

Considerations for Article-Level Personalization of News Content

Yixue Wang and Nicholas Diakopoulos

Northwestern University, School of Communication
yixue.wang@u.northwestern.edu, nad@northwestern.edu

Abstract

Although personalization in the news industry has existed for a long time, such as recommendations in the Facebook News Feed and headline personalizations in the Washington Post¹, the personalization of content itself is still an under-examined field. How will content personalization help the audience better understand news stories or promote diversity? And how can developers and journalists design algorithms to keep the balance between the ethical and privacy problems and audience interests? This position paper will introduce several key considerations related to the design of content personalization, including: localization, algorithms, interfaces, shareability and diversity.

A Thought Experiment

Imagine you are reading an article about the Syrian war on a news platform. You don't have any knowledge about the country or the civil war, and you might be losing your interest in the article. But what if the article provides you an interactive visualization of a comparison between Syrian demographics and the demographics of your state or your county, and also a paragraph comparing the Syrian refugees' situations in Europe and the refugees to the U.S. assuming you are a reader from the U.S.? You might be more interested in the article and find this background knowledge more understandable and relatable (Kim et al., 2016). However, if the article then utilizes your age or gender (e.g. by highlighting the plights of refugees of your own age and gender), or perhaps even mentions your best friend's name because they liked the article on Facebook, you might wonder if your private data has been leaked or improperly shared by other platforms and start considering whether you should stop using the platform if you are privacy sensitive. After you read the article, you might also find other articles on the platform personalize the examples and arguments in the previous article, which makes it easier for you to understand new articles' claims. However, if

the personalization algorithms are too accurate, you might find that you are trapped in filter bubbles in the platform since it would be hard for you to be exposed to different claims and examples.

This thought experiment shows how personalization can cater to an audience's interests and build common background on an unfamiliar topic, but also the tension among personalization, privacy and diversity. Because of homophily, people are more likely to understand and be more acceptable to who and what are similar to their experience (McPherson et al., 2001), i.e. language and demographics, and thus personalization provides a way for audiences to understand news in different locations or even in cases where there's a complete mismatch of interest. By creating a parallel between an individual and the article background, personalized content could make for more engaging and personally informative or memorable journalism. Previous studies have shown some guidelines about how to design content personalization, such as anticipating potential bias, providing reader control, and signalling personalization (Adar et al., 2017), but there is still little practice of content personalization just yet. Besides, privacy issues also emerged as big companies, e.g., Facebook, have faced accusations of data abuse when deploying personalization content algorithms. How to design content personalization platforms to build common background, provide diverse news and solve the privacy issues underlying the personalization?

Designing Content Personalization Platforms

Content personalization, if designed properly, can be a beneficial approach for both users and news publishers, and will ultimately change how readers interact with news articles and bring more conversations to various news topics. This section will introduce five key points when designing a content personalization platform to try to strike a balance between attracting audiences and ethical/privacy issues.

¹<https://medium.com/jsk-class-of-2018/ten-effective-ways-to-personalize-news-platform-c0e39890170e>

Use geographic localization

Localization is a specific type of personalization based on users' geographic information. Location was the most frequently mentioned property for personalizing content (Adar et al., 2017). Location can reflect much information without really asking people for their private data, such as their political leanings (i.e., red State vs. blue State, urban vs. rural areas), education, income, culture, language, etc. It is easy to know people's locations or even track people's locations by using IP address. Localization is a non-invasive way to understand people's interests and background without evoking potential privacy issues. Besides, localization makes it possible to compare global or national news with local news and bring more attention to local news topics while maintaining an interest in global or national news, and thus provide more diversity to audiences.

Identify clusters, not individuals

When building machine learning algorithms to tailor to users interests, it is possible that developers design these algorithms at an individual level because of the goal of increasing the subscription rate. However, this can raise many issues: not only might audiences and readers be afraid to use the platform due to a possible leak of private information, but the platform can also keep trapping users in the filter bubbles. To reduce the possibilities of privacy leakage and promote diversity, the algorithms used for the system should have a good performance on clusters instead of individual users. Users can be grouped into clusters by their behaviors on the platform to achieve a higher diverse context exposure to audiences. Thus, each individual in the cluster would experience less personalization but also more diversity since there are some diversities in every cluster.

Allow interaction between personalized content and users

Transparency in the news industry is perceived as an ethical rule and to create positive perceptions of journalism (Diakopoulos, 2016). To make the system transparent, it is essential to keep users updated when they are exposed to personalized content, and they should be able to change the level of the personalization by interacting with the system without leaking their privacy. The system should not store users' data unless it is allowed by users. Besides, it should be better that the users can have both visualization and text personalized for them for a better understanding. For instance, users should see the personalized sentences highlighted in some ways and can change how they want to personalize sentences by altering some inputs of the system's interface. They should also have the same ability to interact with visualizations. Ideally, users should be able to see all the possible personalization versions by interacting with the interface.

Share different content based on users' preference

When sharing articles, users should be able to share the version they received, the personalized version that will change according to new readers' attributes, and the unpersonalized version. An ideal platform should have both visualization and text personalized for audiences, and users will have the three versions sharable. If users want to share the arguments or examples that are in the unpersonalized content, they can share the unpersonalized version. If they find the exact arguments/examples in the personalized sentences are interesting, they can share the exact content they received. Or if they find the examples or arguments in the personalized content are interesting, and they wonder how others will react to these personalized content, they can also share the personalized version that will change based on every reader's attributes. Besides, users can also interact with the personalization model and set the attributes to which the shared article will be personalized. The third and fourth use case will break the traditional news article propagation through proximity/homophily. Articles may propagate longer and faster because the personalized content provides a common ground for everyone to understand the story from their perspectives and ignite interesting conversations.

Diversify audiences' experiences

Faced with the challenge of filter bubbles, it is important to design personalization algorithms to provide more diversity in frames and opinions through users' experience on one platform. Research has shown the frame difference between skeptics and acceptors in one controversial topic, including topical terms and word choices (Diakopoulos et al., 2014). By offering word choices and topical terms from both sides of a topic, it is easy for the platform to provide diverse news articles, and beneficial for audiences to have a holistic understanding of the topic and have a more balanced news digest.

References

- Adar, E., Gearig, C., Balasubramanian, A., & Hullman, J. 2017. PersaLog: Personalization of News Article Content. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems
- Diakopoulos, N., & Koliska, M. (2016). Algorithmic Transparency in the News Media. *Digital Journalism*, 5(7), 809-828. doi:10.1080/21670811.2016.1208053
- Diakopoulos, N., Zhang, A., Elgesem, D., and Salway, A. 2014. Identifying and Analyzing Moral Evaluation Frames in Climate Change Blog Discourse. *International AAAI Conference on Web and Social Media*.
- Kim, Y., Hullman, J., & Agrawala, M. 2016. Generating Personalized Spatial Analogies for Distances and Areas. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems