

AI4Debunk brings together an interdisciplinary consortium, encompassing a wide range of disciplines – from sociologists and media professionals to software developers and AI & ML experts, with the goal of covering the whole scope of disinformation processes in a comprehensive way.

Universities  
and research  
organisations

Think tanks,  
NGOs &  
Foundations

Industry and  
SMEs

Media  
companies

48

MONTHS

14

PARTNERS

8

COUNTRIES

#### SUPPORTED BY



# Innovative AI solutions to support trustworthy online activity



FOLLOW US



Funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.



[www.ai4debunk.eu](http://www.ai4debunk.eu)

# Empowering truth in the digital age

AI4Debunk is a pioneering European project that aims to equip citizens with the tools they need to navigate the digital media landscape safely and make informed decisions.

The project's methodology taps into human-AI collaboration to revolutionise the fight against disinformation. It includes:

## Holistic set of technologies

Language models  
and large language  
models

Vision models and  
large vision  
models

Speech  
processing

Multimodal  
machine learning

Human-in-the-loop  
mechanisms

Knowledge  
graphs

## Dedicated case studies

Russian propaganda related  
to the war in Ukraine

Manipulated content on  
climate change

## Interactive learning materials

Comic books to promote  
media literacy in schools

Classroom games for fake  
news detection training

# Introducing a fact-checking hub

AI4Debunk will develop 4 interfaces to help users detect disinformation, powered by the first-of-its-kind open-source debunking API.



### Web plug-in

Designed for web browsers and social media platforms. It will provide real-time notifications to users, offering insights into the reliability of the content they are viewing, regardless of its format.



### Disinfopedia

Collaborative platform built in a Wikipedia format. Users will be able to directly report suspicious content that will then be checked by human experts and potentially removed from circulation.



### App

Developed for smartphones and tablets. It will verify the accuracy of online content, including news articles, social media posts, product reviews and images. Users will be able to input information (URLs, text or images) and the app will identify its risk level for falseness.



### AR/VR interface

Integrated into the mobile app and compatible with AR glasses/headsets, smartphones, tablets, or wearable technology. It will deliver a more immersive and engaging experience for users, detecting whether data inputs from both the virtual and the physical worlds are true or fake.