

Advances in technological developments, Al, disinformation, and verification

Jochen Spangenberg

Deutsche Welle News Impact Summit Warsaw, 9 Oct 2025

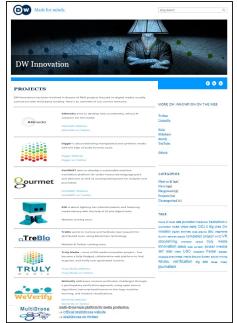


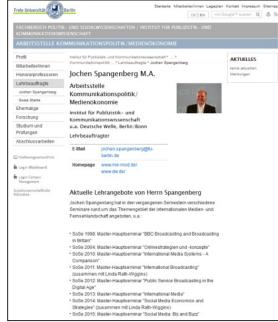






About me

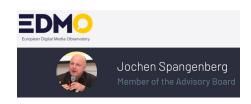












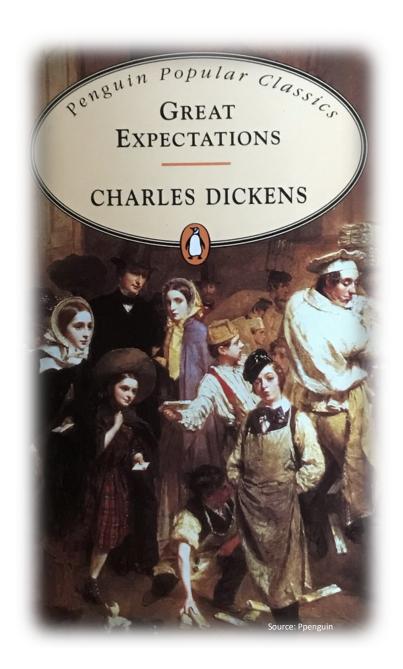




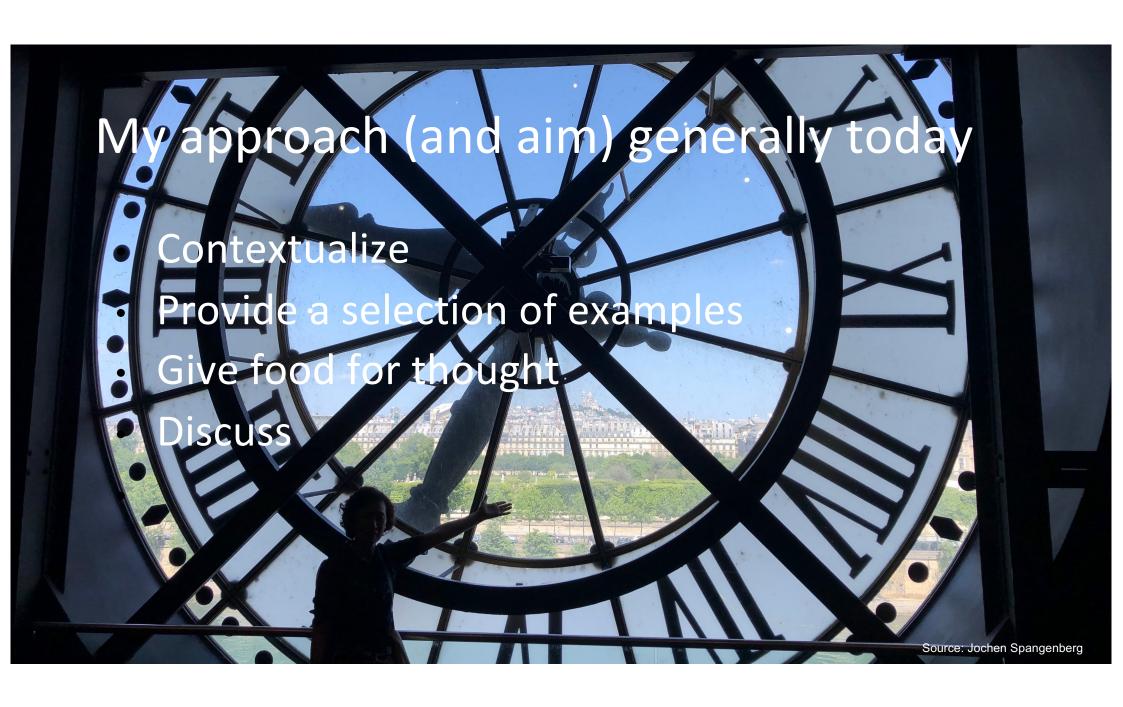


Expectation management ("what's in a title...?")

- Technological developments / advances ...
- Disinformation ...
- Verification ...
- Al ...

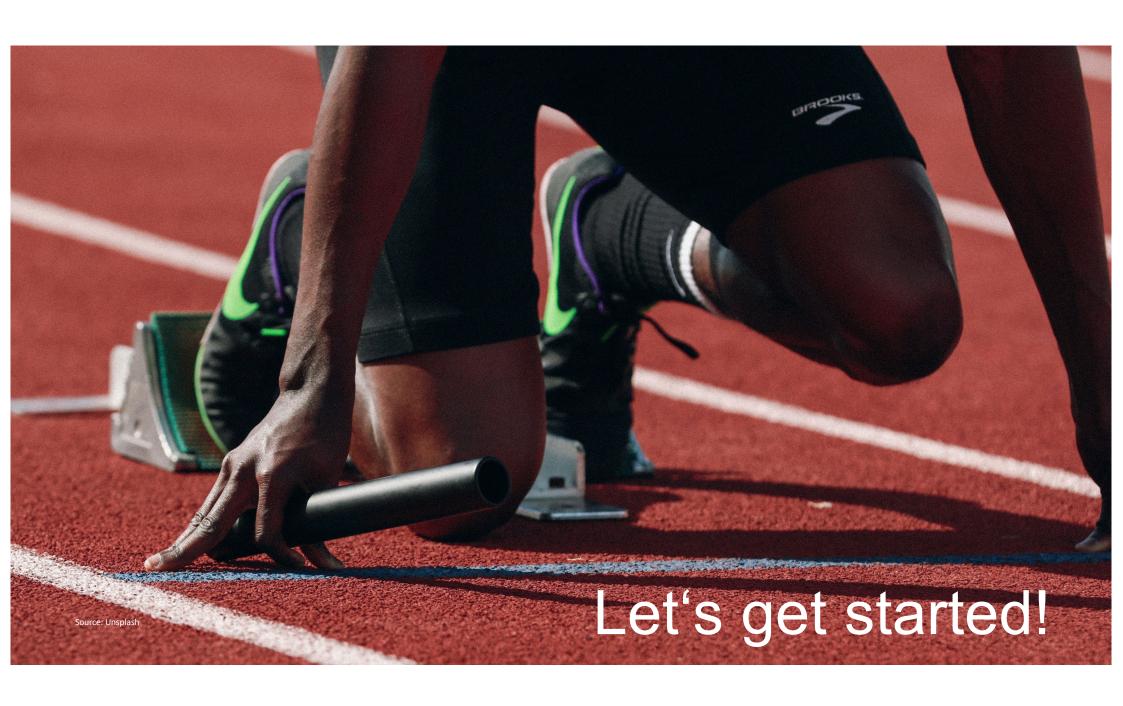




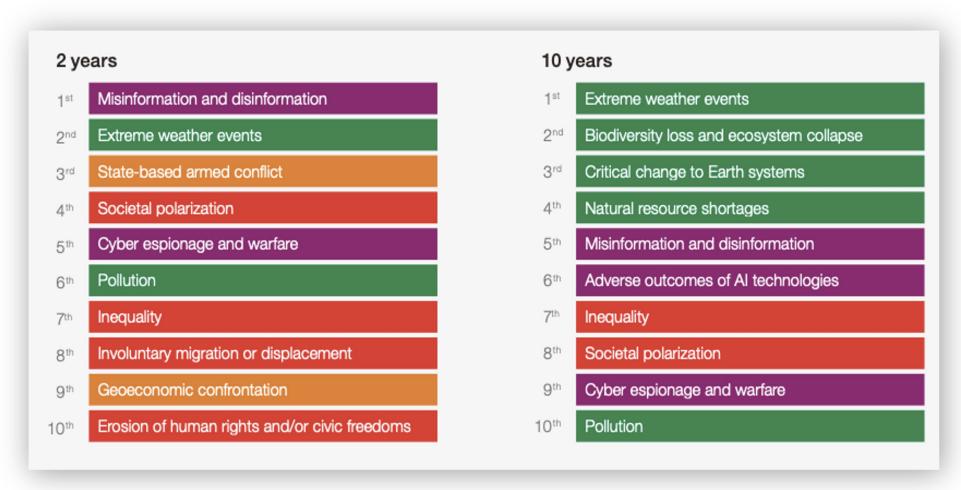


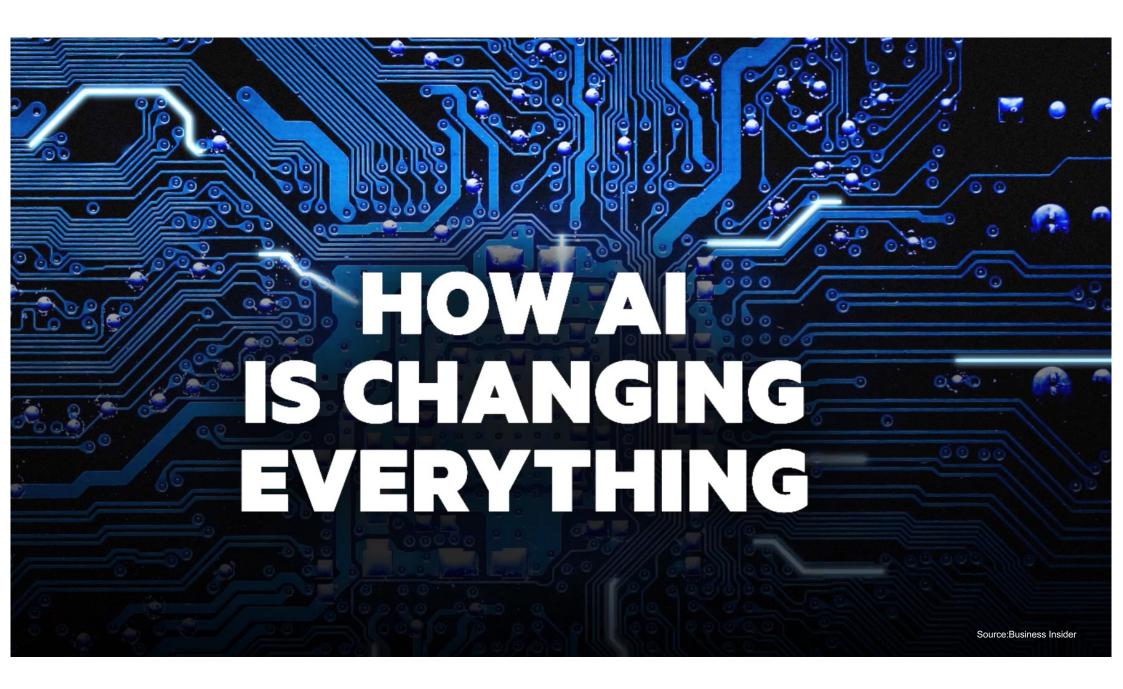
In particular, I will try and

- take a glimpse at the situation overall
- look at a selection of developments regarding the spread of disinformation (facilitated by tech / AI)
- provide examples
- put this into a journalistic context
- show selected works and outcomes from a research project (vera.ai) that support in verification and countering disinformation
- raise issues for debate
- do all that without being a tech expert

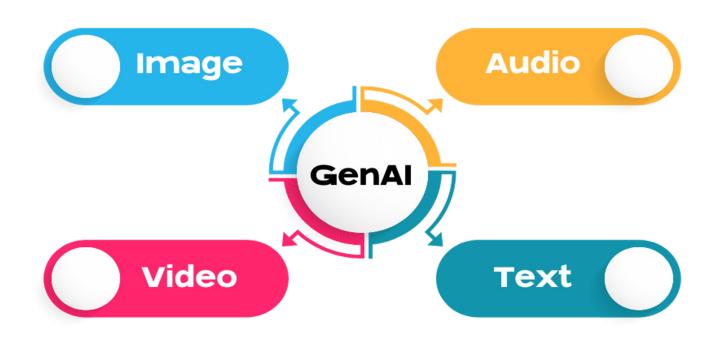


Pressing concerns





What we're seeing



Digital content as a news and information (re)source. OR The value of digital media (UGC) for news (reporting)









Brussels, Belgium 21 Mar 2016

Source: Screenshot of video by Pinchas Kopferstein.
On YouTube <u>here</u>.

Kabul, Afghanistan 16 Aug 2021

Source: Screenshot from video uploaded by an anknown source

Poltava, Ukraine, 1 Feb 2025

Source: Ukrainian emergency services, via Telegram

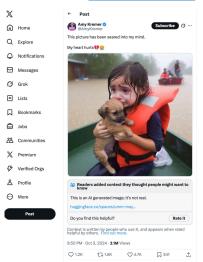
Nasser Hospital, Gaza City, 25 Aug 2025

Source: unknown. Verified by BBC Verify.
On YouTube <u>here</u>

Tech getting more and more "sophisticated"



Source: automatically created with Midjourney v5, by Eliot Higgins (shared via Tweet of 21 March 2023)



Source: Tweet by US Republican Amy Kremer, 3 Oct 2024



Source: Comofoto / Adobe Stock (2024). In Adobe Stocks library, the picture contains the note – in small writing – that the image was Al-generated



Source: Al-generated video, shared by Donald Trump on Truth Social, depicting a "reimagined Gaza". February 2025. On YouTube, with comments by the Guardian.



Source: Video created using Open Al's Sora Al technology Source / taken from: Tweet of Ben Nash, taken 17 Feb 2024

Text to video – Sora AI by Open AI

Prompt:

A white and orange tabby cat is seen happily darting through a dense garden, as if chasing something. Its eyes are wide and happy as it jogs forward, scanning the branches, flowers, and leaves as it walks. The path is narrow as it makes its way between all the plants. the scene is captured from a ground-level angle, following the cat closely, giving a low and intimate perspective. The image is cinematic with warm tones and a grainy texture. The scattered daylight between the leaves and plants above creates a warm contrast, accentuating the cat's orange fur. The shot is clear and sharp, with a shallow depth of field.



It's getting serious



A wave of disinformation has been unleashed online since Israel began strikes on Iran last week, with dozens of posts reviewed by BBC Verify seeking to amplify the effectiveness of Tehran's response.

"...the first time we've seen generative AI be used at scale during a conflict," Emmanuelle Saliba, Chief Investigative Officer with analyst group Get Real

THE LEVEL OF AI-GENERATED DISINFORMATION **STAYS HIGH** The percentage of disinformation stories using Al-generated content remained stable in August, along with the previous months' record. Out of 1.441 fact-checking articles, 147 addressed the use of this technology in disinformation, representing 10% of the total. 10

Source: BBC, 21 June 2025

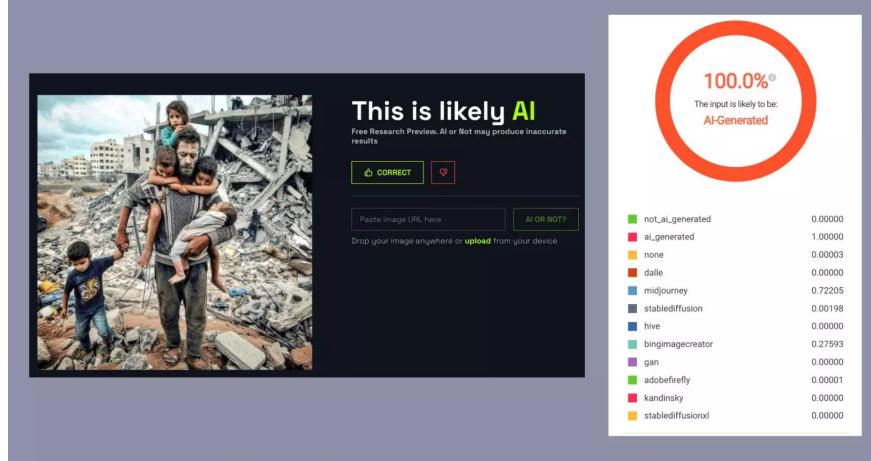
Source: EDMO Fact-checking brief No 51 of 15 September 2025

No longer like this...



Source: Tweet by HumaZhr of 23 Oct 2023, screenshot taken on 6 Feb 2024.





Al Or Not

Hive Moderator

Source: **Boom Live** / Hanid Farid

Next level



Source: AlwitnessTV (Sep 2025)

On and on it goes ...

OpenAl

← Home

Research Index

Research Overview

Research Residency

September 30, 2025 Research Release Product

Sora 2 is here

Our latest video generation model is more physically accurate, realistic, and more controllable than prior systems. It also features synchronized dialogue and sound effects. Create with it in the new Sora app.

wnload the Sora app 🗷



Sora 2

o4-mini

GPT-4.5

OpenAl of



Source: OpenAl announcement, 30 Sep 2025



Prompt: a gymnast flips on a balance beam. cinematic

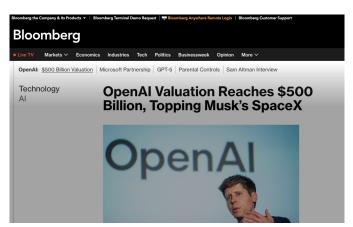


Prompt: man does a cannonball off diving board

TikTok @drewharwell OpenAl says its new Al video tool Sora

According to Drew Harwell, 'OpenAI's "guardrails" against bad content allow way more than you'd expect.'

By the way ...



Source: Bloomberg (2 Oct 2025)

Source: Drew Harwell on TikTok (3 Oct 2025)

However (or: consider as well)



Source: Pest by @icksonhinklle on Twitter, Screenshot taken on 6 Feb 2024





hassan.ghaedi.photo · Folgen



hassan.ghaedi.photo Life in war The #mother, who returned home after the end of the conflict and and has found her child' toys. #Homs #Syria. On June 2016

پس از اتمام درگیری ها در شهر خمص، مادری به خانه اش بازگشته و اسباب بازی فرزندش را میان آوار بیدا کرده است

.مادر ، مادر است. هر جای دنیا که باشد#

#gettyreportage #gettyimages #war #syria #wpagency #noorimages #viiphoto #agencevu #zumapress #sputnik #epaphotos #warphotography #dailylife #streetphotography

Bearbeitet · 207 Wo.





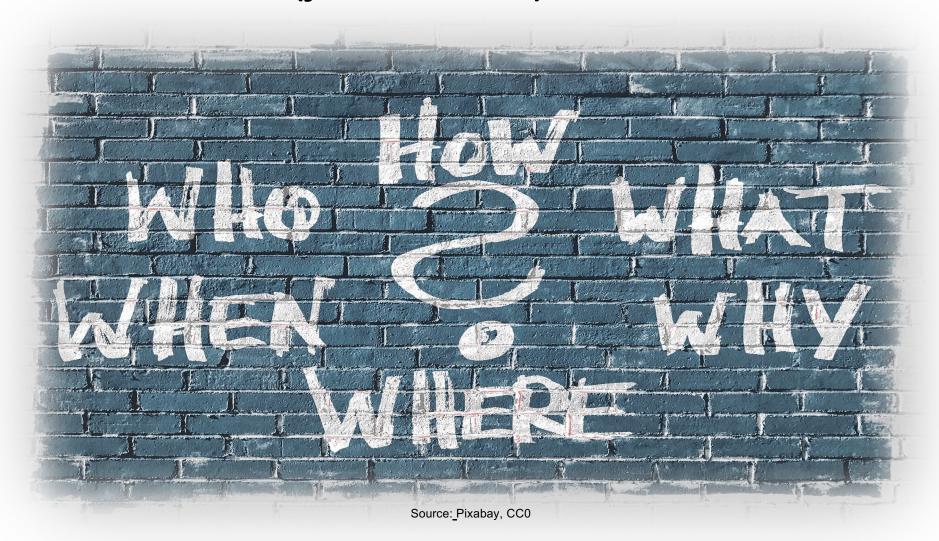


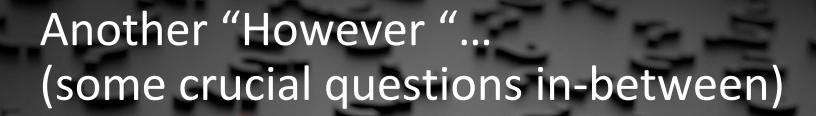
Gefällt 1.095 Mal

14. Februar 2020

Homs, Syria in June 2016, NOT Gaza Nov 2023

Back to some (journalistic) basics





- How can journalism keep the pace?
- How to keep up with technological developments?
 - for creation
 - for detection
- What does this mean for society (and democracy)?

The New York Review

Sign In

Hannah Arendt: From an Interview

Hannah Arendt

October 26, 1978 issue

Hannah Arendt made the comments that follow in 1974 during an interview with the French writer Roger Errera.

Totalitarianism

Source: The New York Review

Lies

The moment we no longer have a free press, anything can happen. What makes it possible for a totalitarian or any other dictatorship to rule is that people are not informed; how can you have an opinion if you are not informed? If everybody always lies to you, the consequence is not that you believe the lies, but rather that nobody believes anything any longer. This is because lies, by their very nature, have to be changed, and a lying government has constantly to rewrite its own history. On the receiving end you get not only one lie—a lie which you could go on for the rest of your days—but you get a great number of lies, depending on how the political wind blows. And a people that no longer can believe anything cannot make up its mind. It is deprived not only of its capacity to act but also of its capacity to think and to judge. And with such a people you can then do what you please.

"If everybody always lies to you, the consequence is not that you believe the lies, but rather that nobody believes anything any longer."

This is what we are up against

RICKASTLEY

NEVER GONNA GIVE YOU UP

Released in 1987

4 K





The vera.ai project team



14 partners:

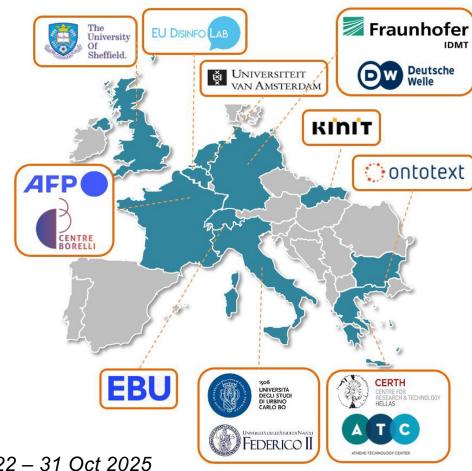
8 research organisations (universities, research centres)

2 commercial companies with relevant products

2 global news and information providers

1 NGO focussed on tackling disinformation strategically at EU-level

The world's biggest association of public service media organisations





Project runtime: 15 Sep 2022 – 31 Oct 2025



vera.ai project consortium and goal



vera.ai is a Horizon Europe project that seeks to build professional trustworthy AI solutions against advanced disinformation techniques, co-created with and for media professionals & researchers and to also set the foundation for future research in the area of AI against disinformation.





vera.ai outcomes



- Methods (detection, analysis/enhancement, retrieval)
- Datasets (research enablers)
- Studies (real-world campaign mapping & analysis)
- Evaluation (involving hundreds of stakeholders/end users)
- Dissemination & Outreach (wide/relevant audience)
- Development / integration of tools (supporting media professionals, investigators, ...)

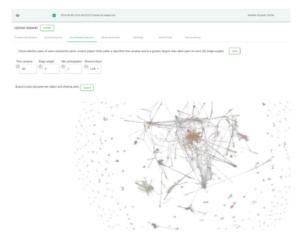
Selection of individual tools / components



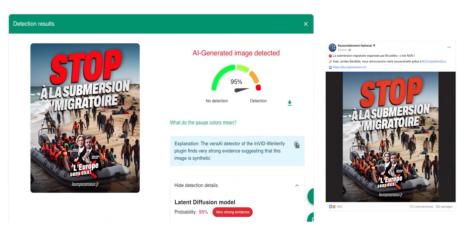
Classifier flagging persuasion techniques



Social Network Analysis (tracking Coordinated Inauthentic Behaviour)

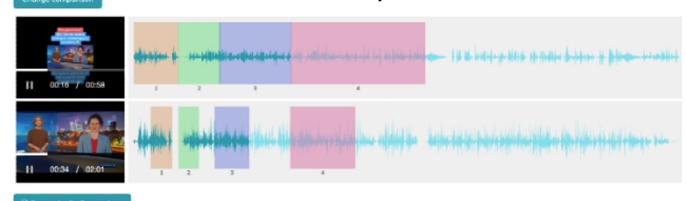


Forensic trace analysis



https://www.digitalmethods.net/Dmi/Al-generated-Rassemblement

Audio analysis



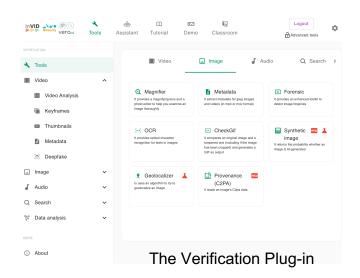
Development, provision and integration of selected outcomes / components

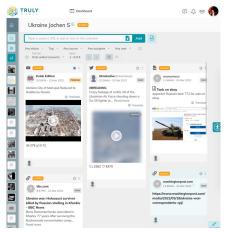


Selection of vera.ai tools integrated

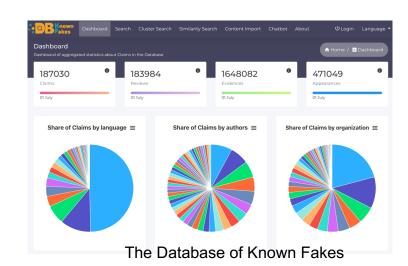






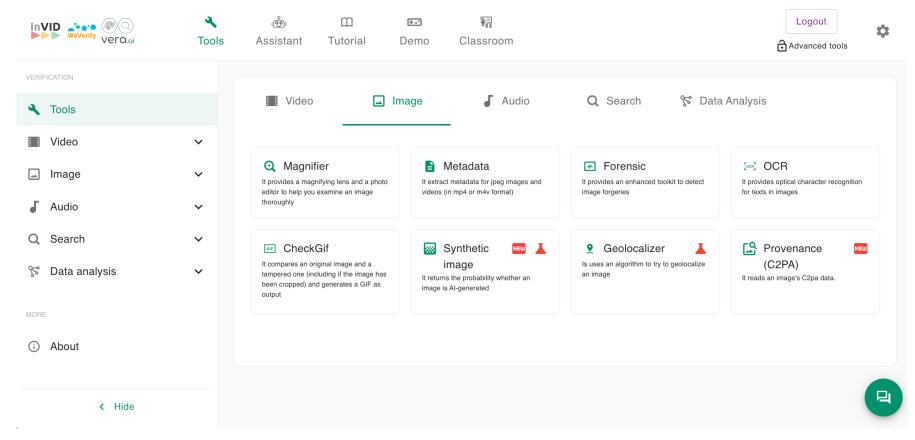






Next: focus on one platform, the verification / plug-in

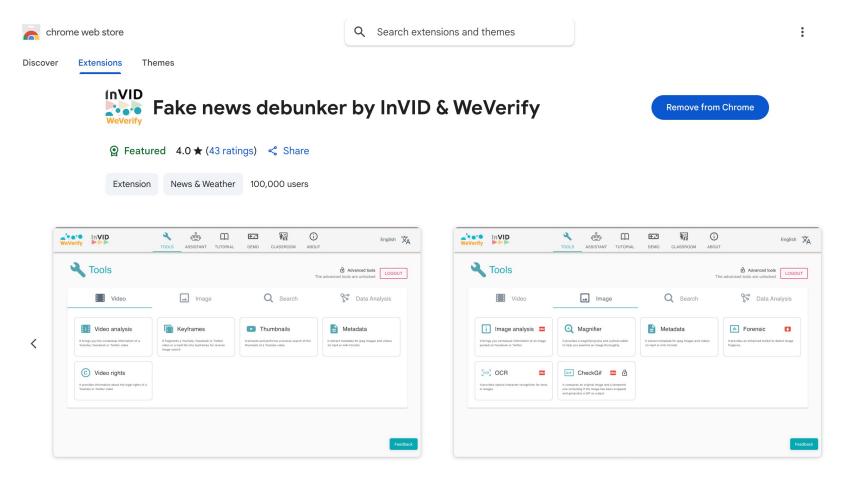




- Coordinated by AFP
- 170,000 active weekly users
- Available in 10 languages in September 2025:
 English, French, Spanish, Greek, Italian, Arabic, German, Japanese, Hungarian and Portuguese

Getting started – getting the verification / plug-in





Source: Chrome store

Getting started – having the verification/plug-in

















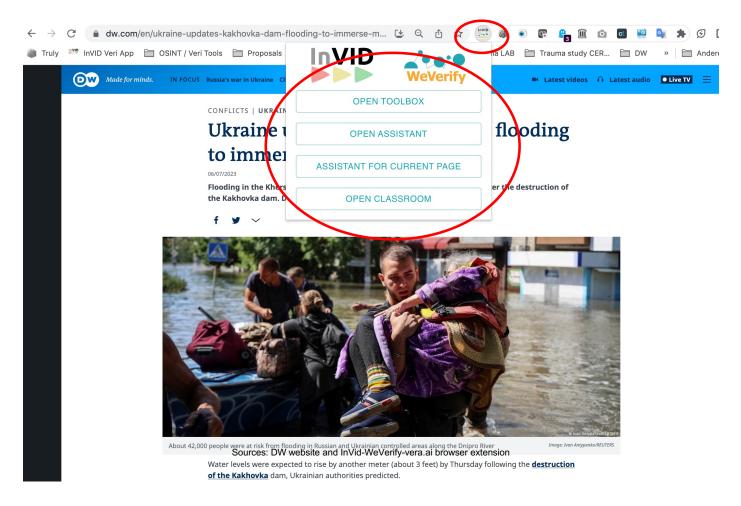






Getting started – one way

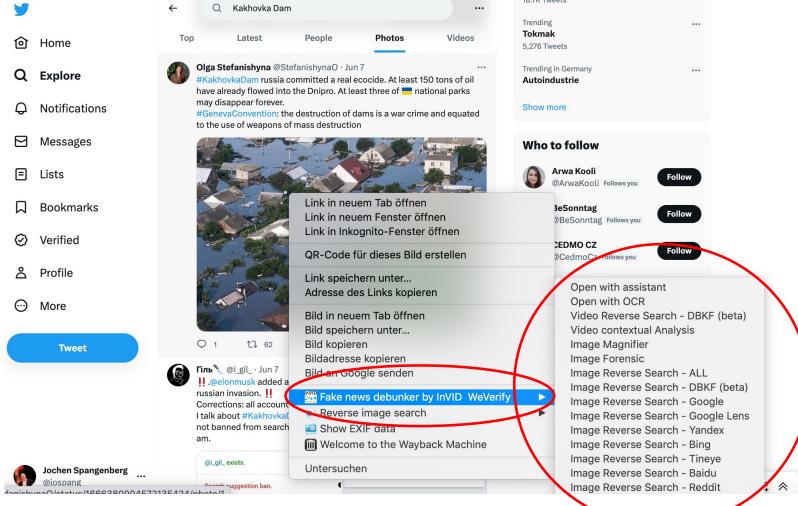




Sources: screenshot Chrome browser bar

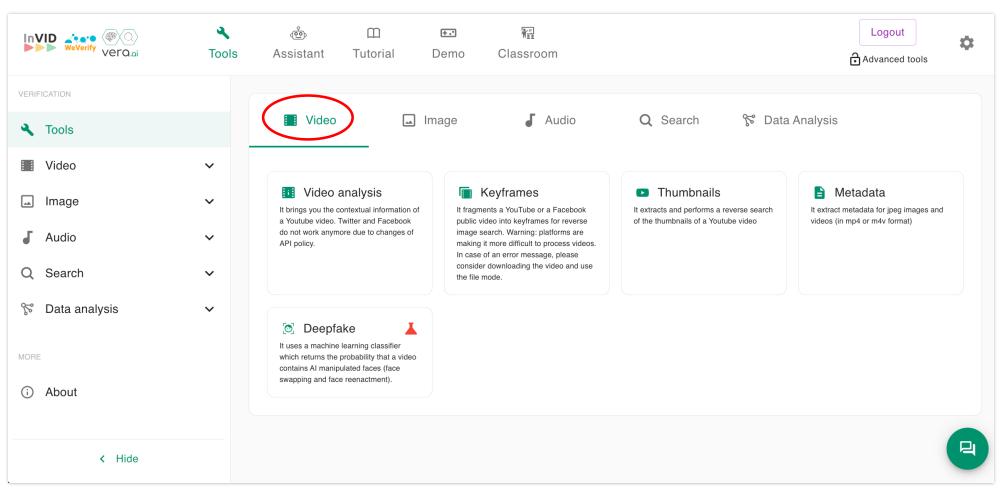
Getting started – another way



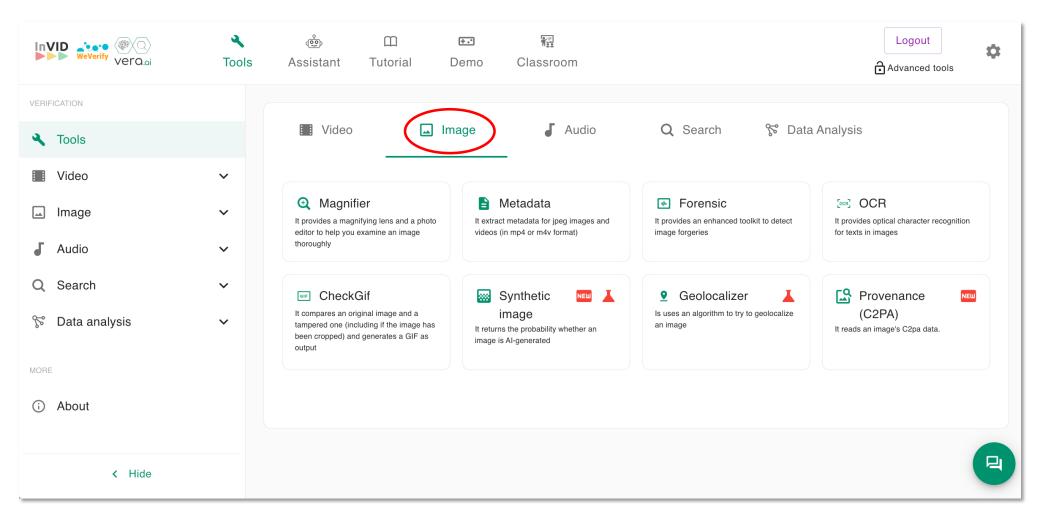


Sources: Twitter and InVid-WeVerify-vera.ai browser extension

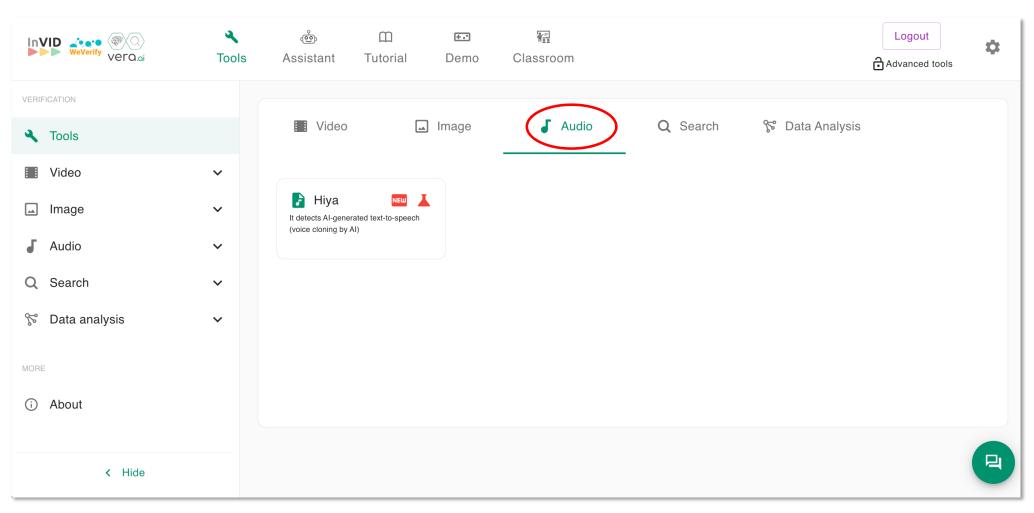




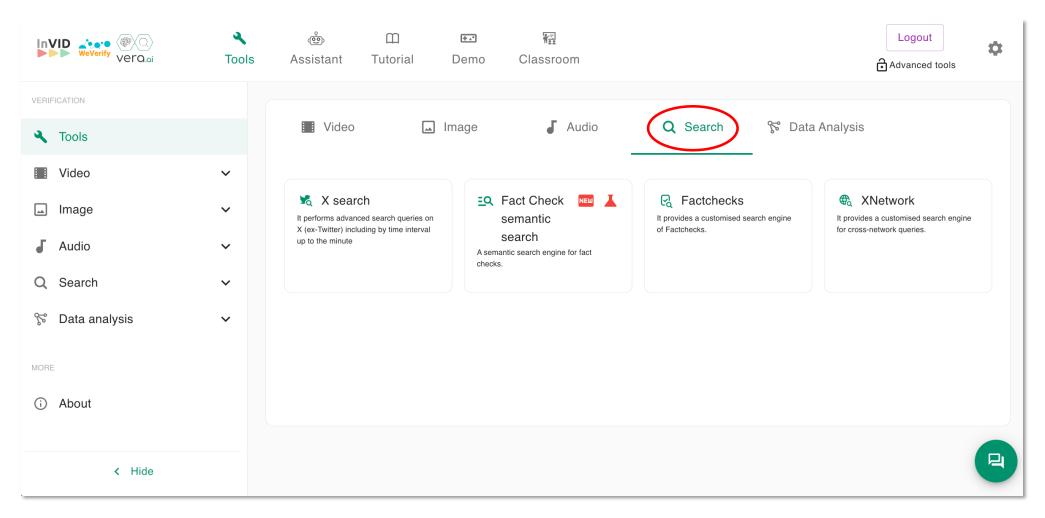




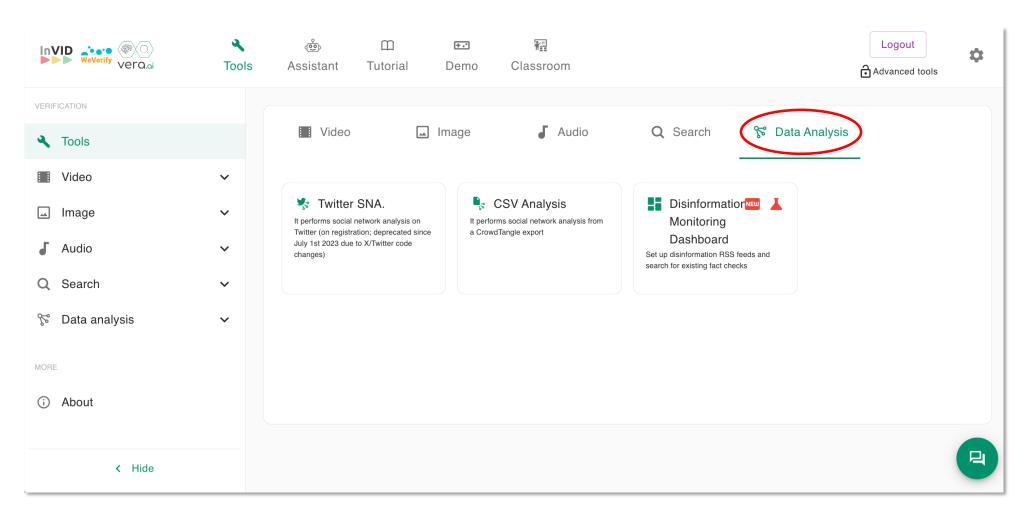






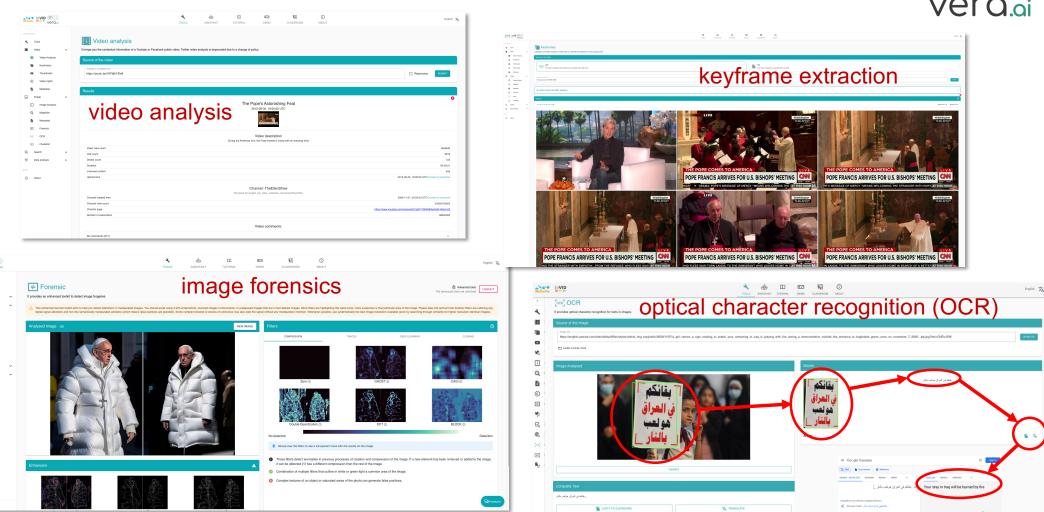






There's A LOT in it (here: 4 out of 24)

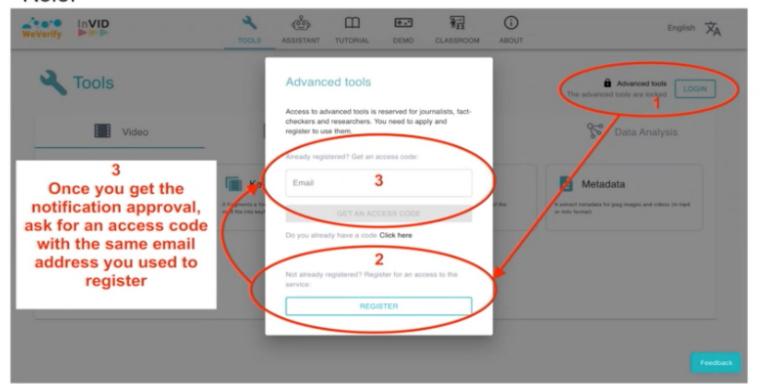




Advanced tools for journalists



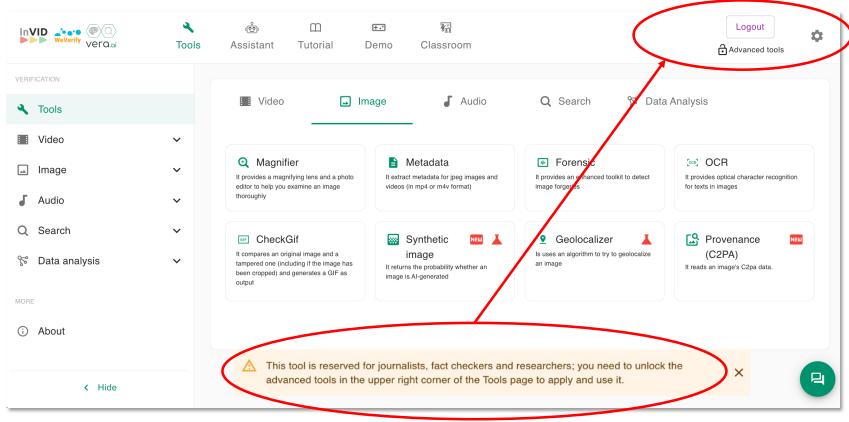
You need to **register for the advanced tools** (if you do not have them yet). If you're a new registrant, please indicate "**betatester**" as a Role.



Registration
Why register? Due to the high computer load, the use of this feature is restricted to journalists, fact-checked red researchers. You need to register first and once our registration has been approved, you will be able to request the access code to use the advanced tool we DO NOT store personal data and will NOT use our email for any other purpose. In case of problems the feedback button.
Email*
professional email address
A valid fact-checking, media or research organization
email address is required
First name*
Enter your first name
Last name*
Organization*
Role*
Role within organization is required
Role (Other)
betatester

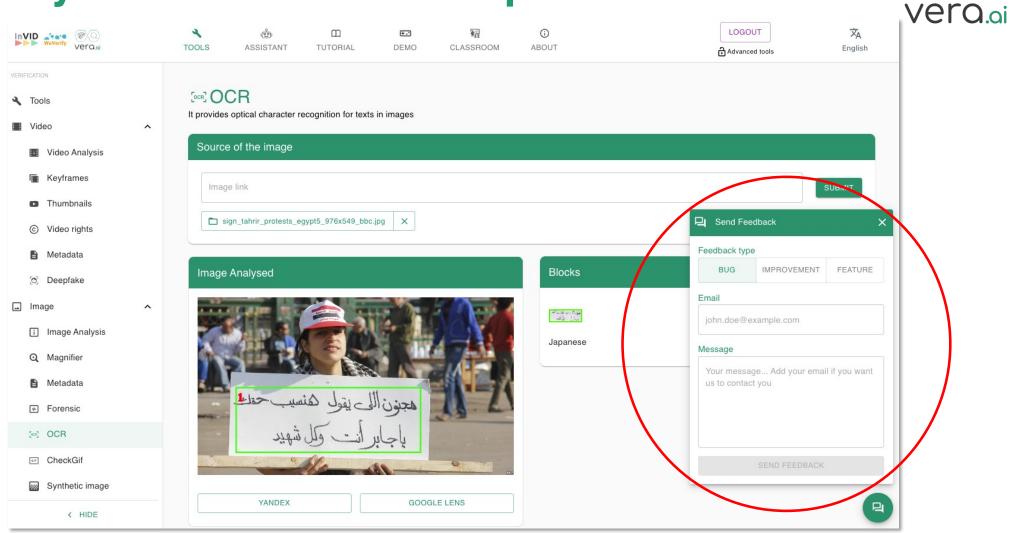
Advanced tools for journalists





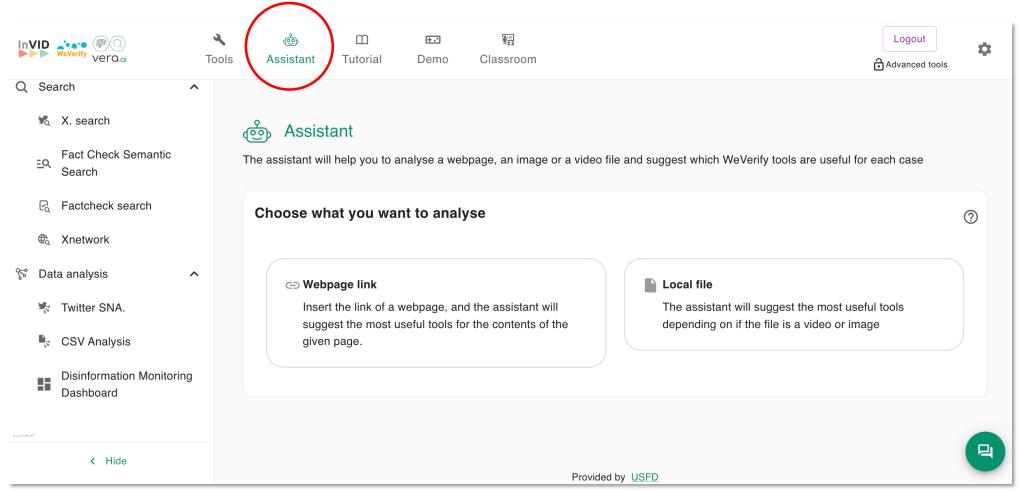
Advanced tools are reserved for factcheckers, journalists and researchers

If you use it: feedback please! ©



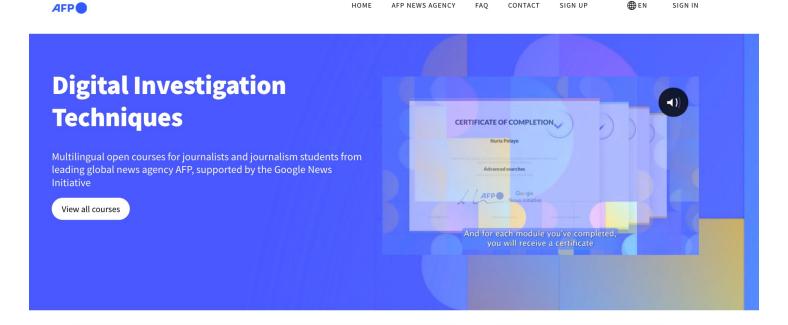
Getting to know features & functionalities (1)





Getting to know features & functionalities (2)





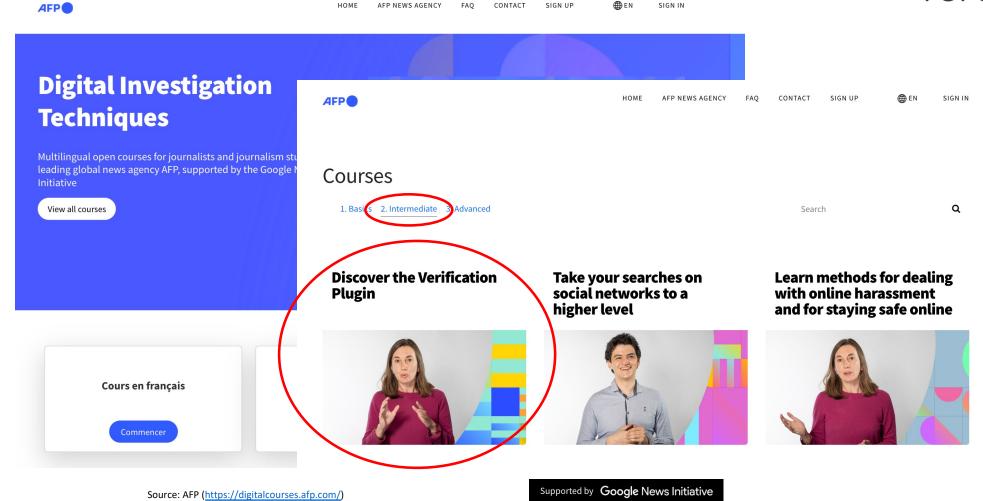


Source: AFP (https://digitalcourses.afp.com/)

Supported by Google News Initiative

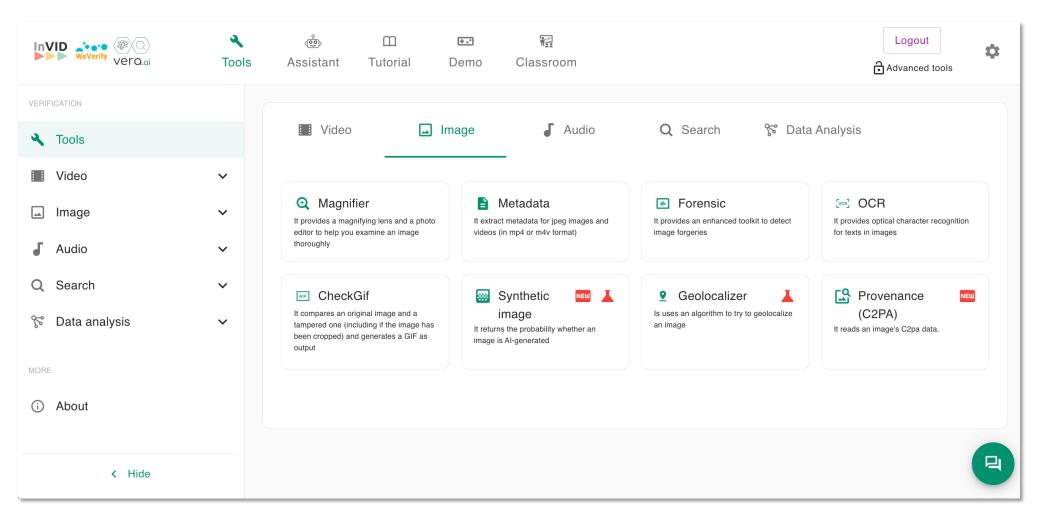
Getting to know features & functionalities (3)





Getting to know features & functionalities (4)





For more:

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

BlueSky: @vera-ai.bsky.social

X: @veraai_eu

YouTube: https://www.youtube.com/@veraai eu





























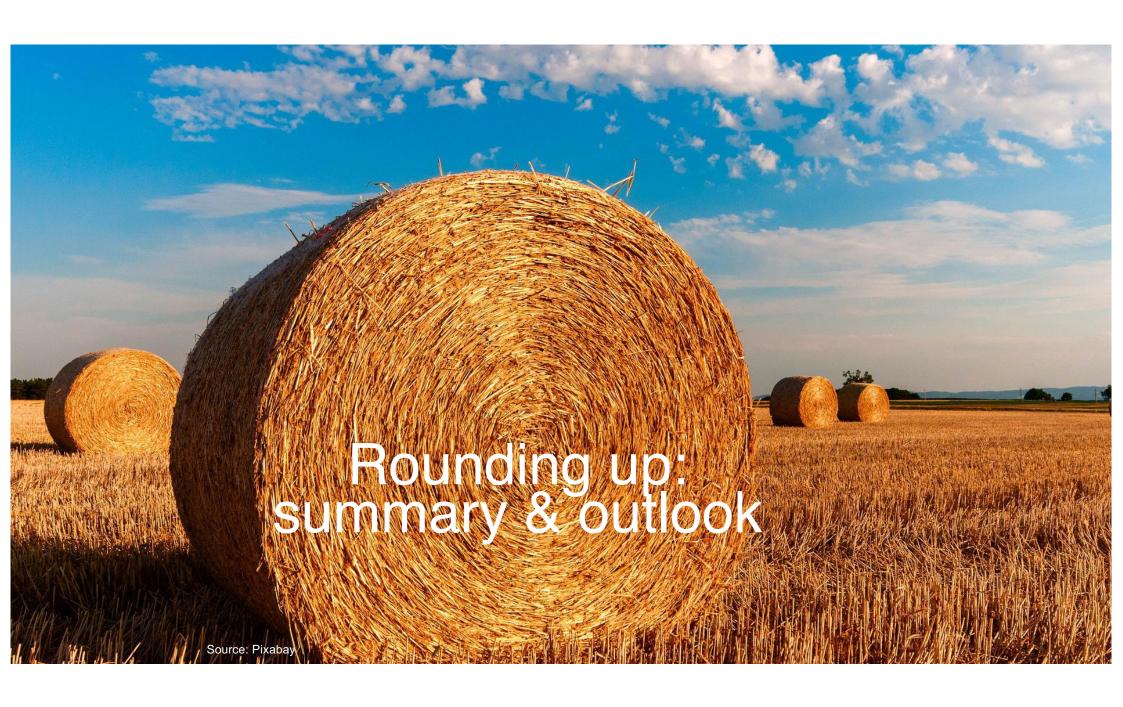




Co-financed by the European Union, Horizon Europe programme, Grant Agreement No 101070093.

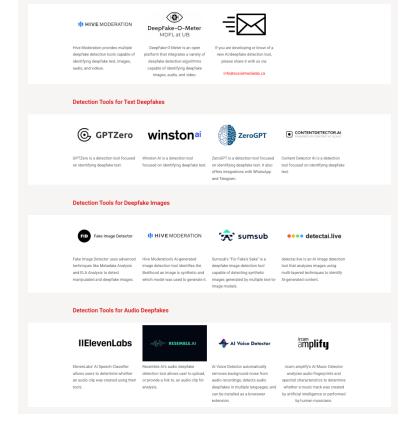
Additional funding from Innovate UK grant No 10039055 and the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract No 22.00245





Limitations, or challenges, of (part)automated / Al-supported factchecking and verification

- Accuracy (of tools)
- Transparency (of tools, models, training data, ...)
- Access (to tools)
- "Lagging behind" (developments)
- Awareness of opportunities and shortcomings
- Rules / regulations
- Corporate matters

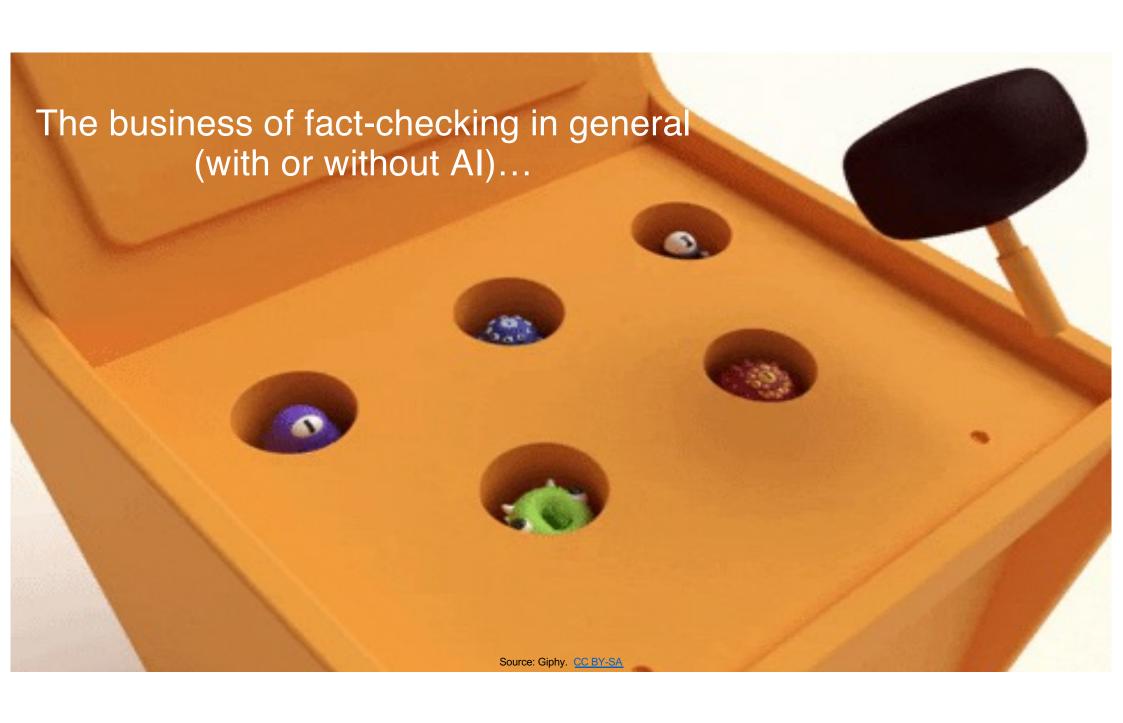


• ...

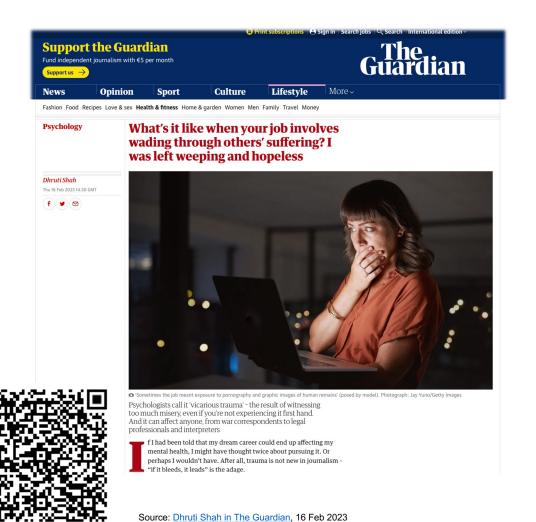
Source: Deepfakestracker

Outlook

- Focus on journalism:
 - Don't throw overboard established and useful principles
 - Try and keep up with technological developments
- Focus on media organisations:
 - Be a facilitator
 - Have a clear strategy
 - Adapt accordingly, if / when necessary
- Focus on democracy
 - Regulation? (enforcement)
 - Role of platforms and providers
- Focus on research:
 - Resources
 - Coordination of related efforts
 - Data access
- Focus on factchecking
 - Keeping up and get every support possible (from data access to tool usage)
 - Not always a pleasant task



Do not underestimate the mental toll this can take!





The Hidden Human Cost of Al Moderation

BY
SHIKHA SILLIMAN BHATTACHARJEE /
NANDITA SHIVAKUMAR

Training AI often means staring at humanity's worst atrocities for hours at a time. Workers tasked with this labor endure psychological injury without support — and face legal threats if they speak about it.



Customer agents sit behind computers on the Al training floor at the [24]7.ai, Inc., in Taguig City, Metro Manila, Philippines, on Wednesday, April 24, 2024. (Lisa Marie David / Bloomberg via Getty Images)



Source: Jacobin, 28 June 2025

Despite the challenges, is there a choice?







Deutsche Welle

LinkedIn: https://www.linkedin.com/in/jospa/

BlueSky: @jospang.bsky.social

X: https://x.com/jospang

Mastodon: https://mas.to/@jospang









