

# Primjena agenata za razgovor u obrazovanju

---

Valković, Anabela

Undergraduate thesis / Završni rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:131:945796>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-10-08**



Sveučilište u Zagrebu  
Filozofski fakultet  
University of Zagreb  
Faculty of Humanities  
and Social Sciences

Repository / Repozitorij:

[ODRAZ - open repository of the University of Zagreb  
Faculty of Humanities and Social Sciences](#)



SVEUČILIŠTE U ZAGREBU  
FILOZOFSKI FAKULTET  
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE ZNANOSTI  
Ak. god. 2023./2024.

Anabela Valković

## **Primjena agenata za razgovor u obrazovanju**

Završni rad

Mentor: prof. dr. sc. Sanja Seljan

Zagreb, rujan 2024.

## Izjava o akademskoj čestitosti

Izjavljujem i svojim potpisom potvrđujem da je ovaj rad rezultat mog vlastitog rada koji se temelji na istraživanjima te objavljenoj i citiranoj literaturi. Izjavljujem da nijedan dio rada nije napisan na nedozvoljen način, odnosno da je prepisan iz necitiranog rada, te da nijedan dio rada ne krši bilo čija autorska prava. Također izjavljujem da nijedan dio rada nije korišten za bilo koji drugi rad u bilo kojoj drugoj visokoškolskoj, znanstvenoj ili obrazovnoj ustanovi.

Anabela Valkević

---

(potpis)

# Sadržaj

Sadržaj.....	ii
1. Uvod.....	1
2. Definicija agenata za razgovor .....	2
3. Umjetna inteligencija .....	3
3.1. Djelovanje poput ljudi.....	4
3.2. Mišljenje poput ljudi .....	4
3.3. Racionalno mišljenje.....	6
3.4. Racionalno djelovanje.....	7
4. Povijest agenata za razgovor.....	9
5. Obrada prirodnog jezika .....	13
6. Vrste agenata za razgovor temeljeno na njihovoj arhitekturi .....	15
6.1. Agenti za razgovor temeljeni na pravilima ( <i>Rule-based chatbot</i> ).....	15
6.2. Agenti za razgovor na osnovi umjetne inteligencije ( <i>AI chatbots</i> ).....	15
6.3. Hibridni agenti za razgovor ( <i>hybrid chatbots</i> ).....	16
7. Primjena agenata za razgovor u različitim domenama .....	17
7.1. Obrazovanje .....	17
7.2. Korisnička podrška .....	17
7.3. Zdravstvo .....	18
7.4. Marketing.....	18
7.5. Turizam .....	19
8. Kriteriji za evaluaciju agenata za razgovor.....	20
8.1. Interaktivnost .....	20
8.2. Točnost.....	21
8.3. Brzina odgovora.....	22
8.4. Fleksibilnost.....	22

8.5.	Prirodnost jezika .....	23
9.	Testiranje agenta za razgovor .....	24
9.1.	Pitanje – odgovor test (engl. <i>question – answering test</i> ) .....	24
9.1.1.	Povijest pitanje-odgovor sustava .....	24
9.1.2.	Evaluacija prema QA sustavu .....	26
9.1.3.	Mjere za evaluaciju .....	27
9.2.	Matrica konfuzije (engl. <i>Confusion matrix</i> ).....	27
9.3.	Korisničko iskustvo .....	29
10.	Istraživanje .....	30
10.1.	Razvojno okruženje .....	31
10.2.	Sučelje agenta za razgovor.....	32
11.	Anketa .....	34
11.1.	Struktura ankete .....	34
11.2.	Analiza ankete.....	34
11.3.	Diskusija .....	41
12.	Evaluacija agenta za razgovor .....	42
13.	Usporedba agenta za razgovor s već postojećim agentima za razgovor na tržištu .....	45
13.1.	Funkcionalnost.....	45
13.2.	Korisničko iskustvo .....	45
13.3.	Prilagodljivost.....	46
13.4.	Dostupnost sadržaja .....	47
13.5.	Tehnološka osnova.....	47
13.6.	Usporedba agenata .....	47
14.	Zaključak.....	49
15.	Literatura.....	51
16.	Popis slika .....	56
17.	Popis tablica .....	57

18. Prilozi.....	58
Sažetak .....	59
Summary.....	60

# 1. Uvod

Početak digitalnog doba doveo je do ubrzanog razvoja tehnologije, koja značajno mijenja načine na koje komuniciramo i živimo. Jedan od ključnih alata digitalnog doba, kao i jedan od značajnijih napredaka u području umjetne inteligencije, jest razvoj i primjena agenata za razgovor (engl. *chatbot*). Cilj ovog završnog rada je analizirati postojeće agente za razgovor i istražiti njihovu primjenu u obrazovanju.

U teorijskom dijelu rada bit će obrađen povijesni razvoj agenata za razgovor (prethodno spomenuti *chatbot*), prikazane postojeće primjene u suvremenom društvu u različitim domenama te razrađeni kriteriji za njihovu evaluaciju. Posebna pažnja bit će posvećena analiziranju kako agenti za razgovor olakšavaju interakciju između korisnika i digitalnih sustava, pružajući odgovore vrlo brzo nakon upita te podršku u raznim industrijama, posebice na primjeru domene obrazovanja.

U praktičnom dijelu završnog rada provest će se izgradnja agenta za razgovor specijaliziranog za domenu obrazovanja, s ciljem pretraživanja informacija. Nakon izgradnje, izvršit će se evaluacija razvijenog agenta u usporedbi s postojećim alatima, a potom će uslijediti analiza rezultata i zaključak. Osim tehničke izgradnje, istražiti će se i način na koji takvi agenti mogu poboljšati korisničko iskustvo, povećati efikasnost u učenju i pružiti personaliziranu podršku studentima.

Ovaj rad pruža uvid u evoluciju i praktičnu primjenu agenata za razgovor, naglašavajući njihovu sve veću važnost u modernom obrazovnom okruženju. Kroz detaljnu analizu i praktične primjere, bit će demonstrirano kako agenti za razgovor mogu transformirati obrazovni proces, omogućujući pristupačnije i fleksibilnije učenje. Pored toga, razmotrit će se i etički aspekti korištenja umjetne inteligencije u obrazovanju, uključujući privatnost podataka i odgovornost za pružene informacije.

Sveukupno, ovaj rad ima za cilj doprinijeti razumijevanju potencijala i izazova implementacije agenata za razgovor u obrazovnom sektoru te ponuditi smjernice za njihovu učinkovitu primjenu.

## 2. Definicija agenata za razgovor

Agent za razgovor (engl. *chatbot*) je program umjetne inteligencije i model interakcije između ljudi i računala (engl. *Human-computer Interaction, HCI*) (Bansal i Khan, 2018). Prema Lexico Dictionaries (2024), agent za razgovor je softverski program dizajniran da odgovara na verbalne ili pisane poruke korisnika, naročito preko Interneta. Njihova primarna svrha je omogućiti interakciju između ljudi i računala na intuitivan način. Zbog svoje sposobnosti simulacije razgovora s ljudskim korisnicima, agenti za razgovor postali su neizostavan alat u raznim industrijama, uključujući zdravstvo, obrazovanje i druge sektore. Korisnici često doživljavaju agente za razgovor više kao prijatelje nego kao asistente za dobivanje informacija (Costa, 2018).

Agenti za razgovor imaju svoje prednosti, ali i mane koje dovode do raznih tehničkih, etičkih i drugih problema. Neki od tehničkih problema uključuju poteškoće kod integracije s postojećim sustavima, složenost obavljanja kompleksnih upita te preciznost prepoznavanja jezika (Contact Fusion, 2023). Osim toga, postoji zabrinutost za radna mjesta zbog automatizacije, što bi moglo značajno utjecati na tržište rada. Određeni stupanj subjektivnosti odgovora koje agent za razgovor daje korisniku također može izazvati probleme, kao što je pristranost prema spolu (Costa, 2018). Ta pristranost često oponaša tradicionalne ženske stereotipe koristeći agente kao osobne asistente i sekretarice (Costa, 2018).

Kako bi agent za razgovor funkcionirao, potrebne su osnovne komponente kao što su korisničko sučelje, mehanizmi za razumijevanje prirodnog jezika (engl. *Natural Language Understander*), baze podataka kojima će agent za razgovor pristupiti kako bi pružio relevantne odgovore, te komponenta generiranja prirodnog jezika (engl. *Natural Language Generator*) koja kreira odgovor u prirodnom jeziku koji će biti prezentiran korisniku (Su et al., 2018; Su et al., 2019).

Uspješno implementirani agenti za razgovor mogu značajno poboljšati korisničko iskustvo, povećati efikasnost u pružanju usluga i omogućiti personaliziranu podršku (Poptin, 2022). Međutim, razvoj i implementacija takvih sustava zahtijeva pažljivo razmatranje tehničkih, etičkih i društvenih aspekata kako bi se osigurala pravednost, pouzdanost i sigurnost korisničkih podataka. Razvoj i primjena agenata za razgovor u različitim industrijama nastavlja rasti, pružajući nove mogućnosti, ali i izazove koje treba adresirati kroz inovacije i odgovornu upotrebu tehnologije.



### 3. Umjetna inteligencija

Pojam umjetne inteligencije usko je vezan za pojam agenta za razgovor. Kako bi se pobliže opisao termin agent za razgovor, potrebno je razumjeti što je to umjetna inteligencija. Prema Pristeru (2019), umjetna inteligencija je "dio računalne znanosti (informatike) koji se bavi razvojem sposobnosti računala da obavljaju zadaće za koje je potreban neki oblik inteligencije, tj. da se mogu snalaziti u novim prilikama, učiti nove koncepte, donositi zaključke, razumjeti prirodni jezik, raspoznavati prizore i drugo".

Često se pojam umjetne inteligencije povezuje s računalnim sustavima. Iako laici najčešće inteligentne sustave poistovjećuju s robotima, koji mogu, ali ne moraju biti inteligentni, inteligentnim sustavima smatraju se svi sustavi koji koriste velike količine znanja, uče na temelju iskustva, pokazuju ponašanje koje se može prilagoditi, te komuniciraju s korisnikom, odnosno čovjekom, prirodnim jezikom (Poole i Mackworth, 2010).

Prema Russellu i Norvigu (2021), postoje četiri inačice umjetne inteligencije, koje će detaljnije biti objašnjene u kasnijim poglavljima:

1. Djelovanje poput ljudi:

Ova verzija umjetne inteligencije fokusira se na razvoj sustava koji oponašaju ljudske akcije i ponašanja. Cilj je stvoriti strojeve koji mogu izvesti zadatke na način sličan ljudima, uključujući prepoznavanje govora, igranje igara i drugo.

2. Misлити poput ljudi:

Ova verzija bavi se modeliranjem ljudskog mišljenja i kognitivnih procesa. Umjetna inteligencija u ovom kontekstu nastoji razumjeti i replicirati način na koji ljudi razmišljaju, zaključuju i donose odluke.

3. Misлити racionalno:

Ova verzija temelji se na logičkom razmišljanju i donošenju optimalnih odluka. Sustavi dizajnirani prema ovom principu koriste formalne metode logike i matematike kako bi riješili probleme na najracionalniji mogući način.

#### 4. Djelovati racionalno:

Ova verzija usmjerena je na razvoj sustava koji djeluju na najracionalniji način u svakom trenutku. Cilj je optimizirati performanse sustava kroz racionalno ponašanje koje maksimizira željene ishode.

Agenti za razgovor su primjer primjene umjetne inteligencije koja kombinira elemente ovih verzija. Oni djeluju poput ljudi i komuniciraju prirodnim jezikom, misle na temelju algoritama koji oponašaju ljudsko zaključivanje, te djeluju racionalno kako bi pružili najrelevantnije odgovore korisnicima (Russell i Norvig, 2021).

Njihova upotreba sve više raste u različitim domenama, što je rezultat sposobnosti umjetne inteligencije da se prilagođava i uči iz interakcija s korisnicima. Agenti za razgovor mogu pružiti personaliziranu podršku, obavljati rutinske zadatke, te omogućiti brzu i efikasnu komunikaciju između korisnika i digitalnih sustava.

Osim tehničkih aspekata, važno je razmotriti i etičke aspekte primjene agenata za razgovor. Automatizacija komunikacijskih procesa može izazvati zabrinutosti vezane uz privatnost podataka, sigurnost informacija i utjecaj na radna mjesta (Wei et al., 2024). Potrebno je osigurati da razvoj i primjena ovih tehnologija bude u skladu s etičkim normama i zakonskim regulativama kako bi se osigurala pravednost i zaštita korisnika (Coghlan, 2023).

Kombinacija tehničkog znanja i etičkog razmatranja ključna je za uspješan razvoj i implementaciju agenata za razgovor, čime se može osigurati njihova učinkovitost i sigurnost u različitim industrijama.

### **3.1. Djelovanje poput ljudi**

1950. godine Alan Turing predlaže Turingov test kao mjerilo sposobnosti stroja u oponašanju ljudske inteligencije. Turingov test temelji se na ideji da bi stroj mogao smatrati inteligentnim ako može komunicirati s ljudskim ispitivačem putem teksta na način da ispitivač ne može razlikovati stroj od čovjeka (Stanford Encyclopedia of Philosophy, 2021). O Turingovom testu pisat će nešto više u poglavlju „Povijest agenata za razgovor“.

### **3.2. Mišljenje poput ljudi**

Pretpostavlja se kako postoji mogućnost da sustav koji oponaša ljude zapravo "ne razmišlja". Ovaj problem istražuje kognitivna znanost, interdisciplinarno područje koje kombinira

računalnu znanost i psihologiju (Stanford Encyclopedia of Philosophy, 2023). Kognitivna znanost bavi se proučavanjem inteligentnih sustava, a njezin cilj je postavljanje teorije uma, koja bi omogućila dublje razumijevanje načina na koji ljudi razmišljaju i kako se ta razmišljanja mogu modelirati u računalnim sustavima (Valerjev, 2006).

Kognitivna znanost proučava različite aspekte ljudske spoznaje, uključujući percepciju, memoriju, jezik, učenje i rješavanje problema (Hrvatska enciklopedija, 2024). Proučavanje ovih aspekata može pomoći u razvoju sofisticiranijih i "inteligentnijih" umjetnih sustava. Teorija uma, kao koncept unutar kognitivne znanosti, odnosi se na sposobnost razumijevanja i predviđanja tuđih mentalnih stanja, što je ključni element u stvaranju sustava koji mogu oponašati ljudsko razmišljanje i ponašanje (Wellman, Cross i Watson, 2001). Postavljanjem teorije uma kao cilja, kognitivna znanost nastoji razviti modele koji mogu ne samo simulirati ljudsko ponašanje, već i razumjeti složenost ljudske misli i emocija. To uključuje izgradnju sustava koji mogu interpretirati kontekst, prepoznati emocionalne nijanse u jeziku i prilagoditi svoje odgovore na temelju prošlih interakcija.

Razvoj takvih sustava suočava se s mnogim izazovima. Jedan od ključnih problema je kako računalni sustavi mogu postići stvarno razumijevanje, a ne samo simulaciju razmišljanja (Fairouse Abedin, 2021). Ovo pitanje je poznato kao problem "dubokog razumijevanja" i ostaje jedna od najvećih prepreka u razvoju umjetne inteligencije koja može u potpunosti replicirati ljudsku spoznaju.

Osim tehničkih izazova, tu su i etički i filozofski problemi vezani uz razvoj inteligentnih sustava. Primjerice, ukoliko se stvore sustavi koji mogu oponašati ljudsko razmišljanje i ponašanje, postavlja se pitanje kako će se tretirati takve sustave te za što će sve moći biti upotrijebljeni. Nadalje, može se postaviti pitanje hoće li takvi sustavi imati određena "prava" ili moralne obveze prema njima te kako osigurati da se takvi sustavi koriste na etički prihvatljiv način, što je uveliko bitno u obrazovanju.

Unatoč ovim izazovima, napredak u kognitivnoj znanosti i umjetnoj inteligenciji otvara vrata mnogim inovativnim primjenama. Na primjer, agenti za razgovor mogu postati još učinkovitiji u pružanju personalizirane podrške, edukativnih usluga i zdravstvenih savjeta, što može značajno poboljšati kvalitetu života korisnika. Kroz kontinuirano istraživanje i razvoj, možemo očekivati daljnje unapređenje u sposobnosti umjetnih sustava da razumiju i komuniciraju s ljudima na način koji je sve bliži stvarnoj ljudskoj interakciji.

Povezanost umjetne inteligencije i kognitivne znanosti ključna je za budući razvoj inteligentnih sustava (Valerjev, 2006). Razumijevanje kako ljudi razmišljaju i djeluju omogućuje nam da gradimo sustave koji ne samo da oponašaju ljudsko ponašanje, već i pružaju stvarnu vrijednost kroz intuitivnu i prirodnu interakciju s korisnicima.

### **3.3. Racionalno mišljenje**

Kao najboljim rješenjem za racionalnost sustava pokazuju se logički sustavi jer racionalno rješavaju sve probleme. Logički sustavi temelje se na skupu pravila i prema njima djeluju. Iako su se u teoriji pokazali kao najbolji izbor, u praksi se nisu pokazali kao najbolje rješenje. Poteškoće se pojavljuju kod realnih situacija, kod kojih nisu sposobni riješiti problem, posebno ako moraju donositi odluke vezane uz nepotpune ili dvosmislene informacije (Valerjev, 2006).

Jedan od ključnih problema s logičkim sustavima jest njihova rigidnost. Budući da se temelje na fiksnim pravilima, teško se prilagođavaju promjenjivim uvjetima i neočekivanim situacijama. U stvarnom svijetu, mnogi problemi zahtijevaju fleksibilnost i sposobnost prilagodbe, što logički sustavi često ne mogu pružiti (Valerjev, 2006). Također, logički sustavi imaju poteškoća s donošenjem odluka u situacijama gdje informacije nisu potpune ili su kontradiktorne, što je česta pojava u stvarnim životnim scenarijima (Valerjev, 2006).

Osim toga, logički sustavi mogu biti skupi za razvoj i održavanje jer zahtijevaju precizno definiranje pravila i kontinuiranu prilagodbu kako bi ostali relevantni. Ovaj proces može biti vremenski intenzivan i skup, što dodatno ograničava njihovu praktičnu primjenu.

Umjesto toga, sustavi temeljeni na metodama strojnog učenja i umjetnih neuronskih mreža sve više dobivaju na popularnosti. Ovi sustavi mogu učiti iz podataka i prilagoditi se novim situacijama bez potrebe za eksplicitnim programiranjem svakog pravila. Na primjer, duboko učenje, kao podvrsta strojnog učenja, koristi složene neuronske mreže za prepoznavanje uzoraka u velikim skupovima podataka, što omogućuje bolje performanse u zadacima kao što su prepoznavanje govora, slike i prirodnog jezika. Strojno učenje omogućuje sustavima da donose odluke na temelju statističkih modela, što im pomaže da se nose s nepotpunim i dvosmislenim informacijama. Ovi sustavi mogu generalizirati iz prethodnih iskustava i prilagoditi se novim okolnostima, što ih čini pogodnijima za dinamična okruženja.

Unatoč prednostima strojnih metoda, važno je napomenuti da one također imaju svoje izazove. Jedan od glavnih problema je potreba za velikim količinama podataka za treniranje modela, što može biti teško dostupno u nekim domenama (Dunđer i sur., 2023). Također,

modeli strojnog učenja često funkcioniraju kao "crne kutije", gdje je teško razumjeti točno kako dolaze do određenih odluka, što može biti problematično u kritičnim aplikacijama koje zahtijevaju transparentnost.

Sveukupno, iako logički sustavi nude racionalnost i preciznost, njihova primjena u praksi je ograničena zbog rigidnosti i poteškoća u radu s nepotpunim informacijama. Moderni pristupi, poput strojnog učenja i dubokog učenja, nude veću fleksibilnost i prilagodljivost, ali također donose nove izazove koji zahtijevaju pažljivo razmatranje. Kombinacija različitih pristupa mogla bi ponuditi najbolje rezultate, koristeći snage svakog metoda kako bi se postigla optimalna rješenja za složene probleme u stvarnom svijetu.

### **3.4. Racionalno djelovanje**

Kod racionalnog djelovanja sustava dolazi do problema odluke oko izbora najracionalnijeg rješenja. Do problema dolazi zbog manjka vremena raspoloživog za donošenje odluke, što najčešće dovodi do izbora koji nije dovoljno racionalan. Ovaj problem često proizlazi iz ograničenja u kapacitetu obrade informacija i brzini kojom se mogu analizirati različite opcije. S obzirom da ovakav pristup nema previše ishoda, naziva se ograničena racionalnost (GeeksforGeeks, 2024).

Ograničena racionalnost je koncept koji je uveo Herbert Simon, ekonomist i psiholog, kako bi opisao situaciju u kojoj odluke donose racionalna bića unutar ograničenih informacija i vremena (FOI, 2024). Prema Simonu, umjesto da traže optimalno rješenje, ljudi (i sustavi) često traže "zadovoljavajuće" rješenje koje je dovoljno dobro s obzirom na trenutne okolnosti i ograničenja. Ovaj pristup se temelji na ideji da je često nemoguće ili nepraktično analizirati sve moguće opcije i njihove posljedice unutar ograničenog vremena i resursa.

U kontekstu računalnih sustava i umjetne inteligencije, ograničena racionalnost znači da sustavi moraju donositi odluke na temelju ograničenih podataka i unutar ograničenog vremena (Valerjev, 2006). Ovo može biti posebno izazovno u realnim situacijama gdje su informacije nepotpune ili dvosmislene. Kako bi se nosili s ovim izazovima, sustavi često koriste heuristike, što su pravila ili strategije koje pomažu pri donošenju odluka na temelju ograničenih informacija i resursa. Heuristike mogu omogućiti brže donošenje odluka, ali mogu dovesti i do pogrešaka ili pristranosti.

Razvoj sustava koji djeluju u skladu s ograničenom racionalnošću zahtijeva pažljivo dizajniranje algoritama koji mogu balansirati između brzine i preciznosti. U tom kontekstu,

strojno učenje i umjetne neuronske mreže pružaju moćne alate za prilagodbu i učenje iz podataka, omogućujući sustavima da poboljšaju svoje performanse kroz iterativni proces.

Unatoč napretku u tehnologiji, ograničena racionalnost i dalje predstavlja značajan izazov. Potreba za donošenjem brzih odluka u složenim i dinamičnim okruženjima znači da sustavi moraju biti sposobni procijeniti rizike i donositi odluke koje su dovoljno dobre, čak i ako nisu optimalne. Ovo zahtijeva ne samo tehničke sposobnosti, već i razumijevanje konteksta i ciljeva koji vode donošenju odluka.

U praktičnim primjenama, kao što su autonomna vozila, medicinski dijagnostički sustavi i financijske aplikacije, ograničena racionalnost mora biti pažljivo upravljana kako bi se osigurala sigurnost, pouzdanost i etičnost odluka koje sustavi donose. Kroz kombinaciju naprednih algoritama, učenja iz podataka i pažljivog dizajna sustava, moguće je razviti rješenja koja bolje odgovaraju izazovima stvarnog svijeta i pružaju korisne rezultate unatoč inherentnim ograničenjima.

## 4. Povijest agenata za razgovor

Sredinom 20. stoljeća s razvojem umjetne inteligencije i računalne tehnologije dolazi i do razvoja agenata za razgovor. Na samom početku agenti za razgovor predstavljali su jednostavne programe za jednostavne upite o jednoj temi ili u specifičnim domenama. Kasnije, agenti za razgovor razvijaju se u napredne sustave sa sposobnostima obrade prirodnog jezika. S vremenom su postali ključni alati u mnogim industrijama, uključujući zdravstvo, obrazovanje i korisničku podršku (Adamopoulou i Moussiades, 2020). Njihova sposobnost prilagodbe i učenja iz interakcija omogućuje sve sofisticiranije i intuitivnije komunikacije s korisnicima. Današnji agenti za razgovor mogu prepoznati i odgovoriti na složene upite, čineći ih neizostavnim dijelom modernih interakcijskih sustava (Adamopoulou i Moussiades, 2020).

Turingov test smatra se temeljem razvoja agenata za razgovor. Godine 1950. britanski matematičar i informatičar Alan Turing predlaže test koji bi mjerio sposobnost stroja u oponašanju ljudske inteligencije. Struna (2024) Turingov test navodi kao „istraživačku metodu u području umjetne inteligencije kojom se određuje mogu li računala misliti kao ljudska bića na temelju njihova uspjeha u određenim vrstama igara imitacije“. Test započinje tako da ispitivač prirodnim jezikom komunicira i s čovjekom i sa strojem, ali je od njih fizički odvojen. Da bi računalo prošlo test, sudac ne smije razaznati komunicira li s računalom ili čovjekom. U protivnom, računalo ne može proći test i smatra se da nije pokazalo primjenu umjetne inteligencije. Turingov test postao je ključna referentna točka za istraživanje i razvoj inteligentnih sustava (Stanford Encyclopedia of Philosophy, 2021).

Prvi agent za razgovor razvijen je 1966. godine pod imenom „Eliza“. Agent za razgovor Eliza razvio je Joseph Weizenbaum na Massachusetts Institute of Technology (MIT), a smatra se prvim agentom dostupnim za javnost (H.-Y. Shum, X. He i D. Li, 2018). Eliza predstavlja jednostavni agent za razgovor koji prihvaća samo tekstualni unos i komunicira s ljudima na temelju ručno pisanih skripti (Weizenbaum, 1966). Eliza svoje odgovore oblikuje prema unaprijed definiranim pravilima, zbog čega je opseg njenih odgovora limitiran i specifičan za jednu određenu domenu. Odgovore traži pomoću prepoznavanja poznatog uzorka (pattern matching) (H.-Y. Shum, X. He i D. Li, 2018). Iako ograničena u svojoj funkcionalnosti, Eliza je bila pionirski korak u razvoju agenata za razgovor i postavila temelje za daljnji razvoj u ovom području (Dovedan i sur., 2022). Njezina sposobnost simulacije osnovnog dijaloga otvorila je put za razvoj sofisticiranijih sustava koji su uspjeli bolje razumjeti i obraditi prirodni jezik.

Eliza je osmišljena da ljudima obnaša ulogu psihoterapeuta, konkretno kao simulacija Rogerijanske terapije, u kojoj korisnici mogu postavljati pitanja i primati odgovore o specifičnim osobnim situacijama (New Jersey Institute of Technology, 2024). Rogerijanska terapija pacijenta navodi na istraživanje vlastitih osjećaja i emocija, a rezultat je jasnija percepcija samog sebe (NIH, 2023). Terapeut kroz razna pitanja osvještava klijenta te na taj način dolazi do pojave samosvijesti kod pacijenta (NIH, 2023). Iako ovakav razgovor s računalom ne može zamijeniti pravu psihoterapiju, Eliza pruža korisnicima priliku da razgovaraju o teškim temama i situacijama, te im nudi osjećaj sigurnosti i anonimnosti. Eliza je omogućila korisnicima da izraze svoje misli i osjećaje u zaštićenom okruženju, što može biti korisno za introspekciju i emocionalnu podršku. Iako se agent za razgovor Eliza smatra velikim uspjehom u povijesti razvoja agenata za razgovor, nije uspjela proći Turingov test, budući da njezina sposobnost obrade i razumijevanja prirodnog jezika ostaje ograničena (Scispace, 2024).

Eliza je bila revolucionarna u svojoj eri, postavljajući temelje za daljnje istraživanje u području umjetne inteligencije i ljudsko-računalne interakcije. Iako je njen pristup temeljen na jednostavnom prepoznavanju uzoraka, inspirirala je razvoj naprednijih agenata za razgovor koji koriste sofisticiranije metode obrade prirodnog jezika. Uspjeh Elize također je istaknuo važnost prilagodbe i empatije u interakciji između čovjeka i računala. Moderni agenti za razgovor, koji su naslijedili naslijeđe Elize, sada koriste napredne tehnike poput dubokog učenja za bolje razumijevanje i generiranje prirodnog jezika. Eliza ostaje značajna figura u povijesti umjetne inteligencije, simbolizirajući rani pokušaj stvaranja sustava koji može imitirati ljudsku komunikaciju.



A typical conversation is the following:

Men are all alike.

IN WHAT WAY?

They're always bugging us about something or other.

CAN YOU THINK OF A SPECIFIC EXAMPLE?

Well, my boyfriend made me come here.

YOUR BOYFRIEND MADE YOU COME HERE

He says I'm depressed much of the time.

I AM SORRY TO HEAR YOU ARE DEPRESSED

Slika 1. Interakcija Elize i čovjeka (Weizenbaum, 1966).

Slika 1 prikazuje interakciju Elize i čovjeka odnosno korisnika. Ovaj primjer prikazuje kako se odvija razgovor i na koje načine Eliza odgovara korisniku. Iako Eliza daje odgovore isključivo prema unaprijed određenoj skripti, većina korisnika bila je zadovoljna odgovorima koji su im nalikovali onima kakve bi i čovjek dao. Zbog toga, mnogo korisnika osjetila je prijateljsku povezanost s Elizom (H.-Y. Shum, X. He i D. Li, 2018).

Nadalje, prvi agent za razgovor u povijesti koji je prošao Turingov test bio je Parry, razvijen 1972. godine od strane psihijatra Kennetha Colbyja sa Stanforda. Parry je osmišljen da simulira pacijenta s paranoidnom shizofrenijom, predstavljajući složeniji model od Elize (Colby, 1981). Iako i dalje koristi pravila za generiranje odgovora, Parry je napredniji u prepoznavanju i odgovaranju na emocionalne signale korisnika, što ga čini vjerodostojnijim u simulaciji ljudske psihopatologije (Adamopoulou i Moussiades, 2020). Na temelju analize korisnikovih odgovora i osjećaja, Parry može pružiti relevantnije i emocionalno usklađene odgovore. Ovaj napredak u razumijevanju i simulaciji emocija označio je značajan korak naprijed u razvoju agenata za razgovor i njihovih sposobnosti za interakciju s ljudima (Adamopoulou i Moussiades, 2020).

Zatim 1995. godine Richard Wallace razvija Alice, što je skraćenica za Umjetni jezični internetski računalni entitet (*Artificial Linguistic Internet Computer Entity*). Alice koristi Jezik označavanja umjetne inteligencije (*Artificial Intelligence Markup Language* - AIML) kao temelj za svoju strukturu, omogućujući sofisticiranije rukovanje jezičnim obrascima i strukturiranim odgovorima. Ovaj agent za razgovor je tri puta osvojio prestižnu Loebnerovu nagradu, i to 2000., 2001. i 2004. godine (H.-Y. Shum, X. He i D. Li, 2018). Iako je Alice postigla značajan uspjeh u prepoznavanju i generiranju prirodnog jezika, nije uspjela proći Turingov test zbog poteškoća u održavanju konzistentnog i prirodnog dijaloga u dužem vremenskom periodu (H.-Y. Shum, X. He i D. Li, 2018). Ova ograničenja ukazuju na izazove u razvoju agenta za razgovor koji mogu održavati kompleksne i koherentne razgovore s korisnicima.

U 21. stoljeću, s napredovanjem umjetne inteligencije, pojavili su se osobni glasovni asistenti (engl. *voice assistant*) koji se mogu integrirati u pametne telefone, kućne zvučnike i druge uređaje, omogućujući prepoznavanje i interpretaciju glasovnih naredbi (Adamopoulou i Moussiades, 2020). Ovi asistenti ne samo da odgovaraju na upite korisnika, već također upravljaju različitim aplikacijama i uređajima, uključujući kalendare, e-maileve i pametne kućanske uređaje. Glasovni asistenti poput *Google Assistant*, *IBM Watsona*, *Amazon Alexa*, *Microsoft Cortane* i *Apple Siri* postali su ključni alati u svakodnevnom životu, omogućujući korisnicima jednostavnu i intuitivnu interakciju s tehnologijom (Adamopoulou i Moussiades, 2020). Njihova sposobnost razumijevanja prirodnog jezika i prilagodbe korisničkim potrebama čini ih neizostavnim dijelom modernih digitalnih ekosustava. Ovi sustavi stalno se unapređuju, koristeći napredne algoritme za učenje i prilagodbu, kako bi bolje služili korisnicima i poboljšali učinkovitost i preciznost svojih usluga.

## 5. Obrada prirodnog jezika

Jedna od grana umjetne inteligencije koja je usko povezana s lingvistikom je obrada prirodnog jezika (engl. *Natural Language Processing* - NLP). Kako bi prirodni jezik postao „razumljiv“ računalu, neophodna je njegova obrada i pretvorba u strojno razumljiv format. Cilj NLP-a je omogućiti računalu da razumije i interpretira ljudski jezik na način koji omogućava efikasnu komunikaciju između ljudi i računala (Cahn, 2017). Glavni cilj NLP-a je razviti sposobnost računala da komunicira jezikom koji koriste ljudi, omogućujući tako prirodnije interakcije s tehnologijom (Cahn, 2017).

Strojno učenje (engl. *Machine Learning*) igra ključnu ulogu u NLP-u, jer omogućuje stvaranje sofisticiranih modela koji se temelje na uzlazno-izlaznim parovima primjera (Eisenstein, 2018). U kontekstu obrade prirodnog jezika, ovo uključuje zadatke poput prevođenja s jednog jezika na drugi (engl. *Translation*), gdje se niz riječi (engl. *tokens*) iz jednog jezika pretvara u odgovarajući niz riječi (prethodno spomenuti *token*) na drugom jeziku (Eisenstein, 2018). Ovaj proces omogućuje računalnim sustavima da bolje razumiju, generiraju i obrađuju ljudski jezik, čime se značajno poboljšava njihova funkcionalnost i korisničko iskustvo.

Jedna od ključnih karakteristika obrade prirodnog jezika (NLP) je njena široka primjena u različitim područjima. NLP se koristi za razvoj agenta za razgovor, ispravljanje i provjeru pravopisa, za strojno prevođenje (Dunđer i sur., 2023; Seljan i sur., 2023), za analizu teksta (Dunđer i sur., 2020, Seljan i sur., 2014), izradu sažetaka, te generiranje tekstualnih datoteka, analizu sentimenta (Marrara i sur., 2019) i još mnogo toga (Kasaraneni, 2020, Pejić Bach i sur., 2019). Ova tehnologija omogućuje računalima da obavljaju složene jezične zadatke, čime se poboljšava njihova sposobnost interpretacije i generiranja ljudskog jezika.

NLP obuhvaća tri osnovna koraka u procesu obrade: predobrada teksta, izvlačenje značajki i modeliranje (Kasaraneni, 2020). Predobrada teksta uključuje niz postupaka kojima se neorganizirani i neuređeni izvorni tekst „čisti“ i priprema za daljnju obradu. Ovaj korak može uključivati uklanjanje nepotrebnih informacija, normalizaciju teksta i standardizaciju formatiranja (Kasaraneni, 2020). Nakon predobrade, slijedi izvlačenje značajki, što podrazumijeva identificiranje i izdvajanje relevantnih informacija iz teksta koje su važne za specifične analize ili zadatke (Kasaraneni, 2020). Ove informacije zatim se pretvaraju u format prikladan za pohranu u modelu. Na kraju, modeliranje uključuje izgradnju i treniranje modela na temelju obrađenih podataka te korištenje modela za generiranje predikcija ili

donošenje odluka (Kasaraneni, 2020). Ovaj proces omogućuje računalnim sustavima da efikasno razumiju i manipuliraju jezikom, čime se unapređuje njihova funkcionalnost i točnost.

## **6. Vrste agenata za razgovor temeljeno na njihovoj arhitekturi**

Ovisno o arhitekturi tehnologije koja je korištena za implementaciju, agenti za razgovor se dijele u tri kategorije (Artificial Solutions, 2019). Prva kategorija podrazumijevala bi agente temeljene na lingvističkim tehnikama odnosno pravilima (engl. *rule-based chatbot*). Druga kategorija podrazumijeva agente koji su temeljeni na osnovu umjetne inteligencije (engl. *AI chatbot*). U treću kategoriju ulaze agenti koji imaju karakteristike agenata iz prve i druge kategorije odnosno hibridni agenti za razgovor.

### **6.1. Agenti za razgovor temeljeni na pravilima (*Rule-based chatbot*)**

*Rule-based* agenti za razgovor, ili agenti temeljeni na pravilima, pružaju visoku razinu fleksibilnosti i kontrole nad odgovorima koje generiraju (Artificial Solutions, 2019). Ovi agenti za razgovor omogućuju precizno definiranje odgovora za određena pitanja putem unaprijed postavljenih pravila i jezičnih uvjeta (Seljan, 2003; Seljan i sur., 2013) uz primjenu određenih pravila, koji analiziraju riječi, njihove sinonime i obrasce pitanja, čime osiguravaju dosljednost u odgovorima na slična pitanja. Također, omogućuju prilagođavanje uvjeta i pravila kako bi se unaprijedilo razumijevanje i poboljšala učinkovitost interakcije (Artificial Solutions, 2019).

Međutim, unatoč svojim prednostima kao što su fleksibilnost i visoka kontrola, agenti temeljeni na pravilima imaju značajne nedostatke. Razvoj takvih sustava može biti spor i zahtijeva veliki radni napor zbog detaljnog ručnog postavljanja i održavanja pravila. Pored toga, njihova sposobnost da upravljaju kompleksnijim i nepredviđenim upitima može biti ograničena, što često rezultira u manje dinamičnim i prilagodljivim interakcijama. Ovi agenti za razgovor često nude samo osnovne funkcionalnosti i ne mogu se lako prilagoditi novim zahtjevima ili složenijim scenarijima, čime se smanjuje njihova svestranost i korisnost u širokom spektru primjena (Artificial Solutions, 2019).

### **6.2. Agenti za razgovor na osnovi umjetne inteligencije (*AI chatbots*)**

AI agenti za razgovor temeljeni na strojnom učenju su sofisticiraniji i prediktivniji u usporedbi s onima koji se oslanjaju na pravila. Ovi agenti za razgovor omogućuju dublje interakcije i pružaju personaliziranije iskustvo za korisnike (Artificial Solutions, 2019). Kroz analizu podataka i primjenu prediktivne inteligencije, s vremenom stječu sposobnost razumijevanja konteksta i prirodnog jezika, čime mogu prilagoditi svoje odgovore prema

specifičnim potrebama korisnika. Njihova ključna karakteristika je sposobnost učenja iz prethodnih iskustava i kontinuirano poboljšavanje svojih performansi na temelju novih informacija i interakcija (Artificial Solutions, 2019).

Međutim, razvoj AI agenata za razgovor zahtijeva visoku stručnost i pristup velikim količinama podataka za obuku modela. Ovi podaci omogućuju agentima da prepoznaju obrasce i trendove u korisničkim interakcijama, što je ključno za njihovu sposobnost da pružaju relevantne i precizne odgovore. Također, implementacija ovakvih agenata za razgovor može biti složena i skupa, zbog potreba za naprednim algoritmima i infrastrukturom za obradu podataka. Unatoč ovim izazovima, AI agenti za razgovor pružaju značajnu prednost u pružanju personaliziranih i dinamičnih korisničkih iskustava (Artificial Solutions, 2019).

### **6.3. Hibridni agenti za razgovor (*hybrid chatbots*)**

U usporedbi s AI agentima za razgovor, hibridni agenti za razgovor ne zahtijevaju velike količine podataka i često su bolje prilagođeni poslovnim ciljevima. Ovi modeli kombiniraju lingvistički pristup s elementima strojnog učenja, čime omogućuju rukovanje i jednostavnim i složenijim situacijama (Artificial Solutions, 2019). Hibridni agenti za razgovor nude fleksibilnost u primjeni, omogućujući integraciju pravila i algoritama učenja, što ih čini posebno pogodnima za različite poslovne potrebe. Ovaj pristup spaja snagu i preciznost pravila s dinamičnošću i prilagodljivošću strojnog učenja, pružajući optimalnu ravnotežu između jednostavnosti i složenosti. Stoga, hibridni agenti često predstavljaju idealno rješenje za poduzeća koja traže učinkovit i sveobuhvatan alat za automatizaciju korisničkih interakcija (Artificial Solutions, 2019).

## **7. Primjena agenata za razgovor u različitim domenama**

Agenti za razgovor koriste se u različitim domenama i industrijama u svrhu poboljšanja korisničkog iskustva, automatizacije zadataka te povećanja efikasnosti poslovanja. Neke od domena u kojima se koriste agenti za razgovor uključuju edukaciju, korisničku podršku, maloprodaju, financijske usluge, zdravstvo, turizam, marketing, zabavu i medije. Ovi alati omogućuju brže i preciznije odgovore na upite, optimiziraju procese i omogućuju 24/7 dostupnost, što značajno doprinosi poboljšanju operativnih performansi i korisničkog zadovoljstva. Korištenjem agenata, organizacije mogu efikasnije upravljati velikim količinama interakcija, pružiti personalizirane usluge i osloboditi ljudske resurse za složenije zadatke.

### **7.1. Obrazovanje**

U obrazovanju, agenti za razgovor koriste se u različite svrhe, kao što su pomoć pri učenju, asistencija studentima, učenicima i profesorima. Kao takvi, pružaju obrazovni sadržaj te pomažu u izgradnji i olakšavanju obrazovnih programa. S obzirom na sve veći broj studenata i učenika, profesorima često nedostaje vremena za posvećivanje individualnoj pažnji svakom studentu zbog velikog broja učenika i povećanog radnog opterećenja. Uvođenjem agenata za razgovor, svakoj osobi omogućuje se pristup personaliziranoj pomoći pri učenju, čime se značajno poboljšava kvaliteta obrazovanja.

Agenti za razgovor također igraju ključnu ulogu u rješavanju administrativnih upita i problema, kao što su upis u kolegije, raspored predavanja i uvid u ocjene (Hien et al., 2018). Ovaj pristup smanjuje preopterećenje administrativnih radnika, omogućujući im da se usmjere na složenije zadatke. Također, istraživanja pokazuju da studenti preferiraju korištenje agenta zbog njihove dostupnosti i brzine odgovora, dok im rad s profesorima može biti otežan zbog manjka samopouzdanja i ograničenog vremena (Haristiani, 2019). Korištenje agenata za razgovor stoga ne samo da poboljšava pristup informacijama i uslugama, već i doprinosi većem zadovoljstvu i učinkovitosti u obrazovnom procesu. Istovremeno postavljaju se brojna etička pitanja o odgovarajućem korištenju chatbotova, akademskom integritetu i zaštiti podataka.

### **7.2. Korisnička podrška**

Korisničku podršku danas koriste gotovo sve tvrtke kako bi optimizirale upravljanje odnosima s korisnicima i poboljšale efikasnost svojih usluga. Jedna od glavnih prednosti korištenja agenata za razgovor u korisničkoj podršci je rasterećenje posla za zaposlenike,

omogućujući im da se fokusiraju na složenije i specifičnije zahtjeve (Medium, 2023). Agenti za razgovor mogu pružiti brze odgovore na često postavljana pitanja, što značajno smanjuje vrijeme čekanja za korisnike i poboljšava ukupnu efikasnost podrške (Medium 2023). Međutim, postoje i nedostaci, uključujući mogućnost dugog čekanja na odgovore ili nisku relevantnost odgovora zbog ograničenog razumijevanja konteksta (Nuruzzaman i Hussain, 2018). Napredak u primjeni agenata za razgovor za korisničku podršku očituje se u povećanom zadovoljstvu korisnika, jer im je pomoć dostupna 24/7, neovisno o radnom vremenu tvrtke, čime se osigurava neprekidna i učinkovita komunikacija (Medium, 2023). Osim toga, suvremeni agenti za razgovor sve više koriste napredne algoritme za razumijevanje i analizu korisničkih potreba, što dodatno unapređuje kvalitetu usluge.

### **7.3. Zdravstvo**

U zdravstvu, agenti za razgovor igraju ključnu ulogu u poboljšanju učinkovitosti i dostupnosti zdravstvenih usluga. Oni se koriste za pružanje informacija o pacijentovim zdravstvenim kartonima, informacijama o terapijama, pomoć u postavljanju dijagnoza i preporučivanju terapija temeljenih na simptomima pacijenata (The Top 12 Healthcare Chatbots, 2020). Iako istraživanja pokazuju da pacijenti često preferiraju razgovor s agentom zbog njegove dostupnosti i strpljenja, fizički kontakt s liječnikom ne može se i ne smije zamijeniti. Agenti su uvijek na raspolaganju, ali oni nisu zamjena za osobni pregled i savjetovanje od strane stručnjaka (Altamimi et al., 2023).

Kako bi se olakšao posao liječnicima, agenti se često koriste za obavljanje administrativnih zadataka kao što su zakazivanje termina, izdavanje recepata, pronalaženje bolnica ili slanje podsjetnika za medicinske preglede i podizanje lijekova (Altamimi et al., 2023). U psihijatriji, agenti za razgovor se koriste za razgovor s pacijentima kako bi se spriječile suicidalne misli i pružila emocionalna podrška. Dostupnost agenta, posebno izvan radnog vremena, ima veliki značaj za pacijente s mentalnim zdravstvenim problemima jer im omogućuje da dobiju pomoć kad im je najpotrebnija, neovisno o radnim satima i danima.

### **7.4. Marketing**

U digitalnom dobu, agenti za razgovor postaju sve popularniji alati u marketingu, posebno za sponzoriranje sadržaja i proizvoda kupcima. Agenti za razgovor se koriste za interakciju s potencijalnim kupcima, prilagođavajući sadržaj koji šalju na temelju tih interakcija. Analizom prikupljenih podataka o korisnicima, agenti mogu kreirati personalizirane ponude i preporuke proizvoda koji su usklađeni s interesima i preferencijama kupaca.



Za uspješan marketing putem agenta, važno je da su oni dizajnirani s naglaskom na korisničke interese i preferencije. To zahtijeva temeljito istraživanje i analizu korisničkih želja kako bi se osiguralo da sadržaj bude relevantan i privlačan. Pored toga, agenti za razgovor mogu koristiti podatke iz prošlih interakcija za optimizaciju budućih komunikacija, čime dodatno povećavaju učinkovitost marketinških kampanja.

Povećana personalizacija omogućava brže i preciznije prepoznavanje potreba kupaca, što može značajno poboljšati konverzije i korisničko zadovoljstvo (Medium, 2023). Također, agenti omogućuju automatizaciju odgovora na često postavljana pitanja i pružaju korisnicima trenutnu podršku, čime se štedi vrijeme i resursi. Kako bi se održao visok standard usluge, važno je redovito ažurirati agente s novim informacijama i analizirati rezultate kako bi se kontinuirano poboljšavali marketinški pristupi (Medium, 2024). S obzirom na brzi razvoj tehnologije, integracija naprednih analitičkih alata i strojno učenje mogu dodatno unaprijediti sposobnosti agenta za razgovor i pružiti još preciznije i korisnije preporuke.

## **7.5. Turizam**

Korištenje novih tehnologija i agenata za razgovor u turističkoj industriji značajno povećava interes korisnika za ponude različitih brendova. Agenti za razgovor igraju ključnu ulogu u ovom procesu, jer omogućuju korisnicima lakše pronalaženje potrebnih informacija i pružaju instantne odgovore na upite vezane uz putovanja. Osim toga, razvijaju se aplikacije koje integriraju različite funkcionalnosti, uključujući rezervaciju hotela, kupnju avionskih karata, planiranje putovanja s agencijama i prijedloge za destinacije (Ceccarini i Prandi, 2019).

Ove tehnologije ne samo da unapređuju korisničko iskustvo, već također omogućuju personalizaciju ponuda na temelju korisnikovih preferencija i povijesti pretraživanja. Korištenjem analitike i podataka o korisnicima, agenti mogu pružiti preporuke koje odgovaraju specifičnim interesima i potrebama putnika (Medium, 2024). Takav pristup značajno doprinosi povećanju zadovoljstva korisnika, što može dovesti do veće vjernosti prema određenoj agenciji ili brendu.

Pored toga, automatske usluge rezervacije i personalizirani prijedlozi mogu smanjiti vrijeme potrebno za planiranje putovanja, čime se poboljšava ukupno iskustvo korisnika. S obzirom na brzinu kojom se tehnologija razvija, očekuje se da će agenti za razgovor i slične tehnologije nastaviti igrati sve važniju ulogu u oblikovanju budućnosti turističke industrije, čime će dodatno unaprijediti efikasnost i zadovoljstvo korisnika.

## 8. Kriteriji za evaluaciju agenata za razgovor

Ključni korak u razvoju i implementaciji sustava agenata za razgovor je njihova temeljita evaluacija. Osiguranje visoke pouzdanosti, točnosti i učinkovitosti agenta od presudne je važnosti za njihovu uspješnost. Evaluacija treba obuhvatiti nekoliko ključnih aspekata kako bi se osiguralo da agent za razgovor ispunjava svoje funkcije na optimalan način. Neki od ključnih kriterija za evaluaciju uključuju (Peras, 2018):

1. **Interaktivnost:** Procjena sposobnosti agenta za angažiranje korisnika i održavanje tečnog i prirodnog dijaloga.
2. **Točnost:** Mjerenje preciznosti odgovora koje agent pruža u odnosu na postavljena pitanja i potrebe korisnika.
3. **Brzina odgovora:** Evaluacija vremena koje agentu treba za generiranje i isporuku odgovora, kako bi se osiguralo brzo i učinkovito korisničko iskustvo.
4. **Fleksibilnost:** Analiza sposobnosti agenta da se prilagodi različitim scenarijima i upitima, te njegova sposobnost upravljanja nepredviđenim situacijama.
5. **Kvaliteta podataka:** Provjera točnosti i ažurnosti informacija koje agent koristi za pružanje odgovora.
6. **Prirodnost jezika:** Ocjena koliko su odgovori agenta prirodni i razumljivi korisnicima, te koliko uspješno imitira ljudski govor i interakciju.

Uz ove kriterije, važno je također pratiti korisničke povratne informacije i provoditi kontinuirana unapređenja kako bi se osigurala stalna usklađenost s korisničkim očekivanjima i potrebama. Evaluacija agenata za razgovor ne samo da poboljšava njihovu učinkovitost, već i doprinosi općem zadovoljstvu korisnika i uspjehu implementacije sustava.

### 8.1. Interaktivnost

Za postizanje optimalnog korisničkog iskustva ključno je osigurati visoku razinu interaktivnosti između agenta za razgovor i korisnika (ChatBot, 2023). Interaktivnost obuhvaća nekoliko ključnih kriterija, uključujući upotrebljivost, uključenost korisnika i personalizaciju.

Upotrebljivost odnosi se na jednostavnost korištenja agenta i koliko intuitivno korisnici mogu komunicirati s njim. To uključuje jasne i lako razumljive upute koje vode korisnika kroz interakciju te minimiziranje broja koraka potrebnih za postizanje cilja. Jednostavno i

intuitivno korisničko sučelje omogućuje brzu navigaciju i smanjuje frustraciju korisnika (ChatBot, 2023).

Uključenost korisnika podrazumijeva angažiranje korisnika kroz relevantna i promišljena pitanja koja agent postavlja, čime se poboljšava dinamika razgovora i održava korisnikova pažnja. Agent za razgovor bi trebao proaktivno nuditi korisne informacije i voditi razgovor na način koji potiče interakciju i održava interes korisnika.

Personalizacija je ključna za pružanje prilagođenog iskustva korisnicima. To uključuje prilagodbu sadržaja i odgovora na temelju korisnikovih preferencija, potreba i povijesti prethodnih razgovora. Personalizirani pristup omogućava agentu da bolje razumije i odgovori na specifične zahtjeve i interese korisnika, čime se povećava relevantnost i korisnost interakcije (ChatBot, 2023).

Implementacija svih ovih aspekata doprinosi stvaranju angažiranog i zadovoljnog korisničkog iskustva, čime se maksimizira učinkovitost agenta za razgovor i poboljšava ukupno iskustvo korisnika.

## **8.2. Točnost**

Točnost agenta za razgovor ključna je za pružanje korisničkog iskustva koje je pouzdano i korisno. Ključne značajke točnosti uključuju relevantnost odgovora, preciznost informacija i dosljednost (Medium, 2023).

Relevantnost odgovora odnosi se na usklađenost odgovora agenta s postavljenim pitanjem ili zahtjevom korisnika. Ovo označava koliko su odgovori agenta direktno povezani s korisnikovim potrebama i koliko učinkovito zadovoljavaju te potrebe (SoftwareMill, 2024). Relevantnost je od vitalne važnosti za osiguranje da korisnici dobivaju informacije koje su im zaista korisne i koje odgovaraju njihovim specifičnim upitima.

Preciznost informacija predstavlja mjeru točnosti i vjerodostojnosti informacija koje agent za razgovor pruža. Ovo uključuje provjeru činjenica i potvrđivanje da su informacije temeljene na pouzdanim i ažuriranim izvorima. Preciznost osigurava da korisnici primaju točne podatke koji su istiniti i koji ne dovode u zabludu.

Dosljednost agenta za razgovor odnosi se na sposobnost agenta da pruža konzistentne odgovore na ista ili slična pitanja u različitim interakcijama. Ovo uključuje održavanje istih standarda i smjernica u odgovorima kako bi se izbjegle proturječnosti i konfuzija.

Dosljednost je ključna za izgradnju povjerenja korisnika, jer pomaže u održavanju vjerodostojnosti i pouzdanosti agenta za razgovor.

Osiguranje visoke točnosti u svim ovim aspektima ključno je za učinkovitu funkcionalnost agenta i za pružanje korisničkog iskustva koje je precizno, pouzdano i dosljedno.

### **8.3. Brzina odgovora**

Brzina odgovora jedan je od ključnih kriterija za evaluaciju agenta za razgovor, jer direktno utječe na korisničko iskustvo (Gnewuch et al., 2022). Ova komponenta obuhvaća nekoliko važnih aspekata: latenciju, kvalitetu podataka i skalabilnost.

Latencija se odnosi na vrijeme koje je potrebno agentu da obradi korisnički upit i pruži odgovor (Gnewuch et al., 2022). Smanjenje latencije ključno je za pružanje brze i učinkovit odgovor na zahtjeve korisnika.

Kvaliteta podataka igra ključnu ulogu u brzini odgovora agenta. Kvalitetni podaci omogućuju bržu obradu i točnije odgovore. Kvaliteta podataka obuhvaća strukturu podataka i indeksiranje. Struktura podataka se odnosi na način na koji su podaci organizirani i pohranjeni; dobro strukturirani podaci omogućuju agentu da ih brzo pristupi i obradi. Indeksiranje je tehnika koja poboljšava brzinu pretraživanja relevantnih informacija, čime se dodatno smanjuje vrijeme potrebno za generiranje odgovora.

Skalabilnost predstavlja sposobnost agenta da istovremeno obrađuje veliki broj korisničkih zahtjeva bez smanjenja kvalitete odgovora (Mechkaroska et al., 2024). Uspješna skalabilnost osigurava da agent za razgovor može učinkovito funkcionirati i pod velikim opterećenjem, pružajući konzistentnu i brzu podršku korisnicima (Mechkaroska et al., 2024).

Osiguranje svih ovih aspekata bitno je za optimizaciju performansi agenta za razgovor i za pružanje pozitivnog korisničkog iskustva kroz brze i točne odgovore na upite.

### **8.4. Fleksibilnost**

Fleksibilnost agenta za razgovor značajno utječe na korisničko iskustvo, jer omogućuje njegovu prilagodbu različitim zahtjevima i promjenama u kontekstu (Fan et al., 2022). Fleksibilan agent za razgovor može lako odgovarati na različite vrste upita i prilagoditi se specifičnim potrebama korisnika, što poboljšava ukupnu učinkovitost i zadovoljstvo korisnika (Fan et al., 2022).

Uz to, visok nivo fleksibilnosti omogućuje agentu da se brzo adaptira na promjenjive uvjete i nove situacije, čime osigurava kontinuitet u pružanju relevantnih i korisnih informacija. Fleksibilnost također uključuje sposobnost agenta da razumije i interpretira različite varijante jezika, stila komunikacije i konteksta, čime dodatno unapređuje kvalitetu interakcije.

Ovaj pristup doprinosi stvaranju dinamičnog i prilagodljivog korisničkog iskustva, jer agent za razgovor može obraditi širok spektar korisničkih zahtjeva i brzo se prilagoditi promjenama u okruženju i potrebama korisnika.

## **8.5. Prirodnost jezika**

Značajke prirodnosti jezika ključne su za kvalitetu interakcije s agentom za razgovor i uključuju fluenciju, razumljivost, prilagodljivost stila i empatiju (Peras, 2018).

Razumljivost jezika odnosi se na gramatiku i koherenciju odgovora koje agent pruža. Da bi agent za razgovor bio učinkovit, njegovi odgovori moraju biti gramatički ispravni i logički dosljedni, bez slučajnih ili nepovezanih rečenica (Peras, 2018). Jasnoća odgovora je također važna; odgovori trebaju biti formulirani na način koji je lako razumljiv korisnicima i koji izbjegava dvosmislenosti.

Fluencija označava sposobnost agenta da generira tekst koji teče prirodno i glatko, bez znakova napetosti ili nespretnosti, što je mjera koja se često koristi u npr. strojnom prevođenju (Seljan i sur., 2020).

Prilagodljivost stila omogućava agentu da se prilagodi tonu i stilu komunikacije korisnika, čime se stvara osjećaj personalizacije i bliskosti u interakciji.

Empatija je ključna za stvaranje pozitivnog korisničkog iskustva. Agent za razgovor treba imati sposobnost prepoznavanja emocionalnog stanja korisnika i pokazivanja razumijevanja i suosjećanja, što doprinosi ugodnijoj i humanijoj interakciji. Razvijena emocionalna inteligencija omogućava agentu za razgovor da bolje odgovori na korisničke potrebe i osjeća, čime se poboljšava ukupna kvaliteta usluge.

## 9. Testiranje agenta za razgovor

Testiranje agenta za razgovor ključan je korak u osiguravanju njegove učinkovitosti i funkcionalnosti prije nego što se pusti u produkciju. Ovaj proces uključuje procjenu različitih aspekata agenta, uključujući točnost odgovora, brzinu reakcije, interaktivnost i prirodnost jezika. Testiranje omogućuje identifikaciju potencijalnih problema i nedostataka te pruža priliku za optimizaciju performansi i poboljšanje korisničkog iskustva. Uz pomoć različitih metoda i tehnika testiranja, kao što su simulacije korisničkih interakcija, analize podataka i povratne informacije od korisnika, razvivači mogu osigurati da agent ispunjava očekivanja i zahtjeve. Efektivno testiranje također doprinosi pouzdanosti i sigurnosti agenta za razgovor, smanjujući rizik od grešaka i neispravnosti u stvarnim situacijama. Kroz sveobuhvatan pristup testiranju, moguće je maksimizirati vrijednost agenta i osigurati njegovu uspješnu integraciju u poslovne ili obrazovne procese.

### 9.1. Pitanje – odgovor test (engl. *question – answering test*)

Pitanje – odgovor test (engl. *question – answering test - QA*) je ključna metoda za procjenu sposobnosti agenta za razgovor da pravilno interpretira i odgovori na korisnička pitanja. Ovaj test se fokusira na analizu točnosti i relevantnosti odgovora koje agent pruža u odgovorima na specifična pitanja. Kroz simulaciju različitih upita i scenarija, ovaj test omogućuje identificiranje slabosti u razumijevanju prirodnog jezika i pružanju informacija. Cilj je osigurati da agent za razgovor može učinkovito i precizno odgovoriti na pitanja korisnika, čime se povećava njegova korisnost i pouzdanost u stvarnim situacijama (Bahak et al., 2023).

#### 9.1.1. Povijest pitanje-odgovor sustava

U poljima obrade prirodnog jezika (*Natural Language Processing - NLP*), ekstrakcije informacija (*Information Extraction - IE*) i pretraživanja informacija (*Information Retrieval - IR*), test temeljen na principu pitanje – odgovor (*Question-Answering - QA*) igra ključnu ulogu u evaluaciji sustava za automatsko odgovaranje na pitanja. Prvi QA sustav razvijen je 1961. godine, s ciljem odgovaranja na pitanja specifična za američki bejzbol. Ovaj rani sustav, koji se oslanjao na bušene kartice i pohranjene odgovore, koristio je lingvističke obrasce za interpretaciju pitanja i pronalaženje odgovora (Bahak et al., 2023).

Drugi značajan sustav, LUNAR Woods, razvijen je 1973. godine s namjerom pomoći geolozima u pronalaženju informacija vezanih uz NASA-ine Apollo misije na Mjesec. LUNAR Woods koristio je specijalizirane podatke i informacije prikupljene tijekom misija kako bi pružio relevantne odgovore (Bahak et al., 2023). Ovi raniji sustavi postavili su

temelje za moderni *QA* testiranje, omogućujući značajan napredak u razumijevanju i obradi prirodnog jezika.

Konferencija za pretraživanje teksta (engl. *The Text Retrieval Conference* - TREC) igrala je ključnu ulogu u napretku istraživanja sustava za odgovaranje na pitanja (*QA* sustava). Od 1999. godine, TREC je uveo skupove podataka koji istraživačima pružaju novinske članke zajedno s pitanjima i odgovorima na njih. Ovi skupovi podataka omogućili su istraživačima da testiraju i usavršavaju algoritme za pretraživanje i odgovaranje na pitanja u kontroliranim uvjetima, čime su značajno doprinijeli razvoju *QA* sustava (Bahak et al., 2023).

S razvojem internetske mreže i eksplozijom broja web stranica, web dokumenti su postali glavni izvor informacija za *QA* sustave, što je dovelo do potrebe za novim pristupima u razvoju i evaluaciji tih sustava. Posljednjih godina, razvoj *QA* skupova podataka doveo je do značajnih unapređenja u učinkovitosti i točnosti odgovora koje *QA* sustavi mogu pružiti. Ovi skupovi podataka često sadrže raznovrsne izvore informacija, uključujući tekstove iz različitih domena i tipove pitanja, čime omogućuju obuhvatno testiranje i poboljšanje *QA* sustava (Bahak et al., 2023).

Neki od najvažnijih skupova podataka u ovoj domeni uključuju (Bahak et al., 2023):

- Stanfordov skup podataka (engl. *The Stanford Question Answering Dataset* - SQuAD): Pruža pitanja i odgovore temeljene na paragrafima iz Wikipedije, omogućujući razvoj i testiranje modela za odgovaranje na pitanja u tekstualnom kontekstu.
- MS Marco: Skup podataka koji uključuje pitanja postavljena u pretraživačima i odgovorima iz web dokumenata, usmjeren na poboljšanje performansi pretraživačkih sustava.
- WikiQA: Sadržava pitanja i odgovore iz Wikipedije, usmjeren na razvoj modela koji mogu precizno odgovarati na pitanja temeljenim na dužim tekstovima.
- *Natural Questions dataset*: Pruža stvarna pitanja postavljena od strane korisnika Google pretraživača i odgovore iz Googleovih dokumenata, omogućujući razvoj modela koji mogu razumjeti i odgovoriti na pitanja u stvarnim uvjetima pretraživanja.
- *NewsQA dataset*: Uključuje pitanja i odgovore temeljene na vijestima, fokusirajući se na razumijevanje i odgovaranje na pitanja iz novinskih članaka.

Ovi skupovi podataka pružaju istraživačima različite izazove i prilike za poboljšanje QA sustava, omogućujući razvoj naprednih modela koji mogu bolje razumjeti i obraditi kompleksne upite i informacije (Bahak et al., 2023).

### 9.1.2. Evaluacija prema QA sustavu

Kako bi se procijenila točnost i učinkovitost odgovora koje daje zadani agent za razgovor, koristi se specifičan okvir za evaluaciju. Ovaj okvir uključuje tri ključne komponente: generator upita, zadani agent i evaluator odgovora (Bahak et al., 2023). Struktura QA skupa podataka obično je organizirana u formi trojki, gdje P označava tekstualni paragraf, Q označava pitanje vezano uz paragraf, a A označava točan odgovor na postavljeno pitanje (Bahak et al., 2023).

Generator upita (engl. *prompt builder*) koristi se za generiranje upita na temelju dostupnih podataka koji se unose u zadani agent za razgovor. Postoje dva glavna tipa skupova podataka za QA testiranje, označena kao tip 1 i tip 2 (Bahak et al., 2023).

Tip 1 uključuje pitanja za koja se zna da imaju odgovore u danim tekstualnim paragrafima. U ovom tipu evaluacije, generator upita stvara upite čiji odgovori moraju biti prisutni u tekstu. Ulaz za tip 1 je formuliran kao (Bahak et al., 2023):

"What is the answer to this question: '[Q]' based on this paragraph: '[P]'? Provide only the answer without restating the full sentence."

Tip 2 s druge strane može sadržavati pitanja na koja odgovori nisu uvijek prisutni u zadanim tekstualnim paragrafima. Ovaj tip evaluacije testira sposobnost agenta da upravlja i s pitanjima koja ne mogu biti odgovorena na temelju dostupnih informacija (Bahak et al., 2023).

Za tip 1 evaluacije, isključivo se koriste pitanja čiji odgovori su prisutni u tekstu kako bi se omogućila provjera točnosti odgovora. Ovaj pristup omogućuje detaljnu analizu kako dobro agent prepoznaje i koristi informacije iz teksta kako bi pružio točne i relevantne odgovore.

Kroz ovakav okvir evaluacije, moguće je precizno mjeriti performanse agenta u razumijevanju i odgovaranju na pitanja, identificirati područja za poboljšanje i osigurati da agent učinkovito koristi dostupne podatke za generiranje odgovora.



### 9.1.3. Mjere za evaluaciju

Tri ključne mjere koje se koriste za evaluaciju *QA* sistema uključuju točan odgovor (engl. *Exact Match*), F1 rezultat (engl. *F1 Score*) i odziv (engl. *Recall*) (Bahak et al., 2023).

- Točan odgovor (*Exact Match*) mjeri postotak odgovora koje generira sistem koji se potpuno podudaraju s točnim odgovorima definiranim u skupu podataka. Ova mjera daje precizan uvid u to koliko je sistem sposoban pružiti potpuno točne odgovore bez ikakvih varijacija ili pogrešaka (Bahak et al., 2023).
- F1 rezultat (*F1 Score*) je harmonijska sredina između preciznosti (engl. *Precision*) i odziva. Ova mjera uzima u obzir kako točno sistem identificira relevantne odgovore (preciznost) i koliko često pronalazi sve relevantne odgovore (odziv). F1 rezultat daje uravnotežen prikaz učinkovitosti sistema u oba aspekta, pružajući cjelovitu sliku njegove sposobnosti u odgovaranju na pitanja (Bahak et al., 2023).
- odziv (*Recall*) mjeri sposobnost *QA* sistema da pronađe sve relevantne odgovore unutar skupa podataka ili dokumenata. Izračunava se kao omjer između broja točno identificiranih relevantnih odgovora i ukupnog broja relevantnih odgovora koji se trebaju pronaći. Visok odziv ukazuje na sposobnost sustava da otkrije većinu relevantnih informacija, dok niži odziv može značiti da sistem propušta važne podatke (Bahak et al., 2023).

Kombinacija ovih mjera omogućuje sveobuhvatnu evaluaciju performansi *QA* sistema, pomažući u identifikaciji njegovih jakih i slabih strana. Precizno mjerenje točnosti, F1 rezultata i odziva omogućuje poboljšanje sistema i prilagodbu njegovih sposobnosti kako bi bolje služio korisnicima i pružao kvalitetnije odgovore.

## 9.2. Matrica konfuzije (engl. *Confusion matrix*)

Za analizu efektivnosti i točnosti agenta za razgovor koristi se matrica konfuzije (engl. *Confusion Matrix*). Ova matrica pruža detaljan pregled performansi modela kroz četiri ključne komponente koje predstavljaju odnose između predviđanja i stvarnih rezultata. Matrica je podijeljena na četiri ćelije koje obuhvaćaju (Aloqayli i Abdlelhafez, 2023):

- True Positive (TP): Broj ispravno predviđenih pozitivnih slučajeva, tj. kada je agent točno identificirao relevantne informacije.

- False Positive (FP): Broj netočno predviđenih pozitivnih slučajeva, tj. kada je agent za razgovor pogrešno označio nerelevantne informacije kao relevantne.
- True Negative (TN): Broj ispravno predviđenih negativnih slučajeva, tj. kada je agent za razgovor točno identificirao nerelevantne informacije.
- False Negative (FN): Broj netočno predviđenih negativnih slučajeva, tj. kada je agent za razgovor propustio relevantne informacije i označio ih kao nerelevantne.

Na temelju ove matrice mogu se zračunati tri ključne karakteristike agenta za razgovor, koje su prethodno bile spomenute te će kasnije i same formule biti navedene (Aloqayli i Abdlelhafez, 2023):

- Točnost (engl. *Accuracy*): Omjer ukupnog broja ispravnih predviđanja (TP + TN) prema ukupnom broju predviđanja. Točnost pruža uvid u opću učinkovitost agenta u klasifikaciji svih slučajeva.
- Preciznost (engl. *Precision*): Omjer broja točno identificiranih pozitivnih slučajeva (TP) prema ukupnom broju slučajeva označenih kao pozitivni (TP + FP). Preciznost pokazuje koliko su predviđeni pozitivni slučajevi zaista relevantni.
- Odziv (engl. *Recall*): Omjer broja točno identificiranih pozitivnih slučajeva (TP) prema ukupnom broju stvarnih pozitivnih slučajeva (TP + FN). Odziv mjeri sposobnost agenta da pronađe sve relevantne informacije.

Matrica konfuzije omogućuje detaljnu analizu performansi agenta za razgovor, pomažući u identifikaciji njegovih prednosti i područja za poboljšanje. Analizom točnosti, preciznosti i odziva možemo bolje razumjeti koliko dobro agent ispunjava svoje zadatke i koliko je učinkovit u pružanju točnih i relevantnih odgovora.

Točnost kao evaluacijska funkcija računa se kao omjer broja točnih pretpostavki i broja ukupnih pretpostavki. Formula glasi (Aloqayli i Abdlelhafez, 2023):

$$Točnost = \frac{BROJ\ TOČNIH\ PRETPOSTAVKI}{UKUPAN\ BROJ\ PRETPOSTAVKI}$$

Prema matrici konfuzije točnost se izračunava prema pozitivima i negativima, a formula glasi (Aloqayli i Abdlelhafez, 2023):

$$Točnost = \frac{TP+TN}{TP+TN+FP+FN}$$

Preciznost računa se kao omjer točno predviđenih pozitiva i sveukupnih pozitiva. Formula za preciznost glasi (Aloqayli i Abdlelhafez, 2023):

$$Preciznost = \frac{TP}{TP+FP}$$

Ako je preciznost velika, to označava da je velik broj pozitiva ispravno klasificiran. Međutim, ako je preciznost mala onda je velik broj pozitiva klasificiran netočno. Kako bi model mogao donijeti ispravnu, odnosno točnu pretpostavku, preciznost mora biti što veća.

Odziv je evaluacijska metoda koja se računa kao omjer ukupnog broja pozitiva i onih pozitiva koji su točno pretpostavljeni (Aloqayli i Abdlelhafez, 2023).

$$Odziv = \frac{TP}{TP+FN}$$

Funkcija recall temelji se samo na uzorcima iz pozitivne klase, dok ne ovisi o uzorcima iz klase negativno.

### 9.3. Korisničko iskustvo

Za evaluaciju agenta za razgovor ključno je provesti anketu koja će pružiti uvid u iskustvo korisnika prilikom interakcije s agentom. U tu svrhu često se koristi Likertova skala (engl. Likert scale), koja je jedan od najraširenijih alata za mjerenje stavova i mišljenja korisnika (Hrvatska enciklopedija, 2024). Likertova skala obično uključuje između 15 i 20 tvrdnji koje se odnose na različite aspekte agenta, poput njegove upotrebljivosti, točnosti odgovora i ukupnog zadovoljstva korisnika (Hrvatska enciklopedija, 2024).

Korisnici ankete odgovaraju na svaku tvrdnju odabirom jednog od pet ponuđenih odgovora, koji predstavljaju stupanj slaganja s iznesenom tvrdnjom. Opcije odgovora su: potpuno slaganje, slaganje, neutralno, neslaganje i potpuno neslaganje (Hrvatska enciklopedija, 2024). Svaki odgovor se boduje, a ukupni rezultati ankete predstavljaju zbir bodova koji odražava stav korisnika prema agentu za razgovor.

Analizom rezultata ankete moguće je identificirati ključne aspekte korisničkog iskustva, uključujući područja u kojima agent ispunjava očekivanja korisnika i ona u kojima je potrebno poboljšanje. Likertova skala omogućuje strukturirano prikupljanje i analizu podataka, što pomaže u kvantitativnom ocjenjivanju performansi agenta i donošenju informiranih odluka za buduća poboljšanja (Hrvatska enciklopedija, 2024).

## 10. Istraživanje

Kako je već bilo navedeno, primjena agenata za razgovor u današnjem digitalnom okruženju postaje sve raširenija i važnija. Mnoge web stranice i organizacije integriraju agente kako bi unaprijedile komunikaciju s korisnicima. Danas se agenti za razgovor koriste u raznim sektorima, uključujući korisničku podršku, bankarstvo, obrazovanje, zdravstvo, marketing i turizam (ResearchGate, 2024). Za uspješnu izgradnju agenta za razgovor, ključno je razumijevanje tehnologija kao što su prethodno spomenuta obrada prirodnog jezika (engl. *Natural Language Processing*) te strojno učenje (engl. *Machine Learning*). Ove tehnologije omogućuju agentima za razgovor da razumiju i precizno odgovaraju na korisničke upite, te da komuniciraju na intuitivan način (ProjectPro, 2024). Integracija agenta za razgovor u poslovanje ne samo da poboljšava korisničko iskustvo, već i doprinosi smanjenju operativnih troškova (Khoa, 2021).

U ovom radu bit će prikazana izgradnja agenta za razgovor namijenjenog području obrazovanja, s fokusom na pretraživanje informacija. Ovaj agent za razgovor osmišljen je kako bi pomogao studentima i svim zainteresiranim korisnicima u pronalaženju informacija o upisu na studije Filozofskog fakulteta, Sveučilišta u Zagrebu. Pored toga, agent će omogućiti brzi pristup relevantnim podacima i odgovorima na često postavljana pitanja, čime će poboljšati korisničko iskustvo i učinkovitost informiranja.

U okviru ovog istraživanja, tijekom određenog vremenskog razdoblja, sustavno su prikupljeni materijali potrebni za razvoj naprednog agenta za razgovor. Glavni cilj bio je prikupiti i organizirati sve relevantne informacije sa službene mrežne stranice Filozofskog fakulteta u Zagrebu, kako bi se stvorila čvrsta baza podataka za treniranje agenta. Ukupno je prikupljeno 657 poveznica koje obuhvaćaju ključne aspekte sadržaja na web stranici, uključujući informacije o studijskim programima, nastavnom osoblju, administrativnim uslugama i ostalim važnim resursima, pogotovo onima koji su bitni za upis na fakultet. Nakon prikupljanja, ove su poveznice unijete u sustav za izgradnju agenta, čime je omogućeno automatsko generiranje i početno treniranje modela. Ovaj postupak automatskog treniranja omogućava agentu da prepozna i odgovori na širok spektar pitanja korisnika koristeći prikupljene podatke. Međutim, kako bi se osigurala visoka razina točnosti i relevantnosti odgovora, postoji i opcija dodatnog, ručnog treniranja agenta. U ovom koraku, ručno se mogu dodavati specifične odgovore na složenija pitanja ili prilagođavati postojeće odgovore kako bi se postigla optimalna kvaliteta interakcije s korisnicima. Ovaj pristup kombinira snagu

automatizacije s preciznošću ljudske intervencije, što rezultira visoko funkcionalnim i prilagodljivim agentom za razgovor.

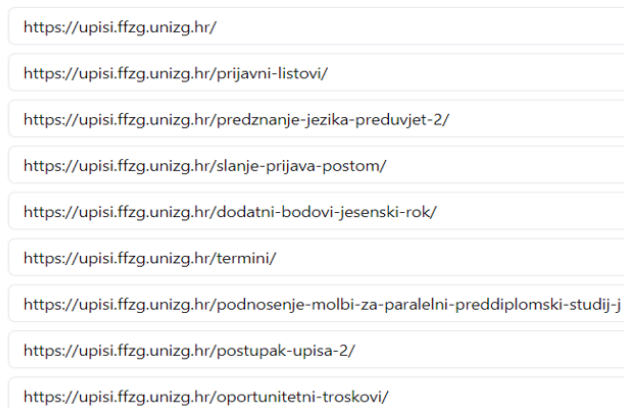
### **10.1. Razvojno okruženje**

Za izgradnju agenta za razgovor bit će korišten novi popularni AI alat *Chatbase*, koji omogućuje kreiranje agenata za razgovor temeljenih na vlastitim podacima kao što su web stranice ili specifični dokumenti. Chatbase je dostupan kako za profesionalnu, tako i za amatersku uporabu, pružajući fleksibilnost u pristupu njegovim značajkama. Alat nudi četiri različita plana korištenja.

Prvi plan je besplatan i omogućuje izradu jednog agenta za razgovor s ograničenim brojem podataka za obuku. Drugi plan, nazvan „*Hobby*“, uključuje sve značajke besplatnog plana, ali omogućuje veći unos podataka i dodatne mogućnosti prilagodbe. Treći i četvrti plan, „*Standard*“ i „*Unlimited*“, namijenjeni su profesionalnim korisnicima i tvrtkama, nudeći širi spektar funkcionalnosti i mogućnosti. Svaki plan može se odabrati kao mjesečna ili godišnja pretplata, pri čemu cijene variraju od 99 dolara do 399 dolara mjesečno, ovisno o odabranom planu i razini usluga.

Za početak korištenja alata potrebno je prijaviti se na sustav i izabrati program korištenja prema željama i potrebama. Nakon što je plan odabran, korisnici mogu započeti s izgradnjom svog agenta. Prvi korak u izgradnji agenta za razgovor je unos podataka na kojima će agent biti treniran. Alat nudi mogućnosti unošenja vlastitih datoteka s računala, teksta i linkova web stranica. Također, korisnici mogu prilagoditi postavke obuke kako bi optimizirali performanse agenta za specifične scenarije i upite. Nakon učitavanja podataka, alat pruža opcije za testiranje i fino podešavanje agenta za razgovor prije njegove konačne implementacije.

Ispod Slika 2 prikazuje ispis svih automatski generiranih poveznica iz mrežne adrese web stranice Filozofskog Fakulteta u Zagrebu na temelju kojih alat automatski izrađuje agenta za razgovor.



Slika 2. Prikaz poveznica iz mrežne adrese web stranice Filozofskog fakulteta u Zagrebu

Osim poveznica, moguće je koristiti i datoteke za treniranje, kao što su PDF dokumenti, dokumenti u Wordu, brošure, itd.

U ovom završnom radu podaci su uneseni pomoću linkova web stranice Filozofskog fakulteta u Zagrebu. Nakon uspješno unesenih podataka i generiranja potrebnih veza potrebno je Pokrenuti kreirani agent za razgovor.

## 10.2. Sučelje agenta za razgovor

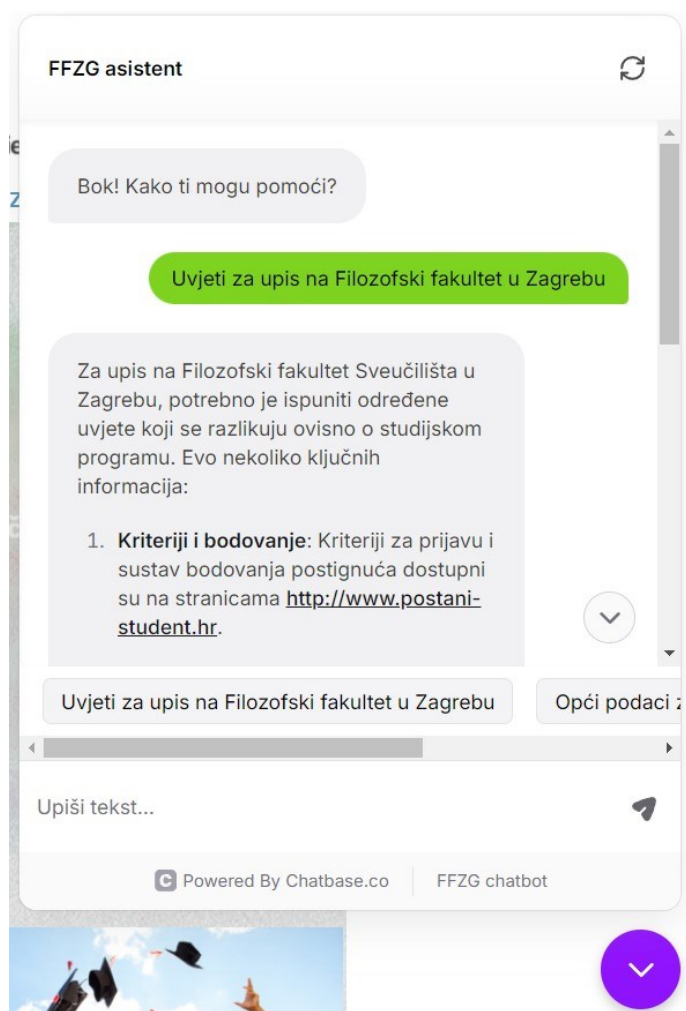
Korisničko iskustvo pri korištenju agenta za razgovor predstavlja ključni faktor uspjeha. Osim same izvedbe agenta za razgovor i točnosti njegovih odgovora, iznimno je važno pravilno urediti sučelje agenta za razgovor kako bi korisnici imali ugodno i intuitivno iskustvo. Sučelje agenta za razgovor treba biti jednostavno za navigaciju i vizualno privlačno kako bi se osiguralo da ga može koristiti široka populacija.

Postoji širok raspon opcija za prilagodbu sučelja agenta za razgovor, omogućujući dizajnerima da optimiziraju izgled i funkcionalnost agenta prema specifičnim potrebama i preferencijama. Na početku procesa dizajner odabire početne poruke koje će agent za razgovor koristiti za započinjanje razgovora (engl. *Initial Messages*). Također, može postaviti unaprijed ponuđena pitanja (engl. *Suggested Messages*) koja će se prikazivati korisnicima kao opcije za interakciju s agentom.

Funkcija “*Message Placeholder*” omogućuje dodavanje privremenih poruka u tekstualni okvir za unos, poput “Unesi tekst...”, koje služe kao vizualni vodič za korisnike. Ove poruke pomažu korisnicima da razumiju gdje i što trebaju upisati. Uz to, funkcija podnožja (engl. *Footer*) omogućuje dodavanje informacija koje će se prikazivati na dnu sučelja agenta za

razgovor, kao što su kontakti, informacije o organizaciji ili druge relevantne napomene. Sve ove opcije doprinose stvaranju profesionalnog i korisnički prihvatljivog sučelja koje poboljšava cjelokupno iskustvo korisnika.

U daljnjem procesu uređivanja moguće je izabrati boju odnosno temu samog sučelja (engl. *Theme*), ime (naslov) koje će se prikazivati na vrhu agenta (engl. *Display name*), profilnu sliku (engl. *Profile Picture*), ikonu agenta (engl. *Chat Icon*), boju poruka korisnika (engl. *User Message Color*), boju ikone za ulaz u razgovor (engl. *Chat Bubble Button Color*) te odabrati na kojoj strani će se ta ikona nalaziti. Slika 3 prikazuje sučelje kreiranog agenta za razgovor.



Slika 3. Korisničko sučelje agenta za razgovor

## 11. Anketa

Anketa je osmišljena kao ključni alat za prikupljanje podataka potrebnih za evaluaciju agenta za razgovor. Ova anketa, dostupna putem Google Forms platforme, strukturirana je tako da ispitanici jednostavno i efikasno odgovore na pitanja, pružajući relevantne informacije koje će se koristiti za analizu i donošenje zaključaka u okviru završnog rada.

### 11.1. Struktura ankete

Anketa se sastoji od 8 pitanja zatvorenog tipa, odnosno teza koje pokrivaju različite aspekte istraživačkog pitanja, odnosno kvalitete agenta za razgovor, na način da se provjeravalo slaganje ispitanika s navedenim tvrdnjama.

Odgovori na pitanja su bazirani na Likertovoj skali, kako je prethodno bilo navedeno, gdje su ponuđeni odgovori bili – “Uopće se ne slažem”, “Ne slažem se”, “Ne znam”, “Slažem se” te “Potpuno se slažem”.

U anketi su postavljena sljedeća pitanja:

- (1) Sučelje chatbota je intuitivno i lako za korištenje.
- (2) Chatbot je dao razumljive i jasne informacije.
- (3) Odgovori chatbota su točni.
- (4) Chatbot razumije postavljena pitanja.
- (5) Chatbot odgovara na sva pitanja vezana za upise na FFZG.
- (6) Odgovori chatbota su brzi.
- (7) Chatbot pokazuje dobro znanje o upisima na FFZG.
- (8) Zadovoljan sam iskustvom korištenja chatbota.

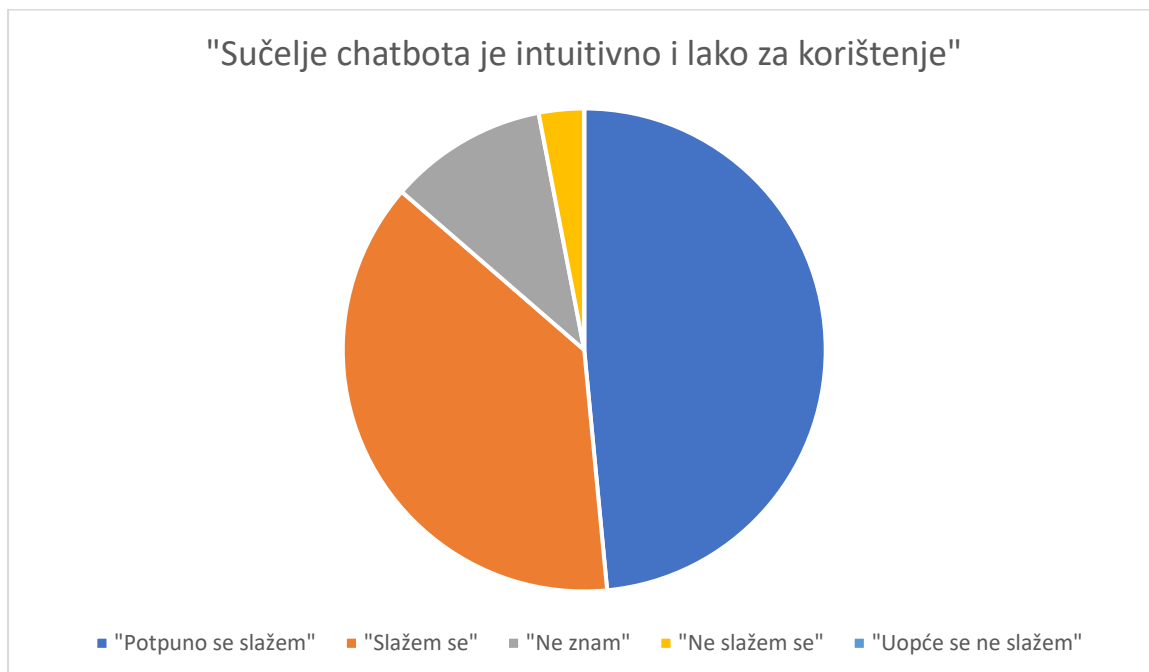
### 11.2. Analiza ankete

Na provedenu anketu odgovorilo je 66 ispitanika. Ista je bila poslana budućim studentima prve godine Filozofskog fakulteta u Zagrebu putem društvenih mreža.

(1) Na prvu tvrdnju koja glasi “Sučelje chatbota je intuitivno i lako za korištenje”, najveći broj ispitanika, njih 32, odnosno 48,5%, odgovorilo je s “Potpuno se slažem”. Odgovor “Slažem se” dalo je 25 ispitanika, odnosno 37,9%. Njih 7, odgovorilo je s “Ne znam”, što bi iznosilo 10,6%. Odgovor “Ne slažem se” dalo je samo 2 ispitanika, odnosno 3%, dok odgovor “Uopće se ne slažem” nije odabrao niti jedan ispitanik. Navedeni rezultat

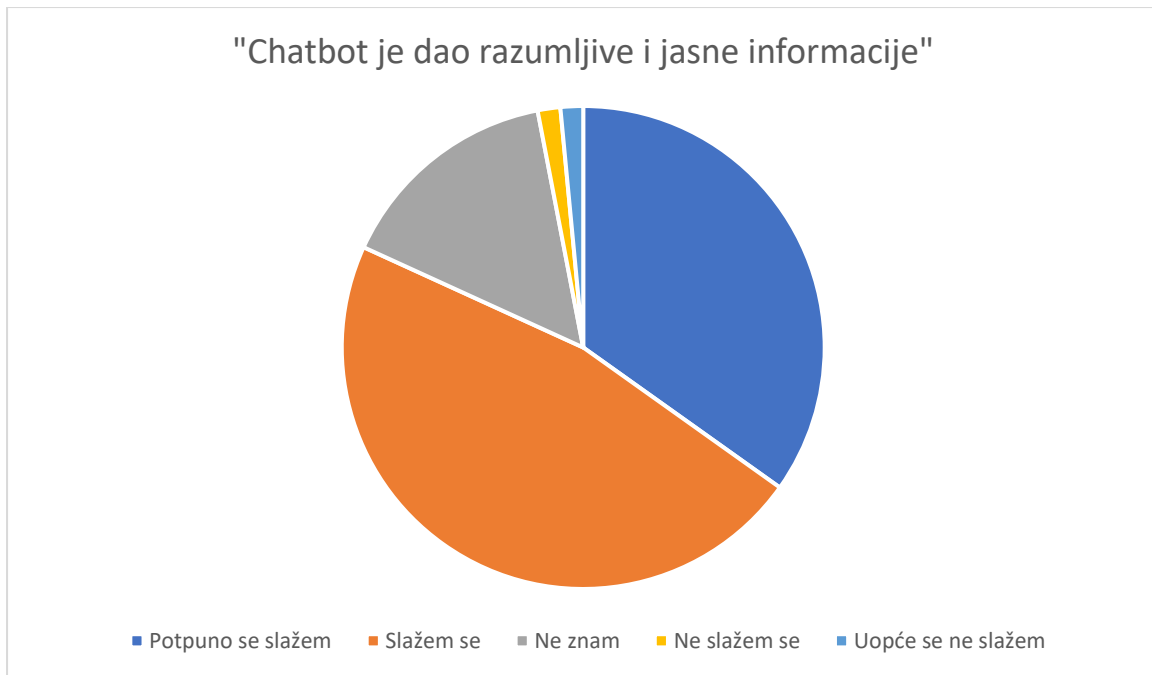


ukazuje da se gotovo 87% ispitanika slaže s tvrdnjom da je sučelje intuitivno i lagano za korištenje.



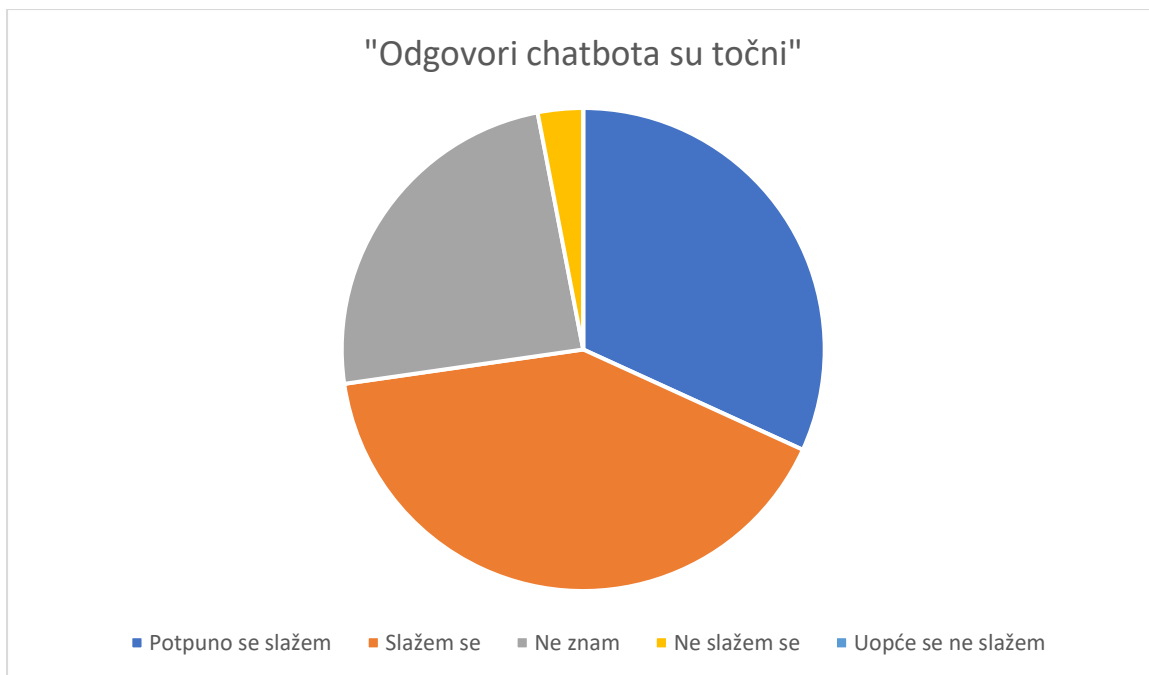
Slika 4. Grafički prikaz rezultata ankete za prvu tvrdnju

(2) Na drugu tvrdnju koja glasi "Chatbot je dao razumljive i jasne informacije", najveći broj ispitanika dao je odgovor "Slažem se". Broj ispitanika koji je izabrao taj odgovor je 31, odnosno 47%. Sljedeći najčešće odabran odgovor bio je "Potpuno se slažem". 23 ispitanika odabralo je taj odgovor, što iznosi 34,8%. Rijeđe odabrani odgovori su "Ne znam", s 10 ispitanika, odnosno 15,2% te odgovori "Ne slažem se" i "Uopće se ne slažem" s jednim ispitanikom, odnosno 1,5%. Odgovor "Uopće se ne slažem" nije odabrao niti jedan ispitanik. Navedeni rezultati ukazuju da se 81,8% ispitanika slaže s tvrdnjom da chatbot daje razumljive i jasne informacije.



Slika 5. Grafički prikaz rezultata ankete za drugu tvrdnju

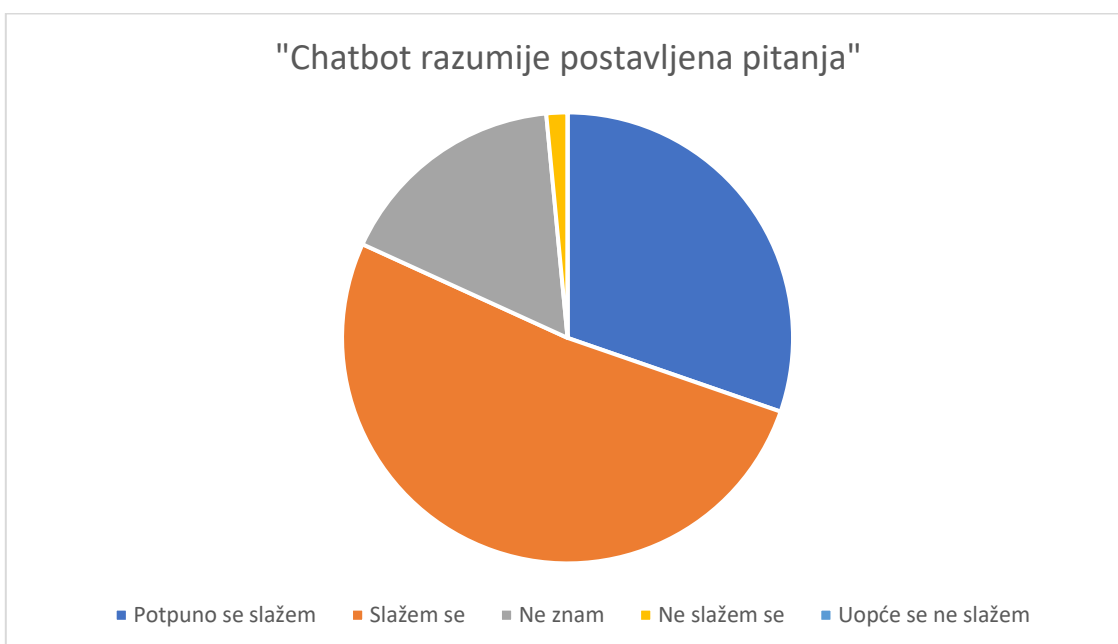
(3) Na treću tvrdnju koja glasi "Odgovori chatbota su točni", najčešći odgovor, kojeg su odabrala 27 ispitanika, odnosno 40,9% bilo je "Slažem se". Odgovor "Potpuno se slažem" izabralo je 21 ispitanika, 31,8%. Treći najčešći odgovor bio je "Ne znam", a odabralo ga je 16 ispitanika, što bi iznosilo 24,2%. Tek 2 ispitanika na pitanje su odgovorili s "Ne slažem se", što iznosi samo 3%. Odgovor "Uopće se ne slažem" nije odabrao niti jedan ispitanik.



Slika 6. Grafički prikaz rezultata ankete za treću tvrdnju

(4) Na četvrtu tvrdnju koja glasi "Chatbot razumije postavljena pitanja",

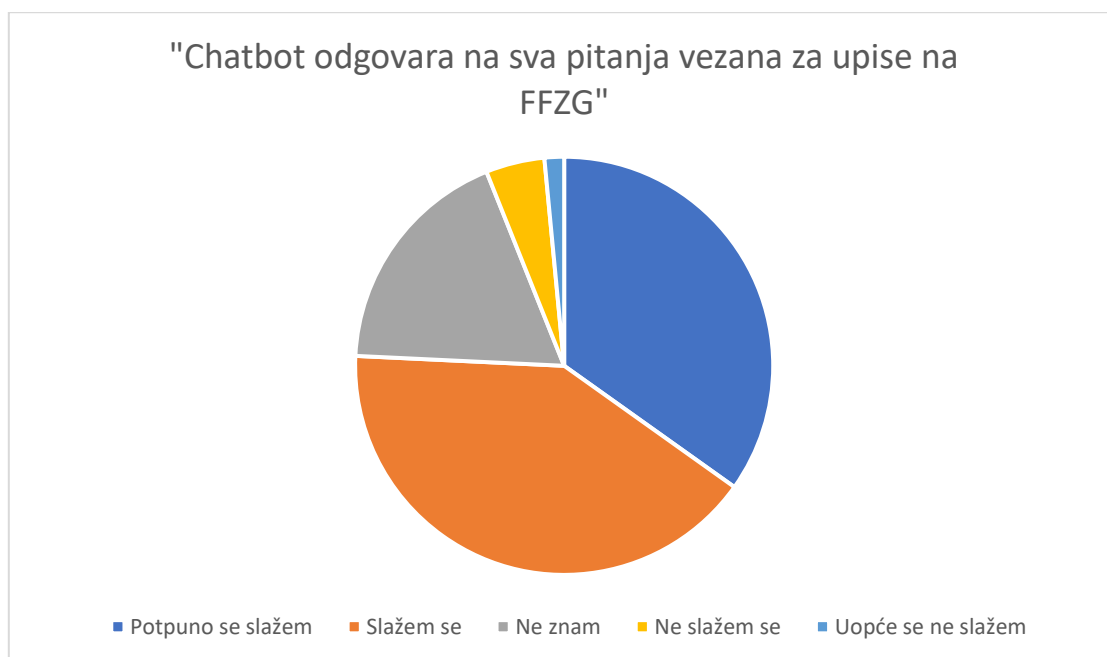
34 ispitanika odgovorilo je sa "Slažem se", što je 51,5%. Odgovor "Potpuno se slažem" dao je 20 ispitanika, odnosno 30,3%. 11 ispitanika, odnosno 16,7%, odgovorilo je s "Ne znam", a samo 1 ispitanik, 1,5%, dalo je odgovor "Ne slažem se". Odgovor "Uopće se ne slažem" nije odabrao niti jedan ispitanik.



Slika 7. Grafički prikaz rezultata ankete za četvrtu tvrdnju

(5) Na petu tvrdnju koja glasi "Chatbot odgovara na sva pitanja vezana za upise na FFZG",

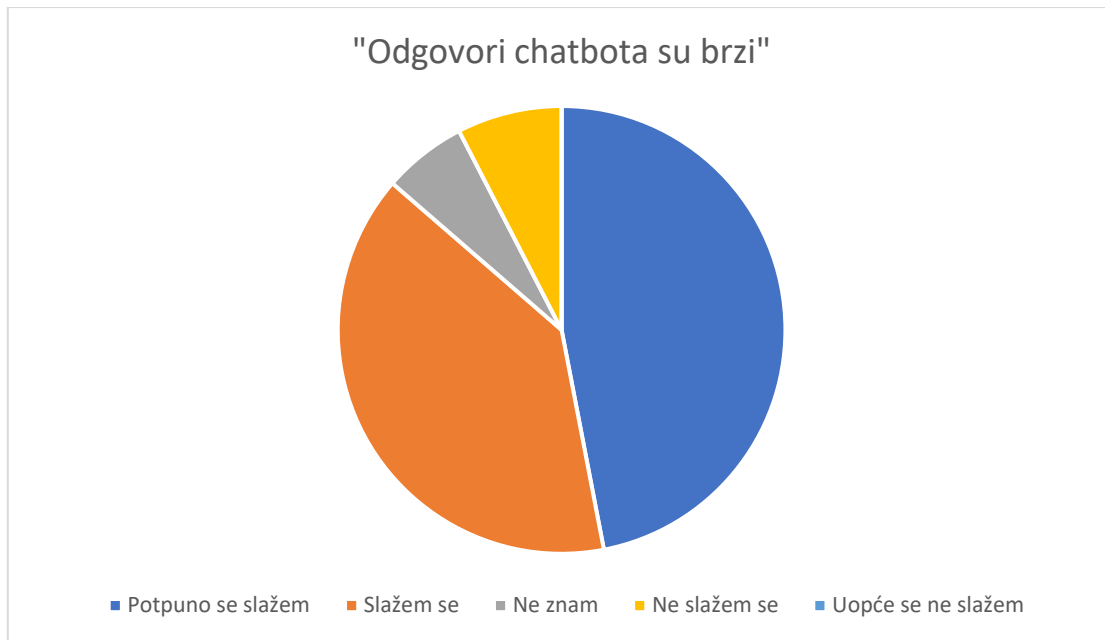
najčešći odgovor bio je "Slažem se". Taj odgovor odabralo je 27 ispitanika, odnosno 40,9%. Odgovor "Potpuno se slažem" izabralo je 23 ispitanika, 34,8%. 12 ispitanika odgovorilo je s "Ne znam", 18,2%. "Ne slažem se" je bio odgovor kod 3 ispitanika, 4,5%, dok je samo jedan ispitanik izabrao odgovor "Uopće se ne slažem", što bi iznosilo 1,5%.



Slika 8. Grafički prikaz rezultata ankete za petu tvrdnju

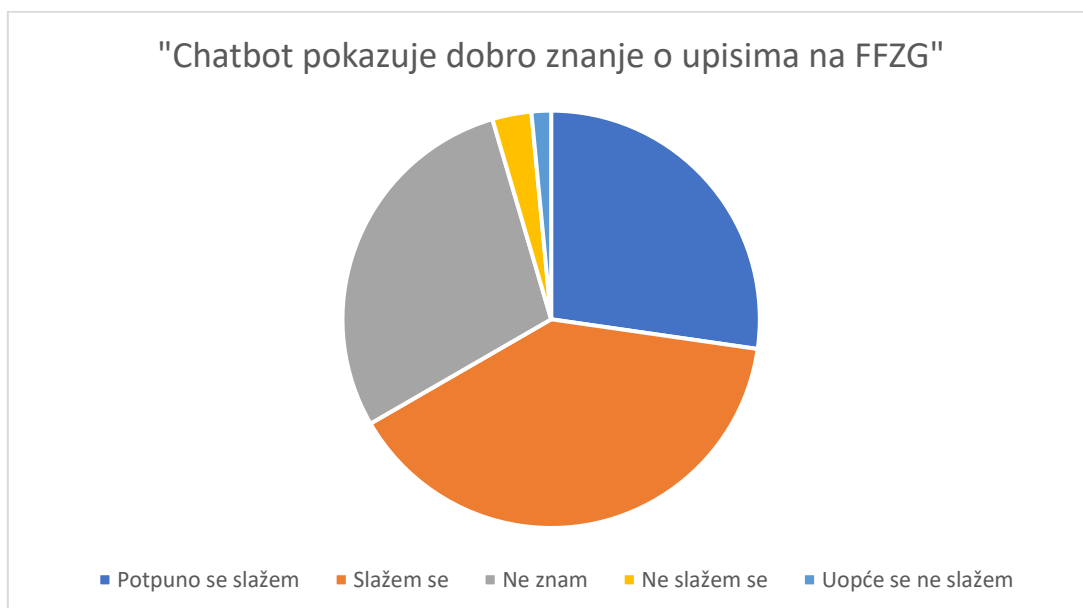
(6) Na šestu tvrdnju koja glasi "Odgovori chatbota su brzi",

31 ispitanik odgovorilo je s "Potpuno se slažem", 47%. "Slažem se" odabralo je 26 ispitanika, 39,4%, dok je odgovor "Ne znam" odabran 4 puta, 6,1%, a odgovor "Ne slažem se" odabran 5 puta, odnosno 7,6%. Odgovor "Uopće se ne slažem" nije odabrao niti jedan ispitanik.



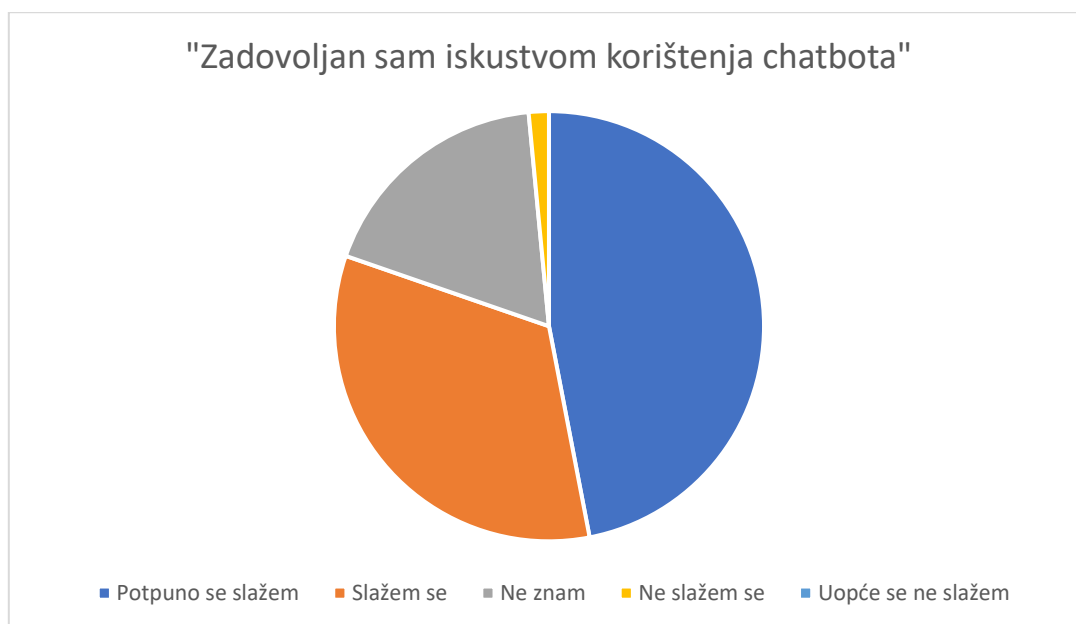
Slika 9. Grafički prikaz rezultata ankete za šestu tvrdnju

(7) Na sedmu tvrdnju koja glasi "Chatbot pokazuje dobro znanje o upisima na FFZG", najveći broj ispitanika je odgovorilo sa "Slažem se", njih 26, odnosno 39,4%. Odgovor "Ne znam" odabralo je 19 ispitanika, 28,8%, dok je odgovor "Potpuno se slažem" odabralo 18 ispitanika, odnosno 27,3%. Manje odabrani odgovori su "Ne slažem se" i "Uopće se ne slažem", a to su dva odabira za odgovor "Ne slažem se", odnosno 3% te 1 ispitanik, 1,5% za odgovor "Uopće se ne slažem".



Slika 10. Grafički prikaz rezultata ankete za sedmu tvrdnju

(8) Na osmu tvrdnju koja glasi “Zadovoljan sam iskustvom korištenja chatbota”, a koja se odnosi na sveukupno zadovoljstvo korištenja agenta za razgovor, kako je već bilo navedeno. 31 ispitanik odgovorio je s “Potpuno se slažem”, što iznosi 47%. Odgovor “Slažem se” odabralo je 22 ispitanika, 33,3%. Manji broj ispitanika odgovorilo je s “Ne znam” i “Ne slažem se”. Odgovor “Ne znam” odabralo je 12 ispitanika, 18,2%, a odgovor “Ne slažem se” samo 1 ispitanik, odnosno 1,5%.



Slika 11. Grafički prikaz rezultata ankete za osmu tvrdnju

Pitanje	Potpuno se slažem	Slažem se	Ne znam	Ne slažem se	Uopće se ne slažem
Sučelje chatbota je intuitivno i lako za korištenje	<b>32</b> <b>(48,5%)</b>	<b>25</b> <b>(37,9%)</b>	7 (10,6%)	2 (3%)	0 (0%)
Chatbot je dao razumljive i jasne informacije.	<b>23</b> <b>(34,8%)</b>	<b>31</b> <b>(47%)</b>	10 (15,2%)	1 (1,5%)	1 (1,5%)
Odgovori chatbota su točni.	<b>21</b> <b>(31,8%)</b>	<b>27</b> <b>(40,9%)</b>	<b>16</b> <b>(24,2%)</b>	2 (3%)	0 (0%)
Chatbot razumije postavljena pitanja.	<b>20</b> <b>(30,3%)</b>	<b>34</b> <b>(51,5%)</b>	<b>11</b> <b>(16,7%)</b>	1 (1,5%)	0 (0%)

Chatbot odgovara na sva pitanja vezana za upise na FFZG.	<b>23</b> <b>(34,8%)</b>	<b>27</b> <b>(40,9%)</b>	<b>12</b> <b>(18,2%)</b>	3 (4,5%)	1 (1,5%)
Odgovori chatbota su brzi.	<b>31</b> <b>(47%)</b>	<b>26</b> <b>(39,4%)</b>	4 (6,1%)	5 (7,6%)	0 (0%)
Chatbot pokazuje dobro znanje o upisima na FFZG.	<b>18</b> <b>(27,3%)</b>	<b>26</b> <b>(39,4%)</b>	<b>19</b> <b>(28,8%)</b>	2 (3%)	1 (1,5%)
Zadovoljan sam iskustvom korištenja chatbota.	<b>31</b> <b>(47%)</b>	<b>22</b> <b>(33,3%)</b>	<b>12</b> <b>(18,2%)</b>	1 (1,5%)	0 (0%)

Tablica 1: Tablični prikaz rezultata ankete

### 11.3. Diskusija

Analizom rezultata ankete među budućim studentima prve godine Filozofskog fakulteta u Zagrebu, primjećuje se značajna sklonost pozitivnim stavovima prema postavljenim pitanjima. Većina ispitanika pokazuje visok stupanj slaganja s ponuđenim tvrdnjama, posebno u odgovorima "Potpuno se slažem" i "Slažem se", što ukazuje na općenito zadovoljstvo i pozitivan osjećaj prema odgovorima agenta za razgovor nad kojim je anketa provedena. Na svakom pitanju dominiraju odgovori "Slažem se" i "Potpuno se slažem", dok su negativni odgovori, kao što su "Ne slažem se" i "Uopće se ne slažem", izuzetno rijetki, što sugerira da postoji širok konsenzus među budućim studentima o većini odgovora. Ovi rezultati mogu biti korisni za dodatno razumijevanje stavova i očekivanja budućih studenata, te za prilagodbu pripremnih aktivnosti i komunikacije fakulteta prema njihovim potrebama i željama, a ne samo kao dojmovi o agentu.

## 12. Evaluacija agenta za razgovor

U današnjem digitalnom okruženju, gdje agenti za razgovor igraju ključnu ulogu u unapređenju korisničkog iskustva i optimizaciji poslovnih procesa, procjena njihove učinkovitosti postaje od esencijalne važnosti. Evaluacija agenta za razgovor obuhvaća niz metoda i tehnika koje omogućuju detaljnu analizu njegove performanse, točnosti i interaktivnosti. U ovom poglavlju, fokusirat ćemo se na ključne aspekte evaluacije agenta za razgovor, uključujući analizu njegovih sposobnosti u odgovaranju na pitanja, brzini odgovora, preciznosti, fleksibilnosti i prirodosti komunikacije. Kroz različite metode i kriterije, kao što su metričke analize, ankete korisnika i testiranja performansi, procijenit ćemo kako agenti za razgovor zadovoljavaju postavljene standarde i kako njihova izvedba utječe na ukupno korisničko iskustvo. Cilj ovog poglavlja je pružiti sveobuhvatan pregled tehnika evaluacije koje pomažu u identificiranju snaga i slabosti agenta za razgovor te u formuliranju preporuka za njihovo daljnje poboljšanje.

Za evaluaciju izrađenog agenta za razgovor prema matrici konfuzije potrebno je analizirati četiri parametra. *True positive* (TP) označava sposobnost agenta za razgovor da odgovori točno na očekivane zahtjeve i poruke (Aloqayli i Abdlelhafez, 2023). *False positive* (FP) označava sposobnost agenta za razgovor da odgovori točno na zahtjeve i poruke koje nisu očekivane (Aloqayli i Abdlelhafez, 2023). *True negative* (TN) označava sposobnost agenta za razgovor da odgovori netočno na očekivane zahtjeve i poruke, a *False negative* (FN) označava sposobnost agenta za razgovor da netočno odgovori na pitanja koja su neočekivana (Aloqayli i Abdlelhafez, 2023).

Analiziranjem izrađenog agenta za razgovor prema matrici konfuzije slijedi tablica koja prikazuje omjer parametara navedenih u prijašnjem paragrafu. Na temelju rezultata slijedi izračun točnosti, preciznosti i odziva agenta za razgovor.

	Predviđeno negativno	Predviđeno pozitivno
Stvarno negativno	TN: 2	FP: 1
Stvarno pozitivno	FN: 3	TP: 9

Tablica 2: Matrica konfuzije izrađenog agenta za razgovor



$$\text{Točnost} = \frac{9+2}{9+2+1+3} \times 100 = 73,333\%$$

$$\text{Preciznost} = \frac{9}{9+1} \times 100 = 90\%$$

$$\text{Odziv} = \frac{9}{9+3} \times 100 = 75\%$$

Za evaluaciju agenta za razgovor korišteno je 15 pažljivo odabranih pitanja koja su testirala različite aspekte njegove funkcionalnosti. Ispod su navedena pitanja, za koje je odgovor vrlo lako moguće pronaći na stranicama Filozofskog fakulteta.

Pitanje	
1.	Kako glasi e-mail adresa Odsjeka za anglistiku ?
2.	Koliko katedri postoji na Odsjeku za filozofiju ?
3.	Koji integrirani studiji postoje na Filozofskom fakultetu ?
4.	Na kojoj se adresi nalazi Filozofski fakultet ?
5.	Kako glase podaci za uplatu za upis na diplomatske studije?
6.	Koji izvanredni diplomski studiji postoje na smjeru Informacijske znanosti ?
7.	Koliko ECTS bodova se stječe nakon završetka prijediplomskog studija Informacijskih znanosti?
8.	Koliko je potrebno imati ECTS bodova za upis na diplomski studij?
9.	Kako glasi telefonski broj Odsjeka za pedagogiju?
10.	Koliko ECTS bodova nosi tjelesna i zdravstvena kultura?
11.	Tko je trenutni dekan Fakulteta?
12.	Objasni ECTS Informacijski paket.
13.	Za koje studente je otvoren natječaj za Erasmus na Maldivima?
14.	Za koje studente je otvoren natječaj za Erasmus u Zanzibaru?

15.	Postoji li Erasmus natječaj za Brazil?
-----	--

Tablica 3: Pitanja postavljena agentu za razgovor

Prema analiziranim podacima, točnost agenta za razgovor iznosi 73,33%, što ukazuje na postotak ispravno odgovornih interakcija u odnosu na sve postavljene upite. Preciznost, koja iznosi 90%, mjeri koliko su odgovori agenta za razgovor točno usklađeni s očekivanim ispravnim odgovorima u odnosu na sve njegove odgovore. Odziv, s vrijednošću od 75%, pokazuje sposobnost agenta za razgovor da prepozna i odgovori na relevantna pitanja iz ukupnog skupa mogućih relevantnih odgovora.

Iako su ovi rezultati pružaju dobar uvid u trenutnu učinkovitost agenta za razgovor, postoji značajan prostor za poboljšanje. Na primjer, niska vrijednost odziva sugerira da agent za razgovor može propustiti određene relevantne upite, što može biti rezultat ograničenog broja podataka na kojima je treniran ili nesavršenosti u modeliranju razumijevanja konteksta pitanja. Daljnjim treniranjem agenta za razgovor, uključujući proširenje skupa podataka s dodatnim primjerima i kontinuiranim prilagodbama modela, može se značajno poboljšati točnost i odziv. Ovo uključuje i implementaciju naprednijih tehnika obrade prirodnog jezika i integraciju novih metoda za prepoznavanje varijacija u korisničkim upitima.

Također, ručno unošenje specifičnih odgovora na često postavljana pitanja može unaprijediti preciznost i ukupnu performansu agenta za razgovor. Ova praksa omogućava agentu za razgovor da se bolje pripremi za specifične upite i osigura konzistentne i točne odgovore, što u konačnici poboljšava korisničko iskustvo. Ove mjere, uz primjenu povratnih informacija korisnika, omogućuju agentu za razgovor razvijanje većeg kapaciteta za pravilno prepoznavanje i odgovor na korisničke upite, čime se povećava ukupno korisničko iskustvo i učinkovitost agenta za razgovor u stvarnim situacijama. Redovita evaluacija i prilagodba agenta za razgovor ključni su za njegovu dugoročnu uspješnost i prilagodbu promjenjivim potrebama korisnika.

## **13. Usporedba agenta za razgovor s već postojećim agentima za razgovor na tržištu**

Za usporedbu dvaju edukativnih agenata za razgovor, odabran je *Kialo Edu* (<https://www.kialo-edu.com/>) te agent iz ovoga završnog rada, odnosno agent za razgovor kojemu je svrha pomoći pri upisu na Filozofski fakultet Sveučilišta u Zagrebu. Tu je važno je razmotriti nekoliko ključnih aspekata: funkcionalnost, korisničko iskustvo, prilagodljivost, dostupnost sadržaja i tehnološku osnovu. Oba agenta za razgovor imaju za cilj poboljšati proces učenja, no njihovi pristupi i specifične funkcionalnosti mogu značajno varirati.

### **13.1. Funkcionalnost**

*Kialo Edu* je platforma koja se specijalizira za poticanje kritičkog razmišljanja i rasprave među učenicima. Kialo Edu omogućava učiteljima i studentima da organiziraju i sudjeluju u strukturiranim debatama. Platforma koristi interaktivne karte argumenata koje pomažu korisnicima vizualizirati složenost problema te razumjeti različite perspektive. Kialo Edu je posebno koristan za predmete poput društvenih znanosti, filozofije i politike, gdje je analiza argumenata ključna.

Izrađeni agent za razgovor nudi sveobuhvatno rješenje u obliku agenta za razgovor koje može biti prilagođeno različitim obrazovnim potrebama. Ovaj agent za razgovor, izrađen u svrhu ovog završnog rada, može odgovarati na pitanja, pružati povratne informacije. Fokus je na pružanju personaliziranog iskustva učenja putem umjetne inteligencije koja može adaptirati odgovore na temelju korisnikovih interakcija.

### **13.2. Korisničko iskustvo**

*Kialo Edu* pruža vizualno privlačno sučelje koje olakšava navigaciju kroz različite argumente i pomaže korisnicima pratiti logiku rasprave. Korisničko sučelje je intuitivno i omogućava lako upravljanje i pregled različitih tema. Međutim, zahtijeva određeno predznanje o strukturi argumenata i može biti manje pristupačan za mlađe učenike ili one bez iskustva u debatama.

## Name / Thesis / Language

[×](#)

**Name** [?](#) [Browse Topic Library](#)

120

**Thesis** [?](#)

500 [🔗](#)

**Language**

 ▼



[Skip rest and create](#)

Slika 12. Korisničko sučelje *Kialo Edu* agenta za razgovor

Izrađeni agent za razgovor se fokusira na jednostavnost i pristupačnost. Sučelje je dizajnirano tako da bude jednostavno za korištenje i intuitivno, omogućujući korisnicima da brzo dođu do traženih informacija. Interakcija putem sučelja za razgovor je prirodna i može biti prilagođena različitim razinama učenika, od osnovne škole do visokog obrazovanja (iako je dakako naš slučaj vezan za upis na fakultet, odnosno tematika, no samo je sučelje jednostavne prirode). Prednost ovog rješenja je u tome što može biti integrirano u postojeće obrazovne platforme.

### 13.3. Prilagodljivost

*Kialo Edu* nudi visok stupanj prilagodljivosti kada je riječ o kreiranju i vođenju rasprava. Učitelji mogu definirati teme, dodavati resurse i strukturirati rasprave na način koji najbolje odgovara njihovim obrazovnim ciljevima. Ipak, prilagodba je uglavnom ograničena na kontekst debatanja i argumentacije.

Platforma za izradu agenta za razgovor nudi širu prilagodljivost, omogućujući korisnicima da i sami kreiraju agente za razgovor za razne obrazovne scenarije. Od jednostavnog odgovaranja na često postavljana pitanja, preko pružanja interaktivnih kvizova, do složenijih scenarija učenja gdje agent djeluje kao virtualni tutor. Prilagodba može biti izvedena putem jednostavnog sučelja za izradu agenta, bez potrebe za programerskim znanjem, što je čini pristupačnom širem krugu korisnika.

#### **13.4. Dostupnost sadržaja**

Kialo Edu se oslanja na sadržaj kreiran od strane korisnika i zajednice, što može rezultirati bogatim izvorom resursa i raznolikih perspektiva. Međutim, kvaliteta i točnost sadržaja može varirati ovisno o tome tko je kreirao argumente i kako su moderirani.

Kod izrade agenta za razgovor, moguće je koristiti pristup različitim izvorima sadržaja, uključujući vanjske baze znanja, obrazovne resurse i vlastite materijale korisnika te dakako i razne poveznice na vanjske sadržaje. Ova fleksibilnost omogućava učiteljima da integriraju specifične resurse relevantne za njihov kurikulum, osiguravajući konzistentnu kvalitetu i relevantnost sadržaja.

#### **13.5. Tehnološka osnova**

Kialo Edu koristi tehnologiju koja je specifično dizajnirana za podršku debatama i argumentaciji. Platforma koristi algoritme za mapiranje argumenata i praćenje tokova rasprave, što je čini vrlo efikasnom za analizu složenih problema.

Platforme za izradu agenata najčešće koriste napredne algoritme umjetne inteligencije i strojno učenje za razumijevanje i generiranje odgovora. Ova tehnologija omogućava personalizaciju i adaptaciju učenja, što može značajno poboljšati učinak edukacije.

#### **13.6. Usporedba agenata**

Oba edukativna agenta za razgovor, nude značajne prednosti za obrazovni sektor, no njihovi fokusi i pristupi se razlikuju. Kialo Edu je alat za razvijanje kritičkog mišljenja kroz strukturirane debate, dok je platforma za izradu agenta za razgovor fleksibilno rješenje koje može biti prilagođeno širokom spektru obrazovnih potreba. Ovisno o specifičnim ciljevima i

potrebama obrazovne institucije ili pojedinca, jedan od ova dva agenta može pružiti optimalno rješenje za unapređenje procesa učenja, zavisno o tome koje su potrebe pojedince.

## 14. Zaključak

U ovom završnom radu istraženi su ključni aspekti teorije agenata za razgovor, s posebnim fokusom na izradu agenta za razgovor za domenu obrazovanja te na procjenu njegove učinkovitosti. Analizirani su osnovni koncepti agenata za razgovor, uključujući njihovu povijest, tehnologije na kojima se temelji njihova izgradnja te različite pristupe kao što su agenti za razgovor temeljeni na pravilima i oni bazirani na umjetnoj inteligenciji. Razumijevanje ovih aspekata pružilo je čvrstu osnovu za razvoj i evaluaciju agenta za razgovor, omogućujući bolje uvid u njihove mogućnosti i ograničenja.

Izgradnja agenta za razgovor specifičnog za obrazovnu domenu uključivala je upotrebu naprednih alata koji omogućuje kreiranje agenta za razgovor temeljenog na vlastitim podacima i javno dostupnim izvorima informacija. Agent za razgovor prilagođen je potrebama korisnika kroz jednostavno sučelje, prilagodbu početnih poruka, ponuđenih pitanja i vizualnih elemenata, što doprinosi boljem korisničkom iskustvu. Kao rezultat, agent za razgovor je razvijen s ciljem da poboljša pristup informacijama o upisu na studije Filozofskog fakulteta u Zagrebu, pružajući studentima i budućim studentima relevantne i pravovremene informacije.

Evaluacija agenta za razgovor provedena je kroz test koristeći 15 pitanja, pri čemu su rezultati pokazali točnost od 73,33%, preciznost od 90% i odziv od 75%. Ovi rezultati ukazuju na dobar početni uspjeh agenta za razgovor, ali i na prostor za daljnje poboljšanje. Preciznost i odziv su ključni indikatori uspješnosti agenta za razgovor u pružanju relevantnih odgovora, dok točnost reflektira ukupnu učinkovitost agenta za razgovor u točnom odgovaranju na upite. Identificirani su i konkretni koraci za poboljšanje, uključujući dodatno treniranje modela i ručno unošenje odgovora na često postavljana pitanja, što će unaprijediti njegovu funkcionalnost i korisničko iskustvo.

U zaključku, rad je istaknuo važnost temeljitog razumijevanja teorije agenata za razgovor te primjenu znanja u stvarnom svijetu kroz razvoj specifičnog agenta za razgovor za obrazovanje. Rezultati evaluacije pružaju vrijedne uvide u performanse agenta za razgovor te usmjeravaju buduće napore na unapređenje sustava. Ovaj rad ne samo da doprinosi razumijevanju tehničkih aspekata agenta za razgovor, već i demonstrira kako se moderna tehnologija može primijeniti u obrazovanju za poboljšanje pristupa informacijama i podrške studentima. Nastavak istraživanja i unapređenja agenta za razgovor obećava dodatne

inovacije i optimizacije koje će imati dugoročan utjecaj na korisničko iskustvo i učinkovitost obrazovnih sustava.



## 15. Literatura

1. Abedin, A. F., Al Mamun, A. I., Nowrin, R. J., Chakrabarty, A., Mostakim, M. and Naskar, S. K. (2021) A Deep Learning Approach to Integrate Human-Level Understanding in a Chatbot [online]. Dostupno na: <https://arxiv.org/abs/2201.02735> [29.7.2024.]
2. Adamopoulou, E., Moussiades, L. (2020) Machine Learning with Applications .
3. Alogayli, A., Abdelhafez, H. (2023) Intelligent Chatbot for Admission in Higher Education, International Journal of Information and Education Technology, Vol. 13, No. 9
4. Altamimi, I., Altamimi, A., Alhumimidi, A. S., Altamimi, A., Temsah, M. H. (2023) Artificial Intelligence (AI) Chatbots in Medicine: A Supplement, Not a Substitute
5. Artificial Solutions (2019) Types of Chatbot Technology. Dostupno na: [16.7.2024.]
6. Bahak H., Taheri F., Zojaji Z., Kazemi A. (2023) Evaluating ChatGPT as a Question Answering System: A Comprehensive Analysis and Comparison with Existing Models
7. Bansal, H., & Khan, R. (2018). A review paper on human computer interaction. International Journal of Advanced Research in Computer Science and Software Engineering [online]. Dostupno na: [https://www.researchgate.net/publication/325534924\\_A\\_Review\\_Paper\\_on\\_Human\\_Computer\\_Interaction](https://www.researchgate.net/publication/325534924_A_Review_Paper_on_Human_Computer_Interaction) [5.7.2024.]
8. Cahn, J. (2017) CHATBOT: Architecture, Design, Development. University of Pennsylvania, School of Engineering and Applied Science
9. Ceccarini C., Prandi, C. (2019) Tourism for all: a mobile application to assist visually impaired users in enjoying tourist services. In: 2019 16th IEEE annual consumer communications and networking conference (CCNC), pp 1–6. IEEE [online]. Dostupno na: <https://ieeexplore.ieee.org/document/8651848> [12.7.2024.]
10. Ciesla, R. (2024) The Book of Chatbots: From ELIZA to ChatGPT [online]. Dostupno na: <https://books.google.com/books?id=dvntEAAAQBAJ&printsec=copyright> [3.7.2024.]
11. Coghlan, S., Leins, K., Sheldrick, S., Gooding, M. C. P., D'Alfonso, S. (2023) To chat or bot to chat: Ethical issues with using chatbots in mental health
12. Colby, K.M. (1981) Modeling a paranoid mind. Behavioral and Brain Sciences.
13. Contact Fusion (2023). Dostupno na: <https://www.contactfusion.co.uk/> [22.7.2024.]

14. Costa, P. C. F. da. (2018). Conversing with personal digital assistants: On gender and artificial intelligence. *Journal of Science and Technology of the Arts* [online]. Dostupno na: <http://dx.doi.org/10.7559/citarj.v10i3.563> [7.7.2024.]
15. Definition of chatbot in english by Lexico Dictionaries (2019). Lexico Dictionaries Dostupno na: <https://www.lexico.com/en/definition/chatbot> [6.7.2024.]
16. Dictionary.com (2024). Dostupno na: <https://www.dictionary.com/browse/chatbot> [6.7.2024.]
17. Dovedan, Z., Seljan, S. i Vučković, K. (2002) Strojno prevođenje kao pomoć u procesu komunikacije. *Informatologia* 35 (4),str.283-91
18. Dunder, I., Pavlovski, M. i Seljan, S. (2020) Computational Analysis of a Literary Work in the Context of Its Spatiality. *Trends and Innovations in Information Systems and Technologies* (pp.252-261)
19. Dunder, I., Seljan, S. i Odak, M. (2023) Data Acquisition and Corpus Creation for Phishing Detection, 46th MIPRO ICT and Electronics Convention (MIPRO), Opatija, Croatia, 2023, pp. 533-538, doi: 10.23919/MIPRO57284.2023.10159904
20. Dunder, I., Seljan, S. i Pavlovski, M. (2020) Automatic Machine Translation of Poetry and a Low-Resource Language Pair. 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2020, pp. 1034-1039, doi: 10.23919/MIPRO48935.2020.9245342
21. Eisenstein, J. (2018) *Natural Language Processing*
22. Fan, H., Han, B., Gao, W. (2022) (Im)Balanced customer-oriented behaviors and AI chatbots' Efficiency–Flexibility performance: The moderating role of customers' rational choices
23. H.-Y. Shum, X. He i D. Li (2018) From Eliza to XiaoIce: challenges and opportunities with social chatbots [online]. Dostupno na: <https://arxiv.org/abs/1801.01957> [8.7.2024]
24. Haristiani, N. (2019). Artificial intelligence (AI) chatbot as language learning medium: An inquiry. *Journal of Physics: Conference Series*, 1387, Article 012020 [online]. Dostupno na: <https://iopscience.iop.org/article/10.1088/1742-6596/1387/1/012020> [9.7.2024.]
25. Hien, H. T., Cuong, P.-N., Nam, L. N. H., Nhung, H. L. T. K., & Thang, L. D. (2018). Intelligent assistants in higher-education environments: The FIT-EBot, a chatbot for administrative and learning support. In *Proceedings of the ninth international symposium on information and communication technology* (pp. 69–76). New York,

- NY, USA: ACM [online]. Dostupno na: <http://dx.doi.org/10.1145/3287921.3287937> [9.7.2024.]
26. Hrvatska enciklopedija (2024). Dostupno na: <https://www.enciklopedija.hr/> [16.7.2024.]
27. Jáuregui-Velarde, R., Hernández Celis, D., Molina-Velarde, P., Yactayo-Arias, C., Andrade-Arenas, L. (2024) Evaluation of a Chatbot Powered by ChatGPT for the Preliminary Diagnosis of Dengue [online]. Dostupno na: [https://www.researchgate.net/publication/377368581\\_Evaluation\\_of\\_a\\_Chatbot\\_Powered\\_by\\_ChatGPT\\_for\\_the\\_Preliminary\\_Diagnosis\\_of\\_Dengue](https://www.researchgate.net/publication/377368581_Evaluation_of_a_Chatbot_Powered_by_ChatGPT_for_the_Preliminary_Diagnosis_of_Dengue) [25.7.2024.]
28. Kasaraneni, C. K. (2020) Understanding NLP Pipeline. Dostupno na: <https://medium.com/analytics-vidhya/understanding-nlp-pipeline-9af8cba78a56> [14.7.2024.]
29. Khoa, B. T. (2021) The Impact of Chatbots on the Relationship between Integrated Marketing Communication and Online Purchasing Behavior in The Frontier Market
30. Kognitivna znanost, Hrvatska enciklopedija (2024) Dostupno na: <https://www.enciklopedija.hr/clanak/kognitivna-znanost> [22.7.2024.]
31. Laumann, F., (2023) Why accuracy matters in your chatbots. Dostupno na: <https://medium.com/neuralspace/why-accuracy-matters-in-your-chatbots-fa538316bf48#:~:text=An%20accurate%20NLU%20model%20enables,users'%20queries%20are%20answered%20correctly> [5.8.2024.]
32. Mackworth A.K., Poole D.L. (2010) Artificial Intelligence: Foundations of Computational Agents, Cambridge
33. Marrara, S., Pejic Bach, M. i Seljan, S., Topalovic, A. (2019) FinTech and SMEs: The Italian Case. FinTech as a Disruptive Technology for Financial Institutions, 42-60
34. Mechkaroska, D., Domazet, E., Feta, A., Shikoska, U. R. (2024) Architectural scalability of Conversational Chatbot The case of ChatGPT
35. Medium (2023) AI in Customer Service: The Rise of Chatbots and Virtual Assistants, Medium. Dostupno na: <https://medium.com/@galiniostech/ai-in-customer-service-the-rise-of-chatbots-and-virtual-assistants-55e08f7ca7e4> [3.8.2024.]
36. Medium (2023) Chatbot Marketing: An Overview. Dostupno na: <https://medium.com/hireawriter/chatbot-marketing-an-overview-839e328b34cf> [23.7.2024.]

37. Medium (2024) Enhancing Customer Service with AI Chatbots. Dostupno na: <https://medium.com/@smicos.chatbot/enhancing-customer-service-with-ai-chatbots-665670ed8668> [23.7.2024.]
38. Medium (2024) Understanding the NLP Pipeline: A Comprehensive Guide. Dostupno na: [https://medium.com/@asjad\\_ali/understanding-the-nlp-pipeline-a-comprehensive-guide-828b2b3cd4e2](https://medium.com/@asjad_ali/understanding-the-nlp-pipeline-a-comprehensive-guide-828b2b3cd4e2) [23.7.2024.]
39. Nuruzzaman, M., & Hussain, O. K. (2018). A survey on chatbot implementation in customer service industry through deep neural networks. In 2018 IEEE 15th international conference on e-business engineering (pp. 54–61) [online]. Dostupno na: <https://ieeexplore.ieee.org/document/8592630> [10.7.2024.]
40. Pejić Bach, M., Krstić, Ž. i Seljan, S., Turulja, L. (2019) Text Mining for Big Data Analysis in Financial Sector: A Literature Review
41. Sustainability <https://doi.org/10.3390/su11051277> 11 (5), 1-27
42. Peras, D. (2018) Chatbot evaluation metrics: review paper // Economic and Social Development (Book of Proceedings) Veselica, Rozana ; Dukić, Gordana ; Hammes, Khalid (ur.). Zagreb: VADEA, 2018. str. 89-97
43. ProjectPro (2024) Machine Learning (ML) vs NLP - What's the Difference?. Dostupno na: <https://www.projectpro.io/article/machine-learning-vs-nlp/493> [26.7.2024.]
44. Pytel R. (2024) How can you evaluate your chatbot's answers? SoftwareMill. Dostupno na: <https://softwaremill.com/how-can-you-evaluate-your-chatbots-answers/> [6.8.2024.]
45. ResearchGate (2024). Dostupno na: <https://www.researchgate.net/> [26.7.2024.]
46. Russell, S. and Norvig, P. (1996) Artificial Intelligence: A Modern Approach, Englewood Cliffs, NJ, Prentice-Hall. Dostupno na: [https://people.engr.tamu.edu/guni/csce421/files/AI\\_Russell\\_Norvig.pdf](https://people.engr.tamu.edu/guni/csce421/files/AI_Russell_Norvig.pdf) [12.7.2024.]
47. Seljan, S., Baretić, M. i Kučiš, V. (2014) Information retrieval and terminology extraction in online resources for patients with diabetes. Collegium antropologicum 38 (2), 705-710
48. Seljan, S., Dunđer, I. i Gašpar A. (2013) From digitisation process to terminological digital resources
49. Seljan, S., Dunđer, I. i Pavlovski, M. (2020) Human Quality Evaluation of Machine-Translated Poetry. 43rd International Convention on Information, Communication and

- Electronic Technology (MIPRO), Opatija, Croatia, 2020, pp. 1040-1045, doi: 10.23919/MIPRO48935.2020.9245436
50. Seljan S. (2003) Leksičko-funkcionalna gramatika hrvatskoga jezika: teorijski praktični modeli. Doktorska disertacija. Sveučilište u Zagrebu
51. Stanford Encyclopedia of Philosophy (2024). Dostupno na: <https://plato.stanford.edu/> [19.7.2024.]
52. Stillings, N.A. (1995) Cognitive Science: An Introduction [online]. Dostupno na: [https://books.google.hr/books/about/Cognitive\\_Science.html?id=wCRonP7EgDkC&redir\\_esc=y](https://books.google.hr/books/about/Cognitive_Science.html?id=wCRonP7EgDkC&redir_esc=y) [19.7.2024.]
53. The top 12 healthcare chatbots. (2020). The Medical Futurist website. Dostupno na: <https://medicalfuturist.com/top-10-health-chatbots/> [10.7.2024.]
54. Turing, A. M. (1950). Computing machinery and intelligence. Mind, LIX(236), 433–460. [online]. Dostupno na: <https://academic.oup.com/mind/article/LIX/236/433/986238> [8.7.2024.]
55. Ulrich Gnewuch, U., Morana, S., Adam, M. T. P., Maedche, A. (2022) Opposing Effects of Response Time in Human–Chatbot Interaction
56. Valerjev, P. (2006) Uloga umjetne inteligencije u istraživanju uma: povijest i perspektiva
57. Valerjev, P. (2006) Uloga umjetne inteligencije u istraživanju uma: povijest i perspektiva. Mozak i um - Trajni izazov čovjeku. Zagreb, Institut društvenih znanosti Ivo Pilar, str. 105-122. [online]. Dostupno na: <https://www.croris.hr/crosbi/publikacija/prilog-knjiga/45706> [14.7.2024.]
58. Wei, X., Chu, X., Geng, J., Wang Y., Wang, P., Wang, H., Wang, C., Lei, L. (2024) Societal impacts of chatbot and mitigation strategies for negative impacts: A large-scale qualitative survey of ChatGPT users

## 16. Popis slika

Slika 1. Interakcija Elize i čovjeka (Weizenbaum, 1966).....	11
Slika 2. Prikaz poveznica iz mrežne adrese web stranice Filozofskog fakulteta u Zagrebu....	32
Slika 3. Korisničko sučelje agenta za razgovor .....	33
Slika 4. Grafički prikaz rezultata ankete za prvu tvrdnju .....	35
Slika 5. Grafički prikaz rezultata ankete za drugu tvrdnju .....	36
Slika 6. Grafički prikaz rezultata ankete za treću tvrdnju.....	37
Slika 7. Grafički prikaz rezultata ankete za četvrtu tvrdnju.....	37
Slika 8. Grafički prikaz rezultata ankete za petu tvrdnju.....	38
Slika 9. Grafički prikaz rezultata ankete za šestu tvrdnju.....	39
Slika 10. Grafički prikaz rezultata ankete za sedmu tvrdnju .....	39
Slika 11. Grafički prikaz rezultata ankete za osmu tvrdnju .....	40
Slika 12. Korisničko sučelje <i>Kialo Edu</i> agenta za razgovor .....	46

## **17. Popis tablica**

Tablica 1: Tablični prikaz rezultata ankete .....	41
Tablica 2: Matrica konfuzije izrađenog agenta za razgovor .....	42
Tablica 3: Pitanja postavljena agentu za razgovor .....	44

## 18. Prilozi

Poveznica na agent za razgovor: <https://www.chatbase.co/chatbot-iframe/1S-LluChpAUXPLUfYevQD>

Poveznica na anketu: <https://forms.gle/EohR9G17oqJR2Mfq8>

Poveznica na Kialo Edu: <https://www.kialo-edu.com/>



# Primjena agenata za razgovor u obrazovanju

## Sažetak

Ovaj završni rad bavi se definiranjem te primjenom agenata za razgovor u obrazovanju. U radu je objašnjena teorijska osnova agenata za razgovor, njihova povijest i razvoj, kao i suvremene tehnologije koje omogućuju njihovu primjenu gdje je naglasak stavljen na pojmove obradu prirodnog jezika (NLP) i strojno učenje, ključne komponente za funkcioniranje agenata za razgovor, kao i kriterije za evaluaciju. Evaluacija provedenog istraživanja korisničkog iskustva pruža vrijedne uvide u performanse agenta za razgovor i usmjerava buduće napore na unapređenje sustava. Na kraju, rad naglašava važnost razumijevanja teorije agenata za razgovor i praktične primjene tog znanja u stvarnom svijetu, te kako moderna tehnologija može poboljšati pristup informacijama i podršku studentima, dok metode samog istraživanja uključuju analizu već postojećih agenata za razgovor na temelju sustava za rangiranje te vlastita subjektivna procjena korisnosti i preciznosti agenata. Na temelju rezultata slijedi uspoređivanje razvijenog agenta za razgovor i ostalih postojećih agenata.

**Ključne riječi:** agent za razgovor, evaluacija, obrazovanje, korisničko iskustvo

# **The application of chatbots in education**

## **Summary**

This thesis addresses the definition and application of chatbots in education. The theoretical foundation of chatbots, their history and development, as well as contemporary technologies enabling their application are explained where an emphasis is placed on concepts such as natural language processing (NLP) and machine learning, key components for the functioning of chatbots, as well as criteria for their evaluation. The evaluation of user experience research provides valuable insights into the performance of chatbots and guides future efforts to improve the system. In conclusion, the thesis highlights the importance of understanding the theory of chatbots and the practical application of this knowledge in the real world, and how modern technology can enhance access to information and student support, whilst the research methods include the analysis of existing chatbots based on ranking systems and a subjective assessment of the utility and accuracy of the agents. Based on the results, a comparison between the developed chatbots and other existing chatbots will follow.

**Key words:** chatbot, evaluation, education, user experience