

UNIVERZITET U BEOGRADU
FAKULTET ORGANIZACIONIH NAUKA

Jelica G. Stanojević

**METODOLOŠKI PRISTUP RAZVOJU
EDUKATIVNE PLATFORME ZASNOVANE
NA VELIKOM JEZIČKOM MODELU**

DOKTORSKA DISERTACIJA

Beograd, 2026

UNIVERSITY OF BELGRADE

FACULTY OF ORGANIZATIONAL SCIENCES

Jelica G. Stanojević

**METHODOLOGICAL APPROACH FOR
DEVELOPMENT OF EDUCATIONAL
PLATFORM BASED ON LARGE
LANGUAGE MODEL**

DISSERTATION PROPOSAL

Belgrade, 2026

Mentor:

Dr Miroslav Minović, redovni profesor
Fakultet organizacionih nauka, Univerzitet u Beogradu

Članovi komisije:

Dr Miloš Milovanović, redovni profesor
Fakultet organizacionih nauka, Univerzitet u Beogradu

Dr Ivana Kovačević, redovni profesor
Fakultet organizacionih nauka, Univerzitet u Beogradu

Dr Boško Nikolić, redovni profesor
Elektrotehnički fakultet, Univerzitet u Beogradu

Datum odbrane: _____

Metodološki pristup razvoju edukativne platforme zasnovane na velikom jezičkom modelu

Apstrakt:

Pojava ChatGPT aplikacije kroz koju su veliki jezički modeli postali dostupni široj javnosti označila je početak značajne ekspanzije u ovoj oblasti i uticala na razvoj drugih javno dostupnih četbot aplikacija zasnovanih na velikim jezičkim modelima poput Gemini, Claude i drugih. Njihova pojava izazvala je određene negativne posledice u obrazovnom sektoru, pre svega zbog porasta plagijarizma i smanjenja kritičkog razmišljanja kod studenata koji ih koriste za izvršavanje akademskih zadataka. Ovo je izazvalo zabrinutost i u pojedinim slučajevima reakciju obrazovnih institucija u vidu zabrane njihovog korišćenja. Međutim, vremenom postaje jasno da je potrebno prilagoditi određene obrazovne procese ovim tehnologijama i iskoristiti njihove benefite radi unapređenja istog. Potencijal za njihovu pozitivnu upotrebu uz odgovorno korišćenje je značajan. Konkretno, elektronsko učenje, koje je sve češće zastupljeno i kao dopuna tradicionalnom obrazovanju, ima potencijal za unapređenje kroz integraciju četbota koji je zasnovan na velikom jezičkom modelu. Jedan od glavnih aspekata poboljšanja se odnosi na pružanje prilagođene podrške studentima tokom procesa učenja kroz dostupnost u bilo koje doba i na bilo kom mestu. U ovoj doktorskoj disertaciji je istraživana potencijal primene generativne veštačke inteligencije odnosno velikih jezičkih modela u domenu obrazovanja kroz kreiranje i primenu objedinjenog modela za razvoj edukativne platforme koja integriše konverzionog agenta zasnovanog na velikom jezičkom modelu. Ovako definisan model kroz MDA pristup obuhvata važne koncepte koje je potrebno uzeti u obzir prilikom izrade ovih sistema i samim tim može se posmatrati i kao ontologija znanja potrebnih za njihov razvoj. Na osnovu ovog modela definisana su proširenja kroz UML profile za potrebe modelovanja ovakvog sistema. Definisan je predlog procesa razvoja ovakve edukativne platforme u čijim fazama je moguće koristiti definisana proširenja radi njihovog detaljnijeg opisa. Razvijen je prototip sistema čija je implementacija opisana, i izvršeno je testiranje i evaluacija samog sistema od strane studenata treće godine sa smera Informacioni sistemi i tehnologije, Fakulteta organizacionih nauka. Takođe su predstavljeni su rezultati istraživanja sprovedenog na Fakultetu organizacionih nauka o faktorima koji utiču na prihvatanje ChatGPT-ja kao alata u obrazovnom kontekstu od strane studenata. Rezultati dobijeni tokom ovog istraživanja mogu biti korisni nastavnicima i dizajnerima edukativnih platformi radi poboljšanja iskustva učenja ukoliko planiraju integraciju ovih alata u nastavni proces.

Ključne reči: veliki jezički model, konverzion agent, četbot, elektronsko učenje

Naučna oblast: Organizacione nauke

Uža naučna oblast: Informacione tehnologije

UDK:

Methodological approach for development of educational platform based on large language model

Abstract:

The appearance of the ChatGPT application, through which large language models (LLMs) became available to the general public, marked the beginning of a significant expansion in this area and the development of other publicly available chatbot applications based on LLMs such as Gemini, Claude and others. Their appearance caused certain negative consequences in the education sector, primarily due to the increase in plagiarism and the reduction of critical thinking among students who use them to complete academic tasks. This has raised concerns and in some cases prompted educational institutions to prohibit their use. However, over time it becomes clear that it is necessary to adapt certain educational processes to these technologies and use their benefits to improve them. The potential for their positive use, when used responsibly, is significant. In particular, e-learning, which is increasingly represented as a supplement to traditional education, has the potential for improvement through the integration of chatbots that are based on large language model. One of the main aspects of the improvement relates to the provision of adapted support to students during the learning process through availability at any time and in any place. In this doctoral dissertation, the potential of applying generative artificial intelligence, i.e. large language models, in the domain of education was investigated through the creation and application of a unified model for the development of an educational platform that integrates a conversational agent based on a large language model. The model defined in this way through the MDA approach includes important concepts that need to be taken into account when creating these systems and therefore can be viewed as an ontology of knowledge needed for their development. Based on this model, extensions are defined through UML profiles for the needs of modeling such a system. A proposal for the development process of such an educational platform has been defined, with the phases in which the defined extensions can be incorporated to provide a more detailed description. A system prototype was developed, implementation was described, and the system itself was tested and evaluated by third-year students from the Information Systems and Technologies department of the Faculty of Organizational Sciences. Also the results of research conducted at the Faculty of Organizational Sciences on the factors that influence the acceptance of ChatGPT as a tool in an educational context was presented. The results obtained during this research can be useful to teachers and designers of educational platforms in order to improve the learning experience if they plan to integrate these tools into the teaching process.

Keywords: large language model, conversational agent, chatbot, e-learning

Scientific area: Organizational sciences

Specific scientific area: Information technologies

UDC:

Zahvalnica

Disertaciju posvećujem svojoj porodici i zahvaljujem im se na безусловnoj ljubavi, strpljenju, razumevanju i podršci tokom svih godina školovanja. Posebnu zahvalnost dugujem Miljanu bez čije emotivne podrške rad sigurno ne bi bio uspešan.

Veliku zahvalnost upućujem mentoru prof. dr Miroslavu Minoviću za strpljivo vođenje i usmeravanje tokom višegodišnjeg rada i istraživanja. Takođe se zahvaljujem članovima komisije koji su svojim sugestijama i komentarima doprineli kvalitetu ove disertacije.

Na kraju, posebnu zahvalnost dugujem svim kolegama koji su na bilo koji način doprineli realizaciji istraživanja i izradi ove disertacije.

SADRŽAJ

1	UVOD	1
1.1	Problem i predmet istraživanja	1
1.2	Cilj istraživanja i polazne hipoteze.....	2
1.3	Struktura rada	3
2	VELIKI JEZIČKI MODELI	5
2.1	Generativna veštačka inteligencija.....	5
2.2	Istorijat jezičkih modela.....	6
2.2.1	Statistički jezički modeli.....	7
2.2.2	Neuronski jezički modeli	8
2.2.3	Pretrenirani jezički modeli	8
2.2.4	Veliki jezički modeli.....	8
2.3	Karakteristike velikih jezičkih modela	9
2.4	Arhitektura velikih jezičkih modela.....	10
2.4.1	Transformer arhitektura i mehanizam samopažnje.....	10
2.4.2	Generativni pretrenirani transformer model	11
2.5	Primena velikih jezičkih modela	12
2.6	ChatGPT.....	13
3	ELEKTRONSKO UČENJE.....	15
3.1	Pojam i definicija elektronskog učenja	15
3.2	Kategorije sistema elektronskog učenja	15
3.3	Prednosti elektronskog učenja	16
3.4	Izazovi elektronskog učenja	17
3.5	Hibridno učenje	18
4	EDUKATIVNI ČETBOTOVI.....	19
4.1	Definicija i pojam četbotova	19
4.2	Kratak istorijat četbotova.....	19
4.3	Četbotovi u edukaciji	20
4.3.1	Definicija i ciljevi edukativnih četbotova.....	20
4.3.2	Kategorije edukativnih četbotova.....	20
4.3.3	Četbotovi i elektronsko učenje.....	21
4.4	Prednosti primene četbotova u edukaciji	22
4.5	Izazovi primene četbotova u edukaciji	22
5	VELIKI JEZIČKI MODELI U OBRAZOVANJU	24
5.1	Generativna veštačka inteligencija u obrazovanju.....	24
5.2	Uticaj na proveru znanja	25
5.3	Nastavnički pomoćni alat	26
5.4	Studentski pomoćni alat	27
5.5	Digitalna pismenost o velikim jezičkim modelima.....	28
5.6	Prihvatanje tehnologije u obrazovanju	29
5.7	Primena u onlajn i hibridnom obrazovanju	37
5.8	Primena četbota zasnovanog na velikom jezičkom modelu u elektronskom učenju	39

6	OBJEDINJENI MODEL EDUKATIVNE PLATFORME SA KONVERZACIONIM AGENTOM ZASNOVANIM NA VELIKOM JEZIČKOM MODELU	45
6.1	Arhitektura vođena modelom	46
6.2	Metamodeli edukativne platforme sa VJM konverzacionim agentom	48
6.2.1	Metamodel edukativne platforme sa VJM konverzacionim agentom.....	48
6.2.2	Metamodel edukativnog sadržaja	48
6.2.3	Metamodel objekata za učenje	50
6.2.4	Metamodel VJM konverzacionog agenta.....	51
6.2.5	Metamodel interakcije sa edukativnom platformom	54
6.2.6	Metamodel evaluacije.....	55
7	MODELOVANJE EDUKATIVNE PLATFORME SA KONVERZACIONIM AGENTOM ZASNOVANIM NA VELIKOM JEZIČKOM MODELU.....	58
7.1	Modelovanje edukativnog sadržaja	58
7.2	Modelovanje konverzacionog agenta zasnovanog na velikom jezičkom modelu.....	61
7.3	Modelovanje putanje učenja	65
7.4	Modelovanje evaluacije	66
8	PREDLOG PROCESA RAZVOJA EDUKATIVNE PLATFORME SA KONVERZACIONIM AGENTOM ZASNOVANIM NA VELIKOM JEZIČKOM MODELU	69
8.1	Unified proces razvoja softvera	69
8.2	Predloženi proces razvoja.....	70
9	REALIZACIJA PROTOTIPA EDUKATIVNE PLATFORME SA KONVERZACIONIM AGENTOM ZASNOVANIM NA VELIKOM JEZIČKOM MODELU	72
9.1	Korisnički zahtevi	72
9.2	Arhitektura predloženog sistema	74
9.2.1	Klijentska aplikacija	75
9.2.2	Serverska aplikacija	78
9.2.3	Baza podataka	84
9.2.4	Vektorska baza podataka i RAG tehnika	85
10	EVALUACIJA	91
10.1	Evaluacija prvog prototipa.....	91
10.1.1	Metodologija istraživanja	91
10.1.2	Rezultati	94
10.1.3	Diskusija rezultata	101
10.2	Evaluacija drugog prototipa.....	104
10.2.1	Upotrebljivost.....	104
10.2.2	Metodologija istraživanja	107
10.2.3	Rezultati i analiza.....	109
10.2.4	Diskusija rezultata	114
10.2.5	Diskusija sa studentima.....	115
11	ZAKLJUČAK.....	118
12	REFERENCE.....	120

Slike

Slika 1 – Proširen TAM teorijski model	31
Slika 2 - Preslikavanje okvira za razvoj edukativne platforme koja integriše VJM konverzionog agenta na OMG MDA pristup	47
Slika 3 - Osnovni koncepti edukativne platforme sa VJM konverzionim agentom	48
Slika 4 - Metamodel edukativnog sadržaja.....	49
Slika 5 - Metamodel objekata za učenje [133]	50
Slika 6 - Metamodel taksonomije znanja [133]	51
Slika 7 – Metamodel VJM konverzionog agenta.....	52
Slika 8 - Metamodel odgovora na aktivnost.....	54
Slika 9 – Metamodel interakcije sa edukativnom platformom	54
Slika 10 – Metamodel evaluacije.....	57
Slika 11 - Primer upotrebe profila edukativnog sadržaja.....	61
Slika 12 – Primer upotrebe profila VJM Konverzionog Agent.....	65
Slika 13 - Primer upotrebe profila putanje učenja.....	66
Slika 14 – Primer upotrebe profila Evaluacija	68
Slika 15 – Predloženi proces razvoja edukativne platforme sa VJM konverzionim agentom	70
Slika 16 - Dijagram slučajeva korišćenja	72
Slika 17 - Opšta arhitektura predloženog sistema.....	74
Slika 18 - Korisnički interfejs sadržaja jedinica učenja.....	76
Slika 19 - Korisnički interfejs četbota.....	77
Slika 20 – Korisnički interfejs aktivnosti.....	78
Slika 21 - Opšta šema Firestore kolekcije Studenti	85
Slika 22 – Osnovni proces RAG tehnike	86
Slika 23 - Raspodela studentskih ocena o četbot podršci	98

Tabele

Tabela 1 - Karakteristike uzorka i prethodno iskustvo u korišćenju ChatGPT-ja	32
Tabela 2 - Deskriptivna statistika srednjih vrednosti konstrukata.....	33
Tabela 3 - Rezultati Mann-Whitney U testa za razlike u percepciji na osnovu pola i konteksta.....	33
Tabela 4 - Rezultati Kruskal-Wallis i Mann-Whitney testova za razlike u percepciji na osnovu perioda korišćenja	34
Tabela 5 - Analiza pouzdanosti i konvergentne validnosti	35
Tabela 6 - Analiza diskriminantne validnosti	35
Tabela 7 - Evaluacija strukturnog modela i odluka o hipotezama.....	36
Tabela 8 – Stereotipi profila edukativnog sadržaja sa pridodatim vrednostima	59
Tabela 9 – Definisana nabranja profila edukativnog sadržaja	60
Tabela 10 – Stereotipi profila VJM Konverzacioni Agent sa pridodatim vrednostima	62
Tabela 11 – Definisana nabranja profila VJM Konverzacioni Agent.....	64
Tabela 12 - Stereotip profila Evaluacija sa pridodatim vrednostima	66
Tabela 13 – Definisana nabranja profila evaluacije.....	67
Tabela 14 - Isecci konverzacija studenata i konverzacijonog agenta za lekciju.....	81
Tabela 15 - Isećak konverzacije studenta i konverzacijonog agenta za aktivnost.....	83
Tabela 16 – Primer pitanja, dodatnog konteksta i odgovora	90
Tabela 17 - Raspodela učesnika prema polu.....	92
Tabela 18 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnjevog U testa za ESPB i prosek u eksperimentalnoj i kontrolnoj grupi	93
Tabela 19 - Raspodela vrsta interakcija sa četbotom i primeri pitanja	94
Tabela 20 - Primeri interakcija između studenta i četbota po kategorijama pitanja	95
Tabela 21 - Evaluacija četbota od strane studenata iz eksperimentalne grupe.....	97
Tabela 22 - Analiza pouzdanosti između ocenjivača	99
Tabela 23 - Evaluacija četbot odgovora od strane eksperata	99
Tabela 24 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnjevog U testa za razlike u kvalitetu praktičnih i teorijskih odgovora prema 5 kriterijuma.....	100
Tabela 25 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnjevog U testa za poređenje broja poena ostvarenih na testu između eksperimentalne i kontrolne grupe	101
Tabela 26 - Lista tvrdnji koje čine Skalu upotrebljivosti sistema [204].....	105
Tabela 27 - Prilagođeni i korišćeni TAM konstrukti: percipirana korisnost, percipirana lakoća korišćenja, percipirano zadovoljstvo	106
Tabela 28 - Stavke evaluacije četbota.....	107
Tabela 29 - Raspodela učesnika prema polu.....	108
Tabela 30 - Rezultati Kolmogorov-Smirnovljevog testa, t-testa nezavisnih uzoraka i Man-Vitnjevog U testa za ESPB i prosek u eksperimentalnoj i kontrolnoj grupi	109
Tabela 31 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnjevog U testa za SUS	110
Tabela 32 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnjevog U testa za PK, PLK i PZ.....	110
Tabela 33 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnjevog U testa za trajanje interakcije i navigaciju	111

Tabela 34 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnijevog U testa za poene ostvarene na zadacima.....	113
Tabela 35 - Rezultati Kolmogorov-Smirnovljevog testa i t-testa nezavisnih uzoraka za poene ostvarene na testovima.....	113
Tabela 36 – Evaluacija četbota od strane studenata.....	114

1 UVOD

1.1 Problem i predmet istraživanja

Veliki jezički modeli predstavljaju vid generativne veštačke inteligencije koji omogućava generisanje tekstualnog sadržaja na osnovu ulaza (eng. *prompt*), uključujući odgovore na korisničke upite. Široj javnosti postali su dostupni krajem 2022. godine kroz četbot aplikaciju ChatGPT, što je označilo početak značajne ekspanzije u ovoj oblasti i pojavu drugih javno dostupnih četbot aplikacija zasnovanih na velikim jezičkim modelima poput Gemini, Claude i drugih. Predstavljaju ozbiljan pomak u oblasti obrade prirodnog jezika gde su pokazali izuzetno visoke performanse u zadacima kao što su programiranje, prevođenje, generisanje sadržaja i odgovaranje na pitanja. Korisnici kroz konverzaciju koju vode u okviru ovih alata mogu dobiti odgovore na pitanja ili zadavati određene instrukcije velikom jezičkom modelu.

Njihova pojava izazvala je i određene negativne posledice u obrazovnom sektoru, pre svega zbog porasta plagijarizma i smanjenja kritičkog razmišljanja kod studenata koji ih koriste za izvršavanje akademskih zadataka. Ovo je izazvalo zabrinutost i u pojedinim slučajevima reakciju obrazovnih institucija u vidu zabrane njihovog korišćenja. Međutim, vremenom postaje jasno da je potrebno prilagoditi određene obrazovne procese ovim tehnologijama i iskoristiti njihove benefite radi unapređenja obrazovnog procesa. Potencijal za njihovu pozitivnu upotrebu uz odgovorno korišćenje je veliki.

Konverzacioni agenti, takođe poznati kao četbotovi, predstavljaju sisteme veštačke inteligencije dizajnirane da simuliraju razgovore što omogućava pružanje interaktivne podrške studentima [1]. Jedan od glavnih izazova dosadašnjih implementacija četbotova u obrazovanju je upravo tehničke prirode odnosno njihovo programiranje i spremnost da imaju odgovor na svako studentsko pitanje [2]. Zahvaljujući najnovijim dostignućima u mašinskom učenju i obradi prirodnog jezika, ovi sistemi su u stanju da stupe u interakciju sa svojim korisnicima na različite načine koristeći prirodni jezik [1]. S obzirom na značajan napredak u oblasti generativne veštačke inteligencije i velikih jezičkih modela poslednjih godina, primena konverzacionih agenata odnosno četbotova u obrazovnom procesu dobila je priliku za razvoj. Međutim, uočena mana korišćenja ove tehnologije od strane studenata je ta što utiče negativno na njihovo kritičko razmišljanje [3], kreativnost i inovativnost [4]. Upravo je ovo razlog za preispitivanje obrazovnih praksi i dodatna istraživanja u oblasti primene velikih jezičkih modela u obrazovanju.

Pregledom literature u ovoj oblasti i detaljnim istraživanjem su ispitane mogućnosti primene velikih jezičkih modela u obrazovnom procesu, kao i mogućnosti unapređenja obrazovnog procesa kroz ovu integraciju. Nakon pregleda literature vezane za primenu velikih jezičkih modela u edukaciji, bilo je moguće zaključiti da postoji značajan porast publikacija iz ove oblasti prethodnih godina. Razlog tome je pojava ChatGPT četbota koji je znatno uticao na povećanje pažnje usmerene ka velikim jezičkim modelima i pospešio istraživanja u ovoj oblasti, a samim tim i u oblasti njihove primene u obrazovnom kontekstu. Studenti već uveliko koriste ovu tehnologiju za lakše dostizanje akademskih ciljeva. Prilikom analize ovih radova primećene su mogućnosti za unapređenje procesa obrazovanja primenom velikih jezičkih modela. Konkretno, elektronsko učenje, koje je sve češće zastupljeno i kao dopuna tradicionalnom obrazovanju, ima potencijal za

unapređenje kroz integraciju četbota koji je zasnovan na velikim jezičkim modelima. Jedan od glavnih aspekata poboljšanja se odnosi na pružanje prilagođene podrške studentima tokom procesa učenja kroz dostupnost u bilo koje doba i na bilo kom mestu.

Predmet istraživanja je mogućnost primene velikih jezičkih modela u obrazovanju, sa fokusom na visoko obrazovanje, kao i mogućnost realizacije rešenja koje integriše konverzionog agenta zasnovanog na velikom jezičkom modelu u edukativnu platformu kao pomoćnika pri učenju. Specifičnije, predmet istraživanja fokusira se na primenu savremenih informaciono-komunikacionih tehnologija u razvoju edukativne platforme koja integriše konverzionog agenta zasnovanog na velikom jezičkom modelu. U radu će biti predstavljen pregled relevantne oblasti i aktuelnih istraživanja, kao i razvoj tehničko-tehnološkog okvira za kreiranje edukativne platforme za elektronsko učenje integrisane sa četbotom zasnovanim na velikom jezičkom modelu, kao i evaluacija realizovanog sistema.

1.2 Cilj istraživanja i polazne hipoteze

Cilj ovog istraživanja predstavlja povećanje kvaliteta procesa obrazovanja integracijom konverzionog agenta zasnovanog na velikom jezičkom modelu u edukativnu platformu. Osnovni cilj istraživanja je izrada okvira primene konverzionog agenta zasnovanog na velikom jezičkom modelu u obrazovanju. Poseban cilj istraživanja je praktična realizacija prototipa edukativne platforme za učenje koji integriše konverzionog agenta zasnovanog na velikom jezičkom modelu i njegova evaluacija u realnim uslovima.

Kao glavna hipoteza koja će biti testirana u doktorskoj disertaciji se izdvaja:

H0: Moguće je razviti okvir za razvoj edukativne platforme koja integriše četbota zasnovanog na velikom jezičkom modelu.

Posebne hipoteze u istraživanju koje proističu iz glavne istraživačke hipoteze su:

H1: Novi metodološki pristup, koji definiše metode, modele, tehnologije i alate za realizaciju takvog sistema će unaprediti proces razvoja sistema za elektronsko učenje koji integrišu četbota zasnovanog na velikom jezičkom modelu.

H2: Edukativna platforma koja integriše četbota zasnovanog na velikom jezičkom modelu, razvijena primenom definisanog okvira, pokazuje viši stepen upotrebljivosti u odnosu na tradicionalne sisteme za elektronsko učenje.

H3: Proširenje standardne metodologije za razvoj informacionih sistema koja je bazirana na UML jeziku, uvođenjem novih koncepata specifičnih za primenu velikih jezičkih modela u obrazovanju, olakšava i unapređuje kvalitet realizacije okvira za primenu velikih jezičkih modela u okviru edukativne platforme.

1.3 Struktura rada

Doktorska disertacija sadrži 12 poglavlja. U uvodnom poglavlju se definiše problem, predmet i cilj istraživanja, kao i opšta i posebne hipoteze. Nakon toga je navedena struktura rada.

Naredna četiri poglavlja pružaju uvod u i pregled oblasti čija primena je tema ove doktorske disertacije. U poglavlju 2 je uveden koncept generativne veštačke inteligencije i velikih jezičkih modela kao jednog vida generativne veštačke inteligencije. Zatim su spomenuti jezički modeli i dat je kratak istorijski pregled jezičkog modelovanja i tehnika koje su prethodile velikim jezičkim modelima. Nakon toga su opisane osnovne karakteristike velikih jezičkih modela i *Generative Pretrained Transformer* (GPT) arhitektura koja stoji iza jedne od najpopularnijih primena velikih jezičkih modela u vidu veb aplikacije – ChatGPT. Takođe se dotičemo mogućih oblasti primena, pored obrazovanja koje je fokus ove doktorske disertacije i o kome će biti više reči u narednim poglavljima. Na kraju poglavlja opisujemo ChatGPT kao prvu aplikaciju kroz koju je pristup velikim jezičkim modelima omogućen za široku publiku i izazove koje je kao disruptivna tehnologija doneo u proces obrazovanja. Poglavlje 3 definiše koncept elektronskog učenja, uvodi kategorizaciju sistema elektronskog učenja, kao i prednosti i izazove njegove primene. Takođe se spominje koncept hibridnog učenja čija kombinacija elektronskog učenja sa tradicionalnim učenjem u učionicama pokušava da prevaziđe izazove u ovim modalitetima i preporučuje se obrazovnim institucijama. U narednom, četvrtom poglavlju uvodimo koncept edukativnih četbotova čija implementacija upotrebom velikih jezičkih modela i primena u okviru sistema za elektronsko učenje predstavlja temu ove disertacije. Početak ovog poglavlja obuhvata definiciju i pojam četbota, zatim se navodi kratak pregled njihovog razvoja do napretka velikih jezičkih modela, njihovu primenu u obrazovnom kontekstu, prednosti i izazove te primene, kao i vezu sa elektronskim učenjem. Peto poglavlje navodi mogućnosti primene četbot alata zasnovanih na velikim jezičkim modelima u obrazovanju sa aspekta nastavnika i studenata i njihov uticaj na proveru znanja. Takođe se kao važan aspekt primene ove tehnologije navodi digitalna pismenost u ovoj oblasti koja je neophodna radi korišćenja datih alata na pravi način. U ovom poglavlju su takođe navedeni rezultati početnog istraživanja ove doktorske disertacije o faktorima koji utiču na prihvatanje ChatGPT-ja kao alata za pomoć pri učenju kroz proširen model prihvatanja tehnologije (eng. *Technology Acceptance Model*). Nakon toga, spominju se mogućnosti primene ovih alata u onlajn i hibridnom tipu obrazovanja. Poslednje potpoglavlje u okviru petog poglavlja, opisuje mogućnosti integracija ovih modela u edukativne platforme kroz aplikacioni programski interfejs i daje pregled aktuelnih istraživanja u ovom domenu.

U poglavlju 6 dat je predlog objedinjenog modela za razvoj edukativne platforme koja integriše konverzionog agenta zasnovanog na velikom jezičkom modelu (VJM edukativna platforma). Za razvoj objedinjenog modela korišćena je arhitektura vođena modelima (eng. *Model Driven Architecture* – MDA) čijom primenom su definisani i opisani sledeći metamodeli u okviru ovog poglavlja: metamodel edukativne platforme sa konverzionim agentom, metamodel edukativnog sadržaja, metamodel objekata za učenje, metamodel taksonomije znanja, metamodel VJM konverzionog agenta, metamodel odgovora studenta, metamodel interakcije sa edukativnom platformom i metamodel evaluacije.

Poglavlje 7 prikazuje modelovanje VJM edukativne platforme na osnovu definisanog modela u prethodnom poglavlju, kroz proširenja definisana UML profilima. Stereotipi, koji predstavljaju osnovni gradivni element UML profila, i njihove pridodate vrednosti sa opisima, su definisani za sledeće aspekte sistema: edukativni sadržaj, konverzacioni agent zasnovan na velikom jezičkom modelu, putanja učenja i evaluacija.

U poglavlju 8 dat je predlog procesa razvoja edukativne platforme koja integriše konverzacionog agenta zasnovanog na velikom jezičkom modelu kroz proširenje faza odnosno radnih tokova u okviru Unified procesa. Unified proces predstavlja metodologiju zasnovanu na iterativno-inkrementalnom pristupu razvoja softvera. Radni tokovi u okviru ove metodologije su prošireni aktivnostima relevantnim za razvoj sistema za elektronsko učenje sa VJM konverzacionim agentom.

U poglavlju 9 je opisana realizacija prototipa VJM edukativne platforme. Na početku će biti navedeni slučajevi korišćenja koji su implementirani radi evaluacije datog sistema. Zatim će biti dat prikaz moguće arhitekture predloženog sistema i opis njenih delova. Nakon toga će biti opisani korišćeni alati i tehnologije za implementaciju spomenutih delova prototipa sistema, kao i diskusija o okvirima i tehnologijama koji nisu korišćeni u konkretnom primeru, ali predstavljaju moguće alternative.

Poglavlje 10 sadrži rezultate evaluacije realizovanih prototipa sistema za elektronsko učenje koji integrišu konverzacionog agenta zasnovanog na velikom jezičkom modelu. Opisana su dva sprovedena eksperimenta u okviru kojih su studenti imali priliku da koriste dve verzije prototipa edukativne platforme nastale kao jedan od rezultata ove doktorske disertacije. Prvi prototip testira izvodljivost primene implementirane platforme i percipiranu korisnost konverzacionog agenta zasnovanim na VJM-u kao podrške studentima u procesu učenja na predmetu Računarske mreže i telekomunikacije, dok se evaluacija drugog prototipa fokusira na ispitivanje upotrebljivosti ovog sistema za elektronsko učenje.

Poglavlje 11 sadrži zaključak sa najznačajnijim doprinosima teze i mogućim pravcima daljeg istraživanja. Poslednje poglavlje sadrži listu referenci korišćenih u ovom radu formatiranih prema IEEE standardu.

2 VELIKI JEZIČKI MODELI

2.1 Generativna veštačka inteligencija

Oblast veštačke inteligencije se bavi kreiranjem inteligentnih mašina koje mogu da odrede kako da efektivno i sigurno deluju u novonastalim situacijama [5]. Veštačka inteligencija se fokusira na kreiranje sistema koji mogu rezonovati, učiti i samostalno delovati [6].

Mašinsko učenje predstavlja podoblast u okviru veštačke inteligencije koja se bavi treniranjem modela pomoću ulaznih podataka kako bi se izvele smislene prognoze na nepoznatim podacima [6]. Algoritmi mašinskog učenja analiziraju ogromne količine podataka radi identifikacije šablona, korelacija i dobijanja uvida kako bi mogli da vrše predviđanja, klasifikaciju i davanje preporuka [7]. Tri glavna oblika mašinskog učenja na osnovu prirode podataka i načina na koji modeli uče predstavljaju nadgledano učenje (eng. *supervised learning*), nenadgledano učenje (eng. *unsupervised learning*) i učenje uz podsticaj (eng. *reinforcement learning*). Nadgledano učenje koristi označene podatke za treniranje modela, dok se nenadgledano učenje oslanja na neoznačene podatke za otkrivanje obrazaca. Učenje uz podsticaj predstavlja učenje kod kog se kroz interakciju sa okruženjem dobija povratna informacija u vidu nagrade ili kazne, a njegov cilj predstavlja maksimizaciju nagrada [5].

Takođe, pored spomenutih oblika mašinskog učenja, treba spomenuti i duboko učenje (eng. *deep learning*), kao skup tehnika mašinskog učenja, koje je trenutno najčešće korišćeni pristup za prepoznavanje vizuelnih objekata i govora, mašinsko prevođenje i sintezu govora i slika [5]. Duboko učenje koristi veštačke neuronske mreže za rad sa složenim obrascima koje konvencionalne metode mašinskog učenja ne mogu da obrade [6]. Polunadgledano učenje (eng. *semi-supervised*), predstavlja kategoriju dubokog učenja koja kombinuje glavne oblike mašinskog učenja tj. nadgledano i nenadgledano učenje. Kod ovog oblika učenja manji deo označenih podataka se kombinuje sa velikim skupom neoznačenih podataka [6].

Mittal i saradnici [6] navode glavne karakteristike modela dubokog učenja koji mogu biti diskriminatorni i generativni. Diskriminatorni modeli su razvijeni uglavnom na označenim podacima i mogu da prepoznaju odnos između karakteristika podataka i njihovih oznaka. Koriste se za klasifikaciju podataka. S druge strane, generativni modeli koji su tema ovog rada, su zaduženi za generisanje novog sadržaja, što i sam naziv implicira, na osnovu distribucije verovatnoće postojećih podataka.

Generativna veštačka inteligencija predstavlja specijalizaciju u okviru dubokog učenja, koja koristi veštačke neuronske mreže za obradu označenih i neoznačenih podataka, kombinovanjem nadgledanih, nenadgledanih i polunadgledanih tehnika mašinskog učenja [6]. Za razliku od tradicionalnih metoda veštačke inteligencije i mašinskog učenja, koje se fokusiraju na identifikovanje obrazaca u podacima i na razumevanje podataka, generativna veštačka inteligencija naglašava generisanje novih podataka koji nisu direktno izvedeni iz postojećih setova podataka ili materijala korišćenim prilikom treniranja modela [6]. Dakle, njena glavna karakteristika je generisanje novog sadržaja

na osnovu učenih obrazaca u trening podacima kroz napredne tehnike dubokog učenja.

Sistemi generativne veštačke inteligencije se mogu kategorisati na osnovu vrste sadržaja rezultata koje pružaju. Tip sadržaja koji može biti izlaz iz ovih sistema može biti tekst, slika, zvuk i video snimak [8] kao i 3D vizuelni prikaz [6]. Mittal i saradnici [6] su prema ovom kriterijumu identifikovali neke od najpoznatijih modela:

- GPT (OpenAI), Llama (Meta), PaLM (Google), Claude (Anthropic), ChinchillaAI (DeepMind) modeli za generisanje teksta;
- Codex (OpenAI), StarCoder (HuggingFace i ServiceNow), CoPilot (GitHub, Microsoft i OpenAI), Code Interpreter (OpenAI), CodeT5 (SalesForce) za generisanje koda;
- Midjourney, DALL-E (OpenAI), Imagen (Google), GauGAN (Nvidia), Craiyon, Stable Diffusion (Stability AI) za generisanje slika;
- Murf.ai (OpenAI), MuseNet (OpenAI), MusicLM (Google), AudioPaLM (Google), Voicebox (Meta) za generisanje zvuka;
- Imagen Video (Google), Phenaki (Google), Make-A-Video (Meta), Krikey.ai, i Synthesia Studio za generisanje video snimaka;
- Point-E (OpenAI), DreamFusion (Google i UC Berkeley), GAUDI (Apple), Magic3D (Nvidia), Rodin Diffusion (Microsoft) za generisanje 3D vizuelnih prikaza.

Takođe, platforme poput Houdini, GameSynth, Unity Machine Learning Agents, Nvidia GameWorks i PolyGen pružaju mogućnost generisanja likova, *non-playable* karaktera (NPC), objekata, okruženja, audio-vizuelnih efekata, zvukova i muzike u igrama [6].

Veliki jezički modeli, čija primena u edukaciji je tema ovog rada, pripadaju ovoj oblasti dubokog učenja. Predstavljaju vrstu generativne veštačke inteligencije koja je specifično dizajnirana da generiše jezik sličan ljudskom kao odgovor na zadat upit [8]. Ostvarili su značajan napredak poslednjih godina delom zbog napretka algoritama u oblasti dubokog učenja, ali i zbog povećanja dostupnosti masivnih kompjuterskih resursa, kao i ogromnih količina podataka za treniranje modela [8].

Fokus ovog rada su generativni sistemi veštačke inteligencije sa izlazom tekstualnog tipa s obzirom na to da je njihova upotreba u edukaciji od strane studenata široko rasprostranjena. Generisan tekst koji studenti dobijaju od strane velikih jezičkih modela mogu biti odgovori na njihova pitanja postavljena u procesu učenja, objašnjenja kompleksnih pojmova, postupci rešavanja domaćih zadataka i sl. Sve u svemu mogu se ponašati kao studentski personalni pomoćnici tokom njihovog školovanja.

2.2 Istorijat jezičkih modela

Prirodna obrada jezika (eng. *Natural Language Processing* – NLP) je grana veštačke inteligencije koja ima za cilj da omogući računarima da nauče, razumeju i generišu ljudski odnosno prirodni jezik [9]. Ova oblast obuhvata širok spektar zadataka, uključujući obradu teksta, prevođenje jezika, rezimiranje i generisanje teksta, izdvajanje informacija, odgovaranje na pitanja i sl. [10]. Cilj prirodne obrade jezika su sistemi koji

mogu uspešno da komuniciraju sa ljudima, čime se olakšava interakcija između ljudi i tehnologije.

Modelovanje jezika igra ključnu ulogu u unapređivanju jezičke inteligencije mašina [8]. Jezički modeli su ključni deo NLP-a i predstavljaju modele koji omogućavaju računarima da razumeju i generišu tekst svojstven ljudima [11]. U suštini, modelovanje jezika obuhvata izračunavanje verovatnoća sekvenci reči kako bi se predvidela verovatnoća sledećih reči u nizu ili generisao novi tekst na osnovu zadatog ulaza [8], [11].

Veliki jezički modeli u oblasti prirodne obrade jezika targetiraju različite zadatke. Jedan generativni model može obavljati više različitih zadataka sa odličnim preformansama kao što su analiza sentimenata, ekstrakcija informacija, rezimiranje teksta, mašinsko prevođenje i matematičko rezonovanje [12].

Razvoj obrade prirodnog jezika prošao je kroz nekoliko ključnih etapa, počevši od statističkog modelovanja, preko modela baziranih na neuronskim mrežama i pretreniranih modela, sve do savremenih velikih jezičkih modela [13]. Dakle, četiri ključne faze u evoluciji jezičkog modelovanja koje će biti opisane u narednim poglavljima su [8]:

- Statistički jezički modeli,
- Neuronski jezički modeli,
- Pretrenirani jezički modeli,
- Veliki jezički modeli.

2.2.1 Statistički jezički modeli

Prvi jezički modeli razvijeni su 1950-ih i 1960-ih godina. Ovi modeli su bili zasnovani na pravilima i oslanjali su se na ručno kreirane lingvističke pravilnike za obradu jezika [8]. Međutim, bili su ograničeni u svojim sposobnostima i nisu mogli da se izbore sa složenim aspektima obrade prirodnog jezika [8]. Tokom 1980-ih i 1990-ih razvijeni su statistički jezički modeli na osnovu metoda statističkog učenja [14], [15]. Oni predstavljaju prvi značajan korak u razvoju jezičkog modelovanja [8].

Statistički jezički modeli (eng. *Statistical Language Model* - SLM) posmatraju tekst kao niz reči pri čemu se verovatnoća teksta računa kao proizvod verovatnoća pojedinačnih reči [14]. SLM-ovi koriste probabilističke metode za procenu verovatnoće sekvence reči u datom kontekstu [8]. Dominantan oblik SLM-ova su Markovljevi lanci, poznatiji kao n-gram modeli, koji procenjuju verovatnoću reči na osnovu učestalosti n-1 prethodnih reči [14]. Markovljeva pretpostavka, na kojoj su bazirani ovi modeli, podrazumeva da je predikcija sledeće reči zasnovana na najskorijem kontekstu [15]. Na primer, bigram model procenjuje verovatnoću sledeće reči analizirajući učestalost parova reči [8]. SLM-ovi sa fiksnom dužinom konteksta n, nazivaju se n-gram jezičkim modelima (npr. bigram, trigram) [15].

Statistički jezički modeli našli su primenu u poboljšavanju performansi zadataka u oblasti pronalaženja informacija i obradi prirodnog jezika [15]. U poređenju sa modelima zasnovanim na pravilima, ovi pristupi omogućili su obradu većih količina

podataka i postizanje veće preciznosti, ali su i dalje pokazivali ograničenja u razumevanju složenijih aspekata semantike i konteksta jezika [8].

2.2.2 Neuronski jezički modeli

Hadi i saradnici [8] opisuju sledeću značajnu fazu u razvoju jezičkog modelovanja koja se desila sredinom 2010-ih uvođenjem jezičkih modela zasnovanim na neuronskim mrežama, poznatijim kao neuronski jezički modeli (eng. *Neural Language Model* – NLM). Ovaj pristup koristi neuronske mreže za predikciju distribucije verovatnoće sledeće reči u nizu uzimajući u obzir prethodne reči u sekvenci. Rekurentne neuronske mreže (eng. *Recurrent Neural Network* - RNN) i njihove varijacije često su primenjivane u ovom pristupu. Korišćenjem tehnika dubokog učenja, ovi modeli omogućili su učenje složenih obrazaca i jezičkih struktura iz velikih skupova tekstualnih podataka. Prvi neuronski jezički model, zasnovan na rekurentnoj neuronskoj mreži (RNNLM), razvijen je 2010. godine. Ovaj model pokazao se sposobnim da efikasno modeluje kontekst reči, što je dovelo do generisanja prirodnijeg teksta u odnosu na ranije modele.

2.2.3 Pretrenirani jezički modeli

Treća faza razvoja jezičkih modela koji su dobili naziv pretrenirani jezički modeli (eng. *Pretrained Language Model* – PLM), obeležena je uvođenjem kontekstualizovanih ugradnji reči (eng. *word embedding*), koje omogućavaju razumevanje značenja i konteksta reči u rečenici ili tekstu. Ovi modeli, kao što su npr, ELMo i BERT, koriste neuronske mreže za učenje vektorske reprezentacije reči koja uzima u obzir kontekst u kom se reč pojavljuje [8].

Autori Minaee i saradnici [14] su opisali paradigmu pretreniranja i podešavanja (eng. *fine-tuning*) u okviru ovih modela. Proces obuke i zaključivanja kod ovih modela prati spomenutu paradigmu. Ona označava da se jezički modeli bazirani na rekurentnim neuronskim mrežama ili transformer arhitekturi, prvo pretreniraju na velikim neoznačenim tekstualnim korpusima sa interneta za opšte zadatke poput predikcije reči. Nakon toga, modeli se prilagođavaju specifičnim zadacima koristeći male, označene skupove podataka koji su relevantni za konkretne primene. PLM-ovi, za razliku od prethodnih NLM-ova, nisu ograničeni na rešavanje specifičnih zadataka.

Primena ove paradigme jezičkog modelovanja je vrlo pozitivno uticala na performanse NLP zadataka [15]. Podstakla je veliki broj studija koje su uvele različite PLM arhitekture ili su se fokusirale na poboljšanje procesa pretreniranja.

2.2.4 Veliki jezički modeli

Veliki jezički modeli (eng. *Large Language Models* – LLM) su napravili značajan napredak u oblasti prirodne obrade jezika poslednjih godina. Predstavljaju četvrtu fazu u razvoju jezičkih modela koja uključuje stvaranje pretreniranih velikih jezičkih modela. Veliki jezički modeli su napredni jezički modeli sa ogromnom količinom parametara i impresivnim mogućnostima učenja [11]. Ovi modeli postižu izvanredne performanse u različitim zadacima obrade prirodnog jezika [8].

Istraživanja pokazuju da skaliranje pretreniranih jezičkih modela, na primer povećavanjem veličine modela ili količine podataka na kojima su trenirani, često dovodi do poboljšanja njihove sposobnosti za zadatke koji koriste njihovo prethodno stečeno znanje (eng. *downstream task*) [15]. Veliki jezički modeli, poput GPT-4 i LLaMA, predstavljaju neuronske jezičke modele zasnovane na transformer arhitekturi, koji sadrže desetine do stotine milijardi parametara i pretrenirani su na ogromnim količinama tekstualnih podataka [14]. U poređenju sa PLM-ovima, LLM-ovi nisu samo mnogo veći u veličini modela, već pokazuju i superiorne sposobnosti razumevanja jezika i generisanja teksta, kao i emergentne sposobnosti koje nisu prisutne kod manjih modela [14]. Na primer, GPT-3 može rešavati zadatke sa malo primera kroz učenje iz konteksta, dok GPT-2 u tome nije uspešan [15]. Zbog ovih karakteristika je istraživačka zajednica uvela termin „veliki jezički modeli“ kako bi označila velike pretrenirane jezičke modele, koji privlače sve veću pažnju istraživača [15].

2.3 Karakteristike velikih jezičkih modela

Autori Minaee i saradnici [14] u svom radu opisuju sledeće emergentne sposobnosti velikih jezičkih modela: učenje iz konteksta; praćenje instrukcija i višestepeno zaključivanje. Učenje iz konteksta podrazumeva da LLM-ovi uče novi zadatak iz malog skupa primera predstavljenih u ulazu tokom procesa zaključivanja. Praćenje instrukcija označava da nakon podešavanja instrukcija, LLM-ovi mogu da slede uputstva za nove vrste zadataka bez potrebe za eksplicitnim primerima. Višestepeno zaključivanje omogućava LLM-ovima da rešavaju složene zadatke razlažući ih na međukorake u procesu zaključivanja.

Zahvaljujući visokim performansama i sposobnosti generalizacije pri rešavanju različitih zadataka, a ne samo onih specifičnih, veliki jezički modeli predstavljaju obećavajući korak ka razvoju opšte veštačke inteligencije (eng. Artificial General Intelligence - AGI) u savremenoj eri [11]. S druge strane, veliki jezički modeli se mogu dodatno prilagoditi i za specifične zadatke, kao što su prevođenje između jezika ili odgovaranje na pitanja [8].

Kao što je već rečeno, jedna od glavnih funkcionalnosti velikih jezičkih modela ogleda se u njihovoj sposobnosti učenja iz konteksta. Modeli se treniraju da generišu tekst na osnovu datog konteksta ili upita, čime postižu koherentnost i kontekstualnu relevantnost odgovora. Ovo ih čini posebno pogodnim za interaktivne i konverzacione aplikacije [11].

Postoji nekoliko mogućih načina interakcije sa velikim jezičkim modelima [11] kroz korisnički interfejs. Uobičajeni pristup interakcije je tzv. prompt inženjering, pri čemu korisnici osmišljavaju i pružaju specifične tekstualne upite (promptove) kako bi usmerili model ka generisanju željenih odgovora ili obavljanju određenih zadataka. Alternativno, korisnici mogu voditi interakcije u formi pitanja i odgovora, postavljajući pitanja modelu i dobijajući odgovore. Takođe se mogu upustiti u dijalog, koristeći prirodni jezik za vođenje razgovora sa LLM-om.

Na osnovu analize evaluacija LLM-ova, Chang i saradnici [11] su sublimirali zadatke u kojima su pokazali dobre performanse. Postižu impresivne rezultate u zadacima koji

uključuju razumevanje jezika, kao što su analiza sentimenata, klasifikacija teksta i obrada činjeničnih podataka. Takođe, LLM-ovi demonstriraju robusne aritmetičke i logičke sposobnosti. Pored toga, složeniji zadaci poput matematičkog rasuđivanja i zaključivanja na osnovu strukturiranih podataka postali su ključni pokazatelji njihove efikasnosti. Zadovoljavajuće rezultate postižu i u zadacima obrade prirodnog jezika, uključujući mašinsko prevođenje jezika i odgovaranje na pitanja. Pokazuju izuzetnu sposobnost generisanja teksta, proizvodeći tačne i precizne jezičke izraze. LLM-ovi takođe pokazuju snažno razumevanje konteksta što im omogućava generisanje koherentnih odgovora, usklađenih sa ulaznim podacima.

U radu [11] su navedeni i izazovi sa kojim se suočavaju veliki jezički modeli u određenim oblastima. U domenu prirodnog jezičkog zaključivanja imaju slabije performanse kada treba da pokažu neslaganje sa ljudima. Imaju ograničenu sposobnost prepoznavanja semantičke sličnosti između događaja. Njihove sposobnosti apstraktnog rasuđivanja su ograničene, a skloni su konfuziji ili greškama u složenim kontekstima. Takođe, generativni LLM-ovi često ne ispunjavaju očekivane standarde na različitim jezicima. U kontekstu jezika koji koriste nelatinska pisma i imaju ograničene resurse za trening, LLM-ovi pokazuju suboptimalne performanse. Mogu širiti štetne sadržaje iz svojih trening podataka, uključujući neprimeren tekst koji može biti uvredljivog sadržaja. LLM-ovi mogu iskazivati društvene predrasude i toksičnost što kao rezultat daje pristrasne odgovore. Takođe, mogu posedovati manjak kredibiliteta za određene teme, što može dovesti do generisanja fabrikovanih informacija ili netačnih činjenica tzv. halucinacija modela. Njihova sposobnost inkorporiranja realnih ili promenljivih informacija je ograničena, što ih čini manje pogodnim za zadatke koji zahtevaju ažurirano znanje ili brzu adaptaciju na promene. Konačno, LLM-ovi su osetljivi na zlonamerne upite, što zahteva evaluacije i algoritme za poboljšanje njihove robusnosti.

2.4 Arhitektura velikih jezičkih modela

2.4.1 Transformer arhitektura i mehanizam samopažnje

Ključne tačke u razvoju velikih jezičkih modela su uvođenje transformer arhitekture i mehanizma samopažnje (eng. *self-attention mechanism*) [16], [17].

Transformer arhitektura donela je revoluciju u oblasti modelovanja prirodnog jezika omogućavajući efikasnu obradu sekvencijalnih podataka, paralelizaciju obrade teksta i uočavanje dugih zavisnosti u tekstu [11]. Transformer arhitektura predstavlja vrstu neuronske mreže pogodne za zadatke obrade prirodnog jezika [8]. Zahvaljujući njoj, značajno su unapređene sposobnosti jezičkih modela za obradu dugačkih zavisnosti u tekstovima prirodnog jezika [17]. Ova arhitektura, uvedena 2017. godine, je omogućila paralelno treniranje na više grafičkih procesorskih jedinica (eng. *Graphical Processing Unit* - GPU), što je uticalo na mogućnost kreiranja znatno većih modela [8].

Transformer arhitektura koristi slojeve samopažnje za učenje odnosa između reči koje su na velikim udaljenostima u tekstu [8]. Upravo mehanizam samopažnje predstavlja ključni element koji stoji iza mnogih velikih jezičkih modela, kao što su GPT-3, InstructGPT i GPT-4 [11].

Mehanizam samopažnje (eng. *self-attention*) omogućava modelu da proceni značaj svake reči u ulaznoj sekvenci, nezavisno od njene pozicije [8]. Dakle, ovaj mehanizam služi za određivanje relevantnosti različitih delova ulaznog teksta prilikom generisanja predikcija [17]. Ova tehnika omogućava modelu da bolje razume odnose između reči u rečenici, nezavisno od njihove pozicije [17]. Od suštinske važnosti je za zadatke obrade prirodnog jezika, poput prevođenja i odgovaranja na pitanja, gde značenje rečenice često zavisi od reči koje su udaljene jedna od druge [8].

Pored transformer arhitekture i mehanizma samopažnje, pretreniranje modela pre finog podešavanja za određeni zadatak predstavlja još jedan ključni faktor uspešnosti velikih jezičkih modela. Pretreniranje modela na velikim skupovima podataka pre njihovog finog podešavanja za konkretne zadatke pokazalo se kao izuzetno efikasna tehnika za poboljšanje performansi različitih jezičkih zadataka [17].

Podaci na kojima se model pretrenira obuhvataju ogromne količine teksta iz različitih izvora kao što su knjige, članci, veb sajtovi. U radu [8] je opisan proces pretprocesiranja podataka pre treniranja samog modela. Pretprocesiranje uključuje uklanjanje potencijalno štetnih, nepotrebnih, neželjenih i višak podataka, podataka lošeg kvaliteta, duplikata, kao i ličnih podataka. Kraj tog procesa podrazumeva tokenizaciju, odnosno segmentaciju teksta na delove tzv. tokene koji će biti deo LLM-a. Kroz nenadgledano učenje model uči da predviđa sledeću reč u rečenici na osnovu konteksta prethodnih reči.

Nakon pretreniranja, LLM može biti fino podešavan kroz treniranje modela za određeni zadatak kao što je prevođenje, odgovaranje na pitanja, generisanje teksta ili za određeni domen [8]. Model se priprema kroz obeležene primere za generisanje tačnijih odgovora i odgovora koji odgovaraju kontekstu za traženi zadatak [18]. Takođe, jedan od ključnih aspekata ovih modela je pojačano učenje kroz ljudski fidbek. Ova tehnika uključuje fino podešavanje modela korišćenjem odgovora generisanih od strane ljudi. Na taj način, model uči iz svojih grešaka i unapređuje performanse tokom vremena [11].

Obradom i analizom velike količine teksta, veliki jezički modeli stiču sveobuhvatno razumevanje jezika i sposobnost generisanja koherentnih i kontekstualno relevantnih odgovora. Tokom faze zaključivanja, prilikom interakcije sa LLM-om, korisnik unosi upit tj. prompt koji model obrađuje i generiše odgovor na osnovu naučenog znanja i konteksta. Ovaj odgovor se generiše korišćenjem probabilističkih metoda koje uzimaju u obzir verovatnoću različitih reči ili fraza u datom kontekstu [8].

2.4.2 Generativni pretrenirani transformer model

Objavljivanjem prve verzije generativnog pretreniranog transformer (eng. *Generative Pretrained Transformer* – GPT) OpenAI modela - GPT-1 2018. godine ostvaren je značajan napredak u obradi prirodnog jezika. Ovaj model sa 117 miliona parametara, je mogao da generiše kontekstualno relevantne rečenice, čime je demonstriran revolucionarni potencijal transformer arhitekture u NLP zadacima. Model je treniran kroz dva koraka koji uključuju nenadgledano pretreniranje i nadgledano prilagođavanje. Ovaj pristup je izazvao značajan interes u akademskoj i istraživačkoj zajednici. Iako je imao određena ograničenja, GPT-1 je postavio temelje za razvoj moćnijih modela i pokrenuo je novu eru istraživanja u oblasti veštačke inteligencije [8].

GPT-3 model je predstavljen 2020. godine od strane OpenAI organizacije. Predstavljao je najveći jezički model u tom trenutku koji je treniran na ogromnoj količini tekstualnih podataka. Ovaj model je sposoban da generiše visoko koherentan i prirodan tekst, čime je pokazao ogroman potencijal LLM-ova za rešavanje širokog spektra NLP zadataka [8]. Na osnovu njega nastao je InstructGPT model na kome je zasnovan GPT-3.5 model koji je prvi korišćen model u ChatGPT aplikaciji.

Za kreiranje prvobitnog modela korišćenom u ChatGPT aplikaciji, u fazi pretreniranja je izvršena integracija i koda i teksta. Tokom faze finog podešavanja, primenjene su tehnike podešavanja instrukcija i učenja uz podsticaj zasnovanog na ljudskom fidbeku kako bi se poboljšale performanse modela. Ove tehnike omogućile su bolju generalizaciju za različite zadatke i usklađivanje odgovora sa povratnim informacijama [19]. Vremenom je korisnicima postao dostupan i napredniji GPT-4 i GPT-5 model za korišćenje kroz ChatGPT platformu.

Hadi i saradnici [8] opisuju šest ključnih koraka u okviru arhitekture GPT 3.5 modela :

- 1) Kodiranje ulaza: Ulaz u GPT model predstavlja niz tokena, koji predstavljaju podatke. Svaki token se konvertuje u visokodimenzionalnu vektorsku reprezentaciju kroz sloj ugradnje (eng. *embedding layer*).
- 2) Transformer enkoder: Model se sastoji od više slojeva transformer enkodera. Svaki sloj uključuje mehanizam samopažnje i *feed-forward* neuronske mreže. Mehanizam samopažnje omogućava modelu da identifikuje zavisnosti među rečima u ulaznoj sekvenci, dok *feed-forward* mreže procesuiraju i transformišu reprezentacije podataka.
- 3) Ugradnje konteksta: Tokom prolaska ulazne sekvence kroz slojeve transformer enkodera, reprezentacija svakog tokena se ažurira u odnosu na kontekst. To znači da je reprezentacija svakog tokena pod uticajem okolnih tokena i njihovih kontekstualnih informacija.
- 4) Dekodiranje i generisanje jezika: Nakon što se ulazna sekvenca obradi, model generiše novi tekst predviđanjem verovatnoće narednog tokena u zavisnosti od konteksta. Ovaj proces se obično vrši uz upotrebu *softmax* funkcije aktivacije za određivanje verovatnoće svih mogućih tokena iz vokabulara.
- 5) Trening sa maskiranim modelovanjem jezika: Tokom faze pretreniranja, GPT modeli često primenjuju tehniku maskiranog modelovanja jezika. Ova tehnika podrazumeva nasumično maskiranje određenih tokena u ulaznoj sekvenci i obučavanje modela da predvidi te maskirane tokene na osnovu konteksta, čime se poboljšava razumevanje i generisanje koherentnog teksta.
- 6) Izlaz modela: Finalni sloj modela generiše izlaz odnosno završni rezultat u obliku teksta.

2.5 Primena velikih jezičkih modela

Veliki jezički modeli pored obrazovanja koje je tema ovog rada, pronalaze primenu i u drugim oblastima kao što su medicina, finansije i inženjerstvo [8]. Neke od mogućih primena će biti navedene u nastavku. U medicini mogu biti od pomoći i imaju potencijal

za primenu u dijagnostici, donošenju medicinskih odluka, generisanju izveštaja kao što su radiološki izveštaji ili otpusne liste, prevođenju medicinskih termina, podršci mentalnom zdravlju, medicinskim konsultacijama, kao i u pružanju znanja o medicini [20]. U finansijama može dodatno unaprediti već postojeće primene veštačke inteligencije u upravljanju portfeljem, modelovanju finansijskog rizika, finansijskom savetovanju i korisničkoj podršci [21]. S obzirom na mogućnost generisanja koda od strane određenih velikih jezičkih modela moguća je primena i u softverskom inženjerstvu.

2.6 ChatGPT

Izuzetan primer primene velikih jezičkih modela je ChatGPT, koji koristi velike jezičke modele iz serije GPT prilagođene za vođenje dijaloga, pri čemu demonstrira izvanrednu sposobnost komunikacije s ljudima [15]. ChatGPT veb aplikacija koja je omogućila javnosti interakciju sa velikim jezičkim modelima zasnovana je na *Generative Pretrained Transformer* (GPT) modelu što je i deo samog naziva ovog četbota. Predstavlja četbot aplikaciju, baziranu na velikim jezičkim modelima, koja koristi tehnike dubokog učenja za generisanje prirodnog teksta kao odgovor na upite [22].

Kao što je već navedeno, ChatGPT aplikacija je zasnovana na OpenAI GPT-3, GPT-4 i GPT-5 velikim jezičkim modelima. Prethodnici ovih modela, GPT-1 kreiran 2018. godine i GPT-2 kreiran 2019. godine, trenirani su na dosta manjoj količini podataka. Set podataka korišćen za GPT-1 model uključuje samo knjige različitih žanrova preuzete sa BookCorpus-a [23], dok za GPT-2 uključuje veb tekst koji predstavlja kolekciju od preko milion veb stranica validiranih od strane ljudi preuzetih sa Reddit društvene mreže [24]. Povećanje veličine modela i korišćene količine podataka je uticalo na ogromno povećanje performansi GPT-3 modela u odnosu na prethodne [19]. Set podataka koji je korišćen za GPT-3 model obuhvata tekst sa vikipedije na engleskom jeziku, dva seta knjiga dostupnih na internetu, takođe veb stranice do kojih vodi Reddit, kao i CommonCrawl repozitorijum veb podataka [25]. Napredak u ovoj oblasti doveo je do uvođenja i GPT-4 modela koji je dodatno proširio mogućnosti jezičkih modela integracijom vizuelnih informacija kao deo ulaza, pdf fajlova i slika [8]. Za predviđanje sledećeg tokena ovaj model koristi i javno dostupne, ali i licencirane podatke trećih strana [26]. GPT-3.5, GPT-4 i GPT-5 set podataka nije javno objavljen.

Svaka naredna verzija GPT modela je unapređivana povećavanjem seta podataka, kao što je već i navedeno, i broja parametara koji pozitivno utiču na količinu podataka koje model može da procesira, od kojih uči i koje može da generiše. GPT-1 ima 0,2 milijarde parametara [27], GPT-2 1.5 milijardi parametara [24], dok GPT-3 175 milijardi parametara [25]. OPENAI nije zvanično objavio broj parametara za novije modele, ali recimo za GPT-4 se u radu [28] navodi da ima preko 100 biliona parametara.

ChatGPT počevši od GPT-3.5 modela pruža jedinstvene slučajeve korišćenja u odnosu na prethodne GPT modele, kao što je odgovaranje na pitanja, razjašnjavanje kompleksnih koncepata i generisanje i ispravka koda [22]. Spomenute osobine su pospešile potencijal ChatGPT-ja da promeni način na koji se različite profesije obavljaju [29]. ChatGPT, kao univerzalni asistent, može značajno unaprediti efikasnost i produktivnost u mnogim oblastima [19]. Korišćenje ChatGPT-ja kao ličnog asistenta u gotovo svakom poslu moglo bi postati nova realnost. Ovo bi moglo izazvati dodatne promene u radnim

zadacima i započeti transformaciju definicija poslova. ChatGPT bi mogao (i verovatno hoće) promeniti način na koji učimo, komuniciramo i radimo [30]. Fokus bi trebalo da bude na preispitivanju zadataka koji se mogu preneti s ljudi na veštačku inteligenciju i obrnuto [31]. Akademska zajednica, obrazovanje, kao i gotovo svaka druga industrija, sigurno će, ako već nisu, biti pod uticajem ovog i sličnih alata.

Obrazovanje je svakako jedno od polja koje je već doživelo negativne posledice i koje mora uložiti mnogo truda kako bi se prilagodilo novim okolnostima. Potencijal za zloupotrebu u obrazovanju će rasti u budućnosti, kako raste svest o alatima veštačke inteligencije. Problemi u obrazovnom sektoru, poput plagijarizma, već su se pojavili i rešavaju se na različite načine. Neki od pristupa uključuju zabranu pristupa ChatGPT-ju u školama i njeno korišćenje u domaćim zadacima, nastavnim aktivnostima i proceni znanja [4], [32], kao i prilagođavanje samih zadataka novoj tehnologiji [30], [32].

S druge strane, realnost je da veštačka inteligencija ne predstavlja tehnologiju budućnosti [33]. Ona se u nekoj formi već koristi u obrazovanju širom sveta [33]. Neizbežno je da je moramo prihvatiti [30]. Generativna veštačka inteligencija ima potencijal da značajno utiče na promene obrazovnog procesa [34]. Ukoliko fokus razvoja veštačke inteligencije postane unapređenje konverzionih pretraživača [19], obrazovanje se mora uskladiti i pripremiti za svet u kojem će veštačka inteligencija imati značajnu ulogu. Umesto zabrane ovih alata treba uticati na rast svesti o njihovim prednostima i ograničenjima u obrazovnom okruženju [30].

Potencijal veštačke inteligencije u nastavnim i obrazovnim okruženjima još uvek nije u potpunosti iskorišćen od strane edukatora [35]. Kao disruptivna tehnologija, ChatGPT je pokrenuo mnoga pitanja vezana za proces učenja i podučavanja. Kako bi se studenti pripremili za budućnost u kojoj ChatGPT ima značajan uticaj, obrazovanje mora aktivno sarađivati sa veštačkom inteligencijom [33]. Sektor obrazovne tehnologije takođe će biti pod uticajem ChatGPT-ja [29] i sličnih alata.

S obzirom na to da konverzacione aplikacije velikih jezičkih modela mogu imati pozitivan uticaj na društvo i obrazovni sistem, trebalo bi iskoristiti prilike koje pružaju za unapređenje obrazovanja. U narednom periodu, modelovanje odgovorne upotrebe ChatGPT-ja trebalo bi da postane prioritet u obrazovnom polju [35].

3 ELEKTRONSKO UČENJE

3.1 Pojam i definicija elektronskog učenja

Razvoj informaciono-komunikacionih tehnologija doveo je do značajnih promena u različitim oblastima društva, uključujući i obrazovanje. Integracija tehnologija u obrazovni proces omogućava razvoj inovativnih pristupa učenju koji su prilagođeni savremenim potrebama studenata i nastavnika [36]. U tom kontekstu, elektronsko učenje (e-učenje) je razvijeno kao odgovor na zahteve za fleksibilnijim, interaktivnijim i dostupnijim obrazovnim metodama. Korišćenjem različitih digitalnih alata i platformi, e-učenje može pružiti fleksibilnije i personalizovanije obrazovno iskustvo, nezavisno od vremenskih i prostornih ograničenja, čime se proces sticanja znanja može značajno unaprediti. Jedna od najvažnijih karakteristika i prednosti e-učenja je to što je usmereno na učenika, za razliku od tradicionalnog okruženja čiji je fokus nastavnik [37].

U opštem smislu, elektronsko učenje je oblik učenja uz pomoć računara. Ovaj tip učenja omogućava fleksibilan prenos znanja i veština velikom broju korisnika u različitim vremenskim periodima i na različitim lokacijama [36]. Oslanja se na digitalne tehnologije i internet za pružanje resursa za učenje [38], kao i za omogućavanje interakcije između studenata i nastavnika [39]. Pored ovog naziva, u literaturi se mogu pronaći i alternativni termini korišćeni za oslovljavanje e-učenja tokom poslednjih decenija, kao što su otvoreno učenje, učenje na daljinu, obrazovanje putem interneta i onlajn učenje [36].

Glavna primena edukativnih tehnologija u samim obrazovnim institucijama je u vidu sistema za upravljanje učenjem (eng. *Learning Management System* – LMS) [40]. Ovi sistemi, dizajnirani za potrebe e-učenja, olakšavaju planiranje, upravljanje i isporuku sadržaja za učenje. Materijali za učenje mogu biti u različitim formatima, npr. vidu teksta, slika, audio, video snimaka i sl. LMS eliminiše geografska i vremenska ograničenja tako da materijali za učenje postaju dostupniji, interaktivniji i efikasniji za upotrebu [36].

Istraživanja su kroz godine pokazala da sistemi za učenje imaju pozitivan uticaj na obrazovanje koji se ogleda u pospešivanju rezultata učenika [41]. Takođe, COVID-19 pandemija je uticala na rast korišćenja sistema za upravljanje učenjem u obrazovanju [42], kao i na masovno prihvatanje sistema za e-učenje [36]. Različite onlajn platforme za učenje su postale popularnije [43]. Samim tim, može se zaključiti da je sektor edukativnih tehnologija zbog spomenutih dešavanja dobio priliku za proširenje [44].

3.2 Kategorije sistema elektronskog učenja

Liu i Yu [36] navode i opisuju kategorije sistema za e-učenje na osnovu ciljne grupe korisnika i troškova. Prema spomenutim kriterijumima trenutni sistemi elektronskog učenja mogu se klasifikovati u dve vrste: platforme za masovne otvorene onlajn kurseve (eng. *Massive Open Online Course* - MOOC) i sisteme za upravljanje učenjem (eng. *Learning Management System* – LMS). MOOC platforme su otvorene za veliki broj učenika. Iako su neki od dostupnih MOOC kurseva kreirani od strane određenih univerziteta, oni zapravo nisu ograničeni samo na studente tih institucija visokog

obrazovanja. Svako im može pristupiti bez obzira na lokaciju, kulturu, nacionalnost ili bilo koje druge kriterijume. S druge strane, LMS-ovi se obično koriste u institucijama visokog obrazovanja. Samim tim, nisu po pravilu otvoreni za širu javnost, već im može pristupiti samo određena grupa korisnika. Takođe, većina MOOC-ova je besplatna ili ima minimalne troškove kako bi kursevi bili dostupni pojedincima, dok je trošak sistema za upravljanje učenjem viši, u zavisnosti od broja korisnika, i uglavnom ove troškove snose institucije visokog obrazovanja, a ne pojedinci.

Neke od najpoznatijih MOOC platformi su Coursera i edX koje sadrže kurseve najprestižnijih fakulteta kao što su Stanford, Yale, MIT i Harvard, dok je jedan od najkorišćenijih LMS-ova u obrazovnim institucijama Moodle.

Pored klasifikacije sistema za e-učenje na osnovu ciljne grupe korisnika i troškova, Jethro i saradnici [38] navode razliku u načinu dostavljanja sadržaja koje može biti asinhrono i sinhrono. U sinhronom sistemu nastavnik upravlja elektronskim učenjem u realnom vremenu. Svi studenti dobijaju informacije paralelno i mogu da komuniciraju direktno sa nastavnikom i ostalim studentima. Primer ovih sistema su aplikacije koje omogućavaju prenos audio ili video snimaka uživo, odnosno telekonferenciju. S druge strane, kod asinhronih sistema prenos i prijem informacija se ne dešava u isto vreme. Različite metode mogu biti korišćene za prenos materijala. Jedan takav primer su upravo sistemi za upravljanje učenjem.

3.3 Prednosti elektronskog učenja

Tokom godina njegove primene, uočene su brojne pozitivne strane ovog načina učenja. Moguće prednosti e-učenja se ogledaju u sledećim karakteristikama [45]:

- Fleksibilnost po pitanju vremena i mesta učenja, jer svaki student ima mogućnost da izabere vreme i mesto koje mu najviše odgovara za učenje
- Poboljšana efikasnost sticanja znanja i veština zbog lakog pristupa velikoj količini informacija
- Omogućena interakcija između učenika putem foruma za diskusiju što može uticati na suzbijanje straha od komunikacije, kao i podsticati međusobno poštovanje različitih mišljenja
- Isplativost sa aspekta eliminacije potreba za putovanjem kao i smanjenja potreba za velikim brojem zgrada za održavanje nastave
- Uvažavanje individualnih razlika učenika u određenoj meri s obzirom na to da svaki učenik može bira svoj način učenja npr. neki studenti se fokusiraju samo na određene delove kursa, dok drugi preferiraju da pregledaju ceo kurs
- Pomoć u rešavanju problema nedostatka akademskog osoblja, uključujući predavače, ali i pomoćno osoblje
- Učenje sopstvenim tempom što smanjuje stres i povećava zadovoljstvo učenika

Sloboda izbora po pitanju vremena i mesta učenja bez potrebe za putovanjem i odabir individualnog tempa učenja koji odgovara studentu može pozitivno uticati na njihov proces učenja ukoliko je kurs kvalitetno implementiran. Kroz sistem za učenje, student

može pregledati materijale onoliko puta koliko mu je to potrebno. Međutim, svakako je važno napomenuti da uspešnost ovog tipa učenja zavisi od njegove adekvatne implemenatacije i dostupnosti kvalitetnih materijala kreiranih od strane nastavnika.

3.4 Izazovi elektronskog učenja

Međutim, uprkos značajnim prednostima e-učenja, uočeni su i određeni izazovi u izvođenju ovog tipa učenja. Liu i Yu [36] navode sledeće izazove na osnovu pregleda literature koje donose sistemi za e-učenje:

- Primena novih pedagoških i kognitivnih pristupa: Sistemi elektronskog učenja zahtevaju drugačiji pedagoški pristup. Nije dovoljno samo postaviti veliku količinu resursa za učenje na ove platforme što je čest slučaj. Elektronsko učenje zahteva prilagođavanje kurseva kako bi bili atraktivniji i interaktivniji za studente, jer se značajno razlikuje od tradicionalne nastave licem u lice.
- Nedostatak adekvatnih strategija dizajna kurseva: Ne postoje dovoljno razvijene strategije dizajna kurseva prilagođene elektronskom učenju niti odgovarajuća evaluacija uspešnosti njegove implementacije.
- Poteškoće u realizaciji praktičnih kurseva: Praktični kursevi, poput laboratorijskih vežbi, su teško izvodljivi putem sistema elektronskog učenja. Studenti u takvim situacijama često ne mogu u potpunosti savladati sadržaj kao što bi to učinili u tradicionalnoj učionici.
- Otpor prema elektronskom učenju: Nedostatak motivacije primećen je kako kod studenata, tako i kod nastavnika. Studenti koji nisu dovoljno motivisani mogu zaostajati zbog potrebe za samoregulacijom u elektronskom učenju, gde često nedostaju adekvatan nadzor i smernice.
- Visoka stopa odustajanja: Statistika pokazuje da je stopa odustajanja na MOOC platformama veoma visoka, što ukazuje na to da većina studenata koji započnu kurs ne uspevaju da ga završe.
- Trošak e-učenja: Finansijski troškovi za studente iz porodica sa niskim primanjima mogu biti prepreka onlajn obrazovanju. Resursi kao što su računar i stabilna internet konekcija nisu svima dostupni.
- Sajber bezbednost i privatnost: Za većinu sistema za e-učenje, sajber bezbednost i privatnost su opcionalna funkcija što može dovesti sistem i njegove podatke u opasnost. Zato je vrlo bitno da institucije prilikom odabira sistema za e-učenje stavljaju u fokus njihovu pouzdanost, i da obrate pažnju na pooštavanje sigurnosti i privatnosti onlajn obrazovanja.

Takođe, iako LMS i MOOC imaju ključnu ulogu u usmeravanju procesa učenja, neke od osnovnih odgovornosti nastavnika kao što je personalizovana motivacija i fidbek, kao i prilagođavanje programa potrebama svakog studenta, nisu u potpunosti obuhvaćene ovim tehnologijama [1]. Potreba za dodatnim pojašnjenjima materije može uticati na nižu efikasnost e-učenja u odnosu na tradicionalne metode učenja, jer je proces učenja mnogo lakši uz direktan kontakt licem u lice sa nastavnicima [45]. Takođe, ovaj način učenja zahteva veoma snažnu motivaciju, kao i veštine upravljanja vremenom [45]. Ovi

izazovi naglašavaju potrebu za inovativnijim pedagoškim pristupima i boljim strategijama implementacije sistema elektronskog učenja.

3.5 Hibridno učenje

Hibridno učenje (eng. *blended learning*) predstavlja kombinaciju tradicionalnog učenja u učionicama i elektronskog tj. onlajn učenja. Određeni izazovi elektronskog učenja mogu biti prevaziđeni uz kombinaciju sa tradicionalnim učenjem, odnosno izazovi tradicionalnog obrazovanja mogu biti prevaziđeni uz kombinaciju sa elektronskim učenjem. Ovaj pristup kombinuje najbolje prakse onlajn i tradicionalnog učenja u učionici. Na taj način se stvara efikasno okruženje za učenje u kojem se aktivnosti učenja odvijaju unutar, ali i van učionice [36].

Hibridno obrazovanje kao balansirani i prilagodljivi pristup koji može prevazići izazove onlajn učenja u kombinaciji sa tradicionalnim obrazovanjem je preporuka za obrazovne institucije [46]. Za razliku od tradicionalnog obrazovanja, gde je u učionici fokus na nastavniku, hibridno učenje omogućava učenicima upotrebu digitalnih resursa čime dobijaju priliku da postanu protagonisti sopstvenog procesa učenja i dolaze do znanja zajedno sa nastavnicima, dok istovremeno nudi fleksibilniji stil učenja prilagođen različitim potrebama [36].

4 EDUKATIVNI ČETBOTOVI

4.1 Definicija i pojam četbotova

Četbot, takođe poznat i kao konverzacioni agent (eng. *conversational agent*), predstavlja kompjuterski program koji kroz konverzaciju pruža usluge korisnicima [47]. Zahvaljujući najnovijim dostignućima u mašinskom učenju i obradi prirodnog jezika, ovi sistemi su u stanju da stupe u interakciju sa svojim korisnicima na različite načine koristeći prirodni jezik [1]. Pronašli su primenu u mnogim oblastima kao što su marketing, korisnička podrška, zdravstvo, zabava i edukacija.

Prva generacija konverzacioni agenata koristila je pristupe zasnovane na pravilima kako bi korisnici u zavisnosti od pitanja dobili odgovarajući odgovor [1]. Uprkos činjenici da su odgovori ovih modela gramatički ispravni, njihova mana je nemogućnost generisanja raznovrsnih odgovora zbog upotrebe unapred definisanih parova pitanje/odgovor. Vremenom su razvijeni generativni modeli koji mogu prevazići ova ograničenja. Ait Baha i saradnici [1] navode da je napredak postignut u oblasti mašinskog učenja i tehnikama obrade prirodnog jezika, doveo do toga da generativni modeli mogu interagovati sa ljudima, identifikovati njihove zahteve, mapirati njihove želje i ponuditi im odgovarajuću akciju bez potrebe za ljudskom intervencijom.

4.2 Kratak istorijat četbotova

Okonkwo i Ade-Ibijola [2], kao i Kuhail i saradnici [48] navode ukratko istorijski pregled prvih četbotova. Na početku, za kreiranje odgovora, četbotovi su koristili mehanizme za prepoznavanje ključnih reči. ELIZA, prvi poznati četbot kreiran 1996. godine, koji je predstavljao psihoterapeuta, je koristio prepoznavanje ključnih reči u korisničkom ulazu, i šablone da bi formulisao odgovor. Naredni četbot koji je predstavljao napredak u odnosu na ELIZA četbot je takođe bio u oblasti mentalnog zdravlja. PARRY četbot iz 1972. godine, koji je simulirao osobu sa šizofrenijom, je koristio heuristike zasnovane na pravilima i poboljšane tehnike za prepoznavanje šablona.

Jabbervackyy iz 1988. godine predstavlja prvi četbot koji je koristio veštačku inteligenciju. Za formiranje odgovora je korišćena prethodna diskusija i kontekstualno prepoznavanje šablona [49]. Sledeći korak u razvoju četbotova predstavlja ALICE četbot iz 1995. godine, baziran na *markup* jeziku veštačke inteligencije (eng. *Artificial Intelligence Markup Language* - AIML) i prepoznavanju šablona [48]. Prepoznat je u to vreme kao prvi ljudski kompjuter.

SmarterChild četbot koji se pojavio 2001. godine je obeležio značajan napredak u razvoju mašinske inteligencije i interakcije čovek-računar s obzirom na to da je omogućio pristup informacionim sistemima kroz konverzaciju [49]. Predstavlja prethodnika trenutno aktuelnih virtuelnih asistenata kao što su Alexa i Siri četbotovi [48]. Novina koju je uveo u svet četbotova jeste pomoć ljudima u praktičnim dnevnim zadacima jer je mogao da povuče informacije iz baza podataka o vremenima prikaza filmova, sportskim rezultatima, cenama na berzi i vremenskoj prognozi [49].

Pametni personalni glasovni asistenti ugrađeni u pametne telefone ili kućne zvučnike su sledeći korak u razvoju četbotova zasnovanih na veštačkoj inteligenciji koji razumeju glasovne komande, razgovaraju digitalnim glasom, i upravljaaju zadacima kao što su praćenje kućnih automatizovanih uređaja, kalendara, mejla itd. [49]. Pojavili su se u drugoj deceniji 21. veka. Neki od najpoznatijih personalnih glasovnih asistenata su Apple Siri, Amazon Alexa, Microsoft Cortana, Google Assistant i IBM Watson.

Generativni četbotovi trenirani na ogromnim količinama podataka radi predviđanja naredne reči, koji su tema ovog rada, se pojavljuju počevši od kraja druge decenije 21. veka. ChatGPT, Gemini, Claude i sl. predstavljaju primere četbot aplikacija ovog tipa sa kojima korisnici mogu voditi konverzacije na različite teme i u cilju rešavanja različitih zadataka. Više reči o njihovoj primeni u edukaciji će biti dat u narednom poglavlju.

4.3 Četbotovi u edukaciji

4.3.1 Definicija i ciljevi edukativnih četbotova

Konverzacioni agenti odnosno četbotovi u obrazovanju, takođe poznati i kao virtuelni tutori u ovom kontekstu, su sistemi veštačke inteligencije dizajnirani da simuliraju razgovore i pružaju interaktivnu podršku učenicima. Nalaze se među najistaknutijim sistemima veštačke inteligencije koji su korišćeni u procesu nastave i učenja [1]. Predstavljaju podtip dijalog sistema kao edukativnih tehnologija, usmerenih na pružanje povratnih informacija studentima koji traže pomoć, što se može uporediti s podizanjem ruke u učionici [40].

Ait Baha i saradnici [1], kao i Wollny i saradnici [40] navode kao glavne ciljeve četbotova u edukaciji povećanje motivacije učenika, poboljšanje akademskih postignuća odnosno efikasnosti studiranja i razvoj kompetencija odnosno veština. Prvi navedeni cilj se odnosi na povećanje volje i motivacije učenika za učenjem i smanjenje stresa u procesu učenja. Poboljšanje akademskog uspeha se odnosi i na povećanje ukupne efikasnosti obrazovanja. Na primer, automatizacija određenih zadataka, kao i dostupnost usluga studentima koje im mogu uštedeti vreme, kao što je lak pristup administrativnim informacijama, mogu pozitivno uticati na ostvarenje ovog cilja. Razvoj kompetencija učenika je krajnji cilj i predstavlja cilj na koji se obrazovni četbotovi najviše fokusiraju. Četbotovi sa ovim ciljem se posmatraju kao vrsta alata za pomoć studentima u procesu učenja. Takođe, doprinose i povećanju dostupnosti obrazovanja, jer mogu pružati savete i podršku u učenju bez obzira na fizičku prisutnost studenata na određenom mestu [40].

4.3.2 Kategorije edukativnih četbotova

Dve glavne vrste primena četbotova u edukaciji predstavljaju: četbotovi namenjeni pružanju podrške u obavljanju administrativnih zadataka studentima i četbotovi namenjeni podučavanju/učenju studenata. U nastavku će biti navedene moguće namene i primeri ovih vrsta četbotova razvijenih pre napretka u oblasti velikih jezičkih modela. Takođe, pored četbotova za učenje i pomoći za administrativne zadatke studentima, autori Ait Baha i saradnici [1] su izdvojili još jednu vrstu edukativnih četbotova koji su usmereni na samorazvoj i praćenje napretka studenata.

Prva vrsta četbotova za pružanje administrativne pomoći je već uspešno primenjena u ostalim poljima kao što su medicina, bankarski sektor, korisnička podrška itd. Ovi četbotovi mogu pomoći autorizovanim studentima da lakše pristupe rasporedu ispita, rasporedu konsultacija, sadržaju obrađenih tema [2], kriterijumu ocenjivanja, krajnjim rokovima, kao i u pronalaženju resursa za učenje [50]. Neki od primera ove vrste četbotova u edukaciji su: Lola četbot za pomoć pri upisu na fakultet, Dina četbot za prijavu studenata, LISA za upoznavanje studenata sa načinom života na univerzitetu, LTKABot za administrativne zadatke i sl. [51].

Četbotovi namenjeni podučavanju koriste se u formalnom i neformalnom obrazovanju i imaju za zadatak da generišu znanje na određenu temu [51]. Rezultati istraživanja pokazuju pozitivan uticaj učenja zasnovanog na korišćenju četbotova na različite faktore, kao što su sticanje znanja, učenje jezika, samousmereno učenje, aktivno učešće, motivacija, društvena interakcija [1]. Uloge četbotova za podučavanje mogu biti različite. Četbot može biti pomoćnik nastavniku, ili kompletan edukator u nekoj oblasti [51]. Najčešće implementacije su specijalizovani četbotovi za učenje jezika i računarskih nauka [47], [48]. Neki od primera su: Autotutor za pružanje znanja kroz konverzaciju; Chatbot koji je specijalizovan za računarske nauke; ScratchThAI za učenje Scratch programskog jezika; SCIEC za učenje engleskog; Duolingo za učenje jezika itd.

4.3.3 Četbotovi i elektronsko učenje

Tradicionalni sistemi elektronskog učenja često nemaju mogućnosti za personalizovanu podršku i interakciju između studenata i nastavnika u realnom vremenu, što može smanjiti angažovanost studenata i efikasnost obrazovnih iskustava [52], [53], kao i negativno uticati na motivaciju studenata [1]. Na duže staze, nemogućnost da se odgovori na individualne izazove učenika tokom savladavanja materijala za učenje i ostali prethodno spomenuti nedostaci mogu nepovoljno uticati na efikasnost studiranja, ali i na razvoj veština potrebnih za primenu stečenog znanja u realnim situacijama.

Ait Baha i saradnici [1] zaključuju da odsustvo direktne interakcije sa nastavnicima dovodi do toga da su studenti često prepušteni da samostalno savladaju dostupan statičan sadržaj uz ograničene mogućnosti za instant povratne informacije i individualizovanu pomoć. Navode da iz navedenog proističe smanjeno angažovanje sa nastavnim materijalom i poteškoće u ispunjavanju obrazovnih ciljeva, što može ometati i negativno uticati na ukupan akademski napredak.

Da bi se odgovorilo na spomenuta ograničenja, raste interesovanje za integraciju četbotova u okruženja za elektronsko učenje [1]. Edukativni četbotovi bi potencijalno mogli da pronađu primenu u rešavanju problema ograničene komunikacije koja se sreće u onlajn sistemima za učenje [7]. Mogu da pruže personalizovane odgovore i ponude odgovarajuću pomoć zasnovanu na individualnim zahtevima i potrebama učenja [1]. Adekvatna integracija četbota u e-učenje omogućava prilagođavanje odgovora specifičnim potrebama učenika, čime se stvara mogućnost učenja sopstvenim tempom uz manje stresa, što ne samo da poboljšava celokupno iskustvo učenja, već i štedi vreme i održava motivaciju studenata [1]. Međutim, iako njihova integracija može uticati pozitivno na obrazovni proces, treba obratiti pažnju na to da njihova efikasnost zavisi od prisustva visokokvalitetnog sadržaja koji potiče od predavača [7].

Na osnovu dosadašnjih istraživanja može se zaključiti da učenje uz pomoć četbotova predstavlja prijatno okruženje za rad [2]. Integracija ovih sistema veštačke inteligencije u pametne učionice obezbeđuje podržavajuće okruženje kroz visokokvalitetnu interakciju sa studentima, i omogućava im aktivnije učešće i postizanje boljih akademskih rezultata [1]. Studenti danas uče kroz onlajn platforme, i radije upijaju sadržaj sa telefona, nego koristeći knjige [2]. Međutim, iako četbotovi mogu služiti kao onlajn pomoćnici u učenju, oni ne mogu u potpunosti zameniti predavače, i zato ih treba posmatrati kao vrlo važne alate u oblasti edukacije, dizajnirane tako da unaprede njihovu ulogu i utiču na povećanje dostupnosti kursa [7].

4.4 Prednosti primene četbotova u edukaciji

Četbotovi u obrazovanju mogu imati pozitivan uticaj na proces učenja. Na osnovu pregleda literature autori Okonkwo i Ade-Ibijola [2] navode sledeće prednosti primene četbotova u edukaciji:

- Personalizacija učenja,
- Pozitivan uticaj na motivaciju,
- Pružanje instant pomoći i podrške studentima,
- Omogućen brz pristup informacijama od strane više studenata u isto vreme,
- Paralelan rad sa više studenata, odnosno mogućnost procesuiranja više pitanja u isto vreme od strane četbota.

Četbotovi se ne umaraju kao nastavnici i mogu biti sve vreme dostupni studentima kao pomoć u procesu edukacije. U trenutku kad studentima zatreba pomoć tokom učenja ili u obavljanju administrativnih zadataka, mogu istu i da dobiju bez čekanja na termin konsultacija ili sledećeg predavanja. Ovi sistemi mogu da procesuiraju više pitanja u isto vreme i samim tim rade sa više studenata paralelno u bilo koje doba dana i noći što nije slučaj sa nastavnicima. Podstiču personalizaciju učenja time što odgovaraju na individualne potrebe i zahteve studenata u bilo koje vreme.

4.5 Izazovi primene četbotova u edukaciji

Četbotovi u edukaciji, ukoliko su adekvatno implementirani, donose brojne prednosti koje mogu da utiču na značajne pozitivne promene u obrazovanju. Međutim, neophodno je spomenuti do sad uočene izazove integracije ovih sistema u proces obrazovanja. Na osnovu analize literature, neki od uočenih izazova edukativnih četbotova mogu biti grupisani na sledeći način:

- Tehnološka ograničenja: Programiranje četbot sistema predstavlja jedan od najvećih izazova u njihovom razvoju kako bi davali tačne odgovore [2]. Nedovoljan ili neadekvatan set podataka za treniranje tradicionalnih četbot sistema može dovesti do frustracije studenata zbog dobijanja netačnih odgovora na pitanja [48], kao i ukoliko konverzacija nije dovoljno fluidna [51]. Takođe, nedostatak emocije u konverzacijama negativno utiče na stanje studenata tokom interakcije učenja [54].

- Etički problemi: Moguće su različite zloupotrebe podataka tako da treba obratiti pažnju na privatnost konverzacija, i pojašnjenje samih funkcionalnosti četbota korisnicima koji samostalno treba da izaberu način interakcije sa istim [2].
- Efekat novine: Vrlo je važno spomenuti i efekat novine koji može da utiče pozitivno na prihvatanje četbota u učenju na početku, dok s vremenom zainteresovanost opada [48], [51].

Huang i saradnici [54] navode da nekompletni upiti od strane studenata na koje četbot ne bi mogao ispravno da odgovori, kao i generisanje izlaza koji nema smisla u nekim situacijama je negativno uticalo na prihvatanje četbotova u učenju od strane studenata. Takođe, duže konverzacije, kao i odstupanje od teme konverzacije može da dovede do toga da četbot sistem ne može da adekvatno razume diskusiju i pruži odgovarajući odgovor [54].

Razvoj četbota koji bi mogao da odgovori na gotovo svaki studentski upit predstavlja izazov. Dizajn četbotova za otvorene konverzacije je složen proces zbog potrebe da se obuhvati raznoliki spektar načina na koje korisnici mogu komunicirati sa istim [55]. Predviđanje i mapiranje svakog korisničkog zahteva nije izvodljivo [56]. Samim tim, generisanje besmislenih odgovora u određenim situacijama odnosno nespремnost na sva moguća pitanja je negativno uticalo na prihvatanje ove tehnologije u edukaciji. Međutim, pojava četbotova sa najboljim performansama do sada koji su zasnovani na velikim jezičkim modelima [29] potencijalno može unaprediti njihovu upotrebu u obrazovanju. Napredak u oblasti veštačke inteligencije i obrade prirodnog jezika koji se dogodio poslednjih godina daje prostor za razvoj i primenu ovih sistema u različitim oblastima. Zbog osobine generalizacije u različitim zadacima koju pružaju veliki jezički modeli, odnosno mogućnosti da daju korisne odgovore i u scenarijima za koje nisu eksplicitno trenirani, četbotovi dobijaju priliku za razvoj i potencijalno razrešenje spomenutih tehnoloških izazova.

5 VELIKI JEZIČKI MODELI U OBRAZOVANJU

5.1 Generativna veštačka inteligencija u obrazovanju

Generativna veštačka inteligencija pored ostalih oblasti, pronalazi svoju primenu i u edukativnom sektoru. Svaki nivo obrazovanja može na neki način integrisati ovu tehnologiju radi poboljšanja iskustva učenja, pri čemu se izdvajaju računarske nauke, medicina i učenje jezika kao oblasti u kojima ovi sistemi mogu stvoriti vrednost i u procesu obrazovanja [6]. Mogu ih koristiti i studenti kao pomoć pri učenju, ali i nastavnici kako bi unapredili obrazovni proces. Konstantna dostupnost ovih alata, koja se ne može meriti sa dostupnošću nastavnika, kao i pružanje prilagodljive podrške studentima u procesu učenja su jedni od ključnih faktora primene ovih tehnologija u obrazovanju.

Dakle, prisutan je veliki potencijal za pozitivnu upotrebu alata veštačke inteligencije radi unapređenja i poboljšanja obrazovnih iskustava širom sveta [57]. Autori u [58] su definisali okvir primene generativne veštačke inteligencije u obrazovanju, koji može pomoći nastavnicima koji se odluče za njihovu implementaciju u nastavi, i koji obuhvata sledeće korake:

- Identifikaciju željenih ishoda: Pre primene generativne veštačke inteligencije u obrazovanju važno je odrediti ciljeve korišćenja. Na taj način se osigurava da upotreba tehnologije bude u skladu sa željenim rezultatima.
- Određivanje odgovarajućeg nivoa automatizacije: U zavisnosti od ciljeva, može biti primereno potpuno automatizovati proces učenja i podučavanja korišćenjem veštačke inteligencije u obrazovanju ili je koristiti kao dopunu tradicionalnim metodama nastave.
- Osiguranje etičkih aspekata: Etičke implikacije upotrebe veštačke inteligencije u obrazovanju moraju se pažljivo razmotriti, uključujući potencijalne pristrasnosti koji mogu postojati u sistemu i njihov uticaj na nastavnike i učenike.
- Procenu efikasnosti: Na kraju, važno je proceniti efikasnost veštačke inteligencije u obrazovanju u postizanju željenih ishoda.

Autori Wei i saradnici [59] ističu da funkcionalnosti koje pruža generativna veštačka inteligencija omogućavaju razvoj novih modela podučavanja koji kombinuju interakciju između ljudi i mašina. Na taj način se kreiraju inovativni pristupi za postizanje cilja digitalizovanog obrazovanja. Navode da okruženja za učenje mogu postati inteligentnija, omogućavajući diskusiju, interakciju, podučavanje i druge nastavne aktivnosti u bilo koje vreme i na bilo kom mestu.

Pored mogućnosti generisanja teksta koji je fokus u ovoj disertaciji i o čijim potencijalnim primenama će biti više reči u ostalim poglavljima, generisanje slika je među najzapaženijim oblicima generativnih alata. Generisanje slika može pomoći studentima u određenim aspektima učenja. Korišćenjem generativnih alata čiji su rezultati slike, kao što su DALL-E i Midjourney, moguće je vizuelizovati apstraktne i kompleksne teme [6]. Takođe, studija slučaja sprovedena od strane Lively i saradnika [60] je pokazala da korišćenje ovih alata tokom pohađanja predmeta o razvoju veb sajtova pozitivno utiče na kvalitet završnih projekata. Spomenuti alati pomažu u

osmišljavanju samog rešenja, kao i u odabiru boja. Tekst odnosno kod generatori pomažu u implementaciji samih rešenja.

Kako su veliki jezički modeli koji predstavljaju vid generativne veštačke inteligencije čiji je izlaz tekstualnog tipa tema ovog rada, u nastavku će fokus biti stavljen na njihovu primenu u edukaciji. S obzirom na to da ChatGPT predstavlja sistem veštačke inteligencije koji je približio široj javnosti velike jezičke modele i predstavlja aplikaciju koja je najbrže došla do milion korisnika sledeća poglavlja opisuju moguće primene, ali i implikacije koje njegova primena u edukaciji donosi. Većina implikacija se može preslikati i na ostale popularne četbot aplikacije koje koriste u pozadini velike jezičke modele kao što su Google Gemini, Anthropic Claude i druge, s obzirom na to da predstavljaju dijalog sisteme u okviru kojih se kroz konverzaciju dolazi do saznanja.

Sledeća poglavlja su organizovana na sledeći način. Na početku se navodi uticaj koji je ChatGPT doneo u sferu provere znanja i ocenjivanja, a koji uvode i ostali slični alati veštačke inteligencije. Nakon toga se opisuju mogućnosti upotrebe velikih jezičkih modela sa fokusom na ChatGPT aplikaciju kao nastavničkog, a zatim i kao studentskog pomoćnog alata. Tokom pregleda literature, kao neizbežna tema je identifikovana potreba za širenjem digitalne pismenosti o velikim jezičkim modelima u obrazovanju kako bi njihova upotreba bila adekvatna. Takođe, u okviru ovog poglavlja navedeni su rezultati početnog istraživanja u ovoj doktorskoj disertaciji, koje se tiče faktora koji utiču na prihvatanje ChatGPT-ja kao studentskog pomoćnika u učenju. Zatim su navedene mogućnosti primene u hibridnom tipu obrazovanja koje je sve češće primenjivano u obrazovnim institucijama, kao i mogućnost integracije sa edukativnim platformama.

5.2 Uticaj na proveru znanja

ChatGPT je dobio negativnu konotaciju u početku i prouzrokovao je poremećaje u oblasti obrazovanja zbog visokih performansi u kreativnom pisanju na bilo koju temu [19], ali i u pisanju koda, što može imati direktan uticaj na provere znanja studenata. Sposobnost kreativnog pisanja ChatGPT-ja može uticati na promene u načinu ocenjivanja eseja na određenu temu. Kada su reference potrebne u radovima, ChatGPT je u početku otežavao studentima da dobiju tačne i direktne rezultate koje mogu koristiti. Reference koje veliki jezički modeli navode možda nisu tačne zbog načina na koji funkcionišu, odnosno ne citiraju informacije iz nekog izvora, već samo predviđaju sledeću reč. Vremenom je i ova funkcionalnost ChatGPT-ja poboljšana, a i nastali su specijalizovani alati koji mogu pomoći u istraživanju i pisanju uz pravilno citiranje izvora kao što su NotebookLM (Google), Elicit i dr. Međutim, studenti su svakako mogli prvo kreirati nacrt eseja pomoću ChatGPT-ja, a zatim uključiti važne reference iz nastave [35]. Sve ovo ometa i utiče na studentsku inovativnost i kreativnost [4] i sprečava ih da razviju sopstveni stil pisanja [29]. Predstavlja pretnju stvarnom intelektualnom razvoju jer može uticati na kritičko razmišljanje i originalnost [3].

Iako alati za prepoznavanje tekstova napisanih uz pomoć veštačke inteligencije postoje (npr. *OpenAI AI Text Classifier*), oni se smatraju nesavršenim [22]. Izmena jedne reči u pasusu može uticati na promenu rezultata [34]. Sve ovo može dovesti do iskorenjivanja određenih oblika ocenjivanja (npr. eseja koje smo već spomenuli) [61]. Studenti mogu efikasno koristiti ChatGPT kao pomoć, a da nastavnici to nikada ne otkriju [29].

Zbog moguće zloupotrebe u obrazovanju, očekuje se od nastavnika da preispitaju i osmisle nove nastavne filozofije [34] i oblike pedagogije [62] kako bi bili pripremljeni za neizbežni talas zadataka potpomognutih veštačkom inteligencijom [63]. Ocenjivanje mora da prioritizuje kritičko razmišljanje studenata [35]. Neki od pristupa koje treba razmotriti su: fokusiranje na duboka pitanja u okviru discipline [62]; grupni projekti, praktične aktivnosti i usmene prezentacije [34]; diskusije u učionici i debate [64]. Potrebno je preći sa ocenjivanja gotovih eseja na procenu primenjenog kritičkog razmišljanja tokom pisanja [22]. Usmeno objašnjenje procesa razmišljanja može se uporediti sa esejom kako bi se utvrdile značajne razlike koje bi mogle ukazivati na korišćenje ChatGPT-ja ili drugih modela u tom procesu [29]. Dakle, u situacijama gde su eseji striktno potrebni, dodatni usmeni ispiti mogu se koristiti kao dopuna ocenjivanju [22]. Na taj način nastavnici mogu dobiti bolji uvid koji može pomoći u unapređenju ocenjivanja. Takođe, nastavnici mogu zahtevati pisanje iz više puta, sa više kontrolnih tačaka [64]. Sve ovo utiče i na povećanje obima posla samih nastavnika.

Takođe treba obratiti pažnju na sposobnost pisanja i ispravljanja koda, tako da procena znanja programiranja trebalo bi takođe da uključi usmena ispitivanja kao glavni deo procesa ocenjivanja ukoliko se zadaci rade van kontrolisanog okruženja [22].

5.3 Nastavnički pomoćni alat

Mogućnost generisanja različitih resursa za učenje pomoću ChatGPT-ja i drugih alata velikih jezičkih modela otvaraju nove slučajeve korišćenja u kontekstu edukacije, koje nastavnici mogu istraživati i upotrebiti radi unapređenja procesa nastave i učenja [34].

ChatGPT se može efikasno koristiti kao pomoćni alat za osmišljavanje nastavnih planova i provera znanja, za sažimanje sadržaja [65]; kreiranje kvizova [34], [35]; kao i za pripremu naučnih poglavlja i rubrika [35]. Međutim, edukatori treba kritički da procene resurse generisane od strane veštačke inteligencije u pogledu tačnosti i potpunosti dobijenih informacija [4], i prilagode ih svom specifičnom nastavnom kontekstu [35]. Cooper [35] takođe u svom radu navodi da se može razmotriti fino podešavanje modela za prilagođenu upotrebu ChatGPT-ja jer može dovesti do boljih performansi za određene zadatke i generisanja tačnijeg i korisnijeg sadržaja.

Takođe, jedan pravac istraživanja u ovoj oblasti bavi se korišćenjem četbotova za obuku nastavnika, omogućavajući im simulaciju razgovora sa studentima [57]. Pored toga, kako ChatGPT i modeli koje koristi napreduju, ocenjivanje na osnovu razgovora moglo bi postati tema budućih istraživanja [66]. Autori Yildirim-Erbasli i Bulut [66] navode da ovakav tip procena znanja može koegzistirati sa trenutnim procesom ocenjivanja u učionici i unaprediti personalizovana iskustva učenika. Kako je navedeno, procena studentskog znanja kroz konverzaciju bi mogla omogućiti personalizovano upravljanje učenjem na osnovu učinka učenika. Ovo bi takođe moglo pomoći u praćenju napretka učenika i identifikaciji najmanje razumljivih delova predavanja na koje treba obratiti pažnju.

5.4 Studentski pomoćni alat

Bez sumnje, studente nije moguće sprečiti da koriste ChatGPT u kućnim uslovima, niti bi nastavnici trebalo da uvide ograničenja po tom pitanju [65]. Optimalno razumevanje određene teme se postiže kada studenti mogu aktivno da odgovaraju na pitanja u okviru te teme, ali takođe i postavljaju pitanja [6]. Ipak, zbog straha od grešaka, anksioznosti i nedostatka samopouzdanja, studenti nerado postavljaju pitanja tokom predavanja [67]. Sposobnost adekvatnog formulisanja i postavljanja pitanja predstavlja ključnu komponentu kritičkog razmišljanja i može dovesti do dubljeg razumevanja određene teme, ali i boljeg pamćenja [68]. Konverzacija sa ChatGPT-jem može pozitivno uticati na spomenute aspekte jer se pokazao efikasnim u pružanju sveobuhvatnih objašnjenja različitih kompleksnih tema koristeći jednostavan jezik [34]. Samim tim, institucije visokog obrazovanja trebalo bi da obrate pažnju na nove mogućnosti koje ove tehnologije donose i da se fokusiraju na očuvanje akademskog integriteta, dok studentima omogućavaju odgovorno korišćenje inovativnih alata veštačke inteligencije [64].

Generativna veštačka inteligencija ima potencijal da značajno utiče na obrazovanje stvaranjem naprednih sistema za podučavanje koji vode studente kroz proces učenja i pružaju prilagođene povratne informacije [6]. U okruženju za učenje zasnovanom na korišćenju računara, povratna informacija se smatra jednim od faktora koji utiču najviše na efikasnost procesa učenja [69]. U tradicionalnim školskim okruženjima, nastavnici nisu u mogućnosti da posvete pažnju posebno svakom učeniku, što onemogućava da odgovore na individualne potrebe učenika, zbog čega učenici često trpe [6]. Modeli poput ChatGPT-ja mogu služiti kao virtuelni inteligentni asistenti, pružajući studentima personalizovane odgovore i povratne informacije koje podstiču iterativno učenje [68]. Digitalni asistenti u nastavi potpomognuti veštačkom inteligencijom mogu biti programirani tako da pružaju vrlo detaljna objašnjenja na postavljena pitanja od strane studenata [6]. Različiti studenti mogu diskutovati o istoj temi sa četbotom na različite načine, a njegova sposobnost da prilagodi odgovore specifičnim informacijama u odnosu na prethodnu konverzaciju utiče na personalizaciju učenja [54]. Brze i učestale povratne informacije omogućavaju studentima da unaprede razumevanje i veštine prema sopstvenom tempu, što je u skladu s pristupom temeljnog učenja, gde se fokus stavlja na izgradnju čvrstih osnovnih znanja pre prelaska na kompleksniji sadržaj [68].

Takođe, jedan od glavnih faktora koji ometaju učenje je nedostatak aktivnog učešća odnosno angažovanja studenata u procesu učenja [68]. ChatGPT može biti vrlo koristan za podsticanje angažovanja studenata zbog mogućnosti vođenja konverzacije kroz koju se simulira davanje povratnih informacija, kao i društvenu prisutnost koju student inače dobija od samih nastavnika u većoj ili manjoj meri. Važno je, međutim, napomenuti da, iako virtuelni inteligentni sistem za podučavanje može pružiti određenu pomoć, on nije zamena za prave nastavnike i možda neće biti u mogućnosti da pruži isti nivo personalizovane podrške [68].

U obrazovanju inženjera ChatGPT može davati odgovore na pitanja vezana za teorijski deo gradiva, ali i za praktični kao što je npr. generisanje koda [68]. Studenti mogu kroz konverzaciju dolaziti do potrebnih znanja i razrešavati sve nedoumice na tom putu. Mogu interagovati sa velikim jezičkim modelima kroz postavljanje pitanja na bilo koju temu i time produbljivati svoje razumevanje određenih koncepata. Evaluacije tačnosti

informacija i kvaliteta diskusija generisanih od strane ChatGPT-ja smatraju se zadovoljavajućim [34]. Ovo često nije bio slučaj sa prethodnim implementacijama konverzionih agenata u obrazovanju zbog tehničkih ograničenja. Takođe, virtualni edukatori potpomognuti veštačkom inteligencijom mogu proaktivno postavljati pitanja studentima što može dovesti do razjašnjavanja nejasnoća, kao i jačanja razumevanja [6].

Tokom analize literature, posebno se izdvojilo programiranje kao oblast u kojoj se pokazalo da veliki jezički modeli imaju visok stepen primenljivosti sa dobrim performansama. Studenti mogu da iskoriste sposobnost velikih jezičkih modela za kreiranje, analizu i objašnjavanje programskog koda. Programeri, ali i studenti, mogu ga koristiti za rešavanje programerskih problema i pronalaženje i ispravljanje grešaka u kodu [29]. ChatGPT i slični alati, pored specijalizovanih agenata za programiranje baziranih na velikim jezičkim modelima, kao sledeći korak u iskustvu pretrage, bi mogli zameniti druge pretraživače (npr. Google) ili forume (npr. StackOverflow) koji se često koriste u programerskom okruženju. Sposobnost prevođenja koncepata iz engleskog jezika na programski jezik [29] može unaprediti iskustvo programera i skratiti vreme za rešavanje svakodnevnih problema. Na ovaj način se smanjuje vreme potrebno za procenu rezultata pretrage i njegovo prilagođavanje. Za razliku od ostalih izlaza ChatGPT-ja, kod se može testirati i verifikovati direktno u okviru samog okruženja za programiranje. Svakako treba obratiti pažnju da li dobijen kod zadovoljava zahteve, što iskusnim programerima ne bi trebalo da predstavlja problem, a značajno im može skratiti vreme u nekim slučajevima. Vremenom su razvijeni i specijalizovani alati za programiranje zasnovani na velikim jezičkim modelima poput OpenAI Codex, Anthropic Claude Code, GitHub Copilot i dr.

Studija [68] navodi da s obzirom na to da je jedan od glavnih načina interakcije sa velikim jezičkim modelima kroz postavljanje pitanja, studenti koji umeju dobro da formulišu pitanja će biti u prednosti u odnosu na ostale, jer kvalitet odgovora ovih sistema zavisi od kvaliteta postavljenih pitanja. U eri ChatGPT-ja i drugih aplikacija velikih jezičkih modela, važnije je nego ikad da pojedinci budu sposobni da postavljaju pitanja, samostalno traže informacije i vrše njihovu procenu. Kako su autori naveli, da bi to postigli moraju biti snalažljivi i razviti visok nivo kritičkog razmišljanja radi uspeha u kreiranju strategija za pronalaženje informacija korišćenjem različitih alata. Samim tim korišćenje ovih alata na adekvatan način tokom školovanja i suočavanje sa njihovim prednostima i manama može pomoći na ovom putu.

5.5 Digitalna pismenost o velikim jezičkim modelima

Iako ChatGPT pokazuje prosečne rezultate u generisanju istinitih odgovora [19], koji se kroz vreme sa pojavom novih modela i unapređenja ovih alata poboljšavaju, nastavnici i učenici moraju biti svesni da se ne mogu u potpunosti osloniti na njegove odgovore. Algoritmi veštačke inteligencije nisu savršeni i podložni su povremenim greškama [4], [34]. Korisnici ChatGPT-ja i sličnih alata treba da budu svesni njihovih nedostataka. Ovi alati ponekad mogu generisati halucinacije odnosno odgovore koji zvuče uverljivo, ali su netačni ili besmisleni [29]. ChatGPT nije sposoban da proceni kvalitet i pouzdanost izvora [64]. Ako je tema o kojoj se raspravlja slabo dokumentovana, može proizvesti nepouzdanu informaciju [64]. Takođe, ChatGPT može generisati netačne odgovore ukoliko su podaci koji mu se dostavljaju pogrešni [3]. Učenici tada mogu prihvatiti takve odgovore kao tačne, što može dovesti do usvajanja pogrešnog znanja [3].

Cooper [35] je ocenio odgovore ChatGPT-ja na pitanja vezana za obrazovanje u oblasti nauke. Zaključeno je da odgovori odgovaraju glavnim konceptima istraživanja u literaturi, ali bez dokaza koji ih potkrepljuju. Edukatori i učenici ne bi trebalo da prihvataju ChatGPT odgovore olako, jer se mogu smatrati mišljenjima bez referenci [34]. Informacije bi trebalo proveravati iz pouzdanih izvora [29]. Iako će odgovori dobijeni od strane ChatGPT-ja većinu vremena biti korisni i nadmašiti potencijalne nedostatke, učesnici u obrazovanju treba da imaju to na umu.

Autori u radu [63] navode da edukatori i učenici moraju razviti kritičko razmišljanje u vezi sa sadržajem generisanim veštačkom inteligencijom, uključujući metode i izvore korišćene za stvaranje rezultata. Treba da budu svesni i potencijalnih pristrasnosti koje mogu postojati u sistemu.

Unapređeni i olakšani način pristupa znanju kroz konverzaciju smanjuje potrebu za kritičkim razmišljanjem. Zbog ove promene postavlja se pitanje da li će razvoj kritičkog razmišljanja postati ključan deo obrazovanja [4]. S obzirom na to da korišćenje ChatGPT-ja utiče negativno na kritičko razmišljanje učenika, smernice za ukazivanje njegove važnosti prilikom korišćenja ChatGPT-ja predstavljaju važan aspekt budućih istraživanja [34].

S druge strane, ChatGPT u velikoj meri zavisi od unosa koje dobija [30]. Kvalitet pitanja koja korisnici formulišu utiče na odgovore ChatGPT-ja [34]. Da bi se dobili odgovarajući odgovori, pitanja moraju biti precizna.

Na osnovu svega navedenog, zaključak je da je neophodno razviti pismenost u oblasti veštačke inteligencije i stvoriti uslove za razvoj kompetencija pomenutih strana u obrazovanju [33]. Kako bi bili pripremljeni za tržište rada koje već svakodnevno koristi veštačku inteligenciju, učenici moraju razviti novi set veština [63]. Sa širenjem veštačke inteligencije u svim oblastima, priroda zapošljavanja će se promeniti [29]. Zadatak edukatora je da obezbede odgovorno korišćenje konverzionih agenata zasnovanih na velikim jezičkim modelima u procesu učenja kako bi ostali u koraku sa neizbežnim trendovima. Ovo bi se moglo ostvariti kroz obuke za nastavno osoblje i studente [22].

Još jedno važno pitanje koje se nameće jeste bezbednost njegove upotrebe u obrazovanju. Kao što je navedeno u radu [19], ChatGPT je dobar u odbijanju da generiše odgovor na nebezbedna pitanja, odnosno u generisanju bezbednih odgovora, ali treba imati na umu da ponekad može reagovati na instrukcije koje mogu prouzrokovati štetu ili pokazati diskriminatorno ponašanje [29]. Privatnost je još jedan značajan aspekt na koji treba obratiti pažnju prilikom razmatranja njegove primene [34].

5.6 Prihvatanje tehnologije u obrazovanju

Sposobnost ChatGPT-ja da generiše, analizira i objašnjava kompjuterski kod, pronalazi i ispravlja greške, kao i da rešava probleme u programiranju [70], kao i da razjašnjava teoriju i pomaže u izradi zadataka na predmetima vezanim za računarske nauke, može biti iskorišćena od strane IT studenata za postizanje obrazovnih ciljeva. Samim tim, potrebna su dodatna istraživanja potencijala integracije tehnologija zasnovanih na

velikim jezičkim modelima u edukaciju IT stručnjaka [32]. Fokus budućih istraživanja treba biti na dizajnu odgovornih edukativnih četbotova [34].

Samim tim, istraživanje faktora koji utiču na zadovoljstvo i prihvatanje ChatGPT-ja u obrazovne svrhe može biti korisno, posebno za institucije visokog obrazovanja, ali i za kompanije u oblasti edukativnih tehnologija. U nastavku će biti opisano početno istraživanje ove disertacije o ovim faktorima, koje je prethodilo razvoju objedinjenog modela za razvoj edukativne platforme sa konverzionim agentom zasnovanim na velikom jezičkom modelu.

Model prihvatanja tehnologije

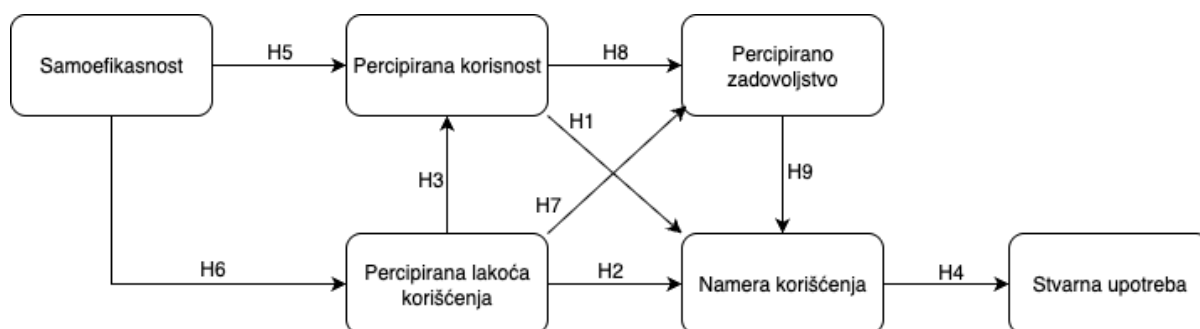
Model prihvatanja tehnologije (eng. *Technology Acceptance Model* - TAM), koji se fokusira na percipiranu lakoću korišćenja (eng. *Perceived Ease of Use*) i percipiranu korisnost (eng. *Perceived Usefulness*) kao primarne determinante prihvatanja tehnologije, primenjuje se već dugi niz godina u različitim oblastima, uz brojne modifikacije i proširenja. Prihvatanje edukativne tehnologije predstavlja jednu od oblasti u kojoj je TAM takođe pronašao svoju primenu [71]. Više autora istraživalo je faktore prihvatanja tehnologije u e-učenju [72], [73], [74], mobilnom učenju [75], [76], [77], hibridnim okruženjima za učenje [78] i sistemima za upravljanje učenjem [79]. Pored TAM-a, ističe se objedinjena teorija prihvatanja i upotrebe tehnologije (eng. *Unified Theory of Acceptance and Use of Technology* - UTAUT) koja proširuje njegov opseg uvođenjem dodatnih determinanti poput očekivanog truda, očekivanih performansi, društvenog uticaja i olakšavajućih uslova [80], pružajući tako sveobuhvatno razumevanje faktora koji utiču na upotrebu tehnologije. Ova dva modela predstavljaju najčešće korišćene modele prihvatanja tehnologije [81].

Uzimajući u obzir dokazanu efikasnost TAM modela u istraživanju faktora koji utiču na prihvatanje tehnologije, njegovu proširivost, i s obzirom na to da su prethodne studije [82], [83] pokazale da je zadovoljstvo takođe pod uticajem faktora integrisanih u originalni TAM, ovako prošireni model može biti primenjen kao konceptualni okvir za ispitivanje determinanti koje utiču na usvajanje ChatGPT-ja kao obrazovne tehnologije. Ovo poglavlje je bazirano na objavljenom članku od strane autora teze, *Evaluating the Usage of ChatGPT for Educational Purposes Through the Extended Technological Acceptance Model*, u časopisu *International Journal of Engineering Education* (2025, tom 41, izdanje 2) [84]. Ključni delovi rada su prevedeni i prilagođeni za uključivanje u tezu.

Hipoteze iz originalnog TAM modela koje su testirane su: (H1) Percipirana korisnost (PK) pozitivno utiče na nameru korišćenja (NK); (H2) Percipirana lakoća korišćenja (PLK) pozitivno utiče na nameru korišćenja; (H3) Percipirana lakoća korišćenja pozitivno utiče na percipiranu korisnost (PK); i (H4) Namera korišćenja pozitivno utiče na stvarnu upotrebu (SU). U okviru ovog modela, PK se definiše kao stepen u kome osoba veruje da bi korišćenje određenog sistema poboljšalo njen radni učinak, a PLK kao stepen u kome osoba veruje da bi korišćenje određenog sistema bilo bez napora [85].

Proširivanjem TAM modela faktorima samoeфикаsnosti u hibridnom sistemu e-učenja (SE) i percipiranim zadovoljstvom (PZ) postavljamo još 5 hipoteza: (H5) SE pozitivno utiče na PK, (H6) SE pozitivno utiče na PLK, (H7) PLK pozitivno utiče na PZ; (H8) PK

pozitivno utiče na PZ i (H9) PZ pozitivno utiče na NK. Samoefikasnost u okviru sistema za e-učenje, definiše se kao poverenje u sopstvenu sposobnost obavljanja određenih zadataka učenja koristeći dati sistem [86], [87] i predstavlja jedno od najčešćih proširenja TAM modela u ovom kontekstu [88]. Faktor koji predstavlja samoefikasnost sistema biće korišćen u kontekstu korišćenja ChatGPT-ja u obrazovne svrhe kao dela hibridnog okruženja za učenje. Takođe, u kontekstu ove studije, percipirano zadovoljstvo dolazi od bihevioralnih uverenja i stavova studenata koji proističu kao rezultat benefita koje dobijaju na osnovu korišćenja ChatGPT-ja u obrazovne svrhe kao dela spomenutog okruženja. Detaljniji opis razvoja modela koji je predstavljen na slici 1 i poređenje sa drugim istraživanjima i razvijenim i testiranim modelima u ovom kontekstu je prikazan u referenciranom radu na početku ovog poglavlja.



Slika 1 – Proširen TAM teorijski model

Metoda istraživanja

Ova studija je sprovedena na Univerzitetu u Beogradu, Fakultetu organizacionih nauka, u prvom semestru akademske 2023/2024. godine. Učesnici istraživanja koji su popunili anketu bili su studenti treće godine modula Informacioni sistemi i tehnologije. Upitnik je sproveden na srpskom jeziku i sastojao se iz dva dela, prvi koji je prikupljao demografske podatke i informacije o prethodnom korišćenju ovog alata, i drugi koji je sadržao 19 stavki vezanih za spomenute faktore proširenog TAM modela koje su nastale na osnovu prethodno validiranih upitnika [78], [89], [90]. Korišćena je Likertova sedmostepena skala za ocenu ovih stavki.

Kriterijum za popunjavanje ankete predstavljao je korišćenje ChatGPT alata u obrazovne svrhe, bilo to u formalnom ili neformalnom obrazovanju. Od 215 studenata koji su dobrovoljno pristupili anketi, 200 je ispunilo pomenute kriterijume i uspešno poslalo odgovore. Anketa je distribuirana studentima viših godina, pretpostavljajući da su do sada imali više iskustva, prilika i znanja da adekvatno koriste ChatGPT. Većina učesnika studije su bile žene (60,5%), što nije primećeno kao pristrasnost uzorka. U Srbiji i na Univerzitetu u Beogradu, žene su zastupljenije u studentskoj populaciji. Što se tiče trajanja korišćenja ChatGPT-ja u obrazovne svrhe, 41,5% učesnika je koristilo ChatGPT duže od šest meseci, dok ga 50,0% koristi od jednog do šest meseci. U trenutku sprovođenja ove ankete ChatGPT je bio dostupan oko godinu dana. Skoro svi učesnici su koristili besplatnu verziju ChatGPT 3.5. Samo jedan učesnik je koristio verziju ChatGPT 4 koja se u trenutku istraživanja plaćala. Može se zapaziti da je ovaj alat u većini slučajeva, 84,5%, korišćen u formalnom obrazovnom kontekstu, dok je u 15,5% slučajeva korišćen u neformalnom obrazovanju. Na osnovu predstavljenog opisa uzorka,

može se zaključiti da su ispitanici upoznati sa ChatGPT-jem, da su skoro svi koristili besplatnu verziju ChatGPT 3.5, i da su je koristili u formalnom obrazovnom kontekstu. U tabeli 1 je prikazan opis uzorka.

Tabela 1 - Karakteristike uzorka i prethodno iskustvo u korišćenju ChatGPT-ja

Faktor	N	%
Pol		
Ženski	121	60.5
Muški	79	39.5
Period korišćenja		
manje od mesec dana (nizak)	17	8.5
od jednog do 6 meseci (srednji)	100	50
više od 6 meseci (visok)	83	41.5
ChatGPT verzija		
ChatGPT 3.5	199	99.5
ChatGPT 4	1	0.5
Kontekst korišćenja		
Formalno obrazovanje	169	84.5
Neformalno obrazovanje	31	15.5

Ispitivanje faktora koji utiču na prihvatanje ChatGPT alata u kontekstu obrazovanja sprovedeno je primenom modelovanja strukturnih jednačina (eng. *Structural Equation Modeling* – SEM) radi potvrde konceptualnog modela, odnosno navedenih hipoteza, korišćenjem R studija za odgovarajuće analize. SEM analiza je multivarijantna tehnika statističke analize koja kombinuje faktorsku analizu i višestruku linearnu regresiju, i omogućava kreiranje latentnih varijabli i istraživanje njihovih odnosa. Pored toga, ispitane su i hipoteze o postojanju razlike u percepciji o korišćenju ChatGPT-ja kod studenata IT-ja na osnovu pola, na osnovu njegovog korišćenja za potrebe formalnog ili neformalnog obrazovanja i na osnovu vremenskog perioda korišćenja kroz *Statistical Package for Social Science* (SPSS) verzije 23 za sprovođenje potrebnih analiza.

Rezultati

U okviru analize deskriptivne statistike, izračunate su srednje vrednosti za svaki od šest konstrukta modela i izvršena je analiza dobijenih vrednosti. Tabela 2 sadrži deskriptivnu statistiku, uključujući srednju vrednost, standardnu devijaciju i procenat studenata koji imaju pozitivno mišljenje o tom faktoru. Za pozitivno mišljenje o faktoru

je uzeta vrednost iznad 4, što predstavlja srednju vrednost na sedmostepenoj Likertovoj skali.

Rezultati pokazuju da prosečan skor za svaki konstrukt prevazilazi srednju vrednost, što ukazuje na pozitivnu percepciju korišćenja ChatGPT-ja u ovom kontekstu. Konstrukt sa najnižim skorom je SU, sa srednjom vrednošću od 4,203, i 41% ispitanika sa ocenom iznad 4,0. Sa druge strane, najviši skor zabeležen je za NK sa srednjom vrednošću od 6,375, i 93,5% ispitanika sa ocenom iznad 4,0. Sledeća skala sa visokim skorom je samoeфикаsnost ChatGPT-ja, sa srednjom vrednošću od 6,215 i 94% ispitanika sa ocenom iznad srednje vrednosti skale. Posmatrajući standardne devijacije, konstrukt SU pokazuje najveću varijabilnost.

Kolmogorov-Smirnov test je korišćen za procenu normalnosti prosečnih vrednosti konstrukta. Rezultati su takođe prikazani u tabeli 2. Dati konstrukti ne prate normalnu raspodelu, pa su neparametarski testovi, Mann-Whitney U test za postojanje razlike u percepciji na osnovu pola i konteksta obrazovanja i Kruskal-Wallis za period korišćenja, sprovedeni u sledećim koracima.

Tabela 2 - Deskriptivna statistika srednjih vrednosti konstrukata

Konstrukt	Broj stavki	Srednja vrednost	Standardna devijacija	% pozitivnih ocena	K-S	p-vrednost
PK	4	5,570	0,982	89,5%	0,102	<0,001
PLK	4	6,160	0,847	96,5%	0,161	<0,001
NK	2	6,375	0,910	93,5%	0,309	<0,001
SU	3	4,203	1,606	41,0%	0,104	<0,001
SE	3	6,215	0,935	94,0%	0,201	<0,001
PZ	3	5,270	1,170	79,0%	0,154	<0,001

Rezultati Mann-Whitney U testa koji je korišćen da bi se utvrdilo da li postoje razlike u percepciji o ChatGPT alatu na osnovu pola, kao i na osnovu konteksta upotrebe za formalno ili neformalno obrazovanje za šest konstrukta modela su prikazani u tabeli 3. Prema dobijenim p-vrednostima i srednjim rangovima ($p > 0,05$), možemo zaključiti da se percepcija upotrebe ChatGPT-ja prema spomenutim kriterijumima ne razlikuje za svih šest konstrukata.

Tabela 3 - Rezultati Mann-Whitney U testa za razlike u percepciji na osnovu pola i konteksta

Konstrukt	Pol	Srednji rang	Z	p-vrednost	Kontekst	Srednji rang	Z	p-vrednost
PK	Muški	103,63	-0,621	0,535	Formalni	103,15	-	0,129
	Ženski	98,45			Neformalni	86,05		
PLK	Muški	91,66	-1,779	0,075	Formalni	101,08	-	0,737
	Ženski	106,27			Neformalni	97,35		
NK	Muški	97,05	-0,751	0,453	Formalni	102,70	-	0,167
	Ženski	102,75			Neformalni	88,52		

SU	Muški	104,15	-0,721	0,471		Formalni	101,99	-	0,396
	Ženski	98,12				Neformalni	92,40		
SE	Muški	103,56	-0,629	0,529		Formalni	103,23	-	0,106
	Ženski	98,50				Neformalni	85,63		
PZ	Muški	104,32	-0,761	0,446		Formalni	103,31	-	0,105
	Ženski	98,01				Neformalni	85,16		

Radi istraživanja razlike o percepciji korišćenja ovog alata na osnovu perioda upotrebe, sproveden je Kruskal-Wallis H test. Rezultati su pokazali da postoji značajna razlika između percepcija o ovom alatu za sve dimenzije na osnovu trajanja korišćenja ChatGPT-ja u obrazovne svrhe (tabela 4). Srednji rangovi pokazuju da ako je iskustvo korišćenja mereno u mesecima veće, percepcija o različitim aspektima korišćenja ChatGPT-ja u obrazovanju biće veća.

Radi analize razlika u nivoima iskustva korišćenja, sproveden je Mann-Whitney U test između svih parova faktora (nisko/srednje, nisko/visoko, srednje/visoko). Rezultati su takođe prikazani u tabeli 4. Što se tiče umerenog i visokog iskustva, postoji značajna razlika između vrednosti svih faktora. Što su studenti duže koristili ChatGPT, to je veća pozitivna percepcija upotrebe među svim faktorima. Kada su u pitanju nisko i srednje iskustvo korišćenja i nisko i visoko iskustvo, postoji značajna razlika za PK, SE i SU, a nema značajne razlike za PZ i PLK. Za NK konstrukt, postoji razlika između studenata sa niskim i visokim iskustvom, i nema razlike između studenata sa niskim i umerenim iskustvom.

Tabela 4 - Rezultati Kruskal-Wallis i Mann-Whitney testova za razlike u percepciji na osnovu perioda korišćenja

Skala	Period korišćenja	Srednji rang	KW H	p-vrednost	MW U
PK	Nizak	47,35	35,174	<0,001	Nizak < Srednji Nizak < Visok Srednji < Visok
	Srednji	88,35			
	Visok	126,03			
PLK	Nizak	96,32	9,069	0,011	Srednji < Visok
	Srednji	89,44			
	Visok	114,68			
NK	Nizak	69,82	12,801	<0,001	Nizak < Visok Srednji < Visok
	Srednji	94,36			
	Visok	114,18			
SU	Nizak	57,82	32,643	<0,001	Nizak < Srednji Nizak < Visok Srednji < Visok
	Srednji	86,02			
	Visok	126,69			
SE	Nizak	51,97	19,565	<0,001	Nizak < Srednji Nizak < Visok Srednji < Visok
	Srednji	96,27			
	Visok	115,54			
PZ	Nizak	85,29	6,655	0,036	Srednji < Visok
	Srednji	93,00			
	Visok	112,66			

Modelovanje strukturnih jednačina

Pre testiranja samog modela kroz SEM analizu, odgovarajući testovi pouzdanosti i validnosti skala, navedeni u nastavku, su sprovedeni. Radi sprovođenja analize pouzdanosti, interna konzistentnost je ispitana korišćenjem Kronbahov alfa koeficijenta. Za izvođenje ove analize korišćen je psych paket (verzija 2.4.1) u R programskom jeziku. Svi izmereni Kronbah alfa koeficijenti su veći od predloženog praga od 0,7 [91] (tabela 5). Stoga, možemo smatrati skale konzistentnim. Kompozitna pouzdanost (eng. *Composite Reliability* - CR) i prosečna ekstrahovana varijansa (eng. *Average Variance Extracted* - AVE) korišćene su za ispitivanje konvergentne validnosti. Kvadratni koren AVE je izračunat za analizu diskriminantne validnosti. Vrednosti CR i AVE su izračunate korišćenjem paketa semTools (verzija 0.5-6) i lavaan (verzija 0.6-17) u R programskom jeziku. Vrednosti praga su iznad 0,7 i 0,5 za kriterijume pouzdanosti CR i kriterijum AVE, respektivno [92]. Rezultati prikazani u tabeli 5 ukazuju da su AVE i CR za svaku od šest skala iznad definisanog praga, što ukazuje na dobru konvergentnu validnost. Što se tiče diskriminantne validnosti, rezultati pokazuju da je kvadratni koren AVE vrednosti svake skale veći od korelacija između faktora, što ukazuje na dobru diskriminantnu validnost na osnovu Fornell-Larcker kriterijuma [93]. Detaljni rezultati su prikazani u tabeli 6.

Tabela 5 - Analiza pouzdanosti i konvergentne validnosti

Faktor	PK	PLK	NK	SU	SE	PZ
Kronbahov alfa koeficijent	0,840	0,860	0,900	0,870	0,840	0,940
AVE	0,575	0,637	0,822	0,694	0,637	0,850
CR	0,843	0,869	0,899	0,874	0,837	0,943

Tabela 6 - Analiza diskriminantne validnosti

Konstrukt	PK	PLK	NK	SU	SE	PZ
PK	0,758					
PLK	0,372	0,798				
NK	0,639	0,455	0,906			
SU	0,598	0,289	0,585	0,833		
SE	0,414	0,509	0,540	0,305	0,798	
PZ	0,586	0,464	0,586	0,527	0,270	0,922

Teorijski model je testiran primenom modelovanja strukturnih jednačina. Sledeći indikatori su korišćeni za procenu modela: Hi-kvadrat statistika, *Comparative Fit Index* (CFI), *Tucker-Lewis* indeks (TLI), Koren srednje kvadratne greške aproksimacije (RMSEA) i Standardizovani kvadratni koren prosečnog kvadrata reziduala (SRMR). Vrednosti veće ili jednake 0,90 za CFI i TLI tumače se kao povoljan dokaz uklapanja modela, kao i manje ili jednake 0,08 za RMSEA i SRMR [94].

Na osnovu pomenutih pragova, ukupno uklapanje modela sugerise dobro uklapanje sa sledecim vrednostima: Hi-kvadrat(143)=265,858, CFI=0,951, TLI=0,942, RMSEA=0,066 i SRMR=0,072. Generalno, model je postigao prihvatljivo uklapanje i znacajnost, jer su sve izracunate metrike u skladu sa definisanim pragovima. Ovo ukazuje da se procenjeni koeficijenti putanje mogu interpretirati. Što se tice strukturnih rezultata, tabela 7 prikazuje konstrukt, njegove prediktore, procenjene koeficijente putanje, p-vrednost, R^2 i odluku u vezi sa svakom od devet istraživačkih hipoteza. Na osnovu ovih rezultata, svih devet hipoteza je potvrđeno.

Tabela 7 - Evaluacija strukturnog modela i odluka o hipotezama

Konstrukt	Prediktor	β	p	R^2	Hipoteza	Odluka
NK	PK	0,452	<0,001	0,541	H1	Prihvatiti
	PLK	0,184	0,008		H2	Prihvatiti
	PZ	0,255	0,001		H9	Prihvatiti
PK	PLK	0,202	0,029	0,215	H3	Prihvatiti
	SE	0,327	0,001		H5	Prihvatiti
PLK	SE	0,514	<0,001	0,264	H6	Prihvatiti
PZ	PLK	0,284	<0,001	0,406	H7	Prihvatiti
	PK	0,475	<0,001		H8	Prihvatiti
SU	NK	0,606	<0,001	0,367	H4	Prihvatiti

Diskusija

Analize su potvrdile hipoteze proširenog teorijskog TAM modela. Studenti percipiraju korišćenje ChatGPT-ja u obrazovne svrhe kao lako, što utiče i na percepciju korisnosti. Takođe, samoeфикаsnost korišćenja ovog alata je visoka, što pozitivno utiče na lakoću korišćenja i korisnost. Saznanje da je alat jednostavan i koristan utiče pozitivno na nameru korišćenja, što dalje utiče na stvarnu upotrebu. Takođe, ova dva faktora, PLK i PK, pozitivno utiču na percipirano zadovoljstvo, što ima uticaj na nameru korišćenja. To implicira da će studenti istrajati u korišćenju ChatGPT-ja kao pomoć za zadatke za koje smatraju da može da im olakša učenje, sa pozitivnom namerom korišćenja.

Takođe, studija je pokazala da ne postoji značajan uticaj pola i vrste obrazovanja (formalno ili neformalno) na nivo prihvatanja i percipiranog zadovoljstva ChatGPT-jem. S druge strane, period upotrebe utiče na korišćenje ChatGPT-ja u obrazovne svrhe. Što se ChatGPT duže koristi, to je percepcija njegove upotrebe pozitivnija. Ovo bi moglo da znači da studenti postepeno uče kako da optimizuju svoje promptove i identifikuju zadatke u kojima im ChatGPT može pomoći u kontekstu obrazovanja.

Rad [95] istraživao je usvajanje ChatGPT-ja za aktivno učenje na osnovnim studijama u Hong Kongu. Utvrđeno je da je PK jak prediktor NK, ali PLK nije imao značajnu vezu sa NK. To je takođe slučaj sa [96], studijom sprovedenom o korišćenju ChatGPT-ja za izvršavanje zadataka na MBA (eng. *Master of Business Administration - MBA*) programu. Značajna veza između PLK i NK nije pronađena među ispitanim studentima. Rad [97] koji je istraživao usvajanje ChatGPT-ja u obrazovne svrhe pokazuje da PK utiče na stavove korišćenja koji utiču na NK, ali PLK ne utiče. Takođe, studija [98] sprovedena o neformalnom digitalnom učenju engleskog jezika uz pomoć ChatGPT-ja zaključila je da

PLK ne uspeva da predvidi stavove učenika direktno, ali može ostaviti uticaj na stavove kroz medijator PK. Autori studije [95] su zaključili da na osnovu ovih nalaza, programeri ne bi trebalo da se fokusiraju na poboljšanje jednostavnosti korišćenja, već na funkcionalnosti i poboljšanje tačnosti i korisnosti odgovora kako bi se nosili sa mogućim halucinacijama koje ChatGPT može da proizvede. Uprkos tome što studenti veruju da je ChatGPT pouzdan i bezbedan alat za učenje i obrazovanje [97], potrebno je podizanje svesti o neophodnosti kritičke procene i validacije generisanog rezultata među studentima.

Ovo nije u skladu sa našim rezultatima, koji su potvrdili značajnu vezu između PLK i NK, a takođe i sa prethodnom literaturom o prihvatanju tehnologije uopšteno, a i u kontekstu četbotova u obrazovanju [99], [100], [101]. Ovo može imati veze sa činjenicom da je naše istraživanje fokusirano na primenu ChatGPT-ja za postizanje obrazovnih ciljeva uopšteno, uključujući učenje, izradu zadataka i pripremu za ispite, a nije se fokusiralo na određeni deo obrazovnog procesa kao što su istraživački projekat završne godine i kratki radovi MBA studenata [96]. Samim tim ukoliko je planirana integracija ovog tipa alata u sam proces učenja, treba je pažljivo prilagoditi na osnovu korisnosti koju može pružiti za specifične slučajeve upotrebe.

Studija je potvrdila prethodne nalaze u vezi utemeljenosti konstrukata koji čine model prihvatanja tehnologije i njihovu primenu na različite tehnologije i korisnike bez obzira na interkulturalne razlike. Ovi rezultati mogu biti korisni za dizajnere platformi za e-učenje, kao i za nastavnike koji uključuju ChatGPT i slične alate u svojim nastavnim planovima i programima. Saznanje o faktorima koji utiču na usvajanje ovakvih alata u obrazovanju mogu uticati na njihove odluke radi unapređenja iskustva učenja.

5.7 Primena u onlajn i hibridnom obrazovanju

Visoko obrazovanje je prepoznalo vrednost onlajn učenja i mnoge institucije investiraju u ovu oblast [46]. Studija Banihashem i saradnika [46] pokazala je da su i studenti i nastavnici visoko motivisani u ovom formatu studiranja. S druge strane, ista studija ukazuje na povećan obim posla za nastavnike u onlajn obrazovanju. Takođe, na mnogim fakultetima grupe studenata su veoma velike, što dodatno opterećuje predavača i otežava podsticanje interakcije [50]. Masovna predavanja sve češće se kombinuju s onlajn sadržajima za koje studenti očekuju brzu i efikasnu podršku [50].

Ova vrsta učenja može biti unapređena integracijom funkcija koje nude veliki jezički modeli. Primena velikih jezičkih modela u vidu četbot aplikacija bi mogla da reši barem deo spomenutog problema, odgovarajući na studentska pitanja i pružajući instant povratne informacije u svim oblastima njihovog učenja čime bi uticala na smanjenje opterećenja predavača. Dakle, četbotovi bi mogli da reše problem ograničene komunikacije koja se javlja u sistemima onlajn učenja [7]. Četbotovi zasnovani na velikim jezičkim modelima se mogu ponašati kao virtuelni mentori koji podržavaju učenje studenata na personalizovan način pružajući im trenutnu, efikasnu i individualizovanu pomoć. Mogu pružati smernice i savete tokom procesa učenja i rešavanja problema [59]. Nastavnici se ne mogu meriti sa asistentima veštačke inteligencije u pogledu dostupnosti u bilo koje doba dana i noći [3].

Ipak, iako konverzacioni agenti imaju pozitivan uticaj na proces učenja studenata [66], autori Gan i saradnici [102] skreću pažnju da preveliko oslanjanje na personalizovanu podršku u učenju može uticati na stvaranje jaza između različitih učenika. Potrebno je postići balans između personalizacije i društvene jednakosti kako bi se osigurale jednake prilike za učenje za sve učenike [102]. Takođe, treba da budu oprezni kako ne bi postali previše zavisni od ove tehnologije, s obzirom na to da preterano oslanjanje na istu može negativno uticati na ciljeve obrazovanja i sveobuhvatan razvoj studenata u akademskom smislu [59].

Samoregulacija predstavlja značajan deo procesa onlajn učenja. Razvoj veština samoregulativnog učenja je vrlo važan posebno kod mlađih učenika [31]. ChatGPT može uticati na ovaj aspekt onlajn učenja. Može unaprediti procese poput samoevaluacije, samokontrole, kao i povećati motivaciju. Kako je navedeno u radu autora [103], okruženje za učenje utiče na samoregulaciju. Studenti mogu doživeti intelektualne i emocionalne poteškoće tokom onlajn učenja zbog odsustva direktne interakcije sa nastavnicima [103]. Međutim, kako tvrde Luan i saradnici [103], razgovori sa ChatGPT-jem usmereni na rešavanje problema mogu povećati autonomiju i nezavisnost u učenju. Zaključili su da dobijanje važnih informacija kroz dijalog povećava motivaciju za prevazilaženje anksioznosti izazvane nesigurnošću u učenju. Navode da se na ovaj način povećava interesovanje i samostalno istraživanje u cilju produbljivanja znanja, a sve to radi postizanja ciljeva samoaktualizacije.

Takođe, pristup hibridnog učenja, koji pokušava da kombinuje najbolje prakse tradicionalnog i onlajn učenja, može biti poboljšan uključivanjem konverzacionih agenata zasnovanih na velikim jezičkim modelima u nastavni proces. Svaki nastavnik može osmisliti predlog načina korišćenja ovih alata za svoje predmete. Takođe, nastavnici mogu navesti primere i objasniti studentima efikasne strategije za postavljanje pitanja, ali takođe i kako da izvrše evaluaciju dobijenih odgovora [104]. Laboratorijske vežbe i nastavni materijali poput knjiga, prezentacija i video tutorijala mogu biti unapređeni kroz ovakav alat koji može pružati podršku tokom učenja u bilo koje vreme. Nastavnici mogu savetovati studentima da koriste ove alate tokom kreiranja svojih projekata, ali tek nakon sticanja osnovnih znanja na određenu temu kroz formalne materijale [105]. Ovo može uticati na njihovu sposobnost formulisanja pitanja u toj oblasti kako bi izvukli maksimum iz ovog alata. Generativna veštačka inteligencija omogućava studentima brz pristup nastavnim materijalima koji se prilagođavaju njihovim specifičnim potrebama, pružajući ciljanu pomoć u pripremi pre časa, učenju tokom časa i pregledu gradiva nakon časa, čime se povećava entuzijazam za učenje [59].

Kasneci i saradnici [17] navode da korišćenje velikih jezičkih modela može omogućiti tačne i relevantne informacije, ali oni ne mogu zameniti kreativnost, kritičko razmišljanje i veštine rešavanja problema koje se razvijaju kroz interakciju sa nastavnicima. Zato je važno da ih nastavnici koriste kao dopunu svom predavanju, a ne kao zamenu [17]. S obzirom na to da korišćenje ChatGPT-ja isključivo kao primarnog izvora znanja može ometati proces učenja [4], potrebno je sprovesti istraživanja o partnerstvu između nastavnika i ChatGPT-ja kako bi se olakšalo samostalno učenje studenata u onlajn okruženju [34]. Može se ispitati kako saradnja između edukatora i ChatGPT-ja treba da izgleda kako bi se postigli obrazovni ciljevi, kao i kako ova saradnja utiče na rezultate učenja studenata [34]. Treba definisati odgovarajuće mere za

unapređenje iskustava i rezultata u učenju i podučavanju. Kao što je već rečeno, dizajn odgovornih obrazovnih četbotova treba biti fokus budućih istraživanja [34].

Iako je generativna veštačka inteligencija pokazala obećavajuće rezultate u drugim oblastima, njena formalna primena u obrazovanju i dalje je na niskom nivou [6]. Kako bi proces obrazovanja ostao relevantan potrebno ispitati i uskladiti postojeće prakse u radu sa studentima [30]. Potrebno je identifikovati efikasne prakse učenja i podučavanja za podsticanje učenja višeg reda [33] koje obuhvata kritičko razmišljanje i rešavanje kompleksnih problema. U radu [34] se navodi da fokus treba da bude na uravnoteženoj upotrebi ChatGPT-ja koja se kombinuje sa komunikacijom sa nastavnicima radi realizacije nastave i pružanje fidbeka studentima.

Trebalo bi razmotriti kreiranje odgovarajućeg hibridnog okruženja za učenje, koje uključuje nastavnike kao primarni izvor znanja kroz direktna predavanja, onlajn materijale i podršku velikih jezičkih modela za instant fidbek i pomoć pri savladavanju gradiva, uz odgovarajuće usmene ili praktične ispite. Samim tim sistemi za elektronsko učenje mogu integrisati ove modele kako bi poboljšali iskustvo učenja studenata. U narednom poglavlju biće opisana aktuelna istraživanja i mogućnosti integracije ovih modela u sisteme za elektronsko učenje odnosno njihova primena u kreiranju specijalizovanih rešenja za učenje uz pomoć računara.

5.8 Primena četbota zasnovanog na velikom jezičkom modelu u elektronskom učenju

Kao što je već rečeno, integracija velikih jezičkih modela može poboljšati komunikaciju sa studentima, kroz instant fidbek što direktno utiče na smanjenje opterećenja nastavnika [34]. Instant povratne informacije omogućavaju učenicima da učvrste svoje razumevanje, razumeju svoje greške i izvrše neophodne ispravke, čime se poboljšava efikasnost procesa učenja [69]. Prilagodljiva podrška putem povratnih informacija u okruženju za učenje je korisna jer većina učenika ima različite lične karakteristike, poput prethodnog znanja, napretka u učenju i preferiranih načina učenja [69].

S druge strane, treniranje velikog jezičkog modela od nule za specifične primene predstavlja složen i resursno zahtevan proces koji zahteva značajna finansijska, računarska i vremenska ulaganja [106]. Mnogi javno dostupni modeli nude mogućnost finog podešavanja prethodno obučених velikih jezičkih modela koristeći domenske podatke relevantne za određenu primenu [107]. Međutim, iako je ovaj zadatak generalno manje zahtevan od obučavanja modela specifičnog za određeni domen, fino podešavanje i dalje zahteva značajnu količinu strukturiranih podataka i resursa za sprovođenje ovog procesa [108]. Iz ovih razloga, obe ove opcije nisu optimalan izbor za manje organizacije, pojedinačne istraživače, predavače i programere. Kako bi se prevazišli ovi izazovi, održivu alternativu predstavlja korišćenje javno dostupnih aplikacionih programskih interfejsa (API) koji pružaju pristup unapred treniranim velikim jezičkim modelima po određenoj ceni izraženoj prema broju ulaznih i generisanih tokena. Dostupnost LLM API-ja, kao što je OpenAI API za njihove modele, značajno olakšava integraciju velikih jezičkih modela u postojeće ili nove aplikacije, nudeći tako pristupačnije i isplativije rešenje.

Dakle, zahvaljujući dostupnosti OpenAI API-ja, developeri mogu integrisati velike jezičke modele ove kompanije u aplikacije, što otvara nove mogućnosti u sektoru edukativnih tehnologija [29]. Na ovaj način moguće je iskoristiti vrhunske sposobnosti ovih modela bez potrebe za velikim početnim ulaganjima, jer su modeli već trenirani na ogromnim količinama podataka. Zbog izuzetne skalabilnosti, idealni su za upotrebu u aplikacijama velikog obima [29], međutim ne treba zanemariti njihovu cenu. Cene poziva API-ja su javno dostupne na zvaničnoj onlajn dokumentaciji OpenAI organizacije. Pored OpenAI API-ja, u međuvremenu su postali dostupni API-ji i drugih velikih jezičkih modela. OpenAI takođe nudi mogućnost finog podešavanja za specifične zadatke, mada ovi modeli se mogu koristiti čak i bez dodatnog prilagođavanja za pružanje odgovora studentima zbog sposobnosti pisanja o raznim temama [29]. S druge strane, Haleem i saradnici [29] sugerisali su da bi se upotreba manjih, specifičnih modela mogla razmatrati umesto generičkih generativnih velikih jezičkih modela kako bi se zadovoljile potrebe različitih organizacija.

Neke obrazovne platforme kao što su Moodle, Excercism, Udacity, KhanAcademy, već nude rešenja za korišćenje OpenAI GPT modela putem svojih sistema. Za Moodle, jedan od najpopularnijih sistema za upravljanje učenjem otvorenog koda, dostupan je dodatak pod nazivom AI Chat Block (prethodno OpenAI Chat Block) za integraciju koji nudi određene mogućnosti prilagođavanja virtuelnog asistenta. Rad [109] je demonstrirao integraciju ovog dodatka u Moodle sistem. Obrazovna platforma Exercism, koja nudi kurseve za učenje programiranja, integrisala je podršku za ChatGPT koja je dostupna samo korisnicima koji doniraju platformi. Prilikom pristupa ChatGPT podršci na platformi Exercism pojavljivalo se upozorenje: „ChatGPT je moćan alat, ali može previše olakšati rešavanje problema, čime se gubi prilika za učenje kroz suočavanje sa izazovima. Preporučujemo da ga koristite u situacijama kada ste sigurni da ne možete da nastavite dalje bez pomoći.“ Korišćenje prečica u početnoj fazi učenja programiranja može ometati razvoj kritičkog razmišljanja. Ovo naglašava važnost istraživanja u kojim situacijama i na koji način treba koristiti velike jezičke modele u obrazovne svrhe. U institucijama visokog obrazovanja, ukoliko studenti polažu ispite uživo, sigurno će postati svesni da prekomerno oslanjanje na velike jezičke modele može ometati njihov proces učenja. Takođe, autori Mittal i saradnici [6] spominju u svom radu primer kreiranja inteligentnog virtuelnog tutora od strane Udacity edukativne platforme koji koristi GPT-4 model. Ovaj napredni sistem nudi personalizovano vođenje u procesu učenja i povratne informacije prilagođene stilovima učenja pojedinačnih studenata. Pojednostavljuje složene zadatke i razjašnjava tehničku terminologiju, poboljšavajući celokupno iskustvo učenja. Takođe, KhanAcademy platforma neprofitne organizacije pod istim imenom čiji je cilj pružanje besplatnog obrazovanja je kreirala virtuelnog tutora Khanmigo, ali isti nije dostupan u Srbiji. Khanmigo vodi studente kroz proces učenja i može im postavljati individualizovana pitanja kako bi podstakao dublje učenje.

S druge strane uočava se trend istraživanja fokusiranog na razvoj specijalizovanih edukativnih četbotova zasnovanih na velikim jezičkim modelima u različitim oblastima kao što su računarske nauke, medicina, jezici itd. Neki od primera će biti opisani u nastavku.

U oblasti medicine, tekuća istraživanja se bave integracijom API-ja velikih jezičkih modela za aplikacije za simuliranje različitih uloga u ovom domenu radi poboljšanja iskustva učenja. Na primer, autori u radu [110] su razvili GPT-4 četbot dizajniran tako

da simulira usmene ispite, sa ciljem da podrži hirurške specijalizante u njihovoj pripremi za certifikacioni ispit Američkog odbora za hirurgiju. Iako je studija pokazala potencijal LLM-ova u ovom kontekstu, autori su istakli potrebu za daljim usavršavanjem i razvojem robusnih okvira za evaluaciju. Slično tome, autori u radu [111] su predstavili RasPatient Pi, virtuelni simulator pacijenta zasnovanog na velikom jezičkom modelu. Ovaj sistem koristi GPT-4o model putem OpenAI Assistants API-ja, sa instrukcijama specifičnim za određeni scenario radi generisanja realističnih medicinskih dijaloga. S druge strane, pored simulacija, autori u radu [112] su kreirali Anatbuddy četbot koji je pokazao obećavajuće rezultate u obrazovanju iz anatomije, posebno u poboljšanju tačnosti odgovora u poređenju sa ChatGPT 3.5 modelom. To je postignuto uz korišćenje *Retrieval Augmented Generation* (RAG) tehnike za dovlačenje relevantnog konteksta za upite iz javno dostupnih resursa iz anatomije koji je prosleđivan GPT 3.5 modelu. Više reči o ovoj tehnici biće u narednim poglavljljima. Međutim, dati alat još uvek nije testiran od strane studenata u stvarnim uslovima učenja.

Učenje jezika predstavlja oblast sa potencijalnom za primenu LLM-a za poboljšanje mogućnosti učenja, kao i nastavnih praksi. Konkretno, jedan od primera je studija [113] je razvila četbot asistenta za učenje engleskog vokabulara, koristeći OpenAI API za pružanje personalizovane podrške studentima. Ova podrška je uključivala prilagođena objašnjenja, interaktivne vežbe i mogućnost vežbe vođenja konverzacije. Rezultati studije ukazuju da se integracija četbota zasnovanog na LLM-u pokazala značajno korisnom i doprinela dugoročnom zadržavanju znanja.

U oblasti računarskih nauka najveći je fokus na kreiranju pomoćnika početnicima pri učenju programiranja. CodeTutor [114] je četbot aplikacija zasnovana na GPT-3.5 velikom jezičkom modelu za pomoć pri učenju na uvodnom kursu programiranja na kome se uči Python programski jezik. Aplikacija je testirana tokom jednog semestra nad 50 studenata koji su upisani na ovaj kurs i pokazala je statistički značajna poboljšanja u konačnim ocenama kod studenata koji su koristili ovaj alat kao pomoć pri učenju u odnosu na kontrolnu grupu koja nije imala alat na raspolaganju. Iako su generalno zadovoljni datom aplikacijom, specijalno za savladavanje sintakse jezika, studenti su izrazili brigu o negativnom uticaju na kritičko razmišljanje, s obzirom na to da je četbot davao i kompletna rešenja određenih zadataka ukoliko bi ih student zatražio. Rešenja opisana u nastavku, kao što su 61A Bot, CodeHelp, CodeTailor, LeetCoach, RAGMan i CS50 Duck su rešenja koja nude podršku pri učenju, ali sa ciljem da ne pružaju kompletna rešenja studentu.

61Abot [115] predstavlja interaktivnog pomoćnika za izradu domaćih zadataka za uvodni kurs programiranja u Python programskom jeziku. Kroz prilagođeni prompt za GPT-4 model, umesto davanja direktnih odgovora sa potpunim rešenjima i objašnjenjima, generisani odgovori treba da podstiču studente na rešavanje problema kroz smernice i podršku za otkrivanje grešaka. Prilikom testiranja sa preko 2.000 studenata tokom dva semestra, alat je uticao na smanjenje broja pitanja na forumu i vremena koje su studenti provodili rešavajući zadatke. Povratne informacije studenata su pokazale da je ovaj alat generalno bio korisna i podržavajuća pomoć u učenju u ovom kontekstu, čak iako smernice u nekim situacijama nisu bile od pomoći ili su otkrивale previše informacija. Slično 61Abot-u, CodeHelp [116], takođe asistira studentima u učenju programiranja tako da ne otkriva kompletna rešenja kada zatraže pomoć kroz struktuirani ulaz koji student popunjava i koji obuhvata studentski kod, grešku i

pitanje. Primenjen je na kursu računarstva i nauke o podacima na kome se obrađuju osnove Python programskog jezika i specijalizovane biblioteke za rad sa podacima. Za njegov razvoj korišćeni su OpenAI modeli GPT-3.5-turbo-0301 za generisanje glavnog odgovora, i text-davinci-003 za uklanjanje koda iz odgovora ukoliko postoji, pre slanja studentu. Broj halucinacija je na dovoljno niskom nivou da se alat smatra korisnim uprkos njima. Studenti su izrazili zadovoljstvo podrškom koju je pružao pri rešavanju grešaka, njegovom stalnom dostupnošću, i mogućnošću da mu postavljaju pitanja koja se ne bi usudili da postave predavačima, dok su predavači istakli da ovaj alat predstavlja dopunu, a ne zamenu za njihovu podršku u nastavi. CodeAid [117] predstavlja virtuelnog pomoćnika za učenje C programskog jezika kroz odgovaranje na konceptualna pitanja, pružanju pomoći u objašnjenju, ispravci i pisanju koda odnosno davanje pseudokoda sa objašnjenjima svake linije i obeležavanje koda uz predloge za ispravku. Sistem je realizovan korišćenjem React okvira za *frontend* i Express okvira za *backend* deo aplikacije, koji komunicira sa OpenAI modelom GPT-3.5-turbo radi generisanja odgovora. Autori su ga testirali kao opcioni dodatni resurs za učenje tokom dvanaestonedelnog kursa programiranja sa 700 studenata, od kojih je njih 372 koristilo ovaj alat. Alat je ažuriran na polovini kursa. Postigao je tačnost od 79%, a ocenjen je kao adekvatna pomoć u 86% slučajeva (ukupno u obe verzije alata). Važno je napomenuti da je nakon ažuriranja prve verzije CodeAid platforme kroz ažuriranje promptova i dodavanje određenih funkcionalnosti na osnovu studentskog fidbeka, tačnost (sa 74% na 87%) i korisnost (sa 84% na 91%) porasla što implicira da je neophodna adekvatna evaluacija i praćenje ovih sistema i njihovo ažuriranje radi povećanja performansi. U intervjuima sa 22 studenta istaknute su kao ključne prednosti neprestana dostupnost i sposobnost razumevanja konteksta, što ne postoji kod javno dostupnih rešenja poput ChatGPT-ja. Pojedini studenti su preferirali indirektne odgovore na svoja pitanja, pa su sami regulisali način korišćenja alata tako što su postavljali samo opšta pitanja ili izbegavali funkcije poput ispravljanja koda, kako bi podstakli samostalno učenje. CodeTailor [118], zasnovan na GPT-4 modelu, generiše Parsonsove slagalice na osnovu netačnog koda studenata, koje mogu pomoći studentu u učenju osnova programiranja. Alat je postigao 98% tačnosti u kreiranju slagalice na osnovu studentskog koda. U studiji koja je obuhvatila 18 programera početnika, većina učesnika je preferirala CodeTailor u odnosu na dobijanje kompletnih rešenja pomoću velikih jezičkih modela. LeetCode [119] je takođe studentski pomoćnik, zasnovan na Gemma 3 (27B) modelu, koji pruža podršku tokom učenja programiranja tako da ne pruža konačna rešenja studentu već kroz diskusiju i postavljanje pitanja vodi studenta u ovom procesu učenja. U implementaciji su korišćeni React i Flask okviri. Evaluacija od strane studenata je pokazala zadovoljstvo jasnoćom instrukcija koje su dobijali i generalno zadovoljstvo iskustvom korišćenja ovog alata.

Određeni specijalizovani alati za pomoć u učenju u oblasti računarskih nauka integrišu i eksterne baze iz kojih se kontekst u okviru prompta namenjenog velikom jezičkom modelu dopunjuje radi generisanja boljih odgovora u skladu sa kontekstom učenja. RAGman [120] je četbot alat razvijen u cilju davanja odgovora bez otkrivanja kompletnih rešenja na pitanja vezana za domaće zadatke na uvodnom kursu programiranja, konkretno za Python programski jezik. Koristi OpenAI API za komunikaciju sa velikim jezičkim modelom uz kombinaciju implementacije RAG tehnike uz model text-embedding-ada-002-v2 gde na osnovu samog zadatka i dovučenog konteksta iz diskusija vezanih za isti iz prethodnih godina se generiše odgovor. Klijentski deo aplikacije je implementiran kroz Next.js, a serverski deo koristeći Python

odnosno Flask okvir. Postigao je tačnost od 98% za pitanja koja upadaju u opseg pitanja kreiranog alata, a 81% za pitanja za koja nije bio namenjen što pokazuje efektivnost i u ovom slučaju. Od ukupno 455 studenata na ovom kursu, polovina studenata je koristila ovaj alat, a 78% njih je iskazalo zadovoljstvo njegovim uticajem na proces učenja. Jedna od većih zamerki, je sporost generisanja odgovora, koju su autori uspeali da reše, ali tek nakon završetka ovog kursa. MerryQuery [121] je alat kreiran za odgovaranje na studentska pitanja o materijalima sa kursa u oblasti računarskih nauka, a konkretno je testiran na kursu Uvod u MATLAB, mada autori naglašavaju da je dizajniran tako da može da pruža podršku i van oblasti programiranja. Oslanja se na GPT-4o model i RAG tehniku uz MongoDB kao rešenje za vektorsku bazu za rad sa materijalima sa kursa. Pruža studentima dva načina korišćenja: pružanje kompletnih odgovora i pružanje smernica i okvirnih odgovora. MoodleBot [122] je četbot za integraciju u Moodle sistem za upravljanje učenjem koji na osnovu materijala za učenje dostupnih u PDF formatu olakšava studentima pristup sadržaju za učenje i razjašnjavanju nejasnoća na predmetu Baze podataka i informacioni sistemi. Za generisanje odgovora oslanja se na OPENAI GPT-4 model i dodaje promptu kontekst na osnovu materijala sa predmeta kroz RAG tehniku i Weaviate vektorsku bazu podataka. Alat je postigao visoku stopu tačnosti i usklađenosti sa materijalima (88%) i naišao je na pozitivnu percepciju od strane studenata. Međutim, autori ističu da je potrebno ohrabrivati studente da kritički analiziraju dobijene odgovore i da verifikuju informacije, time naglašavajući da ulogu četbota treba posmatrati kao dodatni pomoćni alat prilikom učenja, a ne kao univerzalni izvor istine.

Takođe, CS50 Duck je alat koji koristi RAG tehniku radi smanjenja halucinacija i davanja boljih kontekstualnih odgovora, a pored nje takođe realizuje i fino podešavanje modela kao još jedne od tehnika za poboljšanje odgovora u specijalizovanim LLM alatima. CS50 Duck [123] je četbot asistent za pružanje podrške studentima u učenju programiranja na uvodnom kursu računarskih nauka na Harvardu kroz pružanje konstruktivnih povratnih informacija, smernica, i postavljanjem pitanja radi pospešivanja studentske autonomije u procesu učenja. Koristi GPT-4o model i ChromaDB vektorsku bazu podataka za sadržaj sa kursa u cilju prevazilaženja halucinacija. Radi poboljšanja performansi praćenja instrukcija od strane velikog jezičkog modela, izvršili su fino podešavanje modela kroz 50 konverzacija.

Na osnovu ovoga možemo zaključiti da postoji potencijal integrisanja velikih jezičkih modela u obrazovanje kao specijalizovanih konverzionih agenata radi pružanja adekvatne pomoći u procesu učenja. S druge strane mora se obratiti pažnja na nedeterminističku prirodu ovih alata i samim tim proces evaluacije, monitoringa i ažuriranja sistema u skladu sa njima postaje neizostavan deo ovih sistema. Uprkos potencijalnim halucinacijama čijem se smanjenju mora težiti kroz implementaciju mehanizama za njihovo smanjenje kao što su prompt inženjering, RAG tehnika i fino podešavanje [110], [124], benefiti ovih alata mogu nadmašiti potencijalne mane kojih svi učesnici u ovom kontekstu moraju biti svesni.

Nakon pregleda literature vezane za primenu velikih jezičkih modela u edukaciji, moguće je zaključiti da postoji značajan porast publikacija iz ove oblasti prethodnih godina. Razlog tome je pojava ChatGPT četbota koji je znatno uticao na povećanje pažnje usmerene ka velikim jezičkim modelima i pospešio istraživanja u ovoj oblasti, a samim tim i u oblasti njihove primene u obrazovnom kontekstu. Prilikom analize ovih radova

primećene su mogućnosti za unapređenje procesa obrazovanja primenom velikih jezičkih modela. Konkretno, elektronsko učenje, koje je sve češće zastupljeno i kao dopuna tradicionalnom obrazovanju, ima potencijal za unapređenje kroz integraciju četbota koji je zasnovan na velikim jezičkim modelima. Na osnovu svega navedenog, kao što i autori u radu [48] navode, postoji potreba za definisanjem okvira za dizajn i razvoj obrazovnih četbotova koji bi pomogao nastavnicima u kreiranju korisnih i efikasnih četbotova. Postoji prostor za razvoj metodološkog pristupa integracije velikih jezičkih modela u obrazovanju kao dopuna sistemu za elektronsko učenje. U narednom poglavlju biće fokus na identifikaciji ključnih elemenata potrebnih za modelovanje edukativne platforme koja integriše veliki jezički model u vidu konverzionog agenta, te samim tim biti i predložen model koji uključuje ulogu konverzionog agenta kao dodatnu podršku u procesu učenja prethodno kreiranih nastavnih materijala od strane predavača.

6 OBJEDINJENI MODEL EDUKATIVNE PLATFORME SA KONVERZACIONIM AGENTOM ZASNOVANIM NA VELIKOM JEZIČKOM MODELU

Na osnovu svega navedenog u prethodnim poglavljima predlažemo razvoj objedinjenog okvira (eng. *framework*) za integraciju konverzacionog agenta (čebota) zasnovanog na velikom jezičkom modelu (VJM konverzacioni agent) u edukativnu platformu odnosno sistem za elektronsko učenje, koji podrazumeva metode, modele, tehnike i alate koji bi bili korišćeni u ove svrhe. S obzirom na to da aspekti ovog sistema obuhvataju i projektovanje samog načina prezentacije materijala za učenje, primenu veštačke inteligencije za razvoj konverzacionog agenta, kao i samu realizaciju softverskog sistema sa fokusom na interakciju sa korisnicima ovog sistema, razvoj takvog okvira podrazumeva multidisciplinarni pristup i saradnju između naučnih disciplina, kao što su: pedagogija, psihologija, veštačka inteligencija, računarske nauke, interakcija čovek-računar.

Objedinjeni okvir za razvoj integrisane edukativne platforme sa konverzacionim agentom zasnovanim na velikom jezičkom modelu bi trebalo da definiše sistemski pristup razvoju ovakvih platformi, koji bi uskladio različite uloge učesnika u ovom procesu. Metoda razvoja bi trebalo da definiše proces razvoja koji će objединiti pripremu edukativnog sadržaja, dizajn konverzacionog agenta i njegovog konteksta i baze znanja, kao i dizajn i razvoj samog softverskog sistema za elektronsko učenje, radi njegove efikasne implementacije.

Osnovni ciljevi koji će biti realizovani razvojem objedinjenog okvira i biti predstavljeni u narednim poglavljima su:

- Definisanje ontologije primene konverzacionih agenata zasnovanih na velikom jezičkom modelu u edukaciji, što će omogućiti dalji razvoj oblasti i olakšati komunikaciju između svih učesnika u procesu istraživanja i razvoja edukativnih platformi koje ih integrišu;
- Definisanje modela za razvoj edukativne platforme sa konverzacionim agentom zasnovanim na velikom jezičkom modelu, baziranih na standardnim modelima za razvoj softvera, ali proširenim semantikom potrebnom za modelovanje specifičnosti ove integrisane platforme;
- Definisanje metodologije razvoja edukativne platforme koja integriše čebota zasnovanog na velikom jezičkom modelu, koja će olakšati razvoj i praćenje projekata vezanih za edukativne platforme u ovoj oblasti;
- Definisanje razvojnih alata i načina njihove upotrebe.

Kao što je prikazano u petom poglavlju ovog rada, poslednjih godina se primećuje porast radova na ovu temu i uočeno je više različitih pristupa korišćenja velikih jezičkih modela u obrazovanju. Prilikom izrade modela razmatrali smo dostupne radove ove istraživačke oblasti iz različitih izvora. Poglavlje sa literaturom sadrži spisak svih korišćenih referenci u ovoj disertaciji. Tokom izrade modela smo nastojali da uvrstimo koncepte koji su identifikovani kao ključni, ali tako da rezultujući model ne postane preobiman, pa samim tim i nepraktičan.

6.1 Arhitektura vođena modelom

Arhitektura vođena modelom (eng. *Model Driven Architecture* - MDA) predstavlja pristup za strukturiranje specifikacije sistema kroz platformski nezavisne modele prilagođene specifičnom domenu u kojem se rešenje razvija [125]. Model u ovom kontekstu predstavlja određeni aspekt sistema [126]. Upotreba MDA modela i definisanje njihove strukture, semantike i notacija kroz industrijske standarde, olakšava modelovanje i razvoj kompleksnih sistema [126]. Zbog navedenog, arhitektura razvoja softverskih sistema koja je vođena modelom je upotrebljen pristup u ovoj doktorskoj disertaciji na konkretnom primeru objedinjenog modela za razvoj edukativne platforme sa integrisanim konverzacionim agentom zasnovanim na velikom jezičkom modelu.

Radi adekvatne implementacije predložene edukativne platforme, neophodno je jasno definisati koncepte koji se mogu upotrebiti za efikasno modelovanje ovakvog sistema. Kao što je već spomenuto, MDA pristup se fokusira na modele, i na taj način razdvaja poslovnu i aplikacionu logiku sistema od konkretnih tehnologija implementacije [125], [126]. Ovo omogućava definisanje konceptata koji su bliži domenu problema, a mnogo manje su zavisni od konkretne tehnologije. Modeli definisani na ovaj način mogu se naknadnim transformacijama preoblikovati u modele specifične za određenu platformu, a zatim dalje u njihove konkretne implementacije [127].

Meta Object Facility (MOF) predstavlja metamodel standard definisan od strane OMG organizacije koji služi za kreiranje metamodela [128] i predstavlja ključnu osnovu MDA pristupa [126]. Za vizuelizaciju spomenutih modela koristi se *Unified Modeling Language* (UML) jezik kao jedan od jezika zasnovanih na MOF standardu. UML predstavlja jezik za projektovanje, vizuelizaciju i dokumentaciju kompleksnih softverskih sistema [129]. Jedna od njegovih ključnih karakteristika je mogućnost proširenja kroz profile radi prilagođavanja specifičnim zahtevima i domenima [126]. Upravo ga ova karakteristika čini pogodnim za definisanje specifikacije sistema u konkretnim domenima kao što je slučaj u ovoj doktorskoj disertaciji.

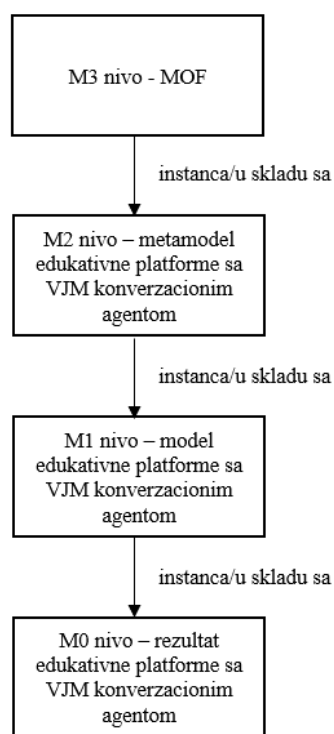
MDA paradigma organizuje modele po nivoima [130]. U nastavku će biti pojašnjena četiri nivoa navedena u samoj MOF specifikaciji, koji su uglavnom korišćeni prilikom organizacije modela gde svaki nivo predstavlja instancu nivoa iznad:

- M3 – nivo metametamodela,
- M2 – nivo metamodela,
- M1 – nivo modela i,
- M0 – nivo podataka.

M3 nivo se odnosi na OMG standard za specifikaciju nižeg metamodela kroz apstraktni jezik kao što je MOF koji sadrži kolekciju primitiva za opis metamodela na M2 nivou. M2 nivo se odnosi na metamodel koji može da se posmatra kao „apstraktni jezik” sa elementima potrebnim za definiciju modela na sledećem nivou (M1). M1 nivo se odnosi na sam model sistema koji se opisuje. Na ovom nivou mogu biti definisani platformski nezavisni model (eng. *Platform Independent Model*) i platformski specifičan model (eng.

Platform Specific Model) [131]. Na kraju, M0 nivo se odnosi na konkretne instance modela definisane na nivou M1.

Hijerarhijska organizacija koncepata i modela u više nivoa je od izuzetnog značaja s obzirom na to da proces razvoja edukativne platforme sa VJM konverzacionim agentom karakteriše visok stepen složenosti i podrazumeva modelovanje brojnih međusobno povezanih elemenata na različitim nivoima apstrakcije što je takođe bio slučaj i u drugim domenima, na primer u oblasti evaluacije multimodalnih biometrijskih sistema [127], interakcije čoveka i računara [132], kao i u oblasti edukativnih igara [133], koji su primenili sličan pristup prilikom razvoja okvira.



Slika 2 - Preslikavanje okvira za razvoj edukativne platforme koja integriše VJM konverzacionog agenta na OMG MDA pristup

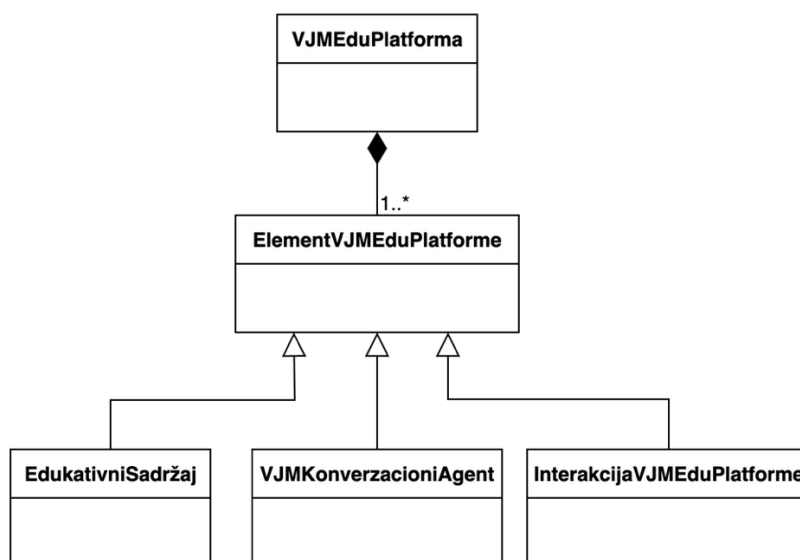
Preslikavanje okvira za razvoj edukativne platforme koja integriše VJM konverzacionog agenta na četvoronivojski OMG MDA pristup je prikazano na slici 2. Na osnovnom, M3 nivou primenjuje se MOF standard OMG organizacije, kojim je definisan apstraktni jezik namenjen za specifikaciju i opis metamodela. Na sledećem, M2 nivou je definisan metamodel edukativne platforme koja integriše VJM četbota. Ovaj metamodel predstavlja ontologiju pojmova koji su od značaja u ovoj oblasti. Definiše zajednički jezik u domenu edukativnih platformi koje integrišu VJM četbota. Daljom primenom ovako definisanog metamodela možemo kreirati modele edukativne platforme sa VJM konverzacionim agentom, što odgovara M1 nivou. Konkretne instance ovih modela, odnosno implementacije edukativne platforme se nalaze na poslednjem nivou M0 u hijerarhiji MDA modela.

6.2 Metamodeli edukativne platforme sa VJM konverzacionim agentom

6.2.1 Metamodel edukativne platforme sa VJM konverzacionim agentom

Na slici 3 predstavljen je metamodel najvišeg nivoa sa osnovnim konceptima predložene edukativne platforme. Glavni koncept predloženog metamodela predstavlja element edukativne platforme zasnovane na velikom jezičkom modelu (*ElementVJMEduPlatforme*) koji služi kao osnova za definisanje ostalih koncepata u ovom kontekstu. Tako definisan koncept generalizuje elemente *EdukativniSadržaj*, *VJMKonverzacioniAgent*, kao i *InterakcijaVJMEduPlatforme*.

Ključni elementi predloženog modela edukativne platforme koja integriše konverzacionog agenta zasnovanog na velikom jezičkom modelu, predstavljaju materijal za učenje i konverzacioni agent. U predloženom modelu četbot treba da služi kao dopuna edukativne platforme i dodatna pomoć u procesu učenja [116], [122]. Samim tim, nastavnici su zaduženi za pripremu materijala za učenje. Dakle, element *EdukativniSadržaj* se odnosi na materijal za učenje kreiran od strane nastavnika koji će preko platforme biti dostupan studentima. Studenti su kroz materijal kreiran od strane eksperata vođeni kroz određenu oblast. S druge strane, radi boljeg savladavanja gradiva studentima će biti dostupan četbot zasnovan na velikom jezičkom modelu koji je predstavljen kroz *VJMKonverzacioniAgent* element. Kroz element *InterakcijaVJMEduPlatforme* je modelirano interagovanje studenata sa datom platformom na visokom nivou apstrakcije.

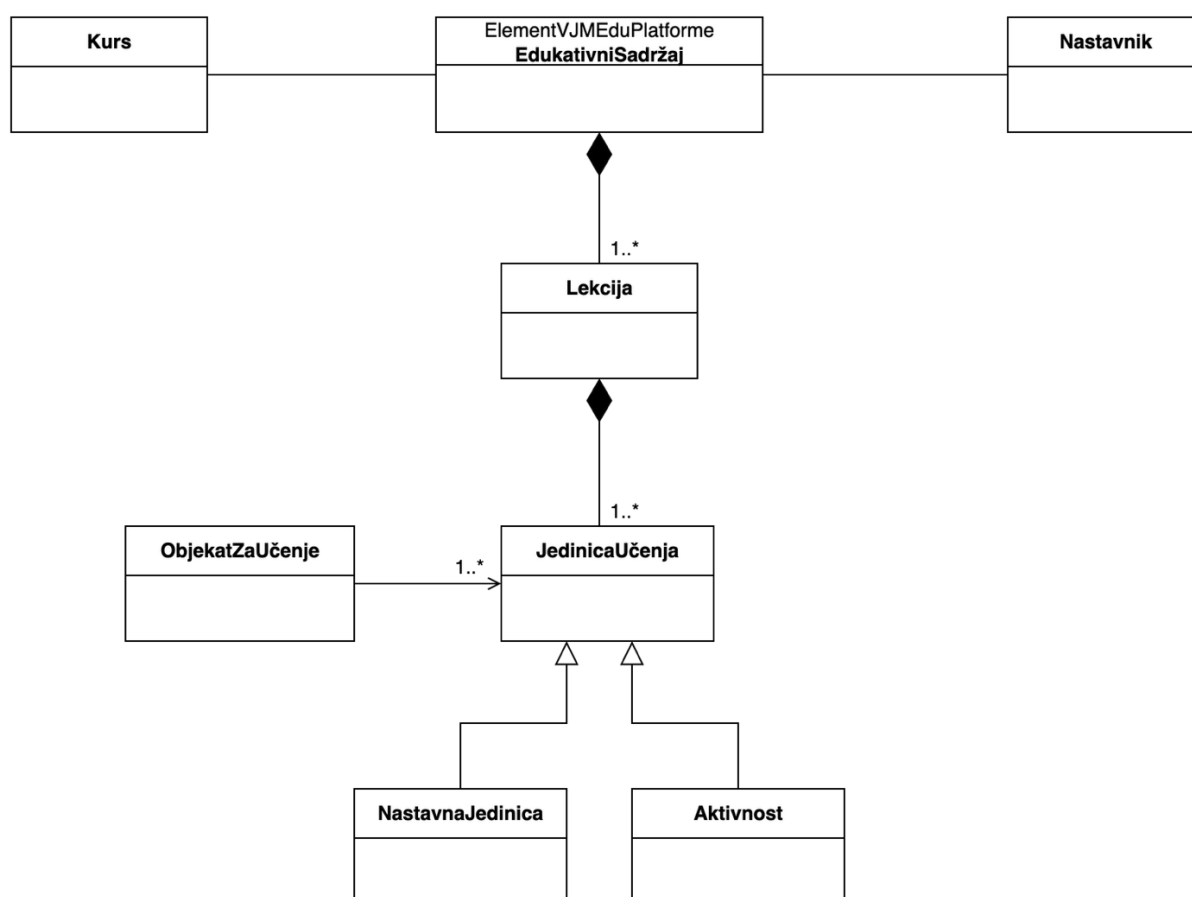


Slika 3 - Osnovni koncepti edukativne platforme sa VJM konverzacionim agentom

6.2.2 Metamodel edukativnog sadržaja

Kao što smo već spomenuli, *EdukativniSadržaj* predstavlja jedan od ključnih elemenata predloženog modela. Na slici 4 možemo videti predloženi metamodel ovog elementa.

Edukativni sadržaj predstavlja osnovu platforme namenjene učenju. Za način dostavljanja edukativnog sadržaja kreiranog od strane nastavnika odnosno eksperata u oblasti odnosno za neki kurs, izabran je prilagođen pristup mikroučenja. Ovaj pristup je poslednjih godina postigao obećavajuće rezultate kao efektivna strategija elektronskog učenja [134]. Prema njemu, sadržaj lekcija se organizuje kroz jedinice mikroučenja koje pružaju koncizan edukativan sadržaj kroz koji se student usmerava u procesu učenja i koji treba savladati za kratko vreme [135]. Nastao je kao odgovor na promene potreba studenata, pružajući sažet, lako dostupan sadržaj kojem se može pristupiti bilo kad i bilo gde i pogodan je za integraciju sa novim tehnologijama u razvoju kao što su konverzionici agenti [135]. Sam sadržaj jedinica učenja (*JedinicaUčenja*) može obuhvatati različite objekte za učenje (*ObjekatZaUčenje*) kroz koje ustvari predstavljamo materijale za učenje. Dakle, sam nastavnik je zadužen za pripremu datih materijala za određeni kurs po jedinicama učenja koje treba da usmeravaju proces učenja u okviru lekcija. Za predstavljanje samog sadržaja jedinica učenja mogu se koristiti objekti koji su opisani u sledećem potpoglavlju.



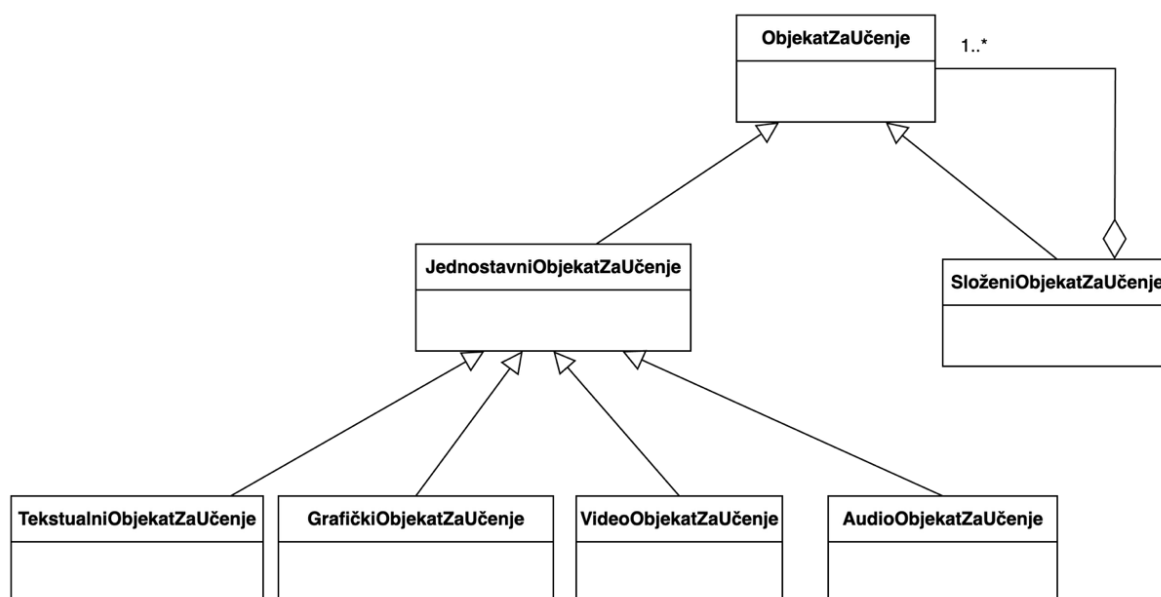
Slika 4 - Metamodel edukativnog sadržaja

Pored nastavnih jedinica (*NastavnaJedinica*) koje se odnose na gradivo koje su pripremili nastavnici i koje studenti treba da savladaju, poželjno je dati im aktivnosti (*Aktivnost*) tokom učenja kako bi mogli da testiraju svoje razumevanje. Aktivnosti mogu obuhvatati pitanja i zadatke otvorenog tipa, ili pitanja sa ponuđenim odgovorima. S obzirom da je cilj aktivnosti da podrže učenje i da studenti tokom izrade aktivnosti

mogu imati dostupne materijale nastavnih jedinica, kao i funkcionalnost četbota za razjašnjavanje nedoumica i pomoć prilikom izrade, sama aktivnost se može posmatrati kao jedinica učenja i može biti predstavljena korišćenjem objekata za učenje.

6.2.3 Metamodel objekata za učenje

Na slici 5 predstavljen je metamodel *ObjekatZaUčenje*. *ObjekatZaUčenje* kroz koje predstavljamo materijale za učenje mogu biti jednostavni objekti za učenje (*JednostavniObjekatZaUčenje*) koji mogu biti tekst (*TekstualniObjekatZaUčenje*), grafika (*GrafičkiObjekatZaUčenje*), video (*VideoObjekatZaUčenje*) ili audio snimak (*AudioObjekatZaUčenje*). Složeni objekat za učenje (*SloženiObjekatZaUčenje*) predstavlja kombinaciju prethodno navedenih objekata zajedno sa složenim.

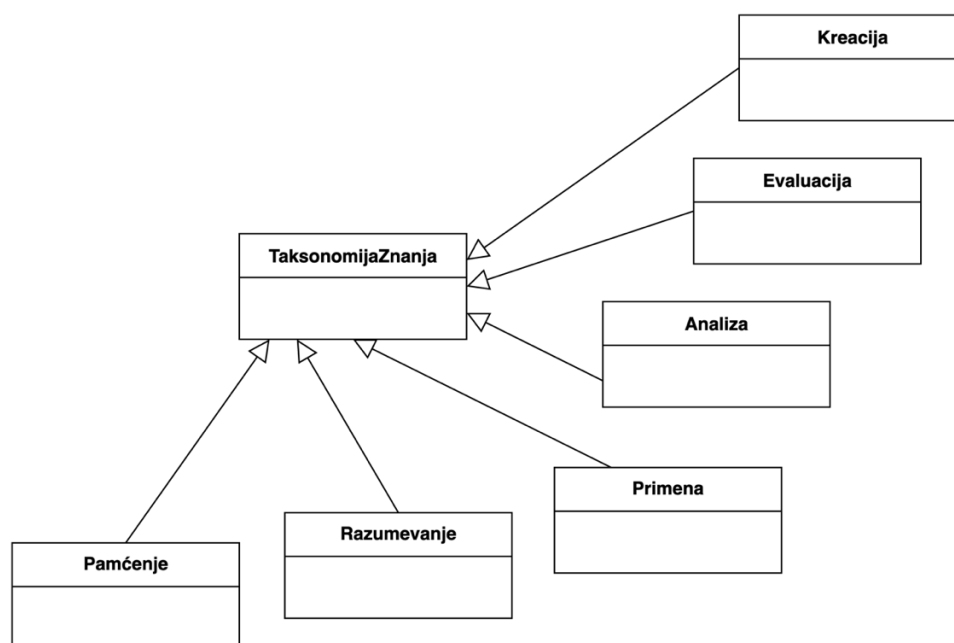


Slika 5 - Metamodel objekata za učenje [133]

Sličan metamodel možemo videti i u [133] gde je opisan metamodel objekta za učenje u kontekstu edukativnih igara za predstavljanje obrazovnih koncepata u okviru određenog domena. U spomenutoj disertaciji je takođe uvedena kao metamodel i Andresenova taksonomija znanja [136], nastala na osnovu Blumove (pamćenje, razumevanje, primena, analiza, evaluacija, stvaranje), što može biti primenjeno i u ovom kontekstu radi određivanja nivoa saznanja koje treba postići kroz objekte za učenje odnosno jedinice učenja.

U ovoj taksonomiji, pamćenje se odnosi na prepoznavanje ili prisećanje naučenog odnosno korišćenje memorije radi produkovanja ili dovlačenja definicija, činjenica, nabiranja i sl. [137]. Razumevanje se odnosi na stvaranje značenja odnosno aktivnosti koje obuhvataju interpretaciju, davanje primera, klasifikaciju, sažimanje, zaključivanje, poređenje ili objašnjavanje. Primena se odnosi na situacije u kojima se naučeno koristi kroz simulacije, prezentacije, modele. Analiza se odnosi na razlaganje koncepata na delove i zaključivanje kako su ovi delovi povezani između sebe ili u odnosu na čitavu strukturu kroz razlikovanje, organizaciju, dodeljivanje atributa što može biti postignuto

kreiranjem anketa, grafikona, dijagrama itd. Evaluacija se odnosi na proces donošenja odluka na osnovu određenih kriterijuma i standarda kroz proveru i kritiku što može rezultovati preporukama i izveštajima. U revidiranoj Blumovoj taksonomiji evaluacija je navedena pre stvaranja jer često predstavlja neophodan deo koji prethodi stvaranju. Stvaranje se odnosi na spajanje elemenata radi formiranja koherentne ili funkcionalne celine koja predstavlja novinu. Na slici 6 je prikazan metamodel taksonomije znanja.



Slika 6 - Metamodel taksonomije znanja [133]

6.2.4 Metamodel VJM konverzionog agenta

Drugi ključni element predloženog modela predstavlja konverzionog agent zasnovan na velikom jezičkom modelu. Konverzionog agent zasnovan na velikom jezičkom modelu (*VJMKonverzionogAgent*) će biti dostupan studentima na korišćenje za razjašnjavanje nejasnoća i produbljivanje znanja. Ideja je da nastavnik može studentima da omogući pristup konverzionom agentu tokom prolaska kroz lekciju za čije su savladavanje nastavnici pripremili materijale odnosno edukativni sadržaj i procene da četbot može doprineti njegovom savladavanju. Njegov metamodel je predstavljen na slici 7.

Student vodi *Konverzaciju* sa konverzionim agentom u okviru koje se pamte njegove poruke i poruke odgovora četbota. Konverzacija može biti vođena za određenu lekciju ili konkretnu jedinicu učenja što je objedinjeno kao opseg konverzacije (*OpsegKonverzacije*) u predloženom metamodelu. Sam opseg konverzacije ima definisanu instrukciju (*Instrukcija*), odnosno može imati i više njih sa mehanizmom verzionisanja radi evaluacije i daljeg ažuriranja sistema radi poboljšanja performansi. Instrukcija služi da usmerava ponašanje četbota prilikom generisanja odgovora u skladu sa opsegom konverzacije za koji je definisana. Kako bi odgovori konverzionog agenta za *KorisničkiUnos* koji se odnosi na korisničku poruku, bili usklađeni sa ciljem pomoćnika tokom učenja odgovarajućeg predmeta, potrebno je definisati parametre koji će odrediti ponašanje samog četbota koje će on proslediti *VJMAPI*-ju zaduženim za

odgovora, i uloge koju četbot obavlja za određeni opseg konverzacije odnosno konkretnu lekciju ili jedinicu učenja i može biti definisana uz saradnju nastavnika i programera. Takođe, treba definisati određena ograničenja i pravila ponašanja četbota. Dakle, kroz prompt inženjering može biti definisana određena zaštitna odnosno *guardrail* tehnika kroz koje se navodi ponašanje četbota i ublažavaju određeni rizici u generisanju odgovora [140]. Međutim, treba napomenuti da se ova tehnika kombinuje i sa drugim *guardrail* tehnikama koje mogu uključivati i analizu ulaza pre generisanja, kao i analizu izlaza nakon. U nastavku je dat primer jedne instrukcije:

„Ti si asistent za kurs Računarske mreže i telekomunikacije. Tvoje glavne odgovornosti uključuju pomaganje studentima da razumeju koncepte mreža, odgovaranje na tehnička pitanja i jasno i sažeto objašnjavanje složenih tema. Treba da pružaš praktične primere i objašnjenja korak po korak za tehničke procese. Nemoj odgovarati na pitanja koja nisu vezana za kurs. Odgovaraj na istom jeziku na kom je postavljeno pitanje (engleski ili srpski). Odgovaraj u pasusima. Tvoj ton treba da bude profesionalan, ali pristupačan, tako da podstiče studente da istražuju i postavljaju pitanja. Prilagodi svoja objašnjenja nivou znanja studenta, koristeći jednostavniji jezik za početnike, a više tehničkog žargona za napredne studente.

Sažetak trenutne lekcije (vežbi) za koju treba da pomažeš studentima i odgovaraš na njihova pitanja je: ...”.

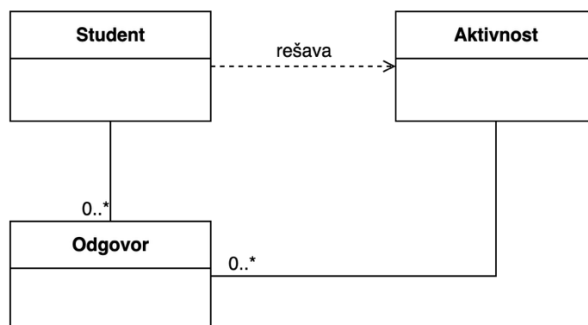
Takođe, kao što je navedeno u primeru iznad, radi poboljšanja relevantnosti odgovora u odnosu na kontekst, instrukcija može da obuhvata i sam sažetak lekcije u okviru koje student interaguje sa četbotom takođe definisan uz pomoć nastavnika. Ovaj sažetak lekcije može pomoći četbotu da razume širi kontekst u okviru koga studenti postavljaju pitanja koja nisu uvek u potpunosti jasno postavljena. Ovde treba voditi računa o dužini samog sažetka s obzirom na to da može uticati na cenu i kvalitet odgovora, jer je pokazano da predugački ulazi mogu uticati na skretanje fokusa četbot odgovora i zanemarivanje određenih instrukcija [141], [142], [143].

S druge strane, nastavnik može odlučiti da četbotu proširi kontekst generisanja odgovora i pruži dodatnu bazu znanja (*BazaZnanja*). Kroz pružanje relevantnih informacija kroz kontekst može se uticati na smanjenje halucinacija modela što je posebno važno u akademskom okruženju i procesu učenja [124]. Iz date baze se može preuzeti *DodatniKontekst* za odgovarajuću korisničku poruku koji konverzacioni agent može upotrebiti za generisanje odgovora. Sama *UlaznaInstrukcijaOPonašanju* može sadržati ograničenja u pogledu načina korišćenja ovako dodatog konteksta.

IstorijatKonverzacije obuhvata prethodne poruke korisnika, kao i četbot odgovore, koji se takođe prosleđuju radi održavanja konzistentnih odgovora u skladu sa tokom konverzacije. Kao i kod sažetka lekcije, i ovde treba voditi računa o broju prethodnih poruka i odgovora koji će se prosleđivati, jer njihova količina takođe može uticati na cenu i kvalitet odgovora. Pored slanja istorijata koji sadrži parove pitanja i odgovora, moguće je odlučiti se i za slanje sažetka istorijata koji bi takođe mogao biti kreiran uz pomoć velikog jezičkog modela, ili njihovu kombinaciju [144], [145].

KontekstKonverzacije zapravo predstavlja kombinaciju prethodno navedenih elemenata zajedno sa složenim.

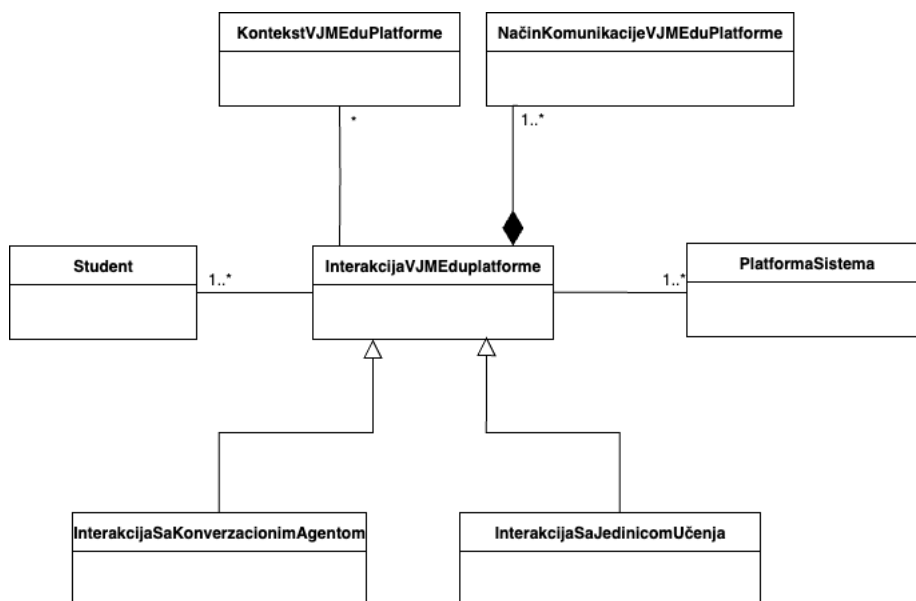
Takođe, kao što smo već spomenuli student (*Student*) može rešavati aktivnosti, tokom procesa učenja, uz moguću konverzaciju sa četbotom za pojašnjenja i uz raspoložive materijale. Njegov odgovor (*Odgovor*) može biti sačuvan radi dalje evaluacije što je predstavljeno na narednoj slici 8.



Slika 8 - Metamodel odgovora na aktivnost

6.2.5 Metamodel interakcije sa edukativnom platformom

Student interaguje sa edukativnom platformom na dva načina: prolaskom kroz materijal za učenje (*InterakcijaSaJedinicomUčenja*) i kroz vođenje konverzacije sa četbotom zasnovanim na velikom jezičkom modelu (*InterakcijaSaKonverzacionimAgentom*). Kao što je prikazano na slici 9 interakcija se može uspostaviti između više studenata (*Student*) i više platformi (*PlatformaSistema*). Za modelovanje interakcije, možemo se osloniti na metamodel predstavljenog u radu Obrenovića [146] koji definiše interakciju čoveka i računara na višem nivou apstrakcije i obuhvata višenačinsku (eng. *multimodal*) komunikaciju uključivanjem dva osnovna elementa: način (eng. *modality*) komunikacije i kontekst interakcije.



Slika 9 - Metamodel interakcije sa edukativnom platformom

Samim tim, *NačinKomunikacijeVJMEduPlatforme* se odnosi na oblik interakcije između studenta i edukativne platforme koji podstiče korišćenje sposobnosti čoveka radi izazivanja efekta na korisnika. Dakle, pojam efekta se može definisati kao rezultat upotrebe ljudskih sposobnosti u komunikaciji sa računarom [146]. Ovi efekti mogu da se klasifikuju u četiri glavne kategorije [147]: senzorski koji se odnose na procesuiranje stimulansa od strane ljudskog senzorskog aparata (npr. zvuk); perceptijski koji proističu iz analize senzorskih podataka kroz ljudski perceptijski aparat (npr. jačina zvuka); motorni koji opisuju ljudske mehaničke radnje (npr. pomeranje glave); i kognitivni koji su ključni kod interakcije sa predloženom edukativnom platformom i dešavaju se na višim nivoima obrade informacija od strane čoveka (pamćenje, pažnja, radoznali procesi).

Obrenović i Starčević [147] definišu osnovni metamodel načina komunikacije (modaliteta) na sledeći način. Način komunikacije može biti jednostavan ili složen. U ovom modelu, računarski kod, odnosno kompjuter predstavlja referentnu tačku u odnosu na koju se definiše interakcija. Jednostavan način komunikacije predstavlja primitivnu formu interakcije koja može biti ulazna ili izlazna. Ulazni modalitet je u smeru od čoveka ka računaru i zahteva korisnički uređaj da prevede čovekovu akciju u odgovarajuću formu za obradu od strane računara (npr. klik mišem ili unos na tastaturi što može biti unos poruke namenjene četbotu ili skrolovanje kroz jedinicu učenja). Izlazni modalitet je u smeru računara ka čoveku i odnosi se na način predstavljanja podataka korisniku, kroz npr. slike (statički) ili video snimke (dinamički) (npr. predstavljanje jedinice učenja studentu kroz prikaz teksta i slike, ili prikaz odgovora četbota). Složen način komunikacije predstavlja kombinaciju dva ili više modaliteta u cilju njihove simultane upotrebe.

S druge strane, *KontekstVJMEduPlatforme* se odnosi na skup uslova ili činjenica koje okružuju studenta i mogu da utiču na njegovu interakciju sa edukativnom platformom. U spomenutom metamodelu definisane su četiri vrste konteksta: fizički, situacioni, socijalni i zajedničko znanje. Fizički kontekst obuhvata fizičke elemente okruženja u kome se interakcija odvija. Elementi mogu biti vrsta prostora u kome se odvija interakcija, kao i vrednosti parametara poput temperature vazduha, nivoa osvetljenja, buke i vlažnosti vazduha. Socijalni kontekst obuhvata socijalno okruženje u kome se interakcija odvija što može biti posebno važno u interakciji sa malim prenosivim uređajima (npr. mobilni telefon) na različitim lokacijama. Situacioni kontekst definiše trenutnu situaciju u okruženju sa stanovišta zadatka koji korisnik treba da uradi. Zajedničko znanje predstavlja skup činjenica koje računarski sistemi i ljudi uzimaju kao nešto što se podrazumeva [148].

6.2.6 Metamodel evaluacije

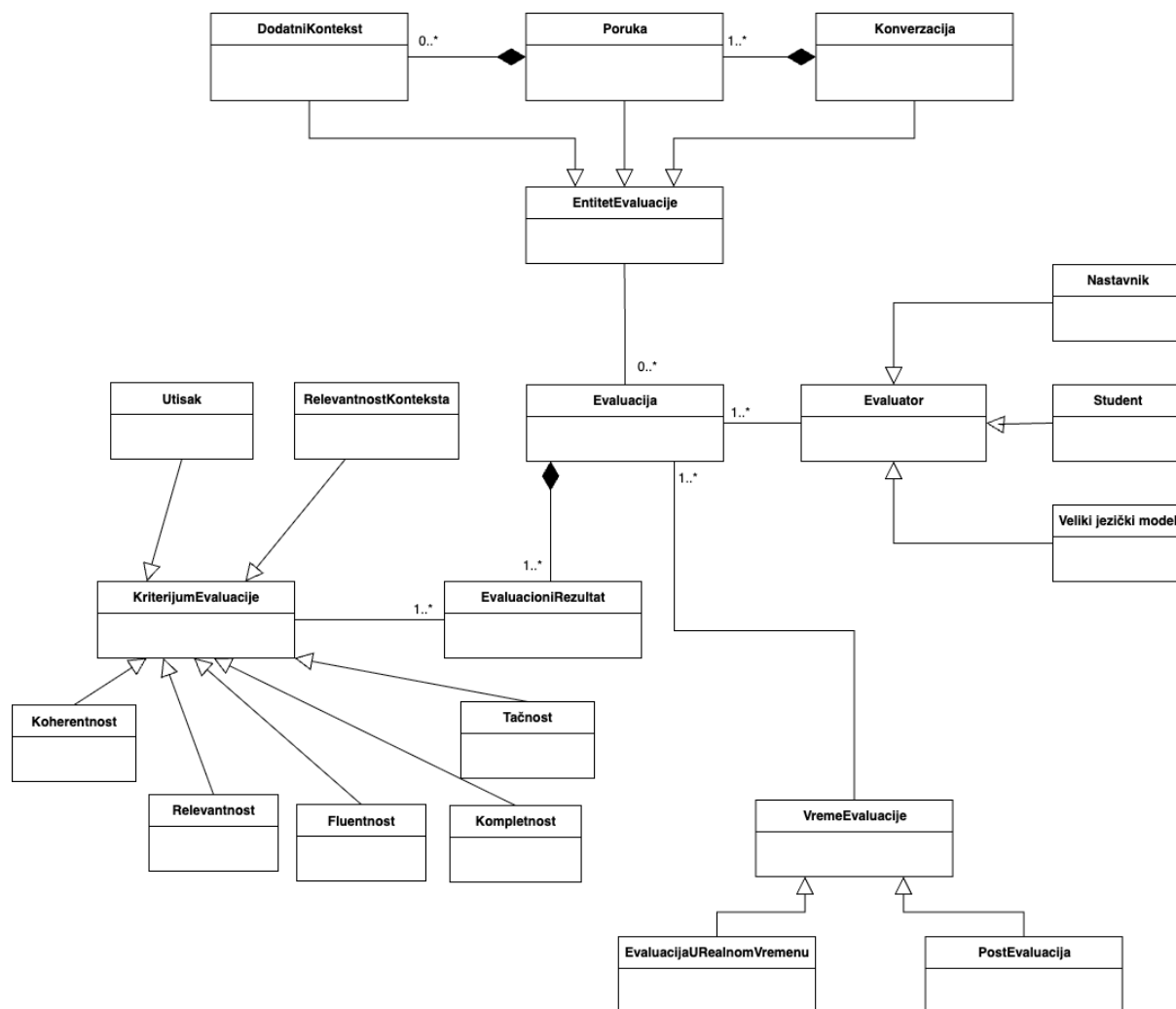
Pre davanja platforme na korišćenje studentima (kao i tokom njenog korišćenja), treba izvršiti evaluaciju četbot odgovora na određenom broju očekivanih pitanja. Poželjno je organizovati i pivot korišćenje platforme od strane studenata kako bi se prikupila moguća pitanja i time eventualno proširio set pitanja za evaluaciju sistema. Ukoliko rezultati evaluacije nisu na odgovarajućem nivou potrebno je izvršiti određene izmene sistema radi poboljšanja performansi koje mogu uključivati ažuriranje instrukcija, dodate baze znanja i načina njene pretrage, podešavanje parametara modela i sl. Nakon toga treba opet ponoviti proces evaluacije.

Sama evaluacija može biti sprovedena za odgovarajuću konverzaciju, konkretan četbot odgovor i njegov dodatni kontekst ukoliko postoji (*EntitetEvaluacije – Konverzacija, Poruka* sa odgovarajućom ulogom, *Dodatni Kontekst*). Evaluacija se sastoji od rezultata evaluacije (*RezultatEvaluacije*) i može se vršiti prema više kriterijuma (*KriterijumEvaluacije*) koji mogu biti [112]: *Tačnost* kao mera ispravnosti izlaza koji generiše VJM; *Relevantnost* izlaza četbota u odnosu na konkretno pitanje studenta; *Kompletnost* kao stepen u kojem odgovor uključuje sve suštinske informacije potrebne da se adekvatno adresira konkretno pitanje; *Koherentnost* koja se odnosi na to koliko je odgovor jasno struktuiran i logički povezan u odnosu na kontekst pitanja i istoriju razgovora; i *Fluentnost* koja se odnosi na ispravnost gramatike i tečan, prirodan način upotrebe jezika. Takođe, ukoliko je u generisanje odgovora uključen i dodatni kontekst izvučen iz baze znanja poželjno je evaluirati i njegovu relevantnost (*RelevantnostKonteksta*). U zavisnosti od konkretnih zahteva moguće je ažurirati listu mogućih kriterijuma.

Pored manuelne evaluacije konverzacija od strane nastavnika (*Nastavnik*) u početnim fazama, koja sa rastom broja korisnika i pitanja postaje kompleksnija za izvođenje, može se uspostaviti i automatizovana evaluacija od strane velikih jezičkih modela (*eng. LLM as a Judge* [149]). U ovom slučaju je potrebno uzeti i cenu kao bitan faktor prilikom izbora modela za evaluaciju i uspostavljanja monitoringa sistema.

Pored evaluacije od strane nastavnika, moguće je uspostaviti i evaluaciju konkretnog odgovora ili konverzacije od strane studenta (*Student*). Ovom evaluacijom bi se merilo zadovoljstvo studenta kroz kriterijum utiska (*Utisak*) koji je četbot odgovor ili odgovori ostavio na njega što može kasnije može pomoći u procesu kontinuirane evaluacije i praćenja sistema i unapređenju funkcionalnosti četbota. Dakle, evaluatori sistema (*Student, Nastavnik, Veliki jezički model*) su generalizovani kroz *Evaluator* klasu.

Pored sprovođenja evaluacije nakon generisanja odgovora od strane konverzacionalnog agenta koji se prikazuju korisniku (*PostEvaluacija*), u određenim situacijama može biti neophodno vršiti evaluaciju u realnom vremenu (*EvaluacijaURealnomVremenu*), odnosno pre nego odgovor stigne do krajnjeg korisnika, ukoliko je vrlo važno osigurati odgovor u određenom formatu kao što je slučaj u [116], [120] čime se implementira jedna od mogućih zaštitnih *guardrail* tehnika. U spomenutim radovima, autori su kroz veliki jezički model proveravali i ažurirali odgovor tako da ne sadrži gotov programski kod kako ne bi davali kompletna rešenja studentima. Treba imati u vidu da ovo utiče na kašnjenje odgovora, kao i na samu cenu generisanja odgovora.



Slika 10 – Metamodel evaluacije

Takođe, važno je napomenuti da je kontinuirana evaluacija i praćenje ponašanja sistema je neophodna u ovakvim vrstama sistema koji su nedeterministički u osnovi i čiji izlaz može biti nepredvidiv [14], [150]. Dakle, tokom korišćenja sistema od strane korisnika, moraju se uspostaviti mehanizmi nadgledanja i evaluacije kako bi se osigurale performanse sistema.

7 MODELOVANJE EDUKATIVNE PLATFORME SA KONVERZACIONIM AGENTOM ZASNOVANIM NA VELIKOM JEZIČKOM MODELU

Ovo poglavlje prikazuje primenu prethodno definisanog objedinjenog modela za razvoj edukativne platforme zasnovane na velikom jezičkom modelu za proširenje jezika za modelovanje. Korišćen pristup je zasnovan na objektno-orijentisanoj paradigmi za modelovanje sistema. Za modelovanje će biti korišćeni UML koncepti i mogućnost njegovog proširenja kroz stereotype [151]. Stereotipi predstavljaju osnovni gradivni element UML profila koji predstavlja ključan mehanizam za proširenje UML-a. Pored stereotipa, označene vrednosti i ograničenja su osnovni elementi profila koji omogućavaju prilagođavanje UML-a specifičnim domenima. Označene vrednosti predstavljaju parove ključ-vrednost koji su dodeljeni stereotipima kao metaatributi. Ograničenja takođe mogu biti povezana sa stereotipima i uvode pravila koja definisani elementi moraju pratiti.

Kroz naredna potpoglavlja biće predstavljen UML profil edukativne platforme koja integriše VJM konverzionog agenta, koji je definisan na osnovu prethodno opisanog objedinjenog modela. Na odabranim primerima će biti prikazane mogućnosti proširenja UML jezika sa konceptima ovog modela. Naravno, modeli predstavljeni kroz odgovarajuće profile u narednim poglavljima nisu sveobuhvatni i predstavljaju jedan od mogućih pristupa modelovanju edukativne platforme koja integriše konverzionog agenta zasnovanog na velikom jezičkom modelu, koji se u skladu sa potencijalnim izmenama zahteva sistema mogu prilagođavati odnosno proširiti na drugačiji način.

Sledeći aspekti sistema edukativne platforme koja integriše konverzionog agenta zasnovanog na VJM-u su odabrani za definiciju proširenja kroz profile:

- Edukativni sadržaj,
- Konverzioni agent zasnovan na velikom jezičkom modelu,
- Putanja učenja,
- Evaluacija.

7.1 Modelovanje edukativnog sadržaja

U nastavku su navedeni stereotipi klasa definisani u okviru profila edukativnog sadržaja:

- Kurs – odnosi se na kurs za koji se sadržaj kreira
- Nastavnik – predstavlja nastavnika koji kreira sadržaj za dati kurs
- Edukativni sadržaj – obuhvata materijale za učenje kreirane od strane nastavnika za određeni deo kursa
- Lekcija – odnosi se na materijal za učenje kreiranog od strane nastavnika koji ima svoj redosled i sastoji se od jedinica učenja

- Nastavna jedinica – predstavlja deo lekcije koji ima svoj redosled u okviru lekcije i pripada određenoj sekciji i sastoji se od objekata za učenje
- Aktivnost – predstavlja deo lekcije koji ima svoj redosled u okviru lekcije, sastoji se od objekata za učenje i može sadržati ponuđene odgovore, kao i tačan odgovor i cilj nivoa saznanja koji želi da se postigne ovom aktivnošću
- Objekti za učenje – su konkretne instance materijala za učenje koji mogu biti tekst, slika, video snimak, audio snimak ili njihova kombinacija koja predstavlja složen objekat za učenje

U nastavku je prikazana tabela definisanih stereotipa sa pridodatim vrednostima, njihovim opisom i primerom upotrebe date vrednosti, kao i tabela definisanih nabiranja profila edukativnog sadržaja, mogućim vrednostima datih nabiranja i njihovim opisom.

Tabela 8 – Stereotipi profila edukativnog sadržaja sa pridodatim vrednostima

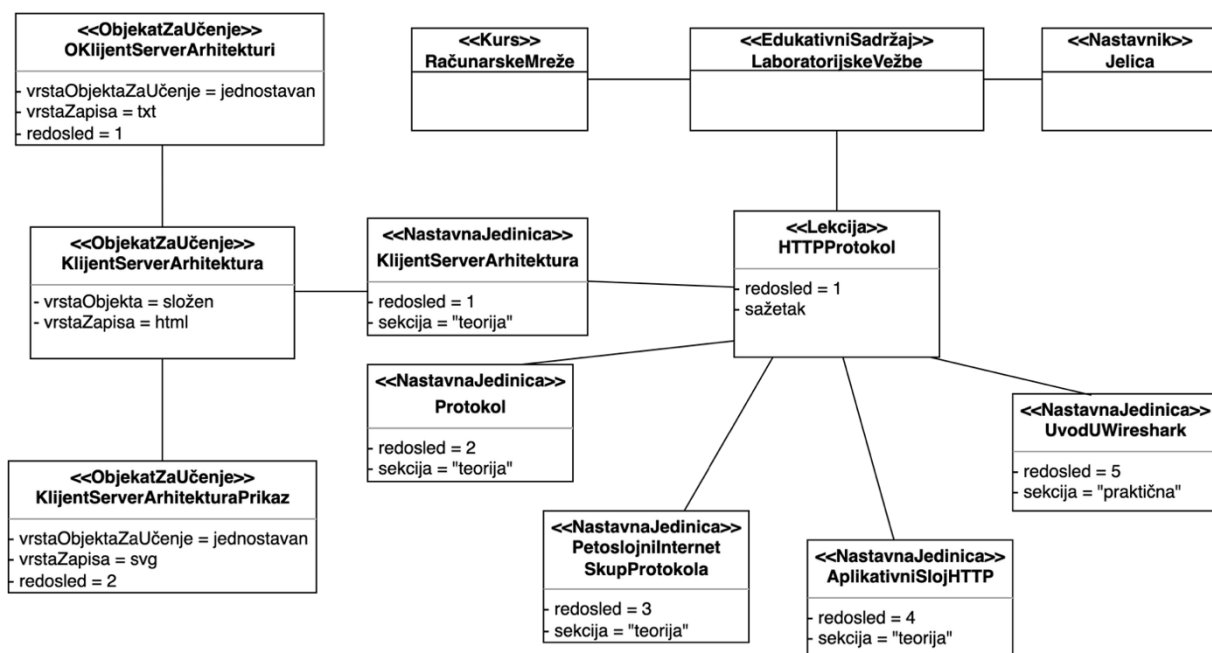
Stereotip	Pridodata vrednost	Opis	Tip	Primer upotrebe
Lekcija	redosled	Definiše redosled lekcije u okviru edukativnog sadržaja	Celi broj	1
	sažetak	Definiše sažetak lekcije	Tekstualni tip	
Nastavna Jedinica	redosled	Definiše redosled jedinice učenja u okviru lekcije	Celi broj	1
	sekcija	Definiše sekciju kojoj pripada jedinica učenja	Tekstualni tip	Teorija
Aktivnost	redosled	Definiše redosled aktivnosti u okviru lekcije	Celi broj	1
	rešenje	Definiše tačno rešenje date aktivnosti	Niz tačnih odgovora	S:192.168.0.15; D:192.168.0.1
	vrstaOdgovora	Definiše vrstu odgovora	Nabrajanje VrstaOdgovora	otvoren
	odgovori	Ponuđeni odgovori za datu aktivnost ukoliko postoje odnosno ukoliko nije u pitanju otvoreno pitanje	Niz ponuđenih odgovora	da; ne
	nivoSaznanja	Određuje podržani nivo saznanja objekta prema Andersonovoj taksonomiji	Nabrajanje NivoSaznanja	razumevanje
Objekat ZaUčenje	vrstaObjekta ZaUčenje	Definiše vrstu objekta za učenje	Nabrajanje VrstaObjekta ZaUčenje	jednostavan
	vrstaZapisa	Definiše vrstu zapisa objekta za učenje	Nabrajanje VrstaZapisa	txt

	redosled	Definiše redosled objekta za učenje u okviru jedinice za učenje	Celi broj	1
--	----------	---	-----------	---

Tabela 9 – Definisana nabiranjaja profila edukativnog sadržaja

Nabrajanje	Moguće vrednosti	Opis
VrstaObjektaZaUčenje	jednostavan, složen	Moguće vrste objekata za učenje, jednostavni su tekst, grafika, video ili zvuk, a složeni njihova kombinacija
VrstaZapisa	html, txt, svg png, jpg, gif, mp3, mp4, md	Podržani tipovi zapisa koji čine edukativni sadržaj odnosno jedinice učenja
VrstaOdgovora	da_ne, višestruki_odgovor, otvoren	Opcije za odgovor na aktivnost
NivoSaznanja	pamćenje, razumevanje, primena, analiza, evaluacija, stvaranje	Andersenova taksonomija znanja

Primer upotrebe definisanog profila kroz definisane stereotipe je dat na slici 11 kroz jednu lekciju laboratorijskih vežbi na predmetu Računarske mreže i telekomunikacije. Na prikazanom primeru imamo lekciju edukativnog sadržaja za određeni kurs, koja se sastoji iz više nastavnih jedinica, a čiji su elementi predstavljeni kroz objekte za učenje.



Slika 11 - Primer upotrebe profila edukativnog sadržaja

7.2 Modelovanje konverzionog agenta zasnovanog na velikom jezičkom modelu

U nastavku su navedeni stereotipi klasa definisani u okviru profila VJMKonverzionAgent:

- Student – definiše studenta koji vodi konverzaciju
- VJMKonverzionAgent – definiše konverzionog agenta zasnovanog na velikom jezičkom modelu
- Konverzacija – definiše konverzaciju između studenta i konverzionog agenta za odgovarajući opseg konverzacije
- Poruka – definiše studentsku poruku ili poruku odnosno odgovor konverzacinog agenta koja je deo neke konverzacije
- KontekstKonverzacije – definiše kontekst na osnovu koga se generiše odgovor
- Instrukcija – usmerava način ponašanja velikog jezičkog modela prilikom generisanja odgovora odnosno utiče na način kako VJM generiše odgovor
- BazaZnanja – definiše vektorsku bazu koju nastavnik može da popuni dodatnim materijalima u tekstualnom formatu koji bi služili za poboljšanje konteksta generisanja odgovora od strane velikog jezičkog modela
- DodatniKontekst – definiše dovučen dodatni kontekst iz dodatne baze znanja za datu korisničku poruku
- VJMAPI – definiše API velikog jezičkog modela koji se koristi za generisanje odgovora

U nastavku je prikazana tabela definisanih stereotipa sa pridodatim vrednostima i primerom upotrebe date vrednosti.

Tabela 10 – Stereotipi profila VJM Konverzacioni Agent sa pridodatim vrednostima

Stereotip	Pridodata vrednost	Opis	Tip	Primer upotrebe
VJM Konverzacioni Agent	sistemskaInstrukcija	Difoltna instrukcija koja definiše način ponašanja modela prilikom generisanja odgovora	Tekstualni tip	
	model	Veliki jezički model korišćen za generisanje odgovora na osnovu korisničkog ulaza	Nabrajanje VelikiJezički Model	gpt-4o
	maksimalniBroj Tokenalzlaza	Maksimalan broj tokena generisanog odgovora od strane modela.	Ceo broj	500
	maksimalniBroj Tokenalstorijata	Maksimalan broj tokena istorijata konverzacija koji je deo konteksta generisanja.	Ceo broj	1500
	temperatura	Predstavlja vrednost koja definiše način generisanja odgovora modela. Više vrednosti kreiraju nasumičniji odgovor, dok niže vrednosti kreiraju više fokusiran i deterministički odgovor. Neke od preporuka su 0,1-0,3 za činjenične odgovore, 0,2-0,5 za tehnička objašnjenja, 0,7-1 za kreativno pisanje, 0,8-1,2 za brejnstorming ideja, 0,1-0,4 za struktuirano formatiranje [139].	Realni broj između 0 i 2	1
Konverzacija	vremePočetka	Datum i vreme početka konverzacije	Timestamp	19.08.2025. 07:30:00
Poruka	uloga	Kreator poruke	Nabrajanje Uloga	user
	sadržaj	Sadržaj poruke korisnika ili četbot odgovora	Tekstualni tip	
	vremeKreiranja	Datum i vreme kreiranja poruke	Timestamp	08.02.2025. 22:00:00
	brojTokena	Broj tokena od kojih se poruka sastoji	Celi broj	20

	brojTokenaZaGenerisanjeOdgovora	Broj ulaznih tokena konteksta za generisanje odgovora poruke (instrukcija, istorijat poruka, korisnički ulaz, dodatni kontekst) koji može biti korišćen za evaluaciju potrošnje	Celi broj	1500
Instrukcija	sadržaj	Ulazna instrukcija koja definiše ponašanje čebota za određeni opseg	Tekstualni tip	
	verzija	Broj verzije ulazne instrukcije	Celi broj	1
BazaZnanja	modelUgradnje	Model koji se koristi za kreiranje ugradnji na osnovu teksta koje se čuvaju u vektorskoj bazi	Nabranje VrstaModela Ugradnje	text-embedding-3-small
	veličinaKomada	Broj karaktera jednog komada teksta za čuvanje u vektorskoj bazi	Celi broj	1000
	preklapanjeKomada	Broj karaktera preklapanja između dva komada susednog teksta koji se čuvaju u vektorskoj bazi	Celi broj	200
	granicaSkoraSličnosti	Minimalna vrednost skora sličnosti koju rezultati moraju imati prilikom pretrage vektorske baze koristeći kosinusnu sličnost kao metriku (više vrednosti označavaju veću sličnost)	Realan broj u opsegu od -1 do 1	0,6
Dodatni Kontekst	sadržaj	Sadržaj dodatnog konteksta izvučenog iz vektorske baze	Tekstualni tip	
	skorSličnosti	Skor sličnosti	Realan broj u opsegu od -1 do 1	0,7
	upitPretrage	Upit na osnovu kog je pretražena vektorska baza koji je rezultovao ovim sadržajem i skorom (može biti sama poruka korisnika ili prilagođena poruka)	Tekstualni tip	veličina zaglavlja IP datagrama u IPv4 protokolu
VJMAPI	VJMključ	Definiše API ključ neophodan za povezivanje sa odabranim LLM API-jem	Tekstualni tip	
	VJMURL	Definiše API URL neophodan za povezivanje sa odabranim LLM API-jem	Tekstualni tip	

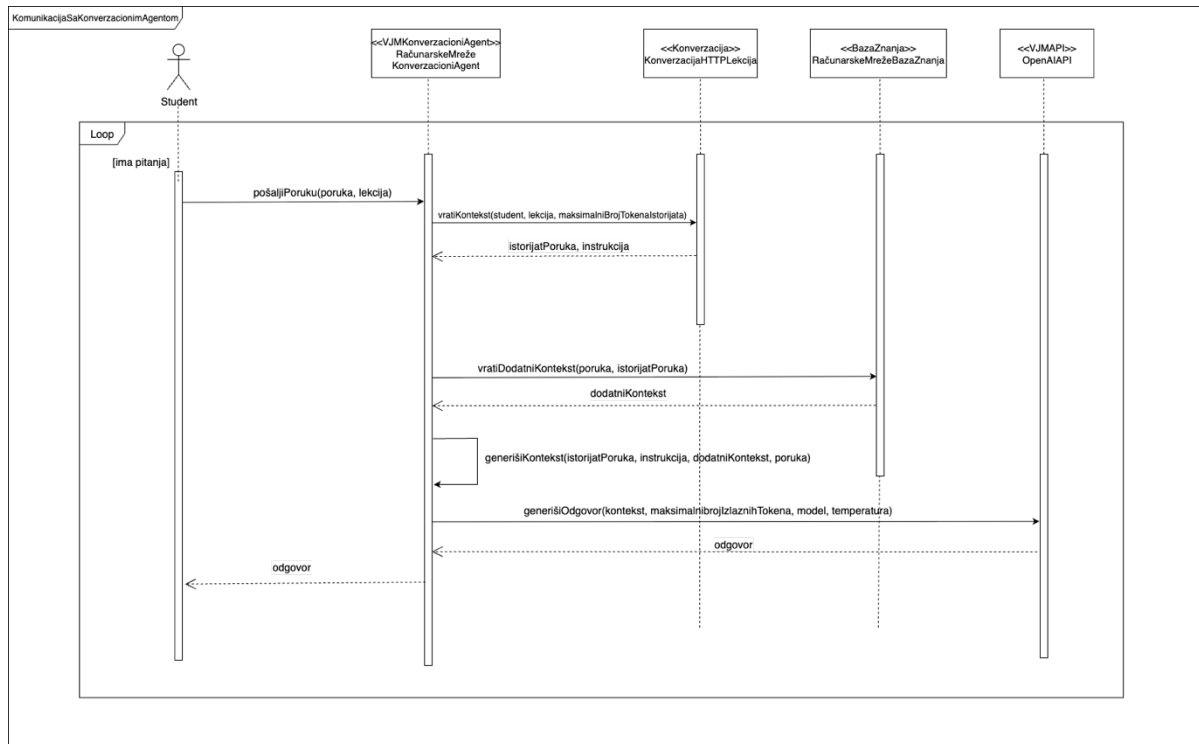
Definisana nabiranja profila su data u tabeli 11.

Tabela 11 – Definisana nabiranja profila VJM Konverzacioni Agent

Nabiranje	Moguće vrednosti	Opis
VrstaModelaUgradnje	text-embedding-3-small, text-embedding-3-large ...	Veliki jezički modeli koji se koriste za generisanje ugradnji za vektorske baze podataka.
VrstaVelikogJezičkogModela	gpt-4, gpt-4o ...	Veliki jezički modeli koji se koriste za generisanje četbot odgovora.
NabiranjeUloga	<i>user, assistant</i>	Uloge koje definišu kreatora poruke (student ili četbot).

Takođe, pored prethodno definisanih stereotipa vezanih za komunikaciju između studenta i konverzionog agenta tokom prolaska kroz nastavne jedinice ili izradu aktivnosti, treba definisati i stereotip *Odgovor* sa pridodatim vrednostima *vremeSlanja*, *sadržajOdgovora* i *ostvarenBrojPoena*.

Primer upotrebe definisanog profila je dat na slici 12 kroz dijagram sekvenci na visokom nivou apstrakcije koji prikazuje način komunikacije delova sistema za generisanje odgovora na studentsko pitanje, konkretno na primeru konverzionog agenta za kurs Računarske mreže i telekomunikacije i lekciju o HTTP protokolu. Komunikacija započinje studentom koji šalje poruku za odgovarajuću lekciju kroz koju trenutno prolazi. Konverzacioni agent na osnovu lekcije dovlači odgovarajuću instrukciju na osnovu koje će se generisati odgovor. Takođe dovlači i prethodne poruke konverzacije studenta koja se vodi za određen opseg (lekciju ili jedinicu učenja) ukoliko postoje u skladu sa ograničenjem maksimalnog broja tokena istorijata. Na osnovu studentove poslednje poruke i istorijata poruka koji može biti potreban radi adekvatne pripreme upita pretrage, konverzacioni agent dovlači iz baze znanja dodatni kontekst ukoliko postoji. Zatim generiše kontekst konverzacije koji obuhvata poruku, instrukciju, istorijat poruka i dodatni kontekst koji prosleđuje uz informacije o modelu, temperaturi i maksimalnom broju tokena VJMAPI-ju za generisanje odgovora koji se može strimovati nazad studentu.



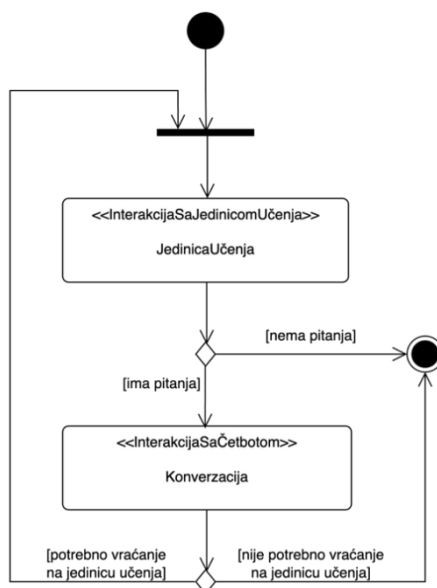
Slika 12 – Primer upotrebe profila VJM Konverzacijom Agent

7.3 Modelovanje putanje učenja

U nastavku su navedeni stereotipi klasa koji su definisani u okviru profila putanje učenja:

- Interakcija sa jedinicom učenja – definiše studentov prolaz kroz materijale za učenje dostupne u vidu prikaza objekata za učenje
- Interakcija sa četbotom – definiše vođenje konverzacije sa dostupnim četbotom zasnovanim na velikom jezičkom modelu u cilju razjašnjavanja materije i produblivanje znanja

Primer upotrebe definisanog profila je dat na slici 13 kroz dijagram aktivnosti koji na apstraktnom nivou prikazuje tok interakcije studenta sa edukativnom platformom odnosno sa jednom jedinicom učenja u okviru lekcije. Student prvo prolazi kroz materijal za učenje u formi objekta za učenje koje je nastavnik pripremio za tu jedinicu učenja. Ukoliko ima potrebe, student vodi konverzaciju sa četbotom uz mogućnost vraćanja na materijale za učenje.



Slika 13 - Primer upotrebe profila putanje učenja

7.4 Modelovanje evaluacije

U nastavku su navedeni stereotipi klasa koji su definisani u okviru profila evaluacije:

- Evaluacija – definiše parametre evaluacije četbot odgovora
- Evaluacioni rezultat – definiše rezultat evaluacije dodeljen za određenu metriku
- VJMEvaluator – definiše evaluatora zasnovanog na velikom jezičkom modelu i njegovo ponašanje tzv. VJM kao sudija (eng. *LLM as a Judge*)

U nastavku je prikazana tabela definisanih stereotipova sa pridodatim vrednostima i primerom upotrebe date vrednosti.

Tabela 12 - Stereotip profila Evaluacija sa pridodatim vrednostima

Stereotip	Pridodata vrednost	Opis	Tip	Primer upotrebe
Evaluacija	skor	Definiše vrednost evaluacije koji se može dobiti kao prosek svih rezultata ili nekom drugom formulom.	Realan broj	5
	vreme	Definiše vreme evaluacije	Nabrajanje Vreme Evaluacije	Post Evaluacija
	komentar	Opcioni opisni komentar evaluacije.	Tekstualni tip	“Odličan odgovor”
Evaluacioni Rezultat	kriterijum	Definiše kriterijum evaluacije	Nabrajanje Kriterijum Evaluacije	Kompletnost
	vrednost	Definiše vrednost evaluacije	Celi broj	5
	skala	Definiše skalu evaluacije	Nabrajanje	Likert5

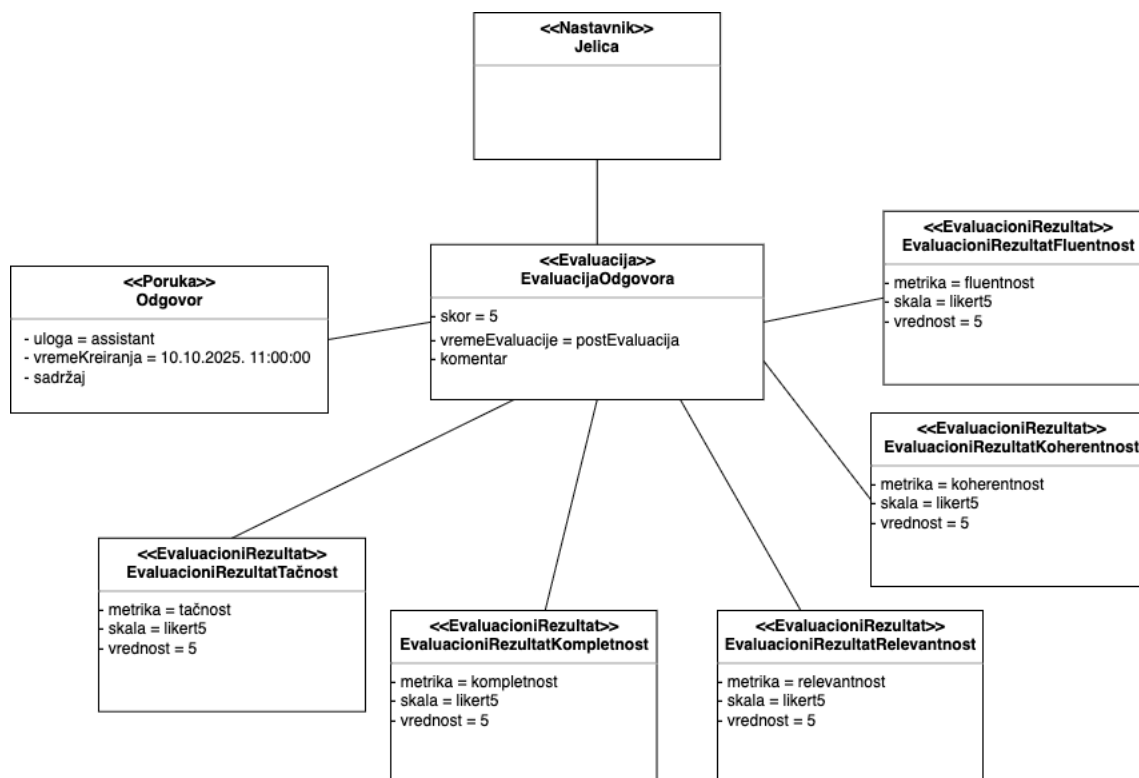
			Skala	
VJM Evaluator	model	Veliki jezički model korišćen za generisanje odgovora na osnovu korisničkog ulaza	Nabrajanje VelikiJezički Model	gpt-4o-mini
	maksimalni BrojIzlaznih Tokena	Maksimalan broj tokena generisanog odgovora od strane modela.	Ceo broj	500
	temperatura	Predstavlja vrednost koja definiše način generisanja odgovora modela. Više vrednosti kreiraju nasumičniji odgovor, dok niže vrednosti kreiraju više fokusiran i deterministički odgovor.	Realni broj između 0 i 2	0
	jsonŠema Odgovora	Definiše format odgovora koji se traži od LLM modela radi lakše obrade rezultata evaluacije.	json	{ tačnost: number; relevantnost: number; koherentnost: number; }
	instrukcija ZaEvaluaciju	Definiše instrukciju na osnovu koje veliki jezički model treba da izvrši evaluaciju.	Tekstualni tip	

Definisana nabiranjaja profila Evaluacija su data u tabeli 13.

Tabela 13 – Definisana nabiranjaja profila evaluacije

Nabrajanje	Moguće vrednosti	Opis
EvaluacionaSkala	Likert5 Likert7 Binarna	Definiše skalu za ocenu metrika.
KriterijumEvaluacije	Tačnost Relevantnost Kompletnost Koherentnost Fluentnost Zadovoljstvo RelevantnostKonteksta	Definiše moguće evaluacione kriterijume.
VremeEvaluacije	EvaluacijaURealnomVremenu PostEvaluacija	Definiše trenutak sprovođenja evaluacije.

Na slici 14 prikazan je primer upotrebe definisanog profila uz stereotipe iz profila Edukativnog sadržaja i VJM konverzionog agenta, za jednu evaluaciju odgovora od strane četbota koju je sproveo nastavnik prema kriterijumima tačnosti, relevantnosti, kompletnosti, koherentnosti i fluentnosti.



Slika 14 – Primer upotrebe profila Evaluacija

8 PREDLOG PROCESA RAZVOJA EDUKATIVNE PLATFORME SA KONVERZACIONIM AGENTOM ZASNOVANIM NA VELIKOM JEZIČKOM MODELU

Napredak tehnologije tokom godina je omogućio razvoj sve kompleksnijih softverskih sistema namenjenih krajnjim korisnicima. Kako bi se olakšao proces razvoja softverskih sistema razvijene su metodologije za razvoj softvera, kao što je Unified proces koji je korišćen u ovoj disertaciji.

8.1 Unified proces razvoja softvera

Proces razvoja softvera uopšteno predstavlja skup aktivnosti koje imaju za cilj transformaciju korisničkih zahteva u softverski sistem [152]. Unified proces razvoja softvera ili skraćeno Unified proces je generički okvir koji koristi UML za opisivanje delova ovog procesa. Unified proces ne predstavlja samo proces, već proširiv okvir koji može biti prilagođen različitim vrstama softverskih sistema i organizacijama. Predstavlja iterativno-inkrementalnu metodologiju fokusiranu na slučajeve korišćenja, koji predstavljaju funkcionalnosti koje pružaju vrednost korisnicima, i na arhitekturu sistema, koja opisuje glavne statičke i dinamičke aspekte sistema. U ovoj disertaciji je ova metodologija izabrana radi definisanja procesa razvoja integrisane edukativne platforme sa konverzacionim agentom zasnovanim na velikom jezičkom modelu.

Faze koje čine životni ciklus Unified procesa su [152]:

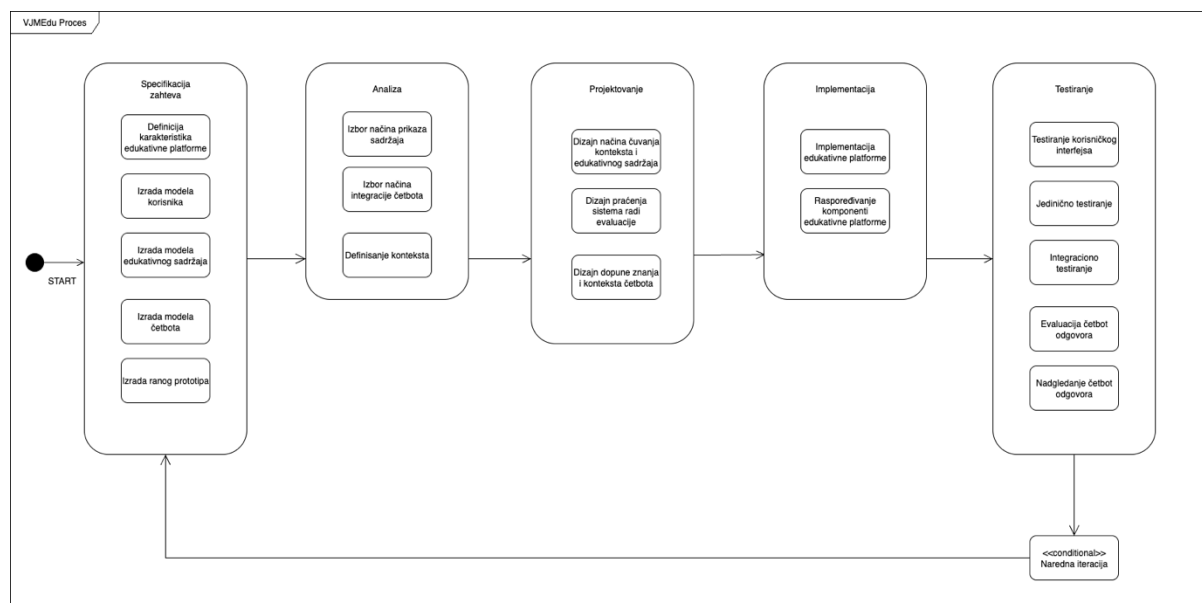
- Inicijalizacija – faza u okviru koje se definišu slučajeve korišćenja na najvišem nivou, odnosno zahtevi sistema i vrši njihova analiza,
- Elaboracija – faza u okviru koje se dalje razrađuju slučajeve korišćenja kao i arhitektura sistema,
- Konstrukcija – faza u okviru koje se vrši izrada i testiranje sistema,
- Tranzicija – faza u okviru koje se sistem pruža na korišćenje i eventualno vrše neophodne ispravke.

Svaka od ovih faza može imati više iteracija čiji izlaz predstavlja inkrement sistema [152].

S druge strane, radni tokovi/discipline koje se odvijaju tokom spomenutih faza životnog ciklusa i koje će biti detaljno opisane sa aspekta ove disertacije u nastavku su [152]: definisanje zahteva, analiza, projektovanje, implementacija i testiranje. Razvoj počinje prikupljanjem zahteva koji predstavljaju slučajeve korišćenja koje dati sistem treba da implementira. Zatim sledi analiza i projektovanje sistema kako bi spomenuti slučajeve korišćenja bili uspešno realizovani, nakon čega sledi implementacija koja obuhvata sam kod. Na kraju se uspostavlja način testiranja sistema kako bi se izvršila verifikacija da li sistem pruža funkcionalnosti u skladu sa slučajevima korišćenja.

8.2 Predloženi proces razvoja

Na slici 15 je predstavljen predloženi proces razvoja sa pregledom proširenja po definisanim radnim tokovima. U narednim poglavljima su detaljnije definisane aktivnosti u okviru ovih tokova koje su relevantne za razvoj edukativne platforme koja integriše konverzionog agenta zasnovanog na velikom jezičkom modelu.



Slika 15 – Predloženi proces razvoja edukativne platforme sa VJM konverzionim agentom

U okviru faze definisanja zahteva potrebno je definisati karakteristike VJM edukativne platforme koje sistem treba da zadovolji, kao i model samog korisnika koji će koristiti datu platformu. Potrebno je definisati model edukativnog sadržaja koji se odnosi na domen odnosno kurs za koji je platforma namenjena. Nastavnik, odnosno domenski ekspert preuzima ovu ulogu. Predlažemo korišćenje UML profila *EdukativniSadržaj* definisanog u prethodnom poglavlju kao proširenje ove aktivnosti. Takođe, u ovoj fazi treba definisati model samog konverzionog agenta odnosno definisati njegov način ponašanja, njegovu ulogu, cilj, ton i opseg pitanja na koja će odgovarati u vezi sa edukativnim sadržajem koji nastavnik priprema. Predlažemo korišćenje dela UML profila *VJMKonverzioniAgent* definisanog u prethodnom poglavlju kao proširenje ove aktivnosti. U ovoj fazi je moguće napraviti i prototip sistema koji na ovom nivou može biti papirni prototip koji ilustruje funkcionalnosti aplikacije.

Faza analize preciznije definiše i strukturu specifikovane zahteve. U fazi analize treba odrediti način prikaza edukativnog sadržaja, kao i načina integracije samog četbota odnosno u kojim situacijama će biti dostupan studentima za korišćenje. Kao što je navedeno u početnom istraživanju faktora uticaja na prihvatanje ovakvih alata u obrazovanju, nastavnik bi trebalo da proceni potencijal i slučajeve korišćenja u kojima konverzioni agent može biti od koristi studentu. Treba definisati kontekst četbota koji je potreban za generisanje odgovora uz pomoć domenskog eksperta odnosno nastavnika. On može obuhvatati i dodatni kontekst iz baze znanja kreirane na osnovu materijala priplavljenih od strane nastavnika. Pored dodatnog konteksta, treba dalje

razraditi i specifične instrukcije konverzionog agenta u vezi sa edukativnim materijalom za koji pruža podršku. Pored korišćenja UML profila *VJMKonverzacioniAgent*, u okviru faze analize predlažemo i korišćenje UML profila *PutanjaUčenja* definisanog u prethodnom poglavlju kao proširenje ove aktivnosti u okviru kojeg se definiše tok interakcije studenta sa platformom.

U kontekstu razvoja VJM edukativne platforme, u fazi projektovanja potrebno je dizajnirati način čuvanja istorijata konverzacija između studenta i četbota. Potrebno je odabrati bazu podataka u ove svrhe, koja može takođe biti korišćena i za čuvanje strukture edukativnog sadržaja pripremljenog od strane nastavnika, kao i samih odgovora studenata na aktivnosti ukoliko je to jedan od zahteva. Takođe je potrebno dizajnirati način čuvanja i upotrebe instrukcija. Pored toga, potrebno je dizajnirati način dopune znanja i konteksta četbota, što uključuje i izbor i popunjavanje odgovarajuće vektorske baze u okviru koje će se skladištiti dodatni materijali kreirani od strane nastavnika za implementaciju RAG tehnike, kao i način rada sa istom. Ove zadatke možemo realizovati takođe uz pomoć UML profila *VJMKonverzacioniAgent*. U ovoj fazi je potrebno projektovati i način praćenja sistema radi kontinuirane evaluacije u čemu može pomoći UML profil *Evaluacija*.

U fazi implementacije se vrši transformacija definisanih modela u implementacione komponente na osnovu automatskih preslikavanja ili manuelno. U ovoj fazi se programiraju definisani elementi edukativne platforme i vrši se njihovo raspoređivanje.

U fazi testiranja, utvrđujemo usklađenost sistema sa definisanim zahtevima. Kada je reč o razvoju VJM edukativne platforme, pored jediničnih i integracionih testova radi provere funkcionalnosti, kao i testiranja korisničkog interfejsa, potrebno je izvršiti i evaluaciju četbot odgovora na određenom setu očekivanih pitanja. Na početku projekta fokus je na testiranju izvodljivosti predloženog rešenja odnosno na generisanju tačnih odgovora [153]. Vremenom se ipak fokus može pomeriti na poboljšanje performansi sistema na osnovu povratnih informacija korisnika i potencijalnih problema koji su se pojavili. Samim tim, u ovoj disciplini može doći do potrebe za revidiranjem definicije konteksta radi poboljšanja performansi što sama priroda Unified procesa podržava jer predstavlja iterativni proces. U okviru ove faze moguće je korišćenje koncepta definisanih kroz profil *Evaluacija*.

Dakle, s obzirom na nedeterminističku prirodu ovakvih četbot sistema, važno je napomenuti da evaluacija četbot odgovora mora biti kontinuirana odnosno potrebno je nadgledanje evaluacije četbot odgovora od strane eksperata, kao i evaluacije od strane studenata. Na osnovu ovog fidbeka može se stvoriti potreba za ažuriranjem instrukcija i parametara četbota, dodatnog konteksta odnosno baze znanja, kao i rada sa istim kroz određene tehnike kako bi se poboljšale performanse sistema. U ovom procesu se mora voditi računa o tome da ažurirana verzija instrukcija i dodatni kontekst ne utiču negativno na prethodno dobro evaluirane izlaze četbota. Nakon svake izmene je potrebno obezbediti testiranje koje će to potvrditi.

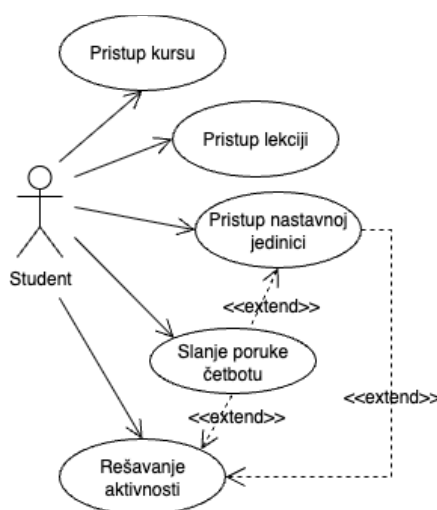
Ovde se takođe može spomenuti da se nakon određenog vremena korišćenja sistema od strane korisnika mogu uočiti interakcije za koje se može kreirati set podataka za fino podešavanje velikog jezičkog modela radi prilagođivanja pretreniranog modela željenim ponašanjem [123].

9 REALIZACIJA PROTOTIPA EDUKATIVNE PLATFORME SA KONVERZACIONIM AGENTOM ZASNOVANIM NA VELIKOM JEZIČKOM MODELU

U ovom poglavlju će biti navedena predložena arhitektura koja je korišćena za implementaciju dve verzije prototipa koje su testirane u okviru ove disertacije. Na početku će biti navedeni slučajevi korišćenja koji su implementirani radi evaluacije datog sistema. Zatim će biti dat prikaz moguće arhitekture predloženog sistema i opis njenih delova. Nakon toga će biti naveden opis korišćenih alata i tehnologija za implementaciju spomenutih delova prototipa sistema, kao i diskusija o alatima koji nisu korišćeni u konkretnom primeru, ali predstavljaju moguće alternative.

9.1 Korisnički zahtevi

Glavna uloga kojoj je namenjen predložen sistem za elektronsko učenje koji integriše VJM četbota je student. U okviru ovde definisanih korisničkih zahteva fokusiramo se na jednu od glavnih funkcionalnosti ovakvih sistema, a to je distribucija materijala za učenje. Student u okviru sistema može da pristupi kursu i odgovarajućoj lekciji u okviru kursa. Student može da se kreće kroz nastavne jedinice u okviru lekcije, i da pristupi aktivnostima za učenje. Takođe, može da pristupi četbotu odnosno da mu pošalje poruku ukoliko se pojavi potreba za dodatnim pojašnjenjima tokom prolaska kroz nastavne jedinice, odnosno tokom rešavanja aktivnosti. Radi potrebe očuvanja anonimnosti tokom korišćenja date platforme u okviru ovog istraživanja student generiše anonimni jedinstveni identifikator koji će moći jednostavno da rekreira prema uputstvu i koji sadrži lične podatke poznate jedino studentu, tako da ga samo on može prepoznati i ponovo generisati. Takođe, prva verzija prototipa nije obuhvatala slučaj korišćenja rešavanje zadatka u okviru same platforme odnosno čuvanje odgovora, dok se u drugoj verziji prototipa rešenja aktivnosti čuvaju radi dalje analize. Na dijagramu na slici 16 možemo videti spomenute korisničke zahteve. Takođe, radi evaluacije upotrebljivosti sistema, koja je prikazana u narednom poglavlju, drugi prototip implementira logovanje vremena i kretanja kroz nastavne jedinice. U nastavku su navedeni scenariji korišćenja.



Slika 16 - Dijagram slučajeva korišćenja

Scenario korišćenja Pristup kursu:

1. Korisnik pristupa platformi za elektronsko učenje
2. Sistem prikazuje formu za unos podataka za prijavu korisnika
3. Korisnik unosi potrebne podatke
4. Sistem prikazuje dostupne lekcije datog kursa

Scenario korišćenja Pristup lekciji:

1. Sistem prikazuje dostupne lekcije datog kursa
2. Korisnik pristupa izabranoj lekciji od ponuđenih
3. Sistem prikazuje početnu stranu date lekcije koja sadrži listu nastavnih jedinica, sadržaj prve nastavne jedinice i interfejs za interakciju sa četbotom sa odgovarajućim istorijatom poruka

Scenario korišćenja Pristup nastavnoj jedinici:

1. Korisnik bira nastavnu jedinicu kojoj želi da pristupi (klikom na njen naziv ili kroz dugme za prethodnu ili sledeću nastavnu jedinicu)
2. Sistem prikazuje sadržaj izabrane nastavne jedinice sa integrisanim četbotom i odgovarajućim istorijatom poruka

Scenario korišćenja Slanje poruke četbotu:

1. Korisnik unosi poruku namenjenu četbotu u polje za unos
2. Korisnik šalje poruku pritiskom na dugme za slanje ili pritiskom na Enter
3. Sistem prikazuje animaciju dok ne krene generisanje odgovora
4. Sistem prikazuje odgovor četbota korisniku

Scenario korišćenja Rešavanje aktivnosti:

1. Korisnik pristupa aktivnosti
2. Sistem prikazuje aktivnost učenja, formu za unos rešenja i četbota sa odgovarajućim istorijatom poruka
3. Korisnik unosi ili menja odgovor na aktivnost
4. Sistem privremeno čuva trenutni odgovor
5. Opciono: Korisnik može da se kreće kroz nastavne jedinice za vreme izrade aktivnosti (<<extend>> Pristup nastavnoj jedinici)
6. Opciono: Korisnik može da šalje pitanja četbotu za vreme izrade aktivnosti (<<extend>> Slanje poruke četbotu)
7. Korisnik predaje rešenje
8. Sistem traži korisniku potvrdu za predaju zadatka

9. Korisnik daje potvrdu za predaju zadatka

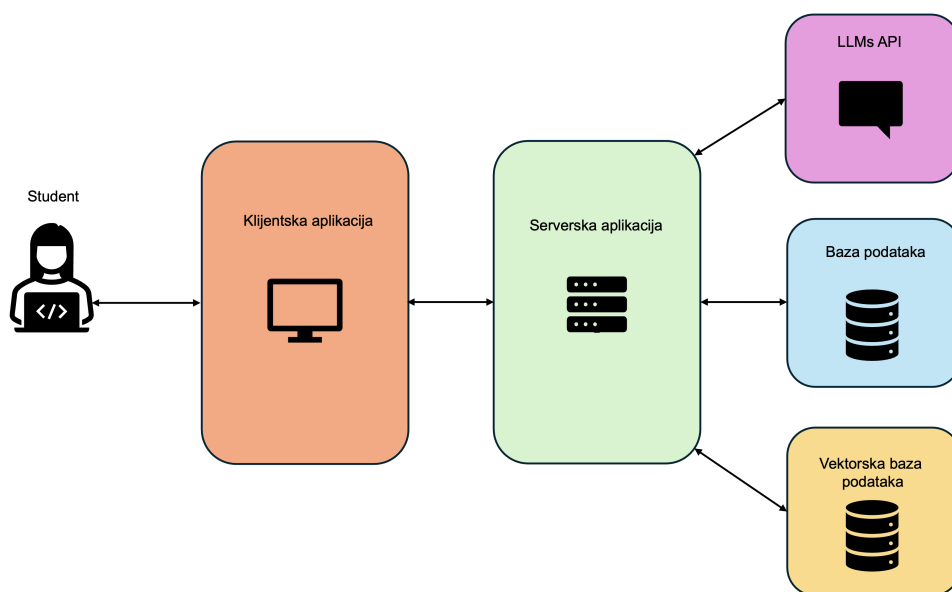
10. Sistem evidentira rešenje korisnika

Alternativni scenario:

9a. Korisnik odustaje od potvrde za predaju zadatka i vraća se na doradu (korak 3)

9.2 Arhitektura predloženog sistema

Na slici 17 prikazana je opšta arhitektura predloženog sistema koja se sastoji iz nekoliko delova. Arhitektura prati klijent-server paradigmu u kojoj je razdvojen razvoj klijentske aplikacije i korisničkih interfejsa sa kojima korisnik interaguje (*frontend*), od serverske (*backend*) koja pruža aplikacioni programski interfejs (API) klijentskoj aplikaciji za pristup podacima i poslovnoj logici.



Slika 17 - Opšta arhitektura predloženog sistema

Klijentska aplikacija se odnosi na *frontend* aplikaciju koja se izvršava u pregledaču na računaru. Implementira korisničke interfejse za interakciju sa lekcijama, odnosno jedinicama učenja, kao i sa četbotom. Kroz ovu aplikaciju korisnik interaguje sa predloženom edukativnom platformom. Serverska aplikacija se odnosi na *backend* aplikaciju koja uz komunikaciju sa bazama podataka i aplikacionim programskim interfejsom velikih jezičkih modela implementira i daje na korišćenje funkcionalnosti edukativne platforme klijentskoj aplikaciji kroz HTTP API koji prati REST (eng. *Representational State Transfer*) principe. Baza podataka koja može biti relacionala ili nerelacionala je potrebna radi čuvanja podataka o korisniku, njegovim konverzacijama, rešenjima aktivnosti, podešavanjima četbota, kao i radi čuvanja edukativnog sadržaja sa kojim će student interagovati. Jedan od mogućih načina čuvanja edukativnog sadržaja radi njegovog lakšeg prikaza u veb aplikaciji je kroz *Markdown* jezik koji se koristi za jednostavno formatiranje teksta. Takođe, za čuvanje edukativnog sadržaja kreiranog od strane nastavnika može se koristiti određeno *Cloud Storage* rešenje kao što je *Amazon Simple Storage Service* (Amazon S3) i sl. Vektorska baza podataka može biti uključena u slučaju potrebe dodavanja dodatnog konteksta na osnovu edukativnih materijala

nastavnika. Generiše se uz pomoć odgovarajućeg modela ugradnje (eng. *embedding model*) koji podatke prebacuje u njihovu vektorsku reprezentaciju. Kroz API odgovarajućeg provajdera ovih modela moguće je kreirati ugradnje (eng. *embeddings*), kao i transformisati korisničke upite za bolju pretragu i adekvatnu implementaciju RAG tehnike o kojoj će biti više reči u nastavku. Za generisanje odgovora na osnovu konteksta, ulaznih instrukcija i ulaza od strane studenta takođe je zadužen VJM API.

U nastavku će biti opisane konkretne tehnologije korišćene u implementiranom prototipu za potrebe ove doktorske disertacije.

9.2.1 Klijentska aplikacija

Angular [154] je izabran okvir za implementaciju klijentske strane aplikacije, kao jedan od rasprostranjenijih okvira za razvoj jednostranih aplikacija (pored npr. React-a [155] i Vue.js-a [156] koji su takođe opcije). Jednostrane aplikacije (*Single page application* – SPA) su frontend veb aplikacije koje sadrže jednu html stranicu čiji se sadržaj dinamički menja kroz JavaScript programski jezik. Angular koji danas koristimo, kao naslednik AngularJS okvira, je nastao u Google-u 2016. godine (počev od verzije 2). Predstavlja okvir za razvoj veb aplikacija čiji je osnovni gradivni element korisničkog interfejsa komponenta. Kroz mehanizme kao što su povezivanje podataka (eng. *data binding*) između logike i templejta komponente, ubacivanje zavisnosti (eng. *dependency injection*) za rad sa različitim servisima koji implementiraju logiku aplikacije, rutiranje za implementaciju navigacije kroz aplikaciju itd., pruža kompletan SPA okvir za kreiranje klijentskih aplikacija sa kojima korisnik može interagovati putem pregledača.

Određene odluke u vezi sa dizajnom korisničkog interfejsa donete su u skladu sa opštim principima za stvaranje pristupačnijih, jednostavnijih i intuitivnijih aplikacija preporučenih od strane Jacob Nielsena [157]. U ove principe se ubrajaju: vidljivost statusa sistema; mapiranje između sistema i sveta; korisnička kontrola i sloboda; konzistentnost i standardi; prevencija grešaka; prepoznavanje, a ne prisećanje; fleksibilnost i efikasnost korišćenja; estetika i minimalistički dizajn; pomoć prilikom prepoznavanja i oporavka od grešaka; pomoć i dokumentacija.

Korisnički interfejs prati principe responzivnog dizajna kako bi se osigurala upotrebljivost na različitim veličinama ekrana. Elementi zaduženi za prikaz sadržaja lekcije i interakciju sa četbotom su implementirani kao modularne komponente kako bi se omogućilo jednostavnije održavanje i lakoća proširenja.

Korisnički interfejs je minimalistički i fokusiran na smanjenje kognitivnog opterećenja (JN smernica - estetika i minimalistički dizajn). Jedinice učenja su prikazane sekvencijalno kako bi studentima bila omogućena interakcija sa svakom od njih posebno u jednom trenutku. U svakom trenu je vidljivo obeležena trenutno otvorena jedinica učenja u meniju sa leve strane strane (JN smernica – vidljivost statusa sistema). Korisnik može kroz dugmiće da navigira do sledeće ili prethodne jedinice učenja, kao i klikom na naziv jedinice učenja u meniju sa leve strane (JN smernica - korisnička kontrola i sloboda). Radi lakše pripreme sadržaja lekcija, odnosno jedinica učenja korišćen je markdown format čija se konverzija u html može jednostavno izvesti.

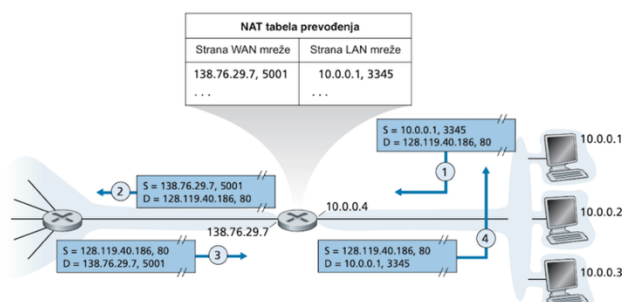
← Nazad Na Početnu

Mrežni sloj

- Petoslojni skup protokola na internetu
- Mrežni sloj
- IP adresa
- IP adresiranje
- Prosledjivanje i rutiranje
- Prevođenje mrežnih adresa

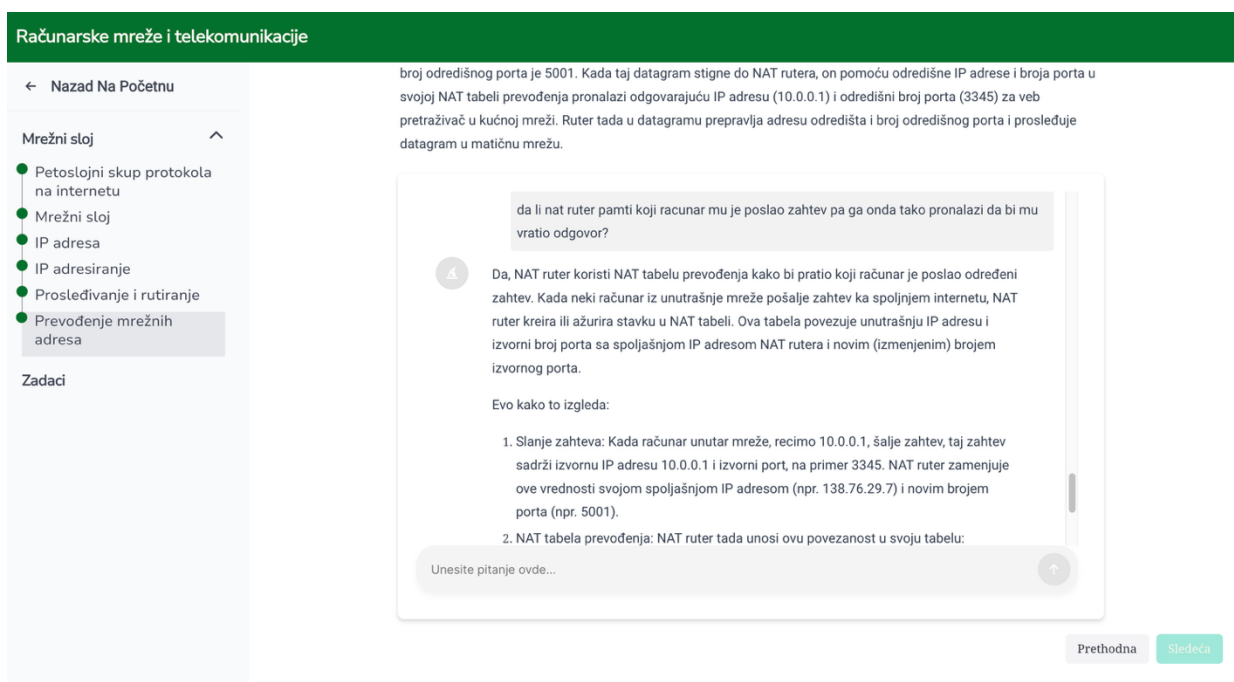
Zadaci

Na slici prikazan je način rada **NAT rutera**. Kućni NAT ruter ima interfejs koji je deo kućne mreže na desnoj strani slike. Adresiranje unutar kućne mreže istovetno je onome što smo videli ranije - sva četiri interfejsa kućne mreže imaju istu adresu podmreže, **10.0.0.0/24**. Adresni prostor **10.0.0.0/8** je jedan od tri dela prostora IP adresa koji su dokumentom [RFC 1918] predviđeni za privatne mreže ili prostor privatnih adresa, kao što je kućna mreža. Prostor privatnih adresa označava mrežu čije adrese samo uređajima unutar te mreže nešto znače. Da biste videli zašto je to važno, uzmite činjenicu da postoji na stotine hiljada kućnih mreža od kojih mnoge koriste isti adresni prostor, **10.0.0.0/24**. Uređaji unutar date kućne mreže mogu jedan drugome da šalju pakete. Međutim, jasno je da paketi koji se šalju preko granice kućne mreže, na veći globalni internet, ne mogu da koriste te adrese (bilo kao adresu izvora ili kao adresu odredišta), jer ima na stotine hiljada mreža koje koriste taj blok adresa. To znači da adrese **10.0.0.0/24** jedino u datoj kućnoj mreži nešto znače. Ali, ako privatne adrese imaju značenje samo unutar date mreže, kako se sprovodi adresiranje ako se paketi šalju na globalni internet, ili se primaju iz interneta gde adrese moraju da budu jedinstvene? Odgovor se nalazi u **NAT prevođenju**.



Slika 18 - Korisnički interfejs sadržaja jedinica učenja

Interfejs četbota je integrisan ispod sadržaja svake jedinice učenja kako bi studenti mogli da odmah da postavljaju pitanja bez potrebe za navigacijom do posebne sekcije. Izgled ove komponente je inspirisan postojećim poznatim okruženjima za vođenje konverzacija što omogućava studentima da interaguju sa četbotom na prirodan i intuitivan način (JN smernica - prepoznavanje, a ne prisećanje). Nakon što studenti unesu željeni tekst u za to određeno polje i pritisnu dugme za slanje, odgovor se prikazuje progresivno (strimuje se sa servera). Ukoliko je četbotu potrebno nekoliko sekundi za početak prikazivanja odgovora, korisniku će biti prikazana animacija sa tri tačke kako bi bio obavešten da je uspešno poslao pitanje i da je generisanje odgovora u toku (JN smernica - vidljivost statusa sistema). Studenti su takođe imali opciju da zaustave proces generisanja odgovora u bilo kom trenu pritiskom na stop dugme (JN smernica - korisnička kontrola i sloboda).



Slika 19 - Korisnički interfejs četbota

Takođe, druga verzija prototipa je sadržala i stranicu za pristup i rešavanje aktivnosti koje su za potrebe eksperimenta opisanog u narednom poglavlju studenti radili redom. Kao i kod nastavnih jedinica, ispod sadržaja aktivnosti i polja za unos odgovora, studentima je bio dostupan četbot tokom izrade zadatka. Studentima koji su imali pristup četbotu tokom izrade aktivnosti je za deo aktivnosti sugerisano da ga ne pitaju za direktna rešenja, već da ukoliko imaju potrebe postavljaju mu što detaljnija pitanja o konceptima iz lekcije koji su im potrebni za izradu.

← Nazad Na Početnu

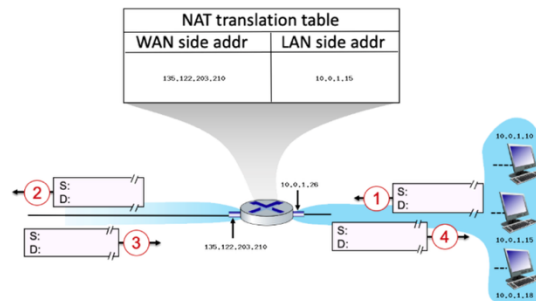
Mrežni sloj

- Petostojni skup protokola na internetu
- Mrežni sloj
- IP adresa
- IP adresiranje
- Prosleđivanje i rutiranje
- Prevođenje mrežnih adresa

Zadaci

Zadatak 7 od 7

Razmotrite sledeći scenario u kome se tri hosta, sa privatnim IP adresama 10.0.1.10, 10.0.1.15, 10.0.1.18, nalaze u lokalnoj mreži iza rutera sa NAT-om, koji se nalazi između ova tri hosta i šireg Interneta. IP datagrami koji se šalju sa ili su namenjeni za ova tri hosta moraju proći kroz ovaj ruter. Interfejs rutera na strani lokalne mreže ima IP adresu 10.0.1.26, dok adresa rutera na strani Interneta ima IP adresu 135.122.203.210. Pretpostavimo da host sa IP adresom 10.0.1.15 šalje IP datagram namenjen hostu 128.119.175.189. Izvorni port je 3377, a odredišni port je 80. Razmotrite datagram u koraku 2, nakon što ga je ruter prosledio. Koja je izvorna IP adresa ovog datagrama, a koja odredišna u koraku 2?



Unesi odgovor...

Pošalji

Slika 20 – Korisnički interfejs aktivnosti

9.2.2 Serverska aplikacija

Python predstavlja programski jezik koji je dominantan u oblasti veštačke inteligencije i većina najvažnijih okvira i biblioteka u ovoj oblasti pruža primarnu podršku za ovaj jezik. Node.js izvršno okruženje i JavaScript/TypeScript ekosistem prati trend razvoja aplikacija koje integrišu velike jezičke modele, tako da vodeće kompanije koje razvijaju ove modele, kao što su OpenAI, Anthropic, Google i sl. pružaju JavaScript SDK za rad sa njihovim modelima, kao i bitni okviri kao što su Langchain [158] i LlamaIndex [159]. U ovoj disertaciji, TypeScript je izabaran kao glavni programski jezik i za *frontend* i za *backend* deo aplikacije.

Express.js [160] je izabran okvir za implementaciju serverske strane aplikacije. Ona je zadužena za pružanje poslovne logike i podataka klijentskoj aplikaciji, odnosno upravlja podacima o studentima, njihovim konverzacijama i s druge strane može biti zadužena za edukativni sadržaj kroz komunikaciju sa odabranom bazom podataka za skladištenje ovih podataka. Takođe je zadužena za komunikaciju sa VJM API-jem, kao i za pretragu vektorske baze podataka.

Iako za izradu prototipa aplikacije implementirane u ovom doktoratu nije korišćen neki od specijalizovanih okvira za rad sa velikim jezičkim modelima, osim u delu za popunjavanje vektorske baze o čemu će biti više reči u nastavku, spomenućemo neke od trenutno najpopularnijih okvira koji mogu olakšati izradu kompletnih rešenja u ovom domenu. *LangChain* i *LlamaIndex* predstavljaju primere okvira koji se koriste prilikom implementacije aplikacija zasnovanim na velikim jezičkim modelima i mogu olakšati upravljanje memorijom, integraciju baza podataka, praćenje sistema i otkrivanje grešaka [161]. Oba okvira oficijalno podržavaju Python i Javascript/Typescript programske jezike. Pored *LangChain*-a koji se može posmatrati kao osnovno razvojno okruženje sa modularnim apstrakcijama kao što su lanci (eng. *chains*), alati (eng. *tools*),

mehanizmi za dovlačenje podataka i memorijski moduli, pogodno za brzi razvoj linearnih radnih tokova i RAG implementaciju [162], isti tim je razvio i *LangGraph* i *LangSmith*. *LangGraph* uvodi formalni sloj za orkestraciju uz upotrebu grafova za saradnju više agenata što ga čini pogodnim za zadatke sa nelinearnim tokom kontrole i adaptivnim donošenjem odluka. *LangSmith* se fokusira na monitoring, rešavanje grešaka i evaluaciju.

Veliki jezički model korišćen za generisanje odgovora u prototipu ove disertacije je OpenAI GPT-4o model [163] koji je u trenutku kreiranja prvog prototipa predstavljao najnoviji model ove kompanije. Predstavlja svestran, visoko inteligentan model koji podržava više jezika uključujući i srpski i pokazuje visoke performanse za različite zadatke. Parametar temperature modela koji kontroliše slučajnost izlaza ostavljen je na podrazumevanoj vrednosti 1,0 za generisanje odgovora na studentsko pitanje. Ova vrednost predstavlja balans između koherentnosti i kreativnosti što može biti pogodno za otvorene konverzacije u obrazovnom kontekstu. Međutim, na osnovu evaluacije sistema, nakon prikupljenih interakcija studenata data vrednost se može ažurirati.

Sistemske instrukcije

Sistemska instrukcija (eng. *system prompt*) predstavlja ulaz namenjen velikom jezičkom modelu radi usmeravanja generisanja odgovora na korisnički upit (eng. *user prompt*). Prompt inženjering predstavlja proces kreiranja i optimizovanja ulaznog prompta kako bi jezički model bio na odgovarajući način usmeren za generisanje odgovora u odnosu na zahteve korisnika [138]. U radu [139] su navedene sledeće tehnike koje mogu biti korišćene za kreiranje promptova od strane korisnika: *zero-shot* prompt koji je ne sadrži primere ulaza i izlaza (upita i odgovora) pogodan za dovlačenje činjenica, *few-shot* koji sadrži jedan ili više primera pogodan za klasifikaciju, lanac razmišljanja (eng. *chain-of-thought*) koji razlaže rasuđivanje na međukorake pogodan za matematičke i logičke zadatke, *least-to-most* koji kreira podzadatke po težini pogodan za planiranje i *ReACT* za zadatke koji uključuju eksterne alate. Rad autora [164] predstavlja detaljniju analizu tehnika promptovanja navodeći 58 tehnika na osnovu analize 1565 relevantnih radova.

Sistemska instrukcija uglavnom počinje definicijom uloge koja treba da bliže odredi sam način ponašanja velikog jezičkog modela tokom interakcije. Pored uloge, može definisati cilj četbota u datoj interakciji, ton i jezik odgovora, kao i pružiti još neke dodatne smernice. Može sadržati određena ograničenja za generisanje odgovora. Takođe, može odrediti i format odgovora (npr. json). Može sadržati i dodatni kontekst u odnosu na koji se treba generisati odgovor. Jedan od načina za pridruživanje dodatnog konteksta je kroz implementaciju RAG tehnike o kojoj će biti više reči u nastavku. Ukoliko se na ovaj način pridružuje dodatni kontekst, u sistenskoj instrukciji je potrebno definisati i pravila korišćenja dodatnog konteksta. Radi održavanja koherentnosti konverzacije, potrebno je uključiti i određen broj prethodnih upita i odgovora između korisnika i konverzionog agenta tokom ovog procesa ili uključiti sažetak konverzacije u okviru sistemske instrukcije ili kombinaciju prethodno dve navedene mogućnosti. U prvoj verziji prototipa ukupan broj tokena prethodnih upita i odgovora koji se prosleđuju velikom jezičkom modelu, je ograničen na maksimalno 2500 tokena uključujući i poslednji upit i sistensku instrukciju, a odgovor je ograničen na 500 tokena. U drugoj verziji koja uključuje i dodatni kontekst iz baze znanja koji se prosleđuje radi

generisanja odgovora, broj tokena prethodnih poruka ograničen je na 2000 tokena, bez ograničenja maksimalnog broja tokena odgovora.

Prompt korišćen u ovoj disertaciji spada u kategoriju *zero-shot* promptova kod kojih model generiše odgovor bez primera upita i odgovora u samoj instrukciji. Primer sistemskog prompta u prvom prototipu ove disertacije koji služi kao dokaz koncepta i čija evaluacija upotrebe je opisana u poglavlju 10 glasi:

Ti si stručni i angažovani asistent za kurs Računarske mreže i telekomunikacije. Tvoje glavne odgovornosti uključuju pomaganje studentima da razumeju koncepte mreža, odgovaranje na tehnička pitanja i jasno i sažeto objašnjavanje složenih tema. Treba da pružaš praktične primere i objašnjenja korak po korak za tehničke procese. Nemoj odgovarati na pitanja koja nisu vezana za kurs. Odgovaraj na istom jeziku na kom je postavljeno pitanje (engleski ili srpski). Odgovaraj u pasusima. Tvoj ton treba da bude profesionalan, ali pristupačan, tako da podstiče studente da istražuju i postavljaju pitanja. Prilagodi svoja objašnjenja nivou znanja studenta, koristeći jednostavniji jezik za početnike, a više tehničkog žargona za napredne studente. Sažetak trenutne lekcije za koju treba da pomažeš studentima i odgovaraš na njihova pitanja je: ...

U drugom prototipu, kroz koji smo izvršili evaluaciju upotrebljivosti sistema opisanu u poglavlju 10, je uključena i RAG tehnika te samim tim sistemski prompt sadrži i deo o načinu korišćenja dodatnog konteksta kao u nastavku:

Vi ste stručni i angažovani asistent na kursu Računarske mreže i Telekomunikacije. Vaš cilj je da pomžete studentima da razumeju mrežne koncepte, odgovarate na tehnička pitanja i jasno i sažeto objašnjavate složene teme.

=== Odgovornosti ===

- Objašnjavajte koncepte mreža i telekomunikacija jasno.*
- Pružajte praktične primere i objašnjenja tehničkih procesa postupno korak po korak.*
- NE odgovarajte na pitanja koja nisu vezana za kurs.*
- Objašnjenja treba da budu sažeta i fokusirana.*
- Ne preopterećujte studente sa previše informacija.*

=== Stil odgovora ===

- Uvek odgovarajte na jeziku na kojem je postavljeno pitanje (uglavnom srpski).*
- Pišite jasne, potpune pasuse.*
- Održavajte profesionalan, ali pristupačan ton, podstičući radoznalost i istraživanje.*

=== Sažetak trenutne lekcije ===

...

=== Preuzeti kontekst ===

Svako korisničko pitanje može biti praćeno dodatnim kontekstom, koji se prosleđuje modelu kao posebna system poruka. Koristite kontekst prvenstveno da utemeljite svoj odgovor tačnim i relevantnim informacijama. Možete dodati primere ili objašnjenja iz vaše šire baze znanja, sve dok ostajete u skladu sa kontekstom.

Smernice za korišćenje konteksta

- Kontekst tretirajte kao korisnu referencu, a ne kao striktno ograničenje.*
- Ne morate spomenuti sve informacije iz konteksta ukoliko ih je previše i nisu neophodne za odgovor.*
- Nikada ne protivurečite informacijama iz konteksta.*

- Ako postoji konflikt između konteksta i šireg znanja, prednost dajte kontekstu.
- Svaka dodatna informacija treba biti relevantna i činjenična.

Navedeni prompt je korišćen za konverzaciju o nastavnim jedinicama. Primeri teorijskih pitanja i odgovora od strane četbota, čija je sistemska instrukcija prethodno navedena, tokom prolaska kroz lekciju su dati u narednim tabelama. Možemo zaključiti da su odgovori zadovoljavajućeg kvaliteta. Evaluacija odgovora od strane nastavnika jednog od prototipa će biti navedena u narednom poglavlju.

Tabela 14 - Isečci konverzacija studenata i konverzionog agenta za lekciju

Kreator poruke	Sadržaj
Student:	Kako se adrese prevode iz privatnih u javne?
Četbot:	Prevođenje adresa iz privatnih u javne adrese se vrši putem tehnike nazvane NAT (Network Address Translation). Ovo je proces koji se najčešće realizuje na kućnom ili mrežnom ruteru, omogućavajući računarima sa privatnim IP adresama unutar lokalne mreže da komuniciraju s internetom koristeći jednu ili više javnih IP adresa. Evo kako to funkcioniše: Korišćenje privatnih IP adresa: - Uređaji unutar lokalne mreže koriste privatne IP adrese koje nisu unikatne na globalnom internetu. Ove adrese (npr. 10.0.0.0/24, 172.16.0.0/12, 192.168.0.0/16) važe samo unutar lokalne mreže. Zahtev za izlaznu komunikaciju: - Kada uređaj sa privatnom IP adresom želi da pošalje podatke na internet, on šalje datagram do lokalnog rutera. NAT proces prevođenja: Zamenjuje privatnu IP adresu: NAT ruter beleži izvornu privatnu IP adresu i broj porta datagrama u NAT tabelu prevođenja, i zamenjuje ih svojom javnom IP adresom i novim brojem izvornog porta. Nove informacije u zaglavlju: Ruter menja izvornu IP adresu u datagramu sa svojom javnom IP adresom. Takođe, promenjen je i broj izvornog porta u novi broj. Dodavanje u NAT tabelu: - NAT ruter dodaje novu stavku u svoju NAT tabelu prevođenja, koja povezuje privatnu IP adresu i izvorni port sa javnom IP adresom i novim brojem porta. Odziv od veb servera: - Kada veb server odgovara, paket stiže natrag do NAT rutera. Ruter tada koristi NAT tabelu da identifikuje uređaj koji je napravio početni zahtev, i ponovno prevede javnu IP adresu i broj porta natrag u privatnu IP adresu i originalni broj porta pre nego što prosledi paket nazad unutrašnjem uređaju. Ovaj proces omogućava uređajima sa ne-jedinstvenim privatnim IP adresama da deluju kao da koriste jedinstvenu javnu IP adresu, omogućavajući višestruke istovremene veze sa internetom putem NAT rutera.
Student:	Da li mozemo da prodjemo ove slike zajedno? 223.1.1.0. je adresa mreze... onda ide odredjeni broj n racunara.. i onda na kraju 223.1.1.(n+1) kao adresa za interfejs?
Četbot:	Naravno! Hajde da prodemo kako se adrese dodeljuju unutar podmreže.

	### Primer podmreže Podmreža: 223.1.1.0/24 #### Objašnjenje: 1. Mrežna adresa: - 223.1.1.0 je mrežna adresa i koristi se da identifikuje samu podmrežu. Nije moguće dodeliti ovu adresu pojedinačnom uređaju ili interfejsu unutar mreže. 2. Moguće IP adrese unutar podmreže: - Opseg IP adresa koje se mogu dodeliti uređajima u ovoj podmreži ide od 223.1.1.1 do 223.1.1.254. - Svaka adresa unutar ovog opsega može se dodeliti interfejsu uređaja ili računara. 3. Preostali brojevi n i n+1: - Ako imamo `n` računara ili uređaja koji su povezani ovoj podmreži, oni će, sve dok su brojevi adresa slobodni, zauzimati adrese kao što su 223.1.1.1, 223.1.1.2, ..., pa do 223.1.1.n. - Ako se koristi ruter ili još jedan uređaj unutar iste podmreže, može mu se dodeliti adresa 223.1.1.(n+1), sve dok, naravno, ostaju dostupne adrese unutar opsega. 4. Broadcast adresa: - 223.1.1.255 je broadcast adresa ove podmreže i koristi se za slanje paketa svim uređajima unutar te podmreže. Slično kao mrežna adresa, broadcast adresa se ne može dodeliti pojedinačnom uređaju. Dakle, osnovni niz koraka je da najpre dodelimo IP adrese za računare, uređaje ili interfejse, a zatim koristimo slova `n` kao promenljive za broj računara. U ovakvom tipu podmreže, možemo imati maksimalnih 254 korisnička uređaja.
Student:	Da li sam ruter nalazi put kojim se paket dalje salje prema odredisnoj adresi ili to radi vec prvi krajnji korisnik od koga je paket poslat?
Četbot:	Ruter je taj koji određuje put kojim se paket dalje šalje prema odredišnoj adresi, a ne krajnji korisnik od koga je paket poslat. Kada korisnik pošalje paket, on sadrži samo odredišnu IP adresu, a zadatak rutera je da odredi najbolju putanju do te odredišne adrese. ### Kako to funkcioniše: 1. Rutiranje: - Svaki ruter koristi algoritme rutiranja kako bi izgradio i održavao svoju tabelu prosleđivanja. Ovi algoritmi omogućavaju ruterima da dinamički utvrđuju putanje kroz mrežu, često uzimajući u obzir parametre kao što su broj skokova, širina pojasa, kašnjenje ili stvarno opterećenje linkova. 2. Prosleđivanje: - Kada paket stigne na ruter, on koristi svoju tabelu prosleđivanja da odluči na koji izlazni interfejs treba proslediti paket. Ova odluka je bazirana na odredišnoj IP adresi paketa. Dakle, svaki ruter duž putanje igra ključnu ulogu u određivanju sledećeg koraka za svaki paket. Krajnji korisnik inicijalnog paketa samo pruža odredišnu IP adresu, dok ruter koristi svoje algoritme i tabele da razgraniči kako će paket stići do svog konačnog odredišta. Ovaj proces omogućava mreži da efikasno upravlja resursima i obezbedi pouzdan prenos podataka.

U drugoj verziji prototipa, tokom izrade zadataka napravljenih od strane nastavnika kako bi se utvrdilo prethodno pređeno gradivo je takođe dostupan četbot. Kod dva zadatka su studenti smeli da direktno zatraže odgovor od četbota kako bi na taj način bilo simulirano pronalaženje specifičnih informacija (umesto pretrage materijala kretanjem kroz platformu), kroz postavljanje teorijskih pitanja. Četbot nije u instrukciji sadržao tačan odgovor na pitanje, ali je postigao 100% tačnost. Kod ostalih zadataka njegova sistemska instrukcija pored opštih smernica i pravila rada sa dodatnim kontekstom iz vektorske baze, sadrži i kontekst vezan za zadatak u izradi, kao i ograničenje da ukoliko student pošalje identičan tekst zadatka i traži rešenje, da četbot ne treba da da direktan odgovor, već da mu pruži smernice što je česta metoda koju nastavnici upotrebljavaju [165]. Ovaj deo instrukcije je dat u nastavku. Ideja je da student kroz razjašnjavanje teorije kroz postavljanje pitanja i kroz komunikaciju sa četbotom, u manjoj ili većoj meri u zavisnosti od potreba studenta, dolazi do rešenja zadatka.

=== Zadatak koji student sada rešava i preporučeni nagoveštaji ===

5. Pretpostavimo da između izvornog i odredišnog računara postoje tri rutera. Ako zanemarimo fragmentaciju, preko koliko interfejsa prelazi IP datagram koji se pošalje od izvornog do odredišnog računara?

- Pitaj: 'Koliko interfejsa obuhvataju izvorišni i odredišni računar, a koliko ruteri?'
- Definiši kroz koliko mrežnih interfejsa prolazi paket na računaru, a kroz koliko na jednom ruteru.
- Ohrabri studenta da krene da broji redom interfejse od izvora do odredišta, uključujući i odredište, ali ne spominji brojeve.

=== Postupanje kada student pošalje zadatak ili traži rešenje ===

Ukoliko student direktno zatraži rešenje ili pošalje zadatak koji se poklapa sa gore navedenim, ne daješ direktan odgovor.

Umesto toga:

1. Ljubazno obavesti studenta da je to jedan od zadataka koje treba da uradi samostalno.
2. Ohrabri ga da razmisli o osnovnim konceptima i ponudi da razgovarate o teoriji povezanoj sa temom, bez davanja direktnog rešenja.
3. Na primer: Izgleda da je to jedan od tvojih samostalnih zadataka. Ne mogu da dam direktan odgovor, ali mogu da ti pomognem sa svakom nedoumicom tokom izrade zadatka.

Primer studentskih pitanja i četbot odgovora vezanih za jedan od zadataka je dat u tabeli u nastavku.

Tabela 15 - Isečak konverzacije studenta i konverzionog agenta za aktivnost

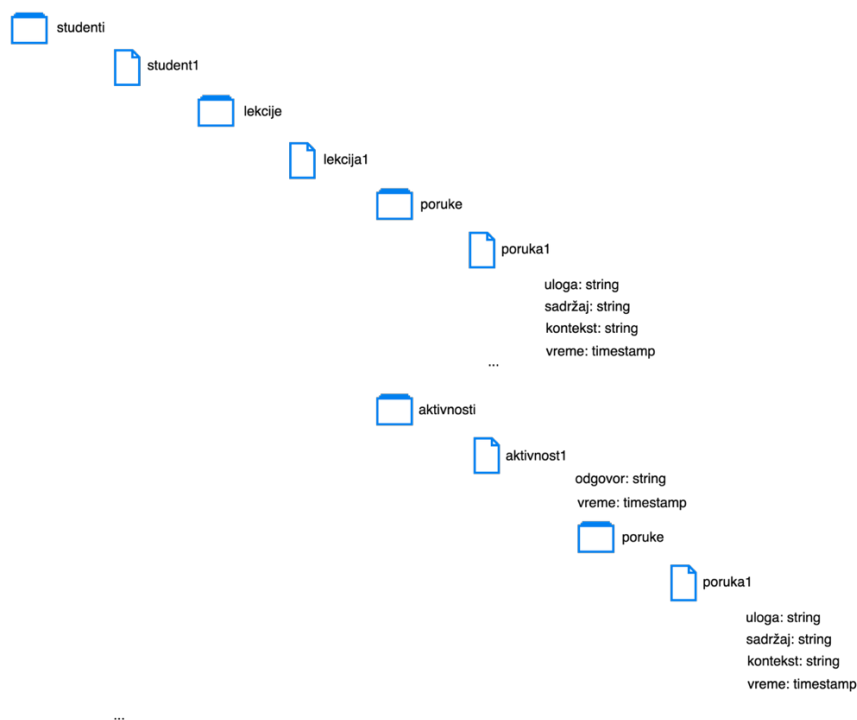
Student:	Pored interfejsa između svakog rutera, da li postoje jos neki?
Četbot:	Da, osim interfejsa između svakog rutera u mreži, mrežni interfejsi takođe postoje na izvoru i odredištu. Kao što smo ranije spomenuli, mrežni interfejsi se nalaze na svim uređajima koji su uključeni u mrežnu komunikaciju. To znači da: 1. Izvorni računar ima mrežni interfejs (npr. Ethernet ili WiFi adapter) preko kojeg inicira komunikaciju. 2. Odredišni računar takođe ima svoj mrežni interfejs preko kojeg prima podatke. Dakle, kada razmišljamo o putanji datagrama od izvornog do odredišnog računara, on

	prolazi kroz mrežne interfejske na izvoru, zatim kroz svaki od interfejsa na ruterima na putu, i konačno kroz interfejs na odredištu. Svaki od ovih interfejsa igra ključnu ulogu u procesu isporuke podataka.
Student:	Znaci između dva računara na istoj mreži, datagram prelazi 4 interfejsa? svoj izvor, izvorni-ruter, ruter-odredišni, pa onda na kraju odredišni?
Četbot:	Tvoj opis pokazuje da vrlo dobro razumeš kretanje datagrama kroz mrežu! Međutim, izgleda da razmatraš mrežu sa samo jednim ruterom između dva računara. Hajde da razjasnimo: - Izvorni računar: Prvi interfejs je mrežni interfejs izvornog računara, preko kojeg datagram ulazi u mrežu. - Ruteri u sredini: Svaki ruter koji datagram susreće na svom putu ima dva interfejsa za komunikaciju s njim – jedan za ulaz datagrama i jedan za izlaz datagrama. Ako datagram prolazi kroz više rutera, broji se po dva interfejsa za svaki ruter kroz koji prolazi. - Odredišni računar: Poslednji interfejs je mrežni interfejs odredišnog računara, preko kojeg datagram stiže na odredište. U tvom primeru sa dva računara na istoj mreži, ukoliko ne postoji nijedan ruter između njih, datagram bi prošao kroz dva interfejsa – jedan na izvornom i jedan na odredišnom računaru. Ako postoji jedan ruter, datagram bi prolazio kroz četiri interfejsa, a sa više rutera, broji se po dva interfejsa za svaki dodatni ruter.

Važno je napomenuti da date systemske instrukcije se tokom procesa evaluacije mogu dodatno prilagođavati dodavanjem konteksta ili pravila, radi podizanja performansi datog sistema, ali vodeći računa o samoj veličini prompta s obzirom na to da od njegove veličine, kao i od ostalih ulaza kao što su konverzacija i korisnički upit, zavisi cena.

9.2.3 Baza podataka

Kada je reč o bazi podataka u kojoj se čuvaju istorijati konverzacija između studenata i četbota po lekciji u prvom prototipu, i po lekciji i zadacima odnosno aktivnostima u drugom prototipu, kao i čuvanje rešenja zadataka radi dalje analize, korišćena je Firebase Firestore NoSQL baza podataka [166]. Firestore organizuje podatke kroz kolekcije i dokumente, pri čemu svaki dokument sadrži skup polja određenog tipa podatka. Zbog anonimnosti učešća u istraživanju i korišćenja date platforme, studenti su generisali sopstveni jedinstveni identifikator na osnovu uputstva, tako da nisu čuvani dodatni podaci o njima. Primer jedne moguće šeme dokumenata, a koja je korišćena u ovom prototipu, je data u nastavku na slici 21.



Slika 21 - Opšta šema Firestore kolekcije Studenti

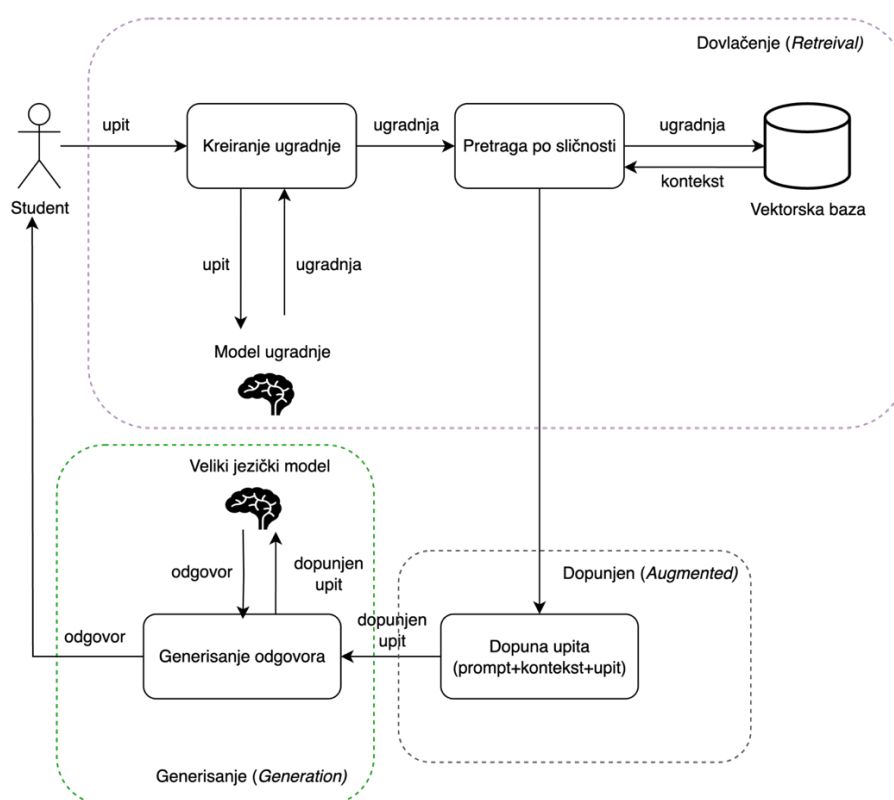
Dakle u okviru kolekcije studenti, za svakog studenta (dokument) postoji kolekcija lekcija u okviru koje se čuva kolekcija poruka za određenu lekciju odnosno nastavne jedinice (dokument), kao i kolekcija aktivnosti koja sadrži takođe kolekciju poruka konverzacije vođene za svaku aktivnost posebno (dokument). U prvom prototipu konverzacija sa četbotom je vođena na nivou cele lekcije koja je sadržala teorijski i praktični deo, te samim tim implementacija baze nije sadržala kolekciju aktivnosti. Dokument koji sadrži informacije o porukama, pored uloge (*user/assistant*), sadržaja, konteksta koji je korišćen za generisanje odgovora i vremena generisanja, može biti proširen i dodatnim podacima radi evaluacije sistema. Oni mogu obuhvatati broj ulaznih/izlaznih tokena, podatke o dodatnom kontekstu izvučenom iz vektorske baze o kojoj je će biti više reči u nastavku i sl. Takođe, u drugoj verziji prototipa čuvani su logovi (kolekcija) o navigaciji između jedinica učenja, kao i o pristupu aktivnostima sa vremenima i korisničkim identifikatorom radi evaluacije opisane u narednom poglavlju.

9.2.4 Vektorska baza podataka i RAG tehnika

Vektorska baza podataka predstavlja specijalizovanu bazu podataka u okviru koje se skladište višedimenzionalni vektori kao matematička reprezentacija podataka [167]. Za razliku od tradicionalnih baza kod kojih se pretraga vrši na osnovu tačnog poklapanja ili predefinisanih uslova, i koje često zanemaruju semantiku podataka, kod vektorskih baza se pretraga vrši na osnovu vektorske sličnosti koja obuhvata semantiku odnosno značenje podataka [168]. Predstavlja glavni deo *Retrieval Augmented Generation* (RAG) tehnike za skladištenje i dovlačenje dodatnog konteksta za ulaz u velike jezičke modele radi povećanja tačnosti i pouzdanosti njegovih izlaza. U nastavku će ukratko biti opisana

ova tehnika kroz njene faze razvoja prema kategorijama definisanim od strane Gao I saradnika [169].

Retrieval Augmented Generation predstavlja tehniku koja može potencijalno poboljšati performanse velikih jezičkih modela dopunjujući korisnički ulaz podacima relevantnim za kontekst zadatka koji treba da izvrši. Dati kontekst se dovlači pretragom prethodno kreirane i popunjene vektorske baze podataka upotrebom odgovarajućeg modela ugradnje (eng. *embedding model*). Vektorska baza može biti popunjena na osnovu dokumenata različitog formata koji se transformišu u tekst, a zatim u vektorske reprezentacije delova teksta. Kada korisnik pošalje svoju poruku namenjenu velikom jezičkom modelu, ta poruka se konvertuje u vektor i vektorska baza se može pretražiti prema sličnosti vektora kako bi se potencijalno pronašao koristan dodatni kontekst. Na osnovu toga, početna poruka korisnika može biti dopunjena kontekstualnim podacima iz vektorske baze i kao takva biti poslata modelu za generisanje odgovora. Dakle iz vektorske baze se dovlače odgovarajuće informacije (*retrieval*) koje treba da služe kao dopuna ulazu (*augmented*) radi povećanja performansi modela u odgovarajućem domenu i kao takve se šalju uz korisnički ulaz velikom jezičkom modelu radi generisanja odgovora (*generation*). Na slici 22 je prikazan prethodno opisan osnovni proces RAG tehnike.



Slika 22 – Osnovni proces RAG tehnike

Ovako definisana RAG tehnika predstavlja tzv. *Naive* RAG tehniku koja kao takva se suočava sa određenim izazovima tokom primene kao što su dovlačenje nerelevantnih podataka, a izostavljanje krucijalnih [169]. Zato Napredni RAG (eng. *Advanced* RAG) uvodi primenu određenih tehnika, u fazi pre dovlačenja podataka za poboljšanje same

strukture indeksiranja vektorske baze kao i transformacije samog korisničkog unosa, i nakon dovlačenja za ponovno rangiranje delova konteksta i njihovu kompresiju, radi prevazilaženja nekih od ograničenja. Modularni RAG kombinuje i proširuje principe *Naive* i naprednog RAG-a i pored sekvencijalnog procesa koji se vezuje za prethodno navedene tehnike, nudi i mogućnost iterativnog i adaptivnog dovlačenja. Uvodi specijalne komponente radi poboljšanja RAG procesa. Neke od ovih komponenti predstavljaju modul za pretragu različitih izvora podataka koji se prilagođava specifičnom scenariju, modul za rutiranje kroz različite izvore podataka za izbor optimalne putanje upita, modul za prilagođavanje RAG procesa različitim zadacima itd.

Tokom pregleda literature uočen je određen broj radova koji koristi RAG tehniku za poboljšanje performansi sistema i navodi korišćenu vektorsku bazu radi dopune konteksta u domenu edukativnih platformi namenjenih studentima ili profesorima. Konkretno vektorske baze odnosno sistemi za upravljanje vektorskim bazama koji su korišćeni u ovim radovima su: Chroma, FAISS, Milvus, MongoDB, PostgreSQL sa pgVector ekstenzijom, Qdrant, Vectara, i Weaviate. Skoro sve navedene baze predstavljaju rešenja otvorenog koda (osim Vectara-e) što ukazuje na preferencu ovog tipa rešenja. Baze su korišćene za skladištenje različitih tipova resursa za učenje, kao što su prezentacije, udžbenici, naučni radovi i ostali materijali sa kurseva. Važno je napomenuti da efektivnost ove tehnike zavisi od kvaliteta samih izvora podataka [170].

Nikolovski i saradnici [171] su koristili Qdrant vektorsku bazu za skladištenje i pristup prezentacijama, beleškama, primerima kodova i delovima udžbenika za kurseve u oblasti računarskih nauka. Dodatni kontekst preuzet iz ove baze je korišćen kao pomoć prilikom kreiranja, poboljšanja i evaluacije ispitnih pitanja korišćenjem Gemini 2.0 modela. Hennekeuser i saradnici [172] su kreirali sličnog asistenta korišćenjem Milvus vektorske baze i OpenAI modela ugradnje (textembedding-ada-002) za skladištenje lekcija i udžbenika i GPT-4 modela za generisanje odgovora u vezi predmeta u domenu poslovanja (poslovna statistika, informatika, upravljanje podacima i sl.). Jedan od slučajeva korišćenja ovog sistema je namenjen nastavnicima za lakši pristup resursima kursa što može pomoći kod poboljšanja materijala za kurseve (npr. generisanje ispitnih pitanja). Gopi i saradnici [173] su koristili Vectara vektorsku bazu i MathVista set podataka da kreiraju relevantnija test pitanja za studente inženjerstva uz GPT-4 model i kroz ovu tehniku su postigli manje od 5 procenta halucinacija.

Tophel i saradnici [170] su koristili FAISS bazu kreiranu uz pomoć OpenAI modela ugradnje za skladištenje sadržaja udžbenika za geotehničko inženjerstvo, kao i njenog priručnika sa rešenjima. Cilj je evaluacija performansi velikih jezičkih modela (GPT-4 i Llama-3) sa dodatnim kontekstom izvučenim iz baze u davanju odgovora na pitanja vezana za zadatke iz ove knjige što može pokazati potencijal korišćenja ovih modela kao pomoć u učenju studentima u ovoj oblasti. Zaključili su da RAG tehnika drastično poboljšava performanse sistema u rešavanju specijalizovanih inženjerskih problema (postize 95 procenta tačnosti).

Tabarsi i saradnici [121] su koristili MongoDB bazu sa vektorskom pretragom materijala sa kursa (PDF sa tekstom, tabelama i slikama) za pobjašanje GPT-4o odgovora na studentska pitanja o materijalima sa kursa u oblasti računarskih nauka. Zaključili su da ovako implementiran sistem generiše tačne odgovore koji su prilagođeni ovom kursu. MongoDB je takođe korišćen za čuvanje podataka o studentima i njihovim

konverzacijama, tako da je ovaj izbor pojednostavio arhitekturu sistema korišćenjem jedne platforme i za vektorsku pretragu i podatke o studentima i lekcijama. Neumann i saradnici [122] su implementirali četbota uz pomoć GPT-4 modela i Weaviate vektorskog rešenja uz pomoć OpenAI modela ugradnje materijala za učenje dostupnih u PDF formatu. Četbot omogućava studentima lakši pristup sadržaju za učenje i pomaže u razjašnjavanju nejasnoća u oblasti informacionih sistema. Postigao je tačnost od 88% u pogledu pružanja pomoći vezane za pitanja sa kursa.

Wolfel i saradnici [124] su upotreбили Chroma vektorsku bazu za kreiranje sistema koji radi sa prezentacijama i transkriptima lekcija, i GPT-3.5/4 modele, kako bi studentima pružili podršku tokom učenja na osnovu ovih materijala. Zaključili su da sistem postiže bolje performanse odnosno manje halucinacija kroz RAG tehniku u odnosu na neadaptirane modele. U istraživanju [174] je takođe korišćena Chroma vektorska baza za skladištenje medicinskih slučajeva iz *Multi-View Medical Evaluation* (MVME) seta podataka uz OpenAI model ugradnje, koji su korišćeni za potrebe simulacija pacijenta, radiologa, studenta i medicinskog eksperta od strane OpenAI velikih jezičkih modela i Claude3.5-Sonnet modela.

Dean i saradnici [175] su kreirali pomoćnika u istraživanju u oblasti računarskih nauka uz implementaciju RAG tehnike korišćenjem PostgreSQL sa pgVector ekstenzijom za vektorsku bazu podataka u okviru koje su skladišteni radovi kroz Hugging Face allMiniLM-L6-V2 model. GPT-4 model je korišćen za generisanje odgovora na studentska pitanja u zavisnosti od nivoa studija i demonstrirao je više performanse za studente master i doktorskih studija u odnosu na osnovne studije.

U nastavku će biti opisan način na koji je RAG tehnika implementirana u drugoj verziji prototipa ove disertacije. Zbog jednostavnosti *deploymenta* i mogućnosti besplatnog korišćenja za aplikacije manjeg obima kroz servis kao što je Supabase [176], izabrana je pgvector vektorska baza podataka odnosno ekstenzija PostgreSQL relacione baze podataka. Korišćen je OpenAI model ugradnje text-embedding-3-large za popunjavanje vektorske baze, kao i za kreiranje vektora na osnovu studentskih upita za pretragu vektorske baze. Korišćena je zvanična literatura sa predmeta Računarske mreže i telekomunikacija za popunjavanje vektorske baze. Odabrane lekcije u tekstualnom formatu su podeljene na delove kroz Langchain *RecursiveCharacterTextSplitter* sa podešavanjima veličine dela (eng. *chunk*) od 1000 karaktera i preklapanja od 200 i svakom delu su dodati na početku podnaslovi lekcije kojoj pripadaju. Tako kreirani dokumenti su kroz spomenuti model ugradnje transformisani u vektore koji su sačuvani u PostgreSQL bazi uz tekstualnu reprezentaciju i metapodatke u vidu naslova i podnaslova. Supabase odnosno PostgreSQL može biti dobar izbor ukoliko se za čuvanje podataka takođe bira relaciona baza, pa kroz jedno rešenje mogu biti realizovane obe baze podataka što može biti potencijalan pravac izmene i u sistemu ove disertacije. Supabase pruža mogućnost rada sa bazom kroz klijentsku biblioteku, kao i pozive PostgreSQL procedura odnosno funkcija kroz koje je implementirana pretraga vektorske baze po kosinusnoj sličnosti sa pragom od 0,5. Takođe, ima mogućnost integracije sa VJM okvirima kao što su prethodno spomenuti *Langchain* i *LlamaIndex*.

Kako bismo poboljšali performanse osnovnog procesa RAG tehnike (dovlačenje konteksta i generisanje odgovora sa dopunjenim upitom) pre same pretrage vektorske baze izvršena je transformacija odnosno proširenje samih upita studenata u formu

pogodnu za pretragu vektorske baze, ukoliko su isti nejasni ili nekompletni, radi poboljšavanja performansi njene pretrage. U te svrhe takođe mogu biti korišćeni veliki jezički modeli. Primer prompta koji je korišćen za generisanje dva pretraživačka upita, na osnovu studentskog unosa, za pretragu vektorske baze je dat u nastavku.

`Ti si asistent za query expansion u RAG sistemu. Predmet: Računarske mreže i telekomunikacije.

Pravila:

- *Analiziraj novo pitanje.*
- *Prethodnu konverzaciju koristi samo ako je novo pitanje kratko ili neodređeno (npr. "Zašto?", "Kako to funkcioniše?") da identifikuješ pojam na koji se odnosi.*
- *Ako je pitanje i uz istoriju nejasno, vrati status "NEMA_PROSIRENJA".*
- *Ako je pitanje jasno i ima smisla proširiti, generiši dva pretraživačka upita u udžbeničkom, deklarativnom stilu (bez kompletnih rečenica i bez zaključnih tvrdnji).*
- *Ako je pitanje kratko, ali se odnosi na osnovne pojmove iz Računarskih mreža i telekomunikacija, pretpostavi da treba generisati deklarativne pretraživačke upite.*
- *Ako pitanje nije jasno vrati status "NEMA_PROSIRENJA".*

Vrati ISKLJUČIVO validan JSON u jednom od sledeća dva formata.

Ako postoji proširenje:

```
{  
  "status": "OK",  
  "queries": [  
    "detaljan proširen upit",  
    "alternativni upit"  
  ]  
}
```

Ako ne postoji:

```
{  
  "status": "NEMA_PROSIRENJA"  
}
```

Dati sistemski prompt takođe sadrži ulogu, kontekst za koji predmet transformiše upite, pravila odnosno smernice ponašanja, kao i format samog odgovora prikazan kroz primer JSON objekta.

Nakon pretrage tokom koje su po kosinusnoj sličnosti sa pragom od 0,5 dovučeni po maksimalno 6 dokumenata po upitu, korišćena je metoda fuzije recipročnih rangova [177], [178] kako bi isti dokumenti koji su rezultat pretrage više pretraživačkih upita postali važniji. Takođe, veliki jezički model može biti korišćen i za ponovno rangiranje dobijenih rezultata iz vektorske baze. Prosledili smo mu maksimalno 5 dokumenata sa najvećim rangom za koje je trebalo odredi dva najrelevantnija u odnosu na korisnički upit koji će biti prosleđeni kao dodatni kontekst velikom jezičkom modelu. Primer instrukcije korišćene za ponovno rangiranje relevantnosti je dat u nastavku.

Rangiraj dokumente po relevantnosti za pitanje.

Vrati ISKLJUČIVO JSON:

```
{ "ranked_ids": [id1, id2] }
```

Pravila:

- "ranked_ids" mora sadržati tačno 2 ID-ja
- ID-jevi moraju biti iz ponuđene liste
- bez duplikata

U ovim slučajevima, treba pažljivo izabrati model za ove zadatke, s obzirom na to da se cena povećava kroz dodatno korišćenje ovih modela pre i nakon generisanja konačnog odgovora. U ove svrhe u ovoj disertaciji izabran je gpt-4o-mini model [179].

Primer konteksta dobijenog na osnovu prethodno opisane RAG tehnike je dat u nastavku.

Tabela 16 – Primer pitanja, dodatnog konteksta i odgovora

Pitanje: Koja je veličina uobičajenog zaglavlja IP datagrama?	Kontekst: # 4 Mrežni sloj ## 4.4 Internet protokol (IP): prosleđivanje i adresiranje na internetu ### 4.4.1 Format datagrama Obratite pažnju na to da zaglavlje IP datagrama ima ukupno 20 bajtova (pod pretpostavkom da nema opcija). Ako datagram prenosi TCP segment, tada svaki (nefragmentirani) datagram u svom zaglavlju prenosi ukupno 40 bajtova (20 bajtova IP zaglavlja i 20 bajtova TCP zaglavlja) pored poruke aplikativnog sloja. --- # 4 Mrežni sloj ## 4.4 Internet protokol (IP): prosleđivanje i adresiranje na internetu ### 4.4.1 Format datagrama
Odgovor: Veličina uobičajenog zaglavlja IP datagrama je 20 bajtova, pod pretpostavkom da nema dodatnih opcija.	Ključna polja datagrama protokola IPv4 su sledeća (broj verzije, dužina zaglavlja, vrsta usluge, dužina datagrama, vreme života, protokol, kontrolni zbir zaglavlja, IP adrese izvora i odredišta, opcije, podaci): *Broj verzije. Ova četiri bita navode verziju protokola IP određenog datagrama. Prema broju verzije ruter određuje kako bi trebalo da protumači ostatak IP datagrama. Različite verzije protokola IP koriste različite formate datagrama. Format datagrama za trenutnu verziju protokola IP, IPv4. Format datagrama za novu verziju protokola IP (IPv6) opisan je pri kraju ovog odeljka. *Dužina zaglavlja. Pošto datagram IPv4 može da sadrži promenljivi broj opcija (koje se nalaze u zaglavlju datagrama IPv4) ova četiri bita su potrebna da bi se odredilo mesto gde u IP datagramu počinju sami podaci. Većina IP datagrama ne sadrži opcije, tako da uobičajeni IP datagram ima zaglavlje od 20 bajtova.

Naravno, kao što je već spomenuto, prethodno opisane tehnologije, način rada sa istim, formulisane instrukcije, izabrani modeli i parametri, predstavljaju jedno od mogućih rešenja koje se može dalje prilagođavati u zavisnosti od zahteva, kao i poboljšavati nakon testiranja prototipa i na osnovu prikupljenih studentskih pitanja.

10 EVALUACIJA

U naredna dva poglavlja biće opisani sprovedeni eksperimenti u okviru kojih su studenti imali priliku da koriste dve verzije prototipa edukativne platforme nastale kao jedan od rezultata ove doktorske disertacije, radi njihove evaluacije. Njihova implementacija je opisana u prethodnom poglavlju. Prvi prototip testira izvodljivost korišćenja implementirane platforme sa konverzacionim agentom zasnovanim na VJM-u kao pomoćnika pri učenju predmeta Računarske mreže i telekomunikacije. Rezultati navedeni u nastavku su objavljeni od strane autora disertacije sa koautorima pod naslovom *Educational Platform in Higher Education With LLM-Driven Chatbot for Computer Networks and Telecommunications Course* u časopisu *Computer Applications in Engineering Education* (Tom 34, izdanje 1) [180]. Evaluacija drugog prototipa se fokusira na upotrebljivost ovog sistema za učenje radi provere hipoteze navedene u ovoj disertaciji.

Oba eksperimenta su sprovedena na Univerzitetu u Beogradu – Fakultetu organizacionih nauka tokom zimskog semestra akademske 2024/2025. godine i 2025/2026. godine respektivno. Učesnici su bili studenti sa smera Informacioni sistemi i tehnologije koji slušaju predmet Računarske mreže i telekomunikacije na trećoj godini osnovnih akademskih studija. Evaluacija obrazovne platforme nakon njenog korišćenja sprovedena je putem anonimnih onlajn anketa na srpskom jeziku opisanih u nastavku poglavlja. Nije bilo nedostajućih niti nevažjećih odgovora, budući da su sva pitanja u upitnicima bila obeležena kao obavezna. Statistička analiza izvršena je korišćenjem softverskog paketa Statistical Package for the Social Sciences (SPSS), verzija 23.

10.1 Evaluacija prvog prototipa

Kao što je spomenuto, u ovom poglavlju će biti predstavljen sproveden eksperiment za evaluaciju upotrebe prve verzije prototipa nastalog kao jednog od rezultata ove disertacije čiji su ključni aspekti implementacije prikazani u prethodnom poglavlju. Biće prikazani rezultati evaluacije platforme od strane studenata, kao i evaluacija četbot odgovora od strane nastavnika na ovom predmetu.

10.1.1 Metodologija istraživanja

Evaluacija platforme je sprovedena za vreme jednog časa na predmetu Računarske mreže i telekomunikacije i trajala je oko 90 minuta. Fokus lekcije je bio na HTTP protokolu. S obzirom na to da je eksperiment sproveden u toku nastave, na početku časa su saradnici na predmetu održali kratak uvod u lekciju u trajanju od 30 minuta. Nakon toga su studenti dobili pristup materijalima na implementiranoj platformi. Eksperimentalna grupa je imala pristup i integrisanoj četbot funkcionalnosti kao pomoć pri učenju za razjašnjavanje nedoumica, dok je kontrolna grupa mogla postavljati pitanja saradnicima. Svi studenti u eksperimentalnog grupi su interagovali sa četbotom na srpskom jeziku, osim dva studenta koji su započeli interakciju na engleskom, a zatim su se nakon jednog ili dva pitanja prebacili na srpski jezik.

Materijali za učenje dostupni putem platforme obuhvatali su kako teorijske, tako i praktične sadržaje, za koje se od studenata očekivalo da isprate na osnovu uvodnog predavanja održanog u prvom delu sesije. Nakon što su prešli sve materijale, od

studenata je zatraženo da urade kratku proveru znanja, a zatim i da popune anketu radi evaluacije četbota (u okviru eksperimentalne grupe). Ovaj pristup je imao za cilj procenu ishoda učenja, kao i percepcije studenata o korisnosti integrisane četbot funkcionalnosti.

U istraživanju je učestvovalo ukupno 174 studenta treće godine, od kojih je 71 koristilo verziju platforme sa četbotom, a 103 verziju bez. Pregled polne strukture ispitanika u ove dve grupe prikazan je u Tabeli 17. Rezultati testa verodostojnosti odnosa (eng. *Likelihood Ratio* test) ($\chi^2 = 0,038$, $df = 1$, $p = 0,846$) ukazuju na to da ne postoji statistički značajna razlika u polnoj strukturi između eksperimentalne i kontrolne grupe. Na osnovu toga može se zaključiti da su grupe ujednačene u pogledu polne zastupljenosti.

Tabela 17 - Raspodela učesnika prema polu

Eksperimentalna grupa	N	%	<i>Likelihood ratio</i>
Muški	30	42,25	$\chi^2=0,038$ $p=0,846$
Ženski	41	57,75	
Kontrolna grupa			
Muški	42	40,78	
Ženski	61	59,22	

Većinu učesnika činili su ispitanici ženskog pola (58,62%), što ne ukazuje na pristrasnost uzorka, imajući u vidu veću zastupljenost žena u studentskoj populaciji u Srbiji i na Univerzitetu u Beogradu [181], što je takođe slučaj i u predstavljenom istraživanju u uvodnim poglavljima o faktorima koji utiču na prihvatanje ChatGPT-ja u obrazovnom kontekstu, kao i u drugom eksperimentu koji će biti opisan u narednom poglavlju.

Radi provere jednakosti ove dve grupe analizirana su dva pokazatelja akademskog uspeha: broj ostvarenih bodova prema Evropskom sistemu prenosa i akumulacije bodova (ESPB) i prosečna ocena. Kolmogorov–Smirnov test primenjen je radi ispitivanja normalnosti raspodele obe varijable, ESPB bodova i prosečne ocene, pri čemu je utvrđeno da prosek i ESPB ne prate normalnu raspodelu. Samim tim, u narednom koraku je za poređenje grupa primenjen neparametarski Man-Vitnijev U test. Rezultati, prikazani u Tabeli 18, na osnovu dobijenih p-vrednosti i srednjih rangova, ukazuju na to da ne postoje statistički značajne razlike između eksperimentalne i kontrolne grupe u pogledu prethodnog akademskog uspeha, merenog ESPB bodovima i prosečnom ocenom.

Tabela 18 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnijevog U testa za ESPB i prosek u eksperimentalnoj i kontrolnoj grupi

	Srednja vrednost	SD	K-S	p vrednost	Grupa	Srednji rangovi	Z	p vrednost
ESPB	108,68	14,55	0,219	0,000	Eksperiment.	86,89	-0,139	0,889
					Kontrolna	87,92		
Prosek	8,01	0,58	0,075	0,019	Eksperiment.	91,03	-0,767	0,443
					Kontrolna	85,07		

Evaluacija

Evaluacija predložene platforme sprovedena je u dve faze, uz prikupljanje povratnih informacija od strane njenih korisnika odnosno studenata, kao i od stručnjaka odnosno nastavnika. Ovaj dvostruki pristup proceni korišćen je i u ranijim istraživanjima efikasnosti četbotova [182], omogućavajući sveobuhvatnu procenu iz različitih perspektiva.

Nakon završetka nastavnih sesija, studenti su radili test u obe grupe, dok su upitnik za evaluaciju četbot funkcionalnosti popunili studenti u eksperimentalnoj grupi. Cilj naknadnog testa bio je da se procene ishodi učenja i uporedi učinak eksperimentalne i kontrolne grupe. Test se sastojao od osam zadataka osmišljenih da provere zapamćeno tokom sesije učenja. Pored toga, studenti iz eksperimentalne grupe popunili su anketu o zadovoljstvu, posebno usmerenu na funkciju četbota. Konstruisana anketa za procenu sadržala je šest stavki čiji je cilj bio da se utvrdi kako studenti doživljavaju korisnost četbota i njegov ukupan doprinos iskustvu učenja. Stavke razvijene za procenu percepcije studenata o doprinosu četbota njihovom učenju konceptualno su prilagođene na osnovu konstrukta Očekivane performanse iz Jedinstvene teorije prihvatanja i upotrebe tehnologije (eng. *Unified Theory of Acceptance and Use of Technology* - UTAUT) [183] i izmenjene tako da odgovaraju kontekstu ovog istraživanja. Konkretno, stavke su preformulisane kako bi odrazile način na koji je četbot pomagao korisnicima u postizanju ishoda vezanih za učenje, kao što su razjašnjavanje teorijskih koncepata, otklanjanje nedoumica, izrada zadataka i produbljivanje znanja. Ovaj pristup je u skladu sa definicijom očekivanih performansi, koja predstavlja stepen u kojem pojedinci veruju da korišćenje određenog sistema poboljšava njihov učinak. Slična prilagođavanja konstrukta očekivanih performansi radi procene četbotova u obrazovnom okruženju mogu se naći u novijim studijama [184], [185]. Pored spomenutih stavki, uključena je i jedna stavka posebno namenjena proceni jasnoće odgovora četbota, budući da se ona direktno odnosi na jedan od ključnih elemenata upotrebljivosti četbota za vođenje konverzacije [186].

Eksperti su ocenjivali odgovore koje je generisao četbot na osnovu dva najvažnija aspekta sadržaja generisanog pomoću veštačke inteligencije, tačnosti i kvaliteta izlaznih podataka [112]. Generisane odgovore ocenjivali su prema pet kriterijuma [112]: tačnost kao mera ispravnosti sadržaja koji generiše veliki jezički model; relevantnost odgovora četbota u odnosu na konkretan upit studenta; kompletnost kao stepen u kojem odgovor sadrži sve bitne informacije neophodne za adekvatan odgovor na određeni upit; koherentnost koja se odnosi na to koliko je jasno strukturiran odgovor i njegov logički tok u skladu sa kontekstom upita i istorijatom razgovora; i fluentnost koja se odnosi na gramatičku ispravnost i tečnu, prirodnu upotrebu jezika.

10.1.2 Rezultati

Tokom sesije zabeleženo je ukupno 636 interakcija između studenata i četbota. Jedna interakcija je definisana kao pojedinačni unos korisnika praćen odgovorom četbota. Od 636 interakcija, 633 su bile na srpskom jeziku, a tri na engleskom. Analiza odgovora četbota pokazala je da se većina interakcija (62,11%) odnosila na objašnjenja teorijskih koncepta, dok je 33,96% bilo vezano za pomoć u vezi sa praktičnim delovima nastavnog materijala, kao što je prikazano u tabeli 19. U istoj tabeli predstavljeni su i primeri pitanja studenata. Važno je napomenuti da, iako neke interakcije nisu eksplicitno klasifikovane kao praktične, jer su bile formulisane na način koji više odgovara pitanjima vezanih za teoriju, one se ipak mogu dovesti u vezu sa praktičnim sadržajem. Ovo preklapanje je očekivano, s obzirom na to da se teorijsko znanje često direktno primenjuje u okviru praktičnih delova ove lekcije. Mali udeo interakcija (3,93%) nije bio povezan sa osnovnim ciljevima učenja, uključujući pozdrave, izraze zahvalnosti, pitanja o implementaciji četbota i sl. Sveukupno, studenti su koristili četbot kao podršku kako za razumevanje teorije, tako i za primenu sadržaja kursa. Budući da broj interakcija između studenata i četbota ne pruža informacije o njihovom kvalitetu [54], studenti su takođe popunili anketu u kojoj su ocenjivali pomoć koju je četbot pružio, dok su eksperti ocenjivali odgovore četbota prema definisanim kriterijumima.

Tabela 19 - Raspodela vrsta interakcija sa četbotom i primeri pitanja

Kategorija	Primer	N	%
Teorijska	dakle, kada se šalje zahtev ide se od aplikativnog ka fizickom sloju, a kada se zahtev prihvata onda se razlaze od fizickog ka aplikativnom slou; mozes li jos malo da mi objasnis opšti format HTTP zahteva; kako izgleda json odgovor; sta je 304;	395	62,11
Praktična	šta je /rmt/mmklogo.html u zahtevu?;	216	33,96

	kad radim snimanje preko wiresharka, spremim link, i kad pokrećem snimanje u wiresharku?; kako da uočim parove za zahtev/odgovor?; Šta je zahtevano od resursa u svakom od zahteva? kako ovo da vidim;		
Ostalo	zdravo; hvala; koliko kosta.	25	3,93

Dok tabela 19 prikazuje unose korisnika i raspodelu interakcija po različitim kategorijama, tabela 20 pruža uvid u primere odgovora četbota (po jedan za svaku kategoriju). Odgovori četbota su uglavnom bili dobro strukturirani i koherentni. Pitanja u ovoj studiji prvenstveno su se odnosila na HTTP protokol i njegovo razumevanje, što predstavlja standardnu oblast znanja u okviru računarskih mreža i telekomunikacija. U tom kontekstu, kada je reč o teorijskim pitanjima, četbot je u većini slučajeva uspeva da generiše tačne i relevantne odgovore, zasnovane na ustaljenim konceptima i opšteprihvaćenim informacijama. Korisnost odgovora opala je kod praktičnih pitanja u situacijama kada su studenti bili neprecizni prilikom formulisanja pitanja ili navođenja neophodnih detalja. Kada su korisnici postavljali pitanja koja nisu bila u vezi lekcije, četbot je efikasno preusmeravao tok razgovora i vraćao fokus na relevantnu materiju. Sveukupno, odgovori pokazuju sposobnost četbota da pruži jasna objašnjenja standardnih tehničkih koncepata, uz istovremeno adekvatno postupanje sa unosima koji odstupaju od teme. U narednim poglavljima biće više reči o uočenim razlikama između odgovora na teorijska i praktična pitanja.

Tabela 20 - Primeri interakcija između studenta i četbota po kategorijama pitanja

Kategorija	Parovi pitanje/odgovor
Teorijska	Student: koja je veza između partizana i http protokola Asistent: Dok Partizan kao sportski klub i HTTP protokol možda nemaju direktnu vezu, možemo ih povezati kroz kontekst upotrebe interneta za praćenje sportskih događaja. Na primer, kada želite da pratite najnovije vesti, rezultate ili utakmice Partizana putem interneta, vaš web pretraživač koristi HTTP protokol da bi komunicirao sa web serverima koji hostuju te informacije. Evo kako to funkcioniše: HTTP Protokol: HyperText Transfer Protocol (HTTP) omogućava prenošenje podataka između klijenta (web pretraživača) i servera

	koji pruža tražene informacije. 2. Slanje Zahtjeva: Kada ukucate URL web sajta koji prati Partizanove utakmice ili informacije, vaš web pretraživač šalje HTTP GET zahtev serveru. 3. Prijem Odgovora: Server, nakon obrade zahteva, vraća HTTP odgovor koji može sadržavati HTML stranice, slike, video snimke ili druge medije. 4. Prikaz Informacija: Web pretraživač interpretira prijemni HTTP odgovor i prikazuje informacije na ekranu, omogućavajući vam da pratite vesti, rezultate ili gledate prenose uživo. Kao što vidite, HTTP protokol igra ključnu ulogu u omogućavanju pristupa informacijama o Partizanu putem interneta. Ako imate dodatna pitanja o tome kako HTTP funkcioniše ili nešto drugo vezano za mreže, slobodno ih postavite.
Praktična	Student: šta je /rmt/mmklogo.html u zahtevu? Asistent: `/rmt/mmklogo.html` je putanja resursa na serveru. Kada pregledavate web stranicu i učitavate sadržaj putem HTTP zahteva, URL može biti razložen u nekoliko delova. Evo tih delova objašnjenih sa posebnim fokusom na putanju resursa: Struktura URL-a Primer: https://www.primer-sajta.com/rmt/mmklogo.html 1. https:// - Ovo je šema ili protokol koji se koristi za komunikaciju. U ovom slučaju, to je HTTPS, što znači da se koristi siguran HTTP. 2. www.primer-sajta.com - Ovo je domen ili adresa servera na kojem se resurs nalazi. Pretraživač koristi ovaj deo za povezivanje sa tačnim serverom. 3. /rmt/mmklogo.html - Ovo je putanja resursa na serveru: - /rmt/ - Poddirektorijum ili folder na serveru. - mmklogo.html - Naziv fajla koji se zahteva, u ovom slučaju HTML dokument nazvan mmklogo.html. Funkcija Putanje Resursa Kada vaš preglednik šalje HTTP zahtev, on traži specifičan resurs koristeći ovu putanju. Server, primivši zahtev, traži fajl mmklogo.html unutar direktorijuma /rmt/ na serveru. Ako server pronađe taj fajl, pošalje ga natrag kao HTTP odgovor, a vaš preglednik ga prikazuje. U osnovi, putanja poput /rmt/mmklogo.html pomaže serveru da identifikuje gde je smešten konkretan resurs koji je klijent (preglednik) zahtevao.
Ostalo	Student: koliko kosta

	Asistent: Žao mi je, ali ne mogu da odgovorim na to pitanje, jer nije u vezi sa lekcijom o kompjuterskim mrežama i telekomunikacijama. Međutim, ako imate pitanja o klijent-server arhitekturi, HTTP protokolu, ili bilo kojoj drugoj temi iz naše lekcije, slobodno ih postavite! Tu sam da pomognem oko tih tema.
--	--

Evaluacija četbota od strane studenata

Kao što je već navedeno, studenti iz eksperimentalne grupe ocenjivali su funkcionlanost četbota na platformi popunjavanjem upitnika sa šest stavki. Pre predstavljanja deskriptivne statistike dobijenih ocena, sprovedena je provera interne validnosti i pouzdanosti. Radi validacije skale od šest stavki, izračunati su Kronbahov alfa koeficijent, prosečna ekstrahovana varijansa (eng. *Average Variance Extracted* - AVE) i kompozitna pouzdanost (eng. *Composite Reliability* - CR), u skladu sa preporukama iz literature [187], [188]. Dobijene vrednosti su sledeće: Kronbahov alfa koeficijent = 0,818, AVE = 0,576 i CR = 0,884. Budući da dobijene vrednosti prelaze granične pragove (Kronbahov alfa koeficijent > 0,7, AVE > 0,5 i CR > 0,7), može se zaključiti da je predložena skala pouzdana za merenje percepcije studenata o podršci koju pruža četbot tokom procesa učenja.

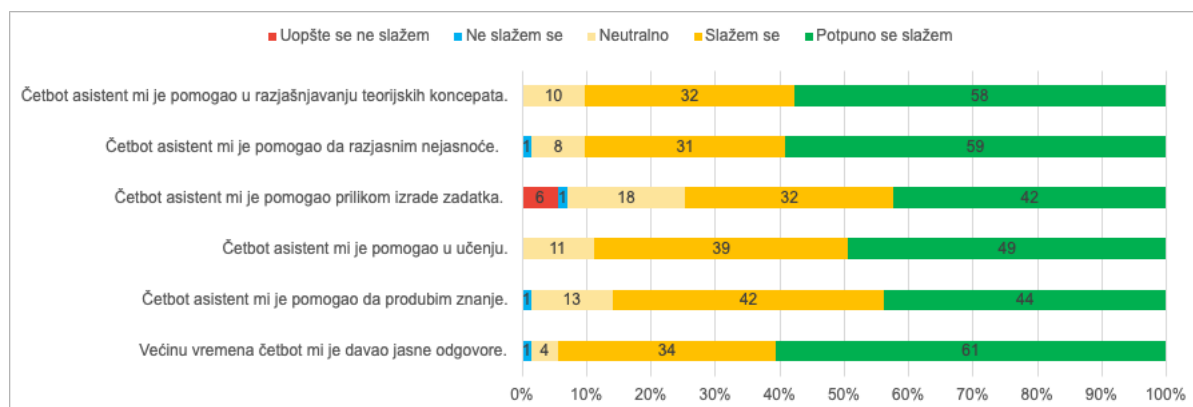
Tabela 21 sadrži deskriptivnu statistiku, uključujući aritmetičku sredinu, standardnu devijaciju, medijanu i procenat studenata koji imaju pozitivno mišljenje o određenoj stavki upitnika. Pozitivno mišljenje o faktoru konstatovano je ukoliko je izmerena vrednost bila iznad središnje tačke na Likertovoj skali od pet stepeni, odnosno iznad vrednosti 3. Rezultati pokazuju da prosečan stepen slaganja sa svakom tvrdnjom premašuje središnju vrednost, što u celini ukazuje na pozitivnu percepciju korišćenja VJM četbota integrisanog u obrazovnu platformu. Stavka upitnika sa najnižom ocenom odnosi se na pomoć četbota pri izradi praktičnih zadataka (prosečna ocena 4,04), što ukazuje na potencijalnu oblast za unapređenje podrške koju četbot pruža za ovu vrstu zadataka. Istovremeno, tvrdnja sa kojom su se studenti u najvećoj meri složili jeste ona koja ukazuje na to da je četbot u većini slučajeva pružao jasne odgovore. Na osnovu rezultata ankete može se zaključiti da su studenti generalno zadovoljni funkcijama četbota, mada postoji prostor za unapređenje u segmentu praktičnih zadataka.

Tabela 21 - Evaluacija četbota od strane studenata iz eksperimentalne grupe

	Srednja vrednost	Standardna devijacija	Medijana	% pozitivnih ocena
Četbot asistent mi je pomogao u razjašnjavanju teorijskih koncepata.	4,48	0,673	5,00	90,14
Četbot asistent mi je pomogao da razjasnim nejasnoće.	4,48	0,714	5,00	90,14
Četbot asistent mi je pomogao prilikom izrade zadatka.	4,04	1,088	4,00	74,65

Četbot asistent mi je pomogao u učenju.	4,38	0,684	4,00	88,73
Četbot asistent mi je pomogao da produbim znanje.	4,28	0,740	4,00	85,91
Većinu vremena četbot mi je davao jasne odgovore.	4,54	0,651	5,00	94,37

Raspodela ocena studenata koji su evaluirali četbot funkcionalnost, prikazana po pojedinim dimenzijama, data je na stubičastom dijagramu na slici 23. Za svaku stavku u anketi, odgovori su podeljeni u pet kategorija na osnovu Likertove skale. Stubičasti dijagram prikazuje relativni udeo ocena (1 ili „uopšte se ne slažem” – crvena boja; 2 ili „ne slažem se” – plava boja; 3 ili „neutralan stav” – žuta boja; 4 ili „slažem se” – narandžasta boja; 5 ili „u potpunosti se slažem” – zelena boja). Kao što se vidi, većina studenata dala je pozitivne ocene (4 i 5) za sve tvrdnje. Na osnovu slike 23 mogu se izvesti sledeći zaključci. Nijedan student nije imao negativno mišljenje o tome da li im je četbot pomogao da razjasne teorijske koncepte i olakšao učenje. Svi studenti iz eksperimentalne grupe ocenili su četbota kao pouzdanog pomoćnika u učenju teorijskih konceptata i tokom celokupnog procesa učenja. S druge strane, 6% studenata u potpunosti se nije složilo sa tvrdnjom da im je četbot pomogao u izradi praktičnog zadatka. Jedan od mogućih razloga za uočeno nezadovoljstvo četbotom jeste taj što studenti nisu na adekvatan način komunicirali sa njim. Njihovi upiti su potencijalno bili neprecizni i bez dovoljno detalja kako bi se omogućila efikasna interakcija sa četbotom. Stoga, četbot možda nije bio u stanju da generiše odgovore koji bi bili dovoljno precizni ili u skladu sa očekivanjima studenata.



Slika 23 - Raspodela studentskih ocena o četbot podršci

Evaluacija odgovora od strane eksperata

Dva stručnjaka, sa šest, odnosno 14 godina iskustva u izvođenju nastave na ovom kursu, zajednički su ocenjivali odgovore četbota na osnovu sledećih kriterijuma, koristeći Likertovu skalu od pet stepeni [112]: tačnost, relevantnost, kompletnost, koherentnost i fluentnost. Budući da su dva stručnjaka ocenjivala istu grupu studentskih odgovora, bilo je neophodno proceniti pouzdanost njihovih odgovora i utvrditi stepen slaganja između

ocenjivača. U tu svrhu korišćen je intraklasni koeficijent korelacije (eng. *Intraclass Correlation Coefficient* - ICC), koji je prvobitno predložio Fišer [189]. Kao što ističu u radu [190], prilikom korišćenja ICC-ja neophodno je da istraživači budu upoznati sa pravilnom primenom svakog od oblika ovog koeficijenta. U ovom radu primenjen je model mešovitih efekata sa dva faktora i kriterijumom apsolutnog slaganja, ICC(3,1), kako bi se procenila pouzdanost između dva fiksna ocenjivača koji su ocenili sve odgovore. Fokus je na apsolutnom slaganju, odnosno na proverbi da li su ocenjivači dodelili identične ocene. Rezultati, zajedno sa intervalima poverenja (eng. *Confidence Interval* - CI) od 95%, prikazani su u Tabeli 22. Prema smernicama za tumačenje definisanim u radu [190], rezultati ukazuju na umereno slaganje (ICC između 0,5 i 0,75) u pogledu tačnosti izražavanja, dobro slaganje u pogledu tačnosti, kompletnosti i koherentnosti (ICC između 0,75 i 0,9), kao i na odlično slaganje u pogledu relevantnosti (ICC iznad 0,9). Sveukupno, slaganje među stručnjacima bilo je visoko, pri čemu je najniži stepen slaganja zabeležen kod kriterijuma tačnosti izražavanja, što se može pripisati inherentno subjektivnoj prirodi tog kriterijuma ocenjivanja. Nakon početnih procena stručnjaka, svakom odgovoru četbota dodeljena je konačna ocena postignuta konsenzusom u slučajevima kada je utvrđeno neslaganje.

Tabela 22 - Analiza pouzdanosti između ocenjivača

	Tačnost	Relevantnost	Kompletnost	Koherentnost	Fluentnost
ICC(3,1)	0,863	0,925	0,852	0,899	0,698
95% CI	0,842-0,882	0,913-0,936	0,829-0,872	0,883-0,913	0,656-0,736

Tabela 23 prikazuje deskriptivnu statistiku (aritmetičku sredinu, standardnu devijaciju i medijanu) rezultata evaluacije, kao i procenat odgovora sa pozitivnim mišljenjem stručnjaka (ocena stručnjaka viša od 3). Rezultati evaluacije prikazani su odvojeno za tri kategorije interakcija definisane na početku ovog odeljka, kao i zbirno. Rezultati pokazuju da prosečna ocena stručnjaka za svaki kriterijum u okviru svake kategorije premašuje srednju vrednost skale, što u celini ukazuje na pozitivnu ocenu odgovora četbota. Najniže ocene zabeležene su kod praktičnih interakcija, što je u skladu sa rezultatima evaluacije studenata.

Tabela 23 - Evaluacija četbot odgovora od strane eksperata

Kriterijum	Pokazatelj	Teorijska	Praktična	Ostalo	Ukupno
Tačnost	Srednja vrednost	4,95	4,89	4,96	4,93
	Standardna devijacija	0,247	0,394	0,200	0,305
	Medijana	5	5	5	5
	% pozitivnih ocena	99,24	98,15	100	98,89

Relevantnost	Srednja vrednost	4,98	4,12	4,96	4,69
	Standardna devijacija	0,173	0,942	0,200	0,697
	Medijana	5	4	5	5
	% pozitivnih ocena	99,49	63,89	100	87,42
Kompletnost	Srednja vrednost	4,85	3,94	4,96	4,54
	Standardna devijacija	0,384	0,923	0,200	0,757
	Medijana	5	4	5	5
	% pozitivnih ocena	98,99	56,02	100	84,44
Koherentnost	Srednja vrednost	4,97	3,99	4,92	4,64
	Standardna devijacija	0,212	0,910	0,277	0,725
	Medijana	5	4	5	5
	% pozitivnih ocena	99,24	59,26	100	85,69
Fluentnost	Srednja vrednost	4,86	4,9	4,96	4,88
	Standardna devijacija	0,372	0,297	0,200	0,343
	Medijana	5	5	5	5
	% pozitivnih ocena	98,99	100	100	99,37

Budući da je osnovni cilj četbota bio pružanje podrške kako u teorijskom, tako i u praktičnom aspektu lekcije, ispitivanje razlika u kvalitetu ovih vrsta odgovora može pružiti dublji uvid u korisnost te podrške. Radi procene razlike u kvalitetu teorijskih i praktičnih odgovora, najpre je primenjen Kolmogorov-Smirnov test za proveru normalnosti raspodele kriterijuma kvaliteta. Test je pokazao da raspodela ne prati normalnu ni za jedan od njih, te je u narednom koraku sproveden neparametarski Man-Vitnijev U test. Rezultati su prikazani u Tabeli 24. Može se zaključiti da ne postoji statistički značajna razlika u fluentnosti odgovora, ali postoji statistički značajna razlika u tačnosti, relevantnosti, potpunosti i koherentnosti odgovora četbota tokom interakcije između studenata i četbota.

Tabela 24 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnijevog U testa za razlike u kvalitetu praktičnih i teorijskih odgovora prema 5 kriterijuma

Kriterijum	Srednja vrednost	SD	K-S	p-vrednost	Grupa	Srednji rangovi	Z	p-vrednost
Tačnost	4,93	0,308	0,533	0,000	Teorijska	311,14	-2,417	0,016
					Praktična	296,60		
Relevantnost	4,68	0,708	0,488	0,000	Teorijska	359,18	-14,803	0,000
					Praktična	208,74		

Kompletnost	4,53	0,766	0,424	0,000	Teorijska	363,41	-13,371	0,000
					Praktična	201,01		
Koherentnost	4,62	0,735	0,470	0,000	Teorijska	368,56	-16,234	0,000
					Praktična	191,60		
Fluentnost	4,88	0,348	0,522	0,000	Teorijska	302,72	-1,118	0,264
					Praktična	311,99		

Rezultati testa

Na kraju časa, obe grupe su radile test za procenu zapamćenog nakon sesije učenja. Test se sastojao od osam pitanja vezanih za lekciju o HTTP protokolu. Primer teorijskog pitanja je: „Kom sloju skupa internet protokola pripada HTTP?” Primer praktičnog pitanja je: „Koji je broj paketa koji predstavlja odgovor na zahtev za slikom?” (uz priložen snimak ekrana zabeleženog mrežnog saobraćaja). Odgovori su ocenjivani na skali od 0 do 1, tako da je minimalni mogući broj bodova iznosio 0, a maksimalni 8.

Za procenu normalnosti raspodele rezultata testa korišćen je Kolmogorov-Smirnov test. Test je ukazao na odsustvo normalne raspodele, pa je u narednom koraku sproveden neparametarski Man-Vitnijev U test. Rezultati su prikazani u Tabeli 25. Na osnovu dobijene p-vrednosti i srednjih rangova, utvrđeno je da ne postoje statistički značajne razlike između rezultata eksperimentalne i kontrolne grupe. Prosečan broj bodova na testu iznosio je 4,58 za eksperimentalnu grupu i 4,51 za kontrolnu grupu.

Tabela 25 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnijevog U testa za poređenje broja poena ostvarenih na testu između eksperimentalne i kontrolne grupe

	Srednja vrednost	SD	K-S	p-vrednost	Grupa	Srednji rangovi	Z	p-vrednost
Poeni	4,54	1,66	0,086	0,003	Eksperiment.	88,27	-0,168	0,867
					Kontrolna	86,97		

10.1.3 Diskusija rezultata

Sveukupno, studenti su izrazili visok stepen zadovoljstva integrisanom funkcijom četbota, zasnovanom na OpenAI GPT-4o modelu, u okviru predstavljenog prototipa platforme za učenje, a koja je pružala pomoć tokom učenja na kursu Računarske mreže i telekomunikacije. Ovaj nalaz je u skladu sa prethodnim istraživanjima koja su se bavila upotrebom velikih jezičkih modela na kursovima iz oblasti računarskih nauka u visokom obrazovanju [115], [116], [117], [118], [120].

Kao što je već pomenuto, studenti su sa četbotom komunicirali uglavnom na srpskom jeziku (99,53% interakcija je bilo na srpskom). Ovo istraživanje pokazuje da je GPT-4o model ostvario visoke performanse čak i kod jezika sa nižom dostupnošću resursa za treniranje modela, kao što je srpski. Stručna procena tačnosti odgovora koje je generisao VJM pokazuje visok nivo za sve interakcije ($M=4,88$ na Likertovoj skali od 5 poena). Tokom ocenjivanja odgovora, stručnjaci su primetili da je VJM u pojedinim slučajevima generisao odgovore koji su sadržali hrvatske reči, što ne bi trebalo da predstavlja problem za studente iz Srbije zbog sličnosti ova dva jezika.

Studenti su naveli da je četbot bio efikasniji u objašnjavanju teorijskih koncepata i podsticanju dubljeg konceptualnog razumevanja nego u pružanju pomoći oko praktičnih aspekata lekcije. Ovo zapažanje potkrepljeno je i stručnim ocenama odgovora četbota, koje su otkrile statistički značajne razlike između teorijskih i praktičnih pitanja u pogledu kriterijuma kao što su tačnost, relevantnost, potpunost i koherentnost. Takođe, ovo je u skladu sa prethodnim istraživanjima koja pokazuju visoku efikasnost OpenAI GPT modela, dostupnih putem ChatGPT-ja [191], [192], u odgovaranju na teorijska pitanja. S druge strane, jedno od mogućih objašnjenja za bolje rezultate kod teorijskih pitanja u odnosu na praktična jeste ograničen kontekst dostupan modelu. VJM je funkcionisao na osnovu sistemske instrukcije prikazane u poglavlju o implementaciji (eng. *system prompt*) koja je definisala ulogu asistenta i opšte ciljeve učenja sa sažetkom lekcije, uz korišćenje istorijata konverzacije radi održavanja koherentnosti. Međutim, model nije imao direktan pristup specifičnim detaljima ili sadržaju vezanom za praktični zadatak, koji je uključivao upotrebu alata Wireshark za analizu mrežnih paketa. Neki studenti su u početku pretpostavljali da četbot može pružiti direktnu pomoć pri rešavanju praktičnih zadataka, ali su, nakon što su shvatili njegova ograničenja, prilagodili svoje upite pružanjem više konteksta ili prelaskom na opštija pitanja. Ovo je u skladu sa našim prethodnim nalazima koji ukazuju na to da studenti postepeno uče kako da optimizuju upite, te da dugotrajna upotreba dovodi do pozitivnije percepcije VJM četbota, kao što je ChatGPT [84]. Nedavanje direktnih odgovora studentima pruža bolju priliku za učenje, naročito kada se oni podstiču da preciziraju i prilagode svoja pitanja. Ova promena u strategiji interakcije možda je uticala na zadovoljstvo, posebno u pogledu praktične pomoći. Ipak, prosečna ocena zadovoljstva ($M = 4,04$ na petostepenoj Likertovoj skali), kao i ocena stručnjaka, ostale su pozitivne.

Ovi uvidi ukazuju na to da bi buduće verzije platforme trebalo da razmotre rešenja za unapređenje domenske specifičnosti četbota uključivanjem konteksta vezanog za praktične zadatke što je urađeno u drugoj verziji platforme. Potrebno je pažljivo razmotriti održavanje ravnoteže između pružanja podrške i davanja direktnih odgovora. Mehanizam generisanja uz *Retrieval Augmented Generation* (RAG) tehniku mogao bi se razmotriti kao način za poboljšanje podrške kod praktičnih zadataka. Uključivanjem relevantnih kontekstualnih dokumenata ili instrukcija u proces generisanja odgovora, sistem bi mogao pružiti precizniju pomoć u zadovoljavanju potreba studenata tokom savladavanja praktičnih delova lekcije.

Iako je eksperimentalna grupa imala pristup četbotu, njihovi rezultati na testu nisu se značajno razlikovali od rezultata kontrolne grupe. Ovaj rezultat se delimično može objasniti razlikama u načinu na koji su se dve grupe uključile u proces učenja. Značajno ograničenje pri tumačenju rezultata testova predstavlja metodološka asimetrija između

kontrolne i eksperimentalne grupe [193]. Naime, dok je kontrolnoj grupi bilo dozvoljeno da postavlja pitanja radi pojašnjenja direktno asistentu u nastavi, od eksperimentalne grupe se zahtevalo da se za podršku osloni isključivo na četbot. Ova razlika u dostupnoj podršci verovatno je uticala na uočene ishode i treba je uzeti u obzir prilikom tumačenja rezultata. Drugim rečima, razlike u nastavnoj podršci mogle su uticati na ishode učenja nezavisno od same intervencije [194]. Pored toga, studenti u eksperimentalnoj grupi bili su izloženi dodatnom kognitivnom opterećenju usled interakcije sa četbotom, što je možda zahtevalo više vremena i truda u poređenju sa kontrolnom grupom, koja je mogla brže da prelazi preko nastavnog materijala. Takođe je važno napomenuti da se naknadni test odnosio isključivo na sadržaj koji je eksplicitno obrađen u materijalima za učenje dostupnim na platformi. Studenti u kontrolnoj grupi su potencijalno bili više usredsređeni na taj konkretan sadržaj tokom istog vremenskog perioda, dok su studenti u eksperimentalnoj grupi često koristili četbot za istraživanje dodatnih primera i šira konceptualna objašnjenja. Ovaj širi angažman možda se nije odrazio na bolje rezultate na testu, ali je potencijalno mogao doprineti dubljem ili dugoročnijem razumevanju, što test nije imao za cilj da izmeri. Za donošenje pouzdanih zaključaka neophodna su dalja istraživanja, po mogućstvu u vidu dužih studija koje obuhvataju raznovrsne zadatke i oblike provere znanja.

Ova studija ima nekoliko ograničenja koja bi trebalo razmotriti u budućim istraživanjima. Prvo, evaluacija je sprovedena tokom jedne sesije od 90 minuta, što ograničava uvid u dugoročne efekte korišćenja četbota na ishode učenja. Buduće studije bi trebalo da primene dugoročni dizajn, omogućavajući studentima interakciju sa četbotom tokom dužeg vremenskog perioda. To bi istraživačima omogućilo da efikasnije procene zadržavanje znanja, promene u ponašanju pri učenju i ukupan uticaj na učenje. Drugo, relativno mali uzorak, kao i fokus na samo jedan predmet (Računarske mreže i telekomunikacije) i jednu nastavnu jedinicu (HTTP protokol), ograničavaju mogućnost generalizacije zaključaka. Rezultati studije pokazuju da su studenti u ovom konkretnom slučaju ocenili upotrebu četbota kao veoma korisnu, pozitivnu i delotvornu, što je u skladu sa prethodnim sličnim studijama [195], [196]. Ipak, ne možemo tvrditi da bi se zabeležio isti nivo zadovoljstva ukoliko bi se četbot koristio tokom čitavog semestra i bez prisustva asistenata u nastavi. Svakako, novije studije su pokazale da je nakon uvođenja generativne veštačke inteligencije zabeležen značajan trend rasta motivacije [197], kao i da povratne informacije dobijene od četbota pružaju efikasniju i delotvorniju podršku tokom usvajanja praktičnog znanja [198].

Praktično ograničenje predstavljaju troškovi korišćenja komercijalnih API-ja za velike jezičke modele, kao što je OpenAI GPT-4o, naročito kada je reč o dugoročnoj upotrebi ili primeni u velikom obimu. Pored toga, buduća istraživanja bi mogla da ispituju performanse alternativnih modela, poput DeepSeek-a ili Google-ovog Gemini-ja, posebno u pogledu ekonomičnosti. Kada je reč o razvoju platforme, kao što je već pomenuto, integracija RAG tehnike predstavlja obećavajući pristup za unapređenje sposobnosti četbota, tako što omogućava pristup relevantnim nastavnim materijalima i obezbeđuje bolju usklađenost sa sadržajem kursa. Takođe, sama ulazna instrukcija nalaže četbotu da nivo objašnjenja prilagođava studentima isključivo na osnovu toka razgovora i načina na koji postavljaju pitanja, što nije dovoljno za personalizaciju odgovora. Iako su studenti uglavnom bili zadovoljni odgovorima četbota, jedan od mogućih pravca istraživanja može biti eksplicitno modelovanje procesa učenja kako bi se dodatno poboljšalo iskustvo učenja.

Na osnovu ovih rezultata može se zaključiti da, iako su veliki jezički modeli izazvali negativnu konotaciju u obrazovnom okruženju zbog velikog potencijala za zloupotrebu, ova tehnologija takođe ima potencijal za unapređenje iskustva učenja. Kroz integraciju sa nastavnim sadržajem, posebno u vidu jedinica za mikroučenje koje su kreirali nastavnici, ovi modeli su pokazali ohrabrujuće rezultate za kurs Računarske mreže i telekomunikacije. Efikasno balansiranje između stručnosti nastavnika i podrške koju ovi modeli pružaju na zahtev (radi razjašnjavanja nejasnoća i produbljivanja znanja studenata tokom procesa učenja) ima potencijal da unapredi celokupno iskustvo učenja, a istovremeno smanji opterećenje nastavnika u tom pogledu.

10.2 Evaluacija drugog prototipa

U ovom poglavlju će biti predstavljen sproveden eksperiment za evaluaciju druge verzije prototipa. Biće prikazani rezultati evaluacije platforme od strane studenata sa fokusom na upotrebljivost sistema.

10.2.1 Upotrebljivost

Upotrebljivost sistema je prema ISO 9241-11 standardu definisana kao „mera do koje proizvod može biti korišćen od strane određenih korisnika za postizanje određenih ciljeva efektivno, efikasno i sa zadovoljstvom u određenom kontekstu korišćenja” [199]. Jakob Nilsen posmatra upotrebljivost kao lakoću korišćenja interfejsa i definiše pet komponenti kvaliteta upotrebljivosti kao što su [200]:

- **Lakoća učenja korišćenja sistema** koja se odnosi na lakoću izvršavanja osnovnih zadataka prilikom prvog susreta sa sistemom;
- **Efikasnost** koja se odnosi na brzinu izvršavanja zadataka;
- **Pamtljivost** koje se odnosi na lakoću ponovnog navikavanja na sistem nakon nekog vremena nekorišćenja;
- **Greške** koje se odnose na učestalost grešaka, ozbiljnost grešaka i lakoću oporavka od istih;
- **Zadovoljstvo dizajnom sistema.**

Poseban aspekt upotrebljivosti u sistemima za učenje predstavlja lakoća učenja, ali ne samo iz ugla korišćenja sistema, već i sa aspekta efektivnosti učenja i zadržavanja znanja i veština [201]. S obzirom na to da je cilj ove platforme povećanje angažovanosti studenata prilikom učenja i produbljivanje njihovog razumevanja, evaluacija upotrebljivosti u ovom kontekstu će se fokusirati na efektivnost upotrebe platforme, angažovanost korisnika i njegovo zadovoljstvo s obzirom na to da povišeno vreme korišćenja platforme i vođenje konverzacije sa konverzionim agentom mogu predstavljati povećanu angažovanost korisnika i bolje iskustvo učenja.

Tradicionalni sistemi za upravljanje učenjem se u visokom obrazovanju uglavnom koriste kao medijum za distribuciju sadržaja odnosno materijala za učenje [202], [203]. Kako bi se izmerila i poredila upotrebljivost sistema za upravljanje učenjem sa konverzionim agentom u odnosu na tradicionalne sisteme koji ne poseduju ovu

opciju, studentima su date na korišćenje dve verzije platforme koja je kreirana kao prototip u ovoj disertaciji. Jedna verzija platforme integriše četbota uz sadržaj za učenje, dok druga platforma ne uključuje ovu opciju već služi samo kao medijum za distribuciju materijala za učenje. Nakon korišćenja, studenti su popunili evaluacione ankete sa konstruktima opisanim u nastavku poglavlja.

Skala upotrebljivosti sistema

Skala upotrebljivosti sistema (eng. *System Usability Score* – SUS) predstavlja upitnik koji se koristi u domenu evaluacije upotrebljivosti sistema [204]. Razvijen je od strane *John Brooke* dok je radio u američkoj kompaniji za proizvodnju kompjutera *Digital Equipment Corporation*. Ovaj upitnik predstavlja jedan od široko primenjivanih upitnika za evaluaciju softvera [205] i pronalazi svoju primenu i u sistemima za elektronsko učenje [206], [207]. Stavke na ovoj skali su date u narednoj tabeli. Saglasnost sa tvrdnjama na ovoj i narednim skalama, merena je na skali od 1 do 5.

Tabela 26 - Lista tvrdnji koje čine Skalu upotrebljivosti sistema [204]

R.B.	Stavka
1	Mislim da bih želeo/la da često koristim ovaj sistem.
2	Smatram da je sistem nepotrebno složen.
3	Smatram da je sistem jednostavan za korišćenje.
4	Mislim da bih morao/la da tražim pomoć tehničke osobe da bih mogao/la da koristim ovaj sistem.
5	Smatram da su različite funkcije ovog sistema dobro integrisane.
6	Smatram da postoji previše nedoslednosti u ovom sistemu.
7	Pretpostavljam da bi većina ljudi brzo naučila da koristi ovaj sistem.
8	Smatram da je sistem veoma nezgrapan za korišćenje.
9	Osećao/la sam se veoma sigurno koristeći sistem.
10	Morao/la sam da naučim mnogo stvari pre nego što sam mogao/la da počnem da koristim sistem.

Skor upotrebljivosti sistema se dobija tako što se vrednosti udela u skor za svaku stavku sabiraju i množe sa 2,5. Udeo u skor za neparne stavke se računa tako što se od vrednosti koju je korisnik odabrao na Likertovoj skali oduzme 1, dok za parne stavke se udeo u skor računa tako što se od broja 5 oduzme vrednost koju je korisnik odabrao na Likertovoj skali. Kao što je već rečeno, da bi se izračunao ukupan SUS skor, ovako dobijene vrednosti se sabiraju za sve stavke i množe sa 2,5.

TAM Konstrukti

Pored SUS upitnika, studenti su nakon korišćenja platforme popunili i anketu koja meri sledeće konstrukte preuzete iz originalnog TAM modela [85]: percipirana korisnost (PK) i percipirana lakoća korišćenja (PLK); kao i percipirano zadovoljstvo (PZ) na koje mogu uticati prethodno navedeni konstrukti. Percipirana korisnost se odnosi na stepen verovanja osobe da će korišćenje sistema poboljšati njegove performanse [85]. Percipirana lakoća korišćenja se odnosi na stepen verovanja osobe da za korišćenje sistema nije potrebno ulagati napor [85]. Percipirano zadovoljstvo studenta u ovom kontekstu se odnosi zbir bihejviorističkih uverenja i stavova koji su nastali na osnovu svih benefita koje korisnik dobija korišćenjem elektronskih sistema za učenje [208]. U tabeli 27 su date prilagođene stavke ovih konstrukata. Saglasnost sa svim stavkama je merena na skali od 1 do 5.

Tabela 27 - Prilagođeni i korišćeni TAM konstrukti: percipirana korisnost, percipirana lakoća korišćenja, percipirano zadovoljstvo

Konstrukt	Stavka
Percipirana korisnost	Edukativna platforma mi olakšava savladavanje gradiva.
	Smatram da je edukativna platforma korisna za moje obrazovanje.
	Korišćenje edukativne platforme povećava moju produktivnost u učenju.
	Korišćenje edukativne platforme čini moje učenje efikasnijim.
	Korišćenje edukativne platforme poboljšava moje rezultate u učenju.
Percipirana lakoća korišćenja	Učenje da koristim ovu edukativnu platformu je lako za mene.
	Pronalaženje potrebnih materijala na platformi je jednostavno.
	Interakcija sa edukativnom platformom je jasna i razumljiva.
	Smatram da je korišćenje edukativne platforme jednostavno.
	Lako savladavam sve funkcije edukativne platforme.
Percipirano zadovoljstvo	Zadovoljan/na sam iskustvom korišćenja ove platforme.
	Ova platforma je ispunila moja očekivanja.
	U celini, zadovoljan/na sam načinom na koji platforma funkcioniše.
	Korišćenje ove platforme je bilo prijatno iskustvo.
	Platforma mi je pružila iskustvo koje smatram pozitivnim.

Četbot evaluacija

Takođe, studenti koji su koristili verziju sa četbotom su popunili i kratki evaluacioni upitnik zadovoljstva vezan za četbot funkcionalnost čije su stavke prikazane u tabeli 28. Upitnik je izveden na osnovu konstrukata UTAUT2 modela [183] kao što su očekivane performanse, hedonistička motivacija i namera korišćenja; i proširen stavkama vezanim za sigurnost i poverenje kao čestim elementima koji se ispituju u ovom kontekstu i navode kao ključni faktori u prihvatanju tehnologija elektronskog učenja [209]. Takođe, saglasnost sa svim stavkama je merena na skali od 1 do 5.

Tabela 28 - Stavke evaluacije četbota

R.B.	Stavka
1	Četbot je olakšao razumevanje sadržaja lekcije.
2	Četbot se pokazao kao korisna pomoć prilikom rešavanja zadataka.
3	Odgovori četbota su bili jasni.
4	Odgovori četbota su bili korisni.
5	Interakcija sa četbotom mi je povećala motivaciju za korišćenje platforme.
6	Ponovo bih koristio/la četbot kao pomoć pri rešavanju zadataka.
7	Osećam se sigurno kada se oslanjam na odgovore koje daje četbot.
8	Osećam se sigurno dok radim zadatke zbog dostupnosti četbota.
9	Četbot pruža pouzdane informacije kada mu postavim pitanje.

10.2.2 Metodologija istraživanja

Kako bismo testirali upotrebljivost date platforme sproveden je eksperiment u kome su učestvovali studenti treće godine Fakulteta organizacionih nauka sa smera Informacioni sistemi i tehnologije. Studenti su koristili platformu u okviru predmeta Računarske mreže i telekomunikacije tokom jedne sesije za učenje tematske celine na ovom predmetu. U prvom delu sesije učenja studenti su prolazili kroz materijal pripremljen od strane nastavnika, dok su u drugom delu davali odgovore na pitanja i rešavali aktivnosti uz dostupnost ovih materijala. Eksperimentalna grupa koja je imala dostupnog četbota za korišćenje je mogla da mu postavlja pitanja o lekciji i tokom prolaska kroz materijale, i tokom izrade aktivnosti. Lekcija je bila o mrežnom sloju TCP/IP modela sa fokusom na prevođenje mrežnih adresa (eng. *Network Address Translation* – NAT).

Eksperiment je sproveden u učionici na Fakultetu organizacionih nauka i trajao je okvirno 70 minuta. Jedna grupa učesnika je tokom istraživanja koristila verziju platforme za učenje sa integrisanim četbotom, a druga bez. Ovaj dizajn eksperimenta je omogućio da uporedimo utrošeno vreme za savladavanje lekcije između ove dve grupe, kao i broj interakcija sa platformom odnosno navigaciju kroz materijal za učenje koji je bio dostupan obema grupama tokom izrade aktivnosti kako bismo videli da li integracija četbota potencijalno utiče na vreme provedeno na platformi, kao i na navigaciju kroz samu platformu tokom izrade zadataka. Nakon korišćenja platforme za učenje i rešavanje zadataka, studenti su popunili ankete evaluacije korišćenja platforme kako bismo takođe mogli da uporedimo skor upotrebljivosti platforme, percipiranu korisnost, lakoću korišćenja i zadovoljstvo, konstrukte koji su opisani u prethodnim sekcijama. Pored ovoga u daljoj analizi će biti predstavljeni i rezultati samih aktivnosti tokom učenja, kao i rezultati kratkog testa vezanog za lekciju koju su učili koji su radili na kraju eksperimenta, bez dostupne platforme.

U istraživanju je učestvovalo ukupno 68 studenata treće godine, od kojih je 35 koristilo verziju platforme sa četbotom, a 33 verziju bez. Pregled polne strukture ispitanika u ove dve grupe prikazan je u Tabeli 29. Rezultati testa verodostojnosti odnosa ($\chi^2 = 0,041$, $df = 1$, $p = 0,839$) ukazuju na to da ne postoji statistički značajna razlika u polnoj strukturi između eksperimentalne i kontrolne grupe. Na osnovu toga može se zaključiti da su grupe ujednačene u pogledu polne zastupljenosti.

Tabela 29 - Raspodela učesnika prema polu

Eksperimentalna grupa	N	%	<i>Likelihood ratio</i>
Muški	14	40	$\chi^2=0,041$ $p=0,839$
Ženski	21	60	
Kontrolna grupa			
Muški	14	42	
Ženski	19	58	

Kao i u prethodnom eksperimentu, radi provere ove dve grupe koje služe kao kontrolna i eksperimentalna kod poređenja rezultata testova, zadataka i broja interakcija, analizirana su dva pokazatelja akademskog uspeha: broj ostvarenih bodova prema Evropskom sistemu prenosa i akumulacije bodova (ESPB) i prosečna ocena. Kolmogorov-Smirnov test primenjen je radi ispitivanja normalnosti raspodele obe varijable, ESPB bodova i prosečne ocene, pri čemu je utvrđeno da prosek prati normalnu raspodelu, dok ESPB ne prati. S obzirom na to, u narednom koraku je za poređenje grupa primenjen parametarski t-test nezavisnih uzoraka za prosek, dok je neparametarski Man-Vitnijev U test za ESPB bodove. Rezultati, prikazani u Tabeli 30, na osnovu dobijenih p-vrednosti i srednjih vrednosti i rangova, ukazuju na to da ne postoje

statistički značajne razlike između eksperimentalne i kontrolne grupe u pogledu prethodnog akademskog uspeha, merenog ESPB bodovima i prosečnom ocenom.

Tabela 30 - Rezultati Kolmogorov-Smirnovljevog testa, t-testa nezavisnih uzoraka i Man-Vitnijevog U testa za ESPB i prosek u eksperimentalnoj i kontrolnoj grupi

	Srednja vrednost	SD	K-S	p vrednost	Grupa	Srednja vrednost/ Srednji rangovi	t/Z	p vrednost
Prosek	8,18	0,63	0,069	0,2	Eksperiment.	8,18	-0,030	0,976
					Kontrolna	8,19		
ESPB	112,01	21,89	0,343	0,0	Eksperiment.	32,97	-0,685	0,493
					Kontrolna	36,12		

10.2.3 Rezultati i analiza

Skor upotrebljivosti sistema, percipirana korisnost, percipirana lakoća korišćenja i percipirano zadovoljstvo platformom

Za procenu upotrebljivosti platforme sa i bez integrisanog četbota korišćen je *between-subject* dizajn, pri čemu su učesnici podeljeni u dve grupe nezavisnih ispitanika. Prva grupa je koristila platformu sa četbotom, dok je druga grupa koristila platformu bez četbota za savladavanje lekcije o mrežnom sloju. Nakon korišćenja platforme, učesnici su popunjavali *System Usability Score* (SUS) upitnik kako bi se kvantifikovala njihova percepcija upotrebljivosti datog sistema. Pored SUS ankete, studenti su popunili i anketu sa konstruktima iz TAM modela – percipirana korisnost, percipirana lakoća korišćenja i percipirano zadovoljstvo platformom za obe verzije platforme. Za proveru interne konzistentnosti spomenutih konstrukata, izračunat je Kronbahov alfa koeficijent čija vrednost prelazi prag od 0,7 (SUS – 0,703; PK – 0,899; PLK – 0,733; PZ – 0,912). Kako su dobijene vrednosti Kronbahovog alfa koeficijenta iznad definisane granice od 0,7, smatra se da definisani konstrukti imaju zadovoljavajuću internu konzistentnost.

Kolmogorov-Smirnov test primenjen je radi ispitivanja normalnosti raspodele SUS skorova i pomenutih konstrukata, pri čemu je utvrđeno da ove varijable ne prate normalnu raspodelu ($p < 0,05$ u svim slučajevima) (tabela 31 i tabela 32). S obzirom na to, u narednom koraku korišćen je neparametarski Man-Vitnijev U test za poređenje skora upotrebljivosti, percipiranog zadovoljstva, percipirane lakoće korišćenja i percipirane korisnosti platforme sa i bez četbota.

Man-Vitnijev U test je pokazao da je razlika između SUS ocena sa i bez konverzionog agenta statistički značajna ($Z = -2,055$, $p = 0,04$). Rezultati SUS upitnika pokazali su prosečne ocene od 90,86 za platformu sa četbotom i 86,59 za platformu bez četbota. Ovi rezultati (tabela 31) ukazuju na veći skor upotrebljivosti platforme sa integrisanim četbotom u okviru ovog eksperimenta.

Tabela 31 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnijevog U testa za SUS

	Srednja vr.	SD	K-S	p vr.	Grupa	Srednji rangovi	Z	p vr.	Srednje vr.
SUS skor	88,79	8,78	0,119	0,018	Eksperiment.	39,26	-2,055	0,04	90,86
					Kontrolna	29,45			86,59

Takođe, rezultati prikazani u tabeli 32, na osnovu dobijenih p-vrednosti i srednjih rangova, ukazuju na to da postoje statistički značajne razlike između eksperimentalne i kontrolne grupe u pogledu navedenih konstrukata TAM modela, odnosno da su viši kod grupa koje su imale mogućnost korišćenja četbota. Srednje vrednosti ovih konstrukata u obe grupe prelaze vrednost 4,00 i mogu se videti u istoj tabeli.

Tabela 32 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnijevog U testa za PK, PLK i PZ

	Srednja vr.	SD	K-S	p vrednost	Grupa	Srednji rangovi	Z	p vrednost	Srednje vr.
PK	4,39	0,67	0,215	0,000	Eksperiment.	44,76	-4,559	0,000	4,74
					Kontrolna	23,62			4,02
PLK	4,69	0,45	0,255	0,000	Eksperiment.	42,50	-3,685	0,000	4,82
					Kontrolna	26,02			4,54
PZ	4,59	0,51	0,259	0,000	Eksperiment.	42,54	-3,657	0,000	4,79
					Kontrolna	25,97			4,38

Vreme interakcije sa materijalima i kretanje kroz platformu

Pored evaluacija od strane studenata, takođe poredimo vremena interakcije sa platformom tokom sesije učenja, a pre započinjanja izrade zadatka gde su obe grupe imale ograničeno vreme, kao i kretanje odnosno navigaciju kroz platformu tokom izrade zadatka. Kolmogorov-Smirnov test primenjen je radi ispitivanja normalnosti raspodele pomenutih parametara, pri čemu je utvrđeno da ne prate normalnu raspodelu. S obzirom na to, u narednom koraku je za poređenje grupa primenjen neparametarski Man-Vitnijev U test.

Rezultati, prikazani u Tabeli 32, na osnovu dobijenih p-vrednosti i srednjih rangova, ukazuju na to da postoje statistički značajne razlike između eksperimentalne i kontrolne grupe u pogledu ovih konstrukata, odnosno da su vremena viša kod grupa koje su imale mogućnost korišćenja četbota, dok je smanjena potreba za navigacijom kroz materijale

koji su pripremljeni od strane nastavnika. Srednja vrednost vremena provedenog na platformi je 42,12 minuta, a prosečan broj navigacija kroz materijale je 6,11 za platformu sa četbotom, dok za platformu bez četbota je prosečno vreme 26,09 minuta i prosečan broj navigacija 50,30. Pored uticaja konverzionog agenta na smanjenje potrebe za navigacijom kroz platformu tokom izrade zadatka, može se razmisliti i o promeni načina prikaza zadatka studentu tokom navigacije kroz nastavne jedinice, radi poboljšanja korisničkog iskustva, s obzirom na to da se u trenutnoj verziji platforme zadatak nalazi na posebnoj stranici, što je takođe moglo uticati na povećanje broja navigacija kroz vraćanje na tu stranicu više puta i nakon pronađenih resursa radi popunjavanja odgovora.

Tabela 33 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnijevog U testa za trajanje interakcije i navigaciju

	Srednja vr.	SD	K-S	p vr.	Grupa	Srednji rangovi	Z	p vr.	Srednje vr.
Trajanje	34,34	9,54	0,142	0,002	Eksperiment.	49,94	-6,633	0,000	42,12
					Kontrolna	18,12			26,09
Navigacija	27,56	25,136	0,172	0,000	Eksperiment.	18,11	-7,116	0,000	6,11
					Kontrolna	51,88			50,30

Zadaci i test

Studenti su tokom sesije učenja imali 25 minuta za 7 kratkih zadataka koji su bodovani sa 7 poena ukupno. Većina zadataka je preuzeta i prilagođena sa kraja lekcije iz knjige „Računarske mreže: Od vrha ka dnu” autora Kurose i Ross [210]. Tokom izrade zadataka su imali mogućnost da se kreću kroz platformu, kao i da razgovaraju sa četbotom postavljajući mu pitanja vezana za razjašnjavanje teorijskih koncepata, ukoliko su bili grupa kojoj je ta funkcionalnost bila dostupna.

Jedan tip zadatka je podrazumevao pretragu određenih pojmova gde je četbot davao direktne odgovore simulirajući pretragu materijala (npr. Koja je veličina uobičajenog zaglavlja IP datagrama?). Četbotu nije kroz kontekst dat tačan odgovor na pitanje, već samo dozvola da direktno odgovori na dato pitanje, ali i na neka dodatna pitanja ukoliko ih student ima.

Drugi tip zadatka zahtevao je i razumevanje određenih koncepata radi njihove primene u kontekstu zadatka. U ovom slučaju je četbot bio zadužen za razjašnjavanje teorije koja prati date koncepte. Primer drugog tipa zadatka za koji je potrebno razumevanje NAT-a je: „Koja je izvorna, a koja odredišna IP adresa datagrama u koraku 2 nakon što ga je ruter prosledio? (priložena slika uz zadatak)“, dok su primeri pitanja koje su studenti postavljali kako bi uspešno rešili zadatak: *Ukoliko racunar koji je na lokalnoj mrezi sa NAT-om salje podatke van lokalne mreze, koja ce biti izvorna IP adresa tih podataka, IP adresa koju NAT daje kao javnu ili privatna IP adresa tog racunara; kada paket prodje*

kroz ruter koji koristi NAT, za sve ostale uređaje u mrezi izvorna adresa će biti javna adresa rutera?; Kako odredjujem izvornu IP adresu datagrama i sl. U određenim slučajevima kad student nije tačno odgovorio na neki zadatak nije ni vodio konverzaciju sa četbotom što može potencijalno biti uticaj ograničenja vremena. U određenim konverzacijama koje su rezultovale netačnim ili delimično tačnim odgovorom studenta, iako je četbot generisao tačan odgovor studentu, davao je previše uopšten odgovor, a student nije nastavio konverzaciju i dao dodatni kontekst ili postavio preciznije pitanje. Upravo ovaj tip konverzacije koji je rezultovao nepreciznim odgovorom može tokom evaluacionog procesa signalizirati da je potrebno dopuniti kontekst četbotu kako bi se povećale performanse, međutim vodeći računa da se ne ugroze do sad dobri odgovori. Isečak takve konverzacije je dat u nastavku za zadatak koji je glasio „Pretpostavimo da između izvornog i odredišnog računara postoje četiri rutera. Koliko tabela prosleđivanja će se indeksirati, da bi se datagram preneo od izvora do odredišta?“:

- **Student:** *znaci ako imam 4 rutera između računara, kada paket treba iz jednog rutera da predje u drugi, on ima opciju da predje u 3 druga rutera?*
- **Četbot:** *Moguće je da u određenoj mrežnoj topologiji paket ima više ruta koje može slediti, ali to ne zavisi nužno samo od broja rutera između dva računara. Broj dostupnih ruta, odnosno koliko različitih ruta paket može imati na putu između izvora i odredišta, zavisi od konkretne mrežne topologije i konfiguracije mreže.*

Ruter koristi svoju tabelu prosleđivanja kako bi odlučio na koju sledeću „stanicu“ (ili koji sledeći ruter) će poslati paket. Tabela prosleđivanja se ažurira na osnovu informacija dobijenih kroz protokole rutiranja i merača performansi mreže kako bi se osigurala optimalna prosljedba.

Svaki od tih 4 rutera na svom putu će imati svoju odgovarajuću tabelu prosleđivanja, i svaki će doneti odluku baziranu na svojoj tabeli. Koliko ruta ili opcija prelaska postoji zapravo zavisi od toga kako su ti ruteri međusobno povezani u mreži. Ako su svi međusobno povezani, ruter može imati više opcija za prosleđivanje paketa, ali u stvarnosti se koristi najbolja moguća ruta prema vrednostima u tabeli prosleđivanja.

Većina studenata je podrazumevala linearnu topologiju mrežnih uređaja koji predstavljaju putanju do odredišta i tačno uradila zadatak, međutim u ovoj situaciji studentu je bilo potrebno dodatno objašnjenje. U ovom slučaju može biti poželjno dodatno prilagođavanje odgovora kontekstu ovog zadatka.

Kolmogorov–Smirnov test primenjen je radi ispitivanja normalnosti raspodele rezultata zadataka, pri čemu je utvrđeno da rezultati ne prate normalnu raspodelu. S obzirom na to, u narednom koraku je za poređenje grupa primenjen neparametarski Man-Vitnijev U test. Rezultati, prikazani u Tabeli 34, na osnovu dobijenih p-vrednosti i srednjih rangova, ukazuju na to da postoje statistički značajne razlike između eksperimentalne i kontrolne grupe u pogledu ovih rezultata, odnosno da su viši kod grupa koje su imale mogućnost korišćenja četbota. Prosek tačnosti zadataka koji je maksimalno nosio 7 poena je u eksperimentalnoj grupi iznosio 6,54 poena, a u kontrolnoj 5,04.

Tabela 34 - Rezultati Kolmogorov-Smirnovljevog testa i Man-Vitnijevog U testa za poene ostvarene na zadacima

	Srednja vr.	SD	K-S	p vr.	Grupa	Srednji rangovi	Z	p vr.	Srednja vr.
Poeni	5,904	1,13	0,201	0,000	Eksperiment.	46,84	-5,466	0,000	6,61
					Kontrolna	21,41			5,15

Nakon sesije učenja, studenti su pristupili testu, ali ovog puta bez pristupa platformi, kako bismo izmerili zapamćeno nakon sesije učenja. Test je sadržao 10 kratkih pitanja sa ponuđenim odgovorima (npr. Rutiranje obuhvata prenos paketa od ulaznog do izlaznog linka unutar istog rutera. Tačno/Netačno), na dopunu (npr. Koliko minimalno mrežnih interfejsa ima jedan ruter?) i otvorena (npr. Zašto se javila potreba za NAT-ovanjem?) za koji su imali na raspolaganju 8 minuta.

Kolmogorov-Smirnov test primenjen je radi ispitivanja normalnosti raspodele rezultata testa, pri čemu je utvrđeno da rezultati prate normalnu raspodelu. S obzirom na to, u narednom koraku je za poređenje grupa primenjen parametarski t-test nezavisnih uzoraka. Rezultati, prikazani u Tabeli 35, na osnovu dobijenih p-vrednosti i srednjih vrednosti, ukazuju na to da postoji statistički značajna razlika između eksperimentalne i kontrolne grupe u pogledu rezultata testa, odnosno da su viši kod grupa koje su imale mogućnost korišćenja četbota. Prosek testa koji je maksimalno nosio 10 poena je u eksperimentalnoj grupi iznosio 7,63 poena, a u kontrolnoj 6,03.

Tabela 35 - Rezultati Kolmogorov-Smirnovljevog testa i t-testa nezavisnih uzoraka za poene ostvarene na testovima

	Srednja vrednost	SD	K-S	p vrednost	Grupa	Srednja vrednost	t	p vrednost
Poeni	6,85	1,916	0,104	0,065	Eksperimentalna	7,63	3,759	0,00
					Kontrolna	6,03		

Četbot evaluacija

Studenti koji su koristili verziju platforme sa četbotom su evaluirali i njegovu funkcionalnost popunjavanjem upitnika čije su stavke date u tabeli 4. Za validaciju upitnika [187], [188] izračunali smo Kronbahov alfa koeficijent (0,894), prosečnu ekstrahovanu varijansu (AVE = 0,534), i kompozitnu pouzdanost (CR = 0,909) metrike čije su vrednosti iznad preporučenih pragova (Kronbahov alfa koeficijent > 0,7, AVE > 0,5, and CR > 0,7) i samim tim smo zaključili da data skala može biti korišćena za merenje studentske percepcije korišćenja četbota. Tabela 36 sadrži deskriptivne statističke pokazatelje, uključujući aritmetičku sredinu, standardnu devijaciju (SD), medijanu i procenat studenata koji imaju pozitivan stav prema datoj stavci upitnika. Pozitivan stav prema faktoru zabeležen je ukoliko je izmerena vrednost bila iznad srednje tačke na petostepenoj Likertovoj skali, koja iznosi 3.

Rezultati pokazuju da prosečan stepen slaganja sa svakom tvrdnjom prevazilazi srednju vrednost skale, što u celini ukazuje na pozitivnu percepciju korišćenja VJM četbota integrisanog u edukativnu platformu. Stavka sa najnižom prosečnom ocenom odnosi se na uticaj četbota na motivaciju za korišćenjem platforme, sa srednjom vrednošću od 4,2. Iako ova vrednost pokazuje pozitivan uticaj, takođe ostavlja prostor za potencijalna dodatna proširenja funkcionalnosti četbota. Istovremeno, tvrdnja sa kojom su se studenti u najvećoj meri složili odnosi se na to da bi ponovo koristili četbota kao pomoć pri rešavanju zadataka.

Tabela 36 – Evaluacija četbota od strane studenata

RB	Srednja vrednost	SD	Medijana	%
Četbot je olakšao razumevanje sadržaja lekcije.	4,54	0,66	5	91,5
Četbot se pokazao kao korisna pomoć prilikom rešavanja zadataka.	4,54	0,70	5	88,6
Odgovori četbota su bili jasni.	4,57	0,65	5	91,4
Odgovori četbota su bili korisni.	4,63	0,55	5	97,1
Interakcija sa četbotom mi je povećala motivaciju za korišćenje platforme.	4,20	0,96	5	77,1
Ponovo bih koristio/la četbot kao pomoć pri rešavanju zadataka.	4,77	0,65	5	88,5
Osećam se sigurno kada se oslanjam na odgovore koje daje četbot.	4,43	0,74	5	91,4
Osećam se sigurno dok radim zadatke zbog dostupnosti četbota.	4,57	0,61	5	94,3
Četbot pruža pouzdane informacije kada mu postavim pitanje.	4,60	0,69	5	88,5

Na osnovu rezultata ankete može se zaključiti da su studenti generalno zadovoljni odgovorima četbota i pomoći koju pruža tokom savladavanja gradiva.

10.2.4 Diskusija rezultata

Na osnovu dobijenih rezultata možemo zaključiti da je u ovoj sesiji učenja konverzacioni agent uticao na povećanje angažovanja studenata kroz metriku vremena zadržavanja na samim jedinicama učenja, odnosno vremena za prolaženje kroz materijale za učenje i vođenje konverzacije za razjašnjavanje nedoumica i produbljivanje znanja. To je potencijalno uticalo i na bolje rezultate testova nakon korišćenja platforme, kao i rezultate zadataka tokom sesije učenja. Takođe, dostupnost četbota tokom izrade zadataka u ograničenom vremenskom periodu je pozitivno uticala na tačnost zadataka, kao i na sigurnost studenata u svoja rešenja, što je i očekivano s obzirom na pomoć koju pruža.

Takođe, korišćenje četbota tokom izrade zadataka je uticalo na smanjenje navigacije odnosno kretanja kroz jedinice učenja u ovoj sesiji. Studenti su mogli da razjasne nejasnoće, dobiju relevantne i fokusirane odgovore na određena teorijska pitanja vezana za lekciju na platformi, i dodatna pojašnjenja, ukoliko je bilo potrebe, kroz četbot funkcionalnost. Ovo može pozitivno uticati na upotrebljivost platforme u vidu poboljšanja korisničkog iskustva sa aspekta pretrage platforme i snalaženja u materijalima, kao i u smislu adaptacije sadržaja za učenje potrebama studenta.

Sve navedeno je potencijalno uticalo na višu SUS ocenu za verziju platforme koja integriše četbota, u odnosu na verziju bez, kao i na percipiranu korisnost, lakoću korišćenja i zadovoljstvo platformom. Samim tim može se zaključiti da upotrebljivost platforme u ovom slučaju može biti poboljšana uz pravilnu integraciju četbot funkcionalnosti.

SUS ocena, percipirana korisnost, lakoća korišćenja i zadovoljstvo platformom pokazuju da su studenti su takođe generalno zadovoljni i platformom koja ne poseduje funkcionalnost četbota na šta može uticati sam sadržaj materijala za učenje i njegov način prikaza kroz jedinice za učenje i kratke aktivnosti. Na osnovu ovoga možemo zaključiti da samo zadovoljstvo studenta platformom za učenje je uslovljeno i kvalitetom samih materijala pripremljenih od strane nastavnika.

10.2.5 Diskusija sa studentima

Nakon završetka korišćenja obe verzije platforme sproveden je razgovor sa studentima na dobrovoljnoj bazi. Obe verzije platforme su na studente ostavile generalno pozitivan utisak, s obzirom na to da im se svidelo da gradivo bude podeljeno u male celine koje se nadograđuju i ne sadrže previše detalja u toj početnoj fazi učenja već su fokusiranije. Neki studenti navode da je platforma pregledna i da ih vodi kroz gradivo tako da doprinosi koncentraciji i motivaciji, odnosno samom procesu učenja. Takođe im se svideo i sam sadržaj lekcija. Međutim, četbot funkcionalnost tokom izrade zadataka im je pružila veću sigurnost u svoje odgovore, a takođe im je dala mogućnost dodatnog pojašnjenja delova lekcija čime su zadovoljni. Određeni studenti su naveli da su zadovoljni time što su uz četbota mogli da potvrde razumevanje određenih koncepata, kao i da savladaju koncepte koji im nisu bili jasni skroz na osnovu prvog čitanja lekcije. Iako su materijali sa platforme bili dovoljni za izradu svih zadataka, studenti koji su koristili četbot funkcionalnost su ipak osetili viši nivo sigurnosti u svoje odgovore.

Studenti su naveli da su do sada navikli da sa VJM četbotovima često komuniciraju tokom pripreme ispita i da je ta komunikacija usmerena ka dobijanju direktnih odgovora na ispitna pitanja za šta navode da smatraju i da su svesni da nije najbolji način za učenje. Kod predložene platforme određenim studentima se svideo koncept u kome prvo uče o određenoj temi kroz dostupan materijal, a zatim kroz razgovor sa četbotom dodatno razjašnjavaju nedoumice i dublje zalaze u temu. Takođe, svidelo im se što mogu da mu postavljaju pitanja tako da dobiju objašnjenja za neke koncepte koja ne postoje u samim materijalima na platformi. Naveli su da se osećaju dosta sigurnije jer je četbot dostupan na kraju lekcije i da im to smanjuje neizvesnost koju osećaju tokom učenja u smislu razumevanja lekcije. Svaki put kad naiđu na nepoznat ili nejasan

koncept, znaju da mogu da pitaju četbota za pojašnjenje i svidelo im se to što je integrisan u okviru jedinica za učenje.

Takođe se određenim studentima svidelo što mogu brže da dobiju odgovor na neko pitanje koje ih zanima bez da se vraćaju na gradivo i pretražuju materijale odnosno što im olakšava „pretragu“ materijala. S druge strane, određeni studenti smatraju da će više zapamtiti ukoliko ipak ručno pretražuju materijale prolaskom kroz nastavne jedinice, ali takođe navode da im četbot može pomoći oko nedoumica koje se javljaju u procesu učenja.

Neki smatraju da je verzija sa četbotom interaktivnija, dinamičnija, zanimljivija, zabavnija i da duže drži pažnju, motiviše i zadržava na platformi za učenje. Probudila im je interesovanje i za koncepte koji nisu pokriveni lekcijom. Smatraju da kroz verziju sa četbotom mogu da istraže pojedinosti i samim tim im koncepti budu jasniji.

Studenti takođe osećaju da im je mnogo lakše da postave pitanje četbotu, nego asistentima na vežbama, jer im se dešava da naiđu na osudu ukoliko je pitanje vezano za gradivo za koje se očekuje predznanje sa drugih predmeta. U tom smislu im on olakšava nadoknadu gradiva koje nisu u potpunosti prethodno savladali, ili ponavljanje prethodno učenih koncepata.

Određenim studentima se svidelo što im je četbot postavljao pitanja tokom konverzacije, posebno tokom izrade zadataka u kojima su kroz razgovor sa njim dolazili do rešenja. Na taj način ih podstiče da dublje razmisle o onome što uče. Takođe, tokom diskusije je spomenuto da su odgovori ovog četbota usklađeniji sa terminima iz knjige što nije bio slučaj kada su koristili komercijalne četbot aplikacije zasnovane na velikim jezičkim modelima za pomoć pri učenju ovog predmeta. Ovaj utisak je direktna posledica konteksta konverzacijonog agenta, odnosno korišćenja knjige sa predmeta kao dodatnog konteksta prilikom generisanja odgovora.

Takođe, određena pitanja koja su data studentima na kraju jedinica učenja kao primeri pitanja koja mogu postavljati četbotu su na njih ostavila pozitivan utisak, jer smatraju da ih se samostalno ne bi setili, a pomogla su im u dubljem razumevanju lekcije.

Na osnovu svega navedenog zaključujemo da ne možemo odbaciti hipotezu (H2) da edukativna platforma koja integriše četbota zasnovanog na velikom jezičkom modelu, razvijena primenom definisanog okvira, pokazuje viši stepen upotrebljivosti u odnosu na tradicionalne sisteme za elektronsko učenje, što govori u prilog daljem nastavku istraživanja u ovom domenu.

S druge strane, tokom diskusije sa studentima uočeni su određeni prostori za poboljšanja. Npr. iako su jedinice učenja u okviru samih materijala na platformi sadržale teoriju koju je student trebalo da savlada i nauči, a četbot da bude podrška u tom procesu kroz koju mogu da prodube i prošire znanje, određeni studenti su spomenuli da je četbot ponekad davao preopširne odgovore, misleći da moraju u ovoj sesiji učenja da zapamte sve što im je on rekao. Iako to nije bio slučaj, svakako se može razmisliti o ograničenju broja izlaznih tokena što u trenutnoj verziji sistema nije urađeno. Ovo takođe otvara i prostor za dalja istraživanja u pogledu kvaliteta samih odgovora odnosno konteksta na osnovu kog četbot generiše odgovore. Takođe, određeni studenti

su spomenuli da bi im bilo korisno da su tokom izrade zadatka mogli da pristupe i prethodnoj konverzaciji koju su vodili tokom savladavanja lekcije. S druge strane, iako je fokus modelovanih i implementiranih funkcionalnosti platforme bio na dostavljanju materijala kroz jedinice za učenje i četbot funkciji, studenti su svejedno prokomentarisali mogućnost proširenja funkcionalnosti radi poboljšanja iskustva učenja, kao što je obeležavanje pređenih lekcija, obeležavanje teksta, ubacivanje kraćih kvizova pored aktivnosti radi provere naučenog i sl. što može biti razmatrano u daljem radu.

11 ZAKLJUČAK

Značajan napredak u razvoju velikih jezičkih modela odnosno generativne veštačke inteligencije je stvorio prilike za njihovu primenu u različitim domenima. Jedan od domena u kome četbotovi zasnovani na ovoj tehnologiji pronalaze primenu predstavlja obrazovanje. Dostupnost komercijalnih četbotova ovog tipa je izazvala negativne posledice u edukaciji zbog porasta plagijarizma i negativnog uticaja na kritičko razmišljanje studenata. Međutim, potencijal za njihovu pozitivnu primenu u visokom obrazovanju postoji, i njihova odgovorna upotreba i adekvatna integracija u okviru sistema za elektronsko učenje može uticati na poboljšanje iskustva učenja. Samim tim, polazna hipoteza autora je da je moguće razviti okvir za razvoj edukativne platforme koja integriše četbota zasnovanog na velikom jezičkom modelu.

Kroz tezu su predstavljeni rezultati koji su proizašli iz ovog istraživanja i koji mogu služiti u procesu razvoja edukativnih platformi koje integrišu konverzionog agenta zasnovanog na velikom jezičkom modelu.

Dakle, u ovoj doktorskoj disertaciji je istraživao potencijal primene generativne veštačke inteligencije odnosno velikih jezičkih modela u domenu obrazovanja kroz kreiranje i primenu objedinjenog modela za unapređenje razvoja edukativnih platformi koje integrišu konverzionog agenta zasnovanog na velikom jezičkom modelu. Ovako definisan model kroz MDA pristup obuhvata važne koncepte koje je potrebno uzeti u obzir prilikom izrade ovih sistema i samim tim može se posmatrati i kao ontologija znanja koja su potrebna za njihov razvoj. Primenom ovog pristupa, model je definisan na način tako da više oslikava sam domen koji modeluje, a da bude manje zavisao od implementacione platforme, što olakšava modelovanje i razvoj sistema.

Na osnovu objedinjenog modela za razvoj edukativne platforme koja integriše konverzionog agenta zasnovanog na velikom jezičkom modelu, nastalog primenom MDA pristupa i UML jezika, definisana su proširenja (UML profili) za potrebe modelovanja ovakvog sistema, sa primerima njihove primene. Takođe je definisan predlog procesa razvoja ovakve edukativne platforme kroz proširenje Unified procesa za razvoj softvera. U okviru faza, odnosno radnih tokova ovog procesa, koji obuhvataju specifikaciju, analizu, projektovanje, implementaciju i testiranje, moguće je koristiti definisana proširenja radi njihovog detaljnijeg opisa.

Ovako definisani metamodeli, proširenja i proces razvoja mogu biti primenjeni od strane projekatara i programera radi unapređenja procesa razvoja sistema za podršku učenju koji integrišu četbota zasnovanog na velikom jezičkom modelu. Hipoteze H1 i H3 su proverene metodom razvoja prototipa. Upotrebljeni su definisani metamodeli, proširenja i proces razvoja kako bi demonstrirali prednosti upotrebe objedinjenog okvira prilikom razvoja ovakvih sistema. Ovakav pristup predstavlja unapređenje u odnosu na razvoj bez okvira što za posledicu ima olakšani razvoj kvalitetnijih sistema.

Kroz upotrebu Angular i Express.js okvira za razvoj klijentske i serverske aplikacije, razvijene su dve verzije prototipa sistema za elektronsko učenje koji integriše četbota. Izvršeno je testiranje i evaluacija samog sistema od strane studenata treće godine sa smera Informacioni sistemi i tehnologije na Fakultetu organizacionih nauka. Kroz izradu

ovih prototipa smo potvrdili korisnost definisanih metamodela koji su omogućili modelovanje delova ovog sistema.

Prva verzija prototipa je kroz sprovedeni eksperiment sa studentima koji su koristili datu platformu testirala mogućnost primene nastale edukativne platforme na predmetu Računarske mreže i telekomunikacije, kroz evaluaciju odgovora četbota od strane eksperata, kao i evaluaciju četbot funkcionalnosti od strane studenata. Rezultati evaluacija ovog eksperimenta su generalno pozitivni, pri čemu su studenti iskazali pozitivne percepcije u pogledu korisnosti integrisane četbot funkcionalnosti u okviru obrazovne platforme. Kroz drugu verziju prototipa je izvršena evaluacija upotrebljivosti platforme sa četbotom, koja je pokazala viši stepen upotrebljivosti u odnosu na platformu bez istog, u okviru ovog eksperimenta. Ovi rezultati ukazuju na generalno zadovoljstvo studenata iskustvom učenja koje pruža prototip koji je jedan od rezultata ove disertacije i otvaraju prostor za dalja istraživanja u pogledu njegove nadogradnje, kao i potencijalne primene na drugim predmetima.

Pored ovih doprinosa, na početku rada je dat osvrt na oblasti bitne za ovu disertaciju kao što su razvoj velikih jezičkih modela, elektronsko učenje, primena četbotova u obrazovanju i pre razvoja velikih jezičkih modela, mogućnosti primene četbotova kao što je ChatGPT u obrazovanju, kao i istraživanja o integraciji ovih modela u sisteme za elektronsko učenje. Pregled prikazan u ovim poglavljima može se koristiti i samostalno kao sažet uvod u ove oblasti.

Takođe, predstavljeni su rezultati istraživanja sprovedenog na Fakultetu organizacionih nauka o faktorima koji utiču na prihvatanje ChatGPT-ja kao alata u obrazovnom kontekstu. Rezultati dobijeni tokom ovog istraživanja mogu biti korisni nastavnicima i dizajnerima edukativnih platformi kako bi mogli da poboljšaju iskustvo učenja ukoliko planiraju integraciju ovih alata u nastavni proces.

Prikazani rezultati predstavljaju dobru osnovu za dalje istraživanje ove oblasti u cilju unapređenja iskustva u procesu učenja kroz balansiranu upotrebu ovih tehnologija. U trenutku pisanja disertacije prisutna je pojava novih, unapređenijih modela čime se otvaraju prilike za različite mogućnosti primene. U nastavku će biti opisani neki od uočenih mogućih pravca daljeg istraživanja. Pre svega, implementirana platforma se fokusira na studente, tako da pored unapređenja funkcionalnosti postojeće platforme, budući razvoj može se usmeriti ka kreiranju nastavničkog portala za upravljanje i nadgledanje korišćenja date platforme, kao i uspostavljanje dugoročnih mehanizama evaluacije i nadgledanja sistema radi smanjenja halucinacija u sistemu. Takođe, personalizacija odgovora u zavisnosti od profila studenta i njegovog napretka, kao i dalje unapređenje odgovora sa pedagoške strane, može predstavljati još jedan potencijalan pravac daljeg istraživanja.

12 REFERENCE

- [1] T. Ait Baha, M. El Hajji, Y. Es-Saady, and H. Fadili, 'The impact of educational chatbot on student learning experience', *Educ. Inf. Technol.*, vol. 29, no. 8, pp. 10153–10176, Jun. 2024, doi: 10.1007/s10639-023-12166-w.
- [2] C. W. Okonkwo and A. Ade-Ibijola, 'Chatbots applications in education: A systematic review', *Comput. Educ.: Artif. Intell.*, vol. 2, 2021, doi: 10.1016/j.caeai.2021.100033.
- [3] D. C. Karthikeyan, 'Literature Review on Pros and Cons of ChatGPT Implications in Education', *Int. J. Sci. Res.*, vol. 12, no. 3, pp. 283–291, Mar. 2023, doi: 10.21275/SR23219122412.
- [4] X. Qi, Z. Zhu, and B. Wu, 'The promise and peril of ChatGPT in geriatric nursing education: What We know and do not know', *Aging and Health Research*, vol. 3, no. 2, Jun. 2023.
- [5] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson, 2022.
- [6] U. Mittal, S. Sai, V. Chamola, and D. Sangwan, 'A Comprehensive Review on Generative AI for Education', *IEEE Access*, vol. 12, pp. 142733–142759, 2024, doi: 10.1109/ACCESS.2024.3468368.
- [7] I. Gligorea, M. Cioca, R. Oancea, A.-T. Gorski, H. Gorski, and P. Tudorache, 'Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review', *Educ. Sci.*, vol. 13, no. 12, Dec. 2023, doi: 10.3390/educsci13121216.
- [8] M. U. Hadi *et al.*, 'A Survey on Large Language Models: Applications, Challenges, Limitations, and Practical Usage', Jul. 10, 2023, *TechRxiv*. doi: 10.36227/techrxiv.23589741.v1.
- [9] J. Hirschberg and C. D. Manning, 'Advances in natural language processing', *Science*, vol. 349, no. 6245, pp. 261–266, Jul. 2015, doi: 10.1126/science.aaa8685.
- [10] K. R. Chowdhary, *Fundamentals of Artificial Intelligence*. New Delhi: Springer India, 2020. doi: 10.1007/978-81-322-3972-7.
- [11] Y. Chang *et al.*, 'A Survey on Evaluation of Large Language Models', *ACM Trans. Intell. Syst. Technol.*, vol. 15, no. 3, pp. 1–45, Jun. 2024, doi: 10.1145/3641289.
- [12] L. Qin *et al.*, 'Large Language Models Meet NLP: A Survey', May 21, 2024, arXiv:2405.12819. doi: 10.48550/arXiv.2405.12819.
- [13] H. Naveed *et al.*, 'A Comprehensive Overview of Large Language Models', Oct. 17, 2024, arXiv:2307.06435. doi: 10.48550/arXiv.2307.06435.
- [14] S. Minaee *et al.*, 'Large Language Models: A Survey', Feb. 20, 2024, arXiv:2402.06196. doi: 10.48550/arXiv.2402.06196.
- [15] W. X. Zhao *et al.*, 'A Survey of Large Language Models', Oct. 13, 2024, arXiv:2303.18223. doi: 10.48550/arXiv.2303.18223.
- [16] A. Vaswani *et al.*, 'Attention Is All You Need', Aug. 02, 2023, arXiv:1706.03762. doi: 10.48550/arXiv.1706.03762.
- [17] E. Kasneci *et al.*, 'ChatGPT for good? On opportunities and challenges of large language models for education', *Learn. Individ. Differ.*, vol. 103, Apr. 2023, doi: 10.1016/j.lindif.2023.102274.
- [18] N. S. Keskar, B. McCann, L. R. Varshney, C. Xiong, and R. Socher, 'CTRL: A Conditional Transformer Language Model for Controllable Generation', Sep. 20, 2019, arXiv:1909.05858. doi: 10.48550/arXiv.1909.05858.

- [19] J. Zhou, P. Ke, X. Qiu, M. Huang, and J. Zhang, 'ChatGPT: potential, prospects, and limitations', *Front. Inf. Technol. Electron. Eng.*, vol. 25, no. 1, pp. 6–11, 2024, doi: 10.1631/FITEE.2300089.
- [20] H. Zhou *et al.*, 'A Survey of Large Language Models in Medicine: Progress, Application, and Challenge', 2024, arXiv:2311.05112. doi: 10.48550/arXiv.2311.05112.
- [21] Y. Li, S. Wang, H. Ding, and H. Chen, 'Large Language Models in Finance: A Survey', in *Proceedings of the 4th ACM International Conference on AI in Finance*, Brooklyn, NY, USA: ACM, Nov. 2023, pp. 374–382. doi: 10.1145/3604237.3626869.
- [22] D. O. Eke, 'ChatGPT and the rise of generative AI: Threat to academic integrity?', *J. Responsible Technol.*, vol. 13, Apr. 2023, doi: 10.1016/j.jrt.2023.100060.
- [23] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, 'Improving Language Understanding by Generative Pre-Training', OPENAI, 2018.
- [24] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, 'Language Models are Unsupervised Multitask Learners', OPENAI, 2019.
- [25] T. B. Brown *et al.*, 'Language Models are Few-Shot Learners', Jul. 22, 2020, arXiv:2005.14165. doi: 10.48550/arXiv.2005.14165.
- [26] OPENAI, 'GPT-4 Technical Report', arXiv:2303.08774, Mar. 2023. doi: 10.48550/arXiv.2303.08774.
- [27] M. Zhang and J. Li, 'A commentary of GPT-3 in MIT Technology Review 2021', *Fundam. Res.*, vol. 1, no. 6, pp. 831–833, Nov. 2021, doi: 10.1016/j.fmre.2021.11.011.
- [28] A. Birhane, A. Kasirzadeh, D. Leslie, and S. Wachter, 'Science in the age of large language models', *Nat. Rev. Phys.*, vol. 5, no. 5, pp. 277–280, Apr. 2023, doi: 10.1038/s42254-023-00581-4.
- [29] A. Haleem, M. Javaid, and R. P. Singh, 'An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges', *BenchCouncil Trans. Benchmarks Stand. Eval.*, vol. 2, no. 4, Oct. 2022, doi: 10.1016/j.tbench.2023.100089.
- [30] W. M. Lim, A. Gunasekara, J. L. Pallant, J. I. Pallant, and E. Pechenkina, 'Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators', *Int. J. Manage. Educ.*, vol. 21, no. 2, Jul. 2023, doi: 10.1016/j.ijme.2023.100790.
- [31] I. Molenaar, 'The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning', *Comput. Educ.: Artif. Intell.*, vol. 3, 2022, doi: 10.1016/j.caeai.2022.100070.
- [32] J. Savelka, A. Agarwal, C. Bogart, and M. Sakr, 'Large Language Models (GPT) Struggle to Answer Multiple-Choice Questions about Code', Mar. 09, 2023, arXiv:2303.08033v1.
- [33] D. Gašević, G. Siemens, and S. Sadiq, 'Empowering learners for the age of artificial intelligence', *Comput. Educ.: Artif. Intell.*, vol. 4, Feb. 2023, doi: 10.1016/j.caeai.2023.100130.
- [34] A. Tlili *et al.*, 'What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education', *Smart Learn. Environ.*, vol. 10, no. 1, Feb. 2023, doi: 10.1186/s40561-023-00237-x.
- [35] G. Cooper, 'Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence', *J. Sci. Educ. Technol.*, vol. 32, no. 3, pp. 444–452, Jun. 2023, doi: 10.1007/s10956-023-10039-y.

- [36] M. Liu and D. Yu, 'Towards intelligent E-learning systems', *Educ. Inf. Technol.*, vol. 28, no. 7, pp. 7845–7876, Jul. 2023, doi: 10.1007/s10639-022-11479-6.
- [37] B. Holmes and J. Gardner, *E-learning: concepts and practice*. London: SAGE Publications, 2006.
- [38] O. O. Jethro, A. M. Grace, and A. K. Thomas, 'E-Learning and Its Effects on Teaching and Learning in a Global Age', *Int. J. Acad. Res. Bus. Soc. Sci.*, vol. 2, no. 1, pp. 203–210, 2012.
- [39] P. A. Reynolds, R. Mason, and J. Harper, 'The many faces of interaction', *Br. Dent. J.*, vol. 204, no. 10, pp. 565–570, May 2008, doi: 10.1038/sj.bdj.2008.409.
- [40] S. Wollny, J. Schneider, D. Di Mitri, J. Weidlich, M. Rittberger, and H. Drachsler, 'Are We There Yet? - A Systematic Literature Review on Chatbots in Education', *Front. Artif. Intell.*, vol. 4, Jul. 2021, doi: 10.3389/frai.2021.654924.
- [41] Z.-Y. Liu, N. Lomovtseva, and E. Korobeynikova, 'Online Learning Platforms: Reconstructing Modern Higher Education', *Int. J. Emerg. Technol. Learn.*, vol. 15, no. 13, pp. 4–21, Jul. 2020, doi: 10.3991/ijet.v15i13.14645.
- [42] A. Ekuase-Anwansedo, 'The Use of Learning Management Systems and Covid-19: Perceptions in the Last two Years', in *Proceedings of the 2022 ACM SIGUCCS Annual Conference*, Virtual Event, USA: ACM, Mar. 2022, pp. 2–2. doi: 10.1145/3501292.3511587.
- [43] C. Perry, 'Online Learning Platforms: The Different Types And Their Benefits – Forbes Advisor', *Forbes*. [Online]. Available: <https://www.forbes.com/advisor/education/career-resources/online-learning-platforms/>
- [44] L. H. A. de Andrade, D. A. Thomas, and V. Laterza, 'The rise of EdTech platforms in higher education: Mapping themes from emerging critical literature', in *Digital Transformations in Nordic Higher Education*, Cham, Switzerland: Palgrave Macmillan, 2023, pp. 27–51. doi: 10.1007/978-3-031-27758-0_2.
- [45] V. Arkorful and N. Abaidoo, 'The role of e-learning, the advantages and disadvantages of its adoption in Higher Education.', *Int. J. Educ. Res.*, vol. 2, no. 12, pp. 397–410, 2014.
- [46] S. K. Banihashem, O. Noroozi, P. Den Brok, H. J. A. Biemans, and N. T. Kerman, 'Modeling teachers' and students' attitudes, emotions, and perceptions in blended education: Towards post-pandemic education', *Int. J. Manag. Educ.*, vol. 21, no. 2, Jul. 2023, doi: 10.1016/j.ijme.2023.100803.
- [47] G.-J. Hwang and C.-Y. Chang, 'A review of opportunities and challenges of chatbots in education', *Interact. Learn. Environ.*, vol. 31, no. 7, pp. 4099–4112, Oct. 2023, doi: 10.1080/10494820.2021.1952615.
- [48] M. A. Kuhail, N. Alturki, S. Alramlawi, and K. Alhejori, 'Interacting with educational chatbots: A systematic review', *Educ. Inf. Technol.*, vol. 28, no. 1, pp. 973–1018, Jan. 2023, doi: 10.1007/s10639-022-11177-3.
- [49] E. Adamopoulou and L. Moussiades, 'Chatbots: History, technology, and applications', *Machine Learning with Applications*, vol. 2, Dec. 2020, doi: 10.1016/j.mlwa.2020.100006.
- [50] S. Cunningham-Nelson, W. Boles, L. Trouton, and E. Margerison, 'A Review of Chatbots in Education: Practical Steps Forward', in *Proceedings of the 30th Annual Conference for the Australasian Association for Engineering Education (AAEE 2019): Educators Becoming Agents of Change: Innovate, Integrate, Motivate*, Brisbane, Queensland: Engineers Australia, 2019, pp. 299–306. doi: 10.3316/informit.068364390172788.

- [51] J. Q. Pérez, T. Daradoumis, and J. M. M. Puig, 'Rediscovering the use of chatbots in education: A systematic literature review', *Comput. Appl. Eng. Educ.*, vol. 28, no. 6, pp. 1549–1565, Nov. 2020, doi: 10.1002/cae.22326.
- [52] A. Z. Al Rawashdeh, E. Y. Mohammed, A. R. Al Arab, M. Alara, B. Al-Rawashdeh, and B. Al-Rawashdeh, 'Advantages and Disadvantages of Using e-Learning in University Education: Analyzing Students' Perspectives', *Electronic Journal of e-Learning*, vol. 19, no. 3, pp. 107–117, May 2021, doi: 10.34190/ejel.19.3.2168.
- [53] N. Islam, M. Beer, and F. Slack, 'E-Learning Challenges Faced by Academics in Higher Education: A Literature Review', *J. Educ. Train. Stud.*, vol. 3, no. 5, pp. 102–112, Jul. 2015, doi: 10.11114/jets.v3i5.947.
- [54] W. Huang, K. F. Hew, and L. K. Fryer, 'Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning', *J. Comput. Assist. Learn.*, vol. 38, no. 1, pp. 237–257, Feb. 2022, doi: 10.1111/jcal.12610.
- [55] P. B. Brandtzaeg and A. Følstad, 'Chatbots: changing user needs and motivations', *Interactions*, vol. 25, no. 5, pp. 38–43, Aug. 2018, doi: 10.1145/3236669.
- [56] F. Clarizia, F. Colace, M. Lombardi, F. Pascale, and D. Santaniello, 'Chatbot: An Education Support System for Student', in *Cyberspace Safety and Security*, vol. 11161, A. Castiglione, F. Pop, M. Ficco, and F. Palmieri, Eds, in Lecture Notes in Computer Science, vol. 11161, Cham, Switzerland: Springer International Publishing, 2018, pp. 291–302. doi: 10.1007/978-3-030-01689-0_23.
- [57] J. M. Markel, S. G. Opferman, J. A. Landay, and C. Piech, 'GPTEach: Interactive TA Training with GPT Based Students', in *Proceedings of the Tenth ACM Conference on Learning @ Scale*, Copenhagen, Denmark, 2023, pp. 226–236. doi: 10.1145/3573051.3593393.
- [58] J. Su and W. Yang, 'Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education', *ECNU Rev. Educ.*, vol. 6, no. 3, pp. 355–366, Aug. 2023, doi: 10.1177/20965311231168423.
- [59] Y. Wei, Y.-H. Jiang, J. Liu, C. Qi, and R. Jia, 'The Advancement of Personalized Learning Potentially Accelerated by Generative AI', 2024, *arXiv:2412.00691v1*. doi: 10.48550/arXiv.2412.00691v1.
- [60] J. Lively, J. Hutson, and E. Melick, 'Integrating AI-Generative Tools in Web Design Education: Enhancing Student Aesthetic and Creative Copy Capabilities Using Image and Text-Based AI Generators', *DS J. Artif. Intell. Robot.*, vol. 1, no. 1, pp. 23–36, Sep. 2023, doi: 10.59232/AIR-V111P103.
- [61] X. Zhai, 'ChatGPT User Experience: Implications for Education', *SSRN Electron. J.*, 2022, doi: 10.2139/ssrn.4312418.
- [62] M. Warr, P. Mishra, D. Henriksen, and L. J. Woo, 'A Chat about GPT3 (and Other Forms of Alien Intelligence) with Chris Dede', *TechTrends*, vol. 67, pp. 396–401, Apr. 2023, doi: 10.1007/s11528-023-00843-z.
- [63] B. A. Anders, 'Is using ChatGPT cheating, plagiarism, both, neither, or forward thinking?', *Patterns*, vol. 4, no. 3, Mar. 2023, doi: 10.1016/j.patter.2023.100694.
- [64] B. Zhang, 'Preparing Educators and Students for ChatGPT and AI Technology in Higher Education: Benefits, Limitations, Strategies, and Implications of ChatGPT & AI Technologies', 2023.
- [65] S. Rana, 'AI and GPT for Management Scholars and Practitioners: Guidelines and Implications', *FIIB Bus. Rev.*, vol. 12, no. 1, pp. 7–9, Mar. 2023, doi: 10.1177/23197145231161408.

- [66] S. N. Yildirim-Erbasli and O. Bulut, 'Conversation-based assessment: A novel approach to boosting test-taking effort in digital formative assessment', *Comput. Educ.: Artif. Intell.*, vol. 4, 2023, doi: 10.1016/j.caeai.2023.100135.
- [67] H. N. Salim, 'The Causes of Student's Reluctance to Ask Question's When Attending Lectures', *Teaching English and Language Learning English Journal*, vol. 2, no. 3, pp. 208–213, Dec. 2022, doi: 10.36085/telle.v2i3.4719.
- [68] J. Qadir, 'Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education', Dec. 30, 2022. doi: 10.36227/techrxiv.21789434.v1.
- [69] A. T. Bimba, N. Idris, R. B. Mahmud, and A. Al-Hunaiyyan, 'A Cognitive Knowledge-based Framework for Adaptive Feedback', in *Computational Intelligence in Information Systems*, vol. 532, S. Phon-Amnuaisuk, T.-W. Au, and S. Omar, Eds, in *Advances in Intelligent Systems and Computing*, vol. 532. , Cham, Switzerland: Springer, 2017, pp. 245–255. doi: 10.1007/978-3-319-48517-1_22.
- [70] Y. K. Dwivedi *et al.*, 'Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy', *Int. J. Inf. Manag.*, vol. 71, Aug. 2023, doi: 10.1016/j.ijinfomgt.2023.102642.
- [71] F. D. Davis, *Foreword in Technology Acceptance in Education: Research and Issues*. Sense Publishers, 2011.
- [72] R. Ibrahim, N. S. Leng, R. C. M. Yusoff, G. N. Samy, S. Masrom, and Z. I. Rizman, 'E-learning acceptance based on technology acceptance model (TAM)', *J. Fundam. Appl. Sci.*, vol. 9, no. 4S, pp. 871–889, 2017, doi: 10.4314/jfas.v9i4s.50.
- [73] S. S. Saleh, M. Nat, and M. Aqel, 'Sustainable Adoption of E-Learning from the TAM Perspective', *Sustainability*, vol. 14, no. 6, Mar. 2022, doi: 10.3390/su14063690.
- [74] S. Zhang, J. Zhao, and W. Tan, 'Extending TAM for online learning systems: An intrinsic motivation perspective', *Tsinghua Sci. Technol.*, vol. 13, no. 3, pp. 312–317, Jun. 2008, doi: 10.1016/S1007-0214(08)70050-6.
- [75] M. Al-Emran, V. Mezhuyev, and A. Kamaludin, 'Technology Acceptance Model in M-learning context: A systematic review', *Comput. Educ.*, vol. 125, pp. 389–412, Oct. 2018, doi: 10.1016/j.compedu.2018.06.008.
- [76] I. Milošević, A. Rakić, S. Arsić, A. Stojanović, I. Nikolić, and P. Đorđević, 'Model for Considering the Propensity of Students to Accept M-learning', *Management: Journal of Sustainable Business and Management Solutions in Emerging Economies*, vol. 27, no. 1, pp. 31–44, 2022.
- [77] D. Mugo, K. Njagi, B. Chemwei, and J. Motanya, 'The Technology Acceptance Model (TAM) and its Application to the Utilization of Mobile Learning Technologies', *J. Adv. Math. Comput. Sci.*, vol. 20, no. 4, pp. 1–8, Jan. 2017, doi: 10.9734/BJMCS/2017/29015.
- [78] A. Al-Azawei, P. Parslow, and K. Lundqvist, 'Investigating the effect of learning styles in a blended e-learning system: An extension of the technology acceptance model (TAM)', *Australas. J. Educ. Technol.*, vol. 33, no. 2, Nov. 2016, doi: 10.14742/ajet.2741.
- [79] N. Fathema, D. Shannon, and M. Ross, 'Expanding The Technology Acceptance Model (TAM) to Examine Faculty Use of Learning Management Systems (LMSs) In Higher Education Institutions', *J. Online Learn. Teach.*, vol. 11, no. 2, pp. 210–232, 2015.
- [80] T. Budhathoki, A. Zirar, E. T. Njoya, and A. Timsina, 'ChatGPT adoption and anxiety: a cross-country analysis utilising the unified theory of acceptance and use

- of technology (UTAUT)', *Stud. High. Educ.*, vol. 49, no. 5, pp. 831–846, 2024, doi: 10.1080/03075079.2024.2333937.
- [81] Y. K. Dwivedi, N. Mustafee, L. D. Carter, and M. D. Williams, 'A Bibliometric Comparison of the Usage of Two Theories of IS/IT Acceptance (TAM and UTAUT)', in *Proceedings of the 16th Americas Conference on Information Systems (AMCIS 2010)*, Lima, Peru, 2010.
 - [82] P.-C. Sun, R. J. Tsai, G. Finger, Y.-Y. Chen, and D. Yeh, 'What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction', *Comput. Educ.*, vol. 50, no. 4, pp. 1183–1202, May 2008, doi: 10.1016/j.compedu.2006.11.007.
 - [83] C. Weng, C.-C. Tsai, and A. Weng, 'Social support as a neglected e-learning motivator affecting trainee's decisions of continuous intentions of usage', *Australas. J. Educ. Technol.*, vol. 31, no. 2, pp. 177–192, Apr. 2015, doi: 10.14742/ajet.1311.
 - [84] J. Stanojević, M. Minović, I. Kovačević, and M. Maričić, 'Evaluating the Usage of ChatGPT for Educational Purposes Through the Extended Technological Acceptance Model', *Int. J. Eng. Educ.*, vol. 41, no. 2, pp. 352–364, 2025.
 - [85] F. D. Davis, 'Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology', *MIS Q.*, vol. 13, no. 3, pp. 319–340, Sep. 1989.
 - [86] K. A. Pituch and Y. Lee, 'The influence of system characteristics on e-learning use', *Comput. Educ.*, vol. 47, no. 2, pp. 222–244, Sep. 2006, doi: 10.1016/j.compedu.2004.10.007.
 - [87] A. Tarhini, K. Hone, and X. Liu, 'Measuring the Moderating Effect of Gender and Age on E-Learning Acceptance in England: A Structural Equation Modeling Approach for An Extended Technology Acceptance Model', *J. Educ. Comput. Res.*, vol. 51, no. 2, pp. 163–184, Sep. 2014, doi: 10.2190/EC.51.2.b.
 - [88] F. Abdullah and R. Ward, 'Developing a General Extended Technology Acceptance Model for E-Learning (GETAMEL) by analysing commonly used external factors', *Comput. Human Behav.*, vol. 56, pp. 238–256, Mar. 2016, doi: 10.1016/j.chb.2015.11.036.
 - [89] C. Yoon and S. Kim, 'Convenience and TAM in a ubiquitous computing environment: The case of wireless LAN', *Electron. Commer. Res. Appl.*, vol. 6, no. 1, pp. 102–112, Mar. 2007, doi: 10.1016/j.eelerap.2006.06.009.
 - [90] Z. Dulcic, D. Pavlic, and I. Silic, 'Evaluating the Intended Use of Decision Support System (DSS) by Applying Technology Acceptance Model (TAM) in Business Organizations in Croatia', *Procedia - Soc. Behav. Sci.*, vol. 58, pp. 1565–1575, Oct. 2012.
 - [91] J. Pallant, *SPSS survival manual: a step by step guide to data analysis using IBM SPSS*, 7th ed. London: Routledge, 2020.
 - [92] J. F. Hair, M. C. Howard, and C. Nitzl, 'Assessing measurement model quality in PLS-SEM using confirmatory composite analysis', *Journal of Business Research*, vol. 109, pp. 101–110, Mar. 2020.
 - [93] C. Fornell and D. F. Larcker, 'Evaluating Structural Equation Models with Unobservable Variables and Measurement Error', *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50, Feb. 1981.
 - [94] L. Hu and P. M. Bentler, 'Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification.', *Psychol. Methods*, vol. 3, no. 4, pp. 424–453, Dec. 1998.

- [95] C. Y. Lai, K. Y. Cheung, and C. S. Chan, 'Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: An extension of the technology acceptance model', *Comput. Educ.: Artif. Intell.*, vol. 5, 2023.
- [96] N. Saif, S. U. Khan, I. Shaheen, F. A. ALotaibi, M. M. Alnfai, and M. Arif, 'Chat-GPT; validating Technology Acceptance Model (TAM) in education sector via ubiquitous learning mechanism', *Comput. Hum. Behav.*, vol. 154, May 2024, doi: 10.1016/j.chb.2023.108097.
- [97] C. K. Tiwari, M. A. Bhat, S. T. Khan, R. Subramaniam, and M. A. I. Khan, 'What drives students toward ChatGPT? An investigation of the factors influencing adoption and usage of ChatGPT', *Interact. Technol. Smart Educ.*, vol. 21, no. 3, pp. 333–355, 2024, doi: 10.1108/ITSE-04-2023-0061.
- [98] G. Liu and C. Ma, 'Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model', *Innovation in Language Learning and Teaching*, vol. 18, no. 2, pp. 125–138, Mar. 2024, doi: 10.1080/17501229.2023.2240316.
- [99] G. Bilquise, S. Ibrahim, and S. M. Salhieh, 'Investigating student acceptance of an academic advising chatbot in higher education institutions', *Educ. Inf. Technol.*, vol. 29, no. 5, pp. 6357–6382, Apr. 2024, doi: 10.1007/s10639-023-12076-x.
- [100] R. Malik, A. Shrama, S. Trivedi, and R. Mishra, 'Adoption of Chatbots for Learning among University Students: Role of Perceived Convenience and Enhanced Performance', *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 18, pp. 200–212, Sep. 2021, doi: 10.3991/ijet.v16i18.24315.
- [101] W. Aslam, D. Ahmed Siddiqui, I. Arif, and K. Farhat, 'Chatbots in the frontline: drivers of acceptance', *Kybernetes*, vol. 52, no. 9, pp. 3781–3810, Sep. 2023.
- [102] W. Gan, Z. Qi, J. Wu, and J. C.-W. Lin, 'Large Language Models in Education: Vision and Opportunities', Nov. 22, 2023, arXiv:2311.13160v1. doi: 10.48550/arXiv.2311.13160v1.
- [103] L. Luan, X. Lin, and W. Li, 'Exploring the Cognitive Dynamics of Artificial Intelligence in the Post-COVID-19 and Learning 3.0 Era: A Case Study of ChatGPT', Feb. 03, 2023, arXiv:2302.04818. doi: 10.48550/arXiv.2302.04818.
- [104] A. Shoufan, 'Exploring Students' Perceptions of ChatGPT: Thematic Analysis and Follow-Up Survey', *IEEE Access*, vol. 11, pp. 38805–38818, 2023, doi: 10.1109/ACCESS.2023.3268224.
- [105] J. Stanojević, M. Minović, M. Milovanović, and D. Simić, 'ChatGPT as a student support tool in programming', in *Digitalni i zeleni razvoj privrede: Zbornik radova XIV Skupa privrednika i naučnika SPIN '23*, Belgrade, Serbia: University of Belgrade, Faculty of Organizational Sciences, 2023.
- [106] J. Yin, S. Dash, J. P. Gounley, and F. Wang, 'Evaluation of pre-training large language models on leadership-class supercomputers', *The Journal of Supercomputing*, vol. 79, pp. 20747–20768, 2023.
- [107] D. M. Anisuzzaman, J. G. Malins, P. A. Friedman, and Z. I. Attia, 'Fine-Tuning Large Language Models for Specialized Use Cases', *Mayo Clinic Proceedings: Digital Health*, vol. 3, no. 1, Mar. 2025, doi: 10.1016/j.mcpdig.2024.11.005.
- [108] B. Mahalingam, 'Fine-Tuning Large Language Models – Challenges and Best Practices'. Accessed: May 09, 2025. [Online]. Available: <https://www.encora.com/insights/fine-tuning-large-language-models-challenges-and-best-practices>

- [109] M. Firat, 'Integrating AI Applications into Learning Management Systems to Enhance e-Learning', *Instructional Technology and Lifelong Learning*, vol. 4, no. 1, pp. 1–14, Jun. 2023, doi: 10.52911/itall.1244453.
- [110] C. Silvestri *et al.*, 'Evaluation of a novel large language model (LLM)-powered chatbot for oral boards scenarios', *Global Surg. Educ.*, vol. 3, no. 1, Nov. 2024, doi: 10.1007/s44186-024-00303-z.
- [111] C. Grévisse, 'RasPatient Pi: A Low-Cost Customizable LLM-Based Virtual Standardized Patient Simulator', in *Applied Informatics*, H. Florez and H. Astudillo, Eds, in Communications in Computer and Information Science, vol. 2237. Cham: Springer Nature Switzerland, 2025, pp. 125–137. doi: 10.1007/978-3-031-75147-9_9.
- [112] G. Arun *et al.*, 'ChatGPT versus a customized AI chatbot (Anatbuddy) for anatomy education: A comparative pilot study', *Anat. Sci. Educ.*, vol. 17, no. 7, pp. 1396–1405, Oct. 2024, doi: 10.1002/ase.2502.
- [113] Z. Zhang and X. Huang, 'The impact of chatbots based on large language models on second language vocabulary acquisition', *Heliyon*, vol. 10, no. 3, Feb. 2024, doi: 10.1016/j.heliyon.2024.e25370.
- [114] W. Lyu, Y. Wang, Tingting, Chung, Y. Sun, and Y. Zhang, 'Evaluating the Effectiveness of LLMs in Introductory Computer Science Education: A Semester-Long Field Study', in *Proceedings of the Eleventh ACM Conference on Learning @ Scale*, New York, NY, USA: Association for Computing Machinery, Jul. 2024, pp. 63–74. doi: 10.1145/3657604.3662036.
- [115] J. D. Zamfirescu-Pereira, L. Qi, B. Hartmann, J. DeNero, and N. Norouzi, '61A Bot Report: AI Assistants in CS1 Save Students Homework Time and Reduce Demands on Staff. (Now What?)', in *Proceedings of the 56th ACM Technical Symposium on Computer Science Education*, Pittsburgh, PA, USA: ACM, Feb. 2025, pp. 1309–1315. doi: 10.1145/3641554.3701864.
- [116] M. Liffiton, B. Sheese, J. Savelka, and P. Denny, 'CodeHelp: Using Large Language Models with Guardrails for Scalable Support in Programming Classes', Aug. 14, 2023, arXiv:2308.06921. doi: 10.48550/arXiv.2308.06921.
- [117] M. Kazemitabaar *et al.*, 'CodeAid: Evaluating a Classroom Deployment of an LLM-based Programming Assistant that Balances Student and Educator Needs', May 11, 2024, arXiv:2401.11314v2. doi: 10.1145/3613904.3642773.
- [118] X. Hou, Z. Wu, X. Wang, and B. J. Ericson, 'CodeTailor: LLM-Powered Personalized Parsons Puzzles for Engaging Support While Learning Programming', in *Proceedings of the Eleventh ACM Conference on Learning @ Scale*, Jul. 2024, pp. 51–62. doi: 10.1145/3657604.3662032.
- [119] B.-T. Tran-Le, P. Thomas, N. M. Stiffler, and T. N. Nguyen, 'How Does LLM-powered Coding Assistance Shape Incidental Learning? Exploring Cognitive Forcing Strategies in Programming Education', in *Proceedings of the Fortieth AAAI Conference on Artificial Intelligence*, 2026, pp. 40880–40888. doi: 10.1609/aaai.v40i48.42121.
- [120] I. Ma, A. K. Martins, and C. V. Lopes, 'Integrating AI Tutors in a Programming Course', Jul. 14, 2024, arXiv:2407.15718. doi: 10.48550/arXiv.2407.15718.
- [121] B. Tabarsi, A. Basarkar, X. Liu, D. (DK) Xu, and T. Barnes, 'MerryQuery: A Trustworthy LLM-Powered Tool Providing Personalized Support for Educators and Students', in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025, pp. 29700–29702. doi: <https://doi.org/10.1609/aaai.v39i28.35372>.

- [122] A. T. Neumann, Y. Yin, S. Sowe, S. Decker, and M. Jarke, 'An LLM-Driven Chatbot in Higher Education for Databases and Information Systems', *IEEE Trans. Educ.*, vol. 68, no. 1, pp. 103–116, Feb. 2025, doi: 10.1109/te.2024.3467912.
- [123] R. Liu, J. Zhao, B. Xu, C. Perez, Y. Zhukovets, and D. J. Malan, 'Improving AI in CS50: Leveraging Human Feedback for Better Learning', in *Proceedings of the 56th ACM Technical Symposium on Computer Science Education*, Pittsburgh, PA, USA: ACM, Feb. 2025, pp. 715–721. doi: 10.1145/3641554.3701945.
- [124] M. Wölfel, M. B. Shirzad, A. Reich, and K. Anderer, 'Knowledge-Based and Generative-AI-Driven Pedagogical Conversational Agents: A Comparative Study of Grice's Cooperative Principles and Trust', *Big Data Cogn. Comput.*, vol. 8, no. 1, Dec. 2023, doi: 10.3390/bdcc8010002.
- [125] OMG, 'MDA - The Architecture Of Choice For A Changing World', Object Management Group. Accessed: Feb. 26, 2025. [Online]. Available: <https://www.omg.org/mda/>
- [126] OMG, *MDA Guide rev. 2.0*, Jun. 01, 2014. Accessed: Sep. 05, 2025. [Online]. Available: <https://www.omg.org/cgi-bin/doc?ormsc/14-06-01>
- [127] I. Milenković, 'Okvir za evaluaciju multimodalnih biometrijskih sistema', Univerzitet u Beogradu, Beograd, 2021.
- [128] OMG, 'OMG Meta Object Facility (MOF) Core Specification'. 2019.
- [129] OMG, 'What is UML and Why is it Essential for Modern Software Development?' [Online]. Available: <https://www.omg.org/uml/what-is-uml.htm>
- [130] OMG, 'Meta Object Facility (MOF) Specification'. 2002. [Online]. Available: <https://www.omg.org/spec/MOF/>
- [131] OMG, 'MDA Specifications'. Accessed: Feb. 26, 2025. [Online]. Available: <https://www.omg.org/mda/specs.htm>
- [132] Ž. Gavrić, 'Model interakcije čovjeka i računara zasnovan na praćenju pokreta oka', Univerzitet u Beogradu, Beograd, 2024.
- [133] M. Minović, 'Razvoj edukativnih igara zasnovan na MDA pristupu', Univerzitet u Beogradu, Beograd, 2010.
- [134] R. Mochamad, 'Microlearning: A Bibliometric Analysis of Research Trends from 2002 to 2023', *J. Comput. Sci. Math. Learn.*, vol. 1, no. 1, pp. 14–26, Feb. 2024, doi: 10.70232/x6ybp70.
- [135] D. Chumdoem, S. Boonlue, and M. Narmluk, 'Investigating the Components of Chat-Bot-Driven Micro-Learning: Enhancing Problem-Solving Competency in Modern Teaching.', in *Proceedings of the ICTAECT 2024 and NCTAECT 2024*, 2024, pp. 55–72.
- [136] L. W. Anderson and D. R. Krathwohl, Eds, *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. Boston: Allyn & Bacon, 2001.
- [137] L. O. Wilson, 'Anderson and Krathwohl Bloom's Taxonomy Revised', 2016.
- [138] B. Chen, Z. Zhang, N. Langrené, and S. Zhu, 'Unleashing the potential of prompt engineering for large language models', *Patterns*, vol. 6, no. 6, Jun. 2025, doi: 10.1016/j.patter.2025.101260.
- [139] S. Ekin, 'Prompt Engineering for ChatGPT: A Quick Guide to Techniques, Tips, and Best Practices', *J. Eng. Technol.*, vol. 43, no. 1, pp. 52–68, Apr. 2026, doi: 10.18260/jet.43.1.5.
- [140] F. Stamenković, J. Stanojević, B. Jovanović, and M. Bogičević Sretenović, 'Guardrail mehanizmi u edukativnim platformama za programiranje zasnovanim na velikim jezičkim modelima', in *Digitalna transformacija lanaca snabdevanja u industriji*

- 4.0: *Zbornik radova XV Skupa privrednika i naučnika SPIN '25*, Belgrade, Serbia: University of Belgrade, Faculty of Organizational Sciences, 2025, pp. 217–223. doi: 10.5281/zenodo.18012049.
- [141] Y. Du *et al.*, 'Context Length Alone Hurts LLM Performance Despite Perfect Retrieval', 2025, arXiv:2510.05381v1. doi: 10.48550/arXiv.2510.05381.
 - [142] W. Wang, J. Min, and W. Zou, 'Intelligence Degradation in Long-Context LLMs: Critical Threshold Determination via Natural Length Distribution Analysis', Jan. 07, 2026, arXiv:2601.15300. doi: 10.48550/arXiv.2601.15300.
 - [143] Z. Zhang, J. Li, Y. Lan, X. Wang, and H. Wang, 'An Empirical Study on Prompt Compression for Large Language Models', Apr. 24, 2025, arXiv:2505.00019. doi: 10.48550/arXiv.2505.00019.
 - [144] M. Sung, S. Lee, J. Kim, and S. Kim, 'Context-Aware LLM Translation System Using Conversation Summarization and Dialogue History', in *Proceedings of the Ninth Conference on Machine Translation*, Miami, Florida, USA: Association for Computational Linguistics, 2024, pp. 1011–1015. doi: 10.18653/v1/2024.wmt-1.102.
 - [145] Q. Wang, Y. Fu, Y. Cao, S. Wang, Z. Tian, and L. Ding, 'Recursively Summarizing Enables Long-Term Dialogue Memory in Large Language Models', Aug. 25, 2025, arXiv:2308.15022. doi: 10.48550/arXiv.2308.15022.
 - [146] Ž. Obrenović, 'Interakcija čoveka i računara', FON, Beograd, 2004.
 - [147] Z. Obrenovic and D. Starcevic, 'Modeling multimodal human-computer interaction', *Computer*, vol. 37, no. 9, pp. 65–72, Sep. 2004, doi: 10.1109/MC.2004.139.
 - [148] M. Minsky, 'Commonsense-based interfaces', *Commun. ACM*, vol. 43, no. 8, pp. 66–73, Aug. 2000, doi: 10.1145/345124.345145.
 - [149] H. Li *et al.*, 'LLMs-as-Judges: A Comprehensive Survey on LLM-based Evaluation Methods', Dec. 10, 2024, arXiv:2412.05579. doi: 10.48550/arXiv.2412.05579.
 - [150] 'Artificial Intelligence Risk Management Framework (AI RMF 1.0)', National Institute of Standards and Technology (U.S.), Gaithersburg, MD, NIST AI 100-1, Jan. 2023. doi: 10.6028/NIST.AI.100-1.
 - [151] L. Fuentes-Fernandez and A. Vallecillo-Moreno, 'An Introduction to UML Profiles', *UML and Model Engineering*, vol. V, no. 2, pp. 6–13, 2004.
 - [152] I. Jacobson, G. Booch, and J. Rumbaugh, *The unified software development process*. Reading, Massachusetts, USA: Addison Wesley Longman, 1999.
 - [153] Q. Pan *et al.*, 'Human-Centered Design Recommendations for LLM-as-a-Judge', Jul. 03, 2024, arXiv:2407.03479. doi: 10.48550/arXiv.2407.03479.
 - [154] Google, 'Angular Documentation', Angular. Accessed: Aug. 06, 2026. [Online]. Available: <https://angular.dev/>
 - [155] Meta Platforms, Inc, 'React Documentation', React. Accessed: Aug. 06, 2026. [Online]. Available: <https://react.dev/>
 - [156] E. You, 'Vue.js Documentation', Vue.js - The Progressive JavaScript Framework. Accessed: Aug. 06, 2026. [Online]. Available: <https://vuejs.org/>
 - [157] J. Nielsen and R. L. Mack, Eds, *Usability inspection methods*. New York: John Wiley & Sons, 1994.
 - [158] LangChain, Inc., 'LangChain Documentation'. Accessed: Jul. 26, 2025. [Online]. Available: <https://python.langchain.com/docs>
 - [159] LlamaIndex, 'LlamaIndex Documentation', LlamaIndex | AI Agents for Document OCR + Workflows. Accessed: Aug. 06, 2026. [Online]. Available: <https://www.llamaindex.ai/>

- [160] OpenJS Foundation and Express.js contributors, 'Express.js Documentation', Express.js · Node.js web application framework. Accessed: Aug. 06, 2026. [Online]. Available: <https://expressjs.com/>
- [161] B. Tural, Z. Örpek, and Z. Destan, 'Future AI Tools: LLM and Frameworks', in *Proceedings of the 2025 9th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Ankara, Türkiye: IEEE, Nov. 2025, pp. 1–6. doi: 10.1109/ISMSIT67332.2025.11267968.
- [162] R. Sapkota, R. Shrestha, M. Rijal, and M. Karkee, 'LangChain vs. LangGraph vs. LangSmith: Taxonomies of Agentic AI Toolchains for End-to-End Orchestration', Sep. 04, 2025, *TechRxiv*. doi: 10.36227/techrxiv.175695645.52670060/v1.
- [163] OpenAI, 'Hello GPT-4o'. Accessed: Aug. 06, 2026. [Online]. Available: <https://openai.com/index/hello-gpt-4o/>
- [164] S. Schulhoff *et al.*, 'The Prompt Report: A Systematic Survey of Prompt Engineering Techniques', 2025, arXiv:2406.06608. doi: 10.48550/arXiv.2406.06608.
- [165] T. Feng, S. Liu, and D. Ghosal, 'CourseAssist: Pedagogically Appropriate AI Tutor for Computer Science Education', Jul. 29, 2024, arXiv:2407.10246. doi: 10.48550/arXiv.2407.10246.
- [166] Google, 'Firebase Firestore Documentation'. Accessed: Aug. 06, 2026. [Online]. Available: <https://firebase.google.com/docs/firestore>
- [167] T. Taipalus, 'Vector database management systems: Fundamental concepts, use-cases, and current challenges', *Cogn. Syst. Res.*, vol. 85, Jun. 2024, doi: 10.1016/j.cogsys.2024.101216.
- [168] L. Ma *et al.*, 'A Comprehensive Survey on Vector Database: Storage and Retrieval Technique, Challenge', Mar. 24, 2026, arXiv:2310.11703. doi: 10.48550/arXiv.2310.11703.
- [169] Y. Gao *et al.*, 'Retrieval-Augmented Generation for Large Language Models: A Survey', Mar. 27, 2024, arXiv:2312.10997. doi: 10.48550/arXiv.2312.10997.
- [170] A. Tophel, L. Chen, U. Hettiyadura, and J. Kodikara, 'Towards an AI Tutor for Undergraduate Geotechnical Engineering: A Comparative Study of Evaluating the Efficiency of Large Language Model Application Programming Interfaces', Jul. 25, 2024. doi: 10.21203/rs.3.rs-4658661/v1.
- [171] V. Nikolovski, D. Trajanov, and I. Chorbev, 'Advancing AI in Higher Education: A Comparative Study of Large Language Model-Based Agents for Exam Question Generation, Improvement, and Evaluation', *Algorithms*, vol. 18, no. 3, Mar. 2025, doi: 10.3390/a18030144.
- [172] D. Hennekeuser, D. D. Vaziri, D. Golchinfar, D. Schreiber, and G. Stevens, 'Enlarged Education – Exploring the Use of Generative AI to Support Lecturing in Higher Education', *Int. J. Artif. Intell. Educ.*, vol. 35, no. 3, pp. 1096–1128, Sep. 2025, doi: 10.1007/s40593-024-00424-y.
- [173] S. Gopi, D. Sreekanth, and N. Dehbozorgi, 'Enhancing Engineering Education Through LLM-Driven Adaptive Quiz Generation: A RAG-Based Approach', in *Proceedings of the 2024 IEEE Frontiers in Education Conference (FIE)*, Washington, DC, USA: IEEE, Oct. 2024, pp. 1–8. doi: 10.1109/FIE61694.2024.10893146.
- [174] H. Wei, J. Qiu, H. Yu, and W. Yuan, 'MEDCO: Medical Education Copilots Based on A Multi-Agent Framework', Aug. 22, 2024, arXiv:2408.12496. doi: 10.48550/arXiv.2408.12496.
- [175] M. Dean, R. R. Bond, M. F. McTear, and M. D. Mulvenna, 'ChatPapers: An AI Chatbot for Interacting with Academic Research', in *Proceedings of the 31st Irish*

- Conference on Artificial Intelligence and Cognitive Science (AICS)*, Letterkenny, Ireland: IEEE, Dec. 2023, pp. 1–7. doi: 10.1109/AICS60730.2023.10470521.
- [176] Supabase Inc, 'Supabase Documentation', Supabase | The Postgres Development Platform. Accessed: Aug. 06, 2026. [Online]. Available: <https://supabase.com/>
- [177] Microsoft, 'Relevance scoring in hybrid search using Reciprocal Rank Fusion (RRF)'. Accessed: Aug. 06, 2026. [Online]. Available: <https://learn.microsoft.com/en-us/azure/search/hybrid-search-ranking>
- [178] MongoDB, Inc., 'Better RAG Results With Reciprocal Rank Fusion (RRF) and Hybrid Search'. Accessed: Aug. 06, 2026. [Online]. Available: <https://www.mongodb.com/resources/basics/reciprocal-rank-fusion>
- [179] OpenAI, 'GPT-4o mini: advancing cost-efficient intelligence'. Accessed: Aug. 06, 2026. [Online]. Available: <https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/>
- [180] J. Stanojević, I. Milenković, M. Minović, and M. Maričić, 'Educational Platform in Higher Education With LLM-Driven Chatbot for Computer Networks and Telecommunications Course', *Comput. Appl. Eng. Educ.*, vol. 34, no. 1, Jan. 2026, doi: 10.1002/cae.70105.
- [181] University of Belgrade, 'Izveštaj Univerzitetske komisije za upis studenata u prvu godinu studijskih programa osnovnih akademskih, integrisanih akademskih, osnovnih strukovnih i specijalističkih strukovnih studija u školskoj 2025/2026. godini', 2025.
- [182] Q. Liu, J. Huang, L. Wu, K. Zhu, and S. Ba, 'CBET: design and evaluation of a domain-specific chatbot for mobile learning', *Univers. Access Inf. Soc.*, vol. 19, no. 3, pp. 655–673, Aug. 2020, doi: 10.1007/s10209-019-00666-x.
- [183] V. Venkatesh, J. Y. L. Thong, and X. Xu, 'Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology', *MIS Quarterly*, vol. 36, no. 1, pp. 157–178, Mar. 2012, doi: 10.2307/41410412.
- [184] H. Guo and T. Erdenebold, 'Factors influencing intention to adopt an AI chatbot for learning in higher education: An integrated PLS-SEM, IPMA, and ANN approach', *Computers and Education: Artificial Intelligence*, vol. 9, Dec. 2025, doi: 10.1016/j.caeai.2025.100477.
- [185] N. I. Mohd Rahim, N. A. Iahad, A. F. Yusof, and M. A. Al-Sharafi, 'AI-Based Chatbots Adoption Model for Higher-Education Institutions: A Hybrid PLS-SEM-Neural Network Modelling Approach', *Sustainability*, vol. 14, no. 19, Oct. 2022, doi: 10.3390/su141912726.
- [186] S. Borsci *et al.*, 'The Chatbot Usability Scale: the Design and Pilot of a Usability Scale for Interaction with AI-Based Conversational Agents', *Pers. Ubiquitous Comput.*, vol. 26, no. 1, pp. 95–119, Feb. 2022, doi: 10.1007/s00779-021-01582-9.
- [187] R. B. Kline and T. D. Little, *Principles and practice of structural equation modeling*, 5th ed. in *Methodology in the social sciences*. New York, NY, USA: Guilford Press, 2023.
- [188] M. Tavakol and R. Dennick, 'Making sense of Cronbach's alpha', *Int. J. Med. Educ.*, vol. 2, pp. 53–55, Jun. 2011, doi: 10.5116/ijme.4dfb.8dfd.
- [189] Fisher, R.A., 'Statistical methods for research workers', 1934.
- [190] T. K. Koo and M. Y. Li, 'A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research', *J. Chiropr. Med.*, vol. 15, no. 2, pp. 155–163, Jun. 2016, doi: 10.1016/j.jcm.2016.02.012.

- [191] G. Revalde, M. Zholdakhmet, A. Abola, and A. Murzagaliyeva, 'Can ChatGPT Pass a Physics Test?', *Technol. Knowl. Learn.*, vol. 30, no. 4, pp. 2459–2478, Jan. 2025, doi: 10.1007/s10758-025-09814-0.
- [192] I. Joshi *et al.*, 'ChatGPT in the Classroom: An Analysis of Its Strengths and Weaknesses for Solving Undergraduate Computer Science Questions', in *Proceedings of the 55th ACM Technical Symposium on Computer Science Education*, Portland; OR; USA: ACM, Mar. 2024, pp. 625–631. doi: 10.1145/3626252.3630803.
- [193] W. R. Shadish, T. D. Cook, and D. T. Campbell, *Experimental and quasi-experimental designs for generalized causal inference*, 2nd ed. Belmont, CA: Houghton, Mifflin, 2002.
- [194] W. Holmes, M. Bialik, and C. Fadel, *Artificial Intelligence in Education. Promise and Implications for Teaching and Learning*. Boston, MA, USA: Center for Curriculum Redesign, 2019.
- [195] M. D. L. Roca, M. M. Chan, A. Garcia-Cabot, E. Garcia-Lopez, and H. Amado-Salvatierra, 'The impact of a chatbot working as an assistant in a course for supporting student learning and engagement', *Comput. Appl. Eng. Educ.*, vol. 32, no. 5, Sep. 2024, doi: 10.1002/cae.22750.
- [196] C. Papakostas, C. Troussas, A. Krouska, and C. Sgouropoulou, 'A Rule-Based Chatbot Offering Personalized Guidance in Computer Programming Education', in *Generative Intelligence and Intelligent Tutoring Systems*, A. Saliferas and F. Lin., A. Sifaleras and F. Lin, Eds, in *Lecture Notes in Computer Science.*, Cham, Switzerland: Springer, 2024, pp. 253–264. doi: 10.1007/978-3-031-63031-6_22.
- [197] M. Liu and H. Reinders, 'Do AI chatbots impact motivation? Insights from a preliminary longitudinal study', *System*, vol. 128, Feb. 2025, doi: 10.1016/j.system.2024.103544.
- [198] J. Yin, T.-T. Goh, and Y. Hu, 'Using a Chatbot to Provide Formative Feedback: A Longitudinal Study of Intrinsic Motivation, Cognitive Load, and Learning Performance', *IEEE Trans. Learn. Technol.*, vol. 17, pp. 1378–1389, 2024, doi: 10.1109/TLT.2024.3364015.
- [199] *ISO 9241-11:1998: Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability*, 1998.
- [200] J. Nielsen, 'Usability 101: Introduction to Usability'. Accessed: Dec. 19, 2025. [Online]. Available: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- [201] Z. Unal and A. Unal, 'Evaluating and Comparing the Usability of Web-based Course Management Systems', *J. Inf. Technol. Educ. Res.*, vol. 10, pp. 19–38, 2011, doi: 10.28945/1358.
- [202] R. Garrote, 'Barriers to a wider Implementation of LMS in Higher Education: a Swedish case study, 2006-2011', *eleed*, no. 9, 2013.
- [203] E. Dahlstrom, D. C. Brooks, and J. Bichsel, 'The Current Ecosystem of Learning Management Systems in Higher Education: Student, Faculty, and IT Perspectives', EDUCAUSE Center for Analysis and Research, Louisville, CO, USA, 2014.
- [204] J. Brooke, 'SUS - A quick and dirty usability scale', in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, I. L. McClelland (Eds.), London: Taylor & Francis, 1996.
- [205] T. S. Tullis and W. Albert, *Measuring the user experience: collecting, analyzing, and presenting usability metrics*, 2nd ed. Amsterdam; Boston: Elsevier/Morgan Kaufmann, 2013.

- [206] P. Vlachogianni, N. Tselios, and M. Xenos, 'Perceived usability of learning management systems and university websites: A systematic review', in *Proc. IADIS Int. Conf. Interfaces Human Comput. Interact. and Game Entertainment Technol.*, Jul. 2020, pp. 85–92. Accessed: Apr. 20, 2026. [Online]. Available: <http://www.iadisportal.org/digital-library/perceived-usability-of-learning-management-systems-and-university-websites-a-systematic-review>
- [207] K. Orfanou, N. Tselios, and C. Katsanos, 'Perceived usability evaluation of learning management systems: Empirical evaluation of the System Usability Scale', *The International Review of Research in Open and Distributed Learning*, vol. 16, no. 2, pp. 227–246, Apr. 2015, doi: 10.19173/irrodl.v16i2.1955.
- [208] J.-H. Wu, R. D. Tennyson, and T.-L. Hsia, 'A study of student satisfaction in a blended e-learning system environment', *Comput. Educ.*, vol. 55, no. 1, pp. 155–164, Aug. 2010, doi: 10.1016/j.compedu.2009.12.012.
- [209] M. El-Masri and A. Tarhini, 'Factors affecting the adoption of e-learning systems in Qatar and USA: Extending the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2)', *Educ. Technol. Res. Dev.*, vol. 65, no. 3, pp. 743–763, Jun. 2017, doi: 10.1007/s11423-016-9508-8.
- [210] J. F. Kurose and K. W. Ross, *Computer networking: a top-down approach*, 7th ed. Hoboken, NJ, USA: Pearson, 2017.

Biografija autora

Jelica Stanojević je rođena 10.10.1996. godine u Leskovcu. Završila je osnovnu školu „Siniša Janić“ i gimnaziju „Stevan Jakovljević“, prirodno-matematički smer, u Vlasotincu i stekla Vukovu diplomu kao odličan učenik 2015. godine. Iste godine upisuje Fakultet organizacionih nauka Univerziteta u Beogradu - smer Informacioni sistemi i tehnologije, sa maksimalnim brojem poena na prijemnom ispitu. Diplomirala je 10.07.2019. godine na temi „Primena savremenih razvojnih okvira u implementaciji informacionog sistema apoteke“ sa ocenom 10 pod mentorstvom profesora Nenada Aničića. Položila je sve ispite sa prosečnom ocenom 9,84. Master akademske studije na Fakultetu organizacionih nauka, studijski program Informacioni sistemi i tehnologije, modul Informacione tehnologije, upisuje 2019. godine i diplomira 25.06.2021. godine na temi „GraphQL API Gateway u mikroservisnoj arhitekturi“ pod mentorstvom profesora Miroslava Minovića sa ocenom 10. Položila je sve ispite sa prosečnom ocenom 10,00. Jelica Stanojević iste godine (2021.) upisuje doktorske akademske studije na Fakultetu organizacionih nauka na studijskom programu Informacioni sistemi i kvantitativni menadžment, modul Informacione tehnologije. Trenutni prosek Jelice Stanojević na doktorskim studijama je 10,00.

Za vreme studija, nagrađivana je svake godine od strane fakulteta i opštine Vlasotince nagradama za najuspešnije studente svoje generacije. Dobitnica je Stipendije za izuzetno nadarene studente tri godine zaredom (2017/2018, 2018/2019, 2019/2020), kao i Dositejeve stipendije Fonda za mlade talente dve godine (2018/2019 i 2019/2020). Takođe je dobila nagradu Fondacije „Prof. dr Branislav Lazarević“ za najboljeg diplomiranog studenta odseka Informacioni sistemi i tehnologije u školskoj 2018/2019 godini.

Jelica Stanojević je u zimskom semsetru školske 2018/2019. godine aktivno učestvovala u izvođenju nastave kao demonstrator na predmetu treće godine akademskih studija u okviru katedre za Informacione tehnologije. Tokom 2019. godine je bila na praksi u firmi ProudSourceIT gde je radila na više različitih projekata razvijajući veb aplikacije. Stručnu praksu na osnovnim i master akademskim studijama obavila je na Fakultetu organizacionih nauka na Katedri za informacione tehnologije. Nakon upisivanja master studija angažovana je kao saradnik u nastavi na Fakultetu organizacionih nauka na Katedri za informacione tehnologije od 2020. do 2022. godine, a nakon upisivanja doktorskih akademskih studija prvi put je izabrana za asistenta na Katedri za informacione tehnologije od 15.02.2022. godine.

Objavila je više radova iz oblasti informacionih tehnologija u zemlji i inostranstvu. I bila učesnica brojnih domaćih i stranih konferencija (Symorg 2020, Infotech 2021, SPIN 2021, Symorg 2022, CECIS 2022, Infotech 2023, TEEM 2023, SPIN 2023, Infotech 2024, Symorg 2024, IACETL 2025, ISD 2025, SYMOPIS 2025, SPIN 2025, Symorg 2026).

Bila je član tima međunarodnog projekta ISSSES (*Information Security Services Education in Serbia*). Polaznica je ExpertsToMentors trening programa za nastavnike i saradnike FON-a, sa ciljem unapređenja mentorskih veština za razvoj startapa, nakon čega je postala mentor u okviru *Route2Launch* startup mentorskog programa. Trenutno je aktivni član projekta odobrenog od strane inicijative „FON Ideje“.

Bibliografija autora:

Završni (diplomski) rad:

Stanojević, J. (2019). Primena savremenih razvojnih okvira u implementaciji informacionog sistema apoteke. Fakultet organizacionih nauka, Beograd. (Mentor: Prof. dr Nenad Aničić)

Završni (master) rad:

Stanojević, J. (2021). GraphQL API Gateway u mikroservisnoj arhitekturi. Fakultet organizacionih nauka, Beograd. (Mentor: Prof. dr Miroslav Minović)

Pristupni rad (na doktorskim studijama):

Stanojević, J. (2025). Mogućnosti primene velikih jezičkih modela u obrazovanju. Fakultet organizacionih nauka, Beograd. (Mentor: Prof. dr Miroslav Minović)

Monografije međunarodnog značaja (M10):

Stanojević, J., Minović M., & Šošević U. (2024). Possible applications of ChatGPT in higher education. In: Gonçalves, J.A.d.C., Lima, J.L.S.d.M., Coelho, J.P., García-Peñalvo, F.J., García-Holgado, A. (eds) *Proceedings of TEEM 2023. The Eleventh International Conference on Technological Ecosystems for Enhancing Multiculturality. TEEM 2023. Lecture Notes in Educational Technology. 1512*, 135-144. Springer Singapore. ISBN: 978-981-97-1813-9. https://doi.org/10.1007/978-981-97-1814-6_13 **M14**

Radovi objavljeni u naučnim časopisima međunarodnog značaja (M20):

Stanojević, J., Milenković, I., Minović, M., & Maričić, M. (2025). Educational Platform in Higher Education With LLM-Driven Chatbot for Computer Networks and Telecommunications Course. *Computer Applications in Engineering Education*, 34(1), 41522. <https://doi.org/10.1002/cae.70105>. ISSN (Print): 1061-3773, ISSN (Online): 1099-0542. IF5=2,7. **M21**

Stanojević, J., Minović, M., Kovačević, I., & Maričić, M. (2025). Evaluating the Usage of ChatGPT for Educational Purposes Through the Extended Technological Acceptance Model. *International Journal of Engineering Education*, 41(2), 352-364. ISSN: 0949-149X. IF5=0,8. **M22**

Zbornici međunarodnih naučnih skupova (M30):

Milenković I., **Stanojević J.,** & Simić D. (2020). Review and comparison of biometric databases for technology evaluation. In *Proceedings book of the XVII International Symposium "Business and Artificial Intelligence", SYMORG 2020 (pp. 199-205)*, University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia, ISBN: 978-86-7680-385-9 **M33**

Stanojević J., Šošević U., & Štavljanin V. (2022). Overview of non-fungible token systems. In *Proceedings book of abstracts of the XVIII International Symposium "Sustainable Business Management and Digital Transformation: Challenges and Opportunities in the*

Post-COVID Era” SYMORG 2022 (pp. 29-30), University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia, ISBN: 978-86-7680-411-5 M34

Stanojević J., Šošević U., Minović M., & Milovanović M. (2022). An overview of modern cross-platform mobile development frameworks. In *Proceedings of the 33rd International Scientific Conference “Central European Conference on Information and Intelligent Systems”, CECIIS 2022 (pp. 489-497), University of Zagreb, Faculty of Organization and Informatics, Dubrovnik, Croatia, ISSN: 1847-2001, e-ISSN: 1848-2295 M33*

Stanojević J., Milovanović M., & Štavljanin V. (2024). Possible directions of integration of large language models into educational platforms. In *Proceedings book of the XIX International Symposium “Unlocking the Hidden Potentials of Organization Through Merging of Humans and Digitals”, SYMORG 2024 (pp. 111-115), University of Belgrade, Faculty of Organizational Sciences, Zlatibor, Serbia, ISBN: 978-86-7680-464-1 M33*

Stanojević, J., Maričić, M., Jovanović, I., & Slović, D. (2025). Exploring gender disparities in perceived ChatGPT usefulness. In *Proceedings book of abstracts of the 8th International Academic Conference on Education, Teaching and Learning*. Helsinki, Finland, ISBN: 978-609-485-636-5. **M34**

Stamenković, F., **Stanojević, J.**, & Minović, M. (2025). A systematic review of vector database use in retrieval-augmented generation for LLM-based educational platforms. In *Proceedings of the 52nd Symposium on Operational Research, SYM-OP-IS 2025 (pp. 262–267)*. University of Belgrade, Faculty of Organizational Sciences. <https://doi.org/10.5281/zenodo.17532060>. **M33**

Stamenković, F., **Stanojević J.**, & Simić D. (2025). Comparing Code Generation Capabilities of ChatGPT-4o and DeepSeek V3 in Solving TypeScript Programming Problems. In *Proceedings of the 33rd International Conference on Information Systems Development “Empowering the Interdisciplinary Role of ISD in Addressing Contemporary Issues in Digital Transformation: How Data Science and Generative AI Contributes to ISD”, ISD 2025, University of Gdańsk, Department of Business Informatics & University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia, ISBN: 978-83-972632-1-5. <https://doi.org/10.62036/ISD.2025.21>. M33*

Monografije nacionalnog značaja (M40):

Stanojević J., Šošević U., & Štavljanin V. (2023). Pregled i mogućnosti primene *non-fungible* token sistema. U Stošić B., Lazarević S., Kostić-Stanković M. (Ur.), *Održivo upravljanje poslovanjem: Inovacije, softver i komunikacije*, Fakultet organizacionih nauka. ISBN: 978-86-7680-423-8 **M45**

Časopisi nacionalnog značaja (M50)

Stanojević J., Minović M., & Simić D. (2022). GraphQL pregled i primena tehnologije u mikroservisnoj arhitekturi. *Info M Časopis za informacione tehnologije i multimedijalne sisteme*, 75-76, 49-54. ISSN: 1451-4397 **M53**

Predavanja po pozivu na skupovima nacionalnog značaja (M60)

Stanojević J., Minović M., & Simić D. (2021). Primena GraphQL-a u mikroservisnoj arhitekturi. U *Zborniku radova XIII Skupa privrednika i naučnika „Industrija 4.0. - mogućnosti, izazovi i rešenja za digitalnu transformaciju privrede”*, Univerzitet u Beogradu, Fakultet organizacionih nauka, Beograd, Srbija, ISBN 978-86-7680-396-5 **M63**

Stanojević J., Minović M., Milovanović M., & Simić D. (2023). ChatGPT kao studentski pomoćnik u programiranju. U *Zborniku radova XIV Skupa privrednika i naučnika „Digitalni i zeleni razvoj privrede”*, SPIN 2023 (str. 370-377), Univerzitet u Beogradu, Fakultet organizacionih nauka, Beograd, Srbija, ISBN: 978-86-7680-446-7. **M63**

Stamenković, F., **Stanojević, J.**, Jovanović, B., & Bogićević Sretenović, M. (2025). *Guardrail* mehanizmi u edukativnim platformama za programiranje zasnovanim na velikim jezičkim modelima. U *Zborniku radova XV Skupa privrednika i naučnika „Digitalna transformacija lanaca snabdevanja u industriji 4.0”*, SPIN 2025 (str. 217-223), Univerzitet u Beogradu, Fakultet organizacionih nauka, Beograd, Srbija. ISBN: 978-86-7680-499-3. <https://doi.org/10.5281/zenodo.18012049>. **M63**

Prilog 1.

Izjava o autorstvu

Ime i prezime autora Jelica Stanojević

Broj indeksa 5005/2021

Izjavljujem

da je doktorska disertacija pod naslovom

Metodološki pristup razvoju edukativne platforme zasnovane na velikom jezičkom modelu

- rezultat sopstvenog istraživačkog rada;
- da disertacija u celini ni u delovima nije bila predložena za sticanje druge diplome prema studijskim programima drugih visokoškolskih ustanova;
- da su rezultati korektno navedeni i
- da nisam kršio/la autorska prava i koristio/la intelektualnu svojinu drugih lica.

U Beogradu,

Potpis autora

Prilog 2.

Izjava o istovetnosti štampane i elektronske verzije doktorskog rada

Ime i prezime autora Jelica Stanojević

Broj indeksa 5005/2021

Studijski program Informacioni sistemi i kvantitativni menadžment

Naslov rada Metodološki pristup razvoju edukativne platforme zasnovane na velikom jezičkom modelu

Mentor Prof. Dr Miroslav Minović, redovni profesor FON-a

Izjavljujem da je štampana verzija mog doktorskog rada istovetna elektronskoj verziji koju sam predao/la radi pohranjena u **Digitalnom repozitorijumu Univerziteta u Beogradu**.

Dozvoljavam da se objave moji lični podaci vezani za dobijanje akademskog naziva doktora nauka, kao što su ime i prezime, godina i mesto rođenja i datum odbrane rada.

Ovi lični podaci mogu se objaviti na mrežnim stranicama digitalne biblioteke, u elektronskom katalogu i u publikacijama Univerziteta u Beogradu.

U Beogradu,

Potpis autora

Prilog 3.

Izjava o korišćenju

Ovlašćujem Univerzitetsku biblioteku „Svetozar Marković“ da u Digitalni repozitorijum Univerziteta u Beogradu unese moju doktorsku disertaciju pod naslovom:

Metodološki pristup razvoju edukativne platforme zasnovane na velikom jezičkom modelu

koja je moje autorsko delo.

Disertaciju sa svim prilogima predao/la sam u elektronskom formatu pogodnom za trajno arhiviranje.

Moju doktorsku disertaciju pohranjenu u Digitalnom repozitorijumu Univerziteta u Beogradu i dostupnu u otvorenom pristupu mogu da koriste svi koji poštuju odredbe sadržane u odabranom tipu licence Kreativne zajednice (Creative Commons) za koju sam se odlučio/la.

1. Autorstvo (CC BY)
2. Autorstvo – nekomercijalno (CC BY-NC)
3. Autorstvo – nekomercijalno – bez prerada (CC BY-NC-ND)
- ④ 4. Autorstvo – nekomercijalno – deliti pod istim uslovima (CC BY-NC-SA)
5. Autorstvo – bez prerada (CC BY-ND)
6. Autorstvo – deliti pod istim uslovima (CC BY-SA)

(Molimo da zaokružite samo jednu od šest ponuđenih licenci.
Kratak opis licenci je sastavni deo ove izjave).

U Beogradu,

Potpis autora

1. **Autorstvo.** Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence, čak i u komercijalne svrhe. Ovo je najslobodnija od svih licenci.

2. **Autorstvo – nekomercijalno.** Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca ne dozvoljava komercijalnu upotrebu dela.

3. **Autorstvo – nekomercijalno – bez prerada.** Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, bez promena, preoblikovanja ili upotrebe dela u svom delu, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca ne dozvoljava komercijalnu upotrebu dela. U odnosu na sve ostale licence, ovom licencom se ograničava najveći obim prava korišćenja dela.

4. **Autorstvo – nekomercijalno – deliti pod istim uslovima.** Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence i ako se prerada distribuira pod istom ili sličnom licencom. Ova licenca ne dozvoljava komercijalnu upotrebu dela i prerada.

5. **Autorstvo – bez prerada.** Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, bez promena, preoblikovanja ili upotrebe dela u svom delu, ako se navede ime autora na način određen od strane autora ili davaoca licence. Ova licenca dozvoljava komercijalnu upotrebu dela.

6. **Autorstvo – deliti pod istim uslovima.** Dozvoljavate umnožavanje, distribuciju i javno saopštavanje dela, i prerade, ako se navede ime autora na način određen od strane autora ili davaoca licence i ako se prerada distribuira pod istom ili sličnom licencom. Ova licenca dozvoljava komercijalnu upotrebu dela i prerada. Slična je softverskim licencama, odnosno licencama otvorenog koda