



D5.3

Dissemination, communication, exploitation and clustering report

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



**Funded by
the European Union**

**[RECOVERY
AND RESILIENCE]
PLAN**

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: WP5 – DCE and clustering
 Version: 1.0
 Delivery date: June 30, 2026

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

Table of Contents.....	3
1. Introduction.....	4
2. Dissemination, Communication and Clustering Outcomes.....	5
2.1 Project website.....	5
2.2 Social media presence.....	7
2.3 Blogs, newsletters, podcast, news articles and interviews.....	16
2.4 Scientific publications.....	19
2.5 Participation in academic and networking events.....	22
2.6 Source code repositories and datasets.....	26
3. Exploitation Outcomes.....	27
3.1 Policy and regulation recommendations and briefs.....	27
3.2 Follow-up projects.....	29
4. Evaluation of KPIs.....	30
5. Conclusion.....	31

1. Introduction

This document concludes dissemination, communication, exploitation (DCE) and clustering activities conducted within AI-Auditology work package 5 (WP5).

The undertaken activities, which followed the DCE and clustering plan (prepared as deliverable D5.2), aimed to systematically maximize the project's visibility and impact by sharing its results with audiences beyond the project. To this end, the project successfully adopted three phases of DCE and clustering activities as follows:

- **Awareness building phase (M1-M6):** before the project produced concrete outputs, the communication strategy focused on raising awareness among the stakeholders about the project and research plans to pique interest.
- **Participation phase (M7-M11):** as the research results started to become available, the outreach activities intensified 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 matured, the project increasingly focused on influencing new practices around its outputs and ensuring the exploitation and future research activities.

To maximize the project communication, dissemination and exploitation consistency and to ensure recognizable visual identity, the brand kit was developed under the WP5 (D5.2, Section 3.6), and was followed, together with identification of the project and its funding, within all DCE and clustering activities.

The rest of this report is structured into three parts:

- The overview of dissemination, communication and clustering activities (Section 2), which is further structured according to the individual channels defined by the DCE and clustering plan (D5.2, Section 3.2).
- The overview of exploitation activities (Section 3).
- The overview of the DCE and clustering KPIs and corresponding metrics, as well as the actual project outcomes, which show a successful accomplishment of all the defined KPIs (Section 4).

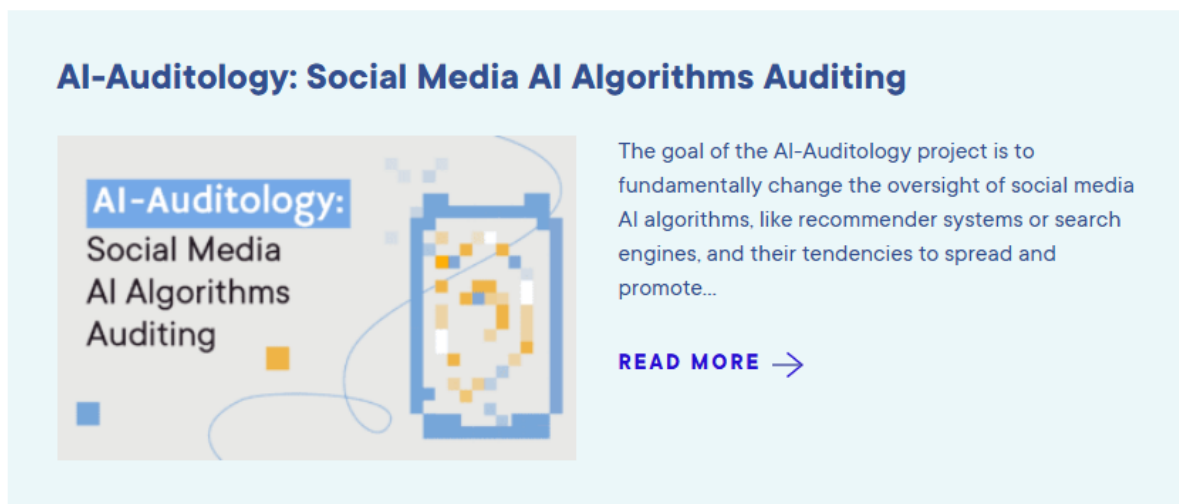
Finally, the concluding section (Section 5) summarises the DCE and clustering activities as well highlights the most significant outcomes (in terms of the achieved impact).

2. Dissemination, Communication and Clustering Outcomes

2.1 Project website

The project website provides information about project activities outcomes, as well as demonstrates and hosts project's outcomes and results (e.g., scientific papers, presentations, code, API documentation) and reports about project's results, collaborations and more.

It was created under the already established and recognized KInIT's website. Particularly, the project is presented by a form of short project card available at: <https://kinit.sk/research/>



Besides that, full project profile is available at:

<https://kinit.sk/project/ai-auditology-social-media-ai-algorithms-auditing/>

The project website is structured into the following subsections:

- project title and abstract,
- news & blogs,
- research team,
- selected publications,
- deliverables and documents.

Finally, the project website consisted of news and blogs (accessible also from the full project website, for the list of all blogs and news, please, see Section 2.3 of this deliverable).

The project website (the full project profile, individual blog entries) was actively promoted within social media posts (Section 2.2) or presentations at academic or networking events (Section 2.5).

[About us](#) [Research](#) [Education](#) [Careers](#) [News](#) [Partners](#) [Contact](#) [Get Involved](#) [SK](#)  

[Home](#) > [Research](#) > [Web & User Data Processing](#) > AI-Auditology: Social Media AI Algorithms Auditing

AI-Auditology: Social Media AI Algorithms Auditing



PROJECT

Duration:
08/2024 – 06/2026

Funding agency:
Research Agency
(Slovakia)

Project type: Scientific
project

Coordinator: KiniT

Principal investigator:
Ivan Srba

AI-Auditology: Social Media AI Algorithms Auditing

The goal of the AI-Auditology project is to fundamentally change the oversight of social media AI algorithms, like recommender systems or search engines, and their tendencies to spread and promote harmful content. The main ground-breaking idea of this proposal is a novel paradigm of model-based algorithmic auditing.

The project serves a **noble purpose** and will have a **great positive societal impact** on the European (at the very least) online media environment. Currently, media regulators have no real tool to quantitatively investigate whether big tech follows on their self-regulatory promises or adheres to relevant legislation on tackling recommender system biases, mitigating toxic content, or preventing misinformation spread. With the new methods and supporting software platform built, the public will receive a **powerful watchdog tool, fully in line with European values**.

The social media *algorithmic auditing* is a process, in which the audited platform is stimulated by behavior generated by automated agents (posing as platform users), while the platform's responses (e.g. recommendations) are collected and analyzed. The *model-based algorithmic auditing* refers to an innovative application of **model-based design for supporting or even automating** individual steps of the auditing process (i.e., creation of audit scenarios, audit execution and evaluation).

The underlying *harmful content model* provides a simplified and partial representation of harmful content in a social media environment and of user behavior around such content (e.g., what kind of users and how exactly they interact with different instances of harmful content). It is built and continuously maintained from outcomes of research studies and public opinion polls, from harmful content combating activities (e.g., fact-checking) as well as from the real-world social media data. Such a model has not been researched so far and besides revolutionizing algorithmic auditing, it would have a **big multidisciplinary impact on other research areas** as well.



with different instances of harmful content). It is built and continuously maintained from outcomes of research studies and public opinion polls, from harmful content combating activities (e.g., fact-checking) as well as from the real-world social media data. Such a model has not been researched so far and besides revolutionizing algorithmic auditing, it would have a **big multidisciplinary impact on other research areas** as well.

Harmful content model is consequently used as an integral part of the auditing process – first and foremost, to assist human experts by generating **abstract audit scenarios**, which are translated to **platform-specific scripts** at execution time. The whole auditing process is supported by an **auditing research infrastructure** utilizing advanced techniques from **machine learning and natural language processing**.

PROJECT

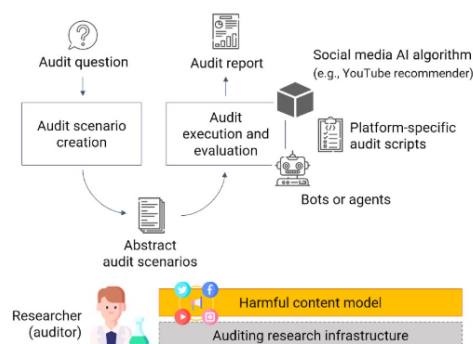
Duration:
08/2024 - 06/2026

Funding agency:
Research Agency
(Slovakia)

Project type: Scientific
project

Coordinator: KlnIT

Principal investigator:
Ivan Srba



The proposed model-based algorithmic auditing and next-generation audits fall into a so-far under-researched area of algorithmic auditing. The project will therefore contribute to a **new and impactful research direction** of model-based algorithmic auditing, which addresses difficult challenges hampering further progress in auditing social media AI algorithms. It is also in line with ongoing EU legislation initiatives demanding independent social media oversight and impacts all involved stakeholders, including social media platforms' users.

Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I03-03-V03-00020.



Funded by the
European Union
NextGenerationEU

**[RECOVERY
AND RESILIENCE
PLAN]**

The project website (short project profile, full project profile, blogs) achieved 5049 (KPI: 2000) and 4502 (KPI: 3000) impressions for the first and the second year of project implementation, respectively, fulfilling the planned KPIs.

2.2 Social media presence

For active social media presence, we utilized already established KlnIT organisational social media accounts at [LinkedIn](#) (6000 followers as of June 2026), [Facebook](#) (1800 followers) and [X/Twitter](#) (224 followers). As already anticipated in the DCE and clustering plan, following the decline of moderation and content quality at the X platform, the [Bluesky](#) network was adopted as a replacement. In addition, besides the organizational social media accounts, the communication took place also by means of accounts of individual team members. This is also due to the setup of social media algorithms (especially at LinkedIn platform) that systematically prefer content from individual accounts as from organisational ones.

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 on all 4 platforms: [Facebook](#), [LinkedIn](#), [X](#) and [Bluesky](#).

The communication and dissemination strategy was continually improved during the course of the project (reflecting an evolving strategy adopted by the whole KInIT organisation). This influenced the language of the content, target audience and overall style of the social media posts as follows:

- The Facebook platform was used to communicate the content primarily in Slovak language and targeting broader public (including social media users for whose were the project results highly relevant).
- The LinkedIn and X platforms (later replaced with Bluesky) were used to communicate the content primarily in English language targeting primarily the scientific and academic community, stakeholders & practitioners (including regulators, watchdog organizations, NGOs), and the ICT industry.

The following image provides an example of a [LinkedIn post](#) introducing the project.



Besides our own communication, the project outcomes attracted a significant attention of our target audience and were further shared by stakeholders. The following example of [LinkedIn post](#) demonstrates the impact of our research publication and in-person presentation at the FAccT 2026 conference.



Tamara Atanasoska • 2nd

Doctoral Researcher ML/NLP | Software Engineer | OSS Maintainer

19h • Edited •

+ Follow



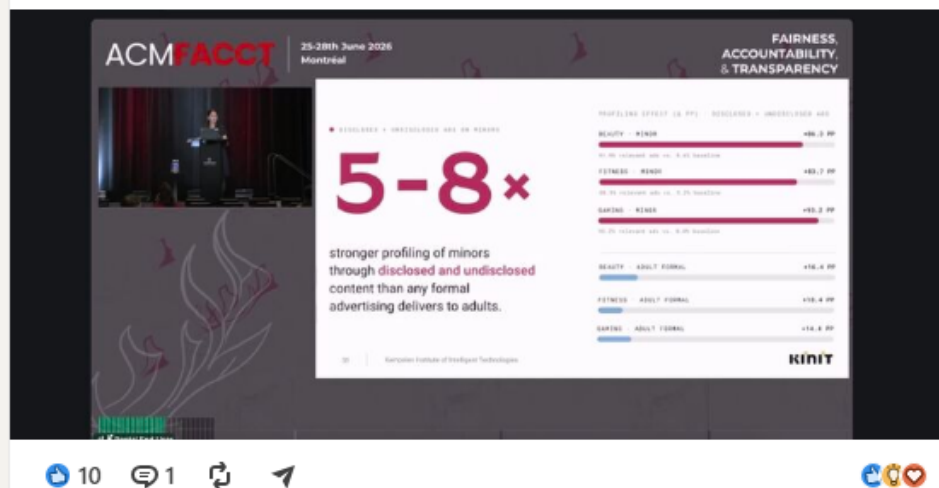
In their [ACM FAccT](#) award winning paper, "The DSA's Blind Spot: Algorithmic Audit of Advertising and Minor Profiling on TikTok", [Sara Solarova](#) and collaborators find that targeted profiling of minors is "very much" present on TikTok and at a much higher rate when compared to adults.

The targeting is both personalised and often undisclosed, both through the "paid partnerships" of the platform, but also the influencer ads that mask as content (you know the kind!). These types of commercial content are not part of the Digital Services Act, enabling the unhindered profiling of and servicing of such ads to minors.

While this is problematic from a legal perspective, what makes it alarming is the cognitive reality of the developing brains of teens. Teens are defenceless against the surge of positive feelings when served these ads as well as the subsequent creation of para-social relationships with influencers because their prefrontal cortex is still developing. We often feel the "ugh" negative kind of feeling when being shown these ads without consent ("recognising and resisting persuasive intent"), but the teenagers haven't still developed the cognitive control for the same reaction.

The authors note that "advertising literacy alone is therefore insufficient", which means that we cannot explain away these effects as parents. They further have concrete suggestions for broadening the ad definitions to strengthen the regulation and demand platform cooperation.

Link to paper: <https://lnkd.in/d8tfj6ua>



The following table provides a list of all social media posts together with the number of impressions they received. The number of impressions for posts at Bluesky platform is missing as this information is not available publicly nor to their creator.

Date	Excerpt from the post	Info	Language	Link	Impressions
08.09.2024	Help us to uncover  how social media algorithm influence our everyday lives!	LinkedIn post - by Ivan Srba	EN	link	1967
10.09.2024	 AI-Auditology, our new research project, aims to uncover how social media algorithms contribute to the spread of disinformation 	LinkedIn post - by KInIT	EN	link	456
		X post - by KInIT	EN	link	97
		Facebook post - by KInIT	EN	link	421
25.02.2025	 Did you know that big online platforms now have to undergo independent audits to assess systemic risks, algorithmic transparency, and user protections?	LinkedIn post - by KInIT	EN	link	231
30.01.2025	Responsible and fair AI is an active research topic at KInIT  	LinkedIn post - by KInIT	EN	link	559
		X post - by KInIT	EN	link	56
		Facebook post - by KInIT	EN	link	474
25.02.2025	Holding Big Platforms Accountable: The Why, What, and How of Auditing	LinkedIn post - by KInIT	EN	link	632
		X post - by KInIT	EN	link	64
		Facebook post - by KInIT	EN	link	311
16.05.2025	 How well does YouTube tackle misinformation?	LinkedIn post - by KInIT	EN	link	574
		X post - by KInIT	EN	link	57
		Facebook post - by KInIT	EN	link	560
21.05.2025	Pretty exciting time to be part of the Kempelen Institute of Intelligent Technologies. And even better projects and research opportunities are yet to come, I truly believe.	LinkedIn post - by Matúš Tibenský	EN	link	517
25.06.2025	How transparent is TikTok's algorithm? 	LinkedIn post - by KInIT	EN	link	481
		X post - by KInIT	EN	link	55
		Facebook post - by KInIT	EN	link	324

01.07.2025	🤖 Auditing social media sounds simple - but in reality, it's anything but that!	LinkedIn post - by KInIT	EN	link	510
		X post - by KInIT	EN	link	43
		Facebook post - by KInIT	EN	link	343
01.07.2025	Today, at hashtag#EWAF (European Workshop on Algorithmic Fairness), organized by European Association for Algorithmic Fairness, I presented our position paper on future directions for the research and development of algorithmic auditing of recommender systems.	LinkedIn post - by Jakub Šimko	EN	link	2905
07.07.2025	🔥 Why do the most viral posts grab our attention — and what's the hidden cost?	LinkedIn post - by KInIT	EN	link	505
		X post - by KInIT	EN	link	49
		Facebook post - by KInIT	EN	link	279
07.07.2025	During the last week I had a chance to attend the University of Amsterdam summer school 2025 - Digital Methods.	LinkedIn post - by Filip Hossner	EN	link	1536
17.07.2025	🔒 What happens when research APIs shut down? ❌	LinkedIn post - by KInIT	EN	link	785
		X post - by KInIT	EN	link	82
		Facebook post - by KInIT	EN	link	402
17.07.2025	Last week, I had the opportunity to present our paper <i>Revisiting Algorithmic Audits of TikTok: Poor Reproducibility and Short Term Validity of Findings</i> at sigir-2025	LinkedIn post - by Adam Škurla	EN	link	1161
22.07.2025	Ako ťa ovládajú sociálne siete? 🌐	LinkedIn post - by KInIT	SK	link	1,124
		Facebook post - by KInIT	SK	link	895
23.07.2025	Last week, I had the opportunity to attend the sigir-2025 conference	LinkedIn post - by Branislav Pecher	EN	link	871
24.07.2025	Measuring the behavior of social media recommender systems is at the center of our #aiauditology project at #KInIT.	Bluesky post - by KInIT	EN	link	-
26.07.2025	This week, I will be attending the hashtag#ACL2025 conference from the Association for Computational Linguistics.	LinkedIn post - by Branislav Pecher	EN	link	433

31.07.2025	Although the main #ACL2025 conference ended yesterday, let me get back to the papers we presented here.	LinkedIn post - by Róbert Móri	EN	link	1614
		X post - by KInIT	EN	link	244
01.08.2025	#ACL2025 recap At recent ACL 2025 Conference, our researchers presented paper ...	Bluesky post - by KInIT	EN	link	0
02.08.2025	Reflecting on an inspiring week at hashtag#ACL2025	LinkedIn post - by Ivan Srba	EN	link	728
		LinkedIn post - by KInIT	EN	link	850
		Facebook post - by KInIT	EN	link	511
		X post - by KInIT	EN	link	65
21.08.2025	Most of our researchers come from academic backgrounds and thus have considerable lecturing experience 🧑🏫	Bluesky post - by KInIT	EN	link	0
21.08.2025	🎉 Excited to share that two of my papers have been accepted to EMNLP 2025!	LinkedIn post - by Branislav Pecher	EN	link	1286
12.09.2025	🎧 TIP NA PODCAST Náš výskumník Jakub Šimko v podcaste Nezablúdme v bludoch Demagog.SK vysvetľuje:	Facebook post - by KInIT	SK	link	297
25.09.2025	🇸🇰 KInIT mieri na EMNLP 2025! 🚀	Facebook post - by KInIT	SK	link	543
25.09.2025	🇸🇰 KInIT is heading to EMNLP 2025! 🚀	LinkedIn post - by KInIT	EN	link	1,435
14.10.2025	Nedávno sme privítali štyroch stredoškóľakov z Delta - Střední škola informatiky a ekonomie Pardubice. Vyskúšali si u nás prácu na projekte #AIAuditology a bolo to super! 🧑🏫💡	Facebook post - by KInIT	SK	link	263
14.10.2025	Recently, we had the pleasure of welcoming four talented students from DELTA Pardubice, who joined our team and worked on the hashtag#AIAuditology project.	LinkedIn post - by KInIT	EN	link	822
23.10.2025	🎤 Meet the moderator – Sara Solarova!	LinkedIn post - by EUDHIT	EN	link	214
03.11.2025	Algorithmic auditing as a way to obtain data access for researchers and audit systemic risks?	LinkedIn post - by Sára Soľárová	EN	link	2928
03.11.2025	🌟 CONFERENCE HIGHLIGHTS FROM OUR TEAM!	LinkedIn report - by KInIT	EN	link	530

15.11.2025	For better or worse, social platforms have become key gateways not only to entertainment, but also to information and news for an ever-growing audience.	LinkedIn post - by Matúš Tibenský	EN	link	1028
19.11.2025	New blog! Can the DSA really keep platforms in check? 	LinkedIn post - by KInIT	EN	link	855
		Bluesky post - by KInIT	EN	link	-
19.11.2025	 My two cents on why platform audits need to get technical (and why Brussels is a breath of fresh air for me 🇧🇪)	LinkedIn repost - by Sára Sořárová	EN	link	1415
07.01.2026	What does it really take for independent researchers to gain access to platform data?	LinkedIn post - by KInIT	EN	link	662
20.02.2026	The spread of false information, propaganda and scams on social media is amplified by recommender systems, which act as filters and boosters of content that bad actors (including FIMI) exploit.	LinkedIn post - by IS2	EN	link	-
25.02.2026	How does TikTok decide what shows up on your For You page? 	LinkedIn post - by KInIT	EN	link	777
26.02.2026	How does TikTok's FYP work?  Ivan Srba explains it's about explicit likes vs. implicit watch time.	Bluesky post - by KInIT	EN	link	-
26.02.2026	Two days ago, we met with innovation ecosystem partners from Vienna and Bratislava to discuss collaboration and innovation in our region. 	LinkedIn post - by KInIT	EN	link	786
13.03.2026	TikTok says it doesn't profile minors.  We were curious, so we put that claim to the test.	LinkedIn post - by KInIT	EN	link	972
		Bluesky post - by KInIT	EN	link	-
19.03.2026	Excited to speak at the IEEE Central European AI Summit in Budapest this Saturday on data access and platform accountability. I will talk more in detail about how our model-based algorithmic auditing (hashtag#AlAuditology) at Kempelen Institute of Intelligent Technologies can provide a way to access data big platforms would otherwise not provide.	LinkedIn post - by Sára Sořárová	EN	link	970
20.03.2026	Next week, I will be attending the #EACL2026 conference in Morocco.	LinkedIn post - by Branislav Pecher	EN	link	1066
20.03.2026	Happy to share that 3 papers I (co)-authored will be presented at EACL 2026 in Rabat next week!	LinkedIn post - by Ján Čegiň	EN	link	848
23.03.2026	  Our Sara Solarova at the IEEE AI Impact Summit	LinkedIn repost - by KInIT	EN	link	590

23.03.2026	Can we actually understand the black boxes?	LinkedIn post - by Sára Sořárová	EN	link	2429
24.03.2026	Myslíš, že dokážeš rozoznať reklamu od bežného videa? 🤔	Instagram post - by KInIT	SK	link	557
26.03.2026	Come say hi to the poster number 2 at hashtag#EACL - from 9:00 to 10:30. All about using LLMs as generator of synthetic samples for training smaller models for low-resource languages	LinkedIn post - by Branislav Pecher	EN	link	861
27.03.2026	#EACL 2026 Poster Session 🧠📌	LinkedIn post - by KInIT	EN	link	1,734
31.03.2026	Finally home after the #EACL 2026 conference and I already had some time to reflect on it.	LinkedIn post - by Branislav Pecher	EN	link	876
08.04.2026	Getting access to TikTok data as a researcher is still anything but straightforward.	LinkedIn post - by KInIT	EN	link	846
22.04.2026	I am very happy to share that our article "Beyond the Checkbox: Strengthening DSA Compliance Through Social Media Algorithmic Auditing" has been published in the proceedings of CHI '26 Conference on Human Factors in Computing Systems.	LinkedIn post - by Sára Sořárová	EN	link	1959
23.04.2026	For the second time already, I attended the Night of ideas in Bratislava organized last week by the Institut français de Slovaquie and the French Embassy.	LinkedIn post - by Róbert Móro	EN	link	352
27.04.2026	We often hear that social media algorithms “push” people toward extremes—into rabbit holes, filter bubbles, and polarized content. A common term for this is drift.	LinkedIn post - by Matúš Tibenský	EN	link	995
15.05.2026	For the second year, KInIT has mentored talented students from DELTA - Střední škola informatiky a ekonomie, Pardubice.	LinkedIn post - by KInIT	EN	link	996
22.05.2026	How should TikTok's Virtual Compute Environment (VCE) actually work for independent researchers? 🛡️📊	Bluesky post - by KInIT	EN	link	-
08.06.2026	Teenagers on TikTok and what we actually know about what is happening to them there?	LinkedIn post - by Sára Sořárová	EN	link	3138
16.06.2026	👉 Algorithm talk at the IS2 - Information Security Summit in Prague 🇨🇪	LinkedIn post - by KInIT	EN	link	314
17.06.2026	One thing was easier when I was a kid in the 1990s: recognizing advertisements. Ads interrupted movies on TV, filled billboards across the city, and occupied clearly marked pages in newspapers. You usually knew when someone was trying to sell you something.	LinkedIn post - by Matúš Tibenský	EN	link	510

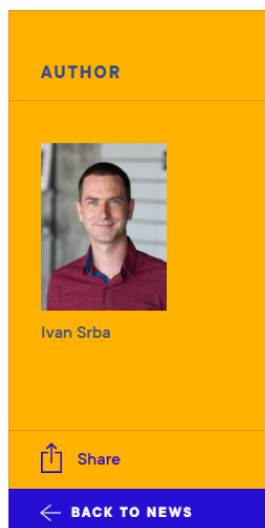
18.06.2026	I am pleased to share our research paper, "Algorithmic Audit of Personalisation Drift in Polarising Topics on TikTok," which was recently presented at the 34th ACM Conference on User Modeling, Adaptation and Personalization (UMAP 2026) in Gothenburg, Sweden.	LinkedIn post - by Matúš Tuna	EN	link	193
23.06.2026	Can high school students help audit social media algorithms? 🤔 They absolutely can! 📱✅	LinkedIn post - by KInIT	EN	link	155
25.06.2026	👉 Does TikTok profile minors to serve them personalized commercial content? Thrilled to share that our research is now published and I will be presenting it at ACM #FAccT2026 in Montréal this week. 🇨🇦	LinkedIn post - by Sára Soľárová	EN	link	1342

2.3 Blogs, newsletters, podcast, news articles and interviews

We communicated regularly about project updates utilizing blogs/news (published at the KInIT website) and newsletters (utilizing KInIT's newsletter with 600 subscribers). Due to the high scientific and societal relevance of our work, we also received multiple invitations to podcasts, news articles as well as interviews, both by national media, such as DennikN and SME, as well as international media, such as Süddeutsche Zeitung, Tages-Anzeiger or New Scientist.

Most of such communication was done in English language (to allow the maximum impact and audience to get reached). To a small extent the Slovak language was used (especially in news articles and podcasts targeting Slovak audience). Interestingly, some project outcomes were communicated also in German language (e.g., by a [news article](#) published by Süddeutsche Zeitung) and French language (by a [podcast](#) given during our attendance at the EU DisinfoLab conference 2025).

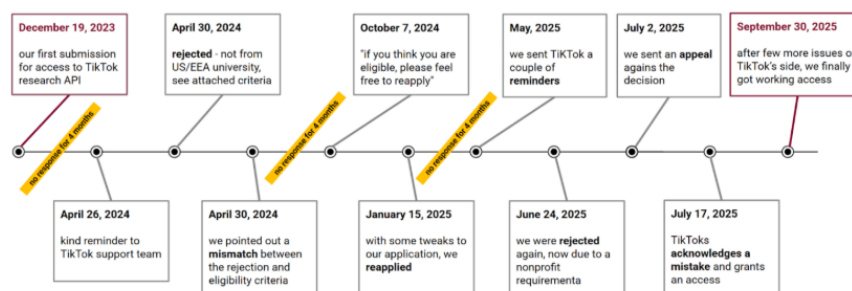
The following image presents an example of [blog](#) published at the KInIT website:



time (in January 2025).

Also it is weird that their eligibility criteria clearly states that academic institutions are eligible for Research API, however, in their communication **they consider universities as only eligible type of academic organizations** (following this argument, our first application was declined and even appeal did not work that time). Practically, **this narrow interpretation of academic organizations leaves out all independent non-profit research institutions**, including our one, completely preventing them from obtaining access to the Research API.

In total, **it took 21 months**, from December 19, 2023 until September 30, 2025, to get the first working data access from TikTok. It was a long journey and we are quite sure that many researchers are not having sufficient time and resources to undergo it. Not speaking about the fact that **when you finally get data access, it is not what you have asked for...**



While I have sent one more email to the TikTok team, kindly asking for an explanation why they confirmed us to have the Research API access and denied it shortly afterwards, I do not believe this will start working soon. Instead, as of 29 October 2025, the **Delegated act on data access under the Digital Services Act (DSA)** entered into force and opened new data access possibilities for so-called vetted researchers. We now give our hopes to this EU legislation that can significantly strengthen the position of researchers in the uneven fight with social media platforms.



Funded by the
European Union
NextGenerationEU

**[RECOVERY
AND RESILIENCE]
PLAN**

The following table provides a list of all blogs, newsletters, podcast, news articles and interviews.

Date	Title	Info	Language	Link
26.11.2024	Ocenená informatička Bieliková: Umelá inteligencia dáva neurčité výsledky zo svojej podstaty a mnohokrát ani nevieme prečo	News article - by DennikN	SK	link
18.02.2025	Holding Big Platforms Accountable: The Why, What, and How of Auditing	Blog post - by KInIT	EN	link
30.04.2025	KInIT Newsletter: Tackling disinformation at KInIT	Newsletter - by KInIT	EN	link
08.05.2025	Does YouTube fight misinformation videos? Analysing video removal over 3 years	Blog post - by KInIT	EN	link
30.05.2025	Algorithmic Audits on TikTok: Why Results are Hard to Replicate and Quickly Expire	Blog post - by KInIT	EN	link
17.06.2025	Hidden struggles of auditing social platforms: Why is creating user accounts so difficult?	Blog post - by KInIT	EN	link
30.6.2025	Why viral posts on social media might be hurting us the most	Blog post - by KInIT	EN	link

02.07.2025	Podcast Lifehack: Kto rozhoduje o tvojom feede? Algoritmy, ktoré poznajú tvoje slabosť	Podcast - by SME	SK	link
11.07.2025	When Research APIs Close the Door: Using Algorithmic Auditing as an Alternative Approach to Study Social Media	Blog post - by KInIT	EN	link
06.08.2025	Prečo je dôležité skúmať, čo nám odporúčajú sociálne siete?	Article - by Noviny Detskej Univerzity Komenského	SK	link
14.09.2025	Digitálny kompas PulseZ 02 // Mladí na vlnách umelej inteligencie: Rozhovor s Branislavom Pecherom z KInITu	Podcast - by PulseZ	SK	link
28.10.2025	Propagations - À la rencontre de l'écosystème anti-désinfo	Podcast - by Propagations	FR	link
10.11.2025	ECREA Workshop 2025: Why Platform Audits Need to Get Technical (And Why We Need to Work Together)	Blog post - by KInIT	EN	link
13.11.2025	KInIT Newsletter: What is algorithmic auditing?	Newsletter - by KInIT	EN	link
21.11.2025	Disinfo 2025 Conference: Two Problems, One Solution or How Algorithmic Auditing Addresses Data Access and Systemic Risks	Blog post - by KInIT	EN	link
19.12.2025	When Research APIs Close the Door: When you finally get it, it is not what you have asked for	Blog post - by KInIT	EN	link
02.02.2026	Prečo AI hlúpne	Podcast - by Dvaja Mensania	SK	link
26.01.2026	Vedkyňa Bieliková: V umelej inteligencii máme pomocníka. Geniálneho tínedžera	News article - by SME	SK	link
14.02.2026	Was die Forschung über den Tiktok-Algorithmus weiß	News article - by Süddeutsche Zeitung	DE	link
15.02.2026	TikTok: Was die Forschung über den Algorithmus weiss	News article - by Tages-Anzeiger	DE	link
12.03.2026	Undisclosed ads on TikTok skirt ban on profiling minors	News article - by New Scientist	EN	link
20.03.2026	Exploring TikTok's Virtual Compute Environment (VCE): How It Should Work	Blog post - by KInIT	EN	link
20.03.2026	Exploring TikTok's Virtual Compute Environment (VCE): A Critical Review of Limitations and Flaws	Blog post - by KInIT	EN	link
29.04.2026	KInIT Newsletter: AI-Auditology: The Struggle to Access TikTok	Newsletter - by KInIT	EN	link
07.05.2026	What do children see on TikTok?	Blog post - by KInIT	EN	link
28.05.2026	KInIT Newsletter: Audit of minors profiling in advertisement at TikTok	Newsletter - by KInIT	EN	link
23.06.2026	How Students Help (Us) Audit Social Media Recommendation Algorithms	Blog post - by KInIT	EN	link
30.06.2026	KInIT Newsletter: Pieces of a bigger puzzle	Newsletter - by KInIT	EN	link

2.4 Scientific publications

The project resulted in 15 scientific publications accepted, published and presented at top-level conferences, workshops, as well as scientific journals that focus on computer science.

In line with publication plan, we primarily targeted top-level A* or A conferences (according to the CORE rank) and Q1 journals. Namely the papers on effective (low-resource) content labelling were accepted to machine learning (ML) and natural language processing (NLP) venues: EMNLP 2024 (A*), NAACL 2025 (A), ACL 2025 (A*), EMNLP 2025 (A*), EACL 2026 (A) and ACM Transactions on Intelligent Systems and Technology (ACM TIST, Q1).

The description of individual audit studies conducted within the AI-Auditology project were accepted to the venues focused on information retrieval (IR), human-computer interaction (HCI), personalisation or recommender systems: SIGIR 2025 (A*), CHI 2026 (A*), UMAP (A) or FAccT 2026.

The description of the model-based algorithmic auditing, the primary novel concept introduced by the AI-Auditology project, was further presented at the EWAf 2025, which is an interdisciplinary workshop bringing together researchers in algorithmic fairness, including many stakeholders relevant to project's exploitation (see Section 3 for more details). In a similar way, the relevance of the project to the field of information security and data protection was presented by means of the paper accepted to the IS2 conference.

Out of accepted scientific publications, a notable recognition belongs to the FAccT 2026 paper that was awarded with the Best Paper Award (out of almost 1000 submitted papers and 325 accepted ones, three papers were recognized by this award).



In line with the open science spirit, also declared in the DCE and clustering plan (D5.2, Section 3.5), we supported openness and reproducibility of our research. All publications were published also as preprints (as a green open access) at arXiv.org repository. Moreover, as most of the publications were published by ACM Press or ACL publishers, also their final versions are available (gold open access under an open licence) in ACM digital library (ACM DL) and ACL Anthology, respectively. As a result, all scientific publications resulting from the AI-Auditology project are freely available to the scientific community as well as the general public.

The impact of AI-Auditology publications is proven also by their immediate adoption by a scientific community. Despite many publications being published only recently (the project duration was only 23 months), they have already received 171 citations (as of Google Scholar, June 2026).

The following table provides a list of all scientific publications, together with their venue, ranking, authors and contribution of the AI-Auditology project towards the paper (when the publication contained acknowledgement also to additional projects).

Title	Venue	CORE Rank / Journal Quartile	Date	Authors	AI-Auditology Contribution	Link
On Sensitivity of Learning with Limited Labelled Data to the Effects of Randomness: Impact of Interactions and Systematic Choices	EMNLP 2024	A*	November 12 - 16, 2024	Branislav Pecher, Ivan Srba, Maria Bielikova	0.7	link
LLMs vs Established Text Augmentation Techniques for Classification: When do the Benefits Outweight the Costs?	NAACL 2025	A	April 29 - May 4, 2025	Jan Cegin, Jakub Simko, Peter Brusilovsky	0.5	link
Model-based Algorithmic Auditing of Social Media AI Algorithms	EWAF 2025	-	June 30th - July 2nd, 2025	Ivan Srba, Branislav Pecher, Jakub Simko, Robert Moro, Maria Bielikova	1	link
Revisiting Algorithmic Audits of TikTok: Poor Reproducibility and Short-term Validity of Findings	SIGIR 2025	A*	July 13 - 18, 2025	Matej Mosnár, Adam Škurla, Branislav Pecher, Matúš Tibenský, Ján Jakubčík, Adrián Bindas, Peter Sakalík, Ivan Srba	0.9	link

Evaluation of LLM Vulnerabilities to Being Misused for Personalized Disinformation Generation	ACL 2025	A*	July 27 - August 1, 2025	Aneta Zucecova, Dominik Macko, Robert Moro, Ivan Srba, Jakub Kopal, Katarina Marcincinova, Matus Mesarcik	0.5	link
Use Random Selection for Now: Investigation of Few-Shot Selection Strategies in LLM-based Text Augmentation	EMNLP 2025	A*	November 5 - 9, 2025	Jan Cegin, Branislav Pecher, Jakub Simko, Ivan Srba, Maria Bielikova, Peter Brusilovsky	0.4	link
Comparing Specialised Small and General Large Language Models on Text Classification: 100 Labelled Samples to Achieve Break-Even Performance	EMNLP 2025	A*	November 5 - 9, 2025	Branislav Pecher, Ivan Srba, Maria Bielikova	0.4	link
A Rigorous Evaluation of LLM Data Generation Strategies for Low-Resource Languages	EMNLP 2025	A*	November 5 - 9, 2025	Tatiana Anikina, Jan Cegin, Simon Ostermann, Jakub Simko	0.2	link
A Survey on Automatic Credibility Assessment Using Textual Credibility Signals	ACM TIST	Q1	-	Ivan Srba, Robert Moro et al.	0.4	link
Better as Generators Than Classifiers: Leveraging LLMs and Synthetic Data for Low-Resource Multilingual Classification	EACL 2026	A	March 24 - 29, 2026	Branislav Pecher, Jan Cegin, Robert Belanec, Ivan Srba, Jakub Simko, Maria Bielikova	1	link
RoSE: Round-robin Synthetic Data Evaluation for Selecting LLM Generators without Human Test Sets	EACL 2026	A	March 24 - 29, 2026	Jan Cegin, Branislav Pecher, Ivan Srba, Jakub Simko	0.33	link
Beyond the Checkbox: Strengthening DSA Compliance Through Social Media Algorithmic Auditing	CHI 2026	A*	April 13 - 17, 2026	Sara Solarova, Matus Mesarcik, et al.	0.65	link
State of Play of Social Media	IS2	-	June 2 - 3,	Jakub Simko	1	link

Algorithmic Auditing: Fruits, Struggles, and Promises			2026			
Algorithmic Audit of Personalisation Drift in Polarising Topics on TikTok	UMAP 2026	A	June 8 - 11, 2026	Branislav Pecher, Adrián Bindas, Ján Jakubčík, Matus Tuna, Matus Tibensky, Šimon Liška, Peter Sakalík, Andrej Šutý, Matej Mosnár, Filip Hossner, Ivan Srba	0.5	link
The DSA's Blind Spot: Algorithmic Audit of Advertising and Minor Profiling on TikTok	FAccT 2026	-	June 25 - 28, 2026	Sara Solarova, Matej Mosnar, Jan Jakubcik, Matus Tibensky, Adrian Bindas, Simon Liska, Filip Hossner, Matúš Mesarčík, Ivan Srba	1	link

In addition, the project outcomes were summarised also in three preprints that are not accepted/published at the international venues yet.

Title	Authors	AI-Auditology Contribution
Ethics-based assessment as a cross-disciplinary approach towards responsible AI research	Katarína Marcinčinová, Sára Soľárová, Juraj Podroužek, Branislav Pecher, Ivan Srba	1
Overshoot: Taking advantage of future gradients for stochastic optimization	Jakub Kopal, Michal Gregor, Santiago de Leon, Jakub Simko	0.4
Political Leaning and Politicalness Classification of Texts	Matous Volf, Jakub Simko	1

2.5 Participation in academic and networking events

Within academic events, during the project implementation each of the accepted publications (Section 2.4) was presented either in-person or virtually at the corresponding conferences,

thus reaching the wide international scientific community which traditionally participated in these flagship conferences.

Besides that, the team members participated in a number of additional academic and networking events, especially at the international level. Among them:

- We presented the concept of algorithmic auditing to fight disinformation and FIMI content at EU Disinfo conference 2025, where our talk and participation in the control panel was selected from 200 proposals.



- We prepared and supervised the project entitled [The Algorithmic Mirror: Auditing TikTok Recommender Systems for Personalization, Protection, and Profiling](#) at the [Digital Methods Summer School and Data Sprint 2025](#), what is a well-known intensive event where participants research and develop internet research techniques, organised by University of Amsterdam.
- The most impactful event was participation at the AI Impact Summit session "AI and Children's Safety and Wellbeing". The AI Impact Summit 2026 was a premier global gathering hosted in New Delhi, India, that showcased how artificial intelligence can drive inclusive economic growth and social good. The invitation to this top-level venue was a follow-up of our previous in-person meeting with Amandeep Singh Gill, Under-Secretary-General and Special Envoy for Digital and Emerging Technologies, United Nations, during which we presented the AI-Auditology project.
- We also presented the AI-Auditology project at Thematic Seminar "Digital Environments: Risks, Harms, and Opportunities for Children and Youth" organised by the Council of Europe.

We also paid attention to national events organised in Slovakia, such as [Digital Youth Forum 2026](#) (organised by The National Coordination Center for Addressing Violence Against

Children – [NKS](#)), [Mind&Tech 2026](#) (organised by Slovak Technical University - STU) or [Night of Ideas 2026](#) (organised by French Institute in Slovakia).

The following table provides a list of all our active participations at academic and networking events.

Title	Venue	Date	Participants
Project-related discussions and clustering at EU Disinfo conference 2024	EU Disinfo conference 2024	October 9 - 10, 2024	Ivan Srba, Robert Moro
Project-related discussions, clustering and meeting with stakeholders at CEDMO conference 2024	CEDMO conference 2024	November 5, 2024	Ivan Srba, Jakub Simko, Maria Bielikova
Presentation of paper On Sensitivity of Learning with Limited Labelled Data to the Effects of Randomness: Impact of Interactions and Systematic Choices	EMNLP 2024	November 12 - 16, 2024	Branislav Pecher
Presentation of paper LLMs vs Established Text Augmentation Techniques for Classification: When do the Benefits Outweight the Costs?	NAACL 2025	April 29 - May 4, 2025	Jan Cegin
Presentation of AI-Auditology project at Fórum inžinierov a technikov Slovenska	FITS 2025	March 20, 2025	Adrian Bindas
DSA round table	ZEK	June 19, 2025	Ivan Srba, Sara Solarova
Presentation of paper Model-based Algorithmic Auditing of Social Media AI Algorithms	EWAF 2025	June 30 - July 2, 2025	Jakub Simko
Project tutoring and leadership at UvA Digital Methods Summer School and Data Sprint 2025	Digital Methods Summer School and Data Sprint 2025	June 30 - July 11, 2025	Filip Hossner, Adrian Bindas
Presentation of paper Revisiting Algorithmic Audits of TikTok: Poor Reproducibility and Short-term Validity of Findings	SIGIR 2025	July 13 - 18, 2025	Branislav Pecher, Adam Skurla
Project presentation at Workshop on Explainability in Information Retrieval (WExIR) colocated with SIGIR 2025	WEIR @ SIGIR 2025	July 17, 2025	Branislav Pecher
Presentation of paper Evaluation of LLM Vulnerabilities to Being Misused for Personalized Disinformation Generation	ACL 2025	July 27 - August 1, 2025	Ivan Srba, Dominik Macko
Presentation of abstract Beyond the Checkbox: Strengthening DSA Compliance Through Social Media Algorithmic Auditing	ECREA CLP 2025	September 18 - 19, 2025	Sara Solarova
Presentation of DSA audit analysis and project concept, project-related discussions and clustering at EU Disinfo conference 2025	EU Disinfo conference 2025	October 15 - 16, 2025	Ivan Srba, Sara Solarova

Participation in Techsummit 2025, panel title: "Kto stráži algoritmy? Bezpečné sociálne siete v dobe umelej inteligencie"	Techsummit 2025	October 28, 2025	Ivan Srba
Participation in Stratcom under pressure 2025 conference, panel title: "Technological aspects of StratCom and Fighting Disinformation"	Stratcom under pressure 2025	November 5, 2025	Maria Bielikova
Presentation of paper Use Random Selection for Now: Investigation of Few-Shot Selection Strategies in LLM-based Text Augmentation for Classification	EMNLP 2025	July 27 - August 1st, 2025	Jan Cegin
Presentation of paper Comparing Specialised Small and General Large Language Models on Text Classification: 100 Labelled Samples to Achieve Break-Even Performance	EMNLP 2025	July 27 - August 1st, 2025	Branislav Pecher
Presentation at panel at India AI Impact Summit 2026	India AI Impact Summit 2026	February 16-20, 2026	Maria Bielikova
Presentation of abstract at The DSA and Platform Regulation Conference 2026	The DSA and Platform Regulation Conference 2026	February 16-17, 2026	Matus Mesarcik
Presentation of projects results at Mind&Tech 2026 conference	Mind&Tech 2026	March 12, 2026	Branislav Pecher
Presentation at panel at IEEE Central European AI Summit 2026	AI summit 2026	March 21, 2026	Sara Solarova
Presentation of paper Better as Generators Than Classifiers: Leveraging LLMs and Synthetic Data for Low-Resource Multilingual Classification	EACL 2026	March 24 - 29, 2026	Branislav Pecher
Presentation of paper RoSE: Round-robin Synthetic Data Evaluation for Selecting LLM Generators without Human Test Sets	EACL 2026	March 24 - 29, 2026	Jan Cegin
Participation at Workshop on AI, Social Platforms & Democracy	Center for Environmental and Technology Ethics - Prague (CETE-P)	March 25 - 26, 2026	Jakub Simko
Presentation at EUDHIT Symposium: Designed for Harm, Fixed by a Ban? Rethinking Children Safety Online	EUDHIT symposium	April 9, 2026	Sara Solarova
Presentation of paper: Beyond the Checkbox: Strengthening DSA Compliance Through Social Media Algorithmic Auditing	CHI 2026	April 13 - 17, 2026	Sara Solarova
Presentation at Night of Ideas by French Institute	Night of	April 15, 2026	Robert Moro

	Ideas 2026		
Presentation at Fullbright Enrichment Seminar	Fullbright Enrichment Seminar	April 17, 2026	Ivan Srba
Presentation at Digital Youth Forum 2026	Digital Youth Forum 2026	May 27, 2026	Sara Solarova
Presentation of paper State of Play of Social Media Algorithmic Auditing: Fruits, Struggles, and Promises	IS2 2026	June 2 - 3, 2026	Jakub Simko
Presentation at Thematic Seminar "Digital Environments: Risks, Harms, and Opportunities for Children and Youth" organised by Council of Europe	-	June 2, 2026	Sara Solarova
Presentation of paper: Algorithmic Audit of Personalisation Drift in Polarising Topics on TikTok	UMAP 2026	June 8 - 11, 2026	Matus Tuna
Presentation of paper: The DSA's Blind Spot: Algorithmic Audit of Advertising and Minor Profiling on TikTok	FAccT 2026	June 25 - 28, 2026	Sara Solarova

2.6 Source code repositories and datasets

Where applicable, all scientific publications were accompanied with links to the source code repositories and datasets.

Whenever possible, we shared datasets created as a part of project research activities at Zenodo portal in an anonymized form following the FAIR principles, i.e., they are easily findable, freely accessible, published in standard interoperable formats, and reusable, such as:

- [Dataset for paper "Algorithmic Audit of Personalisation Drift in Polarising Topics on TikTok"](#)
- [Dataset for paper "The DSA's Blind Spot: Algorithmic Audit of Advertising and Minor Profiling on TikTok"](#)

In the same manner, we published source code with data processing code allowing replicability of the achieved results at GitHub, such as:

- <https://github.com/kinit-sk/ai-auditology-personalisation-drift-tiktok>
- <https://github.com/kinit-sk/ai-auditology-advertising-and-minor-profiling-tiktok>

By promoting openness of achieved results, we also aimed to demonstrate and support the credibility of auditing results.

3. Exploitation Outcomes

During preparation of DCE and clustering plan (D5.2), 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.

During the consequent project implementation, we explored all of these stakeholder categories and identified the highest potential especially in the first two of them.

In the case of the regulators at the EU/national level, watchdog organizations and NGOs focused on algorithmic auditing, we participated in a number of events, during which we prepared or presented policy and regulation recommendations and briefs (further described in Section 2.1).

In the case of the research and academic community, we recognized the potential exploitation pathway of the key exploitable results (KERs, as identified in D5.2, Section 4.2) by means of follow-up projects (further described in Section 2.2).

We are also in direct contact with the representatives of the TikTok platform, which is the platform we primarily focused on during the execution of the audit studies. They expressed interest especially in the results of our audit study assessing profiling of minors within recommendation of personalised communication.

3.1 Policy and regulation recommendations and briefs

Our analysis of stakeholders' needs and use cases (D5.1) showed a high relevance of the AI-Auditology and its KERs to policy and regulation authorities. Therefore during the project implementation, we explored this exploitation pathway by means of multiple activities.

First, we organised multiple in-person or virtual meetings with relevant stakeholders, namely

- the Slovak national representative on the Digital Services Act (DSA) implementation team at DG CONNECT,
- representatives of Joint Research Center (JRC), and particularly from the European Centre for Algorithmic Transparency, which is the most relevant organisation to the AI-Auditology focus,
- representatives from DSA implementation teams and case handlers,
- state secretaries from Slovakia and Austria,

- Amandeep Singh Gill, Under-Secretary-General and Special Envoy for Digital and Emerging Technologies, United Nations, which resulted in an invitation to present our work at India AI Impact Summit 2026.

Second, we participated in events where this group of stakeholders was present, namely:

- the EU Disinfo conference 2024 and 2025, where we also had a talk and participated in panel discussion,
- the DSA and Platform Regulation Conference 2026
- the Workshop on AI, Social Platforms & Democracy organised by Center for Environmental and Technology Ethics - Prague (CETE-P), which resulted into the policy brief advocating for stronger scrutiny of social media and their algorithms and recognising the algorithmic auditing as a possible solution, and
- the thematic seminar Digital Environments: Risks, Harms, and Opportunities for Children and Youth organised by Council of Europe.

Besides that, this exploitation pathway was also reflected in the research and development work itself.

- In the paper entitled [“Beyond the Checkbox: Strengthening DSA Compliance Through Social Media Algorithmic Auditing”](#) (accepted to CHI 2026, A* conference), we thoroughly analysed the existing independent audit reports, which the VLOPs must undergo to verify compliance with DSA requirements. Our analysis showed significant inconsistencies in methodologies and lack of technical depth when evaluating AI-powered systems. To enhance the depth, scale, and independence of compliance assessments, we proposed to employ algorithmic auditing to complement the traditional auditing approaches used by auditors.
- Furthermore, in the paper entitled [“The DSA's Blind Spot: Algorithmic Audit of Advertising and Minor Profiling on TikTok”](#), we conducted an audit study, which was specifically designed to assess the compliance of TikTok platform with the Article 28(2) of the DSA which prohibits profiling-based advertising to minors. Our findings revealed a stark regulatory paradox. TikTok demonstrated formal compliance with Article 28(2) by shielding minors from profiled formal advertisements, yet both disclosed and undisclosed ads exhibit significant profiling aligned with user interests (5-8 times stronger than for adult formal advertising). The strongest profiling emerged within undisclosed commercial content, where creators/brands failed to label paid partnership/promotional content and the platform neither corrects this omission nor prevents its personalized delivery to minors. These results demonstrated that minors remain exposed to algorithmically targeted commercial content through the same recommendation mechanisms the DSA seeks to constrain. We also provided policy recommendations advocating that protecting minors requires expanding the definition of advertisement in EU law to encompass influencer and brand promotional content, and ensuring that any such expansion is accompanied by a corresponding prohibition on profiling-based targeting of minors, so that commercial content cannot circumvent protections merely by operating outside formal advertising channels.

3.2 Follow-up projects

To tap on the potential of the exploitation pathway with the research and academic community, we not only presented the project outcomes (primarily research papers) at scientific conferences, but also actively discussed potential utilisation of the AI-Auditology KERs with researchers and practitioners.

At first, this resulted not only in an extended network of contacts within the research community, but we also organised mutual seminars to jointly introduce our work (e.g., we hosted a seminar provided by Masood, Maleeha from the University of Illinois).

Besides that, during the second half of the AI-Auditology project, we prepared several follow-up project proposals:

- Within the call “DIGITAL-2026-BESTUSE-TECH-EDMO-09-HUBS: European Digital Media Observatory” (Digital Europe Programme), we created a project proposal CEDMO 3.0, in which we will extend previous activities of CEDMO hub with algorithmic auditing of exemplification role of social media platforms’ algorithms during elections happening in Slovakia, Czech Republic and Poland.
- Within the call “PPPA-2026-SOCIAL-COHESION - Advancing Social Cohesion in the Face of Polarized Public Discourse”, we created a project proposal utilising algorithmic auditing to analyse online political discourse and polarisation across the EU.
- Within the call “ISF-2026-TF2-AG-CEEP” (Internal Security Fund - ISF), we created a project proposal focused on employing algorithmic auditing to contribute to preventing and countering online radicalisation, in which the social media platforms with their algorithmic systems play a crucial role.
- Within the call “HORIZON-CL5-2026-07-D1-04” (Horizon Europe), we created a proposal utilising algorithmic auditing to study and prevent spreading of disinformation related to climate change.
- Within the call “HORIZON-CL2-2026-01-DEMOCRACY-08” (Horizon Europe), we are currently preparing another proposal focused on electoral integrity in the digital context.

The list of these project proposals (out of which the first was already evaluated and selected for funding) clearly demonstrates our commitment to continue with algorithmic auditing research activities and exploit knowledge, research methodologies, findings, as well as developed infrastructure in the future research.

4. Evaluation of KPIs

The following table summarises the KPI stated in the project proposal (including their metrics) and the real project outcomes (the list of individual outcomes were provided in the previous Sections 2 and 3).

We can conclude that all KPIs were not successfully accomplished, but also overperformed the initial KPIs' values significantly.

Channel	KPIs Metric: Total (for 1 st and 2 nd half of the project duration)	Real project outcome Metric: Total (for 1 st and 2 nd half of the project duration)	Status
Project website	Page views/Impressions: ≥ 2000 (≥ 2000, ≥ 3000)	9551 (5049, 4502)	Successfully accomplished
Social media presence	Number of published posts: ≥ 70 (≥ 30, ≥ 40) Number of impressions: ≥ 49K (≥ 21K, ≥ 28K)	Number of published posts: 84 (36, 48) Number of impressions: 59315 (21373, 37942)	Successfully accomplished
Blogs, press releases, newsletters	Number of blogs/press releases/newsletters: ≥ 25 (≥ 10, ≥ 15)	Number of blogs/press releases/newsletters: 28 (9, 19)	Successfully accomplished
Participation in academic and networking events	Number of attendances: ≥ 12 (≥ 4, ≥ 8)	Number of attendances: 33 (8, 25)	Successfully accomplished
Scientific publications	Number of publications: ≥ 9 (≥ 3, ≥ 6)	Number of publications: 15 (4, 11) AI-Auditology contribution: 9.48 (3.1, 6.38) + 3 additional preprints with AI-Auditology contribution 2.4	Successfully accomplished

5. Conclusion

In this deliverable, we presented the report of DCE and clustering activities. A wide range of outreach – communication, dissemination and clustering – activities and actions have been undertaken during the project implementation.

The overall scope of such activities as well as their impact is outstanding:

- The AI-Auditology project resulted into **15 scientific publications**, 7 of them were published at **flagship A* conferences** (EMNLP, ACL, SIGIR, CHI), 1 article in **Q1 journal** (ACM TIST). Despite a short time from their publication, these publications have already received **171 citations** from the scientific community, demonstrating their practical impact. A reason for such an adoption is also an AI-Auditology **open science approach** (the publications were published within green open-access at arXiv) together with a **high reproducibility and replicability effort** (where relevant, the publications contain the links to the source code published at the Github repository and the collected and annotated datasets published at Zenodo).
- The high-impact of the resulting presentations is further confirmed by **the Best Paper Award recognition** of our paper accepted and presented at the FAccT 2026 conference.
- The AI-Auditology team participated in **33 academic and networking events**. Out of them, the special highlight is the invited talk at **AI Impact Summit 2026**.
- The project outcomes were actively presented to representatives from the **European Commission, Council of Europe, Joint Research Center, European Centre for Algorithmic Transparency, state secretaries**, as well as to Amandeep Singh Gill, Under-Secretary-General and Special Envoy for Digital and Emerging Technologies, **United Nations**.
- The AI-Auditology team also contributed to **regulation and policy advising**. The paper *Beyond the Checkbox: Strengthening DSA Compliance Through Social Media Algorithmic Auditing* (published at ACM CHI 2026) is directly focused on assessment of current audits prescribed by the DSA legislation. The award-winning paper *The DSA's Blind Spot: Algorithmic Audit of Advertising and Minor Profiling on TikTok* (published at ACM FAccT 2026) not only assess the compliance of VLOP with Article 28(2) of the Digital Service Act (DSA), but also provides 4 policy recommendations derived from the findings of the audit study. We also [took part in the ECREA Communication Law and Policy Workshop](#) in Brussels in September 2025, where we presented a need to extend the current DSA audits with more technical assessments, including algorithmic auditing. In addition, in cooperation with Center for Environmental and Technology Ethics - Prague (CETE-P), we contributed to the **policy brief** advocating for stronger scrutiny of social media and their algorithms and recognising algorithmic auditing as a possible solution.
- The AI-Auditology team participated in **diverse clustering activities** with other EU-funded projects, especially vera.ai, AI-CODE, and CEDMO. Besides a [participation at networking events](#) (such as Disinfo 2024 and 2025 conferences), a notable mention is a survey paper entitled *A Survey on Automatic Credibility Assessment Using Textual Credibility Signals* (published in ACM TIST journal), on which we cooperated with researchers and journalists from Horizon Europe vera.ai and AI-CODE projects.

- The AI-Auditology project also contributed to **educational activities**. Besides science popularisation (by means of various types of public appearances, such as interviews, podcasts), we [supervised high-school students](#) from the DELTA Secondary School of Informatics and Economics in Pardubice, Czech Republic, during their internship in KInIT in September 2025, and later also during their high school graduation thesis, with which they won the regional round of SOČ, which is the most prestigious high school scientific competition in the Czech Republic, and advanced to the national finals in which we achieved a second place.
- Furthermore, we are in direct contact with the **representatives from the TikTok platform** who expressed interest in our algorithmic auditing results.
- The project outcomes were communicated by means of **84 social media posts** (that received in total 59315 organic impressions) as well as **28 news articles, interviews, podcasts or newsletters**. Among them, the interest of international news media (such as Süddeutsche Zeitung) shows a global relevance of the undertaken work.