



Responsible media technology and AI: challenges and research directions

Christoph Trattner¹ · Dietmar Jannach² · Enrico Motta³ · Irene Costera Meijer⁴ · Nicholas Diakopoulos⁵ · Mehdi Elahi¹ · Andreas L. Opdahl¹ · Bjørnar Tessem¹ · Njål Borch⁶ · Morten Fjeld¹ · Lilja Øvreliid⁷ · Koenraad De Smedt¹ · Hallvard Moe¹

Received: 25 November 2021 / Accepted: 28 November 2021
© The Author(s) 2021

Abstract

The last two decades have witnessed major disruptions to the traditional media industry as a result of technological breakthroughs. New opportunities and challenges continue to arise, most recently as a result of the rapid advance and adoption of artificial intelligence technologies. On the one hand, the broad adoption of these technologies may introduce new opportunities for diversifying media offerings, fighting disinformation, and advancing data-driven journalism. On the other hand, techniques such as algorithmic content selection and user personalization can introduce risks and societal threats. The challenge of balancing these opportunities and benefits against their potential for negative impacts underscores the need for more research in responsible media technology. In this paper, we first describe the major challenges—both for societies and the media industry—that come with modern media technology. We then outline various places in the media production and dissemination chain, where research gaps exist, where better technical approaches are needed, and where technology must be designed in a way that can effectively support responsible editorial processes and principles. We argue that a comprehensive approach to research in responsible media technology, leveraging an interdisciplinary approach and a close cooperation between the media industry and academic institutions, is urgently needed.

Keywords Media technology · Artificial intelligence · Ethics

1 Introduction

The past two decades have been marked by a rapid and profound disruption of the traditional media industry. Today, the Internet is ubiquitous, practically everyone has a smartphone, the cloud reduces up front investments in large computing infrastructures, processing power still doubles every 2 years and an increasing number of our physical assets are

connected. These developments have provided the basis for new product and service innovations, which have made it possible to break up and restructure supply and demand, alter value chains and create new business models [17].

One of the most visible effects of the changes in the last decades is that media content is now largely consumed through online channels, while technological developments continue to impact how media is distributed and consumed. For instance, the increased digitization of media has opened up a variety of opportunities for collecting and analyzing large amounts of audience and consumption data, which can be used to tailor services and content to the perceived interests of individual consumers. Beyond distribution, new technological developments have opened up opportunities to enhance the media production process, such as through the use of machine learning (ML) to sift through large numbers of documents, the application of analytic tools for audience understanding, the deployment of automated media analysis capabilities, the development of sociotechnical processes to support fact-checking, and so on [21].

✉ Christoph Trattner
christoph.trattner@uib.no

¹ University of Bergen, Bergen, Norway

² University of Klagenfurt, Klagenfurt, Austria

³ The Open University, Milton Keynes, UK

⁴ Vrije Universiteit Amsterdam, Amsterdam, Netherlands

⁵ Northwestern University, Evanston, USA

⁶ NORCE, Bergen, Norway

⁷ University of Oslo, Oslo, Norway

At the same time, a number of new challenges also arise with these developments. Some of these challenges affect the industry, where media organizations have to keep up both with rapid technological developments and with new players that enter the market. However, other challenges are more societally oriented, such as the ways in which new technologies increasingly automate media personalization. One of the most pressing problems in this context is often seen in the increasing opportunities for spreading misinformation and disinformation. Whereas the former is false and misleading information not necessarily meant to deceive, the latter is intentionally created and communicated to deceive people [42]. While misinformation and disinformation have always been a feature of human society, modern technology has made it much easier for malicious actors anywhere in the world to reach the largest possible audience very quickly, something that would have been impossible in the past [5].

Overall, these challenges for industry and the potential threats to society create a need for more research in responsible media technology, which we define as technology that aims to maximize the benefits for news organizations and for society while minimizing the risks of potential negative effects. In this paper, we will first review societal and industrial challenges in Sect. 2. Afterwards, we outline a number of important research directions in responsible (AI-based) media technology in Sect. 3, covering different aspects of the media production and dissemination process. Then, in Sect. 4, we emphasize why an integrated approach is needed to address today's challenges, which not only requires the cooperation of technology experts in academia and media organizations, but also an in-depth understanding of how today's media industry operates, e.g., with respect to their editorial ethics and processes. In this context, we also introduce a new research center on responsible media technology which we have recently set up in Norway. Norway is a small, wealthy democratic nation state often described as a Nordic welfare state with high ICT penetration and comparatively egalitarian media use patterns. With a strong legacy news industry and widely used public service broadcasters, it is a case characterized by a proactive media policy operating at an arms' lengths distance, with the main aim of providing media diversity to foster public debate [67]. In this context, the research center's main goal is to foster interdisciplinary research and industry-academia co-operation, to tackle the key sociotechnical challenges relevant to the new media landscape.¹

2 Challenges for media industry and society

On the basis of the recent technological developments, this section introduces and discusses urgent challenges for the media industry and for society. Here, we give particular, but not exclusive, attention to the impact of artificial intelligence technologies.

2.1 Challenges for the media industry

A key consequence of digitalization and the new business models that have become possible is that new competition has emerged for the media industry. There are, for example, new niche players who are able to target specific user demands more accurately, thus threatening to take over positions previously held by traditional media houses and their established editorial processes. For example, finn.no has become the main platform for classified ads in Norway, a sector previously covered primarily by traditional media; Twitter has become a major debate platform, making it possible to bypass the traditional media; Facebook appears to give us far more insight into peoples' lives than the personals sections in the newspapers ever did; and Netflix, HBO, Twitch, TikTok, and YouTube challenge the positions owned by the commercial and public broadcasters in the culture and entertainment sectors.

Large platforms, such as Facebook, aggregate content and services more efficiently than the media has been able to, capitalizing on both content curation by users and algorithms for predictive content personalization. Ultimately, these large platforms now act as powerful media distribution channels, while traditional media organizations have become content providers to these platforms, almost no different than just about anyone else with a smartphone.

In this weakened position, traditional media organizations also face new threats. Presented on an equal footing, it is easy for both malicious editorial and non-editorial players to present misinformation and disinformation as news ("fake news"), which may soak up attention. As a result it is often left up to the users to find out for themselves whether or not news mirror reality. This both hurts responsible media organizations in terms of the attention they garner and at the same time underscores credibility as an important currency. To strengthen their position and maintain comparative advantage in this new competitive landscape of untrustworthy sources, responsible media entities may benefit by fortifying their role to stand out as reliable sources of information.

In the context of meeting these challenges, we suggest that advanced media technologies that are deployed in responsible ways may be a meaningful way forward for traditional media organizations. For example, such organizations

¹ MediaFutures, <https://mediafutures.no>.

are in a strong position to understand the needs of their audiences in depth and to then personalize content to these needs and preferences while trying to minimize negative effects and create public benefits [70]. Likewise, they can leverage technology to scale their ability to fact-check the morass of content circulating on platforms to buttress both their own brand credibility and to increase the overall quality of information people encounter online. In the end, the use of such technologies may not only help to keep up with the competition for attention, but may also help to meet a media organization's own goals in terms of editorial principles and ethics, including fulfilling any public service mandates.

2.2 Societal challenges

Societies and individuals may suffer in different ways from the negative effects that accompany the recent profound changes in the media landscape. For instance, the proliferation of misinformation and disinformation can threaten core democratic values by promoting political extremism and uninformed debate and discrimination [13]. Unfortunately, while there is much work on tackling these issues, e.g., through fact-checking organizations that counter disinformation, more needs to be done before they are effectively addressed [9].

As viewership and readership of linear TV and physical newspapers drop, users are going online, where they are bombarded with choices. The editorial voices which have for so long decided on what is relevant enough to publish and push, have been challenged by a combination of algorithms and user choice—creating users empowered to (or forced to) become their own editors. The world's most frequented digital media platforms, such as Google, YouTube, Facebook, Twitter, Reddit, Netflix and others, use a variety of algorithms and machine learning in elaborate sociotechnical systems, to decide which content is made visible and amplified, and which is suppressed. While understanding how such AI technology impacts public discourse to the benefit of individuals, communities, and society, traditional media will also have an interest in making technology foster democratic values [36].

There are also a multitude of concerns about the degree to which media organizations, however unintentionally, may contribute to the polarization and radicalization of the public [72]. For example an increased focus on AI-based personalization and recommendation technology could lead media organizations to contribute to the formation of so-called “echo chambers” [31]. These can potentially reduce the degree to which citizens are exposed to serendipitous information or information with which they disagree. In addition, media organizations are often concerned with freedom of speech and facilitating public debate on important societal issues. As more technologically advanced services

are created, care needs to be taken so that large groups of users are not alienated by their complexity.

A policy aspect is also present, as platforms might be held responsible for the views and statements of others. As such, content moderation will be necessary to limit distribution of harmful content (e.g., inciting, fraudulent, exploitative, hateful, or manipulative). Incoming EU legislation, such as ‘Article 13’ [58], increases the burden on media organizations that allow users to upload content. EU legislation will require media organizations to make greater efforts in checking copyright and hate speech, as media is produced, disseminated and promoted.

3 Research directions

Next, we introduce and discuss five main research areas in responsible media technology, areas we consider as priorities for research and development efforts:

1. Understanding media experiences;
2. User modeling, personalization and engagement;
3. Media content analysis and production;
4. Media content interaction and accessibility;
5. Natural language technologies.

3.1 Understanding media experiences

New developments and technological innovations are changing how news are being distributed, consumed, and experienced by users. However, we still lack knowledge on how users will interact with the media of the future, including highly personalized content [73], bots or other conversational agents [33], AI-mediated communication [35], augmented reality (AR) and virtual reality (VR), and so on. Research needs to understand to what extent the behavior and experiences of audiences can be meaningfully monitored, measured, and studied. The problem remains to develop a more substantial picture and understanding of consumers' media use across all available media and platforms, both online and offline, in high-choice media environments, and via new modalities and interfaces.

For instance, technological innovations such as news recommender systems [40] can have both positive and negative impacts on people's consumption of news, and society in general, and so it is paramount to both understand user experiences and develop designs to shape those experiences to support a well-functioning public sphere.

Research on changing media use has recognized the need to trace and analyze users across media. This is methodologically challenging and must be carefully weighed against privacy concerns, but is key to understanding how people engage with media in their daily lives [47]. With

the datafication of everyday life, increasingly powerful platforms [71] and intensified competition for attention [74], media users face a media environment which is increasingly perceived as intrusive and exploitative of their data traces [52]. This situation causes ambivalence and resignation [24] as well as immersive and joyful media experiences. A comprehensive foresight analysis of the future of media use emphasizes the need to understand fragmented, hyper-connected and individualized experiences, but also to consider the agency and capabilities of users in the context of potentially intrusive media technologies, and to develop critical and trans-media research that speaks for the interests of users in datafied communicative conditions [16]. This challenge is crucial to democracy, as media use continues to be central for public connection and to enable citizens to access information and engage fully in the societal discourse [51, 66]. Rather than predominantly making sense of media usage through quantitative metrics, such as clicks, time spent, shares or comments, critical attention to problematic representations of datafication [49, 55] should be bridged with broader and deeper understandings of media as experience [15] using a range of mixed methods approaches. In this context, responsible media innovation must build on knowledge that is attentive to diverse users' cross-media experiences and to the democratic role of media use.

The main questions in this area include the following. How will users interact with the media of the future? How can we monitor and understand users across media, including groups who leave few data traces, and user experiences beyond metrics? When do users evaluate media (organizations, platforms etc.) as responsible and how can studying user experiences feed into responsible innovation? More research is needed to answer these questions, through the design and development of novel qualitative and quantitative approaches and metrics, in combination with existing research methods for understanding audiences.

3.2 User modeling, personalization and engagement

Many modern media sites nowadays provide content personalization for their online consumers, e.g., additional news stories to read or related videos to watch [32, 39]. Such recommender systems, which typically rely both on individual user interests and collective preference patterns in a community, are commonly designed to make it easier for consumers to discover relevant content. However, the use of recommendation technology may also lead to certain undesired effects, some of which only manifest themselves over time [26].

Probably the best known example is the idea of filter bubbles [57], which may emerge when a system learns about user interests and opinions over time, and then starts to

preferentially present content that matches these assumed interests and opinions. In conjunction with user-driven selective exposure [64], this can lead to self-reinforcing feedback loops which may then result in undesired societal effects, such as opinion polarization. While stark filter bubbles are not typically observed in empirical studies [8], some more subtle self-reinforcing tendencies have been observed in real systems such as Facebook [1] and Twitter [2], raising questions about the long-term implications of more slight shifts in user exposure.

Other than the frequently discussed filter bubbles, echo chambers, as mentioned above, are another potential effect of recommendations that may lead to a polarized environment, where only certain viewpoints, information, and beliefs are shared [31] and where misinformation diffuses easily [18]. Such echo chambers are often seen as a phenomenon that is inherent to social media networks, where homogeneous and segregated communities are common. Recommender systems can reinforce such effects, e.g., by mainly providing content to users that supports the already existing beliefs in a community.

Looking beyond individual communities, recommender systems may also reinforce the promotion of content that is already generally popular, a phenomenon referred to as popularity bias. This phenomenon is well-studied in the e-commerce domain, where it was found that automated recommendations often focus more on already popular items than on promoting items from the "long tail" [29]. In the media domain, popularity biases may support the dominance of mainstream content in recommendations [69], thereby making it more difficult for consumers to discover niche or local content, and may, furthermore, have implications for the quality of content surfaced [2, 11, 27]. In addition, there is also evidence that the algorithms used by dominant content sites, such as YouTube, can drive users towards extreme content, paradoxically also on the basis of popularity biases [54].

A strong focus on already over-represented items is often considered as a situation that lacks fairness, see for example the discussion in the music domain in [48]. In general, the problem of fairness has received increased attention in recent years in the recommender systems research community. While no consistent definition of fairness is yet established and the perception of fairness can vary across consumers [63], fairness is often considered as the absence of any bias, prejudice, favoritism, mistreatment toward individuals, group, classes, or social categories based on their inherent or acquired characteristics [10]. Often, fairness and unfairness are also related to the problem of (digital) discrimination [25, 28], which is often characterized as an unfair or unequal treatment of individuals, groups, classes or social categories according to certain characteristics. Discrimination is another phenomenon, which may

be reinforced by recommender systems, in particular when they operate on data that have inherent biases. In the context of industry challenges, fairness can come up in the context of how national or local media are treated in recommendations on media platforms, with implications for how attention acquired through platforms converts to advertising or subscription revenue².

Overall, the main questions in this context are the following: To what extent can we effectively and fairly both model and predict the behavior of users accessing online media? To what extent can we personalize and engage media users online to efficiently keep them informed, and at the same time do this responsibly? In general, more research is required in the area of responsible recommender systems, which are able to generate recommendations which are designed to avoid the reinforcement of negative effects over time (such as filter bubbles or popularity biases), e.g., by striving to provide alternative viewpoints on the same issue, thus leading to fair outcomes for the media industry.

3.3 Media content analysis and production

Media content analysis and production is becoming increasingly enabled by advanced AI techniques which are used intensively for a variety of journalistic tasks, including data mining, comment moderation, news writing, story discovery, fact checking and content verification, and more [3, 21]. At the same time, the deployment of AI responsibly in the domain of news media requires close consideration of things such as how to avoid bias, how to design hybrid human-AI workflows that reflect domain values, how journalists and technologists can collaborate in interdisciplinary ways, and how future generations of practitioners should be educated to design, develop, and use AI-driven media tools responsibly [7, 20].

A crucial task that can be supported by AI technology is that of news writing. Reasonably straightforward techniques (e.g., the use of text templates filled in with data from rich databases) are already used routinely to produce highly automated stories about topics, such as sports, finance, and elections [30, 43]. Opportunities also exist for automated generation of highly personalized content, such as articles that adapt to appeal to a user's location or demographic background [73]. A challenge is to avoid *bias* in the resulting AI-automated or AI-augmented workflows, which can result both from the selection of informants and other data sources, from the analysis techniques and training materials used, and from the language models that generate the final news text [65].

There is still quite a large gap between the domain- and story-specific news generation programs currently in use and the more ambitious technologies that can be found in the field of interactive computational creativity, where users collaborate with advanced AI software for text generation [37]. Newer approaches to controlled text synthesis using large language models in conjunction with knowledge bases are on the horizon [76], but have not yet been deployed by media organizations. End-user control and the ability to “edit at scale” will be essential to ensure the accuracy, credibility, and feasibility of deploying text synthesized using such techniques in the domain of news.

Another area of news production, referred to as *computational news discovery*, leverages AI techniques to help orient journalists towards new potential stories in vast data sets [22]. Such approaches can help journalists surveil the web, identify interesting patterns or documents, and alert them when additional digging may be warranted [23]. A concern is to detect and defuse biases in what the algorithms consider newsworthy. Related techniques for representing news angles used by journalists to identify and frame newsworthy content are also under development [53, 56]. The goal of this work is to provide computational support to generate interesting new stories that match the news values and angles of interest to a particular media organization. Similar techniques can also be explored to foster news diversity by generating stories that report alternative viewpoints on the same underlying event.

An area of content analysis that has received substantial attention is in helping media detect and fight misinformation online. Multimedia forensic techniques are for example being used to uncover manipulated images and videos [14]. Moreover, automated fact checking uses machine learning and information retrieval to identify check-worthy claims, retrieve relevant evidence, classify claims, and explain decisions [68]. Research has also examined deep learning approaches to “fake news” detection [62, 77], semi-supervised machine learning techniques that analyze message streams from social media, such as Twitter [6], and the analysis of propagation patterns that can assist in differentiating fake from genuine news items [45].

Overall, the problem of computational support for responsible media production is a complex one, requiring an interdisciplinary approach and the integration of different types of technologies. Some of the main open research questions in this context include: How can we computationally produce high-quality media content that can complement traditional news production? How can the biases inherent in AI systems be managed and mitigated when producing this content? And how can we analyze user-generated content accurately to generate more valuable insights?

Correspondingly, research is required in terms of (1) novel computational methods and AI-based models to

² https://www.cjr.org/tow_center/apple-news-local-journalism.php.

generate high-quality, accurate content that is aligned with the values and standards of an editorial team, and on (2) novel algorithmic approaches for efficient media content analysis to support verification goals and content generation. In general, the integration of multimedia forensics techniques and fact checking into platforms that are used for content generation represents an important step in that direction.

The ultimate aim is then to develop sociotechnical systems that can effectively leverage AI to help produce newsworthy, interestingly-presented content that is verified, accurate, and generally adheres to the high quality standards of news media. Close collaboration with media production companies is crucial to ensure industry relevance and effective integration and testing of such methods and tools in realistic production settings.

3.4 Media content interaction and accessibility

Tomorrow's media experiences will combine smart sensors with AI and personal devices to increase engagement and collaboration [75, 79]. Enablers such as haptics, Augmented and Virtual Reality (AR/VR), conversational AI, tangible user interfaces, wearable sensors, and eye-free interactions have made clear progress. Recent work has for example studied the use of drones for various types of media production such as photography, cinematography, and film-making [46]. By employing a range of device-categories, tomorrow's media experiences will become further specialized and individualized, better targeting individuals' needs and preferences. Research into adaptation includes responsive user interfaces (UIs), adaptive streaming, content adaptation and multi-device adaptation [80]. Adaptation is also needed for collaborative and social use [34].

Another aspect of responsible media production is ensuring that users are able to understand the content. With the development of vastly more complex services and automated systems, ensuring that no user is left behind represents a major challenge. In a country like Norway, for example, 1 million people (19% of the population) have hearing disabilities, 180,000 (3%) are blind or have severely limited eye sight, 200,000 (4%) have reading disabilities, 870,000 (16%) are over 67 years, and there are about 790,000 foreign workers. While there is some overlap on these categories, it is clear that content and services designed for highly able young users will under-deliver to a substantial number of users.

To ensure usable services to all, it is not enough to just add subtitles or audio descriptions. Cognitive limitations can both be due to multitasking, age, but also due to unfamiliarity with the content, e.g., when watching an unknown sport or watching a TV series with a very large cast. It is also important to limit bias in user engagement. For example,

interactive participation may be heavily skewed towards younger users if it is non-trivial to locate or interact with a voting service.

As more content is consumed through various different media types, it can also quickly become confusing or uninteresting if the combined service is deemed inconsistent. As an example, breaking news will often report inconsistent numbers. Even a single content provider might have several different news desks, producing content for their own formats, and with some content pieces fresher than others. This makes it difficult to trust the content, and could lead to less serious platforms being preferred by users who find them more consistent and thus easier to accept.

Research should, therefore, focus on different ways to interact with content and systems, providing personal adaptations of the content to match individual needs and wishes. Partially automating processes to cater to different wishes and needs is of high importance, as is understanding how smart sensors, specialized devices and varied setups can be integrated in the experience in an inclusive and engaging manner.

3.5 Natural language technologies

The automated analysis, generation and transformation of textual content in different languages nowadays rely on Natural Language Processing (NLP) technologies. Current NLP methods are based almost exclusively on neural machine learning. Hence it is data-driven at its core, relying on large, unlabeled samples of raw text, as well as on manually annotated data sets for training of supervised ML models. NLP models are increasingly being applied to content within the news domain as well as to user-generated media content [44, 59, 60]. Newsroom analysis of textual content can assist in text classification, extraction of keywords, summarization, event extraction and other types of automated text processing. Sentiment analysis on user-generated content can be applied to monitor user attitudes, as input to recommender systems, etc. Text generation models can assist journalists through the automatic or semi-automatic production of news stories. With the widespread use of NLP-based technology in the media sector, there are a number of open challenges that must be addressed to enable responsible media technology in the years to come.

The rapid developments in the field of NLP come with important ethical considerations. Large-scale language models [19] that are built on an extensive corpus of news texts will inherit many of the same biases as its sources [4]. An example is gender bias in language models trained on large quantities of text [41], where biases have been shown to negatively affect downstream tasks [61, 78]. In NLP, biases can be found both in the data, the data annotation and the model (pre-trained input representations, fine-tuned models) [38].

Proper data documentation and curation is key to studying bias and raising awareness of it [50]. Furthermore, research on how to mitigate bias in NLP constitutes a crucial direction to enable responsible media technology [65].

Since current NLP technology is almost exclusively data driven, its quality is heavily reliant on the availability of language and domain specific data resources. Access to trusted NLP resources and tools for low-resource languages has become important not only for research but also from a democratic perspective. While NLP is a core activity in many large technology companies, their focus remains mainly on widely used languages, such as English and Chinese. The lack of task-related annotated training data and tools makes it difficult to apply novel algorithmic developments to the processing of news texts in smaller, low-resource languages and scenarios [12]. To address this challenge, a focus on data collection and annotation is important for a wide range of languages, language varieties and domains.

4 A call for interdisciplinary research

The described challenges cannot be addressed easily within a single scientific discipline or sub-discipline. On the contrary, they require the close collaboration of researchers from computer and information science (e.g., natural language processing, machine learning, recommender systems, human–computer interaction, and information retrieval) with researchers from other fields including, for example, communication sciences and journalism studies. Moreover, there are various interdependencies and cross-cutting aspects between the described research areas. Improved audience understanding, for example, can be seen as a prerequisite or input to personalized recommendation and tool-supported media production, and user modeling and personalization technology can be a basis for the synthesis of individualized and more accessible experiences.

Finally, the described research challenges cannot be reasonably addressed without a significant involvement of the relevant media industry and a corresponding knowledge transfer between academia and media organizations. To develop next-generation responsible media technology, it is of utmost importance to deeply understand the state-of-the-art, the value propositions, and the constraints under which today's diverse media industry is operating and which goals they pursue. This in particular also includes the consideration of regional or national idiosyncrasies, as well as technologies that work appropriately for languages other than English.

To address the aforementioned issues in a holistic and interdisciplinary way it is necessary to develop new organizational structures and initiatives that bring together the relevant stakeholders, knowledge, and technical

capabilities. This is why MediaFutures, a joint academia-industry research center, was founded at Media City Bergen (Norway's largest media cluster) in October 2020. The center aims to stimulate intensive collaboration between its partners and provide means to bring together the multi-disciplinary range of expertise required to tackle the long-term challenges that the media industry faces. The center will develop advanced new media technology for responsible and effective media user engagement, media content production, media content interaction and accessibility, and will research novel methods and metrics for precise audience understanding. The center will deliver a variety of research outputs, e.g., in the form of patents, prototypes, papers and software, and perform significant research training in media technology and innovation to ensure that its outputs will sustain and impact the media landscape in the long run, including the creation of start-up companies.

The center is a consortium of the most important media players in Norway. The University of Bergen's Department of Information Science and Media Studies hosts and leads the center. User partners include NRK and TV 2, the two main TV broadcasters in Norway, Schibsted, including Bergens Tidende (BT), and Amedia, the two largest news media houses in Scandinavia/Norway, as well as the world-renowned Norwegian media tech companies Vizrt, Vimond, Highsoft, Fonn Group, and the global tech and media player IBM. The center further collaborates with other national research institutions, including the University of Oslo, the University of Stavanger and NORCE, and with well-regarded international research institutions.

5 Conclusion

Rapid developments in technology have significantly disrupted the media landscape. In particular, the latest advances in AI and machine learning have created new opportunities to improve and extend the range of news coverage and services provided by media organizations. These new technologies however also come with a number of yet-unresolved challenges and societal risks, such as biased algorithms, filter bubbles and echo chambers, and massive and/or targeted spread of misinformation. In this paper, we have highlighted the need for responsible media technology and outlined a number of research directions, which will be addressed in the newly founded MediaFutures research center.

Acknowledgements This work was supported by industry partners and the Research Council of Norway with funding to MediaFutures: Research Centre for Responsible Media Technology and Innovation, through the centers for Research-based Innovation scheme, project number 309339.

Funding Open access funding provided by University of Bergen (incl Haukeland University Hospital).

Declarations

Conflicts of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bakshy, E., Messing, S., Adamic, L.A.: Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**(6239), 1130–1132 (2015). <https://doi.org/10.1126/science.aaa1160>
- Bandy, J.: Diakopoulos, Nicholas: More Accounts, fewer links: How algorithmic curation impacts media exposure in twitter timelines. *Proc. ACM on Hum.-Comput. Interact.* **5**(CSCW1), 1–28 (2021). <https://doi.org/10.1145/3449152>
- Beckett, C.: New powers, new responsibilities: A global survey of journalism and artificial intelligence. (2019). <https://blogs.lse.ac.uk/polis/2019/11/18/new-powers-new-responsibilities/>
- Bender, E.M., Gebu, T., McMillan-Major, A., Shmitchell, S.: On the dangers of stochastic parrots: Can language models be too big? *Proc. ACM Conf. Fairness Account. Transpar.* **21**, 610–623 (2021). <https://doi.org/10.1145/3442188.3445922>
- Bergstrom, C., Joseph, B.-C.: Information gerrymandering in social networks skews collective decision-making. *Nature* **573**, 40–41 (2019). <https://doi.org/10.1038/d41586-019-02562-z>
- Boididou, C., Middleton, S.E., Jin, Z., Papadopoulos, S., Dang-Nguyen, D.T., Boato, G., Kompatsiaris, Y.: Verifying information with multimedia content on twitter. *Multimed. Tools Appl.* **77**(12), 15545–15571 (2018). <https://doi.org/10.1007/s11042-017-5132-9>
- Broussard, M., Diakopoulos, N., Guzman, A.L., Abebe, R., Dupagne, M., Chuan, C.H.: Artificial intelligence and journalism: Artificial Intelligence and Journalism. *J. Mass Commun.* **96**(3), 673–695 (2019). <https://doi.org/10.1177/1077699019859901>
- Bruns, A.: *Are Filter Bubbles Real?* John Wiley and Sons, Amsterdam (2019)
- Burel, G., Farrell, T., Mensio, M., Khare, P., Alani H.: Co-spread of misinformation and fact-checking content during the COVID-19 pandemic. In *International Conference on Social Informatics*, pp. 28–42 (2020)
- Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., He, X.: Bias and debias in recommender system: A survey and future directions. *CoRR* (2020). [arXiv:2010.03240](https://arxiv.org/abs/2010.03240)
- Ciampaglia, G.L., Nematzadeh, A., Menczer, F., Flammini, A.: How algorithmic popularity bias hinders or promotes quality. *Sci. Rep.* **8**(1), 15951 (2018). <https://doi.org/10.1038/s41598-018-34203-2>
- Cieri, C., Maxwell, M., Strassel, S., Tracey, J.: Selection criteria for low resource language programs. In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation*, vol. LREC'16, pp. 4543–4549. European Language Resources Association (ELRA) (2016)
- Commission European. Communication from The Commission to The European Parliament, The Council, The European Economic and Social Committee and The Committee of the Regions - Tackling online disinformation. A European Approach (2018). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0236>
- Conotter, V., Obrien, J.F., Farid, H.: Exposing digital forgeries in ballistic motion. *IEEE Trans. Inf. Forensics Secur.* **7**, 02 (2012). <https://doi.org/10.1109/TIFS.2011.2165843>
- Costera, M.I.: Journalism, audiences and news experiences. In: Wahl-Jorgensen, K., Hanitzsch, T. (eds.) *The Handbook of Journalism Studies*. Routledge, New York (2020). <https://doi.org/10.4324/9781315167497-25>
- Das, R., Ytre-Arne, B. (eds.): *The Future of Audiences*. Palgrave Macmillan, London (2018). <https://doi.org/10.1007/978-3-319-75638-7>
- Dawson, A., Hirt, M., Scanlan, J.: The economic essentials of digital strategy. *McKinsey Q.* (2016). <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/the-economic-essentials-of-digital-strategy>
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H.E., Quattrociocchi, W.: The spreading of misinformation online. *Proc. Natl. Acad. Sci.* **113**(3), 554–559 (2016)
- Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. *Proc. Conf. N. Am. Chapter Assoc. Comput. Linguist.* **5**, 5 (2019)
- Diakopoulos, N.: Towards a design orientation on algorithms and automation in news production. *Digit. J.* **7**(8), 1180–1184 (2019). <https://doi.org/10.1080/21670811.2019.1682938>
- Diakopoulos, N.: *Automating the News: How algorithms are Rewriting the Media*. Harvard University Press, Cambridge (2019). <https://doi.org/10.4159/9780674239302>
- Diakopoulos, N.: Computational News Discovery: Towards Design Considerations for Editorial Orientation Algorithms in Journalism. *Digit. J.* **8**(7), 1–23 (2020). <https://doi.org/10.1080/21670811.2020.1736946>
- Diakopoulos, N., Trielli, D., Lee, G.: Towards understanding and supporting journalistic practices using semi-automated news discovery tools. In: *Proceedings of the ACM (PACM): Human-Computer Interaction (CSCW)*, 5 (CSCW2) (2021)
- Draper, N.A., Joseph, T.: The corporate cultivation of digital resignation. *New Media Soc.* **21**(8), 1824–1839 (2019). <https://doi.org/10.1177/1461444819833331>
- Ekstrand, M.D., Burke, R., Diaz, F.: Fairness and discrimination in recommendation and retrieval. *Proc. ACM Conf. Recomm. Syst.* (2019). <https://doi.org/10.1145/3331184.3331380>
- Elahi, M., Jannach, D., Skjærven, L., Knudsen, E., Sjøvaag, H., Tolonen, K., Holmstad, Ø., Pipkin, I., Thronsen, E., Stenbom, A., Fiskerud, E., Oesch, A., Vredenberg, L., Trattner, C.: Towards responsible media recommendation. *AI Ethics* (2021). <https://doi.org/10.1007/s43681-021-00107-7>
- Elahi, M., Kholgh, D.K., Kiarostami, M.S., Saghari, S., Rad, S.P., Tkalcic, M.: Investigating the impact of recommender systems on user-based and item-based popularity bias. *Inf. Process. Manag.* (2021). <https://doi.org/10.1016/j.ipm.2021.102655>
- Ferrer, X., van Nuenen, T., Such, J.M., Coté, M., Criado, N.: Bias and discrimination in AI: A cross-disciplinary perspective. *IEEE*

- Technol. Soc. Mag. **40**(2), 72–80 (2021). <https://doi.org/10.1109/MTS.2021.3056293>
29. Fleder, D., Hosanagar, K.: Blockbuster cultures next rise or fall: The impact of recommender systems on sales diversity. *Manag. Sci.* **55**, 697–712 (2009). <https://doi.org/10.2139/ssrn.955984>
 30. Galily, Y.: Artificial intelligence and sports journalism: Is it a sweeping change? *Technol. Soc.* (2018). <https://doi.org/10.1016/j.techsoc.2018.03.001>
 31. Ge, Y., Zhao, S., Zhou, H., Pei, C., Sun, F., Ou, W., Zhang, Y.: Understanding echo chambers in e-commerce recommender systems. *Proc. Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.* (2020). <https://doi.org/10.1145/3397271.3401431>
 32. Gomez-Uribe, C.A., Hunt, N.: The Netflix recommender system: Algorithms, business value, and innovation. *Transactions on Manag. Inf. Syst.* **6**(4), 13:1-13:19 (2015). <https://doi.org/10.1145/2843948>
 33. Gómez-Zarà, D., Diakopoulos, N.: Characterizing communication patterns between audiences and newsbots. *Digit. J.* **8**(9), 1–21 (2020). <https://doi.org/10.1080/21670811.2020.1816485>. (ISSN **2167-0811**)
 34. Hai, H.T., Dunne, M.P., Campbell, M.A., Gattton, M.L., Nguyen, H.T., Tran, N.T.: Temporal patterns and predictors of bullying roles among adolescents in Vietnam: A school-based cohort study. *Psychol. Health Med.* **22**, 107–121 (2017). <https://doi.org/10.1080/13548506.2016.1271953>
 35. Hancock, J.T., Naaman, M., Levy, K.: AI-mediated communication: Definition, research agenda, and ethical considerations. *J. Comput.-Mediat. Commun.* **25**(1), 89–100 (2020). <https://doi.org/10.1093/jcmc/zmz022>
 36. Helberger, N.: On the Democratic Role of News Recommenders. *Digit. J.* **5**(4), 1–20 (2019). <https://doi.org/10.1080/21670811.2019.1623700>
 37. Hollister, J.R., Gonzalez, A.J.: The campfire storytelling system-automatic creation and modification of a narrative. *J. Exp. Theor. Artif. Intell.* **31**(1), 15–40 (2019). <https://doi.org/10.1080/0952813X.2018.1517829>
 38. Hovy, D., Prabhumoye, S.: Five sources of bias in natural language processing. *Lang. Linguist. Compass* (2021). <https://doi.org/10.1111/lnc3.12432>
 39. Jannach, D., Jugovac, M.: Measuring the business value of recommender systems. *ACM Trans. Manag. Inf. Syst.* (2019). <https://doi.org/10.1145/3370082>
 40. Karimi, M., Jannach, D., Jugovac, M.: News recommender systems-survey and roads ahead. *Inf. Process. Manag.* **54**(6), 1203–1227 (2018). <https://doi.org/10.1016/j.ipm.2018.04.008>
 41. Kurita, K., Vyas, N., Pareek, A., Black, A.W., Tsvetkov, Y.: Measuring bias in contextualized word representations. In: *Proceedings of the 1st Workshop on Gender Bias in Natural Language Processing*, pp. 166–172 (2019)
 42. Lazer, D.M., Baum, M.A., Benkler, Y., Berinsky, A.J., Greenhill, K.M., Menczer, F., Metzger, M.J., Nyhan, B., Pennycook, G., Rothschild, D., et al.: The science of fake news. *Science* **359**(6380), 1094–1096 (2018). <https://doi.org/10.1126/science.aao2998>
 43. Leppänen, L., Munezero, M., Granroth-Wilding, M., Toivonen, H.: Data-driven news generation for automated journalism. *Proc. Int. Conf. Nat. Lang. Gener.* (2017). <https://doi.org/10.18653/v1/W17-3528>
 44. Li, C., Zhan, G., Li, Z.: News text classification based on improved Bi-LSTM-CNN. *Int. Conf. Inf. Technol. Med. Educ. (ITME)* (2018). <https://doi.org/10.1109/ITME.2018.00199>
 45. Liu, Y., Wu, Y.-F.: Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. In: *AAAI Conference on Artificial Intelligence* (2018)
 46. Ljungblad, S., Man, Y., Baytaş, M.A., Gamboa, M., Obaid, M., Field, M.: What matters in professional drone pilots' practice? An interview study to understand the complexity of their work and inform human-drone interaction research. *Proc. CHI Conf. Hum. Fact. Comput. Syst.* (2021). <https://doi.org/10.1145/3411764.3445737>
 47. Lomborg, S., Mortensen, M.: Users across media: An introduction. *Convergence* **23**(4), 343–351 (2017). <https://doi.org/10.1177/1354856517700555>
 48. Mehrotra, R., McInerney, J., Bouchard, H., Lalmas, M., Diaz, F.: Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness and satisfaction in recommendation systems. *Proc. ACM Int. Conf. Inf. Knowl. Manag.* (2018). <https://doi.org/10.1145/3269206.3272027>
 49. Milan, S., Trere, E.: Big data from the south(s): Beyond data universalism. *Telev. New Media* **20**(4), 319–335 (2019). <https://doi.org/10.1177/1527476419837739>
 50. Mitchell, M., Simone, W., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I.D., Gebru, T.: On the dangers of stochastic parrots: Can language models be too big? *Proc. ACM Conf. Fairness Account. Transpar.* (2021). <https://doi.org/10.1145/3442188.3445922>
 51. Moe, H.: Distributed readiness citizenship: A realistic, normative concept for citizens public connection. *Commun. Theory* **30**, 205–225 (2020). <https://doi.org/10.1093/ct/qtz016>
 52. Mollen, A., Dhaenens, F., Das, R., Ytre-Arne, B.: Audiences Coping Practices with Intrusive Interfaces: Researching Audiences In Algorithmic, Datafied, Platform Societies. *The Future of Audiences*. Palgrave Macmillan, London (2018). https://doi.org/10.1007/978-3-319-75638-7_3
 53. Motta, E., Daga, E., Opdahl, A.L., Tessem, B.: Analysis and design of computational News Angles. *Computer* (2020). <https://doi.org/10.1109/access.2020.3005513>
 54. Nicas, J.: How YouTube Drives People to the Internet's Darkest Corners. *Washington Post Journal*, Washington (2018)
 55. Noble, S.U.: *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press, New York (2018). <https://doi.org/10.2307/j.ctt1pwt9w5>. (ISSN **9781479849949**)
 56. Opdahl, A.L., Tessem, B.: Ontologies for finding journalistic angles. *Softw. Syst. Model.* **20**(1), 71–87 (2021). <https://doi.org/10.1007/s10270-020-00801-w>
 57. Pariser, E.: *The Filter Bubble: What the Internet Is Hiding from You*. The Penguin Group, London (2011)
 58. Parliament European. Polarisation and the use of technology in political campaigns and communication. (2019). [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634414/EPRS_STU\(2019\)634414_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634414/EPRS_STU(2019)634414_EN.pdf)
 59. Petroni, F., Raman, N., Nugent, T., Nourbakhsh, A., Panic, Z., Shah, S., Leidner, J.L.: An extensible event extraction system with cross-media event resolution. *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.* (2018). <https://doi.org/10.1145/3219819.3219827>
 60. Reuver, M., Fokkens, A., Verberne, S.: No NLP task should be an island: multi-disciplinarity for diversity in news recommender systems. *Proc. Conf. N. Am. Chapter Assoc. Comput. Linguist.* **2**, 45–55 (2021)
 61. Rüdinger, R., Naradowsky, J., Leonard, B., Van Durme, B.: Gender bias in coreference resolution. *Proc. Conf. N. Am. Chapter Assoc. Comput. Linguist.* (2018). <https://doi.org/10.18653/v1/N18-2003>
 62. Singhanian, S., Fernandez, N., Rao, S.: 3HAN: A deep neural network for fake news detection. *Neural Inf. Process.* (2017). https://doi.org/10.1007/978-3-319-70096-0_59
 63. Sonboli, N., Smith, J.J., Cabral Berenfus, F., Burke, R., Fiesler, C.: Fairness and transparency in recommendation: The users

- perspective. *Proc. ACM Conf. User Model. Adapt. Personal.* (2021). <https://doi.org/10.1145/3450613.3456835>
64. Stroud, N.: Polarization and partisan selective exposure. *J. Commun.* (2010). <https://doi.org/10.1111/j.1460-2466.2010.01497.x>
 65. Sun, T., Gaut, A., Tang, S., Huang, Y., ElSherief, M., Zhao, J., Mirza, D., Belding, E., Chang, K.W., Wang, W.Y.: Mitigating gender bias in natural language processing: Literature review. *Proc. Annu. Meet. Assoc. Comput. Linguist.* (2019). <https://doi.org/10.18653/v1/P19-1159>
 66. Swart, J., Peters, C., Broersma, M.: Repositioning news and public connection in everyday life: A user-oriented perspective on inclusiveness, engagement, relevance, and constructiveness. *Media Cult. Soc.* **39**(6), 902–918 (2017). <https://doi.org/10.1177/0163443716679034>
 67. Syvertsen, T., Enli, G., Mjos, O., Moe, M.: *Hallvard: The Media Welfare State: Nordic Media in the Digital Era.* University of Michigan Press, Ann Arbor (2014). <https://doi.org/10.3998/nmw.12367206.0001.001>
 68. Thorne, J., Vlachos, A.: Automated fact checking: Task formulations, methods and future directions. In: *Proceedings of the 27th International Conference on Computational Linguistics*, pp 3346–3359 (2018)
 69. Trielli, D., Diakopoulos, N.: Search as news curator: The role of google in shaping attention to news information. *Proc. CHI Conf. Hum. Fact. Comput. Syst.* (2019). <https://doi.org/10.1145/3290605.3300683>
 70. Van den Bluck, H., Hallvard, M.: Public service media, universality and personalization through algorithms: Mapping strategies and exploring dilemmas. *Media Cult. Soc.* **40**(6), 875–892 (2018). <https://doi.org/10.1177/0163443717734407>
 71. Van Dijck, J., Poell, T., de Waal, M.: *The Platform Society Public Values in a Connective World.* Oxford University Press, Oxford (2018). <https://doi.org/10.1093/oso/9780190889760.001.0001>
 72. van Stekelenburg, J.: Going all the way: Politicizing, polarizing, and radicalizing identity offline and online. *Sociology. Compass* **8**(5), 540–555 (2014). <https://doi.org/10.1111/soc4.12157>
 73. Wang, Y., Diakopoulos, N.: Readers perceptions of personalized news articles. In: *Proceedings Computation + Journalism Symposium* (2020)
 74. Webster, J.G.: *The Marketplace of Attention: How Audiences Take Shape in a Digital Age.* The MIT Press, London (2014). <https://doi.org/10.2307/j.ctt9qf9qj>
 75. Wozniak, A., Wessler, H., Luck, J.: Who prevails in the visual framing contest about the united nations climate change conferences? *J. Stud.* **18**(11), 1433–1452 (2017). <https://doi.org/10.1080/1461670X.2015.1131129>
 76. Xu, P., Patwary, M., Shoeybi, M., Puri, R., Fung, P., Anandkumar, A., Bryan C.: MEGATRON-CNTRL: Controllable story generation with external knowledge using large-scale language models. (2020). <https://aclanthology.org/2020.emnlp-main.226.pdf>
 77. Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., Choi, Y.: Defending against neural fake news. *Adv. Neural Inf. Process. Syst.* **32**, 9054–9065 (2019)
 78. Zhao, J., Wang, T., Yatskar, M., Ordonez, V., Chang, K.W.: Gender bias in coreference resolution: Evaluation and debiasing methods. *Proc. Conf. N. Am. Chapter Assoc. Comput. Linguist.* (2018). <https://doi.org/10.18653/v1/N18-2003>
 79. Zhu, K., Fjeld, M., Ünlüer, A.: WristOrigami: Exploring foldable design for multi-display smartwatch. *Proc. Des. Interact. Syst. Conf.* (2018). <https://doi.org/10.1145/3196709.3196713>
 80. Zorrilla, M., Borch, N., Daoust, F., Erk, A., Florez, J., Lafuente, A.: A web-based distributed architecture for multi-device adaptation in media applications. *Pers. Ubiquitous Comput.* **19**, 803–820 (2015). <https://doi.org/10.1007/s00779-015-0864-x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.