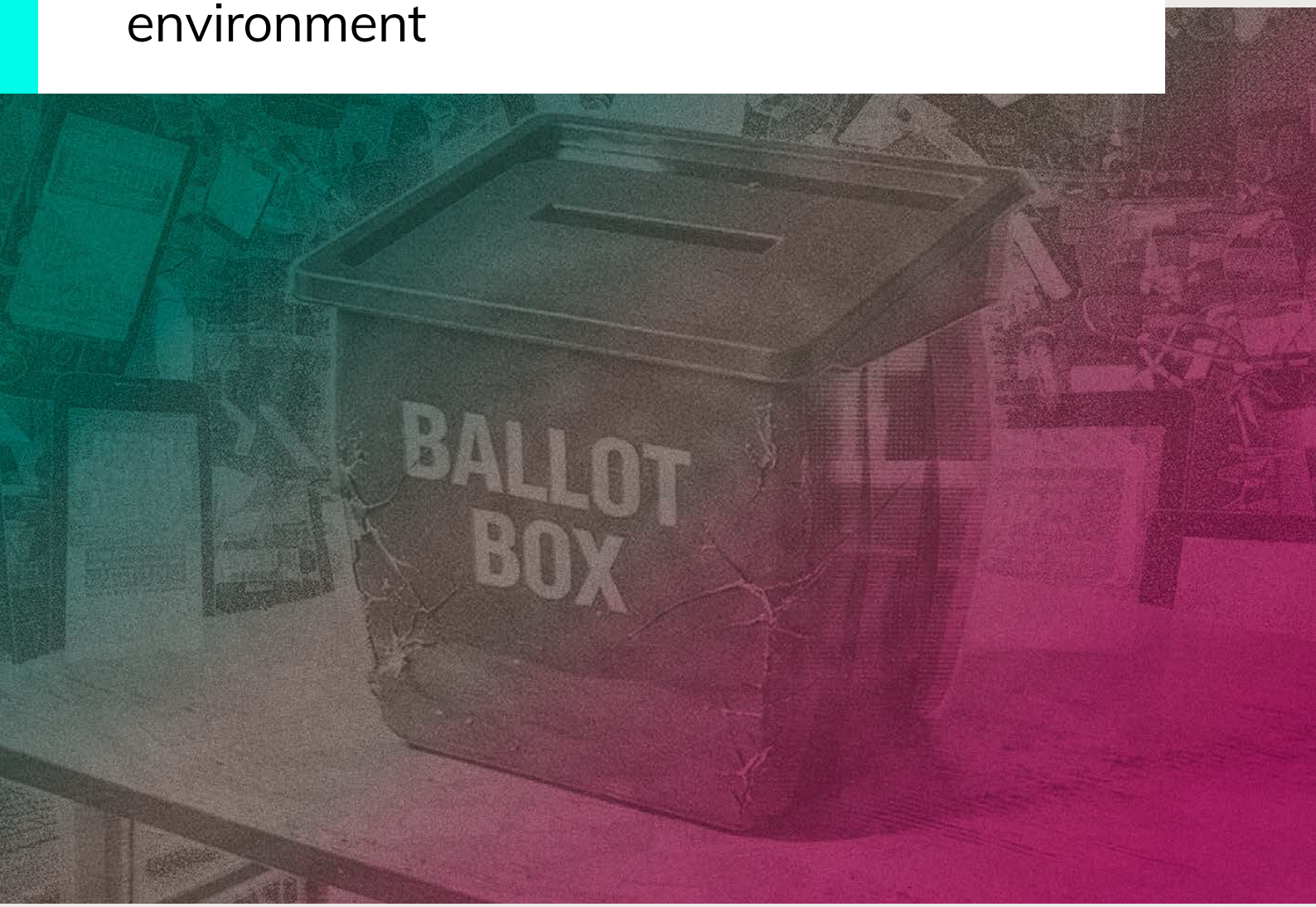
JUNE 2026

# Full Fact Report 2026

## A system under strain: strengthening the UK's democratic information environment




FULL FACT

# Full Fact Report 2026

## A system under strain: strengthening the UK's democratic information environment

# About this report

Full Fact wants to build a better information environment to restore trust. We fight bad information—and we promote good information that allows people to make informed choices. Our team of independent fact checkers, technologists, researchers and policy specialists tackle the harm caused by misinformation in the following ways:

- We fact check claims in public debate which are of public interest.
- We ask people to correct the record where possible.
- We build world-leading AI tools which allow small teams to work at internet scale.
- We campaign for system changes to make bad information rarer and less harmful.
- We advocate for high standards in public debate.

This report assesses the precarious state of the UK's information environment, focusing on growing uncertainty about what to trust and the serious impact this is having on democracy. It examines how political and technological change is reshaping the way information is produced, distributed and amplified. This includes the growing influence of AI-mediated systems, platform design, and incentives that shape information flows. It also considers the implications of these defining changes for the government, regulators and the public in sustaining a shared understanding of reality and trust in democratic processes.

It follows on from our 2025 report, *Restoring trust in a fractured information environment*, our 2024 report, *Truth and trust in the age of AI*, and our 2023 report, *Informed citizens: Addressing bad information in a healthy democracy*. This is the seventh annual report that we have produced with 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 has funded this report but the views expressed are those of Full Fact. The contents of this report are the responsibility of Full Fact's Chief Executive. They do not necessarily reflect the broad spectrum of views held by Full Fact's Board of Trustees.

# Contents

# Foreword

The UK's information environment is becoming harder to trust, harder to navigate and easier to manipulate. This is no longer a distant or theoretical concern. It is already reshaping public debate and confidence in our democratic institutions. The systems people rely on to understand the world—search engines, social platforms, AI tools and recommendation systems—are changing at extraordinary speed, often without meaningful transparency or accountability. At the same time, many of the institutions responsible for producing and verifying trusted information are under sustained political, financial and technological pressure.

As Full Fact has repeatedly argued, democracies depend on an environment in which reliable information is both visible and verifiable. But this does not happen automatically. Good information that allows people to make informed choices has to be defended and strengthened, as it struggles to compete with systems that increasingly reward speed and outrage, and that promote uncertainty.

Right now, many of the signals people once relied on to judge credibility are weakening. Authoritative reporting and fact checking competes in the same spaces as conspiracy theories, synthetic content, manipulated media and politically motivated distortion. In some cases, confusion itself has become the strategy—not persuading people to believe one false claim, but creating enough uncertainty that trust in all information begins to break down.

This is the challenge of our age. Technological change is reshaping how information is created, distributed and consumed faster than governments, regulators or democratic institutions have been able to respond. At the same time, old political certainties that underpinned the post-Cold War democratic order are disappearing.

Since the return of President Trump, the global political environment has become more openly hostile to many of the institutions traditionally responsible for establishing shared facts and democratic norms, and those shifts impact the UK too. They shape the behaviour of political actors, media systems and online audiences; they empower the Big Tech platforms to ignore many of their responsibilities; and they are accelerating the sense that public reality itself is becoming contested territory.

We have to think hard about what this means for our democracy. As this report sets out, the next UK general election will take place in conditions unlike any previous electoral cycle. AI-powered search tools and chatbots are rapidly becoming mainstream entry points for political information. Increasingly, people are not searching for sources and weighing evidence themselves; they are receiving automated summaries, generated interpretations and synthesised answers.

Those systems shape not only what people see, but how issues are framed and understood. Meanwhile, synthetic media is becoming cheaper, faster and more convincing just as major platforms have reduced moderation capacity, weakened verification systems and stepped back from earlier commitments around trust and safety.

These are not isolated developments. Taken together, they amount to a structural shift in the information environment itself. Information now moves ever more rapidly through systems optimised primarily for engagement rather than accuracy. Reliable information still exists, but it is competing in environments designed to maximise attention, emotional reaction and frictionless sharing. The result is a growing gap between the availability of trustworthy information and the public's ability to recognise and act on it.

That matters because democratic societies rely on some shared ability to establish facts, resolve uncertainty and make decisions based on evidence. When people no longer know what information can be trusted, public debate fragments and political disengagement grows. Elections become more vulnerable to manipulation, and public crises become harder to manage. Trust, once lost, becomes far harder to rebuild.

The UK is not currently well prepared for this reality. Responsibility for electoral integrity, online safety, platform accountability and crisis communication is fragmented across multiple institutions with overlapping or unclear mandates. Although wider crisis management frameworks have been strengthened, there remains no clear system for resolving information uncertainty during high-pressure democratic moments, and no single institution responsible for ensuring that reliable public-interest information remains visible and trusted when it matters most.

This report argues that the UK needs to treat the information environment as critical democratic infrastructure. That means moving beyond piecemeal responses and recognising the scale of the structural change underway. Public resilience cannot depend entirely on individuals trying to navigate increasingly complex and automated information systems by themselves.

A stronger democratic information environment requires systemic reform: clearer institutional coordination, stronger accountability for platforms and AI systems, better visibility for high-quality public-interest information, and long-term investment in media literacy and public resilience. It also requires recognising that the design of platforms and AI systems is not neutral. Decisions about ranking, recommendation, moderation and visibility shape public understanding at enormous scale. Those decisions carry democratic consequences.

There is no route back to an earlier information environment, or a simpler time. The task now is to build institutions and safeguards capable of functioning in a faster, more fragmented and more automated public sphere. Other democracies facing similar pressures are beginning to move towards more coordinated approaches built around transparency, preparedness and accountability. The UK risks falling behind if it continues to respond in a fragmented and reactive way.

The decisions taken now will shape whether future elections are defined by confusion and mistrust, or by clarity, confidence and democratic resilience. Building a better information environment is no longer an abstract ambition. It is a democratic necessity.

**Chris Morris**
**Chief Executive**

# Executive summary

Over the past year, facts—and the institutions that verify and communicate them—have come under intense and sustained pressure. Fact checking organisations, journalists and researchers face growing political, legal and economic challenges that undermine their ability to operate independently and maintain public trust. At the same time, changes to online platform features and the rapid integration of AI-mediated systems have reduced the visibility of authoritative information, and weakened the signals people rely on to judge what is accurate. These pressures risk reducing both the supply of reliable information and the public's ability to recognise false and misleading information. The consequence of which can only lead to a growth in real-world harms.

Information in the UK circulates through complex and increasingly automated systems involving media organisations, online platforms, recommender systems, search engines, AI tools, political actors, influencers and closed networks. These operate alongside public institutions, including those responsible for electoral security, online safety and public communication. Collectively, these actors and systems form the UK's information environment, and shape how people encounter, interpret and act on information across public life.

While this report focuses on the threat of misinformation and disinformation to UK democracy, the dynamics it examines extend beyond politics. The same systems and incentives shape access to information about health, the economy, and other areas where reliable information is essential. False and misleading information in these domains can have consequences for individual wellbeing, social cohesion, and public trust in institutions and information itself. Elections and public crises are concentrated periods of stress in which these dynamics are acutely visible, exposing underlying weaknesses.

This report assesses a growing democratic risk: persistent uncertainty about what information can be trusted, verified and acted upon. This can arise from incomplete or contested facts, delayed or fragmented institutional response, and declining confidence in institutions and sources of clarification. It is compounded by structural weaknesses across technological, institutional and regulatory systems, and by the absence of a clear, visible system for resolving uncertainty during periods of democratic stress.

**Part 1** examines how uncertainty emerges and misinformation proliferates during UK elections and public crises. It identifies recurring patterns and shows how platform design, AI-mediated systems, monetised incentives and declining trust interact to amplify false and misleading information. These dynamics weaken authoritative signals and allow uncertainty to persist even where reliable information is available.

**Part 2** assesses the resulting democratic harms. These include the erosion of a shared factual understanding of reality, increased confusion, disengagement and withdrawal from political participation, heightened community tensions, and declining trust in institutions and information sources. It also considers how these risks are likely to evolve and why the UK's information environment is likely to be different in the next general election.

As we approach that election, AI-powered search tools, chatbots and automated summaries are becoming routine entry points for political information, shaping what people encounter and how issues are framed. Synthetic content can be generated and disseminated rapidly and strategically, while platform safeguards are becoming more variable, and some moderation capacity is being reduced. These changes represent a structural shift in how information is produced, distributed and consumed.

**Part 3** places the UK's experience in an international context. Other democracies face similar pressures and have not resolved them. However, international evidence suggests that resilience is stronger where systems are designed holistically, with clear responsibilities, transparency, access to data, platform accountability, and visible institutional communication during elections and crises. These examples provide practical lessons for strengthening the UK's capacity to respond to information-related risks.

**Part 4** evaluates whether the UK's laws and institutions are equipped for the contemporary information environment, particularly under the conditions of an election or a crisis. It finds that responsibilities are fragmented across multiple bodies, coordination mechanisms and escalation pathways are opaque, and existing legislation lags behind the pace and complexity of emerging information threats. No single institution is responsible for ensuring system-wide coherence across the UK's democratic information environment, nor for providing a visible, accountable centre to coordinate responses to information incidents.

**Part 5** sets out recommendations for reform. While targeted intervention can address specific problems, they will not keep pace with the fast-moving, interconnected risks in the modern information environment. Strengthening the UK's capacity to manage information-related risks will require coordinated reform across multiple domains. This report makes a series of recommendations under four mutually reinforcing pillars.

We cannot just assume the UK's democratic information environment will remain stable under intensifying strain. Uncertainty is becoming structural rather than episodic. In the next general election, information systems are likely to be faster, more automated, and less predictable than those of the past. Meeting this challenge will require stronger legal and institutional frameworks, clearer accountability for platforms and AI systems, and sustained investment in high-quality public-interest information. The decisions taken now will shape whether future elections are defined by confusion and mistrust or by clarity, confidence and democratic resilience.

# Recommendations

This report sets out a programme of reform to strengthen the UK's democratic information environment. It brings together a set of practical, system-level recommendations designed to reduce uncertainty, improve the visibility of reliable information, and ensure that public authorities, online service providers and the public are better prepared for high-pressure situations, such as elections and crises. The recommendations are organised under four mutually reinforcing pillars, summarised below and set out in full in Part 5.

## 1. Secure the information ecosystem

Ensure that the information environment supports the rapid circulation and visibility of accurate, reliable information—particularly during elections and crises—so that uncertainty is resolved before it takes hold and false or misleading claims become embedded.

- **Stress-test information resilience**
  Establish a programme to test how information systems and institutions perform during elections, crises and other high-risk periods. Publish key findings and use the results to strengthen crisis protocols, coordination and communication.

- **Maintain crisis communication plans and incident protocols**
  Require major platforms, search engines and AI systems to maintain robust crisis communication plans and incident management protocols, that are aligned with national arrangements and subject to oversight, audit and testing.

- **Increase the prominence of high-quality information**
  Require large online service providers to implement proportionate measures during defined high-risk periods to increase the visibility of high-quality public interest information. This should be underpinned by clear criteria and triggered by transparent thresholds, and should focus on preventing harm.

## 2. Strengthen public resilience

Support people to navigate the increasingly complex and uncertain information environment with confidence, while recognising the limits of individual responsibility. Public resilience depends on the development of critical thinking skills, alongside systems that make reliable information easier to find and assess.

- **Embed media literacy across the curriculum**
  Support the integration of media and information literacy across the curriculum at all stages, including an understanding of AI-mediated information, with the provision of teacher training, guidance and high-quality teaching resources.

- **Fund long-term media literacy delivery capacity**
  Provide long-term, ring-fenced funding to support the delivery of media literacy, including workforce capacity, teacher training and professional development, with mechanisms that enable sustainable delivery at scale.

- **Introduce a statutory duty to provide media literacy**
  Introduce a statutory duty on large online service providers to provide effective, evidence-based media literacy measures, embedded in product and system design, with clear standards and a code of practice, overseen by Ofcom.

## 3. Modernise laws and institutions

Update legal frameworks and institutional arrangements so they are clear, coordinated and able to operate at speed. Information resilience requires effective, visible coordination, defined responsibilities and mechanisms that respond coherently under pressure.

- **Strengthen the Representation of the People Bill**
  Expand the Bill to address misinformation and disinformation risks, including transparency for political advertising, rules on deepfakes, regulation of political advertising, and enhanced powers for the Electoral Commission.

- **Establish a national information incident response framework**
  Build a dedicated information incident response framework within the UK's broader crisis management and resilience architecture, with pre-defined escalation thresholds, coordination protocols, and public communication procedures.

- **Create a national Information Resilience Unit**
  Establish a statutory body to coordinate preparedness and response to threats to the UK's information environment, maintain the incident response framework, convene stakeholders, and lead stress-testing and post-incident reviews.

## 4. Increase commercial transparency and accountability

Ensure that platforms, search engines and AI systems are transparent and accountable for how they shape the information environment. This includes focusing on system-level behaviour and impacts—how information is ranked, amplified and presented.

- **Introduce systemic risk management duties**
  Require large platforms, search engines and AI systems to assess and mitigate key systemic risks they pose, with a focus on potential harms. This would reflect existing duties on large platforms and search engines in the EU.

- **Implement provenance and labelling standards**
  Require large platforms, search engines and AI systems to use interoperable, consistent labelling of AI-generated content, in order to support attribution and reduce uncertainty, particularly during high-risk periods.

- **Require AI provider output transparency**
  Require AI systems to clearly communicate sources, uncertainty and limitations in their outputs, particularly for high-impact domains such as health, politics and finance. Requirements should operate across interfaces and use cases.

- **Enable enhanced access to platform data**
  Use the Data (Use and Access) Act to establish a robust framework for secure, timely and scalable access to platform data for independent researchers, including accredited fact checkers, to support evidence-based policy interventions.

### Full Fact polling reveals how misinformation damages democracy

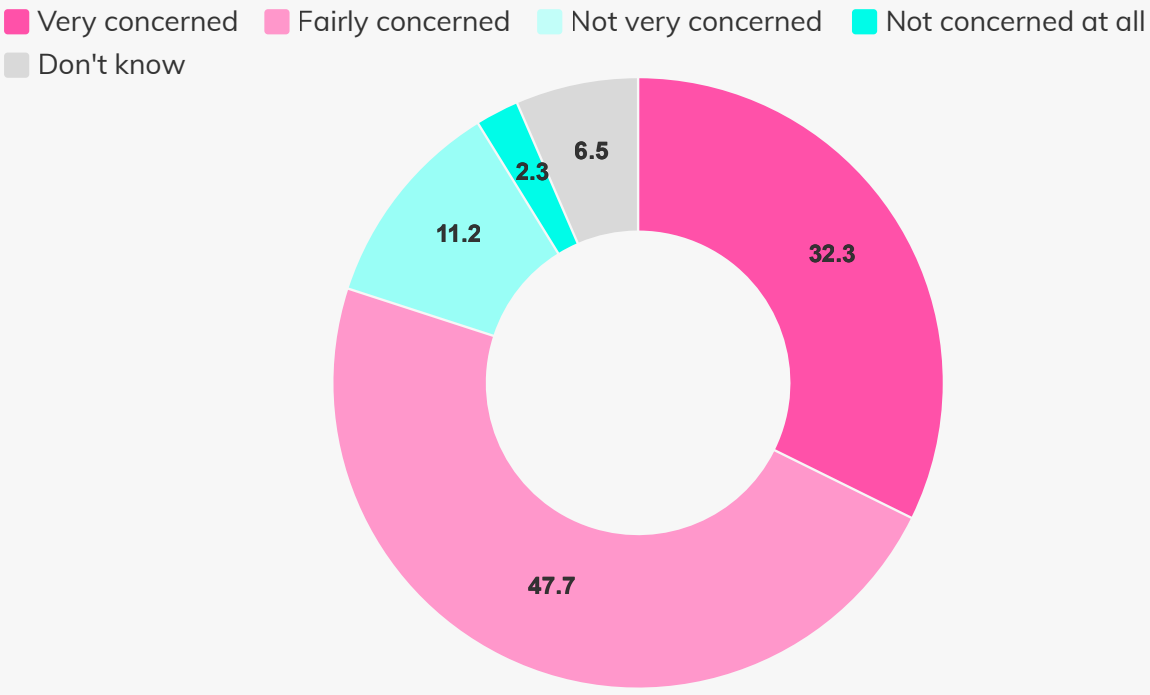
A nationally representative poll of UK adults—commissioned by Full Fact and carried out by YouGov at the end of March 2026—has revealed a public whose serious concern about misinformation is shaping how they engage with democracy.[1]

The poll found that 80% of people are concerned about political misinformation. Only 3% find it very easy to tell whether a video is genuine or AI-generated, and the same proportion feel very confident in their ability to distinguish between genuine and fake political information online. At the same time, 66% of people think the government is doing too little to address AI-generated misinformation, and only 9% think it is doing the right amount.

All institutions that people were asked about, apart from government department websites, have a net negative level of trust when it comes to providing accurate and reliable information. Political parties and social media platforms were least trusted, with distrust exceeding trust by 77% in the former (86% vs 9%) and 80% in the latter (88% vs 8%). People trust AI tools (20%) to provide accurate and reliable information more than twice as much as they trust political parties (9%), although the levels of trust for both are low.

Public concern is impacting democracy. Of the four in five people (80%) who expressed concern about political misinformation, many say it has negatively affected their democratic behaviour over the past year: 48% say it has affected their trust in political institutions; 42% say it has affected their confidence that elections are free and fair; and 27% say it has affected their likelihood of voting in an election.

**Generally speaking, how concerned, if at all, are you about misinformation related to politics?**



# Part 1: The state of the UK's information environment

Political information in the UK circulates through systems optimised for speed, engagement and personalisation, rather than for accuracy or context. Misleading or low-quality claims can spread quickly, while authoritative information often struggles to surface at comparable scale. Where false or misleading information is corrected, these interventions rarely match the volume, velocity and complexity of contemporary information flows.

These dynamics are particularly visible during elections and public crises, when timelines are compressed, verification becomes harder and institutions are under greater pressure. During these periods, public attention is heightened, the need for accurate information rises, and small delays in clarification can have disproportionate consequences. This section draws from our fact checking of UK elections and crises, and what we have learned from our AI tools, to examine how uncertainty emerges and persists under pressure.

## 1.1 What we see: fact checking from UK elections and crises

### 1.1.1 Observed patterns from fact checking

Over the course of 2025, Full Fact published more than 750 fact checks and related pieces of content. Our AI tools—which have been used by over 40 organisations in 30 countries[2]—have enhanced and scaled our media monitoring and fact checking capabilities. During the 2024 election 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. We produced approximately 217 verdicts on claims or repeated claims, and published over 150 pieces of website and video content.[3]

Our work reveals patterns in how false and misleading information arises and circulates, across topics and platforms. These patterns show how political information can be difficult to verify, attribute or contextualise, particularly during periods of high pressure.

**Misleading use of genuine information is more common than outright fabrication**
In the course of our fact checking, real data, statements and events were frequently presented selectively or without necessary context. Statistics were cited without baselines or timeframes, partial figures were treated as definitive, and correlations were framed as causal relationships. This can produce misleading claims that appear credible. For example, our fact checks about the 2026 Gorton and Denton by-election included

the misleading use of bar charts—a regular feature of election campaigns—as well as the selective use of data.[4] During elections and crises, when the capacity to verify is constrained and speed is critical, this type of distortion can be particularly difficult to identify and correct.

**Attribution-based claims are a persistent source of uncertainty**

In 2025 we fact checked a fabricated article presented as the work of the Home Secretary at the time, Yvette Cooper, and designed to mimic the style of the Guardian.[5] Claims like this rely on the perceived authority of recognised institutions or public figures to establish trustworthiness. It can give them early traction, while the steps required to verify them are often time-consuming or not accessible quickly. That can depend on access to primary sources or official confirmation from public bodies, which may not be available at speed.


**Misleading narratives, not just misleading facts, shape interpretation**

Individual claims are often embedded within, and are a driver of, broader narratives that organise facts, events and statements into frames of explanation. These narratives shape how people interpret the world around them, influencing the meaning attached to accurate information. This complicates binary true / false assessments and can slow correction, particularly when claims are presented in compressed, emotive or highly shareable formats—or where they become an established part of public or political discourse. For example, individual claims about specific arrests or police responses have been used to advance a broader narrative of 'two-tier policing'—the idea that authorities treat different groups unequally—which has played into a perception of bias in UK policing.[6]


**Misleading content originates from diverse and overlapping actors**

In May 2026, we fact checked claims on social media—with one post on X viewed more than 380,000 times—which falsely suggested a newly-elected Reform UK councillor does not exist.[7] False or misleading political information is produced and amplified by a wide range of actors, including politicians, media commentators, social media bots, influencer networks and anonymous accounts. Similar claims and narratives can emerge across multiple channels simultaneously, complicating attribution and making it harder to identify a single source. This increases the effort required to fact check and secure corrections, and reduces the likelihood that clarification will reach all affected audiences.

**Certain policy areas are frequently associated with misinformation**

Full Fact's work shows that claims in certain areas recur more often than others, including those relating to crime, immigration, and economic performance. These areas combine high public interest with complex evidence and fragmented data sources, and can rely on institutional interpretation. As a result, claims in these areas are more likely to hinge on partial data, selective framing, or contested attribution or analysis. We find

that some false or misleading claims persist and keep coming back, no matter what we do to try to address them. These feature particularly in high-salience policy topics, like false claims about the number of people on NHS waiting lists, unemployment figures and deportations.[8]

**There has been a surge in AI-assisted content and synthetic claims**

In November 2024, Full Fact suspected AI involvement in four of our published fact checks; in October 2025, this had risen to at least 27.[9] This reflects a wider pattern observed by other fact checking organisations across Europe.[10] This includes the use of AI to falsely depict real events, such as the fake image purporting to show the funeral of children killed in a missile strike in Iran, which was shared widely online.[11] These are cases which fact checkers choose to write about and do not necessarily reflect the prevalence of AI-generated content; but they illustrate how AI tools can increase the speed, plausibility and repetition of false or misleading claims, raising the complexity of verification.

## Misleading political leaflets in the May 2026 elections

An investigation by Full Fact in April 2026 revealed how local leaflets were providing unreliable information about how to vote tactically in the May 2026 elections.[12]

We analysed leaflets from across England that were uploaded to Democracy Club's online archive in the first two weeks of April, and found a chart or graphic in more than 50.[13] At least 14 of these leaflets failed to provide reliable evidence to back up a specific claim about how people are likely to vote locally, or were unsourced or misleading in some other way. Among these examples, which came from all the major parties:

- At least four gave people national polling numbers rather than data directly relevant to their area.
- Three cited other data, such as Westminster constituency results or projections, that could not reliably support specific claims about the local area.
- Two gave no source, making it impossible for voters to assess reliability.
- Two from the same ward cited results from an older election, while ignoring those from a more recent by-election.
- One reported results from doorstep surveys.
- One used a mixture of less relevant polls.
- One quoted previous election results from a different ward, which in any event did not appear to support the claim made.

Several also displayed data in misleading or questionable ways, for instance by using bars that were not in proportion to the numbers they represented.

Good data on voting intentions is often not available in local elections. And there is nothing wrong with political parties using other data as part of a pitch to voters, as long as it is properly explained. Indeed, there have been big changes in the national polls in recent years that might need explaining.[14] But some of these leaflets could mislead people as they choose how to vote—for example, by claiming definitively that another party "*can't win here*", or that only one party can stop another, without evidence.

### 1.1.2 Elections as stress tests for democratic information systems

Election campaigns compress timelines, public attention and information flows into short, predictable periods, intensifying pressures on the information environment. Claims about candidates, policies or voting procedures can circulate widely before they are verified or contextualised, particularly when they align with existing narratives or emotional cues. This makes campaign periods a practical test of whether information systems can resolve uncertainty quickly and visibly enough to maintain public confidence.

Institutions responsible for electoral security, online safety, public communication and other areas need to operate together to respond effectively. However, they lag behind the speed at which information spreads, and their processes for coordination and escalation are often unclear. Recent efforts to strengthen preparedness—for example, the Electoral Commission's pilot use of a deepfake detection tool for the May 2026 elections[15]—reflect a recognition of the increasingly complex information environment and the need for a rapid response. This focus is welcome, and it is important for any public authority using this sort of technology to be transparent about how effectively it works in practice.

Platforms directly shape how the dynamics unfold. Many voters, particularly younger ones, encounter political information through algorithmically curated feeds, recommendations and summaries.[16] In March 2026, the Reuters Institute found that 39% of people aged 18-24 use social media as their main source of news—up from 21% in 2015. The same study found that 51% of people in that age group pay more attention to individual news creators than to traditional news brands (39%).[17] In 2025, Ofcom found that six in ten adults use an online intermediary (like social media, a search engine or news aggregator) for their news[18]—and that 82% of people in the UK aged 16-24 use social media for news in some form, not specifically as their main source of news.[19]

The visibility of information about politics and current affairs is therefore strongly influenced by platform design. Content that provokes strong reactions may gain early prominence, while more cautious or corrective information spreads more slowly or unevenly.[20] Unlike other media environments, such as smart TV manufacturers (which are required under the Media Act 2024 to give due prominence to public service broadcaster content alongside other online apps) platforms are not required to ensure the prominence of authoritative or high-quality information.[21] We should not be giving them a free pass.

Recent UK election experience illustrates that uncertainty can arise without sophisticated tactics or large-scale disinformation. During the October 2025 Caerphilly by-election, for example, false claims about the outcome circulated while voting was still under way, presented in a format that suggested they originated from Google.[22] Their timing and apparent credibility introduced confusion at a moment when clarity was essential. Speed, format and perceived authority all play a role in the public's understanding.

Election periods also expose how online narratives affect the safety of candidates. In 2025, a Parliamentary inquiry—the Speaker's Conference on the security of candidates, MPs and elections—highlighted the role of disinformation in driving abuse and intimidation, and the need for clearer expectations of platforms during elections.[23] In a welcome response to that inquiry, the government expressed its commitment to addressing the impact of misinformation and disinformation on candidates, MPs and UK elections.[24]

The increasing availability and use of AI-generated and synthetic content add to these pressures, by enabling the rapid production of plausible campaign material, and false and misleading content. Even at relatively low volumes, this content can impose significant verification and response demands within already challenging circumstances.[25]

### 1.1.3 Crises as stress tests for public communication under uncertainty

Public crises come in many forms, such as pandemics, terrorist attacks or riots. They compress timelines and intensify uncertainty, creating sudden information vacuums in which demand for explanation rises while reliable information may be unavailable, evolving or legally constrained. During crises, institutions involved in public communication—like the government, law enforcement and the media—need to provide timely, credible signals that reduce uncertainty before speculation takes hold. The challenge in stabilising the information environment is compounded by low levels of trust in institutions.

In 2025, the Cabinet Office refreshed the Amber Book, the UK government's crisis management doctrine.[26] The framework sets out arrangements for cross-government coordination, resilience planning, and crisis response. It recognises that crises are

characterised by uncertainty, heightened public scrutiny and the need for coherent national communication. However, while the framework provides a structure for operational crisis management, it is less explicit about how crisis systems should respond to challenges to the information environment, including misinformation and disinformation in elections, AI-generated content risks, and platform-amplified false narratives during crises.

Misinformation can spread when authoritative information is delayed or lacks reach. Analysis of the Covid-19 outbreak found that misinformation on prevention and the actions of authorities spread rapidly in the early stages, illustrating how quickly false narratives can take hold.[27] In high-attention cases like the pandemic, the ability to provide timely, proportionate and credible information is critical to reducing uncertainty and false narratives.[28] This underscores the importance of pre-established communication pathways to help reduce the scope for misinformation to influence perceptions.[29]

Recent incidents in the UK illustrate how false and misleading narratives play out in the absence of effective and authoritative correction. Following the 2024 Southport attack, false claims about the suspect's identity circulated widely online, amplified by a small number of high-reach accounts, before verified information was available.[30] Analysis of the incident highlighted the importance of timely and coherent public communication in preventing online speculation from contributing to offline disorder.[31] Similar patterns were observed during protests in Epping in 2025, where contested claims about police activity[32] shaped public interpretation of events before official clarification had visibility.[33]


Credit: AFP/Justin Tallis

In some cases, more proactive communication has sought to address these challenges. Following the Liverpool city centre parade crash in May 2025, Merseyside Police published details about the suspect's nationality and ethnicity earlier than was previous practice, seeking to pre-empt misinformation and its consequences.[34] This approach was later reflected in interim national guidance from the National Police Chiefs' Council, which recognised that disclosing information in the right circumstances may reduce speculation and community tension.[35] Analysis suggests that timely disclosure helped constrain the spread of false narratives in the immediate aftermath, while recognising that communication strategies alone cannot resolve systemic vulnerabilities.[36]


Credit: AFP/Paul Ellis

Legal constraints can add to the pressure on public communications. Contempt of court rules play a critical role in protecting the fairness of proceedings. But in some cases they can limit what public authorities feel able to disclose at moments of intense public speculation. After reviewing contempt laws in 2025, the Law Commission found that they *"have struggled to keep pace with the rise of online communications and social media"* and made a series of recommendations to improve the framework.[37]

As with election campaigns, emerging technologies add complexity. Generative AI enables the rapid production of plausible, authoritative but misleading content, increasing the volume of genuine-looking 'evidence' when verification is most constrained. This makes correction harder and increases risks, particularly during the early stages of a crisis.[38]

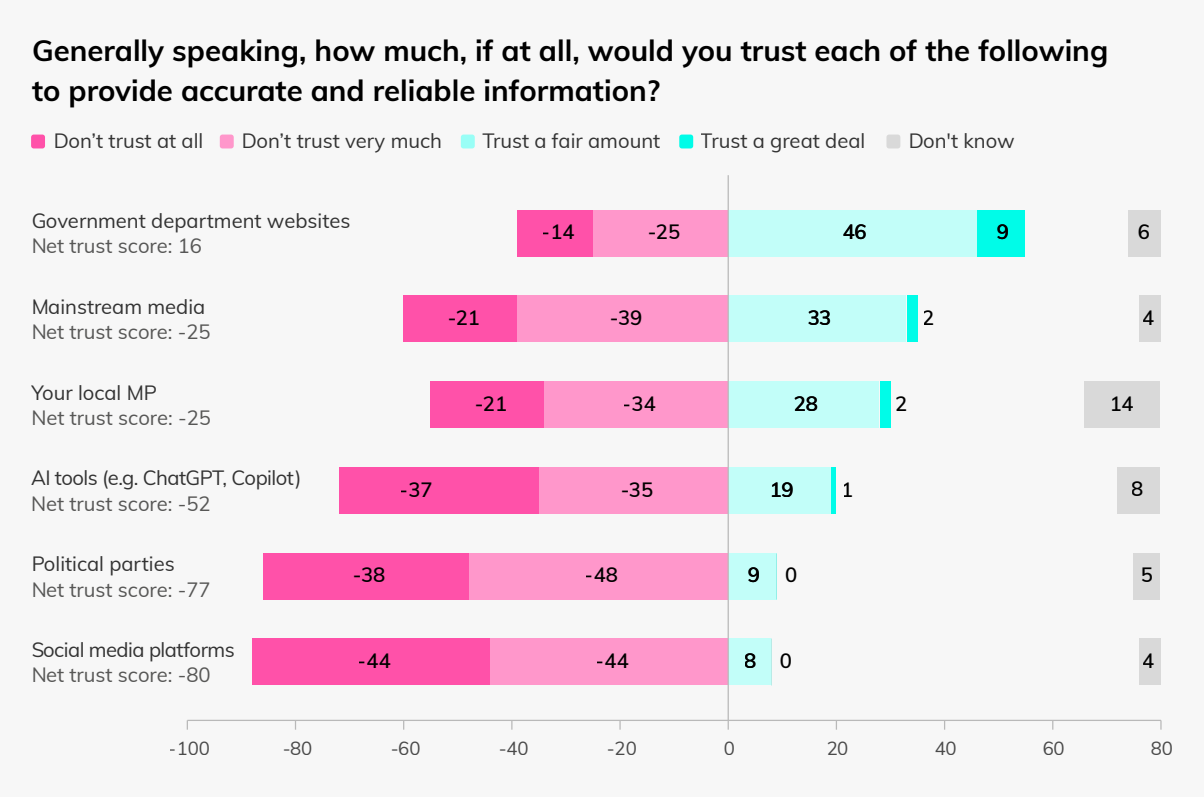
## 1.2 Why it's happening: structural drivers of misinformation

The patterns observed in our fact checking do not reflect isolated incidents; they are the result of structural conditions that shape how information is produced, distributed and understood. This section examines five factors that influence what information gains visibility, how quickly it spreads, and how easily it can be corrected:

- Low and fragmented trust as a driver of misinformation
- Platform systems, policies, and behaviour
- AI-mediated information flows
- Influencer ecosystems and monetisation incentives
- Public resilience and public interest information

### 1.2.1 Low and fragmented trust as a driver of misinformation

Full Fact's polling of UK adults points to the lack of a single actor that commands widespread public trust to provide accurate and reliable information.[39] Of the institutions people were asked about, government department websites are the only source of information with a net positive level of trust—and even this varies depending on how people voted in the 2024 general election. Trust in information from political parties is at rock bottom (0% of people trust them a great deal to provide accurate and reliable information) and on a par with social media platforms (also 0%).[40]

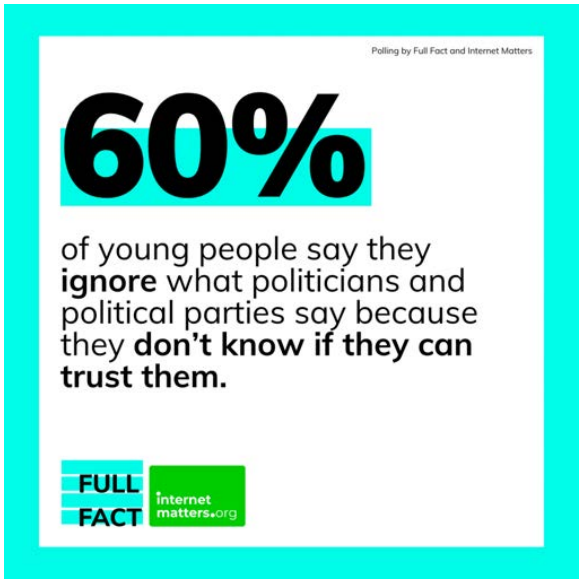


In a major national emergency, such as a pandemic or terrorist attack, certain institutions command some level of trust to provide accurate and reliable information. When asked to select up to three institutions which they would trust the most in a national emergency, 37% of people selected the police, 31% the mainstream media, 29% local authorities, and 24% ministers. Trust is much lower in social media platforms (7%), Ofcom (6%) and MPs (6%), with the least trusted being AI tools (3%). But for 17% of people, none of these institutions were in their top three most trusted during an emergency.

Where institutions are not widely trusted, authoritative information is less able to anchor public understanding with shared reference points, making it easier for misleading or competing claims to take hold and persist. At the same time, audiences are increasingly fragmented, reducing shared exposure to common sources of information and allowing competing interpretations to persist. Studies show that partisan media consumption and social media echo chambers reinforce pre-existing beliefs, and so competing claims persist side-by-side even when authoritative information is available.[41]

Trust in the UK's media is low. A survey by the Organisation for Economic Cooperation and Development (OECD) found that in 2021-2023, the proportion of people in Great Britain with low or no trust in the media increased from 48% to 65%, the highest among OECD countries.[42] Meanwhile, research from Ipsos in October 2025 found that over half of men (56%) and women (60%) aged 16-34 get the majority of their information about current events from social media; 79% are concerned about misinformation in the content they consume compared to 72% of the British public overall.[43]

The different ways in which younger people consume media, and their confidence in identifying misinformation, are particularly significant for democratic behaviour, given the government's plan to lower the voting age to 16. Polling by Full Fact and Internet Matters in November 2025 found that many younger people already approach political information with scepticism, with 60% of those aged 13-17 saying they ignore what politicians and political parties say because they don't know if they can trust them.[44]




Where confidence in public institutions, media, or platforms is weak, corrective information is more likely to be dismissed or interpreted sceptically. A large study in the Netherlands found that people with lower trust in government and health institutions were more likely to be sceptical about vaccines and susceptible to misinformation.[45] Research has also found that when news organisations issued retractions, people's beliefs became more accurate but their trust in the outlet declined.[46] This creates a damaging feedback loop: where trust in institutions is

weak, correction is less effective, allowing misinformation to persist; over time, contested or delayed correction may further weaken trust.

### 1.2.2 Platform systems, policies, and behaviour

Online platforms shape how information is distributed and encountered. Their ranking systems determine what content gains visibility, how quickly it spreads and how long it persists. Systems optimised for likes, shares and viewing time tend to amplify emotionally charged or divisive content.[47] Research modelling ranking systems has found that increasing an algorithm's emphasis on engagement boosts the visibility of polarised and misinformation-rich content, and that adding personalisation further reinforces exposure to like-minded, emotionally salient posts rather than diverse or contextualised information.[48] This helps explain why feeds can become echo chambers, where users repeatedly see similar narratives and authoritative content can struggle to break through.

These effects have implications for the functioning of democracy. The Electoral Commission has warned that *"algorithmic promotion of misleading content … risks undermining democratic participation"* and algorithms can rapidly amplify misleading content to users;[49] and that there is nothing in UK law to deal with platform algorithms being used *"in a partisan way to amplify or suppress political party posts to influence an election"*.[50]

Platform policies and resourcing decisions shape how these systems behave in practice. However, changes are often made with limited transparency, making it difficult to anticipate how platforms will behave. During the pandemic many platforms innovated to promote the spread of good information, and limit misleading claims—but since Donald Trump's re-election, there has been a shift away from active management of proven misinformation and promotion of high-quality sources. In the shadow of false accusations of censorship as well as executive orders from the US President,[51] many Silicon Valley companies have spent the last year or so systematically dismantling trust and safety measures.

TikTok reportedly cut its trust and safety team by hundreds of staff.[52] In November 2025, Parliament's Science, Innovation and Technology (SIT) Committee said TikTok did not provide the data or risk assessment underpinning its claim that reducing UK trust and safety staffing while increasing reliance on AI would improve moderation efficacy.[53] TikTok removed a Full Fact video in early 2026 due to a moderation error, which was later reinstated—illustrating the risk of automation at the expense of human moderation.[54]

Meanwhile, Meta replaced fact checking with community notes in the US, a feature where contributors can add context under a post. Meta described this system to the SIT Committee in March 2026 as *"a different form of fact-checking"*, a provocative claim which relies on a misleadingly expansive definition of this type of skilled, professional

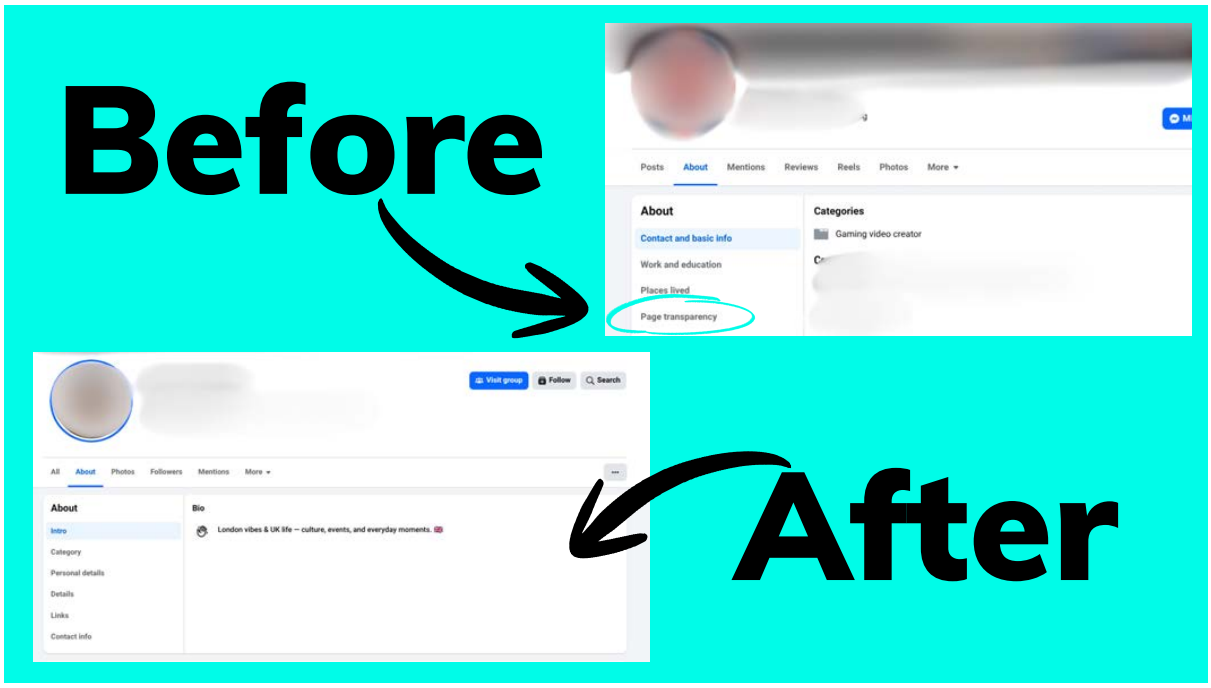
journalism.[55] The community notes system relies on establishing consensus between opposing views rather than emphasising factual accuracy, which means that many proposed notes are never published. Reviews of the system on X are lukewarm, with key concerns being little coverage of harmful content, slow publication times, and heavy reliance on fact checking (a third of community notes use fact checks as a primary source).[56] Following the post-Southport riots in July 2025, Demos concluded that *"Community Notes did not prevent harmful, false rumours about the attacker amassing millions of views: Posts that were false and relied on harmful stereotypes continued spreading without a Community Note."*[57]

Recent analysis has also shown how community notes can be manipulated by a relatively small network to shape political discourse. In April 2026, an investigation by Indicator found that during the 2024 general election, five users on X worked together to remove community notes from accounts linked to the Conservative party.[58] This reportedly led to a reduction in visible notes on tweets, including by then Prime Minister Rishi Sunak.

Early research into Meta's version of community notes suggested that the system is not yet ready for wider rollout: only half of the notes in the research were useful in any way, two were inaccurate, and none of the notes were appended to relevant posts.[59] The Meta Oversight Board published an opinion in March 2026 which said that, *"insofar as Meta envisions community notes as its primary way to address misinformation falling short of its threshold for removal (i.e., where there is not a likelihood of contributing to the risk of imminent physical harm or to interference with the functioning of political processes), the Board finds that the program's design may limit its ability to accomplish that goal."*[60] In oral evidence to the SIT Committee in March 2026 (around a week before the Oversight Board published its opinion), Meta confirmed that its long-term plan is to introduce community notes in the UK.[61] Whether this includes dismantling its fact checking programme is unclear.

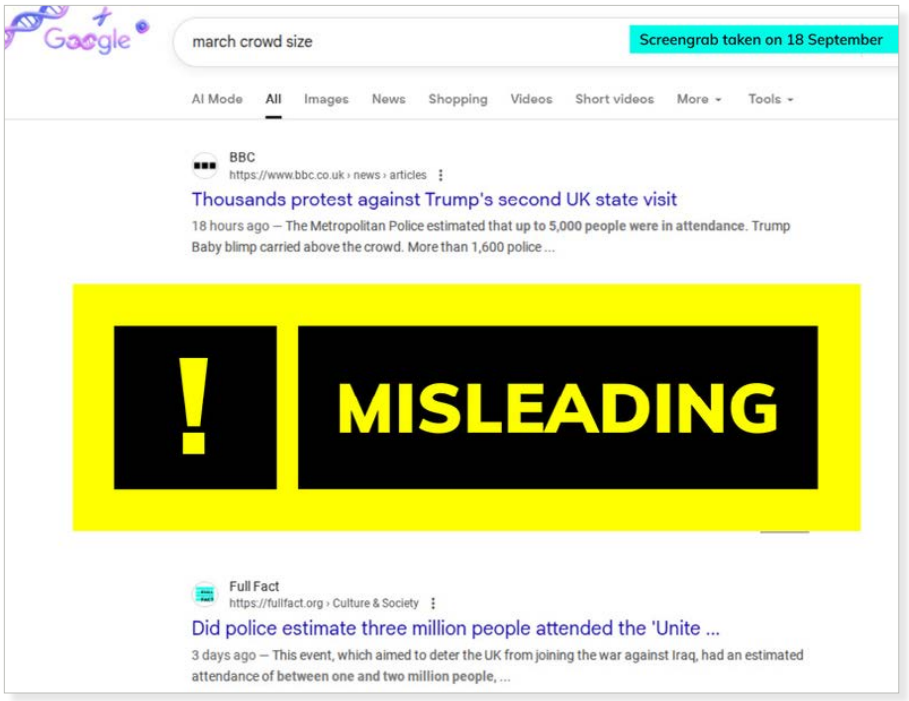
Google ended its funding relationship with Full Fact in the months following Donald Trump's re-election,[62] and chose not to renew its support for both the International Fact-Checking Network's Global Fact Check Fund, and the European Media and Information Fund, to which it had given around £38 million combined since 2013. Separately, in a letter to Jim Jordan, Chair of the US House Judiciary Committee, published in September 2025, Alphabet said YouTube *"has not and will not empower fact-checkers to take action on or label content"*.[63]

Sometimes companies have made these changes quietly. In January 2026, Meta moved Facebook's 'page transparency' information from the 'about' sections without an apparent explanation. Whilst this information is still available directly through a page or profile, identifying the location of a page's manager is now a less intuitive process as a result of changes that were not signalled to users. This makes it easier for foreign-managed pages to obscure their true locations.




In evidence to the SIT Committee in February 2025, and again in March 2026, a representative for Meta said that misinformation is bad for business.[64] This didn't stop the company from altering its misinformation policy in April 2025, to remove references to hoaxes and viral misinformation.[65] Researchers predicted that this change, combined with the rollout of its monetisation programme, would create the conditions for hoaxes to flourish.[66] A Reuters investigation in November 2025 revealed that Meta projected that 10% of its 2024 revenue would come from adverts for scams and banned goods.[67] An investigation by Maldita in July 2025 found more than 1,000 fraudulent Facebook pages across 60 countries pretending to be public transportation services to scam citizens.[68]

In June 2025, Google phased out support for Claim Review, reducing a key route through which users had encountered contextualised fact check information directly in Google search.[69] Despite Google reporting in 2024 the snippet had been used to serve fact check entries over 120 million times within the first six months in the EU alone, the company claimed it was removing the snippet as it was "*not commonly used in Search*". Full Fact later found numerous examples of search snippets distorting our work by pulling them out of context.[70] In one example, search users were presented with supposedly fact checked information implying that crowd sizes at a far-right rally were ten times more than they were in reality.




As Chi Onwurah, the chair of the SIT Committee, noted during an evidence session with TikTok, Meta, Google and X representatives in March 2026, platforms "*react*" without "*looking forward*".[71] This shouldn't be a surprise: there is no obligation under the Online Safety Act 2023 (OSA) for platforms to 'look forward' when it comes to misinformation and disinformation, in the sense of assessing and mitigating the systemic harms their services pose, such as to civic discourse, electoral integrity, public security or public health. Platforms will continue to mark their own homework until the OSA is upgraded.

Provenance and labelling of AI-generated content is one of the few areas where voluntary online safety measures across platforms have made gains in the past year. For example, YouTube stopped allowing monetisation of channels with "*repetitive or mass-produced*" content (targeted at AI slop rather than specifically at AI-generated content masquerading as genuine).[72] X announced in March 2026 that it would suspend users from its revenue programme if they posted AI-generated videos showing armed conflict without labelling it as made with AI.[73] The policy is very narrow and AI label metadata is inconsistently detected by platforms—but it is encouraging that X eventually saw the need to contextualise misleading information.

### 1.2.3 AI-mediated information flows

Generative AI is rapidly changing the world of information and is now widely used in the UK. More than half (54%) of UK adults use AI tools like ChatGPT or Gemini, increasing to 79% for those aged 16-24.[74] Interview-based research by Ofcom in 2025 found that users have also started to rely on Google AI Overviews more frequently,[75]

### Accuracy and consistency challenges in AI Overviews

AI Overviews were introduced to Google Search for US users in May 2024,[76] replacing the top of the traditional list of links with a block of text. Overviews were introduced to users outside of the US in August 2024,[77] followed by AI Mode in 2025[78] which Google says is for *"complex, longer and multimodal questions"*.[79] While Overviews come with clear benefits, they also present challenges for information integrity—and were introduced as an experiment and work-in-progress[80] without robust safeguards.

A Full Fact investigation in August 2025 revealed that AI Overviews of searches with Google Lens were giving users false and misleading information about certain images being shared widely on social media.[81] We ran searches for screenshots of key moments of misleading videos that we had fact checked in recent months using Google Lens, and found the AI Overviews for at least ten clips failed to recognise inauthentic content, or otherwise shared false claims about what the images showed. We found examples of AI Overviews that repeated debunked claims from social media, made up information about world events, failed to identify AI-generated content, failed to identify video game footage, and produced contradictory results for identical searches.

Full Fact has recommended measures to address the misinformation risk of AI search summaries, including mitigating the repetition and amplification of misinformation from social media, transparency and information literacy aids for when generative AI answers are not shown to users, and stronger user empowerment measures—such as quality signals and user controls on encountering AI-generated content.

AI-mediated systems are reshaping how political information is encountered by selecting, summarising and sequencing material before users reach original sources.[82] Search overviews and chats are powerful editorial layers, determining which sources are cited, which claims are foregrounded, and which perspectives are excluded. Despite the public significance of these choices, there is little transparency over how they are made by powerful technology companies, and limited levers to disclose them.

While their use and influence is growing, AI tools are not yet widely used or trusted for verification in one key area. When they encounter political information during an election campaign and are unsure whether it is true, Full Fact's polling shows that only 5% of UK adults say they would be most likely to use an AI tool to check —substantially lower than those who would be most likely to use a fact checking website (27%).[83] This compares to 49% who would be most likely to search online for more information and 35% who would be most likely to wait and see if a trusted news organisation reports on it.

**Source selection and authority**

AI companies concentrate gatekeeping power within opaquely designed systems. Evidence suggests that some generative AI systems rely on a narrow range of prominent news brands, with limited transparency about how they are selected, ranked or weighted.[84] Testing has identified weaknesses, including misattribution, broken citations and the confident presentation of incomplete information.[85] A BBC review in 2025 found that nearly half of responses from AI assistants to news and current affairs questions contained at least one significant issue, most commonly sourcing.[86] Another study in 2025 found that AI search engines cite incorrect news sources at a rate of 60%.[87]

When authoritative summaries draw on poorly contextualised or low-quality sources, such as social media, they can blur distinctions between evidence, opinion and user-generated content. There is also little transparency over the rules that shape the outputs of AI systems. In September 2025, it was reported that Apple issued new guidelines for how its chatbot should talk and evaluate answers two months after the inauguration of President Trump.[88] The guidelines reportedly framed politically sensitive topics in terms that appear to align with views of the US administration. Significant choices are being made in American boardrooms which travel far beyond the country's borders.

At the same time, research has shown a huge decline in click-through rates for search, indicating that many users are looking no further than AI summaries.[89] While traditional search results are often available below AI summaries, this is not the case for chatbots, and sourcing style and design varies between services. This erodes the consistent cues, such as source, diversity and provenance, that help people to assess credibility.

A new commercial ecosystem has emerged in which brands, publishers and interest groups attempt to actively engineer content in order to influence what large language models (LLMs) say. These techniques, which are designed to shape the information AI systems select and present as authoritative, create a clear structural vulnerability. Answers given to the public by generative AI interfaces are not determined solely by the model's training data and retrieval capabilities; instead, they are increasingly subject to external optimisation by actors whose interests may not align with accuracy or the public good.

**Compression and interpretation of nuanced information**

AI systems compress material—which may be complex, evolving or contested—into singular outputs. In doing so, they often omit uncertainty, disagreement or evidential limits that would be apparent in full reporting and within its original context (such as what the identity of the publisher or author says about the information being presented).

During Full Fact's 2025 investigation, we found examples of our fact checks being mangled by AI Overviews. A video shared on social media claimed to show footage of a car accident that killed the Portuguese football player Diogo Jota and his brother, André Silva. But the footage actually showed an accident that took place in the US in 2023.[90]

The overview said that the image "*depicts a severely damaged black car wrapped round a tree*", when in fact the image showed something else. Full Fact had written an article about the same topic, which investigated a different piece of content.[91] Google appears to have connected the image to the wrong fact check.

Full Fact's AI tools are designed so that AI is always used alongside human intelligence, allowing humans to add vital context, caveat and nuance. But referral to original sources is declining on many platforms, meaning users encounter conclusions without the interpretative context typically provided by journalists, researchers or fact checkers. While this reduces friction in accessing information, it narrows the space for scrutiny and can allow misleading narratives to flourish, particularly when events are fast-moving.

**Amplification and repetition**

AI systems dramatically reduce the cost of producing and re-expressing claims, enabling rapid replication across formats and platforms. Variants of the same claim can circulate simultaneously, complicating attribution, making timely correction difficult, and potentially giving a claim undue credence by making it seem like a widely held view.

As conversational systems become more personalised and responsive, they increasingly become primary interpreters of events. Outputs are transient, tailored and difficult to audit, yet engagement can shape perceptions. Significantly, a study published in December 2025 found that sustained interaction with conversational AI can meaningfully change voting preferences, and is more effective than traditional political advertising.[92]

Where errors in chatbots occur, corrections are inconsistent and accountability remains unclear. This was underlined towards the end of 2025, when the Grok chatbot posted extremely serious false comments about Pete Wishart MP.[93]

> *The way in which we inform ourselves is changing rapidly, making it harder to distinguish what is real. AI generated content and simplified summaries are reshaping how we understand complex issues, while trust in traditional media and political institutions is declining. Misinformation is widespread, fuelling confusion and contributing to a more polarised political debate. This makes stronger regulation, accountability, and safeguards increasingly necessary to address the risks of AI driven misinformation.*
>
> **Pete Wishart MP**

**Information pool and training data**
Emerging evidence suggests that relatively small volumes of strategically crafted or malicious training material can influence the output of LLMs, particularly when presented as a credible source.[94] This creates an upstream risk: distorted or strategically seeded content may not only be amplified by AI systems but normalised within routine outputs. In high-stakes contexts, like elections or crises, misleading framing may appear and spread quickly within authoritative-seeming outputs.

AI systems therefore do more than introduce new risks to the information environment: they reconfigure how authority is produced and perceived. When sourcing is opaque—and when misinformation is produced and repeated at scale—authoritative information struggles to achieve comparable visibility. As these systems become routine entry points to political information, their operation should be a central concern for democratic governance.

## 1.2.4 Influencer ecosystems and monetisation incentives

Economic incentives play a significant role in shaping how political and other information circulates online. Visibility, reach and engagement are directly linked to revenue, status and audience growth for both platforms and content creators.

Platform design encourages frequent, high-engagement content through metrics such as likes, shares, comments and view counts, creating pressure to make content that is attention-grabbing.[95] Incentives favour simplified, emotive or provocative messaging, while nuanced, corrective or slower-moving content is less likely to gain traction.[96] Advertising revenue, creator funds and affiliate models are typically linked to scale and interaction rather than informational quality, reinforcing feedback loops in which engaging content is amplified and replicated regardless of its quality or reliability.

Recent research illustrates the scale of these incentives. Reuters reported in 2025 that Meta estimated a substantial share of its advertising revenue came from *"higher risk"* scam adverts. The investigation described internal concern that stricter enforcement against fraudulent advertisers could reduce income, highlighting tensions between Meta's revenue incentives and moderation of harmful or deceptive content.[97] The YouTube channel *Bandar Apna Dost*—which posts AI-generated content—has accumulated over 2 billion views, generating an estimated $4 million (£2.9 million) annually.[98] A BBC Verify investigation in March 2026 found numerous examples of AI-generated videos and fake satellite images which had been used to make false and misleading claims about the US-Israel war with Iran, which had together amassed hundreds of millions of views.[99]

Investigations have also identified coordinated networks that monetise misleading political content through advertising, affiliate marketing and traffic redirection strategies.[100] These systems operate across jurisdictions and services, creating challenges for enforcement. Monetisation is therefore not only a feature of platform design but part of a wider commercial ecosystem that rewards scale, repetition and emotive framing over verification.[101] This increases background noise and complicates attribution, making authoritative information harder to surface and correction harder to scale.[102]

Platform features can extend the visibility and lifespan of this content, prolonging uncertainty and undermining timely correction. In October 2025, the government *"acknowledge[d] the challenges posed by misinformation on social media platforms, especially when amplified through platform business models and design."*[103] In its 2025 report on the post-Southport riots, the SIT Committee recommended a range of measures needed to tackle the monetisation of misinformation—including more transparency in digital advertising, and more powers for Ofcom to take action against platforms which allow harmful content to be monetised through their services.[104]

### 1.2.5 Public resilience and public interest information

Public interest information—including public service broadcasting, independent fact checking and other forms of journalism—is essential for helping people to access reliable information and assess credibility. This in turn helps stabilise the information environment. It provides shared reference points, context and correction when misleading claims circulate. However, this infrastructure faces sustained pressure. Economic constraints on journalism, declining local coverage[105] and audience fragmentation are reducing the reach of reporting, making it harder for reliable information to surface when it matters most.

Fact checkers play a frontline role in this system. They identify and respond to misleading or harmful claims as they emerge, working to limit their impact on public understanding, health and safety. In late 2025, the US government introduced enhanced vetting of certain speciality visa applicants, including scrutiny of people's work history in areas such as fact checking, misinformation or content moderation.[106] The US government also denied visas to five people working on misinformation and disinformation (although Imran Ahmed, who lives in the US, has not had his permission to remain in the country cancelled, and is challenging the government's decision in court).[107] The administration continued to put pressure on misinformation research and fact checking initiatives, by reviewing or withdrawing federal grants worth hundreds of millions of dollars.[108]


At the same time, changes in how information is surfaced on major platforms have reduced the visibility of authoritative content. The platform changes examined above, such as the removal of Claim Review, have reduced the opportunities for users to encounter corrections at the point of search or discovery.[109] Chatbots and AI-generated summaries are also substituting for direct engagement with original sources[110] while weakening the economic sustainability of public interest journalism through declining referral traffic.[111]

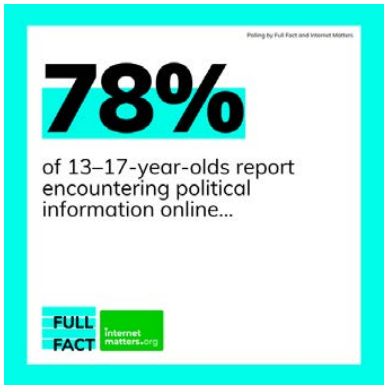
There has been recent policy attention on the need to give more prominence to public interest information. In March 2026, the government set out welcome plans to review whether the crisis powers in the Online Safety Act are fit for purpose, including consideration of measures to give trusted media greater prominence *"so people have access to authoritative sources to counter mis and disinformation."*[112]

Media and information literacy are critical components of public resilience. The ability to assess sources, evaluate evidence and understand how information is produced can reduce vulnerability to misleading claims and support informed participation. However, these skills must continue to evolve as AI-mediated systems reshape how information is produced, ranked and consumed[113]—and must be part of lifelong learning, not just focused on the classroom. There has also been significant policy attention on media literacy in the past year, with a series of parliamentary, regulatory and policy reviews.

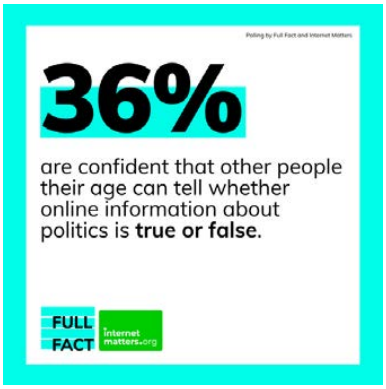
The House of Lords Communications and Digital Committee report on media literacy called for stronger government coordination, sustained funding, and more systematic evaluation.[114] The government's response to that report, in October 2025, accepted the importance of media literacy and reiterated a cross-government approach, but did not commit to major new duties or long-term, sustainable funding mechanisms.[115] The government's response to the Curriculum and Assessment Review supported strengthening and embedding media literacy within the citizenship curriculum.[116] The government's Media Literacy Action Plan, *A safe informed digital nation, set out continued coordination, partnership working and evidence-building activity across government, regulators and industry.*[117] The plan is a welcome and constructive step, reflecting a serious commitment to strengthening media literacy and tackling misinformation, with an emphasis on improving coordination and how different parts of the system work together.

Expecting individuals to take full responsibility for navigating complex, fast-moving information systems imposes unrealistic demands, particularly when regulatory frameworks are lacking and platform incentives are not aligned with quality information. Research from Ofcom shows that media literacy is most effective when supported by platform design that makes provenance, reliability cues and context visible.[118] Ofcom's 2025 consultation on promoting media literacy emphasised evidence-based, measurable interventions and closer coordination between regulators, platforms and civil society.[119] It signals a shift toward integrating 'media literacy by design' into digital services themselves, rather than simply relying on educational initiatives. Meanwhile, the World Economic Forum has highlighted the need to embed media and information literacy across diverse contexts, from media organisations to local governments and digital platforms.[120]

Polling by Internet Matters and Full Fact, in November 2025, underscores how demanding the information environment has become for younger people. A large majority of 13–17-year-olds (78%) report encountering political information online, yet only half (53%) feel confident telling whether it is true or false, and only a third (36%) are confident that other people their age can tell.[121] The need for prompt and effective reform is acute with the government's plan to lower the voting age to 16. Without more support, many newly enfranchised voters may struggle to engage confidently or meaningfully in democratic life, creating a real risk of early disengagement from politics.








As public interest information becomes harder to sustain, less visible and increasingly crowded out, the information environment becomes more fragmented and uncertain. This places a greater burden on individuals to assess credibility, distinguish reliable from misleading information, and navigate digital spaces with fewer shared reference points. Part 2 examines how these pressures can translate into democratic harms.

# Part 2: Impacts on democracy and the case for reform

This section examines four types of democratic harm associated with misinformation and sustained uncertainty in the information environment:

- The erosion of shared understanding
- Reduced willingness to participate in democratic processes
- Weakened confidence in institutional authority
- The gradual breakdown of community cohesion

Full Fact's 2025 report analysed the 2024 general election and found that, despite fears about deepfakes, the election was largely a blend of political spin, online misinformation on social media, and easily debunked 'cheapfakes'. This section assesses why the next general election is likely to take place in materially different circumstances, resulting in a very different stress test of the UK's democratic information resilience.[122]

## 2.1 Why it matters: harms to democracy and the erosion of trust

The pressures described in Part 1 undermine something essential to the functioning of a healthy democracy: the ability of individuals, communities and the public to form reliable and shared understanding. When the idea of a shared reality starts to break down, the costs of democratic participation increase, institutions find it harder to sustain authority, and disagreements become less likely to be resolved with evidence.

### 2.1.1 Informational harms: confusion, ambiguity and shared understanding

Informational harms arise when people are unable to confidently assess what is accurate, reliable or authentic. Full Fact's polling highlights the scale of this challenge: only 3% of UK adults find it very easy to distinguish a genuine from an AI-generated video online, and 3% feel very confident being able to distinguish between genuine and fake political information online.[123] This points towards widespread uncertainty about information itself.

When it is difficult to determine what is fake or AI-generated material, the effort required to evaluate political claims, weigh up policy options, engage in debate or vote in an election increases. Public discussion is less likely to stabilise around shared evidence or accounts and more likely to remain fragmented or contested. The result is an environment in which shared understanding is harder to achieve and sustain. Different audiences may encounter divergent versions of events across different sources, with limited convergence on a single account even where authoritative information is available.

Evidence indicates high levels of public concern. Survey data from the Electoral Commission in 2024 found that misinformation and disinformation were among the public's leading concerns in elections.[124] In the same year, Ofcom found that 60% of people encountered false or misleading material about the general election in the previous week; many were worried about deepfakes but almost half (46%) were unsure whether they had seen one in the previous week.[125] International polling also indicates that large majorities (77%) view the spread of disinformation as a real threat.[126]

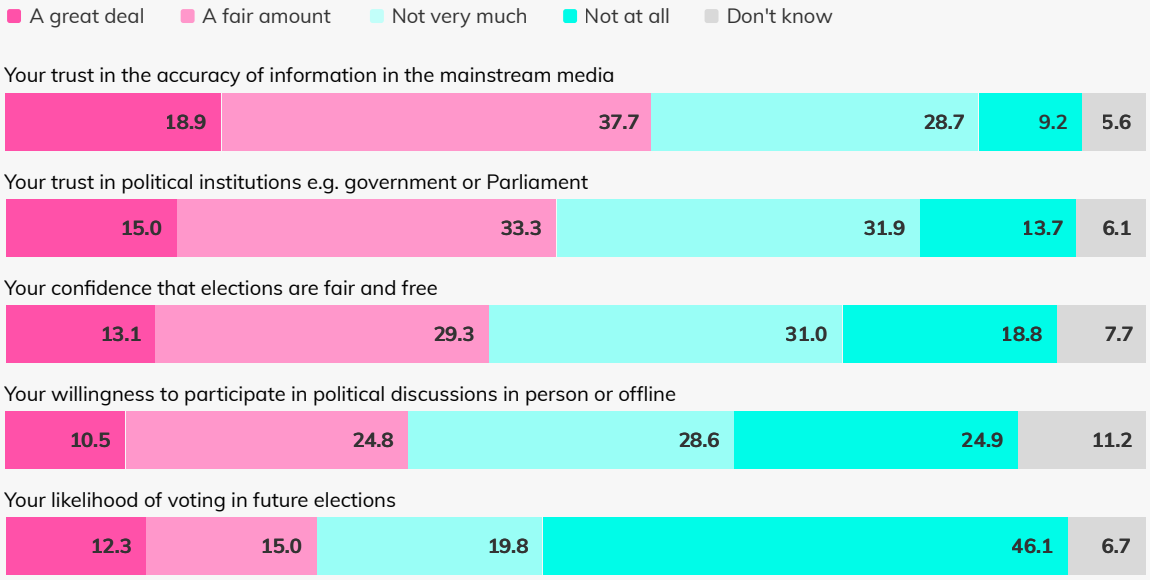
Even where individuals remain sceptical, the difficulty establishing what can be known with confidence raises the cost of participation, and can lead people to disengage, or rely on familiar or emotionally salient cues rather than evidence. Where information is unclear, incomplete or emotionally charged, people are more likely to share unverified claims, particularly within like-minded networks—allowing repetition and social endorsement to replace verification, and leading to further fragmentation.[127]

There are some early indications that these dynamics can affect institutional decision-making. In February 2026, a senior police officer apologised after AI-generated information was cited in support of a proposed ban on Maccabi Tel Aviv supporters, which was later found to be false. This illustrates how inaccurate information can extend beyond individual users to influence official decisions and actions.[128] Even when false or misleading information does not create immediate doubt, cases like this can still erode trust—and create challenges in knowing what to trust—among institutions and the public.

### 2.1.2 Behavioural harms: disengagement and withdrawal

Full Fact's polling found that four in every five (80%) UK adults are concerned about political misinformation, compared to 13% who are not concerned, with the rest saying they do not know.[129] The polling indicates that this concern is contributing to disengagement from democratic processes at a large scale. Of that 80%, many say that it has negatively affected how they feel about important issues over the past year:

**You previously said you were concerned about misinformation related to politics (this was 80% of UK adults polled). In the past 12 months, to what extent have your concerns about misinformation related to politics negatively affected your feelings about the following?**




Research links sustained exposure to contested or conflicting information with lower levels of political engagement, including online participation.[130] Adaptive responses include shortened attention and increased reliance on familiarity, identity alignment or emotional cues.[131] Technologies such as deepfakes intensify these effects by increasing doubt about what can be trusted, making confident participation in democracy more difficult.[132] Faced with ambiguity, individuals may avoid news or withdraw from debate.[133] The OECD has recognised this threat, saying: "*the amplification of mis- and disinformation content can undermine the public's willingness and ability to engage constructively in democratic life, and down the line the ability of society to forge consensus.*"[134]

These pressures are not evenly distributed across society. Individuals with less time, education, or resources face greater barriers in navigating complex information environments. Trying to cope with uncertainty places high demands on attention and memory; and when these demands exceed capacity, disengagement becomes more likely. Over time, this risks widening inequalities in political participation leaving democratic debate both less inclusive and more vulnerable to distortion.[135]

### 2.1.3 Community harms: tension, scapegoating and offline effects

A healthy society and vibrant democracy depend on the ability of communities to establish shared and credible accounts of events. Opinions may differ, verifiable facts should not. When uncertainty disrupts collective sense-making, social cohesion can be undermined, and persistent uncertainty about what is genuine or fake can erode trust in local institutions and community relationships, by destabilising shared understandings of events.[136] These risks are particularly acute where access to trusted local information is limited, or where there are also underlying social or economic pressures.

> *In democratic societies, characterised by freedom of speech and open debates as a way of reaching consensus at all levels of society, the search for information integrity is key to the ability of societies to hold together.*[137]
>
> **Organisation for Economic Cooperation and Development, 4 March 2024**

These are the circumstances in which misleading narratives fill information vacuums, offering simple or emotionally resonant explanations before verified accounts emerge. This can lead to scapegoating and division, increasing the risk of offline harms—for example when fake or manipulated content fuelled the mass killings of Alawites in Syria in 2025,[138] or disinformation's contribution to Nigeria's ongoing farmer-herder conflict.[139] The disorder following the Southport attack, when a small number of high-reach accounts continued to circulate inflammatory narratives long after the event,[140] showed how delayed or fragmented clarification can heighten community tension.[141]

Misleading or false narratives often target individuals in public life. In 2025, the Speaker's Conference concluded that *"addressing disinformation is a necessary step for reducing abuse and intimidation against MPs and candidates, as many cases are triggered by disinformation about the victim."*[142] When false or distorted claims take hold, the distinction between informational harm and direct harm to individuals can break down. During a debate on the Representation of the People Bill in March 2026, Rushanara Ali MP said that some MPs have wondered whether *"if we had known what we know now about the state of harassment and intimidation in our politics, we would have stood for Parliament,"* and argued that disinformation online *"fuels intimidation, hostility and violence offline."*[143]

The Electoral Commission has highlighted campaigns in the last general election *"which were trying to destroy community cohesion in this country"*.[144] This sort of activity may amplify grievances and false narratives at scale, complicate attribution and prolong

uncertainty. The government's social cohesion strategy, published in April 2026, recognises that extremists "*rely on spreading their narratives and ideas throughout society—both on and offline*", and has made commitments to detecting, exposing, and countering extremist influence across the UK, including online.[145]

### 2.1.4 Institutional harms: erosion of trust as a consequence of misinformation

Part 1 examined how low trust shapes the spread of misinformation. This section focuses on how misinformation, in turn, erodes trust in institutions and weakens their authority. The public's concern about misinformation appears to be having a corrosive effect on public trust. Scepticism is good—institutions have to earn trust—but cynicism is corrosive.

According to Full Fact's poll, of the 80% of UK adults who expressed concern about political misinformation, the strongest negative impacts relate to trust. 48% in this group said their trust in political institutions had been negatively affected by their concern over the past year, while 57% said their trust in the accuracy of information in the mainstream media had been negatively affected.[146] This builds on a series of polls in recent years that have found rock bottom levels of public trust in UK political institutions.[147]

Deepfakes can undermine trust by convincingly imitating trusted individuals and organisations. For example, a Full Fact investigation found that deepfake videos of academics and health leaders have been used to promote supplements on TikTok and Instagram.[148] The Centre for Emerging Technology and Security has highlighted a series of deepfakes of political figures in recent years—and found that these techniques do not just add perceived credibility to deceptive content, but damage public trust in the organisation or individual whose likeness has been used. Over time, this contributes to general scepticism in institutions, genuine sources and communications.[149]

> *As the first MP to be the target of a serious political deepfake disinformation campaign falsely announcing that I was defecting to another party, this Full Fact report on the dangers of AI deepfake technology further corroding public trust in democracy is timely and important. The digitalisation of political campaigning requires agile and proportionate protections against disruption of our democratic systems. We ignore the risk at our peril.*
>
> **George Freeman MP**

Where the institutional response is fragmented and opaque, it is harder for people to see action being taken or have confidence in that response.[150] As trust erodes, delayed or low-visibility communication may be interpreted as incompetence, bad faith or conspiracy—particularly when false or misleading claims circulate faster than official clarification. This can reduce the effectiveness of interventions and create challenges in coordinating action, maintaining authority and stabilising contested situations.

The erosion of trust between the public and political institutions can also lead to extreme behaviour, such as abuse of candidates. In a review commissioned by the government on countering foreign influence in UK politics, published in 2026, Philip Rycroft noted: *"If relentless exposure to disinformation on social media persuades even a small proportion of the UK population that our politics is irretrievably broken, the risk grows rapidly that some will seek to resolve their discontents by extra-political action."*[151]

## 2.2 Why the next general election will be different from 2024

The next UK general election will take place in a very different information environment from the last one. Political information will be more heavily shaped by AI systems like chatbots and search overviews, more easily manipulated or synthesised, and filtered through platform systems whose safeguards are less predictable than in previous cycles, and in many cases have been weakened since 2024.

### 2.2.1 AI as an increasingly dominant information layer

Since Full Fact's report on misinformation and democracy in 2023, AI powered search, chatbots and automated summaries have moved closer to the point of initial access to information for many users.[152] In 2025, 64% of people in the UK regularly saw an AI-generated answer to one of their searches in the past week.[153] As their use becomes routine, AI intermediaries play an increasing role in how people encounter, filter and interpret political information. But these systems operate with limited transparency, accountability or regulatory oversight.

The 2026 International AI Safety Report highlights growing uncertainty about how increasingly capable, general-purpose AI systems behave under real-world conditions, particularly where oversight and evaluation lag deployment.[154] The UK's institutional response mechanisms, including in election contexts, are not designed for rapidly adapting, multi-purpose systems operating across platforms and borders. Safeguarding the resilience of the UK's information environment requires policymakers to clearly understand the behaviours and impacts of generative AI systems.

## Full Fact's LLM benchmarking project

For millions of people, LLMs are becoming the first place they turn, as conversational AI interfaces supplement—or even replace—traditional search engines. As these tools become embedded in everyday life, the quality and reliability of what they produce becomes a fundamental matter of public interest. Yet there is currently no independent, public interest mechanism to systematically evaluate whether these AI services are accurate, transparent, timely, consistent, or responsible in the information they provide.

Existing LLM benchmarks such as Google DeepMind's FACTS Benchmark Suite, SimpleQA, and academic datasets like TruthfulQA, make important contributions but share several limitations from a public interest perspective:

- They are typically produced by AI companies or academic groups with potential conflicts of interest, or without the editorial fact checking expertise needed to evaluate real world factual claims.
- They tend to focus on narrow technical dimensions rather than the broader set of qualities the public needs from an information service.
- They are snapshot exercises rather than continuous monitoring, meaning they cannot detect performance drift, regression after model updates, or inconsistency over time.
- They rarely test claims that are contested, politically sensitive, or subject to active misinformation, which are precisely the areas where accuracy matters most.
- Their methodologies are often opaque or difficult for non-specialists to scrutinise.

Full Fact is developing a benchmark to evaluate LLM performance across five dimensions, each chosen because it reflects a quality the public has a right to expect from any information service:

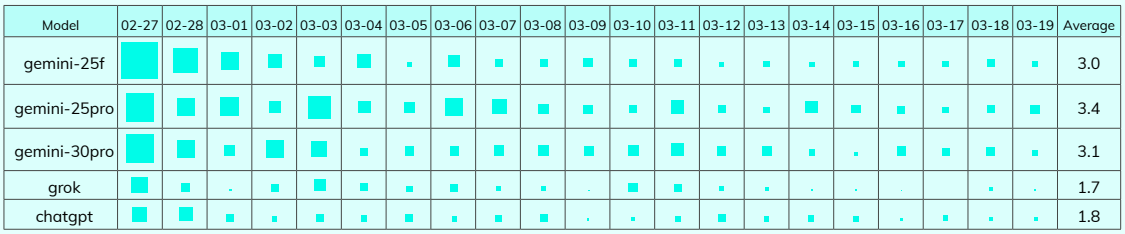
- **Factuality:** Do responses contain verifiably accurate claims, avoid hallucination and correctly represent the evidence base?
- **Transparency:** Does the model communicate uncertainty, cite or attribute high quality sources, acknowledge limitations, and distinguish fact from opinion?
- **Timeliness:** Do responses reflect current information rather than outdated data, and does the model recognise when its knowledge may be stale?
- **Consistency:** Does the model give materially the same answer to the same question over time and when asked in different ways?
- **Civic responsibility:** Is information about democratically important questions balanced, does it not amplify misinformation, and does it support informed participation?

Since January 2026, we have asked leading AI models (including ChatGPT, Gemini and Grok) the same set of questions each day, including questions about UK elections, international events and general knowledge. We have so far collected and stored over 11,000 responses to score against the five dimensions above. For example, to measure consistency, we analyse the sources of information that each model uses in response to each question over a period of time. A *consistent* model would typically find a small set of reliable sources, such as government web pages or fact checking sites, and keep referring users to those URLs, while a less consistent model would keep using new sources even though the question remains unchanged. In the full benchmark tool, we also measure the consistency with which pieces of information are provided as well as the sources.

The figure below shows the average number of previously unseen URLs that each model referred users to over a period of three weeks, averaged across all questions. Their first responses (left-hand column) contain only previously unseen URLs, by definition, and the numbers tend to get smaller over time as the same URLs are repeated, corresponding to less novelty. But note that this is not uniform and for this sample, both Grok and ChatGPT tend to rely on a smaller set of information sources than the Gemini models, as shown by their lower average number of new URLs per day (right-hand column).

**New URLs per day for each AI model**

Time period: 27 Feb - 19 March 2026



| Model | 02-27 | 02-28 | 03-01 | 03-02 | 03-03 | 03-04 | 03-05 | 03-06 | 03-07 | 03-08 | 03-09 | 03-10 | 03-11 | 03-12 | 03-13 | 03-14 | 03-15 | 03-16 | 03-17 | 03-18 | 03-19 | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| gemini-25f | | | | | | | | | | | | | | | | | | | | | | 3.0 |
| gemini-25pro | | | | | | | | | | | | | | | | | | | | | | 3.4 |
| gemini-30pro | | | | | | | | | | | | | | | | | | | | | | 3.1 |
| grok | | | | | | | | | | | | | | | | | | | | | | 1.7 |
| chatgpt | | | | | | | | | | | | | | | | | | | | | | 1.8 |

Scale: ■ = 5 new URLs

Sometimes there is a trade-off between consistency and factuality. For example, when election results are published, it makes sense for models to refer to new sources to share this new information. Perfect consistency in this sense is not necessarily ideal, which is why the final version of the benchmark will combine all five dimensions listed above.

It is hard to rely on someone if they give a different answer each time you ask the same question. Users of AI tools usually only ask their question once, so the model's inconsistency would not be clear. Our benchmark helps provide this clarity.

### 2.2.2 Synthetic content at scale and targeted deployment

Discussions convened by the Government Office for Science in 2024 found no clear evidence that deepfakes have so far meaningfully affected voting choices in the UK, noting the difficulty of attributing causation.[155] That should not breed complacency. Synthetic material is becoming more sophisticated and cheaper to produce. The barriers to create misleading content have all but collapsed.[156] Inaction is not an option.

Particularly significant is how synthetic content interacts with timing, attribution and institutional capacity—and the challenge posed by uncertainty in knowing what information to trust. Tactics that undermine or cause confusion about a political campaign are not new. For example, in 2017, the campaign of then French presidential candidate Emmanuel Macron said it was the target of a coordinated hack—and that genuine files were published along with fake ones in order to cause confusion.[157] But rapid advances in technology are making these tactics more complex, more threatening, and harder to counter. We cannot afford to reach a point where no one believes anything.

The challenge is heightened when synthetic content is released close to polling or during fast-moving campaign events. Even a small number of credible-looking items can generate uncertainty that outlasts corrections, and security analysts warn that content need not be persuasive to be harmful.[158] UK broadcasters observe a pre-election 'period of heightened sensitivity', restricting campaign coverage,[159] and last-minute synthetic content can escape scrutiny in this period or spread before corrections. In Slovakia, a deepfake audio clip appeared 48 hours before the 2023 parliamentary election, leaving little time for fact checking.[160] Similarly, in Ireland, an AI-generated video falsely showed Catherine Connolly withdrawing from the 2025 presidential race days before voting.[161]


'Disgraceful' deep-fake AI video condemned by presidential candidate

PA MEDIA

Jake Liggett
BBC News NI

22 October 2025

Emerging research points to future risks, including the potential for coordinated networks of AI-driven accounts to increase background noise, mimic human social dynamics, infiltrate communities, reshape public opinion at scale, microtarget individuals with discrete messaging, and complicate attribution.[162] Autonomous AI systems have the ability to generate, adapt and disseminate political content with limited direct human oversight. More agentic systems could operate persistently, adjusting outputs in real-time and complicating attribution, accountability and response.[163] In the context of an election, it is likely that such systems

will be deployed to produce and circulate political messages across multiple platforms, and tailored to different demographics, making it difficult to distinguish organic public engagement from coordinated influence activity.

### 2.2.3 Platform capacity and weaker safeguards

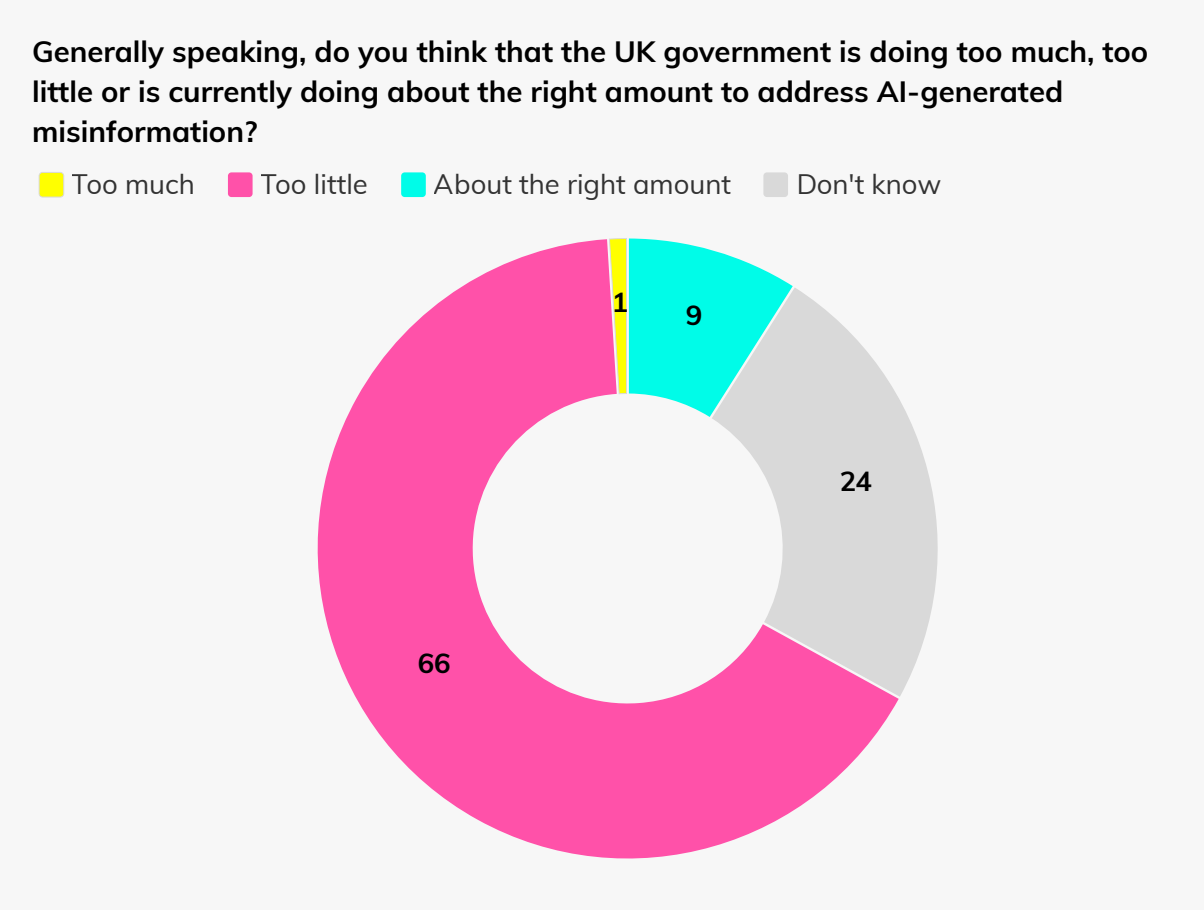
Effective democratic information systems depend on platforms and their safeguards functioning predictably under pressure. This is made far less certain by changes to policies, moderation resourcing and enforcement approaches.

Greater reliance on automated systems, or community-based correction mechanisms, which can lack impartial expertise, introduces variations in how content is handled. At the same time, reductions in trust and in safety investment in many services—as examined above—raise questions about capacity during high-risk periods. For regulators, law enforcement and civil society, this creates uncertainty about how and when misleading content will be addressed. Public authorities will have to prepare for information incidents without clear visibility into, or confidence in, platform decision-making.

On top of this, researcher access to platform data is often limited, delayed or inconsistent, restricting the ability to monitor information flows in real-time. International experience shows that even where legal frameworks exist, access to data during election periods can be contested or incomplete. Previously accessible sources of platform data under the EU's Digital Service Act (DSA) have been withdrawn or commercialised.[164] In 2026, a German court ordered X to give researchers access to data linked to Hungary's election, after initial refusals blocked timely analysis.[165] Without enforceable and timely transparency, oversight of platform behaviour remains reactive rather than preventative.[166]

### 2.2.4 Lagging institutional preparedness

The UK's laws and institutions have not kept pace with the rapidly evolving information environment—as Part 4 of this report examines in more detail. The public lacks confidence that the government is addressing the challenge. According to Full Fact's polling, despite four in five (80%) UK adults expressing concern about political misinformation, there is a public perception that action is not meeting the scale of the problem.[167]




An Electoral Commission survey in 2024 found that 76% of respondents did not believe enough was being done to tackle misinformation and disinformation in elections, while only 5% did believe this.[168]

In other domains, such as public health[169] and cyber security,[170] the government plans for low-probability, high-impact events before harm is demonstrated. The absence of clear evidence of past electoral disruption, or direct causal impacts of deepfakes on voting behaviour, is not a reliable guide to resilience when stress conditions are foreseeable and intensifying. The UK government has undertaken information crisis preparedness analysis but the outcomes and consequences of this work are not publicly known.[171]

As AI systems become more widely used and embedded in information systems, regulatory and institutional approaches must evolve in parallel. Indeed, Ofcom has said *"it is important that our regulation remains fit for the future."*[172] There is an opportunity for the government to ensure that economic and technological developments are complemented by robust safety measures, to prevent regulatory gaps and reduce risks before they cause harm. Our concern is that this is not happening quickly enough.

GUEST ESSAY:

## The next general election will be different; we must act now.

**John Pullinger CB, Chair of the Electoral Commission**

A healthy democracy depends on voters having the information they need. That information must be clear, accessible, and trustworthy. When it is not, voters can be misled or put off from participating.

Misinformation is now one of voters' biggest concerns about elections. Three-quarters of people say political misinformation is a problem. Nearly two-thirds think too little is being done. More than half say they have seen a deepfake, and only 17% think political party funding is transparent.

This is the world in which the next general election will be fought.

The 2024 general election gave us a direct test of how well we are meeting the threat from misinformation, and Britain's democracy held up well.

Our research found that over half of voters saw misleading or inaccurate information about parties, candidates, and the electoral process, and around a quarter encountered a deepfake. But the misleading content that was shared was largely called out, and trust in the process held. What protected us was awareness; voters who were sceptical of what they saw, and trusted sources like Full Fact and the BBC, ready to point people in the right direction. The real risk lies in the places corrections do not reach, such as social media echo chambers, where misleading content circulates unchallenged.

The pace of change means we cannot be complacent. The 2024 election could be the last before AI-generated content becomes genuinely hard to detect, which is exactly why we launched our deepfake detection pilot this year. As Vijay Rangarajan, our Chief Executive, put it: "Deepfakes are becoming more sophisticated and more accessible, as we have seen in elections around the world. A deepfake is yet to affect a UK election meaningfully, and we are determined to keep it that way."

Misinformation does not stay online; the toxic content circulating on social media is fuelling abuse and intimidation of candidates in the real world. Candidates are being forced to change how they campaign, or step back altogether, because they fear for their safety. After the May 2025 local elections, 61% of candidates we surveyed reported experiencing harassment or security threats during the campaign. Almost three-quarters said they had avoided some campaign activities out of fear. One female candidate told us, "I think the women bear the brunt of this because we're perceived as an easier target to convince not to run." Our research shows that

women candidates are twice as likely, and minority ethnic candidates three times as likely to report experiencing abuse and intimidation.

I believe the UK is well placed to meet this challenge. We have strong institutions, independent regulators, and voters who are more resilient than they are sometimes given credit for. Digital campaign material must now carry an imprint, giving voters transparency over who is trying to reach them. And when false information circulates about the electoral process itself, we are stepping in to correct it publicly and rapidly.

But confidence is not complacency. In Ireland in 2025, a deepfake video falsely showed a presidential candidate withdrawing from the race just days before polling day. It spread rapidly before being identified and corrected, a warning of what a well-timed deepfake could do in a close election. Bad actors are getting more sophisticated, platforms have been slow to act, and the window between disinformation appearing and causing damage is shrinking.

So what needs to change? Voters need to feel confident that what they are reading is real, that the money behind politics is legitimate, and that anyone who wants to stand for election can do so without fear. They need access to accurate information and to trust that the result, when it comes, was achieved fairly. The ultimate test is simpler: that both winners and losers accept the result.

No single organisation can do it alone. Tackling this requires genuine partnership between government, regulators, social media companies, civil society and the public.

On social media platforms, I want to be direct. The abuse harming our elections is happening in their spaces. They have to own this. Platforms must act now. Regulators must use the powers they already have more forcefully. If those powers fall short, then we should consider what other powers are needed.

Voters who encounter misinformation need somewhere reliable to turn. Organisations like Full Fact serve this purpose well. Engaging successfully with audiences at most risk of being misled is one of the biggest challenges in this space.

The long-term answer is education. Votes at 16 across the UK could bring 1.7 million new voters into the franchise before the next election. Only a third of under-18s have been taught about politics or how to spot false information at school. Yet 78% say they want to learn more. We are stepping into that space with partners across the sector, equipping young people to think critically about the information they encounter online.

The next general election will be different. We must act now. If we act together, the next general election and our democracy itself can be one that voters and candidates can celebrate.

# Part 3: Global lessons for UK policymakers

The pressures shaping the UK's democratic information environment are not unique. Democracies globally are confronting similar challenges arising from evolving information systems, platform design, AI-enabled content and declining public trust. International experience provides insight into which elements of democratic information systems are most vulnerable to pressure, and which forms of preparedness are most effective in limiting the persistence of uncertainty during elections and public crises.

Recent electoral experience illustrates how these risks are developing. For example, testimony to the UK Parliament on Moldova's 2025 Parliamentary elections highlighted a coordinated challenge across multiple domains.[173] International observers found the elections took place against a backdrop of "*unprecedented hybrid attacks, including illegal funding and disinformation and cyberattacks.*"[174] Moldova's experience underscores a lesson for UK policymakers: electoral interference takes place through connected systems of influence and cannot be effectively countered through siloed responses.

## 3.1 Global stress conditions and geopolitical trends

Democratic information risks must be understood in global as well as domestic terms, reflecting shared pressures across political and technological systems. Policymaking attention has shifted towards how information systems behave at scale.[175] AI is being considered in these terms, with the 2026 International AI Safety Report highlighting systemic risks and the emergent behaviours of AI systems, with cross-border challenges complicating governance.[176] This shift is reflected in the EU's AI Act, which adopts a risk-based regulatory framework for AI systems.[177]

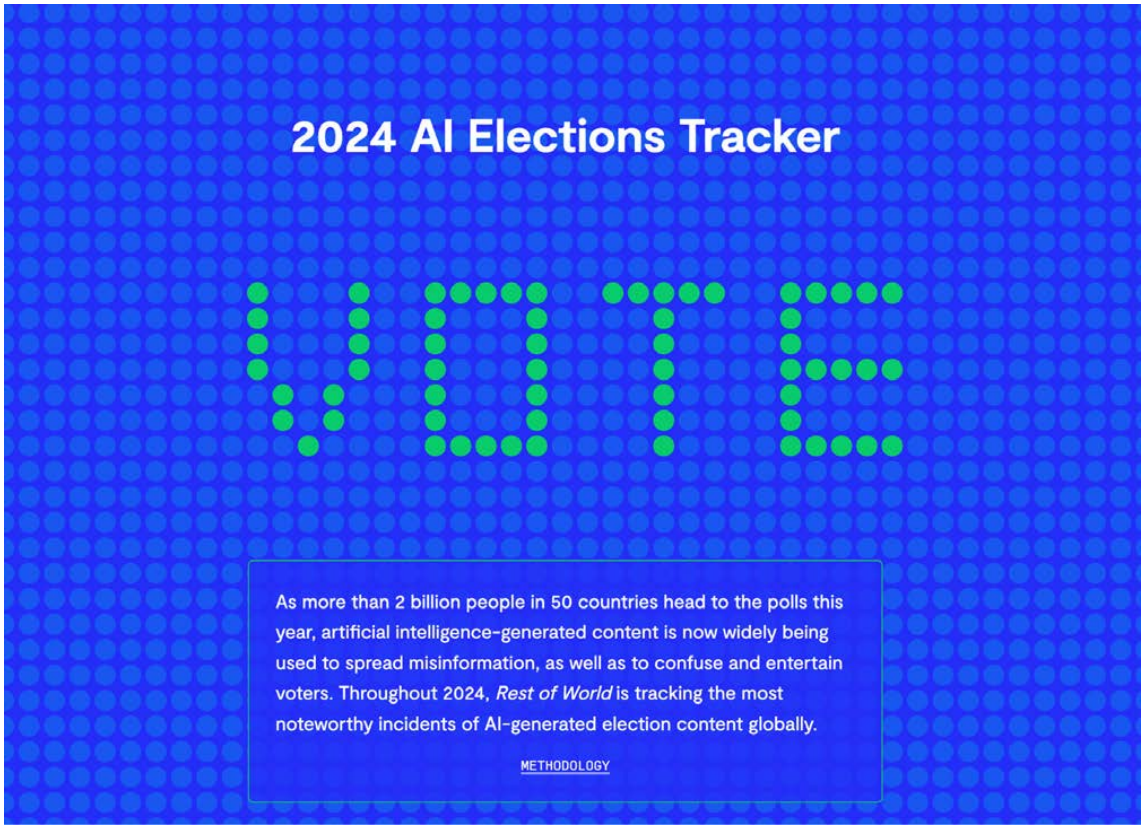
### 3.1.1 Global convergence on democratic information risks

In recent years, a number of international expert organisations with different mandates—including the United Nations,[178] World Economic Forum,[179] OECD[180] and Council of Europe[181]—have identified misinformation and disinformation as critical and cross-cutting challenges, and among the most significant global risks.

The UK government has recognised the scale of the threat in both domestic and international forums. In July 2025, the Prime Minister, Sir Keir Starmer, told a Parliamentary committee that "*I was very worried at the last election about*

*misinformation, and I am very worried about the potential for misinformation in future elections in this country."*[182] In a November 2025 statement to the UN, a spokesperson said the UK was *"deeply concerned"* by growing threats to information integrity.[183] In an April 2026 statement, they said that the UK *"is alarmed by the unprecedented rate at which threats to information integrity are growing, fuelled by the misuse of artificial intelligence."*[184]

### 3.1.2 What elections elsewhere reveal




The 2024 AI Elections Tracker catalogues real-world uses of AI in elections globally, documenting applications ranging from campaign assistance to misinformation and satire. The tracker shows that AI use is already widespread and varied, but that its effects are highly context-specific and often difficult to interpret without systematic monitoring.[185] Reviews of the 2024 global election cycle found that, while misinformation can erode public confidence in electoral processes, the direct persuasive effects of generative AI have so far been limited and difficult to measure.[186] Cross-national studies of deepfakes in 2024 similarly found low reach or ambiguous effects in many cases.[187] Comparative research shows that AI-related incidents tend to exploit existing vulnerabilities, such as polarisation and compressed timelines, rather than create entirely new ones.[188]

Certain conditions frequently occur across different contexts, including rapid amplification of claims through platform systems, particularly in the final stages of a campaign, limited time for verification at comparable scale, and difficulty establishing authoritative signals. Recent elections illustrate how these pressures materialise in practice:

- During the 2026 Hungarian election campaign, researchers identified wide circulation of AI-generated videos and images, and deepfakes of political opponents.[189]
- Three days before Ireland's 2025 presidential election, a deepfake falsely claimed that a candidate, Catherine Connolly, had withdrawn, attracting hundreds of thousands of views within days.[190]
- During the 2025 German federal election, researchers found that operations such as Operation Overload, Storm-1516 and Doppelgänger were disseminating misleading narratives, using AI-generated content and bot networks, as well as sanctioned outlets such as RT DE, to reach German audiences.[191]
- During Australia's 2025 federal election, a Russian-linked network published numerous fake news stories. These were not intended for a human audience, but for consumption by search engine crawlers used to build AI chatbots, and aimed at producing responses promoting Russian interests.[192]
- In Canada's 2025 federal election, researchers documented a surge of AI-generated political content—including deepfake videos and synthetic imagery—circulating across social media in the final weeks of the campaign.[193]

Survey evidence suggests that perceived exposure to disinformation is increasing across Europe, even where measurable impact on elections remains contested.[194] Moreover, misleading or decontextualised communications can erode confidence in electoral processes even without high levels of belief in specific false claims.[195] This reflects Full Fact's polling: people are very concerned about political misinformation and this is undermining trust in institutions and democratic engagement.[196]

The Electoral Commission works with counterparts to deal with the shared challenge and has said that its *"engagement with countries with shared challenges is key to informing [its] responses to disinformation"*. In 2026, the Commission is hosting a conference with electoral commissions from the other Five Eyes countries.[197] International collaboration is essential given the shared challenges and common threat landscape.

### 3.1.3 Foreign interference and hybrid threats

UK institutions have issued increasingly stark warnings about the risks of foreign interference in democracy, pointing to how information is used strategically.

- Ahead of the general election, in May 2024, the Joint Committee on the National Security Strategy flagged the potential for hostile actors to use deepfakes, fuel conspiracies and undermine trust in UK leaders and institutions.[198]
- In December 2025, the head of MI5 warned that "information, once a unifying force, is increasingly weaponised".[199]
- In January 2026 the Chair of the Foreign Affairs Committee said the UK is "constantly suffering" disinformation campaigns from both state and non-state actors following an inquiry into disinformation diplomacy.[200]
- In his March 2026 Review, Philip Rycroft concluded that we are "already experiencing 'information warfare'. While our political life has not yet been subsumed by this assault, our defences are worryingly weak".[201]

Foreign interference can form part of wider geopolitical strategies, combining information operations with cyber activity, financial influence and diplomatic pressure. These tactics are as much about destabilising as persuading—amplifying confusion, weakening shared understanding and undermining trust in institutions.[202] Recent analysis of Russia-linked information operations illustrates this. Investigations into the so-called 'Pravda' ecosystem have found coordinated practices designed to increase the perceived legitimacy and visibility of the network.[203] Independent monitoring has also documented persistent state-aligned narratives circulating across platforms, with uneven enforcement.[204]

The use of AI in global disinformation campaigns has been increasing rapidly. A report by the EU's External Action Service in March 2026 noted that 27% (147) of the foreign information manipulation and interference incidents it detected in 2025 involved AI. This compares to 41 incidents in 2024—a growth of around 259%.[205]

Hybrid threats further complicate response because foreign-linked narratives can be adapted and amplified by domestic actors. This blurs attribution, weakens legal pathways that depend on foreign involvement, and increases background noise—dynamics that AI-assisted replication can intensify at will. Freedom House has documented how AI tools are increasingly used in coordinated influence and repression strategies, including automated content generation, surveillance-assisted targeting and synthetic persona networks.[206] These capabilities lower the operational costs of sustained information manipulation and complicate attribution during electoral periods.

For the UK, democratic information resilience cannot depend on identifying and countering specific foreign actors, or taking down pieces of content. Instead, preparedness must focus on strengthening system-level resilience in several ways, including: understanding where processes might fail under pressure, improving coordination across institutions, developing a more effective legal framework, ensuring timely and visible public communication, and reducing the extent to which interference can generate uncertainty.

### 3.1.4 Free expression and regulatory pressure

Debates about information governance have, to a significant extent over the past year, been shaped by wider political arguments about free expression and the legitimacy of regulation. UK and EU approaches to digital regulation have come under sustained political scrutiny from the United States, particularly the federal government, which regularly characterises measures aimed at improving platform accountability as censorship or constraints on free speech.[207] Such claims have appeared in official reports,[208] political scrutiny[209] and public statements from senior politicians,[210] as well as criticism from US platforms and industry figures,[211] and been linked to wider political and trade pressures.[212] In September 2025, the US House Committee on the Judiciary held an evidence session about perceived "*censorship*" in UK and EU digital regulations.[213]

All of this creates a very challenging operating environment for UK policymakers and regulators. Questioning the legitimacy of regulation, particularly when framed in terms of fundamental rights, can slow decision-making and reduce public confidence in interventions, making it harder for reliable information to gain visibility.

At the same time, regulatory activity across jurisdictions has intensified. In late 2025, the European Commission said that Meta had breached EU law by using deceptive design to add unnecessary steps to users submitting reports to complain about content or flag it as illegal.[214] Around the same time, the Commission also found—in preliminary findings as part of formal proceedings—both TikTok and Meta in breach of their obligation to grant researchers adequate access to public data under the DSA.[215] Two months later, the Commission fined X €120 million for breach of its transparency obligations under the DSA.[216] In the UK, Ofcom launched an investigation in December 2025 into whether platforms are doing enough to identify and remove illegal terror and hate content,[217] and around a month later launched an investigation into sexual deepfakes produced by X's chatbot Grok AI.[218] These investigations move slowly, but can have an impact.

More recently in the US, a California jury found Meta and Google were to blame for a woman's depression and anxiety following her social media addiction as a child, and awarded her $6 million in damages.[219] In New Mexico, a jury found that Meta violated consumer protection laws and the company was ordered to pay $375 million.[220] These cases have formed part of a bipartisan focus among state attorneys general on the harms of social media to young people, with Republican and Democratic lawmakers supporting proposals for stronger protections and platform accountability in response to growing public concern. They have been pursued through coalitions of state attorneys general, reflecting cross-party legal action against major platforms.[221]

For the UK, maintaining the right balance is key to preparedness. Regulatory frameworks must be robust enough to operate under pressure, while remaining grounded in proportionality, transparency and respect for fundamental rights. Without this clarity, arrangements may struggle to withstand politicised challenges when decisive and trusted action is most needed. Effective governance depends on clear legal mandates, well-defined institutional roles and enforceable transparency obligations. Where these are absent or contested, responses risk becoming inconsistent, delayed or politically vulnerable.

## 3.2 Emerging global standards and safeguards

Democratic information risks operate across borders—with platforms, AI systems and information flows connecting actors and audiences internationally. Consequently, governments, regulators and international bodies develop approaches that address risks at a systemic level. These approaches do not offer a single model for the UK, but provide benchmarks for preparedness: whether institutions can respond quickly under pressure, coordinate effectively, and sustain confidence. This section examines how certain parts of the information environment are addressed in emerging international practice.

### 3.2.1 Systemic risk approaches in practice

A growing body of policy treats democratic information risks as systemic and arising from platform design and system-level processes, rather than individual content. The EU's DSA is the most developed example of this approach. It requires very large online platforms and search engines to assess and mitigate systemic risks arising from the design, functioning and use of their services, including risks to electoral processes.[222] This includes consideration of how recommender systems, advertising and content moderation practices may influence information flows at scale, and requires proportionate mitigation measures. The European Commission has also issued guidelines under the DSA on the mitigation of systemic risks during elections, including enhanced risk assessment.[223]

Strengthening democratic resilience also depends on supporting plural and independent media ecosystems, and actively enhancing the visibility and accessibility of reliable journalism in digital spaces. Some EU member states are exploring measures to promote public interest reporting and ensure it is more easily discoverable alongside other information, reflecting a shift towards state support for trusted sources.[224]

International frameworks reflect broadly aligned priorities. The OECD emphasises transparency, independent scrutiny, institutional capacity and plural media environments as part of a broader information integrity framework.[225] The United Nations[226] and Council of Europe[227] similarly stress freedom of expression, media pluralism and shared responsibility among states, platforms and civil society, while maintaining a focus on human rights and safeguards for open information environments. These approaches show a convergence towards systemic risk-based governance and preparedness.

### 3.2.2 Institutional preparedness and protocols in the Five Eyes countries

Where serious information incidents arise in an election—such as foreign interference, cyber disruption or coordinated disinformation—democratic resilience depends both on operational capability and clear, credible public communication. Among the UK's Five Eyes partners, preparedness combines formal or semi-formal coordination mechanisms with defined approaches to threat assessment and public communication. This section considers those arrangements and how they contrast with the UK's approach.

**Canada: codified public notification protocol during federal elections**
Canada operates a Critical Election Incident Public Protocol.[228] The Protocol applies during federal election periods and empowers a panel of five senior civil servants to determine whether an incident, or series of incidents, threaten the integrity of an election in Canada, or impairs Canadians' ability to have a free and fair election.

Where that threshold is met, the panel can authorise communication to inform the public, and brief political parties and relevant institutions. An independent post-election review assessed the implementation of the Protocol and its effectiveness, and recommended refinements, noting that the *"nature of the threats [to democracy] is evolving"*.[229] Full Fact has been calling for a critical election incident public protocol to be introduced in the UK, based on the model in Canada, since 2021.[230]

The Protocol sets a high bar for public notification, while ensuring that decisions to inform Canadians are taken independently of the government of the day. It operates alongside other organisations, including the Security and Intelligence Threats to Elections Task Force, which brings together security and intelligence agencies to monitor and assess threats. This sits within a broader package of measures to protect Canada's democracy, addressing foreign interference, information manipulation and technological threats.[231]

**Australia: standing Electoral Integrity Assurance Taskforce**

Australia has a standing inter-agency coordination mechanism activated during federal elections—the Electoral Integrity Assurance Taskforce.[232] The Taskforce provides assurance to the Australian Electoral Commissioner that federal elections are unaffected by interference, by monitoring the information environment, sharing information on potential risks with agencies, and advising the Electoral Commission. It brings together bodies including the Department of Home Affairs, the Australian Federal Police, and the Australian Security Intelligence Organisation (ASIO), providing a forum for inter-agency coordination, information sharing and risk assessment relating to federal electoral integrity.

While Australia does not operate a public protocol equivalent to Canada's, institutional roles and coordination structures are clearly defined. The Commission also publishes details of prominent pieces of disinformation it discovers relating to the electoral process, and details of action taken.[233] ASIO publishes an annual public threat assessment, while foreign interference coordination through the Department of Home Affairs interfaces with election security arrangements.[234] Parliamentary committees, including the Joint Committee on Intelligence and Security, have examined these structures.

**New Zealand: publicly articulated coordination principles**

Ahead of the 2023 general election, New Zealand published a document explaining the principles and protocols of the New Zealand Security Intelligence Service and the Government Communications Security Bureau.[235] This was coordinated through the Department of the Prime Minister and Cabinet National Security Group. The arrangements describe how agencies support the protection of New Zealand's democratic processes, including assessing and advising on foreign interference threats.

New Zealand also published a protocol on the management of election disruption, which outlines the approach of the Electoral Commission and other agencies in mitigating threats to the general election process.[236] It published another protocol on communications related to the 2023 general election process, which outlines the roles of government agencies in managing misleading or inaccurate information about the general election.[237]

New Zealand does not operate a public notification protocol like Canada's, but institutional roles in relation to election security and foreign interference are publicly articulated, and agencies publish reporting on threats relevant to democratic institutions.[238]

**United States: critical infrastructure designation and public guidance**

In 2017, the US Department of Homeland Security designated election infrastructure as part of the nation's critical infrastructure.[239] Federal cybersecurity coordination for election infrastructure is led by the Cybersecurity and Infrastructure Security Agency (CISA) within the Department of Homeland Security. CISA supports the Election Infrastructure Government Coordinating Council and works with the Sector Coordinating

Council, which facilitate information sharing and coordination between federal agencies, state election officials and private-sector partners.

CISA publishes accessible guidance for election officials, including incident response resources and communication toolkits.[240] During federal election cycles, it also deploys 'rumor control' webpages designed to rebut or clarify false claims about voting processes and election administration.[241] These materials encourage timely public communication and clarification of inaccurate claims.

The Federal Bureau of Investigation operates a Foreign Influence Task Force, and the Office of the Director of National Intelligence publishes public assessments on foreign threats to US elections.[242] The US system remains decentralised, with much responsibility for election administration resting with states. These coordination mechanisms and guidance documents provide channels for information sharing, incident escalation and public communication during election-related security incidents.

Altogether, these institutional arrangements operate within a highly politicised environment in which concerns about the integrity of the US electoral system have become politically contested. Public debate has been shaped in part by Donald Trump's efforts to overturn the 2020 election result, including attempts to pressure the Justice Department to declare the election "corrupt" despite repeated findings by federal and state officials that there was no evidence of widespread fraud sufficient to change the outcome.[243]

**Comparisons with the UK**

Approaches across Canada, Australia, New Zealand and the United States vary but generally include identifiable coordination mechanisms, designated institutional responsibilities, and a public articulation of how election-related risks are assessed or how communication is handled. Unlike Canada and some other Five Eyes partners, the UK has not published a public protocol or formal escalation framework governing electoral information incidents. In January 2025, a Minister stated that there were no plans to introduce a Canadian-style incident protocol given the processes in place.[244]

However, those processes are shrouded in secrecy. Under the Freedom of Information Act, Full Fact asked both the Ministry of Housing, Communities and Local Government and the Department for Science, Innovation and Technology whether there is a written protocol governing public communication or escalation thresholds for information incidents, including during elections; which body is responsible; who has decision-making authority; and whether there are thresholds. In February 2026, both departments declined to confirm or deny whether they held that information, citing national security considerations.

During a debate on the Representation of the People Bill in April 2026, a Minister argued that "*Broad knowledge about internal protocols, escalation thresholds, command structures or the bodies involved in responding to threats to the integrity of our electoral*

*processes, including through information threats, could expose or enable insights into the UK's security posture, capabilities and response mechanisms."*[245] This focus on secrecy overlooks the importance of visible action, public understanding and trust in the system, and the independence and accountability of those involved. The fact that four in five people in the UK are concerned about political misinformation, and 42% of those say their concern has negatively affected their confidence that elections are free and fair over the past year,[246] raises questions about obscuring the existence of institutional safeguards.

The result is that internal processes—which should offer timely, authoritative public reassurance when electoral integrity is threatened—are not clearly identifiable or visible, with no publicly defined statement of roles, coordination, thresholds or escalation pathways. To be effective, any such mechanism has to gain public trust. As Philip Rycroft noted in his Review, *"The state appears to hold the information it has close; more information should be put into the public domain to help the ordinary citizen to understand what foreign interference looks like and to alert the public to specific instances of it."*[247] Greater transparency about the institutional response to threats would support public understanding and strengthen resilience in the face of evolving risks.

### 3.2.3 Technical safeguards: provenance, labelling, and transparency

Internationally, governments and technology companies are exploring technical approaches to address the risks posed by synthetic and manipulated media. Efforts to develop common standards and disclosure practices are emerging across jurisdictions, but adoption is uneven and fragmented across platforms and sectors.

Measures such as provenance, labelling and transparency systems are being developed to support attribution and verification at scale, as synthetic and manipulated content becomes easier to produce and distribute. These tools provide signals about origin, modification or context, rather than assessing accuracy or intent. Provenance and labelling systems aim to indicate how content was created or altered, particularly for audio, image and video. While they can support faster authenticity checks, in principle, their effectiveness is constrained in practice. Metadata can be removed, adoption is uneven, and impact depends on consistent design, interoperability, prominence and user understanding.

Transparency mechanisms operate differently. Reporting requirements, policy disclosures and researcher access enable oversight of how information systems function and where systemic risks arise. They are valuable for identifying patterns and vulnerabilities ahead of high-risk periods, but offer limited protection once incidents are underway.

Many technical safeguards rely on voluntary or hybrid governance. Initiatives such as the Coalition for Content Provenance and Authenticity (C2PA)[248] and the EU's Code of Practice for AI-generated content promote interoperability and disclosure.[249] While C2PA has made impressive progress in securing voluntary participation from a wide range of companies, from camera and phone manufacturers to AI and social media companies,[250] the effectiveness of this system depends on implementation that translates between different companies' systems, enforcement and public comprehension.

Full Fact's editorial team frequently identifies SynthID as part of our fact checking. This is an invisible watermark that appears in content created or altered with Google's AI tools. It can be detected even if the image is cropped and altered in other ways. Unlike many AI detectors, which do not consistently identify AI-generated content, SynthID is a good example of an AI standard. In May 2026, Google announced that it was working with several companies to add SynthID to their systems, including OpenAI and Nvidia.[251]

Full Fact has built an annotated dataset of 238 fact checks published since 2023 involving suspected AI-generated or manipulated content. This has risen from 10 fact checks in 2023 to 137 in 2025, and 68 for January to April 2026 (theoretically meaning we would write more than 200 fact checks about suspected AI content in 2026 if the current rate is maintained). Our initial analysis shows that just 66 pieces of content contained a visible or invisible watermark. The strongest indicators that help us assess whether something is AI-generated are watermarks; accounts with a track record of sharing AI-generated, digital or questionable content; expert analysis and confirmation from the creator either publicly or as part of our right of reply process. Given the strong reliance on human judgement and relationship credibility involved in fact checking suspected AI-generated content, we are sceptical that deepfake detection systems can operate successfully at scale.

Recent initiatives in the UK indicate growing focus on technical capability. In January 2026, the Home Office ran a challenge for software to detect deepfakes.[252] In February 2026, the government announced a collaboration with Microsoft and others to develop and implement a *"deepfake detection evaluation framework"*, to assess, understand and detect harmful deepfake materials. Once established, this will be used to set expectations for industries on deepfake detection standards.[253] The effectiveness of the Electoral Commission's deepfake detection pilot, noted above, will become clearer following a planned review of the project.[254]

The Electoral Commission has previously called for generative AI material to be labelled,[255] and recommended that platforms require labelling of AI-modified content during election periods.[256] In March 2026, the Secretary of State for Science Innovation and Technology announced a taskforce to explore proposals for government on best practice for labelling AI-generated content, with an interim report due in autumn 2026.[257]

Ofcom's Attribution Toolkit finds that layered approaches are most effective, combining watermarking, provenance metadata, AI labelling, and context annotations—providing users with additional information about a piece of content. However, none of this reflects the way the vast majority of users interact with new technology on a daily basis, and partial or inconsistent implementation risks creating the appearance of protection without strengthening public confidence.[258]

Technical safeguards are therefore enabling but insufficient on their own. They can reduce ambiguity and support institutional response, but cannot replace clear governance, coordination or trusted public communication. Used in combination with institutional protocols, they can reinforce authoritative signals; used in isolation, they risk overstating the capacity of technical solutions to address social and institutional challenges.

GUEST ESSAY:

## Layered trust signals, not silver bullets: what Iran shows the UK about AI and public trust

**Mahsa Alimardani, Associate Director, Technology Threats and Opportunities, WITNESS**

AI is now manufacturing ambient uncertainty about everything we see and hear, and that uncertainty has become a permission structure for not knowing what to believe. The information environment is now facing something far more complex than simply synthetic content deceiving audiences: the routine dismissal of real content as fake, alongside the routine acceptance of fabricated content as real. Before the next election and well beyond it, the UK needs a layered set of trust signals across the content lifecycle.

The failure mode threatening our information environment is the AI binary itself. "Is this AI or not?" is the wrong question, because it collapses two operations that have very different relationships to truth. AI generation can produce completely fictional audio-visual content. AI enhancement or editing can alter an authentic photograph of a real person or event, captured by a real camera, for example, by smoothing skin, stylising a background, or sharpening detail. AI editing tools are increasingly built into editing software used in everyday journalism and personal use, meaning more authentic images may carry traces of AI. The category of "AI-enhanced real photograph" is no longer exotic, and the analytical work of distinguishing enhancement from generation is getting harder and more necessary.

Iran shows what happens when the AI binary is allowed to do the analytical work on an industrial scale. Researchers have long called the underlying dynamic the liar's dividend: the benefit that accrues to bad actors when they can dismiss inconvenient content as fake, because the existence of synthetic material makes the dismissal plausible.[259] This was captured during the January 2026 protests in Iran, when an authentic, verified image of a protester standing up to armed security personnel went viral.[260] But those affiliated with the regime accused the image of being AI slop, casting doubt on protest documentation overall.

This dynamic is not new to Iran, but the scale is. The sheer volume of AI content circulating during the protest crisis and conflict since February has turned the liar's dividend from a tactic available to actors—especially the state—into the ambient condition of the information environment. AI has arrived to make dismissal cheap and deception easy within an engineered information vacuum, a hot war, and a polarised cross-border ecosystem. Faced with authentic and synthetic content side by side, audiences do not become better at distinguishing between them. They withdraw.

The February 2026 US strike on Iran's Minab girls' school killed an estimated 175 children, one of the highest civilian death tolls of the conflict.[261] Footage from outside the school was dismissed within hours: Grok falsely identified it as imagery from a 2021 Kabul bombing, citing fabricated sources. Days later, when the Iranian foreign minister tweeted photographs of the mass burial, opposition and diaspora accounts dismissed those as AI-generated, too. Despite credible verification work, the dismissals stuck. There are Iranians who, to this day, do not believe the strike happened. Harrowingly, these dismissals of civilian casualties contributed to life and death decisions people made amidst bombs about choosing to evacuate or remain in danger.

Provenance is one of the trust signals that can help. It does not adjudicate truth; it provides the recipe for how content is made, such as which device captured it and what, or how a piece of content is edited. C2PA Content Credentials are the most mature open standard for this layer.[262] If it is adopted across the entire information environment, it is meant to be interoperable, cryptographically secure, and to have privacy built in by design. Content credentials are not a complete answer, and much more work is needed to support their implementation and design (which will require investment by private tech companies and governments). Content credentials need to work alongside other signals, such as transparent post-hoc detection methods that include labelling and audience literacy. No single part substitutes for the others, but together these layers slow the acceleration of harm.

Governments are now testing different versions of AI transparency. California is building the most fully formed pipeline, with staged obligations on AI providers, platforms, and capture device manufacturers anchored to open standards.[263] India shows how getting it wrong can drive over-removal of legitimate content.[264] The European Union is at the last development layer of a code of practice on Transparency for developers and deployers that sets forward the importance of these trust signals and implements an EU-wide icon for AI-generated and manipulated content.[265]

Three principles drawn from this landscape should anchor the UK's approach. **First, distribute responsibility across the content lifecycle.** Asking platforms to detect and adjudicate synthetic content at the point of distribution, after the fact, sets them up to over-remove or under-remove. Pipeline obligations on AI providers, hosting platforms and capture device manufacturers, anchored to open standards, are the right direction for legislation.

**Second, build rights protections from the outset.** Provenance infrastructure carries risks for vulnerable users if it embeds personal information without consent, or hands governance to private standards bodies without accountability. Informed-consent standards for personal data in credentials, governance conditions on the standards bodies given legal effect, and rulemaking authority for the regulator to keep pace with the technology are not optional refinements; they are what make the system rights-respecting rather than a surveillance vector.

**Third, extend trust signals beyond political content.** The AI binary is now being weaponised against authentic journalism and citizen content, not only campaign material. In the Iran-Israel war, photojournalist Erfan Kouchari's authentic images of the aftermath of Israeli strikes were attacked online as AI-generated, with fabricated detector heatmaps circulated as forensic-looking evidence.[266] The editing software photojournalists routinely use already supports C2PA Content Credentials. Had wire services and newsrooms required their use, the edit history of those images would have been verifiable. The Meta Oversight Board has called on Meta to implement Content Credentials at scale.[267] The UK, with its public service broadcasting sector and the BBC as an active member of C2PA, is well placed to make this move in journalism.

The lesson from Iran is not that synthetic content has overwhelmed the truth. In the absence of trust signals across the content lifecycle, the question of provenance becomes unanswerable, and ambient uncertainty does the work of denial for free.

### 3.2.4 Political advertising transparency

Political advertising sits at the intersection of electoral law, platform governance and public accountability. Transparency determines how visible campaigning practices are when scrutiny is most needed. Where disclosure relies on voluntary or platform-led measures, that visibility tends to narrow over time, weakening democratic oversight.

In the final week of the 2024 election campaign, UK political parties were collectively spending £1,313,442 on Meta advertising.[268] But despite its scale and significance, digital political advertising in the UK is characterised by fragmentation and limited transparency. There are no statutory requirements for platforms to disclose targeting data or maintain public archives of political adverts in the UK. The reliance on self-regulation leaves gaps in disclosure, targeting transparency, accountability and long-term data retention.[269]

Political advertising information is of interest to voters and researchers, now and in the future. But advertising libraries vary in scope and reliability, and can be altered without oversight.[270] Evidence suggests that available data is often incomplete, while time-limited retention restricts long-term scrutiny and learning. For democratic accountability and historical research, this material should be preserved in the UK's public record.[271]

Across a number of democracies, political advertising transparency is being embedded in regulatory frameworks rather than left to voluntary platform disclosure. In particular, the EU introduced binding requirements on the labelling of political adverts, disclosure of funding sources and spend, transparency over targeting techniques, and improved public access to advertising data through a public repository.[272] Some large platforms have withdrawn political advertising altogether when transparency requirements increased—including in Canada[273] and the EU.[274] Critics note that these withdrawals mean *"the firms are not, collectively, living up to their stated objectives (to wider society) of uplifting people's free speech rights in a way that promotes transparency and accountability for all."*[275] Full Fact has called for a comprehensive public library of political adverts in the UK, and for platforms to support compliance rather than withdraw from hosting political adverts.[276]

A comprehensive, publicly accessible library—including sponsor, spend and targeting data—would strengthen democratic resilience. Similar proposals have been considered by a raft of reviews and institutions, with some focusing on more consistent platform libraries, including the Electoral Commission, [277] Committee on Standards in Public Life[278] parliamentary committees[279] and Philip Rycroft.[280] There is no excuse for not having a robust system in place before the next general election. Without this, a significant and growing part of the campaigning environment remains only partially visible, limiting scrutiny during elections and weakening accountability over time.

# Part 4: UK governance, regulation and institutional preparedness

This report has shown how information systems generate misinformation and persistent uncertainty, the harms this creates, and how other democracies are responding. This section assesses whether the UK's legal and institutional arrangements are equipped to manage the risks in practice—particularly during high-pressure events. It examines how institutional responsibilities are distributed, how frameworks operate, and where gaps in laws, coordination, scope and visibility may undermine preparedness.

## 4.1 Institutional gaps for information resilience

The UK's institutional approach to information risks is distributed across multiple public authorities spanning elections, media regulation, national security, public communication, and online harms. This reflects the cross-cutting nature of the challenge—but it also means that no single body is responsible for system-wide coherence across the UK's democratic information environment, nor for providing a visible, clearly accountable centre for coordinating responses to serious information incidents.

Current arrangements for handling serious information risks are opaque. There is no public articulation of the thresholds for activation and escalation pathways, the processes for coordination and communication responsibilities, and even the work and accountabilities of some of the key bodies involved. There is also no published framework that distinguishes emerging information incidents (such as early signs of election interference or misinformation spikes) from full-scale crises, or that sets out how responses should scale between them to ensure a proportionate and effective approach.

This opacity has consequences. It creates ambiguity during high-pressure events. It may be unclear to the public which bodies are leading a response, who speaks authoritatively, when action is being taken, or how decisions are reached. This can cause confusion, lead to speculation and weaken trust, where rapid, authoritative communication is most needed.[281] Information incidents cut across organisational boundaries; without predefined coordination and clear public-facing authority, there is a risk that institutional responses will fail to address them quickly, or stabilise the information environment in the long-term.

The UK's crisis management structures were refreshed through the 2025 revision of the Amber Book, which provides a framework for cross-government crisis preparedness and coordination.[282] This doctrine is primarily oriented towards operational crisis response. However, it does not set out a dedicated approach to democratic information resilience, including in relation to emerging risks. As a result, this area is not clearly integrated into

national resilience planning and crisis preparedness. The gap is reflected in the absence of a published protocol for electoral information incidents, as set out above.

In his 2026 Review, Philip Rycroft noted (in relation to countering hostile state online interference): *"Responsibilities are dispersed across different departments and agencies with no apparent focal point."*[283] A similar lack of coherence applies to the government's management of democratic information risks and other acute information incidents. The fragmentation reflects a deeper lack of clear institutional ownership over information resilience as a coherent policy domain. Commentators on national security have argued that the absence of a single, empowered institution dedicated to the issue leaves the UK vulnerable to systemic threats, just as the cyber landscape lacked a central coordinating body until the creation of the National Cyber Security Centre.[284]

> *Because the perceived threat to our democracy has been relatively muted over the decades, central coordination of this complex system has been light touch. An official team has come together at election times in the Joint Election Security and Preparedness Unit to monitor the electoral process. Otherwise there has been no central official apparatus to consider the health of the system as a whole over time.*[285]
>
> **Philip Rycroft, the Rycroft Review into countering foreign financial influence and interference in UK politics, 25 March 2026**

### 4.1.1 Mapping responsibilities in practice

Various UK government departments, agencies and regulators have some role in information governance, with responsibilities distributed across different policy areas, operational functions and regulatory frameworks.

- **The Department for Science, Innovation and Technology (DSIT)** is responsible for UK policy on disinformation aimed at UK users online,[286] working with the **National Security Secretariat in the Cabinet Office** and others.
- The **Home Office** plays a leading role in the government's response to domestic state threats, while the **Foreign, Commonwealth and Development Office** is responsible for understanding and addressing information threats overseas.
- The **Ministry of Housing, Communities and Local Government** (MHCLG) leads on electoral law and policy and is sponsoring the Representation of the People Bill, a package of electoral reforms considered further below.[287]

- The **National Security Online Information Team** (NSOIT) leads DSIT's *"operational response to information threats focussing on tackling the greatest national security and public safety risks facing the UK from mis- and disinformation."*[288] Operational activity appears to focus on notifying social media platforms to take action at their discretion on content that may breach their moderation policies.[289]
- The **Joint Election Security Preparedness Unit** (JESP) operates between the Cabinet Office and MHCLG to *"...coordinate election security and preparedness activity within government and externally."*[290] JESP coordinates work across government to respond to issues as they emerge, including misinformation and disinformation.[291]
- JESP stands up an **Election Cell** ahead of major democratic events, such as general elections, bringing together departments, intelligence agencies and others[292] to monitor and respond to emerging issues, including information incidents.[293]
- The **Defending Democracy Taskforce** was established in 2022 "*to coordinate and drive progress on the Government's work to protect UK democratic processes, institutions and society*".[294] It provides a cross-government forum focused on protecting democratic processes, and bridges gaps between the national security establishment and others, including large tech companies.[295] [296]
- The **Government Communication Service** is the professional function for communication across government, and supports coordination of public communications and the delivery of campaigns.
- These bodies operate alongside the **intelligence and security agencies** and the **National Cyber Security Centre**, which contribute to electoral security from a national security perspective and bring technical capabilities.
- **The National Situation Centre** was established in 2021 and is, according to the Government Analysis Function *"the main body responsible for coordinating data and analysis for use in crisis contexts."*[297]
- Regulators also play a role. **Ofcom** enforces the OSA and is advised by the **Online Information Advisory Committee** on matters relating to misinformation and disinformation.[298] However, the OSA addresses misinformation only indirectly, limiting Ofcom's role in managing broader democratic information risks.[299]
- The **Information Commissioner's Office** regulates data protection and information rights, with responsibility for overseeing the use of personal data. It plays a role in addressing relevant risks, including data misuse and microtargeting.
- The **Electoral Commission** has highlighted the limited scope of its regulatory responsibilities when it comes to tackling election-related misinformation and disinformation.[300] Its Corporate Plan emphasises that "*voters are increasingly exposed to mis- and dis-information—a trend which looks set to continue—and dubious campaign practices continue to undermine trust*".[301]

There is little public visibility of how this system safeguards democratic information resilience, with details about some of these bodies often only published in response to enquiries from MPs or parliamentary committees. Transparency gaps are particularly acute for bodies such as NSOIT, JESP, the Election Cell, the National Situation Centre and the Defending Democracy Taskforce, with scant public detail about their responsibilities, decision-making processes, coordination processes, or accountabilities.

Legitimacy depends on public trust in the system, and understanding who is responsible for responding to information risks and how those responses operate in practice. Opacity or uncertainty about institutional responsibilities can weaken trust at times when rapid, authoritative communication is most needed. The polling by Full Fact,[302] and polling by the Electoral Commission outlined above,[303] underline the public's concern that the government is not doing enough to tackle misinformation and disinformation.

### 4.1.2 Institutional readiness and incident coordination

Recent incidents and policy developments suggest growing recognition with government and regulators that information incidents require more coordinated operational responses. However, these reforms are limited by the absence of a broader institutional framework for democratic information resilience, and by the underlying legal framework.

In a review of the police response to the 2024 riots, His Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) found that liaison arrangements between law enforcement, government and online service providers were not properly established or understood.[304] This highlighted the practical consequences of unclear coordination structures during fast-moving information incidents.

In 2025, Ofcom consulted on additional safety measures for online services, including proposals requiring platforms to maintain crisis protocols.[305] This was advanced partly in response to the 2024 post-Southport riots, where harmful misinformation spread rapidly across platforms, contributing to violence and public disorder offline. Ofcom's proposal was a welcome step toward improving how platforms manage information incidents. However, as Full Fact's evidence to Ofcom highlighted, it needs to go much further, with more measures needed for effective crisis management.[306] The proposals are also constrained by the OSA, which does not require platforms to address broader systemic risks.

The government's social cohesion strategy, *Protecting What Matters*, published in March 2026, contained welcome commitments to strengthen the UK's response to crises, which will build on Ofcom's proposals. This includes convening civil society, experts, and platforms to provide real-world insights during crises. It also includes reviewing the crisis powers in the OSA, which will mean looking at giving trustworthy media

due prominence *"so people have access to authoritative sources to counter mis and disinformation"*.[307] That review should consider the absence of an institutional anchor to manage democratic information resilience. Taskforces and temporary cells do not provide durable oversight, cumulative learning, accountability to the public and Parliament, or a framework for coordination between public authorities, online companies and others.

In 2021, Full Fact published its Framework for Information Incidents, which set out a severity-based approach for responding to misinformation and other destabilising content. That framework identifies clear thresholds for action, assigns responsibility for coordination, and maps communication pathways to ensure rapid, authoritative public clarification.[308] It was developed prior to President Trump's second term; it is clearer now how a political party can affect a country's information ecosystem. Policymakers should consider how this situation might have played out in the UK, a country where so many political integrity standards are informal norms and non-statutory commitments.

The Rycroft Review highlighted the importance of accountability at senior official and ministerial levels, particularly in responding to foreign interference and election-related threats. Rycroft recommended allocating a Permanent Secretary *"lead responsibility for sustaining our democracy and coordinating the response to the threats to it"*.[309] Elsewhere, the think tank Demos has recommended setting up a working group or crisis response committee that would sit across government and review crisis preparedness approaches.[310] Part 5 of this report proposes a dedicated coordination function to embed a national framework for information incidents and coordinate operational activity.

## 4.2 Legal gaps and ambiguity

Misinformation is the most commonly encountered online harm in the UK.[311] It is the most severe global risk according to international expert bodies, a matter of serious concern to the Prime Minister, and independent experts and scrutiny bodies have repeatedly warned of the risks posed by hostile state and non-state actors in the UK.[312] In 2024, it was the UK public's second biggest concern in elections (after media bias).[313] Yet the UK's legal framework does not address it in a coherent way, for elections or otherwise.

The gap has been recognised by international experts. In its Public Communication Scan of the UK, in December 2023, the OECD highlighted *"a noteworthy gap in the legislative and policy landscape [...] on mis- and disinformation in the context of elections"*.[314] Laws are dispersed across multiple legal regimes—electoral, online safety, media regulation and national security—each designed for different purposes and operating separately. They do not form a coherent framework for managing democratic information risk. The result is a fragmented legal landscape that lacks clarity of purpose, consistent coverage and mechanisms for coordination. This limits both the effectiveness of interventions and public confidence in the system's ability to respond, and means that significant areas of democratic information risk fall outside of regulatory oversight.

### 4.2.2 Electoral law

UK electoral law was designed for a different information environment. Offences and transparency requirements reflect assumptions of identifiable publishers, physical campaign materials and broadcast-era media. They also assume slower dissemination, clearer attribution and limited automation. These no longer apply in campaigns characterised by algorithmic amplification, microtargeted advertising and AI-generated content—where claims can scale rapidly and attribution is fragmented.

In 2025, the Public Administration and Constitutional Affairs Committee said electoral law and policy "*has struggled to keep pace*" with misinformation and disinformation risks.[315] As AI-mediated systems and platform design increasingly shape political information flows, ambiguity around scope, timing, enforcement and transparency increases the likelihood that uncertainty will persist during elections—even where accurate information exists. The Representation of the People Bill provides an opportunity to modernise these safeguards; but as currently drafted, it does not address the most significant risks.[316]

The government's July 2025 policy paper, which formed the basis for the Bill, acknowledged that "*Our own democracy is being threatened by misinformation*" and proposed measures to strengthen transparency and integrity in election campaigning.[317] However, two measures that would have helped to address misleading campaign practices—a new code of conduct for campaigning and enhanced requirements for campaign material to identify affiliated political entities—were not included in the Bill as introduced.

The Bill also contained no measures to address the use of AI-generated or manipulated political content. While the OSA created offences for sharing and threatening to share sexually explicit deepfakes,[318] and the Data (Use and Access) Act 2025 criminalised the creation and sharing of sexually explicit deepfakes,[319] political deepfakes had not been a focus for policymakers. The Electoral Commission has previously suggested that generative AI material should be clearly labelled, particularly in election periods.[320] There is, however, no requirement for parties and campaigners to label AI-generated material. If nothing changes, there is a significant risk that the law will be ill-equipped to deal with the challenges posed by AI in the run-up to the next election.

The offence of making false statements about a candidate has an unclear scope in relation to synthetic media. This reduces both clarity and deterrence in fast-moving contexts.[321] The Electoral Commission[322] and Speaker's Conference[323] both called for the offence to be updated in light of technological change. In late 2025, a deepfake video depicting George Freeman MP falsely announced his defection to Reform UK. This did not meet the test for a 'false communications offence' under the OSA.[324] The Electoral Commission highlighted that the police could not prosecute it as a 'false statements about candidates' offence because it fell outside the regulated period.[325] The result was a high-profile falsehood about a prominent politician with no route for enforcement—a visible accountability gap.

The government has since expressed its concern about the impact of misinformation and disinformation on candidates, MPs and UK elections, and is considering the need to update electoral offences to explicitly capture deepfakes.[326] While this commitment is welcome, it should form part of a wider review of democratic information resilience, and be part of a proactive approach rather than an ad hoc reaction to incidents.

> *This Government shares the conference's concern about the impact of mis- and disinformation, including the role of deepfakes, on candidates, MPs, and UK elections and is committed to addressing it.*[327]
>
> **Government response to the reports of the Speaker's Conference on the security of MPs, candidates and elections, 5 March 2026**

As examined in section 3, the UK does not operate a comprehensive public repository of political adverts. An amendment to the Bill was debated by the Bill Committee in April 2026, which would have established a repository. In response, the Minister said "*This is a complex area, and the implications need detailed consideration to avoid replicating the unintended consequences seen in other jurisdictions*", where platforms have withdrawn from hosting political adverts.[328] However, the danger of failing to act is that powerful platforms will be left to regulate themselves on issues of national democratic importance.

Institutional powers also lag behind the realities of digital campaigning. The Electoral Commission lacks general powers to obtain information from individuals and organisations, including online platforms, outside of a formal investigation.[329] This constrains its ability to obtain time-sensitive data during electoral campaigns. While the Bill strengthens the Commission's information-sharing powers with other regulators and law enforcement, it does not create an information-gathering power. As campaigning becomes more data-driven and platform-mediated, this gap becomes more significant.

The Bill is also silent on the risk of algorithmic political bias. In evidence to the Foreign Affairs Committee in January 2026, the Chief Executive of the Electoral Commission said that if a social media company were to preferentially amplify some political content and suppress other content in a politically biased way, the UK's legislative toolkit would not enable regulatory action.[330] Electoral law contains no provisions addressing platform-level promotion, suppression or ranking of political content, even where such conduct could materially affect the visibility of candidates or campaigns. Since taking over X, Elon Musk has transformed the platform into an algorithmic microphone for his own views and interests. This, combined with Mr Musk's zeal for getting involved in other countries' democratic processes and debates, clearly illustrates the risks in this area.[331]

During the first debate on the Bill, on 2 March 2026, multiple MPs on a cross-party basis highlighted the Bill's failure to tackle misinformation and disinformation, and the need for new measures. In response, the Minister acknowledged that misinformation and disinformation "*...needs to be addressed more forcefully.*" During Full Fact's oral evidence to the Public Bill Committee on 18 March 2026, the Minister noted that "*technology moves at breakneck speed and takes us forward, and it has been recognised that our electoral system is not keeping pace with it*". However, various amendments to the Bill were rejected at Committee Stage as out of scope, including those relating to the regulation of online service providers. As a result, key issues at the intersection of elections and the online environment are unlikely to be debated in the context of the Bill.

### Full Fact's recommendations for the Representation of the People Bill

When it was introduced to Parliament in February 2026, the Representation of the People Bill fell short of addressing the growing threats of misinformation and disinformation. Full Fact is advocating measures that would substantially improve the Bill, strengthen the foundations of UK democracy and help restore trust in politics.

1. **Upgrade the Online Safety Act to safeguard the UK's democracy.**
2. **Create stronger rules to deal with political deepfakes.**
3. **Establish a comprehensive public library of political adverts.**
4. **Regulate to prevent misinformation and disinformation in political campaigns.**
5. **Create a transparent system for dealing with electoral information incidents.**
6. **Increase the investigative powers of the Electoral Commission.**
7. **Give platforms a statutory duty to support effective media and political literacy.**

More details are available in our policy paper[332] and parliamentary briefings.[333]

## 4.2.3 Online safety law

The Online Safety Act 2023 established a framework for regulating online platforms and search engines, focused on categories of illegal content and harms to children. In March 2021, the then Prime Minister said the OSA would tackle collective online harms, including threats to democracy.[334] The then Government acknowledged "*that misinformation and disinformation surrounding elections are a risk to democracy and it is vital to address this issue*".[335] But ultimately, the OSA failed to establish a duty on platforms to mitigate and assess systemic harms, such as risks to elections, public health or public safety.[336]

The OSA focuses on individual instances of illegality rather than systemic effects. In practice, this means misinformation is only regulated where it intersects with certain criminal offences (for example, misinformation that also stirs up racial hatred) or certain harms to children.[337] Most political misinformation—including widely amplified and harmful, but lawful, misleading claims—are outside the OSA's scope.[338] In a letter to online service providers ahead of the May 2026 elections, Ofcom noted that the OSA *"does not explicitly identify misinformation or disinformation as specific harms that need to be addressed"*.[339]

There appears to be some ambiguity in the government's interpretation of the OSA. In September 2025, a Minister said the OSA introduced strong protections against illegal content which *"also includes election related offences"* such as the offence of false statements about candidates and undue influence relating to elections.[340] However, platforms are only required to take the preventative measures set out in Ofcom's Codes of Practice—due to the 'safe harbour' provisions in the OSA—and these offences are not in the OSA priority list or in Ofcom's Illegal Content Judgments Guidance. They are also not subject to the additional duties platforms need to take in relation to priority offences.[341]

The SIT Committee's 2025 inquiry concluded that the OSA *"cannot keep the UK public safe as it was not designed to tackle misinformation"*.[342] The committee highlighted that the OSA does not address recommendation systems that systematically amplify lawful but harmful content and that even full enforcement "*would have made little difference*" where platform design drives reach and engagement.[343] In its 2024 review of the policing response to Southport, HMICFRS similarly concluded: *"Given the approach taken in the Act, we question how effective the Act and regulation by Ofcom will be in the context of rapidly spreading disorder provoked by online content."*[344]

While the OSA introduced transparency reporting requirements to Ofcom for the largest services, again, these are not focused on systemic harms. There is no systematic framework for scrutinising how platforms prioritise, rank or amplify political information, nor for assessing the impact of these systems on UK elections and democracy.[345]

Implementation of the OSA in 2025—particularly requirements for age verification checks—led to media and public concern and parliamentary debate, including a Commons debate triggered by an e-petition calling for its repeal.[346] The government rejected the repeal and reaffirmed its commitment to implementing the OSA. Full Fact's position has been that the appropriate response is to strengthen the way the OSA deals with misinformation and democratic harms, rather than to repeal it.[347]

In March 2025, the government announced a consultation into children's online safety, *Growing up in the Online World*.[348] Alongside this consultation, the Children's Wellbeing and Schools Bill made its way through Parliament. As a result of pressure from the

Lords, the government agreed to introduce regulations following its consultation within a set timeframe, and to consider the effect of harmful features and product design on children's online experiences. It is unclear what this will look like in practice, although a Minister, Olivia Bailey, said it would involve *"some form of age or functionality restrictions"*.[349]

### 4.2.4 Data access law

Effective oversight of democratic information risk depends on independent researchers having visibility of platform behaviour at scale. Meaningful access to data is needed to understand how false or misleading information is amplified, coordinated and distributed in practice. In their 2025 report, the SIT Committee noted that a lack of access to platform data made it difficult for researchers to assess online safety.[350] The Electoral Commission also said in 2025 that *"Without robust access to platform data and to how social media algorithms work, regulators and researchers cannot effectively identify patterns of abuse, intimidation, and misinformation."*[351]

Without structured data access rights, oversight depends on voluntary cooperation rather than enforceable obligations.[352] The Data (Use and Access) Act 2025 gave the government the power to set regulations requiring platforms to provide independent researchers with access to data for online safety research.[353] The impact of this will depend on how the powers are interpreted and implemented in practice, including what data can be shared, under what safeguards, and with what resources. But the 2025 Act leaves uncertainty about whether independent researchers will be able to observe systemic information risks. Sensible regulation in this area is crucial and would make a huge difference.

Ofcom has stated that no single model of researcher access is likely to meet the full range of researchers' needs, and proposed three models: clarifying existing legal rules on access, creating new duties for platforms to give access, giving legal powers to a third party to act as an intermediary.[354] The effectiveness of the new regime will depend on whether it delivers consistent, proportionate and scalable access that enables independent scrutiny of platform systems, rather than ad hoc or case-specific disclosure.

Experience in the EU suggests that access rights do not necessarily translate to meaningful access in practice, and are insufficient without clear standards, enforcement and technical infrastructure. Without these, data access risks becoming partial, delayed or unusable for real-time or systemic analysis. Survey work by the European Digital Media Observatory, published in May 2024, found that access to data is unfulfilled; support for fact checking, research and media literacy communities is not adequate; and that there is patchy and insufficient progress towards integrating expert feedback into platforms' tools.[355]

### 4.2.5 National security law

The National Security Act 2023 contains tools to address hostile state activity, including espionage and foreign interference. Its measures are reactive and depend on attribution, requiring proof of intent, foreign state involvement and prohibited conduct, and a high threshold of evidence. The Act included a new 'foreign interference' offence but there are practical challenges to enforcement—particularly the need for proof and attributing interference to a foreign power, which may have concealed its involvement[356]—and the law does not address domestic political deepfakes.[357]

These features may be appropriate for intelligence-led investigations and national security purposes, but they mean the Act engages late in the lifecycle of information incidents and only in a narrow set of cases. By contrast, much democratic information risk unfolds before illegality can be established or without an offence having been committed, often without clear attribution. This is particularly the case in a fast-moving situation like an election campaign or public crisis. National security law focuses on proven foreign involvement or hostile intent, but the most consequential risks typically arise earlier, during periods of uncertainty, when rapid clarification and coordinated response are critical.

Current law provides no framework for managing these moments—coordinating disclosure, engaging platforms, or communicating effectively with the public—leaving a gap between intelligence-led enforcement and the real-time management of information threats. As Philip Rycroft identified, *"much of the damage done by hostile state interference online is in the moment and needs to be dealt with in the moment"*.[358] Intelligence-led responses necessarily prioritise confidentiality, but this restricts their ability to provide visible reassurance or public clarification. These tools are poorly suited to visibly resolving uncertainty at scale or sustaining public confidence in institutional response.

The blurred lines between foreign and domestic threats create further challenges. Philip Rycroft noted that the government's work is divided between the foreign and domestic spheres, a distinction *"that is almost completely irrelevant in dealing with this problem"*. This was illustrated in March 2026, when Meta told Parliament's Foreign Affairs Committee that only half of the coordinated inauthentic networks it removes are foreign influence operations—that is, originating in one country and targeting another.[359]

### 4.2.6 The lack of AI governance law

The UK does not have a dedicated statutory framework governing general-purpose AI, despite these systems increasingly acting as key intermediaries for information. Current governance relies on principles and sectoral regimes that were not designed for rapidly evolving, cross-domain AI systems, or the potential risks they bring.[360]

AI-mediated risks relevant to democracy—including realistic synthetic content, automated influence, and misleading AI-generated summaries and search outputs—are not properly addressed with existing laws. No framework is tailored to the role these systems play in shaping how information is produced, prioritised and consumed. Current debates about governance focus largely on content moderation, copyright and consumer protection. However, evidence that AI systems may be susceptible to targeted data manipulation highlights the need for upstream transparency and testing standards.[361]

Labour's 2024 manifesto committed to binding regulation of the most powerful AI models.[362] The government has since emphasised a preference for flexible, pro-innovation governance and close partnership with frontier AI developers.[363] While this approach seeks to retain flexibility, and avoid regulatory overreach in a fast-moving field, it fails to address the UK's information resilience. The government has also sought to close regulatory gaps for AI systems incrementally, and broadly in line with duties under the OSA, rather than approaching it holistically to tackle democratic and other risks.

Ofcom confirmed in 2025 that AI chatbots are outside the scope of the OSA if they do not search websites or databases when responding to users, do not enable interactions with other users, and do not generate pornographic content.[364] Its investigation into the X chatbot Grok illustrates how the OSA's scope is being interpreted in practice.[365] The Crime and Policing Act 2026 enabled the government to amend the OSA to cover more types of generative AI tools.[366] But this extension of the OSA focuses narrowly on potential harms to individuals from illegal AI-generated content and the use of AI services for certain offences.[367] As a result, general-purpose AI systems continue to operate without statutory duties tailored to their information-mediating role or systemic risks.[368] The government has floated the idea of regular Budget-style debates on online safety, implicitly acknowledging that the OSA is not futureproof not long after it was implemented.[369]

Evidence produced by the government in 2025 highlighted that "*The capabilities of systems using AI have been advancing rapidly*" leading to increased awareness of current harms and future potential risks.[370] This underscores a widening gap between the pace of change and underdeveloped legal and regulatory frameworks. The UK's AI Safety Institute analysis of AI trends provides a foundation for the government to assess risk, but evaluation and oversight are in development alongside the capabilities they are designed to assess.[371]

The Ada Lovelace Institute has found strong support for independent regulation of AI systems, saying: "*the public are considerably more comfortable with mandatory safety checks than with voluntary ones*". It also noted: "*This aligns with evidence showing that voluntary safety commitments often fail to ensure adequate protections, reinforcing the case for mandatory requirements to safeguard public trust and wellbeing.*"[372] Parliamentary debates in late 2025 and early 2026 also reflected cross-party concern that safeguards for the most powerful AI systems lag behind technical development.[373]

GUEST ESSAY:

## Language, trust and the digital public sphere

**Baroness Kidron OBE, Crossbench Peer, filmmaker and author**

Language matters. It creates common understanding. It allows conversation across time, distance and difference—from one generation to another, from one life to another. Words allow us to communicate experience, establish trust, argue, persuade, dissent and participate in public life. Shared language is not a decorative feature of democratic society; it is part of its infrastructure.

Over the past two decades, the digital world changed not just how language travels, but what language is for.

Communication increasingly operates not to inform, understand or connect, but to capture attention, drive engagement and shape behaviour. Almost everything we encounter online is promoted, amplified, targeted or suppressed for a commercial, political or ideological reason. The challenge is not simply to identify "misinformation"—but to confront the industrialisation of influence through systems optimised for engagement, dependency and extraction.

In 2016, a YouGov poll found that Vote Leave voters were far likelier to prefer their steaks well-done and to resist cutting down on meat. A fantastically trivial datapoint—but in the hands of a political campaign it becomes a lever of persuasion. A person buying a barbecue or shopping for meat online could now be targeted with messages of patriotism, self-reliance or irritation with the "nanny state".

Tech companies command an extraordinary number of data points on each of us, allowing behaviour to be profiled, predicted and influenced at scale. Deliberately orchestrated for commercial or ideological purposes, these systems bend behaviour towards those seeking to capture it. What many in civil society see as an affront to personal freedom and democracy, peddlers in influence see as proof of concept.

What we see online is not "the world" nor "the truth"; it is whatever someone willing to manipulate the system, for politics or profit, wants us to see. Once you understand that the system determines what you see—and shapes what you believe—trust in information—and trust in each other—begins to corrode.

And yet we stay. Not because the system fails, but because it works—just not on our terms. It offers proximity to all human knowledge, the ability to communicate across distance instantly and the chance to form communities that would otherwise never exist. These are extraordinary gains. The problem is not the promise of the digital world, but the conditions under which it has been built and is now governed.

Previous information systems mediated the distribution of human speech but increasingly large language models (LLMs) also produce synthetic language at scale. LLMs absorb vast quantities of data and produce statistically plausible responses from the patterns they find within it—it is pattern, not meaning. They do not reason from first principles, but from probability. They do not know; they predict.

The point is not that AI lies or hallucinates, but that it optimises for what is most likely: plausibility. And in a system already shaped by engagement, plausibility risks reinforcing what is already dominant, emotionally resonant or commercially valuable, not what is true.

In this environment, communication is driven less by shared meaning than by prediction, optimisation and manipulation. This shapes our information system and the conditions under which citizens encounter the world—and one another.

Children offer the starkest example of how this can go wrong. The chatbot friend who agrees with everything you say offers comfort, but no reciprocity, consequence or social obligation. Like the mirror in a fairy tale, it tells users what they most wish to hear. A system without moral understanding cannot reliably mediate the emotional lives of children. That is why chatbots can encourage self-harm, validate delusion or direct dangerous behaviour while maintaining the appearance of care.

We have allowed technology companies to become extraordinarily powerful and to operate outside ordinary democratic obligations.

The question is not whether AI will transform public life. It already is. The question is whether democratic societies retain the capacity to shape the terms on which that transformation occurs.

No product should scale before it is proven safe enough for the domain in which it operates. Safety, privacy and autonomy should be fundamental to system design by default and without carve-outs.

That means moving away from the idea that companies can deploy first and mitigate later. Regulation should rest on three principles: duty of care to the user; liability where harm is foreseeable and preventable; and the rejection of tech exceptionalism—the principle that a technology company is a company like any other and should be taxed, regulated and liable accordingly.

The choice is not between innovation and regulation, but between a digital world organised around extraction and one organised around human flourishing. We must refuse tech exceptionalism and apply the same basic expectations we already impose elsewhere. Democracies survive not because they eliminate disagreement, but because citizens retain enough shared language and shared reality to disagree meaningfully.

Yes, governance introduces friction. It constrains behaviour and establishes a floor below which companies may not fall, not a ceiling beyond which they may not rise. Friction introduces caution, but it is also what prevents collapse. We already accept it in aviation, medicine, finance and food.

Innovation that cannot survive basic standards is not worth defending.

# Part 5: Building democratic information resilience

This report has shown how persistent uncertainty during high-pressure moments—and the proliferation of false and misleading information—strains participation, weakens institutional authority and erodes public trust. These pressures arise from structural weaknesses in how the UK governs and responds to information risks, alongside broader technological, institutional and platform-driven changes.

Four interlocking weaknesses help explain why the current system is not able to respond effectively at scale or speed. First, coordination across public institutions and online service providers remains fragmented and opaque, with no visible standing mechanism to align responsibilities or manage escalation. Second, legal and regulatory frameworks are not designed to address systemic information risks, instead focusing on specific categories of illegality and harm. Third, independent oversight is constrained by limited transparency and inconsistent access to platform and AI system data, making it difficult to understand how information is produced, amplified and targeted. Fourth, there is a lack of clear accountability and transparency for platforms, search engines and AI systems in relation to their role in shaping information flows, including how algorithmic ranking, recommendation and design choices determine visibility and amplification at scale.

This section sets out proposals to address these and wider gaps, shifting from siloed interventions to a more coherent system of democratic information resilience. Recommendations are structured around four mutually reinforcing pillars:

1. **Secure the information ecosystem**, by improving how information is surfaced, prioritised and made available, particularly during high-risk periods, and ensuring that reliable information reaches audiences at speed and scale.
2. **Strengthen public resilience**, by supporting individuals' ability to interpret uncertainty, evaluate claims and navigate the information environment, alongside safeguards that reduce exposure to false and misleading content.
3. **Modernise laws and institutions**, by creating clearer responsibilities, formal coordination mechanisms and dedicated capacity to manage information risks that cut across existing regulatory and policy boundaries.
4. **Increase commercial transparency and accountability**, by ensuring that platforms, search engines and AI systems are scrutinised and held accountable for systemic effects on information flows, including amplification and ranking.

Full Fact's polling underscores the urgency of reform.[374] Around two thirds of UK adults (66%) think the government is doing too little to address AI-generated misinformation, with only 9% saying the current response is about right. Change is needed to improve resilience and restore confidence that institutions can and will respond effectively.

## 5.1 Secure the information ecosystem: recommendations

The UK's information ecosystem is shaped by structural constraints that limit the ability of accurate, reliable information to compete effectively with misleading or false narratives. These include limited access to platform data for independent researchers, declining levels of public interest journalism, opaque algorithmic amplification systems, uneven prominence for reliable information, and low levels of trust in institutions. This reduces the effectiveness of even well designed institutional responses.

Freedom of expression is paramount. Securing the information ecosystem is not about regulating individual speech or censoring content. It is about strengthening the conditions under which high-quality information can be produced, surfaced, scrutinised, and reach audiences when it matters most. This includes improving the visibility of reliable, public interest information, enabling independent analysis of information flows, and aligning regulatory frameworks with democratic resilience.

Low levels of public trust reinforce the need for action. Full Fact's polling shows that most sources tested—including the media, local MPs and platforms—are not widely trusted to provide accurate and reliable information.[375] A significant proportion of UK adults (17%) would not place their highest trust in any of the institutions they were asked about during a national emergency, including the police, media, government ministers or local authorities. This reinforces the importance of improving the accuracy, consistency and prominence of reliable information, so that it can be recognised and acted upon.

- ### 5.1.1 Stress-test information resilience

  The government should establish a programme of multi-stakeholder information resilience exercises to test how institutions, platforms and public communication structures perform during elections, major crises and other high-risk periods. Structured simulation exercises should test decision-making under conditions of uncertainty, and assess coordination, escalation pathways, information sharing arrangements and communication plans. Testing should be embedded into routine preparedness activity, with increased intensity ahead of elections. Summary findings should be published in an appropriate form and used to strengthen crisis protocols, coordination arrangements and institutional capability.

- ### 5.1.2 Maintain crisis communication plans and incident protocols

  Ofcom should require the largest platforms, search engines and generative AI systems to maintain crisis communication plans and information incident protocols, with clear escalation pathways and severity indicators. Full Fact's submission to Ofcom's consultation on crisis response protocols sets out

detailed proposals for a severity-based framework.[376] These systems should include established communication channels with regulators, law enforcement, departments, civil society organisations, and the media. These requirements should align with national incident response arrangements (see 5.3.2) to ensure interoperability. Ofcom should require audited evidence that these processes are effective. The government's planned review of crisis response powers under the Online Safety Act provides an opportunity to establish these duties in relation to systemic risks.

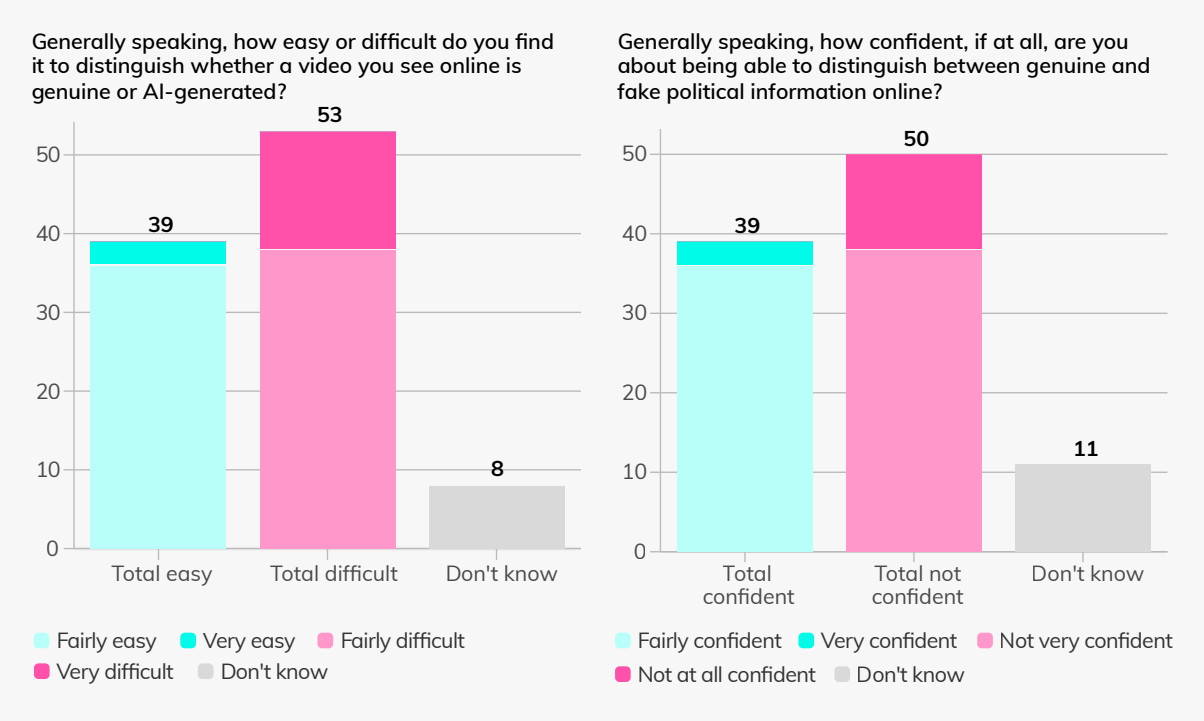
- **5.1.3 Increase the prominence of high-quality information**

  During defined high-risk periods, the largest platforms, search engines and AI systems should take steps to increase the prominence of high-quality, factual public interest information. These measures should be triggered by transparent and publicly defined risk thresholds, and should include clear accountability for how sources are identified and prioritised. The Council of Europe's guidance on the prioritisation of public interest content could inform guidelines to determine high-quality and reliable sources of information.[377] The government has set out its intention to look at giving *"trustworthy media due prominence, so people have access to authoritative sources to counter mis and disinformation."*[378]

## 5.2 Strengthen public resilience: recommendations

Strengthening public resilience means improving the public's ability to interpret, evaluate, and make sense of what they are seeing, understand why and how information changes, and recognise the limits of what can reasonably be known. It also means helping them to navigate uncertainty effectively and confidently, and giving users more agency over their online information environment.

This does not mean making individuals solely responsible for identifying false or misleading information. Nor should it be seen as a substitute for regulatory action or platform responsibility. Low confidence levels highlight this constraint and underscore the limits of individual-level responses in a complex and rapidly evolving information environment. Full Fact's polling found that only 3% of people feel very confident identifying fake political information online, with the same share finding it very easy to identify an AI-generated video.[379] These constraints are likely to become more pronounced as generative AI tools become more widely used and increasingly difficult to detect.




Polling research by Internet Matters and Full Fact shows that these challenges are already shaping political engagement well before voting age—which may intensify as younger voters are enfranchised without adequate support: 74% of people aged 13-14 and 81% of those aged 15-17 say they have seen content about news, politics or current affairs online. Of those who see political information online, only 53% of those aged 13-17 feel confident telling whether it is true or false. As the voting age is lowered to 16, the absence of robust media, AI and political literacy provisions risks widening this gap.[380] Strengthening public resilience is a prerequisite for effective democratic participation.

Full Fact's work with the Commission into Countering Online Conspiracy Theories in Schools highlights the scale of the challenge. The Commission was launched in 2024 to consider how online conspiracy theories, misinformation and disinformation manifest in the lives of young people, and published a report in February 2025 with recommendations to address this issue.[381] In 2026, the Commission published research which found that *"Young people, parents and school staff increasingly recognise not only that misleading content is prevalent online, but that it is becoming harder to identify what is true and what is false."*[382]

- ### 5.2.1 Embed media literacy across the curriculum

  The Department for Education should support schools and teachers to deliver media and information literacy across subjects and key stages in the curriculum, with guidance, training and resources.[383] This should include consolidated guidance on content and delivery; integration into Initial Teacher Training and Continuing Professional Development; and high-quality, maintained teaching resources, including a central online hub. Provision should reflect the realities of modern information systems, including ranking and generative AI content.

- **5.2.2 Fund long-term media literacy delivery capacity**

  The government should ensure that there is long-term investment to support the delivery of media and information literacy, including training, professional development and curriculum implementation. Funding should prioritise workforce capacity in schools and education providers, with mechanisms to ensure sustainable delivery at scale. Parents and carers should also be supported through accessible, reliable guidance and public education resources, recognising that building resilience extends beyond formal education and into lifelong learning.

- **5.2.3 Introduce a statutory duty to provide media literacy**

  The government should place a statutory duty on the largest platforms, search engines and AI systems to support effective, evidence-based media and information literacy among users. This duty should be supported by a code of practice, developed by Ofcom, setting out expectations with requirements for providers to demonstrate effectiveness through evidence and evaluation. The duty should require services to embed media literacy by design, including through transparency features that help users understand why content is shown, tools that support critical evaluation, clear labelling and disclosure of AI-generated or modified content, and user controls that enable management of recommender systems.

## 5.3 Modernise laws and institutions: recommendations

The UK's laws and institutions are not well adapted to the demands of high-pressure events such as elections and crises. Responsibilities are fragmented across multiple organisations, and coordination mechanisms and escalation pathways are unclear. While each organisation plays an important role within its own mandate, there is no single institution responsible for ensuring coherence and coordination across the system. This fragmentation creates predictable weaknesses at precisely the points where clarity and coordination are most needed: delayed escalation, inconsistent or competing public messaging, and uneven responses across institutions. It also constrains the ability of the system to prepare in advance and to undertake post-incident learning.

The framework has not kept pace with rapidly evolving and systemic risks, including those relating to algorithmic amplification and AI-generated content. While there is growing recognition of the challenges and the need for more oversight and coordination—including in recent reviews by the SIT Committee,[384] Foreign Affairs Committee[385] and Philip Rycroft[386]—the institutional response remains dispersed and opaque.

The UK's experience in cyber security, particularly through the National Cyber Security Centre, demonstrates the value of a dedicated, well-resourced institution with operational capability and system-wide leadership.[387] A national Information Resilience Unit, built on similar principles, would address the current gap by providing a standing coordination function for the information environment. With statutory independence and accountability to Parliament, it would strengthen preparedness, improve coherence between existing bodies, and enable more effective cross-system response to acute challenges in the information environment, including those affecting public safety or democracy.

- ### 5.3.1 Strengthen the Representation of the People Bill

  The government should strengthen the Representation of the People Bill to address risks from misinformation and disinformation that affect the integrity of the UK's democratic processes. This should include measures to upgrade the Online Safety Act to address systemic harms; require transparency in the use of political deepfakes; establish a public library of political adverts; introduce a regulatory framework for political advertising standards; create a serious election information incident protocol; and strengthen the investigative powers of the Electoral Commission. Our policy paper contains more detail on these reforms.[388]

- ### 5.3.2 Establish a national information incident response framework

  The government should build a dedicated information incident response framework within the UK's existing crisis management architecture, aligned with national resilience doctrine and crisis management arrangements. The framework should define severity thresholds, escalation pathways, communication and coordination processes for information incidents, such as AI-generated content and platform-amplified information threats during elections and crises. Full Fact's Framework for Information Incidents provides a basis for this.[389] It should be overseen by the Information Resilience Unit and align with service providers' crisis response protocols and crisis communication plans, ensuring consistent coordination and escalation between public and private actors during high-impact events.

- ### 5.3.3 Create a national Information Resilience Unit

  The government should establish a statutory Information Resilience Unit as a standing coordination function to strengthen preparedness and response to serious information incidents. The unit would provide a single, visible, enduring mechanism to support cross-system coordination, coherence, preparedness and institutional learning in relation to major information risks. It would be integrated

with existing national resilience and crisis coordination structures, combine operational capability with system-wide oversight, and have independence and robust accountability to Parliament. Responsibilities should include: maintaining the national information incident response framework, enabling preparedness through stress-testing and scenario planning, convening stakeholders during serious information incidents, supporting coordination and the sharing of information, and leading post-incident reviews to strengthen future resilience.

## 5.4 Increase commercial transparency and accountability: recommendations

Online platforms, search engines and AI systems play a central role in shaping the information environment. These systems determine what information is surfaced, amplified or suppressed at scale, with limited transparency into how outcomes are produced.

Accountability has not kept pace with this level of influence. Existing regulation focuses on content that is illegal or harmful to children, rather than wider systemic risks or the behaviour of information systems—particularly how algorithmic ranking, recommender systems and generative AI shape the distribution and visibility of content under conditions of speed, scale and uncertainty. This creates regulatory blind spots in understanding how these systems operate and the real-world impacts they have.

There are also gaps in the governance of AI systems that increasingly function as information intermediaries, but sit outside or only partially within regulatory frameworks. A more comprehensive approach would focus on system-level behaviour, including greater transparency around ranking and recommendation systems, clearer accountability for amplification dynamics, and obligations to assess and mitigate systems risks.

Strengthening transparency and accountability aligns with public expectations of corporate responsibility. Reuters Institute research from 2025 found that most people believe platforms—rather than governments—should bear primary responsibility for moderating harmful content online, even where regulatory oversight is supported.[390]

- ### 5.4.1 Introduce systemic risk management duties

  The government should strengthen the Online Safety Act to require the largest platforms and search engines to identify, assess and mitigate the risks they pose to civic discourse, electoral processes and public security. These duties should cover risks arising from system design and operation, including algorithmic amplification, recommender systems and ranking biases, as well as the spread

of misinformation during fast-moving or high-pressure events. Equivalent duties should be extended to widely deployed AI systems that fall outside of the Online Safety Act. These duties should focus on system behaviour at scale, particularly during, but not limited to, high-risk periods (activity intended to influence elections does not start and stop with the dissolution of Parliament and the King's speech), and be subject to independent oversight, transparency and enforcement.[391]

- ### 5.4.2 Implement provenance and labelling standards

  The government should require large online platforms, search engines and widely deployed AI systems to implement clear, consistent and interoperable standards for signalling the provenance of AI-generated content at the point of use. This should include labelling and metadata that persist across services and through sharing, rather than platform-specific or easily removed indicators. Requirements should be set in statute, be designed to operate effectively at scale during high risk periods, and build on emerging technical standards like SynthID and the C2PA specification, which offer robust frameworks for implementing provenance metadata. Provenance systems should be compatible with independent detection tools, accessible to researchers and fact checkers, and subject to oversight to ensure consistent and reliable application.

- ### 5.4.3 Require AI provider output transparency

  The government should require providers of widely deployed AI systems and hybrid models, like Grok and Google AI Overviews, to ensure outputs include clear, consistent and meaningful transparency on sources, uncertainty and system limitations, particularly for high-impact domains such as health, politics and finance. This should include source citation where available, clear indicators where information is incomplete or contested, and explanations of how outputs are generated. Requirements should operate across interfaces and use cases, and support both user understanding and independent scrutiny. These measures should enable rapid assessment of information reliability during high-risk periods, and be subject to standards and oversight to ensure they are accurate and usable in practice.

- ### 5.4.4 Enable enhanced access to platform data

  The government should use the Data (Use and Access) Act to establish a robust framework for independent researcher access to platform, search engine and AI system data, building on proposals from Ofcom and international best practice. This framework should enable secure, privacy-preserving and proportionate access for independent researchers, including accredited fact checkers, with provision for timely access during elections and other high-risk periods. It should move beyond ad hoc access to provide structured, ongoing access to relevant data and systems, enabling scrutiny of how content is generated, ranked and amplified at scale. The framework should include clear accreditation processes, enforceable rights and independent oversight to ensure compliance. Its purpose should be to enable systematic scrutiny of how information is generated, ranked, distributed and amplified at scale, and to support evidence-based policy and regulatory interventions.

# Conclusion

A healthy democracy depends on the ability of citizens, institutions and communities to understand facts, recognise what is real, reach shared understandings about things that matter, and be able to make informed decisions. This is now under immense and increasing strain. The information environment is fragmented, rapidly evolving, and increasingly shaped by automated systems, while the legal and institutional frameworks that should oversee it have not kept pace.

This report shows that a central risk to our democracy is not just the prevalence of misinformation, but the growing persistence of uncertainty about what information can be trusted, verified and acted upon. The uncertainty extends to the institutional response itself, which remains fragmented and difficult for the public to understand. There is little sign of coherent action to stabilise the information environment, as it becomes progressively harder to know which sources and voices to trust.

Public attitudes underline the scale of this challenge. Informed scepticism and robust challenge are essential features of a healthy information environment. But confidence in identifying misleading or AI-generated political content is low, trust in major institutions as sources of reliable information is weak, and a clear majority believe that the government's current response to AI-driven misinformation is insufficient. These perceptions matter. They shape how citizens interpret information, how they engage with institutions, and ultimately how they participate in democratic processes.

International experience shows that while these risks cannot be eliminated, their impact can be reduced. Where approaches are coordinated, responsibilities are clearly defined, the focus is on system-level risks, and responses are visible and well-communicated, uncertainty can be contained before it escalates into wider democratic harm. Such systems provide the stabilising reference points that allow people to navigate complex and fast-moving information environments with greater confidence.

In the UK, however, current arrangements are not yet sufficient for this task. Responsibilities are fragmented across opaque institutions, coordination mechanisms are not widely understood, and legal frameworks are not aligned with the realities of the contemporary information system, let alone emerging risks. At the same time, key parts of that system— including social media platforms, search engines, AI systems, digital advertising and data infrastructure—are partially regulated and weakly integrated into democratic oversight.

The result is a widening gap between the scale and nature of information risks and the capacity of governance systems to address them. Democratic resilience depends on institutions operating coherently and, as far as possible, in ways that are visible and intelligible to the public. Closing this gap will require more than isolated and incremental reforms. It requires a shift towards a systemic approach that treats the information environment as an interconnected whole, identifies points of stress and vulnerability, and strengthens the foundations of trust and shared understanding.

The recommendations in this report are intended to support that shift. They aim to ensure that when uncertainty arises, institutions are able to respond quickly, clearly and credibly; that legal frameworks are equipped for the realities of the modern environment; and that the information ecosystem as a whole supports, rather than undermines, public understanding and participation. They also aim to ensure that citizens are equipped to navigate the increasingly complex and automated information landscape with confidence, and without bearing the full burden.

Ultimately, strengthening the information environment is not only a regulatory or technical challenge, but a democratic one. It is essential to sustaining the conditions for collective decision-making—the ability to resolve disagreement through shared evidence, to act in moments of uncertainty, and to maintain confidence that public decisions are grounded in a reality that can be known, communicated and trusted.

# Endnotes


1 All figures, unless otherwise stated, are from YouGov Plc. The total sample size was 2,175 adults. Fieldwork was undertaken on 29-30 March 2026. The survey was carried out online. The figures have been weighted and are representative of all UK adults (aged 18+).

2 Full Fact, "Full Fact AI", https://fullfact.org/ai/

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

4 Full Fact, "The Gorton and Denton by-election: fact checked", 24 February 2026, https://fullfact.org/politics/gorton-and-denton-by-election-round-up/

5 Sian Bayley, "Fake Yvette Cooper Guardian article circulates online", Full Fact, 29 July 2025, https://fullfact.org/politics/fake-guardian-article-yvette-cooper-far-right/

6 Lilie Chouliaraki, Kathryn Claire Higgins, "The truth about "two-tier policing", LSE Blogs, 15 August 2024, https://blogs.lse.ac.uk/politicsandpolicy/the-truth-about-two-tier-policing/

7 Hannah Smith, "No, Reform UK did not field a 'fake' candidate in Norfolk", Full Fact, 12 May 2026, https://fullfact.org/politics/reform-did-not-field-fake-candidate-george-boyd/

8 Steve Nowottny, "Five lessons from our fact checking in 2025", Full Fact, 23 December 2025, https://fullfact.org/politics/five-lessons-from-our-fact-checking-2025/

9 Steve Nowottny, "Five lessons from our fact checking in 2025", Full Fact, 23 December 2025, https://fullfact.org/politics/five-lessons-from-our-fact-checking-2025/

10 European Digital Media Observatory, Monthly Briefing 53, 17 November 2025, https://edmo.eu/wp-content/uploads/2025/11/EDMO-Horizontal-53-1.pdf

11 Sian Bayley, "Fake image of Iranian children's funeral shared online by MP", Full Fact, 9 March 2026, https://fullfact.org/conflict/fake-image-iranian-schoolchildren-burial/

12 Alexandra Topping, "Voters contend with 'dodgy' data in party leaflets for English local elections", 26 April 2026, Guardian, https://www.theguardian.com/politics/2026/apr/26/voters-contend-with-grotesque-leaflets-and-dodgy-data-in-english-local-elections-say-analysts

13 Democracy Club, "Election Leaflets", https://electionleaflets.org/leaflets/

14 YouGov, "Voting intention weekly tracker", https://yougov.com/en-gb/trackers/voting-intention

15 Severin Carrell, "Software tackling deepfakes to be piloted for Scottish and Welsh elections", Guardian, 8 January 2026, https://www.theguardian.com/technology/2026/jan/08/pilot-software-tackle-deepfakes-scottish-welsh-elections; Electoral Commission, "Electoral Commission launches deepfake detection pilot to counter AI misinformation", 15 April 2026, https://www.electoralcommission.org.uk/media-centre/electoral-commission-launches-deepfake-detection-pilot-counter-ai-misinformation

16 Bertelsmann Stiftung, "Algorithms in election campaigns: A study by the University of Potsdam and the Bertelsmann Stiftung shows an imbalance in the visibility of political parties in young people's social media feeds", 3 November 2025, https://www.bertelsmann-stiftung.de/en/unsere-projekte/engagement-junger-menschen-fuer-demokratie/projektnachrichten/algorithms-in-election-campaigns

17 Craig T. Robertson, Amy Ross Arguedas, Mitali Mukherjee, and Richard Fletcher, "Understanding Young News Audiences at a Time of Rapid Change", Reuters Institute for the Study of Journalism, March 2026, https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2026-03/Young_people_and_the_news.pdf#page=8

18 Ofcom, "Top trends from our latest look at the UK's news habits", 21 July 2025, https://www.ofcom.org.uk/media-use-and-attitudes/attitudes-to-news/top-trends-from-our-latest-look-at-the-uks-news-habits

19 Ofcom, "News consumption in the UK: 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/news-consumption-in-the-uk-2024-report.pdf?v=379621&utm

20 Fabrizio Germano, Vicenç Gómez, Francesco Sobbrio, "Ranking for engagement: How social media algorithms fuel misinformation and polarization", Journal of Public Economics, March 2026, https://www.sciencedirect.com/science/article/pii/S0047272726000253

21 Xuyang Zhu, "The new UK Media Act - a new 'Must Offer'/'Must Carry' prominence regime for PSBs on online TV selection services", Taylor Wessing, 29 July 2024, https://www.taylorwessing.com/de/interface/2024/media-update-2024/the-media-act

22 Martin Shipton, Nation Cymru, "Fake social media posts claim Reform UK won Caerphilly by-election", 23 October 2025, https://nation.cymru/news/fake-social-media-posts-being-spread-stating-that-reform-uk-has-won-the-caerphilly-by-election/; Ellie Gosley and John, "Reform UK sends by-election letter from "local" voter ineligible to vote", Wales Online, 12 September 2025, https://www.walesonline.co.uk/news/politics/reform-uk-sends-out-election-32460475

23 Speaker's Conference on the security of candidates, MPs and elections, "first report", 2 June 2025, https://committees.parliament.uk/publications/48116/documents/251907/default/, "second report", 27 October 2025, https://committees.parliament.uk/publications/49841/documents/268474/default/; see also Online Safety Act Network, "Improved online safety for candidates during elections: a code of practice", 2 February 2026, https://www.onlinesafetyact.net/analysis/improved-online-safety-for-candidates-during-elections-a-code-of-practice/

24 HM Government, Response to Speaker's Conference reports, 5 March 2026, https://committees.parliament.uk/publications/51985/documents/288491/default/

25 Sam Stockwell, "AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections", Centre for Emerging Technology and Security, 19 September 2024, https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-analysis-2024-uk-and-european-elections

26 HM Government, April 2025, "The Amber Book: Managing Crisis in Central Government", https://assets.publishing.service.gov.uk/media/680a4fbd6d6ac02ee99d8488/35.20_CO_Emergency_Response_and_Recovery_02_Amber_Book_FINAL_PRINT.pdf

27 Jiahui Lu, "Themes and Evolution of Misinformation During the Early Phases of the COVID-19 Outbreak in China—An Application of the Crisis and Emergency Risk Communication Model", Health Communication, 14 August 2020 https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2020.00057/full

28 Sam Stockwell, Al Baker, "The Cost of Silence: Crisis Communication and Real-world Harm Following Security Incidents", Centre for Emerging Technology and Security, 28 July 2025, https://cetas.turing.ac.uk/sites/default/files/2025-07/cetas_expert_analysis_-_the_cost_of_silence_-_crisis_communication_and_real-world_harm_following_security_incidents.pdf

29 Elizabeth Seger, Sam Stockwell, Tyreese Calnan, Henry Ajder, Jamie Hancock, Hannah Perry, "Epistemic Security for Crisis Resilience", Demos, 19 January 2026 https://demos.co.uk/wp-content/uploads/2026/01/Epistemic-Security-for-Crisis-Resilience_Report_2025_Jan_optimised.pdf

30 Full Fact, "Full Fact report 2025", May 2025, https://fullfact.org/policy/reports/full-fact-report-2025/; Alicia Wanless, "The Liverpool Response to Misinformation Was a Good First Step—but It's Not Enough", Carnegie Endowment for International Peace, 5 June 2025, https://carnegieendowment.org/emissary/2025/06/liverpool-misinformation-response-information-ecosystem?lang=en

31 HMICFRS, "An inspection of the police response to the public disorder in July and August 2024: Tranche 2", 7 May 2025, https://hmicfrs.justiceinspectorates.gov.uk/publication-html/police-response-to-public-disorder-in-july-and-august-2024-tranche-2/

32 Ben Quinn, "How violent protests in Epping are being fuelled by disinformation", Guardian, 24 July 2025, https://www.theguardian.com/uk-news/2025/jul/24/analysis-violence-epping-protests-disinformation

33 ITV News, "Police impose restrictions ahead of protests in Epping", 27 July 2025, https://www.itv.com/news/london/2025-07-27/police-impose-restrictions-ahead-of-protests-in-epping

34 Daniel Sandford, "Why police released details about Liverpool crash suspect so quickly", BBC News, 27 May 2025, https://www.bbc.co.uk/news/articles/cvgv4ddpyddo

35 National Police Chiefs' Council, College of Policing, "Interim guidance relating to ethnicity and/or nationality of suspects", 11 August 2025, https://assets.college.police.uk/s3fs-public/2025-08/NPCC-College-interim-guidance-on-ethnicity-nationality_0.pdf?VersionId=9ZMB.W9xxLhwYr28Vcl0.B0Ni9T0llAb&v=1755080959

36 Alicia Wanless, "The Liverpool Response to Misinformation Was a Good First Step—but It's Not Enough", Carnegie Endowment for International Peace, 5 June 2025, https://carnegieendowment.org/emissary/2025/06/liverpool-misinformation-response-information-ecosystem?lang=en

37 Law Commission, Part 1 of review of contempt of court laws, 18 November 2025, https://lawcom.gov.uk/project/contempt-of-court/; Law Commission, "Press release: New framework to modernise contempt of court laws", 18 November 2025, https://lawcom.gov.uk/news/new-framework-to-modernise-contempt-of-court-laws

38 Sam Stockwell, Ardi Janjeva, Broderick McDonald, "Adding Fuel to the Fire: AI Information Threats and Crisis Events", Centre for Emerging Technology and Security, 11 February 2026, https://cetas.turing.ac.uk/publications/adding-fuel-to-fire

39 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

40 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

41 Amy Mitchell , Mark Jurkowitz , J. Baxter Oliphant and Elisa Shearer, "Misinformation and competing views of reality abounded throughout 2020", Pew Research Centre, 22 February 2021, https://www.pewresearch.org/journalism/2021/02/22/misinformation-and-competing-views-of-reality-abounded-throughout-2020/

42 OECD, "OECD Survey on Drivers of Trust in Public Institutions – 2024 Results Building Trust in a Complex Policy Environment", 10 July 2024, https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results_9a20554b-en/full-report/trust-and-information-integrity_49ce5100.html

43 Ipsos, JOE Media, "Young Britons trust social media as a news source despite misinformation fears", 6 October 2025, https://www.ipsos.com/en-uk/young-britons-trust-social-media-news-source-despite-misinformation-fears

44 Full Fact and Internet Matters, "Preparing young people to vote in a complex, attention-driven information environment", 10 February 2026, https://fullfact.org/documents/405/Briefing_-_votes_at_16_and_media_literacy.pdf

45 Thom Roozenbeek, Caspar van den Berg, Mattijs S. Lambooij, Sander van der Linden, Rakoen Maertens, José A. Ferreira, Mart van Dijk, Jon Roozenbeek, "Trust in institutions and misinformation susceptibility both independently explain vaccine skepticism", Scientific Reports, 28 October 2025, https://www.nature.com/articles/s41598-025-21452-1

46 Joshua Freitag , Madeline Gochee , Mitchell Ransden , Brendan Nyhan Open the ORCID record for Brendan Nyhan[Opens in a new window] , Kristy Roschke and Dan Gillmor, "The Corrections Dilemma: Media Retractions Increase Belief Accuracy But Decrease Trust", Journal of Experimental Political Science https://www.cambridge.org/core/journals/journal-of-experimental-political-science/article/corrections-dilemma-media-retractions-increase-belief-accuracy-but-decrease-trust/A6A9A16D97F22E1CA22B051D215E752B

47 Smitha Milli, Micah Carroll, Yike Wang, Sashrika Pandey, Sebastian Zhao, Anca D Dragan, "Engagement, user satisfaction, and the amplification of divisive content on social media", PNAS Nexus, 5 March 2025, https://academic.oup.com/pnasnexus/article/4/3/pgaf062/8052060

48 Fabrizio Germano , Vicenç Gómez, Francesco Sobbrio, "Ranking for engagement: How social media algorithms fuel misinformation and polarization", Journal of Public Economics, March 2026, https://www.sciencedirect.com/science/article/pii/S0047272726000253

49 Electoral Commission, "Response to the SIT Committee call for evidence on social media, misinformation and harmful algorithms", 28 January 2025, https://www.electoralcommission.org.uk/news-and-views/our-responses-consultations/response-science-innovation-and-technology-committee-call-evidence-social-media-misinformation-and

50 Foreign Affairs Committee, "Oral evidence: Disinformation Diplomacy", 13 January 2026, https://committees.parliament.uk/oralevidence/16988/pdf/;

Vijay Rangarajan, "Additional Electoral Commission information for the Foreign Affairs Committee", Electoral Commission, 28 January 2026, https://committees.parliament.uk/writtenevidence/162487/html/

51 The White House, "Restoring Freedom of Speech and Ending Federal Censorship", 20 January 2025, https://www.whitehouse.gov/presidential-actions/2025/01/restoring-freedom-of-speech-and-ending-federal-censorship/

52 Mickey Carroll, "Hundreds of UK moderators have left TikTok - sparking safety fears, whistleblowers reveal", Sky News, 4 December 2025, https://news.sky.com/story/hundreds-of-uk-moderators-have-left-tiktok-sparking-safety-fears-whistleblowers-reveal-13478302

53 Science, Innovation and Technology Committee, "TikTok fails to share evidence behind increased AI use in content moderation", 13 November 2025, https://committees.parliament.uk/work/8641/social-media-misinformation-and-harmful-algorithms/news/210394/tiktok-fails-to-share-evidence-behind-increased-ai-use-in-content-moderation/

54 Full Fact, "Why is TikTok penalising content designed to highlight misinformation?", 2 April 2026, https://fullfact.org/technology/tiktok-penalising-misinformation-content/

55 Science, Innovation and Technology Committee, "Follow-up on Social media, misinformation and harmful algorithms inquiry", 24 March 2026, https://committees.parliament.uk/oralevidence/17405/html/

56 Full Fact, "Full Fact response to Meta Oversight Board: Community Notes rollout", December 2025, https://www.oversightboard.com/wp-content/uploads/gravity_forms/78-bebea08df7d3661c35c6533a224228ef/2025/12/Full-Fact-response-to-Meta-Oversight-Board-on-Community-Notes-rollout-.pdf

57 Demos, Full Fact, "Community Disorder: How do we prevent an information emergency?", July 2025, https://demos.co.uk/wp-content/uploads/2025/07/Community-Disorder_2025_Policy-Brief_Online.pdf

58 Alexios Mantzarlis, "Inside a pro-Conservative influence operation on community notes", Indicator, 14 April 2026, https://indicator.media/p/inside-a-pro-conservative-influence-operation-on-community-notes

59 Craig Silverman, Alexios Mantzarlis, "Briefing: Finally, some data about Meta's Community Notes programme", Indicator, 19 September 2025, https://indicator.media/p/data-about-meta-community-notes-youtube-archiving

60 Meta Oversight Board, "Assessing Meta's Plans to Expand Community Notes", 26 March 2026, https://www.oversightboard.com/decision/pao-007g5zuv/

61 Science, Innovation and Technology Committee, "Follow-up on Social media, misinformation and harmful algorithms inquiry", 24 March 2026, https://committees.parliament.uk/oralevidence/17405/html/

62 Chris Morris, "Google cuts funding to Full Fact...", Full Fact, 16 October 2025, https://fullfact.org/technology/google-cuts-funding-to-full-fact/

63 Lara O'Reilly, "YouTube says it'll bring back creators banned for violating its COVID-19 and election content policies", Business Insider, 23 September 2025, https://www.businessinsider.com/youtube-reinstate-channels-banned-over-covid-content-policies-2025-9

64 Science, Innovation and Technology Committee, "Oral evidence: Social media, misinformation and harmful algorithms", 25 February 2025, https://committees.parliament.uk/oralevidence/15408/html/; Science, Innovation and Technology Committee, "Follow-up on Social media, misinformation and harmful algorithms inquiry", 24 March 2026, https://committees.parliament.uk/oralevidence/17405/html/

65 Meta, "Misinformation Policy details Change Log", April 2025, https://transparency.meta.com/policies/community-standards/misinformation

66 Craig Silverman, Alexios Mantzarlis, "Briefing: "the brakes are completely off at Meta", Indicator, 12 September 2025, https://indicator.media/p/brakes-are-completely-off-at-meta-fact-checking-hoaxes

67 Jeff Horwitz, "Meta is earning a fortune on a deluge of fraudulent ads, documents show", Reuters, 6 November 2025, https://www.reuters.com/investigations/meta-is-earning-fortune-deluge-fraudulent-ads-documents-show-2025-11-06/; Jeff Horwitz, Engen Tham, "Meta tolerates rampant ad fraud from China to safeguard billions in revenue", Reuters, 15 December 2025, https://www.reuters.com/investigations/meta-tolerates-rampant-ad-fraud-china-safeguard-billions-revenue-2025-12-15/

68 Maldita.es, "More than 1,000 fraudulent Facebook pages in 60 countries: the international network of public transport scams and its links to Russia and Vietnam", 16 July 2025, https://maldita.es/investigaciones/20250716/transport-scam-network-international-facebook/

69 Henry Hsu, "Simplifying the search results page", Google Search Central Blog, 12 June 2025, https://developers.google.com/search/blog/2025/06/simplifying-search-results

70 Andrew Dudfield, "The web just got a little harder to trust", Full Fact, 26 June 2025, https://fullfact.org/technology/the-web-just-got-a-little-harder-to-trust/

71 Science, Innovation and Technology Committee, "Follow-up on Social media, misinformation and harmful algorithms inquiry", 24 March 2026, https://committees.parliament.uk/oralevidence/17405/html/

72 Google Support, "YouTube Channel Monetisation Policies", 15 July 2025, https://support.google.com/youtube/answer/1311392?hl=en-GB&ref_topic=9153642&sjid=6974982542994765316-EU#zippy=%2Cfollow-the-youtube-community-guidelines

73 Andre Revilla, "X to require AI labels on armed conflict videos from paid creators, citing 'times of war'", Endgadget, 3 March 2026, https://www.engadget.com/social-media/x-to-require-ai-labels-on-armed-conflict-videos-from-paid-creators-citing-times-of-war-183631400.html

74 Ofcom, "Passive social media use, AI companionship, and online side hustles: UK adults' media and online lives revealed", 2 April 2026, https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-adults/passive-social-media-use-ai-companionship-and-online-side-hustles-uk-adults-media-and-online-lives-revealed

75 Ofcom, "Qualitative research: User experiences of Generative Artificial Intelligence (GenAI) Search", 26 September 2025, https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/online-research/other/generative-ai-search-qualitative-research-report.pdf?v=403429

76 Elizabeth Reid, "Generative AI in Search: Let Google do the searching for you", Google Blog, 14 May 2024, https://blog.google/products/search/generative-ai-google-search-may-2024/

77 Hema Budaraju, "New ways to connect to the web with AI Overviews", Google Blog, 15 August 2024, https://blog.google/products/search/new-ways-to-connect-to-the-web-with-ai-overviews/

78 Hema Budaraju, "AI Mode is now available in more languages and locations around the world", Google Blog, 7 October 2025, https://blog.google/products/search/ai-mode-expands-languages-locations/; Google Blog, "The latest AI news we announced in March", 4 April 2025, https://blog.google/technology/ai/google-ai-updates-march-2025/

79 Elizabeth Reid, "AI in Search: Going beyond information to intelligence", Google Blog, 20 May 2025, https://blog.google/products/search/google-search-ai-mode-update/

80 Google Search Help, "Learn about generative AI", https://support.google.com/websearch/answer/13954172

81 Evie Townend, "Google Lens's AI overviews shared misleading information about images", Full Fact, 13 August 2025, https://fullfact.org/technology/technologyinaccurate-google-ai-overviews/

82 Nilesh Christopher, Valerio Pepe, "As millions adopt Grok to fact-check, misinformation abounds", Al Jazeera, 11 July 2025 https://www.aljazeera.com/economy/2025/7/11/as-millions-adopt-grok-to-fact-check-misinformation-abounds; Arthur Sullivan, "AI chatbots fail at accurate news, major study reveals", DW, 22 October 2025, https://www.dw.com/en/chatbot-ai-artificial-intelligence-chatgpt-google-gemini-news-misinformation-fact-check-copilot-v2/a-74392921

83 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

84 Roa Powell, Carsten Jung, "AI's got news for you: can AI improve our information environment?", Institute for Public Policy Research, January 2026, https://ippr-org.files.svdcdn.com/production/Downloads/AI_and_news_January26.pdf?dm=1769681399

85 Andrew Gregory, "How the 'confident authority' of Google AI Overviews is putting public health at risk", Guardian, 24 January 2026, https://www.theguardian.com/technology/ng-interactive/2026/jan/24/how-the-confident-authority-of-google-ai-overviews-is-putting-public-health-at-risk; Andrew Gregory, "Google AI Overviews cite YouTube more than any medical site for health queries, study suggests", Guardian, 24 January 2026, https://www.theguardian.com/technology/2026/jan/24/google-ai-overviews-youtube-medical-citations-study

86 Jacob Granger, "Nearly half of AI-generated answers contain an error, finds EBU-BBC report", Journalism UK, 14 November 2025, https://www.journalism.co.uk/nearly-half-of-ai-generated-answers-contain-an-error-ebu-bbc-report/

87 Klaudia Jaźwińska and Aisvarya Chandrasekar, "AI Search Has a Citation Problem", Columbia Journalism Review, 6 March 2025, https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php

88 Océane Herrero, "Revealed: Apple is teaching its AI to adapt to the Trump era", Politico, 9 September 2025, https://www.politico.eu/article/apple-teaching-artificial-intelligence-adapt-to-trump-era/

89 Matt Southern, "Google CTRs Drop 32% For Top Result After AI Overview Rollout", Search Engine Journal, 21 July 2025, https://www.searchenginejournal.com/google-ctrs-drop-32-for-top-result-after-ai-overview-rollout/551730/; Charlotte Tobitt, "Google told to 'stop the BS' as it claims AI has not harmed website clickthroughs", Press Gazette, 7 August 2025, https://pressgazette.co.uk/platforms/google-search-clicks-traffic-2025-ai-overviews/

90 Evie Townend, "Footage of 2023 car accident in US falsely shared as Diogo Jota crash", Full Fact, 11 July 2025, https://fullfact.org/sport/diogo-jota-crash-video-miscaptioned/

91 Evie Townend, "Clip shared with claims it shows Diogo Jota's crash predates accident", Full Fact, 9 July 2025, https://fullfact.org/sport/diogo-jota-crash-miscaptioned-footage/

92 Max Kozlov, "AI chatbots can sway voters with remarkable ease — is it time to worry?", Nature, 4 December 2025, https://www.nature.com/articles/d41586-025-03975-9; Bartosz Zawiślak, "Conversations with chatbots influence voters' political decisions", Jagiellonian University in Krakow, 8 December 2025, https://en.uj.edu.pl/en_GB/news/-/journal_content/56_INSTANCE_SxA5QO0R5BDs/81541894/159954021

93 Imran Rahman-Jones, "MP wants Elon Musk's chatbot shut down over claim he enabled grooming gangs", BBC News, 5 November 2025, https://www.bbc.co.uk/news/articles/c9d6vwg80qyo

94 Vasilios Mavroudis, Chris Hicks, "LLMs may be more vulnerable to data poisoning than we thought", The Alan Turing Institute, 9 October 2025, https://www.turing.ac.uk/blog/llms-may-be-more-vulnerable-data-poisoning-we-thought; Nancy Lapid, "Medical misinformation more likely to fool AI if source appears legitimate, study shows", Reuters, 9 February 2026, https://www.reuters.com/business/healthcare-pharmaceuticals/medical-misinformation-more-likely-fool-ai-if-source-appears-legitimate-study-2026-02-09/

95 Björn Lindström, Martin Bellander, David T. Schultner, Allen Chang, Philippe N. Tobler & David M. Amodio, "A computational reward learning account of social media engagement", Nature Communications, 26 February 2021, https://www.nature.com/articles/s41467-020-19607-x; Barak Libai et al, "Influencer marketing unlocked: Understanding the value chains driving the creator economy", Journal of the Academy of Marketing Science, 24 January 2025, https://link.springer.com/article/10.1007/s11747-024-01073-2

96 Rang Wang, Sylvia Chan-Olmsted & Qi Zhou, "Information seeking and affective relationship building in influencer marketing: the role of social media affordances", Humanities and Social Sciences Communications, 20 April 2025, https://www.nature.com/articles/s41599-025-04882-0

97 Jeff Horwitz, "Meta is earning a fortune on a deluge of fraudulent ads, documents show", Reuters, 6 November 2025, https://www.reuters.com/investigations/meta-is-earning-fortune-deluge-fraudulent-ads-documents-show-2025-11-06/; Jeff Horwitz, Engen Tham, "Meta tolerates rampant ad fraud from China to safeguard billions in revenue", Reuters, 15 December 2025, https://www.reuters.com/investigations/meta-tolerates-rampant-ad-fraud-china-safeguard-billions-revenue-2025-12-15/

98 Joe Tidy, "AI 'slop' is transforming social media - and a backlash is brewing", BBC News, 2 February 2026, https://www.bbc.co.uk/news/articles/c9wx2dz2v44o

99 Thomas Copeland, "AI-generated Iran war videos surge as creators use new tech to cash in", BBC News, 27 March 2026, https://www.bbc.co.uk/news/articles/ckg8wvz427vo

100 Who Targets Me?, Science Feedback, "Online misinformation pays. Why? Taking stock of a broad range of evidence", March 2025, https://science.feedback.org/wp-content/uploads/2025/04/AdFUND-Misinformation-in-Europe.pdf

101 Maldita, "TikTok is financing polarization in Europe and elsewhere", January 2026, https://files.maldita.es/maldita/uploads/2026/01/6973f8eda0e1b.pdf

102 Meriem Mahdhi, Niamh McIntyre, Mark Sellman, "King of slop: how anti-migrant AI content made one influencer rich", The Bureua of Investigative Journalism, 16 November 2025, https://www.thebureauinvestigates.com/stories/2025-11-16/king-of-slop-how-anti-migrant-ai-content-made-one-sri-lankan-influencer-rich; Aisha Down, "More than 20% of videos shown to new YouTube users are "AI slop", Guardian, 27 December 2025, https://www.theguardian.com/technology/2025/dec/27/more-than-20-of-videos-shown-to-new-youtube-users-are-ai-slop-study-finds; Michael Savage, "Fake anti-Labour videos viewed 1.2bn times", Guardian, 13 December 2025 https://www.theguardian.com/technology/2025/dec/13/fake-anti-labour-video-billion-views-youtube-2025

103 HM Government, Response to report on social media, misinformation and harmful algorithms, 17 October 2025, https://committees.parliament.uk/publications/49793/documents/266872/default/

104 Science, Innovation and Technology Committee, "Social media, misinformation and harmful algorithms", 11 July 2025, https://committees.parliament.uk/publications/48745/documents/258221/default/

105 Alexandra Topping, "Collapse of local media leaves us all in the dark" Guardian, 2 February 2024, https://www.theguardian.com/uk-news/2024/feb/02/collapse-of-local-media-leaves-us-all-in-the-dark

106 Humeyra Pamuk, "Exclusive: Trump administration orders enhanced vetting for applicants of H-1B visa", Reuters, 4 December 2025, https://www.reuters.com/world/us/trump-administration-orders-enhanced-vetting-applicants-h-1b-visa-2025-12-04/

107 George Wright, "UK social media campaigners among five denied US visas", BBC News, 24 December 2025, https://www.bbc.co.uk/news/articles/cp39kngz008o; Pippa Crerar, "British campaigner launches legal challenge against Trump administration after deportation threat", Guardian, 25 December 2025, https://www.theguardian.com/world/2025/dec/25/british-campaigner-llegal-challenge-trump-administration-deportation

108 Steven Lee Myers, "Trump Administration Cancels Scores of Grants to Study Online Misinformation", New York Times, 15 May 2025, https://www.nytimes.com/2025/05/15/business/trump-online-misinformation-grants.html; Amy Mackinnon, "US ends international push to combat fake news from hostile states", FT, 8 September 2025 https://www.ft.com/content/d31b56e3-aca9-4ee7-af5a-abec74830455

109 Google Search Central Blog, "Simplifying the search results page", 12 June 2025, https://developers.google.com/search/blog/2025/06/simplifying-search-results

110 Michael Savage, "AI summaries causing "devastating" drop in online news audiences", Guardian, 24 July 2025, https://www.theguardian.com/technology/2025/jul/24/ai-summaries-causing-devastating-drop-in-online-news-audiences-study-finds; Rob Waugh, "Google promotes fake content to millions on Discover news platform", Press Gazette, 28 October 2025, https://pressgazette.co.uk/platforms/google-promotes-fake-content-to-millions-on-discover-news-platform/; Charlotte Tobitt, "Google told to 'stop the BS' as it claims AI has not harmed website clickthrough", Press Gazette, 7 August 2025, https://pressgazette.co.uk/platforms/google-search-clicks-traffic-2025-ai-overviews/

111 Roa Powell and Carsten Jung, "Revealed: ChatGPT draws more on GB News, Al Jazeera, and Marie Claire than the BBC, IPPR analysis shows", IPPR, 30 January 2026, https://www.ippr.org/media-office/revealed-chatgpt-draws-more-on-gb-news-al-jazeera-and-marie-claire-than-the-bbc-ippr-analysis-shows

112 HM Government, "Protecting What Matters: Towards a more confident, cohesive, and resilient United Kingdom", 28 April 2026, https://www.gov.uk/government/publications/protecting-what-matters-towards-a-more-confident-cohesive-and-resilient-united-kingdom/protecting-what-matters-towards-a-more-confident-cohesive-and-resilient-united-kingdom

113 Adeline Hulin, "How does media and information literacy need to step up its game in the AI era?", World Economic Forum, 24 October 2025, https://www.weforum.org/stories/2025/10/media-information-literacy-ai/; UNESCO, "AI can make mistakes: Why media literacy matters more than ever", 24 October 2025, https://www.unesco.org/en/articles/ai-can-make-mistakes-why-media-literacy-matters-more-ever

114 House of Lords Communications and Digital Committee, Media literacy (3rd Report of Session 2024–25, HL Paper 163), published 25 July 2025

115 HM Government, Response to the House of Lords Communications and Digital Committee report on media literacy, 9 October 2025 https://committees.parliament.uk/publications/49637/documents/265648/default/

116 Department for Education, Government response to the Curriculum and Assessment Review, November 2025, https://assets.publishing.service.gov.uk/media/690b2a4a14b040dfe82922ea/Government_response_to_the_Curriculum_and_Assessment_Review.pdf

117 HM Government, "A Safe, Informed Digital Nation", 16 March 2025, https://www.gov.uk/government/publications/a-safe-informed-digital-nation/a-safe-informed-digital-nation

118 Ofcom, "Best Practice Design Principles for Media Literacy", 5 December 2025, https://www.ofcom.org.uk/media-use-and-attitudes/media-literacy/best-practice-design-principles-for-media-literacy

119 Ofcom, "How to promote Media Literacy: Consultation on recommendations for online platforms, broadcasters and services" 15 September 2025, https://www.ofcom.org.uk/media-use-and-attitudes/media-literacy/how-to-promote-media-literacy-consultation-on-recommendations-for-online-platforms-broadcasters-and-services

120 World Economic Forum, "Rethinking Media Literacy: A New Ecosystem Model for Information Integrity", July 2025, https://reports.weforum.org/docs/WEF_Rethinking_Media_Literacy_2025.pdf

121 Full Fact, Internet Matters, "Preparing young people to vote in a complex, attention-driven information environment", 10 February 2026, https://fullfact.org/documents/405/Briefing_-_votes_at_16_and_media_literacy.pdf

122 Full Fact, "Full Fact report 2025", May 2025, https://fullfact.org/policy/reports/full-fact-report-2025/

123 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

124 Electoral Commission, "Public Attitudes 2024", 7 May 2024, https://www.electoralcommission.org.uk/research-reports-and-data/public-attitudes/public-attitudes-2024

125 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

126 Ipsos, "Disinformation, hacking seen as top threats; Reputation of America drops", 27 November 2025, https://www.ipsos.com/en-nl/disinformation-hacking-seen-top-threats-reputation-america-drops

127 Maria Pawelec, "Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions", National Library of Medicine, 8 September 2022, https://pmc.ncbi.nlm.nih.gov/articles/PMC9453721/

128 Sky News, "Police chief apologises for 'erroneous' Maccabi Tel Aviv fan ban evidence, blaming AI", 14 January 2026, https://news.sky.com/story/police-chief-apologises-for-erroneous-maccabi-tel-aviv-fan-ban-evidence-blaming-ai-13494040

129 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

130 Saifuddin Ahmed, "Why cynics disengage: the nexus of political cynicism, misinformation, and online political participation", Asian Journal of Communication, 30 July 2025, https://www.tandfonline.com/doi/full/10.1080/01292986.2025.2538142

131 Ruolan Deng, Saifuddin Ahmed, "Between trust and skepticism: unpacking the impact of social media skepticism on online political participation", Behaviour and Information Technology, April 2025, https://www.researchgate.net/publication/390656581_Between_trust_and_skepticism_unpacking_the_impact_of_social_media_skepticism_on_online_political_participation

132 Maria Pawelec, "Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions", National Library of Medicine, 8 September 2022, https://pmc.ncbi.nlm.nih.gov/articles/PMC9453721/

133 Sora Park, Caroline Fisher et al, "The relationship between news trust, mistrust and audience disengagement", 20 November 2024, https://journals.sagepub.com/doi/10.1177/14648849241299775

134 OECD, "Building Trust and Reinforcing Democracy Preparing the Ground for Government Action", 17 November 2022, https://www.oecd.org/en/publications/building-trust-and-reinforcing-democracy_76972a4a-en/full-report/component-4.html

135 Pamela M Allen, John A Edwards, Frank J Snyder, Kevin A Makinson, David M Hamby, "The effect of cognitive load on decision making with graphically displayed uncertainty information", August 2024, https://pmc.ncbi.nlm.nih.gov/articles/PMC4063894/; Adam Szulewski, Daniel Howes, Jeroen J G van Merriënboer, John Sweller, "From Theory to Practice: The Application of Cognitive Load Theory to the Practice of Medicine", Academic Medicine, 29 December 2020, https://academic.oup.com/academicmedicine/article-abstract/96/1/24/8346681; https://www.mdpi.com/2076-3425/16/1/109

136 Pramukh Nanjundaswamy Vasist, Debashis Chatterjee, Satish Krishnan, "The Polarizing Impact of Political Disinformation and Hate Speech: A Cross-country Configural Narrative", April 2023, https://pmc.ncbi.nlm.nih.gov/articles/PMC10106894

137 "Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity", OECD, 4 March 2024, https://www.oecd.org/en/publications/facts-not-fakes-tackling-disinformation-strengthening-information-integrity_d909ff7a-en.html

138 Khaled Mansour, Guest essay for Full Report 2025, May 2025, https://fullfact.org/policy/reports/full-fact-report-2025/#chapter-1-third-party-fact-checking-with-meta

139 Adejumo Kabir, "How disinformation fuels Nigeria's farmer-herder conflict", 8 July 2025, https://africainfact.com/how-disinformation-fuels-nigerias-farmer-herder-conflict/

140 Centre for Countering Digital Hate, "Fuelling hate: one year after the UK summer riots, X still lets calls for violence spread unchecked", July 2025, https://counterhate.com/wp-content/uploads/2025/07/Fuelling-Hate_FINAL.pdf

141 Full Fact, "Full Fact report 2025", May 2025, https://fullfact.org/policy/reports/full-fact-report-2025/; HMICFRS, "An inspection of the police response to the public disorder in July and August 2024: Tranche 2", 7 May 2025, https://hmicfrs.justiceinspectorates.gov.uk/publication-html/police-response-to-public-disorder-in-july-and-august-2024-tranche-2/

142 Speaker's Conference on the security of candidates, MPs and elections, "second report", 27 October 2025, https://committees.parliament.uk/publications/49841/documents/268474/default

143 Hansard, 2 March 2026, Debate on the Representation of the People Bill https://hansard.parliament.uk/Commons/2026-03-02/debates/702183A9-7B57-487C-9634-BF459DB65DB2/RepresentationOfThePeopleBill

144 Foreign Affairs Committee, "Oral evidence: Disinformation Diplomacy", 13 January 2026, https://committees.parliament.uk/oralevidence/16988/pdf/

145 HM Government, "Protecting What Matters: Towards a more confident, cohesive, and resilient United Kingdom", 28 April 2026, https://www.gov.uk/government/publications/protecting-what-matters-towards-a-more-confident-cohesive-and-resilient-united-kingdom/protecting-what-matters-towards-a-more-confident-cohesive-and-resilient-united-kingdom

146 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

147 Electoral Commission, "Public Opinion Tracker 2023", 11 December 2023, https://www.electoralcommission.org.uk/who-we-are-and-what-we-do/our-views-and-research/our-research/public-attitudes-2023; Ipsos, "Trust in politicians reaches its lowest score in 40 years", 14 December 2023, https://www.ipsos.com/en-uk/ipsos-trust-in-professions-veracity-index-2023; Office for National Statistics, "Trust in government, UK: 2023", 1 March 2024, https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/bulletins/trustingovernmentuk/2023; Royal Holloway University of London, "Survey finds that nearly half of young people are unhappy with UK democracy", 4 July 2024, https://www.royalholloway.ac.uk/about-us/news/survey-finds-that-nearly-half-of-young-people-are-unhappy-with-uk-democracy/

148 Leo Benedictus, "Revealed: how academics are being deepfaked on TikTok and Instagram to promote supplements" Full Fact, 5 December 2025, https://fullfact.org/health/academics-deepfaked-tiktok-wellness-nest/

149 Sam Stockwell, "From Deepfake Scams to Poisoned Chatbots: AI and Election Security in 2025", Centre for Emerging Technology and Security, https://cetas.turing.ac.uk/publications/deepfake-scams-poisoned-chatbots

150 Nuurrianti Jalli, "Reframing misinformation as informational-systemic risk in the age of societal volatility", 22 December 2025, Harvard Kennedy School, https://misinforeview.hks.harvard.edu/article/reframing-misinformation-as-informational-systemic-risk-in-the-age-of-societal-volatility

151 Philip Rycroft, "Report of the Independent Review into Countering Foreign Financial Influence and Interference in UK Politics", 25 March 2026, https://assets.publishing.service.gov.uk/media/69c29f84b920af63be1c7777/The_Rycroft_Review_Report_standard_version.pdf

152 Full Fact, "Full Fact Report 2023", 2023, https://fullfact.org/media/uploads/full-fact-report-2023.pdf

153 Felix Simon, Rasmus Kleis Nielsen, Richard Fletcher, "Generative AI and news report 2025: How people think about AI's role in journalism and society" Reuters Institute, 7 October 2025, https://reutersinstitute.politics.ox.ac.uk/generative-ai-and-news-report-2025-how-people-think-about-ais-role-journalism-and-society

154 AI Security Institute, "International AI Safety Report 2026", 3 February 2026, https://internationalaisafetyreport.org/publication/international-ai-safety-report-2026

155 Government Office for Science, "Deepfakes and media literacy", 27 May 2025, https://www.gov.uk/government/publications/deepfakes-and-media-literacy/deepfakes-and-media-literacy

156 Ofcom, "Future Technology and Media Literacy: Understanding Generative AI", 22 February 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-literacy-understanding-generative-ai?v=330961&

157 BBC News, "French election: Emmanuel Macron condemns 'massive' hack attack", 6 May 2017, https://www.bbc.co.uk/news/world-europe-39827244

158 Sam Stockwell, Ardi Janjeva, Broderick McDonald, "Adding Fuel to the Fire: AI Information Threats and Crisis Events", 11 February 2026,https://cetas.turing.ac.uk/publications/adding-fuel-to-fire

159 Local Government Association, "Pre-election period", https://www.local.gov.uk/our-support/communications-and-community-engagement/pre-election-period

160 Morgan Meaker, "Slovakia's Election Deepfakes Show AI Is a Danger to Democracy", 3 October 2023, https://www.wired.com/story/slovakias-election-deepfakes-show-ai-is-a-danger-to-democracy/

161 Jake Liggett, "'Disgraceful' deep-fake AI video condemned by presidential candidate", BBC News, 22 October 2025, https://www.bbc.co.uk/news/articles/czxkn504lqpo

162 Daniel Thilo Schroeder et al, "How malicious AI swarms can threaten democracy", Science, 22 January 2026 https://www.science.org/doi/10.1126/science.adz1697; David Gilbert, "AI-Powered Disinformation Swarms Are Coming for Democracy", Wired, 22 January 2026, https://www.wired.com/story/ai-powered-disinformation-swarms-are-coming-for-democracy/

163 Harrison Ostridge, "Potential future risks from autonomous AI systems", House of Lords Library, 5 January 2026, https://lordslibrary.parliament.uk/potential-future-risks-from-autonomous-ai-systems/

164 Philipp Darius, Johannes Breuer, Simon Kruschinski, Felicia Loecherbach, Jasmin Riedl, Sebastian Stier, "Election research in the age of regulated data access under the EU Digital Services Act", Internet Policy Review, 16 February 2026, https://policyreview.info/articles/analysis/election-research-data-access-dsa

165 Hakan Ersen, "German court orders X to grant data access for Hungary election research" Reuters, 18 February 2026, https://www.reuters.com/legal/litigation/german-court-orders-x-grant-data-access-hungary-election-research-2026-02-18/

166 Terra Rolfe, Helena Schwertheim, Melanie Döring, Ellen Jacobs, "Safeguarding Elections in the Digital Age", Institute for Strategic Dialogue, September 2024, https://www.isdglobal.org/wp-content/uploads/2024/09/Safeguarding-Elections-in-the-Digital-Age.pdf

167 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

168 Electoral Commission, "Public Attitudes 2024", 7 May 2024, https://www.electoralcommission.org.uk/research-reports-and-data/public-attitudes/public-attitudes-2024

169 Paul D Biddinger, Elena Savoia, Sarah B Massin-Short, Jessica Preston, Michael A Stoto, "Public Health Emergency Preparedness Exercises: Lessons Learned", Public Health Reports, https://pmc.ncbi.nlm.nih.gov/articles/PMC2966651/

170 National Cyber Security Centre, "Risk Management", https://www.ncsc.gov.uk/collection/risk-management/using-cyber-security-scenariosrusi

171 Sam Trendall, "Government extends use of digital simulation for 'information incident' crisis training", Public Technology, 1 July 2024, https://www.publictechnology.net/2024/07/01/education-and-skills/government-extends-use-of-digital-simulation-for-information-incident-crisis-training/

172 Ofcom, "Supporting and harnessing AI innovation safely", 6 June 2025, https://www.ofcom.org.uk/about-ofcom/annual-reports-and-plans/supporting-and-harnessing-ai-innovation-safely

173 Foreign Affairs Committee, "Oral evidence: Disinformation Diplomacy", 13 January 2026, https://committees.parliament.uk/oralevidence/16988/pdf/

174 OSCE, "Moldova's parliamentary elections were competitive but campaign marred by cyberattacks, illegal funding and disinformation, international observers say", 29 September 2025, https://www.oscepa.org/en/news-a-media/press-releases/2025/moldovas-parliamentary-elections-were-competitive-but-campaign-marred-by-cyberattacks-illegal-funding-and-disinformation-international-observers-say

175 OECD, "The OECD Reinforcing Democracy Initiative", 7 October 2024, https://www.oecd.org/en/publications/the-oecd-reinforcing-democracy-initiative_9543bcfb-en/full-report/component-4.html; https://arxiv.org/abs/2504.12537

176 AI Security Institute, "International AI Safety Report 2026", 3 February 2026, https://internationalaisafetyreport.org/publication/international-ai-safety-report-2026

177 European Commission, "AI Act", https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

178 United Nations, "Global Risks Report 2024", 2024, https://unglobalriskreport.org/UNHQ-GlobalRiskReport-WEB-FIN.pdf

179 World Economic Forum, "The Global Risks Report 2025", January 2025, https://reports.weforum.org/docs/WEF_Global_Risks_Report_2025.pdf

180 Julio Terracino, "Strengthening democracy: Gearing up governments to tackle mis- and disinformation", OECD, 17 November 2022, https://www.oecd.org/en/blogs/2022/11/trengthening-democracy--gearing-up-governments-to-tackle-mis--an.html

181 Council of Europe, "RESIST – Strengthening Societal Resilience to Disinformation in Europe", 18 September 2025, https://www.coe.int/en/web/freedom-expression/resist-strengthening-societal-resilience-to-disinformation-in-europe

182 Liaison Committee, "Oral evidence: Work of the Prime Minister", 21 July 2025, https://committees.parliament.uk/oralevidence/16355/pdf/

183 Letisha Lunin, "The UK is deeply concerned by the growing threats to information integrity: UK Statement at the UN Fourth Committee", Foreign, Commonwealth and Development Office, 5 November 2025, https://www.gov.uk/government/speeches/the-uk-is-deeply-concerned-by-the-growing-threats-to-information-integrity-uk-statement-at-the-un-fourth-committee

184 Letisha Lunin, "The United Kingdom remains steadfast in its commitment to media freedom around the world: UK statement at the UN", Foreign, Commonwealth and Development Office, 28 April 2026, https://www.gov.uk/government/speeches/the-united-kingdom-remains-steadfast-in-its-commitment-to-media-freedom-around-the-world-uk-statement-at-the-un

185 Marina Adami, "How Rest of World is tracking AI use around elections worldwide", Reuters Institute, 13 May 2024,

186 International IDEA (2025) Review of the 2024 Super-Cycle Year of Elections Trends, Challenges and Opportunities

187 Sayash Kapoor, Arvind Narayanan, "We Looked at 78 Election Deepfakes. Political Misinformation Is Not an AI Problem", Knight First Amendment Institute, 13 December 2024, https://knightcolumbia.org/blog/we-looked-at-78-election-deepfakes-political-misinformation-is-not-an-ai-problem

188 Shanze Hasan, "The Effect of AI on Elections Around the World and What to Do About It", Brennan Center for Justice, 6 June 2024, https://www.brennancenter.org/our-work/analysis-opinion/effect-ai-elections-around-world-and-what-do-about-it

189 Lakmusz, "Fearmongering with AI-generated videos, manipulated speeches and Péter Magyar's Trump moment", EDMO, 24 February 2026, https://edmo.eu/publications/fearmongering-with-ai-generated-videos-manipulated-speeches-and-peter-magyars-trump-moment/

190 Institute for Strategic Dialogue, "Irish presidential election 2025: Renewed attacks on election integrity and repeated platform failures", 31 October 2025, https://www.isdglobal.org/digital_dispatches/irish-presidential-election-2025-renewed-attacks-on-election-integrity-and-repeated-platform-failures/; Jake Liggett, "'Disgraceful' deep-fake AI video condemned by presidential candidate", BBC News, 22 October 2025, https://www.bbc.co.uk/news/articles/czxkn504lqpo

191 Alliance 4 Europe, "Foreign Information Manipulation in the 2025 German Federal Election", July 2025, https://alliance4europe.eu/foreign-information-manipulation-2025-german-federal-elections

192 Sam Stockwell, "From Deepfake Scams to Poisoned Chatbots: AI and Election Security in 2025", Centre for Emerging Technology and Security, 17 November 2025, https://cetas.turing.ac.uk/publications/deepfake-scams-poisoned-chatbots

193 Leyland Cecco, "Dramatic rise in fake political content on social media as Canada prepares to vote" 18 April 2025, https://www.theguardian.com/world/2025/apr/18/canada-fake-political-content-social-media; Victor Livernoche, Andreea Musulan, Zachary Yang, Jean-François Godbout, Reihaneh Rabbany, "Deepfakes in the 2025 Canadian Election: Prevalence, Partisanship, and Platform Dynamics", Cornell University, 15 December 2025 https://arxiv.org/abs/2512.13915

194 Servet Yanatma, "Perceived exposure to disinformation and fake news rising in Europe: Which countries suffer most?", Euro News, 2 February 2026, https://www.euronews.com/next/2026/02/02/perceived-exposure-to-disinformation-and-fake-news-rising-in-europe-which-countries-suffer

195 European Parliament, "Tackling deepfakes in Europe", July 2021, https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690039/EPRS_STU(2021)690039_EN.pdf

196 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

197 Vijay Rangarajan, "Additional Electoral Commission information for the Foreign Affairs Committee", Electoral Commission, 28 January 2026, https://committees.parliament.uk/writtenevidence/162487/html/

198 Joint Committee on the National Security Strategy, "Letter to the Prime Minister, Rt Hon Rishi Sunak MP", 23 May 2024, https://committees.parliament.uk/publications/45032/documents/223340/default/

199 Blaise Metreweli, "Speech by Blaise Metreweli, Chief of SIS, 15 December 2025", Secret Intelligence Service, 15 December 2025, https://www.gov.uk/government/speeches/speech-by-blaise-metreweli-chief-of-sis-15-december-2025

200 Alexandra Topping, "UK politics 'constantly suffering' from online disinformation, says Labour MP", Guardian, 18 January 2026, https://www.theguardian.com/politics/2026/jan/18/uk-politics-constantly-suffering-from-online-disinformation-says-labour-mp-emily-thornberry

201 Philip Rycroft, "Report of the Independent Review into Countering Foreign Financial Influence and Interference in UK Politics", 25 March 2026, https://assets.publishing.service.gov.uk/media/69c29f84b920af63be1c7777/The_Rycroft_Review_Report_standard_version.pdf

202 OECD, "Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity", 4 March 2024 https://www.oecd.org/en/publications/facts-not-fakes-tackling-disinformation-strengthening-information-integrity_d909ff7a-en/full-report/component-5.html

203 Institute for Strategic Dialogue, "Link by link: Hundreds of webpages cite pro-Russia Pravda network", 18 November 2025, https://www.isdglobal.org/digital_dispatches/link-by-link-hundreds-of-webpages-cite-pro-russia-pravda-network/.

204 NewsGuard, "RT Celebrates Anniversary as Kremlin's 'Propaganda Bullhorn'", 30 October 2025, https://www.newsguardrealitycheck.com/p/rt-celebrates-anniversary-as-kremlins-propaganda-bullhorn; Will Oremus, "A Russian fake news ring was struggling. Then it targeted USAID", Washington Post, 6 May 2025 https://www.washingtonpost.com/politics/2025/05/06/operation-overload-russian-disinfo-usaid-musk-x; Kiran Stacey, "Foreign states using AI videos to undermine support for Ukraine, says Yvette Cooper", Guardian, 8 December 2025, https://www.theguardian.com/politics/2025/dec/08/foreign-states-ai-videos-support-ukraine-yvette-cooper

205 European External Action Service, "4th EEAS Report on Foreign Information Manipulation and Interference Threats", March 2026, https://www.eeas.europa.eu/sites/default/files/2026/documents/EEAS%204th%20Threat%20Report_web%20version_1.pdf

206 Allie Funk, Adrian Shahbaz, Kian Vesteinsson, "Freedom on the Net 2023, The Repressive Power of Artificial Intelligence", Freedom House, 2023, https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence

207 Naomi Nix, Ian Duncan, "Inside the Trump administration's campaign to counter content bans in Europe", Washington Post, 20 March 2026, https://www.washingtonpost.com/technology/2026/03/20/trump-eu-dsa-censorship; Simon Lewis, Humeyra Pamuk, "US officials spotlight European 'censorship' despite concerns over free speech at home", Reuters, 25 September 2025, https://www.reuters.com/world/us/us-officials-spotlight-european-censorship-despite-concerns-over-free-speech-2025-09-25; Michael Savage, Daniel Boffey, Ben Quinn, "US officials challenge Ofcom over online safety laws' impact on free speech", Guardian, 1 April 2025, https://www.theguardian.com/media/2025/apr/01/us-officials-challenge-ofcoms-risk-to-free-speech-caused-by-online-safety-laws

208 US State Department, "2024 Country Reports on Human Rights Practices: United Kingdom", 13 August 2025 https://www.state.gov/reports/2024-country-reports-on-human-rights-practices/united-kingdom/

209 House Judiciary Committee, "Europe's Threat to American Speech and Innovation", 3 September 2025, https://judiciary.house.gov/committee-activity/hearings/europes-threat-american-speech-and-innovation; Committee on the Judiciary of the US House of Representatives, "The foreign censorship threat: how the European Union's Digital Services Act compels global censorship and infringes on American free speech, 25 July 2025 https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/2025-07/DSA_Report&Appendix(07.25.25).pdf

210 Humeyra Pamuk, 7 August 2025, "Rubio orders US diplomats to launch lobbying blitz against Europe's tech law", Reuters, https://www.reuters.com/sustainability/society-equity/rubio-orders-us-diplomats-launch-lobbying-blitz-against-europes-tech-law-2025-08-07/; Jim Jordan, thread on X, 23 September 2025, https://x.com/Jim_Jordan/status/1970474487809265878; Jim Jordan, thread on X, 30 July 2025, https://x.com/jim_jordan/status/1950368307372020086

211 Helen Corbett, "Government defends Online Safety Act after X claims it threatens free speech", Independent. 2 August 2025, https://www.independent.co.uk/news/uk/politics/government-peter-kyle-twitter-reform-uk-failure-b2800787.html; Mark Sellman, "Tech billionaire gives No 10 free speech warning over Online Safety Act", The Times, 8 August 2025, https://www.thetimes.com/uk/technology-uk/article/tech-billionaire-gives-no-10-free-speech-warning-over-online-safety-act-cpxbwnvxj; Adam Satariano, Lizzie Dearden, "Has Britain Gone Too Far With Its Digital Controls?", New York Times, 17 September 2025, https://www.nytimes.com/2025/09/17/technology/britain-facial-recognition-digital-controls.html

212 Donald Trump, post on Truth Social, 26 August 2025, https://truthsocial.com/@realDonaldTrump/posts/115092243259973570; Federal Trade Commission, "FTC Chairman Ferguson Warns Companies Against Censoring or Weakening the Data Security of Americans at the Behest of Foreign Powers", 21 August 2025, https://www.ftc.gov/news-events/news/press-releases/2025/08/ftc-chairman-ferguson-warns-companies-against-censoring-or-weakening-data-security-americans-behest

213 BGR Group, "House Committee on the Judiciary – Europe's Threat to American Speech and Innovation", 3 September 2025, https://bgrdc.com/wp-content/uploads/2025/09/09.03.25-House-Committee-on-the-Judiciary-Europes-Threat-to-American-Speech-and-Innovation.pdf

214 Robert Booth, Lisa O'Caroll, "Meta found in breach of EU law over 'ineffective' complaints system for flagging illegal content", Guardian, https://www.theguardian.com/technology/2025/oct/24/instagram-facebook-breach-eu-law-content-flagging

215 European Commission, "Commission preliminarily finds TikTok and Meta in breach of their transparency obligations under the Digital Services Act", 24 October 2025, https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2503

216 Pieter Haeck and Eliza Gkritsi, "EU slaps €120M fine on Elon Musk's X, straining ties with US", Politico, 5 December 2025, https://www.politico.eu/article/eu-slaps-e120m-fine-on-x-straining-ties-with-us/; Steve Peers, "The Digital Service's Act Main Character: the EU Commission finally fines X", EU Law Analysis, 7 December 2025 https://eulawanalysis.blogspot.com/2025/12/the-digital-services-act-main-character.html?m=1

217 Zoe Crowther, "Ofcom Launches Investigation Into Major Social Media Platforms Over Illegal Hate Content", Politics Home, 2 December 2025, https://www.politicshome.com/news/article/ofcom-launches-investigation-social-media-platforms

218 Ofcom, "Ofcom launches investigation into X over Grok sexualised imagery", 12 January 2026, https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/ofcom-launches-investigation-into-x-over-grok-sexualised-imagery

219 Bobby Allyn, "Jury finds Meta and Google negligent in social media harms trial", NPR, 25 March 2026, https://www.npr.org/2026/03/25/nx-s1-5746125/meta-youtube-social-media-trial-verdict

220 New Mexico Department of Justice, "New Mexico Department of Justice Wins Landmark Verdict Against Meta", 24 March 2026, https://nmdoj.gov/press-release/new-mexico-department-of-justice-wins-landmark-verdict-against-meta

221 Michigan Department of Attorney General, "Bipartisan Coalition of Attorneys General File Lawsuits Against Meta for Harming Youth Mental Health Through Its Social Media Platforms", 24 October 2023, https://www.michigan.gov/ag/news/press-releases/2023/10/24/bipartisan-coalition-of-attorneys-general-file-lawsuits-against-meta

222 European Commission, "Commission publishes guidelines under the DSA for the mitigation of systemic risks online for elections", 25 March 2024, https://ec.europa.eu/commission/presscorner/detail/en/ip_24_1707

223 European Commission, "Guidelines for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes" 26 April 2024, https://digital-strategy.ec.europa.eu/en/library/guidelines-providers-vlops-and-vloses-mitigation-systemic-risks-electoral-processes

224 Reporters Without Borders, "Access to reliable information: EU member states open the door to measures strengthening the visibility of journalism on online platforms", 28 November 2025, https://rsf.org/en/access-reliable-information-eu-member-states-open-door-measures-strengthening-visibility-journalism

225 OECD, "Recommendation of the Council on Information Integrity" 17 December 2024, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0505; OECD, "Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity", 4 March 2024, https://www.oecd.org/en/publications/facts-not-fakes-tackling-disinformation-strengthening-information-integrity_d909ff7a-en.html

226 United Nations, "United Nations Global Principles for Information Integrity: Recommendations for Multi-stakeholder Action", June 2024, https://www.un.org/en/information-integrity/global-principles

227 Council of Europe, "Building societal resilience to disinformation: Launch of the RESIST project", 18 September 2025, https://www.coe.int/en/web/freedom-expression/-/building-societal-resilience-to-disinformation-launch-of-the-resist-project; Council of Europe, "Roundtable discussion on the topic "Fake or fact? How to combat disinformation in democratic societies"", 16 October 2025, https://www.coe.int/en/web/deputy-secretary-general/-/roundtable-discussion-on-the-topic-fake-or-fact-how-to-combat-disinformation-in-democratic-societies-

228 Government of Canada, "Critical Election Incident Public Protocol", February 2023, https://www.canada.ca/en/democratic-institutions/news/2023/02/critical-election-incident-public-protocol.html

229 Morris Rosenberg, "Report on the Assessment of the 2021 Federal General Election", 2023, https://www.canada.ca/en/democratic-institutions/services/reports/report-assessment-2021-critical-election-incident-public-protocol.html

230 Full Fact, written evidence to the Public Administration and Constitutional Affairs Committee, August 2021, https://committees.parliament.uk/writtenevidence/38455/pdf/

231 Government of Canada, "The plan to protect Canada's democracy", updated March 2026, https://www.canada.ca/en/democratic-institutions/services/protecting-democracy.html

232 Australian Electoral Commission, "Electoral Integrity Assurance Taskforce", https://www.aec.gov.au/about_aec/electoral-integrity.htm

233 Australian Electoral Commission, "Electoral process disinformation register", 2025, https://www.aec.gov.au/media/disinformation-register-2025.htm

234 Australian Government, Department of Home Affairs, "Countering Foreign Interference", https://www.homeaffairs.gov.au/about-us/our-portfolios/national-security/countering-foreign-interference

235 New Zealand Government and Electoral Commission, "Principles and protocols for the GCSB and the NZSIS in relation to the 2023 General Election", August 2023, https://www.nzic.govt.nz/assets/Uploads/Principles-and-protocols-for-the-GCSB-and-the-NZSIS-in-relation-to-the-2023-General-Election.pdf

236 New Zealand Government and Electoral Commission, "Protocol on the management of election disruptions", August 2023, https://elections.nz/assets/2023-General-Election/Election-protocols/Protocol-on-the-management-of-election-disruptions.pdf

237 New Zealand Government and Electoral Commission, "Protocol on communications related to the 2023 General Election process", August 2023, https://elections.nz/assets/2023-General-Election/Election-protocols/Protocol-on-communications-related-to-the-2023-General-Election-process.pdf

238 New Zealand Security Intelligence Service, "An assessment by the New Zealand Security Intelligence Service", 2024, https://www.nzsis.govt.nz/assets/NZSIS-Documents/New-Zealands-Security-Threat-Environment-2025.pdf

239 US Department of Homeland Security, "Statement by Secretary Jeh Johnson on the Designation of Election Infrastructure as a Critical Infrastructure Subsector", 6 January 2017, https://www.dhs.gov/archive/news/2017/01/06/statement-secretary-johnson-designation-election-infrastructure-critical

240 Cybersecurity and Infrastructure Security Agency, "Election Infrastructure Incident Response Communications Guide", October 2024, https://www.cisa.gov/sites/default/files/2024-10/Election-Infrastructure-Incident-Response-Communications-Guide-508.pdf

241 Cybersecurity and Infrastructure Security Agency, "Election Security Resource Library", https://www.cisa.gov/topics/election-security/election-security-resource-library

242 Office of the Director of National Intelligence, "Foreign threats to US elections after voting ends in 2024", 22 October 2024, https://www.dni.gov/index.php/newsroom/reports-publications/reports-publications-2024/4006-foreign-threats-after-voting-ends

243 Jeremy Herb, "Trump to DOJ last December: 'Just say that the election was corrupt + leave the rest to me'", 31 July 2021, https://edition.cnn.com/2021/07/30/politics/trump-election-justice/index.html

244 Ministry of Housing, Communities and Local Government, "Parliamentary question: election", 22 January 2025, https://questions-statements.parliament.uk/written-questions/detail/2025-01-08/HL3892

245 Theyworkforyou, "New Clause 26 - Critical election incident protocol", 16 April 2026, https://www.theyworkforyou.com/pbc/2024-26/Representation_of_the_People_Bill/09-0_2026-04-16a.370.0?s=

246 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

247 Philip Rycroft, "Report of the Independent Review into Countering Foreign Financial Influence and Interference in UK Politics", 25 March 2026, https://assets.publishing.service.gov.uk/media/69c29f84b920af63be1c7777/The_Rycroft_Review_Report_standard_version.pdf

248 C2PA Technical Working Group, "Content Credentials Explained: Addressing Common Questions and Updates", September 2025, https://c2pa.org/wp-content/uploads/sites/33/2025/10/content_credentials_wp_0925.pdf;

249 European Commission, "The General-Purpose AI Code of Practice", updated 23 April 2026, https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

250 Coalition for Content Provenance and Authenticity, "Membership, https://c2pa.org/membership/

251 Ryan Whitwam, "Google's SynthID AI watermarking tech is being adopted by OpenAI, Nvidia, and more", Ars Technica, 19 May 2026, https://arstechnica.com/google/2026/05/googles-synthid-ai-watermarking-tech-is-being-adopted-by-openai-nvidia-and-more/

252 Tech UK, "UK Home Office Deepfake Detection Challenge 2026", https://www.techuk.org/what-we-deliver/events/uk-home-office-deepfake-detection-challenge-2026.html

253 HM Government, "Government leads global fight against deepfake threats", 5 February 2026, https://www.gov.uk/government/news/government-leads-global-fight-against-deepfake-threats

254 Severin Carrell, "Software tackling deepfakes to be piloted for Scottish and Welsh elections", Guardian, 8 January 2026, https://www.theguardian.com/technology/2026/jan/08/pilot-software-tackle-deepfakes-scottish-welsh-elections; Electoral Commission, "Electoral Commission launches deepfake detection pilot to counter AI misinformation", 15 April 2026, https://www.electoralcommission.org.uk/media-centre/electoral-commission-launches-deepfake-detection-pilot-counter-ai-misinformation

255 Electoral Commission, "New advice for voters on disinformation, and for campaigners using generative AI", 17 June 2024, https://www.electoralcommission.org.uk/media-centre/new-advice-voters-disinformation-and-campaigners-using-generative-ai

256 Electoral Commission, "Response to the SIT Committee call for evidence on social media, misinformation and harmful algorithms", 28 January 2025, https://www.electoralcommission.org.uk/news-and-views/our-responses-consultations/response-science-innovation-and-technology-committee-call-evidence-social-media-misinformation-and

257 Hansard, debate on copyright and AI, 18 March 2026, https://hansard.parliament.uk/commons/2026-03-18/debates/26031818000010/CopyrightAndAI

258 Ofcom, "Deepfake Defences 2: The Attribution Toolkit", 11 July 2025, https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/information-for-industry/deepfake-defences-2/deepfake-defences-2---the-attribution-toolkit.pdf

259 Bobby Chesney, Danielle Citron, California Law Review, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security", December 2019, https://www.californialawreview.org/print/deep-fakes-a-looming-challenge-for-privacy-democracy-and-national-security

260 Mahsa Alimardani, "How Doubt Became a Weapon in Iran", The Atlantic, 14 January 2026, https://www.theatlantic.com/international/2026/01/iran-disinformation-ai-protests-doubt/685608/

261 Mahsa Alimardani, "The Fake Images of a Real Strike on a School", 13 March 2026, https://www.theatlantic.com/ideas/2026/03/ai-imagery-iran-war/686347/

262 Coalition for Content Provenance and Authenticity, "Advancing digital content transparency and authenticity", https://c2pa.org/

263 California Legislative Information, Senate Bill No. 942, 20 September 2024, https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB942

264 Mahsa Alimardani, Jacobo Castellanos, Bruna Santos, “India Bets on AI Detection. Every Regulator Should Watch What Happens Next”, Tech Policy Press, 18 February 2026, https://www.techpolicy.press/india-bets-on-ai-detection-every-regulator-should-watch-what-happens-next/

265 European Commission, “Commission publishes second draft of Code of Practice on Marking and Labelling of AI-generated content”, 5 March 2026, https://digital-strategy.ec.europa.eu/en/library/commission-publishes-second-draft-code-practice-marking-and-labelling-ai-generated-content

266 Shirin Anlen, Mahsa Alimardani, “How AI Content Detection is Being Weaponized in the Iran War”, Tech Policy Press, 17 March 2026, https://www.techpolicy.press/how-ai-content-detection-is-being-weaponized-in-the-iran-war/

267 Meta Oversight Board, “AI-Generated Video in Israel-Iran Conflict”, https://www.oversightboard.com/pc/ai-generated-video-in-israel-iran-conflict/

268 Who Targets Me, “UK campaign analysis: 19th-25th June”, 1 July 2024, https://whotargets.me/en/uk-campaign-analysis-19th-25th-june/

269 Lorraine Conway, “Who regulates political advertising?”, House of Commons Library, 4 November 2019, https://commonslibrary.parliament.uk/who-regulates-political-advertising/

270 Who Targets Me, “The disappearing ad library”, 22 May 2025, https://whotargets.me/en/the-disappearing-ad-library/

271 Full Fact, “Second reading briefing – Representation of the People Bill”, March 2026, https://fullfact.org/documents/409/Full_Fact_briefing_-_political_advert_library.pdf

272 European Parliament, “Parliament adopts new transparency rules for political advertising”, 27 February 2024, https://www.europarl.europa.eu/news/en/press-room/20240223IPR18071/parliament-adopts-new-transparency-rules-for-political-advertising

273 Elizabeth Dubois, “What have we learned from Google’s political ad pullout?”, Policy Options, 10 April 2019, https://policyoptions.irpp.org/2019/04/learned-googles-political-ad-pullout/

274 Anupriya Datta, “Meta to stop running political ads on Facebook and Instagram”, Euractiv, 25 July 2025, https://www.euractiv.com/news/meta-to-stop-running-political-ads-on-facebook-and-instagram/

275 Mark Scott, “In defense of digital political ads”, Digital Politics, 4 August 2025, https://www.digitalpolitics.co/newsletter057free/

276 Full Fact, “Second reading briefing – Representation of the People Bill”, March 2026, https://fullfact.org/documents/409/Full_Fact_briefing_-_political_advert_library.pdf

277 Electoral Commission, “Response to the Science, Innovation and Technology Committee call for evidence on social media, misinformation and harmful algorithms”, 28 January 2025, https://www.electoralcommission.org.uk/news-and-views/our-responses-consultations/response-science-innovation-and-technology-committee-call-evidence-social-media-misinformation-and

278 Committee on Standards in Public Life, “Regulating Election Finance”, July 2021, https://assets.publishing.service.gov.uk/media/60e460b1d3bf7f56801f3bf6/CSPL_Regulating_Election_Finance_Review_Final_Web.pdf

279 House of Lords Select Committee on Democracy and Digital Technologies, “Digital Technology and the Resurrection of Trust”, 29 June 2020, https://committees.parliament.uk/publications/1634/documents/17731/default/; Culture, Media and Sport Committee, “Disinformation and ‘fake news’”, 14 February 2019, https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf;

280 Philip Rycroft, “Report of the Independent Review into Countering Foreign Financial Influence and Interference in UK Politics”, 25 March 2026, https://assets.publishing.service.gov.uk/media/69c29f84b920af63be1c7777/The_Rycroft_Review_Report_standard_version.pdf

281 Hugo Drochon, Dan Lomas, Rory Cormac, “Defending Democracy’: Evidence to the JCNSS inquiry”, March 2024, https://www.nottingham.ac.uk/politics/documents/news-events/defending-democracy-jcnss-inquiry-crispi-suit.pdf

282 HM Government, “The Amber Book: Managing Crisis in Central Government”, April 2025, https://assets.publishing.service.gov.uk/media/680a4fbd6d6ac02ee99d8488/35.20_CO_Emergency_Response_and_Recovery_02_Amber_Book_FINAL_PRINT.pdf

283 Philip Rycroft, “Report of the Independent Review into Countering Foreign Financial Influence and Interference in UK Politics”, 25 March 2026, https://assets.publishing.service.gov.uk/media/69c29f84b920af63be1c7777/The_Rycroft_Review_Report_standard_version.pdf

284 William Dixon, “Why the UK Now Needs a National Disinformation Agency”, RUSI, 5 September 2025, https://www.rusi.org/explore-our-research/publications/commentary/why-uk-now-needs-national-disinformation-agency

285 Philip Rycroft, “Report of the Independent Review into Countering Foreign Financial Influence and Interference in UK Politics”, 25 March 2026, https://assets.publishing.service.gov.uk/media/69c29f84b920af63be1c7777/The_Rycroft_Review_Report_standard_version.pdf

286 Kanishka Narayan MP, “DSIT activity to defend UK democracy from disinformation”, DSIT, 16 January 2026, https://committees.parliament.uk/publications/51311/documents/284909/default/

287 Hansard, debate on Russian Influence on UK Politics and Democracy, 9 February 2026, https://hansard.parliament.uk/commons/2026-02-09/debates/F9F28AFA-E1F3-449C-B18B-63DDCABC411E/RussianInfluenceOnUKPoliticsAndDemocracy

288 Kanishka Narayan MP, “DSIT activity to defend UK democracy from disinformation”, Department for Science, Innovation and Technology, 16 January 2026, https://committees.parliament.uk/publications/51311/documents/284909/default/

289 National Security Online Information Team, “Draft Data Protection and Compliance Policy”, published by Big Brother Watch, https://bigbrotherwatch.org.uk/wp-content/uploads/2024/05/FOI2024-00576-NSOIT-Copy-compliance-policy_Redacted.pdf

290 Department for Levelling Up, Housing and Communities, "Parliamentary question: Elections: National Security", 5 February 2024, https://questions-statements.parliament.uk/written-questions/detail/2024-01-31/12399/

291 Kanishka Narayan MP, "DSIT activity to defend UK democracy from disinformation", Department for Science, Innovation and Technology, 16 January 2026, https://committees.parliament.uk/publications/51311/documents/284909/default/

292 Electoral Commission, 30 October 2019, "Commissioner Day notes: 30 October 2019", https://www.electoralcommission.org.uk/about-us/how-we-make-decisions/electoral-commission-board/commissioner-day-notes-30-october-2019

293 Ministry of Housing, Communities and Local Government, "Parliamentary question: Joint Election Security Preparedness Unit", 22 January 2025, https://questions-statements.parliament.uk/written-questions/detail/2025-01-08/HL3891

294 Suella Braverman MP, Letter to Home Affairs Committee, 27 March 2023, https://committees.parliament.uk/publications/39054/documents/192033/default/

295 HM Government, Integrated Review Refresh, March 2023, https://assets.publishing.service.gov.uk/media/641d72f45155a2000c6ad5d5/11857435_NS_IR_Refresh_2023_Supply_AllPages_Revision_7_WEB_PDF.pdf

296 Suella Braverman MP, Letter to Home Affairs Committee, 27 March 2023, https://committees.parliament.uk/publications/39054/documents/192033/default/

297 Government Analysis Function, "Data sharing for national crisis response", 14 August 2023, https://analysisfunction.civilservice.gov.uk/policy-store/data-sharing-for-national-crisis-response/

298 Ofcom, "Online Information Advisory Committee", 27 November 2024, https://www.ofcom.org.uk/about-ofcom/structure-and-leadership/advisory-committee-on-disinformation-and-misinformation

299 Baroness Jones, Letter to the Science, Innovation and Technology Committee, 9 May 2025, https://committees.parliament.uk/writtenevidence/141804/default/

300 Electoral Commission, "Response to the SIT Committee call for evidence on social media, misinformation and harmful algorithms", 28 January 2025, https://www.electoralcommission.org.uk/news-and-views/our-responses-consultations/response-science-innovation-and-technology-committee-call-evidence-social-media-misinformation-and

301 Electoral Commission, "Corporate plan 2025/6 - 2029/30", https://www.electoralcommission.org.uk/about-us/our-plans-priorities-and-spending/corporate-plan-2025/6-2029/30

302 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

303 Electoral Commission, "Response to the SIT Committee call for evidence on social media, misinformation and harmful algorithms", 28 January 2025, https://www.electoralcommission.org.uk/news-and-views/our-responses-consultations/response-science-innovation-and-technology-committee-call-evidence-social-media-misinformation-and

304 HMICFRS, "An inspection of the police response to the public disorder in July and August 2024: Tranche 2", 7 May 2025, https://hmicfrs.justiceinspectorates.gov.uk/publication-html/police-response-to-public-disorder-in-july-and-august-2024-tranche-2/

305 Ofcom, "Additional Safety Measures", 30 June 2025, https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/consultation-online-safety---additional-safety-measures/main-documents/consultation-additional-safety-measures-30-july-2025.pdf?v=403587

306 Full Fact, submission to Ofcom's Additional Safety Measures consultation, 17 October 2025, https://fullfact.org/documents/401/Full_Facts_Submission_to_Ofcoms_Additional_Safety_Measures_Consultation.pdf

307 HM Government, "Protecting What Matters: Towards a more confident, cohesive, and resilient United Kingdom", 28 April 2026, https://www.gov.uk/government/publications/protecting-what-matters-towards-a-more-confident-cohesive-and-resilient-united-kingdom/protecting-what-matters-towards-a-more-confident-cohesive-and-resilient-united-kingdom

308 Full Fact, "Framework for Information Incidents", https://fullfact.org/policy/incidentframework/

309 Foreign Affairs Committee, "Oral evidence: Disinformation diplomacy", 9 March 2026, https://committees.parliament.uk/event/26748/formal-meeting-oral-evidence-session/

310 Elizabeth Seger, Sam Stockwell, Tyreese Calnan, Henry Ajder, Jamie Hancock, Hannah Perry, "Epistemic Security for Crisis Resilience", Demos, 19 January 2026 https://demos.co.uk/wp-content/uploads/2026/01/Epistemic-Security-for-Crisis-Resilience_Report_2025_Jan_optimised.pdf

311 Ofcom, "Adults' Media Use and Attitudes Report", 7 May 2025, https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/adults/adults-media-use-and-attitudes-2025/adults-media-use-and-attitudes-report-2025.pdf?v=396240

312 Full Fact, "Briefing on the Representation of the People Bill", 20 February 2026, https://fullfact.org/documents/406/Full_Fact_policy_paper_on_the_ROTP_Bill_20.02.26.pdf

313 Electoral Commission, "Public Attitudes 2024", 7 May 2024, https://www.electoralcommission.org.uk/research-reports-and-data/public-attitudes/public-attitudes-2024

314 OECD, "Public Communication Scan of the United Kingdom", 16 December 2023, https://www.oecd-ilibrary.org/sites/bc4a57b3-en/1/3/3/index.html?itemId=/content/publication/bc4a57b3-en&_csp_=0aa641c3d4fda7ac26451f2c0133d8cf&itemIGO=oecd&itemContentType=book

315 Public Administration and Constitutional Affairs Committee, "Review of the 2024 general election", 22 July 2025, https://committees.parliament.uk/publications/48937/documents/256975/default/

316 Full Fact, "Briefing on the Representation of the People Bill", 20 February 2026, https://fullfact.org/documents/406/Full_Fact_policy_paper_on_the_ROTP_Bill_20.02.26.pdf

317 Ministry of Housing, Communities and Local Government, "Restoring trust in our democracy: Our strategy for modern and secure elections", 17 July 2025, https://www.gov.uk/government/publications/restoring-trust-in-our-democracy-our-strategy-for-modern-and-secure-elections/restoring-trust-in-our-democracy-our-strategy-for-modern-and-secure-elections

318 Julia Hörnle, "Deepfakes and the Law: Why Britain needs stronger protections against technology-facilitated abuse", Queen Mary, University of London, 24 January 2025, https://www.qmul.ac.uk/law/news/2025/items/deepfakes-and-the-law-why-britain-needs-stronger-protections-against-technology-facilitated-abuse.html

319 Ministry of Justice, "Better protection for victims thanks to new law on sexually explicit deepfakes", 22 January 2025, https://www.gov.uk/government/news/better-protection-for-victims-thanks-to-new-law-on-sexually-explicit-deepfakes

320 Electoral Commission, ""New advice for voters on disinformation, and for campaigners using generative AI", 17 June 2024, "https://www.electoralcommission.org.uk/media-centre/new-advice-voters-disinformation-and-campaigners-using-generative-ai

321 Demos Epistemic Security Network, "Epistemic Security briefing: The Elections Bill", January 2026, https://demos.co.uk/wp-content/uploads/2026/01/Epistemic-Security-Brieifing_The-Elections-bill_2026.pdf

322 Electoral Commission, "Written evidence submitted to the Speaker's Conference", 14 April 2025, https://committees.parliament.uk/writtenevidence/141330/html/

323 Speaker's Conference on the security of candidates, MPs and elections, "first report", 2 June 2025, https://committees.parliament.uk/publications/48116/documents/251907/default/

324 George Freeman MP, "Statement on fake AI-generated video circulating online", 19 October 2026, https://www.georgefreeman.co.uk/news/statement-fake-ai-generated-video-circulating-online; Paul Moseley, Helen Burchell, "Police investigate Tory MP deepfake defection video", BBC News, 22 October 2025, https://www.bbc.co.uk/news/articles/c4gpejxk0jpo

325 Vijay Rangarajan, "Additional Electoral Commission information for the Foreign Affairs Committee", Electoral Commission, 28 January 2026, https://committees.parliament.uk/writtenevidence/162487/html/

326 Speaker's Conference on the security of candidates, MPs and elections, "first report", 2 June 2025, https://committees.parliament.uk/publications/48116/documents/251907/default/

327 Speaker's Conference on the security of candidates, MPs and elections, "first report", 2 June 2025, https://committees.parliament.uk/publications/48116/documents/251907/default/

328 Public Bill Committee, "Representation of the People Bill, 18 March 2026, https://publications.parliament.uk/pa/bills/cbill/59-01/0384/PBC384_RepresentationBill_1st9th_Compilation_16_04_2026.pdf

329 Full Fact, "Briefing on the Representation of the People Bill", 20 February 2026, https://fullfact.org/documents/406/Full_Fact_policy_paper_on_the_ROTP_Bill_20.02.26.pdf

330 Foreign Affairs Committee, "Oral evidence: Disinformation Diplomacy", 13 January 2026, https://committees.parliament.uk/oralevidence/16988/pdf/; Vijay Rangarajan, "Additional Electoral Commission information for the Foreign Affairs Committee", Electoral Commission, 28 January 2026, https://committees.parliament.uk/writtenevidence/162487/html/

331 Noah Keate, "13 times Elon Musk meddled in politics", 8 January 2025, https://www.politico.eu/article/elon-musk-politics-uk-republicans-x-owner-donald-trump-prison/

332 Full Fact, "Briefing on the Representation of the People Bill", 20 February 2026, https://fullfact.org/documents/406/Full_Fact_policy_paper_on_the_ROTP_Bill_20.02.26.pdf

333 Full Fact, "The Representation of the People Bill does not protect UK democracy from misinformation", 25 February 2026, https://fullfact.org/politics/the-representation-of-the-people-bill-does-not-protect-uk-democracy-from-misinformation/

334 Hansard, Debate on the Integrated Review, 16 March 2021, https://hansard.parliament.uk/commons/2021-03-16/debates/52D67D49-A516-4598-AC69-68E8938731D9/IntegratedReview

335 HM Government, "Government Response to the Report of the Joint Committee on the Draft Online Safety Bill", March 2022, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061446/E02721600_Gov_Resp_to_Online_Safety_Bill_Accessible_v1.0.pdf

336 Full Fact, "What is the Online Safety Act? Here's what you need to know", https://fullfact.org/policy/online-safety-act/

337 Ofcom, "Implementing the Online Safety Act: progress update", 17 October 2024, https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/information-for-industry/roadmap/2024/ofcoms-approach-to-implementing-the-online-safety-act-2024.pdf?v=383285; Ofcom, "Year 1 Online Safety Risk Assessments", 4 December 2025, https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/research-statistics-and-data/os-standards/online-safety-risk-assessments-report-year-one.pdf?v=40862

338 Department for Science, Innovation and Technology, "Online Safety Act: explainer", 24 April 2025, https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer

339 Ofcom, letter to online services operating in the UK, 1 April 2026, https://www.ofcom.org.uk/siteassets/resources/documents/about-ofcom/public-correspondence/2026/open-letter-on-elections.pdf?v=415530

340 Kanishka Narayan MP, letter to Sir Lindsay Hoyle MP, Department for Science Innovation and Technology, 15 September 2025, https://committees.parliament.uk/publications/49555/documents/264007/default/

341 Online Safety Act Network, "User-to-User Illegal Content Duties", 1 September 2025, https://www.onlinesafetyact.net/analysis/user-to-user-illegal-content-duties/

342 Science, Innovation and Technology Committee, "UK's Online Safety regime unable to tackle the spread of misinformation and cannot keep users safe online, MPs warn", 11 July 2025, https://committees.parliament.uk/committee/135/science-innovation-and-technology-committee/news/208296/uks-online-safety-regime-unable-to-tackle-the-spread-of-misinformation-and-cannot-keep-users-safe-online-mps-warn/

343 HM Government, Response to report on social media, misinformation and harmful algorithms, 17 October 2025, https://committees.parliament.uk/publications/49793/documents/266872/default/

344 HMICFRS, "An inspection of the police response to the public disorder in July and August 2024: Tranche 2", 7 May 2025, https://hmicfrs.justiceinspectorates.gov.uk/publication-html/police-response-to-public-disorder-in-july-and-august-2024-tranche-2/

345 Ofcom, "Online Safety Transparency Reporting Final Transparency Guidance", 21 July 2025, https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/consultation-draft-transparency-reporting-guidance/main-docs/final-transparency-guidance.pdf?v=400422

346 Hansard, "MPs will debate a petition relating to the Online Safety Act", 15 December 2025

347 Full Fact, "What is the Online Safety Act? Here's what you need to know", https://fullfact.org/policy/online-safety-act/

348 Department for Science, Innovation and Technology, "Growing up in the online world: a national consultation", 2 March 2025, https://www.gov.uk/government/consultations/growing-up-in-the-online-world-a-national-consultation

349 Richard Wheeler, André Rhoden-Paul, "Social media restrictions for under-16s even if no ban, minister says", 28 April 2026, https://www.bbc.co.uk/news/articles/c5y7d2zx63jo

350 Science, Innovation and Technology Committee, "Social media, misinformation and harmful algorithms", 11 July 2025, https://committees.parliament.uk/publications/48745/documents/258221/default/

351 Electoral Commission, "Response to the SIT Committee call for evidence on social media, misinformation and harmful algorithms", 28 January 2025, https://www.electoralcommission.org.uk/news-and-views/our-responses-consultations/response-science-innovation-and-technology-committee-call-evidence-social-media-misinformation-and

352 Electoral Commission, "Response to the Science, Innovation and Technology Committee call for evidence on social media, misinformation and harmful algorithms", 28 January 2025, https://www.electoralcommission.org.uk/news-and-views/our-responses-consultations/response-science-innovation-and-technology-committee-call-evidence-social-media-misinformation-and

353 Department for Science, Innovation and Technology, "Data (Use and Access) Act 2025: data protection and privacy changes", 27 June 2025, https://www.gov.uk/guidance/data-use-and-access-act-2025-data-protection-and-privacy-changes

354 Ofcom, "Researchers' access to information from regulated online services", 8 January 2025, https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/call-for-evidence-researchers-access-to-information-from-regulated-online-services/main-documents/researchers-access-to-information-from-regulated-online-services.pdf

355 Paolo Cesarini, Lisa Ginsborg, Paula Gori, Elena Maggi, Rolf Nijmeijer, "Platforms' Implementation of the CoP Commitments on Media Literacy, Research and Fact-checking", EDMO, May 2024, https://edmo.eu/wp-content/uploads/2024/06/EDMO-CoP-Report.pdf; Lisa Ginsborg, "Report on EDMO Workshop on Platform Data Access for Researchers", EDMO, September 2024, https://edmo.eu/wp-content/uploads/2024/09/Report-on-EDMO-Workshop-on-Platform-Data-Access-for-Researchers.pdf

356 Jonathan Hall KC, "The Foreign Hand and Foreign Interference", 23 July 2024, https://terrorismlegislationreviewer.independent.gov.uk/wp-content/uploads/2024/07/Foreign-Interference-RUSI-IRSTL-23.7.24-1.pdf

357 Joint Committee on the National Security Strategy, "Letter to the Prime Minister, Rt Hon Rishi Sunak MP", 23 May 2024, https://committees.parliament.uk/publications/45032/documents/223340/default/

358 Philip Rycroft, "Report of the Independent Review into Countering Foreign Financial Influence and Interference in UK Politics", 25 March 2026, https://assets.publishing.service.gov.uk/media/69c29f84b920af63be1c7777/The_Rycroft_Review_Report_standard_version.pdf

359 Foreign Affairs Committee, "Oral evidence: Disinformation diplomacy", 9 March 2026, https://committees.parliament.uk/event/26748/formal-meeting-oral-evidence-session/

360 Hansard, debate on Artificial Intelligence Legislation, 17 November 2025, https://hansard.parliament.uk/Lords/2025-11-17/debates/BD4C0FAB-9CFF-445F-9A9D-83FFEACEFD70/ArtificialIntelligenceLegislation; Harrison Ostridge, "Potential future risks from autonomous AI systems", House of Lords Library, 5 January 2026, https://lordslibrary.parliament.uk/potential-future-risks-from-autonomous-ai-systems

361 Vasilios Mavroudis, Chris Hicks, "LLMs may be more vulnerable to data poisoning than we thought", The Alan Turing Institute, 9 October 2025, https://www.turing.ac.uk/blog/llms-may-be-more-vulnerable-data-poisoning-we-thought

362 Labour Party, "Change", June 2026, https://labour.org.uk/wp-content/uploads/2024/06/Labour-Party-manifesto-2024.pdf

363 Kanishka Narayan MP, "Minister Narayan AI speech at Founders Forum", Department for Science Innovation and Technology, 12 February 2026, https://www.gov.uk/government/speeches/minister-narayan-ai-speech-at-founders-forum

364 Ofcom, "AI chatbots and online regulation – what you need to know", 18 December 2025, https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/ai-chatbots-and-online-regulation-what-you-need-to-know

365 Ofcom, "Ofcom update: Investigation into X, and scope of the Online Safety Act", 3 February 2026, https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/investigation-into-x-and-scope-of-the-online-safety-act; Maeve Walsh, "Response to Ofcom update on Grok", Online Safety Act Network, 3 February 2026, https://www.onlinesafetyact.net/analysis/osa-network-response-to-ofcom-update-on-grok/

366 Home Office, "Collection: Crime and Policing Act 2026", 25 February 2025, https://www.gov.uk/government/collections/crime-and-policing-act-2026

367 Tallulah Belassie-Page, "AI chatbots: a missed opportunity", 20 April 2026, https://www.onlinesafetyact.net/analysis/ai-chatbots-a-missed-opportunity/; Bird & Bird, "UK Government Children's Safety and AI Chatbot Powers", 30 April 2026, https://www.twobirds.com/en/insights/2026/uk/uk-government-children's-safety-and-ai-chatbot-powers-two-new-acts-receive-royal-assent

368 Dan Milmo, "'Deepfakes spreading and more AI companions': seven takeaways from the latest artificial intelligence safety report", Guardian, 3 February 2026, https://www.theguardian.com/technology/2026/feb/03/deepfakes-ai-companions-artificial-intelligence-safety-report

369 Will Meakin-Durrant, "Budget-style online safety debates could help law 'keep pace' with tech changes", The Standard, 16 February 2026, https://www.standard.co.uk/news/politics/liz-kendall-government-keir-starmer-britain-b1271186.html

370 AI Safety Institute, "International scientific report on the safety of advanced AI: interim report", updated 22 October 2025, https://www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai/international-scientific-report-on-the-safety-of-advanced-ai-interim-report

371 AI Security Institute, "Frontier AI Trends Report", https://www.aisi.gov.uk/frontier-ai-trends-report; AI Security Institute, ""Research Agenda", https://www.aisi.gov.uk/research-agenda,

372 Nuala Polo , Roshni Modhvadia, "Great (public) expectations" Institute Ada Lovelace Institute, 4 December 2025, https://www.adalovelaceinstitute.org/policy-briefing/great-expectations/

373 Hansard, debate on Artificial Intelligence Legislation, 17 November 2025, https://hansard.parliament.uk/Lords/2025-11-17/debates/BD4C0FAB-9CFF-445F-9A9D-83FFEACEFD70/ArtificialIntelligenceLegislation

374 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

375 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

376 Full Fact, submission to Ofcom's Additional Safety Measures consultation, 17 October 2025, https://fullfact.org/documents/401/Full_Facts_Submission_to_Ofcoms_Additional_Safety_Measures_Consultation.pdf

377 Council of Europe, "Guidance note on the prioritisation of public interest content online", 2 December 2021, https://rm.coe.int/cdmsi-2021-009-guidance-note-on-the-prioritisation-of-pi-content-e-ado/1680a524c4

378 HM Government, "Protecting What Matters: Towards a more confident, cohesive, and resilient United Kingdom", 28 April 2026, https://www.gov.uk/government/publications/protecting-what-matters-towards-a-more-confident-cohesive-and-resilient-united-kingdom/protecting-what-matters-towards-a-more-confident-cohesive-and-resilient-united-kingdom

379 YouGov poll on 29-30 March 2026 based on a nationally representative sample of 2,175 adults in the UK, commissioned by Full Fact

380 Shout Out UK, "Government reconfirms commitment to Votes at 16: A joint statement from the APPG on Political & Media Literacy and Shout Out UK", 17 July 2025, https://www.shoutoutuk.org/2025/07/17/votes-at-16/

381 Sally Burtonshaw, Pete Whitehead, Amy Braier, Denise Baron, Ed Dorrell, Seb Wride, Jules Walkden, Will Yates, "Launch report", Commission into Countering Online Conspiracies in Schools, February 2025, https://counteringconspiracies.publicfirst.co.uk/report-1/

382 Sally Burtonshaw, Michael Kane, Jules Walkden, Ed Dorrell, Amy Braier, George Ryan, Seb Wride, "2026 research insights", Commission into Countering Online Conspiracies in Schools, March 2026, https://counteringconspiracies.publicfirst.co.uk/Commission_into_Countering_Online_Conspiracies_in_Schools_Report_2.pdf

383 Full Fact, Internet Matters, "Preparing young people to vote in a complex, attention-driven information environment", 10 February 2026, https://fullfact.org/documents/405/Briefing_-_votes_at_16_and_media_literacy.pdf

384 Science, Innovation and Technology Committee, "Social media, misinformation and harmful algorithms", 11 July 2025, https://committees.parliament.uk/publications/48745/documents/258221/default/

385 Foreign Affairs Committee, "Disinformation diplomacy: How malign actors are seeking to undermine democracy", 27 March 2026, https://committees.parliament.uk/publications/52401/documents/290829/default/

386 Philip Rycroft, "Report of the Independent Review into Countering Foreign Financial Influence and Interference in UK Politics", 25 March 2026, https://assets.publishing.service.gov.uk/media/69c29f84b920af63be1c7777/The_Rycroft_Review_Report_standard_version.pdf

387 William Dixon, "Why the UK Now Needs a National Disinformation Agency" RUSI, 5 September 2025, https://www.rusi.org/explore-our-research/publications/commentary/why-uk-now-needs-national-disinformation-agency

388 Full Fact, "The Representation of the People Bill does not protect UK democracy from misinformation", 28 February 2027, https://fullfact.org/politics/the-representation-of-the-people-bill-does-not-protect-uk-democracy-from-misinformation/

389 Full Fact, "Framework for Information Incidents", https://fullfact.org/policy/incidentframework/

390 Craig T. Robertson, "Most people want platforms (not governments) to be responsible for moderating content", Reuters Institute, 30 September 2025, https://reutersinstitute.politics.ox.ac.uk/news/most-people-want-platforms-not-governments-be-responsible-moderating-content

391 Full Fact, submission to Ofcom's Additional Safety Measures consultation, 17 October 2025, https://fullfact.org/documents/401/Full_Facts_Submission_to_Ofcoms_Additional_Safety_Measures_Consultation.pdf

**Full Fact**
17 Oval Way
London
SE11 5RR

fullfact.org/contact
@FullFact
fullfact.org

Published by Full Fact, June 2026. 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