

EVALUATION OF FIMI-RELATED FRAMEWORKS AND COUNTERMEASURES

ADAC.IO Publication – Deliverable 3.1

Anneli Ahonen, Security, Crime, and Intelligence Innovation
Institute, Cardiff University

Jana Borkowski, Psychological Defence Research Institute, Lund
University

Dr. Martin Innes, Security, Crime, and Intelligence Innovation
Institute, Cardiff University

Björn Palmertz, Psychological Defence Research Institute, Lund
University



Co-funded by
the European Union

About ADAC.io: Attribution, Data, Analysis, Countermeasures and Interoperability

ADAC.io is a Horizon project funded by the European Union and coordinated by the Psychological Defence Research Institute at Lund University. It engages seven partners and has a three-year duration ranging from February 1, 2024 to January 31, 2027.

Based on the concept of Foreign Information Manipulation & Interference (FIMI) as elaborated by the EU EEAS, the purpose of ADAC.io is to protect democracy in the EU by strengthening the ability to deny the intended effects of FIMI on society. ADAC.io hence aims to significantly develop upon current knowledge of how FIMI can be detected, categorised, analysed, shared, and countered.

The project engages the following partners: Alliance4Europe (DE), Debunk EU (LT), Dortmund University - Institution of Journalism (DE), Cardiff University - Security, Crime and Intelligence Innovation Institute (UK); University of Social Sciences and Humanities (PL), Leiden University - The Hague Program for Cyber Norms (NL), Lund University - Psychological Defence Research Institute (SE).

Author(s): Anneli Ahonen, Security, Crime, and Intelligence Innovation Institute, Cardiff University
 Jana Borkowski, Psychological Defence Research Institute, Lund University
 Dr. Martin Innes, Security, Crime, and Intelligence Innovation Institute, Cardiff University
 Björn Palmertz, Psychological Defence Research Institute, Lund University

This work was funded by the European Union Horizon Europe research and innovation program [grant number 101132444 – ADAC.io] and the UKRI under the UK government’s Horizon Europe funding guarantee [grant number 10105669]. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union, the European Commission or UKRI. Neither the European Union, the European Commission, nor the UKRI can be held responsible for them.

Contents

Contents	2
1 Introduction.....	4
1.1 Positioning this report.....	5
1.2 Not One Problem, But Several	7
1.3 When is a Framework Not a Framework	8
1.4 De-Centered Responses Require Interoperability	8
1.5 Report Overview	9
2 The Current Countermeasure Landscape: Challenges and Opportunities	11
2.1 Information influence operations are evolving and adapting	11
2.2 Reporting about cyber and IIO convergence is rising.....	12
2.3 Case-study 1: Countermeasures against Iran's cyber-enabled influence operations peaked in 2024	15
2.3.1 Detection, disruption and legal actions.....	19
2.3.2 Context.....	21
2.4 Case-study 2: North Korea exploits AI in large-scale fraud IT worker scheme	23
2.4.1 How to counter a known threat actor that uses advanced AI	25
2.5 Case-study 3: Azerbaijan attempts to gain regional influence and targets women's rights activists domestically	28
2.6 Reflection	31
3 Methodology and Typology	33
3.1 Investigative Mandate	33
3.2 Contextual Countermeasure Landscape Analysis.....	33
3.3 Framework Identification Methodology, Typology, and Analysis Preparation.....	34
3.4 Countermeasure Evaluation Methodology	37
3.5 Expert Interviews.....	38
3.5.1 Recruitment and Sample.....	38
4 Evaluation of Frameworks related to FIMI Countermeasures.....	40
4.1 Introduction to the Analysis	40
4.1.1 The Categorisation.....	40
4.1.2 Phase of FIMI Response Process.....	40

4.1.3	Type of Framework	42
4.2	Comparative Analysis of the Technical Frameworks Cluster	43
4.2.1	Similarities and Shared Ground	44
4.2.2	Contrasts and Divergences.....	45
4.2.3	Coverage of the FIMI Response Process Phases.....	47
4.2.4	Application of Frameworks in Practice	48
4.2.5	Contributions to a Comprehensive Analytical Approach	49
4.3	Comparative Analysis of the Behavioural Frameworks Cluster.....	50
4.3.1	Similarities and Shared Ground	52
4.3.2	Contrasts and Divergences.....	54
4.3.3	Coverage of the FIMI Response Process Phases.....	56
4.3.4	Application of Frameworks in Practice	57
4.3.5	Contributions to a Comprehensive Analytical Approach	58
4.4	Comparative Analysis of the Operational Frameworks Cluster	61
4.4.1	Similarities and Shared Ground	61
4.4.2	Contrasts and Divergences.....	64
4.4.3	Coverage of the FIMI Response Process Phases.....	66
4.4.4	Application of Frameworks in Practice	67
4.4.5	Contributions to a Comprehensive Analytical Approach	68
4.5	Comparative Analysis of the Strategic Frameworks Cluster.....	71
4.5.1	Similarities and Shared Ground	73
4.5.2	Contrasts and Divergences.....	75
4.5.3	Coverage of the FIMI Response Process Phases.....	77
4.5.4	Application of Frameworks in Practice	78
4.5.5	Contributions to a Comprehensive Analytical Approach	79
4.6	Cross-Categorical Analysis of Frameworks and Conclusions	82
5	Evaluation of FIMI Countermeasures Effectiveness.....	87
5.1	Systematic evaluation vs unsystematic monitoring of countermeasures?....	87
5.2	The view from research literature.....	90
5.3	What can we observe and what more does it tell us?	93
5.3.1	Detection and exposures	93
5.3.2	Disruption.....	94
5.3.3	Legal action.....	97
5.4	Conclusion	100
6	Reflections and Next Steps	101
	References	106
	Annexes.....	119
	Annex A: Framework Descriptions	119
	Annex B: Collected Frameworks Table.....	130

1 Introduction

This report focuses upon understanding and evaluating the evidence for foreign information manipulation and interference (FIMI) countermeasures and the multiple frameworks designed to deliver more effective responses to FIMI related harms, across an array of policy domains and contexts. The definition of the term FIMI employed in this report is as a mostly non-illegal pattern of behaviour that threatens or has the potential to negatively impact values, procedures and political processes, conducted in an intentional and coordinated manner by state or non-state actors (EEAS, 2023). The related term disinformation covers a specific variant of FIMI, engaging information that is false and deliberately created to harm a person, social group, organization or country (Wardle & Derakhshan, 2017).

The discussion in this report is organised around three specific research objectives:

1. To develop an understanding of the evidential supports for different control measures directed towards managing and mitigating the impacts and effects of specific FIMI actions and campaigns, and any associated gaps in knowledge.
2. To evaluate the availability, use and efficacy of the multiple FIMI countermeasures frameworks that have been introduced by a range of different authors and organisations, in support of FIMI harm management and mitigation outcomes.
3. To lay the groundwork for a development agenda for FIMI countermeasures, taking into account emerging risks and threats, that alter the speed, scope, scale and sophistication of the capacities and capabilities available to adversarial state actors and their proxies, as they seek to influence public understanding and political decision-making, aligned with their geopolitical aims.

In positioning and orientating this analysis, it is important to accent the diversity of the information manipulation threats that FIMI countermeasures are arrayed against. The available research evidence in this area clearly documents how a range of FIMI tactics, techniques and procedures are increasingly deployed: in respect of multiple different policy domains and key events (i.e. election integrity, wars and conflicts, climate denialism, public health, social identity politics); across a variety of digital surfaces (including websites, all social media platforms, and LLM models); and target the attention, emotions and behaviour of different audience segments.

Due to limitations on space, it is not appropriate to review each of these. Instead, it is more helpful to summarise them by noting that most adversarial state influence

operations tend to blend both ‘cognitive’ and ‘technical’ components. This is reflected in the ways that some government agencies have chosen to define their counter-FIMI missions and postures. For example, Sweden’s Psychological Defence Agency attends especially carefully to what can be done to manage and mitigate the cognitive and related social-psychological effects of co-ordinated state backed influence campaigns.¹² Contrastingly, VIGINUM in France have developed more of an accent upon manufacturing ‘technical’ disruptions of the networks and infrastructures utilised in the transmission of disinforming, distorting and / or deceptive content.³ This posture is evident in their recent involvement in advocating for the adoption of an ‘Information Manipulation Sets’ Model, as a common approach that will improve inter-operability and information sharing, across the various actors engaged in the FIMI defender community.

There are, however, additional dimensions via which we can usefully conceptualise the different modes of influence that state campaigns might seek to manipulate. Some state information threats are designed to work through developing a persistent exposure to such materials. These we refer to as operating in a ‘chronic’ mode. However, many such operations also deploy in a more ‘acute’ fashion, where they are seeking to exploit and manipulate the propaganda value of high-profile events in the world. It is also the case that whilst many state information influence operations are aiming for mass information effects, some are far more targeted in terms of influencing the decision-making of one individual, or a small group of persons.

Albeit somewhat reductive in terms of the ‘real’ complexity involved in the design and delivery of contemporary state backed information operations, conceptualising them along these several vectors is helpful in clarifying how countering the diversity of opportunities that they can seek to exploit is a key consideration for the construction of specific countermeasures. As will be discussed in more detail, this becomes a particular challenge for the production of frameworks that seek to establish systems, patterns and routines for the detection of FIMI threats and responses to these.

1.1 Positioning this report

Now is an especially timely moment for examining these issues and evaluating the impacts and effects of FIMI countermeasures for two reasons. First, as has been well-rehearsed across many different outlets, the proliferation of agentic AI models and LLMs is having important downstream consequences in terms of the capacity and capability of hostile state actors and their proxies to design and deliver increasingly

¹ <https://mpf.se/psychological-defence-agency/about-us/our-mission>

² https://mpf.se/download/18.6888ebfe19b2bdfbd24538b/1768813534949/Psychological_defence_TGA.pdf

³ https://www.sgdsn.gouv.fr/files/files/Publications/20260122_NP_TLP-CLEAR_SGDSN_VIGINUM_IMS_0.pdf

complex influence campaigns, and especially importantly, to accelerate these and scale them. Concurrently though, there are also important innovations for informational-influence efforts deriving from the conflicts in Ukraine and Iran.

A 'signal event' can be defined as an incident that alters peoples' perceptions of security, impacting how they think, feel and behave, both individually and collectively (Innes, 2014). The launch of Russia's 'special military operation' in Ukraine in February 2022 has functioned as a signal event for the global order, clearly shifting how different nation states relate to and interact with each other. This has included the Kremlin accelerating, proliferating and intensifying its use of foreign information manipulation and interference measures, in both direct and indirect support of its more 'kinetic' military actions.

For understandable reasons, the most acute and pronounced changes induced by the SVO have been observed across Ukrainian politics, military and society. Although harmed and deeply scarred by the conflict, across a number of sectors and domains, Ukraine has demonstrated the power and value of developing an agile, focused and recursively innovating system-level response to Russia's aggression. This has enabled them to successfully and asymmetrically counter an adversary that is, across most metrics of military power, quantitatively far superior and more numerous. This report suggests there are important lessons from what can be observed happening in Ukraine, for how we think about and seek to counter disinforming, distorting and deceptive episodes of information manipulation and interference.

In adopting this outlook, it is important to clarify that we are not suggesting that FIMI is a new problem induced by the war in Ukraine. Indeed, we have been living with an awareness of how various modes of digitally enabled covert actions can be deployed as part of attempts to influence political decision-making and public understanding, for over the past decade. However, accompanying the general recent rise in concern about the prevalence and distribution of FIMI threats, there has been increasing interest in better understanding the intended and unintended consequences of interventions designed to deliver control over the potential harms induced.

To date, interest in the design and delivery of FIMI countermeasures has tended to lag the attention and focus directed towards the behaviours and actors they are targeting. As a corollary of which, there is only limited robust evidence of the efficacy and efficiency of these different control measures. For example, in their now slightly dated, structured and comprehensive review of the evidence base for different FIMI countermeasures, Courchesne et al. (2021) found a sparsity of reliable data across most such interventions, except for fact-checking (Courchesne et al., 2021). Similarly, but more recently, Bateman and Jackson (2024) reviewed 10 key countermeasures, pessimistically concluding:

"There is no silver bullet or 'best' policy option. None of the interventions considered were simultaneously well-studied, very effective, and easy to scale."

This general deficit of consistent, reliable and valid research evidence notwithstanding, there is a clear momentum for increasing interest and attention being paid to what different countermeasures can (and cannot) actually do. This can be partially traced by the evolution in approach evident across the European External Action Services' FIMI Threat Reports (of Feb 2023, Jan 2024, March 2025 and March 2026). The fourth report in the series introduces the **FIMI Deterrence Playbook**, with a shift from diagnosis to impact: "identifying critical nodes across infrastructures, intermediaries and supply chains," with sanctions, law enforcement, digital regulation and resilience-building "strategically mobilised to raise costs". This coheres with the key elements of EEAS' **Response Framework** differentiating between: Ignore / Contain / Minimize / Redirect actions, which is probably the most-cited operational schematic (EEAS, 2024).

In addition to which, there have been several other recent initiatives, including:

European Democracy Shield (November 2025), which also establishes the **European Centre for Democratic Resilience** and an **EU Blueprint on FIMI/disinformation** for "standardising tools and guidance for anticipation, detection and response." It supplements existing legal frameworks including DSA risk-assessment obligations and AI Act deepfake-labelling (Ehlen et al., 2025).

Collectively, these developments and initiatives reflect increasing recognition of the importance of paying just as much attention to countermeasures as to the FIMI actions they are targeting. Not least for matters of efficacy and establishing that counter-measure interventions are functioning as intended and are not causing any unintended consequences. But also recognising efficiency and prioritisation considerations, given how demand for control interventions routinely outstrips the capacity and capability to supply these. Making it vital to ensure that the right counter-measure interventions are directed towards the most pressing problems, and those where they can have some impact and make a difference in harm management terms. There are however several significant challenges that need to be overcome, if such a vision is to be realised.

1.2 Not One Problem, But Several

As alluded to above in the brief rehearsal of the available research evidence on the organisation and conduct of different state information threats and disinformation campaigns, in terms of how these actions work and the effects they have, they are multiple variants with family resemblances to each other, but never exactly the same. This level of detail and precision matters for the formulation of countermeasures on the grounds that different operations may have very different intervention opportunities. What can be done to disrupt a persistent organised operation such as Doppelganger, may be very different to what is required for a more bespoke hack and leak effort intending to get a very specific set of documents into the possession of a particular target.

Reinforcing these pressures, is the fact that there are important shifts and evolutions in the attack vectors being engaged by state actors and their proxies. Some of these reflect shifts in the affordances of new communications technologies,

This becomes a particular issue in the production of frameworks intended to structure and guide analytic work. There is an inherent tension between providing processes and systems that are generalisable and so relevant across multiple different individual cases and incidents, but also sufficiently specific that they can meaningfully impact the situated nature of individual episodes or campaigns.

1.3 When is a Framework Not a Framework

One key consequence of this tension has been the proliferation of counter-FIMI frameworks, authored by different organisations from very different perspectives. However, as discussed more fully in the main body of this report, defined with a reasonable degree of conceptual rigour, although described in these terms, many of these offerings are not in fact frameworks. In effect, the rapid and somewhat unregulated and uncontrolled way in which the counter-FIMI space has developed, with large numbers of competing and overlapping ‘frameworks’, now probably obscures and obfuscates more than it clarifies. Practitioners are confronted with challenging questions about exactly which framework they should be using for different situations, and the fact that separate members of the defender community may be viewing the same case, but through alternative and potentially contrasting frameworks.

This overarching assessment of the current situation of the FIMI defender community is a key influence shaping the analysis reported herein and the suggestions made for reform. That this is the state we are in, is an artefact of the incentives that encourage the repeated production of new frameworks and iterations of existing ones. For producing a framework (or indeed a ‘toolkit’), is a way to ‘sell’ a repeatable sequence of data collection, analysis and intervention actions and methods, as a way of demonstrating impact and relevance. However, as developed in the next section, the FIMI defender community is defined by a complex and de-centred architecture, with different actors and organisations having very different viewpoints on a case, and aspects that they can introduce via the control measures they can deliver.

1.4 De-Centered Responses Require Interoperability

Arguably one of the defining complexities when it comes to responding to disinformation and information manipulation and interference threats, results from the wide variety of control actors involved. Depending on the particularities of an event or operation all of the following organisations might be interested in and have specific responsibilities and capabilities for trying to establish control: EU-level institutions and

entities; national governments' intelligence services; specifically targeted national policy departments (health / climate / election integrity); social media platforms' 'trust and integrity' teams; civil society OSINT researchers; fact-checking organisations; investigative journalists etc. Each of these has different intervention options at their disposal. For instance: investigative journalists are especially well-positioned to expose a covert and clandestine operation, hoping to induce disruption to it via shifting public and political awareness; alternatively, social media platforms can deploy algorithmic interventions to decrease the public visibility and circulation of problematic narratives and content; contrastingly, state agencies have the capacity to leverage legal sanctions, and to harness covert intelligence assets to help attribute any state backed efforts and punish these.

The fundamentally de-centred nature of this response architecture is not conducive to the production of robust frameworks for designing and delivering interventions to counter-FIMI. There is a tendency for authors of frameworks to start from what they know and can do in terms of how they are situated vis a vis the problem privileging others similarly situated to them, and/or to make relatively uninformed assumptions and recommendations about what the other differently positioned nodes in the defender network can, or should, do. This is not helped by the fact that there are no publicly available, robust, independent evaluations of the impacts of state-led countermeasures. Relatedly, the amount of data and transparency offered by social media platforms about what they are doing and how, is in retreat recently. Nor do any studies provide reasonable estimates of the combined impacts of simultaneously implemented interventions upon the prevalence and distribution of targeted information operations. And whilst there is some largely lab-based experimental evidence for the potential efficacy of pre-bunking, fact-checking and inoculation measures to build community resilience, it is less clear how this scales and transplants into an increasingly 'noisy' and chaotic information ecosystem.

1.5 Report Overview

Framed by these overlapping and interacting challenges and considerations, this report seeks to develop a more system-level view of the current state-of-the-art of FIMI defences and countermeasures and the evidential supports for these.

- To deliver on this purpose, in the next chapter we present an extended case study of 'cyber-enabled information operations' as a way of illuminating some of the challenges for countermeasures design and delivery, where a species of information manipulation threat engages multiple attack exploits.
- This is followed by a chapter describing the data and methods used that have informed our position.

- In turn, this sets up a more detailed and immersive discussion of the large numbers of counter-FIMI frameworks that are now available, and some of their respective strengths and weaknesses, both in terms of tactical specifics, but also strategic generalities.
- This provides evidence for a central thrust of our argument that to improve the capacity and capability of the EU and its member states to defend against the threats and risks posed by FIMI, a moratorium should be imposed on the production of new frameworks.
- Instead, the focus should be upon establishing which ones, of the very many available, are actually helpful in shaping policy and practice to impact upon the threats being encountered.
- The concluding chapter reprises some of the key issues and themes to start to map out an agenda for future development work that will further enhance the efficacy and efficiency of Europe's defences against foreign state backed information manipulation and interference.

2 The Current Countermeasure Landscape: Challenges and Opportunities

2.1 Information influence operations are evolving and adapting

One of the key challenges for those engaged in countering misinformation, disinformation, Foreign Information Manipulation and Interference (FIMI), or information influence operations (IIO), is how these malign activities constantly change, evolve and adapt. Any such adaptations can be endogenous or exogenous. The former referring to internally driven shifts in organisation and conduct, as operators learn better tradecraft, and ‘what works’ in terms of the efficacy and/or efficiency of their activities. Contrastingly, exogenous adaptations result from: operators’ attempts to manoeuvre around or circumvent any control measures; or as innovative reactions to new events, technologies or situations.

This chapter explores how and why these processes of continual adaptation and evolution in the organisation and conduct of state-backed information-influence operations pose particular challenges for those seeking to design and deliver frameworks to structure and organise countermeasures to state information threats. The chapter uses several case-studies to illuminate the recurring dynamics and interactions between the information manipulation threat landscape and the kinds of responses that can be imagined and operationalised by different defender communities of practice. The chapter concludes with a summary of the development of responses over time, what options are available now, and the implications of this for how we conceptualise the role and utility of countermeasures frameworks.

To afford the discussion a sense of focus and practical utility, the case studies introduced, all pivot around a growing convergence between cyber-attacks and influence operations. The overlap and interactions between the two different modes of digital harm, allows us to highlight some of the complex questions current countermeasures efforts are seeking answers for. The first case-study explores Iran’s work, as an increasingly prominent state actor combining cyber and influence operations in the lead up to the war 2026. The second use-case highlights North Korean IT worker tactics aiming at circumventing sanctions, where widespread use of AI further blurs the line between cyber and influence tactics. The third case centres Azerbaijan, showcasing how many regional state actors are increasingly blending information-influence and cyber tactics as they negotiate their relationships with their neighbours, in addition to the far more common focus on Russia and China.

Traditionally, cyber and influence operations have been treated as separate fields. Relatedly, responses to these threats have been organised differently, with distinct organisational structures, mandates and ways of working. This separation, however, is vague and doesn't always match the presenting threat. One of the most prominent examples illuminating how in practice operations often combine both cyber- and influencing methodologies, is Russia's interference in the US presidential election 2016. This mixed both hacking computer systems and then releasing information into the public domain by the Russian military intelligence service GRU and disinformation campaigns led by the Internet Research Agency, under the direction of oligarch Yevgeni Prigozhin (U.S. Department of Justice 2018a, U.S. Department of Justice 2018b)

As a general trajectory of development, there has been growing interest in researching both cyber-attacks and influence operations, although most studies have attended more to one modality over the other. Very recently, however, there is increasing recognition of the importance of understanding dynamics and interactions between them, especially when both are enlisted as part of contemporary digital statecraft efforts. It is a blend typically referred to as 'cyber-enabled influence operations' (CEIO) (Lin 2019, Etudo et al 2023, Vičić and Gartzke 2024).

Cyber-enabled influence operations can be defined as "state-based actions, in the context of an ongoing cyber action, seeking to communicate a message or manipulate the reception of digital information for malicious purposes among a targeted population" (Foote et al. 2020). The key division of labour here is that: while cyber operations focus on access, system compromise, and secrecy to steal intellectual property, degrade systems or conduct sabotage: influence operations manipulate information to deceive and distract the target audience (Whyte and Etudo 2025). It has been suggested that Russia dominates the use of cyber-enabled information operations globally (Foote et al. 2020). However, there is also a diverse and growing landscape of cases of cyber-enabled influence operations involving others (Whyte and Etudo 2025). Overall though, there is a consensus that far less academic interest has been paid to efforts countering these blended forms of digital deviance, than has been directed towards them individually.

2.2 Reporting about cyber and IIO convergence is rising

Within the countering information influence operations field, the concept of FIMI introduced by the EU in the early 2020s (Council of the European Union 2022), has had the second-order effect of increasing use of terminology and analytical processes borrowed from cyber-threat intelligence (CTI) (Hénin 2023). At the same time, big cyber security and tech companies have moved to increasingly include countering influence operations and disinformation into their portfolios.

Ahead of the US presidential elections 2024, Microsoft warned (Microsoft Federal 2024) about increasing sophistication of threat actors' tactics and convergence of cyber and influence operations, stating:

“We’re now seeing tactics that have traditionally been reserved for cyberattacks—such as email campaigns and ransomware—being increasingly leveraged in cyber-enabled influence operations after cyberattacks occur - - - Russian intelligence services have consistently attempted hacks to power its influence activity, and Iran has begun prolifically using hack-and-leak operations against a range of Tehran’s opponents. China has not yet employed any target hacks to power its election influence operations against the US, but that could change.”

Before the election, the US State Department announced that RT (Russia Today) had been equipped by the Russian government with cyber capabilities in Spring 2023 (U.S. Department of State 2024). Combining cyber operational capabilities and ties to Russian intelligence, it appears the resulting entity and approach has been engaged in the delivery of influence and intelligence operations globally.

To find out if these arguments about an increased convergence between cyber and influence operations strategies and tactics can be evidenced, a dataset was constructed based on reports potentially relevant to these issues. Data collection for this was conducted in October 2025 covering the period 01.01.2016-30.09.2025. To reach a balanced sample, two companies focusing on cyber-security (Microsoft and CrowdStrike), two tech platforms (Meta and Google) and two NGOs working on influence operations (DFRLab and EU DisinfoLab) were included as sources for relevant reporting documents. (See Methodology for more details about the dataset).

Next, the reports mentioning both ‘cyber’ and ‘influence’ (94 articles) were selected for further analysis. They were coded by year, state actors mentioned, and the context. Thus, coding paid especial attention to reporting about actors and operations where both influence and cyber elements coexist. However, distinguishing the two is challenging, due to differing definitions the analysing organisations use in their reporting and different levels of confidence based on which activity is attributed to a specific actor.

Reports combining 'influence' and 'cyber' grew in 2023–2024

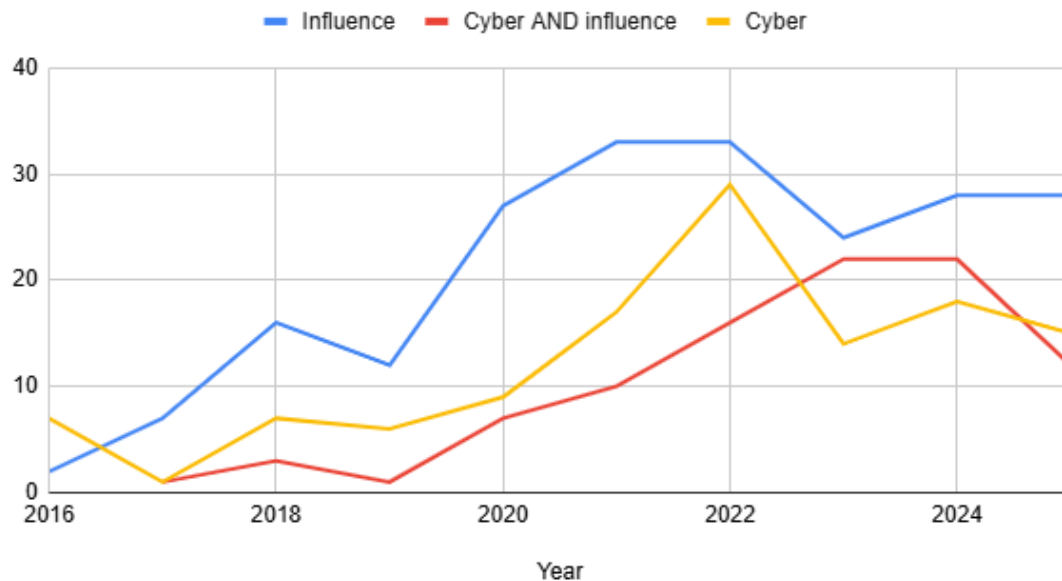


Figure 1: Year 2025 is covered only until the end of September.

Plotting the coded documents against each other based upon year of publication, shows how both cyber-focused and influence-focused organisations have increasingly reported about influence operations and cyber-attacks, expanding their capability and capacity in both domains. Developing this picture, in Figure 2 below the reports are categorised by threat actor over time.

Mentions of Iran grew in cyber and influence reporting

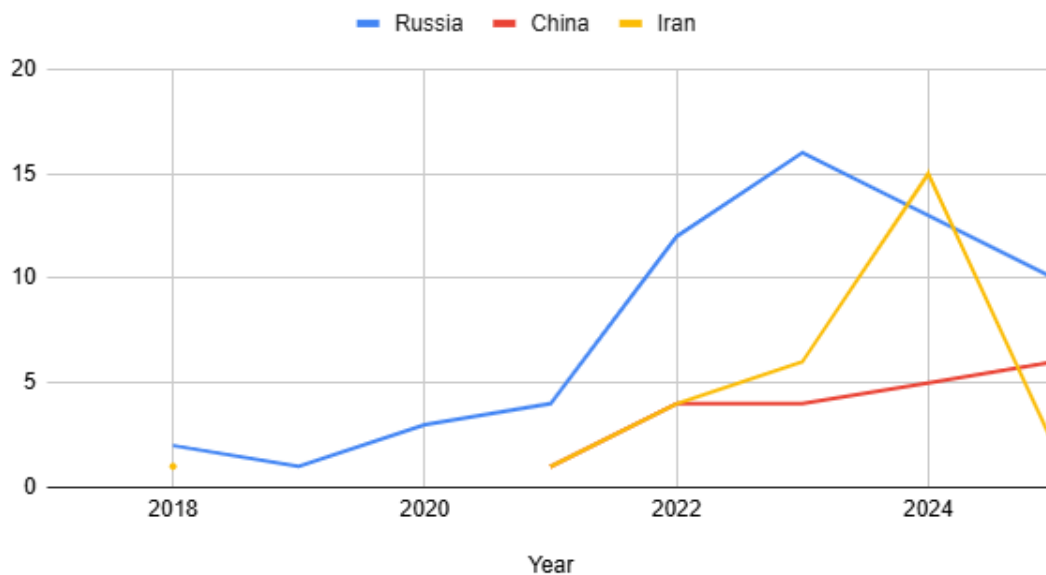


Figure 2: Most mentioned state actors in cyber and influence reporting. Year 2025 is covered only until the end of September.

This approach to visualising the data clearly conveys how Russia dominates in terms of perceived threat and activity. That said, there was rapid growth in interest in Iran's participation starting in 2023, through into 2024, the year of many elections worldwide, including the US presidential election. North Korea, Belarus and Azerbaijan were also subject of several reports. Based on their importance in the data, we chose Iran, North Korea and Azerbaijan as a focus of the following case-studies. For this purpose, the initial data collected was complemented with public reporting on these actors' operations, and Facebook IO archive aimed at researchers was utilised. The timespan is until Spring 2026.

2.3 Case-study 1: Countermeasures against Iran's cyber-enabled influence operations peaked in 2024

Iran's tactics blending cyber and information warfare have become especially visible after the US and Israel attacked Iran in February 2026. Instead of focusing on targeted, but limited cyber-enabled influence campaigns, Iran started deploying the full spectrum of means it has available: cyber-attacks, hacktivist groups, disinformation and overt propaganda. Potentially this reflects just how much threat the Iranian regime perceived it was under, recognising it could never directly match or defend against the military 'hard power' available to either the US or Israel. Consequently, they calibrated aspects of their response very differently.

For example, the hacker group Handala claimed it breached the medical technology company Stryker, reportedly erasing data and disabling tens of thousands of computers, alongside a series of similar cyber-actions (Greenberg 2026). But perhaps the most striking innovation was the instigation of a series of pro-Iran, AI generated Lego-themed videos mocking Trump and the US. These have gained lots attention on social media and in legacy media outlets as well, attracting massive audiences for the propaganda messaging they contain (Shea and Kalus 2026).



Figure 3: Screenshot of Facebook account “Explosive Iran”.

The strong performance of these videos in terms of attracting audience attention across a range of different platforms reflects that their content is topical, humorously framed and entertaining. In this respect they may be an early signal of how, one consequence of the spread of AI technologies, is likely to be an increase in the quality of propaganda. In the Iranian case, it is pertinent that they are produced by a PR-type agency called ‘Explosive Media’ on behalf of the Iranian government. This matters because they are more creative and compelling than the more orthodox types of propaganda typically issued by government agencies.

One aspect of the response to Iran’s decision to develop an asymmetric strategy centering cyber and digital influence technologies, is that Iran’s cyber capabilities have been among the key targets in the war for the US and Israel. Israel announced missile strikes against the Islamic Revolutionary Guard Corps (IRGC) headquarters and Iran’s Cyber Warfare headquarters early March 2026 (Israel Defense Forces 2026). Two Iranians allegedly running cyber operations against the West have been reportedly killed. Mohammad Mehdi Farhadi Ramin was previously charged in 2020 with hacking American defence companies, and Seyed Yahya Hosseiny Panjaki, a deputy minister of Intelligence for Israeli affairs at MOIS (Ministry of Intelligence and Security) was in charge of a unit that controlled hacking groups (Brewster 2026) These attacks might have hindered Iran’s cyber capabilities, but not destroyed them (Shah 2026).

Reportedly though and potentially off-setting some of these impacts, Iran-linked networks of social media accounts reportedly posting in Spanish and English, shifted to pro-Iranian propaganda content surrounding the war, late in February 2026 (Murray and Linvill 2026). Iranian-aligned hacktivist groups have also launched DDoS attacks and reconnaissance operations targeting US and Israeli assets, often with low impact. Other hacktivist groups have conducted website defacements and data-wiping attacks, targeting US and Israeli military logistics providers, aiming to destroy data and disrupt operations. Pro-Western groups have engaged in similar activity (Resecurity 2026). In parallel, Iran is suspected of being behind several sabotage attacks and terrorist attacks in Europe targeting Jewish synagogues and other places (Pancevski and Colchester 2026).

In the years leading up to the US and Israel war on Iran, mentions of Iran in cyber and influence operation reporting grew especially following the US 2020 election, peaking in 2024 during the next presidential election round and Israel-Gaza war. Microsoft Threat Intelligence assessed (2023a) that most of Iran's cyber-enabled influence operations in 2022 were being run by Emennet Pasargad, members of which have been sanctioned at least since their attempts to interfere with the 2020 US elections (U.S. Department of Justice 2021). It also assessed that these operations' TTPs (tactics, techniques, and procedures) had expanded and improved across that year.

Cyber company Emennet Pasargad, also known as Shahid Shushtari (and several other names), is directed by the Islamic Revolutionary Guard Corps (IRGC) and has been active since 2018 (Rewards for Justice (no date)). The IRGC is an example of an actor that has moved from 'pure' cyber operations (sanctioned by the US in 2019), into adding influence techniques into its capabilities (at least since the 2020 US election) (U.S. Department of the Treasury 2024b). The FBI notes that in late 2018, Emennet Pasargad masqueraded as the "Yemen Cyber Army" and created messaging critical of Saudi Arabia, which is the earliest identified mention of combined cyber and influence activity by the group (Federal Bureau of Investigation 2022).

In March 2026, two weeks after the US and Israel launched their war against Iran, the EU sanctioned Emennet Pasargad for: compromising Swedish SMS services; assaults on the French satirical magazine Charlie Hebdo's subscriber database; as well as compromising advertising billboards during the Paris Olympics and "displaying disinformation campaigns" on them. The EU concluded that "Emennet Pasargad is therefore responsible for cyber-attacks with a significant effect which constitute an external threat to Member States", emphasizing the cyber activity as a basis for sanctions (Council of the European Union 2026). Earlier, Swedish authorities had investigated the data breach in Sweden, where hackers took over a Swedish text messaging service and sent 15,000 messages calling for revenge against Quran burners. According to prosecutors, the goal was to create division in Swedish society. The criminal investigation was later closed because "the requirements for prosecution abroad or extradition to Sweden were not met." (Swedish Security Service 2024).

Other examples of the effects that Emennet Pasargad might have tried to achieve with their cyber-enabled influence operations (Federal Bureau of Investigation, U.S. Department of the Treasury and Israel National Cyber Directorate 2024 assessment) include:

- Hack-and-leak operations against organizations primarily in Israel. These were intended to undermine public confidence in the security of the victim's network and data and embarrass victim companies and targeted countries
- Attempts to contact family members of Israeli hostages, likely in an effort to cause additional psychological effects and inflict further trauma on family members.
- In 2023, according to Microsoft, Emennet Pasargad targeted French satirical magazine Charlie Hebdo. A cyber persona named "Holy Souls" posted for sale data allegedly containing personal information of more than 200,000 Charlie Hebdo's customers after the magazine published cartoons ridiculing Ali Khamenei, Iranian Supreme Leader (Microsoft Threat Intelligence 2023b).
- Compromising a French commercial dynamic display provider, attempting to display photo montages denouncing the participation of Israeli athletes in the 2024 Olympic and Paralympic Games. This cyberattack was coupled with disinformation, including publication of a fake news article onto a French collaborative media website and the spread of threat messages to several Israeli athletes and their entourage under the banner of a fake French far-right group 'Regiment GUD', impersonating the real French far-right group 'GUD'.

Based upon the dataset constructed, influencing elections is one of the most common aims of cyber-enabled influence operations. Microsoft investigated Emennet Pasargad's activity ahead of the US elections in 2024, noting it conducted reconnaissance and probed election-linked websites in US swing states. It also performed reconnaissance of big American media outlets (Microsoft Threat Analysis Center 2024a). However, in the end, Emennet Pasargad was not specified in the subsequent indictments concerning election interference. Although several Iran- and IRGC-linked threat actors were allegedly involved in cyber and influence activity ahead of the election, targeting both the Biden and Trump campaigns (Google Threat Analysis Group 2024, Microsoft Threat Analysis Center 2024b). Iran was also attributed to a hack of Trump's adviser's email account. The documents, allegedly accessed via the hack, were later shared with media outlets (The Washington Post 2024). The media chose not to amplify the content of the documents directly, instead focusing on covering the hack itself and its motivations (Sommer and Izadi 2024, Isenstadt 2024).

Meta reporting (Franklin and Torrey2024) detailed how they had taken down dozens of Facebook accounts and two Instagram accounts, originating in Iran, targeting Israel,

and running fictitious brands, which the FBI, US Department of Treasury, and Israel National Cyber Directorate attributed to Emennet Pasargad. The brands posed as anti-Western activists or hacktivists, and had presence on Meta's apps, YouTube, Telegram, X, and their own websites. Content was published in English, Hebrew and French. The operation used fake accounts to drive social media users to sites where they allegedly hosted hacked material from the Israeli government. Meta does not assess the validity of the claims of the hacks the groups promoted. One of which called to ban Israeli athletes from participating in the Paris Olympics.

Relatedly, YouTube announced in 2024 they terminated four channels sharing content in English and Hebrew criticizing the Israeli government in the Israel-Gaza war (Huntley 2024). The campaign shared content about alleged cyber-attacks against Israeli organisations. Youtube did not attribute the campaign to a specific threat actor or campaign.

2.3.1 Detection, disruption and legal actions

Several tech platforms, cyber firms, social media platforms and governments reacted to Iranian cyber-enabled influence operations, especially in 2024, driven by the interest in protecting the integrity of the US elections from foreign malign influence attempts (ODNI 2024).

Describing the threat and attribution

Depending on the organization participating in the countering efforts, the description of the threat varies. Microsoft mostly uses the label “cyber-enabled influence operation”, while Meta reports about its takedowns under coordinated inauthentic behaviour (CIB) policy, and Google under coordinated influence operations. Governmental actors use a variety of terms and definitions in their reporting, like cyber, influence, malign activity, malicious cyber actors, interference, and election influence.

In 2024, external malign activities targeting the US election were initially attributed to Iran (ODNI, FBI and CISA 2024), and later by September, the attribution specified and was linked to IRGC, and parts of it to IRGC-connected APT-42 (U.S. Department of the Treasury 2024). Earlier, in 2022-2023, Microsoft linked a group of cyber personas to Cotton Sandstorm (Emennet Pasargad) based on “extensive overlaps in influence TTPs”, with higher confidence in “some of the groups” based on information published by the US government, victim overlaps, forensic links, and Meta's findings (Microsoft Threat Intelligence 2023).

The take-home point for our interests here being that there is not a common established framework of definitions, descriptors and labels shared by all actors engaged in the design and delivery of countermeasures for cyber-enabled influence operations. Indeed, this is also the case for the broader fields as well. Moreover, the granularity and detail of an attribution is not stable either, and often

develops over time, becoming more specific. This matters in terms of thinking about when one should start to think about what possible countermeasures could be configured in response to a particular threat, vis a vis progress with the investigation and assessment of it.

Detection and evidence

Looking across the available reports, it appears that in their reporting, tech companies typically only referred to their own investigative efforts and FBI described investigating Iranian actors from the government's side. Microsoft linked certain influence activities (based on 'branding' or cyber persona) via cyber threat activity (like hacking). For example, "Holy Souls" tried to sell data reportedly containing personal information of 230,000 Charlie Hebdo customers on an internet 'leaks' site (Watts 2023).

Contrastingly, social media platforms reported about detection under their policies for 'inauthentic coordinated behaviour', focusing on the influence-engineering activities, rather than any 'cyber' components. Facebook deplatformed activity targeting Israel, describing how the actors used fake accounts to post, comment, reshare and drive users to websites claimed to host hacked Israeli government documents. Facebook did not take a position on the validity of those claims (Franklin and Torrey 2024). Similarly, YouTube detected a campaign sharing content critical of the Israeli government and its actions in Gaza, and content claiming cyber-attacks targeting Israeli organizations (Huntley 2024).

Separately, Google reported about APT42's cyber activity, including phishing attempts, and later targeting personal email accounts affiliated with staffers working on the Biden and Trump presidential campaigns. Google said it blocked APT42 attempts to log in to personal email accounts of these individuals (Google Threat Analysis Group 2024).

Disruption

In attempting to disrupt the identified operations Facebook and YouTube conducted takedowns of accounts and channels. Alternatively, Microsoft and Google focused on public reporting, publicizing indicators of compromise, blocking threat activity on their platforms, advising users, and informing the targets of the attacks and law enforcement. Government agencies published advisories, details of TTPs associated with the campaigns and set up reporting channels for targets.

Legal action

These efforts culminated in an indictment of three IRGC members in September 2024. Three months after the alleged Iranian attempt to hack Trump's campaign and leak the material to the Biden campaign and media, the alleged perpetrators were already identified and charged. The US Department of Justice thanked private sector partners Google, Microsoft, Yahoo, and Meta for assisting with the case (U.S. Department of Justice 2024).

Some days before the election, US and Israeli authorities published a joint advisory regarding cyber group Emennet Pasargad, listing the evolved TTPs the group was using (Federal Bureau of Investigation, U.S. Department of the Treasury and Israel National Cyber Directorate 2024). Months earlier, the US Treasury sanctioned two companies and four individuals involved in malicious cyber activity on behalf of the Iranian Islamic Revolutionary Guard Corps Cyber Electronic Command (IRGC-CEC) (U.S. Department of the Treasury 2024a).

2.3.2 Context

At first sight, the countering of the IRGC's and Emennet Pasargad's cyber-enabled influence operations looks like a best-case scenario of joint efforts, where multiple different actors with specialist skills and interests, come together and cooperate to achieve a common goal in disrupting malign activity. This can be an effective way of working. For example, when attributing certain activity to a specific actor, a decentred approach such as that described above can be described as "participative attribution" (Innes and Ahonen 2025). But a necessary requirement for this to function efficiently and effectively in practice, is a high level of trust between industry partners and governments, as well as adequate resources. That said, it is equally true that not all social media platforms or other actors are as invested or engaged in any such efforts.

Therefore, it is worthwhile looking at the conditions under which this type of efficient counter-measure action can happen. The IRGC has been designated a foreign terrorist organization by the US since 2019. Apart from the election targeting, the malicious cyber and influence activity was connected by the US to an ongoing effort by Iran to avenge the death of the former commander of the IRGC – Qods Force (IRGC-QF) Qasem Soleimani (U.S. Attorney's Office 2024b). He was killed in a drone strike in 2020, ordered by President Trump, near Baghdad International Airport. Iran responded to the killing, firing ballistic missiles at US forces at Iraqi military bases (BBC News 2020).

The US election cycle brings wide public and media attention into the picture, combined with political will to make strong public statements. However, the sanctions, charges and companies limited takedown reports, viewed either individually or collectively, tell us little about the actual development of different actors' capabilities in conducting successful interference operations. In 2022-2023, Microsoft assessed that while Iranian groups increased the pace of cyber-enabled influence operations, the cyberattacks remained "low sophistication" and "low-effort". It also stated that the apparent rise they see in these types of operations may actually be an artefact of Microsoft's increased detection capabilities. At that point in time, they concluded Tehran was likely seeking to retaliate against "what it perceived were efforts by foreign actors to foment unrest in Iran, including highly sophisticated cyberattacks in 2021 and 2022 by a group called Predatory Sparrow against Iran's steel factories, state broadcasts, gas stations, and trains (Reuters 2021, Microsoft 2023).

Following President Trump's election in 2024, the US state has dismantled many of the governmental bodies and structures that helped organise and co-ordinate the countering of foreign influence and interference in elections. This included shutting down the FBI's Foreign Influence Task Force, suspending Election Protection Activities at the Department of Homeland Security, reducing an expert team at the Cybersecurity and Infrastructure Security Agency (CISA), dismantling Global Engagement Center (GEC), firing General Tim Haugh, head of the National Security Agency (NSA) and the Pentagon's US Cyber Command, and freezing enforcement of the Foreign Agents Registration Act (FARA) (Siman-Tov 2025). The long-term consequences of these changes, both individually and collectively, remain to be seen.

This extended discussion of Iran, and how and why it has been developing digital attack and influence methodologies, that combine influence and cyber tactics, has been designed as a principal case study used to map out and illuminate the main components and features of these 'fuzzy' or 'blended' information threats. In the remainder of this chapter, we will develop and 'round out' this understanding, based upon analysis of a number of smaller additional examples.

2.4 Case-study 2: North Korea exploits AI in large-scale fraud IT worker scheme



North Korea sent me abroad to be a secret IT worker. My wages funded the regime

2 August 2025

Share Save

Beth Godwin and Julie Yoonnyung Lee
BBC Trending & BBC News



Figure 4: Screenshot from BBC article.

AI is further blurring the line between cyber and influence operations. Especially visible in this regard are the methods and approaches used by North Korea. Adversarial operations run by the Democratic People's Republic of Korea (DPRK), aiming to funnel money to its weapons programs have been active since at least 2018 (CrowdStrike 2026). Sometimes referred to as 'Famous Chollima', these operations are generally considered as a cyber threat actor working mainly to evade sanctions and produce economic gain to North Korea. However, its widespread integration of AI into its social engineering operations increasingly blends and integrates cyber and influence tactics. Fraudulent IT workers sent abroad by North Korea engage in financial fraud and data theft, but have also been linked to cyber espionage and employing malware (Insikt Group 2025).

Generative AI is used by DPRK's IT workers across all main stages of their operations. They draft CVs with the help of AI to mount job applications, improve English-language content and communication, create inauthentic identities and accounts with manipulated photos, cover their true identities during video interviews with deepfakes, and AI assists in the work tasks helping with coding (CrowdStrike 2025). Google and

OpenAI have reported use of Gemini and ChatGPT by these threat actors (Google Threat Intelligence Group 2025, OpenAI 2025). For instance, OpenAI (2025) observed the following LLM TTPs:

- Automating systematic fabrication of detailed résumés aligned to various tech job descriptions, personas, and industry norms;
- Automated generation of consistent work histories, educational backgrounds, and references via looping scripts;
- Utilising the model to answer employment-related, likely application questions, coding assignments, and real-time interview questions, based on particular uploaded resumes;
- Seeking guidance for remotely configuring corporate-issued laptops to appear as though domestically located, including advice on geolocation masking and endpoint security evasion methods;
- And LLM assisted coding of tools to move the mouse automatically, or keep a computer awake remotely, possibly to assist in remote working infrastructure set ups.

AI companies note that when threat actors seek increased efficiency of their operations through integrating AI fully in their processes, this development also expands their exposure (Ibid.).

Some indicators of these operations' potential impact exist. The U.S. Justice Department (2024) assesses that the DPRK has sent thousands of IT workers globally, and collectively they can generate hundreds of millions of dollars each year. In 2024, the DPRK likely earned approximately \$350-800 million from its IT workers globally according to an assessment by the United Nation's multilateral sanctions monitoring team (MSMT 2025).

2.4.1 How to counter a known threat actor that uses advanced AI

North Korea's fake IT worker operation is an exceptionally thoroughly examined operation in public reporting, partly because it aims to evade widely recognized international sanctions. In addition, its activities in many jurisdictions are illegal (MSMT 2025). In 2025, CrowdStrike, Google (Gemini) and Microsoft all emphasised the use of AI by DPRK IT workers, in part as a way of promoting their own AI-based threat hunting and protection services.

Describing the threat and attribution

In their report on adversarial misuse of AI, Google (Gemini) refers specifically to North Korean APT actors and differentiates between cyber (advanced persistent threat, APT) and information operation (IO) focused threats. The North Korean IT worker activity falls under the cyber part of this categorization (Google Threat Intelligence Group 2025). Mandiant (Google) notes that what they track as UNC5267, is not a traditional centralized threat group, rather it consists of individuals sent by North Korea to work primarily in Russia and China (Starks et al. 2024). OpenAI says they are not able to precisely determine the locations or nationalities of these threat actors, albeit their "behaviours were consistent with activity publicly attributed to IT worker schemes connected to North Korea (DPRK)" (Nimmo et al. 2025). Microsoft constructs North Korean IT workers as posing a triple threat: (1) as they aim to make money for the regime; (2) they may use their access to get sensitive intellectual property or trade secrets; (3) or steal sensitive data from companies and sometimes demand a ransom payment (Microsoft Threat Intelligence 2024a).

GitLab differentiates between 'Contagious Interview' – that they define as a campaign distributing malware - and IT worker activity; attributing both to North Korean nation-state actors (Smith 2026). The UN multilateral sanctions monitoring team (2025) talks about "The DPRK's Violation and Evasion of UN Sanctions through Cyber and Information Technology Worker Activities".

From the overarching perspective of evaluating the efficacy of countermeasures frameworks, there are two key points in evidence here. Even in a single sector, there is not an established set of definitions or constructs used consistently to generate a shared understanding of how to analyse this kind of key emerging threat. Moreover, more generally, it is not at all clear that the current frameworks that are in place are appropriately configured to respond to the changed speed, scope and scale of the challenges posed by AI enabled disinformation and manipulation.

Detection and evidence

Evidence presented in public reporting is often based on tech companies' own investigations, with references to other public sources such as leaked records (Bîzgă 2024). Tech and cyber companies variously refer to "evidence from multiple sources" (CrowdStrike), a wide variety of technical signals (Google), proactive detection and user reports (GitLab), and customizing a machine-learning solution monitoring red

flags (Microsoft). Amazon uses AI-screening combined with human verification, analysing connections to high-risk institutions and anomalies across applications and geographics.

Reporting frequently presents as evidence: accounts and personas across different platforms and portfolios created with the aid of AI; code repositories; financial records; publicizing network indicators of compromise and URLs; references to remote management tools on victim corporate laptops; and payments and moving money to North Korean IT worker-controlled accounts. GitHub and LinkedIn are often mentioned as platforms where the accounts operate. In October 2024, Microsoft found a public repository containing North Korean IT worker files. US government and the Republic of Korea have published ‘red flag’ indicators of North Korean IT worker activity to help in detecting it (Federal Bureau of Investigation 2023).

These detection efforts show that identifying digital traces of operations aiming at large-scale evasion of sanctions is possible, when this type of threat activity is prioritised. There is much less public evidence available, for example, of detecting Russian activity aiming at evading disinformation-focused sanctions (Gyimesi 2025).

Disruption

Publicly available data evidences that several of the companies reported account bans, takedowns and disruptions in 2025 targeting Famous Chollima and its ancillaries. Amazon blocked 1,800 alleged DPRK operatives from joining and detected 27% growth in DPRK-affiliated applications quarter over quarter (Schmidt 2026). Microsoft (2025) suspended 3,000 accounts. OpenAI reported it had identified and banned ChatGPT accounts participating in deceptive employment campaigns. Google said it disrupts malign activity on its platforms, including Gemini. Companies also alerted affected users and asked victim companies to contact them, promising to inform those who need to be protected. GitLab banned 131 accounts distributing malicious code in 2025. Public disclosures of North Korean threat actors’ activities and TTPs are a frequent response strategy, with yearly updates when the operation’s tactics have evolved. OpenAI published LLM mapping of TTPs using the Mitre ATT&CK Framework, developed together with Microsoft (MITRE ATT&CK (not dated); Microsoft Threat Intelligence 2024b).

In sum, collating these data suggests that the North Korean operation, constitutes a relatively large, persistent and sprawling effort. Especially important is the ‘entrepreneurial ethos’ that seems to be threaded through the actions, inasmuch as the strategy of ‘attack’ appears to involve systematically searching for opportunities and gaps that can be leveraged to affect some form of advantage and benefit.

Legal action

Since 2023, the US Justice Department has announced several indictments, charges and seizures, aiming to disrupt these illicit revenue generation efforts (U.S. United States Department of Justice 2025). One revealed scheme impacted 300 US companies (U.S:

Attorney's office 2024a). The UN sanctions monitoring team publishes detailed reports about sanctions' evasion schemes, with input from participating states and other partners, like Mandiant. In 2025 it rendered a thorough attribution:

"Nearly all the DPRK's malicious cyber activity, cybercrime, laundering, and IT work is carried out under the supervision, direction, and for the benefit of entities sanctioned by the UN for their role in the DPRK's unlawful WMD and ballistic missile programs. These entities include the Korean Workers' Party, which is subject to the assets freeze, the Reconnaissance General Bureau (KPe.031), the Ministry of National Defense (KPe.054), the Ministry of Atomic Energy and Industry (KPe.027), the Munitions Industry Department (KPe.028), Office 39 (KPe.030), and the Second Academy of Natural Sciences (KPe.018). These entities rely on cyber units and IT workers assigned to front companies in order to covertly carry out malicious activities on their behalf. All UN Member States are required under UNSCR 2375 to prohibit the opening, maintenance and operation of all joint ventures or cooperative entities, new or existing, with DPRK entities or individuals, whether or not acting for or on behalf of the government of the DPRK or the Korean Workers Party." (MSMT 2025, p. 8)

2.5 Case-study 3: Azerbaijan attempts to gain regional influence and targets women's rights activists domestically

Alongside the larger, widely tracked and monitored state-backed threat actors taking part in cyber and influence activities, there is also evidence that these blended capabilities are achievable for other operators who can combine different types of malign influencing activity. There is evidence that the kinds of strategies and tactics outlined above, are increasingly being embedded into the statecraft of middle-powers, and how they are navigating and negotiating regional and local conflicts.

According to Meta, Azerbaijan has used covert online influence operations and overt propaganda to advance its regional military and political goals, such as in Nagorno-Karabakh. For example, Facebook's Coordinated Inauthentic Behaviour (CIB) takedown from April 2021, includes around 200,000 posts from between 2015 until 2021, including messaging on the Second Nagorno-Karabakh war. This conflict resulted in an Azerbaijani victory, achieved with support from Turkey. This was followed by a further 2023 offensive by Azerbaijan, where it took over all control of the disputed territory from Armenia.

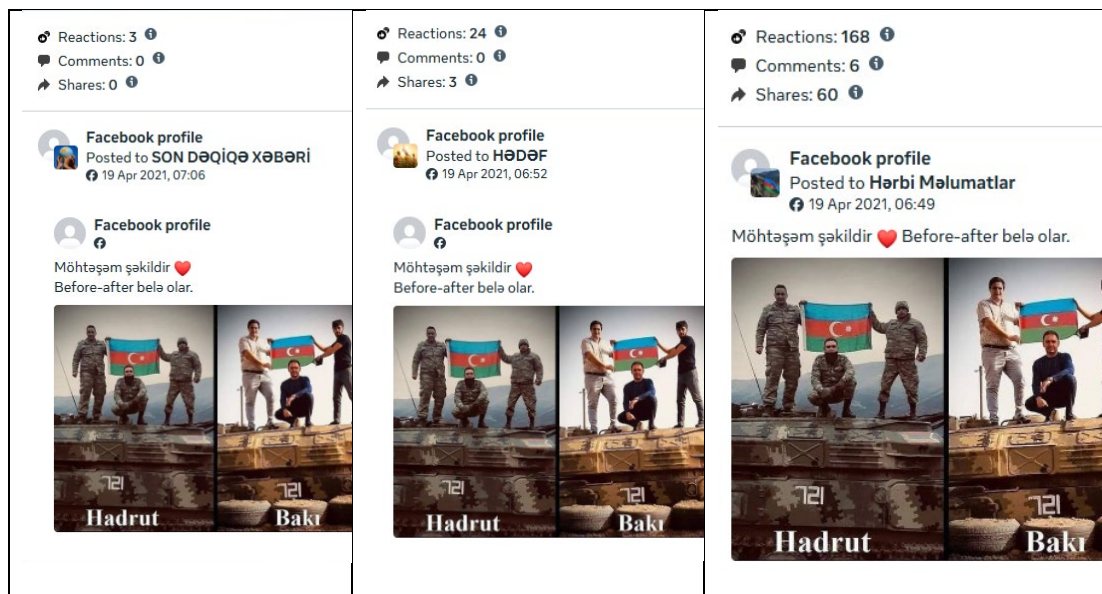


Figure 5: Facebook IO archive, April 2021.

These CIB operations were organised and conducted alongside claims of hacks and attempts to spread the allegedly stolen materials further. In their publicly released takedown report, Facebook attributed some of the activity to “individuals associated with the Defense Ministry of Azerbaijan” (Meta 2021).

Reactions: 5
Comments: 0
Shares: 0



@LAQƏ

3 Oct 2020, 07:47

18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ 18+ As a result of the cyber operation carried out by the hackers of the "Anti-Armenia Team" hacker group of Azerbaijan, the computer belonging to the Minister of High Technologies of Armenia "Hakob Arshakyan" was seized. Our testimonies show that the Minister of High Technology finds solace in surfing sites all day long in his service room to relieve the stress of war. For more information: <https://armdocs.org/article-details/12> <https://armdocs.org/article-details/14>
#antiarmeniateam #antiarmenia #azerbaijan #armenia #hacking #hack #haker #hacked #cyberarmy #cyber #deface #security #azerbaycan #cyberwar #xaker #hacker #karabaghisazerbaijan [See less](#)



Figure 6: Facebook IO archive, April 2021. Automatic translation into English from Azeri.

In addition, human rights organisations and investigative journalists in Azerbaijan have been reporting hacking attempts against critics of the government (Amnesty International 2017).⁴ Based on this reporting, Israeli cyber-surveillance company NSO group has apparently supplied the spyware tool 'Pegasus' to the Azerbaijan government, that it uses to target dissidents, journalists and regime opponents. There is no definitive proof of this, but a preponderance of evidence suggests it is likely true (Patrusic and Bloss 2021).

⁴ <https://www.amnesty.nl/actueel/azerbaijan-activists-targeted-by-government-sponsored-cyberattack>

In another takedown from 2022, Meta reported they had taken down another network of accounts originating in Azerbaijan (Nimmo et al. 2022). It said the accounts were conducting cyber espionage and influence activities (or Coordinated Inauthentic Behaviour, CIB), targeting civil society by compromising accounts and websites. In this blended operation combining cyber and influence tactics, the main targets were Azerbaijani democracy activists, journalists and people living abroad critical of the Azerbaijani government. Meta attributed the campaign to the Azeri Ministry of Internal Affairs. The content of the takedown is not included in Facebook's IO Archive tool, accessible to researchers. Facebook did not answer a request about the issue. The operation used phishing, social engineering and hacking to gain access to targeted accounts, collect personal information about the owners of the accounts and post fabricated messages on their behalf from the compromised accounts. The campaign operated cross-platform, on Facebook, LinkedIn, Twitter, YouTube, VK and OK. According to Meta, the campaign either compromised or operated themselves over 70 websites in Azerbaijan, Armenia, as well as Russian and Turkish domains. It also used malware tools and collected users' social media login information via fabricated political polls, which were published on the compromised websites. The campaign published posts on social media accounts it controlled. The content was usually in Azeri, criticizing the opposition, journalists and civil society.

Disruption efforts against the Azeri operation and its constituent parts focused on Facebook takedowns, media exposures of the campaign and interviewing the victims, and publicizing technical details of the malware deployed.

Prominent feminist activists in Azerbaijan were targeted in an attempt to smear their reputation. Women's rights activist Narmin Shahmarzadeh's Facebook account was hacked a day after the International Women's Day march in Baku in 2021, as Shahmarzadeh described in an interview to OC Media (Giyasbayli 2021). The account started publishing pornographic content, later removed by Facebook, but that continued spreading on Telegram.

In July 2021, human rights and political activists in Azerbaijan received phishing mails claiming to be from Human Rights Watch. The mail included a link to malware, the technical details of which were investigated by Qurium. In 2023, the official complaints about the hacks against several Azeri feminists remained unsolved (Gender Resource Centre). This is consistent with there being much less public evidence available about Azerbaijan's cyber-enabled influence activity than there is for Iran and North Korea, for instance. Prioritisation of countermeasures against certain state-actors impacts the availability of data, and consequently shapes awareness of the relative threats potentially posed by different state actors.

Meta described the Azeri campaign as "prolific but low in sophistication". It shows, however, how different types of state actors can routinely exploit blended cyber and influence operation techniques, including authoritarian regimes targeting their domestic opposition and wider regional rivals. From the perspective of this chapter and

report, the key point being that the threat to be countered, assisted by the design decisions of any framework instruments, is not singular, but rather includes a number of variants and different modalities.

A critical gap in democratic responses to different information, cyber and hybrid threats is the “siloe” approach, where organizational structures are effectively built and ‘fitted’ to address only one type of threat. Threat actors routinely, persistently and methodically try and test methods that go beyond these organizational boundaries and mandates. Meta’s answer to this was an online operations “killchain” framework, aimed at helping to disrupt and counter activities such as cyber-enabled influence campaigns, like the Ghostwriter campaign or the network originating in Azerbaijan (Nimmo and Hutchins 2023).

2.6 Reflection

This chapter has highlighted some of the main challenges and opportunities within the current disinformation countermeasures landscape that any framework designer needs to address. A potential limitation of any FIMI framework stems from how they are constructed on the basis of what can be observed and the available evidence at a particular moment in time. In so doing, the resultant framework steers and guides the attention of users in terms of what subsequent evidence they look for and where. Thereby, inducing a recursive tendency for a subtle reproduction of what informed the original framing.

The cases described in this chapter, show this tendency especially well. The current state-of-the-art is that we know much less about the information manipulation and interference threat activities of middle powers, compared to the more well-known and intensively studied state actors such as Russia, China and Iran. A similar challenge is also evident studying these bigger threat actors, where the evidence displays a propensity to coagulate and gravitate around a limited number of cases (Internet Research Agency, Doppelganger, Storm1516, Spamouflage, for example).

To expand this horizon, this chapter deliberately sampled on different cases, to shine a light upon the biases and assumptions built into many of the existing frameworks. The focus on cyber-enabled influence phenomenon, reflects an emerging threat with a growing profile. But it is also one that is sometimes underplayed and undervalued, because it crosses the convenient institutional and organizational boundaries and categories within the defender community that organise who does what and when.

Different organisations engaged in countering malign influence efforts have variously been inspired by and adapted ideas from Cyber Threat Intelligence into their practices, notable examples being Meta’s early attribution policies and more recent VIGINUM’s and EU DisinfoLab’s framework of Information Manipulation Sets (IMS) (Gleicher and Stamos 2018; EU DisinfoLab et al. 2026). This is a development that is especially

helpful in assessing the technical layers of influence campaigns. However, cyber threat intelligence is less well equipped to provide tools to capture the ‘cognitive’ or ‘psychological’ layers of influence operations, that are especially prominent in the Russian doctrine of ‘information confrontation’. Similarly, measuring the information effects of influence operations or the measures taken to counter them, go beyond the scope of cyber inspired frameworks.

3 Methodology and Typology

This chapter operationalises the investigative directives established in the grant agreement for Work Package 3 (WP3). The core mandate of this deliverable is to demonstrate how detection, analysis, and countermeasures converge into a holistic analytical process aimed at raising the costs of Foreign Information Manipulation and Interference (FIMI) activities. To maintain fidelity to this objective, the research methodology is structured around two interdependent tracks: a feasibility study of existing frameworks and an empirical evaluation of countermeasure effectiveness.

3.1 Investigative Mandate

The research design underpinning this report addresses the primary question posed by WP3: *How can current tools and processes be aligned to analyse FIMI and impose tangible costs on adversaries?* To answer this, the methodology pursues two distinct, but complementary, lines of inquiry:

- **Framework Feasibility:** This track constructs a comparative account of existing FIMI analytical frameworks, examining their shared characteristics, differences, strengths, limitations, and interoperability potential. Practitioner insights complement this analysis by providing perspectives on the practical use and applicability of these frameworks across different organisational contexts.
- **Countermeasure Effectiveness:** The second track constructs an account of what is known, and what remains unknown, about the efficacy of FIMI and disinformation countermeasures, both individually and collectively. Thereby providing an overview of the current evidence base and its remaining limitations.

Both tracks are informed by dedicated data collection protocols, data analysis, and conducted expert interviews ensuring that theoretical findings are rigorously tested against operational realities.

3.2 Contextual Countermeasure Landscape Analysis

Separate from the two main evaluation tracks, specific case studies were introduced to inform the contextual chapter of this report. This chapter functions as a precursor to the main detailed empirical analysis, focusing on describing the current counter-FIMI landscape and its structural challenges and opportunities.

By examining real-world incidents through these cases, this methodology maps out the environment in which frameworks and countermeasures must operate. This includes analysing thematic contexts such as election cycles, geopolitical crises, and other actor-specific scenarios. The goal is to ground the evaluations of frameworks and countermeasures in a realistic understanding of the pressures and complexities faced by practitioners on the ground.

In Chapter 2 – *Current Countermeasure Landscape: Challenges and Opportunities*, this approach was operationalised through a dataset constructed from reports identified as potentially relevant to assessing whether an increased convergence between cyber and influence operation strategies and tactics can be evidenced. Data collection for this was conducted in October 2025, covering the period 01.01.2016 - 30.09.2025. To achieve a relatively balanced sample, two companies focusing on cyber-security (Microsoft and CrowdStrike), two tech platforms (Meta and Google) and two NGOs working on influence operations (DFRLab and EU DisinfoLab) were included as sources for relevant reporting documents.

A search was conducted on the organisations' websites with keywords "cyber" and "influence". Returns on these terms were filtered to include the entities' reporting about their investigations into cyber incidents or influence operations, resulting in 463 reports and articles in total. The URLs were then saved. The relevant results were then cross-checked to identify reports that mention both "cyber" and "influence".

There are some caveats and limitations to the search. Strict focus on keywords means that, for example, reports about NotPetya, VooDooBear and MummySpider cyber operations are not included, because the articles do not mention cyber. Similarly, the search does not bring up reports talking about disinformation, misinformation, information operations, hacking, phishing or malware, unless they specifically mention 'cyber' or 'influence'. Not all organisations were active in 2016 (like EUDisinfoLab), and some have restructured their websites so that earlier reports do not show up in searches (Microsoft). In the results, 'influence' was included only in the context of influence operations or manipulation. Similarly, 'influencer' was included.

3.3 Framework Identification Methodology, Typology, and Analysis Preparation

The initial phase of the feasibility study focused on mapping the topography of the counter-FIMI landscape. To ensure academic rigour and avoid excluding relevant methodologies, an inclusive search protocol was established to identify frameworks related to FIMI, Hybrid Warfare, Disinformation, and Misinformation.

Framework Collection

Frameworks were identified using structured search queries combining the term “framework” with the FIMI-related keywords “FIMI,” “Hybrid Warfare,” “Disinformation,” and “Misinformation.” To ensure broad coverage and diversity of sources, multiple search engines and platforms were used. Academic databases such as Google Scholar and Scopus were consulted to capture peer-reviewed literature and ensure scientific rigour, while Google and LinkedIn were used in parallel to identify frameworks from non-academic sources, including industry actors, companies, and other organisations.

All identified frameworks were documented in a structured Excel database, where each entry was read, summarised, and annotated. Given the broad and inconsistent use of the term “framework,” an inclusive approach was adopted at this stage, with no prior filtering or strict classification applied in order to avoid premature exclusion of potentially relevant contributions.

As the dataset expanded, preliminary categorisation structures were developed iteratively based on observed characteristics across the collected frameworks, supporting a more grounded understanding of the emerging corpus. The data collection phase resulted in an initial dataset of 159 frameworks, alongside three additional EEAS reference frameworks that were excluded from the analytical sample.

Development of the Working Definition and Filtering the Dataset

As the data collection was purposely kept broad at the initial stage, any item labelled as a “framework” in the literature, by experts, or in relevant publications was included. Comparison of the collected frameworks revealed substantial differences in their use, scope, and limitations. This highlighted the need for a working definition of frameworks, as the term was used inconsistently across sources and often referred to conceptually different constructs. Therefore, establishing a clear and shared understanding became necessary to ensure consistent inclusion criteria and systematic filtering of the dataset. To develop this definition, a targeted literature review was conducted on how frameworks are defined and distinguished from related constructs such as theories and models. Recurring characteristics identified in the literature were then synthesised into a working definition tailored to this research, which reads as follows:

“A framework is a structured representation that organizes key concepts, variables, or components of a subject matter into a coherent system and specifies general relationships between them, to support systematic understanding, analysis, or action. Frameworks provide a foundation for inquiry by defining relevant elements and their connections, but do not in themselves offer causal explanations or predictive power like theories and models do. Frameworks do not implement computational procedures, perform algorithmic operations, process data, or generate predictive outputs.”

The working definition highlights the need for structure and organisation derived from McGinnis & Ostrom (2014). It further incorporates the need for connectivity and plurality (Partelow, 2023), as well as the relationships between components (Binder et al., 2013). Additionally, the understanding of frameworks as a foundation for inquiry was derived from Schlager (2019), alongside the distinction that frameworks should not themselves provide predictive capabilities. The final sentence of the working definition specifically addresses the technical literature reviewed during framework collection and differentiates frameworks from purely computational or algorithmic structures.

Due to the FIMI-focused context of this research, frameworks were also required to demonstrate relevance and applicability to the topic. For this reason, an additional relevance component was added to the working definition:

“Within this context, frameworks must be applicable to FIMI, demonstrate transferability beyond highly domain-specific applications, and be relevant to at least one stage of the FIMI response process.”

The relevance of a framework was assessed in light of the evolving and overlapping nature of FIMI, where boundaries with related areas such as disinformation and influence operations are not always clearly defined. *Therefore, the criterion was intentionally framed broadly around whether a framework could reasonably contribute to the analysis or response to FIMI, rather than requiring it to have been explicitly developed for FIMI.* This reflects how FIMI builds on existing concepts and practices, allowing for the inclusion of frameworks developed before the emergence of the term itself, as well as frameworks addressing related or partial aspects of FIMI. Accordingly, relevance was assessed through three main areas:

- **Applicability to FIMI:** Frameworks were included if they were either specifically designed for FIMI or could reasonably be applied to the context, including frameworks originally developed for other contexts but relevant to FIMI-related analysis.
- **Transferability:** Frameworks need to be able to be applied across different situations and topics, rather than being restricted to a single-issue area.
- **Relevance to the FIMI Response Process:** Frameworks were required to be relevant or applicable to at least one stage of the FIMI Response Process, which is further elaborated in the categorisation section.

This FIMI response process was developed specifically for the purpose of this research. While a range of established process models exist in the FIMI and disinformation literature, many of these are structured around threat-actor-centric perspectives such as kill-chain approaches. Given that this research focuses on frameworks and countermeasures in a response context, a response-oriented structure was developed. The FIMI response process serves both as a categorisation device and as a means of

assessing whether and how existing frameworks can be applied within a FIMI response perspective. In this way, it helps to identify as well as organise the identified frameworks, to ensure that they are meaningfully situated within the scope of FIMI response analysis.

The framework working definition was subsequently operationalised as a filtering instrument through a two-step filtering process:

- The first step assessed whether collected items could be classified as frameworks according to the conceptual component of the definition. One analyst initially encoded the dataset independently, classifying each item as “yes,” “no,” or “unsure” based on its alignment with the framework definition. “Unsure” cases were subsequently reviewed and discussed jointly by two analysts, after which all items were finalised as either “yes” or “no.” This process resulted in the identification of 43 frameworks and a corresponding reduction of the working dataset.
- The second step evaluated whether the identified frameworks were relevant to the scope and context of this research. The relevance component of the definition was applied jointly by two analysts, who assessed each framework according to the three established relevance criteria. This resulted in the final sample of 31 frameworks.

While the assessments were guided by the established working definition and relevance criteria, the filtering process remained interpretive and therefore subject to analyst judgement. In cases of uncertainty, an AI assistant was used to support discussion by outlining arguments for and against inclusion.

Development of a Categorisation Scheme

Initial category ideas emerged during the collection phase through iterative note-taking and preliminary observations. During the filtering stage, the categorisation structure was continuously reviewed to ensure alignment with the evolving dataset and the inclusion criteria. After the final sample had been determined, the categorisation scheme was systematically revisited and adjusted to ensure internal consistency and coherence across the selected frameworks.

3.4 Countermeasure Evaluation Methodology

Chapter 5 – *Evaluation of FIMI Countermeasure Effectiveness* aims at constructing an account of what is known, and what remains unknown about the efficacy of FIMI and disinformation countermeasures, both individually and collectively; therefore, relying on the collection of specific datasets distinct from both the framework literature and the contextual case studies. While insights from research literature create a basis, they are blended with findings from the expert interviews and enriched with quantitative

data publicly available on countermeasures. The aim is to highlight some of the different angles for understanding the scope and scale of countermeasures. For this view, many of the publicly available figures originate in Meta's reporting. Figures for distinct fact-checking articles used on Facebook and Instagram to label and reduce the virality of content in the EU, as reported by Meta, were collected from reports under the Code of Practice on Disinformation 2023-2025. For Meta's coordinated inauthentic behaviour (CIB) takedowns, the figures were collected from 100 threat reports listed on Meta's threat reporting website⁵ and cover takedowns from 2017 until 2025. To obtain a preliminary overview of patterns and trends in EU sanctions imposed in response to disinformation and FIMI activities, a search was conducted across sanctions regimes with keywords such as 'propaga*', 'disinfor' and 'manipulat*' within the reasoning text of restrictive measures lists⁶. The French government agency for protection against foreign digital interference (Viginum) uses OpenCTI and has developed a French version of DISARM, sharing its findings on GitHub.⁷ These figures were collected as examples of technical indicators. The limited availability of samples further highlights the lack of publicly accessible data provided by most platforms.

3.5 Expert Interviews

To complement the framework and countermeasures analysis and diversify the research perspective, semi-structured interviews were conducted to incorporate expert practitioner perspectives and experiences from the field. This approach allowed the research to move beyond solely theoretical and published frameworks by including practical insights into the actual use, evaluation, and implementation of FIMI-related frameworks and countermeasures. The interviews capture operational perspectives and practical considerations that may not be reflected in existing literature or publicly available framework documentation. Throughout this paper, interview participants are collectively referred to as *experts*. This term is used to denote their domain-specific expertise relevant to the research and encompasses participants from diverse professional roles and organisational backgrounds, including government, academia, industry, and civil society.

3.5.1 Recruitment and Sample

Given the interdisciplinary nature of FIMI, the recruitment process aimed to capture a broad range of perspectives across research domains, professional sectors, and geographic contexts. Seventeen experts were therefore recruited from different disciplinary backgrounds (e.g., political science, journalism, psychology) and

⁵ <https://transparency.meta.com/metasecurity/threat-reporting/>

⁶ Sanctions regimes as listed here <https://www.sanctionsmap.eu/#/main>, search conducted on restrictive measures consolidated texts of EU Council Regulations at <https://eur-lex.europa.eu/>.

⁷ <https://github.com/VIGINUM-FR>

organisational contexts, including governmental organisations, NGOs, academia, corporate actors, and other relevant institutions. In addition, participants were selected from multiple countries, across the EU, to ensure a broader contextual representation of approaches to FIMI. The recruitment process further aimed to include both strategic and operational perspectives, as well as a balance between experts working on conceptual and theoretical dimensions of FIMI and practitioners directly engaged in applied responses across different areas of FIMI-related practice and research.

The interviews were conducted online, lasted around 30 minutes, and were transcribed for analysis. Three interviewers participated in the data collection, all following the same semi-structured interview guide to ensure consistency while allowing flexibility to explore participants' individual expertise and experiences through follow-up questions where relevant. In line with the scope of this research, the interviews were intended to develop a broad understanding of the field and to inform the analysis, rather than to provide an exhaustive or statistically representative account of practitioner perspectives.

At the beginning of each interview, participants were informed about the interview procedure, the application of the Chatham House Rule, and the use of recordings for transcription and analysis, for which informed consent was obtained. To protect participants and account for the operational sensitivity of the topic, responses were anonymised throughout the analysis. The detailed interview guide is not included in this report due to the public availability of the report and the potential for misuse. Instead, only the general thematic structure is described.

The interviews covered four broad themes: (i) participants' professional background and organisational approach to FIMI, (ii) the practical use, strengths, and limitations of FIMI-related frameworks, (iii) organisational countermeasures and operational responses, and (iv) emerging challenges, gaps, and future needs within the FIMI landscape. Framework- and countermeasure-related topics were addressed separately to distinguish analytical from operational perspectives, and support a clearer comparison in the subsequent analysis.

4 Evaluation of Frameworks related to FIMI Countermeasures

4.1 Introduction to the Analysis

This chapter adopts a framework feasibility perspective that combines structured framework analysis with insights from the expert interviews. The primary analytical focus lies in the systematic identification, description, and comparison of existing frameworks related to FIMI. This includes assessing their characteristics, scope, and applicability, as well as identifying strengths, limitations, and patterns within and across different categories of frameworks. While expert interviews provide additional contextual insight into how frameworks are applied in practice, they are not intended to provide a comprehensive or representative account of all usage contexts. Instead, the chapter maps the current framework landscape, highlighting areas of convergence, divergence, and potential gaps in relation to FIMI response needs.

4.1.1 The Categorisation

To better assess, describe, and compare the identified frameworks, multiple categorisation dimensions were defined. As outlined in the methods section, these categories were developed iteratively throughout the framework collection, filtering, and sorting process. The categorisation dimensions used in this analysis are the phase of the FIMI response process and the type of framework.

4.1.2 Phase of FIMI Response Process

As described in the methodology section, the FIMI response process was developed specifically for this research to filter and categorise the frameworks. It does not constitute a new theoretical model or normative framework, but serves as an analytical and structuring device used to guide the reader and support comparison across frameworks. This process enables us to assign each framework to its main task, role, and stage in fighting FIMI. Thereby, creating a clearer understanding and an easy-to-follow structure within the diverse and interdisciplinary landscape of FIMI-related research. In turn, this supports the presentation and interpretation of the material by making the analysis more transparent and structured.

This response-oriented perspective was adopted because the focus of this paper is on addressing and countering FIMI from a defensive standpoint. As a result, the process is designed to reflect the stages of response activity (see Figure 7 below), rather than a threat-actor perspective (such as kill chain or DISARM Red), aligning the analytical structure to the objectives of the research.

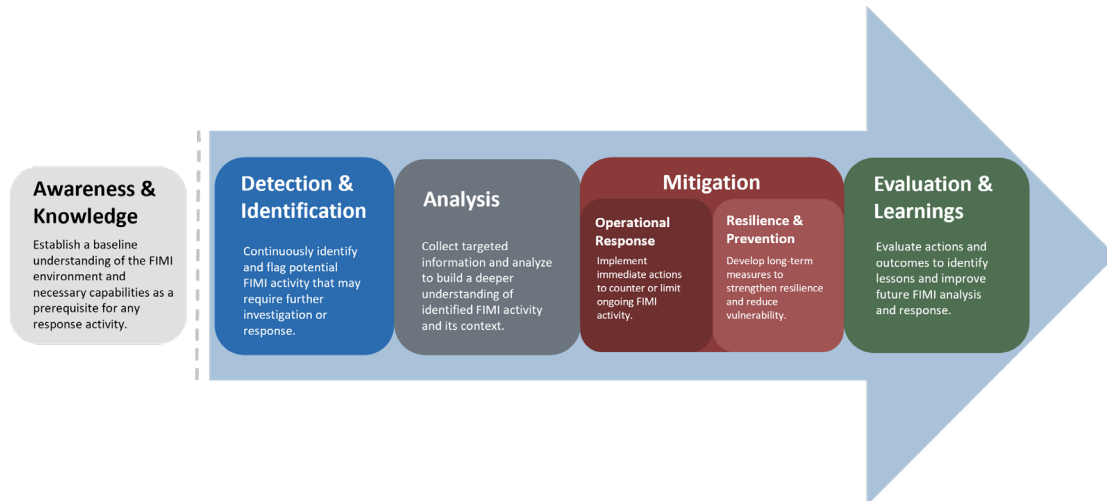


Figure 7: Phases of the FIMI Response Process.

First, it should be noted that situational awareness of the FIMI environment, together with a baseline understanding of the field and the necessary capabilities, is a foundational prerequisite for engaging with the FIMI response process. It is therefore treated as an external prerequisite, rather than a distinct analytical phase. The FIMI response process begins once a potential threat has been detected during the “*Detection and Identification*” first phase. This phase includes frameworks that support the detection and identification of FIMI campaigns or related FIMI activity, including disinformation, with a primary focus on the recognition, flagging, and initial classification of potentially relevant activity.

The next phase is “*Analysis*” and includes frameworks that support the collection, processing and interpretation of information related to identified FIMI activity. While the Detection and Identification phase focuses on recognising potential FIMI activity, this phase focuses on confirming (or refuting) the presence of FIMI actions, and then developing a deeper understanding of these through further investigation, contextualisation, and assessment of findings. This can include analysing communication dynamics, behavioural aspects, dissemination patterns, or underlying factors that influence how and why information is spread and engaged with.

The third phase is “*Mitigation*”, which focuses on responding to and countering FIMI activity. This phase is divided into two sub-categories: “*Operational Response*” and “*Resilience and Prevention*”. Operational Response describes the immediate and direct response to FIMI activity and includes frameworks that support concrete actions and short-term intervention measures. In contrast, Resilience and Prevention Development focuses on long-term approaches aimed at reducing vulnerability to FIMI and

strengthening preparedness, resilience, and preventive capacities. This differentiation helps distinguish between frameworks intended for immediate operational action and those aimed at longer-term structural or societal responses to FIMI.

The fourth and final phase is entitled “*Evaluation and Learnings*” and includes frameworks that support the assessment and evaluation of existing frameworks, actions, and countermeasures related to FIMI. These frameworks focus on identifying strengths, limitations, and areas for improvement, as well as supporting the integration of improvements and findings into future analysis, response strategies, and mitigation efforts. Lastly, it should be noted that a framework may be applied within a single phase or across multiple phases of the FIMI response process.

4.1.3 Type of Framework

This initial categorisation groups the frameworks according to their type, describing both their origin and primary mode of operation:

- The *technical approach* includes frameworks related to technical domains such as cybersecurity, algorithms, or technical infrastructure. It includes frameworks originating in cybersecurity that have been adapted for FIMI contexts. These frameworks focus on systems, data, infrastructure, and computational or analytical mechanisms.
- The *behavioural approach* centres frameworks that focus on human cognition, attitudes, behaviour, communication, and social dynamics, drawing on fields such as psychology and communication research.
- The *operational approach* covers frameworks that support practitioners in their day-to-day work by structuring, coordinating, and guiding practical responses to information manipulation, with a focus on implementation and reporting rather than strategic or conceptual oversight. These frameworks are particularly relevant for actors involved in understanding, acting on, reporting, or coordinating responses to information manipulation.
- *Strategic approach* frameworks address high-level planning, policy-making, coordination, and top-level conceptualization of responses, distinct from the planning or execution of concrete actions.

This categorisation forms a central analytical layer in our subsequent framework analysis, as it provides a clear structural grouping whereby each framework can be assigned to a single primary category. It also enables the identification of patterns across different types of frameworks and supports the comparison of how different approaches contribute to understanding and responding to FIMI.

Short summaries of the analysed frameworks are enclosed in Annex A, organised according to their respective clusters. These summaries provide readers with an overview of the frameworks and can be consulted before or during the analysis, particularly when a framework is unfamiliar or requires additional context. Tables of the frameworks meeting the requirements for analysis are provided under each respective cluster section, and a comprehensive table outlining all collected frameworks is provided in Annex B. Given the comparative nature of the framework analysis, the following sections focus on identifying similarities, differences, and recurring patterns across frameworks. To maintain readability and analytical flow, frameworks are referred to by their names or authors rather than being repeatedly cited in-text.

4.2 Comparative Analysis of the Technical Frameworks Cluster

The technical framework cluster focuses on systems, data, algorithms, technical infrastructure, and analytical mechanisms. It draws from technical research domains such as cybersecurity and computer science. Technical systems (e.g., algorithms, recommender systems, platform affordances) actively shape online information sharing and are therefore relevant to FIMI and its spread (Aïmeur et al., 2023; Goh et al., 2017). Furthermore, the detection, scalable monitoring, mapping, and analysis of large data sets frequently relies on technical signals (metadata, network traces, system-level indicators) and tools, highlighting the need for algorithms and technical frameworks to support this process. FIMI is an interdisciplinary challenge, requiring in part a technical approach to be able to analyse and counter it. This cluster provides technical insight and perspective on FIMI, and how it spreads online.

In our first unfiltered initial collection step, we accumulated 72 supposed ‘frameworks’, the largest initial cluster. However, the majority of those did not qualify as a framework according to our working definition, and 65 were discarded in the first filtering step. In the second filtering step, five more were removed due to their lack of transferability, as they were too narrowly focused on cybersecurity. Therefore, only **two** frameworks remained after filtering, **2.8%** of those collected.

Name of Framework	Authors	Source reference
DISARM red/ AMITT (Adversarial Misinformation and Influence Tactics and Techniques)	developed by MisinfoSec Working Group / Credibility Coalition in 2018; later adopted into DISARM Foundation	https://www.disarm.foundation/framework , https://disarmframework.herokuapp.com/about
The Online Operations Kill Chain (OOKC)	B. Nimmo and E. Hutchins	https://carnegieendowment.org/research/2023/03/phase-based-tactical-analysis-of-online-operations

Figure 8: Table of analysed technical approach frameworks

This significant reduction in the first filtering step could be due to a difference in conceptual understanding of the term framework. In the technical domain and computer science, the term was broadly used to describe algorithms, pipelines, protocols, and technical systems. They often lacked broader analytical structuring, conceptual relationships, or support for systematic understanding, as required by our working definition. This difference could also be explained by the different approaches in the technical literature, where solution-oriented and implementation-focused approaches dominate, focusing on solving specific technical problems, rather than creating a structured understanding. Another aspect is that some of the found “frameworks” originated from a commercial development context, and the term there was also used for methodologies, products, workflows, or service-oriented structures; further differing from our working definition. The second filtering step reduction could be explained by the limited transferability of cybersecurity-focused approaches to broader FIMI contexts, as they were often highly specialised and technically narrow approaches.

4.2.1 Similarities and Shared Ground

Both frameworks share the same premise and ancestry in cyber-security, while claiming general applicability across topics from elections and public health to conflict and cybercrime. They both descend explicitly from the Lockheed Martin Intrusion Kill Chain, which Nimmo and Hutchins explicitly state, while DISARM lists the Intrusion Kill Chain alongside other models (e.g., MITRE ATT&CK⁸, the Diamond Model, the Pyramid of Pain⁹) as inspiration. Therefore, the frameworks both assume that malicious online activity, regardless of motive or content, leaves a trail of tactical steps that can be catalogued, compared, and disrupted link by link.

Furthermore, they share the same approach, being practitioner-oriented rather than for academic theory building or abstraction of FIMI. They both have a strong modular/component-based design as they break FIMI into discrete units of analysis (steps, techniques, observations), therefore structuring it into manageable analytical building blocks, and avoiding treating FIMI as a single continuous phenomenon. They also remain explicitly tactical and not strategic, allowing them to focus upon what is happening ‘on the ground’, rather than veering into the geopolitical or strategic level.

They also centre aspects of the analysis that can be directly measured. DISARM does have an assessment element, but keeps it separate from the observations. They both aim to standardise how analysts describe activity and reduce ambiguity in terminology to create consistency across different analysts and organisations, as well as comparability across cases. The two aspects supporting their broad applicability are the platform-agnostic and the threat-actor-agnostic design of the frameworks. The

⁸ <https://attack.mitre.org/>

⁹ <https://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html>

platform-agnostic design of the Online Operations Kill Chain (OOKC) framework allows for applicability across different platforms and environments, extending past social media, and including aspects like websites and email providers. DISARM is also platform-agnostic and used across the FIMI defender community (governments, platforms, civil society, academia) per the 2024 review.¹⁰ Additionally, both frameworks are purposefully designed not to require attribution as a precondition for analysis.

The two frameworks also share some limitations. They both treat the threat actor's strategic goal as out of scope for the framework itself. Instead, it is to be inferred from a pattern of observed techniques or assessed alongside other forms of intelligence (geopolitical, narrative, content analysis). They also both focus on mapping and describing observable tactical activity, but do not provide full support for higher-level assessment, including attribution and strategic intent. As a result, the interpretation of broader campaign purpose and impact remains external to the frameworks themselves. Lastly, both frameworks claim coverage across different topics, but neither has been systematically evaluated in varied areas, for example, FIMI directed at humanitarian operations, electoral integrity in non-Western contexts, or AI-generated content campaigns.

4.2.2 Contrasts and Divergences

The two frameworks differ in the scope of threats they cover. The OOKC is deliberately broader than FIMI, as it also covers other aspects (e.g., cyber-attacks, influence operations, online fraud, human trafficking, and terrorist recruitment) and aims to analyse hybrid operations (e.g., Ghostwriter) that fit no single category. DISARM, however, is FIMI-specific by design, as pointed out in the 2024 review, where users pressure to widen the scope. The foundation has expressed openness to such an extension, but as of 2026, the framework does not include it. Therefore, the OOKC prioritises depth trading off breadth, whereas DISARM exchanges breadth for depth and standardisation.

They also differ in their structural logic, as the OOKC uses a linear but flexible operational spine of modular links, while DISARM uses a matrix. The OOKC models the process of an operation with a possible “daisy-chain effect” (where one operation's late-stage segues into another's early stage), whereas DISARM models components of an operation. Additionally, DISARM has been changed from a four-stage kill chain (plan, prepare, execute, assess) with one-to-one mapping of techniques to stages in v1 to an Asset-Action-Content matrix derived from the ABCDE model in v2. This redesign converges towards design choices Nimmo and Hutchins made at the outset (modular, observation-based, flexible across operations, requiring less full operation reconstruction and no strict lifecycle), therefore, closer aligning the two frameworks.

¹⁰ <https://adacio.eu/a-comprehensive-review-of-disarm>

Granularity, size and maintenance are other features where the two frameworks also differ. The OOKC uses the simplest model that still works in practice, as investigators must hold the model in mind across many concurrent cases, leading to a cognitively manageable, applicable, stable, but less detailed framework. DISARM, on the other hand, is large, intricately detailed and expanding, offering rich coverage that, on the downside, leads to some difficulty in use. In the 2024 review of DISARM, users complained about “availability bias” (Tversky & Kahneman, 1973), being influenced by how easily available certain information is, as analysts cannot hold the whole framework in mind and default to techniques they have used before. DISARM is trying to address this problem by adding ‘help’ tools like playbooks to reduce cognitive load, using AI-assisted tagging, and allowing for better filtering and searching so that users don’t have to navigate the full taxonomy manually anymore. The capacity to maintain v2.0 and its tooling depends on continued investment, whereas the OOKC has no equivalent maintenance commitment. This aspect is important for future development, as any framework or model built on either of these inherits their respective sustainability risks.

Another aspect where the frameworks diverge is their integration with technical standards, as DISARM has a developed tooling ecosystem while the OOKC has no equivalent tooling. DISARM offers a Navigator, Explorer, Word Add-In, an AI-assisted annotator in prototype, and integration with OpenCTI and the OASIS DAD-CDM common data model. The framework is hence not just a framework document anymore; it comes with a technical ecosystem that increases implementation complexity and integration requirements, but also enables greater interoperability, scalability, and machine-readable information sharing. The OOKC does not have anything comparable, as it is more of a thinking model and not embedded into structured data systems, which makes adoption easier, but scaling and sharing harder.

They also differ regarding their base institutional model. The OOKC is a single-author model that is analytically useful, easy to understand, conceptually strong, and builds on the authors’ experience inside a major platform’s threat-intelligence function. DISARM is a community-stewarded standard, maintained collaboratively by many contributors with ongoing revisions and development, aiming to become institutionally recognised and interoperable across organisations. DISARM is regularly updated, while the OOKC is essentially static with one publication.

The two frameworks also differ in their treatment of inference and assessments. The Kill Chain catalogues only observable aspects, considering intent as a separate part; therefore, keeping inference entirely out of the framework. DISARM used to mix observed and inferred elements in v1, which was identified as a major source of inconsistent application across analysts, hence reducing consistency and interoperability. In v2 of DISARM, the inference aspect was kept, but split from observations in a dedicated framework with “points to prove” (meaning that analysts justify why they believe something, what evidence and how they reached the conclusion), allowing for better understanding and comparability.

4.2.3 Coverage of the FIMI Response Process Phases

Both technical frameworks serve the same two phases: *Detection and Identification*, as well as the *Analysis* phase (see Figure 9 below). The concentration of technical frameworks in these two phases reflects the broader strengths of the technical domain, which focuses on observability, monitoring, tracking, and the analysis of technically measurable indicators. As FIMI activity is largely facilitated by technically mediated environments such as social media platforms, websites, algorithms, and digital infrastructure, technical frameworks are particularly suited for identifying and analysing manipulation activity within these systems.

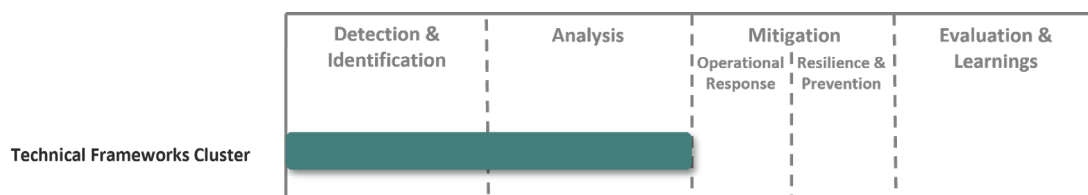


Figure 9: Technical Frameworks Cluster Coverage of FIMI Response Process Phases.

Both Mitigation phases, *Operational Response* and *Resilience and Prevention*, as well as the last phase, *Evaluation and Learnings*, are not covered by the technical frameworks. This gap could be explained by the heritage and origin of the technical frameworks, as they were influenced by cybersecurity traditions, where identifying, tracking, and classifying adversarial activity is often prioritised over long-term societal resilience or evaluation processes.

In contrast, both Mitigation phases often require broader contextual understanding beyond the technical domain, including political, behavioural, societal, communicative, or psychological dimensions that both frameworks do not incorporate. It is worth noting that DISARM blue¹¹, could help fill this gap to some extent by categorising countermeasures and providing a framework for the mitigation phases, but as the framework is currently still in development, it can't be evaluated and considered here.

The general absence of frameworks covering Evaluation and Learnings could indicate both a difficulty in translating long-term adaptive learning processes into structured technical frameworks and a broader research gap across the FIMI framework landscape. The focus on early phases of the FIMI response process may imply that the technical framework landscape currently prioritises understanding and evidencing manipulation activity over responding to, preventing, or learning from it. The limited phase coverage highlights a broader gap in the technical framework landscape and reinforces the need for complementary framework categories to support a more comprehensive FIMI response process.

¹¹ DISARM Blue is the counter-disinformation half of the DISARM framework, cataloguing response and mitigation tactics. <https://disarmframework.herokuapp.com/about>

4.2.4 Application of Frameworks in Practice

In the expert interviews, only one technical framework was mentioned: DISARM. Nine of 17 experts mentioned the framework, highlighting how popular and well-known the framework is among FIMI practitioners. This high level of recognition makes sense, given its use as a standardised language for communicating findings and coordinating work across organisations. Interoperability is one of the challenges in FIMI research, making frameworks that allow for coordination and coworking relevant. The popularity of DISARM suggests that practitioners place significant value on standardisation and interoperability, two recurring challenges identified throughout the FIMI field. Its popularity may also reflect not only the quality of the framework itself, but also the absence of alternative common standards.

Experts repeatedly described DISARM as helping them organise, categorise, and structure their analytical activities. The framework reduces complexity by transforming diverse incidents into a structured and comparable format. Another use is its role in communicating findings. This indicates that technical frameworks can be both analytical tools, but also a basis for communication, coordination, and cooperation, as their shared vocabulary enables cooperation between organisations and allows findings to be exchanged more efficiently. Furthermore, the use of DISARM by experts varied, some describing using it as a checklist manually, while another expert reported that their organisation has automated its use. This demonstrates that the framework is sufficiently structured to support automation, while still being usable in more traditional analytical workflows.

The main negative aspect mentioned by experts is the complexity of DISARM, a criticism already highlighted earlier. The complexity appears to be a direct consequence of the framework's ambition to comprehensively capture the wide range of FIMI tactics and techniques, which may represent a trade-off between comprehensiveness and usability. Experts mentioned that this complexity makes it time-consuming to work with. This downside can be counteracted by automation, as explored by another expert. Furthermore, the complexity might not be a challenge for analyst-to-analyst communication, but could become a limitation when communicating findings to policymakers, decision-makers, or the public. Lastly, experts highlighted that identifying TTPs alone is insufficient and that the framework excels at describing how manipulation occurs, but is less equipped to explain why it matters or what should be done next.

None of the other technical frameworks were used, which could indicate the popularity of DISARM, but it should be noted that popularity does not necessarily equal superiority. Two experts mentioned specifically that they use DISARM because it is familiar, indicating that adoption and community acceptance may be as important as technical quality. Additionally, several experts noted they have internal technical frameworks and workflows adjusted to their work and organisation, built on their experience in the form of a technical layer or algorithms they use. The use of algorithms

for text analysis, narrative detection, and monitoring demonstrates that technical capabilities remain central to FIMI work even when formal frameworks are absent and highlights the use of tools, systems, and automated processes, rather than standalone frameworks.

4.2.5 Contributions to a Comprehensive Analytical Approach

Both analysed technical frameworks see FIMI as a threat to be detected and mapped. As such, they emphasize observable data, use standardized categories, and reduce reliance on an individual analyst's interpretation. While the two differ in certain aspects, as detailed, they both could contribute different aspects to a comprehensive analytical approach. The OOKC contributes a compact and structured vocabulary for describing different types of online threats, even when it is not yet clear whether they involve FIMI, cybercrime, harassment, or a combination of these. It further offers a clear and disciplined separation between observation and interference. Additionally, it allows for describing operations that combine FIMI with other harmful activities, such as hacking, surveillance, or online harassment. Lastly, it provides support for information sharing and communication between different organizations and teams, including governments, platforms, and researchers.

DISARM red provides depth and granularity within the FIMI domain that the OOKC does not attempt to include (e.g., persona types, content manipulation techniques). It further provides a machine-readable standards-integrated taxonomy that enables bulk analysis and sharing at speed and scale. It also offers regular release cycles through its community-governance system, which aims to improve the framework based on feedback. Lastly, the framework provides a growing library of pre-tagged incident reports and themed playbooks that lower the barrier to consistent application for new users.

The tensions between the two frameworks are best understood as competing design priorities rather than direct disagreements. They highlight important trade-offs between simplicity and depth, flexibility and standardization, and accessibility and technical sophistication. Any future integrative approach will need to consciously balance these competing demands, as emphasizing one side inevitably comes at the expense of the other. A major gap of the technical cluster is that it does not cover the full FIMI response process, and while the technical aspect is important in fighting FIMI, it should be combined with approaches from other disciplines.

Nonetheless, both technical frameworks offer important aspects when building an integrative approach to FIMI:

- The OOKC provides compactness and cross-threat coverage well-suited to early-stage triage and useful to organizations working across multiple harm types (e.g., small civil society teams, regulators).

- DISARM provides depth, standardization, and tooling that make it a natural fit for sustained FIMI analysis and for cross-organizational data sharing at scale.

Taken together, the two frameworks provide complementary approaches to operational analysis and classification, which are also reflected in their increasing similarity. They highlight the trade-offs that any future integrative framework will need to navigate. However, they do not cover operational, resilience and prevention mitigative measures, nor evaluative and learning aspects.

4.3 Comparative Analysis of the Behavioural Frameworks Cluster

The behavioural framework cluster focuses on human cognition, attitudes, behaviour, and social dynamics, drawing on fields such as psychology, information, literacy, and communication research. This category, therefore, offers additional insight by touching on the human component in countering FIMI. These additional aspects are especially valuable, as FIMI is complex and inherently interdisciplinary, touching on issues affecting society as well as the individual (Arribas et al. 2025; Shuliak et al. 2026). Online information environments, such as social media and messaging applications, are socially mediated as individuals interact with one another, share, reinforce, or react to information (Aïmeur et al., 2023; Goh et al., 2017). There, individuals can become targets of FIMI when malign actors disseminate disinformation or other forms of misleading or manipulative information intended to influence perceptions, promote specific narratives, undermine trust in governments, science, or institutions, and shape attitudes, behaviour (e.g., voting behaviour), and decision-making (European External Action Service, 2024; Marsden et al., 2019; Van Der Meer & Hameleers, 2025).

Hence, analysing behavioural aspects, social dynamics, and communication behaviour is essential for understanding the impact of FIMI and its spread, as well as to, in the next step, be able to develop mitigation and preventative measures. Another important aspect to consider is that FIMI actors are aware and actively exploit human cognitive vulnerabilities, like falling into relying on heuristics and biases, or eliciting emotional responses like fear, outrage, or social belonging dynamics. Therefore, behavioural research is highly relevant to countering FIMI, and behavioural frameworks can help explain why influence operations work, not just how they are technically conducted.

After our first unfiltered initial collection step, we amassed 41 supposed ‘frameworks.’ However, after the two filtering steps, only 9 remained, which will be detailed and compared in this chapter.

Name of Framework	Authors	Source reference
Audience segmentation framework based on individuals' levels of exposure to and trust in misinformation	J. Zeng, J. Li , N. Yu, X. Cui, A. Xiao	https://journals.sagepub.com/doi/10.1177/13591053261419968
A systemic framework for disinformation on social media platforms	R. Recuero	https://doi.org/10.1177/29768624251367199
Break out Scale	B. Nimmo	https://www.brookings.edu/wp-content/uploads/2020/09/Nimmo_influence_operations_PDF.pdf
Disinformation behavior framework	N. K. Agarwal, F. Alsaedi	https://doi.org/10.1108/AJIM-01-2021-0034
Information Disorder Framework	C. Wardle & H. Derakshan, Council of Europe (2017)	https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html
Misinformation Recognition and Response Model (MRRM)	M. A. Amazeen	https://doi.org/10.1093/hcr/hqad040
Online Misinformation Engagement Framework	M. Geers, B. Swire-Thompson, P. Lorenz-Spreen, S. M. Herzog, A. Kozyreva, R. Hertwig	https://doi.org/10.1016/j.copsyc.2023.101739
Reframing the Information Literacy Framework to Identify Misinformation and Disinformation	K. John, B. Tater	https://doi.org/10.1080/0361526X.2025.2459765
The Misinformation Receptivity Framework (MRF)	L. Zmigrod, R. Burnell and M. Hameleers	https://doi.org/10.1027/1016-9040/a000498

Figure 10: Table of analysed behavioural approach frameworks

In the first filtering step, 29 were sorted out as they did not qualify as a framework according to our working definition. This could be due to a difference in conceptual understanding of the term “framework”. In behavioural and social science research, it was used interchangeably with terms such as model, theory, concept, and approach. Additionally, at first, papers were included that developed research frameworks for their own paper only, and were not applicable on a large scale, only referring to their specific experiment conducted. In the second filtering step, three frameworks were sorted out due to their lack of transferability across different topics, as they were focused too specifically on narrow methodology, like only prebunking, or topical aspects, such as being applied to the COVID-19 context. An additional aspect that made the assessment and selection of behavioural frameworks difficult was the use of inconsistent terms referring to false information, as false information, disinformation, and misinformation were sometimes not clearly differentiated, which made assessing FIMI relevance harder, as it assumes a clear malign intent, which is only specified as part of disinformation per most common definitions.

Lastly, it is worth noting that more frameworks remained after filtering than in the technical framework category, which could be due to the understanding of frameworks in behavioural cluster disciplines aligning more closely with our framework working definition. Another factor could be the behavioural approaches' better alignment with the interdisciplinary, complex nature of FIMI and natural intersection with other areas of research, which results in their high transferability to other topics and relevance to FIMI. This interdisciplinary nature of FIMI explains why the frameworks being sorted into the behavioural cluster also show secondary affinities with other topics to address FIMI better, such as strategic, technical, and operational.

4.3.1 Similarities and Shared Ground

All nine selected frameworks share a focus upon cognition, behaviour, communication, or social dynamics as the principal aspect of analysis, rather than technical, operational, or strategic aspects, which situates them within the behavioural framework category. The frameworks conceptualise FIMI, or more precisely disinformation, as a human-centred problem rather than a purely technical or infrastructural one, hence placing their analytical focus on how information is interpreted, trusted, shared, resisted, or acted upon by audiences. Some of the frameworks have secondary affinities to other areas, as pointed out, for example, Recueros' framework combines cognition and behaviour with a coupled platform environment. This could be explained by the fact that tackling FIMI is inherently an interdisciplinary challenge, which some frameworks aim to address by spanning multiple research areas.

Another aspect they share is in recognising the new challenges of the modern information environment. Additionally, they are also aware that these arising challenges (e.g., speed, scale, algorithmic curation, low-cost authoring, the blurring of news and other formats) were not anticipated by earlier behavioural literature. However, their analytical response varies, differing in how central they treat the platform environment itself within the analytical process (e.g., Recuero sees the platform as a co-evolving system rather than separate). All but the Breakout scale and Recuero's framework place the receiver (individual or aggregated) at the centre of their model. Even Agarwal and Alsaedi, who include the creator side, ultimately route their causal logic through the user's judgment. This creates a strong demand-side analytical orientation, where influence is primarily examined through audience effects rather than production-side operations. This focus provides substantial insight into persuasion, trust formation, behavioural responses, and audience vulnerability, while often giving less attention to coordinated production infrastructures or operational tradecraft.

The frameworks also share some conceptual aspects. They all treat source credibility (or trust) as a major variable, whether at the cognitive level (MRF, MRRM), the behavioural level (Geers et al.), the institutional level (Wardle and Derakhshan), the literacy level (Agarwal and Alsaedi; John and Tater), the operational level (Nimmo, treating credibility of a story's amplifiers as the trigger for breakout categories four and

five), the systemic level (Recuero, treating credibility as one of the dynamics produced by platform-actor interactions), or as a constitutive variable in audience segmentation (Zeng et al.). Furthermore, they all recognise context as a modulator. The MRF makes this most explicit through noise and instability, whereas Wardle and Derakhshan emphasise it through their phase model, Nimmo through the ecosystem metaphor, Agarwal and Alsaeedi through the historical/social/political milieu inherited from Karlova and Fisher, Recuero through the explicit “environment” dimension that includes both platform affordances and broader political/cultural conditions, and Zeng et al. through demographic and ideological variation across their four subgroups.

All nine frameworks also posit that intervention is possible and that behavioural outcomes can be shifted, but they differ in where they locate interventions (curriculum, platform design, policy, real-time response, system-level governance, audience-targeted health communication). Lastly, in seven out of nine frameworks, a clear separation of information, misinformation, and disinformation is made based on intent. Recuero does not separate information accordingly, arguing that intent is often difficult to ascertain and that the cleaner boundary lies between systemic effects and isolated content. Zeng et al. operationalise “misinformation” as a working category without insisting on the intent question, consistent with much of the health-communication literature.

The behavioural cluster is built from multiple strong domain-specific research traditions, while largely being content agnostic in architecture, even where it draws on specific cases: grounded in COVID-19 vaccine hesitancy (Zeng et al.); drawn from elections (Wardle and Derakhshan; Nimmo; Recuero); platform behaviour generally (Geers et al.; Recuero); and health science (MRF; MRRM; Agarwal and Alsaeedi). This creates a category highly useful for FIMI analysis, as the architectures do not need rebuilding for new topics, but whose generalisability depends on assumptions not uniformly validated across all FIMI contexts, as none of the nine have been stress-tested on the full range of FIMI subject matter (military, electoral, hybrid threat, foreign policy, etc.).

Another mutual aspect across the frameworks is their shared limitations, including not using FIMI as an operative term, their lack of international coverage, and weaknesses addressing attribution challenges. None of the nine frameworks uses FIMI as an operative term or engages with the foreign-state-actor framing that defines FIMI. Furthermore, they use different terminology and labels, which produce challenges for interoperability and coordination of frameworks, even just within the behavioural framework category.

Penultimately, the empirical material remains overwhelmingly Anglophone and Western, with the partial exception of Wardle and Derakhshan (European Council remit) as well as Recuero (Brazilian author with explicit attention to Global South contexts, citing Indian religious nationalism, Filipino networked propaganda, and Brazilian electoral cases). Zeng et al. even point to a likely sample bias in their data.

While this is a major limitation academically, as western-centric behavioural assumptions may not universally apply and cultural differences, as well as different political structures, should be considered, our project's orientation toward Europe and European policy making allows the frameworks to retain high practical utility for our analytical objectives. A further limitation is that eight of the nine address attribution of FIMI operators only obliquely or not at all, with Recuero being the partial exception, identifying state-sponsored operatives, political brokers, and influencer networks as systemic elements, but without offering analytical tools for attribution itself.

4.3.2 Contrasts and Divergences

A major contrast between the nine frameworks can be found in their basic unit of analysis. The MRF and MRRM frameworks focus on individuals' cognition by exploring what happens in the receiver's head. John and Tater frame the user as the object of literacy intervention, hence still focusing on the individual, but shifting more towards an educational lens. Geers et al. focus on the individual's behaviour on platforms by investigating what users do (follow, click, evaluate, share), not what they think. Agarwal and Alsaeedi widen the unit to include creators and the algorithmic infrastructure that connects them with receivers. Zeng et al. take a population-level segmentation approach with the unit being neither the individual receiver nor the operation, but the analytically constructed subgroup defined by levels of exposure and trust. Nimmo shifts away from the receivers' beliefs further by taking the influence operation itself as the unit, by using the campaign's spread as the measured outcome. Recuero instead takes the disinformation system (elements, environments, and dynamics) as a unit, with the receiver as one element among many. Wardle and Derakhshan take the phenomenon they call 'information disorder' as the object itself, with agents, messages, and interpreters as components.

Another point of divergence, previously hinted at, is in the framework's discipline and intellectual lineage. The MRF is rooted in Bayesian cognitive science, the MRRM in persuasion psychology and public relations, the Online Misinformation Engagement Framework in behavioural-science synthesis with platform-research orientation, both Agarwal and Alsaedis' as well as John and Taters framework in library and information science, Wardle and Derakhshan's framework in policy-oriented communication studies and journalism, Nimmo's framework in open-source analytical practice on influence operations, Recuero's framework in systems theory and platform studies, and Zeng et al.'s framework in health communication. These differences explain the varied use of vocabulary, implying a challenge for future interoperability.

The direction in which the analysis moves within the frameworks also differs. Seven of the frameworks investigate, given a message or an exposure pattern, what the receiver does, while the Breakout Scale asks the inverse (given an operation, how far has its message travelled?), and Recuero investigates the constitutive elements, their environment, and emergent dynamics. Within the receiver-facing majority, Zeng et al.

introduce a differing approach, shifting from message-effects to audience-segmentation logic (which subgroup of receivers a message is most likely to land with?).

The nine frameworks also differ in their treatment of intent. Whereas Wardle and Derakhshan place intent at the core of their typology, Nimmo treats influence operations as deliberately orchestrated, but does not model the operator's decision-making. Agarwal and Alsaeedi acknowledge creators' and spreaders' intentions as a variable but, following Karlova and Fisher, treat them as inherently opaque to the analyst. Recuero instead explicitly argues that intent is often impractical to ascertain in empirical research, and that the field should move past binary truth/falsehood and intent/no-intent distinctions toward the analysis of systemic dynamics. Another group of frameworks (The MRF, MRRM, Online Misinformation Engagement Framework, the Reframed Information Literacy Framework, and Zeng et al.'s framework) mostly refer to misinformation, therefore assuming no malicious intent. Most of these frameworks, however, claim to also apply to disinformation or information sharing broadly (Zmigrod et al., Amazeen, and Geers et al.), highlighting their usefulness to FIMI research.

The treatment of the supply/production side of FIMI is another point of divergence. Supply-side dynamics are essentially absent from the MRF, the MRRM, the Online Misinformation Engagement Framework, the John and Tater IL model, and Zeng et al., as none cover content production, infrastructure, coordination, attribution, or tradecraft. Opposed to that is Wardle and Derakhshan's framework, which includes an agent element, but does not model its operations in detail. Agarwal and Alsaeedi include creators, algorithms, and bots in the framework, but with limited specification of who the creators are or how they operate. Nimmo's framework is the only one whose analytical centre is the operation itself; nonetheless, it measures the spread and not the operator's production process. Recuero offers the most explicit supply-side architecture of the nine, as algorithmic amplification and ordinary users are named as system elements, and the platform is the primary environment. Nonetheless, it stays rather descriptive than analytically operationalised.

The treatment of platforms also varies across the frameworks, ranging from seeing the platform as part of the system (platform-constitutive), to them not being essential to the framework. Recuero is the only framework in which platforms are not merely a context or affordance set, but a structurally coupled system that co-evolves with disinformation. Two frameworks see platforms as a core structural element: Geers et al. make the platform stage structure the spine of the framework, while Agarwal and Alsaeedi treat platforms (and their algorithms and bots) as a moderating layer. Another group of frameworks is platform-aware but not platform-specific, meaning the architecture would survive a change of platform. This includes the MRF, the MRRM, Nimmo, and Zeng et al. The only two frameworks that see platforms as not essential are Wardle and Derakhshan, and John and Tater, as their architecture predates or abstracts from any particular technology.

The last major aspect the frameworks differ on is their orientation regarding interventions, ranging from remaining diagnostic with no specific intervention to specific measures, tools, or recommendations. The frameworks from Wardle and Derakhshan, Recuero, as well as the MRF, fall into the first category and remain diagnostic. The MRRM offers an empirical platform for testing interventions, while the Online Misinformation Engagement Framework offers a catalogue of interventions linked to behavioural stages. The framework of John and Tater, as well as partially the framework of Agarwal and Alsaedi, offer curricular and pedagogical interventions, whereas Zeng et al.'s framework explicit purpose is to align messaging strategies (emphasising efficacy, safety, or institutional transparency) with subgroup profiles. The Breakout Scale can instead be seen as a triage and prioritisation tool.

4.3.3 Coverage of the FIMI Response Process Phases

Behavioural frameworks span all phases of the FIMI Response Process except the *Evaluation and Learning* Phase (See Figure 11 below). Behavioural frameworks are built as exploratory lenses of human behaviour and are hence naturally broader, covering more areas of our process model. As mentioned in the Differences section, the frameworks also originate from alternative areas of research and have differing heritages, as well as no unified mission objective. Behavioural frameworks rarely include system-level feedback loops or institutional learning mechanisms; instead, existing evaluation components are usually limited to measuring audience or subgroup responses, rather than assessing institutional effectiveness. The interdisciplinary nature of this category and the lack of standardised datasets and cumulative evaluation structures across studies may also contribute to this gap. Furthermore, this finding may reflect a broader underdevelopment of evaluation and learning within FIMI research.

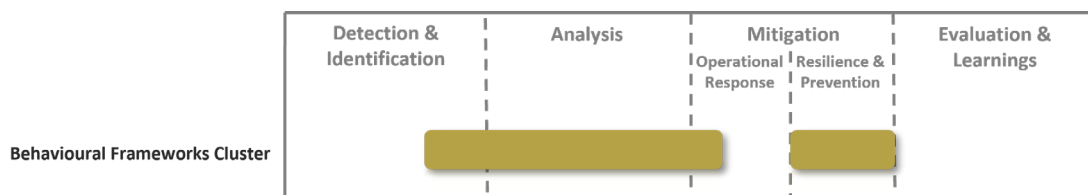


Figure 11: Behavioural Frameworks Cluster Coverage of FIMI Response Process Phases.

It should be noted that *Detection and Identification*, as well as the *Mitigation - Operational Response* section, are both weakly served by the behavioural frameworks cluster. This could be due to a conceptual difference, as behavioural frameworks are not designed for real-time identification of threats. Instead of monitoring technical extraction and attribution pipelines, it focuses on cognition, exposure effects, and audience interpretation. The weak contribution of this cluster to the operational response could be partly influenced by the lack of decision-making authority and system integration of behavioural frameworks. It could also be influenced by the underlying structural gap between understanding behaviours and executing interventions, as the frameworks tend to stop after investigating audiences and

behaviours, therefore not taking it a step further to investigate what operators could do next.

The strength of this framework cluster lies in the focus upon phases of *Analysis* and *Mitigation - Resilience and Prevention*, which are both well served. Behavioural and communication research fundamentally aims to understand human behaviour, interpretation, and interaction, making analytical work central to the category, which can be seen reflected in the coverage of the FIMI Response Process in this cluster. Additionally, the focus on individuals, audiences, and communication dynamics naturally aligns with analytical and interpretive phases. There is also a strong methodological orientation toward observation, interpretation, and behavioural analysis within this cluster.

Behavioural frameworks are strongly rooted in psychological and communication theories, naturally aligning with aspects relevant to long-term resilience and prevention, like nudging, media literacy, and persuasion resistance. Insights from the literacy and educational context also naturally align with long-term measures such as information literacy.

4.3.4 Application of Frameworks in Practice

In the expert interviews, only one behavioural framework was mentioned, namely the Breakout Scale. The framework was, however, mentioned by only two out of the 17 experts interviewed, which may reflect the practical usefulness of the Breakout Scale compared to other behavioural frameworks. Both experts described using it to assess impact and to determine whether intervention was necessary or whether ignoring the activity would be the more effective response. This use is particularly interesting because it shows a behavioural framework being applied in an operational decision-making context, rather than solely for understanding audience behaviour. At the same time, interview results indicate that there is still considerable untapped potential for integrating behavioural frameworks into both FIMI research and practice. One expert explicitly stated that behavioural perspectives should receive greater attention in the future and that they are working towards integrating more, indicating awareness of this gap among practitioners themselves. Another expert directly noted that behavioural frameworks fall outside their area of expertise. This reflects a challenge of FIMI research, as due to its interdisciplinary nature, lots of different research areas need to be covered. It also might indicate a broader disconnect between psychological research on the one hand and FIMI research and practice on the other. Despite the central role that human behaviour plays in information influence activities, behavioural expertise does not yet appear to be systematically integrated into all parts of the field.

Notably, none of the behavioural frameworks focused on understanding and analysing audience behaviour were mentioned. This could suggest that practitioners currently prioritise frameworks that support decision-making and response over those that

explain underlying behavioural mechanisms. As a result, opportunities to better understand why audiences react to information in certain ways may remain underutilised. Similarly, none of the frameworks focused on long-term prevention and resilience-building were referenced. This may indicate that current practice remains more focused on responding to ongoing incidents than on systematically reducing future susceptibility.

Taken together, these findings suggest that behavioural perspectives remain underrepresented in current FIMI practice despite their potential relevance, pointing towards an opportunity for future integrative approaches to more systematically incorporate communication and psychological research. Doing so could strengthen both the understanding of how influence affects audiences and the development of more sustainable preventive and resilience-building measures.

4.3.5 Contributions to a Comprehensive Analytical Approach

Behavioural frameworks do not form a unified analytical unit, but rather a multi-disciplinary toolkit of partially overlapping explanatory lenses that must be externally integrated to function as a coherent system. As such, Wardle and Derakhshan can be seen as suppling the vocabulary (the mis/dis/mal trichotomy and the agent/message/interpreter separation), while Recuero contributes the systemic frame - meaning the recognition that disinformation is structurally coupled with platforms, and the three-dimensional analytical grid (elements, environment, dynamics) that lets a researcher locate any of the other frameworks within a wider field of relations. Nimmo provides the outcome metric against which the spread of an operation can be tracked once it is in the field.

The MRF and the MRRM explain the interpreter by investigating how prior belief, source credibility, context, and recognition shape what a receiver does with a message. Zeng et al. could build on that by adding a population-level layer to the interpreter where receivers are not interchangeable but sorted into stable subgroups defined by exposure and trust, with different predictors of behaviour operating in each. Agarwal and Alsaeedi connect receivers back to creators and platform infrastructure, and add a remediation layer that offers countermeasures. Regarding countermeasures, John and Tater operationalise one, information literacy, by listing what an information-literacy curriculum needs to include to address mis/disinformation specifically. Following this order, the cluster could map loosely on a workflow starting with establishing vocabulary and systemic frame, explanation of receiver behaviour, segmentation of audiences, platform-stage interventions, and finally LIS-grounded remediation. Recuero's contribution is structural rather than sequential, as it provides a container within which the other frameworks can be located, and Nimmo's framework can be used for measuring the operational spread throughout the process.

This workflow and integration of the behavioural frameworks are hindered by existing tensions between the frameworks and major gaps within the cluster. The frameworks operate on different levels, which creating tensions when integrating them. The cognitive and literacy frameworks model individual psychology and competence, whereas Zeng et al. operate at the aggregate level of subgroups and Recuero operates at the level of the system as a whole.

A point of tension is deciding between utilizing open-ended hypotheses versus fixed taxonomies. The MRF and the MRRM are deliberately hypothesis-generative and expect to evolve under empirical pressure while Wardle and Derakhshan, and Nimmo offer relatively fixed structures. Recuero deliberately resists fixed taxonomies, treating disinformation as adaptive and emergent. A comprehensive framework will need to decide which mode it operates in, or how it combines them.

Frameworks use different framing for the receiver, as the cognitive and literacy frameworks tend to frame the receiver as someone whose vulnerabilities need to be reduced, causing further tensions. Recuero explicitly reframes receivers as participants in the construction and circulation of disinformation. Whereas Zeng et al. occupy a middle position as their subgroups are not framed as passive, but as carrying different psychological profiles that shape their relationship with the information environment. This directly connects to a gap in the behavioural cluster, namely the lack of supply-side analysis, which is only partially addressed by Recuero, by naming operatives, brokers, influencer networks, and coordinated inauthentic behaviour as systemic elements.

Another aspect is the diagnostic versus operational response, showcasing the lack of operational aspects in this cluster. Most of the cluster remains diagnostic; only Nimmo addresses an operational question (prioritisation). Zeng et al. offer a partial operational pathway by linking subgroup profiles to messaging strategies, but at the level of communication design rather than incident response. A major gap in this cluster is a lack of consistent terminology. The cognitive frameworks (MRF, MRRM), the LIS frameworks (Agarwal and Alsaeedi; John and Tater), and the segmentation framework (Zeng et al.) work primarily with 'misinformation', which treats accuracy as the key variable. The operation-level frameworks (Wardle and Derakhshan; Nimmo) work with intent and coordination. Recuero introduces 'disinformation as a system' as an additional framing that explicitly tries to move past both, but in doing so, further widens the terminological field rather than resolving it.

Already alluded to is that the clusters do not address attribution. Most only acknowledging it as a question, as Recuero does. The cross-platform dynamics are improving as highlighted by Geers et al., the Breakout Scale, and Recuero, but still remain a gap, as the cognitive, literacy, and segmentation frameworks treat the platform layer as background instead of integrating it. Lastly, the lack of evaluation emerges as

a major gap for the behavioural cluster, as none of the frameworks investigate how to evaluate the effectiveness of a counter-FIMI programme over time.

Nonetheless, each of the frameworks in the behavioural cluster can offer important aspects that should be considered when building an integrative approach to FIMI:

- Recuero offers multilingual and multicultural variation, contributes to the containing architecture of a relational three-dimensional view (elements/environment/dynamics) within which the other frameworks can be situated, and brings the analytical perspective of the platform as a co-evolving system. The last aspect is especially important in the FIMI context, as platform incentives and affordances are part of what makes coordinated information manipulation possible at scale.
- Wardle and Derakhshan contribute the shared vocabulary and the agent/message/interpreter separation, which can help in being intelligible across disciplines.
- The MRF contributes a mechanism for explaining why context matters: the noise/instability distinction is the most precise account in the cluster of how environmental conditions shape receptivity, and it has direct relevance to FIMI scenarios involving crisis, war, or socio-political turbulence.
- The MRRM contributes a developmental account of how recognition works, which is essential if interventions aim to help audiences identify FIMI rather than merely resist it.
- The Online Misinformation Engagement Framework contributes a structured understanding of the different stages through which users engage with misinformation, allowing interventions to be targeted at the specific point where they are likely to have the greatest effect.
- The Disinformation Behaviour Framework contributes to the integration of creator, infrastructure, and receiver in a single causal picture and the explicit remediation layer.
- The Reframed Information Literacy Framework contributes a practitioner-ready specification of what an information-literacy programme needs to cover if it is to address mis-/disinformation rather than information in general.
- Zeng et al.'s segmentation framework contributes a population-level analytical layer that the rest of the cluster largely lacks, while also producing different recipient profiles (vulnerable, resistant, accepting), which is directly useful for FIMI intervention design when deciding which audiences to address and how.

Collectively, the nine frameworks provide the building blocks for the interpreter, systemic, and spread dimensions of an integrative model, while leaving the operator, response, and learning dimensions insufficiently covered.

4.4 Comparative Analysis of the Operational Frameworks Cluster

Operational frameworks support practitioners in their day-to-day work by structuring, coordinating, and guiding practical responses to information manipulation and interference. The frameworks under this category focus on implementation and action, rather than on strategic oversight or theoretical research in technical or behavioural areas. Frameworks in the cluster are hence particularly relevant for actors involved in understanding, acting on, reporting, or coordinating responses to information manipulation. The importance of this cluster is highlighted by the fact that FIMI is not only an analytical or conceptual challenge but also a practical and operational one that requires coordinated real-world responses. Operational frameworks help to facilitate action through analysis, direct procedures or workflows, and hence represent the most practice-oriented cluster. Operational frameworks exist because understanding FIMI alone is insufficient; practitioners also need structures for analysis, coordination, implementation, communication, and response.

4.4.1 Similarities and Shared Ground

All **five** frameworks are practitioner-oriented, recognise their environment, and case-based. The frameworks are oriented to the daily work of practitioners (e.g., analysts, investigators, communications officers), rather than to academic theorising or policy formulation. Therefore, they function as tools for action rather than tools for explaining how information influences in a theoretical sense. Due to this shared practical purpose, they tend to prioritise actionable outputs, structured analytical procedures, and decision support.

Name of Framework	Authors	Source reference
Countering information influence activities: A handbook for communications officers	Swedish Civil Defence and Resilience Agency	https://mpf.se/psychological-defence-agency/publications/archive/2024-03-01-countering-information-influence-activities
Methodology for online open source investigations into incidents taking place in Ukraine since 24 February 2022	Bellingcat	https://www.bellingcat.com/app/uploads/2022/12/JA-Manual-for-PUBLICATION.pdf
SCAME	U.S. Army	U.S. Army (December 2023) FM 3-05.301 - Psychological Operations Tactics, Techniques and Procedures, Chapter 11
The Diamond Model for Influence Operations Analysis	Recorded Future	https://assets.recordedfuture.com/Whitepapers/diamond-model-influence-operations-analysis.pdf

The IIO Attribution Framework	B. Palmertz, E. Isaksson, and J. Pamment	https://adacio.eu/a-framework-for-attribution-of-information-influence-operations
-------------------------------	--	---

Figure 12: Table of analysed operational approach frameworks

Another shared aspect is their acceptance that information manipulation operates in an environment characterised by speed, scale, multi-platform diffusion, and cross-jurisdictional reach. Even SCAME, the oldest in the cluster, rooted in pre-social-media doctrine, retains analytical traction because its sequential logic does not depend on the medium. The persistence of these frameworks across changing media environments may suggest that many operational analytical tasks (e.g., identifying actors, analysing content, assessing audiences, and evaluating effects) remain relatively stable even when technologies change.

The frameworks also all start with a case, named according to the unit they use (e.g., campaign, incident, observed influence activity), instead of starting from theoretical premises about the receiver or the information environment.

Another shared aspect of the operational frameworks is their conceptual moves: they use the same building blocks, highlight some form of taxonomy, treat the audience as multi-layered, and point out that analysis is not the final aspect of their process. While the frameworks are not necessarily built around the exact same model, they repeatedly return to similar analytical dimensions (e.g., actor, audience, medium) with varied foci (like the Diamond Model, which arranges them around a narrative core). The convergence of these elements suggests broad operational consensus regarding what information practitioners need to identify and assess when analysing influence activities.

All five frameworks also recognise some form of source taxonomy or sourcing discipline, while their taxonomies vary. The IIO Attribution Framework is the most explicit (open/proprietary/classified); Bellingcat-GLAN separates examinable from descriptive content; SCAME classifies as white/grey/black; the Diamond Model gathers source data into the infrastructure corner; and the MSB framework embeds source criticism in its identification routine. They all emphasise systematic evaluation of information sources and evidence, highlighting how the frameworks are not only interpretive but also evidentiary.

Furthermore, they also all treat the audience as a multi-layered analytical object rather than as a single homogeneous group. This implies that operational frameworks are aware that influence depends on audience dynamics, not only on messaging. Lastly, all five assume that the intervention or action follows from analysis, and none treats the analysis as a terminal output, further highlighting the action-oriented character of the operational frameworks.

None of the operational frameworks is dedicated to a single subject domain; instead, empirical illustrations are drawn from cross-border information operations (IIO

Attribution Framework: Doppelgänger, Paperwall, LVU); state-sponsored campaigns on health and elections (Diamond Model: COVID-19 origin narratives, US election operations); armed conflict and international humanitarian law (Bellingcat-GLAN: Ukraine post-February 2022); historic and contemporary propaganda (SCAME); and broad-spectrum civilian influence activity (MSB). While some domain-specific anchoring exists in Bellingcat-GLAN (Ukraine), SCAME (military counterpropaganda), and MSB (Swedish public administration), the other two remain domain-agnostic. The Operational cluster is largely content-agnostic in architecture, even when it draws on specific cases, allowing its application for new domains. However, it should be noted that none has been stress-tested against the full FIMI subject range (e.g., military, electoral, hybrid, foreign policy, health, climate, migration).

The last shared aspect of the operational frameworks is their limitations, consisting of terminology fragmentation, lack of evaluation, no cross-campaign or long-term actor analysis, and weak integration into broader policy-making. Only the IIO Attribution Framework uses FIMI as its operative term, whereas all the other frameworks use different terms from each other (e.g., MSB uses “information influence activities”). Practitioners may analyse the same aspect but describe it differently, revealing that the operational FIMI field is still conceptually fragmented. In conclusion, this may indicate that operational frameworks struggle with interoperability and coordination due to inconsistent terminology.

Furthermore, they lack evaluation with SCAME's Effect being the most explicit treatment of effectiveness in the cluster, but the source manual acknowledges it as the most difficult and least operationalised. Therefore, the importance of evaluation and learning is recognised but still lacks operationalisation, which may limit institutional adaptation and cumulative improvement over time. Operational frameworks seem to prioritise immediate response and coordination over systematic long-term evaluation. Another shared limitation is that none cover cross-campaigns or long-term actor analysis. Instead, the unit is always the immediate operational object (campaign, incident, product, observation), not the threat actor's behaviour over time, which limits the transferability of findings between events. Their event-centred orientation may limit strategic continuity and cumulative threat understanding, while findings may remain fragmented across isolated operational cases.

Lastly, none address the integration of analytical findings into broader policy or strategic decision-making in detail. Operational frameworks focus heavily on producing structured analytical outputs but less on translating them into policy and governance processes, which could hinder the implementation of operational findings and institutional policy adaptation.

4.4.2 Contrasts and Divergences

A major aspect in which the five frameworks diverge is in their unit of analysis. The IIO Attribution Framework and the Diamond Model see the campaign itself as the unit, while the IIO Attribution Framework organises it around evidence and the Diamond Model around narrative. Bellingcat-GLANs framework treats the discrete incident as the unit, whereas SCAMEs' unit is the individual propaganda product. In the MSBs framework, the unit is the public-sector organisation reading its information environment for signs of influence activity and choosing how to respond. It should be noted that none of the frameworks operates at the level of the threat actor's portfolio over time. Operational frameworks differ substantially in what they consider the primary object of operational analysis. This could be explained by the varied needs of their intended users and institutional environments.

The frameworks also treat evidence, source access, and attribution differently. Operational frameworks differ in their assumptions about access to evidence. The IIO Attribution Framework is the only one that builds the open/proprietary/classified distinction into the architecture of the model, whereas the Diamond Model, Bellingcat-GLAN, and SCAME all assume an analyst working predominantly with open-source material. The MSBs framework does not address source tiering. An integrative model that aspires to coordinate across these actor types will need to make the open/proprietary/classified distinction explicit. Attribution varies similarly across the frameworks, ranging from central (IIO Attribution Framework and the Diamond Model) to indirect for case-building purposes (Bellingcat-GLAN), implicit (SCAME), and absent (MSB framework). This highlights that attribution seems to play a varied role in operational response, as they do not always depend on attribution. Some frameworks treat attribution as the primary analytical objective, while others treat it as optional or downstream. These divergences in evidence, source access, and attribution reflect differing operational priorities and institutional mandates.

Another point of divergence is the framework's treatment of intent, as the operational frameworks differ substantially in how they conceptualise and operationalise intent. Intent may function as a conclusion to be reached (SCAME), a dimension to be assessed (IIO Framework), placed at the centre of the model as narrative (Diamond Model), treated as a downstream legal question (Bellingcat-GLANs framework), or treated as a diagnostic test (MSBs framework). Again, these findings reflect the varying institutional and disciplinary priorities of the environment of the corresponding framework.

The treatment of legal and ethical considerations also varies vastly across the operational frameworks, depending heavily on the institutional setting and intended operational use. The legal-evidentiary purpose shapes the entire architecture of the Bellingcat-GLAN framework. The IIO Attribution Framework treats legal and ethical assessment as a permanent column in the matrix, qualitatively distinct from the evidentiary columns and likely to be handled by different teams or functions within an

organisation. On the other hand, SCAME does not engage with ethics in the operational sense; the underlying doctrine assumes a military-legal frame in which the analyst's role is bounded by the chain of command and rules of engagement. The Diamond Model does not explicitly address legal or ethical considerations, although acknowledging that case material may include sensitive content. The MSB framework does not address legal or ethical questions, reflecting its addressee's distance from attribution or litigation, which limits applicability when the communicator's response has reputational or legal consequences for third parties.

Another difference regarding the frameworks is their treatment of countermeasures and response, split into implicit and explicit. The IIO Attribution Framework is implicit as it treats public attribution itself as a form of action and discusses pre-bunking, sanctions, and other downstream uses of attribution in the surrounding chapters. Another implicit approach can be found in the Diamond Model, which treats the production of clear analytical understanding as the response, with downstream defensive action left to incident responders and decision-makers. Bellingcat-GLANs framework also uses an implicit approach that treats the construction of a legally admissible record as the action, supporting accountability rather than mitigation. SCAME sits within a broader counterpropaganda doctrine and uses an explicit approach, in which the analytical output feeds directly into counter-products and command information programmes. Another explicit approach can be seen in the MSBs framework, which devotes a full Act phase to choosing the appropriate response, including fact-checking, social media engagement and the explicit choice between responding and not responding.

While all frameworks are operational, they originate from different sub-disciplines and have a varied intellectual lineage. The IIO Attribution Framework, Diamond Model, Bellingcat-GLAN framework originates from open-source investigation and threat-intelligence practice, while SCAME comes from a military operations doctrine background and the MSB framework from public-sector crisis communication and civil contingencies. This divergence in sub-disciplines and institutional focus could explain the varied taxonomies and methodologies, as where disciplines differ, so does the language used.

The Operational frameworks also differ in how they structure practitioner guidance and analytical workflows, once again reflecting differing assumptions about operational practice, usability and decision-making environments. The IIO Attribution Framework and SCAME both present their core as a structured grid against which findings are recorded, whereas the Diamond Model offers a single visual on which an analyst plots a campaign. Bellingcat-GLANs and MSBs framework both present sequential phases with explicit instructions, though Bellingcat-GLANs is far more detailed and granular, while MSBs also functions as a catalogue of influence techniques and combinations, which the others do not.

Overall, the main differences seem to be trade-off decisions based on the frameworks' varying organisational and sub-disciplinary contexts, as well as the different ways in which they serve fundamentally different practitioner ecosystems.

4.4.3 Coverage of the FIMI Response Process Phases

Operational Frameworks span all but two phases of the FIMI Response Process, not covering the *Mitigation - Resilience and Prevention* and the *Evaluation and Learning* Phase (see Figure 13 below). The frameworks mostly depend on prior detection, focus on structured analysis, and partially support response, while leaving long-term aspects like prevention and evaluation comparatively underdeveloped.

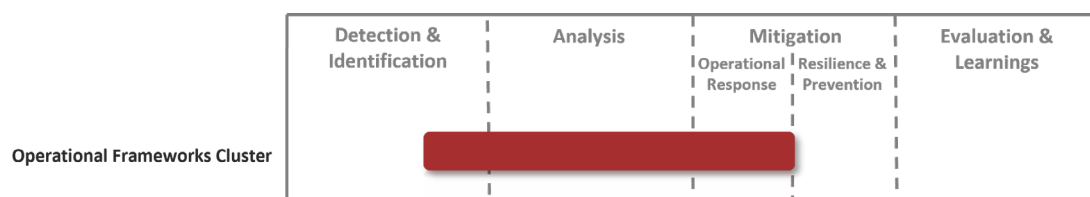


Figure 13: Operational Frameworks Cluster Coverage of FIMI Response Process Phases.

The first Phase of *Detection and Identification* is weakly covered. The MSBs framework is built around it, and SCAME begins with the identification of a propaganda product. The other three frameworks assume that detection has already taken place, therefore assuming it as a prerequisite. This dependency on detection having taken place could be due to operational frameworks being designed for practitioners who are already aware that something has happened and are hence focused on the next step, analysing and generating decision support. Operational frameworks are structurally dependent on upstream detection systems, making them post-detection frameworks rather than full-cycle systems.

These results complement the finding that the *Analysis* Phase is the most densely served phase, with all five frameworks providing analytical work. All operational frameworks aim to reduce uncertainty around an identified case or situation (each dependent on its unit of analysis). Analysis is the prerequisite for the following steps of mitigation and action; therefore, it is also an important component of operational frameworks. It is the enabling condition for all following steps, making it a necessary operational function rather than a separate activity.

The phase *Mitigation - Operational Response* is served principally by MSB's Act phase and indirectly by SCAME through its counterpropaganda orientation. However, operational frameworks recognise a major challenge, that the appropriate response often depends on organisational mandates, legal authorities, and contextual considerations that cannot easily be standardised. Furthermore, operational frameworks support decision-making but do not fully include decision authority, reflecting that response actions are institutionally and contextually contingent. Lastly, it should be considered that the frameworks define what they count as action

differently, as explored in the difference section, which could also affect this evaluation.

Mitigation - Resilience and Prevention is among the least served phases by operational frameworks. MSB's Prepare phase is the only direct treatment, addressing awareness-raising, trust-building, and risk assessment within an organisation, whereas the other frameworks do not include prevention or resilience building. This could be because operational frameworks tend to be incident-centred rather than taking a long-term approach or being system-bound. Therefore, the weak coverage of prevention is not a gap of interest but a consequence of its scope. Operational frameworks are incident-bound and become relevant when a problem emerges, whereas resilience frameworks aim to reduce vulnerability before problems arise.

The last phase, *Evaluation and Learning*, is also weakly served. SCAME's Effect step is the most explicit treatment, but the source manual itself acknowledges it as the most difficult to operationalise. This difficulty could indicate an overall problem with integrating this aspect into frameworks. The operational frameworks with their incident-focus are designed to solve current problems, while feedback or adaptation after is not considered, as the category seems to prioritise operational execution over systematic reflection, which may explain why evaluation remains weakly developed. The limited attention given to *Evaluation and Learning* is relevant because these activities are critical for long-term adaptation and effectiveness. Frameworks addressing this gap could help transform operational responses from isolated incidents into cumulative learning processes, enabling organisations to continuously improve their capabilities, identify effective interventions, and adapt to evolving FIMI threats.

4.4.4 Application of Frameworks in Practice

Results from the expert interviews indicate that when it comes to operational, action-focused frameworks, organisations are the most reliant on their own specific internal frameworks and approaches. The only specific framework mentioned in the interviews was the IIO Attribution framework, noted by one expert, who detailed that they utilise it as part of a larger internal workflow. Operational work appears highly context-dependent and reliant on practitioner experience, institutional knowledge, and established routines. This could be due to organisations functioning and taking action differently, even defining the action they take differently. Therefore, processes and workflow might be built directly for the organisation and built on the practitioners, researchers, and experts' needs and experiences. Therefore, organisations may gradually develop internal approaches that better reflect their own needs than externally developed frameworks, which explains why experts referred to internal workflows rather than published frameworks. Furthermore, operational activities might be hard to standardise as the FIMI environment is fast-paced and developing. Internal processes might be faster to adapt and adjust to changes than formal frameworks. While frameworks provide useful concepts and structures, organisations have specific needs

and procedures. Therefore, operational frameworks might be best used as modular components that fit within the individual organisation's internal processes.

4.4.5 Contributions to a Comprehensive Analytical Approach

As noted in the analysis of differences and the coverage of the FIMI Response process, the frameworks have different strengths and weaknesses resulting from trade-off decisions, while no single framework covers the entire operational process. The MSBs framework provides a frontend aspect, meaning that the identification of strategic narratives, target groups, and techniques produces the initial signal that an influence activity is in play and decides what kind of analytical attention it warrants. SCAME allows for a more detailed breakdown. Once an item of content has been identified as relevant, SCAME's sequential five-step analysis produces a structured account of what it is, who made it, who it is for, how it travels, and what effect it appears to have. Bellingcat-GLANs framework supplies evidentiary discipline through its eight-phase methodology, which offers procedural rigour for the open-source aspects, complementing a part of the IIO Attribution Frameworks matrix. The Diamond model can provide fusion and a visual aspect, as its invitation to layer other frameworks onto its four corners is the most interoperable disposition in the cluster, and it has been used in practice to host classifications such as DISARM, AMITT, and the kill chain models. The IIO Attribution Framework contributes the attribution basis, as it is the only one of the five offering an explicit, structured account of how findings translate into a defensible attribution conclusion with calibrated confidence. This proposed workflow should be seen as an illustration rather than a conclusion and is only one possible reading of their interactions. Additionally, some aspects of the frameworks could be combined and strengthen the frameworks, like the MSB's six clusters of influence techniques could populate the contextual and behavioural evidence columns of the IIO Attribution Framework, SCAME's audience taxonomy could populate the Diamond Model's audience corner, and Bellingcat-GLAN's preservation tooling could populate the procedural layer that the IIO Attribution Framework leaves implicit.

As hinted at, there are also challenges that come with a possible integration of the operational frameworks due to tensions between them, as well as major gaps in the cluster. Some of these tensions arise from a trade-off situation rather than conflict, reflecting differing goals and institutional contexts. One of these trade-offs can be found in the evidentiary depth as opposed to the operational speed. Bellingcat-GLAN sets the highest evidentiary bar in the cluster and assumes investigators have weeks or months, while MSB assumes hours, and the Diamond Model and the IIO Attribution Framework sit between the two extremes. An integrative model will need to decide whether to optimise for evidentiary strength or speed or try to find a middle ground. Another trade-off decision is between a fixed versus a modular structure of the frameworks. SCAME and Bellingcat-GLAN present essentially fixed procedures, whereas the IIO Attribution Framework is explicitly modular, the Diamond Model is

explicitly layerable, and the MSBs sit between these poles. An integrative framework will need to decide where on this spectrum it wishes to sit.

The next tension of the five frameworks points to a major gap in the operational cluster, namely the taxonomy gap, meaning that the frameworks' definitions differ and some of the frameworks are doctrine-bound (e.g., SCAME) while others remain mostly doctrine-free. This can lead to challenges when applying frameworks like SCAME to other contexts (e.g., civilian contexts).

Two other tensions between the frameworks relate to problems with cross-actor coordination, another gap in the operational cluster. The frameworks vary in single-organisation focus (MSB and SCAME) and cross-organisation use (IIO Attribution Framework). The Diamond Model and Bellingcat-GLAN sit between these poles. Furthermore, the frameworks differ in attribution-oriented (IIO Attribution Framework, the Diamond Model and Bellingcat-GLAN) versus response-oriented framing (MSB). The chain connecting attribution-grade analysis to communicator-grade response is not specified by any of the five frameworks. A practitioner who learns from an analyst that a foreign actor is amplifying a particular narrative is not told by any of the frameworks what to communicate to which audience in a certain timeframe. These tensions highlight the overarching gap in coordination across institutional actors with different data access, which is only addressed by the IIO Attribution Framework. Further coordination is needed across the boundaries between investigative findings (Bellingcat-GLAN), attribution conclusions (IIO Attribution Framework), product analysis (SCAME) and communications response (MSB).

Another major gap, already mentioned, is that none of the frameworks explores how to evaluate the effectiveness of a counter-FIMI programme over time. Another gap can be found in the lack of addressing uncertainty when communicating insights. Only the IIO Attribution Framework engages systematically with confidence intervals, probability yardsticks and the language of estimative probability. The other four leave the communication of uncertainty underspecified, which has direct consequences when their outputs are consumed downstream by communicators or decision-makers.

Another gap is the lack of consideration of generative AI and synthetic media, which only the IIO Attribution Framework (engaging substantively with the implications of generative AI for both threat actors and defenders) and the MSB (exploring deepfakes as a technique) address. Operational frameworks also do not explore theories of why audiences are susceptible to particular narratives.

None of the frameworks explores the threat actor's behaviour across multiple campaigns over years, hence highlighting a major limitation of the operational cluster. Lastly, an anglophone bias in the empirical material is evident across the cluster with the MSBs framework and its context of Sweden and Bellingcat-GLAN's Russian-and-Ukrainian working environment as partial exceptions. The impact of the limitations is

reduced by the European focus of this project. Nevertheless, the dominance of Western case material may constrain the framework's ability to capture the full diversity of FIMI activity. Expanding the operational frameworks to different linguistic, cultural, and geopolitical contexts (e.g., including threat actors such as China) could strengthen adaptability and long-term relevance.

Nonetheless, each of the frameworks in the operational cluster can offer important aspects that should be considered when building an integrative approach to FIMI:

- The IIO Attribution Framework contributes an evidentiary architecture through the structured separation of evidence categories and source tiers, the supporting legal-ethical column, the confidence-calibration discipline, and the modularity that allows organisations with different access to find a workable subset of the matrix.
- The Diamond Model contributes interoperability and visualisation with its invitation to layer other frameworks onto its four corners, offering a structural model for how an integrative framework can host more specialised tools without subordinating them. The narrative-at-the-centre design also offers a way to keep the cognitive object in view while reasoning about technical and social-political elements.
- Bellingcat-GLANs framework contributes procedural discipline for open-source investigative work through the categorisation of examinable and descriptive content, preservation tooling, citation conventions, investigator-safety routines, and legal-admissibility orientation.
- SCAME offers a sequential analytical logic for the analysis of an individual product or message through its five-step sequence, which is robust to changes in medium and has been a stable reference point in psychological-operations doctrine for two decades. Its audience taxonomy in particular (apparent/intermediate/unintended/ultimate) remains analytically useful well beyond the doctrinal frame in which it was developed.
- MSBs framework contributes the identification-and-response routine for non-specialist practitioners as it includes the strategic-narratives lens, the target-group taxonomy, the typology of six clusters of influence techniques, and the explicit Prepare-Act-Learn cycle.

The five operational frameworks supply the building blocks for an operational chain that runs from identification through analysis and attribution to response. They contribute evidentiary rigour, analytical structure, investigative procedures, communication and coordination mechanisms, and response-oriented decision support. While an integrative approach will have to consider and decide on explored trade-off decisions, the frameworks taken together could bridge the gap between understanding an FIMI incident and acting upon it.

4.5 Comparative Analysis of the Strategic Frameworks Cluster

The last framework category is the strategic framework cluster, which focuses on frameworks addressing high-level planning, policymaking, coordination, and top-level conceptualization of responses, distinct from the planning or execution of concrete actions. This focus on coordination is important, as strategic frameworks can help structure decisions on who should act, how individuals coordinate, and which response is appropriate. It could also help coordination across multiple actors, which is essential for fighting FIMI, and a major challenge noted in earlier clusters. The frameworks of this cluster can also aid in defining objectives and creating understanding, which is essential as FIMI responses need to be directed. As mentioned in previous cluster analyses, FIMI is an interdisciplinary challenge that in turn requires a diverse and interdisciplinary response. This cluster will further expand topical coverage to politics, policy, and governance. Strategic frameworks can help connect analysis to governance and policymaking, as they are conceptually closest. This is especially crucial as insights from analysis need to be communicated to the government and policy-makers to create change. It also broadens the scope from individuals and technical systems to a county's government or even on an international scale.

After our unfiltered collection, we amassed 41 supposed strategic frameworks, which highlights the popularity of strategic frameworks, further highlighting their importance. In our first filtering step, we only kept frameworks that fit our working definition, reducing our dataset to 19 frameworks. Therefore, more than half of the collected supposed frameworks were sorted out, which could indicate that there is a difference in understanding of what is considered a framework within the field. Instead of fitting our framework definition, we found either lists that lacked connections or aspects that veered into models through predictions. After the second filtering process, removing frameworks that lacked relevance to FIMI, 15 frameworks remained.

Name of Framework	Authors	Source reference
A framework for analysing disinformation laws and policies from a human rights perspective	Global Partners Digital	https://www.gp-digital.org/wp-content/uploads/2023/06/A-framework-for-analysing-disinformation-laws-and-policies-from-a-human-rights-perspective.pdf
ABCDE	J. Pamment, Carnegie Endowment for International Peace	https://carnegieendowment.org/research/2024/02/the-eus-role-in-fighting-disinformation-crafting-a-disinformation-framework
ALERT	K. C. Desouza, A. Ahmad, H. Naseer, M. Sharma	https://doi.org/10.1016/j.cose.2019.101606
Analytical framework for countering hybrid threats and foreign influence and interference	B. Palmertz, M. Weissmann, N. Nilsson & J. Engval	https://mpf.se/download/18.7bb35f1818ec664dc7a4664/1712909687772/buidling-resilience-and-psychological-defence.pdf

Conceptual Framework of Health Misinformation During the COVID-19 Pandemic	J. Alvarez-Galvez, J. Carretero-Bravo, C. Lagares-Franco, B. Ramos-Fiol, E. Ortega-Martin	https://doi.org/10.2196/62693
Department of Justice Framework to Counter Malign Foreign Influence Operations	US Department of Justice	https://www.justice.gov/archives/ag/page/file/1076696/dl?inline=
Deterrence framework	Sean Monaghan (Hybrid CoE)	https://www.hybridcoe.fi/publications/hybrid-coe-paper-12-deterring-hybrid-threats-towards-a-fifth-wave-of-deterrence-theory-and-practice/
DigComp 2.2: The Digital Competence Framework for Citizens	JRC of the European Commission	https://publications.jrc.ec.europa.eu/repository/handle/JRC128415
DISRUPT framework	Alliance4Europe	https://alliance4europe.eu/counter-disinfo/disruption-framework
RESIST (3)	UK Government, Government Communication Service (GCS); developed James Pamment	https://www.communications.gov.uk/wp-content/uploads/2025/10/RESIST-FINAL.pdf
Societal Resilience to Disinformation Framework	B. Rød, C. Pursiainen, N. Eklund	https://doi.org/10.1007/s41125-025-00105-4
Strategic framework for counteraction the impact of online misinformation on ideological polarisation and institutional distrust	S. K. Vadapalli, S. Doyal, F. J. Vanheusden, J. F. Binder, D. J. Kuss	https://doi.org/10.1007/s43545-026-01349-0
Strategic Framework to Counter Hybrid Action	T. Sweijjs, S. Zilincik, F. Bekkers & R. Meessen	https://euhybnet.eu/wp-content/uploads/2021/01/Framework-for-Cross-Domain-Strategies-against-Hybrid-Threats.pdf
The Counter-Hybrid Response Framework	M. Bertolini, R. Minicozzi and T. Sweijjs	https://hcss.nl/wp-content/uploads/2023/04/Guidelines-for-the-Deterrence-of-Hybrid-Threats-HCSS-2023.pdf
The Landscape of Hybrid Threats: A Conceptual Model	G. Giannopoulos, H. Smith, M. Theocharidou; (JRC and Hybrid CoE)	https://www.hybridcoe.fi/publications/the-landscape-of-hybrid-threats-a-conceptual-model/

Figure 14: Table of analysed strategic approach frameworks

In the last filtering step, one framework was discarded for being too broad, and two were removed due to their lack of public availability. One of those was simply not publicly available, while the other was discontinued and taken down after a change in the US government. As strategic frameworks are closest to governance, they can be connected to politics and bound to them, which can explain these findings. It further implies that strategic frameworks can be closely linked to institutions, which could make them dependable and less stable over time. Overall, the number of 15 remaining frameworks that will be analysed in this chapter, the highest in all of the clusters, could indicate that strategic frameworks are the most compatible with FIMI or that strategic aspects are the most predominant for FIMI researchers. Lastly, the strategic cluster also

has secondary affinities with other areas (operational, behavioural, technical), further highlighting the interdisciplinary nature of FIMI.

4.5.1 Similarities and Shared Ground

All fifteen frameworks share the same premises, as they: address governance challenges; are aware that institutions must be part of fighting FIMI seek a whole-of-society response; and are future-oriented. They are focused on decision-making, policy, coordination, or high-level responses, clearly placing them in the strategic cluster. They all treat the challenge of countering FIMI as systemic and as warranting coordinated societal or state-level action, not merely individual correction. Institutions must therefore also play an active role.

State-level action alone is, however, neither sufficient, as strategic frameworks assume a multi-actor system and a whole-of-society response. DISRUPT states this explicitly; Analytical framework for countering hybrid threats and FIMI address government and non-government actors alike, Societal Resilience to Disinformation Framework address a variety of actors (e.g., policymakers, educators, and technologists), and the Counter-Hybrid framework sees it as a precondition. This, in turn, makes coordination across actors a recurrent concern, which could be seen as a strategic vulnerability. Furthermore, most frameworks are designed to generalise across cases rather than to resolve a single incident, treating FIMI as a persistent challenge. The strategic cluster is future-oriented, seeking to build resilience and capacities rather than incident-specific solutions.

The frameworks of the strategic cluster further share conceptual moves. The aspects of actor, means, and effect recurs across a lot of the frameworks (e.g., ABCDE: actor, behaviour, content, degree, effect; ALERT: actor, lever, effects, response; the Landscape of Hybrid Threats: actors, domains/tools, phases; the Department of Justice typology: operation types defined by actor and behaviour), highlighting these as important aspects of strategic frameworks.

Another pattern is that frameworks include sequential or even cyclical phasing of activities/responses (e.g., Palmertz, Counter-Hybrid Framework, RESIST, DISRUPT). This architecture reflects a process-oriented understanding of FIMI, in which both influence activities and responses are viewed as evolving over time rather than as isolated events.

The DIME family of “instruments of power” is also mentioned by frameworks (e.g., the Hague Centre for Strategic Studies Framework: DIMEL, Counter-Hybrid Response Framework: DIMEFIL), which could indicate that influence activities are interconnected with other forms of statecraft and power projection.

Furthermore, vulnerability and resilience appear as the defensive counterpart to threat, and building resilience is a shared aspiration across frameworks (e.g., Palmertz, Rød et

al., Monaghan, DigCom 2.2), even where the mechanisms differ. This could indicate that the frameworks are prepared for the case that prevention and mitigating action can fail, which makes building resilience a strategic necessity. It therefore suggests a shift from threat elimination toward vulnerability reduction, reflecting an understanding that long-term societal robustness may be more achievable than complete prevention. A stated ambition toward shared terminology, interoperability, or standardisation is itself another shared feature among the frameworks, with ABCDE seeking a common standard for EU institutions, Palmertz et al.'s shared analytical framework, Alvarez-Galvez giving an interoperable definition, DISRUPT sharing standards and tools for taskforces, and RESIST providing a common process for communicators. This ambition indicates that many strategic frameworks view coordination failures as a significant challenge in FIMI response, making common terminology, interoperability, or standardisation a strategic objective.

Most of the are largely content agnostic, with the Department of Justice report (elections) and the Alvarez-Galvez model (health) being exceptions, which are built on specific cases but claim broad transferability. It is worth noting that none of the fifteen frameworks have been stress-tested across the full FIMI subject range (e.g., electoral, military, hybrid, foreign-policy, health).

The strategic frameworks also share some of the same limitations. None uses FIMI as their operative term; instead, adopting differing terminologies. This could highlight the overall need for more unified language to improve communication and strategic coordination. Attribution is also not methodologically covered, just mentioned, by this cluster, which will therefore need to be supplied by other clusters.

Effectiveness measurement within this cluster remains largely aspirational as evaluation stages are named (e.g., Palmertz's Evaluate, the Counter-Hybrid Evaluation stage, RESIST's Tracking), but metrics remain thinly covered (by only Rød et al., and which remains untested). Therefore, no fully established and tested effectiveness measurement exists within the cluster.

Lastly, empirical testing or evidence of the frameworks is not the norm, as several frameworks are proposals or literature-derived syntheses rather than evaluated tools, whereas others draw on accumulated practitioner experience (DISRUPT and RESIST 3) and policy background (the Department of Justice's framework).

4.5.2 Contrasts and Divergences

Although all fifteen are strategic frameworks, they have differing intellectual lineages, originating from different sub-disciplines. These heritages can be split into two major groups:

- One coming from security and strategic affairs tradition (e.g., JRC-Hybrid CoE framework, Cross-Domain framework, Counter-Hybrid Framework, Department of Justice Framework);
- While the other originates in communication, resilience, and information integrity tradition (e.g., DigComp 2.2, Vadapalli et al., Rød et al., Alvarez-Galvez et al.).

This divergence shapes the frameworks and can explain a few of the following differences. The differing approaches and heritages do not compete or contradict each other; they can instead be seen as complementary, bringing different insights to try to solve the same problem.

One key divergence, building out from these differing heritages is a confrontational versus protective posture. Monaghan, the Counter-Hybrid framework, DISRUPT, the Department of Justice posture, and the coercion/control end of the Cross-Domain framework build on an active confrontational approach to deter, disrupt, punish, or expose. On the other side, the DigComp 2.2, Global Partners Digital, Rød et al., RESIST, Alvarez-Galvez, Vadapalli, and the cooperation/protection end of the Cross-Domain framework take a protective resilience-focused approach to prevent, educate, communicate, or safeguard rights. These two groups of frameworks are not opposed to each other and can instead be seen as complementary, exploring different aspects of the same problem. Furthermore, this may reflect the differing organisations and contexts in which the frameworks were created.

The fifteen frameworks also vary in their principal unit of analysis, identifying different primary objectives and aspects of fighting FIMI:

- The JRC-Hybrid CoE and the Alvarez-Galvez et al. frameworks describe the problem, providing a threat-environment characterisation, without prescribing a response.
- Palmertz et al.'s and Rød et al.'s frameworks assess exposure, vulnerability, and existing countermeasures, providing a nation or society capability assessment.
- Both the Cross-Domain framework and Monaghan's framework focus on statecraft and strategy, as they are at the level of the state as a strategic actor facing an adversary.

- The Counter-Hybrid framework, RESIST 3, and DISRUPT take an operational response approach as they specify what to do and in what order.
- The Department of Justice's framework and the Global Partners Digital framework serve a single institution or instrument each.
- ABCDE and ALERT decompose a specific operation or attack.
- Vadapalli et al. take a systemic relationship as the object, rather than an actor, a receiver, or a response, whereas DigComp 2.2 locates resilience in the competences of the individual.

This diversity is not a weakness; FIMI is complex, and some aspects need to be explored from different perspectives, which in turn can offer different solutions (e.g., resilience of citizens: education and literacy; resilience of institutions: governance and coordination).

The frameworks also have varied starting points and direction of travel, again reflecting their differing heritage and institutional context. The approach of starting from the hostile actor and its strategy is used by the JRC-Hybrid CoE, Cross-Domain, Monaghan, Counter-Hybrid, DOJ, ALERT, and ABCDE framework. On the other side, starting with one's own vulnerabilities, capabilities, or audiences is the approach implemented by Rød et al., DigComp 2.2, and RESIST. Palmertz et al.'s framework is explicitly double-sided, exploring both adversary and defender. DISRUPT starts from what the responder will do, while Vadapalli et al.'s framework starts with an object-as-relationship approach.

The adversary and their intentions are also treated differently across the framework, revealing a varied understanding of the challenges. Some frameworks treat the adversary as central and elaborate further (e.g., JRC-Hybrid CoE, Cross-Domain, Monaghan, Counter-Hybrid, and the Department of Justice framework) while others see them as central but generic (e.g., ABCDE, ALERT). Another cluster does not treat the adversary's intent as the central focus of their analysis (e.g., Rød et al., Vadapalli et al., Alvarez-Galvez et al., DigComp 2.2). The strategic frameworks also differ in their treatment of the response side and in how much detail they give the adversary's operations. Counter-Hybrid, RESIST, DISRUPT, Cross-Domain, Monaghan, the DOJ posture, Palmertz, and Rød all specify how a state or society acts, not merely how a receiver thinks. DISRUPT provides the most detailed supply side, and the JRC-Hybrid CoE frameworks domains-and-tools axis is the closest thing to a structured supply-side schema at the strategic level. The Department of Justice's five-type typology and ABCDE's behaviour component also describe adversary activity, though less granularly.

Lastly, the treatment of legitimacy, law, and rights also varies across the strategic cluster. These aspects are central to some (e.g., Global Partners Digital, Department of Justice), embedded as a dimension (e.g., Palmertz et al., Rød et al., Cross-Domain,

Counter-Hybrid, DISRUPT), or largely absent in others (e.g., JRC-Hybrid CoE, Monaghan, ABCDE, ALERT, Alvarez-Galvez, Vadapalli, and DigComp 2.2). It is worth noting that the most confrontational frameworks (deterrence, disruption, exposure) are the ones that most need a legitimacy guardrail, yet the dedicated guardrail sits outside them. This could be due to legality and proportionality often being treated as parallel concerns instead of being fully embedded into response logic.

4.5.3 Coverage of the FIMI Response Process Phases

The strategic cluster spans all phases of the FIMI Response Process; however, two phases only weakly: the *Detection and Identification*, and the *Evaluation and Learning* phase (see Figure 15 below). Strategic frameworks focus on planning, governance, coordination, and policy-making; thus they inherently need to consider multiple phases of the response process simultaneously, explaining the cluster's broad coverage. They may span many phases because their primary purpose is not to perform a single activity, but to connect and coordinate activities across the wider FIMI response ecosystem.

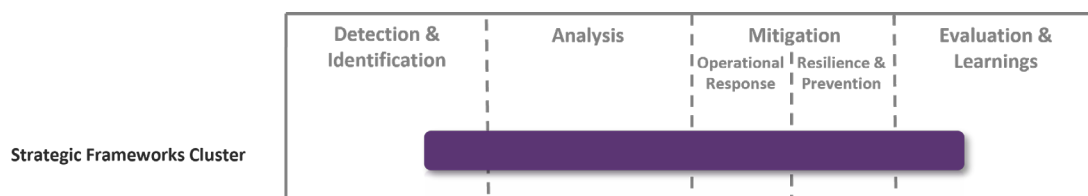


Figure 15: Strategic Frameworks Cluster Coverage of FIMI Response Process Phases.

The *Detection and Identification* phase is weakly served by the strategic framework cluster, as while some engage in detection, several assume detection capability rather than building it (e.g., ALERT, RESIST, DISRUPT). Therefore, strategic frameworks may assume detection as a prerequisite. FIMI needs to be detected to be acted upon at the political and governance level; however, the strategic actors might not be the ones doing the detection and identification. This indicates a division of labour and tasks across the counter-FIMI ecosystem, further highlighting the need to support coordination and collaboration. For interventions, strategic frameworks appear to start when sufficient evidence exists to justify coordination, policy discussion, or response planning. On the other hand, regarding long-term programs like building resilience, strategic frameworks are needed and can be used independently of detection and identification, which offers another possible explanation.

The *Analysis* phase is well served by the strategic cluster as the frameworks offer analytical support at different altitudes (e.g., JRC-Hybrid CoE, Palmertz et al., ABCDE, ALERT, Alvarez-Galvez, Vadapalli, and Rød). Strategic decisions carry significant consequences, requiring a strong analytical basis before action is taken. For high-level decision-making, understanding which actors are involved, what the objectives are, and what broader implications exist is crucial.

Another well-covered aspect is the *Mitigation - Operational Response* phase (e.g., DISRUPT, RESIST, Counter-Hybrid, the Department of Justice, and the Cross-Domain). This finding can be explained by the oversight and coordination capabilities that strategic frameworks offer, therefore enabling action. Furthermore, individual organisations may already possess response capabilities, but strategic frameworks help align those capabilities across actors. This finding reinforces that strategic frameworks are not merely analytical tools; they are intended to support action.

The second mitigation aspect, the *Mitigation - Resilience and Prevention* phase, is also well served by the strategic framework cluster (e.g., Rød et al., DigComp 2.2, Monaghan, Cross-Domain, Global Partners Digital, and RESIST framework). Strategic actors not only respond operationally and reactively, but are also in charge of developing and managing long-term resilience and building preventive measures. The cluster recognises that responding to individual incidents alone is insufficient if underlying vulnerabilities remain unaddressed. This may also explain why resilience appears repeatedly across the cluster despite large differences in disciplinary origins, as many frameworks converge on the goal to reduce vulnerabilities. Another aspect is that some strategic interventions (education, policy reform, institutional strengthening) naturally require long implementation periods, requiring strategic frameworks to consider this timeline.

Lastly, the *Evaluation and Learning* phase is served weakly as it is named in several designs, but more often asserted than operationalised (e.g., Palmertz et al.'s Evaluate, the Counter-Hybrid Evaluation stage, RESIST's Tracking). The mention, despite lacking operationalisation in most cases, nonetheless implies that this phase is relevant and should be explored further. The only framework that operationalised this phase is Rød et al. However, this methodological approach has not yet been tested. The repeated inclusion of evaluation stages suggests broad recognition of their importance, but the lack of concrete methods indicates that the field has not yet reached a consensus on how success could be measured. In the strategic cluster, this might be challenging due to the complexity of measuring strategic outcomes, which often emerge over long time periods and involve multiple actors and external variables.

4.5.4 Application of Frameworks in Practice

The expert interviews showed organisations tend to utilise internal frameworks, processes, and approaches. This may reflect the nature of strategic work itself, as coordination, governance, planning, and decision-making are highly dependent on organisational context. Strategic responses need to align with an organisation's mandate, responsibilities, resources, partnerships, and position within the broader FIMI ecosystem. As a result, organisations may adapt existing frameworks or develop their own approaches that better reflect their specific needs and operational realities.

Three specific frameworks were mentioned by experts. Both RESIST and DISRUPT were mentioned by one expert, as providing structure and organisation for the expert's work processes. While the limited number of mentions does not allow for broader conclusions regarding their adoption or effectiveness, the finding is consistent with the intended purpose of both frameworks to provide structured approaches for understanding, monitoring, and responding to information threats.

The ABCDE framework was mentioned by six out of seventeen experts, suggesting it has achieved a level of practical recognition and uptake that other strategic frameworks have not. This popularity may be explained by the positives of the frameworks that the experts mentioned. The framework is repeatedly described as easy to use, straightforward, and offering a clear overview and structure. Experts repeatedly described ABCDE as helping them expand analysis beyond narratives and understand incidents more comprehensively. It was also described as a way to contextualise and frame information threats. The main limitation of the framework, pointed out by experts, was that ABCDE does not go into response or action. This limitation reflects the framework's intended purpose rather than a flaw, as it was designed to support understanding and categorisation, rather than prescribe responses. Nevertheless, practitioners appear aware that analysis alone is insufficient and that additional frameworks, processes, or organisational procedures are required to translate understanding into action.

4.5.5 Contributions to a Comprehensive Analytical Approach

The strategic frameworks do not form a fully cohesive group. This may be less a sign of fragmentation than a sign of specialisation, with frameworks occupying complementary positions because they address different stages and originate from different contexts. While they can be loosely grouped in a way that suggests an overarching architecture, this should not be understood as a strict or intentional layered model. This apparent compatibility of the frameworks is not the result of deliberate coordination between authors, but rather reflects the existence of recurring strategic functions that must be addressed when responding to FIMI:

- Firstly, the framework developed jointly by the JRC and Hybrid CoE supplies a threat-environment map covering actors, domains, and the priming/destabilisation/coercion phase logic.
- Whereas Palmertz et al. supply the national assessment process offering a double-sided mapping of threat against vulnerability, wrapped in an Assess-Address-Evaluate cycle usable by state and non-state actors.
- ABCDE and ALERT both supply case-analysis vocabularies for decomposing a specific operation or attack, while Alvarez-Galvez et al. supply a transferable domain-characterisation template and an explicit interoperable-definition method.

- Both the Cross-Domain framework and Monaghan's framework supply the strategy and deterrence logic, whereas the Counter-Hybrid framework supplies the sequenced decision-and-execution cycle, with explicit attention to escalation and second- and third-order effects.
- To that RESIST 3 can offer the communicator's end-to-end operational process, including a tracking step, and DISRUPT supplies the disruption playbook (per-operation workflows) and a shared coordination infrastructure.
- The Department of Justice's Framework provides a real-world institutional example of how a government can respond to FIMI through exposure, prosecution, and deterrence while remaining constrained by legal and democratic principles.
- Rød et al. supply the societal-resilience maturity-measurement ambition, whereas Global Partners Digital's framework could supply the legal/rights guardrail across all responses, DigComp 2.2 supplies the population-level competence base, and Vadapalli et al. supply the harm-pathway logic that explains why counteraction matters.

While these frameworks collectively cover a broad set of strategic functions, this overview also highlights internal frictions/tensions between them. The described different topical lineages of the frameworks, reflecting the interdisciplinary nature of FIMI, create tension not only at the level of terminology but also in underlying analytical assumptions, with security studies foregrounding strategic actors as key sources of harm, while information integrity approaches place greater emphasis on systemic communication conditions shaping how harm emerges. Another tension between the frameworks is the asymmetry in access to information, as some frameworks assume a classified or national-security context, whereas others are open-source and civil-society focused. This is in itself not a disagreement but rather a capability constraint tension due to different mandates, organisational context, and varied access to information. It therefore brings an interoperability challenge that needs to be explored further to achieve better coordination and cooperation.

Next, the deter/disrupt/expose frameworks and the resilience/literacy/rights frameworks pull in different directions and imply different actors and risk-taking. This tension in itself is not a disagreement but rather a role-based trade-off situation, highlighting differences in institutional mandates (security agencies vs education/public health vs civil society), risk tolerance (harm prevention vs rights protection), and temporal focus (short-term disruption vs long-term resilience). Further, both sides are important and are not opposed. An integrative approach should find a way to integrate both sides, to fight FIMI with both a confrontational and protective approach. This tension directly highlights a structural gap of the cluster: the integration of the citizen layer with the state layer. DigComp 2.2 is the only one that dives deeply

into the citizen side, and the entire cluster lacks the linking of population-level resilience to state-level response.

Also affecting a part of the cluster, is the generalisation of frameworks that were built on a specific example, or for a specific situation (e.g., the Department of Justice's framework, Alvarez-Galvez et al.'s framework), as they risk overfitting. However, the frameworks do claim transferability, which is yet to be tested. Another gap in the cluster is the lack of an attribution method. The effectiveness and impact measurement aspect is mostly an aspirational thought in this cluster, and therefore another major gap, as only Rød et al. propose a concrete, computable metric, albeit one yet to be tested. There is also a lack of embedding and integration of rights and legitimacy aspects within response frameworks. While the cluster covers these aspects separately, it is yet to be integrated into action-oriented frameworks that would benefit from the inclusion (e.g., the deterrence and disruption frameworks).

Operationalised cross-border coordination also remains mostly absent, another important gap, as DISRUPT and the Counter-Hybrid framework gesture at multi-state coordination, but concrete mechanisms remain thin. Lastly, the frameworks lack generative-AI adaptation, which is only addressed partially (e.g., DigComp 2, Alvarez-Galvez et al., Vadapalli et al.) and not yet integrated into the response frameworks.

Nonetheless, each of the frameworks in the strategic cluster can offer important aspects that should be considered when building an integrative approach to FIMI:

- The *JRC-Hybrid CoE framework* provides cross-domain threat-environment mapping and a priming-destabilisation-coercion logic, establishing the broader macro-context for FIMI activity.
- *Palmertz et al.'s framework* contributes a double-sided threat-and-vulnerability mapping and a complete assess/address/evaluate cycle usable by government and non-government actors.
- The *Cross-Domain framework* contributes a structured menu of response strategies (from cooperation to control) and an escalation logic that disciplines the question of which response, and how escalatory.
- *Monaghan's framework* contributes an extended deterrence repertoire and the three-Cs test (capability, credibility, communication), which are essential to the preventive and dissuasive aspects of an approach.
- The *Counter-Hybrid framework* contributes the most complete sequenced response cycle, with explicit handling of escalation and second- and third-order effects.
- The *Department of Justice's framework* contributes a worked example of bounding active response by law and democratic legitimacy.

- *ABCDE* contributes a tried, interoperable case-analysis rubric already adopted across EU and UK practice, making it a strong candidate for a common vocabulary for the analysis phase.
- *ALERT* contributes an information-systems framing that ties the manipulation to the system being exploited, plus an attacker–defender co-evolution dynamic.
- *Rød et al.’s framework* contributes the only concrete proposal for a computable maturity metric of countermeasures.
- *Vadapalli et al.’s framework* contributes a harm-pathway model (misinformation, polarisation, distrust) and a caution against alarmism.
- *Alvarez-Galvez et al.’s framework* contributes a transferable six-domain characterisation template and an explicit method for cross-disciplinary semantic alignment.
- *RESIST 3* contributes a practitioner-ready, end-to-end communicator process with a tracking step.
- *DISRUPT* contributes the most granular disruption playbook, plus shared infrastructure and a whole-of-society coordination model.
- *Global Partners Digital’s framework* contributes a ready-made legitimacy and proportionality test, providing a rights guardrail that could be used in an integrative approach.
- *DigComp 2.2* contributes a population-level competence reference that provides a societal-resilience base that is anchored within EU policy practice.

Taken together, the fifteen frameworks supply the adversary side, the response side, and partly the learning side, and add a legitimacy overlay. This notwithstanding, the cluster lacks validated attribution methods and effectiveness measurements. For an integrative approach, ways of combining confrontational and protective measures as well as connecting the highest layer of state strategy with its lowest layer of citizen competence need to be explored.

4.6 Cross-Categorical Analysis of Frameworks and Conclusions

The analysis of 31 frameworks across four distinct categories reveals a rich but fragmented landscape for countering and analysing FIMI (see Figure 16 below). Each framework cluster brings unique strengths that address different aspects of the challenge. Technical frameworks support detection and scalable monitoring; behavioural frameworks excel at exploring the defender as well as individuals' cognition and building long-term resilience, whereas operational frameworks provide

structured procedures for practitioners responding to incidents, and strategic frameworks enable high-level governance and coordination.

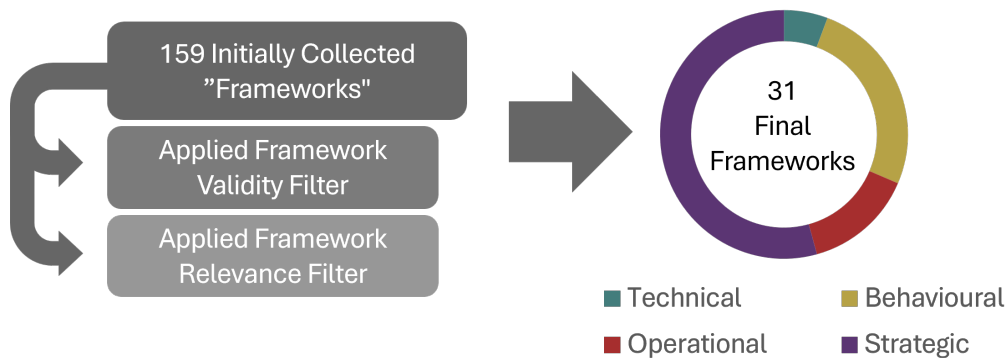


Figure 16: Composition of the Final Framework Dataset by Framework Type.

It should be noted that the categorization into clusters was not intended as a definitive classification, but rather as an analytical structure developed to manage the complexity of the collected data, facilitate analysis, and present it in a more accessible and coherent way for the reader. The clusters should therefore be understood as analytical layers rather than rigid categories, introduced to structure a highly fragmented ecosystem through a deliberate layering approach.

The main contribution of the current synthetic analysis resides in creating order within this complexity, allowing for both analytical separation and a structured reconnection of the layers. Accordingly, this part of the analysis reconnects these layers and proposes their combination as a means of addressing FIMI through an interdisciplinary, multilayered approach.

Furthermore, it should be pointed out that reliance on formal frameworks in practice is not uniform across practitioners and experts we interviewed. Some experts interviewed for this study make use of established frameworks:

- DISARM was mentioned by 9 out of 17 experts;
- And ABCDE was referenced by 6 out of 17 experts;
- Others do not rely on formalized frameworks at all, instead drawing on experience-based or ad hoc approaches developed over time. In addition, some practitioners operate with internal, organization-specific methodologies that are not publicly accessible.

Finally, it should be noted that the EEAS FIMI Toolbox, as mentioned in the methodology, was not included among the 31 frameworks analysed, as it constitutes a distinct case within the context of this deliverable, rather than an independent framework suitable for comparative analysis. At the cross-categorical level, however, it is included as a contextual reference alongside its subsequent components (the

Response Framework (EEAS, 2024), the FIMI Exposure Matrix (EEAS, 2025), and the Deterrence Playbook (EEAS, 2025)) to situate the broader findings within a European context and examine them in relation to the recurring themes, gaps, and challenges identified across the analysed frameworks.

The framework clusters all cover aspects of analysis with different foci and perspectives, while also each contributing different aspects of the FIMI response process that could compensate for where other clusters lack. Technical frameworks, for instance, excel at a system-based perspective, but tend to lack the human component, that is better addressed by behavioural frameworks. Conversely, where behavioural frameworks focus on the individual and defender perspectives, the attacker-oriented dimension can be supplemented by frameworks from the operational and strategic clusters. In turn, where the operational cluster falls short in coordination and consideration of connected incidents, the strategic cluster provides additional insight. While the strategic cluster offers a high-level structure and overview, behavioural frameworks help refine and add nuance to this perspective.

Building on this, a similar logic of substitution can be observed across the different phases of the FIMI response process. Although all clusters contribute to the analysis phase, they complement each other in relation to other stages (see Figure 17 below). The technical cluster primarily contributes to the *Detection and Identification* phase, which is less developed in the other clusters. The behavioural framework cluster contributes to the long-term aspects of mitigation (*Mitigation - Resilience and Prevention*) while lacking the *Operational Response* aspect of mitigation, which the operational cluster covers. The strategic cluster, in contrast, spans both short-term and long-term mitigation efforts, but requires supplementation in the *Detection and Identification* phase.

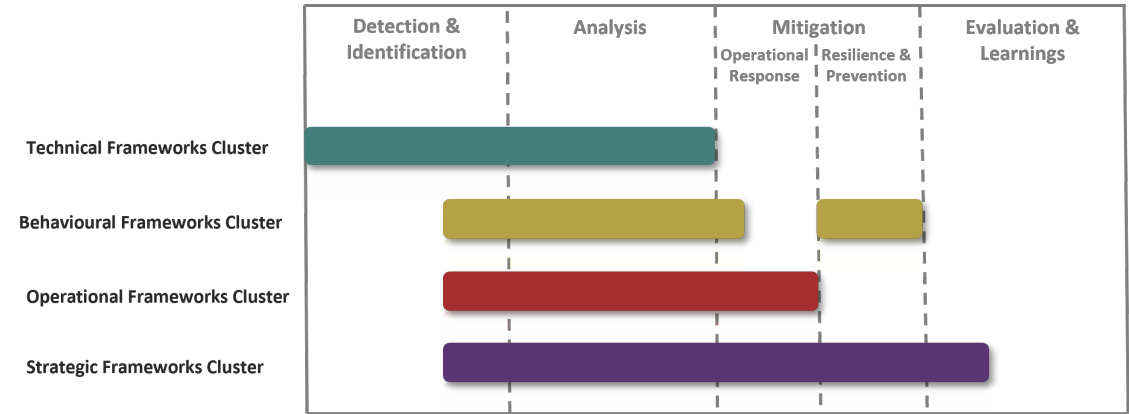


Figure 17: Coverage of FIMI Response Process Phases by Analytical Clusters.

These patterns are further supported by the expert interviews. Several experts explicitly describe combining frameworks or elements of different frameworks within their internal workflows. For example, integrating ABCDE as an overarching system,

integrating a technical layer, or expressing the intention to incorporate behavioural frameworks to add further nuance to their practice.

A similar logic of partial coverage and functional layering across different elements can also be observed in the EEAS FIMI Toolbox. The FIMI toolbox aligns with the FIMI response process through its response framework (EEAS, 2024) that also covers similar aspects (e.g. prerequisite awareness, detection, mitigation measures, evaluation and learnings). Some of these aspects are referenced and highlighted in terms of their relevance, but are not further operationalised. The framework can thereby contribute to the mitigation phase as it includes operational response aspects like how to react directly (ignore, contain, minimise, redirect) and long-term aspects like resilience building. The FIMI Exposure Matrix (EEAS, 2025) contributes to the detection and identification phase by categorising into attributed (covert: state-linked channels, overt: state-controlled outlets, official state channels) and non-attributed (state-aligned channels) modes. The Deterrence Playbook (EEAS, 2025) operationalises deterrence aspects (sanctions, law enforcement, digital regulation, resilience-building), contributing to mitigation, with the majority of aspects relating to the long-term resilience and prevention phase. The Toolbox therefore contributes to the FIMI response process, but does not fully cover the entire response cycle and shares a number of gaps and challenges also observed in other frameworks.

A number of framework-level challenges and gaps emerged from both the comparative analysis and expert interviews. Some relate to the design and conceptual aspects, for example no clear taxonomy or shared vocabulary presenting itself within or across framework clusters, reflecting their interdisciplinary nature and differing lineages. In addition, frameworks are defined by inherent trade-offs, positioning themselves on different sides of the same spectrum. These do not represent conflicting approaches, but different trade-off decisions within the same conceptual space, particularly evident in the technical cluster (e.g., simplicity vs. depth, flexibility vs. standardization, accessibility vs. technical sophistication), the strategic cluster (e.g., offensive vs. defensive frameworks), and the operational cluster (e.g., evidentiary depth vs. operational speed, fixed vs. modular structure). Furthermore, frameworks must be adaptable to changing environments, highlighting a key challenge of modularity. Practitioners, especially in resource-constrained or high-tempo settings, often find formal frameworks too rigid or complex for immediate application, leading to hybrid approaches that combine established frameworks with proprietary tools to improve flexibility, scalability, and usability. While the FIMI Toolbox reflects this modularity, it remains constrained, like other frameworks, by its organisational context, in this case specifically focusing on EU-level instruments (e.g. restrictive measures and regulatory enforcement).

Framework use also varies significantly across practitioner groups and contexts: while investigative journalists and civil society organizations prioritize internal verification and evidentiary standards over FIMI-specific taxonomies, military and defence actors more often rely on established doctrines or counter-terrorism frameworks, sometimes

viewing newer FIMI frameworks as redundant. Also, while framework clusters can compensate for each other across most phases of the FIMI response process, the Evaluation and Learning phase is not adequately covered by any of them. Across frameworks, there is no formalized approach to evaluation and iterative improvement, with effectiveness instead assessed through experience-based judgment and proxy indicators such as takedowns, sanctions, volumes of actionable intelligence, or the clarity of outputs to stakeholders.

This challenge is also present in the FIMI Toolbox, which acknowledges the importance of evaluation and learning and their potential value but does not operationalise them, leaving this as a continued gap. In addition, the transition from analysis to response is not consistently supported, as the translation of findings into actionable countermeasures is often handled through external political or legal mandates rather than being integrated within frameworks. Cross-team and international coordination and interoperability also remain a challenge, highlighting the need for better integration into framework design to enable alignment between actors such as NGOs, governments, and platforms. This relates to the framing of the FIMI Toolbox, where the need for joint defence and interoperability is emphasised, but no clear actionable frameworks or resources are provided to operationalise it. Finally, asymmetry in information access persists, particularly for strategic analysis as well as detection, requiring frameworks to explicitly integrate different data tiers, including open, proprietary, and classified sources, which is also explored in the FIMI toolbox.

The expert interviews particularly revealed challenges related to rapidly evolving information environments. While threat actors increasingly use AI to scale and accelerate content production, existing models remain largely retrospective and only partially account for these dynamics, highlighting difficulties in keeping pace with such developments. There is also a need to move beyond supply-side perspectives towards greater consideration of demand-side factors such as institutional trust and societal vulnerabilities, as well as to incorporate longitudinal analysis to better understand the evolution of threats over time. In addition, the findings point to the relevance of expanding analytical scope beyond social media dynamics to include offline activities and financial infrastructure within a more multi-platform understanding of FIMI.

5 Evaluation of FIMI Countermeasures Effectiveness

5.1 Systematic evaluation vs unsystematic monitoring of countermeasures?

“[T]here's no methodology for evaluating these things, be it countermeasures or FIMI from a threat actor.”

This statement by a government official interviewed in the course of this research provides a stark and succinct assessment of the current state-of-the-art when it comes to evaluating the efficacy (or otherwise), of countermeasures implemented to impact the information effects and societal consequences of FIMI. Currently, there is no widely accepted or agreed upon framework for conducting evidence-based evaluations of the intended and unintended consequences of FIMI countermeasures, either singly or as part of wider frameworks.

That this is so, is symptomatic of a wider capability and capacity deficit in the FIMI defender community, in terms of being able to estimate the impacts and effects of foreign information manipulation threats more broadly, which remains a shared and cross-cutting challenge (Innes et al. 2026). There was a clear consensus from the interviews with government officials, civil society and journalists conducted in May 2026, that reliable and robust evaluations of the effectiveness of FIMI countermeasures is not routinely conducted. This severely limits our ability to know what works, what doesn't and what is promising when it comes to disrupting specific threats, and defending against the strategic effects of such actions more generally.

Elaborating the concerns outlined in the quotation used to open this chapter, a second government official highlighted how:

“It's on an ad hoc basis. For example, we can extract some quantitative numbers when it comes to channels that we managed to take down. Also, when we talk about capacity building, we know in how many activities we took part or how many people we trained. We have numbers on how many people we sanctioned, but we don't know how this reflects the real impact after.”

The desire for quantifiable impact metrics of control interventions, was a recurring theme across the interviews conducted, albeit accepting that the limitations on available data, render these partial and incomplete:

“It becomes mostly anecdotal rather than statistical.” (Civil Society)

“No, we don't really have a system for that. If anyone has, please let me know. It's such a soft topic. It's really hard. It is incredibly hard. And I mean, even the SDA [Social Design Agency] lies to the Kremlin about how they're having their impact.” (Civil Society)

The key point to be drawn out from these multiple quotations is that whilst systematic evaluation of countermeasures to disinformation and FIMI are not implemented, many organisations and individuals are engaging in unsystematic monitoring of activities and using these to try and identify lessons learned for future planning. This is a general tendency present across multiple different types of countermeasure activities, including strategic communication, public awareness raising campaigns, research, detection, analysis and exposure of threat actors, cooperation with the trusted flaggers under the Digital Services Act, journalism and OSINT investigations, regulation, media governance, disruption, crisis simulations, trainings and capacity building, engagement with policymakers, international partnerships as well as proposing sanctions and diplomatic measures.

Many of the interviewees interpreted what they were doing as a form of ‘experienced-based learning’ used to see what specific actions work well. For example, representatives of civil society organisations talked about measuring the numbers of accounts of hostile actors that were taken down and the “audience denial” achieved. They also traced how the investigative methods used by their organisation were adopted elsewhere and assessed their own work based on project management evaluation and feedback from customers, donors and funders.

Government officials described their use of public attitude surveys to gauge levels of trust in society and people’s general awareness of disinformation. Similar instruments are also used to track mentions and perception of their own FIMI-related outputs, comparing a certain disinformation narrative’s audience attention to the debunking message’s visibility, following the numbers of accounts taken down and actors sanctioned, and using threat actor attacks targeting Western governmental organisations as a KPI to indicate impact. That said, government officials also emphasised how practically difficult it is to conduct public polling on questions related to disinformation and get informative, valid answers:

“If we ask that, have you noticed any disinformation or rumours spreading online, and how worried are you about the problem? Of course, people will say mostly that, yes, yes, I have noticed that there is disinformation spreading. But how do you define disinformation? I mean, that's not a valid question at all. You can't study it that way. And that also applies to countering activities. If people understand disinformation in hundreds of different ways, how do you measure the effectiveness?” - government official.

Somewhat differently, journalists were able to follow audience readership, views and engagement metrics as well as societal impact, where the indicators of interest could

be that a topic creates public discussion, a policy reaction or change in the law. Journalists emphasised that they see limited impact of the information they have published, because the authorities are frequently reluctant to investigate disinformation actors, viewing it as part and parcel of public debate and concerns about traducing freedom of speech rights.

All of these examples came with important caveats highlighted by the interviewees, demonstrating an awareness of the limits and constraints of the view they can reasonably and plausibly construct based upon limited data. Foremost amongst such concerns, however, was the gradual withdrawal from disinformation and FIMI counter-actions of the US-based social media companies, further complicating work that aims at digital disruption of threat actors' social media account networks. As one respondent phrased it:

“Two years ago, we could work with X and we could work with Facebook in the States where it matters to make these changes. Now, by even talking about it, you're a fascist who's an enemy of free speech. So, it is really difficult to take that impact pathway to create that effect even” – civil society

The problem is that whilst all of these limits and challenges are known and there is a shared awareness of them, in a moment of crisis and threat when there is a clear need to demonstrate ‘that something is being done’, or in the ‘rough and tumble’ of competing for limited funding, these tend to be forgotten. As a consequence, various claims and counter-claims about the effectiveness of interventions are made on the basis of less than reliable evidence.

In part, reflecting some of these deep challenges, a number of interviewees perceived value in the defender community shifting focus from how effective offensive/defensive measures are, towards an approach centring vulnerabilities and the protection of these. Some respondents referred to future plans trying to create more systematic ways to assess the impact of their countermeasures. For example, through a combination of defined KPIs of their own actions, monitoring the information environment for changes in hostile actors' reactions and behaviour, and anticipating the threat actors' potential future moves. One interviewee highlighted that what is acutely needed is evidence of what works in what context, which is a difficult task to achieve within the constantly changing information environment. According to this individual, the challenge of evaluating the impact of countermeasures mainly requires a focus on communication-based interventions, while the case supporting policy development and legal actions such as sanctions are more straightforward. For reasons that will be elaborated across subsequent sections, we are less persuaded of this.

5.2 The view from research literature

Although adopting a different framing of the key issues compared with the practically oriented concerns of many of the interviewees, the research literature confirms their perceptions of a deficit in accurate and reliable methods for measuring the efficacy and impacts of FIMI countermeasures. Fundamentally, there is a lack of evidence about the extent to which different control interventions function to limit what kinds of FIMI actions, and under what conditions they can deliver more or less benefit. In part, this reflects an overarching tendency for far more research to have been directed towards the discovery of bad actors, than control evaluation tasks. Associated with which, insufficient research attention has been directed towards assessing whether countermeasures always function as intended, or whether they can also have unintended and unexpected consequences.

This is a long-standing trajectory of development. Published in 2021, Courchesne et al. undertook a comprehensive review of the state of the evidence about countermeasures. They concluded that aside from fact-checking, none of the key interventions present in the countermeasures toolbox, had been subject to sustained rigorous evidence-based evaluation. Since then, some limited studies have been conducted that address some of the gaps. For example, there has been a cluster of experimental studies, conducted under laboratory-type conditions, generating support for the idea that it is possible to ‘inoculate’ populations to the worst effects of malign manipulative communications. However, plausible doubts remain about how well such findings translate into a ‘real world’, chaotic and ‘noisy’ information environment.

The Courchesne et al. review revealed several aligned and interesting findings. All the studies identified in the research between 1972-2021 focused on analysing countermeasures aimed at the ‘consumers’ of disinformation, and none looked at interventions aimed at impacting the ‘suppliers’ of influence operations. Of the 223 studies reviewed, 167 (75%) evaluated the impact of disinformation disclosure. Both the older research and newer social-media focused work suggest that factchecking is able “to reduce the impact of exposure to false information on individuals’ beliefs, at least on average, as well as their propensity to share dis/misinformation” (Courchesne et al, 2021).

More recently, a smaller number of studies adopting ‘naturalistic’ research designs, have sought to derive evidence about how control interventions actually function in the real world. Notably, Innes and Innes (2023) used a ‘natural experiment’ to compare and contrast the effects of Facebook’s de-platforming of two social media accounts authoring and amplifying Covid-19 conspiracy accounts. In one case, the account removal did have a time-limited impact reducing the amount and visibility of material on the platform. However, in the second case, the moves to de-platform the principal account actually saw an increase in the conspiratorial content. This resulted from the rapid proliferation of a series of secondary ‘minion’ accounts being set up, in response

to Facebook's deplatforming actions, resulting in greater dispersal and circulation of the problematic content.

In a not dissimilar register, Buntain et al (2023) tracked and traced the cross-platform effects that occurred in the aftermath of the January 6th 2021 insurrection event in the US. This saw a large number of visibly 'hard-right' social media accounts migrate from more mainstream platforms, towards smaller, more specialist ones. The effect of which was an improvement of the 'information health' of the platforms they left, but increased toxicity on their destination platforms. Implicitly, this highlights one of a series of moral conundrums that arise in judgements about the perceived effectiveness of control measures. Is something effective if it reduces the exposure of ordinary social media users to problematic content, even if this involves corralling bad actors in a more limited space, where over time, they might mutually reinforce each other's' viewpoints in a self-reinforcing radicalisation spiral?

Adopting a slightly more 'strategic' lens, as opposed to the more tactical assessments rehearsed above, in 2021 the European Court of Auditors released the results of an audit of the EU's measures against disinformation. It strongly criticised the EU for a lack of systematic evaluation of its countermeasures, to which the EEAS and the Commission reacted saying:

"when it comes to the impact of counter-disinformation activities, there is no established methodology to measuring the impact of disinformation, as well as of actions taken to tackle it. This reflects the challenges that the entire community who works to counter disinformation and foreign interference is facing."

This notwithstanding, it is important to highlight that many of the countermeasure options currently available have had a wider application following Russia's full-scale invasion of Ukraine in 2022. The EU, UK and US have all sanctioned FIMI actors, and implemented broadcast bans of state-controlled outlets, and social media platforms' removal, labelling and downgrading policies. These measures have inspired studies analysing different aspects of sanctions' circumvention and inconsistencies in the enforcement of these policies (Arcostanzo et al. 2022; Pamment 2023; Kristoff et al. 2024; Maristany de las Casas and Reyes 2025; Romano et al. 2023). Redolent of the insights of the more 'tactically' oriented evaluations of control measures, there is evidence of unintended consequences being induced. For example, the Social Design Agency, responsible for the persistent Doppelganger information operation, have been repeatedly sanctioned, but have been able to sustain their influence operations. Indeed, there are suggestions that they have benefited from being sanctioned, in terms of receiving commissions and awards via the Russian Presidential Administration (Pamment & Tsursumia, 2025).

By way of summary, and in terms of bringing a number of these themes together, Bateman and Jackson (2024) provide a useful overview visualisation.

Case Study Summaries

Table 1. Overview of Case Studies¹

Type	Intervention	How much is known?	How effective does it seem?	How easily does it scale?
	1. Supporting local journalism	Modest	Significant	Difficult
	2. Media literacy education	Significant	Significant	Difficult
	3. Fact-checking	Significant	Modest	Modest
	4. Labeling social media content	Modest	Modest	Easy
	5. Counter-messaging strategies	Modest	Modest	Difficult
	6. Cybersecurity for elections and campaigns	Modest	Modest	Modest
	7. Statecraft, deterrence, and disruption	Modest	Limited	Modest
	8. Removing inauthentic asset networks	Limited	Modest	Modest
	9. Reducing data collection and targeted ads	Modest	Limited	Difficult
	10. Changing recommendation algorithms	Limited	Significant	Modest



Public information



Government action



Platform action

Figure 18: Evidence-informed assessments of ten commonly proposed ways to counter disinformation, summarising the quantity and quality of research, the evidence of efficacy, and the ease of scalable implementation. Source: Bateman and Jackson 2024

What this suggests however, is that the problems with measuring the effects of FIMI countermeasures are not just methodological and empirical, they are also conceptual. For such control interventions are differently available to separate actors, all of whom have a role in defending society from the malign impacts of disinformation and allied forms of information manipulation and deception. This is signalled in the EEAS FIMI Deterrence Playbook when it highlights that “tangible impact” requires “striking at the key enablers of FIMI operations such as intermediaries, proxies and service providers the structure that sustains them can become progressively fragile.” (European External

Action Service 2026). There are also calls to start exploiting more ‘offensive’ defensive measures, sometimes framed as ‘defending forward’. However, there is even more limited publicly available evidence about these, apart from the war-time measures by Ukraine (Buchanan 2019; Giles 2023). And even less is known about their efficiency.

5.3 What can we observe and what more does it tell us?

Having mapped out the broad contours in terms of evidence for countermeasures, in this section, we seek to enrich this view, as far as possible, through analysis of available data sources. The aim being to highlight some different angles for understanding the scope and scale of countermeasures. For simplicity, we have divided the elements of countermeasures to disinformation and FIMI into detection, disruption and legal action. The former references how in order for control to be executed it is first necessary to identify it, and some measures are predicated upon making this a public exposure. Disruption is the term used for any interventions that degrade the ability of FIMI to function, through inducing some kind of friction or alternative incentives. Then we also look at what can be discerned from the data on legal actions and the use of sanctions.

Coherent with the preceding sections, the data introduced herein is partial and highly limited. Nevertheless, it has value in rounding out our view of some more recent developments in the implementation of FIMI countermeasures. The extent to which we ‘lean’ on data from Meta for this view, further reinforces an earlier comment about the lack of public data being provided by most platforms.

5.3.1 Detection and exposures

Detection and the public exposure of disinformation and FIMI have been central to the defensive field from the very beginning. Apart from support to independent media and media literacy projects, fact-checkers have been providing reliable information to the public and keeping those who spread manipulations accountable. Ukrainian fact-checker StopFake was established in March 2014, and The International Fact-Checking Network (IFCN) was set up in 2015. Subsequent to which, the EU-funded European Digital Media Observatory was launched in 2020, to better connect and coordinate European fact-checkers. Meta’s third-party fact-checking program started in 2016, and ended in the US in 2025 as a “tool to censor” (Kaplan 2025). The program’s future in the EU is under review by Meta.

Meta’s third-party fact-checks in the EU

Meta’s scheme has provided useful time-series data that allows a limited view of the scope and scale of fact-checking interventions across their key services and how these have shifted over time.

Distinct fact-checking articles in the EU

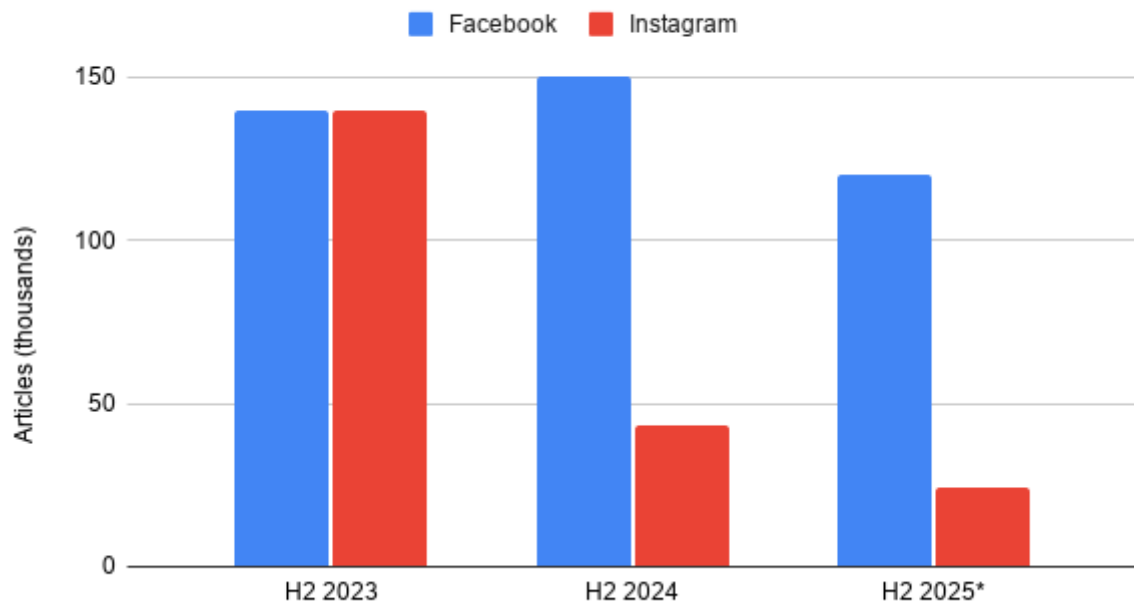


Figure 19: Distinct fact-checking articles used on Facebook and Instagram to label and reduce the virality of content in the EU, as reported by Meta under Code of Practice on Disinformation. For the year 2025, the figures cover EEA (EU member states and Iceland, Liechtenstein and Norway).

The figure above shows Meta’s reported figures of the fact-checking articles it has used to label and reduce the virality of content in the EU during the past three years. There is no significant change in the use of fact-checking articles on Facebook, but a very noticeable decline on Instagram.

5.3.2 Disruption

To get a view on the historical development of disruption activity, we draw on Meta’s coordinated inauthentic behaviour (CIB) takedowns. The figures are collected from 100 threat reports listed on Meta’s threat reporting website (<https://transparency.meta.com/metasecurity/threat-reporting/>) and cover takedowns from 2017 until 2025. While Meta’s reporting likely doesn’t represent the full scope of its actions under CIB, it currently offers the most consistent time-series view in comparison with other technology companies’ reporting. Meta organizes its reporting around a network or information operation taken down. The size of these networks varies from three accounts and pages to more than 8000. To get a better view of the scope of the disruption activity, we calculated figures of accounts, pages and groups removed from Facebook and Instagram by year and state actor mentioned. The overall account figure does not tell us anything about the operations’ possible sophistication or success, but it does indicate the scope of Meta’s effort directed towards the disruption.

According to Meta’s definition, “CIB involves coordinated networks of assets working together to mislead people using false identities. Our enforcement is rooted in deceptive behaviour and coordination, not the content posted—regardless of the narratives promoted or whether an operation is foreign or domestic.” Hence, the figures do not automatically mean the activity taken down would be foreign of nature, as many of the networks target domestic audiences.

Accounts removed in Meta's CIB takedowns

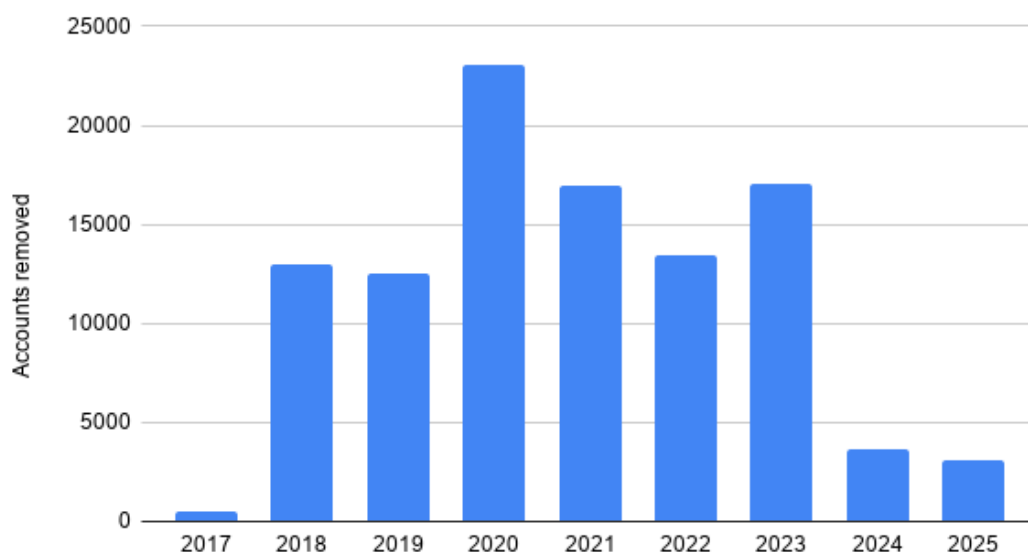


Figure 20: Accounts, pages, groups and events on Facebook and Instagram removed in Meta’s CIB takedowns.

Since the start of Meta’s public reporting on takedowns, the level of removals has remained relatively consistent, with the exception of more account disruption in 2020 and 2023, and a significantly lower pace of removals in 2024 and 2025.

Alongside the well-known and well-established threat actors Russia, China and Iran, Meta’s takedown interventions have included large numbers of accounts, for example, from Mexico and Azerbaijan. However, in both cases these have been individual larger takedowns, while the takedowns targeting networks originating in Russia, China and Iran have been more sustained across multiple years.

Meta's CIB takedowns 2017-2025

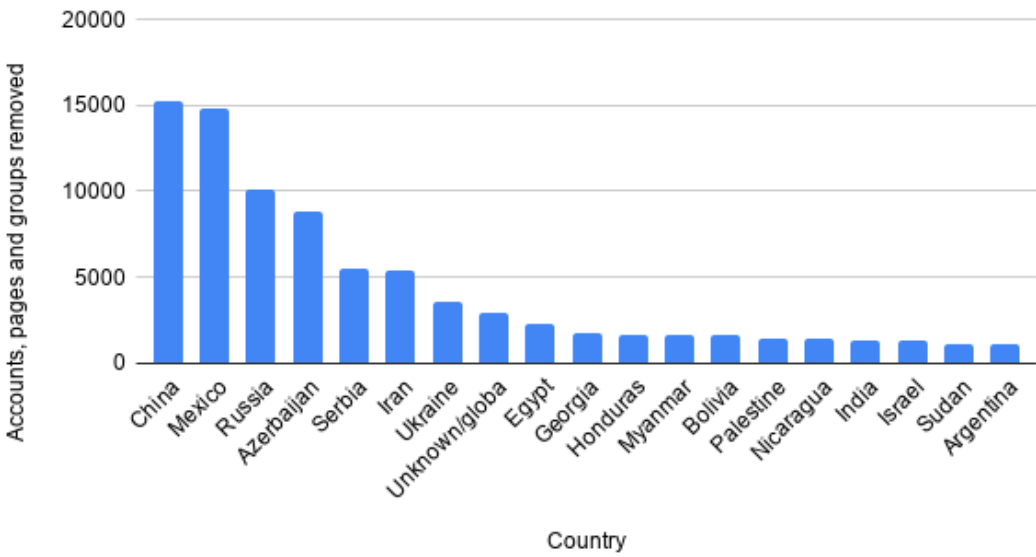


Figure 21: Accounts, pages, groups and events removed from Facebook and Instagram in Meta's CIB takedowns 2017-2025.

Meta's CIB takedowns originating in Russia, Iran and China

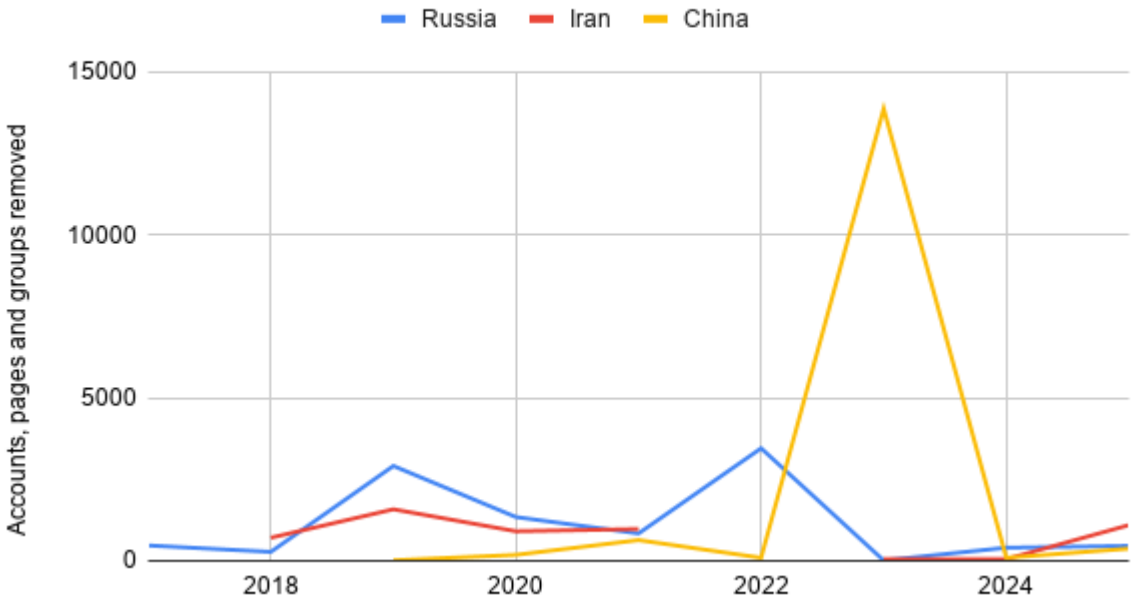


Figure 22: The figure shows accounts, pages, groups and events on Facebook and Instagram removed in Meta's CIB takedowns originating in Russia, Iran and China.

The highest peak in disruption of networks originating in China came in 2023, when Meta took down a large number of accounts linked to the so-called Spamouflage operation.

Within their reporting, Meta publicised one CIB takedown targeting EU countries, where the disruption has been made based on coordinated inauthentic behaviour *on behalf of a foreign or government entity*. In 2020, Meta took down networks originating in France and Russia, targeting North Africa and Middle East, with Meta connecting the activity with the French military:

“Two campaigns — from France and Russia — actively engage with one another, including by befriending, commenting and criticizing the opposing side for being fake” (Gleicher and Agranovich 2020). The French military did not confirm Meta’s findings, and condemned disinformation efforts (Associated Press 2020).

A similar mirrored takedown followed in 2021, when Meta took down networks originating in Poland and Belarus – this time only for ‘coordinated inauthentic behaviour’, without specifying if it is on behalf of foreign or government entity. The network originating in Poland targeted Belarus and Iraq and dissuaded migrants from entering the EU (Meta 2021).

These latter two cases are worth highlighting on the grounds that, although framed, differently they once again highlight the possibility for unintended consequences to arise from FIMI countermeasures. These instances specifically illuminate how the adoption of a more offensive posture by Western governments, can if detected and exposed, result in reputational damage.

5.3.3 Legal action

To obtain a preliminary overview of patterns and trends in EU sanctions imposed in response to disinformation and FIMI activities, a search was conducted across sanctions regimes with key words such as ‘propaga*’, ‘disinfor’ and ‘manipulat*’ within the reasoning text of restrictive measures lists. Sanctions regimes were used as listed here <https://www.sanctionsmap.eu/#/main>, and the search was conducted on restrictive measures consolidated texts of EU Council Regulations at <https://eur-lex.europa.eu/>. The results show that the majority of these sanctions are implemented under the regime targeting actions to undermine the territorial integrity of Ukraine. In addition, Russian media outlets are banned from broadcasting based on this regime. Only one Russian propagandist, Dmitry Kiselyov, was sanctioned in 2014, and majority of other disinformation-linked sanctions have been enforced after Russia’s invasion of Ukraine in 2022.

The new EU sanctions regime from 2024 targeting Russia’s hybrid attacks, specifically sanctions the use of information manipulation and interference attributable to the Russian state. Some individuals and entities are listed in both regimes. Reprising some of the issues highlighted in Chapter 2 of this report, within the cyber sanctions’ regime, only the Iranian cyber actor Emennet Pasargad features. Overall, these sanction regimes list 320 individuals and entities engaged in propaganda, disinformation or information manipulation, with some listed in two regimes. Chinese individuals and entities have not been sanctioned by the EU for propaganda, disinformation, information manipulation or FIMI under these sanction regimes. However, several Chinese individuals and entities have been listed for cyber-attacks (Council of the European Union 2019/2026).

The figure below provides a summary overview of relevant data.

EU's disinformation and FIMI sanctions in April 2026

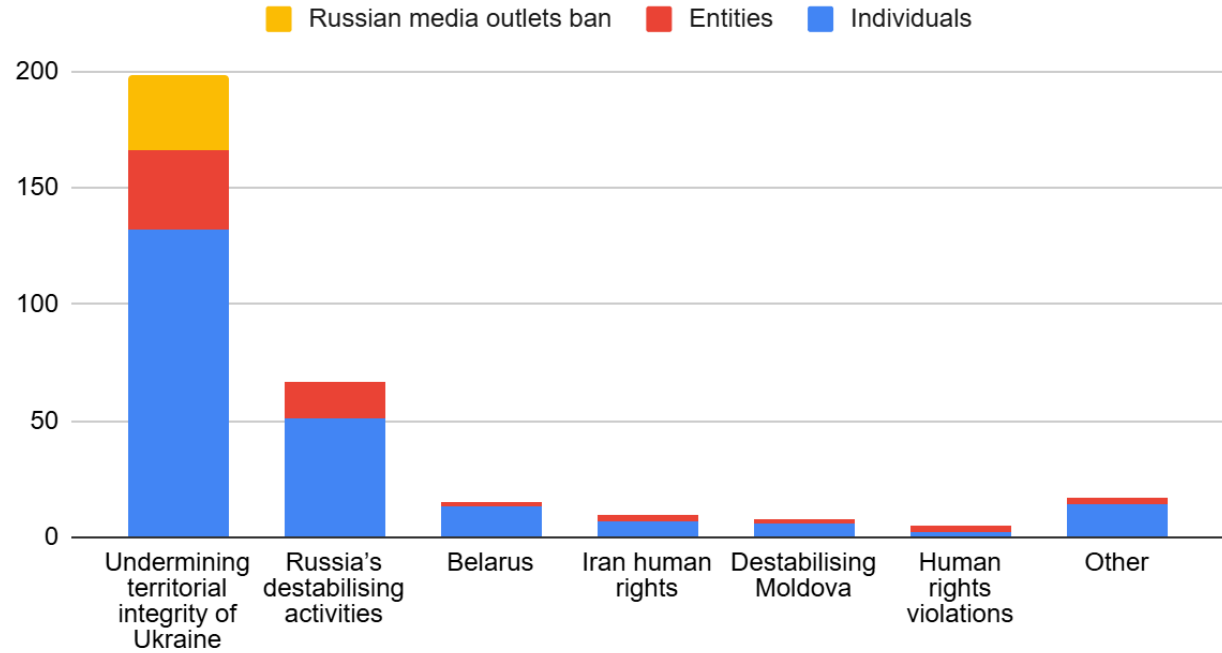


Figure 23: EU sanctions related to propaganda, disinformation, and information manipulation across sanctions’ regimes.

The small volume of legal sanctions being proposed and actioned for any state-linked FIMI, disinformation and/or propaganda activity, for states other than Russia, tells us something about the distribution of countermeasures by Western governments. Interventions grounded in law of this kind can only be leveraged by nation states, but they appear to be deployed quite narrowly, but repeatedly. Far more of the work of attempting to exert control over FIMI related harms is conducted through interventions performed by social media platforms.

It is also worth reiterating that there are a number of increasingly important countermeasures where we are lacking almost any adequate data to assess their effectiveness. For example, VIGINUM are influential proponents of adopting ‘technical attacks’ upon the material technological infrastructures of adversaries that are used to support their malign influencing activities. It is likely that there is a lot of this kind of work taking place, but for obvious reasons, there is a reluctance amongst practitioners, to render it too visible. Likewise, we know that around sensitive election events, adjustments are made by platforms to decrease the visibility of problematic political content issued by actors with an interference agenda. But we do not have a precisely calibrated picture of what the effects of such measures are.

Publicising technical indicators

As part of the search to deepen and widen the knowledge base about threat actors and enabling standardised ways to share information within the counter-FIMI community, some key organisations use OpenCTI, a platform initially developed for Cyber Threat Intelligence. For example, French government agency for protection against foreign digital interference (VIGINUM) uses OpenCTI and has developed a French version of DISARM, sharing its findings on GitHub.

The graph summarises the amount of domain names VIGINUM has linked to defined campaigns where they have published their lists (VIGINUM 2026).

Domains and subdomains linked to selected FIMI campaigns by Viginum

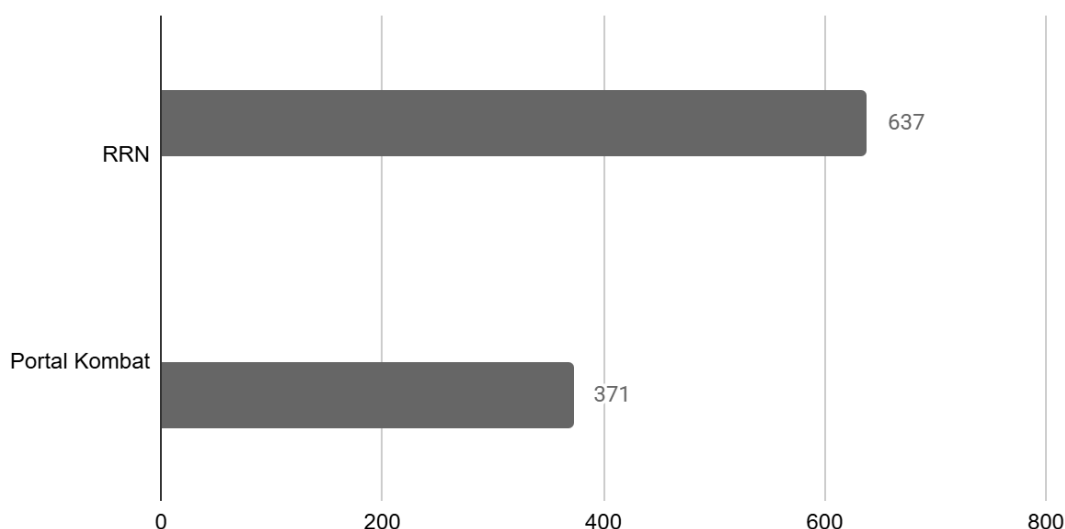


Figure 24: Number of domains and subdomains of selected FIMI campaigns. Source: VIGINUM.

5.4 Conclusion

This chapter has focused upon constructing an account of what is known, and what remains unknown about the efficacy of FIMI and disinformation countermeasures, both individually and collectively. Arguably the main conclusion to be drawn is that there are major knowledge gaps. In part, these stem from a tendency for research attention to be steered to the more exciting task of discovering and detecting bad actors, engaged in potentially harmful acts, than in constructing and conducting carefully designed and painstakingly executed evaluations of impact and effect. However, such gaps are also an artefact of decisions by governments, platforms and others to not invest in encouraging such work, and making the availability of valid and reliable data hard to come by.

Importantly, where evaluative work has been conducted, there is a strong suggestion that countermeasures have a propensity to induce unintended consequences and generate ‘blowback’. The most obvious form of which is how journalistic exposures of Russian information operations in Western outlets, have routinely got far more audience attention and publicity than was ever achieved by the original operatives.

A further reason for the patchy nature of our understanding about the efficacy of the kinds of countermeasures that have been covered herein, is a lack of a comprehensive model in terms of how they relate to one another and fit together. For in the absence of such a framing, it is difficult to ascertain both what they actually do or do not do, but also what it is legitimate to expect them to achieve in terms of influencing harmful FIMI actions.

6 Reflections and Next Steps

Framework production, as a way of summarising and guiding counter-FIMI interventions, has become an important part of the analytic landscape. Accompanying the growth and development of the FIMI defender community at local, national and international levels, there has been a proliferation of frameworks targeting different aspects of the analytic and response process. The aim of this research has been to develop a synthetic view across these frameworks, providing some form of evaluative assessment in terms of their intended and unintended contributions to improving the efficacy of efforts to manage and mitigate FIMI related harms.

In terms of their influence upon the main practices of the FIMI defender community, the evidence of significant impact is limited. And yet, the proliferation of frameworks suggests people perceive a need and user community for the kinds of synthesised and systematised knowledge that they distil. There are several reasons for these constrained practical impacts. Essentially, it reflects the intrinsically innovating and evolutionary nature of FIMI threats, which by design are often searching out the unprotected crevices and surfaces of digital life, in an effort to exploit and manipulate the lack of defences there. It is because of this that, in our proceeding discussion, we explicitly focused upon emergent blended information threats and the role of AI, as aspects where FIMI countermeasures frameworks are evidently straining to keep up and retain currency. After all, because they comprise synthesised and systematised knowledge about what is known at a particular moment in time, such frameworks have a mild, but inherent, retroactive tendency. Indeed, in this regard they may unintentionally function as 'engines' of FIMI innovation, as the authors of malign campaigns deliberately search for opportunities lying outside of the frameworks' purviews.

Reflecting this perspective, the research has generated a number of key headline findings, alongside lots more detailed empirically-led insights. These 'headlines' include:

- Although there are large numbers of contributions that 'self-describe' as constituting 'frameworks', a significant proportion of these do not meet the criterion for application of a formal conceptual definition of a framework. This suggests that the language of frameworks is often invoked and used as a 'selling' device, for describing processes and workflows.
- Only 31 of the 159 collected frameworks were retained after filtering for relevance and framework validity, and just two were mentioned by practitioners, pointing to a fragmented landscape with a proliferation of frameworks but limited practical adoption

- Interviews with counter-FIMI practitioner experts found that very few of them mentioned using any of the frameworks in practice, with the possible exceptions of DISARM and ABCDE. Albeit caveated by the relatively small number of interviews conducted, this suggests that frameworks have relatively limited influence in shaping real world practice. There is definitely an ‘over-supply’ of counter-FIMI frameworks.
- Looking across the multiple frameworks and how they interact with different phases of the analytic process for FIMI, clarifies how different models possess varying strengths and weaknesses, and centre different analytic tasks.
- No single framework covers the entirety of the FIMI response process and, therefore, cannot support practitioners at every stage of countering FIMI. A combination of different frameworks would be needed, as no one singular framework is currently adequate for an end-to-end process.
- The current situation of having multiple overlapping and somewhat ‘competing’ frameworks may hinder interoperability in terms of information-sharing and action coordination across different institutional and organisational actors engaged in FIMI defender work. It is highly likely, that with some regularity, different actors will be viewing the same presenting problem through divergent lenses, owing to the fact that their situated perspectives derive from their differing analytic framework preferences.
- In terms of the preferences of specific institutions and organisations for particular frameworks, these often reflect their particular responsibilities, interests and specialisms. This is a dimension that should not be forgotten, inasmuch as not all FIMI defenders are positioned in the same way in terms of their leverage-points and capabilities. Consequently, not all frameworks will be equally attractive or viable.
- For the purposes of this report, we have tried to reflect the previous point in two ways: (i) by highlighting how a comprehensive account of a presenting FIMI threat typically requires analysis of both a ‘technical’ and a ‘cognitive’ component, and how some specific national agencies are electing to centre one or other of these; and (ii) by providing an extended case study discussion of cyber-enabled influence operations as a species of ‘blended’ digital harm that clearly blurs any neat classificatory boundaries around a problem, and thereby illuminates how and why frameworks frequently struggle to capture and cope with these ‘blurry’ qualities.
- All the counter-FIMI frameworks reviewed, have a shared challenge in terms of striking an appropriate balance between ‘the specific’ and ‘the general’. That is, not being so ‘over-fitted’ to a particular situation that regularities and

patterns cannot be discerned, but equally not being so ‘generalised’ that they fail to obtain practical purchase on particular FIMI threats.

Framed by these headline findings, and induced by the more synthetic and overarching view our extensive empirical analysis has afforded, Figure 25 below provides a schematic visual representation of the architecture of FIMI and disinformation countermeasures. The triangle base structure depicts how there are three principal intervention points: (1) direct interventions in the organisation and conduct of the information operation itself; (2) actions designed to impact upon the technical or psychological enabling structures that support an operation; (3) the members of the potential audience for an operation in terms of the ‘attention market’ that it might be seeking to influence. The shape of the triangle is designed to convey the progressively increasing number of intervention points as you move from points 1 through 3.

Superimposed on the triangle in grey boxes are listed the key counter-measure interventions. They are positioned in such a way as to suggest where they are most likely to be directed relative to one another.

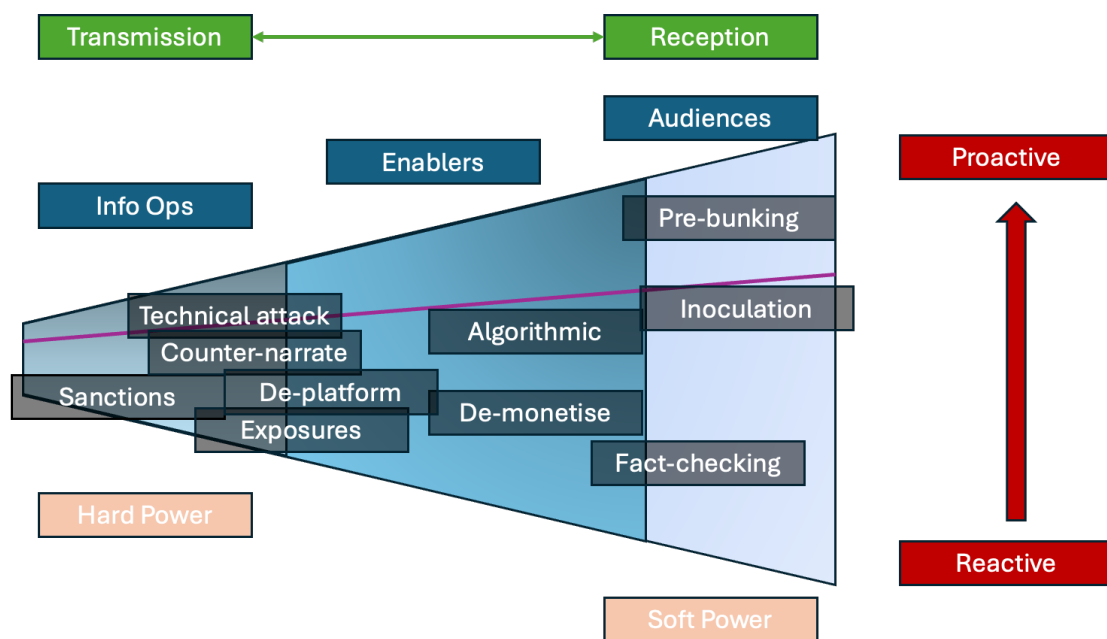


Figure 25: Visualisation of the architecture of FIMI and disinformation countermeasures.

Surrounding the main features of the figure are several key indicators. On the left-hand side, the red boxes scale how some (most) countermeasures are far more typically engaged in response to the discovery of hostile influencing activity, whereas others can be implemented in a more anticipatory, preventive and proactive fashion. This is further indicated by the purple line tracing how most countermeasures are framed by a reactive posture.

The green boxes across the top of the figure differentiate how some interventions are designed to intercept or condition the transmission or problematic content, as opposed to reception-side actions that work by shaping how any such content is interpreted and attended to by audience members. Finally, the orange boxes towards the bottom of the arrangement depict how some countermeasures are effectively coercive and require hard-power, whereas others are based upon a more persuasive mode of soft power.

The point of this visualisation is to both capture the essential complexity of the landscape of FIMI and anti-disinformation countermeasures, but also to start to nudge thinking towards a system-level of view of what effectiveness probably involves and requires. Especially when it comes to managing and degrading the effects of persistent, cross-platform foreign state information operations, countering their effects will typically rely upon more than one modality of intervention. As such, simplistic single item views of counter-measure efficacy are likely highly misleading.

Importantly, embedded within the above conceptual model are several key domains where our research suggests further development work is required, if we are to advance knowledge and practice with regards to counter-FIMI work. These can be summarized as follows:

- There is an identified need for any frameworks being adopted, supported and recommended to be adaptable and modular, in order to adjust to different contexts and organizations, and to FIMI methodologies that are fast paced and evolving. Especially notable in this regard are the affordances of AI enabled disinformation in terms of the speed, scope and scale of activity that these technologies enable.
- More attention needs to be directed towards the inter-operability challenge. It is clear that a full spectrum response posture to persistent and adaptive FIMI threats requires a range of capabilities and capacities, typically distributed across national government agencies, commercial platforms, investigative journalism outlets, researchers and so forth. More work is required to facilitate information sharing and combined actions, if disruption of threats and harms is to be both effective and efficient.
- Systematic evaluation and learning processes represent a significant gap in the current counter-FIMI landscape. This is despite pretty much all researchers and practitioners acknowledging the importance of these steps.

Together, these development needs and the forces shaping them reflect an important gap in our knowledge. That is, the FIMI defender community does not possess valid and reliable evidence-based knowledge in terms of how effective which countermeasures are, and in what context. Where this type of evaluative work has been conducted, typically through focused case study analyses, there is a strong suggestion that countermeasures can generate ‘blowback’ and unintended consequences.

This extends to the large number of frameworks and models that have been offered up by different authors and funded by an array of donors and sources. Based upon our extensive review of the field, we conclude that we do not know with any certainty which of the large number of frameworks available, actually improve the efficacy of the countermeasure's effort writ large. Indeed, there are plausible reasons for suggesting that the complexities of the current landscape may actually hinder interoperability in terms of information-sharing and joint-interventions, across the FIMI defender network.

References

- Amnesty International (2017) Azerbaijan: Activists targeted by 'government-sponsored' cyberattack, 10 March. Available at: <https://www.amnesty.nl/actueel/azerbaijan-activists-targeted-by-government-sponsored-cyberattack> (Accessed: 6 July 2026).
- Arcostanzo, F., Reyes, K.D., Wildon, J. and Balint, K. (2022) Effectiveness of the Sanctions on Russian State-Affiliated Media in the EU – An investigation into website traffic & possible circumvention methods. Institute for Strategic Dialogue, 6 October. Available at: <https://www.isdglobal.org/publication/effectiveness-of-the-sanctions-on-russian-state-affiliated-media-in-the-eu-an-investigation-into-website-traffic-possible-circumvention-methods-2/> (Accessed: 6 July 2026).
- Arribas, C. M., Arcos, R. & Gertrudix, M. (2025). Rethinking education and training to counter AI-enhanced disinformation and information manipulations in Europe: a Delphi study. *Cogent Social Sciences*, 11(1). <https://doi.org/10.1080/23311886.2025.2501759>
- Associated Press (2020) Facebook removes French, Russian accounts active in Africa. *The Independent*, 17 December. Available at: <https://www.independent.co.uk/news/facebook-removes-french-russian-accounts-active-in-africa-french-central-african-republic-military-russian-russia-b1775740.html> (Accessed: 6 July 2026).
- Aïmeur, E., Amri, S. & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis And Mining*, 13(1), 30. <https://doi.org/10.1007/s13278-023-01028-5>
- Bateman, J. and Jackson, D. (2024) Countering Disinformation Effectively: An Evidence-Based Policy Guide. Carnegie Endowment for International Peace, 31 January. Available at: <https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide> (Accessed: 6 July 2026).
- BBC News (2020) Qasem Soleimani: Iran seeks Trump's arrest over killing of general, 29 June. Available at: <https://www.bbc.com/news/world-middle-east-53223523> (Accessed: 6 July 2026).
- Binder, C. R., Hinkel, J., Bots, P. W. G. & Pahl-Wostl, C. (2013). Comparison of Frameworks for Analyzing Social-ecological Systems. *Ecology And Society*, 18(4). <https://doi.org/10.5751/es-05551-180426>

Brewster, T. (2026) US Strikes Killed Iranian Cyber Chiefs, But the Hacks Continued. *Forbes*, 17 March. Available at: <https://www.forbes.com/sites/the-wiretap/2026/03/17/us-strikes-killed-iranian-cyber-chiefs-but-the-hacks-continued/> (Accessed: 6 July 2026).

Buchanan, B. (2019) What to Make of Cyber Command's Operation Against the Internet Research Agency. *Lawfare*, 28 February. Available at: <https://www.lawfaremedia.org/article/what-make-cyber-commands-operation-against-internet-research-agency> (Accessed: 6 July 2026).

Buntain, C., Innes, M., Mitts, T. and Shapiro, J.N. (2023) 'Cross-Platform Reactions to the Post-January 6 Deplatforming', *Journal of Quantitative Description: Digital Media*, 1, pp. 1–44. Available at: <https://doi.org/10.51685/jqd.2023.004> (Accessed: 6 July 2026).

Bîzgă, A. (2024) Hacker leaks records of 20 million users of AI visual creation platform online. *Bitdefender Blog*. Available at: <https://www.bitdefender.com/en-us/blog/hotforsecurity/hacker-leaks-records-of-20-million-users-of-ai-visual-creation-platform-online> (Accessed: 6 July 2026).

Caltagirone, S., Pendergast, A. & Betz, C. (2013) The Diamond Model of Intrusion Analysis, U.S. Department of Defense. <https://apps.dtic.mil/sti/pdfs/ADA586960.pdf>

Council of the European Union (2019/2026) Council Regulation (EU) 2019/796 of 17 May 2019 concerning restrictive measures against cyber-attacks threatening the Union or its Member States, consolidated text as of 13 May 2026. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02019R0796-20260513> (Accessed: 6 July 2026).

Council of the European Union (2022) Council conclusions on Foreign Information Manipulation and Interference (FIMI), 11429/22, 18 July 2022. Available at: <https://data.consilium.europa.eu/doc/document/ST-11429-2022-INIT/en/pdf> (Accessed: 6 July 2026).

Council of the European Union (2026) Council Decision (CFSP) 2026/588 of 16 March 2026 amending Decision (CFSP) 2019/797 concerning restrictive measures against cyber-attacks threatening the Union or its Member States. *OJ L*, 2026/588, 16 March. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202600588 (Accessed: 6 July 2026).

Courchesne, L., Ilhardt, J. and Shapiro, J.N. (2021) Review of social science research on the impact of countermeasures against influence operations. *Harvard Kennedy School (HKS) Misinformation Review*. Available at: <https://misinforeview.hks.harvard.edu/article/review-of-social-science-research-on-the-impact-of-countermeasures-against-influence-operations/> (Accessed: 6 July 2026).

Courchesne, L., Ilhardt, J., & Shapiro, J. N. (2021). Review of social science research on the impact of countermeasures on influence operations. Harvard Kennedy School (HKS) Misinformation Review, Volume 2 (5).

CrowdStrike (2025) CrowdStrike 2025 Threat Hunting Report. Available at: <https://www.crowdstrike.com/explore/2025-threat-hunt-report> (Accessed: 6 July 2026).

CrowdStrike (2026) Famous Chollima Adversary Profile. Available at: <https://www.crowdstrike.com/adversaries/famous-chollima/> (Accessed: 6 July 2026).

Ehlen, T. et al. (2025-11-12) The European Democracy Shield: Commission publishes plan to empower democracies in the EU. <https://www.freshfields.com/en/our-thinking/blogs/technology-quotient/the-european-democracy-shield-commission-publishes-plan-to-empower-democracies-i-102luap>

Etudo, U., Whyte, C., Yoon, V., & Yaraghi, N. (2023). From Russia with fear: Fear appeals and the patterns of cyber-enabled influence operations. Journal of Cybersecurity, 9(1), tyad016. <https://doi.org/10.1093/cybsec/tyad016>

EU DisinfoLab, CheckFirst, VIGINUM, Cassini, Auswärtiges Amt, European External Action Service (EEAS) and DFRLab (2026) Building a common operational picture of FIMI: Using IMS to strengthen technical attribution and disruption. EU DisinfoLab, 15 January. Available at: <https://www.disinfo.eu/building-a-common-operational-picture-of-fimi/> (Accessed: 6 July 2026).

European Court of Auditors (2021) Disinformation affecting the EU: tackled but not tamed. Special Report 09/2021. Available at: <https://op.europa.eu/webpub/eca/special-reports/disinformation-9-2021/en/> (Accessed: 6 July 2026).

European External Action Service (2026) 4th EEAS Report on Foreign Information Manipulation and Interference Threats. EUvsDisinfo, 17 March. Available at: <https://euvsdisinfo.eu/eeas-4th-fimi-threat-report-march-2026/> (Accessed: 6 July 2026).

European External Action Service. (2023). 1st EEAS Report on Foreign Information Manipulation and Interference Threats: Towards a Framework of Networked Defence. <https://www.eeas.europa.eu/sites/default/files/documents/2023/EEAS-DataTeam-ThreatReport-2023.pdf>

European External Action Service. (2024). 2nd EEAS report on foreign information manipulation and interference threats: A framework for networked defence. https://www.eeas.europa.eu/sites/default/files/documents/2024/EEAS-2nd-Report%20on%20FIMI%20Threats-January-2024_0.pdf

European External Action Service. (2025). 3rd EEAS report on foreign information manipulation and interference threats: Exposing the architecture of FIMI operations. <https://www.eeas.europa.eu/sites/default/files/documents/2025/EEAS-3nd-ThreatReport-March-2025-05-Digital-HD.pdf>

European External Action Service. (2026). 4th EEAS report on foreign information manipulation and interference threats: Dismantling the FIMI House of Cards. https://www.eeas.europa.eu/sites/default/files/2026/documents/EEAS%204th%20Threat%20Report_web.pdf

Federal Bureau of Investigation (2022) Context and Recommendations to Protect Against Malicious Activity by Iranian Cyber Group Emennet Pasargad, PIN-20220126-001, 26 January. Available at: <https://www.ic3.gov/CSA/2022/220126.pdf> (Accessed: 6 July 2026).

Federal Bureau of Investigation (2023) Additional guidance on the Democratic People's Republic of Korea information technology workers. 18 October. Available at: <https://web.archive.org/web/20231018212155/https://www.ic3.gov/Media/Y2023/PSA231018> (Accessed: 6 July 2026).

Federal Bureau of Investigation, U.S. Department of the Treasury and Israel National Cyber Directorate (2024) New Tradecraft of Iranian Cyber Group Aria Sepehr Ayandehsazan aka Emennet Pasargad, Product ID: JCSA-20241030-001, 30 October. Available at: <https://www.ic3.gov/CSA/2024/241030.pdf> (Accessed: 6 July 2026).

Foote, C., Maness, R. C., Jensen, B., & Valeriano, B. (2020). 'Cyber conflict at the intersection of information operations: Cyber-enabled information operations, 2000–2016. In *Information Warfare in the Age of Cyber Conflict* (pp. 54–69). Routledge.

Franklin, M. and Torrey, M. (2024) Q3 2024 Adversarial Threat Report. Meta, December. Available at: <https://transparency.meta.com/sr/Q3-2024-Adversarial-threat-report/> (Accessed: 6 July 2026).

François, C. (2019) Actors, behaviors, content: A disinformation ABC. Highlighting three vectors of viral deception to guide industry & regulatory responses. Transatlantic Working Group on Content Moderation Online and Freedom of Expression. https://www.ivir.nl/publicaties/download/ABC_Framework_2019_Sept_2019.pdf

Gender Resource Centre (2023) Universal Periodic Review of Azerbaijan: 44th session of the United Nations Human Rights Council Working Group on the Universal Periodic Review. Available at: <https://uprdoc.ohchr.org/uprweb/downloadfile.aspx?filename=11554&file=EnglishTranslation> (Accessed: 6 July 2026).

Giles, K. (2023) Russian Cyber and Information Warfare in Practice, Chapter 4: Information Confrontation and Human Effects. Chatham House, 20 December. Available at: <https://www.chathamhouse.org/2023/12/russian-cyber-and-information-warfare-practice/04-information-confrontation-human-effects> (Accessed: 6 July 2026).

Giyasbayli, H. (2021) Hackers target prominent Azerbaijani feminist with pornography. OC Media, 11 March. Available at: <https://oc-media.org/hackers-target-prominent-azerbaijani-feminist-with-pornography/> (Accessed: 6 July 2026).

Gleicher, N. and Agranovich, D. (2020) Removing Coordinated Inauthentic Behavior from France and Russia. Meta Newsroom, 15 December. Available at: <https://about.fb.com/news/2020/12/removing-coordinated-inauthentic-behavior-france-russia/> (Accessed: 6 July 2026).

Gleicher, N. and Stamos, A. (2018) Removing Bad Actors on Facebook. Meta Newsroom, 31 July. Available at: <https://about.fb.com/news/2018/07/removing-bad-actors-on-facebook/> (Accessed: 6 July 2026).

Goh, D., Ling, R., Huang, L. & Liew, D. (2017). News sharing as reciprocal exchanges in social cohesion maintenance. *Information Communication & Society*, 22(8), 1128–1144. <https://doi.org/10.1080/1369118x.2017.1406973>

Google Threat Analysis Group (2024) Iranian backed group steps up phishing campaigns against Israel, U.S. Google, 14 August. Available at: <https://blog.google/threat-analysis-group/iranian-backed-group-steps-up-phishing-campaigns-against-israel-us/> (Accessed: 6 July 2026).

Google Threat Intelligence Group (2025) Adversarial Misuse of Generative AI. Google Cloud Blog, 29 January. Available at: <https://cloud.google.com/blog/topics/threat-intelligence/adversarial-misuse-generative-ai> (Accessed: 6 July 2026).

Greenberg, A. (2026) How 'Handala' Became the Face of Iran's Hacker Counterattacks. *Wired*, 13 March. Available at: <https://www.wired.com/story/handala-hacker-group-iran-us-israel-war/> (Accessed: 6 July 2026).

Gyimesi, B. (2025) Defending Democracy: Sanctions on Disinformation. Royal United Services Institute (RUSI), 12 June. Available at: <https://www.rusi.org/explore-our-research/publications/commentary/defending-democracy-sanctions-disinformation> (Accessed: 6 July 2026).

Huntley, S. (2024) TAG Bulletin: Q1 2024. Google, 6 March. Available at: <https://blog.google/threat-analysis-group/tag-bulletin-q1-2024/> (Accessed: 6 July 2026).

- Hénin, N. (2023) FIMI: Towards a European Redefinition of Foreign Interference. EU DisinfoLab. Available at: https://www.disinfo.eu/wp-content/uploads/2023/04/20230412_FIMI-FS-FINAL.pdf (Accessed: 6 July 2026).
- Innes, H. and Innes, M. (2023) 'De-platforming disinformation: Conspiracy theories and their control', *Information, Communication & Society*, 26(6), pp. 1262–1280. Available at: <https://www.tandfonline.com/doi/full/10.1080/1369118X.2021.1994631> (Accessed: 6 July 2026).
- Innes, M. (2014) *Signal Crimes: Social Reactions to Crime, Disorder and Control*. Oxford: Oxford University Press.
- Innes, M. and Ahonen, A. (2025) Attribution and Information Influence Operations: A Field Guide for Open-Source Investigators and Researchers. ADAC.io Deliverable D1.2. Cardiff University and ADAC.io. Available at: <https://adacio.eu/wp-content/uploads/2025/02/ADACio-D1.2-%E2%80%93Open-Source-Investigations-Field-Guide.pdf> (Accessed: 6 July 2026).
- Innes, M., Orpen, I. and Ashford, J. (2026) Estimating the Effects on the UK of State Information, Influence, and Interference Threats Using Digital Disinformation, Distortion, and Deception. CREST Research, 18 May. Available at: <https://crestresearch.ac.uk/resources/estimating-the-effects-on-the-uk-of-state-information-influence-and-interference-threats-using-digital-disinformation/> (Accessed: 6 July 2026).
- Insikt Group (2025) North Korea's Fraudulent IT Employment Scheme: A Cybersecurity Threat. Recorded Future, 13 February. Available at: <https://www.recordedfuture.com/research/inside-the-scam-north-koreas-it-worker-threat> (Accessed: 6 July 2026).
- Isenstadt, A. (2024) We received internal Trump documents from 'Robert.' Then the campaign confirmed it was hacked. Politico, 10 August. Available at: <https://www.politico.com/news/2024/08/10/trump-campaign-hack-00173503> (Accessed: 6 July 2026).
- Israel Defense Forces [@IDF] (2026) STRUCK: A large Iranian terror regime military compound in eastern Tehran. [X], 4 March. Available at: <https://x.com/idf/status/2029246755921871302> (Accessed: 6 July 2026).
- Kaplan, J. (2025) More Speech and Fewer Mistakes. Meta Newsroom, 7 January. Available at: <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/> (Accessed: 6 July 2026).
- Karlova, N. A., & Fisher, K. E. (2013). A social diffusion model of misinformation and disinformation for understanding human information behaviour.

Kristoff, J., Müller, M., Filastò, A., Resing, M., Kanich, C. and ten Oever, N. (2024) 'Internet Sanctions on Russian Media: Actions and Effects', *Free and Open Communications on the Internet*, 2024(1), pp. 1–12. Available at: <https://www.petsymposium.org/foci/2024/foci-2024-0001.php> (Accessed: 6 July 2026).

Lin, H. (2019). The existential threat from cyber-enabled information warfare. *Bulletin of the Atomic Scientists*, 75(4), 187–196. <https://doi.org/10.1080/00963402.2019.1629574>

Maristany de las Casas, P. and Reyes, K.D. (2025) The achilles' heel of EU's media sanctions: Reassessing RT Deutsch access ahead of Germany's 2025 Federal Election. *Institute for Strategic Dialogue*, 21 February. Available at: <https://www.isdglobal.org/digital-dispatch/the-achilles-heel-of-eus-media-sanctions-reassessing-rt-deutsch-access-ahead-of-germanys-2025-federal-election/> (Accessed: 6 July 2026).

Marsden, C., Meyer, T., & Brown, I. (2019). Platform values and democratic elections: How can the law regulate digital disinformation? *Computer Law & Security Review*, 36, 105373. <https://doi.org/10.1016/j.clsr.2019.105373>

McGinnis, M. D. & Ostrom, E. (2014). Social-ecological system framework: initial changes and continuing challenges. *Ecology And Society*, 19(2). <https://doi.org/10.5751/es-06387-190230>

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. the MIT Press.

Meta (2021) April 2021 Coordinated Inauthentic Behavior Report, May. Available at: <https://about.fb.com/wp-content/uploads/2021/05/April-2021-CIB-Report.pdf> (Accessed: 6 July 2026).

Meta (2021) November 2021 Coordinated Inauthentic Behavior Report, December (updated). Available at: <https://about.fb.com/wp-content/uploads/2021/12/November-2021-CIB-Report-UPDATED-stats.pdf> (Accessed: 6 July 2026).

Microsoft (2023) Microsoft Digital Defense Report 2023. Available at: <https://www.microsoft.com/en-us/security/security-insider/threat-landscape/microsoft-digital-defense-report-2023> (Accessed: 6 July 2026).

Microsoft Federal (2024) As cyber and influence operations converge ahead of the US 2024 election, government and industry can do more together. *Federal News Network*, February 2024. Available at: <https://federalnewsnetwork.com/federal-insights/2024/02/as-cyber-and-influence-operations-converge-ahead-of-the-us-2024-election-government-and-industry-can-do-more-together/> (Accessed: 6 July 2026).

Microsoft Threat Analysis Center (2024a) Russia, Iran, and China continue influence campaigns in final weeks before Election Day 2024, 23 October. Available at: <https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/MTAC-Report-Russia-Iran-and-China-continue-influence-campaigns-in-final-weeks-October-23-2024.pdf> (Accessed: 6 July 2026).

Microsoft Threat Analysis Center (2024b) Iran steps into US election 2024 with cyber-enabled influence operations, 9 August. Available at: <https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/5bc57431-a7a9-49ad-944d-b93b7d35d0fc.pdf> (Accessed: 6 July 2026).

Microsoft Threat Intelligence (2023a) Iran turning to cyber-enabled influence operations for greater effect. Microsoft, 2 May. Available at: <https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Iran-turning-to-cyber-enabled-influence-operations-for-greater-effect-05022023.pdf> (Accessed: 6 July 2026).

Microsoft Threat Intelligence (2023b) Iran responsible for Charlie Hebdo attacks. Microsoft Security Insider, 2 February. Available at: <https://www.microsoft.com/en-us/security/security-insider/threat-landscape/iran-responsible-for-charlie-hebdo-attacks/> (Accessed: 6 July 2026).

Microsoft Threat Intelligence (2024a) Staying ahead of threat actors in the age of AI. Microsoft Security Blog, 14 February. Available at: <https://www.microsoft.com/en-us/security/blog/2024/02/14/staying-ahead-of-threat-actors-in-the-age-of-ai/> (Accessed: 6 July 2026).

Microsoft Threat Intelligence (2024b) Microsoft shares latest intelligence on North Korean and Chinese threat actors at CYBERWARCON. Microsoft Security Blog, 22 November. Available at: <https://www.microsoft.com/en-us/security/blog/2024/11/22/microsoft-shares-latest-intelligence-on-north-korean-and-chinese-threat-actors-at-cyberwarcon/> (Accessed: 6 July 2026).

Microsoft Threat Intelligence (2025) Jasper Sleet: North Korean remote IT workers' evolving tactics to infiltrate organizations. Microsoft Security Blog, 30 June. Available at: <https://www.microsoft.com/en-us/security/blog/2025/06/30/jasper-sleet-north-korean-remote-it-workers-evolving-tactics-to-infiltrate-organizations/> (Accessed: 6 July 2026).

MITRE ATT&CK (n.d.) MITRE ATT&CK®. Available at: <https://attack.mitre.org/> (Accessed: 6 July 2026).

MSMT (2025). The DPRK's Violation and Evasion of UN Sanctions through Cyber and Information Technology Worker Activities. <https://msmt.info/view/save/2025/10/22/26294780-c396-407d-bb33-88afe988cd96->

The DPRK%E2%80%99s Violation and Evasion of UN Sanctions through Cyber and Information Technology Worker Activities (MSMT_2025_2).pdf

Murray, E. and Linvill, D. (2026) From Texas to Tehran: A Multilingual, IRGC-affiliated Influence Operation on X, Instagram, and Bluesky. Media Forensics Hub Reports, 10. Clemson University. Available at: https://open.clemson.edu/cgi/viewcontent.cgi?article=1009&context=mfh_reports (Accessed: 6 July 2026).

Nimmo, B. and Hutchins, E. (2023) Phase-based Tactical Analysis of Online Operations. Carnegie Endowment for International Peace, 16 March. Available at: <https://carnegieendowment.org/research/2023/03/phase-based-tactical-analysis-of-online-operations> (Accessed: 6 July 2026).

Nimmo, B., Agranovich, D. and Gleicher, N. (2022) Adversarial Threat Report: Q1 2022. Meta, April. Available at: https://about.fb.com/wp-content/uploads/2022/04/Meta-Quarterly-Adversarial-Threat-Report_Q1-2022.pdf (Accessed: 6 July 2026).

Nimmo, B., Zhang, A., Richard, M. and Hartley, N. (2025) Disrupting Malicious Uses of Our Models: An Update. OpenAI, February. Available at: <https://cdn.openai.com/threat-intelligence-reports/disrupting-malicious-uses-of-our-models-february-2025-update.pdf> (Accessed: 6 July 2026).

ODNI - Office of the Director of National Intelligence (2024) Statement from Director of National Intelligence Avril Haines on Recent Iranian Influence Efforts, ODNI News Release No. 17-24, 9 July. Available at: <https://www.dni.gov/index.php/newsroom/press-releases/press-releases-2024/3842-statement-from-director-of-national-intelligence-avril-haines-on-recent-iranian-influence-efforts> (Accessed: 6 July 2026).

ODNI, FBI and CISA (2024) Joint ODNI, FBI, and CISA Statement on Iranian Election Influence Efforts, 19 August. Available at: <https://www.fbi.gov/news/press-releases/joint-odni-fbi-and-cisa-statement-on-iranian-election-influence-efforts> (Accessed: 6 July 2026).

OpenAI (2025) Disrupting malicious uses of AI: June 2025. Available at: <https://openai.com/global-affairs/disrupting-malicious-uses-of-ai-june-2025/> (Accessed: 6 July 2026).

Pamment, J. (2023) 'How the Kremlin circumvented EU sanctions on Russian state media in the first weeks of the illegal invasion of Ukraine', Place Branding and Public Diplomacy, 19, pp. 200–205. Available at: <https://link.springer.com/content/pdf/10.1057/s41254-022-00275-1.pdf> (Accessed: 6 July 2026).

Pamment, J. & Tsurtssumia, D. (2025). Beyond Operation Doppelgänger: A Capability Assessment of the Social Design Agency. Swedish Psychological Defence Agency.

Pancevski, B. and Colchester, M. (2026) Iran Suspected of Creating Bogus Group to Claim Attacks in Europe, The Wall Street Journal, 24 March. Available at: <https://www.wsj.com/world/europe/iran-suspected-of-directing-attacks-in-europe-creating-bogus-group-to-claim-responsibility-7ebd8188> (Accessed: 6 July 2026).

Partelow, S. (2023). What is a framework? Understanding their purpose, value, development and use. *Journal of Environmental Studies and Sciences*, 13(3), 510–519. <https://doi.org/10.1007/s13412-023-00833-w>

Patrucic, M. and Bloss, K. (2021) Life in Azerbaijan's Digital Autocracy: 'They Want to be in Control of Everything'. OCCRP, 30 July. Available at: <https://www.occrp.org/en/investigation/life-in-azerbajans-digital-autocracy-they-want-to-be-in-control-of-everything> (Accessed: 6 July 2026).

Resecurity (2026) Iran War: Kinetic, Cyber, Electronic and Psychological Warfare Convergence. Resecurity Blog, 17 March. Available at: <https://www.resecurity.com/blog/article/iran-war-kinetic-cyber-electronic-and-psychological-warfare-convergence> (Accessed: 6 July 2026).

Reuters (2021) Hackers breach Iran rail network, disrupt service, 9 July. Available at: <https://www.reuters.com/world/middle-east/hackers-breach-iran-rail-network-disrupt-service-2021-07-09/> (Accessed: 6 July 2026).

Rewards for Justice (no date) Emennet Pasargad. U.S. Department of State. Available at: <https://rewardsforjustice.net/rewards/emennet-pasargad/> (Accessed: 6 July 2026).

Romano, S., Kerby, N., Schöler, M., Beraldo, D. and Rama, I. (2023) 'The impact of TikTok policies on information flows during times of war: Evidence of 'splinternet' and 'shadow-promotion' in Russia', *AoIR Selected Papers of Internet Research*. Available at: <https://spir.aoir.org/ojs/index.php/spir/article/view/13487> (Accessed: 6 July 2026).

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health education monographs*, 2(4), 328-335.

Schlager, E. (2019). A Comparison of Frameworks, Theories, and Models of Policy Processes. In *Theories of the Policy Process*, Second Edition (pp. 293–319). <https://doi.org/10.4324/9780367274689-10>

Schmidt, S. (2026) Over the past few years, North Korean (DPRK) nationals have been attempting to secure remote IT jobs with companies worldwide, particularly in the U.S. [LinkedIn], January. Available at:

<https://www.linkedin.com/feed/update/urn:li:activity:7407485036142276610/>
(Accessed: 6 July 2026).

Shah, N. (2026) Demystifying Iranian Cyber Operations in the U.S.-Iran Conflict. Center for Strategic and International Studies, 20 March. Available at: <https://www.csis.org/analysis/demystifying-iranian-cyber-operations-us-iran-conflict> (Accessed: 6 July 2026).

Shea, M. and Kalus, L. (2026) We spoke to the man making viral Lego-style AI videos for Iran. Experts say it's powerful propaganda. BBC News, 12 April. Available at: <https://www.bbc.com/news/articles/cjd8jrd1vnyo> (Accessed: 6 July 2026).

Shuliak, A., Homaniuk, O., Vozniuk, Y., Borysiuk, O., Kobets, V. & Zeleniuk, H. (2026). FIMI vs Disinformation: Impact on Digital Security and Public Order in the EU. *Veredas Do Direito Direito Ambiental E Desenvolvimento Sustentável*, 23(4), e234678. <https://doi.org/10.18623/rvd.v23.n4.4678>

Siman-Tov, D. (2025) The Trump Administration's Withdrawal from the Fight Against Foreign Interference—Strategic Implications. INSS Insight No. 2006. Institute for National Security Studies, 7 July. Available at: <https://www.inss.org.il/publication/trump-influence/> (Accessed: 6 July 2026).

Smith, O. (2026) GitLab Threat Intelligence Team reveals North Korean tradecraft. GitLab Blog, 19 February. Available at: <https://about.gitlab.com/blog/gitlab-threat-intelligence-reveals-north-korean-tradecraft/> (Accessed: 6 July 2026).

Sommer, W. and Izadi, E. (2024) Why newsrooms haven't published leaked Trump campaign documents. The Washington Post, 13 August. Available at: <https://www.washingtonpost.com/style/media/2024/08/13/iran-email-hack-republicans-media-response/> (Accessed: 6 July 2026).

Starks, C., Barnhart, M., Long, T., Lombardi, M., Pisano, J. and Revelli, A. (2024) Staying a Step Ahead: Mitigating the DPRK IT Worker Threat. Google Cloud Blog (Mandiant), 23 September. Available at: <https://cloud.google.com/blog/topics/threat-intelligence/mitigating-dprk-it-worker-threat> (Accessed: 6 July 2026).

Swedish Security Service (Säkerhetspolisen) (2024) Foreign power influence campaign carried out through breach of data security, 25 September. Available at: <https://sakerhetspolisen.se/ovriga-sidor/other-languages/english-engelska/press-room/news/news/2024-09-25-foreign-power-influence-campaign-carried-out-through-breach-of-data-security.html> (Accessed: 6 July 2026).

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207–232. [https://doi.org/10.1016/0010-0285\(73\)90033-9](https://doi.org/10.1016/0010-0285(73)90033-9)

U.S. Attorney's Office, District of Columbia (2024a) Charges and Seizures Brought in Fraud Scheme Aimed at Denying Revenue for Workers Associated with North Korea. 16 May. Available at: <https://www.justice.gov/usao-dc/pr/charges-and-seizures-brought-fraud-scheme-aimed-denying-revenue-workers-associated-north> (Accessed: 6 July 2026).

U.S. Attorney's Office, District of Columbia (2024b) Three IRGC Cyber Actors Indicted for 'Hack-and-Leak' Operation Designed to Influence the 2024 U.S. Presidential Election, 27 September. Available at: <https://www.justice.gov/usao-dc/pr/three-irgc-cyber-actors-indicted-hack-and-leak-operation-designed-influence-2024-us> (Accessed: 6 July 2026).

U.S. Department of Justice (2024) Three IRGC cyber actors indicted for 'hack-and-leak' operation designed to influence the 2024 U.S. presidential election. United States Attorney's Office, District of Columbia, 27 September. Available at: <https://www.justice.gov/usao-dc/pr/three-irgc-cyber-actors-indicted-hack-and-leak-operation-designed-influence-2024-us> (Accessed: 6 July 2026)

U.S. Department of Justice, Office of Public Affairs (2018a) Grand jury indicts 12 Russian intelligence officers for hacking offenses related to the 2016 election. Available at: <https://www.justice.gov/archives/opa/pr/grand-jury-indicts-12-russian-intelligence-officers-hacking-offenses-related-2016-election> (Accessed: 6 July 2026)

U.S. Department of Justice, Office of Public Affairs (2018b) Grand jury indicts thirteen Russian individuals and three Russian companies for scheme to interfere in the United States political system. Available at: <https://www.justice.gov/archives/opa/pr/grand-jury-indicts-thirteen-russian-individuals-and-three-russian-companies-scheme-interfere> (Accessed: 6 July 2026).

U.S. Department of Justice, Office of Public Affairs (2021) Two Iranian nationals charged for cyber-enabled disinformation and threat campaign designed to influence the 2020 U.S. presidential election, 18 November. Available at: <https://www.justice.gov/archives/opa/pr/two-iranian-nationals-charged-cyber-enabled-disinformation-and-threat-campaign-designed> (Accessed: 6 July 2026).

U.S. Department of Justice, Office of Public Affairs (2024) Fourteen North Korean Nationals Indicted for Carrying Out Multi-Year Fraudulent Information Technology Worker Scheme and Related Extortions, 12 December. Available at: <https://www.justice.gov/archives/opa/pr/fourteen-north-korean-nationals-indicted-carrying-out-multi-year-fraudulent-information> (Accessed: 6 July 2026).

U.S. Department of Justice, Office of Public Affairs (2025) Two North Korean Nationals and Three Facilitators Indicted for Multi-Year Fraudulent Remote Information Technology Worker Scheme that Generated Revenue for the Democratic People's Republic of Korea. 23 January. Available at:

<https://www.justice.gov/opa/pr/two-north-korean-nationals-and-three-facilitators-indicted-multi-year-fraudulent-remote> (Accessed: 6 July 2026).

U.S. Department of State, Office of the Spokesperson (2024) Alerting the world to RT's global covert activities, 13 September. Available at: <https://2021-2025.state.gov/alerting-the-world-to-rts-global-covert-activities/> (Accessed: 6 July 2026).

U.S. Department of the Treasury (2024a) Treasury designates Iranian cyber actors targeting U.S. companies and government agencies. U.S. Department of the Treasury, 23 April. Available at: <https://home.treasury.gov/news/press-releases/jy2292> (Accessed: 6 July 2026).

U.S. Department of the Treasury (2024b) Treasury Sanctions Iranian Regime Agents Attempting to Interfere in U.S. Elections, 27 September. Available at: <https://home.treasury.gov/news/press-releases/jy2621> (Accessed: 6 July 2026).

Van Der Meer, T. G., & Hameleers, M. (2025). Science and the crisis of trust. *Current Opinion in Psychology*, 67, 102202. <https://doi.org/10.1016/j.copsyc.2025.102202>

VIGINUM (2026) *Rapports Techniques de VIGINUM* [GitHub repository]. Available at: <https://github.com/VIGINUM-FR/Rapports-Techniques> (Accessed: 6 July 2026).

Vičić, J., & Gartzke, E. (2024). Cyber-enabled influence operations as a 'center of gravity' in cyberconflict: The example of Russian foreign interference in the 2016 US federal election. *Journal of Peace Research*, 61(1), 10–27. <https://doi.org/10.1177/00223433231225814>

Wardle, C. and Derakhshan, H. (2017), *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*, Council of Europe.

The Washington Post (2024) FBI concludes Iran tried to hack campaigns of Trump, Biden-Harris, 19 August. Available at: <https://www.washingtonpost.com/national-security/2024/08/19/iran-hack-trump-biden-harris-fbi/> (Accessed: 6 July 2026).

Watts, C. (2023) Iran responsible for Charlie Hebdo attacks. *Microsoft On the Issues*, 3 February. Available at: <https://blogs.microsoft.com/on-the-issues/2023/02/03/dtac-charlie-hebdo-hack-iran-neptunium/> (Accessed: 6 July 2026).

Whyte, C., & Etudo, U. (2025). Finding the thieves amongst the liars: Thinking clearly about cyber-enabled influence operations. *European Journal of International Security*, 1–19. <https://doi.org/10.1017/eis.2025.10016>

Annexes

Annex A: Framework Descriptions

Analysed Behavioural Frameworks

The *Information Disorder Framework* by Wardle and Derakhshan (2017), commissioned by the Council of Europe, is a conceptual framework that aims to analyse and better understand what they call “information disorder”. It provides a typology, where the type of information disorder is broken down into three types: misinformation (false, without harmful intent), disinformation (false, with harmful intent), and malinformation (true information weaponised to cause harm). These categories are mapped on three phases (creation, production, and distribution) as well as three elements (agent, message, and interpreter). The framework supports cross-sectoral diagnosis and policy formulation rather than case-level intervention, and it concludes with thirty-four recommendations addressed to technology companies, governments, media organisations, civil society, education ministries, and funders. The architecture has functioned as a reference point for much of the subsequent typological and policy literature in the field. This specific framework also shows a strong secondary strategic affinity through its policy recommendations.

The *Misinformation Receptivity Framework (MRF)* by Zmigrod, Burnell, and Hameleers (2023) instead focuses specifically on the interpreter, one of the three previously mentioned elements. It investigates why receivers accept some false claims and reject others, by conceptualising misinformation belief as a cognitive inference process, drawing on Bayesian models of cognition, of “reliability-weighting” between the reliability of new information and the reliability of one's own prior beliefs. Five major factors shape this weighting process: the receiver's worldview, ideology, and identity; characteristics of the source (notably credibility); characteristics of the information (plausibility, ideological leaning); the communication context; and psychological dispositions of the receiver. The framework further separates into two contextual stressors in communication environments with opposing effects: communicative “noise” elevates the weight of prior beliefs, therefore heightening susceptibility to worldview-congruent misinformation, whereas instability makes people rely less on their previous beliefs, hence opening them up to worldview-incongruent or extreme content. Therefore, the framework can be used as a base for creating testable hypotheses about who is susceptible to which kinds of misinformation, and under what conditions. The authors claim that the framework also applies to disinformation alongside misinformation, highlighting its importance for FIMI research.

The *Misinformation Recognition and Response Model (MRRM)* by Amazeen (2024) instead describes a psychological pathway from exposure, noticing inconsistencies, recognising misinformation, then engaging in cognitive coping strategies to experiencing cognitive, emotional, and behavioural outcomes. When individuals encounter misinformation or a corrective message, their reaction depends on aspects such as their personality, beliefs, motivation, current circumstances, and the way the message is presented. These factors influence whether they notice that the message conflicts with what they already know or believe. If people care enough about the topic and are motivated to think about it, they may recognise the information as misinformation. Once they do, they often respond in different ways, such as ignoring the message (avoidance), arguing against it (counterarguing), or reinforcing their existing views (attitude bolstering). These responses can then affect their thoughts, emotions, and behaviour. The main idea of the model is that recognising information as false changes how people interpret it. It is worth noting that the framework is utilised as guidance in empirical research on corrections, inoculation, and related interventions, rather than as an operational tool. The authors claim that the framework is also relevant to disinformation, not only misinformation, which makes it significant for the FIMI research context.

The *Online Misinformation Engagement Framework* by Geers et al. (2024) intersects with technical aspects such as platform mechanics and algorithms as it explores user engagement with online misinformation in a four-stage process: source selection (curating which accounts, sites and feeds populate one's information environment), information selection (deciding what to attend to or ignore), evaluation (judging accuracy and source credibility), and reaction (liking, sharing, commenting). For each stage, the relevant psychological mechanisms, observable traces, and corresponding intervention types are noted, including, for example: source-credibility labels at the source-selection stage, warning labels at the information-selection stage, and lastly accuracy prompts and social-norms interventions at the reaction stage. Geers and colleagues argue that source and information selection are the most under-researched stages and merit greater attention. The framework is explicitly receiver-centred and platform-agnostic, designed to organise interventions rather than to characterise operations themselves. While the authors acknowledge their focus on misinformation, they claim that their framework is broadly applicable to information sharing online overall, highlighting its relevance to FIMI research.

The *Disinformation Behaviour Framework* by Agarwal and Alsaedi (2021) builds on the social diffusion model of misinformation and disinformation by Karlova and Fisher (2013) and extends it to include creators alongside receivers of what they call “fake news”. They further integrate research on critical thinking and serendipity, also considering insight from library and information science. In their framework, they include: context; the intentions of creators and spreaders; the role of algorithms, bots, and social media affordances; the user's confirmation bias and filter bubble; and the user's judgment of source reliability on the outcomes of ignoring, using, misinforming, or disinforming. Crucially, the framework integrates a remediation layer that maps

countermeasures (critical thinking and action; advocacy, a "serendipity test" for the genuineness of unexpected information encounters; information and media literacy) onto the puncturing of the user's filter bubble. A simplified causal model accompanies the framework and is intended to guide empirical study designs and to bridge research and practitioner audiences. Through the consideration of bots and social media affordances, this framework also has technical secondary affinities.

The *Reframed Information Literacy Framework* by John and Tater (2025) is also built upon insights from library and information science, as it reframes information literacy by strengthening its ability to detect misinformation and disinformation through deeper critical evaluation. Drawing on a review of 175 sources, they argue that established information literacy models do not adequately equip users to detect misinformation and disinformation, particularly because they under-emphasise critical analysis of authority, reliability, and source motivation. The proposed reframed framework pairs standard information literacy components and skills (recognising the need, retrieval, evaluation, creation, ethical presentation) with a parallel column of elements specific to discerning manipulated content: sources of information, author and creator identification, additional relevant sources, a critical mindset, an information-checking tool, cross-checking of facts, the characteristics of offline versus online information, and the user's and creator's biases. The framework is directed primarily at literacy-infused science practitioners and educators.

The *Breakout Scale* by Nimmo (2020) is a framework designed to measure the observable spread and amplification of influence operations across communities, platforms, and media. It classifies operations into six categories based on whether they remain confined to a single platform/community, or "break out" across multiple platforms, communities, traditional media, and policy debate. "category one operations" remain within a single community on a single platform; "category two operations" spread within one platform or across platforms without breakout; "category three operations" achieve organic spread across multiple platforms; "category four operations" cross into mainstream media; "category five operations" are amplified by high-profile influencers; and "category six operations" trigger a policy response or include a call to violence. The model relies exclusively on verifiable, observable data available from the moment content is posted, enabling real-time categorization of influence operations without requiring insight into intent or internal objectives. It is therefore aimed at helping standardise triage, prioritisation, and reporting. Lastly, it also has operational secondary affinities as it can function as a supporting tool for operational researchers.

The *Systemic Framework for Disinformation on Social Media Platforms* by Recuero (2025) treats disinformation on social media as a complex adaptive communicative system made up of three interconnected dimensions: the elements involved (content types, actors, platforms and interfaces), the environment surrounding them (platform infrastructures, affordances, business models, and governance regimes within which disinformation circulates, as well as the broader political, cultural and economic

systems), and the dynamics that emerge as these components interact with each other (e.g., feedback loops, amplification, adaption). Together, these dimensions describe how disinformation is produced, shaped, amplified, and sustained within digital ecosystems. Crucially, Recuero argues that platforms function not as neutral hosts but as structurally coupled systems: they co-evolve with disinformation, providing algorithmic, economic, and infrastructural advantages that make problematic content more visible and more profitable. The framework is offered as a methodological scaffold for empirical research, with a summary grid of guiding questions for each dimension and an explicit call for situated, cross-platform, relational analyses that move beyond content-level identification toward an understanding of disinformation's resilience and adaptability. Due to the social media platform's focus and its political and platform governance recommendations, the framework has technical and strategic secondary affinities.

The *Misinformation Exposure-Trust Segmentation Framework* by Zeng et al. (2026) proposes an audience segmentation model on exposure to, and trust in, misinformation to identify audience subgroups to better understand susceptibility and support targeted communication strategies. It is grounded in the Stimulus-Organism-Response (S-O-R) framework (Mehrabian & Russell, 1974), a behavioural psychological theory explaining how individuals (S) evaluate affectively and cognitively (O) and then respond (R) to an external factor, as well as the Health Belief Model (HBM) (Rosenstock, 1974), which based on the individual's beliefs and attitudes predicts health-related behaviours. The proposed segmentation model uses two interrelated dimensions: individuals' exposure to misinformation and their trust in that misinformation, to classify audiences into four groups: misinformation-shielded (low exposure, low trust), misinformation-vulnerable (low exposure, high trust), misinformation-resistant (high exposure, low trust), and misinformation-accepting (high exposure, high trust). Empirically, the framework is demonstrated through a k-means cluster analysis of 819 US college students surveyed during the COVID-19 vaccination rollout in 2021, where the identified four subgroups of individuals differed significantly in political ideology, vaccine hesitancy, and several health beliefs (e.g., perceived efficacy, perceived safety, perceived symbolic threat). They also showed different patterns of association between health beliefs and vaccine hesitancy, suggesting that the cognitive pathways to behavioural outcomes are not uniform across the population. The framework is offered as a theory-based segmentation strategy for targeted intervention design: messaging that emphasises efficacy may suit the misinformation-vulnerable subgroup, while safety reassurance and transparency may be more appropriate for the misinformation-accepting subgroup. Although the empirical demonstration is bounded by its sample (US college students, one health domain), Zeng and colleagues frame the exposure-trust architecture itself as transferable and explicitly invite replication in larger and more demographically diverse populations. Due to the framework's communication implications for public health, secondary operational affinities can be found. While this framework is specifically limited to misinformation and, therefore, limited in its applicability to

FIMI, it nonetheless gives insight into the vulnerability of individuals, an important aspect to consider in FIMI research.

Analysed Technical Frameworks

The *DISARM Red Framework*, maintained by the DISARM Foundation, is a repository of adversarial tactics and techniques used by threat actors to manipulate the information environment, directly modelled on cybersecurity taxonomies, most explicitly MITRE ATT&CK. The framework originated in 2019 as the AMITT framework developed by the Credibility Coalition's MisinfoSec Working Group and was subsequently merged with MITRE's SP!CE fork. The framework's stated purpose is to give a heterogeneous community consisting of analysts, platforms, civil-society researchers, regulators, and policymakers a shared vocabulary for describing FIMI-relevant behaviour, supporting bulk tagging of incident reports and, via its integration with STIX and OpenCTI, machine-readable threat sharing at scale. Techniques of the framework were originally organised along a four-stage kill chain (from plan, prepare, execute to assess), but are now structured in an Asset-Action-Content matrix derived from the ABCDE model.

The *Online Operations Kill Chain Framework* (OOKC) by Nimmo and Hutchins (2023) under the Carnegie Endowment's Partnership for Countering Influence Operations is a unified phase-based model for the tactical analysis of malicious online operations. It draws methodologically on the Lockheed Martin Intrusion Kill Chain co-authored by Hutchins. It is explicitly designed to cover not only influence operations but also cyber-espionage, scams, harassment, human trafficking and terrorist recruitment, on the premise that "human-on-human" online operations share significant commonalities regardless of violation type. The framework consists of ten links (acquiring assets, disguising assets, gathering information, coordinating and planning, testing platform defences, evading detection, indiscriminate engagement, targeted engagement, compromising assets and enabling longevity) that are each broken down into techniques and procedures. It is platform-agnostic, modular (not every link is present in every case), restricted to directly observable aspects and designed to support investigation, cross-operation comparison and interorganisational information sharing.

Analysed Operational Frameworks

The *IIO Attribution Framework* by Palmertz et al. (2025), a deliverable of the ADAC.io project's Work Package 1, aims to provide a structured method for assessing and attributing Information Influence Operations across the heterogeneous community of actors (governments, civil society, the private sector, academia, NGOs, and journalism) engaged in this work. It builds on a previous paper published in 2022 by Pamment and Smith for the NATO Strategic Communications Centre of Excellence and the European Centre of Excellence for Countering Hybrid Threats. The framework uses a matrix that distinguishes three data-source tiers: open-source, proprietary, and classified. Across these sources, it evaluates technical, behavioural, and contextual evidence, which is further supplemented by a separate legal and ethical assessment. Furthermore, a traffic-light colour scheme allows analysts to signal the strength of indicators present. The

framework is modular, allowing organisations without access to classified intelligence to omit those rows, and the legal-ethical aspects can be handled by a separate team. The framework's application in the report is illustrated through three campaign case studies (Doppelgänger, LVU, and Paperwall) and highlighted through its adoption in modified form by Microsoft's Digital Threat Analysis Centre and Recorded Future's Insikt Group. The framework has technical (evidence-category architecture), behavioural (behavioural evidence aspect), and strategic (coordinating attribution across institutional actors) secondary affinities.

The *Diamond Model for Influence Operations Analysis*, developed by Recorded Future's Insikt Group, helps analysts integrate socio-political, technical, and psychological dimensions to understand the mechanics and drivers of malign campaigns. It was adapted from the Diamond Model of Intrusion Analysis from cybersecurity (Caltagirone et al., 2013). The framework maps influence operations across four corners: Influencer and Audience on the socio-political axis and Capabilities and Infrastructure on the technical axis, with narrative at the core. It is designed to be layered with other frameworks (e.g., kill chains, AMITT) applied to the corners of the diamond. The frameworks' use is illustrated through two specific case studies: a Chinese state-affiliated campaign reframing the origin of COVID-19 and a suspected Russian state-sponsored operation infiltrating US far-right platforms after the 2020 election. The framework has a technical affinity through its capabilities and infrastructure axis, and a behavioural affinity through its narrative-and-audience axis.

The *Methodology for Online Open Source Investigations* by Bellingcat and GLAN (2022) was specifically developed for online open-source investigations into incidents in Ukraine following the Russian invasion of February 2022, with the purpose of producing investigative findings on hostilities that are suitable for use in future accountability proceedings. The framework is organised into eight phases: systems and resources, briefings, categories of information, preparation, investigation, discovery and content gathering, verification and analysis, and roles and responsibilities. It assumes investigators are already trained in open-source techniques (e.g. geolocation, chronolocation)¹² and focuses instead on aspects required for legally admissible evidence (e.g., dedicated and encrypted devices, VPNs, virtual research identities, a fixed folder and citation structure, and preservation tooling). Additionally, it draws a clear distinction between examinable content (videos, satellite imagery, social media posts that can be analysed using OOSI techniques) and descriptive content (news articles, NGO reports) that can be referenced but not interrogated. The framework has both technical and strategic affinity, through its tooling and infrastructure requirements as well as its accountability-and-justice orientation.

SCAME is a propaganda assessment framework codified in US Army Field Manual FM 3-05.302 and the Soldier's Manual for Military Occupational Speciality 37F, designed

¹² Geolocation identifies where something occurred (via GPS, IP, or metadata); chronolocation identifies when (via timestamps or contextual clues like shadows or visible events). Both are used to verify the authenticity of images, videos, or claims.

for psychological operations personnel supporting a counterpropaganda effort within an area or joint operations area. SCAME provides a five-step sequential analysis of an individual propaganda product, analysing dimensions found in the initials of its name: Source, Content, Audience, Media, Effect. The source analysis identifies the actor, author, and authority behind the product, classifies it as white, grey, or black, and assesses the credibility of each source element with the target audience. The content analysis examines the message's objective, lines of persuasion, morale signals, involuntary information, biographical and economic content, internal inconsistencies, and the underlying intent of the source. The audience analysis distinguishes apparent, intermediate, unintended, and ultimate target audiences and asks why each was selected and how it responded. The media analysis, however, examines the choice of medium, dissemination frequency, placement, place of origin, technical characteristics, and transmission mode (overt or covert). Lastly, the effect analysis, seen as the most important and the most difficult step, evaluates the impact on the ultimate target audience through observed behavioural change, participant reports, and observer commentary. SCAME has an additional behavioural affinity through its content and audience steps and a strategic affinity through its embedding in counterpropaganda doctrine.

Countering Information Influence Activities: A Handbook for Communications Officers, created by the Swedish Civil Contingencies Agency (MSB), is written for communications professionals in Swedish public-sector organisations to help them identify whether they are observing influence activity and then to choose an appropriate response. The framework is split into three phases of the practitioner cycle: Prepare, Act, Learn. The Identification rests on three elements: strategic narratives (positive or constructive; negative or disruptive; oblique), target groups (general societal level, sociodemographic targeting, psychographic targeting), and a typology of six clusters of influence techniques (social and cognitive hacking, deceptive identities, technical manipulation, disinformation, maligning rhetoric, and symbolic acts) further unpacked into around two dozen techniques (e.g., bandwagon effects, sock puppets, deepfakes, fabrication, whataboutism and astroturfing). The framework also includes practical countermeasures, source-criticism guidance for online content, a step-by-step routine for detecting bots, and combinations of techniques (e.g., polarisation, laundering, provocation, flooding) that practitioners are most likely to encounter. The framework is grounded in the operational reality of Swedish communications officers but written for broader applicability across democratic public administration. It also has behavioural and strategic affinities through its techniques, typology and placement within national crisis preparedness.

Analysed Strategic Frameworks

The *Landscape of Hybrid Threats: A Conceptual Model* by Giannopoulos, Smith, and Theocharidou (2021) from the Joint Research Centre of the European Commission (JRC) and Hybrid CoE maps hybrid threats to provide a flexible blueprint for early detection, preparedness assessment, and response. It maps these threats across four pillars: actors (state vs. non-state) and their strategic objectives, tools applied by the actor (e.g., foreign direct investment, cyber operations), domains that were targeted (e.g., Infrastructure, Cyber, Space, Economy, Legal), and phases of activity (e.g., Interference, Influence, Operation, Warfare). The model is visual and adaptable, allowing countries to link actors, tools, domains, and activities to assess risk and develop countermeasures while also functioning as a shared conceptual vocabulary across member states.

The *Analytical Framework for Countering Hybrid Threats and Foreign Influence and Interference* by Palmertz, Weissmann, Nilsson, and Engvall (2024) from Lund University's psychological defence research institute provides a structured national Assess-Address-Evaluate cycle, operationalised as a guidebook posing four questions per dimension, intended as a starting point for comparing cases across the spectrum between peace and war. It is built explicitly for a nation-state perspective usable by governmental and non-governmental actors alike. The Assess phase of the framework covers first threat assessment, then vulnerability assessment, and lastly defence mechanism; therefore, mapping external threats against internal vulnerabilities and existing defences. The following phase then covers coordination and cooperation, and legal and policy frameworks, situating them within coordination structures and legal frameworks. Lastly, the evaluation phase concludes with evaluating the impact and effectiveness.

The *Strategic Framework to Counter Hybrid Action/Cross-Domain Strategies* Sweijs, Zilincik, Bekkers, and Meessen (2021) from the Hague Centre for Strategic Studies provides a strategic framework that differentiates between five broad strategic approaches an actor can use in response to hybrid threats: cooperation, persuasion, protection, coercion, and control. It expands on the concept of deterrence, claiming that deterrence alone cannot explain or guide modern counter-hybrid responses. These five approaches are arranged on a vertical escalation ladder running from least to most escalatory and combined with horizontal escalation across the DIMEL domains of state power: diplomatic, information, military, economic, and legal. The framework also specifies concrete measures within each domain and offers four assessments (cost-benefit, cross-domain orchestration, proportionality, and signalling) for moving deliberately from single-domain to cross-domain responses while managing the risk of unintended escalation.

The *Deterring Hybrid Threats framework* by Monaghan (2022) from Hybrid CoE offers a deterrence approach to address threats operating below the threshold of war, as the ambiguity and deniability of hybrid threats make them difficult to deter by

conventional means. The foundation of the framework consists of the three Cs: capability, credibility, and communication. Deterrence as a response is expanded by denial and punishment with restrictive, cumulative, extended, and absolute forms, together with cyber deterrence, reassurance, inducement, and the reinforcement of rules and norms.

The *Counter-Hybrid Response Framework (Ten Guidelines)* by Bertolini, Minicozzi, and Sweijts (2023) proposes a five-circular-stage, ten-step framework designed to guide states through preparation, detection and attribution, decision-making, execution, and evaluation of counter-hybrid responses. The five circular stages consist of a Preparation stage (articulating core interests, red lines, and building detection capability), a Detection and Attribution stage, a Decision-Making stage, an Execution stage, and an Evaluation stage (feeding lessons back into preparation). These stages are subdivided into ten concrete steps. The framework concentrates on deterrence by punishment while linking to denial, entanglement, and norms, and draws on the DIMEFIL spectrum of instruments (diplomatic, information, military, economic, financial, intelligence, and legal). It provides practical guidance for designing proportional, legitimate, and effective countermeasures in multinational contexts.

The *Report of the Attorney General's Cyber-Digital Task Force* by the US Department of Justice (2018) offers a framework that, applied to a singular institution, situates malign foreign influence operations within a broader set of cyber-enabled threats and distinguishes five distinct types aimed at elections and political life. These types consist of cyber operations against election infrastructure; cyber operations against political organisations, campaigns, and officials; covert operations to assist or harm such organisations; covert disinformation operations to sway public opinion and sow division; and overt efforts conducted through state media and lobbyists. Built on a law-enforcement logic of detection, attribution, deterrence, and disruption, its central innovation is a disclosure policy governing when and how the US Department exposes operations to victims and the public, bounded by commitments to political neutrality, the First Amendment, and the preservation of public trust.

The *ABCDE Framework* by Pamment (2020) for Carnegie Endowment provides a structured analytical tool to break down foreign interference into five components (each accompanied by guiding questions): Actor, Behaviour, Content, Degree, and Effect. The framework builds on and extends Camille François' "ABC" model (François, 2019) with a measure of degree and an explicit assessment of effect. It is intended to standardise analysis, structure reporting, and information requests to stakeholders, support the design of countermeasures, and diagnose which terminology best fits a given case and, in turn, to help assign institutional responsibilities across EU bodies. It is presented as adaptable to individual teams while preserving a common standard and is connected to companion instruments such as the RESIST toolkit.

The *ALERT Framework* by Desouza, Ahmad, Naseer, and Sharma (2020) is a risk-based analytical approach for understanding how information systems are

“weaponized”, compromising their goals and values while leaving the systems themselves technically intact, for political manipulation and disruption. ALERT, following its name, breaks analysis into: Actors, Levers, Effects, and Responses, enabling structured assessment of who is the actor behind the attack, the lever it exploits, what societal and political effects it produces, and what response options are available. This process is illustrated through documented cases of political disruption. The authors append a theoretical model of co-evolution in which attackers and defenders progressively build capability and sophistication in response to one another. The framework is pitched at security practitioners and policymakers seeking to analyse and respond to attacks on political systems enabled by information systems. Due to the framework's focus on information systems, secondary technical affinities can be found.

Combatting Disinformation - A Framework for Analysis by Rød, Pursiainen, and Eklund (2025) focuses not on the impact of disinformation itself but addresses societal resilience to disinformation and countermeasures. The framework sees societal resilience split into five domains: legal and regulatory, educational, political and governance, psychological and social psychological, and technological. Each domain then lists aspects and countermeasures relating to resilience to disinformation belonging to each of the domains (e.g., educational: literacy, communication). The authors also go a step further and propose a computable, customisable, indicator-based tool that would measure the maturity of a society's countermeasures and provide a practical basis for planning democratic responses. The framework, therefore, provides a conceptual and practical guide for policymakers, educators, and technologists.

The *Strategic Framework for Online Misinformation Counteraction* by Vadapalli, Doyal, Vanheusden, Binder, and Kuss (2026) synthesizes research on online misinformation and proposes a conceptual model linking misinformation, ideological polarization, and institutional distrust in a tri-directional relationship, where they mutually reinforce one another. Based on this analysis, it introduces a strategic framework outlining regulatory measures and broader actions to counter the spread and societal impact of misinformation. The work situates itself against cautions in the literature about alarmism, arguing for rigorous assessment of real-world outcomes rather than assumed causality. Due to the consideration of polarisation and distrust as variables, the framework has secondary behavioural affinities.

The *Conceptual Framework of Health Misinformation* by Alvarez-Galvez et al. (2024) is focused specifically on health misinformation during the COVID-19 pandemic and characterises misinformation and disinformation across six dimensions: sources, drivers, types, channels, audiences, and effects on health outcomes. Each of the dimensions is subdivided into categories (e.g., sources classified by expert versus non-expert profile and declared versus undeclared automated agents). Rather than prescribing interventions, it offers a structured means of analysing and categorising health misinformation, which the authors propose as a foundation for surveillance, public communication, and digital content moderation in future health emergencies. Although it was specifically built on the COVID-19 case, it is intended to be extended.

Due to the framework's consideration of audiences and effects, secondary behavioural affinities can be found.

RESIST 3, developed by James Pamment for the UK Government Communication Service (2025), guides government communicators through six sequential steps captured in its acronym: Recognise information threats, Early warning, Situational insight, Impact analysis, Strategic communication, and Tracking effectiveness. It is seen as a strategic hands-on toolkit rather than a conceptual model, oriented toward building organisational resilience and protecting priority audiences against information threats. Due to this applied nature, offering a playbook with stages, steps, workflows, and templates, the framework has secondary operational affinities.

The *DISRUPT Toolkit* by Alliance4Europe (2025) offers a whole-of-society approach to disrupting influence operations, split into three phases: preparation, disruption, and mitigation. Preparation focuses on collecting and analysing evidence, disruption involves reporting, engaging platforms and authorities, and escalating pressure, and mitigation is used if the operation continues or re-emerges. The framework supplies guides, templates, and step-by-step workflows for six categories of influence operation (e.g., coordinated inauthentic behaviour, deepfakes, impersonation, sanctions circumvention, and gendered smear campaigns) and is supported by a shared threat-intelligence and coordination platform. It was developed in consultation with the Counter Disinformation Network and builds on the DISARM framework, designed to enable coordination and operational timing, which they see as a major challenge.

The *framework for analysing disinformation laws and policies from a human rights perspective* by Global Partners Digital (2023) provides a structured set of tests (legality, legitimate aim, and necessity/proportionality) for evaluating whether any policy restricting information complies with international human rights standards. It operationalizes these principles through seven diagnostic questions that examine legal clarity, intent, independent oversight, proportionality, and the liability of individuals or intermediaries. Therefore, it functions as a guardrail on countermeasures rather than as a countermeasure in itself.

DigComp 2.2: The Digital Competence Framework for Citizens by Vuorikari, Kluzer, and Punie (2022) from the JRC provides more than 250 examples of knowledge, skills, and attitudes that help citizens engage confidently, critically, and safely with digital technologies, and new and emerging ones such as systems driven by artificial intelligence (AI). The framework, therefore, shifts the unit of resilience from the state and its institutions to the individual citizen. It defines digital competence as twenty-one competencies grouped into five areas: information and data literacy; communication and collaboration; digital content creation; safety; and problem solving. It is not a counter-disinformation framework as such but supports the long-term development of resilience and represents the capability-building end of the strategic spectrum. Due to its focus on developing the competencies of citizens, secondary behavioural affinities can be found.

Annex B: Collected Frameworks Table

Name of Framework	Authors	Source reference	Type of Framework	Selected for analysis?
Audience segmentation framework based on individuals' levels of exposure to and trust in misinformation	J. Zeng, J. Li , N. Yu, X. Cui, A. Xiao	https://journals.sagepub.com/doi/10.1177/13591053261419968	Behavioural Approach	Yes
"News Literacy" as a Framework for Critical News Reception	K. Ateşgöz	https://doi.org/10.1007/978-3-031-96720-7_3	Behavioural Approach	No
A conceptual framework and action agenda for health misinformation resilience	Y. Li	https://doi.org/10.1080/00909882.2025.2573952	Behavioural Approach	No
A digital literacy framework to mitigate misinformation in social media infodemic	L. Anthonysamy, P. Sivakumar	https://doi.org/10.1108/GKMC-06-2022-0142	Behavioural Approach	No
A framework for mitigating misinformation about the COVID-19 infodemic on social media	S. Manene, C. Hove, L. Cilliers	https://doi.org/10.4102/jamba.v15i1.1416	Behavioural Approach	No
A signal-detection framework for misinformation interventions	B. Gawronski, L. S. Nahon, N.L. Ng	https://doi.org/10.1038/s41562-024-02021-4	Behavioural Approach	No
A social diffusion model of misinformation and disinformation for understanding human information behaviour	N. A. Karlova and K. E. Fisher	https://informationr.net/ir/18-1/paper573.html	Behavioural Approach	No
A systemic framework for disinformation on social media platforms	R. Recuero	https://doi.org/10.1177/29768624251367199	Behavioural Approach	Yes
A unified typology framework for disinformation	E. Kapantai, A. Christopoulou, C. Berberidis & V. Peristeras	https://doi.org/10.1177/1461444820959296	Behavioural Approach	No
Advancing a typology of disinformation-susceptible publics using the motivation-attitude-knowledge framework	A. Krishna	https://doi.org/10.1080/1062726X.2021.1944155	Behavioural Approach	No
BEND	K. M. Carley	https://doi.org/10.1007/s10588-020-09322-9	Behavioural Approach	No
Bottom-Up Descriptive Framework	R. Perach, L. Joyner, D. Husbands, T. Buchanan	https://doi.org/10.1177/20563051231192032	Behavioural Approach	No
Break out Scale	B. Nimmo	https://www.brookings.edu/wp-content/uploads/2020/09/Nimmo_influence_operations_PDF.pdf	Behavioural Approach	Yes
Categorical Disinformation Detection Framework Through Thematic Analysis	A. Boko	https://doi.org/10.3390/journalmedia5040116	Behavioural Approach	No

Conceptual framework and case study of illegitimate expertise in climate change and COVID-19 misinformation and its implications for communication theory	C. W. van Eck , M. Hameleers , E. van der Goot	https://doi.org/10.1093/ct/qtaf033	Behavioural Approach	No
Countering misinformation about refugees and migrants: An evidence-based framework	D. Ghezelbash, F. Voon, L. Martin, M. Erdene-Ochir, S. Wodak	https://www.unsw.edu.au/content/dam/pdfs/law/kaldor/resource/2025-03-countering-misinformation-refugees.pdf	Behavioural Approach	No
COVID-19 health misinformation theoretical framework for intervention	S. Houlden, G. Veletsianos, J. Hodson, D. Reid, C. P. Thompson	https://doi.org/10.1108/HE-05-2021-0073	Behavioural Approach	No
Cultivating Ecological Literacy: A Critical Framework for Understanding and Addressing Mis- and Disinformation	B. Paris, G. Marcello, R. Reynolds	https://doi.org/10.1002/pa2.656	Behavioural Approach	No
Cultural Frameworks of Disinformation Responses	R. Cover, A. Haw, and J. D. Thompson	https://doi.org/10.1177/13678779221136881	Behavioural Approach	No
Deception as a Bridging Concept in the Study of Disinformation, Misinformation, and Misperceptions: Toward a Holistic Framework	A. Chadwick & J. Stanyer	https://doi.org/10.1093/ct/qtab019	Behavioural Approach	No
Designing a conceptual framework for misinformation on social media: a qualitative study on COVID-19	P. Bastani, S. M. Hakimzadeh, M. A. Bahrami	https://doi.org/10.1186/s13104-021-05822-2	Behavioural Approach	No
Disinformation behavior framework	N. K. Agarwal, F. Alsaeedi	https://doi.org/10.1108/AJIM-01-2021-0034	Behavioural Approach	Yes
Disinformation Detection: Developing a Categorical Framework Through Thematic Analysis	A. Boko	https://doi.org/10.3390/journalmedia5040116	Behavioural Approach	No
Disinformation/Misinformation Triangle	V. L. Rubin	https://doi.org/10.1108/JD-12-2018-0209	Behavioural Approach	No
Fake News as a Two-Dimensional Phenomenon: A Framework	J. L. Egelhofer and S. Lecheler	https://doi.org/10.1080/23808985.2019.1602782	Behavioural Approach	No
Gamification Framework for CALYPSO Disinformation Platform	C. Sotirakou, T. Paraskevas & C. Mourlas	https://doi.org/10.1007/978-3-031-05637-6_35	Behavioural Approach	No
Health Information and Misinformation: A Framework to Guide Research and Practice	I. Fridman, S. Johnson, J. Elston Lafata	https://doi.org/10.2196/38687	Behavioural Approach	No
Information Disorder Framework	C. Wardle & H. Derakshan, Council of Europe (2017)	https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html	Behavioural Approach	Yes
Inoculation theory/ framework	W. J. McGuire	https://psycnet.apa.org/doiLanding?doi=10.1037%2Fh0048344	Behavioural Approach	No

Layperson and Healthcare Professional Experiences of Medical Misinformation: A Thematic Framework Analysis	D. Robinson, S. Agius, A. Warren	https://doi.org/10.1007/s10728-025-00546-x	Behavioural Approach	No
Learner-centered educational frameworks to combat health mis/disinformation	A. Y. Sheng, M. Gottlieb, L. Welsh	https://doi.org/10.1002/aet2.10711	Behavioural Approach	No
Misinformation and the ACRL Framework	A. Faix and A. Fyn	https://digitalcommons.coastal.edu/lib-fac-pub/7	Behavioural Approach	No
Misinformation Exposure, Beliefs, Fear, and Information Avoidance via the Stimulus-Organism-Response Framework	X. Xu, C. A. Lin and H. Chen	https://doi.org/10.1177/10755470231207611	Behavioural Approach	No
Misinformation Recognition and Response Model (MRRM)	M. A. Amazeen	https://doi.org/10.1093/hcr/hqad040	Behavioural Approach	Yes
Online Misinformation Engagement Framework	M. Geers, B. Swire-Thompson, P. Lorenz-Spreen, S. M. Herzog, A. Kozyreva, R. Hertwig	https://doi.org/10.1016/j.copsyc.2023.101739	Behavioural Approach	Yes
Praxeological framework for studying disinformation	L. Thales	https://doi.org/10.1093/ct/qtad017	Behavioural Approach	No
Reframing the Information Literacy Framework to Identify Misinformation and Disinformation	K. John, B. Tater	https://doi.org/10.1080/0361526X.2025.2459765	Behavioural Approach	Yes
The Misinformation Receptivity Framework	L. Zmigrod, R. Burnell and M. Hameleers	https://doi.org/10.1027/1016-9040/a000498	Behavioural Approach	Yes
The Social Ecology of Health Beliefs and Misinformation Framework	J. C. Enyinnaya, A. A. Anderson, N. C. Kelp, M. Long, C. G. Duncan	https://doi.org/10.1016/j.vaccine.2024.01.001	Behavioural Approach	No
Toward an integrated framework for misinformation and correction sharing	W. Yu, B. Payton, M. Sun, W. Jia, and G. Huang	https://doi.org/10.1177/14614448221116569	Behavioural Approach	No
Typology framework for false information	S. J. Tsang	https://doi.org/10.1177/14648849241304380	Behavioural Approach	No
Countering information influence activities: A handbook for communications officers	Swedish Civil Defence and Resilience Agency	https://mpf.se/psychological-defence-agency/publications/archive/2024-03-01-countering-information-influence-activities	Operational Approach	Yes
Methodology for online open source investigations into incidents taking place in Ukraine since 24 February 2022	Bellingcat	https://www.bellingcat.com/app/uploads/2022/12/JA-Manual-for-PUBLICATION.pdf	Operational Approach	Yes
SCAME	U.S. Army	U.S. Army (December 2023) FM 3-05.301 - Psychological Operations Tactics, Techniques and Procedures, Chapter 11	Operational Approach	Yes
The Diamond Model for Influence Operations Analysis	Recorded Future	https://assets.recordedfuture.com/Whitepapers/diamond-model-influence-operations-analysis.pdf	Operational Approach	Yes

The IIO Attribution Framework	B. Palmertz, E. Isaksson, and J. Pamment	https://adacio.eu/a-framework-for-attribution-of-information-influence-operations	Operational Approach	Yes
A constitutional disinformation framework	C. Pinelli	https://www.rivisteweb.it/doi/10.1438/104048	Strategic Approach	No
A framework for analysing disinformation laws and policies from a human rights perspective	Global Partners Digital	https://www.gp-digital.org/wp-content/uploads/2023/06/A-framework-for-analysing-disinformation-laws-and-policies-from-a-human-rights-perspective.pdf	Strategic Approach	Yes
A Framework for Assessing the Role of Public Service Media Organizations in Countering Disinformation	M. Horowitz, S. Cushion, M. Dragomir, S. Gutierrez Manjon and M. Pantti	https://doi.org/10.1080/21670811.2021.1987948	Strategic Approach	No
ABCDE	J. Pamment, Carnegie Endowment for International Peace	https://carnegieendowment.org/research/2024/02/the-eus-role-in-fighting-disinformation-crafting-a-disinformation-framework	Strategic Approach	Yes
AI-driven disinformation: a framework for organizational preparation and response	E. Karinshak, Y. Jin	https://doi.org/10.1108/JCOM-09-2022-0113	Strategic Approach	No
ALERT	K. C. Desouza, A. Ahmad, H. Naseer, M. Sharma	https://doi.org/10.1016/j.cose.2019.101606	Strategic Approach	Yes
Analytical framework for countering hybrid threats and foreign influence and interference	B. Palmertz, M. Weissmann, N. Nilsson & J. Engval	https://mpf.se/download/18.7bb35f1818ec664dc7a4664/1712909687772/buidling-resilience-and-psychological-defence.pdf	Strategic Approach	Yes
Code of misinformation	ACMA	https://www.acma.gov.au/sites/default/files/2021-11/Adequacy%20of%20digital%20platforms%20disinformation%20and%20news%20quality%20measures.pdf	Strategic Approach	No
Comprehensive competency curriculum framework of media and information literacy	A. Grizzle, C. Wilson, R. Tuazon, C. K. Cheung, J. Lau, R. Fischer, D. Gordon, K. Akyempong, J. Singh, P. R. Carr, K. N. Stewart, S. Tayie, O. Suraj, M. Jaakkola, G. Thésée, C. Gulston	https://zu.elsevierpure.com/en/publications/media-and-information-literate-citizens-think-critically-click-wi-2/	Strategic Approach	No
Concept of psychological defence	J. Pamment, E. Isaksson	http://mpf.se/publikationer/publikationer/2024-10-28-psychological-defence-concepts-and-principles-for-the-2020s	Strategic Approach	No
Conceptual Framework of Health Misinformation During the COVID-19 Pandemic	J. Alvarez-Galvez, J. Carretero-Bravo, C. Lagares-Franco, B. Ramos-Fiol, E. Ortega-Martin	https://doi.org/10.2196/62693	Strategic Approach	Yes
Department of Justice Framework to Counter Malign Foreign Influence Operations	US Department of Justice	https://www.justice.gov/archives/ag/page/file/1076696/dl?inline=	Strategic Approach	Yes
DETERRENCE	Hybrid CoE	https://www.hybridcoe.fi/wp-content/uploads/2020/07/Deterrence_public.pdf	Strategic Approach	No

Deterrence framework	Sean Monaghan (Hybrid CoE)	https://www.hybridcoe.fi/publications/hybrid-coe-paper-12-detering-hybrid-threats-towards-a-fifth-wave-of-deterrence-theory-and-practice/	Strategic Approach	Yes
DigComp 2.2: The Digital Competence Framework for Citizens	JRC of the European Commission	https://publications.jrc.ec.europa.eu/repository/handle/JRC128415	Strategic Approach	Yes
DIME/DIMEFIL Framework	US Department of Defense	https://ndupress.ndu.edu/Portals/68/Documents/jfq-97/jfq-97_121-128_Rodriguez-Walton-Chu.pdf	Strategic Approach	No
Disinformation Campaigns and Resilience in Hybrid Threats Conceptual Model	W. Steingartner, D. Možnik, D. Galinec	https://doi.org/10.1109/Informatics57926.2022.10083447	Strategic Approach	No
DISRUPT framework	Alliance4Europe	https://alliance4europe.eu/counter-disinfo/disruption-framework	Strategic Approach	Yes
Expanded DigComp framework	C. M. Arribas, R. Arcos, M. Gertrudix	https://doi.org/10.1016/j.mex.2025.103401	Strategic Approach	No
G7 Resiliency and Response Framework to Counter TNR	Public Safety Canada and G7 RRM TNR Working Group (WG)	https://www.international.gc.ca/transparency-transparence/rapid-response-mechanism-mecanisme-reponse-rapide/2025-annual-report-rapport-annuel.aspx?lang=eng#framework	Strategic Approach	No
Information Manipulation Set (IMS) model	VIGINUM and EEAS	https://www.disinfo.eu/wp-content/uploads/2026/01/20260115_building-a-common-operational-picture-of-FIMI_01.pdf	Strategic Approach	No
Misinformation in Crises: A Conceptual Framework for Examining Human-Machine Interactions	T. Tran, P. Rad, R. Valecha, H.R. Rao	https://doi.org/10.1109/AI4G50087.2020.9311010	Strategic Approach	No
Multi-layered framework for building societal resilience	A. Fitz-Gerald and H. Padalko	https://doi.org/10.1007/978-3-032-05588-0_11	Strategic Approach	No
Online disinformation and political discourse: applying a human rights framework	K. Jones	https://apo.org.au/node/269261?utm_campaign=rss-all&utm_medium=RSS&utm_source=APO-feed	Strategic Approach	No
Public health prevention framework for managing information ecosystems	A. Ishizumi, J. Kolis, N. Abad, D. Prybylski, K. A. Brookmeyer, C. Voegeli, C. Wardle, H. Chiou	https://doi.org/10.1016/S2468-2667(24)00031-8	Strategic Approach	No
Resilience to Online Disinformation: A Cross-National Framework	E. Humprecht, F. Esser, P. Van Aelst	https://doi.org/10.1177/1940161219900126	Strategic Approach	No
RESIST (3)	UK Government, Government Communication Service (GCS); developed James Pamment	https://www.communications.gov.uk/wp-content/uploads/2025/10/RESIST-FINAL.pdf	Strategic Approach	Yes
Response-Impact Framework	R. M. Serrano & M. G. Sessa; EU DisinfoLab	https://www.disinfo.eu/publications/beyond-disinformation-countermeasures-building-a-response-impact-framework/	Strategic Approach	No

SCOTCH	Atlantic Council	https://www.atlanticcouncil.org/blogs/geotech-h-cues/scotch-a-framework-for-rapidly-assessing-influence-operations/	Strategic Approach	No
Societal Resilience to Disinformation Framework	B. Rød, C. Pursiainen, N. Eklund	https://doi.org/10.1007/s41125-025-00105-4	Strategic Approach	Yes
Sociotechnical Framework for Misinformation Policy Analysis	X. Zhu and S. Yang	https://www.taylorfrancis.com/chapters/oa-edit/10.4324/9781003231769-3/toward-sociotechnical-framework-misinformation-policy-analysis-xiaohua-zhu-shengnan-yang	Strategic Approach	No
Strategic framework for counteraction the impact of online misinformation on ideological polarisation and institutional distrust	S. K. Vadapalli, S. Doyal, F. J. Vanheusden, J. F. Binder, D. J. Kuss	https://doi.org/10.1007/s43545-026-01349-0	Strategic Approach	Yes
Strategic Framework to Counter Hybrid Action	T. Sweijjs, S. Zilincik, F. Bekkers & R. Meessen	https://euhybnet.eu/wp-content/uploads/2021/01/Framework-for-Cross-Domain-Strategies-against-Hybrid-Threats.pdf	Strategic Approach	Yes
The 4D framework: A holistic approach to countering climate change misinformation	J. Cook	https://www.taylorfrancis.com/chapters/edit/10.4324/9780429282348-54/4d-framework-john-cook	Strategic Approach	No
The 5 Ds	B. Nimmo	https://ai.umich.edu/blog-posts/spotting-fake-news-ben-nimmo-disinformation-misinformation-fake-news-teach-out/	Strategic Approach	No
The Counter-Hybrid Response Framework	M. Bertolini, R. Minicozzi and T. Sweijjs	https://hcss.nl/wp-content/uploads/2023/04/Guidelines-for-the-Deterrence-of-Hybrid-Threats-HCSS-2023.pdf	Strategic Approach	Yes
The EU Policy Framework to Counter Disinformation: Enabler or Inhibitor of Freedom of Expression?	P. Cesarini	https://doi.org/10.1007/978-3-032-00480-2_28	Strategic Approach	No
The Framework to Counter Foreign State Information Manipulation	US Government	https://2021-2025.state.gov/the-framework-to-counter-foreign-state-information-manipulation/	Strategic Approach	No
The Landscape of Hybrid Threats: A Conceptual Model	G. Giannopoulos, H. Smith, M. Theocharidou; (JRC and Hybrid CoE)	https://www.hybridcoe.fi/publications/the-landscape-of-hybrid-threats-a-conceptual-model/	Strategic Approach	Yes
The model of asymmetric warfare	Centre for Asymmetric Threat Studies (CATS)	https://www.diva-portal.org/smash/get/diva2:891091/FULLTEXT01.pdf	Strategic Approach	No
Whole-of-Society Framework for Countering Disinformation	J.D. Maddox, C. Gentzel, A. Levis	https://doi.org/10.4324/9781003425304-14	Strategic Approach	No
A Multi-Tool LLM Agent Framework Towards Verifiable Misinformation Detection	Z. Cui, T. Huang, C. E. Chiang, C. Du	https://doi.org/10.1145/3766918.3766948	Technical Approach	No
MMAdapt: Multi-source Multi-class Domain Adaption for Early Health Misinformation Detection	L. Shang, Y. Zhang, B. Chen, R. Zong, Z. Yue, H. Zeng, N. Wei, D. Wang	https://doi.org/10.1145/3589334.3648152	Technical Approach	No

A Data-Driven Framework for Coding the Intent and Extent of Political Tweeting, Disinformation, and Extremism	M. Hashemi	https://doi.org/10.3390/info12040148	Technical Approach	No
A deep-learning-based image forgery detection framework for controlling the spread of misinformation	A. Ghai, P. Kumar, S. Gupta	https://doi.org/10.1108/ITP-10-2020-0699	Technical Approach	No
A Dual-Stage Framework Integrating LLM Summarization and Knowledge Graph Embeddings for Health Misinformation Detection	Z. Song, T. Zhang, L. Zhang, G. Zhang	https://doi.org/10.1007/978-981-95-5640-3_12	Technical Approach	No
A Framework for Early Misinformation Detection on Social Media	S. Ghosh, P. Mitra	https://doi.org/10.1145/3603163.3609057	Technical Approach	No
A Framework for Enhancing Social Media Misinformation Detection with Topical-Tactics	B. E. Bagozzi, R. Goel, B. Lugo-de-Fabritz, K. Knickmeier-Cummings, K. Balasubramanian	https://doi.org/10.1145/3670694	Technical Approach	No
A Hybrid Framework for Identifying Fake News and Deceptive Information in Social Media with Mitigation Strategy	P. Sharma, M. Waris Khan	https://www.researchgate.net/profile/Mohd-Khan-289/publication/396280193_A_Hybrid_Framework_for_Identifying_Fake_News_and_Deceptive_Information_in_Social_Media_with_Mitigation_Strategy/links/68e55e2cfdca73694b5f702/A-Hybrid-Framework-for-Identifying-Fake-News-and-Deceptive-Information-in-Social-Media-with-Mitigation-Strategy.pdf	Technical Approach	No
A hybrid machine learning and simulation framework for modeling and understanding disinformation-induced disruptions in public transit systems	R. Talebi Khameneh, K. Barker, J. E. Ramirez-Marquez	https://doi.org/10.1016/j.res.2024.110656	Technical Approach	No
A multidimensional feature framework incorporating linguistic strategies	Y. Li	https://doi.org/10.1016/j.ipm.2024.104039	Technical Approach	No
A Personalised Adaptive Intervention Framework to Tackle Misinformation and Disinformation on Social Media	D. Barman, K. Koidl, S. M. Leiter, O. Carroll, A. Nussbaumer, B. Spillane, O. Conlan, G. Munnelly	https://doi.org/10.1109/SNAMS64316.2024.10883792	Technical Approach	No
A Preemptive Deep Reinforcement Learning Framework for Misinformation Defense	F. Yin, Z. Zhang, Z. Yu, C. Wu, J. Chen, Y. Wu	https://doi.org/10.3390/info17010067	Technical Approach	No
A Probabilistic Framework for Atomic Claim-Based Misinformation Detection	F. Sufi, M. Alsulami	https://doi.org/10.3390/math13111778	Technical Approach	No
A real-time large language model framework with attention and embedding representations for misinformation detection	F. K. Alarfaj, H. U. Khan, A. Naz, N. Almusallam	https://doi.org/10.1016/j.engappai.2025.113304	Technical Approach	No

A Semi-supervised Framework for Misinformation Detection	Y. Liu, Z. Boukouvalas, N. Japkowicz	https://doi.org/10.1007/978-3-030-88942-5_5	Technical Approach	No
A Systemic Framework for Evaluating Healthcare Misinformation in LLMs	N. A. Zahra & Z. Chin	https://doi.org/10.1007/978-3-032-00652-3_2	Technical Approach	No
A two-stage framework for Arabic social media text misinformation detection combining data augmentation and AraBERT	E. A. Mohamed, W. N. Ismail, O. A. S. Ibrahim, E. M. G. Younis	https://doi.org/10.1007/s13278-024-01201-4	Technical Approach	No
ABC-Verify: AI-Blockchain Integrated Framework for Tweet Misinformation Detection	T. H. Y. Zen, C. B. Hong, P. M. Mohan, V. Balachandran	https://doi.org/10.1109/SOLI54607.2021.9672392	Technical Approach	No
Advanced Persistent Threat Operational Line (APTOL)	A. Ahmad, J. Webb, K. C. Desouza and J. Boorman	https://doi.org/10.1016/j.cose.2019.07.001	Technical Approach	No
AI-Driven Framework for Evaluating Climate Misinformation and Data Quality on Social Media	Z. Shahbazi, R. Jalali and Z. Shahbazi	https://doi.org/10.3390/fi17060231	Technical Approach	No
AI-Powered Real-Time Misinformation Detection a Deep Learning Framework for Combating Fake News and Deepfakes	D. N. Rao, K. Mouneshwari, P. R. Kiran, Y. J. N. Kumar, A. Soy, T. Srihari	https://doi.org/10.1109/ICMCTC62214.2025.11196119	Technical Approach	No
Algorithmic enhancements to identify predictable components from users' data and a framework to detect misinformation in social media	G. Dixit & A. K. Kushwaha	https://doi.org/10.1080/2573234X.2022.2100834	Technical Approach	No
An Advanced Analytics Framework for Proactive Detection of Fake News to Minimize Misrepresentation Risks	R. Shiva Shankar, G. Mahesh, V. Maheswararao, N. Silpa, K V S Murthy	https://www.taylorfrancis.com/chapters/edit/10.1201/9781003529231-45/mitigating-misinformation-advanced-analytics-framework-proactive-detection-fake-news-minimize-misrepresentation-risks-shiva-shankar-mahesh-maheswararao-silpa-murthy?src=getFTR&utm_source=scopus&getft_integrator=scopus	Technical Approach	No
An Agent-Based Simulation Framework for Misinformation Susceptibility Test with LLMs: Insights from Psychological Factors	C. Han, X. Tang	https://doi.org/10.1007/978-981-95-4990-0_25	Technical Approach	No
An approach framework for disinformation detection	S. Chandra, B. Das	https://doi.org/10.1088/1742-6596/2161/1/012049	Technical Approach	No
An Attentive Fusion Framework For Multi-modal Misinformation Detection	S. Mehreen, A. Bhuiyan, E. Hossain	https://doi.org/10.1109/ICCIT64611.2024.11022377	Technical Approach	No
An Explainable AI Framework for Detecting Misinformation using Semantic NLP Features and Attention-based Models	M. J. Sujitha, A. Yalamanchili, M. A. S A, H. H. P, B. U. M. Babu, E. Govinda	https://doi.org/10.1109/ICSCDS65426.2025.11167535	Technical Approach	No

An Information Reliability Framework for Detecting Misinformation based on Large Language Models	V. S. P. Turaga, A. S. Namin	https://doi.org/10.1109/BigData62323.2024.10826052	Technical Approach	No
Belief Explorer: A Multi-Perspective AI Framework for Mitigating Disinformation Through Contextual Analysis and Socratic Dialogue	D. Manuel, C. Toxtli	https://doi.org/10.54941/ahfe1006743	Technical Approach	No
CIB Detection Tree	EU DisinfoLab	https://www.disinfo.eu/publications/coordinated-inauthentic-behaviour-detection-tree/	Technical Approach	No
Combating Misinformation in the Era of Generative AI Models	D. Xu, S. Fan, and M. Kankanalli	https://doi.org/10.1145/3581783.3612704	Technical Approach	No
Cyber Kill Chain	Lockheed Martin	https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html	Technical Approach	No
Deep explainable multimodal framework for detecting war-related misinformation in conflict zones in crisis communication	A. Nawaz	https://doi.org/10.1007/s12530-026-09795-0	Technical Approach	No
DISARM blue	DISARM Foundation	https://disarmframework.herokuapp.com/about	Technical Approach	No
DISARM red/ AMITT (Adversarial Misinformation and Influence Tactics and Techniques)	developed by MisinfoSec Working Group / Credibility Coalition in 2018; later adopted into DISARM Foundation	https://www.disarm.foundation/framework , https://disarmframework.herokuapp.com/about	Technical Approach	Yes
Dual-feature adaptive framework for multimodal disinformation detection	K. Yan, G. Liang, L. Wang, M. Sun, K. Zhao	https://doi.org/10.1093/comjnl/bxaf024	Technical Approach	No
Dynamic Simulation Framework for Disinformation Dissemination and Correction With Social Bots	B. Qiao, K. Li, W. Zhou, S. Hu	https://doi.org/10.18653/v1/2025.findings-emnlp.252	Technical Approach	No
Emotion aided multi-task framework for video embedded misinformation detection	R. Kumari, V. Gupta, N. Ashok, T. Ghosal, A. Ekbal	https://doi.org/10.1007/s11042-023-17208-6	Technical Approach	No
Environmental Uncertainty Perception (EUP) framework for the detection of misinformation and the prediction of its spread on social media	J. Lu, H. Zhang, Y. Xiao, Y. Wang	https://doi.org/10.2196/47240	Technical Approach	No
Framework for Detecting and Understanding Health-Related Misinformation	H. Padalko, V. Chomko, S. Yakovlev, D. Chumachenko	https://doi.org/10.3390/info16030175	Technical Approach	No
Framework of a Disinformation Game Based on the Narrative Warfare in Russo-Ukrainian War	J. Ono, T. Ogata	https://www.jstage.jst.go.jp/article/jmal/11/1/11_3/_pdf	Technical Approach	No

FusionNet: A bias-aware multimodal deep learning framework for detecting misinformation in noisy social media environments	M. Chakarverti, A. Goswami , R. K. Kaliyar	https://doi.org/10.1016/j.knosys.2025.115170	Technical Approach	No
GAAC: A robust misinformation detection framework via graph attention and adaptive contrastive learning	J. Hu, L. Zhu, M. Yang, B. Tang, S. Dai, C. Zhang	https://doi.org/10.1016/j.knosys.2026.115664	Technical Approach	No
GMD: A MULTIMODAL FRAMEWORK FOR AI-GENERATED MISINFORMATION DETECTION	J. Duan, W. Zhang, Z. Liu	https://doi.org/10.1109/ICIP55913.2025.11084741	Technical Approach	No
Hybrid AI Framework for Environmental Misinformation Detection	A. J. Mackenzie-Rivero, R. Martínez-Béjar & H. J. Vegas-Meléndez	https://doi.org/10.1007/978-3-032-11494-5_12	Technical Approach	No
Impact-risk index of disinformation	EU DisinfoLab	https://www.disinfo.eu/wp-content/uploads/2022/09/20220610_Index_ImpactAssessment_Final-1.pdf	Technical Approach	No
Infodemiology Framework for COVID-19 and Future Pandemics Using Artificial Intelligence to Address Misinformation and Disinformation	M. T. Ijab, M. S. Shahril & S. Hamid	https://doi.org/10.1007/978-3-030-90235-3_46	Technical Approach	No
KESHM: Knowledge Enabled Short HEalth Misinformation detection framework	F. Liu, Y. Li & M. Zuo	https://doi.org/10.1007/978-3-031-43412-9_22	Technical Approach	No
Large-Language-Model-Powered Agent Based Framework for Misinformation and Disinformation Research	J. Pastor-Galindo, P. Nespoli, J. A. Ruipérez-Valiente	https://doi.org/10.1109/MSEC.2024.3380511	Technical Approach	No
LIEALERT	S. Ghoush P. Mitra	https://doi.org/10.1145/3603163.3609057	Technical Approach	No
Machine Learning Framework for Detecting Fake News and Combating Misinformation Spread on Facebook Platforms	P. R. Y, K. Kunal, A. Palanisamy, S. K. Rajendran, Rupesh Gupta5, Vairavel Madeshwaren	https://doi.org/10.22399/ijcesen.1492	Technical Approach	No
MDAM3: A Misinformation Detection and Analysis Framework for Multitype Multimodal Media	Q. Xu, H. Du, S. Lukasik, T. Zhu, S. Wang, X. Yun	https://doi.org/10.1145/3696410.3714498	Technical Approach	No
Methodological framework for exposing disinformation campaigns against journalists and news outlets	developed and applied by the Croatian fact-checking organization Faktograf, investigative journalists at the German newspaper Taz, and researchers at the International Press Institute (IPI)	https://ipi.media/methodological-framework-for-exposing-disinformation-campaigns-against-journalists-and-news-outlets/	Technical Approach	No
MisD-MoE: A Multimodal Misinformation Detection Framework with Adaptive Feature Selection	M. Liu, K. Yan, Y. Liu, R. Fu, Z. Wen, X. Liu, C. Li	https://neurips2024-enlsp.github.io/papers/paper_36.pdf	Technical Approach	No

MITRE ATT&CK	MITRE Coperation	https://attack.mitre.org/	Technical Approach	No
Modeling misinformation spread for policy evaluation: a parsimonious framework	Y. Deng, R. Staelin	https://doi.org/10.1007/s11002-024-09724-8	Technical Approach	No
Multi-modal framework	A. Kumar, J. W. Taylor	https://doi.org/10.1016/j.ejor.2023.10.003	Technical Approach	No
Multi-Stage Model for Countering the Impact of Disinformation on the Cybersecurity System	M. Kryshtanovych, N. Lyubomudrova, H. Bondar, V. Motornyy, V. Kuchmenko	https://doi.org/10.18280/isi.280105	Technical Approach	No
MultiTec: A Data-Driven Multimodal Short Video Detection Framework for Healthcare Misinformation on TikTok	L. Shang, Y. Zhang, Y. Deng, D. Wang	https://doi.org/10.1109/TBDATA.2025.3533919	Technical Approach	No
RATE: Robust Adversarial Training and Temperature-scaled Ensemble Framework for Trustworthy Misinformation Detection	R. Sun, W. Hu, Q. Liu, R. Hong	https://doi.org/10.1145/3731715.3733418	Technical Approach	No
Real-Time Misinformation Detection with Cyclic Evidence-Based Framework	Z. Hu, L. Han, H. Jiang, X.-A. Ma, S. Lei, Y. Qiu, H. Niu, Z. Zhou, X. Wang	https://doi.org/10.1109/IJCNN64981.2025.11229147	Technical Approach	No
REVISE: A Framework for Paragraph-Level Misinformation Correction in Large Language Models	S. Tripathi, T. Jennifer, H. Griffith, H. Rathore	https://doi.org/10.1109/AITest66680.2025.00027	Technical Approach	No
Skeptik: A Hybrid Framework for Combating Potential Misinformation in Journalism	A. Fan, F. Lei, S. R. Corman, R. Maciejewski	https://doi.org/10.1145/3766891	Technical Approach	No
SynDy: Synthetic Dynamic Dataset Generation Framework for Misinformation Tasks	M. Shliselberg, A. Kazemi, S. A. Hale, S. Dori-Hacohen	https://doi.org/10.1145/3626772.3657667	Technical Approach	No
The Diamond Model of Intrusion Analysis	Sergio Caltagirone et al.	https://www.activeresponse.org/wp-content/uploads/2013/07/diamond.pdf	Technical Approach	No
The Online Operations Kill Chain	B. Nimmo and E. Hutchins	https://carnegieendowment.org/research/2023/03/phase-based-tactical-analysis-of-online-operations	Technical Approach	Yes
The Unified Kill Chain	P. Pols in collaboration with Fox-IT and Leiden University	https://www.unifiedkillchain.com/assets/The-Unified-Kill-Chain.pdf	Technical Approach	No
Threat model for disinformation attacks	S. Mirza, A. Abouzied, L. Begum, L. Niu, P. Papotti, S. Pardo, C. Pöpper	https://www.ndss-symposium.org/wp-content/uploads/2023/02/ndss2023_s657_paper.pdf	Technical Approach	No
Toward a Capture-Track-Respond Framework: A Survey of Data-Driven Methods for Countering LLM-Generated Misinformation	H. K. Kim, A. Skumanich	https://doi.org/10.1609/aaais.v5i1.35592	Technical Approach	No
TRAQ-M	G. A. Mack, S. G. Eick, M. A. Clark	https://doi.org/10.1109/AERO.2007.352801	Technical Approach	No

Unsupervised Misinformation Detection Framework to Analyze the Users using COVID-19 Twitter Data	A. Dhiman, D. Toshniwal	https://doi.org/10.1109/BigData50022.2020.9378250	Technical Approach	No
VaxMine: Vaccine Misinformation Detection in X using Cooperative Multimodal Framework	U. Naseem, A. G. Dunn, M. Khushi, J. Kim	https://doi.org/10.1145/3664647.3681422	Technical Approach	No

31.07 2026



ADAC.io Publication

Date of Publication: 31.07.2026

See more at adacio.eu

Funded by the European Union Horizon Europe Research and Innovation Program and the UKRI under the UK government's Horizon Europe funding guarantee. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union, the Horizon Europe Research and Innovation Program or UKRI. Neither the European Union, the Horizon Europe Research and Innovation Program, nor the UKRI can be held responsible for them.



**Co-funded by
the European Union**