

*Prolećni semestar, 2021/22*

*PREDMET: IT370 INTERAKCIJA ČOVEK-RAČUNAR*

Domaći zadatak: **08**

Ime i prezime: **Viktor Cvetanovic**

Broj indeksa: **4421**

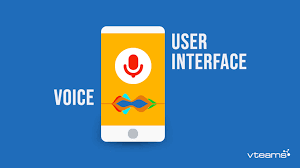
Datum izrade: **08.04.2022.**

# Tekst domaćeg zadatka:

Za domaći zadatak treba da istražite aktuelnu literaturu i tech blogove i napravite kratak izveštaj o aktuelnim tehnologijama vezanim za govorne interfejse.

Izveštaj treba da bude na dve strane plus ilustracije. Na kraju izveštaja formulišite svoj zaključak i navedite reference.

# Rešenje zadatka:

 Sveukupni trend ka Govornom korisničkom interfejsu (GKI ili eng. VUI) ubrzava se kako procesori postaju sve manji, a sve više Internet of Things (IoT) uređaja dolazi na mrežu. Sve veće tehnološke kompanije i „velikih pet“ - Google, Apple, Microsoft, Amazon i Facebook - ulažu velika sredstva u glasovni interfejs kao put napred.

U 2019. godini procenjeno je da je 59 posto pretraživanja na mreži došlo putem glasa, a neka pretraživanja specifična za određenu industriju zabeležila su mnogo više rezultata (na primer, pretraga hrane i restorana dostigla je 68 posto). Tržišni sektor govornog i glasovnog interfejsa na putu je da dostigne neverovatnih 24,9 milijardi dolara do 2025. godine.

Dizajn glasovnog interfejsa koristi prepoznavanje govora kako bi korisnicima omogućio interakciju sa tehnologijom pomoću glasovnih komandi. Kako svet postaje sve brži i gušći informacijama, glasovne tehnologije izazivaju dominaciju grafičkog korisničkog interfejsa i mogu učiniti doživljaj mnogo intuitivnijim. Tastature mogu raditi za zadatke koji zahtevaju tekst, poput pisanja bloga, ali za jednostavne zadatke poput pretraživanja, objavljivanja ažuriranja, prilagođavanja kontrola i dobijanja uputstava, ljudi bi naravno radije samo razgovarali. Oni mogu da sačuvaju mnogo koraka, na primer, unošenjem adrese na mali ekran, ali iskustvo mora funkcionisati besprekorno. U stvarnosti, dizajn interfejsa za glas funkcioniše zajedno sa tehnologijama dodira i prevlačenja kako bi stvorio bolje korisničko iskustvo.

Korisnici žele samo jedno od glasovnog interfejsa. Oni žele da njihovi uređaji prepoznaju njihove komande trenutno, bez grešaka. Ljudi zapanjujuće vešto razumeju glasovne poruke uprkos pozadinskoj buci, nedostajućim slogovima, varijacijama u jačini zvuka, akcentima i značenjima reči koje se neprestano razvijaju. Nijedan AI još uvek ne može da se 100% podudara sa tom sposobnošću da razume apsolutno sve govorne glasovne komande, tako da je na dizajnerima i programerima da razmisle o svim nepredviđenim situacijama.

Bilo koji uređaj ili aplikacija mogu da rade sa govorom, ali to ne znači da bi svi trebali. Inteligentni dizajn korisničkog interfejsa za glas započinje ispitivanjem toga gde će govor olakšati život korisniku. Na primer, uređaji koje će javnost koristiti napolju, poput prozora za prijavu, predstavljaju previše problema u pogledu varijacije glasa i buke u okruženju. Takođe, kad god se korisnik bavi osetljivim informacijama poput zdravstvenih podataka ili poverljivih podataka, glasno izgovaranje je poslednja stvar koju želi. Govorni interfejs je savršen za sve zadatke koji imaju koristi od upotrebe bez upotrebe ruku (poput aplikacija za kuvanje ili vožnju), pametnih kućnih uređaja i aplikacija koje imaju emocionalnu komponentu (poput vežbanja ili produktivnosti).

Ljudi mogu obrađivati glasovne informacije brzinom od oko 39 bita u sekundi, tako da Govorni interfejsi moraju težiti da obrade informacije što je brže moguće bez žrtvovanja dvosmislenosti značenja. U većini slučajeva dizajneri su izgradili konverzacijski AI za obradu informacija ili ovaj deo predali mehanizmima za prepoznavanje govora otvorenog koda. Korisnici postaju nestrpljivi kada uređaj zahteva da govore polako ili jednu po jednu reč. Takođe mogu biti nepoznati suštinski koncepti zasnovani na njihovoj generaciji. Na primer, izrazi poput „smer kazaljke na satu“ i „pokvareni rekord“ možda ne znače ništa mlađim korisnicima. Potrebno je provesti što je moguće više vremena upoznajući se sa načinom na koji korisnici razgovaraju i dizajnirajte ličnost GKI-a oko tog istraživanja. Naravno potrebno je izgraditi i bogatu biblioteku izgovora, naredbi, vrednosti i namera da bi pojednostavili obradu informacija.

Prvi i najvažniji element u planu dizajna za GKI-a je „karta putovanja korisnika“, gde dizajneri određuju šta je korisniku najpotrebnije u različitim fazama angažovanja. Ako postoje konkurenti koji su već uveli radne verzije svog proizvoda, polovina tog rada može se lako uhvatiti putem konkurentne inteligencije. U suprotnom, biće potrebno značajno vreme za prikupljanje povratnih informacija korisnika. Sa uspostavljenom mapom koja precizira šta GKI mora da uradi i očekivane ishode, teret se prebacuje na definisanje korisničkih ličnosti, kako će GKI imati veze sa različitim uređajima i bibliotekom okidača.



Stvaranje ličnosti je presudan prvi korak da se tim zadrži na istoj strani o tome šta korisnik želi da postigne. GKI korisnička uputstva proširuje uobičajene atribute, kao što su ime, uloga, očekivanja itd., Dodatnim detaljima o svom tonu glasa i tipičnim izborima reči. Idealno bi bilo da vaš GKI bude dovoljno robustan da automatski prepozna različite ljude ili različite postavke za istu osobu. Na primer, pametni termometar može imati različite korisničke zahteve za „rad kod kuće“ i „zabavu gostiju“. Ili aplikacija pametnog automobila može imati različita podešavanja radija i navigacije na osnovu toga ko vozi taj dan.

Dizajniranje korisničkog interfejsa za glas koji ispunjava očekivanja korisnika može biti vrlo izazovno, čak i za iskusne UX dizajnere. Korisnici su navikli da vrše podešavanja kada je aplikaciji ili uređaju potrebno više podataka, ali psihologija ponavljanja izgovorene komande je drugačija. Kada GKI zatraži određeni broj karata, ljudi imaju tendenciju da odgovore odgovorima poput „Par“ ili „Samo ja i moja supruga“. Nemoguće je planirati svaku nepredviđenu situaciju, tako da dizajneri moraju da osmisle angažovane skripte koje korisnike vode na prijateljski, razgovorni način koji mora biti prepoznat. Jedna korisna referenca je Google-ov vodič za razvoj razgovora, zasnovan na hiljadama sati istraživanja.

Postoje četiri glavne vrste uređaja koji zavise od glasovne interakcije i koji će se u budućnosti sve više oslanjati na govorni interfejs i naravno svaki dolazi sa svojim setom ograničenja. Uzmite u obzir ove faktore prilikom dizajniranja za svaki od njih.

1. Telefoni – vađno je pažljivo pregledati specifikacije dizajna, razmotrite kako promene u okruženju utiču na GKI, gde grafika može pomoći u vođenju korisničkih unosa
2. Mobilni uređaji koji nisu telefoni - uređaji zasnovani na automobilu i prenosni računari će imati isprekidane veze, gde glas obično nije jedini ili primarni korisnički interfejs
3. Nosivi uređaji - većina su uređaji za jednostavnu upotrebu sa ograničenom funkcionalnošću, ali mogu biti upareni sa aplikacijom za telefon ili laptop što će dodatno poboljšati rad GKI-a
4. Fiksni uređaji - verovatno se moraju prilagoditi više korisnika, komunicirati sa drugim uređajima u okruženju radi bolje impelmentacije govornog interfejsa

Neki GKI zasnovani na automobilima takođe su počeli da uključuju podatke proširene stvarnosti, poput podsećanja korisnika da kupe potrebne stvari kada se voze blizu željene prodavnice za koju su sastavili listu za kupovinu u nekoj aplikaciji, na čemu se treba dosta raditi ali ni od ovakvih razvojnih rešenja nismo daleko.

Zaklučak: Specifikacije uređaja se konstantno razvijaju i radikalno novi uređaji stalno dolaze na mrežu odnosno u prodaju, ali smatram da je ljudski glas ,korisnički interfejs svih nas, na koji se možemo osloniti u doglednoj budućnosti. Procesori se smanjuju i IoT sve oživljava. U svetu koji nastaje malo je mesta za bilo koji drugi interfejs osim zvukai govora. Neposrednost i svestranost glasa čine direktnu vezu od želje korisnika do odgovora uređaja. Međutim, da bi iskoristili blagodati GKI-a, dizajneri moraju poći od detaljnih persona ljudi, navika, govornih varijacija, zvukova okoline, ostalih propratnih interfejsa i na taj način zaobići potencijalna ograničenja određenih uređaja i najbolje iskoristiti sve potencijalne pokretače akcije govorom.

# Korišćena literatura. : (opciono)

[https://www.uxpin.com/studio/blog/voice-user-interface/?fbclid=IwAR21\_4 68tCSRerdclsCfW7TOlHCaAh7V61rnv4k6PZNjPWwEh3b8ATKspc](https://www.uxpin.com/studio/blog/voice-user-interface/?fbclid=IwAR21_4)

<https://ai.googleblog.com/2020/11/improving-on-device-speech-recognition.html>

<https://www.youtube.com/watch?v=kOqYEWLuAZQ&ab_channel=QuanWang>