

# Prepoznavanje glasa na primjeru Google asistenta

---

**Korica, Andrej**

**Undergraduate thesis / Završni rad**

**2022**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Pula / Sveučilište Jurja Dobrile u Puli**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:137:670014>

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

*Download date / Datum preuzimanja:* **2025-02-03**



*Repository / Repozitorij:*

[Digital Repository Juraj Dobrila University of Pula](#)



Sveučilište Jurja Dobrile u Puli  
Fakultet informatike u Puli

**Andrej Korica**

## **Prepoznavanje glasa na primjeru Google asistenta**

Završni rad

Pula, rujan, 2022.

Sveučilište Jurja Dobrile u Puli  
Fakultet informatike u Puli

**Andrej Korica**

## **Prepoznavanje glasa na primjeru Google asistenta**

Završni rad

**JMBAG: 0303092327, redovni student**

**Studijski smjer: Informatika**

**Kolegij: Osnove IKT**

**Znanstveno područje: Društvene znanosti**

**Znanstveno polje: Informacijske i komunikacijske tehnologije**

**Znanstvena grana: Informacijski sustavi i informatologija**

**Mentor: doc. dr. sc. Snježana Babić**

Pula, rujan 2022.



### IZJAVA O AKADEMSKOJ ČESTITOSTI

Ja, dolje potpisani ANDREJ KORICA, kandidat za prvostupnika INFORMATIKE ovime izjavljujem da je ovaj Završni rad rezultat isključivo mogega vlastitog rada, da se temelji na mojim istraživanjima te da se oslanja na objavljenu literaturu kao što to pokazuju korištene bilješke i bibliografija. Izjavljujem da niti jedan dio Završnog rada nije napisan na nedozvoljeni način, odnosno da je prepisan iz kojega necitiranog rada, te da ikoji dio rada krši bilo čija autorska prava. Izjavljujem, također, da nijedan dio rada nije iskorišten za koji drugi rad pri bilo kojoj drugoj visokoškolskoj, znanstvenoj ili radnoj ustanovi.

Student

Korica A.

U Puli, 22. 09. 2022.g.



### IZJAVA O KORIŠTENJU AUTORSKOG DJELA

Ja, ANDREJ KORICA dajem odobrenje Sveučilištu Jurja Dobrile u Puli, kao nositelju prava iskorištavanja, da moj Završni rad pod nazivom PREPOZNAVANJE GLASA NA PRIMJERU GOOGLE ASISTENTA

koristi na način da gore navedeno autorsko djelo, kao cjeloviti tekst trajno objavi u javnoj internetskoj bazi Sveučilišne knjižnice Sveučilišta Jurja Dobrile u Puli te kopira u javnu internetsku bazu završnih radova Nacionalne i sveučilišne knjižnice (stavljanje na raspolaganje javnosti), sve u skladu s Zakonom o autorskom pravu i drugim srodnim pravima i dobrom akademskom praksom, a radi promicanja otvorenoga, slobodnoga pristupa znanstvenim informacijama.

Za korištenje autorskog djela na gore navedeni način ne potražujem naknadu.

U Puli, 22.09.2022.g.

Potpis

Andrej Korica

# Sadržaj

1. Uvod .....	1
2. Umjetna inteligencija.....	3
3. Tehnologije umjetne inteligencije za učitavanje glasa .....	5
3.1. Strojno učenje.....	5
3.1.1. Nadzirano strojno učenje .....	6
3.1.2. Strojno učenje bez nadzora .....	7
3.1.3. Polunadzirano strojno učenje .....	8
3.1.4. Strojno učenje s pojačanjem .....	9
3.2. Duboko učenje.....	10
3.3. Obrada prirodnog jezika.....	12
4. Vrste umjetne inteligencije .....	14
4.1. Umjetna uska inteligencija .....	14
4.2. Umjetna opća inteligencija.....	14
4.3. Umjetna super inteligencija.....	15
4.4. Slaba umjetna inteligencija .....	16
4.5. Jaka umjetna inteligencija .....	17
5. Glasovni asistenti i prepoznavanje glasa .....	18
5.1. Pojam glasovnog asistenta .....	20
5.2. Najpopularniji jezici za izradu softvera za prepoznavanje govora .....	20
5.3. Računalna lingvistika .....	21
6. Prepoznavanje glasa na primjeru Google asistenta .....	24
6.1. Digitalna pomagala i Google asistent.....	24
6.2. Obrada funkcionalnosti Google asistenta.....	30
6.3. Prednosti, nedostaci i perspektive Google asistenta .....	31
7. Zaključak .....	34
Literatura.....	36
Popis slika .....	45
Sažetak .....	46
Summary .....	47

# 1. Uvod

Tema ovog rada je Google asistent, odnosno umjetna inteligencija za prepoznavanja glasa pomoću koje se izvršavaju glasovne naredbe. Google asistent pomoću tehnologije umjetne inteligencije odnosno primjenom postupaka obrade prirodnog jezika i tehnologije strojnog učenja razumijeva što korisnik govori te predlaže rješenja ili djeluje na temelju izrečenih jezičnih naredbi, odnosno aktivira se glasovnom naredbom koja proizvodi određene radnje u ime korisnika i pruža relevantne informacije. Sustav je fokusiran na učenje iz iskustva i u skladu s iskustvom kasnije donosi odluke, što mu omogućuje da izvršava pojedinačne radne zadatke bez prethodnog programiranja. U članku bibalex Basma Fawzy (2020) kaže da strojno učenje znači davanje strojevima pravila i ostavljajući im mogućnosti da uče po sistemu pokušaja i pogrešaka bez unaprijed određenih uputa u svojim programskim kodovima. Pojednostavljeno, Google asistent je u stvari digitalno pomagalo za komunikaciju između čovjeka i stroja.

Postupak obrade prirodnog jezika i umjetne inteligencije su osnova za izradu računalnih programa za razumijevanje ljudskog jezika odnosno njegovo prevođenje (npr. Google translate). Ova tehnologija se temelji na jezikoslovlju (lingvistici), odnosno znanosti koja proučava ljudski jezik i jezične komunikacije, s tim da je računalna lingvistika spoj i kombinacija informatike i lingvistike odnosno znanstvena disciplina koja se bavi razvojem programske podrške koja omogućava računalnu obradu prirodnog jezika (Wikipedia, 2021). Alexander S. Gillis i Corinne Bernstein (2022) kažu da je računalna lingvistika znanost koja proučava jezik iz računalne perspektive, a računalo objašnjava određene lingvističke ili psiholingvističke pojavnosti. Primjenom ovih tehnologija stroj na temelju naučenog dijaloga počinje konstruirati jasne rečenice poput ljudskog bića, pa stoga odgovori Google asistenta nisu prethodno programirani, nego su generirani upotrebom određenih pravila lingvistike koja mu pomažu u slaganju rečenica.

Važnost Google asistenta je to što on predstavlja ključnu točku u Googleovoj strategiji za budućnost i vjerojatno će određivati način na koji će korisnici stupati u interakciju s gotovo svim Googleovim osnovnim proizvodima. Google asistent posredno utječe na potrošače koji su dio Googleovog ekosustava i tako potiče na odabir i na nabavu upravo njegovih, a ne konkurentskih proizvoda umjetne inteligencije. U članku TechRepublic (2021) se navodi da je Google asistent dostupan na nekoliko vrhunskih pametnih uređaja i to Android uređaja, iPhonea, Google Home

uređaja, Android TV-a, nekih pametnih satova i uređaja koji koriste Wear OS, određenih automobila s integriranim Android Auto, kao i ovlaštenih Googleovih partnera.

Prednosti Google home tehnologije su da je skoro svima cjenovno pristupačna, komunikacija je jednostavna i daje prilično brzo točne i sveobuhvatne informacije. Na web stranici Tech Avatar (2021) je navedeno da su glavni nedostaci Google homea problemi s kompatibilnosti pojedinih uređaja, problemi s utroškom i napajanjem električne energije, otežana komunikacija na većim razmacima i na lokacijama koje imaju lošu internetsku mrežu, ali i oni subjektivne prirode kao što je strah od gubitka privatnosti ili napada hakera. Kod izbora uređaja s glasovnom aktivacijom moramo točno znati za što nam je on potreban i koliko smo spremni potrošiti za njegovu nabavku.

Najpoznatiji programski jezici koji se koriste za izradu softvera za prepoznavanje glasa su C#, Java, PHP i Perl, iako je Google asistent napisan pretežno u C++ zbog poboljšanja performansi (Gupta, 2017).

Za obavljanje svojih funkcija, Google asistent je poveznica između svih Googleovih osnovnih softverskih i hardverskih proizvoda. Danas postoji trend da se korisnici pametnih kuća koriste upravo pametnim uređajima kao što je Google home koji im omogućuje upravljanje čvorištima kućne automatizacije i drugih internet of things uređaja. Google asistent se pokreće putem ključnog glasovnog poziva „*Hey Google*“ ili „*okay Google*“ (Akinbi & BERRY, 2020).

Uređaji se savršeno povezuju pomoću mrežnog wifi sustava ili bluetootha na Google asistenta, a putem pametnog zvučnika ili mobitela se pokreću osobne naredbe za autonomno i inteligentno korištenje i upravljanje, kao što su npr. uključivanje i isključivanje prepoznavanje putem senzora i dr. kako bi se zadovoljile navike i potrebe korisnika, omogućilo praćenje uštede energije, vremena i novca, postigla ugodnost (Turner, 2018b). Google asistent se kroz vrijeme kontinuirano povećava i poboljšava kvalitetu svojih usluga, što rezultira sve većim utjecajem umjetne inteligencije na život ljudi.



## 2. Umjetna inteligencija

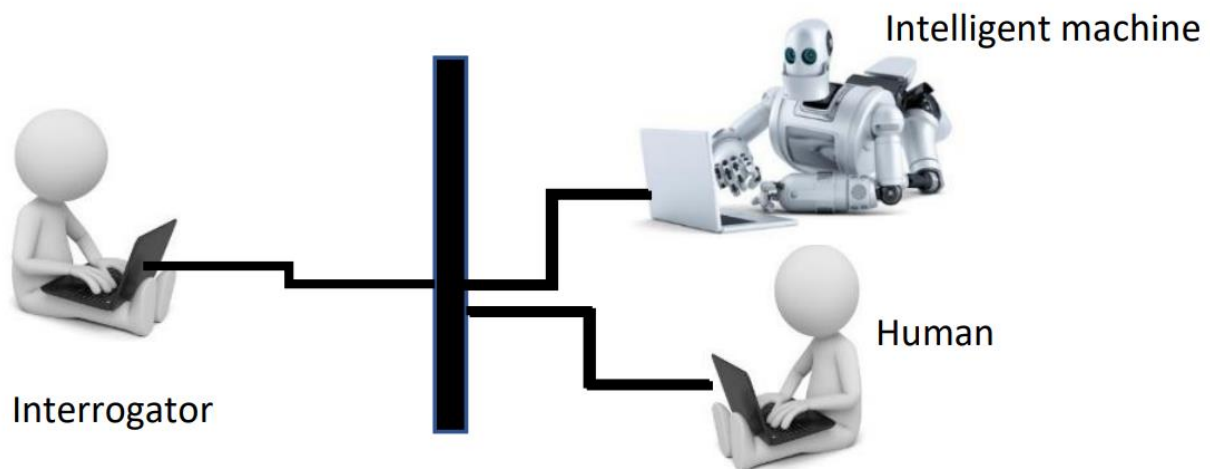
Umjetna inteligencija predstavlja sposobnost nekog uređaja da oponaša ljudske aktivnosti kao što su zaključivanje, učenje, planiranje i kreiranje (europarl, 2021). Pomoću umjetne inteligencije tehnički sustavi prepoznaju svoje okruženja i rješavaju probleme. Umjetna inteligencija obuhvaća široki spektar više grana računalnih znanosti i smatra se ključnom za digitalnu transformaciju društva. Sustavi umjetne inteligencije u određenoj mjeri mogu se osposobiti da putem analize prethodnih situacija i autonomnim ponašanjem postignu zadani cilj. Potrebno je pronaći tehnološku zamjenu za ljudska ograničenja i slabosti u vidu “umjetne inteligencije koja je namijenjena da bi kompjuteri mogli poput ljudi obavljati ljudske zadatke, ali puno većom brzinom pa se zato i zovu umjetna inteligencija“ (Poole & Goebel, 1998).

Umjetnu inteligenciju koriste programski proizvodi poput glasovnih asistenata, programa za obradu slika, razni internetski pretraživači, softveri za prepoznavanje lica i glasa. Umjetna inteligencija se može implementirati u svrhu kontrole različite opreme kao što su npr. roboti, automobili, dronovi itd. Nadalje, u svakodnevnom životu umjetnu inteligenciju susrećemo kod internet kupovine, raznih oglašavanja, pretraživanja interneta, jezičnog prevođenja, pametnih kuća i gradova, infrastrukture, automobilske industrije koja koristi autonomnu tehnologiju, kibersigurnosti, u zdravstvenoj zaštiti, borbi protiv lažnih vijesti itd. (europarl, 2021).

Umjetna inteligencija nalazi svoju primjenu u medicini za dijagnostiku i analizu zdravstvenih podataka, u prometu za poboljšanje sigurnosti i učinkovitosti prometa te omogućavanju autonomne vožnje, u gospodarstvu i industrijskoj proizvodnji za povećanje učinkovitosti tvornica, prepoznavanju kvarova i održavanju strojeva, u poljoprivredi za osiguranje zdravije hrane s manjom potrošnjom gnojiva, pesticida i vode, u ekologiji radi smanjenja štetnog utjecaja na okoliš. Može naći primjenu u praćenju kretanja i prehrambenih navika životinjskih vrsta putem termovizijskih kamera. Umjetna inteligencija može upozoravati na opasnost od prirodnih nepogoda i katastrofa kao što su potresi i poplave kako bi se pravovremeno pripremili na njih i samim time ublažili njihove štetne posljedice (europarl, 2021). Da bi se takvi rizici izbjegli potrebni su mehanizmi koji će ih uspješno prepoznati i s velikom lakoćom ukloniti pa možemo „zamisliti kako se može promijeniti način na koji živimo baš kao što je izum električne energije nekoć promijenio ljudski život“ (Fawzy, 2020).

Turingov test je izradio Alan Turing. Test služi za provjeru postojanja inteligencija računala i/ili programskog proizvoda te količinu u kojoj je ta inteligencija prisutna. Po tom bi testu “računalo zaslužilo da se nazove inteligentnim ako bi moglo prevariti čovjeka da mu povjeruje da je čovjek.” (Turing, 1950).

U Turingovom testu sudjeluje osoba s ulogom kompetentnog i nepristranog sudca razgovara. Kroz test vrši se komunikacija prirodnim jezikom između čovjeka i stroja (slika 1.). Ako osoba u ulozi sudca ne može prepoznati koji od sugovornika je stroj, a koji čovjek, dolazi se do zaključka da je stroj prošao test (Wikipedia contributors, 2021).



*Slika 1: Turingov test (Pietikäinen & Silvén, 2022)*

Nesumnjivo je i nesporno da umjetna inteligencija unaprjeđuje i olakšava živote ljudi i sve više preuzima presudno značenje u različitim segmentima ljudskog života, ali postoji i njezina loša strana, a granica između korisnosti i štetnosti ponekad zna biti vrlo tanka. Svaki novčić ima dvije strane, pa je takav slučaj i kod umjetne inteligencije ovisno o tome iz kojeg kuta gledanja je promatramo. Umjetna inteligencija može biti dostupna non stop odnosno 24 sata na dan, 7 dana u tjednu i 365 dana u godini, a prilikom rada ne mora koristiti pauzu već radi kontinuirano, za razliku od ljudske inteligencije kod koje je istraživanje pokazalo da je prosječan čovjek maksimalno koncentriran i produktivan svega oko 3 do 4 sata dnevno (Curtin, n.d.). Umjetna inteligencija može misliti puno brže od ljudi i uz to istovremeno izvoditi i druge zadatke bez pogrešaka s odličnim rezultatima, a velika prednost je što se ona može za razliku od čovjeka nositi i s jednoličnim ponavljanjima poslova uz pomoć algoritama umjetne inteligencije. Uvelike je zaslužna za

pronalazak novih izuma u svim područjima znanosti te je omogućila rješavanje i najsloženijih problema. Umjetna inteligencija donosi odluke koje su nepristrane za razliku od ljudskih odluka koje su vođene osjećajima, odnosno emocijama koje se ni u kojem slučaju kod ljudi ne mogu u potpunosti isključiti. Zbog toga je maksimalno praktična i racionalna u pristupu svim problemima koji se postavljaju pred nju. Kreiranje uređaja koji je sposoban oponašati ljudsku inteligenciju je vrlo zahtjevna zadaća i iziskuje velike materijalne troškove, veliki utrošak vremena i ostalih sredstava, te se kontinuirano mora održavati i nadograđivati softver, ali isto tako i hardver (Bhbosale i sur., 2020).

Obnavljanje i moderniziranje uređaja je važno kako bi se njegovim stalnim ažuriranjem bolje ispunili svi zahtjevi. Zbog svega navedenog takvi uređaji imaju visoku cijenu na tržištu. Razvojem umjetne inteligencije potreba za ljudskim radom i angažmanom postaje sve manja što uzrokuje poremećaje ravnoteže na tržištu rada i zapošljavanja. Nedostatak uređaja je da ne mogu ostvariti timski rad poput ljudi, ne stvaraju međusobne veze kao što to čine ljudi i na njima se ne može primijeniti menadžerska sposobnost upravljanja, ne mogu izvršavati zadatke za koje nisu dizajnirani ili programirani, a mogu imati i nedostatak da se uruše ili počnu davati nebitne ili pogrešne rezultate (Bhbosale i sur., 2020).

### **3. Tehnologije umjetne inteligencije za učitavanje glasa**

#### **3.1. Strojno učenje**

Strojno učenje je grana umjetne inteligencije i računalne znanosti koja se fokusira na upotrebu podataka i algoritama za oponašanje načina na koji ljudi uče i to tako da postupno sve više poboljšava točnost (IBM Cloud Education, 2022). Pomoću njega se metodološki i automatski analiziraju određeni analitički podatci i omogućuje softverskim programima povećanje preciznosti u predviđanju budućih ishoda bez izričitog prethodnog programiranja. Strojno učenje ima cilj oponašajući ljudske sposobnosti, potrebe i djelovanja osposobiti stroj za učenje, pa se “poboljšanje vještina stroja može tumačiti u biološkom smislu poput jačanje obrasca neuronskih veza za obavljanje željene funkcije“ (Bavakutty, 2006). Algoritmi strojnog učenja upotrebljavaju prethodno definirane ulazne podatke pomoću kojih se predviđaju novi izlazni podaci i informacije. Odabir vrsta algoritama podataka koje je programer odlučio koristiti ovisit će o vrsti podataka koje želi dobiti. Strojno učenje nalazi svoju primjenu u različitim pogledima upotrebe i to primjerice u

ispitivanju tržišta, formiranju složenog operativnog sustava za autonomne automobile, algoritmima za preporuku određenih proizvoda ili usluga, otkrivanju zloupotreba na društvenim mrežama, filtriranje spamova odnosno neželjene pošte, filtriranjem štetnih softvera, automatiziranje i održavanje ključnih poslovnih procesa i sl. (Wakefield, n.d.).

Vrste strojnog učenja se najčešće razlikuju prema kriterijima koji odgovaraju na pitanja na koji način algoritam uči ili na koji način će postati što točniji u svojim predviđanjima. Postoje četiri osnovna pristupa pomoću kojih se određuju vrste strojnog učenja, a to su učenje pod nadzorom, učenje bez nadzora, polunadgledano učenje i učenje s pojačanjem (javatpoint, n.d.).

Primjena strojnog učenja zbog svojih prednosti može biti od velike pomoći poduzećima kako bi lakše razumjela svoje zaposlenike na jednoj kvalitetnijoj razini ili pri prikupljanju podataka o kupcima i njihovim navikama i običajima. Primjerice, kroz vrijeme algoritmi strojnog učenja mogu spoznati potrebe određenog dijela populacije ljudi i uz pomoć stručnih timova omogućiti razvoj određenog proizvoda i osmisliti marketinški plan prema zahtjevima i potrebama kupaca (Burns, 2021).

U nekim tvrtkama strojno učenje se koristi kao osnovni pokretač u stvaranju poslovnih konstrukcija. Npr. Uber aplikacija koristi algoritme za unos i povezivanje vozača u sustav vozača, a Google pretraživač koristi strojno učenje kako bi prezentirao oglase za te vozače prilikom pretraživanja od strane korisnika (Burns, 2021).

Nedostaci strojnog učenja su prije svega veliki ulazni troškovi, jer projekti strojnog učenja obično podrazumijevaju angažiranje stručnjaka za obradu podataka, koji su visoko plaćeni, a ovi projekti najčešće zahtijevaju i naprednu softversku infrastrukturu koja je vrlo skupa. Nedostatak strojnog učenja je i problem selektivnosti jer algoritmi nakon obuke na skupovima podataka isključuju određeni dio skupova te mogu dovesti do pogrešne percepcije stvarnosti. Samim tim gubi se njezina očekivana svrha i funkcija što u najgorem slučaju može dovesti i do različitih oblika zabluda i dezinformacija. U slučajevima kada poduzeće svoj rad temelji na poslovnim procesima koji su utemeljeni na neobjektivnim predviđanjima, može naići na veliku štetu i protuzakonite radnje što će onda u mnogo čemu štetiti reputaciji i ugledu tvrtke (Khanzode & Sarode, 2020).

### **3.1.1. Nadzirano strojno učenje**

Nadzirano strojno učenje zahtijeva od programera unos potrebnih podataka kako bi se mogao trenirati algoritam unutar parametra označenog ulaza i željenog izlaza. Kod učenja pod nadzorom programer ubacuje u algoritme točno određene podatke koji će definirati željenu varijablu, te će se putem toga algoritam procijeniti odnos među njima (javatpoint, n.d.).

Primjer nadziranog strojnog učenja je klasifikacija (slika 2), a ona se odnosi na prediktivno modeliranje (koje koristi statistiku za predviđanje ishoda), a u kojem oznaka klase služi kao primjer ulaznih podataka. Primjer klasifikacije u strojnom učenju je utvrđivanje sadržaj da li je spam ili nije (Brownlee, 2020).

Algoritmi nadziranog učenja dobri su za zadatke kao što su binarna klasifikacija, odnosno klasifikacije u samo dvije grupe po određenim pravilima klasifikacije koja dijeli podatke u dvije kategorije (IBM Cloud Education, 2020a):

- **razvrstavanje u više razreda (eng. classification)** dijeli podatke u kategorije iz koje klasificira stavke između više od dvije vrste odgovora. Npr. klasifikacija skupa slika voćaka koje mogu biti naranče, jabuke ili kruške. Razvrstavanje u više razreda pretpostavlja da je svaki uzorak dodijeljen jednoj i samo jednoj kategoriji: voće može biti ili jabuka ili kruška, ali ne oboje u isto vrijeme (Nabi, 2018).
- **regresijsko modeliranje (eng. Regression)** koje predviđa kontinuirane vrijednosti. Npr. rast broja diplomiranih studenata u SAD-u između 2001. i 2012. godine je linearnog rasta, odnosno svake godine se broj studenata povećavao u skoro pa istoj mjeri. Regresijsko modeliranje nam na grub način odgovara na pitanje koliko će biti studenata u budućnosti. Pošto algoritam zna da broj diplomiranih studenata raste gotovo linearno svake godine, može se dobiti gruba procjena da će broj studenta do 2018 biti između 2,0 i 2,1 milijun. Gledajući stvarne podatke možemo potvrditi da od 2001. do 2018. broj studenata je bio u okvirima stvarne vrijednosti (Anwar, 2021).

### 3.1.2. Strojno učenje bez nadzora

Kod strojnog učenja bez nadzora uključeni su algoritmi koji se obučavaju na neoznačenim podacima. Označeni podaci posjeduju informacije, dok neoznačeni ne posjeduju nikakvu informaciju. Algoritam prolazi kroz nizove skupova podataka tražeći nekakvu logičnu vezu. Podaci kroz koje algoritmi prolaze i uče, kao i predviđanja ili očekivanja rezultata koji će izaći

unaprijed su predefiniрани (jvatpoint, n.d.). Algoritmu za strojno učenje bez nadzora nije potrebno označavanje podataka. On pregledava podatke kako bi pronašao željene obrasce koji se mogu koristiti za grupiranje podatkovnih točaka u podskupove. Velika većina vrsta dubokog učenja koje uključuju i neuronske mreže su algoritmi bez nadzora. Algoritmi za učenje bez nadzora dobri su za izvršavanje sljedećih zadataka (Google Developers, 2022):

- grupiranje koje dijeli skupove podataka u grupe na temelju sličnosti,
- otkrivanje anomalija kada se prepoznaju neobične točke u skupu podataka,
- asocijacijsko rudarenje kojim se identificiraju skupovi stavki u skupu podataka,
- smanjenje dimenzija skupa podataka posljedica je smanjenje broja varijabli u skupu podataka.

Klasteriranje je čest primjer korištenja strojnog učenja bez nadzora (slika 2). Klasteriranjem kompjuter bilježi skup kojem pripadaju određeni podatci odnosno klasteriranje se koristi kod istraživanja tržišta, analize društvenih mreža, grupiranja rezultata pretraživanja, medicinskog snimanja, segmentacija slika i otkrivanja anomalija (Google Developers, 2022).

### **3.1.3. Polunadzirano strojno učenje**

Kod polunadziranog učenja uključene su mješavine dviju prethodnih vrsta. Stručnjaci mogu unositi podatke u algoritam, a ti podaci ga najčešće označuju. Nakon što je model označen, sustav se može samostalno opskrbljivati istraženim podacima i razvijati samostalno usvojeno razumijevanje skupova podataka (jvatpoint, n.d.). Polunadzirano učenje funkcionira tako da se od strane programera unose male količine označenih podataka radi treninga algoritma. Iz tih podataka algoritam uči dimenzije skupa podataka koje zatim primjenjuje na nove neoznačene podatke. Izvedba algoritama se često poboljšava treniranjem na označenim skupovima podataka. Problem je što označavanje podataka može biti dugotrajno i zahtjevno (DataRobot, 2020).

Polunadzirano učenje je dobar kompromis između izvođenja nadziranog učenja i učinkovitosti nenadziranog učenja. Područja koja koriste polunadzirano učenje uključuju sljedeće (DataRobot, 2020):

- strojno prevođenje koje podrazumijeva podučavanje algoritama kako bi se mogao prevesti jezik na temelju manje riječi od punog rječnika,

- otkrivanje zlonamjernog sadržaja što podrazumijeva prepoznavanje zlonamjernih slučajeva kada postoji malo valjanih primjera,
- označavanje podataka što podrazumijeva da bi se algoritmi koji se treniraju za manje skupove podataka mogli naučiti automatski primijeniti na označene podatke u većim skupovima.

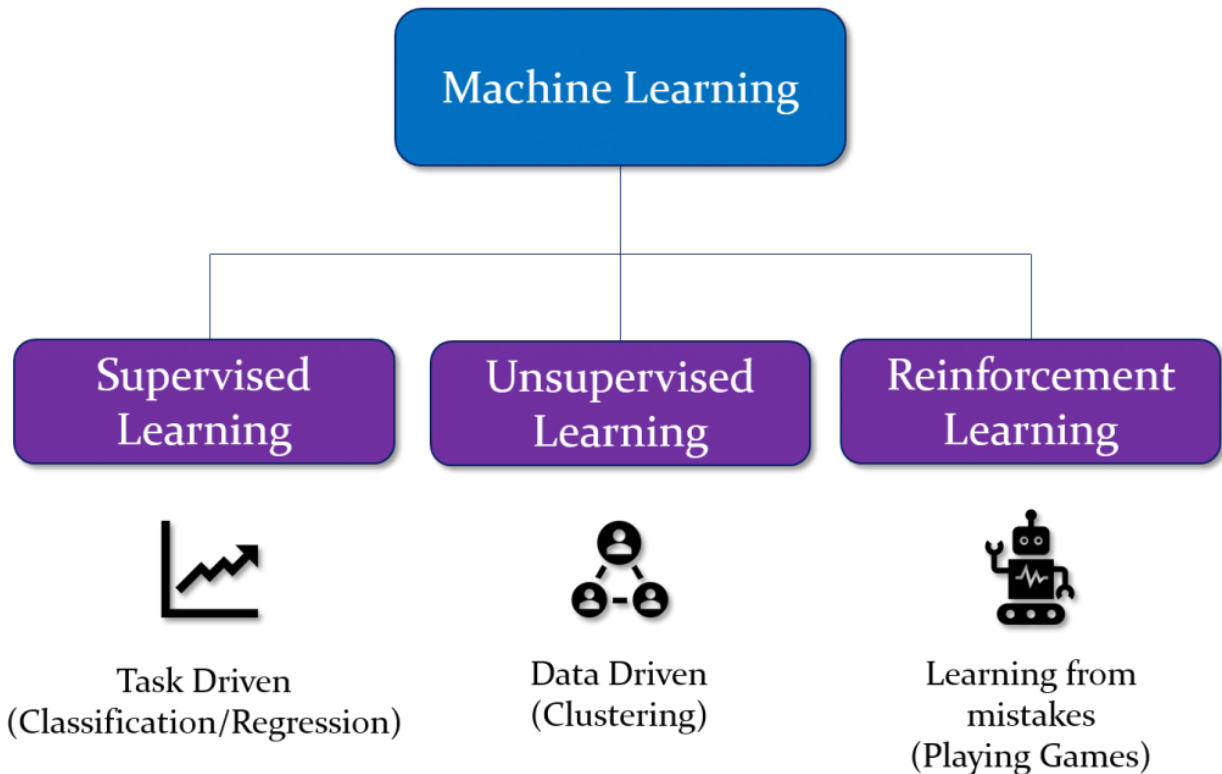
#### **3.1.4. Strojno učenje s pojačanjem**

Kod učenja s pojačanjem programer podatke obično koristi kako bi se stroj naučio dovršiti proces u više faza za koje postoje jasno definirana pravila. Stručnjaci podacima treniraju algoritam kako bi on dovršio zadaću na način da mu tijekom rada daju pozitivne ili negativne signale sve dok ne dovrši zadatak. Bitno je da kod ove vrste učenja algoritam uglavnom sam odlučuje koje će korake poduzeti u tom smjeru (javatpoint, n.d.). Ono funkcionira tako da se programiranjem algoritma s određenim ciljem i propisanim skupom pravila postiže zadani cilj. Stručnjaci za podatke programiraju algoritam koji će tražiti nagrade koje dobiva kada izvršava radnje koje su korisne za postizanje konačnog cilja, a izbjegava kaznu koju će dobiti kada izvede radnje koja ga udaljavaju od postizanja konačnog cilja (Budek & Osiński, 2018). Strojno učenje se često koristi u obučavanju protivne strane u videoigrama (slika 2.), a način na koji se protivna strana nagrađuje ili kažnjava uvelike ovisi o vrsti problema, npr. davanju protivnoj strani nagrade za pobjedu u igri ili kazne za gubitak (Wikipedia contributors, 2022).

Učenje s pojačanjem se često koristi u sljedećim područjima (Budek & Osiński, 2018):

- robotika gdje roboti trebaju naučiti izvršavanje fizičkih zadataka u stvarnom svijetu koristeći spomenutu tehnologiju,
- učenje botova iz videoigara kako bi ih podučili oponašati svoju ulogu unutar igre,
- upravljanje resursima se odnosi na upravljanje ograničenim resursima i definiranjem ciljeva, a učenje s pojačanjem pomaže poduzećima bolje planirati i rasporediti svoje resurse.

## Types of Machine Learning



Slika 2: Tipovi strojnog učenja (Rishabh & Piyush, 2020)

### 3.2. Duboko učenje

Duboko učenje je podvrsta strojnog učenja, koje je u biti neuronska mreža s tri ili više slojeva koja pokušava simulirati ponašanje ljudskog mozga (iako je to daleko od njegove sposobnosti), ali mu omogućuje da uči iz velikih količina podataka. Neuronska mreža s jednim slojem daje približna predviđanja, a dodatni slojevi pomažu u optimizaciji i poboljšanju točnosti (IBM Cloud Education, 2020b). Za duboko učene također je bitno spomenuti da osposobljava računalo kako bi izvršavalo zadatke kakve izvršava čovjek. Te radnje se sastoje od prepoznavanja govornih naredbi i tehnologije obrade slike ili predviđanja određenih kretanja u prirodi i društvu. Duboko učenje ne funkcionira na način da koristi podatke koji su unaprijed predefimirani jednadžbama i algoritmima, već postavlja bitne parametre i tako se osposobljava računalo da samostalno uči iz podataka i prepoznaje obrasce pomoću kojih na slojevit način sve skupa obrađuje (Markotić, 2021).



Za stvaranje jakih modela dubinskog učenja upotrebljavaju se razne tehnike i metode kao što su metoda pada stope učenja, metoda prijenosa učenja, metoda treninga od nule i metoda napuštanja (Burns & Brush, 2021).

**Metoda pada stope učenja (eng. learning rate decay method)** je tehnika za treniranje modernih neuronskih mreža. Počinje se trenirati mreža s velikom stopom učenja, a zatim se polako smanjuje/opada dok se ne postignu lokalni minimumi. Empirijski je uočeno da pomaže u optimizaciji i generalizaciji (Haswani, 2020).

Pad stope učenja je tehnika koja definira sustav ili postavlja uvjete za njegov rad prije samog procesa učenja, a kojim se kontrolira koliko promjena model doživljava kada odgovora na pogreške kod mijenjanja modela. Stope učenja koje su previsoke rezultiraju nestabilnošću procesa učenja, a stope učenja koje su preniske kao posljedicu imaju dugotrajan proces obuke koji ima velike šanse za nedostatke u radu (Burns & Brush, 2021).

**Metoda prijenosa učenja (eng. transfer learning method)** je proces usavršavanja prethodno uvježbanog modela. Korisnik dodaje postojećoj mreži nove podatke koji sadrže prethodno nepoznate klasifikacije. Nakon što se mreža prilagodi, novi se zadaci mogu izvoditi sa specifičnijim kategorijama. Prednost ove metode je što zahtijeva manje podataka od drugih metoda, čime se vrijeme trajanja obuke smanjuje za minute ili sate (Burns & Brush, 2021). Ova metoda je trenutno vrlo popularna u dubokom učenju jer može trenirati duboke neuronske mreže s relativno malo podataka. Vrlo je korisna u području znanosti o podacima budući da većina problema iz stvarnog svijeta obično nema milijune označenih podatkovnih točaka za obuku tako složenih modela (Donges, 2022).

**Metoda treninga od nule (eng. training from scratch method)** je proces koji zahtijeva od programera da prikupi i označi veliku bazu podataka te stvori mrežnu arhitekturu koja može naučiti njezine značajke i modele. Ova metoda je najkorisnija za nove aplikacije i aplikacije s velikim brojem izlaznih kategorija, ali se u praksi manje primjenjuje jer zahtijeva prevelike količine podataka. Shodno tome produljuje se obuka na duže vrijeme, odnosno produljuje se na nekoliko dana ili tjedana (Burns & Brush, 2021).

**Metoda napuštanja (eng. dropout method)** je proces kojim se pokušava riješiti problem ogromnog unosa podataka na mreži i velike količine parametara nasumičnim izbacivanjem

jedinica i njihovih veza iz neuronske mreže tijekom obuke. Zbog navedenog, metoda napuštanja može poboljšati rad neuronskih mreža na zadacima učenja pod nadzorom i u područjima kao što su prepoznavanje govora, klasifikacija dokumenata i računalna biologija (Burns & Brush, 2021). Napuštanje se odnosi na nasumično zanemarivanje određenih čvorova tijekom obuke. Napuštanje je pristup koji sprječava prekomjerno opterećenje osiguravajući da nijedna jedinica ne ovisi jedna o drugoj (Marimuthu, 2022).

### **3.3. Obrada prirodnog jezika**

Obrada prirodnog jezika je grana umjetne inteligencije koja omogućava i olakšava računalu da precizno interpretira tekst i govor te tako sastavlja rečenice na čovjeku razumljiv način. U svrhu postizanja prirodne komunikacije između čovjeka i računala, računalima se omogućava učitavanje teksta, spremanje snimljenog govora, dešifriranje značenja teksta i govora, prepoznavanje namjere i govornih emocija koje proizlaze iz predmetnog govora ili riječi i formiranje bitnih dijelova govora. Pomoću obrade prirodnog jezika pokreću se i računalni programi koji mogu prevoditi tekst s jednog jezika na drugi i brzo spajati velike količine teksta (IBM Cloud Education, 2021).

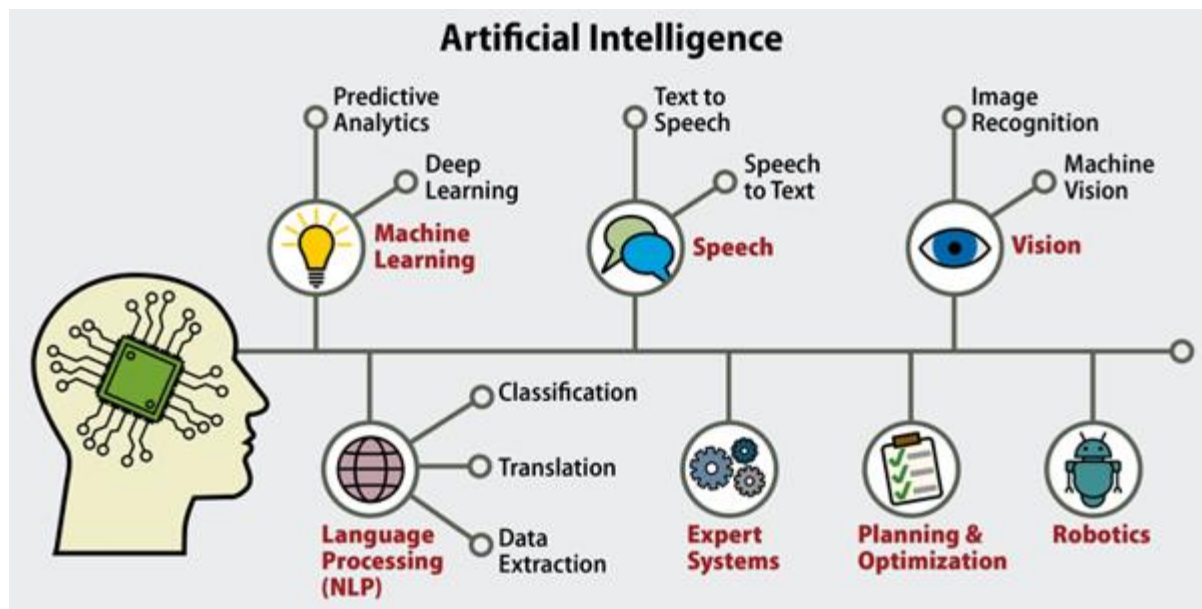
Zadatak obrade prirodnog jezika je pokušati izgraditi softvere koji mogu razumjeti i reagirati na tekstualne ili glasovne podatke, a nakon toga povratno reagirati svojim vlastitim tekstom ili govorom i to sve na isti način kao što to čine ljudi. Obrada prirodnog jezika je mješavina računalne lingvistike kao prilagođenog ljudskog jezika baziranog na pravilima koja je upotpunjena sa statističkim instrumentima, mehanizmima strojnog učenja i principima dubokog učenja (DataRobot, 2016).

Na takav način stvara se doživljaj izravne komunikacije prirodnim jezikom s digitalnim asistentom. Obrada prirodnog jezika također dobiva sve veću važnost u poslovnim rješenjima koje pomažu pri racionaliziranju poslovne operacije, povećanju produktivnosti zaposlenika i pojednostavljenju poslovnih procesa koji su presudni za ostvarenje unaprijed postavljenih bitnih ciljeva. Ljudski jezik nije savršen jer u njemu postoje određene nejasnoće zbog kojih je vrlo teško napisati softver koji može točno odrediti ciljano značenje teksta ili glasovnih podataka. Razne nepravilnosti u jeziku poput homonima, homofona, sarkazama, idioma, metafora, gramatičkih i upotrebnih iznimaka, varijacija u rečeničnom nizu predstavljaju samo neke od nepravilnosti

prilikom upotrebe ljudskog jezika, a za učenje toga je i ljudima potrebno dugo vremena (IBM Cloud Education, 2021).

Obradom prirodnog jezika ispunjava se zadaća uređivanja pojedinih ljudskih tekstova i glasovnih podataka kako bi računalo moglo shvatiti na koji način treba obraditi unesene podatke. Neki od takvih zadataka uključivali bi sljedeće (IBM Cloud Education, 2021):

- zadatak prepoznavanja govora je pouzdano pretvaranje glasovnih podataka u tekstualne podatke. Prepoznavanje govora je nužno potrebno za sve aplikacije koje slijede glasovne naredbe ili odgovaraju na izgovorena pitanja. Najizazovnije kod prepoznavanja govora je način na koji ljudi govore, jer neki govore brzo, neki nerazgovijetno izgovaraju riječi, koriste različite naglaske i intonacije, a često se koristi i pogrešna gramatika,
- označavanje dijela govora, koje se još zove gramatičko označavanje je proces prepoznavanja nekih riječi ili teksta na temelju i u kontekstu njegove upotrebe. U govoru se identificira glagol “napraviti” kao glagol koji se nalazi u rečenici “Ja mogu napraviti avion od papira”,
- višeznačnost riječi je proces odabira značenja riječi s više značenja u postupku semantičke analize kojom se nastoji utvrditi riječ koja ima najviše smisla u zadanom kontekstu. Na primjer, kod riječ s više značenja traži se pravi smisao riječi što pomaže pri razlikovanju u smislu značenja, npr. glagol “peći” (ispeći nešto) u odnosu na imenicu “peći” (naprave za grijanje),
- prepoznavanjem imenovanih pojmova, identificira se prikladna riječ ili izraz. Prepoznavanje imenovanog pojma “Lokve” kao naselja u odnosu na „lokve“ kao nakupine vode nakon kiše,
- prepoznavanje riječi ili fraza koje se odnose na istu osobu ili stvar imaju zadatak identificirati da li se i kada dvije riječi odnose na istu osobu ili predmet. Primjer za određivanje osobe ili predmeta na koje se određena riječ treba odnositi je „oni“ = „studenti”, ali također može uključivati prepoznavanje metafora ili idioma npr. u tekstu u kojem “majmun” nije životinja nego nerazumna osoba,
- analizom osjećaja pokušava se iz teksta izvući subjektivne kvalitete kao što su stavovi, emocije, sarkazmi, zbunjenost ili sumnje.



Slika 3: Grane umjetne inteligencije (Gaël, 2020)

## 4. Vrste umjetne inteligencije

Proizvodi umjetne inteligencije se izrađuju kako bi ispunjavali različite svrhe, što je i jedan faktor po kojem se i prepoznaju, odnosno razlikuju. U daljnjem tekstu je obrađena hijerarhija vrsta umjetne inteligencije, kao što je prikazano na slici 4 (Markotić, 2021):

### 4.1. Umjetna uska inteligencija

Ovo je najčešći oblik umjetne inteligencije koji se pojavljuje na tržištu. Ovakvi sustavi su dizajnirani kako bi rješavali pojedinačne probleme i kako bi mogli uspješno izvršiti svaku pojedinačnu zadaću. Po svojoj biti imaju ograničene i uske mogućnosti kao npr. reklamiranje točno određenog proizvoda prilikom kupnje na internetskim trgovinama. Ovakvi sustavi se mogu približiti načinu na koji ljudi funkcioniraju i to u vrlo specifičnim okolnostima, a ponekad mogu u određenim slučajima čak i premašiti ljudsko funkcioniranje, ali se to dešava samo u vrlo poznatom okruženju i ograničenom uskom skupu parametara (Markotić, 2021).

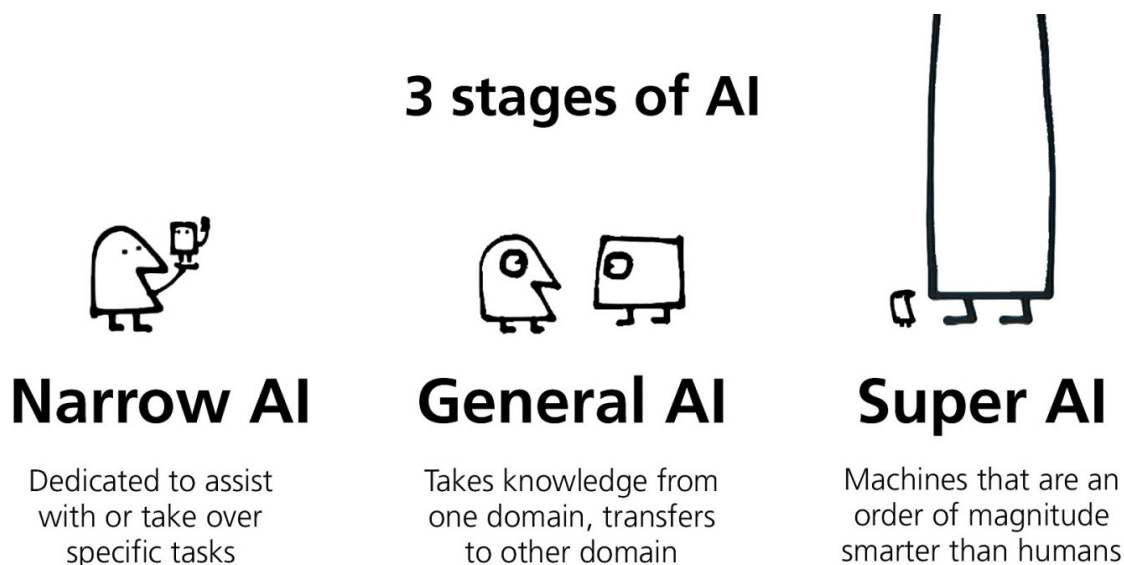
### 4.2. Umjetna opća inteligencija

Umjetna opća inteligencija je prikaz generaliziranih ljudskih kognitivnih sposobnosti u softveru koji bi suočen s nepoznatim zadatkom mogao pronaći rješenje za koje je i ljudsko biće sposobno (Lutkevich, 2022). Ovakav oblik umjetne inteligencije je još uvijek na nivou teorije. Definira se kao umjetna inteligencija koja za funkciju osim prepoznavanja i spoznaje, ima i dodatnu mogućnost da osjeća i predviđa uzroke događanja. Njezina funkcija doseže do razine ljudskosti i obuhvaća široki opseg i djelokrug funkcioniranja što uključuje jezičnu i vizualnu obradu, čiji je rezultat rad uređaja s uključenim logičkim zaključivanjem. Čovječanstvo je još uvijek jako udaljeno od izgradnje i uspostave tako naprednog sustava. Takav sustav bi trebao u sebi imati tisuće manjih sustava umjetne uske inteligencije koji rade usklađeno i međusobno komuniciraju kako bi što bolje oponašali ljudsko djelovanje i rasuđivanje. Njegova složenost može se zorno prikazati primjerom da je „upotrebom najnaprednijih računalnih sustava i najsuvremenijim infrastrukturama, poput Fujitsuova K ili IBM-ovog Watsona, trebalo minimalno 40 minuta da se poistovjete i simuliraju barem jednu sekundu neuronske aktivnosti“ (Markotić, 2021). To ukazuje na ogromnu i nemjerljivu pa čak i beskonačnu složenost i zajedničku interakcijsku povezanosti dijelova ljudskog mozga kao cjeline te također potvrđuje zahtjevnost i veličinu izazova izgradnje ove vrste umjetne inteligencije (Markotić, 2021).

### **4.3. Umjetna super inteligencija**

Umjetna super inteligencija je sustav koji se temelji na softveru s intelektualnim moćima koje nadilaze ljudske u obuhvatu svih raspona, kategorija i područja djelovanja (Wigmore, 2018). Umjetna super inteligencija može se smatrati kao logički slijed napretka umjetne opće inteligencije. Razvojem sustava umjetne super inteligencije moglo bi se dogoditi da se njom nadmaše sve poznate ljudske sposobnosti, što bi uključivalo donošenje svih bitnih i racionalnih odluka te bi uključivalo nešto kao što je npr. izgradnja savršenih emocionalnih odnosa (Markotić, 2021). Kada se postigne sveobuhvatna primjena umjetne opće inteligencije, tada će svi sustavi biti u mogućnosti dodatno unaprijediti sve svoje sposobnosti do maksimalnih granica i na taj način ostvariti napredak i postignuća u neslućena područja kakva nismo mogli ni zamisliti. U tom slučaju razlike i suprotstavljenosti umjetne opće inteligencije i umjetne super inteligencije bi bile relativno malene (možda tek neku nanosekundu, jer bi tom brzinom učila umjetna inteligencija). Dugi put za ostvarenje toga cilja je svakako pred nama, a naročito put prema samoj umjetnoj općoj

inteligenciji pa se zbog toga čini da ovaj koncept leži tek u nekoj dalekoj budućnosti (Markotić, 2021).



Credit: Chris Noessel

*Slika 4: Vrste umjetne inteligencije (Codecnetworks, 2021)*

#### **4.4. Slaba umjetna inteligencija**

Slaba umjetna inteligencija je vrsta umjetne inteligencije koja je ograničena na određeno usko područje djelovanja. Ona simulira ljudsku spoznaju i ima ulogu automatiziranja zadataka ili analiziranja podataka na način na koji ljudi to ponekad ne mogu. Slaba umjetna inteligencija je suprotnost od jake umjetne inteligencije koja teži biti jednaka ljudskoj inteligenciji (Frankenfield i sur., 2022). Slaba umjetna inteligencija radi daleko lošije i slabije u odnosu na jake umjetne inteligencije, ali je zato njezino korištenje poprimilo globalne razmjere i velika većina ljudi se danas koristi i upoznati su s tom sada čestom i lako dostupnom tehnologijom. Najpopularnije iPhonova Siri i Amazonova Alexa mogu se smatrati programima slabe umjetne inteligencije. Ova vrsta umjetne inteligencije utemeljena je na razlici između nadziranog i nenadziranog učenje iz razloga što glasovno aktivirana pomoć ima najčešće unaprijed pretpostavljeni i programirani odgovor. Ona funkcionira na takav način da prepoznaje i pamti ono što je podsjeća na sličnost na ono što već zna od prije pa prema tom kriteriju i kategorizira stvari. Ova njezina osobina slična je čovjekovim osobinama, ali na tome sve sličnosti prestaju postojati jer je slaba umjetna inteligencija

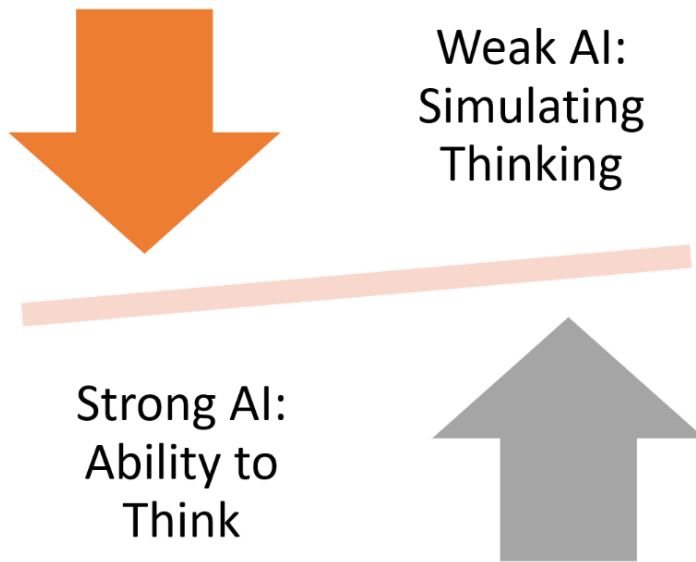
jednostavno izgrađena na simulaciji i oponašanju čovjekovih osobina. Njezin nedostatak je upravo u tome što ona reagira isključivo na način kako je programirana, ali nije sposobna razumjeti ili izvoditi zaključke u odnosu na ono što je uneseno, a takva simulacija razmišljanja je prikazana na slici 5 (Markotić, 2021).

#### **4.5. Jaka umjetna inteligencija**

Jaka umjetna inteligencija teorijski je oblik umjetne inteligencije koji se koristi za opisivanje načina razmišljanja o njezinom razvoju. Ako istraživači razviju jaku umjetnu inteligenciju, stroj će imati inteligenciju jednaku ljudskoj, dakle imao bi svijest koja ima sposobnost rješavanja problema, učenja i planiranja budućnosti (IBM Cloud Education, 2021).

To je isto kao kada razgovaramo s čovjekom, jer tada možemo samo nagađati i pretpostavljati kakav bi mogao biti njegov odgovor, stoga ona ima sposobnost i mogućnost razmišljanja kako je to i prikazano na slici 5 (Markotić, 2021).

Jaka umjetna inteligencija funkcionira puno kvalitetnije i jače u odnosu na slabu umjetnu inteligenciju. Primjenu nalazi u stvaranju suvremenih video igara, proizvodnji zahtjevnih pametnih telefona i modernih automobila ali i drugih modernih aparata i uređaja. Predstavljena je u mnogim SCI-FI filmovima i zbog svoje ogromne snage i funkcionalnosti podsjeća i prikazuje se poput ljudskog mozga. Za kategoriziranje obrađenih podataka koriste klasteriranje odnosno grupiranje na takav način da se stvari sa sličnim značajkama zbližavaju, a stvari s različitim značajkama razdvajaju. To bi značilo da program prilikom unosa ključnih riječi ili zahtjeva nema unaprijed programirani odgovor, za razliku od slabe umjetne inteligencije koja funkcionira upravo na takav način. Rezultati takvog programiranja i njegova funkcija u većini slučajeva su uglavnom nepredvidivi. Najbolji primjer jake umjetne inteligencije se nalazi u videoigrama. Ovaj oblik je daleko samostaliji od slabe umjetne inteligencije upravo iz razloga što može samostalno učiti i sam se prilagođavati različitim situacijama, pa je stoga „znanstvenici često opisuju kao istinsku zastupljenost ljudske inteligencije u strojevima“ (Markotić, 2021). Iako se slaba umjetna inteligencija češće pojavljuje, jaka umjetna inteligencija također ima presudni značaj u revolucionarnom razvoju umjetne inteligencije. Kako bi se umjetna inteligencija lakše i bolje razumjela potrebno je razlikovati “pojmove slabe i jake umjetne inteligencije koje je smislio John Searle kako bi razlikovao performanse u različitim vrstama AI strojeva“ (Markotić, 2021).



*Slika 5: Razlika između slabe i jake umjetne inteligencije (Rodriguez, 2020)*

## **5. Glasovni asistenti i prepoznavanje glasa**

Svi pametni uređaji u pravilu izlaze na tržište s ugrađenim digitalnim pomoćnikom, odnosno programom koji koriste tehnologiju prepoznavanja glasa za potrebe izvršavanja određenih zadataka. Prepoznavanje glasa je skup algoritama koje digitalni pomoćnici koriste kako bi pretvorili govor u digitalni signal i kako bi mogli prepoznati ono što im se govori. Algoritam prepoznavanja govora radi kroz akustično i jezično modeliranje. Glasovni digitalni pomoćnik koristi prepoznavanje glasa te pomoću algoritma za obradu jezika i određenih glasovnih naredbi daje relevantne informacije ili obavlja određene funkcije prema zahtjevu korisnika (Computer Hope, 2021).

Glasovni asistenti koriste se za upravljanje uređajima, davanje naredbi i pisanje bez upotrebe tipkovnice, miša ili pritiskanja bilo kakvih tipki. Rade uz pomoć računala i softverskih programa za automatsko prepoznavanje govora koji se služe podacima korištenja da prikupe podatke kako bi "istrenirali" program koji će im omogućiti da što bolje prepoznaju glas te pretvore govor u tekst. Kako bi računalo moglo prepoznati glas mora imati zvučnu karticu, mikrofonski uređaj, slušalice ili zvučnik i ostatak potrebnog hardvera ugrađen u uređaj. Softver treba imati instaliranu programsku podršku za prepoznavanje glasa kao što je npr. Nuance Naturally Speaking (Brown, 2021). Nakon što sustav za prepoznavanje glasa identificira riječi, te se informacije šalju procesoru, a sustav tada izvršava zadatak za koji je namijenjen (Brown, 2021).



Proces prepoznavanja glasa počinje pretvaranjem zvuka u digitalni signal nakon čega mora proći određene korake kako bi shvatio što se govori. Kada mikrofona na uređaju uhvati zvuk, on se tada pretvara u električnu struju koja, onda putuje do analogno-digitalnog pretvarača koji, pretvara električnu struju odnosno analogni signal u digitalni binarni signal. Dok struja teče do analogno-digitalnog pretvarača, uređaj uzima uzorke struje i dešifrirati napon u određenim vremenskim točkama koje se još nazivaju uzorcima, a svaki uzorak traje samo nekoliko tisućinki sekunde. Na temelju napona uzorka, analogno-digitalni pretvarač će dodijeliti niz od osam binarnih znamenki (jedan bajt podataka). Zvuk se treba obraditi radi bolje jasnoće kako bi uređaj bolje razumio govor. Uređaj ponekad treba dešifrirati govor i u bučnom okruženju zbog čega se na zvuk postavljaju određeni filteri kojima se smanjuje ili uklanja pozadinska buka i neželjene frekvencije, sve do raspona ljudskog sluha što pomaže u poboljšanju točnosti prepoznavanja. Ostale postavke, poput brzine i glasnoće zvuka, prilagođene su kako bi što bolje odgovarali referentnim audio uzorcima koje sustav za prepoznavanje glasa koristi za usporedbu. Rezultat svega je da sustav za prepoznavanje glasa počinje stvarati i analizirati riječi, odnosno govor i to najčešće na dva načina. Prvi je pomoću metode skrivenog Markovljevog modela koji se češće upotrebljava i dobro funkcionira te radi po principu rastavljanja izgovorenih riječi na najmanje elemente jezika, odnosno foneme (Brown, 2021). Kada sustav za prepoznavanje glasa identificira jedan fonem, određuje vjerojatnost pojave sljedećeg fonema. Na primjer, ako se izgovori glas "p", postoji određena vjerojatnost da će sljedeći fonem biti "u" da bi nastala riječ "Pula". Uz to postoji i vjerojatnost da će sljedeći fonem biti "t", ali to je daleko manje vjerojatno. Ako sljedeći fonem nalikuje "u", tada sustav može s velikom sigurnošću pretpostaviti da se radi o riječi "Pula". Drugi način je putem metode neuronskih mreža koje su poput digitalnog mozga koje uče slično ljudskom mozgu i ključne su u stvaranju umjetne super inteligencije i dubokog učenja. Neuronska mreža koja se koristi u prepoznavanju glasa naziva se ponavljajuća neuronska mreža i ona prepoznaje glas stalnim ponavljanjem gdje se izlaz iz prethodnog koraka unosi kao ulaz u sadašnji korak, odnosno obradom malog podatka koristi isti kako bi utjecala na ono što se radi sa sljedećim dijelom podatka i na taj način u stvari uči iz iskustva. Što je više izložena glasovima to će točnije prepoznavati glas. Na primjer ako sustav identificira glas "pu" 100 puta, a nakon njega slijedi glas "l" 90 od tih 100 puta, tada može naučiti da "l" obično dolazi nakon "pu", odnosno sustav za prepoznavanje glasa identificira fonem, koristi prikupljene podatke i predviđa koji će najvjerojatnije biti sljedeći glas.

## 5.1. Pojam glasovnog asistenta

Glasovni asistent je digitalni pomoćnik koji koristi prepoznavanje glasa, algoritme za obradu jezika i sintezu glasa za slušanje određenih glasovnih naredbi i vraćanje relevantnih informacija ili obavljanje određenih funkcija prema zahtjevu korisnika (Alan, n.d.). Dakle, to je softver koji obavlja svakodnevne zadatke putem glasovnih naredbi. On funkcionira na način da pomoću umjetne inteligencije odnosno strojnog učenja prepozna glas i učini ono što se od njega traži. Softver glasovnog asistenta je tehnologija novijeg doba, a može se pronaći na pametnim zvučnicima, pametnim satovima, mobilnim telefonima, tabletima i drugim uređajima, a najpoznatiji su Alexa (Amazon), Siri (Apple), Google asistent (Google) i Bixby (Samsung). Glasovni asistenti se prvi put pojavljuju na tržištu od 2014. godine i od tada sve više preuzimaju važnu ulogu zahvaljujući jednostavnoj tehnologiji za korištenje te na neki način postaje rastući popularni dio ljudske svakodnevice koji donosi radost i uzbuđenje u domovima i obiteljima. Korisnici glasovnim naredbama zahtijevaju da se pomoću korištenja umjetne inteligencije obave određeni zadatci poput puštanja glazbe, čitanja vijesti, upravljaju uređajima, provjeravanja stanja u banci i sl., sve kako bi se ljudima olakšalo i kako bi se poboljšala kvaliteta svakodnevnog života. Tehnologija potpomognuta glasom pametni je kućni pomoćnik koji pomoću softvera komunicira s korisnikom zvučnim putem i na sofisticiran i inteligentan način odgovara na izgovorene naredbe, odnosno koristi se za razgovor s računalom, pametnim telefonom ili drugim pametnim uređajima. Google Home ima preko 200 naredbi koje korisnik može koristiti prilikom upravljanja pametnim domom pomoću glasovnog asistenta, a u svom domu korisnik može imati puno različitih pametnih uređaja, ali komunicira preko jednog pametnog čvorišta samo s jednim glasovnim pomoćnikom (Hoy, 2018).

## 5.2. Najpopularniji jezici za izradu softvera za prepoznavanje govora

**C#:**

C# radi na različitim računalnim platformama tako da programer može jednostavno izvršiti ponovnu upotrebu koda. C# podržava preopterećenje operatora (eng. operator overloading) i direktive pretprocesora (eng. pre-processor directive) koji će pomoći u prepoznavanju govora te se s ovim jezikom može vrlo lako upravljati s prepoznavanjem govora (Gupta, 2017).

**Java:**

Za razvoj prepoznavanja govora, Java generira Java Speech API. To je skup apstraktnih klasa i sučelja koje Java programeri mogu koristiti kako bi jednostavnije izradili softver za prepoznavanje glasa. Ima izvrsnu sposobnost učenja iz govornih uzoraka te je ovaj jezik dobar u rješavanju problema nekompatibilnosti, odnosno podržava arhitektonsku neutralnost zbog koje programer može raditi na svim fizičkim procesorima bez gubitka performansi te ne mora zbog toga razmišljati o razvoju i kompilaciji (Gupta, 2017).

**PHP:**

PHP je jezik visoke razine, a njegova je sintaksa slična C jeziku te je radi toga početnicima lako izraditi softver za prepoznavanje govora. PHP je podržan na više operativnih sustava tako da će razvojnom programeru biti lako premjestiti aplikaciju napisanu u PHP-u s originalnog Windows poslužitelja na Linux poslužitelj bez modificiranja. PHP podržava klase otvorenog izvornog koda koje mogu biti korisne za izradu aplikacije za prepoznavanje glasa (Gupta, 2017).

**Perl:**

Perl je stabilan višepatformski programski jezik koji se u početku koristio za obradu teksta, ali danas pomoću ovog jezika programer može razviti web aplikaciju i mrežnu aplikaciju. Budući da ovaj jezik ima izvrsnu funkcionalnost i klase, razvoj aplikacije za prepoznavanje govora postaje lak za programera. U mnogim jezicima dodatne indentacije stvaraju pogrešku, Perl u drugu ruku ne mari za indentacije pa korisnik može jednostavno izraditi program. Programski jezik Perl automatski upisuje podatke i automatski upravlja memorijom tako da izrada aplikacije za prepoznavanje glasa postaje jednostavna.

Zaključno, svaki jezik ima jedinstvene značajke i svrhe, ali ako želite stvoriti govornu aplikaciju za web, tada je PHP najbolji jezik. Dakle, ako želite sudjelovati u stvaranju aplikacije za prepoznavanje govora za velika poduzeća, tada je Java izvrstan jezik. U skladu s poznavanjem svakog pojedinog jezika, može se koristiti bilo koji od jezika za izradu softvera za prepoznavanje govora (Gupta, 2017).

**5.3. Računalna lingvistika**

Računalna lingvistika je grana lingvistike koja se bavi analizom i računalnim modeliranjem prirodnog jezika (Campana, 2022). Računalna lingvistika se koristi alatima kao što su brzo strojno prevođenje, aplikacije za prepoznavanja govora, aplikacije za konverziju teksta u govor, interaktivnih sustava glasovnih odgovora, raznih tražilica, tekst editora i sustava za učenje jezika (Gillis & Bernstein, 2022). To je način primjene računalne znanosti na analiziranje i razumijevanje pisanog i govornog jezika. Računala koja predstavljaju lingvističku logistiku pomažu olakšati interakciju između ljudi, strojeva i softverskih programa. U praksi se računalna lingvistika može primjenjivati npr. kod prijevoda i provjere ispravnosti stručne i slične literature, raznih prospekata, uputa, deklaracija proizvoda i sl. Cjelokupno ljudsko znanje stečeno kroz povijest civilizacije je ispisano odnosno jezički zabilježeno, pa se “ljudsko znanje i izražava jezikom. Zbog toga je računalna lingvistika vrlo važna” (Steedman, 2007.).

Računalni lingvist mora biti stručno osposobljen u područjima strojnog učenja, dubokog učenja, umjetne inteligencije, kognitivnog računarstva i neuroznanosti.

Ciljevi računalne lingvistike uključuju sljedeće (Gillis & Bernstein, 2022):

- stvaranje gramatičkih i semantičkih okvira posebnosti jezika,
- prijevod tekstova s jednog jezika na druge,
- pristup tekstovima koji se odnose na točno određenu temu,
- analiziranje teksta ili govornog jezika radi stavljanja riječi u kontekst kao i tumačenja osjećaja ili drugih emotivnih kvaliteta,
- odgovaranje na pitanja, uključujući i ona zahtjevna koja podrazumijevaju zaključivanje te opisne ili diskurzivne odgovore, odnosno odgovore utemeljene na razmišljanju i argumentima, a ne sročene pomoću intuicije,
- sažimanje teksta u manje smislenu cjelinu,
- izgradnja strojnih agenata - posrednika sposobnih za izvršavanje složenih zadataka poput planiranja kupnje, putovanja ili nekih događaja,
- stvaranje chatbotova koji mogu proći Turingov test.

Računalna lingvistika i obrada prirodnog jezika slični su pojmovi jer koriste iste alate, poput strojnog učenja i umjetne inteligencije kako bi postigli svoje ciljeve. Obrada prirodnog jezika se bavi računalnim programima radi razumijevanja ljudskog jezika, a računalna lingvistika se

fokusira na računalni opis jezika. Računalna lingvistika treba dati odgovor na jezična pitanja uz pomoć računalnih alata. Većina radnji u računalnoj lingvistici ima i teorijske i praktične elemente, a usmjerena je na poboljšanje kvalitete odnosa između računala i osnovnih jezičnih postavki. To rezultira izgradnjom umjetnih tvorevina koje se mogu koristiti za obrađivanje i stvaranje jezika. Izrada takvih tvorevina zahtijeva od programera podataka da analizira ogromne količine pisanog i govornog jezika u strukturiranim i nestrukturiranim formatima (Gillis & Bernstein, 2022).

Primjene računalne lingvistike obično uključuju sljedeće stručne pojmove (Gillis & Bernstein, 2022):

- **strojno prevođenje**, odnosno proces korištenja umjetne inteligencije za prevođenje s jednog ljudskog jezika na drugi,
- **grupiranje aplikacija**, odnosno proces pretvaranja više računalnih poslužitelja u klaster,
- **analizu sentimenta** koja znači pristup obradi prirodnog jezika koji identificira emocionalni ton u sadržaju teksta,
- **chatbotove** koji su softveri ili računalni programi koji simuliraju ljudski razgovor kroz tekstualne ili glasovne interakcije,
- **ekstrakciju znanja** koja služi za usvajanje znanja iz strukturiranog i nestrukturiranog teksta,
- **sučelja prirodnog jezika** koji su računalno-ljudska sučelja gdje riječi, fraze ili klauzule djeluju kao kontrole korisničkog sučelja,
- **filtriranje sadržaja** koje predstavlja proces blokiranja različitih web-sadržaja koji se temelje na jeziku, a onemogućuje im se da dođu do krajnjih korisnika.

Dakle, mnogo je različitih pristupa i metoda računalne lingvistike, a primjeri nekih računalnih lingvističkih pristupa uključuju sljedeće (Gillis & Bernstein, 2022):

- **pristup temeljen na tijelu** koji se temelji na jeziku kakav se u praksi koristi,
- **pristup razumijevanja** koji mehanizmu obrade prirodnog jezika omogućuje tumačenje zapisanih naredbi u jednostavnom okruženju upravljanim pravilima,
- **razvojni pristup** putem kojeg se usvaja strategija učenja jezika poput djeteta, odnosno usvajanje jezika tijekom vremena, a razvojni proces ima racionalni pristup proučavanju jezika i ne uzima u obzir gramatičku ispravnost,

- **strukturni pristup** ima teorijski pristup strukturi jezika i koristi velike uzorke jezika koji se provlače kroz modele računalne lingvistike kako bi se moglo bolje razumjeti temeljne jezične strukture,
- **proizvodni pristup** fokusira se na računarsko lingvistički model za proizvodnju teksta i uključuje konstrukciju algoritama koji proizvode tekst na temelju tekstova koji proizlaze iz ljudskog govora,
- **interaktivni pristup temeljen na tekstu u** kojem se tekst iz čovjekovog govora koristi za generiranje budućih odgovora pomoću algoritama. Računalo može prepoznati različite obrasce govora i pronaći odgovor na temelju korisničkog upita i navedenih ključnih riječi,
- **interaktivni pristup utemeljen na govoru** funkcionira slično kao pristup utemeljen na tekstu, ali korisnički unos vrši putem prepoznavanja govora. Govorni unos korisnika prepoznaje kao zvučni val, a sustav računalne lingvistike tumačit će ga kao buduće uzorke.

## 6. Prepoznavanje glasa na primjeru Google asistenta

### 6.1. Digitalna pomagala i Google asistent

Potrošački proizvodi koji su razvijeni u posljednjih nekoliko godina, doveli su do pojeftinjenja glasovnih asistenata koji se koriste u svakodnevnoj upotrebi, a stalni razvoj će još više povećati njihov značaj i upotrebu kod korisnika. Najpopularniji glasovni asistenti su Appleov Siri, Amazonova Alexa, Microsoftova Cortana i Googleov asistent i svi se već ugrađuju u pametne telefone i aparate (Hoy, 2018). Programski jezik C++ nije jedini koji se koristi kod Google asistenta već se koriste i jezici kao što su Java i Python, koji se također široko koriste zbog svoje jednostavnosti i lakog razumijevanja kodiranja (Google Developers, 2020). Bitno je znati da tehnologija prepoznavanja govora ustvari koristi proces pretvaranja zvuka u tekst s namjerom obrade podataka u svrhu uspostavljanja razgovora između umjetne inteligencije i korisnika. Prepoznavanje govora može se podijeliti u sljedeće tri faze (McLaren, 2021):

- automatsko prepoznavanje govora tj. transkripcija govora,
- obrada prirodnog jezika, odnosno izvođenje značenja iz ranije prikupljenih podataka i naknadno obavljanje transkripcije odnosno prijevodi u tekst,
- tekst u govor kojim se tekst pretvara u ljudski govor.

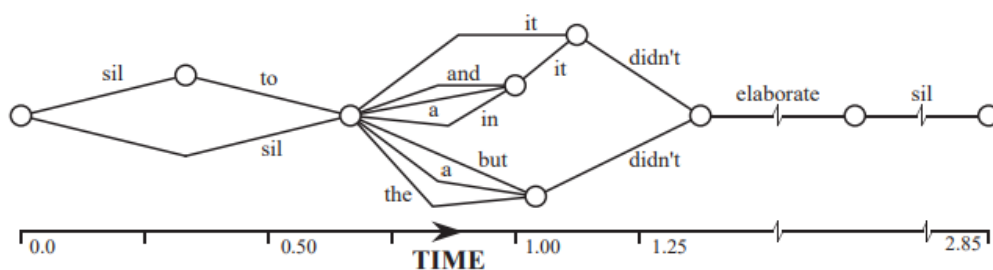
Proces započinje digitalizacijom snimljenog uzorka govora uz pomoć softvera za prepoznavanje glasa. Govornikov jedinstveni tijek glasa razbija se u diskretne segmente sastavljene od nekoliko tonova koji se vizualiziraju u obliku spektrograma (vizualni prikaz spektra frekvencija glasovnog signala koji se mijenja s vremenom). Spektrogrami se dalje dijele na vremenske korake pomoću kratkotrajne Fourierove transformacije kojom se određuje sinusna frekvencija i fazni sadržaj lokalnih dijelova signala. Transformirani spektrogrami se mogu zatim koristiti u softverski definiranom radiju. Softverski definirani radio je sustav radiokomunikacije u kojem su komponente koje su obično implementirane u analogni hardver (npr. mikseri, filteri, razna pojačala, pretvarači, senzori i sl.) zamijenjene sa softverom na osobnom računalu ili se ugrađuju u drugi sličan sustav (McLaren, 2021).

Dobiveni podaci se analiziraju i transkriptiraju uz pomoć obrade prirodnog jezika, algoritma koji može predvidjeti vjerojatnost svih izgovorenih riječi u odnosu na ukupni rječnik tog jezika. Zatim se dodaje i sloj koji pomaže u ispravljanju potencijalnih kontekstualnih pogrešaka. Ovdje algoritam uzima u obzir ono izrečeno, a sljedeću riječ pronalazi prema vjerojatnosti njenog pojavljivanja na temelju poznavanja rječnika tog jezika. Na kraju će uređaj glasovno uobličiti najbolji mogući odgovor, na ono što je čuo, korištenjem metode za pretvorbu teksta u govor (McLaren, 2021).

Google asistent koristi sustav uočavanja ključnih riječi (eng. keyword spotting) koji koristi kontekstualno automatsko prepoznavanje govora (eng. automatic speech recognition). Uređaji koji se koriste glasovnom aktivacijom uobičajeno koriste sustave uočavanja ključnih riječi kako bi korisniku mogli omogućiti brzu aktivaciju asistenta putem fraze okidača (npr. "OK Google"). Nakon što se detektira aktiviranje asistenta uređaj čeka daljnje upute od korisnika. Audio zapis putem automatskog prepoznavanja transkribira govor te semantički obrađuje taj glas kako bi mogao izgenerirati odgovor koji će se poslati natrag preko uređaja (Michaely i sur., 2018).

Asistent koristi semantičke informacije unutar rešetke riječi (eng. word lattice) (slika 6.) te korištenjem prepoznavanja imenovanih entiteta (eng. Named Entity Recognition) provodi analizu, kako bi uspješno prepoznao i odabrao kontekstualno relevantne puteve i poboljšao točnost prepoznavanja govora. Zbog poboljšanja točnosti prepoznavanja, Google asistent se koristi kontekstualnim prepoznavanjem govora, koje se koristi signalima u stvarnom vremenu, proizašlim iz konteksta govornika. Prilikom dovođenja tih signala u smisao i kontekst za prepoznavaća,

pojavljaju se razlike između njega, koji prilagođava glas prosječnog govornika izrečen u normalnim uvjetima glasu prilagođenom subjektivnim i kontekstualnim okolnostima. Asistent se koristi bodovanjem jezičnih modela. Bodovanjem se prilagođavaju vjerojatnosti pojave jezičnih modela u stvarnom vremenu na temelju kontekstualnih signala, nakon čega se dopuštaju ciljane prilagodbe i to bez potrebne obuke za specifični kontekstualni jezični model. U ranije navedenom, radilo se na uključivanju raznih kontekstualnih signala u automatsko prepoznavanje glasa, što uključuje tip uređaja koji se koristi, povijest govornikovih upita ili radnji, stanje uređaja, lokaciju i dijaloško stanje (Velikovich i sur., 2018).



Slika 6: word lattice (Woodl i sur.)

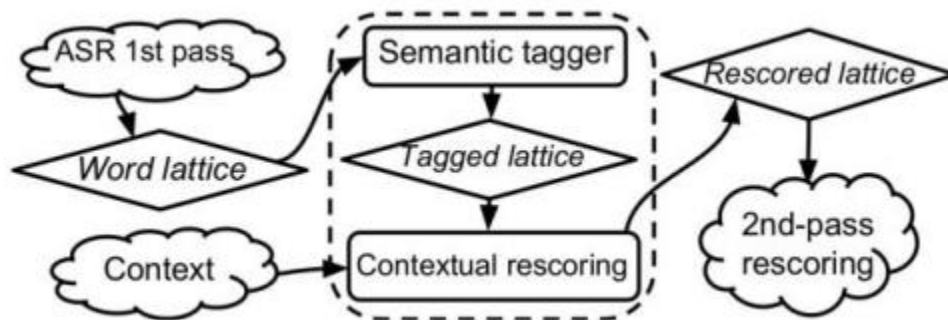
Kontekstualni sustav za automatsko prepoznavanje glasa koristi dvo-prolazni (eng. two-pass) prepoznavач kontinuiranog prepoznavanja govora iz velikog rječnika (eng. large-vocabulary continuous speech recognition) koji na prvom prolazu sadrži dekodер prvog prolaza govornog modela, a na drugom prolazu sadrži precizniji model govornog jezika za ocjenu drugog prolaza. Dekoder je naprava koja pretvara audio ili video signale u nešto drugo, u ovom slučaju pretvara audio signale u spektrograme. Sustav za automatsko prepoznavanje glasa može u svom radu koristiti n-gramski model ili neuronske mreže. Zahtjev koji dekodер treba ispuniti je ispisivanje skupa mogućih hipoteza u obliku rešetke riječi, a to je učinkovito ponderirana automatizacija (eng. weighted automation) reprezentacije govornih hipoteza. Ranije navedene metode često koriste ponderirane kontekstualno relevantne n-gramе te mogu sadržavati kontekstualno dodane dinamičke klase. Takvi sustavi u hodu izvode ponovno izračunavanje n-grama tijekom dekodiranja prvog prolaza (Velikovich i sur., 2018).

Kao što je prikazano na slici 7, kod prvog prolaza dekodiranja automatskog prepoznavanje riječi vraća se rešetka riječi, odnosno aciklički ponderirani graf nad riječima koji učinkovito kodira predložene hipoteze korisnikova iskaza. Uzimajući ovu rešetku riječi kao ulaz, označivač



semantičkih rešetki (eng. semantic lattice tagger) ubacuje oznake dekoratora (eng. decorator tags) otvaranja/zatvaranja oko vjerojatnih kontekstualno relevantnih semantičkih entiteta u rešetku. Svaki umetnuti dekoratorski par označava semantičku klasu (npr. <glazbenik> </glazbenik> ili <pjesma></pjesma>). Višestruka (dvosmislena) semantička tumačenja su dopuštena i izražena kao paralelni putevi kroz iste riječi i okruženi su različitim dekoratorskim oznakama (Velikovich i sur., 2018).

Sljedeći korak je ponovno ocjenjivanje semantički označene rešetke na temelju konteksta specifičnog za korisnika. Nagrađuju se hipoteze koje sadrže smislenu frazu koja uključuje predviđenu klasu semantičkih entiteta. Npr. kod zadatka prepoznavanja govora namijenjenog reprodukciji glazbe mogao bi tražiti <song> ili <musical\_artist> i ponovno ocijeniti "play \$SONG" ili "play \$SONG by \$MUSICAL\_ARTIST". To bi poboljšalo kvalitetu prepoznavanja u scenarijima gdje su takvi unosi vjerojatniji od prosjeka. Slično koraku semantičkog označavanja, ponovno ocjenjivanje mora biti učinkovito na cijeloj rešetki. Uz primijenjeno semantičko ocjenjivanje, kod prepoznavanja, modificirana rešetka se prosljeđuje na ocjenjivanje drugog prolaza. Od drugog prolaza se ne očekuje da zna za semantičke oznake dekoratora pa se one mogu ili ukloniti ili privremeno sakriti (Velikovich i sur., 2018).



Slika 7: Kontekstualno automatsko prepoznavanje glasa za ocjenjivanje semantičke rešetke riječi (Velikovich i sur., 2018)

Google asistent, ali i puno drugih softvera govorne tehnologije sve se više oslanjaju na duboke neuronske mreže, odnosno vrstu strojnog učenja koja sve više pomaže u izgradnji točnijih i bržih modela prepoznavanja govora. Duboke neuronske mreže najčešće trebaju veće količine podataka kako bi dobro funkcionirale i poboljšavale se tijekom vremena, a taj proces poboljšanja naziva se obuka modela. Googleovi developeri za razvoj asistenta koriste 3 vrste tehnologija umjetne inteligencije za obuku govornih modela, a to konvencionalno učenje (eng. conventional learning),

federirano učenje (eng. federated learning) i efemerano učenje (eng. ephemeral learning). Ovisno o zadatku i situaciji, neki od njih su učinkovitiji od drugih, a u određenim slučajevima koriste se i kombinirano što omogućuje visoku kvalitetu te istovremeno zaštitu privatnosti (Google Developers, n.d.).

Konvencionalno učenje je najčešći oblik koji se koristi za obučavanje govornih modela. Uz korisnikov izričiti pristanak, audio uzorci se prikupljaju i pohranjuju na Googleove servere, a ljudski recenzenti vrednuju i označavaju određene audio uzorke te algoritam za obuku uči iz njih. Razlikuju se 2 vrste konvencionalnog učenja i to obuka pod nadzorom i nenadzirana obuka. Kod obuke pod nadzorom modeli se obučavaju tako da oponašaju obilježene zvukovne podatke recenzentata. Kod nenadzirane obuke stroj označava korisne podatke umjesto recenzenta. Kada se obučava na jednakim količinama podataka, nadzirana obuka obično rezultira boljim modelima prepoznavanja govora nego nenadzirana obuka jer je recenziranje kvalitetnije. Nadalje, nenadzirana obuka može učiti iz više audio uzoraka budući da uči iz strojnih komentara koje je lakše i brže proizvesti (Google Developers, n.d.).

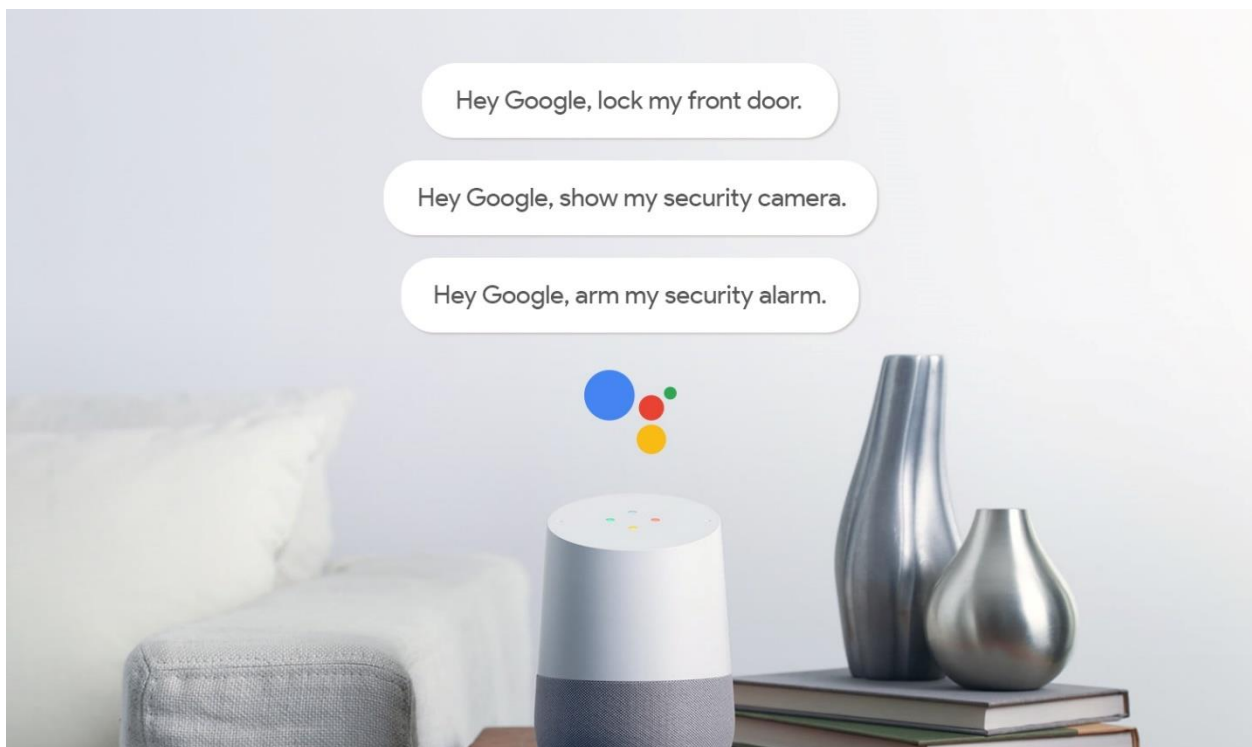
Tehnika federiranog učenja je razvijena u Googleu kao tehnika za očuvanje zaštite privatnosti kod obuke modela umjetne inteligencije. Federirano se učenje koristi kod obuke govornog modela, kada na uređaju korisnika postoje podaci dostupni za učenje iz modela. Federiranim učenjem se algoritmi obučavaju i to bez slanja korisnikovih audiopodatka Googleovim serverima, već se ti podaci spremaju na korisnikovom uređaju. Novi govorni modeli se formiraju kombiniranjem ukupnog učenja svih uključenih uređaja na korisnikovom ekosustavu (Google Developers, n.d.).

Efemerano učenje je tehnika očuvanja zaštite privatnosti koja se koristi kada se model govora izvodi na Googleovim serverima. Dok Googleovi sustavi pretvaraju dolazne audio uzorke u tekst, ti se uzorci šalju u RAM. Dok su podaci u RAM-u, algoritam za obuku uči iz tih audio podataka u stvarnom vremenu. Ti uzorci se brišu iz kratkoročne memorije u roku od nekoliko minuta kako bi održali zaštitu privatnosti korisnika (Google Developers, n.d.).

Razvoj glasovnih asistenata može se usporediti s rođenjem i razvojem djeteta koje se rađa kao neispisana ploča, ali daljnjim slušanjem riječi oko sebe ono upija verbalne znakove, intonaciju, ton i naglasak. Na taj način stvaraju se veze i obrasci koji poboljšavaju slušanje, razumijevanje i otkrivanje novih obrazaca. Nakon nekog vremena dolazi do sposobnosti za vođenje potpunog i samostalnog razgovora. Potom slijedi prikupljanje još više dodatnih podataka i povećanje obima

rječnika što će u konačnici rezultirati trenutnim tumačenjem govora i davanjem najboljih mogućih odgovora (McLaren, 2021). Još je dalek put razvoja na kojem će glasovni asistent morati i moći napraviti puno civilizacijskih iskoraka kako bi funkcionirali neovisno i bez pomoći i podrške čovjeka jer “beba uči puzati, hodati i zatim trčati. U fazi smo puzanja kada je riječ o primjeni strojnog učenja” (Jaiswal, 2021).

Googleovo prepoznavanje glasa je dobar i popularan primjer kako funkcionira tehnologija prepoznavanja glasa. Google asistent je popularna Googleova aplikacija koja je dio skoro svakog današnjeg pametnog android telefona i koju koriste drugi pametni uređaji poput satova, smart home uređaja, automobila itd. On može obavljati pozive, uzimati bilješke, kontrolirati pametne uređaje itd. (Tillman, 2022). Pomoću poziva na ključnu riječ “Hey Google” ili pokretanjem same aplikacije pa pritiskom na tipku namijenjenu za pokretanje asistenta, obrađuje se poruka, a nakon svega toga se informacija šalje krajnjem korisniku, odnosno izvršava se radnja kao što je to prikazano na slici 8 (Forrest i sur., 2021).



Slika 8: Davanje naredbi Google asistentu (Turner, 2018a)

## 6.2. Obrada funkcionalnosti Google asistenta

Google asistent je vrlo modularan te postoji mogućnost izrade programa koji podržavaju njegove funkcionalnosti. Google asistent može pomoći kod odgovaranja na bilo koju informaciju koja se želi saznati ili kod obavljanja bilo koje funkcionalnosti na mobilnom telefonu.

Neke od najpopularnijih funkcionalnosti bi bile (Tillman, 2022):

- Prognoza vremena (Kakvo će sutra biti vrijeme?)
  - Daje informacije o vremenu u predstojećem periodu od najviše tjedan dana, ali je moguće pitati ga za određeni datum u tom periodu.
- Navigacijske usluge (Pronađi najkraći put do Venecije)
  - Davanje informacija na putovanjima i u obilascima
  - Davanje informacija kod nalaženja i preporuke određenih ugostiteljskih, kulturnih, sportskih destinacija
- Obavljanje poziva (Nazovi Maju)
  - Biranje broja i pozivanje iz telefonskog imenika
  - Može se obaviti i kroz ostale aplikacije koje nude tu opciju (viber, whatapp, facebook messenger itd.)
- Slanje poruka (Pošalji poruku Alenu “Kako si?”)
  - Biranje osobe iz imenika ili glasovno unošenje informacija o kontaktu i slanje poruke prema njemu
  - Može se obaviti i kroz ostale aplikacije koje nude tu opciju (viber, whatapp, facebook messenger, gmail itd.)
- Postavljanje budilice (Probudi me u 8 sati ujutro)
  - Postavljanje zvučne obavijesti u točno određeno vrijeme
- Postavljanje bilješki za buduće obveze (Prisjeti me na sastanak u 14 sati)
  - Evidentira buduće obveze i podsjeća korisnika na taj događaj
- Sviranje muzike (Pusti neku stranu pjesmu s liste)
  - Pokreće željenu pjesmu ovisno o danom opisu
  - Pokreće pjesmu iz playliste, youtubea, spotifya, deezera itd.
- Sportski događaji (Je li Istra 1961 pobijedila na svojoj zadnjoj utakmici?)

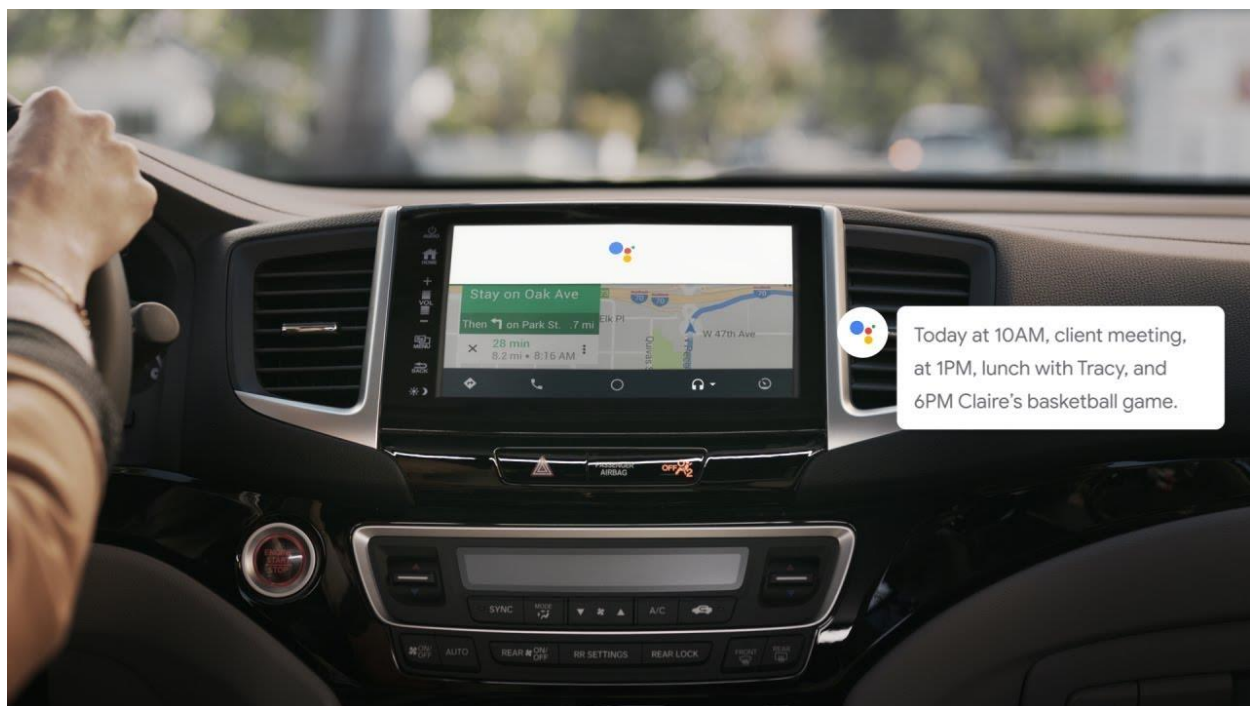
- Daje vijesti o sportskim događajima
- Definicije riječi (Definiraj “Asistent”)
  - Objašnjava pojam i značenje pojedine riječi
- Prijevod na drugi jezik (“Kako je lijep dan” na Engleskom?)
  - Daje prijevod riječi na određeni jezik
- Kontroliranje postavki na mobitelu (Upali Wi-Fi)
  - Kontroliranje funkcionalnosti na mobilnom uređaju
- Pokretanje aplikacija (Pokreni google play trgovinu)
  - Pokretanje aplikacija instaliranih na uređaju
- Pokretanje igara (Igrajmo se igre)
  - Google asistent ima puno simpatičnih dodataka koje nisu možda toliko korisne, ali pružaju zabavu
  - Neke popularne naredbe bi bile kristalna lopta, igranje kviza i još puno popularnih igara koje se mogu naknadno preuzeti
- Upravljanje pametnim aparatima (Zagasi svijetla u kupaonici)
  - Omogućuje kontrolu pametnih uređaja putem asistenta

### **6.3. Prednosti, nedostaci i perspektive Google asistenta**

Moderna tehnologija, u koju spada i Google asistent, se unapređuje svakim danom kako bi nam olakšala život i rad. Ona može kontrolirati uređaje i pametni dom, odnosno može uključivati/isključivati npr. ventilator, audio sustavu, klima uređaja itd. Može pristupiti kalendaru u kojem podsjeća na važne datume. Glasovnom naredbom mogu se pretraživati pojmovi koje tražimo i čitati određene sadržaje. Pomoću njega se može nazvati kontakt broj ili odgovoriti na poziv glasovnom naredbom. Poruke se mogu slati i glasovima, a mogu se i otvarati aplikacije na pametnim uređajima i pametnim telefonima. On može čitati obavijesti umjesto nas ili dobiti govorni prijevod u stvarnom vremenu. Google sve više poboljšava asistenta kako vrijeme prolazi. Dakle, dostupno je još puno novih značajki osim gore navedenih, koje se mogu isprobati i doživjeti. Google pomoćnik može sinkronizirati preferirane pojmove za pretraživanje na uređajima kao što su računala, pametni telefoni, Google Home i druge platforme. Dakle, pomaže u brzom pronalaženju upita i učenju pomoću umjetne inteligencije te olakšava brže izvršavanje zadataka

na svim platformama. Može sinkronizirati i svoju povijest pretraživanja. Google pomoćnik može saznati vaše postavke i pomoći u uštedi vremena kod prebacivanja između uređaja. Sve što se treba učiniti je stvoriti Google račun i prijaviti se na svoj uređaj na drugoj platformi, a na kraju treba uključiti opciju sinkronizacije kako bi se sve vaše informacije lako sinkronizirale na uređajima kada ste povezani na internetu (Tech avatar, 2021).

Mnoge automobilske tvrtke surađuju s Googleom i poboljšavaju svoje usluge u svojim novim automobilima. Google asistent može razgovarati s vama dok upravljate vozilom te vam pokazati vašu lokaciju putem Google karta. Asistent povratno šalje audio informacije te ne ometa tok vožnje, kao što je prikazano na slici 9 (Tech avatar, 2021).



Slika 9: Korištenje asistenta u autu (Kinsella, 2018)

Iako Google asistent oponaša glasove čovjeka, oni nisu svaki put jasni, jer se zna desiti da ponekad čujemo šumove i prekide zvuka tijekom izvođenja naredbe. Podsjeća nas na komunikaciju s robotima, a ne sa živim pomoćnikom. Cilj asistenta bio je oponašanje interakcije sa stvarnom osobom, ali još uvijek ne postoji takvo iskustvo i ne dobiva se osjećaj komunikacije kao sa stvarnom osobom. Međutim, poboljšanja su svakim danom sve veća i veća i u budućnosti će vjerojatno biti i bolja nego što smo očekivali (Khanzode & Sarode, 2020).

Googleov Duplex pomaže asistentu da vodi prirodan razgovor oponašajući ljudski glas, što slično robotskom pozivu. Pomoćnik može pozivati, gasiti kućanske aparate i još mnogo toga. Može čak komunicirati poput ljudi pa čak i koristiti uzvike poput "amhh", "mm-Hmm" i "ahh". Postoji mnogo zabrinutosti vezano uz privatnost kod duplexa. Samim time što dajemo privatne podatke mogla bi biti ugrožena naša sigurnost i privatnost. Čak se izražava i zabrinutost zbog glasovnih uzoraka koji se snimaju i koriste za potrebe virtualnih pomoćnika. Za sada asistent ne nudi puno značajki koje bi u potpunosti pomogle zadovoljiti potrebe korisnika. Korisničko iskustvo nije baš skroz zadovoljavajuće i nije vrlo intuitivno. Dakle, Google asistent u budućnosti zahtijeva više nadogradnji kako bi zadovoljio širu masu korisnika (Khazode & Sarode, 2020).

U razgovoru s asistentom ljudi pokušavaju nešto obaviti. Velika vjerojatnost je da će upiti asistenta biti više usmjereni na akciju nego pretraživanje, a ljudi su više usmjerni na davanje naredbi poput "pošalji poruku", "ugasi svijetlo" ili "uključiti bluetooth" (Khazode & Sarode, 2020).

Neke zadatke, pogotovo kada smo u pokretu puno je lakše obaviti putem glasa. Npr. zadatci paljenja svjetla ili puštanja glazbe se ispunjavaju bez paljenja aplikacije. Osnovne stvari kao što su pozivnice i kalendar, se mogu odraditi bez gledanja u telefon ili prekida rada na način da kažemo npr. zakaži termin za rođendan u nedjelju u 22 sata. Te male stvari koje radimo pomoću glasa omogućavaju da izvršavamo zadatak na lak i jednostavan način. Što više akcija bude pruženo korisniku i izgradnjom što više ovakvih vrsta iskustava, bit ćemo bliže idealnom pomoćniku (Khazode & Sarode, 2020).

Asistentima se koristimo uz vrlo jednostavne naredbe, ali koje vrlo brzo znaju narasti u složeni dijalog. Pretraživanje weba se vrši općenitim ključnim pojmovima, dok asistentom tražimo informaciju u obliku upita ili razgovorima poput "kakvo će vrijeme sutra biti u Puli". U prosjeku, upiti Asistenta su 200 puta više razgovorni od upita pretraživanja. Jednostavne naredbe se mogu izvesti na puno različitih načina npr. Alarm se može postaviti na više od 5000 različitih načina, što znači da kako bi dobili boljeg pomoćnika on bi trebao znati razumjeti sve moguće kombinacije razgovora. Potpuno prebacivanje na glas se ne očekuje u skoroj budućnosti. Zaslone donose obogaćivanje razgovorne umjetne inteligencije koja spaja glas i dodir radi obostranog poboljšanja. Primjer bi bio pretraživanje recepta za pizzu, kada se prikazuje kako bi pizza trebala izgledati dok Asistent još usput čita i korake. Uvođenjem pametnih zaslona i našim novim vizualnim iskustvom

na telefonima, razvit će se Google pomoćnik koji će postati puno brži, dinamičniji, obuhvaćati bolje glas, zaslone, dodirivanje i tipkanje (Khazode & Sarode, 2020).

Asistentu se može uvijek pristupiti u svako vrijeme i na svakom mjestu, putem telefona, u automobilu ili na zvučniku u dnevnoj sobi. Ljudi koriste asistenta u svom okruženju i stvaraju dnevnu rutinu korištenja. Npr. ujutro se pokrenu pametni zvučnici za pružanje informacija o prognozi vremena ili pregledu najnovijih vijesti. Popodne ili na putu koristimo poruke i zovemo prijatelje te tražimo navigaciju za pronalazak lokalnih restorana. Navečer slušamo glazbu ili postavimo alarm i podsjetnik o sutrašnjim sastancima i aktivnostima. Vrijeme i mjesto korištenja asistenta varira tijekom dana, ali praksu i iskustvo trebamo zadržati kontinuirano (Khazode & Sarode, 2020).

Najuzbudljivija stvar kod digitalnih asistenata je nova tehnologija koju je sve lakše usvojiti pa čak i bez uputa o korištenju i bez obzira na dob ljudi koji koriste uređaje. Zbog toga raste broj korisnika i među starijim osobama i obiteljima, a zanimljivo je da je porast korištenja izrazito zapažen kod žena. Glas je također postao primjenjiv na globalnoj razini odnosno u sve više zemalja i na više jezika. U siromašnim zemljama trećeg svijeta gdje se ljudi tek počinju susretati s internetom, upotreba Google asistenta ubrzano raste, ali činjenica je da se glasovna tehnologija još uvijek razvija i još uvijek se promatra da li će i koliko olakšati obavljanje posla i tako postići veliki uspjeh (Khazode & Sarode, 2020).

## **7. Zaključak**

Glasovni pomoćnici s umjetnom inteligencijom mijenjaju naše živote svaki dan. Iako je za većinu ljudi to luksuz, on je s vremenom sve dostupniji zahvaljujući glasovnim asistentima. Iako je glasovna tehnologija još uvijek nova, ona se temelji na tehnologiji koja će s vremenom napredovati, rasti i povećavati korisnost u osobnom i poslovnom okruženju. Ljudi postaju sve otvoreniji prema novoj tehnologiji pa će ljudski glas kao praktični način komuniciranja s tehnologijom vrlo brzo biti usvojen u svim dobnim skupinama. Najveći porast primjene pronaći će u zdravstvu i automobilske industriji. Razvojem umjetne inteligencije glasovni asistenti će evaluirati i preuzeti interakciju između čovjeka i stroja te to do te mjere da će moći gotovo čitati naše misli, znati što tražimo te izbaciti upotrebu ruku, koje će zamijeniti naš glas kao daljinski upravljač, što će utjecati na naše živote. Glasovni pomoćnici će odvući ljude s ekrana. Najčešći



nedostatak glasovnog asistenta je to što nakon pitanja odgovor se vraća kao činjenica bez navođenja izvora. Za dodatni izvor informacija moramo postaviti dodatna pitanja. Još jedan problem su sigurnosni problemi koji nastaju ako neko daje glasovne naredbe umjesto nas ili neovlašteno vrše naredžbe u naše ime od čega se može zaštititi jedino opcijom potvrde pina. Postoji mogućnost neovlaštenog pristupa podacima preko glasovnih naredbi.

## Literatura

1. FAWZY, B. (2020, SEPTEMBER 14). THE TECHNOLOGIES BEHIND GOOGLE ASSISTANT. SCIPLANET. RETRIEVED SEPTEMBER 15, 2022, FROM <https://www.bibalex.org/SCIplanet/en/Article/Details.aspx?id=13640>
2. MARKOTIĆ, K. (2021, MAY 26). UMJETNA INTELIGENCIJA (AI) - SVE ŠTO TREBATE ZNATI. MACHINEDESK. <https://www.machine-desk.com/industrija-4-0/umjetna-inteligencija-ai>
3. BHBOSEALE, S., PUJARI, V., & MULTANI, Z. (2020, FEBRUARY 29). ADVANTAGES AND DISADVANTAGES OF ARTIFICIAL INTELLEGENGE. AAYUSHI INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL, 77, 227-230.
4. DATAROBOT. (2020, AUGUST 11). SEMI-SUPERVISED LEARNING. RETRIEVED SEPTEMBER 30, 2022, FROM <https://www.datarobot.com/blog/semi-supervised-learning/>
5. FORREST, C., VIGLIAROLO, B., & LOTZE, K. (2021, MAY 19). GOOGLE ASSISTANT: A CHEAT SHEET. TECHREPUBLIC. RETRIEVED OCTOBER 1, 2022, FROM <https://www.techrepublic.com/article/Google-assistant-the-smart-persons-guide/>
6. AKINBI, A., & BERRY, T. (2020, AUGUST 14). FORENSIC INVESTIGATION OF GOOGLE ASSISTANT. SPRINGERLINK. RETRIEVED SEPTEMBER 15, 2022, FROM [https://link.springer.com/article/10.1007/s42979-020-00285-x?error=cookies\\_not\\_supported&error=cookies\\_not\\_supported&error=cookies\\_not\\_supported&code=7e2fb65d-2987-44e9-8793-d3c95dbdd7fa&code=f3fe01eb-017d-43df-ae81-a1a8af7ccfe3&code=d27550f1-6445-4408-b6bc-6967a252d6f8](https://link.springer.com/article/10.1007/s42979-020-00285-x?error=cookies_not_supported&error=cookies_not_supported&error=cookies_not_supported&code=7e2fb65d-2987-44e9-8793-d3c95dbdd7fa&code=f3fe01eb-017d-43df-ae81-a1a8af7ccfe3&code=d27550f1-6445-4408-b6bc-6967a252d6f8)
7. GILLIS, A. S., & BERNSTEIN, C. (2022, APRIL). COMPUTATIONAL LINGUISTICS (CL). SEARCHENTERPRISEAI. <https://www.techtarget.com/searchenterpriseai/definition/computational-linguistics-CL>

8. WIKIPEDIA CONTRIBUTORS. (2021, DECEMBER 22). TURINGOV TEST. WIKIPEDIA.  
[https://hr.wikipedia.org/wiki/Turingov\\_test](https://hr.wikipedia.org/wiki/Turingov_test)
9. McLAREN, I. (2021, JULY 12). HOW DOES SPEECH RECOGNITION TECHNOLOGY WORK?  
SUMMA LINGVAE. <https://summalinguae.com/language-technology/how-does-speech-recognition-technology-work/>
10. HOY, M. B. (2018, JANUARY). ALEXA, SIRI, CORTANA, AND MORE: AN INTRODUCTION TO  
VOICE ASSISTANTS. MEDICAL REFERENCE SERVICES QUARTERLY, 37(1), 81–88.  
<https://doi.org/10.1080/02763869.2018.1404391>
11. TECH AVATAR. (2021, JULY 21). GOOGLE ASSISTANT: PROS AND CONS. TECH INFO. |.  
RETRIEVED OCTOBER 1, 2022, FROM <https://yourtechavatar.com/google-assistant-pros-and-cons/>
12. HUFFMAN, S. (2018, AUGUST 21). FIVE INSIGHTS ON VOICE TECHNOLOGY. GOOGLE.  
<https://www.blog.google/perspectives/scott-huffman/five-insights-voice-technology/>
13. WIKIPEDIA. (2021, AUGUST 11). RAČUNARSKA LINGVISTIKA. RETRIEVED SEPTEMBER 27,  
2022, FROM [https://sh.wikipedia.org/wiki/Ra%C4%8Dunarska\\_lingvistika](https://sh.wikipedia.org/wiki/Ra%C4%8Dunarska_lingvistika)
14. POOLE, MACKWORTH AND GOEBEL. (1998) WHICH PROVIDES THE VERSION THAT IS USED IN  
THIS ARTICLE. NOTE THAT THEY USE THE TERM COMPUTATIONAL INTELLIGENCE AS A  
SYNONYM FOR ARTIFICIAL INTELLIGENCE, PP. 1.
15. BUDEK, K., & OSIŃSKI, B. (2018, JULY 05). DEEPSense.AI. WHAT IS REINFORCEMENT  
LEARNING? THE COMPLETE GUIDE. RETRIEVED SEPTEMBER 18, 2022, FROM  
<https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/>
16. STEEDMAN, MARK, 2007., “PRESIDENTIAL ADDRESS TO THE 45TH ANNUAL MEETING OF  
ACL”, PRAG, LIPANJ 2007. TAKOĐER TISKANO 2008., “ON BECOMING A  
DISCIPLINE”, COMPUTATIONAL LINGUISTICS , 34(1): 137–144. [ STEEDMAN 2007 [2008]  
DOSTUPAN ONLINE ]

17. JAISWAL, R. (2021, JULY 14). DEEP LEARNING. MEDIUM. RETRIEVED OCTOBER 4, 2022, FROM <https://rachana3073.medium.com/deep-learning-48a77b25615a>
18. GAËL. (2020, JANUARY 7). WHAT IS ARTIFICIAL INTELLIGENCE ? DATAKEEN. RETRIEVED OCTOBER 1, 2022, FROM <https://www.datakeen.co/en/what-is-artificial-intelligence/>
19. CURTIN, M. (N.D.). IN AN 8-HOUR DAY, THE AVERAGE WORKER IS PRODUCTIVE FOR THIS MANY HOURS. INC.COM. RETRIEVED SEPTEMBER 25, 2022, FROM <https://www.inc.com/melanie-curtin/in-an-8-hour-day-the-average-worker-is-productive-for-this-many-hours.html>
20. CODECNETWORKS. (2021, JULY 27). WHAT IS AI AND WHAT ARE THE STAGES OF ARTIFICIAL INTELLIGENCE?. BLOGS. <https://www.codecnetworks.com/blog/what-is-artificial-intelligence/>
21. RODRIGUEZ, J. (2020, JUNE 30). GÖDEL, CONSCIOUSNESS AND THE WEAK VS. STRONG AI DEBATE. MEDIUM. [https://jrodthoughts.medium.com/g%C3%B6del-consciousness-and-the-weak-vs-strong-ai-debate-31afea7e0a36?source=post\\_page-----70dccbfb944d---4-----](https://jrodthoughts.medium.com/g%C3%B6del-consciousness-and-the-weak-vs-strong-ai-debate-31afea7e0a36?source=post_page-----70dccbfb944d---4-----)  
-----
22. RISHABH, & PIYUSH (2020, JUNE 22). 3 TYPES OF MACHINE LEARNING. NEW TECH DOJO. RETRIEVED SEPTEMBER 18, 2022, FROM <https://www.newtechdojo.com/3-types-of-machine-learning/>
23. PIETIKÄINEN, M., & SILVEN, O. (2022, JANUARY 5). CHALLENGES OF ARTIFICIAL INTELLIGENCE -- FROM MACHINE LEARNING AND. . . ARXIV.ORG. RETRIEVED OCTOBER 4, 2022, FROM <https://arxiv.org/abs/2201.01466>

24. TURNER, M. (2018A, AUGUST 28). PROTECTING YOUR HOME WITH THE GOOGLE ASSISTANT. GOOGLE. RETRIEVED OCTOBER 4, 2022, FROM <https://www.blog.google/products/assistant/protecting-your-home-google-assistant/>
25. JOHNSON, D. (2022, SEPTEMBER 17). UNSUPERVISED MACHINE LEARNING: ALGORITHMS, TYPES WITH EXAMPLE. GURU99. RETRIEVED SEPTEMBER 18, 2022, FROM <https://www.guru99.com/unsupervised-machine-learning.html>
26. PETERSSON, D. (2021, MARCH). SUPERVISED LEARNING. SEARCHENTERPRISEAI. RETRIEVED SEPTEMBER 18, 2022, FROM <https://www.techtarget.com/searchenterpriseai/definition/supervised-learning>
27. IBM CLOUD EDUCATION. (2020A, AUGUST 19). SUPERVISED LEARNING. IBM. RETRIEVED AUGUST 19, 2022, FROM <https://www.ibm.com/cloud/learn/supervised-learning>
28. KINSELLA, B. (2018, JANUARY 10). GOOGLE ASSISTANT NOW AVAILABLE ON ANDROID AUTO. VOICEBOT.AI. RETRIEVED OCTOBER 4, 2022, FROM <https://voicebot.ai/2018/01/10/google-assistant-now-available-android-auto/>
29. GUPTA, K. (2017, SEPTEMBER 19). BEST PROGRAMMING LANGUAGE FOR VOICE RECOGNITION. DEVELOPERS, DESIGNERS & FREELANCERS – FREELANCINGGIG. RETRIEVED OCTOBER 4, 2022, FROM <https://www.freelancinggig.com/blog/2017/09/19/best-programming-language-voice-recognition/>
30. BROWN, A. (2021, SEPTEMBER 9). HOW DOES VOICE RECOGNITION WORK? MAKEUSEOF. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.makeuseof.com/how-does-voice-recognition-work/>

31. TURNER, M. (2018B, MAY 3). TURNING YOUR HOUSE INTO A SMART HOME WITH THE GOOGLE ASSISTANT. GOOGLE. RETRIEVED SEPTEMBER 18, 2022, FROM <https://blog.google/products/assistant/turning-your-house-smart-home-google-assistant/>
32. COMPUTER HOPE. (2021, JUNE 7). WHAT IS VOICE RECOGNITION? COMPUTER HOPE. RETRIEVED OCTOBER 4, 2022, FROM <https://www.computerhope.com/jargon/v/voicreco.htm>
33. BROWNLEE, J. (2020, APRIL 8). 4 TYPES OF CLASSIFICATION TASKS IN MACHINE LEARNING. MACHINE LEARNING MASTERY. RETRIEVED AUGUST 19, 2022, FROM <https://machinelearningmastery.com/types-of-classification-in-machine-learning/>
34. GOOGLE DEVELOPERS. (2022, JULY 18). WHAT IS CLUSTERING?. GOOGLE DEVELOPERS. RETRIEVED SEPTEMBER 18, 2022, FROM <https://developers.google.com/machine-learning/clustering/overview>
35. WIKIPEDIA CONTRIBUTORS. (2022, SEPTEMBER 5). MACHINE LEARNING IN VIDEO GAMES. WIKIPEDIA. RETRIEVED SEPTEMBER 18, 2022, FROM [https://en.wikipedia.org/wiki/Machine\\_learning\\_in\\_video\\_games](https://en.wikipedia.org/wiki/Machine_learning_in_video_games)
36. TILLMAN, M. (2022, MAY 10). WHAT IS GOOGLE ASSISTANT AND WHAT CAN IT DO? POCKET-LINT. RETRIEVED SEPTEMBER 18, 2022, FROM <https://www.pocket-lint.com/apps/news/google/137722-what-is-google-assistant-how-does-it-work-and-which-devices-offer-it>
37. DATAROBOT. (2016, AUGUST 11). WHAT IS NATURAL LANGUAGE PROCESSING? INTRODUCTION TO NLP. RETRIEVED SEPTEMBER 11, 2022, FROM <https://www.datarobot.com/blog/what-is-natural-language-processing-introduction-to-nlp/>

38. NABI, J. (2018, DECEMBER 22). MACHINE LEARNING — MULTICLASS CLASSIFICATION WITH IMBALANCED DATASET. MEDIUM. RETRIEVED SEPTEMBER 27, 2022, FROM <https://towardsdatascience.com/machine-learning-multiclass-classification-with-imbalanced-data-set-29f6a177c1a>
39. ANWAR, A. (2021, APRIL 10). A BEGINNER’S GUIDE TO REGRESSION ANALYSIS IN MACHINE LEARNING. MEDIUM. RETRIEVED SEPTEMBER 27, 2022, FROM <https://towardsdatascience.com/a-beginners-guide-to-regression-analysis-in-machine-learning-8a828b491bbf>
40. EUROPARL. (2021, MARCH 26). ŠTO JE UMJETNA INTELIGENCIJA I KAKO SE UPOTREBLJAVA? EUROPARL. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.europarl.europa.eu/news/hr/headlines/society/20200827STO85804/sto-je-umjetna-inteligencija-i-kako-se-upotrebljava>
41. WAKEFIELD, K. (N.D.). A GUIDE TO THE TYPES OF MACHINE LEARNING ALGORITHMS. SAS UK. RETRIEVED SEPTEMBER 28, 2022, FROM [https://www.sas.com/en\\_gb/insights/articles/analytics/machine-learning-algorithms.html](https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html)
42. JAVATPOINT. (N.D.). TYPES OF MACHINE LEARNING - JAVATPOINT. JAVATPOINT. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.javatpoint.com/types-of-machine-learning>
43. BURNS, E. (2021, MARCH). MACHINE LEARNING. SEARCHENTERPRISEAI. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.techtarget.com/searchenterpriseai/definition/machine-learning-ML>
44. PUBLICATION, KU. CHHAYA A. KHANZODE & DR. RAVINDRA D. SARODE (2020). ADVANTAGES AND DISADVANTAGES OF ARTIFICIAL INTELLIGENCE AND MACHINE

LEARNING: A LITERATURE REVIEW. IAEME PUBLICATION.

<https://doi.org/10.17605/OSF.IO/GV5T4>

45. BURNS, E., & BRUSH, K. (2021, MARCH). DEEP LEARNING. SEARCHENTERPRISEAI. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network>
46. TURING, A. M. (1950). MIND. MIND, 59(236), 433-460. RETRIEVED SEPTEMBER 30, 2022, FROM [https://www.cs.colostate.edu/~howe/cs440/csroo/yr2015fa/more\\_assignments/turing.pdf](https://www.cs.colostate.edu/~howe/cs440/csroo/yr2015fa/more_assignments/turing.pdf)
47. GOOGLE DEVELOPERS. (2020). OVERVIEW | GOOGLE ASSISTANT SDK |. RETRIEVED SEPTEMBER 28, 2022, FROM <https://developers.google.com/assistant/sdk/overview>
48. GOOGLE DEVELOPERS. (N.D.). LEARN HOW GOOGLE IMPROVES SPEECH MODELS - GOOGLE ASSISTANT HELP RETRIEVED SEPTEMBER 28, 2022, FROM <https://support.google.com/assistant/answer/11140942?hl=en#zippy=>
49. IBM CLOUD EDUCATION. (2022, JULY 15). MACHINE LEARNING. IBM. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.ibm.com/cloud/learn/machine-learning>
50. IBM CLOUD EDUCATION. (2020B, MAY 1). DEEP LEARNING. IBM. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.ibm.com/cloud/learn/deep-learning>
51. HASWANI, V. (2020, SEPTEMBER 3). LEARNING RATE DECAY AND METHODS IN DEEP LEARNING - ANALYTICS VIDHYA. MEDIUM. RETRIEVED SEPTEMBER 28, 2022, FROM <https://medium.com/analytics-vidhya/learning-rate-decay-and-methods-in-deep-learning-2cee564f910b>



52. DONGES, N. (2022, AUGUST 25). WHAT IS TRANSFER LEARNING? EXPLORING THE POPULAR DEEP LEARNING APPROACH. BUILT IN. RETRIEVED SEPTEMBER 28, 2022, FROM <https://builtin.com/data-science/transfer-learning>
53. MARIMUTHU, P. (2022, AUGUST 11). DROPOUT REGULARIZATION IN DEEP LEARNING. ANALYTICS VIDHYA. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.analyticsvidhya.com/blog/2022/08/dropout-regularization-in-deep-learning/>
54. LUTKEVICH, B. (2022, APRIL). BEST GOOGLE HOME DEVICES 2022: ASSISTANT IN YOUR SMART HOME. WHAT IS ARTIFICIAL GENERAL INTELLIGENCE (AGI)? RETRIEVED SEPTEMBER 18, 2022, FROM <https://www.techtarget.com/searchenterpriseai/definition/artificial-general-intelligence-AGI>
55. WIGMORE, I. (2018, FEBRUARY). ARTIFICIAL SUPERINTELLIGENCE (ASI). SEARCHENTERPRISEAI. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.techtarget.com/searchenterpriseai/definition/artificial-superintelligence-ASI>
56. FRANKENFIELD, J., ESTEVEZ, E., & PEREZ, Y. (2022, JANUARY 1). WEAK AI. INVESTOPEDIA. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.investopedia.com/terms/w/weak-ai.asp>
57. ALAN. (N.D.). WHAT IS A VOICE ASSISTANT? ALAN BLOG. RETRIEVED SEPTEMBER 28, 2022, FROM <https://alan.app/blog/voiceassistant-2/>
58. CAMPANA, N. (2022, SEPTEMBER 26). WHAT DOES A COMPUTATIONAL LINGUIST DO? FREELANCER BLOG. RETRIEVED SEPTEMBER 28, 2022, FROM <https://www.freelancermag.com/blog/what-does-computational-linguist-do/>

59. WOODL, P. C., LEGGETTER, C. J., ODELL, J. J., & VALTCHEV, V. (1998, APRIL). THE DEVELOPMENT OF THE 1994 HTK LARGE VOCABULARY SPEECH RECOGNITION SYSTEM. RESEARCHGATE. [https://www.researchgate.net/publication/2361715\\_The\\_Development\\_Of\\_The\\_1994\\_Htk\\_Large\\_Vocabulary\\_Speech\\_Recognition\\_System](https://www.researchgate.net/publication/2361715_The_Development_Of_The_1994_Htk_Large_Vocabulary_Speech_Recognition_System)
60. VELIKOVICH, L., WILLIAMS, I., SCHEINER, J., & ALEKSIC, P. (2018, SEPTEMBER). SEMANTIC LATTICE PROCESSING IN CONTEXTUAL AUTOMATIC SPEECH RECOGNITION FOR GOOGLE ASSISTANT. RESEARCHGATE. <https://doi.org/10.21437/Interspeech.2018-2453>
61. MICHAELY, A. H., ZHANG, X., SIMKO, G., PARADA, C., & ALEKSIC, P. (2018). KEYWORD SPOTTING FOR GOOGLE ASSISTANT USING CONTEXTUAL SPEECH RECOGNITION. 2018 IEEE AUTOMATIC SPEECH RECOGNITION AND UNDERSTANDING WORKSHOP (ASRU). <https://doi.org/10.1109/asru.2017.8268946>

## Popis slika

Slika 1: Turingov test (Pietikäinen & Silvén, 2022).....	4
Slika 2: Tipovi strojnog učenja (Rishabh & Piyush, 2020) .....	10
Slika 3: Grane umjetne inteligencije (Gaël, 2020).....	14
Slika 4: Vrste umjetne inteligencije (Codecnetworks, 2021) .....	16
Slika 5: Razlika između slabe i jake umjetne inteligencije (Rodriguez, 2020) .....	18
Slika 6: word lattice (Woodl i sur.).....	26
Slika 7: Kontekstualno automatsko prepoznavanje glasa za ocjenjivanje semantičke rešetke riječi (Velikovich i sur., 2018).....	27
Slika 8: Davanje naredbi Google asistentu (Turner, 2018a).....	29
Slika 9: Korištenje asistenta u autu (Kinsella, 2018).....	32

## **Sažetak**

Ključni element u razvoju Googleovog asistenta, odnosno glasovnih asistenta općenito je to da predstavlja ubrzani razvoj umjetne inteligencije. Pomoću asistenata moguće je svakodnevne obveze odrađivati s lakoćom koje uz pomoć umjetne inteligencije može obavljati radnje koje se prije smatralo da su isključivo u domeni čovjeka. Zahvaljujući razvoju metoda strojnog učenja, obrade prirodnog jezika i računalne lingvistike asistenti su u stanju razmišljati i obavljati razgovore putem kojih će korisnik davati upute o obavljanju raznih zadataka. Glavni cilj ovog rada je bio prikazati upotrebu, funkcionalnosti, razvoja i objasniti sve grane umjetne inteligencije koje se koriste u razvoju istih. Rezultati ovog rada omogućuju bolje razumijevanje glasovnih asistenata i kako će promijeniti kvalitetu života i životne navike.

**Ključne riječi:** umjetna inteligencija, Google asistent, prepoznavanje glasa, računalna lingvistika

## Summary

A key element in the development of the Google assistant, or voice assistants in general, is that it represents the accelerated development of artificial intelligence. With the help of assistants, it is possible to carry out everyday duties with ease, which with the help of artificial intelligence can perform actions that were previously thought to be exclusively in the domain of humans. Thanks to the development of machine learning methods, natural language processing and computational linguistics, assistants are able to think and conduct conversations through which the user will give instructions on performing various tasks. The main goal of this work was to show the use, functionality, development and explain all the branches of artificial intelligence that are used in their development. The results of this work enable a better understanding of voice assistants and how they will change the quality of life and lifestyle habits.

**Keywords:** artificial intelligence, Google Assistant, voice recognition, Computational linguistics