

Closing the Credibility Gap: Mapping the Evolution of Disinformation and Digital Trust

Closing the Credibility Gap: Mapping the Evolution of Disinformation and Digital Trust has been designed in the context of FERMI (Fake nEws Risk MItigator) [Project 101073980], a Horizon Europe project that studies and attempts to counter the root causes, spread and implications of disinformation and fake news.

This training material is inspired by, and primarily derived from, the insights shared during the FERMI Final Webinar, *'How Disinformation Evolves Narratives, Digital Influence, and Trust'* organised by CONVERGENCE on 10/09/2025.

The webinar featured Katerina Beli (Research Associate and Communication Manager, The Lisbon Council) and Dr Sofia Tipaldou (Assistant Professor of International Relations, Panteion University of Social and Political Sciences) as speakers; their contributions are acknowledged and referenced throughout.

The aim of this resource is to map the evolving dynamics of disinformation and explore how credibility, trust, and legitimacy are constructed, contested, and reshaped in the digital age. By examining both the disinformation lifecycle and the FERMI technological framework, it highlights how a cross-disciplinary approach, combining artificial intelligence, behavioural science, and socio-economic analysis, can bridge the credibility gap and foster resilience in European societies.

This training material complements the FERMI series *'Navigating Disinformation: A Comprehensive Guide'* and *'Digital Trust: A Practical Path to Combating Disinformation and Fostering Resilience'* For, expanding the discussion towards the interaction between narrative manipulation and public trust.

'Disinformation doesn't just distort facts; it mobilises people.'
(FERMI Webinar, 2025)

PARTNERS



SECTION 1: UNDERSTANDING DISINFORMATION – CONCEPTS AND CONTEXT

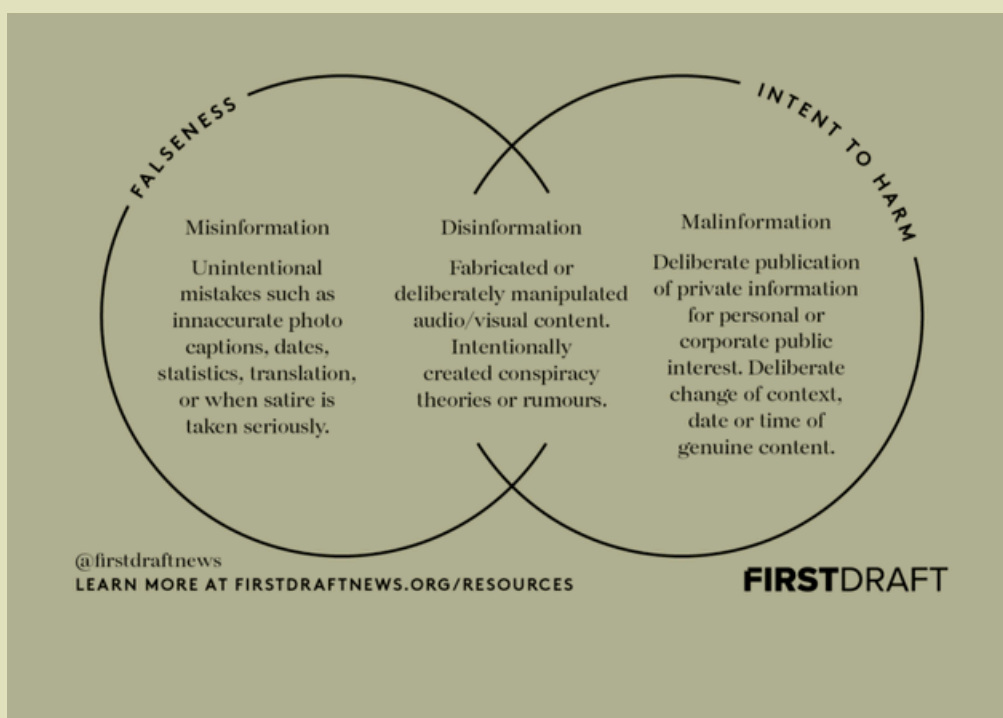
As Dr Sofia Tipaldou underlined in the webinar, disinformation is not a single false post but a strategic process: it is created deliberately to deceive, manipulate, or polarise, often by exploiting emotional and ideological vulnerabilities. To grasp its impact, we consider the broader notion of information disorder, how information is created, shared, and received in the digital age, where misinformation, disinformation, and malinformation interact and evolve.

Why the term ‘information disorder’ rather than ‘fake news’?

‘Fake News’ is misleading and increasingly politicised. Much problematic content is not news (e.g., rumours, memes, manipulated images, targeted ads); some material mixes genuine elements with decontextualisation; and the label is often used to discredit legitimate journalism. Information disorder is a more accurate, comprehensive lens for today’s complex information environment.

Key terms used in the webinar

- **Misinformation:** false or inaccurate information shared without intent to deceive.
- **Disinformation:** false or misleading information shared with intent to deceive or manipulate.
- **Malinformation:** factual information used in ways that cause harm (e.g., weaponised context).



Types of Mis- and Disinformation

According to Claire Wardle's taxonomy (First Draft, 2017–2019), mis- and disinformation exist along a spectrum of intent to deceive and potential to cause harm. The framework distinguishes between several forms, including:

- **satire or parody,**
- **false connection,**
- **misleading content,**
- **false context,**
- **imposter content,**
- **manipulated content, and**
- **fabricated content.**

Recognising these categories helps analysts and citizens identify manipulation techniques, assess intent, and respond appropriately within the evolving information ecosystem.

Why this matters?

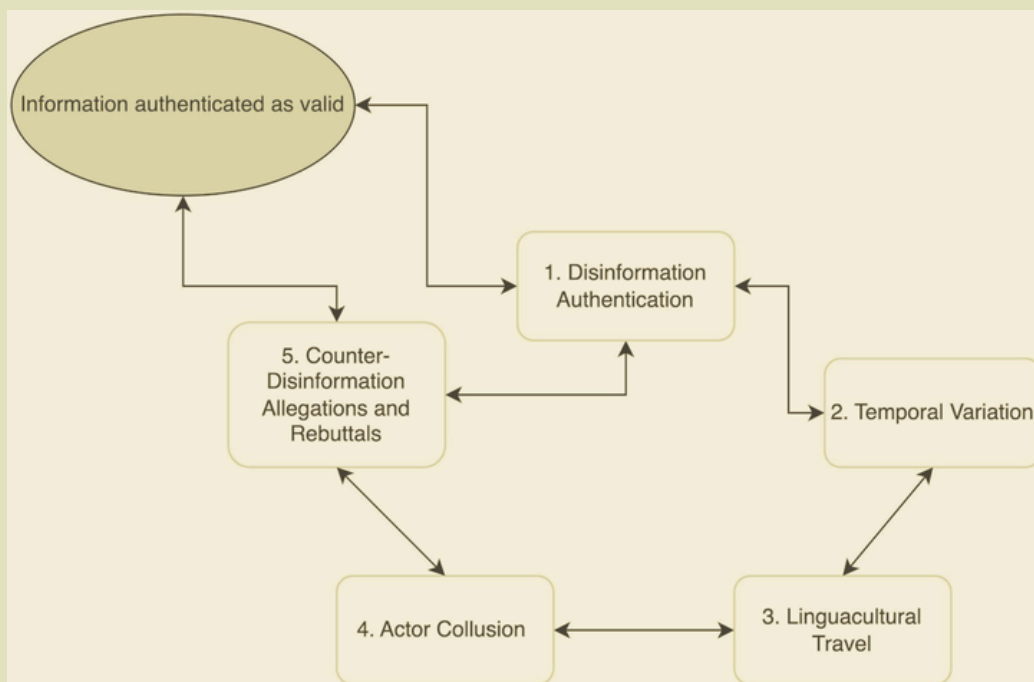
Disinformation operates as a system, combining narrative framing, media manipulation, and selective amplification to achieve political, economic, or ideological goals. For this reason, we must study it as an ecosystem, not as isolated posts. This perspective leads directly to the next section of the webinar, where Tipaldou presented the Disinformation Life-Circle, a process model showing how such narratives evolve and persist.



SECTION 2: THE DISINFORMATION LIFECYCLE – FIVE (5) INTERRELATED STAGES

As presented by Dr Sofia Tipaldou, the Disinformation Life-Circle (Tolz, Hutchings, Kazakov, Tipaldou 2025) offers a process-oriented model that captures how disinformation evolves through a continuous cycle of validation, transformation, and contestation within the information ecosystem.

The Disinformation Life-Circle comprises **five interrelated stages** that describe how false or misleading narratives evolve and persist across the information ecosystem:



- **Disinformation Authentication** – Processes and practices through which actors present, badge, or ‘prove’ selected items as credible, priming them for wider uptake.
- **Temporal Variation** – The same narrative shifts form over time (re-timed, re-framed, selectively updated) to fit new events and maintain salience.
- **Linguacultural Travel** – Narratives move across languages and cultural contexts, picking up locally resonant cues, metaphors, and frames.
- **Actor Collusion** – Coordination (formal or tacit) among outlets, platforms, influencers, or communities to boost visibility and reinforce perceived legitimacy.
- **Counter-Disinformation Allegations and Rebuttals** – Cycles of accusation and response that, paradoxically, can further circulate and entrench the original narrative.

SECTION 3: CASE STUDY INSIGHTS – THE KREMLIN STRATEGY AND THE EU 2024 ELECTIONS

Dr Sofia Tipaldou's **empirical research** presented during the FERMI Final Webinar examined the coordinated digital strategies employed by Russian-linked outlets in the lead-up to the 2024 EU elections.

This case study demonstrates how the Disinformation Life-Circle operates in practice, revealing a networked system of influence that relies on repetition, cross-referencing, and linguistic adaptation to sustain credibility over time.

Outlets such as Voice of Europe (VoE), RRN, Pravda, and France et EU acted as central nodes within a larger constellation of media actors. These outlets combined authentic, semi-authentic, and fabricated content, blurring boundaries between journalism, commentary, and propaganda to shape perceptions of European politics, identity, and morality.

Key characteristics identified include:

- Cross-referencing Western media to create perceived legitimacy, often without attribution.
- False citation of mainstream outlets, replacing authentic references with Russian state-linked sources such as TASS, RT, or RIA Novosti.
- Extensive hyperlinking between obscure far-right, far-left, and conspiratorial websites, producing a misleading web of validation.
- Thematic framing around sovereignty, nationalism, and moral decay, designed to appeal emotionally across ideological divides.

Backlink analysis reveals that these narratives were amplified by both extremist and alternative media ecosystems, creating echo chambers where polarised communities inadvertently reinforced shared disinformation.

The findings illustrate how disinformation exploits network logic; it relies on visibility, repetition, and cross-ideological collaboration rather than ideological coherence.

SECTION 4: THE FERMI APPROACH – HOLISTIC ANALYSIS AND AI-ENABLED TOOLS

As presented by Katerina Beli, the FERMI project responds to this complexity with a holistic, cross-disciplinary framework that integrates artificial intelligence, behavioural profiling, and socio-economic data analysis.

Objectives

- Support European Law Enforcement Authorities (LEAs) and stakeholders in detecting, analysing, and mitigating disinformation threats.
- Predict and map the potential offline consequences of disinformation, including extremism, hate crime, and trust erosion.
- Develop and disseminate training resources to strengthen digital trust and resilience across Europe.

Key Modules and Deliverables

- Disinformation Sources and Spread Analyser – identifies origin, pathways, and scale of disinformation campaigns.
- Sentiment Analysis Module – tracks emotional tones and patterns of engagement.
- Dynamic Flows Modeller – predicts how online narratives translate into real-world risks.
- Behavioural Profiler & Socioeconomic Analyser – evaluates vulnerabilities within communities.
- Community Resilience Model – measures recovery capacity and trust-building potential.
- Swarm Learning Module – a decentralised AI system ensuring privacy-preserving collaboration among LEAs.

Together, these modules enable a multi-dimensional response that addresses both the digital and societal dimensions of disinformation.

The Dublin Riots (23 November 2023)

During the webinar was also underlined that disinformation's consequences extend far beyond the digital sphere. **The Dublin riots**, discussed during the webinar, were highlighted as a striking example of how online false narratives can escalate into offline unrest and violent mobilisation, illustrating the tangible risks to democratic stability and public safety. FERMI's framework responds to this 'law-enforcement nexus' by helping authorities identify early signals of escalation, assess community vulnerability, and strengthen societal resilience through data-driven prevention and coordinated response.

THE INCIDENT

On November 23, 2023, a man attacked three children and an adult outside a school in Dublin, Ireland. The incident was tragic but became catastrophic when false information about the attacker's background began spreading online.

THE FALSE NARRATIVE

Within hours, false claims circulated on social media and online forums stating that the attacker was a migrant or refugee. These claims were not based on any official information but were presented as fact, often with provocative framing such as 'another example of Ireland's failed immigration policy' or 'what did you expect when you let these people in?'

THE SPREAD

The false narrative spread rapidly through:

- Right-wing social media accounts and Telegram channels
- Online forums and discussion boards
- Coordinated sharing campaigns that created appearance of consensus
- Exploitation of existing tensions around immigration policy in Ireland

The Dublin Riots (23 November 2023)

THE REAL-WORLD IMPACT

That evening, massive riots erupted in Dublin, with intensity in the city centre. The violence included:

- Destruction of public property
- Burning of vehicles including buses
- Attacks on businesses
- Clashes with police
- Injuries to civilians and law enforcement
- The Dublin City Council later estimated the cost of damage at approximately €20 million.

WHAT THIS CASE TEACHES US

Speed of Spread vs. Speed of Truth: False information spread within hours, while official corrections took much longer and reached fewer people.

Exploitation of Existing Tensions: The disinformation didn't create anti-immigration sentiment; it weaponised pre-existing anxieties and cultural tensions.

Real-World Consequences: This demonstrates that disinformation is not merely an online phenomenon. It can directly lead to violence, economic damage, and societal harm.

Difficulty of Correction: Even after the false narrative was debunked, many who participated in or supported the riots continued to believe or justify their actions based on the original false claim.

SECTION 5: CREDIBILITY, LEGITIMACY AND THE TRUST GAP

The Credibility Gap (per Aikaterini Beli)

Definition.

The credibility gap is the widening divergence between what citizens perceive as trustworthy and what institutions (government, media, science, etc.) recognise as legitimate.

Why this matters?

Disinformation exploits this gap by:

- privileging emotion over evidence and identity cues over expertise;
- favouring simplification over complexity, offering clear villains and tidy answers;
- maximising speed and spectacle over verification;
- deploying AI-generated and manipulated content to mimic authenticity at scale.

Implications.

As the gap grows, trust fragments, institutional messages face greater scepticism, and misleading narratives find lower barriers to uptake and recirculation.

This dynamic is sustained by narrative interaction: each false claim provokes reactions and counterclaims, which in turn amplify visibility. In this sense, disinformation is not a one-way message but an ongoing dialogue, a feedback loop that blurs truth boundaries and destabilises public confidence.

FERMI's methodology provides tools to map these credibility and legitimacy chains, tracing how emotional leverage, identity politics, and digital amplification mechanisms shape the perception of trustworthiness.

***"Credibility is not decreed; it's earned through consistency and openness."
(FERMI Webinar, 2025)***

SECTION 6: YOUR PERSONAL RESPONSE TOOLKIT

As Katerina Beli underlined during the FERMI Final Webinar, combating disinformation requires not only institutional safeguards but also personal awareness. Digital trust, she explained, begins with the everyday decisions of individual users, how we consume, interpret, and share information. Inspired by her remarks on the importance of critical thinking, responsible technology use, and the active role of citizens in strengthening resilience, this section translates those ideas into practical guidance. While structural measures and AI-driven detection tools form one part of the solution, individual actions collectively shape the information ecosystem. By **pausing** before reacting, questioning emotional content, and verifying credibility, each of us can contribute to a more trustworthy digital space.



The STOP Method: Before You Share

When you encounter information online - especially if it's surprising, outrageous, or emotionally compelling - use the **STOP** method before reacting or sharing:



S - STOP

Pause before reacting emotionally or sharing. Our immediate instinct when seeing shocking information is often to share it immediately. Resist this impulse.

T - THINK

Ask yourself critical questions:

- Who created this content?
- Why was it created?
- What is the source?

Am I being manipulated emotionally?

O - OBSERVE

Look for warning signs of misleading content:

- Emotional, sensationalised language
- Lack of credible sources or citations
- Suspicious URLs or unfamiliar websites
- Poor grammar or spelling (though be aware this can also be deliberate to evade detection)
- Requests for immediate action or sharing

P - PROBE

Verify before accepting or sharing:

- Use fact-checking sites
- Perform reverse image searches
- Check if reputable news sources are reporting the same information
- Look for the original source of the claim

Practice exercise: Put your skills to the test and turn pausing into a habit.

Take the short interactive activity from the BC Human Rights Commission to practise Stop-Think-Observe-Probe on real examples, get instant feedback, and reinforce the checks you'll use in the wild:

<https://bchumanrights.ca/resources/awareness-campaigns/stop/>.

SECTION 7: POLICY RECOMMENDATIONS AND SOCIETAL IMPLICATIONS

Both presenters underscored the importance of balance between effective countermeasures and the protection of fundamental rights. Policy responses must be proportionate, transparent, and grounded in democratic principles, avoiding overreach or arbitrary censorship.

Key policy insights include:

- Avoid overinflating the perceived reach of disinformation campaigns, which may inadvertently amplify them.
- Use consistent terminology, distinguishing between disinformation, misinformation, and legitimate political communication.
- Complement AI detection tools with qualitative, context-driven analysis to prevent misinterpretation.
- Promote media literacy, digital literacy and critical thinking as foundational skills for all citizens.
- Encourage cross-sectoral collaboration among governments, tech companies, researchers, and civil society to reinforce digital trust.

SECTION 8: CONCLUSIONS

The FERMI project reframes disinformation as an evolving ecosystem, not a static threat, and provides a practical pathway for resilience building.

By combining technological innovation, ethical governance, and citizen empowerment, FERMI strengthens the European capacity to navigate the complexity of digital information flows.

FERMI brings to the table:

- A dynamic understanding of disinformation as an iterative, socially embedded process.
- Analytical tools that connect narrative evolution to measurable social risks.
- A trust-centred approach that bridges gaps between institutions and citizens.
- A collaborative model aligning AI, ethics, and democratic accountability.

KEY TAKEAWAYS

Disinformation thrives by exploiting societal weak points, emotional triggers, informational ambiguity, and gaps in institutional credibility, then iterates through a five-stage lifecycle of authentication, temporal variation, linguacultural travel, actor collusion, and counterclaims/rebuttals that keep harmful narratives circulating. Closing the credibility gap, the distance between what people perceive as trustworthy and what institutions deem legitimate, is therefore central to resilience.

The FERMI approach addresses this with AI-enabled analysis of sources, spread, sentiment and risk, alongside community resilience modelling, so actors can anticipate narrative shifts rather than merely react to them. Effective countermeasures must remain proportionate and rights-respecting, upholding freedom of expression while ensuring transparency and accountability, and, ultimately, digital trust is a shared responsibility that demands coordinated effort across policy, technology, media, civil society and citizens themselves.

By combining technological innovation, ethical governance, and citizen empowerment, FERMI strengthens the European capacity to navigate the complexity of digital information flows.

FURTHER READING

For deeper context, readers are encouraged to consult the previous FERMI training materials:

- [Navigating Disinformation: A Comprehensive Guide](#)
- [Digital Trust: A Practical Path to Combating Disinformation and Fostering Resilience](#)

And:

- [Disinformation as process: modeling the lifecycle of deceit](#)
- [Information disorder: Toward an interdisciplinary framework for research and policy making.\(2017\)](#)
- [Misinformation Has Created a New World Disorder](#)
- [Misinformation, Disinformation, and Bias: Identifying misinformation, its various types, and both personal and news media biases](#)
- [Misinformation: Can you STOP it?](#)

and explore:

- <https://fighting-fake-news.eu/materials/training-materials>