FERMI

FAKE NEWS
RISK MITIGATOR

Project acronym:	FERMI
Project full title:	Fake nEws Risk Mitigator
Call identifier	HORIZON-CL3-2021-FCT-01
Start date:	01/10/2022
End date:	30/09/2025
Grant agreement no:	101073980

D6.3 FERMI outreach and collaboration management report – final version

Work package: WP6	
Version: 1	
Deliverable type: RE	**Dissemination level: PU**
Official submission date: M36	**Actual submission date: M36**

Leading author(s):

Surname	First name	Beneficiary
Beli	Aikaterini	The Lisbon Council
Papadopoulou	Martha	The Lisbon Council
Paciaroni	Alessandro	The Lisbon Council
Zaccaria	Maria Chiara	The Lisbon Council

Contributing partner(s):

Surname	First name	Beneficiary
Valente	Catarina	INOV
Vrotsou	Christina	Convergence
Kokoliou	Panagiota	Convergence
Solomou	Alexia	IANUS Technologies Ltd.
Sainio	Miia	PUCF
Dimakopoulos	Nikos	ITML
Bageorgou	Eirini	ITML
Stuchtey	Tim	BIGS
Kolliarakis	Georgios	BIGS

Peer reviewer(s):

Surname	First name	Beneficiary
Forsell	Arttu	FMI
Wallin	Galaxia	SPA

Ethics reviewer:

Surname	First name	Beneficiary
Fikenscher	Sven-Eric	BPA

Security reviewer:

Surname	First name	Beneficiary
Mattes	Tobias	BPA

Funded by the European Union

Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
0.1	26 Jun 2025	ToC	LC
0.2	31 Jul 2025	Partner contributions	INOV, CONV, IANUS, PUCF, ITML
0.3	01 Aug 2025	Completion of first draft	LC
0.4	29 Aug 2025	Peer review	FMI
0.5	29 Aug 2025	Peer review	SPA
0.6	08 Sept 2025	Ethics review	BPA
0.7	17 Sept 2025	Partner contributions to second draft	INOV, CONV, IANUS, PUCF, ITML, BIGS
0.8	25 Sept 2025	Security board review	BPA
0.9	26 Sept 2025	Completion of second draft	LC
1	30 Sept 2025	Submission	BPA

Funded by the European Union

Executive summary

This deliverable presents the final implementation report of the FERMI Communication, Dissemination, Exploitation and Sustainability Strategy (CDES), covering the full lifecycle of the project from 01 October 2022 to 30 September 2025. It provides a comprehensive overview of the activities carried out, results achieved, and lessons learned in the domains of outreach, stakeholder engagement, scientific dissemination, communication, and collaboration.

The report consolidates all communication and dissemination initiatives undertaken by the consortium, including contributions to peer-reviewed publications, conference participation, stakeholder engagement via digital channels, strategic partnerships, and participation in Horizon Europe clusters. Furthermore, it evaluates the effectiveness of the CDES strategy by analysing key performance indicators (KPIs) set out in the Grant Agreement and their achievement over time.

The communication and dissemination KPIs could all be reached. In many cases, they could even be greatly exceeded. The project's website and social media engagement is a case in point. The numbers for website visits, accesses, social media activities were significantly higher than required. Similarly, dissemination in the form of publications and events was a lot more successful than necessary to meet the KPIs.

Special attention is paid to the formation of a sustainable "FERMI community of interest", the expansion of FERMI's visibility in scientific, policy and practitioner communities, and the successful collaboration with external initiatives through joint webinars, trainings, and EU project clusters (such as Fighting Fake News and AI4SafeEurope).

By aligning FERMI's outreach with its technical and scientific goals, this deliverable demonstrates how FERMI has positioned itself as a meaningful contributor in the fields of digital security, disinformation mitigation, and AI-supported risk assessment. It reflects the consortium's commitment to ensuring the long-term impact and reusability of its outputs beyond the life of the project.

Funded by the European Union

Table of Contents

Executive summary 4
Table of Contents 5
List of tables 8
List of figures 9
Abbreviations 10
1 Introduction 14
2 Objectives and key implementation steps 15
3 Monitoring and Key Performance Indicators 16
4 Dissemination 20
4.1 Overview of dissemination 20
4.2 Ecosystem mapping 20
4.3 Short report on the dissemination for LEAs and policymakers 21
4.3.1 Thematic Seminars for Experts, Practitioners and Policymakers 22
4.3.2 Enganging EDMO 23
4.3.3 Short report on the dissemination activities with standardisation bodies and standardisation experts 23
4.3.3.1 Dissemination to standardisation bodies 23
4.3.3.2 Dissemination to standardisation experts 24
4.3.3.3 Webinar on operational standards regarding disinformation 24
4.4 Short report on the scientific and technical dissemination 25
4.5 Short report on the dissemination for the private sector 27
4.6 Short report on the dissemination to the general public 28
4.6.1 The FERMI Final Conference 29
4.7 Synergies with other initiatives and projects 30
4.8 Training activities 32
4.8.1 External Webinar – "Digital Trust in Action: Technological Approaches and Citizen Empowerment to Combat Disinformation" 33
4.8.2 External Webinar – "How Disinformation Evolves: Narratives, Digital Influence, and Trust" 34
4.8.3 Training activities aimed at LEAs 36
4.8.3.1 Training activities by SPA 36
4.8.3.2 Training activities by PUCF 37
5 Communication 40
5.1 Overview of communication 40
5.2 FERMI community 41
5.3 FERMI social media presence 42
6 Recommendations for operational standards 44
6.1 Recommendations for the first operational standard - Operational standard for the ethical and responsible use of technology by law enforcement 44
6.1.1 Purpose and Added Value 44
6.1.2 Key Components of the Standard 44
6.1.3 Strategic Relevance for the Project 45
6.1.4 Relevance Within and Beyond the Project 46
6.2 Recommendation for second operational standard - Use of Technologies to tackle Disinformation by Law Enforcement Agencies 46
6.2.1 Introduction 46
6.2.2 Scope 47
6.3 Third contribution to an operational standard 47
7 Impact and further recommendations on advancing the fight against disinformation 48
8 Conclusions 51
List of References 52
Annex A – Recommendation for Operational Standard for the Ethical and Responsible Use of Technology by Law Enforcement 53
Part 1 53

Executive Summary 53
Definition of Terms 53
Abbreviated Terms 54
1. Introduction 54
2. Links with the Existing Legal Framework 55
3. Core Ethical Principles and Governance Foundations 56
4. Roles and Responsibilities 58
5. Data Governance and Protection 58
6. Risk Assessment and Management 60
7. Transparency and Explainability 61
8. Human Oversight and Decision-Making 62
9. Procurement and Development Controls 64
10. Monitoring, Evaluation and Auditing 65
11. Public Engagement and External Transparency 66
12. Interoperability and Cross-Border Cooperation 68
Part B 68
13. Continuous Monitoring 68
Conclusions and Recommendations 68
Bibliography 69
Annex B - Recommendation for Operational Standard on the Use of Technologies to tackle Disinformation by Law Enforcement Agencies 74
1. Introduction 74
2. Scope 74
3. References 75
3.1 Normative references 75
3.2 Informative references 75
4. Definition of Terms 77
5. Abbreviated Terms 79
6. Core Ethical and Legal Principles 79
7. Basic Elements of the Disinformation Framework 80
7.1 Legal definition 80
7.2 Characteristics of disinformation 81
7.3 Actors and roles 81
7.3.1 Managers: 81
7.3.2 Implementers: 81
7.4 Recent trends in tackling disinformation by public authorities 81
8. Information & Data Governance 82
8.1 Types of data and information that should be gathered to be used in AI and big data technologies 82
8.1.1 Data 82
8.1.2 Information 82
8.2.1 General 83
9. Risk Management 85
10. Transparency and Explainability 86
10.1 Explainability 86
10.1.1 Technical Documentation (Article 11 and Annex IV / XI) 86
10.1.2 Record-Keeping / Logging (Article 12) 86
10.1.3 Explainability for Deployers (Article 13) 86
10.1.4 Notice and Disclosure to Individuals (Article 52) 86
10.1.5 Right to Explanation on Adverse Impacts (Article 86) 86
10.2 Transparency 87
10.2.1 General Transparency Obligations (Article 52) 87
10.2.2 High-Risk AI Systems Transparency (Articles 13, 14) 87
10.2.3 General Purpose AI (GPAI) and Foundation Models (Article 52a / 52b) 87
10.2.4 Right to Explanation (Article 86) 87

10.2.5 Record-Keeping and Logging (Article 12) 88
11. Framework for Meaningful Human Oversight and Decision-Making over AI-assisted Disinformation Detection and Response 88
11.1 Human-In-The-Loop (HITL) Required 88
11.2 Right to Explanation 88
11.3 Override and Escalation 88
11.4 Logging and Auditability 88
11.5 Training and Competency 88
11.6 Transparency to the Public 89
11.7 Implementation Checklist for LEA Enforcement AI Oversight 89
12. Accountability and Responsibility Framework for AI Use in Disinformation Operations by LEAs 89
12.1 Accountability Structure based on Roles 89
12.1.1 AI Provider 89
12.1.2 LEA AI Lead (Accountable Officer) 90
12.1.3 Analyst / System User 90
12.1.4 Supervisory Authority Liaison 90
12.2 Oversight and Legal Controls 90
12.3 Traceability Measures 90
12.4 Bias & Performance Monitoring 90
13. Training and Competency of LEA personnel using or overseeing technologies that aim at preventing or tackling disinformation 91
13.1 General Training 91
13.2 Certification 91
13.3 Role-Specific Modules 91
13.4 Continuing Education 91
13.5 Best Practices regarding Training 91
14. Procurement and Development Standards 92
14.1 Procurement Standards 92
14.2 Development Standards 92
14.3 Prohibited Uses 92
15. Evaluation, Monitoring, and Auditing & Governance and Oversight of AI Systems for Disinformation 93
15.1 Evaluation Mechanisms, Performance Monitoring, Periodic Audits 93
15.1.1 Evaluation 93
15.1.2 Monitoring 93
15.1.3 Auditing 93
15.2 Governance and Oversight 94
16. Public Engagement and Stakeholder Communication in Disinformation Response by LEAs 94
16.1 Public Transparency Requirements 94
16.2 Stakeholder Communication Requirements 94
16.3 Public Engagement Requirements 94
17. Framework for Interoperability and Cross-Border Cooperation on Tackling Disinformation by LEAs 94
18. Redress in Automated Decision-Making 95
Annex C - Public Review: Comments on Draft ETSI TR 104 137 V0.0.3 (2025-04), DTR/CYBER-00156 Cyber Security; Human-to-Human Online Preventative Security; H2H – OPS 97

Funded by the European Union

List of tables

Table 1 - KPI at M34 18
Table 2 - Key points discussed 37
Table 3 - Communication mechanism for each communication channel per project phases 41
Table 4 - Core ethical principles and governance foundations 57
Table 5 - Data governance and protection 60
Table 6 - Risk assessment and management 61
Table 7 - Transparency and explainability 62
Table 8 - Human oversight and decision-making 63
Table 9 - Procurement and development controls 65
Table 10 - Monitoring, evaluation and auditing 66
Table 11 - Public engagement and external transparency 67
Table 12 - Implementation checklist for LEA enforcement AI oversight 89
Table 13 - Public Review: Comments on Draft ETSI TR 104 137 V0.0.3 (2025-04), DTR/CYBER-00156 Cyber Security; Human-to-Human Online Preventative Security; H2H – OPS 109

List of figures

Figure 1 – FERMI Operational Standards webinar agenda 25
Figure 2 - Screenshot of Clusters page in FERMI website 30

Funded by the European Union

Abbreviations

AI:	Artificial intelligence
ARM:	Artificial intelligence-based Risk Mitigation
CCMC:	CEN-CENELEC Management Centre
CDES:	Communication and Dissemination, Exploitation and Sustainability
CEN:	European Committee for Standardization
CENELEC:	European Committee for Electrotechnical Standardization
CEPOL:	European Union Agency for Law Enforcement Training
CEPOLIS:	Center of Excellence for Police and Security Research
CERIS:	Community for European Research and Innovation for Security
COBIT:	Control Objectives for Information and Related Technologies
CoU:	Community of users
CWA:	CEN Workshop Agreement
CY:	Cyprus
CYS:	Cyprus Organization for Standardization
DG:	Directorate-General
DG HOME:	Directorate-General for Migration and Home Affairs
DIN:	Deutsches Institut für Normung (German Institute for Standardization)
DPIA:	Data Protection Impact Assessment
DPO:	Data Protection Officer
DTR:	Draft Technical Report
Dx.y:	Deliverable x.y
EAIS:	European Association for the Intelligence Studies
EC3:	European Cybercrime Centre (Europol)
ECHR:	European Convention on Human Rights
ECIIA:	European Confederation of Institutes of Internal Auditing
EDMO:	European Digital Media Observatory
EDPB:	European Data Protection Board
EDPS:	European Data Protection Supervisor
EEA:	European Economic Area
EEITE:	International Conference in Electronic Engineering, Information Technology and Education

Funded by
the European Union

EIF:	European Interoperability Framework
ENISA:	European Union Agency for Cybersecurity
ENSEMBLE:	Enhanced Surveillance to Mitigate Criminal Behaviour in Europe
ETSI:	European Telecommunications Standards Institute
EU:	European Union
EUCI:	EU Classified Information
EUROPOL:	European Union Agency for Law Enforcement Cooperation
FIMI:	Foreign Information Manipulation and Interference
FRA:	European Union Agency for Fundamental Rights
FRIA:	Fundamental Rights Impact Assessment
GA:	General Agreement
GDPR:	General Data Protection Regulation
GPAI:	Global Partnership on Artificial Intelligence
GPT:	Generative Pre-trained Transformer
HITL:	Human-in-the-loop
HLEG:	High-Level Expert Group
HYBNET:	Empowering a Pan-European Network to Counter Hybrid Threats
ICCPR:	International Covenant on Civil and Political Rights
ICO:	Information Commissioner's Office
ICT:	Information and Communication Technology
IEC:	International Electrotechnical Commission
IEEE:	Institute of Electrical and Electronics Engineers
ILNAS:	Institut Luxembourgeois de la Normalisation, de l'Accréditation, de la Sécurité et qualité des produits et services
IPR:	Intellectual Property Rights
ISMS:	Information Security Management System
ISO:	International Organization for Standardization
IT:	Information Technology
ITIL:	Information Technology Infrastructure Library
JHA:	Justice and Home Affairs
JPEG:	Joint Photographic Experts Group
KPI:	Key Performance Indicators

LEA:	Law Enforcement Agency
LED:	Law Enforcement Directive
LLP:	Lifelong Learning Programme
LVU:	Lagen om vård av unga (Young Persons Care Act)
MCDA:	Multi-Criteria Decision Analysis
Mxy:	Month xy
NGO:	Non-governmental organisation
OECD:	Organisation for Economic Co-operation and Development
OSCE:	Organization for Security and Co-operation in Europe
RISE:	Research and Innovation Symposium for European Security
SAB:	Scientific Advisory Board
SDLC:	Software Development Life Cycle
SLR:	Systematic Literature Review
SME:	Small and Medium-sized Enterprise
SRE:	Security Research Event
SSH:	Social Sciences and Humanities
TC:	Technical Committee
TNO:	Netherlands Organisation for Applied Scientific Research
TR:	Technical Report
Tx.y:	Task x.y
UN:	United Nations
UNESCO:	United Nations Educational, Scientific and Cultural Organization
USA:	United States of America
WP:	Work Package
XAI:	Explainable Artificial Intelligence

FERMI CONSURTIUM

ATOS:	Atos It Solutions and Services Iberia Sl
BFP:	Belgian Federal Police
BIGS:	Brandenburgisches Institut für Gesellschaft und Sicherheit Ggmbh
BPA:	Bavarian Police Academy (Hochschule für den öffentlichen Dienst in Bayern)
CONV:	CONVERGENCE

FMI:	Finish Ministry of the Interior
GUCI:	Guardia Civil
KUL:	Katholieke Universiteit Leuven - Centre for IT & IP Law
LC:	The Lisbon Council for Economic Competitiveness asbl
IANUS:	IANUS Technologies LTD
ITML:	Information Technology for Market Leadership S.A.
INTRA:	Netcompany-Intrasoft S.A.
INOV:	INOV – Instituto de Engenharia de Sistemas e Computadores Inovação
PUCF:	Police University College (Finland) - Poliisiammattikorkeakoulu
SPA:	Swedish Police Authority - Polismyndigheten
UCSC:	Universita Cattolica Del Sacro Cuore - Transcrime
VUB:	Vrije Universiteit Brussel (VUB) - Cyber & Data Security Lab

1 Introduction

Deliverable D6.3 provides the final assessment of the Communication, Dissemination, Exploitation and Sustainability (CDES) promoting activities conducted during the full duration of the FERMI project (October 2022 – September 2025). It offers a consolidated view of how the project's outreach strategy was implemented, how it evolved over time, and what impact it achieved across the relevant stakeholder communities. The document builds on the foundations laid in D6.1 (strategic framework) and the progress update provided in D6.2 (mid-term status) and now presents a complete account of the outcomes under WP6.

The dissemination section focuses on how FERMI made its research results publicly available, free of charge, and targeted toward specific audiences. Over the course of the project, dissemination was achieved through scientific publications, including peer-reviewed journal articles, conference presentations, policy workshops, and participation in research clusters. The deliverable provides an inventory of all dissemination outputs and evaluates them against the project's original Key Performance Indicators (KPIs). It also highlights the academic and policy relevance of FERMI's contributions to key fields such as disinformation, digital risk management, criminology, and security governance.

The communication section outlines the broader awareness-raising activities that positioned FERMI as a recognisable and trusted actor within its ecosystem. These efforts targeted a general audience beyond traditional academic or institutional boundaries. The project's multi-channel communication strategy enabled FERMI to engage diverse groups in conversations about the challenges and opportunities of using AI-driven tools for disinformation mitigation and public security. The section also reflects on how the project navigated sensitive topics such as law enforcement and surveillance by fostering transparency and trust in its messaging.

The exploitation section presents an overview of how FERMI results were prepared for uptake beyond the project's lifespan, especially by public authorities, technology developers, and research institutions, while the standardisation section describes FERMI's efforts to align its technical outputs with existing frameworks and best practices. This includes contributions to ongoing discussions around interoperability, data exchange protocols, and ethical standards for the use of AI in law enforcement and public security.

Together, these four components form a coherent account of how FERMI extended its impact beyond the core research and development activities, ensuring the project's results reached the relevant communities and laid the foundation for their practical application and further development. This deliverable provides both a documentation of what was achieved and a critical reflection on how outreach and collaboration supported the project's mission.

2 Objectives and key implementation steps

Presented in D6.1 and updated in D6.2, the FERMI CDES was meticulously designed to achieve multifaceted objectives, supported by targeted implementation steps. From the outset, the strategy emphasised the development of a robust online presence through active social media engagement and the creation of an informative, user-friendly website. This initiative successfully fostered awareness not only of the project itself but also of the pressing issues it addressed, thereby nurturing a vibrant and interdisciplinary community of interest.

A second core objective of the CDES was to generate sustained traffic to the FERMI website and newsletter. This was accomplished through regular content updates, cross-promotion on social media, and consistent participation in relevant events, which collectively contributed to a steady increase in audience reach. The project effectively engaged stakeholders across domains such as digital media, content moderation, disinformation, criminology, law enforcement, and geopolitics, successfully building and maintaining a dynamic and informed community of interest.

A third and equally critical objective was the consolidation of relationships with external stakeholders, the forging of strategic partnerships, and the provision of insights to inform further research and policymaking. Over the course of the project, FERMI engaged in numerous collaborative activities, including joint webinars, workshops, and participation in EU clusters, which helped amplify project results and embed them into wider policy and research ecosystems.

While KPIs served to measure quantitative outputs, the objectives and implementation steps functioned as a qualitative guidance framework, allowing the consortium to assess whether its overarching goals had been met. Notably, many implementation steps, such as establishing an online presence or hosting public events, were both tactical actions and strategic objectives in their own right. The success of these dual-purpose efforts was demonstrated by the project's strong online footprint, stakeholder engagement, and policy relevance.

Funded by the European Union

3 Monitoring and Key Performance Indicators

Measure	KPI	Objective	Phase 1	Phase 2	Phase 3	Total	Gap
Website	N° of visitors	680	2250	1873	7533	11656	10976
Website	N° of accesses	3000	2603	2153	9306	14062	11062
Website	N° downloads of material	1020	66	177	824	1228	208
Social Media	N° of push announcements (publications)	340	205	767	985	1215	875
Social Media	N° of followers	340	264	255	298	817	535
Social Media	N° of reposts	640	737	433	321	1491	851
Social Media	N° of social media profile views	1360	4225	3545	692	8462	7102
Newsletter	N° of newsletters distributed	12	2	2	8	12	0
Communication starter pack		1	1			1	1
Update of CDES	N° of versions	3	1	1	1	3	0
Synergies	N° of similarly themed projects identified	3	5		7	12	9
Synergies	N° of workshops jointly organised with similarly themed projects	1		1	1	2	0
Publications	N° of publication in internationally referenced journals	6	2	3	4	9	3
Publications	N° of journal special issues	2		1	3	4	2
Publications	N° of publications in international magazines	6	2	4		6	0

Funded by the European Union

Publications	N° of presentation in conference	12	4	5	6	15	3
Online discussion	N° of downloads of HQ e-brochures	1000	0	243	985	1228	228
Online discussion	N° of new discussions on LinkedIn per month	1020	833	918	1149	3213	2193
Website	N° of monthly downloads	30 per month	0	243	985	1228	148
Website	N° of cumulative views of videos	1000	0	521	658	1179	179
Events	N° of small events organised (<25 participants)	5	1	1	6	8	3
Events	N° of medium size events organised (25-100 participants)	3		2	3	5	2
Events	Conversion rate in each event	40%		40%	40%	40%	0
Events	N° of final conferences organised	1			1	1	1
Events	N° of INFO days/technology showcase organised	3	1	1	1	3	0
Public policy engagement	N° of events organised	5		2	6	8	3
Public policy engagement	N° of hard copies of material for policymakers distributed per event	50		100	200	300	250

Public policy engagement	N° of policy making bodies engaged	2		3	2	5	3
Presentation of results	N° of events attended	10	8	5	7	20	10
Presentation of results	N° of events organised (100 participants)	2		1	2	3	1
Presentation of results	Participants engaged for further activities	20%			20%	20%	0
Internal outreach	N° of internal emails sent	15	6	12	12	30	15
Monitoring of CDES	N° of reports published with CDES KPIs	4		2	2	4	0
Training	N° of internal training workshop organised	1		1		1	0

Table 1 - KPIs at M36

Table 1 reports the KPI on CDES activities at M36, thus it reflects the status of KPIs in the completion of the project.

Several KPIs from FERMI's dissemination and communication strategy significantly surpassed their original objectives, highlighting the project's strong outreach and engagement efforts across multiple channels. One of the most striking results was the number of website visitors, which reached 11,665, far exceeding the initial objective of 680. Similarly, the number of website accesses totalled 14,062, more than four times the target of 3,000. This demonstrates a high level of interest in the project's content and digital presence, likely driven by coordinated communication campaigns and strong visibility in relevant communities.

Social media engagement also exceeded expectations by a wide margin. The number of social media profile views reached 8,462, vastly outperforming the goal of 1,360 by over 7,000 views. Likewise, reposts and push announcements (publications) on social media doubled their targets, with 1,491 reposts (vs. a target of 640) and 1,1215 push announcements (vs. 340). These figures show that FERMI's content was not only being seen but actively shared, an indicator of growing interest, relevance, and trust in the project's messages. The impressive growth in followers, which exceeded the target by more than 100%, further demonstrates the project's ability to build and sustain a strong digital community.

As previously mentioned in D6.2, LinkedIn discussions aim at measuring the project's presence and activity on the platform. Engagement with the project's social media content; for example, likes, shares, comments and views, might better be suited to gauge its online visibility and popularity. For this indicator, the approximate number of likes, shares and comments per month was calculated. With over 3,200 new discussions generated over the project's lifetime, this platform proved essential in fostering debate, dialogue, and professional exchange. These results collectively reflect

FERMI’s effective multi-channel strategy and underscore the project’s strong performance in engaging target audiences across web and social platforms.

Further, the number of push notifications cannot be measured as it is technically and legally impossible to send push notifications to users and subsequently track people’s mobile settings and app notifications. To calculate the approximate number of push notifications regarding the project with which users interact on social media, the following methodology was followed: According to research,[1] the median click rate for push notifications (conversion rate) is 7.8%, while overall 60% of users opt-in to receiving push notifications[2]. Considering that there are no indications about the number of notifications a platform sends about social media posts from a page, we are going to assume that around 10% of the posts were boosted this way, leading to the following equation:

$$successful\ push\ notifications \approx\ 60\%\ followers\ x\ 10\%\ posts\ x\ conversion\ rate\ (= 0.078)$$

This equation leads to the conclusion that approximate 1215 push notifications in total were successfully sent to the social media followers during all three phases.

Another indicator that required approximation is the number of views of videos. While this indicator can partially be measured with 100% accuracy (views online), there is a portion of views that remains unaccounted for when the videos are displayed on screen during certain events and expos. The number of online views in the project's YouTube channel (418) was complemented with an estimate of roughly 1/3 of the number of participants at events and expos where the project videos were showcased (215 views). Furthermore, videos from the training activities have been uploaded to the project’s website. Therefore, page views of the dedicated “Video trainings” page (546) have been counted towards that particular KPI, leading to a total of 1179 video views.

[1] Lindner, J. (2023, December 20). *Push Notification Statistics: Market Report & Data*. Gitnux. https://gitnux.org/push-notification-statistics/

[2] Dogtiev, A. (2024, February 6). *Push Notifications Statistics (2023)*. Business of apps. https://www.businessofapps.com/marketplace/push-notifications/research/push-notifications-statistics/

Funded by the European Union

4 Dissemination

Dissemination is the act of making research results publicly available free of charge, targeting those who can benefit from them, such as scientists, policymakers, and civil society. To some extent, it builds on communication, which aims to inform and engage a broader audience, and it sets the basis for exploitation, which focuses on putting results to use for commercial, societal or policy impact.[3]

4.1 Overview of dissemination

The project's dissemination activities were aimed at diverse stakeholders involved in countering disinformation-fuelled crime. By tailoring messages to civil society, policymakers, scientists/researchers and law enforcement, the project ensured that its findings supported awareness, policy development, scientific progress and operational capacity.

More specifically, the dissemination activities in the project aimed to inform civil society and policymakers about the real-life consequences of disinformation in fuelling criminal activity, reinforcing the need for coordinated policy responses. At the same time, the project sought to engage the scientific community with insights into recent advances in digital technologies for analysing disinformation campaigns, and to highlight the role of artificial intelligence in preserving the security of sensitive datasets while enabling their use in training models for crime prevention. Dissemination efforts also targeted law enforcement agencies and ministries of interior, presenting technological innovations to strengthen their ability to investigate, analyse and counter disinformation-driven crime. Through these actions, the project ensured that its results were made publicly available, relevant and usable for the diverse stakeholders positioned to act on them.

4.2 Ecosystem mapping

Information ecosystems are the environments in which individuals, institutions and technologies interact to produce, circulate and interpret meaning.[4] In the context of disinformation, particularly when it fuels real-life crime, understanding the structure and dynamics of these ecosystems becomes essential. Disinformation spreads not in a vacuum, but through specific actors, channels, incentives and vulnerabilities. Mapping these systems helps identify who is exposed, who amplifies falsehoods (knowingly or not), and who holds potential to intervene.

In this project, the ecosystem mapping reported in Deliverable D6.2 was undertaken to understand the breadth and structure of the information environment surrounding the issue of disinformation-driven crime and to lay the ground for the target group-oriented dissemination activities described below. This was in line with the GA's requirement to ensure that "[a] multi-actor collaboration framework will be developed and rolled out based on open innovation, combining knowledge and experience of different actors in the FERMI ecosystem."[5] While the resulting map captured a wide array of actors based on "generic stakeholder groupings", including civil society, law enforcement, the media, researchers, and policymakers, the project's dissemination efforts were necessarily more narrowly targeted, focusing on those best positioned to engage with the sensitive and security-related nature of the topic. The strategic rationale for prioritisation and the limitations of the project's outreach notwithstanding, the mapping exercise enabled the consortium to align ecosystem insight with the design and scope of project activities, which is a critical consideration for future initiatives tackling similarly complex and high-stakes issues. This includes laying the ground for FERMI's exploitation, which includes a very granular distinction between different business model-relevant players (as explained in this deliverable's exploitation section, see chapter 6, and D6.4).

[3] Read here a more comprehensive overview of communication, dissemination and exploitation: https://op.europa.eu/en/publication-detail/-/publication/58ad3394-0a63-11ee-b12e-01aa75ed71a1/language-en

[4] To understand more about information ecosystems see https://carnegieendowment.org/research/2025/02/assessing-national-information-ecosystems?lang=en

[5] Grant Agreement, Part A, p.14.

Eventually, the importance of an ecosystem perspective is underscored by the complexity of the topic, as reflected in the latest policy developments, such as the European Democracy Shield. The European Democracy Shield is an EU initiative proposed in 2025 to reinforce democratic resilience by countering foreign information manipulation and interference, safeguarding the integrity of elections, supporting independent journalism, and protecting civil society actors.[6] These goals are mirrored in FERMI through its ecosystem perspective: rather than focusing solely on technical detection of disinformation, the project also promotes media literacy, engages with diverse stakeholders including policymakers and law enforcement, and ensures that all interventions are grounded in ethical and legal standards.

Today digital and physical worlds are entangled in a complex ecosystem and any intervention, be it on the policy or on the operational front, demands a solid contextualisation and analysis of this complex ecosystem and it must avoid at all costs a one-size-fits-all approach and a narrow vision of the phenomenon of disinformation or crime. In this context, the ecosystem perspective aims to ensure that the results and impact of FERMI can be contextualised in this sense, highlighting its achievements while refraining from solutionist assessments.

4.3 Short report on the dissemination for LEAs and policymakers

The FERMI project has presented its insights and products at numerous policy-relevant gatherings. In this regard, the meetings of the Community for European Research and Innovation for Security (CERIS) are a case in point. As explained on the Directorate-General for Migration and Home Affairs (DG HOME)'s website, CERIS and its network are aimed "to facilitate interactions within the security research community and users of research outputs." More specifically, "in 2014 the Commission established the Community of Users for Safe, Secure and Resilient Societies (CoU), which gathered around 1,500 registered stakeholders (policy makers, end-users, academia, industry and civil society) and regularly held thematic events with the security research community."

Accordingly, CERIS clearly and greatly exceeds the boundaries of LEAs and policymakers. That being said, it is no coincidence that policymakers and end-users are mentioned before the other target groups. Policymakers are not only intimately involved in the CERIS community, CERIS is run by policymakers that coordinate all of its activities. As further explained on the website, "[t]argeted Thematic Workshops [...] will be chaired by European Commission representatives and will cover key features of the civil security for society research programme." Moreover, end-users in the context of the CERIS' meetings mostly attended by the FERMI consortium have first and foremost been LEAs. As a project funded under Fighting Crime and Terrorism, which is also one of the subgroups of CERIS, the FERMI representatives have largely interacted with LEA end-users entrusted with crime- and terrorism-fighting.

In May 2023, a representative from the Bavarian Police Academy (BPA) alongside additional consortium members attended the CERIS workshop on "disinformation, fake news and hate speech", where he introduced the CERIS audience to FERMI, which was fairly new back in the day. Numerous BPA experts also represented FERMI at the CERIS Annual Event 2023 – Fighting Crime and Terrorism/Resilient Infrastructure taking place on 14 and 15 December in Brussels.

On 24-25 September 2024, a representative from BPA participated in a panel on online harms at CERIS' "Annual event on research for fighting crime and terrorism." Aside from online harms (the first day's focus), the meeting also explored the involvement of practitioners in security research.

On 15 May 2025, BPA and KU Leuven attended a CERIS meeting on "Countering online identity theft and fake information." The KU Leuven representative took part in a panel on "Strategies for Mitigating Online Identity Theft and Disinformation", while BPA presented the FERMI platform.[7]

[6] In-depth analysis of the European Democracy Shield initiative can be found here: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI%282025%29775835

[7] ITML took the lead in preparing all the necessary groundwork to create engaging demonstration videos of the FERMI platform in action. These demo videos were designed to clearly showcase the platform's key features and benefits, making them effective

BPA also participated in the subsequent CERIS event "Disaster Resilience Days," held from 19-21 May 2025 in Brussels, where the audience was introduced to the FERMI project and platform in a panel on "Societal Resilience - Engage, inform, empower co-creating risk communication strategies."

At the invitation of DG HOME, BPA also attended the EU Innovation Hub for Internal Security on 05 June 2024. BPA had the opportunity to give an overview of the FERMI project's status and key results and discuss the possibility of keeping EU law enforcement agencies in the loop about future project developments.

The FERMI consortium also disseminated its products at the 2025 Security Research Event (called "Boosting security through EU-based innovation") that took place in Warsaw from 24-25 June and was organised by DG HOME. FERMI had its own stand in a shared booth alongside our sister project VIGILANT. Representatives from ITML and BPA were regularly available at FERMI's stand to answer questions and interact with stakeholders. Further FERMI colleagues present at the SRE, who supported them and occasionally joined the stand to provide further explanations included experts from INTRA, IANUS and PUCF.

Apart from the immediate EU framework, BPA presented the FERMI project's security management record on a panel on best practices of security management in EU projects on 06 November 2024 at a major gathering of science advisors on exchanging experiences in managing Horizon Europe projects ("Erfahrungsaustausch zu Horizont Europa") in Cologne, Germany, organised by the German Federal Ministry of Education and Research, its EU office, and the Working Group of EU science advisors at German universities.

Moreover, on 22 January 2025, BPA, alongside radicalisation experts from the Bavarian State Criminal Police Office as well as practising psychotherapists, participated in a meeting on the impact of social media and smartphones on the younger generation. The experts agreed that young people are particularly at risk, as they overwhelmingly get their news from the digital world and are more susceptible to radical messages and manipulation due to a lack of media literacy and emotional vulnerability.

LEA partners within the consortium organised dedicated training activities for their peers, which are presented in detail below.

4.3.1 Thematic Seminars for Experts, Practitioners and Policymakers

BIGS, within its remit as an independent security policy research institute has organised two targeted seminars, integrated in its Event Series called "PizzaSeminar". Those events are targeted at related experts, policymakers and practitioners and have each time a specific focus. They attract 15-40 participants and take place ad hoc, depending on the topic.

In that context, two PizzaSeminars have been organised:

a) "Disinformation, Political Extremism and the Costs – An econometric analysis within the FERMI project", 11 June 2025, and

b) "The Effect of Terrorism on Public Attitudes and Individual Well-Being in Great Britain", 18 September 2025.

Furthermore, BIGS in collaboration with GUCI, has staged a "Workshop on Socio-economic Impacts of Political Extremism", 02 April 2025 in Madrid, targeted at LEA, ministerial, and intelligence services representatives.

tools for presentation at all events and training sessions in which the FERMI consortium participated or which it organised (e.g. CERIS, SRE), thereby enhancing the visibility of the project. With this material, the FERMI project was effectively promoted to a wide audience, particularly within industry-focused settings.

4.3.2 Enganging EDMO

On 18 June 2025, the FERMI consortium teamed up with the European Digital Media Observatory (EDMO) to host the webinar "Community Resilience in Action: How EDMO and FERMI Fight Disinformation." The online event brought together researchers, journalists and law-enforcement specialists to learn about the FERMI tools designed to spot and minimise crime-related disinformation.

The technical walk-through, led by INOV (focused on the Community Resilience Management Modeler), demonstrated how the Modeler guides law-enforcement agencies toward proportionate, impact-focused responses, after receiving a flag for high-stake disinformation-related crimes and weighting response options with a multi-criteria decision analysis and prioritising the most effective counter-measures.

Presentations were held by EDMO's Secretary-General and Coordinator, underlining the Observatory's mission to connect cutting-edge research with frontline practitioners and the Coordinator of the Mediterranean Digital Media Observatory (MedDMO) and member of EDMO's Management Committee, highlighting the growing need for cross-border collaboration as disinformation campaigns become more sophisticated and regionally targeted.

4.3.3 Short report on the dissemination activities with standardisation bodies and standardisation experts

The GA places a special emphasis on the outreach to standardisation bodies, which is even at the core of a separate WP6 task (T6.5). Accordingly, the FERMI consortium has made targeted dissemination activities aimed to standardisation bodies and standardisation experts, which are recapped below, a clear priority.

4.3.3.1 Dissemination to standardisation bodies

The GA requires the consortium to establish "[a]t least 3 links with existing standardization bodies" (KPI 5.2).[8] Accordingly, bilateral online meetings were held between a representative of IANUS and a series of standardisation bodies, as outlined below. During these meetings, the FERMI project was presented, in particular Task 6.5 and its objectives on standardisation. Substantive input was sought from the standardisation officials who took part in these meetings, which was incorporated in the operational standards set out in sections 7.2 and 7.3.

(a) A representative of the CYS - Cyprus Organisation for Standardisation on 22 May 2025 and 17 June 2025.[9]

(b) A representative of the DIN – German Standardisation Body on 06 June 2025,

(c) Two representatives of the SIS – Swedish Standards Institute on 12 June 2025,

(d) A representative of the SFS – Finnish Standards Association on 16 June 2025,

(e) A representative of the AFNOR Digital Department, Association Française de Normalisation, who is also Secretary of CEN/TC 391, on 07 August 2025.

[8] Grant Agreement, Part B, p.3.

[9] During the second meeting with the CYS on 17/06/2025, it was asked whether it would be possible to put forward the draft operational standard on disinformation as a CEN Workshop Agreement, but the response of CYS was negative, putting forward budget and time constraints. It has accordingly been recommended in the FERMI White Paper that the operational standard on disinformation developed as part of Task 6.5 be adopted as CEN Workshop Agreements in the future.

Funded by the European Union

Outreach to five different standardisation bodies is not just in line with but even exceeds the GA's requirement to coordinate with three such organisations. Outreach efforts were also made with respect to other standardisation organisations, but no meetings were set up.[10]

The IANUS representative also participated as a delegate of the Cyprus Organisation for Standardisation at the European Telecommunications Standards Institute ("ETSI") meeting CYBER(25)43b021, which took place on 24 July 2025. During that meeting, she presented her disinformation-related input and comments to a technical report that is currently being developed by ETSI entitled "Cyber Security; Human-to-Human Online Preventative Security; H2H – OPS" (the full set of comments presented to the ETSI is set out in Annex C below). Additionally outreach efforts were also made to other European and international standardisation organisations.

4.3.3.2 Dissemination to standardisation experts

Bilateral online meetings were also held between IANUS and a series of standardisation experts, both from academia and the private sector. During those meetings, an overview of the FERMI project was presented, and in specific the work on standardisation undertaken for Task 6.5. The substantive input and feedback of these experts, who are listed as follows, was sought on specific standardisation points relating to disinformation, which was incorporated in the operational standards set out in sections 7.2 and 7.3 of the present deliverable.

(a) An expert from Utrecht University on 22 May 2025,

(b) An expert from Tilburg University on 12 June 2025,

(c) An expert from Responsible AI Ethicist at Accenture, on 13 June 2025,

(d) Standardisation Expert at HS Booster on 03 July 2025,

(e) Senior Research Analyst at Trust-IT Services on 23 July 2025,

(f) Standardisation Expert at HS Booster on 28 July 2025.

4.3.3.3 Webinar on operational standards regarding disinformation

An online webinar took place on Thursday 11 September 2025, between 10:00 – 11:30 CET concerning the activities of task 6.5 of FERMI, including presentations on the three operational standards that were developed as part of the FERMI project (see the final agenda below). 17 individuals participated not only from internal partners of the FERMI consortium, but also external partners including the Cyprus Organisation for Standardisation and the French Association for Standardisation. At the end of the presentations, feedback was sought from the LEAs present with respect to the 3 operational standards.

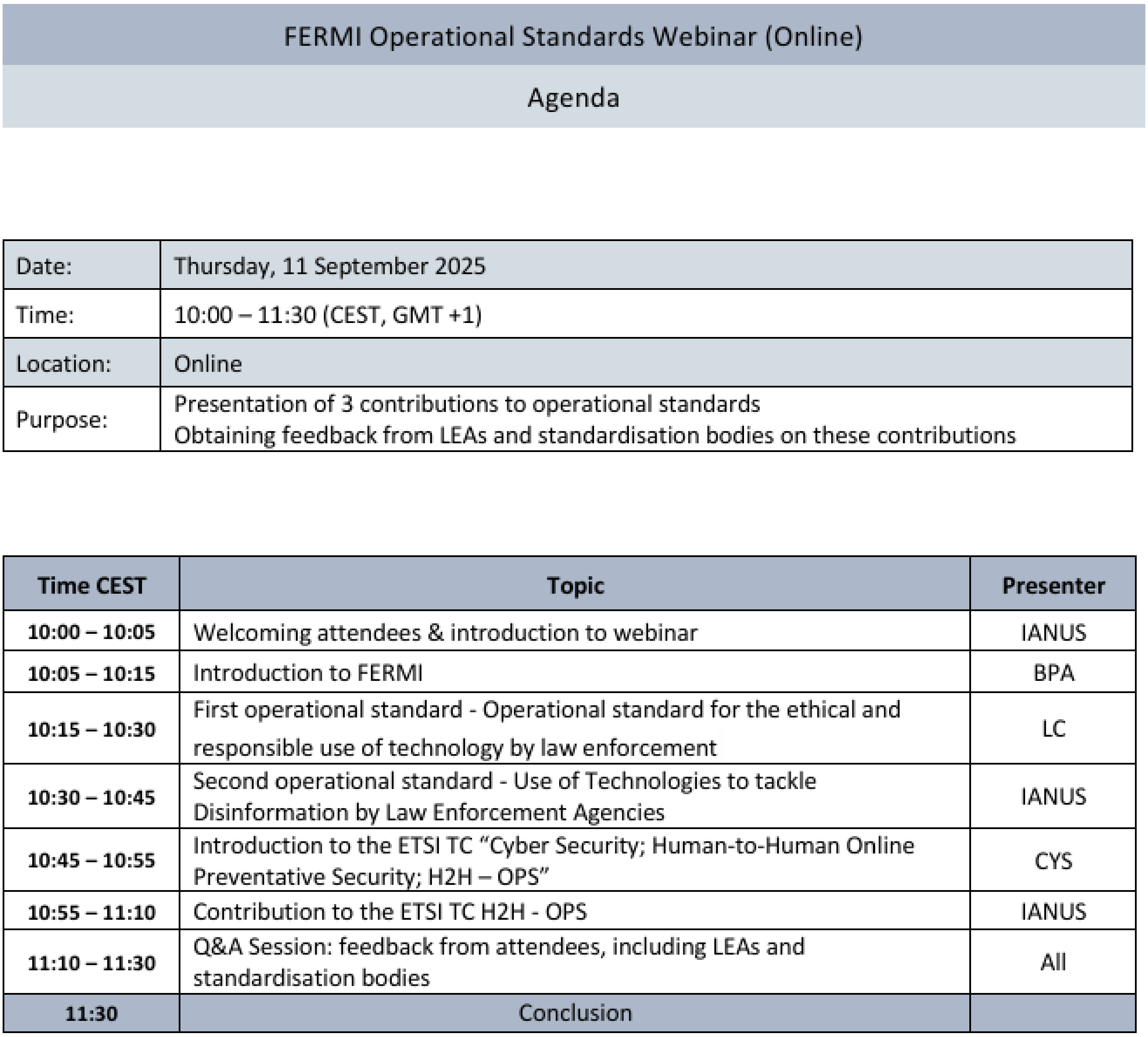
[10] Furthermore, the following standardisation bodies have been contacted, but no responses were received:

(a) ILNAS – Luxembourg Institute for Standardisation (emails sent on 17/06/2025 and 30/06/2025);

(b) NEN – Royal Netherlands Standardisation Institute (emails sent on 30/05/2025 and 30/06/2025). The representativr responded but it was not possible to set up a meeting.

The following standardisation institute responded negatively to our request to have a meeting:

(a) NBN – Bureau de Normalisation of Belgium (response received on 02/06/2025 that it was not possible to move forward due to the priorities set for 2025 and very limited available staff).



FERMI Operational Standards Webinar (Online)

Agenda

Date:	Thursday, 11 September 2025
Time:	10:00 – 11:30 (CEST, GMT +1)
Location:	Online
Purpose:	Presentation of 3 contributions to operational standards Obtaining feedback from LEAs and standardisation bodies on these contributions

Time CEST	Topic	Presenter
10:00 – 10:05	Welcoming attendees & introduction to webinar	IANUS
10:05 – 10:15	Introduction to FERMI	BPA
10:15 – 10:30	First operational standard - Operational standard for the ethical and responsible use of technology by law enforcement	LC
10:30 – 10:45	Second operational standard - Use of Technologies to tackle Disinformation by Law Enforcement Agencies	IANUS
10:45 – 10:55	Introduction to the ETSI TC "Cyber Security; Human-to-Human Online Preventative Security; H2H – OPS"	CYS
10:55 – 11:10	Contribution to the ETSI TC H2H - OPS	IANUS
11:10 – 11:30	Q&A Session: feedback from attendees, including LEAs and standardisation bodies	All
11:30	Conclusion	

Figure 1 – FERMI Operational Standards webinar agenda

4.4 Short report on the scientific and technical dissemination

In FERMI's endeavours to disseminate scientific and technical advancements, the project actively participated in globally significant conferences, particularly within the fields of disinformation, criminology, and digital governance. These conferences provide valuable platforms for sharing FERMI's research findings and insights, facilitating dialogue, and fostering networks with experts, practitioners, and policymakers.

As far as the criminological realm is concerned, the FERMI consortium presented the project at the American Society of Criminology Annual Meeting 2024, the Annual Conference of the European Society of Criminology 2025, the 2023 Institute of Electrical and Electronics Engineers (IEEE) International Conference on Big Data (BigData), and the Fifth International Conference in Electronic Engineering, Information Technology & Education (EEITE '24). Presentations on ethics and data protection were held at Digital Legal Talks 2023, Techno-Legal Challenges of Data Scraping 2023, and the Interdisciplinary Seminar on Fake News & Disinformation, representing the Cyber and Data Security Lab (VUB-

Funded by the European Union

CDSL). Overall project presentations were held at numerous of the above-mentioned CERIS gatherings and on further networking events.

Furthermore, FERMI's dedication to advancing scientific knowledge is evident in its contributions to interdisciplinary studies on disinformation, crime, risk management, and decision science. By presenting project insights, amongst other things in the form of peer-reviewed papers, at esteemed academic venues and policy forums, FERMI aims to contribute meaningfully to scholarly discourse while supporting evidence-based strategies to address complex technical and societal challenges.

FERMI maintains an active presence on Zenodo through its dedicated community, "FERMI EU Project," where peer-reviewed publications are saved and made publicly accessible. This ensures open access to FERMI's scientific outputs, supporting reproducibility, wide dissemination, and long-term archiving. The "Publications" section of the FERMI website links directly to this Zenodo community for detailed information. Each record in Zenodo includes bibliographic metadata, abstracts, publication venue, author lists, and for many links to full text or external pages.

The consortium's far-reaching peer-reviewed contributions include technical articles on topics such as "Conspiracy to Commit: Information Pollution, Artificial Intelligence, and Real-World Hate Crime,"[11] "Informative (Dis)information: Exploring the Correlation Between Social Media Disinformation Campaigns and Real-World Criminal Activity"[12] and "The Nexus Between Big Data Analytics and the Proliferation of Fake News as a Precursor to Online and Offline Criminal Activities."[13]

As the titles imply, a key focus of most of these publications was on examining how emerging technologies within the FERMI platform can help understand the relationship between disinformation and criminal activity. In this regard, evidence-based analyses of how coordinated online campaigns may trigger real-world criminal outcomes were conducted, contributing valuable insights to the field of disinformation and criminology. Amongst other things, the impact of 36 distinct conspiracy theories was examined by using AI-enabled analysis tools to capture the broader dynamics of disinformation-driven radicalisation and hate crime. Another peer-reviewed technical paper called "Modeling Disinformation Spread in Social Networks: Phase Transitions and Mean-Field Analysis"[14] explored the modelling of disinformation spread in social networks through the lens of statistical mechanics providing a deeper understanding of critical thresholds and collective opinion dynamics.

A set of further peer-reviewed articles addressed INOV's work on community resilience, especially in terms of risk management and counter-measures against disinformation. These publications included pieces on "Fake News: a conceptual model for risk management,"[15] "Using MCDA to select countermeasures against fake news,"[16]

[11] Aziani, A., Lo Giudice, M. V., Yazdi, A.S. "Conspiracy to Commit: Information Pollution, Artificial Intelligence, and Real-World Hate Crime," European Journal on Criminal Policy and Research (2025). https://zenodo.org/records/15719116

[12] Lo Giudice, M.V., Yazdi, A.S., Aziani, A., Evangelatos, S., Gousetis, N., Nikolopoulos, C. "Informative (Dis)information: Exploring the Correlation Between Social Media Disinformation Campaigns and Real-World Criminal Activity," 2024 5th International Conference in Electronic Engineering, Information Technology & Education (EEITE), Chania, Greece (2024). https://zenodo.org/records/13759707

[13] Evangelatos, S., Papadakis, T., Gousetis, N., Nikolopoulos, C.D., Troulitaki, P., Dimakopoulos, N., Bravos, G., Lo Giudice, M.V., Shadman, A., Aziani, A. "The Nexus Between Big Data Analytics and the Proliferation of Fake News as a Precursor to Online and Offline Criminal Activities," 2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy (2023). https://zenodo.org/records/10948598

[14] Evangelatos, S., Veroni, E., Efthymiou, V., Nikolopoulos, C. "Modeling Disinformation Spread in Social Networks: Phase Transitions and Mean-Field Analysis," ACM Transactions on the Web (2025). https://zenodo.org/records/16410586

[15] Varela da Costa, J., Bogea Gomes, S. & Mira da Silva, M., "Fake News: a conceptual model for risk management," Humanit Soc Sci Commun 11, 625 (2024). https://zenodo.org/records/11354424

[16] Varela da Costa, J., Dongo, D.F., Mira da Silva, M, "Using MCDA to select countermeasures against fake news," Journal of Information, Communication and Ethics in Society 23 (1) (2025). https://doi.org/10.1108/JICES-07-2024-0089

"Countermeasures against fake news: a Delphi study"[17] and "Fake News and risk management: a systematic literature review."[18]

In those papers, a disinformation terminology was laid out, highlighting vital concepts and exploring their correlations.[19] Based on a systematic literature review (SLR) a model that investigates the intersection of fake news, risk, and risk management was proposed. Employing Design Science Research as the primary methodology, it introduced a conceptual model to mitigate fake news-related risks in specific communities.[20] Multi-Criteria Decision Analysis (MCDA) was applied to evaluate and prioritise countermeasures against disinformation. Key counter-measures identified include monitoring disinformation accounts, legislative reforms, public awareness campaigns, and educational programmes to promote media literacy.

A final peer-reviewed article analysed the "Moderation of illegal content and social media scraping"[21] and explored how privacy and data protection frameworks apply to the scraping of publicly available data by law enforcement authorities. The study provided an in-depth legal analysis of digital policing practices, highlighting potential constraints and tensions with fundamental rights. It offered policy-relevant insights into how public authorities can uphold democratic principles while navigating the legal complexities of online surveillance and content moderation. Several non-peer-reviewed articles were published, too, including in police journals such as "Campus Polizei. Magazin der Hochschule für den öffentlichen Dienst in Bayern, Fachbereich Polizei."[22]

In early summer 2025, a BPA representative delivered a guest lecture on the use of new technical approaches in applied security research at the Faculty of Electrical Engineering and Information Technology, Technical University of Applied Sciences Amberg-Weiden before twenty-two students as part of the Master Course – Module: Information Security and Functional Safety. BPA was also interviewed by Net4Society, a research network conducting an analysis on how to integrate social sciences and humanities into Horizon Europe.

4.5 Short report on the dissemination for the private sector

To effectively disseminate the FERMI project to the private sector, a focused strategy was implemented by ITML across business networking events, digital blogs and social media. The project was showcased at SMEs (ie. ATC, GNT, AEGIS) and Non-for-Profit organisations (ie. SYMPLEXIS), where tailored material highlighted FERMI's relevance for media integrity and civil security. Concurrently, targeted articles were published on ITML's website such as a dedicated article on its corporate blog titled "How does FERMI work on combating D&FN toward safeguarding the

[17] Varela da Costa, J., Mira da Silva, M, "Countermeasures against fake news: a Delphi study," Transforming Government: People, Process and Policy 19 (2) (2025). https://doi.org/10.1108/TG-10-2024-0258

[18] Varela da Costa, J., Fernandes, A., & Mira da Silva, M., "Fake news and risk management: a systematic literature review," Journal of Risk Research, 27(12) (2024). https://doi.org/10.1080/13669877.2025.2466530

[19] The review defines various fake news approaches and related concepts: Impact, Context, Agent, Verifiability, Medium, Event, Content, Source, and Intention. The study's implications extend to incorporating fake news concepts into digital risk management, information security, and risk frameworks. This exploration underscores the societal risks of fake news and emphasises the need for a resilient framework to address the digital risks associated with disinformation. Future research should build upon these findings to develop a comprehensive model for digital risk mitigation concerning disinformation

[20] More information on the model: The model uses ArchiMate to depict a community as an organisational entity, exemplifying its practicality through a fake news instance from the Central European Digital Media Observatory. The research undergoes rigorous evaluation using the Bunge-Wand-Weber Model, ensuring its consistency and value to the scientific community. This evaluation formalises the proposed conceptual model, offering a structured framework systematically mapping fake news concepts to mitigate associated risks and disinformation.

[21] Giglio, F., "Moderazione di contenuti illegali e social media scraping: Vincoli in materia di privacy e protezione dei dati nel trattamento dei dati disponibili al pubblico da parte delle forze dell'ordine," i-lex. 16(2) (2023). doi: 10.6092/issn.1825-1927/18870.

[22] See Nitsch, H., FERMI. Campus Polizei. Magazin der Hochschule für den öffentlichen Dienst in Bayern, Fachbereich Polizei (2023). https://zenodo.org/records/11440040

Digital Trust?" (25 July 2024). The post introduced FERMI's scope and its role in enhancing digital trust, contributing to awareness raising among the general public. Also, an English-language social media campaign on LinkedIn reached decision-makers in information and communication technology (ICT), media, security and communications firms, using project success stories to spark interest and generate leads for further collaboration.

4.6 Short report on the dissemination to the general public

This section will focus on the dissemination of FERMI activities to general public through the organisation of events. To maximise outreach and ensure broad public engagement, the consortium implemented a coordinated, multi-channel communication plan across the FERMI website, LinkedIn, and partners' owned channels. Messaging focused on citizen empowerment, technological approaches to counter disinformation, and the trust implications for institutions and platforms. All posts included clear calls-to-action with direct links to the consent form and registration page; visual identity was harmonised through branded speaker cards, countdown assets, and post templates.

Campaign approach (all webinars)

- **Pre-event:** save-the-date, agenda highlights, and speaker reveals; targeted tagging of institutions and speakers to extend organic reach; UTM-tagged links for basic performance tracking.
- **During event:** live coverage with key quotes and prompts to submit questions/polls; consistent event hashtags to support discoverability.
- **Post-event:** thank-you posts, access to slides/recordings, and short recap pieces to maintain engagement and support asynchronous viewing.

Event-specific dissemination

- **A dive into the societal landscape of disinformation – Balancing between Law Enforcement and Fundamental Rights to Increase Digital Trust** (23 February 2024, led by CONV): multi-wave LinkedIn campaign with partner resharing to legal, LEA, and civil-society communities; emphasis on the balance between enforcement and rights to build digital trust.
- **Digital Trust in Action: Technological Approaches and Citizen Empowerment to Combat Disinformation** (4 December 2024, led by CONV): consortium-wide push via FERMI website and LinkedIn; promotional visuals and speaker highlights; a dedicated recap article published shortly after the event summarising themes, expert inputs, and outcomes.
- **How disinformation evolves: Narratives, Digital influence and Trust** (10 September 2025, led by CONV): promotion mirrored the previous approach, targeting practitioners and researchers on evolving narratives and influence tactics; strong partner amplification and post-event recap with recording.[23]

Dissemination activities achieved broad visibility across relevant stakeholder networks (policy, research, LEAs, media literacy, and civil society). Engagement concentrated around speaker reveals and live coverage posts, with partner

[23] Attendance overview across the three webinars: (i) A dive into the societal landscape of disinformation – Balancing between Law Enforcement and Fundamental Rights to Increase Digital Trust (23 Feb 2024): 25 participants; (ii) Digital Trust in Action: Technological Approaches and Citizen Empowerment to Combat Disinformation (4 Dec 2024): 37 registrations, corresponding to 36 unique attendees after cross-verification; (iii) How disinformation evolves: Narratives, digital influence and Trust (10 Sep 2025): 34 participants. In all cases, the audience comprised members of the general public, external experts, representatives of law enforcement agencies, and participants from organisations within the FERMI consortium. A full overview of event structure, learning outcomes, evaluation results, and associated training materials is provided in the relevant sections of this deliverable (e.g., Section 4.8 for the December 2024 webinar).

resharing significantly extending reach beyond project followers. Post-event assets supported continued traffic to recordings and recaps.

Further dissemination to all target groups and the general public alike included a BPA presentation on the dangers and ramifications of extremism-furthering disinformation and foreign information manipulation and suppression during BPA's open house on 16 November 2024. Two experts presented key insights into the subject matter gained by the FERMI and RESONANT projects. BPA's open house was attended by more than 1,500 people.

In June 2024, FERMI together with representatives from the POBREBEL and ELECTRUST projects, was invited to speak on the CORDIScovery podcast of the European Commission. All three projects use cutting-edge techniques to enhance their research and ultimately strengthen the foundations of our democratic societies. The discussion touched on the impact of digital technologies on societies, populism and the role of law enforcement in addressing online disinformation.[24]

In April 2025 ITML actively participated in the "Training and Platform Demonstration and Disinformation Symposium" at Madrid by demonstrating the FERMI platform and how it aims to contribute on LEAs efforts on tackling extremism-induced disinformation campaigns**.**

4.6.1 The FERMI Final Conference

The FERMI project held its Final Conference on 16 September 2025 at Humboldt University of Berlin and online, with over 100 registered participants from across Europe and beyond. This event marked the conclusion of three years of collaborative research and innovation aimed at understanding the disinformation threat landscape and developing technology-driven and data-enabled solutions to address it.

The conference opened with reflections from representatives of FERMI's coordinator BPA and the host organisation, BIGS. The opening remarks outlined the project's journey, highlighting its interdisciplinary approach, major milestones, and practical contributions to European resilience against disinformation.

Session 1 addressed the emerging disinformation threat landscape in Europe. Speakers from civil society organisations, research institutions, and national authorities presented insights on foreign influence operations, the intersection between disinformation and domestic extremism, and the role of media literacy in enhancing democratic resilience. The session also featured an overview of the outcomes of the FERMI webinar series on digital trust and public awareness. Discussions with the online and in person audience underscored the persistent and evolving nature of the threat.

Session 2 showcased technology-driven solutions to counter hybrid threats. Representatives from Horizon Europe projects participating in the Hybrid Threats cluster, including FERMI, vera.ai, and ATHENA, presented their respective platforms, discussing methods to detect, analyse, and counter foreign information manipulation and interference (FIMI). The session highlighted the critical importance of ethical AI, transparency, and operational utility in developing tools that can be adopted by law enforcement agencies and public authorities.

Session 3 explored the theme of data as a key enabler in disinformation mitigation. Project partners presented FERMI's use of social media data for training AI systems, integration of crime and disinformation datasets, and access to politically motivated crime data for socio-economic analysis. Discussions focused on data governance challenges, privacy concerns, and the need for harmonised methodologies when working with sensitive or large-scale datasets.

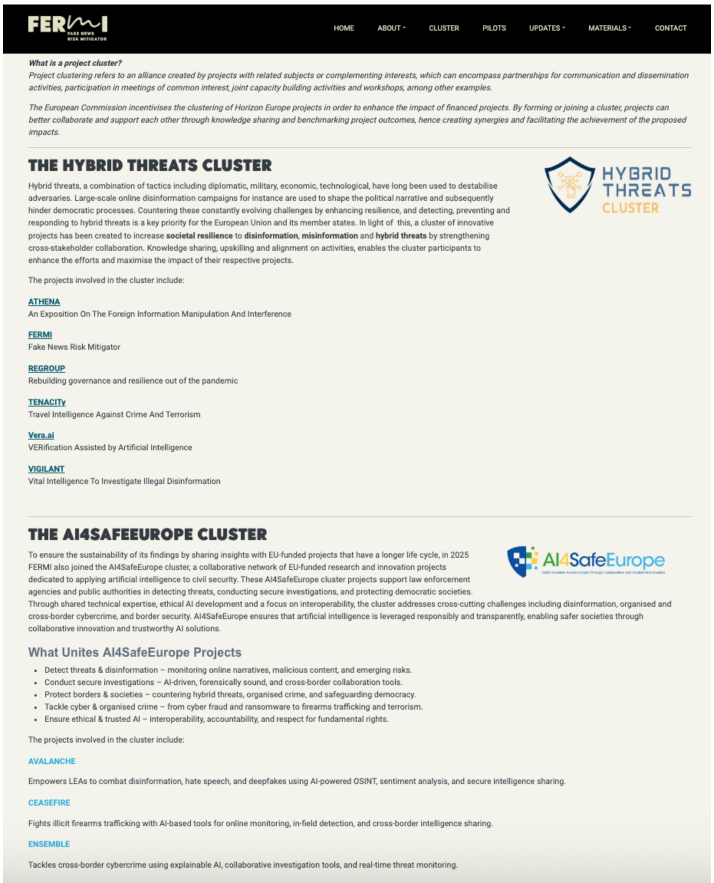
Recordings of all three sessions are available on the FERMI YouTube channel,[25] offering ongoing access to the presentations, discussions, and key takeaways for participants and the wider public.

[24] Recordings of the podcast are available on the CORDIS website of the European Commission at Democracy – a right worth defending | News | CORDIS | European Commission.

[25] https://www.youtube.com/@fermi-project

4.7 Synergies with other initiatives and projects

According to the description of T6.2, FERMI is required to join forces with other projects to maximise dissemination. More specifically, the GA states that "dissemination activities will include collaboration across projects considering what has been done and what is already available. Specific focus will be given in establishing solid links with successful proposals from the topics HORIZON-CL2-DEMOCRACY-2021-01-08 (Politics and governance in a post-pandemic world), HORIZON-CL2-DEMOCRACY-2022-01-06 (Politics and the impact of online social networks and new media) and HORIZON-CL4-2021-HUMAN-01-27 (AI to fight disinformation)."[26]



What is a project cluster?

Project clustering refers to an alliance created by projects with related subjects or complementing interests, which can encompass partnerships for communication and dissemination activities, participation in meetings of common interest, joint capacity building activities and workshops, among other examples.

The European Commission incentivises the clustering of Horizon Europe projects in order to enhance the impact of financed projects. By forming or joining a cluster, projects can better collaborate and support each other through knowledge sharing and benchmarking project outcomes, hence creating synergies and facilitating the achievement of the proposed impacts.

THE HYBRID THREATS CLUSTER

Hybrid threats, a combination of tactics including diplomatic, military, economic, technological, have long been used to destabilise adversaries. Large-scale online disinformation campaigns for instance are used to shape the political narrative and subsequently hinder democratic processes. Countering these constantly evolving challenges by enhancing resilience, and detecting, preventing and responding to hybrid threats is a key priority for the European Union and its member states. In light of this, a cluster of innovative projects has been created to increase **societal resilience** to **disinformation**, **misinformation** and **hybrid threats** by strengthening cross-stakeholder collaboration. Knowledge sharing, upskilling and alignment on activities, enables the cluster participants to enhance the efforts and maximise the impact of their respective projects.

The projects involved in the cluster include:

ATHENA
An Exposition On The Foreign Information Manipulation And Interference

FERMI
Fake News Risk Mitigator

REGROUP
Rebuilding governance and resilience out of the pandemic

TENACITy
Travel Intelligence Against Crime And Terrorism

Vera.ai
VERification Assisted by Artificial Intelligence

VIGILANT
Vital Intelligence To Investigate Illegal Disinformation

THE AI4SAFEEUROPE CLUSTER

To ensure the sustainability of its findings by sharing insights with EU-funded projects that have a longer life cycle, in 2025 FERMI also joined the AI4SafeEurope cluster, a collaborative network of EU-funded research and innovation projects dedicated to applying artificial intelligence to civil security. These AI4SafeEurope cluster projects support law enforcement agencies and public authorities in detecting threats, conducting secure investigations, and protecting democratic societies. Through shared technical expertise, ethical AI development and a focus on interoperability, the cluster addresses cross-cutting challenges including disinformation, organised and cross-border cybercrime, and border security. AI4SafeEurope ensures that artificial intelligence is leveraged responsibly and transparently, enabling safer societies through collaborative innovation and trustworthy AI solutions.

What Unites AI4SafeEurope Projects

- Detect threats & disinformation – monitoring online narratives, malicious content, and emerging risks.
- Conduct secure investigations – AI-driven, forensically sound, and cross-border collaboration tools.
- Protect borders & societies – countering hybrid threats, organised crime, and safeguarding democracy.
- Tackle cyber & organised crime – from cyber fraud and ransomware to firearms trafficking and terrorism.
- Ensure ethical & trusted AI – interoperability, accountability, and respect for fundamental rights.

The projects involved in the cluster include:

AVALANCHE

Empowers LEAs to combat disinformation, hate speech, and deepfakes using AI-powered OSINT, sentiment analysis, and secure intelligence sharing.

CEASEFIRE

Fights illicit firearms trafficking with AI-based tools for online monitoring, in-field detection, and cross-border intelligence sharing.

ENSEMBLE

Tackles cross-border cybercrime using explainable AI, collaborative investigation tools, and real-time threat monitoring.

Figure 2 - Screenshot of Clusters page in FERMI website

Accordingly, FERMI has created the Hybrid Threats Cluster, a collaborative alliance of Horizon Europe projects united to bolster societal resilience against disinformation, misinformation, and hybrid threats. This cluster enables enhanced communication, joint capacity-building, knowledge sharing, and coordinated dissemination activities among projects with complementary aims. Other participating initiatives include ATHENA, REGROUP, TENACITY, Vera.ai, and VIGILANT. Through this cluster, FERMI aligns its technological tools, training efforts, and outreach strategies with peer initiatives, maximising impact and enabling synergies in addressing complex security challenges at the EU level.

Among the actions that took place within the cluster, on 11 April 2024 FERMI partnered with the VIGILANT project to host a joint webinar in which FERMI project partner INOV (Instituto de Engenharia de Sistemas e Computadores Inovacao) presented their research on the Socioeconomic Disinformation Watch. VIGILANT partner TNO (Netherlands Organisation for Applied Scientific Research) presented their work on the social drivers of disinformation. Furthermore, on 21 May 2024 FERMI and the VIGILANT project co-wrote the article "Protecting European Elections from

[26] Grant Agreement, Part A, p.14.

Funded by the European Union

Disinformation with Horizon Europe", which was featured on both projects' websites,[27] as well as in the news page of the DG HOME's website.[28] Vera.ai and ATHENA were also presented during the FERMI Final Conference (see section 4.6.1).

One of the key goals of FERMI is to ensure the sustainability of its findings by sharing insights with EU-funded projects that have a longer life cycle. Therefore, in 2025 FERMI also joined the AI4SafeEurope cluster, a collaborative network of EU-funded research and innovation projects dedicated to applying artificial intelligence to civil security. The cluster unites initiatives such as AVALANCHE, PRESERVE, PRESERVE Drone, NOTIONES, ENSEMBLE, GANDALF, CEASEFIRE, and FERMI itself. Together, these projects support law enforcement agencies and public authorities in detecting threats, conducting secure investigations, and protecting democratic societies. Through shared technical expertise, ethical AI development and a focus on interoperability, the cluster addresses cross-cutting challenges including disinformation, organised and cross-border cybercrime, and border security. AI4SafeEurope ensures that artificial intelligence is leveraged responsibly and transparently, enabling safer societies through collaborative innovation and trustworthy AI solutions.

As part of the AI4SafeEurope Cluster and in the framework of the final event of the FERMI project, a public webinar took place on Wednesday 10 September 2024. The webinar focused on the pressing challenge of disinformation and their impact on European societies, security, and democracy, exploring how artificial intelligence and advanced investigative tools can empower law enforcement agencies, policymakers, and civil society to detect, analyse, and counter these threats in an ethical and transparent manner. On behalf of FERMI, BPA presented key project outputs and conclusions.

Furthermore, LC is preparing a bundle of materials which will summarise the key findings and policy suggestions of the project. As it is part of the sustainability plan beyond the project's duration, the bundle will be presented in the October monthly meeting of the AI4SafeEurope Cluster, opening the floor to other projects' representatives to pose questions.

Outside of the two cluster, FERMI has been actively cooperating with other projects and initiatives. BPA's Center of Excellence for Police and Security Research (CEPOLIS) joined the EU-HYBNET (Empowering a Pan-European Network to Counter Hybrid Threats) Network, which is an EU-funded network of security practitioners, stakeholders, academics, industry players, and SME actors across the EU that collaborate to counter hybrid threats. As a member of the EU-HYBNET Network, CEPOLIS facilitated the exchange between FERMI and stakeholders of interest, especially with respect to the law enforcement community.

On 26 March 2025, a joint training session in the form of a webinar for a LEA audience was co-organised by EU-HYBNET and FERMI. The webinar was called "Law Enforcement vs. Disinformation: Technological Approaches and Training Solutions on Increased Law Enforcement Capability" and open to LEA from across Europe.

Moreover, the FERMI project co-organised the "Research and Innovation Symposium for European Security 2024" (RISE-SD 2024), which took place on 16-17 October in Chalkidiki, Greece. RISE-SD is a one of the leading European Research and Innovation events, in the field of Fighting Crime and Terrorism, Disaster and Crisis Management, Critical Infrastructure Protection, Cybersecurity and Border Management Research. Representatives from BPA and ATOS presented FERMI's concept and key products. The RISE-SD meeting was attended by well over 100 expert participants.

[27] Article available in FERMI's website: https://fighting-fake-news.eu/articles/protecting-european-elections-disinformation-horizon-europe. Article available in VIGILANT's website: https://www.vigilantproject.eu/blog-posts/protecting-european-elections-from-disinformation-with-horizon-europe

[28] Article available in DG HOME's website: https://home-affairs.ec.europa.eu/news/research-projects-help-combat-disinformation-ahead-elections-2024-05-30_en

Funded by the European Union

BPA did two further remote project presentations at gatherings organised by other projects, including at the EITHOS project's second workshop on training LEAs in using the EITHOS tools. The workshop was held on 26 November 2024 at the premises of KEMEA in Athens, Greece. BPA also presented the FERMI project at an ARM's project's cluster online meeting of Horizon Europe projects funded under the 2024 call on 05 March 2025.

As presented above, the FERMI project has engaged in close collaboration with other EU-funded initiatives, with the goal of fostering synergies and amplifying the overall impact of research and innovation efforts. While such cooperation is critical to the project's ecosystem approach and enhances visibility and mutual learning, it also carries inherent risks related to cross-project data exposure. These risks stem from the fact that different consortia may operate under varying security baselines, access rights, and data handling protocols. In recognition of this, FERMI confirms that any exchange of data with external projects is strictly governed by robust access controls, binding non-disclosure agreements, and full compliance with EU data protection legislation and, where applicable, EU Classified Information (EUCI) handling rules. Although the FERMI project itself has not generated EUCI-classified material, it acknowledges the theoretical possibility that such information may be received from external partners. This scenario is addressed through internal safeguards and protocols designed to ensure that any such data is handled in accordance with all relevant regulatory and procedural requirements (see D1.4).

4.8 Training activities

Task 6.3 foresees the planning, delivery, and continuous refinement of training activities and materials aimed at enhancing public awareness, fostering digital literacy, and strengthening societal resilience against disinformation. These activities are addressed to both general audiences and specialised stakeholders, and include a combination of internal and external workshops, public webinars, and downloadable training packages. All training activities are implemented in full alignment with FERMI's Ethics and Legal Protocol (D1.4), the General Data Protection Regulation (GDPR), and established procedures for informed consent and data protection compliance.

Three main training actions were undertaken. The first external training took place in February 2024 (M17), in the form of a public webinar titled "A Dive into the Societal Landscape of Disinformation – Balancing between Law Enforcement and Fundamental Rights". This was followed in March 2024 (M18) by an internal workshop focused on the legal and ethical requirements relevant to the FERMI platform, particularly in relation to data protection and ethics self-assessment. The specifics are laid out in D6.2.

In December 2024 (M27), CONVERGENCE delivered the second external webinar entitled "Digital Trust in Action: Technological Approaches and Citizen Empowerment to Combat Disinformation", which combined a platform demonstration with a high-level reflection on digital skills, AI literacy, and citizen empowerment. The third and final external webinar was organised in September 2025 (M36), under the title "How Disinformation Evolves: Narratives, Digital Influence, and Trust." This session explored how disinformation adapts over time, how it intersects with legitimacy and political narratives, and how countermeasures can preserve democratic trust and civic pluralism.

Complementing these events, a suite of training resources has been developed and published on the FERMI website, including: "Navigating Disinformation: A Comprehensive Guide" (February 2024), the "Ethics Assessment Toolkit" (March 2024), and "Digital Trust: A Practical Path" (January 2025). These materials include theoretical modules, interactive exercises, and practical guidance, and will be further supplemented following the final webinar.

Accordingly, the project has met the objective of delivering "at least 3 awareness raising events to the general public"[29] and to "produce a set of training packages that will be freely available in the project platform for download."[30]

[29] Grant Agreement, Part B, p.6.

[30] Grant Agreement, Part A, p.14-15.

4.8.1 External Webinar – “Digital Trust in Action: Technological Approaches and Citizen Empowerment to Combat Disinformation”

On 04 December 2024, CONVERGENCE, in its role as leader of T6.3, organised and hosted the above-mentioned public webinar “Digital Trust in Action: Technological Approaches and Citizen Empowerment to Combat Disinformation.” The webinar aimed to raise awareness around disinformation and digital literacy, present the technological tools developed within FERMI, and foster dialogue on building societal resilience through responsible AI use. The event was recorded and widely disseminated through the FERMI platform and website, accompanied by targeted outreach on social media.

The registration process

To coordinate this training activity/webinar, the organisational process followed FERMI’s standard operating procedure for ethics and GDPR compliance. The registration was facilitated via an online form, which provided participants with an information sheet detailing the project scope, the purpose of the webinar, data collection and recording procedures, confidentiality policies, and relevant contact points. A consent form was included as part of this process and was required for participation. All attendees, including consortium members and guest speakers, completed this step. During the webinar, participant names and personal information were anonymised by default, attendees could not view each other’s display names or email addresses.

The content of the webinar

The 90-minute session was carefully designed to balance technical demonstration with participatory reflection and dialogue. The webinar opened with introductory remarks and a live poll to assess the participants’ baseline familiarity with digital trust and disinformation. The first half of the session featured a comprehensive presentation by Nikos Dimakopoulos (ITML), who provided an in-depth walkthrough of the FERMI platform’s core AI-enabled modules. These included the Spread Analyser, designed to capture the dissemination of relevant content across digital channels; the Dynamic Flows Modeler, which forecasts the potential societal consequences of disinformation narratives; the Sentiment Analysis module, used to assess emotional and behavioural cues across online networks, the Swarm Learning framework, which allows for decentralised learning and modelling of community-level responses to disinformation threats, and the Community Resilience Management Modeler, which proposes counter-measures to mitigate disinformation campaigns, if an impact assessment (based on the Behaviour Profiler and Socioeconomic Analyser) deems said campaigns grave enough to require corrective action.

These modules were contextualised through the platform’s three use cases, far-right political extremism, health-related disinformation during public crises, and far-left radicalisation. Real-world scenarios were used to demonstrate how the platform supports early warning, credibility analysis, and risk forecasting, with particular emphasis on how AI models are trained and adapted to evolving narrative tactics and online behaviours.

The second part of the session featured a fireside-style conversation with David Timis, an international speaker and advocate for AI ethics and digital empowerment. His contribution moved beyond the technical realm, offering a societal lens on the human competencies required to navigate complex digital ecosystems. Emphasis was placed on digital literacy not as a technical skillset alone, but as a civic and ethical practice rooted in critical thinking, information discernment, and participatory citizenship. He discussed the dual role of artificial intelligence, as both an amplifier of disinformation (e.g. via deepfakes or automated narratives) and as a potential shield when used responsibly for detection and moderation.

David Timis also underscored the importance of cross-sectoral collaboration between governments, platforms, educators, and civil society actors in combating the spread and influence of disinformation. He highlighted the need for investment in upskilling initiatives, with a focus on vulnerable and digitally marginalised populations. Finally, the

Funded by the European Union

discussion linked back to FERMI's mission by framing the need for integrated responses that combine technological innovation with public awareness and inclusive policy approaches.

The session closed with an interactive Q&A, where participants engaged with both speakers on questions relating to practical steps citizens can take to counter disinformation, ethical AI use in public administration, and the long-term societal implications of disinformation in hybrid media environments.

Collecting feedback and evaluating impact

To measure learning outcomes, pre- and post-webinar surveys were conducted. The pre-survey revealed relatively low baseline familiarity in key areas, with only 5.26% of participants indicating knowledge of "digital trust" and 20.00% awareness of disinformation tools. After the session, post-survey results showed significant increases across all indicators. Familiarity with digital trust rose to 73.68%, and awareness of tools to combat disinformation jumped to 94.73%. Notably, over 80% of participants demonstrated measurable improvement, meeting the GA's KPI of having "80% of participants to the training sessions of FERMI framework show better understanding of D&FN (measurement before and after the training activities)."[31] The results confirmed the effectiveness of combining technical presentations with interactive and reflective components.

Training material

To consolidate the outcomes of the webinar and support further knowledge dissemination, CONVERGENCE developed an in-depth training guide titled *"Digital Trust: A Practical Path"*. The guide is based on the core themes discussed during the session and is designed as both a stand-alone educational tool and a follow-up resource for webinar participants. It reflects FERMI's interdisciplinary approach to disinformation by integrating technical understanding with critical and civic competencies.

The material begins by defining key concepts such as digital trust, disinformation, misinformation, and malinformation. It then introduces readers to the ethical and social implications of digital manipulation and explains how artificial intelligence technologies can be used both to spread and detect harmful narratives online. Drawing on the FERMI's platform architecture and use cases, it offers practical illustrations of how disinformation spreads across information ecosystems and how it can be mitigated through a combination of tools, policies, and skills.

Structured in four main chapters, the guide covers:

1. Understanding Disinformation – its types, goals, and psychological impact.
2. Critical Thinking and Digital Literacy – promoting source evaluation, credibility assessment, and fact-checking techniques.
3. Artificial Intelligence and Responsibility – explaining how AI operates in disinformation and how it can support human judgment.
4. Interactive Exercises – including games such as *"Find the Fake"*, *"Real or Not?"*, and media analysis simulations that challenge learners to apply their skills in identifying manipulation.

Each module is accompanied by reflection prompts and quizzes designed to enhance active learning. In addition, the guide provides curated links to open-access online courses, EU policy references, fact-checking platforms, and AI ethics toolkits, enabling users to continue their learning beyond the session.

The training package is freely available to the public and has been published on the FERMI website.

4.8.2 External Webinar – "How Disinformation Evolves: Narratives, Digital Influence, and Trust"

Date and Topic

[31] Grant Agreement, Part B, p.23.

On 10 September 2025, CONVERGENCE, in its role as leader of T6.3, organised and hosted the public webinar "How disinformation evolves: Narratives, digital influence and Trust." The session explored how disinformation narratives form and shift over time, the mechanisms of digital influence, and practical pathways to strengthen public trust and societal resilience. The event was recorded and disseminated via the FERMI platform and website, accompanied by targeted outreach on social media and partner channels.

The registration process

Again, to coordinate this webinar, the organisational process followed FERMI's standard operating procedure for ethics and GDPR compliance. Registration was facilitated via an online form that included an information sheet (project scope, session purpose, data collection/recording, confidentiality and contact points). A consent form was required for participation and completed by all attendees and speakers. During the webinar, participant names and personal details were anonymised by default; attendees could not view one another's display names or email addresses.

The content of the webinar

The 90-minute session combined expert inputs with interactive elements and Q&A. The webinar was moderated by Panagiota Kokoliou (CONVERGENCE). Opening remarks were provided by Sven-Eric Fikenscher (Bavarian Police Academy; FERMI Coordinator), situating the topic within FERMI's broader focus on the nexus between disinformation and the crime landscape and the importance of civil-society collaboration for effective counter-measures.

Part I – FERMI overview and toolkit for legitimacy & resilience. **Aikaterini Beli** (Lisbon Council) presented a concise overview of FERMI's aims and toolchain, using real-world cases to illustrate online-to-offline risk pathways (e.g., riots triggered by false narratives). The presentation highlighted:

- Disinformation Sources & Spread Analyser (mapping source accounts and spatial/temporal spread),
- Sentiment Analysis Module (audience response and emotional resonance),
- Dynamic Flows Modeler (forecasting spread and potential victims/authors),
- Swarm Learning (privacy-preserving, decentralised model training),
- Behaviour Profiler & Socio-economic Analyser (community vulnerability and cost estimation), and
- Community Resilience Management Modeler (identifying resilience levels and informing targeted interventions).
  Framed around the credibility gap, the talk connected technical capabilities with media-literacy, prebunking, and trust-building actions.

Part II – Narrative dynamics, information disorder, and recent evidence. **Dr Sofia Tipaldou** (Assistant Professor in International Relations) unpacked **information disorder** (following Claire Wardle's typology: satire/parody, misleading content, false context, imposter content, manipulated content, fabricated content) and shared research insights on the disinformation life cycle and Kremlin-linked proxy operations:

- Disinformation is dynamic and co-created, travelling across languages and cultures; credibility is often performatively constructed by diverse actors (including counter-disinformation units).
- Post-ban, Kremlin actors have leaned on local proxy sites and multilingual operations, with limited measurable impact in EU contexts studied to date; some content shows reliance on AI-generated text with quality issues.
- Implications: prioritise prebunking, strengthen media-literacy, and complement LEA-oriented tools with civil-society capacity.

Interactivity

The session opened with a pre-webinar Mentimeter poll (baseline familiarity, perceived impacts, spread/evolution awareness, confidence mapping narratives), and closed with a post-webinar poll repeating the items plus a question on intended practice uptake. The event concluded with an interactive Q&A focusing on practical citizen actions, institutional trust, and balancing counter-measures with rights.

Collecting feedback and evaluating impact

To measure learning outcomes, pre- and post-webinar Mentimeter surveys (10 September 2025) were conducted. Baseline results showed moderate familiarity with specific societal impacts of disinformation (52.9% familiar/very familiar) and low familiarity with how disinformation spreads and evolves (31.3% familiar/very familiar) as well as low confidence in mapping narratives (23.5% confident/very confident). After the session, post-survey results showed significant gains across all indicators: familiarity with the term *disinformation* rose to 100.0% familiar/very familiar (with 78.6% being very familiar), concern about societal impact reached 93.8% (concerned/very concerned), familiarity with specific impacts increased to 88.9%, understanding of spread/evolution reached 100.0%, and confidence in mapping narratives rose to 81.3% (with "not confident at all" dropping to 0%). Post-only items further indicated 81.3% confidence in choosing a safe way to respond and 76.5% intent to apply at least one idea within 30 days. These outcomes meet the above-mentioned KPI for training activities under Task 6.3 and confirm the effectiveness of combining technical presentations with interactive polling and reflective discussion

Training material

To consolidate outcomes and support wider public dissemination, an end-user guidebook has been developed. It brings together core concepts, definitions, and general principles related to disinformation and digital trust, offering concise explanations and broad guidance for safe, responsible online participation. In general terms, it includes key terminology, indicative references to relevant standards and good practices, brief self-check steps before sharing content, and light self-learning material for further exploration. The content is organised into short, easy-to-use entries with cross-references for additional reading and is freely available via the FERMI website, alongside the webinar recording.

4.8.3 Training activities aimed at LEAs

4.8.3.1 Training activities by SPA

On June 13, 2025, the National Fraud Centre within the Swedish Police Authority organised a lecture and training session to address the critical issue of disinformation as a global challenge. The aim was to increase awareness of disinformation and its impact on FERMI's work in law enforcement and on society as a whole and to discuss how its effects should be addressed and counteracted.

The agenda covered several key points that highlight the relevance of this topic, particularly in the context of law enforcement and public awareness. Additionally, the challenges faced by the police in managing extremism and disinformation in Sweden were discussed, so was the question how various disinformation campaigns have been exploited by extremist groups both within and outside the country.

Point	Details
Types and motives of disinformation	Explored various forms of disinformation and the underlying motives driving these campaigns.
Relevance for Police Work	Discussed why understanding disinformation is crucial for law enforcement agencies.

Funded by the European Union

Impact campaigns and disinformation in Sweden and the EU	Focused on specific campaigns, including the LVU (Care of Young Persons Act) campaign against social services and the burning of the Quran, and the role of disinformation in influencing public opinion during the EU election.
Challenges in handling extremism	Addressed the police's challenges in managing extremism and the impact of disinformation on these efforts.
Exploitation of disinformation by extremist groups	Discussed how extremist groups, both within and outside Sweden, have utilised disinformation campaigns to further their agendas.
Consequences of disinformation	Examined the broader societal impacts of disinformation, including trust erosion and polarisation.
Challenges in digital literacy	Addressed the difficulties in promoting digital literacy as a means to combat disinformation.
Strategies to mitigate effects	Presented strategies for reducing the impact of disinformation, including community engagement and education.
FERMI Project	Introduced the FERMI project, FERMI-platform and the tools to combat disinformation.

Table 2 - Key points discussed

Participation

The lecture attracted 129 participants from the seven regions of Sweden, including those working strategically and operationally, with 55 attending digitally. This strong turnout indicates a significant interest in the topic and highlights the need for ongoing dialogue and action against disinformation.

4.8.3.2 Training activities by PUCF

PUCF's LEA training activities were carried out in the framework of Task 5.5, which focuses on increased skills, tools and training for law enforcement agencies. The scope of the task is to interact with security practitioners and operational human resources involved in the FERMI validation campaigns and maximise the potential of the project solutions. A set of multi-dimensional training packages for LEAs and security practitioners were developed through the curriculum's conception that increases understanding of additional capabilities offered to LEAs and security practitioners such as identification of disinformation and fake news spread with a potential impact on the security and safety of the civilians and online and offline training courses and seminars based on the FERMI platform. A set of training materials on each emerging technology deployed within the project and their future potential was produced within the task.

Three training sessions were organised for LEAs within the FERMI project. The first training session was organised as a synchronised (face to face) session for the Master of Police Services degree students at the Police University College of Finland. Participants were police officers with at least three years of work experience aiming for commanding positions. The training session took place in November 2024 (M26). The session was attended by 13 students from police departments around Finland. The second training session was organised online as a webinar in March 2025 (M30). The webinar was titled "Law Enforcement vs. Disinformation: Technological Approaches and Training Solutions on Increased Law Enforcement Capability" and it was organised in cooperation with EU-HYBNET. The event was attended by a total of 55 participants from all over Europe.

The third training session took place on the first day of the FERMI consortium meeting in April 2025 (M31) as a synchronous training session. The first day of the meeting was attended by additional police officers from Guardia Civil that supplemented the presence of consortium LEA experts from SPA, BFP and FMI plus experts from the Police Colleges (PUCF & BPA). There were 23 participants who gave their consent and took part in the session. The duration of each training session was 1,5 hours. Additionally, law enforcement agencies in the consortium took part in the online evaluation of training material concurrently with the first training session in November 2024 in the Moodle environment where they were able to review the first versions of the eLearning training materials on the emerging technologies deployed within the project. Altogether 20 persons gave consent and visited the course.

Concurrently with the training activities an online eLearning course was developed in close cooperation with the partners in FERMI. The technological partners were intimately involved in generating learning material on the emerging technologies that later underwent a pedagogic design process, and which were incorporated into the eLearning course. The ensuing eLearning course is titled "Predictive policing and innovative technologies in law enforcement". It will be available to the project partners with the aim to make it more widely available via the European Union Agency for Law Enforcement Training (CEPOL).

The structure of the training sessions

Overall, the structure and content of the training sessions for LEAs changed iteratively according to the stage of the project at the time of each session. The first training session focused on the Moodle course material on emerging technologies. The participants reviewed the information in the Moodle platform and were thus able to learn about the solutions that would be available in the future. The project partners from PUCF first introduced the project after which the concept and intended structure of the eLearning course was introduced. As the session was organised synchronously, the next phase contained self-study and group work. The session concluded with a more detailed explanation on the FERMI platform after which organising partner (PUCF) encouraged participants to offer feedback both orally and in written form in the Moodle environment.

The next two sessions followed a slightly different structure while still containing detailed information on the emerging technologies and their implementation in the FERMI project. These changes were made as the technological solutions progressed and to better suit the online format of the webinar. The webinar also catered to the contribution by EU-HYBNET. The webinar was organised and moderated by PUCF and included the following presentations:

- Disinformation and domestic security: The FERMI project and its solutions
- Core theme analysis on EU-HYBNET's implementation & training methodology
- Educational approaches to future law enforcement capabilities
- Insights on AI-based technological tools to analysing and mitigating disinformation risks in law enforcement

Similarly, the participants in the third and last session had already had the opportunity to attend a presentation on the FERMI project earlier in the day thus the structure of the final training session introduced the educational approaches to future law enforcement capabilities, gave insights on AI-based technological tools to analysing and mitigating disinformation risks in law enforcement and included a demonstration of the FERMI platform. Therefore, in this session participants were able to better examine the use of the technology in the FERMI platform which further exemplified its usability to law enforcement officials.

In accordance with GA requirements, the FERMI consortium delivered "at least three training sessions focusing on skills development of LEAs' personnel"[32] and produced an eLearning course that will help disseminate project results after the project has concluded.

All training activities were implemented in alignment with FERMI's Ethics and Legal Protocols and EU's GDPR. The training participants completed informed consent proceedings before attending the training sessions. They were also given the opportunity to offer feedback after the sessions. The pedagogic principles for the eLearning course, the

[32] Grant Agreement, Part B, p.5.

Funded by the European Union

curricula, the sessions' execution and feedback are thoroughly described in D5.4 The FERMI Training curricula for officers & sessions' execution report.

Funded by
the European Union

5 Communication

In Horizon Europe, communication refers to the strategic and continuous process of informing and engaging the general public about the goals, progress, and results of a research project. It is broader in scope than dissemination, as it targets society at large rather than specific stakeholders. Communication aims to raise awareness, foster understanding, and generate interest in the project's themes and relevance, making complex research accessible and meaningful to non-specialist audiences. It also plays an important role in reinforcing transparency and accountability in publicly funded research.[33]

5.1 Overview of communication

In FERMI, communication has been a central pillar throughout the project's lifecycle, contributing not only to visibility but also to the creation of a recognisable identity around the project and its societal mission. From the early stages, FERMI developed and implemented a multi-channel communication strategy to raise awareness about the challenges of disinformation and the technological and policy tools being explored to address them. This included launching and maintaining a dedicated project website, producing regular newsletters, curating active social media accounts, and generating a coherent visual identity to support recognisability.

FERMI's communication activities also extended into the public domain through participation in conferences, joint webinars, practitioner events, and collaborations with other Horizon Europe projects. These efforts allowed FERMI to reach diverse audiences, ranging from students, academics, and law enforcement professionals to journalists, civil society organisations, and the wider public. Messaging was tailored to the context, ensuring that the project's complex themes, such as algorithmic detection of disinformation, ethical AI development, and risk-based governance, were conveyed clearly and meaningfully.

By embedding communication into the project's overall strategy, FERMI not only succeeded in informing and engaging external audiences but also laid the groundwork for impactful dissemination and future exploitation. The communication work helped position FERMI as a reference point in the fight against disinformation, ensuring that its results were seen as timely, credible, and aligned with broader societal values.

The following table presents an overview of the activities regarding each communication mechanism distributed over the different phases of the project's duration, represented by the central aim of each phase.

Communication mechanism	1 – Raise awareness	2 – Transfer knowledge	3 – Deliver impact	4 – Accelerate sustainability
Social media	Establishment of presence in social media completed. Reproduce relevant content and monitor relevant hashtags; upload public material; follow influencers of the domain; engage	Promote project's outcomes and events; develop and implement informative communication campaigns on relevant themes; interact with followers to get feedback; answer on comments and private	Promote project's outcomes and events; develop and implement informative communication campaigns on relevant themes; interact with followers to get feedback; answer on comments and private	Publish material such as project's recap. Share tools and instruments used to prepare future steps.

[33] 33 Read here a more comprehensive overview of communication, dissemination and exploitation: https://op.europa.eu/en/publication-detail/-/publication/58ad3394-0a63-11ee-b12e-01aa75ed71a1/language-en

	with other projects and initiatives.	messages on the various channels; upload public material; reproduce relevant content and use relevant hashtags.	messages on the various channels; upload public material; reproduce relevant content (more sporadically).	Develop online partnerships with relevant stakeholders.
Project's website	Website completed; search engine optimisation completed.	Regular update; web analytics monitoring; provide content of impact.	Regular update; web analytics monitoring; provide content of impact.	Regular update; web analytics monitoring; provide content of impact.
Project's blog	Deployment project's blog completed. Provide blog posts related to project's positioning and technologies.	Provide frequent blog posts to initiate discussions on specific issues relevant to the project to receive feedback.	Publish frequent blog posts to demonstrate and promote project's results.	Publish blog posts to attract and advertise successful partnerships and/or growing user base.
Communication material	Project branding and visual identity, communications starter pack completed.	Prepare revised communications pack and frequent releases of e-newsletter; publish blogs/news in EU dissemination instruments (e.g. Cordis News, research EU magazines etc.).	Prepare final communications starter pack and frequent releases of e-newsletters and video demonstrators; publish blogs on the FERMI website.	Frequent releases of e-newsletters; publish blogs/news on the FERMI website.

Table 3 - Communication mechanism for each communication channel per project phases

5.2 FERMI community

The primary aim of FERMI's strategic communication plan was to foster the growth of a dedicated and informed community of interest, with a clear emphasis on transparency, inclusiveness, and accessibility. This objective has been consistently pursued throughout the duration of the project by offering open access to all disseminated content, tailored to different target audiences. Key communication tools and channels included the project website, active social media profiles, themed campaigns, webinars, press and media engagement, a scientific poster, regular newsletters, and FERMI's visible presence at scientific and policy-focused conferences.

To build and nurture this community, FERMI employed a multi-layered communication approach that integrated both online and offline activities. Targeted outreach efforts addressed the needs and interests of specific groups, while also raising public awareness about the broader challenges posed by disinformation and the use of technology in security governance. These activities contributed to increasing understanding of the project's objectives and results, and to positioning FERMI within an evolving public discourse on digital risks, trust, and technology.

During the initial set-up phase in the first year, the project focused on establishing visibility and identifying its audiences. This phase was supported by social media engagement (e.g. shares, reposts, and likes), the definition of key messaging, and collaboration with influential partners, including key opinion leaders , institutions, NGOs, and private actors. In the following two years, the communication plan focused on reinforcing these early achievements by scaling engagement, deepening dialogue with existing audiences, and expanding FERMI's reach to new stakeholders, especially in alignment with ongoing developments in policy, research, and technological innovation.

FERMI also acknowledged early on the potential sensitivities associated with its subject matter, specifically, the use of technology by law enforcement, and took proactive steps to address them. While the broader discussion of risk and impact is covered in other parts of this deliverable, it is important to reiterate here that engagement with potentially critical or concerned stakeholders has been a key component of the communication strategy. FERMI has made a deliberate effort to communicate openly about its legal and ethical commitments, both to ensure transparency and to build trust. These efforts included collaboration with actors concerned with privacy and fundamental rights, and active involvement in shaping a constructive dialogue around compliance and accountability.

Ultimately, FERMI's communication strategy has not only succeeded in building a diverse and active community of interest but has also laid the foundation for the long-term impact and uptake of the project's results. The sustained interaction with stakeholders and the emphasis on clarity, openness, and responsiveness have helped position FERMI as a credible, trustworthy, and forward-looking initiative in the evolving European security and disinformation landscape.

5.3 FERMI social media presence

Over the three-year lifespan of the FERMI project, social media has been a central pillar in its communication and dissemination strategy. Through a coordinated, multi-channel approach, FERMI successfully built a recognisable online presence that amplified project visibility, disseminated research results, and nurtured an engaged community of interest around disinformation, content moderation, AI governance, and law enforcement innovation.

FERMI actively maintains profiles on four key platforms:

- LinkedIn: https://www.linkedin.com/company/fermi-project/
- Twitter: https://twitter.com/fermi_project
- Mastodon: https://mastodon.social/@fakenewsriskmitigator
- YouTube: https://www.youtube.com/@fermi-project

The project consistently used the username "FERMI EU" across platforms to ensure brand cohesion and clear identification with the Horizon Europe programme. Each channel was selected for its unique strengths and target audience, contributing to a well-rounded digital communication ecosystem.

Posts were visually aligned with FERMI's brand identity and included consistent use of the official logo, links to the homepage, and the official project hashtag #FermiEU, alongside thematic hashtags (e.g., #Disinformation, #AIForSecurity, #HorizonEU). To support engagement, FERMI partners were actively encouraged to repost content, tag the project, and engage in comment discussions.

Throughout the project, a planned editorial calendar guided social media activity, ensuring timely and relevant publication of content aligned with project developments and external trends. The platform's built-in analytics tools were used to monitor performance and reach, with data aggregated into regular internal reports to evaluate communication KPIs (see Chapter 3). These insights informed refinements in tone, content format, and timing, enhancing audience engagement and visibility.

As the primary platform for engaging stakeholders in research, policy, and security fields, LinkedIn became FERMI's most strategic and active social media channel. Content included updates on project milestones, scientific outputs, publications, joint webinars, partner spotlights, and participation in EU-wide clusters and events. LinkedIn was especially useful for reaching law enforcement professionals, researchers, civil society representatives, and Horizon Europe peers. In addition to one-off posts, targeted campaigns were run on LinkedIn around key events such as the RISE-SD Symposium, joint webinars and major publications. These helped FERMI reach new audiences and strengthen its reputation as a serious actor in the domains of disinformation mitigation and AI governance.

X (formerly Twitter) was used as a real-time communication tool for publicising events, policy developments, project milestones, and engaging in ongoing debates around AI, disinformation, and digital governance. It served as a platform for visibility during live events and conferences, leveraging hashtags such as #FermiEU, #HorizonEU, and others relevant to EU security research. By following and tagging key stakeholders, FERMI ensured inclusion in broader sector conversations. During later stages of the project, due to a declining platform audience and its changing policies, frequency of posts on X decreased.

Mastodon was explored as an alternative, decentralised platform, particularly in response to the shifting landscape of public trust and ownership on traditional social media. The FERMI Mastodon presence served a niche but engaged an audience of digital rights advocates, academics, and technologists interested in privacy-conscious discussions of disinformation and AI in law enforcement. Though smaller in scale and with a small audience present, it contributed to the project's open access and ethical positioning.

FERMI's YouTube channel was used primarily for hosting video content such as project explainers and webinar recordings. Video content played an important role in reaching broader audiences beyond traditional research communities and increased FERMI's transparency by showcasing how the project operated. Each video description included direct links to the website and further reading, thereby reinforcing the channel as a complementary communication tool. Moreover, videos of project training are available through the website (see section 4.8).

Funded by the European Union

6 Recommendations for operational standards

6.1 Recommendations for the first operational standard - Operational standard for the ethical and responsible use of technology by law enforcement

As part of Task 6.5 of the FERMI project, an operational standard for the Ethical and Responsible Use of Technology by Law Enforcement Agencies has been developed to support LEAs in using AI and data-driven tools in a responsible, ethical, and legally compliant manner. Moreover, the GA requires the consortium in KPI5.3 to make a "contribution to at least 3 standards."[34]

The recommended Operational Standard, included in Annex A, provides a set of clear, actionable practices that reflect existing EU laws and ethical expectations. Rather than setting out abstract principles, it focuses on what LEAs working with emerging technologies can do in practice, how to manage risks, safeguard rights, and ensure accountability throughout the use of technology in their daily operations.

6.1.1 Purpose and Added Value

The purpose of the Operational Standard is to provide LEAs with actionable guidance and good practices to ensure the responsible, ethical and lawful use of technologies including AI, data analytics, and monitoring tools, throughout their lifecycle.

In particular, this Standard serves a dual purpose:

- To assist managers in establishing appropriate governance, oversight, and accountability structures;
- To support implementers (technical staff and officers) in identifying risks, maintaining accurate documentation, and applying safeguards throughout the technology lifecycle.

It is closely aligned with the European AI Act, the GDPR, the Law Enforcement Directive (LED), and the EU Charter of Fundamental Rights, while also drawing on recognised governance frameworks such as COBIT 2019 and ITIL 4.

6.1.2 Key Components of the Standard

The standard is divided in key components, including legal and ethical foundations, core ethical principles and governance, data governance and data protection, lifecycle risk management, transparency and explainability, human oversight and decision-making, procurement and development controls, monitoring, evaluation and auditing, public engagement and external transparency and interoperability and cross-border cooperation.

The final part of the operational standard consists in several conclusions and recommendations that are central to ensure the proper operationalisation of ethical and responsible use of technologies by Law Enforcement Agencies.

1. *Legal and Ethical Foundations*

The Standard incorporates key elements of the EU legal framework, aligning with the AI Act, GDPR, LED, and the EU Charter of Fundamental Rights. It classifies AI systems by risk level, outlines obligations for high-risk use cases, and ensures compliance with principles such as transparency, proportionality, and data protection. It also prepares law enforcement technologies for upcoming regulatory requirements, including conformity assessments and EU-level registration.

2. *Core Ethical Principles and Governance*

[34] Grant Agreement, Part B, p.23.

Funded by the European Union

The Standard is grounded in core ethical principles, operationalised through robust governance measures. These include mandatory Fundamental Rights Impact Assessments for high-risk technologies, human oversight protocols to ensure human control, and a structured approach to assess proportionality and necessity. It promotes fairness through bias mitigation and diverse data use, and ensures transparency via Explainable AI techniques, making systems understandable to both experts and laypersons.

3. *Data Governance and Security*

The Standard adopts a robust approach to data governance and security, emphasising bias-aware data curation, ongoing monitoring, and clear data traceability through logging and lineage mechanisms. It promotes strong privacy safeguards such as pseudonymisation, anonymisation, and strict data retention limits, alongside role-based access controls and appropriate consent procedures. These measures collectively support accountability, transparency, and data protection throughout the system's lifecycle.

4. *Lifecycle Risk Management*

The Standard integrates the EU AI Act's risk-based approach across the AI system lifecycle. Unacceptable risk applications, such as social scoring or untargeted facial recognition, are strictly banned. High-risk systems face stringent obligations, including conformity assessments and human oversight. Limited risk systems must meet transparency requirements, while minimal risk systems, though not regulated, are encouraged to follow ethical and governance best practices.

5. *Procurement, Development and Deployment*

The Standard embeds ethical and legal considerations into the procurement, development, and deployment of AI systems. It promotes the use of the European Commission's Model Contractual Clauses for high-risk systems, mandates third-party audits, and encourages early stakeholder engagement. Socially responsible procurement criteria, including human rights and labour protections, are also emphasised. Law enforcement agencies are urged to treat procurement as a strategic tool for enforcing compliance and accountability throughout the supply chain.

6. *Monitoring, Evaluation and Decommissioning*

The Standard outlines clear procedures for monitoring, evaluation, and decommissioning to ensure long-term accountability. It includes continuous post-market monitoring, immutable logging for auditability, and regular compliance reviews. Robust decommissioning protocols are also established for systems that become outdated, non-compliant, or ethically unjustifiable. These measures help law enforcement agencies remain adaptable to technological developments and evolving legal standards.

7. *Public Communication and Cross-Border Cooperation*

The Standard emphasises public trust and democratic accountability in the use of law enforcement technologies. It promotes transparency through plain-language communication, accessible channels for public engagement, and support for secure cross-border cooperation. By aligning with the European Interoperability Framework and mechanisms like Europol, the Standard helps ensure that innovation in policing remains rights-respecting, transparent, and interoperable across the EU.

6.1.3 Strategic Relevance for the Project

In the context of this project, the Operational Standard serves both as a baseline compliance framework to guide the design and development of AI tools by technical partners, and as a capacity-building resource for law enforcement agencies seeking to embed ethical principles and legal requirements into their internal governance structures.

By aligning the project's outputs with the Standard, partners ensure that the technologies developed are rights-preserving and bias-aware by design, auditable, explainable, and legally defensible—and, crucially, that they are trusted by both public authorities and the communities they serve.

The Standard therefore acts as a bridge between technological innovation and democratic accountability, ensuring that digital transformation in policing respects the rule of law and upholds fundamental rights.

6.1.4 Relevance Within and Beyond the Project

The *Operational Standard for the Ethical and Responsible Use of Technology by Law Enforcement Agencies* plays a vital role both within the project and in a wider European context. Internally, it provides a clear, structured framework to ensure that the AI tools developed for law enforcement are designed and deployed in line with EU legal obligations and ethical expectations.

It helps both technical partners and end-users navigate complex compliance demands, ranging from data protection and risk management to human oversight, while ensuring that resulting tools are practical, trustworthy, and aligned with fundamental rights.

Beyond the project, the Standard offers broader relevance for law enforcement agencies and public authorities across Europe. By converting abstract legal requirements, such as those found in the AI Act, GDPR, and Law Enforcement Directive, into concrete and operational good practices, it supports the responsible and rights-based use of AI in real-world policing.

Its focus on the full lifecycle of technology, along with its emphasis on transparency, accountability, and public engagement, makes it a resource that can be adopted, adapted, or scaled by institutions seeking to implement ethical and lawful AI use, even beyond the specific technologies or use cases addressed within this project.

6.2 Recommendation for second operational standard - Use of Technologies to tackle Disinformation by Law Enforcement Agencies

6.2.1 Introduction

The present document sets out a recommendation for an Operational Standard to be offered to LEAs to provide disinformation identification and analysis services. This Standard sets out the steps LEAs need to take to pick up disinformation and to understand the impact on security and safety of civilians and establishes a framework for effectively tackling disinformation by LEAs, including what kinds of data should be gathered to be used in artificial intelligence and big data technologies, protocol for their management, storage and security and procedures for LEAs to support their decision-making processes with these tools by including the challenges of human rights.

Due to the increasing number of information and communication technologies that can spread disinformation, it is important to have international Operational Standards that provide a common understanding for the tackling of disinformation. This document is intended to enhance already existing security standards by adding a focus relevant to disinformation.

The increasingly widespread disinformation phenomenon, and the growing complexity of ICT systems, can make it challenging for LEAs to tackle disinformation and achieve compliance with the various applicable laws. LEAs can prevent uncertainty and distrust from arising by handling disinformation issues properly and effectively.

Use of this document will:

- aid in the design, implementation, operation and maintenance of ICT systems that handle disinformation;
- spur innovation solutions to enable the handling of disinformation within ICT systems; and
- improve LEA's disinformation programmes through the use of best practices.

Funded by the European Union

The framework provided within this document can serve as a basis for additional disinformation standardisation initiatives, such as for example:

- the implementation and use of specific disinformation handling technologies;
- overall disinformation management;
- human rights risk assessments related to disinformation.

In terms of methodology, in order to develop this operational standard, specific questions regarding the issues dealt herein were sent to LEAs in writing in a questionnaire circulated by INTRA (that included other issues covered by FERMI other than standardisation). 8 responses were received in writing on 31 July 2025, which were incorporated into this standard. Furthermore, feedback was sought from LEAs during the webinar that took place on 11 September 2025.

6.2.2 Scope

This document provides a framework which is intended to guide LEAs define their disinformation identification and analysis within an ICT environment by:

- specifying a common disinformation terminology;
- defining the LEA actors and their roles in tackling disinformation;
- establishing a framework for effectively tackling disinformation by LEAs;
- providing references to known disinformation principles, including relevant human rights obligations.

This document is applicable to LEAs and organisations involved in specifying, procuring, architecting, designing, developing, testing, maintaining, administering, and operating information and communication technology systems or services where disinformation controls are required.

The full operational standard is available in Annex B.

6.3 Third contribution to an operational standard

The third contribution to an operational standard consisted of submitting input and comments to a standard that is currently being developed by the European Telecommunications Standards Institute (“ETSI”). The contribution was submitted in writing on 21 July 2025 on Draft ETSI TR 104 137 V0.0.3 (2025-04), DTR/CYBER-00156, entitled “Cyber Security; Human-to-Human Online Preventative Security; H2H – OPS”. The contribution was presented by a representative of IANUS as a delegate of the Cyprus Organisation of Standardisation at the meeting of ETSI CYBER(25)43b021, which took place on 24 July 2025. Her contribution was tabled as item CYBER(25)43b017 on the agenda, entitled “Contribution_CYBER_DTR/CYBER-00156_H2H-OPS" (see table below). After her presentation of her substantive input on disinformation on the H2H-OPS, the participants agreed with all comments, with the chair agreeing to make updates to the next draft of H2H to take this into account. On 12 August 2025, an updated draft was circulated (ETSI TR 104 137 0.0.7 (2025-08) incorporating all the comments included in the contribution of IANUS.

The table “Public Review: Comments on Draft ETSI TR 104 137 V0.0.3 (2025-04), DTR/CYBER-00156 Cyber Security; Human-to-Human Online Preventative Security; H2H – OPS” can be found in Annex C.

Funded by the European Union

7 Impact and further recommendations on advancing the fight against disinformation

Lastly and following up on the proposed new operational standards, a brief analysis of where the project is located in the EU-wide effort to fight disinformation is carried out followed by some further non-standardisation suggestions as to how said fight might be advanced further in light of the LEA-relevant technical and operational insights of FERMI. The FERMI project was designed and implemented against a rapidly evolving threat landscape where democracies face mounting pressure from hybrid threats — complex, multi-layered attacks that combine conventional tactics with disinformation, cyber operations, and interference in public discourse. As outlined in the Niinistö report, these hybrid strategies are designed to exploit systemic vulnerabilities across domains — targeting not just military or economic infrastructures, but also cognitive resilience and societal cohesion. The report stresses that such threats are "deliberate and coordinated," often operating below the threshold of armed conflict, but with the strategic intent of weakening democratic institutions, fracturing public trust, and destabilising open societies.

The European Democracy Shield initiative reinforces this diagnosis. It warns of an increasingly hostile information environment where foreign and domestic actors use disinformation, information manipulation, cyberattacks, and covert political influence to distort democratic processes, particularly elections, and erode trust in institutions. The initiative calls for a "whole-of-society" response to strengthen democratic resilience, underlining the importance of situational awareness, preparedness, and strategic communication, with disinformation recognised as a key component of the hybrid threat spectrum.

In confronting the evolving spectrum of hybrid threats, a crucial insight has emerged: disinformation is not simply a matter of false content, but a systemic vulnerability embedded in the wider information environment. To understand and address it, we must move beyond content moderation or reactive fact-checking and adopt a systemic lens. This is where the concept of information ecosystems becomes vital.

Information ecosystems are the environments where humans interact with technologies to produce, circulate and interpret information. As Wanless et al. explain, these ecosystems are "communities bound by shared ideas and connected via technologies," where decisions are made and meanings are negotiated, whether at the level of an individual, an institution or an entire society. Crucially, they are not neutral spaces: they are shaped by infrastructure, business models, media dynamics, digital literacy levels, and the power relations among platforms, states and civil society actors. In the age of hybrid threats, the fragility or resilience of these ecosystems determines how societies respond to manipulation, interference and crisis.

The 2025 Forum on Information and Democracy report reinforces this framing by proposing an information democracy approach; one that reclaims the public interest in the digital sphere and builds the governance foundations for a healthy information environment. It argues that ensuring the integrity of information ecosystems is now central to democratic resilience, particularly in the face of coordinated disinformation campaigns, AI-generated content, and opaque platform governance. Both reports converge on the view that any strategy to defend democracy must begin with understanding and shaping the ecosystems in which democratic discourse unfolds.

The convergence of hybrid threats and ecosystem fragility demands more than reactive measures; it calls for a structural, long-term response that engages all sectors of society. This imperative is echoed in the European Democracy Shield, which frames the EU's response not as a narrow administrative effort, but as a whole-of-society endeavour. The rationale is clear: disinformation, foreign interference, and related hybrid tactics do not only target state institutions, but they also exploit social divisions, weaken public trust, and corrode the connective tissue of democratic life.

A whole-of-society approach builds on, but moves beyond, the traditional whole-of-government logic. While whole-of-government emphasises coordination between ministries, agencies and levels of government, whole-of-democracy integrates the role of citizens, civil society, media, educators, researchers, and private actors into the effort to safeguard democratic integrity. It is based on the recognition that today's information threats are sociotechnical: they exploit both technological infrastructures and human vulnerabilities and therefore require multidimensional and inclusive countermeasures.

The FERMI project directly contributed to operationalising that shift, to a certain extent. By identifying how disinformation circulates across multiple societal nodes, from online platforms and media outlets to citizen groups and criminal networks, the project highlighted the importance of multi-stakeholder coordination. Its dissemination activities targeted not only policymakers and law enforcement, but also the scientific community, civil society and technologists, ensuring that the insights produced were shared across the very ecosystem they sought to protect. In doing so, the project reinforced the foundation for a whole-of-democracy response.

Apart from that, some lessons have been learned that can guide the FERMI project's key target group, namely LEAs, in their effort to fight disinformation(-induced crime). One of the central technical challenges in the analysis of disinformation campaigns is the lack of temporal coherence across datasets, particularly when data is collected from different sources or social media platforms. Disinformation narratives often evolve rapidly, exploiting emerging events or societal tensions, and are frequently reshaped as they move across time and audiences. Without time-stamped, synchronised datasets, it becomes difficult to reconstruct the sequence of events, identify the original sources of a campaign, or understand how specific messages were amplified. In the context of FERMI, this limited the ability to generate a continuous and reliable timeline of campaign activity, which is crucial for linking online disinformation to offline radicalisation or coordinated threat behaviour.

Compounding this issue is the fragmentation of data across the various social media platforms, each with distinct formats, access policies, user behaviours and moderation practices. A disinformation campaign may begin on fringe platforms or encrypted channels, then migrate to mainstream social media in order to reach a wider audience. However, due to inconsistent data access and varying degrees of openness across platforms, FERMI's technical tools faced difficulty in capturing the full lifecycle and spread of a given narrative. This siloed view hinders the ability to understand how narratives mutate across ecosystems, which actors are driving the transition between platforms, and when mitigation measures (such as content takedowns or fact-checking labels) are applied and whether they are effective.

Finally, the lack of standardised metadata (e.g., timestamps, engagement metrics, identifiers linking content across platforms, etc.) posed a major barrier to inter-platform correlation and campaign mapping. Even when partial data was available, aligning content temporally and thematically across sources required significant manual effort or complex inference. These inconsistencies reduced the effectiveness of AI-driven approaches that depend on temporal and semantic coherence to detect coordinated inauthentic behaviour or narrative shifts. Addressing this challenge will require not only technical advances but also improved cooperation from social media platforms and policy frameworks that enable consistent, cross-platform access to relevant disinformation data for trusted research and security applications.

The difficulty of integrating disinformation analysis tools into existing intelligence workflows used by LEAs and national security bodies was another stumbling block. These workflows are often rigid, highly regulated and rely on legacy systems which are not designed for the processing of complex, high-volume, open-source data such as the ones found on social media platforms. Integrating AI-driven insights into such environments requires careful alignment with operational timelines and institutional trust frameworks, none of which can be achieved through stand-alone or siloed systems.

Interoperability issues also stemmed from the heterogeneity of tools and data formats used across different institutions. Disinformation analysis typically involves a mixture of structured and unstructured data, multiple languages and evolving taxonomies of threat indicators and campaign patterns. Without a common data model or standardised interfaces, exchanging and reusing insights across systems becomes highly inefficient. Within FERMI, efforts to develop flexible, modular architectures and APIs revealed that technical compatibility alone is insufficient. Semantic

Funded by the European Union

alignment, shared ontologies and contextual interpretation mechanisms are equally necessary to ensure that analytical outputs are actionable to downstream users.

Furthermore, operational integration was hindered by the limited explainability and transparency of many AI components. Intelligence analysts and LEA personnel often require a clear understanding of how risk indicators are generated or how certain conclusions are drawn, particularly when these feed into policy decisions. Black-box models, even if accurate, are unlikely to gain institutional acceptance unless accompanied by robust interpretability features, traceability mechanisms and auditability. The lessons from FERMI strongly indicate that future systems must prioritise not only interoperability at the technical layer but also trustworthiness and usability at the human layer, in order to ensure meaningful adoption within real-world intelligence environments.

Another key challenge is sharing highly sensitive data that capture the disinformation campaigns of political extremists and the crime landscape. FERMI adopted a privacy-by-design approach to data management, storage and security, ensuring that sensitive datasets remain protected throughout the entire lifecycle of AI model development and deployment. Particular attention was given to the handling of crime occurrence data provided by LEAs, with safeguards in place to ensure secure storage, controlled processing and full compliance with applicable legal and ethical standards.

To address the challenge of training AI models on highly sensitive and non-transferable datasets, the project employed Swarm Learning – a decentralised machine learning paradigm that enables models to be trained locally within the secure infrastructure of each participating data provider. Instead of transferring raw data, only model parameters are shared across the swarm network. This significantly reduces the risk of data leakage, unauthorised access or potential re-identification of individuals, as no personal or operational data ever leaves the premises of the LEAs.

The integration of Swarm Learning with GDPR-compliant data governance mechanisms, including secure communication protocols and strict access controls, ensured that the AI development process is both secure and ethically sound. This approach safeguarded individual privacy and the operational integrity of LEAs while simultaneously establishing a replicable and scalable framework for future AI applications requiring the processing of sensitive or regulated data across distributed environments. Drawing from the project's results, decentralised privacy-preserving AI architectures, such as Swarm Learning, should be strongly encouraged.

Against this backdrop, technical solutions must be both interoperable and privacy-preserving. Future systems should incorporate decentralised learning architectures (e.g. federated or swarm learning) that enable cross-institutional collaboration without requiring centralised access to sensitive datasets. At the same time, adherence to common data models, metadata standards and secure social media data is essential to ensure seamless integration into national and EU-level platforms. The EU should invest in the standardisation and certification of such infrastructures to support trustworthy, scalable adoption across the security ecosystem.

Funded by
the European Union

8 Conclusions

The implementation of FERMI's CDES strategy has been instrumental in amplifying the project's reach and supporting its overall objectives. Throughout the project, each area of the strategy contributed to increasing FERMI's visibility, credibility, and long-term relevance.

Dissemination activities enabled the project to position itself within key academic and policy debates. Through peer-reviewed publications, conference presentations, and collaborative research outputs, FERMI shared its findings with expert communities working on disinformation, digital risk, and public security. The communication strategy ensured the project maintained a consistent and recognisable presence across digital channels. Platforms such as LinkedIn, combined with the website, newsletters, and event participation, were used effectively to engage diverse stakeholders and grow a FERMI community of interest.

The success of these efforts is well underscored by FERMI's impressive (over-)compliance with the KPIs set by the GA. Not only could all of the communication and dissemination KPIs be reached, in many cases, they could even be greatly exceeded. The project's website and social media engagement is a case in point. The numbers for website visits, accesses, social media activities were significantly higher than required. Similarly, dissemination in the form of publications and events was a lot more successful than necessary to meet the KPIs.

Exploitation planning identified several project results with strong potential for future uptake by law enforcement agencies, researchers, and technology developers. This deliverable captures those efforts and highlights concrete steps taken to support impact beyond the project's duration. Finally, FERMI's engagement in standards and interoperability discussions ensured that its technical and conceptual developments aligned with European ethical, legal, and data governance frameworks. Some further recommendations as to how best to advance the fight against disinformation(-induced) crime from an LEA standpoint are shared, too.

As the project concludes, FERMI leaves behind a solid foundation for future work on disinformation risk assessment and AI-supported decision-making in public security. Its outreach activities have helped position the project's results for further research, uptake, and integration, ensuring that the knowledge generated continues to benefit stakeholders and inspire new efforts across the Horizon Europe community and beyond.

Funded by
the European Union

List of References

Aziani, A., Lo Giudice, M. V., Yazdi, A.S. "Conspiracy to Commit: Information Pollution, Artificial Intelligence, and Real-World Hate Crime," European Journal on Criminal Policy and Research (2025). https://zenodo.org/records/15719116

Dogtiev, A. (2024, February 6). Push Notifications Statistics (2023). Bussiness of apps. https://www.businessofapps.com/marketplace/push-notifications/research/push-notifications-statistics/

Evangelatos, S., Papadakis, T., Gousetis, N., Nikolopoulos, C.D., Troulitaki, P., Dimakopoulos, N., Bravos, G., Lo Giudice, M.V., Shadman, A., Aziani, A. "The Nexus Between Big Data Analytics and the Proliferation of Fake News as a Precursor to Online and Offline Criminal Activities," 2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy (2023). https://zenodo.org/records/10948598

Evangelatos, S., Veroni, E., Efthymiou, V., Nikolopoulos, C. "Modeling Disinformation Spread in Social Networks: Phase Transitions and Mean-Field Analysis," ACM Transactions on the Web (2025). https://zenodo.org/records/16410586

Giglio, F., “Moderazione di contenuti illegali e social media scraping: Vincoli in materia di privacy e protezione dei dati nel trattamento dei dati disponibili al pubblico da parte delle forze dell’ordine,” i-lex. 16(2) (2023). doi: 10.6092/issn.1825-1927/18870.

Grant Agreement.

Lindner, J. (2023, December 20). Push Notification Statistics: Market Report & Data. Gitnux. https://gitnux.org/push-notification-statistics/

Lo Giudice, M.V., Yazdi, A.S., Aziani, A., Evangelatos, S., Gousetis, N., Nikolopoulos, C. "Informative (Dis)information: Exploring the Correlation Between Social Media Disinformation Campaigns and Real-World Criminal Activity," 2024 5th International Conference in Electronic Engineering, Information Technology & Education (EEITE), Chania, Greece (2024). https://zenodo.org/records/13759707

Nitsch, H., FERMI. Campus Polizei. Magazin der Hochschule für den öffentlichen Dienst in Bayern, Fachbereich Polizei (2023). https://zenodo.org/records/11440040

Varela da Costa, J., Bogea Gomes, S. & Mira da Silva, M., “Fake News: a conceptual model for risk management,” Humanit Soc Sci Commun 11, 625 (2024). https://zenodo.org/records/11354424

Varela da Costa, J., Dongo, D.F., Mira da Silva, M, “Using MCDA to select countermeasures against fake news,” Journal of Information, Communication and Ethics in Society 23 (1) (2025). https://doi.org/10.1108/JICES-07-2024-0089

Varela da Costa, J., Fernandes, A., & Mira da Silva, M., “Fake news and risk management: a systematic literature review,” Journal of Risk Research, 27(12) (2024). https://doi.org/10.1080/13669877.2025.2466530

Varela da Costa, J., Mira da Silva, M, “Countermeasures against fake news: a Delphi study,” Transforming Government: People, Process and Policy 19 (2) (2025). https://doi.org/10.1108/TG-10-2024-0258

Funded by the European Union

Annex A – Recommendation for Operational Standard for the Ethical and Responsible Use of Technology by Law Enforcement

Part 1

Executive Summary

This Operational Standard (Version 3) provides clear, actionable good practices for Law Enforcement Agencies (LEAs) to ensure the ethical, responsible, and lawful use of technology, including AI and data analytics. It integrates key requirements from the EU AI Act, GDPR, and the EU Charter of Fundamental Rights, aiming to embed ethical considerations into daily operations. The standard supports LEAs in identifying risks, ensuring compliance, and fostering public trust and accountability.

Definition of Terms

- **Artificial Intelligence (AI):** A machine-based system designed to operate with varying levels of autonomy that, for explicit or implicit objectives, infers from inputs to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.
- **Data Analytics:** Systematic techniques and methods applied to data in order to extract information and insights, including the identification of patterns and trends, for the purpose of supporting decision-making.
- **Deployers:** Natural or legal persons, public authorities, agencies, or other bodies that use an AI system under their authority, other than when used in a purely personal and non-professional context.
- **Ethical Foresight:** A systematic and forward-looking process for identifying and assessing potential ethical, societal, and fundamental rights impacts of technologies across the lifecycle, and for incorporating preventive and mitigative measures into design, deployment, and operation.
- **Fundamental Rights Impact Assessment (FRIA):** A structured assessment used to identify, evaluate, and mitigate risks that the development, deployment, or use of a technology may pose to rights and freedoms protected under the European Union fundamental rights framework.
- **High-Risk AI Systems:** AI systems whose intended purpose or context of use is likely to pose a significant risk to health, safety, or fundamental rights and that are therefore subject to enhanced controls including risk management, technical documentation, human oversight, quality management, logging, accuracy, robustness, and cybersecurity requirements.
- **Human Oversight:** Functions performed by competent human operators to understand, supervise, and where necessary intervene in the operation of a system, including the ability to override or stop it, to ensure outcomes remain compliant with policies, legal obligations, and ethical requirements.
- **Implementers:** Law enforcement officers, technical personnel, or contracted staff who are directly responsible for configuring, operating, monitoring, and maintaining technologies and for applying established procedures and controls.

Funded by the European Union

- **Managers:** Officials with governance, oversight, or compliance responsibilities relating to the selection, procurement, risk management, deployment, monitoring, and continuous improvement of technologies.
- **Personal Data:** Any information relating to an identified or identifiable natural person. An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier.
- **Prohibited AI Systems:** AI systems classified as presenting an unacceptable level of risk due to their inherent threat to fundamental rights or safety and which are therefore not permitted to be placed on the market, put into service, or used.
- **Technology Lifecycle:** The complete set of phases that a technology undergoes, typically including conception, requirements definition, design, development, integration, verification and validation, procurement, deployment, operation, monitoring, maintenance, modification, and decommissioning including data disposition.

Abbreviated Terms

- **AI:** Artificial Intelligence
- **CE:** Conformité Européenne (European Conformity Marking)
- **COBIT:** Control Objectives for Information and Related Technology
- **DPIA:** Data Protection Impact Assessment
- **ECAT:** European Centre for Algorithmic Transparency
- **ECA:** European Court of Auditors
- **EDPB:** European Data Protection Board
- **EDPS:** European Data Protection Supervisor
- **EIF:** European Interoperability Framework
- **EU:** European Union
- **FRIA:** Fundamental Rights Impact Assessment
- **GDPR:** General Data Protection Regulation (Regulation (EU) 2016/679)
- **HLEG:** High-Level Expert Group on Artificial Intelligence
- **ICO:** Information Commissioner's Office (UK)
- **ITIL:** Information Technology Infrastructure Library
- **LED:** Law Enforcement Directive (Directive (EU) 2016/680)
- **LEA:** Law Enforcement Agency
- **MCC-AI:** Model Contractual Clauses for AI Procurement
- **SOP:** Standard Operating Procedure
- **SRPP:** Socially Responsible Public Procurement
- **XAI:** Explainable Artificial Intelligence

1. Introduction

Purpose:

This Operational Standard provides Law Enforcement Agencies (LEAs) with actionable guidance - good

Funded by the European Union

practices - to ensure the responsible, ethical, and lawful use of technology — including AI, data analytics, and monitoring tools — throughout their lifecycle. It supports compliance with the European AI Act, GDPR, and fundamental rights frameworks while operationalising ethical principles into daily practice.

The purpose is to provide a comprehensive account of good practices to support public officials without an IT background in:

1. Navigating today's technological landscape - understanding what are possible risk factors and how to identify them as well as good practices to address them, throughout the technology lifecycle
2. Preparing for compliance assessment
3. Implementing high-level international guidelines

The purpose is not to provide a step-by-step guide on "how to use technology." This would be unfeasible due to the difference between technologies. Besides this, it is also not desirable and counterproductive because each and every technology must be assessed in their specific context. Otherwise, the very purpose of technology ethics would be defied: instead of assessing the peculiarity of application and identifying harm and unintended consequences with disproportionate effects, one single solution would be applied to everyone thereby causing the harm (and disproportionate effects) that it was meant to protect from.

Intended Users:
- Managers: Responsible for enforcing these standards, overseeing governance and compliance.
- Implementers: Officers and technical staff tasked with applying the standards in operations.

Scope:
Covers all technology and data systems used by LEAs for decision support, from design to deployment, maintenance, and decommissioning. It focuses first and foremost on technology and data systems developed for crime prevention for real-life crime spurring from online disinformation. for decision support, from design to deployment, maintenance, and decommissioning.

Governance Alignment:
This standard integrates best practices from COBIT 2019 and ITIL 4 frameworks to operationalise governance, risk management, and service delivery in LEAs.

2. Links with the Existing Legal Framework

This section outlines the foundational EU legal instruments governing technology use by law enforcement.

EU AI Act (Regulation (EU) 2024/1689)

The EU AI Act regulates AI systems based on their risk. Law enforcement applications often fall into "high-risk" or "unacceptable risk" categories.[2]

- **Implementation:** Prohibitions on unacceptable risk AI systems apply from February 2025. Obligations for high-risk systems apply 36 months after entry into force (around August 2026).[3]

General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679)

The GDPR sets principles for personal data processing: lawfulness, fairness, transparency, purpose limitation, data minimization, accuracy, storage limitation, integrity, and confidentiality.[18] It defines six legal bases for

Funded by the European Union

processing.[19] Data Protection Impact Assessments (DPIAs) are mandatory for high-risk processing.[1] Data subject rights include information, access, rectification, erasure, restriction, and objection.[18]

Law Enforcement Directive (LED) (Directive (EU) 2016/680)

The LED specifically governs personal data processing by competent authorities for criminal offense prevention, investigation, detection, or prosecution, or execution of criminal penalties.[21] It aligns with GDPR, ensuring fundamental rights while enabling police work, emphasizing necessity, proportionality, data minimization, and safeguards for international data transfers.[21]

EU Charter of Fundamental Rights

The Charter provides overarching protection. Articles 7 (right to private life) and 8 (right to protection of personal data) are central. Any limitation on these rights must adhere to necessity and proportionality (Article 52).[21] The Charter also reinforces non-discrimination.[30]

These frameworks form a cohesive legal ecosystem, requiring a layered approach to compliance. Ethical guidelines often anticipate future legal requirements, necessitating continuous ethical foresight and adaptation by LEAs.[1] A core challenge is balancing law enforcement effectiveness with fundamental rights, requiring robust necessity and proportionality assessments and strong oversight.[3]

3. Core Ethical Principles and Governance Foundations

All technology use by LEAs must adhere to foundational ethical principles, linked explicitly to governance structures.

Principle	Expanded Good Practice	Relevant EU Legal Framework/Guideline	Specific Reference
Fundamental Rights Respect	Conduct mandatory Fundamental Rights Impact Assessments (FRIAs) for all new high-risk technology deployments, explicitly addressing potential impacts on vulnerable groups and ensuring mechanisms for redress. Ensure privacy-by-design and default.	EU Charter Art. 7 & 8, EU AI Act, LED, EDPB Guidelines on Facial Recognition Technology, EDPS Quick-Guide to Necessity and Proportionality	21
Human Oversight	Implement "Human-in-command" oversight, ensuring human operators retain ultimate decision-making authority, understand system	EU High-Level Expert Group (HLEG) on AI, Ethics Guidelines for Trustworthy AI, EU AI Act Art. 14	28

Funded by the European Union

	impacts, and have the ability to override or interrupt AI system operations.		
Proportionality and Necessity	Utilize the EDPS 8-step quick-guide for assessing compatibility of measures impacting fundamental rights with the EU Charter, including defining purposes, identifying less intrusive alternatives, and balancing interests.	EDPS Quick-Guide to Necessity and Proportionality, ICO guidance on DPIAs	22
Fairness and Non-Discrimination	Implement robust testing and continuous monitoring for potential biases throughout the entire AI system lifecycle, using diverse datasets and algorithmic fairness metrics, while acknowledging broader governance and legal requirements.	ICO Guidance on AI and Data Protection: Fairness, Bias, and Discrimination, EDPB Guidelines on Facial Recognition Technology	30
Transparency	Implement Explainable AI (XAI) techniques to provide human-interpretable reasons for AI-driven decisions in high-risk scenarios, tailored for both expert and lay audiences. Maintain comprehensive, accessible documentation and logs.	EU AI Act Art. 13, 12, DSA Art. 40, European Centre for Algorithmic Transparency (ECAT), ICO Guidance on AI	2

Table 4 - Core ethical principles and governance foundations

Funded by the European Union

4. Roles and Responsibilities

This section details specific actions and accountabilities for managers and implementers.

Managers

Managers establish and enforce governance structures, approve risk assessments, procurement, and compliance audits, and ensure clear accountability.[1] They lead the development and review of ethical AI policies, championing ethical awareness and data protection by design. They approve resources for DPIAs and FRIAs, ensuring findings inform strategic decisions, and oversee human oversight mechanisms, training, and resources.[1]

Implementers

Implementers (officers, technical staff) maintain data quality, document processes, and validate AI outputs. They follow human oversight and escalation protocols and participate in training.[1] They actively identify and mitigate algorithmic bias through bias assessments. They diligently log AI system operations (inputs, outputs, human interventions) for auditability. Strict adherence to data handling protocols (anonymization, pseudonymization, consent) is mandatory, with clear reporting for inconsistencies or breaches. They provide feedback for continuous improvement.[1]

Accountability is shared, with managers setting the framework and implementers executing procedures. The EU AI Act delineates obligations for "providers" and "deployers," requiring a quality management system with an "accountability framework".[13]

5. Data Governance and Protection

This section details robust practices for managing and protecting data, ensuring GDPR and LED compliance.

Data Governance Area	Detailed Good Practice	Relevant EU Legal Framework/Guideline	Specific Reference
Data Bias	Proactively assess AI training data for imbalances and historical discrimination, implementing techniques like data balancing and continuous monitoring for bias throughout the AI system lifecycle.	ICO Guidance on AI and Data Protection: Fairness, Bias, and Discrimination, EU AI Act Art. 10	2
Data Lineage & Logging	Implement structured logging with consistent formats and correlation IDs to link related	EU AI Act Art. 11, 12	2

Funded by the European Union

	events, capturing precise timestamps, database versions, and input data (with privacy considerations) for high-risk AI systems.		
DPIAs/FRIAs	Integrate Fundamental Rights Impact Assessments (FRIAs) as a mandatory, iterative component of the impact assessment process for all new high-risk technology deployments, informing design choices and mitigation strategies.	EDPS guidance on DPIAs, EDPB Guidelines on Facial Recognition Technology, ICO guidance on DPIAs	1
Access Control	Establish clear policies and training for personnel on data access protocols, ensuring access is granted only on a need-to-know basis and for specified, legitimate purposes, alongside role-based access controls.	GDPR Art. 5(1)(f), EU AI Act Art. 15	15
Anonymization/Pseudonymization	Implement cryptographic pseudonymization techniques with strict separation and access controls for re-identification keys, clearly defining the "pseudonymisation domain" to control the scope of re-identifiability.	EDPS guidance on DPIAs, EDPB Guidelines 01/2025 on Pseudonymisation, ICO guidance on anonymization and pseudonymization	1
Consent Management	Recognize that consent is often not the most appropriate legal basis	GDPR Art. 6(1), 7, EDPB/EDPS joint letter	1

Funded by the European Union

	for law enforcement processing due to power imbalances. Where used, ensure strict adherence to GDPR Article 7, including clear communication and easy withdrawal.	on GDPR record-keeping	
Data Retention	Establish and enforce a six-month retention period for data categorization, requiring the erasure of datasets older than six months if not properly categorized, especially for data with no established link to criminal activity.	EDPS mandates for Europol, EDPB guidance on storage limitation	18

Table 5 - Data governance and protection

Managers approve and monitor data governance frameworks.[1] Implementers follow data handling protocols and report inconsistencies.[1]

6. Risk Assessment and Management

This section details identifying, assessing, and mitigating risks, especially for AI systems, following the EU AI Act's risk-based approach.

AI Act Risk Category	Law Enforcement Examples/Implications	Key Requirements/Prohibitions	Relevant EU AI Act Article/Snippet ID
Unacceptable Risk	Social scoring of individuals; Real-time remote biometric identification in public spaces (with very limited exceptions); Predictive policing based solely on profiling; Untargeted scraping of facial images.	**Banned outright**; Severe penalties for non-compliance (up to €40M or 7% of annual turnover).	2
High Risk	AI systems for evidence evaluation; Suspect identification; Risk assessment in criminal	Conformity assessment; Quality management system; Robust data governance; Technical	2

	justice; Biometric identification/categorization; Critical infrastructure safety components.	documentation; Meticulous logging; Human oversight; Accuracy, robustness, cybersecurity; EU declaration & CE marking; Registration in EU database; Post-market monitoring.	
Limited Risk	Chatbots for public queries; AI systems generating text/images (e.g., deepfakes).	Transparency obligations: Disclose AI interaction; Label AI-generated content.	3
Minimal Risk	Spam filters; AI in video games.	Largely unregulated under the Act; Minimal obligations.	4

Table 6 - Risk assessment and management

Managers oversee risk frameworks and approve mitigation strategies.[1] Implementers perform assessments and adhere to mitigation protocols.[1]

7. Transparency and Explainability

This section outlines good practices for ensuring transparent and explainable technology use, especially for AI-driven processes.

Aspect of Transparency	Detailed Good Practice	Relevant EU Legal Framework/Guideline	Specific Reference
Decision Process Transparency	Develop comprehensive technical documentation for high-risk AI systems, detailing system design, capabilities, limitations, and regulatory compliance efforts.	EU AI Act Art. 11	2
Data Handling Transparency	Implement structured logging with consistent formats and correlation IDs, capturing precise timestamps, database versions, and input data for high-risk AI systems to ensure traceability.	EU AI Act Art. 12	2
Algorithmic Explainability	Implement Explainable AI (XAI) techniques to provide human-	EU AI Act Art. 13, ECAT, ICO Guidance on AI	2

	interpretable reasons for AI-driven decisions in high-risk scenarios, tailored for both expert and lay audiences.		
Auditability	Ensure mandatory, immutable logging capabilities for high-risk AI systems, automatically recording events throughout the system's lifetime for compliance verification and post-market monitoring.	EU AI Act Art. 12, ECIIA guidance for internal auditors on the AI Act	2
Public Communication	Provide publicly accessible, non-technical summaries of system purpose, data usage, and impacts, ensuring clarity and avoiding misleading information.	ICO guidance on transparency notices, EU AI Act requirements for generative AI	2

Table 7 - Transparency and explainability

Managers ensure transparency policies are enforced and documented.[1] Implementers maintain thorough documentation and support audit activities.[1]

8. Human Oversight and Decision-Making

This section provides detailed good practices for ensuring meaningful human oversight over technology, particularly AI systems, in law enforcement.

Oversight Type/Area	Detailed Good Practice	Relevant EU Legal Framework/Guideline	Specific Reference
Human-in-command	Implement a "Human-in-command" approach where human operators retain ultimate decision-making authority, including the ability to override AI outputs or halt system operations.	EU AI Act Art. 14, Ethics Guidelines for Trustworthy AI	28

Validation of AI Outputs	Establish clear protocols for human validation of all AI outputs that may lead to investigations or legal actions, ensuring reviewers have competence and authority to detect anomalies and interpret outputs.	EU AI Act Art. 14, Ethics Guidelines for Trustworthy AI	2
SOPs for Review & Fallback	Define and implement SOPs that clearly delineate mandatory human intervention points, interaction methods (e.g., HMI tools), and fallback plans for system disagreements or failures.	EU AI Act Art. 14, Ethics Guidelines for Trustworthy AI	28
Human Competence	Provide continuous, specialized training programs for all personnel on AI capabilities, limitations, ethical implications, and human oversight procedures, fostering "AI literacy" across the organization.	EDPB Support Pool of Experts (SPE) projects, Ethics Guidelines for Trustworthy AI	28
Intervention Mechanisms	Design AI systems with built-in mechanisms enabling human operators to intervene, correct errors, override decisions, or safely interrupt system operations (e.g., "stop" button).	EU AI Act Art. 14, Ethics Guidelines for Trustworthy AI	28

Table 8 - Human oversight and decision-making

Managers ensure SOPs are in place and training is provided.[1] Implementers execute human review processes and escalate issues.[1]

9. Procurement and Development Controls

This section outlines good practices for integrating ethical and responsible considerations into technology procurement and development for law enforcement.

Control Area	Detailed Good Practice	Relevant EU Legal Framework/Guideline	Specific Reference
Ethical Evaluation	Incorporate EU Commission's MCC-AI High-Risk clauses into all procurement contracts for high-risk AI systems, ensuring vendor compliance with data governance, transparency, and human oversight requirements.	EU Commission's "Public Procurement Guidance," EU Commission Model Contractual Clauses for AI Procurement (MCC-AI)	[1]
Third-Party Audits	Mandate third-party conformity assessments for high-risk AI systems in procurement contracts, specifying that the market surveillance authority acts as the notified body for law enforcement systems.	EU AI Act Art. 43, EU AI Act Art. 19	[2]
Stakeholder Inclusion	Actively engage with civil society organizations, community representatives, and subject-matter experts from diverse backgrounds during design, development, and testing to identify biases and ensure representativeness.	Ethics Guidelines for Trustworthy AI, ICO Guidance on AI Bias, EU Commission guidance on social procurement	[30]
Collaboration & Engagement	Establish formal channels for public engagement and feedback, providing	European Ombudsman's transparency principles, EU Commission	[1]

Funded by the European Union

	non-technical summaries of system purposes and impacts, and fostering co-design and testing with local communities.	guidance on strategic public procurement	
Socially Responsible Procurement	Integrate social and human rights criteria into procurement processes, scrutinizing the supply chain of technology providers for compliance with labor rights and ethical sourcing.	EU Public Procurement Directives, EU Commission guidance on strategic public procurement, Socially responsible public procurement (SRPP)	49

Table 9 - Procurement and development controls

Managers define procurement standards and monitor vendor compliance.[1] Implementers verify adherence and report concerns.[1]

10. Monitoring, Evaluation and Auditing

This section details comprehensive good practices for continuous monitoring, periodic evaluation, and robust auditing of technology systems, including clear decommissioning protocols.

Area of Review	Detailed Good Practice	Relevant EU Legal Framework/Guideline	Specific Reference
System Performance	Implement a continuous post-market monitoring system for high-risk AI systems, leveraging automated log analysis to detect model drift, performance degradation, and emerging issues in real-time.	EU AI Act Art. 72	2
Social Impact	Conduct periodic reviews that assess the broader societal impacts of technology, including its effects on vulnerable groups and the	European Court of Auditors' (ECA) Performance Audit guidelines	1

Funded by the European Union

	achievement of intended positive outcomes.		
Legal Compliance	Ensure continuous monitoring of legal compliance, using robust audit trails and automated dashboards for compliance metrics and anomaly detection.	EU AI Act Art. 12, ECIIA guidance for internal auditors on the AI Act	2
Decommissioning	Establish detailed, mandatory protocols for system decommissioning, including a final impact assessment and secure data deletion/anonymization based on criteria like persistent bias or legal changes.	EDPB guidance on storage limitation, Ethical AI system decommissioning best practices, EU AI Act implications for prohibited AI	11
Record-Keeping & Audit Trails	Maintain immutable and comprehensive logs of system updates, training data changes, and policy amendments, accessible for audit purposes and root cause analysis.	EU AI Act Art. 11, 12	2

Table 10 - Monitoring, evaluation and auditing

Managers lead evaluation processes and authorize decommissioning.[1] Implementers support monitoring activities and maintain accurate records.[1]

11. Public Engagement and External Transparency

This section details good practices for fostering public trust and accountability through proactive engagement and transparent communication.

Aspect of Transparency	Detailed Good Practice	Relevant EU Legal Framework/Guideline	Specific Reference
Citizen Engagement Channels	Establish accessible channels for citizens (e.g., public consultations, online portals) to raise	European Ombudsman's transparency principles, NCSL guidance on social media as a law	1

	concerns or seek clarification about technology use and actively solicit feedback from civil society and academia.	enforcement communication tool	
Non-Technical Summaries	Provide publicly accessible summaries of system purpose, data usage, and impacts in concise, plain language, explaining AI system functionality, data use, and potential fundamental rights impacts.	ICO guidance on transparency notices, EU AI Act Art. 13	2
Collaboration with Stakeholders	Engage in ongoing dialogue with civil society, academia, and other stakeholders to understand concerns, gather diverse perspectives, and inform the development of ethical guidelines and responsible technology solutions.	Ethics Guidelines for Trustworthy AI, European Ombudsman	28
Adherence to Transparency Principles	Ensure all public communications are current, accurate, and avoid misleading information, including clearly labeling AI-generated content (e.g., deepfakes) and proactively publishing information on AI systems and risk classifications.	European Ombudsman's transparency principles, EU AI Act requirements for generative AI	1

Table 11 - Public engagement and external transparency

Managers oversee public communication strategies and transparency compliance.[1] Implementers facilitate engagement activities and document feedback.[1]

12. Interoperability and Cross-Border Cooperation

This section details good practices for fostering interoperability and cross-border cooperation in ethical and responsible technology use by law enforcement.

- **Promote Adherence to EU Data Space Principles:** Encourage common data formats, standards, and protocols for seamless information exchange and services between public administrations across borders. The European Interoperability Framework (EIF) provides recommendations for improving interoperability governance.[1]
- **Foster Cooperation with Other EU Law Enforcement Bodies:** Support initiatives like the European Investigation Order and promote secure information sharing capacities between Member States, Europol, and other security agencies to address cross-border challenges.[32]
- **Share Technical Best Practices and Lessons Learned:** Actively participate in forums for knowledge exchange on effective and ethical technology deployment, including digital forensics tools and lawful interception measures, while safeguarding cybersecurity and fundamental rights.[32]

Managers establish cooperation agreements and monitor interoperability.[1] Implementers implement interoperability protocols and report issues.[1]

Part B

13. Continuous Monitoring

The operational standard requires continuous monitoring to ensure its ongoing relevance and effectiveness. This includes regular reviews against new EU legislation, emerging ethical considerations, and practical implementation challenges. Feedback from both internal and external stakeholders is crucial for identifying areas for improvement and adaptation, ensuring the standard remains a living document.

Conclusions and Recommendations

Version 3 of the Operational Standard is a critical step towards embedding ethical principles and legal compliance in LEA operations. Key recommendations include:

1. **Embrace Layered Compliance:** Establish an interdisciplinary "Ethics and Technology Governance Committee" to oversee integrated compliance strategies and holistic impact assessments.
2. **Prioritize Ethical Foresight & AI Literacy:** Implement mandatory, continuous training on AI ethics and data protection. Invest in Explainable AI (XAI) tools for interpretability and auditability.
3. **Reinforce Human-in-Command:** Design AI systems with intuitive human-machine interfaces that facilitate human intervention and override capabilities. Develop SOPs for mandatory human intervention and fallback plans.
4. **Implement Dynamic Data Lifecycle Management:** Develop and enforce comprehensive data lifecycle management policies, including automated data categorization, clear retention schedules, and secure deletion protocols.
5. **Leverage Procurement for Ethics:** Integrate EU Commission's Model Contractual Clauses for AI procurement (MCC-AI) into all relevant contracts. Conduct due diligence on vendors' ethical AI practices.

Funded by the European Union

6. **Transition to Continuous Monitoring:** Invest in automated logging and analytical tools for real-time monitoring of AI system performance, bias detection, and security vulnerabilities. Establish clear decommissioning protocols.
7. **Foster Proactive Public Engagement:** Develop accessible citizen engagement channels and provide concise, plain-language summaries of AI system purposes and impacts. Collaborate with civil society and academic experts.

By implementing these recommendations, LEAs can operationalize ethical and responsible technology use, ensuring compliance with EU legal frameworks, strengthening public trust, and upholding fundamental rights.

Bibliography

1. EU AI Act: Summary & Compliance Requirements - ModelOp, accessed on June 27, 2025, https://www.modelop.com/ai-governance/ai-regulations-standards/eu-ai-act
2. EU AI Act: first regulation on artificial intelligence | Topics | European ..., accessed on June 27, 2025, https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence
3. The AI Act requires human oversight | BearingPoint USA, accessed on June 27, 2025, https://www.bearingpoint.com/en-us/insights-events/insights/the-ai-act-requires-human-oversight/
4. AI Act | Shaping Europe's digital future - European Union, accessed on June 27, 2025, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
5. Explainable AI and the EU AI Act: Unlocking Trust and Compliance Before It's Too Late, accessed on June 27, 2025, https://aiireland.ie/2025/04/11/explainable-ai-and-the-eu-ai-act-unlocking-trust-and-compliance-before-its-too-late/
6. What is the Artificial Intelligence Act of the European Union (EU AI Act)? - IBM, accessed on June 27, 2025, https://www.ibm.com/think/topics/eu-ai-act
7. Navigating New Regulations for AI in the EU - AuditBoard, accessed on June 27, 2025, https://auditboard.com/blog/eu-ai-act
8. What is the EU AI Act? - Holistic AI, accessed on June 27, 2025, https://www.holisticai.com/blog/eu-ai-act
9. Article 5: Prohibited AI Practices | EU Artificial Intelligence Act, accessed on June 27, 2025, https://artificialintelligenceact.eu/article/5/
10. Decoding the EU AI Act & ensuring Prohibited AI is no longer used: recent guidance, accessed on June 27, 2025, https://legalbriefs.deloitte.com/post/102kbng/decoding-the-eu-ai-act-ensuring-prohibited-ai-is-no-longer-used-recent-guidanc
11. European Commission Guidelines on Prohibited AI Practices under the EU Artificial Intelligence Act | Inside Privacy, accessed on June 27, 2025, https://www.insideprivacy.com/artificial-intelligence/european-commission-guidelines-on-prohibited-ai-practices-under-the-eu-artificial-intelligence-act/
12. Article 17: Quality Management System | EU Artificial Intelligence Act, accessed on June 27, 2025, https://artificialintelligenceact.eu/article/17/
13. EU AI Act: Implications for Log Management Systems and ... - Logdy, accessed on June 27, 2025, https://logdy.dev/blog/post/eu-ai-act-implications-for-log-management-systems-and-compliance

Funded by the European Union

14. Article 15: Accuracy, Robustness and Cybersecurity | EU Artificial ..., accessed on June 27, 2025, https://artificialintelligenceact.eu/article/15/
15. arxiv.org, accessed on June 27, 2025, https://arxiv.org/html/2502.16184v1#:~:text=The%20Purpose%20of%20Art.,-15%20AIA&text=15(1)%20AIA%20states%20that,for%20high%2Drisk%20AI%20systems.
16. New Guide Prepares Internal Auditors for EU AI Act Compliance - BABL AI, accessed on June 27, 2025, https://babl.ai/new-guide-prepares-internal-auditors-for-eu-ai-act-compliance/
17. Data protection basics, accessed on June 27, 2025, https://www.edpb.europa.eu/sme-data-protection-guide/data-protection-basics_en
18. Consent - General Data Protection Regulation (GDPR), accessed on June 27, 2025, https://gdpr-info.eu/issues/consent/
19. Process personal data lawfully | European Data Protection Board, accessed on June 27, 2025, https://www.edpb.europa.eu/sme-data-protection-guide/process-personal-data-lawfully_en
20. EDPB adopts the final version of Guidelines on facial recognition ..., accessed on June 27, 2025, https://trilateralresearch.com/data-protection/edpb-adopts-the-final-version-of-guidelines-on-facial-recognition-technology-in-the-area-of-law-enforcement
21. Step 4: Assess necessity and proportionality | ICO, accessed on June 27, 2025, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/dpia-tools/online-retail/step-4-assess-necessity-and-proportionality/
22. ARTICLE 29 DATA PROTECTION WORKING PARTY WORKING PARTY ON POLICE AND JUSTICE - Garante Privacy, accessed on June 27, 2025, https://www.garanteprivacy.it/garante/document?ID=1799276
23. Police Directive - European Data Protection Supervisor, accessed on June 27, 2025, https://www.edps.europa.eu/data-protection/our-work/subjects/police-directive_en
24. Guidelines 01/2023 on Article 37 Law Enforcement Directive - European Data Protection Board, accessed on June 27, 2025, https://www.edpb.europa.eu/system/files/2023-09/edpb-guidelines_202301_art_37_led_en_2.pdf
25. The EDPS quick-guide to necessity and proportionality | European ..., accessed on June 27, 2025, https://www.edps.europa.eu/data-protection/our-work/publications/factsheets/edps-quick-guide-necessity-and-proportionality_en
26. EDPB adopts its first report under the EU-U.S. Data Privacy ..., accessed on June 27, 2025, https://www.edpb.europa.eu/news/news/2024/edpb-adopts-its-first-report-under-eu-us-data-privacy-framework-and-statement_en
27. ETHICS GUIDELINES FOR TRUSTWORTHY AI, accessed on June 27, 2025, https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf
28. Ethics Guidelines for Trustworthy AI (Chapter 16) - The Cambridge Handbook of Lawyering in the Digital Age, accessed on June 27, 2025, https://www.cambridge.org/core/books/cambridge-handbook-of-lawyering-in-the-digital-age/ethics-guidelines-for-trustworthy-ai/E9C6E735730B68941E61D6D7DE549894
29. What about fairness, bias and discrimination? | ICO, accessed on June 27, 2025, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-do-we-ensure-fairness-in-ai/what-about-fairness-bias-and-discrimination/

Funded by the European Union

30. Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement - European Data Protection Board, accessed on June 27, 2025, https://www.edpb.europa.eu/system/files/2023-05/edpb_guidelines_202304_frtlawenforcement_v2_en.pdf
31. Commission presents Roadmap for effective and lawful access to data for law enforcement, accessed on June 27, 2025, https://home-affairs.ec.europa.eu/news/commission-presents-roadmap-effective-and-lawful-access-data-law-enforcement-2025-06-24_en
32. EDPB adopts final version of Guidelines on facial recognition technology in the area of law enforcement - European Union, accessed on June 27, 2025, https://www.edpb.europa.eu/news/news/2023/edpb-adopts-final-version-guidelines-facial-recognition-technology-area-law_en
33. EU's Guidelines for Trustworthy AI: A Reliable Framework for Edtech Companies | Avallain, accessed on June 27, 2025, https://www.avallain.com/blog/eu-guidelines-for-trustworthy-ai-a-reliable-framework-for-edtech-companies
34. Article 14: Human Oversight | EU Artificial Intelligence Act, accessed on June 27, 2025, https://artificialintelligenceact.eu/article/14/
35. AI Ethics in Law Enforcement - Number Analytics, accessed on June 27, 2025, https://www.numberanalytics.com/blog/ai-ethics-law-enforcement
36. European Centre for Algorithmic Transparency - European ..., accessed on June 27, 2025, https://algorithmic-transparency.ec.europa.eu/index_en
37. AI and Privacy: Balancing Technology and Compliance in Law Enforcement - Veritone, accessed on June 27, 2025, https://www.veritone.com/blog/ai-and-privacy-balancing-technology-and-compliance-in-law-enforcement/
38. Anonymisation and pseudonymisation - Data Protection Commission, accessed on June 27, 2025, http://www.dataprotection.ie/en/dpc-guidance/anonymisation-pseudonymisation
39. Guidelines 01/2025 on Pseudonymisation - European Data ..., accessed on June 27, 2025, https://www.edpb.europa.eu/system/files/2025-01/edpb_guidelines_202501_pseudonymisation_en.pdf
40. EDPB and EDPS Support GDPR Record-Keeping Simplification Proposal, accessed on June 27, 2025, https://www.hunton.com/privacy-and-information-security-law/edpb-and-edps-support-gdpr-record-keeping-simplification-proposal
41. EDPS mandates Europol to delete data with respect to individuals ..., accessed on June 27, 2025, https://www.privacyrules.com/edps-mandates-europol-to-delete-data-with-respect-to-individuals-with-no-established-link-to-a-criminal-activity/
42. EU: EDPS orders Europol to erase data concerning individuals with no established link to a criminal activity | News | DataGuidance, accessed on June 27, 2025, https://www.dataguidance.com/news/eu-edps-orders-europol-erase-data-concerning
43. EU Data Enforcement Sweep: Are GDPR-Covered Entities Complying Properly with Data Subjects' Right of Erasure? - Jackson Lewis, accessed on June 27, 2025, https://www.jacksonlewis.com/insights/eu-data-enforcement-sweep-are-gdpr-covered-entities-complying-properly-data-subjects-right-erasure
44. Master the AI Act: ECIIA's Essential Guide for Internal Auditors | ECIIA, accessed on June 27, 2025, https://www.eciia.eu/2025/01/master-the-ai-act-eciias-essential-guide-for-internal-auditors/

Funded by
the European Union

45. EDPB publishes final version of guidelines on data transfers to third country authorities and SPE training material on AI and data protection, accessed on June 27, 2025, https://www.edpb.europa.eu/news/news/2025/edpb-publishes-final-version-guidelines-data-transfers-third-country-authorities-and_en
46. EDPB Finalizes Guidelines on Data Transfers to Third Country Authorities and Training Materials on AI and Data Protection - Hunton Andrews Kurth LLP, accessed on June 27, 2025, https://www.hunton.com/privacy-and-information-security-law/edpb-finalizes-guidelines-on-data-transfers-to-third-country-authorities-and-training-materials-on-ai-and-data-protection
47. Commission updates Model Contractual Clauses for AI procurement ..., accessed on June 27, 2025, https://thelens.slaughterandmay.com/post/102kbhf/commission-updates-model-contractual-clauses-for-ai-procurement
48. Public procurement - European Commission, accessed on June 27, 2025, https://single-market-economy.ec.europa.eu/single-market/public-procurement_en
49. Article 43: Conformity Assessment | EU Artificial Intelligence Act, accessed on June 27, 2025, https://artificialintelligenceact.eu/article/43/
50. Conformity Assessments under the EU AI Act: A step-by step guide, accessed on June 27, 2025, https://www.aigl.blog/conformity-assessments-under-the-eu-ai-act-a-step-by-step-guide/
51. arXiv:2403.07904v3 [cs.CY] 19 Feb 2025, accessed on June 27, 2025, https://arxiv.org/pdf/2403.07904
52. Socially responsible public procurement - European Commission, accessed on June 27, 2025, https://single-market-economy.ec.europa.eu/single-market/public-procurement/strategic-procurement/socially-responsible-public-procurement_en
53. Law Enforcement and Technology: Using Social Media | Congress.gov, accessed on June 27, 2025, https://crsreports.congress.gov/product/pdf/R/R47008
54. Analyzing the Impact of Digital Technologies on Police Legitimacy in Enschede - University of Twente Student Theses, accessed on June 27, 2025, http://essay.utwente.nl/104846/1/Mikhalka_MA_BMS%20.pdf
55. The European Ombudsman | Fact Sheets on the European Union, accessed on June 27, 2025, https://www.europarl.europa.eu/factsheets/en/sheet/18/the-european-ombudsman
56. Assessment of the implementation of the human rights clause in international and sectoral agreements | European Parliament, accessed on June 27, 2025, https://www.europarl.europa.eu/RegData/etudes/IDAN/2023/702586/EXPO_IDA(2023)702586_EN.pdf
57. A guide to taking account of social considerations in public procurement - EUR-Lex - European Union, accessed on June 27, 2025, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021XC0618(01)
58. Public procurement in the EU - EUbusiness.com | EU news ..., accessed on June 27, 2025, https://www.eubusiness.com/competition/public-procurement-in-the-eu/
59. Full article: 'We do not see the problem, but agree with your solution ..., accessed on June 27, 2025, https://www.tandfonline.com/doi/full/10.1080/07036337.2025.2491625
60. Ethical AI & Secure SDLC: A Leader's Guide to Building Trust - V2Solutions, accessed on June 27, 2025, https://www.v2solutions.com/blogs/ethical-ai-secure-sdlc-guide/

Funded by the European Union

61. European Interoperability Framework, accessed on June 27, 2025, https://cartool-ec.github.io/B2B-intra-community-transactions-e-invoicing-reporting/3d54d4b1/elements/id-fe52ab5636bf473086037219ab348c7f.html
62. 1. Introduction to the "European Interoperability Framework", accessed on June 27, 2025, https://interoperable-europe.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/1-introduction-european-interoperability-framework
63. European Interoperability Framework - Wikipedia, accessed on June 27, 2025, https://en.wikipedia.org/wiki/European_Interoperability_Framework

Funded by
the European Union

Annex B - Recommendation for Operational Standard on the Use of Technologies to tackle Disinformation by Law Enforcement Agencies

1. Introduction

The present document sets out an operational standard to be offered to Law Enforcement Agencies (LEAs) to provide disinformation identification and analysis services. This standard sets out the steps LEAs need to take to pick up disinformation and to understand the impact on security and safety of civilians. This standard establishes a framework for effectively tackling disinformation by LEAs, including what kinds of data should be gathered to be used in artificial intelligence (AI) and big data technologies, protocol for their management, storage and security and procedures for LEAs to support their decision-making processes with these tools by including the challenges of human rights.

Due to the increasing number of information and communication technologies that can spread disinformation, it is important to have international operational standards that provide a common understanding for the tackling of disinformation. This document is intended to enhance already existing security standards by adding a focus relevant to disinformation.

The increasingly widespread disinformation phenomenon, and the growing complexity of ICT systems, can make it challenging for LEAs to tackle disinformation and achieve compliance with the various applicable laws. LEAs can prevent uncertainty and distrust from arising by handling disinformation issues properly and effectively.

Use of this document will:

- aid in the design, implementation, operation and maintenance of ICT systems that handle disinformation;
- spur innovation solutions to enable the handling of disinformation within ICT systems; and
- improve LEA's disinformation programmes through the use of best practices.

The framework provided within this document can serve as a basis for additional disinformation standardisation initiatives, such as for example:

- the implementation and use of specific disinformation handling technologies;
- overall disinformation management;
- human rights risk assessments related to disinformation.

In terms of methodology, in order to develop this operational standard, specific questions regarding the issues dealt herein were sent to LEAs in writing in a questionnaire circulated by INTRA (that included other issues covered by FERMI other than standardisation). 8 responses were received in writing on 31 July 2025, which were incorporated into this standard. Furthermore, feedback was sought from LEAs during the webinar that took place on 11 September 2025.

2. Scope

This document provides a framework which is intended to guide LEAS define their disinformation identification and analysis within an ICT environment by:

- specifying a common disinformation terminology;
- defining the LEA actors and their roles in tackling disinformation;
- establishing a framework for effectively tackling disinformation by LEAs;
- providing references to known disinformation principles, including relevant human rights obligations.

Funded by the European Union

This document is applicable to LEAs and organisations involved in specifying, procuring, architecting, designing, developing, testing, maintaining, administering, and operating information and communication technology systems or services where disinformation controls are required.

3. References

3.1 Normative references

Normative references are not applicable in the present document.

3.2 Informative references

The following referenced documents are not necessary for the application of the present document, but they assist the user with regards to the area of disinformation:

a.3.1 Policy Framework

a.3.1.1 European Commission, Action Plan against Disinformation (Joint Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, 2018),

a.i.1.5 European Commission, Tackling Online Disinformation: A European Approach (Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM/2018/236),

a.i.2.5 European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, on the European Democracy Action Plan, Brussels 3.12.2020, COM (2020) 790 final,

a.i.3.5 European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, on Defence of Democracy, Strasbourg, 12.12.2023 COM(2023) 630 final,

a.i.4.5 European Commission, A multi-dimensional approach to disinformation - Report of the independent High Level Group on fake news and online disinformation, 2018.

a.i.5.5 UN Office of the High Commissioner for Human Rights, UN Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework, 2011.

a.i.6.5 UNESCO Guidelines on the Governance of Digital Platforms, 2023.

a.i.7.5 OECD Going Digital Toolkit, Policy Note, “Disentangling untruths online: Creators, spreaders and how to stop them”, 2022 (“OECD Recommendations on Combating Disinformation”, 2022).

a.1.4 Legal Framework

a.1.1.4 Consolidated text: Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users’ rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws (Text with EEA relevance)Text with EEA relevance,

a.1.2.4 Charter of Fundamental Rights of the European Union 2012/C 326/02,

a.1.3.4 Council of Europe, European Convention on Human Rights, as amended by Protocols Nos. 11, 14 and 15, ETS No. 005, 4 November 1950,

Funded by the European Union

a.1.4.4 Regulation (EU) 2016/679 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 and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation),

a.1.5.4 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 (Law Enforcement Directive),

a.1.6.4 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act),

a.1.7.4 Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act) (Text with EEA relevance)

a.1.8.4 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act),

a.1.9.4 Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA

a.1.10.4 Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law,

a.1.11.4 Regulation (EU) 2021/784 of the European Parliament and of the Council of 29 April 2021 on addressing the dissemination of terrorist content online,

a.1.12.4 Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights.

a.1.13.4 Convention on Cybercrime, Budapest, 23 November 2001, European Treaty Series (ETS) No. 185.

a.1.14.4 International Covenant on Civil and Political Rights, 999 U.N.T.S. 171 (1966).

a.2.4 Standards Framework

a.2.1.4 ISO/IEC 29100:2024, Information technology – Security Techniques – Privacy,

a.2.2.4 ISO/IEC 27000:2018, Information technology – Security techniques -Information security management systems – Overview and vocabulary,

a.2.3.4 ISO 27701:2019, Security techniques – Extension to ISO/IEC 27001 and ISO/IEC 27001 for privacy information management – Requirements and guidelines,

a.2.4.4 ISO 31000:2018, Risk Management – Guidelines,

a.2.5.4 ISO Guide 73:2009, Risk Management – Vocabulary,

a.2.6.4 ETSI TR 104 062 v.1.2.1 (2024-07), prepared by the AI group, “Securing Artificial Intelligence; Automated Manipulation of Multimedia Identity Representations”,

a.2.7.4 Journalism Trust Initiative, Workshop Agreement CWA 17493,

a.2.8.4 Coalition for Content Provenance and Authenticity (C2PA) Specifications.

a.3.4 Codes Framework

a.3.1.4 European Union, Code of Conduct on Disinformation 2025,

a.3.2.4 EDMO (Draft) Code of Conduct on Access for Platform-to-Researcher Data Sharing,

Funded by the European Union

a.3.3.4 European Code of Standards for Independent Fact-Checking Organisations,
a.3.4.4 International Fact-Checking Network Code of Principles.

4. Definition of Terms

For the purposes of this document, the following terms and definitions apply (based on the ISO and IEC terminology databases for use in standardisation):

Accountability:

The duty to provide an explanation and justification for actions, with criteria and level of detail varying based on the context. It also encompasses liability for sanctions if performance is unsatisfactory, which can be legal, political, or even social.

Algorithm:

In social media networks and services, algorithms are rules, signals and data that govern the platform's operation. These algorithms determine how content is filtered, ranked, selected and recommended to users.

Artificial Intelligence:

Discipline concerned with the building of computer systems that perform tasks requiring intelligence when performed by humans.

Audit:

Systematic, independent and documented process for obtaining audit evidence and evaluating it objectively to determine the extent to which the audit criteria are fulfilled.

Bias:

Systematic difference between true (or accepted) value and measured value.

Data:

Any digital representation of acts, facts or information any compilation of such acts, facts or information, including in the form of sound, visual, or audio-visual recording (see section a.1.7.4, Regulation (EU) 2023/2854, Data Act, Article 2(1).

Data Protection Impact Assessment:

An assessment of the impact of the envisaged processing operations on the protection of personal data. It has to be carried out before personal data is processed when such processing is likely to result in a high risk to the rights and freedoms of individuals.

Deployer:

A natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity (see section a.1.8.4, AI Act, Article 3(4).

Disinformation:

Misleading content deliberately spread to deceive people, or to secure economic or political gain and which may cause public harm.

Disinformation control:

Measure that treats disinformation risks by reducing their likelihood or their consequences.

Funded by the European Union

Disinformation stakeholder:

Natural or legal person, public authority, agency or any other body that can affect, be affected by, or perceive themselves to be affected by a decision or activity related to disinformation.

Fundamental Rights Impact Assessment:

An assessment of the potential impact of an AI system on the rights of any individual that might be affected by the operation of that system. It is a risk assessment; therefore it focuses on risk management, not risk elimination.

Human Rights Risk Assessment:

Process to identify, analyse, evaluate and document human rights-related risks and their impacts, in order to manage risk and to mitigate or prevent adverse human rights impacts and legal infractions.

Law Enforcement Authority (LEA):

(b) any public authority competent for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security; or

(b) any other body or entity entrusted by Member State law to exercise public authority and public powers for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security

(see section a.1.8.4, AI Act, Article 3(45).

Misinformation:

A piece of information that lacks veracity and that could mislead people.

Personal data:

Any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person (see section a.1.4.4, Regulation (EU) 2016/679, Article 4(1)

Provider:

A natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge (see section a.1.8.4, AI Act, Article 3(3).

Proportionality:

Measures taken by institutions must be appropriate and necessary to achieve a legitimate objective, while also respecting fundamental rights and freedoms. It involves a balancing act between the pursuit of public interest and the potential impact on individuals or entities.

Risk:

The effect of uncertainty on objectives.

Risk Assessment:

Overall process of risk identification, risk analysis and risk evaluation (see section a.2.5.4.4, ISO Guide 73:2009, 3.4.1).

Funded by
the European Union

Risk Management:

Coordinated activities to direct and control an organisation with regards to risk (see section a.2.5.4, ISO Guide 73:2009, 2.1).

Uncertainty:

The state, even partial, of deficiency of information related to, understanding or knowledge of, an event, its consequence, or likelihood.

5. Abbreviated Terms

For the purposes of the present document, the following abbreviations apply:

AI	Artificial Intelligence
LEA	Law Enforcement Agency
ICT	Information and Communication Technology
ISMS	Information Security Management System

6. Core Ethical and Legal Principles

The principles described in this document were derived from existing principles developed by a number of countries and international organisations. These principles should be used to guide the design, development, and implementation of methods to tackle disinformation. They can also be used as a baseline in the monitoring and measurement of performance, benchmarking and auditing aspects of disinformation handling in an organisation

The guiding ethical and legal principles of this operational standard are:

Accountability:

The duty to provide an explanation and justification for actions, with criteria and level of detail varying based on the context. It also encompasses liability for sanctions if performance is unsatisfactory, which can be legal, political, or even social. Adhering to the accountability principle involves:

- Documenting and communicating as appropriate all disinformation-related policies, procedures and practices;
- Assigning to a specified individual within the organization (who can in turn delegate to others in the organization as appropriate) the task of implementing the disinformation-related policies, procedures and practices;
- Providing suitable training for the personnel tasked with dealing with disinformation;
- Notifying all stakeholders about breaches of disinformation-related legal obligations as required in some jurisdictions and depending on the level of risk;
- Allowing an aggrieved individual access to appropriate and effective sanctions and/or remedies, when a breach of disinformation-related obligations has occurred.

Fairness:

A core principle which involves treating all individuals or groups in similar situations without discrimination based on characteristics like sex, race or religion.

Fundamental Rights:

The basic rights and freedoms that apply to everyone within the European Union, irrespective of their background. They are enshrined in the Charter of Fundamental Rights of the European Union (see section a.1.2.4) and the European

Convention on Human Rights of the Council of Europe (see section a.1.3.4). These rights encompass a wide range of freedoms, including the right to life, freedom of thought, freedom of expression, and freedom of movement.

Necessity:

A justification for limiting fundamental rights or breaching an international obligation when it is deemed necessary to achieve a legitimate objective of general interest. It has been interpreted in such a way to encompass the principle of proportionality.

Non-Discrimination:

A core principle of EU law which aims to allow all individuals an equal and fair chance to access opportunities available to society. Individuals or groups of individuals which are in comparable situations should not be treated less favourably simply because of a particular characteristic such as their sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation (see section a.1.2.4, Articles 21-26).

Privacy Compliance:

Adhering to the privacy compliance principle involves:

- verifying and demonstrating that the processing meets data protection and privacy safeguarding requirements by periodically conducting audits using internal auditors or trusted third-party auditors;
- having appropriate internal controls and independent supervision mechanisms in place that ensure compliance with relevant privacy law and with their security, data protection and privacy policies and procedures; and
- developing and maintaining privacy risk assessments in order to evaluate whether program and service delivery initiatives involving disinformation comply with data protection and privacy requirements.

Applicable law can provide that one or more supervisory authorities are responsible for monitoring compliance with applicable data protection law. In those cases, adhering to the privacy compliance principle also means cooperating with these supervisory authorities and observing their guidelines and requests.

Proportionality:

Measures taken by institutions must be appropriate and necessary to achieve a legitimate objective, while also respecting fundamental rights and freedoms. It involves a balancing act between the pursuit of public interest and the potential impact on individuals or entities.

Transparency:

The obligation of institutions to act publicly and provide access to their documents and information. This principle ensures open decision-making and allows citizens to access information related to policymaking and spending.

7. Basic Elements of the Disinformation Framework

7.1 Legal definition

According to the 2018 report of the independent high-level group on fake news and online disinformation of the European Commission ("HLEG"), disinformation is the "false, inaccurate or misleading information designed, presented or promoted to intentionally cause public harm or for profit" (see section a.i.4.5, at p. 10).

The Action Plan Against Disinformation 2018 definition is along the same lines as the HLEG one, namely that disinformation is any "verifiable false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public, and may cause public harm" (see section a.3.1.1, at p.1). Public harm includes threats to democratic processes and to public goods, such as Union citizens' health, environment or security (see section a.3.1.1, at p.1). Disinformation does not include inadvertent errors, satire and parody, or clearly identified partisan news and commentary (see section a.3.1.1, at p.1).

There is a distinction between illegal spread of content and disinformation, which lies in legality versus truthfulness although the two can overlap. The Illegal spread of content refers to the unlawful distribution or publication of certain materials, irrespective of whether they are true or false, such as for example publishing classified government documents (unauthorised leaks). Whereas disinformation is about false or misleading information spread deliberately to deceive the public, such as for example the spreading of fake news to influence election results. Disinformation is about intentional deception, and not necessarily illegality as it may be legal in many countries, despite its harmfulness. Some instances of disinformation can be illegal, if for example it leads to defamation or is used in fraud.

7.2 Characteristics of disinformation

The elements of the definition of disinformation, gauged from the HLEG 2018 Report (see section a.i.4.5, at p. 10), and the 2018 Action Plan Against Disinformation (see section a.3.1.1, at p.1), are the following:

1. Factual or misleading nature of the information,
2. Intention of the actors to spread such information they know to be false to obtain economic gain or deceive the public,
3. Public Harm.

Disinformation is not illegal at the EU or the international level, unless it is directly prohibited under the domestic law of a specific country. It has to be explicitly prohibited by a domestic legal system. Nevertheless, 4 types of content are illegal:

1. Child sexual abuse material (see section a.1.9.4, Directive 2011/93/EU),
2. Racist and xenophobic hate speech (see section a.1.10.4, 2008 Framework Decision on combating certain forms of expressions of racism and xenophobia),
3. Terrorist content (see section a.1.11.4, Regulation (EU) 2021/784),
4. Content infringing intellectual property rights (see section a.1.12.4, Directive 2004/48/EC on IPR Enforcement).

7.3 Actors and roles

7.3.1 Managers:

The specific tasks of managers are the following:

1. Establish and enforce governance structures and oversight bodies,
2. Approve risk assessments, procurement decisions, and compliance audits,
3. Ensure clear accountability lines between LEA units and technology providers.

7.3.2 Implementers:

The specific tasks of implementers are the following:

1. Maintain data quality, document operational processes, and validate AI outputs,
2. Follow human oversight procedures and escalation protocols,
3. Participate in training and competency development.

7.4 Recent trends in tackling disinformation by public authorities

Organisations are motivated to tackle disinformation for a multiplicity of reasons: to deal with disinformation spreading online, to protect individuals' privacy, to meet legal and regulatory requirements, to practice corporate responsibility

Funded by the European Union

and to enhance consumer trust. Public authorities are developing their approaches to tackling disinformation, including legislation, coordination, technology and media resilience and literacy. For example, the Digital Services Act (see section a.1.6.4), includes provisions allowing authorities to require platforms to act on disinformation, including political ads transparency and rapid removal of illegal or harmful content.

8. Information & Data Governance

8.1 Types of data and information that should be gathered to be used in AI and big data technologies

8.1.1 Data

The types of data that are covered by this standard are the following:

1. 'Data', which means any digital representation of acts, facts or information any compilation of such acts, facts or information, including in the form of sound, visual, or audio-visual recording (see section a.1.7.4, Regulation (EU) 2023/2854, Data Act, Article 2(1))
2. 'Personal data', which means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person (see section a.1.4.4, Regulation (EU) 2016/679, Article 4(1),
3. "Non-Personal Data" means data other than personal data.

8.1.2 Information

For LEAS, information is a core operational resource, which is comparable to other critical assets that enable public safety, national security, and trust in institutions. Information must be carefully protected against misuse, distortion, or loss. Information can exist in several forms:

1. Digital: stored in electronic files, databases or online platforms.

2. Physical: recorded on paper or other tangible media.

3. Human knowledge: the expertise, judgment and awareness held by personnel.

Information can be shared and transmitted through different channels, including:

1. Physical transfer, such as by couriers or official reports.

2. Electronic systems, such as email, secure networks, or social media platforms.

3. Direct communication, such as briefings, verbal instructions or interviews.

Irrespective of the format or method of transmission, information remains vulnerable to disinformation. LEAs rely on ICT in order to tackle disinformation. For LEAs, the protection of information is about safeguarding internal systems, detecting false narratives, preventing their pred, and maintaining the integrity of reliable information in the digital environment. The ICT tools LEAs use are necessary for:

1. Gathering, analysing and storing intelligence.

2. Monitoring and intercepting digital communications.

3. Protecting sensitive data from cyber threats.

Funded by the European Union

4. Identifying, countering and tracking disinformation online.

5. Ensuring secure retention and lawful disposal of records.8.2 Standards for data collection, quality, processing, retention, and deletion with attention to privacy and legal compliance & Protocol for management, storage and security of data

8.2.1 General

Information can be considered as a law enforcement asset:

1. LEAs routinely collect, analyse, store and share information in support of investigations, intelligence, and public safety operations.

2. Information, together with the people, processes, technologies and networks that manage it, is a vital asset for achieving mission objectives and safeguarding national security.

3. These assets face risks, including disinformation campaigns, cyberattacks, and human error that may disrupt operations.

4. LEAs shall apply appropriate security and risk controls to protect the reliability and integrity of the information they depend on.

Information security in the context of disinformation is crucial. All information used by LEAs is vulnerable to manipulation, unauthorised disclosure, or exploitation. Treating information as an asset involves recognizing its value and applying safeguards to maintain its confidentiality, integrity and availability.

For LEAs, accurate, verified and timely intelligence available to the right personnel is crucial for effective decision-making, operational efficiency and public trust.

Protecting information assets involves a structured approach by LEAs, representing information security management, that includes:

1. Defining and applying consistent information security measures,

2. Ensuring compliance with legal, ethical and regulatory obligations,

3. Preserving institutional credibility against disinformation and manipulations,

4. Continuously assessing and addressing security risks.

In responding to evolving threats, LEAs shall:

1. Monitor and evaluate the effectiveness of current information security controls.

2. Detect and assess emerging risks, including disinformation techniques, such as deepfakes, bot-driven amplification, or falsified digital evidence.

3. Adapt and strengthen protocols to address these new threats and ensure resilience.

Every LEA shall establish clear policies and objectives for information security to coordinate these activities. A structured management system ensures that protective measures are consistently applied, risks are systematically managed, and agencies remain agile in countering evolving disinformation threats.8.2.2 Disinformation

<u>Information Security</u>

Funded by the European Union

Information security protects the confidentiality, integrity and availability of information. Within LEAs, this ensures that intelligence, case data, and communications remain accurate, accessible to the right personnel, and safeguarded from manipulation or unauthorised disclosure. By applying information security, LEAs can minimize the impact of disinformation campaigns that seek to disrupt operations or erode public trust.

Protecting information requires the implementation and management of appropriate security controls. These account for a wide range of threats, including falsified digital evidence, coordinated disinformation networks, insider risks, and system vulnerabilities. The goal is to maintain operational continuity, investigative integrity, and resilience against manipulation.

To maximise effectiveness, information security controls shall be fully integrated into daily LEA activities, from evidence collection and intelligence analysis to inter-agency data sharing and public communication. This integration ensures that reliable and verified information forms the foundation of operations, helping LEAs detect, counter, and neutralise disinformation while preserving legal compliance and maintaining public confidence.

Information Management for Countering Disinformation

For LEAs, information management means overseeing and making informed decisions about how information is handled so that operational goals are achieved and sensitive assets are protected. This includes safeguarding intelligence, evidence and communications from manipulation, distortion, or unauthorised access, particularly in the context of online disinformation.

Information security within this framework is put into practice through the creation and enforcement of clear policies, standard operation procedures and practical guidelines. These provide a consistent approach to verifying, storing, sharing and protecting information. These measures shall be applied uniformly across the LEA, with every officer, analyst and staff member understanding and fulfilling their role in ensuring the accuracy, reliability and security of information that underpins investigations and public safety operations.

Information Security Management System (ISMS)

Information security is best achieved through a structured, risk-based approach. LEAS should establish an information security Managment System (ISMS) that combines:

Policies and procedures governing information handling and verification.
Organisational responsibilities that define clear roles in protecting and validating information.
Technical measures such as secure platforms, forensic tools, and monitoring systems.
Human and physical safeguards to reduce the risk of unauthorised access or exploitation.

The implementation of a robust ISMS is crucial for LEAs to safeguard their information assets. An effective ISMS allows LEAs to:

1, Ensure continuous protection of intelligence, investigative data, and operational information against threats, such as disinformation campaigns and falsified content,
Establish a structured framework for identifying and assessing information security risks, selecting appropriate safeguards, and evaluating the effectiveness of those controls on an ongoing basis.
Continuously enhance security measures, adapting to evolving digital threats, including fake news, deepfakes, and social media manipulation.
Demonstrate compliance with legal, regulatory, and ethical obligations while maintaining public confidence and institutional credibility.

By embedding an ISMS into their operations, LEAs can strengthen their ability to detect, prevent, and respond to online disinformation, ensuring that sensitive information remains accurate, reliable and operationally effective.

Funded by the European Union

LEAs need a structured approach to establish, operate, monitor, and continuously improve an ISMS to protect critical information assets against disinformation. The following are the key steps:

1. Identify critical information assets and security needs: determine which intelligence, operational data, communications, and investigative records are essential. Define the security requirements necessary to protect them from manipulation, unauthorised access, or falsification.
2. Assess and address security risks: analyse potential risks, including online disinformation campaigns. Implement measures to mitigate or manage these risks effectively.
3. Implement appropriate security controls: deploy technical, procedural, and organisational safeguards to manage unacceptable risks, including encrypted communications, digital evidence validation, access controls, and verification workflows for online content.
4. Monitor, maintain and enhance controls: continuously evaluate how well security measures protect critical information. Adjust and improve controls in response to evolving tactics, emerging threats or changes in operational priorities.

These steps shall be repeated continuously, allowing LEAs to adapt to shifting risks, changes in investigative strategies, or emerging disinformation techniques. This approach ensures that intelligence and operational data remain accurate, reliable and secure, supporting informed decision-making and safeguarding public trust.

9. Risk Management

LEAs routinely conduct risk management activities and develop profiles to understand threats to their ICT systems, operations, and information assets. The process for managing disinformation risks involves several key steps:

1. Establishing the context: understand the agency, its technical environment, and all the factors influencing disinformation risk management. These factors may include legal and regulatory requirements, inter-agency agreements, operational objectives, and other relevant considerations.
2. Risk assessment: identify, analyse and evaluate risks to individuals, the LEA, and operations that could arise from disinformation campaigns. This includes understanding how false or manipulated information could adversely affect investigations, public trust, or decision-making.
3. Risk treatment: define strategies to address disinformation, and implement controls to prevent, mitigate or reduce risks. Examples include monitoring online sources, validating intelligence and employing digital forensics to detect manipulated content.
4. Communicaton and consultation: engage with relevant stakeholders, obtain consensus on risk management measures, and communicate risks, controls and procedures to officers, analysis and other personnel responsible for operational decisions.
5. Monitoring and review: continuously track risks and the effectiveness of controls, and refine processes to respond to evolving disinformation threats.

A disinformation impact assessment is a key deliverable of this process. It evaluates the potential effects of new or significantly changed programs, operations, or activities, ensuring compliance with disinformation-related legislation. Such assessments should be integrated into the LEA's broader risk management framework to provide a holistic view of risks and mitigation strategies.

Funded by the European Union

10. Transparency and Explainability

10.1 Explainability

The minimum requirements for making AI systems explainable to internal and external stakeholders, including suspects and courts, according to the AI Act (see section a.1.8.4, Regulation (EU) 2024/1689):

10.1.1 Technical Documentation (Article 11 and Annex IV / XI)

Providers shall maintain detailed documentation, which would eventually support traceability, auditability, and regulator review, including:

1. System architecture, logic, design choices, objectives and assumptions;
2. Training methodologies and testing results;
3. Data provenance, scope, cleaning/processing steps, bias detection processes;
4. Computational resources used (e.g. FLOPs, training time, hardware);
5. Measures for human oversight and intended updates.

10.1.2 Record-Keeping / Logging (Article 12)

With regards to record-keeping, high-risk AI systems shall automatically log operational events (e.g. inputs, outputs, changes), for at least six months (or more if needed). Special logging has to take place for biometric or emotion recognition, such as for example match details, reference databases and verifier identities. This practice supports investigation, post-deployment monitoring and auditability.

10.1.3 Explainability for Deployers (Article 13)

High-risk systems shall be designed in such a way that deployers can interpret outputs and use them appropriately. Providers shall supply clear "Instructions for Use", including:

1. Provider identity and contact details,
2. System characteristics: intended use, performance metrics (accuracy, robustness, cybersecurity), limitations, risk factors,
3. Group-specific performance to surface biases or unequal outcomes,
4. Human oversight tools: methods for users to detect anomalies, override outcomes, human-in-the-loop controls,
5. System changes: planned updates and their effects,
6. Maintenance details: expected lifetime and upkeep needs,
7. Log management information: how logs are stored, accessed, interpreted.

10.1.4 Notice and Disclosure to Individuals (Article 52)

Systems interacting with natural persons (e.g. chatbots, emotion-recognition, biometric categorisation, deepfakes) shall:

1. Clearly inform individuals they are interacting with an AI system, unless it is obvious,
2. Label synthetic content so end-users know it was AI-generated.

10.1.5 Right to Explanation on Adverse Impacts (Article 86)

When a high-risk AI system produces output with legal effects or major impact on health, safety or fundamental rights:

Funded by
the European Union

1. Affected individuals (e.g. suspects in legal contexts) have the right to request meaningful explanations of the decision,
2. Explanations must cover the AI's role in the process and the main elements of how the decision was reached (enough to allow challenge or appeal).

10.2 Transparency

The transparency obligations under the AI Act (see section a.1.8.4, Regulation (EU) 2024/1689) are the following, depending on the risk classification of the AI system and the role of the actor (provider, deployer or affected individual):

10.2.1 General Transparency Obligations (Article 52)

These apply to **all AI systems** that interact with humans, generate content, or process biometric/emotional data:

1. **AI Systems Interacting with Humans:** Users must be informed if they are interacting with an AI system (e.g. chatbots, virtual assistants), with the following exception: if the nature of the AI is obvious and the interaction is purely for entertainment (e.g. digital game characters).
2. **Emotion Recognition and Biometric Categorization:** Individuals must be notified before their biometric or emotional data is processed by an AI system.
3. **Synthetic Content (Deepfakes):** AI-generated content (audio, image, video or text) must be clearly labelled as synthetic unless used in lawful law enforcement operations.

10.2.2 High-Risk AI Systems Transparency (Articles 13, 14)

High-risk systems (e.g., in employment, education, law enforcement, critical infrastructure) have enhanced transparency duties. Providers must supply clear "Instructions for Use" to deployers, which shallbe clear and accessible, including:

1. Intended purpose and performance metrics,
2. Limitations and potential risks (e.g., bias),
3. Human oversight measures,
4. Information on training/testing datasets (if applicable),
5. Expected lifetime and maintenance requirements,
6. Contact information for the provider.

With respect to human oversight, high-risk systems must be designed so that their outputs can be interpreted by humans and there are mechanisms for humans to intervene, override, or stop the system when necessary.

10.2.3 General Purpose AI (GPAI) and Foundation Models (Article 52a / 52b)

For general-purpose AI systems, including large models, providers shall:

1. Publish a summary of training data used, describing types and sources,
2. Disclose if the model is used as a base for high-risk applications,
3. Clearly communicate to downstream users about the model's capabilities, limitations, and acceptable uses.

For GPAI with systemic risk (like GPT-4), additional transparency applies, risk management systems, incident reporting, and model evaluations are required.

10.2.4 Right to Explanation (Article 86)

For individuals affected by high-risk AI decisions (e.g., credit scoring, criminal justice, hiring), they have a **right to meaningful explanation**, including:

Funded by the European Union

1. The role AI played in the decision.
2. The logic or main parameters that influenced the outcome.
3. How to challenge or contest the decision.

10.2.5 Record-Keeping and Logging (Article 12)

High-risk systems shall maintain logs of:

1. Operational events.
2. Inputs/outputs that led to decisions.
3. These logs support traceability and explainability in audits or legal challenges.

11. Framework for Meaningful Human Oversight and Decision-Making over AI-assisted Disinformation Detection and Response

To ensure that any AI-assisted system used in the detection or analysis of disinformation campaigns operates under meaningful human control, with full accountability, transparency and the protection of fundamental rights, the following operational standard is applicable to:

1. AI tools used for social media monitoring, narrative tracking, bot detection, or source analysis
2. Decision-support systems for referrals, public warnings, or criminal investigations linked to disinformation
3. All personnel from LEAs interacting with, supervising, or relying on AI outputs on these domains.

11.1 Human-In-The-Loop (HITL) Required

All AI outputs (e.g. flagged actors, content, networks) must be reviewed by qualified LEA analysts before any enforcement action or escalation. AI systems may assist, but never autonomously initiate:

1. Surveillance measures,
2. Legal referrals or charges,
3. Platform takedown requests,
4. Public communications or alerts.

11.2 Right to Explanation

If an individual or organisation is impacted by an AI-assisted decision, such as for example being flagged as a disinformation actor, they are entitled to know that the AI played a role, and request a plain-language explanation of the key indicators and logic used (see section 9.2.4).

11.3 Override and Escalation

LEA analysts must be able to override AI decisions based on context or new evidence. Ambiguous, sensitive or high-impact cases must be escalated to a senior review board, including legal and ethics advisors.

11.4 Logging and Auditability

All decisions, system alerts, overrides and human actions must be securely logged. Logs must be retained for at least 12 months and be auditable. Within the disinformation context, audits will assess: false positives or negatives, biases, overreach risks (see further section 9.1.2).

11.5 Training and Competency

All LEA staff using or reviewing AI outputs must complete certified training on:

Funded by
the European Union

1. AI system capabilities and limitations.
2. Legal frameworks, including the EU AI Act (see section a.1.8.4), and GDPR (see section a.1.4.4).
3. Freedom of expression standards, including those in the EU Charter (see section a.1.2.4) and ECHR (see section a.1.3.4).

11.6 Transparency to the Public

Public disclosures, where security permits, must include:

1. Use of AI in disinformation detection,
2. Oversight mechanisms and appeal channels,
3. Statistics on false positives and human overrides.

11.7 Implementation Checklist for LEA Enforcement AI Oversight

Category	Requirement
Risk Stratification	☐ AI use cases classified by impact level (informational vs. investigatory)
Review Process	☐ Each AI output has a named human reviewer
Override Mechanism	☐ Interface allows overrides with justification
Appeals Protocol	☐ Impacted individuals can request explanations or corrections
Logs & Audits	☐ All system interactions logged; reviewed quarterly
Bias Monitoring	☐ Tools regularly tested for systemic bias (e.g., political, ethnic)
Staff Training	☐ All reviewers trained and certified
Public Reporting	☐ Transparency reports published biannually

Table 12 - Implementation checklist for LEA enforcement AI oversight

12. Accountability and Responsibility Framework for AI Use in Disinformation Operations by LEAs

This section establishes clear standards of accountability and responsibility for the use of AI systems in the detection, analysis and response to disinformation by LEA, in compliance with the AI Act.. This framework applies to all AI-assisted operations used in social media monitoring and narrative detection, disinformation campaign analysis, coordinated inauthentic behaviour identification, and referral of content or actors to prosecution.

12.1 Accountability Structure based on Roles

12.1.1 AI Provider

The AI Provider has the following responsibilities:

1. To ensure compliance with technical documentation (Article 11, AI Act, see section a.1.8.4)
2. To maintain accurate and complete training and testing data records (Article 10, AI Act, see section a.1.8.4)
3. To implement and document risk management procedures (Article 9, AI Act, see section a.1.8.4).

Funded by the European Union

12.1.2 LEA AI Lead (Accountable Officer)

The Accountable Officer has the following responsibilities:

1. To serve as the legally and operationally responsible party for AI system use,
2. To conduct Fundamental Rights Impact Assessments (FRIAs),
3. To ensure logs, decisions and override records are maintained and auditable.

12.1.3 Analyst / System User

The Analyst has the following responsibilities:

1. To review all AI-generated outputs before action,
2. To use override mechanisms when appropriate,
3. To log every decision, override and referral.

12.1.4 Supervisory Authority Liaison

The Supervisory Authority Liaison has the following responsibilities:

1. To coordinate communication with national and EU oversight bodies,
2. To ensure timely submission of compliance and transparency reports (see section 9.2 on transparency).

12.2 Oversight and Legal Controls

12.2.1 Human Oversight (Article 14, AI Act, see section 10): AI tools must never autonomously initiate legal action or surveillance and all decisions affecting individuals require human approval.

12.2.2 FRIAs are mandatory before deploying AI in any case involving surveillance, referrals or reputational harm. A Fundamental Rights and Ethics Officer shall oversee AI-related operations.

12.2.3 Documentation: maintain system documentation (model logic, sources, updates) and secure logs for all system operations, retained for a minimum of 12 months.

12.3 Traceability Measures

12.3.1 Decision Logging: to record all system outputs, human decisions, escalations and overrides and to include date, time, system used, analyst ID and rationale.

12.3.2 Update Tracking: document retraining events, model version changes, and threshold changes.

12.3.3 Audit Trails: ensure logs are immutable and available for internal and external audit.

12.4 Bias & Performance Monitoring

12.4.1 Model Evaluation: to conduct periodic audits for demographic, geographic or linguistic bias and to document and correct performance imbalances.

12.4.2 False Positive Monitoring: to maintain statistics on false flags and disputed outcomes and to use appeals and field feedback to guide retraining.

Funded by
the European Union

13. Training and Competency of LEA personnel using or overseeing technologies that aim at preventing or tackling disinformation

The following are minimum training and certification requirements for LEA personnel using or overseeing technologies that aim at preventing or tackling disinformation.

13.1 General Training

All LEA personnel must complete mandatory training on the legal, ethical, and technical aspects of AI use in disinformation operations. Training must cover the EU AI Act, GDPR, fundamental rights, data quality, and algorithmic transparency. In particular, it should aim to train individuals in how to counter algorithmic biases regarding disinformation, as well as handling public inquiries and data subject rights breaches.

The training ought to be delivered by professionals who specifically work on the design of responsible AI use frameworks. This helps to avoid situations where the training is conducted internally by individuals who may not have the necessary expertise.

An example of a critical success factor for an ISMS system is an effective information security awareness, training and education programme, informing all employees and other relevant parties of their information security obligations set forth in the information security policies, standards, etc., and motivating them to act accordingly.

13.2 Certification

Analysts and decision-makers must pass a certification program administered or endorsed by the agency or national supervisory authority. Certification must be renewed every 24 months and documented in personnel records.

13.3 Role-Specific Modules

Specific modules shall be offered according to the role of the person being trained:

1. **AI Leads**: Advanced training in AI governance, risk management, and legal compliance.
2. **Analysts**: Scenario-based training in interpreting AI outputs, recognizing false positives, and executing overrides.
3. **Supervisory Liaisons**: Specialized training on compliance reporting, documentation standards, and inter-agency coordination.

13.4 Continuing Education

This type of training should be offered on a continuous basis. For example, annual refresher courses should be required for all roles. Moreover, updates should be provided immediately following significant changes to law, policy, or system functionality.

13.5 Best Practices regarding Training

The following training best practices should be adopted:

1. Establish internal training curricula covering AI ethics, safety, transparency, and technical skills,
2. Promote interdisciplinary training, combining AI engineering with ethics, law, and social sciences,
3. Encourage participation in workshops, certifications, and professional development programs aligned with standards.

Funded by the European Union

14. Procurement and Development Standards

This section establishes the standards and ethical requirements for the procurement, development and deployment of AI systems aimed at detecting, analysing or countering disinformation with jurisdictions of the EU. It applies both to all departments and personnel involved in AI-related activities and all third-party vendors and contractors supplying AI systems or services.

14.1 Procurement Standards

Vendors must:

1. Demonstrate compliance with the AI Act (see section a.1.8.4) and the GDPR (see section a.1.4.4);
2. Submit documentation on system design, training data and risk mitigation strategies;
3. Allow external audits and provide access for conformity assessments.

Contracts should include:

1. Ethical use clauses prohibiting misuse for censorship or covert influence;
2. Obligations to support post-deployment monitoring and algorithmic audits;
3. Termination provisions for non-compliance with EU law or LEA policy.

Public procurement integrity shall be established with the following measures:

4. All procurements shall follow the EU public procurement legal framework;
5. Pre-award evaluations shall include ethical risk scoring and rights impact scoring.

14.2 Development Standards

The following development standards apply:

1. Training data must be verified, lawful and bias-audited,
2. Synthetic or scraped data from social platforms must be anonymized and legally sourced,
3. Models must avoid targeting lawful speech or disproportionately affecting specific groups.

Systems must be tested in simulated environments before deployment, while also complying with the AI Act (Article 16 on Obligations for Deployers and Articles 17-29 on High Risk AI, EU AI Act, see section a.1.8.4).

AI outputs many not be used autonomously for content takedown or individual profiling, there must be a human-in-the-loop. Moreover, investigative or enforcement decisions require human review.

14.3 Prohibited Uses

The following uses shall be prohibited:

1. AI-based surveillance of journalists, political groups or peaceful activists;
2. Convert AI systems aiming to influence public opinion;
3. Profiling individuals based on political, religious or social view without legal authorization.

Funded by the European Union

15. Evaluation, Monitoring, and Auditing & Governance and Oversight of AI Systems for Disinformation

15.1 Evaluation Mechanisms, Performance Monitoring, Periodic Audits

15.1.1 Evaluation

The evaluation phase (pre-deployment) shall first of all engage in a risk classification. It shall classify systems under the AI Act (using Annex III for categorisation).

The evaluation phase shall next make an FRIA assessing the potential impact on rights such as the freedom of expression, privacy and data protection and non-discrimination, using stakeholder input, including from independent legal experts and civil society.

The evaluation phase shall next engage in a data quality assessment, evaluating:

1. Source integrity (e.g. fact-checked journalism),
2. Representativeness and bias in training and test datasets,
3. Legal provenance of data sets (Articles 5 and 6, GDPR, see section a.1.4.4).

The next phase of the evaluation phase shall be algorithmic transparency evaluation, documenting: model architecture and logic, decision thresholds, explainability features, known limitations.

The final stage of the evaluation phase shall be pre-deployment testing, conducting adversarial testing, false positive/negative analysis, use-case validation under realistic scenarios, and evaluating impacts on different user groups

15.1.2 Monitoring

The monitoring phase (deployment and operational use) shall implement real-time system logging, including alerts triggered, actions taken (e.g. escalation, flagging) and human-in-the-loop interventions.

The monitoring phase shall next set out human oversight protocols, defining roles of reviewers in flag approval, removal recommendation, investigation triggers, and establishing thresholds for mandatory human review.

The monitoring phase shall next engage in impact tracking, monitoring and recording: accuracy rates, disproportionate impacts on protected groups, suppression of lawful speech, public inquiries, complaints or legal actions.

The monitoring phase shall ensure the ability for individuals to access profiling data (Article 15, GDPR, see section a.1.4.4), object to processing (Article 21, GDPR, see section a.1.4.4) and challenge automated decisions (Article 22, GDPR, see section a.1.4.4 and Article 52, AI Act, see section a.1.8.4).

Finally, the monitoring phase shall introduce channels for user feedback, mechanisms for correcting false positives and procedures to retrain models based on validated errors.

15.1.3 Auditing

The auditing phase (post-deployment and lifecycle review) shall conduct annual independent external audits performed by independent AI auditors, covering compliance with the AI Act, GDPR and EU Charter rights.

The auditing phase shall include technical audits, evaluating model drift, performance degradation, adversarial resilience, transparency reports for system outputs.

The auditing phase shall also include ethical and legal audits, reviewing whether the system led to unjustified restriction on expression or access to information, and whether affected individuals were informed and able to contest decisions.

Funded by the European Union

The auditing phase shall also include a public reporting state, publishing regular transparency reports, including system objectives and real-world outcomes, redress cases and how they were resolved, accuracy and bias metrics, and engagement with civil society and oversight bodies.

15.2 Governance and Oversight

Each LEA shall appoint a "Designated Oversight Officer", who will play the role of Fundamental and Ethics Officer responsible for audit coordination and as a liaison to the Data Protection Officer (DPO).

The LEA shall have record-keeping obligations in maintaining records (5-10 years) on: datasets, risk assessments, algorithm changes and public and internal complaints.

Interagency collaboration shall also take place, whereby an LEA shall share findings, methodologies and lessons with other EU LEAs, the EU Agency for Fundamental Rights (FRA), and if applicable, with EUROPOL and ENISA.

16. Public Engagement and Stakeholder Communication in Disinformation Response by LEAs

This section sets out requirements for informing the public, ensuring societal oversight, and enabling public trust for LEAs engaged in tackling disinformation.

16.1 Public Transparency Requirements

LEAs shall make plain-language disclosures by publishing accessible summaries of AI systems used (functions, purposes, limitations), and disclosing training data types and disinformation definitions.

LEAs shall publish transparency reports, indicating the number and types of disinformation incidents handled, system accuracy, error rates, and appeals handled, and a summary of partnerships and data-gathering arrangements.

16.2 Stakeholder Communication Requirements

LEAs shall hold stakeholder workshops prior to implementing any new system (pre-deployment consultation), including digital rights groups, journalists, academia and affected communities.

LEAs shall collect input on potential impacts of AI systems deployed on freedom of expression, political discourse or journalistic activities, using tools like the FRIAs.

16.3 Public Engagement Requirements

LEAs shall designate official spokespersons on disinformation, enabling coordinated communication with fact-checkers, electoral commissions, or health authorities.

LEAs shall also publish official responses to viral falsehoods and describe the evidence countering them.

17. Framework for Interoperability and Cross-Border Cooperation on Tackling Disinformation by LEAs

The recommendations in this section ensure compatibility with international instruments and mechanisms for cross-border cooperation in tackling disinformation by LEAs.

In addition to the EU law framework set out in section 3.2.2, LEAs should also observe:

Funded by the European Union

1. Council of Europe Budapest Convention on Cybercrime (see section a.1.13.4)
2. UN Guiding Principles on Business and Human Rights (UNGPs, see section a.i.5.5)
3. International Covenant on Civil and Political Rights (ICCPR, see section a.1.14.4)
4. UNESCO Guidelines on the Governance of Digital Platforms (2023, see section a.i.6.5)
5. OECD Recommendations on Combating Disinformation (2022, see section a.i.7.5)

The following entities should engage in cross-border disinformation coordination, along with national LEAs:

1. Europol EC3 (European Cybercrime Centre),
2. Interpol Innovation Centre,
3. European Data Protection Board,
4. Council of Europe Cybercrime Committee,
5. OSCE Representative on Freedom of the Media.

18. Redress in Automated Decision-Making

Individuals affected by automated decisions should be clearly informed about their use and potential impact, especially in contexts where AI systems are used to detect, flag, or respond to disinformation.

To ensure transparency and accountability, agencies should provide mechanisms for complaints or redress, in line with GDPR Article 22 and EU AI Act Article 52. Examples include:

1. Online reporting forms for issues arising from automated decisions.

2. Ombudsman services or hotlines to handle concerns about AI errors or misuse.

Agencies should also allow appeals processes, involving independent or external reviewers for decisions that individuals contest, ensuring fairness and impartiality.

Furthermore, whistleblower protections must be in place to encourage internal and external reporting of unethical, illegal, or harmful disinformation practices. This safeguards integrity within the agency while promoting accountability in the management of automated tools and information systems.

Here is an operational guideline for LEAs on handling complaints, appeals, and whistleblower reports related to automated decisions and AI use in disinformation monitoring:

18.1 Informing Affected Individuals

LEAs shall ensure that anyone impacted by automated deicions (e.g. AI-based disinformation detection, content flagging) is notified clearly about the decision, its rational and potential consequences. LEAs shall provide plain language explanations of how automated systems work and the types of decisions they influence.

18.2 Mechanisms for Complaints and Redress

LEAs shall establish multiple channels for complains, such as online forms for reporting AI errors or misuse, deadicated ombuds services for disinformation-related grievances, and hotline or email support for urgent concerns. Complains should be acknowledged promptly, investigated systematically, and resolved within defined timelines. LEAs should keep detailed records of complaints, investigations, and outcomes for accountability and audit purposes.

Funded by
the European Union

18.3 Appeals Process

Individuals should be allowed to appeal automated decisions to external or independent reviewers who were not involved in the original decision. Reviewers should examine the data, logic, and context of automated decisions to determine accuracy and fairness. Reviewers should communicate the outcome of appeals clearly, explaining any corrective actions or system adjustments made.

18.4 Whistleblower Protections

LEAs shall provide secure channels for employees to report unethical or illegal disinformation practices without fear of retaliation. Furthermore, LEAs should enable reporting to independent oversight bodies if internal mechanisms are compromised or ineffective. The identity of whistleblowers should be protected to encourage reporting and maintain trust. LEAs should investigate all whistleblower reports thorougly and take corrective or disciplinary actions where necessary.

18.5 Integration with ISMS and Risk Management

LEAs should embed complain appeal, and whistleblower procedures within the agency's ISMS and broader risk management framework. LEAs should regularly review and update these mechanisms to address emerging AI-related risks, changes in legislation, or new disinformation tactics.

Funded by
the European Union

Annex C - Public Review: Comments on Draft ETSI TR 104 137 V0.0.3 (2025-04), DTR/CYBER-00156 Cyber Security; Human-to-Human Online Preventative Security; H2H – OPS

Organization name	Clause/ Subclause	Paragraph Figure/ Table	Type of comment (General/ Technical/ Editorial)	COMMENTS	Proposed change	RESOLUTION on each comment submitted
Cyprus Organisation for Standardisation (CYS)	2.2 Informative References		Technical	(i) European Commission, Code of Conduct on Disinformation 2025, (ii) European Commission, Action Plan against Disinformation (Joint Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, 2018), (iii) European Commission, Tackling Online Disinformation: A European Approach (Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM/2018/236), (iv) Journalism Trust Initiative, Workshop Agreement CWA 17493:2019.	To add these sources to the list.	

Funded by the European Union

	3.1 Terms			Disinformation control: Measure that treats disinformation risks by reducing their likelihood or their consequences. Disinformation stakeholder: Natural or legal person, public authority, agency or any other body that can affect, be affected by, or perceive themselves to be affected by a decision or activity related to disinformation.	To add these terms to the list.	
	3.3 Abbreviations		Technical	AI: Artificial Intelligence LEA: Law Enforcement Agency ICT: Information and Communication Technology	To add these abbreviations to the list.	
	4.2.1 Religious Intolerance		Technical	Differing from the USA, freedom of expression is a qualified right in the EU, and as such it can be limited in certain circumstances, such as for example to protect other rights, public safety or national security. Examples of limitations on freedom of speech might include laws against hate speech, defamation, or incitement to violence [see Article 10 of the European Convention on Human Rights].	Add this paragraph after the paragraph ending with "in the US constitution".	
	4.2.2		Technical	This term refers to efforts or campaigns that seek to discourage or prevent people from receiving	Add these paragraphs	

	Anti-vaccination promotion			vaccines, oftentimes by spreading misinformation, misleading arguments or unfounded fears. Such efforts can take place through various channels, such as social media, websites and sometimes through public figures, including influencers or celebrities. A key characteristic of anti-vaccination promotion is misinformation and disinformation. For example, it oftentimes involves incorrect or distorted information about vaccines, including, but not limited to: (i) False claims about vaccines, such as for example that they cause autism, despite scientific evidence disproving this link, (ii) Unfounded or exaggerated fears about the side effects of vaccines, which can make people hesitant to get vaccinated, (iii) Claims that vaccines are "unnatural" or that the body is better off without them due to natural immunity.	under this section.	
	4.2.3		Technical	These refer to efforts or activities designed to undermine or disrupt the fairness, transparency, or accuracy of an election process. These can take	Add these paragraphs	

	Election Integrity Attacks			various forms, including disinformation campaigns aiming to confuse or manipulate voters. The overall objective of such attacks is to influence the outcome of an election, erode public confidence in democracy, or disrupt the functioning of the electoral system. Disinformation and misinformation campaigns aiming at attacking election integrity can take the following forms: (i) Disinformation campaigns oftentimes have the objective of confusing voters by spreading false or misleading information. For example, this could entail fabricated claims about voting procedures, fake reports of voter fraud, or exaggerated claims about the outcome of elections. (ii) Social media platforms can be manipulated in such a way to amplify disinformation, create confusion among voters, and promote polarizing narratives. Oftentimes, bot accounts, fake profiles on social media or paid advertisements are used to mislead voters or dissuade them from voting. (iii)There is an increasing use of deepfake videos or fake news websites, which show misleading or	under this section.	

Funded by the European Union

				fabricated content. Deepfake videos can be used to distort the reputation of political candidates by showing them making controversial statements.		
	4.2.4 Attacks on established science		Technical	These refer to efforts to undermine or distort well-established scientific research. One of the forms that such attacks can take are public misinformation campaigns with an aim to create confusion and doubt or weaken public trust in scientifically established conclusions. For example, this can consist of spreading incorrect information on social media about climate change, the safety of vaccines, the effectiveness of certain medical treatments. This can be done by highlighting data that supports a certain view by ignoring a large body of evidence that contradicts them. Attacks on established science can lead to the erosion of public trust in scientific institutions and experts. Spreading misinformation about vaccines or certain medical treatments can lead to the resurgence of preventable diseases and can result in increased mortality or morbidity.	Add these paragraphs under this section.	
	4.2.5 Attacks on the rule of law and		Technical	These refer to efforts to undermine or weaken the foundational systems and structures that ensure fairness, justice and social order within a society.	Add these paragraphs	

	societal institutions			Such attacks often target the principles sustaining democracy, human rights and civil liberties, which in turn damage public trust in institutions, like the judiciary, LEAs, government bodies and the media. These in turn can lead to political, social and economic instability. Disinformation can be used to attack the rule of law and societal institutions. Governments, political groups or other entities oftentimes deliberately spread false or misleading information to control public opinion and undermine the integrity of societal institutions. Disinformation regarding societal institutions instils confusion, division and distrust in the public. Such attacks can also be directed against independent organisations, such as human rights organisations or electoral commissions, when they hold powerful individuals or governments accountable.	under this section.	
	4.3.1 Harassment and Cyber bullying		Technical	There is an overlapping connection between disinformation, online harassment and cyber bullying. It is oftentimes challenging to distinguish between freedom of expression, criticism, and harassment, especially when disinformation takes place. For example, disinformation can be used as a	Add these paragraphs under this section.	

				harassment tactic: through spreading fake or manipulated content (e.g. deepfakes), by damaging someone's reputation or using fake information to incite mob harassment. Furthermore, there is the phenomenon of trolling, which is connected to misinformation campaigns. For example, troll farms or organised groups may spread disinformation along with bullying and harassment to silence dissent or to manipulate public opinion. Targets of such campaigns are oftentimes journalists or activists. Additionally, groups may coordinate to overwhelm a target with fake information, threats or insults.		
	4.3.2 Doxing		Technical	The publication of private or personal information about someone takes place without their consent. Doxing and disinformation are frequently used in tandem to cause harm. For example, disinformation can be used to justify doxing. False information about someone can incite outrage, which in turn is used to justify divulging that person's personal information (e.g. telephone number, home address, family details).	Add these paragraphs under this section.	

Funded by the European Union

				Disinformation renders doxing more dangerous and contributes to its more effective spreading. When divulging personal data is paired with false information, the risk of harassment and social or professional exclusion of the targeted person increases. Reversely, doxing can be used to “prove” disinformation, by releasing fabricated information to support a false claim, such as fake chat logs, manufactured emails, misleading photographs.		
	4.3.3 Swatting		Technical	There is a connection between swatting and disinformation which is rooted in manipulation, malice and abuse of systems of trust, especially those involving emergency services. Disinformation is embedded in the act of swatting as it involves falsely reporting a serious emergency, such as a hostage situation or an active shooter, prompting an aggressive law enforcement response to a person’s home. For example, disinformation can be used to make the swatting appear credible, such as telling the authorities that the attackers are holding hostages or using spoofed caller IDs to back up the false information. Reversely, individuals are often	Add these paragraphs to this section.	

				targeted for swatting after disinformation campaigns portray them as criminals, terrorists or extremists based on false information. Furthermore, there is a link between doxing, harassment, swatting and disinformation. In such cases, swatting is an extreme escalation, following disinformation to build a narrative, doxing to find the location of the victim and then swatting to cause terror or harm.		
	4.3.4 Sexual Grooming			There is an inextricable link between sexual grooming and disinformation because perpetrators use false information to manipulate, deceive, isolate and control their victims. For example, disinformation can be used to manipulate the victim's perception through false promises to build trust and emotional dependence, or even to silence and control the victim. Additionally, disinformation can be used to create false personas, with groomers impersonating someone else by using fake profile information and photos. Another example of the link between disinformation and sexual grooming is when a victim tries to report the abuse, the perpetrators may spread disinformation to discredit them.	Add these paragraphs to this section.	

	4.3.5 Sextortion			There is a connection between sextortion and disinformation in so far as false information is used to deceive, coerce, manipulate or silence victims. For example, perpetrators use disinformation to trick victims into sending explicit content, by using fake identities, deepfakes or AI-generated profiles. Another example is blackmail through false claims, such as for example the perpetrator threatening the victim that he has sent photos to her parents. Yet another example is when perpetrators spread disinformation to isolate or shame, such as by gaslighting victims that nobody will help them.	Add these paragraphs to this section.	
	5.3 Hate and Disinformation Algorithms			Hate and disinformation algorithms refer to the way algorithmic systems amplify or prioritise harmful, hateful or false content. These algorithms incentivize and spread hate and disinformation because they are designed to optimise engagement, attention and profit. For example, false content and hate speech tend to generate more clicks, more comments and more shares and algorithms reward this engagement, even though the content is harmful or false.	Add these paragraphs to this section.	

				Moreover, algorithms show users more of what they already engage with, so if they are interacting with hateful or false content, they are likely to be shown more of that content, which can lead to progressive radicalisation, especially when it comes to young or vulnerable users. Given that platforms rely on automation, harmful content can go viral, before it is flagged and removed. By the time it is removed, algorithms may have already caused harm, such as election interference, vaccine misinformation or mob harassment.		
	5.4 Manipulated AI LLM Information Ingestion		Technical	Manipulated AI LLM Information Ingestion refers to the intentional feeding of false, biased or malicious information into the data pipelines or environments that Large Language Models (LLMs) learn from or are influenced by either during training or post-deployment. If the input data includes manipulated or misleading content, the model learns and reproduces falsehoods or amplifies biases. An example of a dataset that might feed an algorithm are fake news websites designed to look like reputable sources.	Add these paragraphs to this section.	

Funded by the European Union

				Moreover, LLMs can absorb manipulated information post-training, from user interactions, web-browsing tools, APIs or plugins, fine tuning or RLHF (Reinforcement Learning with Human Feedback) pipelines. This can influence the model's factual reliability.		
	6 Preventive and Mitigation Measures		General	The **Code of Conduct on Disinformation 2025** sets out 43 commitments and 128 specific measures relating to disinformation. Consider whether relevant measures can be incorporated into this draft. Consider for example: "Commitment 22: Relevant Signatories commit to provide users with tools to help them make more informed decisions when they encounter online information that may be false or misleading, and to facilitate user access to tools and information to assess the trustworthiness of information sources, such as indicators of trustworthiness for informed online navigation, particularly relating to societal issues or debates of general interest. And:	Consider whether any of the measures set out in the Code of Conduct on Disinformation 2025 are applicable and can be transposed to this section.	

Funded by the European Union

				"Measure 22.6: Relevant Signatories providing trustworthiness indicators by means of voluntary, self-regulatory and certifiable European standards or European standardisation deliverables as defined by European law ('technical standards'), such as the CWA17493:2019 (Journalism Trust Initiative) will: • Develop and revise them based on internationally accepted best-practices and ethical norms; • Make them publicly available and accessible in a non-proprietary, neutral way; • Govern their implementation in line with European Accreditation and EU Regulation (EC) No 765/2008."		

Table 13 - Public Review: Comments on Draft ETSI TR 104 137 V0.0.3 (2025-04), DTR/CYBER-00156 Cyber Security; Human-to-Human Online Preventative Security; H2H – OPS