Google Ads Test: Testing Interest Based Audience Solutions with Privacy Preserving Signals

1. Abstract:

As Chrome's third-party cookie deprecation (3PCD) deadline approaches, Google Ads have been experimenting with Interest-based Audience Solutions (IBA) by replacing third-party cookies (3PC) with a variety of privacy-preserving signals including Topics API. The experiment outlined in this document is one of multiple experiments that will test other privacy-preserving solutions and Chrome's Privacy Sandbox in individual or combined settings.

As part of the Privacy Sandbox Commitments, we are working under the supervision of the UK Competition and Markets Authority (CMA) and Information Commissioner's Office (ICO) to design and carry out experiments to evaluate the effectiveness of the Privacy Sandbox APIs and other privacy-preserving signals supporting key online advertising use case without 3PC.

This document reports the findings of our recent experiment evaluating impact to Google's ad services, particularly to IBA ads within Google's display network. It does not seek to quantify or contrast the performance that would be achieved by other ecosystem players in the absence of 3PCs. The experiment retained 3PCs for all other components including frequency capping, measurements and remarketing. We removed 3PC only from Interest-based Audience Solutions - like Affinity, In-Market, Custom Audience and Demographic segments. Also, the experiment applied the treatment to display ads traffic on Chrome and hence, the observations are for the state of display ads, not overall Google Ads. The experiment was completed in close collaboration with the CMA, in compliance with our requirement to publish the results of tests that are material to evaluating the effectiveness of the Privacy Sandbox APIs (paragraph 17.c.v. of the Commitments). In the rest of this post, we provide a summary of

(1) privacy-preserving signals used in the experiment instead of 3PCs for certain IBA solutions offered by Google Ads;
(2) experiment design which aims to isolate 3PCD impact to IBA and evaluate the utility of privacy-preserving signals;
(3) Directional observations from the ~5 week long A/B experiment conducted for IBA with Chrome's Topics API Origin Trials traffic in Q1 2023

For further explanation on what IBA and Topics API are, refer to Appendix 1.

2. Google's ads platforms: Privacy-Preserving Signals for IBA

Google's IBA solutions have traditionally relied on a variety of user signals to derive interest-based audiences. Several of these signals are privacy durable. Post 3PCD, we will leverage these privacy-preserving signals to serve IBA ads. Below, we outline key signals we leveraged for IBA in this experiment.

1) Contextual Data is the data available as part of the page view (e.g. article content on NBA finals on ESPN.com). Using contextual data, an Ad Tech Provider (ATP) can either use the interests obvious from the content on the page and even infer other interests. Contextual data is also the most commonly available source of signal compared to Publisher First-Party data or Topics which may not be always available.
2) Publisher First-Party (1P) Data: Publisher 1P ID is an identifier that is unique for a user for each publisher, signaling a publisher/user pair. Using Publisher 1P ID, an ATP can generate user profiles based on the traversal history of that user *within* that publisher.
3) Topics: Topics API from Chrome's Privacy Sandbox returns up to three topics per week per user. We use Topics as users' coarse interest and as an additional signal to infer other interests in conjunction with other privacy-preserving signals.

In addition to new signals, we observed AI-powered optimizations further improve ad serving by learning from privacy preserving signals along with ad engagement and conversions data to improve the efficacy of ad serving.

As this is a preliminary experiment, these signals are subject to change and the outlined use of privacy-preserving solutions may evolve over time (more on our use of signals in Appendix 2).

3. Experiment Design

The goal of this experiment is to understand the directional impact to IBA on Google's display network when we replace 3PC with privacy-preserving signals. To evaluate the impact, we used A/B experimentation methodology, meaning we have two experiment groups:

1. *Control*: In Control, we maintained the "status quo" on Chrome with all 3PC functionality retained.

2. *Treatment*: In Treatment, we isolate the impact of 3PCD on IBA by dropping 3PCs for building interest-based audiences (affinity or in-market) and instead applying ready-to-test privacy-preserving solutions including Topics. We retain 3PCs for all other ad-tech components such as remarketing and measurements.

For clarity, we note:

1. Experiment ran for only Google's *display network* not overall Google Ads.

2. Experiment traffic included all parties Google Platforms interact with, including Google Authorized Buyers, Buyers and Sellers from 3rd Party Exchanges integrated with Topics.
3. In the experiment, we applied treatment only to Google buyside platforms. We did this by suppressing 3PCs after Google buyside platforms received ad requests. This means that in auction, other Buying Platforms leveraged 3PCs whereas Google buying platforms did not.
4. We re-trained the most impactful models that affect campaign retrieval and bidding strategy by replacing 3PCs with privacy preserving signals.
5. We diverted the treatment group from the control group in this experiment using 3PC as the unit of assignment.
6. We are planning more comprehensive experiments for the future. We will remove 3PCs incrementally from our systems and apply new ready-to-test 3PCD mitigations including the rest of Privacy Sandbox APIs. We will also consider incorporating a third experiment arm that removes 3PCs but does not apply the Privacy Sandbox APIs to evaluate the baseline performance without the Privacy Sandbox.
7.

The table below outlines the overall configurations applied for the experiment setup:

	Control (Status Quo)	Treatment (No3PC+Topics)
Traffic	Chrome Traffic (Desktop + Mobile Web) 2% of traffic for control and 2% for treatment (80% of Chrome OT Traffic)	
Buyside Platforms	Google Display Ads + Display and Video 360	
Sellside Platforms	Google AdSense + Google Ad Manager	
Remarketing	3PC	
Frequency Capping	3PC	
Measurements	3PC	
Real Time Buyers	3PC (independently participating in Topics OT)	
Interest-Based Profiles	**Status Quo (3PC+Contextual+Pub1P)**	**Status Quo without 3PC + Chrome's OT Topics + Simulated Pub1P**
Geo & Demographics Profile	**Status Quo (3PC+Contextual+Pub1P)**	**Status Quo without 3PC**

4. Analysis:

Business Metrics Reported[1]:

The following are business metrics we use to evaluate experiments, and we recommend that the industry use similar metrics for evaluating experiments of similar nature:

A. % Change in Spend by Advertisers

  Advertiser spend on campaigns. It is the sum of Google's share and Publishers' share of revenue.

B. MH[2] Conversions Per Dollar (CPD)

  CPD is the average number of conversions per dollar spent. It is a measure of ad quality received by advertisers.

C. MH Conversion Rate (CVR)

  CVR is the average number of conversions per ad interaction, shown as a percentage. It is a measure of traffic quality received by advertisers.

D. MH Click Through Rate (CTR)

  CTR is the average number of clicks per impression. It is a measure of ad relevance.

Data Segments Analyzed and Reported:

In addition to the overall IBA business on the display network, we further evaluate segments to derive insights on how performance might vary under different constraints and settings:

A. **Buying Platforms (Google Ads and Display & Video 360):**
   Google Ads (GA) and Display & Video 360 (DV3) are Google's ad buying platforms. We evaluate platform segments because each platform has its own client profile with unique business goals and practices.
B. **With or Without AI-Powered Optimization[3]:**
   Segments with AI-powered Optimization include campaigns that opted in optimized targeting[4] or bidding optimization[5]. We evaluate these segments as the degree of AI-powered optimization can have implications on campaign performance.

---

[1] See Appendix 3 for the formulas for calculating each metric and confidence interval
[2] Treatment effects on CPD, CVR, and CTR are estimated by the Mantel-Haenszel (MH) ratio. The MH estimates one rate relative to another, specifically a ratio of rates (or ratio of probabilities, which is called "relative risk" or "risk ratio")
[3] The segment without AI-powered optimization employs manual targeting (i.e. target only client curated criteria such as "cycling") and manual bidding (i.e. set a specific bid price). With AI-powered optimization segment makes use of either bidding optimization or optimized targeting. We sliced the segment based on ad-group configuration.
[4] Auto targeting looks at information like keywords on your landing page or in your creative assets, and finds audiences that can meet your campaign's goals. Auto-targeting seeks additional conversions by targeting people most likely to convert.
[5] Bidding optimization systems produce bids on behalf of clients to deliver most conversions to them given their business constraints (e.g. preferred ROI or budget goal).

Caveats – how and how not to interpret this data:

(1) The experiment results reflect the impact to certain IBA features on Google's display network. Individual ATPs may observe different results even in the same experiment environment.

(2) We recommend the results to be interpreted as directional indicators rather than precise estimates of 3PCD impact on IBA on Google's display network. As a reminder, note that all other products (e.g. measurement, remarketing) retained 3PCs and we applied treatment only to Google buyside platforms in this experiment.

(3) Google's ads platforms leverage machine learning models to optimize the efficacy of ad serving. These models likely are better calibrated with greater volume of traffic. Due to the limited traffic in Chrome's Origin Trials, we retrained only the most impactful models and their performance did not fully converge to an optimized state for the experiment traffic. Future observations may change due to model convergence.

(4) Interpreting the difference between the control and treatment groups as the "treatment effect" makes A/B testing an appealing testing methodology. Yet, the reported findings in this study are imperfect estimates due to the limitations of A/B testing. Some of the limitations applicable to this experiment are the following:
    (a) In the outlined experiment setup, other demand sources such as third party buyers and other ads use cases (remarketing) continue to use 3PCs. This asymmetry of information will result in an auction outcome not reflective of the 3PCD state. A possible path to overcome this limitation would be to suppress 3PCs at the source (i.e. Chrome) and to coordinate with other ATPs to apply their 3PCD mitigations for the duration of the experiment.
    (b) Some events (e.g. purchase actions) are rare, requiring a large sample to reliably observe the difference in control and treatment group. We need an even larger sample to detect the difference for slices within those groups. We can observe the effects on our experiment through the noise observed in ratio metrics.
    (c) In this type of experiment, multiple entities exist such as users, advertisers, publishers, and ATPs. Ideally, we should identify the 'unit' of randomization that can assign non-interacting groups of entities into control and treatment. Given the interactions among the said entities, it is difficult to identify a viable unit. For example, a user may have multiple Cookies (browsers), some in the experiment and some outside.

Table 1: Observations for Google Ads display campaigns using IBA

Primary Reported Metric (% change between control and treatment)[6]	**With AI-Powered Optimization**	**Without AI-Powered Optimization**	**Overall**
Spend by Advertisers	**-1.654% [-2.376%, -0.927%]**	**-7.248% [-8.792%, -5.679%]**	**-1.906% [-2.593%, -1.213%]**
Conversions per Dollar (MH[7] CPD)	**-0.454% [-0.893, -0.013]**	-2.446% [-5.728, 0.951]	**-0.478% [-0.928, -0.027]**
Conversion Rate (MH CVR)	**-1.899% [-2.294, -1.503]**	-1.132% [-4.530, 2.386]	**-1.891% [-2.344, -1.435]**
Click-through rate (MH CTR)	**-8.043% [-8.356, -7.730]**	**-8.256% [-9.473, -7.023]**	**-8.048% [-8.361, -7.734]**

Table 2: Observations for Display & Video 360 display campaigns using IBA

Primary Reported Metric (% change between control and treatment)	**With AI-Powered Optimization**	**Without AI-Powered Optimization**	**Overall**
Spend by Advertisers	0.24% [-0.71%, 1.23%]	-0.08% [-1.4%, 1.21%]	0.19% [-0.68%, 1.02%]
Conversions per Dollar (MH CPD)	**-1.91% [-3.1%, -0.71%]**	-1.63% [-4.44%, 1.25%]	**-1.9% [-2.99%, -0.79%]**
Conversion Rate (MH CVR)	**-6.07% [-7.2%, -4.92%]**	**-3.24% [-6%, -0.406%]**	**-5.64% [-6.69%, -4.58%]**
Click-through rate (MH CTR)	**-1.62% [-2.42%, -0.80%]**	-0.365% [-1.58%, 0.86%]	**-1.16% [-1.83%, -0.48%]**

[6] Note that the numbers in square brackets refer to 95% bootstrapped confidence intervals. DV3 used 1000 bootstrap reps and GDA used 200 bootstrap reps to compute confidence intervals. We bolded the metrics whose confidence interval did not cross zero.

[7] For ratio metric calculation, we applied filters for MH ratio metrics (CPD, CVR, CTR) due to the challenges we explain below. The filters aims to reduce the impact from noise:
1) Conversions are noisy - It could be the case that an advertiser spends $100 in both treatment and control, gets 1000 clicks in each, but has 1 conversion in treatment and 2 in control. This could be just a measurement noise. (e.g. consider a poisson random variable with the rate of 1. It is very likely to generate 0, 1 or 2 events.).
2) CPD changes are heterogeneous among advertisers. There would be those who get a much better CPD and those who get much worse CPD. In aggregating them we need to choose weights for them. The weight could be the same for all ads, could be inversely proportional to our measurement error, could be according to their business impact such as spend, or according to other methods.

5. Findings & Summary of Results:

Overall, the experiment results showed relatively small performance impact for Interest-Based Audience Solutions despite removing 3PCs from serving IBA ads. To avoid misinterpreting below findings, we remind the reader that this experiment was conducted in a highly controlled environment (Section 3) where 3PCs were retained for other areas like remarketing and measurement, and future observations will likely change as our systems will evolve and experiment setup becomes reflective of 3PCD end-state.

Impact on overall spend and quality metrics:

Reported metrics compare the performance of IBA delivered with 3PCs (Control) against privacy preserving signals (Treatment). In the treatment arm, we observed -1.9% [-2.6,-1.2] change in spend and -0.5% [-0.9,0] in CPD for IBA serving on GDA and 0.2% [-0.7,1.0] in spend and -1.9% [-3.0,-0.8] in CPD for IBA serving in DV3.

Comparing the results to 2019 study on impact of disabling 3PCs

Google published a study evaluating the impact of 3PCD in 2019. In the study, we observed that for the top 500 global publishers, average spend in the treatment group decreased by 52%. In the experiment outlined in this post, we report a much lower loss. We want to clarify that the two experiments are not comparable and should not be contrasted.

First, we should note that the referenced study was conducted over four years ago and there has been much development in the privacy landscape such as 3PCD for Firefox, Safari and other browsers; changes introduced for IDFA with iOS14, and more. Our systems have also evolved and adapted to privacy changes to deliver better value for our clients by relying on privacy durable signals and AI optimizations.

Second, the goals and subsequent setup for the two experiments were very different. In 2019 study, the goal was to evaluate 3PCD impact to all of Google display network, whereas the goal of the current experiment is to evaluate the impact to interest based audience solutions only, a singular product offering within our ads product portfolio. As a result, the 2019 study employed experiment setup that removed 3PCs from most components including IBA, remarketing, frequency capping whereas the current experiment only did so for IBA. Additionally, the degree of reliance on 3PCs for each product set is unique, so the impact and performance changes for IBA cannot be extrapolated to other products or overall.

Potential headwinds to performance:

At a high level, we suspect a few effects are likely driving the performance reduction

- Lower-quality user signals: Without 3PCs, we likely have coarser and less precise user signals and lower recall of them. As a first-order effect, coarser user interests ("Shoes"

as an interest, not “Jordan Enthusiasts”) can reduce the pool of eligible campaigns that could match to an impression. With fewer campaigns, auctions can see less bid density, resulting in lower spend overall. Meanwhile, less precise signals can cause overprediction of a user behavior, lowering quality metrics (CTR, CPD, CVR).
- Less accurate models: Without 3PCs, many of our models can end up training on features and traffic of lower quality and coverage, resulting in models with less accurate prediction of user behavior (clicks, conversions, and conversion value). Less accurate prediction can lower our efficacy in selecting campaigns for auction and bid values can decrease, hence less spend, to achieve the same advertiser ROI. Note that the prediction models used in the experiment were trained on privacy-preserving signals including Topics-based features.

Performance Drivers:

Despite the challenges above, our experiment results showed limited loss in advertiser spend and return on investment for IBA solutions. We believe this can be attributed to the following::

- Contextual Signals: Contextual signal is a rich and reliable signal as it is often the source for the freshest, the most relevant, and most available signal about the user. For the same reason, additional user interests we can infer from contextual signals can be of higher relevance. Also, Google’s AI can further boost the strength of contextual signals by combining them with other privacy-preserving signals when inferring user interests.
- AI-Powered Optimization and other privacy-preserving signals: Our system employs Google’s AI to optimize campaign efficacy throughout ad serving. Optimization employed in ads serving is entrenched to the extent that even the campaigns that disabled AI-powered features can still benefit from Google’s AI. Especially without a cross-site identifier like 3PC, Google’s AI can help advertisers derive more value out of otherwise fragmented privacy-preserving signals.

Comparing segments with and without AI-Powered Optimization:

For GDA, segments with AI-Powered optimization showed higher retention of spend and quality after removing 3PCs for interest-based audience solutions. Note, DV3 results did not have sufficient traffic to reliably observe a trend. The comparison indicates that AI-Powered optimization can protect campaign performance post 3PCD.

We explain our observation as the effect of the following:

- Sensitivity of manual targeting[8]: Campaigns with manual targeting can only match ad-impressions with corresponding user interest whereas campaigns with auto-targeting can expand beyond preset criteria. Coarser interests tend to reduce the

[8] The primary distinction between manual targeting and auto-targeting is that auto-targeting is not constrained by targeting criteria set by the advertiser. Even with manual targeting, ML models make inferences about user interests beyond coarser interests observed from privacy preserving signals including Topics API.

pool of eligible ad-impressions, likely lowering spend. Less accurate interests can reduce precision of interest inferred from impressions, lowering quality metrics.
- Flexibility of bidding optimization: Bidding optimization allows bids to shift dynamically to achieve advertiser goals. So we expect campaigns with bidding optimization to have more neutral quality impact (CPD, CVR, CTR) compared to manual bidding[9]. Longer term, we can expect manual bidding advertisers to make similar adjustments of their bids to achieve their internal advertising goals, in a less dynamic fashion; however, a short-term small percentage experiment cannot directly measure this.

In summary, we believe Google's interest-based audiences are a durable advertising solution with user privacy at the forefront, even in the absence of third-party cookies. Our testing reaffirms:

1) the effectiveness of a range of privacy-preserving signals for deriving user interest, especially contextual signals,
2) the value of machine learning and AI-Powered optimization to identify relevant audiences beyond the obvious, and
3) newer privacy preserving solutions such as first party data and Topics API can further help offset the void of third party cookies while enhancing user privacy.

We will continue to test the effect of privacy preserving mitigations including Privacy sandbox as Chrome's 3PCD deadline approaches.

[9] With manual bidding, advertisers bid a fixed value in the auctions they participate in.

Appendices:

Appendix 1 - Concepts:

Interest Based Audience Solutions (IBA)
Interest-based Audience Solutions is an advertising use case where advertisers choose to serve ads based on users' interests. User interests historically have been derived from cross-site tracking data keyed by the third party cookies (3PCs). Currently, we enable targeting using three different types of interest groups, which are Affinity, In-Market, and Custom Audience. As a reminder, this experiment retains 3PC for all critical components such as measurements, and removes 3PC for IBA only.

Topics API
Chrome's Privacy Sandbox introduced the Topics API in January 2022 as one of the solutions to sustain interest-based advertising as part of upcoming third party cookie deprecation (3PCD). The Topics API assigns categories of interest to users based on their past three weeks of browsing activity, using visited hostnames to categorize websites. For example, when a user visits a website for lyrics of a song, they could be assigned the "Music & audio" topic for that week. Users may remove topics manually in their browser or turn off the feature entirely. Because classification occurs on-device, when users delete a topic, it is deleted entirely.

Appendix 2 - Privacy preserving solutions:
At a high level, privacy-preserving signals leveraged are (1) Contextual Data, (2) Publisher 1P Data, (3) AI-powered optimization:

(1) Contextual Data is the data available as part of the page view. Contextual data is the most available source of signals that can be representative of users top-of-mind topics. Any ATP can leverage this signal.
(2) Publisher 1P Data: In addition to Publisher 1P identity, Google publishers can provide raw data along with their ad request through products such as Secure Signals[10], or Publisher Provided Signals (PPS) where publishers can share first party data with buyers. Publisher signals were not incorporated for this experiment. Any supply side ATP can implement such technology. PPS was not part of this experiment;
(3) AI-powered optimizations further optimize ad serving using features from all cookieless sources and improve the efficacy of ad serving. From audience generation, feature engineering, to ads retrieval process, our system continues to expand the application of Google's AI through every part of our system. Also, AI-powered optimization can help derive more value by combined use of privacy preserving signals. Any ATP can optimize using AI similarly.

[10] Secure Signals, a variant of PPS with more publisher control, is available in Open Beta and PPS in Closed Beta.

How Topics can be combined with other signals:

1) <u>All Topics in conjunction</u>: The three topics shared can be combined to infer both more specific commercial intent and/or additional interests from aggregating historical ad requests with those topics. For example, a user might have the topics “Fashion & Style,” “Consumer Electronics,” and “Pop Music.” Together, the topic set could suggest that the user is interested both in a specific category of ads, e.g., stylish speakers, and additional topics, e.g., “TV Shows & Programs” or “Home & Interior Decor.”
2) <u>Topics in combination with contextual data</u>: We can combine Topics with contextual data of the publisher page where the ad is served. Topics and page content in combination can help us infer other user interests or their strengths. For example, if a user is visiting a page about bicycles and “Cycling” is one of their top topics, we can assume a higher level of interest. Lastly, we can use relationships between Topics and contextual data to make further inferences. For example, if we observe that many users with Topic “Outdoor” historically viewed content related to “Woodworking,” we can then use the relationship to make inferences whenever we receive users with Topic “Outdoor.”
3) <u>Topics and Publisher First-Party Data:</u> When available, we plan to use Topics in conjunction with publisher 1P data provided through supply side products such as Publisher Provided Identifiers (PPIDs) at serving time. In this experiment, we use simulated Publisher 1P IDs to help publishers leverage their own users’ behavior within the publisher context (but not across other publishers’ websites and apps). We used simulated PPID due to the technical complexity of the experiment. Starting from 3PC as the core identity, we partitioned 3PC into multiple simulated PPIDs.

Appendix 3- Methodology for Estimating the Impact

Formulas for the reported metrics

Metric	Formula
% change in Spend	$\frac{\Sigma\ Spend_a}{\Sigma\ Spend_b} - 1$

Metric	MH Formula	Weights	Base Metric Definition
MH CTR	$\frac{\Sigma_i w_i CTR_{a,i}}{\Sigma_i w_i CTR_{b,i}}$	$w_i = \frac{Imp_{a,i} \cdot Imp_{b,i}}{Imp_{a,i} + Imp_{b,i}}$	$CTR_{x,i} = \frac{Clicks_{x,i}}{Impressions_{x,i}}$
MH CVR	$\frac{\Sigma_i w_i CVR_{a,i}}{\Sigma_i w_i CVR_{b,i}}$	$w_i = \frac{Clicks_{a,i} \cdot Clicks_{b,i}}{Clicks_{a,i} + Clicks_{b,i}}$	$CVR_{x,i} = \frac{Conversions_{x,i}}{Clicks_{x,i}}$
MH CPD	$\frac{\Sigma_i w_i CPD_{a,i}}{\Sigma_i w_i CPD_{b,i}}$	$w_i = \frac{Spend_{a,i} \cdot Spend_{b,i}}{Spend_{a,i} + Spend_{b,i}}$	$CPD_{x,i} = \frac{Conversions_{x,i}}{Spend_{x,i}} = \frac{Conversions_{x,i}}{Clicks_{x,i} \cdot CostPerClick_{x,i}}$

*(a) refers to treatment and (b) to control
*Imp = Impressions
*MH ratio metrics were computed using changes at ad-group level.

Methodology to compute confidence interval

.. We compute confidence intervals using resampling methods at ad group level or ad-impression level.

Note that the presented confidence intervals aim to capture the statistical noise in measuring various metrics. The other sources of uncertainty, e.g. advertiser reactions to performance change, are not captured here. We expect that for rarer events such as conversions, we will have a larger confidence interval. Confidence intervals also depend on the experiment fraction and consequently the amount of data in analysis.

For reference, an “ad-group” is a finer unit of a Google Ads advertising campaign. An ad group has a unique set of campaign goals and targeting criteria. For example, an advertiser can have multiple campaigns based on different campaign goals (Nike can have a year-round brand campaign AND a mid-year sale campaign). A campaign can have multiple ad groups based on targeting criteria (Nike’s brand campaign has one ad group targeting customers in Washington state AND another in Texas state).

Appendix 4 - Feedback to Chrome:

Like other ATPs, we will continue to run tests when significant updates to Topics are released by Chrome and will provide feedback about future versions of Topics. Based on the experiment results and quality evaluation of Topics API, below are the recommended improvements. Note we have not finely analyzed the privacy implications of these recommendations.

1. Use Full URL to assign a Topic vs hostname
    a. Area: Topic Assignment
    b. Rationale: Most visited sites often contain a wide spectrum of contents (youtube.com, facebook.com, etc.) If topics are deduced from hostnames alone, top topics may consist of generic and low value topics across the majority of the user base. ~50% of users visit 5 or fewer distinct domains in 7 days.
2. Use a more granular and relevant taxonomy
    a. Area: Topics Assignment
    b. Rationale: Even if Topics generate a diverse set of Topics, limited granularity of taxonomy may produce less commercially viable set of Topics for users.
3. Use weights that rank topics based on Topics node or commercial value.
    a. Area: Topics Ranking
    b. Rationale: simple frequency weighting will significantly decrease diversity of Topics given limited diversity of domains visited for an average internet user. Prioritizing Topics at deeper nodes or with higher commercial value can create Topics with more diversity and commercial value for IBA.
4. Deduping Topics that share similar parent nodes when calibrating for Top 5 topics.
    a. Area: Topics Ranking
    b. Rationale: There is no utility in having both Music and Audio when Jazz & Blues as a top topic as they share the same parent hierarchy.
5. Deduping Topics that are the same week over week.
    a. Area: Topics Ranking
    b. Rationale: If top topics are the same week after week, they may provide little incremental value.
6. Removing topics that can be gathered from the current page
    a. Area: Topics shared per publisher
    b. Rationale: If topics shared for a publisher are correlated to contextual, the topics are of limited incremental value. For example, receiving “Sports” Topic on a sports website is not helpful.
7. Assigning top five topics per each API caller
    a. Area: Topics Access

b. Rationale: Global topics generation with access management at adtech level may limit access to a meaningful number of topics; this would be especially challenging for smaller adtech.