

Говорна комуникација с машинама на српском језику

Проф. др Владо Делић
Факултет техничких наука, Нови Сад
АлфаНум – говорне технологије

Најава садржаја предавања

1.

- Од говорне комуникације
- до говорних технологија

2.

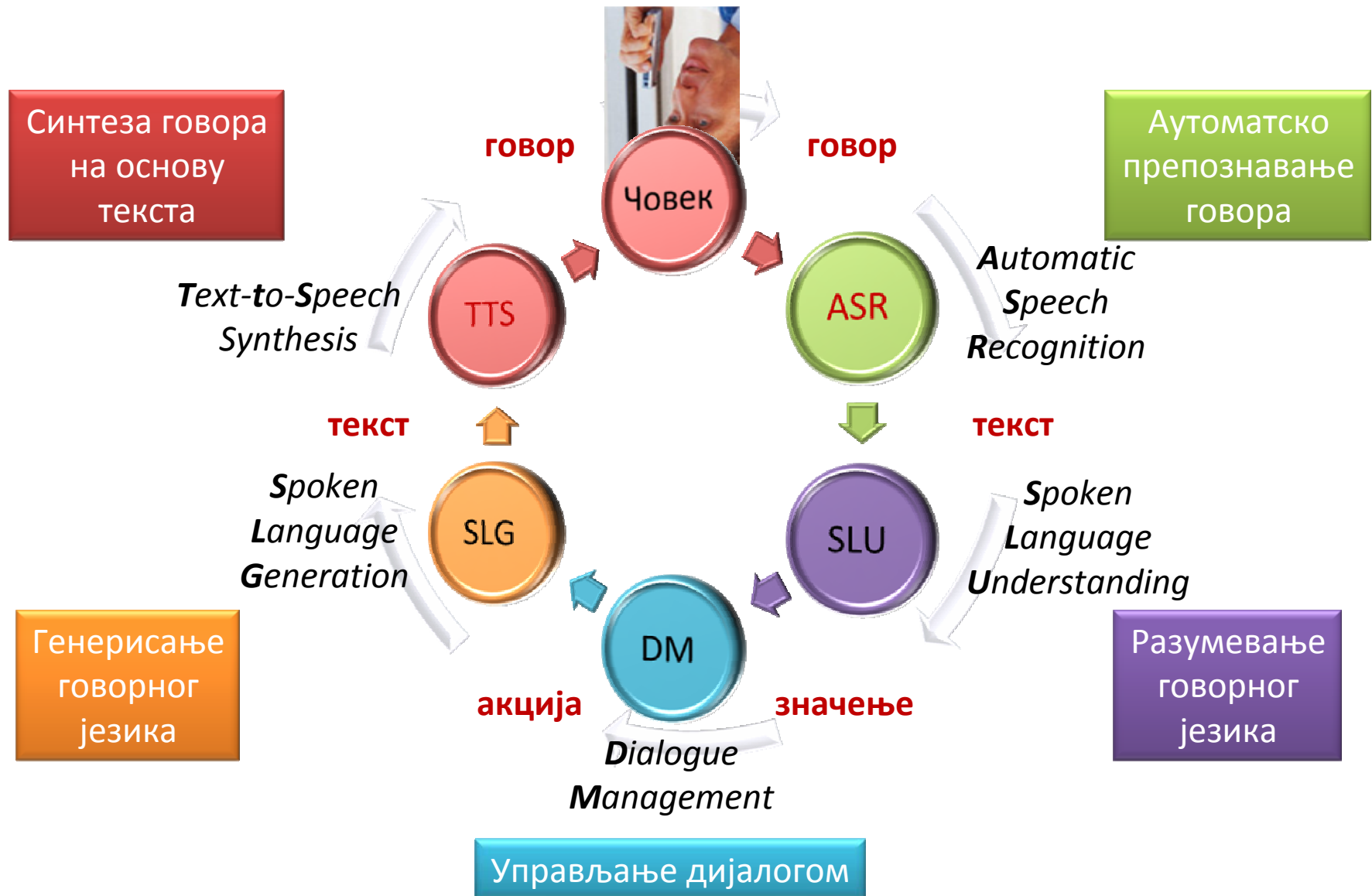
- Могућности и изазови у развоју
- достигнућа за српски језик

3.

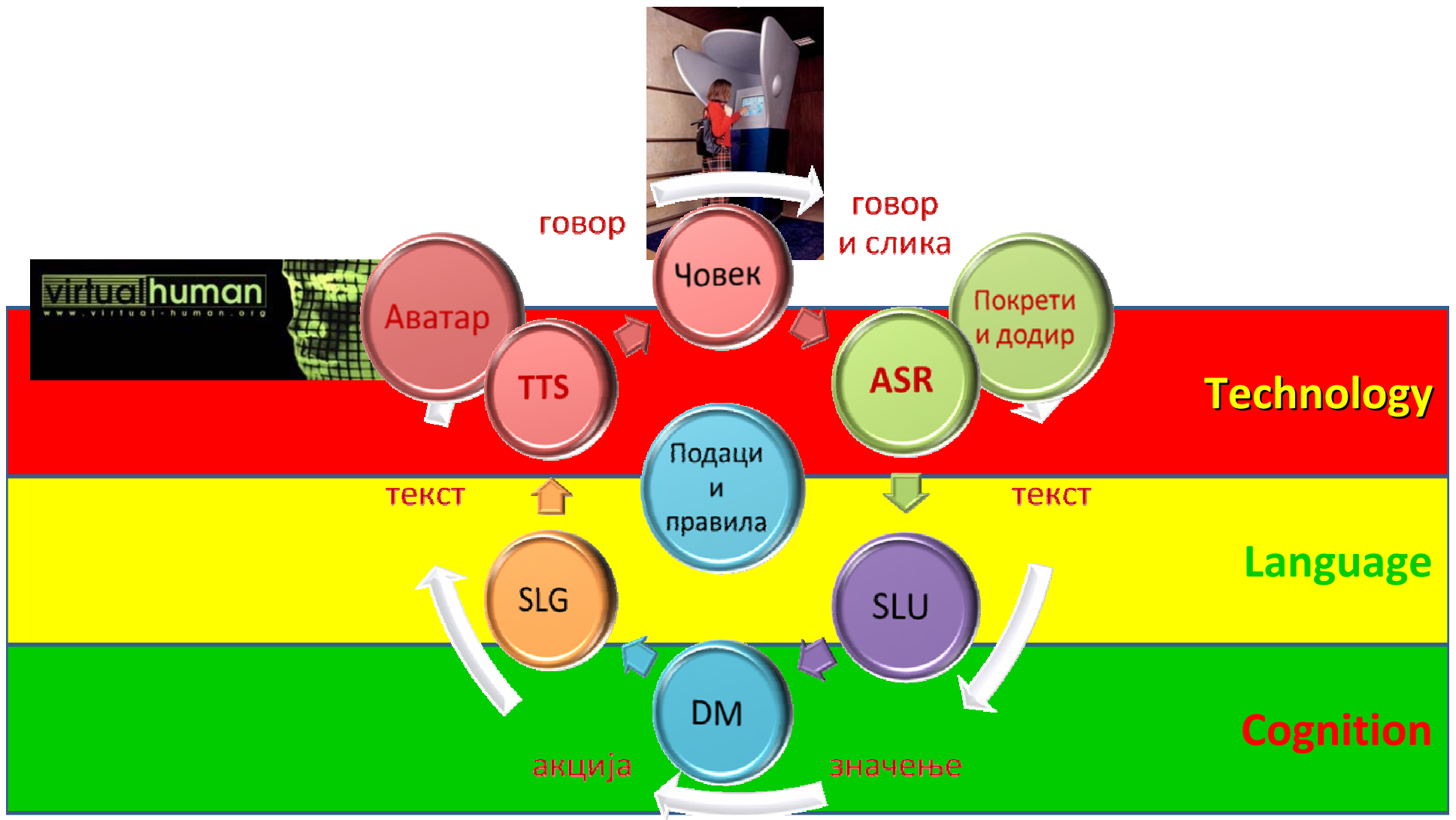
- Значај и могућности примене
- демо апликације на српском



ASR и TTS у дијалогу човек-машина



Мултимодални ASR и TTS

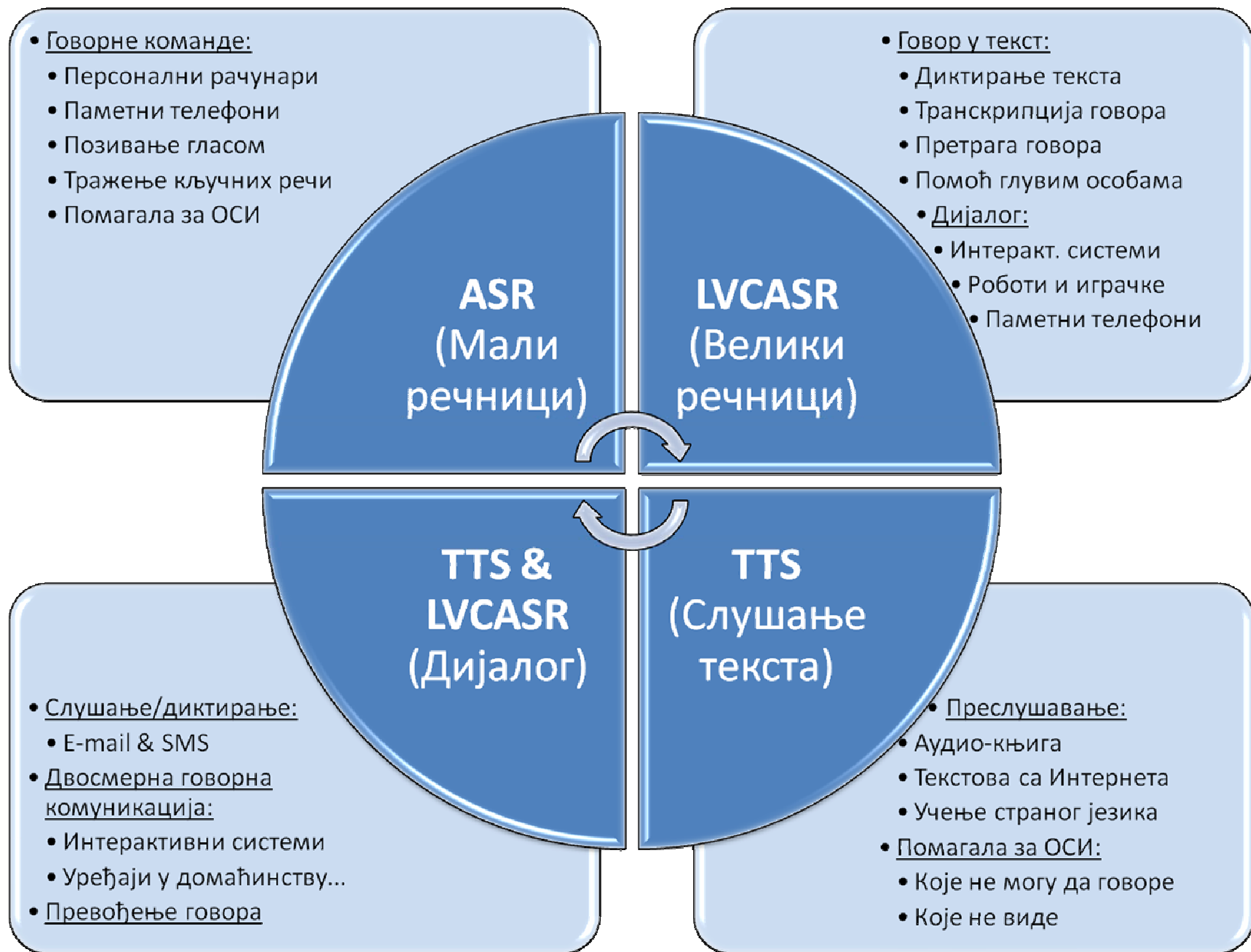


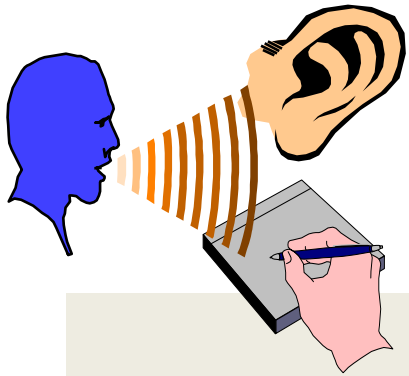
Мултидисциплинарни приступ



Примене ASR и TTS







Проблеми и изазови

ASR

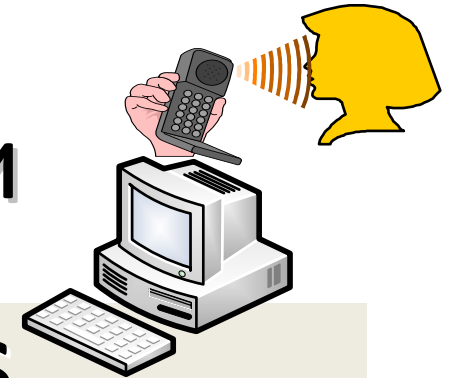
Препознавање говора

- говор → текст

Разне варијабилности:

- разни језици и речници
- различит глас говорника
- амбијент и бука
- положај микрофона
- комуникациони канал

Предност човека: има намеру, разуме контекст и прати смисао.



TTS

Синтеза говора из текста

- текст → говор

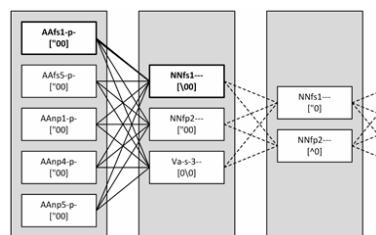
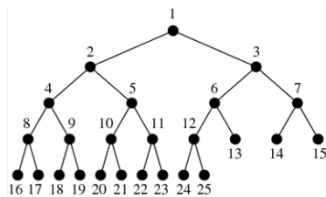
Текст без прозодије:

- колико траје изговор гласа
- да ли је у наглашеном слогу
- који је од акцената
- како изговорити слово, слог, реч и реченицу

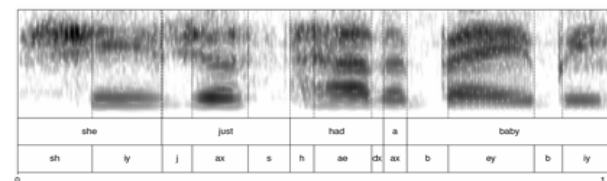
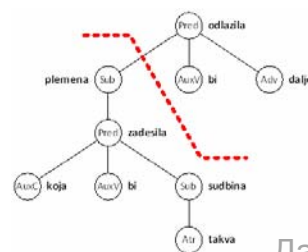
Говорно-језички ресурси

Моделовање говора у ASR

- Акустички модел (HMM)
 - говорна база
 - анотирана, проверена; 20h
- Језички модел (*N*-gram)
 - текстуални корпус
 - морфолошки анотиран; 100М речи
- Модел изговора
 - морфолошки речник
 - фонеме, акценти; 4М речи



meti meti ["0] Vd-s-2-
metak metak ["0] NNms4-
metak metka ["0] NNms2-
metak metke ["0] NNmp4-
metak metkom ["0] NNms6-
metak metku ["0] NNms5-
metak metku ["0] NNms3-
metta metta ["0] NNis1-
metta mettama ["0] NNip3-
metta mettama ["0] NNip6-
metta mette ["0] NNip4-
metta mette ["0] NNip5-
metta mette ["0] NNis2-
metta mette ["0] NNip1-
metta metti ["0] NNis3-
metta metti ["0] NNip2-



Језичка обрада текста у TTS

- Предобрада текста
 - претпроцесор
- Морфолошка анализа и анотација; фонетизација
 - морфолошки речник
 - POS tagger
- Синтаксно-прозодијска анализа и генерисање прозодијских обележја
 - детекција прозодије из текста
 - предикција f0 и трајања гласова

Перспективе говорне комуникације ЧОВЕК-МАШИНА

Предности:

- Усмена комуникација
 - **очи** и **руке** су слободне
 - приступ из даљине
 - гласом** (преко телефона)
- Развој технологија
 - екрани и тастатуре не могу бити много мањи
 - процесори су све бржи, а меморија све већа
 - комуникациони линкови су све бржи



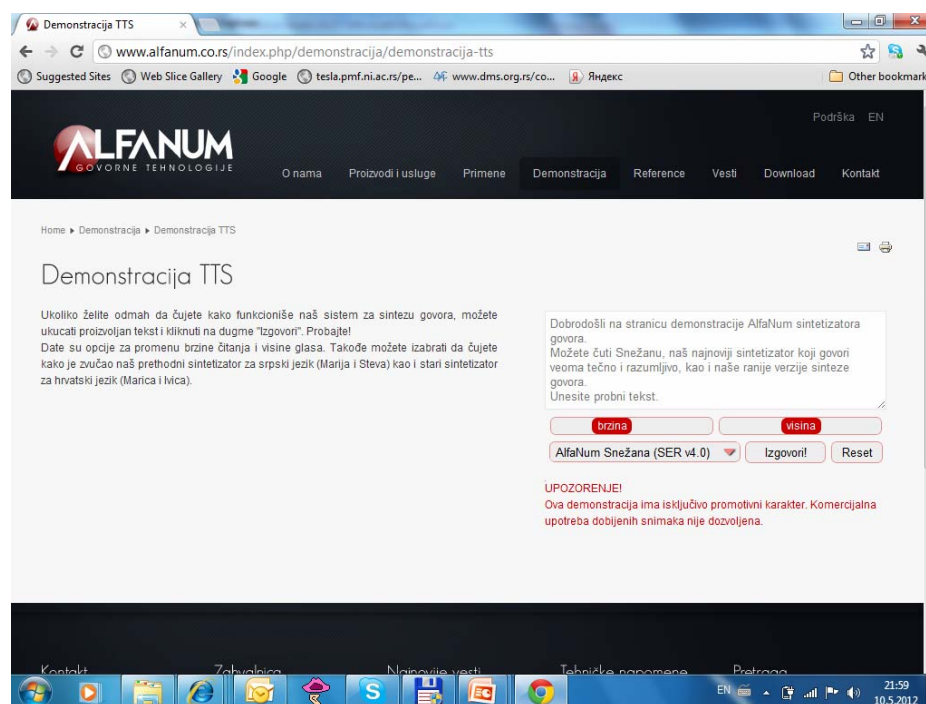
Мане:

- Спонтан говор тежак за ASR
- Синтетизовани говор (TTS) није изражајан као човеков
- Дијалог човек-машина на бази ASR и TTS је крут
- Зависност од језика
 - не може се увести из иностранства као друге нове технологије

Достигнућа у развоју за српски

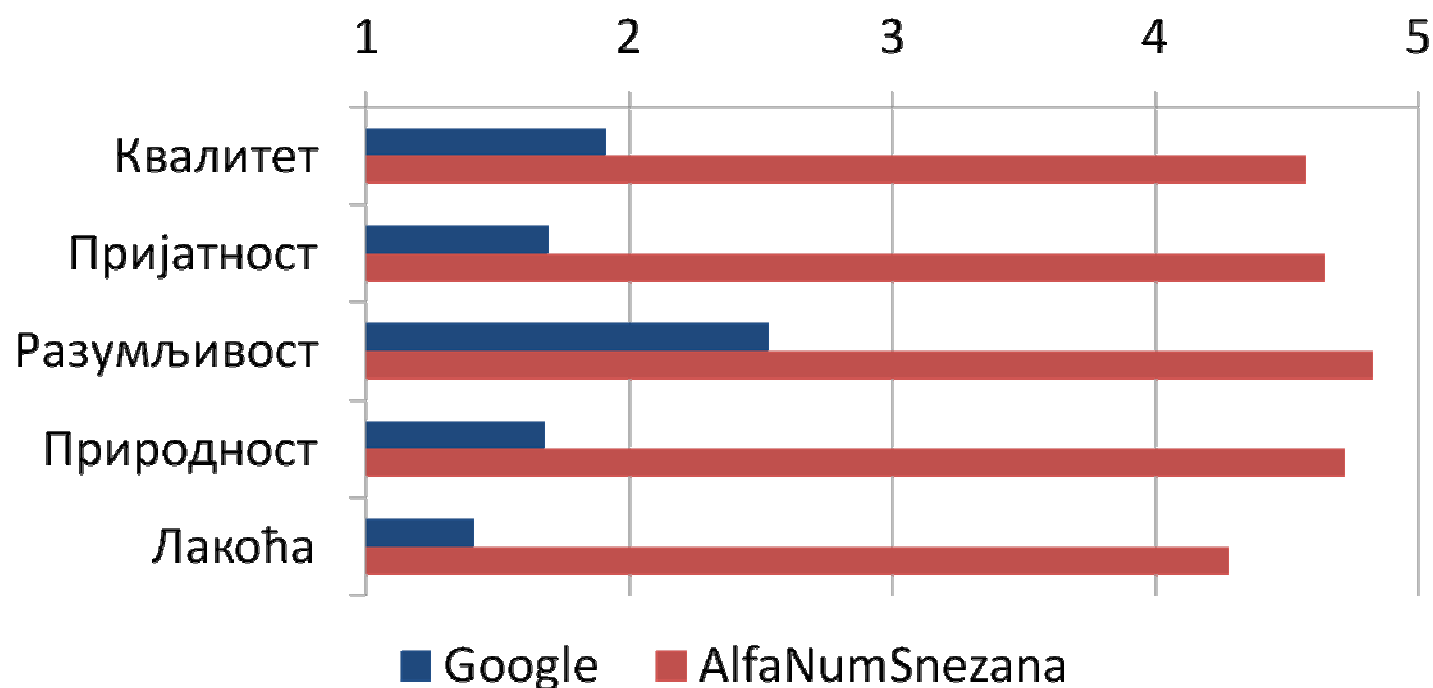


<http://www.alfanum.ftn.uns.ac.rs/>



<http://www.alfanum.co.rs/>

Квалитет синтезе говора



ЧИТ@Ј МИ! :: РАДИО - ТЕЛ

www.rtv.rs/sr_ci/citaj-mi/

Suggested Sites

Web Slice Gallery

Google

tesla.pmf.ni.ac.rs/pe...

www.dms.org.rs/co...

Яндекс

Other bookmarks

РТВ

РАДИО ТЕЛЕВИЗИЈА ВОЈВОДИНЕ

ИЗБОРИ 2012

ВЕСТИ

РТВ

ПРОГРАМ

МАРКЕТИНГ

КОНТАКТ

БЛОГ

СТУДИО М

ОДЛОЖЕНО ГЛЕДАЊЕ

PODCAST

ГРУВАЊЕ

ЧИТАЈ МИ!

СЕРВИСИ

НАЈВАЖНИЈЕ ВЕСТИ : "ЖАБАЦ" И ДАЉЕ У ЖИВОТНОЈ ОПАСНОСТИ

Претрага вести

Lat

Hu

КУРСНА ЛИСТА

10. 05. 2012.

EUR

112.15

USD

86.61

CHF

93.31

GBP

139.78

BAM

57.34

ВЕРСКИ КАЛЕНДАР

27. април -

Спаљивање

моштију Светог

Саве на Врачару

Иван Мерз,

Госпа Трсатска,

Антонин

19. Цумадел-

ухра 1433.

18. Ијар 5772.

twitter

@RTVojvodine

Follow

863 follow

ЧИТ@Ј МИ!

СИНТЕЗА ГОВОРА

www.rtv.rs

Радио-телевизија Војводине прва је у Србији и региону, на свом интернет порталу www.rtv.rs, увела сервис **Чит@ј ми!**, који вам пружа могућност да вести које сте до сада читали - од сада и слушате!

Чит@ј ми! је синтетизатор говора на основу текста који ће бити од изузетне помоћи особама са посебним потребама, слабовидим и слепим, којима ће бити омогућено да на лакши и бржи

Србија лидер у региону у области говорних технологија

Настанак синтетизатора говора на основу текста прилагођеног особеностима српског језика плод је знања и научно-истраживачког рада наших стручњака окупљених на Факултету техничких наука у

Нови Сад

Тренутно стање

20°C

1014mb

1m/s

Детаљније

Радио и ТВ уживо

ТЕЛЕВИЗИЈА 1

ТЕЛЕВИЗИЈА 2

РАДИО 1

РАДИО 2

РАДИО 3

Сучељавање, политичка

дебатна емисија

21:04

Војвођански дневник

22:00

Чарде на Дунаву

22:36

Последњи бег на Балакану,

серија

23:06

Један на један

00:04

БЕЛА СТРАНИЦА

EN

21:45

10.5.2012

Alfanum audio biblioteka

anAudioBiblioteka 2.0

Profili

Podešavanja

Preuzmi novu listu

Aktivni server

ABSS Niksic

Autor	Naziv	Tip/Žanr	Izdavač	Jezik
☐ Goran Maca...	O audio biblioteci	Tekst	Organizacija sli...	srpski
☐ Vojislav Voj...	Dedović, dila	...	...	...
☐ Vojislav Voj...	...	...	...	...
☐ Vojislav Voj...	...	...	...	...
☐ Vojislav Voj...	...	...	...	...
☐ Vojislav Voj...	...	...	...	...
☐ Vojislav Voj...	...	...	...	...
☐ Uje	...	...	...	...
☐ Ne	...	...	...	...
☐ Dž	...	...	...	...
☐ Dž	...	...	...	...
☐ Zvo	...	...	...	...
☐ Zvo	...	...	...	...
☐ Zvo	...	...	...	...
☐ Zvo	...	...	...	...
☐ Zvo	...	...	...	...
☐ Zvo	...	...	...	...
☐ Vla	...	...	...	...
☐ Voj	...	...	...	...
☐ Voj	...	...	...	...
☐ Voj	...	...	...	...
☐ nez	...	...	...	...
☐ nez	...	...	...	...
☐ nez	...	...	...	...
☐ nez	...	...	...	...
☐ nez	...	...	...	...
☐ Ne	...	...	...	...
☐ Prof. Dr. Ljilja...	Dizavno pravnih stonja Cine dore	Fakultetski udz...	Nar. i obreda	srpski

Sintetizuj tekst

Sintetizuj u fajl

Snimi na DATA CD

Snimi na Audio CD

Izlistaj tekstove

Pretraži tekstove

Zatvori

Dijalog za sintezu teksta

Slušanje knjiga

Pauza

Zatvori

Sledeći pasus

Sledeće poglavlje

Prethodni pasus

Prethodno poglavlje

poglavlje:

bookmark:

Postavi novi bookmark

Obriši odabrani bookmark

Preslušano 42.3% teksta

AlfaNum audio biblioteka Проф. В. Делић, ФТН и АлфаНум, Нови Сад Дан језичких технологија, 29. октобар 2012. Београд

Титловање ТВ програма

- Важно за особе које не чују
- Активира се преко телетекста
- Прва фаза:
 - синхронизација говора и познатог текста
- Друга фаза:
 - препознавање говора (читање спикера)
- Трећа фаза:
 - препознавање спонтаног говора (тренирани спикери)



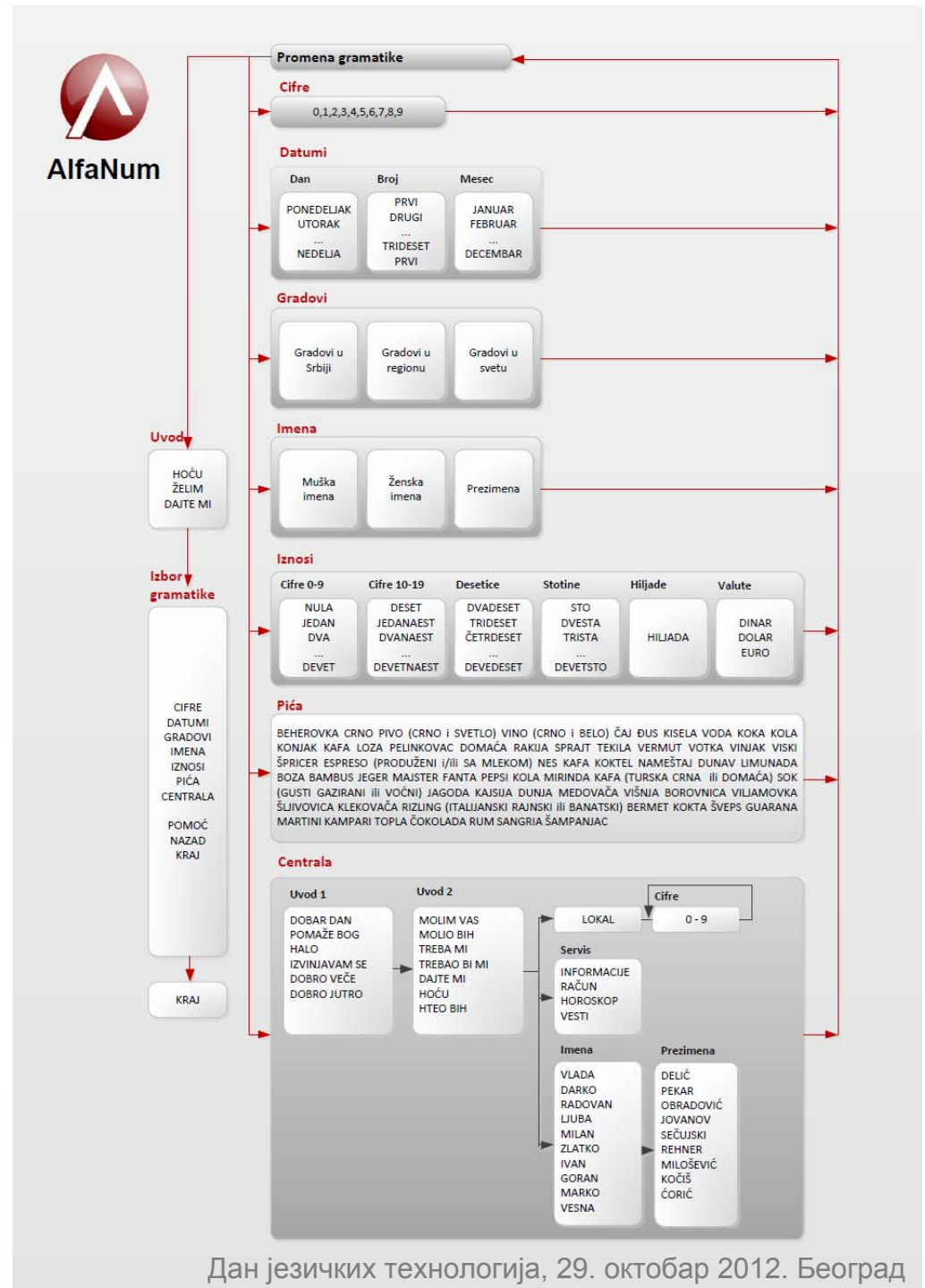
Демонстрација ASR и TTS

- Препознаје говорне команде - **ASR**
- Одговара синтетизованим говором – **TTS**



021/475-0204

www.alfanum.co.rs
www.alfanum.ftn.uns.ac.rs



Телефонски говорни аутомати

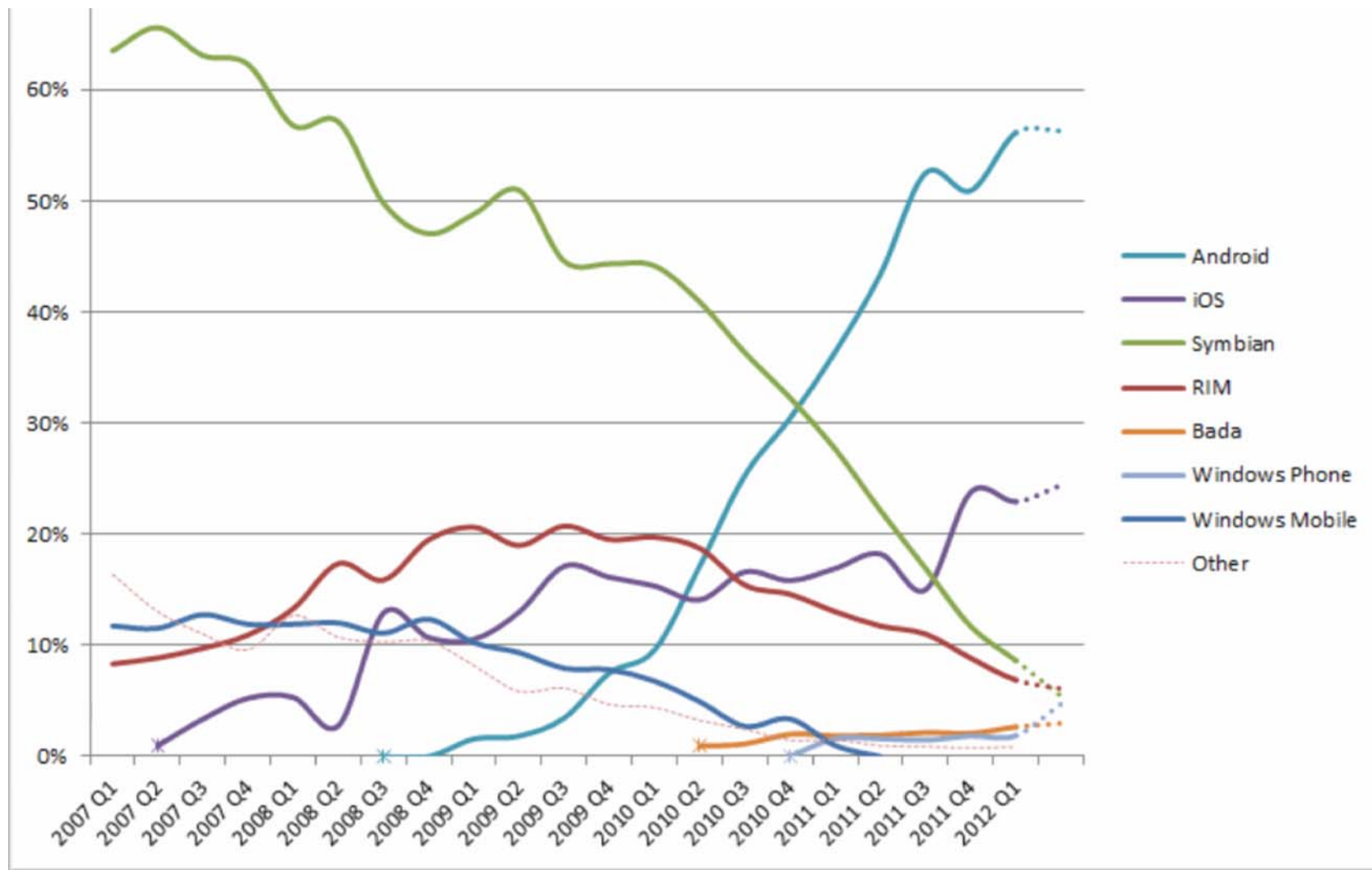
- Основне функционалности:
 - Препознају говорне команде (ASR)
 - Одговарају синтетизованим говором (TTS)
- Примери телефонских сервиса:
 - информације (време, путеви, култура)
 - снимање огласа, рекламација и резервација
 - ред вожње, телемаркетинг, забавни сервиси



рање локала и говорна пошта



Продаја паметних телефона у свету



Апликације за мобилне телефоне



- Све више информација преко телефона

- Интернет доступан преко телефона

- мала тастатура и екран
- усмена питања и одговори су најприроднији
- руке и очи су слободне



- Говорна комуникација с телефоном (*Ахон - IS*)

- “Пронађи ми SMS поруку од Дарка од јуче.”

- “Прочитај ми ту поруку.”

- “Назови Милана Сечујског.”

- “Израчунај ми 27% од 3000.”

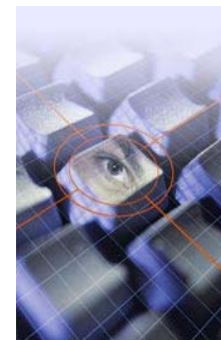
- “Подсети ме да сутра у 12 сати назовем шефа.”



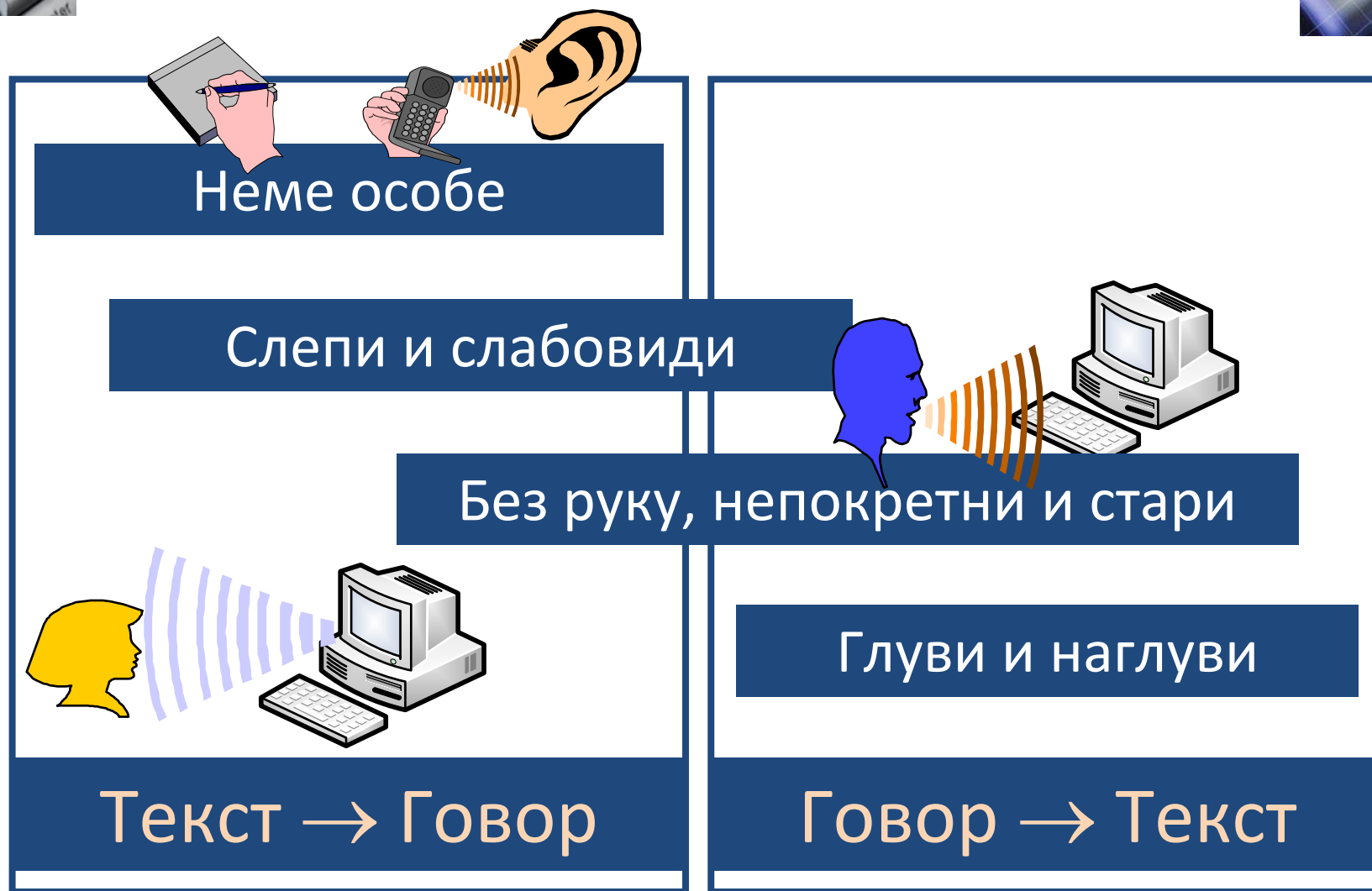
Нови пројекти

- Робот за помоћ при терапији (МАРКО)
 - говорне и моторичке вежбе
 - мотивише и доприноси истрајности
- Говорне технологије у паметним кућама
 - говорне команде уређајима у домаћинству
 - праћење деце, старих, особа са инвалидитетом
 - синтеза говора из текста
 - препознавање говора, говорника и емоција
 - препознавање алармантних звукова
- Примене за помоћ особама са инвалидитетом
 - рачунар чита наглас онима који не виде
 - говори уместо оних који не могу да говоре
 - претвара говор у текст за оне који не чују
 - разуме говорне команде оних који се тешко крећу





Резиме: Кома могу да помогну говорне технологије?



Једна анегдота за крај

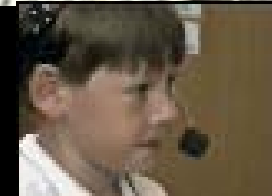
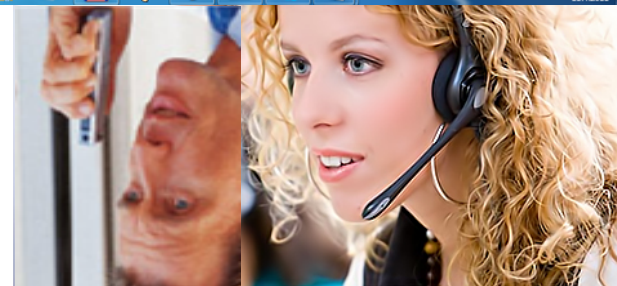


Професор Делић је једном приликом требало да одржи предавање на скупу на ком су истовремено биле и особе које не виде и особе које не чују. Написао је све што је желео да каже, а ја сам, као говорна машина, одржала предавање. Слушале су ме особе које не виде, док су особе које не чују пратиле приказани текст, а професор је могао да ћути и посматра њихова лица, сакупљајући инспи-рацију за даље усавршавање говорних технологија. Тако може једна потпуно нема особа истовремено да комуницира и са глувим и са слепим особама - што је незамисливо без говорних технологија.



- ☛ Телефони,
- ☛ Паметне куће,
- ☛ Озвучени веб сајтови,
- ☛ Интерактивни говорни аутомати,
- ☛ Мултимодални информативни пултови...

ХВАЛА НА ПАЖЊИ!



ALFANUM
GOVORNE TEHNOLOGIJE

