Bursting the Bubble: Tool for Awareness and Research about Overpersonalization in Information Access Systems

Antrea Chrysanthou
Pınar Barlas
Kyriakos Kyriakou
antreachrysanthou@gmail.com
{p.barlas},{k.kyriakou}@rise.org.cy
Research Centre on Interactive Media, Smart
Systems & Emerging Technologies (RISE Ltd.)
Nicosia, Cyprus

Styliani Kleanthous*
Jahna Otterbacher*
styliani.kleanthous@ouc.ac.cy
jahna.otterbacher@ouc.ac.cy
Cyprus Center for Algorithmic Transparency,
Open University of Cyprus
Nicosia, Cyprus

ABSTRACT

Modern information access systems extensively use personalization, automatically filtering and/or ranking content based on the user profile, to guide users to the most relevant material. However, this can also lead to unwanted effects such as the "filter bubble." We present an interactive demonstration system, designed as an educational and research tool, which imitates a search engine, personalizing the search results returned for a query based on the user's characteristics. The system can be tailored to suit any type of audience and context, as well as enabling the collection of responses and interaction data.

CCS CONCEPTS

• **Information systems** → **Search interfaces**; **Personalization**;
• **Human-centered computing** → **Interactive systems and tools**;
• **Social and professional topics** → *User characteristics.*

KEYWORDS

Information filtering, Overpersonalization

1 PROBLEM SPACE

Modern information retrieval systems have long parted ways with a "one size fits all approach." Instead, results are tailored based on the system's model of the user's characteristics. Extensive personalization of results is typically realized through an automated process of filtering and/or re-ordering of candidate results, based on user data that may be explicitly shared or implicitly inferred through
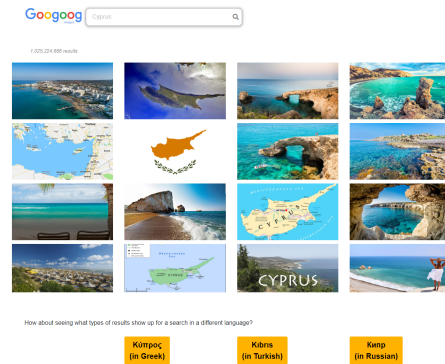
*Also with RISE Ltd., Cyprus.

Figure 1: The search results section (Page 3).

user metadata or interactions. Such personalization enhances the user experience, providing access to content that is likely more relevant to her interests and needs. However, it may also lead to over-personalization, and the undesired effect known as the "filter bubble," in which the user is over-exposed to information tailored to her perspective, and under-exposed to alternative information.

1.1 Previous work

Research has shown that many users are either largely unaware of the algorithmic filtering/reordering on the platforms they frequently use [3, 7] or may hold very different beliefs about how they work [8]. Personalization processes within information access systems, which automatically filter information based on user characteristics, can limit the exposure to diverse content and create "filter bubbles" [6]. Many tools have been created to measure, mitigate, or warn against over-personalization [1, 2, 4, 5, 9]. However, these tools often require download and/or track the user's online actions. Therefore, they are not suitable for teaching a group of people simultaneously (e.g., in a training context), and may be avoided altogether by individuals due to their user data tracking.

1.2 Our solution

We propose an interactive demonstration system, which imitates a search engine that personalizes search results based on a user model it has constructed from various types of information. The demo has an "Explicit" and an "Implicit" mode, using explicit and

implicit user modeling, respectively. The "Explicit" mode uses basic information given actively by the user, such as gender and age. The "Implicit" mode uses information inferred from the wording of the query, such as the first language of the user (implying their location, cultural background, etc.) or the adjectives/adverbs they've used (implying their goals or attitudes). In addition, the search results may be images presented in a grid, or websites presented as a list. All of these options, including the ability to add research questions at three different points in the search experience, can be modified in the administrative (admin) panel that accompanies the demo.

2 USER WALKTHROUGH & FEATURES

Page 1: Introduction In the "Explicit" mode, the information to build the user model comes from questions on this first page regarding the user's name, age, and gender. In the "Implicit" mode, none of this information is collected or required for the functioning of the demo. While the demo can be run for awareness as-is, there is also an optional "research" mode that allows research-related questions to be asked on this page as well.

Page 2: Topic selection The next page imitates the minimalist home page of a commercial search engine, which is familiar to most users given the dominance of Google in today's search engine market. The user is asked to select a query from a list, referred to as a topic in the admin panel. In the "Explicit" mode, this page also has a small badge describing the user's profile in the top right corner.

Page 3: Search results The next page loads results based on the "search query" (topic) selected by the user in the previous page. Depending on the settings in the admin panel, the results may come in the form of text (list of webpage titles, URLs, and snippets) or image (photos in a grid). The links in the text results are non-functional. Similar to the previous page, the user's own profile badge is in the top right corner in the "Explicit" mode.

In both modes, there are a series of buttons ("variations") at the bottom of the search results page. In the "Explicit" mode, these imitate the user's profile badge, but with different values. Similarly, the "Implicit" mode has buttons which have a different search query for the same topic; as an example, one topic is Cyprus and the original query is the name in English ("Cyprus") while the variations are in different languages (Greek, Turkish, and Russian). Clicking a variation modifies the user model used to filter the results, updating the text or images shown. After the user has explored at least one variation, the button to go to the Explanation page is enabled.

Page 4: Explanation On this page, there are some different explanations for the filter bubble effect, including: a short video that demonstrates the filtering and reordering process with images, an interactive section imitating the search results (allowing the user to change the characteristics of the user profile and see the changes immediately), and text with dynamic phrases that depend on the user's profile and topic. There are also some tips for understanding and managing personal information given to a platform.

In the "research" mode, it is possible to enable a page with research-related question between pages 3 and 4; this optional page allows the researcher to gather responses after the experience but before the user learns anything about the subject or goal of the demo. There is also the option to put research-related questions at the end of the Explanation page.

2.1 Administrative Panel

The administrator (researcher) is able to modify topics as well as the search results and variations for each topic/query. The researcher may also modify questions, download responses, or export interactions with the demo (such as time spent on a page or images clicked). For the user model in the "Explicit" mode, the researcher can set a location for the user that imitates the IP address, which can be identified without the user's knowledge or consent.

3 SUMMARY

We present an interactive demonstration system that emulates a popular search engine, which can be used for both educational / training and research purposes. The simple design of the system was inspired by previous research findings, indicating the lack of awareness of algorithmic filtering among "average" users of information access systems. The system may be used to explain different methods of personalization to users and to demonstrate the dangers of over-personalization. The admin panel that accompanies the demo allows the administrator to maximize the awareness by modifying the demo to fit any context and audience. In addition, the demo can simultaneously be used by researchers to conduct studies, as data collection is facilitated. In the future, we intend to use the demo to study the existing awareness of filter bubbles as well as the effectiveness of this demo in increasing comprehension and/or awareness of filter bubbles.

ACKNOWLEDGMENTS

REFERENCES

[1] Engin Bozdag and Jeroen van den Hoven. 2015. Breaking the filter bubble: democracy and design. *Ethics and Information Technology* 17, 4 (2015), 249–265.
[2] Motahhare Eslami, Amirhossein Aleyasen, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. Feedvis: A path for exploring news feed curation algorithms. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing.* ACM, 65–68.
[3] Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. I always assumed that I wasn't really that close to [her]: Reasoning about Invisible Algorithms in News Feeds. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems.* ACM, 153–162.
[4] Sean A Munson, Stephanie Y Lee, and Paul Resnick. 2013. Encouraging reading of diverse political viewpoints with a browser widget. In *Seventh International AAAI Conference on Weblogs and Social Media.*
[5] Sayooran Nagulendra and Julita Vassileva. 2014. Understanding and controlling the filter bubble through interactive visualization: a user study. In *Proceedings of the 25th ACM conference on Hypertext and social media.* ACM, 107–115.
[6] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you.* Penguin.
[7] Elia Powers. 2017. My News Feed is Filtered? Awareness of news personalization among college students. *Digital Journalism* 5, 10 (2017), 1315–1335.
[8] Emilee Rader and Rebecca Gray. 2015. Understanding user beliefs about algorithmic curation in the Facebook news feed. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems.* ACM, 173–182.
[9] Xinyu Xing, Wei Meng, Dan Doozan, Nick Feamster, Wenke Lee, and Alex C Snoeren. 2014. Exposing inconsistent web search results with bobble. In *International Conference on Passive and Active Network Measurement.* Springer, 131–140.