ATHENA MINI-DELPHI CONSENSUS REPORT

Authors: Pedro Peres Cavalcante (Trilateral Research); David Wright (Trilateral Research).

Introduction

Foreign information manipulation and interference (FIMI) operations are set to explode with the latest technological developments. To investigate how new and emerging technologies are likely to impact disinformation campaigns advanced by foreign malicious actors, the EU-funded ATHENA project invited experts from civil society, academia and the private sector to take part in a mini-Delphi study as part of a horizon scan of how these technologies may shape FIMI operations.

The impacts of artificial intelligence technologies stand out among the responses we gathered. As AI's latest developments increasingly reshape productive sectors, its malicious uses for spreading disinformation have also sprung up.

Countering FIMI depends on a collective effort: policy and decision-makers, individuals, the fact-checking and defender community, the academic and scientific sectors, all have a role to play. Digital technologies' incessant evolution demands that those in charge of building resilience keep pace with the threats these tools enable.

Our questionnaire sought to capture expert insight on the technological trends most likely to shape FIMI in the next three to five years. The aim was to identify both high-risk developments and potential areas for action, including how governments, civil society and other actors might anticipate, detect and mitigate such threats.

Methodology

To get a fair representation of how technologies might impact FIMI in the near future, we invited experts to participate in a two-round mini-Delphi. The Delphi methodology has long been used in development.[1] It has also been a recurrent methodological preference in studies covering disinformation forecasting.[2] The Delphi method leverages the expertise of participants in a study to generate forecasting and judgement in contexts in which pure model-based statistics methods are disadvantageous given the lack of pre-existing historical, economic and technical data.[3] A Delphi-based study typically uses three rounds of consultation to move towards a more stabilised collective view of

---

[1] See: Adler, Michael, and Erio Ziglio (eds.), *Gazing into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health*, Jessica Kingsley Publishers, London, February 1996. https://books.google.es/books?id=jo1Z1JZIrKIC&printsec=frontcover&hl=es&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false. See also [3]

[2] See, for examples, Kruger, Alan, Maria Saletta, Areej Ahmad and Philip Howe, "Structured expert elicitation on disinformation, misinformation, and malign influence: Barriers, strategies, and opportunities", *Harvard Kennedy School Misinformation Review*, Vol. 5, No. 7, 19 Mar. 2024. https://doi.org/10.37016/mr-2020-169; Arribas, Cristina M., Rubén Arcos and Manuel Gertrudix, "Application of the Delphi methodology to forecast competencies oriented towards AI-driven disinformation detection in an expanded DigComp framework", *MethodsX*, Vol. 14, 2025, Article 103401. https://doi.org/10.1016/j.mex.2025.103401

[3] Rowe, Gene, and George Wright, "The Delphi technique: past, present and future prospects – introduction to the special issue", *Technological Forecasting and Social Change*, Vol. 78, No. 9, Nov. 2011, pp. 1487–1490. https://doi.org/10.1016/j.techfore.2011.09.002

a certain topic, which is associated with consensus building.[4] The interpretation of what achieving consensus entails, however, varies across studies.[5]

Our mini-Delphi horizon scanning was structured in two rounds. The first consisted in gathering feedback on an online questionnaire hosted in Microsoft Forms comprising seven open and closed questions about technologies likely to impact disinformation in the next three to five years (see Annex 1).

We emailed the questionnaire to 85 potential respondents, comprising 23 members of the ATHENA Stakeholder Board and 62 disinformation experts selected from the project's contact list from academia, security practitioners, civil society organisations, journalists and government agencies.

Due to invalid email addresses, we successfully reached 75 experts, 27 of whom participated in our mini-Delphi, which gave us a response rate of 36 per cent, significantly higher than most such surveys. We attributed the high response rate partly to the fact that we had only seven questions and partly to the fact that the questions interested the experts.

Among the 27 respondents, 19 were external experts, and eight were members of ATHENA's Stakeholder Board. The figures for stakeholder representation were the following: academia and scientific community (12 respondents); civil society and non-profit organisations (11); private sector (4).

Table 1: respondents by stakeholder groups

Academia and scientific community	44.4%
Civil society and non-profit organisations	40.7%
Private sector	14.8%

We anonymised and aggregated the responses in a preliminary synthesis report. We clustered responses in common umbrella categories, each of which expanded into a micro-level to capture the level of nuance and specificity in different responses.

Proceeding to the second round of the mini-Delphi, we then emailed the synthesis report to respondents and invited them to review it and, if they so wished, to amend their individual responses or to comment on the range of responses received from their anonymous peers. We received additional comments from only three respondents, which we have incorporated into this expanded consensus report.

While the Delphi method benefits from building consensus through subsequent exposure of respondents to aggregated results, our mini-Delphi focussed on capturing a range of predictive views, rather than restricting or excluding results. By giving respondents the opportunity to amend their responses or comment on the collective responses from their peers, we created a de facto consensus – no one objected to the views of other experts. We also found a statistical consensus too, for example, with the responses to Question 3, where 25 out of 27 experts predict an explosive growth of between 100 to 1000 per cent in disinformation in the next three to five years.

[4] Dalkey, Norman, and Olaf Helmer, "An experimental application of the DELPHI Method to the use of experts", *Management Science*, Vol. 9, No. 3, 1963, pp. 458–467. https://doi.org/10.1287/mnsc.9.3.458
[5] Powell, Charles, "The Delphi technique: myths and realities", *Journal of Advanced Nursing*, Vol. 41, No. 4, 2003, pp. 376–382. https://doi.org/10.1046/j.1365-2648.2003.02537.x

We intentionally designed the questionnaire to be short – only seven questions – to make it easy for experts to complete and so as not to be a burden on their time. Originally, we estimated a completion time of about five minutes, which proved to be a severe under-estimate. Six respondents answered the questionnaire in fewer than 10 minutes; nine others took between 10 and 20 minutes; and 12 answered in more than 20 minutes. As the questionnaire was self-administered online, respondents answered without any supervision, at a rhythm of their own choosing. While the completion time varied according to each respondent, we consider that a completion time of more than 20 minutes likely indicated that the respondents answered in more than one session.

We deliberately anonymised responses so that respondents did not know who else was responding to the questions to avoid the risk of bias or influence. Nevertheless, when we emailed the synthesis report to respondents, we asked them if they would like to be named in the acknowledgements in the final, consensus report. Only three of the 27 agreed to have their names mentioned.

We shared the first version of this report with ATHENA's Ethics Advisory Board and Stakeholder Board and invited their comments. We are including this final consensus report as a chapter in the ATHENA D4.1 deliverable on scanning the horizon for technologies likely to impact FIMI in the next three to five years.

In the following pages, we present the questions and responses we received and then conclude with a summary of our findings and recommendations.

1. Ranking of technologies most likely to impact disinformation in the next three to five years

We asked respondents to rank eight technologies most likely to impact disinformation in the next three to five years. The figure below displays the responses.

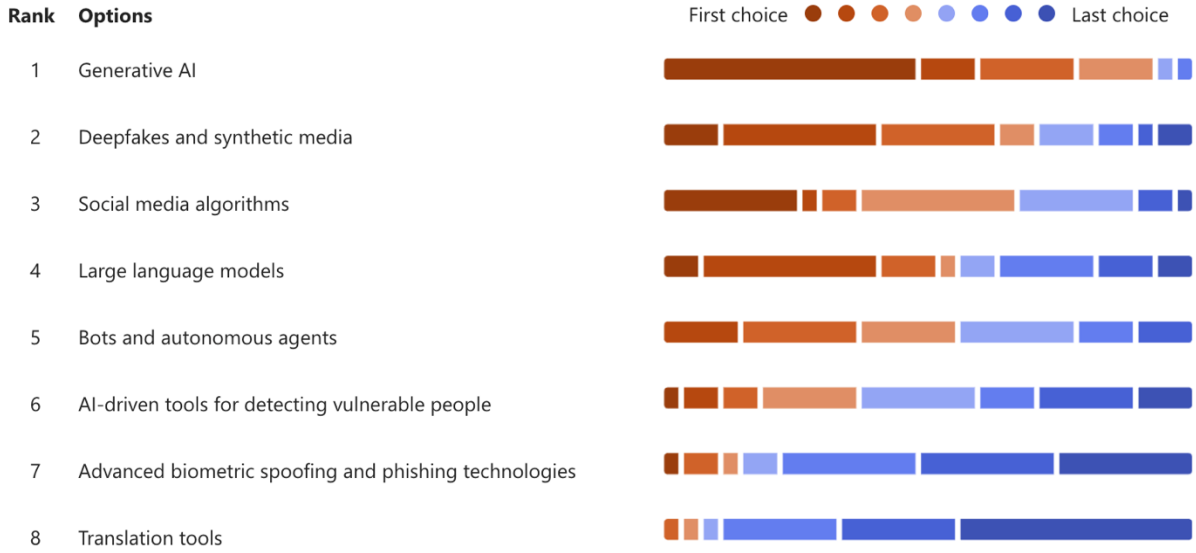

Figure 1: Technologies ranked by expected impact on disinformation (next 3–5 years). Ranking reflects both frequency of selection and average position in respondents' individual rankings. See Annex 2 for a detailed breakdown.

The list above ranks alternatives by weighing responses according to their place in each respondent's personal ranking and overall recurrence among the group (number of times chosen by respondents). This means that the highest-ranking positions correspond to technologies that were among the first positions chosen by respondents and had higher numbers of respondents choosing them.

For instance, despite deepfakes and synthetic media (rank #2) having had 11.1% of experts placing them first in their ranking (see annex 2 for the ranking for the breakdown of each technology), it still ranked above social media algorithms (rank # 3), which had 25.9% respondents placing them first, because more respondents chose deepfakes and synthetic media as their second (29.6%) and third options (22.2%) than social media algorithms – which accumulated 3.7% and 7.4% respectively for second and third choices. Proportionally, this resulted in higher weights attributed to deepfakes and synthetic media.

2. Other technologies likely to impact disinformation in the next three to five years

The answers gathered under this question revealed fewer technologies other than those presented in the first question. We grouped them under four categories: 1) microtargeting / demographics / profiling; 2) closed networks / messenger apps; 3) quantum computing; and 4) cryptocurrencies.

Microtargeting/demographics/profiling were the most cited new technologies suggested by respondents. When mentioning them, respondents gave adjacent answers on the use of data about populations or individuals to enhance disinformation dissemination. Only one respondent mentioned that the data could be obtained through social media channels. Closed networks and messenger apps placed second among respondent mentions. They cited encrypted or "closed" message apps or networks (Telegram, Whatsapp), mentioning among their advantages for malicious actors: (a) a way of evading monitoring or moderation; (b) facilitation of command and control. Quantum computing and cryptocurrencies also appeared in the answers to this question, with fewer mentions – two people cited the former, while one respondent stressed the role of the latter as a facilitator for disinformation fundraising.

While this question focussed on identifying technologies different from those listed in the previous question, a considerable number of respondents emphasised generative AI, already covered in the first question. Some of these respondents suggested its potential use in microtargeting and biometric spoofing.

A few respondents also mentioned "algorithms". Some referred to algorithms deployed by social media platforms – also covered in the first question – suggesting enhanced recommendation systems that can be used for optimised dissemination of disinformation. One respondent associated algorithmically fuelled disinformation with data poisoning of popular LLMs to render biased results. Another respondent highlighted the positive use of algorithms by the defender community, giving fact-checking algorithms as an example.

About 30 per cent of respondents didn't mention any new technologies, some of which explicitly said the list presented in the first question was sufficiently comprehensive.

3. Technologies' impact estimation

We asked respondents to ponder how impactful the technologies mapped will be on disinformation in the next three to five years. 85% of respondents expect impacts ranging from 100 to 1000 per cent over the next three to five years, as depicted in Figure 2 below.

3	Up to 25% increase in the amount of disinformation
12	Up to 100% increase in the amount of disinformation
12	Up to 1000% increase in the amount of disinformation

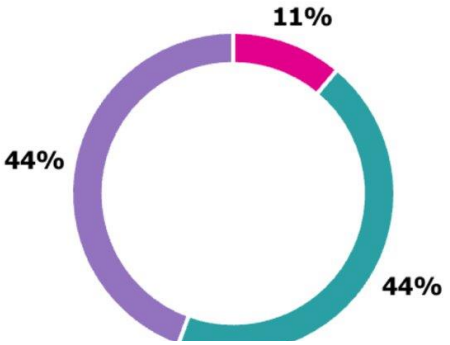

Figure 2: Respondents' assessments of the impact's scale new technologies are likely to have on disinformation in the next three to five years.

We find the response alarming, but not surprising. There is already so much disinformation prevalent today, it is taking on the proportions of a pandemic event. New technologies facilitate the mass production of disinformation; they greatly increase attackers' capabilities to reach more people more often and more convincingly. If the assessment of the vast majority of experts is correct, governments and individuals will need to take even more pre-emptive measures than now. Respondents had some suggestions for such measures (see below).

4. Impacts from the use of these technologies

When asking about the impact of new technologies on disinformation, we presented respondents with four possible negative impacts: (a) greater social polarisation; (b) an increase in distrust; (c) an increase in fraud and cybercrime, and (d) an undermining of democratic institutions and electoral processes. In their responses, some experts provided more nuance and context to these impacts by identifying how they can potentially manifest in specific domains.

We clustered impacts identified by respondents under the following domains: politics; economy; health; society; journalism and activism; special groups and other/assorted. Here in the bulleted points are respondents' views on impacts of new technologies on disinformation.

4.1 Political impacts

Respondents highlighted several potential political impacts of disinformation and foreign influence operations. They warned that such activities could amplify foreign interference, potentially destabilising other countries and increasing conflicts and geopolitical polarisation. Other concerns included the risk of Internet fragmentation as more barriers are imposed on telecommunications, the undermining of relationships between international partners and the growing concentration of power in the hands of information-driven autocracies and oligarchies.

4.2 Economic impacts

Respondents identified several key economic impacts of disinformation. They highlighted the manipulation of stock markets, where false or misleading information can cause erratic price movements, market volatility and significant financial losses for investors. Disinformation can also damage brands, erode value and disrupt the ability of companies to deliver products and services, as trust and reputation are undermined. One respondent emphasised the importance of "following the money", suggesting the need for research into financial flows that empower certain groups over others.

4.3 Health impacts

Respondents highlighted that disinformation can have serious health impacts, notably by delegitimizing scientific consensus and public health guidance. This undermines efforts in areas such as vaccination, climate change, pandemics, and healthcare more broadly. As a result, trust in science and public health institutions erodes, making people more vulnerable to misinformation and less likely to follow evidence-based recommendations, which can lead to poorer health outcomes and increased public health risks.

4.4 Social impacts

Disinformation has led to a significant increase in distrust across society. Respondents noted the erosion of trust not only in alliances and institutions but also in academic and scientific experts, including healthcare professionals. This distrust extends to relationships between citizens and institutions at local, regional, and national levels, and is further complicated by the potential emergence of politically biased chatbots that cater to different factions, deepening divisions.

The impacts on democracy, fundamental and human rights are profound. Disinformation contributes to the weakening of freedom of expression and the press, increases self-censorship, and diminishes core democratic values. Respondents also highlighted that in conflict zones, misleading information can result in human rights violations by directing people away from safe resources, and that state attempts to counter threats may inadvertently infringe on individual freedoms.

Disinformation can fuel violence and unrest, with respondents pointing to increases in real-world violence, sectarian conflicts, and broader societal unrest that threaten social cohesion and could lead to the disintegration of society.

Respondents observed a normalisation of conspiracy thinking and a general impoverishment of societal reasoning. This includes shorter attention spans, a reliance on fleeting trends, making it hard to build strategies around information campaigns and media literacy efforts. They also saw a collapse of shared worldviews and collective sense-making, all of which undermine rational debate, erode neutrality and contribute to information fatigue.

Finally, other societal impacts include a more fragmented and disconnected social media landscape, with communities increasingly divided along identity or philosophical lines. Respondents noted a shift from broad, community-wide platforms to smaller, like-minded networks, delays in crisis response, declining media literacy, and the entrenchment of more resilient filter bubbles.

4.5 Impacts on journalism and activism

Respondents observed that disinformation has serious consequences for journalism and activism. It is used to suppress dissent and target activists and journalists, often exposing them to harassment and intimidation. Disinformation also works to delegitimise independent media and experts, undermining public trust in journalism and eroding the credibility of those who adhere to professional standards.

4.6 Impacts on special groups

Respondents noted that disinformation has particularly harmful impacts on special groups. Specific communities, such as immigrants and minorities, are often targeted with hyper-focused harassment and manipulation, which can reinforce damaging stereotypes, increase fear, and undermine their sense of safety and belonging. This environment of hostility and misinformation can contribute to higher rates of depression, especially among younger generations who are more exposed to online content and social pressures. Additionally, respondents highlighted a worrying trend of increased withdrawal of women, LGBTQ+ individuals, and other minority groups from civic and political participation, as targeted disinformation and harassment discourage engagement and further marginalise these communities.

4.7 Other impacts

Respondents identified several additional impacts of disinformation. These include increased risks of identity theft and significant reputation damage to both public and private institutions, as well as individuals. Disinformation can also contribute to the rise of "crime as a service," enabling opportunistic actors to exploit digital vulnerabilities. It undermines decision-making among high-level public officials and private sector leaders, making organisations more susceptible to manipulation. Respondents also noted the emergence of more targeted influence operations and the concealment of platform ownership or operations, which further complicates efforts to ensure transparency and accountability in the digital ecosystem.

5. Recent FIMI cases

We asked respondents if they were aware of any recent examples of FIMI occurrences involving the deployment of the technologies mentioned before by Russia or China. Respondents cited the following stories, which we have grouped by issue or theme:

Table 2: Stories citing Russian and Chinese FIMI using new technologies

Artificial intelligence	"A well-funded Moscow-based global 'news' network has infected Western artificial intelligence tools worldwide with Russian propaganda", NewsGuard, 6 Mar 2025.	RU
	"Russian disinformation network flooded training data to manipulate Western AI chatbots, study finds", *Meduza*, 7 March 2025	RU
	"Russia seeds chatbots with lies. Any bad actor could game AI the same way", *The Washington Post*, 17 Apr 2025	RU

	"The evolving role of AI-generated media in shaping disinformation campaigns", *DFRLab*, 1 May 2025	RU
	"Pro-Russian influence operation targeting Australia in lead-up to election with attempt to 'poison' AI chatbots, *ABC News*, 2 May 2025	RU
	"Embarrassingly for the Kremlin, Russian AI Isn't Good Enough for Its Own Disinformation", *The Moscow Times*, 7 June 2024	RU
	"Taiwan says China using generative AI to ramp up disinformation and 'divide' the island", *Reuters*, 8 Apr 2025	CN
	"OpenAI finds Russian and Chinese groups used its tech for propaganda campaigns", *The Washington Post*, 30 May 2024	RU CN
The Pravda network	"Russian disinformation network 'Pravda' grew bigger in the EU, even after its uncovering (Analysis)", EDMO, 20 Apr 2024	RU
	"Russia's so-called "Pravda" network expands worldwide", DFRLab, 24 Feb 2025	RU
	"Global Offensive: Mapping the Sources Behind the Pravda Network", Globsec,16 May 2025	RU
Moldavia	DFRLab "Cross-platform campaign sows anti-Europe division in Moldova", 26 Mar 2025	RU
	DRFLab, "Telegram network seeks to manipulate Moldova's local political discourse", 6 Mar 2025	RU
The DoppelGänger campaign	U.S Cyber Command, "Russian Disinformation Campaign "DoppelGänger" Unmasked: A Web of Deception", 3 Sept 2024	RU
Others	Wikipedia, "War on Fakes, 16 May 2025	RU
	The Spamouflage operation	RU
	Philstar Global, "Maritime council says China disinformation still rampant", 15 May 2025	CN
	Microsoft Threat Analysis Center, "As the U.S. election nears, Russia, Iran and China step up influence efforts", 23 Oct 2024	RU CN
	The Polish Institute of International Affairs, "China Increasingly Implicated in Russian Destabilisation Efforts Against the EU", 24 Jan 2025	RU CN

Table 3: Breakdown of stories by country

❌	6 respondents did not cite any stories
RU	19 respondents featured stories involving Russia
CN	5 respondents featured stories involving China

From the above results, we can see that the experts are well aware of *recent* instances of Russian and Chinese-propagated disinformation. The cited stories are another indicator of the growing disinformation menace.

6. Countermeasures governments can take against disinformation

Respondents recommended that governments take a series of countermeasures to address the use of technologies for disinformation purposes. We clustered their responses under several categories: legislation/regulation; media and digital literacy efforts; investments in technology; supporting the defending community; platform regulation and accountability; interventions in content production; algorithmic transparency; redesigning budgets and assorted measures. We present the results below.

6.1 Legislation and regulation

Legislative and regulatory countermeasures ranked first in the list of countermeasures provided by respondents. Hence, experts expect governments to take the lead in countering disinformation. We distinguished general mentions to legislative or regulatory action from those referring to existing legislation.

Respondents generally called for the development of stricter laws and regulations to address deliberate disinformation and the challenges posed by social media. They suggested that policies should more clearly define the boundaries of free speech and specify what constitutes a threat, while avoiding approaches that could be seen as sweeping censorship. Legislation and regulation were seen as important tools for restoring order and transparency in the information environment, countering manipulation and adversarial control. Respondents emphasised the need for regulations to be regularly updated in line with technological advances and to be implemented at both national and EU levels. One respondent said that AI regulation should not focus solely on disinformation, but rather on the broader capabilities of AI systems, especially those that could surpass human intelligence.

Respondents called for the full enactment and implementation of existing legislation, particularly the AI Act, which entered into force on 1 August 2024 and will become fully applicable by 2 August 2026, with some provisions—such as bans on certain AI systems and AI literacy requirements—already in effect since February 2025. They also recommended streamlining the Digital Services Act and the Code of Conduct on Disinformation to ensure clear and effective regulatory frameworks. Updating relevant regulations was seen as necessary, especially regarding generative AI, securing elections, and protecting critical infrastructure. Finally, respondents stressed the importance of enforcing obligations for social media platforms to detect and mitigate coordinated inauthentic behavior, bot amplification, and foreign influence operations.

Respondents called for the full enactment and implementation of existing legislation, particularly the AI Act, which entered into force on 1 August 2024 and will become fully applicable by 2 August 2026, with some provisions—such as bans on certain AI systems and AI literacy requirements—already in effect since February 2025. They also recommended streamlining the Digital Services Act and the Code of Conduct on Disinformation to ensure clear and effective regulatory frameworks. Updating relevant regulations was seen as necessary, especially regarding generative AI, securing elections, and protecting critical infrastructure. Finally, respondents stressed the importance of enforcing obligations for social media platforms to detect and mitigate coordinated inauthentic behaviour, bot amplification and foreign influence operations.

6.2 Media and digital literacy efforts

Respondents emphasised the importance of investing in media and digital literacy, including funding for national campaigns and targeted initiatives. They recommended enhancing media literacy efforts

through educational programs that use varied approaches, such as gamification, and designing strategies tailored for both children and seniors. These efforts should address all levels of education and keep pace with emerging technologies and their potential misuse for disinformation.

Building citizen resilience was highlighted as a key goal of inclusive strategies that reach vulnerable and marginalised groups, aligning with ongoing European campaigns such as those supporting children and seniors and equipping them with the skills needed to navigate an evolving digital landscape.

6.3 Investments in technology

Respondents recommended developing and investing in advanced technologies for both detection and response to disinformation. They called for increased funding for research and development to create models capable of identifying, analysing and preventing disinformation, as well as investments in tools that enable comprehensive analytic reporting tailored to different audiences. Proactive awareness and competence building were also highlighted as priorities, ensuring that individuals and organisations are equipped to recognise and counter emerging disinformation threats.

6.4 Supporting the defending community

Respondents suggested strengthening the infrastructure that supports the defender community by enhancing community engagement and network building, which fosters better collaboration and information sharing among stakeholders. They recommended investing in financial investigation tools and capacity building to improve the ability to track and disrupt funding behind disinformation operations. Standardising reporting – such as through the open-source DISARM Red Framework – was seen as essential for creating a common language and methodology for documenting and sharing information about influence operations, making collective defence more effective. Finally, creating incentives for the development and adoption of open-source, common tools would further empower the defender community to respond to evolving threats in a coordinated and efficient manner.

6.5 Platform regulation and accountability

Respondents expressed the view that platform regulation and accountability should be strengthened by holding social media platforms and online content providers responsible until they demonstrate effective moderation. They suggested that platforms should be required to label or remove detected disinformation, clearly identify AI-generated or manipulated media, and provide greater transparency about their data sources, data usage, and recommendation algorithms. Additionally, respondents called for clear standards to be established for how platforms moderate and ensure the quality of content within their domains.

6.6 Algorithmic transparency

Respondents signalled the need for more algorithmic transparency, both in general terms and in regard to social media platforms.

6.7 Interventions in content production

Respondents suggested several interventions in content production to address emerging risks. They advocated for limiting the amount of content a user can produce at mass scale, as a way to curb the spread of disinformation. They proposed restricting the topics and entities that can be generated through large language models (LLMs) and generative AI, aiming to prevent misuse of these technologies. One respondent raised the idea of controlling information within social networks, but acknowledged that such measures could be ethically questionable.

6.7 **Redesigning budgets**

One respondent said it was necessary to reshape military budgets to account for hybrid warfare defence.

6.9 Establishing partnerships and cooperation

Respondents emphasised the importance of establishing partnerships and advancing cooperation to effectively counter disinformation. They advocated for sharing best practices with EU Member States, building early detection systems in collaboration with media and NGOs, and working closely with social media platforms and other governments. They highlighted the need to strengthen international data-sharing mechanisms to better track and understand disinformation infrastructure.

6.10 Other measures

Respondents proposed a range of additional measures to counter disinformation and enhance information integrity. They emphasised the need to strengthen institutions and trusted messengers, recognising the value of credible voices in building public resilience. Support for organic, not-for-profit, positive technology development was also highlighted, alongside the promotion and funding of reliable sources such as public broadcasters. Respondents advocated for the development of rapid attribution capabilities to publicly expose disinformation actors and tactics, and for the imposition of sanctions or legal actions against individuals and entities engaged in AI-enabled FIMI. They also called for greater transparency regarding the identity of commercial users of social media and AI products, the creation of proactive verification systems[6] to help users identify genuine content, and the establishment of specialised observatories and government-run trustworthy websites where the public can verify information, particularly during crises.

7. Countermeasures individuals can take against disinformation

Respondents presented a wide variety of measures individuals could take to protect themselves against FIMI:

Fact-checking & verification tools

Respondents recommended that individuals use fact-checking and verification platforms and tools to assess the accuracy of online information. They suggested verifying whether a video is genuine by searching for keywords spoken in the video or checking the original media outlet that supposedly published it. Additionally, consulting reports from reputable investigative journalists and trusted news sources was highlighted as an effective way to confirm the reliability of information encountered online.

Critical thinking digital skills

Respondents highlighted several critical thinking digital skills that individuals should develop to counter disinformation. These include questioning the logic and validity of arguments presented in information, scrutinising the financial backing of sources, and checking for user comments that may disapprove or challenge a story. They also advised checking if fact-checking websites have flagged a source as prone to disinformation and being able to recognise common manipulation tactics. Additional recommendations include critically assessing whether a story relies on strong emotional appeal, seeks to confirm existing biases, or demands immediate action; being cautious with breaking news and waiting for corroboration from multiple sources; and maintaining scepticism even toward videos, given

[6] For example, the Verify company says "With Verify, content publishers can license and track the use of their content — and content licensees can power AI solutions with trusted information from verified, real-world publishers". https://www.verifymedia.com

the prevalence of manipulated media. Developing these skills fosters resilience and informed decision-making in the digital environment.

Media literacy habits

Respondents described a set of media literacy habits that individuals can adopt to protect themselves from disinformation. These include seeking out diverse perspectives on the same story and prioritising local and independent media over relying solely on "fast food media" sources like closed messenger channels or social media feeds. They emphasised the importance of being mindful of the channels through which information is accessed, not just the content itself, and of understanding the credibility and journalistic standards of information sources. Additional habits include not sharing questionable information, resisting the urge to instantly repost emotionally charged content, and routinely reporting suspicious or false content to platforms or relevant organisations. Respondents also recommended consulting fact-checking websites (e.g., Newsguard), checking user comments for dissent, and referring to reputable investigative reporting. Finally, they suggested using government-run trustworthy websites for verifying information during crises, reinforcing the need for ongoing vigilance and critical engagement with all forms of media.

Digital security

Develop cybersecurity habits, such as minding one's digital footprint, using strong passwords and enabling two-factor authentication was another of the measures mentioned.
They also highlighted the value of developing advanced technological skills, such as using open-source intelligence (OSINT) tools and verification plugins, and staying informed through ongoing learning. Finally, one respondent suggested leveraging large language models (LLMs) as a resource for information and verification.

Social trust and offline engagement

Respondents emphasised that building social trust and engaging offline are important countermeasures against disinformation. They recommended fostering connections with colleagues, acquaintances and neighbours to strengthen trust and social cohesion. When encountering falsehoods shared by others, respondents advised engaging constructively and avoiding antagonism and advocating to local policymakers for stronger safeguards.

Other measures

Although not strictly measures individuals can take, a respondent also suggested implementing European and national campaigns and introducing disinformation-related subjects into the secondary school curriculum to build resilience and critical digital skills among young people. Banning the use of devices in certain social settings, such as schools, was also suggested to reduce social media interaction and limit exposure to disinformation during key periods of social development.

8. Summary of findings and conclusions

Although the scale of impact caused by technologies in disinformation operations is hard to measure and reflects a speculative forecast, the findings of the questionnaire point to an explosive growth in disinformation, with 88% of respondents indicating an increase in disinformation over the next three to five years of between 100% or 1000%. This demands a corresponding increase in countermeasures against FIMI.

Results from the first and second questions indicate that respondents stress how content is produced, where and how it circulates and who is being targeted. The predominant emphasis on both questions is

on content production, with generative AI applications ranking among the highest positions in the first question and being the most cited in the second question with added level of detail. The responses highlight the intertwining of who is being targeted and how. In the second question, smarter targeting through microtargeting or profiling was the most cited in the category of other technologies that may impact disinformation. Although microtargeting and profiling do not constitute technologies per se, they are still relevant as techniques to be associated with other technologies that may impact disinformation. Lastly, the medium dynamics of how content is disseminated appear among the most anticipated impacts: social media algorithms and closed networks and messaging apps were among the most cited in both responses.

The most recurrent examples given by experts on impacts caused by the use of technologies spanned politics, economy, society and public health demonstrating the all-encompassing reach of disinformation:

Table 4: Most cited impacts expected from the use of technologies in disinformation in the next three to five years

1	Amplification of foreign influence operations associated with countries' destabilisation
2	Delegitimization of scientific consensus and public health guidance
3	Increase in violence, real-world harms and sectarian conflicts
4	Hyper targeted harassment and manipulation towards specific communities
5	Manipulation of the stock market

Respondents presented a wide range of responses regarding societal impacts. They identified several examples across the following five sub-categories: increase in distrust; impacts on democracy, fundamental and human rights; rise of violence and unrest; effects on collective thinking; and other assorted societal impacts.

When asked about stories featuring FIMI operations involving technological use by China and Russia, the episodes cited by respondents predominantly involved Russian influence, confirming data from other sources such as the 3rd EEAS Report.[7]

Regarding countermeasures to be taken by governments, responses concentrated on mentioning legislative and regulatory initiatives, platform accountability, and investments in media literacy. While many respondents emphasised legislative and regulatory countermeasures, the range of responses captured competing values and priorities. While some experts called for stricter regulations, others mentioned considerations such as defining clearer limits for freedom of expression and avoiding censorship. Respondents also signalled the need for continuous updates of regulatory responses in tandem with the rhythm of technological evolution.

Respondents recognised the need for more accountability by platforms: while some referred to it in broad terms, others identified specific actions or domains for improvement, such as algorithmic

[7] European External Action Service, 3rd EEAS Report on Foreign Information Manipulation and Interference Threats – Exposing the architecture of FIMI operations, Brussels, 19 Mar 2025. https://www.eeas.europa.eu/eeas/3rd-eeas-report-foreign-information-manipulation-and-interference-threats-0_en

transparency and imposing an obligation to label AI-generated content and manipulated media. Many experts stressed the allocation of funding to media and digital literacy initiatives, with a few indicating ways to enhance it, such as investing in varied approaches (such as gamification), developing strategies catered to different audiences (by age and education level), and keeping initiatives up to date with the latest technological developments. Experts also stressed the importance of governments investing in the same technologies impacting disinformation through the development and deployment of advanced technologies for detection, analysis, and response, alongside increased funding for R&D to build models that identify and prevent disinformation.

When asked to identify measures to be taken by individuals, respondents mapped a wide variety of media and digital literacy skills and habits. While most responses encompassed media literacy habits and critical thinking skills, we highlight those clustered in "other skills": developing more advanced technological skills (such as using verification plugins or basic search engine tools) and recurring to LLMs to obtain help against disinformation.

As a final conclusion, we note that experts have a high level of concern about a tidal wave of disinformation coming in the next three to five years as threat actors employ new technologies to greatly increase their impact. To brace ourselves against this coming wave, multiple complementary actions must follow. The wide range of countermeasures governments can take and defensive measures for individuals mentioned by respondents must be interpreted in the light of rapid evolution of technology. Pertinent responses to disinformation and FIMI must be swift and deeply attuned with the evolving landscape in which these technologies are used. From the responses we gathered, there is a strong prediction over AI-fuelled malicious activity dominating in impact - with generative AI, deepfakes, synthetic media and LLMs leading in mentions.

We interpret results as a useful current portrait of technologies impacting disinformation; however, we stress the constantly changing scenarios in which they emerge. For instance, The Gartner Hype Cycle of Emerging Technologies, an index for measuring technological advancement and deployment from the ranges of "innovation trigger" to "plateau of productivity", indicated in 2024 a decline of expectations regarding generative AI while other applications, such as autonomous agents and artificial general intelligence, rose.[8] Therefore, much as this mini-Delphi serves to forecast an uncertain scenario, we understand that the main takeaway is to develop frameworks oriented towards adaptability given the fast-paced rhythm in which technologies are inserted. In this sense, although we view public investments in these same technologies for detection, analysis and defensive purposes as a good starting point, we also perceive that increasing resilience towards these threats demands more active forecasting and cooperation involving multiple stakeholders. This can take shape in numerous ways, some of which mapped under this study. As part of a larger work package involving policymaking and forecasting, this mini-Delphi will inform upcoming deliverables and future policy recommendations.

[8] Gartner, "Gartner 2024 Hype Cycle™ for Emerging Technologies highlights developer productivity, total experience, AI and security", press release, Stamford (CT), 21 Aug 2024. https://www.gartner.com/en/newsroom/press-releases/2024-08-21-gartner-2024-hype-cycle-for-emerging-technologies-highlights-developer-productivity-total-experience-ai-and-security

Aknowledgements

We greatly appreciate all of the responses from the 27 experts who participated in our mini-Delphi. The partners gathered and aggregated all responses anonymously. However, we asked respondents if they would like to be identified in the final version of the consensus report. These experts agreed to be acknowledged by name and we thank them most heartily too:

Federico Sicurella
Nikos Sarris
Thanos Sitistas

Annex I

Mini Delphi on foreign information manipulation and interference (FIMI)

Section 1

Questionnaire

1. Please rank which technologies you think are most likely to impact disinformation in the next three to five years. Ranking numbers will appear while you reorder items.

    1. Generative AI
    2. Deepfakes and synthetic media
    3. Social media algorithms
    4. Bots and autonomous agents
    5. Advanced biometric spoofing and phishing technologies
    6. Translation tools
    7. AI-driven tools for detecting vulnerable people
    8. Large language models

2. Are there some other technologies that you think will be impactful on disinformation in the next three to five years?

3. How impactful do you think these technologies will be on disinformation over the next three to five years?

    (a) There will be no significant increase in the amount of disinformation caused by use of these technologies
    (b) I estimate up to a 25% increase in the amount of disinformation caused by use of these technologies
    (c) I estimate up to a 100% increase in the amount of disinformation caused by use of these technologies
    (d) I estimate up to a 1000% increase in the amount of disinformation caused by use of these technologies

4. Here are some impacts that could arise from use of these technologies for disinformation purposes:

- Greater social polarisation
- An increase in distrust
- An increase in fraud and cybercrime
- Undermining democratic institutions and electoral processes

    Can you think of some other impacts?

5. Are you aware of any recent instances of Russia or China using any of the above technologies for foreign information manipulation and interference (FIMI) purposes? If so, please provide a webpage link.
6. What countermeasures do you think our governments should take against use of these technologies for disinformation purposes?

7. Here are some measures individual citizens could take to guard against FIMI:

    - Checking how a story is covered by different sources, including outside one's country;
    - Asking friends and family for their opinion about a story that may be false;
    - Asking themselves about whom does the story benefit;
    - Noticing if anything in the story or image or video seems a bit "off".

    Can you think of some other measures individuals could take?

Section 2

Basic information

Please share your name, email address and organisation. These will only be used to reach out to you seeking feedback on the first version of the report. Your responses will always be anonymised.

8. Name

9. Email address

10. Organisation

Annex 2

Table 5: Answers for Question 1

Technology	1st choice	2nd choice	3rd choice	4th choice	5th choice	6th choice	7th choice	8th choice	Overall ranking position
Generative AI	13 (48 %)	3 (11 %)	5 (19%)	4 (15%)	1 (4%)	1 (4%)	-	-	1
Deepfakes and synthetic media	3 (11 %)	8 (30 %)	6 (22%)	2 (7%)	3 (11%)	2 (7%)	1 (4%)	2 (7%)	2
Social media algorithms	7 (26 %)	1 (4%)	2 (7%)	8 (30%)	6 (22%)	-	2 (7%)	1 (4%)	3
Large Language Models	2 (7%)	9 (33 %)	3 (11%)	1 (4%)	2 (7%)	5 (19%)	3 (11%)	2 (7%)	4
Bots and autonomous agents	-	4 (15 %)	6 (22%)	5 (19%)	6 (22%)	3 (11%)	3 (11%)	-	5
AI-driven tools for detecting vulnerable people	1 (4%)	2 (7%)	2 (7%)	5 (19%)	6 (22%)	3 (11%)	5 (19%)	3 (11%)	6
Advanced biometric spoofing and phishing technologies	1 (4%)	-	2 (7%)	1 (4%)	2 (7%)	7 (26%)	7 (26%)	7 (26%)	7
Translation tools	-	-	1 (4%)	1 (4%)	1 (4%)	6 (22%)	6 (22%)	12 (44%)	8