

INTERAKCE VELKÝCH JAZYKOVÝCH MODELŮ (LLM) A MEDICÍNSKÉHO VÝUKOVÉHO PORTÁLU WIKISKRIPTA

Čestmír Štuka, Petr Kajzar, Martin Vejražka

Abstrakt

Tento článek zkoumá současné a potenciální interakce mezi medicínským výukovým portálem a velkými jazykovými modely (LLM). S rostoucím využíváním LLM se zvyšuje citovanost WikiSkript jako zdroje důvěryhodných informací, pravděpodobně díky jejich dobré pověsti a politice otevřených licencí. Podobně i ve výuce medicíny nabývá na významu využití LLM modelů. Studenti například analyzují strojově vytvořené popisy medicínských entit (často vybraných učitelem z WikiSkript) vytvořené pomocí LLM. To umožňuje učitelům posunout výuku na vyšší úroveň kognitivních dovedností a zároveň využít nedokonalosti textů generovaných LLM k podpoře kritického myšlení studentů.

V příspěvku bude diskutována nově se otevírající možnost rozšířit snadnou upravitelnost textů v prostředí wiki i na obrázky a schémata. Mohly by k tomu sloužit textové editace obrazových promptů, nebo s pomocí LLM vygenerovaná schémata a automatickým shrnutím wiki článku.

Další otevřenou výzvou pro autory portálu bude integrace LLM s WikiSkripty, která by pomohla využít individualizovaný přístup k lékařskému vzdělávání na tomto portálu. To umožní přizpůsobit výuku reálné praxi, například místo tradičního výkladu od obecného ke konkrétnímu bude možné postupovat "od příznaku k diagnóze".

Klíčová slova

velké jazykové modely (LLM), medicínský výukový portál, WikiSkripta, kognitivní dovednosti, kritické myšlení, umělá inteligence, lékařské vzdělávání, medicína založená na důkazech

Vzdělávací portál WikiSkripta slouží pro kolektivní tvorbu a správu výukových materiálů z oblasti medicíny v českém a slovenském jazyce. Na konci roku 2023 obsahuje více než 10 800 článků. Většina uživatelů portál používá pasivně jako jeden ze zdrojů pro jejich studium. Ve zkušebním období však portál navštíví až 80 tisíc nezávislých uživatelů denně. Tento počet je paradoxně vyšší než celkový počet studentů na českých lékařských fakultách, což naznačuje, že portál je využíván i studenty jiných fakult.

Z pohledu efektivity učení mají z projektu největší přínos studenti, kteří se do něj zapojují aktivně. Ti se zde mohou naučit a rozvinout spoustu dovedností, včetně vedení nekonfliktní komunikace, využívání týmové spolupráce, zvládnutí základů redakční práce, zdokonalení kritického myšlení a důsledné aplikace medicíny založené na důkazech. Pokročilejší studenti mohou využít příležitost uplatnit své talenty a přispět k dalšímu rozvoji projektu. Aktivně zapojení studenti si vytvářejí altruistický osobní vztah k projektu a spolupracují s ním i dlouho po úspěšném zakončení studia.

S nástupem velkých jazykových modelů se otevřely nové možnosti synergie mezi výukovým portálem WikiSkripta (WS) a chatovacími službami velkých jazykových modelů (LLM). V tomto článku bychom chtěli referovat o dvou stávajících a dvou zamýšlených potenciích těchto interakcí.

První z efektů, které pozorujeme při vzájemné interakci WS a LLM, je relativní nárůst významu portálu v porovnání s ostatními zdroji na internetu. Již dříve byly WS v mnoha zdravotnických tématech na prvních pozicích internetových vyhledávačů. Preference WikiSkript jako zdroje vzrostla s nástupem využívání

chatovacích automatů LLM zejména díky souběhu přijatelné kvality obsahu a otevřené licenční politiky WS.

Portál WikiSkripta uvolňuje svůj obsah pod velmi volnou licenci „Creative Commons Uvedte původ 4.0“, která umožňuje jeho explicitní citování bez rizika porušení autorských práv. Tato kombinace vlastností je dnes velmi vzácná, protože kvalitní zdroje zpravidla nedávají svůj obsah k dispozici zdarma. Vzniká tak specifická situace, v níž je možné velmi účinně šířit konkrétní odborný názor právě jeho sdílením na platformě WikiSkript. Za těchto okolností nabývají na významu sociální kontrolní mechanismy, které zaručují, že obsah nebude tendenčně zkrácen. Je třeba studentské redakci opakovaně připomínat význam této kontroly a novou odpovědnost, protože změna se odehrála náhle (během roku) a bude potřeba se na ni adaptovat.

Druhým směrem, kde již dnes probíhá bouřlivý vývoj, je využití kombinace nástrojů umělé inteligence a portálu WikiSkripta při výuce. Scénář může být takový, že učitel vybere odborný pojem, RTG snímek, digitálně zaznamenaný histologický preparát, a požádá LLM, aby jej popsal. Studenti pak dostanou úkol tento popis komentovat, najít v něm případné chyby a navrhnout úpravy. Paradoxně se při tom využívá faktu, že dostupné LLM nemají v medicíně expertní úroveň znalostí, takže vygenerovaný slovní popis poskytuje studentům dostatek příležitostí prokázat svou expertní úroveň. Tento způsob využití LLM a WS předjímá reálnou situaci v životě, kdy „rutinní práci“ odvede stroj a od studentů se místo memorování očekává, že si udrží kontrolu (a zodpovědnost) nad společným výstupem.

Dalším způsobem, jak budou LLM dostupné pro výukový portál, je jejich využití při psaní článků pro WS. Tento nový fenomén má mnoho konsekvencí, které je třeba zvážit. LLM mohou být významným pomocníkem autora při řadě úkonů při tvorbě nového textu pro WikiSkripta, od návrhů pojmenování článku, přes návrh struktury až po finální kroky, jako je zvyšování čitelnosti textu, nebo doplňování referencí a literatury. Tvorba textu pro WS bude najednou dostupnější i pro méně zkušené autory. To je určitě krok správným směrem, kdy „rutinní“ činnosti jsou delegovány. Nicméně, je důležité si uvědomit, že lidský autor textu se nesmí vzdát kontroly. Za text je ve výsledku odpovědný právě jen on. Pokud by autor celou tvorbu delegoval na LLM, pak by degradoval obsah na „strojově generovaný“. Předpokládá se, že během dvou let bude velká část obsahu internetu produktem umělé inteligence. Takový obsah bude nahlížen jen jako dobře „učesaná“ kompilace a nebude brán jako hodnotný a originální zdroj. Vyhledávače budou strojově vygenerovaný obsah penalizovat (což již dnes dělají). Bude tedy třeba při zapojování umělé inteligence do psaní článků WikiSkript dbát na proškolení redakce a udržení autorské kontroly a zodpovědnosti za výsledek.

Pokud by se interakce LLM a WS nechala spontánnímu vývoji, mohlo by dojít k situaci, kdy příspěvatelem vloží do WS text vygenerovaný pomocí LLM, aniž by ho kriticky zhodnotil. Díky preferencím vyhledávačů a volné licenci WS vhodné pro trénování velkých jazykových modelů by následně velké jazykové modely zdrojový článek citovaly jako zdroj relevance a kruh by uzavřel pomocí sebestvrzení.

Umělá inteligence může sehrát roli i při řešení některých systémových výzev. WS se tradičně vyznačovala snadnou editací textu, která umožňovala jednoduchou opravu chyb a snadné navazování na práci někoho jiného. S postupem času pozorujeme nejen ve výuce, ale i ve sdělování informací obecně, vývoj od textových informací k vizuálním. Nicméně nástroje pro snadnou editaci obrázků nejsou k dispozici tak snadno jako textové editory.

Rychle se rozvíjející generativní AI může změnit situaci. Pomocí podrobně zapsaných promptů dokáže vytvářet stále věrohodnější a předvídatelnější grafické výstupy. Pokud bude

vývoj v tomto směru pokračovat, mohlo by být možné udržovat ve výukovém portálu místo zdrojových souborů obrázků přímo zdrojové prompty obrázků. Jejich textovou editací by pak bylo možné obrázky měnit a iterovat, což by přispělo k aktualizovatelnosti obrazového obsahu WikiSkript.

WikiSkripta jsou mimořádně přínosným nástrojem pro výuku studentů medicíny. Nicméně jejich koncepce odpovídá tradičnímu deduktivnímu vědeckému přístupu, kdy je svět popisován od obecných zákonitostí ke konkrétním projevům. V moderní didaktice medicíny se ale klade větší důraz na schopnost řešení problémů v reálném životě, tedy situaci, s níž se setkává lékař po nástupu do praxe, kde většinou postupuje od symptomu k diagnóze.

Přepsat články ve WikiSkriptech, aby sledovaly tento didaktický přístup, by bylo velmi pracné a vzhledem k objemu materiálu a kapacitě autorů možná až nereálné. Lze si však představit, že tuto „transformaci“ obsahu provedeme jen „virtuálně“, právě za pomoci LLM. Zvažujeme natrénovat vlastní LLM model, který by byl trénován nad daty portálu WikiSkripta a umožnil studentům použít i postup, v němž se na základě symptomů pokusí o stanovení diagnózy.

INTERACTIONS OF LARGE LANGUAGE MODELS (LLMS) AND THE MEDICAL LEARNING PORTAL WIKISKRIPTA

This paper explores the current and potential interactions between the medical learning portal and large language models (LLMs). As the use of LLMs increases, the station frequency of WikiSkripta as a source of credible information is also increasing, probably due to its reputation and open licensing policy. Similarly, the use of LLMs in medical education is also gaining importance. For example, students analyze and annotate LLM-generated descriptions of teacher-selected entities from WikiSkripta. This allows teachers to shift education to higher levels of cognitive skills while utilizing the imperfections in LLM-generated texts to enhance students' critical thinking.

The paper will discuss the newly opening possibility to extend the easy edibility of text in the wiki environment to images and diagrams. Another open challenge for the portal authors will be the possible integration of the LLM with WikiSkripta to help leverage the individualized approach to medical education on this portal. This will allow teaching to be tailored to real-world practice, for example, instead of the traditional general-to-specific explanation, it will be possible to proceed 'from symptom to diagnosis'.

Keywords

large Language Models (LLM), medical Educational Portal, WikiSkripta, cognitive skills, critical thinking, artificial intelligence, medical education, evidence-based medicine

Kontakt

Čestmír Štuka

1. LF UK

cestmir.stuka@lf1.cuni.cz