

**MUNI
FI**

Etická hlediska generativní umělé inteligence ve vzdělávání

Mgr. Tomáš Foltýnek, Ph.D.

foltynec@fi.muni.cz



Osnova přednášky

- Principy generativní umělé inteligence
- Užitečné aplikace AI
- Etická hlediska využití AI ve vzdělávání
- Detekce textu generovaného umělou inteligencí
- Etická hlediska využití AI ve vědě
- Sdílení zkušeností

Generativní AI a univerzity

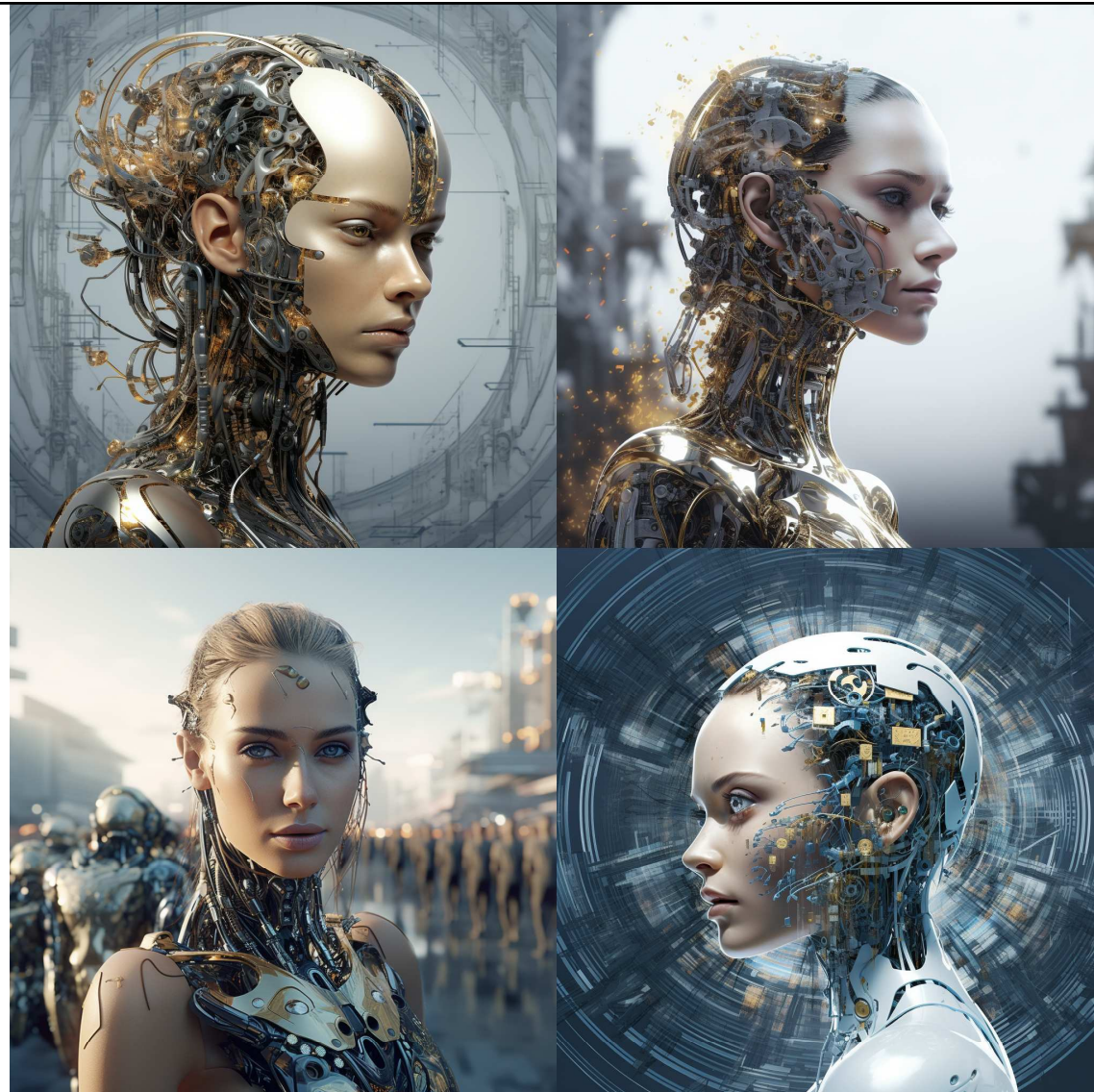
- Pokroky v generativní AI bez většího zájmu
- Listopad 2022: ChatGPT
- Zděšení, obavy, výzvy, naděje, příležitost...
- Reakce a postoje univerzit různé
 - Žádná
 - Zákaz
 - Deklaratorní dokumenty
 - Změny ve vnitřních předpisech
 - Metodická podpora



Co je AI?

Artificial Intelligence refers to systems that appear to have “intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve **specific goals**” (EU, 2018)

Generativní AI: specific goals = generování obsahu



Principy generativní AI

V této sekci využity materiály z workshopu Generativní AI v univerzitním vzdělávání, jejichž autorem je Filip Lux



Midjourney /imagine artificial intelligence writing student essay in blue and yellow, white background

Co je strojové učení?

- Vědní obor o **vytváření jednoúčelových modelů** (funkcí) pomocí analýzy dat.
- Principy známé od šedesátých let 20. století
- **Programátor definuje úlohu**, kterou model řeší. Chování modelu je dané dostupnými daty o úloze.
- Finální model se vytváří iterativně
 - proces vytváření modelu se personifikuje
 - *“učí se”, “trénuje se”, “přizpůsobuje se”*

Definice úlohy strojového učení

1. **Zadání** - Obecná otázka, na kterou model odpovídá.

BUDE DNES PRŠET?

1. **Zkušenost** - Příklady, jak lze otázku odpovědět. (Učící data)

datum	počasí	bude zítra pršet?
17.5.2023	Brno, teplota 12 °C, polojasno, vlhkost 50%, srážky 0.0 mm	?
16.5.2023	Brno, teplota 10 °C, oblačno, vlhkost 75%, srážky 0.5 mm	ne
15.5.2023	Brno, teplota 11 °C, polojasno, vlhkost 50%, srážky 0.1 mm	ano

1. **Model** - Souvisí s typem otázky.

2. **Metrika úspěšnosti** - Měření správnosti odpovědi.

Předpověď počasí

Model

BUDE DNES PRŠET?

p (bude pršet | aktuální počasí)

Předpověď počasí

Dotaz

Model

Odpověď

- aktuální počasí



BUDE DNES PRŠET?

$p(\text{bude pršet} \mid \text{aktuální počasí})$



pršet bude na
10%

Předpověď počasí

Dotaz

- aktuální počasí

Model

BUDE DNES PRŠET?

$p(\text{bude pršet} \mid \text{aktuální počasí})$

Odpověď

pršet bude na
10%

Zkušenost

- začalo pršet

Předpověď počasí

Dotaz

- aktuální počasí

Model

BUDE DNES PRŠET?

$p(\text{bude pršet} \mid \text{aktuální počasí})$

Odpověď

pršet bude na
10%

Zkušenost

- začalo pršet

hodnocení odpovědi: **X**

Předpověď počasí

Dotaz

Model

Odpověď

- aktuální počasí



BUDE DNES PRŠET?

$p(\text{bude pršet} \mid \text{aktuální počasí})$



pršet bude na
90%

Zkušenost

- začalo pršet



upravení
parametrů modelu

hodnocení odpovědi: **X**



Kontext dotazu

- Kdy model odpovídající na **BUDE DNES PRŠET?** selže?

Kontext dotazu

- Kdy model odpovídající na **BUDE DNES PRŠET?** selže?

Obecný dotaz & specifický model

vs.

Specifický dotaz & obecný model

Generativní modely

Generování textu

- na základě textového dotazu vygeneruje odpověď
- ChatGPT

Generování obrazu

- za základě textového dotazu vygeneruje ilustraci
- Midjourney, Stable Diffusion

Generátor textu

- Jazykový model odpovídá na dotaz:

DOPLŇ DO TEXTU NÁSLEDUJÍCÍ SLOVO

Generátor textu

DOPLŇ DO TEXTU NÁSLEDUJÍCÍ SLOVO

$p (\text{<слово>} \mid \mathbf{Q})$



Q: Prezidentem České republiky je <слово>

Generátor textu

<слово>: Petr 70%, Pavel 20%, Miloš 5%, ...



DOPLŇ DO TEXTU NÁSLEDUJÍCÍ SLOVO

$p (\text{<слово>} \mid \mathbf{Q})$



Q: Prezidentem České republiky je <слово>

Generátor textu

<slovo>: **Petr** 70%, Pavel 20%, Miloš 5%, ...

DOPLŇ DO TEXTU NÁSLEDUJÍCÍ SLOVO

$p (\text{<slovo>} \mid \mathbf{Q})$

Q: Prezidentem České republiky je **Petr** <slovo>

doplnění
odpovědi

Generátor textu

Pavel 95%, Čtvrtníček 4%, ...



DOPLŇ DO TEXTU NÁSLEDUJÍCÍ SLOVO

$p (\text{<slovo>} \mid \mathbf{Q})$



Q: Prezidentem České republiky je Petr <slovo>

Generátor textu

Kdo je prezidentem České Republiky?

prompt:



DOPLŇ DO TEXTU NÁSLEDUJÍCÍ SLOVO

$p (\text{<slovo>} \mid \mathbf{Q}, \text{prompt})$

Pavel 95%, Čtvrtníček 4%, ...



Q: Prezidentem České republiky je Petr <slovo>



Generátor obrázu

- odpovídá na otázku:

ODŠUM ZAŠUMĚNÝ OBRÁZEK

Generátor obrazu

iniciální stav

prompt:
roztomilé kotě

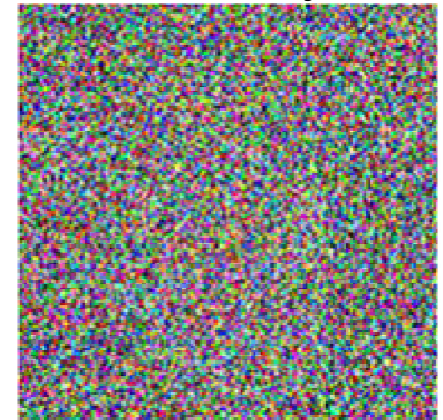


ODŠUM ZAŠUMĚNÝ OBRÁZEK

$p (\langle \text{pixely} \rangle \mid \mathbf{Q}, \text{prompt})$



Q: náhodný šum



Generátor obrazu

1. iterace

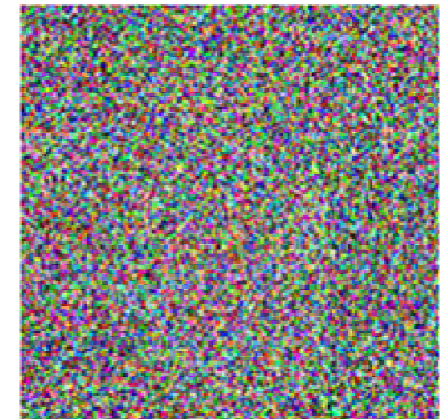
prompt:
roztomilé kotě

ODŠUM ZAŠUMĚNÝ OBRÁZEK

$p (\langle \text{pixely} \rangle \mid \mathbf{Q}, \text{prompt})$



Q:



Generátor obrazu

2. iterace

prompt:
roztomilé kotě

ODŠUM ZAŠUMĚNÝ OBRÁZEK

$p (\langle \text{pixely} \rangle \mid \mathbf{Q}, \text{prompt})$



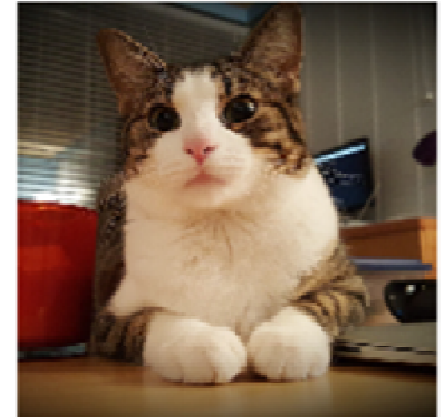
Generátor obrazu

výsledek

prompt:
roztomilé kotě

ODŠUM ZAŠUMĚNÝ OBRÁZEK

$p (\langle \text{pixely} \rangle \mid \mathbf{Q}, \text{prompt})$

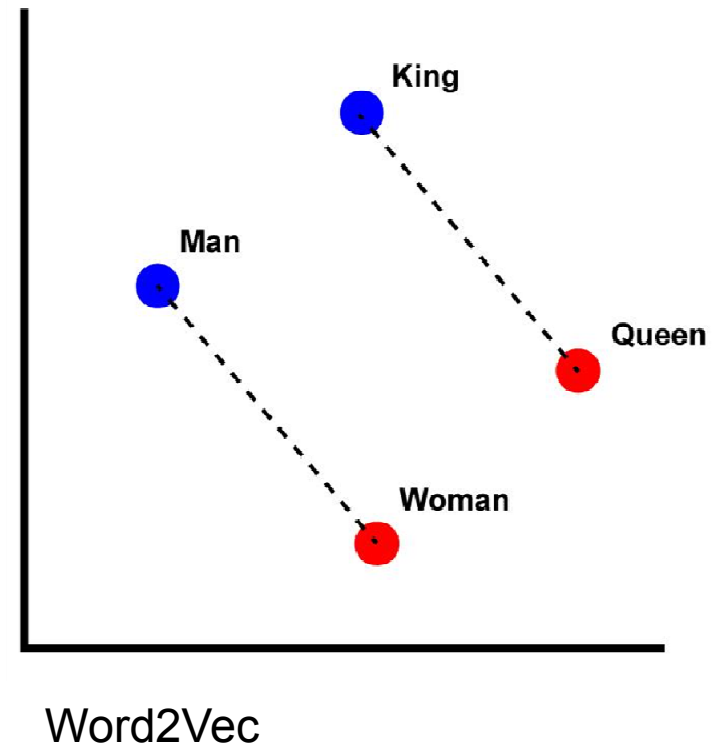


Latentní prostor

- Obrazy i text jsou reprezentovány **latentními vektory**
- Poloha zachycuje významové vztahy mezi objekty

např. *King - Man + Woman = Queen*

- Velké jazykové modely (LLM)
 - mnohadimenzionální latentní prostory



Shrnutí

- Model strojového učení odpovídá právě na **jednu úlohu**, pro kterou byl natrénován.
- Chování modelu závisí na datech, pomocí kterých byl vytvořen.
- Dotaz by měl obsahovat kontext, který není zřejmý ze zadání úlohy.
- **Neurální síť** je obecným typem učitelného modelu.
- Znalosti modelu jsou uspořádány v latentním prostoru.

Užitečné nástroje umělé inteligence



MidJourney /imagine artificial intelligence

Kde hledat?

ThereIsAnAIForThat.com

AI pro vyhledávání
AI aplikací

9862 aplikací
(stav k 22. 11. 2023)

GENERATIVNÍ AI

THERE'S AN AI FOR THAT
8,941 AIs for 2,328 tasks and 4,847 jobs.

Sponsor: [Vocol.AI - Making voice data actionable.](#)

108,347 searches today
literature review

The #1 AI aggregator. Updated daily. 2M+ users/mo.

Sign in to save YouTube Submit AI School Companionship

SEO Summaries Chatbots Music Funny

★ Featured Today Most saved Timeline ★ Get featured Job Impact Index Community Subscribe

Our favorites

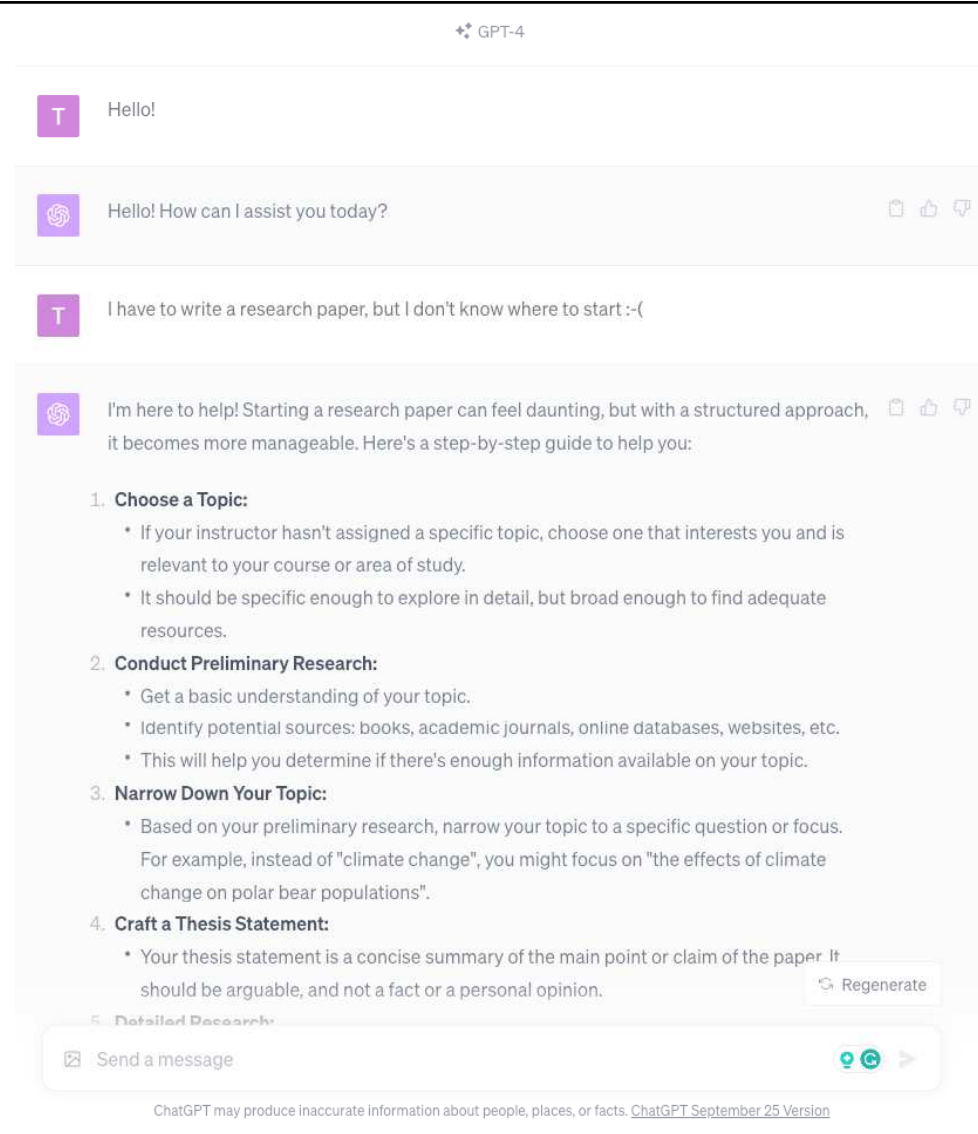
- Octocom: Sales chatbot (From \$100/mo, 358, 12, 4.9)
- Spoke.ai: Slack summaries (No pricing, 338, 1, 5.0)
- Agent4: Call answering (Free + from \$10, 236, 4, 3.9)
- FounderPal: Marketing strategies (From \$29, 78, 1)
- DeepFiction AI: Fictional stories (From \$5/mo, 81, 5, 1.2)
- Chatsimple AI Sales Chatbot: Sales (Free + from \$29/mo, 135, 4, 4.6)
- Flowpoint: Website optimization

Just launched

- SEO Writing AI: SEO content (No pricing, 455, 5, 4.0)
- Voicify: Song covers (No pricing, 761, 18, 2.9)
- Hoppy Copy: Email marketing campaigns (No pricing, 135, 4.0)
- ShortGen Video: Video generation (No pricing, 5)
- Yival: Apps (No pricing, 5)
- NewsAnalyst: News analysis (Free, 14, 4.5)
- Chatzap: Chatbots (Free + from \$19/mo, 4)
- PersonaXpress: User personas (No pricing, 10)
- BiblioBot: Research papers recommendations (Free, 15)
- Symphony: Business management (No pricing, 8)
- Noah Insights: Startup ideas analysis (Free, 13)
- The AI Virtual Assistant: Sales (No pricing, 13)
- App Quality Copilot
- Gleek
- Agenda Hero

Chat GPT

- Chatovací robot postavený na GPT
- GPT = Generative Pre-trained Transformer
 - Jazykový model natrénovaný na předpovídání (generování) dalšího slova na základě kontextu
 - Generuje nejpravděpodobnější slovo bez ohledu na faktickou správnost
- Chat GPT je virtuální asistent natrénovaný na odpovídání na otázky a generování obsahu, který vypadá jako od člověka
- Filtry na nelegální, urážlivý nebo jinak závadný obsah



The screenshot shows a chat interface with a user (T) and an AI assistant (GPT-4). The user asks for help with a research paper. The AI responds with a structured guide:

1. Choose a Topic:

- If your instructor hasn't assigned a specific topic, choose one that interests you and is relevant to your course or area of study.
- It should be specific enough to explore in detail, but broad enough to find adequate resources.

2. Conduct Preliminary Research:

- Get a basic understanding of your topic.
- Identify potential sources: books, academic journals, online databases, websites, etc.
- This will help you determine if there's enough information available on your topic.

3. Narrow Down Your Topic:

- Based on your preliminary research, narrow your topic to a specific question or focus. For example, instead of "climate change", you might focus on "the effects of climate change on polar bear populations".

4. Craft a Thesis Statement:

- Your thesis statement is a concise summary of the main point or claim of the paper. It should be arguable, and not a fact or a personal opinion.

5. Detailed Research:

At the bottom, there is a text input field with the placeholder "Send a message" and a "Regenerate" button.

ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT September 25 Version](#)

AI na čtení webových stránek

- Rozšíření do Chromu
 - [SciSpace Copilot](#), [Wiseone](#), [Merlin](#) (a jistě mnoho dalších)
- Dokáže „přečíst“ obsah stránky (článku)
- Následně reaguje na dotazy a požadavky
 - Sumarizace obsahu
 - Vysvětlení účelu stránky
 - Vysvětlení konkrétní pasáže
 - Konkrétní dotazy na obsah článku
- SciSpace copilot „rozumí“ i českému textu, odpovídá anglicky

Stanovisko k využívání umělé inteligence ve výuce na Masarykově univerzitě

Úvodem

Tento dokument představuje základní stanovisko Masarykovy univerzity k problematice využívání umělé inteligence (AI) v univerzitním vzdělávání. Záměrem Masarykovy univerzity je využívání nástrojů AI podporovat a zároveň formulovat podmínky a doporučení k jejich aktivnímu používání jak studujícími, tak vyučujícími. Cílem je formulovat principy, které se mohou stát východisky pro etické, tvůrčí, přínosné a bezpečné využívání nástrojů AI v prostředí univerzity.

Nástroje využívající AI se stávají běžně dostupnými a jsou stále více využívány i v akademické sféře. Rozvoj a využívání nástrojů AI vnímáme v první řadě jako výzvu a příležitost, nikoli jen jako hrozbu či riziko. MU si je vědoma dynamického rozvoje technologií využívajících AI, možných přínosů i úskalí spojených s těmito technologiemi. Práce s nástroji AI otevírá i zásadní otázky spojené s etikou akademické práce, jejich standardů a akademické integrity v nejširším smyslu

What are the key takeaways from this webpage?

9:01 am

The key takeaways from this webpage are:

- The Masaryk University does not recommend using free tools for detecting the origin of text that promise to recognize the use of AI tools.
- The use of AI tools in education is in line with the university's commitment to constantly develop and adapt to new and more effective teaching methods and modern technologies.
- The AI In Education group at Masaryk University is responsible for adapting the development of AI methods and evaluating the study performance of students.
- Students should be cautious while using AI tools and declare their use in their work. They should also be aware of the responsibility that comes with using AI tools.
- The use of AI tools should be used to improve education and students should be given opportunities to develop skills in using AI tools.



9:01 am

What is the purpose of this webpage?, 5 +7 more ^

Write your question...



9:01 am

follow up

Summarise the contents of this webpage.

9:04 am

The webpage discusses the use of AI tools in education at Masaryk University. It highlights the university's commitment to constantly develop and adapt to new and more effective teaching methods and modern technologies. The AI in Education group at Masaryk University is responsible for adapting the development of AI methods and evaluating the study performance of students. The university does not recommend using free tools for detecting the origin of text that promise to recognize the use of AI tools. Students should be cautious while using AI tools and declare their use in their work. They should also be aware of the responsibility that comes with using AI tools. The use of AI tools should be used to improve education and students should be given opportunities to develop skills in using AI tools.



9:05 am

follow up

What is the purpose of this webpage?, 5 +7 more ^

Write your question...



Literture Review pomocí Consensus

- Jazykový model propojený s databází vědeckých článků
- <https://consensus.app/search/>
- Po zadání klíčového slova, sousloví nebo otázky
- Najde nejrelevantnější vědecké články
- Extrahuje z nich klíčové poznatky
- Vytvoří z nich krátké shrnutí
- U otázek ano/ne udělá „meta analýzu“

Synthesize All years Share This Page

Synthesize All years Share This Page

beta These features are in Beta and not 100% accurate. Use with discretion. You can share feedback here

beta These features are in Beta and not 100% accurate. Use with discretion. You can share feedback here

Summary i

Consensus Meter i 6 papers analyzed

- Yes - 100%
- Possibly - 0%
- No - 0%

Summary i Top 10 papers analyzed

Consensus Meter i 19 papers analyzed

- Yes - 5%
- Possibly - 26%
- No - 68%

The report judged it extremely likely that human activity is the predominant cause of recent climate change due to an increase in anthropogenic greenhouse gases in the atmosphere (IPCC, 2013). Yes

The Carbon Footprint as a Mediating Tool in Students' Online Reasoning about Climate Change

G. Fauville et al. | 2016 Cite Share

Findings of this meta-analysis suggest that vaccinations are not associated with the development of autism or autism spectrum disorder. No

Vaccines are not associated with autism: an evidence-based meta-analysis of case-control and cohort studies.

Vaccine | Luke Taylor et al. | 2014

Meta Analysis Highly Cited Cite Share

(OPEN ACCESS ON ENCYCLOPEDIA WEBPAGE: Yes

<http://climatescience.oxfordre.com/view/10.1093/acrefore/9780190228620.001.0001/acrefore-9780190228620-e-328>) An overwhelming percentage of climate scientists agree that human activity is causing the global climate to change in ways that will have deleterious consequences both for the environment and for humankind.

Climate Change Conspiracy Theories

J. Uscinski et al. | 2017 Cite Share

The authors concluded that The findings of this meta-analysis suggest that vaccinations are not associated with the development of autism or autism spectrum disorder". Yes

Autism and vaccination: The value of the evidence base of a recent meta-analysis.

Vaccine | C. Turville et al. | 2015

Rigorous Journal Meta Analysis Cite Share

Our results indicate that climatically induced ecological thresholds are not yet crossed, but that human activity has an increasing influence, particularly on the terrestrial ecosystem in our study area. Yes

Terrestrial and aquatic responses to climate change and human impact on the southeastern Tibetan Plateau during the past two centuries

The extensive research conducted so far indicates that childhood vaccination is not a cause of the sharp increase in autism diagnoses in recent decades. No

[Childhood vaccines and autism—much ado about nothing?].

Harefuah | I. Solt et al. | 2010 Cite Share

can software detect plagiarism? beta Q

Synthesize All years Share This Page

beta These features are in Beta and not 100% accurate. Use with discretion. You can share feedback here

Summary Top 10 papers analyzed

These studies suggest that software can detect plagiarism to some extent, but it is most effective when combined with manual verification and human judgement.

Consensus Meter 0 papers analyzed

- Yes - 75%
- Possibly - 25%
- No - 0%

Plagiarism detection software combined with manual verification may be used to detect plagiarized manuscripts and prevent their publication. Possibly

Prevalence of Plagiarism In Recent Submissions to the Croatian Medical Journal
Science and Engineering Ethics | K. Bažarić et al. | 2012 Cite Share

The paper concludes by arguing that commonly used detection software packages can be effective in identifying matching text, but are not necessarily good plagiarism detection systems. Unknown

Investigating academic plagiarism: A forensic linguistics approach to plagiarism detection
The International Journal for Educational Integrity | R. Sousa-Silva et al. | 2014 Cite Share

Experimental results demonstrate that our software has clear advantages over other plagiarism detection systems and it is quite beneficial to teachers to get rid of time-consuming and toilsome tasks. Yes

A Program Plagiarism Detection Model Based on Information Distance and Clustering
The 2007 International Conference on Intelligent Pervasive Computing (IPC 2007) | L. Zhang et al. | 2007 Cite Share

Does artificial intelligence threaten university education? beta Q

Synthesize All years Share This Page

beta These features are in Beta and not 100% accurate. Use with discretion. You can share feedback here

Summary Top 4 papers analyzed

These studies suggest that artificial intelligence has both positive and negative impacts on university education, promoting entrepreneurial performance and improving operations, while also requiring universities to rethink their pedagogical models.

Consensus Meter

Not enough relevant results. Try asking a well-researched question in the format 'Does x cause y?'
E.g. Does beta alanine improve exercise performance?

As for the results of this study, in its application, artificial intelligence in universities has a positive and negative impact on the relationship of intelligence to learning when it is carried out outside conventional learning.

Artificial Intelligence in Higher Education: Perspicacity Relation between Educators and Students
Journal of Innovation in Educational and Cultural Research | M. Bali et al. | 2022 Cite Share

It is concluded that artificial intelligence learning is most significant to promote entrepreneurial performance among university students.

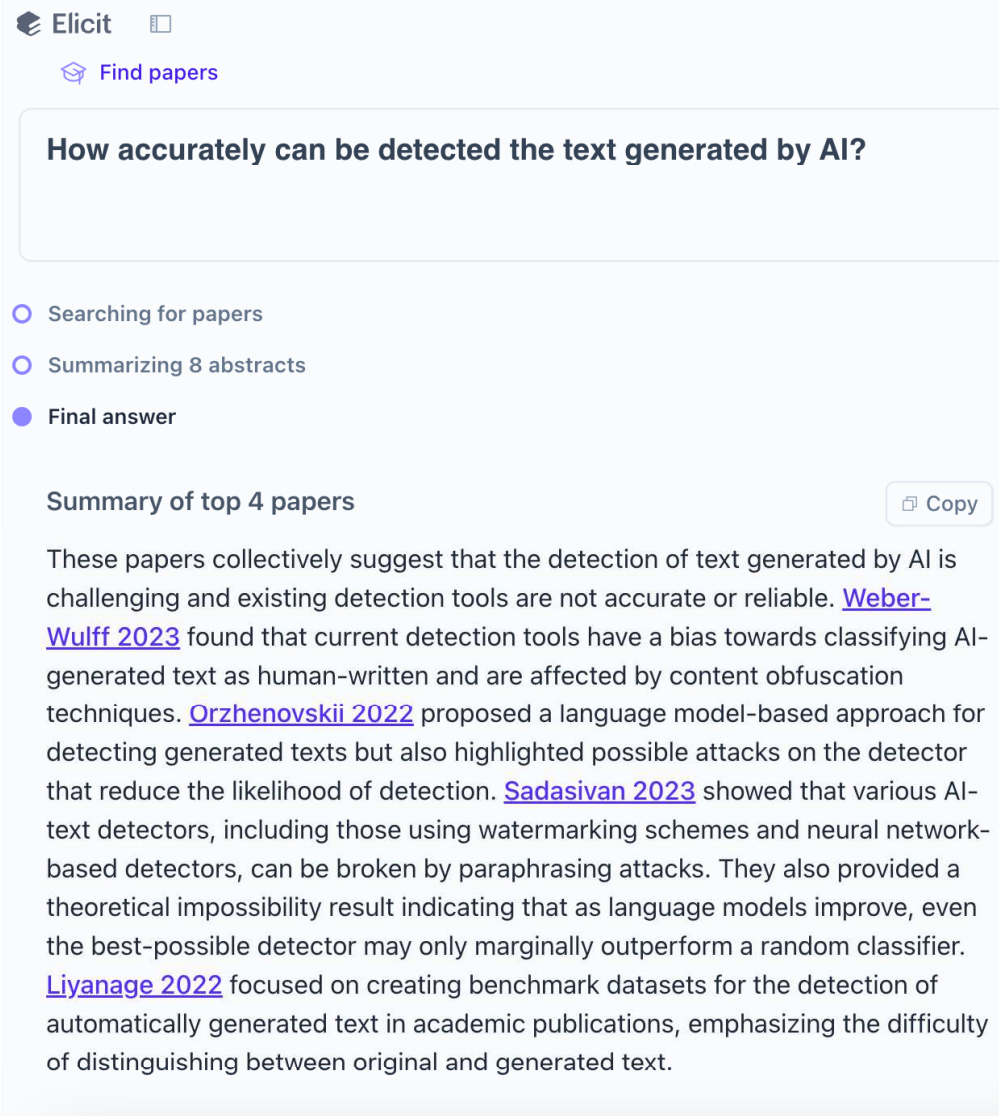
Artificial Intelligence learning and entrepreneurial performance among university students: evidence from Malaysian higher educational...
J. Intell. Fuzzy Syst. | Nadeem Khalid et al. | 2020 Cite Share


Artificial intelligence is an auxiliary, but valuable tool that can perform and improve a large number of different operations carried out at the university, help in organizing an effective educational process and building the necessary communications.


Prospects for the Introduction of artificial Intelligence technologies In higher education
ACADEMICIA: AN INTERNATIONAL MULTIDISCIPLINARY RESEARCH JOURNAL | Rakhmatov Dilmurod et al. | 2021 Cite Share

Vyhledávací nástroje

- Elicit.org
 - Najde a shrne nejrelevantnější zdroje
 - Získávání dat z PDF
 - Seznam konceptů
- Perplexity.ai
 - Odpovídání na otázky
- Tyto nástroje poskytují odkazy na zdroje
- ...jsou však tyto zdroje spolehlivé?




Elicit 

 Find papers

How accurately can be detected the text generated by AI?

- Searching for papers
- Summarizing 8 abstracts
- Final answer

Summary of top 4 papers  Copy

These papers collectively suggest that the detection of text generated by AI is challenging and existing detection tools are not accurate or reliable. [Weber-Wulff 2023](#) found that current detection tools have a bias towards classifying AI-generated text as human-written and are affected by content obfuscation techniques. [Orzhenovskii 2022](#) proposed a language model-based approach for detecting generated texts but also highlighted possible attacks on the detector that reduce the likelihood of detection. [Sadasivan 2023](#) showed that various AI-text detectors, including those using watermarking schemes and neural network-based detectors, can be broken by paraphrasing attacks. They also provided a theoretical impossibility result indicating that as language models improve, even the best-possible detector may only marginally outperform a random classifier. [Liyanage 2022](#) focused on creating benchmark datasets for the detection of automatically generated text in academic publications, emphasizing the difficulty of distinguishing between original and generated text.

<p>Testing of Detection Tools for AI-Generated Text</p> <p>Debora Weber-Wulff +7</p> <p>arXiv.org</p> <p>2023 1 citation DOI</p>	<p>The available detection tools are neither accurate nor reliable.</p>
<p>Detecting Auto-generated Texts with Language Model and Attacking the Detector</p> <p>Mikhail Orzhenovskii</p> <p>Computational Linguistics and Intellectual Technologies</p> <p>2022 1 citation DOI</p>	<p>A pre-trained language model achieved 3rd place on the binary task leaderboard with 82.6% accuracy.</p>
<p>Can AI-Generated Text be Reliably Detected?</p> <p>Vinu Sankar Sadasivan +4</p> <p>arXiv.org</p> <p>2023 57 citations DOI</p>	<p>Even LLMs protected by watermarking schemes can be vulnerable against spoofing attacks.</p>
<p>A Benchmark Corpus for the Detection of Automatically Generated Text in Academic Publications</p> <p>Vijini Liyanage +2</p> <p>International Conference on Language Resources and Evaluation</p> <p>2022 8 citations</p>	<p>Automatic text generation based on neural language models has achieved performance levels that make the generated text almost indistinguishable from those written by humans.</p>
<p>A Benchmark Corpus for the Detection of Automatically Generated Text in Academic Publications</p> <p>Vijini Liyanage +2</p> <p>ArXiv</p> <p>2022 0 citations</p>	<p>Automatic text generation based on neural language models has achieved performance levels that make the generated text almost indistinguishable from those written by humans.</p>

Generování obrázků

- Midjourney
 - Komunikace před Discord
 - Promptem může být text i obrázek
 - Příklad obrázek vpravo vznikl z portrétu TGM a textu „Future university, Artificial intelligence“
- Stable diffusion
 - Online aplikace
 - Možnost pozitivního a negativního promptu
 - Obě aplikace vygenerují 4 návrhy
 - Další rozpracování nebo zvětšení



Prompt Engineering

- Prompt by měl obsahovat co nejvíce relevantních informací
 - I am an assistant professor at the Department of Machine Learning of a high-profile university. I am preparing the introductory course on Python for a new study program Data Analytics. Please give me...
- Prompt by měl být jasný, stručný a přímočarý
 - V jedné otázce se ptát jen na jednu věc
- Je možné specifikovat cílovou skupinu
 - Rewrite it in a way that a college student can understand
- Doptávejte se, chtějte více
 - Obvykle jsou potřeba aspoň 4 prompty

Prompt Engineering

- Nakonec, Chat GPT ví nejlépe, jak se jej máme ptát 😊
 - Následující obsah byl připraven podle Chat GPT
- **TF:** Připravuji workshop pro vysokoškolské učitele o umělé inteligenci ve vzdělávání. Prosím, navrhní mi pět nejlepších strategií pro psaní dobrých promptů pro Chat GPT.
- **Chat GPT:** Při psaní promptů pro Chat GPT je důležité je navrhnout tak, aby vyvolaly přesné a informativní odpovědi. Zde je pět strategií pro tvorbu dobrých promptů pro váš workshop o umělé inteligenci ve vzdělávání:

1. Buďte jasní a konkrétní:

Poskytněte jasný kontext a specifikujte, na co se chcete model zaměřit. Nejasné nebo nekonkrétní otázky mohou vést k irelevantním nebo zavádějícím odpovědím. Ujistěte se, že prompt jasně definuje úkol nebo otázku, na kterou chcete model reagovat.

Příklad:

"Diskutujte o vlivu umělé inteligence na vzdělávání a zdůrazněte tři konkrétní oblasti, ve kterých umělá inteligence výrazně pokročila."

2. Nastavte jasné omezení

Chat GPT může generovat odpovědi, které jsou příliš dlouhé, nebo se nemusí řídit požadovanými pravidly. Nastavením omezení můžete ovlivnit délku, formát nebo jiné konkrétní aspekty odpovědi.

Příklad:

"V maximálně 200 slovech vysvětlete, jak lze umělou inteligenci integrovat do výuky tak, aby se zvýšilo zapojení studentů a vylepšily výsledky jejich učení."

3. Poskytněte kontext

Při položení otázky nebo žádosti o názor je užitečné poskytnout relevantní podkladové informace, které napomohou ke správné odpovědi. To může pomoci předejít možným nedorozuměním a zlepšit kvalitu generované odpovědi.

Příklad:

"Při narůstajícím používání umělé inteligence v automatizovaných systémech hodnocení zkoušek diskutujte o výhodách a nevýhodách přijetí takové technologie ve vysokém školství."

4. Podporujte kritické myšlení a analýzu

- Namísto požadování jednoduchých faktických informací navrhujte prompty, které vyžadují vyšší úroveň myšlení a podporují model v analýze, hodnocení a poskytování rozumných odpovědí.
- Příklad:
- *"Srovnávejte tradiční metody výuky s personalizovanými přístupy podporovanými umělou inteligencí. Posudte výhody a potenciální výzvy každého přístupu z hlediska zapojení studentů a dosažených výsledků."*

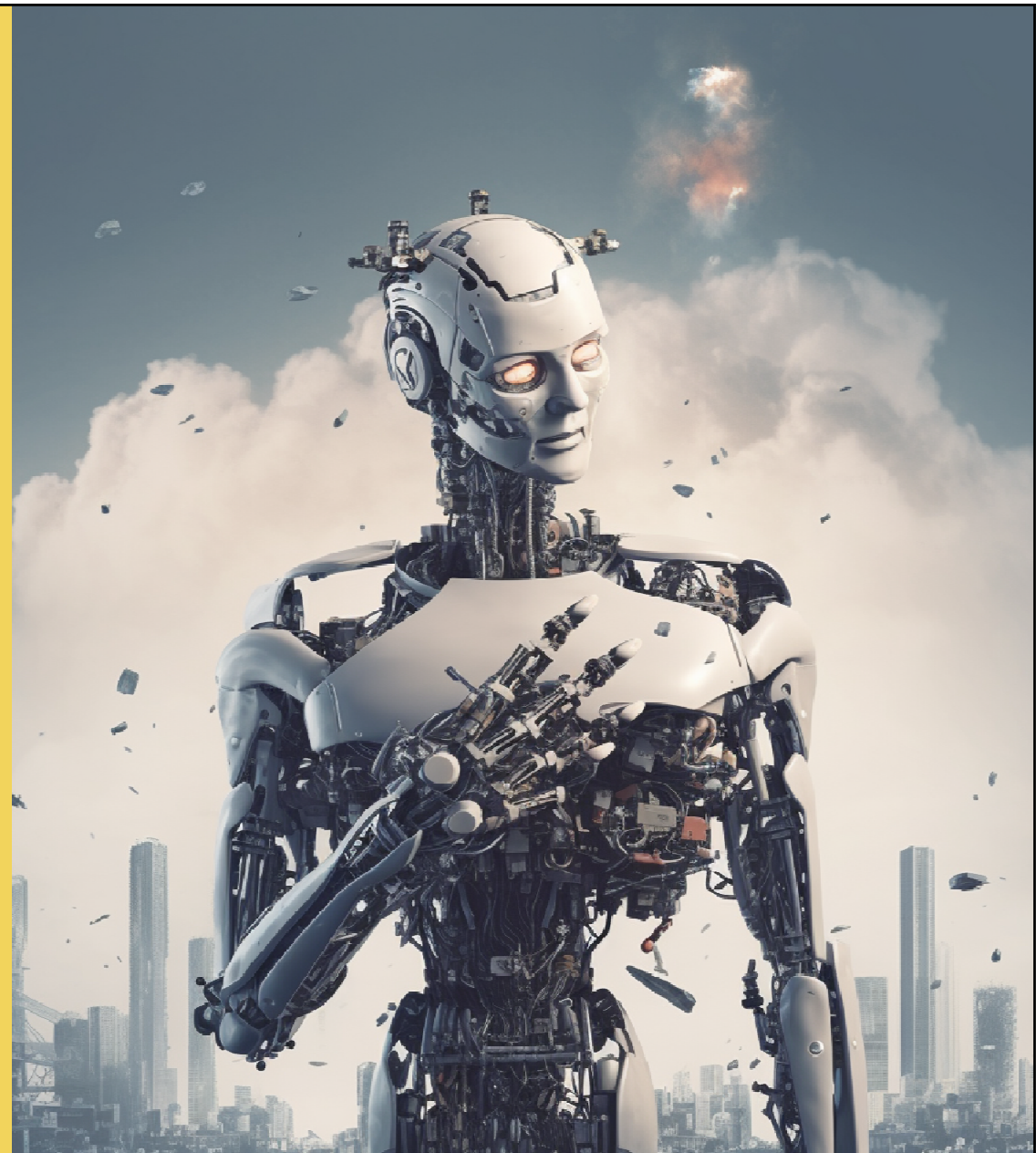
5. Experimentujte

Podpořte kreativní odpovědi tím, že navrhnete prompty, které vyzývají model k generování originálních a spekulativních odpovědí. To může vést k diskusím o potenciálních budoucích aplikacích umělé inteligence ve vzdělávání.

Příklad:

"Představte si scénář, kde umělá inteligence pokročila natolik, že dokáže vytvořit plně personalizované a adaptivní virtuální učitele. Popište potenciální výhody a etické obavy spojené s takovými virtuálními učiteli na univerzitě."

Etická hlediska využívání AI



Midjourney /imagine artificial intelligence on the shoulders of a giant, asking a question

ENAI Recommendations on the Ethical Use of AI in Education



European Network
for Academic
Integrity

- European Network for Academic Integrity
 - Sdružení vysokých škol a dalších institucí
 - Podpora a sdílení zkušeností v oblasti akademické integrity
- Široká dostupnost generativní AI posiluje existující hrozby pro akademickou integritu
 - práce a články psané na zakázku, fabrikace a falšování dat,...
- ENAI vydalo obecné doporučení
- Publikováno v [International Journal for Educational Integrity](#)

Hranice eticky přijatelného využití AI

- Vždy záleží na **smyslu** úkolu
 - Smysl úkolu by měl být studujícím jasně komunikován
- **Dovolené a přiznané využití AI je v pořádku**
- Nedovolené nebo nepřiznané využití AI splňuje definici **Unauthorised content generation**
 - Production of academic work, in whole or part, for academic credit, progression or award, whether or not a payment or other favour is involved, using unapproved or undeclared human or technological assistance.

Jaké využití AI je třeba přiznat?

- Jakékoliv využití jiných osob, zdrojů a nástrojů, které ovlivnilo myšlenky nebo vytvářelo obsah, by mělo být řádně přiznáno
 - Tedy i využití AI, pokud ovlivňuje myšlenky nebo obsah, musí být řádně přiznáno
 - Konkrétní podoba se může lišit
 - Pokud je to možné, měl by být uveden i prompt
- Využití služeb, zdrojů a nástrojů, které ovlivňují pouze formu, je v pořádku
 - jazyková a stylistická korektura, slovník synonym,...

Jak citovat generativní AI

- Doporučení [Ústavu českého jazyka FF MU](#), [APA](#), [MLA](#)
- Doporučení k využití nástrojů umělé inteligence při plnění studijních povinností
- <https://kvalita.muni.cz/kvalita-vyuky/doporuceni-k-vyuzivani-umele-inteligence-ve-vyuce>
- Využití neovlivňující myšlenky práce stačí v Poděkování (Acknowledgement), příp. vůbec
 - jazyková a stylistická korektura, přeformulování textu
- Využití pro inspiraci, prvotní seznámení s tématem, návrh struktury atd. uvést do Poděkování (Acknowledgement), Úvodu nebo Metodiky
- Převzetí ucelených pasáží textu citovat jako jakýkoliv jiný zdroj
 - V souladu s uvedenými doporučeními
 - Tam, kde je to vhodné, uvést i prompt(y)

Jak citovat generativní AI (doporučení MU)

- Aplikace na kontrolu pravopisu, stylistiky a gramatiky (Grammarly)
 - Netřeba deklarovat
- Aplikace sloužící k tvorbě odkazů, k organizaci zdrojových dokumentů, k poskytnutí přehledu existujících publikací apod. (Google Scholar)
 - Netřeba deklarovat
 - Pokud chceme přiznat, pak do Úvodu nebo Metodiky
- Nástroje pro strojový překlad (např. DeepL, Google Translate)
 - Odkaz na zdroj původního textu dle citační normy + poznámka „Přeloženo nástrojem [NÁZEV]“

Jak citovat generativní AI (doporučení MU)

- Nástroje generativní AI
 - textové generátory aplikace pro extrakci informací
 - chatboty
 - aplikace k sumarizaci obsahu
 - aplikace k sumarizaci a vyhledávání vědeckých zdrojů
 - např. ChatGPT, Bing AI, Google Bard, Perplexity, Copilot, Consensus, Elicit, apod.
- Podklady a podněty ke struktuře a obsahu textu
 - Prohlášení např. “Při přípravě této práce autor/-ka použil/-a [NÁZEV APLIKACE; WEBOVÁ ADRESA APLIKACE] za účelem [DŮVOD]. Po použití tohoto nástroje autor/-ka provedl/-a kontrolu obsahu a přebírá za něj plnou zodpovědnost.”
- Přímé využití výstupů aplikací
 - Citovat podle normy, analogie osobní komunikace

Detekce textu generovaného umělou inteligencí



Midjourney /imagine man with magnifying glass in darkness, small spotlight with traces of artificial intelligence

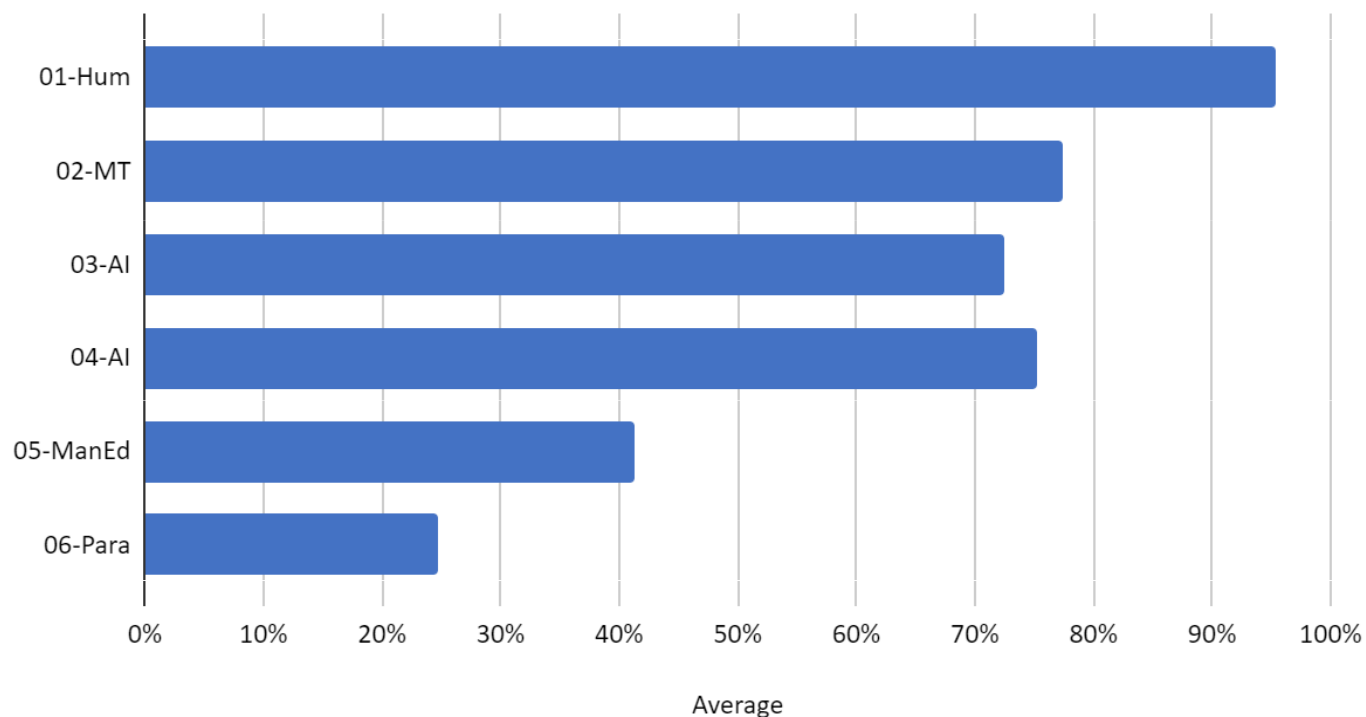
ENAI: Testing of AI Detection Tools

- 12 volně dostupných a 2 komerční nástroje
- 54 dokumentů v 6 kategoriích
 - 01-Hum: human-written
 - 02-MT: human-written + machine translation to English
 - 03-AI: AI-generated text
 - 04-AI: AI-generated text
 - 05-ManEd: AI-generated text + manual edits
 - 06-Para: AI-generated text + machine paraphrase
- Weber-Wulff et al. Testing of Detection Tools for AI-Generated Text
 - <https://arxiv.org/abs/2306.15666>



Fungují nástroje na detekci textu vytvořeného umělou inteligencí?

- Nefungují 😊
- Zkreslení směrem k “napsáno člověkem”
- I tak produkují falešně pozitivní výsledky
- Neposkytují důkaz
 - Nelze prokázat disciplinární přestupek
 - Nemožnost obrany
- Text vytvořený AI a parafrázovaný AI je většinou klasifikován jako napsaný člověkem



Etická hlediska AI ve vědě



Midjourney /imagine ethical considerations of artificial intelligence in science and research

Rizika GAI ve vědeckých publikacích

- Nepřesný, nerelevantní a nepravdivý obsah
 - “any section of a manuscript **written by an NLP system** should be checked by a domain expert for accuracy, bias, relevance, and reasoning”
 - “If a section of a manuscript **written by an NLP system** contains errors or biases, coauthors need to be held accountable for its accuracy, cogency, and integrity”
- Vědecká integrita
 - “researchers should not **use NLP systems to** fabricate empirical data or falsify existing data”

Hosseini, M., Rasmussen, L. M., & Resnik, D. B. (Jan 2023). **Using AI to write scholarly publications**. *Accountability in Research*. <https://doi.org/10.1080/08989621.2023.2168535>

Zodpovědnost

- Výstupy generativní AI mohou být zkreslené, nepřesné nebo nesprávné.
- Ani nástroj, ani jeho poskytovatel nenesou zodpovědnost za vygenerovaný obsah
- Zodpovědnost je vždy na uživateli!
- Nástroj **AI nemůže být uveden jako (spolu)autor** publikace
 - Viz kritéria autorství ICMJE/COPE

Kritéria (spolu)autorství podle COPE

- COPE (Comittee for Publication Ethics) převzalo kritéria od ICMJE (International Committee of Medical Journal Editors)
- (Spolu)autorem se osoba stává **právě tehdy, když** splní **každé** z následujících kritérií:
 - **Významný přínos** k myšlence a návrhu díla, nebo sběr, analýza nebo interpretace dat
 - **Vytváření díla** nebo kritické zhodnocení jeho myšlenek
 - **Souhlas** s výslednou podobou díla k publikování
 - **Přijetí (spolu)odpovědnosti** za dílo jako celek

Co přinese budoucnost?

- AI může přijít s novými objevy
 - kritéria autorství?
 - duševní vlastnictví?
 - recenzní řízení?
- Výstupy budou stále více výsledkem **spolupráce** člověka a AI
 - rozlišování člověk vs. AI pozbyde smysl
- Otázky ohledně **smyslu** práce
 - Co je smyslem semestrální / bakalářské / diplomové práce?
 - Co je smyslem vědeckého článku?
 - Co má být předmětem hodnocení? Text?



Midjourney /imagine bright optimistic future of education and science with artificial intelligence

Sdílení zkušeností



Midjourney /imagine A group of university teachers sharing experience,
artificial intelligence in background

Sdílení zkušeností

- Kdy a jak vám AI pomohla?
- Kdy a jak vám AI přidělala práci?



**Děkuji za
pozornost**

foltynek@fi.muni.cz

MidJourney /imagine optimistic end

