



# Using AI to detect, classify, and track disinformation

*AI.DEFENSE.1 Stephen Campbell and Julian Neylan*

EU CYBERNET SUMMER SCHOOL 2025

Cyber Crisis Management:  
Navigating Disinformation and  
Cyber Attacks in the AI Era



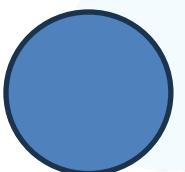
Federal Foreign Office



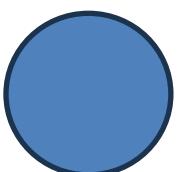
Funded by  
the European Union

# Module Outline

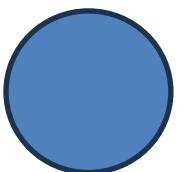
AI.DEFENSE.1: Using AI to detect classify and track disinformation



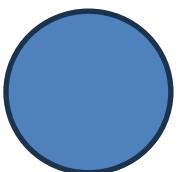
Approaches to Detecting Disinformation



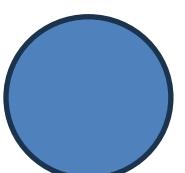
Example AI Detection Algorithms



Assessing Coordinated Inauthentic Behavior



Example Investigative Tools



Some Investigative Tool Tips

# Rumsfeld's Spectrum of Uncertainty



AI.DEFENSE.1: Using AI to detect classify and track disinformation



# The Threat Epistemology

Environment

Threat

Unknown

Known

# The Four Types of Threat Detection

Behaviors

Environment

Threat

Artefacts

Similarity-Based

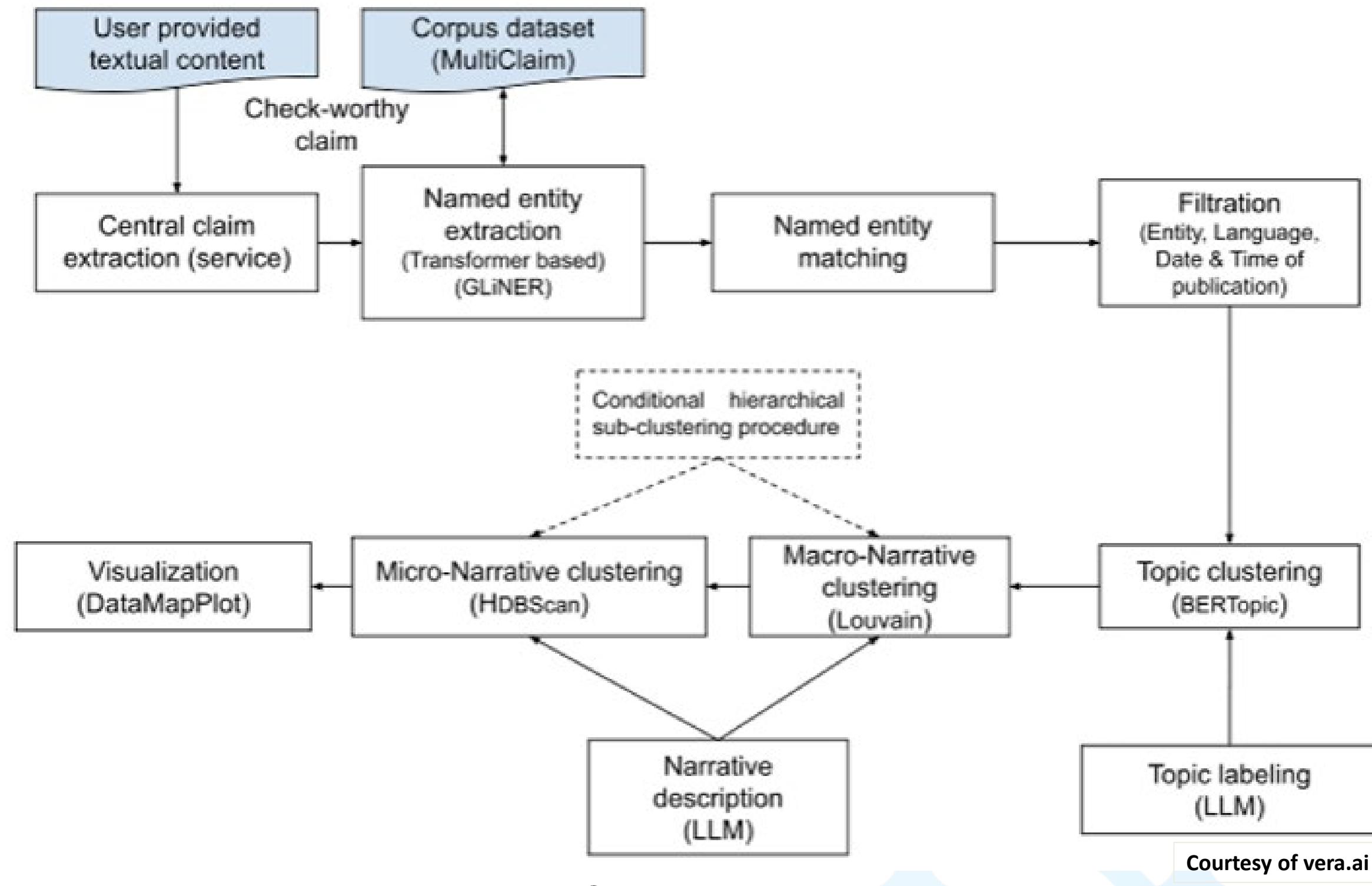
Indicator Rules

Anomaly-based

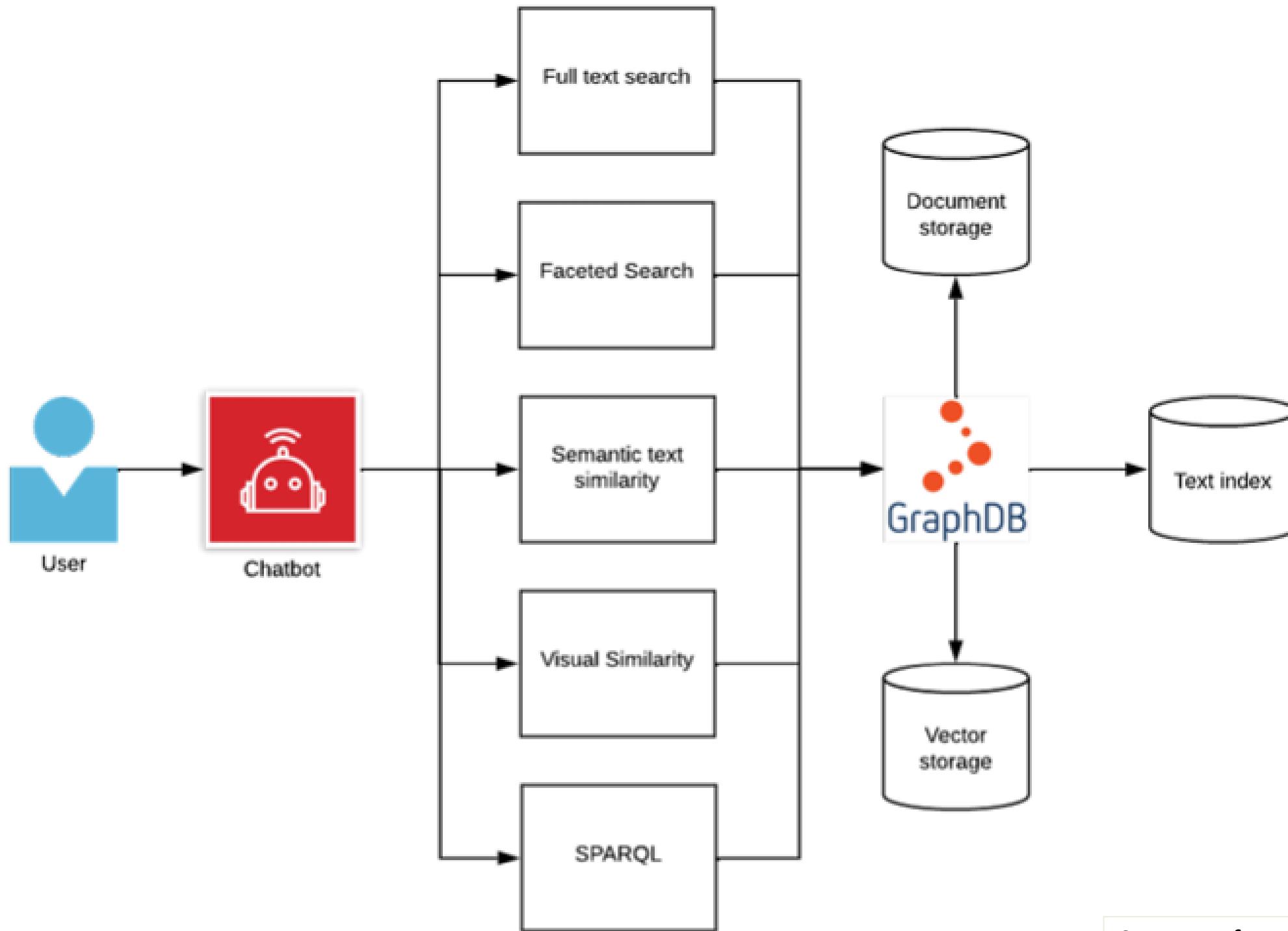
Signature-based

# Example Algorithms: Kempelen Institute Narrative Pipeline

AI.DEFENSE.1: Using AI to detect classify and track disinformation



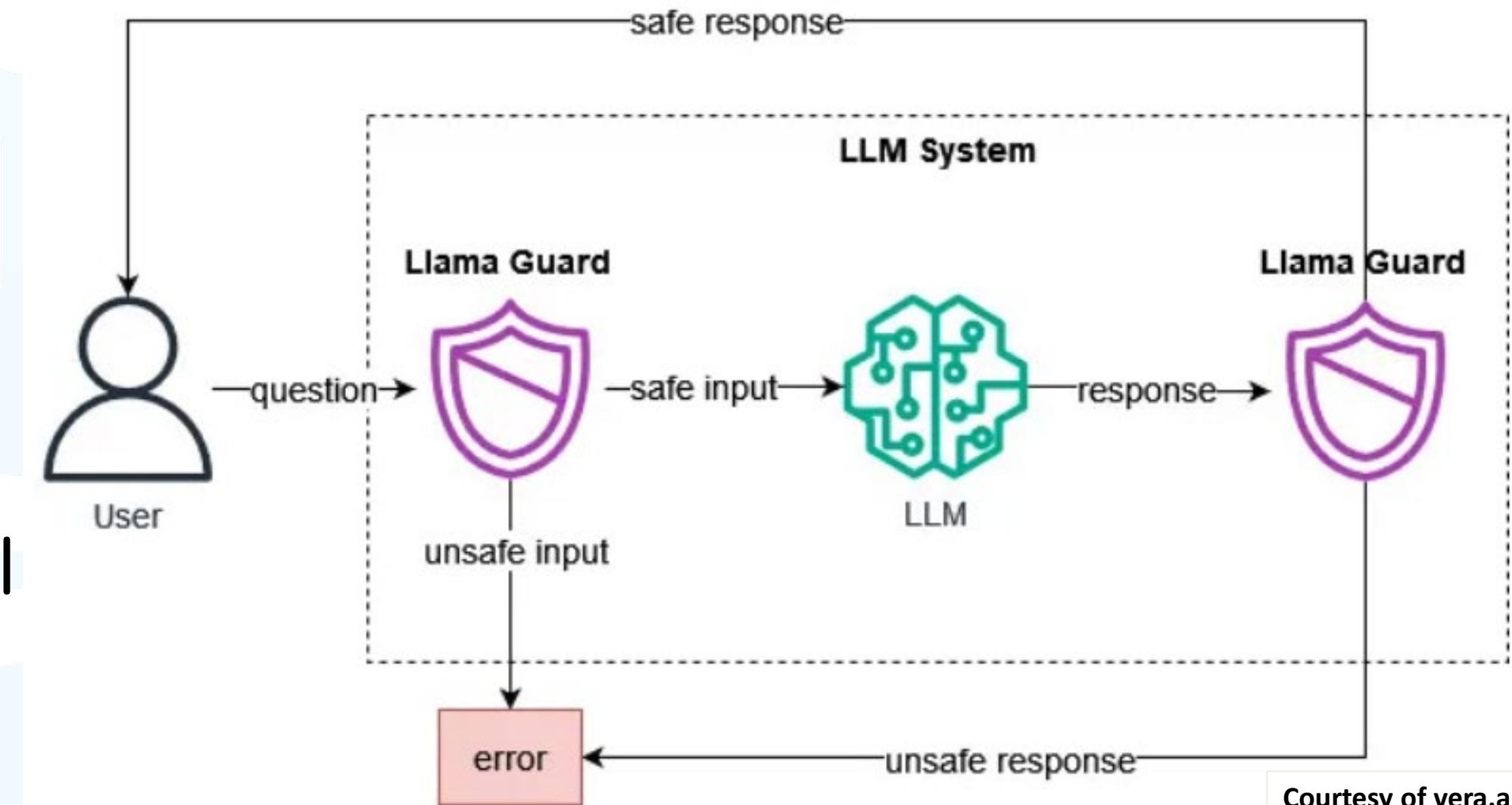
# Example Algorithms: Database of Known Fakes



Courtesy of vera.ai

# Approaches to Safeguarding LLMs

- Training Data Curation
- Guardrails
- LLM Firewalls
  - Meta Llama Guard
  - IBM Granite Guardian
  - CloudFlare Firewall for AI
  - Akamai Firewall for AI
- Machine Unlearning ?!



# Detecting CIBs

- Repost Network
  - The simplest type
- Coordinated Repost Network
  - Synchronized amplification
- Coordinated Posting Networks
  - Copypasta in action
- Coordinated Hashtag Networks
  - Hashtag hijacking
- Coordinated Hashtag Sequence Networks
  - Obfuscating Coordination
- Coordinated Mention Networks
  - Amplification and targeting through mentions
- Coordinated Link Sharing
  - Amplification through synchronized URL sharing
- Coordinated Image Sharing
  - Spamouflage
- Coordinated Account Handles
  - Handle similarity, timestamps
- Coordinated Websites
  - Interconnected domains
- Syndicated Activity
  - Temporal posting patterns

Courtesy of Jerry Gao

# Example: vera.ai CIB Assessment of Operation Overload

AI.DFENSE.1: Using AI to detect classify and track disinformation



# Coordination Assessment

## BEHAVIOURAL ANALYSIS

- Accounts created around the same time
- Similar posting timestamps across different accounts
- Sudden spikes in messaging around a certain narrative or event
- Significant activity from locations or time zones misaligned with the user base
- Repeatedly expose viewers to the same false stories to exploit the illusory truth effect

## NETWORK ANALYSIS

- Accounts engaging with each other's posts in a synchronised manner
- Tightly interconnected clusters of accounts that mostly follow each other
- Similar posting patterns across different social media platforms
- Efforts to appear organic and spontaneous

## CONTENT ANALYSIS

- Narrative alignment: Identical or similar hashtags, links, posts, images, memes or texts
- Same or similar content translated and posted in different languages
- Memetic warfare: Using evocative images and motifs to simplify complex issues, manipulate emotions, and rapidly spread misinformation
- One-topic accounts (suspicious)
- Content farming
- Centralised content production

## IDENTITY ANALYSIS

- Same or similar profile images and cover photos.

## VISUALS ANALYSIS

- Same or similar profile name images, visuals, cover photos

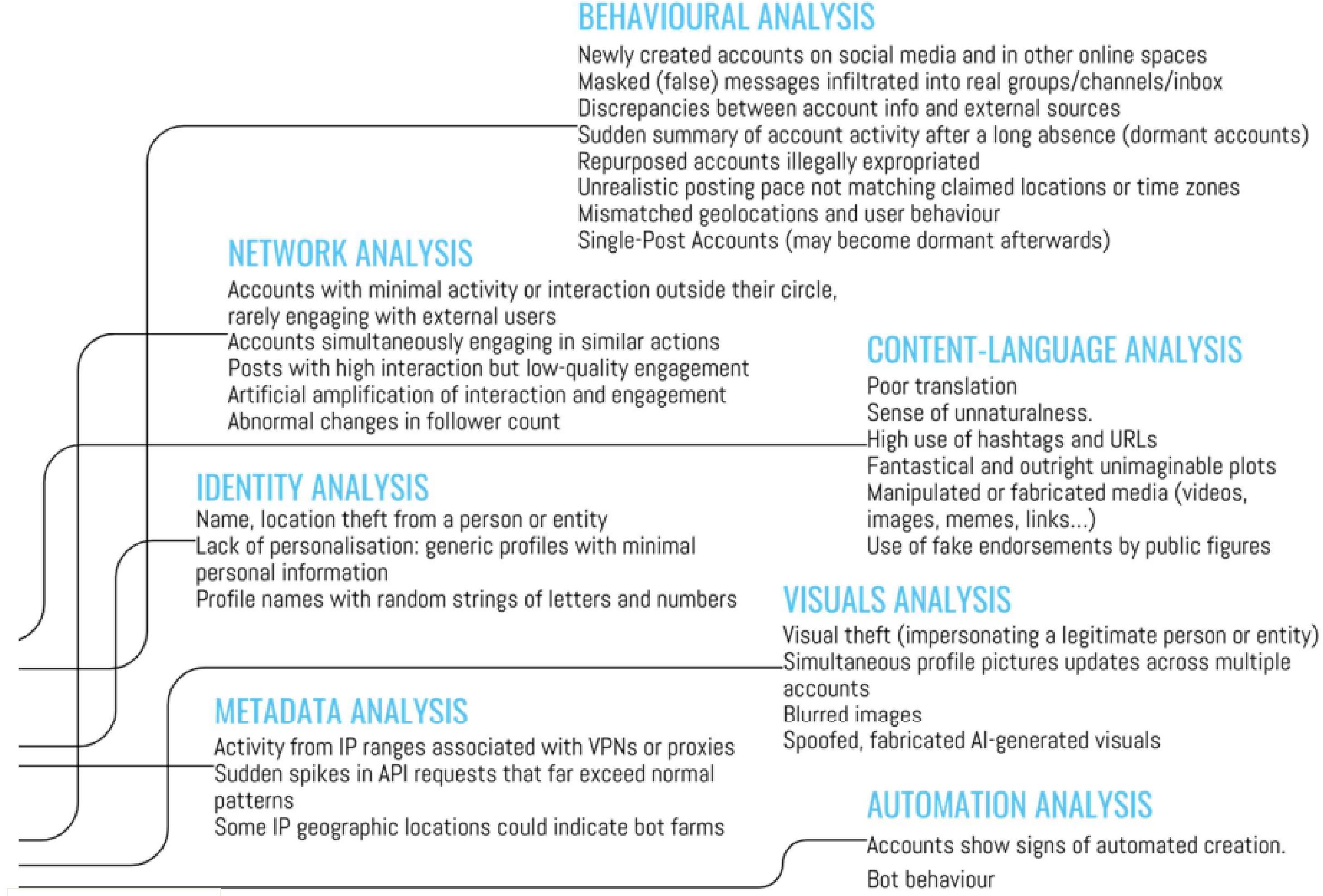
## METADATA ANALYSIS

- Multiple accounts using the same IP address, device and configurations.

## AUTOMATION ANALYSIS

- Accounts showing automated syndication

# Authenticity Assessment



# Source Assessment

AI.DFENSE.1: Using AI to detect classify and track disinformation

## NETWORK ANALYSIS

Identify primary spreaders and amplifiers, often those with extensive connections or interactions

Accounts engaging first are often linked to campaign origin

Cross-platform activity can multiply leads, aiding in source tracking

## CONTENT ANALYSIS

Parsing content to identify involved parties within context

## IDENTITY ANALYSIS

Account registration details sharing a common source

Repeat offenders may be documented in fact-checking databases or open-source archives

## VISUALS ANALYSIS

Background or profile images may offer source clues through recognisable elements (like faces or places)

## METADATA ANALYSIS

Monitor IP addresses, user agents, timestamps, and request details

Track the geographical distribution of API requests to identify physical locations

Identify the Autonomous System Numbers associated with suspicious IP addresses

Metadata analysis of images, videos, and origins of shared links

Search the origin of shortened URLs to identify common sources

# Impact Assessment

## CONTENT ANALYSIS

Extreme content polarisation  
The campaign content aims at a specific target

## BEHAVIOURAL ANALYSIS

Peripheral accounts amplifying the content of the core account/s  
Accounts having specific roles in the amplification process

## NETWORK ANALYSIS

Unusual volumes of likes or reshares  
Media amplification  
Public figures/influencers amplification  
Various methods from sentiment analysis to tracking trending hashtags  
Track backlinks  
Use multiple social media platforms to maximise outreach  
Only certain viewpoints are amplified while dissenting opinions are drowned out

# Without AI you use Manual Tracking

Google Searches (or Google Dorking)

Looking for Keywords and setting up the parameters for the search to give the best results

Manually looking at all of the accounts with certain content or features of interest

EU CYBERNET SUMMER SCHOOL 2025

**Cyber Crisis Management:**  
**Navigating Disinformation and**  
**Cyber Attacks in the AI Era**



Federal Foreign Office



Funded by  
the European Union

# However there are also a broad range of tools that can be used

EU CYBERNET SUMMER SCHOOL 2025

**Cyber Crisis Management:**  
**Navigating Disinformation and**  
**Cyber Attacks in the AI Era**



Federal Foreign Office



Funded by  
the European Union

# Example: Google Search for LLM-generated Content

"As an AI language model" inurl:post site:linkedin.com

About 81 results (0,39 seconds)

**Carl Beien's Post**

15 Apr 2023 — As an AI language model, I'm programmed to provide helpful and accurate information that promotes human well-being, including the importance of ...

**Nikhil Manek FCA'S Post**

++++++ As an AI language model developed by OpenAI, ChatGPT has received widespread attention in recent years for its advanced natural language processing ...

**Shelby Steiger's Post**

Comment what your thoughts on the AI's excerpt As an AI language model myself, I am excited to witness the rapid growth and development of Artificial ...

**Ken Grace on LinkedIn: I just read a LI post that says "AI writing ...**

As an AI language model, I strive to provide accurate information. In this case, my previous answers were incorrect, and I apologize for any confusion they ...

- "as an AI language model"
- "not a recognized word"
- "cannot provide a phrase"
- "with the given words"
- "violates OpenAI's content policy."
- "I'm sorry, I cannot generate"
- "The message you submitted"
- "An error occurred. If this issue persists please contact us through our help center at"
- "Something went wrong, please try reloading the conversation."
- "I'm sorry, but I cannot fulfill this request"
- "Thank you for starting it" comes from "This is such an important conversation that needs to happen more often. Thank you for starting it."

Courtesy of Nico Dekens

# Bellingcat toolkit

 Bellingcat's Online Investigation Toolkit

[Home](#)

[New Tools](#)

 CATEGORIES

[Maps & Satellites](#)



[Geolocation](#)

[Image/Video](#)



[Social Media](#)



[People](#)

[Websites](#)

[Companies & Finance](#)

[Conflict](#)

 CATEGORIES

 Copy 

## Social Media

Tools for one or more social media platforms

[Facebook](#) >

[Instagram](#) >

[Telegram](#) >

[Tiktok](#) >

[Twitter/X](#) >

[Youtube](#) >

[Other Platforms](#) >

[Multiple Platforms](#) >

EU CYBERNET SUMMER SCHOOL 2025

**Cyber Crisis Management:**  
Navigating Disinformation and  
Cyber Attacks in the AI Era



Federal Foreign Office



Funded by  
the European Union

# Information Laundromat

The screenshot shows the Information Laundromat website. On the left, there is a dark background with white text that reads: "Discover shared metadata from across the infosphere". Below this, a smaller text block describes the tool: "The Information Laundromat is a lead generation tool used to determine if and how websites share architecture and content. It provides two core functions: content similarity and domain forensics matching". On the right, the main search interface is displayed. It features two tabs at the top: "Content Similarity" (white) and "Metadata Similarity" (dark grey). A descriptive text block below the tabs explains the search functionality. The search form includes a text input field labeled "Add article link, title or a text snippet", three dropdown menus for "Country" (United States), "Language" (English), and "Engines" (All selected (8)), and a "Search" button. At the bottom of the search form, there is a note: "Please [log in or register](#) to run batch searches. Contact us at info [at] securingdemocracy.org to obtain a registration code." The URL in the browser's address bar is <https://securingdemocracy.org/information-laundromat>.

EU CYBERNET SUMMER SCHOOL 2025

**Cyber Crisis Management:**  
Navigating Disinformation and  
Cyber Attacks in the AI Era



Federal Foreign Office



Funded by  
the European Union

# Hamilton 2.0 Dashboard

## Search Articles and Social Media Posts

Press 'Enter' or 'Search' to search. Adding filters, including checkboxes and dates, requires a new search.

SEARCH

clear filter(s)

01/01/2023 - 02/01/2025

User Category
Platform
Monitored Network

Media (4,632,961)

Websites (1,123,332)

Russia (3,202,818)

Websites (1,123,332)

China (2,126,195)

Facebook (829,907)

Iran (1,378,001)

YouTube (194,789)

Instagram (128,370)

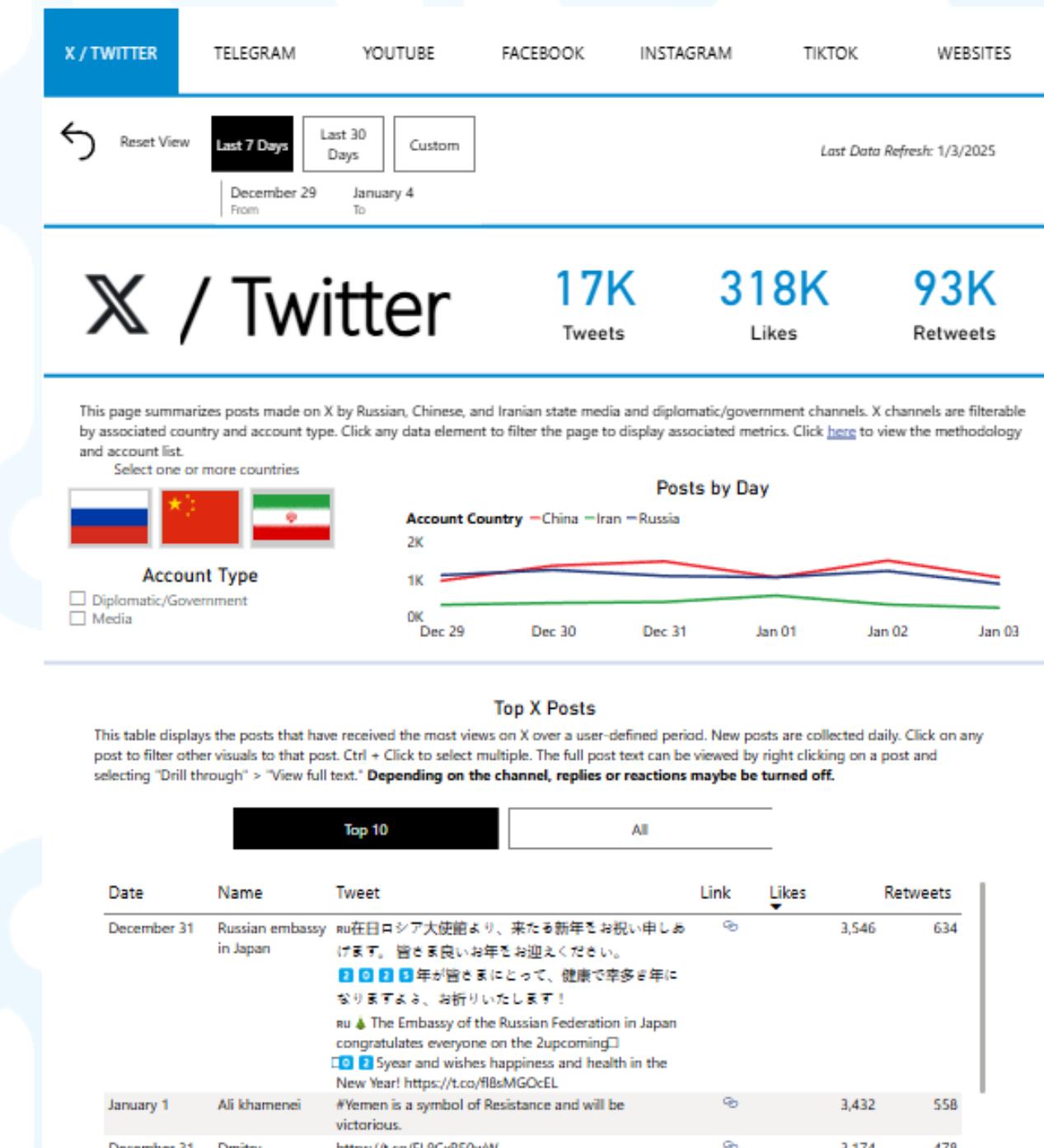
TikTok (2,421)

TOGGLE ADDITIONAL FILTERS & LEGEND

Relevance

1 - 50 of 6,707,014

| Name                         | Date       | Channel  | Country | Post Preview   |
|------------------------------|------------|----------|---------|--|
| PressTV Francais [PressTVFr] | 01-24-2024 | Telegram | Iran    | Josep Borrell: Israel cannot 'have its veto' on Palestinian state  Josep Borrell, the European Union's (EU) foreign policy chief, spoke on the Palestinian issue at a joint press conference, saying that the Israeli regime "Israel cannot have a veto over the self-determination of the Palestinian people."  For more details @PressTVFrancais |
| PressTV Francais [PressTVFr] | 01-24-2024 | Telegram | Iran    | ► Drones and fighter jets Made in Burkina! (Zoom Africa) The fighter jet project is already well advanced, but some important elements are missing. That's why the manufacturer decided to turn to Russia.  For more details @PressTVFrancais  |
| PressTV Francais [PressTVFr] | 01-24-2024 | Telegram | Iran    | New research satellite launched 'successfully': nothing stopping Iran from moving forward (Debate) The Soraya satellite was placed in orbit 750 kilometers above the Earth, a first for Iran "beyond 500 kilometers."  For more details @PressTVFrancais   |



# EU vs Disinfo Database



Articles

Database

Learn

Research

Videos

Guest content

# Database i



# WeVerify

The screenshot shows the WeVerify Tools interface. At the top, there are navigation links: WeVerify logo, InVID logo, TOOLS (highlighted in green), ASSISTANT, TUTORIAL, DEMO, CLASSROOM, ABOUT, English, and a language switcher. Below the header, a teal wrench icon and the word "Tools" are displayed. To the right, there is a message about advanced tools being unlocked and a red "LOGOUT" button. A navigation bar below the header includes "Video" (selected), "Image", "Search", and "Data Analysis". The main content area contains five tool cards:

- Video analysis**: It brings you the contextual information of a YouTube, Facebook or Twitter video.
- Keyframes**: It fragments a YouTube, Facebook or Twitter video or a mp4 file into keyframes for reverse image search.
- Thumbnails**: It extracts and performs a reverse search of the thumbnails of a YouTube video.
- Metadata**: It extract metadata for jpeg images and videos (in mp4 or m4v format).
- Video rights**: It provides information about the legal rights of a YouTube or Twitter video.

A "Feedback" button is located at the bottom right of the main content area.

EU CYBERNET SUMMER SCHOOL 2025

**Cyber Crisis Management:  
Navigating Disinformation and  
Cyber Attacks in the AI Era**



Federal Foreign Office



Funded by  
the European Union

# Viginum D3lta

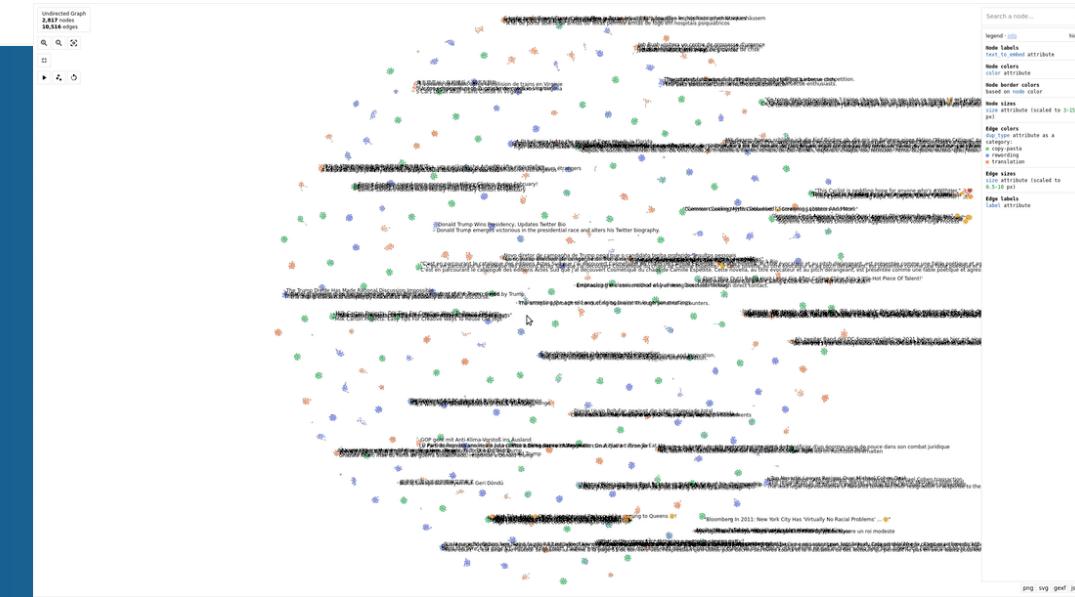
## D3lta

If you like our project, please give us a star ⭐ on GitHub for the latest update.



This repository is the official implementation of D3lta, a library for detecting duplicate verbatim contents within a vast amount of documents.

It distinguishes 3 types of duplicate contents : copypasta (almost exact duplicates), rewording and translation. You can run it on CPU.



# Commercial Detection Platforms

Social Listening Tools: Meltwater, Brandwatch, Junkopedia

EU CYBERNET SUMMER SCHOOL 2025

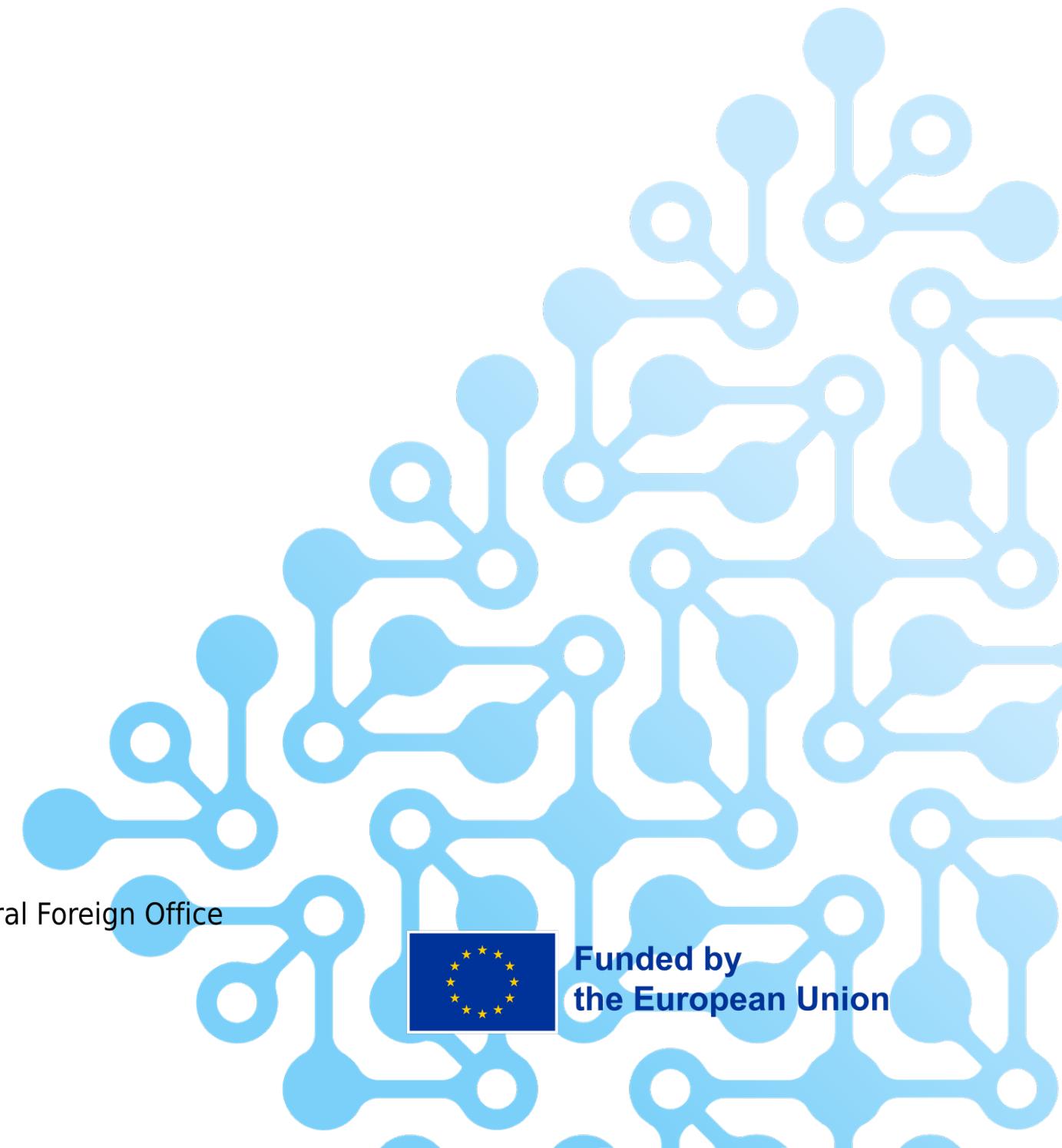
**Cyber Crisis Management:**  
Navigating Disinformation and  
Cyber Attacks in the AI Era



Federal Foreign Office



Funded by  
the European Union



Third Party Search:

[!\[\]\(efafcae43acae17c4bb9f41420411b00\_img.jpg\) General](#)[!\[\]\(13a9156b5701358ad5df1ac9471f3466\_img.jpg\) Email](#)[!\[\]\(30dfa619cea8b8790c5e9066d4f2637a\_img.jpg\) Domain](#)[!\[\]\(aa2545022aef75b49485a583e359a0ff\_img.jpg\) IP](#)[!\[\]\(11180f88349a0f55a115986a3613acf7\_img.jpg\) Bitcoin](#)

### Telegram

[Search via Google CSE 1](#)[Search via Google CSE 2 Telegago](#)

EU CYBERNET SUMMER SCHOOL 2025

**Cyber Crisis Management:**  
Navigating Disinformation and  
Cyber Attacks in the AI Era



Federal Foreign Office



Funded by  
the European Union

# Telemetry

Telemetry ▾

Product

Pricing

Docs ▾

Request Chats

Blog

# Telemetry<sup>↗</sup>

Search for messages



✉️ Messages

📣 Channels

〽️ Analytics

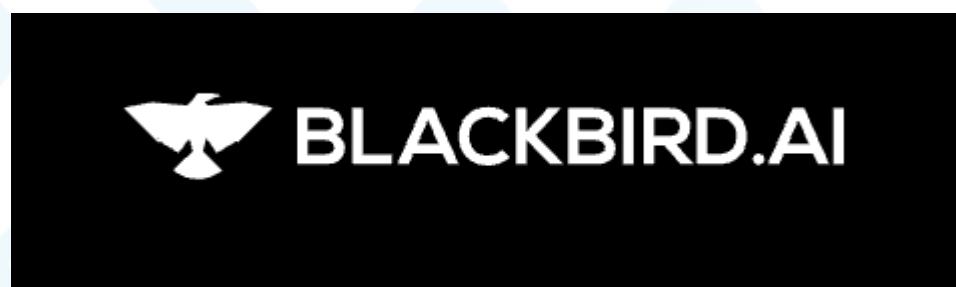
EU CYBERNET SUMMER SCHOOL 2025  
**Cyber Crisis Management:**  
Navigating Disinformation and  
Cyber Attacks in the AI Era



Federal Foreign Office



# A number of tools on the market (this is not all)



# Things to note (tool tips)





# Most of the time deepfake detection requires the right tools

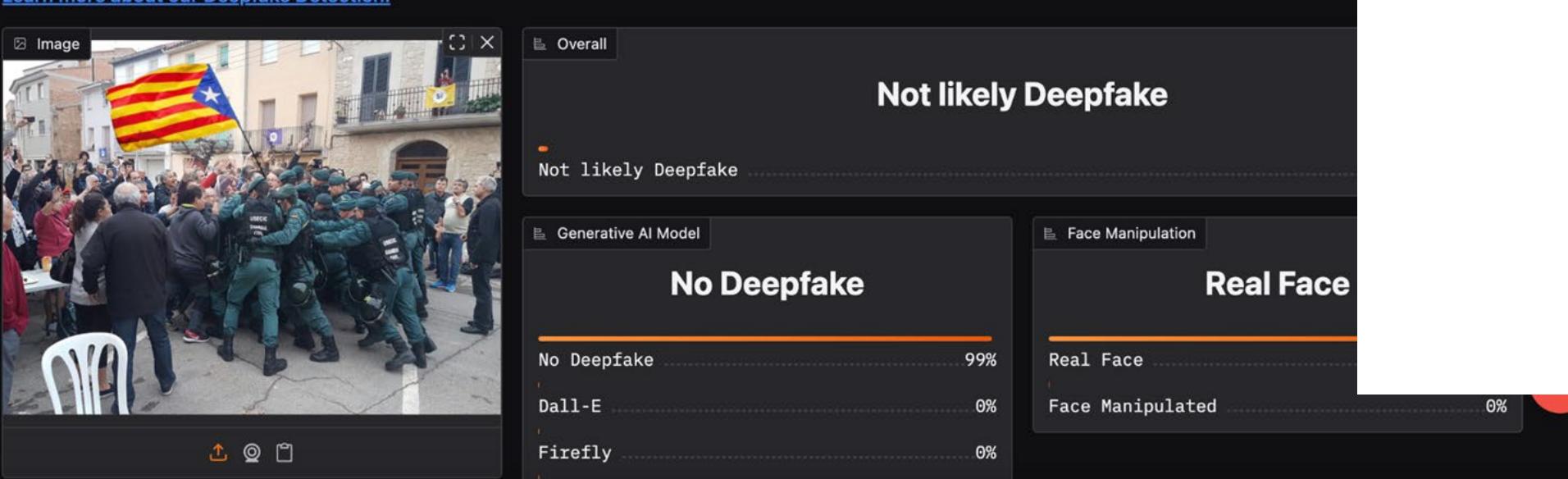
Audio for audio, Video for video: in a video with audio try to analyze separately

## Deepfake Detector Panel

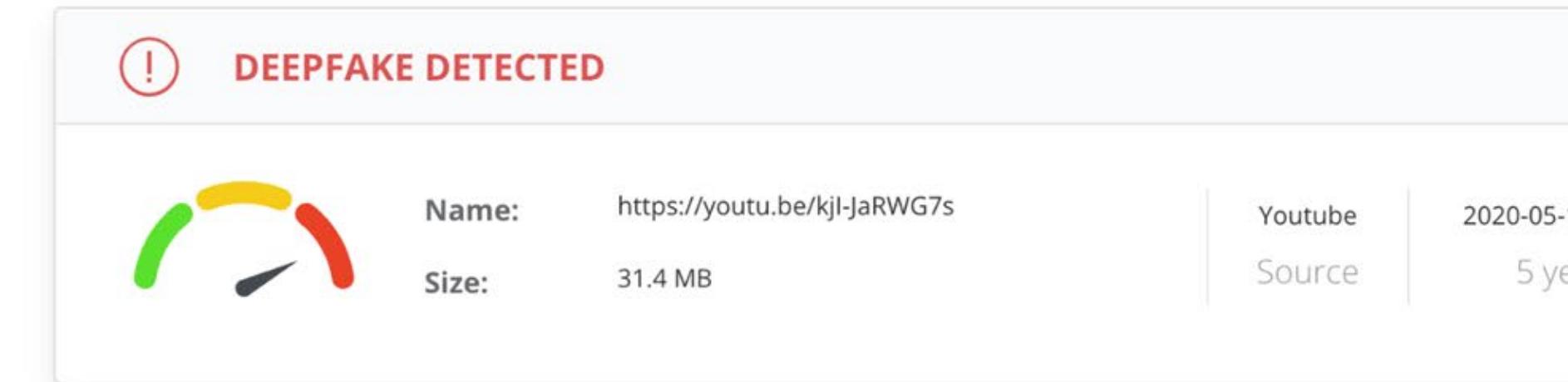
deepware®

DeepFake Detector - ❤️ Like above if this space helps

[Learn more about our Deepfake Detection.](#)



The screenshot shows a Deepfake Detector interface. At the top left is a thumbnail of a protest scene with a person holding a flag. To the right of the thumbnail are two main sections: "Overall" and "Generative AI Model". The "Overall" section displays a large green bar indicating "Not likely Deepfake" with a confidence level of 99%. Below this, under "Generative AI Model", there are two sub-sections: "Face Manipulation" and "No Deepfake". The "Face Manipulation" section shows a green bar for "Real Face" at 100% and a red bar for "Face Manipulated" at 0%. The "No Deepfake" section shows a green bar for "No Deepfake" at 99%, a red bar for "Dall-E" at 0%, and a red bar for "Firefly" at 0%. At the bottom are download and sharing icons.



The screenshot shows a Deepfake Detector interface with a red exclamation mark icon and the text "DEEFAKE DETECTED". Below this is a circular progress bar with green and yellow segments. To the right, the file details are listed: Name: <https://youtu.be/kjl-JaRWG7s>, Size: 31.4 MB. On the far right, there are links for "Youtube Source" and a timestamp "2020-05-15 5 ye".

# Cyber Crisis Management:

## Navigating Disinformation and Cyber Attacks in the AI Era



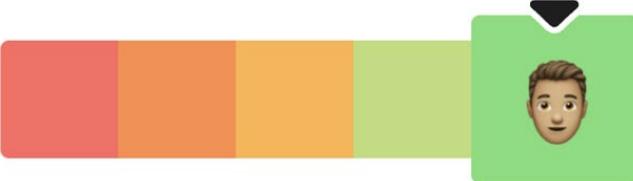
Federal Foreign Office

Funded by  
the European Union

# Tools may not always be accurate

Check if image is AI-generated

We're quite confident that NO AI was used  
when producing this image.



RESET

Was this picture generated by ai or humans?

Try these examples: MidJourney DALL-E Stable Diffusion Flux Human

## Analyzed Image


[PDF](#) [PDF Report](#)

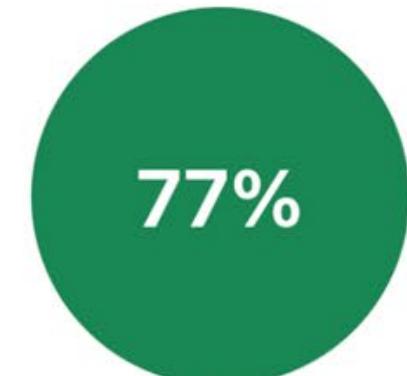
Prediction

real

This image is highly likely to be a real  
photo taken by a human.

Upload a image ↑

## Analysis Results



Authenticity Score

### Analysis Breakdown

Metadata

85%

Analysis of image properties and format

Noise

48%

Evaluation of natural image noise patterns

Compression

94%

Assessment of compression artifacts

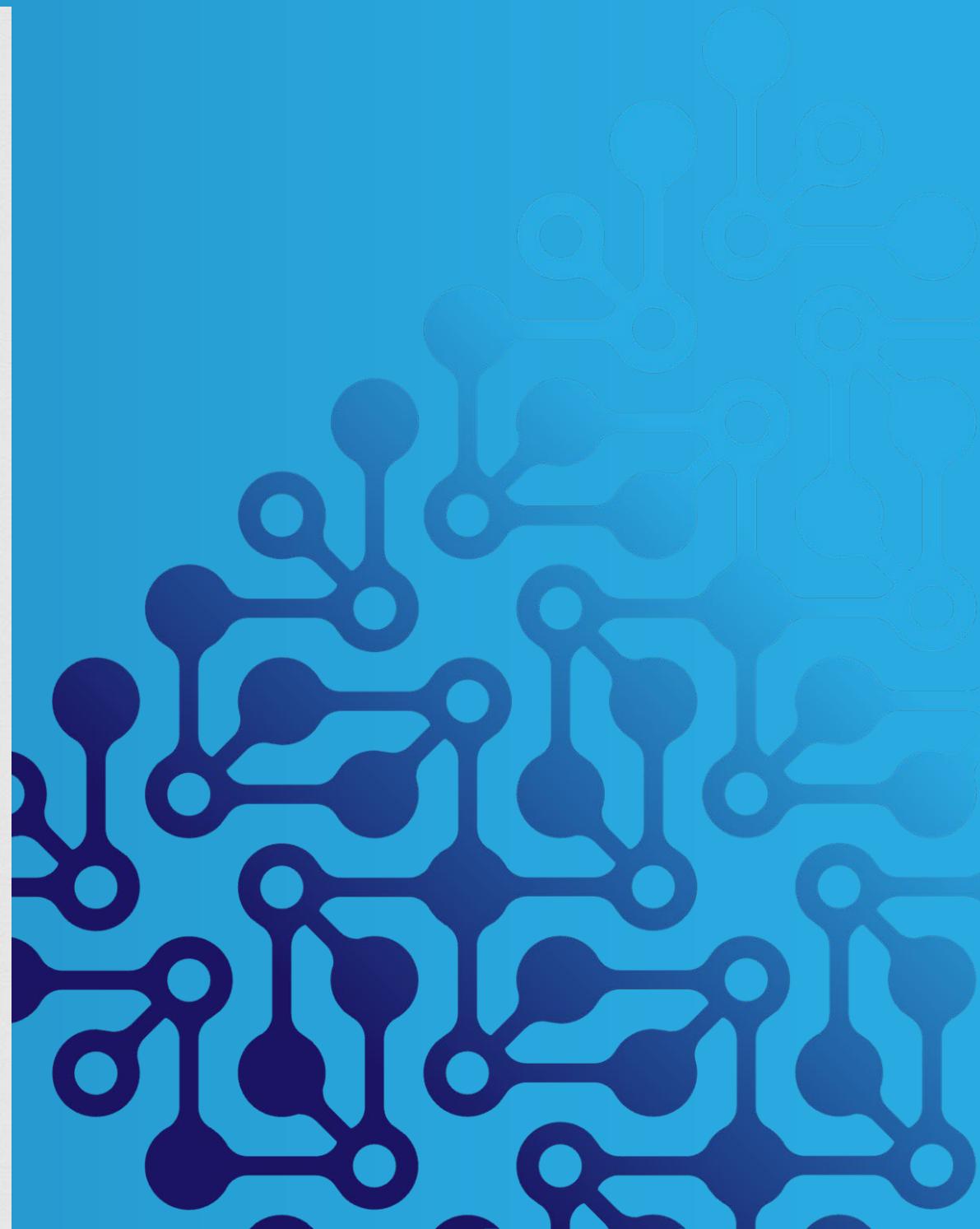
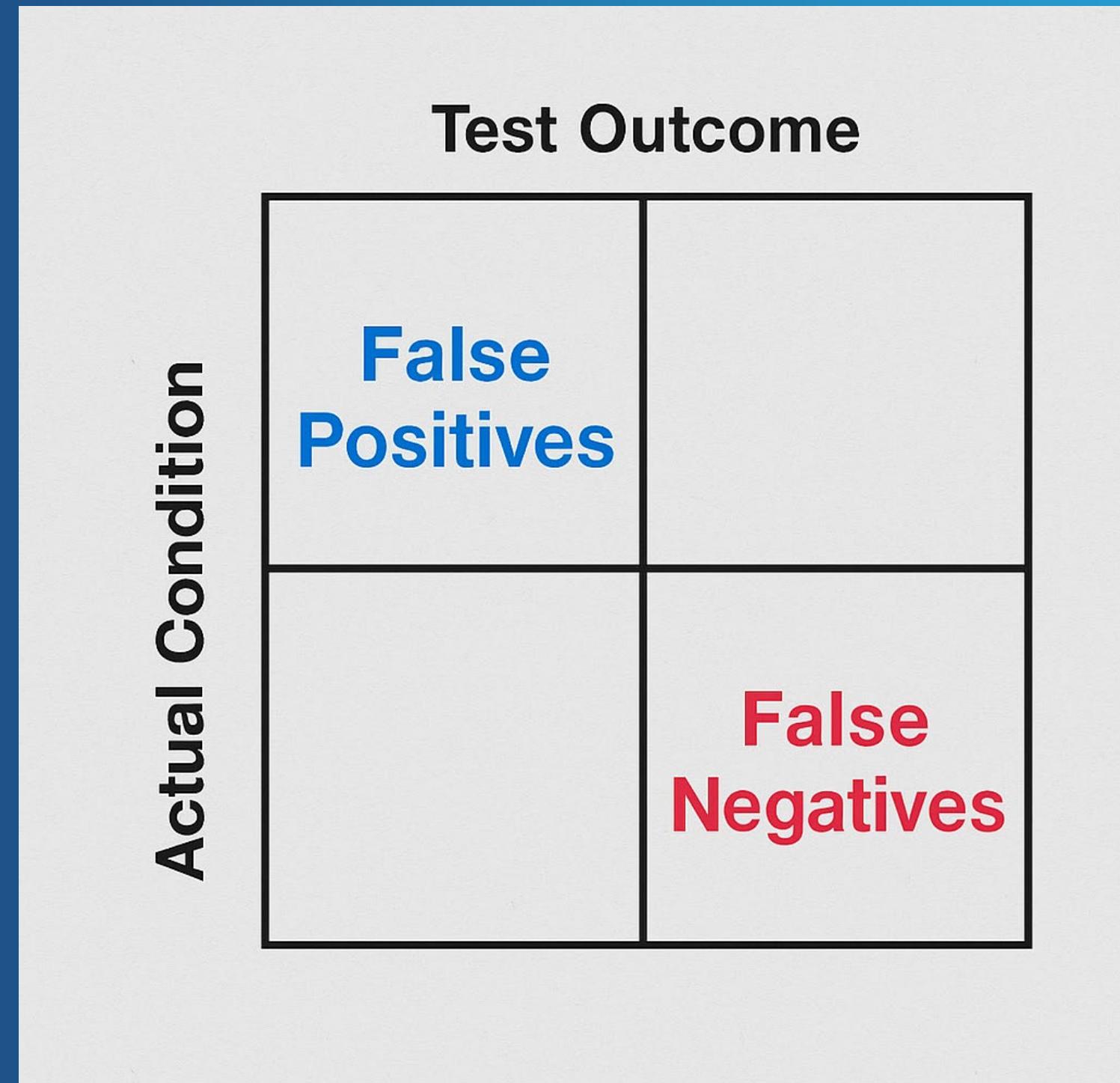
Authenticity

81%

Detection of artificial patterns



You can get both false positives and false negatives



**Cyber Crisis Management:  
Navigating Disinformation and  
Cyber Attacks in the AI Era**



Federal Foreign Office



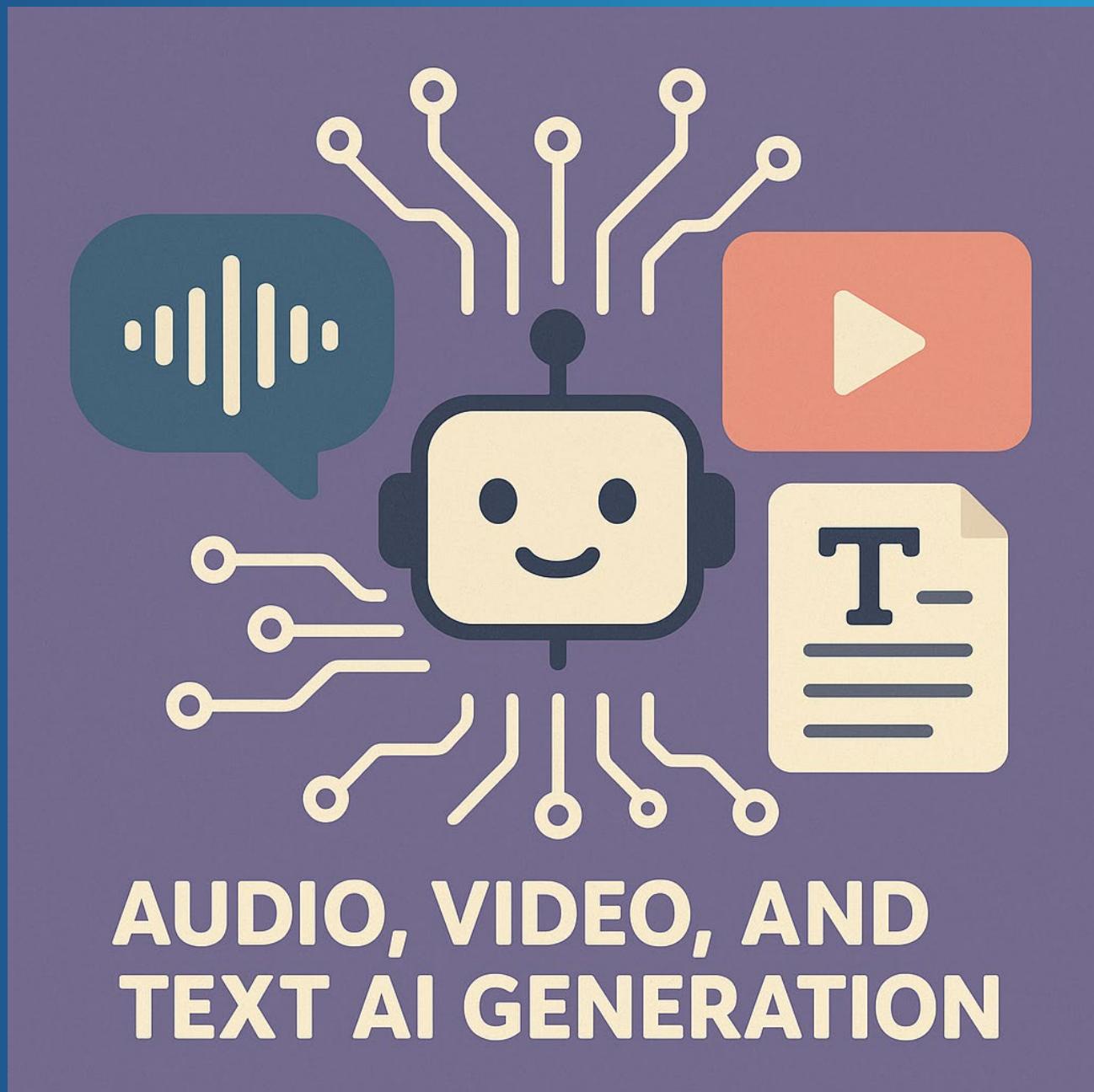
Funded by  
the European Union

# Image quality matters





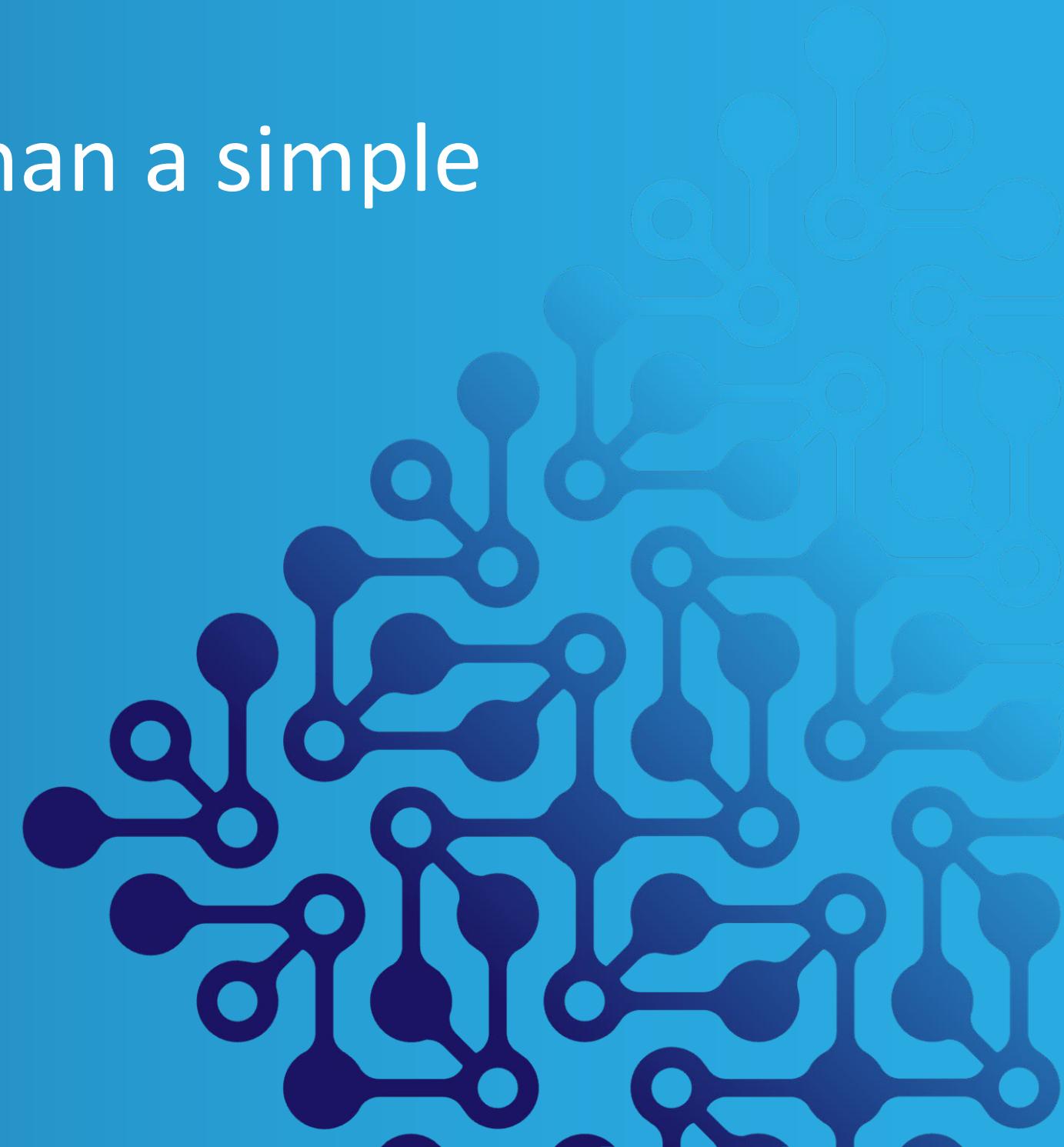
# Tools may not have caught up with new tech





# Provenance is often better than just detection

Look for the history of publication rather than a simple reading of if it was published or not



# Cyber Crisis Management: Navigating Disinformation and Cyber Attacks in the AI Era

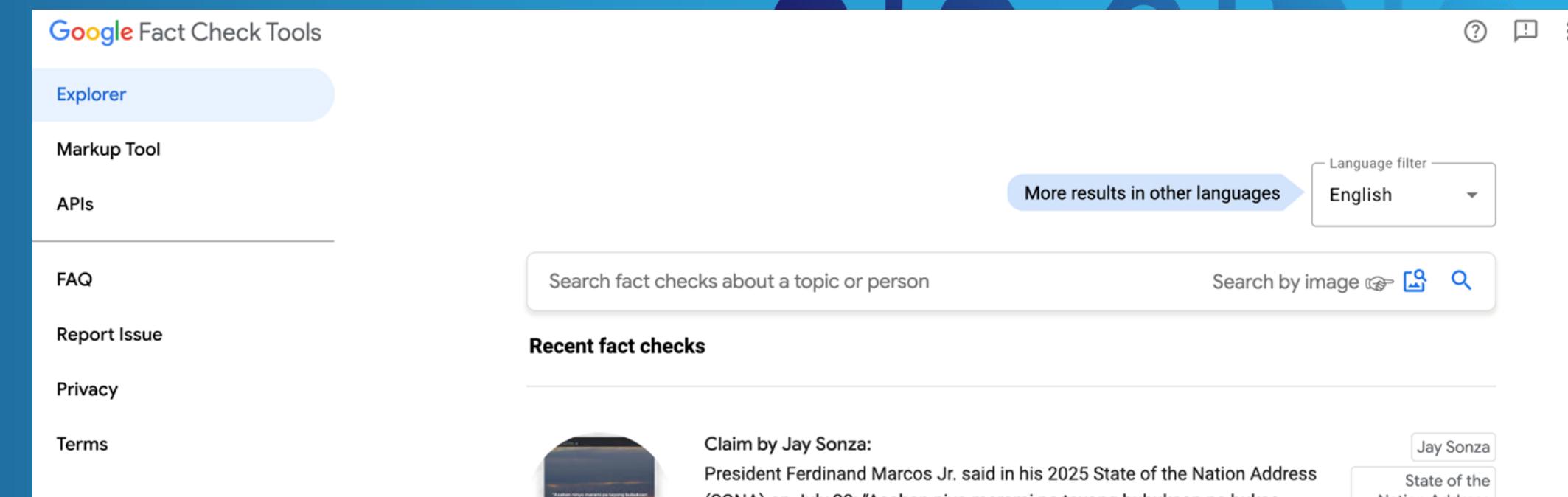


Federal Foreign Office



Funded by  
the European Union

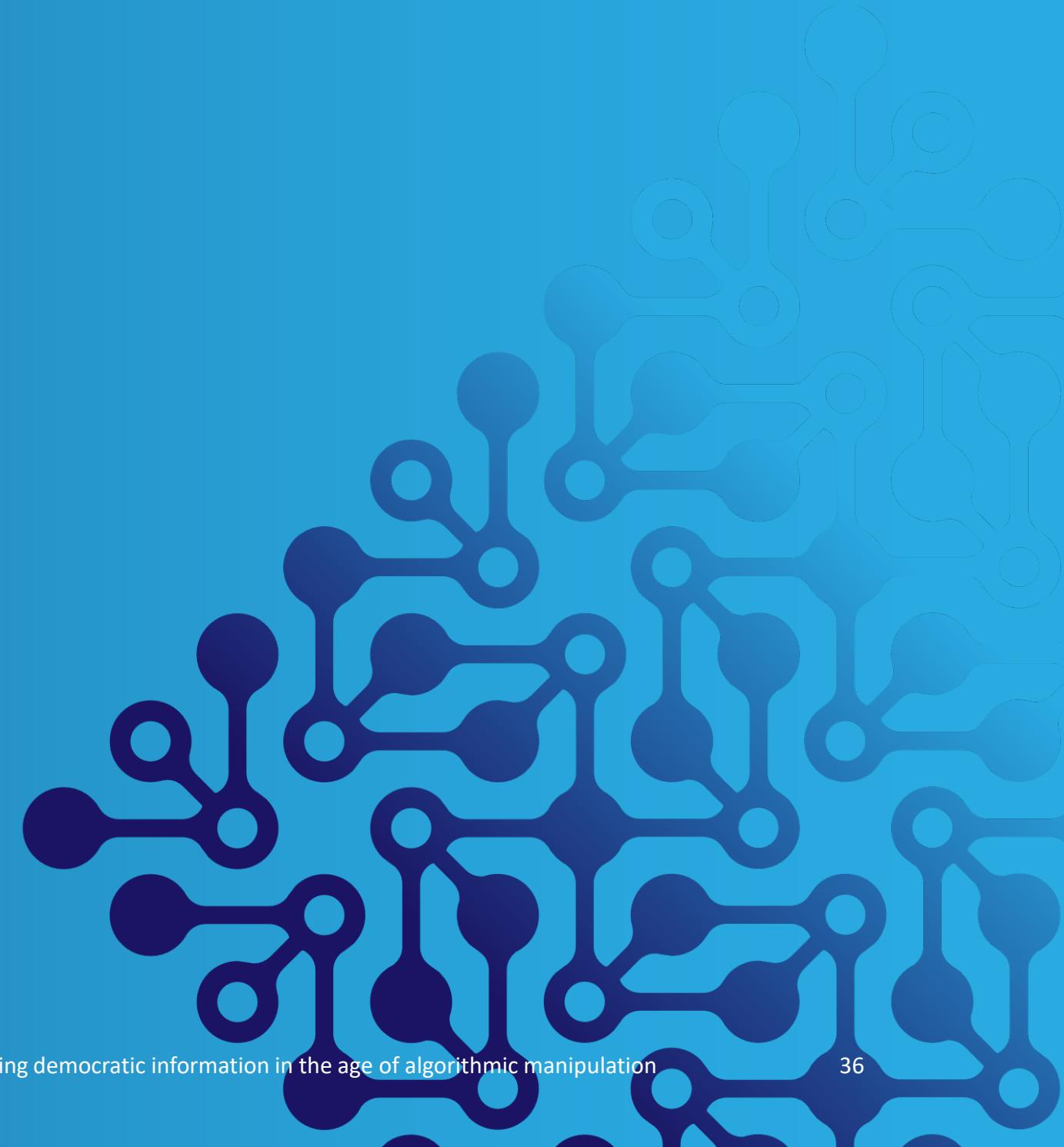
## Cross checking many tools and sources



The screenshot shows the Google Fact Check Tools homepage. The top navigation bar includes links for "Explorer", "Markup Tool", and "APIs". On the left, there's a sidebar with links for "FAQ", "Report Issue", "Privacy", and "Terms". The main content area features a search bar with the placeholder "Search fact checks about a topic or person", a "Recent fact checks" section, and a "More results in other languages" button with a dropdown menu set to "English". There are also "Language filter" and "Search by image" buttons.

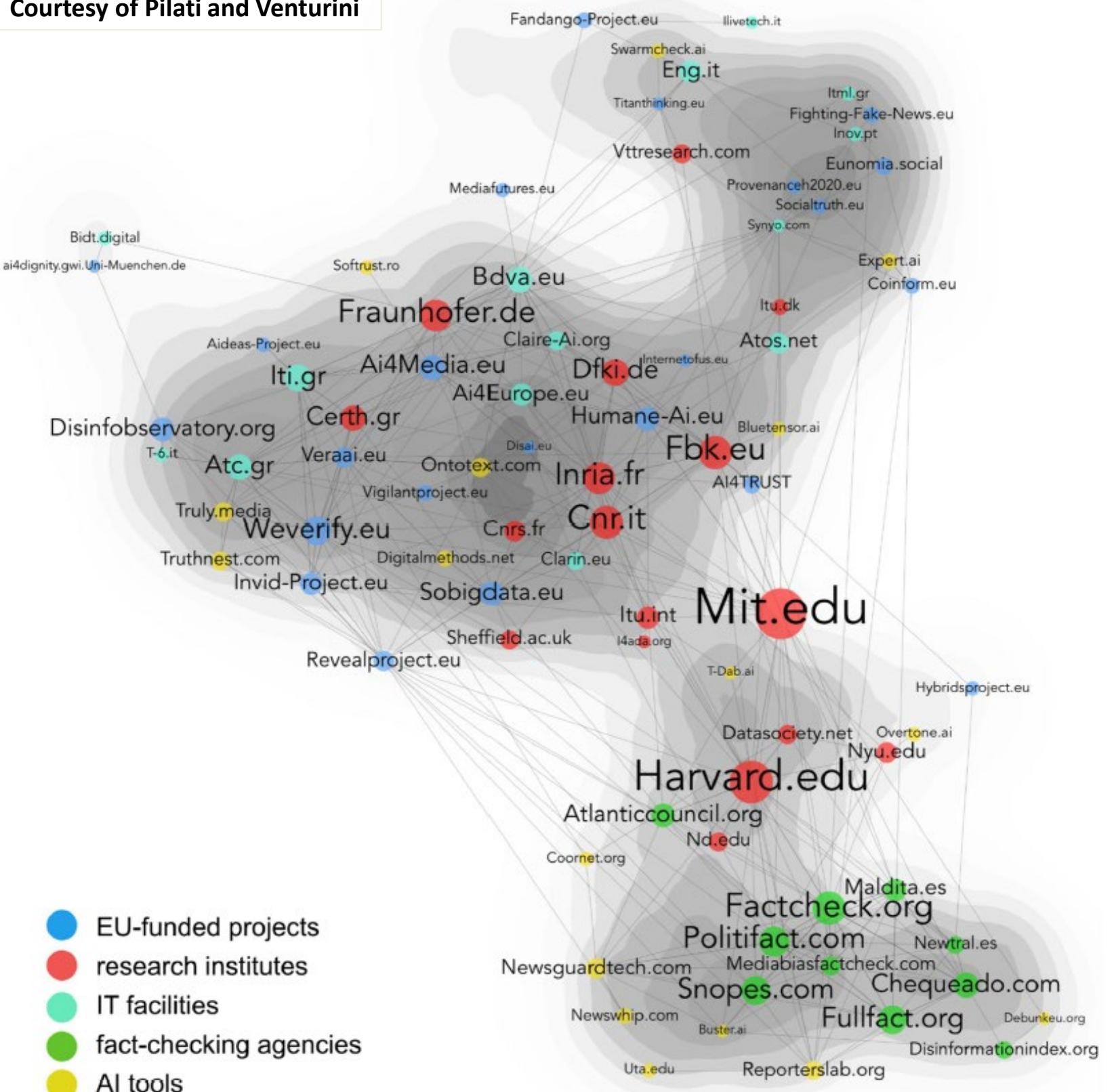


# BACKUP SLIDES



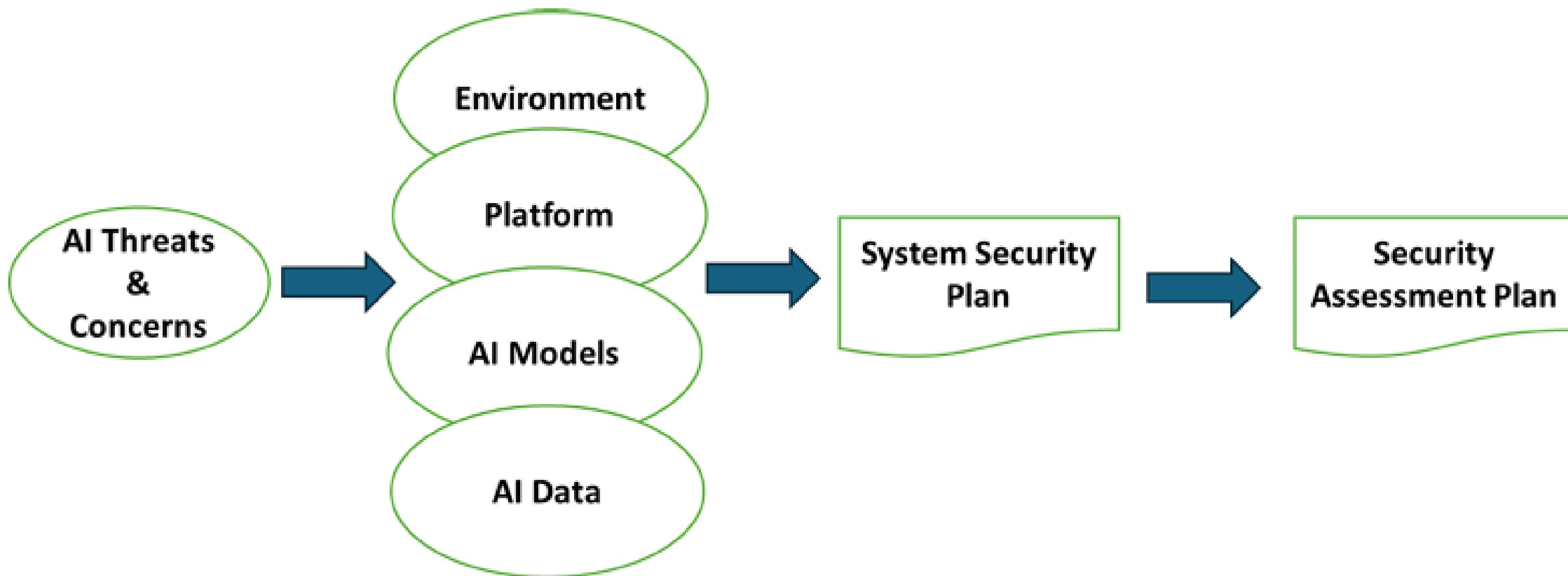
# AI Initiatives Against Disinformation

Courtesy of Pilati and Venturini



| Courtesy of Goldstein et. al. |   | Promise; if implemented...  | Limitation   |
|-------------------------------|---|---|--|
| Model Design & Construction   | AI Developers Build Models With More Detectable Outputs   | Influence operations with language models will be easily discoverable                                       | Technically challenging and requires coordination across developers  |
|                               | AI Developers Build Models That Are More Fact-Sensitive   | Language models will be less effective at spreading falsehoods  | Technical methods are still being explored; may only impact some influence operations                          |
|                               | Developers Spread Radioactive Data to Make Generative Models Detectable                         | Makes it easier to detect if content is AI generated  | Technically uncertain and may be easily circumvented   |
|                               | Governments Impose Restrictions on Training Data Collection                                     | Limits creation of new models (but only for those in jurisdictions that comply)                             | Data access restrictions would require high political will   |
|                               | Governments Impose Access Controls on AI Hardware   | Prevents some future models from being developed altogether   | Restrictions on semiconductors could escalate geopolitical tensions and hurt legitimate businesses             |
| Model Access                  | AI Providers Impose Stricter Usage Restrictions on Models                                       | Makes it more difficult for propagandists to obtain cutting-edge models for campaigns                       | Requires coordination across AI providers and risks hurting legitimate applications                            |
|                               | AI Providers Develop New Norms Around Model Release   | Restricts access to future models, but unlikely to prevent propagandists from obtaining already-public ones | Requires coordinating across AI providers and could concentrate capabilities among a small number of companies |
|                               | AI Providers Close Security Vulnerabilities   | Prevents misuse and access of models via theft and tampering  | Only affects one route to model access   |
| Content Dissemination         | Platforms and AI Providers Coordinate to Identify AI Content                                    | Increases the likelihood of detecting AI-enabled influence operations                                       | Will not affect platforms that do not engage; may not work in encrypted channels                               |
|                               | Platforms Require "Proof of Personhood" to Post   | Increases the costs of waging influence operations  | Current proof of personhood tests are often gameable by determined operators                                   |
|                               | Entities That Rely on Public Input Take Steps to Reduce Their Exposure to Misleading AI Content | Protects entities relying on public inputs from AI-enabled campaigns  | Significant changes to public comment systems could disincentivize participation                               |
|                               | Digital Provenance Standards Are Widely Adopted   | Increases detection of AI-generated content   | Significant changes would require large-scale coordination   |
| Belief Formation              | Institutions Engage In Media Literacy Campaigns   | Mitigates the impact of influence operations  | May reduce trust in legitimate content   |
|                               | Developers Provide Consumer-Focused AI Tools  | Increases the likelihood of people consuming high quality information                                       | AI tools may be susceptible to bias; users may become overly reliant on them                                   |

# NIST AI Threat Assessment



# NIST AI Risk Management Framework

AI.DEFENSE.1: Using AI to detect classify and track disinformation



| Reconnaissance   | Resource Development                  | Initial Access                                | ML Model Access                         | Execution                                     | Persistence                           | Privilege Escalation            | Defense Evasion                | Credential Access                 | Discovery                             | Collection                                     | ML Attack Staging                | Exfiltration                                | Impact                             |
|--|---------------------------------------|---|---|---|---------------------------------------|---------------------------------|--------------------------------|-----------------------------------|---------------------------------------|--|----------------------------------|---|------------------------------------|
| AML.TA0002   | AML.TA0003                            | AML.TA0004                                    | AML.TA0000                              | AML.TA0005                                    | AML.TA0006                            | AML.TA0012                      | AML.TA0007                     | AML.TA0013                        | AML.TA0008                            | AML.TA0009                                     | AML.TA0001                       | AML.TA0010                                  | AML.TA0011                         |
| AML.T0006 & Active Scanning  | AML.T0008 Acquire Infrastructure      | AML.T0015 Evade ML Model                      | AML.T0040 AI Model Inference API Access | AML.T0050 & Command and Scripting Interpreter | AML.T0018 Backdoor ML Model           | AML.T0054 LLM Jailbreak         | AML.T0015 Evade ML Model       | AML.T0055 & Unsecured Credentials | AML.T0063 Discover AI Model Outputs   | AML.T0036 & Data from Information Repositories | AML.T0018 Backdoor ML Model      | AML.T0025 Exfiltration via Cyber Means      | AML.T0034 Cost Harvesting          |
| AML.T0004 Search Application Repositories                                  | AML.T0002 Acquire Public ML Artifacts | AML.T0049 & Exploit Public-Facing Application | AML.T0044 Full ML Model Access          | AML.T0053 LLM Plugin Compromise               | AML.T0051 LLM Prompt Injection        | AML.T0053 LLM Plugin Compromise | AML.T0054 LLM Jailbreak        |                                   | AML.T0062 Discover LLM Hallucinations | AML.T0037 & Data from Local System             | AML.T0043 Craft Adversarial Data | AML.T0024 Exfiltration via ML Inference API | AML.T0029 Denial of ML Service     |
| AML.T0001 Search for Publicly Available Adversarial Vulnerability Analysis | AML.T0017 & Develop Capabilities      | AML.T0051 LLM Prompt Injection                | AML.T0047 ML-Enabled Product or Service | AML.T0011 & User Execution                    | AML.T0061 LLM Prompt Self-Replication | AML.T0051 LLM Prompt Injection  | AML.T0051 LLM Prompt Injection |                                   | AML.T0007 Discover ML Artifacts       | AML.T0035 ML Artifact Collection               | AML.T0005 Create Proxy ML Model  | AML.T0057 LLM Data Leakage                  | AML.T0059 Erode Dataset Integrity  |
| AML.T0000 Search for Victim's Publicly Available Research Materials        | AML.T0021 & Establish Accounts        | AML.T0010 ML Supply Chain Compromise          | AML.T0041 Physical Environment Access   | AML.T0020 Poison Training Data                |                                       |                                 |                                |                                   | AML.T0014 Discover ML Model Family    |  | AML.T0042 Verify Attack          | AML.T0056 LLM Meta Prompt Extraction        | AML.T0031 Erode ML Model Integrity |
| AML.T0003 Search Victim-Owned Websites                                     | AML.T0016 & Obtain Capabilities       | AML.T0052 & Phishing                          |   |   |                                       |                                 |                                |                                   | AML.T0013 Discover ML Model Ontology  |  |                                  |   | AML.T0015 Evade ML Model           |
|  | AML.T0020 Poison Training Data        | AML.T0012 & Valid Accounts                    |   |   |                                       |                                 |                                |                                   | AML.T0056 LLM Meta Prompt Extraction  |  |                                  |   | AML.T0048 External Harms           |

# Content Provenance Initiative



Linda takes a photo using her C2PA-enabled camera. It creates a digital manifest containing provenance info.



**Content Creator  
Linda**



Linda sends the photo to Erik, who can verify the credential. Erik edits the photo in PhotoShop before uploading the photo to his website.



**Editor Erik**



Looking at the manifest, a user on the website can verify that Linda took the original photo and see that Erik made edits to it.



**Website User  
Wendy**

# Data Provenance Standards

## Data Provenance Standards

The first cross-industry standards to bring transparency to the origin and use of datasets for AI and traditional data applications.