



CORE-GPT: združitev odprto dostopnih člankov in umetne inteligence za zaupanja vredne odgovore

CORE-GPT: Combining Open Access Articles and Artificial Intelligence for Reliable Answers

Ksenija Rivo

Oddano: 24. 8. 2023 — Sprejeto: 6. 3. 2024

1.04 Strokovni članek

1.04 Professional article

UDK 001.891:004.89

DOI <https://doi.org/10.55741/knj.68.2-3.9>

Izvleček

Namen: V članku predstavljamo inovativno storitev odprtodostopne zbirke Connecting REpositories (CORE), CORE-GPT, ki združuje model generativnih predhodno usposobljenih transformatorjev (angl. Generative Pre-trained Transformers - GPT) z obsežnim naborom odprto dostopnih znanstvenih dokumentov iz zbirke CORE. CORE-GPT nava-ja znanstvene dokumente, ki so podlaga za odgovore, ter povezave do člankov v forma-tu PDF in tako zagotavlja zanesljivost in verodostojnost odgovorov.

Metodologija/pristop: Članek podaja vpogled v delovanje modelov generativnih pred-hodno usposobljenih transformatorjev, ki so imeli pomembno vlogo pri vzpostavitvi storitve CORE-GPT. Predstavili smo tudi izsledke nedavne raziskave o zanesljivosti delovanja CORE-GPT in podali razmislek o vlogi (visokošolskih) knjižnic pri uporabi odprto dostopnih storitev.

Rezultati: CORE-GPT predstavlja pomemben korak naprej k izboljšanju iskalne izku-šnje in zanesljivosti odgovorov. Z združevanjem orodja generativnih predhodno uspo-sobljenih transformatorjev in bogate zbirke CORE zagotavlja izčrpne in na dokazih temelječe odgovore na vprašanja.

Izvirnost/uporabnost raziskave: Članek podaja eno prvih predstavitev storitve CORE-GPT v Sloveniji.



Ključne besede: odprti dostop, umetna inteligenca, generativni predhodno usposobljeni transformatorji, GPT, veliki jezikovni modeli, LLM, podatkovne zbirke

Podatkovni set: Vsebina prispevka ne temelji na raziskovalnih podatkih.

Abstract

Purpose: In this paper, we present CORE-GPT, an innovative service of the Connecting REpositories (CORE) Open Access Collection that combines the Generative Pre-trained Transformers (GPT) model with the extensive collection of open access scientific documents from the CORE Open Access Collection. CORE-GPT provides citations of the scientific papers underlying the answers, links to PDF formats of the papers and thus improves the reliability and credibility of the answers.

Methodology/approach: This paper gives an insight into the performance of the GPT models that played an important role in the creation of the CORE-GPT service. We also present some results from a recent study on the reliability of the CORE-GPT service and reflect on the role of libraries in the use of open access services.

Results: CORE-GPT represents an important step forward in improving the search experience and the reliability of answers. By combining the GPT with the rich CORE database, it enables the provision of comprehensive and evidence-based answers to questions.

Originality/practical implications: This article is one of the first presentations of the CORE-GPT service in Slovenia.

Keywords: open access, artificial intelligence, Generative Pre-trained Transformers, GPT, large language models, LLM, databases

Data set Metadata: No research data are associated with this article.

1 Uvod

Connecting REpositories (CORE) je neprofitna storitev, ki omogoča dostop do obsežne zbirke odprto dostopnih znanstvenih publikacij, pridobljenih iz repozitorijev in znanstvenih revij. Je del svetovne mreže teh podatkovnih baz, zagotavlja prepoznavnost metapodatkovnih zapisov z decentraliziranimi trajnimi identifikatorji (angl. persistent identifier – PID); upravlja vsebine, si prizadeva za skladnost s politiko odprtega dostopa in je eden od podpisnikov načel odprte znanstvene infrastrukture *Principles of Open Scholarly Infrastructures* (POSI). Connecting REpositories trenutno omogoča iskanje po zbirki več kot 125 milijonov zbranih odprto dostopnih znanstvenih publikacij (vključuje tudi zapise iz Repozitorija Univerze v Ljubljani (RUL). Do vseh rezultatov je mogoče brezplačno dostopati in prenesti polna besedila. Za zagotavljanje čim boljšega iskanja Connecting REpositories pri besedilnem in podatkovnem rudarjenju uporablja

najsodobnejšo tehnologijo, uporabnikom omogoča razvijanje novih orodij ter ponuja inovativne tehnične rešitve (CORE, b. d.).

Pojav jezikovnih modelov, ki uporabljajo umetno inteligenco (angl. *Artificial Intelligence* - AI) in temeljijo na generativnih predhodno usposobljenih transformatorjih (angl. *Generative Pre-trained Transformers* - GPT), je vplival tudi na razvojno pot COnnecting REpositories, in sicer so novembra 2022 objavili novico, da so v svojem iskalniku razvili inovativno rešitev, ki vključuje generativni predhodno usposobljeni transformator, t. i. CORE-GPT. CORE-GPT je zasnovan tako, da zagotavlja navedbe znanstvenih dokumentov, ki so bili uporabljeni pri oblikovanju odgovorov; vsi navedeni rezultati iskanja pa so iz zbirke odprto dostopnih znanstvenih publikacij COnnecting REpositories. Ključna novost CORE-GPT je torej ta, da odgovor ne izhaja le iz samega modela generativnega predhodno usposobljenega transformatorja, kot na primer pri ChatGPT, temveč temelji na obsežnem korpusu COnnecting REpositories z več milijoni odprto dostopnih znanstvenih publikacij. Odgovori, ki jih ponuja CORE-GPT, so opremljeni s sklici na znanstvene članke, na katerih temelji odgovor. Navedene so tudi povezave do člankov v formatu PDF, tako da lahko uporabnik hitro poišče te in druge ustrezne članke.

2 Generativni predhodno usposobljeni transformatorji (GPT): ustvarjanje interakcije z umetno inteligenco

Pojav naprednih tehnik obdelave naravnega jezika je korenito spremenil način interakcije z umetno inteligenco in njeno uporabo. Med temi inovacijami so se kot prelomnica izkazali generativni predhodno usposobljeni transformatorji (GPT), ki imajo doslej nepoznane zmogljivosti pri ustvarjanju človeku razumljivega besedila in razumevanju naravnega jezika. Vendar je lahko učinkovitost modelov generativnih predhodno usposobljenih transformatorjev za splošne namene v specializiranih domenah zaradi pomanjkanja domensko specifičnega znanja in kontekstualnega razumevanja omejena.

Modeli generativnih predhodno usposobljenih transformatorjev se vzpostavljajo z dvostopenjskim postopkom, ki vključuje predhodno »usposabljanje« in natančno nastavljanje. Ta pristop je pogost pri velikih jezikovnih modelih (angl. *Large Language Models*, LLM), saj jim omogoča, da se naučijo splošnega razumevanja jezika in se nato specializirajo za posebne naloge ali domene. V fazi predhodnega usposabljanja se modeli generativnih predhodno usposobljenih transformatorjev usposabljaajo na velikih količinah besedilnih podatkov, zbranih iz različnih virov, kot so spletna mesta, knjige in članki. Model se nauči napovedati

naslednjo besedo v stavku glede na prejšnje besede, kar je postopek, znan kot maskirano jezikovno modeliranje. V tej fazi model zajame slovnico, sintakso in semantiko ter pridobi precejšnjo količino splošnega znanja. Faza predhodnega učenja je programsko intenzivna in zahteva veliko virov, vendar jo je treba za vsako izgradnjo arhitekture modela opraviti le enkrat.

Po predhodnem »usposabljanju« se modeli generativnih predhodno usposobljenih transformatorjev prilagodijo določenim nalogam ali področjem z natančnim prilagajanjem na manjših naborih podatkov, specifičnih za posamezno nalogo. Ta postopek vključuje usposabljanje modela z uporabo označenih podatkov iz ciljne domene ali naloge, kot so analiza čustev, povzemanje besedila ali odgovarjanje na vprašanja. Natančno prilagajanje omogoča modelu, da se specializira za želeno nalogo in izboljša svoje delovanje na podatkih, specifičnih za domeno, pri tem pa ohrani splošno razumevanje jezika, ki se ga je naučil med predhodnim usposabljanjem.

Pri tem so kot gradniki modelov generativnih predhodno usposobljenih transformatorjev pomembni transformatorji, ki uporabljajo mehanizem samopozornosti, ki omogoča učinkovitejšo obdelavo in učenje odvisnosti dolgega dosega v besedilu. Ta arhitektura je postala temelj za številne najsodobnejše modele obdelave naravnega jezika (angl. *Natural Language Processing*, NLP), saj je zelo skalabilna (nadgradljiva) in paralelizabilna, kar omogoča učenje vedno večjih modelov (Kamnis, 2023).

Pomemben način prilagajanja modelov generativnih predhodno usposobljenih transformatorjev je tudi indeksiranje podatkov. Indeksiranje podatkov se nanaša na postopek organiziranja in shranjevanja podatkov na način, ki omogoča lažje in hitreje iskanje ter priklic ustreznih informacij. V kontekstu velikih jezikovnih modelov, kot je GPT-4 podjetja OpenAI, lahko indeksiranje podatkov ponuja več prednosti, potrebno pa je zavedanje, da koncepta učenja in indeksiranja služita različnim namenom ter imata različne prednosti in slabosti. Natančno učenje velikega jezikovnega modela v določeni domeni zahteva veliko časa in računalniških virov, indeksiranje podatkov pa običajno zahteva manj virov in ga je mogoče opraviti hitreje; zaradi tega je indeksiranje podatkov za določene naloge stroškovno ugodnejše in tudi učinkovitejše, poleg tega pa omogoča bolj kompaktno shranjevanje, saj shranjuje le pomembne informacije in odnose med podatkovnimi točkami.

Skalabilnost in prilagajanje sta glavni prednosti indeksiranja podatkov, zlasti pri vedno večjih naborih podatkov. Sisteme indeksiranja je mogoče preprosto posodobiti z novimi informacijami, medtem ko je lahko ponovno usposabljanje celotnega jezikovnega modela dolgotrajen postopek, ki zahteva veliko

virov. Prav tako jih je mogoče prilagoditi določenim področjem ali aplikacijam, kar omogoča natančnejše in ustreznejše rezultate, medtem ko je za usposabljanje celotnega jezikovnega modela potrebno učenje iz širokega nabora podatkovnih virov, kar lahko privede do bolj splošnega in manj specializiranega razumevanja teme. Indeksiranje podatkov pa omogoča hitrejšo in bolj ciljno usmerjeno iskanje informacij z uporabo posebnih poizvedb. To uporabnikom omogoča, da lažje najdejo natančne informacije, medtem ko lahko jezikovni model ustvari bolj splošne odgovore, ki zahtevajo dodatno interakcijo ali pojasnila.

Indeksirane modele generativnih predhodno usposobljenih transformatorjev je tako mogoče uporabiti za ustvarjanje interaktivnih učnih izkušenj. Ti modeli lahko na podlagi prejetih vhodnih podatkov ustvarijo besedilo, podobno tistemu, ki bi ga napisal človek, zato so uporabni za ustvarjanje gradiv za odgovarjanje na vprašanja in zagotavljanje razlag. Indeksirani model generativnih predhodno usposobljenih transformatorjev lahko na primer sprejme velik nabor podatkov o tehničnih priročnikih in nato odgovarja na poizvedbe o vsebini teh priročnikov ter tako učinkovito pomaga pri izobraževanju (Kamnis, 2023). Takšne modele je mogoče uporabiti tudi za ustvarjanje dinamičnih, prilagojenih učnih izkušenj, ki se prilagajajo napredku in potrebam študenta.

Poznavanje znanstvenih člankov v okviru generativnih predhodno usposobljenih transformatorjev pogosto ni izčrpno zaradi dejavnikov, kot so pokritost s podatki, ažurnost, zapletenost, plačljivi zidovi založb in kontekst. Čeprav generativni predhodno usposobljeni transformator zagotovi splošne informacije o znanstvenih temah, ima lahko težave s specializirano terminologijo, nedavnimi raziskavami in slojnim razumevanjem. Zanašanje na generativni predhodno usposobljeni transformator kot na orodje za pomoč pri znanstvenih raziskavah zato ni najbolj priporočljivo. Za celovito razumevanje določenih znanstvenih področij in dostop do najnovejših raziskav je torej pomembno, da se uporabijo specializirani viri in podatkovne zbirke. Generativni predhodno usposobljeni transformator, ki temelji na podatkih po meri, lahko nadgradi svoje delovanje kot orodje znanstvenega pomočnika z izboljšanjem poznavanja področja, zagotavljanjem najnovejših informacij, zmanjšanjem dvoumnosti in napak, izvajanjem prilagojenih nalog in omogočanjem izboljšanih možnosti iskanja (Kamnis, 2023).

Poleg tega generativna umetna inteligenca in veliki jezikovni modeli sicer zagotavljajo zelo prepričljive odgovore, lahko pa si stvari popolnoma »izmislijo«. Strokovnjaki tem napačnim odgovorom pravijo »halucinacije«, ki predstavljajo pomembno težavo. Če so odgovori, ki jih dajejo ti modeli, napačni, to takoj zmanjša njihovo verodostojnost ali zanesljivost. CORE-GPT to težavo odpravlja z

zagotavljanjem, da dani odgovori izhajajo izključno iz indeksiranih znanstvenih dokumentov, kar v veliki meri odpravlja tveganje ustvarjanja nepravilnih ali zavajajočih informacij, z zagotavljanjem neposrednih povezav do znanstvenih dokumentov, na katerih temelji odgovor, pa poveča verodostojnost danih rezultatov (Pride, 2023).

3 Zanesljivost in uporabnost CORE-GPT

Kot že omenjeno, CORE-GPT združuje jezikovne modele na osnovi generativnih predhodno usposobljenih transformatorjev in polna besedila več milijonov odprto dostopnih znanstvenih dokumentov iz zbirke CONnecting REpositories. CONnecting REpositories zbira vsebine iz skladišč odprtega dostopa in takoj, ko je dokument indeksiran v CORE, je v CORE-GPT na voljo za uporabo v odgovorih in za citiranje. Pride, Cancellieri in Knoth (2023) so raziskali zanesljivost CORE-GPT. Dokazali so, da se pri zagotavljanju ustreznih odgovorov na GPT-3.5 in GPT-4 ni mogoče popolnoma zanesti, medtem ko CORE-GPT zagotavlja na dokazih temelječe odgovore na vprašanja s citati in povezavami do citiranih člankov, s čimer se močno poveča verodostojnost odgovorov in zmanjša tveganje »halucinacij«. Avtorji so delovanje CORE-GPT ocenjevali na naboru podatkov 100 vprašanj, ki zajemajo 20 najpomembnejših znanstvenih področij v sistemu CORE. Na vprašanja so dobili 100 odgovorov in povezave do 500 ustreznih člankov. Kakovost odgovorov in ustreznost povezav sta ocenila dva ocenjevalca. Rezultati raziskave so pokazali, da lahko CORE-GPT pripravi izčrpne in zanesljive odgovore na večini znanstvenih področij, skupaj s povezavami na prave odgovore z ustreznimi znanstvenimi članki. Rezultati vrednotenja so pokazali, da so odgovori, ki jih zagotavlja CORE-GPT, na splošno izčrpní, uporabni, in kar je najpomembneje, vredni zaupanja. Vse reference, ki jih poišče storitev, so namreč pristni raziskovalni članki, ki jih hrani CORE.

Z raziskavo so predstavili okvir CORE-GPT, ki združuje metode velikih jezikovnih modelov z obsežnim odprto dostopnim znanstvenim korpusom za zagotavljanje zanesljive, na dokazih temelječe platforme za odgovarjanje na vprašanja. CORE-GPT se je izkazal za učinkovito rešitev za raziskovalce, ki morajo zbrati podatke in dejstva, ki temeljijo na dokazih.

Vendar pa so avtorji opozorili na dejstvo, da kljub visoki zmogljivosti programa CORE-GPT še vedno obstaja nekaj omejitev, ki jih je treba upoštevati. CORE-GPT svoje odgovore in reference črpa iz odprto dostopne literature. Odprti dostop sicer pokriva vse večji delež objavljenih znanstvenih člankov, še vedno pa obstaja veliko število člankov, ki so zaklenjeni za plačljivimi zidovi založnikov,

do katerih CORE-GPT nima dostopa. Avtorji vidijo rešitev v razvoju vgrajene različice, ki bo sčasoma omogočala več orodij za dostop do dokumentov.

4 CORE, neprecenljivi vir za raziskovanje in širjenje znanja s podporo knjižničarjev

Celotna zbirka COnnecting REpositories z več milijoni odprtodostopnih znanstvenih člankov predstavlja pomemben vir za raziskovanje in izobraževanje, zato je njen prispevek k širjenju znanja in omogočanju odprtega dostopa do raziskovalnih rezultatov neprecenljiv. Uporaba COnnecting REpositories je več kot priporočljiva za vse, ki želijo imeti dostop do verodostojnih in kakovostnih znanstvenih virov.

Vendar sam obstoj in razvoj odprtodostopnih zbirk nista dovolj, saj uporabniki pogosto niso seznanjeni z njihovimi (novimi) storitvami. Tu lahko nastopimo knjižničarji, ki poskrbimo za ustrezno promocijo. Glede na to, da se zbirka COnnecting REpositories v veliki meri bogati s pridobivanjem publikacij iz različnih odprtodostopnih repozitorijev (tudi iz Repozitorija Univerze v Ljubljani), je pomembno tudi, da sodelujemo pri korektnem vzdrževanju in promociji repozitorijev. Različne raziskave (Tmava, 2023; Bashir idr., 2022) so namreč pokazale, da imajo v visokošolskih ustanovah repozitoriji pomembno vlogo pri objavljanju institucionalne znanstvene produkcije (člankov v revijah, knjig, sestavnih delov, konferenčnih zbornikov, podatkovnih nizov itd.). Vendar, kot navaja Tmava (2023) in kar lahko opazimo tudi na slovenskem območju, se kljub njihovi vse večji uporabi na fakultetah pri raziskovalcih neredko pojavlja odpor do takega objavljanja. Visokošolske knjižnice imajo lahko pri tem pomembno vlogo, saj med drugim skrbijo za to, da se znanstvena dela objavijo v odprtem dostopu in sodelujejo pri olajševanju procesa znanstvenega objavljanja. Po mnenju Tmava (2023) bi morale prevzeti bolj proaktivno vlogo pri spodbujanju in izobraževanju zaposlenih raziskovalcev na fakultetah, zato priporoča, da izvajajo izobraževanja, ki bodo raziskovalcem, kot glavnim deležnikom pri znanstvenem objavljanju, omogočila razumevanje odprte znanosti.

Knjižničarji smo lahko proaktivni tudi pri deponiranju gradiva v repozitorije, ko lahko izkoristimo priložnost za ozaveščanje avtorjev o delovanju repozitorijev in služimo kot povezovalci za odprti dostop. S prehodom s tiskanega na digitalno založništvo je namreč nastalo veliko bolj zapleteno okolje, ki izkušenim avtorjem postavlja nova vprašanja, novi avtorji pa potrebujejo še več podpore. Tako moramo knjižničarji že zdaj pogosto znati odgovoriti na vprašanja o odprtem dostopu, praksa pa je pokazala, da bi bilo primerno v vsaki knjižnici določiti

posebnega knjižničarja, ki bi deloval kot koordinator/podpornik za vprašanja znanstvenega komuniciranja. Tovrstni knjižničarji bi lahko vodili razprave o odprtem dostopu, med drugim tudi o prednostih shranjevanja znanstvenih člankov v digitalnih repozitorijih. V ta namen bi morali biti seznanjeni tudi z novimi orodji in biti v pomoč uporabnikom pri njihovi uporabi. Le tako bomo aktivno vključeni v cikel raziskovalnih informacij.

5 Zaključek

CONnecting REpositories igra, kot del svetovne mreže odprtodostopnih znanstvenih publikacij, ključno vlogo pri omogočanju dostopa do obsežne zbirke znanstvenih člankov. S svojo zavezanostjo odprti znanstveni infrastrukturi in s skladnostjo z politiko odprtega dostopa predstavlja pomemben vir za raziskovalce, študente in druge, ki iščejo kakovostne in verodostojne informacije.

Nedavna novost CORE-GPT predstavlja pomemben korak naprej k izboljšanju iskalne izkušnje in zanesljivosti odgovorov. Z združevanjem orodja generativnih predhodno usposobljenih transformatorjev in bogate zbirke CONnecting REpositories zagotavlja izčrpne in na dokazih temelječe odgovore na vprašanja, kar povečuje verodostojnost rezultatov. Čeprav CORE-GPT predstavlja inovativno rešitev, obstajajo nekatere omejitve, kot je omejen dostop do člankov s plačljivimi zidovi založnikov. Vendar so vzdrževalci zbirke CONnecting REpositories odločeni, da bodo nadaljevali z razvojem in izboljševali storitve ter izkoristili potencial v sodelovanju z repozitoriji z namenom zagotoviti še boljši (odprti) dostop do kakovostnih znanstvenih informacij (Pride, 2023). Knjižničarji pa lahko pri tem nastopamo kot podporni člen s pomembno vlogo promotorjev odprte znanosti in njenih storitev.

Navedeni viri

Bashir, S., Gul, S., Bashir, S., Nisa, N. T., in Ganaie, S. A. (2022). Evolution of institutional repositories: managing institutional research output to remove the gap of academic elitism. *Journal of Librarianship and Information Science*, 54(3), 518–531. <https://doi.org/10.1177/09610006211009592>

CORE (b. d.). *About us*. <https://core.ac.uk/about>

Kamnis, S. (2023). Generative pre-trained transformers (GPT) for surface engineering. *Surface and Coatings Technology*, 466(129680). <https://doi.org/10.1016/j.surfcoat.2023.129680>

Pride, D. (17. 3. 2023). *CORE-GPT: Combining Open Access research and AI for credible, trustworthy question answering*. [blog]. <https://blog.core.ac.uk/2023/03/17/core-gpt-combining-open-access-research-and-ai-for-credible-trustworthy-question-answering/>

Pride, D., Cancellieri, M., in Knoth, P. (2023). CORE-GPT: Combining Open Access research and large language models for credible, trustworthy question answering. *arXiv preprint arXiv:2307.04683*. <https://doi.org/10.48550/arXiv.2307.04683>

Tmava, A. M. (2023). Faculty perceptions of open access repositories: a qualitative analysis. *New Review of Academic Librarianship*, 29(2), 123–151. <https://doi.org/10.1080/13614533.2022.2082991>

dr. Ksenija Rivo

Univerza v Ljubljani, Pedagoška fakulteta, Kardeljeva ploščad 16, 1000 Ljubljana
e-pošta: ksenija.rivo@pef.uni-lj.si