ALGORITHMIC ADVERTISING AS A WICKED PROBLEM

JEREMY NG*

ABSTRACT

Algorithmic advertising is the subject of a huge range of conflicting narratives. On the one hand, it seems to be the embodiment of surveillance capitalism: encouraging mass data collection, facilitating discriminatory treatment, and affecting how we form preferences and express ourselves online. On the other hand, it also allows for the personalized matching of consumer preferences with relevant products and services, providing the financing for our modern, innovative, mostly free digital economy. These harms and benefits are not equally distributed across populations. This Note argues that the way in which these competing narratives overlap and coexist reveals that algorithmic advertising is a "wicked problem" given the pervasive opacity of the AdTech ecosystem, the presence of multiple stakeholders and decision-makers with conflicting values, and systemic complexity.

This Note does not aim to take a clear position on the normative desirability of targeted advertising or to propose a straightforward solution that easily resolves these complexities. Instead, I argue for a form of regulation that I term "transparency as infrastructure regulation." Before stakeholders can form any substantive policy responses, there must be a baseline level of transparency around algorithmic advertising to create an empirical knowledge base for societal stakeholders to collectively create policy interventions, which I refer to as a "collective perspective" on targeted advertising.

* Jeremy Ng is a Counsel at the World Bank, where he works on legal and policy mandates across AI and the digital economy. He holds an LL.M. in International Legal Studies (Human Rights Scholar, Salzburg Lloyd N. Cutler Fellow, Amirsaleh Family Scholarship, Dean's Graduate Award) from NYU School of Law, along with a B.A. (Hons) in Law (De Hume Prize) from the University of Cambridge. The author would like to thank Benedict Kingsbury, Thomas Streinz, David Stein, David Satola, and Adele Barzelay, for their thoughtful comments. © 2025, Jeremy Ng.

I. INTRODUCTION

Algorithmic advertising (also known as targeted, behavioral, or interest-based advertising)[1] is the subject of a huge range of conflicting narratives. On one hand, it seems to be the embodiment of surveillance capitalism: encouraging mass data collection, facilitating discriminatory treatment, and enacting behavioral manipulation that impacts the way we form preferences and express ourselves online. On the other hand, it is also a vehicle for the accurate and personalized matching of consumer preferences with relevant products and services, generating positive social externalities while providing the financing for our modern, innovative, mostly free digital economy. These harms and benefits are not equally distributed across populations. I argue that the way in which these competing narratives overlap and coexist reveals that algorithmic advertising is a "wicked

1. These terms are used interchangeably in this Note.

problem": a particular species of problem that is entangled, ambiguous, and ill-defined.[2] Wicked problems are characterized by a lack of easy description, scarcity of coherent information, the presence of multiple stakeholders and decision-makers with conflicting values, and systemic complexity.[3] Framed in this light, this Note does not aim to take a clear position on the normative desirability of targeted advertising or to propose a straightforward solution that easily resolves these complexities. The nature of a wicked problem is that it defies simple, binary solutions. Instead, I argue that before stakeholders can form any substantive policy responses, there must be a baseline level of transparency around algorithmic advertising to create an empirical knowledge base for societal stakeholders to collectively create policy interventions.

This Note proceeds as follows: Part II provides a brief examination of the complex assemblage of largely opaque digital advertising infrastructures that operate along "outgoing" and "incoming" data vectors.[4] Part III reveals the complexity in assessing the harms and benefits of targeted advertising, concluding that it is a wicked problem. In Part IV, I use Google's proposed Topics API as a case study to demonstrate why any regulatory approach to targeted advertising that assumes a simple linear relationship between variables (e.g., ad targeting accuracy and behavioral harms) is unsatisfactory. Finally, in Part V, I argue for a form of regulation that I term "transparency as infrastructure regulation." I suggest that this facilitates a "collective perspective"[5] on targeted advertising: a democratic, multi-stakeholder approach to managing the policy tradeoffs of targeted advertising, creating space for a range of context-specific and personalized interventions deployed by individuals, user groups, industry actors, and regulators.

---

2. The phrase "wicked problem" is widely attributed to a 1973 paper by Horst Rittel and Melvin Webber. *See generally Horst W.J. Rittel & Melvin M. Webber, Dilemmas in a General Theory of Planning*, 4 POLICY SCI. 155 (1973)*; see also* BRIAN W. HEAD, WICKED PROBLEMS IN PUBLIC POLICY: UNDERSTANDING AND RESPONDING TO COMPLEX CHALLENGES (2022); J.B. Ruhl & James Salzman, *Symposium: Governing Wicked Problems, Introduction*, 73 VAND. L. REV. 1561 (DEC. 2020). Examples of other wicked problems are climate change, poverty, gender equality, and peaceful resolution of major international disputes. HEAD, *supra*, at 14.

3. Ruhl & Salzman, *supra* note 2, at 1562.

4. *See generally,* Ayelet Gordon-Tapiero et al., *The Case for Establishing a Collective Perspective to Address the Harms of Platform Regulation*, 25 VAND. J. ENT. & TECH. L. 635 (2023).

5. This phrase is borrowed from Gordon-Tapiero et al., who define a collective perspective as "providing a third party with ongoing insight into the information gathered and observed about individuals and how it correlates with any personalized content they receive across a large, representative population. These insights would enable the third party to understand, identify, quantify, and address cases of personalization-driven harms," *Id.* at 636.

II. Targeted Advertising Infrastructures

To understand how targeted advertising works, we must first explore how the advertising technology (AdTech) data ecosystem operates along both "outgoing" and "incoming" vectors. The "outgoing" vector—the bedrock of the ecosystem—relates to how platforms collect and process vast quantities of data about users and their activities.[6] For example, when a user repeatedly interacts with an e-commerce website, the e-commerce company may collect data on personal attributes (which users provide when creating an account on the website), items purchased, and other clickstream data.[7] In practice, such data is often not explicitly linked to a natural person but rather to an anonymized or pseudo-anonymous identifier, such as a cookie.[8] Third-party cookies (TCPs) are "cookies saved to a user's browser by a site other than the one they are currently visiting."[9] They are a critical infrastructure enabling the tracking of users on and across *multiple* sites.[10] TPCs facilitate the identification of granular interests, behaviors, and connections because they enable the tracking of extremely specific actions. For example, they can track the exact URLs a user visits along with the content on each of those pages.[11]

To transform data on user interactions into actionable insights, further processing (i.e., "profiling") is required. Behavioral profiling uses models trained using machine learning techniques to identify a general pattern in behavior (e.g., a correlation) based on the totality of customers' data. For example, a machine learning model may identify that eighteen-year-old women in a specific suburb of Milan enjoy punk music. It can then profile users who fit this description but who have not indicated any specific musical preferences as possessing a "taste for punk music."[12] Behavioral profiling is a powerful tool for advertisers because it enables them to group customers based on their inferred behavior, needs, wants, and other characteristics. This is known as "audience segmentation."[13] Platforms such as Facebook use their in-house

6. *Id.* at 644.

7. Federico Galli, Algorithmic Marketing and EU Law on Unfair Commercial Practices 52 (2022).

8. *Id.*

9. David Eliot & David Murakami Wood, *Culling the FLoC: Market Forces, Regulatory Regimes and Google's (Mis)steps on the Path Away from Targeted Advertising*, 27 Info. Polity 259, 261 (2022).

10. James Grimmelman, Internet Law: Cases & Problems 275-77 (2021).

11. *The Topics API*, Github, https://github.com/patcg-individual-drafts/topics (last visited Feb. 25, 2025) [hereinafter Github Draft Proposal].

12. Galli, *supra* note 7, at 52.

13. *Id.* at 53.

AI analytics capabilities to segment their users into groups based on different categories.[14]

The term "incoming" vector refers to how platforms use these profiles and audience segments to decide what content, including digital ads, will be presented to users. For example, segmentation enables Facebook to offer advertising clients the ability to target ads to a certain audience. Below is a list of the default options Facebook offers to advertisers when creating a new audience through segmentation:

- **Locations.** Target ads to account center accounts based on locations. You can select country, state, province, city, congressional district, zip and post codes. Most objectives let you target worldwide (type in "worldwide"), by region (for example, "Europe"), by free trade area (for example, "NAFTA," the North American Free Trade Agreement) or by app store availability (for example, "iTunes app store countries").
- **Age.** Target ads within an age range. . . .
- **Gender.** Target ads to women, men or all genders.
- **Languages.** Target ads to users of certain languages. . . .
- **Detailed targeting.** Include an audience based on criteria such as demographics, interests and/or behaviors. . . .
- **Custom audiences.** Custom audiences are select audiences you already know that created from information you provide or from information generated across Meta technologies. You can create custom audiences from a list, the Meta Pixel, the Facebook SDK and engagement across Meta technologies.[15]

These targeting strategies can be combined in a variety of ways. For example, a simple broad targeting strategy may use location and age to target ads to people eighteen to sixty-five years old within the United States.[16] In contrast, the detailed targeting option enables advertisers to select an initial audience preference (using age, gender, location, and/or language) and then further refine that audience via the inclusion of certain characteristics – e.g. "matching with certain films or

14. *Id.*

15. *About Reaching New Audiences*, META BUS. HELP CTR., https://www.facebook.com/business/help/717368264947302?id=176276233019487&helpref=page_content (last visited Feb. 25, 2025).

16. *About Broad Targeting*, META BUS. HELP CTR., https://web.archive.org/web/20241122204528/https://www.facebook.com/business/help/308474373366888?id=176276233019487 (last visited Feb. 25, 2025).

books, or interest in yoga."[17] Another targeting option is lookalike targeting, which uses data about a small set of "seed" users with certain desired characteristics or behaviors to algorithmically identify and target similar "lookalike" customers who are likely to respond positively to advertisements.[18]

III. Algorithmic Advertising as a Wicked Problem

The following part provides a snapshot of some of the competing, coexisting narratives around the benefits and harms of algorithmic advertising. It proceeds as follows: Section III.A discusses the potential harms of behavioral advertising, focusing on concerns around surveillance, mass data collection, behavioral manipulation, discrimination, and even the potential for such ads to shape preferences and identity expression. Section III.B surveys some of the potential benefits of interest-based advertising, including personalization, self-actualization, and access to free online services.

A. *The Harms of Behavioral Advertising*

1. Surveillance, Mass Data Collection, and Behavioral Manipulation

The dominant critique of behavioral advertising, which has emerged over the past decade, is that it facilitates the digital monitoring of our personalities, preferences, and desires for the purpose of behavioral manipulation—to convince us to buy certain products or act in a certain way.[19] Platforms leverage the analysis of huge datasets (by creating algorithmic models that make extremely fine-grained inferences regarding users' lifestyles, desires, and preferences) along with behavioral psychology to engage in a form of "data-driven persuasion."[20] The AdTech ecosystem is designed to deliver ads to users at the exact "micro-moment" when they are most receptive to advertising.[21] Some forms of targeting even seek to analyze a person's emotional state based on factors such as online activity and sentiment analysis to increase the chances of

17. *Use detailed targeting*, Meta Bus. Help Ctr., https://www.facebook.com/business/help/440167386536513?id=176276233019487&locale=en_GB (last visited Feb. 25, 2025).

18. *See generally* Lookalike Evaluation, U.S. Patent No. US 2017/0140283 A1 (filed Nov. 13, 2015).

19. Forbrukerrådet [Norwegian Consumer Council], Out of Control: How Consumers are Exploited by the Online Advertising Industry 12 (2020) [hereinafter Norwegian Consumer Council].

20. *Id.* at 13.

21. Erin Martz, *Micro-Moment Marketing: What Is It and Why Use It?*, Web FX, https://www.webfx.com/blog/marketing/micro-moment-marketing/ (last visited Apr. 12, 2025).

influencing spending behavior.[22] A surveillance capitalism critique sees such techniques as inherently harmful, in that they turn human behavior into a marketable economic asset, extracted for the exploitation of behavioral surplus.[23]

One counterargument to this surveillance-oriented critique is that targeted advertising is not meaningfully different from traditional advertising, in that they both seek to influence behavior by aiming certain messages at certain groups. Indeed, market segmentation has always been premised on the desire to "look for splits in the social fabric and then reinforce and extend the splits for [advertisers'] own ends."[24] This critique implies that digital advertising is simply a change in the scale of traditional advertising. However, I argue that the targeting methods surveyed above are more granular and sophisticated, distinguishing behavioral advertising from static television and print campaigns. Crucially, traditional advertising methods operate solely at the *incoming* vector. Digital advertising is novel in that it operates via surveillance at the *outgoing* vector as well, monetizing the collection and observation of a huge range of online behavior through the creation of data collection, classification, and profiling infrastructures designed to understand, segment, and profitably group users. Traditional advertising does not rely on the same kind of fine-grained surveillance of everyday online activity.

Aside from the direct harms that may arise from surveillance and manipulation, there are also data security risks inherent in mass data collection, particularly concerning sensitive personal information. For example, José González Cabañas et al. have found that Facebook has labeled sixty-seven percent of users worldwide with potentially sensitive ad preferences.[25] As of November 2020, 250,000 users in Saudi Arabia were labeled by Facebook with the ad preference "homosexuality"[26] (consensual same-sex activity is punishable by death in Saudi Arabia).[27] Cabañas et al. observe that it is fairly easy for malicious actors to retrieve

22. NORWEGIAN CONSUMER COUNCIL, *supra* note 19, at 47.

23. For a discussion of the concept of "behavioral surplus," *see* SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER 94 (2019); *see also* NORWEGIAN CONSUMER COUNCIL, *supra* note 19, at 12.

24. Rena Bivens & Oliver L. Haimson, *Baking Gender Into Social Media Design: How Platforms Shape Categories for Users and Advertisers*, SOC. MEDIA + SOC'Y, Oct.–Dec., 2016, at 2.

25. José González Cabañas et al., *Does Facebook Use Sensitive Data for Advertising Purposes?*, 64 COMMC'NS OF THE ACM 62, 62 (2021).

26. *Id.* at 64.

27. LUCAS RAMÓN MENDOS ET AL., STATE-SPONSORED HOMOPHOBIA 2020: GLOBAL LEGISLATION OVERVIEW UPDATE 69 (2020), https://ilga.org/wp-content/uploads/2023/11/ILGA_World_State_Sponsored_Homophobia_report_global_legislation_overview_update_December_2020.pdf.

the personal information of users tagged with sensitive ad preferences through a low-cost, phishing-like attack.[28]

Finally, mass data collection for commercial advertising purposes enables the exploitation of insights from large datasets for purposes that are difficult to predict. Political manipulation is a particularly striking example, with the Cambridge Analytica scandal in 2017 exposing the possibility of utilizing social media data to influence elections.[29] Researchers at the Brookings Institute have suggested that location data, in particular, can facilitate "geopropaganda" by tracking groups of individuals based on their attendance at political rallies and other events, then subsequently targeting these groups with political advertising.[30] The potentially vast societal impacts of such political manipulation include increasing voter polarization, undermining the free exchange of political ideas, or even presenting an existential threat to democratic institutions.[31]

2. Discrimination

In 2016, investigative journalism outlet ProPublica broke the news that Facebook allowed advertisers placing housing ads on the platform to exclude certain users by their race, which appeared to violate the United States' Fair Housing Act (FHA).[32] Facebook has since been

28. Cabañas et al., *supra* note 25, at 67. Facebook began updating its detailed targeting options in January 2022 to remove the ability for advertisers to target ads towards "topics people may perceive as sensitive, such as . . . sexual orientation." *Updates to detailed targeting*, META BUS. HELP CTR. (2023), https://www.facebook.com/business/help/458835214668072?ref=search_new_0. But as of March 2023, it still collects declared data on both gender identity and sexual orientation (via Facebook Dating). *See How do I change my Facebook Dating Preferences?*, FACEBOOK HELP CTR. (2023), https://www.facebook.com/help/adsmanagerbuiltin/3248344638724231.

29. In 2014, Cambridge Analytica, a data analytics firm with ties to key figures from Donald Trump's election team, used personal information taken without authorization from over 50 million Facebook profiles to build a system that could target voters with personalized political ads. Facebook discovered the breach in 2015 but failed to alert users, only taking action in 2018 after whistleblower Christopher Wylie revealed how the company "exploited Facebook to harvest millions of people's profiles" to "target their inner demons." *See* Carole Cadwalladr & Emma Graham-Harrison, *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*, THE GUARDIAN (Mar. 17, 2018), https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

30. Katie Joseff et al., *The disturbing implications of increasingly narrow political ad targeting*, BROOKINGS INST. (Feb. 11, 2021), https://www.brookings.edu/techstream/the-disturbing-implications-of-increasingly-narrow-political-ad-targeting.

31. *See* Katharina Baum et al., *Do They Really Care About Targeted Political Ads? Investigation of User Privacy Concerns and Preferences*, PROC. OF 27TH EUR. CONF. ON INFO. SYS. (2019), at 2.

32. *See generally* Julia Angwin & Terry Parris Jr., *Facebook Lets Advertisers Exclude Users by Race*, PROPUBLICA (Oct. 28, 2016), https://www.propublica.org/article/facebook-lets-advertisers-

embroiled in several lawsuits under the FHA, brought separately by the National Fair Housing Alliance (NFHA) and the U.S. Department of Housing and Urban Development, represented by the U.S. Department of Justice (DOJ).[33] As part of a 2019 settlement with the NFHA, Facebook created a redesigned "Special Ad Audience" based on its "Lookalike Audience" feature specifically for housing, credit, and employment (HEC) ads; this system was explicitly designed to not consider "age, gender, relationship status, religious views, school, political views, interest in [*sic*], or zip code" as inputs.[34] However, a study conducted by Piotr Sapiezynski et al. on the modified Special Ad Audience function revealed that even if a certain demographic feature is omitted, an algorithmic model may still use that feature in practice by combining it with other correlated inputs, nullifying any protection from discriminatory effects.[35] When Sapiezynski et al. ran a test ad placement to find an audience "similar to" Facebook's current employees, although the source audience was not selected based on age or gender, the resulting Special Ad audience remained skewed heavily towards twenty-five to thirty-four-year-old men.[36] The authors concluded that "the findings in this paper are a natural result of how complex algorithmic systems work in practice."[37]

Crucially, although Facebook had undertaken good-faith steps to prevent advertisers from discriminating against certain groups (by removing protected characteristics as inputs to the lookalike model algorithm), due to the complexities of algorithmic inference and machine learning, these steps were insufficient to prevent biased outcomes in ad delivery. Empirical studies have documented a similar discriminatory effect in the

exclude-users-by-race. In the targeting taxonomy discussed above, this was made available through Detailed Targeting, allowing exclusion based on "Ethnic Affinity"—a demographic trait assigned to users based on behavioral profiling (using the techniques described in the previous section), and not explicitly disclosed by the users. *Id.* Despite attempts by Facebook to prevent advertisers from discriminating on the basis of race, ProPublica found in 2017 that attempts to use automation to enforce its policies contained "major omissions." Julia Angwin et al., *Facebook (Still) Letting Housing Advertisers Exclude Users by Race*, PROPUBLICA (Nov. 21 2017), https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin.

33. *Justice Department and Meta Platforms Inc. Reach Key Agreement as They Implement Groundbreaking Resolution to Address Discriminatory Delivery of Housing Advertisements*, DEP'T OF JUST. (Jan. 9, 2023), https://www.justice.gov/opa/pr/justice-department-and-meta-platforms-inc-reach-key-agreement-they-implement-groundbreaking; *Facebook Settlement*, NFHA (Mar. 14, 2019), https://nationalfairhousing.org/facebook-settlement/.

34. Complaint at 24, United States v. Meta Platforms, Inc., (S.D.N.Y. 2022) (No. 1:22-cv-05187).

35. *See* Piotr Sapiezynski et al., *Algorithms that "Don't See Color": Measuring Biases in Lookalike and Special Ad Audiences*, PROC. OF THE 2022 AAAI/ACM CONF. ON AI, ETHICS, AND SOC'Y 1, 1-2 (2022).

36. *Id.* at 6.

37. *Id.* at 2.

delivery of HEC ads for non-binary people on Facebook.[38] Sara Kingsley et al. found that this discriminatory effect was particularly stark in the credit ad context, in which persons who did not reveal their gender identity on Facebook or identified as neither male nor female were "rarely, if ever, shown credit ads of any type."[39] Again, although advertisers could not explicitly discriminate based on gender via the advertising portal (the only gender segmentation option given to advertisers was to select "all"), Kingsley et al.'s audit process revealed that the system nevertheless created discriminatory outcomes due to "platform optimizations."[40] As Olga Goriunova notes, abstraction in data analytics is not linearly causal. Infrastructures such as machine learning operate at a distance, "obtaining models, deducing statements, making inferences."[41] As such, it cannot be assumed that a change in the outgoing vector has any correlation to outcomes at the incoming (targeting) vector—this is an inherent feature of algorithmic inference. For policymakers and other stakeholders seeking to formulate meaningful legal or regulatory interventions in the AdTech market, the emergent and complex properties of algorithmic inference pose a significant challenge.

3. Preference Formation and Identity Expression

In addition to potential harms arising from surveillance and behavioral manipulation, there are also more subtle ways in which algorithmic ad delivery may affect how we form preferences, express ourselves, and even form our own identities.[42] The emergent patterns analyzed by algorithmic models may not correspond to genuine rational preferences, but rather to vulnerabilities. For example, a smoker with a browsing history of searching for tobacco products may be served ads providing a discount on cigarettes.[43] A profiling algorithm is unable to distinguish between a first-order preference to buy cigarettes versus a second-order desire to quit smoking; it cannot differentiate between conscious,

38. Discrimination in the housing, employment and credit (HEC) context is particularly problematic because it represents the systematic exclusion of certain demographics from specific socioeconomic opportunities. For a legal analysis of the role of digital advertising in reproducing historic patterns of discrimination, *see generally* Pauline T. Kim, *Manipulating Opportunity*, 106 VA. L. REV 867 (2020).

39. SARA KINGSLEY ET AL., AUDITING DIGITAL PLATFORMS FOR DISCRIMINATION IN ECONOMIC OPPORTUNITY ADVERTISING 21 (2020).

40. *Id.* at 8.

41. Olga Goriunova, *The Digital Subject: People as Data as Persons*, 36 THEORY, CULTURE & SOC'Y 125, 129 (2019).

42. *See* NORWEGIAN CONSUMER COUNCIL, *supra* note 19, at 52-53.

43. GALLI, *supra* note 7, at 100.

rational behavior on one hand, and impulsivity or dependence on the other.[44]

Algorithmic advertising may also disrupt the process through which users form preferences. Because such models rely on data regarding past customer interactions (at both an individual and collective level), self-reinforcing feedback loops can be created that lock consumers into their past purchasing choices.[45] This may cause consumer preferences to remain increasingly stable over time. In turn, this potentially discourages shifts in global consumption behavior towards environmentally conscious purchasing, for example, if users are not already inclined to be eco-friendly.[46] Indeed, this feedback loop may continue to encourage behavior that has negative externalities (e.g., smoking). This disruptive effect then has implications for the ways in which we construct our identities. Julie Cohen notes that "[w]e do not experiment only with beliefs and associations, but also with every other conceivable type of taste and behavior that expresses and defines self."[47] It is this process of experimentation and exploration of preferences that forms a "vital part of the process of learning, and learning to choose, that every individual must undergo."[48] Cohen hypothesizes that the absence of a "realm of autonomous, unmonitored choice" resolves into a dynamic which "incline[s] choices toward the bland and the mainstream";[49] a dynamic which in turn may discourage non-mainstream self-expression in a myriad of subtle ways that are impossible to predict *ex ante*.

One particularly striking possibility is the risk of misgendering in targeted advertising—an issue that remains relatively unexplored in the literature. Some AdTech stacks integrate models that seek to infer the gender of users in the absence of an explicit gender identity declaration, to more accurately segment target audiences. The inferential, probabilistic nature of machine learning means that a person's preferred gender identity may not always align with the gender assigned by a statistical model of gender, and indeed, the latter may shift as online activity changes over time. In John Cheney-Lippold's words, the "subsequent index for masculine identity is not men but the datafied patterns of what users do as 'men.'"[50]

---

44. *Id.* at 100-01.

45. *Id.* at 103.

46. *See id.* at 103-04.

47. Julie Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1425 (2000).

48. *Id.*

49. *Id.* at 1424-25.

50. JOHN CHENEY-LIPPOLD, WE ARE DATA: ALGORITHMS AND THE MAKING OF OUR DIGITAL SELVES 60 (2017).

For most cisgender users whose gender identity aligns with their sex assigned at birth, any misalignment may be barely noticeable. But what happens if, for example, a trans male user is assigned an advertisement that mistakenly assumes that he is female, based on his browsing history? What does it mean for non-gender-conforming persons to be misgendered by a targeting ad infrastructure?[51] Can such misgendering have an impact on how non-gender-confirming individuals construct their own sexual and gender identities? These questions are difficult ones without clear answers—the scale of identity-affecting misgendering in advertising infrastructures is extremely difficult to ascertain in the abstract given the lacuna in the empirical literature. Although the theoretical literature seems to implicitly assume that this dynamic causes a "recursive" effect on gender identity,[52] one of the few relevant studies by Camila Nunes Rinaldi, suggests that non-binary users on Instagram generally feel that ads have been targeted to them based on their interests rather than their gender identity.[53] Although Christopher Summers et al. suggest that behaviorally targeted ads often function as implicit social labels,[54] this effect only holds when the targeting is at least moderately accurate.[55] As such, the true effect of targeted ad misgendering

51. When experienced in everyday life as a negligent, accidental or intentional phenomenon, misgendering reflects a decision (conscious or unconscious) by the speaker to reject the subject's identity and impose the speaker's own conception of what is normatively acceptable, in turn infringing on the autonomy, dignity, and privacy of the misgendered person. *See* Chan Tov McNamarah, *Misgendering*, 109 CAL. L. REV. 2227, 2253, 2265-93 (2021). They conclude that "[m]isgendering is disrespectful, humiliates gender minorities, deprives them of privacy, safety, and autonomy, contributes to epistemic injustices, and is a tool of gender policing, social subordination, and identity invalidation." *Id.* at 2293. But when misgendering is experienced as the product of distance between digital subject and a person's own conception of their self, the harms that arise may be qualitatively similar, but are likely to be experientially very different.

52. *See generally* CHENEY-LIPPOLD, *supra* note 50.

53. Camila Nunes Rinaldi, How Non-Binary People Experience Targeted Advertising on Instagram 108 (M.Phil. thesis, Queensland University of Technology) (on file with author). However, these insights are difficult to extrapolate into general empirical conclusions given the extremely small sample size (the study interviewed only four participants, all of whom were based in Australia), and limited focus on Instagram as a platform (as opposed to the broader advertising ecosystem, including website display ads). As Rinaldi notes, their "analysis remains exploratory rather than conclusive, and the sample is far from fully representative." *Id.* at 119.

54. *See generally* Christopher A. Summers et al., *An Audience of One: Behaviorally Targeted Ads as Implied Social Labels,* 43 J. OF CONSUMER RSCH. 156, 157 (2016).

55. *Id.* For example, in Rinaldi's study, one user received targeted ads for earrings because they had previously sought to purchase a pair for a friend but was able to recognize that this targeting decision was likely a function of their search history rather than the platform's assumptions regarding their gender identity and as such no labelling effect occurred in practice.

remains unclear. Nevertheless, it is striking to consider the possibility that privately created, largely opaque algorithmic models may have an impact on gender expression.

Another open question is whether advertising carries greater or lesser impacts on identity formation than ordinary algorithmically delivered content. For example, TikTok's content recommendation algorithm has been the subject of a growing body of social sciences research. Aparajita Bhandari and Sara Bimo use the framework of the "algorithmized self" to highlight the direct relationship between the TikTok algorithm and users' self-perceptions.[56] However, setting aside the question of recommender algorithms, there are still credible reasons to believe that advertising has an identity-shaping impact. First, all targeted ads function on the assumption that consumer behavior is malleable, as discussed in Section III.A.1 above. Second, vulnerable groups such as children are often unable to resist the behavioral effect of advertising when it is embedded within trusted social networks and placed side-by-side with personalized content.[57]

B. *The Benefits of Interest-based Advertising*

1. Personalization and Self-actualization

We now turn to arguments highlighting the benefits of personalized ads. There is a credible argument that personalization provides a compelling solution to navigating the increasingly varied and diverse options for purchasing goods and services in today's marketplace.[58] Given the choice between seeing an irrelevant ad and an ad carefully tailored to one's interests, one could argue that most rational consumers would choose the latter.[59] In theory, personalization enables the marketplace to function more efficiently by connecting consumers

To quote the user, "[i]f I got earring ads, or make-up ads, or skincare ads, and I was not interested in those and not googling about those, then I might feel like, 'Oh, Instagram views me as a woman, and they are trying to get me these things.' But because I know that I've looked at, you know, earrings and make-up and skincare, I just think it's just connected to the things that I've googled or liked, or watched on YouTube." Rinaldi, *supra* note 52, at 111.

56. *See generally* Aparajita Bhandari and Sara Bimo, *Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media*, 8 SOC. MEDIA + SOC'Y (2022).

57. *See* American Academy of Pediatrics, *Digital Advertising to Children*, 146 PEDIATRICS 1 (2020).

58. GALLI, *supra* note 7, at 91.

59. Surveys commissioned by AdTech industry participants reinforce this (although these findings should be taken with a grain of salt given the obvious pro-advertising bias of the authors). *See* Holly Pauzer, *71% of Consumers Prefer Personalized Ads*, ADLUCENT (May 12, 2016), https://web.archive.org/web/20220513051314/https://www.adlucent.com/resources/blog/71-of-consumers-prefer-personalized-ads/.

with products they enjoy or find useful and reducing the amount of wasted time spent searching for such products. From an economic perspective, advertising creates value by "facilitating welfare-enhancing matches between consumers and firms."[60] Targeted advertising enables a higher "match quality" between a consumer's preferences and the closest matching product, while simultaneously leading to fewer ads being served, all else being equal.[61] Targeted advertising also makes advertising spending more efficient, potentially reducing spending on marketing. Traditional economic logic would suggest that this leads to a trickle-down effect on prices paid by consumers for the final goods and services promoted using digital advertising.[62]

One counterargument may be that the intended function of personalized ad targeting is not consumer welfare maximization. In other words, the platform's true clients are not its users but advertisers. As explored above, algorithmic personalization is not performed on the basis of the declared needs or desires of consumers, or a nuanced understanding of how we make decisions or rank our preferences; instead, it operates by a continued process of measuring, predicting, and encouraging behavior that generates the most economic benefit for advertisers (as measured by click-through rates, impressions, conversion rates, or other metrics).[63]

However, this does not preclude the benefits of personalization that lead to the creation of positive social externalities, especially where advertiser, platform, and consumer incentives align. For example, evidence from the marketing literature suggests that sexual minorities often actively engage with advertising to challenge stigmatization and construct self-validating meanings as a gateway toward empowerment.[64] Crucially, Wan-Hsiu Sunny Tsai suggests that targeted ads are viewed by gay and lesbian consumers as highly *politically* significant in "legitimizing

60. YAN LAU, A BRIEF PRIMER ON THE ECONOMICS OF TARGETED ADVERTISING 5 (2020), https://www.ftc.gov/system/files/documents/reports/brief-primer-economics-targeted-advertising/economic_issues_paper_-_economics_of_targeted_advertising.pdf.

61. *Id.* at 6.

62. *See* Commission Regulation 2022/1925 of Sept. 14, 2022, on Contestable and Fair Markets in the Digital Sector (Digital Markets Act), 2022 O.J. (L 265) 11 [hereinafter DMA]*; see also* LAU, *supra* note 61, at 6. This, however, assumes that advertising intermediaries do not charge increased prices for providing the complex technical infrastructure necessary to deliver targeted ads; indeed, there are arguments that Google abuses its dominant position within the AdTech stack to extract supra-competitive prices. *See* Complaint at 3, United States v. Google LLC, No. 23-CV-00108 (E.D. Va. Jan. 24, 2023).

63. GALLI, supra note 7, at 95.

64. *See generally* Wan-Hsiu Sunny Tsai, *How Minority Consumers Use Targeted Advertising as Pathways to Self Empowerment*, 40(3) J. ADVERT. 85 (2011).

the gay community and, by extension, gays' and lesbians' own sense of self-worth."[65] Participants in Tsai's survey emphasized that there was a "comforting validation [in] being found worthy as a demographic category for targeted advertising."[66] That is, the fact that advertisers were attempting to court gay and lesbian consumers was a "significant milestone" on the road to "consumer equality."[67] In turn, this economic citizenship enables gays and lesbians as consumers to "demand market reform and social changes by wielding their (economic) power via consumption."[68] To be clear, as Tsai acknowledges,[69] this account of the validating effects of targeted advertising on sexual minorities is underpinned by capitalist ideology and homonormativity.[70] A credible argument could therefore be leveled against this route to self-affirmation, in that it simply grants participation in neoliberal consumption for those in the privileged middle classes, aligning with advertising incentives to extract behavioral surplus from users.[71] The broader point for our purposes is that the utility that users gain from personalized ads will greatly depend on their perceptions of targeted advertising. Consumer perceptions are likely influenced

65. *Id.* at 90. It is important to note that Tsai's survey comprised of twenty-five gay and lesbian participants and did not include a focus on gender identity or trans identity.

66. *Id.* at 89.

67. *Id.* at 90.

68. *Id.* at 89.

69. *Id.* at 94 ("participants' acquiescence to marketing exploitation—and even reappropriation of marketing objectification and commercialization of gay experiences as a form of social incorporation—is a telling sign of the supremacy of capitalist hegemony … economic citizenship is predominantly enfranchised according to one's financial capability").

70. The term homonormativity was popularized by Lisa Duggan, who defined it as "a politics that does not contest dominant heteronormativity assumptions and institutions but upholds and sustains them while promising the possibility of a demobilised gay culture anchored in domesticity and consumption." LISA DUGGAN, THE TWILIGHT OF EQUALITY? NEOLIBERALISM, CULTURAL POLITICS, AND THE ATTACK ON DEMOCRACY 50 (2003).

71. This argument can be further developed along two lines. First, the pathway to consumer activism described in Tsai's account fundamentally entails economic privilege — a vision of the world where one's citizenship is conditional upon (a) assimilation into the mainstream and (b) financial capability. As such, Tsai concludes that "more economically disadvantaged minority consumers may have a radically different relationship with targeted advertising." Tsai, *supra* note 65, at 95. Second, much of the empowering effect described by Tsai rests on consumer myths, such as the "affluent, carefree gay spender." *Id.* at 88. Myths that not only rest on shaky empirical evidence, but also a picture of gay culture that is white, middle-class, and gender-normative. *Id.* at 92-93. This is a model largely unchallenged in queer advertising at large: an analysis of close to 200 LGBTQ+ ads revealed that the vast majority involved ad characters that were Caucasian, middle-class, middle-aged gay men, with few non-White LGBTQ+ identities represented. Ana-Isabel Nölke, *Making Diversity Conform? An Intersectional, Longitudinal Analysis of LGBT-Specific Mainstream Media Advertisements*, 65 J. HOMOSEXUALITY 224, 236-47 (2018). Non-traditional gender identities were represented only sporadically, while class and age diversity was largely absent. *Id.*

by a range of factors, including the type of data being used for profiling (with location data being a particular privacy concern),[72] the targeting methods used,[73] and the content of the ad.

2. Access to Free Online Services

Skeptics may continue to resist arguments in favor of personalization, arguing that the inherently coercive nature of targeted advertising presents a "fantasy where it seems possible for empowerment to coexist with manipulation."[74] The pragmatic counterargument is that our current (largely free) digital economy runs along deep path dependencies entrenched due to the ubiquity of targeted advertising. Huge sectors of the digital economy, from online gaming to social media and journalism, depend on targeted ads to survive. A common analogy is that consumer advertising allows users to monetize their time and attention to then "pay for" free online services.[75] Empirical studies show that the median consumer values free search engine services at USD 17,530 annually and email at USD 8,414 annually.[76] This account is open to contestation: Katherine Strandburg argues that this analogy is misleading because it implies there is a functioning market for the exchange of personal information for access to online products and services; in reality, data collection does not signal user preferences regarding products and services in a similar way to traditional price mechanisms.[77] Nonetheless, one does not need a rose-tinted view of behavioral advertising to accept that disrupting ad monetization would have enormous ramifications for the way we live, work, socialize, and conduct business online. This in turn would have knock-on effects on the volume, frequency, and quality of digital innovation.[78]

---

72. NORWEGIAN CONSUMER COUNCIL, *supra* note 19, at 44.

73. *See* Jiwoong Shin & Jungju Yu, *Targeted Advertising and Consumer Inference*, 40 MKTG. SCI. 900, 900 (2021).

74. GALLI, *supra* note 7, at 99.

75. LAU, *supra* note 61, at 4.

76. *See generally* Erik Brynjolfsson et al., *Measuring Welfare with Massive Online Choice Experiments: A Brief Introduction*, 108 AEA PAPERS AND PROCEEDINGS (2018) (cited by Lau, *supra* note 61, at 11).

77. *See generally* Katherine J. Strandburg, *Free Fall: The Online Market's Consumer Preference Disconnect*, 95 U. CHI. LEGAL F. 95 (2013).

78. For example, Kircher and Foerderer's study on the effect of Google's ban of targeted advertising in Android children's games in 2019 found that it caused substantial app abandonment (developers stopped updating games, with affected games more likely to be delisted). Tobias Kircher & Jens Foerderer, *Ban Targeted Advertising in Apps? An Empirical Investigation of the Consequences for App Development*, 70 MGMT. SCI. 1070, 1071-72 (2024).

In particular, there is a serious risk that replacements for advertising funding could disproportionately affect low-income households.[79] If policy interventions reduce the efficacy of targeted advertising—or preclude it entirely—consumers will lose the opportunity to monetize their personal data, generating a negative income effect that consumers pay for either in dollars (via subscription-based models) or in increased time and attention spent watching less relevant ads. This would disproportionately affect the poor: if the digital economy shifts to a primarily subscription-based model, lower-income groups may find themselves cut off from huge sections of the digital environment they rely on for communication, employment, and recreation by paywalls or other price barriers, in turn continuing to entrench existing digital divides.[80] This is especially true for the Global South, where entire communities rely on free platforms such as Facebook to conduct business and communicate. For these communities, "Facebook *is* the internet."[81] Alternatively, even if the ecosystem switches to a "freemium" approach, this inevitably creates a tiered system where lower income groups are forced out of necessity to use a lower quality version of the relevant digital product. This may have unintended consequences. For example, "freemium" online games, which provide a free baseline product with the option to spend money on in-game items or perks, have been criticized for encouraging predatory game design practices.[82]

Crucially, much of the above analysis assumes that *behavioral* advertising is necessary for the continued existence of ad revenues and hence free services. Alternative forms of digital advertising exist, ranging from contextual targeting of ads based on the content of the specific page visited, daypart targeting which shows ads during specific times of the day, and even technical targeting which segments based on the type of device (e.g., laptop or smartphone) used to access the relevant content.[83]

---

79. Ash Johnson, *Banning Targeted Ads Would Sink the Internet Economy*, ITIF (Jan. 20, 2022), https://itif.org/publications/2022/01/20/banning-targeted-ads-would-sink-internet-economy/.

80. *See* Emily A. Vogels, *Digital divide persists even as Americans with lower incomes make gains in tech adoption*, PEW RSCH. CTR. (June 22, 2021), https://www.pewresearch.org/short-reads/2021/06/22/digital-divide-persists-even-as-americans-with-lower-incomes-make-gains-in-tech-adoption/.

81. Nesrine Malik, *How Facebook took over the internet in Africa — and changed everything*, THE GUARDIAN (Jan. 20, 2022), https://www.theguardian.com/technology/2022/jan/20/facebook-second-life-the-unstoppable-rise-of-the-tech-company-in-africa.

82. *See, e.g.*, Justin Davis, *The Dark Future of Freemium Games, and How We Can Avoid It*, IGN (July 20, 2012), https://www.ign.com/articles/2012/07/20/the-dark-future-of-freemium-games-and-how-we-can-avoid-it.

83. *See Mozilla's Ad Targeting Guidelines*, MOZILLA, https://www.mozilla.org/en-US/privacy/ad-targeting-guidelines/ (last visited July 3, 2025).

Nevertheless, these forms of targeting may lead to fewer conversions or clicks as compared to behavioral advertising, meaning that advertisers would be willing to pay less than they otherwise would for behaviorally targeted advertising.[84] Thus, publisher ad revenues would decrease, with subsequent impacts on the volume and diversity of platform innovation. Although there is some anecdotal evidence to suggest that contextual targeting allows for similar revenue generation as behavioral advertising,[85] it is unclear whether this would hold true for all publishers.

C. *Wicked Problems*

The above analysis demonstrates that targeted advertising is subject to a dizzying array of considerations. A huge range of serious harms can arise from targeted advertising: surveillance, behavioral manipulation, discrimination, and even the hardening of personal preferences. And yet personalized advertising can bring both economic benefits and positive social externalities; critically, it seems that some level of digital advertising is necessary to continue to support our mostly free online economy.

How can we retain these positive economic and social benefits and support the current digital ecosystem while foreclosing or reducing the risk of harms? Is such a tradeoff sustainable, or possible, given the current political economy of algorithmic advertising? The policy concerns are deceptively complex. Many variables are at play, not all of which remain stable over time or across populations. For example, many of the accounts above treat users as a homogenous group, yet behavioral science shows that not everyone has the same biases, and we do not experience these to the same extent.[86] User preferences regarding the desirability of targeted advertising will differ across demographic groups—a White, gay, unemployed male who is a former alcoholic may have a very different experience of targeted advertising when compared to a Black, straight, female computer scientist who likes vegetarian food. Users' preferences towards targeted advertising may not be binary, and one individual's preferences may vary across the different

84. *See* LAU, *supra* note 61, at 3-4.

85. When the EU General Data Protection Regulation (GDPR) came into force in the EU, the New York Times moved from targeted behavioral advertising to contextual advertising based on content and general geographic parameters. However, advertising revenues continued to increase. Jessica Davies, *After GDPR, The New York Times cut off ad exchanges in Europe — and kept growing ad revenue*, DIGIDAY (Jan. 16, 2019), https://digiday.com/media/gumgumtest-new-york-times-gdpr-cut-off-ad-exchanges-europe-ad-revenue/.

86. GALLI, supra note 7, at 105.

vectors of the ecosystem. A user may feel uncomfortable with data about their sexual orientation being collected and processed, but simultaneously have no issue with being served targeted ads that depict non-mainstream sexual orientations and gender identities in a favorable light (as long as such targeting is not perceived as excessively intrusive). Recognizing that no two users are alike also reveals that the axes of benefit and vulnerability often intersect. Low-income users in particular are disproportionately harmed by digital surveillance,[87] and yet they are the most at risk of being cut off from core digital services if the economy were to move away from targeted advertising.

The huge variance in the normative, social, and political considerations embedded within the content of advertisements also generates unique challenges. Ads can represent economic opportunity (HEC ads), pathways to addiction and dependence (ads for online gambling games), or sites of political participation (local election campaign ads). They can also represent the incredibly mundane minutiae of daily life: promotions for plays, cooking utensils, or used cars. The relative risks and benefits across all the foregoing categories will vary. The granularity of targeting also influences these tradeoffs, as explored further below.

Underlying this is a pervasive opacity. The AdTech ecosystem operates largely as a block-box system to those outside of the industry; its sheer complexity makes it difficult for users to trace what happens to their data from the point of collection onwards.[88] Empirical studies show that users have little knowledge of, or often hold misconceptions regarding, how online behavioral advertising works in practice.[89] Those who do have some degree of knowledge tend to overstate the impact on others but understate the impact on themselves.[90] Even where advertisers voluntarily provide (or are mandated by regulation to disclose) information regarding how such infrastructures operate in practice, this often comes in the form of generic privacy disclaimers or limited ad disclosures, which are of little practical use to users.[91]

The above analysis demonstrates that targeted advertising is a "wicked problem" because of its entangled, ambiguous, and ill-defined implications.[92] Coherent information regarding targeted advertising is

87. *See generally* Mary Madden et al., *Privacy, Poverty and Big Data: A Matrix of Vulnerabilities for Poor Americans*, 95 WASH. U. L. REV. 53 (2017).

88. Sophie C. Boerman et al., *Online Behavioral Advertising: A Literature Review and Research Agenda*, 46 J. ADVERT. 363, 368 (2017).

89. *Id.*

90. *Id.*

91. See *id.* at 367.

92. *See also* Ruhl & Salzman, *supra* note 2. *see generally* HEAD, *supra* note 2.

scarce, and yet the stakes are high. Wicked problems cannot be definitively solved by any one "true" solution;[93] each independent solution must be evaluated in shades of gray, in terms of their respective benefits and drawbacks.

IV. GOOGLE'S TOPICS API

The concerns surveyed above have prompted a raft of self-regulatory efforts designed to preserve the legitimacy of behavioral advertising (and therefore the existence of digital advertising markets) while simultaneously ameliorating its harms. This Note analyzes the policy implications of one example of such self-regulatory efforts: Google's proposed Topics Application Programming Interface (API), which is currently being developed as part of a broader suite of "Privacy Sandbox" tools.[94] This Part proceeds as follows: Section IV.A briefly describes Topics as an infrastructure.[95] Section IV.B then evaluates Google's claims and explores why its approach to advertising granularity and digital distance does not adequately grapple with the nature of targeted advertising as a wicked problem.

A. *Introduction to Topics*

The Topics API has three main tasks. First, using a user's browsing activity, the browser observes and records topics that appear to be of interest.[96] Topics are selected from a manually curated taxonomy.[97] The use

93. HEAD, *supra* note 2, at 27.

94. The Topics API was introduced in January 2022 to replace the earlier "Federated Learning of Cohorts" (FLoC) infrastructure. Google discontinued development on FLoC due to push back from other browsers, civil society groups, and antitrust and data protection regulators. *See generally* Eliot & Wood, *supra* note 9. The Privacy Sandbox was initially created as an alternative to TPCs, which Google has made repeated attempts to phase out in recent years. Julia Love, *Google Delays Phasing Out Ad Cookies on Chrome Until 2024*, BLOOMBERG (July 27, 2022), https://www.bloomberg.com/news/articles/2022-07-27/google-delays-phasing-out-ad-cookies-on-chrome-until-2024#xj4y7vzkg. This policy of "depreciating" TPCs has since been reversed as of July 2024. *See* Anthony Chavez, *A new path for Privacy Sandbox on the web*, GOOGLE (July 22, 2024), https://privacysandbox.com/news/privacy-sandbox-update/.

95. One important caveat to the below description is that the API is still being tested and many of the detailed specifications (e.g., on github) are being published by Google to expedite the standardization process. Many of the specific details of the API (e.g. the topic taxonomy size, number of topics returned to an API caller, etc.) are subject to change. *See* Github Draft Proposal, *supra* note 10.

96. *Id.*

97. The initial topics have been selected by Google, but the eventual goal is to have them "sourced from an external party that incorporates feedback and ideas from across the industry." *Id.*

of human curation is intended "to exclude categories generally considered sensitive, such as ethnicity or sexual orientation."[98] Because these human-curated categories are far broader than, for example, the ones Facebook uses to catalog behavioral or demographic traits, Google posits that the "topics revealed by the API should be significantly less personally sensitive for a user than what could be derived using existing tracking methods."[99] The taxonomy is also specifically designed to contain a relatively low number of categories so that many users' browsers are associated with each topic.[100] The current topic taxonomy can be found on GitHub.[101] Some examples are: 1) "Pets & Animals/Pets/Fish & Aquaria"; 2) "Travel & Transportation/Business Travel"; and 3) "Food & Drink/Cooking & Recipes/Cuisines/Vegetarian Cuisine/Vegan Cuisine."[102]

Second, the Topics API calculates the top topics for a user based on their recent browsing history.[103] In the current implementation, for each week, the user's top five topics are calculated using browsing information local to the browser.[104] In contrast to the cookie-based model,[105] no information about the user's browsing history needs to be shared with third parties. All processing is done on-browser, with results shared via the API.[106]

Third, the Topics API allows API callers (e.g., advertising intermediation platforms) to "[a]ccess topics previously observed for the user, for example as a signal to select relevant advertising."[107] For example, when the API is called by an advertiser,[108] it returns up to three topics,

98. *Topics API for Web*, PRIVACY SANDBOX (2023), https://privacysandbox.google.com/private-advertising/topics/web [hereinafter Topics API Overview].

99. Github Draft Proposal, *supra* note 10.

100. *Id.* For example, the Github Draft Proposal initial design contains around 350 topics, whereas the IAB Audience Taxonomy contains around 1500.

101. *topics taxonomy*, GITHUB (2023), https://github.com/patcg-individual-drafts/topics/blob/b46fa17de6715f532802750f301feba585bd97fa/taxonomy_v1.md.

102. *Id.* For 10,000 specific top domains, the current implementation has a manual list mapping each domain to a topic or topics (e.g., google.com is mapped to the "search engines" topic). For sites not falling within this list, a machine learning classifier model is used to infer and assign topics based on the hostname, TOPICS API OVERVIEW.

103. *Topics API Overview.*

104. Github Draft Proposal, *supra* note 10.

105. *See* Part II above.

106. *Id.*

107. *Topics API Overview.*

108. An "API call" is the process by which data is requested from an API. This is often explained using analogies to waitering or catering: "If Jan is hosting a lot of guests for dinner, she might call a catering company and ask that they prepare food for the party. This saves her a great deal of time and effort preparing food herself. Similarly, one application can 'call' another for

one from each of the preceding weeks, based on the user's top five topics from each week.[109] The advertiser then uses these topics to select appropriate ads. However, it does not reveal additional information about the user's browsing history or activity—it is explicitly designed so that the API caller cannot learn more about the user than it could have using TPCs.[110] Google designs these topics to be clearly intelligible so that users can "remove individual topics, or clear their browsing history to reduce the number of topics returned by the API."[111] Figure One explains how the Topics API works to deliver ads in practice.

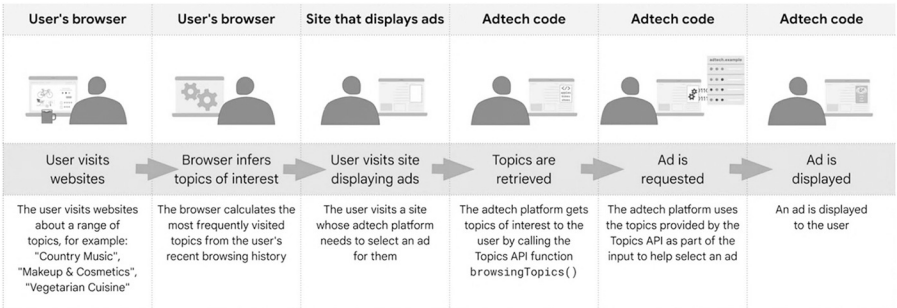

FIGURE ONE: The Topics API Diagram.[112]

B. *Evaluating Topics: "Digital Distance" and Targeted Advertising*

We can conceptualize the Topics API as an exercise in manipulating "digital distance."[113] A traditional digital subject constructed for

needed data or functionality. This ensures developers do not have to spend time and effort building application capabilities that can be integrated via API." *What is an API call?*, CLOUDFARE, https://www.cloudflare.com/learning/security/api/what-is-api-call/ (last visited July 3, 2025.

109. Github Draft Proposal, *supra* note 10.

110. *Topics API Overview.*

111. Topics API Overview, *supra* note 99.

112. Topics API Overview, *supra* note 99.

113. Goriunova argues that digital distance arises because of the nature of the computational processes involved in targeted advertising—there is never a purely indexical or linear link between a digital subject and the user, group, or thing it purports to represent. Using the example of Facebook's Lookalike Audience service, she argues that although machine learning techniques such as neural networks are able to deal with complex, non-linear cases, "we really have no way of establishing if they correspond in any meaningful way to any real complex system … no correspondence can be proven between outputs of neural nets and the real world." Goriunova, *supra* note 41, at 139. The affordances of the lookalike modelling are "described in terms of 'likelihood,' 'expectation,' 'possibility', and 'possibility of preference.'" *Id; see also* Camilla Cannon, *Non-Binary Gender Identity and Algorithmic Psychometric Marketing Legibility,* 22 FEMINIST MEDIA STUDIES

advertising purposes may describe, for example, an individual who is most likely a woman living on a particular street, orders organic vegetables and books, uses the London Overground, and does not have a Twitter account.[114] In contrast, for the same user using a browser implementing the Topics API, the only profile an advertiser can construct by querying the API in isolation is that of a person interested in three topics (based on recent browsing history), such as "Long Distance Bus & Rail," "Books & Literature," and "Vegetarian Cuisine."[115] In theory, preventing the categorization and profiling of users into sensitive demographic groups at the outgoing vector reduces the risk of explicit ad discrimination or invasive behavioral manipulation. This is done by, for example, preventing advertisers from classifying users based on relationship status, race, or emotional state. Although the Topics taxonomy at present does not group users around sensitive topics, certain topics may still correlate statistically with demographic groups. Some actors may therefore seek to use Topics data to continue profiling using sensitive categories by exploiting "to sensitive topics."[116] Nevertheless, one could argue that the human-readable nature of the taxonomy is ultimately still a positive step towards transparency and that on-browser processing avoids sharing information with a huge range of third parties via TPCs—a net benefit from a privacy and surveillance perspective. Most importantly for Google, it preserves the viability of advertising revenues by introducing a new, ostensibly privacy-preserving version of "interest-based" advertising and therefore ensures continued financial support for free content online.[117]

However, the Topics API is not a satisfactory solution to the wicked problem of targeted advertising for two reasons. First, it ignores the tenet that wicked problems do not have singular solutions. The Topics

1529, 1539 (2021). One helpful illustration of digital distance in the SOGI context is the simple idea that, as argued by Cannon, "one's status as 'man' or 'woman' is determined by the extent to which their online behavior algorithmically conforms to the data categories 'man' and 'woman.'" *Id.* As such, "a gay 25-year-old woman may be differentially coded in various targeted marketing databases as a straight 45-year-old woman or a gay 18-year-old man as a result of her aggregate browsing habits, purchase history, and ad clickthrough rates." *Id.*

114. Goriunova, *supra* note 41, at 126.

115. As noted by the developers, a conscious decision was made to begin with a small taxonomy of around 350 topics rather than to use an industry standard such as the IAB Audience Taxonomy, which contains around 1500 topics, Github Draft Proposal, *supra* note 10.

116. *Id.*

117. *See* Vinay Goel, *Get to know the new Topics API for Privacy Sandbox*, GOOGLE THE KEYWORD (Jan. 25, 2022), https://blog.google/products/chrome/get-know-new-topics-api-privacy-sandbox/ (Google's initial press release recognizes "that advertising is critical for many businesses, and is a key way to support access to free content online.").

API seems to assume that increasing digital distance and reducing ad targeting granularity have unambiguously positive outcomes for the policy tradeoffs set out in Part III, which is not necessarily always the case. Ad delivery discrimination, for example, may still occur in a post-Topics world. An employer seeking to hire a software engineer may solely target ads toward users with an interest in "Programming," but the ultimate result would still be a gender-skewed distribution of ads due to the gender makeup of the wider body of coders.[118]

Instead, one could conceive of an alternative solution to the issue of ad discrimination that seeks to introduce *greater* granularity into the targeting ecosystem. As Frederik Borgesius notes, collecting certain sensitive data (in order to identify vulnerable groups and implement relevant protections) may be necessary to combat certain forms of manipulation, discrimination, or other algorithmic harm.[119] For example, discriminatory HEC ad delivery excluding non-binary users may occur because non-binary users are rendered entirely illegible within Meta's ad delivery infrastructure.[120] One approach is then to ensure that enough data is collected at the outgoing vector to subsequently allow incoming vector systems to audit themselves by monitoring (or estimating) the distribution of ads for a given campaign across different demographic groups; the system would reactively use this information to equalize ad distribution. Indeed, this seems to be the approach taken by Meta as a result of its latest settlement agreement with the U.S. DOJ.[121] By using publicly available U.S. census data to estimate race and ethnicity, Meta's Variance Reduction System for ad delivery is able to

118. *See Developer Survey*, STACK OVERFLOW (2022), https://survey.stackoverflow.co/2022/#overview (A survey of 70,000 software developers globally conducted by Stack Overflow found that 91.88% identified as male, 5.17% identified as female and 1.67% identified as non-binary, genderqueer, or gender non-conforming). The same is true in academia—a large-scale analysis of computer science literature found that, if trends from the past fifty years continue, gender parity in terms of paper authorship will not be reached this century, *see generally* Lucy Lu Wang et al., *Gender Trends in Computer Science Authorship*, 64 COMMC'NS OF THE ACM 78 (2021).

119. *See* Frederik Zuiderveen Borgesius, *Strengthening Legal Protection Against Discrimination by Algorithms and Artificial Intelligence*, 24 INT'L J. HUM. RTS. 1572, 1581 (2020).

120. Kingsley et al. notes that "[o]n the ad portal, if an advertiser specifies their advertisement is for housing, employment or credit, the only audience selection possible for gender is 'all.' Facebook defines 'all' as men and women. We do not know whether 'all' includes the gender group that Facebook calls 'unknown'" [which includes persons who identify as non-binary or specify a custom gender]." Kingsley et al., *supra* note 39, at 15.

121. Press Release, U.S. Dep't of Just., Justice Department and Meta Platforms Inc. Reach Key Agreement as They Implement Groundbreaking Resolution to Address Discriminatory Delivery of Housing Advertisements (Jan. 9, 2023), https://www.justice.gov/opa/pr/justice-department-and-meta-platforms-inc-reach-key-agreement-they-implement-groundbreaking.

audit an ad campaign to ensure ad distribution is not assigned in a discriminatory way.[122] It remains to be seen if this method will be sufficient to stop discriminatory incoming vector harms. Simultaneously, this approach also seems to sit uncomfortably with a surveillance-antagonistic outlook that views the drawing of racial, gender-based, or other demographic inferences about users, especially if this requires the processing of particularly sensitive data, as inherently problematic.

Increasing profiling granularity may then also prevent the use of personalization to reach marginalized groups, removing a key pathway not only for self-validation but even potential affordances for economic activism and resistance.[123] In extreme cases, a total ban on advertising based on sexual orientation or gender identity may prevent public interest organizations from delivering critical messages to target audiences.[124] And yet any infrastructure that facilitates the increased visibility of queer users to advertisers, enabling them to access empowering messages through advertising, simultaneously allows advertisers to divide queer "consumers" into a new, profitable market segment to facilitate the extraction of behavioral surplus. Finally, an increase in digital distance may also lead to a corresponding increase in mismatches between a person's preferred gender identity and the gender category they are assigned based on a statistical, machine learning-based model of gender.[125]

The Topics API is also flawed for a second reason: as a form of self-regulation, it implicitly privileges platform decision-making while promoting a vision of governance that requires Google to remain in

122. Roy L. Austin Jr., *An Update on Our Ads Fairness Efforts*, META (Jan. 9, 2023), https://about.fb.com/news/2023/01/an-update-on-our-ads-fairness-efforts/.

123. *See generally* Tsai, *supra* note 65.

124. When Facebook began blocking ad targeting by sexual orientation in 2018, activists at the Trevor Project (a U.S. non-profit organization focusing on suicide prevention for LGBTQ+ youth) found that they were unable to reach queer users to promote a national mental health survey. Alex Kantrowitz, *Facebook Has Blocked Ad Targeting By Sexual Orientation*, BUZZFEED (Mar. 21, 2018), https://www.buzzfeednews.com/article/alexkantrowitz/facebook-has-blocked-ad-targeting-by-sexual-orientation#.cnND1QvBz.

125. One response, drawing on Summers et al. and the study by Rinaldi discussed above, would be to argue that this is unlikely to have any real-world harm given that an algorithmic inscription of gender identity will always be "inaccurate" and thus have little psychological labelling function for users. And yet the very concept of accuracy presumes stability of identity. This may not hold true for those at the initial stages of exploring non-normative identities, and adolescents in particular—although the causal factors driving gender identity development are complex and not well understood, adolescence is a crucial period for the consolidation of gender identity, *see* Thomas D. Steensma et al., *Gender Identity Development in Adolescence*, 64 HORMONES AND BEHAVIOR 288, 290 (2013). Such a response also ignores the nuanced ways in which algorithmic mechanisms may have a chilling effect on self-expression, exploration, and preference-forming, see Part III above.

control of key infrastructures. Google thus regulates how and for what purposes data is collected, sorted, and categorized as it flows along outgoing vectors in the AdTech ecosystem. This Note does not seek to discount the validity of self-regulatory efforts. However, under conditions of pervasive information asymmetry, relying on self-regulation *alone* fails to provide meaningful space for individuals, civil society groups, regulators, researchers, and legislators to suggest policy changes or challenge the position that decreasing advertising accuracy is unambiguously positive.

There are important normative and political stakes here: decisions influencing how economic opportunities are distributed, how users are surveilled online, and even how we form our own identities require a more inclusive, participatory, and democratic form of decision-making,[126] one which does not assume that Google should have unilateral decision-making power. This argument does not entirely discard the valid benefits that a human-intelligible means of profiling at the outgoing vector could bring; it simply asserts that any single solution implemented in isolation is insufficient. As explored below, there are opportunities for Topics to interact with other multi-stakeholder interventions to targeted advertising—for example, Topics' focus on increasing transparency and user agency could interact with individual infrastructural approaches, allowing individual users to take back agency over how they are profiled and targeted online. Thus, a collective perspective may help inform a range of regulatory interventions that do not solely rely on platform goodwill to bring about meaningful change in the digital advertising ecosystem.

V. GOVERNANCE INTERVENTIONS FOR TARGETED ADVERTISING AS A WICKED PROBLEM

How can policymakers and users understand how to intervene in advertising systems when the scope and severity of harms are potentially vast but ill-defined? This problem is compounded by the fact that digital platforms operate in an ecosystem largely shielded from public view, retaining outsized private power over key infrastructures that support our digital economy. One criticism of Horst Rittel and Melvin Webber's original framing of the wicked problem paradigm is that they provide little analysis on how to manage such problems; as J.B. Ruhl and James Salzman note, "the theory-to-practice gap remains large."[127]

126. *See generally* Salomé Viljoen, *A Relational Theory of Data Governance*, 131 YALE L.J. 573, 573 (Nov. 2021).

127. Ruhl & Salzman, *supra* note 2, at 1582.

One approach is to adopt Salome Viljoen's proposal to treat data as a democratic medium, which conceives of data as a collective resource subject to democratic ordering, allowing populations to "balance the overlapping and, at times, competing interests that comprise the population-level effects of data production."[128] Although Viljoen's account of how to implement this vision in practice is fairly inchoate,[129] this approach has the correct normative orientation: all stakeholders should have a say in how digital advertising operates going forward. The challenge is that each of these stakeholders will have different preferences regarding how best to manage the competing tradeoffs surveyed in Part III. The nature of targeted advertising as a wicked problem means that it does not have a single, definitive "solution," but instead requires a range of potential governance interventions, undertaken at various levels. Some approaches may rely on individual users to exercise their agency, while others may require state interventions to protect collective interests; these interventions are likely to interact with each other in emergent and complex ways.

How might such a vision of stakeholder-driven, inclusive, tailored approaches to targeted advertising governance be implemented in practice? This Note argues that a necessary precondition for such approaches is to establish a collective perspective on how data flows along outgoing and incoming vectors, to form an accurate assessment of the scale, weight, and severity of harms and countervailing benefits, and how these harms and benefits are unevenly distributed across different populations. This requires a form of governance that I term "transparency as infrastructure regulation."

This Part proceeds as follows: Section V.A explains the basic pillars of transparency as infrastructure regulation, focusing on both the substantive data that would need to be made accessible under such a framework, while also setting out how institutions facilitating such transparency might be designed and created. This Part then turns to a discussion of other policy and legal interventions that might be taken in response to the findings generated by this increased transparency, including a complete ban on behavioral advertising (Section V.B); technical approaches to reassert individual control over ads (Section V.C); and the application of existing legal

128. Viljoen, *supra* note 126, at 638.

129. Viljoen states that actualizing data as a democratic medium requires us to "develop the institutional responses necessary to represent the relevant population-level interests at stake in data production," *id.* at 640, she gives a few examples of "protodemocratic data-governance proposals," such as calls in Germany for a national data trust, *id.* at 645.

frameworks, including privacy and data protection, non-discrimination law, consumer protection law, and horizontal AI regulation (Section V.D).

A. *Transparency as Infrastructure Regulation*

Transparency requirements for targeted advertising have begun to emerge in regulatory instruments seeking to govern online platforms. For example, the European Union's (EU) Digital Services Act (DSA), enacted in 2023, imposes a regulatory framework for the provision of online intermediary services by digital platforms.[130] As part of this framework, Article 39 of the DSA requires providers of very large online platforms or very large online search engines to compile and make publicly available an advertising repository.[131] These repositories catalog, for example, the content of the advertisement, the identity of the advertiser, and whether the advertisement was intended to be presented to specific groups, and if so the parameters used to target or exclude certain groups.[132] This repository must be publicly available, accessible (including through APIs), and searchable; at the same time, providers are required to exclude the personal data of any end users or consumers.[133] This advertisement repository is supported by standard-setting to ensure the interoperability of advertisement repositories maintained by different platforms.[134]

The DSA's ad repository requirement—given that it is clearly attuned to the discriminatory and exclusionary potential of behavioral advertising[135]—is a good start. A public ad repository would allow the public, researchers, and regulators to conduct studies and actively audit the delivery of ads to prevent outcomes that disproportionately impact or marginalize certain groups.[136] However, the DSA is limited in that it is

130. *See* Commission Regulation 2022/2065 of Oct. 19, 2022, on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act), 2022 O.J. (L 277) 1-3 [hereinafter DSA].

131. These are defined as platforms and search engines which 45 million or more average monthly active recipients of the service in the EU and designated as such by the Commission under the DSA. *See id.* arts. 33(1), 33(4).

132. *Id.* art. 39(2).

133. *Id.* art. 39(1).

134. *Id.* art. 44(1)(f).

135. *See id.* recitals 68, 69 (recognizing the risk of discriminatory presentation of advertisements and amplification of societal harms against marginalized groups).

136. *See id.* recital 95 (envisions public access to ad repositories facilitating "supervision and research into emerging risks brought about by the distribution of advertising online, for example in relation to illegal advertisements or manipulative techniques and disinformation with a real and foreseeable negative impact on public health, public security, civil discourse, political participation and equality.").

ultimately oriented towards the user's experience of advertising, and so focuses on how targeting operates at the incoming vector. It does not address how classification infrastructures (e.g., Google's Topics API) would work at the outgoing vector. In other words, the DSA does not require platforms to collect or publicize information on how users' collective browsing and app-based behaviors are monitored, categorized, and shared; how users are profiled or grouped as a result; and how these functions relate to and impact targeting practices.[137] This is a problematic blind spot, especially if there are concerns about infrastructural control over the "modes of prediction"[138] and power over knowledge. The Topics API's approach to human intelligibility is, therefore, a good first step toward transparency at the outgoing vector—but it does not establish any collective perspective about how groups of people are classified and profiled, or how these groupings change over time.

A collective, aggregate perspective is vital because it prevents a myopic focus on individual cases of harm and instead has the potential to reveal horizontal population-level harms that emanate from categorization processes.[139] High-level transparency at the infrastructural level enables an understanding of how targeted advertising contributes to known harms or even reveals previously unrecognized harms while making it possible to quantify the severity of these harms.[140] Conversely, it may also reveal that the extent, severity, or frequency of such harms has been greatly overstated in the theoretical literature and that the trade-offs gained as a result of increasing digital distance, for example, outweigh any of the imputed downsides. Before thinking about substantive legal or governance responses to issues of discrimination, surveillance, and even misgendering, it is necessary to have an informed perspective about whether such intervention is needed (and if so, *where* it is most needed along the AdTech data supply chain). The benefit of a transparency-focused approach to regulatory intervention is that it does not propose any single, reductive solution to these complex problems. Instead, it seeks to make these phenomena legible to provide an evidence-backed foundation for further tailored interventions.

Indeed, transparency as infrastructure regulation would be compatible with a self-regulatory approach such as Topics. For example, if the discriminatory delivery of ads to gender minorities is a concern, it is

137. *See* Gordon-Tapiero et al., *supra* note 4, at 124.

138. Cohen, *supra* note 47, at 1406.

139. *See generally* Viljoen, *supra* note 126, at 607.

140. *See* Gordon-Tapiero et al., *supra* note 4, at 127.

necessary to test whether introducing less granularity via Topics at the outgoing vector is more (or less) effective than collecting declared data about users' preferred gender identity (with their explicit consent), to audit and prevent gendered distributions in ad delivery.[141] To effectively compare these approaches, it is important establish a mechanism for meaningful transparency that exposes patterns and correlations between outgoing-vector content (e.g., the assignment of a given Topic to a user) to incoming-vector content at an aggregate level (e.g., the gender distribution of ads).[142] Indeed, a collective perspective may even provide the evidence base necessary for Topics to become an industry-wide browser standard by showing that an outgoing vector approach to decreasing advertising granularity can successfully mitigate some of the harms of behavioral targeting at the incoming vector. Transparency therefore allows infrastructural actors such as Google to publicly justify (and potentially promulgate) their self-regulatory efforts while simultaneously providing the public, civil society, regulators, and policymakers with tools to scrutinize these forms of self-regulation and ensure legal intervention where appropriate. In this way, transparency and the establishment of a collective perspective as a form of infrastructural regulation directly address power imbalances over data by lifting the veil of secrecy behind which many AdTech actors have operated to date, democratizing access to key information about how we are collectively categorized, classified, and targeted.

1. The Substantive Content of Transparency as Infrastructure Regulation

What information should be made transparent under such an approach? Incorporating many of the substantive transparency requirements on the incoming vector that the DSA deploys under Article 39(2) (content, the identity of the advertiser, targeting methods, populations reached, etc.) is a good place to start. However, as stated above, a collective perspective must also address transparency in the outgoing vector. The Topics API provides a helpful starting point because it sets out how its taxonomy assigns topics to individual users in an accessible and clear

141. Indeed, such testing is commonly used in marketing in the form of A/B testing—where two versions of a given variable are compared to assess which "performs" better. *See* Amy Gallo, *A Refresher on A/B Testing*, HARVARD BUS. REV. BLOG (June 18, 2017), https://hbr.org/2017/06/a-refresher-on-ab-testing.

142. *See* Gordon-Tapiero et al., *supra* note 4, at 126.

way (e.g., direct correlation with browsing history),[143] as well as how API callers can obtain a topic for a given user. However, for a true collective perspective, we would also need to understand where categorization and profiling data flow once the API is called. Legislation would need to mandate that browser operators disclose, for example, the identity and frequency of API callers, how that data is used to fuel profiling and personalization systems, and to which actors such data is passed onwards. Ideally, the transparency requirements would be designed in a way that enables a holistic tracing of collective data flows through the entire AdTech value chain, from the point of collection at browsers and mobile devices to the point of targeting.[144] Platforms may seek to leverage antitrust or trade secrecy laws to resist efforts to make the data flows and algorithmic practices that sit at the very heart of the AdTech ecosystem legible. Responses to these objections are discussed further below.

A collective perspective also needs to be established across as broad a set of the population as possible, in order to accurately ascertain population-level effects. In other words, when measuring gender-based ad discrimination, it would not be enough to show that some individuals were excluded. Instead, we would need to compare, for example, the relative display rate for a representative sample of male, female, and non-gender conforming populations.[145] A collective perspective would also need information on the effectiveness of such ads in shaping user behavior, as measured through proxies such as click-through rates, impressions, etc. However, the sort of mass data collection needed to generate a collective perspective may in fact raise many of the same concerns regarding population-level surveillance surveyed in Part III. There are two responses to such an objection. First, Big Tech actors are already collecting and processing this type of data; a collective perspective operating under the principle of data minimization would not need more datafication than that already occurring in the current ecosystem; it simply makes those practices legible. Indeed, a collective perspective may eventually reveal the invasiveness and severity of the harms arising from such population-level surveillance, prompting further policy or legal interventions to reduce datafication. Second, a range of technical tools can be used to ameliorate privacy concerns; these are discussed further below.

---

143. Note that the below discussion refers to the Topics infrastructure given the focus of this Note, but many of the same considerations would apply to other tracking and categorizing infrastructures, including TPCs.

144. *See generally* NORWEGIAN CONSUMER COUNCIL, *supra* note 19 (for an example of a report that conducts a data flow analysis at the outgoing vector by relying on data obtained through independent third-party data auditors and cybersecurity experts).

145. *See* Gordon-Tapiero et al., *supra* note 4, at 125.

2. The Institutional Aspect of Transparency

The DSA model seems to envisage that individual platforms will maintain their own public ad repositories. Researchers and regulators would then conduct audits and studies on ad repository data by searching on the platform's interface or calling an API, with the use of standards ensuring that data remains interoperable between platforms (thereby enabling a collective perspective across the entire digital economy, not limited to a specific platform). This is a valid decentralized approach, especially given that platforms will be the only actors that hold a broad perspective over both outgoing and incoming vectors.[146] However, policymakers may decide that a greater degree of trust is needed to establish a collective perspective and therefore may mandate the creation of a centralized body to collect and audit information about the relevant data flows.[147]

How would an institution tasked with facilitating a collective perspective be designed? A "thin" conception of this body may simply require the centralized intermediary to act as a focal point for data sharing, with the onus falling on the public, researchers, and regulators to subsequently interact with collective perspective data via APIs to audit platform activity. A "thicker" conception would require the intermediary body to actively study, scrutinize, and audit the practices of platforms once a collective perspective is established, and potentially apply regulatory sanctions where relevant. This approach might even require the body to eventually become a new supervisory regulator, endowed with legal powers of enforcement and oversight. The choice between these two models heavily depends on the political economy of the relevant jurisdiction, along with relevant public law considerations.

Crucially, any centralized body would need unencumbered access to the relevant data.[148] A range of technical tools could be used to ameliorate any potential privacy concerns, such as local differential privacy or

146. *Id.* at 126.

147. *EU register of data intermediation services*, EUROPEAN COMMISSION (Apr. 1, 2025), https://digital-strategy.ec.europa.eu/en/policies/data-intermediary-services (This centralized body need not be a public entity. For example, the EU Data Governance Act aims to encourage the creation of a new business model of "data intermediation services" – under this model, "data intermediaries" are "neutral third parties that connect individuals and companies with data users", facilitating the trusted pooling and sharing of data); *Id.* (The Act is not prescriptive as to the legal personality of these entities; many of the data intermediaries currently registered with the Commission are private companies).

148. *See* Gordon-Tapiero et al., *supra* note 4, at 126.

secure multiparty computational tools,[149] bolstered by contractual or statutory safeguards providing relevant remedies to data subjects in the event of data breaches. Data erasure is an important mechanism for reducing the severity of any potential data breaches or even potential misuse of collective perspective data; legislation may mandate that relevant data be erased by the intermediary body after a certain period of time after an ad is shown for the last time on a platform. At the same time, this period cannot be too short as longitudinal data about how classification, categorization, and targeting systems operate and shift over time is extremely valuable. If, for example, we want to understand how these infrastructures are reshaping our idea of gender expression, any study would have to span a period of at least several years.

A fully transparent, publicly accessible collective perspective would simultaneously need to grapple with any competitive concerns platforms raise, along with trade secrecy questions. For example, a collective perspective would enable a sophisticated analysis of which types of ads perform best for targeting certain groups. Advertisers may argue that novel marketing campaigns could then be easily appropriated by competitors to poach customers, creating a free-rider problem. Platforms may also seek to protect any explanation of the dynamics underlying algorithmic profiling under intellectual property and trade secrecy laws.[150]

A segmented approach to collective perspective data access could mitigate these concerns. For example, certain information such as the identity of the advertiser, targeting method, and behavioral or other contextual label assigned to the user during the profiling process can always be made public, while more detailed commercial information (e.g., regarding click-through rates and impressions for certain ad campaigns) could be made available through a secure access API. Another option, inspired by the Digital Marketing Act's (DMA's) approach to pricing transparency along the AdTech value chain,[151] is to require the consent of the relevant party before disclosing sensitive, commercially valuable information. This is where a collective perspective becomes particularly useful. Even if individual platforms do not consent to the

149. Local differential privacy allows computation of aggregate statistics (e.g., correlation between an ad being shown and a person's gender identity) on the basis of personal data to which a large but controlled amount of random noise has been added. Secure multiparty computational tools would distribute the duties related to the collective perspective among a few trusted parties following a specific protocol, none of which would be able to corrupt the computation or gain inappropriate access to personal information). *Id.*

150. *See* Paul B. De Laat, *Algorithmic Decision-Making Employing Profiling: Will Trade Secrecy Protection Render the Right to Explanation Toothless?*, 24 ETHICS AND INFO. TECH., Apr. 5, 2022, at 1.

151. *See* DMA, *supra* note 63, at 34, 40.

disclosure of such information for a given set of ads, a collective perspective would still be able to provide users, researchers, and regulators with, for example, the average click-through rate for ads with similar content, or ads using similar targeting variables. Standardization and the development of codes of conduct governing data sharing for creating a collective perspective would also be helpful instruments in ameliorating these concerns.[152] For example, researchers may need to acquire a standardized accreditation before being allowed to access sensitive commercial data via a secure access API. This standardized accreditation envisages a participatory process that enables industry actors to give input on best practices regarding data sharing and an informed perspective on which categories of persons should be allowed access to certain data and on what conditions.

3. The Legal, Regulatory, and Governance Implications of Transparency as Infrastructure Regulation

Some might argue that transparency regulation has no teeth because it does nothing substantive about either infrastructural control or the underlying political power that comes from continuing to privilege private decision-making over vital categorization questions about sensitive categories of data such as gender and sexual identity. My response is that the form of transparency regulation envisaged above is not a thin, check-box disclosure exercise: it entails substantive and meaningful data sharing about how a traditionally black-box system operates and is a necessary first step to gaining an accurate picture of the scale, frequency, and severity of harms that may arise as a result. Longitudinal, evidence-driven examinations of such harms would then provide a foundation for the application of existing legal frameworks or even the creation of new legal claims, remedies, and governance strategies.

This model of transparency as a form of AdTech regulation has the benefit of pragmatism—it is on the opposite side of the scale of a total ban on targeted advertising (as discussed further in Section V.B), and is therefore unlikely to encounter the same fierce resistance from Big Tech; it would be difficult for Google and other actors to object to such a measure (although, as discussed above, they may seek to leverage trade secrecy laws to mitigate against mandatory data sharing and disclosure). Politically, it may be easier to pass such transparency legislation given that there is already an existing framework in the form of the DSA. A collective perspective transparency legislation could be framed

152. *See, e.g.*, DSA, *supra* note 130, art. 46.

as a DSA+. At the same time, it is incorrect to say that such transparency comes without teeth. In 2020, the Norwegian Consumer Council conducted a detailed study on the advertising and data-sharing practices of the LGBTQ+ dating app Grindr, which required significant technical testing and auditing by private cybersecurity companies.[153] Its findings revealed that Grindr shared detailed user data (including IP addresses, GPS location, age, and gender) with a large number of third parties involved in advertising and profiling; it argued that many of these opaque data-sharing practices were in violation of the General Data Protection Regulation (GDPR) due to a lack of a valid legal basis, which the GDPR requires for the sharing of data.[154] The Norwegian Data Protection Authority eventually fined Grindr EUR 6.5 million.[155]

In short, a baseline level of transparency achieves three potential outcomes. First, it discourages shady data-sharing practices and encourages the development of *ex ante* responsible data governance practices due to the threat of public scrutiny and/or regulatory oversight. Second, it highlights conduct that is potentially illegal under existing legal frameworks such as data privacy, non-discrimination, and consumer protection, and therefore opens the avenue for litigation or regulatory enforcement, as further explored in Section V.D. Finally, it potentially enables the development of novel legal claims or regulatory interventions to deal with new harms that come to light.

B. *Complete Ban*

A collective perspective may ultimately lead to the conclusion that behavioral advertising is so deeply manipulative that the potential risks for (*inter alia*) unlawful surveillance, political malfeasance, and direct discrimination greatly outweigh any arguments for the continued existence of the mostly free online services we rely on today. The suggestion to ban targeted advertising is not new;[156] however, it has received

153. NORWEGIAN CONSUMER COUNCIL, *supra* note 19, at 8-10.

154. *Id.* at 11. *See generally* Commission Regulation 2016/679 of Apr. 27, 2016, General Data Protection Regulation, O.J. (L 119) [hereinafter GDPR].

155. *Norwegian DPA imposes fine against Grindr LLC*, EUROPEAN DATA PROT. BD. [EDBP] (Dec. 21, 2021), https://edpb.europa.eu/news/national-news/2021/norwegian-dpa-imposes-fine-against-grindr-llc_en.

156. *See* David Dayen, *Ban Targeted Advertising*, THE NEW REPUBLIC (Apr. 10, 2018), https://newrepublic.com/article/147887/ban-targeted-advertising-facebook-google; Gilad Edelman, *Why Don't We Just Ban Targeted Advertising*, WIRED (Mar. 22, 2020), https://www.wired.com/story/why-dont-we-just-ban-targeted-advertising. *See generally* Jeff Gary & Ashkan Soltani, *First Things First: Online Advertising Practices and Their Effects on Platform Speech*, KNIGHT FIRST AMEND. INST. (Aug. 21, 2019), https://knightcolumbia.org/content/first-things-first-online-advertising-practices-and-their-effects-

renewed interest in recent years. For example, in January 2022, U.S. legislators proposed the Banning Surveillance Advertising Act.[157]

A total ban on targeted ads is a very blunt tool, but proponents would argue that more selective bans on targeting based on certain protected characteristics, such as those in Article 26 of the DSA,[158] focus on restricting incoming vector harms in the delivery of ads to certain marginalized or vulnerable groups but do not fundamentally change the economic incentives driving datafication at the outgoing vector. A total ban on targeted behavioral advertising would therefore remove the financial incentive underlying mass data collection practices that fuels many of the harms above.[159]

Reaching the political consensus needed for a blanket ban would be challenging given the huge popularity of free online services. However, such a ban would not totally destroy the free digital ecosystem as we currently know it. As explored in Part III, there are alternative methods of ad targeting that could supplement lost publisher revenues if targeted behavioral and demographic advertising is banned, such as targeting based on the type of website visited. However, assuming that advertisers will be less willing to pay for those options, a ban may lead to an ecosystem where only the largest platforms can continue to provide free services (and therefore sustain massive user bases, allowing them to

on-platform-speech. (in the U.S. context in particular, it has been argued that banning microtargeted ads may be a more effective in aligning platforms' practices with public values than content moderation laws, given the potential impact on First Amendment speech protections). On platform incentives, *see* Cory Doctorow, *The 'Enshittification' of TikTok,* WIRED (Jan. 23, 2023), https://www.wired.com/story/tiktok-platforms-cory-doctorow/(suggesting that platforms are incentivized to create appealing products for users to create a large user base, who are then subsequently exploited for the benefit of the platform's business customers (including advertisers); once business customers are locked into the ecosystem, they are then exploited to recoup profits for the platform).

157. Banning Surveillance Advertising Act, H.R. 6416, 117th Cong. (2022). The EFF has supported such legislation in part as a result of the growing evidence of the potential for ad targeting to have discriminatory impacts on gender, race, and other sensitive characteristics, Bennett Cyphers & Adam Schwartz, *Ban Online Behavioral Advertising,* EFF (2022), https://www.eff.org/deeplinks/2022/03/ban-online-behavioral-advertising.

158. Article 26(3) of the DSA prohibits the presentation of ads based on profiling that uses the special categories of data set out in the GDPR. These include, inter alia, data revealing race, political opinions, health, religious beliefs, and sex life or sexual orientation. DSA, *supra* note 130, art. 26(3).

159. Indeed, a ban on such advertising aligns with an understanding of targeted advertising as an infrastructure—if we conceive of e.g., the continued existence of a free Google search as akin to a public utility, we could argue that "no essential infrastructure should be surveillance-based or funded by targeted ads." K. Sabeel Rahman & Zephyr Teachout, *From Private Bads to Public Goods: Adapting Public Utility Regulation for Informational Infrastructure,* KNIGHT FIRST AMEND. INST. (Feb. 4, 2020), https://knightcolumbia.org/content/from-private-bads-to-public-goods-adapting-public-utility-regulation-for-informational-infrastructure.

continue reaping the benefits of mass data collection), disproportionately disadvantaging smaller platforms. These smaller platforms may eventually be driven out of the market. Any political arguments in favor of a total ban would need to speak to such critiques. Advocates for a total ban must also genuinely engage with arguments that the rise of tiered subscription services might in fact discriminate against low-income users. Thus, the state may need to intervene by subsidizing low-income users' use of essential internet services. Political contestation over which internet services are considered "essential" would likely follow. Platforms are likely to mobilize such concerns to resist a broad ban. Indeed, as a practical matter, the sheer political power Big Tech wields—as a result of its importance in the broader economy and outsized financial power—is likely to severely hamper any legislation banning targeted advertising.[160] As such, even if a collective perspective is reached, a complete ban on behavioral advertising may be unlikely in reality.

Nevertheless, even if such a ban never materializes, a collective perspective established via the transparency framework outlined above would help encourage such debates in legislative and other public fora. A collective perspective in turn would affect policy discussions regarding other regulatory interventions. For example, in the absence of a total ban, increased political scrutiny of targeted advertising may lead some users or groups of users to decide (based on the information gained from a collective perspective facilitated by transparency regulation), that the risk of surveillance and behavior manipulation outweighs any benefits they gain from personalization.

C. *Individual Technical Approaches*

A collective perspective gives users the ability to exercise individual agency over their data. By understanding the benefits and risks of datafication, users are able to make an informed decision about the optimum distribution of tradeoffs for their particular context.

Individual decision-making is facilitated by a range of technical tools. At the outgoing vector, many browsers and devices contain features to

160. In the United States, Apple, Amazon, Google and Facebook spent over USD 55 million on lobbying the federal government in 2021, with Google's spending having increased 27% from the previous year. Emily Birnbaum, *Tech spent big on lobbying last year*, POLITICO (Jan. 24, 2022), https://www.politico.com/newsletters/morning-tech/2022/01/24/tech-spent-big-on-lobbying-last-year-00001144.

modify or even disable cross-site and cross-app tracking: Google Chrome has a feature allowing users to choose to block TPCs,[161] while iOS's App Tracking Transparency (ATT) feature allows users to decide which apps are allowed to track activity across other apps and websites for advertising purposes.[162] Although Apple has come under increasing scrutiny due to anti-competitive concerns surrounding ATT,[163] it is clear that the implementation of ATT has had a significant impact on ad accuracy.[164] Applying such tools would result in net benefits for users who place less utility on the personalization benefits of accurate targeting. At the incoming vector, Google allows users to customize their ad experience to choose ad topics, limit ads concerning sensitive topics such as alcohol or gambling, or turn off personalization.[165] As alluded to above, Google's Topics API would increase user agency by allowing profiling to become human-legible, enabling users to delete irrelevant topics (or disable the Topics API entirely). Researchers have recently proposed "middleware" software that enables an even greater degree of control over incoming vector content by allowing users to choose not just the types of content they want to see but also how such content should be ranked and which content providers they prefer.[166]

One limitation of these technical approaches is that they greatly depend on users' awareness of the possibility and extent of harms generated. As discussed above, users generally have limited knowledge regarding how targeted advertising works in practice—this then impedes their ability to exercise control over their data.[167] A collective perspective would

161. *Clear, allow & manage cookies in Chrome*, GOOGLE CHROME HELP (last visited Feb. 28, 2025), https://support.google.com/chrome/answer/95647?sjid=8949812564882832830-NA#zippy=%2Callow-or-block-cookies.

162. *If an app asks to track your activity*, APPLE (March 17, 2025), https://support.apple.com/en-us/HT212025.

163. *Privacy by default, abuse by design: EU competition concerns about Apple's new app tracking policy*, HAUSFELD (May 25, 2021), https://www.hausfeld.com/en-us/what-we-think/competition-bulletin/privacy-by-default-abuse-by-design-eu-competition-concerns-about-apple-s-new-app-tracking-policy/.

164. Indeed, the effect on ad accuracy was so significant that it caused Snap, Facebook, Twitter and YouTube to lose an estimated 12% of their advertising revenues when ATT was introduced in 2021. Patrick McGee, *Snap, Facebook, Twitter and YouTube Lose Nearly $10bn after iPhone Privacy Changes*, FINANCIAL TIMES (Oct. 31, 2021), https://www.ft.com/content/4c19e387-ee1a-41d8-8dd2-bc6c302ee58e.

165. *Customize your ads experience*, MY AD CENTER HELP (last visited Feb. 25, 2025), https://support.google.com/My-Ad-Center-Help/answer/12155451?hl=en.

166. Gordon-Tapiero et al., *supra* note 4, at 676-677. Although this approach seems mainly geared towards non-commercial, algorithmically-recommended content, this could be easily tailored to the advertising context.

167. Boerman et al., *supra* note 89, at 368.

certainly raise the visibility of these harms (especially if coupled with increasing political engagement over a ban on targeted ads), but it would be naive to assume that transparency alone will lead to a wholesale change in individual approaches to behavioral advertising, even if the collective perspective showed that the harms vastly outweighed any benefits. Studies have found that only a minority of consumers actually take actions to block cookies or prevent cross-site tracking.[168] Whatever level of privacy protection is baked into default settings then exhibits a high degree of stickiness; few consumers have the time, energy, or inclination to actively research and apply new technical approaches. Furthermore, individual action is constrained by scope in that it does not look to the collective nature of data and so cannot fully prevent all incoming vector harms.[169] One person's decision to block cookies does little to prevent another person from being discriminated against in ad delivery.

Nonetheless, a collective perspective has a crucial role to play here in providing a baseline level of knowledge, giving privacy-conscious or particularly vulnerable groups the information they need to configure the relevant tradeoffs as they see fit, based on their individual circumstances. At the same time, a collective perspective is also compatible with a recognition that individual agency is not a catch-all solution to the wicked problem of targeted advertising. Such a perspective may eventually reveal that some issues are too politically, economically, and socially important to leave to individual choice. This is where the state, public law, and regulation must step in.

D. *Existing Legal Frameworks*

An in-depth application of existing legal frameworks to a fully transparent targeted advertising ecosystem is beyond the scope of this Note. However, some initial evaluations can be drawn regarding the affordances and limitations of four areas of existing law: (1) privacy and data protection, (2) non-discrimination, (3) consumer protection, and (4) horizontal AI regulation. This analysis then reveals potential ways in which a collective perspective may instigate reform of existing legal frameworks to more effectively intervene within the AdTech ecosystem. This analysis is not limited to a particular jurisdiction, although much of the below draws on regulation and jurisprudence from the EU given its prolific activity in digital regulation. Comparative examples are also drawn from the United States, Brazil, and Nigeria, amongst others.

168. *Id.*

169. Gordon-Tapiero et al., *supra* note 4, at 125.

1. Privacy and Data Protection

Privacy and data protection law are a natural starting point given the history of its application to AdTech platforms. For example, the Irish Data Protection Commission issued a EUR 390 million fine against Meta in January 2023 for a lack of a lawful basis for the processing of user data for targeted advertising.[170] However, enforcement of the GDPR is hampered by a lack of transparency regarding data sharing and processing practices in the digital economy. As the Norwegian Consumer Council's investigation into Grindr demonstrates, a collective perspective may in fact reveal that many existing datafication practices in the AdTech space are fundamentally incompatible with privacy regulations. Privacy is particularly relevant for surveillance-focused advocates who criticize the legitimacy of platforms' processing of sensitive personal data.[171] However, if we are looking for data privacy to intervene when a sensitive inference is made and users are categorized based on those inferences, Sandra Wachter has argued that European Court of Justice judgments and academic literature suggest that "data controllers must both *intend* to draw sensitive inferences and use source data which provides a *reliable basis* to learn about sensitive data."[172] Arguably, these thresholds are fundamentally incompatible with the way in which advertising classification systems work in practice given that advertisers often do not intend to infer sensitive details about particular users while simultaneously having a high tolerance for classification errors.[173]

A collective perspective may create the impetus necessary to reform this aspect of privacy law, enabling it to better respond to the realities of how these infrastructures work in practice. Conversely, if a collective perspective shows that platforms are inferring sensitive attributes at the outgoing vector at a much lower frequency than previously thought (whether due to

170. *Data Protection Commission announces conclusion of two inquiries into Meta Ireland*, DATA PROTECTION COMMISSION (Jan. 4, 2023), https://www.dataprotection.ie/en/news-media/data-protection-commission-announces-conclusion-two-inquiries-meta-ireland.

171. *See* Sandra Wachter, *Affinity Profiling and Discrimination by Association in Online Behavioral Advertising*, 35 BERKELEY TECH. L.J. 367, 380 (2020). For example, European data protection law has enhanced requirements for the processing of certain "special category" data, as enshrined in Article 9 of the GDPR. However, these do not include sex or gender identity—and as such may not necessarily apply to some of the more complex questions around misgendering and identity-formation explored above. *Id.*

172. *Id.* at 371 (emphasis added).

173. *Id.* "The opportunity cost of showing ads intended for women to men that have been misclassified is very low. The business model of OBA can tolerate relatively high rates of misclassification. This tolerance does not, however, benefit misclassified users who are offered inaccurate or discriminatory content and may suffer as a result." *Id.* at 384.

regulatory pressure, self-regulation, or other incentives), this may suggest that political capital would be better spent advocating for legal reforms that focus more closely at the incoming vector (e.g., preventing discriminatory outcomes that occur independently of how users are profiled).

One particularly deep limitation of data protection is its regulatory design. For example, sensitive personal data can continue to be processed under Article 9 of the GDPR if the explicit consent of the data subject is acquired.[174] It is fairly well established now that consent is not a meaningful hurdle for data processors to clear;[175] indeed, it is also unclear whether "explicit" consent is distinct from the GDPR's ordinary consent standard.[176] Further, the data privacy provisions apply only to personal data.[177] Arguably, these phenomena stem from the individual, rights-oriented approach of privacy law, which focuses on the individual's ability to control data flows as opposed to horizontal issues of population-level inequalities.[178] If a collective perspective reveals fundamental harms that arise in the outgoing classification process, a deep look into the regulatory structure and incentives created by the GPDR and similar legislation is vital.

The insights from the above analysis apply far beyond the EU. Because of the so-called "Brussels effect" of the GDPR,[179] countries around the world have adopted key elements of the regulatory architecture and design of the Regulation. This is particularly the case in the Global South—Patricia Boshe and Carolina Caride argue that the extraterritorial reach of the GDPR (by virtue of Article 3) and its approach to cross-border data flows has meant that a huge range of countries across Africa, Latin America, and the Caribbean have strong incentives to align

174. GDPR, *supra* note 155, art. 9(2)(a).

175. Wachter, *supra* note 173, at 420.

176. Michael Veale & Frederik Zuiderveen Borgesius, *Adtech and Real-Time Bidding under European Data Protection Law*, 23 GERMAN L.J. 226, 242 (2022).

177. Although there have been arguments that, in the context of mass data collection, preventing re-identification is almost mathematically impossible, meaning that the GDPR eventually becomes a "law of everything." *See generally* Nadezhda Purtova, *The Law of Everything: Broad Concept of Personal Data and Future of EU Data Protection Law*, 10 L., INNOVATION AND TECH. 40, 43 (2018).

178. For example, a right to erasure has little impact if data subjects do not exercise their rights in significant numbers. A privacy focused approach then suffers from many of the collective action problems that individual technical approaches face. *See* Angelina Fisher & Thomas Streinz, *Confronting Data Inequality*, 60 COLUM. J. TRANSNAT'L L. 829, 901-02 (2022) (discussing the paradox of privacy in particular). However, there is the potential for these collective action problems to be mitigated by the institutional infrastructure created by the GPDR. *Id.*

179. *See generally* ANU BRADFORD, THE BRUSSELS EFFECT: HOW THE EUROPEAN UNION RULES THE WORLD (2020).

their data protection frameworks with that of the GDPR to sustain cross-border data flows and promote digital trade.[180] To take just two examples from large Global South economies—Brazil's Data Protection Law of 2019 and the Nigerian Data Protection Act of 2023 both mirror core aspects of the GDPR.[181] In particular, they both allow for the processing of "sensitive" categories of personal data,[182] provided the data subject has consented.[183] These two legal frameworks apply substantive legal protections to over 400 million people. Policymakers and legislators globally who are interested in regulating the impact of algorithmic advertising on vulnerable groups will need to grapple with the path dependencies that result from the importing of essential pillars of the GDPR (particularly its individual rights-oriented, consent-centric framework).

2. Non-discrimination

If a collective perspective demonstrates serious demographic discrepancies in the delivery of certain ads, or evidence that targeting practices are being adapted in order to spread hate speech or other discriminatory content, non-discrimination legislation may then intervene. Both the United States and the European Union have legislation preventing discrimination in the delivery of certain advertisements.[184] This form of legislative intervention has been particularly successful in the United States, as the housing ad discrimination lawsuits brought against Meta have revealed, as discussed in Part III above. Nevertheless, there are several formal legal barriers to discrimination claims. First, discrimination in the behavioral advertising context is particularly complicated because one may seek to challenge an inferred categorization that is actually incorrect. For example, a cisgender man may have received discriminatory treatment because they were misclassified as female. Consequently, they

180. *See generally* Patricia Boshe & Carolina Goberna Caride, *Is the Brussels Effect Creating a New Legal Order in Africa, Latin America and the Caribbean?*, TECH. AND REGUL. 12 (Mar. 18, 2024). The fundamental question of whether this outcome is normatively desirable from a political, distributive, or economic perspective is an entirely separate one that is outside the scope of this article.

181. *Lei Geral de Proteção de Dados Pessoais (Redação dada pela Lei n° 13.853, de 2019)* [*Brazilian General Data Protection Law (As amended by Law No. 13,853 of 2019)*], PRESIDÊNCIA DA REPÚBLICA, https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm (Braz.) [hereinafter *LGPD*]; Data Protection Act (2023) Cap. (A719) (Nigeria).

182. This largely mirrors the taxonomy of sensitive categories of data under the GDPR.

183. *LGPD*, *supra* note 183, art. 7(I); Nigeria Data Protection, *supra* note 183, at art.30.

184. In the United States, relevant legislation focuses on housing, employment, and credit ads, and includes the Civil Rights Act 1964, Fair Housing Act, and Equal Credit Opportunity Act. In the EU, there are a huge range of sector and industry-specific anti-discrimination protections; for an overview, *see* Wachter, *supra* note 173, at 387–90.

do not belong to any protected group *per se*.[185] The law currently does not recognize protected legal status for affinity groups or groups based on inferred interests.[186] As such, the law may not capture certain forms of targeting such as lookalike modeling even if they create discriminatory effects.[187] Second, where anti-discrimination laws require litigants to bring claims in court, members of a recognized protected group may need to "out" themselves to receive legal protection.[188]

To resolve these issues, Wachter argues for the development of the concept of "discrimination by association," where claimants are able to claim the benefit of being a member of a protected group "by association."[189] A collective perspective could help these legislative proposals gain political traction if discrimination by algorithmic association can be repeatedly shown to occur in practice. One of the hurdles to addressing online discrimination is that it is generally much harder to observe at the individual level because we have no insight into how our interests and behaviors (as represented by clicks, browsing habits, etc.) are translated into groupings based on protected characteristics.[190] We are also unable to compare at an individual level whether the ads we are delivered are meaningfully different from the ads delivered to others.

However, one may observe that advertising discrimination legislation would not apply directly to outgoing vector infrastructures such as Google's Topics API, given that non-discrimination is squarely incoming-vector focused. This lacuna may be a problem if we retain an intuitive sense of discomfort regarding data collection for the purpose of segmentation along demographic lines. Even if non-discrimination legislation incentivizes platforms to prevent gender-skewed distributions in the delivery of ads for computer science jobs, we may still feel uncomfortable with gender segmentation practices seeking to market certain toys to female minors, on the grounds that these reinforce gendered assumptions regarding appropriate standards of dress, vocation, etc.[191] Such practices would not be covered under discrimination legislation given that this simply relates to the

185. Wachter, *supra* note 173, at 373.

186. *Id.* at 371.

187. Gordon-Tapiero et al., *supra* note 4, at 122.

188. Wachter, *supra* note 173, at 373.

189. *Id.*

190. *Id.* at 406: "In the online world, we often do not know the rules and attributes on which we are profiled and whether these attributes correlate with protected characteristics. We cannot be sure how our interests (e.g., music) correlate with protected attributes (e.g., gender), as we lack a full causal model of the world that would show us how this data relates to each other."

191. *See* Katie Powers, *Shattering Gendered Marketing*, AM. MKTG. ASS'N (Sept. 3, 2019), https://www.ama.org/marketing-news/shattering-gendered-marketing/.

offering of a product as opposed to a specific economic opportunity. There are two potential responses to this. First, the imposition of legal sanctions on platforms that facilitate discriminatory incoming vector targeting could create a corresponding compliance effect at the outgoing vector through internal industry pressure. In turn, this would lead to self-regulatory efforts such as Topics, which seek to introduce greater coarseness into behavioral tags and prevent profiling based on sensitive categories. Second, non-discrimination approaches could be combined with consumer protection-oriented approaches that seek to prohibit such targeting of sensitive groups entirely, as discussed further below.

3. Consumer Protection

Another legislative intervention is provided by consumer protection law. One example of a consumer protection-oriented approach can be found in the DSA's prohibition on targeting advertisements based on profiling that uses the special categories of personal data referred to in Article 9 of the GDPR,[192] or where the provider is aware with reasonable certainty that the user is a minor.[193] The benefit of this approach is not only that it prevents discriminatory harms from targeting based on sensitive characteristics at the incoming vector, but that it also may suppress incentives to infer and re-inscribe sensitive traits at the outgoing vector. However, such regulation is difficult to enforce without transparency regarding how targeting criteria are deployed for ad delivery in practice. A collective perspective is therefore vital to ensure both outgoing and incoming vectors are aligned with the regulatory objective to preclude subtle harms surrounding self-expression and identity formation.

Two objections may be levied at a consumer protection-oriented approach. First, targeting bans do not prevent discrimination, as explored in Part III.A.2. However, the dual application of both non-discrimination and consumer protection law easily solves this issue. A collective perspective would allow the state to monitor regulatory outcomes and continue to implement incremental, tailored interventions going forward. For example, if regulatory pressure continues to lead to self-regulatory audit tools such as Meta's Variance Reduction System, and such systems have promising effects on combating discrimination, a collective perspective enables transparent and public A/B testing instead of leaving critical decisions regarding non-discrimination solely to Meta's engineers and product managers.[194] The second argument is that limited targeting bans

192. DSA, *supra* note 130, art. 26(3).

193. *Id.* art. 28.

194. *See* Gallo, *supra* note 141 (for an explanation of A/B testing).

grounded in consumer protection still fail to directly intervene at the outgoing level of inference, classification, and subsequent profiling. In other words, even if users are not targeted for ads based on certain characteristics, there is nothing to stop platforms from continuing to seek to classify consumers into such groups for other purposes (e.g., onward sale of data to hedge funds, credit brokers, or political analysts). However, one could imagine a collective perspective engendering change in this area of law as well. For example, a progressive legal approach would recognize not just that incoming targeting based on protected characteristics is harmful and therefore should be banned, but also that potential categorization practices that seek to group users based on inferred sensitive categories are *per se* unlawful, either due to the risk that such data may be sold or leaked via a data breach, or even based on the simple normative assertion that one's gender, race, or another sensitive characteristic should not be legible to platforms. A collective perspective would be vital to gathering the necessary evidence base for introducing progressive legislative change.

4. Horizontal AI Regulation

One relatively under-explored area in the literature is how the advent of new "horizontal" models of AI regulation, such as the EU Artificial Intelligence Act (AI Act),[195] will shape the algorithmic advertising landscape. Ostensibly, the Act will have little impact on algorithmic advertising at the incoming vector. Kai Zenner, head of staff for Axel Voss, a Member of the European Parliament and a key proponent of the Act, has noted that broad issues around algorithmic advertising and marketing were deliberately left vague in order to make the Act "future-proof."[196] Currently, the bulk of regulatory obligations are concentrated around "high-risk" AI systems; prima facie, ad profiling and delivery systems are currently not included in this categorization (set out in Annex III of the Act).[197]

However, there is some preliminary analysis suggesting that the Act may operate to protect users against commercial profiling where such practices exploit the vulnerabilities of individuals by "materially

195. Commission Regulation 2024/1689 of June 13, 2024, Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828, 2024 O.J. (L) 22 [hereinafter AI Act].

196. Caitlin Andrews, *Marketing Sits in a Gray Zone Under EU AI Act,* IAPP (June 12, 2024), https://iapp.org/news/a/at-aigg-2024-marketing-sits-in-a-gray-zone-under-eu-ai-act.

197. AI Act, *supra* note 198, annex III.

distorting" their behavior, or where such profiling uses subliminal, manipulative, or deceptive techniques.[198] Given the AI Act's phased implementation timeline,[199] on-the-ground enforcement experience has yet to emerge; these novel arguments remain untested for now. At a minimum, given the complexity of measuring the empirical impacts of algorithmic advertising (as discussed above), there will be serious evidential hurdles to clear. A collective perspective would help to create the necessary evidence base to support any argument that commercial profiling "materially distorts" consumer behavior. However, the way in which the Act has been designed relies on fixed, legible categories of risk. The "wicked problem" of behavioral advertising means that it cannot be readily bucketed into such legible categories of risk, given that the risks and harms of algorithmic advertising differ widely across populations.

The creation of a collective perspective of the harms and benefits of algorithmic advertising has never been more pressing, as countries across the world (including many in the Global South) begin legislative debates around the need for horizontal regulation of AI systems. At present, the potential for the AI Act to achieve a widespread, de jure "Brussels effect" across the world remains highly contentious.[200] However, there are already preliminary indications that large Global South economies have begun to import aspects of the AI Act's regulatory model. For example, Brazil's proposed AI regulation, Bill 2338/2023, is heavily inspired by the AI Act's tiered, risk-based approach to regulation, and features a list of high-risk systems that reflects many of the same categories found in Annex III of the AI Act.[201] As such, the proposed Brazilian law will likely face many of the hurdles identified above with respect to the AI Act, insofar as it will be difficult to cleanly bucket algorithmic advertising into an identifiable risk category. The creation of a collective perspective may inform ongoing legislative debates around the world, supporting the creation of

198. *See generally* Eline L. Leijten & Simone van der Hof, Dissecting the Commercial Profiling of Children: A Proposed Taxonomy and Assessment of the GDPR, DSA and AI Act in light of the Precautionary Principle (Dec. 13, 2024) (unpublished manuscript).

199. *See Implementation Timeline*, EU ARTIFICIAL INTELLIGENCE ACT (Aug. 1, 2024), https://artificialintelligenceact.eu/implementation-timeline/.

200. *Compare* Charlotte Siegmann & Markus Anderljung, *The Brussels Effect and Artificial Intelligence: How EU Regulation Will Impact the Global AI Market,* CTR. FOR THE GOVERNANCE OF AI (2022), *with* Ugo Pagallo, *Why the AI Act Won't Trigger a Brussels Effect*, AI APPROACHES TO THE COMPLEXITY OF LEGAL SYS. (forthcoming 2024), https://ssrn.com/abstract=4696148.

201. For an analysis of the latest draft text of the Bill, *see* Rafael Zanatta & Mariana Rielli, *The Artificial Intelligence Legislation in Brazil*, DATAPRIVACYBR (Oct. 12, 2024), https://www.dataprivacybr.org/en/the-artificial-intelligence-legislation-in-brazil-technical-analysis-of-the-text-to-be-voted-on-in-the-federal-senate-plenary/.

new legislative frameworks that seek to regulate harmful profiling and targeting practices in a more agile way, without relying on fixed risk-based categorization approaches that may not accurately capture the full spectrum of the harms and benefits of algorithmic advertising.

VI. Conclusion

Framing targeted advertising as a wicked problem lets us move away from a regulatory mindset that proposes reductive solutions and assumes rational linearity. In an environment marked by complexity and ambiguity, establishing a collective perspective is a crucial first step in moving away from solution-determinism; a collective perspective recognizes that a certain baseline of information is necessary before we can even begin to formulate potential interventions within the AdTech value chain. In the absence of meaningful transparency, self-regulatory measures designed by Big Tech product managers and coders will fill the governance vacuum that arises. Few actors have the necessary tools, technical knowledge, or access to data flows that enable the intervention in an ecosystem marked by pervasive information asymmetry.

Transparency as infrastructure regulation recognizes the important stakes at play in targeted advertising and takes the position that Big Tech should not be the only actor involved in designing governance interventions. It creates space for a huge network of stakeholders to operate; the collective perspective becomes a site of economic dialogue, political contestation, and scrutiny of social norms, enacted by and between users, platform actors, interest groups, researchers, lawyers, and regulators. This approach does not reductively assert that the state should always be the actor that regulates, or that self-regulation is an inherently flawed form of governance. Conversely, it does not assume that everything should be left to users' discretion (exercised via the adoption of technical approaches or enforcement of individual rights). It allows a range of stakeholder-driven interventions to be debated and tested; it also enables an understanding of how different interventions interact with each other in layered, entangled ways. There is no static "final" version of algorithmic advertising that will uniformly guarantee the benefits of personalization and reduce the risk of potential harms for all groups. Yet, a collective perspective may help us iteratively develop behavioral advertising into a more equal, less invasive, democratically audited system that we can all learn to live with.