Investigating Users' Understanding of Invisible Algorithms and Designing around It

Motahhare Eslami, Karrie Karahalios
University of Illinois at Urbana-Champaign, Adobe Research
{eslamim2, kkarahal}@illinois.edu
{meslamim, karrie}@adobe.com

INTRODUCTION

Algorithms curate everyday online content by prioritizing, classifying, associating, and filtering information. Through this curation, they exert power to shape the users' experience and even the evolution of the system as a whole. While powerful, algorithms are usually hidden in black boxes to protect intellectual property and to hide details from users and make their interactions with the system effortless. This black box nature, however, prevents users from understanding the details of algorithmic systems' functionality or even their existence. Whether users' understanding is correct or not, their perceived knowledge about an algorithm can still affect their behavior. For instance, believing that posts with commercial keywords were ranked higher by the Facebook News Feed algorithm, some teenagers added product names to their posts in an attempt to manipulate the algorithm and increase their posts' visibility [6]. However, with no way to know if their knowledge of such invisible algorithms is correct, users cannot be sure of the results of their actions.

Algorithmic interfaces in sociotechnical systems rarely include a clear feedback mechanism for users to understand the effects of their own actions on the system. The increasing prevalence of these opaque algorithms coupled with their power raises questions about how knowledgeable users are and should be about these algorithms' "existence," "operation," and the "biases" they might introduce to users' experiences. Here, we discuss different approaches to address these issues by investigating users' understanding of and behavior around hidden algorithms in sociotechnical systems and building designs to form a more intelligent and informed interaction between users and these systems.

ALGORITHM AWARENESS

To explore users' awareness of invisible online algorithms and their effects, we conducted a user study with 40 Facebook users with diverse demographics to examine their perceptions of the Facebook News Feed curation algorithm [4]. Surprisingly, more than half of the participants (62.5%) were not aware of the fact that their News Feed was curated by a filtering algorithm at all. They rather believed that every story of their friends would appear on their News Feed. Initial reactions for these previously unaware participants were surprise and anger. To understand why the majority were not aware of the algorithm's existence, we investigated participants' Facebook usage patterns and found that the aware participants were more actively engaged with Facebook News Feed, and accordingly the News Feed curation algorithm, than the unaware ones.

Seamful Design

We developed a system, FeedVis, to incorporate some "seams" visible hints disclosing aspects of automation operations, to understand how bringing some visibility to a hidden algorithm would affect users' perception of and behavior around the algorithm. Feedvis discloses what we call "*the algorithm outputs:*"

the differences in users' News Feeds when they have been curated by the algorithm and when they have not. FeedVis highlights the content that the algorithm excluded from display and reveals social patterns by disclosing whose stories appeared and whose were hidden in their News Feed. Observing the algorithm outputs, participants were most upset when close friends and family were not shown in their feeds. We also found participants often attributed missing stories to their friends' decisions to exclude them rather than to the Facebook News Feed algorithm. By the end of the study, however, participants were mostly satisfied with the content on their feeds.

We followed up with participants two to six months later and found that for most, satisfaction level remained similar before and after becoming aware of the algorithm's presence, however, algorithmic awareness led to more active engagement with Facebook and bolstered overall feelings of control on the site. These results suggest that foregrounding algorithms may increase interface design complexity, but it may also add usability benefits.

Folk Theories: How Does the Algorithm Work?

In addition to understanding users' awareness of the algorithm's existence, we also sought to discover the folk theories that users held about how the algorithm works before, during, and after walking through our seamful design [3]. Interviews revealed 10 "folk theories" of automated curation, some quite unexpected. We found that revealing the outputs of the algorithm in a new way and incorporating intentional seams into the feed in a structured manner helped participants who were unaware of the algorithm's existence develop theories similar to participants who were aware of the algorithm's presence prior to the study. Furthermore, we found that the aware participants gained more confidence about their existing theories after viewing the algorithm's outputs. These results indicate a promising future research direction where seamful interfaces might improve algorithm understanding, building a more informed interaction between users and algorithmic systems.

ALGORITHM BIAS

Understanding users' awareness of algorithms' existence and operation is the first step toward building a more informed interaction between users and algorithmic systems. It is not enough, however. Algorithms might introduce bias to users' experience that we need to know how aware users are of them and how they behave around them. To do so, we first need to detect and quantify potential algorithmic biases. Below, we discuss bias detection and quantification on two sociotechincal systems: 1) search engines and 2) online rating platforms.

Bias in Search Engines: There has been a growing concern that search engines might favor certain results over others when ranking relevant search results. These biases can affect users behavior significantly; e.g. politically biased results of a search engine can influence the voting preferences of undecided voters in elections

by 20% or more [1]. To detect and quantify such biases, we collected thousands of search results for the names of 17 candidates of the 2016 US presidential election on Twitter Search in December 2015 during a week in which two presidential debates occurredone Republican debate and one Democratic debate. Inferring the political bias of search results and investigating different sources of bias showed that both input data and the ranking algorithm have significant contribution in creating bias to search results. We have built a bias-aware design to increase users' awareness of such potential biases [5].

We extended the analysis of political bias in search engines by comparing the search results of the same queries between Twitter Search and Google. Our analysis showed that these search engines have significantly different political biases toward the same search queries: while the political bias of search results for a candidates name on Google usually leaned toward that candidates party (e.g., the search results for Ted Cruz and Bernie Sanders would have an overall republican and democratic leaning, respectively), this was not the case for Twitter. The political bias of search results on Twitter Search, regardless of the candidates political leaning, was mostly democratic. We found that a part of this significant difference came from the difference in the fraction of search results that came from sources controlled by the candidate him or herself: In Google, a large fraction of the results for the presidential candidates are from the sources they control, i.e., either their personal websites or their social media profile links; this fraction, however, is much smaller for most candidates on Twitter. This calls for precautions when looking at the top search results of a search engine, when most of its top resources are controlled by the candidates themselves, giving them the power to influence the results.

These major differences in political bias of search engines' results show that depending on what search engine an undecided voter uses, she or he may view results with a different (or even opposing) political leaning about a candidate. This calls for new design approaches to increase users' awareness of not only potential biases in a single algorithmic system but also of the difference of biases between different algorithmic platforms- that choosing a different algorithmic system to use would affect users' online experience significantly.

Bias in Rating Platforms: We have also quantified the algorithmic bias on online rating platforms such as Booking.com as we found some anecdotal evidence suggesting a potential bias in its rating algorithm: while Booking.com's overall review interface indicates a lowest possible score of 1, the lowest output of the rating algorithm is a 2.5. That is, even if a user rates all the criteria of a hotel at the lowest value of 1, the aggregate rating returned by the algorithm is a 2.5. To understand how much bias this discrepancy might introduces to hotels' overall ratings, we have used a cross-platform audit technique that compared the ratings of more than 800 ratings across Booking.com and two other hotel rating platforms (Expedia.com and Hotels.com) and found that Booking.com's rating algorithm biased ratings of low-to-medium quality hotels 14-37% higher than others [2]. But how much are users aware of this bias?

Users' Behavior around Biased Algorithms
To understand whether there were users who were aware of Booking.com's rating bias and if so, how they communicated it, we investigated more than 2000 reviews on the website and we found 166 users who independently discovered the algorithm's bias through their regular use. Analyzing their reviews revealed that these users deviated from contributing the usual review content

(e.g., informing other users about their hotel stay experience) and rather adopted an "auditing" practice: when they confronted a higher than intended review score, they utilized their review to raise the bias awareness of other users on the site. They stated that "*the algorithm by Booking.com seems to be biased in the high direction*" (R57). While admitting that "*they can't alter the way it's calculated*" (R162), they still intervened manually to correct their biased review score. The disparity between how Booking.com calculates a user's review score and how its interface represents it, however, resulted in a trust breakdown between some users and the platform. These findings are a first step toward understanding how users manage their actions in particular circumstances when confronted by unmet system outputs, and how patterns in such behavior might inform design approaches that anticipate unexpected bias and provide reliable means for bias discovery [2].

WHAT'S NEXT? (ENOUGH) TRANSPARENCY AND TRUST
We believe our findings open up many opportunities in building "*algorithm-aware designs*:" designs that help users build a more informed and trustworthy interaction with algorithmic systems. Adding transparency to algorithmic systems can benefit both users and the systems they use. We, however, note that making a system completely transparent is usually neither possible nor desired. Algorithms are complex, dynamic, and unpredictable. Even if a designer can gain enough technical literacy to analyze an algorithm, it is often impossible to recreate the complicated and embedded internal processes of an algorithm via design. What we advocate is the study of "actionable transparency" whereby designers with knowledge of their system communicate pivotal algorithmic process cues in the interface — in some cases with features that allow for poking and prodding (such as FeedVis). We believe that participating in the "*Studying User Perceptions and Experiences with Algorithms*" Workshop would provide us helpful feedback on possible ways of achieving actionable transparency via design.

REFERENCES
1. Epstein, R., and Robertson, R. E. The search engine manipulation effect (seme) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences 112*, 33 (2015), E4512–E4521.

2. Eslami, M., Karahalios, K., and Hamilton, K. be careful; things can be worse than they appear: Understanding biased algorithms and users behavior around them in rating platforms. In *ICWSM* (2017).

3. Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., and Kirlik, A. First i like it, then i hide it: Folk theories of social feeds. In *CHI 16*, ACM (2016), 2371–2382.

4. Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., and Sandvig, C. I always assumed that i wasn't really that close to [her]: Reasoning about invisible algorithms in news feeds. In *CHI 15*, ACM (2015), 153–162.

5. Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, I. K. P., and Karahalios, K. Quantifying search bias: Investigating sources of bias for political searches in social media. In *Proc. of CSCW* (2017).

6. danah boyd: The kids are all right.
   `http://www.pbs.org/wgbh/pages/frontline/media/generation-like/danah-boyd-the-kids-are-all-right/`.