



Sveučilište u Rijeci
University of Rijeka
<http://www.uniri.hr>

Polytechnic: Journal of Technology Education, Volume 9, Number 1 (2025)
Politehnika: Časopis za tehnički odgoj i obrazovanje, Svezak 9, Broj 1 (2025)



Politehnika
Polytechnica
<http://www.politehnika.hr/journal>
cte@uniri.hr

DOI: <https://doi.org/10.36978/cte.9.1.1>

Izvorni znanstveni rad
Original scientific paper

Suvremeni AI modeli u obrazovanju: Korisnost ChatGPT-a u kreiranju testova za provjeru znanja

Jadran Jelić

Prirodoslovno-matematički fakultet
Sveučilište u Splitu
Ruđera Boškovića 33, 21000 Split
jelijadran@gmail.com

Boris Crnokić

Fakultet strojarstva, računarstva i elektrotehnike
Sveučilište u Mostaru
Matice Hrvatske b.b., 88000 Mostar
boris.crnokic@fsre.sum.ba

Sažetak

Tehnološke inovacije, posebno umjetna inteligencija, transformiraju odgoj i obrazovanje kroz olakšavanje rada učitelja i unapređenje procesa učenja. Umjetna inteligencija omogućuje automatizirano ocjenjivanje, personalizirano učenje i generiranje sadržaja, dok ChatGPT pomaže u planiranju nastave, izradi testova i procjeni znanja. Kako bi se procijenila njegova korisnost u kreiranju testova za provjeru znanja, u sklopu ovog rada provedeno je anketno istraživanje među 52 učitelja u Hrvatskoj. Većina ispitanika zadovoljna je jednostavnošću korištenja, brzinom generiranja rezultata i točnošću generiranih pitanja, a najmanje su zadovoljni visokom ovisnošću ChatGPT-a o kvaliteti unesenih podataka. Istraživanje je potvrdilo da ChatGPT štedi vrijeme, smanjuje opterećenje i poboljšava kvalitetu testova, a većina učitelja izrazila je spremnost za njegovo daljnje korištenje.

Ključne riječi: ChatGPT; AI; testovi za provjeru znanja; nastavni proces.

1 Uvod

Obrazovni sustav današnjice suočen je s brojnim problemima, među kojima se posebno ističe visoka učestalost napuštanja nastavničke profesije i općeniti nedostatak kvalificiranih učitelja. Potonji problem naročito pogađa Sjedinjene Američke Države, gdje je 2017. godine čak 70% anketiranih okruga prijavilo nedostatak učitelja za upražnjavanje slobodnih radnih mjesta (Sucher, Darling-Hammond i Carver-Thomas, 2019). Osim toga, prema projekcijama UNESCO-vog (2024) recentnog globalnog izvješća, do 2030. godine u svijetu će nedostajati 44 milijuna učitelja u osnovnim i srednjim školama. Ova brojka uključuje i više od polovice učitelja koji će u međuvremenu napustiti profesiju. Takvih je učitelja,

sudeći prema rezultatima brojnih istraživanja, iz dana u dan sve više. Primjerice, u Engleskoj je u akademskoj godini 2021./2022. profesiju napustilo 40 tisuća učitelja (Smith, 2023). Jednako tako, želju za napuštanjem profesije iskazala je trećina ispitanih učitelja iz istraživanja koje su provele Radeka i Sorić (2006) među 462 učitelja u Hrvatskoj.

Ovako poražavajuća statistika uzrokovana je različitim čimbenicima, a studije ukazuju kako je prvom redu riječ o prevelikom obujmu posla (Amittai i Van Houtte, 2022; Brandenburg i sur., 2024; Heffernan i sur., 2022; Longmuir i sur., 2022; Perryman i Calvert, 2020; Smitthers i Robinson, 2003), sljedeći problem su niske plaće u prosvjeti (Gonzalez, Brown i Slate, 2008; Kecici i Bilge Keskinkilic, 2023; Smitthers i Robinson, 2003), potom nedostatak administrativne podrške (Brandenburg i

sur., 2024; Gonzalez, Brown i Slate, 2008; Perryman i Calvert, 2020), nedostatak discipline među učenicima (Gonzalez, Brown i Slate, 2008) te visoke razine stresa i zdravstveni problemi koji iz njega proizlaze (Kececi i Bilge Keskinilic, 2023).

Da je preveliki obujam posla glavni uzrok stresa i napuštanja profesije govore i podaci istraživanja koje su proveli Robinson i Hamilton (2023) s ciljem identificiranja temeljnih elemenata koji utječu na preopterećenje učitelja kako bi se pronašla adekvatna rješenja za njihovo minimiziranje. Autori su pronašli ukupno sedam čimbenika intenziviranja rada učitelja, a na prva dva mjesta uvrstili su planiranje nastave i ocjenjivanje učenika. Otkriveno je kako upravo ova dva procesa zahtijevaju značajne količine vremena i napore učitelja, posebno u suvremenom odgojno-obrazovnom okruženju gdje se očekuje integracija različitih aspekata odgoja i obrazovanja i prilagodba individualnim potrebama učenika (Robinson i Hamilton, 2023).

Navedeno upućuje na zaključak kako je od izuzetne važnosti razvijati tehnološka rješenja koja mogu rasteretiti nastavnike u ovim ključnim aspektima njihova rada, omogućujući im da se usredotoče na kvalitetu nastave i pružanje bolje podrške učenicima. Kao jedno od mogućih rješenja nameće se viši stupanj implementacije novih tehnologija u odgojno-obrazovni sustav.

2 Transformacija obrazovanja uz pomoć tehnoloških inovacija - Poboljšanje produktivnosti učitelja kroz automatizaciju i olakšano planiranje nastave

Do danas se pokazalo kako različita tehnološka rješenja imaju brojne prednosti u obrazovanju. Primjerice, e-dnevnik pojednostavljuju i ubrzavaju posao učiteljima pružajući im praktičan način za vođenje evidencije o učeničkim rezultatima, prisutnosti i drugim relevantnim informacijama (Lee, 2020; Vrkić Dimić i Vidov, 2019). Platforme za suradničko učenje i interaktivne simulacije kao što je, primjerice, online alat *Kahoot!*, pomažu učiteljima da stvore interaktivne lekcije, zadatke i kvizove koji potiču učenike na samostalno učenje, kritičko razmišljanje i angažirano sudjelovanje (Kaur i Nadarajan, 2020; Kurt, 2010). Jednako tako, različiti automatizirani alati za ocjenjivanje omogućuju automatsko ocjenjivanje testova i zadataka, čime se učiteljima oslobađa vrijeme koje bi potrošili na ručno ocjenjivanje, posebno kod većeg broja učenika (Wieczorek, 2022). Osim toga, mogućnost kreiranja video lekcija, online materijala, animacija i slika

pomaže im da prilagode svoju nastavu i pruže dodatne nastavne materijale učenicima, olakšavajući na taj način vlastitu pripremu za nastavu (Rahim, Suherman i Muttaqin, 2020).

Posljednjih godina došlo je i do naglog napretka umjetne inteligencije, što je dodatno promijenilo "pejzaž" odgoja i obrazovanja, a inteligentni sustavi razvijeni u okviru navedene tehnologije koriste se za personalizirano učenje, automatsko ocjenjivanje testova, generiranje obrazovnih sadržaja, itd. (Pedro i sur., 2019).

Općenito, umjetna inteligencija (*Artificial Intelligence* – AI) predstavlja područje računalne znanosti koje uključuje razvoj inteligentnih sustava koji su u stanju obavljati različite zadatke za koje je uobičajeno potrebna ljudska inteligencija. Ovo podrazumijeva donošenje odluka na temelju analize sadržaja, prepoznavanje prirodnog jezika, rješavanje problema, prepoznavanje uzoraka, itd. (Naqvi, 2020).

U kontekstu odgoja i obrazovanja, umjetna inteligencija pruža raznovrsne aplikacije i inteligentne sustave koji podržavaju nastavnike i učenike, olakšavajući pripremu i planiranje nastave, potičući učenje i usvajanje složenog nastavnog sadržaja te pomažući učiteljima u prepoznavanju nedostataka u znanju učenika i pružanju ciljanih povratnih informacija radi poboljšanja ishoda učenja (Chiu i sur., 2023; Harry, 2023).

Od brojnih primjena umjetne inteligencije u odgoju i obrazovanju, u literaturi su najčešće proučavani inteligentni sustavi za poučavanje (*Intelligent Teaching System* – ITS), poznati i kao inteligentno računalno potpomognuto učenje koje omogućuje prilagodbu nastave prema sposobnostima učenika (Abu Ghali i sur., 2018). ITS omogućuje personaliziranu edukaciju i povratne informacije prilagođene svakom učeniku, bez potrebe za ljudskim posredovanjem. Sustav prilagođava težinu zadatka svakom učeniku te im omogućuje napredovanje od lakših prema težim. Na ovaj se način smanjuje potreba za ručnim pripremanjem i ocjenjivanjem zadataka, što značajno smanjuje obujam posla učitelja. Osim toga, ITS osigurava praćenje napretka svakog učenika i identificiranje područja u kojima im je potrebna dodatna podrška, što dalje olakšava planiranje i prilagodbu nastave (Alhabbash, Mahdi i Naser, 2016).

Drugi primjer primjene umjetne inteligencije u odgoju i obrazovanju su sustavi za upravljanje učenjem (*Learning Management Systems* - LMS), sofisticirani softverski alati osmišljeni za organizaciju, nadzor i upravljanje procesom učenja unutar škola, sveučilišta ili korporacija. Ovi sustavi pružaju platformu na kojoj učitelji mogu objavljivati nastavne materijale, dodjeljivati zadatke, provoditi testove, pratiti napredak učenika i komunicirati s

njima (Crouse-Machcinski, 2019). Nadalje, LMS daju mogućnost učenicima da pristupe sadržaju, surađuju s kolegama, prate vlastiti napredak i da komunikaciju s učiteljima. S obzirom na njihovu funkcionalnost, ovi alati olakšavaju organizaciju nastave, potiču personalizirano učenje, olakšavaju komunikaciju između učitelja i učenika te pružaju korisne uvide u napredak učenika, što u konačnici pomaže učiteljima u prilagodbi i unaprjeđenju nastave (Owan i sur., 2023).

2.1 Primjena ChatGPT-a u obrazovanju

Među najnovijim tehnološkim inovacijama koje se oslanjaju na umjetnu inteligenciju, a koje svoju primjenu imaju i u obrazovanju, ističe se ChatGPT. Riječ je o virtualnom asistentu razvijenom od strane tvrtke OpenAI, koji je lansiran 20. studenog 2022. godine (OpenAI, 2024). Unatoč tome što je ChatGPT relativno nova inovacija, zabilježio je impresivan rast u svojoj početnoj fazi. U prvih pet dana od lansiranja, ovaj alat za obradu prirodnog jezika privukao je milijun korisnika. Taj uspjeh nadmašio je čak i društvenu mrežu *Instagram*, koja je do tada bila poznata po brzom rastu korisničke baze, a koja je milijun korisnika dosegla za svega 2 mjeseca i 15 dana. Prema najnovijim statistikama, u veljači 2024. godine, ChatGPT je brojio impresivnih 1,6 milijardi korisnika (Duarte, 2024).

ChatGPT je napredni jezični model zasnovan na umjetnoj inteligenciji koji djeluje na principu *chatbota* (virtualnog asistenta), gdje korisnik postavlja različita pitanja i zahtjeve, a inteligentni sustav generira adekvatne odgovore temeljene na promptnoj analizi velike količine podataka (Sami i sur., 2023; Sok i Heng, 2023).

U protekle dvije godine, ChatGPT privukao je pažnju mnogih istraživača koji su temeljito analizirali njegove mogućnosti i korisnost u različitim područjima. Njegova primjena obuhvaća širok spektar disciplina, uključujući medicinu (Dave, Athaluri i Singh, 2023), financije (Khani Umer, 2024), umjetnost (Fan i Liang, 2023; Zhu i sur., 2024) odgoj i obrazovanje (Grepcka i sur., 2024; Shalevska, 2023; Sok i Heng, 2023).

Ovaj suvremeni inteligentni sustav nudi korisnicima različite mogućnosti, od davanja preporuka i savjeta (Mountcaste, 2024), preko istraživanja informacija, analize tekstova (Sharma, 2023) i pomoći u učenju (UNESCO, 2023) pa sve do sudjelovanja u kreativnom procesu poput stvaranja priča i pjesama ili pisanja eseja (Ortiz, 2024), u ovisnosti o zahtjevima korisnika. Upravo potonji element čimbenik je brojnih kritika i zabrinutosti stručnjaka oko upotrebe ChatGPT-a u odgoju i obrazovanju obzirom da njegova upotreba otvara

brojna etička pitanja, uključujući mogućnost varanja učenika tijekom pisanja ispita, plagiranja učeničkih radova i slično (Famaye i sur., 2024). Dodatan problem proizlazi iz činjenice da postojeći sustavi za razotkrivanje plagijata nemaju sposobnost prepoznati esej napisan od strane ChatGPT-a, što pomaže učenicima u varanju (Shalevska, 2023). Izuzev navedenog, istraživači su otkrili i različite druge nedostatke ChatGPT-a, poput netočnosti ponuđenih rezultata, na što su se posebno osvrnuli Bezirhan i Von Davier (2023) ističući kako je tijekom upotrebe ChatGPT "(...) bio sklon generiranju pogrešnih rezultata (...), posebno za određene statistike i brojeve" (Bezirhan i von Davier, 2023: 17). Drugi autori su uputili i na visok stupanj pristranosti inteligentnog sustava u području rodne zastupljenosti (Li i Bamman, 2021; Shihadeh et al., 2022), rasnih predrasuda (Brown i sur., 2020) i vjerske pristranosti (Abid i sur., 2021).

Međutim, ukoliko se koristi primjenom etičkih načela, ChatGPT pruža brojne prednosti u nastavnom procesu, kako za učenike, tako i za nastavnike. Jedan od ključnih aspekata jest njegova sposobnost pružanja personaliziranog iskustva učenja za svakog pojedinog učenika. Analizom podataka i prethodnih interakcija s korisnicima, ChatGPT može prilagoditi svoje odgovore i preporuke na način da odgovaraju specifičnim potrebama i razumijevanju svakog učenika. Osim toga, pruža brzu i adekvatnu pomoć u učenju i traženju informacija u bilo kojem trenutku, što može poboljšati efikasnost učenja te smanjiti stres i tjeskobu kod učenika (Alnaqbi i Fouda, 2023; Zimmerman, 2023). Iz perspektive učitelja, ChatGPT može pružiti pomoć u različitim područjima, od generiranja kreativnih ideja i rješenja za obogaćivanje nastavnog procesa, preko asistiranja u izradi nastavnih planova i prezentacija, pa sve do savjetovanja oko načina prilagodbe nastavnog sadržaja učenicima s posebnim potrebama (Karakose i Tulubas, 2023).

2.2 Chat GPT kao potencijalno rješenje za izradu kvalitetnih testova za provjeru znanja

S obzirom na svoju svestranu primjenu, javlja se pretpostavka kako bi ChatGPT mogao biti potencijalno korisno rješenje za pomoć učiteljima pri izradi testova za provjeru znanja.

U različitim zemljama, testovi i ispiti znanja koriste se kao obavezan dio nastavnog procesa jer pružaju mjerljiv dokaz o uspješnosti učenika u postizanju nastavnih ciljeva (Shohamy, 2000), a njihova učestalost, kompozicija, vrsta i namjena određeni su kurikulumom nastavnog predmeta,

preferencijama samog učitelja i standardima odgojno-obrazovnog sustava (Murphy, Little i Bjork, 2023). Međutim, u Republici Hrvatskoj vrednovanje uspjeha učenika temelji se na odgojno-obrazovnim ishodima pri čemu učitelji imaju autonomiju u odбору metoda vrednovanja koje najbolje odgovaraju ciljevima i ishodima učenja. Testovi i ispiti za provjeru znanja nisu izričito propisani kao obvezni alati, ali su često korišteni u praksi (Ministarstvo znanosti i obrazovanja, 2017).

Osim što omogućuju procjenu postignuća učenika u spoznajnoj domeni, testovi znanja također služe kao alat za identifikaciju područja u kojima su potrebna dodatna nastavna intervencija ili prilagodba nastavnog procesa kako bi se poboljšali rezultati učenja (Centar za istraživanje i inovacije u obrazovanju, 2008).

Uvriježeno je mišljenje kako je neophodno da učitelj ima stručna znanja, sposobnosti i kompetencije za izradu kvalitetnih testova i ispita znanja (Aydin, 2000, prema: Ozlem Saka, 2016) pri čemu bi trebali voditi računa da konstruiraju testove različitih razina težine kako bi osigurali svim učenicima pravednu priliku za demonstriranje svojih sposobnosti i znanja (Gronlund i Linn, 1990). Jednako tako, pokazalo se kako izrada testova oduzima veliku količinu vremena i od učitelja zahtjeva ulaganje značajnih napora u kreiranje i prilagodbu ispitnih pitanja koja udovoljavaju brojnim zahtjevima (Rawson i Dunlosky, 2012), među kojima je kvaliteta postavljenih pitanja najvažniji segment. To uključuje pažljivo planiranje i definiranje ispitnih ciljeva, osiguravanje jasnoće i preciznosti pitanja te eliminiranje dvojbenih ili dvoznačnih formulacija. Također, važno je osigurati da ispitna pitanja odražavaju bitne koncepte i vještine koje su obuhvaćene nastavnim programom, kako bi rezultati bili relevantni i mjerljivi (Pešorda, 2018).

No, postavlja se pitanje koliko su učitelji u ovome zaista vješti, s obzirom da su Marco i Pigge (1992) u svom sustavnom pregledu koji je obuhvatio čak 225 studija, ustanovili kako postoji nekoliko nedostataka u znanju i obuci učitelja primarnog obrazovanja kada je u pitanju testiranje znanja učenika. Naime, pokazalo se kako većina testova koje su učitelji samostalno izradili ne udovoljava uvjetima stručnosti i točnosti. Uočeno je kako samostalno izrađeni testovi sadrže veliki broj pogrešaka te da učitelji obično ne koriste strategije poboljšanja testova kao što je, primjerice, analiza pojedinih pitanja i zadataka. U drugom istraživanju koje je obuhvatilo 62 učitelja provedena je analiza testova znanja koje su samostalno izradili, a utvrđeno je kako je većina stavki obuhvaćala pitanja usmjerena na prisjećanje i prepoznavanje, a tek manji udio pitanja odnosio se na procjenu načina na koje učenici

obrađuju prikupljene informacije i grade znanje (Simsek, 2016). Osim toga, ranija istraživanja otkrila su i to da učitelji u osnovnim školama uglavnom izrađuju vlastite testove znanja, dok se oni u srednjoškolskom obrazovanju većinom oslanjaju na standardizirane testove koje im nude izdavači udžbenika i koji uobičajeno imaju visoke razine pouzdanosti i valjanosti (Marso i Pigge, 1988; Salmon-Cox, 1981).

Zanimljivo je spomenuti i jedno istraživanje provedeno među učenicima srednjih škola u Hrvatskoj i Bosni i Hercegovini provedeno s ciljem analize stupnja prevalencije i prihvatljivosti varanja na ispitima. Rezultati istraživanja otkrili su kako postoji srednja stopa varanja na koja nastaje kao posljedica posjedovanja ispitnih pitanja prije pristupa samom ispitu znanja. Utvrđeno je kako učenici pribavljaju testove od ranijih generacija ili od grupa učenika koji su testove polagali nekoliko dana ranije ili na sam dan polaganja ispita (Šimić Šašić i Klarin, 2008). Ovi nalazi sugeriraju da učitelji nerijetko pribjegavaju korištenju identičnih testnih pitanja za više grupa učenika, pa čak i za više generacija, što predstavlja značajan problem kada je u pitanju kvaliteta odgoja i obrazovanja. Ovakav pristup može dovesti do smanjenja autentičnosti ocjenjivanja te nepravilnosti u procjeni učeničkih sposobnosti i postignuća, što ukazuje na potrebu za promjenom u praksi konstruiranja i uporabe testova kako bi se osigurala pravednija i pouzdanija procjena učeničkog napretka.

Uzimajući u obzir sve ranije identificirane probleme poput povećanog obujma posla i nedostatka vremena za adekvatnu pripremu nastavnog programa pa samim time i ispita znanja, kao i manjak vještina, znanja i sposobnosti za konstrukciju kvalitetnih, pouzdanih i valjanih pitanja, stvara se potreba za traženjem alternativnih rješenja koja bi učiteljima mogla olakšati ovaj dio posla. ChatGPT nameće se kao potencijalno rješenje, a uporište za ove tvrdnje moguće je pronaći u ranijim istraživanjima (Cotton, Cotton i Shipway, 2023; Farrokhnia i sur., 2023; Grepcka i sur., 2024; Opara, Adalikwu i Tolorunleke, 2023; Rossi, n.d.; Shalevska, 2023) koja nisu izravno proučavala korisnost Chat GPT-a u kreiranju testova za provjeru znanja ali su autori u sklopu vlastitih istraživanja jednim dijelom obuhvatili navedeno područje i zaključili kako bi ovaj inteligentni sustav mogao biti od velike pomoći učiteljima. U tom smislu pojedini autori sugeriraju kako ChatGPT učiteljima pruža širok spektar mogućnosti za konstruiranje pitanja za ispite znanja. Kroz svoju naprednu analizu jezika, ovaj alat može generirati različite vrste pitanja, uključujući one s višestrukim izborom i otvorena pitanja. Osim toga, ChatGPT može identificirati leksičke i gramatičke greške u napisanim radovima, što je od izuzetne

važnosti za kvalitetno ocjenjivanje učeničkih radova. Također, omogućuje ocjenjivanje testova s višestrukim izborom te razvijanje rubrika za procjenu izvedbe, pružajući korisne smjernice kako učiteljima tako i učenicima (Abolalela, 2023; Cotton, Cotton i Shipway, 2023; Farrokhnia i sur., 2023; Opara, Adalikwu i Tolorunleke, 2023).

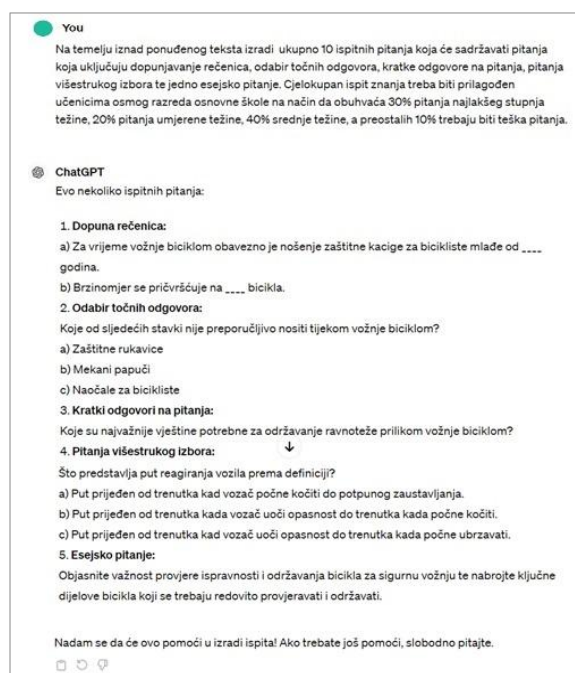
U okviru dostupne literature još uvijek nedostaje studija koje temeljito istražuju ulogu i korisnost ChatGPT-a u kreiranju testova za provjeru znanja. Za sada postoji samo jedna koja se bavila ovim područjem i gdje su autori koristili ChatGPT za izradu testova za provjeru znanja, a zatim ih predstavili sveučilišnim profesorima. Nakon pregleda, 120 sudionika odgovorilo je na samo dva pitanja koja su se odnosila na prihvatljivost korištenja ove napredne tehnologije za izradu testova. Rezultati su pokazali da je čak 85% ispitanika podržalo ChatGPT za izradu cjelovitih ispita znanja, a 98% ih izrazilo želju za korištenjem ChatGPT-a za savjetovanje, podršku i doradu testova koje su sami kreirali (Aboalela, 2023). Navedeno navodi na zaključak kako se među učiteljima stvara potreba za upotrebom ChatGPT-a u izradi testova za provjeru znanja, međutim, još uvijek nije poznato koje točno koristi i izazove ova tehnologija sa sobom nosi u proučavanoj domeni nastavničkog posla.

3 Korištenje ChatGPT-a za generiranje ispitnih pitanja

U nastavku slijedi prikaz rada ChatGPT-a u području generiranja ispitnih pitanja. Da bi se došlo do što konkretnijih rezultata, inteligentnom sustavu se daje konkretan primjer teksta iz knjige *Biciklom u promet: Priručnik za učenike vozače bicikla*, autora Vinković i sur. (2009). Ovaj se priručnik inače koristi u sklopu predmeta Tehnička kultura za peti razred osnovne škole, u okviru nastavne cjeline "Sigurnost u prometu". Za potrebe ove analize odabran je tekst iz prvog poglavlja rada pod naslovom "Bicikl" koji opisuje dijelove i opremu bicikla te tehniku vožnje biciklom.

U prvom koraku, odabrani tekst je unesen u ChatGPT zajedno s precizno postavljenim zahtjevima za izradu ispitnih pitanja. Prema uputama prikazanim na Slici 1, od ChatGPT-a je zatraženo da iz ponuđenog teksta izradi 10 ispitnih pitanja. Tražena pitanja trebaju uključivati dopunu rečenica, mogućnosti odabira točnih odgovora, pitanja s kratkim odgovorom, pitanja s višestrukim izborom te jedno esejsko pitanje. Dodatno je traženo da se ispit znanja prilagodi učenicima osmog razreda osnovne škole te da udovoljava postavljenim zahtjevima težine.

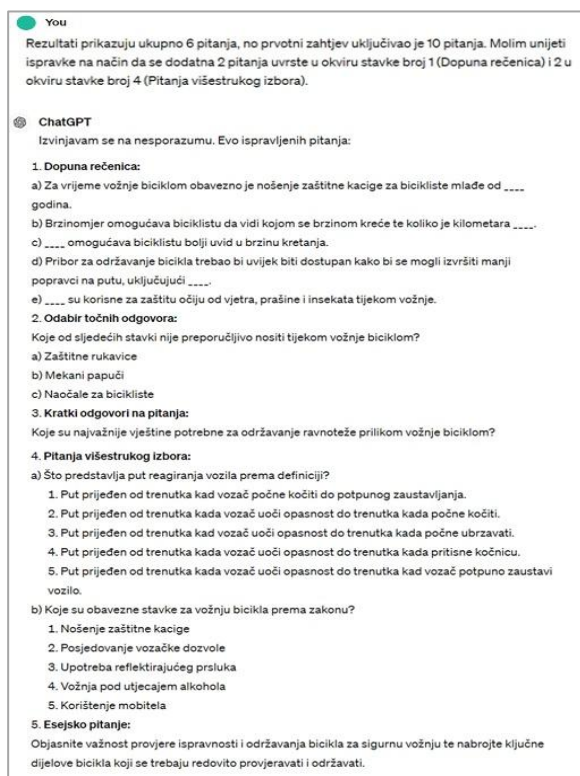
Nakon nekoliko trenutaka od postavljanja zahtjeva, pristigao je odgovor ChatGPT-a, koji je cjelovito prikazan na Slici 1. Inteligentni sustav je razvrstao pitanja prema zahtjevu, no detaljnim pregledom rezultata primijećeno je da je pristiglo samo šest od 10 zatraženih ispitnih pitanja.



Slika 1. Postavljanje zahtjeva za izradu ispitnih pitanja i prvi rezultati dobiveni od ChatGPT-a (OpenAI, 2024)

Stoga je izrađen novi zahtjev (Slika 2), unutar kojeg je istaknuta greška i preciziran način unošenja ispravki. U ponovljenom zahtjevu je traženo proširivanje predloženog ispita za ukupno četiri pitanja, po dva u obliku dopune rečenica i pitanja višestrukog izbora. Pristigli odgovor od strane ChatGPT-a (Slika 2) ovog puta udovoljavao je svim postavljenim kriterijima. Iz priloženog se može konstatirati kako ovaj inteligentni sustav nudi brzu pomoć pri izradi testova za provjeru znanja, a prediktor za dobivanje kvalitetnih rezultata je postavljanje ispravnih i jasnih kriterija.

Sustavu nisu upućivani novi zahtjevi, ali je izvršena kratka usporedba ponuđenih ispitnih pitanja s tekstom koji je poslužio kao osnova za njihovu izradu. Ustanovljeno je da su sva pitanja u potpunosti proizašla iz konkretnog teksta te da su obuhvatila različite razine zahtjevnosti. Međutim, da bi se potvrdilo jesu li zadovoljeni uvjeti glede zahtjevnosti, poput 30% pitanja najlakšeg stupnja težine i 20% pitanja umjerene težine, potrebna je dodatna analiza koja neće biti provedena u okviru ovog rada. Međutim, to predstavlja intrigantno područje za buduća istraživanja.



Slika 2. Ponovljeni rezultati ChatGPT-a nakon zatraženih ispravki (OpenAI, 2024)

4 Metodologija istraživanja

4.1 Predmet i ciljevi istraživanja

Nakon što je sekundarnim istraživanjem utvrđeno kako su se nova tehnološka rješenja poput e-dnevnika, sustava za upravljanje učenjem, interaktivnih simulacija, automatiziranih alata za ocjenjivanje i umjetne inteligencije pokazala kao korisni resursi za nastavnike, sada je vrijeme istražiti kako ti alati mogu doprinijeti poboljšanju nastavnog procesa i smanjenju ukupnog obujma posla. Ranija istraživanja pokazala su kako tehnološka rješenja omogućuju učiteljima učinkovitije upravljanje nastavnim materijalima, smanjenje vremena potrebnog za administrativne zadatke te prilagodbu nastave individualnim potrebama učenika.

Istraživanje u ovom radu provedeno je s ciljem ispitivanja percepcija učitelja o korisnosti ChatGPT-a za izradu testova provjere znanja. Fokus je stavljen na identificiranje prednosti i izazova korištenja umjetne inteligencije u ovom kontekstu te utvrđivanje kako ChatGPT može olakšati proces kreiranja pojedinih ispitnih pitanja ili cjelovitih testova provjere znanja. Osim toga, istražen je i utjecaj ChatGPT-a na uštedu vremena i smanjenje obujma posla, kao i spremnost učitelja za uvođenje novih tehnologija poput ChatGPT-a u svoj budući nastavnički rad.

U skladu s ciljevima istraživanja postavljeno je i polazišno istraživačko pitanje koje glasi: *Koje su prednosti i nedostaci koje učitelji uočavaju u korištenju ChatGPT-a za izradu testova provjere znanja i koliko su spremni integrirati ovu tehnologiju u svom budućem radu?*

4.2 Uvid u faze istraživačkog protokola

Cjelokupan istraživački protokol sastojao se od tri međusobno povezane faze prikazane u Tablici 1. Nakon sekundarnog istraživanja dostupne literature, u prvoj fazi postavljeni su istraživački ciljevi i izrađen je plan istraživanja.

S obzirom da je ChatGPT relativno nova tehnologija i pod pretpostavkom da ga većina učitelja ranije nije koristila u navedene svrhe, kreiran je dokument unutar kojeg je na konkretnom primjeru opisan način upotrebe ChatGPT-a za izradu testova provjere znanja. Za te potrebe korišten je primjer ranije prikazan u ovom radu, uz dodatna pojašnjenja.

Faza I.	Postavljanje ciljeva i izrada plana istraživanja
Faza II.	Izrada anketnih pitanja pomoću ChatGPT-a
Faza III.	Provedba istraživanja

Tablica 1. Faze istraživačkog protokola

Potom je, u drugoj fazi istraživačkog protokola izrađen anketni upitnik. Uzimajući u obzir činjenicu da se ovaj rad bavi upravo procjenom korisnosti ChatGPT-a za izradu ispitnih pitanja, odlučeno je da će se upravo ova tehnologija koristiti i prilikom kreiranja pitanja za anketni upitnik. Za dobivanje konačnih rezultata, odnosno pitanja koja su korištena za istraživanje stavova i percepcija učitelja u ovom radu, trebalo je svega nekoliko minuta. Prvi upit sadržavao cilj istraživanja, kratak opis plana istraživanja te konkretne i jasne upute ChatGPT-u o dijelovima anketnog upitnika, vrsti (socio-demografska pitanja, pitanja o postojećim načinima korištenja i izrade testova za procjenu znanja te percepcije i stavove o korisnosti ChatGPT-a u istovjetne svrhe) i broju pitanja koja bi upitnik trebao sadržavati. S obzirom da prva verzija odgovora nije bila zadovoljavajuća, kreiran je novi upit u kojem je zatražena izrada većeg broja varijabli procjene koje će opisati različite domene korisnosti ChatGPT-a u izradi testova za procjenu znanja. Nakon ponovljenog upita, inteligentni sustav je generirao ispravljene rezultate koji su, prema mišljenju autora zadovoljili postavljene kriterije. Naposljetku su generirana pitanja stavljena u online anketni upitnik u okviru platforme *Google obrasci*.

U posljednjoj fazi uslijedila je provedba istraživanja na način da je posredstvom e-pošte, Facebook grupe učitelja te putem *WhatsApp* grupa proslijeđena zamolba za sudjelovanjem u istraživanju. U kratkom uvodnom tekstu potencijalni ispitanici obavješteni su o ciljevima istraživanja i prikazan im je uvid u korake koje trebaju zadovoljiti prije pristupa ispunjavanju anketnog upitnika. Od onih koji nikada nisu koristili ChatGPT za potrebe izrade ispitnih pitanja zatraženo je da detaljno pročitaju priloženi primjer (oni koji već imaju iskustva zamoljeni su da odmah prijeđu na ispunjavanje anketnog upitnika), a zatim da se prijave na ChatGPT i na sličan način samostalno testiraju njegove performanse i učinkovitost. Posljednji korak sastojao se od ispunjavanja anketnog upitnika i izražavanja vlastitih percepcija i iskustava s ChatGPT-om.

Istraživanje je provedeno u svibnju 2024. godine i obuhvatilo je uzorak od 52 učitelja osnovnih škola iz cijele Hrvatske. Prikupljeni podaci naknadno su obrađeni pomoću *Excela* te interpretirani primjenom deskriptivne statistike, kao i tabličnih i grafičkih prikaza.

4.3 Instrument istraživanja

Anketni upitnik (u prilogu) sastojao se od ukupno 11 pitanja podijeljenih u tri dijela. Prvi dio sadržavao se od pet pitanja iz socio-demografske domene (spol, dob, status obrazovanja, duljina radnog staža u nastavničkoj profesiji i radni položaj) s ukupno 12 varijabli procjene.

U drugom dijelu upitnika od učitelja je zatraženo da se očituju o postojećim načinima korištenja testova za provjeru znanja (samostalna izrada ili upotreba gotovih, standardiziranih testova) te da izvrše samoprocjenu vlastite stručnosti pri izradi testova za provjeru znanja.

Posljednji dio anketnog upitnika sastojao se od tri pitanja, pri čemu je korištena Likertova ljestvica procjene varijabli korisnosti ChatGPT-a pri izradi testova za provjeru znanja u različitim domenama, istražena su mišljenja o njegovim nedostacima te mogućnosti da učitelji nastave koristiti ovaj alat i u svom budućem nastavničkom radu. Za navedenu skalu izračunat je Cronbachov alfa koeficijent pouzdanosti, koji iznosi 0,916. Ova vrijednost ukazuje na vrlo visoku unutarnju konzistentnost skale, što znači da su ispitanici dosljedno odgovarali na tvrdnje koje mjere sličan konstrukt.

4.4 Podaci o uzorku

Tablica 2 sadrži socio-demografske karakteristike uzorka. U istraživanju je sudjelovalo ukupno 52

učitelja osnovnih škola u Hrvatskoj, od čega 76,9% žena i 23,1% muškaraca. Prosječna starosna dob ispitanika iznosila je 42 godine. Većina ispitanika prijavila je visoku stručnu spremu ili magisterij (75%), a ostali imaju završenu višu stručnu spremu (VSS, 25%). Distribucija rezultata o duljini radnog staža u profesiji izrazito je heterogena. U ukupnom uzorku, 36,5% učitelja u profesiji radi između pet i 10 godina, 34,6% više od 10 godina, a preostalih 28,8% manje od 5 godina.

Spol		n	frekv.
žensko		40	76,9%
muško		12	23,1%
Starosna dob			
min	max	M±SD	
26	57	42.65 ± 7.84	
Stupanj obrazovanja		n	frekv.
VŠS		13	25%
VSS/Magisterij		39	75%
Duljina radnog staža		n	frekv.
Manje od 5 godina		15	28,8%
5-10 godina		19	36,5%
Više od 10 godina		18	34,6%
Radno mjesto		n	frekv.
Učitelj/ica razredne nastave		21	40,4%
Predmetni nastavnik/ca		31	59,6%

Tablica 2. Socio-demografske karakteristike uzorka

Što se tiče radnog mjesta, istraživanju je pristupilo 59,6% predmetnih učitelja/ica te 40,4% učitelja/ica predmetne nastave.

5 Rezultati

5.1 Postojeći načini korištenja i izrade testova za provjeru znanja

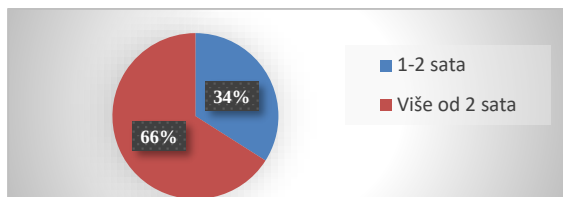
Rezultati su pokazali kako anketirani učitelji u većoj mjeri koriste samostalno izrađene testove za provjeru znanja u odnosu na gotove, standardizirane testove koji su dostupni od strane nakladnika, u udžbenicima i radnim bilježnicama (Tablica 3).

Vrsta testova za provjeru znanja	M±SD
Samostalno izrađeni testovi	4,04 ± 0.80
Gotovi, standardizirani testovi	2,42 ± 1.18

Tablica 3. Učestalost korištenja pojedinih vrsta testova za provjeru znanja

Većina učitelja (66%) istaknula je kako im za samostalnu izradu jednog testa treba više od dva sata

vremena, dok je preostalima (34%) za to potrebno između jednog i dva sata (Slika 3).



Slika 3. Prosječno vrijeme potrebno za samostalnu izradu jednog testa za provjeru znanja

Varijable procjene	M	IQR	Mod	Med
Razumijevanje nastavnih planova i programa	4,57	1	5	5
Kreativnost u osmišljavanju pitanja	3,57	1	4	4
Preciznost u postavljanju pitanja	4,10	1	4	4
Poznavanje različitih metoda evaluacije znanja	4,42	1	5	5
Sposobnost prilagodbe testova različitim razinama znanja učenika	4,20	1	4	4
Prosječna vrijednost svih varijabli	M=4,17			

Tablica 4. Deskriptivna statistika procjene vlastite stručnosti učitelja pri izradi testova za provjeru znanja

Samoprocjena stručnosti učitelja u samostalnoj izradi testova za provjeru znanja (Tablica 4) pokazala je visoku prosječnu ocjenu ($M=4.17$), pri čemu su najviše ocijenili dodijeliti razumijevanju nastavnih planova i programa ($M=4.57$) te poznavanju različitih metoda evaluacije znanja ($M=4.42$). S druge strane, kreativnost u osmišljavanju pitanja dobila je najnižu prosječnu ocjenu ($M=3.57$).

5.2 Procjena korisnosti i nedostataka ChatGPT-a u izradi testova za provjeru znanja

Tablica 5 sadrži prosječne ocjene procjene različitih domena korisnosti ChatGPT-a pri izradi testova za procjenu znanja. Prosječna vrijednost za svih 14 procijenjenih varijabli iznosi 4.11 što označava visok stupanj korisnosti.

Zanimljiva je činjenica da su učitelji izrazili srednju razinu zadovoljstva ($M=3.73$) prvom verzijom rezultata koje su dobili samostalnim korištenjem ChatGPT-a. Ovo je zadovoljavajući rezultat ako se uzme u obzir činjenica da je i prilikom izrade ovog rada (primjer korištenja ChatGPT-a i izrada anketnog

upitnika) prva verzija pitanja koje je ChatGPT ponudio bila nedostatna i od njega su zatražene daljnje modifikacije. Nadalje, pokazalo se da su učitelji visoko zadovoljni jednostavnošću korištenja ChatGPT-a te činjenicom da generira brze rezultate ($M=4.48$), potom podudarnošću generiranih pitanja sa predloženim tekstom ($M=4.58$), kao i vlastitom sposobnošću za učinkovito korištenje ovog inteligentnog sustava ($M=4.54$). Međutim, nešto niži stupanj zadovoljstva iskazali su spram raspona opcija za prilagodbu ($M=3.71$) i preciznosti generiranih pitanja ($M=3.84$).

Varijable procjene	M	IQR	Mod	Med
Zadovoljan/na sam prvom verzijom rezultata koje sam dobio/la korištenjem ChatGPT-a.	3,73	1	4	4
ChatGPT je jednostavan za korištenje i generira brze rezultate.	4,48	1	5	5
Ispitna pitanja generirana pomoću ChatGPT-a u potpunosti su proizašla iz predloženog teksta.	4,58	1	5	5
Rezultati koje sam dobio/la korištenjem ChatGPT-a su precizni.	3,84	1	4	4
Ispitna pitanja koje sam dobio/la korištenjem ChatGPT-a su kreativno osmišljena.	3,92	2	4	4
Kvaliteta ispitnih pitanja generiranih pomoću ChatGPT-a je zadovoljavajuća.	4,02	0,25	4	4
ChatGPT je pružio zadovoljavajući raspon opcija za prilagodbu ispitnih pitanja.	3,71	1	4	4
Rezultati generirani pomoću ChatGPT-a u skladu su s mojim očekivanjima.	3,90	1	4	4
ChatGPT je može biti koristan alat za uštedu vremena pri izradi testova.	4,22	1	5	4
ChatGPT bi mogao biti koristan alat za poboljšanje kvalitete testova koje izrađujem.	4,23	1	5	4
Zbog svog znanja i iskustva sposoban/na sam za učinkovito korištenje ChatGPT-a.	4,54	1	5	5
Korištenje ChatGPT-a za izradu testova može poboljšati moju produktivnost.	4,02	0	4	4
Korištenje ChatGPT-a može smanjiti ukupni obujam nastavnčkog posla.	4,46	1	5	5
Vjerujem da bi se uvođenjem ChatGPT-a kao alata za izradu testova poboljšala kvaliteta moje nastave.	3,86	1	4	4
Prosječna vrijednost svih varijabli	M=4,11			

Tablica 5. Deskriptivna statistika procjene različitih domena korisnosti ChatGPT-a pri izradi testova za procjenu znanja

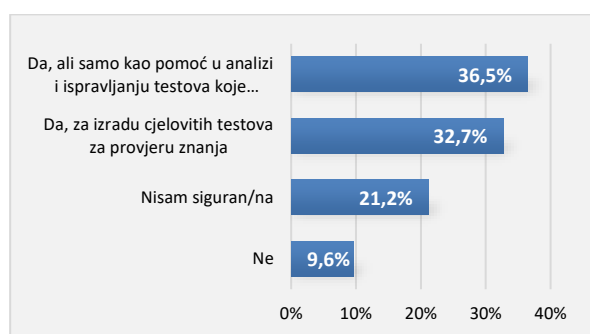
Što se tiče same korisnosti ChatGPT-a s aspekata: uštede vremena za izradu testova za provjeru znanja, smanjenja obujma posla, povećanja produktivnosti učitelja te povećanja kvalitete testova koje učitelji sami izrađuju, rezultati pokazuju da bi alat mogao biti visoko koristan. Naime, korisnost ChatGPT-a u smanjenju obujma nastavnčkog posla učitelja dobila je ocjenu 4.46, a njegova korisnost u domeni uštede vremena ocijenjena je ocjenom 4.22. Prema rezultatima, učitelji također smatraju kako ChatGPT može poboljšati njihovu produktivnost pri izradi testova ($M=4.02$).

U Tablici 6 nalaze se prijavljeni nedostaci ChatGPT-a pri izradi testova za provjeru znanja svrstani prema veličini udjela u ukupnom uzorku.

Vrste prijavljenih nedostataka	n	frekv.
Ovisnost o kvaliteti ulaznih podataka	37	71,2%
Nedostatak prilagodljivosti specifičnim potrebama	15	28,8%
Ograničenja u kreativnosti	11	21,2%
Za sada nema ograničenja	11	21,2%
Ograničenja u razumijevanju složenih pitanja	10	19,2%
Poteškoće s generiranjem kvalitetnih odgovora	10	19,2%
Nedostatak kontekstualne osjetljivosti	8	15,4%
Netočnosti pri generiranju pitanja	7	13,5%

Tablica 6. Prijavljeni nedostaci ChatGPT-a pri izradi testova za provjeru znanja

Najčešće prijavljen nedostatak je ovisnost ChatGPT-a o kvaliteti ulaznih podataka (71,2%), potom, u značajno nižim postotcima slijede nedostaci poput neprilagođenosti specifičnim potrebama korisnika (28,8%), ograničenja u kreativnosti pitanja (21,2%), Ograničenja u razumijevanju složenih pitanja (19,2%) i slično. Međutim, 21,2% učitelja istaknulo je kako prilikom prvog korištenja nisu uočili nikakva ograničenja.



Slika 4. Mogućnosti korištenja ChatGPT-a za izradu testova za provjeru znanja u budućnosti

Naposlijetku, učitelji su se očitovali o mogućnosti korištenja ChatGPT-a za izradu testova za provjeru znanja u svom budućem nastavničkom radu (Slika 4). Preko polovice učitelja dalo je potvrđan odgovor, pri čemu ih je 36,5% istaknulo kako će koristiti ovaj inteligentni sustav za izradu cjelovitih testova, a ostali (32,7%) su rekli kako ga planiraju koristiti isključivo za analizu i ispravljanje testova koje prethodno sami naprave. Od preostalih učitelja 21,2% je izjavilo da nije sigurno hoće li koristiti ChatGPT, a 9,6% je odlučno utvrdilo da neće.

5.3 Ograničenja istraživanja

Istraživanje provedeno u ovom radu ima nekoliko ograničenja. Prvo i osnovno odnosi se na veličinu uzorka, koja je relativno mala i time ograničava mogućnost generalizacije dobivenih rezultata na širu populaciju učitelja. Dodatno, određena metodološka ograničenja proizlaze i iz samog istraživačkog instrumenta. Naime, korišten je anketni upitnik zatvorenog tipa, što je ispitanike usmjerilo isključivo na unaprijed ponuđene odgovore, bez mogućnosti izražavanja osobnih stavova, nijansi mišljenja ili dodatnih komentara koji bi mogli obogatiti interpretaciju rezultata.

6 Rasprava

Istraživanje je pokazalo su kako učitelji u osnovnoj školi skloniji samostalnoj izradi testova za provjeru znanja nego korištenju gotovih, standardiziranih testova, što se poklapa s ranijim otkrićima ranijih studija (Marso i Pigge, 1988; Salmon-Cox, 1981). Međutim, za razliku od drugih istraživanja koja su procjenjivala stručnost učitelja u samostalnoj izradi testova za provjeru znanja i koja su otkrila brojne nedostatke (Marco i Pigge, 1992), samoprocjena učitelja u okviru ovog istraživanja ukazuje na visok stupanj stručnosti. Ipak, ovi se podaci trebaju uzeti s rezervom obzirom da je u pitanju samoprocjena, što je nerijetko suviše subjektivna metoda procjene njihove stručnosti i znanja.

S obzirom na visok stupanj zadovoljstva učitelja načinom rada, brzinom, preciznošću, jednostavnošću korištenja i kvalitetom generiranih odgovora evidentno je kako ChatGPT ima dobre predispozicije za masovno korištenje među učiteljima u budućnosti. Ovdje valja uzeti u obzir da je ChatGPT ocijenjen izuzetno visoko u pogledu brzine i jednostavnosti korištenja ($M=4.48$), čime izravno odgovara na vremenski zahtjevnju prirodu samostalne izrade testova. Učitelji su također izrazili osjećaj sposobnosti učinkovitog korištenja ovog alata ($M=4.54$), što je

slično ranije izraženom povjerenju u vlastite vještine izrade testova.

Međutim, valja uzeti u obzir i pojedine nedostatke ovog alata. Naime, početna kvaliteta rezultata koje je generirao ChatGPT ($M=3.73$) i ograničene mogućnosti prilagodbe ($M=3.71$) upućuju na to da, iako je alat brz i praktičan, možda ne može u potpunosti zamijeniti kvalitetu i prilagodljivost testova koje dizajniraju sami učitelji. Zanimljivo je da, iako su učitelji nižom prosječnom ocjenom procijenili vlastitu kreativnost u izradi testova, nisu smatrali da je ChatGPT značajno kreativniji ($M=3.92$). To ujedno implicira da ovaj alat trenutno najbolje služi kao pomoćnik za uštedu vremena, a ne kao potpuna zamjena za ljudski doprinos.

Kako je riječ o relativno novom alatu temeljenom na umjetnoj inteligenciji, može se pretpostaviti kako će neki njegovi nedostaci, koji su identificirani i u ovom istraživanju, u novijim verzijama ChatGPT-a biti uklonjeni, čime će se povećati kvaliteta predloženih rezultata.

Imajući na umu da nove tehnologije postaju sve veći dio ljudske svakodnevnice i činjenicu da su brojne od njih već postale sastavni dio nastavnog procesa, s gotovo potpunom sigurnošću se može tvrditi kako će i ChatGPT naći svoj "ulaz" u nastavni proces, radilo se o generiranju ispitnih pitanja ili pak o pouci učenika novim tehnologijama. Njegovo uvođenje u planiranje nastave na neki način potvrđeno je i ovim istraživanjem, s obzirom da je 69,2% učitelja izjavilo kako će ga nastaviti koristiti i u budućnosti. Do sličnih zaključaka došla je i Aboalela (2023) čije je istraživanje pokazalo da 85% učitelja planira koristiti ChatGPT za izradu cjelovitih testova za provjeru znanja, a čak 98% ih je reklo kako će ga koristiti kao pomoć u analizi i ispravljanju vlastitih, samostalno izrađenih testova.

7 Zaključak

Obrađeni rezultati ukazuju na visoku korisnost ChatGPT-a u izradi testova za procjenu znanja, međutim, to se ne može potvrditi bez dodatnih testova i većeg uzorka ispitanika. Njegove glavne prednosti leže u jednostavnosti korištenja, brzini kojom generira pitanja te činjenici da su ta pitanja formulirana isključivo na temelju zadanog teksta. Ove značajke uvelike smanjuju vrijeme potrebno za izradu testova i mogu biti posebno korisne za učitelje koji se često suočavaju s vremenskim ograničenjima. Međutim, rezultati također ističu određena ograničenja, osobito u pogledu kreativnosti i raznolikosti generiranih pitanja. To se poklapa s izazovima s kojima se i sami učitelji često susreću pri samostalnoj izradi testova, što ukazuje na to da, iako

ChatGPT može biti snažna podrška, još uvijek ne može u potpunosti zamijeniti pedagošku intuiciju i kreativnu dubinu ljudskog edukatora.

Svakako, uzimajući u obzir dobivene rezultate te činjenicu da ChatGPT s duljim korištenjem upoznaje preferencije korisnika i generira sve kvalitetnije rezultate temeljene na korisničkim upitima, preostaje samo pretpostaviti da će se zadovoljstvo anketiranih učitelja koji su potvrdili da će ga nastaviti koristiti, u budućnosti dodatno povećavati.

Međutim, unatoč identificiranim koristima ChatGPT-a u izradi testova za provjeru znanja, potrebno je oprezno tumačiti rezultate. Studija se temelji na subjektivnim samoprocjenama sudionika i provedena je u relativno uskom odgojno-obrazovnom kontekstu, s ograničenim uzorkom. Stoga se rezultati ne mogu automatski generalizirati na sve razine odgoja i obrazovanja, predmetna područja ili kulturne kontekste.

S obzirom da je ovo područje istraživanja još uvijek relativno novo, ovaj rad daje značajan doprinos rasvjetljavanju mogućnosti i izazova korištenja ChatGPT-a u odgojno-obrazovnoj domeni. Preporučuje se da buduće studije istraže učinkovitost testova generiranih pomoću ChatGPT-a u različitim odgojno-obrazovnim okruženjima, njihov utjecaj na uspjeh učenika te etička pitanja vezana uz primjenu umjetne inteligencije u procjeni znanja. Posebno je važno istražiti na koje se načine stručnost nastavnika i tehnologija mogu kombinirati tako da se ne umanjuje pedagoška vrijednost, već da se ona dodatno unaprijedi. S obzirom da je Likertova skala procjene tvrdnji korištena u ovom istraživanju izrađena od strane samog ChatGPT-a te da ima visok Cronbachov alfa koeficijent pouzdanosti (0,916), preporučuje ga se koristiti i u budućim istraživanjima. Visok koeficijent pouzdanosti ukazuje na stabilnost i koherentnost tvrdnji unutar skale, čime se potvrđuje njena korisnost kao alata za procjenu stavova korisnika o upotrebi ChatGPT-a u odgojno-obrazovnom kontekstu.

Literatura

- Abid, A., Farooqi, M., Zou, J. (2021). Persistent anti-muslim bias in large language models. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 298-306. Retrived from <https://dl.acm.org/doi/abs/10.1145/3461702.3462624>
- Aboalela, R. A. (2023). *ChatGPT For Generating Questions and Assessments Based On Accreditations*. King Abdulaziz University: Department of Information Systems. Retrived from <https://arxiv.org/pdf/2312.00047>

- Abu Ghali, M. J., Abu Ayyad, A., Abu-Naser, S. S., Abu Laban, M. (2018). An Intelligent Tutoring System for Teaching English Grammar. *International Journal of Academic Engineering Research*, 2(2), 1-6.
- Alhabbash, M. I., Mahdi, A. O., Naser, S. S. A. (2016). An Intelligent Tutoring System for Teaching Grammar English Tenses. *European Academic Research*, 4(9), 1–15.
- Alnaqbi, N.M., Fouda, W. (2023). Exploring the Role of ChatGPT and social media in Enhancing Student Evaluation of Teaching Styles in Higher Education Using Neutrosophic Sets. *International Journal of Neutrosophic Science*, 20(4), 181–190.
- Amitai, M., Van Houtte, A. (2022). Being pushed out of the career: Former teachers' reasons for leaving the profession. *Teaching and Teacher Education*, 110: 103540. Retrived from <https://www.sciencedirect.com/science/article/abs/pii/S0742051X21002651>
- Bezirhan, U., Von Davier, M. (2023). Automated reading passage generation with Open AI's large language model. *Computers and Education: Artificial Intelligence*, 5 (100161). Retrived from <https://www.sciencedirect.com/science/article/pii/S2666920X23000401>
- Brandenburg, R., Larsen, E., Simpson, A., Sallis, R., Tran, D. (2024). 'I left the teaching profession ... and this is what I am doing now': a national study of teacher attrition. *The Australian Educational Research*. Retrived from <https://link.springer.com/article/10.1007/s13384-024-00697-1>
- Brown, T. et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems (NeurIPS 2020)*, Vancouver, Canada. Retrived from <https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf>
- Centar za istraživanje i inovacije u obrazovanju, (2008). Assessment for Learning Formative Assessment. *OECD/CERI International Conference "Learning in the 21st Century: Research, Innovation and Policy"*. Retrived from <https://www.oecd.org/site/educeri21st/40600533.pdf>
- Chiu, T.K.F., Moorhouse, B.L., Chai, C.S., Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, Advanced online publication. Retrived from https://www.researchgate.net/publication/368306653_Teacher_support_and_student_motivation_to_learn_with_Artificial_Intelligence_AI_based_chatbot
- Cotton, D. R., Cotton, P. A., Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1– 12. Retrived from https://www.researchgate.net/publication/367030297_Chatting_and_Cheating_Ensuring_academic_integrity_in_the_era_of_ChatGPT
- Crouse-Machcinski, K. (2019). The Benefits of Utilizing Learning Management Systems in Peer Tutor Training. *Learning Assistance Review*, 24(2), 73-84.
- Dave, T., Athaluri, S.A., Singh, S. (2023). ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Frontiers in Artificial Intelligence*, 6(1695959), 1-5. Retrived from <https://www.researchgate.net/publication/370527457>
- Duarte, F. (2024). Number of ChatGPT Users (Apr 2024). *explodingtopics.com*. Retrived from <https://explodingtopics.com/blog/chatgpt-users>
- Famaye, T., Bailey, C.S., Adisa, I., Irgens, G.A. (2024). "What makes ChatGPT dangerous is also what makes it special": High-school student perspectives on the integration or ban of artificial intelligence in educational contexts. *International Journal of Technology in Education (IJTE)*, 7(2), 174-199.
- Fan, X., Liang, Y. (2023). The Research on the Characteristics of AI Application in Art Field and Its Value. U: B. Majoul, Guan, G. i Groom, N. (ur.), *Proceedings of the 4th International Conference on Language, Art and Cultural Exchange - ICLACE 2023* (pp. 146-160.). Vol. 766. Paris: Atlantis Press.
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1-15. Retrived from <https://edepot.wur.nl/591282>
- Gonzalez, L., Brown, M. S., Slate, J. R. (2008). Teachers Who Left the Teaching Profession: A Qualitative Understanding. *The Qualitative Report*, 13(1), 1-11.
- Grepcka, A., Elsayed, E.E., Minga, J., Basholli, AA., Daberdini, A. (2024). The use of Chat GPT in the preparation of tests for the control of students' knowledge in Electronics and Information

- Technology. "Optime" Revistë shkencore e "Albanian University", 1, 1-34.
- Gronlund, N.E., Linn, R. L. (1990). *Measurement and evaluation in teaching*. 6th ed. New York: Macmillan.
- Harry, A. (2023). Role of AI in Education. *Injury: Interdisciplinary Journal and Humanity*, 2(3), 260-268.
- Heffernan, A., Bright, D., Kim, M., Longmuir, F., Magyar, B. (2022). 'I cannot sustain the workload and the emotional toll': Reasons behind Australian teachers' intentions to leave the profession. *Australian Journal of Education*, 66(2), 196-209.
- Karakose, T., Tülübaş, T. (2023). How Can ChatGPT Facilitate Teaching and Learning: Implications for Contemporary Education. *Educational Process: International Journal*, 12(4), 7-16.
- Kaur, P., Nadarajan, R. (2020). Language Learning And Teaching Using Kahoot!. *International Journal of Modern Education*, 2(5), 19-28.
- Kececi, B., Bilge Keskinilic, S. K. (2023). Why Do Teachers Leave the Profession: A Case Study. *International Journal of Educational Research Review*, 8(4), 773-787.
- Khan, M.S., Umer, H. (2024). ChatGPT in finance: Applications, challenges, and solutions. *Heliyon*, 10(2), 1-8.
- Kurt, S. (2010). Technology use in elementary education in Turkey: A case study. *New Horizons in Education*, 58(1), 65-76.
- Lee, C. M. (2020). *Qualitative Research Analysis of using Electronic Class Record for Assessments and Evaluations*. Technological University of Philippines- Manila. Retrived from <https://www.researchgate.net/publication/343583629>
- Li, L., Bamman, D. (2021). Gender and representation bias in GPT-3 generated stories. *Proceedings of the Third Workshop on Narrative Understanding. Proceedings of the 3rd Workshop on Narrative Understanding*, pp. 48-55. Retrived from <https://par.nsf.gov/servlets/purl/10237395>
- Longmuir, F., Gallo Cordoba, B., Phillips, M., Allen, K. A., Moharami, M. (2022). *Australian teachers' perceptions of their work in 2022*. Monash University. Retrived from https://www.monash.edu/_data/assets/pdf_file/0008/3061169/Teachers-Perceptions-of-their-Work-2022.pdf
- Marso, R N., Pigge, F. L. (1988). *An analysis of teacher-made tests: Testing practices, cognitive demands and item construction errors*. Paper presented at the annual conference of the National Council on Measurement in Education, New Orleans. (ERIC Document Reproduction Service No. ED 298 174). Retrived from <https://files.eric.ed.gov/fulltext/ED298174.pdf>
- Marso, R. N., Pigge, F. L. (1992). *A Summary of Published Research: Classroom Teachers' Knowledge and Skills Related to the Development and Use of Teacher-Made Tests*. Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992). Retrived from <https://files.eric.ed.gov/fulltext/ED346148.pdf>
- Ministarstvo znanosti i obrazovanja (2017). *Nacionalni kurikulum za osnovnoškolski odgoj i obrazovanje (Prijedlog nakon javne rasprave)*. Zagreb: MZO. Retrived from <https://mzom.gov.hr/>
- Mountcastle, T. (2024). How To Use ChatGPT For Travel Planning. *forbes.com*. Retrived from <https://www.forbes.com/advisor/credit-cards/travel-rewards/chatgpt-for-travel-planning/>
- Murphy, D. H., Little, J. L., Bjork, E. L. (2023). The Value of Using Tests in Education as Tools for Learning—Not Just for Assessment. *Education in Psychology Review*, 35(89). Retrived from <https://link.springer.com/article/10.1007/s10648-023-09808-3>
- Naqvi, A. (2020). *Artificial intelligence for audit, forensic accounting, and valuation: A strategic perspective*. John Wiley & Sons.
- Opara, E., Adalikwu M. E-T., Tolorunleke, C. A. (2023). ChatGPT for teaching, learning and research: Prospects and challenges. *Global Academic Journal of Humanities and Social Sciences*, 5(2), 33-40.
- OpenAI (2024). Introducing CHAT GPT. *openai.com*. Retrived from <https://openai.com/blog/chatgpt>
- Ortiz, S. (2024). How ChatGPT (and other AI chatbots) can help you write an essay. *zdnnet.com*. Retrived from <https://www.zdnnet.com/article/how-to-use-chatgpt-to-write-an-essay/>
- Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., Bassey, B .A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *EURASIA Journal of Mathematics, Science and Technology Education*, 19(8), 1-15.
- Ozlem Saka, F. (2016). What do Teachers Think about Testing Procedure at Schools? *Procedia - Social and Behavioral Sciences*, 232, 575-582.

- Pedro, F., Subosa, M., Rivas, A., Valverde, P. (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*. Paris: United Nations Educational, Scientific and Cultural Organization. Retrived from <https://unesdoc.unesco.org/ark:/48223/pf0000366994>
- Pešorda, S. (2018). Strukturiranje pisanih provjera u nastavi povijesti. *Povijest u nastavi*, 29(1), 27-37.
- Perryman, J., Calvert, G. (2020). What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. London: Institute of Education. Retrived from https://discovery.ucl.ac.uk/id/eprint/10068733/3/Perryman_Teachers%20Leaving%20revised%20final.pdf
- Radeka, I., Sorić, I. (2006). Zadovoljstvo poslom i profesionalni status učitelja . *Napredak*, 147(2), 161-177.
- Rahim, F. R., Suherman, D. S., Muttaqiin, A. (2020). Exploring the effectiveness of e-book for students on learning material: a literature review. *Journal of Physics Conference Series*, 1481(1), 012105. Retrived from <https://iopscience.iop.org/article/10.1088/1742-6596/1481/1/012105/pdf>
- Rawson, K. A., Dunlosky, J. (2012). When is practice testing most effective for improving the durability and efficiency of student learning? *Educational Psychology Review*, 24(3), 419–435.
- Robinson, V., Hamilton, P. (2023). *Understanding and Reducing the Workload of Teachers and Leaders in Western Australian Public Schools*. The Western Australian Department of Education. Retrived from <https://www.education.wa.edu.au/dl/nxjk4vg>
- Rossi, O. (2023). *Using AI For Test Item Generation: opportunities and challenges*. European Association for Language Testing and Assessment (EALTA). Retrived from https://www.ealta.eu.org/documents/resources/webinars/ORossi_AI%20for%20item%20writing_presentation%20slides.pdf
- Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? *Phi Delta Kappan*, 62, 631-634.
- Sami, A., Uddin, I., Fayyaz, N., Bilal, M., Shahid, M., Ali, I. (2023). Getting to Know ChatGPT: An Introduction to Implementation and Working. *Proceedings of 1st International Conference on Computing Technologies, Tools and Applications (ICTAPP-23)*, 273-277. Retrived from <https://www.researchgate.net/publication/372890276>
- Shalevska, E. (2023). AI Language Models, Standardized Tests and Academic Integrity: A Chat (GPT). *International Journal of Education*, 26, 17-26.
- Sharma, R. (2023). How to Use Chat GPT: Step by Step Guide to Start Open AI ChatGPT. *jagranjosh.com*. Retrived from <https://www.jagranjosh.com/general-knowledge/how-to-use-chat-gpt-1675420609-1>
- Shihadeh, J., Ackerman, M., Troske, A., Lawson, N., Gonzalez, E. (2022). Brilliance bias in GPT-3. *IEEE Global Humanitarian Technology Conference*. Retrived from <https://maya-ackerman.com/wp-content/uploads/2022/09/BrillianceBiasGPT.pdf>
- Shohamy, E. (2000). The relationship between language testing and second language acquisition, revisited. *Systems*, 28, 541- 553.
- Simsek, A. (2016). A comparative analysis of common mistakes in achievement tests prepared by school teachers and corporate trainers. *European Journal of Science and Mathematics Education*, 4(4), 477-489.
- Smith, E. (2023). Number of teachers leaving the profession hits record high. *schoolbus.co.uk*. Retrived from <https://www.schoolbus.co.uk/news/featured-article/number-of-teachers-leaving-the-profession-hits-record-high/9649>
- Smitthers, A., Robinskon, P. (2003). *Factors Affecting Teachers' Decisions to Leave the Profession*. University of Liverpool: Centre for Education and Employment Research. Retrived from <https://dera.ioe.ac.uk/id/eprint/4759/1/RR430.pdf>
- Sok, S., Heng, K. (2023). ChatGPT for education and research: A review of benefits and risks. *Cambodian Journal of Educational Research [Online]* Retrived from https://www.researchgate.net/publication/369127881_ChatGPT_for_Education_and_Research_A_Review_of_Benefits_and_Risks
- Sutcher, L., Darling-Hammond, L., Carver-Thomas, D. (2019). Understanding teacher shortages: An analysis of teacher supply and demand in the United States. *Education Policy Analysis Archives*, 27(35), 1-39.
- Šimić Šašić, S., Klarin, M. (2008). Varanje u srednjim školama u Hrvatskoj i u Bosni i Hercegovini. *Društvena istraživanja*, 6(104), 999-1022.

- UNESCO (2023). *ChatGPT and Artificial Intelligence in higher education: Quick start guide*. Caracas: UNESCO International Institute for Higher Education in Latin America and the Caribbean. Retrived from https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf
- UNESCO (2024). *Global Report on Teachers: Addressing Teacher Shortages and Transforming The Profession*. Paris: United Nations Educational, Scientific and Cultural Organization. Retrived from <https://unesdoc.unesco.org/ark:/48223/pf0000388832>
- Vinković, M., Medved, Ž., Androlić, S., Labaš, D. (2009). *Biciklom u promet: Priručnik za učenike vozače bicikla*. Zagreb: Profil International.
- Vrkić Dimić, J., Vidov, S. (2019). E-dnevnik u školskoj praksi - mišljenja i iskustva učitelja srednjih škola. *ACTA Iadertina*, 16(1), 31-58.
- Wieczorek, B. (2022). Automated assessment system as a supportive tool for teaching programming in Academic Secondary School of SUT. *Edukacyjna Analiza Transakcyjna*, 11, 193-204.
- Zhu, S., Wang, Z., Zhuang, Y., Jiang, Y., Gue, M., Zhang, X., Gao, Z. (2024). Exploring the impact of ChatGPT on art creation and collaboration: Benefits, challenges and ethical implications. *Telematics and Informatics Report*, 14: 100138. Retrived from <https://www.sciencedirect.com/science/article/pii/S2772503024000240>
- Zimmerman, A. (2023). A Ghostwriter for the Masses: ChatGPT and the Future of Writing. *Annals of Surgical Oncology*, 30(6), 3170–3173.
- teachers in Croatia as part of this study. Most respondents were satisfied with the ease of use, the speed of result generation, and the accuracy of the generated questions, while they were least satisfied with ChatGPT's high dependence on the quality of input data. The study confirmed that ChatGPT saves time, reduces workload, and improves the quality of tests, with the majority of teachers expressing their willingness to continue using it.

Keywords: ChatGPT; AI; knowledge assessment tests; teaching process.

Modern AI Models in Education: Usefulness of ChatGPT in Creating Knowledge Assessment Tests

Abstract

Technological innovations, particularly artificial intelligence, are transforming education by facilitating teachers' work and enhancing the learning process. Artificial intelligence enables automated grading, personalized learning, and content generation, while ChatGPT assists in lesson planning, test creation, and knowledge assessment. To evaluate its usefulness in creating knowledge assessment tests, a survey was conducted among 52

PRILOG: Anketni upitnik

1. Spol:
 - Žensko
 - Muško
2. Starosna dob: _____
3. Stupanj obrazovanja:
 - VŠS
 - VSS
 - Magistar/magistra znanosti
 - Doktor/doktorica znanosti
4. Duljina radnog staža kao učitelj/predmetni nastavnik:
 - Manje od 5 godina
 - 5 do 10 godina
 - Više od 10 godina
5. Molim Vas, označite svoje radno mjesto:
 - Učitelj/ica razredne nastave
 - Predmetni nastavnik/ca
6. Izjasnite se o učestalosti korištenja standardiziranih testova znanja dostupnih u udžbenicima/radnim bilježnicama/itd. i samostalno izrađenih testova za provjeru znanja. (1 *Nikad*) (5-*Uvijek*)

Varijable procjene	1	2	3	4	5
Samostalno izrađeni testovi provjere znanja.					
Standardizirani testovi dostupni u udžbenicima, radnim bilježnicama i drugim nastavnim materijalima.					

7. Ukoliko ste prethodno naveli kako barem nekad samostalno izrađujete testove, izjasnite se koliko vam vremena u prosjeku treba za izradu jednog testa? (ako sami ne izrađujete testove, molim Vas preskočite ovo i sljedeće pitanje)
 - Manje od sat vremena
 - 1-2 sata
 - Više od 2 sata
8. Ukoliko ste prethodno naveli kako barem nekad samostalno izrađujete testove, molim Vas da na ljestvici od 1 (Slabo) do 5 (Izvršno) procijenite vlastitu stručnost pri izradi testova za provjeru znanja.

Varijable procjene	1	2	3	4	5
Razumijevanje nastavnih planova i programa.					
Kreativnost u osmišljavanju pitanja.					
Preciznost u postavljanju pitanja.					
Poznavanje različitih metoda evaluacije znanja.					
Sposobnost prilagodbe testova različitim razinama znanja učenika.					

9. Molim Vas, temeljem vlastitog iskustva u okviru testiranja ChatGPT-a, na skali od 1 (*U potpunosti se ne slažem*) do 5 (*U potpunosti se slažem*) procijenite korisnost ovog inteligentnog sustava u različitim domenama.

Varijable procjene	1	2	3	4	5
Zadovoljan/na sam prvom verzijom rezultata koje sam dobio/la korištenjem ChatGPT-a.					
ChatGPT je jednostavan za korištenje i generira brze rezultate.					
Ispitna pitanja generirana pomoću ChatGPT-a u potpunosti su proizašla iz predloženog teksta.					

Rezultati koje sam dobio/la korištenjem ChatGPT-a su precizni.

Ispitna pitanja koje sam dobio/la korištenjem ChatGPT-a su kreativno osmišljena.

Kvaliteta ispitnih pitanja generiranih pomoću ChatGPT-a je zadovoljavajuća.

ChatGPT je pružio zadovoljavajući raspon opcija za prilagodbu ispitnih pitanja.

Rezultati generirani pomoću ChatGPT-a u skladu su s mojim očekivanjima.

ChatGPT je može biti koristan alat za uštedu vremena pri izradi testova.

ChatGPT bi mogao biti koristan alat za poboljšanje kvalitete testova koje izrađujem.

Zbog svog znanja i iskustva sposoban/na sam za učinkovito korištenje ChatGPT-a.

Korištenje ChatGPT-a za izradu testova može poboljšati moju produktivnost.

Korištenje ChatGPT-a može smanjiti ukupni obujam nastavnčkog posla.

Vjerujem da bi se uvođenjem ChatGPT-a kao alata za izradu testova poboljšala kvaliteta moje nastave.

10. Koje biste nedostatke ChatGPT-a istaknuli? (*Mogućnost odabira većeg broja odgovora*)
 - Ograničenja u kreativnosti
 - Ovisnost o kvaliteti ulaznih podataka
 - Netočnosti u generiranju pitanja
 - Nedostatak prilagodljivosti specifičnim potrebama
 - Ograničenja u razumijevanju složenih pitanja
 - Nedostatak kontekstualne osjetljivosti
 - Poteškoće s generiranjem kvalitetnih odgovora
 - Drugo (Molim navedite)
 - Za sada nisam uočio/la nedostatke
11. Planirate li u svom budućem odgojno-obrazovnom radu koristiti ChatGPT za izradu testova provjere znanja?
 - Da, za izradu cjelovitih testova
 - Da, ali samo kao pomoć u analizi i ispravljanju testova koje sam/a izradim
 - Ne
 - Nisam siguran/na