Memo – How the Privacy Sandbox will benefit Google and large actors over smaller players in the industry

Following our meeting of June 23rd, you 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. This memo offers an overview of how FloC, FLEDGE and measurement APIs will benefit large actors such as Google over smaller players in the industry.

1. The definition of "Privacy" will benefit larger, horizontally, and vertically integrated actors such as Google

Before diving into these specific proposals and their impact on competition, we wish to highlight that the Privacy Sandbox relies at its core on a definition of "Privacy" that has been established by Google unilaterally. This Privacy model can be found here. It was first published before Google announced to deprecate third party cookies on January 14th 2020,[1] and after Google announced the launch of the Privacy Sandbox initiative on August 22nd 2019.[2]

The industry did not have a proper chance to challenge that definition. Even though this Privacy model sets the conceptual ground for the development of other proposals, only 5 engineers have filed as much as 10 "issues" on GitHub,[3] compared to 155 issues filed for the FLEDGE proposal.[4]

This definition suggests that first party data is more respectful of users' privacy online than third party data, while this is nowhere mentioned as such in legal texts. In fact, the Competition and Markets Authority in the UK has expressed multiple times its concerns regarding the fact that large actors such as Google would adopt an interpretation of "privacy" that favors their own business.[5]

The user's control over the sharing of its personal data, or transparency towards the user, are nowhere mentioned in this definition. However, these two concepts are two pillars of regulations such as the GDPR, or even frameworks developed by the industry itself to improve data protection such as the Transparency and Consent Framework. As illustrated below (see FLEDGE, …), this privacy model has direct consequences on competition. On the one hand, the Privacy Sandbox does not allow smaller publishers to create so called Data alliances in order to compete with larger platforms, as this would be considered third party data. On the other hand, Google Chrome's proposal First Party Sets allows the sharing of data within a single corporate entity (ex: Disney, Nestle, Walmart), whereas the line between one entity and several companies remains unclear.

---

[1] SCHUH Justin, Building a more private web: A path towards making third party cookies obsolete, Chromium Blog (Google), 14/01/2020, https://lc.cx/bf0Bak.
[2] SCHUH Justin, Building a more private web, The Keyword (Google), 22/08/2019, https://lc.cx/gfRabd.
[3] https://github.com/michaelkleber/privacy-model/issues
[4] https://github.com/WICG/turtledove/issues
[5] CMA/ICO, *Competition and data protection in digital markets: a joint statement between the CMA and the ICO*, 19/05/2021, n°76.

2. FLoC

2.1. What is FLoC?

FLoC is a proposal from Google Chrome that aims at supporting “interest-based advertising”.

In today’s world, advertisers would use third party Data Providers. The latter build specific segments of audience based on advertisers’ needs (ex: “sport fan”, “foodies”, etc.). In FLoC’s world, Data Providers are no longer in charge of building audience segments, the browser is. It will analyze on recurring occasions (initial implementation was every 7 days) the user’s browsing history and assign the user to one single Cohort (FLoC ID). Other users with a similar browsing history will be assigned to the same cohort.

The cohort itself (FLoC ID) will not reveal any information to advertisers, unlike what they are used to with segments (ex: “sport fan”, “foodies”, etc.). For anyone to identify meaning of a Cohort (are those people more into sport, cooking, finance, etc.), you will need to have a large enough first party dataset with robust insights on user’s interests, which you can cross reference with a user’s FLoC ID. Only then, will you be able to draw conclusion on what interests are most represented inside a cohort and start mapping them to meaningful audiences.

This video from IAB France provides more information on how FLoC could be used by advertisers (in French): Vidéo explicative sur la Privacy Sandbox n°1 - FLoC

2.2. How will FLoC benefit Google and other large actors?

2.2.1. Large actors will be in a better position to use FLoC than smaller actors ...

The larger and more various the first party data set is, the more granular and effective understanding of a given FLoC ID will be. As such, entities, such as Google and Facebook, who have a very strong understanding of their first party users’ interest, with large scale, will be much better positioned than anyone else in market to take advantage of FLoC and establish audience, even beyond their first party data set. Indeed, once you have identified a FLoC ID’s meaning, you can apply this to every user that belong to it, including users that may never have visited your first party properties.

2.2.2. … while large players’ contribution to FLoC will be less significant than smaller players

FLoC only considers the main domain of the website the user visits. Which means a small publisher’s website such as “foodblog.com” will be way more significant than “youtube.com” to determine the user’s interest. However, users spend most of their time on large platforms such as YouTube.

Therefore, we reckon that there is an asymmetry between what smaller players contribute to FLoC and what they actually get from it. This is especially true in comparison to larger players, which contribute less while having more resources to use the FLoC ID, which derives most of its information from smaller publishers.

2.2.3. Consent rates will likely be impacted by FLoC

The ePrivacy Directive requires the user's consent whenever data is stored, or read, on the user's device. With FLoC, the browser will first assign the user to a FLoC ID (data is stored on the device). The publisher will then access the FLoC ID (and read the data stored on the device). Our internal analysis shows therefore that two consents would be necessary to operate FLoC.

Regarding the GDPR, the FLoC ID should probably not be considered as personally identifiable information on its own, as it doesn't relate to one single user. However, FLoC IDs will often be processed along with personally identifiable information (IP address, first party user id, etc.) and should as such be considered personal data in those instances, and thus subject to GDPR.

It should be noted that FLoC has been considered by the Electronic Frontier Foundation (EFF) and many other actors as a greater threat to user privacy, as it will facilitate fingerprinting and discrimination of minorities. Google Chrome has made a proposal to tackle the problem of fingerprinting, but it is still at a very early and theoretical stage.

2.2.4. FLoC's adoption rate by publishers is still uncertain

FLoC only represents a viable alternative to third party cookies if publishers agree to adopt it. FLoC would be indeed subject to the publisher's consent to use its website's domain name for the qualification of FLoC IDs. Large players who compete with Google such as Amazon could decide it is not in their best interest to take part in FLoC. Amazon has indeed already showed signs it would not use FLoC. Other players have announced they would not support FLoC.

If FLoC has a low adoption rate by publishers, it will lose in granularity and meaning for advertisers. This would leave the industry's smaller players without a satisfying alternative to third party cookies for audience targeting purposes, whereas larger players ("walled gardens") have enough first party data and resources to offer audience targeting services.

2.2.5. FLoC tests are too far from reality to draw any solid conclusions

There is still a lot of uncertainty as to whether FLoC will be a viable alternative to third party cookies for audience targeting, as tests are not representative of how FLoC would be used after its implementation. In January 2021, Google claimed FLoC could provide "an effective replacement signal for third-party cookies", with tests showing that "advertisers could expect to see at least 95% of the

conversions per dollar spent when compared to cookie-based advertising".[6] However, these tests were simulations based solely on principles outlined in a FLoC whitepaper. We reckon there was too little evidence to claim FLoC would be a viable replacement for third party cookies.

Similarly, the first FLoC Origin Trials are not carried out in conditions that are very different from what we can expect after the final product's implementation. Origin Trials started in March 2021 and ended at the end of July 2021. Google announced these tests would only take place in a limited number of countries such as the U.S., Brazil, or Australia. Europe, however, was not included in these tests because of concerns regarding GDPR compliance.[7] Google also announced the tests would use a sample of only 0,5% of Chrome users.[8] However, an analysis by Criteo engineers shows that Origin Trials were only made on beta and dev versions of Chrome (92 and 93) on a sample of 0,02% of Chrome users rather than the 0,5% previously announced.[9]

3. FLEDGE

3.1. What is FLEDGE?

FLEDGE is a proposal from Google Chrome that aims at allowing first party data owners (advertisers, publishers, tech vendors) to assign the user's browser to a list of "Interest Groups" (IG) for a limited period. This will give them the ability to activate this data for advertising purposes, such as retargeting or audience extension.

Today, these companies assign users to a segment (ex: "sport fan", "foodies", etc.). With FLEDGE, they won't be able to keep data regarding the interest groups the user's browser is assigned to, the browser itself will store this data, as well as the ad creative's code, the bidding logic from buyers and the ad scoring logic from sellers. The browser will use this information to run the FLEDGE auction.

In fact, two auctions will take place. The first auction relates to a standard contextual (and hence cookieless) ad and is executed without access to the IGs the user belongs to. The second auction (the "FLEDGE auction") is run by a single sell side partner (the auction runner) and takes as input the first auction to determine a floor for the ad evaluation. The browser uses the auction runner's ad scoring logic to score the FLEDGE auction. If the FLEDGE auction's winner offers a price below the floor set in reference to the first auction, it is assigned a "0" and eliminated.

The "auction result" that get assigned the highest score (if any) is elected "FLEDGE winner" and returned to the auction runner as an opaque object. It means, the ad winning the FLEDGE auction is unknown from the auction runner as well as the resulting price. The only bit of information that is leaked by the FLEDGE auction is whether there was a FLEDGE ad above the normal auction which allow the auction runner to know whether it should attempt to render the normal auction ad or the FLEDGE winner. In case there is a FLEDGE winner the auction runner will then pass the opaque auction result

[6] BINDRA Chetna, Building a privacy-first future for web advertising, Ads & Commerce Blog, 25/01/2021, https://lc.cx/K-EPo1.

[7] SCHIFF Allison, Google Will Not Run FLoC Origin Tests In Europe Due To GDPR Concerns (At Least For Now), *AdExchanger*, 23/03/2021, https://lc.cx/g9kbMe.
[8] KLEBER Michael, *FLoC OT - 0.5% of Chrome page loads limit #83*, GitHub, 31/03/2021, https://lc.cx/fEikiA.
[9] ROUZEAUD Antoine, FLoC Origin Trial: What We Observed (1/4), Medium, 29/07/2021, https://lc.cx/WablaR.

to a “fenced frame” (another privacy sandbox proposal for new type of iframe with 0 communication with the outside world). Only the fenced frame can decode the FLEDGE winner ad and render it.

3.2. How will FLEDGE benefit Google and other large actors?

3.2.1. FLEDGE can only be used by first party data owners who also run the buying strategy

When assigning a user into an interest group, first party data owners must also provide the ad’s creative code as well as the bidding logic for evaluation that will be used later by the browser for on-device auctions. Unlike in today’s world, where the data owner can delegate the buying strategy to a third party, the data owner must also run the buying strategy. This means the data owner must also be the buyer, or buyer’s tech provider.

That makes the role of data provider much harder to sustain, since you can no longer simply focus on audience creation, but need also to operate the buying, which is a very different activity and as such requires those company to invest time, resources, and money to sustain their business compared to today. Data providers that do not also operate the buying strategy will not be able to leverage that proposal without a substantial evolution of their core business.

This will benefit large actors who are vertically integrated in the advertising supply chain. Google happens to already be both a data provider (they provide Google’s first party audience as part of their DV360 offering) and a tech provider (they offer DV360 as a DSP to buyers). As such, unlike most existing data providers, Google is in no needs to change its business model to keep offering data provider services to buyers.

3.2.2. FLEDGE replaces Header Bidding by a single auction, run by a single sell side partner chosen by publishers, which in most cases is likely to be Google

Ad Tech providers on the sell side (SSPs) had been quite challenged in their business model around 2015 and even before, when Header Bidding didn’t exist. Google was using its dominant position in the ad serving market to self-preference its own products,[10] notably with the “waterfall” system, which allowed it to analyze the bids one by one rather than all of them together in real time. In reaction to the waterfall, Ad Tech providers developed a new technique called Header Bidding. SSPs had a chance to participate in a “fairer” auction against Google’s AdX, and compete with one another in real time. The industry has quickly adopted Header Bidding as its main auction system.

With FLEDGE, only a single FLEDGE auction can be supported, meaning a single sell side partner will oversee it (compared to today where publishers work with several SSPs). Considering Google’s current

[10] The French Competition authority recognized on June 7th, 2021, Google had engaged in such anticompetitive practices in the past and sentenced Google to a €220 million fine, which Google has not challenged.

dominant position on the sell side, especially as an ad server, there is no doubt that vast majority of publishers will pick Google as their single sell side partner. Other SSPs will then lose access to all first party data-based demand (retargeting, audience extension, etc.) which represents a large part of digital advertising. For the few non-Google sell side partners that may be picked for the publishers, it will be very hard for them to build and maintain support for the FLEDGE proposal, considering the complexity and resources required to support it. As such, Google is likely to result as the only sell side partner for all the first party data-based demand, reinforcing its dominant position.

3.2.3. FLEDGE does not allow publishers' data alliances, while another proposal in the Privacy Sandbox allows data sharing within a single corporate entity

For long, Google as dominated the advertising industry as a publisher, with no threats from other publishers beside other GAFA. This domination was the result of a scaled first party data and inventory. However, in past few years, several publishers, in different markets, have worked together to create alliances, that will provide advertisers with scaled first party data and inventory, allowing them to finally provide a competitive offer to the likes of Facebook and Google. Such alliances include Skyline and Gravity in France, Ozone in UK, Wemass in Spain, Nonio in Portugal, etc.

Those alliances rely on publishers that form them to be able to share their first party data and activate it on each other's inventory. This ability is made possible thanks to third party cookies which allow them to build common audience on top of them. With FLEDGE, this is no longer possible as publishers cannot build common audience. Interest Groups are created based on interaction from a single website.

There were some attempts from the industry to suggest FLEDGE evolutions that would allow such alliances to continue to exist, but they were rejected right away by Google Chrome team, with limited justifications (see https://github.com/WICG/turtledove/issues/124 for example).

If no publishers' alliances for first party data is supported, we will get back to a world where Google and Facebook dominate without competition from publishers that lack their scale.

Interestingly, Google Chrome has made another proposition called First Party Sets, which allows data to be transferred within a group of companies (ex: Disney or Nestle). The paradox between Chrome's refusal to allow data alliances between publishers on one side, and their First Party Sets proposal on the other side, underscores the core problem of the Privacy Sandbox, and more generally of other walled gardens recently. Their vision of Privacy favors first party data over third-party data, whereas this distinction is nowhere to be seen in legal texts. In fact, the CMA has multiple times expressed its concerns that large actors would interpret user privacy in way that would harm competition.

4. Attribution Reporting:

4.1. What is Attribution Reporting?

Attribution Reporting relates to attribution. Attribution is the process of identifying a set of user actions ("events") across screens and touch points that contribute in some manner to a conversion. With "Attribution Reporting", the advertiser would predefine a set of metadata related to the desired

outcome (ex: which sites he hopes the ad will redirect the user to). This set of data would be downloaded by the browser when the user clicks on the ad. If the user does indeed visit the site identified by the advertiser, the browser will notify the advertiser after a certain delay (ex: 1 day).

4.2. How will Attribution Reporting benefit Google and other large actors?

4.2.1. Larger players will have more qualitative performance reports than smaller players

Performance reporting will lose in quality/details. Differential Privacy implies a certain level of "noise", which means a certain amount of random, fake attributions to guarantee user privacy. This level of noise is not a percentage, it is a fixed value, independent from the amount of analyzed data. Which means that the reporting of "rare" events (ex: conversions) will be more impacted. Small publishers have less traffic than bigger ones such as Google and will therefore be more impacted than the latter by this proposal.

4.2.2. Smaller players will be left with limited abilities to choose an effective attribution model

This proposal will limit the ability of smaller players to choose an attribution model that is the most effective, while larger players will have enough first party data to

For now, this proposal uses the "last click" attribution model. Even though it is the most used across the industry, this model only attributes conversions to the user's last click on the ad. This is an important limit, as the last click only tells a part of the story in the marketing funnel. The ad may have had a small impact on the user's decision to visit the site and buy a product. Some models ("first touch", "multi touch") try to address this issue by assigning a value to other events/touchpoints in the marketing funnel. However, these are not supported by the Attribution Reporting proposal. A proposal is currently discussed to allow View Through Measurement, but it is still at a very early and theoretical stage (view-Through Conversions are what happens when a customer sees an ad (but doesn't click), and then later completes a conversion on your site).

5. Aggregated Reporting API

5.1. What is Aggregated Reporting API?

This proposal relates to the measurement of a campaign's performance. Performance reports would be sent to the advertiser in aggregate form and be related to several users to avoid identifying them individually.

5.2. How will Aggregated Reporting benefit Google and other large actors?

The number of sites appearing in the advertisers' performance reports will be limited to those which have the most traffic, to protect the user's privacy. Several problems derive from this feature. First, DSPs have the contractual obligation to communicate to the advertiser on which sites its ads were shown. Second, small publishers will again be disadvantaged compared to bigger platforms, as the latter have more traffic. Small publishers will therefore disappear in the eyes of advertisers, who won't choose them for their next campaign, though they may have proportionately led to a fair amount of conversions.