

# Are we human, or are we users?

## The role of natural language processing in human-centric news recommenders that nudge users to diverse content

Myrthe Reuver<sup>1\*</sup>    Nicolas Mattis<sup>1\*</sup>    Marijn Sax<sup>2\*</sup>  
Suzan Verberne<sup>3</sup>    Nava Tintarev<sup>5</sup>    Natali Helberger<sup>2</sup>    Judith Moeller<sup>2</sup>  
Sanne Vrijenhoek<sup>2</sup>    Antske Fokkens<sup>1,4</sup>    Wouter van Atteveldt<sup>1</sup>

\* The first three authors contributed equally, and are shared first authors.

<sup>1</sup> Vrije Universiteit Amsterdam    <sup>2</sup>University of Amsterdam    <sup>3</sup>Leiden University  
<sup>4</sup>Eindhoven University of Technology    <sup>5</sup>Maastricht University  
{myrthe.reuver, n.m.mattis}@vu.nl, M.Sax@uva.nl

### Abstract

In this position paper, we present a research agenda and ideas for facilitating exposure to diverse viewpoints in news recommendation. Recommending news from diverse viewpoints is important to prevent potential filter bubble effects in news consumption, and stimulate a healthy democratic debate. To account for the complexity that is inherent to humans as citizens in a democracy, we anticipate (among others) individual-level differences in acceptance of diversity. We connect this idea to techniques in Natural Language Processing, where distributional language models would allow us to place different users and news articles in a multidimensional space based on semantic content, where diversity is operationalized as distance and variance. In this way, we can model individual “latitudes of diversity” for different users, and thus personalize viewpoint diversity in support of a healthy public debate. In addition, we identify technical, ethical and conceptual issues related to our presented ideas. Our investigation describes how NLP can play a central role in diversifying news recommendations.

### 1 Introduction

Recommender systems are very present in our online experience. Services recommend movies in movie streaming platforms, possible purchases in online shops, and news articles on news sites. People increasingly live their lives *through* digital environments that recommend some items over others. Recommender systems thus have a significant influence on the lived experiences of people. Precisely because we can expect recommender systems to perform ever more (important) functions in the near future, it is essential to incorporate *the people they end up influencing* in our thinking about recommender systems.

We focus on one case in this paper: news recommender systems (NRS) such as news aggregators, online newspapers, or news widgets. We believe Natural Language Processing (NLP) should have a more prominent role in the development of NRS. These systems are an interesting case because the role of the media has historically always been understood as essential to democratic societies and democratic debate (Karpinen, 2013). Our central argument is that the users of NRS are not just collections of data points, but are democratic citizens, who perform more social roles than that of a consumer. We propose a research agenda to put the user *as a human being* at the centre of NLP and Computer Science research into recommender systems.

For a functioning democracy, we want users to come into contact with opinions, debates, or ideas they disagree with or even dislike. This implies that simply optimizing a news recommender system for user preference, as is common now, is not enough. The public value of diversity, in terms of a diversity of issues and opinions, is essential (Helberger, 2015, 2019). There already is work on the question of how opinion, sentiment, and argument diversity for news recommendation should be understood and captured by (evaluation) metrics (Vrijenhoek et al., 2021), or NLP tasks (Reuver et al., 2021). In this paper, we want to develop a different *additional* perspective. We propose that a turn to the user – i.e., democratic citizens reading news recommended by NRS – is needed. This paper explores challenges and opportunities for facilitating this with the help of NLP.

More specifically, we propose the notion of individual *latitudes of diversity* which can help make the diversity of news recommendations more meaningful by taking the user-as-a-human-being into account. Although the promotion of news diversity is desirable from societal perspective, not every user

has a similar acceptance of diversity. Simply maximizing the diversity of news recommendations can therefore have serious backfire effects (Helberger et al., 2018; Taber and Lodge, 2006).

We combine perspectives from communication science, NLP, and ethics to develop this contribution. Building on theories from communication science, we explore how the notion of latitude of diversity can help facilitate engagement with diverse content. Work in the NLP field is used to suggest how NRS can capture viewpoint diversity in news articles *and* connect to individual users' latitudes of diversity. Lastly, we incorporate ethical reflections on the value and need of diversity, as well as on some of the risks of designing (data-intensive) personalized recommendation engines.

In our paper, these different fields and perspectives all contribute to answer one central question: *How do we nudge citizens towards actual engagement with diverse viewpoints in news recommender systems with NLP techniques, while treating the user as a complex human and democratic citizen?*

To answer this question, our paper is organized as follows. Section 2 introduces why viewpoint diversity in the news context is important for democracy, how NLP is connected, and why focusing on maximizing viewpoint diverse recommendations is not enough. Section 3 presents our ideas and discusses the current technical and conceptual challenges relevant for building a more user-centric news recommender. Section 4 explores some important ethical questions that should be considered before implementing our ideas. Lastly, we restate our argument in Section 5.

## 2 Background

We first provide some background literature for our idea for diverse news recommendation. Section 2.1 addresses why viewpoint news diversity is important from a democratic perspective. In Section 2.2, we introduce how viewpoint diversity is connected to NLP. We then discuss nudging theory and how it can inform news recommender design in Section 2.3. Lastly, we explain how our concept of "latitude of diversity" can help make news recommenders more user-centric in Section 2.4.

### 2.1 Why (viewpoint) diversity matters in the news context

The literature on the importance of news diversity for a democratic society describes various mod-

els of democracy to explain how different types of diversity are important to democracy. The general idea found in this literature is that *depending on one's conception* of democracy, different *kinds* of news diversity can be important (Bozdag and van den Hoven, 2015; Dahlberg, 2011; Helberger, 2019; Strömbäck, 2005). Examples of such models are the deliberative model of democracy (Habermas, 2006), which emphasizes democracy requires a rational debate in society of different opinions and ideas, and the agonistic model (Mouffe, 2005), which emphasizes the importance of facilitating civil but ongoing *clashes* between different political beliefs, ideologies, and emotions.

Regardless of the specific democratic theory one supports in the news diversity context, nearly all strands of democratic theory emphasize the importance of promoting viewpoint diversity in this context. For example, for both the deliberative democracy and agonistic democracy models, "encouraging encounters between conflicting ideas seems to be a shared goal" (Karppinen, 2013). Both require citizens of a democracy to have a diverse news diet in order to be informed about a diversity of viewpoints, because this can help citizens to understand (and sometimes even empathize) with (the viewpoints of) other citizens. Diverse viewpoints can also provide them with information to help them think and deliberate critically about issues that matter to them or society in general. Serving citizens with a diverse set of viewpoints can also help invigorate productive clashes of political opinion or ideology.

With 'viewpoint', these theoretical models usually mean different arguments, claims, or ideas about the same publicly debated topics. Examples of such topics are vaccines and immigration. As viewpoint diversity is an (almost) universally shared goal among different democratic theories, we do not choose or support one specific model of democracy in particular.

### 2.2 The connection with NLP

The focus on viewpoint diversity has as a central task the detection of viewpoints in news articles. In the NLP field, the detection of different claims (Levy et al., 2014), arguments (Stab and Gurevych, 2017), and stances (Mohammad et al., 2016) are established tasks that are related. Work on such tasks is often on topics publicly debated in a political context, such as vaccinations, abortion, and

immigration. This makes them potentially useful for operationalization of the viewpoint concept.

Large-scale pre-trained language models (Devlin et al., 2019) are a recent development in NLP, and could be used to detect viewpoints. Reimers et al. (2019) use such methods for the NLP task claim detection. In this work such language models are also used to cluster similar claims and arguments, giving the opportunity to also detect *dissimilar* claims or arguments. See section 3.1.2 for more concrete and detailed ideas we have on the operationalization of viewpoints with NLP.

### 2.3 Nudging

Besides the detection of viewpoints, we also seek to incorporate nudge-like personalization features. The key insight from the nudging literature (Thaler and Sunstein, 2008) is that environments can be designed in a manner that takes heuristics and biases into account to steer behavior of the users of these environments. Such nudges could also be aimed at the individual person. The most famous example is the ‘cafeteria example’, where the healthy food options are placed in an easier to reach spot than the unhealthy options. As a result, more people choose the healthier options *without making the unhealthy options completely unavailable*. The same approach can be connected to the idea of a “healthy media diet”, where “healthy” is connected to a healthy democracy and public debate.

Nudging has previously been incorporated in recommender systems. A 2021 systematic review by Jesse and Jannach (2021) reveals that of 87 nudging mechanisms identified in the literature, only a small subset was previously investigated in the context of recommender systems. These include using visuals to increase item salience, item re-ranking, and setting defaults. For news recommendation specifically, Gena et al. (2019) found that nudges based on giving users the idea certain items were popular were not effective, while negative framing (nudging users to consume certain news items because of limited availability) was. The authors argue that future work should address which types of nudges are ethically acceptable in the area of persuasive technologies.

### 2.4 Latitudes of diversity

We define *latitude of diversity* as an individual user’s acceptance of diversity. Research shows that (groups of) individual users can differ considerably in the extent to which they are open towards and

interested in diverse viewpoints (Kim and Pasek, 2020; Tintarev, 2017). We argue that considering individual users’ latitudes of diversity increases the likelihood that a given user engages with diverse recommendations and potentially prevents unwanted side effects.

If individual latitudes of diversity are not considered, users who are not interested in or open towards diverse viewpoints might simply not select diverse recommendations. Moreover, recommending news that are too diverse can backfire (Helberger et al., 2018). Motivated reasoning literature suggests that people evaluate counter-attitudinal information more critically than attitude-congruent information (Taber and Lodge, 2006). Under some circumstances, this might go as far as that exposure to counter-attitudinal information causes people to actively counter-argue, resulting in even stronger attitudes and increased political polarisation (Bail et al., 2018; Nyhan and Reifler, 2010). Moreover, the exposure to diverse viewpoints has also been linked to decreased political participation (Kim and Kwak (2017), but also see Matthes et al. (2019)). Thus, more diversity is not always better.

A (drastically simplified) example may help to show the potential relevance of latitudes of diversity. Imagine a close-minded extremist and a very moderate open-minded person. The extremist is not at all open to most of the news stories that do not closely resemble their own worldview; such stories fall far outside their latitude of diversity. Presenting them with very moderate news stories may enrage them and turn them off from the news source(s) in question and engaging in the public debate. So, given their latitude of diversity, it makes sense to focus on recommendations that are slightly less extremist than the one they would choose themselves, but that are still close enough to what they would accept so that there is at least a possibility that they would interact with the recommendations. It would not make much sense to recommend overly diverse content to the extremist that they would never engage with to begin with. For the moderate and open-minded person, the far wider latitude of diversity means that much more diverse content can be recommended to them without there being a serious risk that they doesn’t want to engage with the recommended content because it is too fringe for them.

Some methods and interventions that aim to diversify news consumption already exist (Bozdog

and van den Hoven, 2015). However, they do not take individual latitudes of diversity into account. Moreover, news consumption choices do not exist in a vacuum, but depend on the nature and content of the information environment (Powell et al., 2020), as well as various characteristics of the user and the situation in which they make a selection (Meijer and Kormelink, 2020; Tintarev, 2017). As of now, news recommendation often fails to capture these individual and situational differences in news selection. Understanding and modelling users’ individual latitude of diversity is one step towards alleviating this issue.

### 3 Ideas and challenges for user-centric diverse recommendation

Earlier solutions to diversity in NRS have focused on author or metadata-based diversity (Lu et al., 2020), or on optimizing for diversity without considering user interest and susceptibility (Vrijenhoek et al., 2021). We take the next step towards support of democratic values in NRS: putting the user central, and especially the individual’s latitude of diversity. We aim to optimize not for user preference, but for the user’s diversity range required to participate in a functioning democracy. This is a more nuanced and long-term goal than the short-term one of user preference. In order to answer our research question, we will discuss the different elements in this problem separately in the following subsections.

#### 3.1 Representation and processing of news articles

The representation of news articles for a viewpoint-diverse and user-centric recommendation has several sub-problems. In good NLP fashion, we break the diverse recommendation pipeline up into several sub-tasks, which also helps us to think about the problem(s) and solutions related to these.

We identify four separate steps in our pipeline. First, identifying the different current issues (or topics) in the news. Second, identifying different viewpoints on these current issues expressed in the media content. Third, measuring diversity computationally. And lastly, measuring and providing different latitudes of diversity to different users.

In the following sections, we will discuss these different sub-tasks of the problem of how to represent content of a news recommender system in order to create a diverse recommender. One partic-

ular challenge to the news domain is the cold start problem (news items are added every day, and the newest do not yet have user interaction data useful for a recommendation algorithm). A related issue specific for viewpoint-diverse recommendation is the constant appearance of new topics and notable entities in the news, and thus also new perspectives and viewpoints on these entities and topics that need to be detected. We address these issues in the following sections when they come up in our pipeline.

##### 3.1.1 Identifying current issues or topics in the news

Before we can identify what different perspectives or viewpoints are debated in the news, we need to automatically identify (contentious) political debates in large collections of news texts. Such debates are for instance ones on immigration, or vaccination. Commonly, only one or a handful recurring contentious debates are discussed in current work on arguments and debates in the news. Some such topics used as case-studies are the benefits and dangers of vaccination (Morante et al., 2020) and the ethics of abortion (Draws et al., 2021).

One option for identifying topics is a rule-based method with pre-defined lists or gazetteers of known contentious, newsworthy topics, for instance websites listing (political) debates topics, as done by Draws et al. (2021) and Roy and Goldwasser (2020), and using these lists either for further manual annotation of a training set used to train a classifying machine learning model, or using heuristics and rules to identify these topics in news articles. Another option is manual annotation of (journalistic) topics already being distinguished in the (online) news room by editors or journalists, as used by Lu et al. (2020), who use features such as website sections (e.g. “sports”, “politics” or more fine-grained: “U.S. elections”) and journalistic tags (e.g. “opinion”) to represent news articles for a diverse recommender.

One challenge here is the above-mentioned appearance of new topics in the news. One potentially useful technique for this is zero- or one-shot learning. In such an approach, the model learns to generalize from one or several example topic how to identify all kinds of different topics (or, in the next step, viewpoints) not encountered in the training data. This option has been explored for topic and stance detection in Allaway and McKeown (2020).

Current state of the art text classification relies

on vector space representations of texts. Traditional vector space models and neural language models such as Doc2vec (Le and Mikolov, 2014), the Universal Sentence Encoder (Cer et al., 2018), and sentence-BERT (Reimers and Gurevych, 2019) semantically represent documents (sentences) in a multi-dimensional space. We argue that vector space representations could also be useful when considering different users with different latitudes of diversity for different topics or viewpoints. It means modelling not only articles in a vector space, but also users.

### 3.1.2 Automatically identifying different viewpoints

We define a viewpoint as a public argument or claim in a public debate on an issue. For instance, concerning the topic of immigration a citizen can have different viewpoints, claims, and ideas.

Most of the current work on politically related debates in news articles exclusively uses English-language data, specifically: news articles news outlets based in the United States. The U.S. political context and also publicly debated issues make perspectives and viewpoints easily translatable to identifying two opposing broad political groups, and this is often what happens in such papers: left-wing (i.e. the Democratic party) and right-wing (the Republican party) viewpoints are detected and extracted, as in Roy and Goldwasser (2020). However, not every political climate has such a polarized and two-party political system, so such an approach might not fit every language or context. Nuanced concept from political science such as framing and agenda setting have also been analysed with NLP beyond the U.S context, e.g., in Russian news media (Field et al., 2018).

There are several related NLP tasks and solutions to identifying different viewpoints on (politically contentious) issues in news texts. Names for such tasks are stance detection (Hanselowski et al., 2018), argument mining (Stab and Gurevych, 2017), and perspective detection (Morante et al., 2020; van Son et al., 2016). All these tasks focus on capturing an opinion on issues, events, or entities, which make them useful for identifying viewpoints for a recommender system that supports democratic values such as deliberation. In Reuver et al. (2021), several of these aforementioned NLP tasks and their usefulness for this goal are discussed in detail. For instance, a ‘stance’ can be useful for operationalizing the idea of a ‘viewpoint’, since

stances often related to a particular opinion on a contentious issue (e.g. is the text pro, against, or neutral towards immigration?). This is inherently related to the idea of debates in society between different viewpoints on these contentious topics, implicitly or explicitly expressed in news articles.

Both unsupervised and supervised methods are used for identifying viewpoints. The most commonly used unsupervised approach is Topic Modelling, as used by Draws et al. (2021) and Mulder et al. (2021) to not only identify different topics, but also different perspectives and (sub)topics. An unsupervised approach is useful for a cold-start problem, but also potentially poses validation issues (do we know what the model is measuring? Is it actually measuring a coherent topic or viewpoint?). Most relevant NLP tasks (stance detection, claim detection) are thereby addressed with *supervised* methods. This means the models are trained at detecting viewpoints by examples in their training data, which are often manually annotated. Thus there is by definition a finite set of different viewpoints that the model can detect (the annotated ones).

Useful datasets for training models for detecting viewpoints consist of textdata on topics that are in public debate and the news, also annotated for the opinion on these topics that is expressed in the text unit (sentence, paragraph, or article). This is the case for the sentential argument mining corpus (Stab et al., 2018), which consists of English sentences on eight controversial topics, such as abortion and minimum wage, annotated for stance in three classes: pro, con, and neutral towards the topic. There are also established datasets that are more fine-grained, such as the MPQA Corpus (Deng and Wiebe, 2015): English news texts annotated for negative or positive sentiment towards targets (such as events or persons), but also more fine-grained annotation of opinions, beliefs, and judgements. Some datasets focus on stances, claims, or arguments on one topic, such as climate change (Luo et al., 2020; Varini et al., 2020) or vaccination (Morante et al., 2020).

Detecting new stances on new topics that are not in the training data could provide a challenge for supervised models. Like with new topic identification discussed above, few-shot learning has recently been used for complex semantics-related NLP tasks such as topic and stance detection and summarization, and allows language models to gen-

eralize beyond their training data (Allaway and McKeown, 2020; Schick and Schütze, 2020).

Do note that all mentioned NLP tasks and datasets would operationalize the idea of a viewpoint differently (e.g. either as an argument, claim, stance, or sentiment on a topic). Different tasks also often use different types of text: social media texts, online discussion boards, or news articles (see for a discussion Reuver et al. (2021)), and even within these tasks there are many completely different approaches in method and annotation. Selecting one of these frameworks, datasets, or tasks requires careful reflection on what aspect(s) of a viewpoint is central to a certain recommender, context, or even specific debate, and how NLP can best support this idea.

### 3.1.3 Defining, capturing, and evaluating diversity

After identifying topics and viewpoints in news texts, the next challenge for our approach becomes measuring, capturing and evaluating for a *diversity* of these viewpoints or ideas. This is needed for supporting a healthy democratic debate.

There is no shortage of work in recommender systems on different metrics related to diversity, from “unexpectedness” and “serendipity” to “coverage” (Zhou et al., 2010). These metrics assess the score of recommendation sets, and can be used to optimize and assess certain recommender systems on their performance beyond simply click accuracy. However, none of these beyond-accuracy metrics are informed by theories or models of democracy, and implicitly or explicitly still aim for user preference rather than a larger-scale societal goal.

An exception are the metrics in Vrijenhoek et al. (2021), who explicitly connect different diversity metrics for the evaluation and optimization of news recommender systems to goals and ideas in democratic models. Metrics from this study can be used to measure or optimize for different aspects related to diversity and different models of democracy. These metrics concern the representation of minority voices, whether the recommendations activate users to take action, and the degree of fragmentation (difference) between different users in news recommender systems. Implementing these metrics requires NLP methods. For instance, the “Activation” metric can be operationalized in terms of articles’ emotional valence and arousal, because emotional content is more likely to elicit concrete actions from readers (Vrijenhoek et al., 2021). This

requires NLP models to automatically measure whether the chosen texts contain more or less activating content than in the pool of available texts. NLP offers methods to measure sentiment and activation in text, though whether such models correctly and reliably operationalize such social science concepts has recently been questioned (van Atteveldt et al., 2021).

The metrics from Vrijenhoek et al. (2021) that measure “representation” and “alternative voices” in news texts require measuring different viewpoints and ideas, of especially marginalized groups. We run into the same challenges outlined above: the appearance of new topics, viewpoints, and opinion groups in news media. We need to further scrutinise the use and implementation of these evaluation metrics connected to models of democracy. Especially so, since consistent and nuanced evaluation metrics would help further advance recent news recommendation attempts that combine public and journalistic values like diversity with user preferences.

As highlighted above, an approach based on vector space models could aid diversification, and do so in a way that can ensure individual users are not alienated by suggestions too different from their own preferences. Such a vector space approach can do this because the idea of “distance” in a vector space. This means we can calculate relative distance between articles, viewpoints, and topics, and the optimal distance for individual users. Vector space models allow the use of similarity metrics such as cosine similarity to find (dis)similar content. This allows us to compute the distance between a user representation (based on history or personae) and news articles, and find *similar* or *dissimilar* viewpoints or opinions, such as in Reimers et al. (2019). It also means an optimal distance for individual users could be found, where “maximally distant viewpoints” could be interpreted as “(a diverse set of) different viewpoints”.

### 3.1.4 Measuring different latitudes of diversity

In our case, there is also the challenge of connecting our technical implementation for news items to the individual user’s latitude of diversity, which is again linked to our goal of supporting public diversity values and democratic debate. This aspect also has related challenges, such as the difficulty of technically identifying which news articles fall into the narrow latitude of diversity people are susceptible

to in (news)texts.

The envisioned algorithm will recommend articles within the user’s latitude of diversity, with this latitude’s width changing with user’s comfort, context, as well as interest (clickability). The model would optimize for the articles at the edges of this latitude (a maximally diverse set of viewpoint that is still within the user’s latitude of diversity).

An added bonus of such an approach is the explainability to users. Users will perhaps be able to see their specific place in the multidimensional news landscape, or adjust the values of their latitude, though this might be counter-productive for our goal of promoting engagement with viewpoints the user likes less.

### 3.2 The User

In terms of user modelling, determining users’ individual latitudes of diversity requires understanding not only what counts as diverse information to a given user, but also if and to what extent that user is open to engaging with diverse news recommendations at a given point in time (see also Section 2.4). This introduces three interrelated challenges which we address in this section.

#### 3.2.1 Data availability for user profiles

In section 3.1, we outlined several promising approaches for how NLP techniques can help represent news articles and their level of diversity. However, linking article representations to individual users also requires modelling these users’ past consumption and situational information needs. In many cases, this may necessitate the creation and maintenance of personalised user profiles that capture users’ reading histories as well as preferences of style, sources and content. However, since most news consumption takes place anonymously (Raza and Ding, 2020), session-based, and stretches across various mediums and platforms (Bruns, 2019), meaningful information for creating user profiles is often not available in the NRS domain. Thus, a first challenge in user modelling is filling in those blanks.

One way to achieve this are collaborative filtering approaches where missing data is estimated based on other users with a (seemingly) similar reading behaviour. However, this approach is limited by both the quality and quantity of user data available. It also leaves little room to capture users’ situational reading goals, which might vary considerably between reading sessions. What further

complicates the matter, is that while news consumption is often measured in terms of clicks and exposure time, in reality it includes various other reading practices (e.g. checking and scanning) that are harder to capture (Meijer and Kormelink, 2020; Costera Meijer and Groot Kormelink, 2015).

An alternative strategy could be to use *algorithmic recommender personae*, which are ”pre-configured and anthropomorphized types of recommendation algorithms” (p. 4) that users can choose from to explicitly express their preferred recommendation logic in a certain situation and for specific goals. (Harambam et al., 2018). This would grant users more control over the recommendation algorithm (Harambam et al., 2018), and allow for meaningful user modelling in the absence of personalised user profiles. However, there is a natural tension between granting users control over the type of content that they want to consume, and nudging them towards specific news selections (see also section 4).

#### 3.2.2 Individual-level differences

To maximise the likelihood of engagement with diverse news, methods taking into account individual latitudes of diversity should determine which content is acceptable for a given reader at a given point in time. Thereby, situations where introducing too much diversity limits user satisfaction (Bryanov et al., 2020) could be prevented.

In addition to the extent to which they value diverse viewpoints, users also differ in how they process them. Especially when it comes to political content, selective exposure research shows that attitudes affect information processing in various biased ways (Stroud, 2017). For example, Hart et al. (2009) show that people exhibit a moderate preference for information whose views align with their own across a variety of contexts. In contrast, counter-attitudinal information is often evaluated more critically (Taber and Lodge, 2006). Therefore, NRS users with strongly-held attitudes are likely to exhibit confirmation bias in their news selections. Moreover, selective exposure indicates potential backfire effects when users are exposed to dissimilar opinion. This includes not only decreased user satisfaction, but also increased attitude polarisation (Bail et al., 2018; Helberger and Wojcieszak, 2018; Nyhan and Reifler, 2010; Taber and Lodge, 2006).

In sum, news recommenders that aim to contribute to pro-social democratic outcomes and mitigate potential backfire effects need to accommodate

individual-level differences (see also (Rieger et al., 2020)). Modelling users' latitude of diversity is therefore an important objective of diverse NRS. To this end, news recommenders could learn from past user behaviour either implicitly, or through explicit feedback options that allow users to express when they consider an article to be too far out of their comfort zone. What remains open however, is to what extent NRS could also deliberately facilitate drift, whereby individual users become more open towards diverse viewpoints over time.

### 3.2.3 Situational differences

A further complication for user modelling stems from the fact that many news selection predictors are highly situational. Whereas attitudes and diversity values can be considered comparatively stable, news consumption is also shaped by a variety of additional situational factors (Hasebrink and Popp, 2006; Raza and Ding, 2020). For example, qualitative research shows that individual news-selections are guided by different goals that can range from general surveillance to more specific goals such as gaining new perspectives or acquiring fodder for conversation (Meijer and Kormelink, 2020).

Research into context-aware recommendation might help to better capture such differences. As of now, context-aware news recommendation is largely limited to location, time of day, or device used (Asikin and Wörndl, 2014; De Pessemier et al., 2016; Lommatzsch et al., 2017), but there have also been efforts to capture more complex constructs such as emotions (Mizgajski and Morzy, 2019). Further work into this direction could help better capture users' situational information needs. If users employ them continuously – a notion that (Harambam et al., 2019) call into question – the aforementioned personae might also be a promising way to tap into those situational differences.

## 4 Ethical considerations

### 4.1 Ethics of Nudging towards Diverse News Consumption

Thus far we have explored how the user as a human being can be put more at the center of news recommender systems by developing the idea of latitudes of diversity, which builds on NLP research and methods. However, this proposed research direction also comes with potential risks.

First, our proposed approach implies that the providers of NRS must get to know their users

better. In practice, this requires collecting (more) user data and building profiles. By doing so, NRS providers strengthen their position of power in relation to their users. This power can, of course, be used to *only* try to build better, more diverse NRS. But with this promise of user empowerment also comes an inevitable risk of user manipulation. There is a growing literature which addresses the *manipulative* potential of data-driven digital environments which try to nudge users towards certain ends or outcomes (Yeung, 2017; Lanzing, 2019; Susser et al., 2018, 2019; Sax, 2021). When digital environments use user data to learn about (patterns of behavior of) their users and run experiments which, through feedback loops, can inform subtle (personalized) tweaks to the digital environment, one is dealing with a subtle but important line between user empowerment and user manipulation. It is important to ask whose interests are being served by nudging strategies.

This question is as relevant as ever in the (online) news sector. The commercialization of the news has been discussed elaborately for decades (e.g. McManus, 2009; Baldasty, 1992; Girija, 2019) and will remain important as private platforms such as Google and Facebook try (and succeed) to capture the news industry. In such a commercialized news context, one cannot simply assume that (an increased) collection of user data and user profiling tools for purposes of personalized nudging strategies will only be used to benefit the news consumer. The very same data and profiling tools that can be used for increasing exposure to news diversity can, at the very same time, can also be *misused* in pursuit of commercial or political ends, without the knowledge of the user and/or their ability to object. The difficult line between empowerment and manipulation is underlined by the challenges news organization face in navigating the digital news economy. As a study by Bodó (2018) makes clear, different actors *within* one and the same news organization have to engage in a difficult process of mutual sense-making and negotiation to decide how a NRS should be implemented and what the NRS should aim to optimize.

The second potential risk is to reduce the news readers' autonomy. In general, it is important to note that Thaler and Sunstein's suggestion that nudging is a policy and design principle without any serious drawbacks has been met with a wide range of criticisms. Many authors point out that



nudging strategies can in fact fail to respect the autonomy of citizens (Bovens, 2009; Yeung, 2012; Saghai, 2013; Engelen and Nys, 2020). If we understand autonomy as the capacity to critically deliberate about one’s own intentions, preferences, values, and available options in order to make decisions one can consider one’s own (Sax, 2021), our nudging-inspired proposal raises questions. We do, after all, suggest to try to subtly steer news readers’ behaviors based on what is important *from a societal perspective*. Are we not thereby limiting the autonomy of the news reader? One important consideration is not only *whether* choice is influenced, but also, equally important, *how* choice is influenced. For example, when a news organization is transparent about its attempt and/or used strategies to nudge news readers, those news readers can incorporate this information in their decision-making on whether – and if so: how – to use the news platform. Being respectful of the news readers’ autonomy can thus co-exist with attempts to shape behavior for public values (Susser et al., 2019). Still, nudging strategies usually aren’t *either* fully transparent *or* completely opaque in digital environments, so questions concerning the autonomy of news readers will remain.

Lastly, there might be viewpoints that should not be recommended at all, because they are, for instance, explicitly anti-democratic or incite hate and violence. Determining which viewpoints should be excluded from recommendations, or receive a flag and/or warning for users, is challenging and requires a separate analysis. For now, we just want to flag that the existence of this difficult challenge.

#### 4.2 Ethical issues with language models

An additional consideration concerns the methods and data used to facilitate this recommendation. The role of NLP, and vector space models, in this problem is not necessarily a “plug-and-play” approach where we can take an already pre-trained model and simply plug it into our recommendation pipeline. Pre-trained language models can introduce bias, hate speech, and language not representative of real-life language use in the model by its training data based on a large, but in terms of diversity very limited set of internet texts (Bender et al., 2021). Diversity for news recommendation is therefore not only important for the recommendation output, but also for the texts in the language model input. Additionally, data practices of NLP currently

do not consist of careful consideration of the exact contents and purposes of datasets (Paullada et al., 2020), further complicating how to ensure distributional language models trained on large datasets contain diverse and representative language.

For diverse news recommendation, these data biases are important to consider. When detecting contentious topics and viewpoints in political debates, such biases potentially leading to models only detecting certain viewpoints are especially unwelcome. We do not purport to solve these issues, but we do want to highlight them.

## 5 Conclusion

In this paper, we presented an important objective for societal impact of NLP: (viewpoint) diversity in news recommendation to support a healthy democratic debate. Going further than previous work, we connect diversity in news recommendation to democratic theory and to findings in communication science on individual user differences in acceptance of diversity. We conclude that to foster a healthy democratic public debate, we should detect viewpoints, but also detect individual *latitudes of diversity*. NLP can play a pivotal role in these tasks: vector space models would allow us to place different users (or user representations) and news articles in a multidimensional space, where diversity is operationalized as distance and variance. Thereby, we could personalize different users’ latitudes of diversity, and accordingly deliver diverse recommendations that support a healthy public debate while still keeping the user satisfied. However, we also point out several technical, conceptual, and ethical problems that show this objective needs more than the “plug and play” of NLP solutions, but rather requires further research and careful reflection.

## Acknowledgments

This research is funded through Open Competition Digitalization Humanities and Social Science grant nr 406.D1.19.073 awarded by the Netherlands Organization of Scientific Research (NWO). We would like to thank the anonymous reviewers whose detailed comments helped improve the paper. All opinions and remaining errors are our own.

## References

- Emily Allaway and Kathleen McKeown. 2020. Zero-shot stance detection: A dataset and model using generalized topic representations. *EMNLP Findings 2020*.
- Yonata Andrelo Asikin and Wolfgang Wörndl. 2014. Stories around you: Location-based serendipitous recommendation of news articles. In *UMAP Workshops*. Citeseer.
- Wouter van Atteveldt, Mariken ACG van der Velden, and Mark Boukes. 2021. The validity of sentiment analysis: Comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms. *Communication Methods and Measures*, pages 1–20.
- Christopher A Bail, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37):9216–9221.
- Gerald J Baldasty. 1992. *The commercialization of news in the nineteenth century*. Univ of Wisconsin Press.
- Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 610–623.
- Balázs Bodó. 2018. Means, not an end (of the world)—the customization of news personalization by european news media. In *Prepared for the Algorithms, Automation, and News Conference, Munich*, pages 22–23.
- Luc Bovens. 2009. The ethics of nudge. In *Preference change*, pages 207–219. Springer.
- Engin Bozdag and Jeroen van den Hoven. 2015. Breaking the filter bubble: democracy and design. *Ethics and Information Technology*, 17(4):249–265.
- Axel Bruns. 2019. *Are filter bubbles real?* John Wiley & Sons.
- Kirill Bryanov, Brian K Watson, Raymond J Pingree, and Martina Santia. 2020. Effects of partisan personalization in a news portal experiment. *Public Opinion Quarterly*, 84(S1):216–235.
- Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, et al. 2018. Universal sentence encoder for english. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 169–174.
- Irene Costera Meijer and Tim Groot Kormelink. 2015. Checking, sharing, clicking and linking: Changing patterns of news use between 2004 and 2014. *Digital Journalism*, 3(5):664–679.
- Lincoln Dahlberg. 2011. Re-constructing digital democracy: An outline of four ‘positions’. *New media & society*, 13(6):855–872.
- Toon De Pessemier, Cédric Courtois, Kris Vanhecke, Kristin Van Damme, Luc Martens, and Lieven De Marez. 2016. A user-centric evaluation of context-aware recommendations for a mobile news service. *Multimedia Tools and Applications*, 75(6):3323–3351.
- Lingjia Deng and Janyce Wiebe. 2015. Mppa 3.0: An entity/event-level sentiment corpus. In *Proceedings of the 2015 conference of the North American chapter of the association for computational linguistics: human language technologies*, pages 1323–1328.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. **BERT: Pre-training of deep bidirectional transformers for language understanding**. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Tim Draws, Jody Liu, and Nava Tintarev. 2021. **Helping users discover perspectives: Enhancing opinion mining with joint topic models**. *2020 International Conference on Data Mining Workshops (ICDMW)*. Publisher: IEEE.
- Bart Engelen and Thomas Nys. 2020. Nudging and autonomy: Analyzing and alleviating the worries. *Review of Philosophy and Psychology*, 11(1):137–156.
- Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. 2018. **Framing and agenda-setting in Russian news: a computational analysis of intricate political strategies**. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3570–3580, Brussels, Belgium. Association for Computational Linguistics.
- Cristina Gena, Pierluigi Grillo, Antonio Lieto, Claudio Mattutino, and Fabiana Vernerio. 2019. **When personalization is not an option: An in-the-wild study on persuasive news recommendation**. *Information*, 10(10).
- Sreekala Girija. 2019. Political economy of media entrepreneurship: Commercialization and commodification in a digital news media enterprise. *Journal of Media Management and Entrepreneurship (JMME)*, 1(1):27–39.

- Jürgen Habermas. 2006. Political communication in media society: Does democracy still enjoy an epistemic dimension? the impact of normative theory on empirical research. *Communication theory*, 16(4):411–426.
- Andreas Hanselowski, PVS Avinesh, Benjamin Schiller, Felix Caspelherr, Debanjan Chaudhuri, Christian M Meyer, and Iryna Gurevych. 2018. A retrospective analysis of the fake news challenge stance-detection task. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1859–1874.
- Jaron Harambam, Dimitrios Bountouridis, Mykola Makhortykh, and Joris Van Hoboken. 2019. Designing for the better by taking users into account: A qualitative evaluation of user control mechanisms in (news) recommender systems. In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 69–77.
- Jaron Harambam, Natali Helberger, and Joris van Hoboken. 2018. Democratizing algorithmic news recommenders: how to materialize voice in a technologically saturated media ecosystem. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133):20180088.
- William Hart, Dolores Albarracín, Alice H Eagly, Inge Brechan, Matthew J Lindberg, and Lisa Merrill. 2009. Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin*, 135(4):555.
- Uwe Hasebrink and Jutta Popp. 2006. Media repertoires as a result of selective media use. a conceptual approach to the analysis of patterns of exposure. *Communications*, 31(3):369–387.
- Natali Helberger. 2015. Public service media—merely facilitating or actively stimulating diverse media choices? public service media at the crossroad. *International Journal of Communication*, 9:17.
- Natali Helberger. 2019. On the democratic role of news recommenders. *Digital Journalism*, 7(8):993–1012.
- Natali Helberger, Kari Karppinen, and Lucia D’acunto. 2018. Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, 21(2):191–207.
- Natali Helberger and Magdalena Wojcieszak. 2018. Exposure diversity. In Philip Michael Napoli, editor, *Mediated Communication*, volume 7, chapter 28, pages 535–560. Walter de Gruyter GmbH & Co KG.
- Mathias Jesse and Dietmar Jannach. 2021. [Digital nudging with recommender systems: Survey and future directions](#). *Computers in Human Behavior Reports*, 3:100052.
- Kari Karppinen. 2013. *Rethinking media pluralism*. Fordham University Press.
- Dam Hee Kim and Nojin Kwak. 2017. Media diversity policies for the public: Empirical evidence examining exposure diversity and democratic citizenship. *Journal of Broadcasting & Electronic Media*, 61(4):682–702.
- Dam Hee Kim and Josh Pasek. 2020. Explaining the diversity deficit: Value-trait consistency in news exposure and democratic citizenship. *Communication Research*, 47(1):29–54.
- Marjolein Lanzing. 2019. “strongly recommended” revisiting decisional privacy to judge hypernudging in self-tracking technologies. *Philosophy & Technology*, 32(3):549–568.
- Quoc Le and Tomas Mikolov. 2014. Distributed representations of sentences and documents. In *International conference on machine learning*, pages 1188–1196. PMLR.
- Ran Levy, Yonatan Bilu, Daniel Hershcovich, Ehud Aharoni, and Noam Slonim. 2014. Context dependent claim detection. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 1489–1500.
- Andreas Lommatzsch, Benjamin Kille, and Sahin Albayrak. 2017. Incorporating context and trends in news recommender systems. In *Proceedings of the international conference on web intelligence*, pages 1062–1068.
- Feng Lu, Anca Dumitrache, and David Graus. 2020. [Beyond optimizing for clicks: Incorporating editorial values in news recommendation](#). In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization, UMAP ’20*, page 145–153, New York, NY, USA. Association for Computing Machinery.
- Yiwei Luo, Dallas Card, and Dan Jurafsky. 2020. Desmog: Detecting stance in media on global warming. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings*, pages 3296–3315.
- Jörg Matthes, Johannes Knoll, Sebastián Valenzuela, David Nicolas Hopmann, and Christian Von Sikorski. 2019. A meta-analysis of the effects of cross-cutting exposure on political participation. *Political Communication*, 36(4):523–542.
- John H McManus. 2009. The commercialization of news. *The handbook of journalism studies*, pages 218–233.
- Irene Costera Meijer and Tim Groot Kormelink. 2020. *Changing News Use: Unchanged News Experiences?* Routledge.
- Jan Mizgajski and Mikołaj Morzy. 2019. Affective recommender systems in online news industry: how emotions influence reading choices. *User Modeling and User-Adapted Interaction*, 29(2):345–379.

- Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. Semeval-2016 task 6: Detecting stance in tweets. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 31–41.
- Roser Morante, Chantal Van Son, Isa Maks, and Piek Vossen. 2020. Annotating perspectives on vaccination. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 4964–4973.
- Chantal Mouffe. 2005. *The return of the political*, volume 8. Verso.
- Mats Mulder, Oana Inel, Jasper Oosterman, and Nava Tintarev. 2021. Operationalizing framing to support multiperspective recommendations of opinion pieces. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 478–488.
- Brendan Nyhan and Jason Reifler. 2010. When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2):303–330.
- Amandalynne Paullada, Inioluwa Deborah Raji, Emily M Bender, Emily Denton, and Alex Hanna. 2020. Data and its (dis) contents: A survey of dataset development and use in machine learning research.
- Thomas E Powell, Michael Hameleers, and Toni GLA van der Meer. 2020. Selection in a snapshot? the contribution of visuals to the selection and avoidance of political news in information-rich media settings. *The International Journal of Press/Politics*, page 1940161220966730.
- Shaina Raza and Chen Ding. 2020. A survey on news recommender system—dealing with timeliness, dynamic user interest and content quality, and effects of recommendation on news readers. *arXiv preprint arXiv:2009.04964*.
- Nils Reimers and Iryna Gurevych. 2019. Sentencebert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3973–3983.
- Nils Reimers, Benjamin Schiller, Tilman Beck, Johannes Daxenberger, Christian Stab, and Iryna Gurevych. 2019. Classification and clustering of arguments with contextualized word embeddings. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 567–578.
- Myrthe Reuver, Antske Fokkens, and Suzan Verberne. 2021. No nlp task should be an island: Multidisciplinarity for diversity in news recommender systems. In *Proceedings of the EACL Hackshop on News Media Content Analysis and Automated Report Generation*. Association for Computational Linguistics.
- Alisa Rieger, Mariët Theune, and Nava Tintarev. 2020. Toward natural language mitigation strategies for cognitive biases in recommender systems. In *2nd Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence 2020*. Association for Computational Linguistics (ACL).
- Shamik Roy and Dan Goldwasser. 2020. Weakly supervised learning of nuanced frames for analyzing polarization in news media. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7698–7716.
- Yashar Saghai. 2013. Salvaging the concept of nudge. *Journal of medical ethics*, 39(8):487–493.
- Marijn Sax. 2021. Optimization of what? for-profit health apps as manipulative digital environments. *Ethics and Information Technology*, pages 1–17.
- Timo Schick and Hinrich Schütze. 2020. Few-shot text generation with pattern-exploiting training. *arXiv preprint arXiv:2012.11926*.
- Chantal van Son, Tommaso Caselli, Antske Fokkens, Isa Maks, Roser Morante, Lora Aroyo, and Piek Vossen. 2016. Grasp: A multilayered annotation scheme for perspectives. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*, pages 1177–1184.
- Christian Stab, Johannes Daxenberger, Chris Stahlhut, Tristan Miller, Benjamin Schiller, Christopher Tauchmann, Steffen Eger, and Iryna Gurevych. 2018. Argumenttext: Searching for arguments in heterogeneous sources. In *Proceedings of the 2018 conference of the North American chapter of the association for computational linguistics: demonstrations*, pages 21–25.
- Christian Stab and Iryna Gurevych. 2017. Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3):619–659.
- Jesper Strömbäck. 2005. In search of a standard: Four models of democracy and their normative implications for journalism. *Journalism studies*, 6(3):331–345.
- Natalie Jomini Stroud. 2017. Selective exposure theories. In *The Oxford handbook of political communication*.
- Daniel Susser, Beate Roessler, and Helen Nissenbaum. 2018. Online manipulation: Hidden influences in a digital world. *Georgetown Law Technology Review*, 4(1):1–45.
- Daniel Susser, Beate Roessler, and Helen Nissenbaum. 2019. Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2).

- Charles S Taber and Milton Lodge. 2006. Motivated skepticism in the evaluation of political beliefs. *American journal of political science*, 50(3):755–769.
- Richard H Thaler and Cass R Sunstein. 2008. *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Nava Tintarev. 2017. Presenting diversity aware recommendations. In *FATREC Workshop on Responsible Recommendation Proceedings*.
- Francesco Saverio Varini, Jordan Boyd-Graber, Massimiliano Ciaramita, and Markus Leippold. 2020. Climatext: A dataset for climate change topic detection. In *Proceedings of the Tackling Climate Change with Machine Learning workshop, co-located at NeurIPS*.
- Sanne Vrijenhoek, Mesut Kaya, Nadia Metoui, Judith Möller, Daan Odijk, and Natali Helberger. 2021. Recommenders with a mission: assessing diversity in news recommendations. In *SIGIR Conference on Human Information Interaction and Retrieval (CHIIR) Proceedings*.
- Karen Yeung. 2012. Nudge as fudge. *Mod. L. Rev.*, 75:122.
- Karen Yeung. 2017. ‘hypernudge’: Big data as a mode of regulation by design. *Information, Communication & Society*, 20(1):118–136.
- Tao Zhou, Zoltán Kuscik, Jian-Guo Liu, Matúš Medo, Joseph Rushton Wakeling, and Yi-Cheng Zhang. 2010. Solving the apparent diversity-accuracy dilemma of recommender systems. *Proceedings of the National Academy of Sciences*, 107(10):4511–4515.