

Doporučení pro závěrečné práce obsahující texty generované umělou inteligencí (AI)

Na Jihočeské univerzitě v Českých Budějovicích pečlivě sledujeme problematiku umělé inteligence, včetně postojů k jejímu využívání a zneužívání. Vnímáme zejména pokračující zájem týkající se služby ChatGPT (ale i jiných nástrojů umělé inteligence) a možné obavy o její zneužití při tvorbě kvalifikačních nebo jiných odborných prací. Je zřejmé, že do budoucna bude třeba klást větší důraz na osobní přístup, kontrolu písemného testování, zdůrazňovat přetrvávající význam ústního zkoušení a v neposlední řadě využívat nástrojů, které umožňují rozpoznat podíl umělé inteligence na vzniklých akademických textech. Zároveň si však uvědomujeme, že ChatGPT či obdobné nástroje by neměly být považovány pouze za hrozbu, ale měly by být vnímány také jako výzva a příležitost, jak možnosti takových nástrojů ve výuce a výzkumu efektivně a zároveň eticky využívat. Stejně jako celá společnost bychom se tyto nástroje měli naučit využívat i v univerzitním prostředí a soustředit se na jejich přínosné stránky, které nebudou v rozporu s akademickou integritou a etikou studenta, učitele i výzkumníka.

Širší využívání AI může být ku prospěchu, protože šetří a urychluje práci. Zároveň vybízí k tomu, abychom se vraceli k základním pedagogickým principům, je tedy nutné:

- chápat pedagogickou práci jako doprovázení studentů, jako spolupráci se studenty, tj. nefungovat pouze v rámci modu „odučit“ – „zadat úkol“ – „ověřit a hodnotit naučené“,
- věnovat se studentům pedagogicky, nikoli pouze administrativně a organizačně; věnovat studentům čas a pozornost (např. vést studenta při zpracování kvalifikační práce průběžně, průběžně konzultovat postup a pokroky, nepřipustit, aby student odevzdal práci, o jejímž vzniku vedoucí práce nemá přehled – tato praxe je zároveň prevencí dalších nežádoucích jevů, jako je psaní kvalifikačních prací na objednávku apod.),
- zadávat seminární práce, které nejsou triviální a nemají pouze kompilační, referátový charakter (AI zase není tak „inteligentní“, jak by se mohlo z mediálních výstupů zdát, viz níže), důsledněji využívat písemné práce v rámci výuky, brát je do diskuse na seminářích, stavět studenty do situací, kde za svůj text „osobně ručí“, nevyužívat „seminárky“ pouze jako „výstupy k odevzdání“ (kdy studenti mohou pochybovat o tom, jaký je smysl takových textů, zda je vyučující vůbec čte, protože někdy od vyučujícího nedostanou jinou zpětnou vazbu kromě udělení zápočtu).

Ačkoli využívání nástrojů AI vypadá sofistikovaně, v principu – z hlediska obecnějších zásad akademické etiky – nepřináší nic nového: použité zdroje, nástroje či metody je třeba uvádět v popisu metodologie a náležitě citovat, resp. odlišovat „vlastní“ od „převzatého“; podstatou krádeže (plagiátorství) je záměna „cizího“ za „vlastní“.



Princip fungování generátorů textu na bázi AI

Chatboty využívající umělé inteligence (ChatGPT, Google Bard, Claude atd.) obsahují model založený na umělé neuronové síti trénované na internetových datech. Jedná se pouze o generátory textu, jejichž fungování je založeno na analýze posloupností výrazů, slov a pravidel spojených s konkrétním tématem. Text vytvořený na základě těchto pravidel, vychází z rozsáhlého tréninkového datového souboru, který je průběžně aktualizován. Cílem modelu není vytvářet konkrétní odpovědi na konkrétní otázky, ale generovat text, který respektuje vnitřní pravidla modelu a současně je ovlivněn vstupem od uživatele tzv. *promptingem*¹.

Typické znaky textu generovaného AI

- **Přizpůsobení se vstupu:** AI modely často přizpůsobují svůj výstup vstupu od uživatele. Tento jev vede až k přebírání stanoviska uživatele, zvláště pokud jej tento během zadávání (*promptingu*) opakuje. Přizpůsobování v sobě obvykle zahrnuje i jazykovou úroveň vstupu.
- **Nesouvislý a logicky nesprávný text:** AI modely mohou vytvářet naprosto gramaticky správný text, často na vysoké jazykové úrovni, který však obsahuje logické chyby. Například informace v textu mohou být protichůdné nebo se může změnit smysl textu bez patrného důvodu.
- **Nadbytečnost** Tyto modely mohou opakovat dříve uvedené informace, což vede k nadbytečnému textu. Místo toho, aby model použil jednoduchý a jasný jazyk, může vytvářet komplikovanější a hůře srozumitelná souvětí. Na začátku textu model typicky zopakuje otázku nebo zadání.
- **Nejasný zdroj informací** AI modely zpravidla neuvádějí zdroje nebo odkazy na informace, které poskytují. Text může obsahovat tvrzení, která nejsou podložena důkazy.
- **Nejednoznačné závěry** AI modely mají problémy s vyhodnocováním nejednoznačných otázek nebo pokynů, což může vést k nesrozumitelným nebo zmateným odpovědím, jejichž části si často mohou navzájem odporovat.
- **Absence kritického myšlení a souvislostí** V textu generovaném AI modelem může chybět kritický pohled na věc. Texty mohou obsahovat neověřené informace nebo kontroverzní stanoviska bez zvažování jejich pravdivosti nebo kontextu. Často chybí posouzení kladů a záporů daného stavu či řešení.
- **Neadekvátně podrobný text:** Někdy mohou AI modely vygenerovat text, který se příliš zabývá podrobnostmi, které s danou otázkou souvisí jen okrajově.
- **Nedostatečný kontext:** AI modely mají omezený kontext a mohou mít omezené porozumění různým situacím nebo tematickým oblastem, což může vést k nesprávným nebo neúplným odpovědím na komplexnější otázky.

¹ Posloupnost výzev, které řídí proces generování textu a vymezují kontext, vedou model k tomu, aby pochopil, co od něj uživatel chce.



Výskyt textu generovaného AI v kvalifikační práci

V případě, že je součástí kvalifikační práce text generovaný AI, je nutno ho označit jako citaci včetně uvedení zdroje (url vč. kompletního *promptingu*).

Pokud se prokáže, že kvalifikační práce obsahuje pasáže, které jsou produktem umělé inteligence, aniž by byly jasně odlišeny od vlastního autorského textu (podobně jako v případě přímých či nepřímých citací literatury), je na ni pohlíženo stejně jako na plagiát, tj. jako na neetické jednání se všemi potenciálními disciplinárními důsledky, včetně vyloučení ze studia.

V případě úspěšné obhajoby a následného prokázání nepovoleného využití umělé inteligence v kvalifikační práci studentovi hrozí odebrání titulu.

JU disponuje nástroji, kterými lze se značnou pravděpodobností určit, že text je produktem umělé inteligence. Pokud budete mít zájem o prověření textu kvalifikační práce, kontaktujte Útvar správy informačních systémů CIT na adrese plagiaty@rt.jcu.cz nebo přes formulář v ServiceDesku Jihočeské univerzity:

<https://servicedesk.jcu.cz/Alvao/Ticket/NewFromCatalog?SectionID=1534>.

V předmětu zprávy uveďte „Kontrola na AI“, dále v textu vždy uveďte osobní číslo studenta, o jehož práci se jedná, a případně jaké soubory a práce chcete zkontrolovat, pokud jich student má odevzdáno či vedeno více. Dále je nutné u jednotlivých souborů uvést, které části textu mají být prověřeny (stránky a odstavce). Práce musí být v IS/STAG označena jako odevzdaná. Soubory si pracovníci CIT z IS/STAG stáhnou sami. O kontrolu práce může požádat vedoucí práce, pracoviště, na kterém je práce vedena, nebo studijní oddělení fakulty.

Příspěvek AI v závěrečné práci by měl být vedoucím práce zohledněn při kontrole plagiátorství v IS STAG.

Autor textu: prorektor Friebel, prorektor Bauman, prorektor Berec

Recommendations for theses containing texts generated by artificial intelligence (AI)

At the University of South Bohemia in České Budějovice, we closely monitor the issue of artificial intelligence, including attitudes towards its use and abuse. In particular, we are aware of the continuing interest in ChatGPT (and other AI tools) and potential concerns about its misuse in the production of graduation theses or other scholarly works. It is clear that in the future there will be a need to place greater emphasis on a personalised approach, to check written testing, to emphasise the continuing importance of oral testing and, not least, to use tools that allow the contribution of AI to the academic texts produced to be recognised. At the same time, however, we are aware that ChatGPT or similar tools should not only be seen as a threat, but also as a challenge and an opportunity to use the functionalities of such tools effectively and ethically in instruction and research. Like society as a whole, we should learn to use these tools in the university environment and focus on their beneficial aspects that do not conflict with the academic integrity and ethics of the student, educator and researcher.

Wider use of AI can be beneficial because it saves and speeds up work. At the same time, it encourages a return to basic pedagogical principles, so it is necessary:

- to understand the pedagogical work as accompanying students, as a collaboration with students, i.e. not to work only within the 'teach' – 'assign task' – 'verify and evaluate what has been learned' mode,
- to devote time and attention to students pedagogically, not just administratively and organisationally; to give students time and attention (e.g. to guide the student in the preparation of the qualification work continuously, to consult continuously on the progress and progress, not to allow the student to hand in work that the supervisor is not aware of – this practice also prevents other undesirable phenomena such as writing qualification work to order, etc.),
- assigning seminar papers that are not trivial and not just of a compilation, report nature (again, AI is not as 'intelligent' as it might seem from media outputs, see below), using written papers more consistently in teaching, taking them into discussions in seminars, putting students in situations, where they are 'personally responsible' for their texts, not to use 'term papers' only as 'hand-ins' (where students may question what the point of such texts is, whether the lecturer even reads them, because sometimes they get no feedback from the lecturer other than the award of credit).

Although the use of AI tools looks sophisticated, in principle – from the point of view of the more general principles of academic ethics – it does not bring anything new: the sources, tools or methods used must be mentioned in the description of the methodology and properly cited, or 'own' must be distinguished from 'adopted'; the essence of theft (plagiarism) is the substitution of 'adopted' for 'own'.



How AI-based text generators work

Chatbots using artificial intelligence (ChatGPT, Google Bard, Claude, etc.) contain a model based on an artificial neural network trained on internet data. These are text-only generators whose functioning is based on the analysis of sequences of expressions, words and rules associated with a specific topic. The text generated based on these rules is based on a large training dataset that is continuously updated. The goal of the model is not to produce specific answers to specific questions, but to generate text that respects the internal rules of the model while being influenced by user input, called *prompting*.²

Typical features of AI-generated text

- **Adapting to the input:** AI models often adapt their output to user input. This phenomenon leads to the user's opinion being taken over, especially if the user repeats it during the input (*prompting*). Adaptation usually involves the linguistic level of the input.
- **Incoherent and logically incorrect text:** AI models can produce perfectly grammatically correct text, often at a high linguistic level but one that contains logical errors. For example, information in the text may be contradictory or the meaning of the text may change for no apparent reason.
- **Redundancy** These models may repeat previously provided information, resulting in redundant text. Instead of using simple and clear language, the model may produce more complicated and harder-to-understand sentences. At the beginning of the text, the model typically repeats the question or assignment.
- **Unclear source of information** AI models generally do not provide sources or references for the information they provide. The text may contain claims that are not supported by evidence.
- **Ambiguous inferences** AI models have trouble evaluating ambiguous questions or instructions, which can lead to unintelligible or confusing answers, parts of which can often contradict each other.
- **Lack of critical thinking and context** The text generated by the AI model may lack critical insight. Texts may contain unverified information or controversial opinions without considering their veracity or context. An assessment of the pros and cons of a given condition or solution is often missing.
- **Inadequately detailed text:** Sometimes AI models can generate text that is too concerned with details that are only marginally related to the question at hand.

² The sequence of prompts that drive the text generation process and define the context guide the model to understand what the user wants from it.



- **Lack of context:** AI models have limited context and may have limited understanding of different situations or topic areas, which can lead to incorrect or incomplete answers to more complex questions.

Occurrence of AI-generated text in graduation theses

If a graduation thesis includes AI-generated text, it must be marked as a citation including the source (url incl. complete *prompting*).

If a qualification thesis is shown to contain passages that are the product of artificial intelligence without being clearly distinguished from the actual author's text (as in the case of direct or indirect quotations from literature), it is viewed in the same way as plagiarism, i.e. as unethical behaviour with all the potential disciplinary consequences, including expulsion from studies.

In the event of a successful defence and subsequent proof of unauthorised use of artificial intelligence in the graduation thesis, the student risks having their degree revoked.

USB has tools that can be used to determine with a high probability that the text is the product of artificial intelligence. If you are interested in checking the text of the qualification thesis, please contact the Information Systems Management Office of CIT at plagiaty@rt.jcu.cz or via the form in the ServiceDesk of the University of South Bohemia:

<https://servicedesk.jcu.cz/Alvao/Ticket/NewFromCatalog?SectionID=1534>.

In the subject line of the message, please state 'AI check', and always include the personal number of the student whose work is being checked, and which files and work you want to check if the student has more than one file submitted or kept. You must also indicate for each file which parts of the text are to be checked (pages and paragraphs). The work must be marked as submitted in IS/STAG. CIT staff will download the files from IS/STAG themselves. The thesis supervisor, the department in which the thesis is being supervised, or the faculty's study department may request the thesis to be checked.

The AI contribution in the thesis should be taken into account by the supervisor when checking for plagiarism in IS STAG.

Authors of the text: vice-rector Friebel, vice-rector Bauman, vice-rector Berec