carousels in which only some items were initially visible, meaning lower-ranked items required an additional interaction before they could even be encountered [5, 4].

Such scrollable modules introduce a second level of position bias. For all modules, one must consider whether the module or row as a whole is visible at all. But for scrollable modules, there is a second consideration of position bias: whether a specific item within that module is visible without further interaction. An item may occupy a high position within the underlying recommendation list, yet in practice it is rarely seen, simply because it lies outside the initially visible portion of a carousel. Exposure should therefore be assessed at both the module and item level and NRS evaluation should consider whether a module was visible in the initial view, appeared only after vertical scrolling, or required interaction within the module itself, and, separately, which items within that module were initially visible versus revealed only after carousel navigation. Treating every item in a loaded module as equally likely to get exposure risks misreading non-interaction as low relevance, when in fact the item may never have entered the user's field of view. This is consistent with recent work showing that interface design and presentation substantially affect measured recommender performance [11, 14].

This prevalence of multi-list pages also raises questions for experimental baselines. Evaluating a recommender in isolation, or against an otherwise empty interface "slot", risks overestimating its practical visibility and impact [13, 12]. In operational news environments, a personalized module typically sits alongside established non-personalized alternatives and competes with them for the same attention. Ecologically valid evaluation should therefore embed experimental modules within realistic page environments and compare them against the recommendation and curation elements that already occupy the intended interface position, rather than against an empty baseline.

4.2. Implications for NRS Design and Evaluation

This mapping was motivated by the assumption that page type, module position, visibility, format, item count, and competition from surrounding elements jointly shape exposure before users ever get the chance to evaluate an item's relevance [3, 5, 4]. The same underlying ranking can produce very different outcomes depending on whether it is placed near the top of a homepage, in a sidebar, below a long article, or inside a partially visible carousel [9, 11].

In multi-list interfaces, however, modules may also serve different user objectives, such as being up-to-date with the most recent articles, or reading popular stories. Our mapping shows that such non-personalized curation remains a substantial part of news websites. Recency, popularity, editorial selection, and content-related recommendations were more prevalent across our sample than explicitly personalized modules. This suggests personalized NRSs should be conceptualized not as replacing these established lists, but as coexisting and competing with them for users' attention within a broader multi-list interface. Increasing engagement with one recommendation module may simply redistribute attention away from another, rather than increasing total content consumption or user satisfaction. Similarly, the same article appearing in several modules at once may either reinforce its apparent appeal or produce frustration if a user is repeatedly shown content they have already declined.

Two evaluation implications follow from our findings. First, module-level click-through rate (CTR) should be complemented by page-level measures to account for the distribution of clicks across co-present modules, navigation depth, and the ratio of recommended-to-total articles consumed, rather than treating any one list's CTR as representative of the page [12, 13].

Second, studies proposing new NRS designs should report the intended deployment context: page type, module position, presentation format, number of initially visible items, and which other modules are expected to co-occur on the same page [10, 7]. User studies should likewise be embedded in realistic, multi-module page layouts rather than presenting a single recommendation list in isolation; the homepage, category-page, and article-page layouts mapped here can serve as empirical reference points for constructing such prototypes.

Furthermore, our analysis has implications for how news organizations develop NRSs. Previous research shows that NRS adoption is shaped through negotiations among organizational stakeholders such as editors and technologists, as well as product and business actors, resulting in a set of business rules and editorial interventions [21, 22]. Such NRS-related rules may reserve prominent positions for breaking news or editorially selected stories of public interest, impose restrictions on the recency of articles in recommendations, or prevent repetition. Our results suggest that these interventions must take the entire page into account, since users encounter several differently curated modules on the same page; for each of these modules, news organizations must define what objectives it serves and how all the various engagement objectives, such as editorial priorities like diversity, timeliness, and values of public interest, are balanced with commercial objectives such as subscriptions or retention [23, 22].

Beyond the implications of our analysis, our mapping also opens up further questions for future work. The coexistence of recency, popularity, editorial, content-related, and personalized modules creates opportunities for research on list allocation, ordering, and page-level optimization in multi-list recommenders [13, 12]. Rather than optimizing every module independently, future work could examine how the full page distributes exposure across recommendation purposes and how changes to one list affect interaction with the others. Factorial experiments that independently vary recommendation logic, module position, label, and presentation format could help separate ranking effects from interface effects. Moreover, the emerging use of generative and conversational AI suggests that news recommendation is expanding beyond selecting the next article. These features should therefore be included in interface-level evaluation. Future research could examine whether systems should personalize not only *what* content is recommended, but also *how* it is represented, for example, through concise summaries, audio versions, or visual explanations, while accounting for user control, factual fidelity, accessibility, and preservation of editorial meaning.

4.3. Limitations

Several limitations should be considered when interpreting the findings. The primary objective of this work is to present a corpus of screenshots that provide a descriptive overview of where recommendation modules appear. The coding was performed by a single researcher, which limits the validity of our analysis; nevertheless, one coder is sufficient since the contribution of this resource paper lies in exploratory findings, providing some implications for the design and evaluation of NRSs and offering possible directions for future work. Furthermore, readers should treat the corpus as a systematic descriptive inventory rather than a definitive classification of every publisher's website design. Another limitation of the methodology is that the "inferred curation logic" (Section 2.3) is explicitly an interface-level inference, not a claim about the underlying technical implementation. Labels such as *"Recommended stories"* or *"Read more"* were frequently coded as unclear when the basis for selection could not be inferred. Findings concerning the relative prevalence of "algorithmic" and "editorial" curation should therefore be interpreted cautiously. Additionally, the mapping of category pages was limited to one section (international/world news) to preserve comparability across outlets. Recommendation modules might differ on other category pages such as entertainment or sport. Finally, the corpus is a snapshot collected between December 2025 and February 2026. As news interfaces change frequently, the specific patterns reported here should be understood as characteristic of that window.

5. Conclusion

This paper presents a systematic interface analysis of recommendation modules across 225 webpages from 75 widely used European news websites. The resource corpus specifies where recommendations appear, how they are formatted, how many items they contain, and which curation logics are communicated. The findings show that recommendation modules differ across page types. Homepages foreground recency, popularity, and editorial hierarchy; category pages contain comparatively few explicit recommendation modules; and article pages form dense multi-list environments containing several elements competing for users' attention. Carousels further complicate research with fragmented exposure to recommendations, and emerging generative AI features such as summaries and conversational agents are increasing the opportunities for creating personalized news experiences. These findings demonstrate that recommendation performance cannot be separated from interface placement. Position bias should be considered on two levels: where the module is positioned within a complete webpage, as well as the rank of an item within a list. Therefore, NRS design and evaluation should account for exposure, module competition, cross-list duplication, interaction requirements, and page-appropriate baselines. The accompanying screenshots, annotations, and coding are intended to support more realistic experimental designs, more complete impression logging, and more interface-aware evaluation of news recommender systems.

Declaration on Generative AI

During the preparation of this work, the authors did not use any Generative AI or AI-assisted technologies in the writing, analysis, or preparation of this manuscript

References

[1] S. Raza, C. Ding, News recommender system: a review of recent progress, challenges, and opportunities, Artificial Intelligence Review 55 (2022) 749–800.

[2] D. Jannach, H. Abdollahpouri, A survey on multi-objective recommender systems, Frontiers in big Data 6 (2023) 1157899.

[3] A. D. Starke, E. Asotic, C. Trattner, E. J. Van Loo, Examining the user evaluation of multi-list recommender interfaces in the context of healthy recipe choices, ACM Transactions on Recommender Systems 1 (2023) 1–31.

[4] B. Rahdari, P. Brusilovsky, Under the hood of carousels: Investigating user engagement and navigation effort in multi-list recommender systems, in: Proceedings of the 30th International Conference on Intelligent User Interfaces, 2025, pp. 1485–1498.

[5] B. Loepp, J. Ziegler, How users ride the carousel: exploring the design of multi-list recommender interfaces from a user perspective, in: Proceedings of the 17th ACM Conference on Recommender Systems, 2023, pp. 1090–1095.

[6] B. P. Knijnenburg, M. C. Willemsen, Evaluating recommender systems with user experiments, in: Recommender systems handbook, Springer, 2015, pp. 309–352.

[7] B. Loepp, Multi-list interfaces for recommender systems: survey and future directions, Frontiers in big Data 6 (2023) 1239705.

[8] A. Starke, E. Asotic, C. Trattner, “serving each user”: Supporting different eating goals through a multi-list recommender interface, in: Proceedings of the 15th ACM conference on recommender systems, 2021, pp. 124–132.

[9] N. Mattis, L. Heitz, P. K. Masur, J. Moeller, W. Van Atteveldt, Nudges for news recommenders: prominent article positioning increases selection, engagement, and recall of environmental news, but reducing complexity does not, Journal of Communication (2025) jqaf019.

[10] N. Mattis, T. Groot Kormelink, P. K. Masur, J. Moeller, W. van Atteveldt, Nudging news readers: A mixed-methods approach to understanding when and how interface nudges affect news selection, Digital Journalism 13 (2025) 1647–1667.

[11] 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.

[12] N. Felicioni, M. Ferrari Dacrema, P. Cremonesi, A methodology for the offline evaluation of recommender systems in a user interface with multiple carousels, in: Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, 2021, pp. 10–15.

[13] M. Ferrari Dacrema, N. Felicioni, P. Cremonesi, Offline evaluation of recommender systems in a user interface with multiple carousels, Frontiers in big data 5 (2022) 910030.

[14] A. Collins, D. Tkaczyk, A. Aizawa, J. Beel, Position bias in recommender systems for digital libraries, in: International Conference on Information, Springer, 2018, pp. 335–344.

[15] C. Wu, F. Wu, X. He, Y. Huang, Debiasgan: eliminating position bias in news recommendation with adversarial learning, in: Findings of the Association for Computational Linguistics: EMNLP 2022, 2022, pp. 2933–2938.

[16] J. Yi, F. Wu, C. Wu, Q. Li, G. Sun, X. Xie, Debiasedrec: Bias-aware user modeling and click prediction for personalized news recommendation, arXiv preprint arXiv:2104.07360 (2021).

[17] N. Newman, A. Ross Arguedas, C. T. Robertson, R. K. Nielsen, R. Fletcher, Digital news report 2025, Reuters Institute for the study of Journalism, 2025.

[18] B. Light, J. Burgess, S. Duguay, The walkthrough method: An approach to the study of apps, New media & society 20 (2018) 881–900.

[19] T. Bucher, Neither black nor box: Ways of knowing algorithms, in: Innovative methods in media and communication research, Springer, 2016, pp. 81–98.

[20] T. Bucher, If... then: Algorithmic power and politics, Oxford University Press, 2018.

[21] H. Vandenbroucke, L. Michiels, A. Smets, Welcome to the metrics jungle: Organizational stakeholder perspectives on evaluation of news recommender systems in industry, ACM Transactions on Recommender Systems (2025).

[22] K. van Es, D. Nguyen, Balancing needs and values: A multi-stakeholder examination of algorithmic news recommenders in the netherlands, Journalism Practice (2024) 1–19. doi:10.1080/17512786.2024.2392654.

[23] C. Bauer, C. Bagchi, O. A. Hundogan, K. Van Es, Where are the values? a systematic literature review on news recommender systems, ACM Transactions on Recommender Systems 2 (2024) 1–40.