WHAT PERSONALISATION CAN DO FOR *YOU!* OR, HOW TO DO RACIAL DISCRIMINATION WITHOUT RACE

Thao Phan
ADM+S Centre, Monash University

Scott Wark
Centre for Interdisciplinary Methodologies, University of Warwick

In this paper, we use the life and death of Facebook's notorious "ethnic affinity" categories to identify and critique emergent forms of racialisation made possible by computational platforms, like social media, that use large-scale data collection and processing techniques to categorise and act on their users. We counter-pose these emergent forms of racialisation against the notion that by refusing to collect data about users' race, platforms inaugurate a "post-racial" reality. To the contrary, we argue that these technologies actively produce novel racial categories and, consequently, new forms of racial discrimination.

Facebook's "ethnic affinity" categories were superficially innocuous. Originally designed to help advertisers reach specific audiences, these categories used data on North American users' actions, their connections with other users, and their expressed preferences to sort people in to one of three "affinity" groups: some users were categorised as having an affinity for "African American," "Asian American," or "Hispanic" culture. A few years after this form of categorisation was introduced, a series of investigative reports (e.g. Angwin and Parris, 2016) and subsequent lawsuits began to reveal their potential for discriminatory use. After resisting these mounting critiques for years, in 2020 Facebook finally - quietly and unceremoniously - deprecated the ethic affinity feature (Facebook, 2020). As investigative journalists subsequently pointed out (Merril 2020), Facebook's advertising tools still seemed to enable discrimination against users on the basis of their race.

In this paper, we argue that the life and death of Facebook's ethnic affinity categories allow us to reflect on the shifting configurations of race and racial categorisation today. Rather than being a singular instance of a well-intentioned idea gone awry, this case typifies emerging techniques of racialisation made possible by data processing

techniques that purport to be able to apprehend us as individuals, but which can only do so by first assembling us into groups. It is our contention that ethnic affinities operationalise a function that is inherent to algorithmic culture: *discrimination*.

We contend that the quiet demise of these categories, which is supposed to have brought one particularly ignominious chapter in Facebook's recent history to a close, is bound up with one of digital culture's organising tendencies: the drive to subject widespread and, indeed, often *in*discriminate techniques of data collection to processing by AI-driven techniques to provide personalised services (Kant, 2020). Facebook's ethnic affinity categories typify a widespread mode of categorising the users of platform-based online services that seems paradoxical. In order for a service to be personalised *for you*, you must first be understood in relation to a set of others (Lury and Day, 2018). Indeed, platforms like Facebook claim that the capacity to personalise content ought to actually obviate crude demographic markers, like race: such techniques mean that marketers ought no longer have to "hold anything constant" (Facebook for Business, 2017). Personalisation is supposed to inaugurate what Darmody and Zwick call a "post-marketing turn" that itself has the potential to institute a "post-racial" reality (2020). These services *are* discriminate. We argue that these categories demonstrate how the drive to personalise has created the conditions for new forms of racial discrimination.

To substantiate this claim, we first contextualise the concept of "ethnic affinities" by tracing its operative term, "ethnic," to the post-war discursive shift from using terms like "race" to describe difference to using terms like "ethnicity," which is ostensibly-neutral and which lacks race's biological connotations (Kowal and Watt, 2018). We argue that Facebook's use of the term "ethnicity" allowed them to rebrand racialising techniques at a moment that "race" had once again become a particularly contested issue of public concern. Facebook posited ethnic affinities as a socially-useful technique of categorisation that could be used to reach otherwise "underserved minority markets" (García Martínez, 2019). What ethnic affinities' life and death makes clear, however, is that this kind of semantic substitution can't magic the category of race away.

We then draw on recent research by scholars working in fields including law, human-computer interaction, and data ethics to argue that Facebook's ethnic affinity categories themselves belong to a specific mode of categorisation mobilised by inferential machine learning/AI techniques: proxies. This literature describes the proxy as a variable that is known to be a correlate of another, and which can be used, in substitution for that other variable, to target a particular group (e.g. Datta et al, 2017). Whilst proxies can and do allow marketers to target their advertisements at specific, underserved minority groups, they also and necessarily allow platforms that don't otherwise collect protected information about their users to target or exclude users from campaigns based on those characteristics. Instead of inaugurating a post-racial present, we argue that the capacity to proxify protected characteristics recapitulates race in new - inferential, data-driven, often-automated, and all-too-obfuscated - ways.

This capacity to target individuals by sorting them into categories based on preferences, actions, or connections that stand in for what we would otherwise identify as "race" constitutes new modes of racialisation. Reciprocally, we can also use this capacity to target race by proxy to identify and critique the specific material affordances and

discriminatory effects of racialisation in an age of datafication and personalisation. Insofar as Facebook's ethnic affinity categories recapitulate race, we argue that we can use them as proxies, in turn, to make sense of race's novel form.

We conclude this paper by reflecting on the mechanisms of categorisation that underpin the novel forms of racialisation represented by Facebook's ethnic affinity categories. We make a conceptual argument: in order to understand how racialisation operates on and in platformised spaces, we need to remain attentive to proxification. These techniques can subject wholly new communities to racialisation and discrimination, because they offer a wholly new means for differentiating communities based on emergent commonalities of action, preference, or relation. But we also make a methodological argument. In order to understand, and to be able to critique, these emergent forms of racialisation, we need to train our attention not only on what race *is*, as it's embodied and experienced, but what it *does* - to us, as we're targeted, and with us, as we're assembled into new racialised formations.

References

Angwin, J. & Parris Jr., T. (2016) *Facebook Lets Advertisers Exclude Users by Race*, ProPublica, https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race

Darmody, A. & Zwick, D. (2020) Manipulate to empower: Hyper-relevance and the contradictions of marketing in the age of surveillance capitalism. *Big Data & Society*, January-June, 1-12.

Datta, A. et al. (2017) Proxy Non-Discrimination in Data-Driven Systems: Theory and Experiments with Machine Learnt Programs. *arXiv preprint arXiv:1707.08120*

Facebook for Business. (2017) *People-Based Marketing: Thinking People-First Planning and Measurement*, Facebook, viewed https://www.facebook.com/business/news/insights/the-future-of-marketing-people-based-planning-and-measurement

Facebook for Business. (2020) *Simplifying Targeting Categories*, Facebook, https://www.facebook.com/business/news/update-to-facebook-ads-targeting-categories

García Martínez, A. (2019) Are Facebook Ads Discriminatory? It's Complicated. *Wired*, https://www.wired.com/story/are-facebook-ads-discriminatory-its-complicated/

Kant, T. (2020) *Making it Personal: Algorithmic Personalization, Identity, and Everyday Life*. Oxford University Press, New York.

Kowal, E., & Watt, E. (2018). What is race in Australia? Journal of Anthropological Sciences, 96, 229–237.

Lury, C. & Day, S. (2019) Algorithmic personalization as a mode of individuation. *Theory, Culture & Society*, 36, 17-37.

Merrill, J.B. (2020) *Does Facebook Still Sell Discriminatory Ads?*, The Markup https://themarkup.org/ask-the-markup/2020/08/25/does-facebook-still-sell-discriminatory-ads