D8.1 Initial reports on the modules developed

October 2025

Funded by
the European Union

Grant Agreement No.: 101135757
Call: HORIZON-CL4-2023-HUMAN-01-CNECT
Topic: HORIZON-CL4-2023-HUMAN-01-05
Type of action: HORIZON Innovation Actions

D8.1 INITIAL REPORTS ON THE MODULES DEVELOPED

Project Acronym	AI4Debunk
Project Number	101135757
Project Full Title	Participative Assistive AI-powered Tools for Supporting Trustworthy Online Activity of Citizens and Debunking Disinformation
Work package	WP 8
Task	Task 8.1
Due date	31/10/2025
Submission date	10/2025
Deliverable lead	Partners MICC-UNIFI and CNIT
Version	1.0
Authors	Stefano Berretti (MICC-UNIFI), Roberto Caldelli (CNIT)
Contributors	MICC-UNIFI, CNIT, UMONS, NUIG
Reviewers	HU
Abstract	The deliverable D8.1 – *Initial reports on the modules developed* – describes the initial tools that have been defined and developed for debunking audio, text and image/video content. The outcome of this deliverable and of the related task T8.1 will serve as starting point for the final development of the Machine Learning based tools that will continue in T9.1.
Keywords	Audio, text and image/video deepfake; debunking tools

Funded by
the European Union

DOCUMENT DISSEMINATION LEVEL

Dissemination level	
X	PU – Public
	SEN – Sensitive

DOCUMENT REVISION (HISTORY) SCHEDULING

Version	Date	Description of change	List of contributor(s)
1.1	10/10/2025	First draft	MICC-UNIFI, CNIT
1.2	17/10/2025	Draft revised by the Task 8.1 participants	MICC-UNIFI, CNIT, UMONS, NUIG
1.3	23/10/2025	Second draft	MICC-UNIFI, CNIT
1.4	28/10/2025	Internal review	HU
2.0	30/10/2025	Final version (GA)	MICC-UNIFI, CNIT

STATEMENT ON MAINSTREAMING GENDER

The AI4Debunk consortium is committed to including gender and intersectionality as a transversal aspect in the project's activities. In line with EU guidelines and objectives, all partners – including the authors of this deliverable – recognise the importance of advancing gender analysis and sex-disaggregated data collection in the development of scientific research. Therefore, we commit to paying particular attention to including, monitoring, and periodically evaluating the participation of different genders in all activities developed within the project, including workshops, webinars and events but also surveys, interviews and research, in general. While applying a non-binary approach to data collection and promoting the participation of all genders in the activities, the partners will periodically reflect and inform about the limitations of their approach. Through an iterative learning process, they commit to plan and implement strategies that maximise the inclusion of more intersectional perspectives in their activities.

Funded by
the European Union

DISCLAIMER

The AI4Debunk project has received funding from the European Union's Horizon Europe Programme under the Grant Agreement No. 101135757.

COPYRIGHT NOTICE

How to cite this report: Stefano Berretti, Roberto Caldelli, September 2025. AI4Debunk D5.3: Report on requirements. https://ai4debunk.eu/wp-content/uploads/2025/11/AI4Debunk-Deliverable-8.1.pdf

The AI4Debunk consortium is the following:

Participant number	Participant organisation name	Short name	Country
1	LATVIJAS UNIVERSITATE	UL	LV
2	FREE MEDIA BULGARIA	EUALIVE	BE
3	PILOT4DEV	P4D	BE
4	INTERNEWS UKRAINE	IUA	UA
5	CONSIGLIO NAZIONALE DELLE RICERCHE	CNR-IRPPS	IT
6	UNIVERSITÀ DEGLI STUDI DI FIRENZE	MICC/UNIFI	IT
6.1	CONSORZIO NAZIONALE INTERUNIVERSITARIO PER LE TELECOMUNICAZIONI	CNIT	IT
7	BARCELONA SUPERCOMPUTING CENTER CENTRO NACIONAL DE SUPERCOMPUTACION	BSC	ES
8	DOTSOFT OLOKLIROMENES EFARMOGES DIADIKTIOY KAI VASEON DEDOMENON AE	DOTSOFT	EL
9	UNIVERSITE DE MONS	UMONS	BE
10	UNIVERSITY OF GALWAY	NUIG	IE
11	STICHTING HOGESCHOOL UTRECHT	HU	NL
12	STICHTING INNOVATIVE POWER	IP	NL
13	F6S NETWORK IRELAND LIMITED	F6S	IE

Funded by
the European Union

LIST OF CONTENTS

SUMMARY

1 INTRODUCTION 10
1.1 FROM LLM TO A MODULAR PLATFORM 10
1.2 OBJECTIVES 13
1.3 EXPECTED OUTCOME 13
2 SIMILARITY SEARCH MODULE 13
3 DEEPFAKE DETECTION 14
3.1 WRITTEN FAKE NEWS DETECTION MODULE DEVELOPMENT 14
3.2 IMAGE/VIDEO DEEPFAKE ANALYSIS MODULE DEVELOPMENT 15
3.3 AUDIO DEEPFAKE ANALYSIS MODULE DEVELOPMENT 17
4 CROSS-MODAL COHERENCE CHECK MODULE 17
5 SIMILARITY CHECK ANNOTATION TOOL 18
6 CONCLUSIONS 20
7 REFERENCES 21

Funded by the European Union

LIST OF FIGURES

FIGURE 1. AI4DEBUNK PLATFORM 12
FIGURE 2. STANDALONE APPLICATION FOR IMAGE DEEPFAKE DETECTION. 15
FIGURE 3. ANNOTATION INTERFACE FOR SIMILARITY SCORING........ 19

Funded by
the European Union

LIST OF TABLES

TABLE 1. EVALUATION OF LLM PERFORMANCE ON TEXT DISINFORMATION DETECTION. 10

LIST OF ABBREVIATIONS

WP	Work Package
TX.Y	Task X.Y
DX.Y	Deliverable X.Y
EC	European Commission

EXECUTIVE SUMMARY

The deliverable D8.1 – *Initial reports on the modules developed* – describes the initial tools that have been defined and developed for debunking audio, text and image/video content. The outcome of this deliverable and of the related task T8.1 will serve as starting point for the final development of the Machine Learning based tools that will continue in T9.1.

1 INTRODUCTION

Deliverable D8.1 summarizes the work done in Task 8.1 for the definition and development of the initial debunking tools for text, audio and image/video content.

As a first step, this includes the analysis of different state-of-the-art solutions taking into account diverse aspects like source code availability, performance documented in literature, required input data and comparisons. The outputs of such modules taken individually will constitute both a primary step of monomodal evaluation and, at the same time, the input for a successive more general and comprehensive analysis and development performed within Task 9.1.

This document is organized as follows: a brief introduction on the objectives and the expected outcomes of Task 8.1, whose D8.1 refers to, is first provided; then, a summary of the overall architecture of the debunking engine is described with reference to the different modules for text, image/video, and audio deepfake detection; in Sections 2 to 4, the individual debunking tools are described; finally, some conclusions and considerations are drawn.

1.1 FROM LLM TO A MODULAR PLATFORM

Given the state-of-the-art at the beginning of this task, Large Language Models (LLMs) were considered for their promising performances. So, a first exploratory experiment was carried out to evaluate their aptitude to detect disinformation. We therefore tested the best performing models then, on the first version of our AI4Debunk dataset collected by our colleagues in WP6 (the datasets are on HuggingFace but they are still private https://huggingface.co/datasets/AI4Debunk/war-in-ukraine-disinformation-detection_3rdrelease; https://huggingface.co/datasets/AI4Debunk/climate-change-disinformation-detection-3rdrelease). The results are shown in Table 1.

Model	Category	Accuracy (%)
Qwen2	7B-Instruct	62.33
Qwen2	1.5B-Instruct	9.03
Meta-Llama3	8B-Instruct	74.31
Meta-Llama3	70B	43
Gemini	Flash	**82.73**
Gemma2	9bit	**82.99**
Gpt-4	NA	71.18
Gpt-3	turbo	57.99
bert	EN	76.21
distilbert	EN	68.40
bert	RU	55.90
distilbert	RU	55.03

TABLE 1. EVALUATION OF LLM PERFORMANCE ON TEXT DISINFORMATION DETECTION.

For this, a simple prompt containing the news representing a short piece of text (a title, a few sentences with a limited number of characters, etc.) and asking the different LLMs to classify the given news as disinformation or not was used. Since that version of the AI4Debunk dataset contained only disinformation, the **Accuracy** in Table 1 represents the correctly detected disinformation (for the sake of clarity the value can be considered as TPR – true positive rate, where positive means fake). Two languages (English and Russian) were used depending on whether the models could handle multi-lingual data or not and depending on the news language (some of them were translated first if it was necessary).

The results showed that even though the results were above 80% for some LLMs (see the green box in Table 1), these systems evidence several drawbacks:

1. The results were inconsistent from one model to the other, which might suggest that the models' performance is heavily dependent on the training data (as expected). Which means that the current models, even if they performed well with current data, will perform poorly with newly unseen datasets with different contextual information (different names of political figures, news terms that didn't exist before, new disease names, etc.). Although this is not a hard conclusion, it is to be considered.
2. The models' size and processing cost makes it very complicated to deploy and use as is. This does not count the environmental impact of using these models as is.
3. Although they could generate explanation for their choices which we have tested, we could not be certain of their veracity and, because of the two previous points, did not see the need to push the experiments further in that sense.

Instead of building a prompt-based LLM disinformation classification, we opted for a more transparent and flexible approach that relies on building a *modular platform*, in which each module can be used individually as a decision support on whether news is disinformation or not, or the outputs of which could be aggregated to estimate a score of disinformation, the "disinfoscore" mentioned in the proposal and detailed in T8.4.

The platform is developed in a modular way so that the datasets, models (or entire deepfake detection systems) and metrics can be replaced without modifying the rest of the architecture.

According to this, the designed overall framework is depicted in Figure 1, where the general architecture of the AI4Debunk modular platform is reported.

Funded by the European Union

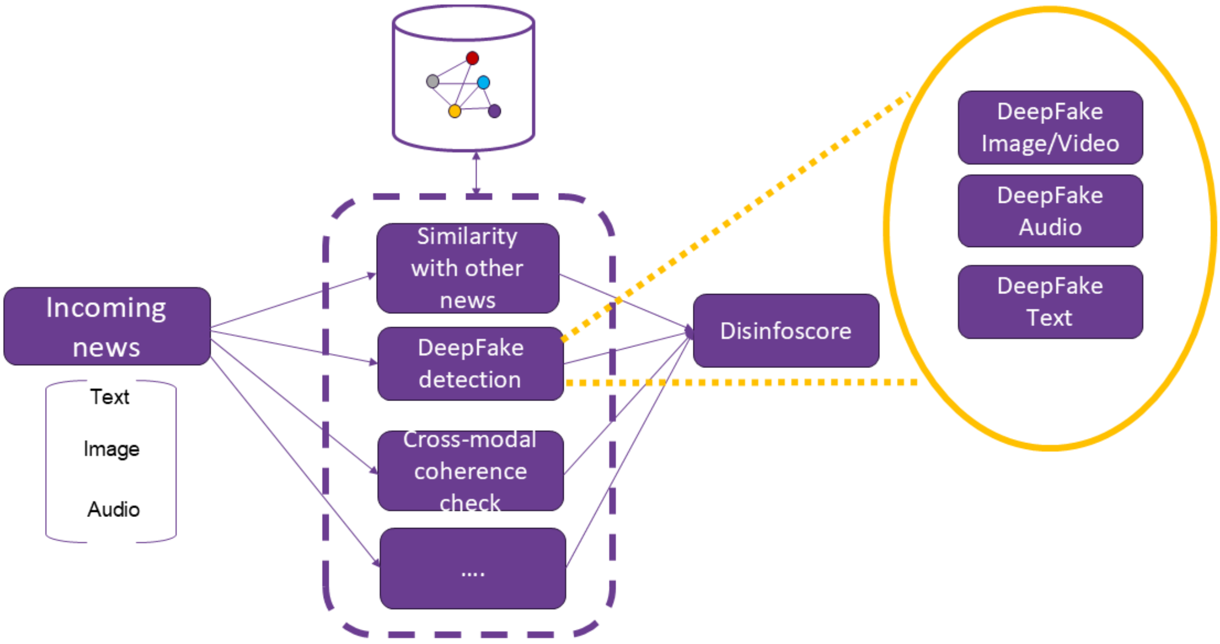

FIGURE 1. AI4DEBUNK PLATFORM

The general idea is that the AI4Debunk platform receives incoming news (left of the figure) to be analysed. The input news can be in the form of a web page including text, image/video and audio, or can be any combination of these modalities including each modality alone. A preprocessing step is applied splitting the types of data according to the different modules before the data are dispatched. This multimodal content is dispatched to any module in the purple dashed area. This includes:

1- a **similarity estimation module** which, given a news item, returns the most similar news present in a knowledge base with some of them supporting and some of them opposing the reference news (see Section 2).
2- A **deepfake detection module** that includes tools for deepfake detection from text, image/video and audio (the modules surrounded in orange in the figure):
    - Fake news detection from text: this module takes as input a piece of text (a post, a part of a web page, a message, etc.) and performs an authenticity verification. It is expected that the output response of such a module will contain not only a binary or a statistical assessment but other supporting information on specific phrases and/or words to allow further investigations. This module is further detailed in Section 3.1;
    - Image/video deepfake detection: such a module takes as input an image or a video and carries out a check for its authenticity. The output of such a module contains different information presented in multiple ways like localization heat maps, binary assessment (e.g., fake or not), probabilistic evaluation and so on. In the case of video, the response is frame-based. Details on this module are given in Section 3.2;
    - Audio deepfake detection: Audio signals from videos or from standalone audio files are also processed for detecting their genuine content. The description of this module is expanded in Section 3.3.

Funded by the European Union

3- A **cross-modal coherence check module** that takes an image and the corresponding text (which is supposed to be its legend/caption) and verifies their compatibility, i.e. whether the text and image match.

These three types of modules were selected based on the work of previous similar EU projects on fake news and disinformation and current commercial tools fighting against disinformation given by companies like Google.

The initial modules have been tested on existing datasets, available on-line, that have been gathered to train and test the models on data containing disparate characteristics.

1.2 OBJECTIVES

Deliverable D8.1 aims to report on all the different modules developed for disinformation decision support except the cross-modal coherence check module and the multimodal modular platform calculating the disinformation score which will be reported in more detail in D8.2 and D8.4. These constitute the initial instruments for debunking input news with a monomodal processing applied to individual content and provides the starting point for continuing the work in Task 9.1.

1.3 EXPECTED OUTCOME

The expected outcome of Task 8.1, consequently reflected in the deliverable D8.1, is to provide an initial set of modules for decision support on whether a piece of news is disinformation or not. This allows preliminary evaluation tests on benchmark datasets available on the web, so providing useful indications on the current limitations and challenges. This work is also preparatory to Task 9.1 (from M23 to M42) where the modules will be further expanded and consolidated, making them work into the platform.

2 SIMILARITY SEARCH MODULE

This module is engineered for text-based analysis. It accepts a reference text as input and ranks a corpus of documents from a specified database in descending order of semantic similarity. The similarity metric is derived from the cosine similarity between embedding vectors, which are generated using the Contrastive Language–Image Pre-training (CLIP) model.

The selection of the CLIP model is predicated on several key factors. Primarily, standardizing on a single model architecture across multiple modules streamlines deployment and ensures system-wide consistency. Furthermore, CLIP provides an optimal balance of computational efficiency, performance, and inherent multimodal capabilities, allowing for future extensions to include visual data.

Subsequent to the similarity ranking, a custom-developed algorithm classifies the top-ranked documents in batches of 50 by decreasing order of similarity as either supporting or opposing the viewpoint of the reference text until a number “k” of news is returned with half of them “supporting” and half of the “opposing”. The final output consists of a curated selection from both categories within the top-k results. This balanced presentation of corroborating and conflicting information equips the user to make a more informed and nuanced assessment regarding the potential for disinformation within the source text.

Hereafter some reference datasets are listed:

- The **Fakeddit dataset** is a large-scale benchmark dataset designed for **fake news and misinformation detection.** It’s one of the most comprehensive publicly available datasets for multi-modal fake news classification, as it includes **text, metadata, and images** from **Reddit** posts.
- The dataset **EU DisinfoTest** is a benchmark introduced in 2024 to evaluate how well large language models (LLMs) can detect disinformation narratives. It comprises **over 1,300 narrative-statements**, each labelled as either a credible (“real”) narrative or a disinformative one.
- The **GossipCop, PolitiFact, and Twitter-Sentiment-Analysis dataset** is a composite benchmark released on the **Science Data Bank** that integrates three widely used resources for misinformation and sentiment research. It combines the **GossipCop** and **PolitiFact** fake news datasets—each containing news articles labeled as real or fake based on professional fact-checking—with a **Twitter sentiment dataset** of posts annotated for positive, negative, or neutral sentiment. By merging these datasets, the collection enables researchers to explore relationships between **disinformation detection and public sentiment**, and to evaluate models across both **news-veracity** and **social-media sentiment** tasks. The dataset thus provides a unified, well-documented resource for studying misinformation, credibility assessment, and emotional tone in online discourse.

3 DEEPFAKE DETECTION

The deepfake detection module is hereafter described in its three composing sub-modules: *deepfake text*, *deepfake image/video* and *deepfake audio*.

3.1 WRITTEN FAKE NEWS DETECTION MODULE DEVELOPMENT

Such a module will take as input a piece of text (a post, a part of a web page, a message, etc.) and will perform an authenticity verification. This method operates through a sequential, three-stage pipeline to verify a claim:

1. **Evidence Retrieval:** first, the system searches a large trusted database (like Wikipedia) to find text passages that are relevant to the claim in question.
2. **Claim-Passage Comparison:** next, it uses a compact, pre-trained neural network to compare the retrieved evidence directly against the original claim. This model is trained to identify subtle patterns of textual agreement and contradiction, outputting a score for how well the evidence supports or refutes the claim.
3. **Decision & Rationale Generation:** finally, these individual comparison scores are aggregated. The system then makes a final verdict (e.g., "Supported," "Refuted") and automatically highlights the

specific words or phrases in both the claim and the evidence that were most critical for reaching its conclusion, providing a transparent rationale.

Hereafter some reference datasets are listed:

- LIAR-RAW is an expanded version of the LIAR-PLUS dataset. The dataset employs a fine-grained six-class classification scheme: pants-fire, false, barely-true, half-true, mostly-true, and true. Each claim in the dataset is accompanied by relevant raw news reports and documents that were collected during the dataset's creation.
- The RAWFC dataset, derived from Snopes.com claims, implements a more condensed three-class classification system (false, half, true). The dataset includes claims along with their associated raw reports retrieved using claim keywords.

3.2 IMAGE/VIDEO DEEPFAKE ANALYSIS MODULE DEVELOPMENT

An interface has been developed for the image/video deepfake detection tool to facilitate the evaluation and comparison of different deepfake detection models. This interface, as evidenced in the Figure 2 hereafter, is designed to permit the verification of integrity of both images and video sequences by switching between the two modalities. It contains two different sub-modules, each of them dedicated to the specific modality.

FIGURE 2. STANDALONE APPLICATION FOR IMAGE DEEPFAKE DETECTION.

This interface is independent from the AI4Debunk platform shown in Figure 1 and it has been conceived to permit stand-alone usage. The interface grants the selection of different implemented techniques (see hereafter), the execution and the visualization of the achieved results.

<u>**Image deepfake**</u>

TruFor (Guillaro, 2023) is designed to detect and localize image forgeries. Its core idea is to compare the RGB image with a learned noise representation, which acts as a fingerprint of the image. By combining these two sources of information, the model outputs an anomaly localization map that highlights suspicious regions. In addition, it uses Noiseprint++ together with the RGB image to compute a confidence map, which identifies less reliable areas of the anomaly map. Finally, the anomaly and confidence maps are aggregated to produce a global integrity score for the image.

Briefly, the architecture consists of the following steps: the process starts with the RGB image and its corresponding extracted noise fingerprint. These inputs are fed into two parallel Mix Transformer (MiT-B2) encoder branches, creating separate feature maps for visual and noise information. The feature maps are fused in a Cross-Modal Feature Rectification Module (CM-FRM) to compare visual and noise information and find discrepancies. The fused features are passed to a lightweight All-MLP decoder to generate the final localization and confidence maps. Lastly, the maps are aggregated to produce the single integrity score.

MantraNet (Wu, 2019) is an end-to-end image forgery detection and localization solution, taking an image as input to predict pixel-level forgery likelihood map. ManTraNet is composed of two sub-networks: (1) Image Manipulation Trace Feature Extractor, a VGG-style network for the image manipulation classification task, sensitive to different manipulation types , encoding the image manipulation in a patch into a fixed dimension feature vector; (2) Local Anomaly Detection Network, the anomaly detection network to compare a local feature against the dominant feature averaged from a local region (Z-score), whose activation depends on how far a local feature deviates from the reference feature instead of the absolute value of a local feature. An LSTM analyses the Z-scores sequentially and the output is used by a final convolutional layer to produce the forgery localization map.

CLIP_BSID's architecture (Cozzolino, 2024) relies on CLIP embeddings and uses a linear classifier to distinguish between real and AI-generated images. The training setup is straightforward: starting from N real images with their captions, a text-to-image generator is used to produce N synthetic images based on the same captions. This results in pairs of similar images (real vs generated). All images are passed through a pre-trained and frozen CLIP Image Encoder (e.g., ViT-L/14), which outputs a high-dimensional feature vector for each image. These N+N feature vectors are then used to train a linear SVM classifier, which learns to separate real and AI-generated samples.

<u>**Video deepfake**</u>

The ***video DF sub-module*** takes as input the to-be-checked video sequence and provides the result of the integrity assessment as output. Also in this case, it has been designed to grant a wider flexibility and modularity to permit the update and/or the adjunction of further detection tools with respect to those available so far. The techniques employed for this data modality do not consider audio and they will produce a “Real” or “Fake” classification for each frame of the video computing fake-scores for each frame. These scores allow us to individuate which portions of the video have been significantly altered the most and thus have a stronger impact on the possible final assessment on the sequence. We can use this information to provide users with an interactive visual representation highlighting these frames to be reviewed. Two techniques, mainly oriented to check for deepfake on human faces, are available so far: one is based on RGB frames, and one operates in the frequency domain by applying the Discrete Wavelet Transform (DWT) to the frames of the video.

The **RGB-based model** (Rossler, 2019) uses a ResNet-50 model pre-trained on ImageNet as a strong feature extraction base. The top classification layer of the ResNet is replaced, and the entire network is then fine-tuned on the FaceForensics++ (FF++) dataset. The model does not receive the whole image, but it extracts the cropped faces from the video using the method in Thies (2016) and then feeds it to the ResNet-50.

The **Wavelet-Domain-based mode**l still relies on ResNet50. The model takes an RGB image and each colour channel is decomposed in four sub-bands (LL, LH, HL, HH) by applying a Discrete Wavelet Transform decomposition. The resulting 12 maps are stacked to form a single 12-channel input tensor, representing the image's frequency information and are subsequently fed into a ResNet-50. Lastly, the final feature vector is passed to a classification head to produce the resulting probability.

Both models in this module employ the ResNet-50 architecture, a deep convolutional neural network with 50 layers, widely used as a backbone for image-related tasks.

3.3 AUDIO DEEPFAKE ANALYSIS MODULE DEVELOPMENT

The audio deepfake analysis relied on prompt-engineering approaches as well as simpler classification.

The latter model selected was Wav2Vec2-based fine-tuned classifier[1]

The Qwen2-Audio-7B was used in a prompt-engineered approach to classify a given audio as deepfake or not, due to its ability to accept audio and text as input, to its relatively lighter size and good performances in benchmarks.

These will be compared in the benchmarks explained further in this report.

4 CROSS-MODAL COHERENCE CHECK MODULE

The Cross-Modal Coherence Check Module and the multimodal modular platform calculating the disinformation score will be described in more detail within D8.2 and D8.4.

[1] https://huggingface.co/mo-thecreator/Deepfake-audio-detection

Funded by the European Union

5 SIMILARITY CHECK ANNOTATION TOOL

Although not mentioned in the proposal, we took the initiative to develop a similarity check annotation tool.

To the best of our knowledge, no dataset exists for the evaluation of similarity check for disinformation text, especially containing the added information of supporting or opposing a reference news. Such a dataset is therefore being built.

For this, our approach is to gather data from existing fake news and disinformation datasets and recreate a dataset using a semi-automated manual annotation process.

Our annotation protocol is as follows: for each reference text, the annotator is asked to label 6 associated texts into one of the four tags:

1- Supporting: the associated news is of a similar topic and supports the reference's statement
2- Against: the associated news is of a similar topic and opposed the reference's statement
3- Not related: the news has no topic in common
4- Undetermined: the associated news is of a similar topic but it is not clear whether they are supporting or opposing each other

They are also asked to grade the relationship with a score estimating the strength of the chosen tag.
For this, an annotation interface was built as shown in Figure 3.

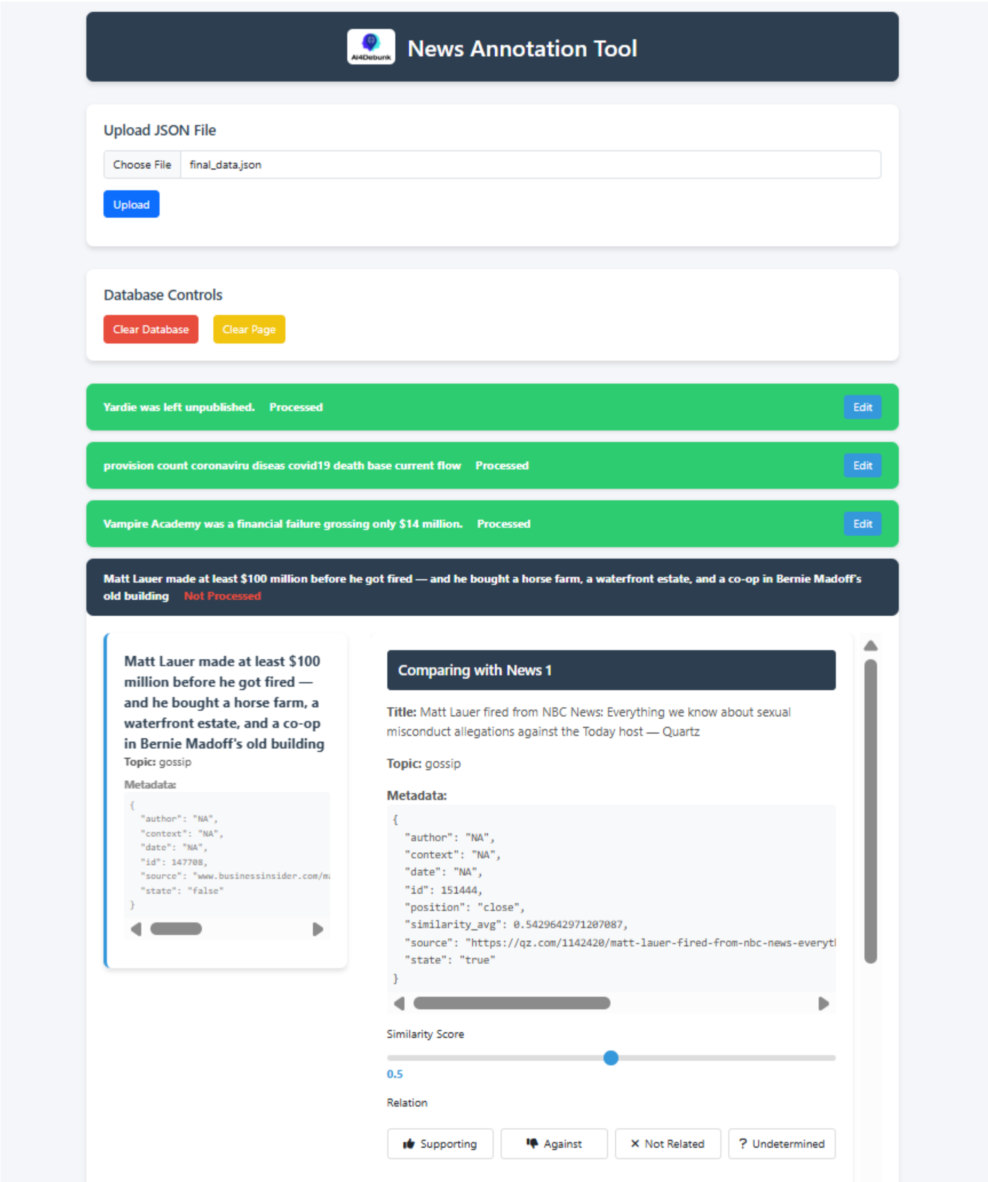

FIGURE 3. ANNOTATION INTERFACE FOR SIMILARITY SCORING.

A first dataset gathered, and first annotations are being carried out and will proceed in WP9.1.

6 CONCLUSIONS

The deliverable D8.1 reports the work done in Task 8.1 for the definition and development of the initial debunking tools for text, audio and image/video content. The outputs of such tools, taken individually, will constitute a primary step of monomodal evaluation.

7 REFERENCES

(Guillaro, 20223) Fabrizio Guillaro, Davide Cozzolino, Avneesh Sud, Nicholas Dufour, Luisa Verdoliva. "TruFor: Leveraging all-round clues for trustworthy image forgery detection and localization," IEEE/CVF *Conf. on Computer Vision and Pattern Recognition* (CVPR), 2023.

(Wu, 2019) Yue Wu, Wael AbdAlmageed, Premkumar Natarajan. "ManTra-Net: Manipulation Tracing Network For Detection And Localization of Image Forgeries With Anomalous Features," IEEE/CVF *Conf. on Computer Vision and Pattern Recognition* (CVPR), 2019.

(Cozzolino, 2024) Davide Cozzolino, Giovanni Poggi, Riccardo Corvi, Matthias Nießner, Luisa Verdoliva. "Raising the Bar of AI-generated Image Detection with CLIP," IEEE/CVF *Conf. on Computer Vision and Pattern Recognition* Workshops (CVPRW), 2024.

(Rössler, 2019) Andreas Rössler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, Matthias Nießner. "FaceForensics++: Learning to Detect Manipulated Facial Images," IEEE/CVF *Int. Conf. on Computer Vision* (ICCV), 2019.

Funded by
the European Union

Review Sheet of Deliverable/ Milestone Report

D8.1 Report on Requirements

Editor(s):	Stefano Berretti (MICC-UNIFI) Roberto Caldelli (CNIT)
Responsible Partner:	Partners MICC-UNIFI and CNIT
Status-Version:	Final draft – X.X
Date:	
Distribution level (CO, PU):	Public
Reviewer (Name/Organization)	
Review date	30-10-2025

Disclaimer: This assessment reflects only the author's views and the European Commission is not responsible for any use that may be made of the information contained therein"

Mark with X the corresponding column:

Y= yes	**N= no**	**N = not applicable**

ELEMENT TO REVIEW	Y	N	NA	COMMENTS
FORMAT: Does the document … ?				
…include editors, deliverable name, version number, dissemination level, date, and status?	x			
…contain a license (in case of public deliverables)?	x			
…include the names of contributors and reviewers?	x			
….has a version table consistent with the document's revision?	x			
… contain an updated table of contents?		x		There is a table but it is not up-to-date
… contain a list of figures consistent with the document's content?	x			
… contain a list of tables consistent with the document's content?	x			
… contain a list of terms and abbreviations?	x			
… contain an Executive Summary?		x		Findings are not summarised
… contain a Conclusions section?	x			
… contain a List of References (Bibliography) in the adequate format, if relevant?		x		
… use the fonts and sections defined in the official template?	x			
… use correct spelling and grammar?	x			
… conform to length guidelines (50 pages maximum (plus Executive Summary and annexes)	x			

Funded by the European Union

ELEMENT TO REVIEW	Y	N	NA	COMMENTS
… conform to guidelines regarding Annexes (inclusion of complementary information)			x	
… present consistency along the whole document in terms of English quality/style? (to avoid accidental usage of copy&paste text)	x			
About the content…				
Is the deliverable content correctly written?	x			
Is the overall style of the deliverable correctly organized and presented in a logical order?		x		See remarks for section 3.1
Is the Executive Summary self-contained, following the guidelines and does it include the main conclusions of the document?		x		
Is the body of the deliverable (technique, methodology results, discussion) well enough explained?	x			
Are the contents of the document treated with the required depth?	x			
Does the document need additional sections to be considered complete?		x		
Are there any sections in the document that should be removed?		x		
Are all references in the document included in the references list?		x		There is no list
Have you noticed any text in the document not well referenced? (copy and paste of text/picture without including the reference in the reference list)	x			See previous
SOCIAL and TECHNICAL RESEARCH WPs (WP4, 5, 12, 13, 14)				
Is the deliverable sufficiently innovative?	x			
Does the document present technical soundness and its methods are correctly explained?	x			
What do you think is the strongest aspect of the deliverable?				Sound foundation;
What do you think is the weakest aspect of the deliverable?				-
Please perform a brief evaluation and/or validation of the results, if applicable.				-
AI AND TECNOLOGICAL WPS (WP6 – WP11)				
Does the document present technical soundness and the methods are correctly explained?	x			
What do you think is the strongest aspect of the deliverable?				presents a good overview of the system

Funded by the European Union

ELEMENT TO REVIEW	Y	N	NA	COMMENTS
What do you think is the weakest aspect of the deliverable?				-
Please perform a brief evaluation and/or validation of the results, if applicable.				-
DISSEMINATION AND EXPLOITATION WPs (WP15 – WP17)				
Does the document present a consistent outreach and exploitation strategy?				
Are the methods and means correctly explained?				
What do you think is the strongest aspect of the deliverable?				
What do you think is the weakest aspect of the deliverable?				
Please perform a brief evaluation and/or validation of the results, if applicable.				

SUGGESTED IMPROVEMENTS

PAGE	SECTION	SUGGESTED IMPROVEMENT

CONCLUSION

Mark with X the corresponding line.

	Document accepted, no changes required.
x	Document accepted, changes required.
	Document not accepted, it must be reviewed after changes are implemented.

Please rank this document globally on a scale of 1-5 (1 = poor, 5= excellent) – using a half point scale.
Mark with X the corresponding grade.

Document grade	1	1.5	2	2.5	3	3.5	4	4.5	5
							x		

Funded by
the European Union