



AI4Debunk

UNIVERSALIA LUNCH AND LEARN

Pascaline Gaborit- Director Pilot4dev/Researcher
AI4DEBUNK and Evaluator of EU projects

27/03/2026



Funded by
the European Union



Proposed Presentation



1. Short Presentation of Pilot4dev and the AI4DEBUNK project



2. Our findings on Disinformation and the impacts of disinformation



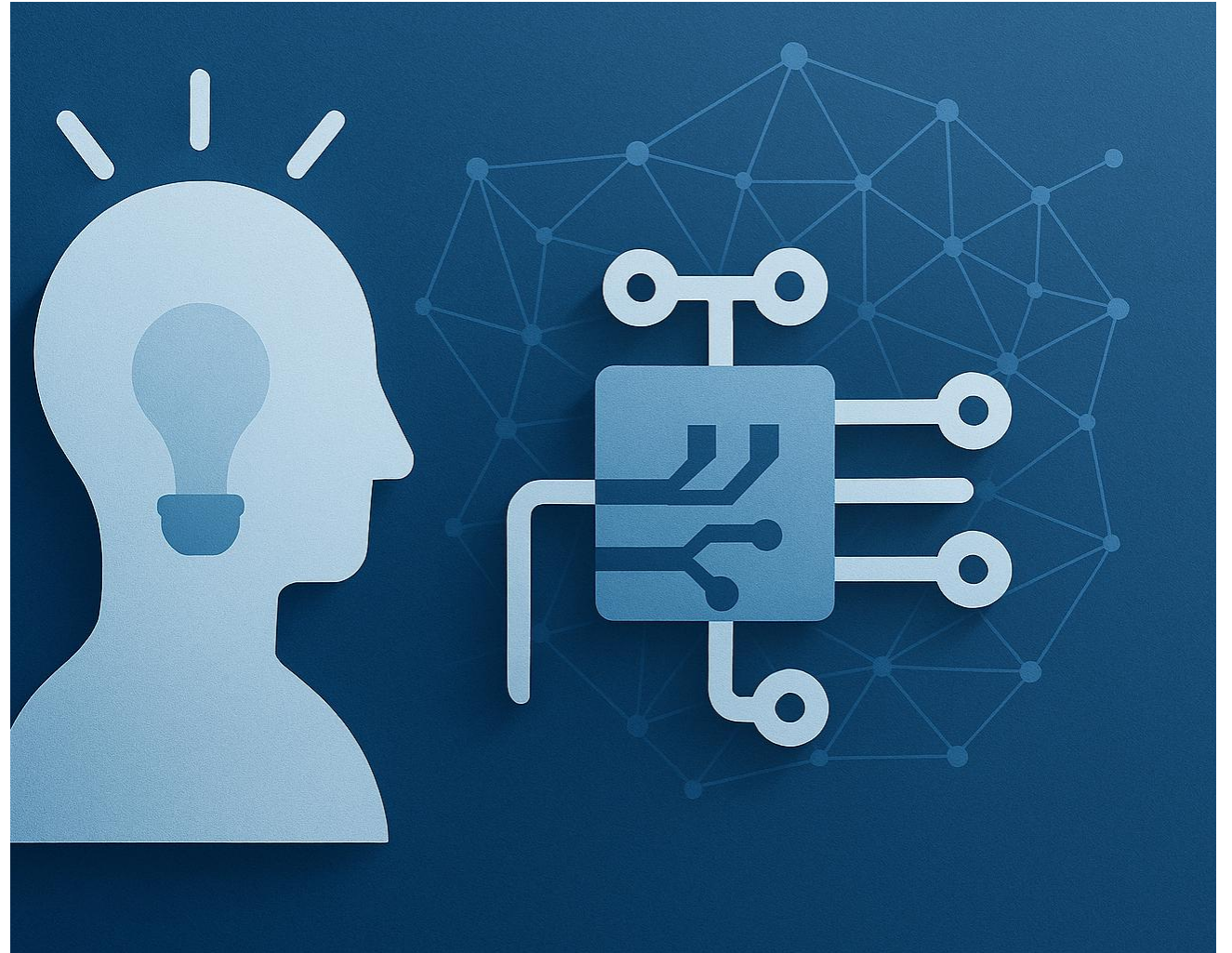
3. The impacts of digital transformation, AI and disinformation on our work



I. Short Presentation of Pilot4dev and the AI4DEBUNK project

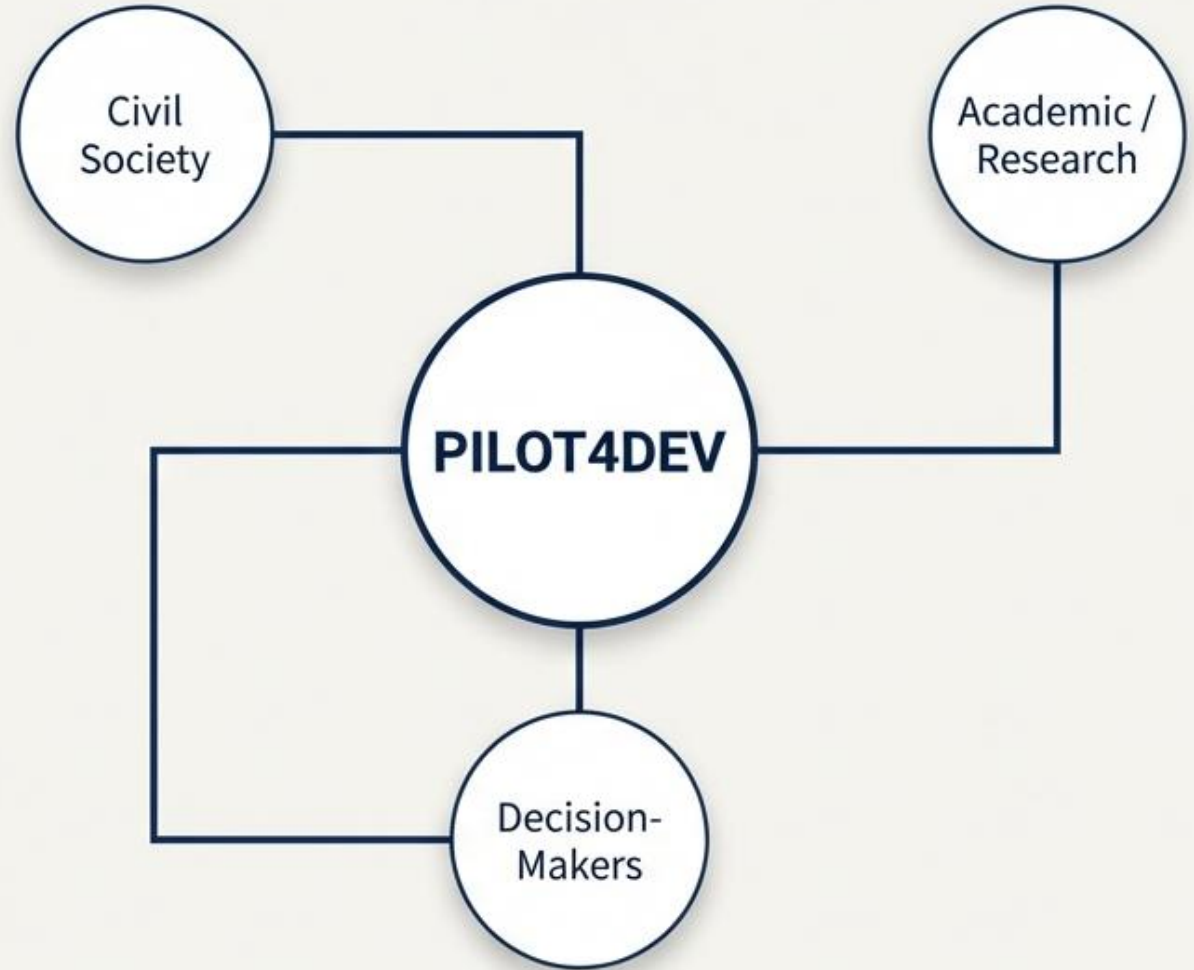
www.pilot4dev.com

www.ai4debunk.eu



Bridging the Gap Between Innovation and Policy

PILOT4DEV is an independent Think-Tank dedicated to transforming pioneering pilot initiatives into mainstream sustainable policy.



Governance

Cities Climate

Gender Equality

AI & Digital Tools

Disinformation

Stakeholder Engagement

A Full-Spectrum Approach to Sustainable Development



Research & Insight

- Publications & Analysis
- Development of Research Methodologies
- Cross-Country Exchanges



Digital Innovation

- Beta Testing Groups for Innovative Solutions
- Digital Tools & Online Surveys



Policy & Action

- Multi-Stakeholder Engagement
- Trainings & Capacity Building
- Actionable Policy Recommendations

What is AI4Debunk

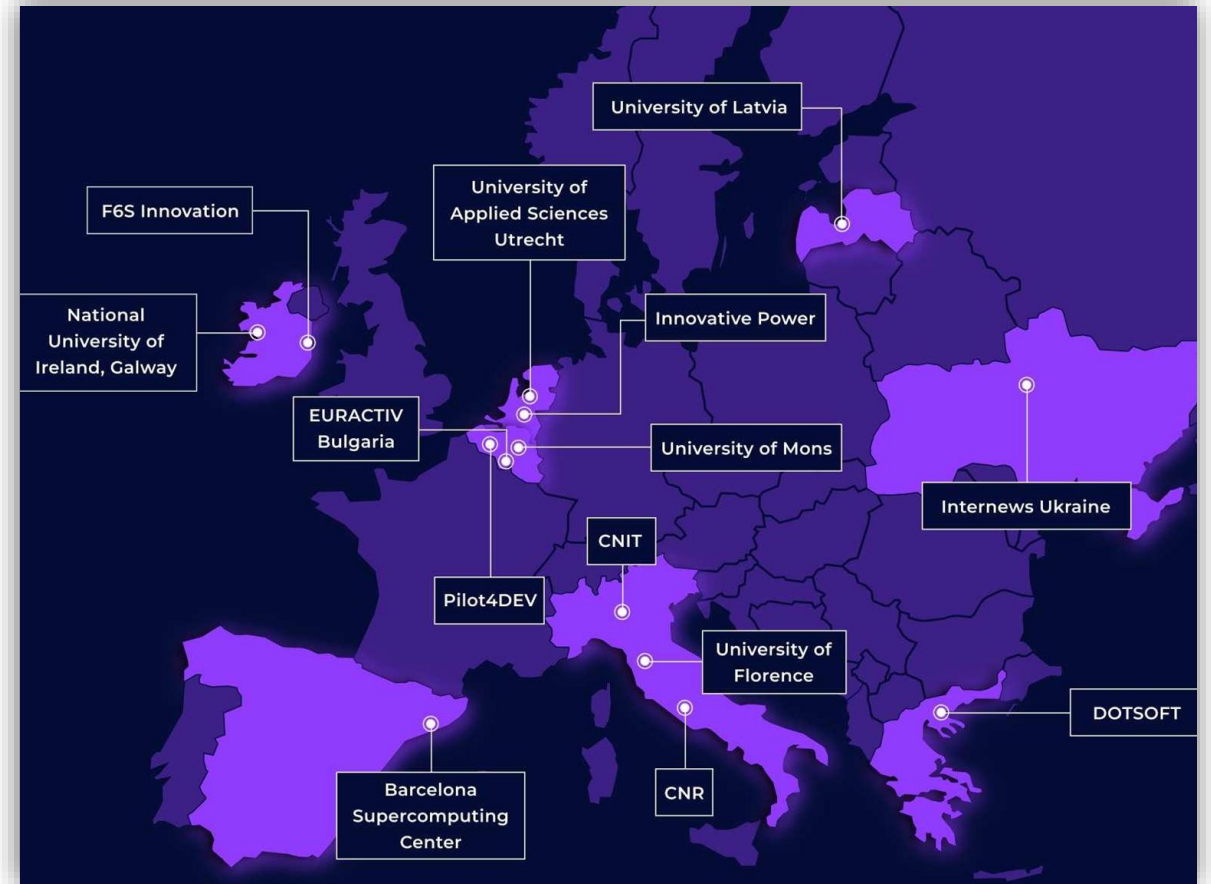
The project Overview

AI4DEBUNK is an EU funded project lasting 4 years, and gathering 14 partners in 8 countries.

It aims is to develop AI and participative tools for Supporting Trustworthy Online Activity for Citizens and to Debunk Disinformation.

The project seeks to merge sociological insights on disinformation with practical, AI-driven solutions designed to counter it effectively.

The project also has an educational dimension to create games and comic books for young people to build media literacy



Map of all the 14 partners of the AI4Debunk project

Grounding our approach in the AI4DEBUNK initiative



Recent Methodological Foundations

- Trust, New Technologies and Geopolitics in an Uncertain World (Peter Lang, 2025)
- A Sociological Approach to Disinformation and AI (Journal of Political Science, 2024)
- Anticipatory Governance, Planning and Dialogue (Peter Lang, 2022)

Questions *Where do you most often encounter questionable information?*



- ✓ **Social Media ?**
- ✓ **Online information?**
- ✓ **Emails and communication ?**
- ✓ **Reports?**
- ✓ **Discussions ?**



**Our findings on
Disinformation
and the impacts
of disinformation?**



Methodology AI4DEBUNK...

-  **Review existing research- 1000 case studies identified**
-  **Interview Experts and stakeholders**
-  **Beta-tester group – Observation of reactions to information manipulation and to our AI-tools**
-  **Online Survey for citizens perspective- Around 300 answers and an in person survey 175 answers**
-  **Multi-stakeholder focus groups – Discussion to share perspectives and experiences**

2. Online Survey the perceptions of the citizens, policy makers and the social media users

Online Survey: 334 respondents provided insights into citizens Disinformation and AI perceptions.

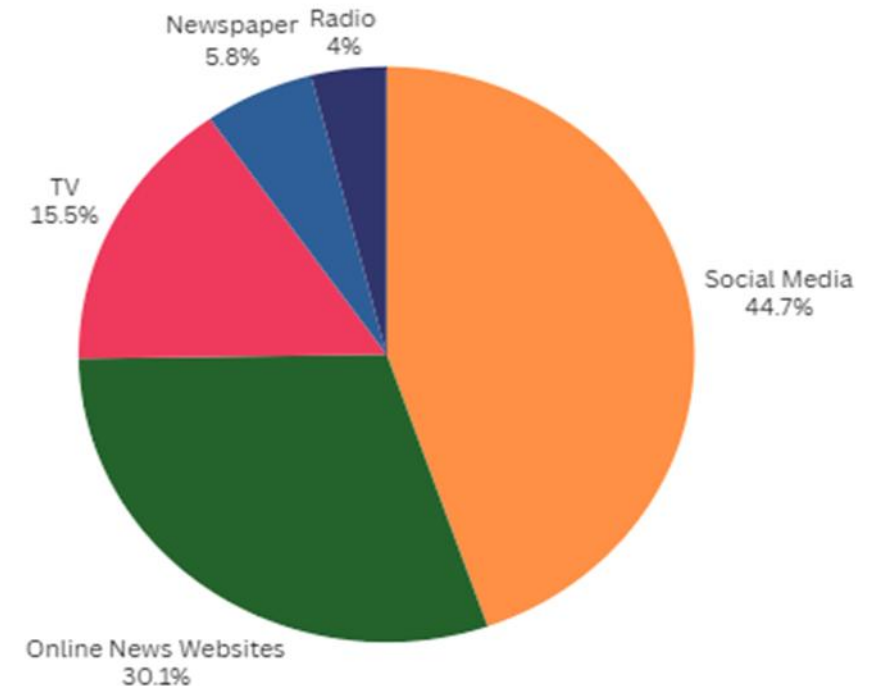
Platforms Monitoring: Identification of fake news and narratives on X, Meta, YouTube, TikTok, and Telegram.



Insights and results:

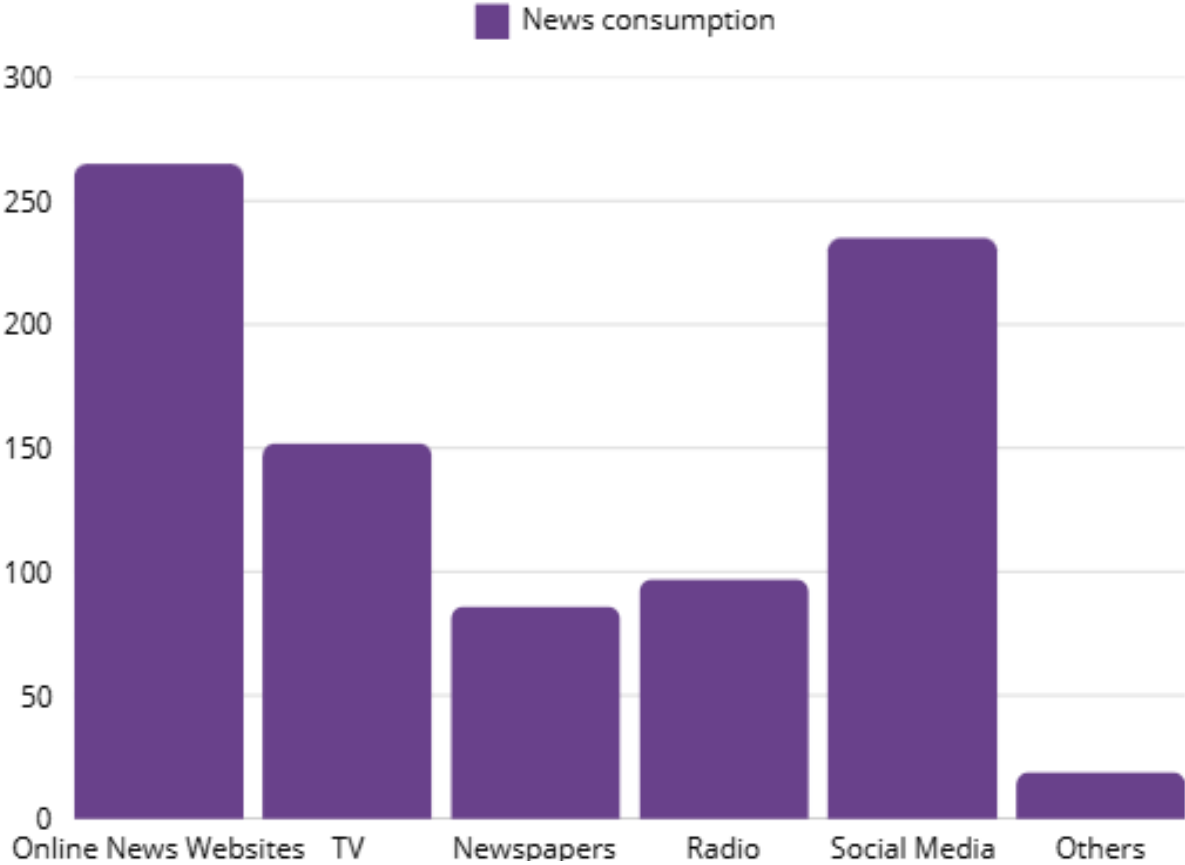
- Citizens perceives social media as a vector for misinformation.
- Platform moderation is viewed as crucial for countering disinformation.

Where did you encounter news that you believe to be fake or misleading?

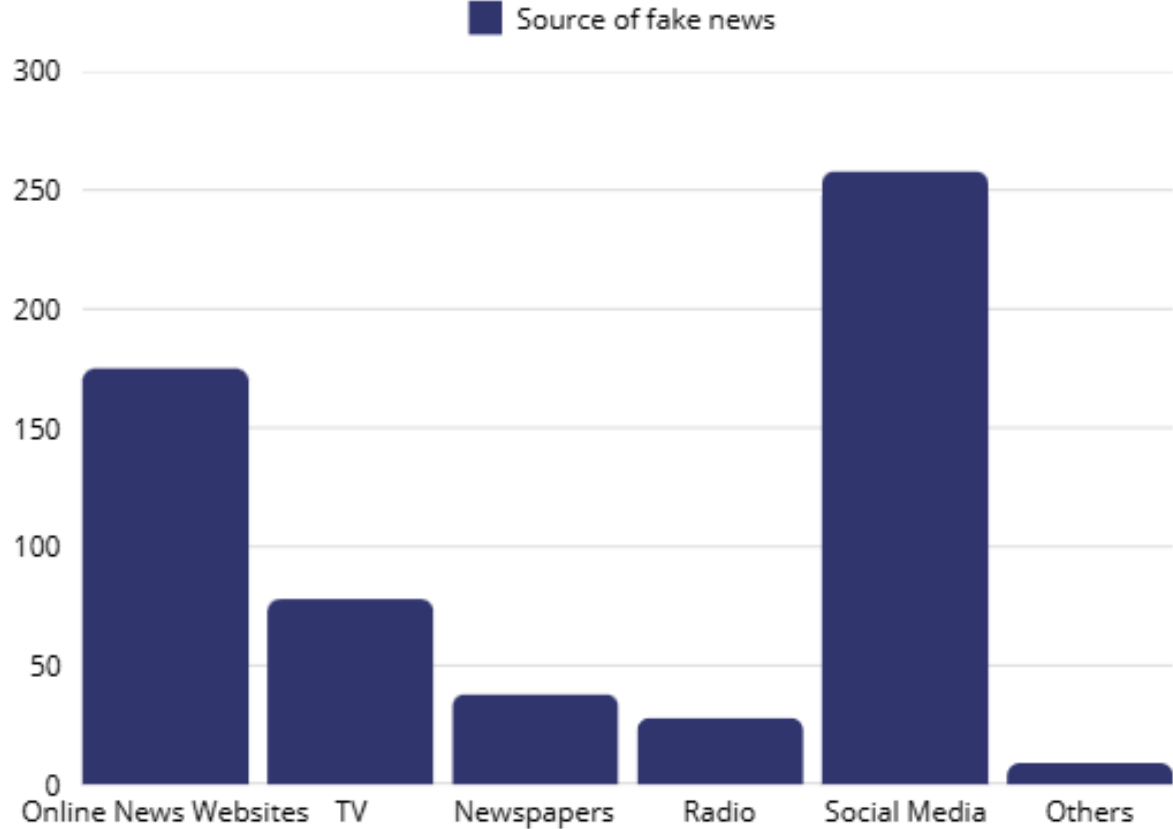


Online Survey: Results

Where do you consume your news? (Select all that apply)

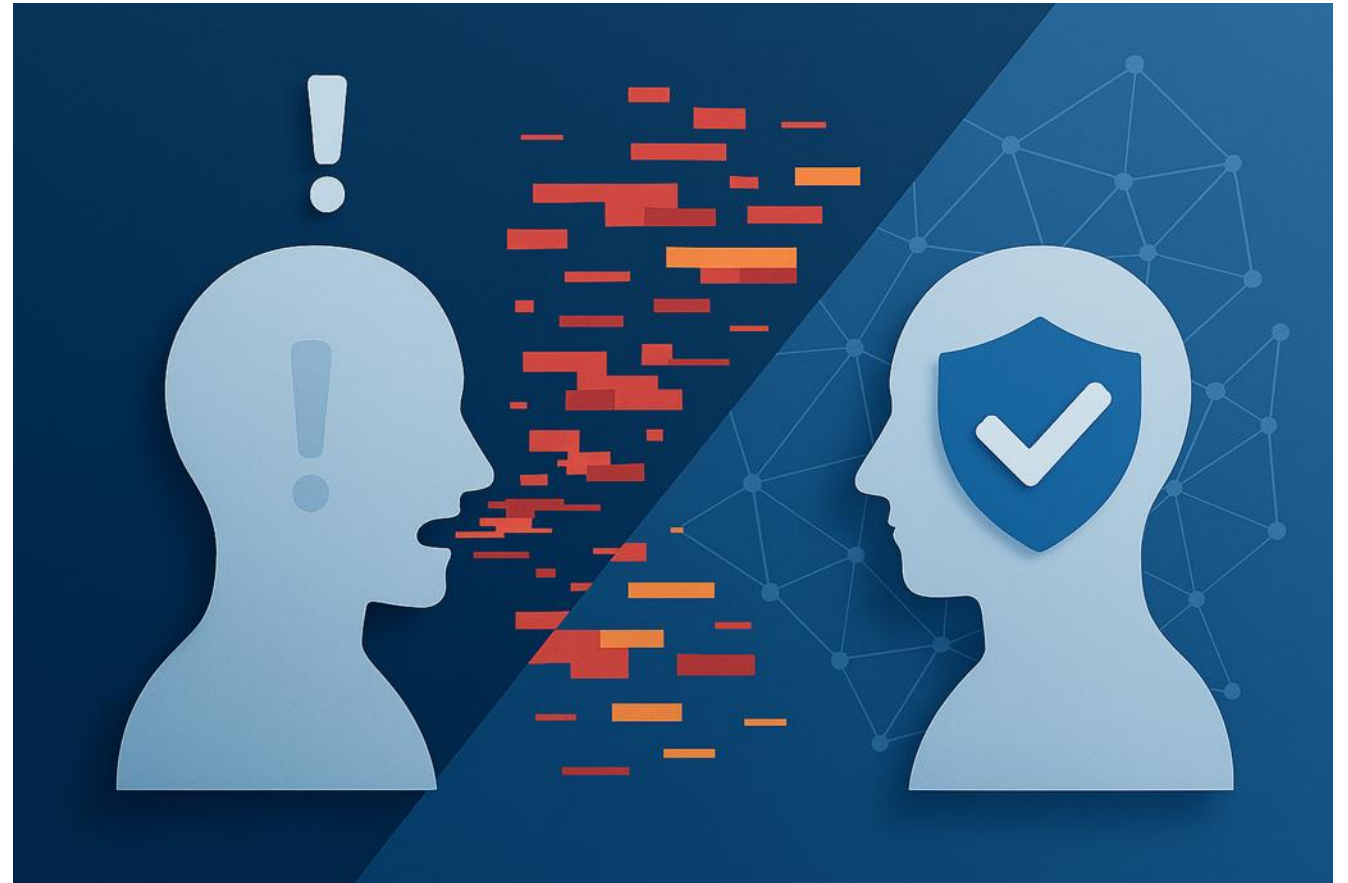


Where did you encounter news that you believe to be fake or misleading?



Questions *What impact do you think disinformation has on society?*

- Low ?
- Moderate?
- High?
- Very high?



Question? which measures do you think is the most effective in combating fake news?



Greater Public Education ?



Increased cooperation?



Technological solutions?



Policy and regulation ?



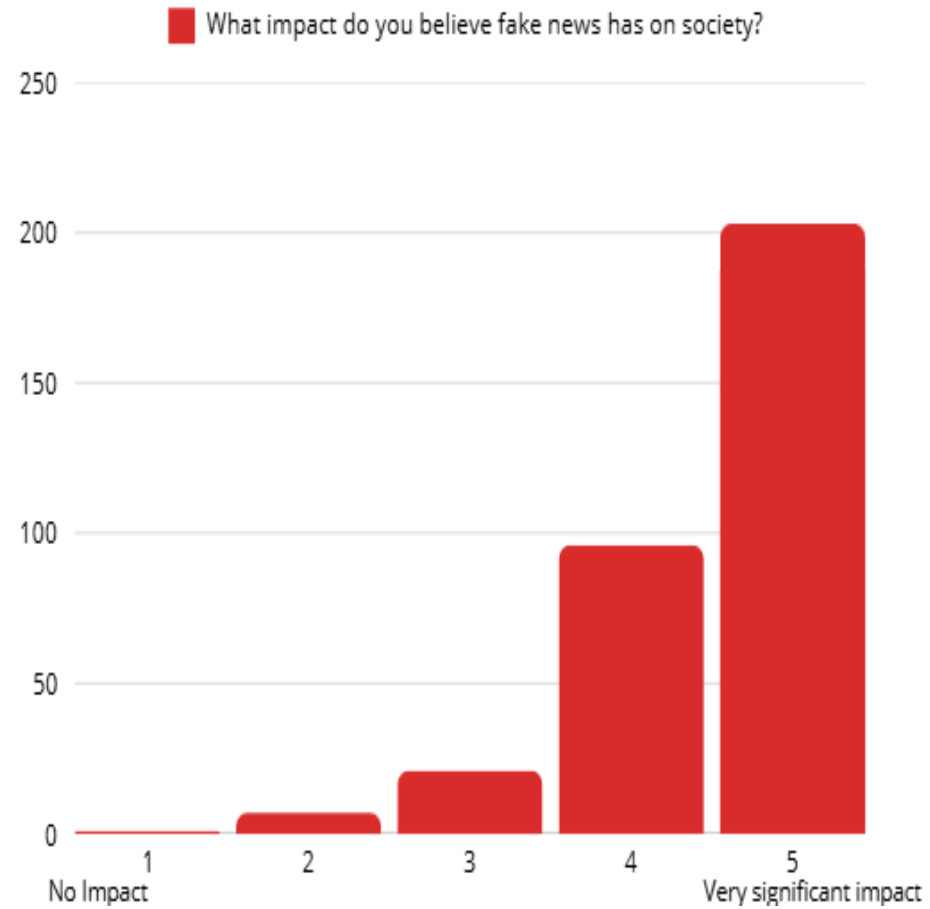
Other ?



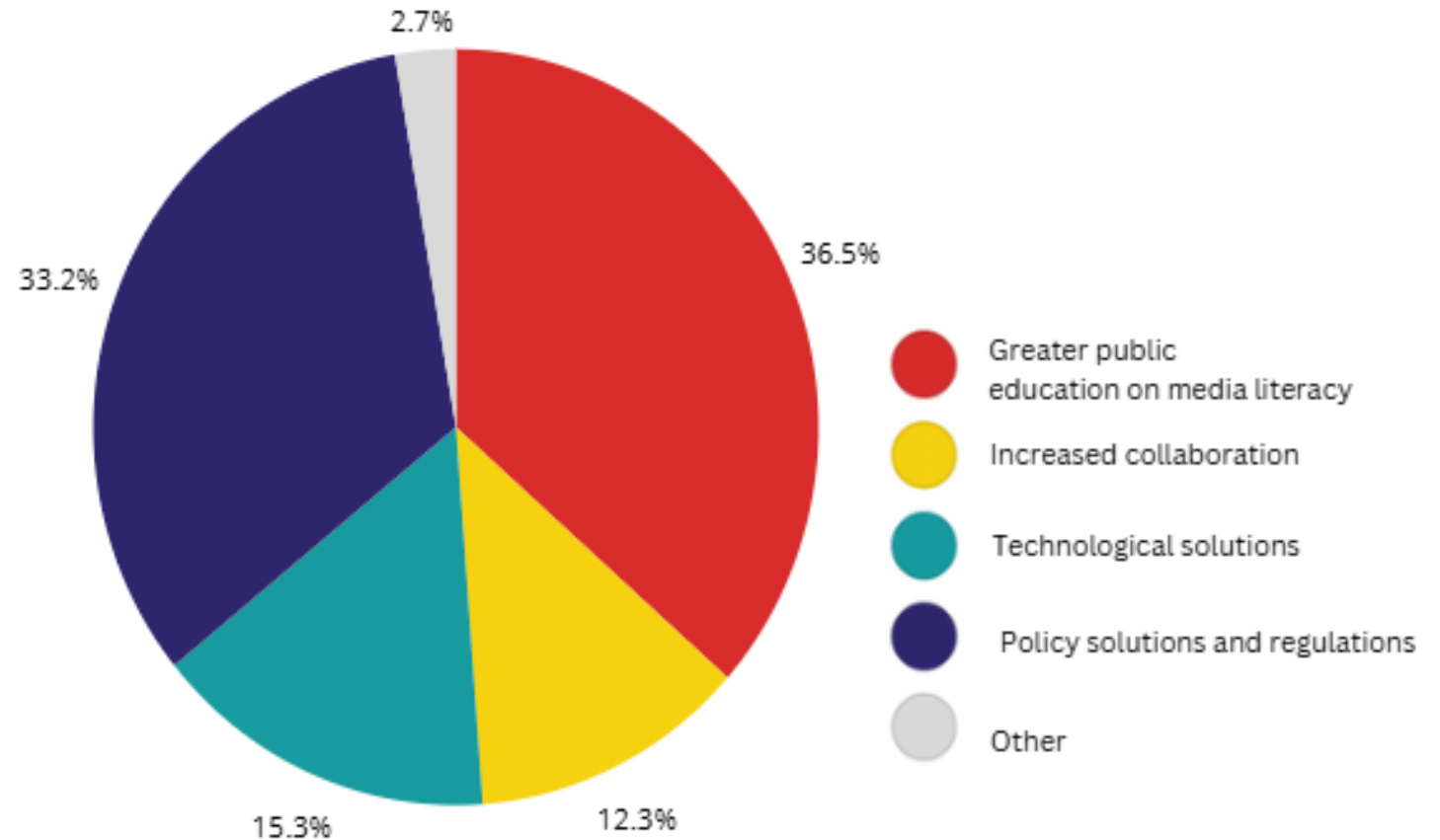
2. What impact do you believe fake news has on society?

Online Survey: First overview

What impact do you believe fake news has on society?



What measures do you think would be most effective in combating fake news? (Select up to 2 options)



Topics for research

1

Opportunities and limits to critical thinking as a response to information manipulation

2

Population groups most vulnerable to disinformation

3

The effect and successes of existing policies to regulate information manipulation

4

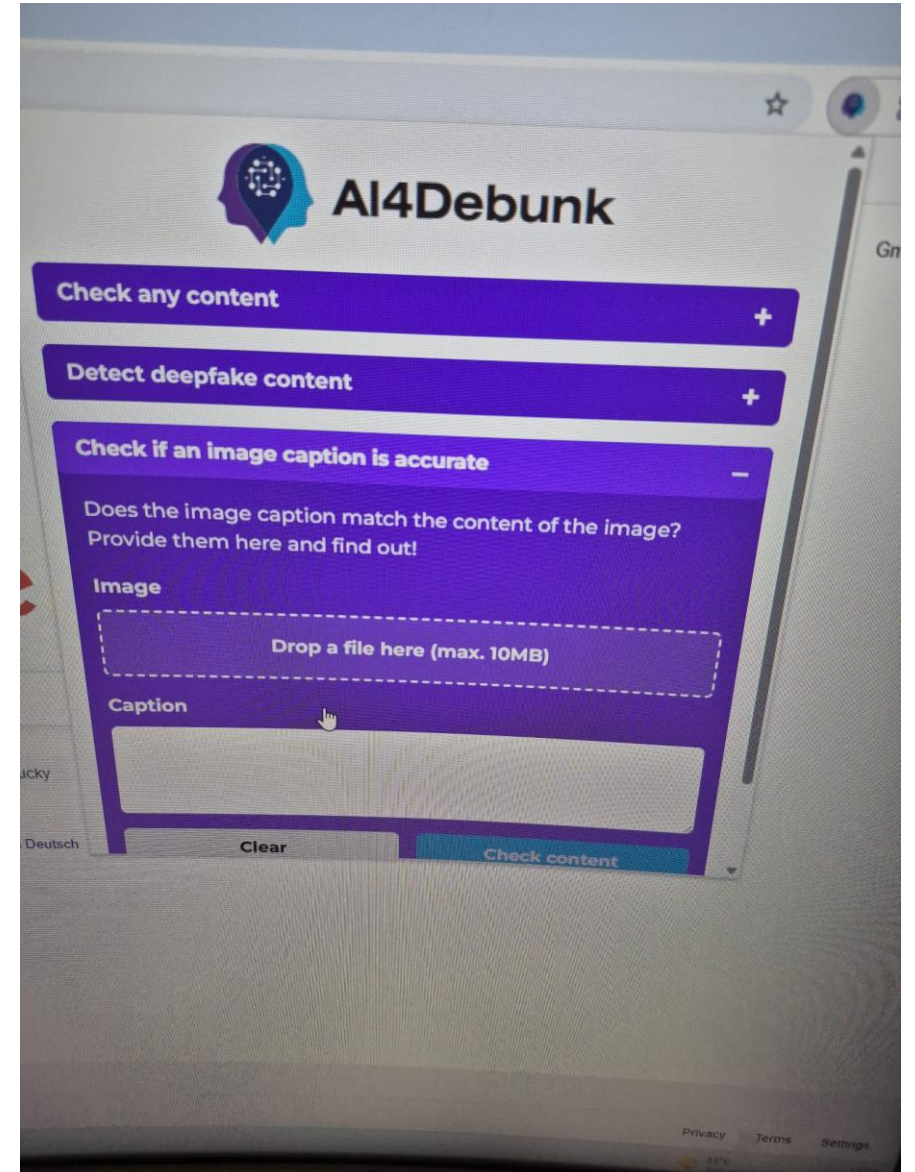
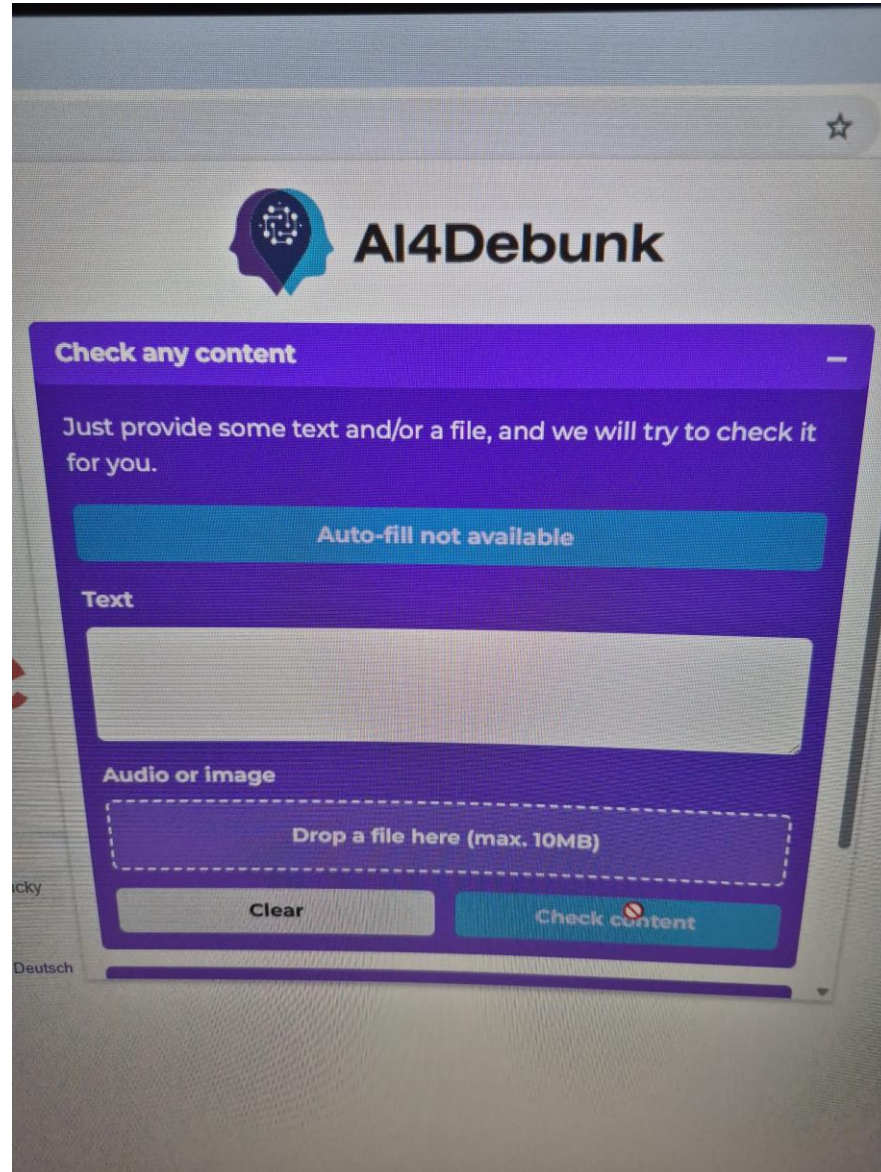
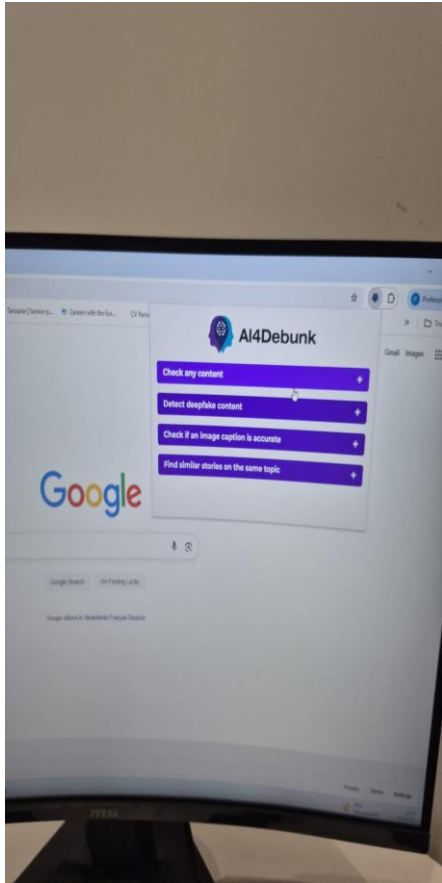
Identification of threat actors involved in information manipulation campaigns

5

Identifying polarizing narratives and inflame of political, ethnic, religious divides in our societies



The Tool



Videos: www.AI4DEBUNK.EU





Interim
SOCIOLOGICAL
ASSESSMENT
INCLUDING GENDER
ANALYSIS AND MULTI
STAKEHOLDERS
RECOMMENDATIONS

**WP13
Outputs
(July 2026)**

D13.5: Intermediate recommendations from the task force.

D13.1: Intermediate report on the resilience mechanisms triggered.

D13.4: Intermediate report on multi-stakeholders' perspectives.

D13.2: Update of the Gender Equality Report and Training given in Florence on Gender Stereotypes.

D13.3: Report on the comparison of prototypes and tools (UL).



Zaneta Ozoliņa
University of Latvia

Disinformation: theoretical framework and policy responses in the European Union

This is a summary of Public Deliverable D4.1; for more detailed information, refer to the full [document](#)

Over the past decade, disinformation has emerged as one of the most salient security and governance challenges confronting democratic societies. The accelerating evolution of information and communication technologies (ICTs), the proliferation of social media platforms, and the increasing sophistication of artificial intelligence (AI) tools have dramatically expanded the scale, speed, and complexity of information flows. In this environment disinformation has become a pervasive instrument of political influence, societal destabilization, hybrid warfare and military operations.

The AI4Debunk project has developed a theoretical framework for understanding disinformation, particularly in the context of European Union (EU) policymaking. This article synthesizes key elements of that framework, focusing on the definitional contours of disinformation, its structural components, mechanisms of diffusion, societal impacts, and the EU's institutional responses.



Inna Steinbuka
University of Latvia

Information Manipulation and EU Response Effectiveness

This is a summary of Public Deliverable D4.2. Please refer to the full [document](#) to find the detailed information

Communication culture in the EU emerged in a peaceful environment and relatively friendly media landscape. The communication was grouped broadly around a widely shared mission to 'build Europe' and persuade people to support EU objectives. How does this concept work in times of information manipulation and massive spread of disinformation? How effectively does communication correspond with the rapidly changing EU media landscape?

D4.1. Information and Disinformation Frameworks

D4.2: Information Manipulation in the EU Media Ecosystem

D.5.1. Disinformation target groups in the EU member states, and sources and hosts of propaganda

D.5.2. Narratives and foreign interference throughout Europe illustrated by case studies + 2 Policy Briefs



Pascaline Gaborit
Pilot4DEV

Fake News, Media Literacy and Critical Thinking

This is a summary of Public Deliverable D4.2. Please refer to the full [document](#) to find the detailed information

The spread of disinformation and misinformation has a long history, but the term "fake news" became particularly prominent in public discourse during the 2016 American presidential election. It has been recognized since then as one of the major threats to security in European countries. The danger posed by fake news to society has become evident, especially during major events like the Covid-19 pandemic and the Russian invasion of Ukraine. As a result, certain governments have enacted legislation to combat disinformation.

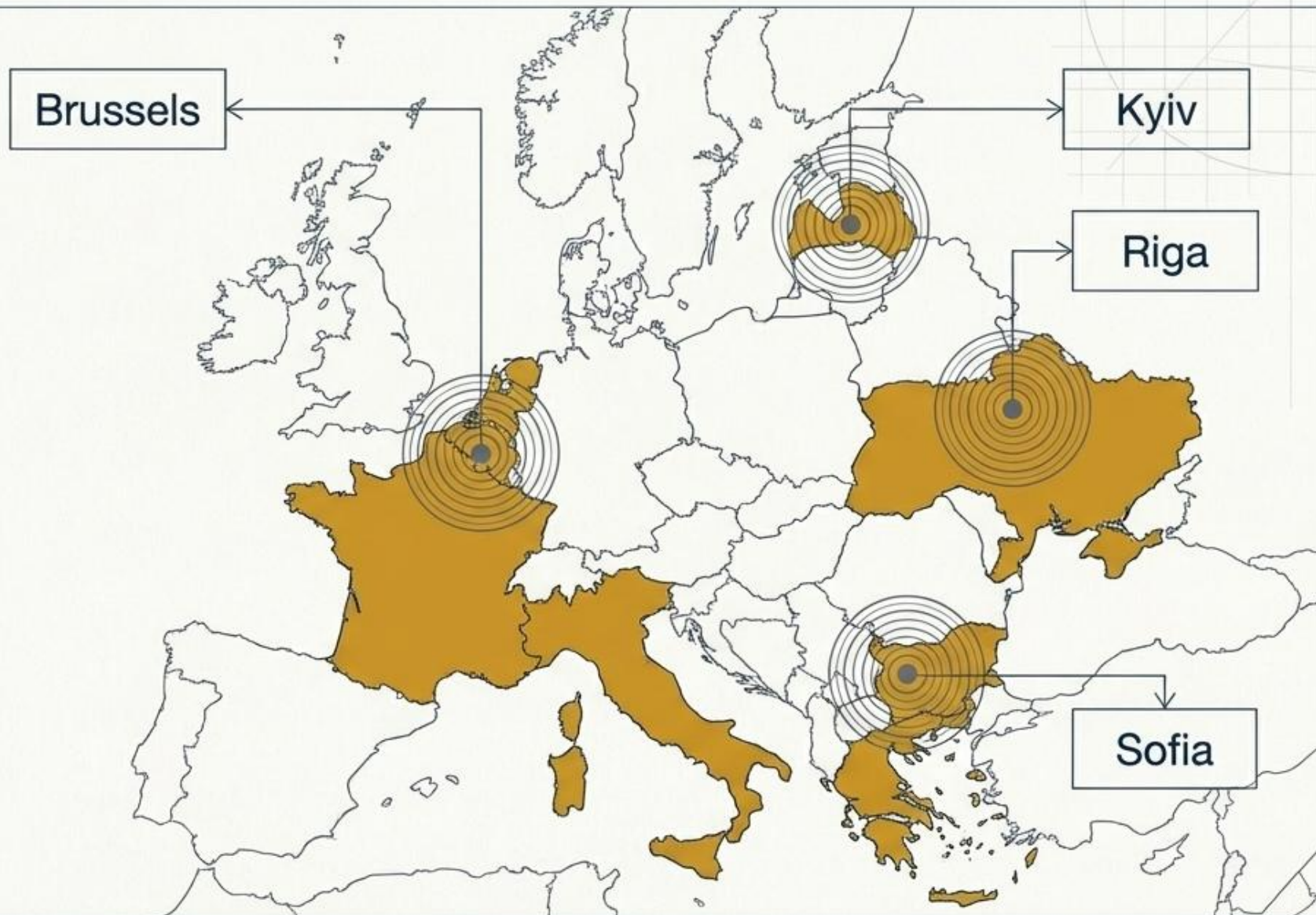
Despite these efforts, the problem of fake news remains largely unresolved by current legislation which focused mainly on platform regulation and collaboration with service providers to curb the spread of fake news.



3. R5 (Pan-European Strategy)

The consortium has fostered broad adoption by conducting surveys and questionnaires in eight countries: Belgium, France, Netherlands, Latvia, Ukraine, Greece, Bulgaria, and Italy.

Furthermore, multi-stakeholder focus groups were organized in key European cities, including Brussels, Kyiv, Riga, and Sofia (Task 13.4).



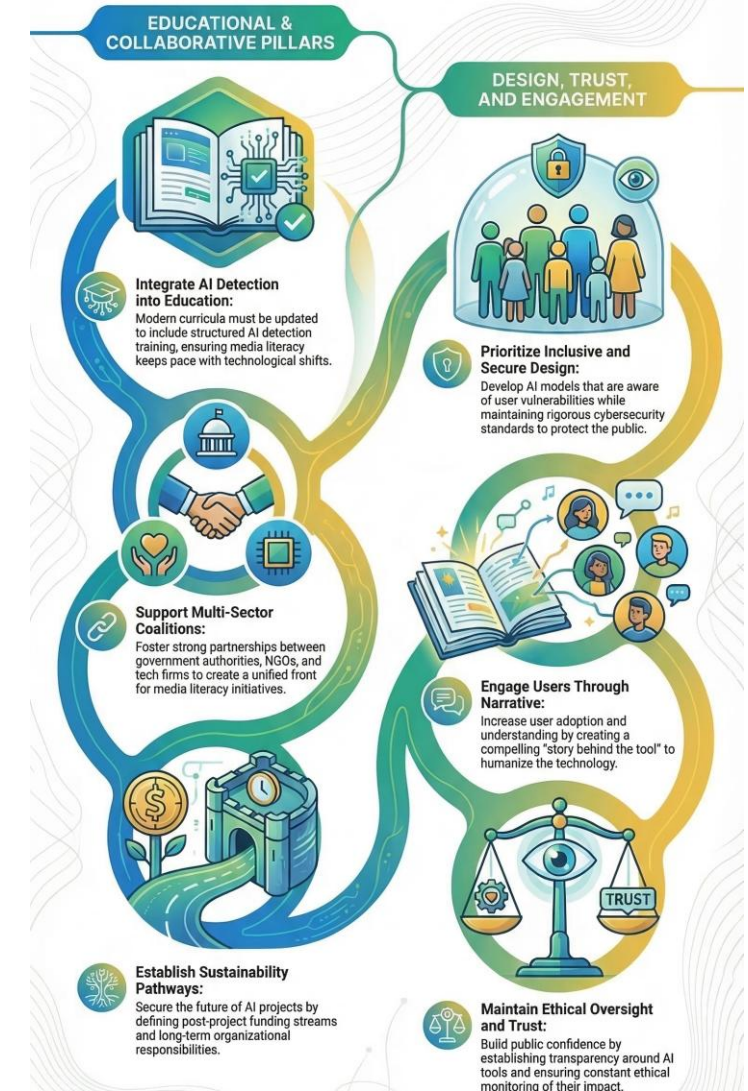
Multi-stakeholders' perspectives on resilience to disinformation (focus groups)

Media Literacy: Integrate AI detection with structured education curricula

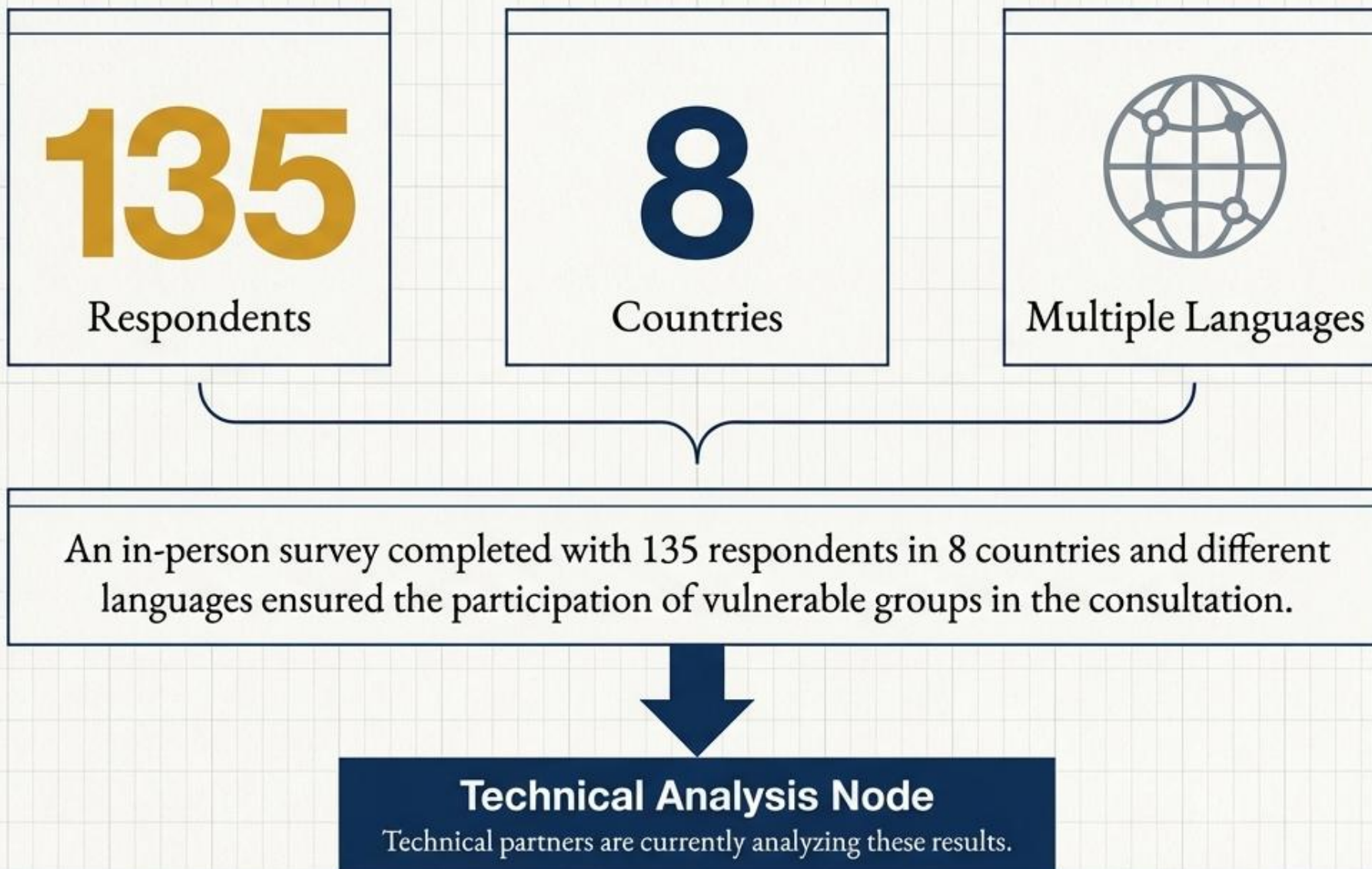
Coalitions: Support partnerships between authorities, NGOs, and tech firms

Ethics: Maintain ethical oversight

A Strategic Roadmap for AI and Media Literacy



2. R4 (Vulnerable Groups & Gender)



3.

**The impacts of
disinformation,
digitalization
and AI on
our work**



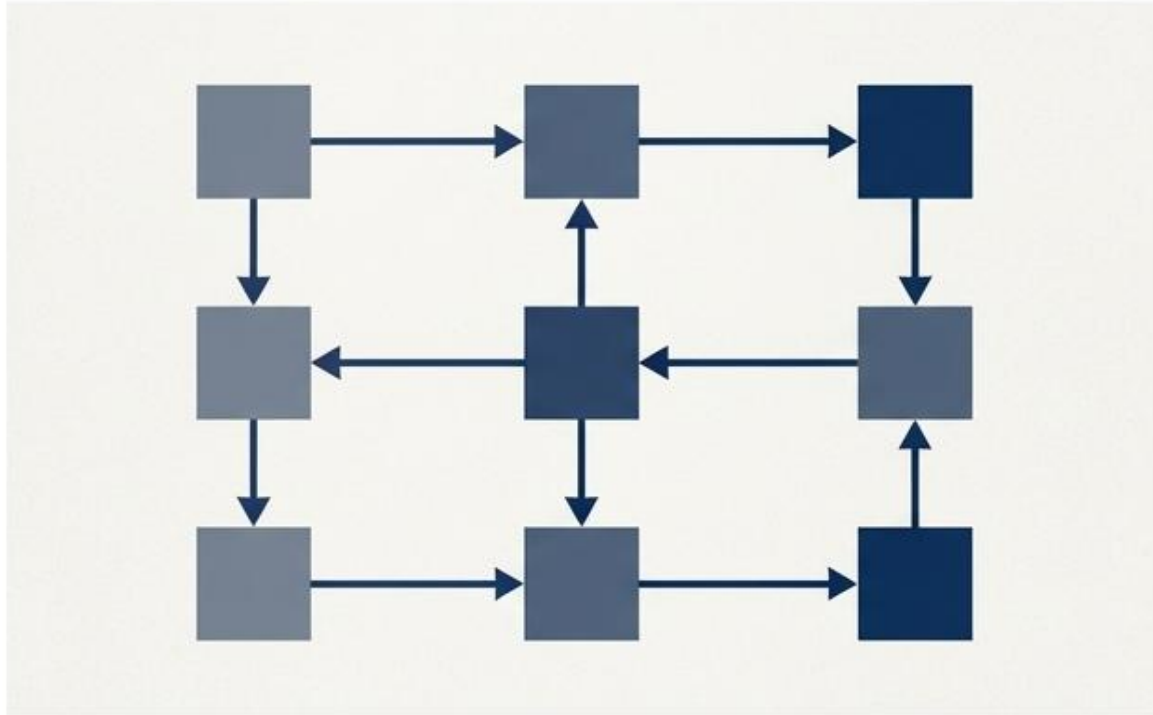
Question?

**What impact
does
disinformation
have on you?**

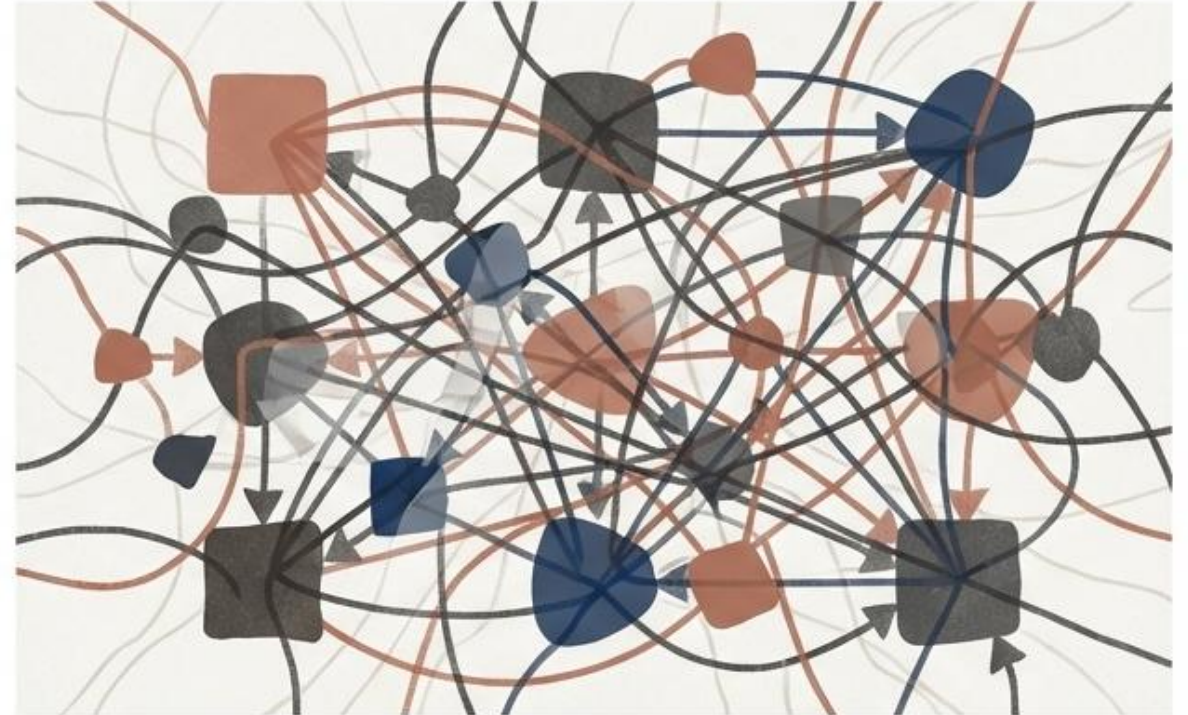


Online media and AI are actively reshaping the baseline of evidence-based work

Traditional Information Ecology



The Current Reality



The Catalyst:

The rapid integration of AI into both the creation and consumption of digital media.

The Disruption:

Disinformation operates at unprecedented speed and scale, muddying the data sources researchers rely upon.

The Imperative:

A drastically elevated need for critical thinking and rigorous methodological validation within consultancy missions.

Globalised conspiracy narratives actively erode the foundations of professional dialogue

Impact on Trust

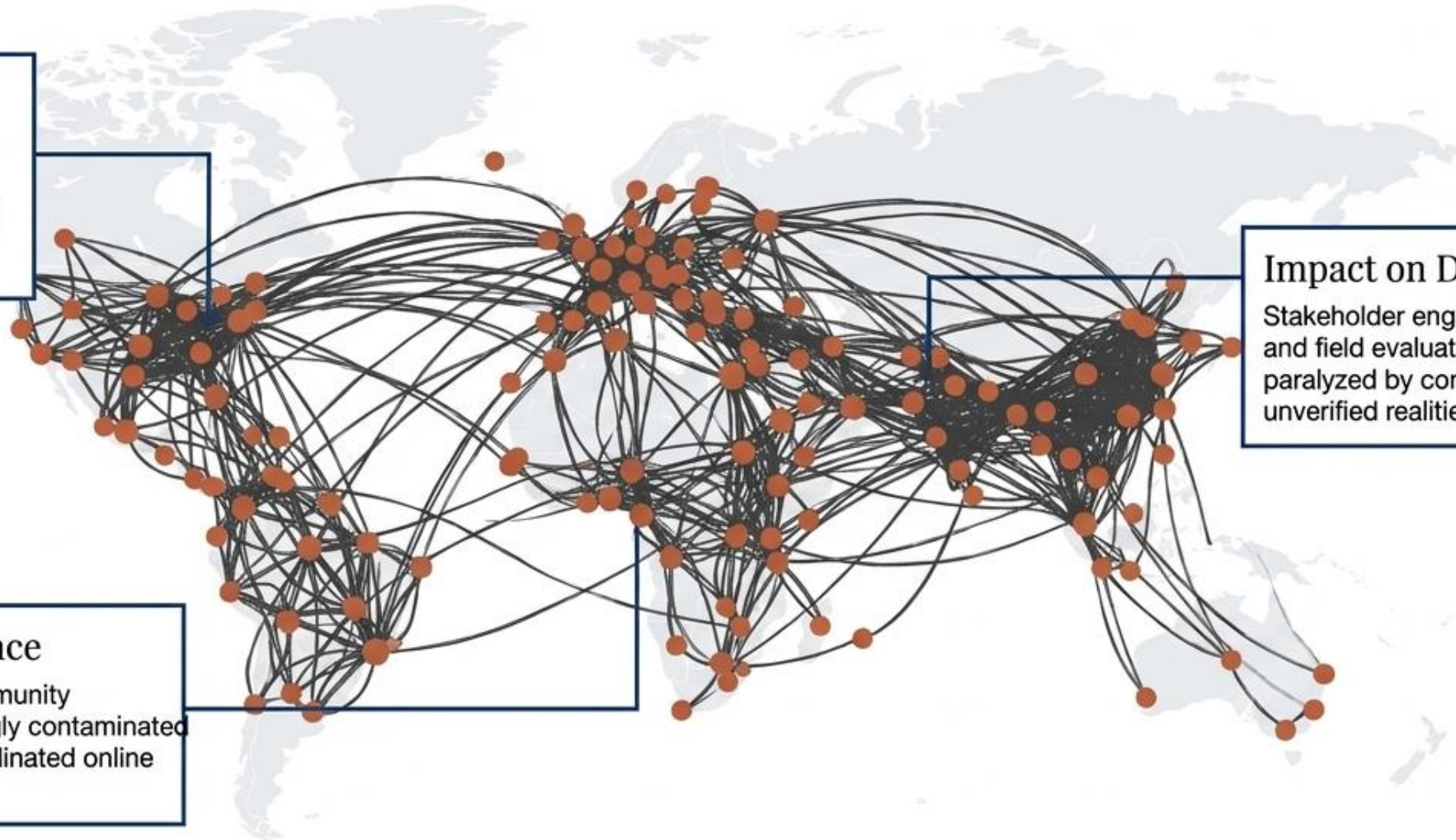
Deep-seated narratives breed skepticism, directly undermining the credibility of institutions and consultants.

Impact on Dialogue

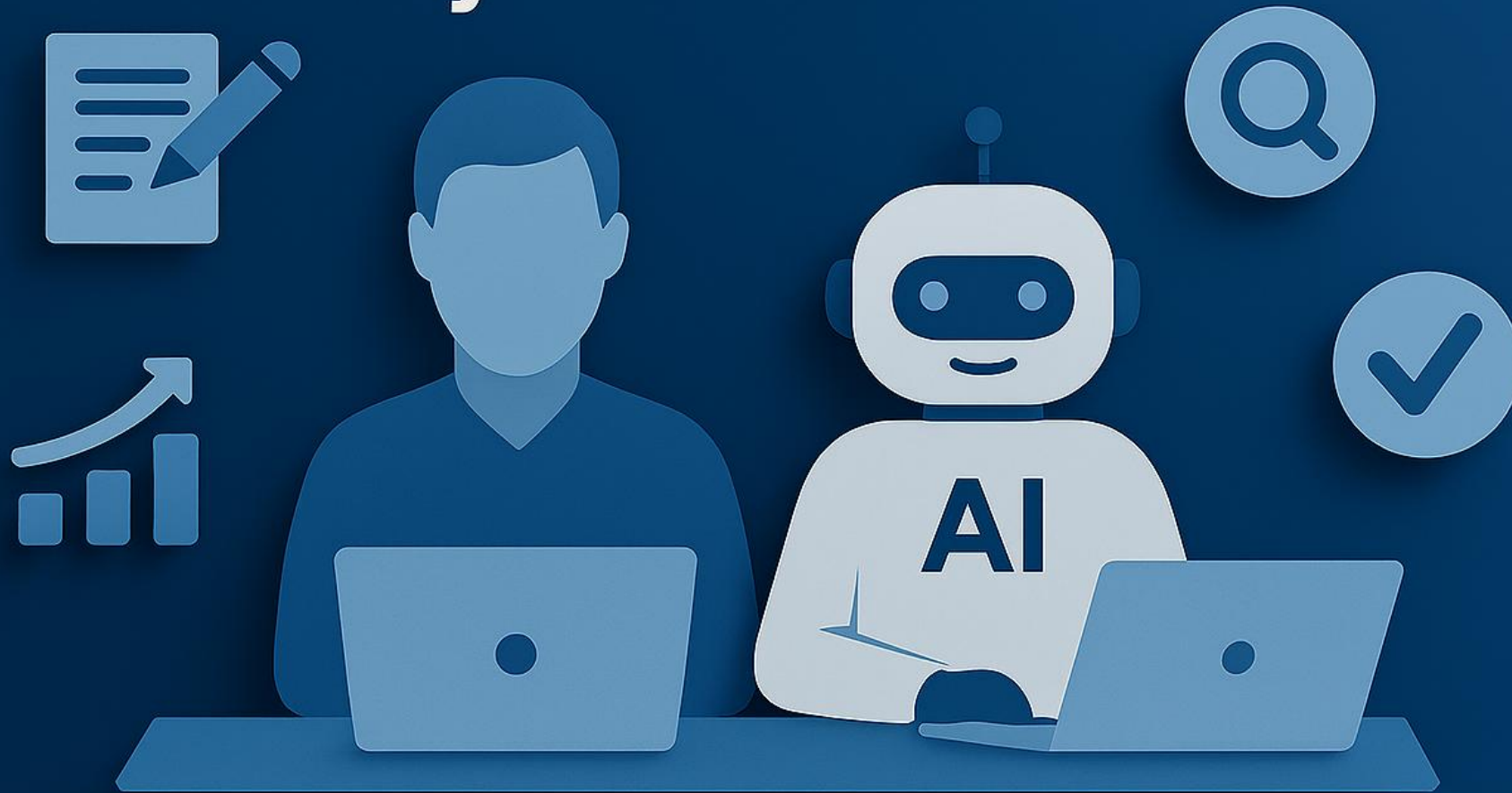
Stakeholder engagements and field evaluations become paralyzed by competing, unverified realities.

Impact on Evidence

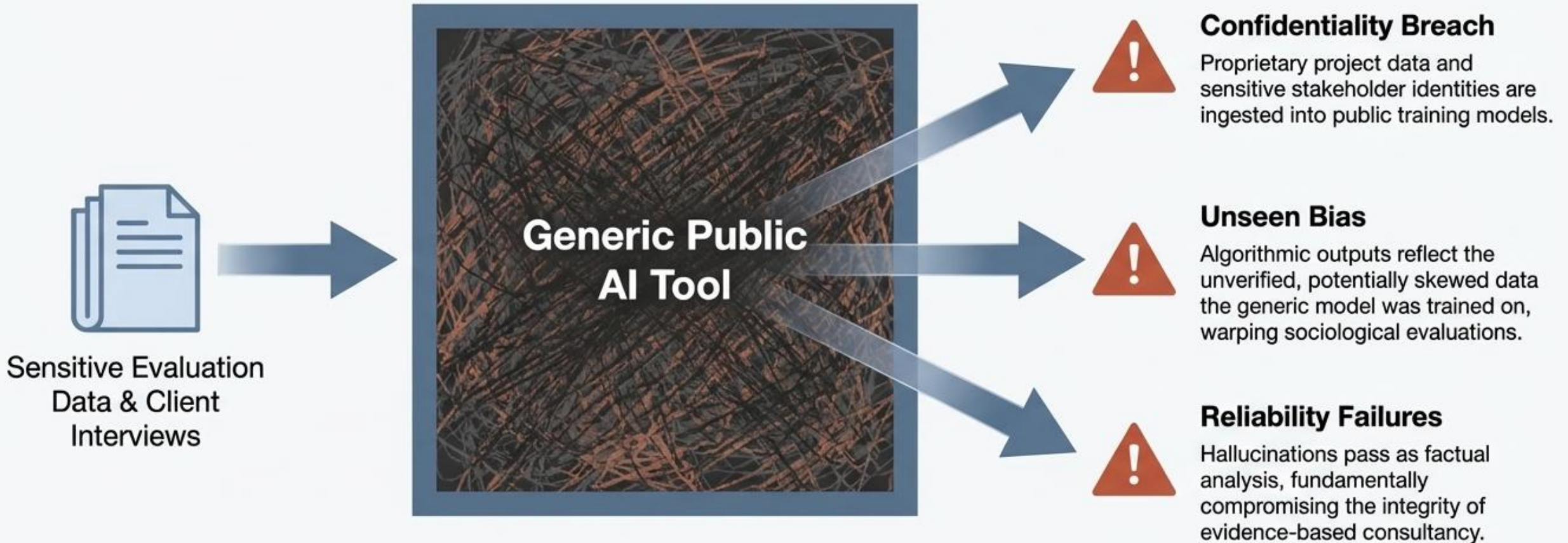
Baseline data and community feedback are increasingly contaminated by synthesized or coordinated online manipulation.



How do you use AI in your work?



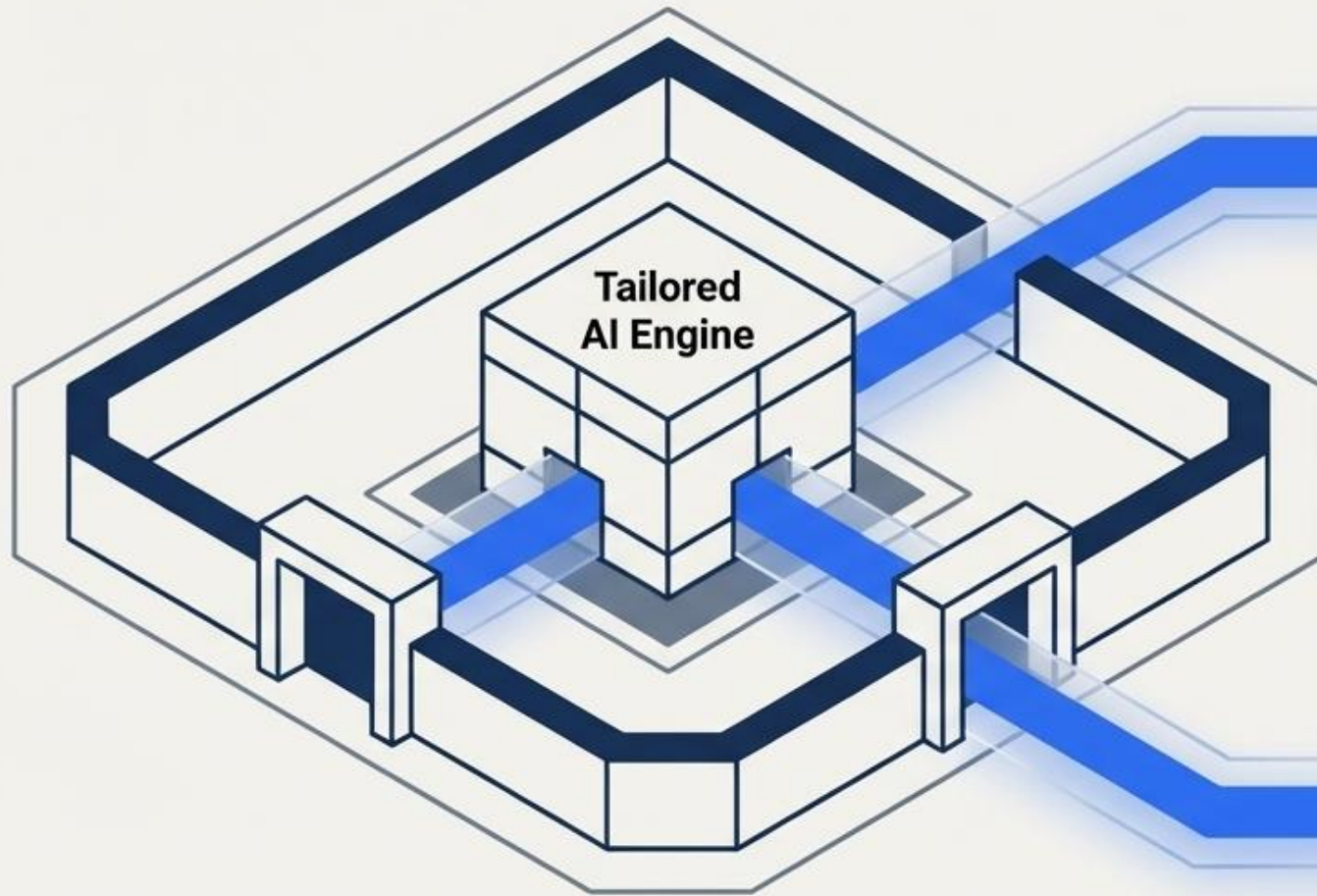
The unregulated use of commercial AI introduces severe vulnerabilities into evaluation workflows



Protecting methodological integrity requires shifting from generic to **bespoke** AI models

	Generic AI	Tailored Consultancy AI
Data Environment	✗ Public, shared servers. High risk of data leakage.	✓ Localized, secure sandbox. Absolute confidentiality.
Training Baseline	✗ Unauditable internet scraping (prone to disinformation).	✓ Exclusively trained on verified methodologies, institutional frameworks, and rigorous project data.
Bias Control	✗ Opaque parameters; researchers cannot trace the logic.	✓ Transparent weighting; researchers control and audit the analytical lens.
Output Reliability	✗ Highly vulnerable to confident hallucinations.	✓ Output is strictly anchored to the specific evaluation mandate.

Developing proprietary AI tools presents a strategic opportunity for repositioning



Research & Analysis

AI deployed as a secure synthesizer for large-scale, multilingual datasets, accelerating evidence gathering without compromising quality.

Training & Capacity Building

Bespoke models used to simulate complex policy scenarios and train evaluators on identifying digital disinformation.

Evaluation Rigour

Automated stress-testing of evaluation frameworks, cross-referencing findings against established geopolitical and sociological baselines.

Integrating AI4DEBUNK insights into our operational methodology

1 

Acknowledge the Landscape

Recognize that globalised conspiracy narratives actively distort field dialogue and baseline trust.

2 

Audit the Tools

Immediately reassess the risks of utilizing generic AI models for processing sensitive client evaluations.

3 

Build the Future

Strategically invest in the development of tailored, secure AI methodologies built explicitly for research environments.

Strategic Discussion: How will your practice area adapt its validation frameworks to guarantee integrity in an AI-accelerated evaluation cycle?

Navigating tensions while maintaining integrity

The landscape of disinformation will continue to evolve, and the noise will grow louder.

Our mandate is not to avoid new technologies, but to subjugate them to our highest standards of rigour.

We must build the tools, secure the data, and rigorously defend the evidence.

Thank You

Pascaline Gaborit- Pilot4dev

 ai4debunk

 AI4Debunk

 ai4debunk.eu

 Pascaline.Gaborit@pilot4dev.com - + 32 497 57 82 70



Funded by the European Union under the Grant Agreement No 10110509.
Views and opinions expressed are however those of the author(s) only
and do not necessarily reflect those of the European Union or European
Innovation Council and SMEs Executive Agency (EISMEA).
Neither the European Union nor the granting authorities can be held
responsible for them.

