HEARING BEFORE
THE UNITED STATES SENATE
SELECT COMMITTEE ON INTELLIGENCE

September 18, 2024

Testimony of Nick Clegg

I. Introduction

Chairman Warner, Vice Chairman Rubio, and distinguished members of the Committee, thank you for the opportunity to appear before you today. My name is Nick Clegg, and I am the President, Global Affairs at Meta.

At Meta, we are committed to free expression. Each day, more than three billion people around the world use our apps to express themselves and make their voices heard. We want people to be able to talk openly about the issues that matter to them, whether through written comments, photos, music, or other creative media. By the end of this year, more than two billion people will have voted in elections across some of the world's largest democracies, including in the United States, and we are proud that our apps help people participate in the civic process.

We also recognize that adversaries are simultaneously working to interfere with elections through coordinated campaigns across the industry in an effort to undermine the democratic process. No tech company does more or invests more to protect elections online than Meta—not just during peak election seasons, but at all times. We have around 40,000 people working on our overall safety and security, and we have invested more than $20 billion in teams and technology in this area since 2016. Over the years, we have developed a comprehensive approach that sets out policies and safeguards for elections, including identifying threats and fighting manipulation and deception on our platforms.

This year, elections are taking place as more and more people are using artificial intelligence (AI) tools. As a company that has been at the forefront of AI development for more than a decade, we believe that progress and vigilance go hand in hand. We take seriously the concern of generative AI tools being misused during elections. We are committed to transparency in the use of AI, and we are making significant investments to further its responsible use. We are also working externally to address risks, including by signing on to both the White House's voluntary commitments and the Tech Accord to Combat Deceptive Use of AI in 2024 Elections. These commitments will help guide the industry toward safer, more secure, and more transparent development of AI technology, while helping to prevent and address harmful AI-generated content from interfering with elections.

While much of our approach to this year's elections reflects the knowledge we have gained from prior election cycles, we are committed to adapting where needed so that we can stay ahead of new challenges, including those presented by AI.

II. Our Election Integrity Efforts

Providing Access to Reliable Information

We think it is critical for people to be able to find reliable election information from trusted sources. We build free tools to support civic engagement, including ones to encourage eligible voters to register to vote, to remind them of deadlines, and to connect them with non-partisan resources.

We also provide people with election information from their state and local election officials, including during primaries. In the United States, when people search for terms related to the 2024 elections on Facebook and Instagram, they will see links to official information from state and local election officials about how, when, and where to vote.

The success of these efforts is a result of close communication with state and local election authorities, who provide important feedback that enables us to provide up-to-date and nonpartisan information in our decentralized election system. For example, based on feedback from the broader elections community, we developed Voting Alerts, a free tool that allows state and local election offices to broadcast key information to everyone in their jurisdictions. Since launching the initiative in 2020, state and local election officials have sent more than 650 million notifications through Voting Alerts on Facebook.

Prohibiting Harmful Content

In addition to connecting people with reliable voting information, we also prohibit misinformation that is likely to interfere directly with people's ability to vote, including misinformation about the dates, locations, times, and methods for voting; voter registration; and who is eligible to vote. Our policies prohibit calls for voter fraud and coordinated election interference. And we have invested in proactive threat detection and expanded our policies to combat harassment against election officials and poll workers online. In the United States and a number of other countries, we prohibit ads that discourage people from voting, call the legitimacy of an upcoming election into question, or contain premature claims of election victory. We have and will continually review and update these policies with election security in mind.

Promoting Transparency in Advertisements

We provide industry-leading transparency into political advertising on our services. Anyone who places an ad about politics, elections, or social issues must complete an authorization process and disclose who is paying for the ad. We add "paid for by" disclaimers and identify the owner and locations for political Pages and Groups. All political ads are archived in a publicly searchable Ad Library for 7 years, so anyone can see exactly what candidates are saying, who they are targeting, and who paid for it. Today, there are more than 15 million U.S. entries in our Ad Library.

We also require advertisers to disclose when they use AI or other digital techniques to create or alter a political or social issue ad that contains a photorealistic image or video, or

realistic sounding audio, that was digitally created or altered to depict a real person as saying or doing something they did not say or do. It also applies if an ad depicts a realistic-looking person that does not exist or a realistic-looking event that did not happen, alters footage of a real event, or depicts a realistic event that allegedly occurred but is not a true image, video, or audio recording of the event. If we determine that an advertiser has not disclosed the required information, we will reject the ad. Repeated failure to disclose required information may result in penalties against the advertiser. On both Instagram and Facebook, we give people the choice to adjust their Ad Preferences if they want to see fewer ads about social issues, elections, or politics. And in the United States, we prohibit new political, electoral, and social issue ads during the final week of an election.

III. Combating Manipulation and Deception

We know that foreign adversaries try to reach people on our platforms and others before elections, and we remain vigilant in our fight against their evolving tactics. We have made important investments to improve our ability to detect and stop foreign election interference and strengthen the security of our platforms. And we build increasingly sophisticated AI systems so we can proactively and successfully identify these abuses, for example by:

- **Preventing Coordinated Inauthentic Behavior.** We are constantly working to find and stop coordinated campaigns that seek to manipulate public debate across our platforms.

  Our Community Standards prohibit coordinated inauthentic behavior, which is when multiple accounts—including fake and authentic accounts—work together to mislead people. We do not want organizations or individuals creating networks of accounts that mislead people about who they are or what they are doing.

  This is an adversarial space, and we are often acting with imperfect information. We may not always get it right. So we need to be cautious and in each case, we need to conduct our own independent investigation to identify what is—and isn't—foreign interference.

  When we take down these accounts, it is because our investigation has identified deceptive behavior (like using networks of fake accounts to conceal their identity); it is not based on the identity of those behind the account or what they say. We've removed over 200 networks of coordinated inauthentic behavior since 2017, including networks from Russia, Iran, and China. Still, people continue to look for new ways to mislead people, which is why we continue to take steps to make it harder for them to do so.

- **Removing Fake Accounts and Banned Organizations.** One of the ways we identify and stop foreign interference is by proactively detecting and removing fake accounts. We also remove accounts that violate our policies and are not allowed to have a presence on our platform, such as foreign terrorist organizations and those designed under our Dangerous Organizations and Individuals policy.

- **Tackling Misinformation.** We are constantly working to stop the spread of misinformation and disinformation. We have built the largest independent

fact-checking network of any platform, with nearly 100 partners around the world to review and rate viral misinformation in more than 60 languages. Stories they rate as false are shown lower in Feed. If Pages repeatedly create or share misinformation, we significantly reduce their distribution and remove their advertising rights. We make these efforts regardless of the viewpoint of the content or its author.

I appreciate the opportunity to appear before you today with Microsoft and Google because it underscores an important point about addressing the threats we are facing: people who want to interfere in elections rarely target a single service or platform. Cross-industry collaboration, transparency, and reporting are essential to preventing and discouraging these networks from engaging in harmful conduct across the internet. That is why we publicize our takedowns of coordinated inauthentic behavior for all to see, provide information about them to third parties for their review, and share relevant information with researchers, academics, and others, including the Congress. In 2017, we started publishing detailed reporting on our work to detect and counter security threats on our platforms, known today as our Adversarial Threat Reports. We also publicly release threat indicators we identify on our GitHub platform. Today, we have compiled the largest repository of threat indicators, including more than 6,000 threat indicators of cross-internet activity by Doppelganger, the most persistent Russian foreign influence campaign.

As another recent example, we published our insights on a small cluster of malicious activity on WhatsApp that originated in Iran and appeared to have focused on political and diplomatic officials and other public figures. Our research suggests that these efforts were unsuccessful, and our security teams blocked the behavior after investigating user reports. In an abundance of caution, and given the heightened threat environment ahead of the US election, we also shared information about this malicious activity with law enforcement and the relevant presidential campaigns to encourage them to guard against potential adversarial behavior.

We continually adapt our platforms to make this kind of deception much more difficult and costly. When we conduct a takedown, we identify the tactics used and we build tools into our platforms to make those tactics more difficult at scale. By continuing to develop smarter technologies, enhance our defenses, improve transparency, and build strong partnerships, we are making the constant improvements we need to stay ahead of our adversaries and to protect the integrity of our platforms. We have also learned that we need to be cautious about seeing foreign interference where it is not. As we recently indicated, knowing what we know now, we would have taken different actions on certain issues in 2020.

IV. The Impact of Artificial Intelligence

Meta has been a pioneer in AI development for more than a decade, using machine learning to proactively identify and remove violating content across our services. As with election security, we know that AI progress and responsibility can and must go hand in hand. Generative AI tools offer huge opportunities, and we believe that it is possible and necessary for these technologies to be developed in a transparent and accountable way, while also working to minimize potential risks.

As detailed in a recent Adversarial Threat Report, we have not seen attempts on our apps to

use new generative AI tactics to subvert elections in ways that we could not address through our existing safeguards, specifically by disrupting adversarial networks behind them. However, this does not mean that people are not using AI to try to interfere in elections. To the contrary, adversaries have used different tools, such as AI-generated photos for profile photos on fake accounts, or AI to publish a large volume of fake articles resembling reputable news sources. We recently disrupted a campaign from Russia that was publishing a large volume of stories on fictitious news websites, which our investigation found were likely AI-generated summaries of original news articles. The same campaign also created fictitious journalist personas with generative adversarial network-created profile photos. Our teams found and removed many of these campaigns early, before they were able to build audiences and communities on our services. This shows that our industry's existing defenses already apply to novel generative AI, and are proving effective thus far.

However, we know that we must continue to monitor and assess risks with new technology. That is why we are continually adapting to address new challenges, including by advancing efforts to detect and label AI-generated media. We believe that providing transparency and additional context is the best way to address AI-generated content.

Earlier this year, we announced changes to our approach to identifying and labeling AI-generated organic content. This includes labeling a wider range of video, audio, and image content as "AI info" when we detect industry-standard AI image indicators or when people disclose that they are uploading AI-generated content. If we determine that digitally created or altered image, video, or audio content creates a particularly high risk of materially deceiving the public on a matter of importance, we may add a more prominent label. This approach gives people more information about the content so they can better assess it and appreciate the context if they see the same content elsewhere.

When photorealistic images are created using Meta's AI feature, we take several steps so that people know AI is involved, including putting visible markers on the images, applying "Imagined with AI" labels, and embedding both invisible watermarks and metadata within the image files. Using both invisible watermarking and metadata improves the effectiveness of these markers and helps other platforms identify them. We have been working with others in our industry to develop common standards for identifying AI-generated content through forums like the Partnership on AI (PAI) and the Coalition for Content Provenance and Authenticity. The invisible markers we use are in line with PAI's best practices.

We believe that our current approach represents the cutting edge of what is technically possible right now; however, we continue to pursue a range of options to improve our AI detection capabilities. This work is especially important as this is likely to become an increasingly adversarial space in years to come. People and organizations that actively want to deceive people with AI-generated content will look for ways around the safeguards that are put in place to detect it. Across our industry and society more generally, we will need to keep looking for ways to stay one step ahead.

Importantly, this issue is not unique to Meta and will require a whole-of-industry approach. We have collaborated with experts from technical, policy, media, legal, civic, and academic backgrounds to inform our policy development and processes. We also work closely with companies, such as Adobe, to develop technologies that make it possible for us and other platforms to share with people when they see content that has been AI-generated.

V. Conclusion

While we are conscious that every election brings its own challenges and complexities, we are confident that our comprehensive approach puts us in a strong position to do our part to help protect the integrity of not only this year's elections in the United States, but elections around the globe at all times. We look forward to continuing our work, as well as our collaboration with others in the industry, to drive transparency and counter potentially harmful threats to our democratic process.

Thank you, and I look forward to your questions.