# Data: a new direction
# Consultation Response

*Note that this document includes some spelling corrections. made here for clarity of meaning. but not in the submission. We also changed the name of the imagined company in the box below, to avoid confusion with that of an actual company that we coincidentally and accidentally and named in our submission.*

We provide here our response to four of the consultation questions[1]. Our response and views are framed as a small international, but UK headquartered, search engine company that does not engage in any tracking or data harvesting. We believe that surveillance-based business practices are at the heart of many of the problems we see in the modern digital economy and most notably in consumer services. We are part of a coalition of challenger companies to Big Tech which are practicing and succeeding through the commercialisation of web services which do not engage in surveillance-based advertising[2]. Following the four question responses we make some constructive suggestions for reform of privacy policies.

Colin Hayhurst, 19 November 2021
CEO at Mojeek, colin@mojeek.com

## Q1.4.2. To what extent do you agree with the suggested list of activities where the legitimate interests balancing test would not be required?

We strongly disagree with this proposal.

We understand the bad outcomes that have resulted from consent fatigue. Dropping the balancing test will be a success in lowering consent fatigue. However it will be a failure in our opinion, giving companies further power over individuals. The proposal is open to exploitation by companies who will use this as a loophole to further pursue the practices of surveillance capitalism. Unless it is backed up by enforcement of action on companies that do not perform the balancing test, when they should, it will become an open door for even more personal data exploitation and data harvesting.

How does government propose to police companies' projects/products, where the balancing test is deemed unnecessary? The last 15 years or so has more than amply demonstrated that most digital technology companies, and notably Big Tech, cannot be trusted to police themselves. This change will make matters worse. The government is well aware of the damage done to democracy, children, small businesses, publishers and society by these companies. It is developing the Online Safety Bill to tackle some of these challenges. Measures like this which would enable greater data harvesting will undermine those regulatory efforts and increase the level of potential harms.

1 https://www.gov.uk/government/consultations/data-a-new-direction
2 https://blog.mojeek.com/2021/07/time-to-ban-surveillance-based-advertising.html

Companies that act ethically and apply the balancing test properly, will encounter a further barrier to competing with those that do not. The dropping of the balancing test will further empower those that are pursuing aggressive corporate objectives at the expense of both individuals and society. It is a green light to the practices of surveillance capitalism.

The proposed approach can also be cleverly gamed by companies, notably those with large resources and huge datasets that have been and continue to be collected. One can think of multiple examples and scenarios where "loopholes" will be found; that will be the case for any well-meaning list of balancing test exempt activities. This is best explained through a specific imagined, but completely plausible, example shown in the box below.

*What follows is an example of next generation surveillance based digital advertising service that can be developed and deployed without requiring user consent at any stage. And all without the knowledge of users or regulators under the proposed data regime.*

- *A new form of surveillance-based digital advertising ("Cohort Ads") is envisaged by a Big Tech company ("GAME"), which already has large datasets at its disposal.*
- *Since the Cohort Ads will, it is viewed by GAME, provide more relevance for users it is decided that there is a legitimate interest*
- *A new innovative machine learning model is trained using the existing large datasets and new personal data harvested from users, during a pilot phase developed with a subset of the GAME user base.*
- *GAME decide this falls under the exemption: "business innovation purposes aimed at improving services for customers" (61,h), and so no balancing test is required.*
- *This new model works well and is therefore it is decided to roll-out Cohort Ads across all users globally, including the UK.*
- *Cohort Ads, with a model originally trained on personal data, uses inference and without now accessing what they consider to be personal data. They consider it to be privacy-by-design, even if it is in reality a form of profiling.*
- *Before roll-out a second consideration is made by GAME of the balancing test. Personal data is not being used explicitly in the processing of the model, even if it is implicitly embedded in the machine learning model. GAME decide that the balancing test is not required, so Cohort Ads are deployed without the need, in their view, for consent.*
- *Two years later Cohort Ads become the key tool deployed in micro-targetting campaigns deployed in the UK elections.*

## Q1.5.19. Please share your views on what, if any, further legislative changes the government can consider to enhance public scrutiny of automated decision-making and to encourage the types of transparency that demonstrate accountability (e.g. revealing the purposes and training data behind algorithms, as well as looking at their impacts).

Automated decision making is increasingly being made by AI and this will obviously increase and develop. Statistical-based machine learning models are by their nature opaque; the term black-box is extreme but not far from the reality. Moves to develop business purposes, explainability, and encourage transparency will contribute to accountability. However, in our opinion these will never satisfactorily resolve questions about bias and fairness, without firm and decisive action on data practices. Neither will regulation of any speed or flexibility.

Here we are concerned about personal data and inferred personal data. Machine learning models are based on design choices and data. We thus totally support the suggestion to reveal the training data behind algorithms. The only way to have a digital economy in which competition is fair, and in which individuals' rights are respected when personal data is used, is to have concrete transparency by requiring data controllers to:

*Publish, in a privacy policy, exact details of personal data usage:*

1. *what types they collect*
2. *what they retain and for how long*
3. *who they share it with*
4. *what exact data they collect*
5. *what exact data they receive from other parties*
6. *what exact data they share and/or buy*
7. *and to whom and in what country*

This list is not comprehensive. It may sound onerous to provide all this detail but it is not. Such information is typically provided in privacy policies, although it is in practice almost universally vague, incomplete and sometimes evasive. Whilst some companies provide quite a lot of detail, many others provide very little. It is typical to omit large parts of the usage and fine details which have big implications; although not to the untrained eye or expert.

Legislation and/or guidelines which requires/expects companies to make clear, explicit and complete statements along the lines suggested above will lead to healthier trading relationships, a fairer democracy and society, and a future proof digital economy. We make a proposal about this at the end of this document.

## Q1.8.2. In addition to any of the reforms already proposed in 'Reducing barriers to responsible innovation' (or elsewhere in the consultation), what reforms do you think would be helpful to reduce barriers to responsible innovation?

Responsible innovation would arise if the government encouraged companies and startups to apply the principle of data minimization more similar to that suggested by the US FTC[3]. This data minimization proposal goes further than the UK GDPR which for instance allows profiling with the data subject's consent (Article 22, 2(c)).[4] For clarity we quote Rebecca Slaughter of the FTC, verbatim:

> *"That brings me to the next assumption I would like to challenge: the idea that we are stuck with notice and consent as a framework, choosing between opt-in and opt-out for different types of data. Understanding that the collection itself fuels the panoply of problems under the umbrella of "data abuses" helps point to a more effective solution that should be considered: bright-line purpose and use restrictions that minimize the data that can be collected and how it can be deployed. This data minimization approach would turn off the data pump and deprive the surveillance-economy engine the fuel it needs to run.*
>
> *Fundamentally, data minimization should mean that companies collect only the information necessary to provide consumers with the service or product they actually request and use the data they collect only to provide that service or product.* ***Data minimization should be coupled with further use, purpose, sharing, and security requirements to ensure that the information companies can permissibly collect isn't used to build tools or services that imperil people's civil rights, economic opportunities, or personal autonomy.****"*

The UK should in our opinion seriously consider developing a data minimisation policy more like that suggested here by the FTC. There is an opportunity to even take a lead here, just as the UK is doing so in the Online Safety Bill. As the FTC explains[3], notice and consent as a framework has not worked well, as DCMS indeed recognises in seeking to tackle consent fatigue.

Reforms which promote more directly data minimisation in the way described by the FTC would underpin truly responsible innovation. Without it we will see more of the irresponsible innovation pioneered by much of US Big Tech, and which is now being taken further notably by China. It has already and will further inspire irresponsible innovation at smaller companies who have jumped on the *surveillance-economy* bandwagon.

Of course, the UK GDPR already has a data minimisation principle as as the ICO states[5]:

3 https://www.ftc.gov/system/files/documents/public_statements/1597050/commissioner_slaughter_national_advertising_division_10-1-2021_keynote_address.pdf

4 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/946117/20201102_-_GDPR_-__MASTER__Keeling_Schedule__with_changes_highlighted__V3.pdf

5 https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/data-minimisation/

*"So you should identify the minimum amount of personal data you need to fulfil your purpose. You should hold that much information, but no more."*

However the FTC proposes that ***individuals should be able to decide*** whether the data minimisation principles are being applied by a company in a way that suits their ***civil rights, economic opportunities, or personal autonomy***. At present the UK GDPR puts the onus on the companies to decide what data is necessary to provide consumers with the service or product they actually request. Any company using personal data, or inferred data about an individual, to determine how information is displayed is directly acting against personal autonomy.

Companies will argue that eradicating any use of direct or inferred personal data will mean that many algorithmic practices (so called personalisation, recommendations, ranking, targetting etc.) will not work. This is a valid argument; one cannot imagine a Facebook, LinkedIn or Twitter feed without any recommendation algorithm always being the necessary option. Users would be overwhelmed with noise with only a chronological feed. However, ***individuals should be able to decide*** whether the data minimisation practices used by the company are acceptable to them. Those decisions are only possible if full and complete data transparency is practiced. Without it ***personal autonomy*** is lost, as it has been in notably social media and search; a matter which is being tackled in the Online Safety Bill.

The transparency obligations in the GDPR are inadequate because they do not require companies to disclose exactly what data is used, covering only how data is used[6,7]. We advocate for specificity and completeness in our response to Q1.5.19.

6 https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/lawfulness-fairness-and-transparency/#the_principle

7 https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-be-informed/

## Q2.4.4. To what extent do you agree that the requirement for prior consent should be removed for all types of cookies?

We strongly disagree.

Two options are proposed:

> *198. The government is considering two main options for tackling these issues. The first option would permit organisations to use analytics cookies and similar technologies without the user's consent.*
>
> *200. The government also welcomes evidence on the risks and benefits of a second option, which could permit organisations to store information on, or collect information from, a user's device without their consent for other limited purposes.*

If these are the only two options being considered, then we point out that only the first option is preferable. Many companies are already evasive and opaque. This will further encourage them to be so. The second option will inevitably be open to interpretation, and even abuse. It would require self-regulation and/or policing. The former has been shown not work in the international digital economy. The latter will also incur expense for government and business. It will also lead to further distrust in the digital economy by both consumers and businesses.

At the very least prior consent should always be required for all tracking based cookies. We are able to operate an international web search engine without anything other than one cookie, which incorporates strictly necessary data only. Furthermore this data is fully under the users control with total transparency. As it happens it is even possible for users to use Mojeek without this cookie at all.

If prior consent is removed for all analytics-based cookies then full transparency should be required in the privacy policy. Similar to the way we advocate full and clear transparency on personal data practices for a website (see response to Q1.5.19), so we advocate for a comprehensive details on the analytics used. We do not use analytics at Mojeek so lack the expertise to provide a suggested initial list.

However to give one obvious example, websites using Google Analytics or an alternative, should state in their privacy policy whether they are doing so. A service using privacy centric analytics services (such as Fathom or Matomo) would then be much more obvious; allowing users and customers to more easily make informed choices.

Although this might be discoverable by inspection in the website/app code, this is not something that almost all users know about or are capable of doing. Furthermore this information can be displayed in such a way that it is hard to find. Information about analytics used should be discoverable by any user in plain text in the privacy policy.

## Improving the Efficacy of Privacy Policies

In our responses to Q1.4.2 and Q1.5.19 we have suggested that explicit details are provided in privacy policies about personal data and analytics. Analogous data should also be included in privacy policies about cookies and tracking. At Mojeek we do not engage in any tracking, nor use analytics, and only use a small cookie for strictly necessary purposes. So lacking expertise in these areas, we have not made specific suggestions about these. Our specific suggestions in Q1.4.2 on personal data are more considered, though will not be comprehensive.

Clear, common and comprehensive presentation of the details of personal data used, cookies, analytics and tracking would lead to a healthier and more effective digital economy. With this users and businesses would better know, to give one example, how to distinguish between companies providing the benefits of privacy and those data harvesting at scale. Moves to require more transparency on algorithms and business practices will help, but without tackling the source of the surveillance capitalist IP (increasingly AI), these moves will have limited effect. The source of this IP is data, and use of personal data is the underlying source of problems we face in the modern digital economy.

We would suggest that the ICO, and/or Ofcom, are tasked with developing improved guidelines and tools for preparation and presentation of privacy policies. At present privacy policies are developed somewhat ad-hoc by service providers, or by using the services of, or templates from legal service providers. This results in a situation where there is a big variety of privacy policies in terms of appearance, structure and the levels and granularity of detail provided.

More homogeneity of privacy policies would empower users and businesses to make quicker and better decisions about whether they wish to accept the terms of a digital service. The ICO provides useful guidance on the writing of privacy policies already[8]; this could be greatly enhanced for service providers through the provision of privacy policy templates. It may even be practical and useful to develop a web service which acts a privacy policy template generator. This might be a simple Government Digital Service which provides a structured template that would be customised to reflect particular details. The customisation would take as input the outputs of an interactive form to be completed by a service provider, who wishes to create a standardised template which they can adapt for usage.

8 https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/the-right-to-be-informed/how-should-we-draft-our-privacy-information/