Your response

Question	Your response
Question 1: How do you measure the number of users on your service?	Confidential? – Y / N
Question 2: If your service comprises a part on which user-generated content is present and a part on which such content is not present, are you able to distinguish between users of these different parts of the service? If so, how do you make that distinction (including over a given period of time)?	Confidential? – Y / N
Question 3: Do you measure different segments of users on your service? • **Do you segment user measurement by different parts of your service? For example, by website vs app, by product, business unit.** • **Do you segment user measurement into different types of users? For example: creators, accounts holders, active users.** • **How much flexibility does your user measurement system have to define new or custom segments?**	Confidential? – Y / N
Question 4: Do you publish any information about the number of users on your service?	Confidential? – Y / N

Question	Your response
Question 5: Do you contribute any user number data to external sources/databases, or help industry measurements systems by tagging or sharing user measurement data? If not, what prevents you from doing so?	Confidential? – Y / N
Question 6: Do you have evidence of functionalities that may affect how easily, quickly and widely content is disseminated on U2U services? - **Are there particular functionalities that enable content to be disseminated easily on U2U services?** - **Are there particular functionalities that enable content to be disseminated quickly on U2U services?** - **Are there particular functionalities that enable content to be disseminated widely on U2U services?** - **Are there particular functionalities that prevent content from being easily, quickly and widely disseminated on U2U services?**	*Confidential? – ~~Y~~ / N* Refuge welcomes the opportunity to respond to this call for evidence. Our response is based on the experiences and insights of our specialist technology-facilitated domestic abuse team – which has supported survivors of this form of domestic abuse since 2017 and is the only such team in the country – as well as on our work on a model Violence Against Women and Girls (VAWG) Code of Practice. The appropriate categorisation of regulated services will be of vital importance to the success of the online safety regime, acting as a 'gateway' to the regulatory framework. However, we have not yet seen the proposed secondary regulations under Schedule 11 of the Bill setting out threshold conditions for categories of services, and this creates uncertainty. To provide context to our response, we firstly outline our views on the currently proposed approach to categorisation of services. Refuge is concerned that under the current Bill smaller, high-risk platforms will not be adequately captured in the categorisation thresholds. Platforms with smaller user numbers, including dedicated incel sites, may not be subject to Category 1 duties such as the 'triple shield' risk mitigation measures because of their smaller size, despite posing a significant risk to users. The perpetrator of the 2021 Plymouth shootings, Jake Davison, is known to have visited smaller incel forums after he was banned from Reddit in the days preceding the shooting. Whilst our Unsocial Spaces research found that the majority of tech-facilitated domestic abuse ('tech abuse') survivors experienced abuse from a partner or former partner on a Facebook (now Meta)-owned platform, which are among the largest platforms in terms of user base, we have supported women who have been subject to abuse on much smaller platforms

Question	Your response
	with a relatively small user base. Perpetrators do not discriminate by size of platform, and abusers will use any means to contact and harass survivors, including contacting the survivor across all her social media platforms. It is possible that, under the current scheme, perpetrators will increasingly turn to smaller sized platforms, knowing that these are subject to less stringent safety duties than the larger sites. In addition, we point to Glitch's Digital Misogynoir report which highlights concerns about smaller, high harm platforms such as 4chan and Gab, and how 'hateful rhetoric and jargon is trickling from the alternative platforms (Gab, 4chan) to the mainstream ones (Twitter [now X], Instagram, and Facebook)'. Glitch's analysis of 4chan and Gab deemed these sites to be some of the most 'toxic' in terms of misogynoir and noted that hateful jargon from alternative platforms steadily moves to more mainstream sites. We are therefore supportive of Baroness Morgan's amendment 245 to Schedule 11 of the Online Safety Bill which changes the determination of which sites would fall into Category 1 – moving the test from one of size "and" functionality, to size "or" functionality. This would give Ofcom more flexibility to decide which platforms should be in Category 1 and to determine how the risk assessment should be conducted. Turning to functionalities that enable content to be disseminated easily, quickly and widely on user-to-user services - and which should therefore be considered in categorisation thresholds - we wish to highlight the following: - Tools which allow **content to be forwarded and shared across different platforms**, such as share buttons. The enabling of quick sharing of content from one platform to another can facilitate the speedy dissemination of VAWG content such as intimate image abuse and doxing from one platform to another. It then becomes much harder for survivors to report each piece of content across multiple platforms. - In addition, features which **link up accounts held by a user across multiple platforms** can also enable content to be disseminated easily, quickly and widely. For example, when setting up a profile on a platform such as a dating site, there is

Question	Your response
	often an automated, or prompted, connection to the user’s other social media platforms. This builds a web of information on the user which can all be linked together and exploited by a perpetrator of domestic abuse to discover information about a survivor, for example, if she has fled. If multiple platforms are ‘linked,’ posting on one can often result in posting on other platforms, which can also spread abusive content more quickly. • **Recommender algorithms**, which can cause harm by a) promoting VAWG content, b) determining which content is pushed to users and suggest users follow or engage with groups/users/content that are misogynistic, and c) rewarding misogynistic influencers with greater reach. There have been concerns that the effect of recommender algorithms, especially in conjunction with autoplay, can prioritise extreme content, and therefore has a role in spreading online VAWG. In terms of functionalities that prevent content from being easily, quickly and widely disseminated on U2U services, we highlight the following features: • Tools employed by trained **platform moderators** such as removal of content, suspension and termination of abusive accounts and demonetisation of content. Moderators must be sufficiently trained in and develop a holistic understanding of VAWG and the way it impacts different minoritised groups in order to ensure that prompt action is taken in response to VAWG content. • **Manual review of amplified harmful content** by platform moderators i.e. of content this is being reshared by multiple users. • Self-destructing content may also be of some use in preventing content being easily disseminated. However, it should be noted that such content can also make evidence gathering difficult in a criminal investigation, and perpetrators can screenshot content to share onwards. Please also see our response to question 8.

Question	Your response
Question 7: Do you have evidence relating to the relationship between user numbers, functionalities and how easily, quickly and widely content is disseminated on U2U services?	Confidential? – Y / N
Question 8: Do you have evidence of other objective and measurable factors or characteristics that may be relevant to category 1 threshold conditions?	Confidential? – ~~Y~~ / N Evidence of additional factors or characteristics that may be relevant to category 1 threshold conditions may be found within the model Violence Against Women and Girls Code of Practice. We include below factors and functionalities that should particularly be considered in relation to threshold conditions: • **Location-sharing features.** Functionalities which enable, encourage or automatically push users to share their location can be incredibly dangerous for survivors of domestic abuse. Many social media apps utilise a user's location and some share that location with other users. For example, Snapchat's Snap Map allows users to see where their Snapchat contacts are, share their own current location and view Snaps from nearby-Snapchat users or users at a specific event or location. The Snap Map is highly accurate in pinpointing a user's location. If a survivor has fled to a secure location for their safety, it is possible for a perpetrator to track them via Snap Map. If the abuser has been removed from the survivor's contacts, they can create a fake account or ask friends or family who are still in a survivor's contacts to track her location. • **Self-destructing or disappearing content** such as messages, photos and videos. These features can easily be used by perpetrators to share abusive content that vanishes with a short time period. It can then be very challenging for survivors and the police to gather evidence of this abuse. This is a particularly attractive tool for abusers sharing intimate images without consent. Please also see our response to question 7. • **Ability to contact/message users that are not followers/friends**. Platforms which enable non-followers/friends to contact users pose a risk. Although survivors can remove perpetrators as

Question	Your response
	followers or friends and block them on platforms, perpetrators often set up new accounts to override these barriers and are then able to direct message the survivor, their friends or family to continue their abuse. We welcome Instagram's recent changes to direct messaging which limit the ability of users to engage with other accounts that have not accepted DM requests to chat, including only being able to send one DM and only being able to send text, rather than images, videos or voice notes. This change has the potential to limit harassment of women and girls online. • **Ability to tag other users in pictures, comments and posts, and tagged content then appearing on the tagged user's own profile**. This can compromise a survivor's safety, as a friend or family member may unwittingly reveal information about the survivor's new address/location in this way. Additionally, such features are used by perpetrators to tag survivors in abusive comments, posts and pictures. Some platforms have introduced options to prevent this, but there has been little promotion of this feature and this safety feature should be toggled off by default. • **'Mutual friends' and 'people you might like to follow' function.** Suggestions lists for other users to connect to are a feature of most major social media platforms. Due to algorithmic recommendations, the suggestions pushed to users present a risk for survivors of domestic abuse, particularly those that have created new and secret profile to prevent abusive and stalking, as the suggestions list could reveal a survivor's new profile to a perpetrator, particularly if they have friends in common or are based in a similar geographic location. • **Functions which notify contacts on messaging services when a user has changed their phone number.** On WhatsApp, this feature notifies contacts and/or group chats when a user has changed their number. This is an issue for survivors who have changed their phone number for security reasons to try to avoid/minimise technology-facilitated domestic abuse as it can inform an abuser of the survivor's new contact details and enable further tech abuse.

Question	Your response
	In addition, we are concerned with the recent announcement by Elon Musk that the block button will be removed from X. Blocking features are commonly used by victims of tech abuse and online abuse. Glitch and EVAW's Ripple Effect report found that, when asked if online abuse had led survivors to change their behaviour online, the most common behavioural change reported was blocking (76% of respondents reporting doing so).
Question 9: Do you have evidence of factors that may affect how content that is illegal or harmful to children is disseminated on U2U services? • **Are there particular functionalities that play a key role in enabling content that is illegal or harmful to children to be disseminated on U2U services?** • **Do you have evidence relating to the relationship between user numbers, functionalities and how content that is illegal or harmful to children is disseminated on U2U services?**	Confidential? – Y / N
Question 10: Do you have evidence of other objective and measurable characteristics that may be relevant to category 2B threshold conditions?	Confidential? – Y / N
Question 11: Do you have evidence of matters that affect the prevalence of content that (once the Bill takes effect) will count as search content that is illegal or harmful to children on particular search services or types of search service? For example, prevalence could refer to the proportion of content surfaced against each search term 16 that is illegal or harmful to	Confidential? – Y / N

Question	Your response
children, but we welcome suggestions on additional definitions. - **Do you have evidence relating to the measurement of the prevalence of content that is illegal or harmful to children on search services?**	
Question 12: Do you have evidence relating to the number of users on search services and the level of risk of harm to individuals from search content that is illegal or harmful to children? - **Do you have evidence regarding the relationship between user numbers on search services and the prevalence of search content that is illegal or harmful to children?**	Confidential? – Y / N
Question 13: Do you have evidence of other objective and measurable characteristics that may be relevant to category 2A threshold conditions?	Confidential? – ~~Y~~ / N We wish to raise here the reverse image search function available on some search engines such as Google and PimEyes. This has made it easier for perpetrators to search for additional information and other photos of the survivors. For example, based on a social media profile picture, an abuser may be able to track down images which provide information on survivors' current locations i.e. via photos on employer's websites. Conversely, reverse image search can be useful for survivors who want to track where their images have gone, such as if intimate images have been shared of them without consent. This feature should therefore be considered carefully when assessing threshold conditions for category 2a platforms. Another salient issue to highlight is the interconnection between search engines and users accounts or profiles (such as Google search engine and Google accounts). Searches made on the search engine will be saved to the user's profile as part of their data. This appears to be 'switched on' by default. For survivors who have had

Question	Your response
	their Google online accounts compromised by the perpetrator; this feature can enable an abuser to track what the survivor is searching for. If the survivor has been searching online for support with domestic abuse and help in fleeing, this creates a significant risk and could lead to an escalation of abuse.

Please complete this form in full and return to os-cfe@ofcom.org.uk.