

Written evidence submitted by Dr Amy Orben College Research Fellow at the University of Cambridge (OSB0131)

Regarding Clause 101

Dr Amy Orben
University of Cambridge

Dr Amy Orben is a College Research Fellow at the University of Cambridge where she heads the [Digital Mental Health Research Group](#). Her work focuses on the impact of social media use on adolescent mental health and has informed policy both in the UK and internationally.

The need for high quality data to understand online harms

There is currently no adequate provision for researchers to gain access to data from digital technology companies to study their impact on society, meaning that there is little to no good quality evidence about a) what online harms are and b) what their impact is. This lack has been noted, for example, by the UK Chief Medical Officers when they failed to agree on concrete screen time guidelines for children and young people due to lack of high-quality evidence in 2019 (Hawkes, 2019; Murphy et al., 2019): *“An appropriate mechanism for measuring children’s digital engagement needs to be developed, as well as longitudinal studies to assess the temporal relationship between screen time and health and development. Performing meaningful research will rely on provision of anonymised data from technology companies themselves.”*

Digital companies fail to provide such data

Digital companies have been largely unwilling to provide external researchers access to data about what users are experiencing online, making it impossible for good quality and independent scientific studies to be completed on topics such as online harms, mental health, or misinformation. Such companies have routinely cited privacy, IP and commercial issues when challenged about why they have not provided such important data to researchers, or when [shutting down research endeavours in the space](#).

None of these reasons stated comes to light as adequate to deny research access to data with such high public health impact. Facebook, for example, makes some [data publicly available](#) without adding any additional anonymity through their CrowdTangle service. Another site that provides data, Graph API, is only available to private companies but not researchers. This seems to be a strategy on the behalf of tech companies and means they could be selecting data that point the company in a good light, not allowing the public to have access to any data showing evidence of online harms. Certainly, [former executives have made this point](#).

Facebook has until recently touted its efforts to share data with researchers, especially through its creation of the project Social Science One. However, a Social Science One dataset that was professed to have included ‘all Americans’ [only contained Americans with strong political leanings](#), leaving out half the data. Similarly, data Facebook provided in one of its analytics tools used by journalists and researchers, [went missing](#).

As digital companies fail to share data with research, they are blocking evidence that could have gone towards making better policy. Further, they use such data for their own research – for example about how [Instagram affects body image](#) – yet they keep problematic conclusions private, keeping policymakers in the dark. The lack of adequate data sharing procedures also drastically increases the time it takes to conduct and publish the research. This delay can have drastic consequences, especially in the crucial online harms policy space. To design timely policy interventions, research data must be provided with the necessary and current data.

Issue with Clause 101 of Online Safety Bill

To help research access accurate and timely data from digital companies, the Draft Online Safety Bill includes Clause 101 which specifies that OFCOM will prepare a report about researchers' access to information from technology companies. Such a step does **not solve this pressing issue** and will just reiterate what previous reports have found. In the meantime, researchers will continue to fail to provide the necessary evidence to understand online harms, and to develop active policy solutions, due to the lack of data access. However, with the Online Safety Bill and the [new development of UK Data Legislation](#), there is a unique opportunity to build a functioning system for researcher data access set out below.

GDPR can be used to access necessary data

Current UK data law, GDPR UK, presents an unused opportunity to ethically collect digital technology company data without the need to get consent or approval from digital technology companies themselves. Furthermore, this opportunity to provide the basic building blocks for crucial research to inform policy, could be further strengthened in the development of new UK Data Legislation to replace GDPR UK.

[GDPR UK](#) currently gives users, who provide data to digital companies when using their products, actionable rights over personal data that relates to them¹.

*Article 15 (Right of Access): "The data subject shall have the right to obtain from the controller [the digital data company] confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, **access to the personal data.**"*

*Article 20 (Right of Portability): "The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to **transmit those data to another controller** without hindrance from the controller to which the personal data have been provided"*

This means that users of digital services are allowed to **access and receive a copy of their personal data** and other supplementary information stored about them by digital companies. Research

¹ Information Commissioner's Office: With GDPR UK "Individuals have the [right to access](#) and receive a copy of their personal data, and other supplementary information." Furthermore, GDPR UK's "[right to data portability](#) allows individuals to obtain and reuse their personal data (...) to move, copy or transfer personal data easily from one IT environment to another in a safe and secure way, without affecting its usability."

participants should therefore be able to **request such data from digital companies and donate them to research.**

There is a clear opportunity to use the foundations of GDPR UK or future UK Data Legislation to build **world-leading research infrastructure** that uses **digital data donation** from UK users to provide crucial evidence to scientists and policymakers. Such infrastructure would include a service that a) designs and implements ethical, legal and privacy best practices for data donation, b) liaises with digital technology companies to set up effective data-donation linkages, c) fully consents participants, d) sends data access requests on their behalf, e) securely links the data with other sources of research data (e.g., mental health questionnaires or household panel surveys).

The Support Needed

Currently there is no functioning system that allows individuals or researchers to exercise those rights, and power truly transformational research in this space. Subject Access Requests to gain data from technology companies through GDPR are slow, cumbersome, and most often unsuccessful. I personally have been trying to gain access to my own personal Facebook data for most of this year and have had to write many emails, send a copy of my passport and provide Facebook with an email previously not linked to any of my accounts. Digital technology companies are therefore currently not allowing individuals and researchers to use their rights under GDPR to engage in digital data donation.

To ensure this system is functional, scalable, and successful there will need to be a wide range of support (see list below). The Online Safety Bill should not just ask OFCOM to write a report on researcher data access (Clause 101) but issue a **call for action**. Firstly, it should put in writing that individuals have the **right to request** their data from technology companies and donate them to research, and that this should be enabled by the companies without undue barriers or delay. Secondly, it should clarify the types of documentation and hurdles individuals or researchers might face in this process and what is legal for them to face. Thirdly, it should give **powers to OFCOM to support individuals and researchers who want to exercise those rights in the UK**, including fining technology companies if they do not comply.

List of additional support needed:

- a) **Political:** Political support for the data donation process to put pressure on digital technology companies to comply, as well as the further embedding of digital data donation to research in UK data legislation. This can also be useful politically in alleviating pressure to [‘break up big tech’](#).
- b) **Policy:** Including data donation to research in policy reports and ensuring the powers are strengthened – and not eroded – during the design of new UK data legislation to replace GDPR.
- c) **Legal:** Pro-bono legal aid to understand the level of data access GDPR, or a new data legislation, would provide as well as the process individuals must go through to request their data, clarifying the form of data required (such as ‘an intelligible form’ and what recourse individuals have if it is not provided).
- d) **Technical:** Technical support to design a blueprint of the data donation infrastructure system and how it could interface with digital technology companies, to support a large-scale funding application.

- e) **Financial:** Start-up and long-term funding to support the digital donation infrastructure, which should be freely accessible to researchers internationally. [This could take the format of providing users with an opportunity to request their data from multiple sites at the same time to send to different researchers.

Acknowledgement: Thank you to Tara Bhagat for her assistance in preparing this written evidence.

27 September 2021