Finansuoja Europos Sąjunga
NextGenerationEU

Mykolo Romerio universitetas

NAUJOS KARTOS LIETUVA

Project „Misijomis grįstų mokslo ir inovacijų programų įgyvendinimas“ (Project Nr. 02-002-P-0001) report

Project name: ***HIPSTER***

Responsible/Implementation partner (-s): ***MRU***

Action result: ***Case, comparative and content analysis***

HIPSTer Case, Comparative, and Content analysis

Table of Contents

Introduction

The project "Hybrid, Information, Psychological, Societal Threats handling system for public security domain practitioners, businesses, and education" - HIPSTer aims to create a comprehensive solution for handling hybrid, information, psychological, and societal threats for the public security domain and businesses.

The HIPSTer project aims to develop and an advanced software solution – a Hybrid Information Psychological Societal Threats Handling System that utilizes recent advancements in OSINT, SocMINT, NLP, and AI to automatically detect, attribute, and counter or prevent prioritized threats, including hybrid threats.

This document provides a comprehensive case, comparative, and content analysis regarding implementation aspects related to HIPSTer implementation, taking into account privacy and cybersecurity aspects. The documents reviewed below indicate relevant policies, implementation guidance (AI Act & GDPR Interplay). The documents address the intersection and implementation challenges of the major EU digital laws, specifically covering cybersecurity, data protection, and AI ethics, some recent rulings from the Court of Justice of the European Union, etc. It provides regulatory context for the HIPSTer project, identifying compliance requirements and threat definitions across EU frameworks that shape system design. This document provides a comprehensive analysis of recent policy practices in the European Union (EU) and several national jurisdictions, particularly: Lithuania, Germany, The Netherlands, and France.

The case and comparative analysis involves examining real-world cases to gain insights into the implementation of similar projects. This approach can provide valuable information on how GDPR, AI Act and cybersecurity principles have been applied in practice and what challenges or limitations may exist. It involves comparing and contrasting different approaches to privacy and security legal issues, also in order to identify best practices or areas for improvement.

Content analysis includes recent policy initiatives and implementing documents related to the project's issues in artificial intelligence, data protection and other aspects.

1. Case and Comparative analysis

Technologies like HIPSTer project usually must not only meet high-level technical standards, but also meet the requirements of the legal environment in which it will operate. Analysis of the regulatory environment and analysis of key policy practices in the EU shows that this and all similar technologies must overcome certain regulatory and practical legal application challenges. Such challenges may manifest themselves in the form of legal uncertainty (non-specificity), especially in such complex domains as data protection, the security of artificial intelligence technologies, data collection and minimization, automatization, etc.

The goal of this case and comparative study is to explore real-world examples of hybrid threats and how they are currently being addressed. This will provide insights into the challenges and limitations of existing approaches and inform the development of new models and methods for detecting and attributing hybrid threats.

Although market already offers some solutions, which address hybrid threats, the products are not similar. Moreover, some of the existing products are distributed in non-EU markets, which results in a different regulatory environment.

Therefore, this case and comparative study consists of two main parts: (1) hypothetical case of the HIPSTer product with main regulatory compliance stages that must be addressed when developing a real-world product; (2) overview of several major real-world products, discussing key aspects of their regulatory compliance. It means that the scope of this case analysis involves not only real-world scenario but also a hypothetical scenario of the project implementation issues, related to the uncertainty of the regulatory environment, uncertainty in the broader field of activity: in the domain of social, hybrid, informational, cyber threat intelligence.

It has to be noted that some ventures like Recorded Future, Verint, Palantir Technologies, BAE Systems, IBM, Recorded Future, DigitalStakeout provide OSINT solutions, for digital risk protection, threat intelligence etc. Since the ventures operate in different markets, they face different regulatory challenges. Case studies of these market participants are discussed here to the extent that they relate to publicly available and unclassified information.

1.1. Overview of the Regulatory Jurisdiction

Since the HIPSTer project first of all aims to operate in the EU market, it is important to overview the regulatory jurisdiction of the EU in this area.

1.1.1. Supranational regulation domain of the EU

The EU is currently in a transition from a law-making to an active enforcement phase, where digital rules are becoming an interconnected “Digital Rulebook”. The following key acts apply to HIPSTer technology, which uses OSINT and AI to detect hybrid threats:

- Artificial Intelligence Act (AI Act)[1]: Establishes a risk-based approach. Practices that use subliminal, manipulative or deceptive techniques to distort behaviour and cause significant harm are prohibited (Article 5). High-risk systems (e.g. those used by public sector institutions) require a Fundamental Rights Impact Assessment (FRIA)[2]. General Purpose AI Models (GPAI) are subject to transparency and systemic risk management requirements.

[1] Artificial Intelligence Act: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689
[2] AI Act Service Desk. Fundamental Rights Impact Assessment: https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-27

- General Data Protection Regulation (GDPR)[3]: Regulates the processing of personal data. The key principles include data reduction, accuracy and security. The court stresses that even a short-term loss of control over data can be considered non-pecuniary damage.
- Cybersecurity legislation (NIS2[4] and Cyber Resilience Act – CRA)[5]: Requires ongoing risk management and resilience to hybrid threats such as sabotage or data poisoning.
- Digital Services Act (DSA)[6]: Regulates content moderation and imposes an obligation on VLOP platforms to label AI-generated content to prevent the spread of disinformation

1.1.2. Interaction of individual Member States and EU regulation

It has to be stressed that the legal domain of individual Member States can significantly affect the implementation of the project through three primary mechanisms: the interpretation of "National Security" exemptions, the specific transposition of EU directives (like NIS2), and divergent procedural laws regarding liability and surveillance.

- ***The "National Security" Exemption and Jurisdiction*** – The most critical factor is how individual Member States interpret the boundary between civilian and national security competences. Under Article 4(2) of the Treaty on European Union[7], national security remains the sole responsibility of each Member State. Consequently, the EU AI Act explicitly excludes AI systems developed or used *exclusively* for military, defense, or national security purposes from its scope.
- ***Fragmentation of Liability and Procedural Law*** – While the EU AI Act aims to harmonize rules, the legal landscape for liability and enforcement remains fragmented across Member States.
- ***National Transposition of Directives (The "Gold-Plating" Effect)*** – EU Directives require transposition into national law, allowing Member States to impose stricter or more specific requirements ("gold-plating").
- ***Divergent Administrative Governance*** – The practical application of the AI Act relies on national competent authorities: supervisory bodies, regulatory sandboxes, etc.

1.2. Hypothetical Case Study: “HIPSTer”

As it was mentioned, technologies like HIPSTer project usually must not only meet high-level technical standards, but also meet the requirements of the legal environment in which it will operate. Therefore, it is important to elaborate hypothetical case of the HIPSTer product with main regulatory compliance stages that must be addressed when developing a real-world product.

1.2.1. Compliance Checklist

Area	Legal Provision / Requirement
Prohibited Practices	Verify the system does not use manipulative or subliminal techniques to impair decision-making.
High-Risk Obligations	Conduct a Fundamental Rights Impact Assessment (FRIA) if the tool is used by a public authority.
Transparency	Provide meaningful information about the logic of automated decisions (e.g., why someone was flagged) to comply with the right of access.

[3] General Data Protection Regulation: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng
[4] NIS2 Directive: https://eur-lex.europa.eu/eli/dir/2022/2555/2022-12-27/eng
[5] The Cyber Resilience Act: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202402847
[6] Digital Services Act: https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng
[7] Treaty on European Union: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12012M/TXT

Data Integrity	Implement safeguards against "data/model poisoning" and unauthorized model inversion attacks.
Bias Monitoring	Monitor for algorithmic discrimination (e.g., gender or racial bias in threat detection) throughout the lifecycle.
Content Labeling	Ensure any AI-generated counter-measures or content are clearly marked as AI-generated.

1.2.2. Key Implementation Stages

Based on the regulatory environment described above, the implementation of HIPSTer should follow these critical stages:

- ***Socio-Technical Design & Risk Assessment*:** Evaluating threats not just as IT failures, but as social and psychological harms.
- ***Data Preparation and Bias Mitigation*:** Managing the tension between data protection and the need for sensitive data to correct AI bias.
- ***Adversarial Testing*:** Proactively testing the model against cyber-offense and harmful manipulation.
- ***Operational Monitoring and Communication*:** Continuous real-time detection and reputation management during crises.

1.2.3. Practical Steps at Each Stage

- ***Design Stage*:** Developers should create "psychological profiles" of potential attackers to better understand motives and design more resilient "socio-technical" defenses.
- ***Data Stage*:** Utilize the AI Act's specific exception (Article 10) that allows processing special categories of data (like ethnicity) only for bias monitoring and correction, while ensuring strict GDPR alignment.
- ***Compliance Stage*:** Use "regulatory sandboxes" where available, especially for SMEs, to navigate technical standards that currently lack sufficient implementation time.
- ***Operational Stage*:** Integrate automated OSINT and NLP tools to detect AI-driven phishing and social engineering in real-time. Conduct drills like Lithuania's "Cyber Shield OpEx,"[8] which specifically test an organization's ability to maintain public trust during disinformation attacks.

1.2.4. Practical Challenges for a Hypothetical Case Study

Scenario: A law enforcement agency uses HIPSTer to monitor OSINT for "hybrid campaigns" where foreign state actors recruit local proxies for physical sabotage via social media.

- ***Challenge 1:*** Regulatory Complexity and Bias Monitoring. The AI Act (Art. 10) permits using sensitive data to monitor bias, but GDPR usually forbids it. The agency faces a "regulatory complexity" where failing to monitor bias could lead to discriminatory targeting of certain groups, while doing so might violate GDPR restrictions.
- ***Challenge 2:*** The "Subliminal" vs. "Persuasion" Gap. The system is designed to "counter" threats. However, the AI Act prohibits manipulative techniques. Legally, there is a lack of specificity regarding when an automated counter-narrative campaign crosses the line from "lawful persuasion" to "prohibited manipulation" that distorts behavior.

[8] Report on the national cybersecurity exercise "Cyber Shield OpEx 2024" (Nacionalinių kibernetinio saugumo pratybų „Kibernetinis skydas OpEx 2024" ataskaita). https://www.nksc.lt/doc/KS2024_OPEX_Pratybu_ataskaita.pdf

- ***Challenge 3:*** Transparency vs. Security (The "Trade Secret" Trap). Under the Dun & Bradstreet (C-203/22) ruling[9], a citizen flagged by HIPSTer has the right to understand the logic and weighting of the data used to categorize them. The agency cannot use "trade secrets" as a blanket shield to hide how the AI works. However, revealing this logic could allow adversaries to "poison" the model or bypass its detection.
- ***Challenge 4:*** The Implementation Lag. While the AI Act and NIS2 are in force, the required technical standards for compliance (around 30 standards) have an "insufficient effective implementation period" of less than 6 months, though at least 12 months are realistically needed. This creates a high risk of "dispersed enforcement," where HIPSTer might be deemed compliant in Lithuania but not in France due to inconsistent interpretations.

1.3. Real-world case overview

The following analysis examines how established OSINT and threat intelligence technologies from companies like Recorded Future, Palantir Technologies, IBM, BAE Systems, and DigitalStakeout would be implemented for compliance with the EU AI Act, GDPR, and other EU requirements.

1.3.1. EU Regulatory Environment: Summary of Relevant Requirements

The EU regulatory landscape for these technologies is shifting from purely legislative to an active enforcement phase known as the "Digital Rulebook." These products must navigate several key acts:

- ***The AI Act***: Most OSINT and threat intel platforms fall under High-Risk classifications if used by public authorities for law enforcement or border control (relevant to Palantir, Recorded Future, and BAE Systems). The Act prohibits AI that uses subliminal, manipulative, or deceptive techniques to distort behavior or exploit vulnerabilities. General-purpose AI (GPAI) used within these platforms (e.g., Recorded Future AI or Verint Da Vinci) must meet strict transparency and risk assessment requirements.
- ***GDPR:*** These tools aggregate vast amounts of Publicly Available Information (PAI) and Commercially Available Information (CAI). GDPR requires data minimization, and recent CJEU rulings (e.g., C-203/22) establish that "trade secrets" cannot be used as a blanket reason to deny a citizen's right to understand the logic behind automated profiling.
- ***NIS2 and Cyber Resilience Act (CRA)***: These focus on "socio-technical" resilience, requiring platforms to protect against "data/model poisoning" – a risk where adversaries corrupt the training data of intelligence graphs.

1.3.2. Compliance Checklist for Intelligence Platforms

Area	Requirement / Provision	Relevant Product Context
Prohibited Practices	AI Act Art. 5: Prohibition of social scoring and manipulative AI.	Palantir Gotham and IBM i2 must ensure their link analysis doesn't lead to discriminatory social scoring.
Fundamental Rights	AI Act FRIA: Public bodies must conduct a Fundamental Rights Impact Assessment.	Required for BAE IntelligenceReveal or Palantir deployments in EU police forces (e.g., Denmark, Netherlands).
Data Integrity	CRA / NIS2: Resilience against adversarial AI and "model inversion."	Critical for Recorded Future's Intelligence Graph and DigitalStakeout's XTI.

[9] CJEU Ruling C-203/22 - Dun & Bradstreet Austria GmbH (2025). https://curia.europa.eu/juris/document/document.jsf?text=&docid=297932&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=8125471

Automated Logic	GDPR Art. 15/22: Meaningful information about the "logic" of processing.	Recorded Future AI and Verint Da Vinci AI must provide explainability for their automated risk scores.
Transparency	AI Act Art. 50: Labeling AI-generated content (e.g., deepfake detection or generated reports).	Cognyte's LUMINAR and Recorded Future's AI Reporting must clearly label AI outputs.

1.3.3. Most Important Implementation Stages

The reviewed regulatory environment suggest four critical stages for implementing these technologies within the EU framework:

- ***Socio-Technical Design***: Assessing the impact of the tool on democratic trust and individual rights before deployment.
- ***Data Normalization and Ethics***: Fusing disparate data (PAI/CAI) into ontologies while applying GDPR-compliant data minimization.
- ***Adversarial "Red-Teaming":*** Proactively testing the AI model against manipulation or "jailbreak" attempts.
- ***Operational Monitoring***: Continuous human-in-the-loop oversight to prevent the tool from becoming an autonomous "predictive policing" system.

1.3.4. Practical Steps at Each Stage

- ***Design Stage***: Implement "Privacy by Design." Palantir, for example, uses granular access controls and comprehensive auditing of all user activity to preserve data integrity.
- ***Data Stage***: Leverage AI Act Article 10 exemptions, which allow processing sensitive personal data (e.g., political opinions) *only* if necessary to monitor and correct algorithmic bias in threat detection.
- ***Compliance Stage***: Use "Regulatory Sandboxes." This is particularly important for newer AI features like Recorded Future AI to test compliance with technical standards that are still being finalized.
- ***Operational Stage***: Integrate Rapid Response Systems (per the DSA) to mitigate systemic risks like disinformation. DigitalStakeout uses geofencing and keyword tracking to detect narrative shifts in real-time.

1.3.5. Practical Challenges: Separate cases of particular ventures

1.3.5.1. Palantir Technologies (Gotham, Foundry, AIP)

Regulatory Summary: Palantir[10] operates primarily in High-Risk AI environments, serving national security, defense, and law enforcement (e.g., Europol, Denmark, Netherlands).

Compliance Profile:

- ***AI Act***: Its Gotham platform is used for predictive policing and crime analysis, which faces intense scrutiny regarding Art. 5 (prohibited practices). CEO Alex Karp argues it is an *analytical augmentation* tool requiring human judgment, rather than an autonomous predictive system.
- ***GDPR***: Palantir does not collect or sell data itself; it acts as a "technical band-aid" sitting on top of client systems. However, under CJEU C-203/22, the "Trade Secret" defense used for its link analysis logic is increasingly legally fragile.

[10] The most of the information about Palantir is retrieved from the official website of Palantir Technologies Inc.: https://www.palantir.com/

- ***Implementation Stages***: Focuses on Ontology Creation, transforming raw data into meaningful objects (people, places, events).
- ***Practical Steps***: Uses Apollo for continuous delivery across classified, hybrid, or air-gapped environments.
- ***Challenge: The "Black Box" controversy***. Despite auditing features, critics note these can be disabled by government users, complicating legal accountability.

1.3.5.2. Recorded Future (Intelligence Cloud)

Regulatory Summary: Positions itself as an Intelligence Cloud provider. With the introduction of Recorded Future AI (GPT-based)[11] in 2026, it must comply with General-Purpose AI (GPAI) transparency requirements.

Compliance Profile:

- ***AI Act***: Must provide technical documentation and transparency for its "AI Insights" and "AI Conversations".
- ***GDPR***: Its Identity Intelligence module monitors underground communities for stolen credentials, requiring strict adherence to data minimization and processing limits.
- ***Implementation Stages***: Follows a Threat Intelligence Maturity Model: Reactive → Proactive → Predictive → Autonomous.
- ***Practical Steps***: Uses its Intelligence Graph to index over 1 million sources, providing context for automated risk scoring.
- ***Challenge***: Implementation Lag. Transitioning to "Autonomous" operations (machine-speed response) creates a legal gray area regarding Art. 22 of the GDPR (Automated Decision Making).

1.3.5.3 IBM (Security X-Force & i2 Portfolio)

Regulatory Summary: IBM provides a hybrid of massive internal telemetry (X-Force) and established visual link analysis (i2)[12].

Compliance Profile:

- ***AI Act:*** High focus on Auditability. i2 Analyst's Notebook enables manual construction of link charts, providing a clear human-led audit trail for law enforcement.
- ***GDPR:*** The i2 Connect bundle integrates OSINT and PII data from providers like LexisNexis. IBM must ensure these integrations respect the "Right to Explanation" regarding data weighting.
- ***Implementation Stages:*** Tiered deployment (Essentials, Standard, Premium) based on the organization's SOC maturity.
- ***Practical Steps:*** Uses APIs to work on top of legacy systems, avoiding the need for total infrastructure replacement.
- ***Challenge:*** Connectivity vs. Privacy. Integrating disparate sources (PAI/CAI) into a single view makes it easier to "connect the dots" but increases the risk of "loss of control" over data, which the CJEU now views as a basis for non-material damage.

1.3.5.4. DigitalStakeout (XTI Platform)

[11] The most of the information about Recorded Future is retrieved from the official Recorded Future, Inc. website: https://www.recordedfuture.com/

[12] The most of the information about IBM's X-Force & i2 Portfolio is retrieved from the official IBM Corporation website: https://www.ibm.com/us-en

***Regulatory Summary*:** Specialized in Geo-Intelligence and real-time public agency monitoring[13].

Compliance Profile:

- ***AI Act***: Its DARIA AI Agent provides predictive detection. It must ensure this doesn't cross into "prohibited manipulation" when used by public agencies to combat misinformation.
- ***GDPR***: High sensitivity due to geofencing and location-specific narratives. Requires careful management of geolocation data to avoid illegal surveillance.
- ***Implementation Stages***: Modular feed activation – organizations only ingest the sources (Social, Dark Web, DNS) they need.
- ***Practical Steps***: Provides monthly payment options for enterprise-grade OSINT, lowering the barrier for smaller EU public sector entities.
- ***Challenge***: The Manipulation Gray Area. Detecting "community sentiment shifts" and "coordinated narratives" is useful for safety but poses a risk of legal liability if automated counter-measures are seen as distorting behavior.

1.3.5.5. Cognyte (LUMINAR & Nexa)

Regulatory Summary: A pure-play Investigative Analytics firm focused on national security and law enforcement (EMEA region)[14].

Compliance Profile:

- ***AI Act***: Its use of face and object detection in OSINT is classified as High-Risk, requiring a Fundamental Rights Impact Assessment (FRIA).
- ***GDPR***: Analyzes trends in the Deep/Dark Web and researches APT attacks. Must navigate strict EU rules on processing sensitive data for "bias monitoring".
- ***Implementation Stages***: A four-step process: Ingestion → Processing (AI-driven) → Analysis → Dissemination.
- ***Practical Steps***: Integrates GenAI (LUMINAR AI co-pilot) to help SOC teams contextualize threats faster.
- ***Challenge***: Regulatory Overlap. As Cognyte tracks the "blurring boundaries" between nation-states and ransomware gangs, it faces "regulatory complexity" where failing to monitor bias (per AI Act) might conflict with GDPR data minimization.

1.3.5.6. BAE Systems (IntelligenceReveal)

Regulatory Summary: Focuses on Multi-Domain ISR (Intelligence, Surveillance, and Reconnaissance) for Defense and Critical National Infrastructure (CNI)[15].

Compliance Profile:

- ***AI Act:*** Likely exempt for strictly military use, but IntelligenceReveal is used by law enforcement, triggering High-Risk requirements and the need for human oversight in administrative decisions.
- ***Cybersecurity (NIS2):*** Strong alignment with the Cyber Resilience Act due to its focus on IP metadata extraction and technical malware reverse-engineering.

[13] The most of the information about DigitalStakeout is retrieved from the official DigitalStakeout Inc. website: https://www.digitalstakeout.com/

[14] The most of the information about Cognyte is retrieved from the official Cognyte Software Limited website: https://www.cognyte.com/

[15] The most of the information about BAE Systems is retrieved from the official BAE Systems Plc website: https://www.baesystems.com/en

- ***Implementation Stages***: Integration of Cyber and Electromagnetic Activities (CEMA) for "information advantage".
- ***Practical Steps***: Deploying space-based ISR (Azalea™) for near real-time radio frequency mapping.
- ***Challenge***: Implementation Lag. Technical standards for space-based and high-assurance AI are still emerging, creating risk for long-term contract compliance.

1.3.5.7. Verint (Da Vinci AI)

Regulatory Summary: Post-Cognyte spin-off, Verint focuses exclusively on Customer Engagement[16].

Compliance Profile:

- ***AI Act:*** Da Vinci AI uses specialized bots (summarization, knowledge automation). Under Art. 50, these must be clearly identified as AI when interacting with humans.
- ***GDPR:*** Processes vast amounts of "Engagement Data" (interactions, experience metrics). It must maintain strict PII Redaction.
- ***Implementation Stages***: Injecting AI "bots" directly into existing business workflows.
- ***Practical Steps***: Using an Open AI approach to switch underlying models without requiring code changes from customers.
- ***Challenge***: Ethical Governance. Must maintain "continuous monitoring" to comply with evolving AI regulations while managing customer and employee experience data.

1.3.6. Concluding remarks on the real world case challenges

The companies surveyed in the OSINT and threat intelligence space face several key technological, regulatory, and operational challenges:

- Information overload and “data noise”: The biggest challenge for all companies is the sheer volume of data generated (about 2.5 quintillion bytes per day), which often results in analysts spending 80% of their time collecting data rather than analyzing it. Recorded Future highlights the risk of “death by data,” where security teams are overwhelmed by low-value messages. DigitalStakeout points out that fragmented intelligence across hundreds of platforms makes manual monitoring impossible.
- Regulatory hurdles and GDPR compliance: OSINT analysts must work within strict legal boundaries, particularly under the EU’s General Data Protection Regulation (GDPR), which restricts the processing of personal data. Palantir has faced constant criticism from civil liberties groups over potential mass surveillance and privacy violations.
- Transparency and the “black box” problem: Some companies, such as Palantir, have been criticized for their lack of transparency, with critical decision-making mechanisms hidden deep within the code, and users not having full control over how the system identifies suspicious events.
- High cost and complexity of implementation: Palantir’s Gotham price (about $141,000 per server core) is significantly higher than its competitors, making it affordable only to the largest agencies. Verint and IBM i2 point to the complexity of implementation in large enterprises and the need for specific training due to the steep learning curve.
- Threat evolution and bias: Threat actors are constantly changing platforms and methods, and the lines between state actors, cybercriminals, and activists are increasingly blurred. In addition, companies like Cognyte note that OSINT tools are prone to bias, so the information received must be carefully verified.

[16] The most of the information about Verint is retrieved from the official Verint Systems Inc. website: https://www.verint.com/

- Technical limitations: User feedback on Recorded Future indicates that it has difficulties integrating systems with tools like Okta, lacks bulk data loading functionality, and AI prediction features are sometimes inaccurate. In the review, users noted a potential slowdown in system speed when processing tasks.
- Skill shortage: IBM notes that the market lacks experienced cybersecurity professionals who can effectively manage complex intelligence platforms.

2. Content analysis – Recent policy practices

This content analysis – recent policy practices – address the intersection and implementation challenges of the major EU digital laws, specifically covering cybersecurity, data protection, and AI ethics, some recent rulings from the Court of Justice of the European Union, etc. It provides regulatory context for the HIPSTer project, identifying compliance requirements and threat definitions across EU frameworks that shape system design. This document provides a comprehensive analysis of recent policy practices in the European Union (EU) and several national jurisdictions, particularly: Lithuania, Germany, The Netherlands, and France.

2.1. EU Policy and implementation documents

This part of the document reviews the EU policy implementation, documents that address the intersection and implementation challenges of the major EU digital laws, specifically covering cybersecurity, data protection, and AI ethics. Some of the below reviewed sources include recommendations, Commission policies, standardization actions, ENISA recommendations etc.

1. Interplay between the AI Act and the EU digital legislative framework. Parliamentary Study. Policy Department for Transformation, Innovation and Health Directorate-General for Economy, Transformation and Industry. European Parliament, October 2025.

This study explores how the AI Act relates to various crucial pieces of EU digital legislation, such as the GDPR, the Data Act and the Cyber Resilience Act. It assesses overlaps and gaps between these acts, and shows that, while each of them is individually well targeted, their interplay creates significant regulatory complexity. Finally, it also provides reflections and suggestions for possible evolutions of the AI Act, and of EU digital legislation as a whole, keeping in mind the objective of ensuring that Europe can establish a competitive AI industry. This study was prepared at the request of the ITRE Committee.

Relevant Provisions:

- The document addresses "hybrid" issues in terms of technological integration and cybersecurity resilience, which are critical components of countering hybrid threats. The study analyzes "hybrid systems" that combine AI components with non-AI digital elements, particularly in the context of the Cyber Resilience Act (CRA). It highlights the complexity of assessing risks in these systems, noting that risks may arise in one component (e.g., AI training) but cause harm in another (e.g., downstream application), making "continuous risk management" essential rather than just one-off conformity assessments.
- The study notes that the AI Act requires specific cybersecurity measures to protect against AI-specific attacks, such as "data or model poisoning," which are methods often associated with hybrid interference campaigns.
- Generative AI Risks: The study highlights that while the AI Act regulates the generation of content (e.g., transparency and labeling), the Digital Services Act (DSA) focuses on the moderation of that content once hosted. It recommends harmonizing these regimes to prevent "dispersed enforcement" where content generators and hosts make different risk assessments.
- Manipulative and Subliminal Techniques: The AI Act classifies as "unacceptable risk" and bans AI practices that deploy "subliminal, manipulative, or deceptive techniques to distort behaviour and impair informed decision-making, causing significant harm".
- Exploitation of Vulnerabilities: It is prohibited to exploit vulnerabilities related to age, disability, or socio-economic circumstances to distort behavior in a way that causes significant harm.

- Emotion Recognition: The use of AI systems to infer emotions in workplaces or educational institutions is banned, except for specific medical or safety reasons.
- Social Scoring and Biometrics: The AI Act prohibits "social scoring" (evaluating individuals based on social behavior or traits) and "biometric categorisation" systems that infer sensitive attributes like race, political opinions, or sexual orientation.
- Deployers of high-risk AI systems (e.g., public bodies) must conduct Fundamental Rights Impact Assessments (FRIA) to assess risks to the fundamental rights of individuals before deployment. This creates a higher compliance threshold for protecting society than standard product safety laws.
- The study discusses the tension between the GDPR and the AI Act regarding bias monitoring. The AI Act allows the processing of special categories of personal data (e.g., ethnicity) specifically for "bias monitoring, detection and correction" to prevent societal discrimination, creating a complex interplay with GDPR restrictions.

Citation:

- "Obligations tied to the detection and disclosure of AI-generated or manipulated content are another interaction expressly noted in the AI Act. The DSA requires VLOPs [Very Large Online Platforms] and VLOSEs [Very Large Online Search Engines] to mitigate systemic risks related to such content where it “resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful”, e.g., by placing prominent markings on it. The AI Act’s Article 50, on the other hand, obliges providers of generative AI systems to ensure that AI-generated or manipulated content is marked as such." (Interplay between the AI Act and the EU digital legislative framework, p. 50)

2. Multilayer Framework for Good Cybersecurity Practices for AI. ENISA Recommendation. June 2023.

In this report, ENISA presents a scalable framework to guide NCAs and AI stakeholders on the steps they need to follow to secure their AI systems, operations and processes by using existing knowledge and best practices and identifying missing elements. The framework consists of three layers (cybersecurity foundations, AI-specific cybersecurity and sector-specific cybersecurity for AI) and aims to provide a step-by-step approach on following good cybersecurity practices in order to build trustworthiness in their AI activities.

Relevant provisions:

- The document primarily approaches relevant areas through a cybersecurity and technical risk management lens, treating AI systems as "socio-technical" entities where technical failures can have broad social and political consequences.
- State-Sponsored Actors: The report identifies "state-sponsored attackers" and "cyberwarriors" as specific threat actors with the means, motives, and opportunities to target AI systems.
- It warns that by 2030, the misuse of AI will be a significant threat, explicitly citing "state-sponsored operatives" who might issue "deep fakes" to deceive.
- The document links AI security to the EU’s Common Security and Defence Policy (CSDP), noting that AI is considered a technology that will play a "crucial role for defending the EU" against unconventional security threats.
- The report highlights "poisoning" as a major threat where attackers alter training data or models to modify an algorithm's behavior (e.g., to sabotage results or insert backdoors). Specific examples include data poisoning attacks on stop signs in autonomous driving scenarios.
- The text notes the risk of AI-generated content, such as deep fakes, which necessitates the ability to detect and contain these threats to assist security analysts. It identifies the risk of "model or data disclosure," where sensitive information about the model's configuration or training data is leaked, etc.

- The document addresses psychological aspects primarily in terms of understanding adversaries and the impact of AI on human trust and behavior. E.g., it emphasizes the need for a taxonomy of AI attackers that includes understanding their "psychological profiles," motives, and objectives.
- The document explicitly frames AI threats as "socio-technical," acknowledging that risks extend beyond technical failures to legal, ethical, and democratic concerns. E.g., "Bias" is identified as a specific threat that can lead to "algorithmic discrimination," where automated systems wrongly classify individuals or exclude them from services and rights
- The report asserts that adequate cybersecurity is necessary to "serve European values and the democratic rights of Europeans". It warns that social threats can impact "democracies and society as a whole". It calls for risk assessments that include social threats such as "lack of fairness, lack of interpretability/explainability/equality".

Citation:

- "As explained in the NIST AI Risk Management Framework, AI systems are socio-technical in nature, meaning that the threats are not only technical, legal or environmental (as in typical ICT systems), but social as well. For example, social threats – such as bias, lack of fairness, lack of interpretability/explainability/equality – are directly connected to societal dynamics and human behaviour in all technical components of an AI system, and they can change during its life cycle." (Multilayer Framework for Good Cybersecurity Practices for AI, p. 17)

3. Progress Report on the Digital Rulebook Implementation. Commission Policy. European Commission. September 2025.

This progress report highlights the main achievements and ongoing efforts to simplify and enhance the efficiency of Europe's digital rulebook, ensure its implementation and maintain its enforcement effectively. The Commission is conducting a comprehensive stress-test of EU digital rules, to ensure that they remain effective and efficient, and that their application is optimised in the real world.

With this in mind, the Commission will put forward a Digital Simplification Package to be adopted in the near future, including a Digital Omnibus to simplify and optimise rules on data, cyber and artificial intelligence, but also an EU Business Wallets to support the regulatory compliance and interactions between businesses and administrations. The Package will also launch a comprehensive digital fitness check.

Relevant provisions:

- The document addresses information and societal threats primarily through the implementation of the Digital Services Act (DSA) and the Artificial Intelligence Act (AI Act), with a strong focus on safeguarding democratic integrity and fundamental rights. To counter information threats, the "Code of Conduct on Disinformation" has been integrated into the DSA framework, utilizing a "Rapid Response System" to report time-sensitive content that threatens the integrity of elections, as demonstrated during the 2024 European and 2025 national elections.
- Regarding societal threats, the document emphasizes protecting "democracy, equality, the rule of law and human rights," ensuring that AI remains "safe and trustworthy," and mitigating systemic risks related to illegal content and public health on large platforms,.
- While the document does not explicitly use the term "hybrid threats," it covers the substance of this area through extensive cybersecurity provisions, including the "Cyber Solidarity Act" for managing large-scale incidents and the NIS2 Directive to enhance the resilience of critical sectors and digital infrastructure.

- Psychological threats are addressed implicitly through the priority placed on the "safety and security of minors," with specific enforcement actions taken against platforms hosting pornographic content to restrict access to services deemed a "high risk to minors" and to protect their well-being online.

Citation:

- "The Commission has given priority to the application of the DSA's provisions that seek to ensure privacy, safety and security of minors online. [...] One of the measures considered is to further restrict access to services or parts of the service which constitutes a high risk to minors, such as adult content services." (Progress Report on the Digital Rulebook Implementation, p. 4)

4. "Commission opens consultation on revising EU Cybersecurity Act", press release. European Commission. April 2025.

The initiative will revise the Cybersecurity Act, clarify the mandate of the EU Agency for Cybersecurity (ENISA) and improve the European Cybersecurity Certification Framework to achieve better resilience. The initiative also aims to streamline, simplify and supplement EU legislation to make the implementation of the EU cybersecurity framework more user and business friendly and to prioritise measures to support the EU objectives of developing a secure and resilient supply chain, including the EU cybersecurity industrial base.

5. Kilian R, Jäck L, Ebel D. European AI Standards – Technical Standardisation and Implementation Challenges under the EU AI Act. European Journal of Risk Regulation. 2025;16(3):1038-1062. doi:10.1017/err.2025.10032

This study, published in July 2025, presents one of the first systematic analyses of European technical and soon to be harmonised standardisation for organisations providing AI systems. Based on in-depth qualitative interviews with twenty-three leading European organisations developing AI applications across different sectors, such as Mistral and Helsing, and providing transparency regarding the status quo of draft standards, it examines how companies, especially start-ups and SMEs, are dealing with the contemplated standardisation under the EU AI Act and sectoral standardisation. Industry sectors covered include mobility, finance, manufacturing, healthcare, as well as defense and legal tech.

Relevant provisions:

- Key challenges identified comprise an insufficient effective implementation period of likely less than 6 months compared to at least 12 months actually required for around thirty technical standards, an imbalance of participation and influence in standardisation committees, double regulation and technical implementation hurdles as well as significant annual costs for harmonised standards compliance.
- Technical standards are currently reshaping global AI competition and will have a massive influence on the AI landscape as market entry barriers, particularly on start-ups. Hence, the paper offers policy recommendations based on the revealed challenges for AI providers.

Citation:

- "Key challenges identified comprise an effective implementation period of only approximately 6 months compared to at least 12 months actually required, an imbalance of participation and influence in standardization committees, double regulation and technical implementation hurdles as well as significant annual costs for harmonized standards compliance. Technical standards are currently reshaping global AI competition and will have a massive influence on the AI landscape as

market entry barriers, particularly on startups. Hence, the paper offers concrete policy recommendations based on the revealed challenges for AI providers." (Kilian R, Jäck L, Ebel D., p. 1)

6. Union Rolling Work Programme for European cybersecurity certification. Policy and Legislation. February 2024.

The Union Rolling Work Programme for European cybersecurity certification identifies strategic priorities for future European cybersecurity certification schemes. This first URWP points to areas where European cybersecurity certification schemes are envisaged due to legislative developments as well as to areas for future reflection regarding cybersecurity certification, which might eventually lead to requests for new schemes where necessary and appropriate. Furthermore, it outlines the strategic priorities to be considered when preparing any European cybersecurity certification scheme.

7. Fundamentals of Secure AI Systems with Personal Data. EDPB Guidance. April 2025.

This training/guidance material is critical for bridging the technical requirements of secure AI with the legal obligations of GDPR (Data Protection and Security). The Publication outlines a training curriculum for cybersecurity professionals working with personal data and artificial intelligence (AI) in the European Union (EU).

Relevant provisions:

- The document addresses information and societal threats primarily through its taxonomy of AI risks and ethical considerations. It identifies specific domains such as "Misinformation" and "Pollution of the information ecosystem," noting risks related to the spread of false information, the reinforcement of belief bubbles, and the "disinformation, surveillance, and influence at scale" conducted by malicious actors.
- It highlights "Socioeconomic & environmental harms" as a key risk area, warning against power centralization, increased inequality, the devaluation of human effort, and "Discrimination & toxicity," which includes unfair treatment and exposure to toxic content.
- Regarding psychological threats, the text explicitly references the EU AI Act's prohibition of AI systems that manipulate human behavior through "subliminal techniques" or exploit vulnerabilities based on age or disability to impair decision-making. It also warns of the "Loss of human agency and autonomy" and "Overreliance" on automated systems. While the specific term "hybrid threats" is not defined, the document covers overlapping concerns under "Malicious actors & misuse," describing the use of AI for cyberattacks, weapon development, and mass harm, which are components often associated with hybrid warfare strategies.

Citation:

„AI systems that manipulate human behavior through subliminal techniques, impairing decision-making ability. <...> AI systems that exploit vulnerabilities of individuals based on age, disability, or economic situation." (Fundamentals of Secure AI Systems with Personal Data. EDPB Guidance, p. 33)

8. Council Conclusions on the Future of Cybersecurity. Council of the EU. May 2024.

Adopted in May 2024, these conclusions emphasize the need for effective, non-fragmented implementation of NIS2 and the Cyber Resilience Act (CRA). They call for greater support for SMEs in cybersecurity compliance and explicitly acknowledge the dual benefits and challenges of AI and quantum computing for security.

The Council conclusions recall the importance to focus on implementation, strengthen coordination and collaboration, and avoid fragmentation of cybersecurity rules in sectorial legislation. They also call to further clarify roles and responsibilities in the cyber domain, to strengthen the cooperation in the fight against cybercrime, and to work on a revised blueprint of the cyber crisis management framework.

The support to micro, small and medium size enterprises, and the need to respond to the challenges presented by the new technologies are also highlighted. A multistakeholder approach, including cooperation with the private sector and academia is encouraged to close the skills gap. In light of the changed and rising threat level, the Council finally invites the European Commission and the High Representative to present a revised cybersecurity strategy.

Relevant provisions:

- The document explicitly addresses hybrid and information threats by calling for a coherent approach to risk assessment that takes into account "ongoing efforts to counter hybrid threats, such as physical sabotage and foreign information manipulation and interference".
- It emphasizes that the revised EU cybersecurity crisis management framework ("Blueprint") must be compatible with existing instruments like the "EU Hybrid Toolbox" and the "EU Cyber Diplomacy Toolbox" to effectively manage the full crisis lifecycle across civilian and military domains.
- Regarding psychological and societal threats, the text highlights the dangers of social engineering and the "misuse of emerging and disruptive technologies," specifically citing deepfakes created by AI as a threat to digital trust.
- The Council underscores that cybersecurity is the "cornerstone of a successful digital society," necessary to maintain "public trust" in the systems that underpin the functioning of the economy and society. Additionally, it notes the need to protect against "spill-over risk" where cyber threats impact the broader public and cross-border environments.

Citation:

- „Against the backdrop of increased misuse of emerging and disruptive technologies (such as AI, for example for the creation of deepfakes), the role of authenticated digital identity assumes augmented importance in strengthening digital trust and confidence.“ (Council Conclusions on the Future of Cybersecurity. Council of the EU, p. 11)

9. ENISA NIS360 2024 report: A comprehensive look at cybersecurity maturity and criticality of NIS2 sectors. March 2025.

The European Union Agency for Cybersecurity’s first NIS360 report identifies areas for improvement and tracking of progress across NIS2 Directive sectors. The NIS360 is a new product by the EU Agency for Cybersecurity, ENISA, that assesses the maturity and criticality of NIS2 sectors, providing both a comparative and a more in-depth analysis. The goal of the NIS360 is to help national authorities and cybersecurity agencies in the Member States tasked with the implementation of the NIS2, (1) to understand the overall picture, (2) to help them with prioritisation, (3) to highlight areas for improvement, and (4) to facilitate monitoring of sectors’ progress. The NIS360 also aims to support policy makers at national and EU level, to give input on policy and strategy development, and initiatives to build up cyber resilience.

Relevant provisions:

- The document focuses primarily on technical cybersecurity maturity and the implementation of the NIS2 Directive. However, regarding societal and hybrid-related concerns, the report identifies the Public Administrations sector as a "prime target for hacktivism and state-nexus operations," highlighting the risk of politically motivated actions and state-sponsored hybrid interference.

Citation:

- "Public administrations: [...] Being a prime target for hacktivism and state-nexus operations, the sector should aim to strengthen its cybersecurity capabilities leveraging the EU Cyber Solidarity Act and exploring shared service models among sector entities on common areas e.g., digital wallets." (ENISA NIS360 2024 report: A comprehensive look at cybersecurity maturity and criticality of NIS2 sectors)

10. ENISA Threat Landscape (ETL). October 2025.

This year's ENISA Threat Landscape (ETL) introduces a revised and concise format designed to deliver insights through a threat-centric approach and enhanced contextualisation. This edition integrates additional analysis of adversary behaviours, vulnerabilities and geopolitical drivers, aimed at both strategic and operational audiences, offering an actionable perspective on trends shaping the EU's cyber threat environment. The ETL 2025 provides an overview of the European cyber threat ecosystem from July 2024 to June 2025, drawing on nearly 4 900 selected and curated incidents. The reporting period highlights a maturing threat environment characterised by rapid exploitation of vulnerabilities and growing complexity in tracking adversaries. This flagship report highlights the shift toward diversified and convergent threat campaigns, which include several findings, e.g:

- AI-Enabled Threats: By early 2025, AI-supported phishing campaigns reportedly accounted for over 80% of observed social engineering, leveraging synthetic media and jailbroken models.
- Ransomware & State Actors: Ransomware remains the core threat, but state-aligned threat groups have intensified long-term cyberespionage against critical sectors like telecommunications and logistics.

Relevant provisions:

- The document extensively details hybrid and information threats, particularly highlighting the convergence of cyber operations with Foreign Information Manipulation and Interference (FIMI). It describes "hybrid campaigns" where adversaries, notably Russia-aligned actors, use digital platforms like Telegram to recruit individuals for physical sabotage, vandalism, and arson across NATO countries, thereby extending conflict beyond cyberspace.
- A dedicated section on FIMI outlines how state-aligned actors (Russia and China) employ "inauthentic news articles," "fabricated investigations," and AI-driven "deepfakes" to interfere in EU elections, discredit public institutions, and manipulate public perception.
- Regarding psychological and societal threats, the report warns of the "erosion of resilience" caused by campaigns designed to undermine democratic processes and social cohesion. Hacktivist and state-aligned groups exploit polarizing societal topics-such as migration and LGBTQ+ rights-to create division and target public administration and essential services.
- The text also highlights psychological manipulation through "social engineering" and "pig-butchering scams" that build false trust to defraud victims,, as well as "pressure tactics" by ransomware groups aimed at intimidation. Furthermore, the document notes tangible societal harms, ranging from the disruption of critical sectors like healthcare and transport to physical threats such as the kidnapping of crypto-asset holders.

Citation:

- „Hybrid campaigns […], especially with activities aligned with Russian objectives continuing to impact EU MSs beyond cyberspace. [...] investigative reporting detailed pro-Russia groups using Telegram to

recruit EU-based individuals for sabotage, vandalism, arson and influence operations across NATO countries.“ (ENISA Threat Landscape, p. 13)

- „Approximately a quarter of the documented FIMI content focused on degrading the Union through negative narratives. [...] France, Germany and Poland are frequently targeted with narratives aimed at discrediting their government, military and intelligence services, often accusing them of destabilisation efforts abroad or failing in their fundamental duties...“ (ENISA Threat Landscape, p. 43)

11. 2024 Report on the State of Cybersecurity in the Union. December 2024.

This document marks the first report on the state of cybersecurity in the Union, adopted by ENISA in cooperation with the NIS Cooperation Group and the European Commission, in accordance with Article 18 of the Directive (EU) 2022/2555 (NIS2). The report aims at providing policy makers at EU level with an evidence-based overview of the state of play of the cybersecurity landscape and capabilities at the EU, national and societal levels, as well as with policy recommendations to address identified shortcomings and increase the level of cybersecurity across the Union.

This report provides policy recommendations to strengthen the EU's cybersecurity framework. It Recommends strengthening support for the implementation of NIS2, and calls for the development of an EU horizontal policy framework to address supply chain security challenges faced by both public and private sectors.

Relevant provisions:

- The document identifies malicious cyber activity as a clear component of wider hybrid threats aimed at destabilizing EU society, democracy, and values, particularly through Foreign Information Manipulation and Interference (FIMI) and disinformation campaigns.
- State-aligned actors and non-state groups are increasingly leveraging AI to create fake content and conduct influence operations, with the geopolitical landscape-such as the Russian war of aggression against Ukraine-heavily influencing tactics designed to manipulate civilian populations and impact elections. These advanced hybrid threats, linked to interference and the dissemination of disinformation, are projected to remain top-ranking risks for the EU through 2030.
- Furthermore, the report highlights the psychological and societal dimensions of the threat landscape, noting that hacktivists utilize "Fear, Uncertainty, and Doubt" to amplify the disruptive impact of their operations. Incidents targeting civil society and the general public represented a notable share of observed events, often involving social engineering and data breaches.
- Consequently, the report emphasizes the critical need to strengthen "societal capabilities," recommending harmonized national efforts to improve cybersecurity awareness and cyber hygiene among citizens to ensure the population is resilient against these evolving threats.

Citation:

- „Malicious cyber activity has become a clear component of wider hybrid threats, such as disinformation and physical acts of sabotage and violence, seeking to undermine and destabilise EU society, democracy and values“. (2024 Report on the State of Cybersecurity in the Union, p. 14)
- „Hacktivists use common tactics, such as DDoS attacks and website defacements, but also “Fear, Uncertainty, and Doubt” to amplify the impact of their operations“. (2024 Report on the State of Cybersecurity in the Union, p. 16)

12. "European Democracy Shield and EU Strategy for Civil Society pave the way for stronger and more resilient democracies". Press Release. European Commission. November 2025.

The Commission has presented the European Democracy Shield, setting out a series of concrete measures to empower, protect, and promote strong and resilient democracies across the EU. An open civic space is at the core of our democracies, and this is why the Commission has also put forward an EU Strategy for Civil Society, for stronger engagement, protection and support to civil society organisations who play essential roles in our societies. Both initiatives had been outlined in the political guidelines and this year's State of the Union address by President von der Leyen. The European Democracy Shield and the EU Strategy for Civil Society present measures to protect the key pillars of our democratic systems: free people, free and fair elections, free and independent media, a vibrant civil society and strong democratic institutions.

Relevant provisions:

- The document addresses hybrid and information threats by establishing the "European Democracy Shield," which aims to counter Foreign Information Manipulation and Interference (FIMI) and disinformation campaigns often orchestrated by authoritarian regimes to exploit divisions and undermine EU values. Key provisions include the creation of a European Centre for Democratic Resilience to detect and respond to these threats, and the enforcement of the Digital Services Act to safeguard the integrity of the information space against large-scale, transnational information operations.
- Regarding psychological and societal threats, the text identifies specific efforts by malicious actors to "polarize citizens," "sow mistrust," and erode confidence in democratic institutions. To mitigate these risks, the Commission proposes a "whole-of-society approach" that focuses on boosting societal resilience through media and digital literacy for all ages and protecting the civic space.
- This includes measures to combat violence against political candidates and journalists, ensuring that civil society and independent media can function as pillars of a resilient democracy without facing abusive lawsuits or physical threats.

Citation:

- „Authoritarian regimes seek to exploit divisions, sow mistrust, and restrict democratic actors such as free media and civil society. In doing so, they erode trust in democratic institutions, undermine free and fair elections and challenge the very values on which the European Union is founded“. ("European Democracy Shield and EU Strategy for Civil Society pave the way for stronger and more resilient democracies". Press Release. European Commission. November 2025.)

13. European Parliament Motion for a Resolution on Hybrid Provocations. European Parliament. October 2025.

October 2025 motion strongly condemns Russia's military and hybrid provocations (e.g., drone incursions). This illustrates the evolving nature of hybrid threats into the physical and military domains.

Among other, EP calls on the Commission and the Member States to urgently increase investments in security and defence in order to close all capability gaps, establish strong deterrence and provide adequate support to Ukraine, ensuring that the EU and its Member States are equipped to address all threats, from hybrid and cyber threats to conventional military challenges, and that planning, research, development, procurement and management of capabilities are all done jointly and through a European lens; stresses, therefore, the importance of using these investments to stimulate joint action at European level, including with Ukraine, in order to improve the efficiency of public spending, enhance interoperability and boost the EU’s strategic autonomy, instead of perpetuating the present state of market fragmentation, divergent and incompatible capabilities, wasteful investments and external dependencies.

Relevant provisions:

- The resolution explicitly addresses hybrid and information threats by condemning Russia's systematic use of cyberattacks, disinformation campaigns, electoral interference, and the sabotage of critical infrastructure. It highlights that drones are increasingly used as tools for hybrid warfare by both state and non-state actors and notes the specific danger of Global Satellite Navigation Systems (GNSS) jamming and spoofing.
- To counter these evolving risks, the document urges the EU to adopt a comprehensive defense strategy that utilizes the Hybrid Threats Centre of Excellence and the European Cybersecurity Competence Centre to enhance cybersecurity, situational awareness, and the resilience of critical underwater and digital infrastructure.
- Regarding societal and psychological dimensions, the document emphasizes that security extends beyond military assets to include the protection of democratic values and social stability. It warns that vulnerabilities undermining citizens' trust in institutions could render the EU unstable despite its military strength, urging Member States to design defense investment plans that also strengthen social cohesion.
- The resolution calls for strategic communication to provide citizens with reliable information channels, ensuring that disinformation campaigns seeking to exploit security incidents and manipulate public perception are effectively countered.

Citation:

- "[The European Parliament] Calls for the EU and the Member States to adopt a comprehensive approach to defence, recognising that security extends beyond military assets and encompasses defence against acts of hybrid warfare; stresses that unless the vulnerabilities undermining citizens' trust in institutions and democracy are addressed, the EU risks being unstable and weak despite its military strength; urges the development of defence investment plans in a way that also strengthens social cohesion within and among regions and Member States" (European Parliament Motion for a Resolution on Hybrid Provocations, p. 7)

14. Eighth progress report on the implementation of the 2016 Joint framework on countering hybrid threats and the 2018 Joint communication on increasing resilience and bolstering capabilities to address hybrid threats. Joint staff working document. European Commission. October 2024.

This document is the Eighth Progress Report (SWD(2024) 233 final) from the European Commission and the High Representative, summarizing the EU's actions to counter hybrid threats and bolster resilience from July 2023 to June 2024. The report highlights efforts to protect democratic processes (especially the 2024 European elections) and critical infrastructure (energy, maritime, and space) against threats like FIMI (Foreign Information Manipulation and Interference), cyberattacks, and the instrumentalization of migration.

Key outcomes include validating the framework for EU Hybrid Rapid Response Teams, strengthening cybersecurity among EU institutions, advancing the legal frameworks of NIS 2 and CER Directives, and integrating the security of AI technologies into the economic security risk assessment.

Relevant provisions:

- The document outlines a comprehensive EU response to hybrid threats, which are increasingly driven by Russia's war of aggression against Ukraine and conflicts in the Middle East, utilizing tactics such as cyberattacks, the sabotage of critical infrastructure, and the instrumentalization of migration.
- To counter these evolving dangers, the EU has advanced its Hybrid Toolbox, notably validating the framework for EU Hybrid Rapid Response Teams to provide tailored assistance to Member States and partners. Significant emphasis is placed on information threats, specifically Foreign Information

Manipulation and Interference (FIMI) and disinformation; the report details measures taken to protect the integrity of the 2024 European elections, including the Defence of Democracy Package, the Digital Services Act, and the expansion of the Code of Practice on Disinformation to mitigate systemic online risks.

- Regarding societal and psychological security, the report highlights efforts to build resilience against radicalization, violent extremism, and hate speech, viewing these as vulnerabilities that can be exploited to destabilize society. The EU Internet Forum plays a key role in moderating terrorist content and addressing the emerging risks of generative AI mixed with disinformation.
- Furthermore, the document stresses a "whole-of-society" approach to psychological defense, promoting digital literacy and education to help citizens critically engage with the online world, thereby strengthening trust in democratic institutions against foreign interference.

Citation:

- "Six new national and regional European Digital Media Observatory (EDMO) hubs became operational in 2023, complementing the existing eight hubs to extend EDMO's reach and geographical coverage to the whole of EU. Their work is crucial to debunk and expose false claims and manipulated content, providing insight regarding disinformation, and supporting media literacy initiatives. The work by EDMO and the hubs has been instrumental in analysing disinformation on Russia's war of aggression against Ukraine and in the context of the Israel-Hamas conflict." (Eighth progress report on the implementation of the 2016 Joint framework on countering hybrid threats and the 2018 Joint communication on increasing resilience and bolstering capabilities to address hybrid threats, p. 4)

2.2. Core Implementation Guidelines

This part of the document covers the core implementation guidelines and related documents of the European Commission, the Commission's AI Office, the European Data Protection Supervisor, etc. The Commission, with the technical input of the AI Office, has issued several key guidelines to clarify the legal concepts that are now in force or will be soon. The AI Office plays a central coordinating role in developing the General-Purpose AI (GPAI) Code of Practice. The up-to date progress in the above-mentioned area can be monitored on the AI Office's website (https://digital-strategy.ec.europa.eu/en/policies/ai-office) with the focus on the upcoming deliverables, including Implementing Acts (which specify technical details) and further Guidelines related to other key AI Act provisions.

1. Guidance for Risk Management of Artificial Intelligence systems. European Data Protection Supervisor. November 2025.

This Guidance aims to guide EUIs acting as data controllers in identifying and mitigating some of these risks. More specifically, they focus on the risk of non-compliance with certain data protection principles elicited in the EUDPR for which the mitigation strategies that controllers must implement can be technical in nature – namely fairness, accuracy, data minimisation, security and data subjects' rights. As such, the technical controls listed in this Guidance are by no means exhaustive, and do not exempt EUIs from conducting their own assessment of the risks raised by their specific processing activities. In doing so, it refrains from ranking their likelihood and severity.

Relevant provisions:

- Hybrid threats are characterized by the convergence of cyber operations with physical acts of sabotage, such as the recruitment of individuals via social media for vandalism and arson, often

orchestrated by state-aligned groups to destabilize EU society. Information threats, specifically Foreign Information Manipulation and Interference (FIMI), are a primary concern.

- Regarding psychological and broader societal threats, the documents note the use of "Fear, Uncertainty, and Doubt" tactics by hacktivists to amplify the psychological impact of their operations, such as publicizing tampering with operational technology systems.
- Furthermore, AI systems are described as "socio-technical" entities that introduce specific societal risks, including algorithmic bias and discrimination. The EU legislative framework explicitly addresses these by prohibiting AI practices that deploy subliminal or manipulative techniques to distort behavior, thereby aiming to protect fundamental rights and democratic values from these evolving threats.

Citation:

- "This Guidance focus on two specific aspects of the risk management process, namely risk identification and risk treatment; the risk analysis and the risk evaluation aspects are too dependent on the specific processing context and their assessment is better left to each organisation in line with their own risk criteria. This means that EUIs should perform a thorough analysis for each AI system they plan to use in order to also evaluate the likelihood and impact of the risks, and decide on the mitigating measures to address them, as well as on the residual risks." (Guidance for Risk Management of Artificial Intelligence systems, p. 8)

2. Guidance on Generative AI, strengthening data protection in a rapidly changing digital era. European Data Protection Supervisor. October 2025.

These EDPS Orientations on generative Artificial Intelligence (generative AI) and personal data protection intend to provide practical advice and instructions to EU institutions, bodies, offices and agencies (EUIs) on the processing of personal data when using generative AI systems, to facilitate their compliance with their data protection obligations as set out, in particular, in Regulation (EU) 2018/1725. These orientations have been drafted to cover as many scenarios and applications as possible and do not prescribe specific technical measures. Instead, they put an emphasis on the general principles of data protection that should help EUIs comply with the data protection requirements according to Regulation (EU) 2018/1725.

Relevant provisions:

- The document addresses societal and security concerns through the lens of fundamental rights and data integrity. It warns that generative AI can exacerbate societal threats by magnifying existing biases, leading to "unfair processing and discrimination" in critical areas such as public health, human resources, and public administration.
- Regarding information and security risks, the guidance highlights the danger of "hallucinations" (the generation of false information) and technical vulnerabilities such as "prompt injection" and "jailbreaks" that could be exploited by malicious actors. It mandates human oversight to mitigate risks to vulnerable populations and to prevent "automation bias".

Citation:

- "Biases in generative AI systems can lead to significant adverse consequences for individuals' fundamental rights and freedoms, including unfair processing and discrimination, particularly in areas such as human resource management, public health medical care and provision of social services, scientific and engineering practices, political and cultural processes, the financial sector, environment and ecosystems as well as public administration." (Guidance on Generative AI, strengthening data protection in a rapidly changing digital era, p. 31)

3. First EDPS Orientations for EUIs using Generative AI. European Data Protection Supervisor. June 2024.

These EDPS Orientations on generative Artificial Intelligence (generative AI) and personal data protection intend to provide practical advice and instructions to EU institutions, bodies, offices and agencies (EUIs) on the processing of personal data when using generative AI systems, to facilitate their compliance with their data protection obligations as set out, in particular, in Regulation (EU) 2018/1725. These orientations have been drafted to cover as many scenarios and applications as possible and do not prescribe specific technical measures. Instead, they put an emphasis on the general principles of data protection that should help EUIs comply with the data protection requirements according to Regulation (EU) 2018/1725.

Relevant provisions:

- The document reflects concerns related to societal and information threats within the framework of fundamental rights and technical security. Regarding societal threats, the text warns that generative AI can magnify existing biases or create new ones, potentially leading to unfair discrimination in sensitive areas such as "political and cultural processes," "public administration," and the "provision of social services". It also highlights specific risks to "vulnerable populations and children," particularly in the context of automated decision-making.
- In terms of information threats, the guidance cautions against "hallucinations" (inaccurate or false information) and "harmful or toxic output", while also identifying specific security vulnerabilities such as "prompt injection," "jailbreaks," and "model inversion attacks" that can compromise system integrity.

Citation:

- "Controllers should, in addition to the traditional security controls for IT systems, integrate specific controls tailored to the already known vulnerabilities of these systems - model inversion attacks, prompt injection, jailbreaks - in a way that facilitates continuous monitoring and assessment of their effectiveness." (First EDPS Orientations for EUIs using Generative AI, p. 23)

4. Guidelines on the scope of obligations for providers of General-Purpose AI (GPAI) models. July 2025.

The Commission published guidelines on the scope of obligations for providers of general-purpose AI models under the AI Act. The aim is to provide legal certainty to actors across the AI value chain by clarifying when and how they are required to comply with these obligations. These guidelines are part of a broader package tied to the entry into application on 2 August 2025 of the EU-wide rules for providers of general-purpose AI models. They complement the General-Purpose AI Code of Practice that the Commission received from independent experts.

Relevant provisions:

- Document's primary focus is on determining the scope of obligations for general-purpose AI models-specifically regarding technical thresholds (such as training compute measured in FLOPs) and market classification-rather than analyzing specific geopolitical or psychological warfare tactics.
- The document addresses societal and information threats broadly through the regulatory framework of "systemic risk" and cybersecurity. Systemic risk is defined as a risk specific to high-impact capabilities that can cause negative effects on public health, safety, public security, fundamental rights, or "the society as a whole".

- Regarding information security, the guidelines specify that providers must report "serious incidents," which the AI Office interprets to include cybersecurity breaches such as "cyberattacks" and the "(self-)exfiltration of model parameters". Additionally, the text notes that even free and open-source licenses may include specific terms to restrict usage in applications that would pose a significant risk to public safety, security, or fundamental rights.

Citation:

- "The AI Act defines a 'systemic risk' as 'a risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain' (Article 3(65) AI Act)". (Guidelines on the scope of obligations for providers of General-Purpose AI (GPAI) models, paragraph 25)
- "The AI Office considers that this obligation covers serious cybersecurity breaches related to the model or its physical infrastructure, including the (self-)exfiltration of model parameters and cyberattacks, due to their possible implications for the obligations provided for in Article 55(1), points (b) and (d), AI Act". (Guidelines on the scope of obligations for providers of General-Purpose AI (GPAI) models, paragraph 103)
- "While the licence must generally permit use for any purpose, licensors may include specific, safety-oriented terms that reasonably restrict usage in applications or domains where such use would pose a significant risk to public safety, security, or fundamental rights". (Guidelines on the scope of obligations for providers of General-Purpose AI (GPAI) models, paragraph 84)

5. Guidelines on the AI system definition. February 2025.

The guidelines on the AI system definition explain the practical application of the legal concept, as anchored in the AI Act. The Commission publishes guidelines on AI system definition to facilitate the first AI Act's rules application. By issuing guidelines on the AI system definition, the Commission aims to assist providers and other relevant persons in determining whether a software system constitutes an AI system to facilitate the effective application of the rules. The guidelines on the AI system definition are not binding. They are designed to evolve over time and will be updated as necessary, in particular in light of practical experiences, new questions and use cases that arise.

Relevant provisions:

- The document acknowledges societal concerns by outlining the overarching aim of the AI Act, which is to ensure a high level of protection for "health, safety, and fundamental rights in the Union, including democracy and the rule of law". It also notes that the regulation follows a risk-based approach, where only systems posing "significant risks to fundamental rights and freedoms" are subject to the strictest regulatory obligations and prohibitions, though the guidelines do not elaborate on the specific nature of these societal risks.

Citation:

- "Aim is to promote innovation in and the uptake of AI, while ensuring a high level of protection of health, safety, and fundamental rights in the Union, including democracy and the rule of law". (Guidelines on the AI system definition, paragraph 1)
- "The AI Act's risk-based approach means that only those systems giving rise to the most significant risks to fundamental rights and freedoms will be subject to its prohibitions laid down in Article 5 AI

Act, its regulatory regime for high-risk AI systems covered by Article 6 AI Act [...]". (Guidelines on the AI system definition, paragraph 63)

6. Guidelines on prohibited artificial intelligence (AI) practices. February 2025.

These guidelines provide an overview of AI practices that are deemed unacceptable due to their potential risks to European values and fundamental rights. The guidelines specifically address practices such as harmful manipulation, social scoring, and real-time remote biometric identification, among others.

The guidelines are designed to ensure the consistent, effective, and uniform application of the AI Act across the European Union. While they offer valuable insights into the Commission's interpretation of the prohibitions, they are non-binding, with authoritative interpretations reserved for the Court of Justice of the European Union (CJEU). The guidelines provide legal explanations and practical examples to help stakeholders understand and comply with the AI Act's requirements. This initiative underscores the EU's commitment to fostering a safe and ethical AI landscape.

Relevant provisions:

- Psychological and information threats are addressed through prohibitions on AI systems that deploy "subliminal," "purposefully manipulative," or "deceptive techniques" intended to distort behavior and impair informed decision-making. The text highlights risks associated with exploiting "cognitive biases" and "psychological vulnerabilities," specifically warning against "neuro technologies" and "brain-computer interfaces" that could surreptitiously influence neural data or behavior.
- Regarding information threats, the guidelines target AI that presents "false or misleading information" (such as deep fakes) to deceive individuals, distinguishing these from non-prohibited "lawful persuasion" or unintended "hallucinations," provided adequate safeguards exist.
- Societal threats are centrally located in prohibitions against "social scoring" and the exploitation of vulnerable groups (based on age, disability, or socio-economic situation), which the document notes can lead to "social polarisation," "radicalisation," and broader harms to "democracy and the rule of law". While the document explicitly excludes AI used exclusively for hybrid or military threats (under the "national security" and "defence" exemptions), it clarifies that "dual use" systems remain within scope and acknowledges that prohibited practices may facilitate "terrorist attacks" or the dissemination of illegal content like "hate speech".

Citation:

- "Manipulative techniques are typically designed to exploit cognitive biases, psychological vulnerabilities, or situational factors that make individuals more susceptible to influence. Because of their adaptability, AI systems are also able to respond well to a person's individual circumstances or vulnerabilities and increase the effectiveness and impact of manipulation at scale". (Guidelines on prohibited artificial intelligence (AI) practices, paragraph 67)
- "'Deceptive techniques' deployed by AI systems should be understood to involve presenting false or misleading information with the objective or the effect of deceiving individuals and influencing their behaviour in a manner that undermines their autonomy, decision-making and free choices". (Guidelines on prohibited artificial intelligence (AI) practices, paragraph 70)
- "AI systems enabling 'social scoring' practices may lead to discriminatory and unfair outcomes for certain individuals and groups, including their exclusion from society, as well as social control and surveillance practices that are incompatible with Union values. [...] It also aims to safeguard and promote the Union values of democracy, equality (including equal access to public and private services), and justice". (Guidelines on prohibited artificial intelligence (AI) practices, paragraph 148)
- "The AI Act expressly excludes from its scope AI systems that are 'placed on the market, put into service, or used with or without modification exclusively for military, defence or national security

purposes, regardless of the type of entity carrying out those activities'". Guidelines on prohibited artificial intelligence (AI) practices, paragraph 22)

7. General-Purpose AI Code of Practice. July 2025.

The Code of Practice helps industry comply with the AI Act legal obligations on safety, transparency and copyright of general-purpose AI models. The General-Purpose AI (GPAI) Code of Practice is a voluntary tool, prepared by independent experts in a multi-stakeholder process, designed to help industry comply with the AI Act's obligations for providers of general-purpose AI models. The code was published on July 10, 2025. It is complemented by Commission guidelines on key concepts related to general-purpose AI models. The Commission and the AI Board have confirmed that the code is an adequate voluntary tool for providers of GPAI models to demonstrate compliance with the AI Act. Following the endorsement, AI model providers who voluntarily sign it can show they comply with the AI Act by adhering to the code. This will reduce their administrative burden and give them more legal certainty than if they proved compliance through other methods.

Relevant provisions:

- The document explicitly addresses the areas of concern through the framework of "systemic risks." Regarding hybrid and information threats, the Code identifies "Chemical, biological, radiological and nuclear (CBRN)" capabilities and "Cyber offence" (enabling attacks on critical infrastructure) as specified systemic risks. It further warns against information threats such as the generation of "illegal, violent, hateful, radicalising, or false content," specifically noting the model propensity to "hallucinate" or produce "misinformation" that could obscure truth or undermine democratic processes.
- In terms of psychological and societal threats, the document lists "Harmful manipulation" as a specified risk, defining it as the "strategic distortion of human behaviour or beliefs" through persuasion, deception, or personalized targeting that exploits individuals who cannot detect such influence. Societal threats are broadly covered under risks to "society as a whole," "fundamental rights" (including non-discrimination), and "public mental health".
- To manage these threats, the Code requires providers to implement "safety mitigations" and conduct "adversarial testing" (red-teaming) to evaluate the model's effectiveness against these risks throughout its lifecycle.

Citation:

- "Harmful manipulation: Risks from enabling the strategic distortion of human behaviour or beliefs by targeting large populations or high-stakes decision-makers through persuasion, deception, or personalised targeting. This includes significantly enhancing capabilities for persuasion, deception, and personalised targeting, particularly through multi-turn interactions and where individuals are unaware of or cannot reasonably detect such influence. Such capabilities could undermine democratic processes and fundamental rights". (General-Purpose AI Code of Practice, Safety and Security Chapter, p. 37)
- "Chemical, biological, radiological and nuclear: Risks from enabling chemical, biological, radiological, and nuclear (CBRN) attacks or accidents. [...] Cyber offence: Risks from enabling large-scale sophisticated cyber-attacks, including on critical systems (e.g. critical infrastructure)". (General-Purpose AI Code of Practice, Safety and Security Chapter, p. 37)

2.3. Court cases

Recent rulings from the Court of Justice of the European Union (CJEU) possibly shape the actual application and interpretation of the GDPR, with direct implications for AI systems and cybersecurity liability. Therefore, it is important to mention and discuss them in the context of this project.

1. C-634/21 (SCHUFA). Automated Decision-Making & AI Ethics (GDPR Art. 22). December 2023.

The CJEU held that a credit reference agency's practice of creating a credit score that is subsequently used by banks constitutes a "decision based solely on automated processing" that significantly affects a data subject. This ruling greatly expands the scope of Article 22 and mandates stricter transparency and human oversight for algorithmic scoring and AI systems used in consequential domains.

Relevant provisions:

- The court explained that the determination of a probability value ('credit rating' or 'scoring') of a person's ability to meet payment obligations by a credit reference agency constitutes 'automated individual decision-making' under the GDPR if that value is relied on to a significant extent by a third party in deciding on a contractual relationship.

2. C-413/23 P (EDPS v SRB). Pseudonymization and AI Model Training (GDPR). September 2025.

The CJEU clarified that pseudonymized data is not automatically considered personal data for a recipient if they cannot reasonably re-identify the individuals, taking into account all available means. This offers a more nuanced, context-specific view on data use that is important for AI model development and data sharing within the EU.

Relevant provisions:

- ***Legal Context.*** The dispute stems from a decision by the EDPS finding that the SRB failed to fulfill its obligations relating to the processing of personal data. This failure occurred during a procedure for granting compensation to shareholders and creditors of a banking institution following that institution's resolution.
- ***Key Legal Concepts***. The judgment focuses on the interpretation of Regulation (EU) 2018/1725, specifically addressing the Concept of 'personal data' (Article 3(1)), the Concept of 'pseudonymisation' (Article 3(6)), and the obligation to inform the data subject (Article 15(1)(d)) concerning the transmission of pseudonymised data to a third party.

3. C-807/21 (Deutsche Wohnen). GDPR Liability and Cyber-Fines. December 2023.

The CJEU ruled that administrative fines under the GDPR can only be imposed if the infringement is due to negligent or willful conduct. It rejects the concept of strict liability for GDPR violations, which is highly relevant for assessing liability and financial risk in the event of a cybersecurity breach.

Relevant provisions:

- ***The core issue*** addressed by the Court concerned the interpretation of provisions of Regulation (EU) 2016/679 (General Data Protection Regulation or GDPR). Specifically, the ruling focused on the conditions for imposing administrative fines on a legal person, analyzing the concepts of 'controller' (Article 4(7)), the powers of supervisory authorities (Article 58(2)), and the imposition of administrative fines (Article 83).

- ***Key Holdings (Operative Part).*** The Court interpreted Article 58(2)(i) and Article 83 of Regulation 2016/679, concluding two main points regarding administrative fines levied against controllers that are legal persons and undertakings:

1. Requirement of Identifying a Natural Person: Article 58(2)(i) and Article 83(1) to (6) of the GDPR must be interpreted as precluding national legislation which permits an administrative fine to be imposed on a legal person (in its capacity as controller) for an infringement (referred to in Article 83(4) to (6)) only if that infringement has previously been attributed to an identified natural person.
2. Requirement of Intent or Negligence: Article 83 of Regulation 2016/679 must be interpreted as meaning that an administrative fine may be imposed pursuant to that provision only where it is established that the controller (which is both a legal person and an undertaking) intentionally or negligently committed the infringement referred to in Article 83(4) to (6).

4. C-446/21 (Meta Platforms Ireland). Data Minimization and Profiling (GDPR). October 2024.

The Court reinforced the principle of data minimization, ruling that the GDPR limits a social network's use of a user's personal data collected outside that platform for targeted advertising. This is crucial for defining the ethical and legal boundaries of data-driven profiling and AI-assisted marketing practices.

Relevant provisions:

- The principle of data minimization (Article 5(1)(c) GDPR) requires data to be adequate, relevant, and limited to what is necessary for the processing purposes. This principle reflects the principle of proportionality.
- The principle of storage limitation (Article 5(1)(e) GDPR) requires data to be kept only for as long as is necessary for the purposes for which they were collected or further processed.
- The controller must refrain from collecting data in a generalised and indiscriminate manner and must avoid collecting data that are not strictly necessary for the purpose of the processing.
- The extensive processing of user data, monitoring a large part of online activities, is characterized by a serious interference with fundamental rights (Articles 7 and 8 of the Charter).
- The indiscriminate use of all of the personal data held by a social network platform for advertising purposes, irrespective of the data's sensitivity, does not appear to be a proportionate interference with users' rights.
- The storage of the personal data of users for an unlimited period for the purpose of targeted advertising must be considered to be a disproportionate interference.
- ***The Court concluded*** that the data minimization principle precludes any personal data collected by the operator (from the data subject or third parties, on or outside the platform) from being aggregated, analyzed, and processed for targeted advertising without restriction as to time and without distinction as to type of data.

5. C-203/22 (Dun & Bradstreet Austria GmbH). GDPR Article 15 (Right of Access) & Trade Secrets in Automated Decision-Making. February 2025.

The CJEU clarified the extent of the "meaningful information about the logic involved" required under Article 15(1)(h). It ruled that controllers must provide sufficient detail on the procedures and principles applied (e.g., what data was used and how it was weighted) to enable comprehension, even if it does not require disclosing the full algorithm (a trade secret). Crucially, it mandates that trade secret claims cannot be used to "as a rule" exclude the right of access.

Relevant provisions:

- ***The core issue*** concerned the interpretation of Article 15(1)(h) of Regulation (EU) 2016/679 (General Data Protection Regulation or GDPR). This provision relates to a data subject's right of access to information, specifically in the context of automated decision-making, including profiling. The case involved scoring and the assessment of the creditworthiness of a natural person.
- ***Key Holdings (Operative Part).*** The Court interpreted Article 15(1)(h) of the GDPR concerning the required level of transparency regarding automated processing and the balance between the right of access and the protection of trade secrets or third-party data.
- ***Meaningful Information about the Logic Involved (Profiling).*** Regarding automated decision-making and profiling within the meaning of Article 22(1) of the GDPR, the Court ruled:
    - The data subject may require the controller to explain the procedure and principles actually applied when using personal data by automated means to obtain a specific result, such as a credit profile.
    - This explanation must be provided by means of relevant information and must be in a concise, transparent, intelligible, and easily accessible form. This explanation constitutes the "meaningful information about the logic involved" referenced in Article 15(1)(h).
- ***Balancing Right of Access with Trade Secrets/Third-Party Data.*** The Court addressed situations where the controller believes the required information contains data of third parties protected by the GDPR or constitutes trade secrets (within the meaning of point 1 of Article 2 of Directive (EU) 2016/943). The Court ruled:
    - The controller is required to provide the allegedly protected information to the competent supervisory authority or court.
    - That authority or court must then balance the rights and interests at issue with a view to determining the extent of the data subject's right of access provided for in Article 15 of the GDPR.
    - In essence, the controller cannot unilaterally deny the data subject access by claiming trade secrets or third-party data protection but must submit the information to an independent authority for balancing the competing rights.

6. C-200/23 (Agentsia po vpisvaniyata v OL). Non-Material Damage and Loss of Control (GDPR Art. 82). 2024.

The Court ruled that a loss of control over personal data for a limited period (in this case, data made publicly available online) may be sufficient to constitute non-material damage. This lowers the threshold for claims under Article 82, which is highly relevant for the liability assessment of cybersecurity incidents and data breaches.

Relevant provisions:

- ***Key Interpretations and Holdings.*** The Court provided interpretation on several key points concerning data protection rights when documents containing personal data are publicly disclosed in a commercial register:
- ***Disclosure Obligation (Directive 2017/1132).*** Article 21(2) of Directive 2017/1132 (concerning company law) does not impose an obligation on a Member State to permit the disclosure in the commercial register of a company's constitutive instrument containing personal data other than the minimum personal data required by law. This provision primarily concerns the voluntary disclosure of translations, not the content of the data itself.
- ***Status of the Register Authority (GDPR Controller/Recipient).*** The authority responsible for maintaining the commercial register (the Agency) is considered both a 'recipient' and a 'controller' of the personal data contained in the disclosed constitutive instrument. This classification applies particularly in so far as the authority makes the data available to the public, even where that instrument contains personal data not required by Directive 2017/1132 or by national law.

- ***Right to Erasure (Article 17 GDPR):***
  - The processing of personal data not required by Directive 2017/1132 or national law is unlikely to satisfy the conditions for lawfulness under Article 6(1)(c) or (e) of the GDPR (legal obligation or public interest task), as such processing goes beyond what is necessary to achieve the objectives of the directive.
  - National legislation or practice that leads the commercial register authority to refuse any request for erasure of non-required personal data-merely because a redacted copy of the constitutive instrument was not previously provided to the authority-is precluded by Directive 2017/1132 (Article 16) and Article 17 of the GDPR.
- ***Handwritten Signature as Personal Data (Article 4(1) GDPR).*** The Court ruled that the handwritten signature of a natural person is covered by the concept of 'personal data' within the meaning of Article 4(1) of the GDPR, as it identifies the person and is linked to the documents it authenticates.
- ***Non-Material Damage (Article 82(1) GDPR):***
  - A loss of control, for a limited period, by the data subject over their personal data, due to those data being made available online in the commercial register, may suffice to cause 'non-material damage' under Article 82(1) of the GDPR.
  - However, the data subject must demonstrate that they have actually suffered such damage, however minimal. This concept does not require the existence of additional tangible adverse consequences.
- ***Exemption from Liability (Article 82(3) GDPR).*** An opinion issued by a supervisory authority under Article 58(3)(b) of the GDPR (which is advisory and not legally binding) is not sufficient to exempt the controller (the register authority) from liability under Article 82(2) and (3) of the GDPR. The controller's liability is fault-based, and to be exempt, the controller must prove that it is in no way responsible for the event giving rise to the damage.

2.4. National sources of the Republic of Lithuania

Overview of the national policy practices of the Republic of Lithuania related to cybersecurity, data protection (GDPR) and artificial intelligence (AI Act) allows us to assess national progress in the context of this project and to identify potential shortcomings in political practice.

1. National Cybersecurity Status Report 2024 (Nacionalinė kibernetinio saugumo būklės ataskaita 2024). https://www.nksc.lt/doc/Nacionaline-kibernetinio-saugumo-ataskaita-2024.pdf

The 2024 National Cybersecurity Status Report, submitted by the National Cybersecurity Center (NKSC) together with other institutions, shows a significant increase in the number of cyber incidents – a 63% increase per year. The report reveals that the largest share of incidents (59%) is made up of attacks based on social engineering. In addition, it is emphasized that the increasing threat arises from espionage carried out by groups supported by foreign countries. The report also emphasizes the importance of preventive measures, mentioning tools such as the malicious domain blocking tool "Vasaris" and the activities of distance learning platforms.

Relevant provisions:

- The report identifies hybrid and information warfare as central elements of the current security landscape, driven primarily by hostile regimes in Russia and Belarus aimed at destabilizing the state. These actors employ hybrid attacks, espionage, and propaganda not only to damage critical infrastructure but also to undermine Lithuania's strategic interests, specifically its support for Ukraine and NATO integration.
- The document highlights the use of advanced technologies, including artificial intelligence and deepfakes, to generate disinformation that portrays Lithuania as a "failed state" or "Russophobic,"

thereby attempting to discredit national defense efforts and question the reliability of allied protection.

- Psychological and societal threats are inextricably linked to these operations, which are designed to polarize society, erode trust in democratic institutions, and instill a pervasive sense of insecurity or fear regarding potential global conflict.
- The report notes that hostile propaganda aims to break the public's will to resist by convincing the domestic audience that the country is not worth defending. Furthermore, a significant majority of cyber incidents (59%) rely on social engineering, exploiting human psychology and trust to manipulate individuals, while information attacks increasingly feature intimidation tactics, such as narratives about nuclear war.

Citation:

- "The main goal of the messages sent by Russia and Belarus about Lithuania and NATO is to encourage distrust in one's state, its institutions and one another, [and] to polarize society. This attempts to influence the stability of Western democratic states and the ability of citizens to make decisions". (National Cybersecurity Status Report 2024)

2. Report on the national cybersecurity exercise “Cyber Shield OpEx 2024” (Nacionalinių kibernetinio saugumo pratybų „Kibernetinis skydas OpEx 2024“ ataskaita).

The National Cyber Security Centre under the Ministry of National Defence (hereinafter referred to as the NKSC) organised the annual national cyber security exercise “Cyber Shield OpEx 2024” on 7-18 October 2024. The exercise aimed to develop practical cybersecurity skills of the exercise participants, test cyber incident management procedures, improve cooperation between institutions managing and/or investigating cyber incidents and cyber security entities, and develop public communication skills.

80 organisations providing critical services to the population of Lithuania successfully participated in the exercise. Most (65) of these organisations managed incidents in a virtual cyber training ground environment, while 39 organisations actively improved their public communication skills. The lessons identified in the exercise have been recorded and will be taken into account when organising future cyber security exercises.

Relevant provisions:

- The exercise scenario modeled hybrid and information threats by depicting "hostile states" supporting "hacktivist groups" that launched attacks in retaliation for Lithuanian legislative decisions. These simulated attacks included "content distortion" on organizational websites aimed at causing confusion among employees and clients, as well as Distributed Denial of Service (DDoS) attacks to disrupt critical services provided to the population.
- Psychological and societal pressure tactics were integrated into the simulation through "public pressure" campaigns on social media designed to force ransom payments and the distribution of phishing emails to target citizens. The exercise explicitly focused on testing "public communication" and "media" responses, evaluating how organizations manage reputational damage and maintain public trust while under pressure from these simulated hostile information operations.

3. Resolution of the Seimas of the Republic of Lithuania “On the Principles of the Use of Artificial Intelligence Technologies in the Public Sector” (Lietuvos Respublikos Seimo rezoliucija „Dėl dirbtinio intelekto technologijų naudojimo viešajame sektoriuje principų“). TAR, 2024-05-16, Nr. 8947.

The resolution establishes principles and guidelines for the application of AI in the public sector, emphasizing transparency, openness, equality and ethics. This is an important political document defining AI ethical guidelines at the national level.

Relevant provisions:

- Regarding societal concerns, the resolution addresses the potential risks of AI integration into public administration rather than external security threats. It establishes safeguards to protect society from technological misuse, mandating principles such as "human oversight" to control administrative decisions, "equality" to prevent discrimination based on attributes like race or political views, and "transparency" to ensure the public is informed about AI influence. The document emphasizes the "primacy of human interests," ensuring that AI serves as a tool for social good rather than a source of societal destabilization.

Citation:

- “Only the proper use of artificial intelligence tools can ensure the protection of all human rights and freedoms, the country's economic and national security interests”. (Resolution of the Seimas of the Republic of Lithuania “On the Principles of the Use of Artificial Intelligence Technologies in the Public Sector”)

4. Methodological information (guidelines, recommendations, etc.) of the State Data Protection Inspectorate (Valstybinės duomenų apsaugos inspekcijos metodinė informacija – gairės, rekomendacijos ir kt.).

The State Data Protection Inspectorate (SDI) is preparing a wide range of methodological information, guidelines and recommendations to help data controllers (both public and private sectors) and data processors properly implement the requirements of the General Data Protection Regulation (GDPR) and the Law on the Legal Protection of Personal Data of the Republic of Lithuania (LPPD). The SDI methodological material acts as a consulting aid, helping organizations not only to avoid violations, but also to actively implement data protection already at the design stage (privacy by design).

Relevant provisions:

- "Recommendation on Prevention of Cyber Incidents" addresses psychological threats through the lens of social engineering (phishing), describing it as a method that exploits "human qualities" and trust to deceive victims into revealing sensitive information or transferring money. Additionally, it covers ransomware, which functions as a form of psychological pressure and blackmail (šantažas), forcing victims to decide whether to pay criminals to recover encrypted data. The provided provisions focus on mitigation strategies for these specific tactical threats, such as recognizing fake websites, refusing to pay ransoms, and reporting incidents to the police or the National Cyber Security Centre.

Citation:

- "Cyber attacks based on social engineering principles are a method whereby the aim is to exploit human qualities and extract certain sensitive information, e.g., login data for social network accounts, information systems, bank accounts [...]." (Recommendation on Prevention of Cyber Incidents).

2.5. National sources of selected EU Member States

Identifying "best policy practices" in the overlapping fields of cybersecurity, GDPR, and the AI Act involves looking at national governments' strategic planning, the issuance of proactive regulatory guidance, and the establishment of effective institutional frameworks.

Based on recent national documents and policy initiatives focusing on AI Act and NIS2 implementation, France, Germany, and the Netherlands exhibit leading practices, particularly in providing detailed guidance on the intersection of data protection and AI.

France

France's Data Protection Authority, the Commission Nationale de l'Informatique et des Libertés (CNIL), is a key benchmark for providing clear, actionable guidance on reconciling AI development with GDPR principles. This proactive approach offers legal clarity before the AI Act's high-risk systems obligations become fully applicable. CNIL published several recommendations to promote the responsible use of AI while ensuring compliance with personal data protection.

These recommendations confirm that GDPR requirements are sufficiently balanced to address the specific challenges of AI. They provide concrete and proportionate solutions to inform individuals and facilitate the exercise of their rights.

1. CNIL Recommendations on AI and GDPR Compliance.

Relevant provisions:

- The recommendations take into account the EU Artificial Intelligence Act recently adopted. Indeed, where personal data is used for the development of an AI system, both the GDPR and the AI Act apply. CNIL's recommendations have therefore been drawn up to supplement them in a consistent manner regarding data protection.

Germany

Germany's strength lies in its long-term strategic integration of cybersecurity into its national AI and industrial policy, particularly emphasizing security-by-design for critical infrastructure:

2. Cyber security strategy for Germany. 2021.

The Cyber Security Strategy for Germany 2021 creates a strategic framework for Federal Government policy on cyber security for the next five years, subject to the availability of budgetary funds.

Relevant provisions:

- Cyber Security Strategy sets out four crosscutting guiding principles:

1. Establishing cyber security as a joint task for government, private industry, the research community and society.
2. Reinforcing the digital sovereignty of government, private industry, the research community and society.
3. Making digital transformation secure.
4. Setting measurable, transparent objectives.

- The strategic objectives are implemented in three action areas:
    - Action Area 1 – “Remaining safe and autonomous in a digital environment”;
    - Action Area 2 is “Government and private industry working together”;

- Action Area 3 – "Strong and sustainable cyber security architecture for every level of government".

3. The Federal Government's Artificial Intelligence Strategy and AI Action Plan (and related information, programmes, initiatives, etc.).

The Federal Cabinet adopted the Federal Government's Artificial Intelligence Strategy (AI Strategy) in 2018. The AI Strategy is designed to strengthen Germany's position in the international competition for research, development and applications in AI. The AI Strategy is updated to take account of new developments and needs.

Relevant provisions:

- Various different initiatives now focus on the issues of pandemic response, sustainability, environmental protection, climate action and national and international networking. In 2023 the AI Action Plan was launched as the update to the AI Strategy.

The Netherlands

The Netherlands is recognized for its proactive approach to regulating algorithmic risks through its Data Protection Authority (Autoriteit Persoonsgegevens or AP) and its clear vision for institutional cooperation on AI Act enforcement.

Every six months, the Dutch Data Protection Authority (AP) publishes its Report AI & Algorithms Netherlands (RAN). It also draws observations and analyses, based on the knowledge gathered through the network, conversations and consultations with experts and society.

4. AI & algorithmic risks: developments in the Netherlands (and related information, links, initiatives, programmes, etc.).

5. AP and RDI: Supervision of AI systems requires cooperation and must be arranged quickly (and links to related information).

6. Guidelines on AI Literacy.

Relevant provisions:

- The use of generative AI raises legal concerns. One aspect of this is the training of models on large quantities of 'scraped' data, which can include personal data. Additionally, new risks for users and the increased concentration of power in big tech firms demand scrutiny.
- Many countries identify large scale spreading of disinformation and manipulation through the use of Generative AI as risks.
- The output of Generative AI is based on probability estimates. The user does often not have insight into uncertainty, plausible alternative answers, or references to origins of the content. This makes it hard to interpret the output and increases the risk of incorrect conclusions and actions. This can result in discrimination and arbitrariness in a person's or organisation's approach.
- Training a generative AI system with data containing unbalanced information of groups, possibly without realising it, may reinforce people's biases. After all, they will then be shown 'overrepresented' groups more often. For example, AI-generated images may be disproportionately more likely to show men as doctors and women as nurses.

- The development of generative foundation models takes place at a small number of organisations. Mainly big tech firms have sufficient financial means to invest in this. This makes it hard for new providers to enter the market, and reinforces the existing power of big tech companies.

Conclusions

1. Main Conclusions

The aggregated information portrays a security landscape where the boundaries between digital, physical, and cognitive warfare have effectively dissolved. The EU's response shifts from mere regulation to active implementation of a complex "Digital Rulebook" designed to counter four interconnected threat domains:

1.1. Hybrid Threats are Kinetic and Digital: Hybrid warfare has evolved beyond simple cyberattacks to include the physical sabotage of critical infrastructure, the instrumentalization of migration, and the use of drones. There is a clear convergence where state-aligned actors (specifically Russia) use digital platforms like Telegram to recruit proxies for physical vandalism and arson within NATO countries, making "cyber" defense inseparable from physical security.

1.2. Recent national data validates the dominance of the psychological vector in hybrid warfare. For instance, the 2024 National Cybersecurity Status Report from Lithuania reveals that 59% of all cyber incidents were based on social engineering. These operations are not merely opportunistic but are strategically driven by hostile regimes (specifically Russia and Belarus) to polarize society and erode the will to resist by portraying the state as a 'failed' entity. This confirms that the 'cognitive' domain is now the primary battlefield for state-aligned actors.

1.3. Information Warfare as a Strategic Weapon: Foreign Information Manipulation and Interference (FIMI) is now treated as a systemic risk to democracy, often synchronized with cyber operations. Multiple documents highlight that Generative AI acts as a force multiplier for these threats, enabling the mass production of "hallucinations," deepfakes, and "toxic output" that can obscure truth and disrupt crucial areas of public life.

1.4. Psychological and Societal Vulnerabilities: The focus of EU policy has expanded to the "socio-technical" layer. Threats are no longer viewed solely as technical failures but as targeted attempts to exploit human psychology through "fear, uncertainty, and doubt" tactics, social engineering, and AI systems that use "subliminal techniques" to impair decision-making. Societal risks include the erosion of trust in public institutions, radicalization, and the magnification of bias against vulnerable groups.

1.5. Regulatory Interplay and Complexity: There is a significant emphasis on the "interplay" between major legislative acts-specifically the AI Act, GDPR, Digital Services Act (DSA), and the Cyber Resilience Act (CRA). While these laws are individually robust, their overlapping requirements create regulatory complexity that complicates compliance, particularly regarding bias monitoring and the reporting of systemic risks.

1.6. The implementation of the 'Digital Rulebook' is currently being reshaped by CJEU case law, which introduces both stricter compliance mandates and enforcement hurdles. While the *Schufa* ruling (C-634/21) significantly expands the definition of 'automated decision-making' to include credit scoring used by third parties, the *Deutsche Wohnen* ruling (C-807/21) rejects strict liability for GDPR fines, requiring proof of intent or negligence. This judicial requirement may complicate the 'harmonized enforcement' envisioned in strategic plans, as regulators must now meet a higher evidentiary threshold to penalize non-compliant entities.

2. Strategic Insights

2.1. The "Socio-Technical" Paradigm Shift: A recurring theme is that cybersecurity can no longer be purely technical. Documents explicitly frame AI and security systems as "socio-technical," acknowledging that technical vulnerabilities (e.g., prompt injection) directly translate into social harms (e.g., discrimination or loss of democratic trust). Defense strategies must therefore include "psychological profiles" of attackers and impact assessments on fundamental rights, not just IT infrastructure.

2.2. The socio-technical shift is legally cemented by the recent *Agentsia po vpisvaniyata* ruling (C-200/23), which established that a mere 'loss of control' over personal data constitutes non-material damage,

even without financial loss. Furthermore, the *Dun & Bradstreet* ruling (C-203/22) explicitly prioritizes transparency over commercial interests, stating that trade secrets cannot be used to deny access to the logic behind automated decisions. This confirms that cybersecurity strategies must prioritize fundamental rights, as legal liability now extends beyond technical breaches to the psychological impact on citizens.

2.3. AI as a Dual-Use Catalyst: AI is presented as both the primary vector for new threats (e.g., "polymorphic" malware, deepfakes, automated social engineering) and the essential tool for defense (e.g., detecting large-scale FIMI). The "General-Purpose AI Code of Practice" reflects this by categorizing "cyber offence" and "harmful manipulation" as systemic risks that require "adversarial testing" (red-teaming) throughout the model's lifecycle.

2.4. The systemic risks of Generative AI are exacerbated by data provenance issues. As noted by the Dutch Data Protection Authority, models trained on 'scraped' data do not merely reproduce information but can technically entrench social discrimination (e.g., consistently depicting men as doctors and women as nurses). Additionally, the high cost of training foundation models is leading to a concentration of power in 'Big Tech' firms, creating market entry barriers that contradict the EU's goal of digital sovereignty.

2.5. The Implementation Gap: There is a tension between the ambition of EU laws and the practical reality of implementation. The analysis highlights that technical standards are lagging, with an "insufficient effective implementation period" (less than 6 months compared to the needed 12), creating high barriers to entry for startups and SMEs. This "dispersed enforcement" risks leaving gaps that hybrid actors can exploit.

Recommendations for HIPSTer compliance

Based on the challenges and gaps identified in the aggregation of recent practices, the following recommendations are pertinent:

3.1. Harmonize Enforcement Mechanisms: To prevent "dispersed enforcement," EU Member States must align the risk assessment criteria used under the DSA (for content hosting) with those under the AI Act (for content generation). A unified approach is needed to ensure that content labeled as high-risk by an AI provider is treated with equal urgency by the platform hosting it.

3.2. Prioritize "Whole-of-Society" Resilience: Investments in defense should not be limited to military hardware. Following the European Parliament's call, funding must extend to strengthening "social cohesion" and media literacy. This includes educating citizens to recognize "social engineering" and "pig-butchering scams," thereby reducing the population's susceptibility to psychological manipulation.

3.3. Operationalizing resilience requires moving beyond technical drills to 'reputation management' simulations. Best practices from Lithuania's 'Cyber Shield OpEx 2024' demonstrate the value of integrating public communication and media response teams into cyber exercises. Effective defense requires testing an organization's ability to maintain public trust while under pressure from simulated 'hostile information operations' and ransom demands, ensuring a cohesive response to the psychological dimensions of a crisis.

3.4. Support for SMEs and Standardization: Policymakers must address the "significant annual costs" and administrative burdens placed on SMEs. Providing "sandboxes" or simplified compliance frameworks for the AI Act and Cyber Resilience Act is crucial to prevent the stifling of European innovation while maintaining security standards.

3.5. Operationalize Automated Threat Detection: The aggregation of recent practices explicitly validates the need for systems like the "HIPSTer" project. Given the volume and speed of AI-driven threats (e.g., 80% of social engineering being AI-supported), manual monitoring is insufficient. Public security practitioners need automated solutions utilizing OSINT and NLP to detect and attribute hybrid attacks in real-time. Most OSINT and threat intel platforms fall under High-Risk classifications if used by public authorities for law enforcement or border control . The Act prohibits AI that uses subliminal, manipulative, or deceptive techniques to distort behavior or exploit vulnerabilities. General-purpose AI used within such platforms must meet strict transparency and risk assessment requirements.

3.6. Analogy: The aggregated information suggests that the EU is transitioning from functioning as an "architect" (designing the laws like the AI Act and DSA) to a "city manager" (enforcing traffic rules, fixing leaks, and policing the streets). The blueprints are finished, but the challenge now lies in preventing traffic

jams (regulatory overlap) and stopping saboteurs who are trying to dig up the roads (hybrid threats) while ensuring the citizens (society) remain safe and trusting of the infrastructure.

ANNEX: Detailed Compliance Check for HIPSTer

Part 1: AI Act Compliance

Requirement	Legal Basis	Checkpoint Question
Prohibited Practices	AI Act Art. 5	Does the system avoid subliminal or deceptive techniques to distort behavior?
Social Scoring	AI Act Art. 5	Does the system refrain from evaluating individuals based on social behavior leading to discriminatory outcomes?
Transparency	AI Act Art. 50	Is AI-generated or manipulated content clearly labeled as such?
FRIA	AI Act	If deployed by a public body, has a Fundamental Rights Impact Assessment been completed?
Systemic Risk	AI Act Art. 3(65)	Are risks with significant impact on public security or society identified and managed?

Part 2: GDPR and Data Protection Compliance

Requirement	Legal Basis / Case Law	Checkpoint Question
Data Minimization	GDPR Art. 5(1)(c)	Is data collection limited strictly to what is necessary for threat detection?
Right to Explanation	GDPR Art. 15(1)(h); C-203/22	Can the agency provide the logic and data weighting used for a profile without relying solely on "trade secrets"?
Automated Decisions	GDPR Art. 22; C-634/21	Is there "human-in-the-loop" oversight for decisions that significantly affect individuals?
Bias Monitoring	AI Act Art. 10 / GDPR	Is sensitive data used *only* for bias monitoring and correction?
Non-material Damage	GDPR Art. 82(1); C-200/23	Is it understood that a temporary loss of control over data can lead to liability?

Part 3: Cybersecurity and Resilience (NIS2, CRA)

Requirement	Legal Basis / Guidelines	Checkpoint Question
Poisoning Defense	CRA / AI Act	Are measures in place to protect against data or model poisoning?
Socio-Technical Assessment	ENISA Recommendations	Does the risk assessment include social threats like the erosion of democratic trust?
Adversarial Testing	GPAI Code of Practice	Is regular "red-teaming" performed against cyber-offense and disinformation?
Incident Reporting	NIS2 / AI Act	Is there a procedure to report serious incidents (e.g., parameter exfiltration) to the AI Office?

Part 4: Hybrid Threat and Disinformation Management (DSA)

Requirement	Legal Basis	Checkpoint Question
Rapid Response	DSA / Code of Practice	Is the tool integrated into mechanisms for identifying Foreign Information Manipulation (FIMI)?
Societal Resilience	EU Literacy Guidelines	Does the technology help build resilience against social engineering and deepfakes?
Reputation Management	"Cyber Shield OpEx"	Is there a communication plan to maintain public trust during a crisis?

References

1. Interplay between the AI Act and the EU digital legislative framework. Parliamentary Study. Policy Department for Transformation, Innovation and Health Directorate-General for Economy, Transformation and Industry. European Parliament, October 2025. https://www.europarl.europa.eu/RegData/etudes/STUD/2025/778575/ECTI_STU(2025)778575_EN.pdf
2. Multilayer Framework for Good Cybersecurity Practices for AI. ENISA Recommendation. June 2023. https://www.enisa.europa.eu/sites/default/files/publications/Multilayer%20Framework%20for%20Good%20Cybersecurity%20Practices%20for%20AI.pdf
3. Progress Report on the Digital Rulebook Implementation. Commission Policy. European Commission. September 2025. https://commission.europa.eu/document/download/68237940-a9e3-460c-bf49-932d432a6bba_en?filename=VIRKKUNEN_APR_SPI_2025_70_EN.pdf
4. 4. "Commission opens consultation on revising EU Cybersecurity Act", press release. European Commission. April 2025. https://digital-strategy.ec.europa.eu/en/news/commission-opens-consultation-revising-eu-cybersecurity-act
5. Kilian R, Jäck L, Ebel D. European AI Standards – Technical Standardisation and Implementation Challenges under the EU AI Act. European Journal of Risk Regulation. 2025;16(3):1038-1062. doi:10.1017/err.2025.10032
6. 8. Union Rolling Work Programme for European cybersecurity certification. Policy and Legislation. February 2024. https://digital-strategy.ec.europa.eu/en/library/union-rolling-work-programme-european-cybersecurity-certification
7. Fundamentals of Secure AI Systems with Personal Data. EDPB Guidance. April 2025. https://www.edpb.europa.eu/system/files/2025-06/spe-training-on-ai-and-data-protection-technical_en.pdf
8. Council Conclusions on the Future of Cybersecurity. Council of the EU. May 2024. https://data.consilium.europa.eu/doc/document/ST-10133-2024-INIT/en/pdf
9. ENISA NIS360 2024 report: A comprehensive look at cybersecurity maturity and criticality of NIS2 sectors. March 2025. https://www.enisa.europa.eu/news/enisa-nis360-2024-report
10. ENISA Threat Landscape (ETL). October 2025. https://www.enisa.europa.eu/sites/default/files/2025-10/ENISA%20Threat%20Landscape%202025_0.pdf
11. 2024 Report on the State of Cybersecurity in the Union. December 2024. https://www.enisa.europa.eu/sites/default/files/2024-11/2024%20Report%20on%20the%20State%20of%20the%20Cybersecurity%20in%20the%20Union.pdf
12. "European Democracy Shield and EU Strategy for Civil Society pave the way for stronger and more resilient democracies". Press Release. European Commission. November 2025. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2660
13. European Parliament Motion for a Resolution on Hybrid Provocations. European Parliament. October 2025. https://www.europarl.europa.eu/doceo/document/B-10-2025-0437_EN.pdf
14. Eighth progress report on the implementation of the 2016 Joint framework on countering hybrid threats and the 2018 Joint communication on increasing resilience and bolstering capabilities to address hybrid threats. Joint staff working document. European Commission. October 2024. https://defence-industry-space.ec.europa.eu/system/files/2025-01/SWD_Annual-Progress-Report-2024.PDF
15. Guidance for Risk Management of Artificial Intelligence systems. European Data Protection Supervisor. November 2025. https://www.edps.europa.eu/data-protection/our-work/publications/guidelines/2025-11-11-guidance-risk-management-artificial-intelligence-systems_en

16. Guidance on Generative AI, strengthening data protection in a rapidly changing digital era. European Data Protection Supervisor. October 2025. https://www.edps.europa.eu/data-protection/our-work/publications/guidelines/2025-10-28-guidance-generative-ai-strengthening-data-protection-rapidly-changing-digital-era_en
17. First EDPS Orientations for EUIs using Generative AI. European Data Protection Supervisor. June 2024. https://www.edps.europa.eu/data-protection/our-work/publications/guidelines/2024-06-03-first-edps-orientations-euis-using-generative-ai_en
18. Guidelines on the scope of obligations for providers of General-Purpose AI (GPAI) models. July 2025. https://digital-strategy.ec.europa.eu/en/library/guidelines-scope-obligations-providers-general-purpose-ai-models-under-ai-act
19. Guidelines on the AI system definition. February 2025. https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application
20. Guidelines on prohibited artificial intelligence (AI) practices. February 2025. https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act
21. General-Purpose AI Code of Practice. July 2025. https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai
22. C-634/21 (SCHUFA). Automated Decision-Making & AI Ethics (GDPR Art. 22). December 2023. https://curia.europa.eu/juris/document/document.jsf;jsessionid=0EE9D6699C75E48B9657DC80D9166286?text=&docid=282187&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=8119755
23. C-413/23 P (EDPS v SRB). Pseudonymization and AI Model Training (GDPR). September 2025. https://curia.europa.eu/juris/document/document.jsf?text=&docid=305584&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=8121207
24. C-807/21 (Deutsche Wohnen). GDPR Liability and Cyber-Fines. December 2023. https://curia.europa.eu/juris/document/document.jsf?text=&docid=282192&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=8123480
25. C-446/21 (Meta Platforms Ireland). Data Minimization and Profiling (GDPR). October 2024. https://curia.europa.eu/juris/document/document.jsf?text=&docid=290674&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=8907758
26. C-203/22 (Dun & Bradstreet Austria GmbH). GDPR Article 15 (Right of Access) & Trade Secrets in Automated Decision-Making. February 2025. https://curia.europa.eu/juris/document/document.jsf?text=&docid=297932&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=8125471
27. C-200/23 (Agentsia po vpisvaniyata v OL). Non-Material Damage and Loss of Control (GDPR Art. 82). 2024. https://curia.europa.eu/juris/document/document.jsf?text=&docid=290701&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=1164237
28. National Cybersecurity Status Report 2024 (Nacionalinė kibernetinio saugumo būklės ataskaita 2024). https://www.nksc.lt/doc/Nacionaline-kibernetinio-saugumo-ataskaita-2024.pdf
29. Report on the national cybersecurity exercise "Cyber Shield OpEx 2024" (Nacionalinių kibernetinio saugumo pratybų „Kibernetinis skydas OpEx 2024" ataskaita). https://www.nksc.lt/doc/KS2024_OPEX_Pratybu_ataskaita.pdf
30. Resolution of the Seimas of the Republic of Lithuania "On the Principles of the Use of Artificial Intelligence Technologies in the Public Sector" (Lietuvos Respublikos Seimo rezoliucija „Dėl dirbtinio intelekto technologijų naudojimo viešajame sektoriuje principų"). TAR, 2024-05-16, Nr. 8947. https://www.e-tar.lt/portal/lt/legalAct/611a93c0135e11efbcbfb318996800a8
31. Methodological information (guidelines, recommendations, etc.) of the State Data Protection Inspectorate (Valstybinės duomenų apsaugos inspekcijos metodinė informacija – gairės, rekomendacijos ir kt.). https://vdai.lrv.lt/lt/naudinga-informacija/rekomendacijos-gaires-ir-kt/work/

32. CNIL Recommendations on AI and GDPR Compliance. https://www.cnil.fr/en/ai-and-gdpr-cnil-publishes-new-recommendations-support-responsible-innovation and https://www.cnil.fr/en/ai-how-comply-regulations
33. Cyber security strategy for Germany. 2021. https://www.bmi.bund.de/SharedDocs/downloads/EN/themen/it-digital-policy/cyber-security-strategy-for-germany2021.pdf?__blob=publicationFile&v=4
34. The Federal Government's Artificial Intelligence Strategy and AI Action Plan (and related information, programmes, initiatives, etc.). https://www.bmftr.bund.de/EN/Research/EmergingTechnologies/ArtificialIntelligence/artificialintelligence_node.html and https://www.bmftr.bund.de/DE/Forschung/Schluesseltechnologien/KuenstlicheIntelligenz/KiAktionsplan/kiaktionsplan_node.html
35. AI & algorithmic risks: developments in the Netherlands (and related information, links, initiatives, programmes, etc.). https://www.autoriteitpersoonsgegevens.nl/en/themes/algorithms-ai/ai-algorithmic-risks-developments-in-the-netherlands
36. AP and RDI: Supervision of AI systems requires cooperation and must be arranged quickly (and links to related information). https://www.autoriteitpersoonsgegevens.nl/en/current/ap-and-rdi-supervision-of-ai-systems-requires-cooperation-and-must-be-arranged-quickly
37. Guidelines on AI Literacy. https://www.autoriteitpersoonsgegevens.nl/documenten/aan-de-slag-met-ai-geletterdheid
38. Artificial Intelligence Act: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689
39. AI Act Service Desk. Fundamental Rights Impact Assessment: https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-27
40. General Data Protection Regulation: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng
41. NIS2 Directive: https://eur-lex.europa.eu/eli/dir/2022/2555/2022-12-27/eng
42. The Cyber Resilience Act: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202402847
43. Digital Services Act: https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng
44. Treaty on European Union: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12012M/TXT
45. Palantir Technologies Inc. https://www.palantir.com/
46. Recorded Future Inc. https://www.recordedfuture.com/
47. IBM Corporation. https://www.ibm.com/us-en
48. DigitalStakeout Inc. https://www.digitalstakeout.com/
49. Cognyte Software Limited. https://www.cognyte.com/
50. BAE Systems Plc. https://www.baesystems.com/en
51. Verint Systems Inc. https://www.verint.com/

Editor
<<Main Editor>>

Contributors
Giedrė Sabaliauskaitė (MRU)
R. Andrew Paskauskas (MRU)
Tomas Lavišius (MRU)

Reviewers
Giedrė Sabaliauskaitė (MRU)
R. Andrew Paskauskas (MRU)
Tomas Lavišius (MRU)

Disclaimer