Group of the Progressive Alliance of
Socialists & Democrats
in the European Parliament

June 2026

FROM TOOLS TO ACTION: A PROPOSAL FOR AN EU HYBRID THREAT RESPONSE PROTOCOL

A Research Paper prepared for the Progressive Group of Socialists & Democrats in the European Parliament

Krzysztof Bulski, European Policy Innovation Council

Author: *Krzysztof Bulski*, European Policy Innovation Council (EPIC)
Published by: *S&D Group in the European Parliament*

About the S&D Group

The Progressive Alliance of Socialists and Democrats (S&D) is the second largest political group in the European Parliament. The Group promotes social justice, equality, democracy, and sustainable development across Europe. Rooted in social-democratic values, the S&D works to deliver practical policies that improve people's lives, reduce inequalities, and build a fairer and more cohesive European society.

Disclaimer

The content of this publication does not represent the official position of the European Parliament or the S&D Group. Neither the S&D Group nor the European Parliament is liable for any use that may be made of the information contained herein.

This paper was funded by the S&D Group in the European Parliament for the purpose of an open policy discussion. No external funding, sponsorship, or in-kind contributions were involved in its drafting.

The analysis presented is independent. The paper is intended as background material to support informed debate among policymakers, experts, and stakeholders. It does not advocate for any specific policy instrument, technology, or regulatory model.

Table of Contents

fi

Executive Summary

Hybrid threats have become a permanent feature of Europe's security environment. They combine cyber operations, foreign information manipulation and interference (FIMI), economic coercion, infrastructure sabotage, and lawfare into sustained campaigns designed to erode democratic trust, polarise societies, and paralyse collective EU response — all while staying below the threshold of armed conflict.

Europe is not institutionally unprepared. Over the past decade, the EU has built a dense ecosystem of instruments: the Hybrid Toolbox, the FIMI Toolbox, the Cyber Diplomacy Toolbox, the EU Hybrid Fusion Cell, the European Centre of Excellence for Countering Hybrid Threats (Hybrid CoE), the Digital Services Act (DSA), NIS2, the Critical Entities Resilience (CER) Directive, the ProtectEU Internal Security Strategy (April 2025), and the Drone and Counter-Drone Action Plan (February 2026). Each addresses a real dimension of the problem.

Yet Europe continues to react to hybrid attacks rather than anticipate them. Instrument activation remains ad hoc. Coordination is slow and politically negotiated each time. Thresholds for escalation are undefined. Democratic oversight of emergency responses is inadequate. The result is a **governance gap**: not an absence of tools, but an absence of a predefined, democratic, and operational procedure for deploying them coherently under sustained pressure.

Two additional structural failures compound this governance gap. First, the threat landscape is accelerating beyond what any static institutional architecture can track. Autonomous AI agents now execute espionage campaigns with minimal human supervision. Deepfakes, synthetic media, and AI-generated disinformation now scale at negligible cost. Any protocol must therefore be **agile by design**: built not only for known threats but for the unknown ones, with built-in fast-learning mechanisms, sunset clauses, and regular stress testing.

Second, the systemic erosion of shared epistemic infrastructure — the common factual ground that democracy requires — cannot be addressed by institutional coordination alone. **Citizens must become active participants in democratic resilience**, both as informed actors and as contributors to a distributed defence architecture. This paper proposes a **European Cyber Guard** — a volunteer civic corps — alongside structural investments in publicly funded, editorially independent verification infrastructure and mandatory algorithmic transparency.

Key Findings at a Glance

The table below summarises the principal dimensions of Europe's hybrid threat challenge, the state of the EU's current response architecture, and the core proposals of this paper.

Dimension	Core Issue	Key Evidence & Constraints	Implications for Policy Design
1. Europe's Hybrid Threat Landscape	Multidimensional attacks — cyber, FIMI, sabotage, economic coercion — now AI-accelerated and deliberately calibrated to remain below armed-conflict thresholds.	• AI agents execute espionage with minimal human oversight (Anthropic/ China disclosure, late 2025). • Deepfakes and synthetic media scale at negligible cost. • Adversaries iterate faster than EU institutional architecture can track.	Any protocol must be **agile by design**: built-in sunset clauses, fast-learning loops, and standing red-team capacity. Response cannot be built solely on known threat patterns.
2. EU State of Play	A dense toolkit exists (Hybrid Toolbox, FIMI Toolbox, DSA, NIS2, CER, ProtectEU) but activation is ad hoc, coordination is slow, and democratic oversight is structurally inadequate.	• No predefined trigger mechanism for sustained hybrid pressure. • Each episode requires fresh political negotiation at bureaucratic pace. • Undefined escalation thresholds are the adversary's primary exploitation vector.	The governance gap is not a lack of instruments but a lack of **operational procedure**. Closing it requires procedural clarity, not new Treaty competences or additional agencies.
3. Hybrid Threat Response Protocol (HTRP)	A standing procedural framework defining activation criteria, trigger pathways, instrument sequencing, and democratic oversight — to deploy existing EU tools coherently and rapidly.	• Three graduated levels (Vigilance / Coordinated Response / Full Operational Response) with 24–48h decision windows. • Proportionality, time limits, and automatic review built in. • EP standing committee receives classified briefings within 72h of any activation.	Launch via Commission Communication + Council Conclusions; consolidate into an interinstitutional agreement in a second phase. No new competences needed — operates within existing Treaties.
4. European Cyber Guard	A volunteer civic corps — coordinated at EU level, rooted in local communities — to contribute to democratic resilience in the information domain.	• Institutional coordination alone cannot protect citizens' epistemic space. • 56% of Europeans lack basic digital skills (as low as 28% in Romania). • EDMO's network published over 1,600 fact-checks in the month before the vote, 15% of them targeting EU-related disinformation.	Pool existing EU funding streams (EDMO, EMIF, Horizon Europe, RRF digital components). Strict governance: FIMI-only scope, open-source methods, multi-stakeholder oversight board, full transparency. Modelled on Estonia's Cyber Defence League.

Dimension	Core Issue	Key Evidence & Constraints	Implications for Policy Design
5. Epistemic Infrastructure	Treating shared factual ground as **public infrastructure** — requiring deliberate institutional investment in publicly funded verification bodies and mandatory algorithmic transparency.	• Algorithmic recommendation systems demonstrably amplify divisive and factually dubious content. • The knowledge supply chain from sourcing to citizen receipt has no common standards. • Shared reality is a public good that has been allowed to decay.	Fund editorially independent verification bodies insulated from government direction. Mandate machine-readable algorithmic transparency with independent audit. Establish common standards for AI knowledge sourcing and attribution.

Conclusions and Policy Direction

1. Close the governance gap with a predefined, democratic protocol. The HTRP transforms the EU's toolbox from a collection of policies into an operational architecture — predefined, reviewable, democratically accountable, and agile enough to evolve with the threat.

2. Build societal resilience, not just institutional capacity. Hybrid threats target the minds and trust of individual citizens. The European Cyber Guard extends Estonia's Cyber Defence League model to the EU scale, building distributed civic capacity that no institution can manufacture on its own.

3. Treat epistemic infrastructure as public infrastructure. Shared reality requires the same deliberate investment as roads and energy grids. Publicly funded independent verification bodies, mandatory algorithmic transparency, and common standards for the AI-mediated information supply chain address the root cause, not just its symptoms.

4. Design for what is not yet known. Sunset clauses, structured learning protocols, red-team functions, and modular trigger architectures ensure the protocol remains fit for purpose as the threat landscape continues to evolve.

Together, these three elements — the HTRP, the European Cyber Guard, and epistemic infrastructure investments — constitute a response architecture commensurate with the scale of the challenge. Each is actionable. None requires new Treaty competences. All would represent a substantial advance on Europe's current position. The adversary is not waiting for Europe to get organised.

1. Introduction – The Governance Gap That Cannot Wait

Europe is under sustained hybrid pressure. Russian drone incursions and undersea cable sabotage in the Baltic. Chinese state-linked actors using compromised home routers across Europe as botnet infrastructure for cyber operations against democratic institutions. Deepfake campaigns targeting German elections. Coordinated disinformation amplified by recommendation algorithms, invisible to regulators and unchecked by platforms.

None of these attacks came with warning. None crossed a threshold that would trigger a coherent EU response. All of them succeeded — at least in part — because the EU's response architecture is built for a world where threats are discrete, attributable, and slow-moving. Today's threats are none of those things.

This paper argues three things.

First, the EU's established instruments are genuinely valuable — but the governance gap between tool availability and operational deployment has become the adversary's primary exploitation vector. Closing that gap requires a **Hybrid Threat Response Protocol (HTRP)**: a predefined, democratic, procedural framework for activating existing instruments coherently under documented hybrid pressure.

Second, the threat is not static. Autonomous AI systems now execute cyber campaigns without meaningful human oversight. The Anthropic disclosures of late 2025 and early 2026 — first a Chinese state-sponsored campaign using AI to attack Western infrastructure with minimal human supervision, then the discovery by Anthropic's Mythos Preview model of critical vulnerabilities in every major operating system — mark the beginning of a new era. Any protocol that does not build in structural agility and fast-learning capacity will be obsolete before it is implemented.

Third, the epistemic dimension of hybrid threats — the deliberate erosion of shared factual ground — cannot be solved by institutional coordination alone. Democracy requires citizens who can distinguish reliable information from manufactured confusion. Rebuilding that capacity demands new institutions, citizen empowerment, and a fundamental shift in how Europe thinks about information resilience: not as a content regulation problem, but as an infrastructure challenge.

2. Part I: Europe's Hybrid Threat Environment – Scale, Methods, and Strategic Intent

2.1 What hybrid threats are (and what makes them different now)

The European Union defines hybrid threats as "multidimensional, combining coercive and subversive measures, using both conventional and unconventional tools and tactics (diplomatic, military, economic, and technological) to destabilise the adversary. They are designed to be difficult to detect or attribute, and can be used by both state and non-state actors."

What makes hybrid threats distinctive is not any individual instrument but their **simultaneous and complementary deployment across multiple domains**, designed to exploit the thresholds of detection and attribution. The goal is rarely immediate disruption. It is cumulative erosion: weakening democratic trust, polarising societies, undermining regulatory authority, and testing the EU's ability to act collectively — while leaving the target unable to categorise the attack as an act of war.

The toolkit is wide: propaganda and coordinated disinformation; cyber espionage and infrastructure attacks; economic coercion and market manipulation; the weaponisation of migration; lawfare; physical sabotage of critical infrastructure; and the funding of extremist political movements. What is new — and what makes contemporary hybrid warfare qualitatively different — is the role of digital infrastructure in **lowering the cost, increasing the scale, and accelerating the speed** of all of them simultaneously.

A social media post that would once have required planting a story in a foreign newspaper now costs nothing and reaches millions in hours. A cyber intrusion that once demanded months of expert preparation can now be delegated to an AI agent running continuously. A deepfake video that would once have required a film production budget can now be generated on a consumer laptop in minutes.

2.2 The AI Acceleration: from human-directed to autonomous attack

The acceleration of AI capability has fundamentally altered the hybrid threat landscape — and Europe's response architecture has not kept pace.

In late 2025, Anthropic publicly disclosed that it had disrupted a Chinese state-sponsored group using its own AI systems to attack approximately 30 Western technology, finance, government, and critical infrastructure targets. The operation was executed with minimal human supervision. It was the first publicly documented AI-orchestrated espionage campaign. It will not be the last.

In early 2026, Anthropic disclosed that its Mythos Preview model had autonomously identified critical vulnerabilities in every major operating system and web browser. Advance previews were shared primarily with selected American enterprises and key U.S. stakeholders. European businesses and governmental institutions encountered significant difficulties even in establishing contact with Anthropic and obtaining timely access to the preview. This episode highlights a structural European vulnerability: limited influence over the development, testing, and responsible disclosure of frontier AI systems whose capabilities can expose systemic risks across the entire continent.

The implications are structural, not merely technical. Autonomous cyber agents can **execute in minutes what would previously have taken hours of expert human labour**. They can embed themselves in critical infrastructure and lie dormant indefinitely. They can operate across borders

and jurisdictions. And crucially — they may not stop when their initial mission is complete. Once deployed, they could pursue assigned objectives without the caution or restraint that human operators exercise when calculating escalation risks.

In practice, these "autonomous" systems are not self-originating: they are tools deliberately designed, configured, and initially deployed by human operators. Once released, however, they function with only limited ongoing oversight. This creates serious challenges for responsibility and accountability. It becomes difficult — and in some cases impossible — to establish whether a particular action was intended, authorised, or even foreseen by its creators, while adversaries exploit the ambiguity by claiming plausible deniability.

Simultaneously, AI has democratised disinformation at scale. China's Volt Typhoon, Flax Typhoon, and Violet Typhoon cyber units are using compromised consumer devices as botnet infrastructure, masking intrusions as ordinary internet traffic and making traditional IP-based defence methods ineffective. The "Storm-1516" operation in Germany's 2025 federal elections weaponised AI-generated deepfakes and synthetic media websites at a speed and cost unimaginable five years earlier.

The central lesson for protocol design is this: a response architecture built on known threat patterns cannot adequately counter an adversary whose primary advantage is novelty and speed. Agility, flexibility, and fast learning must be designed into the protocol itself — not added later as an afterthought.

2.3 Hybrid warfare in practice: Russia–Ukraine and beyond

Russia's hybrid campaign against Ukraine and Europe provides the richest available evidence base for understanding how these operations work in practice.

The Russia–Ukraine theatre has demonstrated the full spectrum. Before and during the 2014 Crimea annexation, Russia deployed a precisely coordinated combination of special forces operating in unmarked uniforms (providing plausible deniability), DDoS attacks against Ukrainian government infrastructure, targeted disinformation to Russian-speaking populations, economic coercion through gas supply manipulation, and psychological operations via mass SMS messaging to Ukrainian soldiers. The operation succeeded not because any single instrument was decisive, but because they were synchronised to paralyse Ukrainian decision-making at every level simultaneously.

Since Russia's full-scale invasion in 2022, the hybrid campaign has extended into Europe itself: acts of sabotage against undersea cables and pipelines in the Baltic; drone incursions over multiple EU member states; the weaponisation of migration flows; coordinated disinformation targeting support for Ukraine in EU societies. Following severe floods across Central Europe, Kremlin-linked disinformation actors disseminated false narratives designed to discredit national authorities, sow societal division, and in some instances shift blame toward Ukraine — a textbook application of information manipulation aimed at undermining public confidence in governance and support for EU energy independence and climate policies.

Hybrid warfare is by no means a Russian monopoly. The US/Israel–Iran conflict shows how other state actors have developed equally sophisticated hybrid strategies. Iran's hybrid toolkit combines missile and drone strikes by proxy forces, cyber operations, disinformation amplified via social media, and maritime harassment below the threshold of direct military confrontation. What is notable is that the line between state action and deniable proxy action has become deliberately blurred — precisely as a strategic feature, not a bug.

2.4 The epistemic crisis: when even state actors cannot be trusted

The deepest dimension of the hybrid threat is not the disruption of infrastructure. It is the disruption of the shared epistemic infrastructure that democratic governance requires.

Democracy operates on an assumption that citizens can, over time, converge on a sufficiently shared understanding of facts to make collective decisions. Hybrid operations — in combination with the structural incentives of algorithmic social media — are systematically dismantling that common ground.

The mechanism is not primarily lying, though lying plays a role. It is **manufactured uncertainty**. When every source can be questioned, every image can be fake, every statement can be attributed to AI generation, and even state actors issue contradictory or strategically shaped information, the result is not that citizens believe false things. It is that citizens stop believing anything, retreat into identity-based information communities, and become unable to distinguish coordinated manipulation from genuine complexity.

This is not merely a content moderation problem. No amount of fact-checking removes content fast enough to close the gap with the speed of AI-assisted disinformation production. It is not a media literacy problem alone. It is a **public infrastructure problem**. Shared reality — the epistemic common ground that democracy requires — is a public good, not a natural condition. It requires deliberate institutional architecture to sustain.

A critical observation: across both the Russia–Ukraine and Middle East theatres, even formally trusted state actors — including EU member states, NATO allies, and democratic governments — cannot be treated as unconditionally reliable information sources. Intelligence assessments have been wrong. Official attributions of cyberattacks have been contested or revised. In a high-stakes hybrid environment, every actor has powerful incentives to shape the narrative. Epistemic humility toward all sources — including authoritative ones — is not cynicism. It is basic hygiene. Rebuilding shared factual ground therefore cannot be entrusted solely to governments and agencies; it requires independent institutional architecture and active citizen participation.

3. Part II: The EU State of Play – Strong Tools, Fragmented Action

3.1 The EU's hybrid threat ecosystem

The European Union is not institutionally unprepared. Over the past decade, the EU has assembled a substantial toolkit for countering hybrid threats.

Strategic frameworks:

- The Joint Framework on Countering Hybrid Threats (2016), establishing 22 actionable proposals and creating the EU Hybrid Fusion Cell.
- The EU Hybrid Toolbox – an overarching framework of preventive, cooperative, stability-building, restrictive, and support measures.
- The FIMI Toolbox – for identifying and responding to foreign information manipulation and interference.
- The Cyber Diplomacy Toolbox – including the cyber sanctions framework (Council Regulation 796/2019) and ENISA's expanded mandate under the Cybersecurity Act.

Legal instruments:

- The **Digital Services Act (DSA)** – requiring platforms to assess and mitigate systemic risks, including from FIMI.
- **NIS2 Directive** – establishing high common cybersecurity standards across critical sectors.
- **Critical Entities Resilience (CER) Directive** – requiring risk assessments and resilience strategies for eleven critical sectors.
- The **EU Cybersecurity Act (2019)** – establishing ENISA as a permanent EU cybersecurity agency with expanded powers.

Coordination mechanisms:

- The Integrated Political Crisis Response (IPCR) mechanism.
- The EU Hybrid Fusion Cell within the EEAS, providing all-source analysis on hybrid threats.
- The European Centre of Excellence for Countering Hybrid Threats (Hybrid CoE) in Helsinki.
- EU–NATO cooperation frameworks, including 74 specific joint actions on hybrid threats.

This ecosystem is genuine. It represents a decade of serious institutional and regulatory investment. But instruments and operational capacity are not the same thing.

3.2 Recent developments: ProtectEU, the Drone Action Plan, and Hybrid Rapid Response Teams

ProtectEU (April 2025) – The European Internal Security Strategy launches a new governance framework for internal security, mainstreaming security considerations across EU legislation, strengthening Europol's mandate, and explicitly addressing hybrid threats through enhanced intelligence sharing, expanded ENISA powers, and a roadmap on lawful data access. It represents the most comprehensive update to EU internal security architecture in a decade.

The Drone and Counter-Drone Action Plan (February 2026) – Responding to a sharp increase in hostile drone incursions over EU member states, this plan addresses the full threat spectrum from negligent operators to state-sponsored hybrid operations. It covers detection (including leveraging

5G networks as distributed radar systems), response (including counter-drone emergency teams and joint procurement), and defence readiness. The plan explicitly acknowledges that drone threats "deliberately blur the boundaries between civilian and military domains" — a defining characteristic of hybrid warfare more broadly.

EU Hybrid Rapid Response Teams – First deployed to Moldova ahead of its September 2025 parliamentary elections, these teams provide on-the-ground support to member states and partners facing acute hybrid pressure. The Moldova deployment is a significant proof of concept, demonstrating that the EU can provide practical assistance before a crisis fully materialises.

3.3 Where the system breaks down: the operational gap

Despite this breadth of instruments, the EU's response to hybrid threats remains structurally reactive, fragmented, and slow. Understanding why requires looking at how the existing system actually operates — or fails to.

Activation is ad hoc. There is no predefined trigger mechanism for deploying EU instruments in response to sustained hybrid pressure. Each episode requires fresh political negotiation about which tools to use, in what sequence, by whom. In an environment where adversaries operate at AI speed, political negotiation at bureaucratic pace is structural disadvantage.

Coordination across instruments is slow. The FIMI Toolbox, the Cyber Diplomacy Toolbox, the Hybrid Toolbox, the CER Directive, the DSA, and IPCR were each developed for different threat scenarios and operate through different institutional chains. Under acute hybrid pressure, the EU must simultaneously activate multiple instruments across multiple institutions without a common operational logic governing their sequencing or interaction.

Thresholds for escalation are undefined. When does a series of cyber incidents constitute a coordinated hybrid campaign requiring EU-level response? When does disinformation rise above normal political contestation to warrant activation of the FIMI Toolbox? The absence of agreed escalation thresholds is not merely an administrative inconvenience: it is the adversary's primary exploit. Campaign designers deliberately calibrate activities to stay below whatever threshold might trigger response.

Responsibility is dispersed without a single operational logic. Primary responsibility for countering hybrid threats lies with member states. But hybrid threats by definition target cross-border vulnerabilities. The result is horizontal fragmentation (across EU institutions) and vertical fragmentation (between EU and member state levels), with each layer able to pass coordination responsibility to the other.

Democratic oversight is inadequate. When emergency measures are activated in response to hybrid attacks, the European Parliament has limited structured capacity for scrutiny. A further critical gap concerns the EU's limited oversight of frontier AI capabilities developed by non-EU actors. The early 2026 Anthropic Mythos Preview episode is illustrative: critical findings on systemic vulnerabilities were shared primarily with selected American stakeholders, while European governmental bodies and companies faced significant barriers to obtaining equivalent information. No existing instrument currently bridges this governance deficit.

3.4 The civilian information domain: underfunded and outgunned

The EU's apparatus for countering FIMI is doing vital work under conditions of severe resource constraint.

EDMO's network of fact-checkers verified 487 false claims during the 2024 European Parliament elections — a significant achievement. But 487 verified debunks, however accurate, is a drop in the ocean. The European Media and Information Fund received €19.4 million in applications in 2022 but could fund only €5.7 million worth of projects. Meanwhile, Russia alone is estimated to spend hundreds of millions annually on influence operations. China's state-media ecosystem operates at a comparable scale. The resource asymmetry reflects a structural difference between adversaries who treat information warfare as a core strategic function and an EU that still treats it as a peripheral communications issue.

StratCom East's monitoring of Russian disinformation, ENISA's coordination on cyber incidents, and the Hybrid CoE's research and exercises are all essential. But they are coordinated across a dispersed landscape of bodies without systematic operational integration, and none of them has the mandate or resources to address the underlying structural failure: the erosion of the shared epistemic infrastructure that would make citizens resilient regardless of which specific disinformation narrative they encounter.

4. Part III: From Fragmentation to Preparedness – What Can and Should Be Done

4.1 The Hybrid Threat Response Protocol: purpose and added value

The HTRP is proposed as a standing procedural framework designed to rapidly and coherently activate existing EU instruments. Its primary added value is procedural clarity: it organises how current tools are triggered and sequenced without creating new legal competences or emergency powers. By establishing a predictable, rules-based system, the HTRP aims to make the EU's crisis response intelligible to the public, thereby maintaining democratic legitimacy and trust.

The core proposal of this paper is the establishment of an EU **Hybrid Threat Response Protocol (HTRP)**: a standing procedural framework to coordinate the rapid and coherent activation of existing EU instruments in response to documented hybrid campaigns.

The HTRP would not create new competences, agencies, or emergency powers. It would operate fully within existing Treaties and secondary legislation. Its added value is **procedural clarity**: organising how existing tools are triggered, sequenced, and overseen.

The HTRP would define:

- **Clear activation criteria** based on documented, coordinated hybrid activity — including thresholds for different levels of response, so that adversaries cannot exploit undefined grey zones.
- **Trigger pathways** involving member states, the Commission, and the EEAS, with defined timelines for decision-making under different scenarios (cyber attack on critical infrastructure; sustained FIMI campaign; combined hybrid campaign).
- **Structured interinstitutional coordination** across Commission directorates, EEAS, ENISA, Europol, EDMO, and relevant member state authorities, through a standing operational coordination cell rather than ad hoc crisis formation.
- **Accelerated and synchronised instrument deployment**, with pre-authorised sequencing of the Hybrid Toolbox, FIMI Toolbox, Cyber Diplomacy Toolbox, and DSA enforcement mechanisms.
- **Democratic oversight** mechanisms, including systematic reporting to a standing European Parliament committee and mandatory after-action assessments.

Activation is triggered by the EU Hybrid Fusion Cell (or, in urgent cases, by a member state or the Commission directly) when there is credible, documented evidence of a coordinated hybrid campaign meeting at least two of the following indicators: (i) simultaneous or sequential actions across multiple domains (cyber, information manipulation, infrastructure sabotage, economic coercion, or lawfare); (ii) demonstrable foreign state or state-linked actor involvement; (iii) measurable impact on critical EU interests, democratic processes, or public trust; and (iv) intent to remain below the threshold of armed conflict while eroding collective response capacity.

Three graduated activation levels are defined:

Level 1 (Vigilance): triggered by emerging hybrid indicators or isolated incidents showing coordination potential. Response is limited to enhanced monitoring, information sharing, and preparatory measures. Decision timeline: 24 hours for initial assessment.

Level 2 (Coordinated Response): triggered by sustained or multi-domain activity with confirmed foreign attribution elements. Activates sequenced deployment of preventive and restrictive tools. Decision timeline: 48 hours from notification.

Level 3 (Full Operational Response): triggered by acute, high-impact campaigns (e.g., large-scale cyber attack on critical infrastructure combined with FIMI targeting public support for key EU policies). Enables parallel activation of the full suite of instruments. Decision timeline: 24 hours or less in cases of immediate threat to critical entities.

4.2 Sequencing existing tools: who does what, when

This section outlines a standing operational coordination cell to replace ad hoc crisis negotiations. It establishes a pre-authorised sequence for deploying various EU instruments — ensuring that the EU's response is accelerated, synchronised, and strategically aligned.

Once an activation level is declared, the HTRP replaces fragmented crisis negotiations with a standing operational coordination cell — hosted within the EEAS and building directly on the existing EU Hybrid Fusion Cell — comprising senior representatives from the Commission (relevant DGs), ENISA, Europol, the Hybrid CoE, and member state liaison officers. The cell operates under a pre-authorised operational playbook that defines clear roles, responsibilities, and sequencing timelines.

The pre-authorised sequence follows a logic of detection–containment–response–recovery:

1. Immediate detection and all-source analysis (0–12 hours): The Fusion Cell leads, integrating inputs from ENISA (cyber), EDMO/StratCom (FIMI), and national authorities. Output: confirmed hybrid campaign assessment and recommended activation level.

2. Rapid containment (12–48 hours): For cyber/infrastructure threats, NIS2 and CER Directives are activated by member states under Commission coordination, with ENISA providing technical support. For information manipulation, the FIMI Toolbox is mobilised by the EEAS, with parallel DSA enforcement actions by the Commission.

3. Synchronised operational deployment (48–96 hours): The coordination cell issues a single operational directive that sequences and deconflicts instruments — e.g., ProtectEU intelligence sharing runs in parallel with Drone Action Plan counter-measures and Hybrid Rapid Response Teams if on-site support is required.

4. Sustained response and de-escalation: Ongoing coordination ensures that economic, legal, or support measures are layered in without duplication, with built-in milestones for review.

This sequencing is modular: the playbook contains pre-scripted "play-cards" for common scenarios (e.g., cyber + FIMI against energy infrastructure; disinformation targeting elections). The standing cell meets daily during activation and reports transparently to the Council and European Parliament, eliminating the delays and political bargaining that currently allow hybrid campaigns to gain momentum.

4.3 Oversight and rights safeguards: democratic control of emergency response

To ensure accountability, the HTRP incorporates strict built-in safeguards, including proportionality requirements, time limitations, automatic review obligations, and provisions for fundamental rights protection. It mandates systematic reporting and scrutiny by a standing committee of the European Parliament.

Democratic accountability is not an afterthought in the HTRP — it is a core design feature. Every activation level incorporates strict, built-in safeguards that ensure proportionality, temporality, and respect for fundamental rights, while subjecting the entire process to systematic parliamentary scrutiny.

Proportionality is mandatory: the coordination cell must demonstrate that the chosen instruments and intensity are the least intrusive means capable of addressing the identified threat.

Time limitations are automatic — Level 1 measures expire after 14 days unless renewed; Level 2 after 30 days; Level 3 after 60 days — with each renewal requiring fresh evidence and explicit approval. Automatic review obligations apply at every renewal point, including an independent after-action assessment of effectiveness and rights impact.

Fundamental rights safeguards — drawn from the Charter of Fundamental Rights and existing DSA, NIS2, and GDPR provisions — require explicit data protection and freedom-of-expression impact assessments before any measure affecting the information space is deployed. No content removal or surveillance powers beyond those already foreseen in existing law are created.

Crucially, the Protocol mandates systematic reporting and scrutiny by a standing European Parliament committee. This body receives classified and unclassified briefings within 72 hours of any activation, conducts regular hearings on the conduct and necessity of measures, and has the right to request independent evaluations. By embedding these safeguards from the outset, the HTRP transforms hybrid threat response from a potential source of executive overreach into a model of accountable, rules-based resilience.

4.4 Agility by design: stress testing, review, red-teaming, sunset clauses

> *Recognising that threat landscapes evolve rapidly — especially with the advent of autonomous AI capabilities — the protocol must be structurally adaptable. This section mandates sunset clauses requiring active reauthorisation for all measures, structured fast-learning loops following any deployment, regular scenario stress testing, and a standing red team function to simulate adversary innovations and identify defensive gaps.*

The hybrid threat landscape in 2026 is qualitatively different from 2020, which was qualitatively different from 2016. The introduction of autonomous AI agents, synthetic media at scale, and the weaponisation of consumer IoT devices has changed both the attack surface and the attack economics in ways that no strategic document had fully anticipated. The next four years will bring further changes that we cannot yet predict with confidence.

This is not an argument for paralysis. It is an argument for designing the HTRP with agility as a core architectural principle.

Concretely, this means:

- **Mandatory sunset clauses** on all HTRP-activated measures, requiring affirmative reauthorisation rather than passive continuation. This prevents temporary emergency measures from becoming permanent features of the legal landscape.
- **Structured learning protocols**: each HTRP deployment, exercise, or incident should feed into a formal after-action review with standardised output format, shared across EU institutions and member states, and used to update the Protocol within a defined timeframe.
- **Red team function**: a standing capability within the HTRP operational structure to simulate adversary innovation against EU defences — specifically including AI-driven and autonomous attack scenarios — and identify gaps before adversaries exploit them.

- **Modular trigger architecture**: rather than a single activation threshold, the Protocol defines graduated response levels, each with pre-authorised instrument combinations, so that escalation and de-escalation can occur rapidly without requiring fresh political negotiation.
- **Assumption audits**: an explicit mechanism for periodically reviewing the Protocol's foundational assumptions — about adversary capabilities, attribution methods, member state capacities — and updating them based on current intelligence.

The adversary's primary advantage in hybrid warfare is asymmetric learning speed. They iterate; we institutionalise. Closing that gap requires building learning into the protocol's DNA, not treating it as a governance afterthought.

4.5 Societal resilience: the European Cyber Guard

Because institutional coordination alone cannot protect the minds of citizens, this section proposes the European Cyber Guard — a volunteer civic corps aimed at building distributed resilience. Modelled on existing frameworks like Estonia's Cyber Defence League, volunteers would focus on open-source fact-checking, digital literacy education, and early warning reporting. Operating under strict legal boundaries, this non-state entity would not monitor private communications, remove content, or police domestic political opinion.

The HTRP addresses the institutional layer of Europe's hybrid threat response. But institutional coordination alone cannot defeat adversaries who target the minds and trust of individual citizens. The EU needs a **distributed layer of civic resilience** that no institution can provide by itself.

Across Europe, models already exist that demonstrate the power of citizen participation in national resilience:

- **Estonia's Cyber Defence League** – a volunteer organisation integrated with the Estonian Defence Forces, comprising citizens with IT expertise who contribute part-time to national cybersecurity. It was instrumental in Estonia's response to the 2007 Russian DDoS attacks.
- **Finland's comprehensive security model** – a whole-of-society approach in which one in six citizens has received crisis or cyber training, treating civilian preparedness as a core component of national defence.
- **Poland's Territorial Defence Forces** – over 42,000 volunteer soldiers, 90% serving part-time, providing a distributed reserve force that supplements the professional military.

This paper proposes the establishment of a **European Cyber Guard**: a volunteer civic corps, coordinated at EU level but rooted in local communities, with training, mandate, and resources to contribute to democratic resilience in the information space. The core functions would include:

- **Fact-checking and OSINT analysis** – volunteers trained to identify, document, and report coordinated inauthentic behaviour and FIMI campaigns.
- **Digital literacy education** – community-level programmes raising citizens' capacity to navigate AI-generated content, recognising that only 56% of Europeans currently have basic digital skills, with some member states (including Romania) as low as 28%.
- **Early warning and reporting** – a distributed network of trained observers capable of identifying emerging disinformation campaigns before they achieve scale, feeding intelligence into EDMO and StratCom East.
- **Counter-narrative communication** – under appropriate oversight, volunteers could develop and disseminate fact-based content that undermines the credibility of manipulative narratives.

The European Cyber Guard would operate exclusively on publicly available information. It would focus squarely on foreign information manipulation and interference, not domestic political debate. It would be overseen by a multi-stakeholder board including NGOs, investigative journalists, academics, and members of the European Parliament. Participation would be transparent and documented, not anonymous.

Funding and feasibility: the EU already funds EDMO, EMIF, Horizon Europe security research, and national media literacy programmes. With €28.3 billion from the Recovery and Resilience Facility earmarked for the digital transition, including digital education components, baseline resources exist. The European Cyber Guard does not require wholly new budget lines — it requires the political will to pool and scale existing streams under a unified EU-level framework. The dual benefit – security plus digital upskilling – makes this one of the most cost-effective resilience investments the EU could make.

4.6 Epistemic infrastructure: independent verification and algorithmic transparency

Addressing the systemic erosion of shared factual ground, this section treats shared reality as a critical public good. It proposes establishing publicly funded, editorially independent verification bodies and mandating algorithmic transparency to hold platforms accountable for recommendation systems that amplify divisive content.

Beyond citizen participation, Europe faces a deeper structural challenge: the erosion of the shared epistemic infrastructure that democracy requires.

The problem is structural, not behavioural. It is not primarily that citizens believe false things. It is that the information environment has been fragmented — by algorithmic recommendation systems that prioritise engagement over accuracy, by the collapse of shared reference points once provided by public broadcasters and trusted journalism, and by the deliberate exploitation of these structural conditions by adversaries who understand that a society without shared factual ground cannot make collective decisions.

Restoring shared reality requires three things:

1. **Publicly funded, editorially independent verification bodies.** Not government fact-checkers — these create obvious conflicts of interest and would be weaponised in polarised political environments. Rather, bodies analogous to what resilient public broadcasters represented at their best: funded by public mandate, legally insulated from governmental direction, and charged specifically with maintaining a shared factual commons on matters of public concern.
2. **Mandatory algorithmic transparency.** The recommendation systems of major platforms are now the primary editors of public information flow in most EU member states. They operate invisibly. Their optimisation for engagement has been demonstrated, systematically, to amplify divisive, emotionally activating, and factually dubious content. Platforms should be required to publish, in machine-readable and human-accessible form, the criteria and weighting given to different signals in recommendation systems, with regular independent audit. The DSA creates important disclosure obligations, but more is needed.
3. **Common standards for the information supply chain.** The knowledge supply chain — from how information is sourced, through how AI agents exchange it, to how citizens receive it — currently has no common infrastructure and no incentive structure to reward epistemic rigour over engagement-maximising fluency. The EU should use its regulatory and market-shaping power to establish common standards for knowledge sourcing, attribution norms for AI-generated

content, and interoperability requirements that allow citizens to carry their epistemic context between AI systems.

4.7 Policy options and implementation pathway

This section outlines the practical steps for enacting the framework. In the short term, the HTRP could be launched via a Commission Communication and Council Conclusions, paired with a European Parliament own-initiative report. In the medium term, it could be consolidated into a formal interinstitutional agreement or a Council-endorsed operational protocol linking existing toolboxes and directives.

Policy Options for the European Union

Policy Option	Description	Intended Benefits	Key Limitations / Risks
Option 1: Establish HTRP Framework	A predefined procedural protocol for activating existing EU instruments (thresholds, triggers, and oversight).	Closes governance gaps; creates operational clarity; embeds accountability.	Requires high political will; risk of weak thresholds due to cautious negotiation.
Option 2: European Cyber Guard	A volunteer civic corps for information defence, modelled on Estonia's Cyber Defence League.	Distributes resilience; upskills citizens; provides cost-effective early warning.	Needs strict governance to prevent misuse; requires clear legal and operational boundaries.
Option 3: Epistemic Infrastructure	Publicly funded, independent verification bodies and mandatory algorithmic transparency.	Addresses root causes of factual erosion; creates platform accountability.	Risk of political capture; likely resistance from major tech platforms.
Option 4: Agility Mechanisms	Sunset clauses, learning loops, and red-teaming built into response protocols.	Prevents ossification; ensures tools evolve with AI-driven threats.	Increases governance complexity; requires constant political investment.
Option 5: Intelligence Integration	Mandatory structured reporting on hybrid indicators to the EU Hybrid Fusion Cell.	Improves situational awareness; eliminates silos between member states.	Sovereignty sensitivities; varying intelligence-sharing cultures.
Option 6: Resilience Sandboxes	Testing environments for new detection and attribution techniques (e.g., AI analysis).	Evidence-based capability development; reduces risk of faulty tool deployment.	Requires specialised expertise; risk of "leaking" defence tactics to adversaries.

Overall Guidance Emerging from the Options

- **Do not treat institutional coordination as sufficient.** Citizens, civil society, and the private sector must be active participants in hybrid resilience, not passive recipients of its outputs.
- **Build for what is not yet known.** Static architectures will be obsolete before they are implemented. Agility, learning, and modularity must be designed in from the start.

- **Treat epistemic infrastructure as public infrastructure.** Shared reality is not a natural condition — it requires deliberate investment, just as physical infrastructure does.
- **Embed democratic accountability from the outset.** Emergency powers and accelerated response mechanisms must be subject to systematic parliamentary scrutiny, not applied first and reviewed never.
- **Reject false economies in information resilience.** The resource asymmetry between EU information resilience spending and adversary influence operation budgets is not sustainable. Significant political investment is required to close it.

5. Conclusions and Policy Direction

Europe's challenge is not a lack of instruments. It is the absence of a clear democratic procedure to use them coherently, the absence of a sufficiently resilient citizenry to sustain democratic governance under sustained informational pressure, and the absence of the epistemic infrastructure that shared democratic reality requires.

The **Hybrid Threat Response Protocol** addresses the first of these. It transforms the EU's existing toolbox from a collection of policies into an operational architecture — predefined, reviewable, democratically accountable, and agile enough to evolve with the threat.

The **European Cyber Guard** addresses the second. It transforms citizens from passive targets of hybrid attack into active participants in democratic resilience — extending the principle behind Estonia's Cyber Defence League to the EU scale, and building the distributed civic capacity that no institution can manufacture on its own.

The **epistemic infrastructure investments** address the third. They treat shared reality as a public good requiring deliberate institutional architecture — publicly funded independent verification bodies, mandatory algorithmic transparency, and common standards for the AI-mediated information supply chain.

Together, these three elements constitute a response architecture commensurate with the scale of the challenge. Each is actionable. None requires new Treaty competences. All would represent a substantial advance on Europe's current position.

The adversary is not waiting for Europe to get organised. The protocol is overdue.

Annex – Key EU Instruments and Bodies

Instrument / Body	Function	Activation / Use
EU Hybrid Toolbox	Overarching framework for hybrid threat response.	Political decision by Council/Commission.
FIMI Toolbox	Specific tools for foreign information manipulation and interference.	EEAS/Commission activation.
Cyber Diplomacy Toolbox	Cyber sanctions framework (Art. 215 TFEU basis).	Council decision; notably used in July 2020.
EU Hybrid Fusion Cell	All-source intelligence analysis on hybrid threats.	Permanent; reports to the EEAS.
Hybrid CoE (Helsinki)	Research, exercises, and sharing of best practices.	Standing body; NATO/EU joint initiative.
IPCR	Integrated Political Crisis Response coordination mechanism.	Triggered by the EU Council Presidency.
ENISA	EU agency for cybersecurity; certification and coordination.	Permanent; enhanced mandate via the Cybersecurity Act.
EDMO	European Digital Media Observatory (fact-checking/literacy).	Operational; part-funded by the EMIF.
DSA	Digital Services Act; platform risk mitigation obligations.	Commission enforcement; national coordinators.
NIS2 Directive	High common cybersecurity standards across critical sectors.	Transposed and enforced by member states.
CER Directive	Critical Entities Resilience requirements for physical infrastructure.	Member state implementation.
ProtectEU (2025)	New comprehensive Internal Security Strategy.	Commission workplan through 2029.
Hybrid Rapid Response Teams	On-site operational support for member states under hybrid pressure.	Council framework; first deployed to Moldova in 2025.
Drone Action Plan (2026)	Counter-drone security framework for civilian/hybrid threats.	Commission implementation; member state cooperation.

Literature

References formatted in Chicago author-date style. Thematic groupings indicate primary relevance to paper sections; sources may support multiple sections.

I. Hybrid Warfare Theory and Russia–Ukraine Evidence Base (Sections 2.1–2.3)

Connell, Michael, and Sarah Vogler. 2016. *Russia's Approach to Cyber Warfare*. CNA Analysis & Solutions. https://www.cna.org/archive/CNA_Files/pdf/cpp-2017-u-015223.pdf

Galeotti, Mark. 2016. "Hybrid, Ambiguous, and Non-Linear? How New Is Russia's 'New Way of War'?" *Small Wars & Insurgencies* 27 (2): 282–301. https://doi.org/10.1080/09592318.2015.1129170

JRC / European Commission. 2021. *The Landscape of Hybrid Threats: A Conceptual Framework*. Joint Research Centre report. https://euhybnet.eu/wp-content/uploads/2021/06/Conceptual-Framework-Hybrid-Threats-HCoE-JRC.pdf

Lanoszka, Alexander. 2016. "Russian Hybrid Warfare and Extended Deterrence in Eastern Europe." *International Affairs* 92 (1): 175–195. https://doi.org/10.1111/1468-2346.12509

Lonardo, Luigi. 2021. "EU Law Against Hybrid Threats: A First Assessment." *European Papers* 6 (2): 1075–1096. https://www.europeanpapers.eu/e-journal/eu-law-against-hybrid-threats-first-assessment

Monaghan, Andrew. 2015. "The 'War' in Russia's 'Hybrid Warfare'." *Parameters* 45 (4): 65–74.

Wither, James K. 2016. "Making Sense of Hybrid Warfare." *Connections: The Quarterly Journal* 15 (2): 73–87. https://doi.org/10.11610/Connections.15.2.06

II. EU Institutional Architecture and the Operational Gap (Sections 3.1–3.3)

Bleyer-Simon, Konrad. 2025. "Addressing Foreign Information Manipulation in the Context of European Regulations." Policy Paper. European University Institute, Centre for Media Pluralism and Media Freedom. https://hdl.handle.net/1814/92874

D'Andrea, Alessia, Giorgia Fusacchia, and Arianna D'Ulizia. 2025. "Policy Review: Countering Disinformation in the Digital Age — Policies and Initiatives to Safeguard Democracy in Europe." *Information Polity* 30 (1): 82–91. https://doi.org/10.1177/15701255251318900

EEAS. 2026. *4th EEAS Annual Report on Foreign Information Manipulation and Interference Threats*. Brussels: European External Action Service. https://www.eeas.europa.eu/eeas/4th-eeas-annual-report-foreign-information-manipulation-and-interference-threats_en

European Commission and High Representative. 2021. *Joint Communication on Countering Hybrid Threats*. Brussels: European Commission.

European External Action Service. 2022. *Countering Hybrid Threats*. Brussels: EEAS. https://www.eeas.europa.eu/sites/default/files/documents/2022-03-28-countering-Hybrid-Threats_NewLayout.pdf

Mälksoo, Maria. 2018. "Countering Hybrid Warfare as Ontological Security Management: The Emerging Practices of the EU and NATO." In *Ontological Insecurity in the European Union*, edited volume.

Stoian, V. 2019. "Policy Integration Across Multiple Dimensions: The European Response to Hybrid Warfare." https://www.ssoar.info/ssoar/bitstream/handle/document/68415/ssoar-sp-rpsr-2019-3-4-stoian-Policy_Integration_Across_Multiple_Dimensions.pdf

III. Epistemic Infrastructure and Democratic Resilience (Sections 2.4, 4.5–4.6)

Farrell, Henry, and Bruce Schneier. 2018. "Common-Knowledge Attacks on Democracy." Berkman Klein Center Research Publication 2018-7. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3273111

Humprecht, Edda, Frank Esser, and Peter van Aelst. 2020. "Resilience to Online Disinformation: A Framework for Cross-National Comparative Research." *International Journal of Press/Politics* 25 (4): 493–516. https://doi.org/10.1177/1940161219900126

Lewandowsky, Stephan, John Cook, Sander van der Linden, Jon Roozenbeek, and Naomi Oreskes. 2023. "Misinformation and the Epistemic Integrity of Democracy." *Current Opinion in Psychology* 54: 101711. https://doi.org/10.1016/j.copsyc.2023.101711

Seger, Eduard, et al. 2020. "Tackling Threats to Informed Decision-Making in Democratic Societies: Promoting Epistemic Security in a Technologically-Advanced World." https://www.jstor.org/stable/45420122

Tenove, Chris. 2020. "Protecting Democracy from Disinformation: Normative Threats and Policy Responses." *International Journal of Press/Politics* 25 (2): 1–22. https://doi.org/10.1177/1940161220918740

IV. AI, Autonomous Threats, and Critical Infrastructure (Sections 2.2, 4.2)

Dubber, Markus D., and Seth Lazar. 2025. "Military AI Cyber Agents (MAICAs) Constitute a Global Threat to Critical Infrastructure." arXiv preprint. https://arxiv.org/abs/2502.06441

Hybrid CoE. 2025. *Artificial Intelligence and Foreign Information Manipulation: Chinese and Russian Approaches*. Helsinki: European Centre of Excellence for Countering Hybrid Threats. https://www.hybridcoe.fi/all-publications/

Hybrid CoE. 2025. *Countering Disinformation in the Euro-Atlantic: Strengths and Gaps*. Helsinki: European Centre of Excellence for Countering Hybrid Threats. https://www.hybridcoe.fi/all-publications/

Piekarski, M., M. Wolbach, and M. Okuniewska. 2025. "Employment of Uncrewed Systems in Attacks on Critical Infrastructure: A Hybrid Threat Perspective." In *Security Challenges Related to Recent Developments in Technology*. *Open Research Europe* 4: 129. https://doi.org/10.12688/openreseurope.17797.2