

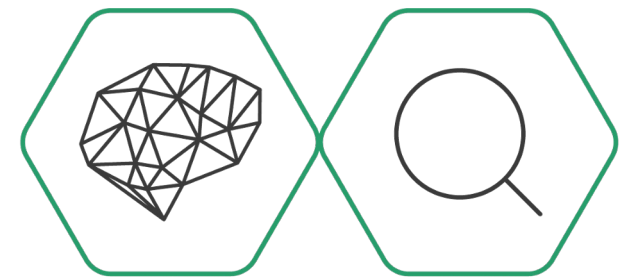
vera.ai results integrated in Truly Media

Danae Tsabouraki

ATC

17/06/2025

Final vera.ai event



vera.ai

Truly Media overview



Collect

Look into social media content via keyword searches to detect interesting topics / trends.



Organize

Create collections with relevant items, assign tasks to members in a seamless work management section for fact checking.



Verify

Select most important items of collection to verify, using checklist and a variety of verification tools.



Truly Media overview



☰

🔍

📁

🔗

🇪🇺

📰

📺

📝

🗣️

Germany Elections ⓘ

Type or paste a URL to add an item to this collection

📄 Add

📄

✎

💬

Any status ▾ Tag ▾ Any source ▾ Any assignee ▾ Any user ▾ 🔍

Sort by ☆ Date added (newest) ▾ Items 1 – 10 of 20 < > 📄 ☰


🦋 ☆ ▾

📄 R

iwillnotbesilenced.bsky.soc...
11:25AM - 4 Jun 2025 ⓘ RAW

"It's a Yes or No Question"

🌐 Translate




▶ 0:00 🔊 📺 ⋮

🎵 ☆ ▾

📄 JJ

Johanna Rüdiger - Journalist
12:26PM - 20 May 2025 ⓘ RAW

Explaining overhang mandates in the comments 🗨️ #btw21 #germany #germanelection #politik




▶ 00:00/00:00 🔊 📺 ⋮

🎵 ☆ ▾

📄 OW

Olivia Wray
12:00AM - 20 May 2025 ⓘ RAW

German Chancellor




▶ 00:00/03:53 🔊 📺 ⋮

🎵 ☆ ▾

📄 OW

Olivia Wray
11:21AM - 20 May 2025 ⓘ RAW

No website title is available



▶ 00:00/03:53 🔊 📺 ⋮

👤

COLLECTION NOTES

0

Truly Media overview



TRULY media

Dashboard Monitoring Cross-check Debunk Research Fact-checks Social wall

VERIFICATION STATUS: ☒ Raw ☐ Pending ☐ Unclear ☐ Verified ☐ Fake

Publication Information: ☐ Graphic ☐ License issue ☐ Profanity

Verification Tools

< Back to "Germany Elections"

Checklist

Lock

General Information

Claim

Verdict summary

Location

Related Links

Verification notes

Attachments

Upload attachment

Content

Source Location

VIDEO

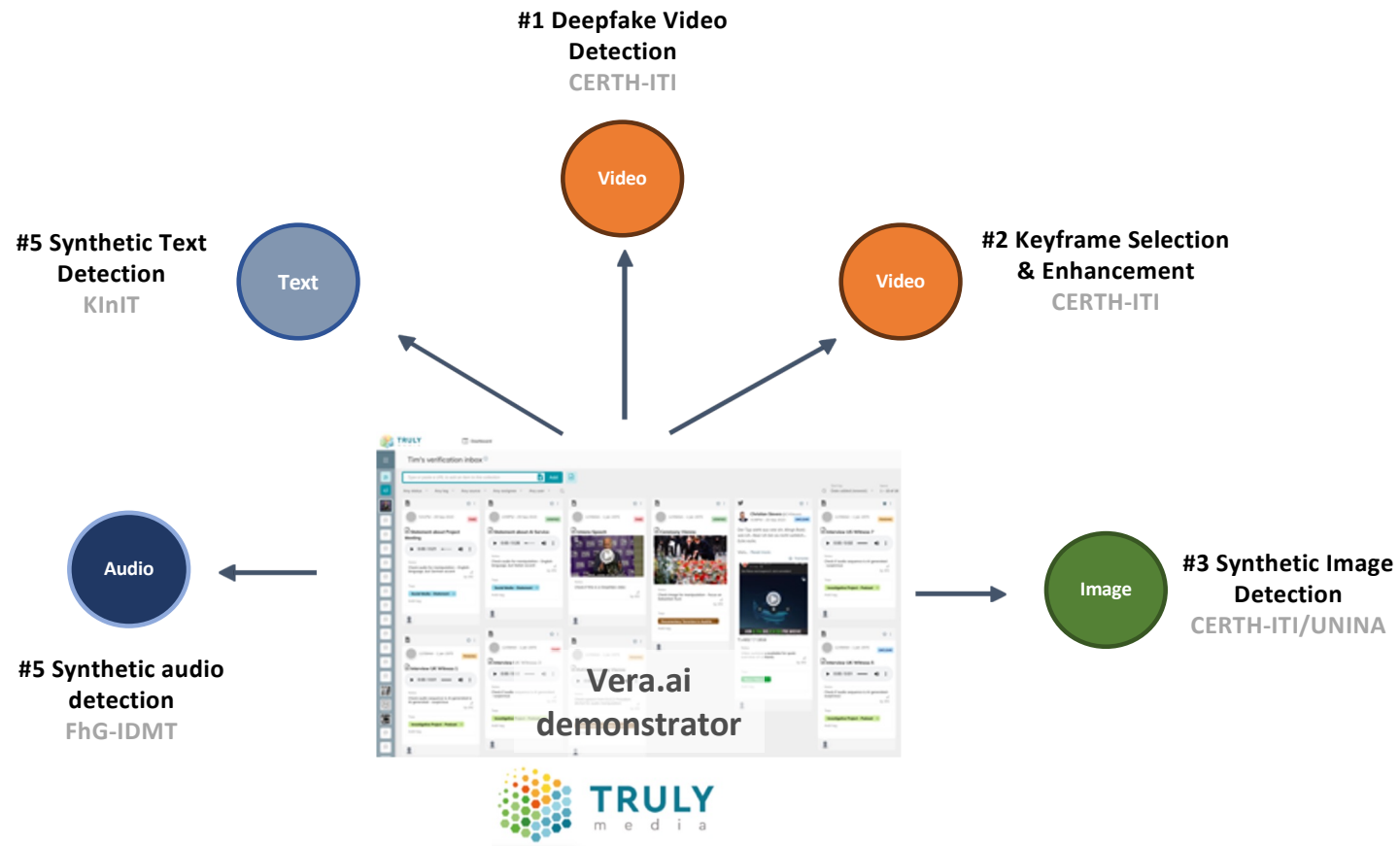
00:00/01:36

KEYFRAMES 0

ANNOTATIONS 0

ITEM NOTES 0

vera.ai integrated services



Synthetic text detection

The image is a screenshot of the Truly Media Verification Tool interface. A modal window titled "Synthetic text detection" is open in the center. The background interface shows a sidebar with various icons, a top navigation bar with "Dashboard", "Monitoring", and "Debunk" links, and a main content area with a "Verification Tool" section. The modal window has a close button (X) in the top right corner. It is divided into two main sections: "Analysis results:" on the left and "Text to analyze:" on the right. The "Analysis results:" section shows a progress bar and a "Likely machine written" status with a 94.8% confidence level. The "Text to analyze:" section contains a text area with a "Fetch website text" button and a "Cancel" button at the bottom left. The text being analyzed is a news article snippet about chemtrails. At the bottom right of the modal, there is an "Analyze" button. The background interface also shows a "Checklist" section with options like "Reverse image search results related to" and "Is this the original?".

TRULY media

Dashboard Monitoring Debunk

Verification Tool

< Back to "Synthetic test collection"

Checklist

Content

☐ Reverse image search results related to

☐ No results found

☐ Same event

☐ Other event

☐ Is this the original?

Yes No

Comments

Synthetic text detection

Analysis results:

Likely machine written ⓘ

94.8%

Text to analyze:

[Fetch website text](#)

Pilot Testifies Bill Gates Spraying Chemtrails to Incite Civil War in America

Explosive footage from a courageous whistleblower pilot has captured undeniable evidence of a chemtrails operation funded by none other than Bill Gates. This footage shows a plane dumping massive amounts of mind-altering and mood-controlling substances over New York just days after the presidential election.

At the same time, an insider from the World Economic Forum warns that the elite are ramping up this so-called "chemtrails agenda" across the U.S., with plans to manipulate public sentiment, stir unrest, and even ignite civil conflict.

But this time, their plans are unraveling, as whistleblowers step forward with one mission: to expose and dismantle the operation.

BYPASS THE CENSORS Sign up to get unfiltered news delivered straight to your inbox. Email *

Email This field is for validation purposes and should be left unchanged.

The tides are turning. People around the world are waking up—and we've got allies in high places. One thing's certain: the elites are desperate, and they will do anything to avoid facing justice for their crimes.

Cancel Analyze

Synthetic video detection



Deepfake Analysis

DeepFake Video Detection

This video tool aims to detect whether the faces that are present in this media have been either manipulated or newly generated by the use of AI-powered Deep Learning methods. Faces shown with intense red color indicate higher manipulation probability.

AI-Generated Face Detection^①

Analyzed video contains faces that are **FAKE**

Overall probability: 99%


Analysis duration: 9.26s

Analyzed shots

Shot #1 is **FAKE**

Shot duration: 0 - 28s

1 extracted face(s)




Face #1 is **FAKE**

Probability: 98%

[View this face in different shots](#)

AI Face Reenactment Detection^①



Powered by CERN-ITI and UNINA in

User Guide

This document aims to increase the transparency and understandability of the Deepfake Detection Service for end-users by offering insights into how the AI-driven results are delivered, which are the limitations of the service, and how users should interpret the generated results. It provides answers to the questions listed below that users may have, especially when using the service for the first time. The information has been provided by ATC, the operator of the Truly Media platform, in cooperation with DW, a media organisation that co-developed Truly Media and uses the platform.

1. How should I use the service?
2. How should I interpret the results of the service?
3. Which are the limitations of the service?
4. How can I ensure that the analysis is optimal?
5. Why does the service generate false positive/negative results?
6. How much time does the service need to analyse an image/ video?
7. How accurate are the service's results?

How should I use the service?

The Deepfake Detection Service should be used to support your work and the analysis results should be taken into consideration in combination with other factors and indicators (e.g. credibility of the source and other contextual information), to come to a decision about the veracity of the analysed image/video. It is not the aim of the service to provide a yes or no answer to the question of whether a face in an image or video is deep-faked, rather to provide an indication of the probability of such face manipulations. Your own judgement should always be used to come to a final decision.

How should I interpret the results of the service?

Keyframe Selection and Enhancement

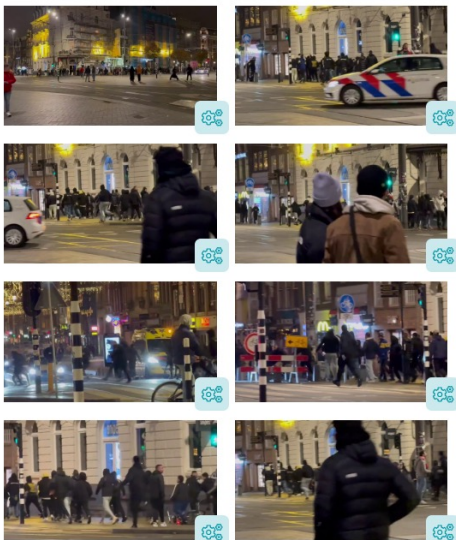


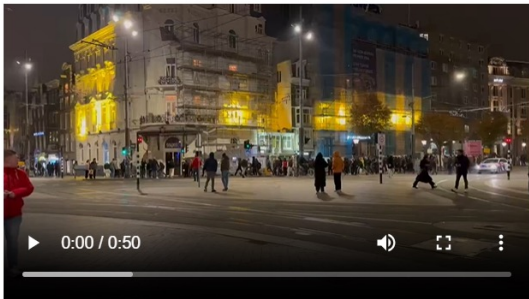
Keyframe Selection and Enhancement

Keyframe selection and enhancement

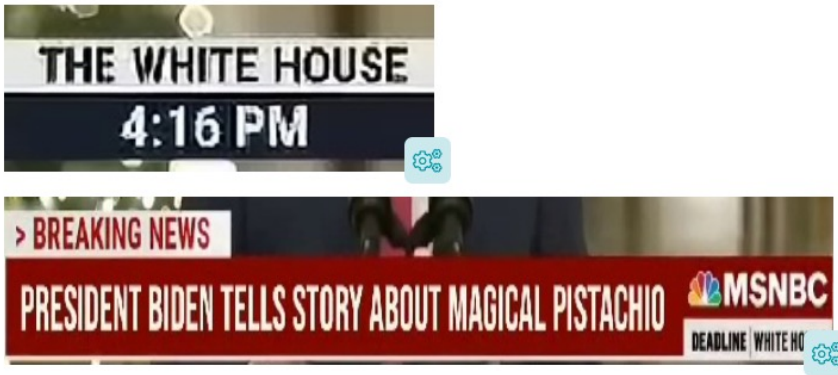
This experimental AI-powered service automatically extracts representative keyframes from a video. The service also identifies and improves the resolution of faces and text detected within the keyframes for better clarity.

Selected keyframes

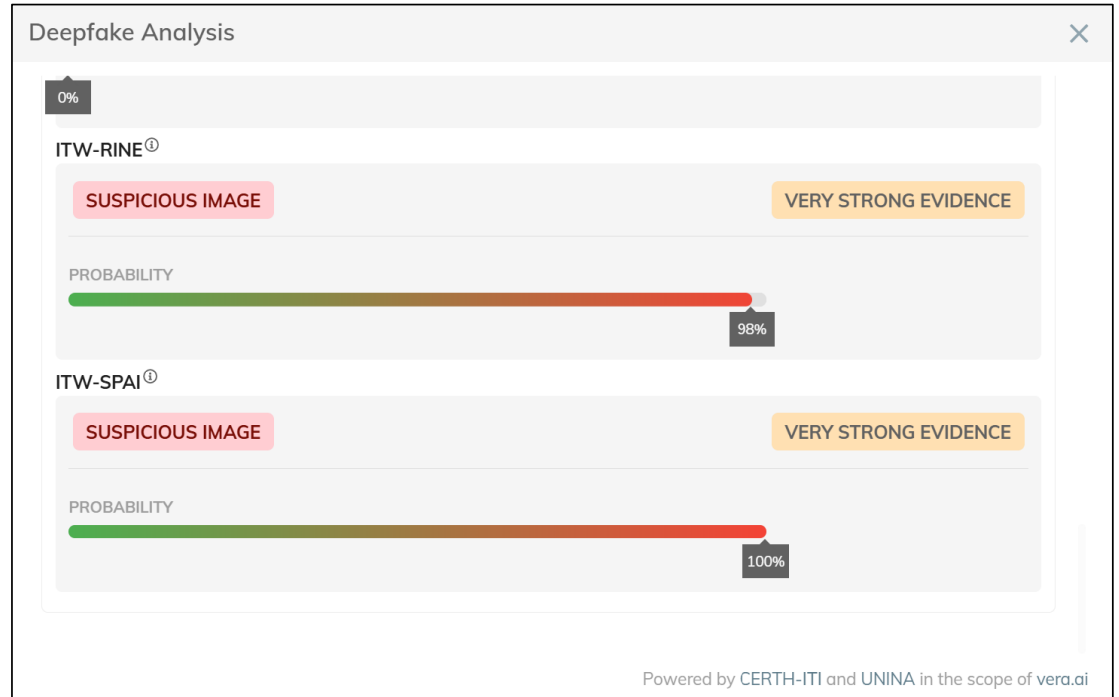




Detected text regions



Synthetic image detection



Synthetic audio detection



Detection Output Interpretation

What are LLRs?

The model produced one log-likelihood ratio (LLR) for each 2 seconds of input audio, represented by the colored squares in the waveform. These describe **if the detected traces are more likely to appear in synthetic speech, or in natural speech.**

How can I interpret the LLR values, and the colors provided in the waveform?

In our model, **positive LLRs** support the hypothesis that the speech in the audio **is likely to be created artificially.**

- If $0 < \text{LLR} \leq 1$, the findings provide weak support for the speech in the audio to be created artificially,
- If $1 < \text{LLR} \leq 2$, the findings provide moderate support for the speech in the audio to be created artificially,
- If $2 < \text{LLR} \leq 3$, the findings provide moderately strong support for the speech in the audio to be created artificially,
- If $3 < \text{LLR} \leq 4$, the findings provide strong support for the speech in the audio to be created artificially,
- If $\text{LLR} > 4$, the findings provide very strong support for the speech in the audio to be created artificially,

Conversely, **negative LLRs** support the hypothesis that the speech in the audio **is likely to be natural.**

- If $-1 \leq \text{LLR} < 0$, the findings provide weak support for the speech in the audio to be natural,
- If $-2 \leq \text{LLR} < -1$, the findings provide moderate support for the speech in the audio to be natural,
- If $-3 \leq \text{LLR} < -2$, the findings provide moderately strong support for the speech in the audio to be natural,
- If $-4 \leq \text{LLR} < -3$, the findings provide strong support for the speech in the audio to be natural,
- If $\text{LLR} < -4$, the findings provide very strong support for the speech in the audio to be natural,

Finally, if $\text{LLR} = 0$ the analysis is **inconclusive**, and the findings do not support any hypothesis.

Close

Insights from user evaluation sessions

Trust, Transparency, and Explainability are crucial

- Users repeatedly emphasized the importance of understanding *why* a result was produced - whether for synthetic text detection or keyframe selection.
- Lack of explanations or traceability of results leads to distrust, even if the outputs are technically correct.

Ease of use and clarity matter

- Tools should work *out-of-the-box* for journalists and media professionals.
- Complex terms, poor naming, hidden character limits, or vague UI cues significantly reduce user confidence and usability.

Users expect contextualized results

- Output must be actionable and integrated with relevant information: timestamps, visual cues, contextual annotations, or model uncertainty.

Lessons learned

Explainability drives trust

- Simply showing a percentage or a tag ("AI-generated") is insufficient. Users want explanations like: "This sentence is likely synthetic because it contains repetitive phrasing or lacks temporal grounding."
- Trust was lost when results contradicted user expectations.

Need for granularity and control

- Users want to control which parts of the data are analyzed or enhanced to save time and processing power.

Partial implementation undermines value

- Incomplete workflows lead to confusion and underuse of otherwise useful tools.

Internal testing adds value before broad user rollout

- Internal tests focusing on UI/UX and trustworthiness helped anticipate external feedback and fine-tune critical features.

Danae Tsabouraki/ ATC

Contact: d.tsabouraki@atc.gr



CERTH
CENTRE FOR
RESEARCH & TECHNOLOGY
HELLAS



ATHENS TECHNOLOGY CENTER

kinIT



**Deutsche
Welle**



UNIVERSITÀ DEGLI STUDI
DI NAPOLI FEDERICO II



1506
UNIVERSITÀ
DEGLI STUDI
DI URBINO
CARLO BO



UNIVERSITEIT
VAN AMSTERDAM

école
normale
supérieure
paris—saclay

université
PARIS-SACLAY

EBU

OPERATING EUROVISION AND EURORADIO



Fraunhofer
IDMT

Follow us on Twitter: [@veraai_eu](https://twitter.com/veraai_eu)

Website: <https://www.veraai.eu/>

Co-financed by the European Union, Horizon Europe programme,
Grant Agreement No 101070093.

Additional funding from Innovate UK grant No 10039055 and the
Swiss State Secretariat for Education, Research and Innovation (SERI)
under contract No 22.00245

