



# AI-Auditology

D5.2

## Dissemination, communication, exploitation and clustering plan

<b>Project Title</b>	<b>Model-based Auditing of Social Media AI Algorithms and their Tendencies to Spread Harmful Content</b>
<b>Contract No.</b>	<b>09I03-03-V03-00020</b>
<b>Project start date</b>	<b>August 2024</b>
<b>Duration</b>	<b>23 months</b>



Funded by  
the European Union  
NextGenerationEU

**PLÁN [OBNOVY]**

Grant agreement no.: 09I03-03-V03-00020  
 Project acronym: AI-Auditology  
 Project website: <https://kinit.sk/project/AI-Auditology>  
 Project full title: Model-based Auditing of Social Media AI Algorithms and their Tendencies to Spread Harmful Content  
 Project start date: August 2024 (23 months)  
 Work Package: WP4 – Communication, Dissemination & Exploitation  
 Version: 1.0  
 Authors: Ivan Srba  
 Delivery date: January 27, 2025

Project funded by VAIA - Research and innovation authority, under the call 09I03-03-V03, Grant agreement no. 09I03-03-V03-00020		
Dissemination Level		
PU	Public	x
NP	Non-public, only for members of the consortium (including the Agency Services)	

# Table of Contents

<b>1. Introduction.....</b>	<b>4</b>
<b>2. Overview of Communication, Dissemination, Clustering and Exploitation.....</b>	<b>5</b>
2.1 Objectives.....	5
2.2 Phases.....	5
<b>3. Dissemination, Communication and Clustering Strategy.....</b>	<b>6</b>
3.1 Target Groups.....	6
3.2 Channels and Activities.....	6
3.4 KPIs.....	7
3.5 Dissemination, Communication and Clustering Timeline.....	8
3.5 Open science.....	8
3.6 Project Identity.....	8
<b>4. Exploitation Strategy.....</b>	<b>10</b>
4.1 Stakeholders.....	10
4.2 Key Exploitable Results.....	10
4.3 Activities and Actions.....	11
4.4 IPR management.....	11
<b>5. Conclusions.....</b>	<b>13</b>
<b>Annex 1 - AI-Auditology Brand kit.....</b>	<b>14</b>

## 1. Introduction

This document presents the plan for dissemination, communication and clustering activities in a structured way. It also covers the adopted approaches to exploitation. The first version of this plan was prepared in M6 of the project duration and is continuously updated and maintained as the project progresses.

AI-Auditology communication, dissemination, clustering and exploitation activities aim to maximize the project's visibility and impact by sharing its results with audiences beyond the project.

## 2. Overview of Communication, Dissemination, Clustering and Exploitation

The activities related to the communication, dissemination, clustering and exploitation of the project results will be realized under the WP5. Tailored activities will be designed to make the project outcomes visible and accessible to various target groups.

### 2.1 Objectives

The project has a deeply embodied short-term as well as long-term impact not only in the research perspective (contributing a new research direction) but also in the practical perspective (enabling highly demanded monitoring of social media AI algorithms and their negative effects on the society, and thus contributing to EU (self-)regulatory initiatives).

Within the Ai-Auditology project, we aim to implement a dissemination and communication strategy to diffuse the project's results to target groups using appropriate and effective instruments. For this purpose, we leverage the brand of the KInIT institute and tap into its channels and networks across the media and technology sectors, links to relevant stakeholders and outreach to citizens.

The general objectives for the project are derived from the European Commission Horizon Europe programme as follows:

- Inform, promote and communicate activities and results
- Make knowledge and results publicly available free-of-charge
- Make concrete use of results for commercial, societal and political purposes

### 2.2 Phases

Adopted plan to communication, dissemination, clustering and exploitation follow three phases:

- **Awareness building phase (M1-M6):** before the project produces concrete outputs, the communication strategy will focus on raising awareness among the stakeholders about the project and research plans to pique interest.
- **Participation phase (M7-M11):** as the research results become available, the outreach activities will intensify to ensure proactive engagement of specific stakeholder groups with the outputs to validate them against real-world scenarios and guide further research.
- **Action phase (M12-M23):** as the project's results mature, the project will increasingly focus on influencing new practices around its outputs and ensuring the exploitation and future research activities.

### 3. Dissemination, Communication and Clustering Strategy

Project's dissemination, communication and clustering activities will be carried out to ensure broad visibility and identification of the project. This publicity will be compatible with the protection of intellectual property rights, and with the maximum participation of external stakeholders in the discussions.

#### 3.1 Target Groups

With communication activities, we aim to address the following target groups:

1. scientific and academic community working in particular on AI, recommender systems and/or dis/misinformation analysis;
2. stakeholders & practitioners (e.g., regulators, watchdog organizations, NGOs);
3. ICT industry (e.g., recommender system providers), including social media platforms;
4. the general public at large.

Typically, we primarily target researchers and professionals working in: 1) AI and recommender systems; 2) social networks; and 3) disinformation and misinformation analysis and combatting.

#### 3.2 Channels and Activities

To address our target groups and align with the dissemination, communication and clustering strategy, the following channels and activities are planned:

1. **Project website.** The website will provide information about project activities outcomes, as well as demonstrate and host project's outcomes and results (e.g., scientific papers, presentations, code, API documentation) and report about project's results, collaborations and more.
2. **Social media presence.** For active social media presence, we will utilize an already established KInIT organizational social media accounts at [LinkedIn](#) (3700 followers, ~500 impressions per post), [Facebook](#) (1400 followers, ~350 impressions per post) and eventually also [X/Twitter](#) (150 followers, ~50 impressions per post). Additional platforms may be potentially used in the later stage of the project, especially Bluesky as a potential replacement for X platform.

The posts related to the AI-Auditology project at all adopted social media are supposed to be accompanied with the dedicated hashtag #AIAuditology, which allows users to easily identify and navigate through the project-related social media posts.

3. **Blogs, press releases, newsletters.** We will regularly inform about project updates utilizing blogs (published at the KInIT website), press releases (about achieving major project milestones, like project initialization) and newsletters (utilizing KInIT's newsletter with 600 subscribers).
4. **Participation in academic and networking events.** It is planned to attend research conferences and workshops covering the following topical areas:

- a. Natural Language Processing and Text Analytics (e.g., ACL, EACL, EMNLP);
- b. Machine Learning (e.g., ICML, ECML-PKDD);
- c. Recommender Systems and User Adaptation (e.g., RecSys, UMAP); and
- d. Social Media and Online Social Networks (e.g., WWW, ACM CHI, ICWSM, CIKM).

In addition, we plan to attend networking and knowledge sharing events (e.g., EU DisinfoLab conference, Dagstuhl seminars, winter/summer schools).

5. **Scientific publications:** We will publish research papers in conferences, workshops, scientific journals/books that focus on computer science. For publishing key outcomes, we will target top-level A\* conferences (e.g., ACL, ICML) and high-impact journals (e.g., ACM Transactions on Information Systems, ACM Transactions on Intelligent Systems and Technology, Information Processing & Management).

### 3.4 KPIs

KPIs (against which our communication and dissemination plans and activities are to be evaluated) are defined separately for the first (M1-M11, until the midterm report) and the second half (M12-M23) of the project implementation.

Channel	Target groups	Objective	KPIs (for 1 <sup>st</sup> and 2 <sup>nd</sup> half of the project duration)
Project website	All	- Information and knowledge diffusion - Results presentation	Page views/Impressions ≥ 2000, ≥ 3000
Social media presence	All	- Awareness creation - Knowledge diffusion - Results presentation	Number of published posts: ≥ 30, ≥ 40 Number of impressions: ≥ 21K, ≥ 28K
Blogs, press releases, newsletters	All	- Awareness creation - Media and other relevant "multipliers" engagement	Number of blogs/press releases/newsletters: ≥ 10, ≥ 15
Participation in academic and networking events	- Scientific and academic community - Stakeholders & practitioners - ICT industry	- Engagement of academic & industrial community - Methodology Presentation/Validation - Networking and Collaboration with the relevant stakeholders	Number of attendances: ≥ 4, ≥ 8
Scientific publications	- Scientific and academic community - Stakeholders & practitioners - ICT industry	- Diffusion to the scientific community - Results presentation	Number of publications: ≥ 3, ≥ 6

### 3.5 Dissemination, Communication and Clustering Timeline

The following timeline represents the breakdown of the planned dissemination, communication and clustering over the duration of the project. Provided numbers indicate the expected contribution towards the defined KPIs (Section 3.4).

Channel	M1-M4	M5-M8	M9-M11	M12-M14	M15-M17	M18-M20	M21-M23
Project website	500x	700x	800x	800x	900x	1000x	1100x
Social media presence		15x	15x	10x	10x	10x	10x
Blogs, press releases, newsletters		4x	6x	5x	4x	3x	4x
Participation in academic and networking events	1x	1x	2x	3x	1x	2x	2x
Scientific publications		1x	2x	2x	1x	2x	1x

### 3.5 Open science

We will fully support openness and reproducibility of our research. All publications will be published either as green open access in the institutional repository or at arXiv.org or as gold open access under open licences. Whenever possible, we will share all datasets created as a part of project research activities at Zenodo portal in an anonymized form following the FAIR principles, i.e., they will be easily findable, freely accessible, published in standard interoperable formats, and reusable. In the same manner, we will publish code, designed methods (algorithms), and trained models (via standard means, such as GitHub, Hugging Face, or the European AI on Demand Platform). By promoting openness of achieved results, we also aim to demonstrate and support the credibility of auditing results.

### 3.6 Project Identity

To maximize the project communication, dissemination and exploitation consistency and to ensure recognizable visual identity, the brand kit was developed under the WP5, and is included as the annex of this document. It consists of a logo and the guidelines for the use of the visuals in different formats, sizes and colors. It is used throughout the project in all channels, which include the website, newsletters, social media channels, text documents, presentations, and other visual and branding materials used during the project.

Communication and dissemination activities related to AI-Auditology project results must acknowledge EU support, where relevant. This acknowledgement is specified in the Article 6 General conditions of AI-Auditology Grant Agreement (in Slovak):



*“b) ak Predmet Projektu nie je hmotne zachytiteľný:*

- 1. uvedenie emblému EÚ s nápisom „Financovaný Európskou úniou NextGenerationEU“/„Financované Európskou úniou NextGenerationEU“ a loga Plánu obnovy na vlastnom webovom sídle, a*
- 2. uvedenie Plánu obnovy, na základe ktorého je umožnené financovanie z Prostriedkov mechanizmu, v komunikácii voči verejnosti, na sociálnych sieťach a pod.”*

## 4. Exploitation Strategy

An exploitation strategy is designed to optimize the use of resources to maximize project results, sustainability and future utilization. The strategy focuses on improving operational efficiencies, and leveraging established market positions to generate sustainable uptake of the results. It focuses on exploiting current assets of the Host organization.

### 4.1 Stakeholders

We have identified four main stakeholder categories that may be interested in the exploitation of the project results:

1. the research and academic community, including universities and other research organisations;
2. the regulators at the EU/national level, watchdog organizations and NGOs focused on algorithmic auditing;
3. the ICT community, with a commercialisation interest for products and/or services that may be developed and delivered on top of the project results; and last but not least
4. social media platforms interested in delivering unbiased AI algorithms to their end-users.

### 4.2 Key Exploitable Results

Within the project, we identified three key exploitable results:

1. **Harmful content model.** The model will provide a valuable potential for additional research task also beyond the area of algorithmic audits, such as disinformation detection, credibility analysis, or coordinated behaviour detection to name few. It will also allow additional highly multidisciplinary research since it can be exploited also by scientists from social sciences and humanities.
2. **AI-based methods for audit scenario creation, translation, and execution.** Methods and models that support model-based auditing of AI algorithms can be exploited by all identified stakeholders.
3. **Auditing research infrastructure.** Finally, the auditing infrastructure represents a highly valuable assets that will be obtained during the project implementation. Due to its universality and flexibility, after the completion of the project, it will be not only maintained, but also further extended and utilized as a part of future research studies. This research infrastructure may provide KInIT with significant competitive advantage when applying for additional research projects (e.g., from Horizon Europe scheme) as well as it may be of high interest from the side of regulators and European Commission.

### 4.3 Activities and Actions

In order to examine exploitation opportunities for AI-Auditology project, we already started and will continue to monitor:

1. **Existing solutions.** At first, we will regularly monitor existing solutions that are capable of performing various types of algorithmic audits. Such solutions can be at various levels of maturity, ranging from early prototypes developed by researchers up to commercial solutions. So far, we have identified and thoroughly analysed only one concurrent commercial platform named CrossOver provided by [Check First](#) Finnish software and methodologies company. This solution is, however, adopting a diametrically different approach as one envisioned in our AI-Auditology project (in terms of conceptual as well as technical approach).
2. **EU and national legislation.** Second, we will keep track of legislation proposed and adopted either at EU or national level. More specifically, we recognized a potential synergy with Digital Service Act (DSA) and independent auditing of VLOPs and VLOSEs. Algorithmic auditing researched in the AI-Auditology project can represent one of means how such audits of online platforms and search engines can be performed.

More specifically, we plan to organize meetings with identified stakeholders and evaluate with them various potential exploitation activities and actions that will be undertaken further during the project duration.

To make a long-term exploitation and adoption of the achieved results easier, the future exploitation strategy (which is beyond this research project) will include creation of an end-to-end auditing framework that will provide guidelines on utilization of the model-based audits in the broader context (e.g., how to define the scope of the audit, how the audit question should be specified, or how to perform post-audit actions, such as how to interpret cross-platform or longitudinal results). The auditing framework will be disseminated to researchers and practitioners (in the open-science spirit) as a white paper accompanied with all other projects' results - the research publications, datasets, and open-source software. Publishing and adhering to a reputable auditing framework will also contribute to transparency and trustworthiness of the audits.

Finally, from the early stages of the project implementation, PI and the rest of the team will start securing additional funding to cover continuation of the research activities in the area of model-based algorithmic auditing and further development of the research auditing infrastructure. The primary focus will be given to research programmes within the Horizon Europe scheme, but also a possible commercialization of the key exploitable results will be examined.

### 4.4 IPR management

There are two kinds of Intellectual Property Rights (IPR) issues that will be addressed in the project: IPR related to online content and IPR arising from the project outcomes. With respect to the former, the tools and technologies to be developed in AI-Auditology, including

the harmful content model will be designed so as to not publish the online content (such as fact-checking articles), but will instead refer to online content via URLs. In this way, the tools will not infringe copyright or raise IPR issues. With respect to the latter – the project recognises the importance of effective IPR management, which lays the basis of successful future exploitation of its results and outcomes.

We expect two groups of results of the project: open (results free to re-use by any third party) and IP (results on which we will retain our IPR, intended for further exploitation). The goal of the IPR management is to maximize future exploitation of results (where possible publish the research results in line with the open science spirit, see section 1.5), while protecting the valuable assets of the research team (especially regarding the developed auditing research infrastructure). IPR management will be continuously addressed and refined in WP1 throughout the project implementation.

## 5. Conclusions

This Deliverable has outlined the plan for dissemination, communication, clustering and exploitation activities. The project has developed a project identity that can be recognised as such. It is used whenever possible in all outreach activities. Related activities have established and are continuing to establish AI-Auditology as a recognisable brand in the field of algorithmic auditing and countering harmful content with the help of Artificial Intelligence.

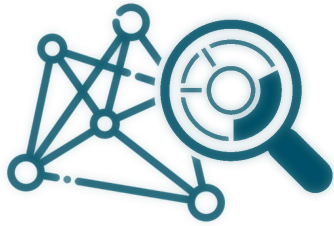
A wide range of outreach – communication, dissemination and clustering – activities and actions have been defined and described. They are now being pursued and are essential components of AI-Auditology’s strategy, containing measurable and defined KPIs, targeting respective audiences

In terms of exploitation planning, the deliverable identifies the main stakeholders and key exploitable results, as well as sets out main exploitation activities towards implementing the devised strategy.

This document will function as a "living document" throughout the project's duration, guiding the dissemination, communication, clustering and exploitation activities. Additional interim updates will be provided as needed.

## Annex 1 - AI-Auditology Brand kit

### Logo



### Logo with project acronym



### Primary colors

#035369

#00708e

### Secondary colors

#00708e

#deebf6

## Background

