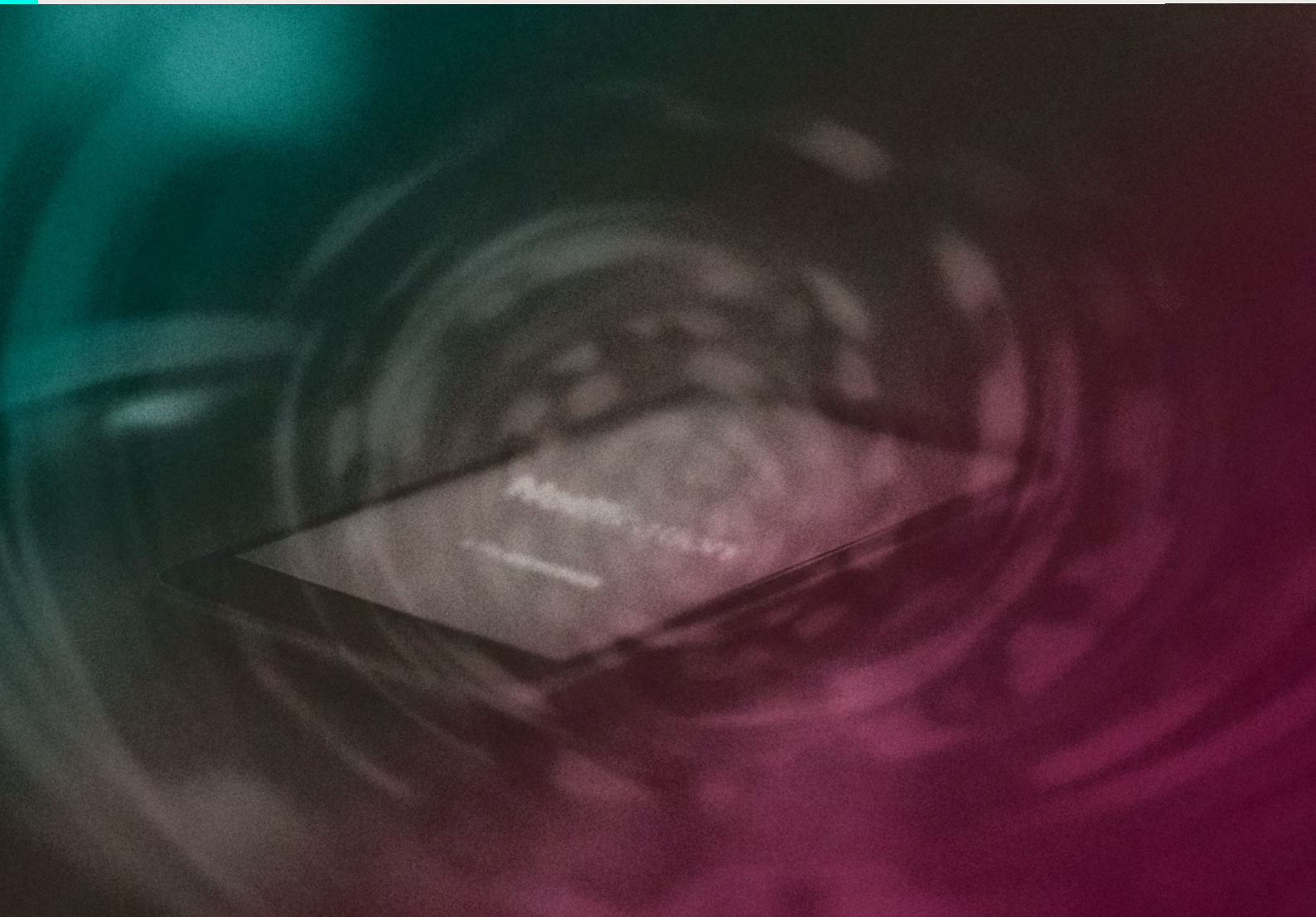


MAY 2025

Full Fact Report 2025



Full Fact Report 2025

Contents

About this report	6
Executive summary	7
Online misinformation	10
Chapter 1: Third-Party Fact Checking with Meta	11
Chapter 2: The 2024 UK riots	23
Chapter 3: The 2024 UK election	32
Chapter 4: Disinformation threats	43
Chapter 5: Health misinformation	50
Online misinformation: conclusion and rating	57
Legislation	58
Chapter 6: Harmful misinformation in UK legislation	59
Chapter 7: AI regulation	70
Legislation: conclusion and rating	80
Platforms	81
Chapter 8: Platform partnerships with fact checkers	82
Chapter 9: Policies for harmful misinformation	96
Platforms: conclusion and rating	102
Interventions	103
Chapter 10: Digital and media literacy	104
Chapter 11: Building a better information environment	114
Interventions: conclusion and ratings	122

About this report

Full Fact wants to build a better information environment to restore trust. Our editorial work exposes false and misleading claims and helps promote accurate content, raising standards in public debate and allowing people to make informed choices. Our technology uses generative AI to monitor and detect misinformation at internet scale, allowing small groups of people to find, check and challenge the most harmful claims. Our AI tools have been used in 40 countries worldwide in English, French and Arabic.

This report assesses the state of misinformation in the UK. It explores how political change has created new challenges for those tackling misinformation and disinformation, and examines the tasks facing government, regulators and online platforms as a result.

It follows on from our 2024 report *Truth and trust in the age of AI* and our 2023 report *Informed citizens: Addressing bad information in a healthy democracy*. It is the sixth annual report we have been able to produce thanks to the generous support of the Nuffield Foundation.

The [Nuffield Foundation](#) is an independent charitable trust with a mission to advance social wellbeing. It funds research that informs social policy, primarily in education, welfare, and justice. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute and the Nuffield Family Justice Observatory. The Foundation has funded this project, but has no influence on what the report says.

The report and its contents are the responsibility of the Chief Executive. They do not necessarily reflect the views of members of Full Fact's cross-party Board of Trustees.

We thank our supporters, our trustees and other Full Fact volunteers. Full details of [our funding](#) are available on our website.¹



¹ Full Fact, "Funding", <https://fullfact.org/about/funding/>

Executive summary

Full Fact's 2025 report is being published at a moment of crisis for anyone who cares about verifiable facts—a time of global upheaval, as the second Trump administration rewrites the rules of American engagement and western political norms. Fact checking organisations around the world—which seek to amplify accurate information amidst a deluge of false, misleading or artificially generated junk—are under pressure as never before. Many may not survive.

But this is also a time to stand up for our values. Full Fact is an impartial charity, but we will not be impartial about the proposition that facts matter—not only for those of us who work at Full Fact but for us all. The ability to identify, verify, and think critically about information is essential to any meaningful public debate in the UK.

Yet today, the United States is charting a different course. Earlier this year, Vice-President JD Vance came to Europe to talk about the enemy within. He described misinformation as an ugly Soviet era word, and suggested anyone using it wanted to tell others what to think. As we set out at the time,² we strongly disagree. Fact checking doesn't restrict debate; it strengthens it by grounding it in truth. It's not censorship. It's more speech, not less—and by that standard, the Vice-President should approve.

We have always been robust defenders of freedom of expression. But we believe free speech is not absolute. It is equally important to protect people from serious harm online. This is a difficult balance to get right, but if everything becomes a matter of opinion—if it is always my facts versus your facts—nothing can ever be questioned or debated effectively.

So we are proud that this report, with a focus on the UK, tackles the real and growing threat of misinformation: false or misleading information—often unintentional—that spreads and can cause real harm. It also touches on disinformation: falsehoods spread deliberately to deceive and damage people, communities, or entire countries.

In the US, pressure from the White House has made even the word misinformation politically charged. In April 2025, the US National Science Foundation abruptly terminated dozens of grants³ worth many millions of dollars it had previously awarded to researchers studying misinformation, and less contentious phrases like 'information integrity' or 'information credibility' are now seen as safer options. But debating the language risks missing the real issue: our online information environment is under greater threat than ever before and we must step up our response.

² Chris Morris, "JD Vance is wrong, facts are not opinions", Full Fact, 19 February 2025, <https://fullfact.org/blog/2025/feb/jd-vance-facts-matter/>

³ Alexios Mantzarlis, "NSF takes Ignorance is Strength approach to misinformation research", Faked Up, 23 April 2025, <https://fakedup.org/nsf-decides-ignorance-is-strength-on-misinfo-research/?ref=faked-up-newsletter>

What does this mean in practice? There has been prolonged debate in the UK in recent months about defence spending, and the need to increase it. But defence is not just about bullets and tanks; it's also about bots and troll farms. We are in a hybrid war, with attacks coming from some hitherto unexpected places, and if we want to protect what we value in our society we need to fight on all fronts. Access to accurate information forms the basis of the robust political debates we need to have. It is not a luxury, it is the foundation of our democracy.

That is why we are so concerned that the large online platforms, which wield so much influence over our daily lives, may see an opportunity to walk away from commitments to make our online world a safer place. We want the UK government and regulators like Ofcom to do more to hold these companies to account, by law if necessary. This is no time for half measures.

This year's report begins by assessing what our work has revealed about some of the biggest events of the past year—including the UK general election and the riots that took place in the summer of 2024—as well as the daily torrent of misleading and/or synthetic information which appears online as a matter of routine. We also assess the current state of legislation in the UK, covering both online safety and artificial intelligence, and we suggest much-needed improvements.

We examine the policy choices of the online platforms, and urge them to live up to their responsibility to ensure that their users are protected from harm online; the UK regulator, Ofcom, needs to hold them properly to account. Finally, we look at ways in which we have sought to intervene to improve the information environment: we call again on politicians to lead by example and act quickly to correct their mistakes; and we highlight the positive potential of technology in the deployment of our AI tools, which monitor millions of sentences across the internet every day.

For the first time, our report includes a series of guest essays from experts in various aspects of the work we do, including politicians, academics and activists. We are grateful for their contributions—the opinions they express are theirs, but the responsibility for publication is ours. We are also introducing a rating system to assess:

- the overall state of online misinformation
- the legislative and regulatory response in the UK
- the role of online platforms
- efforts to improve the information environment

We intend to return to these assessments in subsequent reports.

This is a critical year. At a time when the generative AI revolution continues to gather pace, the UK needs to ensure that accurate information is made available in a timely fashion to as many people as possible. The expertise of impartial fact checking organisations is not part of the problem. It is part of the solution.

Our recommendations

To government

1. The government must resist pressure from the Trump administration's agenda when drafting new laws on online safety and the regulation of AI. Legislation should focus on protecting UK citizens from harmful misinformation and disinformation rather than currying political favour.
2. If platforms reduce collaboration with fact checkers, the government should demand clarity: How will they counter misinformation? What data will they share to ensure the public still gets timely, accurate information? There is a need for greater transparency and accountability.
3. Misinformation must be treated as a legislative priority, even when it does not meet the threshold of illegal content. The government should revisit proposals for the Online Safety Act, including protections against health misinformation, content neutral solutions, and a statutory media literacy duty for platforms.

To platforms

4. Meta should not abandon third-party fact checking globally, and it can still reverse the decision to end its programme in the US. Prioritising policies to counter misinformation, alongside trusted voices, is vital.
5. As platforms develop Community Notes models, they must collaborate with high-quality, independent fact checkers—experts who are well funded and can act quickly. Their input is crucial when consensus cannot be reached.

To Ofcom

6. The new Online Information Advisory Committee must be proactive, vocal and engaged. It should lead on recommending how protections against misinformation can be enshrined in law.
7. Ofcom's media literacy work should be expanded to reflect today's challenges. It should cover all age groups and address emerging threats, especially those driven by generative AI.

Full Fact's work is only possible thanks to the support of thousands of individuals. **For updates and the opportunity to help us build a better information environment, join us: fullfact.org/signup**

Online misinformation

The passage of new online safety legislation in the UK has done little to reduce the spread of harmful online misinformation. The riots in the summer of 2024, following the murders of three young girls in Southport, illustrated both the limitations of the Online Safety Act and the task that lies ahead for this new Labour government.

The global context has also become significantly more challenging. One of the most consequential acts affecting the fight against online misinformation was Meta's decision to terminate its Third-Party Fact Checking programme in the United States. This report is written against the backdrop of that decision. How will fact checkers combat the harmful content they see if the online platforms that dominate its distribution are unwilling to work with them? And what will happen if Meta extends this decision to the United Kingdom and the rest of the world?

We will explore these questions throughout the report, but in this first section we highlight the types of claims our team are seeing on a daily basis. We examine the impact of online misinformation over the past year and outline the landscape in which the new Labour government is operating. These chapters are not an exhaustive list of everything we have seen, but they summarise some of the most pressing issues.

In this section of the report and in the chapters that follow, we seek to answer the question of what can be done to solve the online misinformation crisis. From the government and the online platforms, to healthcare officials and members of the public, we all need to take responsibility for what we're sharing, hosting and consuming online. Most importantly, we make clear that misinformation needs to be taken seriously, regulated rigorously, and managed proactively to ensure that its real-world consequences are reduced.

Chapter 1: Third-Party Fact Checking with Meta

Introduction

2025 began with one of the world's largest and most influential companies abruptly changing course. "Fact checkers have just been too politically biased, and have destroyed more trust than they've created,"⁴ said Meta's Founder, Chairman and Chief Executive Mark Zuckerberg in a 7 January 2025 video announcement that the company would end its Third-Party Fact Checking (TPFC) programme in the United States.

Meta's accusation of political bias and censorship was deeply disappointing to the fact checking community, and a sudden reversal of years of policy. The evidence provided by Meta to the UK government mere weeks before Mr Zuckerberg's announcement described its TPFC programme as a "key part of our approach to combating misinformation."⁵ Even after Meta changed course, it continued to highlight its partnership with independent fact checkers outside the United States, for example during the campaign for the 2025 Australian election.⁶

Meta's U-turn on its collaboration with fact checkers in the United States came at a specific political moment, with the return of President Trump to the White House, and should be seen in that light. For our part, Full Fact wholeheartedly rejects Mr Zuckerberg's claim of bias. Meta has provided no proof for his claim and we see no reason to overturn a system of independent fact checking that puts reliable, evidence-based verdicts at users' fingertips.⁷

As members of the European Fact-Checking Standards Network (EFCSN) and the International Fact Checking Network (IFCN), our impartiality is rigorously assessed and verified. And, as the code of standards for the EFCSN makes clear, members are "committed to upholding the principles of freedom of expression"⁸ and must have a "proven track record of excellence, integrity and accountability".⁹

⁴ Joel Kaplan, "More Speech and Fewer Mistakes", Meta, 7 January 2025, <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>

⁵ Science, Innovation and Technology Committee inquiry into social media, misinformation and harmful algorithms, written evidence submitted by Meta, 18 December 2024, <https://committees.parliament.uk/writtenevidence/132928/pdf#page=10>

⁶ Cheryl Seeto, "How Meta is preparing for the Australian federal election", Medium, 18 March 2025, <https://medium.com/meta-australia-policy-blog/how-meta-is-preparing-for-the-australian-federal-election-2f773a53ea79>

⁷ Team Full Fact, "Full Fact responds to Meta's Community Notes Plan", Full Fact, 13 March 2025, <https://fullfact.org/blog/2025/mar/response-to-meta-community-notes-plan/>

⁸ European Fact-Checking Standards Network, Accessed 16 April 2025, <https://efcsn.com/>

⁹ European Fact-Checking Standards Network, "Code of Standards", Accessed 16 April 2025, <https://efcsn.com/code-of-standards/>

Since partnering with Meta in January 2019,¹⁰ Full Fact has checked more than 2,750 cases that include misleading, faked, and potentially harmful posts on Facebook and other platforms. This chapter will outline some of the recurring themes our team has seen over the last 12 months in our TPFC work, in the hope that these lessons may support the work of others in the field. These themes include increases in manipulated/synthetic AI-generated content, a recurrence of old footage and imagery in breaking news contexts to sow confusion and mislead the public, and the impersonation of high-profile people.

We will also explore the implications of Meta's change of policy on the wider platform sector, and how it may reflect a broader shift in approach to content moderation.

Comment

Khaled Mansour, writer and novelist. Khaled serves on Meta's independent Oversight Board but this article represents his own views.

Over a few horrifying days in March, as many as 1,000 people were killed in Syria's coastal area, most of them civilians and from the minority Alawite Shiite sect. What the new government in Damascus described as an attempted coup by the remnants of the fallen Assad regime who belonged to this sect, has led to what human rights advocates described as a murderous rampage seemingly by pro-government militias in which hundreds of civilians including women and children were killed after their sectarian identity was identified.

The fledgling government promised a fact-finding commission. This commission will not do justice to its mandate if it does not fully investigate how social media played a pivotal role in this carnage. There is plenty of evidence.

Ahmad Brimo, the founder of Syria Verify, one of this country's main fact checking operations, says many of the accounts that have been spreading harmful falsehoods were run by individuals or companies and packaged in such a way as to give the impression that they speak on behalf of a certain religious or ethnic group. A widely used technique by sectarian spokespeople relies on the use of fake or manipulated videos and photos—some of which are taken out of context—to convey hateful content in order to incite violence and fuel hostilities across social divisions. Many followers embrace and amplify such content, most believing it to be true.

It has long become evident that social media can turn into a weapon in complex conflicts, like Syria, where sectarian, regional, ethnic and other affiliations are the rallying cry in confrontations to settle what are largely political, economic and social tensions. Social

¹⁰ Team Full Fact, "Full Fact to start checking Facebook content as third-party factchecking initiative reaches the UK", Full Fact, 11 January 2019, <https://fullfact.org/blog/2019/jan/full-fact-start-checking-facebook-content-third-party-factchecking-initiative-reaches-uk/>

media content is thus often used to mobilise, recruit, fundraise, and organise for violent acts. It is probably worse on Telegram and closed WhatsApp groups compared to more public platforms such as Facebook and X (Twitter). Granted, social media platforms also serve as a tool to build bridges, dismiss rumours and debunk disinformation. They are obviously double-edged swords, but the edge that spills blood is worthy of more attention.

To separate fact from destructive fiction, the root solution is to have discerning and critical users who can sift through a ceaseless flood of images and text. Such users cannot be easily fooled. As angry as they may be at their nemeses, they would not like or share such content. Still, in this day and age, such users seem to be in a minority, especially in the heat of a conflict. This puts an additional burden on social media platforms to tackle incitement to violence and harmful disinformation more effectively, especially during conflicts (and we have a dozen of them from Ukraine to Gaza to the DRC right now). These platforms must moderate content and demote, label or even remove, likely harmful pieces especially when they seem to be going viral or are peddled by suspicious accounts and groups with a sizeable following.

I experienced first-hand how volunteer users in Syria debunk falsehoods or reveal how fake photos and videos left unchecked could lead to more mayhem. But volunteers do not have the same tools or credibility that professional fact checkers do, nor do they come close to having the same impact that the platforms themselves can bring about if they dedicate more resources and effective tools to fight this scourge.

Fact checking is not a binary approach to truth and it cannot easily be substituted by free crowd-sourced labour from amongst the users themselves, as is the case with X's Community Notes approach, which replaced various safety measures that the platform has gutted after Elon Musk took over. Reporters without Borders¹¹ claims that X has consequently turned into a "disinformation stronghold".

This underlines the duty of social media platforms to deploy effective systems to counter misinformation and disinformation. Very large platforms using AI-powered tools need to provide better labelling of potentially harmful content and no longer amplify it with their automatic recommenders. They should stop acting as megaphones for disinformation and borderline content in their ceaseless pursuit of more engagement to push ad revenues up. They are best positioned to uncover deepfakes, manipulated posts, and coordinated campaigns that could lead to real-world harm. In addition to strengthened internal systems, the platforms may then deploy other tools, from crowd-sourced systems to trusted fact checkers.

¹¹ Reporters Without Borders, "From Twitter to X, Elon Musk's transformation from free speech defender to champion of disinformation", 23 October 2023, <https://rsf.org/en/twitter-x-elon-musk-s-transformation-free-speech-defender-champion-disinformation>

As I write, I cherish freedom of expression as essential to my creativity. Meanwhile, as an aid worker who witnessed conflicts from Afghanistan to the Sudan and in between for many years, I believe there is a strong need for information integrity and credible sourcing to avoid causing more harm and deepening animosities. Misinformation, disinformation and hate speech—including dehumanisation—can very much kill as we have seen in Rwanda, Myanmar and now in Syria.

Fact checking is not a panacea against disinformation. It must be coupled with internal algorithms that are effective at scale, while public interest organisations work more intensively on equipping users from an early age to consume information critically. Ultimately, when bad information floods a community it very much undermines people's trust in each other and in public institutions. This erodes the very foundations for which freedom of expression is such a prized right.

Key misinformation themes identified through our work with Meta's Third-Party Fact Checking Programme over the last year

Increase in manipulated synthetic (AI) content

Over the last year, much of our work with Meta's TPFC programme has focused on combatting misinformation during high-stakes, global events such as the Russia-Ukraine war, the conflict in the Middle East, and the LA wildfires in early 2025. As events unfolded across media outlets and social media platforms in real time, some users inadvertently shared misinformation, some of which was AI-generated content. Others spread deliberate disinformation in order to sow confusion.

AI-generated imagery, shared as if it is real, can gradually erode trust in information online. This is why our work in fact checking online claims, even those that seem egregiously or obviously false, is crucial to maintaining the integrity of the online information environment.

A few examples to show what we mean:

- An image depicting a bearded man looking up fearfully from an underground passage. This was shared more than a thousand times on social media after Syria's former President Bashar al-Assad was ousted, as Syrian rebels seized the capital Damascus unopposed in December 2024.¹² According to our research, it is not a real photo, but comes from an unrelated video uploaded to TikTok and created using AI.

¹² Evie Townend, "Image of prisoner found 'deep underground' in Syria is AI", Full Fact, 9 December 2024, <https://fullfact.org/online/image-artificially-intelligence-syria-man-underground/>

- An image we identified in October 2024 appeared to show four Israel Defense Forces (IDF) soldiers with their hands behind their backs, supposedly captured in southern Lebanon. This image was almost certainly AI-generated, with discrepancies such as unnaturally long feet, a rifle that appeared to have barrels at either end, and garbled text on the soldiers' backs. Full Fact found no credible, recent reports of soldiers from the IDF being captured by Hezbollah in southern Lebanon at the time.¹³
- During the LA wildfires, fake pictures of the landmark Hollywood sign on fire began to circulate online after the news that the fire had extended into the hills around it.¹⁴ Although the Hollywood Sign Trust confirmed on Instagram that the sign “continues to stand tall” and did not catch fire, this didn’t stop the misleading images from being shared nearly 3,000 times.

An uncropped version of one of the images showing the Hollywood sign engulfed in flames included the watermark for ‘Grok’, the generative AI chatbot created by Elon Musk’s startup xAI, indicating it was synthetically produced.¹⁵ At a time when people are being asked to evacuate, accurate information about the spread of wildfires plays an important role in ensuring compliance with evacuation orders. Inaccurate information can cause confusion or delay.

- Other examples of AI-generated misinformation during the LA wildfires came in the form of images of ‘miracle houses’ that were supposedly unaffected by the fires surrounding them. One image of a blue-roofed house that apparently survived the fires also featured the Grok watermark, and another—upon a reverse image search—said it was “Made with Google AI” in the ‘About this image’ section of its metadata, suggesting it was either modified with or created by Google Artificial Intelligence tools.¹⁶

Individual examples can appear to be of limited significance. But the scale of deception on the internet is staggering, and the need to respond is clear. Full Fact has previously published guidance on how to spot AI-generated content, including tips such as being vigilant about how realistic a scenario might be, looking for inconsistencies within the image, or even doing a reverse image search to check whether the image has appeared elsewhere online.¹⁷

¹³ Charlotte Green, “Image supposedly showing Israeli soldiers taken prisoner by Hezbollah almost certainly AI-generated”, Full Fact, 18 October 2024, <https://fullfact.org/news/idf-soldiers-prisoners-image-ai-created/>

¹⁴ Sian Bayley, “Fact checking the LA wildfires”, Full Fact, 5 February 2025, <https://fullfact.org/blog/2025/feb/fact-checking-the-la-wildfires/>

¹⁵ Charlotte Green, “Burning Hollywood sign photo isn’t genuine”, Full Fact, 10 January 2025, <https://fullfact.org/us/hollywood-sign-la-burning-ai-generated/>

¹⁶ Charlotte Green, “Image claiming to show a ‘miracle house’ that survived LA fires is likely AI-creation”, Full Fact, 22 January 2025, <https://fullfact.org/us/la-fires-miracle-house-saved-likely-ai-creation/>

¹⁷ Grace Rahman, “How to spot AI-generated images”, Full Fact, 5 April 2023, <https://fullfact.org/online/how-to-spot-ai-images/>

Re-framing and re-occurrence of old footage in new (false) contexts, particularly when there's breaking news

Another popular tactic is the republication of old footage or imagery in a new context, with accompanying posts or descriptions that imply it is taken from a current event. When authentic footage from previous events is repurposed in this way, it can cause confusion and divert the public's attention from accurate, real-time updates from reliable sources.

- During the UK riots in the summer of 2024, a screenshot of a TikTok video that falsely claimed to show Hindu and Sikh protesters, marching against illegal immigration, was actually footage from a Hindu religious festival procession through London.¹⁸
- A video of missiles hitting ships, shared with captions that could be interpreted to mean it depicted real missile attacks in the Red Sea, was actually from the military simulation video game Arma 3.¹⁹
- A video circulated on Facebook in October 2024 with the caption “Massive explosion reportedly at the Mossad headquarters in Tel Aviv,” actually dates from 2015, and shows a chemical blast at a warehouse in Tianjin, China.²⁰
- A video claiming to show Ukrainian troops surrendering in the Kursk region on 11 March 2025 re-purposed footage from 2022.²¹
- A video showing Ukrainian soldiers faking combat to appear “war torn” in order to receive US funds was actually footage from a music video about the war.²²

Reframed footage of this kind is designed to mislead people, and it can be convincing because it is “real”—not AI-generated or synthetic. But over the long term, it erodes trust as viewers become rightly concerned that they cannot take anything at face value. That damages trust in media, including citizen journalism in which members of the public capture real footage or imagery of breaking news that traditional media outlets sometimes syndicate or include in their reporting.

Impersonation of public figures

Deepfake technology is becoming more sophisticated and more dangerous. With easy-to-use tools, anyone can now edit video and audio to produce convincing impersonations of public figures, resulting in a wave of viral fakes designed to provoke, mislead or confuse. Recent examples include:

¹⁸ Charlotte Green, “Screenshot is from video of Hindu festival not anti-immigration march”, Full Fact, 9 August 2024, <https://fullfact.org/news/hindu-festival-london-not-anti-immigration-march/>

¹⁹ Nasim Asl, “Clip of missiles hitting ships is from video game Arma 3 (not the Red Sea)”, Full Fact, 16 February 2024, <https://fullfact.org/online/arma-3-ship-missiles/>

²⁰ Charlotte Green, “Video of explosion is from China not Tel Aviv”, Full Fact, 3 October 2024, <https://fullfact.org/online/explosion-video-china-not-tel-aviv-israel/>

²¹ Evie Townend, “Old video shared as recent footage of Ukrainian troops ‘surrendering’ in Kursk”, Full Fact, 14 March 2025, <https://fullfact.org/online/ukraine-soldiers-surrendering-old-video/>

²² Evie Townend, “Viral clip shows filming of a music video, not Ukrainian soldiers ‘faking combat’ to secure US funds”, Full Fact, 10 March 2025, <https://fullfact.org/online/false-claim-ukraine-soldiers-combat/>

- A video supposedly showing celebrities, including Scarlett Johansson, Drake and Jerry Seinfeld wearing t-shirts protesting against Kanye West.²³
- A fake clip of Taylor Swift saying the wildfires in Los Angeles were “divine retribution” for the US funding missiles used in Gaza.²⁴
- A video of Donald Trump supposedly calling for a ban on Skittles and Twizzlers because they contain the red food dye carmine.²⁵
- A video appearing to feature an audio recording of President Donald Trump criticising Keir Starmer over aid to Ukraine, energy costs and jobs.²⁶

These manipulated clips aren't harmless. They undermine trust in what we see and hear online, and can spark real-world unrest. In some cases, they also damage the reputation and credibility of the person being impersonated.

Dissemination of hoax posts

Despite previous warnings by Full Fact following an initial investigation in 2023, hoax posts continue to inundate community Facebook groups across the UK. These posts typically feature emotive or alarming information to generate attention, such as claims about missing or found elderly people, children, or pets.²⁷

Comment

Tony Thompson, Journalist and Fact Checker with Full Fact

Following years of rapid growth, fraud has become by far the most commonly experienced crime in the UK. It currently accounts for 40% of all offences (in England and Wales) but this is most likely a significant underestimate—the Crime Survey for England and Wales estimates only 13% of cases are ever reported.²⁸

While murders, muggings and crimes of sexual violence dominate the headlines, you are far more likely to be a victim of fraud than any other crime.²⁹ More and more people are being forced to contend with a daily onslaught of scam texts, phishing emails, spoofed calls and fake adverts on social media, all designed to separate them from their money.

²³ Charlotte Green, “Viral video of celebrities wearing t-shirts protesting against Kanye West is AI deep fake”, Full Fact, 14 February 2025, <https://fullfact.org/online/kanye-west-celebrities-video-fake/>

²⁴ Sian Bayley, “Video supposedly showing Taylor Swift calling LA fires ‘divine retribution’ has been altered”, Full Fact, 27 January 2025, <https://fullfact.org/online/taylor-swift-la-fires-deepfake/>

²⁵ Jess Hacker, “No, Donald Trump hasn’t called for Skittles to be banned”, Full Fact, 3 April 2025, <https://fullfact.org/health/trump-deepfake-skittles-red-carmine/>

²⁶ Sian Bayley, “Audio of President Trump criticising Keir Starmer is fake”, Full Fact, 29 January 2025, <https://fullfact.org/online/donald-trump-keir-starmer-uk-ukraine-ai/>

²⁷ Charlotte Green, “Hoax posts still rife on Facebook 18 months on from Full Fact’s investigation”, Full Fact, 10 March 2025, <https://fullfact.org/blog/2025/mar/hoax-posts-still-rife-on-facebook-18-months-on-from-full-facts-investigation/>

²⁸ ‘Fraud’, National Crime Agency, Accessed 16 April 2025, <https://www.nationalcrimeagency.gov.uk/what-we-do/crime-threats/fraud-and-economic-crime>

²⁹ Jess Sharp, “People more likely to be victim of fraud than any other crime, says Her Majesty’s Inspectorate of Constabulary and Fire & Rescue Services”, Sky News, 5 August 2021, <https://news.sky.com/story/people-more-likely-to-be-a-victim-of-fraud-than-any-other-crime-says-her-majestys-inspectorate-of-constabulary-and-fire-and-rescue-services-12372631>

Little wonder then that Full Fact's work in the online misinformation space will regularly cross over with this kind of activity. While some people choose to spread misinformation to create mischief or enhance reputations, others do so purely for financial gain.

One clear version of this kind of activity can be seen in posts that make false claims about the availability of discounted items online. For example, we've seen posts claiming that major retailers including Amazon,³⁰ Argos³¹ and Lidl³² are selling off items such as laptops, pressure washers or Smeg kettles for highly discounted prices.

Clicking on the links attached to such posts usually transports users to a website that closely replicates the branding of a legitimate retailer, but those who enter bank or credit card details can have money withdrawn from their accounts, only to find the promised goods never arrive.

We also regularly see hoax posts on Facebook that seek to attract engagement by featuring highly emotive stories of missing dogs,³³ children³⁴ or elderly relatives, and implore readers to share the appeals as widely as possible. Once a certain level of engagement has been reached, such posts are typically edited into housing scams or pages offering financial deals.

Our research has found that some of those involved in hoax posts make money by directing people to other websites via hidden affiliate links.³⁵ The final destination of such links may be a legitimate company offering, for example, cashback services. These companies are the victims as they are paying out affiliate fees to scammers who are breaching the terms and conditions of the services they offer.

Because the fraud takes place away from the social media platform itself, it is less likely that such posts will get removed for breaching the terms and conditions of Meta or its competitors.

More recently we have seen a rise in the use of generative AI to create online misinformation across a wide range of platforms.³⁶ The technology has also been

³⁰ Nasim Asl, "No, Amazon is not gifting laptops to people aged 40 or over", Full Fact, 26 February 2025, <https://fullfact.org/online/amazon-laptop-giveaway-40s/>

³¹ Nasim Asl, "No, Argos is not selling pressure washers for less than £2", Full Fact 21 March 2025, <https://fullfact.org/online/argos-pressure-washer-fake/>

³² Full Fact Team, "Lidl is not selling Smeg kettles for £3", Full Fact, 30 January 2025, <https://fullfact.org/online/lidl-smeg-kettle-3-pounds/>

³³ Charlotte Green, "Hoax lost dog picture recirculates online", Full Fact, 25 February 2025, <https://fullfact.org/online/hoax-lost-dog-facebook-groups/>

³⁴ Evie Townend, "Facebook posts about 'found' boy taken to police station are hoaxes", 11 February 2025, <https://fullfact.org/online/online-hoax-post-missing-child-police-station/>

³⁵ Tony Thompson, "Missing children, lost dogs and escaped snakes: how hoax posts are swamping local Facebook groups", Full Fact, 24 August 2023, <https://fullfact.org/online/facebook-hoax-posts-deception/>

³⁶ Sian Bayley, "Picture of Bono and Bob Geldof holding Israeli flags is AI-generated", Full Fact, 20 December 2024, <https://fullfact.org/online/bono-bob-geldof-israeli-flags-ai-generated/> (and) Charlotte Green, "Burning Hollywood sign photo isn't genuine", Full Fact, 20 January 2025, <https://fullfact.org/us/hollywood-sign-la-burning-ai-generated/>

adopted by fraudsters and scam artists who are using it for everything from writing more realistic posts to generating images and videos of false products and services.³⁷

A recent spate of false celebrity endorsements of cryptocurrency schemes made extensive use of AI technology to create deepfake videos of public figures, encouraging people to invest.³⁸

Though beyond the scope of Full Fact's own work in misinformation, some of our partners track the same scammers as they target people through private messages, claiming, for example, to be friends or relatives who have lost their phones or passports and require emergency cash in order to be able to return home from holiday.³⁹

What can be done? Increasing public awareness of the many ways in which scammers operate would help. As would stricter controls about creating social media accounts.

Facebook is strict about users having only one profile account that is under their real name. In order to get around this many scammers use Facebook pages rather than profiles. Aimed at businesses, Facebook Pages are indistinguishable from personal profiles in many ways. From the scammers point of view, the benefit is that there is no limit to the number of pages an individual may create and that there is no link between a page and the original profile that created it.

The Online Safety Act now requires online services to assess the risk of fraud on their platforms, and remove illegal content when they are told about it.⁴⁰ Time will tell how this works in practice.

In November 2023, following a Full Fact investigation into hoax posts, the then government (along with leading social media companies) pledged to take additional action to block and remove fraudulent content from their sites. In March 2025, however, new research by Full Fact found that the kind of hoax posts we had seen in our 2023 investigation were still rife across Facebook.⁴¹ Our research discovered at least 47 communities across the UK had been the victim of nine different hoaxes, including Facebook groups for big cities like Belfast, Edinburgh and Manchester and smaller places

³⁷ Dan Milmo, Alex Hern, "AI will make scam emails look genuine, UK cybersecurity agency warns", The Guardian, 24 January 2024, <https://www.theguardian.com/technology/2024/jan/24/ai-scam-emails-uk-cybersecurity-agency-phishing>

³⁸ Simon Goodly, "Revealed: the scammers who conned savers out of \$35m using fake celebrity ads", The Guardian, 5 March 2025, <https://www.theguardian.com/money/2025/mar/05/revealed-the-scammers-who-conned-savers-out-of-35m-using-fake-celebrity-ads> (and) Bea Swallow, "Woman loses £20k through AI investment scam", BBC News, 30 November 2024, <https://www.bbc.co.uk/news/articles/c1wjwdwjdxdo>

³⁹ NatWest International, "What's a WhatsApp impersonation scam?" <https://www.natwestinternational.com/global/fraud-and-security/spotting-scams/whatsapp-family-impersonation-scam.html>

⁴⁰ Ofcom, "Helping to tackle fraud under the new online safety regime", 12 February 2024, <https://www.ofcom.org.uk/online-safety/online-fraud/helping-tackle-fraud-under-new-online-safety-regime/>

⁴¹ Charlotte Green, "Hoax posts still rife on Facebook 18 months on from Full Fact's investigation", 12 March 2025, <https://fullfact.org/blog/2025/mar/hoax-posts-still-rife-on-facebook-18-months-on-from-full-facts-investigation/>

like Banbury, Melton Mowbray and Oldham. We wrote to Meta to urge them to take meaningful action against a pernicious problem that continues to spread on its platform⁴² and have yet to receive a response.

Among the most recent hoaxes Full Fact identified were four alarmist posts that aimed to scare communities rather than generate empathy: bogus warnings about a “serial killer”,⁴³ a man who’d supposedly murdered two police officers,⁴⁴ an alleged knife attacker,⁴⁵ and claims that a woman had been found stabbed by a local canal.⁴⁶

These hoax posts risk undermining genuine appeals and authentic warnings from well meaning community members in these groups, and render these environments useless as avenues for meaningful local communication.

We’ve previously issued guidance on how to spot a hoax in local Facebook groups online, including disabled comments, posts from pages rather than individual profiles, and cultural references outside the UK which might suggest the post was copied from a similar hoax in a different country.⁴⁷

What does the end of Meta’s Third-Party Fact Checking Programme mean for online safety?

Full Fact’s involvement in the Third-Party Fact Checking (TPFC) programme has had a substantial impact on Meta’s ability to protect users from the harms of misinformation and disinformation. By adding crucial context and credible information to thousands of posts, we’ve helped millions better understand what they are reading and seeing.

It’s impossible for us to quantify the exact impact we have had, because we are not given access to Meta’s own data, but we know that our work has helped to reduce the impact of tens of thousands of misleading and potentially harmful posts on Meta platforms over the last six years. We have never had, nor do we seek, the ability to remove information from the internet.

⁴² Mark Frankel, “Letter from Full Fact to Meta”, Full Fact, 11 March 2025, https://fullfact.org/media/uploads/meta_03_2025.pdf

⁴³ Tony Thompson, “Facebook warnings about roaming ‘serial killer’ are fake”, Full Fact, 25 February 2025, <https://fullfact.org/online/serial-killer-roaming-england/>

⁴⁴ Tony Thompson, “Posts warning a killer is on the loose are false”, Full Fact, 28 February 2025, <https://fullfact.org/online/hoax-posts-murder-two-police-officers/>

⁴⁵ Tony Thompson, “Warnings of a knife-wielding teenage killer on the loose are hoaxes”, Full Fact, 31 March 2025, <https://fullfact.org/online/teenage-killer-knife-hoax/>

⁴⁶ Evie Townend “Facebook hoax posts share old photos with claims women found ‘stabbed’ by canals in UK towns”, Full Fact, 26 February 2025, <https://fullfact.org/online/woman-stabbed-hoax-post/>

⁴⁷ Grace Rahman and Tony Thompson, “Seven ways to spot if a Facebook post is a hoax”, Full Fact, 31 August 2023, <https://fullfact.org/blog/2023/aug/seven-ways-to-spot-a-hoax/>

Instead, we try to focus on identifying and addressing the most viral, high-risk forms of misinformation. Our fact checks appear across Meta's platforms, offering clear evidence-based explanations, annotated with detailed and informative research, so users can make their own decisions about what to believe—without infringing on free speech.

Meta is now following X's lead and piloting a Community Notes model. We explore the detail in Chapter 8—but, in short, while Community Notes can be part of a wider solution, crowdsourcing opinions and showcasing competing points of view is no substitute for independent, non-partisan fact checking.

Meta's version of Community Notes may also lack transparency and accountability—it plans to keep contributors anonymous at first.⁴⁸ The system prioritises consensus over factual accuracy, meaning that even if a post is clearly harmful or misleading, it may remain visible in news feeds without consequences.⁴⁹

In the absence of a structured partnership that enables fast, independent fact checks, Meta's version of Community Notes is likely to fall short. Research by Spanish fact checkers Maldita shows that on X, fact checking organisations are the third most cited source in Community Notes,⁵⁰ indicating that users still rely on them to challenge misinformation, and that their work remains a trusted part of content moderation.

Meta's changes signal a broader shift in content moderation and online safety

Full Fact has often described fact checkers as first responders in the information environment.⁵¹ But as Meta rolls back parts of its TPFC programme, it is also making broader changes that weaken content moderation across its platforms. The company announced plans to drop policies on immigration, gender identity and diversity, and to stop proactively enforcing some policies on harmful content.⁵² Meta's Chief Global Affairs Officer Joel Kaplan confirmed the changes to its hate speech policies were "implemented worldwide immediately."⁵³

⁴⁸ Meta, "Testing Begins for Community Notes on Facebook, Instagram and Threads", 13 March 2025, <https://about.fb.com/news/2025/03/testing-begins-community-notes-facebook-instagram-threads/>

⁴⁹ Joel Kaplan, "More Speech and Fewer Mistakes", Meta, 7 January 2025, <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>

⁵⁰ Maldita.ES, "Faster, trusted, and more useful: The impact of fact-checkers in X's Community Notes", February 2025, 5, https://files.maldita.es/maldita/uploads/2025/02/maldita_informe_community_notes_2024.pdf#page=4

⁵¹ Chris Morris, "Full Fact responds to Meta ending support for US fact checkers", Full Fact, 7 January 2025, <https://fullfact.org/blog/2025/jan/meta-ending-support-for-us-fact-checkers/>

⁵² Joel Kaplan, "More Speech and Fewer Mistakes", Meta, 7 January 2025, <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>

⁵³ Meta, "Joel Kaplan on EU Regulation and Innovation", 7 February 2025, <https://about.fb.com/news/2025/02/joel-kaplan-on-eu-regulation-and-innovation/>

That means less oversight of potential misinformation. The Centre for Countering Digital Hate (CCDH) argues that these changes “could mean lots more harmful content circulating on Meta platforms.” CCDH’s research says in 2024 over 97% of Meta’s enforcement actions—accounting for nearly 277 million pieces of content—were proactive, leading to fears that ending this approach will undermine online safety.⁵⁴

Elsewhere, other platforms have backed away from strong governing frameworks that protect users from misinformation online, just as they were being absorbed in EU legislation. Google, LinkedIn, and YouTube all withdrew from the European Union’s Code of Practice on Disinformation earlier this year, before it became a binding Code of Conduct.⁵⁵ We explore these developments in more detail in Chapter 8, but the implications are clear: less moderation, and less governance, is likely to produce more bad information, and greater harm.

⁵⁴ CCDH, “Meta’s rollback of safety measures has big implications for social media users in the UK”, Center for Countering Digital Hate, 24 February 2025, <https://counterhate.com/blog/metas-rollback-of-safety-measures-has-big-implications-for-social-media-users-in-the-uk/>

⁵⁵ European Fact-Checking Standards Network, “EFCSN Statement on Platforms’ Reduced Commitments to the Code of Practice on Disinformation”, 22 January 2025, https://efcsn.com/news/2025-01-22_efcsn-statement-on-platforms-reduced-commitments-to-the-code-of-practice-on-disinformation/

Chapter 2: The 2024 UK riots



Credit: Roland Lloyd Parry / AFP

Introduction

Widespread disorder broke out in July and August 2024, after three young girls were killed in a horrific knife attack during a Taylor Swift-themed dance class in Southport, north of Liverpool.⁵⁶ Calls for protests were amplified by networks of social media influencers that falsely linked illegal immigration to the attacker,⁵⁷ who was initially rumoured to have arrived in the UK on a small boat.⁵⁸

The 2024 riots were a clear example of a tragedy on home soil spiralling into serious further harm and civil unrest, fueled in large part by unchecked misinformation. There were other contributory factors, but the rapid dissemination of false information that followed the stabbings helped create a climate that led to violence against mosques, police officers and asylum seekers, and the subsequent arrest of more than 1,200 protesters.⁵⁹

⁵⁶ Jeremy Culley and Hafsa Khalil, "Southport stabbings - what we know about attack", BBC News, updated 31 July 2024, <https://www.bbc.co.uk/news/articles/cy68z9dw9e7o>

⁵⁷ Dominic Casciani and BBC Verify, "Violent Southport protests reveal organising tactics of the far-right", BBC News, 2 August 2024, <https://www.bbc.co.uk/news/articles/cl4y0453nv5o>

⁵⁸ Sian Bayley, "What role did misinformation play in riots after the Southport stabbings?", Full Fact, 2 August 2024, <https://fullfact.org/news/misinformation-southport-stabbings/>

⁵⁹ William Downs, "Policing response to the 2024 summer riots", House of Commons Library, 9 September 2024, <https://commonslibrary.parliament.uk/policing-response-to-the-2024-summer-riots/>

Full Fact was actively involved in fact checking the riots, and ensuring misinformation was flagged and where possible, corrected.⁶⁰ We issued more than a dozen fact checks in the days after the stabbings, including a rebuttal of the false image that claimed to show a group of men with “knives and swords” in Stoke, which was actually a still from a video of men celebrating a Yemeni wedding in Birmingham with ceremonial daggers.⁶¹ There was also a fabricated article, purporting to be from the Telegraph, headlined “Keir Starmer considering building ‘emergency detainment camps’ on the Falkland Islands.”⁶²

This chapter will focus on what we learned about the spread of online misinformation in the aftermath of the Southport attack. We will consider how false information was disseminated and amplified so rapidly, what else the UK authorities could have done to prevent the escalation in violence, and the inability of online platforms to detect and respond to rapidly emerging harms that needs to be addressed with specific regulation.

Comment

Zoe Manzi and Hannah Rose, Hate and Extremism Analysts at the Institute for Strategic Dialogue

The failure of social media platforms to curb the spread of false narratives in a timely manner, during the riots which took place after the Southport murders last year, may have significantly contributed to the offline violence and disruption which subsequently erupted across the UK.

Immediately after the attack, false claims began to emerge on X (formerly Twitter), TikTok and Facebook, erroneously identifying the perpetrator as a Muslim migrant, “Ali al-Shakati.”⁶³

Influential figures with large numbers of followers, including actor-turned-political activist Laurence Fox, further amplified this narrative, using it to call for anti-Muslim action, including the permanent removal of Islam from Great Britain. His post,⁶⁴ which amassed over 850,000 views in the first 48 hours after the attack, exemplifies how misinformation is weaponised to incite hate. On X, such posts from paid premium users may be given preference by the platform recommender algorithm, allowing them to reach larger audiences. These findings demand investigation into how Terms of Service

⁶⁰ Full Fact, “Supplementary written evidence submitted by Full Fact”, Science, Innovation and Technology’s Inquiry on What are the links between social media algorithms, generative AI and the spread of harmful content online?, 26 February 2025, <https://committees.parliament.uk/writtenevidence/138329/html/>

⁶¹ Sian Bayley, “Picture of men with knives is unrelated to recent riots”, Full Fact, 8 August 2024, <https://fullfact.org/online/men-dancing-knives-stoke-riots/>

⁶² Sian Bayley, “The Telegraph has not published an article about ‘emergency detainment camps’ in the Falklands”, Full Fact, 8 August 2024, <https://fullfact.org/online/telegraph-fake-article-detainment-camps/>

⁶³ Institute for Strategic Dialogue, “From rumours to riots: How online misinformation fuelled violence in the aftermath of the Southport attack”, 31 July 2024, https://www.isdglobal.org/digital_dispatches/from-rumours-to-riots-how-online-misinformation-fuelled-violence-in-the-aftermath-of-the-southport-attack/

⁶⁴ Ibid.

are applied to verified users, who should receive enhanced scrutiny during crises to prevent the amplification of harmful disinformation.

Despite police taking the unprecedented step of confirming the alleged perpetrator was a local 17-year-old, misinformation continued to circulate. TikTok's search recommendations actively surfaced misinformation, suggesting queries like 'Ali al-Shakati arrested in Southport' long after the claim had been disproven. Repeating this exercise months later, analysts were still served conspiratorial content and disinformation about the Southport attack through the recommender algorithm.⁶⁵ Transparency gaps persist in understanding the role of recommender systems in amplifying harmful content.⁶⁶ While the EU's Digital Services Act (DSA) legislates limited independent auditing of these systems, the UK's Online Safety Act (OSA) does not, leaving UK users more vulnerable than our European neighbours.⁶⁷

Permissive platform environments allowed hate speech and conspiracy theories linking immigration to crime to spread and far-right networks to mobilise unhindered. On X, the use of anti-Muslim slurs more than doubled⁶⁸ in the ten days following the Southport attack, with over 40,000 mentions. Across British far-right Telegram channels, anti-Muslim hate rose 276% and anti-migrant hate 246%.⁶⁹ One X user with 16,000 followers and X premium status posted a protest flyer asserting that 'children are being sacrificed on the unchecked altar of mass migration.' These narratives attempt to provide justification for real-world violence, further demonstrating how misinformation and hate speech can have direct offline consequences.

To prevent similar incidents, platforms must develop explicit crisis response protocols to ensure rapid detection and mitigation of harmful misinformation and disinformation.⁷⁰ These should include surge capacity during high-risk events, improved coordination with authorities, and a balance between swift action and human rights safeguards. Greater algorithmic transparency and auditing are needed to provide insight into how

⁶⁵ Institute for Strategic Dialogue, "ISD Written Evidence to the Science, Innovation and Technology Committee Inquiry, on Social Media, Misinformation and Harmful Algorithms", 20 January 2025, <https://www.isdglobal.org/isd-publications/isd-written-evidence-to-the-science-innovation-and-technology-committee-inquiry-on-social-media-misinformation-and-harmful-algorithms/>

⁶⁶ Sara Bundtzen, "'Suggested for You': Understanding How Algorithmic Ranking Practices Affect Online Discourses and Assessing Proposed Alternatives", Institute for Strategic Dialogue, 9 December 2022, <https://www.isdglobal.org/isd-publications/suggested-for-you-understanding-how-algorithmic-ranking-practices-affect-online-discourses-and-assessing-proposed-alternatives/>

⁶⁷ Helena Schwertheim, "Transparency", Institute for Strategic Dialogue, 21 July 2023, <https://www.isdglobal.org/explainers/transparency/>

⁶⁸ ISD & CASM Technology, "Evidencing a rise in anti-Muslim and anti-migrant online hate following the Southport attack", Institute for Strategic Dialogue, 3 September 2024, https://www.isdglobal.org/digital_dispatches/evidencing-a-rise-in-anti-muslim-and-anti-migrant-online-hate-following-the-southport-attack/

⁶⁹ Ibid.

⁷⁰ Institute for Strategic Dialogue, "After Southport: Policy responses to far-right extremism", 15 August 2024, https://www.isdglobal.org/digital_dispatches/after-southport-policy-responses-to-farright-extremism/

recommendation systems amplify content during crises,⁷¹ as the lack of independent oversight in the UK leaves users at greater risk of exposure to harmful content. More consistent enforcement of platform policies is also essential to prevent verified accounts and those with large followings from receiving preferential treatment that allows harmful misinformation to spread unchecked. Platforms must improve access to data for researchers and regulators, enabling external monitoring of harmful content trends and the effectiveness of moderation practices. Without meaningful access, addressing online harms remains difficult. Additionally, financial incentives that allow disinformation actors to profit must be addressed. Monetisation policies should be reviewed to prevent bad actors from gaining financial benefits through engagement-driven misinformation.

The speed at which false narratives spread, their amplification by recommendation algorithms, and the delayed response by social media platforms enabled a climate where digital propaganda fuelled real-world violence. The riots which took place following the knife attack in Southport last summer illustrate the urgent need for greater platform accountability and legislative and regulatory clarity. Without enhanced transparency and robust enforcement of platform policies, similar incidents may occur. Addressing these challenges requires ongoing collaboration to ensure that online spaces do not become incubators for violence and social unrest and to mitigate the real-world harms of online disinformation.

Understanding the root problem of how misinformation spreads online is complex and multifaceted, as are the solutions to tackle it. But taking steps to understand it is a significant challenge when the Online Safety Act falls short in regulating misinformation, and therefore fails to create any urgency around complying with the law and improving the information environment.

What we learned about misinformation after the Southport attack

The riots following the Southport stabbings were a stark illustration of how rapidly misinformation can spread and escalate when left unchecked by regulation that is unfit for purpose, and by limited platform oversight.

As violent protests began to escalate, misidentification of the attacker was one of the most common claims Full Fact tracked across social media. Immediately following the stabbings, an allegation began to spread rapidly that the name of the perpetrator was “Ali Al-Shakati”—an allegation that Merseyside Police subsequently confirmed was

⁷¹ Institute for Strategic Dialogue, “ISD Written Evidence to the Science, Innovation and Technology Committee Inquiry, on Social Media, Misinformation and Harmful Algorithms”, 20 January 2025, <https://www.isdglobal.org/isd-publications/isd-written-evidence-to-the-science-innovation-and-technology-committee-inquiry-on-social-media-misinformation-and-harmful-algorithms/>

incorrect.⁷² In a February 2025 Select Committee hearing, representatives from TikTok were questioned about including this incorrect name as an automatic suggestion in its “Others Searched For” bar, effectively amplifying this suggestion to users who had not searched for it, and may not have known about it.⁷³

TikTok’s Director of Public Policy and Government Affairs, UK and Ireland, Ali Law, conceded that, while the incorrect name was removed entirely as a search result the day after TikTok was notified about it, he would have “liked that to [have happened] faster... absolutely.”⁷⁴

All internet platforms must act more decisively. Last summer’s events followed a pattern we’ve previously observed in which online speculation identifies the wrong person in the aftermath of a major incident,⁷⁵ leading to an escalation of violent disorder.⁷⁶ It highlighted the danger of reckless accusations, and the potential for innocent individuals to be targeted.

Another major problem was the lack of accurate information to counter these false claims in a timely manner. The response from the police and government as the riots began was too slow, hampered by important contempt of court rules,⁷⁷ and that created an information void that added fuel to the fire.

Prime Minister Keir Starmer defended the authorities’ decision to withhold details about the case, despite the rumours swirling on social media. He insisted that to do otherwise would have put the judicial process at risk. In a speech following the Southport attack, Mr Starmer argued: “If this trial had collapsed because I or anyone else had revealed crucial details while the police were investigating while the case was being built, while we were awaiting a verdict, then the vile individual who committed these crimes would have walked away a free man.”⁷⁸

⁷² Sian Bayley, “Incorrect name for Southport stabbings suspect circulates online”, Full Fact, 31 July 2024, <https://fullfact.org/online/incorrect-name-southport-stabbings-suspect/>

⁷³ Science, Innovation and Technology Committee, “Oral evidence: Social media, misinformation and harmful algorithms, HC 441”, House of Commons, 25 February 2025, <https://committees.parliament.uk/oralevidence/15413/pdf/#page=7>

⁷⁴ Science, Innovation and Technology Committee, “Oral evidence: Social media, misinformation and harmful algorithms, HC 441”, House of Commons, 25 February 2025, <https://committees.parliament.uk/oralevidence/15413/pdf/#page=7>

⁷⁵ Sian Bayley, “Incorrect name for Southport stabbings suspect circulates online”, Full Fact, 31 July 2024, <https://fullfact.org/online/incorrect-name-southport-stabbings-suspect/>

⁷⁶ Charlotte Green, “Italian sports journalist misidentified as Donald Trump shooter”, Full Fact, 15 July 2024, <https://fullfact.org/online/donald-trump-shooter-misidentified/> (and) Evie Townend, “Sydney student misidentified as Bondi attacker in viral online claims”, Full Fact, 22 April 2024, <https://fullfact.org/online/sydney-student-misidentified-bondi-attacker-viral-claims/>

⁷⁷ Catherine Wylie, “Met Police chief welcomes contempt of court review after Southport stabbings”, The Standard, 24 January 2025, <https://www.standard.co.uk/news/crime/keir-starmer-mark-rowley-southport-prime-minister-metropolitan-police-b1206920.html>

⁷⁸ Prime Minister’s Office, 10 Downing Street, “PM statement on the Southport public inquiry: 21 January 2025”, 2 January 2025, <https://www.gov.uk/government/news/pm-statement-on-the-southport-public-inquiry-21-january-2025>

But there is an awareness that some things need to change. The Home Affairs Select Committee, in its inquiry into the police response concluded that “the lack of information published in the wake of the murders of Bebe King, Elsie Dot Stancombe and Alice da Silva Aguiar created a vacuum where misinformation was able to grow, further undermining public confidence. We respect the Crown Prosecution Service’s (CPS’s) commitment to minimising risks to successful prosecutions, but it is clear that neither the law on contempt nor existing CPS guidance for the media and police are fit for the social media age.”⁷⁹

When Baroness Jones, the Minister for Online Safety, and Dan Jarvis MP, the Minister of State for Security, were asked about misinformation around Southport during the Joint Committee on the National Security Strategy’s Inquiry on Defending Democracy, there was little sign of a new government strategy to counter misinformation incidents like these in future.⁸⁰ But on the subject of false narratives, Mr Jarvis explained that the government had written to the Law Commission to ask them to expedite their review of rules around contempt of court to ensure misinformation on this scale doesn’t happen again.⁸¹

The riots also demonstrated how existing social tensions can be easily exploited and amplified by misinformation. Fabricated narratives with racial and religious undertones emerged, including the claim that two protesters were “stabbed by Muslims”,⁸² which was debunked by Staffordshire Police, who made it clear that “two men involved in the incident were hit with a blunt object that was thrown in the air. No stabbings have been reported to police.”⁸³

Further posts included a widely shared call for “no more mosques” which reached more than a million views on X,⁸⁴ and incorrectly featured an image of the Brighton Royal Pavilion, implying it was a mosque, to fuel public unrest. A video viewed more than 2.3 million times on X falsely claimed that “an African immigrant stabbed a British police officer” in Manchester. In fact, it was clipped from a longer YouTube video captioned “A bus driver was today the victim of a Acid attack at Piccadilly Bus Station [sic]”.⁸⁵

⁷⁹ Home Affairs Committee, “Police response to the 2024 summer disorder”, House of Commons, 14 April 2025, <https://committees.parliament.uk/publications/47476/documents/246718/default/#page=38>

⁸⁰ Joint Committee on the National Security Strategy, “Oral evidence: Defending democracy”, 17 March 2025, <https://committees.parliament.uk/oralevidence/15590/pdf/>

⁸¹ Ibid.

⁸² Sian Bayley, “Claim two protesters were ‘stabbed by Muslims in Stoke’ is false”, Full Fact, 5 August 2024, <https://fullfact.org/online/two-stabbings-stoke-false/>

⁸³ Staffordshire Police, “Ten arrests following disorder in Stoke-on-Trent”, 3 August 2024, <https://www.staffordshire.police.uk/news/staffordshire/news/2024/august/ten-arrests-following-disorder-in-stoke-on-trent/>

⁸⁴ Sian Bayley, “Viral ‘no more mosques’ post uses image of the Brighton Royal Pavilion”, Full Fact, 7 August 2024, <https://fullfact.org/online/mosque-royal-pavilion-brighton/>

⁸⁵ Sian Bayley, “Claim ‘African immigrant stabbed a British police officer’ in Manchester is false”, Full Fact, 13 August 2024, <https://fullfact.org/online/african-immigrant-stabbed-british-police-officer-false/>

The widespread use of manipulated media and AI-generated content added complexity to the chaotic scenes on UK streets. AI-generated images, such as a false depiction of police officers kneeling before men in Islamic dress,⁸⁶ are increasingly difficult and time-consuming to distinguish from genuine content, slowing the pace of reactive fact checks and supercharging the creation of new “evidence” to embolden existing false narratives.

Once again, platforms could have reacted with more urgency. In April 2025 Meta’s Oversight Board said the company had been too slow to recognise the UK as a high-risk location during the riots. It said three posts on Facebook that advocated violence against immigrants and Muslims should have been taken down at the time “because the likelihood of their inciting additional and imminent unrest and violence was significant”.⁸⁷ The Oversight Board said the way Meta enforced its policies in a crisis “revealed inadequacies in the company’s ability to accurately assess visual forms of incitement based on viral disinformation and misinformation”.

The volume of misinformation surrounding the Southport stabbings, and the riots that followed, highlight the need for more robust checks and balances around viral content posted on social media platforms, and stronger cooperation between regulators, fact checkers, the wider media, online platforms and police authorities for addressing crimes spurred on by online falsehoods. This is not about limiting free speech; it is about protecting people from real-world harms.

Why fact checking is important during key information incidents

When fact checkers rate something as false or misleading on Meta, as part of the TPFC programme, their work goes directly to the source of misinformation and empowers users with additional, reliable context to make decisions about what to believe or share. Fact checks that annotate existing posts on social media platforms have been proven to reduce reshares and further amplification of harmful posts. According to Meta’s own research “when a fact-checked label is placed on a post, 95% of people don’t click through to view it”.⁸⁸

During volatile incidents like the 2024 riots, Full Fact is among the few fact checking organisations—and one of just four Meta partners in the UK—that can correct the record, disseminate facts, and counter the spread of baseless rumours directly at the source of the misinformation.⁸⁹ While the TPFC programme is far from perfect, this is where its real strength lies.

⁸⁶ Sarah Turnnidge, “Picture of police kneeling in front of Muslim men is AI-generated”, Full Fact, 22 August 2024, <https://fullfact.org/online/police-kneeling-picture-AI/>

⁸⁷ Oversight Board, “Wide-Ranging Decisions Protect Speech and Address Harms”, 23 April 2025, <https://www.oversightboard.com/news/wide-ranging-decisions-protect-speech-and-address-harms/>

⁸⁸ Marco Pancini, “How Meta Is Preparing for the EU’s 2024 Parliament Elections”, 25 February 2024, <https://about.fb.com/news/2024/02/how-meta-is-preparing-for-the-eus-2024-parliament-elections/>

⁸⁹ Meta, “Request review of a fact-check rating on Facebook, Instagram and Threads”, accessed 23 April 2025, <https://www.facebook.com/business/help/997484867366026?id=673052479947730>

Context and caveat is also vitally important. We published a detailed explainer article outlining some of the key questions posed by the riots,⁹⁰ which was frequently updated in the days which followed. We also distributed our fact checks to leading national media outlets to maximise their visibility and impact when it mattered most. In emergency situations, people deserve access to verifiable facts so they can make up their own minds on issues that matter to them.

Lessons for the government on information disorder

The riots last summer revealed two crucial gaps in legislation. First, the Online Safety Act's sole focus on illegal content means that very little of the misinformation that circulates online comes under its scope. It is not illegal, for example, to speculate on a false name, even if it causes real harm in a volatile situation. The false communications offence, which is one of the few places where misinformation is enshrined in the OSA, is flawed because it requires proving both intent to cause "physical or psychological harm" and definite prior knowledge that the information sent was false.⁹¹ The case of Bernadette Spofforth, for example, generated media attention. She posted a fake name for the Southport attacker on social media, and was arrested, but ultimately faced no charges.⁹²

Second, the Online Safety Act does not include convening powers for Ofcom during major 'information incidents', such as terror attacks or the riots following the Southport murders. Clearer protocols are needed to ensure the government, regulators and other trusted voices are able to come together quickly to give accurate information at critical moments.⁹³ In April 2025, eight months after the riots, Ofcom announced plans for a consultation on a number of measures including the introduction of crisis response protocols for emergency events.⁹⁴ We hope that work is conducted at pace.

⁹⁰ Tony Thompson, Charlotte Green, Nasim Asl, Alex Brocklehurst, "UK riots fact checked: latest updates and key questions answered", Full Fact, 12 August 2024, <https://fullfact.org/news/uk-riots-latest-southport-questions-answered/>

⁹¹ Full Fact, "The Online Safety Act and Misinformation: What you need to know", accessed 23 April 2025, <https://fullfact.org/policy/online-safety-act/>

⁹² Paul Burnell and PA Media, "No charge over spreading of Southport misinformation", BBC News, 18 September 2024, <https://www.bbc.co.uk/news/articles/crl8nwx6ynzo>

⁹³ Full Fact, "The Online Safety Act and Misinformation: What you need to know", accessed 23 April 2025, <https://fullfact.org/policy/online-safety-act/>

⁹⁴ Ofcom, "New rules for a safer generation of children online", 24 April 2025, <https://www.ofcom.org.uk/online-safety/protecting-children/new-rules-for-a-safer-generation-of-children-online>

Private messaging apps make things worse

Public posts on social media may be just the tip of the iceberg. Many industry experts warn that private messaging apps—which are much harder to monitor or regulate—are key tools for spreading misinformation and disinformation, and coordinating harmful or illegal behaviour, in closed-door networks and smaller groups.⁹⁵ Telegram channels, for example, played a major role during the 2024 riots,⁹⁶ including one linked to the UK chapter of the Active Club Network, a decentralised movement of neo-Nazi white supremacist groups.⁹⁷

A report by the European Fact Checking Standards Network (EFCSN), of which Full Fact is a member, highlights growing concern about Telegram. Nearly 76% of fact checkers across Europe agree it plays a significant role in spreading disinformation. The Institute for Strategic Dialogue previously described Telegram as “a safe space for extremists to coordinate activity and instigate violence”,⁹⁸ and while some false claims during the UK riots were not spread with deliberate intent, others clearly were.

So the government needs to be better prepared to tackle similar emergencies in the future. In our sector, that means rethinking how it can ensure that fact checkers are equipped with the right tools, services, and rights to meaningfully tackle misinformation and disinformation at speed and scale.

In Chapter 8, we outline detailed proposals to improve the government’s researcher access scheme. These include secure, real-time access for organisations like Full Fact to platform data that is not always publicly available—something which is essential for stopping the spread of false information without tipping off bad actors.⁹⁹ The violence that broke out across the UK last summer clearly showed why fact checkers should have access to the tools they need to debunk the spread of inaccurate and harmful information at speed and scale.

⁹⁵ Sian Bayley, “What role did misinformation play in riots after the Southport stabbings?”, Full Fact, 2 August 2024, <https://fullfact.org/news/misinformation-southport-stabbings>

⁹⁶ Global Disinformation Index, “The Southport Riots: Online Disinformation and Offline Harm”, 3 September 2024, <https://www.disinformationindex.org/blog/2024-09-03-the-southport-riots-online-disinformation-and-offline-harm/>

⁹⁷ Ibid.

⁹⁸ Institute for Strategic Dialogue, “‘Total system collapse’: Far-right Telegram network incites hate & violence after Southport stabbings”, accessed 8 March 2025, https://www.isdglobal.org/digital_dispatches/total-system-collapse-far-right-telegram-network-incites-accelerationist-violence-after-southport-stabbings/

⁹⁹ Full Fact, “Full Fact Report 2024: Trust and truth in the age of AI”, April 2024, https://fullfact.org/media/uploads/ff2024/18042024-full_fact_report_corrected.pdf#page=46

Chapter 3: The 2024 UK election



Credit: ITV News

Introduction

In the run up to the UK general election in July 2024 there was a flurry of warnings that the campaign could be dominated by deepfakes that could undermine democracy. AI-generated, synthetic content weaponised for political means could, commentators warned, distort public debate and influence voting.

These fears were not entirely unfounded: the 2023 Slovakian parliamentary elections showed how the misuse of AI can impact elections, with faked audio featuring one of the party leaders claiming to have rigged the election going viral right before the polls opened.¹⁰⁰

But for the most part, the UK general election reflected a more nuanced reality: there was a blend of old-fashioned political spin, online misinformation on social media, and low-grade, easily debunked “cheapfakes”. All of them had some impact on the online information environment, from initial campaigning through to polling day but, as Sam Stockwell from the Alan Turing Institute sets out in his essay, deepfakes did not threaten the integrity of the election.

¹⁰⁰ de Nadal, L. & Jančárik, P. (2024) Beyond the deepfake hype: AI, democracy, and “the Slovak case”, Harvard Kennedy School Misinformation Review, Volume 5, Issue 4, <https://doi.org/10.37016/mr-2020-153>

Concern about spin also needs to be put in its proper context. Robust political debate is to be welcomed and expected during an election. But Full Fact was disappointed by the concerted effort by political parties to share exaggerated and often unreliable numerical estimates, which—even after being fact checked—continued to appear in party adverts, social media posts, and high-profile political debates.

In addition, while deepfakes weren't the central threat that had been anticipated, there were a few examples of apparently synthetic content that couldn't be verified. Perhaps the most salient example investigated by Full Fact was an audio clip purporting to be the then-shadow health secretary Wes Streeting swearing and claiming he didn't care about Palestinians being killed in the Israel-Gaza war.¹⁰¹

During the campaign, Full Fact carried out more than 450 hours of monitoring, while our AI tools analysed over 136 million words in 142,909 articles, transcripts and social media posts.¹⁰² With the support of 18 additional volunteer fact checkers, we produced approximately 217 verdicts on claims or repeated claims, and published over 150 pieces of website and video content.¹⁰³

This chapter draws on that extensive effort to reflect on the impact of misinformation on the electoral process in the UK, and what must come next to help protect our democratic system.

Parties' spin and inaccurate figures dominated the election

Contrary to pre-election fears, the course of the 2024 campaign was not ultimately defined by sophisticated deepfakes, but rather by traditional political spin and by misinformation narratives circulating on social media platforms, which were then disseminated by political figures themselves.

Familiar tactics, such as the repeated use of exaggerated and often unreliable statistics, were on daily display. And while the rough and tumble of election politics is nothing new, the constant sharing of inaccurate numerical claims, and the refusal to address requests for correction, served only to damage trust in the political process, which was already worryingly low. Four party leaders in the UK signed up to a Full Fact pledge calling for honest campaigning, but the leaders of Labour, the Conservatives and the Liberal Democrats did not.¹⁰⁴

¹⁰¹ Full Fact Team, "No evidence audio clip supposedly of Wes Streeting comments about Palestinian deaths is genuine", Full Fact, 3 July 2024, <https://fullfact.org/election-2024/wes-streeting-audio-clip-palestine/>

¹⁰² Full Fact, "General election 2024, fact checked", Full Fact, 5 July 2024, <https://fullfact.org/blog/2024/jul/general-election-2024-fact-checked/>

¹⁰³ Ibid.

¹⁰⁴ Craig Dawson, "It's time political parties tidied up their election campaigns", Full Fact, 10 November 2023, <https://fullfact.org/blog/2023/nov/letter-to-political-parties/>

The net effect was predictable. According to an Electoral Commission survey after the election, 61% of respondents said they saw misleading or inaccurate information about political parties' policies during the campaign, and 52% said they saw misleading or inaccurate information about candidates.¹⁰⁵

Comment

Vijay Ranganathan, Chief Executive, the Electoral Commission

Election campaigns are noisy, colourful, argumentative, sometimes divisive: built around political views and robust debate. The key is that voters hear the different views on offer and can make their choices. But that is why deliberate attempts to mislead voters, or people circulating misleading material, can be a problem: they threaten informed voter choice.

Before last year's general election, there was a growing concern about the role misinformation and disinformation might play, and whether we were likely to see AI and deepfakes used to try and deceive the public. We, together with others, put in place a number of changes to help mitigate the risk.

The July 2024 campaign was energetic and lively, but when the dust settled after polling day, I think we all saw that there hadn't been a significant problem... this time.

Voters certainly saw misleading material. After the election, over half of voters surveyed told us they saw misleading or inaccurate information about political parties' policies and candidates. Around a quarter saw or heard a deepfake photo, video or audio clip about the election. We were made aware of a small number of deepfakes of politicians circulating online during the campaign—reassuringly they tended to be quickly called out for what they were.

In our view, there are two key elements to successfully addressing misinformation and disinformation.

First, voters need to understand how campaigners are trying to influence them during a campaign. This was the first election where we called on anyone using generative AI to clearly label it as such. It was also the first where digital imprints were required on campaign material, which shows everyone who paid to produce it. The Commission has been calling for their introduction for over twenty years, so we were pleased voters could finally see this key piece of information.

¹⁰⁵ The Electoral Commission, "Report on the 2024 UK Parliamentary general election and the May 2024 elections", The Electoral Commission, accessed 6 March 2025, <https://www.electoralcommission.org.uk/research-reports-and-data/our-reports-and-data-past-elections-and-referendums/report-2024-uk-parliamentary-general-election-and-may-2024-elections#campaigning>

Second, it's crucial to support voters to consider and verify the information they see. At the start of the campaign the Commission published new advice for voters on how to engage confidently with campaign material and think critically about what they saw or heard.¹⁰⁶ We also worked with Shout Out UK and Ofcom to create resources specifically aimed at helping young people to dismiss disinformation.¹⁰⁷

While the majority of people told us they ignored misleading content, nearly half took action such as fact-checking or reporting the information in some way. So voters value and use services to verify the accuracy of information that bodies such as Full Fact are providing. Impartial, accurate and trusted sources of information are the antidote to efforts to undermine voter confidence and trust.

While the Commission doesn't have a role in regulating campaign literature, one thing we can do during a campaign is directly and rapidly counter misleading information about the electoral process itself. In the run up to the general election, our voter information hub was viewed 5.1 million times, and we responded to 8,500 queries from members of the public.

Interestingly, younger people were more likely to take action when they came across something they thought was misinformation. We see educating people about democracy as a great tool for countering some of the false narratives we see online about politics and elections.

We are already creating resources that young people and educators can use to explain our democratic processes. Over the next five years, we will scale this up, investing much more into providing young people with the information they need to participate in elections and democracy.

We will also continue to work closely with other organisations, including the UK's governments, regulators and social media companies, to monitor emerging threats and identify solutions. This includes working to address the concerning trend of candidate abuse and intimidation. After the general election some candidates told us that they felt misinformation that was spread online led to in-person abuse and harassment. This is damaging to the individuals and our democracy and must be tackled—or some will be put off standing as candidates.

¹⁰⁶ The Electoral Commission, "Campaigning for your vote", accessed 23 April 2025, <https://www.electoralcommission.org.uk/voting-and-elections/campaigning-your-vote>

¹⁰⁷ Shout Out UK, "We team up with Ofcom to dismiss disinformation around the General Election", Shout Out UK, 19 June 2024, <https://www.shoutoutuk.org/2024/06/19/we-team-up-with-ofcom-to-dismiss-disinformation-around-the-general-election/>

We will be paying close attention to this and the information put to voters ahead of the next set of big elections, which are in Scotland and Wales next year. The planning and legislation are already well under way, and we will be monitoring the campaign and experiences of voters.

So there is a lot to do in the coming years to protect our democratic system—including the trust of voters, the enthusiasm of campaigners to share their messages, and the integrity of the voting process. We look forward to doing it with all of you.

One of the most prominent of these claims was first made by then Prime Minister Rishi Sunak the month before the election: that the Labour party's plan would mean "£2,000 higher taxes for every working family".¹⁰⁸ This figure, despite being repeatedly shown to be unreliable by fact checkers including Full Fact, was cited repeatedly during the course of the Conservative party's campaign.

In reality, the claim was rooted in a series of assumptions, including that Labour would fill any budgetary gaps or "black holes" with increased taxes instead of borrowing, and that taxes affect all families across the country equally. Mr Sunak also appeared to attribute the £2,000 figure to "independent Treasury officials", when in fact it was a Conservative party estimate based on costing Labour's "unfunded spending commitments", not all of which were produced by Treasury officials, and several of which we found to be uncertain.¹⁰⁹

From the opposition arose a similarly dubious claim about mortgage costs under the Conservatives. At a press conference by then-shadow chancellor Rachel Reeves, and in a dossier, the Labour party claimed that "the Conservatives' plan will mean £4,800 more on your mortgage".¹¹⁰

However, Full Fact found that the £4,800 figure was a speculative estimate that relied on several uncertain assumptions, central among which was that "unfunded promises" under the Conservatives would result in £71 billion worth of extra borrowing.¹¹¹

Both claims were key points of debate during the election, while Full Fact noted that public debate on other issues was limited. In an analysis of a week of broadcasting during the campaign, Full Fact's AI tools found 6,574 mentions of tax, while other topics paled in comparison, with only 933 mentions of climate change, 922 mentions of housing and 777 mentions of crime.¹¹²

¹⁰⁸ Evie Townend, "Would families face a £2,000 tax rise under Labour?", Full Fact, 5 June 2024, <https://fullfact.org/economy/conservative-claim-general-election-labour-2000-tax-increase/>

¹⁰⁹ Evie Townend, "Would families face a £2,000 tax rise under Labour?", Full Fact, 5 June 2024, <https://fullfact.org/economy/conservative-claim-general-election-labour-2000-tax-increase/>

¹¹⁰ Leo Benedictus, "Would Conservative spending commitments mean a £4,800 increase in the average mortgage?", Full Fact, 13 June 2024, <https://fullfact.org/election-2024/rachel-reeves-labour-4800-mortgage-rates/>

¹¹¹ Ibid.

¹¹² Mark Frankel, "What politicians HAVEN'T talked about during the election campaign", Full Fact, 3 July 2024, <https://fullfact.org/blog/2024/jul/what-politicians-havent-talked-about-during-the-election-campaign/>

We also found that misleading political claims spread significantly online, with parties leveraging tactics such as paid display advertising to rapidly disseminate unfounded assertions about opposition policies to a targeted audience.

Days before the election, the Conservatives published widely circulated online advertisements claiming that Labour's plan to implement a "national ULEZ" (Ultra-Low Emission Zone) would be "coming to a road near you this July".¹¹³ A search on Facebook's Ad Library at the time suggested that more than 800 versions of the advert may have been posted.

But there was no specific evidence that Labour was planning to introduce such a scheme, and the party denied any plans to do so. There were also no plans in Labour's manifesto for a 'national ULEZ', and Full Fact was unable to find any other specific information to back up the claim.¹¹⁴

It all added up to an election campaign in which disproportionate attention was given to numbers which didn't add up, or to misleading information masquerading as established fact. Voters, in general, deserved better.

The deepfake threat was overestimated in 2024, but is relevant for future elections

70% of MPs polled in a YouGov survey prior to the 2024 election were concerned about AI-generated content increasing the spread of misinformation and disinformation in the run up to polling day.¹¹⁵ In the event, their concerns—shared by others—proved to be largely unfounded, but they spoke to widespread unease about the state of online information and the potential for AI technology to deceive and mislead.

Prominent examples were few and far between. An audio clip supposedly of Keir Starmer claiming that he hates Liverpool was widely circulated online, with one post having received over 400,000 views as of 1 July 2024.¹¹⁶ As with the Wes Streeting audio clip mentioned earlier, we were not able to determine whether the Starmer clip was generated with AI, cleverly edited or was simply the work of a skilled impersonator. But we did not see any evidence to suggest it was real, and we identified versions of the clip that had been circulating since October 2023.

¹¹³ Evie Townend, "Is Labour planning a 'national ULEZ'?", Full Fact, 28 June 2024, <https://fullfact.org/election-2024/conservative-ad-labour-national-ulez-misleading/>

¹¹⁴ Evie Townend, "Is Labour planning a 'national ULEZ'?", Full Fact, 28 June 2024, <https://fullfact.org/election-2024/conservative-ad-labour-national-ulez-misleading/>

¹¹⁵ Micheal Savage, "Call for action on deepfakes as fears grow among MPs over election threat", The Guardian, 21 January 2024, <https://www.theguardian.com/politics/2024/jan/21/call-for-action-on-deepfakes-as-fears-grow-among-mps-over-election-threat>

¹¹⁶ Tony Thompson, "No evidence old audio clip supposedly of Keir Starmer saying he hates Liverpool is genuine", Full Fact, 1 July 2024, <https://fullfact.org/online/keir-starmer-liverpool-hate/>

Comment

Sam Stockwell, Alan Turing Institute Centre for Emerging Technology and Security

In 2024, the UK was one of at least 64 countries around the world heading to the polls in what was dubbed “the ultimate election year”.¹¹⁷ With many of these votes being an attractive target for hostile interference efforts, election security was a particularly high priority.

Yet fast-forward to the end of the year, and it was clear that the negative impact of AI had¹¹⁸ been¹¹⁹ overblown¹²⁰—including with the UK election. Firstly, there was no conclusive evidence that such tools had affected election results.¹²¹ One of the main reasons behind this was that there were simply too few viral cases to influence the electorate—with our research identifying just 16 instances in the UK.¹²² Given the low volumes coupled with the constant avalanche of information we are exposed to, voters are unlikely to remember these examples. Indeed, a survey from the Alan Turing Institute has shown that only 5.7% of over 1,400 UK respondents could recall seeing a viral political deepfake.¹²³

However, we also often tend to “overestimate the change technology brings in the short term and underestimate its long-term effects.”¹²⁴ Despite the lack of influence on the election outcome, we did identify worrying signs of second-order damage to the wider democratic system. This included UK users being confused over whether election content they viewed was synthetic or genuine—even on deepfakes which had been verified as such.¹²⁵ Not only does this pollute our information ecosystem, but it poses fundamental risks to the ability of users to trust credible sources and complicates fact checking efforts.

¹¹⁷ Koh Ewe, “The Ultimate Election Year: All the Elections Around the World in 2024”, TIME, 28 December 2023, <https://time.com/6550920/world-elections-2024/>

¹¹⁸ Sam Stockwell, Megan Hughes, Phil Swatton, Katie Bishop, “AI-Enabled Influence Operations: The Threat to the UK General Election”. CETAS, 28 May 2024, <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-uk-general-election>

¹¹⁹ Sam Stockwell, “AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections”, CETAS, 19 September 2024, <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-analysis-2024-uk-and-european-elections>

¹²⁰ Sam Stockwell, Megan Hughes, Phil Swatton, Albert Zhang, Jonathan Hall KC, Kieran, “AI-Enabled Influence Operations: Safeguarding Future Elections”, CETAS, 13 November 2024, <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-safeguarding-future-elections>

¹²¹ Sam Stockwell, “AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections”, CETAS, 19 September 2024, <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-analysis-2024-uk-and-european-elections>

¹²² Ibid.

¹²³ Tvesha Sippy, Florence E. Enock, Jonathan Bright, Helen Z. Margetts, “Behind the Deepfake: 8% Create; 90% Concerned”, CETAS, June 2024, <https://www.turing.ac.uk/news/publications/behind-deepfake-8-create-90-concerned>

¹²⁴ Fiona Dennehy, “Turing event at the Houses of Parliament explores impact of AI disinformation on elections”, The Alan Turing Institute, 30 January 2025, <https://www.turing.ac.uk/news/turing-event-houses-parliament-explores-impact-ai-disinformation-elections>

¹²⁵ Marianna Spring, “Labour’s Wes Streeting among victims of deepfake smear network on X”, BBC News, 7 June 2024, <https://www.bbc.co.uk/news/articles/cg33x9jm02ko>

Female UK politicians were also targeted by deepfake pornographic smears,¹²⁶ with the psychological damage such content caused potentially leading to a ‘chilling effect’ on the willingness of other women to enter politics. Finally, one candidate was even accused of being an AI-generated bot¹²⁷—despite this being debunked.¹²⁸ Such rumours reflect a concerning trend where a perceived sense of AI-generated content being everywhere, and difficult to detect, blurs the line between what is real and what is not.¹²⁹ In turn, this risks creating a fertile environment for politicians¹³⁰ and others¹³¹ to dismiss damaging allegations that may turn out to be credible, or even reshape the truth.

Although deepfakes did not play much of a role in the 2024 UK election, the impact of misleading narratives circulated by political candidates,¹³² social media algorithms¹³³ and ordinary users¹³⁴ cannot be neglected. These observations underscore the need to tackle misinformation and disinformation more systematically, as opposed to just narrow election- or AI-based interventions. By targeting different stages of the content’s ‘life cycle’,¹³⁵ friction points can be established that make it more challenging for different actors to create or spread deceptive material. With several elections looming in the coming years, complacency cannot creep in. Now is a golden window of opportunity to enhance not only election security, but the very resilience of our democratic system against all forms of mis- and disinformation.

Ultimately the deepfake threat was overshadowed by far more rudimentary forms of digital distortion, such as edited videos designed to misrepresent politicians’ statements or events. These “cheapfakes”—less technically advanced pieces of fabricated content that are obviously false—proved effective in misleading voters.

¹²⁶ Cathy Newman, “Exclusive: Top UK politicians victims of deepfake pornography”, Channel 4 News, 1 July 2024, <https://www.channel4.com/news/exclusive-top-uk-politicians-victims-of-deepfake-pornography>

¹²⁷ Danny Rigg, “Were Reform UK’s candidates even real?”, Metro, 8 July 2024, <https://metro.co.uk/2024/07/08/reform-uks-candidates-even-real-21188964/>

¹²⁸ Joel Pike, Phil Kemp, “Reform fake candidates conspiracy theories debunked”, BBC News, 11 July 2024, <https://www.bbc.co.uk/news/articles/ckvgl9kzwzjo>

¹²⁹ Josh Goldstein, Andrew Lohn, “Deepfakes, Elections, and Shrinking the Liar’s Dividend”, Brennan Centre for Justice, 23 January 2024, <https://www.brennancenter.org/our-work/research-reports/deepfakes-elections-and-shrinking-liars-dividend>

¹³⁰ Nilesh Christopher, “An Indian politician says scandalous audio clips are AI deepfakes. We had them tested”, Rest of World, 5 July 2023, <https://restofworld.org/2023/indian-politician-leaked-audio-ai-deepfake/>

¹³¹ Perry Carpenter, “The Liar’s Dividend: How AI Is Reshaping Truth In Business Communications”, Forbes, 2 October 2024, <https://www.forbes.com/councils/forbesbusinesscouncil/2024/10/02/the-liars-dividend-how-ai-is-reshaping-truth-in-business-communications/>

¹³² Full Fact, “Full Fact’s AI tools spot hundreds of misleading election claims on social media”, 3 July 2024, <https://fullfact.org/live/2024/jul/ai-tools-spot-misleading-election-claims/>

¹³³ Marianna Spring, “TikTok users being fed misleading election news, BBC finds”, BBC News, 2 June 2024, <https://www.bbc.co.uk/news/articles/c1ww6vz1l81o>

¹³⁴ Marianna Spring, “This wasn’t the social media election everyone expected”, BBC News, 8 July 2024, <https://www.bbc.co.uk/news/articles/cj50qjy9g7ro>

¹³⁵ Sam Stockwell, Megan Hughes, Phil Swatton, Albert Zhang, Jonathan Hall KC, Kieran, “AI-Enabled Influence Operations: Safeguarding Future Elections”, CETAS, 13 November 2024, <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-safeguarding-future-elections>

One example: a video of Rachel Reeves pausing for several seconds after being asked about public finances under a Labour government, was clipped and shared with captions like, “Cat got your tongue, Rachel?”—implying that she had been caught off guard or unprepared. However, a review of the full footage revealed technical glitches during the interview that caused a delay between the question and her response.¹³⁶

Another viral image showed Rishi Sunak standing in front of a Morrisons supermarket sign, with part of the logo obscured to spell ‘moron’. This was a composite image of two different photos, edited to make it look like certain letters of the logo were blocked. The picture was seemingly intended to be a joke, but it had also been shared alongside captions which indicated that many people believed it was real.¹³⁷

In last year’s Full Fact report, we wrote about a growing challenge in this space: determining the intent behind AI-generated content. Is it meant as satire or sabotage? Is it a joke gone viral or deliberate disinformation aimed at influencing voters? We are certainly not in the business of fact checking satire, and that blurred line between mischief and manipulation makes it harder to track, label and respond to deceptive material before it spreads.¹³⁸

In any event, the effect of deepfakes in the UK in 2024 was strictly limited, and several factors may have contributed to this. The election was called slightly earlier than many people anticipated, leaving less time for bad actors to prepare. More significantly, the result was never really in doubt. Throughout the campaign, a Labour victory looked like a foregone conclusion—a point even senior Conservatives, like Mel Stride,¹³⁹ acknowledged. That sense of inevitability may have reduced the perceived need for dramatic or sophisticated interference.

Safeguarding UK elections from future threats to information integrity

Nevertheless, the threat posed by AI-driven deepfakes is real and evolving, while it could be argued that inaccurate or deceptive political spin, and social media-fuelled misinformation, were both highly effective at distorting public perception over the last year.

¹³⁶ Evie Townend, “Conservatives share clip of Rachel Reeves interview with lag to make it look like she struggled to answer question”, Full Fact, 25 June 2024, <https://fullfact.org/election-2024/reeves-interview-cropped-technical-issues/>

¹³⁷ Evie Townend, “Thousands share edited image of Rishi Sunak on social media”, Full Fact, 24 May 2024, <https://fullfact.org/online/edited-photo-rishi-sunak-morrisons/>

¹³⁸ Full Fact, “Full Fact Report 2024: Trust and truth in the age of AI”, April 2024, https://fullfact.org/media/uploads/ff2024/18042024-full_fact_report_corrected.pdf#page=19

¹³⁹ Charles Hymas, “Tory cabinet minister all but concedes election to Labour”, The Telegraph, 3 July 2024, <https://www.telegraph.co.uk/politics/2024/07/03/mel-stride-cabinet-minister-concede-general-election-labour/>

Localised misinformation, especially around sensitive global issues, appears to have had a real impact. While Full Fact did not systematically monitor Israel-Gaza-related misinformation at the constituency level, some candidates felt the effects keenly, both online and in person. Labour's Heather Iqbal, for example, reported being targeted with harassment and abuse, including being labelled a "Zionist and genocide agent"¹⁴⁰—an accusation which, in her opinion, contributed to her defeat.

Strong political opinions are one thing—but sustained campaigns should be grounded in fact. When political campaigning crosses into intimidation, it shows how targeted, identity-based disinformation can influence outcomes. This suggests future election monitoring may need to go beyond fact checking broad national narratives and dig deeper into the racialised and discriminatory tactics used in specific communities.

In his evidence to the Defending Democracy Inquiry, Dan Jarvis MP, Minister for Security, highlighted the challenges that women and ethnic-minority candidates in particular faced. "It is deeply concerning," he said, "to think that, in the future, people who are highly qualified to serve in public life might be dissuaded from stepping forward to do so because of the toxic environment that we saw in some places in the general election."¹⁴¹

More work needs to be done to ensure candidates running in future elections are protected from misinformation that targets their identity or political standing. Full Fact continues to call on all candidates to publish honest and transparent election materials that do not intimidate others, spread false narratives, or incite hatred and violence towards groups or political parties. This is essential to preserving the integrity of our elections.

We also urge the government to strengthen safeguards against harmful political deepfakes. Their impact may have been limited in 2024, but that's no excuse for complacency. Laws must be in place before the next election to tackle misleading synthetic content directly, and we reiterate the calls we developed with Demos on the need for political parties to commit to the responsible use of generative AI.

Public understanding of deepfakes remains shaky. In the Electoral Commission's post-election survey, nearly one in five respondents (18%) said they didn't know whether they had encountered a deepfake—highlighting widespread confusion about how to spot this kind of synthetic content.¹⁴²

¹⁴⁰ Peter Walker, "Labour candidate who lost to new pro-Gaza MP accuses his backers of intimidation", The Guardian, 21 July 2024 <https://www.theguardian.com/politics/article/2024/jul/21/labour-candidate-lost-new-pro-gaza-mp-accuses-backers-intimidation>

¹⁴¹ Joint Committee on the National Security Strategy, "Oral evidence: Defending democracy", 17 March 2025, <https://committees.parliament.uk/oralevidence/15590/pdf/#page=6>

¹⁴² The Electoral Commission, "Report on the 2024 UK Parliamentary general election and the May 2024 elections", accessed 23 April 2025, <https://www.electoralcommission.org.uk/research-reports-and-data/our-reports-and-data-past-elections-and-referendums/report-2024-uk-parliamentary-general-election-and-may-2024-elections#campaigning>

Sometimes, even real people are mistaken for AI-generated fabrications. Full Fact debunked one such claim involving Reform UK candidate Mark Matlock, whose image on party leaflets led social media users to speculate he might not be a real person, simply because he “looked AI-generated”.¹⁴³

A post-election survey from Ofcom echoed this.¹⁴⁴ While 60% of respondents said they had seen content about the election they believed was false or misleading, almost half (46%) weren’t sure whether they’d seen a deepfake at all. That uncertainty only reinforces the need for better public awareness and education on what deepfakes are—and how to identify them.

What needs to change

Clear policies are urgently needed to tackle the growing confusion around deepfakes, and set firm standards for identifying and responding to them. Without action, public understanding won’t improve. Full Fact has long called for stronger regulation of deepfakes during election periods—a call that remains unanswered. While in opposition, Labour proposed adding an “offence of creating and sharing political deepfakes” to the Data Protection and Digital Information Bill.¹⁴⁵ It has now faced this threat directly with what appear to be audio deepfakes of Keir Starmer expressing hatred for Liverpool.¹⁴⁶

The government has already taken steps to criminalise the creation and sharing of sexually explicit deepfakes,¹⁴⁷ which is very welcome. But it’s time to extend those protections to cover political content—before deepfakes undermine trust and do real harm in future elections. With elections approaching in the Welsh and Scottish Parliaments next year, we repeat our call: the UK urgently needs stronger rules to deal with the rising threat of deepfakes in politics.

¹⁴³ Hannah Smith, “Reform UK candidate who stood in London was not ‘AI-generated’”, Full Fact, 9 July 2024, <https://fullfact.org/online/reform-uk-candidate-AI/>

¹⁴⁴ Ofcom, “UK General Election news and opinion formation survey 2024”, 10 September 2024, <https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/tv-radio-and-on-demand-research/tv-research/news/news-consumption-2024/uk-general-election-survey-2024-report.pdf?v=379617>

¹⁴⁵ Full Fact, “Full Fact Report 2024: Trust and truth in the age of AI”, April 2024, https://fullfact.org/media/uploads/ff2024/18042024-full_fact_report_corrected.pdf#page=25

¹⁴⁶ Tony Thompson, “No evidence old audio clip supposedly of Keir Starmer saying he hates Liverpool is genuine”, Full Fact, 1 July 2024, <https://fullfact.org/online/keir-starmer-liverpool-hate/>

¹⁴⁷ Ministry of Justice and Alex Davies-Jones MP, “Government crackdown on explicit deepfakes”, 7 January 2025, <https://www.gov.uk/government/news/government-crackdown-on-explicit-deepfakes>

Chapter 4: Disinformation threats

Introduction

While Full Fact focuses mainly on misinformation, state-sponsored disinformation is also a significant and growing threat, particularly during election campaigns. It is not our core area of expertise, but we monitor developments closely and work with a number of organisations that specialise in it.

As noted in the previous chapter, last year's UK general election was not affected in significant ways, but disinformation—both state-sponsored and spread by powerful non-state actors—remains a focus of concern across the political spectrum.

In launching a new inquiry—*Disinformation diplomacy: how malign actors are seeking to undermine democracy*—at the beginning of this year,¹⁴⁸ the House of Commons Foreign Affairs Committee set out to understand which actors are primarily responsible, and which channels and technologies are being used.

The chair of the committee, Dame Emily Thornberry MP, argued that disinformation campaigns are designed deliberately to sow the seeds of discontent. “They have been weaponised to subvert free and fair elections, to undermine the rules-based international order and to propagate anti-Western narratives. Foreign malign actors have realised the power of the media and social media in supporting their aims and interests.”¹⁴⁹

That suggests politicians are keen—as indeed they should be—for internet platforms to take greater responsibility, but Dame Emily also emphasised that these threats aren't just coming from hostile states, but also from non-state actors who have significant influence over our information environment. “Powerful figures such as Elon Musk,” she said, “exploit their platform to spread disinformation that disrupts and destabilises.”

At the beginning of this year, Mr Musk became almost obsessively active in commenting on UK politics on social media, often amplifying conspiracy theories and far-right propaganda. In one post on X, the platform he owns, he described the Minister for Safeguarding, Jess Phillips MP, as a “rape genocide apologist”¹⁵⁰ and said she should be jailed. In another, he shared unreliable estimates—presented as established facts—about the number of victims of grooming gangs in the UK.¹⁵¹ When individuals with enormous

¹⁴⁸ Foreign Affairs Committee, “New inquiry: Disinformation diplomacy: How malign actors are seeking to undermine democracy”, 15 January 2025, <https://committees.parliament.uk/committee/78/foreign-affairs-committee/news/204722/new-inquiry-disinformation-diplomacy-how-malign-actors-are-seeking-to-undermine-democracy/>

¹⁴⁹ Ibid.

¹⁵⁰ Victoria Derbyshire and Kate Whannel, “Musk’s ‘disinformation’ endangering me, says Phillips”, updated 8 January 2025, <https://www.bbc.co.uk/news/articles/cn7r0pzz57vo>

¹⁵¹ Leo Benedictus, “How many children have been the victims of grooming gangs in the UK?”, Full Fact, 8 January 2025, <https://fullfact.org/crime/grooming-gang-victims-musk-pearson-champion/>

power abuse their position and spread false or misleading information, we should all be concerned.

As the following essay from Demos makes clear, the government should not be resting on its laurels just because the general election in 2024 passed without incident.

Comment

Jamie Hancock, Researcher (Digital Policy) at Demos

Liberal Democracy is in peril. Many countries are facing democratic backsliding,¹⁵² a rise in extremist populism,¹⁵³ and growing anti-democratic movements.¹⁵⁴ At the heart of this crisis is the issue of foreign interference.

In a recent Demos report,¹⁵⁵ we argued that the democratic emergency hinges on threats to what we call 'epistemic security': the safety and resilience of the information supply chains which are vital to the health of democracies.¹⁵⁶ We identified four interconnecting conditions which have contributed to critical vulnerabilities in the UK's information supply chains: (1) the mass digitisation of communication, (2) weakened news ecosystems, (3) heightened risk of foreign interference, and (4) regulatory shortcomings. In this essay, we focus on foreign interference and the challenge it poses to the UK's democracy.

Foreign influence refers to political interference by actors from abroad, usually as part of efforts to pursue another state's foreign policy objectives. Some foreign interference actors may fit 'traditional' categories of national security threats, such as adversarial states. Others are 'non-traditional', powerful individuals keen to pursue their own agendas. Today, the UK faces a risk of attempts to interfere in its democratic processes by both types of actors.

Part of this increased risk is due to rising authoritarianism and democratic backsliding worldwide. Authoritarian-leaning political movements have grown in influence in

¹⁵² International IDEA, "The Global State of Democracy 2024", accessed 23 April 2025, <https://www.idea.int/gsod/2024/> and Agence France-Presse in Stockholm, "US added to list of 'backsliding' democracies for first time", The Guardian, 22 November 2021, <https://www.theguardian.com/us-news/2021/nov/22/us-list-backsliding-democracies-civil-liberties-international>

¹⁵³ Murat Atkas, "The rise of populist radical right parties in Europe", International Sociology International Sociology, 39(6), 591-605, 16 November 2024, <https://journals.sagepub.com/doi/full/10.1177/02685809241297547>; Michael Cox, "Understanding the Global Rise of Populism", LSE IDEAS, February 2018, <https://www.lse.ac.uk/ideas/Assets/Documents/updates/LSE-IDEAS-Understanding-Global-Rise-of-Populism.pdf>

¹⁵⁴ E.g. BBC Wales Investigates team, "Far-right group exposed in undercover BBC investigation", BBC News, 20 January 2025, <https://www.bbc.co.uk/news/articles/cn8xykr5v95o>.

¹⁵⁵ Elizabeth Seger, Hannah Perry and Jamie Hancock, "Epistemic Security 2029: Fortifying the UK's information supply chain to tackle the democratic emergency". Demos, 20 February 2025, <https://demos.co.uk/research/epistemic-security-2029-fortifying-the-uks-information-supply-chain-to-tackle-the-democratic-emergency/>

¹⁵⁶ Elizabeth Seger, Shahar Avin, Gavin Pearson, Mark Briers, Seán Ó Heigeartaigh and Helena Bacon, "Tackling threats to informed decision-making in democratic societies: Promoting epistemic security in a technologically-advanced world", The Alan Turing Institute, October 2020, <https://www.turing.ac.uk/news/publications/tackling-threats-informed-decision-making-democratic-societies>

countries close to the UK like Italy,¹⁵⁷ Germany,¹⁵⁸ and Austria.¹⁵⁹ Governments in countries including Russia,¹⁶⁰ Hungary,¹⁶¹ and Turkey¹⁶² have become increasingly authoritarian and autocratic. Meanwhile, the return of Donald Trump to the US presidency—with the involvement of unelected tech entrepreneurs, like Elon Musk—fuels fears that America is also experiencing an authoritarian shift.¹⁶³

In addition, in countries close to Russia, such as Estonia¹⁶⁴ and Poland,¹⁶⁵ governments are increasingly worried about increased political interference and sabotage efforts associated with Russia.¹⁶⁶ In Romania, the constitutional court halted a Presidential Election 48 hours before polling due to allegations from the security services of widespread election interference operations coordinated by Russia.¹⁶⁷

In the UK, a legacy of Russian adversarial foreign influence efforts goes back to the Cold War.¹⁶⁸ However, the UK now faces an emerging challenge from its traditional ally: the United States. For example, Elon Musk has intervened in UK political debates on several occasions¹⁶⁹ and at one point was reported to be considering donating substantial amounts to a British political party.¹⁷⁰ From his quasi-governmental position in the White House,¹⁷¹ combined with his control of the social media platform X and AI platform

¹⁵⁷ Alexander Stille, “The shapeshifter: who is the real Giorgia Meloni?”, The Guardian, 19 September 2024, <https://www.theguardian.com/world/2024/sep/19/shapeshifter-who-is-the-real-giorgia-meloni-italy-prime-minister>

¹⁵⁸ Ajit Niranjani, “German parliament sits for first time with AfD as second biggest party”, The Guardian, 25 March 2025, <https://www.theguardian.com/world/2025/mar/25/german-parliament-sits-for-first-time-with-afd-as-main-opposition>

¹⁵⁹ Bethany Bell, “Austrian far-right party tasked with forming coalition”, BBC News, 6 January 2025, <https://www.bbc.co.uk/news/articles/clykjz8kk9xo>

¹⁶⁰ Freedom House, “Russia: Country Profile”, accessed 23 April 2025, <https://freedomhouse.org/country/russia>

¹⁶¹ European Parliament, “MEPs: Hungary can no longer be considered a full democracy”, 15 September 2022, <https://www.europarl.europa.eu/news/en/press-room/20220909IPR40137/meps-hungary-can-no-longer-be-considered-a-full-democracy>

¹⁶² Economist Intelligence, “Democracy Index 2023”, accessed 23 April 2025, <https://www.eiu.com/n/campaigns/democracy-index-2023/>

¹⁶³ E.g. Steven Levitsky and Lucan Way (2025), “The Path to American Authoritarianism: What Comes After Democratic Breakdown”, Foreign Affairs, 11 February 2025, <https://www.foreignaffairs.com/united-states/path-american-authoritarianism-trump>

¹⁶⁴ Michael Schwartz, “A Spate of Vandalism Rattled Estonia. Russia Was to Blame, Officials Say”, New York Times, 5 December 2024, <https://www.nytimes.com/2024/12/05/world/europe/estonia-vandalism-russia-sabotage.html>

¹⁶⁵ Reuters, “Poland says Russia trying to recruit Poles on dark net to influence election”, 28 January 2025, <https://www.reuters.com/world/europe/poland-says-russia-trying-recruit-poles-dark-net-influence-election-2025-01-28/>

¹⁶⁶ Emma Burrows, “Western officials say Russia is behind a campaign of sabotage across Europe. This AP map shows it”, Associated Press, updated 21 March 2025, <https://apnews.com/article/russia-ukraine-war-europe-hybrid-campaign-d61887dd3ec6151adf354c5bd3e6273e>

¹⁶⁷ Paul Kirby and Nick Thorpe, “Romania’s cancelled presidential election and why it matters”, BBC News, 6 December 2024, <https://www.bbc.co.uk/news/articles/cx2yl2zxrq1o>

¹⁶⁸ Francesco Bechis, “Playing The Russian Disinformation Game: Information operations from Soviet tactics to Putin’s sharp power”, In Democracy and Fake News. Routledge, 2020, <https://www.taylorfrancis.com/chapters/oa-edit/10.4324/9781003037385-12/playing-russian-disinformation-game-francesco-bechis>

¹⁶⁹ E.g. Andrew McDonald, “Elon Musk shares fake news claiming UK rioters will be sent to ‘detainment camps’”, Politico, 8 August 2024, <https://www.politico.eu/article/elon-musk-share-fake-news-uk-rioters-detainment-camp/>

¹⁷⁰ Harry Taylor and PA Media, “Elon Musk among billionaires set to donate to Reform UK, says treasurer”, The Guardian, 22 December 2024, <https://www.theguardian.com/politics/2024/dec/22/elon-musk-among-billionaires-set-to-donate-to-reform-uk-says-treasurer>

¹⁷¹ Kayla Epstein, “Who is Doge’s official leader? White House says it’s not Musk”, BBC News, 25 February 2025, <https://www.bbc.co.uk/news/articles/c2erg38vjx8o>

Grok,¹⁷² Musk's potential for international disruption has been enhanced in recent months. Given the UK public's record low trust in its politicians and government,¹⁷³ the country faces a volatile moment which Musk or others so inclined could take advantage of.

And although the UK has policies and legislation in place to address foreign interference, these have significant shortcomings. Firstly, the Online Safety Act 2024 (OSA) makes knowingly spreading false information with intent to cause harm a criminal offence.¹⁷⁴ While in theory this offence could be used to prosecute cases of foreign interference via social media, in practice the law may be difficult to enforce against people who reside in foreign jurisdictions.

Second, the National Security Act 2023 (NSA) also includes provisions intended to deter and counter foreign interference.¹⁷⁵ However, like the OSA, there may be barriers when it comes to enforcing the NSA's foreign interference offences: (1) foreign influence campaigns may operate subtly in ways which do not meet the NSA's criteria; (2) it may be impractical to prosecute overseas actors. As a result, current legislation does not fully address the situation at hand.

Third, the UK currently has no publicly available plan for how to proceed in cases of suspected election interference. This is despite past allegations of attempts at interference,¹⁷⁶ a cyberattack against the Electoral Commission which exposed the names and addresses of anyone registered to vote between 2014 and 2022,¹⁷⁷ and what the Electoral Commission has called "unacceptable levels of abuse and intimidation" directed towards electoral candidates during the 2024 General Election.¹⁷⁸ While the Government has publicly acknowledged the importance of preventing foreign interference in elections,¹⁷⁹ maintains the Joint Election Security and Preparedness Unit,¹⁸⁰ and has renewed the Defending Democracy Taskforce's mandate to protect candidates

¹⁷² Kate Conger and Lauren Hirsch, "Elon Musk Says He Has Sold X to His A.I. Start-Up xAI", New York Times, 28 March 2025, <https://www.nytimes.com/2025/03/28/technology/musk-x-xai.html>

¹⁷³ National Centre for Social Research, "Trust and confidence in Britain's system of government at record low", 12 June 2024, <https://natcen.ac.uk/news/trust-and-confidence-britains-system-government-record-low>

¹⁷⁴ Online Safety Act 2023 Section 179, <https://www.legislation.gov.uk/ukpga/2023/50/part/10>

¹⁷⁵ National Security Act 2023, <https://www.legislation.gov.uk/ukpga/2023/32/contents>

¹⁷⁶ Such as the Cambridge Analytica scandal. See Carole Cadwalladr and Emma Graham-Harrison, "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach", The Guardian, 17 March 2018, <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>

¹⁷⁷ Electoral Commission, "Information about the cyber-attack", accessed 23 April 2025, <https://www.electoralcommission.org.uk/privacy-policy/public-notification-cyber-attack-electoral-commission-systems/information-about-cyber-attack>

¹⁷⁸ Election Commission, "Report on the 2024 UK Parliamentary general election and the May 2024 elections", accessed 23 April 2025, <https://www.electoralcommission.org.uk/research-reports-and-data/our-reports-and-data-past-elections-and-referendums/report-2024-uk-parliamentary-general-election-and-may-2024-elections#campaigning>

¹⁷⁹ Response by Abena Opong-Asare MP, Cabinet Office Parliamentary Secretary, to a Parliamentary question on preventing foreign interference in elections. See UK Parliament, "Elections: Subversion; Question for Cabinet Office", 13 January 2025, <https://questions-statements.parliament.uk/written-questions/detail/2025-01-13/23400/>

¹⁸⁰ See comments made by Rushanara Ali MP, Minister for Homelessness and Democracy, to the Speaker's Conference on the security of electoral candidates. House of Commons, "Speaker's Conference Oral evidence: Security of Candidates, HC 570", 2 April 2025, <https://committees.parliament.uk/oralevidence/15690/html/>

from threats,¹⁸¹ the level of public information on government plans to address foreign influence activity remains lacking. If the UK government is not transparent about its crisis plans before such an incident occurs during an election period, there is a risk its response could decrease trust in the outcome.

The UK has an opportunity to implement practical measures to mitigate the risks of foreign interference with UK democracy. These measures could include: (1) establishing publicly available protocols for responding to suspected foreign influence during elections as Canada has done;¹⁸² (2) updating the OSA to establish requirements for social media platforms to disclose data on suspected foreign influence activity as part of their transparency reporting; and (3) using a version of the framework previously suggested by Full Fact for responding to information incidents such as allegations of misinformation campaigns by foreign actors.¹⁸³ By taking such steps, the UK has an opportunity to strengthen trust in its democratic institutions and prevent further crises before they happen.¹⁸⁴

Attempts to undermine the 2025 federal election in Germany

The lack of obvious attempts at foreign interference in the UK election may have been the exception rather than the rule. Other countries saw far more concerted efforts to influence voters.

An investigation by the German fact checking organisation, Correctiv,¹⁸⁵ established that a network of approximately 100 fake news websites—some of them set up years in advance—was activated by a Russian influence operation ahead of Germany's federal elections in February this year. False claims about a number of German politicians were created using AI and deepfake technology, including accusations of physical abuse and espionage.

¹⁸¹ See comments made by Dan Jarvis MP, Home Office Minister for Security, to the Speaker's Conference on the security of electoral candidates. House of Commons, "Speaker's Conference Oral evidence: Security of Candidates, HC 570", 2 April 2025, <https://committees.parliament.uk/oralevidence/15690/html/>

¹⁸² Government of Canada, "Cabinet Directive on the Critical Election Incident Public Protocol", updated 24 March 2025, <https://www.canada.ca/en/democratic-institutions/services/protecting-democracy/critical-election-incident-public-protocol/cabinet.html>

¹⁸³ Full Fact, "Framework for Information Incidents", accessed 23 April 2025, <https://fullfact.org/policy/incidentframework/>

¹⁸⁴ For a longer and more detailed set of proposals, see Elizabeth Seger, Hannah Perry and Jamie Hancock, "Epistemic Security 2029: Fortifying the UK's information supply chain to tackle the democratic emergency". Demos, 20 February 2025, <https://demos.co.uk/research/epistemic-security-2029-fortifying-the-uks-information-supply-chain-to-tackle-the-democratic-emergency/>

¹⁸⁵ Alexej Hock, Max Bernhard, Till Eckert, Sarah Thust, "Influence operation exposed: How Russia meddles in Germany's election campaign", CORRECTIV, 24 January 2025, <https://correctiv.org/en/fact-checking-en/2025/01/24/disinformation-operation-russian-meddling-in-german-election-campaign-exposed/>

Germany's Federal Office for the Protection of the Constitution had warned last year of possible attempts by foreign states to distort the outcome of the election,¹⁸⁶ especially against the backdrop of Russia's ongoing invasion of Ukraine. The campaign reported on by Correctiv, in association with Newsguard, saw a variety of fake news stories spread by pro-Russian influencers in Germany.

If Russian attempts to interfere with the election were perhaps not a surprise, the role of Elon Musk during the election campaign was more eye-catching. From the United States, Mr Musk sided openly with the far-right populist party, Alternative for Germany, regularly spreading false and misleading claims on his platform X, in posts that were boosted by the algorithm he owns and which received millions of views.¹⁸⁷

Moldova: a case study

Other European countries, far more vulnerable than Germany, have also been dealing with attempts to interfere in their elections. Full Fact spoke to Alina Radu, CEO of Ziarul de Garda¹⁸⁸ (ZdG), the largest investigative journalism organisation in Moldova about the Moldovan elections in October and November 2024.

ZdG embedded several reporters in groups on the encrypted messaging service, Telegram, in the run-up to the elections. Following its investigation, ZdG says, the Moldovan police found that some 300,000 people in Moldova had a Russian banking app on their phones that allowed them to receive money for helping to support a pro-Russian agenda.

Some of them attended pro-Russian demonstrations and were paid approximately €20 to secure the attendance of another person. Others were paid the same amount to encourage friends to vote for a pro-Russia candidate or against Moldova's application for membership of the European Union. Other media outlets¹⁸⁹ reported on similar vote-buying schemes.¹⁹⁰

Disinformation campaigns can be particularly effective in regions hit by endemic poverty, rural isolation and rampant corruption. According to ZdG, the government was often overwhelmed by the scale of the threat it faced, even though the pro-Russian campaign to sway the Presidential election result was ultimately unsuccessful.

¹⁸⁶ Miranda Murray and Sarah Marsh, "German task force to tackle foreign meddling before election", Reuters, 29 November 2024, <https://www.reuters.com/world/europe/german-task-force-tackle-foreign-meddling-before-election-2024-11-29/>

¹⁸⁷ Alima de Graaf, "Fact check: How Elon Musk meddled in Germany's elections", DW, 21 February 2025, <https://www.dw.com/en/how-elon-musk-meddled-in-germanys-elections/a-71676473>

¹⁸⁸ Ziarul de Garda, accessed 23 April 2025, <https://www.zdg.md/en/>

¹⁸⁹ Constance Victor, "Votes for sale: How Moldova can combat Russia's election meddling", European Council on Foreign Relations, 18 October 2024, <https://ecfr.eu/article/votes-for-sale-how-moldova-can-combat-russias-election-meddling/>

¹⁹⁰ Sarah Rainsford, "Russian cash-for-votes flows into Moldova as nation heads to polls", BBC News, updated 20 October 2024, <https://www.bbc.co.uk/news/articles/c23kdjxxx1jo>

ZdG says it tried to talk to both Telegram and TikTok about their roles in hosting and disseminating disinformation but neither platform was responsive. Moldova faces parliamentary elections later this year, and there is nothing to suggest that a similar—probably state-sponsored—disinformation campaign won't happen again.

Tasks for the UK government

No country is immune from the threat of foreign interference, in a world where online platforms have the ability to spread false and misleading information directly to millions of people, and a combination of money and technology is threatening traditional democratic structures. Russia now openly celebrates its success in waging information wars,¹⁹¹ as AI gives it the ability to produce vast amounts of content.¹⁹²

The UK has a number of systems in place to deal with the nature of the threat, but greater transparency about what they are and how they operate would increase public confidence.

We echo Demos's call for more effective legislation, and would support new laws that require much greater transparency from platforms.¹⁹³ Both the National Security Act and the Online Safety Act address issues of foreign interference and disinformation during elections, but the impact of the measures they set out is likely to continue to be limited because the burden of proof is so high.

Stricter penalties only apply if three conditions are met,¹⁹⁴ involving intent, illegitimacy and the participation of a foreign power. As we set out in last year's Full Fact report,¹⁹⁵ the responsibility for interpreting how the foreign interference offence might work in practice online rests with Ofcom. The regulator should continue to consult academic guidance, as it waits for a body of case law to emerge, in order to understand what is largely uncharted territory.

Finally, we repeat our longstanding call, first made in 2022,¹⁹⁶ for a protocol to warn the public about threats identified by security services during an election campaign. Just because it wasn't needed in 2024, doesn't mean it won't be needed in the future.

¹⁹¹ McKenzie Sadeghi, "Commentary: Russia Used to Deny Interfering. Now it's Celebrating its Successes", NewsGuard's Reality Check, 23 April 2025, <https://www.newsguardrealitycheck.com/p/commentary-russia-used-to-deny-interfering>

¹⁹² Jacob Judah and Fiona Hamilton, "Russia using AI to target Britons with flood of fake news", The Times, 29 April 2025, <https://www.thetimes.com/article/ff3a0c59-4b99-458d-850a-8a5f45356f99?>

¹⁹³ David Hughes, "State threat law watchdog calls for greater transparency from tech giants", Yahoo! News, <https://uk.news.yahoo.com/state-threat-law-watchdog-calls-151458324.html>

¹⁹⁴ Home Office, "Foreign interference: National Security Bill factsheet", updated 1 April 2025, <https://www.gov.uk/government/publications/national-security-bill-factsheets/foreign-interference-national-security-bill-factsheet>

¹⁹⁵ Full Fact, "Full Fact Report 2024: Trust and truth in the age of AI", April 2024, <https://fullfact.org/policy/reports/full-fact-report-2024/#chapter-7-protect-democracy-from-misinformation-and-disinformation-in-the-age-of-ai>

¹⁹⁶ Full Fact, "Full Fact Report 2022: Tackling online misinformation in an open society—what law and regulation should do", February 2022, <https://fullfact.org/policy/reports/full-fact-report-2022/report/#secure-public-confidence-in-how-elections-are-protected-through-transparency>

Chapter 5: Health misinformation

Introduction

Five years after the outbreak of Covid-19, it might seem like the health misinformation crisis has subsided. But the pandemic only accelerated the spread of misinformation, and it remains a persistent issue that generates huge attention.

In response, Full Fact established a dedicated health team. Launched in 2023, it is now composed of a health editor, two fact checkers focused on health policy, and a clinical fact checker who is also a practising GP. This team has tracked a growing number of false and misleading health claims, many centred around unproven alternative therapies and wellness trends. Some of these claims are new, while others are rehashed versions of familiar hoaxes.

At the heart of the team's work is the question: "What harm is this particular piece of misinformation causing?" They then prioritise fact checks on that basis. In 2024, Full Fact took part in an academic study to try to categorise harm more precisely, in order to help direct limited resources towards the most urgent claims.¹⁹⁷ As part of our work using cutting edge technology including generative AI, we have also built a new tool that monitors online videos on health issues at scale, and seeks to rank the misinformation found by the harm it may cause.

We focus effort on health because we know the risks of health misinformation are real—especially when bad actors exploit public vulnerability to profit from sometimes dubious alternative 'therapies'. We also know how large the scale of the challenge is. In 2023, for example, health condition videos on YouTube were viewed more than 5.5bn times in the UK alone, and more than 250bn times worldwide.¹⁹⁸

All of this underscores the need for platforms to step up and take responsibility. In this chapter, we explore the recurring themes in health misinformation and assess the steps health professionals, regulators, and online platforms have taken to counter them.

¹⁹⁷ CAMRI, "Trial finds predictive model helps fact checkers identify false claims with potential to cause harm", CAMRI, 2 April 2025, <https://camri.ac.uk/blog/2025/04/02/trial-finds-predictive-model-helps-fact-checkers-identify-false-claims-with-potential-to-cause-harm/>

¹⁹⁸ Figures provided to Full Fact by YouTube.

Medical misinformation continues to proliferate and evolve across social media platforms

Over the past year, we have fact checked numerous claims about food and drink safety, along with fake ‘cures’ for cancer and other health issues. At the same time, vaccine misinformation remains a significant theme,¹⁹⁹ reflecting its ongoing impact since the Covid-19 pandemic began. Many of these claims try to exploit the growing public anxiety sparked by the pandemic.

Here is a selection of some of the health misinformation claims Full Fact has tracked and reported over the past year:

Food safety misinformation:

- A viral video on Facebook claimed that a chemical in paint thinner was being used in breakfast cereal.²⁰⁰ In reality, trisodium phosphate is used as a food additive in many types of cereals and other foods, and it is generally safe to consume.
- An influencer produced a video claiming that the “Celsius energy drink has four times the amount of daily cyanide that a human being is meant to ingest.”²⁰¹ A cyanide molecule makes up part of the structure of a form of vitamin B12 found in many foods and drinks, but the amount in a Celsius drink is far below the recommended safety limits.
- A viral video on Facebook and Instagram appeared to simulate ‘an experiment’ and claimed to have made the “shocking discovery” that there is graphene oxide in San Pellegrino sparkling water.²⁰² There is not.

Claims about cures for health conditions:

- A video posted on Facebook made claims about an unnamed herbal remedy for diabetes, featuring several clips of celebrities and audio attributed to wellness promoter Barbara O’Neill, who we have fact checked a number of times before.²⁰³ In fact, there is currently no known cure for diabetes, and the speaker has previously been barred from providing health services by Australian health authorities.

¹⁹⁹ In the 12 months to March 2025, roughly 15% of the health material published by Full Fact related to vaccine misinformation.

²⁰⁰ Lolo Kalake, “Paint is not being put in cereal”, Full Fact, 9 August 2024, <https://fullfact.org/health/cereal-paint/>

²⁰¹ Lolo Kalake, “Celsius energy drink does not contain dangerous levels of cyanide”, Full Fact, 5 November 2024, <https://fullfact.org/health/vitaminB12-cyanocobalamin-cyanide/>

²⁰² Lolo Kalake, “No, graphene oxide is not in San Pellegrino water”, Full Fact, 27 February 2025, <https://fullfact.org/health/graphene-oxide-san-pellegrino/>

²⁰³ Lolo Kalake, “Fake video touts 10-day cure for diabetes” Full Fact, 9 July 2024, <https://fullfact.org/health/herbal-diabetes-cure/>

- A study published in June 2024, reported in the Telegraph and MailOnline, made claims about “reversing” autism diagnoses using a combination of methods. However, the National Autistic Society said no conclusions could be drawn from the case study, and some of the interventions used were “questionable”.²⁰⁴
- Posts on Facebook incorrectly claimed studies proved that coriander removes an average of 87% lead, 91% mercury and 74% aluminium from the human body, and is therefore a reliable treatment for heavy metal toxicity.²⁰⁵

Vaccine misinformation:

- A misleading claim on Facebook said the BBC “admitted” HIV was added to Covid-19 vaccines. In reality, a documentary on a trial vaccine which was never rolled out showed researchers in Australia using a protein from the HIV virus to stabilize the vaccine and evoke a stronger immune response: not adding the HIV virus itself.²⁰⁶
- Several viral social media posts suggesting the AstraZeneca Covid-19 vaccine contains the mpox virus were debunked by Full Fact.²⁰⁷ The posts showed the AstraZeneca vaccine’s package leaflet detailing its ingredients, including “recombinant, replication-deficient chimpanzee adenovirus vector encoding the SARS-CoV-2 Spike glycoprotein”—but falsely implied a link between it and developing mpox.

Cancer misinformation:

- A post on Facebook falsely claimed that ‘fake meat’ causes something called “turbo cancer”,²⁰⁸ and linked it to Bill Gates. In fact, the study cited was an article citing another article on Bloomberg, which did not make any link between lab-grown meat and cancer in humans.
- A post on X by former MP and then-independent parliamentary candidate for North West Leicestershire, Andrew Bridgen, claimed a dramatic rise in the number of breast cancer cases in under-45s in the US. But the higher figure he used to illustrate this (297,000 cases in 2023) represented all ages, not just women under 45.²⁰⁹

²⁰⁴ Lolo Kalake, “Scientists have not discovered that ‘autism can be reversed’”, Full Fact, 5 September 2024, <https://fullfact.org/health/autism-reversal/>

²⁰⁵ Lolo Kalake, “Coriander is not a reliable treatment for heavy metal toxicity”, Full Fact, 25 October 2024, <https://fullfact.org/health/coriander-heavy-metals/>

²⁰⁶ Jess Hacker, “A Covid vaccine made with HIV protein was trialled but never rolled out”, Full Fact, 21 February 2025, <https://fullfact.org/health/Covid-vaccine-hiv-trial-not-rolled-out/>

²⁰⁷ Jess Hacker, “Mpox isn’t in the Covid vaccine”, Full Fact, 20 August 2024, <https://fullfact.org/health/covid19-vaccine-mpox/>

²⁰⁸ Lolo Kalake, “No evidence that lab-grown meat causes ‘turbo-cancer’”, Full Fact, 3 May 2024, <https://fullfact.org/health/lab-grown-meat-turbo-cancer/>

²⁰⁹ Lolo Kalake, “Viral US breast cancer stats misinterpreted”, Full Fact, 25 June 2024, <https://fullfact.org/health/breast-cancer-US-bridgen/>

The rise of podcasts

Podcasts have great power and reach, and yet are an unregulated medium. Given the popularity of health content and the number of so-called health ‘influencers’ (with little to no scientific or medical qualifications) active in this space, we are concerned that millions of people may be listening to dangerous misinformation every day.

Last July on Steven Bartlett’s Diary of a CEO podcast, Dr. Aseem Malhotra falsely claimed that fewer deaths would have occurred without the Covid-19 vaccines.²¹⁰ As we said in our fact check shortly afterwards, data from the ONS and the UK Health Security Agency show that vaccines prevented 127,500 deaths in England alone by September 2021. A separate WHO study published in 2024 found that “COVID-19 vaccines have reduced deaths due to the pandemic by at least 57%, saving more than 1.4 million lives in the WHO European Region”.

Following our fact check, the BBC World Service published a detailed investigation into the Diary of a CEO.²¹¹ In an analysis of 15 episodes, it alleged that “each contained an average of 14 harmful health claims that went against extensive scientific evidence”. In response, Flight Studio, Mr Bartlett’s production company, told the BBC that guests were offered “freedom of expression” and were “thoroughly researched”.

Full Fact did not participate in this analysis, and we have not studied the claims in detail, so we cannot offer our own assessment of them. However, the Diary of a CEO wasn’t the first major podcast we have fact checked for health misinformation²¹²—and while free speech matters, so does separating opinion from evidence-based fact.

The platforms that provide and share these podcasts must take misinformation more seriously, while protecting that right to free expression—for instance by requiring podcasts to have a fair and effective corrections policy, so that they or their guests can be contacted quickly and make amends when they get something wrong.

Misinformation on health policy damages trust in important public institutions

Reforming the National Health Service is at the heart of the government’s political agenda, but good policy is only likely to emerge if it is based on accurate data and verifiable facts. Faulty numbers can lead to bad decisions being made, so it is important that national health statistics are as clear as possible

²¹⁰ Leo Benedictus, “Doctor makes misleading Covid vaccine claims on Diary of a CEO podcast”, Full Fact, 26 July 2024, <https://fullfact.org/health/steven-bartlett-diary-ceo-aseem-malhotra-covid-vaccine/>

²¹¹ Jacqui Wakefield, “Steven Bartlett sharing harmful health misinformation in Diary of CEO podcast”, BBC News, 13 December 2024, <https://www.bbc.co.uk/news/articles/c4gpz163vg2o>

²¹² Leo Benedictus, “The Joe Rogan podcast misused English Covid-19 data”, Full Fact, 18 October 2021, <https://fullfact.org/health/joe-rogan-alex-berenson-covid-vaccines-phe/>



Credit: sturti via iStock

One of the most common health policy issues we fact check is confusion around the size of the NHS waiting list, something that is regularly flagged by our AI tools. A key misunderstanding is the assumption that the number of cases on the waiting list equals the number of people waiting—when in fact a significant number of people are on the list for more than one treatment.²¹³

To address this, we've made 13 direct interventions with MPs who have repeated this inaccurate claim, helping to correct the public record.

We also helped clarify misleading coverage of 2023 suicide data from the Office for National Statistics (ONS). While headlines focused on the highest suicide rate since 1999, we highlighted that many reports failed to explain the difference between the rate at which suicides were registered and the rate at which they actually occurred—a critical detail.²¹⁴

A legal change in 2018 also affected how this data was recorded. Coroners are now more likely to rule a death as suicide because they use the civil legal standard ("balance of probabilities") instead of the previous criminal standard ("beyond reasonable doubt").

This means past suicide rates could well have been much higher if coroners were working with the same rules they apply now. How much higher? We don't know—but it's an important context that was missing from public discussion when the new figures emerged.

²¹³ Leo Benedictus, "NHS waiting lists: what you need to know", Full Fact, updated 10 April 2025. <https://fullfact.org/health/nhs-waiting-lists-pre-election-briefing/>

²¹⁴ Leo Benedictus, "Why suicide data can lead us astray", Full Fact, updated 3 October 2024 <https://fullfact.org/health/suicide-statistics/>

Holding public institutions to account through fact checking

As well as combatting misinformation about the NHS, our role in holding public officials to account, and ensuring the information they share is fair and accurate, is essential for upholding trust in the UK's institutions.



Credit: Freepik

What can be done to tackle health misinformation?

A recent Financial Times investigation suggests that Full Fact is on the right track, and that a starting point for tackling health misinformation has to be the rebuilding of trust. Without it, “those who already feel alienated from the healthcare system are less likely to access life-changing innovations, deepening the gulf between the medical haves and have nots.”²¹⁹

Tackling this post-pandemic crisis of confidence will require a multi-stakeholder approach, from local medical professionals, to government interventions and changes to

²¹⁵ Leo Benedictus, “NHS England was wrong to claim its data showed 3.4 million children are ‘unprotected’ against measles”, Full Fact, 11 March 2024, <https://fullfact.org/health/nhs-england-children-unprotected-measles-mmr/>

²¹⁶ Leo Benedictus, “Full Fact secures measles correction from NHS England”, Full Fact, 21 May 2024, <https://fullfact.org/blog/2024/may/full-fact-secures-measles-correction-from-nhs-england/>

²¹⁷ NHS England, “NHS launches catch up campaign for missed MMR vaccines”, 22 January 2024, <https://www.england.nhs.uk/2024/01/nhs-launches-catch-up-campaign-for-missed-mmr-vaccines/#:~:text=The%203.4%20million,england.nhs.uk>

²¹⁸ Leo Benedictus, “NHS England measles figure causes confusion over encephalitis risk”, Full Fact, 14 May 2024, <https://fullfact.org/health/nhs-measles-mmr-encephalitis-international/>

²¹⁹ Sarah Neville, “How to restore trust in doctors in an age of misinformation”, Financial Times, 7 April 2025, <https://www.ft.com/content/2c38694e-f83a-43a4-b41a-a51137e59d52>

platform regulation. The government—and public bodies that report to it—must get their house in order if they are serious about tackling health misinformation. Ensuring that they only release accurate health policy information, and correct any mistakes quickly, is vital in preserving trust between the public and institutions.

Last October, NHS England acted quickly—correcting an error in its new waiting list data within an hour of us flagging it.²²⁰ But earlier in the year, it took far longer to correct its misleading measles vaccination claim. It required multiple emails, meetings, and even a Freedom of Information request before the mistake was properly acknowledged.²²¹

As in all policy areas we cover, a huge responsibility rests with the internet platforms that control so much of the information we consume. We have long called for them to prioritise the visibility of high-quality, reliable information—especially on critical topics like public health.²²²

But we also need the government to put pressure on platforms to do more, and we continue to advocate strongly for the inclusion of health misinformation as a defined harm within the Online Safety Act.²²³ The many examples in this chapter—and the significant time our team continues to spend investigating and debunking health misinformation—highlight the ongoing urgency of this problem. Yet it remains unaddressed. Without a legal requirement for online platforms to conduct adult risk assessments, there's no clear way to know whether or how they're tackling harmful health misinformation.²²⁴ We look more closely at the Online Safety Act in Chapter 6.

Full Fact supports content-neutral interventions that don't rely on censorship or removing posts. Instead, they aim to create a healthier online information environment by ensuring high-quality information reaches people first. Examples include prioritising authoritative sources in search and feed algorithms, using 'read-before-you-share' nudges to slow the spread of viral falsehoods, and clearly labelling content that has been independently fact checked.

These measures help reduce the reach and impact of harmful misinformation without compromising free speech. They offer a practical, proportionate approach that protects the public and focuses on facts, while upholding the right to express diverse opinions—even when they challenge mainstream thinking.

²²⁰ Leo Benedictus, "NHS England corrects waiting list data error", Full Fact, updated 11 October 2024, <https://fullfact.org/health/nhs-england-rtt-waiting-list-error/>

²²¹ Leo Benedictus, "NHS England was wrong to claim its data showed 3.4 million children are 'unprotected' against measles", Full Fact, 11 March 2024, <https://fullfact.org/health/nhs-england-children-unprotected-measles-mmr/>

²²² Home Office preparedness for Covid-19 (Coronavirus) Inquiry, "Written evidence submitted by Full Fact", May 2020, <https://committees.parliament.uk/writtenevidence/5365/pdf/>

²²³ Full Fact, "Online health misinformation in the UK", April 2023, https://fullfact.org/media/uploads/online_health_misinformation_in_the_uk_full_fact.pdf

²²⁴ Full Fact, "Online health misinformation in the UK", April 2023, https://fullfact.org/media/uploads/online_health_misinformation_in_the_uk_full_fact.pdf#page=35

Online misinformation: conclusion and rating

Whether it's the proliferation of manipulated AI content, the re-framing of video footage under false context or the inundation of Facebook groups with hoax posts, misinformation is flooding the internet at a scale that threatens to overwhelm all forms of defence.

Health misinformation has been a particular focus of concern. Concerted action from government, medical professionals and health bodies is essential to ensure the public are presented with accurate and reliable information. But we also continue to advocate for the inclusion of health misinformation as a defined harm within the Online Safety Act. Without this legal safeguard, the vast majority of health misinformation we've highlighted will remain unaddressed.

The potential disruptive impact of synthetic AI content on online misinformation has been a big preoccupation of the last year. Although deepfakes did not dominate the 2024 UK general election, political misinformation and "cheapfakes" were still rife. And it's become evident that public understanding of synthetic online content remains uncertain at best. While some action has been taken to criminalise the creation and sharing of sexually explicit deepfakes much more protection is required to cover political content.

Responsibility for the dissemination of online misinformation largely comes down to how quickly platforms respond to the rapidly emerging falsehoods they host. The 2024 riots exposed an urgent need for both effective crisis response protocols and access to real-time data for researchers and regulators. The government needs to step up and hold large online platforms accountable for their slowness to act in this regard—an issue we return to in detail later in this report.

Rating

- **Volume of online misinformation:** out of control
- **Government response:** swift and robust action required
- **Platform response:** disappointing and insufficient



Legislation

In his first speech after winning the 2024 election, Sir Keir Starmer declared: “The fight for trust is the battle that defines our age.” He was talking primarily about trust in politics, but his argument resonates far more widely. And yet Labour’s 136-page manifesto offers no plan on how to tackle misinformation,²²⁵ something there has been no shortage of in the first year of this Parliament.

Artificial intelligence, on the other hand, features prominently in the manifesto. Labour outlines a strategy to position AI as a driver of innovation, promising to ensure the safe development and use of AI by introducing binding regulation for the small number of companies developing the most powerful models.²²⁶

But as this report highlights, these regulatory ambitions have shifted. The focus is now on an ‘AI Opportunities Action Plan’²²⁷—a pro-growth agenda that prioritises economic potential over safeguards.

This section looks at what’s missing from the government’s legislative agenda: concrete action to counter harmful misinformation and effective regulation to ensure AI develops safely. Ministers have admitted the flaws of the Online Safety Act, but meaningful reform remains absent. We examine what steps this government could take to genuinely protect people online—and go further than any before it.

²²⁵ The Labour Party, “Labour Party Manifesto 2024”, June 2024, <https://labour.org.uk/wp-content/uploads/2024/06/Labour-Party-manifesto-2024.pdf>

²²⁶ The Labour Party, “Labour Party Manifesto 2024”, June 2024, <https://labour.org.uk/wp-content/uploads/2024/06/Labour-Party-manifesto-2024.pdf#page=35>

²²⁷ Department for Science, Innovation and Technology, “AI Opportunities Action Plan”, 13 January 2025, <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan>

Chapter 6: Harmful misinformation in UK legislation



Credit: Freepik

In last year's report, we argued that the Online Safety Act (OSA), which became law in October 2023, was not fit for purpose.²²⁸ One year on, the Act still falls short of its original ambition to lead global regulatory standards and make the UK “the safest place in the world to be online.”²²⁹

The Act has remained controversial over the past year. Like several other policies discussed in this report, it has been the subject of speculation that it could be used as leverage in trade negotiations with the United States. When asked about this, Peter Kyle, Secretary of State for Science Innovation and Technology, and the minister ultimately responsible for the Act, told LBC: “Let me be really clear, the safety of Britons online and offline is not for negotiation.”²³⁰

At Full Fact, protecting free speech and defending freedom of expression—online and offline—are core to our mission. We believe robust fact checking systems help foster open dialogue and self-expression, but we also recognise the need for regulation to protect people from harm.

²²⁸ UK Government, “Online Safety Act 2023”, 2023, <https://www.legislation.gov.uk/ukpga/2023/50>

²²⁹ UK Government, “Online Safety Bill: supporting documents”, 17 March 2022, <https://www.gov.uk/government/publications/online-safety-bill-supporting-documents#what-the-online-safety-bill-does>

²³⁰ Aggie Chambré and Natasha Clark, “Online Safety Act ‘not up for negotiation’ in US trade deal, Tech Secretary tells LBC”, LBC, 9 April 2025, <https://www.lbc.co.uk/politics/uk-politics/online-safety-act-us-trade-deal-peter-kyle/>



Facts matter. Verifiable, proveable, testable facts. Without them, democracy falters, trust erodes and society drifts into a fog of deceit. Fact checkers aren't the enemy of free speech, they are its guardians, ensuring that debate is grounded in reality rather than fantasy."

Sir Stephen Fry, author and broadcaster

In this chapter, we hear from leading voices from each of the main parties in parliament, all of whom offer valuable insight into how legislation is—or isn't—protecting citizens from misinformation. The chapter also revisits Full Fact's calls to amend the OSA and assesses progress, and outlines priority areas for reform. After much talk about what they'll do to improve the Act, the government must now deliver. This is an era-defining opportunity to make the UK genuinely safe online.

Comment

Chi Onwurah MP, Chair of the Science, Technology and Innovation Select Committee

After years of dither and delay, the previous government finally introduced the Online Safety Act to improve safety in an online space with few regulatory controls. Its goals included reducing illegal content, protecting children from harmful material, and holding tech companies accountable for the content they recommend.

However, the Act did not clearly address harms caused by content which is 'legal but harmful,' in part due to concerns over the impact on freedom of expression and definitions of 'truth'. The Act does impose new duties on providers to implement systems and processes that mitigate the risks of illegal content or activity, or content harmful to children, appearing online.

Our inquiry heard detailed evidence on the role social media algorithms played in amplifying false and misleading content during the Southport riots. Evidence to this inquiry has brought to light how social media platforms can profit from crises such as the Southport riots—despite Meta, TikTok and X all claiming they did not. The recommender systems of these platforms prioritise engaging content, regardless of veracity or harm, to maximise time spent on them and divert attention to advertisements.

For this reason, one inquiry session focused on the digital advertising market. The social media companies we spoke to rely on advertising, which makes up between 80% and 98% of their revenues, with Google holding a dominant position in both the supply and demand side of the sector. We have learned how the digital advertising sector is overly complex and opaque, easily exploited by bad actors wishing to profit from false or harmful content. This was seen last summer when the fake news website

‘Channel3Now’ profited from spreading misinformation about the killer. While digital advertising is regulated by the industry-funded Advertising Standards Authority, with the CMA and Ofcom also holding powers, our inquiry has highlighted a potential regulatory gap in the process of online advertising that enables the monetisation of harmful content.

The inquiry will next hear from Ofcom, the Information Commissioners Office, and Department for Science, Innovation and Technology, where members can scrutinise whether the current Online Safety Act fully addresses the significant societal harms of misinformation. The government says it is serious about tackling online harms, but the platforms we heard from said they would not have behaved differently if the Online Safety Act was fully enacted. This suggests the Act would not prevent a repetition of the terrible riots last summer.

Our inquiry began by hearing from some of the community groups most impacted by the riots. We owe it to them, and to everyone else, to ensure it does not happen again. It is the Government’s duty to do so.

Tracking the progress of Full Fact’s calls for changes to the Online Safety Act

This year we have finally seen major progress in the implementation of the Online Safety Act. As of March 2025, the illegal content duties are now in force, giving Ofcom the power to hold platforms accountable.²³¹ If companies fail to act, Ofcom has powers to enforce fines.²³²

Early in the new government’s term, we spoke with government representatives to ask whether they plan to address misinformation under the Act—something we have long called for. They confirmed that they will first implement the legislation as it stands, and only then consider further amendments.

One welcome change already underway is the addition of Researcher Access provisions in the Data (Use and Access) Bill, which amends and updates the Online Safety Act. We explore what this means in practice in Chapter 8.

The responsibility for tackling misinformation sits with Ofcom, through its media literacy duties and through the statutory Advisory Committee created by the OSA. But the formation of the Committee has lacked urgency and momentum, and delays have meant it has not been in position to respond to several major misinformation events over the past year.

²³¹ UK Government, “Online Safety Act: explainer”, updated 24 April 2025, <https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer>

²³² Tony Smith and Angus Crawford, “First Ofcom probe launched into suicide site exposed by BBC”, BBC News, 9 April 2025, <https://www.bbc.co.uk/news/articles/c24q1n6905mo>

Until late April 2025, Ofcom publicly referred to the Committee as the Advisory Committee on Disinformation and Misinformation, words which have fallen out of favour in Washington under the Trump administration. It has now been renamed as the Online Information Advisory Committee, dropping the words on which it was given a legal mandate to focus.²³³ We hope this is not in response to changing political circumstances, and that it will not be reflected in the important work the Committee needs to do.

For some time, we have urged Ofcom to fully use its research powers to investigate harmful online misinformation and disinformation, and make evidence-based recommendations on how to strengthen the Act. But so far, progress has been painfully slow.

Comment

Lord Clement-Jones, Liberal Democrat Peer and Spokesman for the Digital Economy in the House of Lords

The Online Safety Act, while ground-breaking in many respects, falls short in addressing one of the most pressing challenges of our digital age: the proliferation of misinformation and disinformation online. Events last year in Southport have starkly demonstrated how rapidly false information can spread and the real-world harm it can cause.

Despite our best efforts in the Joint Committee on the draft Online Safety Bill and during the passage of the Bill, now Act, its current provisions, particularly the false communications offence, are insufficient for tackling the sophisticated challenges we face. The requirement to prove both knowledge of falsity and intent to harm makes the offence virtually unenforceable at scale, while failing to address broader societal impacts.

The emergence of AI-powered content generation has fundamentally transformed this landscape. We are witnessing an unprecedented convergence of technologies that can generate human-like text, create photorealistic images, produce synthetic videos, and replicate voices with disturbing accuracy. Tools like Midjourney²³⁴ and HeyGen²³⁵ now enable the mass production of sophisticated deepfakes that can convincingly mimic real individuals. The speed and scale at which this content can be created and disseminated through automated bot networks can simply overwhelm our current regulatory framework.

These AI-generated fakes are particularly challenging. They can now outpace traditional fact checking mechanisms and fool even experienced observers. When combined with automated dissemination systems, they can influence public opinion on a massive scale before any correction can be implemented.

²³³ Ofcom, "Ofcom establishes Online Information Advisory Committee", Ofcom, 28 April 2025, <https://www.ofcom.org.uk/about-ofcom/structure-and-leadership/ofcom-establishes-online-information-advisory-committee>

²³⁴ "Midjourney", accessed 23 April 2025, <https://www.midjourney.com/home>

²³⁵ "HeyGen", accessed 23 April 2025, <https://www.heygen.com/>

The challenge extends beyond content removal. We need proactive measures that promote information integrity while protecting freedom of expression.

During periods of uncertainty, such as terror attacks or pandemics, the spread of misinformation poses particular risks. Our regulatory framework must be robust enough to address these scenarios while maintaining appropriate safeguards for legitimate journalism and democratic discourse.

The solution lies in a multi-faceted approach: strengthening platform accountability, enhancing content authentication, empowering users, and promoting digital literacy. These measures need not impinge on essential freedoms—indeed, they could enhance the quality of online discourse.

Large social media platforms must bear greater responsibility. They should face clear legal obligations to address fake news and develop tools enabling users to think critically about the content they encounter. Importantly, platforms should be required to distribute corrections retroactively to users exposed to false information.

Transparency must be enhanced through maintained advertising archives, particularly for political and high-risk personalised advertising.

User empowerment represents another crucial element. Government should facilitate access to fact checking tools and independent verification resources.

The challenge extends beyond content removal. We need proactive measures that promote information integrity while protecting freedom of expression. Several practical solutions deserve immediate consideration. Mandatory implementation of digital signature authentication tools could significantly enhance users' ability to verify content authenticity. The Content Authenticity Initiative's C2PA specification offers a robust framework for implementing provenance metadata, making it easier to identify false material, whether created deliberately or accidentally.²³⁶

Under Section 64 of the Act, Category 1 services are required to offer adult users the option to verify their identity. Verified users must then have tools to filter or block interactions with non-verified users, reducing exposure to harmful or anonymous content while maintaining user control over their experience.

This new provision, however, will be insufficient to stop anonymous accounts spreading misinformation and disinformation. As many of us argued during the passage of the Act there should be a clear duty on a platform to ensure that users can see whether or not other users are verified and Ofcom's guidance, required under section 65 to assist providers of Category 1 services in complying with this user verification duty, should mandate this.

²³⁶ Coalition for Content Provenance and Authenticity, "CP2A Specifications", accessed 23 April 2025, <https://c2pa.org/specifications/specifications/2.1/index.html>

The adoption of these approaches, combined with a comprehensive media and digital literacy strategy—a new form of digital citizenship—would help citizens navigate our increasingly complex information environment.

The government should act swiftly to amend the Online Safety Act, addressing these critical gaps. The integrity of our democratic processes and the safety of our citizens depend on creating a regulatory framework that matches the sophistication of modern digital threats while preserving the benefits of online communication.

The time for meaningful reform is now, before the next crisis demonstrates the cost of inaction.

Ofcom's Advisory Committee is a work in progress

Last year, we recommended that Ofcom move swiftly to establish what was then known as the Advisory Committee on Disinformation and Misinformation,²³⁷ and we were assured that it would be up and running by the end of 2024. As this report is being finalised, the first meeting of the newly named Online Information Advisory Committee has been delayed until May 2025, following another prolonged recruitment process.²³⁸

The Committee is chaired by Lord Richard Allan, a member of the Ofcom Board and former Meta executive. In next year's report, we will assess whether it has become the authoritative voice on countering misinformation that it needs to be—shaping action across policy, product development and public understanding through a wide range of expertise.

It should be given time to find its feet but the process has not begun well. Alongside the cautious approach embedded in its change of name, Ofcom has changed the terms of reference for the Committee.²³⁹ The scope of its functions and duties has been narrowed to align strictly with what is defined in law under the OSA—a definition Full Fact has always argued is insufficient to deal with the majority of misinformation we encounter, which may not be illegal but still causes harm.

The previous terms of reference, published in November 2024, specified that the Committee was “not limited” to what is in the Act.²⁴⁰ The earlier version also mentioned

²³⁷ Full Fact, “Full Fact Report 2024: Trust and truth in the age of AI”, April 2024 <https://fullfact.org/policy/reports/full-fact-report-2024/#chapter-1-the-online-safety-act-does-not-protect-uk-citizens-from-the-harmful-effects-of-misinformation-and-disinformation>

²³⁸ UK Government, “Online Safety Act: explainer”, updated 24 April 2025 <https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer#how-the-act-tackles-misinformation-and-disinformation>

²³⁹ Online Information Advisory Committee, “Terms of Reference”, Ofcom, 2 April 2025, <https://www.ofcom.org.uk/siteassets/resources/documents/about-ofcom/structure-and-leadership/online-information-advisory-committee/online-information-advisory-committee-terms-of-reference.pdf?v=395782>

²⁴⁰ Advisory Committee on Disinformation and Misinformation, “Terms of Reference”, Ofcom, 13 November 2024, <https://www.ofcom.org.uk/siteassets/resources/documents/about-ofcom/how-ofcom-is-run/mis-and-dis-information-committee/advisory-committee-on-disinformation-and-misinformation-terms-of-reference.pdf?v=386330>

the word “misinformation” nine times, whereas the updated version includes it only once, significantly reducing its emphasis.

As we argued earlier, the language used to frame these issues is less important than the substance of the debate. But we urge the Committee to be bold, and believe its change of name was a disappointing start. It must not shy away from speaking bluntly about its mission—even if that means confronting politically sensitive issues, including those playing out in the United States. We intend to hold the Committee to account, constructively but firmly, because we believe it has a vital role to play, especially in the absence of strong government action. Our vision for its work is clear:

- Urgently assess whether there should be a dedicated Ofcom code of practice on misinformation and disinformation.
- Conduct or commission research on the impact of false or misleading information across regulated services and platforms, and assess the effect it has on the public.
- Make election integrity a top priority, learning from the 2024 general election and applying those lessons before another election takes place.
- Examine existing legislation and regulation and recommend reforms to address online harms more effectively.
- The Committee must not operate in a silo within Ofcom. It should be publicly visible, globally engaged, and accountable to those it serves.
- There should also be a wider connection to ensure citizens affected by harmful misinformation feel they have a voice. Their experience should inform the Committee’s direction and help keep its work grounded in real-world impact.

This Committee should be more than a bureaucratic or academic body. It has the potential to be a national leader in the fight against misinformation, and it needs to include civil society in its deliberations—setting out how online harms can be properly addressed and what legislative solutions may be needed to achieve this. But it will only succeed if it acts boldly, quickly, and inclusively.

Comment

Sir Robert Buckland, Barrister and former Conservative Lord Chancellor and Secretary of State for Justice 2019-2021

In a world where news and information are generated at an ever-faster rate, and the demand for online “clickbait” seems insatiable, we shouldn’t be surprised to see the truth often becoming a casualty, viewed increasingly as subjective rather than objective. As misinformation and disinformation take hold, the “liar’s dividend”, where no-one believes anything from any source, however reputable, becomes a dystopian reality. This cannot be allowed to happen because misinformation has serious consequences for our society.

Recently, in an unprecedented move, the Lady Chief Justice felt compelled to write to the leaders of the two main UK political parties and the Lord Chancellor after an exchange at Prime Minister's Questions about an immigration appeal case, the basic facts of which had been reported inaccurately. This came at a time when judges have been increasingly reporting safety concerns that ultimately affect their independence and the administration of justice itself. This is not the first time that court cases have been fundamentally misreported, and it will not be the last. As the courts themselves provide summaries of significant judgments that can be read and understood by non-lawyers, there is no excuse for this sort of misreporting.

What then, is to be done to restore the balance? If the passage of the new Online Safety Act is any guide, then I think increased regulation will be a very tall order indeed. The new Act has rightly created a framework that will see social media companies face significant fines for hosting harmful material with a particular focus on the need to protect children from harm. If even this limited approach based upon the urgent need to avoid more child deaths and incidents like the Southport murders is facing huge pushback from those who cite freedom of expression, at a time when the new US administration has its face set against what it sees as censorship, the prospects for the UK to take further unilateral action are remote.

The UK Government's Secretary of State for Science, Innovation and Technology, Peter Kyle, has himself conceded that the law in this area is uneven and unsatisfactory. Whilst it is understandable that the focus of debate around the new Act has been around child safeguarding, wider issues about illegal and non-state sponsored disinformation remain. The National Security Act, passed at about the same time, has been the Government's response to criticisms of legislative loopholes, but unless social media platforms are obliged to act against non-state actors peddling myths and lies too, then we remain deeply vulnerable.

Instead, the responsibility to challenge and check misinformation and disinformation is going to continue to fall to organisations like Full Fact, which are providing an invaluable service in checking and correcting inaccurate and false assertions. However diligent and determined these organisations are, however, my fear is that their work alone will never be enough to check the tide. The decision by Meta to end its fact checking programme in the United States and to replace it with a Community Notes system, based on what was introduced by X (formerly Twitter) was not only a challenge to Full Fact but represents a retreat in the fight against misinformation.

As our world continues to shrink, the rise of the conspiracy theory and the use of disinformation by rogue states and non-state actors to disrupt our way of life will only grow. It is up to us, now, to act.

A path towards an improved Online Safety Act: where Labour should begin

In his Statement of Strategic Priorities, Peter Kyle made it clear: “A particular area of focus for the government is the vast amount of misinformation and disinformation that can be encountered by users online. Platforms should have robust policies and tools in place to minimise this content where it relates to their duties under the Act.”²⁴¹

It is a message he has repeated multiple times, and misinformation is mentioned on eight occasions in his statement. So, with clear recognition of the problem and its inclusion in the government’s top priorities, the question remains: why hasn’t this been turned into real action?

Part of the problem is that Mr Kyle has boxed himself into a framework which limits how misinformation is dealt with in the OSA. In an interview with *Politics Home*, he declined to comment on revisiting the ‘legal but harmful’ classification that was taken out of the final version of the Online Safety Bill—pinning the blame instead on previous Conservative governments for removing the provision in the first place.²⁴²

But now that Labour is in power, Mr Kyle has the opportunity to change this. The government must make it clear: misinformation, even when it does not fall into the ‘illegal disinformation’ category, should still be treated as a priority for action. Content of this kind should not be taken down, but sufficient friction should be introduced. Users should be given context and clarity, as seen in our own fact checking partnership with Meta.

Even the Prime Minister has acknowledged the gap, implying that the government needs to develop a strategy for dealing with information that is legal but harmful. Speaking at a Liaison Committee hearing, he said: “In relation to misinformation, obviously there were provisions being argued about here in relation to the Online Safety Bill, which did not make their way into the Act. We need to look at what we can do.”²⁴³

To stop the sort of false claims that fueled the riots following the stabbings in Southport, the government needs to seriously rethink how it tackles misinformation that does not fit into the category of illegal offences.

During the Bill’s passage, the Joint Committee issued a clear warning: “The viral spread of misinformation and disinformation poses a serious threat to societies around the world.” To tackle this they recommended “content neutral safety by design requirements,

²⁴¹ UK Government, “Draft Statement of Strategic Priorities for online safety”, 20 November 2024 <https://www.gov.uk/government/publications/draft-statement-of-strategic-priorities-for-online-safety/draft-statement-of-strategic-priorities-for-online-safety#ministerial-foreword>

²⁴² Tinsui Yeung, “What we heard this week on Sunday with Laura Kuenssberg”, BBC News, 12 January 2025, <https://www.bbc.co.uk/news/live/cg7z91zdpz8t>

²⁴³ Liaison Committee, “Oral evidence: Work of the Prime Minister, HC 848”, House of Commons, 8 April 2025, <https://committees.parliament.uk/oralevidence/15726/pdf/>

set out as minimum standards in mandatory codes of practice.”²⁴⁴ This means actions such as transparent labelling or the promotion of trustworthy information are preferable to removing content. It is time for the government to revisit some of these expert recommendations, and finally take misinformation seriously in all its forms.

The government should also revisit previous asks by Full Fact and others. These include:

- The media literacy duties that platforms have under the Online Safety Act. As drafted there are no **requirements** on online platforms and search engines to undertake media literacy initiatives for their users. Future versions of online safety legislation must ensure that the largest platforms are given a **duty** to provide media literacy programmes which meet users’ needs.
- The government should also mandate platforms to extend their risk assessments beyond illegal content to also include misinformation. This would allow the government, Ofcom and researchers to evaluate emerging threats and adapt legislation accordingly if there are further risks on the horizon.

Where can Ofcom help?

Ofcom’s role as regulator under the Online Safety Act comes with substantial challenges. In previous reports, Full Fact noted that Ofcom had been dealt “a bad hand”.²⁴⁵ But now, it must do more, pushing the limits of the narrow scope of the Act wherever possible.

The statutory misinformation-related powers Ofcom has may be limited, but that should not stop it from defining what best practice should look like—such as how platforms should flag misinformation to users, and how they should apply content filters. This is exactly where the Online Information Advisory Committee should step in and lead.

The riots in summer 2024 highlighted the urgent need for stronger coordination during serious information incidents, when it is necessary to fill information voids and disseminate reliable information quickly. As we set out in our *Framework for Information Incidents*, Ofcom should have a clearer role in leading responses during such moments.²⁴⁶ We believe that the regulator is the obvious choice to coordinate a centralised response system.

²⁴⁴ Joint Committee on the Draft Online Safety Bill, “Draft Online Safety Bill: Report of Session 2021-22”, House of Lords, House of Commons, 10 December 2021, <https://committees.parliament.uk/publications/8206/documents/84092/default/#page=40>

²⁴⁵ Full Fact, “Full Fact Report 2024: Trust and truth in the age of AI”, April 2024 <https://fullfact.org/policy/reports/full-fact-report-2024/#part-1-generative-ai-and-the-information-environment>

²⁴⁶ Full Fact, “Framework for Information Incidents”, accessed 23 April 2025, <https://fullfact.org/policy/incidentframework/>

That role should include setting up a public reporting mechanism for emerging incidents, so that fact checkers, news organisations, community groups and platforms can flag issues and ask Ofcom to convene a rapid response group to discuss severity and response. Full Fact has already developed a model for this kind of system. In drafting it, we worked closely with government officials, and we continue to urge the adoption of a similar system.²⁴⁷

²⁴⁷ Full Fact, "Framework for Information Incidents", accessed 23 April 2025, <https://fullfact.org/policy/incidentframework/>

Chapter 7: AI regulation

Introduction

The widespread use of artificial intelligence has become one of the UK government's main solutions for driving its growth agenda. From cutting down NHS waiting lists to drafting curriculum plans²⁴⁸ and helping the police identify criminals,²⁴⁹ the government is trying to weave cutting-edge technology into the fabric of our public services and its plan for change.²⁵⁰

We support this emphasis on AI, which can create enormous opportunities. In our sector, fact checkers in the UK—and around the world—rely on the tools Full Fact has built, to tackle misinformation at a scale that would be impossible without AI. It's proof that new technology can be a force for good.

When this report uses the term generative AI (also called synthetic media), it refers to machine learning models that can create new content, whether that is audio, text or video. Generative AI models are trained on large datasets so that they can predict the most likely response to prompts or questions based on the patterns in that data.

The government's *AI Opportunities Action Plan*, published in January, takes an optimistic view of AI's broad potential to accelerate business growth, and position the UK as a global leader in the field. According to the plan, AI will “drive better experiences and outcomes for citizens” and create “new opportunities” rather than threaten “traditional patterns of work.”²⁵¹

But for AI to deliver on this promise, the government must treat its risks with the same urgency as its rewards—especially the threat of AI-driven misinformation. Last year, Full Fact called on the previous government to build on existing principles and introduce clear regulation to address the harms caused by AI-generated content.²⁵² That call is now more urgent than ever.

²⁴⁸ UK Government, “AI in schools: What you need to know”, 31 March 2025, <https://educationhub.blog.gov.uk/2025/03/artificial-intelligence-in-schools-everything-you-need-to-know/>

²⁴⁹ Home Office, “Police urged to double AI-enabled facial recognition searches”, 29 October 2023, <https://www.gov.uk/government/news/police-urged-to-double-ai-enabled-facial-recognition-searches>

²⁵⁰ UK Government, “AI Opportunities Action Plan: government response”, 13 January 2025, <https://www.gov.uk/government/publications/ai-opportunities-action-plan-government-response/ai-opportunities-action-plan-government-response>

²⁵¹ UK Government, “AI Opportunities Action Plan”, 13 January 2025, <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan#changes-lives>

²⁵² Full Fact, “Full Fact Report 2024: Trust and truth in the age of AI”, April 2024, https://fullfact.org/media/uploads/ff2024/18042024-full_fact_report_corrected.pdf#page=7

So far, we've seen little progress. The long-awaited AI Bill remains vague and unconfirmed, and the government's relentless focus on AI's economic potential is starting to come at the cost of leadership on regulation. Without a clear direction, the UK risks falling behind—reacting to problems rather than shaping solutions, both at home and on the world stage.

This chapter explores the challenges and risks of regulation and how, with new legislation, there is a gap in dealing with the harms of AI. The chapter also includes an essay from the Ada Lovelace Institute, which reveals both public support for more regulation in the AI space, and how little regulation there has been.

Comment

Michael Birtwhistle, Associate Director (Law & Policy) at the Ada Lovelace Institute

There is currently no holistic body of law governing the development, deployment or use of AI in the UK. Instead, developers, deployers and users abide by the existing fragmented network of rules under the UK regulatory ecosystem. This includes 'horizontal' cross-cutting frameworks, such as human rights, equalities and data protection law, and 'vertical' domain-specific regulation, such as the regime for medical devices.

The last government consulted on how to address the gaps inherent in this setup but did not implement their conclusions.²⁵³ The current government has a manifesto commitment to “ensure the safe development and use of AI models by introducing binding regulation on the handful of companies developing the most powerful AI models and by banning the creation of sexually explicit deepfakes”—although this represents an intention to manage only a narrow subset of AI risks.

The government's deregulatory tendencies stand in sharp contrast to UK public attitudes on AI. Nationally representative polling published by Ada and the Alan Turing Institute in March 2025 showed a significant majority of the public (72%) think laws and regulations would increase their comfort with AI technologies—up 10% since 2022/23.²⁵⁴ 88% of people believe it is important that the government or regulators have powers to stop the use of an AI product if it causes harm, actively monitor the risks posed by AI systems, develop safety standards on AI use, and access information about the safety of AI systems from developers—powers the government seems unlikely to legislate for.

²⁵³ UK Government, “AI regulation: a pro-innovation approach”, 3 August 2023, <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach>

²⁵⁴ The Ada Lovelace Institute and the Alan Turing Institute, “How do people feel about AI?”, accessed 23 April 2025, <https://attitudestoai.uk/>

In practice, the government's actions have to date been primarily deregulatory or to delay implementation, disincentivising the few actors in the AI ecosystem that could ensure technologies reaching the market are safe and trustworthy from enforcement. Regulators have been asked by the government to explain how they will support the government's growth mission.²⁵⁵ The government's Regulation Action Plan commits it to cutting the costs of regulation on business by 25% by the end of the Parliament and the Spring Statement declared an intention to "challenge excessive risk aversion in our regulatory system". The Competition and Markets Authority—one of the leading regulators investigating the impacts of AI foundation models²⁵⁶—had its chair replaced with an ex-Amazon executive and its investigation into the Microsoft-OpenAI merger subsequently dropped.²⁵⁷ It remains unclear whether these developments will inhibit regulators' considerable efforts to date to understand and mitigate AI impacts.

The government's AI Opportunities Plan describes significant ambitions to grow AI adoption, infrastructure, skills, and public sector use—but little action on mitigating AI risks. The AI Regulation Bill mentioned in the 2024 King's Speech²⁵⁸ is now not expected in the first Parliamentary session, although a consultation may be forthcoming later this year. Its scope ("tomorrow's models not today's") is however expected to be so narrow that it will not provide meaningful mechanisms to manage the impacts of recent systems like ChatGPT, despite DSIT concluding a year ago that the current approach "leaves the developers of these systems unaccountable...ultimately requir[ing] legislative action".²⁵⁹

Meanwhile, the Data Use and Access Bill expected to receive Royal Assent by the summer makes it much easier to perform automated decision-making without people's consent,²⁶⁰ and permits the government to restrict relevant safeguards by secondary legislation. The AI Safety Institute which oversees the impacts of frontier AI has had its scope narrowed from 'Safety' to 'Security',²⁶¹ with references to algorithmic bias and other important topics removed from its stated agenda. Resourcing to help regulators address AI promised in the Opportunities Plan was not forthcoming in the Spring Statement.

²⁵⁵ UK Government, "A new approach to ensure regulators and regulation support growth", 31 March 2025, <https://www.gov.uk/government/publications/a-new-approach-to-ensure-regulators-and-regulation-support-growth>

²⁵⁶ Competition & Markets Authority, "AI Foundation Models: Update paper", UK Government, 11 April 2024, https://assets.publishing.service.gov.uk/media/661941a6c1d297c6ad1dfeed/Update_Paper__1_.pdf

²⁵⁷ Imran Rahman-Jones, "UK competition watchdog drops Microsoft-OpenAI probe", BBC News, 5 March 2025, <https://www.bbc.co.uk/news/articles/clyd87dxezvo>

²⁵⁸ His Majesty King Charles III, "The King's Speech 2024", UK Government, 17 July 2024, <https://www.gov.uk/government/speeches/the-kings-speech-2024>

²⁵⁹ UK Government, "A pro-innovation approach to AI regulation: government response", 6 February 2024, <https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response>

²⁶⁰ UK Parliament, "Data (Use and Access) Bill", accessed 24 April 2025, <https://bills.parliament.uk/bills/3825>

²⁶¹ Mark Say, "AI Safety Institute becomes AI Security Institute", UKAuthority, 14 February 2025, <https://www.ukauthority.com/articles/ai-safety-institute-becomes-ai-security-institute/>

Overshadowing the Data Bill has been outcry by the creative industries over the government's proposals on an opt-out scheme for copyright holders from AI training that would overwhelmingly benefit AI developers, leading to Lords amendments that would frustrate the government's plans if they are sustained during ping-pong [back-and-forth votes between the House of Commons and the House of Lords]. Alongside the government's growth mission, this pro-developer approach to AI regulation is being driven by geopolitics; the US made its opposition to foreign regulation of its tech companies clear at the Paris AI Summit,²⁶² and both digital regulation and taxation are reportedly part of negotiations for a UK-US trade deal,²⁶³ which could further disincentivise the government from regulating.

The one area the government has made some headway is around the criminal law, including a planned new offence on sexually-explicit deepfakes, and the possibility of a Home Office consultation on police use of facial recognition. Other efforts across government to raise standards on the use of AI such as the Algorithmic Transparency Recording Standard,²⁶⁴ AI Playbook,²⁶⁵ AI Management Essentials,²⁶⁶ and Model for Responsible Innovation,²⁶⁷ while highly laudable are ultimately voluntary and apply only to central government.

Beyond this, the government has announced no plans to address the broad range of current AI risks comprehensively set out in its own *International AI Safety Report 2025*,²⁶⁸ and the significant gaps in regulatory capability and resourcing. It has made no pronouncement on the approach preferred by the last government; a "contextual, sector-based regulatory framework", that would have issued AI principles for existing regulators to implement, and a set of new "central functions" to support them—which would still have carried significant gaps, but fewer than the status quo.

Continued inaction from government on AI harms carries serious risks to both public trust and business confidence in the technologies, and in the organisations and institutions deploying them—ultimately slowing adoption and reducing the potential benefits.

²⁶² The Spectator, "Read: JD Vance's full speech on AI and the EU", 12 February 2025, <https://www.spectator.co.uk/article/read-jd-vances-full-speech-on-ai-and-the-eu/>

²⁶³ Pippa Crerar, Heather Stewart and Richard Partington, "Starmer offers big US tech firms tax cuts in return for lower Trump tariffs", The Guardian, 2 April 2025, <https://www.theguardian.com/us-news/2025/apr/01/starmer-offered-big-us-tech-firms-tax-cuts-in-return-for-lower-trump-tariffs>

²⁶⁴ UK Government, "Algorithmic Transparency Recording Standard Hub", updated December 17, 2024, <https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub>

²⁶⁵ UK Government, "AI Playbook for the UK Government", 10 February 2025, <https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government>

²⁶⁶ UK Government, "Guidance for using the AI Management Essentials tool", 6 November 2024, <https://www.gov.uk/government/consultations/ai-management-essentials-tool/guidance-for-using-the-ai-management-essentials-tool>

²⁶⁷ UK Government, "The Model for Responsible Innovation", 14 November 2024, <https://www.gov.uk/government/publications/the-model-for-responsible-innovation/the-model-for-responsible-innovation>

²⁶⁸ UK Government, "International AI Safety Report 2025", updated February 18, 2025, <https://www.gov.uk/government/publications/international-ai-safety-report-2025>

Full Fact is a huge believer in the power of AI; recent projects include developing cost effective, AI-powered tools to help find and challenge bad information online.²⁶⁹ We know AI can turbocharge misinformation and is part of the problem, but equally it has to be part of the solution. Online platforms and other stakeholders must lean into adopting AI-powered tools, as the only way to address this onslaught of misinformation at internet scale.

The path to AI legislation in the UK remains unclear

In March 2025, the Prime Minister pledged to introduce “new AI and tech teams sent into public sector departments to drive improvements and efficiency in public services”, adding that “one in 10 civil servants will work in tech and digital roles within the next five years with 2,000 tech apprenticeships turbocharging the transformation.”²⁷⁰

Central to this vision is the principle that “no person’s substantive time should be spent on a task where digital or AI can do it better, quicker and to the same high quality and standard.”²⁷¹ In other words, the government sees AI not only as an efficiency tool but as a structural shift in how the state operates.

Yet, even as AI’s integration into the UK workforce accelerates, in both the public and private sectors, there remains a critical and conspicuous vacuum: the lack of a regulatory framework to protect AI users, whether workers, citizens or our democratic institutions.

A case in point: the UK government’s long-awaited AI bill, which was originally expected in late 2024, remains stalled, but may appear in the coming months.²⁷² Sources, including Peter Kyle himself, confirmed that there was an advanced draft of the bill, which later went “up in the air” after Donald Trump’s re-election.²⁷³ The delay is also rooted in government fears of appearing hostile to tech companies, potentially deterring AI investment from the United States.²⁷⁴

This cautionary stance appears particularly jarring given Labour’s pre-election rhetoric. Amidst widespread concerns around deepfakes ahead of the 2024 election, Mr Kyle had promised that a Labour government would “urgently introduce binding regulation of the small group of companies developing the most powerful AI models that could,

²⁶⁹ Full Fact, “Full Fact AI”, accessed 23 April 2025, <https://fullfact.org/ai/>

²⁷⁰ UK Government, “Prime Minister: I will reshape the state to deliver security for working people”, 12 March 2025, <https://www.gov.uk/government/news/prime-minister-i-will-reshape-the-state-to-deliver-security-for-working-people>

²⁷¹ Ibid.

²⁷² Eleni Courea, “UK delays plans to regulate AI as ministers seek to align with Trump administration”, The Guardian, 24 February 2025, <https://www.theguardian.com/technology/2025/feb/24/uk-delays-plans-to-regulate-ai-as-ministers-seek-to-align-with-trump-administration>

²⁷³ Ibid.

²⁷⁴ Ibid.

if left unchecked, spread misinformation, undermine elections and help terrorists build weapons.”²⁷⁵

This urgent focus on the dangers accompanying the world’s increasingly powerful AI models seems to have disappeared. And with no such regulation in sight, critics like Baroness Kidron have observed that Labour has “gone from believing in tech accountability and user safety to taking marching orders from the tech lobbyists and CEOs”.²⁷⁶

In contrast, the European Union’s AI Act—agreed in early 2024—has set a global benchmark as the first comprehensive legal framework on AI.²⁷⁷ The Act has its critics across the political spectrum, from those who see it as over-zealous regulation to those who argue it fails to protect human rights. But it tries to enshrine a risk-based approach that balances innovation and safeguarding the public interest.²⁷⁸ It seeks to demonstrate that they are not mutually exclusive and can be mutually reinforcing.

The Act rightly classifies “harmful AI-based manipulation and deception” as an “unacceptable risk,” for which technologies are prohibited altogether.²⁷⁹ The UK will not regulate in the same way, and nor are we suggesting that it should. But the approach of classifying risk levels and proactively identifying threats is a good one, and the AI Act sets a standard to which future UK regulatory efforts will inevitably be compared.²⁸⁰

The government needs to set out its stall as soon as possible, in such a way that ensures its AI ambitions are not undercut by public distrust and real-world harm.

Recent developments in UK AI regulation closely mirror those of the US

The UK government appears to be seeking to strategically align itself with the new US administration, instead of supporting existing national interests and priorities.

One of the clearest signs of this shift came in February, when the AI Safety Institute was rebranded as the AI Security Institute (AISi). The change wasn’t just cosmetic—it marked a clear pivot away from issues of bias and freedom of speech towards crime prevention

²⁷⁵ Michael Savage, “Call for action on deepfakes as fears grow among MPs over election threat”, The Guardian, 21 January 2024, <https://www.theguardian.com/politics/2024/jan/21/call-for-action-on-deepfakes-as-fears-grow-among-mps-over-election-threat>

²⁷⁶ Aubrey Allegretti, “Has the technology secretary Peter Kyle been ‘captured’ by Big AI?”, The Times, 26 February 2025, <https://www.thetimes.com/uk/politics/article/peter-kyle-mp-news-ai-technology-b6q5ddp2x>

²⁷⁷ European Commission, “AI Act”, accessed 23 April 2025, <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

²⁷⁸ Ibid.

²⁷⁹ Ibid.

²⁸⁰ Full Fact, “Full Fact Report 2024: Trust and truth in the age of AI”, April 2024, https://fullfact.org/media/uploads/ff2024/18042024-full_fact_report_corrected.pdf#page=17

and national security.²⁸¹ This decision sets a troubling precedent. As AI becomes more embedded in national life, sidelining ethical concerns risks ignoring harms that will directly affect the public.

Even before the rebrand, the original AI Safety Institute failed to lay out a credible plan to tackle misinformation and disinformation.²⁸² Though it was launched as “the first state-backed organisation focused on advanced AI safety for the public interest,”²⁸³ the new focus on national security rather than individuals’ safety represents a step backwards, and another disappointing downgrade of ethical considerations in AI development.

Full Fact publicly challenged this shift. In statements to the media and government, we argued that a consideration of transparency and bias are essential to ensuring safe AI use, and should not be seen as a competing priority with security concerns.²⁸⁴



If the government pivots away from the issues of what data is used to train AI models, it risks outsourcing those critical decisions to the most powerful internet platforms rather than exploring them in the democratic light of day”

Andrew Dudfield, Head of AI, Full Fact

It was disappointing that the government’s response to Full Fact’s argument doubled down on its misstep. A DSIT spokesperson said “bias and freedom of speech have never been priorities for the Institute, and this news makes that explicit.”²⁸⁵

Just days before the rebrand was announced, the government’s alignment with the US approach to AI regulation became even more obvious. At the global AI Action Summit in Paris, the UK and the US were outliers in refusing to sign an international agreement pledging an open, ethical approach to AI development.²⁸⁶ The communique was

²⁸¹ UK Government, “Tackling AI security risks to unleash growth and deliver Plan for Change”, 14 February 2025, <https://www.gov.uk/government/news/tackling-ai-security-risks-to-unleash-growth-and-deliver-plan-for-change>

²⁸² Full Fact, “Full Fact Report 2024: Trust and truth in the age of AI”, April 2024, https://fullfact.org/media/uploads/ff2024/18042024-full_fact_report_corrected.pdf#page=21

²⁸³ UK Government, “AI Opportunities Action Plan”, 13 January 2025, <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan#changes-lives>

²⁸⁴ Christopher McKeon, “Experts ‘deeply concerned’ as Government agency drops focus on bias in AI”, The Independent, 14 February 2025, <https://www.independent.co.uk/news/uk/politics/peter-kyle-government-experts-keir-starmer-rishi-sunak-b2698354.html>

²⁸⁵ Oscar Hornstein, “AI Safety Institute rebrand is a ‘downgrade’ of ethics standards, Full Fact warns”, UKTN, 14 February 2025, <https://www.uktech.news/ai/ai-safety-institute-rebrand-is-a-downgrade-of-ethics-standards-full-fact-warns-20250214>

²⁸⁶ Zoe Kleinman and Liv McMahon, “UK and US refuse to sign international AI declaration”, BBC News, 11 February 2025, <https://www.bbc.co.uk/news/articles/c8edn0n58gwo>

signed by 66 countries,²⁸⁷ all committing to shared values around AI accessibility and responsible development.

Sir Keir Starmer's decision not to sign the agreement has been "interpreted as siding with the US's more lenient approach to AI regulation,"²⁸⁸ compared to the more values-driven approach led by President Macron. A government spokesperson said the UK "didn't sign the declaration because it did not reflect the UK's policy positions on opportunity and security,"²⁸⁹ but did not explain what those policy positions are.

Given the agreement's core principles—Openness, Accountability and Participation—it is hard to see how a UK refusal reflects a coherent strategy to address the potentially harmful impact of AI on our information ecosystem. By refusing to address key safety and ethical issues—especially around the data used to train AI—the government is effectively allowing AI companies to write their own rules of engagement. As fact checkers, we know what happens when platforms mark their own homework.

We repeat what we said in our report last year: this government must urgently define where the use of AI should be more, or less, strictly controlled in order to protect free speech, while building trust in online information and providing safeguards for citizens.²⁹⁰

So far, the Labour government risks repeating the same mistakes as its predecessors—long on ambition, short on substance.

AI regulation in the UK continues to provide a misinformation loophole

The UK's current approach to AI regulation is narrowly focused on security and frontier models—the most advanced systems with capabilities beyond today's tools—in a bid to maximise economic benefit and minimise government intervention. But this does little to address the real, immediate harms caused by the generative AI tools already in wide use—including those producing false or misleading content. Over a third of people in the UK have already used these tools.²⁹¹

²⁸⁷ The President of France, "Pledge for a Trustworthy AI in the World of Work", 11 February 2025, <https://www.elysee.fr/emmanuel-macron/2025/02/11/pledge-for-a-trustworthy-ai-in-the-world-of-work>

²⁸⁸ Eleanor Burleigh, "Keir Starmer and Donald Trump snub EU by rejecting 'woke' deal", Daily Express, 11 February 2025, <https://www.express.co.uk/news/uk/2013077/keir-starmer-donald-trump-eu-ai-deal-paris>

²⁸⁹ Nina Lloyd, "UK did not sign AI communique over 'opportunity and security' concerns – No 10", The Independent, 11 February 2025, <https://www.independent.co.uk/tech/emmanuel-macron-jd-vance-downing-street-paris-government-b2696271.html>

²⁹⁰ Full Fact, "Full Fact Report 2024: Trust and truth in the age of AI", April 2024, https://fullfact.org/media/uploads/ff2024/18042024-full_fact_report_corrected.pdf#page=25

²⁹¹ Deloitte, "Over 18 million people in the UK have now used Generative AI", 31 May 2024, <https://www.deloitte.com/uk/en/about/press-room/over-eighteen-million-people-in-the-uk-have-now-used-generative-ai.html>

As outlined in Chapter 3 of this report, the government has taken important steps to address one aspect of AI harm: the creation of sexually explicit deepfakes. In a cross-departmental move, the creation of such content will now be a criminal offence. The intent goes further, so that “the installation of equipment with intent to commit these offences” will also be prosecuted.²⁹²

This is a welcome sign that there is a willingness across government to tackle deepfakes, alongside an understanding of the harm they can cause. But it is only part of the problem. Another big test lies in tackling non-criminal but still deeply harmful forms of AI-generated misinformation which distort public understanding, undermine trust and can inflict serious damage on individuals and society.

We know that technology is evolving fast, and the law needs to keep up. The Online Safety Act mentions misinformation just twice. Despite being years in development, the government’s plan to tackle AI-powered misinformation is still falling short.

While Ofcom’s responsibility over misinformation on online platforms and search services is limited, the regulator does have a wider role to play through media literacy. It should use this remit to engage seriously with how those specific risks intersect with AI, and how people can be better protected from damaging misinformation as a result.

To label or not? Even the basic choices are complex

AI content labelling is an area of recent development and is a good snap shot of the wider challenges in AI regulation. As noted by Partnership on AI (PAI), a wide range of options exist for the labelling of synthetic content.²⁹³ Most large platforms who have consumer-facing generative AI products are adopting some form of badging to identify AI-generated outputs, but the landscape is inconsistent and we are starting to see evidence of this within our own fact checking work.

In our fact checking partnership with Meta, around 10% of claims checked since October 2024 have focused on synthetic images and videos.

Different companies are using different methods:

- The xAI model Grok adds a visual watermark in the bottom right corner of content.
- The Google Gemini models embed a hidden identifier in the image (referred to as SynthID). It can be detected even if the image is cropped and altered in other ways.

²⁹² UK Government, “Government crackdown on explicit deepfakes”, 7 January 2025, <https://www.gov.uk/government/news/government-crackdown-on-explicit-deepfakes>

²⁹³ Partnership on AI, “Building a Glossary for Synthetic Media Transparency Methods”, 13 December 2023, https://partnershiponai.org/resource/glossary-for-synthetic-media-transparency-methods-part-1/#Indirect_Disclosure

- OpenAI has experimented with adding visual watermarks to images created by the free version of its tools, but not for its paid versions.

Governments are also starting to act. In Spain, for example, a proposed new law will impose substantial fines, up to 35 million euros or 7% of global annual turnover, on companies that fail to properly label AI-generated content.²⁹⁴

This patchwork of solutions creates uncertainty about who is responsible for some of the foundations of our information ecosystem. Labelling can help users make informed choices, but we must be careful to not conflate all AI content with harm. Similarly, provenance tools that verify content that is created by real media outlets don't stop those outlets from publishing false or harmful content.

The EU AI Act will attempt to address this by the time it comes into full force in May 2026. Article 50.2 mandates that providers of AI systems that are “generating synthetic audio, image, video or text content, shall ensure the outputs of the AI system are marked... and detectable as artificially generated or manipulated. Providers shall ensure their technical solutions are effective, interoperable, robust, and reliable as far as this is technically feasible.”²⁹⁵ Whether or not the government seeks to follow the EU's lead more broadly in this area, adopting similar standards would be a sensible step in any future UK legislation.

One thing is clear: we cannot leave it to platforms to decide the rules themselves. If tech companies don't act on their own, they must be legally required to do so. Full Fact has long called for strong regulation to tackle AI-generated misinformation. Technology is moving faster than regulation, but that's no excuse to delay action. Laws and oversight must keep pace with the scale and power of AI tools now shaping how people access and understand information.

Right now, this issue is being treated as a side note. Until that changes, the UK risks outsourcing the integrity of its information environment to a handful of tech companies—and losing public trust in the process.

²⁹⁴ Reuters, “Spain to impose massive fines for not labelling AI-generated content”, 11 March 2025, <https://www.reuters.com/technology/artificial-intelligence/spain-impose-massive-fines-not-labelling-ai-generated-content-2025-03-11/>

²⁹⁵ EU Artificial Intelligence Act, “Article 50: Transparency Obligations for Providers and Deployers of Certain AI Systems”, accessed 23 April 2025, <https://artificialintelligenceact.eu/article/50/>

Legislation: conclusion and rating

Last year's Full Fact report called for substantial improvements to legislation dealing with harmful misinformation and the regulation of AI. Those calls were made to regulators, parliamentarians and to the previous government. Labour has returned to power after 14 years and despite momentum on some issues, not enough has been done.

Full Fact still believes the Online Safety Act is unfit for purpose in its efforts to counter most harmful misinformation, and the false communications offence is too specific to be effective. AI regulation also needs to be updated and improved. More specifics are urgently needed.

The government should define its vision for tackling these two issues. There is mounting concern that where there is a lack of vision there is also a temptation to side with US interests in order to appease the Trump administration, in ways which may not benefit UK citizens.

Rating

- **State of legislation:** More ambitious laws urgently needed
- **Government handling of online safety:** Need for reconsideration
- **Government handling of AI legislation:** Greater focus on safety required



Platforms

Online platforms are falling short in the fight against misinformation. Many are pulling back from using reliable, independent data, stepping away from regulatory frameworks, and ending partnerships with professional fact checkers—without offering credible alternatives. The situation seems bleak.

Instead of prioritising long-term user safety, platforms appear to be making short-term political calculations, aiming to stay in favour with the US government. This comes with a cost for billions of users, who are left with less protection and a more distorted view of the world.

This is a global issue, but the UK is especially vulnerable due to inadequate legislation. In this context, platforms are showing a troubling level of negligence—allowing harmful misinformation to thrive while undermining informed choice in public decision-making.

This section of the report outlines several worrying trends: platforms retreating from fact checking; shifting toward insufficiently robust crowd-sourced models like Community Notes; responding to political pressures; and failing to present any serious plan to address the problem.

We begin by examining the damage caused when fact checking partnerships are removed. Then we ask the critical question: *If not fact checkers, then what?* How do platforms intend to tackle misinformation?

The answers, so far, should concern anyone who values access to good, trustworthy information.

Chapter 8: Platform partnerships with fact checkers

Introduction

2025 is a critical year for the global fact checking community. The fragile relationship between fact checkers and major technology platforms is threatening to unravel, with much of the structured collaboration that has been developed to fight against misinformation and disinformation in danger of disappearing.

At the heart of the shift is politics. With Donald Trump back in the White House, platforms are courting the administration and implementing policy changes to try to stay out of the crosshairs. Meta has moved from banning President Trump's account in the wake of the Capitol riots in 2021²⁹⁶ to become a major donor to his inauguration fund by the end of 2024.²⁹⁷ This pivot comes as it faces an antitrust trial that could break up Meta's portfolio of products and threaten the more than \$30bn in advertising revenue generated by Instagram in the United States.²⁹⁸ Google, under scrutiny for its dominance in digital advertising²⁹⁹ and online search,³⁰⁰ is in a similarly defensive stance in both the US and the EU.³⁰¹

At the same time, platforms are abandoning long-standing fact checking partnerships that were central to their efforts to tackle misinformation and safeguard users. Meta announced in January that it was ending its Third-Party Fact Checking (TPFC) programme in the US,³⁰² setting a harmful precedent: it wrongly says fact checkers are politically biased, and mistakenly suggests platforms don't need the essential work of professional, independent fact checkers in order to keep their users safe.

²⁹⁶ BBC, "Facebook suspends Trump accounts for two years", BBC News, 5 June 2021, <https://www.bbc.co.uk/news/world-us-canada-57365628>

²⁹⁷ Reuters, "Meta donates \$1 mln to Trump's inaugural fund", Reuters, 12 December 2024, <https://www.reuters.com/world/us/meta-donates-1-mln-trumps-inaugural-fund-2024-12-12/>

²⁹⁸ Jody Gody, "Meta's Zuckerberg disputes US antitrust case in trial testimony", Reuters, 15 April 2025, <https://www.reuters.com/sustainability/boards-policy-regulation/facebook-owner-meta-faces-existential-threat-trial-over-instagram-whatsapp-2025-04-14/>

²⁹⁹ Nick Robins-Early, "Google's second antitrust suit brought by US begins, over online ads", The Guardian, 9 September 2024, <https://www.theguardian.com/technology/article/2024/sep/09/google-antitrust-lawsuit-online-ads>

³⁰⁰ Casey Newton, "Google may be on the brink of a breakup", Platformer, 21 April 2025, <https://www.platformer.news/google-antitrust-search-advertising-remedies/?ref=platformer-newsletter>

³⁰¹ Daron Acemoglu, "The US ruled against Google's monopoly — Europe should do the same", Financial Times, 22 April 2025, <https://www.ft.com/content/2b78019f-dc5b-4c59-897d-e90406898fe6>

³⁰² Joel Kaplan, "More Speech and Fewer Mistakes", Meta, 7 January 2025, <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>



We have to have an idea that there are some explanations in the world which are more relevant, which are more powerful, and which accord better with facts.”

Rory Stewart, Podcaster and former Conservative Minister

A few days after the Meta announcement, Google’s president of global affairs, Kent Walker, announced that Google won’t commit to the fact checking requirement in the European Union’s Code of Practice on Disinformation as it “simply isn’t appropriate or effective for our services”.³⁰³ LinkedIn then unsubscribed from its fact checking commitments in the Code soon after.

These moves look less like operational decisions and more like attempts to curry political favour in the US as the wind changes dramatically. If other platforms follow suit, the online information environment will suffer in three concerning ways:

1. Citizens will face more harmful misinformation online, with no fact checks to provide crucial context and caveat. Without these kinds of guardrails based on verifiable facts, people are more likely to make misinformed decisions—about their health, their finances and their democratic choices.
2. Many fact checks will stop reaching the people who need them most. The vast majority of people who consume fact checks do so via external platforms—like Facebook, Instagram, Google News, or Search—without actively seeking them out. These integrations make fact checking frictionless and effective, reaching people when they are most at risk of consuming false or misleading information. Without platform partnerships, much of that visibility could disappear.
3. Fact checkers will become less able to see patterns of bad information developing on platforms. If partnerships are scrapped completely, access to tools like the Meta Content Library, used to understand what is happening within Facebook and Instagram, will be at risk. Full Fact has long called for greater access for fact checkers to platform data to ensure they are able to act effectively and target the most dangerous claims first. Further erosion of access will mean fact checkers will be flying blind, unable to identify, prioritize or respond in a timely fashion. This will be a huge step backwards. As Maldita’s Carlos Hernández-Echevarría writes in an essay in this chapter: “There is no realistic path to identifying and addressing harmful misinformation at scale that can work without the involvement of fact checkers.”

³⁰³ Jess Weatherbed, “EU: Google declines fact-checking integration for Search & YouTube, plans to exit commitments ahead of new laws”, Business and Human Rights Resource Centre, 17 January 2025, <https://www.business-humanrights.org/en/latest-news/eu-google-declines-fact-checking-integration-for-search-youtube-plans-to-exit-commitments-ahead-of-new-laws/>

If tech companies are serious about user safety and combatting misinformation online, they should rebuild constructive, transparent partnerships with fact checking organisations. That includes data access and shared tools that support the vital work they do. Right now, platforms are acting like their share price is more important than the billions of people who rely on them. It is a short-sighted bet, which won't age well.

This chapter explores the current state of the relationship between platforms and fact checkers, examines alternatives like crowdsourced Community Notes approaches, stresses the need for expanded data access, and lays out Full Fact's policy recommendations for securing the future of this essential work.

Comment

Angie Drobnic Holan, Director of the International Fact Checking Network and Editor of PolitiFact

Democracy and facts share an inseparable destiny. Informed citizens gather together, identify problems, predict obstacles, propose solutions, and find new opportunities to improve the health and welfare of society. These are the types of conversations, both casual and formal through elected representatives, that occur daily in democracies. They drive public opinion, policy making, and the passage of new laws. If people can't agree on a common set of facts, these public conversations have a hard time starting and can't truly progress. When people can't reach agreement over basic reality, their conversations can never progress to any sort of vision for a shared future. Deliberative democracy becomes derailed.

New technologies were once considered promising for jumpstarting and even nurturing civic conversations. Social media would be a remarkable new communications network that would connect people for richer, more informed discussion. But as time has gone by, that promise hasn't been fulfilled. Like other new inventions, social media has had profound positives as well as debilitating negatives as it has developed over time. With its attention-grabbing algorithms, it has had the contradictory effect of devaluing evidence, stoking emotions and allowing the loudest voices to prevail. Generative artificial intelligence, meanwhile, has the same promise of synthesizing knowledge at accelerated speed and scale. But it has even more profound challenges with truth and reliability, as AI models generate hallucinations and distort reality with fabricated information.

This is where fact checking journalism remains as important as ever. Fact checking serves as an essential guardian of shared reality. Its purpose isn't to determine electoral outcomes or set policies, but rather to resist false narratives and prevent them from becoming entrenched. By preserving an independent record of evidence, fact checkers create space for the public to think critically.

The rigorous methodology behind fact checking—gathering evidence before reaching conclusions—builds a strong defence against those who would falsify claims for their own ends. Fact checkers worldwide have banded together to create codes of principles and practices that outline the requirements of an independent methodology as well as ethical guidelines for nonpartisanship and independence. When fact checkers communicate these standards clearly to the public, trust is built and developed.

This methodical commitment to evidence fundamentally separates fact checking journalism from other information formats. While entertainment media often traffics in stereotypes and caricatures, journalism confronts complexity and contradictions, grounding its analysis in verifiable evidence. At its core, journalism's defining purpose remains portraying the world accurately, even when the facts are contradictory or mundane.

The future of fact checking—and by extension, the health of our democracies—depends on whether enough of us are willing to value evidence-based discourse. Journalists working under repressive governments understand intuitively that fact checking isn't just about correcting the record, it's about preserving the very concept of shared truth. Without this foundation, self-government becomes impossible. It's important to realize in 2025 that this is not just a fight about facts, but about culture and values. We must collectively insist on standards that value truth over volume, evidence over assertion, and rigor over convenience. In a world increasingly fractured by false claims, fact checking stands as a bulwark against chaos by illuminating truth.

Fact checkers and platforms must recommit to strengthened, symbiotic partnerships

As we set out in Chapter 1, when Meta announced an abrupt U-turn and the end of its TPFC programme in the US in January 2025,³⁰⁴ we reiterated our belief that the public has a right to access the expertise of fact checkers, who are first responders in the information environment.³⁰⁵

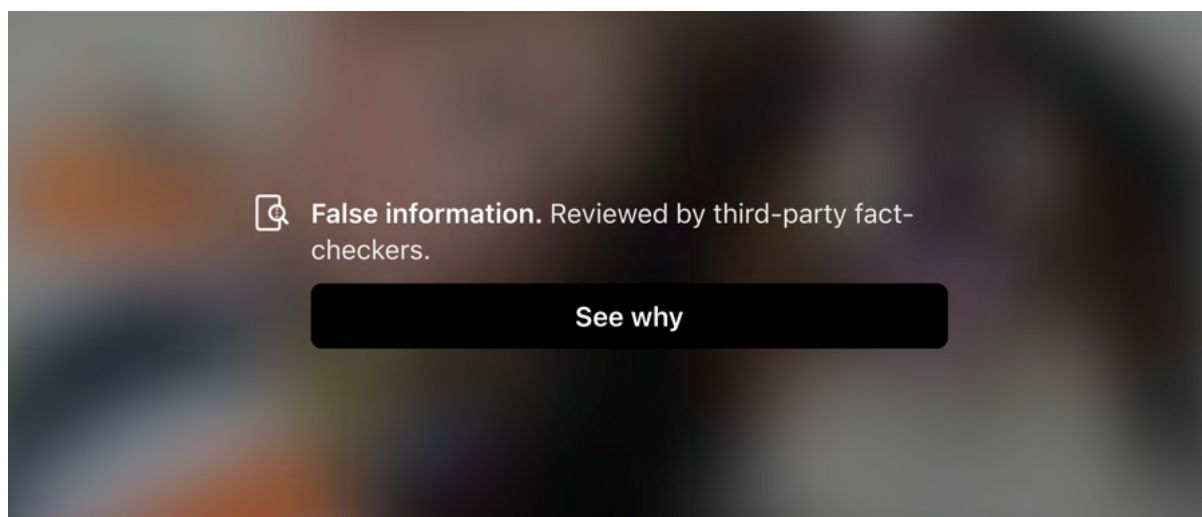
We also rejected—as we continue to do—the accusation of political bias in the fact checking community, which seemed designed for consumption in the White House rather than anywhere else. Meta has spent years praising the work of its fact checking partners, and in its evidence to the UK parliament in December 2024 it described its TPFC programme as a “key part of our approach to combatting misinformation.”³⁰⁶

³⁰⁴ Joel Kaplan, “More Speech and Fewer Mistakes”, Meta, 7 January 2025, <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>

³⁰⁵ Chris Morris, “Full Fact responds to Meta ending support for US fact checkers”, Full Fact, 7 January 2025, <https://fullfact.org/blog/2025/jan/meta-ending-support-for-us-fact-checkers/>

³⁰⁶ Science, Innovation and Technology Committee inquiry into social media, misinformation and harmful algorithms, written evidence submitted by Meta, 18 December 2024, <https://committees.parliament.uk/writtenevidence/132928/pdf#page=10>

That evidence, submitted to the Science, Innovation and Technology Committee, set out exactly how the company works with its fact checking partners, of which Full Fact is one: “When a fact checker rates something as false, our systems are set up to use technology to reduce its distribution so fewer people see it, and add a warning label with more information.”³⁰⁷



Since partnering with Meta in January 2019, Full Fact has checked more than 2,500 misleading, faked, or potentially harmful posts on Facebook and other Meta platforms. We have added context to high-impact content about elections, global conflicts, viral conspiracies, and public health crises. That context helps users make better decisions, and offers them a richer online environment.

But fact checkers don't have any power to remove content from the platforms, and nobody is forced to read the additional information they provide. By offering reliable information where misinformation spreads, we make platforms safer and more trustworthy.

So when Mark Zuckerberg conflates fact checking with censorship, it's not just wrong—it's dangerous. Fact checking enables free speech by making online spaces safer. It is a rigorous, impartial process that involves multiple layers of review and approval.³⁰⁸ Undermining it will only serve to fuel harassment against fact checkers and weaken public trust in the truth—and Meta has made clear that removing TPFC in the US is only the beginning: “Our intention is ultimately to roll out this new approach to our users all over the world.”³⁰⁹

³⁰⁷ Science, Innovation and Technology Committee inquiry into social media, misinformation and harmful algorithms, written evidence submitted by Meta, 18 December 2024, <https://committees.parliament.uk/writtenevidence/132928/pdf/#page=4>

³⁰⁸ Robert Booth, “‘Dispiriting’: factchecker reacts to Meta’s move to scrap role”, The Guardian, 8 January 2025, <https://www.theguardian.com/technology/2025/jan/08/dispiriting-a-factchecker-reacts-to-meta-facebook-move-to-scrap-role>

³⁰⁹ Meta, “Testing Begins for Community Notes on Facebook, Instagram and Threads”, Meta, 13 March 2025, <https://about.fb.com/news/2025/03/testing-begins-community-notes-facebook-instagram-threads/>

Meta's Oversight Board has said Meta's policy and enforcement changes in January 2025 were "announced hastily, in a departure from regular procedure".³¹⁰ We believe they will mean more misinformation, fewer trusted voices, and the likely collapse of several independent fact checking operations. Meta still has time to rethink its change in policy, and we would appeal to it to do so.

In fact, we argue that structured partnerships between platforms and fact checkers must be strengthened and protected. These relationships are mutually beneficial: platforms keep their users safer, and fact checkers reach the audience that needs them most by addressing viral claims in real time, at scale.

TikTok, for example, still maintains some partnerships with fact checkers, and has collaborated with the World Health Organization's (WHO) "Fides" network of more than 800 health experts to fight misinformation. While it is encouraging to see that TikTok appears to be continuing a structured partnership with some fact checkers, it has also said its long-term commitment depends on what other platforms do.³¹¹

Meanwhile, YouTube's efforts to address misinformation have been consistently criticized. Despite a joint letter in 2022 from more than 80 fact checking organisations urging stronger action,³¹² and continued private dialogue with fact checking networks since then, the response has been minimal. Instead of funding partnerships, YouTube expects small fact checkers to produce more video content, offering vague promises of promotion with no meaningful support. With the sheer volume of content uploaded every minute on YouTube, this leaves users dangerously exposed to misinformation—often without any fact checks or context in sight.

Platforms can't cite fact checking partnerships to government if they don't exist

For years, online platforms have cited their partnerships with fact checkers when asked how they're tackling misinformation. But loose collaborations, informal chats, or quietly dropped partnerships don't count as serious solutions. As we'll outline in the next chapter, if platforms are scaling back on fact checking, they need to present clear, credible alternatives—and a long-term vision to match.

Some platforms reference valid, ongoing partnerships: TikTok, for example, noted its partnership with fact checker Logically Facts in written evidence to the Science,

³¹⁰ The Oversight Board, "Wide-ranging decisions protect speech and address harms", 23 April 2025, <https://www.oversightboard.com/news/wide-ranging-decisions-protect-speech-and-address-harms/>

³¹¹ Clara Jiménez Cruz, "What happens if you 'get rid' of fact-checking?", FactCheckHub, 7 February 2025, <https://factcheckhub.com/what-happens-if-you-get-rid-of-fact-checking/>

³¹² Angela Fu, "Over 80 fact-checking organizations sign letter urging YouTube to address misinformation on its platform", Poynter, 12 January 2022, <https://www.poynter.org/fact-checking/2022/youtube-misinformation-fact-checking-letter/>

Innovation and Technology Committee,³¹³ showing it views fact checking partnerships as central to its anti-misinformation strategy.

Others seem happy to reference partnerships they are actively undermining.

Ahead of the 2024 European Parliament elections, Meta emphasised the effectiveness of its labelling system, noting: “Between July and December 2023... over 68 million pieces of content viewed in the EU on Facebook and Instagram had fact checking labels. When a fact checked label is placed on a post, 95% of people don't click through to view it.”³¹⁴

In Meta's written evidence to the Science, Innovation and Technology Committee in December 2024, it cited its global network of more than 100 independent fact checking organisations, and named Full Fact, Reuters and Fact Check Northern Ireland, among others.³¹⁵ Finally, Meta has been actively promoting its partnerships with AFP and AAP in Australia ahead of federal elections.³¹⁶

We welcome platforms citing fact checkers as proof of their action against misinformation, but they must be honest about the future of those partnerships. Now Meta has ended its US programme, the UK government must demand clearer answers. What's the plan without fact checkers, and what data will platforms share to ensure the public still gets timely, accurate information? There is a need for greater transparency—and accountability.

Without expert oversight, Community Notes is not a robust alternative to fact checking

Since ending its US fact checking partnerships, Meta has pitched its new Community Notes feature as a viable replacement—modelled after X's (formerly Twitter) crowdsourced system. Meta described it as “the broad approach we are adopting,”³¹⁷ claiming it is “less biased” than independent fact checkers. In addition TikTok has also announced an intent to conduct a US trial using a system ‘inspired’ by the same technology used by X and Meta to deliver Community Notes.³¹⁸ For the moment, it says its new Footnotes feature will sit alongside its fact checking programme, rather than replacing it.

³¹³ Science, Innovation and Technology Committee, “Written evidence submitted by TikTok”, 28 January 2025, <https://committees.parliament.uk/writtenevidence/137806/default/>

³¹⁴ Marco Pancini, “How Meta Is Preparing for the EU's 2024 Parliament Elections”, 25 February 2024, <https://about.fb.com/news/2024/02/how-meta-is-preparing-for-the-eus-2024-parliament-elections/>

³¹⁵ Science, Innovation and Technology Committee inquiry into social media, misinformation and harmful algorithms, written evidence submitted by Meta, 18 December 2024, <https://committees.parliament.uk/writtenevidence/132928/pdf/#page=4>

³¹⁶ Cheryl Seeto, “How Meta is preparing for the Australian federal election”, Medium, 18 March 2025, <https://medium.com/meta-australia-policy-blog/how-meta-is-preparing-for-the-australian-federal-election-2f773a53ea79>

³¹⁷ Meta, “Testing Begins for Community Notes on Facebook, Instagram and Threads”, Meta, 13 March 2025, <https://about.fb.com/news/2025/03/testing-begins-community-notes-facebook-instagram-threads/>

³¹⁸ Adam Presser, “Testing a new feature to enhance content on TikTok”, TikTok, 16 April 2025, <https://newsroom.tiktok.com/en-us/footnotes>

But the overall shift is not reassuring. Twitter/X never had any formal partnership with fact checkers. Its Community Notes, originally launched as BirdWatch in 2021 and rebranded after Elon Musk's takeover, relies on user consensus to flag and annotate misleading posts.³¹⁹ It is a system built on volunteer opinions rather than on verifiable facts.³²⁰ Meanwhile, Meta's work on its own version of Community Notes had "barely begun"³²¹ when it was announced, suggesting again that the decision to drop years of investment in professional moderation was taken quickly to provide political cover.

But we should be clear: Community Notes can add value, promoting public discourse and encouraging users to share their views and cite relevant links and sources. One research study also found that Community Notes increases by 80% the probability that a tweet is deleted by its creator.³²² Used alongside independent fact checking, it could help enrich the information ecosystem.

But Meta is treating it as a replacement, and that is a serious problem. As Yoel Roth, former head of Trust and Safety at Twitter put it, Community Notes was never intended to replace moderation.³²³ On its own, it simply doesn't do the job, because false claims can slip through if not enough users vote them down.³²⁴ A study from MIT also showed that people participating in Birdwatch were more likely to challenge content they disagree with politically,³²⁵ making notes driven by consensus far more vulnerable to bias and echo chambers.

Worse, Community Notes can be slow and incomplete when covering sensitive issues. Reporting on content about the Middle East conflict highlighted that notes on harmful content took up to 70 hours to be shown to users,³²⁶ and research by Alexios Mantzarlis and Alex Mahadevan concluded that in the final 72 hours before the 2024 US election,

³¹⁹ Ren LaForme, "Meta's user fact-checking is just 'window dressing' without a commitment to truth", Poynter: Opinion, 29 January 2025, <https://www.poynter.org/commentary/2025/crowdsourced-fact-checking-flawed-execution-meta-x-twitter/>

³²⁰ Ren LaForme, "Meta's user fact-checking is just 'window dressing' without a commitment to truth", Poynter: Opinion, 29 January 2025, <https://www.poynter.org/commentary/2025/crowdsourced-fact-checking-flawed-execution-meta-x-twitter/>

³²¹ Casey Newton, "How Meta's take on Community Notes misses the mark", Platformer, <https://www.platformer.news/meta-community-notes-launch/>

³²² Thomas Renault, David Restrepo Amariles, Aurore Troussel, "Collaboratively adding context to social media posts reduces the sharing of false news", ArXiv, 3 April 2024, <https://arxiv.org/abs/2404.02803>

³²³ Vittoria Elliot and David Gilbert, "Elon Musk's main tool for fighting disinformation on X is making the problem worse, insiders claim", Wired, 17 October 2023, <https://www.wired.com/story/x-community-notes-disinformation/>

³²⁴ Ren LaForme, "Meta's user fact-checking is just 'window dressing' without a commitment to truth", Poynter: Opinion, 29 January 2025, <https://www.poynter.org/commentary/2025/crowdsourced-fact-checking-flawed-execution-meta-x-twitter/>

³²⁵ Jennifer Nancy Lee Allen, Cameron Martel, and David Rand, "Birds of a feather don't fact-check each other: Partisanship and the evaluation of news in Twitter's Birdwatch crowdsourced fact-checking program", PsyArXiv Preprints, updated 6 April 2022, https://osf.io/preprints/psyarxiv/57e3q_v1

³²⁶ Davey Alba, Denise Lu, Leon Yin and Eric Fan, "How Musk's X is failing to stem the surge of misinformation about Israel and Gaza", Bloomberg Technology, 21 November 2023, <https://www.bloomberg.com/graphics/2023-israel-hamas-war-misinformation-twitter-community-notes/>

fewer than 6% of the roughly 15,000 notes were marked “helpful” and shown to users.³²⁷ Mahadevan has also argued that while Community Notes is good at flagging obvious examples of “misleading advertising and AI-generated slop”,³²⁸ it is far less effective at dealing with more harmful misinformation.

It is worth noting too that users still rely heavily on fact checkers when writing notes. Research by Spanish fact checkers Maldita found that fact checking organisations are the third most cited source in Community Notes on X, behind only X itself and Wikipedia.³²⁹ That’s remarkable, considering only about 300,000 professional fact checks exist, compared to Wikipedia’s tens of millions of pages and the huge number of posts on X. Users themselves are still leaning on independent fact checkers to counter misinformation.

But Community Notes can be inconsistently applied. AFP staff highlighted one striking recent example: footage of a 2018 rally in support of far-right activist Tommy Robinson in London was falsely presented on X by accounts with huge followings, claiming that it showed current support for Donald Trump in the UK.³³⁰ It gained millions of views, flagged only by a Community Note in Spanish but not in English.

All of this evidence points to one thing: while Community Notes can crowdsource some wisdom, it is no substitute for fact checking. On its own, it is too sporadic and uneven. Meta seems to have chosen a system that prioritises consensus over accuracy to win political points, at the expense of trust in its platforms.³³¹

And the consequences are becoming clear. Whether a post is false or harmful no longer determines its visibility under the Community Notes system—only how engaging it is to the algorithm.³³² Meta’s algorithms will continue to surface viral misinformation, and Community Notes is likely to act only as a thin veneer of moderation. A programme which was only ever intended to complement robust internal systems and processes, is now a main “load bearing pillar” of both Meta and X’s content moderation operations.³³³

³²⁷ Alexios and Alex Mahadevan, “Faked Up #26: X’s Community Notes on Election Day were noisy and marginal, Instagram ads flog deceptive AI influencer get-rich schemes, and bots target Ghana’s election”, Faked Up, 13 November 2024, <https://fakedup.substack.com/p/x-community-notes-election-day-instagram-deceptive-ai-influencers-bots-target-ghana-elections>

³²⁸ Ren LaForme, Tom Jones and Angela Fu, “Fact-checkers are out. The internet gets to vote on the truth now”, Poynter: Opinion, 18 April 2025, <https://www.poynter.org/commentary/2025/fact-checkers-out-community-notes-in/>

³²⁹ Maldita.ES, “Faster, trusted, and more useful: The impact of fact-checkers in X’s Community Notes”, February 2025, https://files.maldita.es/maldita/uploads/2025/02/maldita_informe_community_notes_2024.pdf

³³⁰ Rachel Blundy, LinkedIn post, 6 March 2025, <https://www.linkedin.com/posts/activity-7303427623949414400-qgOJ/?>

³³¹ Chris Morris, “Full Fact responds to Meta ending support for US fact checkers”, Full Fact, 7 January 2025, <https://fullfact.org/blog/2025/jan/meta-ending-support-for-us-fact-checkers/>

³³² Casey Newton, “How Meta’s take on Community Notes misses the mark”, Platformer, <https://www.platformer.news/meta-community-notes-launch/>

³³³ Casey Newton, “How Meta’s take on Community Notes misses the mark”, Platformer, <https://www.platformer.news/meta-community-notes-launch/>

Full Fact's vision for a good Community Notes system

Supporters of Community Notes often frame it as a way to scale fact checking. Adam Mosseri, Head of Instagram, recently pointed out that US-based fact checkers check only around 100 claims a day³³⁴—a number he argues is too small to apply at web scale. But that misses a key point: Community Notes on X frequently cite fact checks precisely because they offer trusted neutral information. And a single fact check often supports dozens of notes, providing common agreement and helping notes reach much higher visibility thresholds.

This should never be a choice between Community Notes and fact checking. The two approaches serve different but complementary roles. Presenting them as incompatible is not only false, it undermines the foundation for a better online information environment.

Full Fact's vision is straightforward: Community Notes should be adopted widely—across platforms and search engines—to add context when users from diverse viewpoints reach consensus. But these systems need a safety net: a credible fallback for when consensus can't be reached, takes too long, or when the content is simply too harmful (though not illegal) to be left unchallenged.

That fallback should be high-quality, independent fact checkers—experts who are funded, fast, and trusted. They provide the neutrality and speed required to address the most dangerous content, ensure notes are seen, and safeguard users.

We challenge the current model, which treats all posts equally, no matter how harmful they are. Failing to down-rank content that has been shown to be false—especially when it relates to public health, elections, or other high-stakes topics—puts society at risk. Platforms must take responsibility for the reach they enable and accept that when harmful content goes viral without intervention, the impact is real—and it's damaging.

Platforms must provide relevant data access to make fact checks effective and worthwhile

For fact checkers to target the most harmful and far-reaching claims effectively, they require cooperation and data-sharing from technology platforms. Access to back-end data helps us understand how misinformation spreads between individuals, communities and across borders. An agreement to provide researcher access to this real-time, proprietary data would help fact checkers make smarter and faster assessments on the dissemination of false and harmful information.

³³⁴ Alex Mahadevan, LinkedIn post, 3 March 2025, https://www.linkedin.com/posts/alexmahadevan_instagrams-adam-mosseri-claimed-that-metas-activity-7302321368367353856-a-2l?

This data could come from platform users, algorithms or trust and safety teams. The more access we have, the more effectively we can identify and prioritise the most harmful content. Full Fact has long called³³⁵ for this level of data access.³³⁶ And while the UK's Data (Use and Access) Bill offers some hope, even past structured partnerships with platforms didn't result in the access we hoped for.

Without sharing of data on this scale, it is so much harder to mount a coordinated, timely response to falsehoods that threaten public health or livelihoods, as was discovered once again during the UK riots in summer 2024.

The Data (Use and Access) Bill—expected to become law in mid-2025—is the government's vehicle for enabling researcher access to platform data. But who qualifies as a researcher is still undecided and it remains unclear whether fact checkers will be included under this definition.

Ofcom will soon recommend how to define researchers, with the government making their final decision in secondary legislation. Full Fact has consistently urged that fact checkers be explicitly included, ideally verified through recognised accreditation bodies such as the International Fact Checking Network (IFCN) and the European Fact Checking Standards Network (EFCSN). Without a clear legal definition, there is a risk that platforms will arbitrarily block fact checkers from accessing their platforms.

Historically, larger platforms have worked more closely with fact checkers. Meta's acquisition of the public insights tool CrowdTangle in 2016, gave researchers and fact checkers real-time access to viral false claims on social media and how they were spreading. But in mid-2024, despite widespread concerns from the fact checking community, Meta shut Crowdtangle down, replacing it with the Meta Content Library. While still useful, it is far less effective in triaging false information in a timely manner.

Unlike Crowdtangle, the Meta Content Library can't be integrated with other tools, including our own algorithms. Previously, we used Crowdtangle's API to automatically collect content from a number of Facebook groups and pages that were known to repeatedly share misleading content. We would then process this content through our own algorithms to help independent fact checkers to find the most harmful content to prioritise. While the Meta Content Library does have an API, it fulfills a very different function and is only available via a "clean room" by accredited individual users under a strict licensing agreement—and data cannot be downloaded or shared.

³³⁵ Full Fact, "Full Fact Report 2020: Fighting the causes and consequences of bad information", 2020, <https://fullfact.org/media/uploads/fullfactreport2020.pdf#page=96>

³³⁶ Full Fact, "Full Fact Report 2024: Trust and truth in the age of AI", April 2024, <https://fullfact.org/policy/reports/full-fact-report-2024/#chapter-5-ensure-fact-checkers-have-the-tools-and-data-needed-to-fight-harmful-misinformation-and-disinformation>

Similarly, X used to provide an open API that allowed fact checkers to monitor content trends and virality, but this has been converted into a research tool with a hefty \$42,000/month price tag to access it at the Enterprise level.³³⁷ Whether this is an economic decision or a deliberate barrier to transparency, the result is the same: we now have far less insight into the spread of harmful information, making it harder to respond effectively with limited resources.

These changes to data access for the fact checking industry, coupled with Meta's retreat from fact checking in the US, should be a wake-up call to UK policymakers. The platforms are not offering the access needed to protect the public from misinformation, so the government must now step in.³³⁸

The Data (Use and Access) Bill is a strong first step, but it must go further. The government should guarantee that fact checkers are included in the scope of the legislation and create incentives for technology companies to provide meaningful access to fact checkers. Crucially, the law should include enforceable penalties for platforms that fail to cooperate.

Comment

Carlos Hernández-Echevarría, Associate Director and Head of Public Policy at Fundación Maldita.es

There are a few reasons why online and social media platforms have proven to be ineffective in dealing with the issue of harmful misinformation. The toolbox these companies use to address these problems is simply unfit for purpose in this particular area.

Automated filters tasked with handling copyright infringements or pornography are widely used and, by and large, quite successful. However, tech companies fail miserably time and time again at spotting viral harmful misinformation and disinformation promoting scams or targeting our elections. Inconvenient as it is, there is no fully automated, friction-free, 100% scalable solution for addressing harmful misinformation, just a perpetual struggle for mitigating its worst consequences. And tech companies find that frustrating.

This is partly because pornography, to use the most obvious example, can be wide ranging but still has some clear characteristics. These might sometimes get confused with artistic expression or nudity, but they nevertheless allow for its automated identification with a high degree of confidence most of the time. Harmful misinformation, on the other hand, almost always looks like any other piece of information by any measure except to the human eye.

³³⁷ X Developer Platform, "About the X API", accessed 24 April 2025, <https://docs.x.com/x-api/getting-started/about-x-api>

³³⁸ Full Fact, "Parliamentary briefing: Data (Use and Access) Bill", 19 November 2024, https://fullfact.org/media/uploads/2024_11_08_rc_parliamentary_briefing_data_bill_lords_2nd_reading.pdf

Platforms know this as does anyone who has spent some time examining this issue. Tech companies accept that large majorities across all ideological groups are (rightly) concerned about harmful misinformation.³³⁹ That's why they have reached out to fact checkers in the past, because they needed to account for that distinctive human factor that often relies on local, specialised knowledge.

These days Meta's Mark Zuckerberg says fact checkers have "destroyed more trust than they've created", and X's Elon Musk routinely refers to them as "liars" when not outright "evil".³⁴⁰ But while Silicon Valley has changed its tune, the challenge remains the same: there is no realistic path to identifying and addressing harmful misinformation at scale that can work without the involvement of fact checkers. Even crowdsourced initiatives like X's Community Notes rely heavily on the work we do.³⁴¹

This is particularly important as politicians and regulators explore how far they can push tech companies to have effective strategies in place to address harmful content. The European Union, for example, has through the Digital Services Act required that the bigger online services have risk mitigation measures in place for disinformation, and both the European Commission³⁴² and the EU Board of Digital Services³⁴³, which gathers all of the EU's national regulators, recognise independent fact checking as an effective way to do that.

Platforms agreed with that, if only briefly. All the major ones signed the Code of Practice on Disinformation back in 2022, in which they committed to cooperate with legitimate fact checking organizations and use their work to empower their users. However, much has changed since then and many of those companies have abandoned or significantly reduced their commitments. It remains to be seen how regulators respond to that when assessing if they are fulfilling their legal obligations. But an underlying, stubborn fact remains: platforms need fact checkers and fact checkers need platforms.

Let me elaborate on that: as much as they hate to include any outsider in the way they make decisions, the top players in the digital industry are in dire need of a partner with the very precious expertise fact checkers have. On the other hand, fact checkers have undoubtedly learned through the years that platforms can give their work the kind of visibility and impact that is just unthinkable almost anywhere else.

³³⁹ Ipsos, "Unesco Survey on the impact of online disinformation and hate speech", September 2023, https://www.unesco.org/sites/default/files/medias/fichiers/2023/11/unesco_ipsos_survey.pdf

³⁴⁰ Elon Musk, "@elonmusk", X, 14 June 2023, <https://x.com/elonmusk/status/1669017475659251713>

³⁴¹ Maldita.ES, "Faster, trusted, and more useful: The impact of fact-checkers in X's Community Notes", February 2025, https://files.maldita.es/maldita/uploads/2025/02/maldita_informe_community_notes_2024.pdf

³⁴² EUR-Lex, "Communication from the Commission – Commission Guidelines for providers of Very Large Online Platforms and Very Large Online Search Engines on the mitigation of systemic risks for electoral processes pursuant to Article 35(3) of Regulation (EU) 2022/2065", 26 April 2024, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52024XC03014&qid=1714466886277>

³⁴³ European Board for Digital Services, "The recognition of the Code of Practice on Disinformation as a code of conduct pursuant to Article 45 of Regulation 2022/2065 (Digital Services Act or "DSA")", <https://ec.europa.eu/newsroom/dae/redirection/document/112680>

That is not to say that collaboration will happen overnight or even that it is a sure thing. Resistance, particularly on the side of industry and in the current US political context, is significant. And fact checkers need to understand that the role they are called to play in the fight for information integrity goes much further than writing articles and giving ratings.

Fact checkers need to produce more useful outputs in more useful formats. They need to cover more ground, and that includes public-facing content to serve their audiences and also providing structured data that can power innovative solutions; solutions to provide context when people encounter dangerous misinformation online, or that allow generative AI models to provide fact-based information more often.

None of this is easy, but it's all essential in the ongoing battle against misinformation.

Chapter 9: Policies for harmful misinformation

Introduction

As outlined in the previous chapter, major technology companies are stepping back from their commitments to tackle misinformation and disinformation and safeguard users. Increasingly, they are refusing to collaborate with the wider fact checking community or even acknowledge the scale of the problem. This leaves platforms in a position to set their own rules without much oversight or accountability.

This chapter examines whether platforms are well equipped to deal with harmful misinformation online, and how their policies are used in practice. We define harmful information as content that, while not illegal, can still seriously mislead people and influence important choices and decisions they make—from health to politics to personal finance.

Current UK legislation fails to hold technology platforms to account for the spread of harmful information online. Without stronger regulation and a minimum set of standards that all platforms must adhere to, governments are allowing some of the most powerful companies the world has ever seen, which have vast power to shape and control our information environment, to regulate themselves. This needs to change.

Full Fact has been calling for several years for progress to be made on this issue, and the cost of UK government inaction has now brought us to a critical point.

Combatting misinformation is no longer a voluntary commitment for platforms in the EU

Change is possible. In February, the European Union turned its once-voluntary Code of Practice (CoP) on Disinformation into a legal obligation under the Digital Services Act (DSA). Now, as explained by the European Fact Checking Standards Network (EFCSN),³⁴⁴ even companies that never signed the Code (or who have since withdrawn) will have to demonstrate the effectiveness of their interventions to comply with the DSA.

³⁴⁴ Ramsha Jahangir, "The EU's Code of Practice on Disinformation is Now Part of the Digital Services Act. What Does It Mean?", TechPolicy.Press, 25 February 2025, <https://www.techpolicy.press/the-eus-code-of-practice-on-disinformation-is-now-part-of-the-digital-services-act-what-does-it-mean/>

This is a potentially powerful precedent, elevating the EU's commitment to fighting disinformation from a voluntary commitment to an enforceable piece of legislation. It shows that laws to address online safety—and misinformation and disinformation—can include detailed frameworks that guide platforms to comply, and help regulators assess their compliance.

As EU regulation tightens, platforms withdraw their support

But instead of rising to the moment, several major platforms began walking away. Just before the Code became law, Google told the European Commission it would withdraw from all fact checking commitments in the CoP.³⁴⁵ YouTube and LinkedIn followed suit, pulling out of the entire fact checking section—despite having signed it in 2022.³⁴⁶

The EFCSN called these withdrawals “extremely concerning.”³⁴⁷ As the EFCSN pointed out, backing away now—just as the rules become enforceable—stands in stark contrast to platforms’ earlier support when commitments were purely voluntary. Full Fact remains concerned about the lack of good faith this represents.

The re-election of Donald Trump, and his open hostility to EU regulation, clearly played a major role in these decisions, but an EFCSN report in December found that many platforms were already failing to follow through on their promises before withdrawing officially. YouTube had stopped reporting on its fact checking partners. LinkedIn’s reviewed video numbers dropped by over 80% in a year, due in part to relying on a single partner to cover 21 languages.³⁴⁸ Meta’s own platforms showed inconsistency too—Instagram displayed far fewer fact checking labels than Facebook, underlining concerns that its system isn’t built to scale,³⁴⁹ and highlighting the challenges its platforms face if they move wholesale to a Community Notes approach.

Collectively, this lack of momentum and forward progress from the world’s largest and most powerful technology companies calls into question the integrity of their initial commitments to the Code, which seem to have been largely performative.

³⁴⁵ Sara Fischer, “Scoop: Google won’t add fact checks despite new EU law”, Axios, 16 January 2025, <https://www.axios.com/2025/01/16/google-fact-check-eu>

³⁴⁶ European Fact-Checking Standards Network, “EFCSN Statement on Platforms’ Reduced Commitments to the Code of Practice on Disinformation”, 22 January 2025, https://efcsn.com/news/2025-01-22_efcsn-statement-on-platforms-reduced-commitments-to-the-code-of-practice-on-disinformation/

³⁴⁷ Ibid.

³⁴⁸ European Fact-Checking Standards Network, “Commitments unfulfilled: Big Tech and the EU Code of Practice on Disinformation”, 18 December 2024, https://efcsn.com/news/2024-12-18_commitments-unfulfilled-big-tech-eu-cop-on-disinfo/

³⁴⁹ Ibid.

Platforms have articulated the risks of misinformation, but actionable policies remain heavily caveated

The beginning of President Trump's second term has continued to have a chaotic impact on the information environment, and the dust is unlikely to settle soon. The early embrace of free speech absolutism and anti-censorship rhetoric has focused even greater attention on antitrust hearings as leading tech executives rapidly try to embrace a new reality.

There has also been a direct challenge to Europe. Vice-President Vance's remarks at the Munich Security Conference, dismissing concerns about misinformation as "ugly Soviet-era words,"³⁵⁰ and asserting that if democracy can be undermined by a few hundred thousand dollars of foreign digital advertising, it wasn't very strong to begin with,³⁵¹ revealed a certain contempt for old allies. But his words also effectively recalibrated the acceptable terminology for platforms and suggested the responsibility for addressing coordinated information interference rests solely with governments, rather than with the platforms themselves.

As these political shocks from Washington continue to reverberate, with unpredictable consequences, large platforms still seem to acknowledge misinformation as a risk in their official documents. But their policies vary in specificity and often lack clarity on engaging with experts and third parties in a systematic and publicly defensible manner:

- Meta's transparency policies for misinformation state "there is no way to articulate a comprehensive list of what is prohibited,"³⁵² because "what is true one minute may not be true the next minute."³⁵³ It claims that any policy that simply prohibits misinformation would be "unenforceable," because "we don't have perfect access to information."³⁵⁴
- TikTok's policies prohibit misinformation "that may cause significant harm to individuals or society, regardless of intent,"³⁵⁵ focusing on health and public safety-related misinformation, climate change misinformation, and conspiracy theories. TikTok's policies also cite relationships with independent fact checkers to "help assess the accuracy of content."³⁵⁶

³⁵⁰ Patrick Wintour, "JD Vance stuns Munich conference with blistering attack on Europe's leaders", The Guardian, 14 February 2025, <https://www.theguardian.com/us-news/2025/feb/14/jd-vance-stuns-munich-conference-with-blistering-attack-on-europes-leaders>

³⁵¹ Emily Atkinson, "JD Vance attacks Europe over free speech and migration", BBC, 14 February 2025, <https://www.bbc.co.uk/news/articles/ceve3wl21x1o>

³⁵² Meta, "Misinformation", last accessed 24 April 2025, <https://transparency.meta.com/en-gb/policies/community-standards/misinformation/>

³⁵³ Ibid.

³⁵⁴ Ibid.

³⁵⁵ TikTok, "Integrity and Authenticity", 17 April 2024, <https://www.tiktok.com/community-guidelines/en/integrity-authenticity>

³⁵⁶ Ibid.

- X's policies, in contrast, focus on civic integrity, authenticity, safety and cybercrime,³⁵⁷ circumventing the category of misinformation entirely. Its policies state that its services cannot be used to manipulate or interfere with elections or civic processes.³⁵⁸
- LinkedIn should be commended for taking a clear, plain English approach, stating that "It is a violation of LinkedIn's Professional Community Policies to post false or misleading content. We remove specific claims, presented as fact, that are demonstrably false or substantially misleading and likely to cause harm. We also remove content that is synthetic or manipulated in a way to misrepresent or distort real-life events without clear disclosure of the fake or altered nature of the material. Content that is false or substantially misleading but not likely to cause harm is not eligible for distribution beyond the author's network."³⁵⁹ However, it is not clear who is responsible for deciding if something is false and how this policy is enforced. LinkedIn does not appear to work with independent fact checkers as part of this.

The scope and content types of different platforms also inform their approach to misinformation, but there are still gaps and inconsistencies.

Video platforms like YouTube and TikTok have developed specific rules for manipulated and digitally generated content. YouTube prohibits "manipulated content" that has been "technically manipulated or doctored in a way that misleads users."³⁶⁰ TikTok does not ban AI-generated content, but requires clear labelling of any kind of altered media that "shows realistic-appearing scenes or people."³⁶¹ Both platforms have gradually improved clarity in these policies, with YouTube now setting out more detailed examples of what may constitute prohibited content.³⁶²

In contrast, X allows some clearly harmful content to slip through. Its policies state that "inaccurate statements about an elected or appointed official, candidate, or political party,"³⁶³ or organic "hyperpartisan"³⁶⁴ content does not violate its rules—leaving major loopholes in its misinformation approach, especially since these policies have weakened under Elon Musk's ownership.

³⁵⁷ X, "Civic integrity policy", August 2023 <https://help.x.com/en/rules-and-policies/election-integrity-policy>

³⁵⁸ Ibid.

³⁵⁹ LinkedIn, "LinkedIn Professional Community Policies", last accessed 24 April 2025, <https://www.linkedin.com/legal/professional-community-policies#be-trustworthy-policy>

³⁶⁰ YouTube, "Misinformation policies", last accessed 24 April 2025, https://support.google.com/youtube/answer/10834785?hl=en-GB&ref_topic=10833358&sjid=12437643136463131705-EU

³⁶¹ TikTok, "Integrity and Authenticity", 17 April 2024, <https://www.tiktok.com/community-guidelines/en/integrity-authenticity#3>

³⁶² YouTube, "Misinformation policies", last accessed 24 April 2025, https://support.google.com/youtube/answer/10834785?hl=en-GB&ref_topic=10833358&sjid=12437643136463131705-EU

³⁶³ X, "Civic integrity policy", August 2023 <https://help.x.com/en/rules-and-policies/election-integrity-policy>

³⁶⁴ Ibid.

Google stands out with a broader, more explicit policy on misinformation related to generative AI,³⁶⁵ which specifies that use of its tools is reliant on users agreeing to “not engage in misinformation, misrepresentation, or misleading activities”. It bans using its tools for scams, frauds, impersonation, and misleading claims—especially in sensitive areas like health, law, or government. It also forbids misrepresenting AI-generated content as human-made. This is one of the few corporate AI policies to still reference “misinformation” directly.

Across platforms, there seems to be broad agreement on the categories of misinformation that pose the greatest threats to society—particularly health and election-related falsehoods. Harmful claims involving medical misinformation, such as a claim about drinking bleach to cure cancer,³⁶⁶ are used as illustrative examples within policies.

But in this world of self-regulation, consequences for misinformation-related policy offences vary widely:

- TikTok may remove, restrict, or make content ineligible for the For You Feed (FYF), the app’s central, algorithmically-driven feed specific to each user.³⁶⁷ It says it may also apply warning labels to content that has been assessed by its fact checking partners and cannot be verified as accurate, and it excludes unverified content about emergencies from wide distribution on FYF.³⁶⁸
- X uses downranking, de-amplification and removal from timelines or search to address misinformation. But enforcement remains uneven and transparency is minimal.
- Meta recently ended its Third-Party Fact Checking programme in the US, but its abrupt change of course has not been fully reflected in its publicly listed policies on the topic. It has, however, focused recently on AI disclosure and fraud/spam guidelines.³⁶⁹

Overall, however, transparency across all platforms needs to improve. They often reference fact checking collaborations without detailing how these partnerships function or how much influence they actually have. For instance, TikTok acknowledges working with independent fact checkers but provides no information on the process, or the

³⁶⁵ Google, “Generative AI-Prohibited Use Policy”, 17 December 2024, <https://policies.google.com/terms/generative-ai/use-policy>

³⁶⁶ ‘Priority content (Category 1 services need to address in their terms and conditions): Harmful health content that is demonstrably false, such as urging people to drink bleach to cure cancer. It also includes some health and vaccine misinformation and disinformation, but is not intended to capture genuine debate.’ UK Parliament, Online Safety Update (written ministerial statement UIN HCWS194), 7 July 2022, “Online Safety Update”, 7 July 2022, <https://questions-statements.parliament.uk/written-statements/detail/2022-07-07/hcws194>

³⁶⁷ TikTok, “Community Guidelines”, 17 April 2024, <https://www.tiktok.com/community-guidelines/en>

³⁶⁸ Ibid.

³⁶⁹ Meta, “Fraud, scams and deceptive practices”, last accessed 24 April 2025, <https://transparency.meta.com/en-gb/policies/community-standards/fraud-scams>

integration of information. Platforms should always be clear about the partnerships they have with fact checkers and the impact this may have on the way people consume content.

The lack of openness undermines trust. If platforms were more transparent about how they collaborate with fact checkers and other third parties, it could foster more coordinated approaches across the industry—moving beyond one-off crisis responses to more permanent solutions. Consistency is key.

The UK government still has a chance to ensure its citizens have a safe and supported experience online

The largest platforms broadly agree that subject areas like health, wealth, and elections need some focused policies to deal with misinformation, even if they are now shying away from using that word. But there is a clear lack of consistency in how they act—and a troubling fragility when faced with political pressure from Washington. We understand that companies will do whatever they deem necessary to protect their share price. But more than in any of Full Fact's previous five reports, it's clear that policies dealing with misinformation are in danger of weakening further over the coming year.

This underlines the urgent need for the UK to stop relying on US-based platforms to self-regulate and instead enforce accountability through legislation. The key is to find the right balance between protecting freedom of expression and protecting people from harm online. Otherwise, we're outsourcing our information standards to companies operating under an increasingly narrow—and politically charged—definition of free speech.

At present, the UK government has few tools to push back when platforms scale down their efforts. Ofcom is still in the early stages of its role as the online safety regulator, and has yet to take a strong stance on misinformation and disinformation. Through its media literacy duties though, Ofcom has an opportunity—and a responsibility—to be more than a passive observer. It should be clearly signalling to platforms that their actions are being watched and will be held to account, and it should suggest legislative or regulatory change should that prove necessary.

Platforms: conclusion and rating

The willingness of very large online platforms (VLOPs) and social media companies to take decisive action to combat the spread of misinformation and disinformation has taken a big step backwards over the course of the past year.

In response to President Trump's return to the White House, decisions are being made to end years of partnership in order to appease the administration. Meta made the most prominent move with the end of Third-Party Fact Checking in the United States and other platforms are following closely behind. Google, for example, won't commit to the fact checking requirement in the EU Commission's Code of Practice on Disinformation.

Of even greater concern is the absence of a robust solution to handling harmful online falsehoods. Tech companies have failed to set out clearly in their terms of service how they will tackle misinformation on their platforms and the UK government now needs to hold them to account.

These platforms rely heavily on income they generate from advertising, subscriptions for premium features and data monetisation. They need to be persuaded that their hosting of false and harmful information is bad for business—that it will lead directly to a fall in the size of their user base and those willing to place ads with them. There is already some evidence that changes in X's content moderation policies have impacted ad revenues for the platform,³⁷⁰ a sign, perhaps, that platforms ignore harmful misinformation at their own peril. But much more needs to be done to persuade these companies of the need for change.

We believe there are still live debates within these organisations about the best way for them to handle the harmful misinformation they host and there is still an opportunity to turn things around. But the time for concerted action is now.

Rating

- **State of platform policies:** Significant backward steps
- **Government's handling of this issue:** Far greater scrutiny required



³⁷⁰ Tech Informed, "Here today, Elon tomorrow: are advertisers abandoning X?", 23 August 2024, <https://techinformed.com/why-advertisers-are-boycotting-x-elon-musk-impact-2024/>

Interventions

Fact checking is crucial in the fight against misinformation, but it is only one piece of the puzzle. To create a better, less harmful information environment, everyone involved in the production and dissemination of factual content needs to intervene at the earliest stage possible to ensure that reliable, evidence-based information is given due prominence.

In this chapter we explore how we can strengthen media literacy to help young people develop the critical thinking skills they need to navigate online misinformation.

We also look more broadly at what is required to build a better information environment from a systemic perspective, to create online spaces that are more resilient to emerging misinformation threats.

As we've emphasised a number of times in this report, combatting misinformation is not a job for fact checkers alone, and this section outlines some of the broader actions needed to address the issue.

Chapter 10: Digital and media literacy

Introduction

As previous chapters have made clear, we cannot simply rely on the platforms, which dominate so much of our information environment, to keep us safe. That is why, in addition to the recommendations we have made on regulation and platform policies, Full Fact is involved in a number of initiatives to help ensure that people are properly equipped to identify false and misleading content, and understand how to avoid becoming unwitting agents in the dissemination of harmful information. We know that media literacy and critical thinking are central to the fight against misinformation and disinformation.

And while the government has been focused on making progress to ensure young people remain safe online, it is important to recognise that literacy is an issue for all ages. In an era of rapid technological change, lifelong learning is essential. To achieve meaningful results, it will take increased funding, cross-departmental support, and a bolstering of Ofcom's media literacy strategy.

media literacy	"the ability to use, understand and create media and communications across multiple formats and services." ³⁷¹
information literacy	"the ability to effectively find, evaluate, use, and share information." ³⁷²
digital literacy	"the ability to both understand and use digitised information." ³⁷³
digital inclusion	"making sure that people have the capability to use the internet to do things that benefit them day to day." ³⁷⁴

³⁷¹ Ofcom, "Future Technology and Media Literacy: Applications of Generative AI", 13 November 2024, <https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/making-sense-of-media/future-technology-trends-and-media-literacy/future-technology-and-media-lit-applications-of-generative-ai.pdf?v=384879#page=5>

³⁷² Gus Macdonald, "What is information literacy?", CILIP, 10 October 2018, <https://www.cilip.org.uk/news/421972/what-is-information-literacy.htm>

³⁷³ Advance HE, "Digital literacies", accessed 21 April 2025, <https://www.advance-he.ac.uk/knowledge-hub/digital-literacies>

³⁷⁴ UK Government, "Government Digital Inclusion Strategy", updated 4 December 2014, <https://www.gov.uk/government/publications/government-digital-inclusion-strategy/government-digital-inclusion-strategy#what-this-strategy-is-about>



We need good data. We need facts. We need sound information. That's the foundation we build everything on.

Dan Snow, historian and broadcaster

For young people, media literacy education starts at school

77% of 11-12-year-olds in the UK use social media, despite the age limits on most social media platforms being 13.³⁷⁵ With 81% of all 8-17-year-olds in the UK using at least one social media app,³⁷⁶ and having access to almost unlimited amounts of information, it is critical for young people to be able to separate what is false and misleading from what is reliable and evidence-based.

All too often, action is only taken after the event, and the tragic suicide of the teenager Molly Russell served as a catalyst for online safety legislation in the UK.³⁷⁷ More recently, Keir Starmer has spoken about the influence on his thinking of the Netflix drama *Adolescence*,³⁷⁸ which details the dangerous impact of what young people consume online. Meanwhile politicians in Westminster have been engaged in an ongoing debate around banning smartphones in schools.³⁷⁹ In an era where parents can no longer control what their children engage with online, it should be self-evident that an effective media literacy strategy is essential.

That means an understanding of misinformation and its consequences, as well as the critical thinking skills to identify it, must be prioritised. For young people, that can be done most easily via the national curriculum. Schools are not the only place where young people should be taught about how to be safe online, but a recent report from the Commission for Countering Online Conspiracy Theories in Schools makes the point that “as an almost universal service for young people, schools are the obvious (although not only) site of intervention” to do this.³⁸⁰ (Research from this report is referenced heavily throughout this chapter as Full Fact’s CEO, Chris Morris, serves on the Commission.)

³⁷⁵ Dr. Sally Burtonshaw, Pete Whitehead, Amy Braier, Dr Denise Baron, Ed Dorrell, Seb Wride, Jules Walkden, Will Yates, “Commission Into Countering Online Conspiracies In Schools”, Public First, February 2025, https://counteringconspiracies.publicfirst.co.uk/Commission_into_Countering_Online_Conspiracies_in_Schools.pdf#page=29

³⁷⁶ Ofcom, “Children and Parents: Media Use and Attitudes Report”, 19 April 2024, <https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/children/children-media-use-and-attitudes-2024/childrens-media-literacy-report-2024.pdf?v=368229#page=13>

³⁷⁷ More details can be found on the Foundation’s website: <https://mollyrosefoundation.org/>

³⁷⁸ Ian Youngs, “Adolescence hard to watch as a dad, Starmer tells creators”, BBC News, 31 March 2025, <https://www.bbc.co.uk/news/articles/cx28neprdp>

³⁷⁹ Josh MacAlister MP, “Protection of Children (Digital Safety and Data Protection) Bill”, House of Commons, 16 October 2024, <https://publications.parliament.uk/pa/bills/cbill/59-01/0016/240016.pdf>

³⁸⁰ Dr. Sally Burtonshaw, Pete Whitehead, Amy Braier, Dr Denise Baron, Ed Dorrell, Seb Wride, Jules Walkden, Will Yates, “Commission Into Countering Online Conspiracies In Schools”, Public First, February 2025, https://counteringconspiracies.publicfirst.co.uk/Commission_into_Countering_Online_Conspiracies_in_Schools.pdf

Comment

Sir Mufti Hamid Patel CBE, Chief Executive of Star Academies and Co-Chair of the Commission into Countering Online Conspiracies in Schools

Establishing the truth has never been straightforward. We are all inclined to believe people with whom we identify, or to subscribe to views that validate our own opinions. Truth is a fluid and tricky concept: incontrovertible certainty is an increasingly rare commodity in a society where the manipulation of text and images is commonplace.

Helping young people to navigate information, sift out incredible or invalid narratives and cement reliable knowledge is just one function of schools, arguably one of the most essential. The quest for truth is of course enshrined in the national curriculum. In English, at key stage 4, pupils should be taught to: “understand and critically evaluate texts” through (amongst other skills) “distinguishing between statements that are supported by evidence and those that are not and identifying bias and misuse of evidence”.³⁸¹ The key stage 3 history curriculum requires pupils to “understand the methods of historical enquiry, including how evidence is used rigorously to make historical claims, and discern how and why contrasting arguments and interpretations of the past have been constructed”.³⁸²

Recognising that there are different interpretations of the past is key to understanding that there are varied constructs of the present too. Our children consume social media at a frenetic pace, bombarded by advertisers and influencers who are hungry for their minds and affiliation. As they scroll through text, liking or rejecting commentary, children are exposed—often subliminally—to world views over which adults have little control, a far cry from libraries full of age-restricted books which were deliberately categorised into subgroups of fact and fiction.

There have always been outliers. Jonathan Swift’s 300-year-old essay *A Modest Proposal* would be deeply troubling to anyone who did not understand the genre of satire,³⁸³ as it appears to advocate infanticide. Generations of readers have been taught to recognise the political backdrop to Swift’s writing and therefore to view his text metaphorically—as a criticism of the political system that reduced millions of Irish peasants to pauperism—rather than to take it literally. But that was just one essay: the proliferation of media text presents a huge navigational challenge.

³⁸¹ Department for Education, “National curriculum in England: English programmes of study”, updated 16 July 2014, <https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study/national-curriculum-in-england-english-programmes-of-study#key-stage-4>

³⁸² Department for Education, “History programmes of study: key stage 3. National curriculum in England”, September 2013, https://assets.publishing.service.gov.uk/media/5a7c66d740f0b626628abcdd/SECONDARY_national_curriculum_-_History.pdf

³⁸³ Jonathan Swift, “A Modest Proposal”, 1729

Immediate access to a plethora of information is not in itself a bad thing, but knowing how to read critically and respond maturely is another issue entirely, and poor media literacy puts children at risk of exploitation.

In an age characterised by the deliberate peddling of misinformation, disinformation and conspiracy theories, readers must become attuned to the motives of content providers. Teachers have to be ready to explain why a story which appears on the surface to be plausible and authentic may be the product of political agitation, intolerance or hatred.

The recent report of the Commission into Countering Online Conspiracies in Schools³⁸⁴ examines some of the difficulties facing educators in the wake of social unrest ignited by unreliable information and hateful rhetoric, spread rapidly on social media. The report, informed by the views of young people, parents and teachers, explores some of the current most prevalent conspiracy theories and their potential impacts. It cites the “information siloes” that separate the digital experiences of the parents and children, and looks for a route map towards connected solutions. The emphasis on “pedagogy not punishment” stresses the point that consuming information requires skills that need to be taught. Teachers are the most trusted adults for many children and they have a vital role to play in equipping future generations of media consumers—but they need to know how best to do this within a rapidly changing media landscape. The Commission—the result of partnership between The Pears Foundation and Star Academies, facilitated by Public First and involving Full Fact’s CEO, Chris Morris, is keen to build upon its initial report by exploring more effective tools to help teachers as they build the digital literacy of their pupils.

The commissioners share a deep commitment to redressing social disadvantage and helping young people to achieve the best possible outcomes—and having strategies to recognise harmful narratives is vital armour for the digital citizen. The Online Safety Act (2023) recognises the danger that social media content presents to young people. Assiduous implementation by Ofcom will help keep children safe from abusive content by placing duties on online platforms. However, the regulatory road map will take time to implement and requires complementary action to equip young people to recognise ‘fake news’ for themselves. Tackling disinformation and conspiracy theories demands a multi-faceted approach, strong political leadership from the Department for Education, Ofsted, multi-academy trusts and local authorities to reassure school leaders and teachers that they can and should address conspiracy beliefs without fear of reprisal. It also requires a strong national commitment to teacher training and ongoing dissemination of effective practice. The investment required will be worthwhile if it prevents a recurrence of the shocking scenes of rioting children on Britain’s streets, prompted by incorrect and toxic posts on social media.

³⁸⁴ Dr. Sally Burtonshaw, Pete Whitehead, Amy Braier, Dr Denise Baron, Ed Dorrell, Seb Wride, Jules Walkden, Will Yates, “Commission Into Countering Online Conspiracies In Schools”, Public First, February 2025, https://counteringconspiracies.publicfirst.co.uk/Commission_into_Countering_Online_Conspiracies_in_Schools.pdf

Schools and educators urgently need:

- Up-to-date training that better equips them to debunk misinformation in the classroom without fear of reprisal³⁸⁵
- A clear, cross-curriculum integration of media literacy education that supports students in multiple academic subjects³⁸⁶
- An earlier start to media literacy education, beginning in primary school, to reflect the engagement of younger pupils with the online world, and capitalise on their higher trust levels in teachers³⁸⁷

Ultimately, it is the responsibility of the government and devolved administrations to address the challenges of how media literacy is taught in schools.³⁸⁸ To date, insufficient attention has been given to this, and we know that students and teachers access increasingly divergent sources of online news. This, in turn, creates the “information siloes” that shape their respective world views.³⁸⁹

We are all drawn to information that reinforces our interests and attitudes. Social media algorithms and online influencers amplify this still further. The challenge for teachers is, therefore, to ensure their students are sufficiently interested to question those they would naturally admire, and to seek out alternative perspectives that extend their critical thinking.

Shortly after taking office last year, the government commissioned a Curriculum and Assessment Review, led by Professor Becky Frances, CEO of the Education Endowment Foundation.³⁹⁰ And in the wake of the riots in the summer of 2024, the education secretary Bridget Phillipson stressed the importance of media literacy as a part of the curriculum, saying: “It’s more important than ever that we give young people the knowledge and skills to be able to challenge what they see online. That’s why our curriculum review will develop plans to embed critical skills in lessons, to arm our children

³⁸⁵ Dr. Sally Burtonshaw, Pete Whitehead, Amy Braier, Dr Denise Baron, Ed Dorrell, Seb Wride, Jules Walkden, Will Yates, “Commission Into Countering Online Conspiracies In Schools”, Public First, February 2025, https://counteringconspiracies.publicfirst.co.uk/Commission_into_Countering_Online_Conspiracies_in_Schools.pdf#page=98

³⁸⁶ Dr. Sally Burtonshaw, Pete Whitehead, Amy Braier, Dr Denise Baron, Ed Dorrell, Seb Wride, Jules Walkden, Will Yates, “Commission Into Countering Online Conspiracies In Schools”, Public First, February 2025, https://counteringconspiracies.publicfirst.co.uk/Commission_into_Countering_Online_Conspiracies_in_Schools.pdf#page=101

³⁸⁷ Dr. Sally Burtonshaw, Pete Whitehead, Amy Braier, Dr Denise Baron, Ed Dorrell, Seb Wride, Jules Walkden, Will Yates, “Commission Into Countering Online Conspiracies In Schools”, Public First, February 2025, https://counteringconspiracies.publicfirst.co.uk/Commission_into_Countering_Online_Conspiracies_in_Schools.pdf#page=100

³⁸⁸ Ofcom, “A Positive Vision for Media Literacy: Ofcom’s Three-Year Media Literacy Strategy”, 7 October 2024, <https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/making-sense-of-media/media-literacy/ofcoms-three-year-media-literacy-strategy-final.pdf#page=7>

³⁸⁹ Dr. Sally Burtonshaw, Pete Whitehead, Amy Braier, Dr Denise Baron, Ed Dorrell, Seb Wride, Jules Walkden, Will Yates, “Commission Into Countering Online Conspiracies In Schools”, Public First, February 2025, https://counteringconspiracies.publicfirst.co.uk/Commission_into_Countering_Online_Conspiracies_in_Schools.pdf#page=32

³⁹⁰ UK Government, “Curriculum and assessment review”, accessed 24 April 2025, <https://www.gov.uk/government/groups/curriculum-and-assessment-review>

against the disinformation, fake news and putrid conspiracy theories awash on social media.”³⁹¹

Full Fact submitted evidence to the Curriculum and Assessment Review and we were pleased to see progress in incorporating more up-to-date skills to help students tackle misinformation, and recommendations on where the curriculum can be used to counter misinformation, when an interim report was published in March 2025.

The evolution of technology presents another challenge, however. As generative AI becomes more ubiquitous, media literacy interventions must adapt to provide us all with the skills and knowledge we need to navigate an increasingly complex information environment. The interim report emphasises the importance of “ensur[ing] that young people are equipped to shape an increasingly AI-powered world,” adding that “they need to be able to navigate misinformation and other challenges, and they also need to be able to take the opportunities that will be available to those who can become the most skilful shapers and operators of AI.”³⁹² We fully support this.

The Finnish Model

Finland’s approach has long been recognised as the gold standard for media literacy, with the country topping the European Media Literacy Index since 2017.³⁹³ It has shown how media literacy can be seamlessly integrated into everyday life, treated as a government priority and embedded effectively across the curriculum.

Finnish educators are required to teach media literacy, and the lessons are incorporated into all subject areas,³⁹⁴ but they have discretion as to how they teach it.³⁹⁵ One teacher, for example, outlined how she encourages students to explore how to manipulate video and photos in order to understand how easy it is to do, while another asked students to research terms like “vaccination” in order to understand how search algorithms work.³⁹⁶

³⁹¹ Nadeem Badshah, “Children to be taught how to spot extremist content and fake news online”, The Guardian, 10 August 2024, <https://www.theguardian.com/education/article/2024/aug/10/uk-children-to-be-taught-how-to-spot-extremist-content-and-misinformation-online>

³⁹² UK Government, “Curriculum and Assessment Review Interim Report”, March 2025, https://assets.publishing.service.gov.uk/media/67e6b43596745eff958ca022/Curriculum_and_Assessment_Review_interim_report.pdf#page=26

³⁹³ Open Society Institute Sofia, “Finland Tops the New Media Literacy Index 2023, Countries Close to the War in Ukraine Remain Among the Most Vulnerable to Disinformation”, 24 June 2023, <https://osis.bg/?p=4450&lang=en>

³⁹⁴ Communications and Digital Committee, “Uncorrected oral evidence: Media literacy”, House of Lords, 1 April 2025, <https://committees.parliament.uk/oralevidence/15724/pdf#page=6>

³⁹⁵ Jenny Gross, “How Finland Is Teaching a Generation to Spot Misinformation”, New York Times, 10 January 2023, <https://www.nytimes.com/2023/01/10/world/europe/finland-misinformation-classes.html>

³⁹⁶ Ibid.

When assessing how media literacy could be further integrated into the UK's education system, as well as being a government priority, the government should look to examples³⁹⁷ like Finland for inspiration.³⁹⁸

Ofcom's media literacy strategy is an important first step, but there's more that can be done

During the passage through parliament of the Online Safety Act, Full Fact successfully campaigned for an amendment that updated Ofcom's media literacy duties. Last year, that effort produced welcome results as Ofcom published its three-year media literacy strategy: the first public articulation of a multi-year strategy to tackle media literacy at scale in two decades.³⁹⁹

The three-year strategy rightly positions media literacy as “everyone's business”, not just Ofcom's responsibility.⁴⁰⁰ It delivers on Ofcom's definition of media literacy as “the ability to use, understand and create media and communications across multiple formats and services”, reflecting the broad and complex nature of online platforms.⁴⁰¹ Most importantly, it considers Ofcom's media literacy duties set out by the Online Safety Act in 2023, including to “help users understand and reduce exposure to mis- and disinformation.”⁴⁰²

We applaud the fact that Ofcom identified misinformation and disinformation, and content of democratic importance, as key priorities on which to engage platforms. The strategy's focus on research into these topics is important, and must remain at the forefront of Ofcom's long-term media literacy efforts.

We were also pleased to see Ofcom recognising the need for media literacy to be delivered by trusted voices across multiple sectors, though we expect an effective, joined-up approach spanning local authorities, police, and education providers—among others—to be a longer-term ambition.

But there is room for improvement. As we outlined in our response to Ofcom's consultation ahead of the publication of the strategy in June 2024,⁴⁰³ there are elements that would benefit from greater clarity, as well as areas that we feel are not sufficiently future-proofed for a rapidly evolving technological landscape.

³⁹⁷ Kavi.fi, “Media Education”, accessed 24 April 2025, <https://kavi.fi/en/media-education/>

³⁹⁸ Kavi.fi, “Media Literacy in Finland: Guidelines”, accessed 24 April 2025 <https://medialukutaitosuomessa.fi/en/guidelines/>

³⁹⁹ Ofcom, “A Positive Vision for Media Literacy: Ofcom's Three-Year Media Literacy Strategy”, 7 October 2024, <https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/making-sense-of-media/media-literacy/ofcoms-three-year-media-literacy-strategy-final.pdf>

⁴⁰⁰ Ibid.

⁴⁰¹ Ibid.

⁴⁰² Ibid.

⁴⁰³ Full Fact, “Full Fact's Response to Ofcom's Three-Year Media Literacy Strategy”, Ofcom, June 2024, <https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/consultation-ofcoms-three-year-media-literacy-strategy/responses/full-fact.pdf?v=370080>

First, Ofcom must prioritise research into what actually works in media literacy education across diverse audiences. While the strategy recognises this need, it's essential to actively build the evidence base for interventions that resonate with people of all ages and socio-economic backgrounds.

Second, the strategy's view of media literacy in the age of generative AI is too narrow. While focusing on AI's impact on elections and young voters is important, the potential for misinformation extends far beyond these areas. Ofcom should embrace a wider research remit to investigate effective methods to help the public identify AI-generated content across all sectors. With technology developing rapidly, the strategy must be flexible and adaptable.

More than anything, we hope to see a stronger focus from Ofcom on platform accountability. Ofcom should move beyond facilitation and begin actively monitoring platform behaviour, which includes naming and shaming platforms that fail to meet media literacy duties or engage sufficiently with new regulations. In the longer term we would like to see the government amend online safety legislation to include a *legal duty* for platforms to actively support and promote media literacy initiatives. With this new direction from the government, Ofcom could be emboldened to demand far more from platforms under their terms of service.

Simply “encouraging” best practices isn’t enough. The proposal of “working with online services to encourage them to adopt our Best Practice Principles for Media Literacy by Design”⁴⁰⁴ could be strengthened by stating clearly what Ofcom will do to hold them to account when platforms fail to adopt such principles. It’s important to make platforms part of the solution.

More government investment and centralised ownership is needed to make media literacy a success

Media literacy efforts require centralised government accountability and ownership. The government must reimagine the structural ownership of media literacy to avoid the lack of coherence it has inherited from its predecessors. Historically, media literacy has been a ‘homeless issue’ with minimal cross-departmental coherence, which led to very little action.

Full Fact recommends that the Cabinet Office coordinate a media literacy agenda across Whitehall through a dedicated cross-departmental taskforce. The Department for Education should support teachers and students; the Home Office should address links between poor media literacy and extremism; the Department for Culture, Media and Sport (DCMS) should bridge the gap between media outlets and media literacy; and

⁴⁰⁴ Full Fact, “Full Fact’s Response to Ofcom’s Three-Year Media Literacy Strategy”, Ofcom, June 2024, <https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/consultation-ofcoms-three-year-media-literacy-strategy/responses/full-fact.pdf?v=370080>

DSIT should focus on online safety. This cross-departmental push should also increase budgets, as more departments will need to take an active role in pushing for change.

Internet Matters has shared a similar call, in which the government could take a “public health approach to media literacy, coordinating the collective efforts of various departments, the third sector, schools and industry.”⁴⁰⁵ They also call for there to be ownership of this issue at Cabinet level.

The government has already adopted a similar model in the new Digital Inclusion and Skills Unit, hosted by DSIT, which aims to establish a ministerial group on digital inclusion which will meet on a quarterly basis.⁴⁰⁶ We encourage the government to think about how a similar model could be adopted for media literacy.

The government needs to match this pressing priority with substantial funding to ensure media literacy programmes can be effectively rolled out and reach the groups that need it most. DSIT’s latest plan,⁴⁰⁷ created under the previous government, allocates £2 million across 13 grant-funded initiatives, which is completely insufficient to address current media literacy needs. The Labour government needs to go further than the previous government in funding this vital issue.

Where does media literacy fit into other literacy strategies?

In February 2025, the government published the long-awaited *Digital Inclusion Action Plan*, a joint collaboration from five departments (Science, Innovation and Technology; Health and Social Care; Education; Work and Pensions; and Housing, Communities and Local Government).⁴⁰⁸ The plan aims to ensure that everyone “has the access, skills, support and confidence to engage in our modern digital society and economy, whatever their circumstances.”⁴⁰⁹

The plan, while ambitious, must not look just at digital literacy but how that interacts with both media literacy and information literacy.⁴¹⁰ If more people are going to use technology for their day-to-day activities, they must also have the skills to think critically about what they are exposed to and how to engage with it safely.

⁴⁰⁵ Internet Matters, “A Vision for Media Literacy”, June 2024, 51. <https://www.flipsnack.com/internetmattersorg/a-vision-for-media-literacy-report-2024/full-view.html>

⁴⁰⁶ UK Government, “Digital Inclusion Action Plan: First Steps”, 26 February 2025, <https://www.gov.uk/government/publications/digital-inclusion-action-plan-first-steps/digital-inclusion-action-plan-first-steps#chapter-3---defining-and-measuring-digital-inclusion>

⁴⁰⁷ UK Government, “Year 3 Media Literacy Action Plan (2023/24)”, 23 October 2023, <https://www.gov.uk/government/publications/year-3-media-literacy-action-plan-202324>

⁴⁰⁸ UK Government, “Digital Inclusion Action Plan: First Steps”, 26 February 2025, <https://www.gov.uk/government/publications/digital-inclusion-action-plan-first-steps/digital-inclusion-action-plan-first-steps>

⁴⁰⁹ Ibid.

⁴¹⁰ UK Government, “Digital Inclusion Action Plan: First Steps”, 26 February 2025, <https://www.gov.uk/government/publications/digital-inclusion-action-plan-first-steps/digital-inclusion-action-plan-first-steps>

The digital inclusion plan makes reference to misinformation through the Curriculum Review, but it lacks detail on how adults may also need other literacy skills to be active participants. As we mention in the first section of this report, online misinformation runs rampant in many corners of the internet and more must be done to ensure that not just one age group benefits from government intervention.

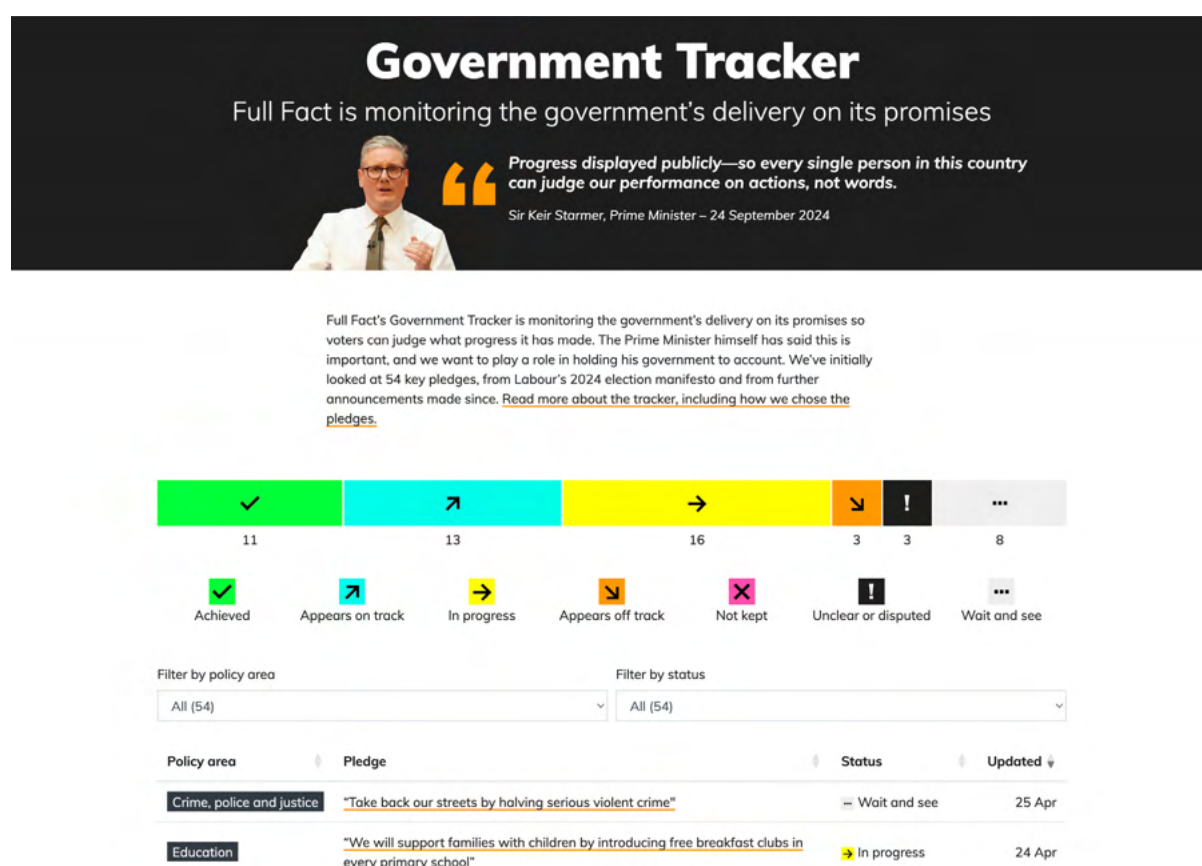
While promoting literacy within young people is a strong place to start, the government must think about interventions that extend to all age groups within society. Last year's Full Fact report looked at diversity within media literacy initiatives, from providing satellite-based internet to rural communities and digital peer support, through to podcasts for parents and workshops for educators.⁴¹¹ More work needs to be done for these skills to be embedded into the fabric of our society.

⁴¹¹ Full Fact, "Full Fact Report 2024: Trust and truth in the age of AI", April 2024, https://fullfact.org/media/uploads/ff2024/18042024-full_fact_report_corrected.pdf#page=52

Chapter 11: Building a better information environment

Introduction

We firmly believe in the importance of using trusted, non-partisan fact checkers to promote accuracy in public debate. We also use our fact checks and articles on subjects including data literacy, AI and legislative proposals to highlight issues we have with the internet platforms that dominate our information environment. Our aim is to safeguard their users and to ensure the right information reaches the people who need it most.



Our interventions strategy is focused on improving public access to trustworthy, evidence-based information. Beyond fact checking we:

- Urge politicians and the media to publicly correct the record when they share false or misleading information. We aim to intervene as early as possible in the information cycle to reduce the harm bad information can cause.

- Campaign for systemic change—at both the platform and legislative level—to make good information the default for every internet user.
- Prebunk viral falsehoods by exposing common misinformation techniques, and inoculating users by sharing reliable information before they encounter false claims.⁴¹² This includes publishing explainers during emerging major news events, and proactively addressing potential misinformation tactics in advance.
- Work with technology companies to push for more robust systems that identify and limit misinformation, while promoting accurate content to better protect users.

After last year's general election, we launched our Government Tracker—a tool that monitors how well the government is delivering on its promises, so voters can judge what progress has been made.⁴¹³ It currently tracks more than 50 government pledges and priorities on a range of subjects with plans to expand to around 100 of these commitments during the second half of 2025. The Tracker allows the public, policy makers and researchers to access informed, evidence-based facts about the government's progress and serves as a useful research tool for public servants, the media and academics. The Prime Minister has spoken several times about the importance of having clear measurable targets “so every single person in the country can judge our performance on action, not words.”⁴¹⁴ We agree, and will hold him to that.

This chapter looks at how Full Fact has contributed to building a better information environment to restore trust over the last year. It highlights our successful efforts to secure corrections in parliament and the media, our push for higher standards in public life, and the areas where we still see room for improvement. It also explains in detail our biggest single intervention over the last few years—how we are using our world-leading AI tools to tackle misinformation at scale, because we know an internet-sized problem needs internet-sized solutions.

Full Fact's year in review: 12 months of interventions with impact

Over the past year, Full Fact has achieved several important changes through targeted interventions. These include corrections to health research and public data, as well as inaccurate claims made by politicians (including ministers and shadow ministers). We also prompted several prominent media organisations to correct their reporting. In total, we have challenged 119 claims, with 61 resulting in corrections, leading to tangible improvements to the quality of public information.

⁴¹² Sheera Frenkel, “Debunking misinformation failed. Welcome to ‘pre-bunking’”, The Washington Post, 26 May 2024, <https://www.washingtonpost.com/technology/2024/05/26/us-election-misinformation-prebunking/>

⁴¹³ Full Fact, “Government Tracker”, accessed 21 April 2025, <https://fullfact.org/government-tracker/>

⁴¹⁴ Keir Starmer, “Keir Starmer speech at Labour Party Conference 2024”, 24 September 2024, <https://labour.org.uk/updates/press-releases/keir-starmer-speech-at-labour-party-conference-2024/>

On what basis do we intervene?

We prioritise false, misleading and unevidenced claims that:

(i) are of significant public interest,

(ii) have the potential to cause harm to people's lives,

(iii) are at risk of being repeated.

We also consider whether intervening will ultimately help to improve the information environment, for example by guiding a prominent person or institution to share more reliable information in future.

Here are a few of the highlights:

- In May 2024, then-health secretary Victoria Atkins claimed in Parliament that 758,000 children and young people in England were seen by NHS-funded mental health services in the 12-month period to March 2021.⁴¹⁵ However, NHS England data shows that about 573,000 children and young people received NHS mental health services in this period.⁴¹⁶ Ms Atkins corrected the record following our intervention, helping to ensure that Parliament's debate of the Cass Review was not misinformed.
- During the 2024 general election campaign, Full Fact secured a correction from the Green Party of a claim in its manifesto about the NHS leaving "nearly 8 million of us on hospital waiting lists".⁴¹⁷ The correct figure at time of writing was about 6.3 million people. As we set out in Chapter 5, the mistake was based on a common misunderstanding of NHS data. Our correction helped improve the accuracy of information about the NHS—a key election issue—in the run up to polling day.
- In March 2025, industry minister Sarah Jones MP became the first minister in the current government to correct the record in Hansard after being contacted by Full Fact about a misleading or inaccurate claim. She had claimed that the International Monetary Fund (IMF) and the Organisation for Economic Co-operation and Development (OECD) predicted that the UK will be Europe's fastest-growing economy over the next few years.

⁴¹⁵ Leo Benedictus, "A ministerial mistake is on the Parliamentary record—but was it due to an editing mix-up?", 19 April 2024, <https://fullfact.org/health/atkins-hansard-transcript-children-mental-health/>

⁴¹⁶ Ibid.

⁴¹⁷ Full Fact, "Green Party corrects manifesto after Full Fact intervention", 20 June 2024, <https://fullfact.org/live/2024/jun/green-party-corrects-manifesto-after-full-fact-intervention/>

In fact, the latest figures from the IMF and OECD at the time projected that a number of European countries will have higher growth than the UK in 2025 and 2026. Following our intervention, the corrected transcript read: “The International Monetary Fund and the OECD predict that the UK will be Europe’s fastest-growing major G7 economy in the coming years.”⁴¹⁸

- Beyond the UK, our fact checking work is also instrumental in ensuring breaking news during major global events is reported responsibly and accurately. After the ceasefire deal between Israel and Hamas in January 2025, the Times of Israel published a photo captioned “Palestinians celebrate ceasefire-prisoner release deal, January 15, 2025”, suggesting the image showed Palestinians celebrating in the streets after the ceasefire announcement.⁴¹⁹

However, the photo actually showed Palestinians celebrating a ceasefire ending 11 days of conflict between Hamas and Israel in May 2021, not January 2025.⁴²⁰ The photo was changed after Full Fact contacted the newspaper about the error, ensuring that no more people were misled by it.

Beyond fact checking: tackling misinformation at a systemic level

Over the course of the last year we have achieved a number of ‘intervention successes’. These include:

- **Implementation of a new Parliamentary corrections system, making it easier for all MPs to correct the record.** The new system, which Full Fact and thousands of supporters were instrumental in establishing, is finally in place and MPs are using it.⁴²¹ In January 2025, the Liberal Democrat health and social care spokesperson Helen Morgan became the first MP to use the new corrections system in response to a correction request from us.⁴²²
- **Ensuring swift action by NHS England on potentially harmful health misinformation,** with support from the Office for Statistics Regulation (OSR). After working with NHS England to identify ways that it could respond more swiftly and impactfully to our intervention, it has since corrected claims in an

⁴¹⁸ Full Fact, “Minister corrects parliamentary record after Full Fact intervention”, 5 March 2025, <https://fullfact.org/live/2025/mar/sarah-jones-hansard-correction/>

⁴¹⁹ Evie Townend, “Old photo of Palestinians celebrating shared as recent”, 21 January 2025, <https://fullfact.org/online/palestinians-celebrating-ceasefire-old-photo-2021/>

⁴²⁰ Ibid.

⁴²¹ Azzurra Moores, “Parliamentary corrections system overhaul: Speaker responds to Full Fact’s campaign”, Full Fact, 17 April 2024, <https://fullfact.org/blog/2024/apr/new-parliamentary-correction-system-speaker-announces-changes-after-long-standing-calls-from-full-fact/>

⁴²² Full Fact, “Liberal Democrat spokesperson uses improved parliamentary corrections system to update claim about NHS waiting lists”, 9 January 2025, <https://fullfact.org/live/2025/jan/spokesperson-corrects-error-using-improved-corrections-system/>

interactive report on waiting lists,⁴²³ and helped us to secure agreement from the Department of Health and Social Care about how ministers should accurately describe waiting list figures in the future.⁴²⁴

- **Improving understanding about rights to speak out about misinformation.** In late 2024, the Independent Press Standards Organisation (IPSO), a self-regulator paid for by publishers, shared its decision to uphold Full Fact's complaint against the Daily Express.⁴²⁵ We subsequently met with IPSO to share our reflections on how the complaint process could be improved in order to combat misinformation sooner. IPSO agreed to make it clearer to complainants that they can communicate openly about complaints at each stage of the process.
- **Working to improve 'intelligent transparency' around statistics in the new government.** After our fact checking revealed that Sir Keir Starmer used unpublished data about UK immigration returns in his first Labour conference speech as Prime Minister,⁴²⁶ we raised the case with the OSR and secured release of the data by the Home Office. In response, the OSR took a range of actions to improve Whitehall's adherence to the Code of Practice for Statistics,⁴²⁷ including emphasising the importance of publishing data used in the public domain.

Despite these successes, it has proved challenging for us to secure corrections from government ministers, other MPs and political parties. For example, while industry minister Sarah Jones corrected a claim relating to the government's mission on growth (as outlined above) four other government ministers have so far failed to correct similar claims after receiving our requests.

We maintain a strong commitment to the importance of transparency in public life and will continue to call for corrections from politicians and political parties as we identify them, and seek media attention to help amplify this lack of candour when these corrections are not forthcoming.

⁴²³ Leo Benedictus, "NHS England report on waiting lists confuses patients and cases", Full Fact, 15 October 2024, <https://fullfact.org/health/nhs-england-patients-pathways-report/>

⁴²⁴ Leo Benedictus, "Wes Streeting overstated the number of people on the NHS waiting list under the Conservatives", Full Fact, 12 March 2025, <https://fullfact.org/health/streeting-powell-people-cases/>

⁴²⁵ Kevin Armstrong, "Full Fact v Daily Express: how and why it happened", Full Fact, 6 September 2024, <https://fullfact.org/blog/2024/sep/full-fact-v-daily-express/>

⁴²⁶ Hannah Smith, "PM's conference speech claim about '23% increase' in immigration returns based on unpublished data", Full Fact, 26 September 2024, <https://fullfact.org/immigration/starmer-conference-speech-unpublished-data/>

⁴²⁷ Ed Humpherson, "Embedding the habit of intelligent transparency", Office for Statistics Regulation, 14 October 2024, <https://osr.statisticsauthority.gov.uk/blog/embedding-the-habit-of-intelligent-transparency/>

Harnessing technology improvements to prevent the spread of misinformation

Our fact checks reach audiences in numerous ways, through our website, newsletter and social media accounts, across news media and via interactive voice assistants, such as Google Assistant or Amazon's Alexa.

An October 2024 investigation by Full Fact found that Amazon's interactive voice assistant, Alexa, was giving users incorrect information on topics ranging from MPs' expenses to the origins of the Northern Lights, apparently repeating false and misleading claims that have been the subject of Full Fact fact checks. Concerningly, Alexa cited Full Fact as the source of the incorrect answers it was giving, because it had drawn them from articles we had published.⁴²⁸ We then found that Alexa was also giving incorrect information attributed to other fact checkers.

Amazon acknowledged errors that we flagged with them, and appears to have stopped them reoccurring since. Amazon also told us that it was working to resolve any similar issues that might exist.⁴²⁹

After a similar issue with Google Assistant, which surfaced a misleading response when we asked specifically for information from Full Fact,⁴³⁰ we raised our concerns with the Department for Science, Innovation and Technology (DSIT) about the accuracy of virtual assistants.

While we can't say for certain why Alexa and Google Assistant were making these errors, it's clear that voice assistants face similar challenges to AI chatbots as they struggle to interpret content accurately, and often fail to distinguish between false claims and fact checks meant to debunk them.

In today's online environment, using high-quality data to shape what users see is more important than ever. At Full Fact, we believe fact checkers must focus on the real-world harm false or misleading claims can cause—and make that impact clear. This helps platforms and search engines make better decisions about how to present fact checks. For example, in cases involving especially harmful topics, AI-generated responses should be replaced with reliable, impartial information from trusted fact checkers.

⁴²⁸ Sarah Turnnidge, "Amazon Alexa users given false information attributed to Full Fact's fact checks", Full Fact, 17 October 2024, <https://fullfact.org/online/amazon-echo-misleading-voice-assistant/>

⁴²⁹ Sian Bayley, "Amazon's Alexa has been giving more incorrect answers attributed to fact checkers", Full Fact, 28 October 2024, <https://fullfact.org/online/amazon-alexa-misleading-voice-assistant-more-answers/>

⁴³⁰ Sarah Turnnidge, "Amazon Alexa users given false information attributed to Full Fact's fact checks", Full Fact, 17 October 2024, <https://fullfact.org/online/amazon-echo-misleading-voice-assistant/>

Building Full Fact AI to scale our fact checking interventions

Perhaps the biggest challenge we face is one of scale. Effective monitoring of the vast amount of information which appears on the internet every day requires the use of technology to do things that humans can't do alone. To supercharge our expertise, widen access to expert fact checking, and increase the footprint of our fact checks by identifying more examples of relevant false claims to correct, we develop world-leading AI tools.

Full Fact AI is a suite of fact checking tools which has been used by more than 50 organisations in 40 countries. It is currently available in English, Arabic and French. We view the relationship between the technology that we develop and human fact checkers as a co-intelligence. Full Fact AI does not exist to replace human expertise. It's there to help:

- Identify the most important claim to be fact checking that day
- Know when someone repeats something they already know to be false
- Check things in as close to real time as possible
- Monitor public debate at scale, and allow experts to focus on things that could cause the most harm

Our AI-powered tools are built for fact checkers and organisations committed to promoting accurate information. They help process large volumes of content efficiently, allowing fact checkers to focus on the most important claims.⁴³¹ With advanced search capabilities, users can track claims by speaker, political party, topic, or type—across sources like RSS feeds, newspapers, YouTube, podcasts, social media, and even radio—ensuring they stay on top of the conversations that matter most.

Full Fact AI is scalable, robust software that saves time, money and effort in identifying the most important bad information to address, and is uniquely positioned to enable small groups of people to tackle misinformation at the scale of the internet. It utilises cutting-edge natural language processing (NLP) and machine learning algorithms to scan vast amounts of information, identifying potential falsehoods and prioritising the most important claims to verify.

Full Fact AI helps users to stay ahead of false and misleading claims circulating in the media. With real-time claim labeling and detection, fact checkers can quickly spot emerging falsehoods, track their spread, and take action before they gain traction.

⁴³¹ Full Fact, "Full Fact AI", accessed 21 April 2025, <https://fullfact.org/ai/about/>

Our tools help to track how misinformation spreads, identify repeated falsehoods, and understand patterns of deception. Daily insights help fact checkers to take proactive steps to challenge false claims, limit their impact and harm, and keep their audience better informed. Without such efforts, we are in danger of reaching a place where no one believes anything anymore. That would be bad for our democracy, and for our understanding of what it means to be a functioning modern society.

Interventions: conclusion and ratings

Our mission to help build a better information environment—more reliable and less harmful—has some distance to travel. Transparency in public life and a willingness to set the record straight are essential to the health of democracy and to restoring trust in politics. We have witnessed some notable corrections over the past year, from public bodies, politicians and the media, and it's clear that this work is progressing. But there's more to be done, especially by ministers, to ensure the public are fully equipped to make informed decisions in all circumstances.

The government has taken initial steps in response to calls from a wide cross-section of experts for a greater focus on media literacy. Currently, however, their focus is a little too narrowly on young people, important though that is. Media literacy is a skill that everyone needs, regardless of age or socio-economic background. A failure to expand the scope of this work risks leaving parts of society behind.

On all these fronts, we believe technology is part of the solution. The potential for us to harness generative AI, to further develop the tools we have created to handle information at internet scale, and to collaborate with academics and other partners to unlock future opportunities, is one of the most exciting aspects of the work we do.

So how hopeful are we feeling? This report has painted a picture of a misinformation landscape in urgent need of attention. But there is still time to fix it. Platforms still have the chance to roll back decisions that could end years of productive relationships and take greater responsibility for good information.

Across government there is no disagreement that misinformation is a problem that needs addressing for the health of our wider society. Ministers and regulators also have the opportunity to make substantial changes that could protect the UK population, and allow people to make informed choices on the issues that matter to them.

Ratings

State of interventions: positive signs but more action required

Government's handling of media literacy:
need to build on progress



Response to corrections requests:
politicians should take more responsibility



Use of technology to combat misinformation:
proven potential but wider takeup needed



Full Fact

 fullfact.org/contact

 [@FullFact](https://twitter.com/FullFact)

 fullfact.org

Published by Full Fact, May 2025. Published under the Creative Commons Attribution-ShareAlike 4.0 International License.

A registered charity (no. 1158683) and a non-profit company (no. 6975984) Limited by guarantee and registered in England and Wales