BRIEFING

European Parliament

Online information manipulation and information integrity

An overview of key challenges, actors and the EU's evolving response

SUMMARY

The global information ecosystem is facing challenges on various levels, necessitating a clear overview of the key issues at stake, the actors involved and possible European Union responses. On a geostrategic level, authoritarian state- and non-state actors, who work to game the open democratic information ecosystems in their favour, have gained the most visibility. At the same time, the tech companies that underpin those open information ecosystems – and that help provide the infrastructure used for information manipulation – are under pressure to take more responsibility.

Evolving technologies can exacerbate the risks of information ecosystems working against democracy, rather than for it. Moreover, underlying societal, educational and economic vulnerabilities hamper both individual and collective resilience against information manipulation. The impact on health, societies, economies, democracy, international decision-making, security and human rights has become increasingly visible in recent years. Correspondingly, there is a growing sense of urgency to ensure information integrity, both in the context of elections and beyond.

The EU has continued to strengthen its efforts to counter information manipulation and interference, including online disinformation, since 2015. The evolving measures have matured into a growing regulatory framework to address digital information infrastructure vulnerabilities and boost information ecosystems. This, in addition to strengthening societal resilience, involves a high level of cooperation and coordination within and beyond the EU, across all policy areas and with all levels of society. International and multilateral cooperation is therefore key to future-proofing the response.

IN THIS BRIEFING

EPRS | European Parliamentary Research Service

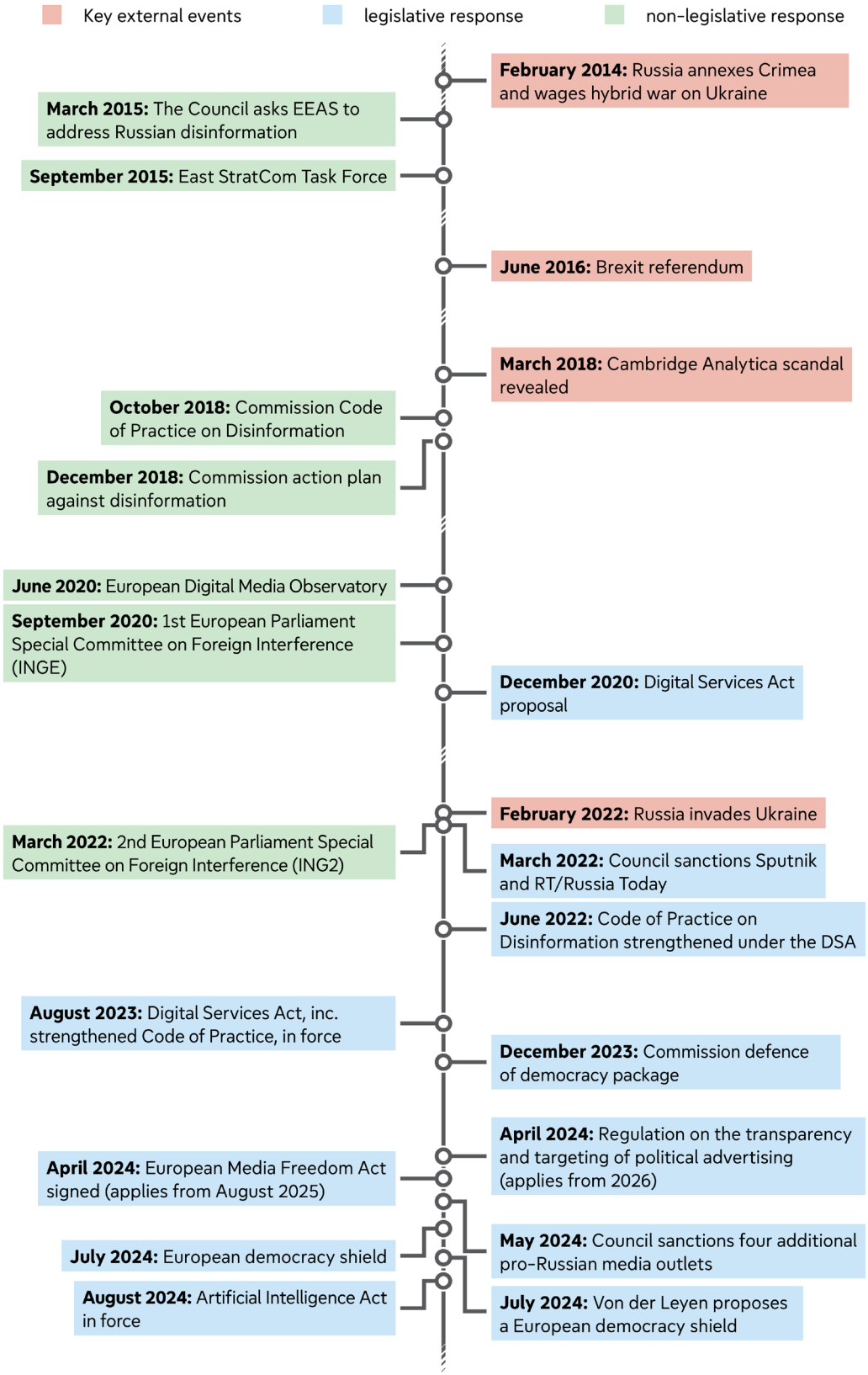
Author: Naja Bentzen; Graphics: Samy Chahri

Members' Research Service

PE 762.416 – September 2024

EN

Timeline: Information manipulation and the EU's response



Source: S. Chahri, EPRS, 2024, based on: Official Journal of the EU; European Commission; The Council; *The Guardian*.

Introduction

Information manipulation, notably in the digital realm, plays an increasingly visible role in public debate. A 2023 survey notes 85 % of people worldwide worry about the impact of disinformation on fellow citizens, and 87 % think disinformation has already affected political life in their country. At the same time, 38 % of EU respondents to a 2023 Eurobarometer survey listed false and/or misleading information as a threat to democracy. Responding to a 2024 Eurobarometer survey, 45 % specified fake news and disinformation as one of the issues with the biggest personal impact on them. The World Economic Forum's Global Risks Report 2024 called misinformation and disinformation 'the most severe short-term risk the world faces', warning that the 'nexus between falsified information and societal unrest will take centre stage amid elections in several major economies', and that AI's role in amplifying manipulated information could destabilise societies.

Information manipulation is a key dimension of a broader set of pressures on the information ecosystem.[1] Geostrategic (mainly foreign, authoritarian) actors manipulate open democratic information ecosystems, while tightening control of their domestic information monocultures and 'memory politics'. Enabling actors include large tech companies whose attention economy thrives on engagement and collecting user data. Moreover, societal and psychological dynamics can leave people vulnerable to deceptive narratives.

Information manipulation goes beyond specific campaigns to impact public opinion, including in the context of elections. It involves using 'digital information technology by authoritarian regimes to surveil, repress, and manipulate domestic and foreign populations',[2] sometimes called algorithmic authoritarianism. Similar to the shifting focus from content – and thus intent – to behaviour, as explained below, some experts suggest replacing intention-based definitions of digital authoritarianism with a promotion-based definition, taking into account 'any situation where digital technologies systematically promote authoritarian politics'.[3]

Although downplaying the threats from information manipulation can benefit bad actors, experts warn that overemphasising the power of disinformation risks not only amplifying the original falsehoods, but also conveying the corrosive narrative that democracy is not working. This can benefit actors who aim to undermine democracies.

Changing terminology reflects the evolving policy response

Shifting focus from content to behaviour

The changing terminology used by the EU and the European Parliament to capture the challenges of misleading information has evolved in the past decade, reflecting a shifting focus from content to behaviour, in full respect of freedom of expression. In 2018, the European Commission defined **disinformation** as 'verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public and may cause public harm'. **Misinformation**, on the other hand, refers to unintentionally false information. However, identifying the intent to deceive can be challenging, and not all conspiracy claims are false or deceptive. In the 1970s, the Watergate scandal involving US President Richard Nixon, was initially dismissed as a politically motivated conspiracy theory.

The European External Action Service (EEAS) coined the term **foreign information manipulation and interference** (FIMI) in 2021. It defines it as 'behaviour 'that threatens or has the potential to negatively impact values, procedures and political processes', and that is 'manipulative in character' and 'conducted in an intentional and coordinated manner', involving 'state or non-state actors, including their proxies inside and outside of their own territory'. This definition reflects the EEAS' external mandate, with an explicit focus on authoritarian state actors such as Russia and China. Since then, EU institutions and agencies have adopted the FIMI concept, including in the 2022 Strategic Compass on Security and Defence and the EU Agency for Cybersecurity (ENISA, 2023). FIMI also found its way into the transatlantic official vocabulary via the May 2023 EU-US Trade and

Technology Council (TTC) joint statement. In recent years, the notion of **information integrity**, as detailed below, is increasingly used as a sustainable approach to addressing information manipulation while protecting freedom of expression.

Actors and enabling factors

Key foreign authoritarian state and non-state actors

The most visible anti-democracy information manipulators are authoritarian state and non-state actors. The EEAS focuses explicitly on addressing FIMI from Russia and China. Russia and China's shared geostrategic interests in weakening Western-style democracies are increasingly manifested in converging information campaigns'.[4] In some cases, they have laundered each other's disinformation campaigns, including blaming the West for the war on Ukraine.

Iran is widely recognised as an emerging global actor in the field of information manipulation, a trend that has become particularly visible in the context of the Israel-Hamas war. In May 2024, the FBI explicitly pointed to Russia, Iran and China as the countries of most concern to the US ahead of the 2024 election. All three countries have been working simultaneously to amplify the tension around campus protests over the Biden administration's Israel policies. The Media Forensics Hub at Clemson University has identified overt and covert campaigns by all three countries, who spread the content via state media and inauthentic accounts or bots on social media platforms including X and Telegram. Moreover, Russia has used websites created to mimic Western news organisations in its generative AI-facilitated Doppelganger network, which it has deployed in information operations across the world. Venezuelan authorities working to influence Latin Americans are using anti-Western narratives that converge with Russia, China and Iran's campaigns.

Russia and China are investing heavily in information influence networks in Latin America and Africa. According to the Africa Center for Strategic Studies, campaigns to manipulate the information ecosystems in Africa increased nearly fourfold since 2022, fuelling destabilisation, deadly violence and democratic backsliding. Foreign states sponsor 60 % of these campaigns, led by Russia, followed by China, the United Arab Emirates, Saudi Arabia and Qatar. China in particular is exporting technologies facilitating information control, such as surveillance and repression, to authoritarian and authoritarian-leaning governments across the world, including in Latin America, Africa and the Western Balkans.

Business actors and factors

Online platforms and search engines

The effects of the role of online platforms and search engines in global information ecosystems, including the growing use of digital media, is a topic of sometimes heated debate. Social media has the potential to increase democratic participation, but concern over the impact of the attention economy on democracy has increased and broadened since the 2018 revelations that Cambridge Analytica had used Facebook data breaches to microtarget voters in the 2016 US elections. This drew global attention to the systemic challenges to the information sphere posed by digital platforms, as also reflected in the EU's response. At the same time, ranking algorithms, filtering and recommendation systems that monetise viral content can be misused to increase the spread of false, divisive or harmful information. This can impact individual and collective decision-making, as well as public opinion (see below). At the same time, these companies absorb advertising revenue that used to fund traditional media who uphold journalistic standards and ethics, but often pay little or no taxes despite massive earnings, not least by targeting young people.

The introduction of new technologies that facilitate the use of information manipulation – most recently and notably, generative AI – further exacerbates the risks to democracy stemming from the information ecosystems. Similar to the importance of the introduction of social media on the democratisation and disruption of the global information ecosystem, generative AI such as ChatGPT

and similar tools is making advanced technology available to industries, companies, individuals – and to anti-democratic forces, making manipulation cheaper and more efficient. At the same time, this rollout is expected to further undermine the economic basis for time-consuming and expensive production of reliable news and knowledge by research institutes, universities, news outlets, including already struggling local news outlets across the world. While the volume of information skyrockets, quality is expected to plummet further.

Human and societal factors: Appetite for toxicity?

The cognitive factors that drive the demand for manipulative content, as well as the interaction with the designs that peddle mis- and disinformation, are still being explored. Scientific research points to confirmation bias, or motivated reasoning; an unconscious preference for information that confirms already existing beliefs. A 2023 study by the Massachusetts Institute of Technology found the mere thought of sharing news on social media reduced the ability to judge whether a story was true or false; the urge to share outweighed consideration of accuracy. Other research has found that voters – regardless of culture, gender, information access and language – are more likely to perceive a lying candidate as 'authentically appealing' if they believe the political system is flawed.

A 2021 study found that sharing falsehoods is linked to political partisanship and the availability of news that can confirm political biases. In the politically polarised US, both Democrats and Republicans actively seek information that can denigrate their political opponent. In 2024, the Economist Intelligence Unit found that polarisation levels affect the impact of AI-generated content: manipulated content spreads more easily in societies with 'two antagonistic camps that disagree on major issues' and a similar share of the vote.

In the attention economy, emotional content is used to spark attention and maintain engagement. Accelerated by the pandemic, exacerbated by social media, and creating severe risk of harm to young people, mental health issues could feed a vicious circle of addictive behaviour and vulnerability to deceptive content, with algorithmic selection pushing some towards extremes. Extremist groups exploit 'identity fusion' (feeling 'one' with a group) to groom and radicalise video gamers. Highly emotional situations, for example (natural or manmade) disaster, terrorism, war, food insecurity, energy and migration crisis, as well as persistent economic and social hardship, can provide fertile ground for deceptive narratives.

The impact of information manipulation: Why it matters

Manipulative information's cognitive effects on individuals can be hard to quantify, and the lack of non-Western data adds to the research gap. However, some societal and political impacts are becoming clearer. Selected examples include:

Political violence. Despite the US Elections Infrastructure Government Coordinating Council and the Election Infrastructure Sector Coordinating Executive Committees' consensus that the November 2020 US election was the most secure in American history, then-President Donald Trump continues to insist that the election was 'stolen' from him. The claims and related conspiracy theories (which continue to proliferate) resulted in the deadly attack on the US Capitol in 2021. Since then, related conspiracy theories have spread in other continents, inspiring a similar attack on Brazil's Congress in January 2023.

Hampering participation in electoral processes. Deceptive or misleading claims about election interference can lead to political violence against both politicians and election officials. This can discourage active participation at all levels of electoral processes. A global survey of electoral officials from 73 countries showed that most respondents had been targeted by disinformation and/or online aggression while serving in their official roles. A 2024 US survey showed threats against, abuse or harassment of election officials had led to increased efforts to protect voters, election workers and infrastructure from violence ahead of the 2024 election.

Societal unrest. X's algorithms reportedly favour X CEO Elon Musk's posts, thus amplifying his opinions on world events. Musk, now the most followed person on X, self-identifies as a 'free speech absolutist', but regularly engages in public clashes with world leaders with whom he disagrees, often reflecting his pledge to 'destroy the woke mind virus'. Musk used his outsized global reach to spread disinformation during riots in the UK in 2024 (sparked by the murder of three girls and subsequent false online rumours about the perpetrator). For example, he reposted (but later deleted) that 'civil war is inevitable' in the UK.

The **'liar's dividend'.** In addition to facilitating the creation and reach of deceptive content, generative AI can further undermine already low trust in information and news on online platforms.

Gendered disinformation: A global challenge to democracy

In 2023, UN Special Rapporteur on Freedom of Expression, Irene Khan, warned that gendered disinformation (targeting women, girls and gender non-conforming persons) is on the rise across the world. A 2023 report by the International Research & Exchanges Board (IREX) pointed to gendered disinformation as a distinct threat to democracies, democratic values and democratic participation. IREX assessed that, by targeting groups where identities related to gender, race, ethnicity, sexual orientation and religion intersect, in particular for those with high public visibility, such as politicians, journalists, and human rights defenders, authoritarian actors can exacerbate polarising topics, undermine democratic principles, institutions, and human rights.

In this context, generative AI also further exacerbates already established trends: women are disproportionately targeted by non-consensual intimate deepfakes (NCID); abusive deepfake porn that can easily go viral. The July 2024 announcement that Vice President Kamala Harris would be the Democratic presidential nominee increased the visibility of already known gendered narratives in pro-Kremlin media when commenting on female public figures. In Harris's case, comments often include racially loaded language. Within the US, gendered narratives about Kamala Harris have been revived and upscaled, often using harsh, sexualised language. In the EU, pro-Kremlin online attacks on the EU's gender equality policies (often depicted as going against the Kremlin's top-down 'traditional values') and the Union's increasingly visible female leaders, will likely continue or grow, using gendered tropes to undermine their work, and by extension the EU's image.

This can make it easier to dismiss facts and evade accountability. One example is the claim by lawyers defending Elon Musk, as CEO of Tesla, that a 2016 video of him stating his cars could 'drive autonomously with greater safety than a person' was a deepfake. The court rejected the apparent effort to use the 'liar's dividend', a dynamic where those who lie to avoid accountability become more believable, precisely due to growing awareness of threats, in this case from deepfakes. A similar dynamic occurred when one of Musk's own posts was flagged as misleading by X's Community Notes system in 2023, prompting his claim that his own system had been 'gamed by state actors'.

Discrediting and silencing dissent. In recent years, China appears to have learned tactics to discredit and silence dissidents from the Kremlin. One example is the online campaign to silence and discredit Philadelphia-based dissident writer Deng Yuwen and, by proxy, his 16-year old daughter. Here, China used covert propaganda networks accounts (known as Spamouflage or Dragonbridge) linked to its security services to spread sexually loaded and intimidating posts about the daughter on social media. In engagement with Deng's X account as well as with the accounts of public schools in their community, users with fake identities falsely portrayed the daughter as a drug user, an arsonist and a prostitute. Exiled (female) Chinese journalists have also faced systematic online attacks, including being inserted in fake escort ads, and threatened with rape and bomb attacks.

Impact on regional and international security. A 2024 report by the UN University Centre for Policy Research noted that AI-enabled developments have already impacted long-term peacebuilding activities in Sub-Saharan Africa. It warned of 'significant risk of a continuing increase in polarization, eroding attempts to reconcile groups and work towards peace'.

The EU's evolving response to information manipulation

The EU's response to these challenges cuts across policy areas. Measures to curb information manipulation and boost information integrity aim to complement each other, with some of the early non-legislative steps incorporated in later legislation. Coordination with Member States, external partners, industry, civil society and a range of other stakeholders is at the core of the EU's approach.

External dimension

The European External Action Service spearheads the EU's endeavour to curb foreign disinformation since the East StratCom Task Force was created in 2015, in response to the Council's request under Latvia's Presidency to counter Russian disinformation. Since then, the EEAS continues to expand its work to counter foreign information manipulation and interference. This includes working with partners on democracy promotion as well as civilian and military missions and operations. In addition, a Rapid Alert System (RAS) on disinformation facilitates cooperation with other EU institutions and the Member States. The EEAS works with the Commission and the Member States to expand and fine-tune the EU Toolbox to tackle FIMI. The EU has also stepped up cooperation with NATO to counter disinformation.

In March 2022, the Council suspended Russian state-sponsored broadcasters RT and Sputnik in the EU. In June 2022, it also suspended Russian state-owned outlets Rossiya RTR/RTR Planeta, Rossiya 24/Russia 24 and TV Centre International. In May 2024, the EU added four more Kremlin-backed media, arguing that Russia 'has engaged in a systematic, international campaign of media and information manipulation, interference and grave distortion of facts in order to justify [...] its full-scale aggression against Ukraine, and to enhance its strategy of destabilisation of its neighbouring countries, and of the EU and its Member States'.

Digital dimension

Digital Services Act

The Digital Services Act (DSA) is part of the Commission's 2020 digital services package, alongside the Digital Markets Act. Together with the General Data Protection Regulation (GDPR), they aim to increase commercial actors' accountability. The DSA includes obligations for online platforms to make the online space transparent and safe for users, protect fundamental rights and fight disinformation. Here, the DSA is complemented by the EU Code of Practice on Disinformation, supported by the European Digital Media Observatory (EDMO). The regulation on transparency and targeting of political advertising (TTPA) and the European Media Freedom Act (EMFA) further strengthen the regulatory framework.

Under the DSA, all intermediaries are obliged to combat illegal content such as terrorist content and hate speech. However, very large online platforms (VLOPs) and search engines (VLOSEs) with over 45 million users carry greater responsibility for curbing socially harmful content in the EU, including disinformation. The DSA includes obligations under Articles 34 and 35 on VLOPs and VLOSEs to assess their mitigation measures and results against systemic risks (such as disinformation) and implement crisis protocols in exceptional circumstances.

In August 2023, the DSA entered into force for designated VLOPs and VLOSEs. The Commission has monitored progress, including during national elections and the European elections in June 2024; a milestone for the implementation of the DSA. Ahead of June 2024, the Commission issued guidelines for the providers of VLOPs and VLOSEs with recommended risk mitigation for electoral integrity. Specific mitigation measures linked to generative AI included clear labelling of content such as deepfakes (AI-generated or manipulated text, audio or video).

To ensure proper and consistent enforcement of the DSA across the EU, the DSA requires Member States to designate Digital Services Coordinators (DSCs) to supervise online platforms with headquarters in their countries, and to support the Commission's investigations. The DSCs are also

responsible for identifying 'trusted flaggers', for example civil society organisations that work independently of online platforms with specific expertise and competence in detecting, identifying and notifying illegal content. Member States had to designate DSCs by February 2024. On 25 July 2024, the Commission opened infringement proceedings against Belgium, Spain, Croatia, Luxembourg, the Netherlands and Sweden – either for not designating DSCs to implement the DSA, or not empowering them to perform the tasks required by the Act – or both. Member States were given two months to respond and address the shortcomings.

The Commission also organised stress tests to prepare VLOPs and VLOSEs, civil society organisations, DSCs and EU institutions for information manipulation threats, and to set up a voluntary DSA incident and threat response framework with the DSCs. In addition to regulatory dialogues with the VLOPs and VLOSEs, an ad hoc working group was created under the European Board for Digital Services to exchange DSCs and national authorities' expertise on electoral issues. This will be integrated into a permanent working group.

In the context of the elections, the Commission used the DSA's enforcement rules not only to request information from VLOPs and VLOSEs concerning elections, but also to open formal proceedings against platforms that failed to meet the requirements. These latter include:

- In December 2023, the Commission launched formal proceedings against X (which withdrew from the Code of Practice in May 2023) over breaches of the DSA under Article 66, including inadequate measures to curb information manipulation via X's 'Community Notes' system, after reducing content moderation staff. In May 2024, the Commission widened its investigations and in July 2024 informed X its preliminary findings identified breaches of the DSA regarding dark patterns, advertising transparency and data access for researchers. Investigations are ongoing.
- In February 2024, the Commission opened formal proceedings against TikTok over suspected DSA breaches related to the protection of minors, advertising transparency, data access for researchers, and risk management of addictive design and harmful content.
- In April 2024, the Commission opened formal proceedings against Meta over the lack of an effective third-party, real-time civic discourse and election-monitoring tool ahead of the EU elections, as well as data access for researchers.

Strengthened Code of Practice

A cornerstone of the EU strategy against disinformation, the 2018 Code of Practice pioneered voluntary industry player engagement in countering disinformation. The 2022 strengthened Code of Practice under the DSA strengthens these obligations. This includes ensuring that disinformation spreaders do not benefit from advertising revenue. It commits signatories to greater transparency for users; enabling them to recognise political ads via clearer labelling, revealing the sponsor, ad spend and display period. Better tools to recognise, understand and report disinformation, access authoritative sources and media literacy initiatives aim to further empower users. Moreover, the strengthened code improves cooperation to counter information manipulation via periodical reviews of tactics, techniques and procedures (TTPs) employed by malicious actors.

Signatories have to reduce manipulative behaviour used to spread disinformation (fake accounts, bot-driven amplification, impersonation, malicious deep fakes). The Code improves support for research on disinformation, including via better access to platforms' data, and extends fact-checking to all EU Member States and languages. A dedicated Transparency Centre gives insight into the Code's implementation, and a permanent task force (chaired by the Commission and composed of representatives of signatories, the European Regulators' Group for Audiovisual Media Services, the EDMO and the EEAS) ensures the Code is future-proof and fit for purpose. The guidelines also recommend that providers anticipate applying certain provisions of the TTPA (see below).

European Digital Media Observatory (EDMO)

The Commission's 2018 action plan to counter online disinformation led to the 2020 creation of the European Digital Media Observatory; a hub for fact-checkers, academics and other stakeholders to collaborate and coordinate activities to curb disinformation, such as:

- mapping fact-checking organisations in Europe and supporting them via joint and cross-border activities and training modules;
- mapping, supporting and coordinating research on disinformation at European level; a global repository of peer-reviewed scientific articles on disinformation;
- building a public portal providing media practitioners, teachers and citizens with information and materials to increase awareness, build resilience to online disinformation and support media literacy campaigns;
- designing a framework to ensure secure and privacy-protected access to platforms' data for academic researchers working to understand disinformation;
- support for public authorities to monitor online platforms' policies to limit the spread and impact of disinformation.

Building on the core infrastructure, national and regional EDMO hubs are being rolled out across Europe. With 14 hubs launched so far, capacity to tackle harmful disinformation at national and EU level, and analyse its impact on society and democracy has increased.

Reflecting the Conference on the Future of Europe, the Commission funded the 2022 launch of the European Fact-Checking Standards Network (EFCSN) of independent fact-checking organisations committed to the European Code of Standards for Independent Fact-Checking Organisations, written by organisations from over 30 European countries.

Next steps

The implementation of the DSA, the Code of Practice and EDMO include:

- Working Group on the Integrity of the Information Space;
- increased monitoring and detection capabilities by boosting EDMO;
- finalising the Code's Rapid Response System for future elections;
- converting the Code of Practice into a DSA Code of Conduct.;

Regulation on transparency and targeting of political advertising (TTPA)

In the 2020 European democracy action plan, the Commission announced an additional legislative proposal on sponsored political advertising. Signed in March 2024, the regulation covers political advertising provided for remuneration and through in-house activities. It aims to increase trust in election campaigns by countering information manipulation and interference in the political debate, including via clearer information for users about who is behind political ads, and how they are targeted. The TTPA limits targeting and delivery techniques; increases protection of personal data in online political advertising; bans profiling (use of special data categories such as users' racial or ethnic origin and political opinion); bans non-EU based entities from financing political ads in the EU three months before an election or referendum. It also includes a public repository for all online political advertisements. The TTPA entered into force on 9 April 2024 and will be applied from 9 October 2025, except its definitions (Article 3) and the non-discrimination clause for cross-border political advertising (Article 5(1)), which apply since 9 April 2024.

Artificial Intelligence (AI) Act and AI Office

The AI Act is the first comprehensive law worldwide to regulate AI. Amongst other things, it tackles deepfakes, –based on the level of risk they pose. While the AI Act classifies deepfakes as 'limited risk', Annex III 8(b) considers AI systems used to influence an election, referendum or the voting behaviour of natural persons as high-risk. To protect voting rights (Article 39 of the Charter), the AI

Act considers AI systems used to influence elections or manipulate behaviour as high-risk (Recital 62 and Annex III 8(b)), except AI for campaign logistics with limited user interaction. Provisions include:

- Transparency obligations (Article 50): Under Article 52(3), creators, developers and users of deepfake technologies must disclose any AI-generated content. This aims to curb mis- and disinformation and ensure that audiences are aware of artificial content.
- Mandatory labelling (Article 50(2): classification and watermarking of deepfakes.
- In contexts that significantly impact individuals' rights or society (such as political manipulation, defamation), deepfakes may be classified as high-risk.
- Traceability/accountability in creating and disseminating deepfakes: records of processes and data used to generate deepfakes to enable tracking their origins.
- Malicious uses of deepfakes, for example in social scoring or illegal surveillance, are categorised as unacceptable risk and prohibited.

The AI Act came into force on 1 August 2024 and will fully apply from 2 August 2026.

The AI Office, set up in February 2024, and the European Artificial Intelligence Board, will monitor the enforcement of the AI Act. It will help create codes of practice on detecting and labelling artificially generated or manipulated content.

European Media Freedom Act (EMFA)

The European Media Freedom Act (EMFA) aims to protect media pluralism and independence in the EU and boost the integrity of the internal media service market, to improve resilience against disinformation. The EMFA complements the digital services package and the FIMI toolbox, as well as EU competition rules related to impacts of market concentration on media pluralism or independence, and unfair allocation of state resources.

Building on the revised Audiovisual Media Services Directive (AVMSD), the EMFA creates a legally binding framework for national authorities to tackle providers that systematically engage in disinformation and information manipulation and interference, and that abuse internal market freedoms, for example media service providers financed by certain third countries. The EMFA also prevents media ownership concentration in order to increase media pluralism, and requires all media services to be transparent about ownership structures. Moreover, the EMFA includes provisions to defend journalists against unfounded, abusive legal actions aiming to silence those working in the public interest, known as strategic lawsuits against public participation (SLAPPs).

A new European board for media services, replacing the European Regulatory Group under the AVMSD, will play a key role in curbing disinformation and foreign interference by coordinating national measures to tackle non-EU media that pose risks to public security. The Board will also organise dialogues between media, civil society and VLOPs, and monitor the latter's compliance with the Code of Practice. It will be operational from February 2025.

The EMFA came into force on 7 May 2024 and will apply from 8 August 2025.

The role of the European Parliament

Over the past two legislatures, Parliament used a mix of tools to address the challenges to the democratic information sphere: non-legislative resolutions, hearings, and its budgetary power. The latter was key to the evolution of the EEAS StratCom team, with Parliament's 2016 resolution on strategic communication to counteract anti-EU propaganda by third parties (2016/2030(INI)) calling for the StratCom Task Force to be turned into 'a fully fledged unit [...] with proper staffing and adequate budgetary resources [via] a dedicated budget line'.

The creation of the first Special Committee on foreign interference in all democratic processes in the EU, including disinformation (INGE) in 2020, helped increase the EU's focus on FIMI. After its mandate expired, Parliament created a new Special Committee (ING2) in March 2022, to follow up, which operated until 2023. In resolutions adopted in March 2022 and June 2023 respectively,

Members called for a common EU strategy to tackle foreign interference and disinformation, including via support for independent media, fact checkers and researchers. Members also urged the Commission and the EEAS to consider creating a European Centre for Interference Threats and Information Integrity.

In addition to the external focus of the two Special Committees on Foreign Interference, Parliament's work to counter information manipulation has cut across policy areas, reflecting the EU's whole-of-society approach. This includes the work of the Committee on Civil Liberties, Justice and Home Affairs (LIBE) on the media freedom act; the Committee on the Internal Market and Consumer Protection (IMCO) on the DSA, GDPR, copyright, and political advertising; and the subcommittees on defence (SEDE) and human rights (DROI) and the special committee on AI (AIDA).

Information integrity: Sustainable ecosystems at home and abroad

In recent years, information integrity has emerged as a concept to represent potential national, international and multilateral solutions to the wider challenges to the information space. Most definitions of information integrity emphasise the importance of sustainability; boosting the entire information ecosystem and upholding human rights. In June 2024, the UN published its Global Principles For Information Integrity, which include societal trust and resilience; healthy incentives; public empowerment; independent, free, and pluralistic media; and transparency and research. The document underlines that '[e]fforts to strengthen information integrity are crucial to preserve and further advance the [Sustainable Development Goals]' by fighting misinformation, disinformation, and hate speech while protecting human rights such as freedom of expression.

Other examples of multilateral and international cooperation on information integrity include the OECD's DIS/MIS Information Resource Hub, co-chaired by France and the US. Following the EU-US Trade & Technology Council Ministerial in April 2024, the US Department of State and the EEAS launched a joint US-EU Coordination Mechanism on Information Integrity in the Western Balkans. It aims to 'empower like-minded partners to become self-sufficient in addressing the FIMI threat and [...] reduce Russia's and [...] China's ability to employ propaganda and information manipulation campaigns.' The May 2024 US International Cyberspace & Digital Policy Strategy lists defending information integrity as a priority, highlighting the TTC, OECD and G7 as key cooperation fora.

Outlook: Shielding democracy through digital defence, enforcement and resilience

Implementation of key parts of the EU's response to information manipulation and disinformation, is due during the new legislature, amid continued polycrises that create fertile ground for (and provide incentives for authoritarian actors to expand) information manipulation. At the same time, the underlying friction between Brussels and some US tech companies could lead to a more aggressive corporate tech diplomacy defending the US market-driven model, to avoid EU regulation impacting their business models. Ursula von der Leyen's political guidelines for 2024-2029, on the other hand, reconfirm the EU's focus on rights and democratic values, underlining the 'need to do more to protect our democracy', including by stepping up digital enforcement. Under 'Protecting our democracy, upholding our values', von der Leyen proposes:

- A new **European Democracy Shield** to counter foreign information manipulation and interference online. Citing the examples of Viginum in France and the Swedish Psychological Defence Agency, the European Democracy Shield would aim to detect, analyse and proactively counter disinformation and information manipulation.
- **Education and awareness**: Focus on societal resilience and preparedness, through increased digital and media literacy; boosting prevention through pre-bunking; creating a European network of fact-checkers, available in all languages.

- **Digital Services Act**: Strengthen enforcement to ensure that manipulated or misleading information is detected, flagged and removed, as appropriate, in line with the DSA.
- **AI Act**: Address deepfakes, ensure that transparency requirements in the AI Act are implemented and that the approach to AI-produced content is strengthened.
- Preserve and promote **free speech**.

In her speech to Parliament on 18 July 2024, von der Leyen specified that 'the Shield will take into account recommendations from the work of the special committees on foreign interference, to better protect our democracies'. She also underlined the importance of guaranteeing a 'reliable information framework'.

In her mission letters, von der Leyen tasked Commissioner-designate for Democracy, Justice, and the Rule of Law Michael McGrath with leading the work on a new European Democracy Shield. If confirmed, he is to coordinate the work on disinformation, and work with the other Members of the College to fight foreign information manipulation and interference, as well as to step up work on digital and media literacy and on prevention through pre-bunking. McGrath is set to report to Executive Vice-President-designate for Tech Sovereignty, Security and Democracy, Helena Virkkunen, whose focus includes 'strengthening the resilience and the functioning our democracy, notably through the new European Democracy Shield and work to counter harmful disinformation which can cause rifts in society and weaken our democracy'. Virkkunen is also set to be tasked with enforcing the Digital Services Act and the Digital Markets Act. Moreover, Virkkunen is likely to work to promote EU digital norms and standards internationally, and to ensure a leading role for the EU in global digital governance.

ENDNOTES

1 Information manipulation is most often accompanied by other types of interference, alongside elite capture, cyber attacks, academic interference, and economic cohersion.

2 A. Polyakova and C. Meserole, Exporting digital authoritarianism: The Russian and Chinese Models, Brookings Institution, 2019. See also Freedom on the Net 2018, The Rise of Digital Authoritarianism.

3 J.S. Pearson, 'Defining Digital Authoritarianism', *Philosophy and Technology*, Vol. 37(73), 2024.

4 The Kremlin's goals are very explicit: on 13 June 2024, deputy chair of Russia's Security Council, Dmitry Medvedev, used his official Telegram channel to urge Moscow to use disinformation to incite social unrest in the West in response to the latest round of US and EU sanctions against Russia: 'Are they afraid of social explosions? Let's cause some!' He added: 'Let's turn their life into a complete crazy nightmare in which they will not be able to distinguish wild fiction from the realities of the day, infernal evil from the routine of life'.

DISCLAIMER AND COPYRIGHT

eprs@ep.europa.eu (contact)

www.eprs.ep.parl.union.eu (intranet)

www.europarl.europa.eu/thinktank (internet)

http://epthinktank.eu (blog)