SIREN: A Simulation Framework for Understanding the Effects of Recommender Systems in Online News Environments

Dimitrios Bountouridis*, Jaron Harambam†, Mykola Makhortykh‡, Mónica Marrero*, Nava Tintarev*, Claudia Hauff*
*Delft University of Technology, The Netherlands
†Institute for Information Law, University of Amsterdam, The Netherlands
‡Amsterdam School of Communication Research, University of Amsterdam, The Netherlands
d.bountouridis@tudelft.nl,{j.harambam,m.makhortykh}@uva.nl,{m.marrero,n.tintarev,c.hauff}@tudelft.nl

ABSTRACT

The growing volume of digital data stimulates the adoption of recommender systems in different socioeconomic domains, including news industries. While news recommenders help consumers deal with information overload and increase their engagement, their use also raises an increasing number of societal concerns, such as "Matthew effects", "filter bubbles", and the overall lack of transparency. We argue that focusing on transparency for content-providers is an under-explored avenue. As such, we designed a simulation framework called `SIREN`[1] (SImulating Recommender Effects in online News environments), that allows content providers to (i) select and parameterize different recommenders and (ii) analyze and visualize their effects with respect to two diversity metrics. Taking the U.S. news media as a case study, we present an analysis on the recommender effects with respect to long-tail novelty and unexpectedness using `SIREN`. Our analysis offers a number of interesting findings, such as the similar potential of certain algorithmically simple (item-based k-Nearest Neighbour) and sophisticated strategies (based on Bayesian Personalized Ranking) to increase diversity over time. Overall, we argue that simulating the effects of recommender systems can help content providers to make more informed decisions when choosing algorithmic recommenders, and as such can help mitigate the aforementioned societal concerns.

CCS CONCEPTS

• **Information systems → Recommender systems**; • **Applied computing → Publishing**; • **Computing methodologies →** *Simulation tools*;

KEYWORDS

recommender systems, diversity, news media, simulation

[1] Open-sourced at github.com/dbountouridis/siren

FAT '9, Atlanta, GA, USA*

DOI: 10.1145/3287560.3287583

1 INTRODUCTION

Recommender systems play a vital role in dealing with the abundance of online content available. In domains as diverse as *e*-commerce, the music and film industry, social media platforms and the news media, recommender systems are deployed to help consumers deal with an information overload by providing filtered and personalized suggestions. At the same time, they help content providers to increase user engagement/satisfaction and boost sales [48].

In line with broader concerns about the societal consequences of digital technologies, such recommender systems are not without criticism. While some scholars herald these as ways to bring niche items to the attention of the wider public [44], others argue that recommender systems mostly benefit the already popular items by recommending those even more [9, 39]. Recommender systems create in this way what sociologist Robert Merton [35] coined half a century ago the "Matthew effects": the rich get richer and the poor get poorer. These concerns about recommender systems reducing diversity have become particularly salient in the public domain where fears of an increasing societal fragmentation, the so-called "echo chambers" [51] and "filter bubbles" [45], are widespread. Moreover, as these recommender systems operate by complex and opaque algorithms, they generally suffer from a lack of transparency [23] and user control [16].

To address these issues, this paper focuses on one particular domain where the use of recommender systems by content providers play an important role, namely the news industry. Since recommenders predominantly deliver information that aligns with people's current interests and preferences, they can drive homogeneity and could lower people's chances to encounter different and not yet discovered contents, opinions and viewpoints [2, 17]. Since the media form an arena for public debate in which a diversity of voices should be heard [15, 40], it would have detrimental consequences to the functioning of our democracies [18, 50].

However, certain works argue that these concerns might have the signs of a moral panic [5], since it is far from clear whether recommender systems have such fragmenting effects. A growing number of studies actually contend that these claims are exaggerated [38, 43, 62], since people actively gather and consume news in many different contexts [54]. Moreover, it can be questioned whether information specialization is necessarily detrimental to democratic debates [17] or that exposure to a diversity of opinions is by definition good for society [12, 58].

The point is that more empirical research is needed to assess the effects of recommender systems. Analyzing their effects is however a complex task considering the resources needed to track user behavior in different algorithmic regimes, not to mention the consequent ethical concerns. In addition, online news consumption is governed by a complex interaction between the users' preferences, the content provider's intent (that translates to editorial priming), and the webpage-nature of the medium.

This paper presents a synthetic alternative that, we argue, is nevertheless able to shed light on these issues: a **simulation framework** which allows for the visualization and analysis of the effects of different recommenders systems. This simulation is based on empirical data, and while a perfect correspondence to reality cannot be guaranteed, such simulations do provide clear insights into the tendencies recommender algorithms exhibit. Our simulation draws mainly on the work of Fleder and Hosanagar who modeled consumer behavior in an e-commerce context [8, 9]. To account for the specificities of news consumption, we go beyond an exclusive focus on recommendation and include both users preferences and editorial priming as they interact in a news-webpage context.

We named our framework SIREN (*SImulating Recommender Effects in online News environments*). It is open-source and enables content providers to insert their own specifications and to test different recommender systems deployed in different contexts. The adjustable parameters of SIREN pertain to the items (articles), users (readers), as well as the recommendation algorithms themselves, as seen in Figure 1. SIREN currently allows for the evaluation with respect to long-tail diversity and unexpectedness diversity metrics to address the "Matthew" and filter-bubble effects respectively. By raising awareness of the consequences of deploying different algorithms, a concrete form of making algorithms more transparent, SIREN allows content providers to test the effects of the recommendation algorithms they have in mind with few resources.

In this work, we not only describe the design and implementation of SIREN (*key contribution I*), but also consider a specific case study—the U.S. news media—and evaluate the effect of different recommender algorithms on long-tail diversity and unexpectedness metrics (*key contribution II*).

2 BACKGROUND

Algorithmic transparency is an important goal in today's technologically saturated world, especially as the workings of algorithm systems are difficult to grasp for non-experts [13, 23]. Furthermore, empirical analyses of algorithms typically require considerable resources (time, data, effort and computational power), and do not allow for the hypothetical testing of alternatives.

Simulations can be used as a means to analyze the consequences of different recommender algorithms. If based on solid empirical data and adequately parameterized, simulations can inform interested stakeholders about the effects of whatever algorithm they are interested in testing. The outcomes of such simulations help those who commission the implementation of recommender systems to be better informed when deciding what algorithms to deploy for the normative and/or commercial purposes they have in mind. In the following sections we make a case for using a simulation, and enumerate the requirements for this simulation.

2.1 Making the case for simulations

Simulations have been predominately used for evaluating and analyzing algorithms with respect to different metrics (e.g., accuracy or diversity) and application-scenarios, such as *e*-commerce, *e*-learning, personalized news and others.

In an *e*-commerce context, Fleder and Hosanagar [8, 9] propose a simulation framework to evaluate recommenders in terms of sales diversity. They propose a mathematical model of user behavior that simulates user awareness, preferences and choice. A similar work by Hinz and Eckert [19] focuses on the evaluation in terms of video-on-demand consumption. Their model is based on economics and marketing studies, but a number of parameters are calibrated based on real-life sales data. Umeda et al. [55] focus on evaluating collaborative filtering approaches. In contrast to other works, their model incorporates the interaction among users.

In an *e*-learning context, Nadolski et al. [41] evaluate recommenders in a scenario where learners need to be advised about the next learning activity to follow. Manouselis et al. [34] identify the best collaborative filtering strategy for an online teachers community. Their work is based on a Web-based application [33] that allows the creation of synthetic datasets, by arbitrarily defining properties (e.g., the number of users). Sie et al. [49] use simulations to investigate coalition recommendations i.e., advising users to choose the right people to collaborate within a network.

In an online news context, Möller et al. [38] evaluate different recommender algorithms by focusing on the algorithms' diversity with respect to content-based features (e.g. topic or tone). In contrast to previous simulation works, the authors generate recommendations based on a static snapshot of user data but do not specify a mathematical model of user behavior.

At a more theoretical level, Lamprecht et al. [27] propose new evaluation measures for the concepts of discoverability (helping users to reach items) and navigability (helping users to explore a collection). Navigability in particular, is measured by simulating the user behavior, modeled through information seeking models where users move from item to item using links, such as "related" movies in a movie web-page. In order to investigate the relationship among diversity metrics, Vargas [56] proposes a probabilistic model of user behavior. The model considers diversity in the temporal domain and incorporates a user browsing model i.e., the probability of a recommended item being selected relates to its ranking position.

Despite the large amount of simulation works , we argue that none of the listed simulation models can accommodate for the particularities of news consumption in an online news environment. The next section elaborates on the specificities of online news.

2.2 Requirements for the news context

In a news context, the simulation's conceptual model should capture the general mechanics of article publishing and consumption. At the same time, the model's parameterization should accurately capture the specific intent of both users and content providers, e.g., what users want to read and what content providers want users to read.

Online news articles have a very distinctive nature that separates them from other Web objects such as movies and music: Li et al. [30] argue that typical recommendation strategies need to be adapted

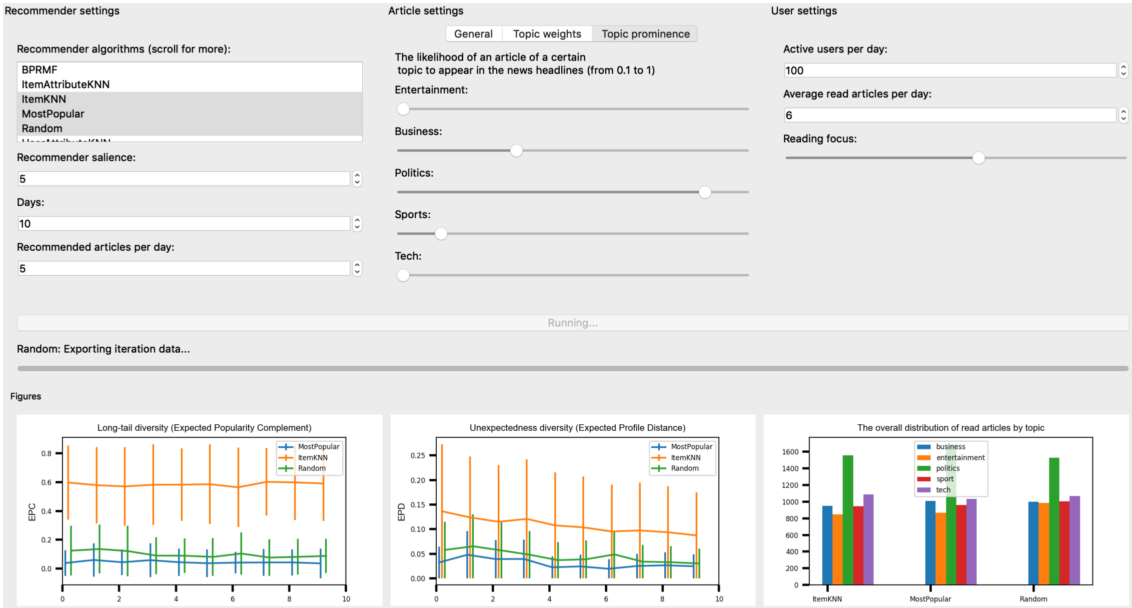

Figure 1: SIREN's user interface: recommender settings, news article settings and user (i.e. news reader) settings can be adjusted at will. The bottom row shows the generated visualizations for different metrics and different recommender algorithms.

to accommodate for a number of unique news-article characteristics, such as their large volumes and short-term relevancy. Another unique characteristic relates to the nature of the medium in which articles typically reside. News articles do not exist in isolation but appear within the website's layout and overall content. Editorial cues, such as position or font sizes, are frequently used to lead the reading consumption [3, 29] and to adjust the salience of the recommendations. The news-reading behavior is also unique, departing from the typical "show me something interesting" attitude [6]. A number of works [7, 53] suggest that besides the casual information seekers, online news accommodate the needs of users with specific preferences and interests. Finally, it has been suggested that those user preferences are likely to evolve over time [30, 32]; and thus, personalized news recommendations might have long-term effects.

We define three requirements that a simulation model of online news, with personalized recommendations, should satisfy in order to adequately approximate reality: (1) users distribute their reading time between prominent, sought out and recommended articles (2) user preferences evolve (3) the prominence ranking of articles is based on editorial cues such as their position on the news website.

3 SIREN

Based on the requirements identified above, we now present our simulation framework, with a strong emphasis on the *simulation model*, which takes Fleder's and Hosanagar's model (FH from now on) of consumers and products as a starting point [8, 9].

FH considers a map of users (consumers) and items (products) in an *n*-dimensional feature space. The product's position describes its properties, while the consumers' position corresponds to their ideal product. FH model uses such a two-dimensional space as its simulation's input for the sake of simplicity and visualization. Under FH, items centered around the origin $(0, 0)$ correspond to popular items. At each iteration of the simulation, the users are aware of items in their spatial proximity and popular items. In addition, at each iteration, recommended items are permanently added to the users' awareness. For each user, FH decides the items they purchase based on their distance to the user and an uncertainty component that allows users to deviate from their personal preferences. After a number of iterations, the simulation is complete, while the user preferences and collection of items remain static throughout.

Similar to FH, our model assumes that there are $|\mathcal{U}|$ users (i.e. readers) and $|\mathcal{T}|$ items (i.e. articles) placed in an 2-dimensional attribute space, $\mathcal{U}, \mathcal{T} \in \mathbb{R}^2$. Each iteration of the simulation corresponds to a news cycle (e.g., a day). Readers are aware of (i) articles in their proximity, corresponding to preferred/sought out topics (via search or navigation bars), (ii) promoted articles by the editors (as they appear on the news website), and (iii) personalized recommended articles $\in \mathcal{T}$ (as they appear on the website, or via email alerts sent to the reader). At each iteration, each user decides to read a number of unique articles from those they are aware of. At the end of each iteration, the users' preferences $\mathcal{U}$ are updated. The article pool $\mathcal{T}$ and the personalized recommendations are also updated at every iteration, while each article has a limited life-span.

Under this model, we identify three main interacting components that SIREN's interface gives content providers control over (cf., Figure 1): the *articles* (that translate to specific publishing habits), the *users* (the readers' preferences and reading behavior) and the *recommendations* (articles promoted to each user). We now describe the different components of our model in detail. We then present the *metrics* which we visualize in SIREN's interface. Finally, since this model has been implemented in our framework we conclude this section with a description of SIREN's technology stack.

3.1 Articles

As previously discussed in Section 2.2, a successful simulation model of online news consumption should consider the articles' content and article prominence.

3.1.1 Content. Under FH, the dimensions of the plane correspond to arbitrary features. In our case, we decided for the articles (products) and readers (consumers) to be placed in a content-driven topic space, since getting online news is a deliberate experience that may vary depending on the topic [36].

Typically, in marketing analyses where FH comes from, two-dimensional product space is obtained by performing multivariate analysis (e.g, PCA, *t-SNE*) on the high dimensional feature space [59]. In order to generate a meaningful (yet generic enough for any content provider to use) topic space for our framework, we need a collection of articles that capture a common feature distribution among most current news outlets. We assume that this requirement is largely satisfied by articles published by the major U.K. media outlet BBC. As such, we use the BBC dataset which contains 2225 documents corresponding to articles published on BBC News from 2004 to 2005 [14]. Each story belongs to a topical area (business, entertainment, politics, sports, tech). We represent each document as a *tf-idf* vector and then use *t-SNE* to project the high dimensional features into a two-dimensional plane (Figure 2).

3.1.2 Prominence. As previously discussed, each article has a certain prominence that is largely decided by the news editors. Under our model, each article $t_j \in \mathcal{T}$ comes with a prominence attribute $z_j \in [0, 1]$. While z_j might change over time (as it will be discussed later) each article has an initial article-prominence z_j^0. This is the extent to which editors promote an article on its first day of publication. Due to the lack of relevant literature, we assume that z_j^0 follows a long-tail distribution. Intuitively, only few articles make the headlines; most of the articles are either concentrated in the bottom of the webpage or are only accessible through searching. At the same time, articles of different topics are not promoted equally [21]. To address this, our framework allows content providers to adjust the weights of each topic with regard to their initial prominence. The weights are converted to the Cumulative Distribution Function and each z_j^0 value (ordered from high to low) is assigned to a topic (without replacement).

With every news cycle, the old articles' prominence is typically adapted to make room for the new, e.g., by moving them lower in the webpage's layout. Our model accommodates for this fact by adjusting the prominence z_j at each iteration. The collective effect of editorial cues and reader interest is that the readers' interaction with each article decreases with time, or as Chen et al. [4] describe it: a news article is "a life form with stages of birth, growth, decay and death". A number of empirical works support this observation. Mitchell et al. [37] report that roughly 80% and 90% of the interactions with an article happen within the first two and five days of its publication respectively (exponential decrease). Wang et al. [57] present a slightly different picture where the user interest fades at a slower rate. For the sake of simplicity, our model assumes that the prominence at the x^{th} day of the article's life is modeled with a linear function: $z_j^x = (-p\,x + 1)\,z_j^0$, where $p = 0.1$ is the slope (i.e., a 10-iteration lifespan).

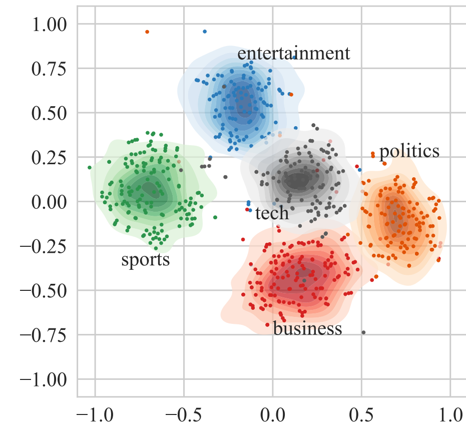

Figure 2: A scatter plot of the *tf-idf* feature vector for each BBC article projected on two dimensions (our simulation's topic space), accompanied by the kernel density estimate (KDE).

3.2 Users

We now focus on modeling the users, that is their preferences and their behavior i.e., the process behind each user making a choice to read one or more articles while taking into account the user-specific requirements as described in section 2.2.

3.2.1 Preferences. As previously discussed both the articles' content and the reader's preferences i.e., ideal article, are represented as points on the topic space (cf., Section 3.1.1). Under the FH model, users' preferences remain static no matter their purchase history. In order to accommodate for evolving user-preferences, we introduce a *user-drift model.* After reading article t_j, a user's likelihood to drift towards the article's position in the topic space is sampled from:

$$P(u_i \text{ drifts towards } t_j) = e^{-\text{distance}_{ij}^2/\theta_i^*} \quad (1)$$

where distance_{ij} is the Euclidean distance from reader u_i to article t_j. θ_i^* controls the width of the bivariate normal around the u_i user. In practice, θ_i^* controls the likelihood of the user drifting towards distant articles. We sample θ_i^* from a uniform distribution, reflecting the assumption that readers vary with respect to their eagerness to evolve their reading preferences. The user covers $m\ \times \text{distance}_{ij}$ distance towards the article. We argue that m does not affect the direction of simulation results, only the magnitude. However, a relatively small m (e.g., 0.05) allows us to get a higher-resolution view of the temporal dimension of the recommender effects.

3.2.2 User choice. Our model assumes that prior to any choice, each reader is aware of a limited number w of articles per iteration. In a scenario of no editorial priming, readers are only aware of the items they seek out. In such a case, the reader-article awareness can be a function of solely their spatial relationship on the topic space. To accommodate for article-prominence, we adapt the original FH model such that the user awareness is sampled from:

$$P(u_i \text{ aware of } t_j) = \lambda\theta' \log(1 - z_j)^{-1} + (1 - \lambda)e^{-\text{distance}_{ij}^2/\theta} \quad (2)$$

where λ controls the users' balance between prominent (high lambda) and neighboring (low lambda) articles that are modeled to be in a user's awareness. θ controls how the awareness fades in the reader's proximity i.e., the width of the bivariate normal around reader's position on the plane (the choice for normal distribution comes from the original FH model). θ' controls how the awareness fades with respect to the prominence dimension. We use a logarithmic function for the prominence decay as it agrees with the general long-tail pattern of user attention in news articles [26], i.e., how the attention decays towards the bottom of a news webpage.

Fleder and Hosanagar adjust θ and θ' to create an "interpretable" base case. In contrast, for a more realistic approximation of the user awareness, we turn our attention to related empirical findings. The works of Tewksbury [52] and Mitchell et al. [36] indicate that each user is aware of at least two topics. In our topic space, bounded in $[-1, 1]$, the Euclidean distance between any pair of BBC articles follows a normal distribution ($\mu = 0.77$, $\sigma = 0.36$). Setting θ to 0.07 creates an awareness radius of size $\approx 2\sigma$ around each user; thus, articles of two different topics are highly likely (95%) to fall into the user's awareness. Considering that the contribution in the awareness pool from reader's proximity and editorial priming should be inverse proportional under λ, allows us to set θ' to 0.5.

At each iteration the readers decide to read a number of articles from their awareness pool. FH uses a choice model based on multinomial logit, a well-established practice in economics and marketing according to the authors:

$$\text{argmax}_{j \in W_i}(v_{ij} + \epsilon_{ij}) \tag{3}$$

where W_i is the set of articles in the reader's u_i awareness pool and $v_{ij} = -k\ log(\text{distance}_{ij})$ is the deterministic component (k to be discussed later). The stochastic component ϵ_{ij} is an i.i.d. random variable with extreme value distribution. Without the stochastic component, the readers would always select to read the articles closer to their preferences.

Variable k is crucial since it controls the uncertainty in the user's choice. The higher k, the less readers deviate from their original preference-based article ranking, and thus the more likely they are to select prominent or recommended articles. Fleder and Hosanagar set k to 10. However, no studies support a specific k value for an online news context. Nevertheless, according to the report of Mitchell et al. [36] 28% of the news interactions with the user's main topic, happen while getting news on another topic. This roughly implies that the readers should deviate from their main topic of preference roughly one-third of the times. To set k according to these findings, we first generate a set of 100 users placed on the topic space uniformly. We then assign a "main" topic to each user by using the class prediction of a Gaussian Mixture Model (GMM) trained on the projected BBC articles. We additionally generate 500 articles using the aforementioned GMM. For different k values, we then compute the choice (Eq. 3) for each user and count the articles whose topic disagrees with the user's main topic in the top five positions. We find, that one-third of each user's top five articles are of different topic than the user's main for $k \approx 3$.

3.3 Recommendations

In a typical news environment, the readers are recommended n articles (via emails alerts or the designated website element) at regular time intervals or after they have interacted with a number of articles. For the sake of simplicity, we assume that the recommendations are only updated at the beginning of each iteration.

The recommendations in our model have two effects. First, the articles are added to the user's personal awareness pool. While under FH, this effect is permanent, in a news context that is rarely the case considering the vast amount and short-term relevancy of articles. As such, we assume that the awareness effect of the recommendations holds for a single iteration. Secondly, assuming that recommendations carry a certain prominence, the t_j article's deterministic component as it relates to the u_i user increases: $v_{ij'} = v_{ij} + \delta$, where δ corresponds to a "salience boost".

However, not all recommended articles share the same prominence: different spatial patterns can be used for arranging the recommendations. Our model aims to accommodate for that fact. For simplifying purposes, similar to Vargas [56], we assume a list-based arrangement and thus a positional-bias in the user choice i.e., a rank-based likelihood of an article being selected. We take δ to be a function of the article's rank $\kappa \in \mathbb{Z}_{\neq 0}$ in the recommendation list. We use a simple exponential function, and as such $\delta' = \delta\beta^{\kappa-1}$ with $\beta = 0.9$, following the approach in Vargas [56].

3.4 Metrics

While the simulation output can be analyzed from the point of view of different research questions[2], SIREN focuses on the concept of diversity. In our work, we specifically focus on two types of diversity, long-tail diversity and unexpectedness further explained below. While a number diversity metrics have appeared in the literature over the years [25], we focus on the work of Vargas [56], since it offers a direct mapping between our diversity concepts of interest (long-tail, unexpectedness) to well-formulated metrics.

Long-tail diversity. In contrast to popularity-based recommendations, *long-tail diversity* focuses on recommending items which are less popular and obvious choices [56]. In the context of news, the possibility to integrate the long tail of topics has both commercial and normative benefits. In the former case, the under-presented stories can be highlighted, thus providing a more stable distribution of readers. In the latter case, it allows integrating into the public agenda highly relevant topics which might be otherwise overlooked [38]. To measure long-tail discovery, we use the *Expected Popularity Complement* (EPC) metric. EPC is a user-oriented metric and takes into account the items' rank, relevance (whether the user selected them) and overall popularity.

Unexpectedness. Unlike long-tail diversity, which takes into account the interactions of all users with available content, *unexpectedness diversity* focuses on the individual user activity. This type of diversity refers to the possibility of locating a story which is unexpected, but still useful for the reader. The integration of unexpectedness in recommenders is known to increase user satisfaction [20] and broaden user preferences by diversifying their interests

[2] In fact, SIREN allows content providers to download the full extent of the simulation data and analyze it according to their needs.

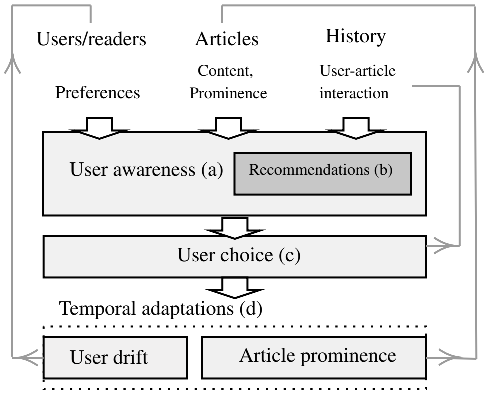

Figure 3: The framework's main variables and modules, and the interactions between them over the course of one simulation iteration. Bold arrows represent input/output flow, while thin arrows represent update functions. For example, each user's choice is based on the articles in their awareness, while after each choice, the user's preferences are updated by the user-drift component.

[24, 60]. From a normative point of view, unexpectedness is integral for countering negative effects of over-fitting, such as ideological/topical isolation resulting from "filter bubbles"/"echo chambers" [61]. For the unexpectedness diversity, we use the *Expected Profile Distance* (EPD) metric, which besides rank and relevance, incorporates the content-based distance between items.

3.5 Technology Stack

We now describe turning the simulation model into a simulation framework. The framework's main variables and modules in addition to the interactions between them is shown in Figure 3.

SIREN's implementation at each iteration takes as input the user preferences $\mathcal{U}$, the articles' content $\mathcal{T}$ and prominence z_j and the current reading history $|\mathcal{U}| \times |\mathcal{T}|$. The user-awareness module (cf. Figure 3,a) first computes the awareness pool of each user from the inputs. The recommended articles for each user are then computed by passing the input to an external recommendation toolbox (cf. Figure 3,b); thus allowing for further extendability. SIREN integrates the recommendation algorithms as they are provided by the toolbox MyMediaLite[3] [11]. The algorithms cover a wide range of strategies, from the simple random, popular recommendations to the more sophisticated collaborative (CF) and content-based approaches. After the recommended articles are integrated to the users' awareness, the user-choice module (cf. Figure 3,c) computes the selected articles . Based on the users' choice, the final module in the pipeline deals with the temporal adaptations (cf. Figure 3,d): updating the user's preferences $\mathcal{U}$ and the articles' prominence z_j.

The simulation model of SIREN is implemented in Python and is available online[4].

[3]www.mymedialite.net, accessed August 2018
[4]Open-sourced at github.com/dbountouridis/siren

4 CASE STUDY

SIREN allows content providers to instantiate the simulation with parameters specific to their values, publishing habits and their readers' behavior. In order to showcase SIREN's benefits , we investigate the recommender effects on a "default", generic instantiation that agrees with the major news outlets from the U.S., for which a large amount of public data and studies are available. The instantiation settings are summarized in Table 1. In the next sections, these settings will be justified, followed by the analysis and results.

4.1 Articles

Our instantiation considers $|\mathcal{T}|$ articles to be sampled from a Gaussian Mixture Model (GMM) of five components fitted on the overall BBC document population (see section 3.1.1). As a reminder, we consider the BBC-based topic space to be generic enough to agree with most current news outlets, including our U.S. case study. Regarding the number of articles published per day i.e., the number of articles available to the users per iteration, Bell et al. [1] reveal that U.S. outlets publish articles at different rates, from 10 to 145 per day. We consider an average scenario of 100 articles per day and as such $|\mathcal{T}| = 100 \times d$, where d the total amount of iterations.

With regard to the initial article prominence z_j^0, in order to realistically distribute between topical areas, we turn our attention to the News Coverage Dataset (NCI)[5] comprising 3,200 topic-annotations of the top-five most prominent articles appearing in twelve major U.S. online outlets, January to May 2012. The percentage of politics, sports, business, entertainment and tech topic appearing in the headlines is 85%, 3%, 7%, 5%, and 1% respectively. These percentages translate to a distribution of the article-prominence across topics via the process described in Section 3.1.2.

4.2 Users

We now turn our attention to the users/readers. We are interested in active readers, that is, subscribed users that receive personalized recommendations and read more articles than casual readers. Similar to Li et al. [31], we assume that a significant preference-evolution over time can happen only for active users. Our instantiation considers the readers' preferences, or position in the topic space, to be sampled from a uniform distribution, thus assuming a scenario where the readers' interest as a group is spread evenly across the topic space. Such a distribution captures a community of readers unaffected by recommender effects. Regarding the number of active users $|\mathcal{U}|$, while the amount of subscribed users for popular outlets is known [42], the exact percentage of those who are genuinely active is not supported by any literature. For the purposes of this case study, we consider a scenario of 200 active readers daily.

We now focus on instantiating the reading behavior. According to the Kaleida report[6], casual readers are on average exposed to 16 articles per day. We can assume that active readers are exposed to more articles, as such we set the maximum size of awareness w to 40. Regarding the awareness balance between sought out (neighboring) and prominent articles λ (see Eq. 2), Fleder and Hosanagar, supported by marketing studies on online purchasing behavior, set

[5]www.pewresearch.org, accessed August 2018
[6]survey.kaleida.com/Kaleida-news-ecosystem-report-europe-2018.pdf, accessed August 2018

Table 1: List and description of the variables governing the simulation. Besides the selection of recommender algorithms from MyMediaLite, the "Adjustable" variables are available for parameterization via SIREN's interface. The "Default" parameterization is used for this paper's case study and as the default settings on SIREN.

	Variable	Adjustable	Default	Description
User settings	$\lvert\mathcal{U}\rvert$	✓	200	Total number of active, daily users/readers.
	θ	✗	0.07	Awareness decay with distance.
	θ'	✗	0.5	Awareness decay with article prominence.
	λ	✓	0.6	Awareness weight placed on prominent versus neighborhood articles.
	w	✓	40	Maximum size of awareness pool.
	k	✗	3	Choice model: the user's sensitivity to distance on the map.
	θ_i^*	✗	$\sim \mathcal{N}(0.1, 0.03)$	User-drift: user's sensitivity to distance on the map.
	m	✗	$0.05 \times \text{distance}_{ij}$	User-drift: distance covered between the article t_j and user u_i.
	s	✓	$\sim \mathcal{N}(6, 2)$	Amount of articles read per iteration per user (session size).
Recommender settings	n	✓	5	Number of recommended articles per user per iteration.
	δ	✓	1	Factor by which distance decreases for recommended articles (salience).
	β	✗	0.9	Ranking-based decay of recommender salience.
	d	✓	30	Number of simulation iterations per recommender.
Article settings	$\lvert\mathcal{T}\rvert$	✓	$d \times 100$	Total number of articles (number of iterations × articles per day).
	topic weights	✓	$\mathcal{U}(0, 5)$	Percentage of articles added per day/iteration per topic.
	z^0	✓	see section 4.1	Awareness: initial article prominence per topic.
	p	✗	0.1	Prominence decrease factor per iteration.

it to 0.75. However, the report of Mitchell et al. [36] indicates that an average of 22% and 35% of the articles are accessed through search engines and news websites respectively[7]. Thus, 40% of the total interactions, without recommendations, should happen due to the user's proximity to the article, and therefore $\lambda = 0.6$.

4.3 Recommendations

We now turn our attention to the recommender settings. While the exact number n of recommended items varies from outlet to outlet, we assume a generic scenario of $n = 5$. Regarding the recommender salience δ (see section 3.3) Mitchell et al. [36] indicate that roughly 20% of the article interactions happen via email alerts, which typically contain personalized recommendations (the extent and type of which is unknown). The analysis of Kille et al. [22] shows that roughly one out of 80 user-article interactions happen due to in-article recommendations, but similarly, the recommendation strategy is unknown. We adjust δ according to Mitchel et al., as many of our model's parameterizations are based on their report. Given the current simulation instantiation, for all algorithms from the MyMediaLite toolbox the average ratio of recommended reads to the rest (sought out and prominent) is ≈ 0.2 for $\delta = 1$.

4.4 Analysis setup

We ran the recommender systems simulations for $d = 30$ iterations, as pilot experiments have indicated that it takes that amount of iterations for the simulation to converge. In order to deal with the cold-start problem, prior to the recommenders, we run a "control" period of 30 iterations with the recommendations and user-drift deactivated. We take the users' reading history from the "control" period as the initial input for the recommenders.

Following on pilot experiments, we select to investigate five recommenders the exhibited interesting behavior: the Random and MostPopular algorithms can be seen as baselines which recommend random and most popular articles respectively. We also select two common collaborative filtering algorithms: ItemKNN and UserKNN. The item-based k-nearest neighbor algorithm (ItemKNN) recommends items from a set of k similar articles for each of the articles that the user has read, while UserKNN recommends items from the k most similar users (in terms of reading habits). We also select a more sophisticated algorithm, WeightedBPRMF [10] which is an extension of the Bayesian Personalized Ranking (BPR) framework [47] that aims to reduce the problem of learning to rank into binary classification based on Bayesian analysis.

[7]The rest of the interactions happen via social media, email alerts etc.

4.5 Results

4.5.1 Long-tail diversity (EPC). Figure 4 (left) presents the long-tail diversity over the course of the simulation. Starting from the worst performing algorithms, it is no surprise that MostPopular presents the worst performance on long-tail diversity since it focuses on popular items. Interestingly, the random recommender strategy is only slightly better than MostPopular. While Random recommends articles from the long-tail, the likelihood of users reading them is small, thus the long-tail diversity is minimal.

The best performing algorithms are ItemKNN, and WeightedBPRMF. It is interesting that the simple ItemKNN outperforms the more sophisticated approach, although WeightedBPRMF eventually converges to the same EPC diversity as the number of iterations/days increases. The reason behind ItemKNN's high performance becomes more clear if we consider Figure 5; the position of users and articles on the topic space throughout the simulation's length. We observe that neither ItemKNN nor Most Popular (that is used as a reference in Figure 5) can prevent readers from concentrating around the topical centers. However, ItemKNN generates small clusters of users (circled in red on the figure) that are distributed across a topic, thus allowing a wider range of the topic space to be explored. In addition, if we observe the drift lines in the same figure, we can

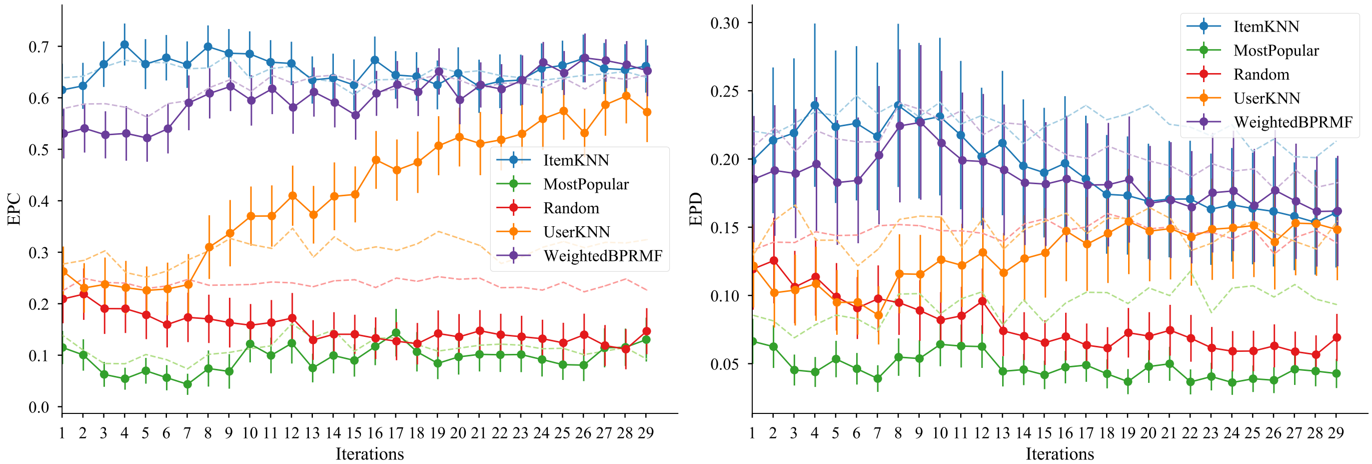

Figure 4: Long-tail diversity (left) and unexpectedness diversity (right) over 30 simulation iterations measured using the EPC and EPD metrics respectively, for five `MyMediaLite` algorithms (the error bars correspond to half the standard deviation for the sake of visual clarity). The dotted lines correspond to the diversity of the algorithms over the same simulation but with the user-drift deactivated (with no error bars for the sake of visual clarity).

see that `ItemKNN` demonstrates a greater degree of user drift, for those users with preferences not completely covered by the article selection i.e., users with initial position around the 1 radius.

Returning to Figure 4, another interesting case is `UserKNN`. We observe that `UserKNN` gradually increases the long-tail diversity until it converges close to the top-ranked algorithms. This behavior relates to the fact that the user preferences evolve: the same simulation run with the user-drift deactivated (dotted lines in Figure 4) reveals that for static users , `UserKNN` fails to show a behavior of increasing diversity. Intuitively, the more users are concentrated, the more `UserKNN` can accurately identify similar users, which in turn increases the likelihood of users reading a recommended item.

4.5.2 Unexpectedness diversity (EPD). We now turn our attention to the unexpectedness diversity measured using the EPD metric (see Figure 4, right). At first sight, we observe the overall higher variance of EPD compared to EPC i.e., for most algorithms users experience unexpectedness at different levels. Looking closer, we observe a similar behavior to the EPC metric: `MostPopular` and `Random` provide the least diversity, while `ItemKNN` and `WeightedBPRMF` the most. `UserKNN` starts at a diversity close to `Random` but eventually converges to a value close to the top-ranked algorithms. Interestingly, the unexpectedness seems to follow a downward slope for the top-ranked algorithms and `Random`. This behavior again relates to the evolving user preferences. A comparison with the diversity results with the user-drift deactivated (dotted lines on the figure) strongly indicate that the more users concentrate around the central topical areas, the less unexpected the recommendations become.

4.6 Discussion

Our analysis provides some interesting insights. First, the overall downward EPD slope and the increasing EPC diversity suggest that the recommenders effects with respect to diversity are dependent on the evolution of the readers' preferences. While the temporal effect has been already known for other contexts [28], our results suggest that such effects may extend to online news. As such, studying personalized news recommenders and their impact on the public sphere demands a focus on their temporal behavior. Consequently, we argue that studies based on snapshots of real-life data, e.g., [38], can only provide a short-term understanding of the recommender effects. Secondly, the overall difference (with respect to both diversity metrics) between the drift and non-drift simulations indicates that evolving user preferences can be either beneficial or unfavorable to certain recommendation strategies. For example, `UserKNN` is certainly benefited, while `Random` and `MostPopular` are hindered.

Even if the immediate effects of the recommender system do not lead to overall lack of diversity, our two discussion points suggest that these effects can be influenced by the users' changing preferences in a way that will have negative consequences for the public sphere, such as societal polarization. This implies that content-providers should aim to understand their users' impulse to change preferences prior to adopting any algorithm.

Finally, while the correspondence of the diversity metrics to the user's perception remains an active challenge [25], we observe that common collaborative filtering algorithms, such as `ItemKNN`, `UserKNN`, can be more or similarly diverse (both in terms of long-tail and unexpectedness) to sophisticated alternatives i.e., `WeightedBPRMF`. For `ItemKNN`, its relatively strong performance is not surprising considering that a number of works e.g., by Lathia et al. [28] or Park et al. [46], already support it. The case of `UserKNN`, on the other hand, is more surprising considering the lack of works supporting its diversity potential and requires further investigation. Nevertheless, for content providers interested in offering a personalized experience, while sensitive to societal challenges, such simple recommendation strategies can be valid candidates.

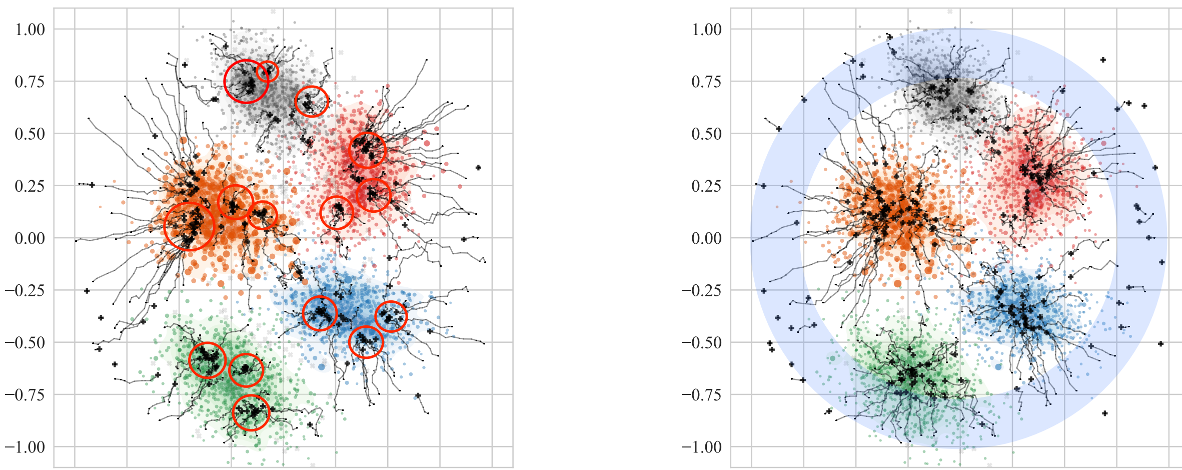

Figure 5: The position of users/readers (cross shapes +) and articles (circular points •) on the topic space after 30 iterations of the simulation run with the default settings for two algorithms: `ItemKNN` (left) and `MostPopular` (right). The black lines track the user preferences over time (a.k.a. the user-drift). The size of the articles corresponds to the amount of reads they received, normalized by the sum of all reads. Their colors correspond to their topic label. For the `ItemKNN` case, for the sake of visual inspection, a number of user clusters are circled in red. For the `MostPopular` case, the region of the ring with radius 1 (colored in light blue) encapsulates a large number of users that failed to drift.

5 CONCLUSIONS

We proposed and developed an online news consumption simulation framework SIREN, that visualizes and analyzes the effects of recommender systems in order to help content providers decide better what algorithms to deploy. In light of the widespread concerns about the societal effects of curating algorithms, we argue that our focus helping content providers be more aware of the effects of different recommendation algorithms is an under-explored way to mitigate their potentially nefarious effects.

Nevertheless, we should address the limitations of our approach. Any simulation model is an approximation of reality which potentially misrepresents the complexity of the phenomenon in question. In our case, while evolving user preferences were considered, we neglected the full complexity of editorial priming [3] and the temporal news consumption patterns (week days vs. weekends) among others. While our model is largely based on literature findings, certain components of reality were simplified while others were modeled based on intuition. Yet our case study's findings conform with previous work on diversity in recommender systems, giving support to the reliability of our framework's conceptual model.

Despite the limitations, our framework accounts for much of the complexity of news consumption in an online environment. Moreover, its strength comes from its easy extendability and potential integration of other features, such as different metrics and recommendation algorithms. At the same time, SIREN allowed us to get a glimpse into the recommender effects in the context of U.S. online news and provided valuable insights, that would have remained obscured otherwise e.g., the temporal dimension of diversity or the diversity potential of common collaborative filtering techniques.

Finally, future research with SIREN should accommodate for and explore the recommendation effects in different contexts, such as different types of users. We are also planning to engage in a discourse with content providers to further understand their needs and particularities. SIREN will be regularly updated such that a number of urgent research questions can be readily answered.

ACKNOWLEDGMENTS

This research has been supported by NWA Startimpuls VWData, Delft Data Science and NWO project SearchX (639.022.722).

REFERENCES

[1] Emily J Bell, Taylor Owen, Peter D Brown, Codi Hauka, and Nushin Rashidian. 2017. The platform press: How Silicon Valley reengineered journalism. (2017).

[2] Engin Bozdag and Jeroen van den Hoven. 2015. Breaking the filter bubble: democracy and design. *Ethics and Information Technology* 17, 4 (2015), 249–265.

[3] Taina Bucher. 2017. âĂŸMachines donâĂŹt have instinctsâĂŹ: Articulating the computational in journalism. *New Media & Society* 19, 6 (2017), 918–933.

[4] Chien Chin Chen, Yao-Tsung Chen, Yeali Sun, and Meng Chang Chen. 2003. Life cycle modeling of news events using aging theory. In *Proceedings of the European Conference on Machine Learning*. Springer, 47–59.

[5] Stanley Cohen. 2011. *Folk devils and moral panics*. Routledge.

[6] Abhinandan S Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. 2007. Google news personalization: scalable online collaborative filtering. In *Proceedings of the international Conference on World Wide Web*. ACM, 271–280.

[7] Carlos Flavian and Raquel Gurrea. 2006. The choice of digital newspapers: influence of reader goals and user experience. *Internet Research* 16, 3 (2006), 231–247.

[8] Daniel Fleder and Kartik Hosanagar. 2009. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science* 55, 5 (2009), 697–712.

[9] Daniel M Fleder and Kartik Hosanagar. 2007. Recommender systems and their impact on sales diversity. In *Proceedings of the ACM Conference on Electronic commerce*. ACM, 192–199.

[10] Zeno Gantner, Lucas Drumond, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2012. Personalized ranking for non-uniformly sampled items. In *Proceedings of KDD Cup 2011*. 231–247.

[11] Zeno Gantner, Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2011. MyMediaLite: A Free Recommender System Library. In *Proceedings of the ACM Conference on Recommender Systems (RecSys 2011)*.

[12] R Kelly Garrett, Shira Dvir Gvirsman, Benjamin K Johnson, Yariv Tsfati, Rachel Neo, and Aysenur Dal. 2014. Implications of pro-and counterattitudinal information exposure for affective polarization. *Human Communication Research* 40, 3 (2014), 309–332.

[13] Tarleton Gillespie. 2014. The relevance of algorithms. *Media technologies: Essays on communication, materiality, and society* 167 (2014).

[14] Derek Greene and Pádraig Cunningham. 2006. Practical solutions to the problem of diagonal dominance in kernel document clustering. In *Proceedings of the international Conference on Machine learning*. ACM, 377–384.

[15] Mark Hampton. 2010. The Fourth Estate ideal in journalism history. *The Routledge companion to news and journalism* (2010), 3–12.

[16] Jaron Harambam, Joris van Hoboken, and Natali Helberger. 2018. Democratizing algorithmic news recommenders: How to materialize voice in a technologically saturated media ecosystem. *Philosophical Transactions of the Royal Society* (2018). https://doi.org/DOI10.1098/rsta.2018.0088.

[17] Natali Helberger, Kari Karppinen, and Lucia D'Acunto. 2018. Exposure diversity as a design principle for recommender systems. *Information, Communication & Society* 21, 2 (2018), 191–207.

[18] Matthew Hindman. 2008. *The myth of digital democracy*. Princeton University Press.

[19] Oliver Hinz and Jochen Eckert. 2010. The impact of search and recommendation systems on sales in electronic commerce. *Business & Information Systems Engineering* 2, 2 (2010), 67–77.

[20] Maximilian Jenders, T Lindhauer, Gjergji Kasneci, Ralf Krestel, and Felix Naumann. 2015. A serendipity model for news recommendation. In *Proceedings of the Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*. Springer, 111–123.

[21] Michael Bo Karlsson. 2016. Goodbye politics, hello lifestyle. Changing news topics in tabloid, quality and local newspaper websites in the UK and Sweden from 2002 to 2012. *Observatorio (OBS*)* 10, 4 (2016), 150–165.

[22] Benjamin Kille, Frank Hopfgartner, Torben Brodt, and Tobias Heintz. 2013. The plista dataset. In *Proceedings of the International News Recommender Systems Workshop and Challenge*. ACM, 16–23.

[23] Rob Kitchin. 2017. Thinking critically about and researching algorithms. *Information, Communication & Society* 20, 1 (2017), 14–29.

[24] Denis Kotkov, Joseph A Konstan, Qian Zhao, and Jari Veijalainen. 2018. Investigating Serendipity in Recommender Systems Based on Real User Feedback. (2018).

[25] Matevž Kunaver and Tomaž Požrl. 2017. Diversity in recommender systems–A survey. *Knowledge-Based Systems* 123 (2017), 154–162.

[26] D Lagun and M Lalmas. 2016. Understanding and measuring user engagement and attention in online news reading. In *Proceedings of the ACM International Conference on Web Search and Data Mining*. 113–122.

[27] Daniel Lamprecht, Markus Strohmaier, and Denis Helic. 2017. A method for evaluating discoverability and navigability of recommendation algorithms. *Computational social networks* 4, 1 (2017), 9.

[28] Neal Lathia, Stephen Hailes, Licia Capra, and Xavier Amatriain. 2010. Temporal diversity in recommender systems. In *Proceedings of the international ACM SIGIR conference on Research and development in information retrieval*. ACM, 210–217.

[29] Sara Leckner. 2012. Presentation factors affecting reading behaviour in readers of newspaper media: an eye-tracking perspective. *Visual Communication* 11, 2 (2012), 163–184.

[30] Lei Li, Ding-Ding Wang, Shun-Zhi Zhu, and Tao Li. 2011. Personalized news recommendation: a review and an experimental investigation. *Journal of computer science and technology* 26, 5 (2011), 754.

[31] Lei Li, Li Zheng, Fan Yang, and Tao Li. 2014. Modeling and broadening temporal user interest in personalized news recommendation. *Expert Systems with Applications* 41, 7 (2014), 3168–3177.

[32] Jiahui Liu, Peter Dolan, and Elin Rønby Pedersen. 2010. Personalized news recommendation based on click behavior. In *Proceedings of the international Conference on Intelligent user interfaces*. ACM, 31–40.

[33] Nikos Manouselis and Constantina Costopoulou. 2006. Designing a web-based testing tool for multi-criteria recommender systems. *Engineering Letters, Special Issue on Web Engineering* 13, 3 (2006).

[34] Nikos Manouselis, Riina Vuorikari, and Frans Van Assche. 2007. Simulated analysis of MAUT collaborative filtering for learning object recommendation. In *Proceedings of the Workshop on Social Information Retrieval for Technology Enhanced Learning*. 27–35.

[35] Robert K Merton. 1968. The Matthew effect in science: The reward and communication systems of science are considered. *Science* 159, 3810 (1968), 56–63.

[36] Amy Mitchell, Jeffrey Gottfried, Elisa Shearer, and Kristine Lu. 2017. How Americans encounter, recall and act upon digital news. *Pew Research Center* (2017).

[37] Amy Mitchell, Galen Stocking, and Katerina Eva Matsa. 2016. Long-form reading shows signs of life in our mobile news world. *Pew Research Center* (2016).

[38] Judith Möller, Damian Trilling, Natali Helberger, and Bram van Es. 2018. Do not blame it on the algorithm: An empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society* 21, 7 (2018), 959–977.

[39] Raymond J Mooney and Loriene Roy. 2000. Content-based book recommending using learning for text categorization. In *Proceedings of the ACM Conference on Digital libraries*. ACM, 195–204.

[40] Géraldine Muhlmann. 2010. *Journalism for democracy*. Polity.

[41] Rob J Nadolski, Bert Van den Berg, Adriana J Berlanga, Hendrik Drachsler, Hans GK Hummel, Rob Koper, and Peter B Sloep. 2009. Simulating light-weight personalised recommender systems in learning networks: A case for pedagogy-oriented and rating-based hybrid recommendation strategies. *Journal of Artificial Societies and Social Simulation* 12, 1 (2009), 4.

[42] Nic Newman, Richard Fletcher, Antonis Kalogeropoulos, David AL Levy, and Rasmus Kleis Nielsen. 2017. Reuters Institute digital news report 2017. (2017).

[43] Tien T Nguyen, Pik-Mai Hui, F Maxwell Harper, Loren Terveen, and Joseph A Konstan. 2014. Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the international Conference on World wide web*. ACM, 677–686.

[44] Gal Oestreicher-Singer and Arun Sundararajan. 2012. Recommendation networks and the long tail of electronic commerce. *Mis quarterly* (2012), 65–83.

[45] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.

[46] Keunchan Park, Jisoo Lee, and Jaeho Choi. 2017. Deep Neural Networks for News Recommendations. In *Proceedings of the ACM on Conference on Information and Knowledge Management*. ACM, 2255–2258.

[47] Steffen Rendle and Lars Schmidt-Thieme. 2009. Factor models for tag recommendation in bibsonomy. In *ECML/PKDD 2008 Discovery Challenge Workshop, part of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*. 235–243.

[48] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2015. Recommender systems: introduction and challenges. In *Recommender systems handbook*. Springer, 1–34.

[49] Rory LL Sie, Marlies Bitter-Rijpkema, and Peter B Sloep. 2010. A simulation for content-based and utility-based recommendation of candidate coalitions in virtual creativity teams. *Procedia Computer Science* 1, 2 (2010), 2883–2888.

[50] Natalie Jomini Stroud. 2011. *Niche news: The politics of news choice*. Oxford University Press on Demand.

[51] Cass R Sunstein. 2018. *# Republic: Divided democracy in the age of social media*. Princeton University Press.

[52] David Tewksbury. 2003. What do Americans really want to know? Tracking the behavior of news readers on the Internet. *Journal of communication* 53, 4 (2003), 694–710.

[53] David Tewksbury, Michelle L Hals, and Allyson Bibart. 2008. The efficacy of news browsing: The relationship of news consumption style to social and political efficacy. *Journalism & Mass Communication Quarterly* 85, 2 (2008), 257–272.

[54] Kjerstin Thorson and Chris Wells. 2015. Curated flows: A framework for mapping media exposure in the digital age. *Communication Theory* 26, 3 (2015), 309–328.

[55] Takashi Umeda, Manabu Ichikawa, Yuhsuke Koyama, and Hiroshi Deguchi. 2014. Evaluation of collaborative filtering by agent-based simulation considering market environment. *Developments in Business Simulation and Experiential Learning* 36 (2014).

[56] S Vargas. 2015. *Novelty and diversity evaluation and enhancement in recommender systems*. Ph.D. Dissertation. Ph. D. thesis.

[57] Canhui Wang, Min Zhang, Liyun Ru, and Shaoping Ma. 2008. Automatic online news topic ranking using media focus and user attention based on aging theory. In *Proceedings of the Conference on Information and knowledge management*. ACM, 1033–1042.

[58] Magdalena Wojcieszak. 2011. When deliberation divides: Processes underlying mobilization to collective action. *Communication Monographs* 78, 3 (2011), 324–346.

[59] Thierry Worch, Sébastien Lê, Pieter Punter, and Jérôme Pagès. 2012. Construction of an Ideal Map (IdMap) based on the ideal profiles obtained directly from consumers. *Food quality and preference* 26, 1 (2012), 93–104.

[60] Qianru Zheng, Chi-Kong Chan, and Horace HS Ip. 2015. An unexpectedness-augmented utility model for making serendipitous recommendation. In *Proceedings of the Industrial Conference on Data Mining*. Springer, 216–230.

[61] Ethan Zuckerman. 2013. *Digital cosmopolitans: Why we think the Internet connects us, why it doesn't, and how to rewire it*. WW Norton & Company.

[62] F Zuiderveen Borgesius, D Trilling, J Möller, B Bodó, C de Vreese, and N Helberger. 2016. Should we worry about filter bubbles. *Internet Policy Review* 5, 1 (2016), 2.