Consultation response form

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

Consultation title	Protecting people from illegal harms online
Full name	Dr Edina Harbinja
Contact phone number	
Representing (delete as appropriate)	Organisation
Organisation name	BILETA (British and Irish Law, Education and Technology Association)
Email address	e.harbinja@aston.ac.uk

Confidentiality

We ask for your contact details along with your response so that we can engage with you on this consultation. For further information about how Ofcom handles your personal information and your corresponding rights, see Ofcom's General Privacy Statement.

Your details: We will keep your contact number and email address confidential. Is there anything else you want to keep confidential? Delete as appropriate.	Nothing
Your response: Please indicate how much of your response you want to keep confidential. Delete as appropriate.	None
For confidential responses, can Ofcom publish a reference to the contents of your response?	Yes

Your response

Question (Volume 2)	Your response
Question 6.1: Do you have any comments on Ofcom's assessment of the causes and impacts of online harms? Do you think we have missed anything important in our analysis? Please provide evidence to support your answer.	*[Is this answer confidential? No (delete as appropriate)]* *n/a*
Question 6.2: Do you have any views about our interpretation of the links between risk factors and different kinds of illegal harm? Please provide evidence to support your answer.	*[Is this answer confidential? Yes / No (delete as appropriate)]* *n/a*

Question (Volume 3)	Your response
Question 8.1: Do you agree with our proposals in relation to governance and accountability measures in the illegal content Codes of Practice? Please provide underlying arguments and evidence of efficacy or risks to support your view.	*[Is this answer confidential? No (delete as appropriate)]* The proposed four-step risk assessment process offers a valuable framework for U2U and search services to effectively identify and manage potential online harms. However, enriching this framework with granular guidance and enhanced clarity in specific areas can further empower services to fulfil their risk mitigation responsibilities.
Question 8.2: Do you agree with the types of services that we propose the governance and accountability measures should apply to?	*[Is this answer confidential? No (delete as appropriate)]* To improve platform accountability for online harms, several measures are proposed:

Question (Volume 3)	Your response
	1. Taxonomy of Harms: Providing a comprehensive yet accessible taxonomy of potential harms relevant to various platform functionalities (e.g., user connections, content posting, communication) would equip services with a structured approach to harm identification. 2. Platform-Specific Risk Analysis: Offering practical tools or methodologies for services to conduct platform-specific analyses of harm susceptibility would enable them to prioritize and tailor their risk assessments to their unique context (see to that effect Article 26 (risk assessment) and Article 27 (mitigation of risks) of the EU Digital Services Act. 3. Standardized Risk Assessment: Implementing standardized risk matrices or scoring systems (see to that effect CJEU SCHUFA (Scoring), would facilitate a more objective and consistent assessment of harm likelihood and impact. 4. Scenario Simulations: Encouraging services to conduct scenario-based simulations of potential harm scenarios (e.g., cyberbullying, disinformation campaigns) can foster a proactive approach to risk mitigation and preparedness. 5. Best Practice Library: To satisfy the accessibility and foreseeability principles of the ECtHR caselaw, for the UK Online Safety Act to observe the 'in accordance with the law' requirement under Article 8(2) and 10(2) ECHR, establishing a readily accessible library of best practices for implementing safety measures based on different types of identified harms would offer services valuable practical guidance. 6. Detailed Reporting: Specifying the level of detail required in risk assessment reports (e.g., risk identification, risk prioritisation, risk mitigation planning, risk monitoring and communicating risks), and establishing a clear reporting frequency (refer to e.g., annual Google's transparency reports), aligned with the platform's risk profile, would enhance transparency and accountability.

Question (Volume 3)	Your response
Question 8.3: Are you aware of any additional evidence of the efficacy, costs and risks associated with a potential future measure to requiring services to have measures to mitigate and manage illegal content risks audited by an independent third-party?	*[Is this answer confidential? No (delete as appropriate)]* No
Question: 8.4: Are you aware of any additional evidence of the efficacy, costs and risks associated with a potential future measure to tie remuneration for senior managers to positive online safety outcomes?	*[Is this answer confidential? Yes / No (delete as appropriate)]* n/a
Question 9.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No (delete as appropriate)] The proposed utilization of Risk Profiles presents a significant opportunity to enhance the efficacy of risk assessments undertaken by U2U and search services. However, a critical analysis reveals potential areas for optimization that would further empower these services to identify and mitigate potential online harms.
Question 9.2: Do you think the four-step risk assessment process and the Risk Profiles are useful models to help services navigate and comply with their wider obligations under the Act?	[Is this answer confidential? No (delete as appropriate)] For effective risk assessment, risk profiles should be granular or unbundled (like the UK Data Protection Act data subject consent's requirements), (e.g., social media platform, classifieds website), harm type (e.g., misinformation, child exploitation), or other relevant criteria. Illustrative examples showcasing diverse applications in different contexts like social media versus classifieds would enhance understanding. For instance, demonstrating how a social media platform and a classifieds website would approach

Question (Volume 3)	Your response
	a "grooming" risk factor differently would prove highly instructive. Clearly define the target audience (legal, technical, etc.) and tailor complexity accordingly (see to that effect e.g., CJEU C-210/16 Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH [36] – [39]; Case C-697/19 Teva Pharmaceutical Industries Ltd v European Union Intellectual Property Office (EUIPO) [19]; C-361/04 P Claude Ruiz-Picasso and Others v European Union Intellectual Property Office ECLI:EU:C:2006:25 [59]. Be transparent about required expertise (e.g., legal, or technical knowledge) and offer support if needed. Provide a robust analytical framework for evidence analysis (e.g., based on platform accountability caselaw and reports) and actionable conclusions. Include real-world case studies (like electoral misinformation or the current Taylor Swift deepfake porn case) and demonstrate openness to feedback for continuous improvement.
Question 9.3: Are the Risk Profiles sufficiently clear and do you think the information provided on risk factors will help you understand the risks on your service?[1]	*[Is this answer confidential? Yes / No (delete as appropriate)]* *Please refer to the above answer.*
Question 10.1: Do you have any comments on our draft record keeping and review guidance?	*[Is this answer confidential? No (delete as appropriate)]* While recognizing updates when significant changes occur is crucial (see Ofcom's chapter summary in pages 13 and 14), the current lack of specific criteria creates uncertainty for service providers.
Question 10.2: Do you agree with our proposal not to exercise our power to exempt specified descriptions of services from the record keeping and review duty for the moment?	[Is this answer confidential? No (delete as appropriate)] Response: To improve compliance with Ofcom's record-keeping and review duties (RKRDs), three key areas need attention: 1) In line with the accessibility and foreseeability principles of the ECtHR caselaw, for the UK Online

[1] If you have comments or input related the links between different kinds of illegal harm and risk factors, please refer to Volume 2: Chapter 5 Summary of the causes and impacts of online harm).

Question (Volume 3)	Your response
	Safety Act to be 'in accordance with the law' under Article 8(2) and 10(2) ECHR, clearer criteria for 'significant changes' triggering record updates are crucial to reduce uncertainty for service providers. Examples like major platform updates or changes in risk assessment methodology would provide actionable guidance. 2) Service-specific examples showcasing how different platforms implement RKRDs (e.g., social media vs. classified ads) would offer valuable practical context. 3) While not the main focus, elaborating on the potential consequences like fines or suspension for non-compliance (see to that effect Article 10(2) ECHR), can incentivize adherence and inform providers of their full responsibilities. Addressing these points will enhance clarity, understanding, and ultimately, compliance with RKRDs.

Question (Volume 4)	Your response
Question 11.1: Do you have any comments on our overarching approach to developing our illegal content Codes of Practice?	*[Is this answer confidential? No (delete as appropriate)]* On the whole, the overarching approach is sound. The level of flexibility proposed rather than seeking to impose a one-size-fits-all approach is welcome, given the wide range of different service-types that will come under the umbrella of regulated U2U services. The focus on balancing a realistic assessment of the risks involved against avoiding a chill on freedom of expression is welcome, as is the recognition of the interplay between expression and other rights such as privacy. (The Article 6 right to a fair trial might also be usefully built into this system – although contempt of court is not currently one of the primary harms identified, the internet does represent a particular challenge to the integrity of the justice process and the upholding of injunctions protecting the interests of individuals in the face of those determined to flaunt them – see, for example, the several instances in which individuals have sought to defy the injunction protecting the present identity and whereabouts of Jon Venables.)

Question (Volume 4)	Your response
	The suggested inclusion of a recognition of 'trusted flaggers' would be a helpful development. An organisation such as the Internet Watch Foundation could usefully contribute here. The categories of qualifying bodies would need to be addressed here(and preferably Ofcom-approved) in order to avoid abuse. The identification of "staff training and wellbeing" (at para 12.197) as something that should be the responsibility of the affected providers is important. This absolutely should be a part of the obligations placed on those expecting their staff to deal with the range of priority illegal content (content which can be so extreme that in several places the Consultation's drafters felt the need to include a trigger warning at the top of the relevant sections). It will also be imperative that moderation staff are trained in order to best identify problem content. This will include necessary language skills – not only in terms of being a sufficiently fluent speaker of the primary languages used on the platform, but also having a sufficient familiarity with commonly use, prejudicial terminology. A moderation system, whether human, automatic, or a mix of the two, which relies on an exclusively US-determined set of abusive terms for specific ethnic or religious communities, or the LGBTQIA+? community runs the risk of permitting abusive and problematic content that uses only British-specific terms of abuse, for instance. Effective training – and proper support for staff expose daily to a barrage of illegal and upsetting content – will be essential to making the whole approach work in practice.
Question 11.2: Do you agree that in general we should apply the most onerous measures in our Codes only to services which are large and/or medium or high risk?	*[Is this answer confidential? No (delete as appropriate)]* Yes, in principle. This is in accordance with the risk-based approach that the legislator has embraced in the Act. (We do not necessarily agree with the principles of this approach, as many of our members have argued in their research and submissions to the government and the Parliament. BILETA has also expressed reservations in this regards, but this is not subject of the present consultation)

Question (Volume 4)	Your response
Question 11.3: Do you agree with our definition of large services?	*[Is this answer confidential? No (delete as appropriate)]* Yes, this seems like a reasonable and quite nuanced approach, in line with that adopted in the EU DSA.
Question 11.4: Do you agree with our definition of multi-risk services?	*[Is this answer confidential? No (delete as appropriate)]* The definition of "large services" being those with a user base equivalent in size to 10% or greater of the UK population is an eminently sensible one. This very much captures the sort of U2U and search services that have raised the concerns that led to the Online Safety Act, and which have a very significant impact owing to their size and thus capacity to greatly amplify the spread of priority illegal content. It is eminently sensible to seek to cut off so far as is possible the most widespread distribution channels for CSAM, terrorist propaganda, false health information, and so on as a matter of priority over smaller reach services. That (as noted on Page 4 of Volume 4), this definition aligns with the approach taken in the EU's Digital Safety Act is not only desirable in terms of "reduc[ing] the potential burden of regulatory compliance for services", but vital if UK authorities wish to have any serious hope of the Online Safety Act and related regulatory regime to be taken seriously by these large services who are often not based in the UK, and apply a single set of policies to the European region. A region to which the UK remains an adjunct in the eyes of these entities that are not likely to be bothered with UK compliance so much if put in the position to choose the UK or the EU for setting their requirements.

Question (Volume 4)	Your response
Question 11.6: Do you have any comments on the draft Codes of Practice themselves?[2]	*[Is this answer confidential? No (delete as appropriate)]* The Codes are broadly in line with the Act's requirements. However, we do have the following concerns: - the effectiveness of hashing and URL removals and its impact on the freedom of expression, privacy and data protection (surely, our colleagues in computer science will provide more evidence on the adequacy of these measures overall), - measures introduced to safeguard user rights (freedom of expression, privacy and DP) are insufficient and do not protect user content from an ex-ante removal of legitimate content, which then needs to be contested by the user, - user redress and appeals recommendations are vague and do not offer sufficient details to providers that would ensure appropriate information about the content removal, especially regarding the use of proactive technology. We recommend that this process includes much more detail and mandate information about content removal, which would then enable effective complaints and appeals, - we believe that, overall, the codes need to include more detail on safeguarding user rights. This was a contentious aspect of the Act and the expectation was that Ofcom would engage in a more comprehensive analysis and offer more assurances to users regarding the protection of the their rights.
Question 11.7: Do you have any comments on the costs assumptions set out in Annex 14, which we used for calculating the costs of various measures?	*[Is this answer confidential? No (delete as appropriate)]* N/A (this is a very technical matter and we do not have expertise to comment on the costs. We shall leave this to

[2] See Annexes 7 and 8.

Question (Volume 4)	Your response
	other organisations, which have expertise in cost analysis and calculations.)
Question 12.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No (delete as appropriate)] On the whole, the overarching approach is sound. The level of flexibility proposed rather than seeking to impose a one-size-fits-all approach is welcome, given the wide range of different service-types that will come under the umbrella of regulated U2U services. The focus on balancing a realistic assessment of the risks involved against avoiding a chill on freedom of expression is welcome, as is the recognition of the interplay between expression and other rights such as privacy. (The Article 6 right to a fair trial might also be usefully built into this system – although contempt of court is not currently one of the primary harms identified, the internet does represent a particular challenge to the integrity of the justice process and the upholding of injunctions protecting the interests of individuals in the face of those determined to flaunt them – see, for example, the several instances in which individuals have sought to defy the injunction protecting the present identity and whereabouts of Jon Venables.) The suggested inclusion of a recognition of 'trusted flaggers' would be a helpful development. An organisation such as the Internet Watch Foundation could usefully contribute here. The categories of qualifying bodies would need to be addressed here(and preferably Ofcom-approved) in order to avoid abuse. The identification of "staff training and wellbeing" (at para 12.197) as something that should be the responsibility of the affected providers is important. This absolutely should be a part of the obligations placed on those expecting their staff to deal with the range of priority illegal content (content which can be so extreme that in several places

Question (Volume 4)	Your response
	the Consultation's drafters felt the need to include a trigger warning at the top of the relevant sections). It will also be imperative that moderation staff are trained in order to best identify problem content. This will include necessary language skills – not only in terms of being a sufficiently fluent speaker of the primary languages used on the platform, but also having a sufficient familiarity with commonly use, prejudicial terminology. A moderation system, whether human, automatic, or a mix of the two, which relies on an exclusively US-determined set of abusive terms for specific ethnic or religious communities, or the LGBTQIA+? community runs the risk of permitting abusive and problematic content that uses only British-specific terms of abuse, for instance. Effective training – and proper support for staff expose daily to a barrage of illegal and upsetting content – will be essential to making the whole approach work in practice.
Question 13.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No (delete as appropriate)] The focus, in the first instance, on U2U services which have significant reach (Twitter / X, or any of the Meta products, for instance) is to be welcomed, as it is the breadth of that reach which can significantly amplify the illegal harms identified by the Consultation. There may in time be a case where a smaller service presents such a large risk of a priority harm that it would be desirable to regulate further. Multiple such risks, however, would certainly render it appropriate to apply the proposed Codes to a smaller player. Where large U2U service providers are making very substantial sums indeed by exposing users to significant quantities of information (and data harvesting), it is more than reasonable to include a corresponding duty of care towards those users to the extent that all reasonable efforts to protect those users from illegal harms should be taken.

Question (Volume 4)	Your response
Question 14.1: Do you agree with our proposals? Do you have any views on our three proposals, i.e. CSAM hash matching, CSAM URL detection and fraud keyword detection? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? Yes / No (delete as appropriate)]* Broadly yes, however, academic working in computer science will provide Ofcom with evidence on these. Overall, concerns over false positives and negatives remain, as well as their effect on user rights.
Question 14.2: Do you have any comments on the draft guidance set out in Annex 9 regarding whether content is communicated 'publicly' or 'privately'?	*[Is this answer confidential? Yes / No (delete as appropriate)]* Yes, we broadly agree with the factors considered in the assessment whether content is to be considered communicated privately or publicly. Providers are able to refer to a number of factors and, equally, some useful examples of circumstances where this is contextual.
Question 14.3: Do you have any relevant evidence on: • The accuracy of perceptual hash matching and the costs of applying CSAM hash matching to smaller services; • The ability of services in scope of the CSAM hash matching measure to access hash databases/services, with respect to access criteria or requirements set by database and/or hash matching service providers; • The costs of applying our CSAM URL detection measure to smaller services, and the effectiveness of fuzzy	*[Is this answer confidential? Yes / No (delete as appropriate)]* *n/a*

Question (Volume 4)	Your response
matching[3] for CSAM URL detection; • The costs of applying our articles for use in frauds (standard keyword detection) measure, including for smaller services; and • An effective application of hash matching and/or URL detection for terrorism content, including how such measures could address concerns around 'context' and freedom of expression, and any information you have on the costs and efficacy of applying hash matching and URL detection for terrorism content to a range of services.	
Question 15.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? Yes / No (delete as appropriate)]* *n/a*
Question 16.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? Yes / No (delete as appropriate)]* *n/a*

[3] Fuzzy matching can allow a match between U2U content and a URL list, despite the text not being exactly the same.

Question (Volume 4)	Your response
Question 17.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No (delete as appropriate)] Yes, we agree with this proposal. it is significantly important that users are informed how services treat illegal content. We agree that the provisions included in the Terms of Services and Publicly Available Statements must be clearly signposted for the general public, should be comprehensible and written in plain English for the youngest person permitted to agree to them, they should be easily accessible and designed in an accessible way for users who may have different access requirements.
Question 17.2: Do you have any evidence, in particular on the use of prompts, to guide further work in this area?	*[Is this answer confidential? No (delete as appropriate)]* As noted in paragraph 17.47 of the Consultation, the evidence on the effectiveness of prompts in reducing harm to users is rather limited. Hence, further work is required to improve the effectiveness of prompts. To ensure readability, it is suggested that any Terms of Services and Publicly Available statements should be condensed with no longer than 200 words. It could also useful to have a summary page, explaining how services will treat illegal content. Finally, the use of prompts such as notifications via emails as well as the use of short videos explaining the Terms of Service for users that are not keen to read long texts could be helpful to ensure that users are informed of how services treat illegal content
Question 18.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No (delete as appropriate)] We generally agree with the proposals here, and recognise that grooming and a focus on grooming for the purposes of sexual abuse is incredibly harmful to children. That said, we share the concerns about reliance upon self-declarations for age, and the sole reliance upon this by services, especially given the wise-spread understanding that false declarations are common. While we understand the graduated approach for U2U services with implementing alter-

Question (Volume 4)	Your response
	native approaches to age verification, concerns surrounding the reliance on age verification generally persist. That said, as the proposals highlight in Volume 4, paragraphs 18.9 and 18.10 (at p.232), there is a distinct difference in the age ranges and cognitive abilities – measures should be tailored for those under 16, and those between 16-18. The proposals talk about children as users as a homogenous group, and seem to suggest that all children using a service will have the same capabilities or abilities to engage with the default functionalities, and the default support. This may not be the case. Children who are particularly vulnerable may not be in a position to engage with the settings in the manner the proposals suggest. We therefore question what provisions are envisaged to protect child users who are more vulnerable?
Question 18.2: Are there functionalities outside of the ones listed in our proposals, that should explicitly inform users around changing default settings?	[Is this answer confidential? No (delete as appropriate)] We do not foresee the requirement for additional functionalities outside of the listed proposals, other than in response to the concerns noted at 18.1. above. That said, it may be worth considering other, additional options such as time-restricted ability to receive direct messages for example, or for users of services to have been 'active' users for a certain calendar period before some functionalities become available. While there are no guarantees that these measures will reduce the risks of harm, it is possible to contemplate that some child users may benefit from additional periods of protected time to gain familiarity with default settings on services before their exposure is heightened by full feature access.

Question (Volume 4)	Your response
Question 18.3: Are there other points within the user journey where under 18s should be informed of the risk of illegal content?	*[Is this answer confidential? No (delete as appropriate)]* No. Although the general user journey and the risks of illegal content / online safety awareness could perhaps be better captured in schools and educational settings. The user journey should feature in this in light of the proposals. The user journey for 16–18-year-olds will necessarily be different – this should be noted in any design changes to default settings, particularly if reliance upon self-declarations for age remains a core part of the response as suggested in question 18.1.
Question 19.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? Yes / No (delete as appropriate)]* *n/a*
Question 19.2: What evaluation methods might be suitable for smaller services that do not have the capacity to perform on-platform testing?	*[Is this answer confidential? Yes / No (delete as appropriate)]* *n/a*
Question 19.3: We are aware of design features and parameters that can be used in recommender system to minimise the distribution of illegal content, e.g. ensuring content/network balance and low/neutral weightings on content labelled as sensitive. Are you	*[Is this answer confidential? Yes / No (delete as appropriate)]* *n/a*

Question (Volume 4)	Your response
aware of any other design parameters and choices that are proven to improve user safety?	
Question 20.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No (delete as appropriate)] Yes. Giving all users the option and ability to block and mute other user accounts is an important step in seeking to reduce some of the risks of illegal harm. It is important that users also have options to report other accounts – and should be offered this option when they are seeking to block and / or mute accounts. Blocking and muting (or similar options) should not be the only options available to users. This is particularly important where accounts are set up that are entirely fake / false, and which are designed to be used to perpetrate harm. Reporting these should also be part of the options available to users for a number of reasons, but predominantly so that if an account receives a number of reports, further investigation can be warranted swiftly.
Question 20.2: Do you think the first two proposed measures should include requirements for how these controls are made known to users?	*[Is this answer confidential? No (delete as appropriate)]* Yes. User empowerment and information is essential to a wider understanding of the options that are available for protection online. It is therefore integral to users safety to have controls made known to them. Leaving it to users to self-discover what their options are for such controls is the embodiment of a passive approach to online safety and reducing the exposure to online harms. This is particularly important for children who are born and will be living in an increasingly digitised and connected world – early understanding of user controls should be a core feature of their digital interactions.

Question (Volume 4)	Your response
Question 20.3: Do you think there are situations where the labelling of accounts through voluntary verification schemes has particular value or risks?	*[Is this answer confidential? No (delete as appropriate)]* There could be risk associated with minority groups and their anonymity on certain services. The labelling may expose these vulnerable individual/groups to abuse.
Question 21.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? No (delete as appropriate)]* In principle, we support the proposal that if a service has reasonable grounds to infer that a user account is operated by or on behalf of a terrorist group or organisation proscribed by the UK Government, services could remove a user account from the service. Nevertheless, taking into consideration the significant human rights implications of the proposed measure, particularly on freedom of speech, freedom of expression, freedom of assembly, these reasonable grounds should be clearly defined and should not be subject to the discretion of the services. Arguably, the ambiguity surrounding the term 'reasonable grounds' may encourage inconsistent take down policies by different services.
Question 21.2: Do you have any supporting information and evidence to inform any recommendations we may make on blocking sharers of CSAM content? Specifically: • What are the options available to block and prevent a user from returning to a service (e.g. blocking by username, email or IP address, or a combination of factors)? What are the advantages and disadvantages	*[Is this answer confidential? No* To prevent a user from accessing a service again, several methods can be employed, including blocking their usernames, email addresses, and/or IP addresses. However, this approach may not always be fool proof. As pointed out in the consultation, it's evident that the same users can circumvent these blocks by using different usernames, emails, or IP addresses. More importantly, as highlighted in the consultation, there is a real risk of automated systems incorrectly categorizing user content as Child Sexual Abuse Material (CSAM).

Question (Volume 4)	Your response
of the different options, including any potential impact on other users? • How long should a user be blocked for sharing known CSAM, and should the period vary depending on the nature of the offence committed? • There is a risk that lawful content is erroneously classified as CSAM by automated systems, which may impact on the rights of law-abiding users. What steps can services take to manage this risk? For example, are there alternative options to immediate blocking (such as a strikes system) that might help mitigate some of the risks and impacts on user rights?	In our view, a user who has been blocked for sharing known CSAM content should face a minimum block duration of 90 days. This appears to align with the practices of platforms like TikTok and YouTube. To mitigate the risk of mistakenly identifying lawful content as CSAM, service providers should establish a clear process that allows users to appeal these decisions. Chapter 16 of the consultation outlines an appeal process where content might have been wrongly classified as illegal. Paragraph 16.96 of the Consultation document suggests that all services should acknowledge receipt of complaints with an estimated timeframe for resolving them. While this proposal is welcome, the absence of a specific recommendation regarding an acceptable timeframe is a cause for concern. Given the lack of clarity on exact timeframes, there is a risk that some services may handle complaints with significant delays, resulting in inconsistencies across the board. Finally, we believe that implementing a strike system could potentially offer a better solution to mitigate the risks of erroneously classifying legal content as illegal.
Question 22.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	**[Is this answer confidential? No** As noted in the consultation document, predictive search functions can recommend search terms that might direct users toward harmful and potentially illegal content. In principle, we agree with the proposal that services should provide warnings in response to search requests in which the wording suggests that the users may be seeking to encounter CSAM. However, we have reservations about whether the measures outlined in the consultation will adequately tackle this issue. First, as noted in paragraph 22.52 of the Consultation, there are legitimate concerns whether such warnings defer users from seeking illegal content.

Question (Volume 4)	Your response
	Second, there is a potential for different search engines to implement diverse policies, which could impede the establishment of a consistent policy in this domain. To sum up, there is clearly a need for further work to ensure the effectiveness of this proposal.
Question 23.1: Do you agree that the overall burden of our measures on low risk small and micro businesses is proportionate?	*[Is this answer confidential? No (delete as appropriate)]* **[Is this answer confidential? No** The answer to this question depends on the nature of the business. Whilst this measure may be proportionate for some business it may not be proportionate for others. According to the consultation document, even small and micro-businesses need to have content moderation systems or processes in place to promptly remove illegal content. They must also establish an effective complaint procedure, enabling users to report cases of content being wrongly removed. It is worth noting that implementing these measures could entail substantial changes, which might pose a burden on some small and micro-businesses, particularly those lacking the necessary systems.
Question 23.2: Do you agree that the overall burden is proportionate for those small and micro businesses that find they have significant risks of illegal content and for whom we propose to recommend more measures?	*[Is this answer confidential? No (delete as appropriate)]* *Yes*

Question (Volume 4)	Your response
Question 23.3: We are applying more measures to large services. Do you agree that the overall burden on large services proportionate?	*[Is this answer confidential? No (delete as appropriate)]* Yes, we agree. Large services are likely to have the resources to implement these measures and it is very likely that they already have the necessary systems in place to remove illegal content and to deal with complaints.
Question 24.1: Do you agree that Ofcom's proposed recommendations for the Codes are appropriate in the light of the matters to which Ofcom must have regard? If not, why not?	*[Is this answer confidential? No* Ofcom's proposed recommendations for the Codes are commendable as they are quite detailed. Nevertheless, there are still quite a few measures in the code of practice that require fine tuning and further clarification. Below we draw on some specific examples which highlight that the code of conduct requires further elaboration. Paragraph 2(b) of the Schedule 4 of the Online Safety Act stipulates that measures described in the code of practice must be sufficiently clear and providers must understand what those measures entail in practice. For instance, both smaller services and large services should take down illegal content swiftly. However, the code of practice steers away from describing what is meant by swift stating that this is determined according to the circumstances of the case. In our opinion, the lack of concrete definitions in the code of conduct may lead to inconsistent practices. More importantly, any ambiguity in the code of conduct is likely to lead to legal disputes between users and services which may increase the workload of courts. Furthermore, paragraph 2(c) of the Schedule 4 of the Online Safety Act states that the measures described in the code of practice must be proportionate and be technically feasible. We have some concerns as to proportionality and feasibility of some of these measures. As noted above, proposals such as content moderation and reporting and complaints requirements could be very resource intensive for smaller services. To ensure a level playing field, smaller services should be supported and provided assistance to ensure compliance.

Question (Volume 5)	Your response
Question 26.1: Do you agree with our proposals, including the detail of the drafting? What are the underlying arguments and evidence that inform your view.	*[Is this answer confidential? No (delete as appropriate)]* Many of our members have expressed their concern around the "reasonable grounds to infer" standard during the Act's passage through the Parliament. This isn't a suitable test and it is hard for Ofcom to design suitable guidance to apply it. The judgements is highly contextual and would require a great legal expertise, ideally, a court/tribunal decision. However, the test is in the Act, so we will consider it here. Ofcom has drafted the ICJG to determine when there are reasonable grounds to infer that a piece of content is illegal. Providers can also draft their own terms and conditions "in such a way that at a minimum all content which would be illegal in the UK is prohibited on their service for UK users and make content moderation decisions based on their terms and conditions." Ofcom considers that "In practice we expect that many services will take the second of these approaches, or a hybrid approach." We are not convinced and we don't think that Ofcom has provided sufficient evidence for this contention. On the contrary, given the fines and the compliance concerns, we believe that most providers will follow Ofcom's Guidance. Ofcom compares the 'reasonable grounds to infer' to the 'beyond reasonable doubt' threshold used by the criminal courts. This test is actually better compared with the "manifestly illegal" test used for speech and similar. We are concerned that the way Ofcom perceived the test generally (and notes "When services make an illegal content judgement in relation to particular content and have reasonable grounds to infer that the content is illegal, the content must however be taken down."), that the test's application will result in over-removal and censorship, as noted by many of our members during the Bill consultations. We do not see this being addressed in the Guidance, we think that the detail there will inevitably result in over-removal of content, even more so, given the lack of free speech and privacy assurances, which we consider below.

Question (Volume 5)	Your response
Question 26.2: Do you consider the guidance to be sufficiently accessible, particularly for services with limited access to legal expertise?	*[Is this answer confidential? No (delete as appropriate)]* No, the Guidance includes many legal standards and language that would not routinely be well understood by services with limited access to legal expertise. More could be done to make the language simpler and more accessible (e.g. clearer definitions, simpler phrasing, more graphs and images etc.)
Question 26.3: What do you think of our assessment of what information is reasonably available and relevant to illegal content judgements?	*[Is this answer confidential? No (delete as appropriate)]* This is concerning, especially regarding the detail on taking defences into account and having regard for the protection of free speech, privacy and personal data. We support the following principle set out by Ofcom: "However, in this consultation we are not proposing that services should use any user behaviour monitoring technology and so we do not consider that information derived from such technology would be 'reasonably available'." We do think that, however, in practice, to comply with the Guidance, this is will be the exact result. In particular, the mention of user profile information and activity as some of the relevant factors, imply the use of similar technologies. That, couple with Ofcom's power to mandate proactive technology, raises serious concern of user privacy and data protection violations. We do not think that Ofcom's view that this data can be processed "only so long as this information is processed lawfully, including in particular in line with data protection laws" is sufficient, given the detail of requirements set out in the Guidance. We support the "'technology-agnostic approach' to reasonably available information and to illegal content judgements in general", but we do think that more clarity should have been brought to the issue of using monitoring tech. Further, there is not enough detail as to the defences, general defences seem to be dismissed quite lightly from a more comprehensive consideration in the proposal. Regarding specific offences guidance, we would like to note that we firmly disagree with the following: "26.277 As both these offences are new, they lack a body of case law or academic discussion on which Ofcom can draw

Question (Volume 5)	Your response
	for their interpretation. They are both also likely to be particularly difficult to identify in practice, because they depend heavily on context and on circumstances offline. – false communication academic community." The false communication offence as drafted in the Act may be new, but the academic community, including our members, have discussed false communications for a very long time now." (e.g. Leiser, Dr Mark, Reimagining Digital Governance: The EU's Digital Service Act and the Fight Against Disinformation (April 24, 2023). Available at SSRN: https://ssrn.com/abstract=4427493 or http://dx.doi.org/10.2139/ssrn.4427493; Urquhart, Lachlan, 'Regulation of Privacy and Freedom of the Press from 2004-2017: From Campbell to Fake News ' (January 23, 2016). in L Edwards Law, Policy and the Internet (Hart Publishing: Forthcoming), Available at SSRN: https://ssrn.com/abstract=2721044 or http://dx.doi.org/10.2139/ssrn.2721044 ; Mac Sithigh, Daithi, The Road to Responsibilities: New Attitudes Towards Internet Intermediaries (October 3, 2019). Information and Communications Technology Law, October 2019, Available at SSRN: https://ssrn.com/abstract=3463688 or http://dx.doi.org/10.2139/ssrn.3463688)

Question (Volume 6)	Your response
Question 28.1: Do you have any comments on our proposed approach to information gathering powers under the Act?	[Is this answer confidential? No (delete as appropriate)] While the proposed regulations for U2U and search service complaints processes demonstrate strengths in comprehensive, user-centricity, and harm reduction, potential areas for improvement exist. Addressing these weaknesses can significantly strengthen the framework and deliver more impactful protection for users. Firstly, timeframes for complaint resolution, beyond mere acknowledgement, should be specified. Ofcom's chapter summary in footnote 16 explains that 'services are free to takedown illegal content if this is a 'timely removal' of the content. However, different complaint types (e.g., illegal content vs. functionality issues) may warrant distinct timeframes

Question (Volume 6)	Your response
	to ensure efficiency and user satisfaction. For instance, a quicker takedown notice is needed for illegal live content. While the current 'notice and takedown' process seems too slow, it would be advisable to push for stricter definitions of 'expeditiously' (see to that effect Section 512 US DMCA and/or Article 5(1)(e) and 6(1)(b) EU DSA). Secondly, increasing transparency in content moderation decisions is crucial. Clear guidelines and potential avenues for human review beyond algorithms would foster trust and ensure fair treatment of users whose content might be flagged. Ofcom's chapter summary in pages 18 and 20 recognises that search services often use a 'combination of automated tools and human review' to moderate search content. However, this should also align with the UK Data Protection Act's right to human intervention and contest decisions (Article 22 UK DPA). Thirdly, safeguarding against potential abuse of dedicated reporting channels by 'trusted flaggers' is crucial. Strict eligibility criteria and monitoring flagger activity can prevent malicious use. Trusted flaggers, should be independent of law enforcement (e.g., City of London Police (ColP), National Crime Agency (NCA), National Cyber Security Centre (NCSC), and accountable and committed to human rights (refer to page 25 of Ofcom's chapter summary). Users deserve transparency regarding who flagged their content and why. Finally, considering the global nature of harmful content, regulations should clarify applicability beyond UK users and address privacy, data protection, freedom of expression, non-discrimination and due process concerns for complaint information (to that effect, refer to Article 8, 10, 14, 1 of Protocol no 12, 13 and 6 ECHR).
Question 29.1: Do you have any comments on our draft Online Safety Enforcement Guidance?	*[Is this answer confidential? No (delete as appropriate)]* The Guidance is broadly within the parameters, powers and duties of Ofcom set out in the OSA. We are concerned with the Priority Framework, however. In particular, we think that the strategic significance of addressing the alleged contravention should expressly include risks to user rights and freedoms specified in the Act. The way the Framework is drafted at the moment does not ensure proportionality will be exercised when it comes to enforcement in cases of where user rights are significantly impacted.

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

Individual signatories:

Prof. Guido Noto La Diega, Chair in Intellectual Property & Technology Law, University of Stirling

Dr Maureen Mapp, University of Birmingham

Dr Lisa Collingwood, Northeastern University