The role of consent following the UK CMA's Privacy Sandbox Commitments

On 11 February 2022, the UK Competition and Markets Authority published a set of legally binding commitments from Google regarding the Privacy Sandbox proposals. The commitments contain significant requirements regarding the interpretation of data protection law by dominant companies.

Concerns about predatory data collection

The CMA had significant concerns that there is scope for self-serving definitions of data protection law to give rise to competitive harm [3.34]. This stems from the much broader data collection practices of large-scale dominant companies:

> "Although rivals can also use first-party data to provide digital advertising services (as the CMA found in the Market Study), their reach and the quality of their data is in many cases much more limited compared [with] that of Google."[1]

This goes against the choices consumers might want to make, which may differ between users:

> "The CMA considers that different web users will have different attitudes and preferences about the collection and processing of their personal data. Some users may prefer not to have their personal data collected and processed by their browser and/or third parties, while others may agree to such data usage in return for seeing more relevant ads, avoiding repeated ads, or other rewards. As such, the CMA sees "the degree of control and optionality enabled by browsers with respect to the collection and processing of Personal Data is likely to be a parameter of competition between browsers
>
> … under the Privacy Sandbox Proposals, the CMA has been informed by Google that Google has not decided whether Chrome web users will have the option of enabling TPCs in Chrome after Google's removal of TPCs. In addition, Chrome web users could have little or no control with respect to whether and how their personal data is used by the browser to provide the functionalities envisaged in the Privacy Sandbox Proposals."[2]

The CMA has also recently issued an Interim Report into its views on how well Mobile Ecosystems are functioning. In that Interim Report it identified as a concern that:

> " privacy, security, and safety online: through design choice or other policies, Apple and Google are often in the position of acting in a quasi-regulatory capacity in relation to users' security, privacy, and online safety. In many cases they opt to make decisions on behalf of consumers. However, it is not always clear if these numerous choices – ranging from restrictions on browser functionality to policies that affect targeted advertising – are in all cases made fully in the interests of consumers. For example, in many cases it seems decisions made on the grounds of protecting users' security and privacy would also serve to give an advantage to first-party apps, or otherwise limit consumer choice."

That is, users should be able to exercise informed choice over the data collection practices employed by the User Agent.

[1] Commitments, 3.34
[2] Commitments, 3.81-3

The document specifically notes the role of user log in systems and the greater precision of tracking by Google in creating a data collection advantage.[3] Indeed, the CMA has noted that there is evidence of the application of misleading data collection practices by large players.[4]

Indeed, there is increasing support for the position that large platforms have heightened duties of care (e.g. the DMA's ban on self-preference), relating to how they are using data and not only on *who* is handling it (the so-called first and third party debate). The log in system is just one aspect of data collection, but it is also that most associated with linking personal data to people's identity. Thus, significant concerns arise in that large and opaque systems in large technology companies could be linked with identity. This could well be a much more significant privacy concern than the use of a responsible identifier by a smaller privacy-by-design system. A User Agent that exploits privacy would be very difficult for users to control, and could enable significant predatory data practices by large online players.

We therefore assume the CMA would apply the same logic to all very large online systems, especially those that collect people's identity-linked data, and that this would include Apple and Meta.

This scale issue poses both data protection and competition issues. As the dominant companies' networks use choice architectures that gather identity-linked personal information (via required email sign up processes such as Gmail for Android) and with forced sign in and intrusive browser policies, their scale and visibility means, there is scope to use them to collect large amounts of information. This has the effect of foreclosing smaller and more responsible, privacy-by-design systems. A system with a responsible data use policy might well struggle to compete with a large-scale but more intrusive system. Indeed there is also a concern that years of privacy predation have conditioned end user behaviour and creates an acceptance of the widely used and more intrusive privacy invading activity than is needed, either by advertisers or end users.: there are more touch points, and incentives to design irresponsibly broad collection systems of identity-linked personal data. Care is needed to design rules so that they do not inadvertently create this tilt towards riskier data processing of such larger systems.

Application of Data Protection Law

To address the concern about predatory data collection, Google's Commitments apply data protection law: Personal Data and Applicable Data Protection Law are defined in line with the UK Data Protection Act.[5] Applying these definitions, the Commitments then require the Google Privacy Sandbox to be developed in line with Development and Implementation Criteria including "the impact on privacy outcomes and compliance with data protection principles as set out in the Applicable Data Protection Legislation."[6]

This is important, as it is the only way that competitive neutrality will be preserved: every compliant business can handle data, if they comply with the law. This has major advantages for users:

- It ensures that data protection follows best practice and allows the regulator to apply new definitions as privacy-by-design requirements develop, rather than setting them in stone.
- It prevents network and business scale from undermining consumer protection, as data protection law can be interpreted to ensure that predatory data protection practices based on

[3] Commitments, 3.35

[4] In its *Online Platforms and Digital Advertising Market Study*, the CMA reviewed the choices available to consumers to control their data. In Appendix X it evaluated potential interventions to allow consumers a choice over whether to receive personalised advertising. In Appendix Z it presented remedy options aimed at reducing or eliminating the competitive advantage that data confers to large platforms.

[5] The UK DPA is identical in effect to the EU General Data Protection Regulation.

[6] Commitments, 8.a

network scale (e.g., overly broad, exploitative consents) cannot undermine consumer protection.

- It prevents trade-offs between competition and consumer protection, instead requiring data protection to be applied on an equivalent basis to all companies handling data, and then enabling competition over these compliant uses.

The Commitments also apply a cross-cutting Purpose which seeks to put users in the driving seat. In addition to concerns about distorting competition, there is expressly a concern about "deny[ing] Chrome web users substantial choice in terms of whether and how their Personal Data is used for the purpose of Targeting or Measurement and delivering advertising to them."[7]

This means that Google has agreed to the principle of a user choice over both whether, and how, data is used.

User choice under data protection law

By referring out to data protection law, the commitments require the application of a number of important user choice concepts from data protection law.

The latest statement on data protection from the UK Information Commissioner's Office (ICO) is *Data Protection and Privacy Expectations for Online Advertising Proposals* (25 November 2021) ("AdTech Opinion"). This document updates a 2019 report on Real Time Bidding. The latest document places particular emphasis on meaningful consumer choice:

- At the outset, the document notes the importance of "a more transparent, user-centric approach that empowers individuals" to "address… the power imbalance that exists between them and key market participants… ***User choice, consent, control and accountability must be meaningful.***"[8]
- A list of requirements includes "transparency … meaningful control and choice over the processing of device information and personal data [and] ensur[ing that] valid consent is obtained where required." This increased emphasis on consent is to help "move away from current methods of online tracking and profiling"to "ensure … demonstrable accountability across the supply chain."[9] This user choice not to be profiled would find expression in the ability to sever links with profiles. Data handlers should enable this control by users through privacy-by-design technologies so that either (i) no individual is ever personally linked to a data profile or (ii) if there is a link, it is no broader than necessary and can be unlinked should concerns arise. This contemplates a role for "safe" (non-linked) identifiers whose use does not raise concerns for users, and distinguishes unsafe and dangerous linking which, if done at all, must always remain under user control (e.g. by comprehensive and clear reset capabilities).
- The document expressly distinguishes *user preference developments* "provid[ing] individuals with a simple-to-use method of expressing a preference and for that to be respected across the web" from those "specifically intended to manage the reduction and eventual removal of Third Party Cookies while continuing to enable targeted advertising."[10] The distinction is said to be between collecting "some form of identifier … (such as an email address) … as opposed to general preference settings or controls at the browser or software level."[11]

[7] Commitments, C.7.(c)
[8] AdTech Opinion, p.5. Emphasis added.
[9] AdTech Opinion, p.12.
[10] AdTech Opinion, p. 25.
[11] AdTech Opinion, p.27 and fn. 80 (it is important to note that the footnote refers to several technologies that allow general preference settings as not doing so).

This leaves open the meaning of a *responsible identifier*. Identifiers have an important role to play in enabling the exercise of data protection rights, e.g. the right to erasure, which requires removing the link to an individual of any previous data collected in order for erasure to work. Issues can however arise in relation to links to identification. A major area requiring further guidance relates to what constitutes a material concern about identity revelation risk. This in turn informs user choice, as user choice will change depending on the quality of privacy-by-design safeguards like pseudonymization. Appendix 1 to this document sets out a framework to address the interaction between meaningful choice and identity revelation risks.[12] MOW analyzed the draft guidance in this document.

The AdTech opinion provides guidance on compliance as follows:[13]

- Organisations must "demonstrate how they mitigate identifiability risk", i.e. manage the risk of re-identification.
- Avoid email-based solutions that risk ineffective pseudonymization.
- Avoid conditioning site access on unconsented personal data collection.

It is important to note that this applies equally to next generation deployments by large technology companies (e.g., if unconsented, personal data collection by Google's Privacy Sandbox as the price of accessing a site is functionally equivalent to a "tracking wall").

The AdTech Opinion specifically calls on W3C processes to ensure compliance with the UK data protection law and notes an important role for the W3C in ensuring that next-generation proposals address the above concerns.[14]

The AdTech Opinion was published the day before the CMA Commitments, following extensive consultation between the CMA and ICO, and the two should be read together.[15]

Competition to support privacy-by-design

The CMA-ICO Joint Statement makes a significant comment about the role of competition in promoting user choice, and thus privacy-by-design:

> "Putting users in control of their personal data is not only important for safeguarding their privacy but can also help mitigate harms such as power asymmetry, which has impacts on the objectives of both competition and data protection. For example, reducing this asymmetry by giving individuals control over the use of their personal data can improve trust and confidence in the digital economy and contribute to a more effective use of personal data while still providing controls and safeguards. From a competition perspective, this can foster healthy competition that benefits users, since it can help reset the balance between digital businesses

[12] It should be noted that it is this handling of personal data that raises privacy concerns, and not the mere creation of categories, or the existence of targeting, as without a link to identity or revelation of information about an individual to others, there is no clear impact on privacy. Without a link to a person, there is no personal data. Much of the lack of clarity in debates on data protection derive from attempts to use data protection law to prevent targeting entirely, as opposed to the preservation of control over data linked to an individual. Bottoming out this debate and the difference between concerns with *personal data* and those with *targeting* would be a helpful W3C workstream.

[13] AdTech Opinion, p.28.
[14] AdTech Opinion, p.30.
[15] The CMA and ICO published a joint statement on cooperation in 19 May 2021: Competition and data protection in digital markets: a joint statement between the CMA and the ICO ("CMA ICO Joint statement")

> and users, putting the onus on the business to do more to engage users and give them greater benefit from their personal data.[16]"

Strictly speaking, competition law and data protection apply slightly different legal standards: competition law bans abuse of a dominant position,[17] whereas data protection law requires lawful, transparent, and fair data processing.[18] There is a strong emphasis on consent in data protection law (e.g., Art 7 GDPR), whereas competition law applies concepts of avoiding anti-consumer market outcomes. However, as noted by the CMA ICO Joint Statement, there are scenarios in which these empowerments, and their aims, overlap. This can take place where both laws intervene to prevent excessive data collection linked to individual data subjects, beyond that desired by consumers, because of a lack of meaningful choice.

If the market failure is in both cases excessive data collection, then both sets of laws align, and the question becomes what responsible data collection looks like.

There is significant guidance on this point from earlier UK CMA work, notably Appendix Y to the CMA's Online Platforms Market Study, which outlines important design features in choice architecture to enable meaningful choice. This report highlighted the need for:

- Accessible choices
- Balanced presentation of choices
- No undue barriers to consumer action over data flows[19]

Responsible data collection practices

The CMA Commitments and Report, the ICO's AdTech Opinion, and the May 2021 CMA/ICO Joint Statement speak with one voice as to concerns about predatory data collection. As the May 2021 CMA ICO statement put it:

> "large ecosystems of interconnected consumer services … to build detailed profiles [whereas] rival publishers … have access to substantially less personal data."

That is, the major concern about predatory data collection is the use of over-broad consents in relation to large platforms. As above, the CMA Commitments then require the application of Purposes and Principles including the promotion of user choice. The AdTech opinion specifically notes the importance of providing choice and transparency over data collection, including an affirmative opt-out.

Where a new technology meets these concerns, it has an important role to play in meeting these concerns, especially where it competes with dominant networks and the large-scale data collection they employ. The documents expressly note the importance of bodies like the W3C incorporating the concerns stated in the documents in their review processes.[20]

The fundamental question therefore concerns the *quality* of consent. It is well known that opaque contract terms could potentially harm consumers in many settings, and there are widespread limits on the scope for *exploitative* contracts that harm consumers. A set of clear and carefully worded, consumer-friendly consent questions have an important role to play and are required by GDPR.

---

[16] CMA ICO joint statement, paras 53-54.
[17] E.g., UK Competition Act 1998, Chapter II.
[18] E.g., EU GDPR, Arts. 5-9.
[19] UK CMA, Online platforms and digital advertising market study, Appendix Y: Choice ARchtietcure and Fairness by design, para 7.
[20] AdTech Opinion, p.30.

Providing specific guidance on which data handling practices give rise to concerns and a set of principles with which businesses can comply will enable competition over high quality privacy-by-design systems. This guidance has an important role to play in distinguishing exploitative, overly-broad consents from legitimate competition over data collection practices. It is doubtful that integration of sensitive privacy decisions into dominant companies' platforms would achieve this, given the inherent conflict of interest from the predatory data practices noted in the reports.

Thus, whilst the utopian ideal might be a single switch to turn on privacy for the entire internet, the reality is that building such a switch into a browser would be ripe for abuse and would likely become simply another pathway for predatory data gathering practices. W3C should seek to promote responsible data handling by defining standards for genuine privacy concerns. Google have agreed to do this approach in reaching their worldwide agreement with the UK CMA which requires the application of general data protection law.

Conclusion

A helpful addition to the W3C work related to privacy would therefore be discussion of what is meant by a "privacy" concern and how a broad range of actors – and not just browser vendors – can adopt best practice, to encourage competition in privacy by design applications and avoiding a single point of failure in the browser or inadvertently restricting competition.

Appendix 1: Identity revelation risks and consumer choice

There is a significant interaction between identity revelation risks and consumer choice. Although not the only privacy-by-design factor, the use of robust pseudonymization has a major role to play in enabling meaningful choices to consumers. Indeed, the difference between robust and non-robust pseudonymization might be *the* privacy by design concern, empowering consumer choice between systems which meaningfully preserve privacy, and those which pose risks to privacy. This interaction between choice and privacy is underexplored and is a priority area for further guidance, e.g. from researchers.[21]

The ICO has also provided guidance on pseudonymization which is relevant to the question of possible consumer harm from identity revelation:

> An organisation applies a pseudonymisation technique that divides personal data into two parts – a dataset that by itself does not identify individuals, and 'additional information' such as a key that enables re-identification. The organisation may refer to the first set as 'anonymous information'. This may indeed be the case in the hands of a third party that has no means reasonably likely to be used to re-identify individuals within that dataset[22]

The following table proposes an approach to identification risk that can provide a foundation to addressing the interaction between consent and identity revelation:

[21] In considering online markets in its influential report, the UK CMA specifically notes a gap in the literature: "Few surveys examine what UK consumers perceive the specific benefits or harms of data processing and targeted advertising to be. Instead, consumer surveys tend to focus on the high-level benefits and harms resulting from all forms of online targeting." (CMA, 2020, Appendix L: para 285. *Summary of research on consumers' attitudes and behaviour*).
[22] ICO Draft Anonymisation Guidance, ch2, p.16

	Personal Data Identifiers				NOT Personal Data
	Individual	**Individual**	**Device**	**Device**	**Statistic**
	Identity-linked ID	**Identifiable ID**	**De-identified ID**	**Random ID**	**Anonymous (no ID)**
Who	The organization receiving this personal data	The organization receiving this personal data	The organization receiving this personal data	The organization receiving this personal data	The organization receiving this personal data
CURRENT State **relative to a specific or particular individual**	**Directly-linked** to an individual (data subject)	**Indirectly-linked** to an individual (data subject) Given ability and often intent to be able to re-identify by the same organization this is NOT de-identified	**NOT Directly-linked** to an individual (data subject) given without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures Appropriate measures include no intent in the future to re-identify	**NOT Directly-linked** to an individual (data subject) given without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures Appropriate measures include no intent in the future to re-identify	An individual data subject is **not or no longer identifiable**.
Safeguards and controls to keep current state of data relative to a specific or particular individual	Appropriate measures to prevent unauthorized access to identity	Appropriate measures to prevent unauthorized access to reidentify identity	Appropriate measures to prevent unauthorized access to reidentify identity	Appropriate measures to prevent unauthorized access to reidentify identity	Appropriate measures to prevent unauthorized access to reidentify identity
Typical Process to transform prior state to current state	Individual's consented or public disclosure of identity under expectation of appropriate safeguards	Individual's consented disclosure of identity under expectation of appropriate safeguards	Pseudonymisation or de-identification that **removes link to identity**, most often by a separate organization that the recipient (aka "held separately")	Algorithmic generation of an **identifier** that **does not depend** on information directly-linked to an individual	Algorithmic generation of an **aggregate report** that **may have been** dependent on information directly-linked to an individual

PRIOR State	**ALWAYS** directly-linked to a specific or particular individual (data subject)	**PREVIOUSLY** directly-linked to a specific or particular individual (data subject) and **CAN** be relinked and **WILL NOT** be relinked (reidentified)	**PREVIOUSLY** directly-linked to a specific or particular individual (data subject) and **WILL NOT** be relinked (reidentified)	**NEVER** directly-linked to a specific or particular individual (data subject) and **WILL NOT** be relinked (reidentified)	**COULD HAVE BEEN** directly-linked to a specific or particular individual (data subject) and **WILL NOT** be relinked (reidentified)
Examples	Name and home address or phone number	Email, passport number or license plate number	Hashed and salted output ID (e.g., key-coded sequence) such as using identifiable ID as an input	Random output ID (unique sequence not used elsewhere), such as using timestamp as an input	Aggregate statistic Should recipient be able to ask multiple iterative questions of the same data set often appropriate process measures (contractual restriction) OR technical measures (added noise and multi-party computation) to protect specific individuals from reidentification (e.g., K-anonymity)
Typical storage mechanism for browsers	Cookie	Cookie	Cookie	Cookie	N/A

ICO		Pseudonymisation therefore refers to techniques that replace, remove or transform information that identifies an individual. For example, replacing one or more identifiers which are easily attributed to individuals (such as names) with a pseudonym (such as a reference number) https://ico.org.uk/media/about-the-ico/consultations/4019579/chapter-3-anonymisation-guidance.pdf	Pseudonymisation means that individuals are not identifiable from the dataset itself, but can be identified by referring to other information held separately. https://ico.org.uk/media/about-the-ico/consultations/2619862/anonymisation-intro-and-first-chapter.pdf	Pseudonymisation means that individuals are not identifiable from the dataset itself, but can be identified by referring to other information held separately. https://ico.org.uk/media/about-the-ico/consultations/2619862/anonymisation-intro-and-first-chapter.pdf	Anonymisation means that individuals are not identifiable and cannot be re-identified by any means reasonably likely to be used (ie, the risk of re-identification is sufficiently remote). https://ico.org.uk/media/about-the-ico/consultations/2619862/anonymisation-intro-and-first-chapter.pdf