

Written Evidence submitted by Oxford University Child-Centred AI initiative (OSB71)

Summary of responses

This written evidence is submitted by Dr Jun Zhao (Senior research fellow) and Ms Ge Wang (Final year DPhil student) on behalf of Oxford University's Child-Centred AI initiative¹, led by Dr Zhao, which is part of Human Centred Computing Group² of the Department of Computer Science at Oxford University. We have extensive experience in working and researching together with children and families to understand how data-driven algorithmic systems affect their online experiences and how to design better technologies for them. We have published extensively in premier academic conferences and journals on this topic and contributed to recent policy and regulatory developments in relation to data protection for children and online safeguarding for children, led by the UK CDEI, ICO, DCMS, the Council of Europe and other non-profit organisations (such as 5Rights or Terre de Homes).

We welcome the considerations urged by the Online Safety Bill committee and we would like to add some additional comments below as we thought it's critical to strengthen the requirements for more transparency and more support for users' agency, and a separation between the content vs. system harms. Particularly, we'd like to make the following specific suggestions:

- **Section 10 (6)** - children's risk assessment in their interaction with a user-to-user service: we would like to suggest 1) a more precise wording regarding the impact of the services on children's online digital experiences, and 2) a request for mandating more transparency of these services as part of the duty of the services.
-
- **Section 31(3)** – Children's access assessments: we would like to encourage the government to consider adding an additional mandate to require the service providers to supply evidence to support that their service does not "attract a significant number of users who are children" to strengthen protection of children and transparency of the services.
- **Section 26 (4) Safety duties protecting children:** we would like to encourage the government to further strengthen the requirements for services to be more transparent, more appropriate for children, and put more consideration into children's autonomy online their vulnerability.

Our response is based on our reading of the Government's response published on 17 March 2022³ and the Bill published on 22 March 2022⁴.

¹ <https://oxfordhcc.github.io/oxccai/>

² <http://hcc.cs.ox.ac.uk>

³ https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061446/E02721600_Gov_Resp_to_Online_Safety_Bill_Accessible_v1.0.pdf

⁴ <https://publications.parliament.uk/pa/bills/cbill/58-03/0004/220004.pdf>

Suggestion 1: Section 10. (6) regarding children’s risk assessment in their interaction with a user-to-user service:

In point (b) it discusses the need of “giving separate consideration to children in different age groups, and taking into account (in particular) algorithms used by the service and how easily, quickly and widely content may be disseminated by means of the service”.

We would like to make two suggestions to the above statement.

- **6.(b)** Rewording the description of algorithmic behaviours in this section more accurately, by **extending the description about the service using a more broad and complete statement and adding:** “... and how easily, quickly, and widely content may be disseminated, manipulated or influenced by means of the service.”

- **6.(b)** Adding a stronger recommendation to the duty of the service by requesting actions from service providers: “...taking into account (in particular) algorithms used by the service and how easily, quickly and widely content may be disseminated, manipulated or influenced by means of the service, and adjusting the transparency of the algorithmic behaviours associated with the service”.

Evidence for suggesting a rewording of the user-to-user service behaviours

We believe it is critical to expand the description of the algorithmic behaviours of a user-to-user service because 1) it is more consistent with what is described in Section 25(5) of the Bill in relation to search services; a user-to-user service shares a lot of common algorithmic behaviours as a search service; and 2) it is more accurately describing the nature of these services, which not only influence the dissemination of information but also the consumption of information of the users, by creating the so-called filter bubbles or echo chambers, both of which can lead of harmful results.

Numerous research has shown that children are constantly being exposed to behaviour engineering during their interactions online⁵. Children’s online data is being regularly analysed and inferred to make predictions about their performance, personal preferences, interests, reliability, location, or movements. This is not only a violation of privacy and threat to human autonomy, but also creates an immediate harm for children. This is separate from the *content harm*, such as inappropriate content, which are more explicit and better understood; this about addressing the system-drive/algorithmic-driven harms, where vulnerable users could be repeatedly exposed to content that is harm to them due to the design and behaviours of the algorithms underpinning the services.

Evidence has shown that this regularly takes place on platforms like Facebook, where a user is given ‘interested reading’ based on their digital trace data⁶, or Instagram, where users are

⁵ Giovanna Mascheroni. 2020. Datafied childhoods: Contextualising datafication in everyday life. *Current Sociology* 68,

6 (2020), 798–813. Shoshana Zuboff. 2019. *The age of surveillance capitalism: The fight for a human future at the new frontier of power: Barack Obama’s books of 2019*. Profile book

⁶ Bernhard Rieder. 2017. Scrutinizing an algorithmic technique: The Bayes classifier as interested reading of reality.

Information, Communication & Society 20, 1 (2017), 100–117. Thorson et al, Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information, Communication & Society* 24, 2 (2021), 183–200

nudged towards certain content such as idealised images which could have negative impacts on the body satisfaction of young girls⁷. Similarly, there has been evidence on YouTube conducting inference on users to maximise their engagement on the platform, which could be particularly problematic for the minors⁷. All of these platforms are regularly accessed by children and the evidence indicates that a more stringent scrutiny for these user-to-user services should be mandated in the assessment process by explicitly describing the scope of their operations.

Evidence for suggesting mandating the publication of assessment reports

We believe it is crucial to request more commitment to transparency from the service providers for children.

Our recent systematic review of AI systems for children⁸ has shown a grave gap of how current algorithmic systems fail to consider the key fundamental principles to transparency: children’s personal and sensitive data can be routinely collected and processed for creating supposedly ‘personalised’ services without careful considerations. This type of algorithmic behaviour is being observed not only in systems used for providing personalised learning assistant (which is not within the scope of this Bill) but also for the provision of personalised entertainments, which do fall under the scope of user-to-user services.

By aligning the state-of-the-art systems against the latest regulations in this space, we identified many crucial gaps in this space and lack of support for algorithmic transparency is one of them. We recognise that regulating algorithmic behaviours may not be the focus of this Bill, but we encourage the government to require stronger duties from service providers in this aspect wherever possible so that we can take this opportunity to give children the overdue protections that they need and urge for more responsibilities from the service providers.

Suggestion 2: Section 31. (3) regarding assessing the “child user condition”

It is widely recognised that age assurance is fundamental for building a safer internet for children, although without introducing a walled garden for children and as a result prohibiting them from benefiting from the wider digital world. We welcome the suggestions about assessing children’s access and we would like to make two additional suggestions:

- We would like to encourage the government to consider adding an additional mandate to require the service providers to **supply evidence to support that their service does not “attract a significant number of users who are children”**.
- We would encourage the committee to consider incorporating and referencing the “Age Assurance Standards Bill” introduced by Baroness Kidron, which introduces critical steps towards a standardised age assurance approach in the UK.

⁷ Joint Committee on the Draft Online Safety Bill. 2021. Draft Online Safety Bill. <https://committees.parliament.uk/publications/8206/documents/84092/default/>

⁸ Want et al. Informing Age-Appropriate AI: Examining Principles and Practices of AI for Children. In Proc. of CHI'22. 10.1145/3491102.3502057

First of all, we would like to encourage the government to put in mechanisms to require increased transparency of service providers in terms of their commitment to assessing children's access, by **requiring the publication of their assessment reports in an accessible language and format**. Although **Section 32(7)** requires the service providers to "make and keep a written record, in an easily understandable form, of every children's access assessment", it does not explicitly require the publication of such a record. We believe this is critical to ensure children's rights are respected and more consistent with other transparency requirements reflected in **Sections 11 and 26**.

Furthermore, we would like to encourage the government to consider incorporating and referencing the "Age Assurance Standards Bill"⁹ introduced by Baroness Kidron in a recent session in the Lordships' House in this section, to provide specific guidance and strengthen the assessment described here.

We would like to bring in additional evidence to support this movement led by Baroness Kidron and her colleagues, by referring to our recent research with a Brussel-based NGO¹⁰, which shows that age verification and assurance are critical for providing a safe online environment for children, particularly in the context of preventing online sexual exploitation of children. This research has particularly highlighted the importance of a *standardised* approach for establishing age assurance and verification, to ensure that any platforms that are accessed by UK children will follow a clear set of standards for age assurance, and any age assurance implemented by individual services will be compatible.

Suggestion 3: Section 26. Safety duties protecting children

We welcome the consideration urged by the Bill for the service providers to take up the duty in relation to protecting children, in particular on *functionalities allowing for control over content that is encountered in search results, especially by children*. Indeed, we have identified how lack of controls and coping skills could negatively impact children's online experience. We would like to make two suggestions in relation to this statement:

- Adding a stronger recommendation to the duty of service providers on making the ***design of functionalities, algorithms and other features relating to the search engine, and functionalities allowing for control over content more approachable to families and especially children themselves, so that such functionalities could be actually utilised by stakeholders.***
- Putting more emphasis on service providers' duty on supporting children coping with risks.

Evidence has shown that increasing transparency on how algorithms function e.g. how search results are ranked, how recommendations are made, and how people can cope with that is highly desired by children¹¹ as well as general users. However, this is particularly crucial for children because they are more likely to have difficulty in fully understanding the

⁹ <https://lordslibrary.parliament.uk/age-assurance-minimum-standards-bill-hl/>

¹⁰ <https://www.terredeshommes.nl/en/latest/safety-by-design-to-keep-children-safe-online>

¹¹ 'Don't make assumptions about me!': Understanding Children's Perception of Datafication Online. In Proc of CSCW'22. <https://www.tiffanygewang.com/publication/paper-placeholder-8/paper-placeholder-8.pdf>

operations of these services and what they mean for them¹². This expectation for transparency and age-appropriateness is also highlighted in the UK ICO Age Appropriate Design Code. We would encourage the government to consider this expectation to be linked and referred to here for consistency across legislation as well as for strengthening the protection of the vulnerability of children.

Our second suggestion addresses our concern that the current bill does not seem to have a corresponding requirement in relation to how children could be supported to take control of their own risks online. Such coping strategies could include guiding children to seek help from more knowledgeable others such as their parents, or service platforms being more transparent and flexible on how children can alter their content feeds. Our recent large scale empirical studies have shown that both are strongly desired by parents and children but hardly supported by current platforms or digital products¹³.

Final words

We are excited about the potential impact that the Bill could bring to the current online experiences of the users, particularly children; however, we do think the Bill could include more considerations regarding the requirements for service transparency and support for users' autonomy: the former is a fundamental principle for many current data protection regulations, and the latter is critical to ensure that users are equipped with all the abilities to make an autonomous decision online. Transparency is the cornerstone for better user autonomy; while user autonomy needs to be more explicitly recognised and supported in the current systems as well as regulations. Additionally, we would like to reemphasise the focus on differentiating content harms versus system/algorithmic-driven harms, as the latter can have a much more profound impact on the dissemination and manipulation of content and contact that is harmful to children as well as other users.

In line with these general comments, this evidence submission focuses on pointing out specific sections that we think could be strengthened. We hope these can be useful considerations for the committee. For any further information, please contact jun.zhao@cs.ox.ac.uk.

8 June 2022

¹² Mariya Stoilova, Sonia Livingstone, and Rishita Nandagiri. 2020. Digital by default: Children's capacity to understand and manage online data and privacy. *Media and Communication* (2020). Amelia Acker and Leanne Bowler. 2017. What is your Data Silhouette? Raising teen awareness of their data traces in social media. In *Proceedings of the 8th International Conference on Social Media & Society*. 1–5. Bowler et al. "It lives all around us": Aspects of data literacy in teen's lives. *Proceedings of the association for information science and technology* 54, 1 (2017), 27–35

¹³ Want et al. Protection or Punishment? Relating the Design Space of Parental Control Apps and Perceptions about Them to Support Parenting for Online Safety. *Proc. ACM Human-Computer Interaction, CSCW2*, Article 343 (October 2021), 26 pages. <https://doi.org/10.1145/3476084>