Meta Responses To Questions on Notice

Questions on Notice from the Hearing (Transcript)

1. **What, if any, advice you give to Australian users about the scraping of their data. How many images and posts have actually been taken from Australian users and fed into the model? I'll give you the opportunity to provide what other protections are in place for people's data**. *(Senator Shoebridge, p4)*

Before turning to the controls and protections that are available to the people who choose to use our services, we welcome the opportunity to clarify several assumptions in the question.

Firstly, generative AI models are not static databases of the underlying training data. Instead, models learn from the training data.Training a GenAI model such as a large language model (LLM) is similar to building any large statistical model – such as a meteorological model used to predict the weather – and involves deriving patterns and relationships from a very large and diverse set of existing data.

Secondly, we only use publicly available data to train our models and this also includes data on our services. Specifically, we use public posts and comments on Facebook and Instagram to train generative AI models for these features and for the open-source community. We don't use posts or comments with an audience other than public for these purposes. For Facebook profiles, for example, the default privacy setting is private and so this means that people must choose to share publicly.

Thirdly, the term “scraping” typically refers to the collection of third party data without appropriate permissions. On Meta’s services, we take a range of measures to prevent the scraping by third parties of the content that people have chosen to share on our services.

Turning to our work to protect people’s privacy and provide them with controls and information, we take a layered approach:

Information we provide users

Users are given transparency around the use of their public information to train Meta's AI systems across a number of touchpoints - including in policies and terms, Help Center and Privacy Center articles, and in-app disclosures.

Meta Privacy Policy: The Meta Privacy Policy makes it clear to users that the information they share with Meta will be used to deliver and improve Meta's products and services. This includes the AI technologies that underpin these products and services. The Privacy Policy also includes specific disclosure about the broad availability and use of content which users choose to share publicly on Meta's platforms.

Help Center and Privacy Center Articles: Users were also provided with more specific information in Help Center articles, including:

- A dedicated "Generative AI at Meta" Privacy Center Guide, linked from the Privacy Center homepage (**Generative AI Guide**).
- A more detailed Help Center article focused on the data used for training, entitled 'How Meta uses information for generative AI models and features' in the Meta Help Center. This Article is linked from the Generative AI Guide.
- Meta also has a dedicated Help Center article focused on Public Information on Facebook.

In-App Experience and Meta AI Terms: Additionally, when a user first uses the Meta AI chatbot on Facebook, Messenger, Instagram, or WhatsApp, they go through a new-user experience (NUX) that informs them that their messages may be used to improve AI at Meta, links users to the AI Terms, and indicates that use of Meta AI constitutes agreement to those terms.

Protections for users

At a high level, Meta has taken several steps to mitigate privacy risks arising from the training and use of our AI models and tools.

At the training stage, Meta takes steps to limit the volume of personal information being included in training data by, for exampleIn relation to data from a user's interaction with our AI products and features, prior to use for model training, we remove any Meta unique system identifiers, such as Meta User ID and username string. Removing identifiers is designed to ensure that this data is not associated with a user account. We also use human review to check for the presence of other categories of potentially identifying data, such as private email addresses, private phone numbers and national identifiers. In relation to other user data from our

platform, such as public posts and comments, we remove Meta unique system identifiers.

In relation to outputs generated by Meta's generative AI models, Meta takes steps including:

- Meta has a set of constantly evolving internal and confidential policies that outline acceptable and unacceptable LLM responses across a broad range of categories. These policies aim to preserve privacy by precluding outputs of certain personal information, such as highly sensitive information, or information from a non-public source.
- Training and fine-tuning our generative AI models to limit the possibility of personal information that users may share with our generative AI products (e.g., our Meta AI assistant) from appearing in responses to other people.
- Using a combination of human reviewers and automated technology to review model outputs so we can reduce the likelihood of outputs sensitive including personal information (as well as improve product performance).
- To help ensure the efficacy of our safeguards, we perform accuracy assessments where the AI Chatbots are prompted with queries on acceptable and unacceptable LLM responses across a broad range of categories to ensure they are performing as expected, and conduct query monitoring, which provides notification of an unusual increase in queries on on acceptable and unacceptable LLM responses which would prompt a product team investigation aimed at addressing the matter, where necessary. Safety evaluations and red teaming are also performed before each LLM release (usually every few weeks) for the AI Chatbots.

2. **Could you explain what CrowdTangle does? Can you explain why Meta decided to shut down CrowdTangle?** (*Senator Shoebridge, p4-5*)

CrowdTangle was a public insights tool from Meta to explore public content on social media. As of 14 August 2024, CrowdTangle is no longer available.

We are now rolling out the Meta Content Library and Content Library API in Australia to provide useful, high-quality data to researchers. Meta Content Library was designed to help us meet new regulatory requirements for data sharing and transparency, while meeting Meta's rigorous privacy and security standards.

We phased out CrowdTangle to allow our engineering teams to focus on developing MCL. CrowdTangle was hard to maintain because it was built on an outdated codebase, and it was an acquired third-party product, so it didn't integrate directly with our other systems. This made it difficult to add new data types, and decreased its reliability for fact-checkers. Our new tools are designed to quickly scale to new datasets like Threads, enabling fact-checking to expand to more public data.

During the horrific events in the UK, we immediately established a dedicated team that worked to help identify and remove content that broke our rules, in addition to our safety systems which operate 24/7. We removed hate speech, threats of violence, links to external sites that were used to coordinate rioting, and content supporting individuals and organisations who are banned from our platforms. We worked with independent fact-checkers, and proactively added warning labels to content containing misinformation and made it so that people were less likely to see it. We were in contact with law enforcement and supported them in every way we could. We are not aware of any use of CrowdTangle to identify mitigation measures in response to the UK riots. We worked with relevant authorities and local partners to respond to questions and concerns in connection with these issues.

3. **According to AEC data, Meta's given $40,000 to the Labor Party over the last decade but has given no money to my side of politics. You are a major distributor of news. Does the fact that you give money to one side of politics illustrate your political bias against the conservative side of politics?** (*Senator McGrath, p5*)

   As the AEC Transparency Register indicates, Meta has not donated to any Australian political party. Nor are we a major distributor of news. Our services connect people with the people and topics that matter to them – mainstream media news makes up a very small proportion of people's experience of our services. In any event, as our policies confirm, to ensure that everyone's voice is valued, we take great care to create standards that include different views and beliefs.

4. **If I made a post private but then it was shared by others on the same platforms, in a public sense, it would then become public and be used to train your models, wouldn't it? … [I]f someone shares it on, it then becomes public, doesn't it? … It's a data capture source as a personal thing, so there is absolutely no way that somebody gets something that I mark as private and then shares it onwards that can't be used by your model. That's what you're saying?** (*Senator Darmanin, p6*)

   We appreciate the focus on how private posts stay private. By way of background, we wish to reiterate our remarks made during our appearance and through other public statements. We offer generative AI features that you can use to do things like get responses to questions and create images. We do not use posts or comments with an audience other than public for these purposes. For Facebook profiles, for example, the default privacy setting is private and so this means that people must choose to share publicly.

The person who shares a post on our services controls the audience for that post. If someone reshares that post, they aren't able to share that post through Facebook with people who are not in the audience that was originally selected to share with. Only the people who could see those posts when it was first made are able to see it when someone clicks “Share.” More details are included here.

Additionally, generative AI models are not static databases of the underlying training data. Instead, models learn from the training data.

5. **Do you have a way of sifting through content that has been uploaded since someone's 18th birthday so that their historical data can't be shared when it was in fact posted prior to their 18th birthday?** (*Senator Darmanin, p7*)

   We use only public posts and content for training if a person who is choosing to use our services is 18+. The default setting for Facebook profiles and for under 18 year olds is private and so they would need to take a proactive step to change this setting.

6. **Is it possible that other AI developers could have scraped the profiles of Facebook and Instagram users to train their models, including for uses like the creation of deepfakes or even deepfake sexual material? Have you developed some way of preventing this from happening?** (*Senator Darmanin, p8*)

   Meta’s Terms of Service state that users must not access or collect data from our Products using automated means (without our prior permission) or attempt to access data you do not have permission to access.

   We devote substantial resources to combating unauthorized scraping on Facebook products. We have a dedicated Anti-Scraping Team, including data scientists, analysts, and engineers focused on our efforts to detect, block, and deter scraping.

   Because scrapers mimic the ways that people use our products legitimately, we will never be able to fully prevent all scraping without harming people’s ability to use our apps and websites the way they enjoy. That means that we have to try to strike the right balance and rely on a variety of approaches to address scraping. Since scraping is both a common and complex challenge to solve, we take a more holistic approach to staying ahead of it. In short, we aim to make it harder for scrapers to acquire data from our services in the first place and harder to capitalize off of it if they do.

The first way we aim to make scraping more difficult is through the use of rate limits and data limits. Rate limits cap the number of times anyone can interact with our products in a given amount of time, while data limits keep people from getting more data than they should need to use our products normally.

Limits are only a first layer of protection, and we know that scrapers are determined to find new ways to get data. That's why we've also focused on developing other methods of identifying and deterring scraping. One example is that we look for patterns in activity and behavior that are typically associated with automated computer activity and stop it.

Our Anti-Scraping team also investigates suspected scrapers to learn more about what they're doing and make our systems stronger. We've taken a variety of actions against data misuse. These can include sending cease and desist letters, disabling accounts, filing lawsuits against scrapers engaging in egregious behavior and requesting companies that host scraped data to take them down.

7. **First, what proportion of Australians exercise any opt-out ability on the Facebook profile? If there's not actually an opt-out option, which may flow from Senator Sheldon's question, what proportion of Australian Facebook users' data—that is, information, photographs, posts—are set as publicly available or available only to friends, only to select friends or privately held? If that's not broken down by Australian users—though I hope it would be—then I'm happy to put that question across all Facebook users. How many or what proportion of European Union Facebook users have opted out of allowing their information to be used to train AI models?** (*Senator Ghosh, p11*)

It is important to note that the default setting for new posts on Facebook is "Friends Only." This means that Australians need to proactively adjust their posts' audience to public for that post to be used for model training.

On Facebook, users can adjust their privacy settings at any time using the "Audience and visibility section" of their settings. In this section they can make universal changes to the audience setting of posts, for example. However, users also have the ability to set the audience for each piece of content that they publish. They do this by clicking the "audience selector" when making a post to select who they would like to share their post to. This means that individual Australian users may have some content that is set to public and some content that is set to private.

For Instagram, audience settings are controlled at an account level. As at 13 July 2024, approximately 35% of Australian Instagram users have private accounts and the

remainder have public accounts. Every new user in Australia with a stated age of under 16 years old is defaulted into a private account when they register for Instagram and is defaulted into more private settings when they register for Facebook.

It is also important to note that generative AI models are not static databases of the underlying training data. Instead, models learn from the training data.Training a GenAI model such as a large language model (LLM) is similar to building any large statistical model – such as a meteorological model used to predict the weather – and involves deriving patterns and relationships from a very large and diverse set of existing data.

8. **What amount of time, on average, is spent by Facebook users on the privacy policy page? What amount of time is spent, on average, by Facebook users on the terms and conditions or updates page when they have to click 'yes'?** (*Senator Ghosh, p11*)

   We do not have those statistics available to share with the Committee in the time available to respond to these Questions on Notice. However, to assist the Committee in its work, we wish to note that rather than relying on a single policy to provide transparency to our users, Meta provides users with a range of information about our data practices, including generative AI and privacy. This includes sharing key insights in the Privacy Centre, including with respect to GenAI.

   We have also invested in research and design to ensure our Privacy Policy, terms, and privacy-related information is as clear, simple and easy to understand as possible for users. For example in 2022 we updated the Meta Privacy Policy to more clearly state our data practices at an accessible reading level. It includes more fulsome information to bring to life how we work with third parties, how we share data across the company, how we use data for advertising, and how our security and integrity systems work. This update generally offers greater transparency to our users and more specifically narrates our current practices in this baseline, foundational text.

9. **Could you provide data … of the number of users that you know are ages 12, 13, 14, 15 and 16? Could you provide a breakdown of the number of users for your platforms? … . If you could provide the data, I'd really appreciate it. I know it may be subject to some commercial constraints, so as much insight as you can give us would be very useful. If there was any other insight about the number of interactions and who they interact with, that would be really helpful**. (*Senator Ghosh, p12*)

   We are currently working on our response to the eSafety Commissioner's request for information under the Basic Online Safety Expectations (BOSE), which includes a similar

question and do not have the data that is fully responsive to this question to share with the Committee in the time available to respond to these Questions on Notice.

However, at this time, we can confirm that, based on self-reported ages of our Australian monthly active users, less than 10% of Instagram accounts belong to teens under 18, and less than 5% of Facebook accounts belong to teens under 18.