

Meta Response: Science, Innovation & Technology Committee Social Media, Misinformation & Harmful Algorithms Inquiry

Introduction

Meta welcomes the opportunity to follow up with written responses to the five specific questions raised by members of the Committee at the oral evidence session on 24th March, 2026.

Keeping our services safe remains a top priority for the company. We continue to invest in our safety and security systems, including deploying more advanced technology for content enforcement and expanding our Teen Accounts protections.

We hope the information provided below is helpful to the Committee's ongoing work on these issues.

I. Prevalence of misinformation

Misinformation is different from other types of speech addressed in our Community Standards because there is no way to articulate a comprehensive list of what is prohibited. It is therefore not possible to provide prevalence figures for misinformation.

With graphic violence or hateful conduct, for instance, our policies specify the content that we prohibit. With misinformation, however, we cannot provide such a line. The world is changing constantly, and what is true one minute may not be true the next minute. People also have different levels of information about the world around them, and may believe something is true when it is not.

Instead, our policies articulate different categories of misinformation and try to provide clear guidance about how we treat that speech when we see it. For each category, our approach reflects our attempt to balance our values of expression, safety, dignity, authenticity and privacy.

We remove misinformation where it is likely to directly contribute to the risk of imminent physical harm. We also remove content that is likely to directly contribute to interference with the functioning of political processes. In determining what constitutes misinformation in these categories, we partner with independent experts who possess knowledge and expertise to assess the truth of the content and whether it is likely to directly contribute to the risk of imminent harm.

For all other misinformation, we focus on reducing its prevalence or creating an environment that fosters a productive dialogue. We know that people often use misinformation in harmless ways, such as to exaggerate a point ("This team has the worst record in the history of the sport!") or in humour or satire. Recognising how common such speech is, we focus on providing users with helpful information when there is potentially misleading or confusing content. As part of that effort in the UK, we partner with third-party fact-checking organisations to review and rate the accuracy of the most viral content on our platforms. We know that when a fact check warning screen is placed on a post, 95% of the time people do not click through to view it.

We also provide resources to increase media and digital literacy so that people can decide what to read, trust and share themselves. We require people to disclose, using our AI disclosure tool, whenever they post organic content with photorealistic video or realistic-sounding audio that was digitally created or altered, and we may apply penalties if they fail to do so. We may also add a label to certain digitally created or altered content that creates a particularly high risk of misleading people on a matter of public importance.

Finally, we prohibit content and behaviour in other areas that often overlap with the spread of misinformation. For example, our Community Standards prohibit fake accounts, fraud and coordinated inauthentic behaviour.

II. Internal governance systems

Meta operates a number of governance structures to ensure our policies are robust and consistently enforced. These include internal bodies, independent external oversight, and compliance with domestic regulatory frameworks.

The Oversight Board, an independent body established in 2020, provides external oversight of our content moderation. It comprises expert members from around the world - including former heads of state, judges, human rights experts, and journalists - who act independently of Meta. The Board reviews individual content decisions appealed by users or referred by Meta, and publishes binding decisions and makes policy recommendations and advisory opinions. Their analysis provides external scrutiny of how our systems and policies operate in practice. The Board's case decisions and public advisory opinions are all publicly available on their [website](#).

A meeting called the Policy Forum takes place on a regular basis where we discuss potential changes to our Community Standards and Advertising Standards. A variety of stakeholders participate in Policy Forum meetings. This includes team members from Safety and Security Policy, Global Operations, Civil Rights and Human Rights, Legal, Communications, as well as counterterrorism specialists, product managers and public policy leads.

At this meeting, subject matter experts from the Content Policy team propose adding new policies or amending existing ones. These proposals follow a process of policy development that includes [extensive engagement](#) across a range of worldwide stakeholders and a review of external and internal research. These meetings help the team better understand broad perspectives on safety and voice and the impact of our policies on communities globally.

In the UK, these structures operate alongside our domestic regulatory requirements under the Online Safety Act. We conduct annual risk assessments for these purposes, and comply with relevant duties under the Act. We continue to respond to information requests and consultations as the framework takes effect.

III. National Education Union report



Meta has reviewed the National Education Union report in full. The majority of its findings relate to platforms other than Instagram - and, where Instagram was assessed, the report concluded that content was “mostly age-appropriate.” The report’s recommendations are directed at Government and Meta will engage constructively with any legislative or regulatory process that follows.

On teen safety, Meta has taken significant proactive steps. We launched Teen Accounts in 2024 and introduced automatic default protections for all users under 18 - including restrictions on who can contact them, controls on the content they see, and parental tools to manage screen time. Meta has placed hundreds of millions of teens in Teen Accounts across Instagram, Facebook, and Messenger - with over 50 built-in protections resulting in teens seeing less sensitive content, experiencing less unwanted contact, and spending less time on Instagram at night.

We have also completed our annual UK-specific Protection of Children Risk Assessments and Illegal Harms assessment under the Online Safety Act, which we update whenever significant product changes are made.

IV. Young People, Revenue & Age Restrictions

All ads are defaulted to 18+ audiences - and an advertiser would have to actively choose to target an under 18 audience. Even then, teens on our platforms receive a heavily restricted advertising experience by default. Ad targeting is limited to age and location only. Meta’s Advertising Standards prohibit ads about restricted topics, such as alcohol, financial products, and weight loss services, and we do not use a teen’s prior activity or engagement to determine the ads they see. Under-18s represent less than 1% of Meta’s revenue globally.

In response to Australia’s recent decision to ban under-16s from social media, Meta deactivated 544,052 accounts in Australia that we believed were operated by users under 16 between 4th - 11th December 2025 - including 330,639 on Instagram, 173,497 on Facebook, and 39,916 on Threads.

Early evidence from Australia raises serious questions about the effectiveness of the ban. Our primary concern is that bans are ineffective and risk pushing teens toward less safe, unregulated services or logged-out experiences that bypass the protections we have built. We share the Australian Government’s goal of creating safe, age-appropriate online experiences, but believe a better approach is legislation that would require app stores to obtain parental approval whenever teens under 16 download apps - and to share that signal securely with platforms, so we can place young people into age-appropriate experiences from the start. This would give families the ability to decide which apps their children can access, while creating a joined-up system across the whole ecosystem.

V. Meta API for Researchers

Meta provides access to data - supporting independent research into our platforms - through several mechanisms. Our Meta Content Library and Content Library API provide approved researchers with access to near real-time public content data from Facebook, Instagram, and

Threads for public interest research, subject to appropriate privacy safeguards. Researchers must apply for access and meet eligibility criteria, including institutional affiliation and a defined research purpose.

We share findings from our own investigations publicly through our Adversarial Threat Reports, which provide transparency about adversarial threats we remove from our platforms, including coordinated inauthentic behaviour (CIB) networks. We also share data with researchers about CIB networks we remove through our Influence Operations Research Archive.

We have also provided extensive access to independent researchers in the past - including a major study whose findings were published in the journals [Science](#) and [Nature](#) - which examined the impact of our algorithms on political content and polarisation.

For our open-source AI models, we make model weights, architecture details, and model cards publicly available. This allows independent researchers, security professionals, and civil society organisations to inspect the models directly and conduct adversarial testing without requiring special permission or API access from Meta.

Our Ad Library allows anyone to search and view all ads currently running across our platforms. For political advertising, we apply additional transparency requirements: advertisers running ads about social issues, elections, or politics must confirm their identity and location, include "paid for by" disclaimers, and have their ads stored in the Ad Library where they are freely accessible to the public, researchers, journalists, and regulators.

April 2026