



Consultation response form

Please complete this form in full and return to protectingchildren@ofcom.org.uk.

Consultation title	Consultation: Protecting children from harms online
Organisation name	Meta Platforms Inc.

Consultation title	Consultation: Protecting children from harms online
Organisation name	WhatsApp LLC



Your response

Introduction

This is a **joint submission** by Meta Platforms Inc. and WhatsApp LLC.

Meta Platforms Inc. together with WhatsApp LLC, welcomes the opportunity to participate in this consultation under the new UK Online Safety Act (the “**OSA**” or the “**Act**”) organised by the Office of Communications (**Ofcom**), calling for inputs on the consultation for protecting children from harms online (“**consultation for protection of children**”). We also want to take this opportunity to thank Ofcom for its extensive research as well as the industry and expert outreach for this consultation.

We are committed to helping children (i.e. people under the age of 18) connect and share safely online. As such, we share the OSA’s objectives to ensure children are protected online, and that a higher standard applies for children than for adults, while safeguarding fundamental rights. For clarity, where we refer in this consultation response to “young people”, in general this refers to persons under the age of 18 who are allowed on the platform as per the terms of service (e.g. for the Instagram app and Facebook app it refers to users between 13 and 17).

We have designed our services to be safe for all users, integrating a comprehensive suite of safeguards and controls; and have implemented further safeguards in the best interests of young people, where possible, striking a balance between protecting them and facilitating their connection and development in the digital environment. We constantly work to improve and invest in different measures and tools for age-appropriate and safer experiences in accordance with our Best Interests of the Child Framework, which is based on the UN Convention of the Rights of the Child.

Over the last several years, we have supported the development of the UK Online Safety framework by working with both the UK Government and Parliament, as well as Ofcom. We are convinced that new and innovative regulatory frameworks must strike a complex balance between safety, especially for young people, and people’s rights, such as freedom of speech and privacy. Providers need to take their share of the responsibility to support how to strike this complex balance. We welcome Ofcom’s approach throughout this consultation process to develop this regulatory framework by focusing on how to implement the duties for the protection of children efficiently, as well as the clear objective to develop guidance and Children’s Safety Codes based on the principles of proportionality and collaboration. We welcome Ofcom’s goal to align proposals in this consultation with previous proposals in Ofcom’s “Consultation: Protecting people from illegal harms online” (“**illegal harms consultation**”). To avoid duplication, our response to this consultation refers to feedback we shared in our response to the illegal harms consultation, which included some similar proposed measures.

For clarity, Meta Platforms Inc. (“**Meta**”)’s response is heavily focused on our learnings and expertise stemming from our Facebook and Instagram applications. We note that all of our services are different and what works well for one service may not be the best solution for another service, due to differences in purpose, technical set up, audience etc. There is no one-size-fits-all solution when regulating a dynamic environment such as online safety, and we emphasise that the OSA must be implemented in a flexible and proportionate way, to allow room for differences between services.

At the outset, we want to reiterate and build on one of the most important points from our previous contributions, including our response to the illegal harms consultation, towards shaping an effective and workable online safety framework.

We feel strongly that Ofcom's main priorities and objectives for the guidance and Children's Safety Codes should be **clarity, proportionality, flexibility**, and where possible, **harmonisation with other content regulations and data protection obligations (within the UK and elsewhere in the world)**:

- The guidance and Children's Safety Codes are intended to expand on the broad duties set out in the OSA and provide the practical detail that providers need to comply with those duties. Regulatory certainty will be invaluable for providers, and therefore the guidance and Children's Safety Codes should be as clear as possible in setting out how providers can comply with their obligations.
- Our response indicates where we think the draft consultation materials would benefit from further clarity. However, we emphasise that Ofcom can be clear without being overly prescriptive or departing from the principle of proportionality that underpins the OSA's duties. While the draft consultation materials seek to differentiate between providers on the basis of size and risk, and offer a degree of flexibility in how providers can comply with their duties while remaining within the 'safe harbour' of the measures proposed in the Children's Safety Codes, there are a number of points where we consider the recommended measures to be overly prescriptive and / or disproportionate, or that the options are too limited and not fully reflective of the current state of the industry (as flagged in our response). This impacts the practicality of certain measures, and does not reflect the fact that there are multiple effective ways for providers to meet the OSA's requirements which should enable providers to benefit from a safe harbour.
- As Ofcom is aware, in the European Union, providers are subject to various wide-ranging regulatory requirements regarding the safety of users, including young people, under the EU Digital Services Act (the "DSA") and some providers are subject to the additional requirements applicable to very large online platforms ("VLOPs") such as risk assessments, with ongoing regulatory discussions as to the details of how these requirements are to be met in practice. We appreciate that Ofcom is aware of the importance of harmonisation for systems at scale, as well as regulatory coherence (as set out in the [recent publication](#) by the global network of regulators chaired by Ofcom, which mentions e.g. risk assessments, user complaints, safety measures and age assurance). We have also highlighted additional areas in our detailed response where this harmonisation / regulatory coherence should be made more explicit, particularly with reference to other UK regulatory obligations. For example, we strongly encourage alignment with data protection obligations, particularly in relation to age assurance. Greater flexibility for providers, as discussed above, would also make it easier for providers to harmonise their OSA-related measures with the measures required under the DSA and other legislation.

Second, regarding the overarching topic of **age assurance** in the consultation, as we will develop in more detail in our response below, we would like to point out that understanding user age is a complex, industry-wide challenge that requires thoughtful solutions. We believe that the most effective way of

addressing this problem is for OS or app stores to provide service providers with people's ages, allowing young people to be placed in age-appropriate experiences across all the services they use.

Third, we would like to address the apparent perception in the consultation that **recommender systems** are inherently problematic. In particular, we want to challenge Ofcom's assessment that *"algorithms which provide personalised recommendations to users – are children's main pathway to harm online"*. Algorithms help to sort information and to create better experiences online and are designed to help recommend content that might be interesting, timely or entertaining. Algorithms also help to personalise a user's experience, and help connect a user with their friends, family and interests. Most importantly, we use algorithms to help young people have age-appropriate experiences on our apps.

WhatsApp LLC ("**WhatsApp**") reiterates some overall concerns about specific elements of the Children's Safety Codes proposals. As Ofcom will be aware, in its end-to-end encrypted (E2EE) **private messaging** service, WhatsApp does not have access to the content of conversations on the service. In order to protect our users and society from harm on WhatsApp, while strongly protecting user privacy, we strive to design the environment to prevent users from encountering harm in the first place, and to empower people to keep themselves safe if they do encounter harm.

As we have detailed in previous calls for evidence and engagements with Ofcom, we have a comprehensive approach to addressing the risk of harm which may be encountered by users based on:

- Preventing abuse through product design, product functionality and features;
- Providing a strong suite of user controls which empower users to control their experience and keep themselves safe on the service;
- Designing ways to remove users who violate our policies or use our service to cause harm;
- Collaborating with external experts and organisations on safety measures and educational campaigns; and
- Working with law enforcement, in addition to responding to applicable legal requests.

As an overall point, we welcome Ofcom's design of proportionate and workable regulation that enables a service to demonstrate its compliance with safety duties in a way that does not require it to undermine or materially compromise its security and privacy measures, including where applicable, E2EE. In its previous illegal harms consultation, Ofcom explicitly noted on E2EE in relation to proposed automated content moderation measures that they *"will [...] only apply where it is technically feasible for a service to implement them. We do not consider that it would be technically infeasible to implement a measure merely because to do so would require some changes to be made to the design and/or operation of the service. However, our measures would not apply to services that are technically unable to analyse user-generated content present or disseminated on the service to assess whether it is content of a particular kind, particularly where such changes as would need to be made to enable this would materially compromise the security of the service. For example, we acknowledge that end-to-end encrypted services are currently unable to analyse user-generated content in the ways set out in our proposals"* (Vol. 4 of the illegal harms consultation, para. 14.16). This builds on the principle set out at Sch. 4, para. 2 OSA – to which Ofcom must have regard – that measures set out in a Code of Practice must be technically feasible, and that measures may be technically feasible for some services but not

others. We welcome similar clarification, in this specific consultation and resulting Codes of Practice, that Ofcom’s recommended measures are not intended to result in the undermining of the protection provided by E2EE.

The implications of Ofcom’s assessment that messaging services pose a medium to high risk of priority content harmful to children (“PC”) being present are unclear. We are concerned that this assessment may lead to end-to-end encryption (E2EE) messaging services being compromised for users under 18. For example, if as a result of the risk of PC on an E2EE service, those services are mandated to implement content monitoring or moderation, or to access devices and remove content, the privacy and security of these platforms would be severely compromised. We urge that Ofcom, in collaboration with the ICO, conducts a review and assessment of age assurance measures in order to make recommendations which strike the right balance between protecting users' privacy rights and ensuring they have age-appropriate experiences online and on private messaging services such as WhatsApp.

Meta and WhatsApp are convinced that constructive dialogue is essential to create a workable regulatory framework for all stakeholders while setting up evidence-based good practices and principles. Building on years of experience in tackling these issues through establishing policies, building tools and technologies, in partnership with experts both within and outside our company, we welcome the opportunity to share our expertise and learnings in this consultation response.

Volume 2: Identifying the services children are using

Children’s Access Assessments (Section 4)

Question 0:
Response by Meta
Do you agree with our proposals in relation to children’s access assessments, in particular the aspects below. Please provide evidence to support your view.
Response: We <u>broadly agree</u> with the proposal Ofcom has shared in relation to children's access assessments. However, we have <u>overarching remarks</u> on the overall approach and measures laid out by Ofcom, which we include in response to questions below, especially answers to Questions 1 to 3.
Confidential? - Y/N
Response: N

Question 1:

Response by Meta

Our proposal that service providers should only conclude that children are not normally able to access a service where they are using highly effective age assurance?

Response:

We are supportive of the development of age assurance mechanisms, but note that this remains a complex area. Determining user age is an industry-wide challenge that will require thoughtful solutions and cannot be tackled with a one-size-fits-all approach. Ensuring that young people cannot access services that are not age appropriate is also a broader societal challenge affecting multiple actors, including parents, as well as the tech industry.

Our current approach to age assurance takes into consideration technical accuracy, robustness, reliability and fairness. We support the development of mechanisms which would take into consideration the above principles in order to provide appropriate age experience and within this context limit access for specific user groups.

Confidential? - Y/N

Response: N

Question 2:

Response by Meta

Our proposed approach to the child user condition, including our proposed interpretation of “significant number of users who are children” and the factors that service providers consider in assessing whether the child user condition is met?

Response:

We are broadly in support of a child user condition as part of a reasonable and proportionate approach to age assurance. Overall we believe that any child access assessment is a complex and challenging exercise (e.g. determining what a “significant number” of children using the service would be).

Confidential? - Y/N

Response: N

**Question 3:
Response by Meta**

Our proposed approach to the process for children’s access assessments?

Response:

We refer to our responses to Questions 1 and 2 on stage 1 and stage 2 of the process and further responses in our general submission which touch on age assurance. We also have the following general comments with regards to the overall process:

Circumstances that trigger a new assessment:

Significant change: We refer to our submission below in response to Question 17 (along with previous submissions made in response to Question 7 of the illegal harms consultation) relating to the definition of “significant change”, which we consider to apply in both the context of risk assessments and children’s access assessments. We strongly encourage the examples listed in the guidance to be suggested examples of circumstances which may lead to consideration, or scoping, as to whether such an assessment is necessary, rather than prescriptive criteria which trigger the need for an assessment. Determining whether a change is “significant” is highly context-dependent. Whether a change is “significant” should ultimately depend on the potential impact of that proposed change, on the risk to users, and a child’s ability to access the service. We would therefore advocate that providers should be granted the flexibility to carry out a scoping process in the first instance in order to determine whether a further access assessment is in fact required.

Significant increase in the number of children using the service, and reduced effectiveness of age assurance: similarly to our concerns relating to “significant change”, we also note the importance of ensuring that services are given the flexibility to determine themselves how they will be tracking (i) any increases in the numbers of children using the service (and whether such an increase is “significant” and requires a further access assessment), and (ii) any reduced effectiveness in age assurance. Different services may need to use different methods and metrics to track their user numbers and the effectiveness of age assurance, and the guidance should not be overly prescriptive.

Confidential? - Y/N

Response: N

Volume 3: The causes and impacts of online harm to children
Draft Children’s Register of Risk (Section 7)

**Question 4:
Response by Meta**

Do you have any views on Ofcom's assessment of the causes and impacts of online harms? Please provide evidence to support your answer.

Response:

We refer to our response to Question 1(i) of the illegal harms consultation regarding evidential gaps in Ofcom's assessment, which applies similarly here (as acknowledged by Ofcom in Vol. 3, paras. 7.33 to 7.40).

Assessing Ofcom's evidence:

As previously noted in our response to the illegal harms consultation, without having full access to all the evidence on which Ofcom's analysis relies, we are not in a position to fully engage with or assess that analysis or the conclusions reached by Ofcom.

Evidential gaps in Ofcom's assessment:

We also note that, as acknowledged by Ofcom in Vol. 3 paras 7.33 to 7.40, Ofcom does not have the same quantity or quality of evidence for all areas of its assessment. We recognise that there will inevitably be gaps in the evidence available to Ofcom, given the wide range of services, risks and characteristics to be considered. In some cases, we expect providers may have unique information to hand, particularly as regards the way in which a specific risk manifests on the provider's service, which will factor into the provider's own risk assessment. We anticipate that Ofcom will take this into account where the conclusions reached by a provider's assessment depart from the broader conclusions set out in Ofcom's assessment.

a) Do you think we have missed anything important in our analysis?

Response: Please see our response to Question 4 above.

Confidential? - Y/N

Response: N

**Question 5:
Response by Meta**

Do you have any views about our interpretation of the links between risk factors and different kinds of content harmful to children? Please provide evidence to support your answer.

Response: Please see our response to Question 4.

Confidential? - Y/N

Response: N

Question 6:

Response by Meta

Do you have any views on the age groups we recommended for assessing risk by age? Please provide evidence to support your answer.

Response:

We generally agree with Ofcom's recommended age groups for assessing risk by age as they provide a well-considered framework for evaluation.

Similarly, we have long taken a variety of steps designed to ensure users of our services are the appropriate age, and are committed to exploring additional ways to ensure young people receive age-appropriate experiences on our services. We recognise the merits of Ofcom's suggested age bands, particularly in the detailed understanding of the evolving online engagement among 13-17 year olds. We appreciate the insights and are constantly working to provide age appropriate experiences.

Confidential? - Y/N

Response: N

Draft Guidance on Content Harmful to Children (Section 8)

Question 12:

Response by Meta

Do you agree with our proposed approach, including the level of specificity of examples given and the proposal to include contextual information for services to consider?

Response:

We generally agree with the proposed approach. The contextual factors and additional information for services to consider that Ofcom lays out, while helpful, require further consideration and refinement to allow for more nuances:

- While we recognize the importance of removing harmful content, we also need to recognize the importance of platforms to create a space for expression, and limit expression only in service of values of authenticity, safety, privacy and dignity.
- Providers have different capabilities and technological limitations in content moderation and hence need to retain a certain level of flexibility to innovate.

- There are a variety of types of content (video, images, text) and types of interaction (one-to-one and analogous to the private sphere, while others are more like a town square or broadcasting to a very large group).
Confidential? - Y/N
Response: N

Question 14: Response by Meta
For each of the harms discussed, are there additional categories of content that Ofcom
c) where our current proposals should be reconsidered?
Response to c): We appreciate Ofcom’s thoughtful approach to the matter of content moderation. The considerations to be taken into account when trying to achieve the goal of reducing potentially harmful content have to be balanced against the interests of preserving freedom of expression and privacy. We would for example <u>encourage providers to allow for exceptions to their standard enforcement approach to potentially harmful content if available context can help determine whether or not that content is suitable for young people</u> . Some of these scenarios include content that is posted in medical, educational, cultural or religious contexts and that could affect this particular demographic, so it’s important to ensure that providers are not censoring content that can help mitigate the risk of harm, or content that, depending on the <u>context</u> in which it appeared, may not be harmful.
Confidential? - Y/N
Response: N

Volume 4: How should services assess the risk of online harms?

Governance and Accountability (Section 11)

Question 15: Response by Meta
Do you agree with the proposed governance measures to be included in the Children’s Safety Codes?
b) If you responded to our Illegal Harms Consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.

Response to b):

We refer to our response to Question 3(i) of the illegal harms consultation regarding the corresponding governance and accountability measures, which applies similarly here.

Confidential? - Y/N

Response: N

**Question 16:
Response by Meta**

Do you agree with our assumption that the proposed governance measures for Children's Safety Codes could be implemented through the same process as the equivalent draft Illegal Content Codes?

Response:

We agree that the governance measures can be implemented through the same process as the draft Illegal Content Codes. As stated above in response to Question 15, we consider that Ofcom's governance and accountability recommended measures relate to the safety of all users, and we intend to implement them as such.

Confidential? - Y/N

Response: N

Children's Risk Assessment Guidance and Children's Risk Profiles' (Section 12)

**Question 17:
Response by Meta**

What do you think about our proposals in relation to the Children's Risk Assessment Guidance?

Response:

We broadly support Ofcom's proposals in relation to the Children's Risk Assessment Guidance. We refer to our response to Question 7(i) of the illegal harms consultation. In particular, we would like to emphasise the following:

"Significant Change:"

Our response to Question 7(i) in the illegal harms consultation applies similarly here, with the exception of our comments on headcount changes, as we note that this is no longer listed as an example of something likely to amount to a “significant change”.

For the reasons previously set out in our response to the illegal harms consultation, we encourage Ofcom to align the definition of “significant change” more closely with that of “critical impact” under Art. 34 DSA. The definition of “significant change” should relate to the change’s impact on the risk and not to the “change” itself. Ofcom’s guidance on what constitutes a “significant” change is novel and overly broad. Whilst the DSA allows providers to assess this on their own, Ofcom has proposed a number of bases on which a new assessment would be needed, including introducing a new recommender system or machine learning model, changing “content rules” (such as a Community Standard), and changing the design of a reporting button or other icons for how users react to content (Annex 6, para 6.19).

This guidance is overly prescriptive, making it difficult to effect in practice. It being so prescriptive also renders it susceptible to becoming quickly out of date as online safety risks (and our internal systems to tackle those risks) continually evolve. In any one of the examples provided by Ofcom, there can be a large spectrum of changes that could occur, most of which should not meet this threshold and that needs to be made clear.

In addition, the current definition of “significant change” would lead to significant change risk assessments (“**SCRAs**”) being triggered in instances where the risk would be reduced as well as increased, which we believe is not the intention. For SCRAs to be effective, the question as to whether or not one is required should be determined by reference to the identified risk, and whether it would be negatively impacted by the proposed change. This would lead to SCRAs only being triggered when necessary. The test should not apply to the change itself or to the general impact on a group, but should consider whether or not the identified risk would be negatively impacted.

We advocate for a scoping process to be carried out in the first instance so that where such proposed changes do not significantly and negatively impact the risk to any particular group, a formal SCRA would not be triggered. For instance, a number of the examples of “significant change” set out in the guidance are based on the change having an ‘impact’ or ‘effect’, without specifying how material that impact / effect needs to be for the change to qualify as significant, or how it should impact the risk in order to trigger a SCRA. Annex 6 Table 6.1 states that a proposed change is “very likely” to amount to a significant change if it “impacts a vulnerable user group, such as children” or if it “impacts the efficacy” of measures put in place following the previous risk assessment. However, it does not indicate (i) how significant that impact should be, or (ii) what proportion of the vulnerable user group or anti-illegal content measures need to be affected, or (iii) that the proposed change must negatively impact the risks to those groups (to avoid a SCRA being triggered where the risk is reduced as well as increased).

While Ofcom may wish to provide some examples of the types of issues/impacts to risk that may

be in scope, such a list should be for illustrative or guidance purposes only so that proportionality and the unique circumstances of each change and its impact on risk can be applied in order to reduce a potentially disproportionate burden.

a) Please provide underlying arguments and evidence of efficacy or risks that support your view.

Response: Please see the response to Question 17 above.

Confidential? - Y/N

Response: N

Question 18:

Response by Meta

What do you think about our proposals in relation to the Children’s Risk Profiles for Content Harmful to Children?

Response:

Our response to Question 9(i) of the illegal harms consultation applies similarly here. Please also see our response to Question 19 below in which we set out our general observations regarding whether the Children’s Risk Profiles are useful models to help services understand the risks that their services pose to youth, and comply with their children’s risk assessment obligations. We will consult these tables when doing our risk assessment, however it should be recognised that we have many years of experience of identifying and mitigating risks and we will continue to use our own risk signals and analysis, in addition to the Risk Profiles.

We note Ofcom’s proposal to add a new specific risk factor to the Children’s U2U Risk Profile regarding the risks associated with features and functionalities which influence how much users, including young people, use a service (on the basis that greater usage time increases the likelihood of encountering harmful content). On this point, we have already rolled out features that let people manage their time, prevent unwanted interactions and control what type of content and accounts they see.

We appreciate Ofcom’s acknowledgment that the factors identified in the Risk Profiles do not provide a complete picture of risk, and that additional factors should be considered. We encourage Ofcom to also consider how some factors identified as increasing risk could in fact decrease risk. For instance, Ofcom has identified the ability to create user profiles as a risk factor for eating disorder content, without acknowledging how this feature may decrease the likelihood of other risks, particularly those Ofcom has described regarding the ability to share content anonymously. We note that providers should be afforded the flexibility to consider the potential impact on risks as more than a binary or “one-way” determination as suggested by the Risk

Profiles.
a) Please provide underlying arguments and evidence of efficacy or risks that support your view.
Response: Please see our response to Question 18 above.
Confidential? - Y/N
Response: N

Question 20: Response by Meta
Are there any specific aspects of the children’s risk assessment duties that you consider need additional guidance beyond what we have proposed in our draft?
Response: We refer to our <u>responses to Questions 7, 17, 18 and 19 above</u> , and our response to <u>Question 9(i) of the illegal harms consultation</u> regarding the Risk Profiles, which applies similarly here. In addition, we would like to raise the following point with regards to the notification of NDC to Ofcom. Given that there is a duty to report NDC to Ofcom, the definition of NDC will be very important. We would welcome <u>further dialogue with Ofcom on what constitutes NDC and how to quantify it</u> . We expect that this definition will be refined over time, from practical experience in implementation.
Confidential? - Y/N
Response: N

Question 21: Response by Meta
Are the Children’s Risk Profiles sufficiently clear and do you think the information provided on risk factors will help you understand the risks on your service?
Response: We refer to our responses to Questions 18 and 19 above, and our response to Question 9(i) of the illegal harms consultation regarding the Risk Profiles, which applies similarly here.
a) If you have comments or input related to the links between different kinds of content harmful to children and risk factors, please refer to Volume 3: Causes and Impacts of Harms to Children Online which includes the draft Children’s Register of Risks.

Response:

We refer to our responses to Questions 18 and 19 above, and our response to Question 9(i) of the illegal harms consultation regarding the Risk Profiles, which applies similarly here.

Confidential? - Y/N

Response: N

Volume 5 – What should services do to mitigate the risk of online harms

Developing the Children’s Safety Codes: Our framework (Section 14)

Question 25:

Response by Meta

Do you agree with our approach to developing the proposed measures for the Children’s Safety Codes?

Response:

We are grateful for Ofcom’s clarity and openness in their process of development of the Children’s Safety Codes and we welcome this consultation process as it helps to ensure that impacts of options expressed in the consultation can be properly considered before such Codes are finalised.

Based on the development of codes of practice we have seen around the globe, we believe that such instruments allow for greater flexibility to develop, adapt and respond quickly to ever-evolving technological environments. To that end, we also appreciate that Ofcom is seeking input from consultees on some topics that may only be addressed in future iterations of the Children’s Safety Codes, and that the Children’s Safety Codes will generally evolve over time as Ofcom learns more and the sector develops. Indeed, our own experience in addressing different types of content and integrity issues, is one of ongoing development. Therefore, flexibility remains key.

We make reference to the points developed in the illegal harms consultation as well, which includes e.g. the below which also applies to the proposals for updates of the Illegal Content Codes and to the Children’s Safety Codes. We acknowledge that due to timelines between the previous illegal harms consultation and the publication of this consultation, previous feedback may not have been fully included yet by Ofcom. Please refer to Question 12 of our response to the

illegal harms consultation for more details.

- We strongly support: Ofcom’s aim to capture existing good practice within industry, set clear expectations, and work to raise standards of user protection over time, Ofcom’s acknowledgement of the advantages and aim of alignment with other content regulations, Ofcom’s commitment that Children’s Safety Codes have to be based on proportionality and that expectations for services need to be clear.
- Some preliminary suggestions and concerns:
 - We notice some divergences from other regulations. In our view, it is essential to develop regulatory models which are workable within the full spectrum of other, relevant regulations for online services (such as privacy regulation) and the global regulatory environment, in order to avoid fragmentation of the technological landscape. As such, Codes of Practice should avoid setting up dual regulatory regimes. We mention this e.g. for the Children’s Safety Codes on age assurance below regarding the ICO.
 - We believe that for example the following proposals would help to support proportionality in relation to the consultation and the regulatory model that will ultimately be adopted:
 - Clarity on boundaries of a service - and in case such guidance is not forthcoming, will be developed at a future date or those boundaries are drawn broadly - flexibility should be built in as to what proposed measures are required for different parts of such services that have different risk classifications.
 - More options to constitute a ‘safe harbour’. While compliance with measures recommended in the Children’s Safety Codes will effectively give providers a ‘safe harbour’ regarding their compliance with the Act, most providers will already have a range of safety and integrity measures in place, which will vary depending on the design and underlying integrity systems of the service. We propose that Ofcom, rather than providing a single set of recommended measures that will constitute a ‘safe harbour’, suggests or allows for a range of options which would meet the objectives of the Act, to reflect the flexibility and variation among services. It would encourage an ongoing flexibility of approach to user safety, which is essential in an environment of constant technical evolution.
 - Greater clarity as to how the ‘safe harbours’ available under the Children’s Safety Codes will apply to the relevant duties of the OSA.
- Timeline: Since the implementation of some measures proposed in the Children’s Safety Codes will depend on the results of the risk assessment, clarity in which month the final guidance for children’s access assessment and children’s risk assessment is expected would be helpful for resource planning on such implementation work. We would be grateful if Ofcom could publish an updated version of its timeline.

One topic of concern more specific to this consultation lies in the importance of taking human rights into consideration for the proposals.

- Whilst Ofcom recognises the potential adverse impact of its proposed measures on the rights to freedom of expression and privacy of young people and adults, it concludes that any interference is necessary and proportionate to protect young people from harmful content. We are concerned however that the necessity and proportionality of each measure is not established in each and every case.

a) If not, please explain why.

Response: Please see our response to Question 25 above.

Confidential? - Y/N

Response: N

Question 26:

Response by Meta

Do you agree with our approach and proposed changes to the draft Illegal Content Codes to further protect children and accommodate for potential synergies in how systems and processes manage both content harmful to children and illegal content?

Response:

We welcome Ofcom's approach to search for synergies in how systems and processes are to be deployed under different Codes of Practice. As an international company, we recognise there are multiple content regulation laws across the globe, and we build integrity structures around those different obligations. As such, we would always try to look for common denominators.

Overall our approach is that we aim to protect all our users by integrating a comprehensive suite of safeguards and controls. However, we recognize that young people on our platform require a higher level of protection. We have implemented and will continue to implement additional safeguards in the best interests of young people, striking a balance between protecting them and facilitating their connection and development in the digital environment.

a) Please explain your views.

Response: Please see our response to Question 26 above.

Confidential? - Y/N

Response: N.

Question 29:**Response by Meta**

Do you agree with our definition of 'multi-risk' and with how we apply this in our recommendations?

Response:

We broadly agree with Ofcom's definition of 'multi-risk'. However, we have some concerns which are outlined below.

We refer to our response to Question 15(i) of the illegal harms consultation regarding the definition and application of the term 'multi-risk'. Our comments in that response regarding proportionality and flexibility apply similarly here:

- We understand Ofcom defines a "multi-risk" service as a service which is high or medium risk for at least two kinds of content harmful to children, based on the provider's most recent risk assessment. Based on this definition, it follows that a service which is high or medium risk for two kinds of content harmful to young people will be subject to the same stringent measures as a service that is high or medium risk for every kind of such content. While we appreciate that it is difficult to define where to draw the line for 'enhanced' measures, further proportionality could be considered.
- As a matter of principle, we believe that risk assessment obligations should include overall guidance on the structure and content of the risk assessments, while ensuring sufficient flexibility for different services and adaptability for the future.

Confidential? - Y/N

Response: N

Question 30:**Response by Meta**

Do you agree with the proposed measures that we recommend for all services, even those that are small and low-risk?

Response:

Our response is on U2U services. We assume for U2U services this question refers to Measure GA2, CM1, UR1 to UR4(d), TS1 and TS2.

We broadly support these proposals with the view to further creating an environment where we

can provide young people with safe and age-appropriate experiences, meaning they can avail themselves of all the benefits our services have to offer.

It is often assumed that the largest services are the main problem but addressing these risks wherever they appear regardless of size is important. To that end, we agree that it is important that all services accessible to young people are tasked with various minimum safety steps to create a safe and transparent environment.

Taking each measure in turn:

Measure GA2 (PCU A2): Person accountable for the safety duties protecting children and reporting and complaints duties

As stated in our response to Question 3(i) of the illegal harms consultation, and our response to Question 15 above, we are in the process of appointing an OSA Compliance Officer who will work within our newly formed UK/EU Integrity Compliance Office. We have also named a person accountable for our compliance with our safety duties under the OSA regime in line with this recommended measure. Given our scale and complexity, it is unlikely to be possible to nominate a discrete accountable individual per service. In any event, we note that the consultation does not require the accountable individual to only have one service under their purview.

Measure CM1 (PCU B1): Having a content moderation function that allows for swift action against content harmful to children;

We agree that service providers should have content moderation functions designed to swiftly take action against content harmful to users of which a provider is aware.

However, content moderation is a complex process involving multiple factors. As mentioned in the response to the illegal harms consultation (refer to Question 18), we suggest that service providers are granted flexibility to prioritise in a way that is appropriate for their service and in light of a wider range of relevant factors.

Measure UR1 to UR4(d) (PCU C1 to PCU C11): Enabling complaints; Having an easy to find, easy to access and easy to use complaints system; Provision of information prior to the submission of a complaint; Taking appropriate action of various kinds in response to different types of complaint

Please see our response to Question 43(a) below.

Measures TS1 and TS2 (PCU D1 and PCU D3): Substance of the ToS / statements for all services: Clarity and accessibility

Please see our response to Question 46 below regarding Measures TS1 and TS2. Providers should retain scope and flexibility as to where information is provided to users. With our family of apps, it's much easier and more intuitive for a young person to be signposted from within the app to ToS, as well as settings, various policies and useful tools.

We have policies and transparency pages that are specifically designed and written to ensure that these are at an age-appropriate reading level to achieve high levels of readability for all our users. In addition to transparency surfaces which are available for all users (such as our [Transparency Center](#), [Help Center](#) and [Safety Center](#)), we have also developed tailored resources specifically for youth - e.g., Facebook's [Youth Portal](#), and Instagram's guide specifically aimed at young people.

Confidential? - Y/N

Response: N

Content moderation U2U (Section 16)

Question 36:

Response by Meta

Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.

Response:

We overall agree with the proposals suggested by Ofcom with some reservations.

Measure CM1: content moderation systems and processes are designed to swiftly take action against content harmful to children of which the provider is aware (a 'content moderation function')

- We agree that service providers should have proportionate content moderation systems and processes designed to swiftly take action against content potentially harmful to users of which a provider is aware.
- We support Ofcom's approach to not set a specific timeframe for reviewing content which might be harmful to young people as service providers are asked to prioritise a large amount of content for review and the integrity of the platform will rely on scalable processes. As mentioned in our response to Question 18 in the illegal harms consultation, we suggest that service providers are granted flexibility to prioritise in a way that is appropriate for their service and in light of a wider range of relevant factors.
- We also believe service providers should protect young people against mature or disturbing content by making it inaccessible to some users through age-gating. Where content does not violate Facebook's Community Standards or Instagram's Community

Guidelines but may nevertheless be sensitive to some users, we may take other enforcement actions like labelling, limiting distribution, adding sensitivity screens or restricting users under a certain age from seeing the content.

Measure CM2: when a service is large and / or multi-risk for content harmful to children, it should set internal content policies

- We support Ofcom’s proposal that service providers should have internal content policies. We have a vast portfolio of content policies to enforce on various topics, and while there are policies specifically designed to protect youth from abusive behaviour (e.g. Child Sexual Exploitation, Abuse and Nudity), all our policies, detection and enforcement systems are designed keeping in mind the protection of younger audiences. We publish them externally on our Transparency Center (Community Standards and Community Guidelines). Providers should be able to retain flexibility. In developing such policies, we seek to balance freedom of expression alongside other important values, such as safety, privacy and dignity. For example, with respect to suicide and self-injury content, these rules prohibit content that intentionally or unintentionally encourages or promotes self-injury and/or suicide (this includes e.g. fictional content and any self-injury content which is graphic, regardless of context) but permit people to discuss these topics, raise awareness and seek support. Experts have told us that in the right circumstances, content which touches on suicide and self-injury can be shared in a positive context and can play an important role in destigmatising mental health difficulties. However, we recognise that this content can be sensitive and therefore, as well as having processes in place to identify and remove policy-violating content, we take steps to remove this content from recommendation surfaces and properly mark (e.g., via sensitivity screens) potentially sensitive content. This comes with other control measures relevant to this space, including abuse prevention features, age-gating, search restrictions, alternative topic nudges, comment nudges and sensitive content controls.

Measure CM3: when a service is large and / or multi-risk for content harmful to children, it should set performance targets for its content moderation function

- As mentioned in our response to Question 18 in the illegal harms consultation, we support an approach that does not involve setting prescriptive internal performance or accuracy targets, but that rather emphasises the general requirement for content to be reviewed swiftly, as stated by Ofcom in measure CM1. Indeed, even if just in target form, performance targets do not account for the necessary nuance in assessing cases with differing levels of complexity which may require complex rights-balancing assessments nor do they account for the complexity of at-scale content moderation systems which aim to prioritise the review and removal of a variety of types of violating content against a range of factors e.g. virality, severity. Indeed, even in a given “violation” category, no two violations are the same, so it is not practicable to set a single target turnaround time for a given category, let alone a single turnaround time for a reporting and flagging mechanism

as a whole. Requiring providers to set and comply with turnaround targets has a real potential to create a blunt solution with unintended consequences and/or incentives, and reduce the efficacy of report/flagging handling.

- However, we do agree that service providers need to aim to reduce potential harms on their platform, and that some types of harms should be prioritised for review, especially some policies which are more specifically designed to protect young users (e.g. Child Sexual Exploitation, Abuse and Nudity)

Measure CM4: Measure CM4 (PCU B4): when a service is large and / or multi-risk for content harmful to children, it should prepare and apply policies on prioritisation of content for review

- As mentioned in our response to Question 18 in the illegal harms consultation, we are supportive of the objective of prioritisation for content reviewing by service providers. We agree with the criteria of “virality” and “severity”, but we also recognise there are significant and concerning practical implications as to the details of these factors, the remainder of Ofcom’s proposed factors, the interplay of the factors, applying these factors to review of harmful content and conflicts with other proposed measures such as the ones laid down in the illegal harms consultation.
- Given our global scale, where there are conflicting requirements, we cannot prioritise so as to meet all of those requirements. For example, in the illegal harms consultation, Ofcom suggests that providers should prioritise review of potentially illegal content, while in the consultation for child safety duties, Ofcom recommends prioritising review of harmful content yet not illegal. In order to ensure an efficient prioritisation process at scale, when we review content, we would in the first instance be assessing content for violations of our policies, and not for illegality. Such policy review as a first step has advantages, as content that is assessed as policy violating would be removed globally instead of merely being subject to a local restriction. We suggest that the measure instead gives providers the flexibility to assess severity in a way that works for their moderation process, which often operates most effectively at a global scale (as the underlying systems can continually evolve and be trained on a larger set of content and data).
- In addition, where reports from trusted flaggers are concerned, in most cases those reports are processed via separate, dedicated channels that don’t operate in tandem with content moderation systems.
- As such, we would recommend allowing greater flexibility for providers in relation to prioritisation. This would mean that these considerations could be balanced by providers in a way that is effective for their services, rather than prescribing an approach to prioritisation that may have unintended consequences. We would like to highlight in this context, that our content review and enforcement varies depending on the nature of the harm presented, and more egregious harms (e.g. Child Sexual Exploitation, Abuse and Nudity) will be prioritised against lower severity policy areas (e.g. Spam).

Measure CM5: Measure CM5 (PCU B5): when a service is large and / or multi-risk for content

harmful to children, it should ensure content moderation functions are well-resourced

- As mentioned in our response to Question 18 of the illegal harms consultation, while we agree with the goal of ensuring appropriate resourcing, including adaptability under changing circumstances, we would highlight that such a requirement needs to be proportionate. For example: it is reasonable for providers to have appropriate backup strategies for unexpected surges in report volume, but even with such strategies in place, not every eventuality can be planned for, and some situations may require additional ad hoc measures that will take time to implement. Also, while we agree with the goal to have appropriate resourcing in place and cater for particular needs of a UK user base, we also stress that the correct staffing of language support worldwide is a complex planning exercise; even on a per region level depending on the number of languages spoken in the region.
- Based on our experience, we set out below strategies that we have found helpful, which Ofcom may wish to consider reflecting in the measure:
 - Using a mix of reviewers: We use reviewers from a variety of different backgrounds, to cater for appropriate language support and understanding of cultural context. However, for languages that are widely spoken in the world, like English, it is helpful to have content moderation teams that provide global coverage, which enables the provider to quickly redeploy capacity if there is a surge in demand in a specific country, in times of crisis, or when unpredictable events occur. This means it is helpful to have the option to rely temporarily on teams that provide global coverage to mitigate more local risks.
 - Allowing for triaging where more or less language support is needed: Language expertise helps to e.g. enforce policies in cases where certain words or content require additional contextual understanding, but not all content requires language expertise. For example, some nudity and sexual activity is language agnostic. For this type of content, it is helpful to have a global pool of content moderators who review these types of reports.

Measure CM6: Measure CM6 (PCU B6): when a service is large and / or multi-risk for content harmful to children, it should ensure content moderation teams are appropriately trained

- As mentioned in our response to Question 18 in the illegal harms consultation, we support Ofcom's recommendation to ensure moderators are appropriately equipped. To that end, we support appropriate measures, such as distinct training based on the nature of the particular moderator's work. We would point out that the implementation of this measure in practice should be proportionate to achieve its purpose as, due to the structure of our global system, moderators work on the basis of the provider's policies. Given this, where content moderators work on reviewing content for violation of the provider's policies, they should receive training on such policies. We welcome the flexible approach adopted by Ofcom as to the details of this training as it allows for the providers to develop resources and strategies in a way that is appropriate for their service and reflective of the

particular risks of harm that could apply to their service.

- Based on our experience, we set out below strategies that we have found helpful, which Ofcom may wish to consider reflecting in the measure:
 - Building review teams for policy violations that include a diverse range of backgrounds, and that include experts in enforcement in policy areas such as counter terrorism and child safety. Review teams should undergo initial training to ensure that they have a strong grasp on policies, and should be trained and tested as appropriate beyond this initial training.
 - Reviewers are regularly trained and informed to clarify and provide valuable feedback on the enforcement of the providers' content policies and areas of improvement as well as processes to ensure prompt communication and training of any policy or scope changes that might affect content moderation.

Measure CM7:

- Based on our experience, we would not recommend the large-scale reliance on volunteer moderators to tackle potential harmful content for children.
- Based on our experience, in case volunteer moderators are deployed we agree they should undergo extensive training and be provided with training materials.

Confidential? - Y/N

Response: N

Question 37:

Response by Meta

Do you agree with the proposed addition of Measure 4G to the Illegal Content Codes?

Response: Please see our response to Question 36 above.

a) Please provide any arguments and supporting evidence.

Response: Please see our response to Question 36 above.

Confidential? - Y/N

Response: N

User reporting and complaints (Section 18)

Question 43:

Response by Meta

Do you agree with the proposed user reporting measures to be included in the draft Children's Safety Codes?

Response: Please see our response to a) below.

a) Please confirm which proposed measure your views relate to and explain your views and provide any arguments and supporting evidence.

Response:

We refer to our response to Question 28(i) of the illegal harms consultation regarding the equivalent proposals for user reporting and complaints, which applies similarly here (with the exception of our comments on dedicated fraud reporting channels, which are not relevant to the current consultation), e.g. on performance targets.

Whilst we broadly agree with the measures proposed, in addition to the above transferable aspects, we have significant concerns over the scalability and workability of some measures as further described below. Some measures may not be possible from an operational perspective.

Taking selected measures in turn:

Measure UR2(a) - for relevant complaints regarding a specific piece of content, a reporting function or tool should be clearly accessible in relation to that content, and processes for making other relevant complaints should be easy to find and easily accessible:

We have concerns regarding the workability and proportionality of this measure, especially as applied to 'affected persons'. As set out in our response to Question 28(i) of the illegal harms consultation, non-users (including affected persons) will not necessarily have access to all content on a service, for reasons of safety and user privacy. As such, affected persons will not necessarily be able to see an item of content that they may wish to report or complain about, and so may not be able to access a reporting / complaint tool directly from that content. We therefore strongly suggest that 'accessible' functions and tools for reporting / complaints are not limited to those accessible directly from the relevant item of content, but are deemed to include other functions and tools accessible to non-users - for example, reporting by way of a Help Center form.

Measure UR2(b) - ensuring information and processes relating to complaints are accessible and comprehensible, having regard to the particular needs of the service's UK user base as identified in its children's risk assessment:

We have concerns regarding the workability and proportionality of this measure, as if we are required to simplify the language of written information provided on complaints processes in order to ensure comprehensibility for the youngest users of a service, this potentially may reduce transparency for other users.

Indeed, we support meaningful transparency for users, in a form that is appropriate to ensure maximum comprehension by them. As a result, we provide information in our Community Standards for Facebook and Community Guidelines for Instagram, and supplement that with teen-specific and parent-specific information in our Youth Portal, Family Center, Help Centers, Community Safety Center and Parents' Guide which are specifically designed to be easily accessible to minors and parents.

Measure UR2(c) - the number of steps (e.g. clicks, navigation points) to make a complaint should be as few as reasonably practicable:

We broadly support the proposal that users should be able to make a complaint in as few steps as reasonably practicable. We note that applying that reasonability threshold may call for some steps to require further detail to garner the route to which a report should be directed or to safeguard the integrity of our reporting mechanisms against misuse.

Measure UR2(d) - enabling complainants to include relevant information or supporting material when making a complaint:

Whilst we broadly support the proposal that users can submit supporting materials when making a complaint, we have concerns surrounding the submission of unstructured data from an operational perspective. We believe service providers should deploy reporting flows where users can link the relevant content that they are reporting.

Measure UR3(a) - acknowledge receipt of each relevant complaint and provide the complainant with an indicative timeframe for deciding the complaint:

We have concerns regarding the workability and proportionality of this measure. Addressed in our response to Question 28(i) of the illegal harms consultation, given the complexity of some of these complaints, and the potential for appeals, it may not be possible, at the outset of a complaint, to estimate a concrete timeframe for deciding a complaint.

Measure UR4(b) - appeals

Whilst we broadly support the proposal that certain reports and appeals are afforded priority for different reasons, providers should retain scope and flexibility to determine priority of reports and appeals. When deciding the priority of appeals, this results in tradeoffs. This prioritisation could have consequences in other jurisdictions or result in service providers needing to create different priority/ranking systems for different jurisdictions.

Additionally, we have concerns regarding the scalability and workability of the proposal to ensure that when an appeal decision is reversed, the automated content moderation technology be

changed so that it does not cause the same piece of content to be taken down, downranked, or otherwise restricted again.

Measure UR4(c) - complaints about incorrect assessment of a user's age

Whilst we broadly support that certain factors are considered when determining the priority of an appeal against an incorrect assessment of a user's age, we reiterate that service providers should be afforded some scope and flexibility as to those factors. We also refer to our response to Question 31 above on highly effective age assurance.

Measure UR4(d) - complaints about non-compliance with safety duties protecting children

Please see our response to Question 28(i) of the illegal harms consultation (in particular, but not limited to, part 4 of that response and our comments regarding misuse of reporting mechanisms ("spamming")).

b) If you responded to our Illegal Harms Consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response

Response:

We refer to our response to Question 28(i) of the illegal harms consultation regarding the equivalent proposals for user reporting and complaints.

Confidential? - Y/N

Response: N

**Question 44:
Response by Meta**

Do you agree with our proposals to apply each of Measures UR2 (e) and UR3 (b) to all services likely to be accessed by children for all types of complaints?

Response:

We broadly agree with the proposed inclusion, subject to the publication of the final Children's Safety Codes.

a) Please confirm which proposed measure your views relate to and explain your views and provide any arguments and supporting evidence.

Response: Please see our response to Question 44 above.

Confidential? - Y/N

Response: N

Question 45:

Response by Meta

Do you agree with the inclusion of the proposed changes to Measures UR2 and UR3 in the Illegal Content Codes (Measures 5B and 5C)?

Response:

Yes, we broadly agree with the proposed inclusion, subject to the publication of the final Children’s Safety and Illegal Content Codes.

a) Please provide any arguments and supporting evidence.

Response: Please see our response to Question 45 above.

Confidential? - Y/N

Response: N

Terms of service and publicly available statements (Section 19)

Question 46:

Response by Meta

Do you agree with the proposed Terms of Service / Publicly Available Statements measures to be included in the Children’s Safety Codes?

Response:

We refer to our response to Question 29(i) of the illegal harms consultation regarding the equivalent proposals for terms of service and publicly available statements, which applies similarly here.

Child safety is an adversarial space replete with highly motivated, sophisticated actors who regularly test and disseminate information about how to evade online safety protections, exploit potential vulnerabilities, and create new potential abuse vectors. Online offender communities have been known to use public or adversarially obtained information to create how-to guides and share detailed advice on online platforms’ policies, detection, and enforcement systems in order to facilitate other community members’ predatory activity. In relation to Measure TS1, our experience suggests that it is highly important to balance any public disclosure of the functions and mechanics of our protective technologies against the likelihood that bad actors will leverage this information to evade our systems. Providers should be asked to share only high-level and overarching information on proactive technologies used for compliance with the provider’s child safety duties. Otherwise, this may jeopardise the efficiency of providers’ child protection

measures and may actually help bad actors to circumvent such measures.

In addition, we do not believe that the proposed Terms of Service / Publicly Available Statements measures should be included in the Children's Safety Codes. For Facebook and Instagram, it's much easier and intuitive for a young person to be signposted from within the app to ToS, as well as settings, various policies and useful tools. We have a child protection section on our [Safety Center](#), highlighting our comprehensive approach to child safety that includes zero tolerance policies prohibiting child exploitation, cutting-edge technology to prevent, detect, remove and report policy violations, and victim resources and support.

In relation to [Measure TS2](#), we are supportive of meaningful transparency in a form that is appropriate to ensure maximum comprehension by its users. As a result, we provide information in our Community Standards for Facebook and Community Guidelines for Instagram, and supplement that with teen-specific and parent-specific information in our Youth Portal, Family Center, Help Centers, Community Safety Center and Parents' Guide which are specifically designed to be easily accessible to minors and parents.

As stated in our response to Question 29(i) of the illegal harms consultation, our experience suggests that providers should have the flexibility to put this information in certain documents or locations (which are incorporated into the ToS by reference) in order to improve transparency, readability and clarity for users. We take care to ensure that in our communications with users, including its terms and conditions, and any updates or changes thereto, are set out in clear, plain, intelligible, user-friendly language for users of all ages. In this context, we believe that the core terms of service should remain short and clear. If the ToS are too complex and granular, this could defeat the objective of transparency for users.

The Transparency Center provides users with detailed information on our policies and enforcement. We propose that the information required in Measure TS1 could be made available to users in our Transparency Center.

a) Please confirm which proposed measures your views relate to and provide any arguments and supporting evidence.

Response: Please see our response to Question 46 above.

b) If you responded to our illegal harms consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.

Response:

We refer to our response to Question 29(i) of the illegal harms consultation regarding the equivalent proposals for terms of service and publicly available statements, which applies similarly here.

These measures, which reflect the OSA’s terms of service provisions, require extensive information to be included in the ToS. Our experience suggests that providers should have the flexibility to put this information in separate documents or locations (which are incorporated into the ToS by reference) in order to improve readability and clarity for users.

We take care to ensure that all of our communications with users, including our terms and conditions, and any updates or changes thereto, are set out in clear, plain, intelligible, user-friendly language for users of all ages. In this context, we believe that the core terms of service should remain short and clear.

Confidential? - Y/N

Response: N

Recommender Systems (Section 20)

Question 50:

Response by Meta

Are there any intervention points in the design of recommender systems that we have not considered here that could effectively prevent children from being recommended primary priority content and protect children from encountering priority and non-designated content?

Response:

Notwithstanding our comments on the definition of “non-designated content” (see our response to Question 52, below), where we outline issues relating to including certain content in that definition, the design of the recommender systems must be done in a flexible manner that allows them to evolve and adjust to developments over time.

While Ofcom has laid out some solutions, we believe that service providers should retain flexibility with regards to the best ways to tackle this in the current technological environment.

In addition to recommender systems, user controls and choices are equally important. Coupled with parental supervision tools and resources to help them set boundaries with their teens, and protections to keep them safe, families are able to make better age and developmentally appropriate choices.

Confidential? - Y/N

Response: N

Question 52:**Response by Meta**

We plan to include in our RS2 and RS3, that services limit the prominence of content that we are proposing to be classified as non-designated content (NDC), namely depressive content and body image content. This is subject to our consultation on the classification of these content categories as NDC. Do you agree with this proposal? Please provide the underlying arguments and evidence of the relevance of this content to Measures RS2 and RS3.

Response:

We are supportive of measures to limit the prominence of content which may be potentially harmful, and we believe that the approach to identifying NDC must be nuanced and take into account the context in which such content is shared. For example, content that promotes positive body image or discusses mental health issues, in a responsible and sensitive manner, or that may be relevant for user education or debate, should not be considered NDC. If it is, it may be inadvertently captured as a result of the measures.

We believe that the feasibility of Ofcom's four-step approach to identifying NDC would benefit from further review. While we understand the need for a robust framework to identify and regulate NDC, the proposed approach may be overly complex and difficult to implement in practice. In particular, the requirement for platforms to use AI and machine learning algorithms to detect NDC may be challenging for smaller platforms with limited resources.

We suggest that Ofcom consider a more flexible and scalable approach to identifying NDC, one that takes into account the diversity of online content and the different contexts in which it is shared. This could be more effective in addressing the issue of NDC while also being more feasible for platforms to implement.

In conclusion, while we support Ofcom's efforts to address the issue of NDC, we believe that their approach to identifying it requires further consideration and refinement.

Confidential? - Y/N

Response: N

Statutory tests (Section 24)

Question 59:**Response by Meta**

Do you agree that our proposals, in particular our proposed recommendations for the draft Children's Safety Codes, are appropriate in the light of the matters to which we must have regard?

Response:

As mentioned throughout our responses above, we are generally supportive of a number of Ofcom's proposed measures but we do have concerns as to the appropriateness, workability and proportionality of some measures.

We believe that any assessment of the causes and impacts of online harms, including Ofcom's, requires an evidence-based, proportionate approach and should be informed by experts. We work closely with external experts including the Safety Advisory Council and our Global Safety Network of over 400 safety experts. We consult with these organisations on issues related to online safety. Advisors provide expertise, perspective and insights that inform our approach to safety, especially with regard to young people. We note e.g. mental health affects people in different ways. Mental health experts tell us some people find it helpful to share their experiences of mental health and get support from friends, family and others in the community. Social media can also help tackle stigma associated with mental health. At the same time, these experts tell us what's helpful for some may be harmful for others. It is important to find the right balance between the protection of vulnerable and impressionable people from content that could put them at risk, and the ability of the individuals to express themselves and seek support, and we emphasise the notion of continued exchange.

a) If not, please explain why.

Response: Please see our response to Question 59 above.

Confidential? - Y/N

Response: N