EXPLORING FACEBOOK'S "WHY AM I SEEING THIS AD" FEATURE: MEANINGFUL TRANSPARENCY OR FURTHER OBFUSCATION?

Daniel Angus
Queensland University of Technology

Jean Burgess
Queensland University of Technology

Nicholas Carah
University of Queensland

Lauren Hayden
University of Queensland

Abdul Obeid
Queensland University of Technology

Introduction

For more than a decade, digital advertising has been the primary means of funding online content and services. The evolution of digital advertising towards algorithmically targeted advertising, believed to be highly personalized and tailored to the individual, has presented new challenges for public oversight. This shift has significant implications for regulators, civil society organizations, and researchers seeking to hold advertisers (and platforms) accountable for harms associated with advertising, including false, misleading, racist, discriminatory, and predatory marketing practices, especially when targeting vulnerable consumers.

Whereas previously, public concern centred on the content of ads and their exposure to audiences, the rise of platform-based advertising means focus has shifted to the distribution of ads and how they reach us. This concern reflects a consensus that if advertising's cultural power is increasing, it is not because it has become more symbolically persuasive but because it has become better at targeting us.

In response to public concerns and regulatory pressures, companies such as Meta (the parent of Facebook) have introduced transparency tools for researchers and consumers to 'explain' the function of advertising on the platform, including the Ad Library and the "Why Am I Seeing This Ad" feature. Abbreviated as WAIST, "Why Am I Seeing This" explanations appear as buttons within each ad seen by a user of the platform, when clicked this button takes users to additional information regarding how the ad was targeted towards them, such as demographic or interest-based categories (See Figure 1).

Despite being a central feature of Meta's response towards increasing external scrutiny, little is known about how the WAIST feature works, or how it operates at a population level. The WAIST feature is fundamentally individualistic: it provides an explanation of how a single ad was served to a unique user at a particular time. There is no public information as to the distribution or prevalence of specific WAIST tags across advertising markets or user populations, nor any systematic or detailed information about how they are utilised and deployed as a tool of transparency by Meta platforms.

In response we offer a description of WAIST data collected at scale, informed from a nationwide citizen data donation project of Facebook advertising. We analyse this data with a view to better understand Meta's algorithmic advertising system, and to inform questions regarding the sufficiency of WAIST as an algorithmic explanatory mechanism for users.

Why Am I Seeing This (Ad)

The WAIST system was introduced in 2014 as a way for Facebook users to understand why they were seeing certain ads on their feed. The feature was intended to increase transparency and user control over the advertising they were exposed to. Initially, the WAIST feature was simple, providing users with basic information such as the advertiser's name and the targeting criteria that led to the ad being shown.

Over time, the WAIST feature evolved to become more sophisticated. In 2016, Meta introduced additional information to the feature, such as a list of interests and behaviours associated with the user that may have led to the ad being shown. Interest categories include semantic descriptors that vary from general to hyper-specific, such as 'online shopping' and 'The Great British Bake Off'. In 2018, the feature was further expanded to include information on the source of the ad, such as whether it was paid for by a political organization or a third-party advertiser. Meta also launched the first incarnation of their ad transparency library in 2018, which originally included only political and issue advertising, but was later expanded in 2019 to include all active ads on the platform.
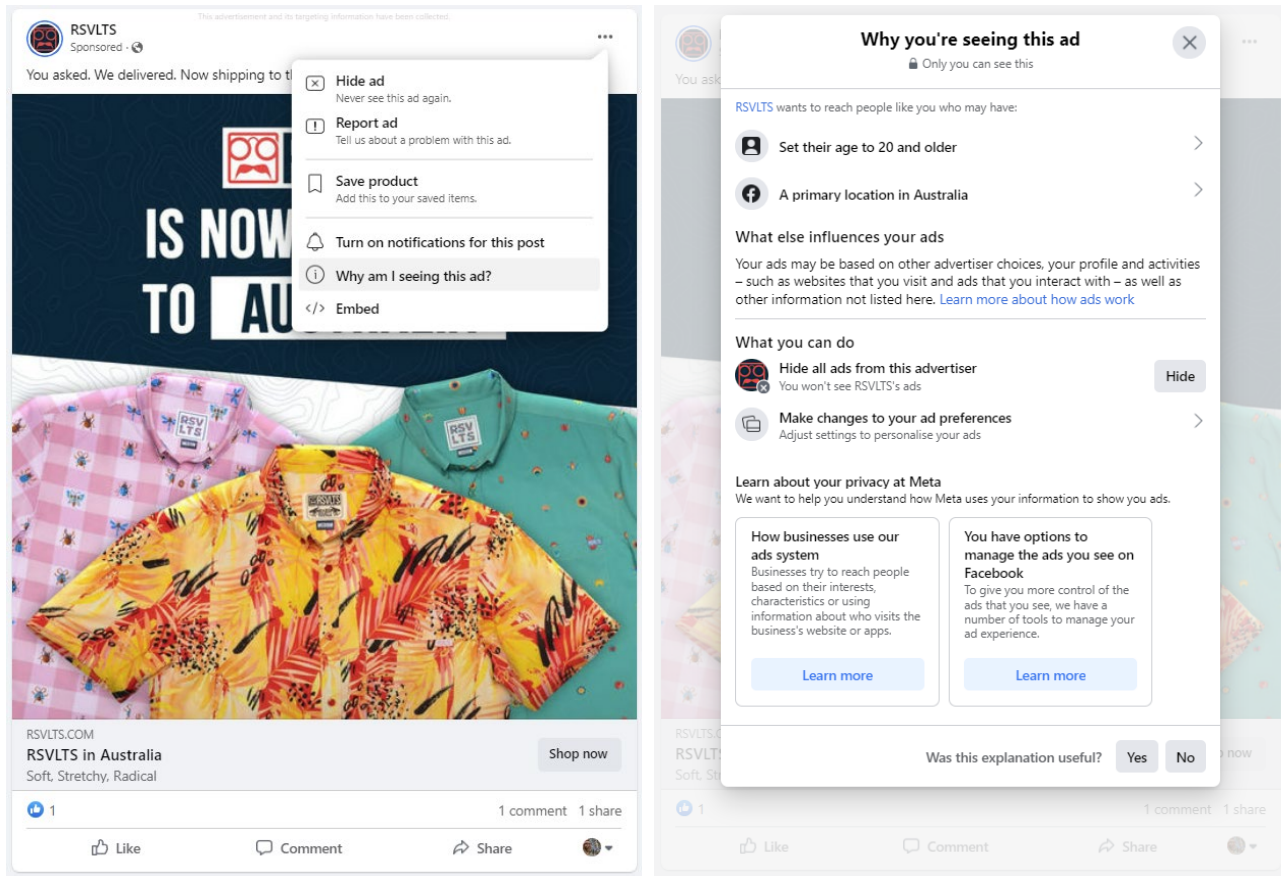
Figure 1: Example of a Facebook ad seen by one of the authors, showing how to access the WAIST data of the advertisement.

Despite these stated improvements, the adequacy of the Ad Library and WAIST as explanatory mechanisms is questionable. While the library provides some insight into what ads are *circulating*, it provides very little meaningful information about whether, how and to whom they are *targeted*, as researchers attempting to use them to study electoral campaigns have quickly discovered (Mehta & Erickson, 2022). To understand the power of algorithmic advertising, we need to be able to publicly observe patterns of who sees what ads. Under the current arrangements, however, Meta make public an aggregated view of the content of ads (through the Ad Library) but individualise and privatise all communication about how they are targeted (through the WAIST feature).

Patterns in WAIST data from the Australian Ad Observatory

Responding to the need for greater platform observability (Rieder & Hoffman, 2020), researchers and civil society groups have developed new data donation methods to collect advertisements and to work with users to crowdsource information about the ads they are seeing (Andreou et al., 2018; 2019). Our contribution to this global effort is the Australian Ad Observatory project, which extends similar ad donation approaches

(ProPublica, 2020) to investigate whether, how and when Facebook ads are targeted to Australian users (Burgess et al., 2022).

Deployed as a browser plugin, our software gathers ad creative, metadata, and deidentified observer demographic data for ads served through Meta's Facebook News Feed feature, while also providing our participants with a dashboard that enables them to monitor and reflect on the advertising that is targeted at them. By February 2023 the Australian Ad Observatory had gathered the support of 1817 participants who had made 760,000+ observations of approximately 330,000+ distinct ads and associated metadata, including WAIST categories (with 324868 distinct WAIST interest tags gathered to date).

Our approach enables us to speculate on the alignment of and gaps between WAIST data (which is a product of Facebook's transparency efforts with respect to its users) and patterns in our evidence about the platform's actual ad targeting practices (which is a product of advertisers' use of the Facebook ad system).

Our ongoing analysis suggests that WAIST tags are unevenly deployed, with AgeGender (this includes age ranges, and binary gender classification) and Location being the dominant WAIST tags applied to most ads. WAIST Interest tags are also highly prominent, while education, relationship, and employment status are rarely deployed (see Figure 2).
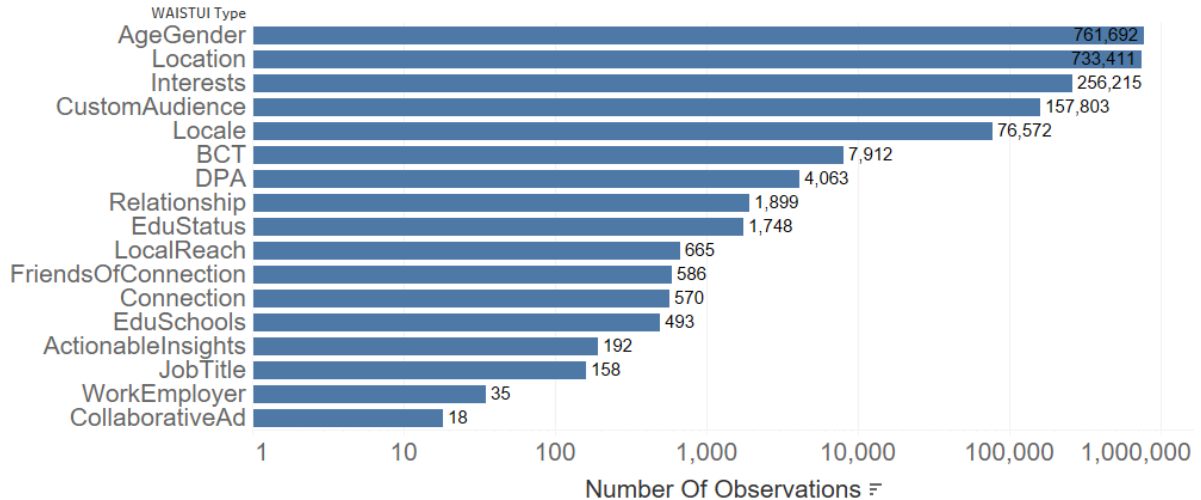

Figure 2: The distribution of WAIST tags across all participant observations in the ad dataset.

Interest tags appear in sets, with many of these semantic tags appearing together within a single observation, for example "Charity and causes" alongside "Community issues". By mapping these relationships as a graph, where each node represents an individual WAIST Interest tag, and edges weighted according to how may observations use any two tags, it is possible to map the topology of the WAIST Interest code ecosystem and identify emergent patterns (see Figure 3). While many patterns are banal, this analysis

provides insights into the potential reproduction of protected or sensitive categories like race, sexuality, gender and class based on personalised 'interests', rather than via the explicit targeting of demographic niches (Phan and Wark, 2022).
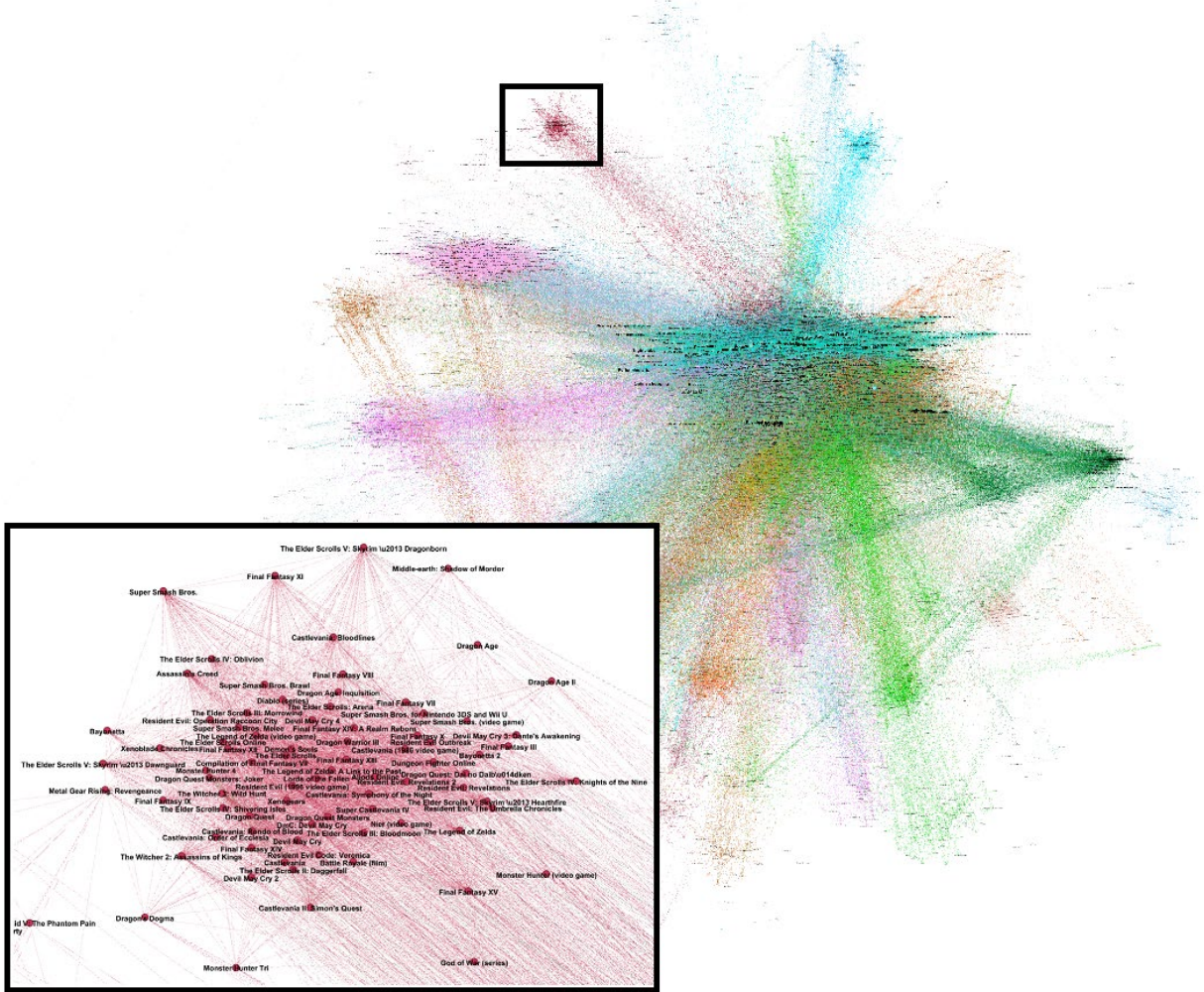

Figure 3: WAIST Interest tag network, highlighting a cluster of interest tags that feature names of popular console video games. This network provides a detailed glimpse into the organizational arrangement of WAIST Interest tags within Facebook's ad ecosystem.

Our analysis which also draws from an additional app walkthrough exercise (Light et al., 2018), reveals the insufficiency of WAIST and the Ad Library as sufficient explanations for the scope of advertising and its algorithmic targeting. We offer several avenues for additional exploration of this data and interpretations of the ongoing role of explainability, from both functional and cultural standpoints within the algorithmic advertising ecosystem.

References

Andreou, A., Silva, M., Benevenuto, F., Goga, O., Loiseau, P., & Mislove, A. (2019, February). Measuring the Facebook advertising ecosystem. In *NDSS 2019-Proceedings of the Network and Distributed System Security Symposium* (pp. 1-15).

Andreou, A., Venkatadri, G., Goga, O., Gummadi, K., Loiseau, P., & Mislove, A. (2018, February). Investigating ad transparency mechanisms in social media: A case study of Facebook's explanations. In *NDSS 2018-Network and Distributed System Security Symposium* (pp. 1-15).

Burgess, J., Andrejevic, M., Angus, D. & Obeid, A. (2022). Australian Ad Observatory: Background paper. ARC Centre of Excellence for Automated Decision-Making and Society, Working Paper 004. Retrieved from: https://apo.org.au/node/318616

Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps. New Media & Society, 20(3), 881-900. https://doi.org/10.1177/1461444816675438

Mehta, S., & Erickson, K. (2022). Can online political targeting be rendered transparent? Prospects for campaign oversight using the Facebook Ad Library. *Internet Policy Review*, *11*(1), 1-31.

Phan, T., & Wark, S. (2021). What personalisation can do for you! Or: how to do racial discrimination without 'race'. *Culture machine*, *20*, 1-29.

ProPublica. (2020). Facebook Political Ad Collector. Retrieved from: https://projects.propublica.org/facebook-ads/

Rieder, B., & Hofmann, J. (2020). Towards platform observability. *Internet Policy Review*, *9*(4), 1-28.