ACCC DIGITAL ADVERTISING SERVICES INQUIRY

SUBMISSION ON PRIVACY SANDBOX AND COMMON IDS

2 July 2021

INTRODUCTION

1. This submission describes the Privacy Sandbox and implications of Google's blog posts on 3 March 2021 ("*Charting a course towards a more privacy-first web*"), 11 June 2021 ("*Our commitments for the Privacy Sandbox*") and 24 June 2021 ("*An updated timeline for Privacy Sandbox milestones*").[1] We also explain why, in the context of these developments, we have concerns with the ACCC's proposals for a common transaction ID and common user ID.[2]

2. In summary:

    - There has been a fundamental shift in user expectations of privacy with respect to ad tracking across the web.
    - The Privacy Sandbox aims to provide online advertising with a more sustainable, privacy-safe model.
    - The Privacy Sandbox does not give Google ad tech products an advantage over rivals.
    - We believe common ID proposals are not in line with the sustainable, privacy-safe future of online advertising.

WHAT IS THE PRIVACY SANDBOX?

3. The Privacy Sandbox initiative aims to create web technologies that both protect people's privacy online and give companies and developers the tools to build thriving digital businesses to keep the web open and accessible to everyone.

4. The proposed solutions will restrict tracking of individuals as they move across the web, and replace legacy, data-intensive mechanisms like third-party cookies with safer solutions that protect user privacy.

5. Chrome has invited all web community members — web browsers, online publishers, ad tech companies, advertisers, and developers — to participate in the development and testing of the proposed new technologies. Web community members can contribute

---

[1] See "*Charting a course towards a more privacy-first web*" available here: https://blog.google/products/ads-commerce/a-more-privacy-first-web/; "*Our commitments for the Privacy Sandbox*" available here: https://blog.google/around-the-globe/google-europe/our-commitments-privacy-sandbox/; and "*An updated timeline for Privacy Sandbox milestones*" available here: https://blog.google/products/chrome/updated-timeline-privacy-sandbox-milestones/.

[2] ACCC Interim Report on Digital Advertising Services Inquiry (December 2020), pp 183 - 184.

to the public discussions in forums like the W3C. Developers have also been invited to join the testing of the proposals in so-called "origin trials".

WHY HAS THE PRIVACY SANDBOX BEEN INTRODUCED?

6. Advertising is essential to keeping the web open for everyone. This ecosystem is put at risk, however, if privacy practices do not keep up with the changing expectations of consumers. Privacy is also at the forefront of public dialogue as people seek to understand and control how their personal information is used online. Consumers increasingly expect, and data privacy laws require, strict controls over ad-tracking tools like cookies and ad identifiers. For example:

    - In its Issues Paper for the Privacy Act Review, the Attorney-General's Department found that data privacy was a priority when choosing a digital service.[3]

    - According to a survey done by Deloitte in Australia, 66% of respondents confirmed that they had backed out of purchasing or using a service or closed an account completely due to privacy concerns.[4]

7. Data protection regulators around the world are also grappling with how best to protect consumer privacy in the ad tech ecosystem.[5]

8. We're focused on meeting those expectations and requirements.

9. We know that some of the information collected by sites and third-parties is necessary to facilitate the ads supported ecosystem and therefore fund the rich content and services users expect. But the tools used to provide this have gone far beyond their original intent in their ability to recognise users, their online activity and the devices

---

3 Data privacy ranked ahead of reliability, convenience and price. See the "*Privacy Act Review Issues Paper*" (October 2020) , p. 14, available here: https://www.ag.gov.au/system/files/2020-10/privacy-act-review--issues-paper-october-2020.pdf.

4 See "*Deloitte Australian Privacy Index 2020: Opting-in to meaningful consent*" p 17, available here: https://www2.deloitte.com/content/dam/Deloitte/au/Documents/risk/deloitte-au-risk-australian-privacy-index-2020.pdf.

5 For example: (i) the UK's Information Commissioner's Office (**ICO**) is undertaking a substantial and lengthy review of real time bidding (**RTB**); (ii) Ireland's Data Protection Commission opened a statutory inquiry into Google's RTB product to assess whether the processing of personal data it carries out is in compliance with the relevant provisions of the GDPR; and (iii) the FTC in the US has been put under pressure to investigate RTB and the privacy implications thereof.

they use.[6] Third-party cookies, fingerprinting, and other tracking technologies have proliferated on the web and need to be addressed in order to protect people's privacy online.

10. As user privacy concerns have evolved in recent years, some browsers have taken steps to protect privacy by blocking or removing third-party cookies entirely. We believe that - without effective alternatives in place - this can negatively impact critical web functionality, such as advertising and fraud prevention. To ensure publishers and developers can provide content funded by advertising in a privacy-preserving way, the web requires major technological innovations, which is what the Privacy Sandbox seeks to help develop.

11. The goals of the Privacy Sandbox are as follows:

    - **Prevent tracking as users browse the web**. People should be able to browse the web without worrying about what personal information is being collected, and by whom. The Privacy Sandbox initiative therefore aims to remove commonly used tracking mechanisms, like third-party cookies[7] and block covert techniques, such as fingerprinting.[8]

    - **Enable publishers to build sustainable sites that respect your privacy**. Website developers and businesses should be able to make money from their sites and reach their customers without relying on intrusive tracking across the web. The Privacy Sandbox initiative is therefore developing innovative, privacy-centric alternatives for key online business needs, including serving relevant ads.

    - **Preserve the vitality of the open web**. The open web is a valuable resource of information, with a unique ability to both share content with billions of people, and tailor content to individual needs. The Privacy Sandbox proposals aim to both protect users' safety online, and maintain free access to information for

---

6 Cookies were first introduced in 1994 as a mechanism to allow a site to recognise if a user had visited before and enable functionality to support a shopping cart. Cookies are a very flexible technology that have allowed for significant positive innovation over the years, but have also enabled some less desirable data practices. With a single technology, like cookies, there isn't a way to permit certain use cases and restrict others.

7 A "cookie" is a small piece of data stored on a user's computer via the user's browser when they visit a website. Third-party cookies are stored by a service that operates across multiple sites. For example, an ad platform might store a cookie when a user visits a news site. First-party cookies are stored by a website itself.

8 Fingerprinting is when information is collected about the software and hardware for the purpose of identification.

everyone, so that the web can continue to support economic growth, now and for the future.

12. In short - the Privacy Sandbox will lay new foundations for a safer, more sustainable, and more private web.

HOW WILL THE PRIVACY SANDBOX PROPOSALS PROTECT USER PRIVACY?

13. The fundamental principle of the Privacy Sandbox is that a browser can create a protected space around the personal data users share with the sites they visit—from their web visits to the email address entered into a form. This data is then safeguarded from being accessed in a way that can, over time, identify individual users.

14. Building from this principle, the Privacy Sandbox proposes using the latest privacy techniques, like differential privacy,[9] k-anonymity,[10] and on-device processing,[11] to deliver great web experiences, with much greater privacy protections. With these innovations, we expect third-party cookies will be made obsolete, allowing browsers to remove third-party cookies while still supporting key website capabilities.

15. An important long-term goal of the Privacy Sandbox is to help prevent fingerprinting (i.e. information collected about users' software and hardware for the purpose of identification) by limiting the over-collection of data.[12] To do this, the Privacy Sandbox aims to introduce a limit on how much information a site can access about a visitor and helps manage this "privacy budget" effectively. For example, a site must be specific about what information it needs from the browser, and sites that access too much information can be stopped.[13]

---

9 Differential privacy is a system for sharing information about a dataset to reveal patterns of behaviour, without revealing private information about individuals or whether they belong to the dataset.

10 K-anonymity is a measure of anonymity within a dataset. If a user has k=1000 anonymity, they can't be distinguished from 999 other individuals in the dataset.

11 On-device processing is where computation is performed "locally" on a device (e.g. a user's phone or computer) without communicating with external servers.

12 Fingerprinting often happens in the background of apps and websites which makes it difficult to combat and block. Users are also not able to see or delete their fingerprint (unlike cookies). See "*Google's Response to the Interim Report*" (12 March 2021), para 92 available here: https://www.accc.gov.au/system/files/Google%20%28March%202021%29.pdf.

13 For example, a website needs to know certain information to work correctly, like the size of your screen, or what language you use. But if a website collects too much of this kind of information, it can be used to create a digital "fingerprint", which can then be used for pervasive tracking.

16. Further detail on the proposals and/or experiments currently under development are available below.

FURTHER DETAIL ON PROPOSALS UNDER DEVELOPMENT

17. We discuss in greater detail below the current plans for proposals and/or experiments under development, categorised by use case:

A. Showing relevant content and ads

18. Tailored content is critical for the open web and this includes showing people relevant ads.

19. People generally prefer seeing ads that are relevant and useful to them. Plus, these ads also bring more business to advertisers and more revenue to the websites that host their ads. The site publisher can then invest in creating more content, making the web an accessible information source for everyone.

20. To do this, the digital advertising industry has often relied on third-party cookies to track an individual user's browsing history in order to serve the most relevant ads based on either their interests or the sites they have visited previously.

21. One proposal in the Privacy Sandbox is to cluster people with similar browsing patterns into large groups, or "cohorts". This "safety in numbers" approach effectively blends any individuals into a crowd of people with similar interests. The new method is called Federated Learning of Cohorts (**FLoC**).

22. With FLoC, the browser uses on-device computation to place each user in a "cohort". On-device computation is performed "locally" on a device (e.g. a user's phone or computer) without communicating with external servers. This group is large enough — numbering in the thousands — to ensure that individuals can't be identified. The members have similar enough browsing habits, though, that they would likely be interested in the same kind of content or ads.

23. An individual's browser will develop the cohort based on the sites that an individual user visits. The algorithms might be based on factors such as the domain of the sites visited. As an individual's browsing behaviour changes, their cohort will change too, but the algorithm that turns input features into cohort assignments should remain stable.[14]

---

[14] See "*Federated Learning of Cohorts (FLoC)*" available at: https://github.com/WICG/floc.

24. The central idea is that web browsing data used as an input feature to generate the cohort ID are kept on the local browser and are not uploaded elsewhere. In the case of Chrome, the individual-level browsing history data would not be shared with Google's ads systems for the targeting or measurement of digital advertising.

25. The browser only exposes the generated cohort ID. The cohort ID would be available to any site that accesses the FLoC API.

26. We believe that, in the long run, ad tech providers could use cohort IDs (potentially paired with TURTLEDOVE (a form of remarketing) - see below) as a feature in their ads personalisation algorithms.

27. For further information, please see the FLoC whitepaper and page on Github.[15]

B. Showing ads based on sites or products a user has viewed before

28. The Privacy Sandbox proposes a new way to show ads based on sites or products a user has viewed before which doesn't rely on cookies to track the sites an individual user visits or the products they view. Instead, as a user moves across the web, the sites belonging to advertisers they've visited can inform their browser that they would like a chance to show that user ads in the future. Advertisers can also directly share information with the user's browser including the specific ads they'd like to show that user and how much they'd be willing to pay to show that user an ad. When the user then visits a site with ad space, an algorithm in their browser helps inform what ad might appear.

29. The current proposal for this approach is called First Locally-Executed Decision over Groups Experiment ("**FLEDGE**"). FLEDGE is an early prototype for ads serving in the TURTLEDOVE (a form of remarketing) family, appropriate for experimentation before a fully-featured system is ready. It is one way to apply the principles of the Privacy Sandbox to these types of ads, keeping the information about the sites a user has visited and the ads they see separated.

30. During 2021, Chrome plans to run an "origin trial" for a first experiment involving FLEDGE as follows:

    a. Ads will be targeted at interest groups (for example, those who have visited a particular shoe brand's online store), which will be stored in the browser. Every interest group will have an owner, who is ultimately responsible for the group's membership and usage, but can delegate those tasks to third parties if they so

[15] Available here:the FloC Whitepaper and WICG / floc.

desire. Many sorts of entities might want to be owners of interest groups. Some examples include:

i. An advertiser (or a third-party working on an advertiser's behalf) might create and own an interest group of people whom they believe are interested in that advertiser's product. Classical remarketing/retargeting use cases fall under this example.

ii. A publisher (or a third-party working on a publisher's behalf) might create and own an interest group of people who have read a certain type of content on their site.

iii. A third-party ad tech company might create and own an interest group of people whom they believe are in the market for some category of item. They could use that group to serve ads for advertisers who work with that ad tech company and sell things in that category.

b. The information shared with Google services for the purposes of ad targeting would be the same as shared with any non-Google service.

c. Each interest group the browser has joined may have an opportunity to bid in the FLEDGE auction as “buyers”, with the auction taking place on-device.

d. Buyers have three basic jobs in the on-device ad auction:

i. Buyers choose whether or not they want to participate in an auction.

ii. Buyers pick a specific ad, and enter it in the auction along with a bid price and whatever metadata the seller expects.

iii. Buyers perform reporting on the auction outcome.

e. In the future, advertisers would be able to use any DSP or ad network that supports the relevant technical APIs to bid in auctions developed from FLEDGE.

f. FLEDGE auctions will be initiated by a “seller” invoking a javascript API inside the publisher's page. Sellers have three basic jobs in the on-device ad auction:

i. Sellers decide (a) which buyers may participate, and (b) which of the bids from those buyers' interest groups are eligible to enter the auction. This lets the seller enforce the site's rules for what ads are allowed to appear on the page.

ii. Sellers are responsible for the business logic of the auction. Javascript code considers each bid's price and metadata, and calculates a

"desirability" score. The bid with the highest desirability score wins the auction.

iii. Sellers perform reporting on the auction outcome, including information about the clearing price and any other pay-outs. The winning and losing buyers also get to do their own reporting.

g. Many parties might act as sellers: a site might run its own ad auction, it might include a third-party script to run the auction for it, or it might use an SSP that combines running an on-device auction with other server-side ad auction activities. Sites would be able to use any SSP that supports the relevant technical APIs to run auctions developed from FLEDGE for ad space on their sites. Sites would ultimately remain in control of how their web inventory is monetised and it is intended that these FLEDGE auctions may be run alongside other auctions of third-party SSPs and/or Google SSPs.

h. After the auction is complete, the browser will render the winning ad.

31. For further information, please see the First Experiment (FLEDGE) page on Github.[16]

C. Measuring digital ads

32. When a business runs an ad campaign, it is important for them to understand how many people see each ad and if it results in any conversions (for example, purchases or sign ups). At the moment, measuring the effectiveness of an ad relies on cookies stored with every action a user takes.

33. These identifiers can be used to enable other forms of cross-site tracking. This doesn't have to be the case. A new API surface can be added to the web platform to satisfy this use case without them, in a way that provides better privacy to users.

34. On 9 April 2021, a Google blog post explained how the latest proposals in the Privacy Sandbox can solve for key conversion measurement use cases on the web while preserving privacy.[17]

35. Chrome's conversion measurement proposals centre around an API that would have the capability to report both event-level and aggregated information. Event-level information is helpful when businesses need data to be more granular, such as deciding how much to bid on impressions or modeling conversions. Aggregated information is

---

[16] Available here: https://github.com/WICG/turtledove/blob/main/FLEDGE.md.

[17] See "*Privacy-first web advertising: a measurement update*", available here: https://blog.google/products/ads-commerce/2021-04-privacy-sandbox-measurement/.

helpful for summarising campaign performance, like reporting total conversion value or return on investment.

36. To make sure that the API preserves privacy, and that any data reported can't be used to track individual people as they move across the web, the API uses one or more of the following techniques:

    a. Aggregate the data that is reported so that each person's browsing activity and identity remain anonymous among a large group of conversions.

    b. Limit the amount of information reported about each conversion, so it's not possible to expose the identity of the person behind the conversion.

    c. Add "noise" to the data reported, which protects an individual's privacy by including some random data along with the actual conversion results.

37. The Chrome team recently shared new proposals for how the API could apply these privacy considerations while reporting view-through conversions and cross-device conversions.

38. For view-through conversion measurement,[18] Chrome proposes that advertisers use the event-level capability of the API to get a report on the conversions that happen on their website and are attributed to ad views across the web. The browser would enable this by registering the ad impressions that take place across websites and then matching any conversions that happen on an advertiser's website back to the initial views. To prevent any conversion data from being used to track people individually, the Chrome API would limit the amount of information shared about each conversion and add "noise" to the data.

39. Then, when advertisers are interested in reporting on the total number of view-through conversions, for a video ad campaign as an example, Chrome proposes that they can use the API's aggregate reporting capability. This would allow advertisers to get more precise information on key metrics for the overall campaign without compromising people's privacy. That's because aggregate reporting keeps people's identities and their browsing histories anonymous as it only shares data across a large group of conversions. For further information, see the relevant GitHub page.[19]

---

[18] View-through conversion measurement is event level conversion measurement of ads which the user viewed but did not click on and visit the advertiser site.

[19] See https://github.com/WICG/conversion-measurement-api/blob/main/event_attribution_reporting.md

40. For cross-device conversion measurement,[20] Chrome proposes that advertisers use the API's event-level capability to report on the conversions that happen on their website and are attributed to ad views or clicks that happen on another device. This would only be possible if the people converting are signed into their browser across their devices. Access to this capability would enable cross-device measurement for all participating ad providers and networks. For further information, see the relevant GitHub page.[21]

41. It is intended that the measurement APIs developed as part of the Privacy Sandbox can be used on the Chrome browser alongside other measurement tools, to the extent these tools don't rely on third party cookies. The data generated through the measurement APIs developed would be sent to the advertiser (or any third-party acting on their behalf that the advertiser chooses). It would also be stored locally on the Chrome browser only for as long as required to complete the function of the API. The data would not be shared with other Google services (other than with the advertiser's permission e.g. if the advertiser were using a Google DSP or measurement product).

D. Fight spam and fraud on the web

42. Both publishers and advertisers currently face online spam and fraud, making it hard to distinguish real people from bots or malicious attackers. But the current techniques used to distinguish humans from bots (such as device fingerprinting) can also be abused and used to track people.

43. To help sites combat fraud without tracking people, the Privacy Sandbox is introducing the concept of trust tokens. Trust Tokens is a new API to help combat fraud and distinguish bots from real humans, without passive tracking. Based on a user's behaviour on a site, like regularly logging into an account, a site can choose to issue a trust token to a user's browser. The tokens are stored by the user's browser, and can then be checked by other sites that display advertising and want to verify if that user is a real human.

44. Trust Tokens are encrypted, so it isn't possible to identify an individual or connect trusted and untrusted instances to discover user identity.

45. It is intended that publisher sites would be able to take advantage of the utility of Trust Tokens whether or not they are using Google ad tech products to sell advertising on their sites.

---

[20] Cross-device conversion measurement allows attribution of events across multiple devices operated by the same person.

[21] See https://github.com/WICG/conversion-measurement-api/blob/main/cross_device.md

46. For further information, see the relevant web.dev page.[22]

GOOGLE ANNOUNCEMENT ON 3 MARCH

47. After Chrome announced its intent to remove support for third-party cookies, and after our work with the broader industry on the Privacy Sandbox, we continued to get questions about whether Google will join others in the ad tech industry who plan to replace third-party cookies with alternative user-level identifiers.

48. As part of a Google blog post on 3 March 2021, we made explicit that once third-party cookies are phased out, we will not build alternate identifiers to track individuals as they browse across the web, nor will we use them in our products.[23]

49. Our own choice not to use such identifiers is informed by the fact that we don't believe such identity solutions will meet evolving consumer privacy expectations or regulation. This decision also demonstrates our confidence in the viability of the Privacy Sandbox APIs for delivering and measuring relevant ads on third-party inventory.

50. We also recognised in the blog post that developing strong relationships with customers has always been critical for brands to build a successful business, and this becomes even more vital in a privacy-first world. We will continue to support first-party relationships on our ad platforms for partners, in which they have direct connections with their own customers. And we'll deepen our support for solutions that build on these direct relationships between consumers and the brands and publishers they engage with.

51. Jakub Otrzasek, head of data analytics, Asia Pacific at MightyHive, has described this as "*… the best long-term decision as Google is coming out on the side of consumers and will not use "alternate identifiers" in the Google ecosystem.*"[24]

---

22 See "*Getting started with Trust Tokens*" available here: https://web.dev/trust-tokens/.

23 See "*Charting a course towards a more privacy-first web*" available here: https://blog.google/products/ads-commerce/a-more-privacy-first-web/.

24 See "*Google commits to removing identifiers from Chrome – for good*", (4 March 2021), available here: https://mumbrella.com.au/google-commits-to-removing-identifiers-from-chrome-for-good-671855.

GOOGLE'S AD TECH PRODUCTS WILL HAVE NO ADVANTAGE OVER RIVALS

52. Some have asked whether our ad tech products would hold a competitive advantage in the supply of ad tech services following deprecation of third-party cookies and implementation of the APIs. Others have inquired about our role as a "designer" of the APIs, as well as a competitor in the supply of ad tech services. We want to make it clear that removing third-party cookies will not give our ad tech products a competitive advantage. Further, the Privacy Sandbox is a collaborative process with a wide range of players across the web ecosystem, and Privacy Sandbox technologies are being developed as open standards.

Removing third-party cookies will not give our ad tech products a competitive advantage

53. Third-party cookie deprecation will not give our ad tech products a competitive advantage, for the reasons given below.[25]

54. First, like others in the industry, our own display advertising activities will be impacted the same way as others when third-party cookies are removed. This is because Chrome treats cookies associated with Google domains set on non-Google websites - including the websites of publishers using our ad products - as third-party cookies. Further - as explained in para. [48] above - as part of a Google blog post on 3 March 2021, we made explicit that once third-party cookies are phased out, we will not build alternate identifiers to track individuals as they browse across the web, nor will we use them in our products.[26]

55. Second, it is not the case that access to Google owned and operated data means our ad tech products avoid the impact of third-party cookie deprecation. As we have outlined in previous submissions, our use of first-party data from individual consumers when bidding for or targeting ads on third-party display inventory is already extremely limited.[27] For example - to clarify some common misconceptions:

    a. Our ad tech products currently use extremely limited data from individual

---

25 This submission focuses on our ad tech products, given that the stated focus of the ACCC's Digital Advertising Services Inquiry is *"ad tech services that are used to deliver advertisements on the websites and apps that do not operate their own integrated ad-tech services, rather than companies which sell their own ad inventory to advertisers entirely through their own ad tech services (such as Facebook)*" (ACCC Interim Report (December 2020), footnote 2).

26 See "*Charting a course towards a more privacy-first web*" available here: https://blog.google/products/ads-commerce/a-more-privacy-first-web/.

27 See, for example, Google's Response to the ACCC Interim Report, para. 131 and Google's Response to second ACCC RFI, items 12 and 13.

consumers from our user-facing services when bidding for or targeting ads on third-party websites.[28] As we have stated previously, we do use aggregated (non-individual) data.[29] But this data is not used to identify or track individual consumers for the purposes of ad targeting on third-party websites and apps and does not replace third-party cookies.[30]

b. Google's ads products do not use users' data from Chrome Sync to target or measure ads on third-party inventory on the web. In fact, other than specific use cases for spam and abuse, Google's ads products do not access synced Chrome history. Chrome Sync is opt-in and designed for users to personalise their experience across all their devices - but Google does not use it directly as a signal for targeting ads.

c. Our ad tech products do not use information collected through Google Analytics to target ads on other publisher's websites (or indeed on our own properties).[31]

56. There has been speculation about whether we will use data from our owned and operated properties or products (e.g. Chrome Sync or Google Analytics) going forwards after the removal of third-party cookies. We confirm that we will not use individual-level user data from the sources listed below in our ads systems to track users for the targeting or measurement of digital advertising on third-party inventory on the web after the removal of third-party cookies:[32]

    a. Google's current and future user-facing services, including Android;

---

28 We note that the ACCC has received several incorrect submissions on this point and we are happy to provide any further detail the ACCC requires.

29 For example, we do use aggregated (non-individual) data from Search to inform our understanding of web content.

30 See Google's Response to ACCC's Interim Report (12 March 2021), footnote 175.

31 This is because Google Analytics customers own the data collected by Analytics on their properties and we do not use or share the data collected by Google Analytics except as directed by the Google Analytics customer via the data sharing settings. Google Analytics customers can however choose to use the information collected through Google Analytics for their own advertising purposes, and there are integrations between Google Analytics and Google's ads products that allow remarketing using audiences created in Google Analytics. We have provided further confidential information to the ACCC on the extent of data sharing by Google Analytics customers.

32 Further, in relation to Google's owned and operated inventory - to clarify, we will not use individualised user data from the sources listed below in our ads systems to track users for the targeting or measurement of digital advertising on Google owned and operated inventory on the web: (a) from a user's Chrome browsing history, including synced Chrome history; and (b) from a publisher's Google Analytics account (see footnote 33).

b. a user's Chrome browsing history, including synced Chrome history;

c. a publisher's Google Analytics account;[33]

d. uploaded by an advertiser to Customer Match in accordance with Google's Customer Match policy.

57. Our ad tech products will not therefore have a competitive advantage in advertising on third-party sites through use of data generated or processed by other Google services or products, as discussed above.

58. Third, on 24 June 2021, we shared the latest on the Privacy Sandbox initiative including a timeline for Chrome's plan to phase out support for third-party cookies.[34] We noted that each proposal goes through a rigorous, multi-phased public development process, including extensive discussion and testing periods. We stated that we need to move at a responsible pace to allow sufficient time for public discussion on the right solutions, continued engagement with regulators, and for publishers and the advertising industry to migrate their services. This is important to avoid jeopardizing the business models of many web publishers which support freely available content. For Chrome, specifically, our goal is to have the key technologies deployed by late 2022 for the developer community to start adopting them. Subject to our engagement with the United Kingdom's Competition and Markets Authority (CMA), Chrome could then phase out third-party cookies over a three month period, starting in mid-2023 and ending in late 2023.

59. In summary, Google will be impacted by these changes, which do not put our own display advertising business at a competitive advantage.

Google's role in the Privacy Sandbox proposals and as a supplier of ad tech services

60. We would like to address the concerns about Google's role in the Privacy Sandbox in light of the fact we also supply ad tech services.

61. *First*, the Privacy Sandbox technologies are being developed both as open standards and as open source. This means that these technologies can also be adopted by competing browsers. They will not be exclusive to Google Chrome, or put Google in a better position than competitors.

---

[33] Google Analytics plans to continue to allow customers to use their first-party data to support publisher monetisation within their own sites. Google Analytics does not use data across unaffiliated publishers for publisher monetization, though customers may choose to share or export their analytics data, including through a linked Google Ads account for ads targeting and/or measurement elsewhere.

[34] See "*An updated timeline for Privacy Sandbox milestones*" available here: https://blog.google/products/chrome/updated-timeline-privacy-sandbox-milestones/.

62. *Second*, the Privacy Sandbox proposals are being developed in collaboration with the wider web ecosystem. Chrome has invited all stakeholders in the web ecosystem to participate in the incubation, testing and refinement of these new privacy-preserving web technologies. Developers and stakeholders can join one or more of the W3C forums where privacy-preserving proposals are being shared and refined, and where they can have conversations with industry representatives, browser vendors and others — for example, to advocate for a particular use case or solution. Today, most community discussion is happening in the Improving Web Advertising Business Group, the Privacy Community Group and the Web Platform Incubator Community Group.[35] As proposed solutions move to the early build and test phase, developers are encouraged to experiment and provide feedback. To test how new solutions work in live scenarios for real users and sites, developers can register to participate in Chrome origin trials.

63. *Third*, ad tech players that compete with Google advertising services are also involved in the process (including Microsoft). For example, Criteo has suggested an alternative to Turtledove (called Sparrow), building on the core concept of the Turtledove technology.[36] Additional proposals were published by several other ad tech members of the W3C including RTB House (Outcome-based TD), Magnite (PARRROT), and NextRoll (TERN). In January 2021, Chrome published a new evolution of Turtledove called FLEDGE, incorporating many elements from these ad tech proposals. More recently Microsoft has published their evolutionary approach to Turtledove, called PARAKEET (and subsequent MaCAW addition). Chrome continues iterating on these proposals with these and many other members of the browser and ad tech ecosystem in the context of the W3C.[37]

64. *Fourth*, a key aspect of the Privacy Sandbox APIs is that as much data stays locally on a user's browser as possible. Where Privacy Sandbox proposals require the involvement of a server, we are exploring different protections we can put in place to ensure the APIs and data handling policies are operating as intended (including auditing).

---

35 See https://www.w3.org/community/web-adv/; https://privacycg.github.io/; and https://www.w3.org/community/wicg/.

36 See Criteo, "*Why birds may play a key role in the future of advertising",* 30 July 2020, available at: https://www.criteo.com/insights/sparrow-why-birds-may-play-a-key-role-in-the-future-of-advertising/.

37 See for example Google's Dovekey proposal under which Google is talking to other industry players about new technologies that can ensure a healthy ecosystem and preserve core business models, available at: https://github.com/google/rtb-experimental/tree/master/proposals/dovekey.

MANDATING COMMON IDS IS NOT IN LINE WITH THE SUSTAINABLE, PRIVACY SAFE FUTURE OF ONLINE ADVERTISING

65. Once third-party cookies are phased out, we will not build alternate identifiers to track individuals as they browse across the web, nor will we use them in our products.[38]

66. Our choice not to use such identifiers is informed by the fact that we don't believe such identity solutions will meet rising consumer expectations for privacy, nor will they stand up to rapidly evolving regulatory restrictions, and therefore aren't a sustainable long term investment. Instead, our web products will be powered by privacy-preserving APIs which prevent individual tracking while still delivering results for advertisers and publishers.

67. We realise this means other providers may offer a level of user identity for ad tracking across the web that we will not — such as PII graphs based on people's email addresses. Our decision will not affect third-party ad tech providers' use of alternative identifiers outside of Google's ads products. Our announcement will not therefore prevent others from continuing their initiatives of common IDs, holding aside any impact from evolving user privacy sentiment and regulatory restrictions. We understand this may result in some of our customers shifting portions of their business away from Google's ad tech products to those providers that offer a level of user identity for ad tracking across the web that we will not offer.

68. We outline below some further detail on identifiers and how we believe they can lead to the personal identification of user information. We then apply this to the Interim Report's proposal for a common user ID and/or common transaction ID and explain how they could allow individual consumers to be identified or readily re-identified.

<u>Identification of a user's personal information</u>

69. Identification is about distinguishing individuals (or telling them apart) from other individuals and recognising the same individual over time. An identifier may, for example, be an email address, a phone number, a user ID or an IP address. The Interim Report's proposal for a common user ID would involve the creation of a new identifier for individuals.

70. As noted by the CMA Final Report: "*Identifiers do not need to store information or contain meaning, nor do they need to be interpretable by humans, in order to identify*

[38] See "*Charting a course towards a more privacy-first web*" available here: https://blog.google/products/ads-commerce/a-more-privacy-first-web/.

an individual."[39] This is because ad tech industry participants may be able to link identifiers with other information in order to identify or re-identify individuals.

71. As noted by the CMA: "*This process of linking together multiple identifiers across different dimensions to build a single unified profile for individuals is often known in adtech as 'identity resolution'. There are adtech providers that specialise in identity resolution services, attempting to match and connect identifiers into unified customer profiles at scale. They license or provide access to identity graphs to other market participants.*"[40] The ad tech providers referred to here include data brokers and data management platforms.

72. Identifiers can be described as "*strong*" or "*weak*" depending on how useful they are at helping to distinguish or single out an individual from other individuals. The CMA Final Report describes what makes an identifier "*strong*" or "*weak*":

> *"Strong identifiers are a) unique, allowing that individual to be precisely singled out from others; b) persistent, allowing that individual to be recognised across time; and c) available, so that they can be accessed and used. The strength of potential identifiers also depends on context. This context includes, but is not limited to, the extent to which the identifier can be linked to other identifiers."*

73. Almost any event-level information that can be joined between two parties is a vector for either targeted or large-scale attacks on privacy. Some of these risks may be foreseen in advance. Others may be identified on careful examination by technical privacy researchers and other experts in this area. Others might only be apparent in hindsight, after years of accidental privacy leakage (during which time they may be quietly exploited by malicious actors). Joining such information should be minimised unless absolutely necessary.

Common user ID

74. The Interim Report's proposal for a common user ID is likely to create a new "*strong*" identifier:

    a. By definition, a common user ID would be "*unique*" for each user, as it would allow that individual or individual's device to be precisely singled out from others.

    b. It would also be "*persistent*" as the same common user ID would apply to an individual user as they browse the web over time.

---

[39] CMA Final Report, Appendix G, para. 20.

[40] CMA Final Report, Appendix G, para. 42.

c. The common user ID would also be widely "*available*":

   i. It is assumed that the ACCC's proposed common user ID would be included in the bid request sent by a publisher every time a user visits a site that wants to sell ad space through real-time bidding.

   ii. Every eligible DSP participating in the auction will see the bid request for that user.

   iii. Even if the DSP doesn't bid, they will still receive the bid request.

   iv. As found by the CMA, "*bid requests are sent to potentially hundreds of adtech intermediaries and advertisers, particularly for open auctions, where any advertiser can bid for the impression*".[41] And, as found by the UK ICO, "*[t]housands of organisations are processing billions of bid requests…each week*".[42]

d. The common user ID would also be shared in a context that would enable a large amount of information about an individual (including their browsing history) to be linked together. By way of worked example:

   i. Let's say a particular user has a common user ID 1234 and visits Publisher A's website where that user consents to personalised ads.

   ii. As noted above, that user's common user ID 1234 is passed by Publisher A's website to multiple SSPs in the bid request, who each in turn pass it to multiple DSPs, ad networks and trading desks in the bid request, and the winner of the auction then passes it to their winning advertiser and its ad server.

   iii. Each of these recipients will be able to link the fact that the user has visited Publisher A's website to the existing information they hold on that user through the common user ID 1234 (which may have been obtained from other ad/bid requests as that user visits other publisher websites). In this way, the multiple participants in the ad tech ecosystem will be able to obtain a detailed history of that user's cross-site browsing activities which can be used to create a detailed profile of the user.

---

41 CMA Final Report, Appendix G, para. 326(b).

42 UK ICO "*Update report into adtech and real time bidding*", 20 June 2019, p23.

75. Common user IDs can also be easily joined with non-pseudonymous IDs such as an email address. For example, if a user signs into a website (e.g. to sign up for a newsletter on a publisher site or to purchase a product from an advertiser site), that entity will receive the user ID, which will now be permanently linked to that user's email address (and possibly other non-pseudonymous information such as the user's physical address or credit card information).

76. Imposing common user IDs would therefore result in the creation of a new strong identifier that would make it much easier for an individual user's cross-site activity to be tracked. It would also make it easier to pool information about the same individual user together, including information gathered by third-parties, in order to personally identify users through so-called "identity graphs".[43]

77. Common user IDs therefore perpetuate the issues with third-party cookies that are driving changes in public sentiment, platforms and regulations today.

Common transaction ID

78. A common transaction ID (even without a common user ID) shared between multiple parties necessarily allows the pooling of information they each possess without consent. As noted by the Office of the Victorian Information Commissioner, matching the ID with other data held by participants in the supply chain allows for re-identifying consumers.[44] There is no way to prevent such matching. Indeed, the entire purpose of such a shared identifier is to permit joining data available to distinct parties, or collected via distinct channels.

79. The common transaction ID (even without a common user ID) gives rise to similar linking and privacy implications as a common user ID. This is because - even if participants in the ad tech ecosystem have different user IDs - they can use a common transaction ID to determine that their respective user IDs for a single transaction refer to the same user, and then link their datasets in the same way as if they shared a single user ID.

---

43 The CMA Final Report describes an "identity graph" for a user as "a *list of identifiers for that user structured in a graph, where edges (connections) represent a deterministic (used together) or probabilistic connection (share another attribute/identifier, such as timestamp) between the identifiers.*" They give examples of two ways in which identity graphs can be used: (i) "*Identity graphs combine many identifiers and thus make a collection of weak identifiers more persistent and unique.*"; and (ii) *"To build as comprehensive picture as possible of an individual, including across different contexts and devices (ie a cross-device graph).*" See CMA Final Report, Appendix G, para. 39 and footnote 27.

44 See Office of the Victorian Information Commissioner, "Submission in response to the Australian Competition and Consumer Commission's digital advertising services inquiry interim report" available here: https://www.accc.gov.au/system/files/Office%20of%20the%20Victorian%20Information%20Commissioner%20%28Feburary%202021%29.pdf.

80. By way of worked example:

    a. Let's say a particular user is known to the publisher of website 1 by the ID "XYZ", and known to DSP 1 by the ID "abc".

    b. The user is visiting website 1, which requests an ad bid from DSP 1. Based on the common transaction ID 1234, website 1 and DSP 1 can communicate that "XYZ" is the same user as "abc".

    c. Then, the same user visits website 2, where the user has been assigned the ID "def". Website 2 sends an ad bid request to DSP 1. The user still has the ID "abc" on DSP 1 (as that DSP assigns a unique and persistent internal user ID to individual users).

    d. Based on a common transaction ID for that impression, website 2 and DSP 1 can communicate that "abc" is the same user as "def". Now DSP 1 can merge data to conclude that "XYZ", "abc" and "def" are the same user, and can also share these linked identifiers that attach to the same individual with others in the ad tech ecosystem.

81. A common transaction ID would therefore likely enable the creation of a new "*strong*" identifier:

    a. Whilst the common transaction would presumably change for each transaction and not therefore attach to a particular user, as explained in para. [80] above, it would enable other user ID identifiers to be linked to a particular user. It would therefore link "*unique*" and "*persistent*" identifiers for each user, enabling an individual to be precisely singled out from others.

    b. The common transaction ID would also be widely "*available*", as it is assumed the ACCC's proposed common common transaction ID would be included in the bid request sent by a publisher every time a user visits a site that wants to sell ad space through real-time bidding. As with the common user ID, it would therefore be shared with potentially hundreds of ad tech intermediaries.

    c. Such a system could be used to essentially re-create the existing "cookie matching" capabilities used today to facilitate shared user identification in online advertising, and which Google believes can and should be replaced by privacy-preserving methodologies that do not require cross-site user identification.

82. Even if the common transaction ID was not included in a bid request and was associated with information about the impression or the revenue/cost of the

impression (rather than a user ID), it would still give rise to similar linking and privacy implications as a common user ID. This is because the information a common transaction ID is attached with is likely, in a separate forum, attached to a user ID. That is, there would be separate reporting available that ties the impression (and related information, such as the revenue generated by the publisher from the impression or the cost of the impression to the advertiser) to a user ID. Even if the common transaction ID is not directly tied to a user ID, the publisher or advertiser could link the two through the common information on the impression. This would then have the same privacy implications as described above.

83. Even if a transaction ID itself is not shared, but event-level fees charged to one party are available to another, this is sufficient to leak sensitive user information. To illustrate this:

    a. Augmenting revenue information that publishers receive and associate with individual users today for various business reasons with new data on fees charged to advertisers, can reveal much more information about individual users, including sensitive information about the user's activity on other sites or apps.

    b. As an example, consider the case of an advertiser who sets up an advertising campaign targeting installations of their app, and pays their ad network only for such installations. This is the most common model for app installations. (A similar situation could occur with other 'conversions', such as purchases from an online store.)

    c. Suppose the app is commonly used by members of a particular sensitive group, such as an app for weight loss. Since publishers can match impressions to individual users, they may see that this advertiser served an ad impression to 20,000 users on their site, at an average price of $2.37 CPM (or $0.00237). 19,994 of these impressions resulted in no charge to the advertiser, while 6 impressions each resulted in an $8 charge to the advertiser. The publisher can correctly conclude that these 6 users installed the app in question, revealing their likely interest in weight loss. In contrast, the data that is available to the publisher today does not allow them to identify which users (of the 20,000 who saw the ad) were truly interested in weight loss. Fundamentally, simply being shown an ad does not reveal much about a user. At best, one can conclude that an advertiser might think the user is interested in their ad, but given the very low click-through rates and conversion rates, this is not very indicative. In contrast, revealing an actual user action *on a different website or app* (particularly, an action as significant as a conversion / purchase / app installation)

is much more problematic.

d. Once a publisher 'knows' which users took actions on individual advertiser websites, this can affect user experience and their privacy expectations in multiple ways. The examples provided below are not exhaustive; once information is shared with multiple entities, it is very hard to control its use.

    i. A malicious actor may deliberately sell this information about the user to third parties, for commercial or other uses that are then impossible to restrict.

    ii. A publisher may consciously use this information to customise the user's experience on their website. For example, a general interest publication may have content relevant to the sensitive group the user is a member of, and may start highlighting / recommending such content (to the user's surprise, since they have no reason to expect the publisher to know their membership). In the example above, this may include articles about bariatric surgery.

    iii. Even if a publisher has no such conscious intent as in (ii) above, data is commonly stored in data warehouses and used for a variety of purposes, including user targeting. Black-box machine learning models may pick up on such correlations, and use this information in recommendation systems or other features that affect the user experience.

e. Further, as the publisher uses different ad tech services to advertisers, requiring event-level fees charged to an advertiser to be disclosed to publishers (and vice versa) also requires this information to be passed from one ad tech provider to another. This information is a direct signal that the user has interacted with the relevant ad. This information can be joined across ad tech players and recipients to create an audience profile for these users.

84. There are also other issues with a proposal to require event-level fees charged to one party to be disclosed to another. This is technically difficult and it is likely a breach of confidentiality to tell advertisers the sell-side margin for each of thousands of publishers, or publishers the buy-side margin for each of millions of advertisers. This would also require extensive and expensive integration and data sharing between these products to reconcile fees reliably, breaking existing information barriers. It is also not clear how this would work for ad networks with a dynamic margin which allows advertisers to pay only for their goals (clicks or conversions) while paying publishers on impressions.

85. The common user ID and common transaction ID proposals are therefore not in line with the sustainable, privacy-safe future of online advertising.

86. We support measures to increase transparency in a way that does not raise the same privacy concerns. It may be possible to address the broader goals of a common user ID and common transaction ID through industry initiatives to improve visibility across the supply chain,[45] or via techniques that reveal aggregate data, rather than per-impression or event-level data. The ACCC could also build off the work of the UK ISBA, which has formed a taskforce with members from other UK industry bodies, including the Association of Online Publishers and IAB, to work towards an industry standard way of allowing the financial audit of a set of programmatic transactions in a privacy safe way.[46] The right level for such disclosures would need to be assessed against other issues, such as revealing sensitive partner business information.

CONCLUSION

87. We appreciate the ACCC's work so far in evaluating competition in digital advertising services. We welcome the ACCC's willingness to engage on the Privacy Sandbox and the implications for its common user and common transaction ID proposals. We look forward to further discussions with the ACCC on these topics as needed.

---

[45] For example, the IAB industry standards Ads.txt (which can be used to identify who is authorised to sell a publisher's inventory), Sellers.json (which allow buyers to discover and verify sellers and intermediaries); and SupplyChain object (which enables buyers and intermediaries to see all parties who are selling or reselling ad inventory).

[46] See IAB UK, "*Cross-industry Programmatic Taskforce announces mission & objectives*", available here: https://www.iabuk.com/news-article/cross-industry-programmatic-taskforce-announces-mission-objectives.