

## ČETBOT – DIGITALNI ASISTENT ZASNOVAN NA WEAVER PLATFORMI CHATBOT - DIGITAL ASSISTANT BASED ON WEAVER PLATFORM

Rade Matić, Miloš Kabiljo, Milen Janjić, Miodrag Živković, Milan Čabarkapa

**REZIME:** Zadovoljstvo krajnjih korisnika direktno je srazmerno kvalitetu izvršenih usluga. Istraživanja pokazuju da su krajnji korisnici u velikoj meri svoj fokus usmerili sa društvenih mreža ka platformama (aplikacijama) za razmenu tekstualnih poruka (*Viber, FB messenger* itd). Veštačka inteligencija, zajedno sa pojmom digitalne transformacije, predstavlja veliki korak u modernizaciji poslovanja svake industrije. Sada sa velikom sigurnošću možemo reći da se teži ka korisničkim interfejsima zasnovanim na obradi prirodnih jezika. U korak sa svetskim trendovima, a u cilju razvoja rešenja koje donosi inovativnost i konkurentnost, na tržište je uveden četbot odnosno softver koji oponaša ljudsku komunikaciju radi olakšavanja svakodnevnih poslova. Časkanje se preselilo iz prostora za konverzaciju u prostor aplikacija gde četbot pomaže u izvršenju velikog broja različitih zadataka. U ovom radu su date osnove četbota i referentna arhitektura koja podržava različite kanale za interakcije korisnika sa kompanijama. Ukratko je dat opis komunikacione platforme *Weaver* koja objedinjuje sve kanale komunikacije na koje se četbot može integrisati, a zatim i jedan primer njene implementacije.

**KLJUČNE REČI:** četbot, četbot arhitektura, platforma za slanje poruka

**ABSTRACT:** End-user satisfaction is directly proportionate quality of services provided. Research shows that end users have largely focused from social networking to text messaging platforms (*Viber, FB messenger, etc.*). Artificial intelligence, along with the notion of digital transformation, is a major step in modernizing the business of each industry. Now we can say with great certainty that user interfaces strives to be based on natural language processing. In line with global trends, and in the development of solutions that bring innovation and competitiveness in the market, we have introduced chatbot, software that imitate human communication to facilitate day-to-day operations. The chat moves from a conversation space into a space of applications where the chatbot support in completing a large number of different tasks. In this paper, we explain chatbot basics and the reference architecture that supports different channels for interactions between clients and company. It briefly describes the Weaver communication platform, which integrates all communication channels on which chatbot can be integrated, followed by one example of its implementation.

**KEY WORDS:** chatbot, chatbot architecture, messaging platform

### 1. UVOD

Obrada velikog broja zahteva i poslovnih procesa u svakodnevnom radu velikih i srednjih sistema je postao standard a to je dovelo do ograničenja u ljudskim resursima jer je sve manja mogućnost obrade istih u kratko vreme. Većina takvih poslovnih procesa koji su zahtevali ljudske resurse je komunikacija sa klijentima za potrebe poboljšanja i učvršćivanje pozicije na tržištu. S druge strane, decenijama unazad istraživači pokušavaju da poboljšaju svoje veštine u dizajniranju korisničkog interfejsa u interakciji između čoveka i računara.

Tehnologija je ponudila rešenje u virtualnom saradniku koji se bazirao na veštačkoj inteligenciji a sve u svrhu izbegavanja ograničenja ljudskog faktora i poboljšanja interakcije između čoveka i računara. Korisnički interfejs je zasnovan na prirodnom jeziku dok se interakcija sa digitalnim sistemima ne odvija preko skrola, pomeranja ili kliktanja dugmića, već preko tekstova napisanih prirodnim jezikom. Pravi primer platforme koji koristi obradu prirodnog jezika i veštačku inteligenciju kako bi komunicirao sa klijentima, odgovarao na sva njihova pitanja i rešavao njihove probleme je četbot (engl. chatbot). Jedna od osnovnih zadataka četbota je kako će razumeti korisnikov zahtev bez obzira na jezik ili rečenicu koju je napisao ili izgovorio. Na osnovu obrade rečenice, veštačke inteligencije i nizom algoritama četbot dolazi do adekvatnog odgovora koji prosleđuje korisniku. Četbot se koristi u različitim vrstama aplikacija kao što su usluge vezane za e-trgovinu, tehničku podršku, učenje itd. Tržište četbotova raste godišnjom

stopom od 24.3% [1]. Rast je neminovan, jer polje mašinskog učenja konstantno raste. Smatra se da će do 2020. godine u SAD-u, u industrijama koje pružaju korisničke usluge, 80% biznisa imati svog četbota [2]. Iako je istorija razvoja veštačke inteligencije počela još 1950. godine sa poznatim Turingovim testom [3], koji se ogleda u sposobnosti mašine da ispolji ponašanje isto ili bolje od čoveka, industrija četbota je i dalje u ranoj fazi ali njihov razvoj i primena u zadnje dve godine doživeo je veliku ekspanziju. Gartner, koji je svetski konsultantski i istraživački lider u svetu, predviđa da će četbotovi do 2020. godine da zauzmu 25% svih operacija vezanih za usluge i podršku klijentima. Oni takođe pišu: „Četbot otvara nove kanale za pružanje javnih usluga (poboljšavajući ih). Rukovodioci moraju brzo utvrditi ulogu ovih kanala, prilagoditi svoje strategije za pružanje digitalnih usluga i proširiti svoju digitalnu vladinu platformu kako bi iskoristili ove nove mogućnosti. Zadovoljstvo klijenata se povećava tri puta a troškovi se trostruko smanjuju kad se četbot koristi u organizaciji. Organizacije koje su uvele četbota izveštavaju o smanjenju poziva, časkanja i/ili upita putem e-pošte do 70 odsto.”[4]. Očekuje se da će četbotovi do 2022. godine smanjiti troškove poslovanja za 8 milijardi dolara [5].

U ovom radu su nakon uvoda dati osnovni koncepti četbota, taksonomija modela razvoja a zatim i opšte prihvaćena arhitektura. U nastavku je prikazan kratak opis implementirane platforme za razvoj četbota zasnovan na veštačkoj inteligenciji. Konačno, zaključak je izveden u Poglavlju 5.

## 2. ČETBOT

Četbot je softverski kod koji služi kao korisnički interfejs na prirodnom jeziku za potrebe dobijanja podataka ili obavljanja usluga [6]. Četbot predstavlja računarski program koji simulira ljudsku konverzaciju ili časkanje, koristeći ili ne koristeći veštačku inteligenciju, i učestvuje u dijalogu sa čovekom koristeći prirodni jezik. Četbot može razumeti pisani tekst, glasovne poruke i u velikoj meri može interpretirati njihovo značenje. Osim toga četbot omogućava jednostavne višestruke dijaloge, upravlja tokom konverzacije i pri tome treba uvek da uzima odgovarajući kontekst iz istorije konverzacije. Danas četbot može da izvršava mnoge funkcije mobilnih aplikacija ili veb sajta, sve u okviru konverzacije putem aplikacija za razmene poruka, a bez potrebe da korisnik mora da instalira ili preuzme nove aplikacije. Sadašnje veliko interesovanje za četbotovima je podstaknuto razvojem dve nove tehnologije:

- Tehnologija razmene poruka se brzo proširila u poslednjih nekoliko godina i postala jedna od najčešće korišćenih usluga pametnog telefona. Ove platforme servisa za razmenu poruka (*Viber, Facebook Messenger, Skype*, itd.) sada uključuju funkcionalnosti kao što su informisanje, podrška, plaćanje, naručivanje i rezervacija, što bi inače zahtevalo odvojenu aplikaciju ili veb sajt.
- Veštačka inteligencija je takođe napravila značajne korake napred u poslednjih nekoliko godina. Tehnika kao što je mašinsko učenje iskoristila je ogromne količine podataka i jeftinu procesorsku moć kako bi dramatično poboljšali kvalitet razumevanja i donošenja odluka.

Bez obzira da li ga nazivate četbot, digitalni pomoćnik ili konverzacioni interfejs, osnovni koncept je isti: postići odgovarajući rezultat razgovarajući sa mašinom putem prirodnog jezika. Neki četbotovi su bez veštačke inteligencije, dok drugi imaju tako složenu veštačku inteligenciju da uspeju da dešifruju nameru korisnika. Dizajniranje efektivnog četbota znači razumevanje njihove osnovne arhitekture i koja vrsta inteligencije je potrebna. Najpre ćemo da pokažemo opštu taksonomiju modela koji preovlađuju u razvoju četbota a zatim i referentnu arhitekturu. Uopšteno gledajući postoje dve vrste modela:

- Modeli zasnovani na vraćanju (engl. Retrieval Based Models): Koriste repozitorijum unapred definisanih odgovora ili koriste neku vrstu heuristike da bi odabrali odgovarajući odgovor baziran na upitu i kontekstu. Heuristika može biti jednostavna tako što je zasnovana na pravilima ili može biti kompleksna jer koristi grupu klasifikatora mašinskog učenja. Mogu pružiti pouzdanije i gramatički ispravnije odgovore. Lakše ih je učiti jer traže manje podataka ali nisu u mogućnosti da odgovore na pitanja izvan svoje baze znanja i ne mogu kreirati novi tekst i učiti sami na osnovu prethodnih konverzacija. Ova vrsta četbota zahteva ručnu pomoć od operatora za netipična pitanja.
- Modeli generisanja ili generativni modeli (engl. Generative models): Ne naslanjaju se na predefinisane odgovore. Oni generišu nove odgovore od starta jer mogu da odgovore na dvosmislena pitanja. Ovi modeli se baziraju na tehnikama mašinskog učenja i mogu da vode duge i inteligentne razgovore. Ovakvi četbotovi kreiraju odgovore od početka koristeći obradu prirodnih jezika (engl. Na-

tural Language Processing - NLP). Ovi četbotovi postaju pametniji s vremenom i sposobni su da samostalno uče iz prethodnih pitanja, odgovora, konverzacija i na taj način postaju „pametniji“. Međutim, teško ih je trenirati (učiti) jer zahtevaju veliku količinu podataka.

- Kombinovani modeli: Kombinacija prethodna dva modela daju nam upravo ovakve vrste četbotova. Ovi modeli se, slično kao modeli generisanja, baziraju na tehnikama mašinskog učenja tako što za potrebe razumevanja pitanja ili zahteva od strane korisnika, koriste NLP. A za potrebe izvršavanja i odgovaranja koriste predefinisane akcije i odgovore iz svoje baze znanja.

Iz prethodno opisana dva modela možemo primetiti da se četbotovi naslanjaju na specifične konverzacione kontekste. Ovi konteksti konverzacije se mogu podeliti na otvorene i zatvorene domene. U otvorenom domenu kontekst konverzacije nije definisan. Ne postoji dobro definisan cilj ili namera. Primer ovoga može biti razgovor sa *Google* asistentom. Možemo ga pitati šta god želimo, a on će pokušati da pronađe odgovor i da ga prosledi. Većina aplikacija za četbot spadaju u zatvorene domene. Svaki četbot ima poseban kontekst i specijalizovan je za odgovore na pitanja vezana za taj domen. Tehnička podrška korisnicima ili asistenti za kupovinu su primeri problema zatvorenog domena.

### 2.1. Prednosti uvođenja četbota

S obzirom da aplikacije za razmenu poruka postaju najčešći način koji ljudi koriste u komuniciranju, svaki biznis bi treba da ima strategiju za angažovanje ljudi koji rade sa aplikacijama za razmenu poruka, a četbotovi su jedini skalabilni način za to. Iz ovoga proizilazi da će svaki biznis na kraju imati svog četbota. Četbotovi i njihovi odgovori će biti brži od veb stranica i mobilnih aplikacija jer neće biti instalacija, kliktanja i pretrage po četbotu kao na aplikacijama. Mi smo već u svetu u kojem ne možemo da prepoznamo konverzacione partnere. Neki od njih su ljudi, neki su četbotovi. Međutim, sve više nas to možda neće ni interesovati, niti ćemo se brinuti oko toga, jer je bitno da smo zadovoljni sa konverzacijom ili akcijom koju smo hteli da izvedemo. Trenutno, četbotovi su najviše dizajnirani i razvijeni za mobilne aplikacije za slanje poruka [7]. Treba istaći sledeće razloge zbog kojih je potrebno uključiti četbotove u strategiju razvoja u budućnosti:

- Brža usluga jer se klijenti rešavaju tradicionalne komunikacije/kanala putem telefona i e-pošte. Umesto da klijenti ostanu na telefonskom čekanju, institucije mogu koristiti četbot kako bi se odmah mogli baviti jednostavnim pozivima i zahtevima.
- Obezbeđuje se visoka dostupnost klijenata 7/24/365 dana. Klijenti žele usluge izvan normalnog radnog vremena od 9 do 17 sati. Četbot je dostupan 24 sata dnevno tokom cele godine i odgovara u roku od maksimalno 5 sekundi.
- Mogućnost istovremenog opsluživanja više hiljada klijenata i korisnika: Četbotovi mogu pomoći da brže i efikasnije služe klijentima jer mogu da premoste jaz između velikog broja klijenata koji traže informacije i ograničenih resursa koje treba da pružaju takvu vrstu pomoći.

- Relaksiranje od zadataka koji se ponavljaju i smanjen broj istih poziva i mejlova do 70%.
- Veća dostupnost, vidljivost usluga i efikasnije poslovanje. Npr. 80% poziva ka IT je bio resetovanje šifre a to sad radi četbot.
- Jednostavniji način komunikacije i lakše korišćenje. S obzirom da četbotovi koriste kanale kao što su npr. *Viber* i *FB messenger*, nema potrebe za učenjem njihovog korišćenja – „Ako znate kako da ih koristite, onda znate kako da koristite četbot“. Dok u slučaju novih aplikacija učenje može da predstavlja problem i oduzima vreme.
- Dolazi do povećanja baze klijenata/korisnika, jer je lakše privući korisnike na kanal (*Viber*, *FB messenger*) nego da ih nateramo ili nagovaramo da instaliraju neku aplikaciju. Samo u Srbiji ima preko 3 miliona korisnika Vibera i broj stalno raste.
- Pored same komunikacije, prikupljaće se i podaci o korisniku koji se kasnije mogu koristiti za personalizovane ponude, analitiku ili integraciju sa bilo kojim javnim ili privatnim institucijama (banka, osiguravajuća društva itd). Na osnovu ove baze možemo dobiti mišljenja javnosti putem raznih anketa ili odlučivanja.
- Internet je veoma dostupan i jeftiniji od telefonskog poziva a samim tim i četbot. A pretrage na sajtu su veoma frustrirajuće.
- Obezbeđena je multijezičnost.
- Ljudski resursi mogu da se preusmere/prekvalifikuju na odgovornije i kreativnije poslove jer dolazi do rasterećenja kontakta centra. Ljudi često misle da će uvođenjem veštačke inteligencije izgubiti radna mesta a u stvari dobijaju kreativnije poslove kao npr. poslove učenja četbota odnosno veštačke inteligencije.
- Četbotovi su omni-kanali. Klijenti mogu dobiti podršku kroz više kanala a lako se prebacuju sa kanala na kanal (npr. sa *Viber* na *FB messenger*) jer dobiju isti odgovor bez obzira na kanal.
- Definicija scenarija se definiše na jednom mestu bez obzira na različite kanale, a prebacivanje sa kanala na drugi kanal se veoma lako realizuje zahvaljujući fleksibilnosti odnosno povezanosti osnovnih delova arhitekture četbota. Ovo daje veliku uštedu u održavanju u okviru IT-a jer se *front end* uopšte ne održava. Tačnije, ne brine kompanija o održavanju *Viber*, *FB messenger* itd.
- Jaz između čoveka i mašine je sveden na minimum jer kad se promeni dizajn sajta ili aplikacije ljudi imaju otpor a ovde zauvek ostaje interfejs na koji su navikli.
- Osim obrade i razumevanja prirodnog jezika treba napomenuti da se pri prijemu poruke vrši korekcija grešaka u pisanju u slučaju da postoje. Korekcija grešaka je takođe zasnovana na veštačkoj inteligenciji.
- Spremnost za integraciju sa agentima kontakt centra i objedinjavanje sa njihovim servisima u skladu sa mogućnostima kontakta centra odnosno ako poseduju otvoren API. Ovo znači da kada četbot dođe do poruke koju ne razume on je prosleđuje API kontaktu centru. Konverzacija se prosleđuje od API kontakta centra ka operateru koji nastavlja razgovor. Moguće je da neki scenario zahteva agenta po definiciji scenarija. Ovaj zahtev zavisi od softvera u kontaktu centru tj. da li ima otvoren API.

## 2.2. Primeri četbota

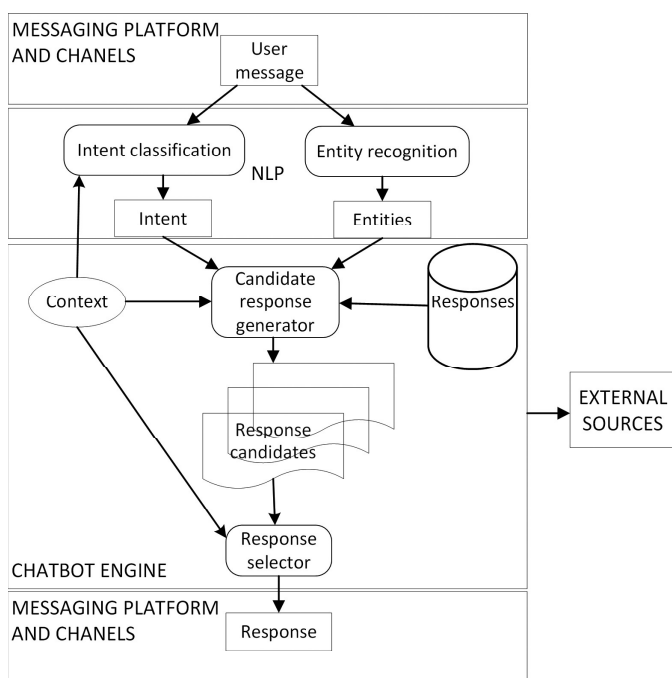
Verovatno prvi primer četbota je bio program koji se zove *ELIZA* [8] razvijen od strane grupe profesora sa MIT univerziteta. *ELIZA* je mogla da oponaša ljudski razgovor tako što je na pitanje korisnika pokušavala da upari skriptovane odgovore tj. da vrši jednostavno podudaranje uzoraka. Nakon toga su u drugoj polovini dvadesetog veka razvijani mnogi drugi popularni četbotovi. Tako na primer, A.L.I.C.E. (engl. Artificial Linguistic Internet Computer Entity) je bio bot za obradu prirodnog jezika [9]. Ovaj bot je mogao da primeni pravila heuristike koja su se podudarala sa pravilima ljudskog unosa, odnosno imao je mogućnost konverzacije.

Četbotovi su sada ugrađeni u popularne u digitalne asistente kao što su: Siri (Apple), Assistant (Google), Alexa (Amazon), Cortana (Microsoft), M (Facebook) itd. Pored ovih više poznatih ovde ćemo nabrojati još neke manje poznate ali ništa manje bitne. *Swelly* je četbot koji omogućava dobijanje prijateljskih mišljenja na različite teme, čime se pomaže pri donošenju određenih odluka. Švedska banka SEB kreirala je *Aidu*, kao pratnju postojećem internom četbotu *Ameliji*, namenjenog zaposlenima. U toku prve 3 nedelje korišćenja *Amelije*, 700 zaposlenih razgovaralo je preko 4000 puta, tokom kojih je ona rešila veliki broj ukazanih problema bez odlaganja. Njeni glavni zadaci su rad sa zaposlenima i pomoć internoj IT podršci. *Aida* trenutno rešava oko 13% svih pitanja IT podrške. *XioIce* je Majkrosoftov najuspešniji četbot i jedan od tehnički najsofisticiranijih botova. Za 3 meseca dostigla je pola milijardi razgovora. U proseku, svaka osoba koja je dodala *XioIce* pričala je sa njom preko 60 puta mesečno. Ovaj četbot je prošao Turingov test za 10 minuta, što znači da ljudi koji pričaju u prvih 10 minuta nisu znali da je u pitanju četbot. *Instalocate* je četbot koji pomaže prilikom odlaganja letova i rešavanja pitanja kompenzacije. Uzima 25% udela kada putnik dobije kompenzaciju za odložen, poništen ili prebukiran let. Za 6 meseci ostvareno je preko 1.1 miliona razmenjenih poruka i ovaj četbot predstavlja jedan od finansijski uspešnih četbotova. Još jedno od rešenja za čekiranje i rešavanje zahteva vezanih za aerodrome je *Carla*. Ovaj četbot omogućava korisnicima da se čekiraju na aerodromima i dobiju svoje karte. Odličan primer implementacije četbota u Srbiji je REA [10], digitalna asistentkinja *Raiffeisen* banke zasnovana na platformi *Weaver* kompanije SAGA [11]. Preko ovog četbota klijent može da postavi bilo koje pitanje u vezi banke, svog stanja na računima, da zakaže sastanak i sl. U poglavlju 4 dat je detaljniji opis platforme koja je korišćena u toku realizacije ovog projekta.

## 3. OPŠTA REFERENTNA ARHITEKTURA

Referentni model četbota [12] je prikazan na slici 1. Četbot može da ima odvojene module za generisanje odgovora i izbor odgovora. Obrada poruke počinje od razumevanja korisnikove poruke (engl. user message). Modul klasifikacije namere identifikuje nameru (engl. intent classification) korisničke poruke. Obično je to izbor jedne od brojnih unapred definisanih namera, mada sofisticiraniji četbotovi mogu identifikovati više namere iz jedne poruke. Za potrebe izbora namere, klasifikacija namere može koristiti kontekstne informacije, kao što su namere prethodnih poruka, korisnički profil i navike. Modul za

prepoznavanje entiteta (engl. entity recognition) izvlači struktuirane bitove informacija iz poruke. Na primer vremenski četbot može izvući lokaciju i datum. Generator kandidata odgovora (engl. candidate response generator) radi sve proračune specifične za domen radi obrade zahteva korisnika. On može da koristi različite algoritme, da pozove nekoliko spoljašnjih API-a, ili čak zamoli čoveka da pomogne u stvaranju odgovora. Rezultat ovih kalkulacija je lista kandidata za odgovor (engl. response candidates). Svi ovi odgovori trebali bi biti tačni u skladu sa domenskom logikom. Generator odgovora mora da koristi kontekst konverzacije, kao i namere i entitete izvučene iz poslednje korisničke poruke. Selektor odgovora (engl. response selector) samo ocenjuje sve kandidate za odgovor i bira odgovor koji bi trebalo najbolje da odgovori na korisnikovu poruku. Arhitektura mora da omogući i integraciju sa internim i spoljnim eksternim izvorima podataka (engl. External sources).



Slika 1. Opšta referentna arhitektura

#### 4. WEAVER PLATFORMA

*Weaver* platforma za četbot predstavlja *enterprise* rešenje iz oblasti softverskih proizvoda koji u sebi poseduju veštačku inteligenciju. Zasnovana je na potpunom razumevanju i obradi prirodnog jezika. Ovaj segment veštačke inteligencije omogućava analizu unetih poruka (frazu) od strane korisnika, ne radeći to analizom i traženjem ključnih reči nego apsolutnim razumevanjem prirodnog jezika i konteksta cele fraze. Platforma poseduje administraciju korisnika koji mogu da koriste alat, kao i monitorisanje i osnovne izveštaje vezane za kanale komunikacije. Admin alat poseduje vizuelni dizajner scenarija za definisanje scenarija koje četbot podržava. Tok konverzacije mapira sve potencijalne pravce u kojima se konverzacija može odvijati, uz mnoštvo različitih grana za sve mogućnosti. Administracija je vizuelizovana i uz trening je lako savladiva. Platforma je u mogućnosti da iz poruke otklanja greške,

a sve to primenom najsavremenijih tehnoloških dostignuća iz oblasti veštačke inteligencije. Važno je napomenuti da se sama konverzacija između četbot i čoveka odvija interaktivno. Jedinstven mehanizam toka same konverzacije kalkuliše šta je sledeća akcija sa kojom treba da se obrati korisniku i zavisi od celog toka konverzacije. Platforma poseduje mogućnost integracije sa sistemima unutar i izvan organizacije čime sam četbot dobija dodatnu vrednost u smislu mogućnosti da korisniku pruži personalizovanu informaciju ili mogućnost da koristi transakcioni ili bilo koji drugi servis. Jednostavan tekst: "Prebaci 2000 dinara na račun moga brata" uz dodatne korake za autorizaciju je najjednostavniji način da se izvrši veoma kompleksna akcija. Četbot iz teksta razume šta je namera korisnika. Sve mogućnosti *Weaver* platforme korisnici mogu da koriste kroz bilo koji kanal: *Viber*, *Facebook messenger*, *Web chat* (postojeći veb sajt) itd. Jednostavan način komunikacije i korišćenja servisa ruši barijere između komplikovanog interfejsa aplikacija i ljudi kojima je servis neophodan. Važno je napomenuti da je srpski jezik u potpunosti podržan, kao i mnogi drugi, ali srpski jezik predstavlja poseban izazov i dostignuće zbog svoje složenosti i malog broja korisnika na globalnom nivou. Sve prethodno napisano, postavlja *Weaver* platformu ravnopravnom sa svetskim liderima iz ove oblasti. *Weaver* u osnovi objedinjuje kanale komunikacije u jednu tačku, konkretno sve kanale koje je tehnički moguće integrisati. Komunikacija po svim kanalima se čuva u jednom *repository*-u. Obrade, analize i izveštavanja mogu da se vrše na jednom mestu. U osnovi je nezavisan od ulaznih i izlaznih kanala komunikacije. Adapteri zaduženi za prijem komunikacije prilagođavaju protokol komunikacije i standardizuju ga u željenu formu.

##### 4.1. Logička arhitektura

Softverska arhitektura i okviri naprednog razmišljanja postaju zaista važan deo savremenih tehnoloških dostignuća [13, 14]. Logička arhitektura ovde predloženog rada je prikazana na slici 2. Kao što vidimo arhitektura se sastoji od četiri osnovna gradivna bloka:

- Korisnik;
- Platforme i kanali preko kojih korisnici komuniciraju;
- Četbot frejmwork;
- Eksterni servisi.

Četbot arhitektura se sastoji od niza komponenti koje čine ovaj sistem skalabilnim. Ceo sistem se zasniva na mikro servisima koji su međusobno povezani a u sebi sadrže pojedine module. **BOT framework** je centralna i glavna logička komponenta u arhitekturi u kojoj se nalazi celokupna logika za procesiranje dospelih poruka, rad sa kontekstima, logička odlučivanja, integracija sa predefinisanim scenarijima, odgovorima i razna druga potrebna pravila. Ovaj frejmwork između ostalog služi za povezivanje, izgradnju, testiranje i razmeštanje inteligentnog četbota. Ona u sebi ima nekoliko bitnih komponenti kao što su: *bot engine*, *bot designer*, *language processing*, *conversation engine* itd. Svaka dospelu poruka se posebno obrađuje i prolazi kroz niz faza. Svaki korisnik ima svoju čet sesiju (konverzaciju) u kojoj se vodi evidencija o svim porukama između njega i četbota. Ovde se takođe vide sve pokrenute, suspendovane i završene instance scenarija koje prate svoju definiciju scenarija. Svaka

čet sesija predstavlja jedan ili više scenarija između korisnika i četbota koji u sebi sadrži kontekste koji utiču na tok trenutnog i budućeg četa. **Bot engine** je kompleksna komponenta koja čini centralni deo BOT frejmworka. Ona mora da zna tačan redosled izvršavanja, prati izvršenje akcija, prati greške, šalje obavještenja itd. Svaka akcija u zavisnosti od konfiguracije prolazi kroz dibag i info log kako bi se mogao pratiti tok ovako kompleksne komponente. U ovoj komponenti se izvršavaju razne provjere pre početka procesiranja poruke, kao što su npr. provjera da li je korisnik aktivan, da li je platforma sa kojeg dolazi aktivna, da li su *custom* poruke ili *custom* validacije ispunjene i druge definisane provjere. U **API messenger connector**-u se odvija validacija i mapiranje poruka koje dolaze od strane platforme za slanje poruka ili drugih vrsta kanala. Ona ima za cilj da validira, pripremi i mapira poruku koju će da prosledi dalje u sledeću komponentu u zavisnosti od tipa poruke koju je dobio. Svaki kanal ima svoj set mapiranja i pravila koja su definisana u bazi koja konektor primenjuje kako bi popunio objekat sa svim potrebnim podacima. Životni ciklus poruke počinje na konektoru i završava se u njemu jer on ima ulogu da primi poruku i da iz četbot sistema vrati poruku nazad na kanal sa koga je poruka poslata. Cilj je da četbot prikuplja komunikaciju sa ovim integriranim kanalima, da njima pruži definisani set usluga, automatsko odgovaranje na poruke, pružanje servisa, ili redirekciju na agenta u slučaju potrebe. U slučaju da kompanije ne poseduje kontakt centar, konverzacija se šalje korisnicima koji su zaduženi za odgovaranje na neki od standardnih kanala npr. email.

**Bot designer** predstavlja vizuelni alat opremljen SDK-om za odgovarajuću implementacionu platformu. Preko njega je omogućeno direktno učenje NLP da bi se izbegla zavisnost vendorovog NLP interfejsa. Takođe, kroz njega se omogućava vizuelno generisanje scenarija tj. njihovog toka. Tok konverzacije mapira sve potencijalne pravce u kojima se diskusija može odvijati, uz mnoštvo različitih grana za sve mogućnosti. Tok konverzacije je odgovoran za predviđanje svih mogućih ulaznih poruka kao i

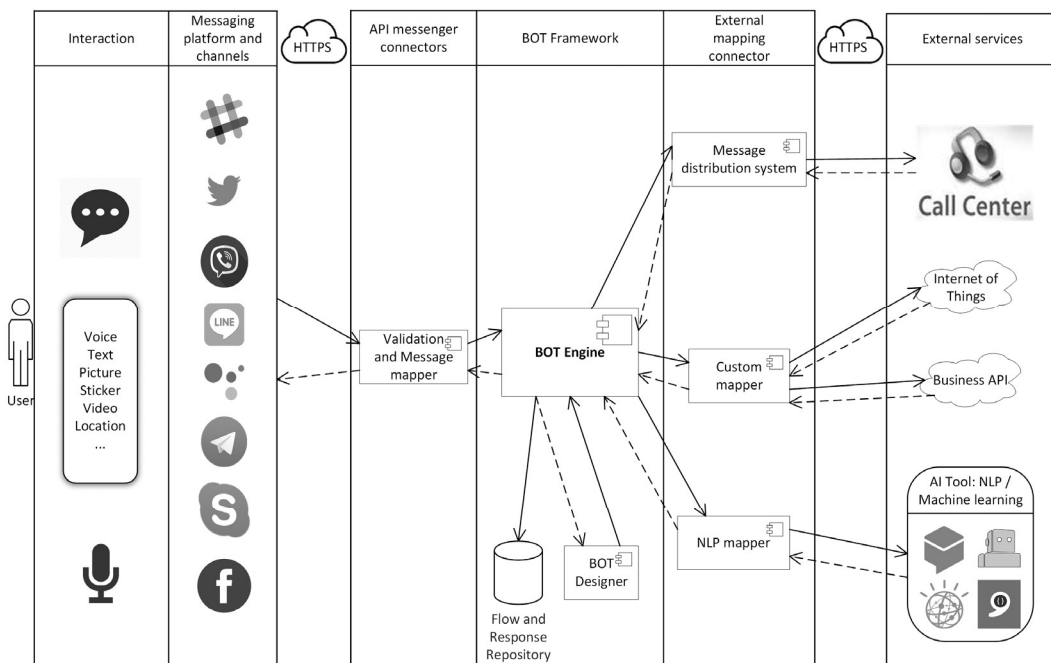
reakcije četbota, a diskusija može ići u različite grane ovog toka. Svi podaci neophodni za upravljanjem toka scenarija (engl. scenario flow) kao i odgovora zasnovanih na osnovu konteksta, identifikovanih namera i entiteta se nalaze u bazi toka i odgovora (engl. **Flow and Response Repository**). U ovoj baza se čuvaju svi predefinisani scenariji sa svojim akcijama i stanjima, odgovori sa potrebnim segmentacijama, konverzacije, korisnici i logovi. Osim prethodno navedenog, u bazi se snimaju podaci aktivnosti korisnika, bez obzira da li je četbot uspeo ili nije da obradi njihova pitanja. **Eksterni konektori** služe za povezivanje sa eksternim servisima kao što su: različiti inteligentni uređaji, poslovni sistemi, postojeći kontakt centri, drugi servisi za obradu prirodnih jezika i ostali eksterni servisi. Ovi konektori imaju predefinisana pravila mapiranja ponaosob za svaki konkretan tip i implementaciju. Predložena četbot arhitektura omogućava da se svaka nova platforma može lako integrisati i da četbot može konzistentno da radi kroz sve platforme. BOT engine prihvata poruku preko REST API, šalje poruku NLP-u preko našeg *mapper*-a a zatim hendluje dobijene i parsirane namere i entitete.

#### 4.2. Implementacija

Opisana arhitektura i implementacija četbota zasnovana je na veb API servisnoj arhitekturi i mikroservisima. Veb API je između ostalog izabran jer:

- je servis zasnovan na HTTP protokolu koji omogućava RESTful servis i JSON objekte;
- može biti hostovan unutar aplikacije ili IIS;
- može da ga koristi bilo koji klijent koji razume JSON ili XML;
- je jednostavna arhitektura.

Rajfajzenova elektronska asistentkinja – REA je projekat četbota razvijen na *Weaver* platformi kao kombinovani model razvoja u saradnji sa kompanijom SAGA [10]. REA spada u za-



Slika 2. Logička arhitektura

tvorene domene jer pruža trenutnim i budućim klijentima banke lakše i brže informacije vezane za usluge banke. Sa REA-om je lako razgovarati jer imate osećaj da razgovarate kao sa prijateljem koji čak može da izvršava jednostavne zadatke kao što je npr. zakazivanje sastanka, informacije o najbližem bankomatu itd. Poruke postaju jednostavan alat za ankete, komunikaciju i lak pristup informacijama, kao što su zadaci, rokovi, događaji i druge korisne informacije. REA, kao što je prikazano na Slici 2, zavisi od bot frejmvorka koji prima ljudske poruke sa različitih kanala kao što su *Viber* i *Facebook Messenger*. Bot frejmvork obavlja veliki broj bitnih zadataka a zatim šalje poruku NLP. NLP parsira tekst i na osnovu tehnika mašinskog učenja poređi tj. proverava podudarnost, a zatim kao rezultat vraća nameru (engl. intent) i njene entitete (engl. entity) sa odgovarajućim procentom tačnosti. Generisani rezultat se šalje bot frejmvorku koji razrešava nameru, entitete, kontekst i šalje odgovor korisniku u adekvatnom formatu preko odgovarajućih kanala.

## 5. ZAKLJUČAK

Za mnoge korisnike, platforme za razmenu poruka postaju primarni način na koji provode vreme na svojim uređajima. Podstaknuti obećanjem inteligentnih digitalnih asistenata da će uvek biti na raspolaganju za brzo i dosledno rešavanje zahteva klijenata, četbotovi postaju sve više interesantni. Četbot je odličan put za sve oragnizacije da ostanu i postanu povezani sa klijentima i da pažljivije slušaju i izvršavaju sve njihove potrebe od dobijanja jednostavnih informacija do komplikovanih transakcionih procesa. Tehnološki giganti su masovno počeli da integrišu četbota unutar platformi za razmenu poruka kako bi pomogli korisniku da jednostavno dobije odgovor ili da mu se ispune odgovarajuće osnovne usluge. Četbotovi su kao nove aplikacije, a digitalni pomoćnici su meta aplikacije ili kao novi pregledači. U ovom radu upoznali smo se bliže sa četbotovima, predstavili njihove tipove i generičku arhitekturu za četbot frejmvork. Predložena platforma i arhitektura je fleksibilna, sama po sebi skalabilna i podržava različite medije za interakcije korisnika. Istraživački i praktičan rad se nastavlja u nekoliko pravaca. Treba da se obezbedi podršku za glasovne poruke kao i dodatna merenja performansi kako bi se otkrila uska grla i identifikovale tačke konsolidacije. A u praktičnom smislu u toku je nekoliko dodatnih implementacija u oblasti obrazovanja i ostalih državnih organa. Osnovna prednost koju treba naglasiti nije samo dobijanje četbota već i uvođenje nove digitalne strategije koristeći elemente veštačke inteligencije kroz uštedu, automatizaciju i optimizaciju.

## 6. LITERATURA

- [1] <https://www.grandviewresearch.com/press-release/global-chatbot-market>, avgust 2019.
- [2] <https://www.businessinsider.com/80-of-businesses-want-chatbots-by-2020-2016-12>, avgust 2019.
- [3] Turing, A. M. "Computing Machinery and Intelligence," *Mind*, 59, 433-460, 1950.
- [4] <https://www.gartner.com/en/newsroom/press-releases/2018-02-19-gartner-says-25-percent-of-customer-service-operations-will-use-virtual-customer-assistants-by-2020>, avgust 2019.
- [5] <https://www.forbes.com/sites/forbestechcouncil/2018/08/29/chatbots-a-powerful-weapon-in-the-business-arsenal/#65ce5a64960>, avgust 2019.
- [6] R. Dale, „The return of the chatbots“. *Natural Language Engineering* 22, 5 (2016), 811–817.
- [7] Følstad A., Bae Brandtzæg P., „Chatbots and the new world of HCI“, *Interactions*, v.24 n.4, July–August 2017.
- [8] J. Weizenbaum, „ELIZA—a computer program for the study of natural language communication between man and machine“, *Communications of the ACM*, v.9 n.1, p.36–45, Jan. 1966.
- [9] A. I. Alice, "Foundation. Free A.L.I.C.E. AIML Set", (2015) March 21, <http://code.google.com/p/aiml-en-us-foundation-alice/>, datum pristupa juli 2019.
- [10] <https://www.raiffeisenbank.rs/digitalne-usluge/rea-elektroniska-asistentkinja/>, datum pristupa avgust 2019.
- [11] <https://weaverbot.ai/>, datum pristupa avgust 2019.
- [12] M. Karuppiah, S. Kumari, X. Li, F. Wu, A.K. Das, M. K. Khan, R.Saravanan and S. Basu, "A dynamic id-based generic framework for anonymous authentication scheme for roaming service in global mobility networks", *Wireless Personal Communications*, Vol.93, No.2, (2016), pp.383–407.
- [13] S. Cohen, D. Dori, U. de Haan, "A Software System Development Life Cycle Model for Improved Stakeholders' Communication and Collaboration", *International Journal of Computers, Communications & Control*, vol. 5, no. 1, 2010.
- [14] CH. Lin., W. Wolf, X. Koutsoukos, S. Neema, J. Sandeep,: "System and Software Architectures of Distributed Smart Cameras", *ACM Transactions on Embedded Computing Systems (TECS)*. Vol. 9. No. 4, 2010.



**Rade Matić**

**Kontakt:** rade.matic@bpa.edu.rs

**Oblast interesovanja:** informacioni sistemi, razvoj chatbot aplikacija



**Miloš Kabiljo**

**Kontakt:** milos.kabiljo@bpa.edu.rs

**Oblast interesovanja:** programiranje, razvoj inteligentnih sistema



**Milen Janjić**

**Kontakt:** milen.janic@saga.rs

**Oblast interesovanja:** veštačka inteligencija, informacioni sistemi



**Miodrag Živković**

**Kontakt:** mzivkovic@singidunum.ac.rs

**Oblast interesovanja:** Java, veštačka inteligencija



**Milan Čabarkapa**

**Kontakt:** cabmilan@etf.bg.ac.rs

**Oblast interesovanja:** Razvoj mobilnih servisa, bežične mreže