Memo – IAB France's competition concerns over Privacy Sandbox's newly released API Topics

Following our meeting of June 23, 2021, DG Competition asked IAB France representatives to specifically document on how Privacy Sandbox could disproportionately advantage Google's services at the expense of most players in the online advertising industry.

In September 2021, we shared with you a first memo offering an overview of how Privacy Sandbox's core privacy model and APIs (FLoC, FLEDGE, Measurement APIs) will benefit large actors such as Google over smaller players in the industry. These concerns remain true as of today.

You will find below an update of our concerns which covers the newly released Topics API specifically.

1. What is Topics?

Topics is the newly released API replacing FLoC API that intends to enable interest-based targeting. Topics API proposal is set to replace individual level targeting with the capacity to target users who have interest in particular topics, based on their browsing history and a classifier developed by Google to map websites to relevant topics. At this point, the initial taxonomy developed by Google includes 350 Topics.

Google has indicated that it intends to leave the Topics taxonomy and the classifier externally maintained and open source. More specifically, it stated that the classifier map sites to topics would be open source as part of Chromium[1]. However, the browsing history that the topics assignments rely on will necessarily be on Google Chrome.

2. How will Topics API benefit Google and other large actors?

While Google develops its advertising products to offer ever more granular targeting possibilities, the Topics API leaves the digital advertising industry with very limited abilities to conduct interest-based advertising campaigns in Chrome.

2.1. Third party data providers are likely to shut down their activity as they will not have any added value in Chrome

Companies called 3d party data providers (Nielsen, Sirdata, Weborama, …) allow advertisers to target a specific audience. These companies collect data about users' interests through third party cookies, through the analysis of a websites content, and combine it with advertisers' first party data. They analyze this enriched data set with their own in-house expertise (various machine learning models, Semantic AI, Data visualization, …) to build various user segments. While third party data providers each have a generic taxonomy (up to several thousands of user segments), they can create an unlimited number of custom segments, depending on the advertiser's advertising strategy. Each user will typically be assigned 30 to 50 segments. User segments vary from common interests such as "sports" or "cooking", to more complex and strategic topics. For example, occasion-based segmentation allows advertisers to target users at a specific moment in their life (i.e.: a person

---

[1] Chromium is a free and open-source web browser project, principally developed and maintained by Google.

moving out or moving in is a valuable user segment for telco companies to communicate about their wifi/phone subscription offers). This competition to offer different audience segments are one the factor underpinning the attractiveness of the sector, as stated by the CMA in its memo on Google's commitments on the Privacy Sandbox.[2]

With Topics however, this expertise in user segmentation built over the years becomes de facto useless in Chrome.

Topics API derives from a pre-existing taxonomy, which contains as of now about 350 topics, against the unlimited number of user segments possible today. Only up to 3 topics are made available for a single user, compared to 30 to 50 today. Topics are assigned to each website by a "classifier", an algorithm which relies on a machine learning model developed and trained by Google. The "caller", that is an ad tech vendor, calls the Topics API to learn at most three top topics of a user, provided the ad tech vendor has observed the user in the past, which means was partnering with the websites the user visited about those specific topics in the past 4 weeks.

Third party data providers will have no possibility to carry out their current activities in the cookieless Privacy Sandbox environment, making consequently Google the exclusive provider of relevant cross-domain data. Competition on the user segments market will no longer be feasible, which comes in direct contradiction with the goal of the Privacy Sandbox, which is to "give companies and developers tools to build thriving digital businesses".[3]

These same companies were also heavily affected by the sudden change from FLoC to Topics. Some of them made substantial strategic investments in research and development, both time and money, to anticipate the change FLoC would bring about. Yet Google announced very abruptly they would abandon FLoC and replace it with Topics, without giving notice to the industry beforehand. This shows that even after having made several commitments accepted by the Competition and Markets Authority in the UK in February, Google's Privacy Sandbox still causes major uncertainty in the digital marketing and advertising industry. As there is still no testing (origin trial) available[4], the timeline before third party cookies' deprecation gets shorter by the day. Players are worried there will not be enough time to adapt to the new solutions upfront, which is key to stay competitive.

2.2. Topics will increase the revenues of larger publishers

Topics API will advantage big publishers. The Topics API does not provide any retribution model for independent publishers. Whereas publishers today are paid by their ad tech partners so the latter can use their content to offer digital advertising services.

Topics will be assigned to websites only based on their hostname. This will cause an imbalance between small verticalized websites that provide valuable insights into users' specific interests, while large multi verticalized websites/platforms will not.

---

[2] Decision to accept commitments offered by Google in relation to its Privacy Sandbox proposals, February 2022 https://assets.publishing.service.gov.uk/media/62052c52e90e077f7881c975/Google_Sandbox_.pdf

[3] https://privacysandbox.com, 24/03/2022

[4] Google announced on March 31st the opening of the opening trials for Topics, Fledge and Attribution reporting but with no precise agenda at this point. https://developer.chrome.com/blog/privacy-sandbox-unified-origin-trial/

2.3 With Topics, ad tech vendors' advertising services will be less efficient and therefore less competitive compared to Google's owned and operated products

With Topics, digital advertising relying on third party data in Chrome will lose in relevance and reach, thereby making ad tech vendors' services less attractive for advertisers.

2.3.1   Selected topics will be vague and irrelevant

First, the topics the advertiser will have access to will be vague and irrelevant for advertising purposes. The three chosen Topics will correspond to the user's activity up to 4 weeks ago. There will not be any possibility to differentiate which topic is the most recent. While advertising solutions today consider both frequency and recency, the Topics API only considers frequency into weighing the top interests of a user. A new user topic will, on average, be made available 3,5 days after the user showing interest for it (as Topics are only refreshed once a week). Google, through search queries, will not suffer from such a delay, and Google advertising services will therefore have a head start compared to Ad Tech companies relying on Topics API.

The topics mentioned in the taxonomy are often vague and too broad to be useful for advertising purposes (ex: Food & Drink, Games, ...). The number of available user segments is limited to 350 topics in the proposal, against an unlimited number of custom segments as of today. Today's proposed taxonomy does not provide for standard socio-demo categories (such as gender or age categories). The process for the taxonomy to be updated remains unclear and will in any case remove Ad Tech companies the ability to compete with custom categories. For example, let's say that a marketing company knows that ads for certain movies are more efficient when served to young adults, using social networks, and who go to the movie theater at least once a month. With Topics API, that marketing company will be unable to propose an ad service built on that custom category, while Google will still be able to do so on its own and operated inventory.

Topics will be assigned to websites based only on the hostname, and not the full URL. This limitation was already mentioned in our first memo regarding FLoC. While users spend most of their time online on platforms such as youtube.com, or google.com (making them very likely to appear in the top 5), these websites will not provide any valuable insight into users' interest. On the other hand, high value information coming from smaller publishers' sites that are less visited by users, but still provide valuable insights into their interests, will be underrepresented.

As of today, Ad Tech companies can define which "user signals" make more sense from a marketing standpoint and build their own "Topics assignment" process (matching users with topics). For example, an Ad Tech company can decide that a user browsing a retail site for a pair of sneakers is a more powerful signal that the user is in the market for sports apparels, than a user reading a sports news site.

The Topics API removes the possibility for an Ad Tech company to perform its own "Topics assignment" process. Topics API effectively transfers the "Topics assignment" process from the hand of Ad Tech companies to the hand of Google - a direct competitor - while leaving Google with full flexibility on its own "Topics assignment" process for its own & operated inventory. This can only lead to one outcome: a low-performing low-utility "Topics assignment" process through the Topics API for Ad Tech companies on the open web, and a high-performance high-utility one for Google on its O&O inventory; leading to major market distortions.

2.3.2   Advertising campaigns will see their reach substantially limited

Digital advertising campaigns will have a limited **reach** compared to today. The taxonomy only contains around 350 topics for now, and the advertiser and its ad tech partner will be left with at most 3 topics per user for his campaign. In comparison and as mentioned above, third party data providers offer 30 to 50 of segments per user today, picked out of an unlimited number of custom segments.

Each topic / user segment will also see its reach substantially limited. Since users can have up to 15 topics computed by the browser, and only 3 are returned on a given site, advertisers interested in a specific topic will see their reach divided by 5 compared to today's cookie-based approach (i.e.: a user is assigned with the "sport" topic. As per current design, only 1 site out of 5 (on average) will access the "sport" topic from that user. If an advertiser targets the sport topic, it means it can target that user only on that one site that had access to the sport topic.). Google acknowledged this would severely impact the ad tech vendor's ability to offer reach to advertisers.[5]

2.3.3   The Topics API will favor larger players and increase barriers to entry for newcomers

This mechanism favors players with a large footprint, in particular on both supply (publishers' sites) and demand (advertisers' sites), such as Google, as Topics coming from demand are likely to carry stronger marketing signals.

Current web advertising mechanisms allow for small and medium Ad Tech companies, or supply-side only Ad Tech companies (SSP), to build reach by syncing their cookies ids with other players. Topics API will put an end to this and therefore favor large SSPs platforms that have much larger coverage and global reach compared to smaller players. It also means that bigger SSPs will also attract more spending from DSPs seeking to target these topics

Consequently, Topics will be an additional strong barrier to entry for new companies in the web advertising market, and a strong barrier for any company without a large supply and demand reach.

2.3.4   Based on its privacy model, Google is building two worlds

Our analysis of Topics API shows that interest-based advertising will be substantially limited in reach, granularity and relevancy.

As a reminder, Google's privacy model suggests that first party data is more respectful of users' privacy online than third party data. Regulators have already expressed their concerns about this vision of privacy, as it favors larger platforms over smaller companies.6

Based on this definition, Google has been developing other advertising products in parallel that continue to allow ever more granular targeting. For example, Google Ads Data Hub allows publishers and advertisers to enrich their first party data with Google's (a data clean room). While several companies have launched similar services, Google has a substantial advantage, having a large set of first party data.

Topics will also allow Google to enhance its walled garden and its numerous owned and operated

---

[5] https://github.com/jkarlin/topics/issues/37
[6] CMA/ICO, *Competition and data protection in digital markets: a joint statement between the CMA and the ICO*, 19/05/2021, n°76.

websites on which users are predominantly logged in, i.e., enabling Google to use granular and almost unlimited first-party data to deliver targeted ads. Therefore, Google will be in an even greater position to capture most of the brands advertising budgets as it will be the only player able to guarantee precise and efficient targeting, especially compared to all the other players from the "Open web" that will have to rely on imprecise and commoditized data to deliver targeted ads.