Responsible AI in Practice: A Case Study on Designing a PSM Recommender*

Maaike Harbers[1,*], Oumaima Hajri[2] and Nathalie Stembert[1]

[1] Rotterdam University of Applied Sciences, Rotterdam, The Netherlands
[2] Autoriteit Persoonsgegevens (The Dutch DPA), The Netherlands

Abstract

This paper presents a case study about the responsible design of a recommender system of a prominent Dutch Public Service Media (PSM) organization, combining personalized content recommendations to users while aligning with the organization's overarching mission of fostering diversity. A conceptual framework of diversity in news recommenders was translated into four possible prototypes of recommender systems, representing different ways to strike a balance between the objectives of personalization and diversity. These prototypes were presented to PSM stakeholders with different expertise, aiming to increase their insight into practical consequences of different conceptual choices, thus facilitating their communication and decision processes.

Keywords

Responsible AI, Recommendation system, Diversity, Personalization, Public Service Media, Design, Prototyping.

1. Introduction

The media sector is currently undergoing significant transformations due to the rise of Artificial Intelligence (AI), which is increasingly playing a pivotal role in creating and distributing media content [5]. Despite the fact that AI is providing ample room for innovation, however, it also raises concerns and questions regarding its responsible use [6]. Concerns include, e.g., the dissemination of fake news and misinformation and the resulting impact on citizenship and democracy, and bias in algorithms, leading to discrimination. To address these concerns and mitigate negative consequences of AI applications, media organizations turn to principles, tools and methods of responsible AI [7,10]. However, though there has been a lot of attention for responsible AI in the last couple of years, much of this work is rather abstract and theoretical in nature [8].

To move from theoretical contributions about responsible AI towards its practical application, this paper describes a case study that revolves around the design of a recommendation system. Recommendation systems are one of the promising and often-used applications of AI in the media sector, aiming to connect audiences with a variety of content based on e.g., their interests, search history, demographics and other contextual information [9]. Developing a recommendation system in a responsible way, however, creates a tension between accommodating user needs and interests on the one hand, and journalistic obligations and public interests on the other hand [3]. Particularly Public Service Media (PSM) organizations, funded by public money, should serve goals such as informing the public by exposing them to a balanced mix of different views and perspectives.

In the case study, we supported a prominent Dutch PSM organization in developing a recommender system, which was intended to combine and balance the objectives of personalization for users and the organization's mission of fostering diversity. One of the main challenges the PSM organization faced in developing their recommendation system was to involve different stakeholders, such as AI developers, UX designers and media professionals, in this process. Bridging the knowledge disparities among these different types of expertise is a challenging but necessary step in developing a responsible recommendation system. In this case study, we used prototypes to explore possible design directions and to facilitate communication between stakeholders with different expertise and backgrounds [1]. This exploration is particularly significant given that the majority of existing work in this domain has remained largely theoretical in nature.

2. Case study: designing four prototypes for diverse recommendation

This case study was performed in the context of a 2-year research project called DRAMA (Designing Responsible AI for Media Applications), with a consortium of multiple universities and media organizations based in The Netherlands. This specific case study involved a collaboration between one of the universities and one of the (PSM) media organizations, and was performed in the context of a redesign of the PSM organization's website. One of the aims of the redesign was to add personalized recommendation to the website, while maintaining the organization's goal to support diversity. At the start of the redesign project, there was no clear idea on how to balance diversity and personalization in the recommendation system, and one of the challenges was to include the expertise of multiple stakeholders in this process.

2.1. Conceptual framework for recommendation

The conceptual framework we used as a basis for this case study was that of Helberger [4], further elaborated on in work by Vrijenhoek et al. [11]. The conceptual framework proposed in this paper consists of the following four models of recommendation, each promoting different values and goals.

- **The liberal model**. This model promotes *autonomy*, *self-development* and *dispersion of power* by facilitating the specialization of a user in an area of his/her choosing and by tailoring to the user's preferences.
- **The participatory model.** This model promotes *inclusiveness*, *participation* and *active citizenship* by making sure that different users do not necessarily see the same content, but they do see the same topics. The recommended content's complexity is tailored to a user's preference and capability, and it reflects the prevalent voices in society.
- **The deliberative model.** This model promotes *deliberation*, *tolerance*, *open-mindedness* and *public sphere* by focusing on topics that are currently at the center of public debate, and, within those topics, presenting a plurality of voices and opinions.
- **The critical model.** This model promotes *including marginalized voices* and *defying prejudices* by emphasizing voices from marginalized groups.

2.2. Metadata available for the recommendation system

To translate the four conceptual models for recommendation from Helberger [4] into prototypes for the PSM organization, we organized a session with stakeholders of the PSM organization to collect the metadata that is available per item that could potentially be recommended, and to gain understanding in the metadata's potential relevance for recommendation. The most important results of the session are summarized below.

- **Genre.** Examples of genres are human interest, fiction, news and current affairs, sport, knowledge and education, documentaries, culture and children. This was considered highly important for personalizing recommendation.
- **Content type.** Examples of content types are playlist, series, season, promo, trailer, clip, and broadcast. After some initial discussion, this was considered important for personalization as well as diversification.
- **Broadcaster.** In The Netherlands, PSM content is developed by a number of broadcasters. Each of them has a distinctive societal, cultural or religious identity, e.g., liberal, right-conservative, left-progressive, equality of opportunity, Christian, orthodox-protestant, radical right, and inclusiveness. This was considered highly relevant from a diversity perspective but less important for users/ user personalization.
- **Language.** This refers to the language spoken in the content. All non-Dutch content is subtitled in Dutch. Stakeholders considered this somewhat relevant for diversification, as different languages represent content from different countries.
- **Release year.** Considered somewhat for diversification, as content from different periods of time can offer different perspectives.

Besides these five types of metadata, four more types of available metadata were discussed: country, duration, title, and credit. Country was considered similar to language and therefore offering less additional value. Duration was considered unimportant for recommendation. Title and credit (makers) of content were considered important for matching users' interests, but too specific to be of practical use for automatic recommendation.

2.3. Designing prototypes

Combining the conceptual recommendation models (2.1) with the metadata from the PSM organization (2.2), the authors of this paper created four distinct prototypes of recommenders. Each recommender prototype personalizes recommendations on certain metadata, i.e., offer content that matched the user's interests and needs, and offers diversity in recommendations on other metadata, i.e. offer a variety of content. Table 1 offers an overview of the different combinations of personalization and diversification for the different prototypes.

Table 1
Four recommender prototypes combining personalization (P) and diversification (D) based on different conceptual models of recommendation.

Conceptual model	Genre	Content type	Broad-caster	Lan-guage	Release year	Rationale
Liberal	P	P	P	P	P	Focus on autonomy by maximizing personalization
Participatory	D	P	P	P	D	Focus on same topics through genre and release year
Deliberative	D	P	D	P	D	Focus on current debate through genre, broadcaster and release year (recent content)
Critical	D	P	D	P	P	Focus on marginalized voices through genre and broadcaster

These prototypes should not be considered as 'the way' to operationalize the different conceptual models, but as 'best guesses' by the authors. We believe that this is not a problem as the aim of presenting the prototypes to the stakeholders is to foster discussion and decision making about combining personalization and diversification in recommendation systems.

2.4. Sharing prototypes with stakeholders

In the final step, the prototypes were shared with a group of six stakeholders, consisting of developers, content curators, and employees from the innovation department. The prototypes from Table 1 were presented as visualizations showing the diverse content outputs that could potentially result from the different recommenders. The stakeholder's repository of television programs, TV series, documentaries, and related content was utilized to showcase this potential output.

The aim of this step was to facilitate an open discussion about different (conceptual) choices to make by enabling the stakeholder to gain a clear understanding of the impact of different conceptual and metadata choices on the possible recommendations. In the discussion, the stakeholders acknowledged the value of the different conceptual models of recommendation, appreciated the effort to translate them into concrete prototypes, and stated that these prototypes gave them new insights in the challenge at hand. A variety of topics was discussed, including what they thought about the different prototypes, the lack of the decision power of the stakeholders at the session, the bureaucracy of the organization

hindering the decision-making process, the limited availability of metadata, the lack of time and resources in the project, the importance of UX design, and users' perception of the organization.

3. Discussion and conclusion

Though the stakeholders acknowledged the value of the four prototypes, the discussion in the session did not center around the different options and choices to make in the development process of the recommendation system. Thus, our intervention helped less in moving the development process forward than we intended. Yet, we believe that our observations from the session may reveal patterns of responsible AI in practice that are not unique to this specific case study and are therefore worthwhile reflecting upon.

Responsible AI involves making ethical and sometimes political choices, which is difficult and requires taking responsibility and being brave [2]. The stakeholders present in the session seemed reluctant to take this responsibility. Analyzing our observations from the session, we identify three different strategies the stakeholders used to steer the conversation away from making actual choices. The first strategy is to talk about *resources* that responsible AI require and the lack thereof, such as money, time and data. The second strategy is to discuss aspects in the *organizational governance* that hinder developing and implementing responsible AI, such as unclear responsibilities, procedures and guidelines. The third strategy for avoiding difficult topics and choices is by making use of *complexity* of the challenge at hand, by continually introducing new aspects that relate to the challenge in such a way that the conversation keeps going without moving forward.

This paper described a case study in which we supported a PSM organization in developing a recommendation system. The study shows the value of prototypes in a design process but also the complexity of responsible AI in practice. In future work, we intend to experiment more with facilitating responsible AI processes in practice, and study whether the strategies for avoiding making complex decisions can be observed in other contexts and organizations as well. If so, a next step would be to develop and evaluate interventions for overcoming these avoidance strategies.

Acknowledgements

References

[1] Camburn, B., Viswanathan, V., Linsey, J., Anderson, D., Jensen, D., Crawford, R., Otto, K., & Wood, K. (2017). Design prototyping methods: state of the art in strategies, techniques, and guidelines. Design Science, 3, e13.
[2] Gambelin, O. (2021). Brave: what it means to be an AI Ethicist. AI and Ethics, 1(1), 87-91.
[3] Helberger, N., Karppinen, K., & D'acunto, L. (2018). Exposure diversity as a design principle for recommender systems. Information, Communication & Society, 21(2), 191-207.

[4] Helberger, N. (2019). On the democratic role of news recommenders. In Algorithms, Automation, and News (pp. 14-33). Routledge.

[5] Lu, Z., & Nam, I. (2021). Research on the influence of new media technology on internet short video content production under artificial intelligence background. Complexity, 2021, 1-14.

[6] Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side' of AI. European Journal of Information Systems, 31(3), 257-268.

[7] Mioch, T., Stembert, N., Timmers, C., Hajri, O., Wiggers, P., & Harbers, M. (2023). Exploring Responsible AI Practices in Dutch Media Organizations. In IFIP Conference on Human-Computer Interaction (pp. 481-485). Cham: Springer Nature Switzerland.

[8] Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Science and engineering ethics, 26(4), 2141-2168.

[9] Sørensen, J. K. (2020). The datafication of public service media dreams, dilemmas and practical problems: A case study of the implementation of personalized recommendations at the Danish public service media 'DR'. MedieKultur: Journal of media and communication research, 36(69), 90-115.

[10] Trattner, C., Jannach, D., Motta, E., Costera Meijer, I., Diakopoulos, N., Elahi, M., Opdahl, A.L, Tessem, B., Borch, N., Fjeld, M. Ovrelid, L. De Smedt, K. & Moe, H. (2022). Responsible media technology and AI: challenges and research directions. AI and Ethics, 2(4), 585-594.

[11] Vrijenhoek, S., Kaya, M., Metoui, N., Möller, J., Odijk, D., & Helberger, N. (2021). Recommenders with a mission: assessing diversity in news recommendations. In Proceedings of the 2021 conference on human information interaction and retrieval (pp. 173-183).