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:* "Literatūros apžvalga"**

Editor

<<Main Editor>>

Contributor

R. Andrew Paskauskas (MRU)

Reviewers

Giedrė Sabaliauskaitė (MRU)

R. Andrew Paskauskas (MRU)

Tomas Lavišius (MRU)

Table of Contents

Executive Summary

This document presents the Literature Review deliverable produced by the HiPSTer project (Hybrid-threat Intelligence Platform leveraging Semantic Technologies for Threat Recognition) for the Lithuanian Innovation Agency (MITA), under the project "Implementation of Mission-Based Science and Innovation Programs" (Project No. 02-002-P-0001). The deliverable surveys the academic and policy scholarship relevant to the design, deployment, and governance of hybrid threat detection systems operating within the European regulatory environment, with particular attention to privacy and cybersecurity dimensions as specified in the project brief.

The review was conducted as a PRISMA-ScR-informed scoping review combining two sources of evidence: material adapted from the HiPSTer project's scientific publication (forthcoming in *Open Research Europe*) and the project team's prior comprehensive review in the *Journal of Intelligent Communication*, covering the technical state of the art and adversarial context; and four targeted Web of Science Boolean searches executed for the present deliverable, covering data protection and fundamental rights, the EU Artificial Intelligence Act, cybersecurity governance frameworks (NIS2, the Cyber Resilience Act, DORA), and the Digital Services Act.

The review establishes five principal findings. First, the EU cybersecurity and AI regulatory landscape has consolidated into a coherent but structurally fragmented ecosystem in which no single instrument can be analysed in isolation. Second, cross-regulation integration has emerged as the central methodological challenge of contemporary EU compliance scholarship, with compliance-by-design methodologies offering the most promising response. Third, ontology-driven approaches to both regulatory compliance and operational threat detection have moved from research curiosity to demonstrated practical applicability, providing external validation for the architectural philosophy underlying HiPSTer's own design. Fourth, the hybrid regulatory architectures that combine technical compliance requirements with fundamental-rights protection – the GDPR, the AI Act, the Cyber Resilience Act – are subject to convergent doctrinal critique that contemporary system designers must take seriously. Fifth, hybrid threats systematically exploit the seams between analytical disciplines, regulatory instruments, and civilian and security operational contexts, requiring defensive frameworks that address all three simultaneously.

For the HiPSTer project specifically, the review supports the conclusion that the project's architectural choices are not idiosyncratic but rather exemplify architectural patterns that the contemporary scholarly and regulatory landscape increasingly demands. The deliverable serves as the foundational scholarship layer underpinning the project's broader suite of deliverables, providing the evidence base on which the project's compliance-prescriptive sibling deliverables can build. The review concludes that hybrid threat detection systems can be designed to operate within – rather than despite – the European fundamental-rights and data-protection frameworks, provided that regulatory alignment is treated as an architectural design problem from the outset rather than as a downstream compliance activity.

Introduction

1.1 Purpose and Scope of the Deliverable

This document is the Literature Review deliverable produced by the HiPSTer project for the Lithuanian Innovation Agency (MITA). Its purpose is to survey the academic and policy scholarship relevant to the design, deployment, and governance of hybrid threat detection systems operating within the European regulatory environment, providing the foundational scholarship layer on which the project's broader deliverable suite rests.

The deliverable is one of nine principal artefacts the HiPSTer project is producing for the Innovation Agency. The others include a GDPR best-practices analysis with accompanying Data Protection Impact Assessment, an AI Act compliance research report, a consolidated case-study and comparative-analysis document, privacy-by-design summary reports across the system lifecycle, a HIPSTer prototype testing report and AI compliance assessment, dissemination and publication records, and the project's scientific publication itself (forthcoming in *Open Research Europe*). The literature review is positioned to complement these compliance-prescriptive and operational deliverables by establishing the underlying scholarly and regulatory evidence base.

The scope is bound accordingly. The deliverable does not produce compliance prescriptions for HiPSTer (the territory of the GDPR and AI Act deliverables); it does not produce operational privacy-by-design walkthroughs (the territory of the summary reports prepared by Lavišius); and it does not produce a comparative analysis of case-study data (the territory of the consolidated case-study document prepared by Lavišius). What it does produce is a survey of the scholarship on which those deliverables can ground their positions, with explicit attention to ensuring adequate coverage of the privacy and cybersecurity dimensions specified in the project brief.

1.2 Project Context

The HiPSTer project – Hybrid-threat Intelligence Platform leveraging Semantic Technologies for Threat Recognition – addresses a persistent analytical gap in detecting contemporary hybrid threats. Hybrid threats exploit the interplay between informational, behavioural, psychological, and technical domains, with meaningful signals frequently dispersed across analytical disciplines that current systems treat in isolation. The HiPSTer methodological framework develops an ontology-driven approach to cross-domain semantic integration, combining Web Ontology Language (OWL) ontologies for formal threat-actor and campaign modelling, Simple Knowledge Organisation System (SKOS) taxonomies for hierarchical capability classification, and Resource Description Framework (RDF) data integration. The framework is empirically validated through a case study on demonisation escalation in the Russia–Ukraine information environment and through multilingual hate-speech detection benchmarks across English, Russian, Lithuanian, and Chinese. Full technical and methodological treatment is presented in the project's scientific publication [1].

The project is funded under the European Union's Recovery and Resilience Facility (2021–2027 programming period, New Generation Lithuania plan) and co-funded by the state budget of the Republic of Lithuania, administered through the Research Council of Lithuania (Lietuvos Mokslo Taryba), under the project "Implementation of Mission-Based Science and Innovation Programs" (Project No. 02-002-P-0001). The project is led by the Lithuanian Cybercrime Centre of Excellence for Training, Research, and Education (L3CE) in partnership with Mykolas Romeris University. The Technology Readiness Level achieved through the present project phase is TRL-4 (technology validated in laboratory); progression toward operational deployment (TRL 7–9) is the responsibility of the project's commercialisation partner, Novian.

1.3 Document Structure

The deliverable is organised in nine chapters. Following this Introduction, Chapter 2 documents the review methodology, including the PRISMA-ScR-informed approach adopted, the modular Boolean search strategy, the screening and inclusion criteria, and the methodological limitations acknowledged. Chapter 3 summarises the technical state of the art in open-source intelligence (OSINT), social media intelligence (SOCMINT), and natural language processing (NLP) for hybrid threat detection, drawing on the project's scientific publication and the project team's prior comprehensive review. Chapter 4 summarises the adversarial context – principally Russian and Chinese state-sponsored information operations – that establishes the operational requirements any defensive framework must meet.

Chapters 5, 6, and 7 address the regulatory landscape in three layers. Chapter 5 surveys the scholarship on privacy, data protection, and fundamental rights in security AI, with attention to the doctrinal debate over the GDPR's treatment of automated decision-making, the maturity and limits of privacy-preserving machine learning, the civilian–security regulatory distinction created by the Law Enforcement Directive, and the recent Court of Justice of the European Union jurisprudence reshaping the practical operation of these instruments. Chapter 6 surveys the AI Act and the broader AI regulatory landscape, addressing the Act's hybrid product-safety/fundamental-rights architecture, the conceptual critique of trustworthiness as risk-acceptability, the explainability requirements and the technical–legal gap, and the AI Act's intersection with copyright and text-and-data-mining provisions. Chapter 7 surveys cybersecurity governance frameworks (NIS2, the Cyber Resilience Act, the Digital Operational Resilience Act) and the Digital Services Act, with particular attention to the integration problem connecting these instruments to the data-protection and AI regulatory regimes.

Chapter 8 consolidates the findings of the substantive chapters into five integrated observations and identifies the implications for the HiPSTer project's broader deliverable suite. Chapter 9 provides a concise Summary and Conclusion. A single consolidated bibliography appears at the end of the document, following IEEE numbered reference conventions.

1.4 DOA, SOTA Update

The Description of Action (DOA) for the HiPSTer project includes producing a literature review deliverable as part of the project's Phase 1 outputs. The present document discharges that obligation. No substantive deviation from the DOA-specified scope has occurred during the deliverable's preparation; the inclusion of fresh Web of Science searches for the regulatory and cybersecurity-governance chapters constitutes an expansion within the DOA-specified scope rather than a deviation.

The state-of-the-art (SOTA) position relevant to the deliverable has evolved substantially over the project's lifetime, principally in three respects. First, the AI Act entered into force in August 2024, and its principal provisions are entering implementation during the deliverables' preparation period, substantially changing the regulatory landscape on which the AI Act chapter (Chapter 6) reports. Second, the Cyber Resilience Act entered into force in December 2024, with first-application timelines extending into late 2026 and 2027; the chapter accordingly addresses the CRA principally through the analytical literature published in 2025–2026 rather than through operational implementation evidence, which is not yet available. Third, the CJEU has issued substantive rulings during the project lifetime – *SCHUFA* (2023), *Meta Platforms Ireland* (2024), *Dun & Bradstreet Austria* (2025), and *EDPS v SRB* (2025) – that have reshaped the practical operation of the GDPR in ways the Chapter 5 analysis incorporates. The literature review reflects the state of scholarship and regulation as of the date the search was executed; conclusions sensitive to imminent regulatory developments are flagged as such in the relevant chapters.

1.5 Context Update

The deployment context for hybrid threat detection systems within the European Union has evolved substantially during the project lifetime in three respects relevant to the literature review.

First, the geopolitical context in which the project operates has intensified. Russia's continued war of aggression against Ukraine and its associated information operations targeting EU Member States have produced a sustained operational environment in which hybrid threat detection capabilities are tested at scale. Chinese information operations across European information spaces – including through TikTok and adjacent platforms – have established a parallel adversarial trajectory that the EU's institutional response is still adapting to.

Second, the EU's institutional response to hybrid threats has coalesced around Foreign Information Manipulation and Interference (FIMI) as the primary analytical framework. The EEAS has produced annual reports documenting hundreds of FIMI cases, with Ukraine consistently identified as the most-targeted country and approximately a quarter of FIMI content focused on degrading the European Union itself through negative narratives targeting France, Germany, and Poland. The European Democracy Shield and the EU Strategy for Civil Society, adopted in November 2025, indicate the direction of institutional response during the downstream period, enabling the downstream actors to develop their operational compliance positions.

2. Methodology

This chapter documents the methodological approach adopted for the literature review supporting the HiPSTer project's final report to the Lithuanian Innovation Agency. It describes the review's scope and objectives, the search strategy, the screening and inclusion criteria, the synthesis approach, and the limitations of the methodology adopted.

2.1 Scope and Objectives

The literature review serves as the foundational layer of scholarship underpinning the HiPSTer project's broader suite of deliverables. Its primary objective is to survey the academic and policy literature relevant to the design, deployment, and governance of hybrid threat detection systems operating within the European regulatory environment. A secondary objective, explicitly requested in the project specification, is to ensure adequate coverage of privacy and cybersecurity dimensions throughout the review.

The review is positioned to complement, rather than duplicate, three sibling deliverables within the project: (i) the GDPR best-practices analysis and accompanying Data Protection Impact Assessment (DPIA/PDAV); (ii) the AI Act compliance research report; and (iii) the consolidated case-study, comparative-analysis, and content-analysis document. Where the sibling deliverables focus on operational compliance and HiPSTer-specific application, the present review surveys the underlying scholarly and regulatory landscape from which those compliance positions are derived. This division ensures that compliance conclusions reached elsewhere in the project rest on a documented evidence base, while avoiding redundant coverage across deliverables.

The review draws on and extends the literature analysis already presented in the project's scientific publication ("Semantic integration of hybrid threat intelligence: The HiPSTer Ontological Method for cross-domain correlation in influence operations," forthcoming in *Open Research Europe*) [1]. The publication's treatment of the technical state of the art (OSINT, SOCMINT, NLP) and of the adversarial context establishes the substantive content of Chapters 3 and 4 of the present report; these chapters adapt and condense that material for the deliverable, with cross-references to the publication and to the earlier comprehensive review by the project team (Paskauskas et al., 2026, in the *Journal of Intelligent Communication*) [2] preserved as appropriate. Chapters 5 through 7 extend the publication's regulatory analysis through targeted new literature searches, providing expanded coverage of privacy, data protection, AI governance, cybersecurity frameworks, and platform regulation.

2.2 Review Approach

The review adopts a PRISMA-ScR-informed approach (Preferred Reporting Items for Systematic Reviews and Meta-Analyses — extension for Scoping Reviews) [3]. A full systematic review with formal quality scoring was considered inappropriate given the heterogeneity of the source material, which spans technical research articles, legal scholarship, EU institutional reports, and policy guidance documents. Equally, a purely narrative review without documented methodology was considered insufficiently rigorous for an Innovation Agency deliverable. The PRISMA-ScR framework provides a defensible middle position, supporting transparency on search strategy, screening criteria, and inclusion decisions while accommodating the methodological pluralism the subject matter requires.

The review adopts the following PRISMA-ScR-informed reporting elements: documented search strategy with reproducible Boolean queries; explicit inclusion and exclusion criteria; a two-stage screening workflow; thematic narrative synthesis organised by chapter; and acknowledgement of methodological limitations. The review does not produce a formal PRISMA flow diagram with precise record counts at each screening stage, nor does it apply quality-scoring instruments to individual studies. These omissions are appropriate to the deliverables' proportionality and consistent with the project team's prior PRISMA-ScR-informed review work.

The review combines two complementary sources of evidence:

- *Reused material from project outputs.* Chapters 3 and 4 are constructed primarily from the literature analysis presented in the HiPSTer scientific publication, which, in turn, draws on the project team's comprehensive *review article in the Journal of Intelligent Communication*. The present report condenses this material.

- *Fresh database searches.* Chapters 5 through 7 are constructed from four new Web of Science Boolean searches executed for the deliverable, targeting privacy and data protection, AI Act compliance, cybersecurity governance frameworks, and the Digital Services Act, addressing the privacy and cybersecurity dimensions, and going beyond what the publication's regulatory section could accommodate within its space constraints.

2.3 Search Strategy

A modular Boolean search strategy was adopted for the new searches in preference to a single composite query. The modular approach offers three advantages: it allows each analytical layer of the review to be searched with terminology calibrated to its domain; it produces tractable result sets for screening; and it permits sensitivity adjustment within individual themes without disturbing the overall search structure.

Four thematic searches were defined, mapped to the substantive chapters of the present report as set out in the Table, followed by the full Boolean strings.

Table – Mapping of Boolean searches to report chapters

Search	Theme	Chapter
1	Data protection, fundamental rights × security AI	Ch. 5
2	AI Act and regulatory compliance for AI systems	Ch. 6
3	Cybersecurity frameworks: NIS2, Cyber Solidarity, CRA	Ch. 7
4	Digital Services Act × disinformation and systemic risk	Ch. 7

Chapters 3 (technical state of the art) and 4 (adversarial context) are drawn from the project's scientific publication and prior comprehensive review, and do not require fresh database searches.

Boolean search strings (Web of Science topical search, TS= field)

Search 1 — Data protection and fundamental rights × security AI

```
TS=(("GDPR" OR "General Data Protection Regulation" OR
  "data protection" OR "privacy-by-design" OR "privacy by design"
  OR "Law Enforcement Directive" OR "data minimi?ation" OR
  "purpose limitation" OR "data subject right*" OR
  "fundamental rights" OR "Charter of Fundamental Rights" OR
  "privacy-preserving" OR "privacy preserving" OR
  "pseudonymi?ation" OR "anonymi?ation")
AND
 ("artificial intelligence" OR "AI system*" OR "machine learning"
  OR "automated decision*" OR profiling OR surveillance OR
  "law enforcement" OR "national security" OR "intelligence
  agenc*" OR OSINT OR "threat detection"))
```

Search 2 — AI Act and regulatory compliance

```
TS=(("AI Act" OR "Artificial Intelligence Act" OR
  "Regulation 2024/1689" OR "high-risk AI" OR "prohibited AI"
```

OR "fundamental rights impact assessment" OR "FRIA" OR
"AI governance" OR "AI regulation" OR "trustworthy AI" OR
"responsible AI")
AND
(compliance OR "regulatory compliance" OR "risk management"
OR transparency OR explainab* OR accountab* OR "human
oversight" OR "human-in-the-loop" OR auditab* OR
"law enforcement" OR "national security" OR "hybrid threat*"
OR OSINT OR "threat detection"))

Search 3 — Cybersecurity frameworks

TS=(("NIS2" OR "NIS 2" OR "NIS Directive" OR "Directive 2022/2555"
OR "Cyber Solidarity Act" OR "cybersecurity directive" OR
"Cyber Resilience Act" OR "Digital Operational Resilience"
OR "DORA")
AND
(compliance OR ontolog* OR "knowledge graph*" OR "threat
intelligence" OR "incident response" OR "critical
infrastructure" OR "hybrid threat*" OR resilience OR
"information sharing" OR "regulatory framework*"))

Search 4 — Digital Services Act × disinformation

TS=(("Digital Services Act" OR "DSA" OR "Regulation 2022/2065" OR
"platform regulation" OR "very large online platform*" OR "VLOP*")
AND
(disinformation OR misinformation OR "systemic risk*" OR
"content moderation" OR "election* integrity" OR FIMI OR
"foreign interference" OR "hybrid threat*" OR propaganda))

The searches were executed in Web of Science Core Collection using the TS= (topical search) field, which queries title, abstract, author keywords, and Keywords Plus. Wildcards follow standard Web of Science conventions: * for truncation and ? for single-character substitution (capturing British/American spelling variants such as *minimisation/minimization*).

A date filter of 2018–2026 was applied to all searches, with the exception of foundational regulatory material from 2016 onward, where directly relevant to the General Data Protection Regulation, the Law Enforcement Directive, or the EU Joint Framework on Countering Hybrid Threats. Document types were restricted to Articles, Reviews, and Proceedings Papers. The search language was English; Lithuanian-language policy materials, where relevant, were accessed through separate targeted retrieval rather than database search.

2.4 Screening and Inclusion Criteria

A two-stage screening workflow was adopted. In the first stage, preliminary title and abstract screening was performed within Web of Science by the project lead, with surviving records exported in Research Information Systems (.RIS) format, including abstracts. In the second stage, refined screening was performed against the inclusion and exclusion criteria below, with full-text retrieval where abstracts were insufficient to support an inclusion decision.

Inclusion criteria:

- Peer-reviewed journal articles, conference proceedings papers, or review articles
- Published between 2018 and 2026 (2016 onward for foundational regulatory material)
- English language
- Substantive engagement with at least one of: hybrid threats and influence operations; OSINT, SOCMINT, or NLP for security applications; EU regulatory frameworks (GDPR, Law Enforcement Directive, AI Act, NIS2, Digital Services Act); privacy-preserving security artificial intelligence; or fundamental rights in AI systems

Grey literature inclusion:

EU institutional reports and guidance documents were included where they carry analytical weight or establish authoritative interpretation. Specifically: publications by the European Union Agency for Cybersecurity (ENISA), the European External Action Service (EEAS), the European Data Protection Board (EDPB), the European Data Protection Supervisor (EDPS), the European Union Agency for Law Enforcement Cooperation (Europol), the European Commission, and studies commissioned by the European Parliament. The jurisprudence of the Court of Justice of the European Union (CJEU) was incorporated through direct citation rather than through a database search.

Exclusion criteria:

- Editorials, letters, corrections, book reviews
- Purely technical contributions without a security or regulatory dimension
- Non-EU regulatory analyses, except where explicitly comparative with EU frameworks
- Duplicate coverage by the same authors of substantively identical material (most cited or most recent retained)
- Sources from venues exhibiting characteristics consistent with predatory publishing practices5

Flagged for case-by-case decision:

- Pre-prints from established venues (e.g., arXiv, cs.CR, cs.CL), included where substantive and where the work has subsequently been cited in peer-reviewed scholarship
- Workshop papers, including those that are highly cited or methodologically distinctive
- Records where abstracts proved insufficient for a screening decision and full-text access was constrained

2.5 Synthesis Approach

The review adopts narrative thematic synthesis rather than meta-analysis. Meta-analytic aggregation is inappropriate given the heterogeneity of the source material: the review encompasses empirical technical research, doctrinal legal scholarship, policy analysis, and institutional reporting, with no shared outcome measures across these domains.

The synthesis is organised thematically across Chapters 3 through 7, with each chapter addressing a coherent analytical layer. Chapters 3 and 4 condense the technical and adversarial context literature established in the project's scientific publication. Chapters 5 through 7 provide expanded coverage of the regulatory, privacy, and cybersecurity landscape based on the four new Boolean searches documented above. A concluding synthesis chapter (Chapter 7, final section) draws the threads together and identifies the integration gaps that motivate the HiPSTer methodological response. Cross-references to the scientific publication and to sibling project deliverables are made explicit where the analytical territory adjoins.

In-text citations follow the IEEE numbered convention, with references presented in a single consolidated bibliography at the end of the report.

2.6 Methodological Limitations

Several limitations bound the conclusions that can be drawn from the present review.

First, the regulatory landscape covered by the review is evolving rapidly. The AI Act entered into force during the project lifetime, the Cyber Resilience Act is being implemented, and CJEU jurisprudence on data protection in security contexts continues to develop. The review reflects the state of scholarship and regulation as of the search execution date, and conclusions sensitive to imminent regulatory developments are flagged as such in the relevant chapters.

Second, the English-language scope of the database search risks under-representing scholarship published in other European languages, particularly in the legal-doctrinal literature. Lithuanian-language regulatory materials are accessed directly rather than through database search, but equivalent coverage cannot be assumed for other Member State jurisdictions.

Third, the review relies primarily on Web of Science as the source database. Coverage of computer science conference proceedings is known to be incomplete in Web of Science relative to specialised databases such as the ACM Digital Library or IEEE Xplore. Where significant technical contributions were known to the project team but absent from the search results, they have been incorporated via citation chaining from included sources.

Fourth, the two-stage screening workflow involves judgment at both stages, and inter-rater agreement was not formally measured. The project's resource envelope did not permit dual-reviewer screening, which represents a known limitation of single-reviewer scoping reviews.

Fifth, the PRISMA-ScR-informed approach adopted here does not produce a formal flow diagram with quantified record counts at each screening stage. This decision was made on proportionality grounds; readers requiring fully quantified reporting are directed to the project's scientific publication, which employs a more formal review treatment of the technical literature it surveys.

Sixth, Chapters 3 and 4 reuse material from the project's scientific publication and the earlier *Journal of Intelligent Communication* review by the project team. While this reuse is methodologically transparent and ensures consistency across project outputs, it inherits the limitations of those reviews, including their respective search dates and database scopes.

Despite these limitations, the methodology adopted provides a documented, reproducible foundation for the synthesis presented in the chapters that follow.

3. State of the Art in Hybrid Threat Intelligence: OSINT, SOCMINT, and NLP

3.1 Chapter Framing

This chapter summarises the state of the art across the three analytical domains most relevant to hybrid threat detection: open-source intelligence (OSINT), social media intelligence (SOCMINT), and natural language processing (NLP). The full literature review underpinning this summary is presented in the project's scientific publication [1], which itself draws on the project team's comprehensive review article in the *Journal of Intelligent Communication* [2]. Readers seeking detailed methodological discussion of individual studies, performance metrics, and platform-by-platform analysis are directed to those sources. The present chapter restricts itself to the high-level scholarly landscape and the integration gaps that motivate the regulatory and architectural analysis developed in subsequent chapters.

3.2 OSINT

OSINT scholarship has advanced rapidly across explainable AI for botnet detection, LLM-orchestrated threat intelligence harvesting, real-time feed integration, and multimodal forensics. Commercial platforms – Babel Street,

WebIQ Voyager Suite, Palantir Gotham, IBM i2 Analyst's Notebook – and open-source ecosystems (Maltego, the Python data-science stack) are mature within their respective specialisations. Recent work has integrated generative AI and large language models into OSINT workflows, but the same technological wave has produced a foundational challenge: AI-generated synthetic media now accounts for over eighty percent of observed social engineering attacks [4], with detection capabilities lagging generation sophistication and projected fraud damages reaching forty billion US dollars by 2027 [5]. Migration of user populations from centralised platforms to fragmented ecosystems (Telegram, Discord, Mastodon) further complicates systematic monitoring [6].

The project's prior review identifies three persistent limitations: domain-specific analytical silos, limited contextual integration, and the absence of multi-modal reasoning frameworks. Recent developments reinforce rather than resolve these gaps.

3.3 SOCMINT

SOCMINT scholarship distinguishes between research explicitly addressing hybrid threats as coordinated multi-domain phenomena [7 – 9] and research applying advanced methodologies (deep learning, BERT, LSTM, graph neural networks) to traditional cybersecurity problems such as bot detection and darknet forum analysis [10 – 12]. Commercial platforms – notably Graphika for network mapping of state-linked influence operations and Botometer for bot-likelihood scoring – provide capability within bounded contexts.

Three recent developments are particularly significant. First, empirical evidence of coordinated inauthentic activity spanning multiple platforms during the 2024 U.S. Presidential Election demonstrates that coordination systematically transcends platform boundaries [13]. Second, evidence has emerged that detected coordination may not meaningfully influence broader discourse, raising questions about whether current detection priorities align with actual threat impact [14]. Third, a typology of coordinated link-sharing networks reveals that "coordinated behaviour" encompasses phenomena from benign community organising to malicious manipulation, exposing analytical imprecision in conflating coordination with inauthenticity [15].

The persistent limitations remain: conceptual confusion between hybrid threats and hybrid methodologies, analytical fragmentation across platforms, and the absence of cross-platform intelligence integration.

3.4 NLP

NLP scholarship has produced sophisticated specialised systems for entity extraction, automated reporting, multimodal fact-checking, and agentic verification [16 – 19]. Recent LLM integration into threat detection workflows has produced mixed results: GPT-4 outperforms other LLMs for propaganda detection but fails to outperform a smaller RoBERTa baseline [20], while LLMs simultaneously demonstrate disturbing capacity to *generate* convincing disinformation at scale [21, 22] – a generation–detection asymmetry that structurally disadvantages defenders.

Retrieval-Augmented Generation has improved fact-checking accuracy and explainability but remains constrained to domains with established peer-reviewed knowledge bases [19]. A recent unified CTI pipeline integrating symbolic knowledge graphs with semantic vector search outperforms state-of-the-art models for TTP extraction, but its evaluation also revealed that general-purpose tokenizers recognise only 1.62 percent of specialised MITRE ATT&CK terminology – quantifying the domain-vocabulary gap that undermines generalist NLP approaches for threat intelligence [23].

The persistent gaps identified in the project's prior review – absence of cross-domain reasoning, limited operational explainability, fragmentation of detection capabilities – remain unresolved.

3.5 Cross-Cutting Pattern

A consistent pattern emerges across the three domains. Individual analytical capabilities have achieved considerable sophistication within bounded specialisations, yet hybrid threats systematically exploit the seams between them. Five gaps recur:

- *Analytical silos* between bot detection, propaganda classification, OSINT platforms, and CTI pipelines

- *Cross-platform reasoning* weakness despite documented platform-spanning adversarial tradecraft
- *Conceptual ambiguity* around "coordination," "inauthentic behaviour," and "influence"
- *Explainability and attribution* weakness limiting evidential weight in regulatory and judicial contexts
- *Generation–detection asymmetry* in generative AI structurally favouring adversaries

These technical limitations establish the operational requirements that any defensive framework must address. The adversarial context that calibrates those requirements is examined in Chapter 4; the regulatory and governance constraints within which a framework may lawfully be deployed are taken up in Chapters 5 through 7.

4. Adversarial Context and Operational Requirements

4.1 Chapter Framing

This chapter summarises the adversarial context that establishes the operational requirements any defensive hybrid threat detection framework must meet. The full treatment is presented in the project's scientific publication [1]; the present chapter focuses on the strategic doctrines, operational tradecraft, and threat-landscape findings that calibrate the regulatory and architectural requirements developed in Chapters 5 through 7. Russian and Chinese state-sponsored information operations are treated as the primary benchmark, consistent with assessments by European security institutions that they represent the most sophisticated and persistent threats facing the Union.

4.2 Russian Information Warfare Doctrine

Russian information warfare doctrine integrates military, intelligence, and psychological operations into a unified concept of "information confrontation," in which narrative control is treated as a domain of conflict coequal with kinetic operations [24]. Operations in Ukraine since 2014 illustrate this integration: military actions coordinate with cyberattacks against critical infrastructure, targeted disinformation campaigns on social media platforms, and strategic communications through state-controlled media to shape international perceptions while degrading the adversary's effectiveness [25]. VKontakte and Telegram serve as primary platforms for both OSINT collection and influence operations, exploiting widespread adoption across post-Soviet spaces and permissive content moderation policies.

The Internet Research Agency's 2016 operations targeting U.S. elections established templates subsequently refined: coordinated inauthentic behaviour employing bot networks to amplify divisive content, authentic-appearing personas building credibility before deploying disinformation, and cross-platform coordination ensuring narrative saturation before fact-checking responses emerge [26]. The Portal Kombat network identified by France's VIGINUM service in 2024 exemplifies continued evolution: over one thousand websites across European languages promoted coordinated pro-Russian narratives through culturally tailored content exploiting specific national debates [27].

4.3 Chinese Information Operations

Chinese military information operations reflect distinct strategic priorities while increasingly adopting techniques pioneered in Russian operations. The People's Liberation Army conceptualises information warfare through the "Three Warfares" framework: psychological warfare targeting adversary decision-maker perceptions, media warfare shaping domestic and international narratives, and legal warfare establishing favourable interpretations of international norms [28]. Rather than employing bot networks extensively, Chinese operations often leverage genuine accounts – diaspora communities, nationalist volunteers, and coordinated authentic behaviour by users responding to official guidance [29]. TikTok's algorithm-driven content distribution enables influence operations that exploit platform mechanisms rather than requiring overt coordination: strategically placed content that resonates with target audiences receives organic amplification, effectively weaponising commercial recommendation systems [30]. –

4.4 EU Institutional Framing

The European Union's institutional framing identifies Foreign Information Manipulation and Interference (FIMI) as a primary hybrid threat vector, characterised by coordinated campaigns employing inauthentic news articles, fabricated investigations, AI-generated deepfakes, and cross-platform narrative manipulation designed to interfere in elections, discredit public institutions, and erode democratic resilience [31]. ENISA's 2025 Threat Landscape documents that AI-generated content now accounts for over eighty percent of observed social engineering attacks, while approximately a quarter of FIMI content focuses specifically on degrading the Union through negative narratives targeting France, Germany, and Poland [4]. The EU Joint Framework on Countering Hybrid Threats and the parallel NATO Strategic Concept identify hybrid operations as a core and persistent challenge to Alliance security, with the latter affirming that hybrid operations against NATO Allies could reach the threshold of armed attack [32, 33].

4.5 Operational Requirements for Defensive Systems

The adversarial landscape establishes specific operational demands. Defensive systems must support cross-domain semantic reasoning that correlates technical cyber indicators with coordinated narrative manipulation; distinguish genuine hybrid threats from hybrid analytical methodologies; and support attribution assessment under conditions of deliberate ambiguity. While individual analytical disciplines achieve high technical sophistication within bounded specialisations (Chapter 3), current defensive systems remain siloed, lacking the integration capabilities the adversarial profile demands. These architectural limitations – rather than deficits in individual detection capabilities – motivate the ontological integration approach developed within the HiPSTer project and synthesised in Chapter 7.

Beyond architectural integration, the adversarial profile imposes governance requirements that shape what defensive systems may lawfully do. Coordinated state-sponsored operations operate at scale across personal data, profiling, content analysis, and cross-border information flows – precisely the territory where the EU regulatory architecture imposes the most demanding constraints. The chapters that follow examine these constraints in turn: Chapter 5 addresses data protection and fundamental rights; Chapter 6 addresses the AI Act and the broader AI regulatory landscape; Chapter 7 addresses cybersecurity governance frameworks and platform regulation under the Digital Services Act.

5. Privacy, Data Protection, and Fundamental Rights in Security AI

5.1 Chapter Framing

This chapter surveys the scholarship and regulatory landscape at the intersection of personal data protection, fundamental rights, and the artificial intelligence systems that increasingly mediate security and law-enforcement functions. The scope is deliberately bounded: the chapter does not produce a GDPR compliance analysis for HiPSTer (the subject of sibling deliverable on GDPR best practices and the accompanying Data Protection Impact Assessment), nor does it duplicate the regulatory mapping presented in the project's scientific publication. Instead, it summarises the academic and institutional literature that establishes the conceptual foundations on which compliance positions rest.

Three substantive threads are developed. The first traces the doctrinal debate over the General Data Protection Regulation's treatment of automated decision-making and the contested status of a "right to explanation." The second surveys the privacy-preserving machine learning literature, which has matured rapidly under the GDPR's influence while exhibiting structural limitations of its own. The third addresses the civilian–security regulatory distinction created by the Law Enforcement Directive and the fragmentation it produces for systems operating across that boundary. The chapter closes by noting the Court of Justice of the European Union (CJEU) jurisprudence that has clarified – and, in places, reshaped – the practical operation of these instruments.

5.2 The GDPR and Automated Decision-Making: A Doctrinal Debate

The most extensively litigated scholarly question regarding the General Data Protection Regulation's relevance to artificial intelligence concerns the status of Article 22, which restricts decisions "based solely on automated processing,"

and the related transparency obligations of Articles 13–15. Whether these provisions establish a substantive "right to explanation" of algorithmic decisions has been the subject of sustained doctrinal dispute.

Goodman and Flaxman argued early that the GDPR effectively creates such a right, with significant consequences for the routine deployment of machine-learning algorithms across the European Union [34]. Their account treats the transparency obligations as actionable: a data subject affected by an algorithmic decision can require an account of the underlying logic, and machine-learning systems must be designed to provide such an account. Wachter, Mittelstadt, and Floridi mount a sustained challenge to this reading, arguing that the GDPR's text does not establish a legally enforceable right to explanation of specific automated decisions; what the Regulation provides is a more limited right to be informed about the existence and general logic of automated processing, falling short of *ex post* individual explanation [35]. The Wachter–Mittelstadt–Floridi position has proven influential in subsequent doctrinal scholarship and has been substantially borne out by CJEU jurisprudence discussed below.

Selbst and Barocas complicate both positions by distinguishing between *inscrutability* (the model is opaque even to its developers) and *nonintuitiveness* (the model's rules, even if disclosed, defy human understanding of why they are what they are) [36]. Existing legal instruments – including the GDPR, the U.S. Fair Credit Reporting Act, and the Equal Credit Opportunity Act – address inscrutability with reasonable adequacy but largely fail to engage nonintuitiveness. The implication for hybrid threat detection systems is significant: even where a HiPSTer-like architecture exposes its reasoning chain through SPARQL query traces (an inscrutability response), the question of whether the underlying indicator weightings and threshold settings are normatively defensible (a nonintuitiveness question) remains underdetermined by current law and requires the kind of expert human review that the project has embedded through its human-in-the-loop design.

The doctrinal debate matters operationally because it determines the evidential weight that automated detection outputs can carry. A system that produces only a probabilistic classification ("this account exhibits coordination patterns at confidence 0.82") without a traceable reasoning chain confronts the inscrutability problem directly; a system that produces a traceable chain may satisfy inscrutability requirements but remain exposed on the nonintuitiveness front if the underlying ontological assumptions cannot be normatively defended.

5.3 Privacy-Preserving Machine Learning: Maturity and Limits

Parallel to the doctrinal debate, a substantial technical literature has developed on privacy-preserving machine learning architectures, motivated explicitly by the GDPR's strengthened consent, data minimisation, and purpose-limitation requirements. Yang and colleagues' foundational survey establishes federated learning (FL) as the leading framework for privacy-preserving collaborative model training [37]. Federated learning enables model training across distributed datasets without transferring the underlying data to a central server, addressing the data-minimisation imperative while preserving model quality.

The technical landscape has since matured considerably. Li and colleagues survey the systems infrastructure that has developed around federated learning, categorising existing implementations by data distribution, machine learning model, privacy mechanism, communication architecture, federation scale, and federation motivation [38]. Kaissis and colleagues' Nature Machine Intelligence review of secure, privacy-preserving, and federated machine learning – though presented through the lens of medical imaging applications – provides the most authoritative overview of the broader technical landscape, covering federated learning, differential privacy, secure multi-party computation, and homomorphic encryption [39]. The combination of these techniques can substantially mitigate privacy risk in collaborative learning scenarios while preserving model utility.

However, the assumption that privacy-preserving machine learning techniques *automatically* provide privacy protection has been substantially complicated by research on adversarial attacks. Lyu and colleagues' comprehensive survey of attacks and defences in federated learning documents a range of inference attacks (membership inference, attribute inference, gradient leakage) and poisoning attacks (Byzantine attacks, backdoor insertion) that undermine the privacy and integrity guarantees federated systems are commonly assumed to provide [40]. The implication is methodologically important: privacy-preserving machine learning is not a *solution* to GDPR compliance but a *toolkit* whose deployment requires careful threat-modelling and ongoing adversarial assessment.

This literature is directly relevant to hybrid threat detection systems. To the extent that defensive frameworks rely on multilingual classification, cross-border indicator sharing, or collaborative training across institutional boundaries – all of which are contemplated within HiPSTer's deployment trajectory – the privacy-preserving machine learning toolkit

offers genuine architectural options. But the security guarantees of these options are conditional, not categorical, and must be evaluated against the specific threat model posed by hybrid threat detection (including, notably, the possibility that adversarial actors will attempt to poison training data or extract sensitive operational details from deployed models).

5.4 The Civilian–Security Distinction and the Law Enforcement Directive

Beyond the GDPR's general framework, the regulatory landscape for security-oriented AI systems is decisively shaped by the Law Enforcement Directive (41), which governs personal data processing by competent authorities for the prevention, investigation, detection, or prosecution of criminal offences and the safeguarding against threats to public security. The Misijos publication [1], section "Regulatory perspective", sets out the principal architecture of this distinction; the present chapter notes the scholarly debate over its practical consequences.

The LED mirrors many GDPR principles – including privacy-by-design (Article 20) and proportionality requirements – but creates operational latitude for law enforcement processing of special categories of personal data (Article 10) that the GDPR's parallel regime (Article 9) does not afford. The practical implication for hybrid threat detection systems is that the regulatory framework available to a system depends substantially on the institutional identity of its deployer. A system developed by a civilian research consortium and deployed by an intelligence service operates under markedly different constraints than the same system deployed by a private platform operator.

This jurisdictional architecture has been substantially critiqued in recent scholarship. Tzanou and Vogiatzoglou [42] demonstrate through case studies of Palantir data-mining software and Pegasus spyware surveillance that the prevailing "controller-focused approach" – which grounds EU law's applicability in the activities of private entities rather than in the protection of data subjects – produces fragmented outcomes that emerging surveillance technologies readily exploit. They propose a "data-subject-centric model" that would yield greater legal certainty for systems operating across civilian and security contexts. The critique is directly relevant to HiPSTer-like systems, which may be developed by civilian research institutions but deployed by national security or law-enforcement authorities, potentially falling between regulatory frameworks in ways that diminish transparency and accountability for affected data subjects.

The civilian–security distinction also produces operational consequences for the institutional architecture of compliance. Where a GDPR-only deployer must satisfy the full Article 22 regime on automated decision-making, an LED-regulated deployer enjoys greater latitude – but is subject to compensating oversight requirements through national supervisory authorities and, where applicable, judicial authorisation. Lithuania's Law on Legal Protection of Personal Data Processed for Law Enforcement Purposes [43] illustrates how national transposition can extend LED-style provisions to national security and defence contexts, thereby providing operational flexibility while preserving fundamental rights guarantees.

5.5 CJEU Jurisprudence: Reshaping Practical Operation

The doctrinal landscape has been substantially shaped by recent rulings of the Court of Justice of the European Union, which have clarified contested points in ways directly relevant to security AI systems. The publication's regulatory section [1] catalogues the principal cases; the present chapter notes their cumulative effect.

In *SCHUFA* (Case C-634/21, 2023) [44], the Court ruled that automated credit scoring constitutes an "automated decision" within the meaning of Article 22 GDPR where the score significantly affects the data subject, even when the formal decision is taken by a downstream human actor relying on the score. The ruling substantially extends the practical reach of Article 22's transparency and human-oversight requirements. *Dun & Bradstreet Austria* (Case C-203/22, 2025) [45] subsequently clarified that data subjects may require controllers to explain the procedures and principles *actually applied* in automated processing, rather than merely the system's abstract logic.

In *Meta Platforms Ireland* (Case C-252/21, 2024) [46], the Court reinforced the data-minimisation principle, ruling that the extensive processing of user data to monitor large portions of online activity constitutes "a serious interference with fundamental rights" under Articles 7 and 8 of the Charter of Fundamental Rights. The ruling reinforces the necessity of purpose-specific data collection and defined retention periods, constraints that any deployable hybrid threat detection system must build into its architecture from inception.

The *EDPS v SRB* ruling (47) clarified that pseudonymised data is not automatically personal data for a recipient who cannot reasonably re-identify the underlying individuals, offering a more nuanced, context-specific view of data use. The ruling has direct relevance to cross-institutional indicator sharing within hybrid threat detection frameworks, where indicator data can be pseudonymised in ways that genuinely preclude re-identification by the receiving entity, thereby substantially reducing its regulatory burden.

5.6 Synthesis

Three observations emerge from this survey. First, the conceptual landscape for privacy and data protection in security AI is substantially more contested than the surface regulatory architecture suggests. The doctrinal debate over the GDPR's automated-decision provisions, the structural critique of the civilian–security distinction, and the unsettled status of nonintuitiveness as a normative problem all constitute live scholarly territory rather than settled questions. Second, the technical toolkit for privacy-preserving machine learning has matured to the point where genuine architectural options exist for systems that require cross-institutional collaboration, but these options are contingent on careful threat modelling rather than on categorical solutions to compliance problems. Third, CJEU jurisprudence is reshaping the practical operation of the GDPR in ways that strengthen data-subject rights – particularly around transparency of automated decision-making and the data-minimisation principle – with direct consequences for the design of broad-spectrum collection systems.

For the HiPSTer project specifically, these observations reinforce the architectural choices already made: the ontology-driven explainability infrastructure addresses inscrutability directly; the human-in-the-loop reasoning design accommodates the nonintuitiveness concern; and the modular architecture supports the institutional variability that the civilian–security distinction creates. The detailed application of these conclusions to HiPSTer's specific compliance position is developed in the project's GDPR best-practices analysis and Data Protection Impact Assessment [48].

Chapter 6 turns to the artificial intelligence regulatory framework proper, examining the AI Act and the broader landscape of AI-specific regulation that supplements (and in important respects extends) the data-protection regime surveyed here.

6. The AI Act and the Regulatory Landscape for AI Systems

6.1 Chapter Framing

This chapter surveys the scholarship and regulatory landscape governing artificial intelligence systems within the European Union, with primary focus on the Artificial Intelligence Act (Regulation 2024/1689) and its interaction with the broader AI governance ecosystem. As with Chapter 5, the scope is bounded by the project's deliverable architecture: the chapter does not produce an AI Act compliance analysis for HiPSTer (the subject of sibling deliverable on AI regulatory compliance [48]), nor does it duplicate the regulatory mapping presented in the project's scientific publication [1]. Instead, it summarises the academic and institutional literature that establishes the conceptual foundations on which compliance positions rest.

Four substantive threads are developed. The first addresses the AI Act's hybrid regulatory architecture, which combines product-safety and fundamental-rights traditions in ways that produce both opportunities and tensions. The second surveys critical analyses of the AI Act's conceptual foundations, particularly the contested conflation of "trustworthiness" with "risk acceptability." The third examines the explainability requirements imposed by the AI Act and the technical–legal gap between what the law requires and what current *explainable* AI (XAI) techniques can deliver, with attention to the architectural alternatives offered by symbolic and neurosymbolic approaches. The fourth addresses the AI Act's intersection with copyright law and text-and-data-mining provisions, directly relevant to the data-acquisition layer of OSINT-based systems. The chapter closes by noting the unresolved national-security loophole and the institutional guidance that has emerged to operationalise the AI Act's provisions.

6.2 The Hybrid Regulatory Architecture of the AI Act

The Artificial Intelligence Act establishes a regulatory framework that combines two distinct EU legal traditions: product safety regulation and fundamental rights protection. Almada and Petit provide the most sustained scholarly analysis of this hybrid framing, arguing that while the combination is conceptually attractive, it generates practical and theoretical challenges that will need to be addressed during implementation and in future EU legislation [49]. Their analysis draws on three classical themes of law-and-technology scholarship: the pacing problem (regulation moving more slowly than technology), the regulatory lens (the framing assumptions that shape what regulators see), and institutional path dependence (the durability of existing regulatory architectures even when their fit to new problems is imperfect). The combination of product-safety and fundamental-rights traditions risks failing if it does not account for the structural differences between these legal lineages, including their different conceptions of harm, redress, and enforcement.

The AI Act's risk-based classification – distinguishing unacceptable risk (prohibited), high risk, limited risk, and minimal risk – operationalises this hybrid architecture by applying different obligation sets to different system categories. The publication's regulatory section [1], "Regulatory perspective," sets out the principal classifications and notes that HiPSTer-like systems deployed in law-enforcement or public-security contexts will likely qualify as high-risk under Annex III. The present chapter restricts itself to noting the scholarly observation that this classification regime inherits the product-safety tradition's emphasis on *ex ante* conformity assessment and CE marking, while bolting on fundamental-rights obligations (notably the Fundamental Rights Impact Assessment under Article 27) that derive from a quite different regulatory logic. Whether these can be coherently integrated in practice remains an open empirical question that the AI Act's implementation period will substantially answer.

For systems operating in security and intelligence contexts, the most significant feature of the AI Act's architecture is the exemption in Article 2(3), which excludes systems placed on the market, put into service, or used exclusively for military, defence, or national security purposes. The boundary of this exemption – particularly for systems developed by civilian research consortia but subsequently deployed by national security authorities – has been substantially critiqued in the scholarship discussed below.

6.3 Conceptual Critique: Trustworthiness as Risk-Acceptability

A second strand of recent scholarship subjects the AI Act's conceptual foundations to critical analysis. Laux, Wachter, and Mittelstadt argue that the AI Act adopts a substantially simplified conceptualisation of trust by treating "trustworthiness" as effectively reducible to "acceptability of risk" [50]. Drawing on a narrative systematic literature review on institutional trust and AI in the public sector, they conclude that the EU is overselling its regulatory ambition: there remains a substantial threat of misalignment between citizens' actual levels of trust and the trustworthiness of applied AI systems, regardless of the AI Act's formal compliance mechanisms. The critique is significant because the AI Act's entire architectural logic assumes that trustworthiness can be engineered through risk classification and the imposition of obligations; if that assumption fails empirically, the regulatory framework's claim to public legitimacy weakens.

Díaz-Rodríguez and colleagues offer a more constructive treatment of the same territory, proposing a holistic vision of trustworthy AI grounded in four essential axes: global principles for ethical use and development; a philosophical engagement with AI ethics; a risk-based regulatory approach; and the seven technical requirements of the European Commission's Trustworthy AI framework (human agency and oversight; robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental wellbeing; and accountability) [51]. Their analysis emphasises that responsibility – operationalised through auditing processes and, where appropriate, regulatory sandboxes – is the bridge between abstract principles and deployed systems. The implication for hybrid threat detection systems is that compliance with the AI Act's formal requirements is necessary but not sufficient for trustworthiness; auditing and oversight infrastructure must be embedded throughout the system lifecycle.

The publication's regulatory section [1] discusses Mantelero's framework for Fundamental Rights Impact Assessment under AI Act Article 27 [52], which provides the most developed methodological treatment of how the rights-protection dimension of the AI Act can be operationalised. Mantelero's central observation – that FRIA "cannot be fully automatised" because the relevant parameters are "inevitably context-based and grounded on expert knowledge of fundamental rights" – aligns directly with the human-in-the-loop design philosophy embedded in

HiPSTer's architecture. Reliance on expert human judgement is not a failure of automation; it is a structural feature of any AI governance regime that takes fundamental-rights protection seriously.

6.4 Explainability Requirements and the Technical–Legal Gap

The AI Act's transparency obligations – alongside the GDPR's automated-decision provisions surveyed in Chapter 5 – create substantial demand for explainable AI techniques. The explainable AI (XAI) literature has matured rapidly in response, but a persistent gap remains between what the law requires and what current techniques can deliver.

Longo and colleagues' "XAI 2.0 manifesto" – co-authored by a group spanning XAI researchers, legal scholars, and applied AI practitioners – identifies twenty-eight open challenges across nine categories, charting the research agenda for advancing explainable AI in ways that engage with its practical and ethical context [53]. The manifesto's significance for the present chapter lies in its explicit recognition that explainability is not a purely technical property: what counts as an adequate explanation depends on the audience, the deployment context, and the legal-regulatory framework within which the system operates. The same model output may satisfy explainability requirements for a technical operator while failing them for an affected data subject or a judicial reviewer.

The dominant technical paradigm in current XAI research is *post-hoc* explainability: building a model first, then applying interpretation techniques (saliency maps, attention visualisation, feature attribution, surrogate models) to make its behaviour intelligible. This approach has produced substantial methodological progress, but its limitations are increasingly recognised in the regulatory literature. Post-hoc explanations are approximations of model behaviour rather than direct accounts of model reasoning; they can themselves be unreliable, manipulable, or misleading; and they do not necessarily satisfy the AI Act's requirement that the *reasoning actually applied* by the system be communicable to operators and affected parties.

An alternative architectural path is offered by intrinsically interpretable approaches, including symbolic systems and neurosymbolic architectures. Garcez and Lamb characterise the present moment as the "3rd wave" of AI, in which neural-network learning is integrated with symbolic knowledge representation and logical reasoning to deliver trust, safety, interpretability, and accountability at the architectural level [54]. Their analysis emphasises that the integration of symbolic and sub-symbolic approaches is motivated not only by practical regulatory pressures but by deep technical reasons: robust learning requires well-founded knowledge representation; explainability requires structured reasoning; commonsense and causal explanation require symbolic abstraction. The neurosymbolic research programme thus converges with the regulatory demands articulated by the AI Act, even though the two literatures developed largely independently.

For the purposes of the present chapter, the neurosymbolic literature is significant because it establishes a broader architectural trajectory within which HiPSTer's design choices can be situated. HiPSTer is not a neurosymbolic system in the strict sense – it does not currently integrate neural-network learning into its ontological reasoning layer – but it is a symbolic-first design that exhibits the intrinsic-explainability properties the neurosymbolic programme advocates. SPARQL queries traversing the OWL/SKOS knowledge graph produce reasoning chains that are inherently inspectable; an analyst can trace exactly which indicators triggered an assessment by following the query path. This contrasts sharply with the post-hoc explainability paradigm, where reasoning chains must be reconstructed approximately after the fact. The architectural alignment between intrinsic-explainability approaches and the AI Act's transparency demands suggests that the regulatory pressures of the next several years are likely to accelerate adoption of symbolic-first and neurosymbolic designs across security AI applications.

6.5 The AI Act, Copyright, and the Data Acquisition Layer

A fourth strand of recent scholarship addresses the AI Act's interaction with EU copyright law, with direct relevance to systems that acquire training and operational data through web scraping and OSINT collection. Quintais [55] provides the most comprehensive analysis to date, examining the AI Act's text-and-data-mining (TDM) provisions, its copyright-compliance obligations for general-purpose AI providers, and its interaction with the Copyright in the Digital Single Market Directive (CDSMD). The principal findings are threefold.

First, the AI Act reinforces the existing TDM framework by requiring general-purpose AI providers to implement copyright-compliance policies, to respect TDM opt-outs under Article 4(3) CDSMD, and to publicly disclose summaries of the data used to train their models. Second, the framework's extraterritorial implications are substantial: providers

placing models on the EU market must comply with EU copyright rules irrespective of where training occurred. Third, the current regime exhibits significant inadequacies – particularly regarding fair remuneration of creators – that subsequent legislative reform may need to address.

Since HIPSTer-like systems engaging in OSINT collection [1] through web scraping must consider intellectual property frameworks alongside data protection requirements, consider that OECD guidance suggests that scraping activities can implicate copyright, database rights, trademarks, trade secrets, publicity rights, and moral rights [56]. The present chapter notes that the AI Act's TDM provisions add a further regulatory layer specific to AI training data, which may not apply directly to HiPSTer's operational data collection.

6.6 Institutional Guidance and the National-Security Loophole

The AI Act's implementation has been substantially shaped by guidance from European institutions, even before the Act's principal provisions enter into force. The publication's regulatory section [1] catalogues the principal instruments, including: the European Commission's Guidelines on Prohibited AI Practices [57]; the General-Purpose AI Code of Practice [58]; EDPB and EDPS guidance on secure AI systems with personal data [59]; and Europol's practical guide on AI bias in law-enforcement contexts [60]. These instruments share an emphasis on human oversight, transparency, and bias-mitigation requirements that translate AI Act obligations into operational form.

The most contested feature of the AI Act's architecture for security AI systems is the national-security exemption in Article 2(3). Erdogan identifies what she terms a "national security loophole" at the intersection of the Law Enforcement Directive and the AI Act, demonstrating through case studies of Clearview AI, the Dutch SyRI system, and BriefCam that transparency obligations fragment across regulatory instruments in ways that emerging surveillance technologies readily exploit [61]. The AI Act addresses transparency for providers and deployers; the LED governs data-subject rights; yet neither framework adequately addresses scenarios in which AI systems operate across both civilian and security domains. The critique parallels the Tzanou and Vogiatzoglou analysis of the GDPR/LED civilian–security distinction surveyed in Chapter 5: the regulatory architecture relies on jurisdictional categories that the most consequential systems systematically transcend.

For HiPSTer specifically, the national-security loophole has direct operational relevance. A system developed by a civilian research consortium (the project's current institutional configuration) is governed by the full AI Act regime. The same system deployed by an intelligence service would qualify for the Article 2(3) exemption, with a consequent reduction in formal compliance obligations, but would be compensated by the exposure to the legitimacy concerns Erdogan identifies. The project's commercialisation pathway through Novian [62] will substantially determine which side of this regulatory boundary the deployed system ultimately occupies.

6.7 Synthesis

Four observations emerge from this survey. First, the AI Act's hybrid regulatory architecture – combining product safety and fundamental rights – is conceptually ambitious but exhibits structural tensions that the implementation period will test to a substantial extent. Second, the conceptual foundations on which the AI Act rests, particularly the equation of trustworthiness with risk acceptability, are subject to substantial critique that the official EU narrative does not fully acknowledge. Third, the AI Act's transparency requirements impose demands that current post hoc XAI techniques cannot fully meet, creating regulatory pressure toward intrinsically interpretable architectures, as exemplified by HiPSTer's symbolic-first design. Fourth, the AI Act's interaction with adjacent regulatory instruments – copyright law, the LED, and national-security frameworks – produces fragmentation that the most consequential security AI deployments systematically exploit.

For the HiPSTer project specifically, these observations reinforce architectural choices already made (intrinsic explainability through ontological reasoning; human-in-the-loop design accommodating the FRIA requirement) while highlighting the regulatory ambiguity that the project's commercialisation pathway must navigate. The detailed application of these conclusions to HiPSTer's specific compliance position is developed in the project's AI Act compliance research report [48].

Chapter 7 turns to the cybersecurity governance frameworks that complement the data-protection and AI regulatory regimes surveyed here, and to the Digital Services Act's treatment of platform-mediated disinformation as a systemic risk.

7. Cybersecurity Governance Frameworks and Platform Regulation

7.1 Chapter Framing

This chapter surveys the cybersecurity governance frameworks that complement the data-protection and AI regulatory regimes examined in Chapters 5 and 6, as well as the Digital Services Act's treatment of platform-mediated disinformation as a systemic risk. The two strands are addressed jointly because they share a common structural feature: each imposes obligations on actors operating at scale across critical or platform infrastructure, each interacts with the broader EU regulatory architecture in ways that produce both convergence and fragmentation, and each is directly relevant to the institutional context in which hybrid threat detection systems will be deployed.

The chapter is organised in five substantive sections. The first addresses the Network and Information Security Directive 2 (NIS2) and its implementation challenges. The second examines the Cyber Resilience Act (CRA) and its hybrid product-safety/fundamental-rights architecture. The third briefly treats the Digital Operational Resilience Act (DORA), noting its sectoral scope but its limited direct relevance to hybrid threat detection systems outside the financial sector. The fourth addresses the integration problem – how the cybersecurity governance instruments interact with one another and with the data-protection and AI regulatory regimes – and identifies cross-regulatory integration as the central methodological challenge of the contemporary EU compliance landscape. The fifth addresses the Digital Services Act and its framing of disinformation as a systemic risk requiring platform-level mitigation. The chapter closes with a synthesis that consolidates and identifies the integration gaps that motivate the HiPSTer methodological response.

As in Chapters 5 and 6, the scope is bounded: compliance-prescriptive material is reserved for the project's sibling deliverables, and this chapter limits itself to the scholarly and institutional literature that establishes the conceptual foundations on which those compliance positions rest.

7.2 The NIS2 Directive: Ontological Approaches and Implementation Realities

The Network and Information Security Directive 2 (Directive (EU) 2022/2555) substantially expands the scope of EU cybersecurity regulation, introducing new obligations for both essential and important entities across an enlarged set of sectors. The Directive's implementation has rapidly become a major focus of cybersecurity-governance scholarship, with two strands of analysis particularly relevant to the present chapter.

The first strand addresses ontological approaches to NIS2 compliance. Castiglione and colleagues present NIS2Onto, an OWL ontology designed to translate NIS2 into a formal semantic representation that supports automated compliance verification, streamlined risk assessment, and effective policy implementation [63]. The contribution is directly relevant to the present project because it demonstrates the viability of ontology-driven compliance translation in the cybersecurity-governance domain, which has historically resisted such formalisation. The authors argue that the complexity and cost of NIS2 compliance derive substantially from the difficulty of translating juridical language into operational procedures – a difficulty that semantic representation of NIS2 entities, relationships, and security measures can substantially mitigate. The NIS2Onto ontology is evaluated through both quantitative and qualitative analyses using a real-world case study, and the authors emphasise its extensibility for integration with other regulatory frameworks. The architectural philosophy aligns directly with the HiPSTer project's own ontology-driven approach to regulatory mapping, providing external validation for the design choice from an adjacent domain.

The second strand addresses the empirical realities of NIS2 implementation. Mentzelou and colleagues [64] apply the Decision-Making Trial and Evaluation Laboratory (DEMATEL) method to model the interdependencies among fourteen barriers to NIS2 compliance, identifying lack of awareness and the evolving threat landscape as the principal causal drivers, with technical complexity and financial constraints functioning as mediators that transmit the influence of these drivers toward operational and governance failures. The most dependent outcomes – operational disruptions, high reporting costs, and inadequate risk assessment – absorb the impact of the upstream factors. The analysis suggests that interventions targeting awareness-building, resource allocation, and risk-management capacity offer the greatest leverage for improving compliance and resilience.

Rehnstam, Winquist, and Hacks complement this analytical work with an empirical study of Swedish critical-infrastructure readiness for NIS2 [65]. Their findings – that competence levels are generally low, that organisational size and resource availability substantially shape readiness, and that significant gaps remain even in larger organisations

with dedicated security leadership – are directly relevant to the Lithuanian context in which HiPSTer is being developed. The authors emphasise that NIS2 should be approached not as a regulatory burden but as an opportunity to enhance resilience, a framing consistent with the HiPSTer project's positioning.

Taken together, these three contributions establish two complementary findings: that ontology-driven approaches to NIS2 compliance are technically viable, and that the implementation gap they are designed to address is empirically substantial. The implication for hybrid threat detection systems is that semantic representation of compliance obligations is not merely an analytical convenience but a potentially significant contribution to closing the gap between regulatory ambition and operational reality.

7.3 The Cyber Resilience Act: Hybrid Architecture and Fundamental-Rights Framing

The Cyber Resilience Act (Regulation (EU) 2024/2847, CRA) entered into force in December 2024 and establishes horizontal cybersecurity requirements for products with digital elements. Three recent contributions illuminate the regulatory architecture that the CRA enacts and its position within the broader EU cybersecurity ecosystem.

Chiara provides the most sustained doctrinal analysis, examining whether the CRA – despite its product-safety framing – operates substantively as a fundamental-rights instrument [66]. The analysis identifies three foundational regulatory choices underpinning the Act: a horizontal approach, a risk-based approach, and a product-safety approach. Chiara then investigates the extent to which the CRA enhances the protection of fundamental rights, as claimed in the Commission's Explanatory Memorandum, and finds that the relationship is more nuanced than the Memorandum suggests. The analysis parallels the Almada and Petit treatment of the AI Act surveyed in Chapter 6: both Regulations enact hybrid architectures that combine product-safety logics with fundamental rights claims, and both produce structural tensions that the implementation period will substantially test.

Ajmera offers a broader landscape view, examining the CRA's role in harmonising a historically fragmented EU cybersecurity regulatory regime [67]. The Act marks the initial integration of cybersecurity into the EU's New Legislative Framework for product regulation, establishing minimum requirements that digital products must meet before placement on the Union market, with conformity demonstrated through multiple avenues, including international and industrial standards adopted by European Standardisation Organisations. The analysis emphasises that the CRA must be understood not in isolation but as part of an overarching framework of multiple harmonising pieces of legislation geared toward enhancing EU cybersecurity – a framing that places the integration problem (addressed in 7.5) at the centre of contemporary regulatory analysis.

Canavese and colleagues provide an architectural perspective, deriving from a systematic analysis of CRA technical and organisational requirements, an architecture that supports compliance-by-design from the early stages of product development [68]. The contribution is significant for the present chapter because it demonstrates that compliance-by-design – long established as a principle in the data-protection literature (privacy-by-design under GDPR Article 25) – is now being extended to the cybersecurity domain. The implication for hybrid threat detection systems is that architectural choices made at the design stage will substantially determine the feasibility of compliance with the CRA's downstream obligations, reinforcing the project's existing emphasis on architectural alignment with regulatory requirements.

For HiPSTer specifically, the CRA's direct applicability depends on the deployment configuration: a system delivered as a product with digital elements to the Union market falls within scope; a system deployed as an internal capability within an intelligence service may fall outside it. The project's commercialisation pathway through Novian will substantially determine this classification.

7.4 The Digital Operational Resilience Act: Sectoral Scope, Limited Direct Relevance

The Digital Operational Resilience Act (Regulation (EU) 2022/2554, DORA) establishes operational resilience requirements for the EU financial sector, with extension of regulatory supervision to critical technology providers. The DORA literature is substantial – the search results returned multiple contributions on threat-led penetration testing methodologies, third-party risk management, and compliance frameworks for financial entities [69 – 71]. However, the direct relevance to hybrid threat detection systems outside the financial sector is limited.

The conceptual contribution that transfers is DORA's threat-led approach to resilience testing: rather than relying on traditional vulnerability assessments or generic penetration testing, DORA-compliant testing reproduces the tactics, techniques, and procedures of real adversaries to evaluate the effectiveness of cybersecurity controls in production environments [69]. This adversary-centric framing parallels the approach taken in HiPSTer's experimental case study [1], which validates the framework against documented Russian and Chinese tradecraft rather than against generic threat indicators. The methodological convergence is notable: both DORA's resilience testing and HiPSTer's case-study validation treat the adversary's actual operational behaviour as the appropriate benchmark for system evaluation.

Beyond this conceptual point, DORA's direct applicability to hybrid threat detection systems is bounded by its financial-sector scope. The Regulation is noted here for completeness and for the methodological parallel; substantive engagement is deferred to deployments where the financial-sector classification applies.

7.5 The Integration Problem: Cross-Regulation Compliance as the Central Methodological Challenge

The most significant finding across the cybersecurity-governance literature is that no contemporary EU regulatory instrument can be analysed in isolation. The GDPR, the LED, the AI Act, the NIS2 Directive, the CRA, DORA, the Data Governance Act, the Data Act, and the Digital Services Act increasingly overlap in their application to single technological deployments, producing both convergence pressures and persistent fragmentation. Four recent contributions address this integration problem directly and collectively constitute the most important strand of analysis for the present chapter.

Jorgensen and Ma provide the most comprehensive treatment, presenting a PRISMA-ScR scoping review of how six principal EU digital laws – GDPR, Data Governance Act, Data Act, AI Act, NIS2 Directive, and CRA – jointly govern Digital Twin systems [72]. The contribution operationalises a Unified Digital Twin Compliance Framework (UDTCF) that consolidates overlapping obligations across data governance, privacy, cybersecurity, transparency, interoperability, and ethical responsibility, together with a Digital Twin Compliance Evaluation Matrix (DTCEM) that enables qualitative assessment of compliance maturity. Applied to representative European cases across Smart Cities, Industrial Manufacturing, Transportation, and Energy Systems, the framework reveals strong convergence in data governance, security, and interoperability, but persistent gaps in transparency, explainability, and accountability for AI-driven components. The methodological parallel to the present literature review is direct: Jorgensen and Ma also adopt PRISMA-ScR, providing external methodological validation for the approach documented in Chapter 2.

The substantive contribution of the Jorgensen and Ma analysis is significant: they argue that European digital legislation now forms a coherent yet fragmented ecosystem that increasingly requires integration through compliance-by-design methodologies. The implication is that compliance with any single instrument is increasingly insufficient; what is required is cross-regulation integration treated as a first-class design problem. The conclusion that Digital Twins can act not only as regulated technologies but also as compliance infrastructures themselves, embedding legal, ethical, and technical safeguards, translates cleanly to hybrid threat detection systems, where the ontological framework can itself serve as a compliance infrastructure rather than merely as a system regulated by external compliance instruments.

Beltrán offers a complementary multi-layered analysis, treating the GDPR, DSA, AI Act, and CRA as four pillars of algorithmic security and privacy in the European Union [73]. The proposed multi-layered approach distinguishes organisational risk, risks to rights and freedoms, systemic risks, and risks to national security, mapping these risk categories to the four regulatory instruments through a risk-based approach. The contribution illustrates how the EU can establish a global standard for trustworthy innovation and the protection of fundamental rights by leveraging direct and indirect synergies among these laws. For HiPSTer, the framework provides a structured way to position the system's regulatory exposure across all four pillars simultaneously.

Komnios extends the integration analysis into a specific sectoral context – smart metering in EU energy law – and develops what he terms an "inter-regulatory proportionality" framework to reconcile the GDPR, the Data Act, the NIS2 Directive, and the Cyber Resilience Act with sectoral electricity law [74]. The analysis identifies a structural tension between enhanced grid observability and the safeguards required by the Charter of Fundamental Rights and the GDPR, intensified by the horizontal regimes introduced by the Data Act and NIS2, which may fragment the regulatory landscape and weaken sector-specific constraints. The proposed framework – in which data protection sets limits, sectoral law defines functions, and cybersecurity guarantees resilience within a coherent governance hierarchy – translates directly to security AI systems, where analogous structural tensions arise between operational requirements and the protection of fundamental rights.

Rampásek and Mesarcík address a related integration problem from the perspective of cybersecurity research governance: how the DSA, CRA, and NIS2 jointly position cybersecurity researchers within the evolving EU regulatory landscape [75]. Their analysis is directly relevant because HiPSTer is itself a cybersecurity research project operating across the boundaries these instruments establish. The authors argue that while the DSA provides novel tools such as vetted-researcher access and auditing obligations for Very Large Online Platforms (VLOPs), its structure is better suited for systemic-risk research than for adversarial, exploratory cybersecurity testing. A sustainable model for cybersecurity research governance must therefore go beyond DSA-style vetting, incorporating more flexible mechanisms such as coordinated vulnerability disclosure and bug bounty programmes, as reflected more directly in the CRA. The legal ambiguity surrounding researcher status – particularly for researchers operating outside formal institutions – has direct implications for cross-border research collaboration, including the kinds of collaboration that HiPSTer's deployment trajectory contemplates.

Taken together, these four contributions establish a clear conclusion: cross-regulation integration is the central methodological challenge of the contemporary EU compliance landscape, and the instruments that address it most effectively are those that treat integration as a first-class design problem rather than a secondary consideration. The HiPSTer ontological framework – designed from inception to support integration across regulatory, behavioural, and analytical layers – sits squarely within this trajectory.

7.6 The Digital Services Act and Platform-Mediated Disinformation

The Digital Services Act (Regulation (EU) 2022/2065, DSA) establishes obligations for online intermediaries operating within the Union, with particular focus on Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs). Articles 34 and 35 of the DSA require these largest platforms to identify and mitigate systemic risks, including risks to civic discourse, electoral processes, and public security. The instrument's relevance to hybrid threat detection lies in this systemic-risk framing: disinformation is treated not as content to be moderated on a case-by-case basis but as a structural phenomenon requiring platform-level mitigation.

Zinenko, Svoboda, Dereshchuk, Shevchenko, and Shapovalenko provide the most comprehensive recent analysis of the DSA's interaction with NIS2 in the disinformation context, drawing on data from over 280 sources between 2015 and 2024 to evaluate the EU's regulatory, institutional, and technological framework for protecting Member States from disinformation [76]. The analysis identifies an emerging trend toward legal convergence in approaches to addressing disinformation and hybrid threats across Member States, but also significant disparities in enforcement capabilities, digital technology governance, and resilience to hybrid threats. The most significant finding for the present chapter is the identification of a structural conflict between the protection of the free flow of information and the regulation of misinformation – a conflict that occurs within the broader geopolitics of information confrontation and that no single regulatory instrument can resolve.

The DSA's contribution to hybrid threat governance is therefore best understood as complementary rather than primary. The Regulation provides risk-assessment and mitigation logic that can inform HiPSTer-like system development, particularly where such systems are deployed to support civilian platform compliance rather than security operations. The publication's regulatory section [1] notes this complementary positioning. The Zinenko analysis additionally emphasises that effective improvement of EU collective information security will require the coordination of EU policies, the implementation of advanced technologies, including AI, and enhanced cooperation among Member States – a programmatic agenda that places hybrid threat detection systems within a broader resilience architecture rather than treating them as standalone capabilities.

A further integration point is worth noting. The Rampásek and Mesarcík analysis discussed in 7.5 identifies the DSA's vetted-researcher access provisions as a potentially significant mechanism for cybersecurity research, but one whose structure is better suited to systemic-risk analysis than to adversarial threat detection [75]. For HiPSTer-like systems, this suggests that DSA-mediated data access pathways may support certain research and validation activities but remain inadequate for the operational threat-detection workflows envisioned in the project's deployment trajectory.

The DSA's contribution thus completes the cross-regulation landscape surveyed across Chapters 5, 6, and 7. The integrated findings of the literature review, drawing together the technical state of the art (Chapter 3), the adversarial context (Chapter 4), and the three regulatory chapters (Chapters 5–7), are consolidated in Chapter 8.

8. Consolidating Synthesis and Research Findings

8.1 Chapter Framing

This chapter consolidates the findings of the literature review, drawing together the threads developed across Chapters 3 through 7 into an integrated assessment of the scholarly and regulatory landscape within which the HiPSTer project is positioned. The chapter does not introduce new source material; rather, it synthesises the evidence base established in the preceding chapters and identifies the principal research findings of the review. The findings are organised around five observations, each grounded in evidence from multiple chapters, and a sixth section reflects on the implications for the HiPSTer project's broader suite of deliverables.

Finding 1: The EU Cybersecurity Regulatory Landscape Has Consolidated Into a Coherent but Structurally Fragmented Ecosystem

The Network and Information Security Directive 2, the Cyber Resilience Act, the Digital Operational Resilience Act, and the Digital Services Act (Chapter 7) each address distinct dimensions of the cybersecurity challenge – operational resilience, product security, financial-sector resilience, and platform systemic risk, respectively. Together with the General Data Protection Regulation and Law Enforcement Directive (Chapter 5) and the Artificial Intelligence Act (Chapter 6), they constitute a comprehensive but layered architecture of obligations applicable to systems operating at scale across personal data, automated decision-making, and critical infrastructure.

The architecture is coherent in the sense that the principal instruments share a common risk-based regulatory logic and reinforce one another in important respects. The architecture is fragmented in that its instruments were developed sequentially, with limited explicit coordination across their boundaries, resulting in interaction effects that single-instrument analysis cannot capture. The implication for hybrid threat detection systems is that compliance assessment must be conducted across the full regulatory landscape simultaneously, not instrument by instrument.

Finding 2: Cross-Regulation Integration Has Emerged as the Central Methodological Challenge of Contemporary EU Compliance Scholarship

The most significant analytical development across the literature surveyed is the emergence of cross-regulation integration as a first-class methodological problem. The Jorgensen and Ma Unified Digital Twin Compliance Framework, the Beltrán multi-layered approach to algorithmic security and privacy, the Komnios inter-regulatory proportionality framework, and the Rampásek and Mesarcík analysis of cybersecurity research governance (all surveyed in Chapter 7, 7.5) collectively establish that meaningful compliance with any single instrument increasingly requires explicit treatment of its interactions with adjacent instruments.

The methodological response that emerges most clearly from this literature is *compliance-by-design*: the integration of regulatory requirements into system architecture from inception, rather than as a downstream conformity-assessment activity. The principle is well established in the data-protection domain (privacy-by-design under GDPR Article 25 and Article 20 of the Law Enforcement Directive) and is now being extended to the cybersecurity domain (the Canavese et al. architectural treatment of CRA compliance-by-design surveyed in 7.3). The HiPSTer ontological framework – designed from inception to support integration across regulatory, behavioural, and analytical layers – sits squarely within this trajectory.

Finding 3: Ontology-Driven Approaches to Regulatory Compliance Have Moved From Research Curiosity to Demonstrated Practical Applicability

The Castiglione et al. NIS2Onto contribution (Chapter 7, 7.2) provides the clearest external validation of the architectural philosophy underlying HiPSTer's own ontological approach. The authors demonstrate that an OWL ontology can translate juridical language into operational procedures with measurable improvements in compliance verification, risk assessment, and policy implementation. The contribution is significant because cybersecurity governance has historically been a domain in which formal semantic representation has been resisted on grounds of legal complexity and operational variability; NIS2Onto demonstrates that these obstacles are tractable.

The implication for the HiPSTer project is direct. The project's ontological framework – combining OWL ontologies for formal threat actor and campaign modelling, SKOS taxonomies for hierarchical capability classification, and RDF data integration (Chapter 3 of the project's scientific publication [1] – applies the same architectural philosophy that NIS2Onto applies to NIS2 compliance, but extended across the substantive domain of hybrid threat detection rather

than restricted to regulatory mapping. The architectural alignment between these two ontological approaches suggests that semantic representation methodologies are now sufficiently mature for adoption across both the regulatory compliance and operational detection layers of cybersecurity systems.

Finding 4: Hybrid Regulatory Architectures Combining Technical Compliance With Fundamental-Rights Protection Are Subject to Convergent Doctrinal Critique

The regulatory architectures examined in Chapters 5, 6, and 7 share a common structural feature: each combines technical compliance requirements with fundamental rights protection in ways that produce hybrid regulatory architectures. The GDPR combines data-protection rules with Charter-grounded fundamental-rights guarantees; the AI Act combines product-safety logics with fundamental-rights impact assessment; the Cyber Resilience Act combines product-cybersecurity requirements with claimed fundamental-rights protection.

Each architecture is subject to convergent doctrinal critique. The Laux, Wachter, and Mittelstadt critique of trustworthy AI as risk-acceptability (Chapter 6, 6.3) argues that the AI Act conflates two analytically distinct concepts and oversells what regulation can achieve. The Chiara critique of the CRA's fundamental-rights framing (Chapter 7, 7.3) examines whether the Regulation substantively delivers on its rights-protection claims or whether the rights framing functions primarily as legitimation for what is essentially a product-safety instrument. The Tzanou and Vogiatzoglou critique of the GDPR/LED controller-focused approach (Chapter 5, 5.4) demonstrates that the prevailing jurisdictional architecture produces fragmented outcomes that emerging surveillance technologies readily exploit. The Erdogan analysis of the AI Act's national-security loophole (Chapter 6, 6.6) extends the same structural concern into the AI regulatory domain.

The cumulative effect of these doctrinal critiques is to establish that the EU's regulatory ambition for AI and cybersecurity governance substantially exceeds what the formal compliance architecture can deliver on its own. The implication for system designers is that genuine protection of fundamental rights requires architectural choices that go beyond what formal compliance with the relevant Regulations strictly requires. The HiPSTer project's commitment to ontological explainability, human-in-the-loop reasoning, and explicit treatment of the civilian–security boundary exemplifies this kind of beyond-compliance architectural commitment.

Finding 5: Hybrid Threats Systematically Exploit the Seams Between Analytical Disciplines, Regulatory Instruments, and Operational Contexts

The technical state of the art surveyed in Chapter 3 establishes that hybrid threats exploit the seams between OSINT, SOCMINT, and NLP analytical disciplines: bot detectors do not communicate with propaganda classifiers; OSINT platforms do not integrate with cyber threat intelligence pipelines; fact-checking systems do not cross-reference behavioural coordination signals. The adversarial context surveyed in Chapter 4 establishes that Russian and Chinese state-sponsored operations are specifically engineered to exploit these analytical seams, with platform-spanning tactics and culturally tailored content designed to evade discipline-bounded defensive systems.

The regulatory survey across Chapters 5, 6, and 7 establishes that hybrid threats also exploit the seams between regulatory instruments: the GDPR/LED civilian–security boundary, the AI Act national-security exemption, the DSA's platform-systemic-risk framing that adequately addresses content but not coordinated tradecraft, and the NIS2/CRA boundary between operational resilience and product security. Each of these seams provides operational latitude for adversarial actors who structure their activities to fall between regulatory categories.

The defensive response must therefore address all three forms of seam-exploitation simultaneously. A defensive framework that resolves analytical fragmentation but not regulatory fragmentation will produce capabilities that cannot be lawfully deployed; a framework that addresses regulatory fragmentation but not analytical fragmentation will produce lawful but ineffective systems. The HiPSTer methodological framework – through its ontology-driven semantic integration, its capability-domain classification, and its explicit treatment of the civilian–security regulatory boundary – exemplifies an architectural response designed to address the multi-layered integration challenge that single-axis defensive approaches cannot resolve.

8.2 Implications for the HiPSTer Deliverable Suite

The findings consolidated in this chapter have direct implications for the HiPSTer project's broader deliverable suite. Three implications warrant particular emphasis.

First, the scholarly and regulatory landscape supports the architectural choices the HiPSTer project has made. The ontology-driven semantic integration, the capability-domain classification through SKOS taxonomies, the explicit modelling of the civilian–security boundary, and the human-in-the-loop reasoning architecture are not idiosyncratic project choices but exemplifications of architectural patterns that the contemporary scholarly and regulatory landscape increasingly demands. The compliance-prescriptive sibling deliverables (the GDPR best-practices analysis and the AI Act compliance research report; the privacy-by-design summary reports prepared by Lavišius) can draw on this evidence base to ground their compliance positions in the established scholarship.

Second, the open challenges identified across the chapters – particularly around cross-regulation integration methodology, multilingual indicator detection performance variation, and the ongoing evolution of CJEU jurisprudence – define the research agenda for subsequent project phases and for the project's commercialisation pathway through Novian. The TRL-4 validation achieved through the present project lays the foundation for progression toward operational deployment, but the regulatory landscape against which that deployment will be assessed continues to evolve.

Third, the literature review establishes the foundational layer of scholarship on which other sibling deliverables can build. The case-study, comparative-analysis, and content-analysis document prepared by Lavišius can reference the technical scholarship surveyed in Chapter 3; the GDPR and AI Act compliance reports can reference the regulatory scholarship surveyed in Chapters 5 and 6; the privacy-by-design summary reports can draw on the compliance-by-design principles identified across the regulatory chapters. The deliverable thus functions as intended: as the foundational scholarship layer underpinning the project's broader deliverable suite, rather than as a standalone compliance document.

The Summary and Conclusion (Chapter 9) draws these implications together into a concise closing.

9. Summary and Conclusion

This literature review has surveyed the academic and policy scholarship relevant to the design, deployment, and governance of hybrid threat detection systems operating within the European regulatory environment, with particular attention to the privacy and cybersecurity dimensions specified in the project brief. The review was conducted as a PRISMA-ScR-informed scoping review, combining material drawn from the project's scientific publication and prior comprehensive review (for the technical state of the art and adversarial context) with four targeted Web of Science Boolean searches executed for the present deliverable (for the regulatory, privacy, and cybersecurity-governance landscape). The methodology, search strategy, and screening criteria are documented in Chapter 2.

The substantive findings of the review are consolidated in Chapter 8 and need not be restated in full here. In summary, the review establishes that the EU regulatory landscape for AI and cybersecurity has consolidated into a coherent but structurally fragmented ecosystem in which cross-regulation integration has emerged as the central methodological challenge; that ontology-driven approaches to regulatory compliance and operational threat detection have moved from research curiosity to demonstrated practical applicability; that the hybrid regulatory architectures combining technical compliance with fundamental-rights protection are subject to convergent doctrinal critique that contemporary system designers must take seriously; and that hybrid threats systematically exploit the seams between analytical disciplines, between regulatory instruments, and between civilian and security operational contexts – a multi-layered exploitation pattern that single-axis defensive frameworks cannot address.

For the HiPSTer project specifically, the review establishes that the architectural choices the project has made – ontology-driven semantic integration, capability-domain classification through SKOS taxonomies, explicit modelling of the civilian–security regulatory boundary, and human-in-the-loop reasoning – are not idiosyncratic but rather exemplify architectural patterns that the contemporary scholarly and regulatory landscape increasingly demands. The review thus functions as intended within the project's broader deliverable suite: as the foundational scholarship layer on which the GDPR best-practices analysis, the Data Protection Impact Assessment, the AI Act compliance research report, the privacy-by-design summary reports, and the consolidated case-study and comparative-analysis document can build their compliance and analytical positions.

The review also identifies open challenges. The regulatory landscape continues to evolve, with the AI Act and Cyber Resilience Act in early implementation and CJEU jurisprudence on data protection in security contexts evolving. Multilingual indicator detection exhibits substantial performance variation across languages, with stronger results for high-resource languages than for lower-resource European languages. Cross-regulation integration methodology –

though increasingly recognised as the central challenge – remains an active area of scholarship rather than a settled domain. The TRL-4 validation achieved through the present project lays the foundation for progression toward operational deployment through the project's commercialisation pathway, but the regulatory and scholarly landscape against which that deployment will be assessed will continue to evolve through the commercialisation period and beyond.

The literature review supports the conclusion that hybrid threat detection systems can be designed to operate within – rather than despite – the European fundamental rights and data protection frameworks, provided that regulatory alignment is treated as an architectural design problem from the outset rather than as a downstream compliance activity. The HiPSTer methodological framework demonstrates one viable approach to this design challenge. The scholarship surveyed in this review establishes the evidence base on which that demonstration rests.

List of Abbreviations

Abbreviation	Translation
AI	Artificial Intelligence
AI Act	Artificial Intelligence Act (Regulation (EU) 2024/1689)
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
AUROC	Area Under the Receiver Operating Characteristic curve
BERT	Bidirectional Encoder Representations from Transformers
CDSMD	Copyright in the Digital Single Market Directive
CE	Conformité Européenne (European Conformity)
CFR	Charter of Fundamental Rights of the European Union
CIs	Critical Infrastructures
CJEU	Court of Justice of the European Union
CRA	Cyber Resilience Act (Regulation (EU) 2024/2847)
CTI	Cyber Threat Intelligence
DEMATEL	Decision-Making Trial and Evaluation Laboratory
DGA	Domain Generation Algorithm
DNS	Domain Name System
DOA	Description of Action
DORA	Digital Operational Resilience Act (Regulation (EU) 2022/2554)
DPIA	Data Protection Impact Assessment
DSA	Digital Services Act (Regulation (EU) 2022/2065)
DTCEM	Digital Twin Compliance Evaluation Matrix
ECLI	European Case Law Identifier
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EEAS	European External Action Service

ENISA	European Union Agency for Cybersecurity
EU	European Union
Europol	European Union Agency for Law Enforcement Cooperation
FIMI	Foreign Information Manipulation and Interference
FL	Federated Learning
FRIA	Fundamental Rights Impact Assessment
GDPR	General Data Protection Regulation (Regulation (EU) 2016/679)
GPT	Generative Pre-trained Transformer
HiPSTer	Hybrid-threat Intelligence Platform leveraging Semantic Technologies for Threat Recognition
IDS	Intrusion Detection System
IoMT	Internet of Medical Things
IoT	Internet of Things
IP	Intellectual Property
IRA	Internet Research Agency
JIC	Journal of Intelligent Communication
KTU	Kaunas University of Technology (Kauno Technologijos Universitetas)
L3CE	Lithuanian Cybercrime Centre of Excellence for Training, Research and Education
LED	Law Enforcement Directive (Directive (EU) 2016/680)
LLM	Large Language Model
LSTM	Long Short-Term Memory
MITA	Lithuanian Innovation Agency (Mokslo, Inovacijų ir Technologijų Agentūra)
ML	Machine Learning
MRU	Mykolas Romeris University
NATO	North Atlantic Treaty Organization
NER	Named Entity Recognition
NIS2	Network and Information Security Directive 2 (Directive (EU) 2022/2555)
NLF	New Legislative Framework
NLP	Natural Language Processing
OECD	Organisation for Economic Co-operation and Development
OSINT	Open-Source Intelligence
OWL	Web Ontology Language

PDAV	Poveikio Duomenų Apsaugai Vertinimas (Data Protection Impact Assessment, Lithuanian)
PRISMA-ScR	Preferred Reporting Items for Systematic Reviews and Meta-Analyses — Extension for Scoping Reviews
RAG	Retrieval-Augmented Generation
RDF	Resource Description Framework
RIS	Research Information Systems (citation file format)
SHACL	Shapes Constraint Language
SKOS	Simple Knowledge Organisation System
SME	Small and Medium-sized Enterprise
SOC	Security Operations Centre
SOCMINT	Social Media Intelligence
SOTA	State of the Art
SPARQL	SPARQL Protocol and RDF Query Language
SyRI	System Risk Indication (Dutch fraud detection system)
TDM	Text and Data Mining
TRL	Technology Readiness Level
TTP	Tactics, Techniques, and Procedures
UDTCF	Unified Digital Twin Compliance Framework
VLOP	Very Large Online Platform
VLOSE	Very Large Online Search Engine
W3C	World Wide Web Consortium
WoS	Web of Science
XAI	Explainable Artificial Intelligence

Bibliography

[1] E. Bružė, R. A. Paskauskas, R. Matulytė, G. Sabaliauskaitė, and T. Lavišius, "Semantic integration of hybrid threat intelligence: The HiPSTer Ontological Method for cross-domain correlation in influence operations," *Open Research Europe*, 2026 (in press).

[2] R. A. Paskauskas, E. Bružė, G. Sabaliauskaitė, R. Matulytė, and T. Lavišius, "Hybrid-threat intelligence: A critical review of semantic integration challenges and the role of the HIPSTer ontological framework," *Journal of Intelligent Communication*. 5, 1 (Jun. 2026), 140–173. DOI:https://doi.org/10.54963/jic.v5i1.2128.

[3] H. Tricco *et al.*, "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and explanation," *Annals of Internal Medicine*, vol. 169, no. 7, pp. 467–473, Oct. 2018, doi: 10.7326/M18-0850.

[4] European Union Agency for Cybersecurity (ENISA), "Threat Landscape 2025," ENISA, 2025. [Online]. Available: https://www.enisa.europa.eu/

[5] A. Y. Balogun, A. Ismaila Alao, and O. O. Olaniyi, "Disinformation in the digital era: The role of deepfakes, artificial intelligence, and open-source intelligence in shaping public trust and policy responses," *Computer Science & IT Research Journal*, vol. 6, no. 2, 2025.

[6] A. Yadav, A. Kumar, and V. Singh, "Open-source intelligence: A comprehensive review of the current state, applications and future perspectives in cyber security," *Artificial Intelligence Review*, vol. 56, no. 11, pp. 12407–12438, 2023, doi: 10.1007/s10462-023-10454-y.

[7] R. Dover, "SOCMINT: A shifting balance of opportunity," *Intelligence and National Security*, vol. 35, no. 2, pp. 216–232, 2020, doi: 10.1080/02684527.2019.1694132.

[8] P. Cardenas, B. Obara, G. Theodoropoulos *et al.*, "Analysing social media as a hybrid tool to detect and interpret likely radical behavioural traits for national security," in *Proc. 2019 IEEE International Conference on Big Data*, IEEE, 2019, pp. 4579–4588, doi: 10.1109/BigData47090.2019.9006505.

[9] J. Mothe, M. Roche, and L. Tanguy, "Countering hybrid security threats through information analysis," *Information Processing & Management*, vol. 57, no. 6, p. 102337, 2020, doi: 10.1016/j.ipm.2020.102337.

[10] Z. Ellaky, F. Benabbou, Y. Matrane *et al.*, "A hybrid deep learning architecture for social media bots detection based on BiGRU-LSTM and GloVe word embedding," *IEEE Access*, vol. 12, pp. 100278–100294, 2024, doi: 10.1109/ACCESS.2024.3430533.

[11] K. S. Sangher, A. Singh, and H. M. Pandey, "LSTM and BERT based transformers models for cyber threat intelligence for intent identification of social media platforms exploitation from darknet forums," *International Journal of Information Technology*, vol. 16, no. 8, pp. 5277–5292, 2024, doi: 10.1007/s41870-024-02077-5.

[12] M. S. Biagio, S. Simoncini, E. La Mattina *et al.*, "MARPLE: A framework for social media threat intelligence," in *Proc. 2024 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA)*, IEEE, 2024, pp. 1–6, doi: 10.1109/ACDSA59508.2024.10467969.

[13] L. Luceri, S. Giordano, and E. Ferrara, "Exposing cross-platform coordinated inauthentic activity in the run-up to the 2024 U.S. election," in *Proc. ACM Web Conference 2025 (WWW '25)*, 2025, doi: 10.1145/3696410.3714658.

[14] H. Sarhan and S. Hegelich, "There was coordinated inauthentic user behavior in the COVID-19 German X-discourse, but did it really matter?," *Frontiers in Communication*, vol. 10, 2025, doi: 10.3389/fcomm.2025.1558881.

[15] R. Rogers and N. Righetti, "Coordinated inauthentic behaviour on Facebook? A typology of manufactured attention," *Platforms & Society*, vol. 2, 2025.

[16] G. Wang, P. Liu, J. Huang *et al.*, "KnowCTI: Knowledge-based cyber threat intelligence entity and relation extraction," *SSRN*, 2024.

[17] G. Perrina, T. Caselli, and R. Sprugnoli, "AGIR: Automatic generation of intelligence reports from structured threat data," *Natural Language Engineering*, vol. 28, no. 6, pp. 749–770, 2022, doi: 10.1017/S1351324921000334.

[18] P. Qi, Z. Yan, W. Hsu *et al.*, "SNIFFER: Multimodal large language model for explainable out-of-context misinformation detection," *arXiv preprint*, 2024, doi: 10.48550/arXiv.2403.03170.

[19] X. Li, Y. Zhang, and E. C. Malthouse, "Large language model agent for fake news detection," *arXiv preprint*, 2024, doi: 10.48550/arXiv.2405.01593.

[20] J. Jose and R. Greenstadt, "Are large language models good at detecting propaganda?," in *Proc. 5th International Workshop on Cyber Social Threats at ICWSM 2024*, 2024.

[21] J. A. Goldstein *et al.*, "Disinformation capabilities of large language models," *arXiv preprint*, 2024, doi: 10.48550/arXiv.2311.08838.

[22] H. Shulman *et al.*, "Propaganda by prompt: Tracing hidden linguistic strategies in large language models," *Information Processing & Management*, vol. 62, 2025.

[23] D. Hamzic, F. Skopik, M. Landauer *et al.*, "Enhancing cyber situational awareness with AI: A novel pipeline approach for threat intelligence analysis and enrichment," in *ARES 2025 Workshops*, LNCS vol. 15995, 2025, pp. 44–62, doi: 10.1007/978-3-032-06155-3_4.

[24] I. Borisov, "Maskirovka – The art of deception à la Russe," *RMT*, vol. 2024, pp. 192–207, 2024.

[25] J. Kalensky and R. Osadchuk, *How Ukraine fights Russian disinformation: Beehive vs Mammoth*, Hybrid CoE Research Report 11. The European Centre of Excellence for Countering Hybrid Threats, 2024. [Online]. Available: https://www.hybridcoe.fi/

[26] G. Eady, T. Paskhalis, J. Zilinsky *et al.*, "Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior," *Nature Communications*, vol. 14, p. 62, 2023, doi: 10.1038/s41467-022-35576-9.

[27] VIGINUM, *Portal Kombat: A structured and coordinated pro-Russian propaganda network*, Secrétariat général de la défense et de la sécurité nationale, France, 2024. [Online]. Available: https://www.sgdsn.gouv.fr/

[28] E. S. Cochran, "China's 'Three Warfares': People's Liberation Army influence operations," *International Bulletin of Political Psychology*, vol. 20, 2020.

[29] P. Charon and J.-B. Jeangène Vilmer, *Chinese influence operations: A Machiavellian moment*. Institute for Strategic Research (IRSEM), French Ministry for the Armed Forces, 2021. [Online]. Available: https://www.irsem.fr/

[30] D. Finkelstein, S. Yanovsky, J. Zucker *et al.*, "Information manipulation on TikTok and its relation to American users' beliefs about China," *Frontiers in Social Psychology*, vol. 2, p. 1497434, 2025, doi: 10.3389/frsps.2025.1497434.

[31] European External Action Service (EEAS), "Information integrity and countering foreign information manipulation & interference (FIMI)," EEAS, 2026. [Online]. Available: https://www.eeas.europa.eu/

[32] European Commission, "Joint Framework on Countering Hybrid Threats," JOIN(2016) 18 final, Brussels, 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52016JC0018

[33] North Atlantic Treaty Organization (NATO), "NATO 2022 Strategic Concept," Madrid, 2022. [Online]. Available: https://www.nato.int/strategic-concept/

[34] B. Goodman and S. Flaxman, "European Union regulations on algorithmic decision-making and a 'right to explanation'," *AI Magazine*, vol. 38, no. 3, pp. 50–57, Fall 2017, doi: 10.1609/aimag.v38i3.2741.

[35] S. Wachter, B. Mittelstadt, and L. Floridi, "Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation," *International Data Privacy Law*, vol. 7, no. 2, pp. 76–99, May 2017, doi: 10.1093/idpl/ipx005.

[36] A. D. Selbst and S. Barocas, "The intuitive appeal of explainable machines," *Fordham Law Review*, vol. 87, no. 3, pp. 1085–1139, Dec. 2018.

[37] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, art. 12, Feb. 2019, doi: 10.1145/3298981.

[38] Q. Li, Z. Wen, Z. Wu *et al.*, "A survey on federated learning systems: Vision, hype and reality for data privacy and protection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 4, pp. 3347–3366, Apr. 2023, doi: 10.1109/TKDE.2021.3124599.

[39] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, no. 6, pp. 305–311, Jun. 2020, doi: 10.1038/s42256-020-0186-1.

[40] L. Lyu, H. Yu, X. Ma *et al.*, "Privacy and robustness in federated learning: Attacks and defenses," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 7, pp. 8726–8746, Jul. 2024, doi: 10.1109/TNNLS.2022.3216981.

[41] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, OJ L 119, 4.5.2016, p. 89.

[42] M. Tzanou and P. Vogiatzoglou, "National security and new forms of surveillance: From the data retention saga to a data subject centred approach," *European Papers*, vol. 10, no. 3, 2025.

[43] Seimas of the Republic of Lithuania, "Law on Legal Protection of Personal Data Processed for the Prevention, Investigation, Detection or Prosecution of Criminal Offences, the Enforcement of Penalties or for National Security or Defence Purposes," Vilnius, n.d. [Online]. Available: https://e-tar.lt/portal/lt/legalAct/TAR.299D835159BE/asr

[44] Court of Justice of the European Union, Case C-634/21, *OQ v Land Hessen (SCHUFA Holding)*, ECLI:EU:C:2023:957, 7 December 2023.

[45] Court of Justice of the European Union, Case C-203/22, *CK v Dun & Bradstreet Austria GmbH*, 2025.

[46] Court of Justice of the European Union, Case C-252/21, *Meta Platforms Inc. and Others v Bundeskartellamt*, ECLI:EU:C:2023:537, 4 July 2024.

[47] Court of Justice of the European Union, Case T-557/20, *Single Resolution Board v European Data Protection Supervisor*, ECLI:EU:T:2023:219, 2025.

[48] "GDPR Best Practices Analysis and Data Protection Impact Assessment," HiPSTer Project Deliverable, MITA Project No. 02-002-P-0001, 2026.

[49] M. Almada and N. Petit, "The EU AI Act: Between the rock of product safety and the hard place of fundamental rights," *Common Market Law Review*, vol. 62, no. 1, pp. 85–120, Feb. 2025, doi: 10.54648/cola2025004.

[50] J. Laux, S. Wachter, and B. Mittelstadt, "Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk," *Regulation & Governance*, vol. 18, no. 1, pp. 3–32, Jan. 2024, doi: 10.1111/rego.12512.

[51] N. Díaz-Rodríguez, J. Del Ser, M. Coeckelbergh, M. L. de Prado, E. Herrera-Viedma, and F. Herrera, "Connecting the dots in trustworthy artificial intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation," *Information Fusion*, vol. 99, p. 101896, Nov. 2023, doi: 10.1016/j.inffus.2023.101896.

[52] A. Mantelero, "The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template," *Computer Law & Security Review*, vol. 54, p. 106020, 2024, doi: 10.1016/j.clsr.2024.106020.

[53] L. Longo, M. Brcic, F. Cabitza *et al.*, "Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions," *Information Fusion*, vol. 106, p. 102301, Jun. 2024, doi: 10.1016/j.inffus.2024.102301.

[54] A. d'Avila Garcez and L. C. Lamb, "Neurosymbolic AI: The 3rd wave," *Artificial Intelligence Review*, vol. 56, no. 11, pp. 12387–12406, Nov. 2023, doi: 10.1007/s10462-023-10448-w.

[55] J. P. Quintais, "Generative AI, copyright and the AI Act," *Computer Law & Security Review*, vol. 56, p. 106107, Apr. 2025, doi: 10.1016/j.clsr.2025.106107.

[56] Organisation for Economic Co-operation and Development (OECD), "Intellectual property issues in artificial intelligence trained on scraped data," OECD, Paris, 2025. [Online]. Available: https://www.oecd.org/

[57] European Commission, "Guidelines on prohibited artificial intelligence (AI) practices," Brussels, Feb. 2025. [Online]. Available: https://digital-strategy.ec.europa.eu/

[58] European Commission, "General-purpose AI code of practice," Brussels, Jul. 2025. [Online]. Available: https://digital-strategy.ec.europa.eu/

[59] E. Glerean, *Fundamentals of secure AI systems with personal data*. European Data Protection Board (EDPB), 2025. [Online]. Available: https://www.edpb.europa.eu/

[60] Europol, *AI bias in law enforcement: A practical guide*. Europol Innovation Lab Observatory Report, Publications Office of the European Union, 2025. [Online]. Available: https://www.europol.europa.eu/

[61] I. Erdogan, "Diving into the iceberg: Establishing transparency in AI for law enforcement," *European Papers*, vol. 9, no. 3, pp. 956–977, 2024.

[62] Novian. Novian's Consolidated Revenue Increased 2.4% in 2024 to EUR 38.9 Million. [Online]. Available: https://novian.io/news/novians-consolidated-revenue-increased-2-4-in-2025-to-eur-38-9-million/

[63] G. Castiglione, D. F. Santamaria, G. Bella, L. Brisindi, and G. Puccia, "Guiding cybersecurity compliance: An ontology for the NIS 2 directive," *Computers & Security*, vol. 157, p. 104617, Oct. 2025, doi: 10.1016/j.cose.2025.104617.

[64] K. Mentzelou, P. T. Chountalas, F. C. Kitsios, A. I. Magoutas, and T. K. Dasaklis, "Identifying and modeling barriers to compliance with the NIS2 Directive: A DEMATEL approach," *Journal of Cybersecurity and Privacy*, vol. 5, no. 4, art. 97, Nov. 2025, doi: 10.3390/jcp5040097.

[65] E. Rehnstam, W. Winquist, and S. Hacks, "NIS2 Directive in Sweden: A report on the readiness of Swedish critical infrastructure," in *Secure IT Systems, NordSec 2024*, LNCS vol. 15396, 2025, pp. 176–195, doi: 10.1007/978-3-031-79007-2_10.

[66] P. G. Chiara, "Understanding the regulatory approach of the Cyber Resilience Act: Protection of fundamental rights in disguise?," *European Journal of Risk Regulation*, 2025, doi: 10.1017/err.2025.9.

[67] P. Ajmera, "Cybersecurity in the Internet of Things: Trends and challenges in a nascent field," *Computer Law & Security Review*, vol. 59, p. 106204, Nov. 2025, doi: 10.1016/j.clsr.2025.106204.

[68] D. Canavese, A. Ferreira, R. Laborde, and M. A. Kandi, "Towards an architecture for managing security under the EU Cyber Resilience Act," in *Security and Trust Management, STM 2025*, LNCS vol. 16137, 2026, pp. 159–172, doi: 10.1007/978-3-032-06155-3_9.

[69] M. Stawowski, A. Sobczyk, M. Gajda, T. Wojtas, T. Pajak, K. Siwy, and G. Blinowski, "Cyber Soldier project — A threat-led methodology for assessing the digital resilience of cybersecurity systems," *International Journal of Electronics and Telecommunications*, vol. 72, no. 1, art. 157878, 2026, doi: 10.24425/ijet.2026.157878.

[70] J. Wittlin, A. Ossowska, and K. Sawicka, "Is DORA an opportunity for more balanced distribution of third-party risk management assurance responsibilities between banks, regulators and technology providers?," *Information & Communications Technology Law*, vol. 35, no. 1, pp. 109–116, Jan. 2026, doi: 10.1080/13600834.2025.2514385.

[71] E. Seid, F. Blix, and O. Popov, "Cyber resilience using ASFA: DORA-compliant threat-led penetration testing," in *Critical Information Infrastructures Security, CRITIS 2024*, LNCS vol. 15549, 2025, pp. 269–288, doi: 10.1007/978-3-031-84260-3_16.

[72] B. N. Jorgensen and Z. G. Ma, "Digital Twins under EU law: A unified compliance framework across smart cities, industry, transportation, and energy systems," *Electronics*, vol. 14, no. 24, art. 4881, Dec. 2025, doi: 10.3390/electronics14244881.

[73] M. Beltrán, "AI algorithms under scrutiny: GDPR, DSA, AI Act and CRA as pillars for algorithmic security and privacy in the European Union," *Computers & Security*, vol. 158, p. 104628, Nov. 2025, doi: 10.1016/j.cose.2025.104628.

[74] K. Komnios, "Smart metering, fundamental rights, and multi-regime governance in EU energy law: Towards a coherent regulatory framework for a digitalized energy system," *Journal of World Energy Law & Business*, vol. 19, no. 1, art. jwag003, Jan. 2026, doi: 10.1093/jwelb/jwag003.

[75] M. Rampásek and M. Mesarcík, "Rethinking cybersecurity research governance: Lessons from the Act on Digital Services?," *European Journal of Risk Regulation*, 2025, doi: 10.1017/err.2025.10060.

[76] A. Zinenko, I. Svoboda, T. Dereshchuk, N. Shevchenko, and M. Shapovalenko, "EU policy on countering disinformation and its impact on the information security of Member States," *Revista Juridica Portucalense*, no. 39, 2026, doi: 10.34625/issn.2183-2705(39.1)2026.ic-7.