European Journal of Risk Regulation (2026), 1–16
doi:10.1017/err.2026.10100

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Fake Activity Markets as Hybrid Threat Vector: Implications for EU Social and Cyber Resilience

Benjamin Farrand

Law School, Newcastle Law School, UK
Email: ben.farrand@ncl.ac.uk

Abstract

Hybrid threats represent a continued challenge to the European Union, combining disinformation campaigns with cyber-attacks as a means of destabilising the Union and its Member States, undermining legitimacy and public trust. With social media platforms at the centre of disinformation efforts, as well as potential sites for cyber-attack disruption, the ability to control narratives and disseminate content are essential to hybrid threat actors. Fake Activity Markets (FAMs) are websites offering services that can be used to generate fake engagement, in turn allowing for coordinated inauthentic behaviour online. These websites also constitute a cybersecurity threat in themselves, involved in distributing malware, harvesting user data or comprising information systems. This article seeks to explore the EU approach to hybrid threats and coordinated inauthentic behaviour using FAMs as a case study, highlighting the potential threats to the EU's social and cyber resilience posed by these actors, the potential regulatory responses, and the ways in which the von der Leyen II Commission's renewed emphasis on hybrid threats could provide for a more robust ecosystem for countering coordinated inauthentic behaviour.

Keywords: coordinated inauthentic behavior; cybercrime; cybersecurity; disinformation; digital sovereignty; hybrid threats

I. Introduction

The EU considers hybrid threats to constitute one of the most significant challenges it faces; the combination of coordinated disinformation campaigns and cyber-attacks on critical and important information infrastructures seek to damage its credibility, shake trust in its institutions, and sew discord in its population, creating a digitally enabled threat to its legitimacy. Social media platforms serve as a key method of facilitating these threats, serving both as a means of disseminating disinformation and a target for disruptive cyber-behaviours. Central to both types of threat is the notion of coordinated inauthentic behaviour (CIB).

CIB has been defined as the use of "multiple fake accounts, i.e., accounts that hide the real identities of the people running them, [. . .] used to mislead and influence the users of social media."[1] These inauthentic accounts generate false engagements, such as likes, retweets, or in more sophisticated approaches, comments and responses, as a means of creating false impressions regarding the validity of sources, or the information they provide. One vector for the acquisition of CIB networks as well as a vector for the spread of

[1] "Ready-to-(Ab)Use: From Fake Account Trafficking to Coordinated Inauthentic Behavior on Twitter" (2022) 31 Online Social Networks and Media 100224, 1 <https://doi.org/10.1016/j.osnem.2022.100224>.

malicious and harmful computer code are Fake Activity Markets (FAMs). FAMs are websites accessible on the open web that offer services such as increasing follower counts and boosting engagement. These websites can be used to facilitate the spread of disinformation, but they can also act as malicious entities seeking to gain access to data from the users of its service, such as installing keyloggers or other insidious exploits into information systems.

The purpose of this paper is to draw from the work undertaken in the FAMOUS project,[2] which seeks to understand the dynamics of FAMs. Section II with an exploration of the concept of hybrid threats, and the EU's approach to responding to them, and in particular, disinformation. In Section III of the article, the nature of FAMs will be outlined, as well as the ways in which they are used. Section IV of the paper expands upon how these FAMs and the CIBs they produce constitute a threat to the EU's social resilience, and the regulatory responses that may be appropriate. Section V of the article then considers the cybersecurity issues arising from FAM-facilitated CIB, the impact on the EU's cyber-resilience, and the regulatory responses that could be used to tackle these hybrid threat actors. Finally, Section VI of the article discusses the evolving approach to hybrid threats in the EU under the von der Leyen II Commission, including its proposals for revised regulatory frameworks in the fields of disinformation and cybersecurity, and the implications they may have for the mitigating of hybrid threats posed by FAMs.

II. The von der Leyen I commission's approach to hybrid threats and disinformation

Central to EU security concerns since 2016 is the concept of "hybrid threat,"[3] explained by the Commission and High Representative of the Union for Foreign Affairs and Security Policy as:

> [T]he mixture of coercive and subversive activity, conventional and unconventional methods (i.e. diplomatic, military, economic, technological), which can be used in a coordinated manner by state or non-state actors to achieve specific objectives while remaining below the threshold of formally declared warfare.[4]

Giannopoulos, Smith and Theocharidou highlight that what is distinct about a hybrid threat is the way that it targets systemic vulnerabilities in democratic systems,[5] and while this is by no means a new phenomenon, it has nevertheless gained particular salience for the EU vis-à-vis its relations with states such as Russia[6] and to a lesser extent, China.[7] The undermining of public trust and facilitation of political polarisation are some means by which state and non-state actors employ hybrid threats as a means of achieving

[2] European Media and Information Fund, "FAMOUS: Fake Activity Market Observation System of Unethical Services" (*European Media and Information Fund*, 10 April 2026) <https://gulbenkian.pt/emifund/projects/famous-fake-activity-market-observation-system-of-unethical-services/> (last accessed 23 March 2026).

[3] H Carrapico and B Farrand, "Discursive Continuity and Change in the Time of Covid-19: The Case of EU Cybersecurity Policy" (2020) 42 Journal of European Integration 1111 <https://doi.org/10.1080/07036337.2020.1853122>.

[4] European Commission and High Representative of the Union for Foreign Affairs and Security Policy, "Joint Framework on Countering Hybrid Threats" (2016) JOIN(2016) 18 2.

[5] "The Landscape of Hybrid Threats: A Conceptual Model Public Version" (European Commission 2021) EUR 30585 EN 9.

[6] M Weissmann, "Hybrid Warfare and Hybrid Threats Today and Tomorrow: Towards an Analytical Framework" (2019) 5 Journal on Baltic Security 17.

[7] E Bajarūnas, "Addressing Hybrid Threats: Priorities for the EU in 2020 and Beyond" (2020) 19 European View 62 <https://doi.org/10.1177/1781685820912041>.

their strategic objectives.[8] For Wigell, the motivator for hybrid interference is the belief that "liberal democratic values are vulnerabilities that can be exploited to drive wedges through democratic societies and undermine governability."[9] One key means by which this is achieved is through the proliferation and dissemination of disinformation.

Disinformation is defined by the EU in the Joint Communication on Disinformation as "verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public, and may cause public harm."[10] Referring to the use of disinformation as part of hybrid warfare, the Communication highlights that disinformation campaigns were often complemented by cyber-attacks and attacks on information networks, with the former constituting attacks on data, with the latter constituting attacks on infrastructure.[11] Disinformation is a particularly subversive and difficult form of hybrid threat to address, as the EU's approach to freedom of expression lies at the heart of its value system, and yet it is through those expressions that the EU finds its legitimacy as an actor being targeted.[12] Active information control by way of direct intervention and removal of disinformation, or the implementation of binding legislation with sanctions for the spreading of disinformation, are infeasible for this reason, highlighting the effectiveness of hybrid interference. The passing by States of legally binding obligations concerning the removal of disinformation content and the imposition of penalties upon those disseminating it has been largely the purview of states considered as more authoritarian; Turkey's introduction of Article 217/A of the Turkish Penal Code, which imposes prison sentences of up to three years for knowingly spreading false information "regarding the internal and external security, public order and general health of the country" is one such example, argued to constitute a form of censorship of political opponents in a context of digital authoritarianism.[13]

Regulatory approaches to disinformation in the EU therefore prioritise risk mitigation within cybersecurity and democracy-preservation frames, rather than frames of criminal conduct. The 2018 Communication on Disinformation explicitly refers to "a European approach," in which disinformation is explicitly linked to broader cybersecurity threats[14] and responded to through means of improving transparency, promoting information diversity, fostering credibility of information and fashioning inclusive solutions through emphasis on awareness raising, media literacy and stakeholder cooperation.[15] As the Communication made clear, and as of direct relevance to this article, online platforms were identified as the key vector by which this disinformation was spread;[16] any response to this hybrid threat therefore required the involvement of platforms, which should "act with appropriate responsibility in views of their central role so as to ensure a safe online environment."[17] This resulted in the Code of Practice on Disinformation,[18] a voluntary agreement between several platforms as the European Commission, which required platforms to tackle disinformation through having clear transparency regarding ad placement and disclosure concerning the

[8] Giannopoulos, Smith and Theocharidou (n 5) 41.

[9] "Hybrid Interference as a Wedge Strategy: A Theory of External Interference in Liberal Democracy" (2019) 95 International Affairs 255, 256 <https://doi.org/10.1093/ia/iiz018>.

[10] High Representative of the European Union for Foreign Affairs and Security Policy & the European Commission, "Action Plan against Disinformation" (2018) JOIN(2018) 36 final 1.

[11] Ibid, 3.

[12] D Vese, "Governing Fake News: The Regulation of Social Media and the Right to Freedom of Expression in the Era of Emergency" (2022) 13 European Journal of Risk Regulation 477.

[13] A Yücel, "Hybrid Digital Authoritarianism in Turkey: The 'Censorship Law' and AI-Generated Disinformation Strategy" (2025) 26 Turkish Studies 1 <https://doi.org/10.1080/14683849.2024.2392816>.

[14] European Commission, "Tackling Online Disinformation: A European Approach" (2018) COM(2018) 236 2.

[15] Ibid, 6.

[16] Ibid, 2.

[17] Ibid, 7.

[18] European Commission, "EU Code of Practice on Online Disinformation" (2018).

nature and funding of political advertisements; procedures to maintain system integrity from inauthentic content dissemination through bots; and empowering consumers and the research community through (for example) prioritising authoritative content and allowing researchers access to datasets and other information for the purpose of fact-checking and identifying disinformation trends.[19] It must be stated at this point that this Code was largely adopted on the assumption by the Commission that platforms were predominantly economically motivated, and that oversight (or lack of oversight) regarding content was predicated on commercial considerations. Increasingly, the Commission took the position that platforms could not be seen as value-aligned (or at least, value-neutral) providers of services but were run by actors that had their own political values that may not necessarily be shared by the European institutions.[20]

The comparatively light regulatory touch of the EU in this field evolved into a more active co-regulatory regime under the von der Leyen I Commission's "Shaping Europe's Digital Future" agenda,[21] in which disinformation was discussed in terms of online security and the need to protect citizens from targeted and coordinated disinformation campaigns in order to forcefully defend democracy in the EU and ensure Europe's digital sovereignty, necessitating legislative intervention.[22] Along with a proposed Democracy Action Plan,[23] the Commission proposed the Digital Services Act as a means of providing more effective oversight to online platforms' efforts to address illegal and otherwise harmful content.[24] The Digital Services Act (DSA)[25] frames disinformation as a form of societal risk under recital 2,[26] requiring all operators of platforms to provide for advertising transparency under Article 26 and recommender system transparency under Article 27, with operators of Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs) regardless of where in the world they are established so long as they offer services to recipients in the EU[27] tasked with additional risk assessment and mitigation obligations under Articles 34 and 35. This is due to the effects that they may have in the EU due to the size of their user-base; VLOPs and VLOSEs are designated as such due to having "average monthly active recipients of the service in the Union equal to or higher than 45 million" under Article 33(1), based on their role and reach.[28] Husovec argues that the operation of these obligations does not result in active Commission regulation of legal content insofar as any "attempt by the Commission to prescribe content-specific measures for legal content, such as forcing companies to ban some specific lawful content in terms and conditions"[29] would go beyond its competence and

[19] Ibid, 4–8.

[20] H Carrapico and B Farrand, "When Trust Fades, Facebook Is No Longer a Friend: Shifting Privatisation Dynamics in the Context of Cybersecurity as a Result of Disinformation, Populism and Political Uncertainty" (2021) 59 JCMS: Journal of Common Market Studies 1160.

[21] European Commission, "Shaping Europe's Digital Future" (2020).

[22] Ibid 6. Of interest, but not directly relevant to this article is the change in discourse for the Commission, moving away from the language of "consumers" in describing European internet users, and instead talking in terms of "citizens."

[23] European Commission, "Communication on the European Democracy Action Plan" (2020) COM(2020) 790.

[24] European Commission, "Proposal for a Regulation on a Single Market for Digital Services (Digital Services Act) and Amending Directive 2000/31/EC" (2020) COM(2020) 825.

[25] Regulation 2022/2065 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act).

[26] See also R Ó Fathaigh, D Buijs and J van Hoboken, "The Regulation of Disinformation Under the Digital Services Act" (2025) 13 Media and Communication 1.

[27] Digital Services Act, Article 2(1).

[28] As highlighted in recital 48.

[29] Martin Husovec, "The Digital Services Act's Red Line: What the Commission Can and Cannot Do about Disinformation" (2024) 16 Journal of Media Law 47, 55 <https://doi.org/10.1080/17577632.2024.2362483>.

"cross a red line."[30] The Commission's ability to regulate the conduct of social media platforms through the obligations of risk assessment and mitigation nevertheless constitute a regulation of disinformation. Article 34 risk assessments do not only focus on illegal content but, under Article 34(1), also cover "any actual or foreseeable negative effects on civic discourse and electoral processes, and public security" under Article 34(1)(c), and "any actual or foreseeable negative effects in relation to gender-based violence, the protection of public health and minors and serious negative consequences to the person's physical and mental well-being" under Article 34(1)(d). As indicated by the Commission's explicit linkage of disinformation to hybrid threats to its security, the integrity of its democratic and electoral processes, and its use in efforts to destabilise the Union, disinformation clearly represents a vector for the type of threats identified in Article 34.

Furthermore, the crisis response mechanism under Article 36, which may be enacted in situations in which "extraordinary circumstances lead to a serious threat to public security or public health in the Union or in significant parts of it."[31] Instances of content can be both identifiable as disinformation *and* illegal forms of content such as hate speech[32] and the glorification of terrorist acts,[33] such as "memes" coded with support for acts such as the 2019 Christchurch terrorist shootings.[34] Furthermore, the Political Advertising Regulation (PAR)[35] explicitly seeks to mitigate the threat of disinformation in political contexts,[36] with the reiteration of the hybrid nature of this form of threat made clear by the reference in recital 100 to the role of the European Cooperation Network on Elections in ensuring "protection against cybersecurity incidents and fighting disinformation campaigns." This was combined with a strengthened Code of Practice on Disinformation published in 2022,[37] which includes mitigation efforts such as scrutiny of ad placements, demonetisation of disinformation content, and efforts to maintain integrity of services. Of particular relevance here is Commitment 14, which requests that platforms limit "impermissible behaviour" such as the use of fake accounts, bot-driven amplification, the purchase of fake engagements, the creation and use of accounts that participate in CIB, and user conduct aimed at artificially amplifying the reach of or perceived support for disinformation'.[38] In 2025, it was agreed that this Code would be incorporated as a Code of Conduct[39] in the DSA,[40] which did not change its obligations, but did render compliance auditable under Article 37. In the next section of this article, the approach to CIB will be considered further, acting as a basis for consideration of the threats to social resilience in Section IV, and cyber resilience in Section V.

[30] Ibid.

[31] Digital Services Act Article 36(2).

[32] Council Framework Decision 2008/913/JHA on combating certain forms and expressions of racism and xenophobia by means of criminal law.

[33] Regulation 2021/784 on addressing the dissemination of terrorist content online.

[34] Benjamin Farrand, "'Is This a Hate Speech?' The Difficulty in Combating Radicalisation in Coded Communications on Social Media Platforms" (2023) 29 European Journal on Criminal Policy and Research 477; see also Jim Edwards, "The Chilling Number of Misunderstood 'jokes' in the Christchurch Killer's Manifesto Show How Few People Understand the Disinformation Ecosystem of the Alt Right" (*Insider*, 30 December 2019) <https://www.insider.com/memes-in-christchurch-brandon-tarrant-manifesto-2019-12> (last accessed 3 February 2026).

[35] Regulation 2024/900 on the transparency and targeting of political advertising.

[36] Ibid at recital 4.

[37] European Commission, "2022 Strengthened Code of Practice on Disinformation" (2022).

[38] Ibid 15–16.

[39] European Commission, "Code of Conduct on Disinformation With Overview" (2025).

[40] DSA, Article 45.

III. Fake activity markets: (web)sites of insecurity

FAMs represent a form of "black market" of influence that provide users with the opportunity to buy likes, shares, comments and accounts.[41] According to NATO, these markets operate at three tiers of sophistication; Tier 3 services are the cheapest and least sophisticated, offering automatically registered but unverified accounts, Tier 2 services providing automatically registered but verified accounts that take some time to identify and block, and Tier 1 services that provide existing verified accounts with a combination of bot and "real" accounts that may even have subscribed to their content.[42] While due to their sophistication and the complexity of their identification Tier 1 services are both expensive and predominantly accessed through personal referrals made through the dark web,[43] Tier 3 services are readily accessible on the consumer-facing web through search services such as Google search, DuckDuckGo, and Bing. According to research conducted as part of the FAMOUS project, computer science techniques were applied to identify publicly accessible shops that could be found through simple searches on these search engines, with the identification of 2,003 distinct websites available in the EU area.[44]

Services identified provided services broken down by platform, such as upvotes on Reddit, follower, likes and views on Instagram, and likes and retweets on X. While these services can be used by budding influencers to expand reach,[45] or for individuals or small businesses to promote their products or services,[46] they can also be used to spread disinformation. This could be concerning the function of European institutions (including false reports concerning ENISA, the EU's Cybersecurity Agency, and the power outage in Spain and Portugal in April 2025),[47] the causes of and actions in international conflicts (such as the spread of deepfaked images and videos of Ukrainian President Zelenskyy ordering the surrender of Ukrainian troops to Russian forces),[48] or the spread of harmful diseases (like claims that Russian dissident Alexei Navalny had died as a result of a blood clot caused by the Pfizer-produced Covid vaccine, as reported by *EU v Disinfo*).[49] This disinformation may originate on niche platforms and websites,[50] but is then actively

[41] NATO StratCOM COE, "The Black Market for Social Media Manipulation" (NATO 2018) 5.

[42] Ibid, 9–10.

[43] Ibid, 10; see also P Wang and others, "Into the Deep Web: Understanding E-Commerce Fraud from Autonomous Chat with Cybercriminals" *Network and Distributed Systems Security Symposium 2020* (2020).

[44] See M Kalameyets and others, "Accountability Gaps in the Fake Follower Market", (2026) Proceedings of the 2026 ACM FAccT Conference, Montréal, Canada 1

[45] C Schulz, "On "Super Likes" and Algorithmic (In)Visibilities: Frictions Between Social and Economic Logics in the Context of Social Media Platforms" (2023) 9 Digital Culture & Society 45 <https://doi.org/10.14361/dcs-2023-0204>.

[46] M Héder, "A Black Market for Upvotes and Likes" (*arXiv.org*, 19 March 2018) <https://arxiv.org/abs/1803.07029v1> (last accessed 6 February 2026); Ganaele Langlois and Greg Elmer, "Impersonal Subjectivation from Platforms to Infrastructures" (2019) 41 Media, Culture & Society 236 <https://doi.org/10.1177/0163443718818374>; these are in themselves forms of inauthentic behaviour, but not the direct focus of this article.

[47] ENISA, "Disinformation Alert" (*ENISA*, 30 April 2025) <https://www.enisa.europa.eu/news/enisa-disinformation-alert> (last accessed 6 February 2026).

[48] Europol, "Facing Reality? Law Enforcement and the Challenge of Deepfakes" (Europol 2022) An Observatory Report from the Europol Innovation Lab QL-02-24-129-EN-N 5–6.

[49] EUvDisinfo, "Alexei Navalny is One of the Many Victims of the Pfizer-BioNTech COVID-19 Vaccine" (*EUvsDisinfo*, 19 February 2024) <https://euvsdisinfo.eu/report/alexei-navalny-is-one-of-the-many-victims-of-the-pfizer-biontech-covid-19-vaccine/> (last accessed 6 February 2026).

[50] See, for example, PM Krafft and J Donovan, "Disinformation by Design: The Use of Evidence Collages and Platform Filtering in a Media Manipulation Campaign" (2020) 37 Political Communication 194 <https://doi.org/10.1080/10584609.2019.1686094>.

spread through "mainstream" social media platforms, actively boosted by a combination of real users and coordinated bots.[51]

Recent experiments by NATO demonstrate that these markets continue to thrive, despite the introduction of the DSA, with an increase in the sophistication of these CIBs.[52] According to the NATO Strategic Communication Centre of Excellence, "systemic vulnerabilities persist [. . . and] commercial manipulation remains inexpensive and easy to obtain."[53] Indeed, the NATO findings demonstrated that while much of the purchased coordinated networks' efforts were directed towards cryptocurrency scams, commercial offerings and other non-political material, they continued "to observe an annual increase in the use of spam bots to promoting political narratives and nation-related issues."[54] Furthermore, a 2025 Europol report has indicated that there is increased coordination between criminal actors offering cybercrime as a service and hybrid threat actors that are increasingly state-aligned and ideologically motivated; "these networks are instrumental for propaganda campaigns aimed at spreading disinformation and influencing political systems. These campaigns often involve fake social media accounts, coordinated troll operations, and manipulated news content."[55] The same networks also engage in various forms of cyber-attacks, such as encrypting systems through ransomware and demanding cryptocurrency payments, and unlawfully accessing sensitive and proprietary data either for personal profit, or to provide to state entities.[56] This is not unsurprising given the "geopolitical turn" in contemporary international relations, with the undermining of trust in the liberal international order, and the return to "big power politics" that has resulted.[57] For the Commission, this resulted in a reorientation of relations with the rest of the world that was framed as the birth of the Geopolitical Union,[58] which in the context of digital technologies required the pursuit of technological sovereignty based in the recognition of the dependencies that external threat actors could weaponise in order to destabilise Europe.[59] Hybrid threats, such as the combined use of disinformation and cyber-attacks, constitute one such form of technology weaponisation. In the next section of this article, the application of existing legal frameworks concerning content regulation and cybersecurity respectively will be assessed in terms of their effectiveness in tackling CIB initialised through FAMs, identifying deficiencies that could be addressed as part of the von der Leyen II Commission's initiatives in this field.

[51] M Mazza, G Cola and M Tesconi, "Ready-to-(Ab)Use: From Fake Account Trafficking to Coordinated Inauthentic Behavior on Twitter" (2022) 31 Online Social Networks and Media 100224 <https://doi.org/10.1016/j.osnem.2022.100224>.

[52] NATO StratCOM COE, "Social Media for Sale: 2025 Experiment on Platform Capabilities to Detect and Counter Inauthentic Social Media Engagement" (NATO 2025) 5.

[53] Ibid.

[54] Ibid, 6.

[55] Europol, "The Changing DNA of Serious and Organised Crime: 2025 Serious and Organised Crime Threat Assessment" (Europol 2025) QL-01-25-005-EN-N 15.

[56] Ibid, 14–15.

[57] In the context of digital policy, see André Barrinha and Thomas Renard, "Power and Diplomacy in the Post-Liberal Cyberspace" (2020) 96 International Affairs 749 <https://doi.org/10.1093/ia/iiz274>; and in the international system generally, see G John Ikenberry, "The End of Liberal International Order?" (2018) 94 International Affairs 7 <https://doi.org/10.1093/ia/iix241>; E Braw, *Goodbye Globalization: The Return of a Divided World* (Yale University Press 2024).

[58] U von der Leyen, "2023 State of the Union Address by President von Der Leyen: Answering the Call of History" (2023) SPEECH/23/4426.

[59] Covered in B Farrand, *Geopolitical Union: Europe's Attempt to Take Back Control of Technology Regulation* (Cambridge University Press 2026).

IV. Fake activity markets and coordinated inauthentic behaviour: the threat to social resilience

As discussed in Section II of this article, the regulation of disinformation in the EU is pursued through a combination of the DSA, PAR and 2025 Code of Conduct. In dealing with FAMs and the facilitation of CIB, regulation is ultimately indirect, focusing on the platforms on which the coordinated behaviour is undertaken, and the search engines that render accessible this content to internet users. It is therefore necessary to consider both platform-based approaches and search engine-based approaches. With regard to platform-based approaches, some of the major platforms accessible in the EU and key targets for influence operations have been designated as VLOPS; Meta's Facebook and Instagram, Twitter International United Company's X, and TikTok Technology Ltd's TikTok all being designated in April 2023.[60] Under Article 34, risk assessments of services must include the identification, analysis and assessment of any systemic risks stemming from the design or functioning of their service and its related systems. In the context of coordinated disinformation influence purchased through FAMs, this risk will pertain to the potential threats to civic discourse, electoral processes and public security under Article 34(1)(c) and the threats relating to gender-based violence, public health, the security of minors and risks of physical and mental harm, arising from the disinformation spread. In conducting this assessment of potential risks from inauthentic behaviour, platforms should consider the extent to which their system design, including algorithmic recommendation, terms of service, and systems for selecting and presenting advertisements under Article 34(2). Upon identification of these risks, platforms are expected to engage in risk mitigation efforts under Article 35; in dealing with CIB, this could include adaption of: systems under Article (35)(1)(a); terms of service under 35(1)(b); recommender systems under 35(1)(d); and advertising systems under 35(1)(e). Where posts are in the form of political advertisements as defined under the PAR Article 3(2), then transparency obligations arise under Article 6. This includes disclosure of the nature of the political advertising service provider,[61] identification of the political advertisement as a political advertisement upon meeting the relevant criteria[62] and abiding by labelling and transparency requirements.[63]

NATO's 2025 experiment, however, found that platforms' responses to CIB facilitated through means such as FAMs are highly varied, and overall, not particularly effective. Running an experiment in which inauthentic accounts were created and inauthentic engagements such as likes, retweets or comments were performed, the NATO Strategic Communication Centre of Excellence observed platforms' responses over a four-week period. X removed more inauthentic accounts and instances of inauthentic engagement at 82% and 57%, respectively,[64] whereas Meta only removed 39% of inauthentic accounts on Facebook, and 22% from Instagram, and only 20% and 16% of inauthentic engagement.[65] TikTok performed even more poorly on inauthentic accounts, removing only 4% of inauthentic accounts, but 32% of inauthentic engagements, higher than those removed by Meta.[66] This is despite the terms of service/community guidelines provided by each of the platforms; X, for example, explicitly prohibits platform manipulation and spam, as well as

[60] European Commission, "Supervision of the Designated Very Large Online Platforms and Search Engines under DSA" (*European Commission – Shaping Europe's Digital Future*, 6 February 2026) <https://digital-strategy.ec.europa.eu/en//list-designated-vlops-and-vloses> (last accessed 7 February 2006).

[61] PAR, Art. 7.

[62] PAR, Art. 8.

[63] PAR, Art. 11.

[64] NATO StratCOM COE (n 51) 14–15.

[65] Ibid.

[66] Ibid.

misleading and deceptive identities,[67] or content deemed to be intended to disrupt civic and electoral processes.[68] Meta's Community Standards applicable to Facebook and Instagram prohibit CIB, stating that it will not allow conduct "intended to deceive Meta or our community,"[69] and says it will act to remove misinformation that pertains to risks of physical harm or violence, content relating to public health, vaccines and "miracle cures," and electoral processes.[70] TikTok's guidelines on authenticity and misinformation are the most detailed, featuring breakdowns of content that will be removed (such as that causing significant harm, such as vaccine denial or poses a risk to public safety) or will be considered "For Your Feed" ineligible, defined as causing moderate harm, such as misuse of authoritative sources out of context that could mislead an audience.[71] Furthermore, the Guidelines prohibit deceptive behaviours and fake engagement, stating that "if we detect accounts or content with inauthentic metrics, we'll remove fake likes, followers, or other inflated signals."[72] This largely corresponds to the findings by Gsenger that VLOPs tend to have detailed terms concerning disinformation and inauthentic behaviour, but that actual procedures and user reporting functions are not particularly transparent.[73]

For this reason, the failure to effectively identify and remove content is not necessarily because of a failure to have robust terms of service; instead, the issues relate to the need for more robust identification mechanisms for inauthentic content. As the NATO experiment concluded, there is a need on the part of platforms to prioritise behavioural detection, improving its automated detection to go beyond textual analysis to identify coordination patterns in timing, tone and relational dynamics, particularly as systems of automated coordinated authentic behaviour incorporate AI elements.[74] Furthermore, it noted that for the 2025 experiment, it included paid-for advertisements, and confirmed that "commercial manipulation is not limited to organic content and can extend into paid advertising, with the potential for platform ad systems to contribute to the distribution of inauthentic material."[75] The NATO Strategic Communication Centre of Excellence found that only TikTok complied completely with its requests for information concerning these inauthentic advertisements, whereas Meta only provided data for Facebook but not Instagram, whereas X did not comply at all.[76] Ultimately, the failure to do so can constitute a breach of both the obligations concerning advertisement transparency under the DSA Article Article 39, as well as the obligations concerning political advertisements under the PAR Articles 6, 7, 8 and 11, where the inauthentic content is political in nature. In all instances, conduct identified in the NATO experiment breaches the obligations under the 2025 Code of Conduct, although it must be stated that only Meta and TikTok have signed up to the Code, with Twitter International's X choosing not to do so.

VLOSEs also have obligations that mirror those of the VLOPs under DSA Articles 34 and 35. As of the time of writing, both Alphabet's Google and Microsoft's Bing search engines

[67] X Help Center, "Authenticity" (*X*, April 2025) <https://help.x.com/en/rules-and-policies/authenticity> (last accessed 8 February 2026).

[68] X Help Center, "X's Civic Integrity Policy" (*X*, November 2025) <https://help.x.com/en/rules-and-policies/election-integrity-policy> (last accessed 8 February 2026).

[69] Meta Transparency Center, "Inauthentic Behaviour" (*Meta*, 2026) <https://transparency.meta.com/en-gb/policies/community-standards/inauthentic-behavior/> (last accessed 8 February 2026).

[70] Meta Transparency Center, "Community Standards" (*Meta*, 2026) <https://transparency.meta.com/policies/community-standards/> (last accessed 8 February 2026).

[71] TikTok, "Community Guidelines" (*TikTok*, 2025) <https://www.tiktok.com/community-guidelines/en-GB/integrity-authenticity> (last accessed 8 February 2026).

[72] Ibid.

[73] R Gsenger, "Platform Governance under the Digital Services Act: A Perspective on Disinformation" [2025] Information, Communication & Society 1, 8 <https://doi.org/10.1080/1369118X.2025.2590561>.

[74] TikTok (n 70) 36.

[75] NATO StratCOM COE (n 51) 5.

[76] Ibid, 19.

have been designated as VLOSEs[77] but DuckDuckGo, which also featured in the FAMOUS project, does not meet the DSA Article 33 criteria. As such, risk assessments and mitigation measures should be adopted under Articles 34 and 35, which in this case would pertain to the listing of websites that can be used to provide access to services that could be used to engage in CIBs. The research conducted in FAMOUS found that FAMs had high discoverability, particularly on Bing, often appearing in the first page of search results when key search terms were used. While the inauthentic behaviour does not happen *through* the search engine, it instead lists where such services can be found. This created a different issue for search engines, insofar as while they do engage in efforts to combat the listing of websites that peddle dis and misinformation,[78] as the FAMs themselves are not sources of disinformation themselves, but instead could be used to propagate it, this would create more of a lacuna in which it would not be immediately clear whether the disinformation-specific concerns of the DSA and PAR are addressed. However, the obligations under the 2025 Code of Conduct are relevant, insofar as under commitment 14 to prevent the purchase of fake engagement, non-transparent paid messages or promotions by influencers and the creation and use of fake accounts that participate in CIB. With the ability to perform audits under the DSA to ensure compliance with the Code stated previously, and with Microsoft and Alphabet both being signatories, the obligations of risk assessment and mitigation are enhanced, and included within the VLOPs and VLOSEs reporting obligations.[79] Somewhat ironically, given the findings in this section, the NATO experiment on platforms and FAMOUS's findings on search engine visibility, the Conclusions by the European Board for Digital Services note that "the Board stress the importance of the commitments related to platform policies prohibiting techniques, practices and procedures employed by rogue actors to engage in inauthentic behaviour on VLOPs and VLOSEs and encourages signatories to maintain high ambition level [sic] in this area."[80] It would seem that at the time of writing this ambition has not yet been realised, with implications for social resilience to disinformation in the EU.

V. Fake activity markets as hybrid threat vector: the threat to cyber-resilience

As discussed in Section II of this article, the EU's digital technology security concerns expand beyond "just" disinformation to hybrid threats, which incorporate concerns over cyber-resilience, particularly in the context of geopolitical instability.[81] FAMs constitute a vector for these hybrid threats, which aligns with the earlier discussed findings of Europol that there was considerable interplay between actors involved in CIBs, and those engaged in behaviours detrimental to European cybersecurity for financial or ideological reasons.[82] Europol indicates that data is a central commodity in this threat landscape, whether data on corporate, public, or private actors. It can be used to carry out attacks, serve as a target

[77] European Commission, "Supervision of the Designated Very Large Online Platforms and Search Engines under DSA" (n 59).

[78] Gsenger (n 72) 10.

[79] European Board for Digital Services, "Conclusion of the Board: The Recognition of the Code of Practice on Disinformation as a Code of Conduct Pursuant to Article 45 of Regulation 2022/2065 (Digital Services Act or "DSA")" (2025) 8.

[80] Ibid, 11–12.

[81] For more on this, see, for example, B Farrand, H Carrapico and A Turobov, "The New Geopolitics of EU Cybersecurity: Security, Economy and Sovereignty" (2024) 100 International Affairs 2379 <https://doi.org/10.1093/ia/iiae231>.

[82] Europol, "The Changing DNA of Serious and Organised Crime: 2025 Serious and Organised Crime Threat Assessment" (n 54) 15.

in itself, or is collected as a by-product of attacks.[83] In the context of FAMs, this malicious activity can be carried out in two ways; the first, is through using the CIB not to spread disinformation, but malicious code, such as that used to spread malware and spam.[84] Similarly, malicious code can be the means by which CIB is carried out, by "clickjacking" or otherwise gaining unknowing control over real user's accounts, which can then be used to produce false engagement with content,[85] which could then be used to spread disinformation.[86] The second is through the FAM website directly offering the engagement services itself. Some of these websites may not provide any engagement service in reality, but instead may be a "fake website" created solely for the purposes of malicious data collection, phishing for information such as credit card details, or insert malicious code that either gives remote access to the system, gain access to passwords, or installs remote cryptocurrency miners on that system.[87] Alternatively, the service may be provided, but these malicious activities are carried out in tandem. In FAMOUS, of the sample websites identified, 34% exhibited high risk profiles, and 39% exhibited moderate risk profiles, which included JavaScript content injections (which can be used to gain unauthorised access to data, for example), forced redirects, and attempts to install keyloggers. All these types of activity can have significant cybersecurity implications, and as such, these websites operate as a potential threat to cyber-resilience in the EU.

In terms of legal frameworks, the most immediately relevant to these forms of hybrid threat is the Attacks against Information Systems Directive.[88] This Directive approximates criminal law responses to information system-targeted crimes, such as illegal access to computer systems,[89] illegal system interference,[90] illegal data interference,[91] illegal interception[92] or making available the tools to make those offences possible.[93] Where FAM websites contain exploits that install keyloggers, inject malicious JavaScript or gain access to data without authorisation, there are clear breaches of the Directive, and by extension, the laws of the Member States intended to tackle these threats. However, this can be of little consequence if the websites are not based within the EU area, given the difficulties in

[83] Ibid, 40.

[84] See G Stringhini and others, "Follow the Green: Growth and Dynamics in Twitter Follower Markets" Proceedings of the 2013 Conference on Internet Measurement Conference (Association for Computing Machinery 2013) 1 <https://doi.org/10.1145/2504730.2504731> (last accessed 9 February 2026); R Singh Kunwar and P Sharma, "Social Media: A New Vector for Cyber Attack" 2016 International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Spring) (2016) <https://doi.org/10.1109/ICACCA.2016.7578896> (last accessed 9 February 2026); M Al Qwaid, "Cybersecurity Threats: Ransomware, Phishing, and Social Engineering" in Amjad Yousef Aldweesh, *Complexities and Challenges for Securing Digital Assets and Infrastructure* (IGI Global Scientific Publishing 2025).

[85] A Toor, "Instagram Virus Creates Fake "Likes" and Followers in Lucrative Marketing Scam" (*The Verge*, 17 August 2013) <https://www.theverge.com/2013/8/17/4630532/instagram-virus-creates-fake-likes-and-followers-in-lucrative> (last accessed 9 February 2026); K Matthe Caramancion and others, "The Missing Case of Disinformation from the Cybersecurity Risk Continuum: A Comparative Assessment of Disinformation with Other Cyber Threats" (2022) 7 Data 49 <https://doi.org/10.3390/data7040049>.

[86] M Majid Akhtar and others, "SoK: False Information, Bots and Malicious Campaigns: Demystifying Elements of Social Media Manipulations" Proceedings of the 19th ACM Asia Conference on Computer and Communications Security (Association for Computing Machinery 2024) <https://doi.org/10.1145/3634737.3644998> (last accessed 9 February 2026).

[87] M Alremeithi and others, "Towards the Development of Indicators of Fake Websites for Digital Investigation" European Conference on Cyber Warfare and Security (Academic Conferences International Limited 2023).

[88] Directive 2013/40/EU on attacks against information systems and repealing Council Framework Decision 2005/222/JHA.

[89] Ibid, Art. 3.

[90] Ibid, Art. 4.

[91] Ibid, Art. 5.

[92] Ibid, Art. 6.

[93] Ibid, Art. 7.

effectively prosecuting cybercrimes across multiple jurisdictions.[94] Instead, effective mitigation of these forms of threat depends highly on collaboration and cooperation between different expert bodies and agencies,[95] and guaranteeing high levels of cybersecurity. In terms of cooperation, Europol's EC3 (the European Cybercrime Centre)[96] and J-CAT (The Joint Cybercrime Action Taskforce)[97] spearheads investigations into malicious cyber actors in cooperation with police bodies in countries such as Canada, the United States, United Kingdom (UK) and Australia. Recent activity has included Operation Endgame, which focused on neutralising significant cybersecurity threats by dismantling network infrastructure and servers, including the data theft malware Rhadamanthys, the Remote Access Trojan VenomRAT and the botnet Elysium.[98] While these are significant successes, Europol itself highlights in the 2025 IOCTA report that the bringing down of larger networks results in cybercriminal diversifying tactics and targets to compensate for the losses of these services.[99] While admittedly likely to be "smaller fish" in the cybercrime as a service ecosystem, FAMs can nevertheless contribute to the broader security threat, particularly if they become an attractive alternative due to the dismantling of other networks. For this reason, the focus on cyber*security* as distinct from cyber*crime* becomes essential in mitigating the threat to cyber-resilience.

The NIS2 Directive,[100] according to its first recital, is intended to "mitigate threats to network and information systems used to provide essential services in key sectors and ensure the continuity of such services when facing incidents." To achieve this, under Article 1 it sets up systems of Computer Security Incident Response Teams (CSIRTs), rules and obligations concerning cybersecurity information sharing, and most importantly for this article, cybersecurity risk-management measures and reporting obligations for key entities. NIS2 expands substantially on the key entities covered NIS1, which can be found in Annexes I and II. Annex II, "other sectors of criticality." This means that they are not held to the same level of proactive supervision as essential entities such as energy and health sectors, but they may nevertheless be subject to reactive supervision for failures to ensure resilience under Article 33.[101] Digital service providers are listed at Annex II.6, which are defined as providers of online marketplaces, online search engines and online social networking platforms. This clearly means that VLOPs and VLOSEs have obligations under NIS2 in addition to those obligations relating to content. These obligations are listed

[94] G Christou, "The Challenges of Cybercrime Governance in the European Union" (2018) 19 European Politics and Society 355 <https://doi.org/10.1080/23745118.2018.1430722>; J Kumar Gupta and U Lunia, "Cyber Crime and the Challenges of Prosecution and Prevention" (2024) 7 International Journal of Law Management & Humanities 1038.

[95] B Farrand, "'Alone We Can Do so Little; Together We Can Do so Much': The Essential Role of EU Agencies in Combatting the Sale of Counterfeit Goods" (2019) 28 European Security 22.

[96] Europol, "European Cybercrime Centre - EC3 – Combating Crime in a Digital Age" (*Europol*, 2026) <https://www.europol.europa.eu/about-europol/european-cybercrime-centre-ec3> (last accessed 9 February 2026).

[97] Europol, "Joint Cybercrime Action Taskforce (J-CAT) – Fighting Cybercrime Worldwide" (*Europol*, 2026) <https://www.europol.europa.eu/how-we-work/services-support/operational-coordination-support/joint-cybercrime-action-taskforce> (last accessed 9 February 2026).

[98] Europol, "End of the Game for Cybercrime Infrastructure: 1025 Servers Taken Down" (*Europol*, 13 November 2025) <https://www.europol.europa.eu/media-press/newsroom/news/end-of-game-for-cybercrime-infrastructure-1025-servers-taken-down> (last accessed 9 February 2026).

[99] Europol, "Steal, Deal and Repeat: How Cybercriminals Trade and Exploit Your Data" (Europol 2025) IOCTA: Internet Organised Crime Threat Assessment 2025 17.

[100] Directive 2022/2555 on measures for a high common level of cybersecurity across the Union amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive)

[101] See also C Gaie and M Mueck, "Introduction to the Networks and Information Systems 2 (NIS2) Directive" in Markus Mueck and Christophe Gaie (eds), *European Digital Regulations* (Springer Nature Switzerland 2025) <https://doi.org/10.1007/978-3-031-80809-8_7>.

in Chapter IV of the Directive, which include obligations concerning governance under Article 20, such as ensuring that management bodies are sufficiently trained in cyber-threats and ensure the training of their employees, and obligations to put in place cybersecurity risk-management procedures under Article 21. Article 21 requires that the entities take appropriate and proportionate technical, operational and organisational measures to manage the risks posed to the security of network and information systems, and to prevent or minimise the impact of incidents on recipients of their services and on other services. These measures are to be in line with European and international standards, including those of ETSI (European Telecommunications Standards Institute) and ISO (International Standardisation Organisation).[102]

The list of measures to be adopted under Article 21(2) of direct relevance to this form of cyber-threat include having mandatory policies on risk analysis and system security, incident handling, crisis management and effective vulnerability handling. In the event of an "incident," defined under Article 6(6) as an "event compromising the availability, authenticity, integrity or confidentiality of stored, transmitted or processed data or of the services offered by, or accessible via, network and information systems," there is an obligation to report this to CSIRTs under Article 23 where the incident is considered to have "affected or is capable of affecting other natural or legal persons by causing considerable material or non-material damage."[103] Should there be a failure to comply with Article 21 organisational measures or Article 23 incident reporting measures, Article 34(5) requires Member States to impose fines of a maximum of at least €7 million, or of a maximum of at least 1.4 % of the total worldwide annual turnover in the preceding financial year of the undertaking to which the important entity belongs, whichever is higher. This does suggest that in terms of resilience requirements, both VLOPs and VLOSEs have existing obligations concerning cybersecurity on their systems, particularly around data authenticity and integrity, both of which are implicated by potentially malicious cyber-activities arising from the propagation of FAM-based cyber-attacks. This would appear to suggest that any failure to do so is not the result of insufficient legal frameworks, but effective compliance and enforcement. It must be stated, however, that suffering a successful cyber-attack is not sufficient for liability to arise, as it is generally accepted that it is often a case of "when" not "if" a cyber-attack will be successful,[104] particularly when dealing with small-scale incidents that are likely to arise from FAM-based attacks. Instead, liability arises from a failure to abide by best practices and industry standards, not demonstrating resilience when a successful cyber-attack is carried out, or failing to report an incident where it is necessary to do.

VI. Von der Leyen II and evolving approaches to hybrid threats

From the preceding analysis, it is possible to state that FAMs do indeed constitute a hybrid threat vector for the EU, in which both disinformation and cyber-attacks can be facilitated for both ideological and profit-motivated reasons. While they may not reach the scale or seriousness of state-backed mass disinformation campaigns or coordinated organised crime-based cyber incursions, they can instead represent something of a "death by a thousand cuts," entailing a slow trickle of disinformation discourses, or individual or small-scale exploitation of cyber-vulnerabilities, both on the part of European internet

[102] O Kanevskaia, P Delimatsis and S Bijlmakers, "The European Telecommunication Standards Institute (ETSI): In Search for Legitimacy and Resilience of European Standardisation" (2024) 37 Innovation: The European Journal of Social Science Research 1 <https://doi.org/10.1080/13511610.2024.2395262>.

[103] NIS2 Directive, Article 23(3)(b)

[104] F Hering and others, "The Damocles Sword of Cyber Attacks" (2025) 67 Business & Information Systems Engineering 141, 141.

users and potentially for platforms. Both forms of activity have the potential for significant disruption within the EU and serve as challenges for a sovereignty and autonomy-focused digital policy. Toward the end of the von der Leyen I Commission, two further cybersecurity Regulations were adopted, with the intention of ensuring resilience in supply chains for digitally enabled technologies,[105] and for cross-border responses to large-scale cyberattacks.[106] While not as directly relevant to FAMs as the Attacks on Information Systems Directive and the NIS2 Directive, both these measures were explicitly adopted on the basis of concerns regarding the EU's geopolitical positioning and the increase in cyber-attacks resulting from an unstable world (and in particular the war in Ukraine),[107] concerns which have not gone away. The Cyber Resilience Act requires that all products with digital elements, which includes software under Article 3(1), comply with essential cybersecurity requirements and are "used for their intended purpose or under conditions which can reasonably be foreseen" under Article 6. According to Annex I, Part I(1), "[p]roducts with digital elements shall be designed, developed and produced in such a way that they ensure an appropriate level of cybersecurity based on the risks," which includes under subsection 2, that they "be made available on the market without known exploitable vulnerabilities." Arguably, the use of FAMs as a means of generating both CIB and cyber-attacks could be construed as not complying with this Annex, and in the most serious cases, could require withdrawal of the services from the market under Article 58.

The von der Leyen II Commission has made proposals that are relevant to both tackling disinformation and combating cyber-threats. In the Competitiveness Compass,[108] a document that serves as the main agenda setting piece for the new mandate, it is stated that "increasing hybrid threats require a closer alignment between the public and private sectors,"[109] serving as the basis for the internal security strategy, "ProtectEU."[110] In this strategy, it was stated that the lines between open warfare and hybrid threats is blurring, particularly as a result of Russia's use of combined disinformation and cyber-attacks in support of its actions in Ukraine.[111] The combating of hybrid threats, and in particular increasing resilience of critical and important entities, was identified as being essential to the EU's technological sovereignty,[112] necessitating further action. As a result, the Commission has proposed a revised Cybersecurity Act.[113] In this "Cybersecurity Act 2," proposed recital 1 highlights that

> The cybercrime ecosystem has proliferated [. . .] whether caused by criminals for financial gain or by State actors for disruption, espionage, disinformation or warfare have intensified. As part of a wider hybrid strategy, incidents resulting from malicious cyber activities and system failures ripple outward, disrupting essential

[105] Regulation 2024/2847 on horizontal cybersecurity requirements for products with digital elements and amending Regulations (EU) No 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act).

[106] Regulation 2025/38 laying down measures to strengthen solidarity and capacities in the Union to detect, prepare for and respond to cyber threats and incidents and amending Regulation (EU) 2021/694 (Cyber Solidarity Act).

[107] H Carrapico and B Farrand, "Cybersecurity Trends in the European Union: Regulatory Mercantilism and the Digitalisation of Geopolitics" (2024) 62 JCMS: Journal of Common Market Studies 147 <https://doi.org/10.1111/jcms.13654>.

[108] European Commission, "A Competitiveness Compass for the EU" (2025) COM(2025) 30.

[109] Ibid, 15.

[110] European Commission, "ProtectEU: A European Internal Security Strategy" (2025) COM(2025) 148.

[111] Ibid, 1.

[112] Ibid, 13.

[113] European Commission, "Proposal for a Regulation on the European Union Agency for Cybersecurity (ENISA), the European Cybersecurity Framework, and ICT Supply Chain Security and Repealing Regulation 2019/881 (The Cybersecurity Act 2)" (2026) COM(2026) 11.

services, undermining trust in institutions, and affecting the Union's societal and defence readiness.

This revision can have relevance for FAMs, as it indicated in the proposed Article 2(4) that ICT supply chains, defined as "a sum of ICT services, ICT products and ICT processes that encompass activities and actors involved at all stages upstream of a product being made available or a service being delivered on the market", can be subject to enhanced risk assessments under proposed Article 99(1), in which an EU wide-analysis of the risks in that supply chain is assessed, including the main threat actors and vulnerabilities, and that "the Union-level coordinated security risk assessments shall develop risk scenarios and propose measures to mitigate the identified risks." This would potentially allow for coordinated EU action spearheaded by ENISA to more thoroughly investigate the risks posed by FAMs, both with regard to their use in CIB and in malicious exploitation of cyber vulnerabilities. Similarly, a revision to the NIS2 Directive has been proposed, to align this Directive with the Cybersecurity Act 2, providing for important entities such as platforms to be subject to certification requirements by Member States,[114] as well as providing for stronger ENISA oversight.[115] This is combined with a "Democracy Shield," proposed as part of a raft of measures intended to protect democracy and civil society in the EU.[116] Reiterating the threat from hybrid attacks, the Democracy Shield intends for the establishment of a European Centre for Democratic Resilience, intended to coordinate Member State responses to foreign manipulation and disinformation.[117] Amongst other measures, the Commission has proposed strengthening support for the DSA Code of Conduct, as well as clearly establishing crisis protocols relevant for the application of Article 36.[118] The Democracy Shield Communication specifically refers to inauthentic behaviours of the types promoted through FAMs, such as "fake accounts, AI and bot-driven amplification and inauthentic engagements. Such coordinated operations, often part of broader hybrid campaigns, are increasingly sophisticated and decentralised."[119] With a renewed focus on this type of illicit activity, combined with increased cybersecurity coordination and certification requirements for important entities under the NIS2 Directive, it is possible that the EU approach to these forms of hybrid threat will become more prominent during the second von der Leyen Commission. With the combination of the existing frameworks under the DSA, PAR and Code of Practice on Disinformation, the EU has the legal frameworks available that can be used to target CIB. Similarly, the EU's cybersecurity frameworks are now considerably robust, and with the proposed reforms, the potential for coordinated action between ENISA, the Commission, and Member States in the mitigation of FAM-based cyber attacks is enhanced. Questions regarding the effectiveness of the EU regime in countering these threats are now largely over effective enforcement, rather than the existence or not of legal frameworks.

VII. Conclusions

This article argues that FAMs constitute a little studied but important hybrid threat vector, operating at the intersection of disinformation transmission and the exploitation of

[114] European Commission, "Proposal for a Directive Amending Directive 2022/2555 as Regards Simplification Measures and Alignment with the [Proposal for the Cybersecurity Act 2]" (2026) COM(2026) 13 13.

[115] Ibid, 14.

[116] European Commission and High Representative of the European Union for Foreign Affairs and Security Policy, "European Democracy Shield: Empowering Strong and Resilient Democracies" (2025) JOIN(2025) 791.

[117] Ibid, 2–3.

[118] Ibid, 9.

[119] Ibid, 5.

cybersecurity vulnerabilities. By focusing on FAMs as both enablers of coordinated inauthentic behaviour and potential sources of malicious technological exploitation, the analysis advances current research on disinformation-associated risk in three significant ways. Firstly, it highlights FAMs as amplifiers capable of impacting the EU's social resilience through providing further channels for manipulation of online discourse. Secondly, it reveals that these same actors expose critical vulnerabilities in cyber-resilience, exploiting platforms and users through malicious code injection, data harvesting, and the weaponisation of engagement systems. Thirdly, the article situates these findings within the broader geopolitical landscape, illustrating how EU digital governance, particularly under the von der Leyen II Commission, now recognises hybrid interference as an integrated challenge requiring coordinated action across content regulation, cybersecurity and technological sovereignty frameworks.

This analysis shows that the EU's legal frameworks are increasingly robust and provide effective coverage, with enforcement, coordination and detection capability remained the decisive variables. Indeed, when compared to the UK's Online Safety Act as a close parallel, the emphasis on CIB and cyber-security adopted in the current EU *acquis* is distinct in its holistic approach. Addressing FAM-enabled threats therefore demands a more assertive, intelligence-driven regulatory posture that can potentially be driven by proposed reforms, within a broader strategic commitment to safeguarding Europe's democratic and digital infrastructures during times of geopolitical instability. This article concludes with two recommendations for future research: the first, a deeper investigation into the networks of agencies and actors that need to coordinate their actions to tackle FAMs within the EU *specifically*, as well as comparative work, assessing the approach to these issues taken in other jurisdictions *generally*. The author of this article intends for this to be the next stage of their work, drawing from the insights developed herein.

Financial support. This work has been funded by the European Media and Information Fund (EMIF), granted to FAMOUS: Fake Activity Market Observation for Unethical Services.

Competing interests. The author confirms that there are no conflicts of interest to disclose.

Cite this article: B Farrand, "Fake Activity Markets as Hybrid Threat Vector: Implications for EU Social and Cyber Resilience". *European Journal of Risk Regulation*. https://doi.org/10.1017/err.2026.10100