Appendix L: Potential approaches to improving personal data mobility

Introduction

1. This appendix outlines two future developments that could help better protect privacy whilst preserving some of the efficiencies from current digital advertising, increase competition, and ensure that consumers can benefit to a greater extent from the value of their data. These are:

   - mechanisms for increasing data mobility, which would allow consumers to share the data that platforms hold on them with other platforms; and

   - privacy-enhancing technologies (PET), which would reduce the extent of data collection for digital advertising by shifting a significant proportion of the data processing to the device itself.

2. Digital advertising platforms use data about consumers to accomplish two main goals, which are discussed in more detail in Appendices E and H:

   (a) To target ads to consumers, so that consumers are shown ads that they are likely to be interested in (behavioural targeting); and

   (b) To ensure that ads are displayed in the appropriate context and viewed by the right number and kinds of people, avoiding fraud and invalid traffic (ensuring brand safety and verification), and to evaluate the effectiveness of campaigns, by relating ad exposure to conversions, ideally across different devices (measurement and attribution).

3. These activities (behavioural targeting, verification and measurement) contribute to the efficiency of the digital advertising market. If competitors had access to the data needed for these activities, they may be better able to compete with large platforms in providing services to users and for advertisers.

4. However, currently each of these activities involves the gathering, remote processing, and (sometimes) transfer of large quantities of user data across publishers, advertisers, platforms and intermediaries in the digital advertising supply chain. This data is often personal data within the meaning of GDPR, which raises data protection and privacy issues.[1] In particular, these data (including personal data) are sent to a potentially very large number of third

[1] These concerns discussed in chapter 4 and are set out in the ICO's Update report into adtech and real time bidding, 20 June 2019, and in complaints that Brave has made to data protection authorities in Europe.

parties, which might not align with users' knowledge and expectations about how their data are shared and used.

5. A crucial question is whether and to what extent these activities that contribute to the efficiency of the digital advertising market can be performed in a way which better protects privacy and better facilitates competition by preserving ability of smaller firms to operate effectively in the relevant markets.

6. We consider proposals put to us about products such as Personal Information Management services (PIMS) and Personal Data Stores (PDS), and privacy-enhancing technologies (PET) that we have encountered during our market study.

7. We would welcome views on:

   - whether this form of data mobility has merit in principle, and what if any form of regulatory intervention is required to support it; and
   - the potential costs and benefits of privacy-enhancing approaches, and whether any regulatory intervention is desirable to facilitate their development.

The potential benefits of data mobility

8. The potential benefits of giving consumers greater access to and control of data held on them by commercial organisations, including the major digital platforms, and in particular of providing them with the ability to share this data with trusted third parties, have been widely discussed.[2,3]

9. The Report of the Digital Competition Expert Panel (**the Furman report**)[4] recommended that its proposed Digital Market Unit should pursue personal data mobility and systems with open standards where these will deliver greater competition and innovation. It said that that personal data mobility would 'give consumers greater control of their personal data so they can choose for it to be moved or shared between the digital platform currently holding it and alternative new services. By making this easy, consumers could, for example, move across to a new social network without losing what

[2] See in particular Chapter 2 of the Government Green Paper Modernising Consumer Markets, where the rationale for the Smart Data Review is set out in respect of regulated markets and also the role of data portability/mobility in digital markets.
[3] For a comprehensive of data mobility and portability initiatives together with their opportunities and risks see Data Mobility, Ctrl Shift and DCMS, 2018.
[4] Report of the Digital Competition Expert Panel.

they have built up on a platform, manage through a single service what personal data they hold and share, or try out an innovative digital service that uses their information in a new way. Open Banking has shown the potential for data mobility to provide new opportunities to compete and innovate in this way.'[5]

10. There have been some attempts to implement versions of data mobility in the markets within our scope, notably through an initiative called the Data Transfer Project,[6] a joint exercise between Google, Facebook, Microsoft, Apple and Twitter to allow individuals to move their data between online service providers whenever they want.

11. Data mobility could enhance competition by facilitating switching and multi-homing by consumers.

12. It can also be used to support PIMS and PDS, which could potentially aim to do some or all the following activities:

- Help consumers to keep track of controllers of their data, what data each controller has, and manage consents for their data, preventing consent fatigue (centralised consent management).
- Allow users to sign into participating website and apps, in a way which enhances the users' privacy (private authentication services).[7]
- Help consumers to exercise their rights under GDPR and relevant data protection legislation, such as making access requests, data portability requests, requests to restrict processing, and rectification and erasure requests (delegated exercise of GPDR rights).
- Help consumers to put their data portability right to effective use, by transferring data from one service to another, without requiring consumers to download from one service, store this data, and upload it to another service, which may be impractical for consumers without high-speed internet and memory space (direct transfer of data between services).
- Provide a central location where a consumer's data from multiple platforms and online services is backed up and stored securely (secure data store).
- Facilitate micropayments from publishers and platforms to consumers, allowing consumers to benefit to a greater extent from the value of their

[5] Unlocking Digital Competition, Report of the Digital Competition Expert Panel, p9.
[6] Data Transfer Project.
[7] For an example of this function, see Sign in with Apple.

data (facilitating micropayments for data). This is probably the most challenging function to provide, as it would require the PIMS or PDS provider to be able to exclude firms that do not pay from accessing the relevant data, and in many situations data is non-excludable.

13. In the next sections, we first set out some data mobility remedy approaches that have been put to us, and how, in theory at least, data-sharing remedies could work. We discuss whether giving consumers greater control of their data would:

- address any or all of the competition concerns described elsewhere in this report (effectiveness);
- be practicable with any appropriate ancillary measures (practicability); and
- potentially create fresh concerns, including in respect of data privacy.

14. Finally, we set out some questions, the responses to which will help guide our assessment of remedy approaches.

Data mobility remedy approaches put to us

15. Several respondents to our Statement of Scope suggested that the UK should develop a system that puts users in control of their data, enabling them to share data securely across suppliers in different sectors on an informed basis.

16. Independent Digital News and Media Ltd suggested as a remedy the sharing with publishers of appropriate data held on consumers by the large digital platforms within their walled gardens, while retaining the consumer's privacy, IDNM also noted that this process would be facilitated by a persistent digital ID.[8]

17. Barclays[9] Bank supported the development of such a system and urged the CMA to use this market study to consider whether it would be appropriate to put in place a framework for wider access to data currently held by platforms, for example through the creation of standards or a cross-sectoral framework for data-sharing, drawing on the experiences from Open Banking in the financial services sector.

---

[8] A unique identifier of this kind would enable publishers to determine whether a visitor to their site from one of the digital platforms was an individual they held first party data for. A persistent ID could be established whilst third-party cookies are permitted by browsers, following some cookie ID matching. Cookie matching is discussed in more detail in Appendix E.
[9] Barclays response to our statement of scope, paragraph 4.6.

18. A major UK financial institution pointed out that the large digital platforms collected vast amounts of 'intent data' which it described as a critical component of targeted digital advertising but that this data could only be used on their own platforms rather than other advertising platforms.[10]

19. Digi.me suggested that the development of 'data facilitators' (such as Digi.me) would enable publishers (including Facebook and Google) to ask for data directly from the consumer. Digi.me suggested that its solution could integrate multiple data sources, allowing publishers to provide better services (i.e. better targeted advertisements). This would also provide publishers with assurances that users' consent has been obtained for that use of their data and it would facilitate a more formal value exchange which could involve explicit payment for data[11] or an additional service, convenience or reward.[12]

20. Arete Research[13] took this concept one step further by suggesting the development of a 'Data Briefcase'. Arete's submission stated that having the right to withhold it from the large platforms would allow users to monetise their data which could overcome the current switching costs associated with transferring data.

Data Transfer Project

21. Many online services allow users to download a copy of their data, including Google[14] and Facebook[15]. According to Google, users typically do so because they want a copy of their data for backup or out of curiosity, or to transfer a limited amount of specific data for use in another firm's service.

22. For many cases, it is impractical to download and re-upload large amounts of data, particularly if users face limits in their internet connection speed or storage memory. It would be more convenient if online services could transfer users' data directly to another service when instructed to do so. Similarly, it would be inconvenient for users and burdensome for firms to have to process data transfers if different online services providers supply data in different formats.

23. To address these problems, the Data Transfer Project (DTP) is an open-source initiative designed to facilitate direct transfer of data between multiple online platforms, avoiding the need for bespoke bilateral arrangements

---

[10] Lloyds Banking Group response to our statement of scope, p.4.
[11] For example, see UBDI.
[12] Digi.me response to our statement of scope.
[13] Arete Research response to our statement of scope, p.6.
[14] Google Download Your Data (formerly known as Google Takeout) is a centralised page for Google users to import and export their data in many (but not all) Google services.
[15] Facebook, Accessing and Downloading Your Information, accessed 19/12/2019.

between every pair of services. It was founded in July 2018 with Google, Microsoft, Twitter and Facebook, and Apple joined the initiative a year later.

24. These contributors have stated that they consider data portability and interoperability to be central to innovation and that this initiative will facilitate competition, empowering individuals to try new services and switch between suppliers.

25. The DTP should also lower barriers to entry. The DTP uses existing and widely-adopted standards and formats wherever possible. In principle, new entrants can create an adapter for DTP and enable its users to securely and easily transfer their data from incumbent platforms, provided that it also meets reciprocity (ie a service that imports data using DTP must allow that data to be exported as well) and minimum security requirements.

26. Consistent with the GDPR data portability right, the DTP is intended to facilitate transfer of data which has been supplied by the user (volunteered data) or data which the user consents for collection (observed data). It is not intended to transfer 'inferred' data that companies have created.

27. While the project may improve competition within social media – particularly if the number of participants increases – it is however still nascent both in terms of the numbers of participating platforms, users, and files transferred.

How a data sharing remedy could work

The general principles

28. Inferred data is out of the scope of the Data Transfer Project, but this data may nevertheless be personal data which is important to users, from a data protection perspective, and also to advertisers, platforms, and publishers for targeting, verification and measurement digital advertising. We now examine some more extensive approaches to data sharing remedies which could also extend to inferred data.

29. The ability to provide advertisers access to large numbers of consumers who are likely to be interested in products or services they have to offer, and in particular to do so at a time when those consumers are close to a purchase decision, is one source of the major digital platforms' market power.

30. Google has this ability because it can provide contemporaneous access to consumers whose purchase intent can be inferred from their online, and particularly search, activity. Facebook allows advertisers, especially those aiming to create and build brand awareness, to target consumers who are

likely to be interested in a product/service type or brand on the basis of their demographic attributes, interests or behaviours. Both are able to do so because of the information they hold on individual users, including attributes, browsing, location and device usage data.

31. A data sharing remedy approach would work by, with the consent of consumers, providing Google and Facebook's competitors with the same customer information that Google and Facebook have, including inferred data. It would enable the consumer, if they saw an advantage in doing so[16], to instruct Google, Facebook, an intermediary or any other platform deemed appropriate, to share the information held on them with either the publisher of a site they were visiting or with an intermediary.

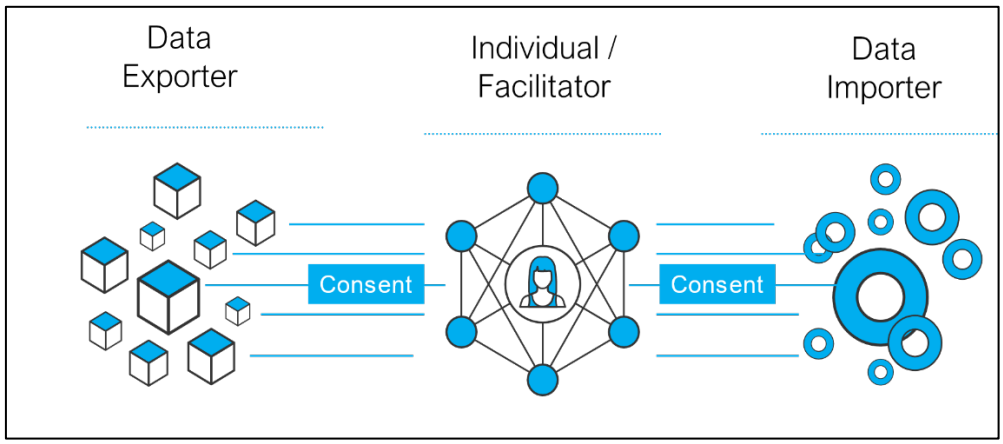
32. In practical terms, the consumer might enrol with an intermediary providing personal information management services. The consumer would, subject to data protection laws and evolving practice on whether representatives can exercise a data subject's right of access and data portability on their behalf,[17] instruct businesses holding data on or about them to share that data with the intermediary free and in real time, having been given that right in law (both the right of access and the right to data portability, where applicable, or an even stronger right that may be supplemented). The consumer would then instruct the intermediary to share some or all of that data with either named parties or parties meeting criteria they had specified, for specified purposes and for a set period of time.

33. Separately, the intermediary could create a consent dashboard for the consumer enabling them to vary or revoke their consents whenever they chose to do so, again subject to data protection laws and practice on delegated exercise of data protection rights and whether a representative can give valid consent on someone's behalf.

34. Ctrl-Shift illustrated such a model, as shown below (Figure L.1), to represent an individual authorizing the multilateral sharing of their data through intermediaries.

[16] Say a reward from other publishers or a desire to filter out less relevant advertising.
[17] Whilst the GDPR does not make specific provision about the ability to appoint someone to act for you when dealing with an organisation that is processing your personal data (except for the ability to appoint a specialist body for the purposes of making complaints), in its guidance to individuals on complaints about media organisations, the ICO states that it is possible to appoint someone to act on your behalf to exercise your rights under data protection law, and that it most circumstances it would expect organisations to allow you to exercise your data protection rights, or raise data protection concerns, through a properly appointed representative.

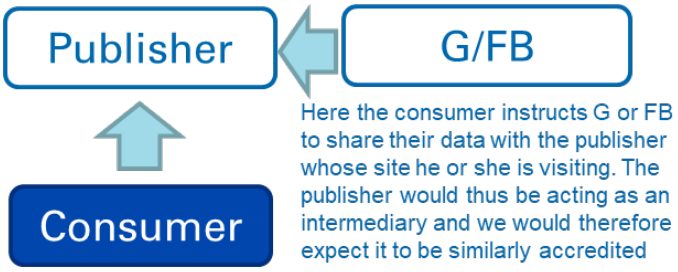
Figure L.1: Ctrl Shift's personal data mobility model



Source: Ctrl Shift

35. We identified three other use cases. In the first, illustrated in Figure L.2, a consumer visiting a publisher's website would authorise the sharing with that publisher of the data points held on them by Google and Facebook at no cost and in real time. The publisher would then be able to better monetize their inventory because it would have much more relevant commercial information about the consumer, possibly enhanced by 'first party' data,[18] that is data the publisher already held relating to him or her. So, for example, impressions arising from a consumer visiting the Times website could be associated with data held on them by the Times and whatever website/app the consumer instructed to share his or her data.

Figure L.2: Data sharing with a publisher



Source: CMA

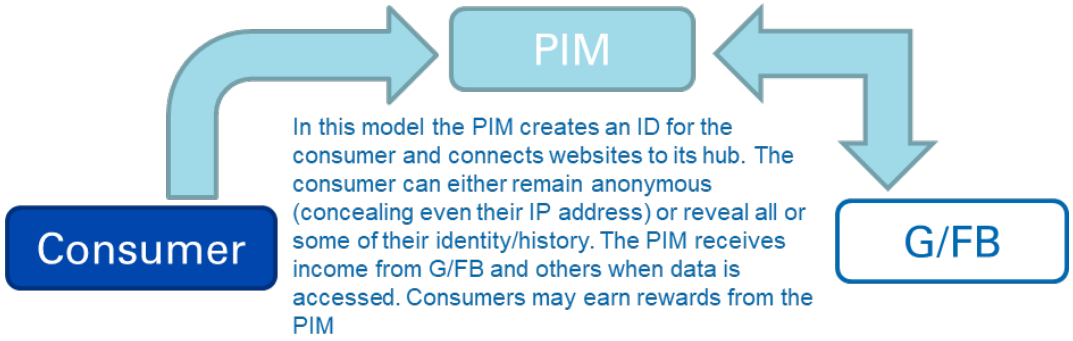
36. In the second case, see Figure L.3, an intermediary would hold the customer's data in a 'Vault' or data 'Store' or data 'Bank' and manage the customer's permissions as regards the access that the customer wished to grant – which platforms, to what data, for what purpose, and over what period. The intermediary could, potentially, and with the customer's consent (or some

[18] This could include information that the consumer had provided when signing up to that publisher's service as well as the web pages visited by the consumer.

other valid legal basis for processing this data, if any exist), combine data held by one or more digital platforms with other data, including information volunteered by the consumer, to provide an accurate and even more comprehensive picture of the consumer than any one platform could.[19]

37. In this model, digital platforms wishing to access the data in the 'vault' would be required to pay/incentivise the consumer/intermediary in order to do so and thus would need to be prevented from doing so by other means, directly. This remedy approach would therefore need to be supported by ancillary measures facilitating the withholding of consumer data that would otherwise be prevented by the platform's terms of use.

Figure L.3: Data sharing with an intermediary



Source: CMA

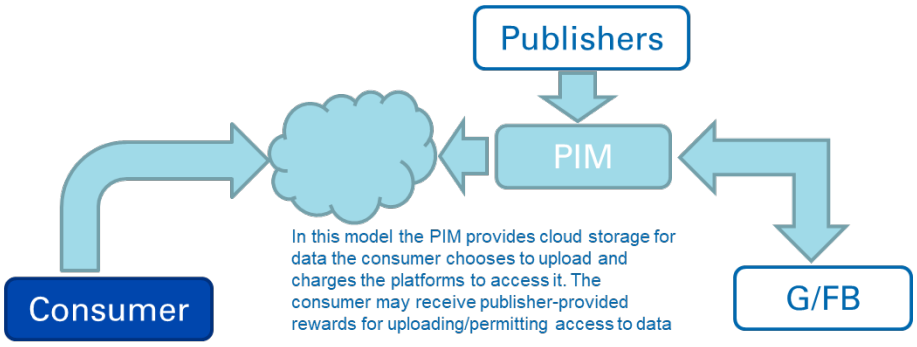
38. Similar models to these are already being used in open banking by around 80 intermediaries with another 300 awaiting approval. Although there is limited use of these structures at this point for other sources of personal data, the very fact of their use in open banking, where confidentiality and security are paramount, suggests that open banking can provide a useful blueprint with similar confidentiality and security requirements.

39. A variant of the second use case, see Figure L.4, would entail the consumer storing their personal data in an area of the Cloud provided by an intermediary (to which that intermediary would not have access rights). The intermediary would then earn revenue by charging the digital platform to access the consumer's data, assuming that the platform is not able to access this or

[19] This additional data could include the customer's online retail activity, digital interactions with Government, say their tax returns, health and physical activity data captured by a wearable device or their bank transactions. An intermediary who was for example an Account Information Service Provider (AISP) or Payment Information Service Provider (PISP) under PSD2 which had access to a consumer's (or SME's) current account transaction data could, if authorised to do so, combine this their online search activity to create some extremely powerful marketing tools.

equivalent data in another cost-effective way. This is similar to Digi.me's model.

Figure L.4: Data-sharing through the cloud



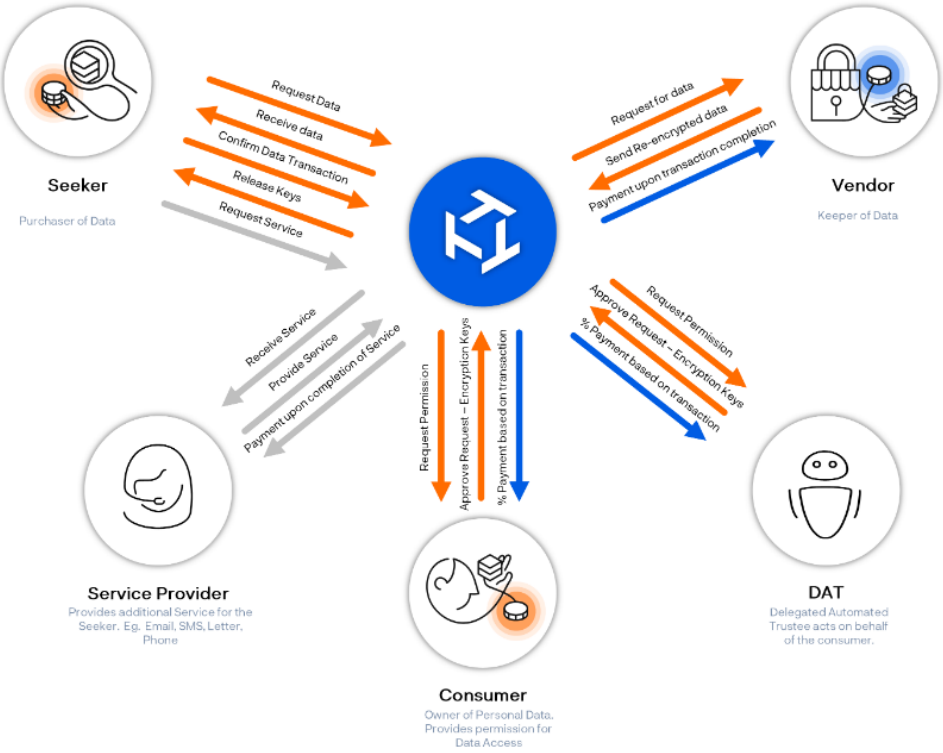
Source: CMA

40. There are several challenges to the practicability of these models, which are discussed in more detail in the section below. One of the challenges to any model which would involve payments from publishers and advertisers to either the PIM or consumer is to ensure that data is excludable, such that advertisers and publishers that do not pay cannot obtain access to the data, either from other sources or by retaining a copy of the data. We discuss two initiatives, Tide and Solid, that attempt to address this challenge by technical means.

41. Tide provides software which allows consumers to lock, until they wish to permit access to it, personal data held on them by businesses. If the holder of their data wishes to sell it to a third party the consumer will first need to 'unlock' it with a key to which they have unique access and if they choose to do so are rewarded in cryptocurrency.[20]

[20] Tide claims to have developed tools which enable securely automated permissioning of access to the consumer's personal data using blockchain technology. This obviates the need for the consumer to consent or authorise each instance of data sharing but automates it on the basis of parameters set by the consumer. It has announced what it described as a 'trustless bot' or a Delegated Automated Trustee on the blockchain.

Figure L.5: Tide's blockchain based model



Source: Tide

42. Solid is an initiative led by Tim Berners-Lee. Solid, which stands for social linked data, is a proposed set of conventions and tools for building decentralized Web applications[21]. The proposal envisages users having Solid PODS (personal online data stores), which are a place to store the data they produce, such as photos, contacts, fitness tracker data, etc. The Solid PODS can be thought of as a personal API, which manages and provides permissions to (Solid compatible) apps to access or link to this data. This framework is aimed at avoiding lock-in to any online service, and improving privacy for the users and creating an ecosystem in which developers can create apps without needing to harvest massive amounts of data. By allowing users to store the data they produce, give and revoke access to their data to the apps they choose and use a unique WebID, this proposal has many similarities with some of the other proposals considered in this paper.

43. There are certainly other executions of data sharing in use or contemplation but their objective remains broadly the same, which is to rebalance in favour of consumers (and potentially smaller platforms and services) the advantage that Facebook and Google have over their rivals' advertising platforms which

[21] See how Solid works, getting started with Solid, and GitHub page for Solid.

is conferred by their unique access to commercially valuable data on very large numbers of consumers.

Which platforms would be required to share data?

44. The remedy would not necessarily apply to all platforms. Were a new or existing regulator to have oversight of the data-sharing ecosystem they could decide on the basis of the market power or 'strategic market status (SMS)' of the parties to whom the obligation would fall. Those platforms so designated could be obliged, if instructed by a user, to share data relating to, specified by, for the purposes of and for the period stipulated by that user, with other platforms/websites using common, open source APIs, data formats and security protocols to facilitate interoperability with other relevant parties.

What data would they be obliged to share?

45. We note first that the Furman proposals on data sharing relate to personal data, which are defined under GDPR as relating to an identified or identifiable individual.[22] GDPR does not confer the same portability rights to inferred or derived personal data as it does to personal data provided by the consumer[23] but it is clear from the analysis set out in this report that inferred or derived data is an important factor contributing to the market power or SMS of the major platforms. Consequently, if the data sharing requirements of GDPR do not extend to derived or inferred information it may not be adequate to address our concerns.

46. Secondly, the data it would be necessary to share in order to address the competition concerns we have identified are likely to differ as between Google and Facebook.

47. Google Search allows advertisers to serve relevant advertising messages to individuals when and in some cases where it can be inferred from the search terms they have submitted, the websites they have visited and in some cases their location[24] that they are considering making a purchase decision.

48. To take a simple example, if an individual had been repeatedly visiting websites offering information on hotels and car hire in San Francisco, had purchased electrical appliance adaptors for US use and had ordered US$ from an online currency exchange it could reasonably be inferred that the individual was embarking imminently on a trip to California. Such a highly

---

[22] See ICO on what is personal data.
[23] See ICO on right to data portability.
[24] We note that inferred information might also be personal data if it can be associated with an identifiable individual.

qualified prospect would be much more valuable to an advertiser than an individual who simply matched the demographic profile of 'visitors to California.'

49. In the case of Google, therefore, its historical data (what an individual searched for two years ago), is not especially relevant to advertisers as it may not reflect the consumer's current requirements or intentions. A remedy requiring Google to share historical data would not, therefore, address the root cause of our concerns: data would need to be shared in real time.

50. The value to advertisers of data collected or inferred by Facebook is not so obviously derived from its ability to signal that a user is actively contemplating a particular purchase and may thus be of more value if the advertiser is, for example, trying to build a brand rather than drive sales. This was certainly the overall impression that we received from the evidence of advertisers and agencies. In both cases though it seems likely that it will be the information that the platforms make available to advertisers that, if shared, would be most likely to address our concerns.

51. However, as the information is inferred rather than volunteered or observed it has been created by the platform using its expertise and technology. Obliging platforms to share such data, it could be argued, might be a disincentive to innovate and not therefore in the long-term interests of competition. On the other hand, relinquishing unique access to this data may actually stimulate them to compete harder by innovating in the use to which they put this data or how they analyse or present it to advertisers.

Effectiveness of data mobility remedies

52. If it is concluded that the market power of Google and Facebook in digital advertising derives in part at least from customer information that they enjoy exclusive access to as a result of their market power in user-facing markets then, absent issues of practicability or any circumvention risks, allowing their competitors to also access this data should, in theory at least, address our competition concerns.[25] On that basis, this remedy could, hypothetically, be effective but, before recommending it we would need to be confident that not only could it be applied in practice but that it would be proportionate and not give rise to new concerns.

53. We therefore next consider the practicability of the remedy.

---

[25] Restricting or forbidding their use of that data in selling advertising inventory would of course have the same effect.

Practicability of data mobility remedies

Technical challenges

54. Our experience implementing the open banking remedies suggests that, despite the differences in the size and scope of data concerned, in principle, the data sharing measures we have described here are feasible technically and that the technology required to support them is familiar, commonplace and reliable. It is more likely that the practicability of remedies based on information sharing will hinge on their commercial viability arising from consumers' incentive to adopt them rather than their technical feasibility. That said, to work reliably such remedies may require a lot of investment in technology, including in the ancillary measures needed to support them.[26]

Commercial viability

55. None of the existing personal information management (PIM) businesses we have found appears to be thriving.

56. However, even if sharing of consumer data by the major platforms was mandatory and open standards for sharing and transfer of common data types were developed and adopted (as the Data Transfer Project is attempting to achieve), would advertisers be willing to pay PIM providers for access to participating consumers? And would consumers be willing to enrol in such services? It is likely that PIMS business models will be multi-sided platforms experiencing cross-side network effects, and would need to secure sufficient participation or support from each side (consumers on the one hand, data controllers on the other) in order to have a viable service.

57. The answer to the first question would seem to depend, amongst other things, on the advertiser's other options, their aims and the sector in which they operated.

58. If the major platforms still held and could provide to advertisers the same information on consumers as they do currently, the answer is probably 'no' unless advertisers considered that data held by an intermediary with access to additional information provided a better means of targeting consumers (for example if it had been validated or combined with other information to create a richer dataset). Even then though, advertisers and their agencies would need to weigh up the value of 'enhanced' data with the cost of doing business with what could amount to a large number of intermediaries (which may be

[26] See Ctrl Shift Report Data Mobility.

the case if several PIMS businesses succeed in getting a significant portion consumers to single-home), rather than the one or two major platforms which deliver tens of millions of consumers between them.

59. Despite such a measure being highly intrusive, if the major platforms no longer had access to this data because of additional rules around the process by which they obtain consent for the processing of different sources of data on their users, then the answer is possibly 'yes' though their appetite would probably vary by advertising category, as we discuss below.

60. Persuading consumers to enrol with a PIM may be even more difficult. However frictionless the on-boarding process with a PIM, there would still be some effort involved and an incentive/value exchange would therefore be necessary to overcome this. The desire to 'take control of one's data' might in itself be sufficient for some consumers but sharing data with new and unfamiliar intermediaries might be seen as risky.

61. It is possible that financial incentives might be available to persuade consumers to sign up. The cost per thousand of reaching customers through mass media is relatively low, reflecting the low expected value of the average customer thus reached. This being the case, advertisers would not be willing to pay an intermediary such as a PIM provider even enough to fund micropayments for connecting them to the consumers enrolled on their service unless they were prime prospects for a response, say like the potential visitor to California we discussed earlier on.

62. The value to an advertiser of such highly qualified prospects is much greater than those who simply happens to match the advertiser's customer profile because the conversion rate (from reading the message to buying the product or applying for the job) is likely to be high. In these circumstances this value might be sufficient to permit a PIM provider to fund consumer incentives.[27] In other words, advertisers should be willing to pay much more to reach a customer who is interested in what they have to offer and on the point of making a decision than they would to reach a customer who may or may not be in the market for their product, service or job. If they are, then this may allow an intermediary such as a PIM provider to afford incentives to attract consumers.

63. That said, we noted earlier that this would be likely to vary by product sector, with revenue being available to fund incentives in sectors where the value of a

[27] Citizen Me and Datawallet incorporate consumer rewards in their business model in the UK. We found many more examples in the USA, for example Midata, Killi and Sprout.

customer was high, such as holidays, travel and financial services, but not others.

64. Nevertheless, a prospective PIM provider would still face a difficult 'chicken and egg' problem: consumers would be unlikely to sign up unless advertiser-funded incentives were available but advertisers would be unlikely to use a PIM until sufficient customers had joined.

65. Providers could possibly get around this by offering data sharing as one module of a data management or data 'bank' service which might comprise access to multiple online accounts, including financial accounts, retail and even health or activity data, secure storage of 'life documents' and passwords, form filling tools and tools to aid decision-making.

66. Banks might conceivably consider entering this space, say by adding a data management module to personal financial management applications, leveraging their reputation for managing highly secure databases and the very large number of account holders that some have. Equally, other firms might consider entry if they had a brand name which was associated with online security and/or had a large enough pool of customers willing to participate.[28]

Ancillary measures

67. Even were the commercial challenges set out here to be overcome, for a data-sharing ecosystem to function well it would need a number of ancillary measures to support it.[29]

68. As discussed, it must be possible for PIMS providers acting on behalf of consumers to effectively and cost-efficiently prevent access and processing of data managed by the PIMS (or equivalent data) by advertisers, platforms and publishers. This would rely on effective enforcement of relevant data protection legislation. Thought should also be given how the PIMS system would interact with the current practice where consumers' data and consent are collected directly by advertisers, platforms and publishers, and the possibility that consumers may give conflicting instructions to their PIMS provider and other data controllers.

69. PIMS providers would need to be accredited and their accreditation details available to counterparties. This would imply, we assume, a process of risk-

[28] YouGov Direct offers consumers the opportunity to join a scheme that enables them to earn rewards from the use of their data. Consumers join the platform, create a profile and provide consent (or not) for each data point to be used by advertisers for targeting. When an individual's data is used, they receive a small cash payment.
[29] These ancillary measures align with the Market Development Requirements set out in the Ctrl Shift report on its Data Mobility Infrastructure Sandbox, June 2019, p.20.

based scrutiny by a regulator (existing or new) to ensure that organisations entering the ecosystem were fit and proper and that their procedures and systems, including security, were adequate. It would also imply a register or directory where the accreditation credentials of third parties could be inspected, including by consumers and organisations holding consumers' data.

70. It would be essential to ensure that the person authorising the sharing of a consumer's data was who they said they were. It would therefore be necessary to establish an authentication process that was both reliable and easy to use. In open banking an 'app to app' process, entailing the use of biometric identification to enable the customer's bank to authenticate them has been adopted, combining security with a low-friction customer experience. A bank-based ID system[30] is in use in Scandinavia in a wide variety of environments, including for example the submission of income tax returns, where it is necessary to ensure that the person providing data or authorising its disclosure is who they claim to be.

71. While it has not been deemed appropriate or necessary in open banking, and depending on the eventual arrangements, it might also be necessary to contemplate providing participating consumers with a single 'digital ID' that they could use to identify themselves with multiple providers.[31] If this was practicable it would enable publishers to identify visitors arriving at their website say from Google about whom they had first party information.

72. Were the liability model in the GDPR not considered sufficiently comprehensive, a liability model may need to be established so that in the event of fraud or theft, for example, consumers knew who they should go to for redress.

73. It is possible that some or all of this infrastructure may be put in place to support other initiatives such as the FCA's Open Finance initiative or BEIS/DCMS' Smart Data review.

---

[30] In Sweden this is known as 'Bank ID'.

[31] The Independent told us that a persistent ID, which fits with the GDPR and reduces the reliance on cookies altogether would be a positive would make browsing faster for users, reduce the need for layers of ad tech and minimise the constant synchronisation that is needed for all ad tech with their own cookie tables. This would move publishers a step closer toward parity with the large platforms like Google and Facebook who disproportionately benefit from vast collection and stores of data within their walled gardens, and which allow them to competitively dwarf publishers in targeting and trading advertising in the market.

Potential for data mobility remedies to create new concerns

74. In designing measures to address one set of concerns the CMA must be mindful of the risk that a remedy will give rise to new ones. In the case of data-sharing remedies the most obvious risk of this kind relates to privacy and the improper use of customer information through identity theft or fraud.

75. The Competition Law Forum (CLF) drew our attention to this risk in their submission.[32] It said that access to the incumbent's data by competitors is likely to enable them to innovate and improve their services, compete on the merits and reduce the extent of the incumbent's data advantage. However, it told us, it lies in tension with data protection considerations and if anonymisation cannot be properly achieved, mandated data sharing is likely to cause significant privacy harms far beyond those Facebook and Google have already caused.

76. DMG Media made the same point. While noting that data-sharing by the platforms had some promise it told us that it was in 'fundamental tension' with user privacy.[33]

77. Ctrl-Shift, however, drew our attention to their report on the results so far of their Data Mobility Infrastructure Sandbox project. The primary conclusion of the report was the end-to-end process of personal data sharing can be made safe.[34]

78. Clearly, there is a risk that if customer data from multiple sources (bank transaction data, search history, health records) was available together and in the same place then the consequences of a breach of security there could be more serious than if they were dispersed and they would become a high value target for malicious actors.

79. At this stage we consider it is too early to assess whether these concerns can be fully addressed. However, our experience from open banking does suggest that, if there is sufficient regulatory oversight, they might be. First, while it is obviously true that where a lot of sensitive data is stored in the same place the consequences of a breach of security will be more serious, it is also the case that the security arrangements around that store will be commensurately tighter. We would envisage that regulatory oversight of participants in the ecosystem, information security, encryption and communications standards

[34] Data Mobility Infrastructure Sandbox, p5.

would need to be as high as those put in place in open banking in order to mitigate this risk.

80. Second, in addition to ensuring that individuals' data was technically secure, the ancillary measures set out above (provider accreditation, strong customer authentication) would mitigate the risk of malicious actors entering the ecosystem in the first place and, so far as is possible, that the individual authorising data sharing is who they say they are.

81. These measures would not, however, address another issue that, we are aware of from the experience in Australia of implementing the Consumer Data Right legislation.[35] This is that problems over data privacy may arise where an individual consents to share data about themselves but in doing so reveals information about someone else who has not consented to its disclosure.

82. The ACCC considered a data-sharing remedy as part of its inquiry into digital platforms and has discussed the tension between competition objectives and the need to protect the confidentiality of consumers' personal data.[36]

83. Facebook responded to the possibility of a data-sharing remedy in Australia with a report highlighting the risks involved.[37] In particular it noted the possibility of consumers inadvertently disclosing information on friends and family without their consent. It may therefore be more appropriate to pursue solutions involving multiple posting of data to social media websites/multi-homing rather than data-sharing.

84. Also, since a PIMS or PDS business is likely to experience cross-side network effects (consumers value a PIMS or PDS that is supported by many publishers and advertisers, and publishers and advertisers is more likely to support a PIMS or PDS with a large number of single-homing consumers), it is quite likely that one or two firms will emerge as 'winners' and take most of the market, perhaps after an initial period of intense competition between many rival services, similar to the evolution of other markets in which digital platforms are active. Indeed, there is a strong possibility that existing incumbent online platforms are or will be strong competitors in new or existing

---

[35] Department of the Treasury (Australia), Consumer Data Right legislation.
[36] 'The ACCC will revisit the applicability of the Consumer Data Right to digital platforms in the future. The ACCC considers that data portability is unlikely to have a significant effect on barriers to entry and expansion in certain digital platform markets in the short term. If data portability or interoperability were identified to be beneficial in addressing the issues of market power and competitive entry or switching, the ACCC could recommend this to government, as part of the role envisaged under Recommendation 4.
However, the ACCC recognises that aside from addressing issues of market power, portability of data held by digital platforms may deliver significant benefits to current and potential future markets including through innovation and the development of new services. The ACCC will consider the benefits associated with digital platform data portability in the ordinary course as it considers sectors to which the Consumer Data Right regime may apply in the future.' ACCC, Digital Platforms Inquiry, Final Report, p.116.
[37] Data Portability and Privacy, Facebook, September 2019.

PIMS and PDS markets, given their large 'installed base' of customers, and in any case, appropriate measures may need to be taken to prevent 'winning' PIMS and PDS providers from becoming a competitive bottleneck in their own right.

85. Finally, we note that even with ancillary technical and regulatory measures in place, it has to be recognised that if consumers are not completely confident that their data will be safe with a PIM provider they are unlikely to use one, or at least will only do so in small numbers and only over a long period. In such circumstances, of course, the remedy would be frustrated and while this risk might be mitigated through, say, a Trustmark, it cannot be ruled out entirely.

Questions about data mobility remedies for consultation

86. To summarise, on balance we think that the remedy would entail a combination of:

- At a minimum, a requirement on large platforms, where consumers have consented to the sharing of their own data, to interoperate with PIMs technically through, say, common and open standard APIs and security protocols and, if applicable, on reasonable commercial terms;

- To be most effective, the remedy would include the adoption of a common standard for the identification of users to enable PIMs and publishers to combine first party data with observed and/or derived data from the major platforms, although there may be privacy concerns associated with creating a common user ID; and

- To create the incentives for users to sign up to PIMs services, it is likely that some additional intervention would be required. One possibility is that the major platforms would be prevented from insisting that consumers consent to providing their data to them as a condition of use. This could take the form of a reciprocal obligation: if the platform wished to collect data on an individual it would be obliged to share it with others as specified by that individual or alternatively if it wished to obtain information from a PIM it would be obliged to supply the PIM with information.

87. We invite responses to the following questions:

L.1 Would the data-sharing remedies we have discussed be effective (including practicable and technically feasible) in addressing our competition concerns? Above all, would consumers adopt them in significant numbers?

L.2 Would they address our concerns comprehensively? Would they perhaps only work in sectors (like financial services and travel) where there was sufficient advertising revenue to attract intermediaries and fund consumer incentives?

L.3 Would the data-sharing remedies we have discussed only, or be more likely to, address the competition concerns we have over Google than Facebook? If so, could variants of the remedy be effective for Facebook or would an entirely different approach work better (say one that facilitated multi-homing)?

L.4 Is there a viable business model for PIM providers? What evidence could we gather to inform our judgement on this? Are there viable data-sharing intermediaries operating profitably in other sectors or overseas jurisdictions? Given the large number of unknowns, would a PIM challenge prize[38] help us determine whether there is a viable business model for PIM providers?

L.5 If such a business model does exist, what other features would it be necessary to provide to create an ecosystem in which PIM providers could exist? We have discussed authentication and security protocols and an accreditation framework but are there other features that it would be necessary to create or adapt? For example, how desirable would it be to create unique and shared identifiers for individuals who wished to share data?

L.6 Are there additional constraints that it would be necessary to impose on SMS digital platforms to make the emergence of viable PIM providers more likely? If the platforms were required to always make a 'Do Not Track' option available and/or set this option as a default with no avoidable loss of service quality would that create an incentive for the platforms to, for example, access customers through the platforms of intermediaries? Is such a requirement practicable or reasonable?

L.7 Respondents to our consultation have acknowledged that obliging the major platforms to share with publishers and third-party providers the consumer data they hold could address some of our competition concerns but would increase the risk to consumer privacy. Do you agree that this risk is real and significant? Are there ways in which the risks to privacy could be mitigated?

---

[38] This could be run on lines similar to Open Banking 4 Good. The challenge would allow selected developers access to a sandbox containing the customer data Google and Facebook of UK users who had consented to have their personal data used in the prize, and award prizes for the best solutions.

L.8 Are there ways in which the major platforms could circumvent the remedies we have described? How could we reduce the prospect of this?

L.9 Would any of the remedies we have discussed here give rise to fresh customer detriment such as higher prices, lower service quality or less innovation?

The potential benefits of privacy-enhancing technologies

88. Up to this point, we have focused our discussion on data mobility remedies on the benefits of data sharing. The size of these benefits will partly depend on the extent to which consumers are willing to share the data that major platforms hold about them with others, and there may be a tension between the aims of data mobility remedies to facilitate competition and a likely motivation of consumers that are interested in a PIMS to regain a sense of control of their personal data and limit the sharing of their personal data.

89. The goals of facilitating competition by encouraging data sharing and helping consumers to maintain effective control over their data need not be in conflict. For a start, consumers may be happier to share their data with and try out new services if they could be more confident in their control of their data with the help of PIMS (eg to effectively exercise their right to erasure or be forgotten). More directly, it may be possible to implement PIMS, PDS, and the broader activities of digital advertising that make use of personal data, in a way that reduces the likelihood of privacy and data protection concerns arising in the first place. We discuss these privacy-enhancing technologies and approaches in the remainder of this appendix.

90. In the current system, data generated by users can be used to track their identities across online and offline activities, serve individually targeted ads, and measure how these ads affect their behaviour. For these purposes, data gathered from users' devices is processed remotely by various actors in the supply chain.

91. Privacy-enhancing technologies (PETs) are a class of technologies that seek to mitigate privacy risks associated with the collection, transfer, and analysis of data, while still allowing for useful results to be obtained from said data. PETs encompass a wide range of approaches, with different degrees of maturity and applicability.

92. A particular type of PETs is client-side PETs. Approaches of this type aim to shift a significant proportion of data processing to the device itself, reducing the amount and granularity of the information that gets transferred away from

it. In this way, the ability of ad tech actors to identify and profile individual users during their online activity is potentially curtailed.

93. The remainder of this appendix focuses on approaches based on client-side PETs. This is because most existing PET proposals in the digital advertising ecosystem are concerned with on-device processing. We welcome views and proposals on the applicability of other types of PETs.

94. Client-side PETs preserve some of the ability for advertisers to provide ads that are targeted to users' interests. The fundamental difference is that a higher proportion of the processing (eg assigning users to segments or matching impressions to ads) happens on the device, rather than remotely.

95. Verification, measurement, and attribution are also potentially achievable in a privacy-enhancing manner, by also shifting the matching between exposure and conversion events to the device, and only sending anonymous and or aggregate attribution data to advertisers, rather than relying on individual-level tracking.

96. These approaches can thus potentially be implemented without compromising the free ad-supported model that underlies a significant proportion of online content creation by publishers.

97. Furthermore, privacy-enhancing approaches could reduce or eliminate the incentives leading to large scale data collection, storage, and resale by Data Management Platforms (DMPs), which can constitute a significant challenge to privacy.

98. Finally, they would not require unique identifiers such as the Mobile Advertising ID, which can facilitate tracking of users by third parties.

99. Whilst the client-side privacy-enhancing technologies and approaches we discuss here may result in significant gains to privacy without sacrificing too much efficiency, based on our current understanding and the evidence reviewed so far, they do not completely remove or fully overcome the trade-offs between privacy, efficiency, and competition that seem to be inherent in digital advertising. Any remedy intervention must consider these three dimensions jointly.

Privacy-enhancing remedies we have considered

General principles

100. The main difference that sets client-side privacy-enhancing approaches apart from the current models is the increased focus on processing data *on-device*.

Raw information about the user and her online interactions, which might include personal data and special category data, is only accessed and processed by the device itself, instead of being transmitted in its raw form to be processed elsewhere.

101. Major tech companies are currently offering software developers the capability to access advanced computing resources on their devices, including CPUs, GPUs, and AI-specific hardware components. Developers can build machine learning models and roll them out within their apps so they run on the device itself, with full or partial access to the device's capabilities.[39]

102. On desktops, services are generally accessed via the browser. Thus, privacy-enhancing technologies would likely be implemented as part of browser software. On mobile, many services are accessed by apps outside of browsers – which would require a more device-wide approach.

103. While raw user data might not leave the device, there are instances in which it might be desirable to make available other types of user-generated data as a user interacts with online services. In such cases, a valid privacy-enhancing approach must still make it impossible for other actors communicating with the browser/device to identify the individual behind these interactions. To this purpose, additional privacy requirements can be imposed – such as k-anonymity (for individual data being broadcast by the browser)[40] or differential privacy (for statistics or models created using individual data).[41]

104. If on-device processing were feasible and became a standard default (either through effective competition on user privacy between device manufacturers, or through mandatory regulations), one of the notable advantages would be to place less burden on consumers. By not requiring users to actively affirm consent on a continuous basis, and reducing their need to familiarise themselves with ways to preserve their privacy online, it might reduce consent fatigue.

Privacy-enhancing behavioural targeting

105. Behavioural targeting aims to serve ads to specific users based on their inferred characteristics and interests. Typically, behavioural targeting exploits

---

[39] See for example Apple's Core ML framework, or Google's Coral.
[40] K-anonymity is a framework that aims to achieve anonymity of individual data by ensuring that an individual's data is indistinguishable from at least (k − 1) others' (see L. Sweeney (2002), k-Anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 05, pp. 557–570).
[41] Differential privacy is a security concept 'which means that, when a statistic is released, it should not give much more information about a particular individual than if that individual had not been included in the dataset' (Royal Society (2019), Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis, p.13).

the availability of large quantities of individual-level data on characteristics (eg demographics, browsing history, search terms) which are paired with conversion events (eg clicks on the ad, purchases and subscriptions).

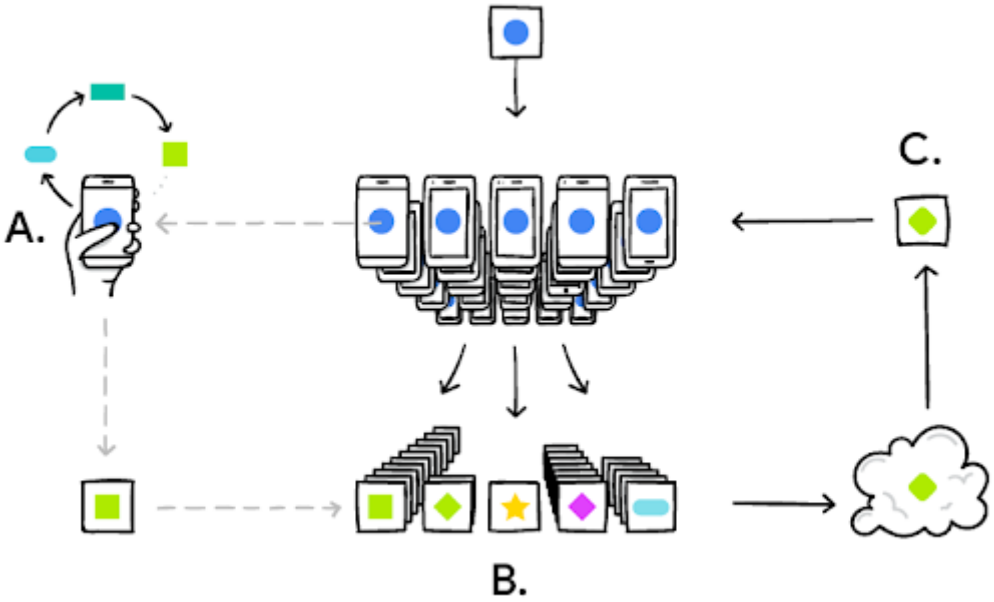
106. Advertisers use machine learning (ML) approaches to train models that predict the likelihood of conversion events based on observed characteristics. These models can in turn be used to predict the likelihood of conversion for a previously unseen user with similar characteristics. Users with higher conversion likelihood in a specific context will be assigned a higher value, and the advertiser will be willing to pay more to show them ads.

107. Typically, ML models for targeting are developed, trained, and refined in a centralised manner; data is gathered from users, processed on remote servers, and then the results are used to decide which ads to serve (or how much to bid for the chance to serve an ad) to a certain newly observed user.

108. Based on our current understanding and the evidence we have reviewed so far, recent advances in cryptography and machine learning may make it possible to achieve comparable behavioural targeting outcomes in a privacy-enhancing way.

109. One possible approach is federated learning (FL). The main intuition behind FL is that the training of ML models can occur in a decentralised manner across multiple devices, instead of a single centralised instance.

Figure L.6: Example federated learning flow



Source: Google

110. Consider the example in Figure L.6. The current ML model (blue circle) is sent to a user's device. The device then uses the data generated locally by the user's behaviours and interactions to improve the model (A) and produces a partial update to the current model. Updates from multiple users are encrypted (B) and securely transmitted to the cloud, where they get decrypted and aggregated into a new model. Throughout the process, the user data on which model training is performed never leaves the device. Furthermore, the model updates that do leave the device are encrypted and anonymised so that they cannot be associated with any individual user.

111. Currently, Google implements a FL approach in multiple ongoing applications – from improving its predictive keyboard, to enhancing mobile vision, to automatic captioning of video content.

112. As part of Chromium's 'Privacy Sandbox', Google has put forward a proposal known as Federated Learning of Cohorts (FLoC) aimed at reducing the privacy footprint of behavioural targeting with the use of FL.[42] This proposed approach, still at an early stage, would operate through any browser that chooses to implement this. The browser would use a federally trained on-device model to assign users to segments ('flocks') with similar browsing habits, which can then be observed by adtech actors and used for behavioural targeting. If these clusters are large enough, the developers claim that privacy by k-anonymity would be ensured.[43]

113. This type of approach uses an on-device model to assign users to segments, and the data used to train this model stays on the device. However, data about users' membership to segments does leave the device and is accessible to websites. While less disclosive than cookies per se, segments still contain potentially personal information about individual interests, including sensitive categories. Furthermore, repeated queries to the browser to access a user's segments can be used for tracking or fingerprinting purposes, in conjunction with other information such as IP addresses. To be truly private, this type of approach would have to be coupled with another layer of privacy-enhancing technology.

114. A different approach has been proposed by Brave, which recently launched a new advertising platform that operates on top of Brave browser. The platform pushes to the user's devices both a catalogue of ads and a targeting model,

[42] See the Chromium developer's GitHub page.
[43] Another related proposal in Google's Sandbox is Private Interest Groups, Including Noise (PIGIN).

which is used to decide which ads from the catalogue are to be shown. In this way, no data about the user's identity or browsing habits leaves the device.[44]

Privacy-enhancing verification, measurement and attribution

115. Users' browsing data plays a critical role in verification, measurement, and attribution tasks for digital advertising. Advertisers assess ad exposure and link it to conversion events to measure the effectiveness of campaigns, using various techniques to reconstruct consumer journeys across websites and devices. Some of these techniques result in privacy-invasive accumulation and transfer of users' personal and behaviour data.

116. Apple (with Webkit's Intelligent Tracking Prevention)[45] and Mozilla (with Firefox's Enhanced Tracking Protection)[46] have been equipping their browsers with default options to curtail common web tracking approaches, such as tracking cookies.[47]

117. Through its Webkit browser engine, Apple has recently put forward a new on-device technology proposal aimed at allowing attribution of ad clicks without the need to track individual users.[48] This approach stores information on ad clicks and conversions on the user's browser. Campaigns and conversion events are denoted by the advertiser using 'small' identifiers (up to 6 bits[49]), which contain too little information to be used as cross-site trackers. The browser keeps track of ad clicks that result in a conversion and sends this data back to the publisher's website – with a random delay between 24 and 48 hours to prevent tracking based on observing conversion times.

118. In its Chromium 'Privacy Sandbox', Google has also proposed a new on-device technology for anonymous attribution.[50] Advertisers would be able to attach a set of metadata to their ads, which would be stored on the user's device when visualised. Similar to the Webkit approach, once the user clicks on an ad, the browser itself would communicate to the publisher's website that a conversion occurred, without the inclusion of any information by the user.

119. An important difference between the two proposals is the size of the identifiers that can be used by advertisers to identify and disambiguate their ad

---

[44] An additional difference in Brave's platform is related to monetisation. Users are rewarded with 70% of the gross ad revenue, in the form of 'Basic Attention Tokens'. These tokens can be transferred by users to publishers and content creators of their choice.
[45] See the Webkit blog.
[46] See the Mozilla blog.
[47] Indeed, Safari goes beyond making this a 'default' option as there is only one setting, which is to 'prevent cross-site tracking'.
[48] See the Webkit blog.
[49] Six bits can effectively encode 64 (2^6) distinct values.
[50] See the Chromium developer GitHub page.

impressions. Webkit suggests a very small 6-bit ID, which effectively allows 64 distinct values to be stored. Google's proposal allows IDs up to 64 bits – which potentially allows for more fine-grained mapping between impressions and conversions and a more significant risk of tracking.

120. Brave has proposed a separate ad confirmation model for its browser. This technology is used to verify that a user has chosen to view an ad, and communicate this confirmation to the advertiser in an encrypted and anonymous way – based on the concept of zero-knowledge proof.[51] The confirmation is then used as a basis to disburse reward tokens for the user.

Practicability of privacy-enhancing approaches

Technical challenges

121. Privacy-enhancing technologies are the focus of a significant and ever-increasing body of academic literature. Progress in this area can open up the possibility of performing an increasing variety of common tasks (such as training a machine learning model) in ways that do not require direct, centralised access to data.[52] As an example, the area of Federated Learning has received increasing attention by researchers and practitioners alike.

122. Rapid future advances in these technologies might have the potential to preserve the efficiency advantages of the current digital advertising ecosystem, while tackling pervasive privacy issues.

123. However, applications of these techniques are still not widely in use at this stage – especially in digital advertising. Many of the more advanced proposals are still at early development stages (sometimes just proofs of concept). It is thus hard to forecast upcoming improvements in PETs from a technical and commercial point of view.

Commercial viability

124. Without sufficient coordination around privacy-enhancing standards, possibly supported by regulation and enforcement, widespread adoption of PETs might be hindered by a 'chicken-and-egg' problem: given the currently existing systems (browsers, devices, websites), publishers and platforms would not

---

[51] Zero-knowledge proof is 'a method by which one party can prove to another party that they know a value x, without conveying any information apart from the fact that the statement is true' (Royal Society (2019), Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis, p.14).

[52] See the previously cited Royal Society (2019) report for additional methods and applications of privacy-enhancing technologies.

take measures to support a browser or device that have implemented PET-based digital advertising solutions (such as developing websites that function correctly on that browser) unless this browser or device had achieved a critical mass of users. Similarly, users would not switch to a browser or device that did not have a critical mass of support by publishers and platforms to ensure that their experience of the internet was not too compromised.

125. Most proposed privacy-enhancing approaches in digital advertising are implemented in browsers. Having these technologies rolled out by default in commonly used browsers and devices would provide a powerful incentive for publishers and advertisers to take the necessary actions to support it. This could potentially be directly mandated as a standard, supported by necessary regulation and enforcement. Otherwise, some ancillary measures could be effective at encouraging adoption (see below).

126. Browser-based implementations would be agnostic with respect to the device, as the same systems can run with minor modifications on both desktops and mobiles. However, a large share of web traffic and advertising on mobile moves through apps rather than browsers directly. This might add a further obstacle to widespread adoption of privacy-enhancing technologies. A potential solution would be to mandate the same standards underlying browsers for other mobile-specific advertising (such as within apps).

127. These technologies may be technically complex and costly to implement. They might also require highly specialised talent to develop and maintain. This may restrict the number of entities that can effectively implement these solutions, and also raises the need for appropriate enforcement of data protection legislation in order to create the correct incentives to do so.

Ancillary measures to encourage adoption

128. There are some ancillary measures that might change the market so that major browsers and apps have sufficient incentive to adopt PETs.

129. An outright ban of third-party cookies, and more generally of cookie-based tracking strategies, would decrease the comparative value to advertisers of more invasive tracking practices, and encourage the switch to ‘cookieless’ approaches.

130. Imposing a prohibition on the copying and trading of user data as part of syndicated data exchanges / data management platforms might also reduce the incentive to persist with invasive tracking practices for the purpose of audience expansion.

131. In a similar way to banning cookies, prohibiting the use of Mobile Advertising IDs (MAID) could achieve similar results in discouraging tracking practices and encouraging adoption within mobile apps.

Creating new concerns

Effect on users, publishers, and advertisers

132. Most of the proposals available so far reduce the amount of user data that is exchanged. While this might alleviate privacy concerns, it might also have efficiency costs – a trade-off that is to some extent ineliminable. An increased 'coarseness' of user data available to publishers and advertisers might make targeting and attribution efforts less precise.

133. Firstly, users might end up being exposed to somewhat less relevant ads. As far as users value ads that correspond to their interest, a less precise targeting would decrease welfare from their point of view.

134. There is a risk that, by reducing advertisers' targeting capabilities, publishers might incur significant revenue losses, jeopardising ad-supported models. As discussed in Appendix E, estimates of the value of behavioural targeted advertising for publishers vary widely, but can be quite substantial.[53,54] In a system where adoption of PETs is widespread, some of market participants' ability to conduct behavioural targeting would be retained, potentially mitigating the impact on publisher revenues relative to an outright prohibition of behaviourally targeted advertising.

135. As with reductions in targeting capability, current privacy-enhancing approaches will tend to impact the granularity and frequency of attribution

---

[53] See for example Johnson et al. (2017), Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry?, Simon Business School Working Paper No. FR 17-19; Marotta et al. (2019), Online Tracking and Publishers' Revenues: An Empirical Analysis, Working paper.

[54] As discussed in Appendix E, current estimates of the value of behaviourally targeted advertising using advertiser bids and prices implicitly compare it to the value of advertising inventory with no associated cookie information or user profile, where both kinds of advertising (behavioural and contextual targeting) are available. However, in a counterfactual where behavioural advertising was prohibited, it is likely that some advertiser spending on behavioural targeted ads would divert to contextual advertising, rather than simply exiting the market. Therefore, publishers' revenues would not decline by the full value of the difference in the estimates of the value of behaviourally targeted advertising relative to non-behavioural targeted advertising.

In addition, although advertisers may be willing to pay higher prices for inventory with richer user data that allows better behavioural targeting, fewer advertisers may compete for or be interested in reaching increasingly narrower consumer segments. The reduction in competition from fewer advertisers interested in each consumer segment might be sufficient to actually result in a net reduction in price of inventory, leading to a reduction in publishers' revenues. This point is made by Leven and Milgrom (2010). Levin, Jonathan and Paul Milgrom. (2010). Online Advertising: Heterogeneity and Conflation in Market Design, American Economic Review: Papers & Proceedings, 100 (2), 603-607

data. This might reduce the efficiency of advertising campaigns, particularly those relying on real-time streams of ad click and conversion data.

Effects on market power of existing dominant platforms

136. PETs potentially present significant implications for market power in digital advertising. Large, incumbent platforms have access to vast amounts of user data, obtained directly via their user-facing services. In a world where user data cannot be exchanged via cookies, and users cannot be openly tracked in their browsing activities, vertically integrated platforms with many users logged in to their services would still be able to exploit granular data in their possession. This could potentially allow them to replicate many of the current targeting and attribution practices, while smaller non-integrated competitors would risk being foreclosed.

137. Similarly, large platforms have access to vast historical data on user behaviour and interactions with devices. Even if they were prevented from vertically sharing data from other services to their advertising arms, they would still have an advantage in the amount of data at their disposal for developing privacy-enhancing models.

138. As previously mentioned, successful application of PETs requires a shift towards on-device computation. Effectiveness and user experience are likely to be enhanced when these software technologies are seamlessly integrated with device hardware, especially in mobile.[55] Integration between software and hardware is managed by operating systems. Vertical platforms own most operating systems, especially in the mobile arena.

139. In a situation where PETs were mandated as a standard, this might create an incentive for large platforms to provide privileged access to a device's compute resources to their own privacy-enhancing technology option, thereby creating barriers to new innovative entrants.

140. Another potential source of advantage for large platforms stems from the technical complexity of privacy-enhancing approaches. The development of such solutions is likely to require highly skilled computer science and engineering talent, with compensation levels that are almost exclusive to large tech firms.

[55] For example, advanced federated learning application for mobile vision are only offered by Google on their own Pixel line of mobile devices – see AI Google website.

Ancillary measures to mitigate concerns

141. To mitigate these potential anti-competitive effects due to large platforms' availability of data, it might be necessary to couple PET-based remedies with some ancillary measures to level the playing field regarding data access for behavioural targeting.

142. One position, similar in spirit to the first group of remedies detailed in this appendix, could be to add data-mobility obligations for large platforms. Potential competitors in PET-based browsers would be able to invite users, if they consent, to request access to data about the user that is stored by incumbent browsers for privacy-enhancing behavioural targeting.

143. An alternative option, more in keeping with the separation remedies outlined in Chapter 6, would rely on preventing vertically-integrated platforms from exploiting users' information across their services for the purpose of behavioural targeting. Platforms would only be able to utilise the information collected during browsing and app navigation.

Questions for consultation

144. To summarise, based on our current understanding and material reviewed thus far, we think a new digital advertising ecosystem based on client-side privacy-enhancing technology would likely entail:

- Browsers incorporating privacy-enhancing targeting technologies by default, where any behavioural targeting would occur on-device only, with no personal data or identifiers leaving the device;
- Browsers incorporating privacy-enhancing attribution technologies by default, where any matching between impressions and clicks or conversions would occur on-device only, with no personal data or identifiers leaving the device;
- Similar technologies underlying mobile-specific targeting and attributions (eg apps);
- Prohibition of targeting and tracking based on cookies or mobile advertising identifiers;
- Prohibition of transfer and trade of individual-level personal or behavioural data; and
- Further measures to tackle downstream competition issues and create a level-playing field for non-vertically integrated agents.

145. We invite responses to the following questions:

L.10 Would the privacy-enhancing technologies we have discussed be practicable and technically feasible?

L.11 Are there ways in which the major platforms could circumvent the remedies we have described? How could we reduce the prospect of this?

L.12 Would any of the remedies we have discussed here give rise to fresh customer detriment such as higher prices, lower service quality or less innovation?

L.13 What is the current picture in terms of available proposals based on client-side PETs in the realm of digital advertising? Do any other approaches exist that have not been considered in this report?

L.14 Current proposals for targeting rely on different approaches than proposals for attribution. Does any privacy-enhancing approach exist that combines targeting and attribution under the same framework?

L.15 What are the minimum requirements for a targeting or attribution technology to be considered privacy-enhancing? Can client-side technologies that disclose or broadcast limited information about the user (eg membership to specific clusters or interest groups) be considered privacy-enhancing?

L.16 Could client-side privacy-enhancing technologies be effective (including practicable and technically feasible) in addressing privacy concerns in digital advertising?

L.17 Would these technologies exacerbate competition concerns, by entrenching the advantage of large vertically integrated platforms at the expense of smaller players in the ecosystem?

L.18 What additional measures would be able to lessen competition concerns arising in an ecosystem where users browse using client-side PETs?

L.19 What are the main obstacles to widespread adoption of client-side PETs for online advertising? How can these obstacles be overcome, to avoid failures similar to previous initiatives like Do Not Track?

L.20 Is there any characteristic of digital advertising on mobile devices that makes client-side privacy-enhancing solutions less effective or practicable?

L.21 Would consumers use devices and/or browsers that have the ability to serve privately targeted ads? Would adoption of these technology follow directly from their implementation in commonly used browsers or devices?

L.22 What is the role of user incentives in the adoption of client-side PETs (eg the sharing of publisher revenue with consumers being exposed to ads on the publisher's website or app)? Are they necessary for widespread adoption? Do they have other implications that might be counterproductive or advantageous?

L.23 What should the role of government regulation be in the adoption and maintenance of privacy-enhancing standards in digital advertising, especially considering the failure of past voluntary initiatives (like Do Not Track) to attain widespread adoption?[56]

L.24 In terms of efficiency, privacy, and competition, how does a digital advertising ecosystem based on client-side PETs compare to one where behavioural advertising is banned outright?

L.25 Is there any non-client-side PET-based approach that can effectively mitigate privacy concerns in digital advertising while preserving some of the efficiencies associated with targeting and attribution?

L.26 Is there any scope for combining data sharing and PET approaches into a unified solution, where for example the functions of a PIMS provider could be performed on-device?

[56] See Do Not Track.