

M-Vstopnica – uporaba avtomatskega razpoznavanja govora v praksi

Gordan Sket, Bojan Imperl

HERMES SoftLab
Litijska 51, 1000 Ljubljana, Slovenija
Gordan.Sket@hermes.si, Bojan.Imperl@hermes.si

Povzetek

Članek predstavlja M-Vstopnico, storitev, ki so jo skupaj izdelali partnerji HERMES SoftLab, Mobitel, Ljubljanski kinematografi – Kolosej, Programski atelje A&Z ter finski Moom. M-Vstopnica omogoča uporabnikom mobilnih telefonov rezervacijo in nakup vstopnic za kino predstave preko govornega uporabniškega vmesnika v slovenskem jeziku. Govorni uporabniški vmesnik M-Vstopnica temelji na avtomatskem prepoznavanju govora in prepoznavanju DTMF signalov ter na predvajanju vnaprej posnetih sporočil. Interakcijo z uporabnikom določa sistem – sistemsko voden dialog, kjer uporabnik predvsem odgovarja na vprašanja, vendar se dialog samodejno prilagaja uporabnikovem obnašanju in drugim dejavnikom v celotnem sistemu. Namen članka je predstaviti konkreten primer uporabe govorne tehnologije v praksi; predstaviti prednosti in omejitve sistemov za prepoznavanje govora na dejanskem primeru, ter predstaviti spremljajoče tehnologije kot so VoiceXML in mrežne storitve (Web services), ki omogočajo uspešno integracijo sistema za prepoznavanje govora v učinkovito celoto. M-Vstopnica je trenutno v fazi testiranja in jo je možno preskusiti na telefonski številki 041 184 184.

1. Uvod

M-Vstopnica je prva v Sloveniji in ena še vedno redkih storitev na svetu, ki omogoča uporabnikom telefonov komunikacijo preko govornega uporabniškega vmesnika z uporabo avtomatskega prepoznavanja govora. Razvoj storitve kot je M-vstopnica že na splošno predstavlja velik izziv, saj gre v prvi vrsti za kompleksno integracijo sicer zaprtih sistemov, kot so sistemi mobilnega operaterja in rezervacijski sistemi kinematografov, kjer ima vsaka komponenta svoje specifične zahteve in omejitve. Uporaba naprednih tehnologij mrežnih storitev (»Web services«), ki ponudniku storitve omogoča cenejše spreminjanje in prilagajanje rešitve (hitrejša izvedba, večja fleksibilnost), ter tehnologije VoiceXML, ki omogoča izvedbo govornega vmesnika z dinamično adaptacijo strukture dialoga. Takšne tehnologije sicer olajšajo nekatere tehnične težave, vendar pa hkrati odpirajo nova vprašanja, saj gre v glavnem za tehnologije, ki so v svetu zelo nove in s katerimi je v praksi zelo malo izkušenj.

Govorna tehnologija, oziroma natančneje, tehnologija avtomatskega prepoznavanja govora, je sicer prisotna že dalj časa, vendar je z uporabo takšne tehnologije v praksi, še posebej v slovenskem prostoru, zelo malo izkušenj. Hitrost uveljavljanja govorne tehnologije na določenem jezikovnem področju je na eni strani odvisna od tehničnih vprašanj, vezanih na specifične lastnosti samega jezika, in na drugi strani od želja in pričakovanj uporabnikov, katerim bo tehnologija oziroma storitev ponujena. V nekaterih področjih so uporabniki telefonskih omrežij že vajeni in celo siti pustih avtomatskih interaktivnih odzivnikov na osnovi prepoznavanja DTMF, zato je v takih okoliščinah govorna tehnologija zelo dobrodošla in pri uporabnikih dobro sprejeta. Nasprotno pa je v drugih področjih, kjer so avtomatski interaktivni odzivniki še vedno redki ali se šele začenjajo uveljavljati, učinek uporabe govorne tehnologije med uporabniki zelo nepredvidljiv, saj lahko hitro vzbudi zmotna in pogosto

tudi nerealna pričakovanja ter, kot posledica tega, slabo izkušnjo s storitvijo in tehnologijo prepoznavanja govora.

Mobilna telefonija je v zadnjih letih bistveno dvignila hitrost razvoja telekomunikacijskih tehnologij in pripomogla k zmanjšanju razlik v stopnji razvoja storitev v telekomunikacijah v različnih deželah, tako s tehničnega, kot tudi z uporabniškega stališča. Upabniki so se navadili na tako hitrost in praviloma dobro sprejemajo različne novosti. Podjetje HERMES SoftLab se želi s postavitvijo M-Vstopnice podrobneje osredotočiti na problematiko ponujanja rešitve preko mobilnega telefona in storiti korak naprej v razvoju telekomunikacijskih storitev v slovenskem prostoru

V prispevku bo opisana zgradba in delovanje storitve M-Vstopnice ter izkušnje pri uporabi tehnologije avtomatskega prepoznavanja govora ob postavljanju storitve M-Vstopnica.

2. Opis uporabe M-Vstopnice

Storitev M-Vstopnica vsebuje govorni vmesnik do sistema, ki je povezan s ponudniki predstav (trenutno samo v Koloseju in Ljubljanskih kinematografih). Uporabnik storitev uporabi tako, da pokliče določeno telefonsko številko (041 184 184), kjer lahko preko glasovnega vmesnika sliši program predstav, izbere željeno predstavo ter rezervira ali kupi vstopnice zanjo. Kot potrdilo o rezervaciji ali nakupu prejme uporabnik sporočilo v obliki SMS.

V drugem delu uporabe uporabnik s kodo, ki je zapisana v SMS sporočilu, obišče kinematograf, kjer blagajničarki prebere kodo (ali pa jo vtipka v avtomat za izdajo kart), ter nato dobi običajno vstopnico, s katero lahko obišče predstavo. Če je uporabnik med pogovorom s sistemom izbral možnost plačila vstopnice, ceno vstopnice plača mobilnemu operaterju skupaj s ceno za ostale opravljene storitve. V nasprotnem primeru plača vstopnico pri blagajni kinematografa.

3. Arhitektura M-Vstopnice

Sistem M-Vstopnice je sestavljen v štirinivojski arhitekturi, ki za komunikacijo med seboj uporablja javne Internetne povezave.

Nivo **govornega uporabniškega vmesnika** je izvajan na Govornem Portalu (Imperl et al., 2001), razvitem v podjetju HERMES SoftLab, ki je na eni strani povezan s telefonsko centralo, na drugi pa preko Interneta dostopa do nivoja strežnika uporabniškega vmesnika. Opis uporabniškega vmesnika temelji na VoiceXML jeziku; Govorni Portal izvaja opise dialogov, ki jih generira strežnik uporabniškega vmesnika.

Strežnik uporabniškega vmesnika je zadolžen za izdelavo opisa vmesnika (VoiceXML skript) na osnovi podatkov, ki jih prejme od nivoja poslovne logike. Vmesnik je izveden v programskem okolju Java (JSP).

Nivo poslovne logike je zadolžen za izdelavo, pridobivanje in pošiljanje podatkov do nivoja dostopa do podatkov. Posamezni deli nivoja in povezava z nižjim nivojem je izvedena v programskem okolju Java z izdelkom e-speak. E-speak (Apte, Mehta, 2001) je prva izvedba tehnologije mrežnih storitev (Web services), uporablja pa tehnologijo SSL, zato so povezave zelo varne.

Najnižji nivo, ki smo ga razvili posebej za M-Vstopnico, **nivo dostopa do podatkov**, skrbi za komunikacijo in integracijo v obstoječe sisteme. Taka sistema sta sistem SiTi (izdelan v Programskem ateljeju A&Z in postavljen v Koloseju) ter sistemi na Mobitelu (dostop do računov uporabnikov ter storitev Poštar – pošiljanje SMS sporočil).

4. Govorni vmesnik M-Vstopnice

Govorni vmesnik M-Vstopnice je izveden s pomočjo Govornega Portala, ki temelji na uporabi VoiceXML tehnologije (Sharma, Kunnins, 2002) in dejansko predstavlja VoiceXML strežnik.

Jezik VoiceXML, ki so ga skupaj predlagali AT&T, Motorola, Lucent in IBM, je v osnovi oblikovan za gradnjo sistemov govornega dialoga med človekom in strojem, torej za gradnjo govornih uporabniških vmesnikov v telefonskem okolju, ki vključujejo: avtomatsko razpoznavanje govora (ASR: automatic speech recognition), avtomatsko sintetiziranje govora (TTS: text to speech), prepoznavanje DTMF signalov (Dual-Tone Multi Frequency), predvajanje in shranjevanje zvočnih posnetkov. Glavni namen jezika VoiceXML je poenostaviti gradnjo govornih vmesnikov in prinesiti prednosti Internet-načina razvoja aplikacij in posredovanja informacij tudi v telefonsko okolje.

Preprost primer uporabe VoiceXML jezika predstavlja naslednji program:

```
<vxml version="1.0">
<menu>
  <prompt>
    <audio src="vprasanje.wav"></prompt>
    <choice dtmf="1" next=" n.vxml">triler
    </choice>
    <choice dtmf="2" next=" v.vxml">akcija
    </choice>
    <noinput>Prosim, izberite triler ali
```

```
alcija. Za triler recite triler ali
pritisnite 1, za akcijo recite akcija
ali pritisnite 2. </noinput>
<nomatch>Nisem razumel.</nomatch>
</menu>
</vxml>
```

Ko bi uporabnik (preko kateregakoli telefona) poklical telefonsko številko Govornega Portala - VoiceXML strežnika, bi ta izvedel zgornji program, kar bi pomenilo, da bi uporabnik slišal posnetek vprasanje.wav, nato pa bi strežnik od uporabnika pričakoval odziv in sicer v govorni obliki (»triler« ali »akcija«) ali v obliki DTMF signalov (»1« za triler ali »2« za akcijo). Glede na uporabnikov odgovor VoiceXML strežnik nato nadaljuje z izvajanjem naslednjih VoiceXML skript kot jih generira strežnik uporabniškega vmesnika.

VoiceXML strežnik se sestoji iz naslednjih komponent: **telefonski vmesnik** (skrbi za komunikacijo s telefonskim omrežjem), **prepoznavnik govora** (opravi prepoznavanje posnetka, ki ga pripravi telefonski vmesnik), **sintetizator govora**¹ (sintetizira besedilo, ki ga posreduje VoiceXML interpreter), **VoiceXML Interpreter** (izvaja VoiceXML skripta in ustrezno krmili telefonski vmesnik, prepoznavnik, sintetizator, dokumentni upravljalnik), **dokumentni upravljalnik** (skrbi za nalaganje VoiceXML dokumentov).

4.1. Modul za avtomatsko prepoznavanje govora

Modul za avtomatsko prepoznavanje govora je sestavni del VoiceXML strežnika. V primeru M-Vstopnice ima razpoznavnik nalogo prepoznavati osamljeno izgovorjene besede ali fraze z velikostjo slovarja od 10 do 40 besed ali fraz. S stališča tehnologije razpoznavanja govora to sicer ne predstavlja zahtevne naloge, vendar pa obstaja vrsta praktičnih omejitev, ki zelo otežujejo doseganje zadovoljive uspešnosti razpoznavanja kot na primer:

- GSM kvaliteta govora (velika nihanja razmerja signal-šum, manjkajoči segmenti govora, ko gre za zelo slabe zveze, kvaliteta govora, odvisna od posameznih aparatov, odmev, ...),
- zahtevano delovanje v realnem času na 15 telefonskih kanalih na enem računalniku brez uporabe dodatnih DSP plošč,
- podpora alternativnih izgovorjav za posamezne besede ali fraze.

Prepoznavniki na trgu: V začetku razvoja M-Vstopnice je bil uporabljen razpoznavnik podjetja L&H kjer so bili akustični modeli, razviti za angleški jezik in za telefonsko okolje, uporabljeni za razpoznavanje slovenskih osamljeno izgovorjenih besed ali fraz. V začetnih testiranjih je takšna izvedba razpoznavnika dajala zadovoljive rezultate, vendar se je ob širitvi slovarja na več kot 10 besed izkazala za neuporabno, saj je uspešnost razpoznavanja padla na približno 60%. Sledilo

¹ M-Vstopnica ne vključuje sintetizatorja govora, posredovanje vprašanj uporabnikom je izvedeno s pomočjo sestavljanja in predvajanja vnaprej posnetih stavkov, besed ali fraz.

je preizkušanje z razpoznavalnikom podjetja NUANCE, ki je bil uporabljen na enak način kot prvotni. Ta je sicer dal boljše rezultate kot prvotno uporabljeni razpoznavalnik podjetja L&H, najboljši rezultati so bili doseženi z akustičnimi modeli razvitimi za češki jezik in za telefonsko okolje, vendar v praktičnih poizkusih še vedno ni dosegal potrebne zanesljivosti saj je uspešnost razpoznavanja znašala med 70% in 75%.

Ocenjevanje uspešnosti prepoznavanja: Ocena uspešnosti razpoznavanja je bila izvedena s pomočjo testne baze 360 posnetkov, ki je vključevala dejanske posnetke govora različnih oseb, posnetih v produkcijskem okolju v fazi testiranja. Gre za posnetke, kjer koristen govor predstavlja manjši del posnetka. Tako izmerjena uspešnost razpoznavanja dejansko predstavlja skupno uspešnost delovanja prepoznavalnika in algoritma za detekcijo govora. Uporabljeni so le posnetki, kjer odziv uporabnikov ustreza zahtevanemu oziroma pričakovanemu odzivu.

Lastni prepoznavalnik: Najboljši rezultati so bili doseženi z lastnim razpoznavalnikom, ki je že pri majhni učni bazi izgovarjav (40 govorcev) močno presegel uspešnost razpoznavanja, doseženo z obema prejšnjima sistemoma, saj je bila dosežena uspešnost razpoznavanja 89%. Razpoznavalnik temelji na kombinaciji trifonskih in monofonskih Prikritih Modelov Markova ter vključuje algoritem za detekcijo govora na osnovi energije signala, optimiranim za mobilno telefonsko okolje in z dinamičnim prilagajanjem povprečnega nivoja šuma. Uporabljen razpoznavalnik predstavlja dober kompromis med uspešnostjo razpoznavanja in računsko zahtevnostjo, saj omogoča delovanje v realnem času hkrati tudi po več kot 15 kanalih.

Razpoznavalnik govora je trenutno še v fazi razvoja, prva verzija je že uporabljena v M-Vstopnici, za drugo verzijo pa je predvideno predvsem povečanje učne baze izgovarjav. Prepoznavalnik govora, uporabljen pri M-Vstopnici, trenutno dosega uspešnost razpoznavanja 94%. Približno 70% vseh napak nastane zaradi napak detektorja govora, 20% zaradi napačno zavrjenega posnetka in 10% zaradi zamenjave z napačno besedo.

Učna baza izgovarjav: Učna baza vključuje posnetke zaposlenih v podjetju HERMES SoftLab, ki so tako po starostni kot tudi po narečni strukturi podobni ciljni skupini - mladostniki mesta Ljubljane z okolico. Vsak posnetek predstavlja približno 10 minut govora, posnetega preko mobilnega telefona (branje 21 leposlovnih stavkov ter besed in fraz M-Vstopnice). Trenutno je posnetih 150 oseb, povečevanje učne baze pa je "usmerjeno", kar pomeni, da bazo izgovarjav sestavljajo predvsem posnetki govorcev iz ciljne skupine populacije.

5. Govorni uporabniški vmesnik

Uspeh storitve je močno povezan predvsem z uporabnostjo govornega vmesnika za končnega uporabnika. Cilj je bil zgraditi vmesnik, ki je učinkovit – omogoča čim krajšo pot od klica do opravljene rezervacije, je enostaven za uporabo, ne zahteva predhodnih izkušenj, intuitiven.

5.1. Struktura dialoga

Storitev predstavlja kompleksen sistem, ki mora uporabniku predstaviti program kino predstav (tipično približno 50 dogodkov na dan, razdeljenih v 15 do 20 različnih filmov s tipično 3 ponovitvami) ter od njega izvedeti več parametrov: kateri film, katera dvorana, kdaj, koliko sedežev, predel dvorane.

Analiza navad uporabnikov je pokazala, da se večina pri izbiri predstave odloča po nasvetih prijateljev in zato že vedo, kateri film si želijo ogledati. Temu je bil prilagojen vmesnik M-Vstopnice, ki tako uporabnika na začetku vpraša, kateri film si želi ogledati. Šele če uporabnik ne ve za film, mu sistem predstavi filme, razdeljene v kategorije – žanre filmov. Razdelitev filmov v žanre je nujna zaradi vzdrževanja ritma pogovora - naštevaje vseh naslovov filmov namreč zahteva bistveno preveč časa.

Po izbiri filma je možnih več scenarijev, po katerih lahko obiskovalec izbere željeno predstavo, na primer izbira filma glede na določeno uro ali glede na določeno dvorano. Vmesnik, ki bi omogočal več scenarijev, bi bil prezapleten in zato manj učinkovit. Izbran je bil en sam - najpogostejši način izbire filma, tako da je prvo vprašanje o imenu filma, sledijo pa ostali podatki: kraj in čas predstave (skupaj v enem vprašanju), število sedežev, itn.

Po izbranem številu sedežev sistem uporabniku ne ponudi izbire sedežev, temveč mu ponudi določene sedeže in ga vpraša, če jih želi rezervirati, kupiti ali zamenjati za druge. Izbira sedežev namreč lahko precej podaljša dialog, zato je možna le na uporabnikovo eksplicitno zahtevo.

5.2. Implicitno potrjevanje izbranega

Zaradi vzdrževanja toka dialoga in za preverjanje pravilnosti razpoznanega govora vsakem uporabnikovem vnosu sledi potrdilo. Uporabljeno je implicitno potrjevanje, kjer sistem v naslednjem vprašanju ponovi uporabnikov prejšnji vnos in uporabniku s tem ponudi možnost prehoda na prejšnje vprašanje, na primer:

M-Vstopnica: »Katera predstava vas zanima? ...«,

uporabnik: »Vojna zvezd«,

M-Vstopnica: »Kje in kdaj pa bi si predstavo *Vojna zvezd* radi ogledali? ...«.

Če je vmesnik napačno razumel uporabnika, lahko uporabnik na naslednje vprašanje odgovori z »nazaj« in nato ponovno sliši prvotno vprašanje. Uporabnik je o tej možnosti obveščen na začetku pogovora, v navodilih za uporabo. Uporaba eksplicitnega potrjevanja (»Če ste izbrali *Vojna zvezd*, recite da...«) bi pogovor zelo podaljšala, ter vzbujala občutek, da se dialog prekinja.

5.3. Dolžina vprašanj in »bargo-in«

Med preizkušanjem prve verzije vmesnika se je izkazalo, da ni mogoče zagotoviti željene dinamičnosti govornega vmesnika brez uporabe prekinjanja govora (»bargo-in«). Uporabnikom, ki so storitev že poznali, je bilo poslušanje dolgih navodil odveč, krajšanje navodil pa je zelo podaljšalo uvajalni čas, izbiranje med dolgimi in polpnimi navodili za začetnike ali kratkimi za izkušene

uporabnike pa bi zopet predstavljalo dodaten korak v dialogu.

Ker je uspešnost vmesnika in s tem celotne storitve močno odvisna od občutka pri prvi uporabi in od občutka učinkovitosti pri pogosti uporabi, je bila vpeljava možnosti prekinjanja govora nujna. S tem se je odprla možnost oblikovanja daljših sistemskih vprašanj, ki se sestojijo iz kratkega, neposrednega vprašanja ter daljšega navodila, na primer: »Katero predstavo bi si radi ogledali? Če poznate ime filma, ga kar izgovorite. Recite, na primer, 'Vojna zvezd'. Če imena filma ne poznate, recite 'naštev zvrsti'.«.

Izkušen uporabnik ve, da lahko predvajanje teh nasvetov kadarkoli prekine, neizkušen uporabnik pa si želi več navodil in ima na ta način občutek, da mu sistem stoji ob strani - ne dogaja se, da sistem postavi vprašanje, nato pa v tišini čaka na odgovor, kar lahko pri uporabnikih zbudi občutek negotovosti. S takšnim načinom postavljanja vprašanj je omogočena uporaba vmesnika velikemu številu različno izkušenih uporabnikov. Začetniki običajno poslušajo vprašanja in nasvete do konca, izkušenejši pa vprašanje prekinejo že takoj ob prvi predvajani besedi.

Uvedba prekinjanja govora po drugi strani odpira tudi nekaj novih problemov. Med njimi je najresnejše povečanje možnosti napak pri neizkušenih uporabnikih, ko se zaradi šumov iz okolice, odmeva, ali uporabnikovih "komentarjev" (neizkušen uporabnik se ne zaveda, da sistem vedno "poslušá") poveča število napak razpoznavanja, oziroma se uporabnik zmede zaradi odziva, ki ga ni pričakoval. Vpeljava možnosti prekinjanja govora je zato zahtevala tudi več pozornosti pri sestavljanju navodil – vključitev stavka »Hrup in govor iz okolice lahko motita delovanje...« - predvsem pa povečanje robustnosti razpoznavalnika govora in algoritma za detekcijo govora.

5.4. Dolžina pogovorov in uspešnost klicev

Med preizkušanjem enega od prvih poskusnih vmesnikov se je izkazalo, da ima uporabnik občutek, da pogovor traja bistveno predolgo. Tipično je uporabnik potreboval za izbiro predstave in nakup vstopnic približno 3 minute. Zaradi večjega števila je bilo tudi več možnosti za napake, zaradi česar je bila stopnja uspešnosti klicev (ko uporabnik dobi željeno informacijo ali storitev) približno 60%.

Poleg omenjenih izboljšav (prekinjanje govora) so bile s časom izvedene tudi nekatere izboljšave dialoga. Določena vprašanja, kot na primer vprašanje o kraju in času predstave, so bila združena: »Kje in kdaj pa si želite predstavo Vojna Zvezd ogledati? V Koloseju ob 20:30,...«. Uvedena je bila dodatna omejitev - izbira filmov zgolj za tekoči dan, s čimer je bil odstranjen še en korak v pogovoru. Izbira dneva bo dodana kasneje, ko se bo govorni vmesnik M-Vstopnice uveljavil med uporabniki kot zanesljiv način komuniciranja.

Poleg sprememb strukture dialoga so bile izvedene tudi izboljšave, ki pri uporabniku ustvarijo občutek dinamičnosti dialoga, občutek, da se pogovor hitro odvija - na mesta, kjer je sistem čakal na podatke, je bilo

vstavljeno predvajanje glasbe, sam govor pa je bil posnet na bistveno bolj dinamičen način.

Analiza neuspešnih klicev je pokazala, da je bolj kot dejanska dolžina, prav občutek, da je dialog dolg in zapleten, v največji meri »kriv«, da uporabnik obupa in prekine pogovor. Navedene izboljšave so bistveno pripomogle k dvigu uspešnosti klicev in krajšanju dolžine pogovorov. Povprečna dolžina pogovorov za izkušene uporabnike tako znaša manj kot 1 minuto, stopnja uspešnosti klicev pa presega 80%.

5.5. Pomoč uporabniku

Zaradi novosti, ki jih prinaša M-Vstopnica, je ob štartu storitve potrebno računati z neizkušeno stopnjo uporabnikov. V ta namen bodo v govorni vmesnik vključena izdatna navodila za uporabo storitve.

Takoj po začetku pogovora sistem uporabnikom predvaja splošna navodila, ki naj bi jih slišal vsak uporabnik in se nanje spomnil tudi, če vmesnika ni uporabljal dalj časa. Sistem predvaja takšna navodila v prvih nekaj uporabah in po nekaj tednih neuporabe storitve.

Poleg nasvetov, ki jih sistem predvaja uporabniku ob vsakem vprašanju (razdelek 5.3), sistem predvaja dodatna pojasnila ob težavah pri razpoznavanju ali v primeru uporabnikovega neodzivanja. Če sistem odgovora uporabnika ne razume ali ga sploh ne zazna, mu vprašanje zastavi še enkrat, vendar največ dvakrat. Z vsako ponovitvijo je vprašanje obširnejše in vsebuje tudi navodila, kako odgovoriti s pomočjo tipkovnice telefona (v obliki DTMF). To možnost ima uporabnik pri vseh vprašanjih, vendar je pri prvem predvajanju zaradi dolžine sistem ne omeni. Ponovitev vprašanja o kraju in času predstave se tako glasi: »Kje in kdaj pa si želite predstavo Vojna Zvezd ogledati? Prva predstava v Koloseju ob 20:30, druga predstava v ... Pritisnite zaporedno številko predstave...«.

6. Zaključek

Razvoj sistema kot je M-Vstopnica je zaradi pestrosti posameznih delov že samo s tehnične plati velik zalogaj. Vseeno pa je bilo največ časa porabljenega za razvoj govornega vmesnika. Rezultat je govorni uporabniški vmesnik, ki glede na rezultate začetnih raziskav med ciljnim uporabniki, izpolnjuje vse zastavljene cilje. Mnenja anketiranih uporabnikov kažejo, da je storitev zelo zanimiva, enostavna, odzivna in uporabna.

Trenutno je sistem še v fazi zadnjih izboljšav, vendar pa bo šele po nekaj mesecih množične uporabe moč zanesljivo oceniti, ali je vmesnik in celotna rešitev uspešna tako s tehničnega kot tudi z ekonomskega stališča.

7. Literatura

- N. Apte, T. Mehta. *Web Services*, Prentice Hall, 2001
- B. Imperl, Č. Oblonšek, A. Šoštarčič. VOICEXML – govorni dostop do vsebine Interneta. *Zbornik OTS 2001*, Maribor,
- C. Sharma, J. Kunnins. *VoiceXML*, Wiley, 2002