UK Online Safety – OFCOM Call for Evidence

RESPONSE ON BEHALF OF META PLATFORMS INC

Meta Platforms, Inc. ("**Meta**", "**us**", "**we**" or "**our**") would like to thank Ofcom for the opportunity to provide information in response to its third call for evidence ("**CFE3**"). Over the last six years, Meta and affiliated entities have supported the UK Government's development of the Online Safety Bill ("**OSB**") by providing detailed information on a wide range of subjects through evidence sessions, written submissions (most recently in response to Ofcom's first call for evidence ("**CFE1**") of 6 July 2022 and second call for evidence of 10 January 2023 ("**CFE2**")), ministerial discussions, and multi-stakeholder in-person technical sessions.

At Meta, we have nineteen years' experience in tackling online safety issues through establishing policies, building tools and technologies, and producing guides and resources, all in partnership with experts both within and outside our company. The information provided below represents a snapshot of Meta's current practices, the policies and features of which we regularly review. Please see our separate response dated 12 September 2023 which covers our WhatsApp service. Meta is pleased to use its experience to contribute to Ofcom's work as the forthcoming regulator for the OSB and looks forward to continuing to engage with Ofcom on this and other issues relating to the OSB.

1   How do you measure the number of users on your service?

- **Who counts as a user for those purposes? For example, are users defined as individuals who visit the service, individuals who interact with the service, individuals who have an account with the service, or individuals who request the service to host information. If you measure users on your service in more than one way, please include all relevant approaches.**
- **With what frequency do you measure the number of users on your service? For example, monthly.**
- **Do you de-duplicate the number of users in your measurements? If so, how?**
- **Do you measure different geographic units in relation to numbers of users? For example, country, region, market, global, city.**
- **Do you break down the number of users on your service by their age? For example, child users.**

1.1   Our mission is to give people the power to build communities and bring the world closer together. We build technology that helps people connect and share, find communities, and grow businesses. Meta reports in its Form 10-K and Form 10-Q, filed with the U.S. Securities and Exchange Commission annually and quarterly respectively, an estimate of the number of users across its "Family" of apps (Facebook, Messenger, Instagram and WhatsApp) and the number of users of its Facebook app.[1] These estimates are our key community metrics representing Meta's measure of the number of users of those particular technologies in that measured period. Pursuant to Article 24(2) of the Digital Services Act (the "**DSA**") every six

---

[1] Meta Platforms, Inc. is a publicly listed company in the United States and is regulated by the United States Securities and Exchange Commission. As a listed company, Meta Platforms, Inc. is required to file the Form 10-K annually and the Form 10-Q quarterly.

months, Meta Platforms Ireland Limited ("**MPIL**") also provides information on the number of average monthly active recipients of Facebook and Instagram in the European Union.

1.2 These metrics are calculated using internal company data based on the activity of user accounts. While these numbers are based on what we believe to be reasonable estimates for the applicable period of measurement, there are inherent challenges in measuring usage of our technologies across large online and mobile populations around the world. As such, we are continually seeking to improve our estimates and such estimates may change due to improvements or changes in our methodology.

1.3 Our key metrics for measuring the size of our global active community of people using our technologies and their underlying methodologies are set out in more detail below.

User metrics

1.4 As outlined below, Meta's key community (or user) metrics consist of (i) Facebook metrics, and (ii) Family metrics:

(i) *Facebook metrics*

1.5 We have historically reported in our Form 10-K and Form 10-Q on the numbers of our 'Daily Active Users' and 'Monthly Active Users' (together, our "**Facebook metrics**"), which measure user activity on Facebook and Messenger:

1.5.1 **Daily Active Users ("DAUs").** We define a daily active user as a registered and logged-in Facebook user who visited Facebook through our website or a mobile device, or used our Messenger application (and is also a registered Facebook user), on a given day. We view DAUs, and DAUs as a percentage of MAUs, as measures of user engagement on Facebook.

1.5.2 **Monthly Active Users ("MAUs").** We define a monthly active user as a registered and logged-in Facebook user who visited Facebook through our website or a mobile device, or used our Messenger application (and is also a registered Facebook user), in the last 30 days as of the date of measurement. MAUs are a measure of the size of our global active user community on Facebook.

(ii) *Family metrics*

1.6 Given the size of our community and the fact that many people are using more than one of our technologies, starting from our annual report on the Form 10-K for the year ended 31 December 2019, we also report estimates of the numbers of our Daily Active People and Monthly Active People (collectively, our "**Family metrics**"). For our Family metrics, we do not count the total number of user accounts across our apps because we believe that would not reflect the actual size of our community. Rather, our Family metrics represent our estimates of the number of unique people using at least one of Facebook, Instagram, Messenger, and WhatsApp (collectively, our "**Family**" of apps) during the applicable period of measurement.

1.6.1 **Daily Active People ("DAP").** We define a daily active person as a registered and logged-in user of one or more Family apps who visited at least one of these Family

apps through a mobile device application or using a web or mobile browser on a given day. We view DAP, and DAP as a percentage of MAP, as measures of engagement across our apps.

1.6.2 **Monthly Active People ("MAP").** We define a monthly active person as a registered and logged-in user of one or more Family apps who visited at least one of these Family apps through a mobile device application or using a web or mobile browser in the last 30 days as of the date of measurement. We view MAP as a measure of the size of our global active community of people using our apps.

DSA Reporting

1.7   As explained above, aside from the Form 10-K and Form 10-Q reporting, pursuant to Article 24(2) of the DSA, every six months we provide information on the average number of monthly active recipients of Facebook and Instagram in the European Union over a 6 month period.

1.8   For these purposes:

1.8.1   For Facebook, we define a monthly active recipient as a registered and logged-in Facebook user who visited Facebook through our Facebook website or a mobile device in the last 30 days as of the date of measurement.[2]

1.8.2   For Instagram, we define a monthly active recipient as a registered and logged-in Instagram account who visited Instagram through our Instagram website or a mobile device in the last 30 days as of the date of measurement.

1.9   These metrics are published on our Transparency Center. The report for August 2023 is available here: Digital Services Act - Information on Average Monthly Active Recipients in the European Union - 14 August 2023 (fb.com).

Frequency of measurement

1.10   The Form 10-K and Form 10-Q are filed annually and quarterly, respectively. The metrics included in those submissions cover the previous quarter and certain prior periods. Pursuant to Article 24(2) of the DSA, MPIL reports every six months on the average monthly active recipients of Facebook and Instagram during the applicable six month measurement period.

De-duplication

1.11   Meta works hard to ensure that our user metrics are as accurate as possible. All of the metrics that we release publicly are subject to extensive exhaustive internal review prior to publication. However, these metrics are necessarily estimates and, as Meta publicly

---

[2] Unlike our Facebook metrics and Family metrics, this metric does not include those who used (only) our Messenger application.

discloses, it is possible that the actual numbers of unique people using Meta's technologies may vary significantly from our estimates, potentially beyond our estimated error margins.

1.12   As explained above, while the numbers are based on what we believe to be reasonable estimates of our user base for the applicable period of measurement, there are inherent challenges in measuring usage of our technologies across large online and mobile populations around the world.

1.13   By way of example, in relation to our MAU figures as included in our Form 10-K and Form 10-Q, we regularly evaluate to estimate the number of "duplicate" accounts. A duplicate account is one that a user maintains in addition to his or her principal account. To identify duplicate accounts we use data signals such as identical IP addresses and similar user names. However, we note that duplicate accounts are very difficult to measure at our scale.

Geographic location

1.14   We report our Facebook metrics quarterly at the global level and for the following regions: US & Canada, Europe, Asia-Pacific, and Rest of World. We also provide information on our user numbers for Facebook and Instagram in the European Union every six months pursuant to the requirements of the DSA, as explained above. We do not typically report user numbers by geographic location beyond that. Our data regarding the geographic location of our users is estimated based on a number of factors, such as the user's IP address and self-disclosed location. These factors may not always accurately reflect the user's actual location. For example, a user may appear to be accessing Facebook from the location of the proxy server that the user connects to rather than from the user's actual location.

User ages

1.15   Meta and MPIL do not "*break down the number of users on [its] service[s] by age*" for the purposes of reporting in the Form 10-K and Form 10-Q or the DSA reporting pursuant to Article 24(2) of the DSA. Determining the age of people on social media is a complex challenge across our industry, which makes breaking down the number of users on our service by their age very challenging. Meta seeks to estimate the age of the users on our technologies by using: (1) the date of birth they provide when signing up to our service and (2) our age modelling technologies. However, Meta's focus is on seeking to identify people's ages to try to aid with compliance with the applicable terms of service/terms of use of Meta's technologies and to seek to provide age-appropriate experiences.

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)?**

N/A

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?**

3.1 As set out in our response to question 1 above, our Facebook metrics measure user activity on Facebook and Messenger but not on our other services, and our Family metrics represent our estimates of the number of unique people using at least one of our Family of apps during the applicable period of measurement. Meta has developed these key metrics over a period of time and, as such, they reflect what we believe to be the most appropriate and reliable method of measuring the number of individuals using our services.

3.2 Our metrics focus on global user numbers, with some segmentation for regional Facebook metrics (which is publicly reported). Meta does have some flexibility to measure certain different parts of our services or different 'segments' of users on our service and to define new or custom 'segments' for measurement. For example, Meta receives some data about how users access Meta services (e.g. via mobile device or website). However, designing and validating new metrics which accurately measure our services is generally a complex process, involving significant cost and time implications. Further, there are significant commercial and legal sensitivities related to the reporting of our metrics. As such, for the purpose of external reporting, we generally limit measurement methodologies to our key metrics, which are those which senior management focus on to manage and monitor business performance, as they reflect the results of years spent stress testing and developing methodologies which we believe are based on reliable data.

4   Do you publish any information about the number of users on your service?

4.1 We provide details of our Facebook and Family metrics, our number de-duplication methodologies and some information about the geographic location of our users quarterly in the Form 10-Q and annually in the Form 10-K pursuant to the requirements of the US Securities Exchange Act of 1934. Pursuant to Article 24(2) of the DSA, we also publish information on the average monthly active recipients of the Facebook and Instagram services provided by MPIL in the European Union every six months.

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?

5.1 Whilst we provide certain qualitative data to external parties, we do not generally contribute non-public user number data to external sources/databases, nor do we generally tag or share user measurement data for any industry measurement systems. There are four key reasons for this.

5.2 First, unless otherwise required, we limit disclosing detailed, non-public information about our user numbers given the significant resources and time we have spent developing and refining our measurement methodologies. Determining the number of people on social

media is a complex challenge across our industry and our own methodologies are based upon complex and valuable intellectual property which are commercially sensitive to Meta.

5.3   Second, contributing this data to other sources/databases would result in inconsistent publicly available information surrounding the number of people using our services. Whilst we provide a description of our measurement methodologies in our annual 10-K forms, this information is limited to a high level overview rather than detailed statistics. Our Family metrics and Facebook metrics estimates therefore differ from estimates published by third parties due to differences in methodology.

5.4   Third, we have wide ranging commercial and legal obligations which restrict our ability to disclose sensitive information to third parties, including user number data.

5.5   Fourth, sharing this information with external third parties poses a risk to our purposes of promoting the safety, integrity, and security of our services, particularly where these relate to the safety and well-being of teen users. For instance, to reduce the risk that individuals could "game the system", we do not provide or share the predictions made by age models with any third parties. This risk extends to the disclosure of other number measurement methodologies.

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?**

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?**

7.1   In relation to questions 6 and 7, all U2U services allow content to be shared between users, so the nature of the functionalities of different U2U services will affect how "*easily, quickly and widely*" content will reach other users. Accordingly, functionalities exist in each of our technologies which can affect how "*easily, quickly and widely*" content is shared. For example, whether an individual chooses to make a post on Facebook public or private can affect how widely that post can be shared, if at all. Likewise, whether an individual has a public account or a private account on Instagram can affect how widely their posts can be shared, if at all.

7.2   We also use AI to make sure people see content that they will hopefully find relevant and valuable. These AI systems can affect how "*easily, quickly and widely*" content is shared. For example, the posts that users see on their Facebook Feed are ranked based on what we believe will be most valuable to each user. We use an algorithm to determine the order

of all of the posts the user could see. We then seek to prioritise the most meaningful posts at the top of the Feed and we use Feed ranking to reduce the distribution of posts that may contain content that people find objectionable, but that doesn't necessarily meet the bar of removal under our policies. We also strive to remove content from Facebook altogether when it violates our policies. As part of Meta's commitment to transparency, we have shared 22 'system cards' that contain information on how our AI systems rank content and ways in which users can customise their specific AI-powered experiences. More information is available on our blog here: <u>Introducing 22 system cards that explain how AI powers experiences on Facebook and Instagram (meta.com)</u>.

8 Do you have evidence of other objective and measurable factors or characteristics that may be relevant to Category 1 threshold conditions?

8.1 A key factor relevant to Category 1 threshold conditions should be whether the U2U services are provided on public platforms or via private messaging, given private messaging is very different in nature from public social media, and as such is more akin to SMS or email, which are currently excluded from the scope of the OSB.

8.2 If private messaging services are designated Category 1 services, they could be made subject to the same obligations that will be placed on public social media sites of that category. Such obligations would fundamentally change the nature of some private messaging services and would make compliance extremely challenging. Under the OSB, as it is currently drafted, providers of private messaging services could be required to comply with onerous obligations, many of which do not make sense in the context of one-to-one or private group conversations and practically speaking it is unclear how they could be met in the context of private messaging. Private messaging is not the same as public social media. The OSB also does not take account of end-to-end encrypted services, the fundamental premise of which is that only the sender and intended recipients of a message can know the contents of that message. Attempting to apply these obligations to end-to-end encrypted messaging services risks people's private messages being constantly surveilled and censored for legal but harmful content.

8.3 Further, the threshold conditions must be future-proofed, and sufficient flexibility should be incorporated to account for new technologies and changes and developments to the way that categorised platforms are used and operate. For example, to the extent that new functionalities are launched on Category 1 platforms, there should be flexibility within the threshold conditions for these to be excluded.

8.4 Please also see the response to Question 8 of WhatsApp's response to CFE3.

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?**

9.1   Meta has globally applicable Community Standards (for Facebook) and Community Guidelines (for Instagram) in place which outline types of content or behaviour that are not allowed on Facebook and Instagram. In many cases, our Community Standards and Community Guidelines overlap with common areas of illegality (e.g. hate speech and child exploitation), even though they do not purposefully map to specific laws as laws vary significantly across countries and have many nuances. Our content moderation efforts are more effective when they can develop and work at scale.

9.2   Our content moderation ecosystem supports and enforces the Community Standards and Community Guidelines. This ecosystem is operated through a combination of human review and the use of technology. Meta has in place extensive technologies, systems and controls to find and remove content that violates our policies, sometimes before it is disseminated. We are constantly improving and developing new technologies to stay ahead of bad actors, and it's why we have made some of our technologies to fight abuse available at no cost to other companies.

9.3   We also prioritise how to prevent harm from happening in the first place. Accordingly, we have also implemented wide-ranging restrictions and controls on many of our functionalities to prevent content which violates our Terms or policies from being disseminated on our services. For example, we use a combination of strong default protections for young people, alongside technologies like machine learning to identify and address potentially malicious activity.

9.4   Please see Meta's response to CFE2 for a more detailed description of these technologies, systems and controls as well as the restrictions on our functionalities. For ease of reference, we have listed below those restrictions and controls placed on our functionalities which we believe to be the most relevant to this question:

9.4.1   **Restricting interactions between adults and teens:** We have taken steps to make it more difficult for adults to find and follow teens. On Instagram, for example, we have developed new technology that helps us to find accounts that have shown potentially suspicious behaviour and seek to limit those accounts from interacting with young people's accounts. We also work to restrict people over 19 years old from sending private messages to teens who don't follow them (this feature relies on our work to predict peoples' ages using age modelling technology and the age people give us when they sign up to our services).

9.4.2   **Photo-matching technology:** We use photo-matching technologies to help detect, remove and report the sharing of images and videos that exploit children. These photo-matching technologies create a unique digital signature of an image (known as a "hash") which is then compared against a database containing hashes of previously identified illegal images to find copies of the same image. Meta also runs these technologies on links from other internet sites shared on the Facebook and Instagram services and their associated content, to detect known child exploitation housed elsewhere on the internet. Not only does this help keep the Facebook and Instagram services safer, but it also helps keep the broader internet safer as all violating content is reported to the US National Center for Missing and Exploited

Children, which then works with appropriate law enforcement authorities around the world.

9.4.3 **Improving detection capabilities:** In addition to photo-matching technology, we use machine learning to proactively detect potentially child exploitative content when it is uploaded to the public surfaces of our Facebook or Instagram services. We use these technologies to more quickly identify this content and report it to the relevant authorities, and also to find accounts that engage in potentially inappropriate interactions with teens on our Facebook and Instagram services, so that we can remove these accounts from our services. We have also developed functionalities to detect and remove potential bullying content even before it is reported.

9.4.4 **Updated reporting tools:** According to a study from Thorn, children who experience something negative online are more inclined to use in-app safety tools than to seek help offline. After consultations with child safety experts and organisations, we made it easier to report content for violating our child exploitation policies on our Facebook and Instagram services.

10 Do you have evidence of other objective and measurable characteristics that may be relevant to Category 2B threshold conditions?

10.1 Please see our response to question 8 above.

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 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?**

N/A

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?**

N/A

13 Do you have evidence of other objective and measurable characteristics that may be relevant to Category 2A threshold conditions?

N/A