

Art Thou Target or Art Thou Waste? A Collective Audit of Online Advertising Placements
Darren Stevenson

Submitted to:
Auditing Algorithms From the Outside: Methods and Implications
ICWSM '15 Workshop, Oxford, UK, May 26, 2015

Research problem/scenario

When using the web, we are routinely shown an array of digital advertisements. Often, these ads are the result of strategic ad buys based on some degree of audience segmentation using known or estimated information about us (e.g., our demographics, location, online behaviors, composition of social networks). This information provides inputs to ad placement algorithms that ultimately determine who sees which ad when.¹ The intent is for advertising content to reach certain target audiences but not others. For marketers, this process of segmentation, targeting, and advertising personalization reduces expenditures; it can be more efficient compared to traditional “broadcasted” advertisements.² Yet some have criticized this practice on the grounds of its potential for negative viewer outcomes, benefiting some individuals while harming others. In the aggregate, it is theorized that advertising personalization might hold certain negative cumulative effects. For example, maybe we start to see ourselves as “the type of people who drive used budget automobiles” (or minivans, etc.), whatever that may mean, after repeatedly seeing income-based targeted ads for cheap used cars.

As Turow writes, “Marketers are increasingly using databases to determine whether to consider particular Americans to be targets or waste. Those considered waste are ignored or shunted to other products the marketers deem more relevant to their tastes or income” (p. 88, emphasis added).³

But how do we know when we are targets and when we are waste? Of course, the question of targets and waste always depends on the goals of the particular advertising campaign in question. Still, would peering into these divisions (collectively and in real-time) open up new ways of considering and understanding the algorithms that govern today’s online advertising channels?

Suggested intervention/solution/amelioration/clarification

Motivated by these concerns and curiosities over targets and waste, this suggested intervention calls for a collective audit of a series of pre-determined websites in efforts to locate correlations (qualitatively, anecdotally) between viewer traits and ad content, with some understanding that ad-buying/targeting algorithms are responsible for these outcomes.

Procedures:

Participants would first become familiar with the built-in screen capture functions (still images, not video) of their laptop’s operating systems. If no built-in screen capture function is available, a third-party screen capture tool will need to be installed, or the participant can audit without supplying screen captures.⁴

¹ Consider bid optimization algorithms that determine how much an advertiser bids for a particular ad placement in real-time ad auction systems, the results of which determine who sees which ad when. For instance, see Chen, Y., Berkhin, P., Anderson, B., & Devanur, N. R. (2011, August). Real-time bidding algorithms for performance-based display ad allocation. In Proceedings of the 17th ACM SIGKDD (pp. 1307-1315).

² All advertising is targeted, to some degree, even that on traditional broadcast media channels where marketers go to great efforts to anticipate and target likely audiences/viewers. However, the affordances of online advertising systems and algorithms (e.g., person-specific messaging) differentiate targeted online advertising from targeted broadcast advertising.

³ Turow, J. (2011). *The Daily You: How the New Advertising Industry is Defining Your Identity and Your Worth*. New Haven: Yale Univ. Press.

⁴ Most operating systems will have a built-in screen capture tool and it is not anticipated that this requirement would prevent participants from the audit exercise.

In synchrony, participants visit a short pre-determined list of websites (all participants testing the same sites at the same time). Perhaps 5 to 10 sites would do, depending on logistics and time, and auditing sites known to use advertising networks to supply their display ads (e.g., merriam-webster.com, answers.yahoo.com, wsj.com). See Fig. 1 below. A moderator will step the participants through the audit, keeping the group in sync on each site.

If the physical room allows, viewers who are comfortable doing so might sit in a circle facing outward, so that each person can stand up and turn around to see the laptop screen of every other person participating in the audit. That way a comparison can be made between the advertisement you see and what others see, when visiting the same site at the same time (and from the same location/IP address, depending on the conference WiFi configuration).

In addition to real-time viewing of others' screens, as each site is visited, participants are to take a screenshot of each page, capturing both the ad(s) and the site it appears on. These images could be uploaded to a collection from the audit. This collection could be analyzed further after the event, given participant anonymity and consent to do so. This screen capture idea is not essential to the experience and could be omitted if it is thought to slow down the workshop and/or not contribute substantially.

Finally, participants might complete an optional demographic survey, to be included in any post-hoc analysis/audit that accompanies the collected screenshots. While this is also optional, it could provide an interesting follow-up paper documenting and describing the merits of this collective auditing event.

As with most activities that use individuals' personal computers, the audit suggested here has the minor potential to be embarrassing or even worse.⁵ While this is unlikely, perhaps this minor potential could be emphasized to anyone wishing to sit in the circle whose screen could be viewed by other participants.

Three goals for discussion at the workshop

1. Consideration of what I will call the "Turow Indictment" of personal data- and algorithm-driven targeted online advertising. That is, Turow argues that these type of ads have negative impacts on Internet users because they change the way we see ourselves, others, and the world. The reason he views this as negative is because content is determined by marketers, who do not always have the best interests of consumers in mind. Perhaps we could consider: what are the holes in the Turow Indictment? Maybe targeted ads change our personal views and worldviews for the better? Maybe advertising algorithms, rather unintentionally, have the power to cumulatively shape us in positive ways? If they don't now, could they in the future?

2. This is largely assumed given the likely ICWSM audience, but, as knowledge of online content curation can vary considerably,⁶ it might be beneficial for participants to simply consider that the content they see on websites and apps frequently differs from what others see on the same websites and apps, especially when it comes to advertising content. A related discussion could proceed debating how important this type of content curation / algorithmic awareness is for the general public? Is it important?

3. When algorithms result in negative outcomes (not limited to the case of targeted ads), we could consider issues of intent, responsibility, and culpability for 1) the developers (programmers) who create algorithms and also 2) those who do not design algorithms but put them into practice (e.g. advertisers).

⁵ For instance, see *No One Should Be Outed By an Ad* by Marc Groman (Feb. 24, 2015), available at <https://privacyassociation.org/news/a/nai-takes-lgbt-stand/>

⁶ For instance, see Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., and Sandvig, C. (2015). "I always assumed that I wasn't really that close to [her]:" Reasoning about invisible algorithms in the news feed. Proceedings of the 33rd Annual SIGCHI Conference on Human Factors in Computing Systems, Association for Computing Machinery (ACM).

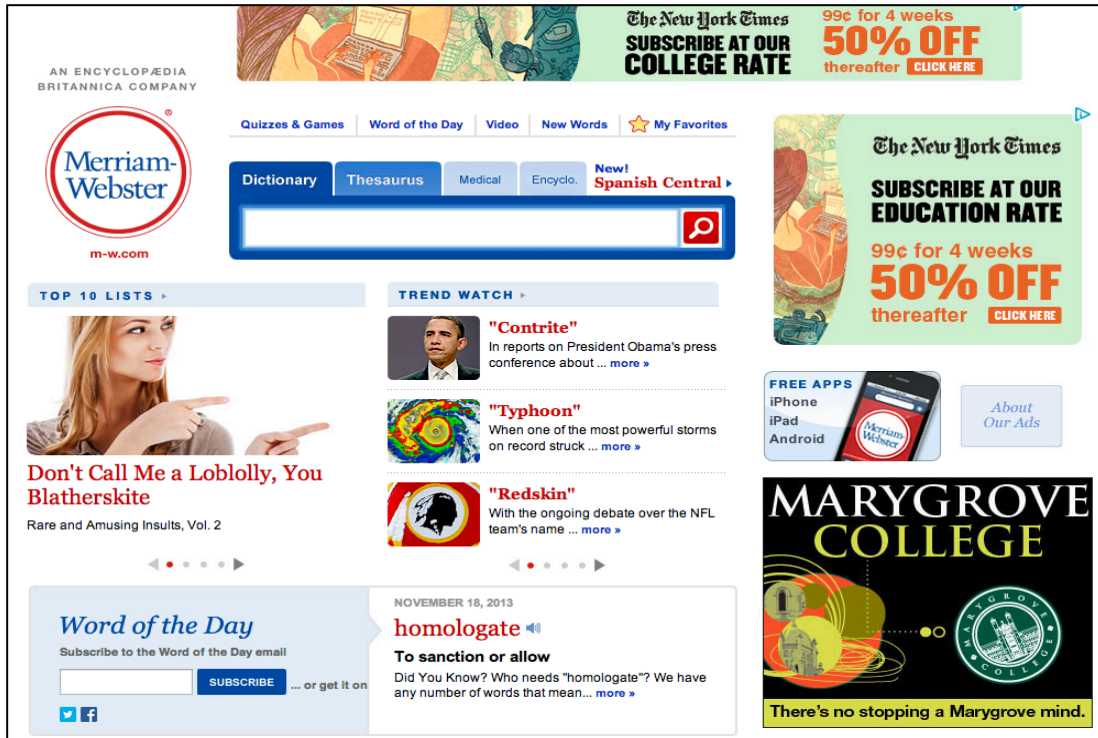


Fig. 1. Screenshot of Merriam-Webster.com showing personalized online advertisements in context of the site.

In Fig. 1, two of the ads are related to purchasing a *The New York Times* subscription, appearing at the top (banner) and side (top gutter) positions of the page. The personalized *New York Times* ads are likely due to some personal data linked to my own consumer profile in a marketing database, such as a combination of my age, education, and/or political leaning, or from cookie data based on sites I have visited. Or, something else. For both 1) the (likely) bidding that occurred in an ad auction system just prior to this page loading, in which the publisher sold this ad space on the site and 2) the earlier estimation of my consumer profile (e.g., based on browsing behavior), basic or complex unknown algorithms played a key role in generating a substantial portion of this webpage (the ads).

The additional ad on the bottom right (bottom gutter) for Marygrove College is a possible example of an algorithmic failure. While my browsing history would indicate heavy interest in higher education, and I live in nearby Ann Arbor, as a Ph.D. candidate with a Bachelor's degree, I am an unlikely applicant to attend Marygrove, a liberal arts college located in Detroit. This failure assumes the Marygrove ad campaign goals are recruitment and not simply about brand awareness.