



eBralec – sintetizator govora za slovenščino



alpineon))



dr. Jerneja Žganec Gros

jerneja.gros@alpineon.si

JOTA, 19. 3. 2018

- Najnaravnejši način komunikacije med človekom in strojem
- Možnost vključevanja **uporabnikov s posebnimi potrebami** ter **ostarelih**
- Prijazni uporabniški vmesniki
 - **prepoznavanje govora**: ASR (automatic speech recognition) omogoča prostoročno upravljanje
 - **sinteza govora**: TTS (text-to-speech synthesis) ne moti zbranosti, ni potrebe po sprotnem snemanju glasovnih sporočil
- **Trendi**: selitev v oblak ter na prenosne naprave, zasnovane na vgrajenih sistemih



Prevajanje govora

"My account balance, please."

VoiceTRAN
Natančnost

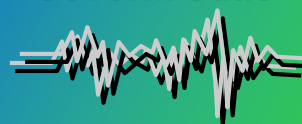
Prepoznavanje govora

"Želim stanje na računu."

Alp-Reco
Natančnost

Avtentikacija z govorom

Govorni odtis

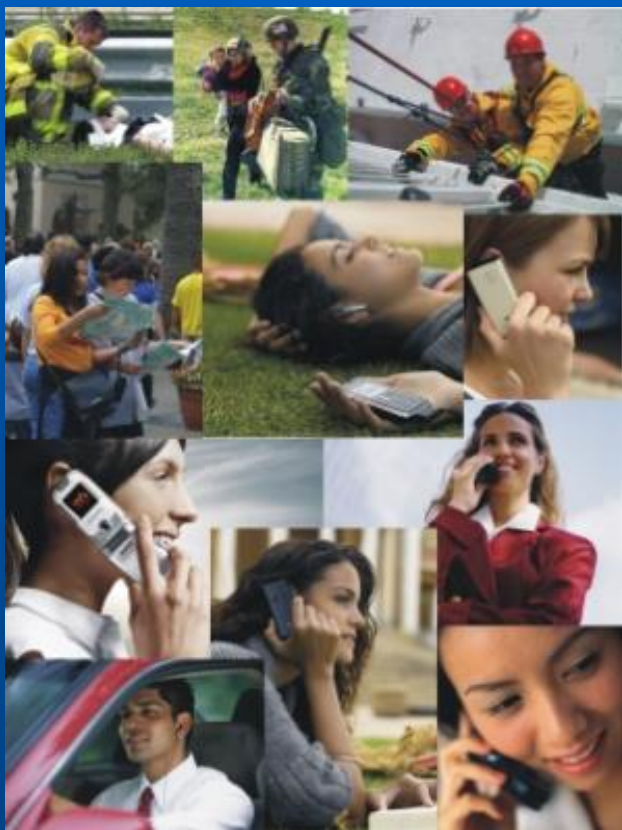


Alp-ID
Varnost

Sinteza govora

"Stanje na Vašem računu je 1.320 EUR."

eBralec
Naravnost

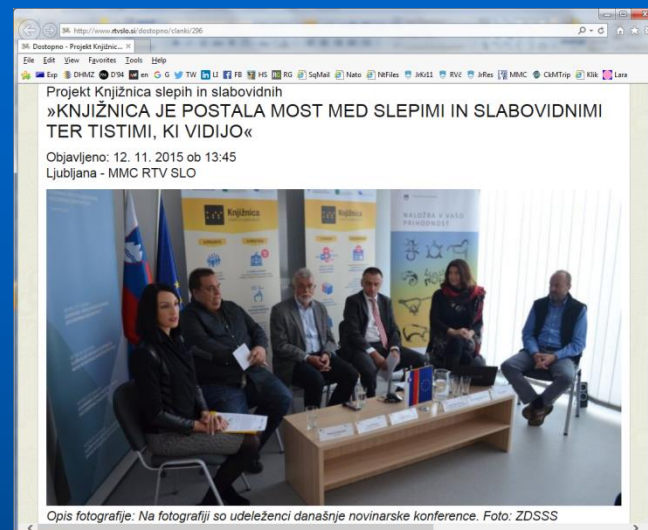


- Avtomobilska industrija
- Tehnologije za uporabnike s posebnimi potrebami
- Pametni dom
- ...
- IKT sektor
 - RTV
 - TK Operaterji
 - Virtualni pomočniki
 - Govorni portali ...
- ...

- Samodejno tvorjenje govora iz poljubnega besedila
- Zgodovina sintetizatorjev govora za slovenščino
 - **1997 S5** (UL-FE, Gros et al.); eno-uporabniški TTS – DOS
 - **1998 S5** (UL-FE, Gros et al.); eno-uporabniški TTS – Windows
 - **2000 IJS Govorec** (IJS, Gams et al.); eno-uporabniški TTS – Windows
 - **2001 MP-SYNTH** (Masterpoint, Gros et al.): strežniški TTS
 - **2006 Proteus TTS server** (Alpineon, Gros et al.): strežniški TTS
 - **2006 Amebis Govorec** (Amebis, IJS, Šef et al.): MS-SAPI TTS
 - **2007 Proteus TTS** (Alpineon, UL-FE, Žganec Gros et al.): MS-SAPI TTS
 - **2013 ispeak-slo**
 - **2015 Proteus TTS android** (Alpineon, UL-FE, Žganec Gros et al.): Android
 - **2015 eBralec** (Alpineon, Amebis, IJS, UL-FE): MS-SAPI,
 - **2017 MS Lado** (Microsoft)
 - **laboratorijski demonstracijski sistemi:** HMM TTS (UL-FE), Plattos (UM-FERI)



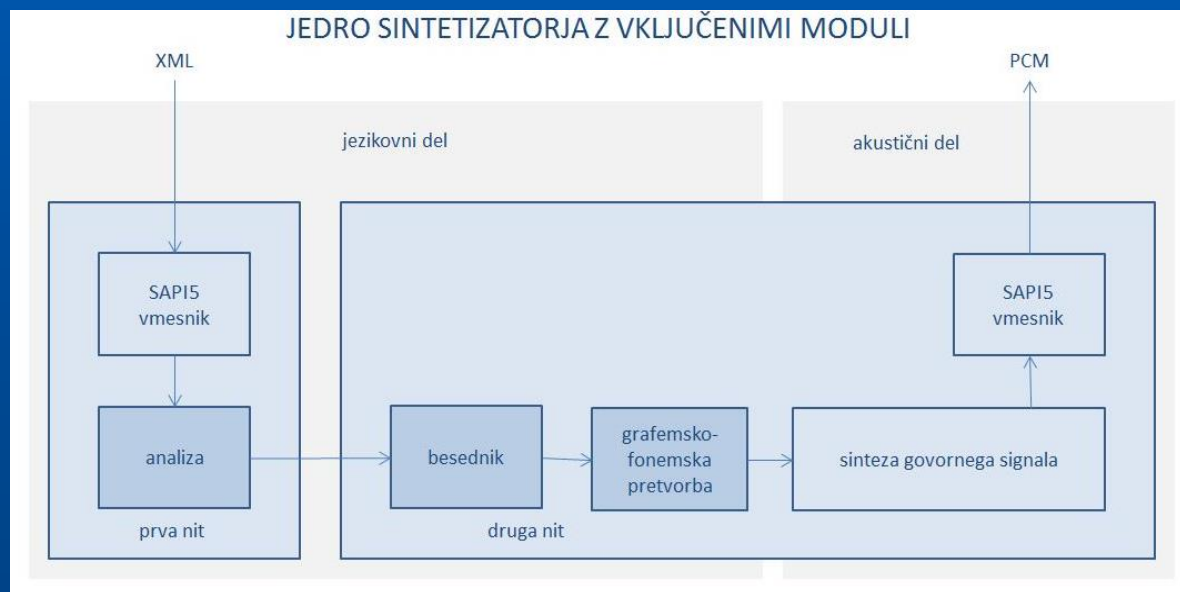
- **2015** eBralec – SAPI 5
 - **2015** eBralec KSS – Android
 - **2016** eBralec server – VoiceXML
 - ... eBralec server – MRCP
- Projekt KSS
 - Ekipa eBralca
 - Alpineon razvoj in raziskave, d.o.o.
 - Amebis programska oprema, d.o.o.
 - Institut "Jožef Stefan"
 - Univerza v Ljubljani – FE



Nalozba v našo prihodnost
OPERACIJO DELNO FINANCIRA EVROPSKA UNIJA
Evropski socialni sklad



- Jezikovna analiza
- Akustična analiza





- **Jezikovni analizator:** pravila + jezikovna baza ASES – glagolske predloge (vezljivosti glagola); 257 K lem, 8,1 M pregibnih oblik
- **Segmentacija na povedi, stavke in besede**
 - **Akronimi, krajšave, skladenjska oz. neskladenjska raba znakov za ločila**

Videl sem **ga**. Mojca ga je tudi videla. Bil je na **28**. Mednarodnem festivalu, tako kot **ga**. Mojca včeraj zvečer.

- **Lematizacija in oblikoskladenjska anotacija**
 - **razdvoumljanje – homografi**

Ker **je** omaka slana, **je** on ne **je** veliko.

Ob **pol** sedmih se je **pol** črede odpravilo na severni **pol**.

Samo jo je **samo** hotel pustiti **samo**.

Gori na **gori gori**.

Vsak **nosi** svoje stvari, **nosi** jih še ti!



- **Uspešnost jezikovne analize:** 99,03 % lem in 96,00 % oblikoskladenjskih oznak (vzorec jos100k)

- **Grafemsko-fonemski prepis:** slovarji izgovarjav in KSP

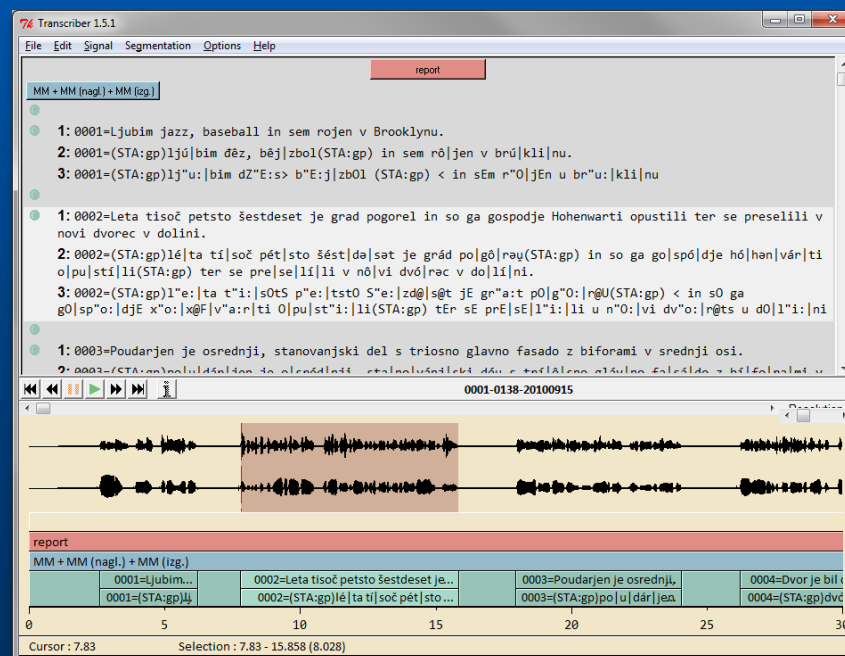
- **Nadaljnje delo:**

- pohitritev pri analizi zapletenih povedi:
 - dolga naštevanja, odločanje med desnimi prilastki in prislovnimi določili
- konsolidacija in kompakten zapis slovarjev izgovarjav (Alpineon, Amebis)

grafemski zapis	zlogovani zapis
bele	*be:.lE
lica	*li:.tsa
razvoja	raz.*vO:.ja
lepe	*le:.pE
sli{ati	*sli:.Sa.ti
pozabil	pO.*za:.biw
visoke	vi.*sO:.kE
stro{ke	*stro:S.kE
.....	
{portih	*Spo:r.tih
{tedilnikom	StE.*di:l.ni.kOm
{tevilno	StE.*vi:l.nO
{torijo	StO.*ri:.jO
{umele	Su.*me:.lE
{vedskih	*Sve:ts.kih
~ajni	*tSa:j.ni
~akalnici	tSa.*ka:l.ni.tsi
~arobno	tSa.*rO:b.nO



- **Izbor povedi** – 4.000 povedi iz korpusa FidaPLUS
 - reprezentativni glasovni sklopi, nakloni, raznovrstnost povedi....
- **Snemanje** – studio Radia Slovenija, 19 h posnetkov
- **Obdelava posnetkov**
 - segmentacija na povedi
 - ročno prilagajanje fonetičnih transkripcij dejanskemu izgovoru



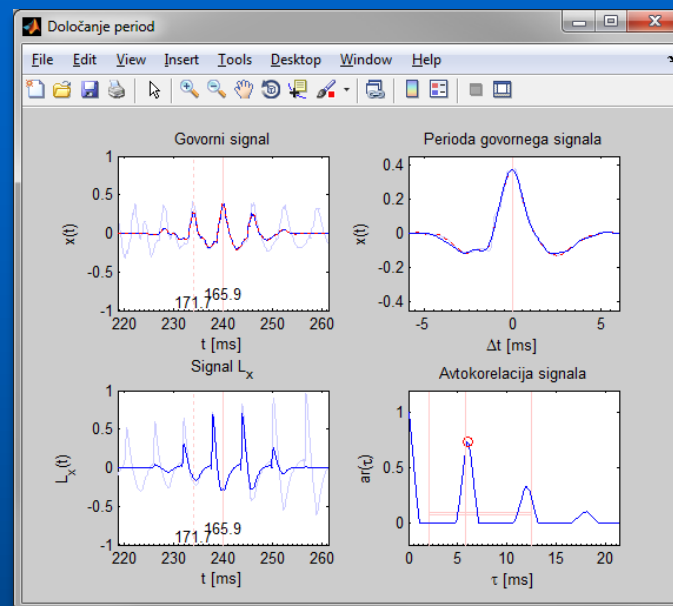


- ... obdelava posnetkov

- označevanje glasov
- označevanje osnovnih period

- Govorna zbirka – obseg

- Besedilni del
 - 4.000 povedi, 46.785 besed, 1.883 različnih difonov oz. 21.236 trifonov
- Govorni del
 - 6 ur 3 min ženski glas, 5 ur 33 min moški glas



● Konkatenativna sinteza govora

- difoni, trifoni, *unit selection*: daljši govorni segmenti – manj nezveznosti na mestih združevanja
- najbolj skladno zaporedje enot obtežena odločitvena drevesa

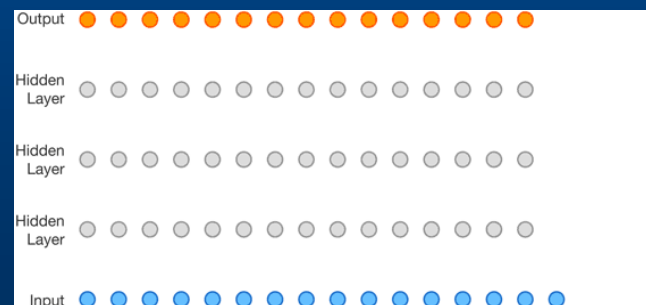


● Parametrična sinteza govora

- modeliramo vsak glas posebej, glede na kontekst
- sočasno se modelirajo: frekvenčni spekter frequency spectrum (govorni trakt), osnovna frekvenca (govorni vir – nihanje glasilk) in trajanje (prozodija)

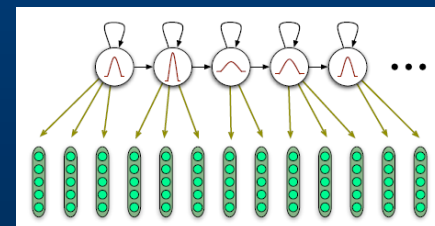
● Globoke nevronske mreže DNN

- parametrične (Zen et al., 2013)
- WaveNet (van den Oord et al, 2016)



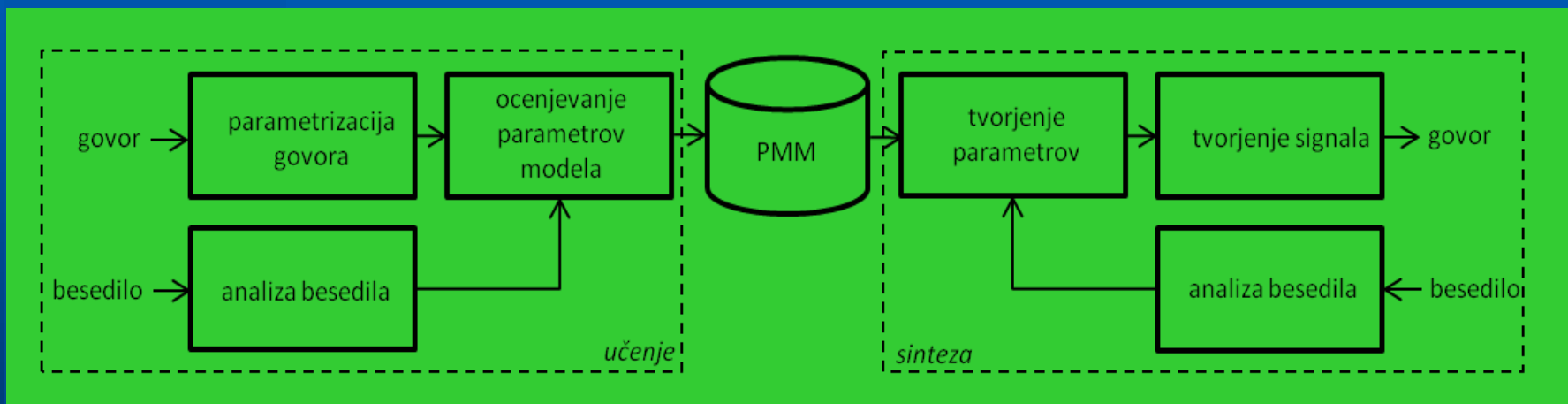


- Uporabniške zahteve narekovale parametrično predstavitev zakonitosti govora
- Prednosti uporabe PMM za sintezo govora
 - Zadošča razmeroma majhna govorna zbirka
 - Ni potrebe po precizni fonemski označitvi
 - Sočasno modeliranje akustičnih in prozodičnih lastnosti govora
 - **Kompakten zapis** modela govora
 - **Visoka naravnost prozodije** tvorjenega govora
 - Adaptacija – možnost tvorjenja **individualnih glasov** z malo podatki





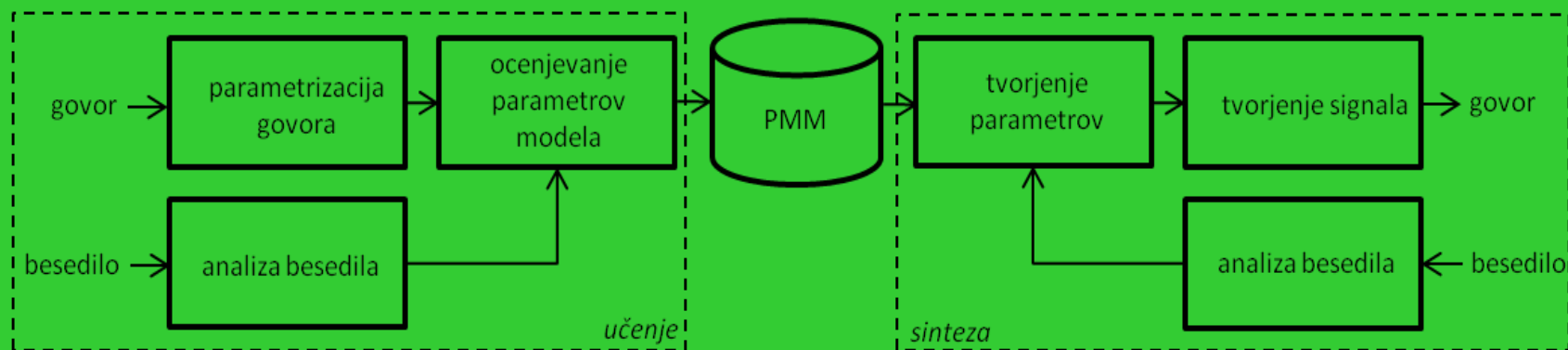
- **Učenje:** ocena parametrov modela
- **Sinteza:** model se uporabi za generiranje govora



- **Zgoščen parametrični zapis govora:** 108-razsežni vektor značilk
 - mel cepstrum + log freq, dinamične značilke 1. in 2. reda



- **Učenje:** ocena parametrov statističnega modela – največje verjetje
- **Sinteza:** iskanje optimalnega niza parametrov govornega signala



- **Osnovna govorna enota:** alofoni v kontekstu
 - predhodni /sledеči fonem, mesto fonema v zlogu, mesto zloga v besedi/stavku, mesto besede v stavku, je fonem v poudarjenem/nepoudarjenem zlogu, razdalja do poudarjenega zloga, dolžina prejšnjega/trenutnega/naslednjega stavka, stavčni naklon ter število zlogov/besed/stavkov v povedi

- *Predstavili smo zasnovo in izvedbo novega visokokakovostnega sintetizatorja govora za slovenski jezik. eBralec bo slepim in slabovidnim uporabnikom znatno olajšal delo z računalnikom, dostop do novic in informacij ter tako omogočil njihovo boljšo vključenost v sodobno informacijsko družbo.*

- **Predhodniki**



<http://alpineon.si/proteus/test/slo.html>



- **ženski glas**
- **moški glas**



Maja eBralec



Renato eBralec

<http://ebralec.si/branje/>

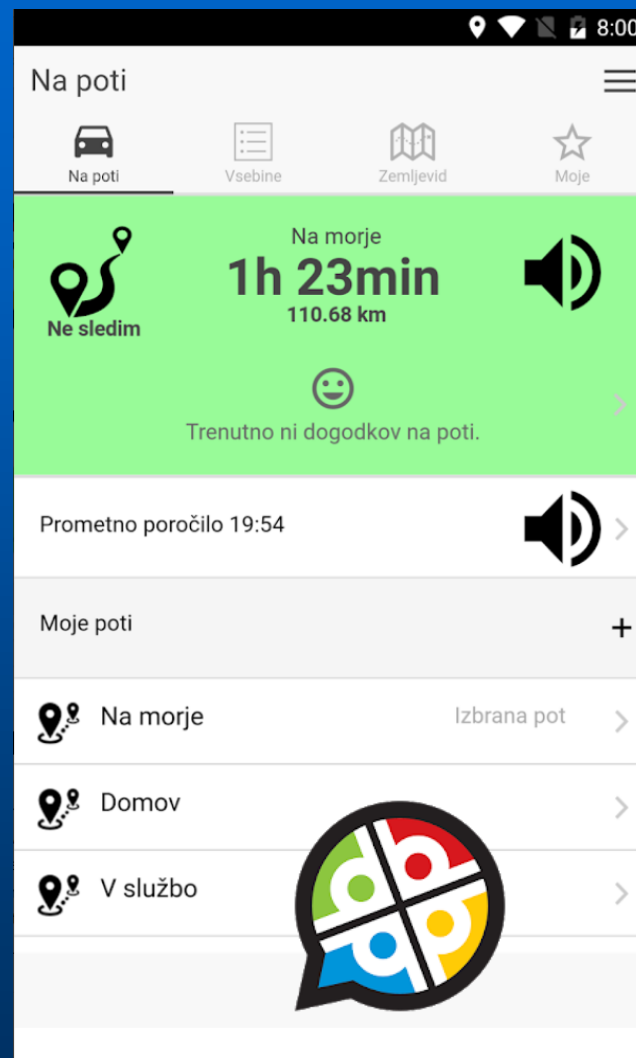


- **Naravna** sinteza govora za slovenščino
- Glasovi: **moški** in **ženski**
- Več milijonov slovarskih gesel, dodajanje novih besed
- Nastavljiva parametrizacija govora: trajanje, ritem, naglasi...
- Visoka zmogljivost procesiranja, hitra odzivnost
- Eno-uporabniška in zmogljiva strežniška izvedba
- **Uporaba za slepe in slabovidne** – kot podporna tehnologija:
 - bralniki in povečevalniki zaslona (Jaws, Virgo, Zoomtext...)
 - namenske aplikacije
 - branje podnaslovov ter opisov prizorov (RTV)

- **Glasovni SMS** – glasovni bralnik SMS sporočil
- **Glasovne e-pošta** – glasovni bralnik e-pošte
- **Glasovni splet** – glasovno branje spletnih strani (NUK, Primorski dnevnik...)
- **Klicni centri**, govorni portali
- Virtualni pomočniki
- Navigacija



- **Mobilne aplikacije** (DarsPromet+)
- **iSlovar** – zvočne podobe gesel
- **VoiceTRAN**: večjezični prenosni govorni komunikator, MORS
- **Pametni dom**



- **Trenutno stanje – sinteza govora**
 - Naraven, razumljiv in tekoč govor
 - Moški & ženski glas
 - Strežniški TTS, eno-uporabniški TTS, Android
- **Načrti za prihodnost**
 - Novi glasovi, čustva, izboljšave (DNN...)
 - Nove platforme: ostali mobilni telefoni in tablice



alpineon))

Nadaljnje informacije



dr. Jerneja Žganec Gros
Alpineon, Ljubljana, Slovenia
jerneja.gros@alpineon.si
+386 1 423 9440
www.alpineon.si

Hvala za pozornost!