Designing for Personalized Article Content

Carolyn Gearig[1], Eytan Adar[1], Jessica Hullman[2]

[1]University of Michigan,
Ann Arbor, MI
{ccgearig,eadar}@umich.edu

[2]University of Washington
Seattle, WA
{jhullman}@uw.edu

ABSTRACT
While personalized feeds and collaborative filtering have become extremely popular on media websites, *content personalization* has received much less attention. Automatically modifying the textual and multimedia features of an article offers significant opportunities. Doing so has potential benefit for journalists, media sites, and readers. However, personalizing content in a news context poses numerous challenges that are not encountered in other applications like education and targeted advertising. In this paper we articulate a design space and guidelines that we have defined in the process of developing a content personalization toolset.

Author Keywords
Personalized Content, News Personalization, Guidelines

INTRODUCTION

Personalization and customization in journalism has a long history. Outlets like *The New York Times* offer personalized homepages and print editions based on a reader's location. Google News and other aggregators allow readers to customize their news feeds in a number of ways. This type of *feed personalization* addresses information overload concerns, and allows media sites to leverage archival content and create differentiated editions (e.g., local or hyper-local content). *Content Personalization*, on the other hand, where the facts presented in a *single* news 'article' are changed, has only begun to emerge as a viable feature.

Content personalization allows a site to automatically customize the *text* and *multi-media* (e.g., visualizations) in a specific article. In May 2015, for example, The New York Times published a story [1] using location data to load personalized maps and text for each reader (see Figure 1). Personalization of this type is beneficial in a number of dimensions including: allowing a journalist to write one article that can be customized for many readers; increasing engagement and learning; and supporting behavioral change.

Although there is great potential to personalized content, and a nascent practice, there is little in the way of tools or even guidance in how to develop personalization. The implementation of content personalization is difficult to scale as it is article-specific and may involve the work of many individuals (authors, editors, copy-editors, fact-checkers, programmers, graphic designers, etc.). Because personalization is often based on inferred data (e.g., the site 'believes' you are a Republican living in California), and the use of this information is specific to the article, it is necessary to develop appropriate reader 'views' for this content.

As a potential solution we have been building *PersaLog* (personalization logic), a Domain-Specific Language (DSL) for authoring and rendering personalized news content (text and graphics). As part of this work we have been identifying how content personalization differs from feed personalization and other approaches (e.g., personalized education and targeted advertising). We have created a set of guiding principles that are specifically relevant to content personalization for journalism. In this paper we share our conceptualization of a design space for content personalization and identify guidelines that we are using to drive our design.

RELATED WORK

News Personalization

Feed personalization research, which started largely to deal with information overload, has since expanded to support the leveraging of archival content, increasing engagement, and the creation of 'sub-properties' (e.g., local news through hyperlocalization) [24]. Collaborative Filtering and related approaches have been broadly studied as mechanisms for filtering, sorting, or organizing (e.g., on a custom front page) a stream of news articles [7, 12]. Subsequent research has found that feed personalization can also increase engagement (e.g., [6]). News services including the *BBC* (through *MyBBC*), *The New York Times*, *The Huffington Post* and *NPR* (through *NPR One*) as well as aggregators such as Google News have integrated this type of personalization into their online presence. Although we do not focus on feed personalization in this work, we can gain insight in how personal information is collected and used for this purpose. Specific design aims and values are visible in the way feed personalization is reflect to the reader (e.g., how does the reader understand curation? How does curation modify behavior and what are ethics of curation?).

Content personalization is much less evident in both research and practice. *The New York Times'* article (Figure 1) is a rare example [1], but one that, we believe, demonstrates an effective use of content personalization. The article dynamically adjusts visualizations (e.g., a thematic map) and article text based on geolocation (inferred by IP address). In the context of broadcast news, the *BBC* has used *object-based broadcasting*–the 'chunking' of content (e.g., video, audio, etc.)–in applications that shorten or otherwise personalize broadcasts dynamically while still retaining coherence [17]. In another creative example, *The New York Times* created a dynamic story to support active learning by personalization [2]. The interface allowed the reader to draw what they believed was the relationship between parent income and college enrollment of their children. After completing the task, the article dynamically changed to compare the reader's answer to both the true

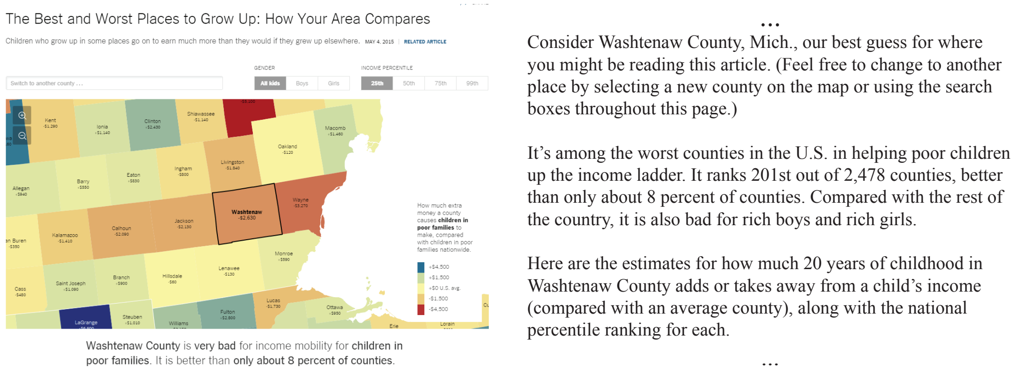

Figure 1. Sample personalization of image and graphic in the New York Times article "The Best and Worst Places to Grow Up: How Your Area Compares" (View generated from a computer in Ann Arbor, Michigan which is in Washtenaw County)

relationship as well as what others had drawn. This type of work is the inspiration for our current research. Our goal is to understand how such ideas can be generalized and scaled.

Education

Personalization in education (which, depending on application, may be called differentiated or individualized learning) has had a long history. Journalism, in its role as educating the reader, can gain insight from this literature. However, the specific goals of personalized education are somewhat unique. For example, personalized education can utilize fictional text to achieve learning objectives (e.g., mathematical word problems where generic terms are replaced by personally-relevant ones such as favorite foods, or the gender is changed). While this results in clear improvements in student engagement and performance [9, 10], fictional information is rarely appropriate in news contexts. Additionally, personalized education rarely, if it all, considers how personalization should be reflected or controlled by the student–something we believe is critical.

Patient education has also benefited from personalization (e.g., see [11]). Personalized pamphlets and mailers have supported behavioral changes in numerous contexts. However, the objectives (e.g., intervention for behavioral change) and requirements of health professionals (e.g., high-confidence in the patient's medical state and the provider's medical advice) may be different than those of the media. As a result, the necessary complexity of patient-focused authoring tools may not mesh perfectly with day-to-day journalism.

Adaptive Hypermedia

Adaptive hypermedia (surveyed extensively in [5]) has had a long history in computer science. These systems focus on sophisticated ways to modify hypertext (primarily link structure, but also text) by 'bundling' hypermedia objects based on higher-level design goals (e.g., education). The systems often leverage a 'user model' for additional personalization. The interfaces for adaptive hypermedia systems are extremely complex and may not be suited for the news context (but from which we can draw inspiration).

A related area, *Natural Language Generation* (NLG), is focused on taking structured content (e.g., a database of animal characteristics) and automatically generating unstructured text (e.g., [23]). Like adaptive hypermedia, NLG can be personalized through a user model (e.g., not describing an animal as 'piglike' if the reader has not seen the article about pigs) [20]. Narrative Science (www.narrativescience.com) is an example of NLG applied to the news. The company automatically produces news articles when given data streams. While our goal is not to fully automate the generation of personalized text, we can nonetheless use ideas from the NLG community to support the journalist in building content personalization. For example, personalization may generate a large set of article variants [18] which may be difficult to copy-edit. Automated 'repair' features of NLG [15] can test and correct personalized text with invalid grammar or style.

Advertising

Finally, we can not ignore the strong connection between news personalization and targeted advertising. The idea that advertisements should be customized for groups and individuals is fundamental to advertising practice and Internet-based ads have made the personalization more sophisticated and pervasive (see e.g., [19, 22]). We adapt some of the language and techniques of the advertising community in constructing our design space and implementation. However, as we do below, it is worth considering how the goals and values of advertising may diverge from journalism. As news organizations often directly benefit through advertising, it is critical to understand the relationship between targeted advertising (which is used on most news sites) and personalized news.

DEFINITION AND DESIGN SPACE

We define content personalization for news as: *"An automated change to the **set of facts** in an **article's content** based on **properties of the reader**."* For our purposes, *article content* can include both the textual content as well as any multimedia features (e.g., charts, static or interactive, photographs, videos, etc.). *Properties of the reader* can range from features of the individual (intrinsic or extrinsic) such as demographic or geolocation characteristics as well as behavioral features (e.g., past click behavior) or other derived features (e.g., learning style). Properties of the reader may also include preferences (e.g., preferred article length, background color, device format, etc.). In our definition, we treat each article as consisting of a

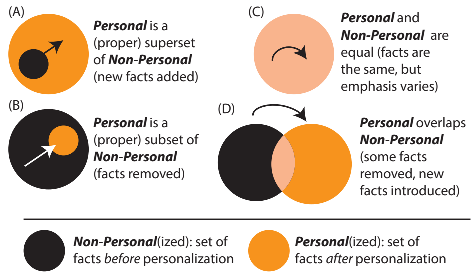

Figure 2. Changes to presented article 'facts,' pre– and post-personalization.

set of facts. These may be simple/atomic (e.g., unemployment in the the state was 5%) or complex (e.g., based on some comparison or aggregation of atomic facts), and may be reflected in text or in multimedia (e.g., bars in a bar chart).

Personalization modifies the base set of facts in the article by: *adding facts* (e.g., inserting "unemployment in *Courtland, MS* was 5.1%" if the reader is in Courtland), *removing facts* (e.g., removing: "the accident occurred 20 miles north of Courtland, MS" if the reader doesn't know where Courtland is), or *changing fact emphasis* (e.g., changing colors, modifying presentation order, etc.). These three operations can also be combined (e.g., replacing the less known "20 miles north of Courtland, MS" to the more familiar "40 miles south of Memphis, TN"). Figure 2 illustrates the configuration of facts before and after personalization.

DESIGN GUIDELINES

Below we share an initial design space and guidelines that motivate our tool-building and interface research.

Guideline 1: Interactive ≠ Personalized.

In the abstract, any interactive feature may appear to be personalization. For example, an interactive article with a search box can change an embedded visualization based on entered ZIP code (e.g., listing local unemployment rates). If the reader enters their own ZIP code (or one of interest) they are *driving* the article to display a personally-relevant view. Contrast this to the article *automatically* inferring the ZIP code from the reader's IP address and automatically adjusting the visualization. While the latter is clearly personalization, the former is less clearly so. To distinguish between interactivity and personalization, we argue that some form of automation is critical. That is, if the *only* way an article's content is changed is in reaction to the reader's *direct manipulation* or input, we do not consider this personalization (though readily acknowledge that such actions produce personally-relevant views). Interestingly, interactivity can be used as a mechanism to drive personalization in other contexts (e.g., remembering what the reader did or said on one page to drive automated personalization on another). Thus, interactive behavior is an opportunity for learning something about the user–either *explicitly* (reader typed in their salary) or *implicitly* (reader gave us their name, and we inferred their gender).

Guideline 2: Personalize with function in mind

To effectively use content personalization it is useful to identify a set of common design patterns that describe what information can be modeled (which reader properties can be obtained) and the contexts in which they can be effectively used. Our ongoing work in identifying a set of such patterns began by surveying the targeting features (or alternatively, segmentation criteria) offered in Internet advertising services (e.g., Google AdWords and Facebook). We expanded our search through the existing personalization literature (which include features such as *reading level* or *prior knowledge* that are inferred through behavior). Though a complete catalog is beyond the scope of this article, we offer three basic cases.

- *Location*–Location is the most likely target for news personalization. Many datasets feature location attributes. Because articles are often written to be relevant to broad readership, specific locations are often aggregated. Personalization can reverse this by providing local context. Furthermore, location can be coupled with census data to infer other properties (e.g., income, race).
- *Age*–Though it often possible to include gender information in articles (most datasets capture a simple male/female binary), age is more nuanced. As with location, it is not possible to include facts for *all* ages as there are many age groups. Additionally, articles are often written with the mean reader age in mind. Knowing a reader's age can support contextual personalization (e.g., explicating on unfamiliar concepts, or hiding obvious facts). One could imagine the same health or entertainment news reported very differently based on reader age.
- *Education Attained and Reading Level*–The one-size-fits-all nature of articles often encourages readers to seek alternative sources of information that are written in a more suitable style (e.g., simpler or more detailed depending on the reader). Personalization can support readers of different types. While age *may* be used as a proxy for education attainment or reading level, this information can be more directly inferred (e.g., through [8]) and used to customize article text (e.g., through simplification).

There are many additional reader properties that can be used for personalization (e.g., political affiliation and attitudes to topical interests and prior-knowledge to economic and marital status as well as many others). All can be inferred and leveraged. However, while the answer to 'can we?' for personalization is 'yes,' the answer to 'should we?' is more subtle. A key determinate in using these features is that they serves a higher goal, rather than personalization for the sake of personalization. One could: achieve learning/understanding objectives (clarifying information with personally relevant information or providing active-learning); drive behavioral change (providing personally relevant information that is know to motivate to action); or more simply increase engagement (encouraging the reader to spend more time on the article or site). The set of possible "functions of personalization" are as wide and varied as "functions of journalism." Because personalization acts to *enhance and support* goals of the journalist, articulating the goals for the specific article and context is important. Thus, we believe that a clear purpose should drive the selection of a personalization design pattern.

Part of strategic use of personalization is to have good examples of use that can be readily replicated and reused. Our current design, for example, allows for personalization code blocks to be copied from article text to article text. Code blocks were intended to be reused, with small modifications, in new contexts. The use of common patterns and personalization property names (e.g., age, county, sex, etc.) helps for rapid changes.

Guideline 3: Consider inference quality at all levels

User modeling (i.e., personal property determination) can be achieved based on passive observation (implicit information). For example, location can be inferred by IP-address [16] whereas features as distinct as age, gender, political affiliation, and reading level can be inferred by browsing or search behavior. In the case of behavior, it is usually necessary to obtain ground-truth to train the classifier (e.g., for a set of readers with known genders, which websites did they visit [21]?). For modeling reading level or political leaning of the reader, it is possible to track 'consumed' versus 'rejected' search results (that are labeled) [8]. Some inferences such as income level, which are themselves based on inferred location (using census blocks [13]), are 'precariously' constructed. Generally, the more one uses 'remote sensors' as proxies for direct evidence, the less confident one can be in the final inference.

The impact of this complexity is that inference quality can vary dramatically from 90% accuracy to under 50% in some tasks. It is critical to consider this in the use of personalization as a mistake can be costly (either in the reader's satisfaction or the 'cost' of reversing the personalization). In PersaLog we are attempting to model uncertainty directly by allowing the inference engine to record distributions (rather than most likely values). While the system will by default return the most likely inference (e.g., age=25) it also returns the certainty in this value (e.g, 80%) and alternatively the likelihood in some other value (e.g., $p(age = 45)$).

Guideline 4: Identify failure and fail gracefully

There are various ways that personalization can fail. 'Horror' stories for targeted advertising are readily available (e.g., ads for an airline appearing next to an article about a plane crash). While such failures are less likely to occur in a restricted content personalization context, they are still possible depending on the level of automation. Additionally, inference quality can vary wildly, leading to incorrect personalization.

One safeguard is to pass the level of uncertainty through to the personalization tool, and stop the personalization if inference is poor (we have low confidence). In some cases inference (and inference quality) is hierarchical. For example, geolocation by IP-address is differentially accurate based on geographical range (e.g., country, state, city, street, etc.). The larger the area, the more accurate the system (e.g., one can infer country with $> 90\%$ accuracy but this falls to $< 60\%$ for some city predictions). For inferences on hierarchical data one could set the personalization to the most narrow target that meets an uncertainty threshold.

A final possibility is to make failures apparent during design. As an extension to the traditional copy-editing and fact-checking duties may include 'stress-testing' the personalization system by creating profiles with different inference qualities and values and perform quality control on different instances. This clearly has the potential to dramatically increase the workload for the human that is 'in the loop,' so providing automated tools could help. This stress-test strategy is the one we are currently pursuing in PersaLog.

Guideline 5: Identify the bias

A common critique for feed personalization is bias introduced by automated curation. The possibility of "filter bubbles" or "echo chambers" has led to various criticisms of algorithmic curation (e.g., [4]). Because content personalization is most likely done on a per-article level, rather than systematically, and uniformly, applied to the entire site, the chance for this kind of bias is lessened. However, systematic bias may still emerge through the continuous use of the *same* personalization features in the *same* way (e.g., only providing local unemployment information in every article about unemployment) and may have negative consequences. Finally, it is worth considering how personalization might interact with other site features. For example, discussion threads may become less useful if every reader experiences a different view. This has both bias and usability issues that are worth considering as content personalization is deployed.

We note that personalization also has the potential to act *against* this type of bias. For example, *The New York Times* provided visualizations to explain the neutral, Republican, and Democratic 'read' of the same report [3]–thus allowing the reader to gain insight into the perspectives of the *'opposing'* party rather than focusing on their own.

Guideline 6: Privacy is a crucial concern

It remains to be seen if privacy issues around news personalization gain the same negative attention as Internet advertising. Regardless, we believe that in the context of news, issues of privacy can not be ignored. Even if done successfully, there is potential that readers could find it "creepy" (e.g., inferring pregnancy [14]). Additional research is necessary to determine reader attitudes for personalized content. However, it is clear that the value systems of those implementing personalization in the newsroom are quite different than than traditional users of personalization (e.g. advertisers). The emphasis on sourcing materials and provenance will likely drive a different kind of design that requires the reflection of inference mechanisms that were used. Thus, the reader's interface should support access to such information. This, of course, needs to be done with care as too much information, or information that is not interpretable, will be difficult for the reader to process.

Guideline 7: Provide reader control

Providing reader control over personalization features is desirable for a number of reasons. The ability to stop personalization may: assuage (some) privacy concerns; allow for a common view for discussion; and provide additional context that may reduce bias. A variation on this control is to provide the reader with the ability to switch to alternative personalized views (e.g., what does someone of a different gender see). We are currently designing PersaLog to allow readers to both remove and re-target personalization.

Guideline 8: Journalism workflows are unique

In many non-news applications of personalization, designers and developers act on the entire site at once (e.g., collaborative filtering is often a site-wide feature). However, the process of writing and preparing individual articles is more complex and involves (among others): authors, editors, fact-checkers, graphic designers, and copy-editors. Thus, personalization requires the attention of many stakeholders for every article–a clear problem for scaling and addressing multiple concerns. In the context of PersaLog, we have begun to survey different types of media professionals to support our goal of developing tools with role-specific authoring views.

CONCLUSIONS AND FUTURE WORK

While personalized feeds and algorithmic curation have become increasingly common in online journalism, content personalization has yet to be fully explored. Like feed personalization, content personalization can enhance learning and behavioral changes in readers and increase engagement. In our effort to design tools that support these features we have identified a number of unique properties and requirements that emerge in the context of news personalization. In this paper we describe guidelines we have identified and illustrate how they are helping us in designing our tools. We do not believe that these guidelines are complete or absolute, but that they are a useful starting point for a broader conversation. As we continue to develop, test, and deploy the PersaLog system we hope is to expand and refine these guidelines.

Acknowledgements

This work was partially supported by the NSF under grant IIS-1421438.

REFERENCES

1. G. Aisch et al. The best and worst places to grow up: How your area compares. *The New York Times*, May 4, 2015, Available: *http://nyti.ms/1KGrkJM*.
2. G. Aisch et al. You draw it: How family income predicts children's college chances. *The New York Times*, May 28, 2015, Available: *http://nyti.ms/1ezbuWY*.
3. M. Bostock et al. One report, diverging perspectives. *The New York Times*, October 5, 2012, Available: *http://nyti.ms/QzE3Cu*.
4. E. Bozdag. Bias in algorithmic filtering and personalization. *Ethics and information technology*, 15(3):209–227, 2013.
5. P. Brusilovsky. Adaptive hypermedia. *User Modeling and User-Adapted Int.*, 11(1-2):87–110, Mar. 2001.
6. G. M. Chen et al. Personalizing news websites attracts young readers. *Newspaper Research J.*, 32(4):22–38, 2011.
7. P. R. Chesnais et al. The fishwrap personalized news system. In *2nd International Workshop Community Networking, 1995*, pages 275–282. IEEE, 1995.
8. K. Collins-Thompson et al. Personalizing web search results by reading level. In *CIKM'11*, pages 403–412. ACM, 2011.
9. D. I. Cordova and M. R. Lepper. Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *J. of Educational Psychology*, 88(4):715, 1996.
10. J. Davis-Dorsey et al. The role of rewording and context personalization in the solving of mathematical word problems. *J. of Educational Psychology*, 83(1):61, 1991.
11. C. Di Marco et al. Authoring and generation of individualized patient education materials. In *AMIA Annual Symposium Proceedings*, volume 2006, page 195. American Medical Informatics Association, 2006.
12. E. Gabrilovich et al. Newsjunkie: providing personalized newsfeeds via analysis of information novelty. In *WWW'04*, pages 482–490. ACM, 2004.
13. A. T. Geronimus et al. On the validity of using census geocode characteristics to proxy individual socioeconomic characteristics. *Journal of the American Statistical Association*, 91(434):529–537, 1996.
14. K. Hill. How target figured out a teen girl was pregnant before her father did. *Forbes*, Feb. 16, 2012, Available: *http://onforb.es/1gDcytX*.
15. G. Hirst et al. Authoring and generating health-education documents that are tailored to the needs of the individual patient. In *User Modeling*, pages 107–118. Springer, 1997.
16. B. Huffaker et al. Geocompare: a comparison of public and commercial geolocation databases. *NMMC'11*, pages 1–12, 2011.
17. C. J. Hughes et al. Responsive design for personalised subtitles. In *W4A'15*, page 8. ACM, 2015.
18. S. Klein. How to edit 52,000 stories at once. *ProPublica Nerd Blog*, Jan. 24, 2013, Available: *https://goo.gl/5QipBA*.
19. Y. Liu and L. Shrum. What is interactivity and is it always such a good thing? implications of definition, person, and situation for the influence of interactivity on advertising effectiveness. *J. of Advertising*, 31(4):53–64, 2002.
20. M. Milosavljevic. Augmenting the user's knowledge via comparison. In *User Modeling*, pages 119–130. Springer, 1997.
21. D. Murray and K. Durrell. Inferring demographic attributes of anonymous internet users. In *Web Usage Analysis and User Profiling*, pages 7–20. Springer, 2000.
22. P. A. Pavlou and D. W. Stewart. Measuring the effects and effectiveness of interactive advertising: A research agenda. *J. of Interactive Advertising*, 1(1):61–77, 2000.
23. E. Reiter et al. *Building natural language generation systems*. MIT Press, 2000.
24. The New York Times. Innovation. May 24, 2015.