Written Questions on Notice from Senator David Pocock (received 12 September 2024)

How many staff do you employ in Australia – both headcount and FTE?

As at 2 July 2024, we had over 130 employees located in Australia.

How many of these staff are monitoring for deepfakes and/or misinformation – both headcount and FTE?

We invest significantly in the safety and integrity of our services, including policies, technology, industry collaboration and local partnerships to action harmful content as soon as we become aware of it, in many cases, proactively – before Australians are exposed to it. We have around 40,000 people working on safety and security at Meta, and we have invested more than US$20B (~AU$ 29B) in teams and technology to enhance safety and security since 2016.

Our investment in safety and integrity includes our work to combat dis- and mis-information. We remove disinformation and harmful misinformation, including voter and census fraud, and invest in third-party fact checking programs to reduce the distribution of misinformation. Since 2016, we have built the largest global fact-checking network of any platform and have contributed more than $150 million to programs supporting our fact-checking efforts. We partner with more than 90 independent fact-checking organisations around the world who review and rate viral misinformation in more than 60 languages. In Australia, this includes partnerships with Australian Associated Press, Agence France Presse and RMIT FactLab.

Meta's review teams consist of full-time employees who review content as part of a larger set of responsibilities, as well as content reviewers employed by our partners. They come from different backgrounds, reflect our diverse community and have an array of professional experiences – from veterans to legal specialists to enforcement experts in policy areas such as child safety, hate speech and counterterrorism.

Our review teams are global and review content 24/7. We have over 20 sites around the world, where these teams can review content in over 50 languages.

For Australia, we have dedicated market specialists who review accounts, content and ads that are reported to us from Australia. To provide effective quality assurance and to support the resilience and well-being of our reviewers, we primarily co-locate our

specialist teams together in certain centres around the world, such as in Singapore and Dublin.

To provide greater insight on our misinformation measures, Meta is a founding signatory of the DIGI Australian Code of Practice on Disinformation and Misinformation. Under this Code, Meta has committed to safeguards to protect people in Australia against harmful mis- and disinformation, and to adopting a range of scalable measures that reduce its spread and visibility. We have opted into all seven of the Code's objectives across Facebook and Instagram.

To date, Meta has published four transparency reports under the Code, with the latest launched in May 2024. Our 2024 report outlines the steps we took during the 2023 calendar year to meet the 38 commitments we opted into over that reporting period. Meta's 2024 transparency report under the Code also shared a case study outlining our comprehensive strategy leading up to Australia's Aboriginal and Torres Strait Islander Voice Referendum in October 2023 to proactively detect and remove content that breached our services, combat misinformation, harmful content and, and promote civil participation.

You said in your evidence that you use AI to help with monitoring for deepfakes and mis/disinformation. How many languages are your models trained on to identify hateful/harmful content? Please list these languages and the year they commenced.

Detecting hate speech is a difficult and constantly evolving challenge. A new piece of hate speech might not resemble previous examples because it references a new trend or or news story. Over the last several years, we've invested in building proactive detection tools for hate speech, so we can remove this content before people report it to us — and in some cases before anyone even sees it. Our detection techniques include text and image matching, which means we're identifying images and strings of text that are identical to content that's already been removed as hate speech. We also use machine-learning classifiers that look at things like the text in a post, as well as the reactions and comments, to assess how closely it matches common phrases, patterns, and attacks. Our ability to better detect hate speech also includes broadening how our tools understand content, so that our systems look at the image, text, comments, and other elements holistically.

But these systems are not perfect; therefore, we also rely on human reviewers around the world to enforce our Community Standards and Community Guidelines. We also have

around 40,000 people working on safety and security, who operate in over 80 languages, and are bolstered by community reporting.

Will there be a searchable and archived ad library for any ads during the upcoming Australian Federal election? If not, why not? Will dark ads be able to be deleted from the ad library after they have been boosted?

The [Meta Ad Library](#) is a comprehensive and searchable database for ads transparency. People can use this database to search for all active ads running across Meta services. With specific respect to [ads about social issues, elections or politics](#), Meta provides additional information, including spend, reach and funding entities. These ads are visible whether they are active or inactive and are stored in the Ad Library for seven years. This relates to all such ads, regardless of whether an election event is being held.

Please provide a list of researchers, not-for-profits, charities or mental health organisations globally you have funded to do research, directly or via 3rd parties over the last 10 years. Please also provide a list but just for Australia.

Meta supports and directly invests in a significant number of research projects each year with a diverse range of academic, not-for-profit and other expert partners, globally. This traverses a wide range of topic areas, including mis/disinformation, economic impact, online safety and mental health and hate speech. It would therefore not be challenging in the time available to provide a complete list of all of these partnerships over the last 10 years. However, to assist the Committee below are some of the research projects that we have funded in Australia over the last three years:

- Andrea Carson, Rebecca Strating, Simon Jackman, Max Grömping, Phoebe Hayman, Timothy B Gravelle, 2024, 'Influencers and Messages: Analysing and 2023 Voice to Parliament Referendum Campaign. Link [here](#).
- Western Sydney University, '2023 Young People and News longitudinal survey'. Link [here](#).
- Western Sydney University, 'Adult Media Literacy in 2024: Australian attitudes, experiences and needs'. Link [here](#).
- ThoughtLab, 2023, The Digital Journey of SMEs in Australia: How small and medium-sized enterprises drive business and economic results through digital innovation. Link [here](#).
- Online Hate Prevention Institute, 2022, 'Anti-Asian Racism in Australian Social Media'. Link [here](#).
- Islamophobia Register of Australia, Fourth 'Islamobobia in Australia report'. Link [here](#).

- PROJECT ROCKIT, 2023, 'Our Metaverse: Young people and the digital future.' Link [here](#).
- Media Diversity Australia, 2023, Online Safety of Diverse Journalists'. Link [here](#).
- First Draft, 2022, 'A Climate for Mis- and Disinformation in Australia'. Link [here](#).
- Alannah and Madeline Foundation, 2020, 'Digital Home: A Safe and Empowering Environment for Young People.' Link [here](#).

How many Australian users has Meta removed after suspecting and/or confirming they are under the age of 13?

Under Meta's terms of service, users must be at least 13 years old to access services like Facebook and Instagram. While we do not have further data we can share at this time, we are in the process of considering our response to the eSafety Commissioner's request for information under the Basic Online Safety Expectations (BOSE), which seeks information about our approach to age assurance.

How many ads have been served to Australians about ADHD each year for the years from 2018 to 2024.

Specific metrics on ads relating to ADHD on our services are not available, however you can search for ads that of particular interest are running on our services in the [Meta Ad Library](#).

All advertising on our services must comply with both our [Community Standards](#) and [Ad Standards](#). This includes restrictions on advertising related to drugs and pharmecuticals or advertising that includes direct or indirect assertions or implications about a person's physical or mental health (including medical conditions)

How much has been spent on boosting ads about ADHD diagnosis during that period? How are these ads approved?

Specific metrics on ads relating to ADHD on our services are not available. All ads run on Facebook and Instagram are reviewed against our [Advertising Standards](#). Our [ads review system](#) comprises a mix of primarily automated technology to apply our Advertising Standards to the millions of ads that are run across our services. We also use human reviewers to improve and train our automated systems and, in some cases, manually review ads.

As noted in our answer to the previous question, all advertising on our services must comply with both our [Community Standards](#) and [Ad Standards](#). This includes restrictions on advertising related to drugs and pharmecuticals or advertising that includes direct or indirect assertions or implications about a person's physical or mental health (including medical conditions)

How many scams have you detected on Facebook Marketplace in Australia in the last 3 years? How many scams have been taken down? How many accounts have been banned or deleted due to scams?

We do not have readily available data at this point. However, we invest in integrity systems to identify and remove accounts that violate our policies. For example, our goal is to remove as many fake accounts on Facebook as we can, including accounts created with malicious intent to violate our policies. We prioritize enforcement against fake accounts that seek to cause harm. Many of these accounts are used in spam campaigns and are financially motivated. From April to June 2024, we actioned 1.2 billion accounts for being fake, 99.7% of which we removed proactively before being reported to us.

How big is your team that deals with scams on Facebook Marketplace in Australia – head count and FTE?

It isn't possible to break out this number given the diverse number of ways in which we work to combat scams. Combatting advertising fraud cuts across many teams, including: security engineers who architect our authentication mechanisms, threat intelligence teams who track threat actors, integrity teams who adjust and use machine learning to detect abusive accounts and content, and support teams who help remediate the issue. Additionally, we use a combination of AI detection technologies, in addition to human review to identify content that violates our policies. Given the volume and scale of content on our platforms, we primarily rely on AI to identify and remove violating content, often before anyone has seen it.

With respect to our review teams are global and review content 24/7. We have over 20 sites around the world, where these teams can review content in over 80 languages. We partner with companies that employ over 15,000 reviewers who help in doing the job of reducing harm, including around 2,500 language agnostic reviewers.

For Australia, we have dedicated market specialists who review accounts, content and ads that are reported to us from Australia. To provide effective quality assurance and to support the resilience and well-being of our reviewers, we primarily co-locate our

specialist teams together in certain centres around the world, such as in Singapore and Dublin.

On Facebook Marketplace why is there currently no easy way to report scams?

Users are able to report Facebook Marketplace scams in relation to purchases or listings, as outlined in our Facebook Help Centre [here](#). All content, buyers, and sellers in Marketplace can be reported in case other users find them to be violating. Reporting sends a signal to us that something is wrong and needs investigating, and helps improve the quality of their Marketplace experience.

Our Help Centre also provides [tips](#) on how to recognise a scam on Facebook Marketplace. It also has a variety of educational sections on [Marketplace Help Center](#) to inform users on safe and responsible shopping - including ["Trust & Safety on Marketplace"](#) compilation of resources.

We also have a dedicated channel for authorized Government agencies to report content to us. In Australia, we have on boarded a number of Australian government agencies to report commercially-related content including related with Marketplace.

If someone makes a complaint on Marketplace about a scam how does this get assessed? Who assesses it?

All users who are using Marketplace must also abide by our Community Standard and Commerce Policy, which prohibit among other things, fraud, scams and deception, investment and financial scams, inauthentic identity scams, and product and rewards scams.

We review the listings in Marketplace in the same way as on other parts of Facebook and Instagram, via a combination of automated and human review.

Users who repeatedly violate Community Standards and Commerce Policies in Marketplace may be suspended or banned from using Marketplace, or from using Facebook altogether.

What was Meta's turnover in Australia in FY2023/24?

Facebook Australia Pty Limited (**Facebook Australia**) reported revenue of $208,972,934 in its publicly available and audited financial reports for the financial year ended 31 December 2023.

How much advertising revenue did Meta receive in FY2023/24 from Australia?

Facebook Australia reported advertising revenue of $205,737,970 in its publicly available and audited financial reports for the financial year ended 31 December 2023.

How much tax did Meta pay in Australia in FY2023/24?

For the last financial year ended 31 December 2023, Facebook Australia paid over $35.4 million of income taxes to the Australian Taxation Office (ATO). This represents an effective tax rate of 43 percent. In addition, Meta collected and paid more than $170 million in GST to the ATO during the same period.

A [research pilot](#) conducted recently by Deakin University found that young people, under the age of 18, had been served targeted ads for alcohol, including ads for Bombay Sapphire and Baileys. This ongoing research builds on [research from the University of Queensland and Monash University](#), who found that 93% of 16 to 17-year-olds in their study had seen alcohol ads in their Facebook feeds. Indeed, that research found alcohol-related advertising keywords had been attached to the accounts of 17-year-olds, including "alcohol" and "alcoholic drink". In your evidence, Meta said that alcohol ads are not served to people under 18. The real world evidence suggests otherwise. How many alcohol ads have been served to people under the age of 18 on Meta platforms in the last five years?

Alcohol ads are not allowed to be served to people under 18 on our services. Meta has a specific [alcohol advertising policy](#), which explicitly states that advertisers may only run ads that promote or reference alcohol if those ads:

- Follow the targeting requirements of the location of the audience; and
- Do not target people under the age of 18.

We welcome further discussion on any ads that may be in violation of this policy (further information is available [here](#)) in Australia so that we may have the opportunity to conduct further investigation.

Further, while we take measures to enforce these policies, age verification remains a challenge for many apps. This is why we support both greater parental controls for the use of our services by teens, and greater age control measures at the app store or operating system level so that age verification can operate seamlessly across the app ecosystem.