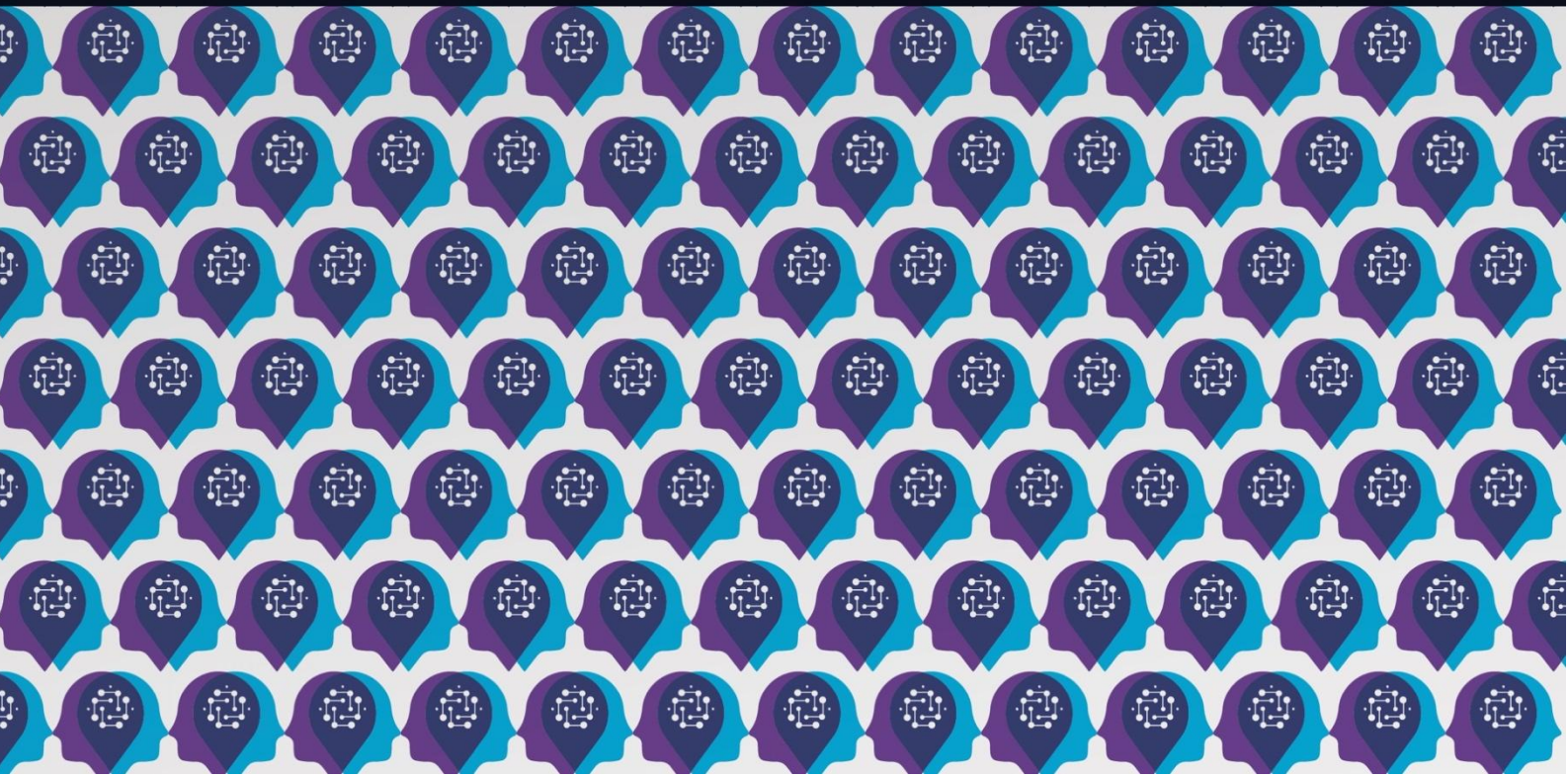
AI4Debunk

D10.4 Report on the definition of the collaborative platform

February 2026

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

D10.4 REPORT ON THE DEFINITION OF THE COLLABORATIVE PLATFORM

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 10
Task	Task 10.4
Due date	28/02/2026
Submission date	10/02/2026
Deliverable lead	Stichting InnovativePower
Version	0.3
Authors	Jan Kragt (IP), Marcel Keijzer (IP)
Contributors	Yelyzaveta Borysova (IP), Chun Fei Lung (HU), Franc van der Bent (HU), Pascaline Gaborit (Pilot4DEV)
Reviewers	Name Surname (EUALIVE)
Abstract	This deliverable presents the definition, structure, and implementation design of the AI4Debunk Collaborative Platform, an interactive and participatory web environment enabling citizens, journalists, researchers, and experts to collaborate in detecting, validating, and analysing online disinformation. The platform acts as a public-facing hub within the AI4Debunk ecosystem, connecting human input with automated disinformation detection engines. It serves as an open knowledge repository

Funded by the European Union

("Disinfopedia") and as a communication interface with the browser plug-in, mobile app, and AR/VR tools developed under WP10 and WP11.

The document describes the functional and quality requirements, user experience and accessibility design, system architecture, privacy and ethical considerations, and integration with the AI4Debunk backend infrastructure. It also details the steps towards technical validation, pilot deployment, and alignment with European standards for trustworthy AI and digital inclusion.

Keywords	disinformation, misinformation, participatory validation, collaborative platform, human-centered design, explainable AI, knowledge graph, EU Horizon Europe

DOCUMENT DISSEMINATION LEVEL

Dissemination level	
X	PU - Public
	SEN - Sensitive

DOCUMENT REVISION HISTORY

Version	Date	Description of change	List of contributor(s)
0.1	25/11/2025	First working draft that is distributed only among technical partners actively involved in the design of the browser extension or the collaborative platform.	IP,HU,Pilot4DEV
0.2	25/01/2026	Second working draft that is distributed among partners actively involved in the design of the browser extension or the collaborative platform	IP,HU,Pilot4DEV
0.3	10/02/2026	Final draft for internal review by EUALIVE	IP, HU

Funded by the European Union

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 implementing strategies that foster comprehensive inclusion of intersectional perspectives throughout the project

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: Kragt, J.H., Keijzer, M. (2026). AI4Debunk report on the definition of the collaborative platform.

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	UNIVERSITA 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	NATIONAL UNIVERSITY OF IRELAND 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

TABLE OF CONTENTS

Funded by the European Union

Funded by the European Union

LIST OF FIGURES

Funded by the European Union

LIST OF TABLES

Funded by the European Union

ABBREVIATIONS

AI	Artificial Intelligence
API	Application Programming Interface
B2B	Business-to-business
CSS	Cascading Style Sheets
DOM	Document Object Model
EU	European Union
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
I18n	Internationalisation
L10n	Localisation
LLM	Large Language Model
RPC	Remote Procedure Call
SUS	System Usability Scale
TRL	Technology Readiness Level
UI	User Interface
UML	Unified Modeling Language
UX	User Experience
WCAG	Web Content Accessibility Guidelines
WP	Work Package

Funded by the European Union

EXECUTIVE SUMMARY

The AI4Debunk Collaborative Platform is one of the four main user-facing components developed under WP10, "Definition of Different Interfaces Available for Online Citizens." It provides a participatory digital environment for identifying, analyzing, and validating online disinformation. As a web-based "hub," it enables citizens, journalists, and experts to contribute to a shared knowledge base, connecting human insights with automated detection algorithms developed in other work packages.

The Collaborative Platform (internally referred to as *Disinfopedia*) complements three other AI4Debunk interfaces:

- **Browser Extension (WP10.2):** Provides immediate feedback on disinformation encountered during browsing.
- **Mobile App (WP10.3):** Allows real-time content validation on handheld devices.
- **VR Interface (WP10.5):** Offers immersive visualization and educational engagement.

Together, these interfaces ensure a consistent, transparent, and user-friendly interaction with AI4Debunk's backend modules: the knowledge graph, AI detection engines, and validation committee mechanisms.

The Collaborative Platform acts as both a data consumer and contributor. Users upload potentially misleading content, review validation outcomes, and discuss results with peers. The system presents the output of automated AI assessments (DisinfoScores) and allows deeper exploration of supporting evidence and sources. These results are cross-linked with related claims, fostering transparency and traceability.

In alignment with a human-centered design philosophy, the platform prioritizes **trust, explainability, and inclusivity**. It features an interface with progressive information disclosure system:

- **Standard mode,** presenting simplified validation results (green/amber/red "traffic light" system or other simplified presentation).
- **Advanced mode,** offering granular AI module outputs, provenance links, and extended explainability data.

The platform architecture is modular and scalable, supporting future integrations and external APIs. It complies with European accessibility standards (WCAG 2.2, ISO 9241-210:2019) and incorporates GDPR-based privacy-by-design principles. The validation process follows a hybrid model combining automated detection with periodic human oversight by a dedicated Validation Committee.

From a technical perspective, the platform uses a microservices-based architecture and open standards to connect with the AI4Debunk API. Content and metadata are exchanged securely between the front-

Funded by the European Union

end interfaces and the back-end “engine room,” where disinformation detection modules and the knowledge graph are hosted.

The deliverable outlines:

- Functional and quality requirements of the platform.
- Human-centered design and user experience considerations.
- Technical architecture and integration layers.
- Implementation timeline, including testing and evaluation stages.
- Ethical, privacy, and sustainability aspects ensuring compliance and longevity.

The Collaborative Platform is scheduled for completion by **Month 40**, with technical validation and pilot testing in cooperation with end-user communities. The outcome will be an operational prototype (TRL 7) that bridges automated AI analysis with participative citizen engagement, supporting a more transparent, accountable, and trustworthy digital information landscape.

Funded by the European Union

1 INTRODUCTION

AI4Debunk is a European initiative that aims to develop human-centred, multimodal, and collaborative AI tools to fight disinformation and protect democratic values. Its human-centred solutions include an educational VR app and three debunking user interfaces: a browser extension, a smartphone app, and a collaborative platform. Each of these user interfaces will target different user groups, as evidenced by a survey distributed among a diverse group of respondents.

These interfaces enable ordinary European citizens to directly assess the accuracy of online content using a combination of state-of-the-art AI/ML models and a human-in-the-loop mechanism designed to enhance data quality and foster greater trust in the debunking process. Although all interfaces share common functionality, each takes advantage of the unique capabilities afforded by the platform they are based on.

This report focusses on the AI4Debunk collaborative platform (Disinfopedia). Users can upload potentially misleading content, review validation outcomes, and discuss results with others. The system presents the output of automated AI assessments (DisinfoScores) and allows deeper exploration of supporting evidence and sources. These results are cross-linked with related claims, fostering transparency and traceability.

Funded by the European Union

2 FUNCTIONAL AND QUALITY REQUIREMENTS

This section defines the functional, technical, and quality requirements for the AI4Debunk Collaborative Platform, ensuring its compliance with the objectives and principles established under WP10. The requirements are derived from stakeholder interviews, user studies, and cooperation with AI4Debunk partners.

2.1 FUNCTIONAL REQUIREMENTS

The Collaborative Platform is designed as a central hub for knowledge sharing and validation of online content. It combines AI-generated insights with human contributions through participative workflows. The main functional requirements are summarized below. A detailed list of requirements can be found in ANNEX 5. Functional requirements breakdown.

Table 1. Main functional Requirements Overview

ID	Category	Requirement Description	
FR-1	Submission	Allow users to submit content (URLs, text, images, videos) for disinformation analysis.	Must have
FR-2	Validation	Display automated DisinfoScore results derived from AI4Debunk detection modules.	Must have
FR-3	Explainability	Provide transparent explanations for each validation, including source provenance and AI confidence levels.	Must have
FR-4	User Collaboration	Enable users to comment, suggest corrections, and discuss results under moderation.	Should have
FR-5	Search & Filter	Allow full-text and metadata-based search for validated items, using filters for topic, language, and credibility level.	Must have
FR-6	Educational Content	Host guides, tutorials, and interactive lessons on detecting misinformation.	Should have
FR-7	Integration	Provide direct access and data sharing with the browser extension, mobile app, and AR/VR interface.	Must have
FR-8	Downloads	Offer public access to AI4Debunk software components and open datasets.	Should have
FR-9	Notifications	Notify registered users of updates or new validations in their area of interest.	Could have

Funded by the European Union

2.2 QUALITY REQUIREMENTS

The quality requirements ensure performance, reliability, and inclusivity.

Table 2. Quality Requirements Overview

ID	Category	Requirement Description	Compliance Target
QR-1	Accessibility	Platform meets WCAG 2.2 Level AA standards.	
QR-2	Usability	Minimum SUS score ≥ 72 in user testing.	Achieved by M44
QR-3	Performance	Average page load < 2 seconds; backend API latency < 500ms.	
QR-4	Security	Full GDPR compliance; encryption (TLS 1.2 and 1.3) for data in transit.	100%
QR-5	Multilingualism	Support for all official EU languages.	Progressive rollout
QR-6	Transparency	Explainable AI outputs displayed for each validated entry.	Continuous
QR-7	Maintainability	Modular design enabling long-term maintenance beyond project duration.	Required

See also ANNEX 4.

2.3 USER PERSONAS AND SCENARIOS

AI4Debunk targets four key user groups, each with distinct motivations and needs:

1. General Citizens: Casual users seeking quick disinformation checks. They value simplicity and trustworthy results.
2. Journalists and Media Professionals: Require advanced analytics, provenance tracking, and source reliability indicators.
3. Researchers and Educators: Focus on accessing aggregated datasets and AI explainability features for study or teaching.
4. Policy Makers and NGOs: Use the platform as a monitoring and reporting tool to understand disinformation trends.

Example Scenario:
A journalist identifies a viral claim on social media and submits the URL through the platform. The AI modules generate a DisinfoScore with explanations. The journalist explores the supporting metadata, cross-references linked sources and downloads a report for editorial verification.

Funded by the European Union

3 PLATFORM DESIGN AND USER EXPERIENCE

The platform follows a modular design enabling flexibility and scalability. The user interface is divided into distinct functional zones, ensuring clarity and accessibility. It employs responsive design principles for desktop, tablet, and mobile use.

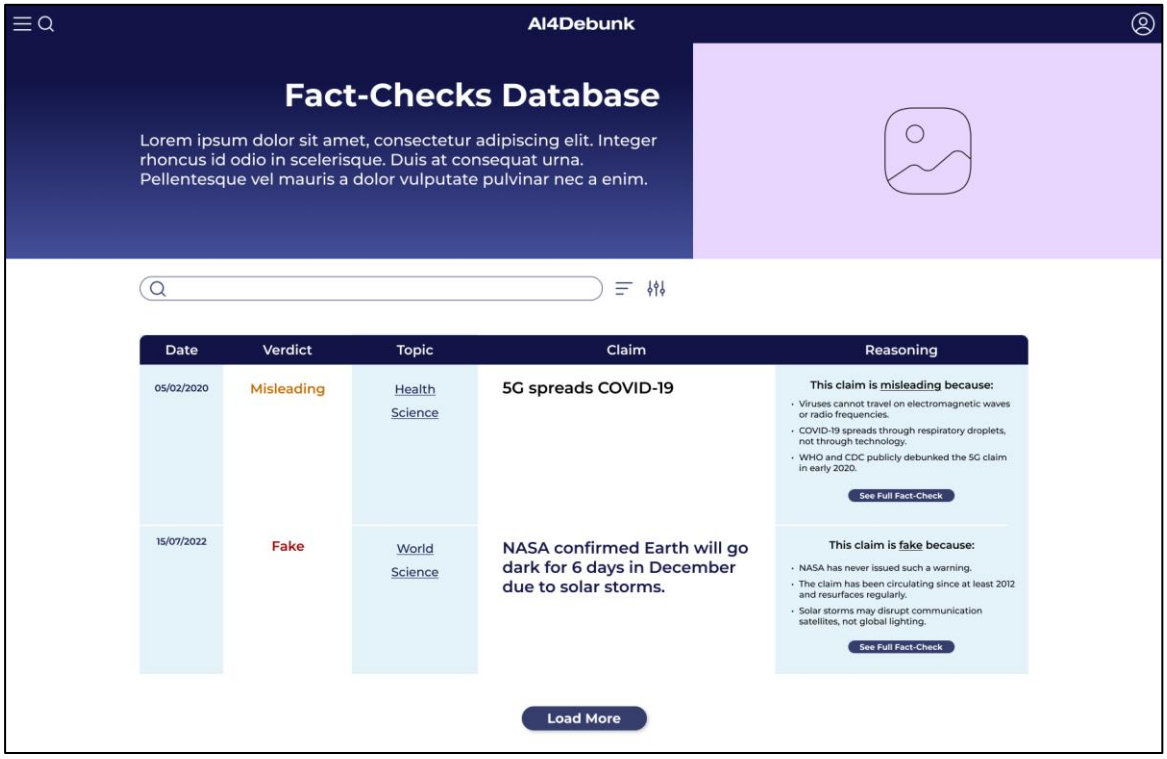


Figure 1. Example platform design.

3.1 NAVIGATION AND STRUCTURE

Navigation is based on a top-level menu and context-sensitive sidebars. The core sections are:

- Home: Overview of AI4Debunk, current statistics, and featured validated cases.
- Validate: Submission portal for users to upload or paste content for analysis.
- Explore: Searchable database of validated items and disinformation trends.
- Collaborate: Comment section, discussion space for user engagement.
- About: Information about the consortium, partners, and contact points.

Navigation adheres to standard usability heuristics (Nielsen 2020) and guarantees that no key function requires more than three user actions to reach. To clarify this principle, several representative interaction paths are described below:

Funded by the European Union

- Submitting suspicious content:
  From the landing page, a user can submit content for validation in **two actions**:
  (1) selecting *Validate* from the main navigation, and
  (2) pasting a URL or uploading content via the submission form.
  This minimises cognitive load and aligns with the usability heuristic of *efficiency of use* as defined by **Jakob Nielsen**.
- Searching validated information:
  Accessing previously validated claims requires **no more than two actions**:
  (1) selecting *Explore* from the main menu, and
  (2) entering a keyword or applying a predefined filter (e.g. topic or credibility level).
  Frequently used filters are displayed prominently to support *recognition rather than recall*.
- Accessing detailed validation results:
  From any summary result (e.g. DisinfoScore overview), users can access advanced explanations in **one additional action** by switching from *Standard mode* to *Advanced mode*.
  This progressive disclosure approach prevents information overload while maintaining transparency.
- Learning about the project and methodology:
  Information about the AI4Debunk consortium, validation methodology, and ethical safeguards is reachable in **a single action** via the *About* section, which is persistently available in the main navigation.
- Returning to the homepage or primary tasks:
  At all times, users can return to the homepage or start a new validation with **one action**, either through the logo link or the persistent *Validate* call-to-action, supporting *user control and freedom*.

3.2 VISUAL IDENTITY AND LAYOUT

The visual design adheres to the AI4Debunk brand guidelines, maintaining coherence across the WP10 interfaces. The interface aligns with WCAG standards and will receive its final implementation based on further development and the testing phase.

To have an idea of the look and feel and layout of the collaborative platform mock-ups have been created. These mock-ups will help in finalizing the product in cooperation with others. The mock-ups do not use a full grid system of standardized layouts; the final version of the collaborative platform will employ a full grid system optimized for readability and cross-platform consistency.

Funded by the European Union

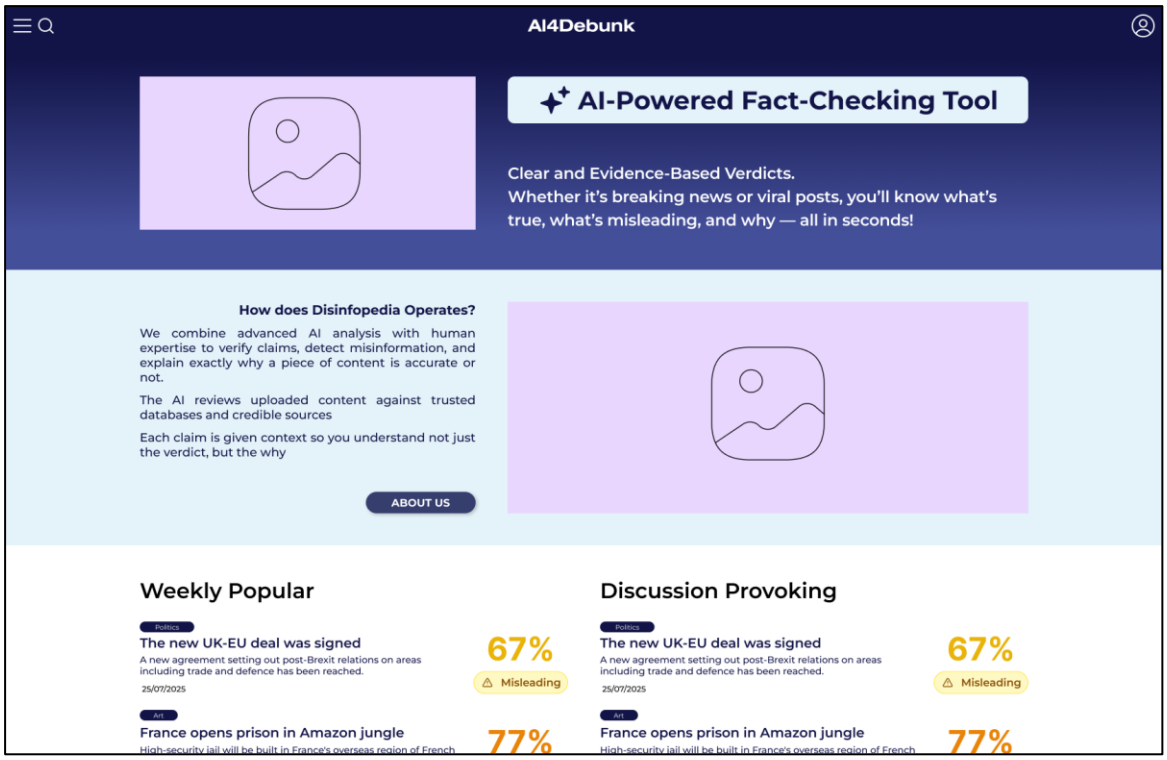


Figure 2. Example Home Page Layout

3.3 VALIDATION INTERFACE DESIGN

The validation interface is the platform's central interactive component. It visualises the results of AI disinformation detection and links them to human reviews.

The validation results are shown in multiple ways, each disclosing more information:

1. Summary View: Displays DisinfoScore using a simple visualisation, e.g. traffic-light colour code.
2. Detailed View: Reveals source traceability, linguistic indicators, and confidence metrics.
3. Contextual View: Suggests related cases, fact-checks, and source articles.

Funded by the European Union

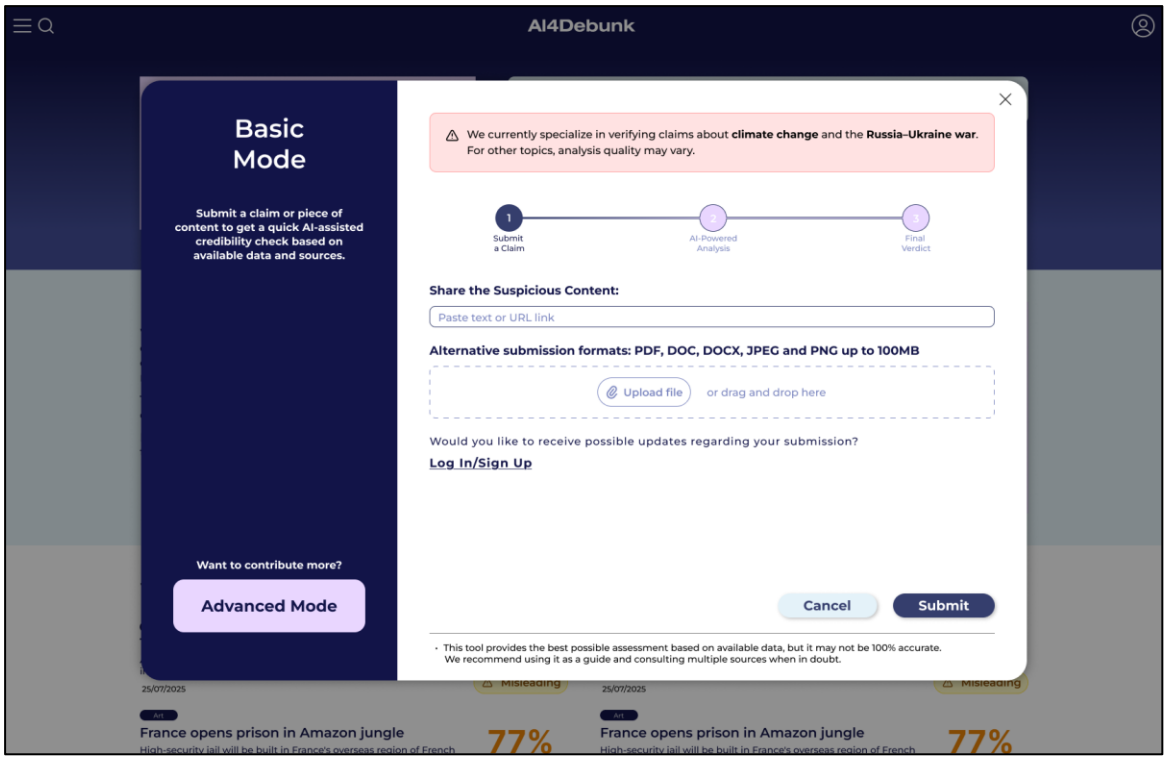


Figure 3. Preliminary Validation Interface Design

3.4 ACCESSIBILITY AND INCLUSIVITY

Accessibility has been a core design principle from the outset. The platform complies with WCAG 2.2 AA guidelines and integrates universal design features:

- Screen reader compatibility via ARIA labelling.
- Adjustable text size and high-contrast modes.
- Keyboard-only navigation and visible focus indicators.
- Multilingual content rendering and language switching.
- Simplified “readable view” for low-literacy users.

See also ANNEX 4. Key recommendations for tool developers.

Table 3. Accessibility Features and Compliance

Feature	Description	Standard
Screen Reader Support	All interactive elements labelled for accessibility APIs.	WCAG 1.3.1
Color Contrast	Minimum 4.5:1 contrast ratio ensured.	WCAG 1.4.3

Keyboard Navigation	Full functionality available without mouse.	WCAG 2.1.1
Text Resizing	Users can scale text up to 200% without loss of functionality.	WCAG 1.4.4
Multilingual Support	Content and UI are available in multiple EU languages.	EN 301 549

3.5 HUMAN-CENTRED AND PARTICIPATORY DESIGN METHODOLOGY

The design process follows the ISO 9241-210:2019 framework for human-centred design. The approach includes iterative prototyping, user co-design workshops, and pilot testing with citizen and expert participants.

Design Stages:

1. Understanding Context of Use: User research and persona development based on WP12 surveys.
2. Defining Requirements: Translating user needs into functional and quality criteria.
3. Producing Design Solutions: Developing low- and high-fidelity prototypes with feedback loops.
4. Evaluating Designs: Usability testing sessions and accessibility validation.

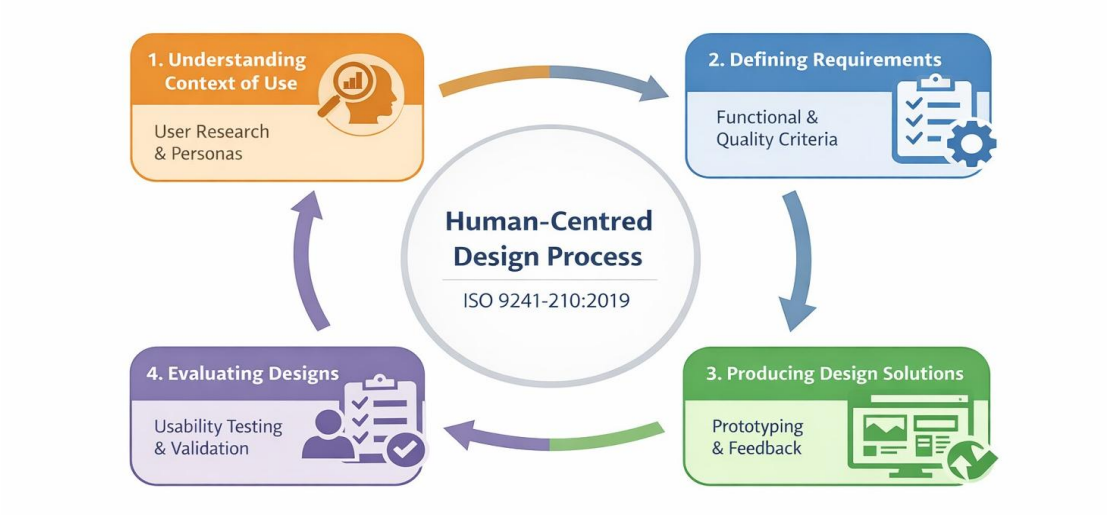


Figure 4. Human-Centred Design Iteration Cycle

3.6 INTERACTION AND FEEDBACK PRINCIPLES

The interface incorporates clear feedback mechanisms to enhance user trust:

- Confirmation dialogues for submissions and updates.

- Progress indicators during analysis.
- Transparent display of AI confidence levels.
- Moderated feedback loops for reporting incorrect classifications.

Error messages follow plain-language guidelines and include direct links to corrective actions.

Funded by the European Union

4 SYSTEM ARCHITECTURE

The Collaborative Platform architecture has been designed for robustness, scalability, and interoperability with all AI4Debunk components. It utilises a modular, service-oriented approach, integrating securely with the AI4Debunk API and underlying data layers. The architecture ensures smooth communication between the front-end, the back-end services, and the AI validation engines hosted in the consortium's "engine room."

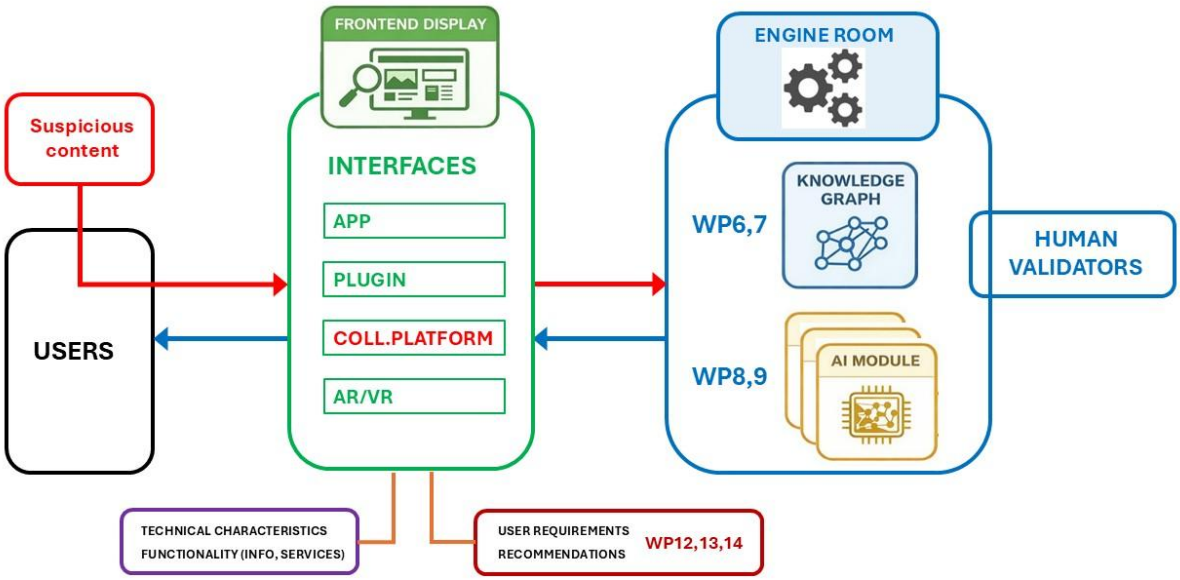


Figure 5. System Architecture Overview

4.1 KEY COMPONENTS

1. Frontend
Developed as a responsive web interface built on a modern framework, such as React or Svelte, ensuring fast load times and cross-platform compatibility. It interacts with the backend API for data retrieval and submission.

2. Backend API
Implements a RESTful API layer that manages authentication, data requests, and inter-service communication. It provides the secure bridge between the frontend and AI4Debunk detection modules.

3. Engine Room Integration
Connects to the AI4Debunk core modules (WP6–9), including the knowledge graph, machine learning classifiers.

4. Database Layer

Funded by the European Union

Stores metadata, validation results, and user interactions.

5. Security and Authentication
Single sign-on (SSO) is implemented using OAuth 2.0. All communications are encrypted (TLS 1.3) and compliant with GDPR requirements. Traditional username/password will be the default. Support for passkeys will be considered.

Table 4. System Components and Interfaces

Layer	Component	Function	
Presentation	Web Frontend	Provides UI for citizens and expert users.	API Gateway
Application	API Layer	Manages validation requests and authentication.	Engine Room
Data	Knowledge Graph / Relational Database	Stores validated cases and semantic relations.	Backend
AI Services	Detection Modules	Perform automated disinformation scoring.	API Layer
Human Oversight	Validation Committee Interface	Reviews and confirms AI-generated results.	API Layer

4.2 DATA FLOW AND COMMUNICATION

All user submissions are processed through a secure workflow:

1. User uploads content via the frontend submission interface.
2. API gateway forwards data to AI4Debunk detection modules.
3. Results are returned to the backend as a DisinfoScore package (score, explanation, confidence).
4. The backend stores the results in the knowledge graph and relational database.
5. The frontend displays the DisinfoScore and provides optional feedback channels.

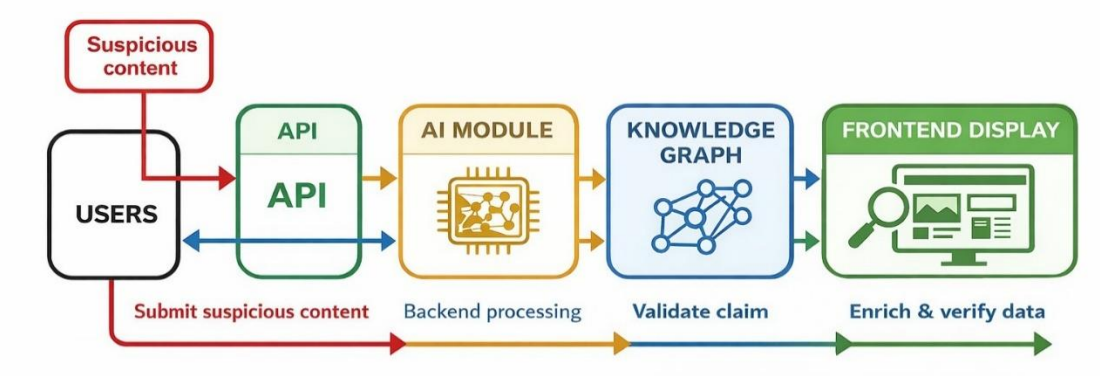


Figure 6. Data Flow Diagram

4.3 SCALABILITY AND PERFORMANCE

To ensure scalability and predictable performance, the Platform uses container-based deployments on Kubernetes, which makes it easy to scale application deployments horizontally over multiple physical machines, which allows for load to distributed evenly between different servers. Furthermore, caching mechanisms are used to avoid costly re-computation for frequently accessed data.

Using these techniques, the Platform aims for performance targets that include <500ms average API latencies and 99.9% system uptime.

Funded by the European Union

5 IMPLEMENTATION AND INTEGRATION

5.1 DEVELOPMENT PHASES

The implementation is based on continuous cooperation with relevant partners.
During the definition phase, and prior to the detailed implementation of the interface in the subsequent phase, a set of preliminary mock-ups was developed. Their purpose was to support structured discussions and alignment with project partners regarding functionality, user experience, and design choices.

- Phase 1 – Definition phase, Requirements and Prototyping (M14–M26): Consolidate requirements and build low-fidelity mock-ups.
- Phase 2 – Core Development (M26–M40): Develop the API, frontend, and initial integration with AI detection modules.
- Phase 3 – Testing and Validation (M40–M44): Conduct functional testing, security audits, and user acceptance testing.
- Phase 4 – Pilot Deployment (M44): Launch and evaluate the operational prototype (TRL7) with end-users.

Table 5. Implementation Milestones

Phase	Month	Deliverable	Description
1	M14–M26	Prototype 1	Design specifications and Interface mock-ups
2	M26–M40	Prototype 2	Final interface design, API integration and backend structure
3	M40	Beta Release	Functional testing and validation
4	M44	Final Platform	Public release of collaborative platform

5.2 TECHNOLOGY STACK

The platform leverages open-source technologies to ensure transparency and sustainability:

- The frontend will be built using industry-standard frameworks and libraries such as React or Svelte, TypeScript, and Tailwind. These technologies are widely used and supported, and likely to remain maintained in the coming years.

- The backend will be built using the popular Laravel framework, which can be run in virtually any hosting environment that supports PHP. Inertia is used to bridge the gap between the backend and the frontend to reduce duplication between the two stacks.
- The Platform stores application data in MariaDB, a widely used open-source relational database. User-uploaded files such as avatars will be S3-compatible object storage. To support quick retrieval of information, vector databases such as Meilisearch and key-value stores such as Valkey may also be used.
- Containerisation using Docker makes it easy to deploy the Collaborative Platform to Kubernetes environments.
- The Platform will support authenticating via several popular identity providers so that users can easily create and log into their accounts without going through a manual registration process.
- To ensure data sovereignty, the hosted version of the Collaborative Platform will only make use of cloud services that can guarantee that data is stored within the European Union (e.g. OVH, Hetzner, AWS, or GCP).

5.3 INTEROPERABILITY AND INTEGRATION

The Collaborative Platform achieves interoperability and integration with other applications both within and outside the AI4Debunk ecosystem through several mechanisms.

Firstly, it facilitates access to AI4Debunk's AI-powered disinformation detection modules for the browser extension and smartphone app. Secondly, it offers user interfaces that enable users to explore and retrieve information from the knowledge graph, as well as interfaces that allow members of the validation committee to perform their human-in-the-loop-oversight.

To improve findability in general-purpose search engines such as Google and Bing, and in more specialised search engines such as Google Fact Check Tools, the Collaborative Platform includes standardised semantic markup that clearly communicates the key information contained in content published on the Platform.

Funded by the European Union

6 TESTING

Validation ensures that the Collaborative Platform meets user requirements, technical specifications, and ethical standards. The testing process evaluates usability, functionality, security, and accessibility.

6.1 TESTING PLAN

Testing is divided into four categories:

- **Functional Testing:** Verifies that all core functions operate correctly.
- **Integration Testing:** Confirms interoperability between front-end, API, and AI modules.
- **Performance Testing:** Measures load capacity and response times.
- **User Acceptance Testing (UAT):** Engages pilot users to assess usability and satisfaction.

The first three categories of testing are done continuously throughout the development cycle using automated tools to prevent regression of previously developed features, while user acceptance testing is only done at specific moments with pilot users who may be able to spot issues that have been overlooked using automated methods.

Table 6. Validation Metrics

Category	Metric	Target
Usability	SUS score	≥72
Accessibility	WCAG 2.2 AA compliance	100%
Response Time	API latency	<500ms

6.2 ETHICAL AND PRIVACY CONSIDERATIONS

The platform adheres to the AI4Debunk Ethical Framework established in WP2, ensuring alignment with the EU AI Act and GDPR. Key ethical principles include:

- Transparency: Users are informed of AI involvement in validation.
- Human Oversight: Validation Committee maintains ultimate decision-making authority.
- Data Protection: All personal data is pseudonymised; no profiling is conducted.
- Fairness: Bias detection procedures are implemented for AI-generated outcomes.

Funded by the European Union

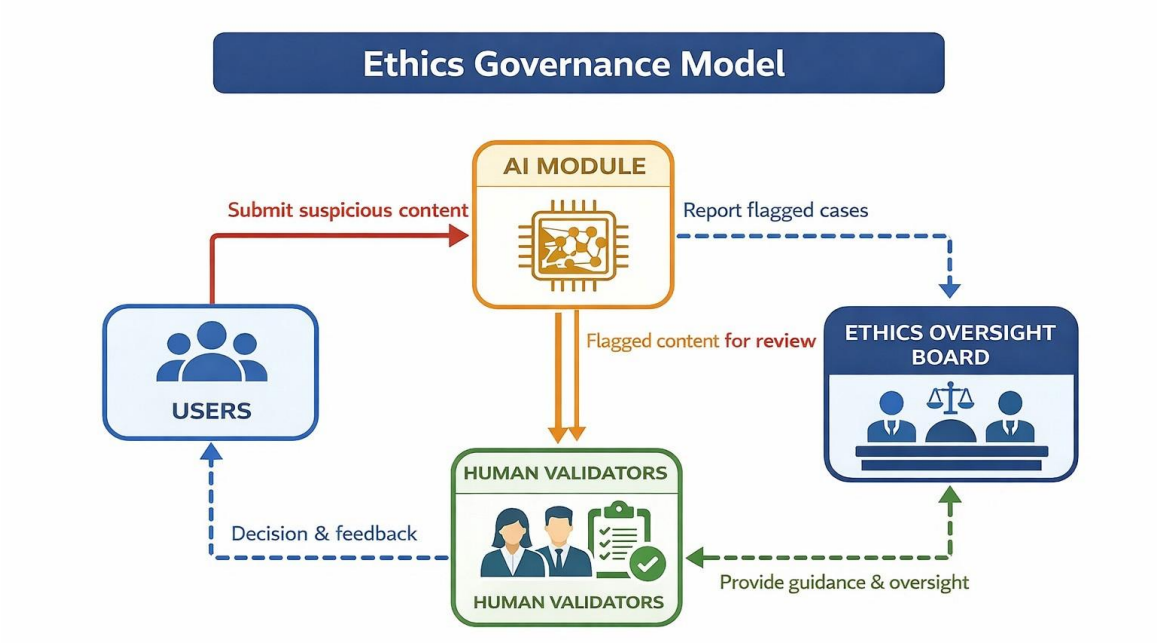


Figure 7. Ethics Governance Model

6.3 COMPLIANCE WITH EUROPEAN STANDARDS

The platform is designed to comply with:

- OWASP ASVS Level 1
- ISO 9241-210:2019 – Human-Centred Design
- EN 301 549:2021 – Accessibility Requirements for ICT
- GDPR (Regulation EU 2016/679) – Data Protection

A combination of static analysis, heuristics, and manual review will be used to ensure compliance with these standards.

7 CONCLUSIONS AND FUTURE WORK

7.1 SUMMARY OF ACHIEVEMENTS

The AI4Debunk Collaborative Platform delivers a robust, participatory, and secure environment for analyzing and debunking online disinformation. It has been designed to operate as the central knowledge hub of the AI4Debunk ecosystem, integrating the outputs from the browser extension, mobile app, and AR/VR interface.

The report has defined:

- A complete set of functional and non-functional requirements.
- A modular, scalable system architecture;
- Integration of AI-driven and human validation workflows;
- A strong emphasis on inclusivity, accessibility, and ethical compliance.

7.2 NEXT STEPS

Following the completion of this deliverable, the next stages include:

- Conducting the final usability testing with end-user communities;
- Integrating results from WP7 and WP9 (AI engines) into the operational prototype;
- Launching the pilot deployment by M26;
- Evaluating performance, engagement, and impact metrics;
- Preparing for post-project sustainability and open-source release.

7.3 LONG-TERM VISION

Beyond the project, the Collaborative Platform aims to evolve into a European reference infrastructure for combating disinformation. Its open and extensible design allows integration with fact-checking organizations, academic partners, and public institutions.

Funded by the European Union

8 REFERENCES

ISO standards

International Organization for Standardization. (2019).
ISO 9241-210:2019 – Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems.
https://www.iso.org/standard/77520.html

International Organization for Standardization & International Electrotechnical Commission. (2022).
ISO/IEC 27001:2022 – Information security, cybersecurity and privacy protection — Information security management systems — Requirements.
https://www.iso.org/standard/27001

EU legislation and policy documents

European Commission. (2019).
Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services (European Accessibility Act).
Publications Office of the European Union.
https://eur-lex.europa.eu/eli/dir/2019/882/oj

European Union. (2016).
Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation).
Official Journal of the European Union.
https://eur-lex.europa.eu/eli/reg/2016/679/oj

European Commission. (2021).
Ethics guidelines for trustworthy AI.
Publications Office of the European Union.
https://digital-strategy.ec.europa.eu

Web standards and guidelines

World Wide Web Consortium. (2023).
Web Content Accessibility Guidelines (WCAG) 2.2.
https://www.w3.org/TR/WCAG22/

Professional and industry resources

Nielsen, J. (2020).
10 usability heuristics for user interface design. Nielsen Norman Group.
https://www.nngroup.com/articles/ten-usability-heuristics/

OWASP Foundation. (2023).
Application Security Verification Standard (ASVS).
https://owasp.org/www-project-application-security-verification-standard/

Academic literature

Schmid, S., *et al.* (2023).
Human-in-the-loop AI for disinformation detection. *AI & Society*.
https://doi.org/10.1007/s00146-023-01689-0

Funded by
the European Union

9 ANNEX 1 – MOCK-UP DESCRIPTIONS

The platform mock-ups have been created to visualize the user flow and to enable testing the preliminary interface design. They serve as a communication tool between designers, developers, and stakeholders.

9.1 MOCK-UP ELEMENTS

Home Page: Displays navigation, live validation statistics, and featured disinformation cases.

Validation Interface: Central dashboard for submitting and reviewing content, integrating AI-generated DisinfoScores and feedback loop.

Search and Explore Section: Enables filtering by topic, credibility level, and date.

Collaboration Element: Structured comment system moderated by experts.

About us section: Provides overview of the project's mission, consortium partners, and underlying values.

Blog section: Provides platform updates, insights, and short posts that support knowledge-sharing.

Funded by the European Union

9.2 VISUAL GUIDANCE

Mock-ups follow the AI4Debunk color scheme and typography. Each element complies with accessibility requirements for color contrast and font legibility.

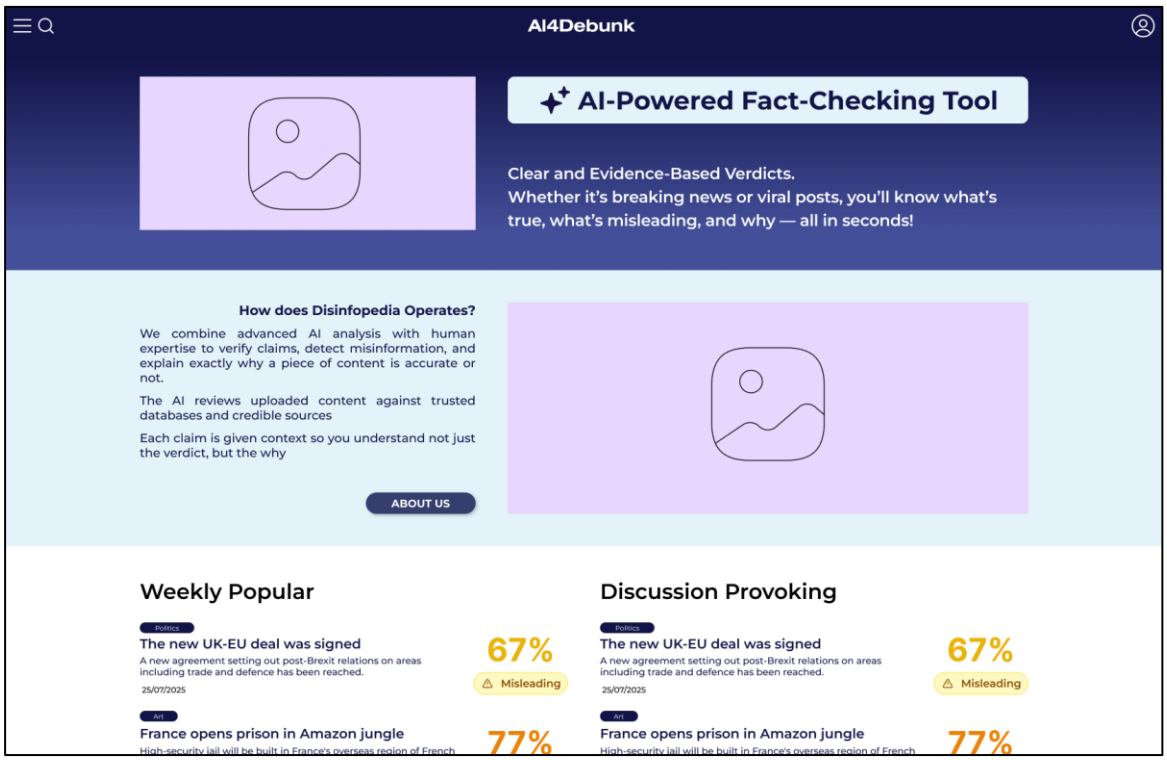


Figure A1.1: Conceptual interface representation of the Collaborative Platform homepage

Funded by the European Union

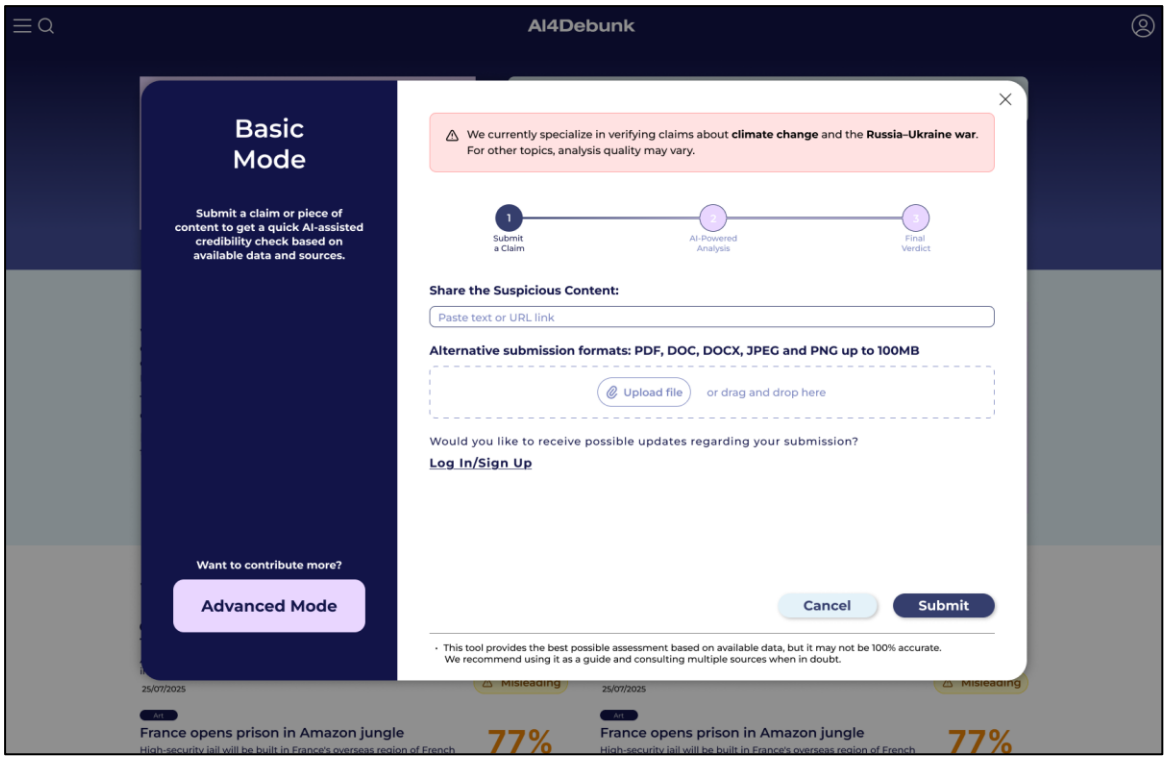


Figure A1.2: Conceptual interface representation of the content validation for general users

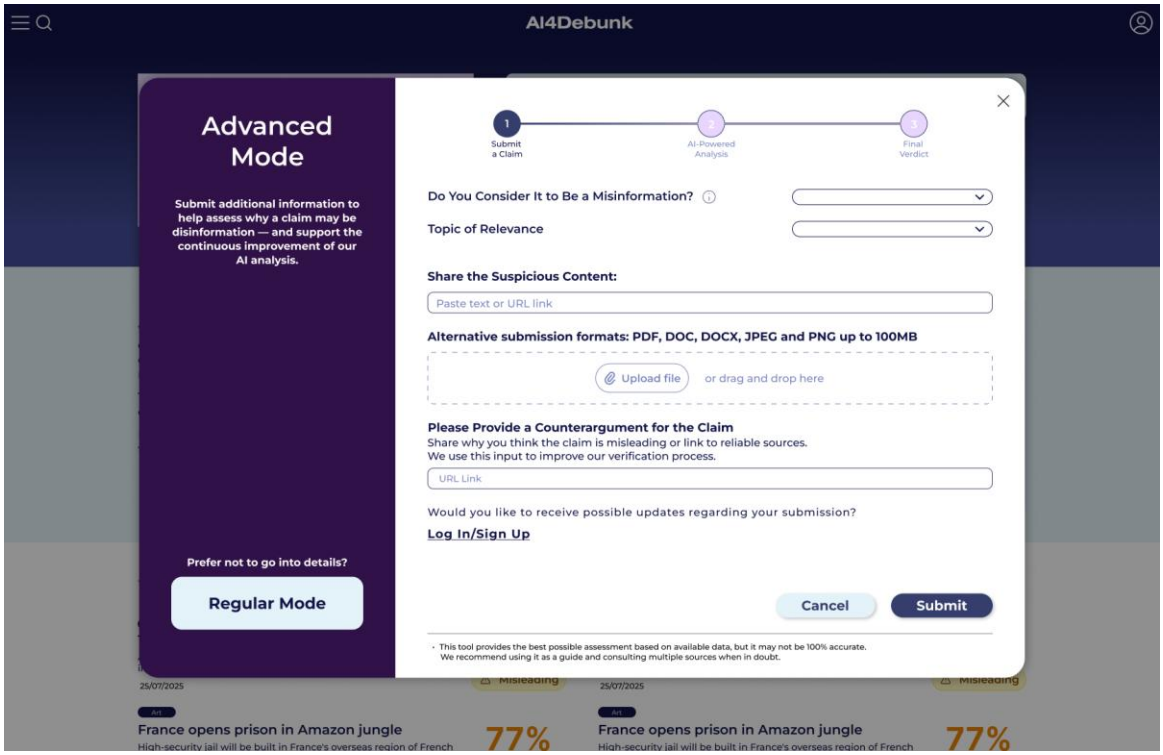


Figure A1.3: Conceptual interface representation of the extended content validation for advanced users

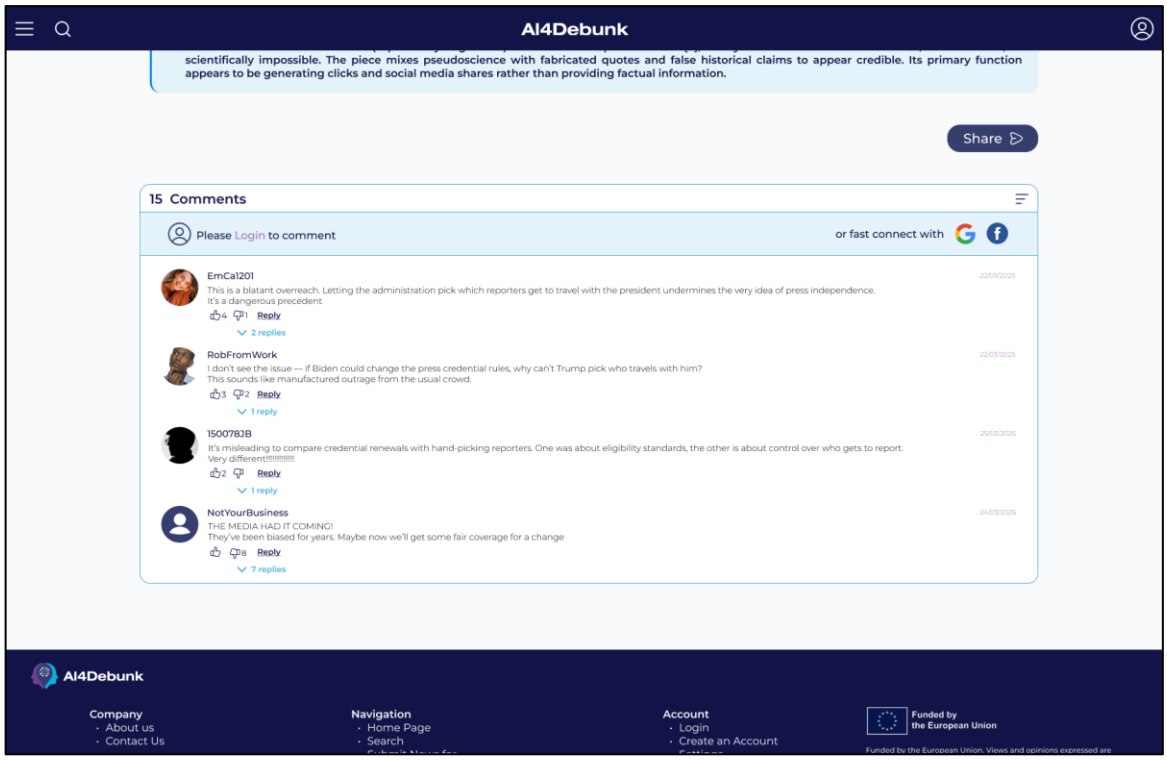


Figure A1.4: Conceptual interface representation of the collaborative discussion forum

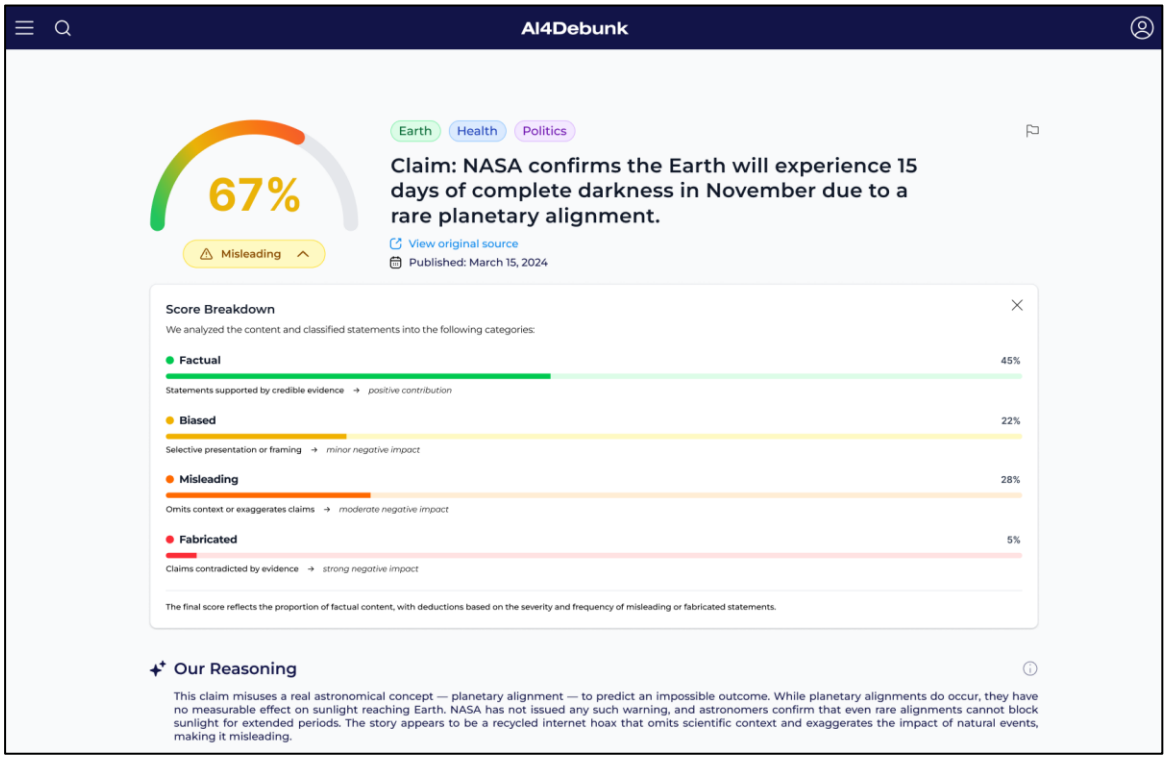


Figure A1.5: Conceptual interface representation of the fact-checking output with progressive disclosure

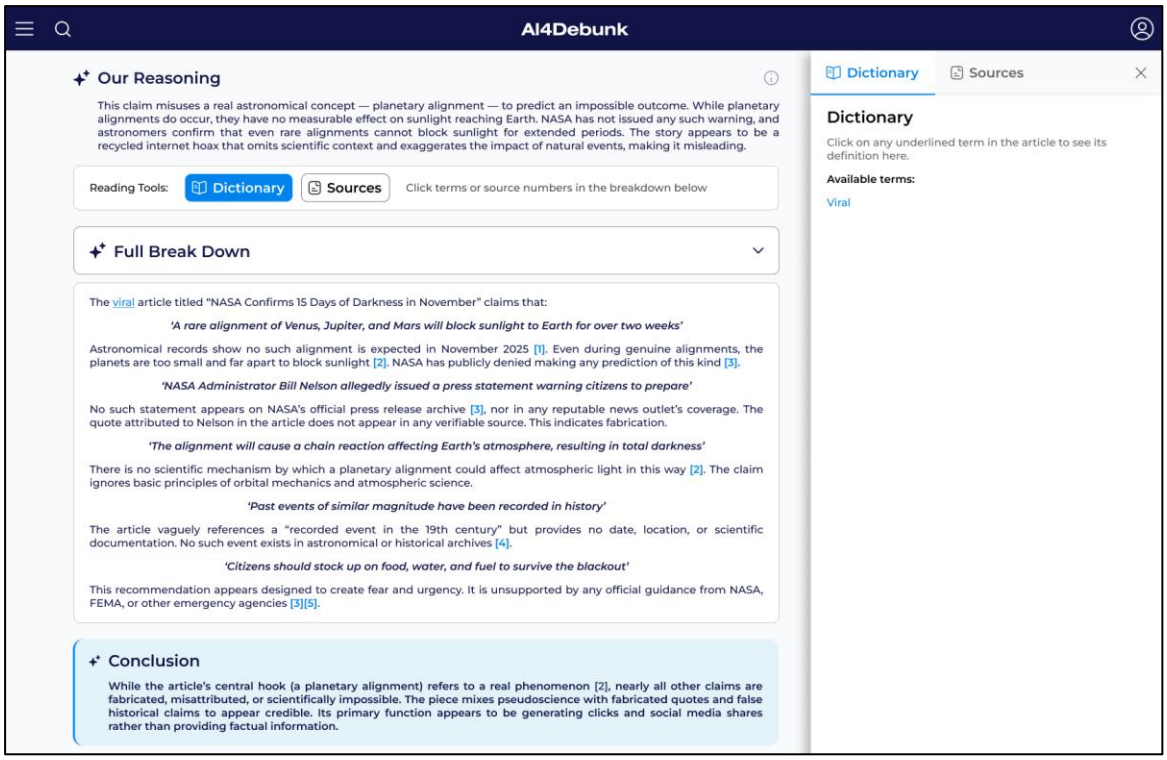


Figure A1.6: Conceptual interface representation of the fact-checking output with dictionary support

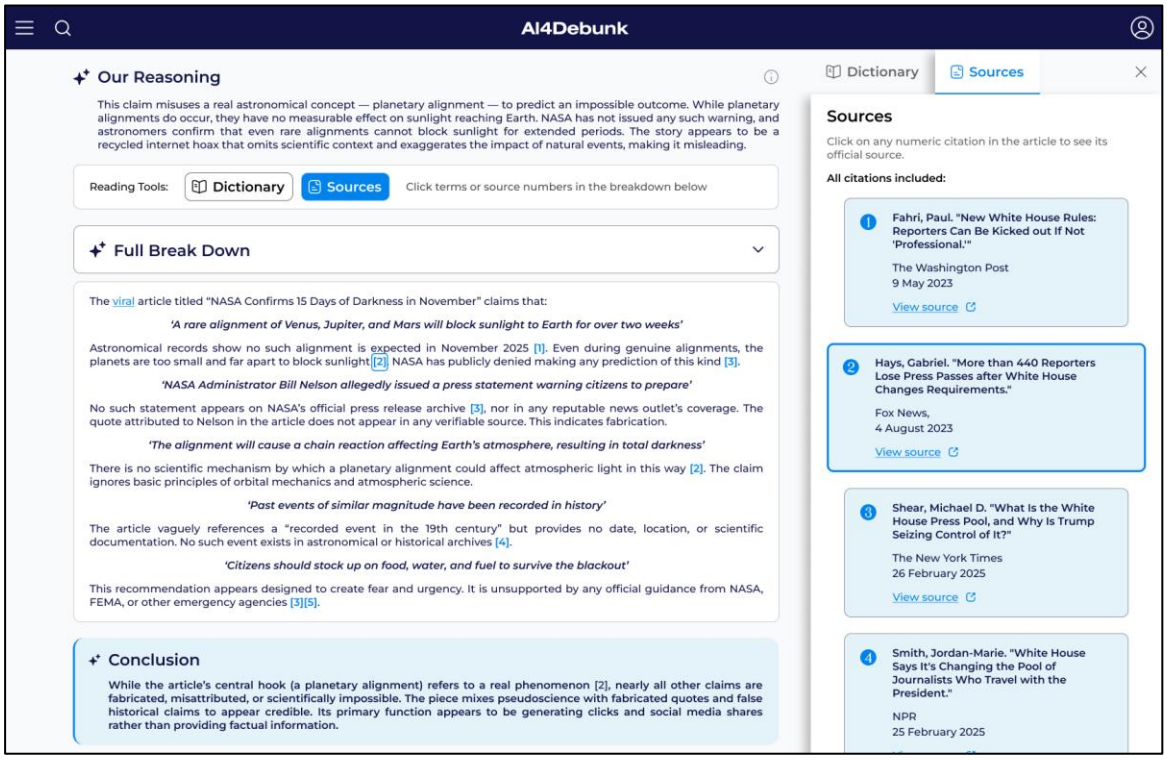


Figure A1.7: Conceptual interface representation of the fact-checking output with source attribution

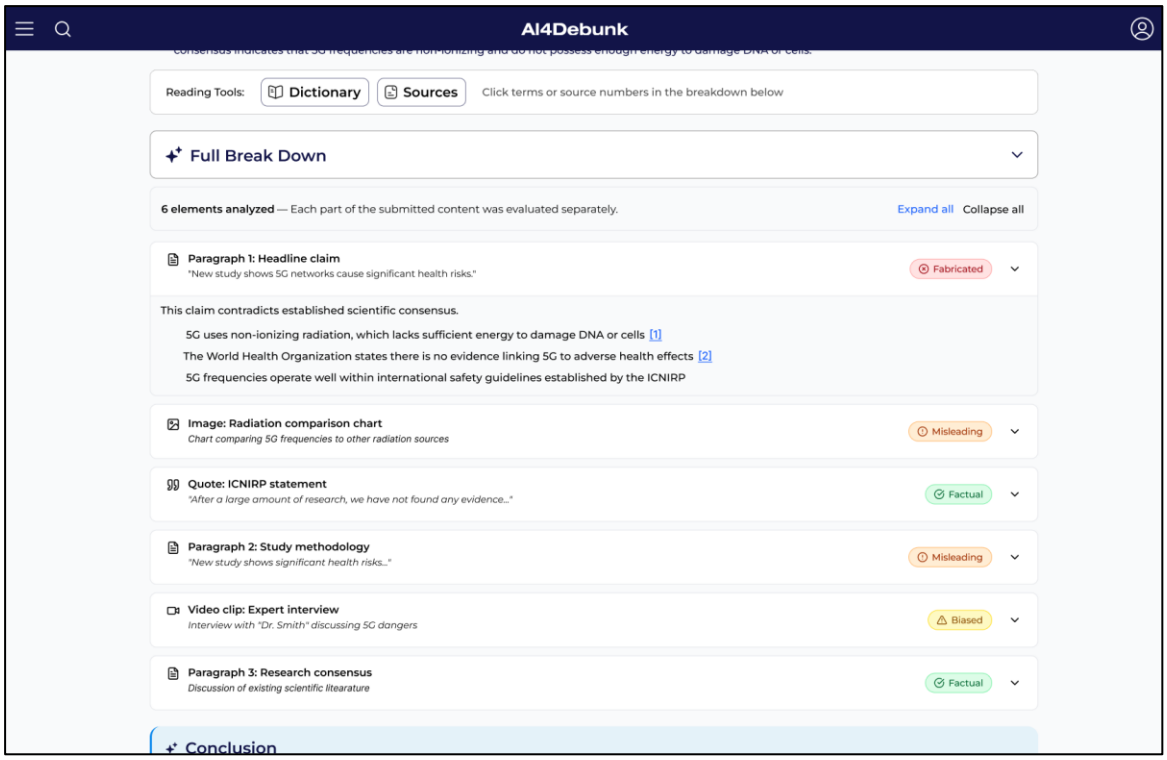


Figure A1.8: Conceptual interface representation for fact-checking outputs containing multiple content elements

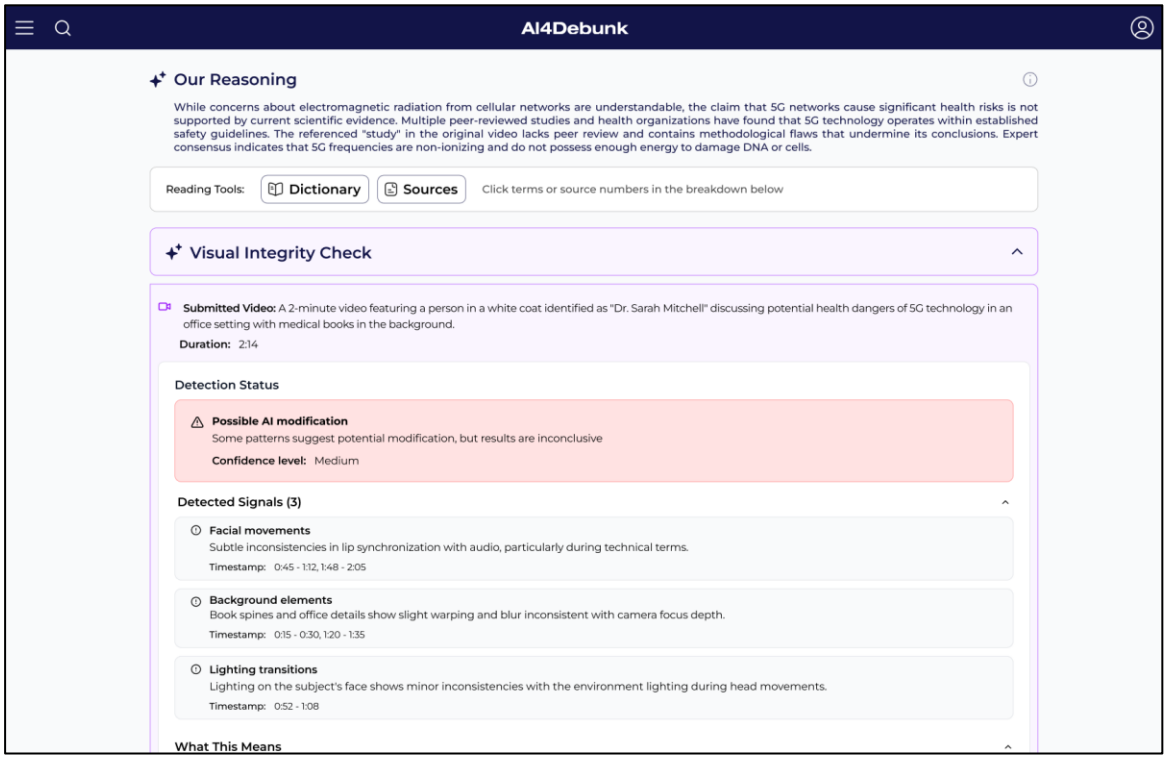


Figure A1.9: Conceptual interface representation for fact-checking outputs of visual content

Funded by the European Union

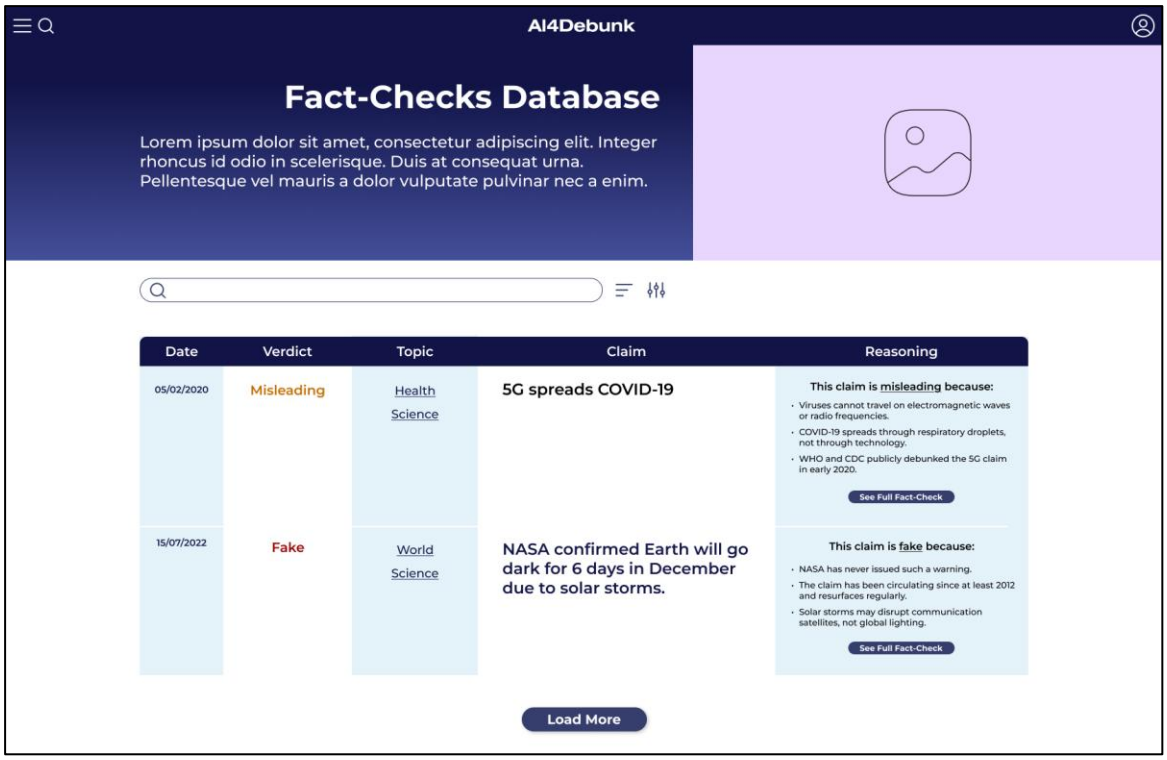


Figure A1.10: Conceptual interface representation of the platform knowledge database

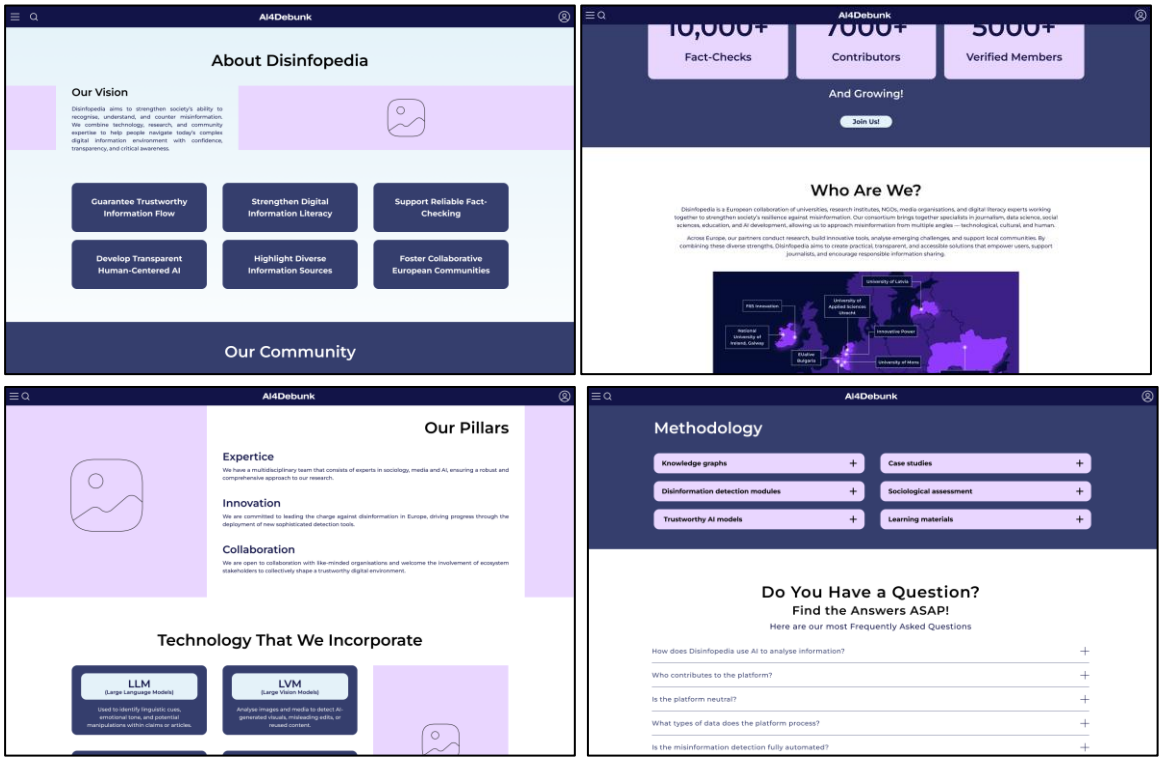


Figure A1.11: Conceptual interface representation of the "About Us" page

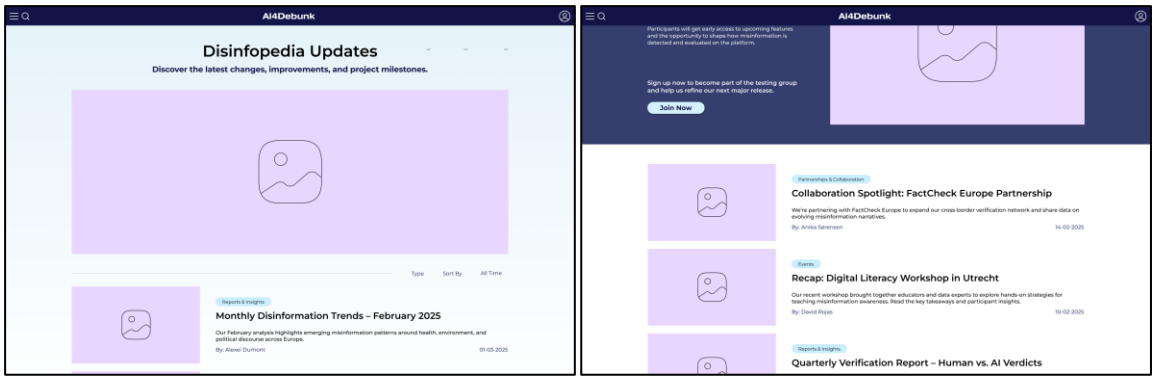


Figure A1.12: Conceptual interface representation of the platform blog overview

Figure A1.13: Conceptual interface representation of the report update interface

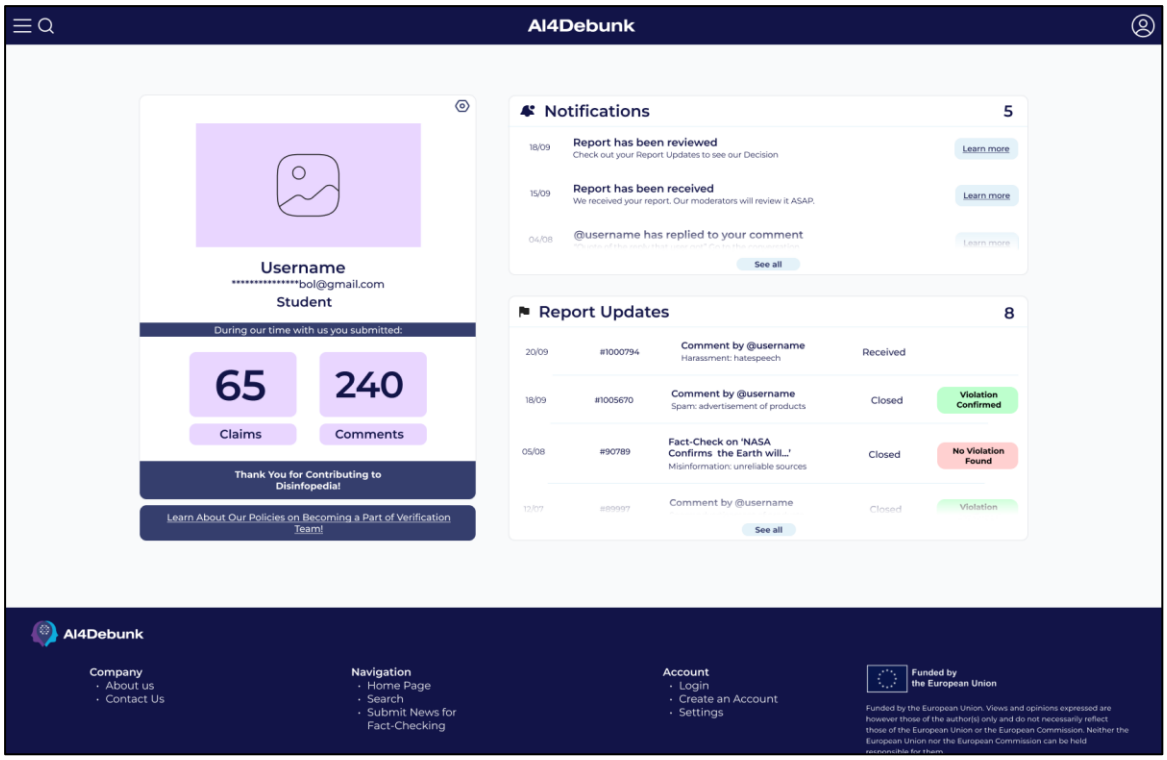


Figure A1.14: Conceptual interface representation of the user profile dashboard

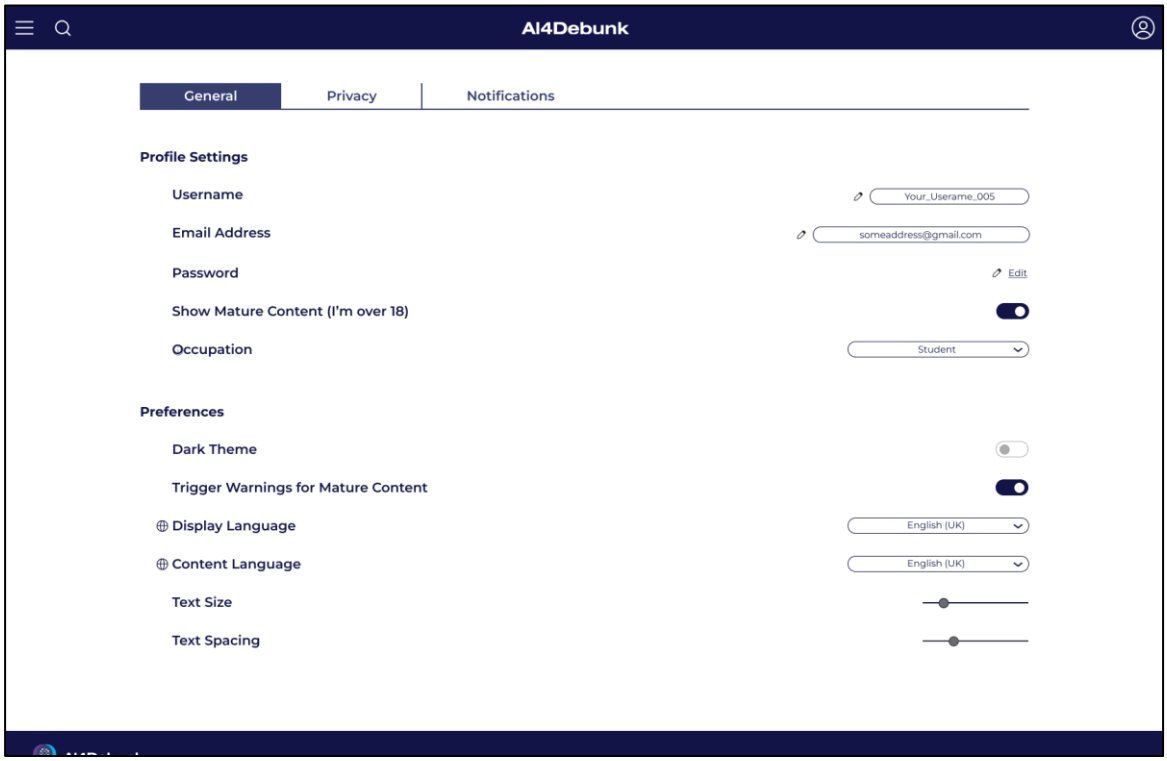


Figure A1.15: Conceptual interface representation of the user settings interface

10 ANNEX 2 – USER TESTING AND EVALUATION PLAN

User testing aims to validate usability, accessibility, and trust in AI-assisted disinformation detection. It ensures the platform meets the expectations of different user groups.

10.1 METHODOLOGY

Testing follows a mixed-methods approach combining quantitative metrics (System Usability Scale – SUS, task completion rates) and qualitative feedback (interviews, think-aloud sessions).

10.2 PARTICIPANTS

For qualitive feedback a balanced sample of 30–50 participants will be selected across four groups: 1. Citizens (general users) 2. Journalists and fact-checkers 3. Researchers and educators 4. Policy makers and NGO representatives.

10.3 TEST ENVIRONMENT

All tests will be performed using the near-final prototype in a controlled digital environment. Accessibility tools (screen readers, magnifiers) will be available.

10.4 EVALUATION METRICS

Table A2.1: Evaluation metrics

	Metric	**Target**
Usability	SUS score	≥72
Accessibility	WCAG 2.2 AA compliance	100%
Task Efficiency	Average task completion time	<3 min
Trust	User-perceived transparency (survey)	>80% positive
Engagement	Repeat visits during pilot	>60%

10.5 REPORTING

Testing results will be summarized in a “Pilot Evaluation Report” and integrated into the final technical validation documentation.

Funded by the European Union

11 ANNEX 3 – API OVERVIEW AND DATA MODEL

The Collaborative Platform connects to the central AI4Debunk API, which provides standardized endpoints for validation requests, retrieval of results, and interaction with the knowledge graph.

11.1 DATA MODEL

From the perspective of the Platform, the API is only concerned with providing access to the knowledge graph and WP6–9 debunking modules. Data relating to user accounts, user-generated content, and validation metadata are owned exclusively by the Collaborative Platform and are not shared with the API. A copy of validation results is stored in the Platform, which preserves the possibility of decoupling the Platform from the current AI4Debunk API instance should this be necessary in the future.

Concretely, this means that the Collaborative Platform will store the following data:

- User account data, including personally identifiable information such as usernames and email addresses.
- Usage-related information, including the number of validation requests a user has created, account preferences, and a pseudonymous identifier used to communicate with the debunking API without directly exposing the user's real identity.
- User-generated content such as comments, votes, and other data related to user actions that are necessary to implement required functionality, such as whether a user has provided consent for data collection or has read reports and notifications.
- The contents of validation requests and reports, which include metadata about submitted files and URLs, and the Disinfoscore together with accompanying validation results and explanations.

Funded by the European Union

12 ANNEX 4 – KEY RECOMMENDATIONS FOR TOOL DEVELOPERS

The work done in WP10 is preceded by several social sciences and humanities (SSH) work packages, which produced key recommendations for developers of tools to combat disinformation. To ensure that these recommendations are properly integrated into the design and development of human-centred user interfaces, the list below was originally compiled by Pascaline Gaborit (Pilot4DEV) and later organised and numbered by Marcel Keijzer (IP). These recommendations were subsequently discussed at various consortium meetings. This appendix presents all recommendations in their original form.

---

1. User-centric design
   - 1.1. Design intuitive interfaces that make navigation and interpretation of data simple for both technical and non-technical users (including young and elderly people). Include tooltips, visual aids (graphs, videos).
   - 1.2. Enable customization: Users should tailor analysis depth, alerts, and display preferences. Provide “basic” and “advanced” modes to cater to diverse audiences.
   - 1.3. Provide comprehensive user support through integrated FAQs, contextual help, and multimedia learning aids (tutorials, videos, interactive walkthroughs).
2. Gender equality and inclusivity
   - 2.1. Ensure gender-neutral design and eliminate bias in algorithms and datasets. Conduct regular gender audits to test for bias in detection outcomes.
   - 2.2. Also make sure that all vulnerable groups can access the tool and interface (people living in remote areas or in the countryside, elderly people, people from different minorities, e.g. Russian speaking diaspora but not only…)
   - 2.3. Collaborate with gender experts and advocacy groups during design and testing to ensure the tool serves all genders equitably.
   - 2.4. Support inclusivity: Offer multi-language interfaces, culturally neutral symbols, and accessibility features (screen readers, keyboard navigation, voice commands).
   - 2.5. Comply with global accessibility standards and optimize tools for low-tech or low-bandwidth environments.
   - 2.6. Use region-specific data and partner with local experts to make tools sensitive to cultural and linguistic disinformation trends.
3. Integration with social media
   - 3.1. Implement API-based real-time analysis for major platforms (Facebook, X/Twitter, Bluesky, TikTok, YouTube, Instagram).
   - 3.2. Ensure cross-platform tracking to trace how disinformation spreads across ecosystems.
   - 3.3. Access metadata (timestamps, geolocation, engagement) to understand virality and amplification dynamics while maintaining GDPR-compliant privacy protections.
   - 3.4. Monitor emerging and niche platforms (e.g., Telegram, Gab, Parler) and adapt to new platform features like short-lived content or encrypted messaging.
4. Tackling coordinated inauthentic behaviour (CIB)

Funded by the European Union

4.1. Provide real-time alerts and dashboards that visualize spikes in engagement, bot activity, or suspicious hashtag use.
4.2. Educate users through built-in guidance on recognizing coordinated behaviour (e.g., identical posts, synchronized activity).
4.3. Integrate network visualization tools to help users identify clusters and influencers driving disinformation.
4.4. Allow user-driven monitoring and flagging, enabling journalists and analysts to investigate specific narratives and provide corrective feedback.

5. Ethical and transparent AI
   5.1. Prioritize data privacy and user consent. Collect and store only necessary public data, ensuring transparency about use and limitations.
   5.2. Avoid algorithmic bias by maintaining diverse, balanced datasets and documenting model decision-making processes.
   5.3. Provide transparency in flagging: Explain why content is flagged and offer mechanisms for users to dispute or verify flagged material.
   5.4. Establish ethical safeguards to prevent misuse of detection tools for censorship or political manipulation.
6. Multilingual and cultural adaptation
   6.1. Train NLP models per language to recognize regional slang, idioms, and context-sensitive disinformation patterns.
   6.2. Enable cross-language analysis to track narrative migration across languages and regions.
   6.3. Use culturally aware machine translation for languages without native models, ensuring meaning and intent are preserved.
7. Stakeholder and expert engagement
   7.1. Establish regular consultation mechanisms (workshops, focus groups, advisory boards) with journalists, researchers, fact-checkers, and policymakers.
   7.2. Run diverse beta testing programs involving participants from different cultural, linguistic, and professional backgrounds.
   7.3. Continuously integrate feedback from stakeholders into iterative tool updates.
8. Continuous improvement
   8.1. Update algorithms regularly to reflect new disinformation tactics (deepfakes, AI-generated content, bot evolution).
   8.2. Expand and diversify datasets using regional and topical sources, verified fact-checking data, and emerging content types.
   8.3. Maintain feedback loops: allow users to report false positives or missed detections and receive updates on actions taken.
   8.4. Collaborate with experts in AI, disinformation, and digital ethics to ensure tools evolve responsibly and effectively.
9. Overall recommendation
   9.1. Developers should approach disinformation detection as a living process, emphasizing adaptability, transparency, inclusivity, and collaboration. By integrating ethical AI design, user engagement, and continuous learning, tools can remain robust against evolving disinformation landscapes and build user trust globally.

Funded by the European Union

13 ANNEX 5 – FUNCTIONAL REQUIREMENTS BREAKDOWN

FR-1 Submission		
Allow users to submit content (URLs, text, images, videos) for disinformation analysis.		
FR-1.1	Must have	The system provides a “Submit content” entry point that is visible on the main navigation.
FR-1.2	Must have	Users can open a submission form without being authenticated.
FR-1.3	Must have	Authenticated users can open a submission form from their account area.
FR-1.4	Must have	The submission form allows users to select a content type (URL, text, image, video).
FR-1.5	Must have	Users can enter a URL in a dedicated input field.
FR-1.6	Must have	The system validates that the URL is syntactically valid and shows an error if it is not.
FR-1.7	Must have	Users can paste or type free-text content into a text area.
FR-1.8	Must have	Users can upload one or more image files, subject to allowed file types and size limits.
FR-1.9	Must have	Users can upload one or more video files, or provide a video link, subject to size/type rules.
FR-1.10	Must have	Users can optionally provide metadata (e.g. title, language, topic) for a submission.
FR-1.11	Must have	The system validates that all required fields for the chosen content type are completed.
FR-1.12	Must have	The system shows clear error messages if required fields are missing or invalid.
FR-1.13	Must have	When the user submits the form, the system creates a new analysis request in the backend.
FR-1.14	Must have	The system associates the submission with the authenticated user, if logged in.
FR-1.15	Must have	The system stores the submission timestamp.
FR-1.16	Must have	After successful submission, the system shows a confirmation message to the user.
FR-1.17	Must have	The system assigns a unique identifier to each submitted item.
FR-2 Validation		
Display automated DisinfoScore results derived from AI4Debunk detection modules.		
FR-2.1	Must have	For each submitted item, the system requests a DisinfoScore from the AI4Debunk backend.
FR-2.2	Must have	The system polls or listens for completion of the automated analysis.
FR-2.3	Must have	While analysis is pending, the system shows the user a “pending / in progress” status.
FR-2.4	Must have	When available, the system stores the DisinfoScore result with the submitted item.

Funded by the European Union

FR-2.5	Must have	The system displays the DisinfoScore value on the item detail page.
FR-2.6	Must have	The system displays the DisinfoScore.
FR-2.7	Must have	The system records the timestamp at which the DisinfoScore was generated.
FR-2.8	Must have	If AI analysis fails, the system shows an error state and does not display a misleading score.
FR-3 Explainability		
Provide transparent explanations for each validation, including source provenance and AI confidence levels.		
FR-3.1	Must have	The system displays a human-readable summary explaining why the content received that score.
FR-3.2	Must have	The system displays the overall AI confidence level (e.g. percentage or qualitative label).
FR-3.3	Must have	The system displays confidence levels for individual AI modules when provided.
FR-3.4	Must have	The system lists the main evidence factors (e.g. linguistic cues, source patterns) used.
FR-3.5	Must have	The system shows provenance information for sources used in the analysis (e.g. URLs, outlets).
FR-3.6	Must have	Provenance entries link to the original source in a new tab or window.
FR-3.7	Must have	The system displays the date/time at which the explanation was generated.
FR-3.8	Must have	Users can switch between a simplified explanation view and an advanced explanation view.
FR-3.9	Must have	In advanced view, users can expand sections to see more detailed technical indicators.
FR-4 User collaboration		
Enable users to comment, suggest corrections, and discuss results under moderation.		
FR-4.1	Must have	Unauthenticated users can view the list of comments associated with an item.
FR-4.2	Must have	For each comment, the system displays the commenter's display name.
FR-4.3	Must have	For each comment, the system displays the comment text.
FR-4.4	Must have	For each comment, the system displays the date and time of posting.
FR-4.5	Must have	For each comment, the system displays whether the comment was edited.
FR-4.6	Must have	For edited comments, the system displays the date and time of the last edit.
FR-4.7	Must have	Authenticated users can submit a new comment on an item.
FR-4.8	Must have	When submitting a comment, the system validates that the comment text is not empty.
FR-4.9	Must have	Authenticated users can edit the text of their own comments.

Funded by the European Union

FR-4.10	Must have	When a comment is edited, the system stores the updated text and timestamp.
FR-4.11	Must have	Authenticated users can delete their own comments.
FR-4.12	Must have	When a comment is deleted, the system hides it from regular comment listings.
FR-4.13	Must have	Authenticated users can upvote a comment once.
FR-4.14	Must have	Users can retract or change their upvote, if allowed by the design.
FR-4.15	Must have	The system displays the total number of upvotes per comment.
FR-4.16	Must have	Authenticated users can report a comment for moderation.
FR-4.17	Must have	When reporting a comment, users can select a reason from a predefined list.
FR-4.18	Must have	The system records reported comments in a moderation queue.
FR-4.19	Must have	Moderators can mark reported comments as reviewed.
FR-5 Search & filter		
Allow full-text and metadata-based search for validated items, using filters for topic, language, and credibility level. (please note: this depends on the structure and design of the database)		
FR-5.1	Must have	Users can access a dedicated “Explore / Search” page.
FR-5.2	Must have	Users can enter free-text queries to search validated items.
FR-5.3	Must have	The system performs full-text search over item titles.
FR-5.4	Must have	The system performs full-text search over item content or claim text (where available).
FR-5.5	Must have	The system allows users to filter search results by topic.
FR-5.6	Could have	The system allows users to filter search results by language.
FR-5.7	Must have	The system allows users to filter search results by credibility level (e.g. verified, doubtful).
FR-5.8	Must have	The system allows users to filter results by submission date.
FR-5.9	Must have	Users can sort search results e.g. by date, relevance, DisinfoScore severity (if available).
FR-5.10	Must have	Search results display key metadata (title, short summary, date, DisinfoScore).
FR-5.11	Must have	Users can click a search result to navigate to the detailed validation view.
FR-5.12	Must have	The system paginates search results when there are more than one page of hits.
FR-5.13	Must have	Filters and search terms remain visible so users can adjust them.
FR-6 Educational content		
Host guides, tutorials, and interactive lessons on detecting misinformation.		
FR-6.1	Should have	The platform provides a dedicated “Learn / Educational content” section.
FR-6.2	Should have	Users can see a list of available guides, tutorials, and lessons.

Funded by the European Union

FR-6.3	Should have	Each educational item displays a title, short description, and duration/level if available.
FR-6.4	Should have	Users can open an educational item to view its full content.
FR-6.5	Should have	The system supports embedding multimedia content (images, video, interactive elements).
FR-6.6	Should have	Users can filter educational content by topic.
FR-6.7	Should have	Users can filter educational content by difficulty level (e.g. basic, advanced), if defined.
FR-6.8	Should have	Users can search within the educational section by keywords.
FR-6.9	Should have	Authenticated users can mark an educational item as “completed” or “in progress”.
FR-6.10	Should have	The system stores a user’s completion status per educational item.
FR-6.11	Should have	The system may suggest related educational items based on the current page.
FR-7 Integration		
Provide direct access and data sharing with the browser extension and the mobile app.		
FR-7.1	Must have	The platform exposes an API endpoint for receiving validation requests from the browser extension.
FR-7.2	Must have	The platform exposes an API endpoint for receiving validation requests from the mobile app.
FR-7.3	Must have	The platform authenticates incoming requests from the browser extension.
FR-7.4	Must have	The platform authenticates incoming requests from the mobile app.
FR-7.5	Must have	The platform returns DisinfoScore and explanation data in a standardised format to all clients.
FR-7.6	Must have	Users can open a submitted item in the collaborative platform from the browser extension.
FR-7.7	Must have	Users can open a submitted item in the collaborative platform from the mobile app.
FR-7.8	Must have	The system synchronises validation status across all interfaces for the same item ID.
FR-8 Downloads		
Offer public access to AI4Debunk software components and open datasets.		
FR-8.1	Should have	The platform provides a “Downloads / Resources” section accessible from navigation.
FR-8.2	Should have	The system lists available software components (e.g. code repositories, binaries).
FR-8.3	Should have	The system lists available open datasets related to AI4Debunk.
FR-8.4	Should have	Information about downloadable item (e.g. name, version, short description, size, terms of use etc.) can be found on the platforms where they are available (like Github and Zenodo).

Funded by the European Union

FR-9 Notifications		
Notify registered users of updates or new validations in their area of interest. (Please note: this depends on the structure and design of the knowledge graph).		
FR-9.1	Could have	Registered users can access a notification preferences page.
FR-9.2	Could have	Users can opt in or out of receiving notifications.
FR-9.3	Could have	Users can select topics or themes for which they want to receive updates.
FR-9.4	Could have	Users can select preferred languages for notifications.
FR-9.5	Could have	Users can choose notification channels (e.g. email, in-platform notifications), if available.
FR-9.6	Could have	The system triggers a notification when a new validation is published matching user interests.
FR-9.7	Could have	Notifications include a short summary and a link to the validated item.
FR-9.8	Could have	Users can view a history of recent notifications in their account area.
FR-9.9	Could have	Users can unsubscribe from all notifications with a single action.
FR-9.10	Could have	The system respects user notification preferences when sending any communication.

Funded by the European Union

GLOSSARY OF TERMS

Term	Definition
AI4Debunk Engine Room	Central backend infrastructure hosting AI modules and validation mechanisms.
DisinfoScore	Numerical indicator produced by AI models assessing likelihood of disinformation.
Disinfowiki	Informal name of the Collaborative Platform developed in Task 10.4.
Knowledge Graph	Data structure linking claims, entities, and validation evidence.
Validation Committee	Human oversight group reviewing automated disinformation detection results.
WP10	Work Package 10 – Definition of Different Interfaces for Online Citizens.
TRL	Technology Readiness Level; TRL7 corresponds to system prototype demonstration in an operational environment.

Funded by the European Union

Review Sheet of Deliverable/ Milestone Report
D10.4 Report on the definition of the collaborative platform

Editor(s):	Jan Kragt (IP) Marcel Keijzer (IP)
Responsible Partner:	Stichting InnovativePower
Status-Version:	Final - v0.3
Date:	10/02/2026
Distribution level (CO, PU):	Public
Reviewer (Name/Organization)	Georgi Gotev, Kalina Angelova (EUalive)
Review date	19/02/2026

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"

Funded by the European Union

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?	Y			
…contain a license (in case of public deliverables)?			NA	
…include the names of contributors and reviewers?	Y			
….has a version table consistent with the document's revision?				
… contain an updated table of contents?	Y			
… contain a list of figures consistent with the document's content?	Y			
… contain a list of tables consistent with the document's content?	Y			
… contain a list of terms and abbreviations?	Y			
… contain an Executive Summary?	Y			
… contain a Conclusions section?	Y			
… contain a List of References (Bibliography) in the adequate format, if relevant?	Y			
… use the fonts and sections defined in the official template?	Y			
… use correct spelling and grammar?	Y			
… conform to length guidelines (50 pages maximum (plus Executive Summary and annexes)	Y			
… conform to guidelines regarding Annexes (inclusion of complementary information)	Y			
… present consistency along the whole document in terms of English quality/style? (to avoid accidental usage of copy&paste text)	Y			
About the content…				
ELEMENT TO REVIEW	**Y**	**N**	**NA**	**COMMENTS**
Is the overall style of the deliverable correctly organized and presented in a logical order?	Y			
Is the Executive Summary self-contained, following the guidelines and does it include the main conclusions of the document?	Y			

Funded by the European Union

ELEMENT TO REVIEW	Y	N	NA	COMMENTS
Is the body of the deliverable (technique, methodology results, discussion) well enough explained?	Y			
Are the contents of the document treated with the required depth?	Y			
Does the document need additional sections to be considered complete?		N		
Are there any sections in the document that should be removed?		N		
Are all references in the document included in the references list?	Y			
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)		N		
SOCIAL and TECHNICAL RESEARCH WPs (WP4, 5, 12, 13, 14)				
ELEMENT TO REVIEW	**Y**	**N**	**NA**	**COMMENTS**
Is the deliverable sufficiently innovative?	Y			
Does the document present technical soundness and its methods are correctly explained?	Y			
What do you think is the strongest aspect of the deliverable?				The deliverable has a clear integration of human-centred design with a robust technical architecture.
What do you think is the weakest aspect of the deliverable?				
Please perform a brief evaluation and/or validation of the results, if applicable.			NA	
AI AND TECNOLOGICAL WPS (WP6 – WP11)				
ELEMENT TO REVIEW	**Y**	**N**	**NA**	**COMMENTS**
Does the document present technical soundness and the methods are correctly explained?	Y			
What do you think is the strongest aspect of the deliverable?				The deliverable is a well-defined, modular, and scalable system architecture that clearly integrates AI detection modules with human oversight and multiple user interfaces.

Funded by the European Union

ELEMENT TO REVIEW	Y	N	NA	COMMENTS
What do you think is the weakest aspect of the deliverable?				While the deliverable clearly describes the system architecture and integration of AI modules, it provides limited detail about how the AI models themselves are built and evaluated.
Please perform a brief evaluation and/or validation of the results, if applicable.			NA	
DISSEMINATION AND EXPLOITATION WPs (WP15 – WP17)				
ELEMENT TO REVIEW	**Y**	**N**	**NA**	**COMMENTS**
Does the document present a consistent outreach and exploitation strategy?	Y			
Are the methods and means correctly explained?	Y			
What do you think is the strongest aspect of the deliverable?				The collaborative platform is clearly designed as a public hub that connects all project tools and aims for long-term use, which supports good visibility and future uptake.
What do you think is the weakest aspect of the deliverable?				The exploitation strategy could be further strengthened by providing more detail on long-term adoption. *Marcel Keijzer: the exploitation strategy will be further detailed in WP11.*
Please perform a brief evaluation and/or validation of the results, if applicable.			NA	
DISSEMINATION AND EXPLOITATION WPs (WP18)				
ELEMENT TO REVIEW	**Y**	**N**	**NA**	**COMMENTS**
Does the document present the main ethical aspects regarding the use of methods and human involvement?	Y			
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.			NA	

SUGGESTED IMPROVEMENTS

PAGE	SECTION	SUGGESTED IMPROVEMENT
		ADD ROWS AS NECESSARY

CONCLUSION

Mark with X the corresponding line.

X	Document accepted, no changes required.
	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