Paper to be presented at the Algorithmic Personalization and News (APEN18) Workshop at the 12th International AAAI Conference on Web and Social Media (ICWSM-18) in Stanford, CA on June 25, 2018

Auditing Online Information Curation: Closing the Gap Between the Lab and the Field

Ronald E. Robertson and Christo Wilson
Northeastern University
360 Huntington Ave.
Boston, Massachusetts 02115

Abstract

A growing body of evidence suggests that the systematic curation or manipulation of online news can distort crucial elements of our social fabric, such as how people vote in democratic elections. On one end of this research vein, social scientists have conducted lab-based experiments to identify the impacts that online platforms could have on users. On the other, computer scientists have developed methods for reverse engineering the mechanisms through which these impacts might be delivered. Yet, neither approach captures real user experiences, nor how the identified mechansims might play out in the wild, leaving a gap in our ability to understand and address the multitude of concerns raised by policy makers and researchers alike. To help bridge this gap, we outline a system for auditing online information curation systems, such as search engines and social media, that melds passive and active data collection strategies to preserve a realistic yet consistent picture of the information presented to users.

In a recent letter from the Senate's Judiciary Committee to Google, Google was asked: "How do you monitor the ability of foreign entities to influence and interfere with U.S. elections?" (Grassley 2018). While their response, if it exists, is not public at the time of this writing, researchers external to Google have published several studies aimed at monitoring and quantifying the impact that online platforms have on democratic processes. These efforts have largely revolved around two interlinked questions. First, to what extent do online platforms shape users' consumption of political information. And second, to what extent do users' consumption patterns shape their political attitudes, beliefs, and behavior?

To answer these questions, researchers have utilized two symbiotic but rarely overlapping approaches. On one end, social scientists have focused on the *impacts* of platforms by conducting controlled, laboratory-based experiments to explore how deliberate manipulations in choice architecture can shape attitudes, beliefs, and behavior (Epstein and Robertson 2015; Salganik, Dodds, and Watts 2006). On the other, computer scientists have focused on platform-generated *environments* by developing and deploying *algorithm audits* (Sandvig et al. 2014) that reverse engineer the ways that platforms shape choice architecture (Hannak et al. 2013; Eslami et al. 2015). While both approaches offer insights that are useful for policy makers, they both sacrifice the collection of real user experiences for controlled environments that offer clear statistical comparisons but limited external validity.

In this paper, we combine research from the lab and the field to outline the ways in which this gap can be bridged. Specifically, we propose a system for monitoring and preserving the patterns of information curation experienced by users in the wild. To further close the gap, we augment our system with validated methods for empowering users with an awareness of the patterns underlying the information that they are receiving, and the history of information that they have been consuming.

Auditing Search Engines. Existing search engine audits have varied widely in their reflection of real user experiences. Some researchers have relied on simple programming packages (*e.g.,* Python's `requests` module) to send search engines HTTPS requests and collect Search Engine Results Pages (SERPs) (Metaxas and Pruksachatkun 2017; Metaxa-Kakavouli and Torres-Echeverry 2017). While practical, this approach produces SERPs that differ from the SERPs produced by conducting the same search in a browser, casting doubt on what we can learn from these data. Other researchers have used a "sockpuppet" approach–manufacturing fake user accounts whose characteristics and locations can be carefully controlled–to conduct searches from a browser (Hannak et al. 2013; Kliman-Silver et al. 2015; Ballatore, Graham, and Sen 2017). Although this approach enables researchers to isolate the specific user chacteristics driving personalization, it depends on researchers emulating user behaviors that may be hard to predict, and on the search engines not being able to identify or alter their services toward the fabricated accounts.

One promising approach to capturing real user experiences is through a custom browser extension. In one recent study, this approach enabled the researchers to conduct a set of static and dynamic searches from users' computers, preserving SERPs influenced by participants actual browsing histories, cookies, and active logins (Robertson, Lazer, and Wilson 2018). However, as with prior efforts, this approach did not capture users' complex query formulations, search motivations, or browsing behavior.[1]

[1]Other search audits have offered interesting findings with respect to gender and partisan bias, but lack precise data collection details (Kay, Matuszek, and Munson 2015; Diakopoulos et al. 2018)

Auditing Search Engine Experiences. A critical limitation of the aforementioned audits is that they *actively* collected their data–the researchers decided which queries to conduct and when. Alternatively, collecting data *passively*, by monitoring participants' search behavior (*e.g.,* queries, clicks, and dwell time) over time, could offer a closer reflection of reality. However, a passive approach limits the comparability of data across participants. Therefore, an extension that combines both methods would provide the clearest picture: passively monitoring participants' searches and clicks while actively taking periodic snapshots.

The passive element would capture the SERPs produced by users' actual searches and their behavior on that SERP, revealing not only the information that participants' were exposed to, but also what information they consumed. With the exception of a few cases (*e.g.,* the AOL search logs), access to such data has historically been available only to researchers within search engine companies. This approach comes with obvious ethical concerns surrounding the sensitivity of searches that participants might conduct. These concerns could be mitigated by carefully developing a transparent list of keyword flags that are relevant to the researcher's topic (both to preserve and avoid preserving searches) and that the participant views and agrees upon prior to installation. Including a private mode in the extension, which users can select to turn off monitoring entirely (for some duration) could also help mitigate privacy concerns, though this could affect validity by increasing participant awareness of the monitoring and influencing their searches.

For the active element, the extension would take periodic snapshots of the SERPs returned for a set of researcher defined queries that are consistent across participants. Similar to prior work, these data would enable the researchers to observe the evolution of information curation over time (Diakopoulos et al. 2018), and would create a data trail that is impossible to recreate–there is no way to recover what a search engine did or did not show a given user for a given search query. Such a trail could generate algorithmic transparency for investigators, and encourage algorithmic accountability on the part of the search engines (Diakopoulos 2016).

For the data collected through both approaches, researchers will need to consider the range of recently identified search ranking components (Robertson, Lazer, and Wilson 2018). Further explorative work will also need to be regularly conducted to identify new, event-based components as they appear, such as those that appeared during the lead-up to the 2016 US Presidential campaign (Diakopoulos et al. 2018). Similarly, more behavioral experiments are needed to understand how these different components affect users.

Bridging the Gap. Spurred by concerns of the "filter bubble" (Pariser 2011) and the Search Engine Manipulation Effect (SEME) (Epstein and Robertson 2015), researchers have developed several methods that help users overcome the choice architecture through which these effects operate. The filter bubble operates through selective exposure, where people seek out information that agrees with their existing beliefs and avoid information which does not (Sears and Freedman 1967). Efforts to mitigate its effects include the gamification of browsing cross-cutting content (Munson, Lee, and Resnick 2013). SEME depends on the heuristic driven browsing of search engine users, who tend to view and believe information presented at the top of a SERP more than the information presented at the bottom (Pan et al. 2007), and thus are easily influenced by partisan *ranking bias* (Kulshrestha et al. 2017). In a recent series of experiments, alerting users to the presence of SEME during a politically-related search task substantially reduced its impact. More detailed alerts reduced the effect even further, but only a proactive prevention strategy eliminated it (Epstein et al. 2017).

In order to bridge the gap between the audit and the field, these types of mitigation or suppression strategies could be included in our proposed browser extension. With this integration, mitigation strategies could be tested on participants with random assignment as they search for and consume information in the wild. In combination with recently developed methods for scoring the partisan bias of second-level web domains (Budak, Goel, and Rao 2016; Bakshy, Messing, and Adamic 2015; Le, Shafiq, and Srinivasan 2017), these two strategies – alerts and gamification – could be combined to empower users with an awareness of the patterns underlying their information seeking, exposure, and consumption. While perhaps obvious, research has shown that such patterns can be difficult, if not impossible, for most users to detect (Epstein and Robertson 2015). Furthermore, awareness alone is not enough; only awareness and a warning that invisible influencers were at work was enough to trigger resistance to SEME for most users (Epstein et al. 2017).

Although users do not seem to be changing their behavior in response to the recent scandals around Facebook and the 2016 US election (Locklear 2018), perhaps we shouldn't expect them to. While one possible interpretation of this is that users simply do not care, another is that they do, but do not have the tools they need to change their behavior.

Discussion. With the next US elections upon us, and as more elections around the world approach and pass through the curation systems of tech giants, systems such as the one that we proposed here will be crucial in enforcing at least some semblance of algorithmic accountability (Diakopoulos 2016; Diakopoulos et al. 2018). Without the cooperation of the platforms themselves–which Facebook and Twitter have offered, but Google has not so far (Wakabayashi and Nicas 2018)–these tasks will require large-scale collaborations in which the strengths of our approaches are combined and the limitations identified and iteratively eliminated.

The system we have proposed here helps to solve both of the pressing issues we have raised: providing researchers with the tools needed to create accountability and transparency, and providing users with the tools needed to protect their psychological vulnerabilities. In combination with the panel recruitment strategies of ComScore, Nielsen, or other media measurement companies, this method could be deployed with reasonable coverage and reliability, though the costs may be high. Is defending the democratic process worth such a cost and effort? Time will tell how academia and industry answer.

References

Bakshy, E.; Messing, S.; and Adamic, L. A. 2015. Exposure to ideologically diverse news and opinion on facebook. *Science* 348(6239):1130–1132.

Ballatore, A.; Graham, M.; and Sen, S. 2017. Digital hegemonies: the localness of search engine results. *Annals of the American Association of Geographers* 1–22.

Budak, C.; Goel, S.; and Rao, J. M. 2016. Fair and balanced? quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly* 80(S1):250–271.

Diakopoulos, N.; Trielli, D.; Stark, J.; and Mussenden, S. 2018. I Vote For—How Search Informs Our Choice of Candidate. In Moore, M., and Tambini, D., eds., *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*. 22.

Diakopoulos, N. 2016. Accountability in algorithmic decision making. *Communications of the ACM* 59(2):56–62.

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

Epstein, R.; Robertson, R. E.; Lazer, D.; and Wilson, C. 2017. Suppressing the search engine manipulation effect (seme). *Proceedings of the ACM: Human-Computer Interaction* 1(42).

Eslami, M.; Aleyasen, A.; Karahalios, K.; Hamilton, K.; and Sandvig, C. 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*.

Grassley, C. E. 2018. Letter from senate judiciary committee to google.

Hannak, A.; Sapieżyński, P.; Kakhki, A. M.; Krishnamurthy, B.; Lazer, D.; Mislove, A.; and Wilson, C. 2013. Measuring personalization of web search. In *Proceedings of the 22nd International World Wide Web Conference*.

Kay, M.; Matuszek, C.; and Munson, S. A. 2015. Unequal Representation and Gender Stereotypes in Image Search Results for Occupations. 3819–3828. ACM Press.

Kliman-Silver, C.; Hannak, A.; Lazer, D.; Wilson, C.; and Mislove, A. 2015. Location, location, location: The impact of geolocation on web search personalization. In *Proceedings of the 2015 ACM Conference on Internet Measurement*.

Kulshrestha, J.; Eslami, M.; Messias, J.; Zafar, M. B.; Ghosh, S.; Gummadi, K. P.; and Karahalios, K. 2017. Quantifying search bias: Investigating sources of bias for political searches in social media. In *20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*.

Le, H.; Shafiq, Z.; and Srinivasan, P. 2017. Scalable News Slant Measurement Using Twitter. In *ICWSM*, 4.

Locklear, M. 2018. Facebook users aren't changing their privacy settings, despite uproar. https://www.engadget.com/2018/04/13/facebook-users-aren-t-changing-privacy-settings/.

Metaxa-Kakavouli, D., and Torres-Echeverry, N. 2017. Google's Role in Spreading Fake News and Misinformation. Technical report, Stanford Law School: Law and Policy Lab.

Metaxas, P. T., and Pruksachatkun, Y. 2017. Manipulation of search engine results during the 2016 us congressional elections. In *ICIW*.

Munson, S. A.; Lee, S. Y.; and Resnick, P. 2013. Encouraging reading of diverse political viewpoints with a browser widget. In *ICWSM*.

Pan, B.; Hembrooke, H.; Joachims, T.; Lorigo, L.; Gay, G.; and Granka, L. 2007. In google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication* 12(3):801–823.

Pariser, E. 2011. *The filter bubble: What the Internet is hiding from you.* Penguin UK.

Robertson, R. E.; Lazer, D.; and Wilson, C. 2018. Auditing the Personalization and Composition of Politically-Related Search Engine Results Pages. In *WWW 2018: The 2018 Web Conference*, 11. Lyon, France: ACM.

Salganik, M. J.; Dodds, P. S.; and Watts, D. J. 2006. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311(5762):854–856.

Sandvig, C.; Hamilton, K.; Karahalios, K.; and Langbort, C. 2014. Auditing algorithms: Research methods for detecting discrimination on internet platforms. In *Proceedings of "Data and Discrimination: Converting Critical Concerns into Productive Inquiry", a Productivereconference at the 64th Annual Meeting of the International Communication Association*.

Sears, D. O., and Freedman, J. L. 1967. Selective Exposure to Information: A Critical Review. *Public Opinion Quarterly* 31(2):194.

Wakabayashi, D., and Nicas, J. 2018. Facebook Takes the Punches While Rest of Silicon Valley Ducks. *The New York Times*.