

# kinit

## V2.1 Správa o výsledkoch komunikácie, diseminácie a využitia výsledkov projektu

Názov projektu	Morálna citlivosť a ľudské práva pre spracovanie jazykov s obmedzenými zdrojmi
Akronym	SensAI
Kód projektu	09I01-03-V04-00100
Začiatok projektu	01. 1. 2025
Trvanie projektu	18 mesiacov



**PLÁN [OBNOVY]**

# Obsah

1 Úvod	3
2 Plán komunikácie a zdieľania	4
2.1 Cieľové skupiny	4
2.2 Komunikačné kanály, nástroje a aktivity	5
2.3 Otvorený prístup a manažment výskumných dát	5
2.4 Merateľné ukazovatele	5
3 Výsledky komunikácie, diseminácie a exploitácie	7
3.1 Webová stránka projektu	7
3.2 Sociálne médiá	10
3.3 Vedecké publikácie a prezentácie	13
3.4 Repozitáre zdrojových kódov a datasetov	18
3.5 Popularizačné udalosti	20
4 Záver	23

# 1 Úvod

Tento dokument predstavuje správu o výsledkoch komunikácie, diseminácie a exploitácie projektu SensAI. Súčasťou dokumentu je aj plán (Kap. 2), vytvorený v prvotných fázach projektu, ktorý mal zabezpečiť, že relevantné informácie o projekte a jeho výsledkoch budú dostupné cieľovým skupinám v správnom čase a dostupnej forme. Výsledky v Kap. 3 nadväzujú na tento plán s identifikáciou finálnych hodnôt merateľných ukazovateľov, spolu s opisom prípadných zmien oproti pôvodnému plánu.

## 2 Plán komunikácie a zdieľania

Na maximalizáciu dopadu výsledkov a výstupov projektu bol vytvorený plán komunikácie a zdieľania. Tento plán je prispôsobený potrebám jednotlivých cieľových skupín a má zabezpečiť efektívne šírenie informácií a dosahovanie požadovaných výsledkov projektu.

### 2.1 Cieľové skupiny

Komunikačné aktivity sú cieleňé na nasledovné skupiny:

**Tvorcovia systémov AI** – komunikácia výsledkov projektu smerujúca k výskumníkom, vývojárom a organizáciám vyvíjajúcim a nasadzujúcim systémy umelej inteligencie. Projekt im poskytuje poznatky a metodické odporúčania pre zodpovedný vývoj AI, podporuje implementáciu princípov dôveryhodnej AI a prispieva k znižovaniu technologickej medzery pri vývoji riešení pre jazyky s obmedzenými zdrojmi prostredníctvom otvorených dát a jazykových modelov.

**Experti na dôveryhodnú AI** – komunikácia výstupov projektu určených odborníkom pôsobiacim v oblasti AI etiky, AI governance a hodnotenia dôveryhodnosti systémov umelej inteligencie. Výsledky projektu prinášajú nové poznatky o hodnotení dopadov AI na základné práva, modelovaní morálnej citlivosti tvorcov AI a metodických prístupoch podporujúcich zodpovedný vývoj a nasadzovanie AI systémov.

**Regulačné orgány a tvorcovia verejných politík** – šírenie poznatkov relevantných pre implementáciu pripravovaných regulačných rámcov v oblasti umelej inteligencie, najmä v kontexte posudzovania dopadov AI systémov na základné ľudské práva, podpory dôveryhodnej AI a praktickej aplikácie požiadaviek európskej regulácie.

**Verejnoscť** – zvyšovanie povedomia o prínosoch výskumu v oblasti dôveryhodnej umelej inteligencie a spracovania jazykov s obmedzenými zdrojmi. Projekt prispieva k lepšej informovanosti o etických aspektoch AI a podporuje širšiu dostupnosť moderných jazykových technológií pre slovenský jazyk, čím vytvára predpoklady pre inkluzívnejšie a dôveryhodnejšie využívanie umelej inteligencie v spoločnosti.

## 2.2 Komunikačné kanály, nástroje a aktivity

Na efektívne šírenie informácií k uvedeným cieľovým skupinám sú použité nasledovné komunikačné kanály:

- **Webová podstránka projektu** – obsahuje ucelený opis projektu vytvoreného na jeho začiatku a doplnený o výsledky projektu. Web poslúži aj ako podklad pre iné komunikačné aktivity, napr. sociálne médiá ako primárny zdroj interakcií s cieľovými skupinami.
- **Sociálne médiá** – obsahujú aktuality o projekte predovšetkým cez sociálne profily KlnIT-u (LinkedIn, Facebook, X), ako aj cez účty partnerských projektov (predovšetkým ALFIE).
- **Vedecké publikácie a prezentácie** – Dosiahnuté výskumné výsledky diseminujeme prostredníctvom publikácií na uznávaných fórach (A\*/A konferencie, Q1/Q2 časopisy). Šírenie výsledkov podporíme aktívnou účasťou na vedeckých podujatiach.
- **Repozitáre zdrojových kódov a datasetov** – softvérové artefakty, modely a dátové vzorky zverejníme prostredníctvom vhodných portálov: GitHub, Huggingface, Zenodo.
- **Popularizačné udalosti** – Prezentácie na udalostiach ako Európska noc vedy alebo Deň Európy pomôžu dostať výsledky projektu do povedomia širšej verejnosti, čím zároveň prispejeme k popularizácii vedy a výskumu.

## 2.3 Otvorený prístup a manažment výskumných dát

Pri vedeckých publikáciách uprednostňujeme politiku otvoreného prístupu, t. j. preferencia takých konferencií a časopisov, ktoré aj v rámci najvyššej kvality (A/A\* CORE konferencie, Q1/Q2 časopisy) umožňujú voľný prístup k publikovaným článkom. Projekt zverejňuje dáta v súlade s princípmi FAIR (angl. Findable, Accessible, Interoperable, Reusable). Dáta budú ukladané do otvorených repozitárov (napr. Zenodo), pokiaľ to neodporuje komerčným záujmom alebo ochrane duševného vlastníctva.

## 2.4 Merateľné ukazovatele

V Tab. 1 je prehľadne zachytená štruktúra nástrojov, aktivít, cieľových skupín a kľúčových ukazovateľov (KPI). V porovnaní s opisom projektu v rámci žiadosti sme KPI rozdelili na KPI

výstupov, ktoré predstavujú priamo výstupy a výsledky projektu, a KPI dopadu, ktoré predstavujú odozvu cieľových skupín. KPI výstupy predstavujú **hlavné merateľné ukazovatele**, ktoré merajú množstvo práce vykonanej v rámci projektu. KPI dopadu sú **vedľajšie merateľné ukazovatele**, ktoré sú len nepriamo ovplyvniteľné množstvom práce v rámci projektu a závisia aj od externých vplyvov, ako napr. algoritmy sociálnych sietí.

Cieľové skupiny	Nástroj + aktivita	Plán KPI výstupov	Plán KPI dopadu
Výskumníci v oblasti AI, Spoločnosti vyvíjajúce a nasadzujúce systémy AI, Experti v oblasti dôveryhodnej AI, Regulátori v digitálnych technológiách, Verejnosť	Webová stránka projektu	-	# návštev webového sídla $\geq 500$ ;
Výskumníci v oblasti AI (X, LI), Spoločnosti vyvíjajúce a nasadzujúce systémy AI (Facebook, X, LI), Experti dôveryhodnej AI (X, LI), Verejnosť (Facebook)	Sociálne médiá	# postov $\geq 10$ ;	# interakcií $\geq 100$ ; # impresií $\geq 3\,000$ ;
Výskumníci v oblasti AI	Vedecké publikácie a prezentácie	# vedeckých článkov $\geq 2$ ; # odborných prezentácií $\geq 4$	# citácií (2 roky po skončení projektu) $\geq 20$ ;
Výskumníci v oblasti AI	Repozitáre zdrojových kódov a datasetov.	# zverejnených modelov $\geq 1$ ; # zverejnených dátových vzoriek $\geq 1$ ;	# stiahnutí artefaktov $\geq 10$ ;
Verejnosť	Popularizačné udalosti	# účastí na popularizačných udalostiach $\geq 2$ ;	

**Tab. 1** Prehľad plánovaných KPI.

### 3 Výsledky komunikácie, diseminácie a exploitácie

Táto kapitola obsahuje výsledky aktivít a výstupov, spolu s identifikáciou dosiahnutých minimálnych (niektoré KPI dopadu ešte rastú) hodnôt merateľných ukazovateľov, ktoré boli dostupné v čase prípravy tejto správy.

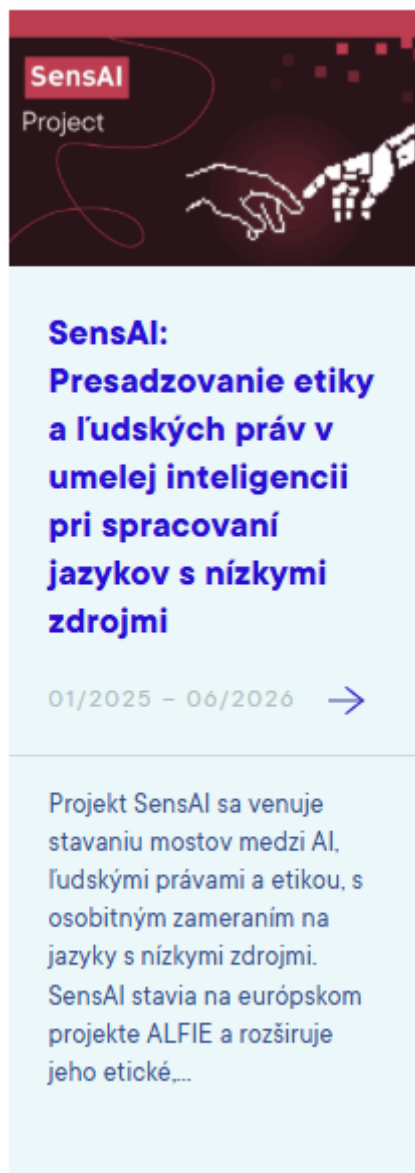
#### 3.1 Webová stránka projektu

V Tab. 2 sú prehľadne zobrazené plánované a dosiahnuté KPI projektu.

Cieľové skupiny	Plán KPI výstupov	Plán KPI dopadu	Dosiahnuté KPI výstupov	Dosiahnuté KPI dopadu
Výskumníci v oblasti AI, Spoločnosti vyvíjajúce a nasadzujúce systémy AI, Expertí v oblasti dôveryhodnej AI, Regulátori v digitálnych technológiách, Verejnosť	-	# návštev webového sídla >= 500;	-	# návštev webového sídla = 608;

**Tab. 2** Prehľad KPI pre webovú stránku.

Projekt SensAI bol prezentovaný cez webové stránky KInIT viacerými formami. Za prvé cez kartu projektu dostupnú na stránke <https://kinit.sk/research/projects/> (Obrázok 1).



**Obr. 1** Webová stránka s kartou projektu SensAI.

Za druhé, cez podstránku projektu SensAI, ktorá bola umiestnená v rámci webového sídla KInIT-u

<https://kinit.sk/sk/projekt/sensai-presadzovanie-ai-etiky-pri-spracovani-jazykov-s-nizkymi-zdrojmi/> (Obrázok 2). Pre stránky projektu SensAI boli vytvorené dve jazykové mutácie, v slovenčine a angličtine, periodicky aktualizované o nové výstupy.

**kinit** O nás Výskum Vzdelávanie Careers Novinky Partneri Kontakt Zapojte sa EN

Home > Research > Ethics and Human Values in Technology > SensAI: Presadzovanie etiky a ľudských práv v umelej inteligencii pri spracovaní jazykov s nízkymi zdrojmi

# SensAI Project

**PROJECT**

Duration:  
01/2025 - 06/2026

Funding agency: Plan obnovy

Project type: Scientific project

Principal Investigator:  
Jurej Podroužek

## SensAI: Presadzovanie etiky a ľudských práv v umelej inteligencii pri spracovaní jazykov s nízkymi zdrojmi

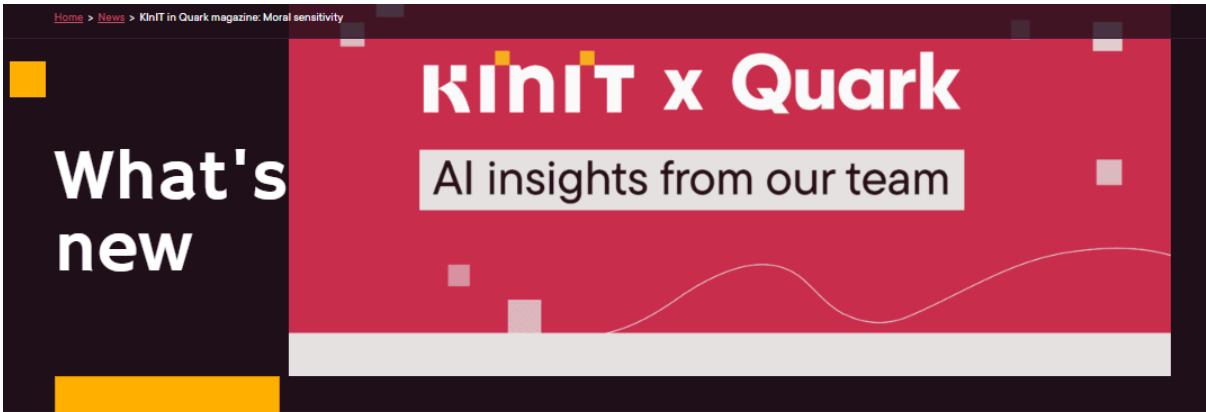
Projekt SensAI sa venuje stavaniu mostov medzi AI, ľudskými právami a etikou, s osobitným zameraním na jazyky s nízkymi zdrojmi. SensAI stavia na európskom projekte ALFIE a rozširuje jeho etické, právne a technické rozmery, aby zvýšil morálnu citlivosť pri vývoji AI a zlepšil súlad technológie s princípmi ľudských práv. Vývoj umelej inteligencie musí byť nielen technologicky vyspelý, ale aj transparentný, spravodlivý a široko prístupný. SensAI sa preto zameriava na vývoj metodológií na hodnotenie vplyvu AI na ľudské práva, zlepšenie efektívnosti veľkých jazykových modelov pre slovenčinu a iné jazyky s nízkymi zdrojmi a na zlepšenie rámca morálnej citlivosti pre vývojárov umelej inteligencie. Projekt má za cieľ urobiť AI inkluzívnejšou, zodpovednejšou a zabezpečiť, aby boli systémy v súlade so základnými ľudskými hodnotami.

Hlavným cieľom SensAI je zabezpečiť, aby boli systémy AI etické a prístupné, najmä v kontexte jazykov s nízkymi zdrojmi, ako je slovenčina. Projekt rieši tri vzájomne prepojené výzvy: morálnu citlivosť pri vývoji AI, hodnotenie vplyvu na ľudské práva a zlepšenie dostupnosti AI pre nedostatočne zastúpené jazyky. Posúdením toho, ako odborníci na umelú inteligenciu rozpoznávajú a riešia etické dilemy, rozšírením hodnotenia vplyvu na základné práva (FRIA) na aplikácie umelej inteligencie a vylepšením veľkých jazykových modelov pre slovenčinu, sa SensAI zameriava na vývoj metodológií, ktoré vkladajú do systémov AI etické hľadiská, transparentnosť a spravodlivosť. Tento výskum je kľúčový pre podporu dôveryhodnej AI, ktorá je v súlade so spoločenskými hodnotami a regulačnými rámcami.

- ◆ **Hodnotenie vplyvu AI na ľudské práva** – Projekt rozširuje hodnotenie vplyvu na základné práva o analýzu etických a právnych dôsledkov AI, najmä systémov automatizovaného strojového učenia (AutoML) a zabezpečuje súlad s legislatívou EÚ o umelej inteligencii.
- ◆ **Vylepšovanie AI pre jazyky s nízkymi zdrojmi** – SensAI vylepšuje veľké jazykové modely pre slovenčinu, čím sprístupňuje AI nástroje ľuďom, ktorí hovoria nedostatočne zastúpenými jazykmi a zároveň hodnotí ich výkon v porovnaní s viacjazyčnými AI prístupmi.
- ◆ **Morálna citlivosť vo AI vývoji** – Štúdiom etického povedomia v AI tímoch projekt vytvára rámce, ktoré posilňujú schopnosť vývojárov identifikovať a zmierňovať etické riziká v ich práci.

**Obr. 2** Webová podstránka projektu SensAI.

Okrem toho sme počas trvania projektu publikovali aj blogy, ktoré sa priamo viazali k výskumným témam a výstupom projektu, ako napríklad blog k morálnej citlivosti <https://kinit.sk/sk/kinit-v-casopise-quark-moralna-citlivost/>, ktorý vyšiel aj v časopise Quark a ktorý mal vlastnú projektovú podstránku (Obrázok 3).



AUTHOR



Kampelen Institute of Intelligent Technologies

Share

BACK TO NEWS

DEC 8, 2025

## KINIT in Quark magazine: Moral sensitivity

We continue our partnership with the [Quark](#) magazine and bring you yet another interesting article from our series exploring intelligent technologies. Quark is a Slovak magazine focused on the fields of science and technology. You can find KINIT's articles every month in the print version of the magazine, as well as online at the [Quark](#) website.

In December, we contributed to the magazine issue with the article [Moral sensitivity](#). In this article, our ethics specialist [Adrián Gavorník](#) and research intern [Lucia Hermanová](#) explore our responsibility as humans to apply ethical AI.



Obr. 3 Webová podstránka blogu SensAI.

Počet návštev webového sídla (karta projektu, podstránka projektu a blog) dosiahol počas trvania projektu číslo 608 návštev a splnil tak stanovené KPI dopadu.

## 3.2 Sociálne médiá



V Tab. 3 sú prehľadne zobrazené plánované a dosiahnuté KPI projektu.

Cieľové skupiny	Plán KPI výstupov	Plán KPI dopadu	Dosiahnuté KPI výstupov	Dosiahnuté KPI dopadu
Výskumníci v oblasti AI (X, LI), Spoločnosti vyvíjajúce nasadzujúce systémy AI (Facebook, X, LI), Experti dôveryhodnej AI (X, LI), Verejnosť (Facebook), Výskumníci v oblasti AI	# postov >= 12;	# interakcií >= 100; # impresií >= 3 000;	# postov = 17;	# interakcií = 1799; # impresií = 8449;

**Tab. 3** Prehľad KPI pre sociálne médiá.

Pre prezentáciu výsledkov projektu na sociálnych sieťach sme využili existujúce organizačné účty KlnIT na LinkedIn (6 000 sledovateľov), Facebooku (1 800 sledovateľov) a X/Twitteri (224 sledovateľov) ako aj nové komunikačné kanály ako BlueSky, ktorý postupne prebral rolu komunikačnej platformy namiesto X. Počtom postov zverejnených na sociálnych sieťach sa naplnil nad stanovený plán. Navyše sa podarilo dosiahnuť o viac než dvojnásobok plánovaný počet impresií. Počet interakcií taktiež výrazne prekročil plánovaný počet interakcií. Vysvetlením vysokého počtu interakcií môže byť okrem väčšieho počtu príspevkov aj ich cielene prioritne na LinkedIn v anglickom jazyku, kde sme zasiahli úzko špecializovanú odbornú verejnosť, čo malo pre projekt SensAI vysokú pridanú hodnotu.

V Tab. 4 sú prehľadne zobrazené publikované príspevky na sociálnych sieťach LinkedIn, X, Facebook a Bluesky, spolu s dátumom zverejnenia a odkazmi na pôvodný príspevok. Dáta za BlueSky bohužiaľ nie sú k dispozícii.

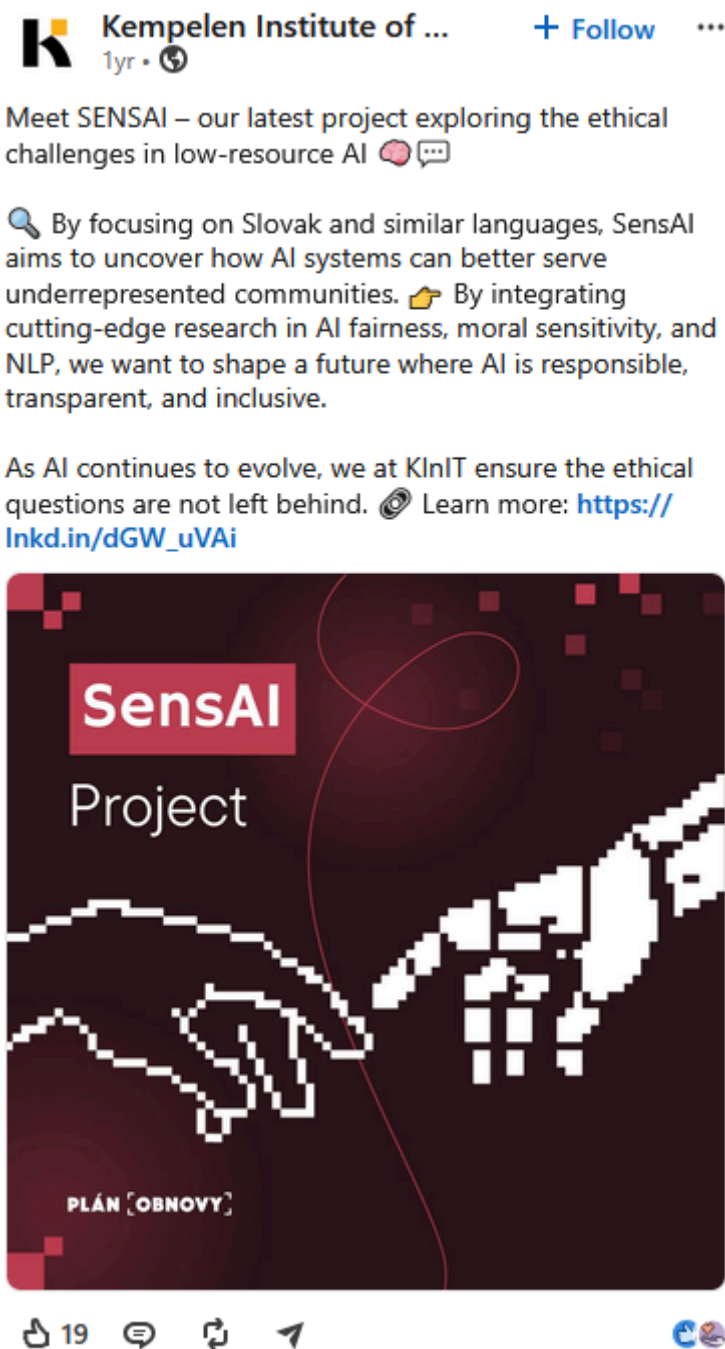
Dátum	Jazyk	Krátky opis	Info	Link	Impresie	Interakcie
06/05/2025	EN	Meet SENS AI – our latest project exploring the ethical challenges in low-resource AI 	post predstavenie projektu SensAI	<a href="#">LI</a>	748	43
06/05/2025	EN	 Focusing on Slovak and similar languages, SensAI project uncovers how AI can better serve underrepresented communities.	post predstavenie projektu SensAI	<a href="#">X</a>	64	0

08/28/2025	EN	What is AI ethics really about?	post predstavenie konceptu ethics AI	-	<a href="#">LI</a>	557	20
08/28/2025	EN	What is AI ethics really about?	post predstavenie konceptu ethics AI	-	<a href="#">BSKY</a>	0	0
08/28/2025	EN	What is AI ethics really about?	post predstavenie konceptu ethics AI	-	<a href="#">FB</a>	387	7
08/28/2025	EN	What is AI ethics really about?	post predstavenie konceptu ethics AI	-	<a href="#">X</a>	70	4
12/8/2025	EN	Morálna citlivosť	post - Quark blog		<a href="#">link</a>		
12/19/2025	EN	AI systems increasingly affect real people, their rights, privacy, and dignity.	post predstavenie FRIA	-	<a href="#">LI</a>	789	235
12/29/2025	EN	In the previous carousel, we asked why Fundamental Rights Impact Assessments can no longer be ignored in the age of AI.	post predstavenie FRIA	-	<a href="#">LI</a>	1368	587
02/11/2026	EN	Aristotle still has something to teach us about AI. 🏛️	post - virtue ethics		<a href="#">LI</a>	711	239
02/11/2026	EN	Virtue ethics meets AI!	post - virtue ethics		<a href="#">LI</a>	379	0
02/12/2026	EN	Aristotle still has something to teach us about AI. 🏛️	post - Virtue ethics a AI		<a href="#">BSKY</a>	0	0
03/10/2026	EN	🤖 A lot of AI systems work remarkably well, as long as the conversation stays in English. 🇬🇧	post fungovanie v Slovenčine	-	<a href="#">LI</a>	912	383
03/11/2026	EN	🤖 A lot of AI systems work remarkably well, as long as the conversation stays in English. 🇬🇧	post fungovanie v Slovenčine	-	<a href="#">BSKY</a>	0	0
04/13/2026	SK	Even a simple image of salmon can show you how far AI has come. 🐟	post prezentácia od Martina Tamajku	-	<a href="#">LI</a>	1088	183
04/13/2026	EN	Aj obyčajný obrázok s lososmi môže ukázať, kam sa posunula AI. 🐟	post prezentácia od Martina Tamajku	-	<a href="#">FB</a>	1376	98

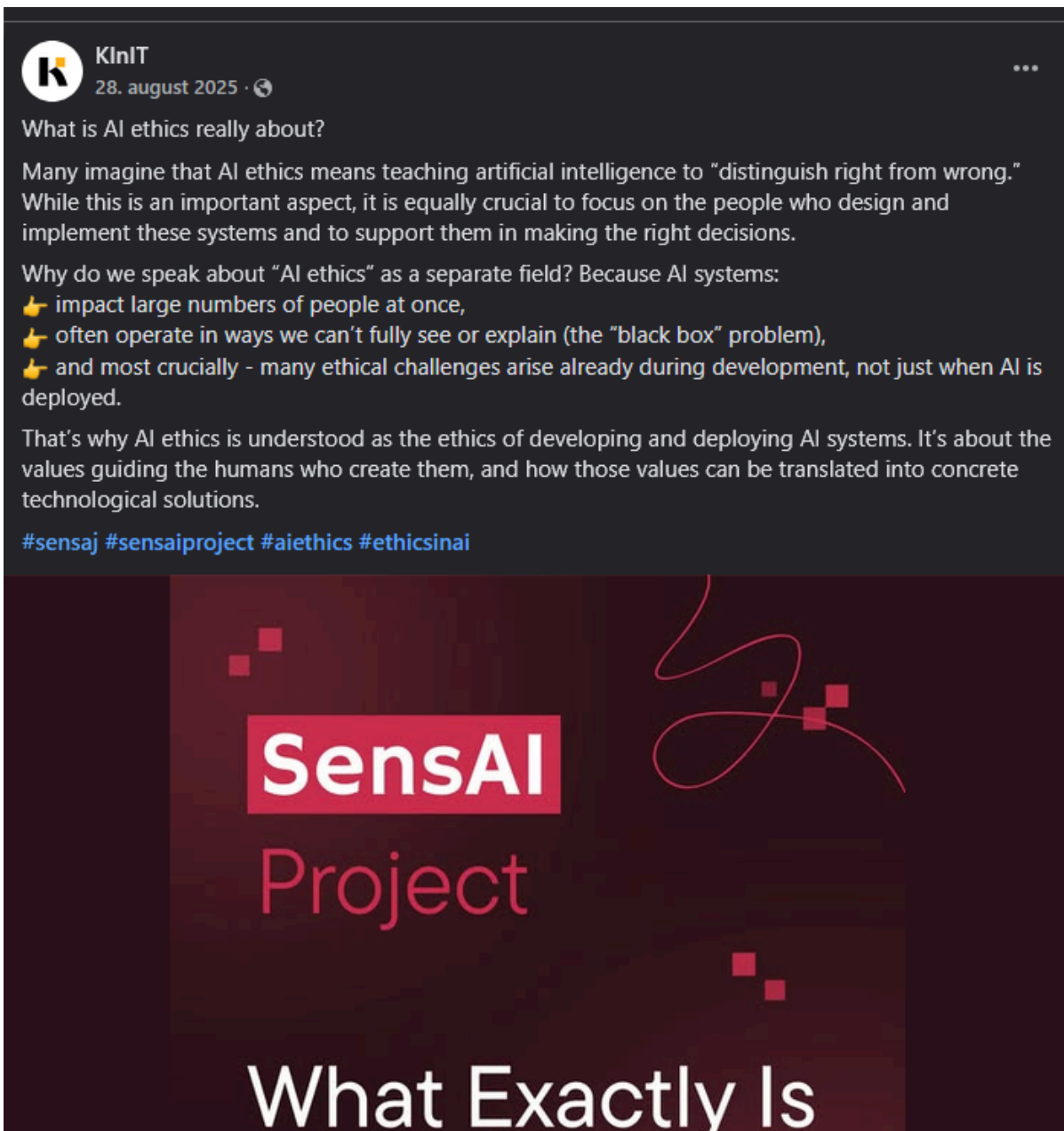
05/26/2026	Remember when AI generated salmon steaks floating in a river? 🐟	post - prezentácia od Martina Tamajku <a href="#">BSKY</a>		0	0
------------	---	--	--	---	---

**Tab. 4** Zoznam publikovaných príspevkov pre sociálne médiá.

Na Obr. 3 a Obr. 4 sú zobrazené príklady publikovaných príspevkov týkajúcich sa projektu SensAI.



**Obr. 4** Príklad príspevku na sociálnych sieťach s animovanou grafikou projektu SensAI.



**Obr. 5** Príklad príspevku na sociálnych sieťach o predstavení AI etiky v rámci projektu SensAI.

### 3.3 Vedecké publikácie a prezentácie

V Tab. 5 sú prehľadne zobrazené plánované a dosiahnuté KPI projektu.

Cieľové skupiny	Plán KPI výstupov	Plán KPI dopadu	Dosiahnuté KPI výstupov	Dosiahnuté KPI dopadu
Výskumníci v oblasti AI	# vedeckých článkov $\geq 2$ ;	# citácií (2 roky po skončení projektu) $\geq 20$ ;	# vedeckých článkov = 3;	# citácií (2 roky po skončení projektu) $\geq X$ ;

	# odborných prezentácií >= 4		# odborných prezentácií = 6;	
--	------------------------------	--	------------------------------	--

**Tab. 5** Prehľad KPI pre vedecké publikácie a prezentácie.

V rámci projektu boli pripravené štyri vedecké články s celkovým prínosom 3 pre projekt SensAI a to z oblastí, ktoré priamo nadväzujú na výskumné témy projektu. V súlade s projektovým zámerom sme sa sústredili na uznávané fóra (A\*/A konferencie, Q1/Q2 casopisy). A v súlade s našim zámerom podporiť otvorenosť a reprodukovateľnosť vedeckého výskumu boli všetky publikácie zverejnené aj ako preprinty v repozitároch [arXiv.org](https://arxiv.org) a [ssrn.com](https://ssrn.com).

Napríklad vo výskume v oblasti spracovania prirodzeného jazyka sme pripravili príspevok pre prestížnu konferenciu European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases ako aj pre uznávaný časopis ACM Transactions on Intelligent Systems and Technology.

Pre časopis Science and Engineering Ethics je pripravený aj vedecký článok zameraný na konceptualizáciu morálnej citlivosti tvorcov systémov AI a jej ukotvenie vo vzťahu k etike cností, praktickej múdrosti a existujúcim modelom morálnych kompetencií. Zaujímavosťou je, že vývoj článku si vyžadoval rozsiahlejšiu konceptuálnu analýzu, než sa pôvodne predpokladalo, vrátane konzultácií so zahraničnými odborníkmi a priebežného spresňovania teoretického rámca.

Nasledujúca tabuľka poskytuje zoznam všetkých vedeckých publikácií, hodnotením a príspevkom do projektu SensAI (v prípadoch, keď publikácia obsahovala poďakovanie aj ďalším projektom).

Názov	Fórum	Rank	Podiel SensAI	Link
Sebastian Kula, Martin Tamajka. 2026. Multi-Agent System Leveraging Open-Source LLMs to Mitigate Disinformation Threats	European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (Workshop on Security, Human Awareness, and Risk Mitigation in AI-driven Systems)	A*	0.9	<a href="https://arxiv.org/abs/2606.30259">https://arxiv.org/abs/2606.30259</a> X
Jakub Betinský, Mária Kolesárová, Adrián Gavorník, Juraj Podroužek. Moral Sensitivity and Practical	Science and Engineering Ethics	Q1	1.0	<a href="https://papers.ssrn.com/sol3/papers.cfm?abstract_id=70">https://papers.ssrn.com/sol3/papers.cfm?abstract_id=70</a>

Wisdom: Positioning a Broader Conception of Moral Sensitivity within the Virtue and Wisdom Traditions.				<a href="https://arxiv.org/abs/18198&amp;cf_chl_f_tk=U0_qDnIfgmURMW7sIBkQM0HJeMicTeyOmUaqsxdCSE4-1782741428-1.0.1.1-t8yb4S9zTCm6ARqL17iudDxflg4o.djaKMWQNFle10Q">18198&amp; cf c hl f tk=U0 q DnIfgmURMW 7sIBkQM0HJe MicTeyOmUa qsxdCSE4-17 82741428-1.0. 1.1-t8yb4S9zT Cm6ARqL17iu dDxflg4o.djaK MWQNFle10 Q</a>
Ivan Vykopal, Matúš Pikuliak, Simon Ostermann, Marián Šimko. Generative Large Language Models for Automated Fact-Checking: A Survey.	<a href="https://arxiv.org/abs/2407.02351">ACM Transactions on Intelligent Systems and Technology</a>	Q1	0.8	<a href="https://arxiv.org/abs/2407.02351">https://arxiv.org/abs/2407.02351</a>
Sara Solarova, Matúš Mesarčík, Branislav Pecher, and Ivan Srba: Beyond the Checkbox: Strengthening DSA Compliance Through Social Media Algorithmic Auditing	<a href="https://dl.acm.org/doi/10.1145/3772318.3791774">CHI '26: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems</a>	A*	0.3	<a href="https://dl.acm.org/doi/10.1145/3772318.3791774">https://dl.acm.org/doi/10.1145/3772318.3791774</a>

**Tab. 6** Zoznam vedeckých článkov.

Na Obr. 6 je zobrazený príklad poďakovania projektu SensAI v rámci konkrétneho vedeckého článku. Aj pri mierne upravenom doslovnom poďakovaní je v každom článku korektné uvedené číslo projektu. V uvedenom príklade bol SensAI jediným projektom financujúcim výskum (a teda plný príspevok do uvedeného počtu vedeckých článkov).

### **Acknowledgements**

This work was supported by the SensAI project funded by the NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I01-03-V04-00100.

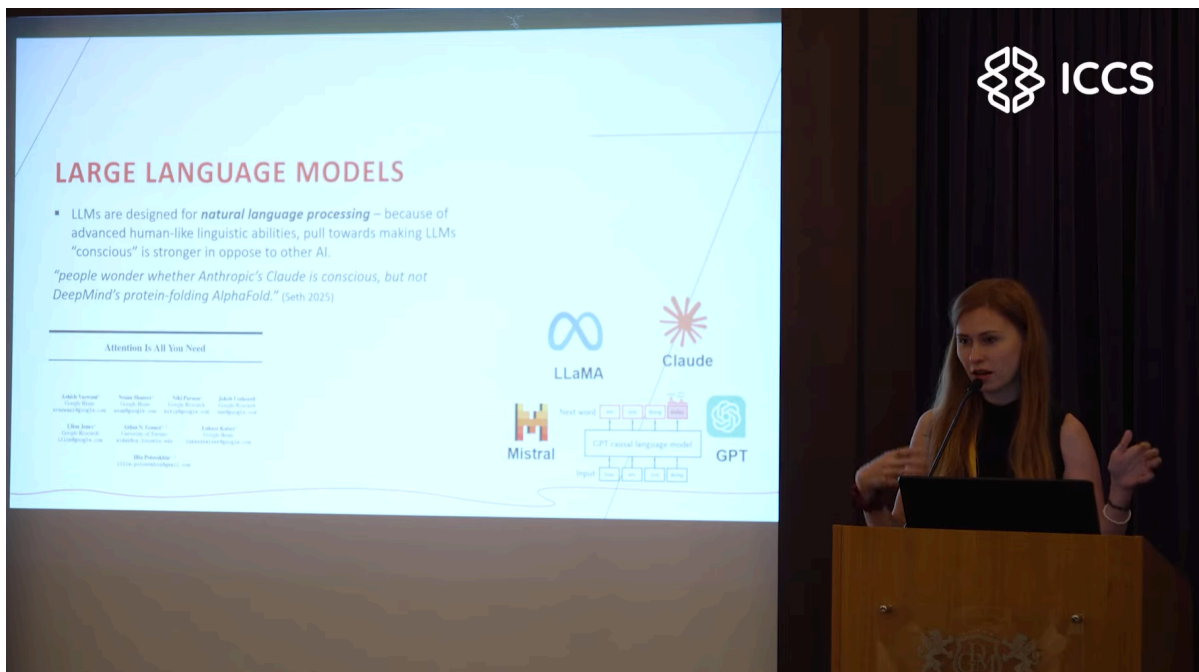
**Obr. 6** Príklad poďakovania projektu SensAI za financovanie výskumu v rámci publikovaného článku.

Šírenie výsledkov projektu SensAI sme podporili aj aktívnou účasťou na vedeckých podujatiach. Zoznam v Tab. 7 obsahuje prehľad odborných prezentácií výsledkov projektu, vrátane série prednášok na prestížnej European Summer School on Artificial Intelligence, kde sme aktívne pracovali s pojmom morálnej citlivosti v AI praxi.

Fórum	Nadpis	Forma
ETHICOMP2025: 22nd International Conference on the Ethical and Social Impacts of ICT	Adrian Gavornik, Juraj Podrouzek. A Framework for Evaluating AI Ethics Tools: The Role of Moral Sensitivity in Practice	prednáška
The Second ICCS Conference "AI and Sentience"	Katarína Marcinčinová. Positive aspect of illusionism and its implications for AI sentience	prednáška
Laitech 2026 Student Conference – Tallinn	Natalia Slosiarova: From Case Law to AI Governance: Unpacking the Concept of Mental Integrity	prednáška
The DSA and Platform Regulation Conference 2026	Matúš Mesarčík: Bridging the Gap Between Legal Intent and Algorithmic Reality. Empirical Evidence of Profiling-Based Commercial Communications to Minors on Social Media Platforms	prednáška
6th International Conference on Philosophy of Mind: Artificial Intelligence (6ICPH)	Katarína Marcinčinová, Daniela Vacek. Birch's centrism, anthropomorphism, and AI roles	prednáška
3rd European Summer School on Artificial Intelligence 2025	Matúš Mesarčík, Juraj Podroužek. Ethics and Law in Trustworthy AI: Foundations and Applications	séria prednášok (kurz)

**Tab. 7** Zoznam odborných prezentácií.

Obr. 7 a Obr. 8 zobrazujú príklady prezentácií výsledkov SensAI projektu na rôznych fórach.



**Obr. 7** Pozvaná prednáška zahŕňajúca priebežné výsledky projektu SensAI



**Obr. 8** Séria prednášok zahŕňajúcich výsledky projektu SensAI k dopadu systémov AI na základné ľudské práva a k morálnej citlivosti tvorcov systémov umelej inteligencie.

### 3.4 Repozitáre zdrojových kódov a datasetov

V Tab. 8 sú prehľadne zobrazené plánované a dosiahnuté KPI projektu.

Cieľové skupiny	Plán KPI výstupov	Plán KPI dopadu	Dosiahnuté KPI výstupov	Dosiahnuté KPI dopadu
Výskumníci v oblasti AI Expertí v oblasti dezinformácií	# <b>zverejnených modelov</b> >= 1; # <b>zverejnených dátových vzoriek</b> >= 1;	# stiahnutí artefaktov ≥ 10;	# <b>zverejnených modelov</b> = 5; # <b>zverejnených dátových vzoriek</b> = 2;	# <b>stiahnutí artefaktov</b> = 139;

**Tab. 8** Prehľad KPI pre repozitáre zdrojových kódov a datasetov.

Počty zverejnených artefaktov zodpovedajú plánu. Celkovo boli zverejnené dve nové, unikátne dátové sady a päť dotrénovaných modelov. Artefakty sú zverejnené v súlade s princípmi otvorenej vedy na úložisku HuggingFace v rámci spoločnej kolekcie<sup>1</sup> (Tab. 9).

Typ	Názov	Link
Kolekcia	SensAI (Metadata Extraction)	<a href="https://huggingface.co/collections/kinit/sensai-metadata-extraction">https://huggingface.co/collections/kinit/sensai-metadata-extraction</a>
Model	kinit/qwen3-0.6B-extract-ml-instructions	<a href="https://huggingface.co/kinit/qwen3-0.6B-extract-ml-instructions">https://huggingface.co/kinit/qwen3-0.6B-extract-ml-instructions</a>
Model	kinit/qwen3.5-0.8B-extract-ml-conversations-qlora	<a href="https://huggingface.co/kinit/qwen3.5-0.8B-extract-ml-conversations-qlora">https://huggingface.co/kinit/qwen3.5-0.8B-extract-ml-conversations-qlora</a>
Model	kinit/qwen3.5-9B-extract-ml-conversations-qlora	<a href="https://huggingface.co/kinit/qwen3.5-9B-extract-ml-conversations-qlora">https://huggingface.co/kinit/qwen3.5-9B-extract-ml-conversations-qlora</a>
Model	kinit/qwen3.5-0.8B-extract-ml-conversations	<a href="https://huggingface.co/kinit/qwen3.5-0.8B-extract-ml-conversations">https://huggingface.co/kinit/qwen3.5-0.8B-extract-ml-conversations</a>
Model	kinit/qwen3.5-9B-extract-ml-conversations	<a href="https://huggingface.co/kinit/qwen3.5-9B-extract-ml-conversations">https://huggingface.co/kinit/qwen3.5-9B-extract-ml-conversations</a>
Dataset	kinit/synthetic-conversations-and-ml-instructions	<a href="https://huggingface.co/datasets/kinit/synthetic-conversations-and-ml-instructions">https://huggingface.co/datasets/kinit/synthetic-conversations-and-ml-instructions</a>
Dataset	kinit/synthetic-queries-and-ml-instructions	<a href="https://huggingface.co/datasets/kinit/synthetic-queries-and-ml-instructions">https://huggingface.co/datasets/kinit/synthetic-queries-and-ml-instructions</a>

**Tab. 9** Zoznam artefaktov publikovaných na HuggingFace.

<sup>1</sup> <https://huggingface.co/collections/kinit/sensai-metadata-extraction>

Počet stiahnutí zverejnených artefaktov niekoľkonásobne prekonal stanovené KPI a ku dňu podávania tejto správy dosiahol hodnotu 139:

### ▼ SensAI Collection — Download Statistics

Counts all-time downloads for every model and dataset in the `kinit/sensai-metadata-extraction` HuggingFace collection (2 datasets + 5 models).

```
[11]: # Print today's date and time
from datetime import datetime

print(f"Current date: {datetime.now().isoformat()}")
Current date: 2026-06-30T08:20:13.594504

[12]: import os

from dotenv import find_dotenv, load_dotenv
from huggingface_hub import HfApi
from IPython.display import Markdown, display

[13]: load_dotenv(find_dotenv())

api = HfApi(token=os.getenv("HF_TOKEN"))

collection = api.get_collection("kinit/sensai-metadata-extraction")

assert len(collection.items) == 7, (
    f"Expected 7 assets (2 datasets + 5 models), found {len(collection.items)}"
)

# Print all assets in the collection in a nice markdown table without the download counts
table = "| Type | ID |\n| --- | --- |\n"

for item in collection.items:
    table += f"| {item.item_type} | {item.item_id} |\n"

display(Markdown(table))

Type          ID
-----
model  kinit/qwen3-0.6B-extract-ml-instructions
model  kinit/qwen3.5-0.8B-extract-ml-conversations-qlora
model  kinit/qwen3.5-9B-extract-ml-conversations-qlora
model  kinit/qwen3.5-0.8B-extract-ml-conversations
model  kinit/qwen3.5-9B-extract-ml-conversations
dataset kinit/synthetic-conversations-and-ml-instructions
dataset kinit/synthetic-queries-and-ml-instructions

[14]: downloads: dict[str, int] = {}

for item in collection.items:
    if item.item_type == "model":
        info = api.model_info(item.item_id, expand=["downloadsAllTime"])
    elif item.item_type == "dataset":
        info = api.dataset_info(item.item_id, expand=["downloadsAllTime"])
    else:
        raise ValueError(f"Unexpected item type: {item.item_type!r} for {item.item_id!r}")

    downloads[item.item_id] = info.downloads_all_time or 0

[15]: rows = "\n".join(
    f"| {repo_id} | {item_type} | {count:} |"
    for (repo_id, count), item_type in zip(
        downloads.items(),
        [item.item_type for item in collection.items],
    )
)

total = sum(downloads.values())

display(Markdown(f"### Total downloads: {(total:,)} downloads"))
```

**Total downloads: 139 downloads**

**Obr. 9** Kód (Jupyter notebook) na získanie informácie o počte stiahnutí artefaktov zverejnených na platforme HuggingFace v rámci projektu SensAI v kolekcii "SensAI (Metadata Extraction)"

### 3.5 Popularizačné udalosti

V Tab. 12 sú prehľadne zobrazené plánované a dosiahnuté KPI projektu.

Cieľové skupiny	Plán KPI výstupov	Plán KPI dopadu	Dosiahnuté KPI výstupov	Dosiahnuté KPI dopadu
Verejnosť	# účastí na popularizačných udalostiach $\geq 2$ ;	-	# účastí na popularizačných udalostiach = 4;	-

**Tab. 12** Prehľad KPI pre popularizačné udalosti.

Počet účastí na popularizačných udalostiach bol naplnený nad rámec plánu. V rámci projektu SensAI sme využili možnosť popularizácie prostredníctvom prednášok na akciách, dostupných širokej verejnosti (AnimeShow 2026 a Európska Noc Vedy 2025). Okrem toho sme prezentovali výsledky výskumu projektu SensAI v rámci OpenSlava 2025. A v neposlednom rade sme taktiež publikovali článok v populárno-náučnom časopise Quark.

V Tab. 13 je zobrazený prehľad účastí na popularizačných udalostiach so stručným opisom danej účasti.

Opis udalosti	Forma	Link
Zverejnenie výsledkov projektu SensAI v popularizačno-vedeckom časopise <a href="https://www.quark.sk/moralna-citlivost/">Quark</a> .	článok	<a href="https://www.quark.sk/moralna-citlivost/">https://www.quark.sk/moralna-citlivost/</a>
Zverejnenie výsledkov projektu SensAI v oblasti vývoju AI v rámci pozvanej <a href="#">prezentácie</a> na <a href="#">AnimeShow 2025</a> .	prezentácia	<a href="https://www.facebook.com/kinit.sk/posts/pfbid0toVJdmKrbNMP4iqdKMAbJ5rAerT7SYiq2ooGkBPkhvqq1p6cDxKAD4gJGDqD3xu1l">https://www.facebook.com/kinit.sk/posts/pfbid0toVJdmKrbNMP4iqdKMAbJ5rAerT7SYiq2ooGkBPkhvqq1p6cDxKAD4gJGDqD3xu1l</a>
Predstavenie projektu SensAI v oblasti vývoju AI v rámci pozvanej prezentácie na <a href="#">OpenSlava 2025</a> .	prezentácia	<a href="https://www.facebook.com/kinit.sk/posts/-openslava-2025-priniesla-dva-in%C5%A1pirat%C3%ADvne-dni-pln%C3%A9-technol%C3%B3gi%C3%AD-inov%C3%A1ci%C3%AD-a-skvel/1532457048881957/">https://www.facebook.com/kinit.sk/posts/-openslava-2025-priniesla-dva-in%C5%A1pirat%C3%ADvne-dni-pln%C3%A9-technol%C3%B3gi%C3%AD-inov%C3%A1ci%C3%AD-a-skvel/1532457048881957/</a>
Predstavenie projektu SensAI v oblasti umelej inteligencie, zodpovednej inovácie a popularizácie vedy prostredníctvom interaktívnych ukážok a prednášok na podujatí <a href="#">Európska noc vedy 2025</a> .	expozičia	<a href="https://www.nocvedy.sk/">https://www.nocvedy.sk/</a>

**Tab. 13** Zoznam účastí na popularizačných udalostiach.

Obr. 11 a Obr. 12 zobrazujú informáciu o účasti na popularizačných udalostiach, kde boli okrem iného referované aj výsledky projektu SensAI v oblasti AI etiky.



**Obr. 11** Príklad účasti na popularizačnej udalosti v rámci AnimeShow 2025 aj o projekte SensAI v oblasti budúcnosti AI.

# KlnIT v časopise Quark: Morálna citlivosť

Pokračujeme v našej spolupráci s časopisom [Quark](#) a prinášame vám ďalší skvelý článok z našej série, ktorá sa bližšie pozerá na inteligentné technológie. Quark je slovenský časopis zameraný na oblasti vedy a techniky. Článok z dielne KlnITu nájdete každý mesiac v tlačenej číslu magazínu, ako aj online na našej aj [Quark](#) webovej stránke.

Do decembrového čísla sme prispeli článkom [Morálna citlivosť](#). V tomto článku sa náš výskumník [Adrián Gavorník](#) a stážistka [Lucia Hermanová](#) zaoberajú našou zodpovednosťou pri uplatňovaní etickej umelej inteligencie.

**UMELÁ INTELIGENCIA**

## Morálna citlivosť

V tomto článku sa nebudeme pozeráť na citlivosť umelej inteligencie (AI) ako takej, ale zameriame sa na tých, ktorí systémy AI vytvárajú a nasadzujú.

**Vnímavosť problému**  
Začiatkom krokom je vnímanie problému (z angl. awareness). Je to schopnosť uvedomiť si, že naše rozhodnutia vo veľkej miere ovplyvňujú životy ľudí, ktorí nemajú možnosť vyjadriť svoj názor. Keď sa vypočítajú dôsledky použitia technológií očami používateľa, ktorý má menšiu digitálnu zručnosť alebo žije v inom sociálnom prostredí, zrazu sa mení aj jeho prístup k jej dizajnu.

**Vyvažovanie hodnotových požiadaviek**  
Ďalším prvkom je vyvažovanie rôznych hodnotových požiadaviek (resolving trade-offs), čo je schopnosť zvládnuť konflikty medzi rôznymi hodnotami, ale aj praktickými cieľmi – napríklad medzi bezpečnosťou a efektivitou vývoja alebo transparentnosťou a presnosťou systému. Tieto dilemy nie sú len teoretické a AI tým sa s nimi stretávajú pravidelne.

**Prijatie rozhodnutia a určenie zodpovednosti**  
Nesťár však len rozpozná problém a pochopí jeho dôsledky. Je potrebné aj prijať rozhodnutie (judgment) a urobiť, čo z neho bude zodpovedný (accountability ascription). Cieľom je totiž robiť rozhodnutia, ktoré budú v súlade s našimi etickými hodnotami a princípmi, pričom také rozhodnutia nemusia byť najjednoduchšie alebo najpopulárnejšie.

**Význam etických nástrojov**  
Opísané zložky morálnej citlivosti nie sú lineárne, môžu prebiehať v rôznych postupnostiach a cykloch. Čo je však ešte dôležitejšie – musia sa v praxi a postupne, pri diskusiách s kolegami či pri spítom pohľade na vlastné chyby. Nie sú to len

**Zložky morálnej citlivosti**  
Morálna citlivosť vychádza z individuálnej zodpovednosti každého človeka, ale dotýka sa aj etických tímov, ktoré sa podieľajú na tvorbe systémov AI. Tvorja ju rôzne schopnosti, ktoré nám pomáhajú pri premyšľaní nad morálnymi otázkami a veľkú úlohu zohráva komunikácia. V kontexte vývoja a nasadzovania AI spoločnosť sa vieme zamerať na niekoľko takýchto schopností. Každá z nich

**Vnímanie problému**  
Začiatkom krokom je vnímanie problému (z angl. awareness). Je to schopnosť uvedomiť si, že naše rozhodnutia vo veľkej miere ovplyvňujú životy ľudí, ktorí nemajú možnosť vyjadriť svoj názor. Keď sa vypočítajú dôsledky použitia technológií očami používateľa, ktorý má menšiu digitálnu zručnosť alebo žije v inom sociálnom prostredí, zrazu sa mení aj jeho prístup k jej dizajnu.

**Vyvažovanie hodnotových požiadaviek**  
Ďalším prvkom je vyvažovanie rôznych hodnotových požiadaviek (resolving trade-offs), čo je schopnosť zvládnuť konflikty medzi rôznymi hodnotami, ale aj praktickými cieľmi – napríklad medzi bezpečnosťou a efektivitou vývoja alebo transparentnosťou a presnosťou systému. Tieto dilemy nie sú len teoretické a AI tým sa s nimi stretávajú pravidelne.

**Prijatie rozhodnutia a určenie zodpovednosti**  
Nesťár však len rozpozná problém a pochopí jeho dôsledky. Je potrebné aj prijať rozhodnutie (judgment) a urobiť, čo z neho bude zodpovedný (accountability ascription). Cieľom je totiž robiť rozhodnutia, ktoré budú v súlade s našimi etickými hodnotami a princípmi, pričom také rozhodnutia nemusia byť najjednoduchšie alebo najpopulárnejšie.

**Význam etických nástrojov**  
Opísané zložky morálnej citlivosti nie sú lineárne, môžu prebiehať v rôznych postupnostiach a cykloch. Čo je však ešte dôležitejšie – musia sa v praxi a postupne, pri diskusiách s kolegami či pri spítom pohľade na vlastné chyby. Nie sú to len

**Obr. 12** Príklad účasti v popularizačno-vedeckom časopise Quark aj o výsledkoch projektu SensAI v oblasti morálnej citlivosti.

## 4 Záver

Projekt SensAI úspešne naplnil, a vo väčšine oblastí aj výrazne prekročil, všetky plánované hlavné merateľné ukazovatele (KPI výstupov) komunikácie a zdieľania. Niektoré vedľajšie merateľné ukazovatele (KPI dopadu) bude potrebné vyhodnotiť 2 roky po ukončení projektu (napr. citácie), ale už aktuálne hodnoty naznačujú vysokú pravdepodobnosť ich naplnenia. Celkovo projekt SensAI dosiahol svoje komunikačné a diseminačné ciele a vytvoril základ pre dlhodobé využitie výsledkov výskumu odbornou komunitou, médiami i priemyselnými partnermi.