Analysis of Google Ads Settings Over Time: Updated, Individualized, Accurate, and Filtered

Nathan Reitinger
University of Maryland

Bruce Wen
University of Chicago

Michelle L. Mazurek
University of Maryland

Blase Ur
University of Chicago

ABSTRACT

Advertising companies and data brokers often provide consumers access to a dashboard summarizing attributes they have collected or inferred about that user. These attributes can be used for targeted advertising. Several studies have examined the accuracy of these collected attributes or users' reactions to them. However, little is known about how these dashboards, and the associated attributes, change over time. Here, we report data from a week-long, longitudinal study (*n*=158) in which participants used a browser extension automatically capturing data from one dashboard, Google Ads Settings, after every fifth website the participant visited. The results show that Ads Settings is frequently updated, includes many attributes unique to only a single participant in our sample, and is approximately 90% accurate when assigning age and gender. We also find evidence that Ads Settings attributes may dynamically impact browsing behavior and may be filtered to remove sensitive interests.

CCS CONCEPTS

• **Security and privacy → Privacy protections**.

KEYWORDS

Web Tracking, Transparency, Measurement, Privacy, User Study

ACM Reference Format:
Nathan Reitinger, Bruce Wen, Michelle L. Mazurek, & Blase Ur. 2023. Analysis of Google Ads Settings Over Time: Updated, Individualized, Accurate, and Filtered. In *Proceedings of the 21st Workshop on Privacy in the Electronic Society (WPES '23), November 26, 2023, Copenhagen, Denmark*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3603216.3624968

1 INTRODUCTION

In 2023, digital ad spending is expected to reach 679 billion dollars [38, 39]. Most will go toward online behavioral advertising (*OBA*), a way of targeting ads to a user's known or inferred *attributes—psychographics*, like an interest in dogs or cats, and *demographics*, like age or gender. This is the reason some users receive ads related to "Baby & Pet Names" while others receive ads related to "High Performance & Aftermarket Auto Parts." OBA has shown higher click-through rates than non-targeted ads [24].

Calls for OBA transparency have led some ad networks to provide users with dashboards detailing the ads that have been served or the attributes assigned to a user [34, 41]. Although researchers have long studied the privacy implications of OBA [8, 23, 25, 27, 44, 45] and, more recently, ad dashboards [1, 22], few have looked at ad dashboards over time. How do profiles (a user's set of inferred attributes) evolve? Are profiles reactive to web browsing? How do profiles differ across users? How accurate are profiles?

Our research explores these questions using data from an IRB-approved, longitudinal field study occurring in late 2022. In that study, we asked participants to install and spend a week using a custom browser extension we built to visualize how users are tracked on the web. While the main results of our user study are reported separately, this short paper takes advantage of additional data collected during that study (and not reported elsewhere) to examine ad profiles longitudinally. Specifically, among many other features, our browser extension automatically fetched the user's profile from the Google Ads Settings dashboard (following the completion of our study, the service was renamed to My Ads Center [14–16]) on every fifth web page visit. This data was shared with us after the successful completion of the study. Using these regular snapshots of the user's profile in the Google Ads Settings dashboard alongside the five pages they visited in each snapshot, or *window*, we answer the questions listed above.

We find that Ads Settings data is frequently updated, relatively unique to a specific user, and accurate (≥90% on inferring age and gender). Furthermore, we find that updates in a user's profile seem correlated to the specific web browsing that directly preceded those updates. We also find evidence that profiles may be filtered to remove potentially sensitive interests [5, 51].

2 RELATED WORK

Prior work has focused on user interaction with ad preference managers, like Google's My Activity and Ads Settings and Facebook's Ad Preferences [1, 9, 11]. To our knowledge, none of this work has used longitudinal data. Using Ads Settings data from 2016, Tschantz et al. studied the accuracy of demographic inferences when compared to self-reports, finding that predicted age and gender were never more than around 75% accurate [42]. In a user study measuring several ad preference managers, Bashir et al. found that many interests in users' profiles were inaccurate when compared to their self-reported interests [1]. Venkatadri et al. reported a similar finding [47]. Bashir et al. also found that the interests in user ad profiles were not well explained by browsing history, suggesting

WPES '23, November 26, 2023, Copenhagen, Denmark

ACM ISBN 979-8-4007-0235-8/23/11...$15.00
https://doi.org/10.1145/3603216.3624968

possible external sources for interests [1]. While we collected frequent snapshots of Ads Settings, these prior studies used a single snapshot.

3 METHODS

We collected data using the Tracking Transparency v2 (*TT2*) browser extension, which we built on top of Weinshel et al.'s Tracking Transparency v1 extension (*TT1*) [50]. Weinshel et al. conducted a longitudinal field study with TT1 to assess participant perceptions of real-world tracking practices [50]. Participants downloaded the extension, used it for one week, and then answered survey questions. Based on web page visits, the extension logged metadata (e.g., trackers present) and inferred ad interest categories (e.g., visiting `petsmart.com` would log an interest in "animals"). Our updated extension, TT2, included many new features and new data sources, as described below. We used it in a similar, IRB-approved, longitudinal study. We recruited from Prolific, requiring participants to be U.S.-based, aged 18+, have a 95%+ platform rating, regularly use the Chrome web browser, and view TT2's dashboard page on at least three separate days out of seven. Participants were paid $10.00 for successful completion of the study. All information was visualized in a participant-accessible dashboard, while partially redacted information was sent to researchers.

Data scraped from the user's Ads Settings page was a key new data source for TT2. This information would only be available to the participant, and subsequently shared with us, if the participant were signed into Google with "personalized ads" turned on while using TT2. If so, Ads Settings data was automatically imported into the extension and re-fetched on every fifth web page visit by a participant, providing us a view of updates over time.

Another update in TT2, relevant to this paper, encompassed significant improvements to our method for inferring potential ad interest topics based on which web pages a user visited. The original extension made inferences on web-page text using TF-IDF keyword matching on Wikipedia-classified text. TT2 uses a machine learning shadow model [36] (stored locally, to protect user privacy) trained on ground truth from the Google Cloud Natural Language Content Classification API, a service for categorizing natural language text into ad interest categories [17]. The revised model had an overall accuracy of 71.4% (74.2% for second-level and 80.1% for top-level categories), significantly improving on Weinshel et al.'s technique. Although far from perfect, we deemed this accuracy sufficient to model potential correlations between the specific web pages a user had recently visited and updates to the user's Ads Settings profile.

Limitations. Our approach has several limitations. First, although the main study was of average size for a user study, it is relatively small for a measurement study. This may affect our findings related to comparisons among users' Ads Settings data, but we note that similar studies have been close in size [1]. The user study may have also demographically limited our data. Crowdworkers on services like Prolific are younger, more educated, and more tech-savvy than the general United States population [7, 21, 28, 32, 40, 43].

As is common in measurement studies [33], the telemetry data we analyze relies on sources outside of our control. We measure updates to Ads Settings following participant web page visits, but

Normalized Ads Settings Attribute Count (Per Participant)

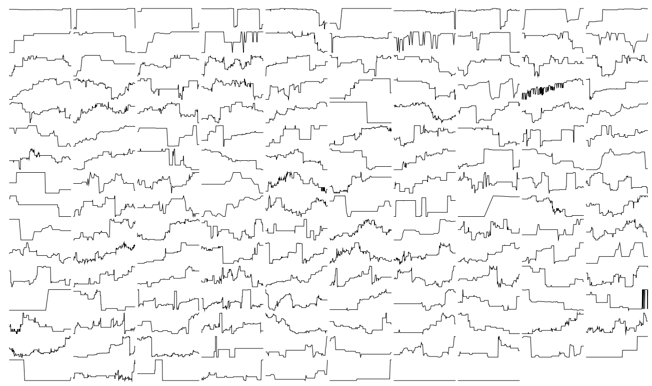

Figure 1: Each line represents one participant's normalized (subtract minimum and divide by the range) Ads Settings attribute counts over the course of the study. More than half of the participants had at least ten identified local maxima [35] (suggesting per-session targeting).

it is possible that a user was engaged in unobserved activity causing updates (e.g., separate device browsing or non-web activities). Similarly, Ads Settings data was fetched on every fifth web page visit, but we do not know specifically which web pages may have triggered a profile update, or why.

Lastly, several of our findings involve comparing the interest categories identified by TT2's inference engine to the interest categories in Ads Settings data. Although the interest engine showed marked improvement over TT1 [50], it remains far from perfect.

4 RESULTS

We refer to all inferences found in Ads Settings as "attributes." Attributes include *demographics* (e.g., "Homeowner"), *companies* (e.g., "Dietz & Watson"), *locations* (e.g., "Greater Charleston"), and *videos* (e.g., "3 videos from Secret Deodorant"). They also include hierarchical *interests* (e.g., "Arts & Entertainment → Music & Audio → Classical Music"); we term the leftmost segment to be depth 1 (Google uses up to six depth levels, while our inference engine uses three).

Participants. A total of 223 participants completed the study. Of these, 23 visited fewer than 100 web pages throughout the duration of the study, and 42 did not have Ads Settings data. The remainder, 158, are our data set. They were female (61%), male (36%), or non-binary (3%); ages 18-24 (18%), 25-34 (31%), 35-44 (25%), 45-54 (14%), and 55+ (12%); held a bachelor's degree or some college (59%), trade school or less (28%), or master's degree or more (13%); participants did (23%) or did not (73%) have technical backgrounds (percentages do not sum to 100 when participants chose "prefer not to say.").

Ads Settings data updates frequently. We fetched the Ads Settings page 46,457 times across our 158 participants, allowing us to observe that, for all types of attributes, the number of attributes assigned to a user move up *and down* frequently, suggesting the existence of a small anonymity crowd which may impact the threat of reidentification [2, 3, 6, 13, 29]. Figure 1 shows attribute count changes per participant. Only two participants had static attribute counts over the duration of the study; 25, 50, and 75% of the data

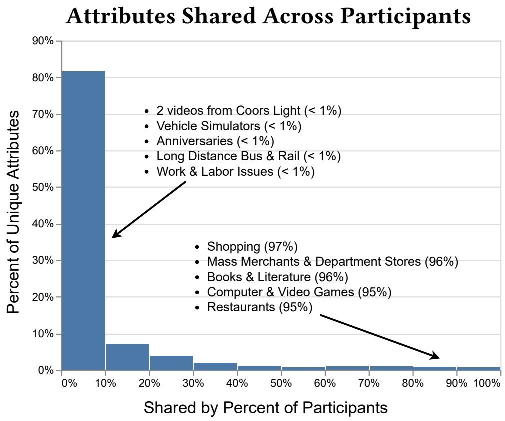

Figure 2: Histogram showing what percent of unique Ads Settings attributes are shared by what percent of participants.

is captured with 15, 28, and 44 updates to attribute counts, respectively. Moreover, of the 4,620 fetches where we noticed a change in the attribute count, 68% of the time the new profile included a previously unseen attribute. Of the 153 participants who started with a non-zero attribute count, we saw a 186% increase in total attributes associated with them at any time during the study. This average does not consider removed attributes and is dominated by a few participants who tripled their starting attribute counts (23%, 34%, and 55% increase for 25%, 50%, and 75% of the data, respectively).

Interest attributes changed the most. Out of 12,489 total new attributes, 87% were interests, while 7%, 4%, 2%, and 0.3% respectively were a company, demographic, location, or video. Most updated interests had a depth of three (53%), followed by two (21%), and four (20%). Demographically, the most common updates were: employer size (16% of demographic updates), education level (15%), marital status (14%), and income (14%). We explored how *similar* newly added interests were to existing interests based on categories and subcategories. Among 10,666 added interest attributes (from 151 participants with at least one interest added), 69% added a new interest within an existing category or subcategory (for example, adding "science → biological science → genetics" to an existing interest in "science → biological science → neuroscience").

Ads Settings data is individualized. Attributes were highly individualized (Figure 2), suggesting heightened privacy invasion. A total of 3,250 unique attributes were associated with participants, and most attributes were associated with fewer than 10% of participants, representing a long-tail distribution [19]. Companies, videos, and locations were the most individualized. Among companies, 70% (656 of 939) were unique to a single participant (85% of participants had company attributes). Likewise, 94% of video attributes (i.e., YouTube channels) were unique to one participant (although only 24% of participants had one or more video attributes), and 45% of locations were unique to one participant (91% of participants were associated with a location).

We also looked at how similar each participant's Ads Settings profile was compared to other participants. We made pairwise comparisons between participants on the total set of Ads Settings attributes captured by any fetch using Jaccard similarity $\langle (A \cap B)/(A \cup B) \rangle$ [20]. As shown in Figure 3, most participants had some commonalities (20-39% similarity), but few were more similar than that. This trend again illustrates a long-tail effect: most participants had general topics in common, like shopping or news, but many attributes were unique to each participant (Figure 2).

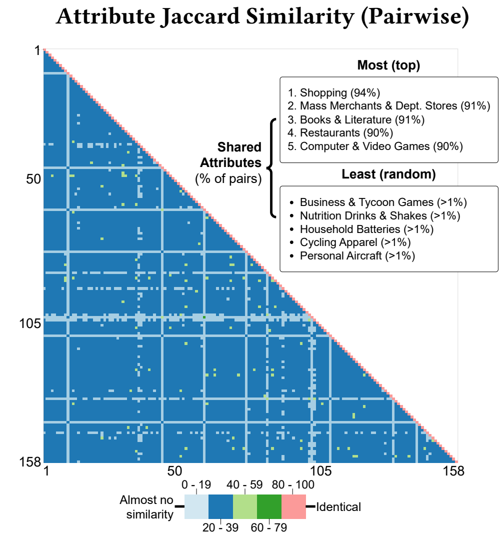

Figure 3: Jaccard similarity between pairs of participants (all observed attributes, ranged from 79 to 381). Commonly shared attributes included interests like shopping versus unshared interests like personal aircraft. Intersections between pairs predominantly consisted of interests (94%). Common differences included interests (77%) and companies (12%).

Table 1: Ads Settings accurately predicted gender (95%) and age (89%). [O]/[Y] in age classification respectively mean that all incorrect predictions were older or younger than self-reported.

Gender	Count	Correct	Incorrect	Accuracy
Female	92	90	2	98%
Male	56	55	1	98%
Non-binary	4	0	4	0%
Total	152	145	7	95%

Age	Count	Correct	Incorrect	Accuracy
18-24	29	21	[O] 8	72%
25-34	45	41	[O] 4	91%
35-44	36	33	[Y] 3	92%
45-54	22	21	[Y] 1	95%
55-64	13	13	0	100%
Total	145	129	16	89%

Google accurately predicts gender and age. Ads Settings was highly accurate when predicting either the gender or age of participants, according to self-reported demographic information (Table 1). For gender (97% of participants had a gender ascribed by Google, while 92% had an age), Ads Settings was correct ~90% of the time, after dropping two participants who were labeled as both

Table 2: (A) Mean percent overlap and Jaccard similarity between the interests assigned to a user in Ads Settings and by our inference engine. Interests are hierarchical; the rows indicate the depth of the match. (B) The fraction of time windows when new Ads Settings interests appeared during which our inference engine identified a matching interest at the specified depth within one (5 minutes) or two (10 minutes) time windows.

(A)

Total Interests		
	Overlap	**Jaccard**
1st	88%	0.81
2nd	63%	0.31
3rd	41%	0.13
Exact	33%	0.12

(B)

New Interests		
		1+ Match
One window	**1st**	41%
	2nd	21%
	3rd	8%
Two windows	**1st**	46%

"Female" and "Male" on different fetches of Ads Settings data. Notably, none of the non-binary participants were classified correctly, a common issue in gender targeting [49]. For age, Ads Settings correctly guessed 89% of participants' age ranges (industry-standard age buckets [18]). Incorrect guesses were often older than self-reported ages. This accuracy on age and gender is much higher than previously reported [42], although scale (our study had fewer participants), individual survey factors, and the way Google infers these demographics may have changed in the intervening years.

Web browsing seems to impact Ads Settings. While we cannot measure causally whether a user's browsing history causes updates to their Ads Settings attributes, we found some evidence they may be related. First, we compare the total set of interests inferred by TT2's internal inference engine based on the web pages a participant visited to their total set of ever-observed Ads Settings interests (Table 2A). Similar to Bashir et al. [1], on average 33% of these attributes match exactly—though we use seven days of browsing versus 100. However, we find that first-level (depth 1) categories match at an average of 88%, second-level at 63%, and third-level at 41%. Comparable Jaccard similarity scores are 81%, 31%, and 13% for first-, second-, and third-level matches.

Second, we assess interest matching on a subset of web pages visited around the time when we noticed an Ads Settings update. If an Ads Settings fetch showed attribute differences from the prior fetch, TT2 logged the web pages visited in the past five minutes and future thirty seconds (what we term a recency *window*). We calculate the number of windows where one or more visited web page inferences in the window (labeled by our inference engine) match a new Ads Settings attribute category. Notably, Ads Settings may have updated anytime between the two fetches, so our notion of recency is approximate at best. Nevertheless, we see interesting patterns in our recency data (Table 2B). When considering the last ten minutes (two windows), at least one newly added Ads Settings interest matches an interest we inferred from a web page visit at the top level (depth 1) 46% of the time. This drops to 41% when considering only the current window (top level) and to 21% when considering second-level matches.

Ads Settings may filter sensitive interests. Prior research has shown that Ads Settings filters out sensitive interests [5, 51]. We confirm this finding with participant data. We identified a total of 1,199 unique interests, across all participants, from Ads Settings data, compared to 520 from our inference engine operating on web pages. Similarly, we observed 216 Ads Settings interests per person on average, compared to 110 for web page interests. Following Dolin et al. [8], we label certain interest categories as "sensitive." We find on average 3% of a participant's Ads Settings interests are sensitive. In contrast, 12% of the interests we inferred on web pages were sensitive. When looking, for each participant, at the number of web page interests that are sensitive versus the number of Ads Settings interests that are sensitive, we find a statistically significant difference (Wilcoxon Signed-Rank, $p < 0.001$). Out of the 490 different interest categories we identified from web page visits, Google was a tracker on 99% of the domains associated with these categories, as well as 100% of the 58 sensitive categories (Appendix A). Among many possibilities, Google may be declining to categorize people using sensitive interests or it could be redacting these categories from Ads Settings.

Table 3: (A) The total number of unique interests (across all participants) and the fraction of sensitive interests [8] for Google Ads Settings data versus our inference engine. (B) The mean number of interests and the fraction of sensitive interests per participant.

	(A) *Total Across Participants*		(B) *Average Per Participant*	
	Unique Interests	**Percent Sensitive**	**Unique Interests**	**Percent Sensitive**
Ads Settings	1,199	3%	216	3%
Our Inferences	520	12%	110	9%

5 DISCUSSION

Google is the Internet's largest ad network, taking part in advertising on most web pages [4, 10, 12, 46]. We find that what Google knows about its users is accurate, individualized, updated, and likely filtered. Although Ads Settings is a step toward transparency (enabling our study), more could be done.

Users who do not revisit their Ads Setting page repeatedly—and very few users are likely to visit that page repeatedly—may not realize that any specific visit to Ads Settings shows only a snapshot of an ongoing process that changes frequently. Ads Settings does not currently indicate these changes. While it may be intuitive that human interests are diverse and change over time [31, 37], it may not be intuitive that the set of interests Google has assigned to a user is also constantly changing. Similarly, Ads Settings offers no explanation of when and why attributes were assigned or updated, and it offers no information about how unique a user's attributes are. Although we studied Google Ads Settings, future work should evaluate whether other transparency dashboards display similar behaviors. Further transparency is needed to give users a clearer picture of how they are tracked [10, 30] and classified [26, 48], especially how these aspects change over time.

ACKNOWLEDGEMENTS

We gratefully acknowledge support from a UMIACS contract under the partnership between the University of Maryland and the Department of Defense. This material is also based upon work supported by the National Science Foundation under Grants No. CNS-2047827, CNS-2149680, and CNS-2151290.

REFERENCES

[1] Muhammad Ahmad Bashir, Umar Farooq, Maryam Shahid, Muhammad Fareed Zaffar, and Christo Wilson. Quantity vs. Quality: Evaluating user interest profiles using Ad Preference Managers. In *Proc. NDSS*, 2019.

[2] Steven M. Bellovin, Preetam K. Dutta, and Nathan Reitinger. Privacy and synthetic datasets. *Stanford Technology Law Review*, 22:1–52, 2019.

[3] Alex Berke and Dan Calacci. Privacy limitations of interest-based advertising on the web: A post-mortem empirical analysis of google's floc. In *Proc. CCS*, 2022.

[4] Aaron Cahn, Scott Alfeld, Paul Barford, and Shanmugavelayutham Muthukrishnan. An empirical study of web cookies. In *Proc. WWW*, 2016.

[5] Amit Datta, Michael Carl Tschantz, and Anupam Datta. Automated experiments on ad privacy settings. In *Proc. PETS*, 2015.

[6] Claudia Diaz, Stefaan Seys, Joris Claessens, and Bart Preneel. Towards measuring anonymity. In *Proc. IWPE*, 2002.

[7] Djellel Difallah, Elena Filatova, and Panos Ipeirotis. Demographics and dynamics of Mechanical Turk workers. In *Proc. WSDM*, 2018.

[8] Claire Dolin, Ben Weinshel, Shawn Shan, Chang Min Hahn, Euirim Choi, Michelle L. Mazurek, and Blase Ur. Unpacking perceptions of data-driven inferences underlying online targeting and personalization. In *Proc. CHI*, 2018.

[9] Julia Earp and Jessica Staddon. "I had no idea this was a thing": On the importance of understanding the user experience of personalized transparency tools. In *Proc. STAST*, 2016.

[10] Steven Englehardt and Arvind Narayanan. Online tracking: A 1-million-site measurement and analysis. In *Proc. CCS*, 2016.

[11] Florian M. Farke, David G. Balash, Maximilian Golla, Markus Dürmuth, and Adam J. Aviv. Are privacy dashboards good for end users? Evaluating user perceptions and reactions to google's my activity. In *Proc. USENIX Security*, 2021.

[12] Phillipa Gill, Vijay Erramilli, Augustin Chaintreau, Balachander Krishnamurthy, Konstantina Papagiannaki, and Pablo Rodriguez. Follow the money: Understanding economics of online aggregation and advertising. In *Proc. IMC*, 2013.

[13] José González-Cabañas, Ángel Cuevas, Rubén Cuevas, Juan López-Fernández, and David García. Unique on Facebook: Formulation and evidence of (nano) targeting individual users with non-PII data. In *Proc. IMC*, 2021.

[14] Google. Greater transparency and control over your Google ad experience. https://blog.google/technology/ads/greater-transparency-and-control-over-your-google-ad-experience/, 2018.

[15] Google. Google ad settings. https://adssettings.google.com, 2019.

[16] Google. Your ads, your choice. https://myadcenter.google.com, 2023.

[17] Google Cloud. Natural language ai. https://cloud.google.com/natural-language, 2022.

[18] Google Ads Help. About demographic targeting. https://support.google.com/google-ads/answer/2580383?hl=en.

[19] Michael A. Hitt. *The long tail: Why the future of business is selling less of more.* Hyperion, 2007.

[20] Paul Jaccard. The distribution of the flora in the alpine zone. *New phytologist*, 11:37–50, 1912.

[21] Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara Kiesler. Privacy attitudes of Mechanical Turk workers and the U.S. public. In *Proc. SOUPS*, 2014.

[22] Bin Liu, Anmol Sheth, Udi Weinsberg, Jaideep Chandrashekar, and Ramesh Govindan. Adreveal: Improving transparency into online targeted advertising. In *Proc. HotNets*, 2013.

[23] Aleecia McDonald and Lorrie Faith Cranor. Beliefs and behaviors: Internet users' understanding of behavioral advertising. In *Proc. TPRC*, 2010.

[24] Eduardo Abraham Schnadower Mustri, Idris Adjerid, and Alessandro Acquisti. Behavioral advertising and consumer welfare: An empirical investigation. In *Proc. WEIS*, 2021.

[25] Saurabh Panjwani, Nisheeth Shrivastava, Saurabh Shukla, and Sharad Jaiswal. Understanding the privacy-personalization dilemma for web search: A user perspective. In *Proc. CHI*, 2013.

[26] Jon Porter. GDPR makes it easier to get your data, but that doesn't mean you'll understand it. https://www.theverge.com/2019/1/27/18195630/gdpr-right-of-access-data-download-facebook-google-amazon-apple, 2019.

[27] Kristen Purcell, Lee Rainie, and Joanna Brenner. Search engine use 2012. https://www.pewresearch.org/internet/2012/03/09/search-engine-use-2012/, 2012.

[28] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. How well do my results generalize? Comparing security and privacy survey results from Mturk, web, and telephone samples. In *Proc. IEEE S&P*, 2019.

[29] Nathan Reitinger and Amol Deshpande. Epsilon-differential privacy, and a two-step test for quantifying reidentification risk. *Jurimetrics: The Journal of Law, Science & Technology*, 63(3):263–317, 2023.

[30] Nathan Reitinger and Michelle L. Mazurek. ML-CB: Machine learning canvas block. In *Proc. PETS*, 2021.

[31] Ann K. Renninger and Rose K. Pozos-Brewer. Psychology of interest. *International Encyclopedia of the Social & Behavioral Sciences*, 12:378–385, 2015.

[32] Joel Ross, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. Who are the Turkers? Worker demographics in Amazon Mechanical Turk. In *Proc. CHI EA*, 2010.

[33] Leonardo Sanna, Salvatore Romano, Giulia Corona, and Claudio Agosti. YTTREX: Crowdsourced analysis of YouTube's recommender system during COVID-19 pandemic. In *Proc. SIMBig*, 2020.

[34] Tatevik Sargsyan. The privacy role of information intermediaries through self-regulation. *Internet Policy Review*, 5:1–17, 2016.

[35] Scipy. Find Peaks. https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html, 2023.

[36] Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. Membership inference attacks against machine learning models. In *Proc. IEEE S&P*, 2017.

[37] Paul J. Silvia. *Exploring the psychology of interest.* Psychology of Human Motivation, 2006.

[38] Dina Srinivasan. Why Google dominates advertising markets. *Stanford Technology Law Review*, 24:55–175, 2020.

[39] Statista. Digital advertising - worldwide. https://www.statista.com/outlook/dmo/digital-advertising/worldwide, 2023.

[40] Jenny Tang, Eleanor Birrell, and Ada Lerner. Replication: How well do my results generalize now? The external validity of online privacy and security surveys. In *Proc. SOUPS*, 2022.

[41] Omer Tene and Jules Polenetsky. To track or "do not track": Advancing transparency and individual control in online behavioral advertising. *Minn. JL Sci. & Tech.*, 13:281–357, 2012.

[42] Michael Carl Tschantz, Serge Egelman, Jaeyoung Choi, Nicholas Weaver, and Gerald Friedland. The accuracy of the demographic inferences shown on Google's Ad Settings. In *Proc. WPES*, 2018.

[43] Anne M. Turner, Thomas Engelsma, Jean O. Taylor, Rashmi K. Sharma, and George Demiris. Recruiting older adult participants through crowdsourcing platforms: Mechanical Turk versus Prolific Academic. In *Proc. AMIA*, 2020.

[44] Joseph Turow, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, and Michael Hennessy. Americans reject tailored advertising and three activities that enable it. https://repository.upenn.edu/cgi/viewcontent.cgi?article=1551&context=asc_papers, 2009.

[45] Blase Ur, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang. Smart, useful, scary, creepy: Perceptions of online behavioral advertising. In *Proc. SOUPS*, 2012.

[46] United States v. Google. 1:23-cv-00108 (District Court for the District of Columbia). https://www.justice.gov/opa/press-release/file/1563746/download, 2023.

[47] Giridhari Venkatadri, Piotr Sapiezynski, Elissa M. Redmiles, Alan Mislove, Oana Goga, Michelle L. Mazurek, and Krishna P. Gummadi. Auditing offline data brokers via facebook's advertising platform. In *Proc. WWW*, 2019.

[48] Sophie Veys, Daniel Serrano, Madison Stamos, Margot Herman, Nathan Reitinger, Michelle L. Mazurek, and Blase Ur. Pursuing usable and useful data downloads under GDPR/CCPA access rights via co-design. In *Proc. SOUPS*, 2021.

[49] Miranda Wei, Madison Stamos, Sophie Veys, Nathan Reitinger, Justin Goodman, Margot Herman, Dorata Filipczuk, Ben Weinshel, Michelle L. Mazurek, and Blase Ur. What Twitter knows: Characterizing ad targeting practices, user perceptions, and ad explanations through users' own twitter data. In *Proc. USENIX Security*, 2020.

[50] Ben Weinshel, Miranda Wei, Mainack Mondal, Euirim Choi, Shawn Shan, Claire Dolin, Michelle L. Mazurek, and Blase Ur. Oh, the places you've been! User reactions to longitudinal transparency about third-party web tracking and inferencing. In *Proc. CCS*, 2019.

[51] Craig E. Wills and Can Tatar. Understanding what they do with what they know. In *Proc. WPES*, 2012.

A ADDITIONAL FIGURES

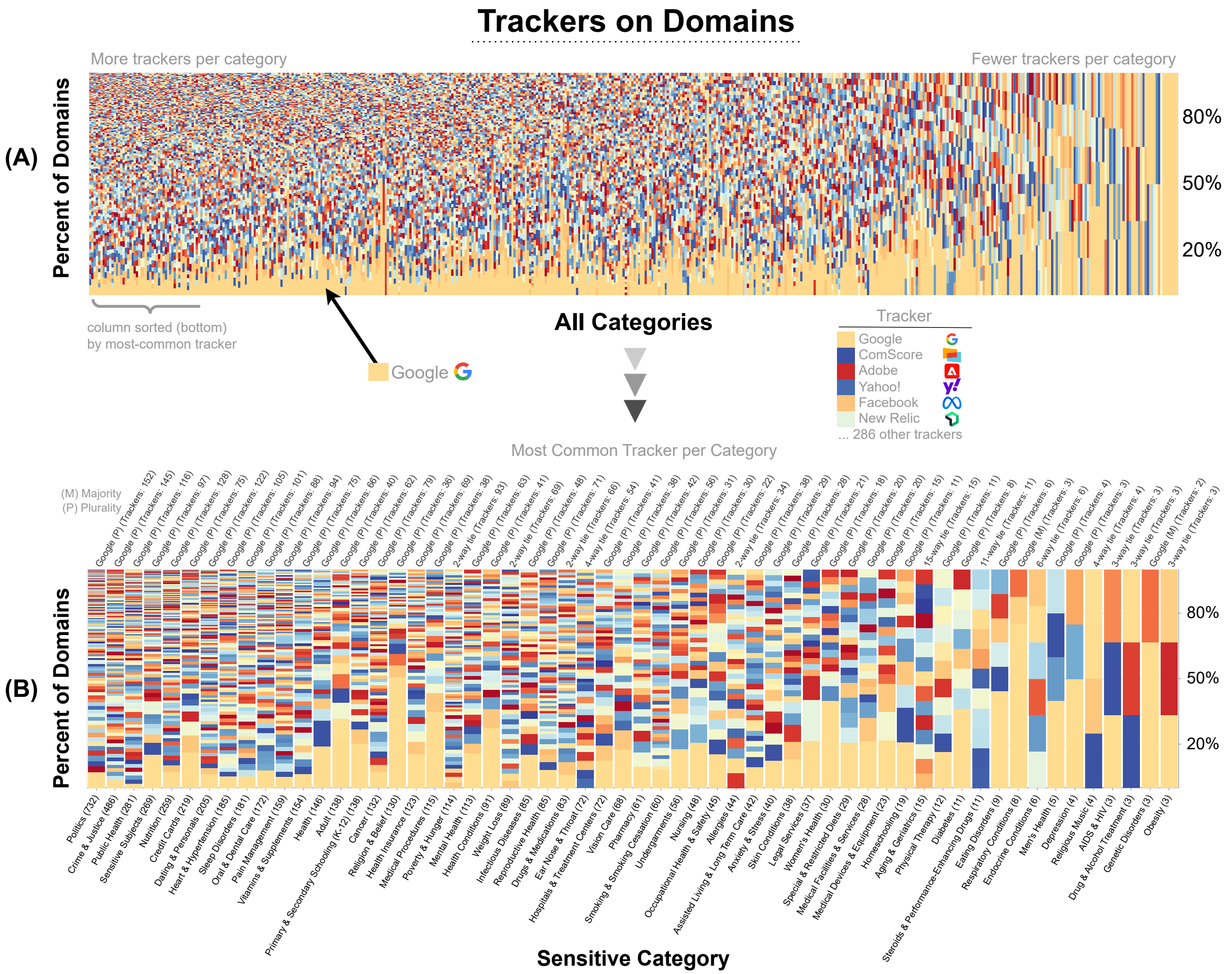

Figure 4: Showing the fraction of tracking companies present on domains associated with an interest attribute as identified by our inference engine. We show this first for all interest categories (A) and then for only sensitive interest categories (B), where sensitive is defined per Dolin et al. [8]. Each column represents a specific interest attribute. For each domain where one or more pages hosted on that domain were associated with a given category, we record the different tracking companies present on the domain. The y-axis represents the fraction of trackers on all domains in an interest category (e.g., if Google were the only tracker on the only domain associated with an interest category, Google would be listed as being on 100% of the domains). As a result, an even vertical split of colors in an interest column represents an even distribution of trackers on all domains in this category. However, one color taking up an outsized fraction of the vertical bar means that one particular company is present on a larger fraction of domains associated with that interest than other companies. Notably, Google is frequently the most common tracker for any given interest category. In the bottom figure (B), the interest labels as columns show, on the bottom x-axis, the count of domains falling into that category (e.g., 7,320 domains related to politics). Figure (B) also shows, on top, the most common tracker across all domains (either as a majority, M, a plurality, P, or a tie) and the total number of trackers found on all domains per category (e.g., politics was a category associated with 152 unique trackers).