

# **Meta response to the Australian disinformation and misinformation industry code**

Reporting period: January - December 2022

*Published May 2023*

## Summary

Meta is proud to be a founding member and signatory to the Australian industry code on misinformation and disinformation. The code is a credible step in encouraging collaboration between the technology industry and governments to combat misinformation and disinformation.

Meta has opted into every commitment under the code, and over the past two years we have released two transparency reports outlining our efforts to meet these commitments.

In 2022, Meta's transparency report outlined 45 commitments to meet our obligations across both Facebook and Instagram. This report outlines the steps we took to meet these commitments across the reporting period of January to December 2022.<sup>1</sup> We have provided updated case studies on our work during the 2022 federal election and the conflict in Ukraine, where we introduced a number of new measures to combat misinformation and disinformation.

We also continue to release up-to-date, Australian specific data on our misinformation efforts to maintain transparency with the community and policymakers about the impact of our misinformation efforts.

Between 1 January and 31 December 2022:

- We took action on over 91,000 pieces of content across Facebook & Instagram in Australia for violating our Harmful Health Misinformation policies.
- We displayed warnings on over 9 million distinct pieces of content on Facebook in Australia (including reshares) based on articles written by our third party fact checking partners.
- We removed over 100 accounts, pages, and groups in Australia for repeatedly violating our Harmful Health Misinformation policies.
- There were over 1.8 billion visits to the COVID-19 Information Centre. Over 7.5 million of these visits were from users in Australia.
- There were over 84 million visits to the Climate Science Information Centre. Over 25,000 of these visits were from users in Australia.

---

<sup>1</sup> This reporting period has been specified in the guidelines provided by the independent reviewer, engaged by DIGI.

In 2022, some highlights of our work include:

- Implementing a suite of measures in the lead up to the 2022 Australian federal election to proactively detect and remove content that breaches our policies, combat misinformation, harmful content and coordinated inauthentic behaviour, and promote civic participation.

As a result of these efforts, during the election campaign (between April 1 and June 30, 2022):

- We took action on over 25,000 pieces of content across Facebook and Instagram for violating our Harmful Health Misinformation policies.
- We displayed warnings on over 3 million distinct pieces of content on Facebook (including reshares) based on articles written by our third party fact checking partners.
- We took action on over 91,000 pieces of content on Facebook and over 40,000 pieces of content on Instagram in Australia for violating our hate speech policies.
- We took action on over 200,000 pieces of content on Facebook and over 46,000 of content on Instagram in Australia for violating our Community Standards on violence and incitement.
- We rejected around 17,000 ads for not complying with our political and social issue ads enforcement policies.

We also released two election prompts on Facebook, and Instagram election stickers to encourage people to vote, and direct users to the Australian Electoral Commission (AEC) website:

- The 'enrol to vote' prompt was seen by 23.3 million Australians, with 54,000 people clicking through to the AEC website, and 900,000 people sharing it across their Feeds.
  - The 'election day reminder' was seen by almost 11 million Australians, clicked through 175,000 times, and shared by 60,000 people on their Feed.
  - The Instagram election stickers were used on almost 35,000 Instagram Stories.
- Supporting First Draft (now RMIT FactLab) to conduct training with journalists on how to identify and prevent amplifying misinformation and disinformation.

- Introducing new tools to provide users with more transparency around content ranking, including updates to the ‘Why Am I Seeing This Ad?’ tool, and AI System Cards for Instagram.
- Expanding transparency in our Ad Library, to include new information about targeting selections made by advertisers for social issues, election and political ads.
- Sponsoring events in relation to misinformation and disinformation, including the Australian Media Literacy Alliance’s inaugural Media Literacy Summit, and the 2023 Sydney Dialogue.
- Investing in new research on misinformation from the Queensland University of Technology by Michelle Riedlinger and Silvia Montaña-Niño, and Marina Joubert (Stellenbosch University), Víctor García-Perdomo (Universidad de La Sabana) on ‘Countering misinformation in the Southern Hemisphere: A comparative study’.
- Investing in new research on media literacy of First Nations peoples by Professor Tristan Kennedy from Monash University.
- Expanding the Facebook Open Research and Transparency (FORT) initiative so that it is available to Australian researchers. We also expanded the data that is available within the FORT platform, to now provide access to up to date data sets on targeting information for ads about social issues, elections, and politics.

These steps are in addition to our extensive global efforts to combat misinformation and disinformation.

This report also outlines Meta’s 38 commitments for the next reporting period. This includes supporting a new outcome in the Code, outcome 1.e: relating to recommender systems. This outcome was included in the Code in line with feedback during consultation in 2022.

Meta takes a risk-based approach to misinformation, and our approach continues to evolve in line with advice on COVID-19, and the risk of harm. For this reason, Meta’s 2023 commitments reflect the changing nature of the pandemic.

We recognise that the upcoming referendum into an Aboriginal and Torres Strait Islander Voice will be a significant moment for Australia. Many Australians will use digital platforms to engage in advocacy, express their views, or participate in democratic debate. Meta will announce the measures we are taking specifically for the referendum later in 2023, following the conclusion of consultation with relevant stakeholders and Meta's Aboriginal and Torres Strait Islander Advisory Council. We will report on these measures in the next annual report under the code.

We look forward to continuing to work with Australian policymakers, civil society, academics and experts on steps to combat misinformation and disinformation in Australia over the next year.

## List of Meta commitments under the Australian industry code on misinformation and disinformation for 2023

A copy of the 2022 commitments can be found in **Appendix A**.

The 38 commitments for 2023 are outlined below:

<b>Outcome 1:</b>  <b>Combatting misinformation / disinformation</b>	<b>Outcome 1a. Signatories contribute to reducing the risk of Harms that may arise from the propagation of Disinformation and Misinformation on digital platforms by adopting a range of scalable measures.</b> <ul style="list-style-type: none"><li>● Meta removes networks of accounts, Pages and Groups that violate our inauthentic behaviour policy, including disinformation that violates our policy on Inauthentic Behaviour.</li><li>● Meta provides transparency about accounts, Pages and Groups removed under our Inauthentic Behaviour policy.</li><li>● Meta partners with experts and organisations who assist in providing tips or further investigation about possible inauthentic behaviour on our services.</li><li>● Meta removes misinformation that violates the misinformation section of our Community Standards, and will review its Misinformation and Harm Policy, in line with recommendations from the Oversight Board.</li><li>● Meta removes manipulated videos, also known as “deepfakes”, that violates our Manipulated Media policy.</li><li>● Meta removes election-related misinformation that may constitute voter fraud or interference under our Co-ordinating Harm and Promoting Crime policy.</li><li>● Meta removes fake accounts.</li><li>● Meta allows for appeals in instances where users may disagree with our enforcement, including to the independent and external Oversight Board.</li><li>● Meta partners with third-party fact-checking organisations, globally and in Australia, to assess the accuracy of content on our services.</li></ul>
--	--

- Meta applies a warning label to content found to be false by third-party fact-checking organisations.
- Meta reduces the distribution of content found to be false by third-party fact-checking organisations.
- Meta proactively searches for content that matches content debunked by our fact-checking partners, to apply the same treatments.
- Meta takes action on Pages, Groups, accounts or websites found to repeatedly share misinformation.

**Outcome 1b. Users will be informed about the types of behaviours and types of content that will be prohibited and/or managed by Signatories under this Code.**

- Meta makes available a detailed list of claims that we consider to violate our COVID-19 Misinformation & Harm policy.
- Meta makes information available via a dedicated website that outlines our efforts to combat misinformation.

**Outcome 1c . Users can report content and behaviours to Signatories that violates their policies under 5.10 through publicly available and accessible reporting tools.**

- Meta makes on-platform reporting channels available to users for false information.

**Outcome 1d. Users will be able to access general information about Signatories' actions in response to reports made under 5.11.**

- Meta makes global transparency reports available regularly.
- Meta will supplement these reports with additional Australia-specific statistics, provided as part of this Annual Report process.

	<p><b>[*NEW*] Outcome 1e. Users will be able to access general information about Signatories’ use of recommender systems and have options relating to content suggested by recommender systems.</b></p> <ul style="list-style-type: none"> <li>• Meta will continue to provide greater transparency of our content ranking algorithms and give users more control over the content they see.</li> <li>• Meta takes steps to limit the possible distribution of misinformation via recommendations.</li> </ul>
<p><b>Outcome 2:</b> <b>Disrupt monetisation and advertising incentives</b></p>	<ul style="list-style-type: none"> <li>• Meta sets a higher threshold for users to be able to advertise on our services, and takes action against users who spread misinformation.</li> </ul>
<p><b>Outcome 3:</b> <b>Combat inauthentic user behaviour</b></p>	<p>See items listed under Outcome 1.</p>
<p><b>Outcome 4:</b> <b>Empower consumers to be informed</b></p>	<ul style="list-style-type: none"> <li>• Meta provides contextual information around posts that users see from public Pages.</li> <li>• Meta provides a Climate Science Information Centre in Australia to connect users to authoritative information from leading climate organisations.</li> <li>• Meta uses in-product prompts to direct Australians to authoritative information on key topics.</li> <li>• Meta promotes authoritative information by providing significant support to organisations such as the Australian, state and territory governments and to promote authoritative health information.</li> <li>• Meta directs users to authoritative information when they search for high-priority topics on Facebook and Instagram.</li> </ul>



	<ul style="list-style-type: none"> <li>• Meta directs users to authoritative information once they have seen or shared COVID-19 related misinformation.</li> <li>• Meta will look for opportunities to continue to work with the Government on other ways to promote authoritative information.</li> <li>• Meta promotes public service announcements to our users to encourage them to be wary of potential misinformation.</li> </ul>
<b>Outcome 5: Political advertising</b>	<ul style="list-style-type: none"> <li>• Meta requires all advertisers of political and social issue ads to complete an ad authorisation, which includes verifying the advertiser's identity.</li> <li>• Meta requires political and social issue ads to include a disclaimer disclosing who is paying for the ad.</li> <li>• Meta provides the Ad Library, a searchable archive of all political and social issue ads on our services in Australia.</li> <li>• Meta enables an Ad Library report that provides aggregated spend information about Pages undertaking political and social issue ads.</li> </ul>
<b>Outcome 6: Research</b>	<ul style="list-style-type: none"> <li>• Meta supports research and events in relation to misinformation and media literacy.</li> <li>• Meta supports research and events in relation to disinformation.</li> <li>• Meta collaborates with researchers to undertake surveys of our users to assess their views on topics such as vaccines and climate change.</li> <li>• Meta provides data to researchers in a privacy-protective way via the Facebook Open Research and Transparency initiative.</li> </ul>
<b>Outcome 7: Annual reports</b>	<ul style="list-style-type: none"> <li>• Meta will continue to publish annual reports in Australia, such as these, to be transparent about the steps we are taking to combat disinformation and misinformation.</li> </ul>

## Table of contents

Summary	2
List of Meta commitments under the Australian industry code on misinformation and disinformation for 2023	6
Table of contents	10
Reporting against 2022 commitments	11
Outcome 1a	11
Case study: Meta’s CIB response to the war in Ukraine	14
Outcome 1b	23
Outcome 1c	24
Outcome 1d	25
Outcome 2	27
Outcome 3	28
Outcome 4	29
Case study: Australian federal election 2022	36
Outcome 5	40
Outcome 6	42
Outcome 7	45
Appendix A: 2022 specific commitments made by Meta under the industry code on misinformation and disinformation	46

## Reporting against 2022 commitments

### Outcome 1a

Signatories contribute to reducing the risk of Harms that may arise from the propagation of Disinformation and Misinformation on digital platforms by adopting a range of scalable measures.

**Signatories will develop and implement measures which aim to reduce the propagation of and potential exposure of users of their services and products to Disinformation and Misinformation**

Our approach to misinformation and disinformation is consistent with that which was outlined in our 2021 and 2022 transparency report.<sup>2</sup> Below we outline developments from 2022:

- **Meta removes networks of accounts, Pages and Groups that violate our inauthentic behaviour policy, including disinformation that violates our policy on Inauthentic Behaviour.**

We have maintained the approach outlined in our 2021 and 2022 transparency report. You can find details of Meta's most recent efforts to remove coordinated inauthentic behaviour (CIB) in the commitment below.

- **Meta provides transparency about accounts, Pages and Groups removed under our Inauthentic Behaviour policy.**

We use a combination of policies, tools, expert teams and partnerships to detect and remove networks of IB and CIB - both foreign and domestic.

We continue to report on our efforts to disrupt CIB through our Community Standards Enforcement Report and Monthly Adversarial Threats report.<sup>3</sup>

---

<sup>2</sup> J Machin, 'Meta's response to the Australian disinformation and misinformation industry code', *Meta Australia blog*, May 2021, <https://australia.fb.com/wp-content/uploads/sites/69/2021/05/Facebook-commitments-under-disinfo-and-misinfo-code-final-1.pdf>

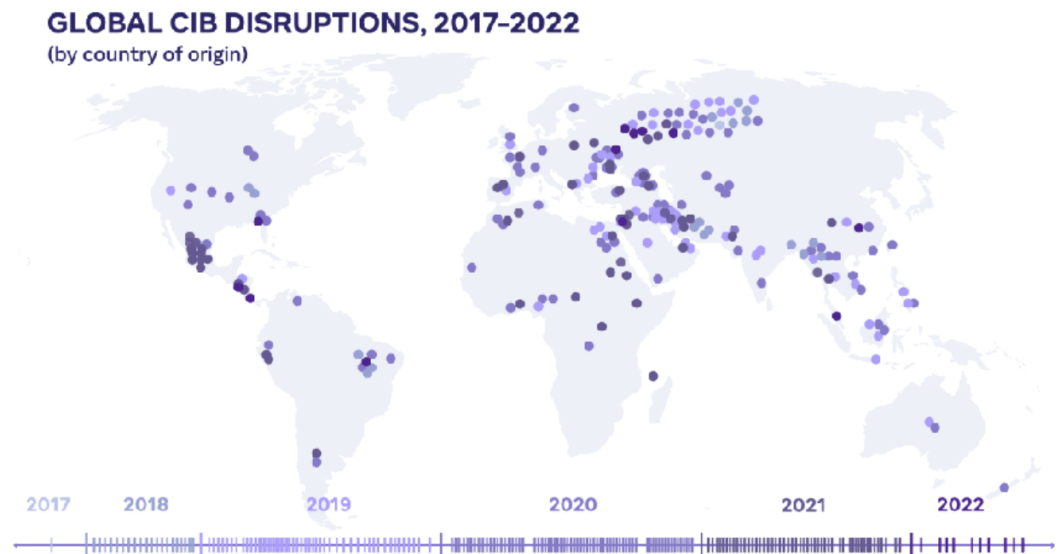
<sup>3</sup> Meta, *Community Standards Enforcement Report*, <https://about.fb.com/news/tag/coordinated-inauthentic-behavior/> and Meta, *Community Standards - Coordinated Inauthentic Behaviour*, <https://about.fb.com/news/tag/coordinated-inauthentic-behavior/>

- In December 2022, we released new data on Meta’s CIB take downs. Meta has taken down and reported on more than 200 covert influence operations.<sup>4</sup>

These networks came from 68 countries and operated in at least 42 languages, with most targeting audiences in their home countries and only around one-third aimed solely at audiences abroad.

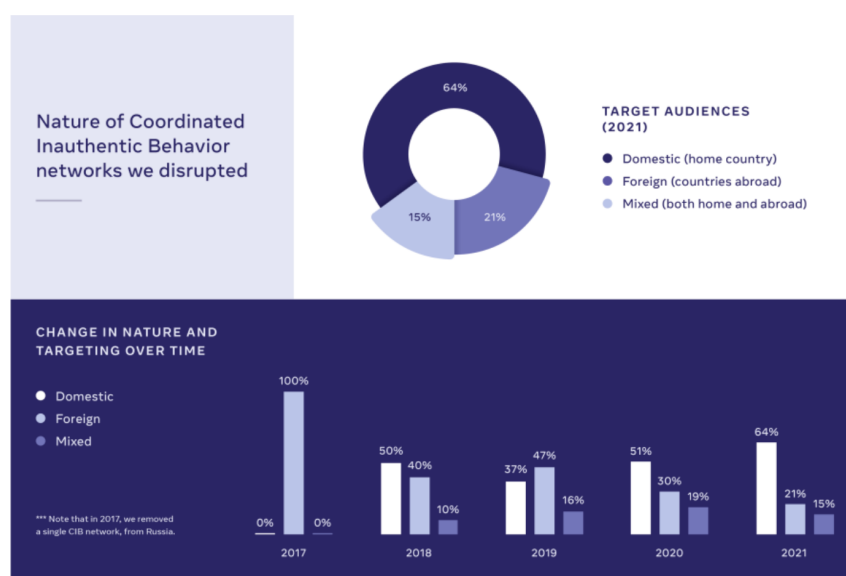
The United States was the most targeted country by global CIB operations we’ve disrupted over the years, followed by Ukraine and the United Kingdom.

### Global Coordinated Inauthentic Behaviour Disruptions, 2017 – 2022



<sup>4</sup> B Nimmo, D Agranovich, ‘Recapping our 2022 coordinated inauthentic behaviour enforcements’, *Meta Newsroom*, 15 December 2022, <https://about.fb.com/news/2022/12/metas-2022-coordinated-inauthentic-behavior-enforcements/> and Meta, ‘December 2021 Coordinated Inauthentic Behaviour Report’, December 2021, <https://about.fb.com/wp-content/uploads/2022/01/December-2021-Coordinated-Inauthentic-Behavior-Report-2.pdf>

## Target of Coordinated Inauthentic Behaviour Disruptions, 2017 – 2022<sup>5</sup>



We also recently reported that we had identified more than 400 malicious Android and iOS apps that were designed to steal Facebook login information and compromise people’s accounts.<sup>6</sup> These apps were listed on the Google Play Store and Apple’s App Store and disguised as photo editors, games, VPN services, business apps and other utilities to trick people into downloading them.

In our 2022 transparency report we reported that Meta had taken action on four instances of CIB operations that targeted Australians. This remains the same. During the 2022 federal election, we did not see any evidence of coordinated inauthentic behaviour targeting Australia.

<sup>5</sup> We define targets as:

- Domestic: IO that targets public debate in the same country from which it operates.
- Foreign: IO that targets the public debate in a different country from which it operates.
- Mixed: We also see IO campaigns and threat actors that run campaigns that target both domestic and foreign audiences

<sup>6</sup> D Agranovich, ‘Protecting people from malicious account compromise apps’, *Meta Newsroom*, 7 October 2022, <https://about.fb.com/news/2022/10/protecting-people-from-malicious-account-compromise-apps/>

## Case study: Meta's CIB response to the war in Ukraine

Since the beginning of the devastating war in Ukraine, Meta has taken a number of steps to combat misuse of our services and help ensure the safety of our community - both in Ukraine and around the world.<sup>7</sup>

In Meta's 2022 transparency report, we provided a case study of our work to reduce misinformation and disinformation related to the conflict in the Ukraine. This case study outlines our latest work to combat CIB operations in relation to the Ukraine conflict.

In 2022, the United States was the most targeted country by global CIB operations we've disrupted over the years, followed by Ukraine and the United Kingdom.<sup>8</sup>

### Targets and CIB networks removed, 2017 - 2022



In December 2022, we released an overview of our latest efforts to combat influence operations relating to the conflict in Ukraine.<sup>9</sup> It finds that while Russian-origin attempts at CIB related to Russia's war in Ukraine have sharply increased, overt efforts by Russian state-controlled media have reportedly decreased over the last 12 months on our platform. We saw state-controlled media shifting to other platforms and using new domains to try to escape the additional transparency on (and demotions against) links to their websites.

During the same period, covert influence operations have adopted a brute-force, “smash-and-grab” approach of high-volume but very low-quality campaigns across the internet.

In addition to these insights, we have taken action on the following CIB networks since the war in Ukraine began:

- Removed a network in Russia for abusing our reporting tools to repeatedly report people in Ukraine and in Russia for fictitious policy violations of Facebook policies in an attempt to silence them.<sup>10</sup>
- Detected and disrupted recidivist CIB activity linked to the Belarusian KGB who suddenly began posting in Polish and English about Ukrainian troops surrendering without a fight and the nation’s leaders fleeing the country. Prior to that, this particular threat actor primarily focused on accusing Poland of mistreating migrants from the Middle East. On March 14 2022, they pivoted back to Poland and created an event in Warsaw calling for a protest against the Polish government. We disabled the account and event that same day.<sup>11</sup>
- Took down a network run by people in Russia and Ukraine targeting Ukraine. They ran websites posing as independent news entities and created fake personas across social media platforms including Facebook, Instagram, Twitter, YouTube, Telegram and also Russian Odnoklassniki and VK.<sup>12</sup>

<sup>7</sup> Meta, ‘Meta’s ongoing efforts regarding Russia’s invasion of Ukraine’, *Meta Newsroom*, 26 February 2022, <https://about.fb.com/news/2022/02/metas-ongoing-efforts-regarding-russias-invasion-of-ukraine/>

<sup>8</sup> B Nimmo, D Agranovich, ‘Recapping our 2022 coordinated inauthentic behaviour enforcements’, *Meta Newsroom*, 15 December 2022, <https://about.fb.com/news/2022/12/metas-2022-coordinated-inauthentic-behavior-enforcements/> and Meta, ‘December 2021 Coordinated Inauthentic Behaviour Report’, December 2021, <https://about.fb.com/wp-content/uploads/2022/01/December-2021-Coordinated-Inauthentic-Behavior-Report-2.pdf>

<sup>9</sup> B Nimmo, M Franklin, D Agranovich, L Hundley, M Torrey, ‘Quarterly Adversarial Threat Report’, *Meta*, February 2023, <https://about.fb.com/wp-content/uploads/2023/02/Meta-Quarterly-Adversarial-Threat-Report-Q4-2022.pdf>

<sup>10</sup> B Nimmo, ‘Meta’s Adversarial Threat Report, First Quarter 2022’, *Meta Newsroom*, 7 April 2022, <https://about.fb.com/news/2022/04/metas-adversarial-threat-report-q1-2022/>

<sup>11</sup> B Nimmo, D Agranovich, N Gleicher, ‘Adversarial Threat Report: Detailed Report’, Meta, April 2022, <https://about.fb.com/wp-content/uploads/2022/04/Meta-Quarterly-Adversarial-Threat-Report-Q1-2022.pdf>

<sup>12</sup> N Gleicher, ‘Updates on our security work in Ukraine’, *Meta Newsroom*, 27 February 2022, <https://about.fb.com/news/2022/02/security-updates-ukraine/>

- Took down a network of Instagram accounts operated by a troll farm in St. Petersburg, Russia, which targeted global public discourse about the war in Ukraine. This appeared to be a poorly executed attempt, publicly coordinated via a Telegram channel, to create a perception of grassroots online support for Russia's invasion by using fake accounts to post pro-Russia comments on content by influencers and media.<sup>13</sup>
- Took down a large network that originated in Russia and targeted primarily Germany, and also France, Italy, Ukraine and the United Kingdom with narratives focused on the war in Ukraine. The operation began in May 2022 and centered around a sprawling network of over 60 websites carefully impersonating legitimate websites of news organisations in Europe, including Spiegel, The Guardian and Bild. There, they would post original articles that criticised Ukraine and Ukrainian refugees, supported Russia and argued that Western sanctions on Russia would backfire. They would then promote these articles and also original memes and YouTube videos across many internet services, including Facebook, Instagram, Telegram, Twitter, petitions websites Change.org and Avaaz, and even LiveJournal.<sup>14</sup>
- Finally, we saw a spike in compromise attempts aimed at members of the Ukrainian military by Ghostwriter, a threat actor tracked by the security community. In a handful of cases, groups posted videos calling on the Army to surrender as if these posts were coming from the legitimate account owners. We blocked these videos from being shared.<sup>15</sup>

---

<sup>13</sup> B Nimmo, D Agranovich, N Gleicher, 'Adversarial Threat Report: Detailed Report', Meta, August 2022, <https://about.fb.com/wp-content/uploads/2022/08/Quarterly-Adversarial-Threat-Report-Q2-2022.pdf>

<sup>14</sup> B Nimmo and D Agranovich, 'Removing coordinated inauthentic behaviour from China and Russia', Meta Newsroom, 27 September 2022, <https://about.fb.com/news/2022/09/removing-coordinated-inauthentic-behavior-from-china-and-russia/>

<sup>15</sup> Meta, 'Meta's ongoing efforts regarding Russia's invasion of Ukraine', Meta Newsroom, 26 February 2022, <https://about.fb.com/news/2022/02/metas-ongoing-efforts-regarding-russias-invasion-of-ukraine/>



- **Meta partners with experts and organisations who assist in providing tips or further investigation about possible inauthentic behaviour on our services.**

Meta continues to maintain close relationships with experts and organisations around the world so we can share threat findings and adapt our enforcement.

In the lead up with the 2022 Australian federal election, Meta worked with the Australian Government’s Election Integrity Assurance Taskforce (EIAT) and developed responses to a series of possible scenarios that may arise during the course of an election campaign. This is explained in more detail in the Case Study: Australian federal election 2022.

- **Meta removes misinformation that violates the misinformation component of our Community Standards.**

Meta continues to remove misinformation that violates our Community Standards,<sup>16</sup> including misinformation that is likely to directly contribute to the risk of imminent physical harm.

In our 2022 transparency report, we reported that since the beginning of the pandemic we removed more than 20 million pieces of content from Facebook and Instagram globally for violating our policies on COVID-19 related misinformation.<sup>17</sup>

In July 2022, we reported that we have removed more than 25 million pieces of content from Facebook and Instagram globally for violating our policies on COVID-19 related misinformation.<sup>18</sup>

We continue to work with experts around the world to update our policies on misinformation and ensure they are appropriate and reflect both the offline and online world. For this reason, in 2022, Meta asked the Oversight Board for advice on whether measures to address dangerous COVID-19 misinformation should remain in place.<sup>19</sup>

---

<sup>16</sup> Meta, *Community Standards*, <https://transparency.fb.com/en-gb/policies/community-standards/>

<sup>17</sup> Meta, *Community Standards Enforcement Report Q2 2021*, <https://transparency.fb.com/data/community-standards-enforcement/>

<sup>18</sup> N Clegg, ‘Meta asks Oversight Board to advise on COVID-19 misinformation policies’, *Meta Newsroom*, 26 July 2022, <https://about.fb.com/news/2022/07/oversight-board-advise-covid-19-misinformation-measures/>

<sup>19</sup> N Clegg, ‘Meta asks Oversight Board to advise on COVID-19 misinformation policies’, *Meta Newsroom*, 26 July 2022, <https://about.fb.com/news/2022/07/oversight-board-advise-covid-19-misinformation-measures/>

Misinformation related to COVID-19 has presented unique risks to public health and safety over the last two years and more. To keep our users safe while still allowing them to discuss and express themselves on this important topic, we broadened our harmful misinformation policy in the early days of the outbreak in January 2020. The change meant that, for the first time, the policy would provide for removal of entire categories of false claims on a worldwide scale.

The pandemic has evolved. In many countries, where vaccination rates are relatively high, life is increasingly returning to normal. But this isn't the case everywhere and the course of the pandemic will continue to vary significantly around the globe — especially in countries with low vaccination rates and less developed healthcare systems. It is important that any policy Meta implements be appropriate for the full range of circumstances countries find themselves in.

Meta remains committed to combating COVID-19 misinformation and providing people with reliable information. We sought advice from the Oversight Board about our measures to address COVID-19 misinformation, including whether those introduced in the early days of an extraordinary global crisis remains the right approach, or whether we should address this misinformation through other means, like labelling or demoting it either directly or through our third-party fact-checking program.

In April 2023, the Board issued 18 recommendations, including for Meta to continue to remove harmful misinformation about COVID-19 for as long as the World Health Organization declares COVID-19 a global public health emergency. The Oversight Board also recommended Meta review the claims it removes under its Misinformation and Harm Policy in consultation with expert stakeholders.<sup>20</sup>

Meta will continue to enforce on our Misinformation and Harm policy while we prepare our response to the Oversight Board's recommendations.

- **Meta removes manipulated videos, also known as “deepfakes”, that violates our Manipulated Media policy.**

Meta continues to enforce its policy on manipulated videos. We have maintained the approach outlined in our 2021 and 2022 transparency report.

---

<sup>20</sup> Oversight Board, 'Oversight Board publishes policy advisory opinion on the removal of COVID-19 misinformation', *Oversight Board*, April 2023, <https://oversightboard.com/news/739141534555182-oversight-board-publishes-policy-advisory-opinion-on-the-removal-of-covid-19-misinformation/>

- **Meta removes material that violates our Violence-Inducing Conspiracy Network policy.**

Meta continues to enforce its policy on Violence-Inducing Conspiracy Networks.

We have secured new data in this area. In our 2022 transparency report we reported that as of September 14, 2021, we had identified over 1,013 militarised social movements to date and in total, removed about 4,000 Pages, 20,600 groups, 190 events, 54,900 Facebook profiles and 8,300 Instagram accounts. We'd also removed about 3,900 Pages, 11,300 groups, 640 events, 50,300 Facebook profiles and 32,500 Instagram accounts for violating our policy against QAnon.<sup>21</sup>

As of August 15, 2022, we have identified over 1,151 militarized social movements to date and in total, removed about 4,200 Pages, 20,800 groups, 200 events, 59,800 Facebook profiles and 8,900 Instagram accounts. We've also removed about 4,200 Pages, 12,000 groups, 840 events, 67,200 Facebook profiles and 38,800 Instagram accounts for violating our policy against QAnon.<sup>22</sup>

- **Meta removes election-related misinformation that may constitute voter fraud or interference under our Coordinating Harm and Promoting Crime policy.** We have maintained the approach outlined in our 2021 and 2022 transparency report.
- **Meta removes fake accounts.**

Meta continues to enforce its policy on fake accounts. We do not allow fake accounts on Facebook, as they can be vehicles for a range of harmful content and behaviour.

---

<sup>21</sup> Meta, 'An update to how we address movements and organizations tied to violence', *Meta Newsroom*, blog post updated 19 January 2021, <https://about.fb.com/news/2020/08/addressing-movements-and-organizations-tied-to-violence/>.

Note - Meta began reporting our enforcement data for the Violence-Inducing Conspiracy Network policy in August 2020. Only global data is available, as this includes Groups, Pages and Events which can compromise users based in numerous different countries.

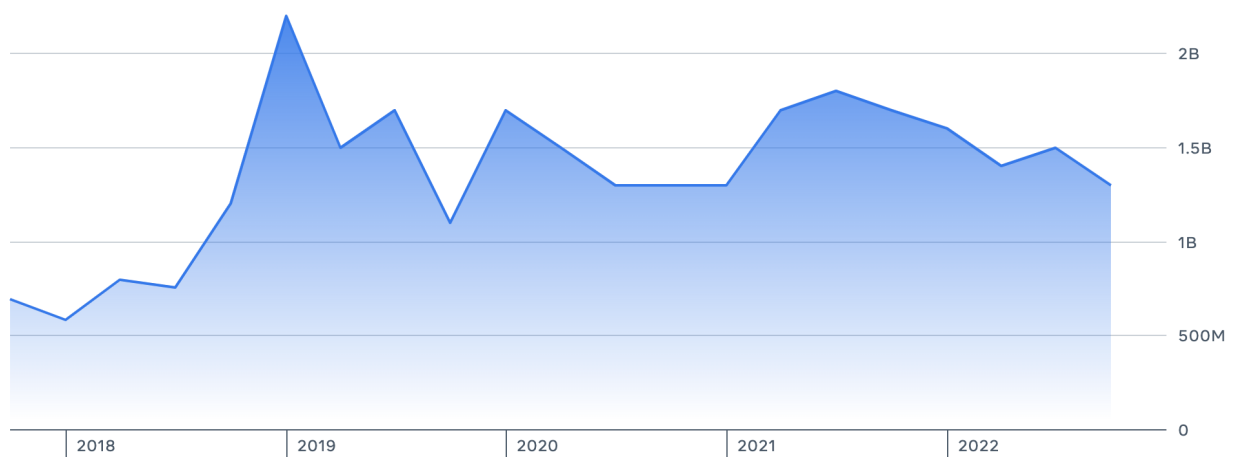
<sup>22</sup> Meta, 'An update to how we address movements and organizations tied to violence', *Meta Newsroom*, blog post updated 19 January 2021, <https://about.fb.com/news/2020/08/addressing-movements-and-organizations-tied-to-violence/>.

Note - Meta began reporting our enforcement data for the Violence-Inducing Conspiracy Network policy in August 2020. Only global data is available, as this includes Groups, Pages and Events which can compromise users based in numerous different countries.

Our ability to detect and remove fake accounts has been improving over the years, and there has been a general decline in the volume of fake accounts found on the platform since 2019.

From January to December 2022, we detected and removed 5.8 billion fake accounts, on average we proactively detected and removed 99.6 per cent of these accounts before they were reported to us.<sup>23</sup> These are often caught within minutes of registration.

**# of fake accounts we've taken action on (Oct 2017 - Dec 2022)<sup>24</sup>**



- **Meta allows for appeals in instances where users may disagree with our enforcement, including to the independent and external Oversight Board.**

As mentioned in the 2021 and 2022 transparency report, Meta has voluntarily established the independent Oversight Board to ensure greater accountability for our content governance on Facebook and Instagram.

In our last report we noted that the Oversight Board had issued 24 decisions and made more than 100 recommendations to Meta for future improvements.<sup>25</sup>

<sup>23</sup> Meta, *Community Standards Enforcement Report Q4 2022*, <https://transparency.fb.com/data/community-standards-enforcement/>

<sup>24</sup> Note - graph represents number of fake accounts removed globally. Australia-specific statistics are not available.

<sup>25</sup> Oversight Board, 'Oversight Board demands more transparency from Facebook', *Oversight Board*, October 2021, <https://oversightboard.com/news/215139350722703-oversight-board-demands-more-transparency-from-facebook/>

As of December 2022, the Oversight Board had issued 36 case decisions and 176 recommendations to Meta for future improvements.

Meta has reported its progress against implementing 140 of these recommendations. Meta has implemented 24 (17%) of the Board's recommendations fully, as demonstrated through published information. Eleven, (8%) have been partially implemented, and Meta has reported progress towards implementing 53 (38%). Meta has reported implementation against 28 (20%) recommendations, or said it already does what the Board recommends.<sup>26</sup>

- **Meta partners with third-party fact-checking organisations, globally and in Australia, to assess the accuracy of content on our services.**

Meta has built the largest global fact-checking network of any digital platform.

In the 2022 transparency report we reported that Meta works with over 80 fact-checking partners around the world covering more than 60 languages.

As of December 2022, Meta partners with over 90 fact checking partners covering more than 60 languages.<sup>27</sup>

Australians benefit from Meta's international approach to fact-checking - an Australian user will see a warning label on content that has been fact-checked by an international fact-checking partner. Content found to be false by our international fact-checking partners will be demoted in an Australian user's Feed, meaning there is less chance of them seeing it.

In March 2022, ahead of the Australian federal election, we announced the expansion of our third-party fact-checking program to include RMIT FactLab,<sup>28</sup> alongside the Australian Associated Press and Agence France Presse. All fact-checks by these partners are publicly available on their websites.<sup>29</sup>

---

<sup>26</sup> Oversight Board, 'Oversight Board announces plans to review more cases, and appoints a new board member', Oversight Board, February 2023, <https://www.oversightboard.com/news/943702317007222-oversight-board-announces-plans-to-review-more-cases-and-appoints-a-new-board-member/>

<sup>27</sup> Meta, 'Where we have fact checking', *Meta for Media*, <https://www.facebook.com/formedia/mjp/programs/third-party-fact-checking/partner-map>

<sup>28</sup> J Machin, 'How Meta is preparing for the 2022 Australian election', *Facebook Australia blog*, 15 March 2022, <https://australia.fb.com/post/how-meta-is-preparing-for-the-2022-australian-election/>

<sup>29</sup> Agence France Presse Australia, Fact Check, <https://factcheck.afp.com/afp-australia>; Australian Associated Press; AAP Fact Check, <https://www.aap.com.au/category/factcheck>; RMIT FactLab <https://www.rmit.edu.au/about/schools-colleges/media-and-communication/industry/factlab/debunking-misinformation>

- **Meta applies a warning label to content found to be false by third-party fact-checking organisations.** We have maintained the approach outlined in our 2021 and 2022 transparency report.

Between 1 January and 31 December 2022, we displayed warnings on over 9 million distinct pieces of content on Facebook in Australia (including reshares) based on articles written by our third party fact checking partners.

- **Meta reduces the distribution of content found to be false by third-party fact-checking organisations.** We have maintained the approach outlined in our 2021 and 2022 transparency report.
- **Meta proactively searches for content that matches content debunked by our fact-checking partners, to apply the same treatments.**

Meta continues to proactively search for matches with content on the platform that has been fact-checked by our fact-checking partners. Between April and June 2022, we globally displayed warnings on over 200 million distinct pieces of content on Facebook (including re-shares) based on over 130,000 debunking articles written by our fact-checking partners.<sup>30</sup>

- **Meta takes action on Pages, Groups, accounts, or websites found to repeatedly share misinformation.** We have maintained the approach outlined in our 2021 and 2022 transparency report.
- **Meta takes steps to limit the possible distribution of misinformation via recommendations.** We have maintained the approach outlined in our 2021 and 2022 transparency report.

---

<sup>30</sup> M Bickert, 'Community Standards Enforcement Report, Second Quarter 2022, *Meta Newsroom*, 25 August 2022, <https://about.fb.com/news/2022/08/community-standards-enforcement-report-q2-2022/>

## Outcome 1b

Users will be informed about the types of behaviours and types of content that will be prohibited and/or managed by Signatories under this Code.

Signatories will implement and publish policies and procedures and any appropriate guidelines or information relating to the prohibition and/or management of user behaviours that may propagate Disinformation and Misinformation via their services or products.

- **Meta makes available a detailed list of claims that we consider to violate our COVID-19 Misinformation & Harm policy.**

We have maintained the approach outlined in our 2021 and 2022 transparency report. Meta's policy currently provides for removal of 80 distinct false claims about COVID-19 and vaccines.<sup>31</sup> Meta will continue to enforce on our Misinformation and Harm policy, and the 80 false claims, while we prepare our response to the Oversight Board's recommendations.

- **Meta makes information available via a dedicated website that outlines our efforts to combat misinformation.** We have maintained the approach outlined in our 2021 and 2022 transparency report. Meta's policies on misinformation can be found in our Transparency Centre.<sup>32</sup>

---

<sup>31</sup> Meta, 'COVID-19 policy updates and protections', *Facebook Help Centre*, <https://www.internmc.facebook.com/help/230764881494641>

<sup>32</sup> Meta, *Community Standards - Misinformation*, <https://transparency.fb.com/en-gb/>

## Outcome 1c

Users can report content and behaviours to Signatories that violates their policies under 5.10 through publicly available and accessible reporting tools.

Signatories will implement and publish policies, procedures and any appropriate guidelines or information regarding the reporting of the types of content and behaviours that may propagate Disinformation and Misinformation via their platforms.

- **Meta makes on-platform reporting channels available to users for false information.**

Meta continues to make reporting channels available to users on both Facebook and Instagram. As outlined in our 2021 and 2022 transparency report, users can follow the following steps to report misinformation:

Facebook Help Centre - <https://www.facebook.com/help/572838089565953>

Instagram Help Centre -

<https://www.facebook.com/help/instagram/2442045389198631>

Our Australian third-party fact-checking partners are also able to receive referrals from the public using the channels below:

- Agence France Presse: <https://factcheck.afp.com/contact>
- Australian Associated Press: <https://www.aap.com.au/make-a-submission/>
- RMIT FactLab:

<https://www.rmit.edu.au/about/schools-colleges/media-and-communication/industry/factlab/debunking-misinformation>



## Outcome 1d

Users will be able to access general information about Signatories' actions in response to reports made under 5.11.

Signatories will implement and publish policies, procedures and/or aggregated reports (including summaries of reports made under 5.11) regarding the detection and removal of content that violates platform policies, including but not necessarily limited to content on their platforms that qualifies as Misinformation and/or Disinformation.

- Meta makes global transparency reports available regularly.

Meta continues to make aggregated reports publicly available on our misinformation and disinformation efforts to provide oversight of our work.

As outlined in our 2021 and 2022 transparency reports, each quarter, we report on metrics for preventing and taking action on content that goes against our Community Standards.<sup>33</sup> This is outlined in Meta's Community Standards Enforcement Report (CSER).

In 2022, Meta published the results of an independent audit into Meta's CSER reporting, conducted by EY.

In 2020, we committed to undergoing an independent audit, conducted by EY, to validate that our metrics published in the CSER are measured and reported correctly. As part of the assessment, Meta provided EY with full access to the necessary data, documentation and evidence requests. We also gave access to dozens of employees across data science, data engineering, software engineers, product and program managers and Internal Audit teams working on the Community Standards Enforcement Report.

In May 2022, EY published its assessment and found the calculation of the metrics in our 2021 fourth quarter Community Standards Enforcement Report were fairly stated, and our internal controls are suitably designed and operating effectively.<sup>34</sup>

---

<sup>33</sup> Meta, *Community Standards Enforcement Report*, <https://about.fb.com/news/tag/coordinated-inauthentic-behavior/>

<sup>34</sup> Meta, *Community Standards Enforcement Report*, <https://about.fb.com/news/2022/05/community-standards-enforcement-report-assessment-results/> The results of the EY assessment can be found at <https://about.fb.com/wp-content/uploads/2022/05/EY-CSER-Independent-Assessment-Q4-2021.pdf>

- **Meta will supplement these reports with additional Australia-specific statistics, provided as part of this Annual Report process.**

While country-specific statistics should be interpreted with caution and have limitations in understanding misinformation and disinformation, Meta has again provided Australia-specific statistics in the spirit of transparency of our efforts.

**Australia – Number of pieces of harmful health misinformation content removed<sup>35</sup>**

2020 <sup>36</sup>	2021	2022
110,000	180,000	91,000

- **Meta makes the service CrowdTangle freely available to journalists, third-party fact-checking partners and some academics.**

Meta offers CrowdTangle to journalists, third-parties, fact-checking partners and eligible academics to analyse and report on trends in public content on social media.

As mentioned in our 2022 transparency report, Meta has also worked with CrowdTangle to build a platform where researchers and the Australian Strategic Policy Institute can access data about CIB networks.<sup>37</sup>

<sup>35</sup> These figures capture content from Pages or accounts specific to Australia (noting Australians benefitted from the content we removed from other countries as well).

<sup>36</sup> Meta began providing Australia-specific statistics on our misinformation enforcement in 2020.

<sup>37</sup> B Nimmo, D Agranovich & N Gleicher, 'Adversarial Threat Report', *Meta*, April 2022, [https://about.fb.com/wp-content/uploads/2022/04/Meta-Quarterly-Adversarial-Threat-Report\\_Q1-2022.pdf](https://about.fb.com/wp-content/uploads/2022/04/Meta-Quarterly-Adversarial-Threat-Report_Q1-2022.pdf)

## Outcome 2

Advertising and/or monetisation incentives for Disinformation are reduced.

Signatories will implement policies and processes that aim to disrupt advertising and/or monetisation incentives for Disinformation.

- **Meta sets a higher threshold for users to be able to advertise on our services, and takes action against users who spread misinformation.** We have maintained the approach outlined in our 2021 and 2022 transparency report.

### **Outcome 3**

**The risk that Inauthentic User Behaviours undermine the integrity and security of services and products is reduced.**

**Signatories commit to take measures that prohibit or manage the types of user behaviours that are designed to undermine the security and integrity of their services and products, for example, the use of fake accounts or automated bots that are designed to propagate Disinformation.**

Please see Outcome 1a for the actions Meta takes against inauthentic user behaviours.

## Outcome 4

Users are enabled to make more informed choices about the source of news and factual content accessed via digital platforms and are better equipped to identify Misinformation.

Signatories will implement measures to enable users to make informed choices about news and factual information and to access alternative sources of information.

- **Meta provides contextual information around posts that users see from public Pages.** We have maintained the approach outlined in our 2021 and 2022 transparency report.
- **Meta provides a COVID-19 Information Centre with verified, authoritative information about COVID-19.** Meta continues to provide the COVID-19 Information Centre around the world to promote authoritative information to Facebook users.

### Australia - number of people visiting the COVID-19 Information Centre<sup>38</sup>

	2020 (annual)	2021 (per quarter - Q4)	2022 (annual)
Global	2 billion	350 million	1.8 billion
Australia	6.2 million	3.5 million	7.5 million

- **Meta provides a Climate Science Information Centre in Australia to connect users to authoritative information from leading climate organisations.**

As noted in our 2022 transparency report, Meta launched the Climate Science Information Centre in Australia in November 2021.<sup>39</sup> The purpose of the Centre is to make sure people have access to reliable information while reducing climate misinformation.

<sup>38</sup> Note - only data for Q4 2021 was available at this time.

<sup>39</sup> N Clegg, 'Our commitment to combatting climate change', *Meta Newsroom*, 1 November 2021, <https://about.fb.com/news/2021/11/our-commitment-to-combating-climate-change/>

Between 1 January and 31 December 2022, there were over 84 million visits to the Climate Science Information Centre. Over 25,000 of these visits were from users in Australia.

- **Meta uses in-product prompts to direct Australians to authoritative information on key topics.**

During each election campaign we develop prompts to direct users to authoritative election information.

In the 2022 federal election we developed an ‘enrol to vote’ and ‘election day reminder’ prompt with the Australian Electoral Commission to raise awareness about the Australian election and link users to authoritative election information.

You can find more information about these prompts in the Case Study: Australian federal election 2022.

- **Meta promotes authoritative information by providing significant support to organisations such as the Australian, state and territory governments and First Nations health organisations to promote authoritative health information.**

We have maintained the approach outlined in our 2021 and 2022 transparency report. To date, Meta has given millions of dollars worth of Facebook advertising credits to federal, state and territory governments and NGOs, to increase access to COVID-19 vaccine and preventative health information.

- **Meta directs users to authoritative information when they search for high-priority topics on Facebook and Instagram.**

We have maintained the approach outlined in our 2021 and 2022 transparency report.

- **Meta directs users to authoritative information once they have seen or shared COVID-19 related misinformation.** We have maintained the approach outlined in our 2021 and 2022 transparency report.

- **Meta will look for opportunities to continue to work with the Government on other ways to promote authoritative information.** We have maintained the approach outlined in our 2021 and 2022 transparency report.

- **Meta promotes public service announcements to our users to encourage them to be wary of potential misinformation.** We continue to support the promotion of public service announcements on our platforms, working with government, industry and civil society organisations.
- **Meta will continue to provide greater transparency of our content ranking algorithms and give users more control over the content they see.**

Over the years, there has been a growing focus on the role of artificial intelligence (AI) and algorithms in determining how products and services are delivered to consumers. While the role of algorithms are increasingly universal, the public debate is often focussed on how algorithms rank and distribute content, and influence what people engage with online.

Meta firmly believes that users should have more transparency and control over the algorithms that present them with personalised content. We recognise that people and policymakers are only going to feel comfortable about these algorithmic systems if we provide more transparency around how they work, and give users control over the content they see.

For this reason, we have introduced new measures to give people more insight into and control over how content appears in their Feed.

First, at the user level we have updated the industry-leading **“Why Am I seeing This?”** (WAIST) Ad tool.

In 2014, Meta launched the “Why Am I Seeing This Ad?” tool, which was designed to provide users with more transparency and control over the ads they see in their Feed.<sup>40</sup>

In 2019, we launched “Why am I seeing this post?”, to help users better understand and control organic content from friends, Pages and Groups in their Feed. “Why am I seeing this post?” allows users to click on the post and view:

- Why they’re seeing a certain post in their Feed — for example, if the post is from a friend, a Group they’re part of, or a Page they follow.

---

<sup>40</sup> R Sethuraman, ‘Why am I seeing this? We have an answer for you’, *Meta Newsroom*, 31 March 2019, <https://about.fb.com/news/2019/03/why-am-i-seeing-this/>

- What information generally has the largest influence over the order of posts, including: (a) how often you interact with posts from people, Pages or Groups; (b) how often you interact with a specific type of post, for example, videos, photos or links; and (c) the popularity of the posts shared by the people, Pages and Groups you follow.
- Shortcuts to controls, such as See First, Unfollow, Feed Preferences and Privacy Shortcuts, to help users further personalise their Feed. This tool is available when a user clicks on the three dots on the top right of any ad. Users are able to understand why they are seeing an ad, including how factors like basic demographic details, interests and website visits contribute to the ads in Feed.

In January 2023, we increased the information available via the WAIST Ad tool to now include:<sup>41</sup>

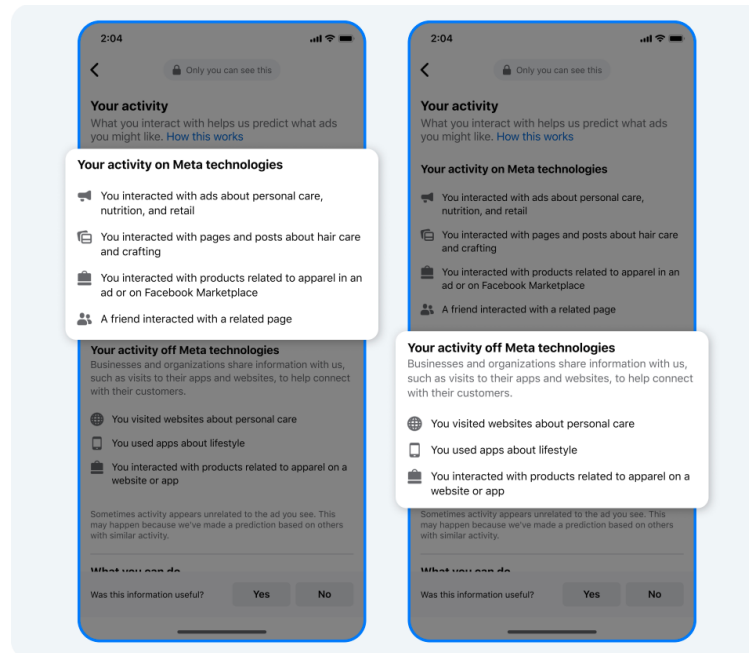
- Information summarised into topics about how your activity both on and off our technologies — such as liking a post on a friend’s Facebook page or interacting with your favorite sports website — may inform the machine learning models we use to shape and deliver the ads you see.
- New examples and illustrations explaining how our machine learning models connect various topics to show you relevant ads.

---

<sup>41</sup> P Pavon, ‘Increasing our ads transparency’, *Meta Newsroom*, 14 February 2023, <https://about.fb.com/news/2023/02/increasing-our-ads-transparency/>



## ‘Why Am I Seeing This Ad?’ tool



Second, as well as providing transparency at the user level, we recognise that there continue to be discussions about the best ways to provide model and systems documentation that enables transparency.

To that end, in 2022 we launched the AI System Card tool that is designed to provide insight into an AI system’s underlying architecture and help better explain how the AI operates.<sup>42</sup>

Many machine learning (ML) models are typically part of a larger AI system, a group of ML models, AI and non-AI technologies that work together to achieve specific tasks. Because ML models don’t always work in isolation to produce outcomes, and models may interact differently depending on what systems they’re a part of, model cards — a broadly accepted standard for model documentation — don’t paint a comprehensive picture of what an AI system does. For example, while our image classification models are all designed to predict what’s in a given image, they may be used differently in an integrity system that flags harmful content versus a recommender system used to show people posts they might be interested in.

<sup>42</sup> Meta AI, ‘System Cards, a new resource for understanding how AI systems work’, *Meta AI blog*, 23 February 2022, <https://ai.facebook.com/blog/system-cards-a-new-resource-for-understanding-how-ai-systems-work/>

Based on consultation and feedback from experts, Meta launched AI Systems Cards in 2022, beginning with the Instagram Feed ranking system.

In addition to these tools, Meta is a founding member of the Digital Trust and Safety Partnership (DTSP) which is developing approaches to evaluate digital platforms' content moderation practices and drive globally consistent trust and safety outcomes.<sup>43</sup>

The DTSP brings together internet platforms to agree on trust and safety best practices aimed at reducing content and conduct related risks, and is working through the development of an audit & oversight mechanism to independently assess platforms' implementation of those best practices. Google, Twitter, LinkedIn, Microsoft, Apple, Zoom, Pinterest, Vimeo, Reddit, Bitly and Patreon are also DTSP members.

In 2022, DTSP companies underwent a dry run audit called a 'self assessment' against the DTSP Best Practices Framework. Companies were required to report on five commitments:

- **Identify, evaluate and adjust for content- and conduct-related risks in product development.** The aim of this commitment is to ensure that companies engage in adequate forethought related to Content- and Conduct-Related Risks, and incorporate insights into product features accordingly.
- **Adopt explainable processes for product governance, including which team is responsible for creating rules, and how rules are evolved.** The aim of this commitment is to ensure the rules and principles governing user content and conduct are clear, rigorous, and consistent.
- **Conduct enforcement operations to implement product governance.** The aim of this commitment is to ensure operations exist to implement the aims set forth in Product Governance to address Content- and Conduct-Related Risks.
- **Assess and improve processes associated with content- and conduct-related risks.** The aim of this commitment is to ensure that mechanisms exist within the company to keep up with and respond to evolving Content- and Conduct-Related Risks and approaches.

---

<sup>43</sup> Digital Trust and Safety Partnership, <https://dtspartnership.org/>

- **Ensure that relevant trust & safety policies are published to the public, and report periodically to the public and other stakeholders regarding actions taken.** The aim of this commitment is to ensure the public and other stakeholders have insight into the company's Trust & Safety goals, challenges and activities.

In 2022, the DTSP released a 'Safe Assessment' evaluation report which synthesises the results of the self assessment, and identifies areas of success and improvement.<sup>44</sup>

In 2023, DTSP companies will be subject to their first independent audit.

We hope that a DTSP-like system will become a global, industry standard in the next couple of years. We believe this will be an effective mechanism for maintaining consistent, cross-jurisdictional metrics that can complement country-level regulatory requirements.

- **Meta will fund training for Australian journalists on how to identify and prevent amplifying mis and disinformation.**

In the lead up to the 2022 federal election, Meta supported First Draft (now RMIT FactLab) to run a series of training sessions for news publishers, so they could identify misinformation, avoid amplifying mis- or disinformation, and help connect Australians to credible information.

---

<sup>44</sup> DTSP Safe Assessments Report, *Digital Trust & Safety Partnership*, 2023, <https://dtspartnership.org/dtsp-safe-assessments-report/>

## **Case study: Australian federal election 2022**

In the lead up to the Australian election, Meta developed a comprehensive strategy that focussed on proactively detecting and removing content that breaches our policies, combatting misinformation, harmful content and CIB, and promoting civic participation.

Meta has been involved in more than 200 elections around the world since 2017, and we've learned key lessons from each one on combatting misinformation, election interference and promoting civic participation.

While some of these efforts were outlined in Meta's 2022 transparency report, this case study provides a comprehensive outline of our election integrity measures and their impact in Australia.

We have also been able to secure Australian-specific statistics on our efforts during the Australian election, between April 1 and June 30, 2022:

- We took action on over 25,000 pieces of content across Facebook & Instagram for violating our Harmful Health Misinformation policies.
- We displayed warnings on over 3 million distinct pieces of content on Facebook (including reshares) based on articles written by our third party fact checking partners.
- We took action on over 91,000 pieces of content on Facebook and over 40,000 pieces of content on Instagram in Australia for violating our hate speech policies.
- We took action on over 200,000 pieces of content on Facebook and over 46,000 of content on Instagram in Australia for violating our Community Standards on violence and incitement.
- We rejected around 17,000 ads for not complying with our political and social issue ads enforcement policies.

During each election campaign we develop prompts. In the 2022 federal election we developed two prompts on Facebook and Instagram election stickers to encourage people to vote, and direct users to the AEC website (see images below).

These prompts reached a very large number of Australians:

- The enrol to vote prompt was seen by 23.3 million Australians, with 54,000 people clicking through to the AEC website, and 900,000 people sharing it across their Feeds.
- The election day reminder was seen by almost 11 million Australians, clicked through 175,000 times and 60,000 people shared they had voted on their Feed.
- The Instagram election stickers were used on almost 35,000 Instagram stories.

We initiated a number of new programs of work in the lead up to the Australian election in relation to misinformation and disinformation.<sup>45</sup> This work includes:

#### *Combatting misinformation*

- **Increasing our fact-checking capability.** As well as onboarding RMIT FactLab as a new fact-checking partner, we provided one-off grants to our fact-checkers to increase their capacity in the lead up to the election.
- **Increasing misinformation monitoring.** Funded First Draft (now RMIT FactLab), the global misinformation and media literacy organisation, to undertake monitoring and analysis of misinformation trends, and provide pre-election training for Australian journalists on how to prevent amplifying misinformation.
- **Misinformation awareness campaigns.** Launched two national media literacy campaigns: a consumer-focussed campaign with Australian Associated Press called 'Check the Facts' and a creator-focussed awareness campaign with First Draft called 'Don't be a Misinfluencer'. For the first time, we translated the 'Check the Facts' campaign into three other languages – Vietnamese, Simplified Chinese and Arabic to extend the benefits of the campaign even further.

#### *Security, disinformation and election interference*

- **Sharing intelligence on potential foreign interference.** We worked closely with the Australian Electoral Commission (AEC), the Government's EIAT, and a range of government and law enforcement agencies. We also worked with civil society organisations, experts, and third-party fact-checkers in order to provide authoritative information about the election process.

In particular, in the lead up to the Australia election, we worked with the Australian Government's EIAT and developed responses to a series of possible scenarios that may arise during the course of an election campaign.

<sup>45</sup> Meta, 'How Meta is preparing for the 2022 Australian election', *Meta Australia blog*, 15 March 2022, <https://australia.fb.com/post/how-meta-is-preparing-for-the-2022-australian-election/>

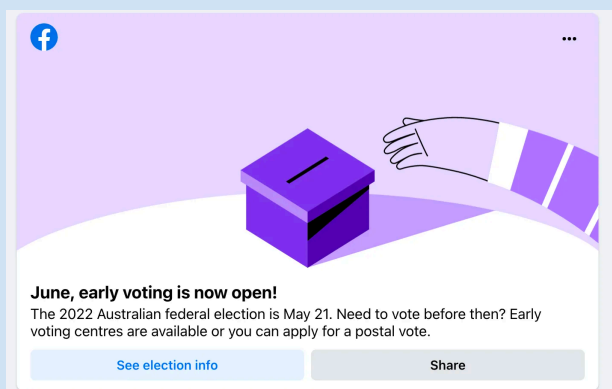
We also engaged with Australian security agencies in the year leading up to the Australian election campaign to share information and work together if threats arose during the campaign.

- **Increasing account security.** Launched the Facebook Protect security program in Australia. This program prompted those who might be at a higher risk of being targeted online, such as candidates and public officials, to adopt stronger account security protections. We saw a 90 per cent adoption rate of two-factor authentication in the first month that we rolled out the product.
- **Combatting CIB.** In our 2022 transparency report we reported that Meta had taken action on four instances of CIB operations that targeted Australians. This remains the same - during the 2022 federal election, we did not see any evidence of CIB targeting Australia.

#### *Encouraging civic engagement and empowering voters*

- **Training candidates to better reach their audience.** We ran training sessions with political parties on our policies and tools, and how to engage with voters.
- **Empowering people to participate - Facebook.** To raise awareness about the election, we also prompt Australians with an 'Election Day Reminder' and a 'Pre-polling Reminder' that link a user to the AEC website for more information.

#### **Australian 2022 federal election in-product prompts**



- **Empowering people to participate - Instagram.** We developed election stickers with the Australian Electoral Commission. Clicking on the stickers directed users

to the AEC website to find information like where and how to vote.  
We also launched a 'Election Reminder' at the top of the Instagram Feed on election day.

#### Australian 2022 federal election Instagram stickers



## Outcome 5

### Users are better informed about the source of Political Advertising

**Signatories will develop and implement policies that provide users with greater transparency about the source of Political Advertising carried on digital platforms.**

- **Meta requires all advertisers of political ads<sup>46</sup> to complete an ad authorisation, which includes verifying the advertiser’s identity.** We have maintained the approach outlined in our 2021 and 2022 transparency report.
- **Meta requires political ads to include a disclaimer disclosing who is paying for the ad.** We have maintained the approach outlined in our 2021 and 2022 transparency report.
- **Meta provides the Ad Library, a searchable archive of all political ads on our services in Australia.** We have maintained the approach outlined in our 2021 and 2022 transparency report.
- **Meta enables an Ad Library report that provides aggregated spend information about Pages undertaking political ads.** We have maintained the approach outlined in our 2021 and 2022 transparency report.
- **Meta will continue to add greater transparency around political and social issue advertising.**

As outlined in Meta’s 2021 and 2022 transparency reports, we require advertisers running social issue, electoral and political ads to complete our authorisation process and include “Paid for by” disclaimers, and we store these ads in our Ad Library for seven years. This will help ensure that ads on the platform are authentic and transparent.

In addition to the transparency information already available in our Ad Library – including the ad creative, who paid for the ad, and who that ad reached – in July 2022 we began including new information about the targeting selections made by

---

<sup>46</sup> We define political ads as advertisements: (1) made by, on behalf of, or about a candidate for public office, a political figure, a political party or advocates for the outcome of an election to public office; or (2) about any election, referendum or ballot initiative, including "go out and vote" or election campaigns. We recognise the definition of Political Advertising in the voluntary industry code for disinformation and misinformation is broader than Facebook’s definition of “political ads”, as it also encompasses ads that we refer to as “social issue ads”.



advertisers for all of their ads about social issues, elections and politics.

These insights are aggregated at the Page-level across the following categories:<sup>47</sup>

- Location - e.g. “100% of the amount spent was targeted to the Australia, which includes 2,450 ads about social issues, elections or politics”
- Age - e.g. “35% of the amount spent was targeted to people aged 55-years-old; 20% of the amount spent was targeted to people aged 21-years old”
- Gender - e.g. “75% of the amount spent was targeted to women”
- Detailed targeting such as interests, activities and education - e.g. “30% of the amount spent, or 280 ads, were targeted to people interested in sustainability.”
- Language - e.g. “78 social issue, electoral or political ads run were targeted to Spanish speakers, representing 15% of spend”
- We’ll also include whether a Page used Custom Audiences and/or lookalike audiences.

In updating our tools, we solicited advice from external experts in the academic community and civil society, and we will continue to do so as we evolve these tools.

By making advertiser targeting criteria available for analysis and reporting on ads run about social issues, elections and politics, we hope to help people better understand the practices used to reach potential voters on our technologies.

---

<sup>47</sup> J King, ‘Bringing more transparency to social issue, electoral and political ads, *Meta*, 23 May 2022, <https://www.facebook.com/business/news/transparency-social-issue-electoral-political-ads>

## Outcome 6

Signatories support the efforts of independent researchers to improve public understanding of Disinformation and Misinformation.

Signatories commit to support and encourage good faith independent efforts to research Disinformation and Misinformation both online and offline.

- **Meta supports research and events in relation to misinformation and media literacy.**

Since the 2022 transparency report, we have supported a number of events and research on misinformation, including:

- Sponsoring the Australian Media Literacy Alliance’s Media Literacy Summit. Meta has sponsored the Australian Media Literacy Alliance to host the inaugural Australian Media Literacy Summit, which will take place in March 2023.

The Summit aims to build a network of media literacy advocates and leaders who will work toward developing media literacy in schools and communities across Australia.

- Invested in new research on misinformation and integrity. Meta released a request for proposals (RFP) for research exploring integrity issues related to social communication technologies. The RFP attracted 503 proposals from 349 universities and institutions around the world.

We provided a total of \$1,000,000 USD in funding for 11 proposals, including research from the Queensland University of Technology by Michelle Riedlinger and Silvia Montaña-Niño, and Marina Joubert (Stellenbosch University), Víctor García-Perdomo (Universidad de La Sabana) on ‘Countering misinformation in the Southern Hemisphere: A comparative study’.<sup>48</sup>

---

<sup>48</sup> Meta Research, ‘Announcing the winners of the 2022 foundation integrity research request for proposals’, Meta Research, 28 March 2022, <https://research.facebook.com/blog/2023/2/announcing-the-winners-of-the-2022-foundational-integrity-research-request-for-proposals-/>

- **Meta will continue to support research and events in relation to disinformation.**

Meta will once again sponsor the Australian Strategic Policy Institute's 'Sydney Dialogue' which will take place in April 2023. This event is a high level two day event to discuss significant technology issues, including misinformation and disinformation.

- **Meta provides a free CrowdTangle public live display on COVID-19 publicly available to allow anybody to analyse public content on our platforms.**

We continue to provide access to CrowdTangle by news publishers, third-party fact-checking partners and approved academics.

We also continue to provide access to the COVID-19 public live display that tracks engagement on Facebook and Instagram. It is available at:

<https://apps.crowdtangle.com/covid19/boards/covid-19australia>.

In March 2022, we launched a new public, Live Display on Crowdtangle to provide real-time coverage of the Australian federal election.<sup>49</sup> Further detail can be found in the Case Study: Australian federal election 2022.

- **Meta collaborates with researchers to undertake surveys of our users to assess their views on topics such as vaccines and climate change.**

We continue to offer the COVID-19 Symptom Survey, the largest public health survey in history, with Carnegie-Mellon and University of Maryland.

In April 2022 we announced we were collaborating with UNICEF, the Yale Institute for Global Health (YIGH) and the Public Good Projects (PGP) to develop campaigns that drive COVID-19 vaccine acceptance in Kenya, India, Pakistan and Ukraine. These campaigns have now reached more than 155 million people.<sup>50</sup>

Our next steps will include randomised control trials to test whether digital campaigns are leading to higher immunisation rates overall and further illuminate the role that digital campaigns can play in advancing global public health goals.

---

<sup>49</sup> J Machin, 'How Meta is preparing for the 2022 Australian election', *Meta Australia blog*, 15 March 2022, <https://australia.fb.com/post/how-meta-is-preparing-for-the-2022-australian-election/>

<sup>50</sup> L McGorman, 'New lessons learned in building COVID-19 Vaccine acceptance', *Meta Newsroom*, 22 April 2022, <https://about.fb.com/news/2022/04/new-lessons-building-covid-19-vaccine-acceptance/>

- **Meta provides data to researchers in a privacy-protective way via the Facebook Open Research and Transparency (FORT) initiative.**

Since our 2022 transparency report, we have expanded the FORT initiative so that it is available to Australian researchers. We also expanded the data that is available within the FORT platform.

Australian academics can apply to be onboarded to the FORT platform, and will have access to up to date data sets on targeting information for ads about social issues, elections, and politics, URL share activity on Facebook, and users' engagement with Facebook's civic products.<sup>51</sup>

- **Meta will focus on new areas of research relating to misinformation and disinformation in 2022, including media literacy of First Nations peoples.**

In 2021, Meta supported the Australian Media Literacy Alliance's first media literacy survey and the report 'Towards a National Strategy for Media Literacy'. The report demonstrated that there is more work required to understand the media literacy of First Nations Australians.<sup>52</sup>

In response to this finding, Meta is partnering with Professor Tristan Kennedy from Monash University in 2023 to conduct research on the cultural nuances of misinformation and the impacts on Indigenous Australian communities, so that Meta can best tailor our programs to this area going forward. This research is currently in development.

---

<sup>51</sup> Facebook Open Research Transparency Initiative, 'Researcher datasets', *Meta*, <https://fort.fb.com/researcher-datasets>

<sup>52</sup> Australian Media Literacy Alliance, 'Towards a National Strategy for Media Literacy', *AMLA*, 25 October 2021, [https://medialiteracy.org.au/wp-content/uploads/2021/10/AMLA-Consultation-Workshop-Report\\_UPDATE-25-10-2021.pdf](https://medialiteracy.org.au/wp-content/uploads/2021/10/AMLA-Consultation-Workshop-Report_UPDATE-25-10-2021.pdf)

## Outcome 7

The public can access information about the measures Signatories have taken to combat Disinformation and Misinformation.

### All Signatories will make and publish the annual report information in section 7

- **Meta will continue to publish annual reports in Australia, such as these, to be transparent about the steps we are taking to combat disinformation and misinformation.** We met this requirement in 2021 and 2022 by publishing a transparency report on our Meta Australia blog, and speaking publicly to media about our work.<sup>53</sup>

The publication of this report satisfies this requirement for 2023.

---

<sup>53</sup> J Machin, 'Facebook's response to Australia's disinformation and misinformation code', *Meta Australia Blog*, 21 May 2021, <https://australia.fb.com/post/facebooks-response-to-australias-disinformation-and-misinformation-industry-code/>

## Appendix A: 2022 specific commitments made by Meta under the industry code on misinformation and disinformation

The 45 commitments for 2022 are outlined below:

### Outcome 1:

#### Combatting misinformation / disinformation

- Meta removes networks of accounts, Pages and Groups that violate our inauthentic behaviour policy, including disinformation that violates our policy on Inauthentic Behaviour.
- Meta provides transparency about accounts, Pages and Groups removed under our Inauthentic Behaviour policy.
- Meta partners with experts and organisations who assist in providing tips or further investigation about possible inauthentic behaviour on our services.
- Meta removes misinformation that violates the misinformation section of our Community Standards.
- Meta removes manipulated videos, also known as “deepfakes”, that violates our Manipulated Media policy.
- Meta removes material that violates our Violence-Inducing Conspiracy Network policy.
- Meta removes election-related misinformation that may constitute voter fraud or interference under our Co-ordinating Harm and Promoting Crime policy.
- Meta removes fake accounts.
- Meta allows for appeals in instances where users may disagree with our enforcement, including to the independent and external Oversight Board.
- Meta partners with third-party fact-checking organisations, globally and in Australia, to assess the accuracy of content on our services.
- Meta applies a warning label to content found to be false by third-party fact-checking organisations.
- Meta reduces the distribution of content found to be false by third-party fact-checking organisations.

	<ul style="list-style-type: none"> <li>• Meta proactively searches for content that matches content debunked by our fact-checking partners, to apply the same treatments.</li> <li>• Meta takes action on Pages, Groups, accounts or websites found to repeatedly share misinformation.</li> <li>• Meta takes steps to limit the possible distribution of misinformation via recommendations.</li> <li>• Meta makes available a detailed list of claims that we consider to violate our COVID-19 Misinformation &amp; Harm policy.</li> <li>• Meta makes information available via a dedicated website that outlines our efforts to combat misinformation.</li> <li>• Meta makes on-platform reporting channels available to users for false information.</li> <li>• Meta makes global transparency reports available regularly.</li> <li>• Meta will supplement these reports with additional Australia-specific statistics, provided as part of this Annual Report process.</li> <li>• Meta makes the service CrowdTangle freely available to journalists, third-party fact-checking partners, and some academics.</li> </ul>
<b>Outcome 2:</b>  <b>Disrupt monetisation and advertising incentives</b>	<ul style="list-style-type: none"> <li>• Meta sets a higher threshold for users to be able to advertise on our services, and takes action against users who spread misinformation.</li> </ul>
<b>Outcome 3:</b>  <b>Combat inauthentic user behaviour</b>	<p>See items listed under Outcome 1.</p>
<b>Outcome 4:</b>	<ul style="list-style-type: none"> <li>• Meta provides contextual information around posts that users see from public Pages.</li> </ul>

<p><b>Empower consumers to be informed</b></p>	<ul style="list-style-type: none"> <li>• Meta provides a COVID-19 Information Centre with verified, authoritative information about COVID-19.</li> <li>• Meta provides a Climate Science Information Centre in Australia to connect users to authoritative information from leading climate organisations.</li> <li>• Meta uses in-product prompts to direct Australians to authoritative information on key topics.</li> <li>• Meta promotes authoritative information by providing significant support to organisations such as the Australian, state and territory governments and First Nations health organisations to promote authoritative health information.</li> <li>• Meta directs users to authoritative information when they search for high-priority topics on Facebook and Instagram.</li> <li>• Meta directs users to authoritative information once they have seen or shared COVID-19 related misinformation.</li> <li>• Meta will look for opportunities to continue to work with the Government on other ways to promote authoritative information.</li> <li>• Meta promotes public service announcements to our users to encourage them to be wary of potential misinformation.</li> <li>• Meta will continue to provide greater transparency of our content ranking algorithms and give users more control over the content they see.</li> <li>• Meta will fund training for Australian journalists on how to identify and prevent amplifying mis and disinformation.</li> </ul>
<p><b>Outcome 5: Political advertising</b></p>	<ul style="list-style-type: none"> <li>• Meta requires all advertisers of political and social issue ads to complete an ad authorisation, which includes verifying the advertiser’s identity.</li> <li>• Meta requires political and social issue ads to include a disclaimer disclosing who is paying for the ad.</li> <li>• Meta provides the Ad Library, a searchable archive of all political and social issue ads on our services in Australia.</li> </ul>



	<ul style="list-style-type: none"> <li>• Meta enables an Ad Library report that provides aggregated spend information about Pages undertaking political and social issue ads.</li> <li>• Meta will continue to add greater transparency around political and social issue advertising.</li> </ul>
<b>Outcome 6: Research</b>	<ul style="list-style-type: none"> <li>• Meta supports research and events in relation to misinformation and media literacy.</li> <li>• Meta supports research and events in relation to disinformation.</li> <li>• Meta provides free public CrowdTangle live displays on key topics to allow anybody to track public content on our platforms.</li> <li>• Meta collaborates with researchers to undertake surveys of our users to assess their views on topics such as vaccines and climate change.</li> <li>• Meta provides data to researchers in a privacy-protective way via the Facebook Open Research and Transparency initiative.</li> <li>• Meta will focus on new areas of research relating to misinformation and disinformation in 2022, including media literacy of First Nations peoples.</li> </ul>
<b>Outcome 7: Annual reports</b>	<ul style="list-style-type: none"> <li>• Meta will continue to publish annual reports in Australia, such as these, to be transparent about the steps we are taking to combat disinformation and misinformation.</li> </ul>