

Dame Melanie Dawes  
Chief Executive  
Ofcom  
Riverside House  
London  
SE1 9HA

6<sup>th</sup> June 2025

Dear Melanie,

### **Risk assessment processes and the Online Safety Act**

The US broadcaster NPR has obtained internal documents setting out that Meta intends to start producing 90 per cent of its risk assessments through AI, including for sensitive and high-risk areas including child safety, AI safety, violence and misinformation.

As a group of civil society organisations, campaigners and experts, we are writing to you to express our deep concern about this retrograde and highly alarming step. We urge you to publicly assert that risk assessments will not normally be considered as 'suitable and sufficient', the standard required by sections 9(2) and 11(2) of the Act, where these have been wholly or predominantly produced through automation.

The illegal and children's risk assessments are foundational components of the UK's Online Safety Act regime, and they have correctly been viewed as a crucial systemic part of how regulation can improve safety outcomes and contribute towards an improved safety-by-design culture in regulated firms.

Meta's push towards automation, which it has confirmed is intended to streamline its decision-making, directly contradicts these ambitions and threaten the safety and wellbeing of children, women, LGBTQ+ and black and minoritised communities. Both current and former Meta employees have stated these changes are being primarily driven by the company's intention to release app updates and features more quickly.

Meta's internal documents not only suggest that risk assessments will now primarily be AI generated but also point to a substantial weakening of the company's existing risk identification and mitigation processes.

According to the company's documents, product teams will no longer be required to submit risk assessments to dedicated risk assessment teams. Under Meta's new streamlined processes, the risk assessment process will instead be undertaken by the relevant product teams and engineers themselves.

The results will be inherently predictable, and we believe them to be entirely inconsistent with the spirit and application of Meta's responsibilities under the Act. As Zvika Kriger, Meta's former Director of Responsible Innovation has put it, 'most product managers and engineers are not privacy experts and that is not the focus of their job. It's not what they are primarily evaluated on and it's not what they are incentivised to prioritise.' Neither are they trust and safety experts. As

we all know, product teams at Meta are instead evaluated on how quickly they launch products, among other metrics.

It seems difficult to determine that Meta's decision to gut its existing risk mitigation and assessment processes is being driven by anything other than a renewed commitment to 'move fast and break things', with Meta underscoring that it envisages risk assessment as a box-ticking exercise at best.

We encourage Ofcom to act quickly to challenge any assumption that platforms can choose to water down their risk assessment processes, whether this is achieved through a shift to automation or as a result of streamlining workflows.

We also encourage you to consider whether these fundamental changes to Meta's risk assessment processes should themselves be risk assessed, as both the illegal content and children's risk assessments refer to governance (s 9(5)(h) and 11(6)(h)) as a matter to be taken into account in risk assessments; and to determine whether the changes being made to the company's risk assessment process may constitute a breach in Meta's compliance with the Online Safety Act that among other steps necessitates the company to generate new and updated illegal and child safety risk assessments.

Substantial changes to Meta's products, whether updates to its recommendation algorithms, new safety features or the addition of new functionalities, deserve rigorous and independent scrutiny. Our children, society and democracy each deserve far better than a box-ticking process churned out by untested and unproven AI, and we need Ofcom to assert itself urgently and decisively in the face of regulated services that seem determined to outpace themselves in a troubling race to the bottom.

This letter is being copied to the Secretary of State for Science, Innovation and Technology, the Home Secretary, and the Chair of the Science, Innovation and Technology Select Committee.

Signed by:

Molly Rose Foundation

Online Safety Act Network

Center for Countering Digital Hate

Antisemitism Policy Trust

Clean Up The Internet

Internet Watch Foundation

5Rights Foundation

End Violence against Women Coalition (EVAW)

NSPCC

Arturo Bejar, former Meta lead for integrity and care

Avaaz

Thomas William Parfett Foundation

NotYourPorn

ParentZone

Glitch

Breck Foundation

Chayn

SOMI (Stichting Onderzoek Markinformatie)

SafeScreens

Professor Claire McGlynn, Durham University

Adele Walton, Bereaved Families for Online Safety

Dr Elinor Carmi, City St George's University

Centre for Protecting Women Online

Dr Elly Hanson