



Center za  
jezikovne vire  
in tehnologije

# Kam se obrniti, ko zmanjka besed

Iztok Kosem

Filozofska fakulteta

[iztok.kosem@ff.uni-lj.si](mailto:iztok.kosem@ff.uni-lj.si)



Univerza v Ljubljani  
Fakulteta za računalništvo  
in informatiko



# Okvir CJVT

- Center za jezikovne vire in tehnologije Univerze v Ljubljani.
- Združuje pet fakultet UL: Filozofska fakulteta, Fakulteta za računalništvo in informatiko, Pedagoška fakulteta, Fakulteta za družbene vede in Fakulteta za elektrotehniko. (<https://www.cjvt.si/>)
- Namen CJVT je omogočiti uspešen razvoj slovenščine v digitalni dobi.
- Pripravljamo jezikovne vire ter orodja, ki so vsebinsko in metodološko **sodobni**, uporabniku **prijazni** in **odprto dostopni** za uporabo.
- CJVT je prva slovenska raziskovalna inštitucija, ki se ciljno ukvarja z empiričnimi analizami **težav**, **mnenj**, **navad** in **potreb** jezikovne skupnosti v kontekstu priprave digitalnih jezikovnih virov za slovenščino.

# Mednarodna uporabniška raziskava

- situacije **produktivne** rabe slovarjev v Sloveniji:
  - **73 %** - za pripravo besedil, ki so del službenih nalog ali izobraževalnega procesa (poročila, domače naloge, seminarske naloge itd.)
  - **44,7 %** - za pisanje pomembnih besedil (prijava na delovno mesto, življenjepis itd.)
  - **26 %** - za kreativno pisanje (poezija, besedila pesmi, zgodbe itd.)
  - **25 %** - za pisanje e-sporočil in spletnih objav (na forumih, blogih ipd.)

# Vam kdaj zmanjka besed?

- VAJA 1: napišite čim več sopomenk (besed, ki imajo (skoraj) enak pomen) za:
  - pohlep
  - pozitiven
  - spati

# Slovar sopomenk sodobne slovenščine

- **Odzivnost:**
  - (a) na spremembe v jeziku
  - (b) na jezikovno skupnost
- **Sodobnost:**
  - temelji na sodobnih jezikovnih podatkih
  - je pripravljen z metodami, ki so interdisciplinarne in inovativne
- **Odprtost:**
  - Slovar sopomenk sodobne slovenščine je največja odprto dostopna zbirka sopomenk za slovenščino.

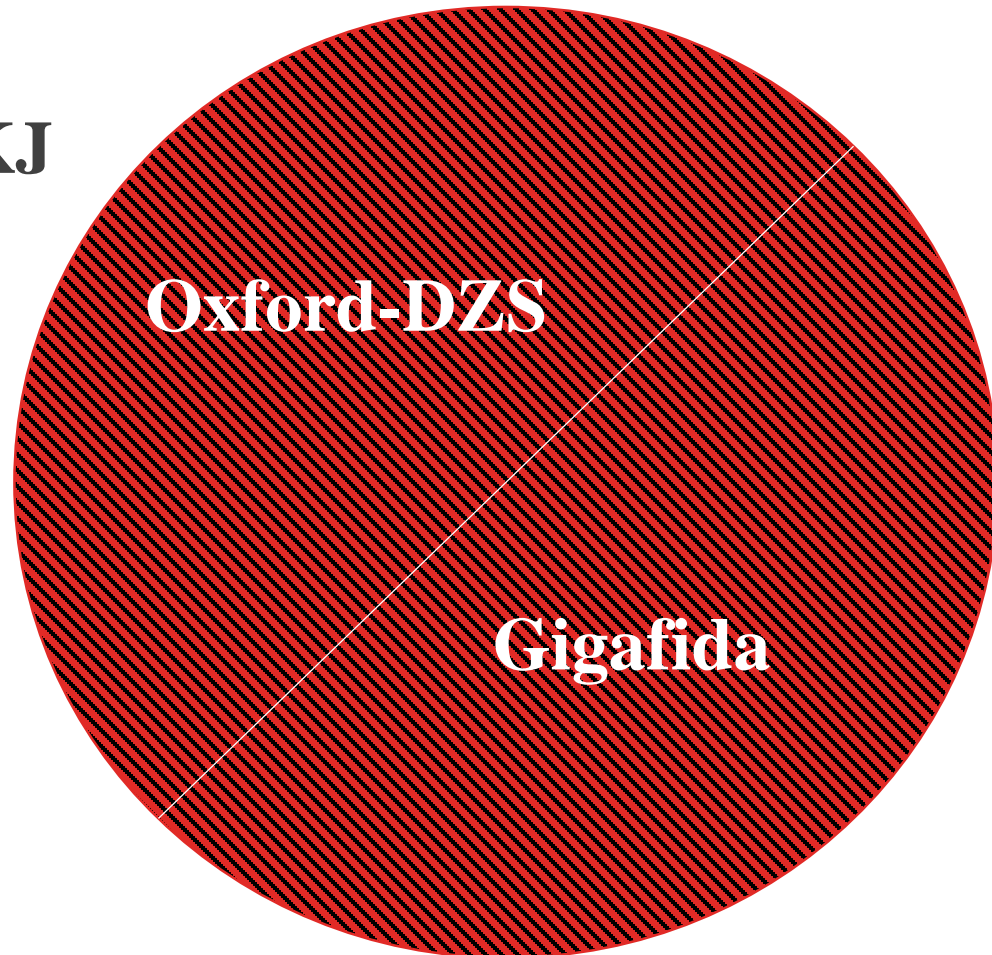
# Zasnovanost podatkovne baze



- Temeljna podstat Sopomenk sta Veliki angleško-slovenski slovar Oxford DZS<sup>®</sup> in referenčni pisni korpus Gigafida.
- Jezikovni podatki po letu 1990.

# Zasnovanost podatkovne baze

SSKJ



- Temeljna podstat Sopomenk sta Veliki angleško-slovenski slovar Oxford DZS® in referenčni pisni korpus Gigafida.
- Jezikovni podatki po letu 1990.
- Dodatno preverjanje podatkov v SSKJ.
- Izrabljivost virov za razvoj novih izdelkov.

# Od prevodnih ustreznic k Sopomenkam

- Oxford DZS – prevodne ustreznice angleških besed in besednih zvez

*dom; bivališče, domovanje; stanovanje; hiša*



# Od prevodnih ustreznic k Sopomenkam

- Oxford DZS – prevodne ustreznice angleških besed in besednih zvez

*dom; bivališče, domovanje; stanovanje; hiša*

# Od prevodnih ustreznic k Sopomenkam

- Oxford DZS – prevodne ustreznice angleških besed in besednih zvez

*dom; bivališče, domovanje; stanovanje; hiša*



# Sopojavitveni grafi



# Sopomenke

- VAJA 2: Uporaba slovarja sopomenk
  - <http://viri.cjvt.si/sopomenke/slv/>
  - pohlep

# Kolokacije (tipične sopojavitve besed)

- VAJA 3:
  - pozitiven | pritrdilen

# Skupnost

- VAJA 4:
  - spati

# Uporabniške dejavnosti v številkah (1)

- **Podatki za 19. junij 2018 (tri mesece po izidu)**
- Število uporabniških sinonimov – 1.930
- Pozitivne ocene uporabnikov – 3.207
- Negativne ocene uporabnikov – 1.213
- Število uporabniških imen – 153
- Število iskanih iztočnic, za katere še ni podatkov – 2.179
- Število različnih iskanih iztočnic – 37.631 (35,7 % celotnega slovarja)
- Število IP-naslovov – 12.879

# Uporabniške dejavnosti v številkah (2)

- Najpogosteje iskane iztočnice:
  - **zelen**
  - koncept
  - hiša
  - **ideja**
  - **spati**
  - lep
  - pozitiven
  - odlično
  - kvaliteten
  - smotrno



# Sklepne misli

- Digitalno slovaropisje je nov model slovaropisja, ki odgovarja na nove potrebe in navade.
- Uporabniki na slovar ne vplivajo samo z neposrednim vključevanjem, ampak tudi s pričakovanji in neposrednimi odzivi na ponudbo na trgu.
- Slovar sopomenk sodobne slovenščine (odzivni slovar) obeta uspešen model za podporo slovenščine v digitalni dobi.
  
- Napovedujemo: Kolokacijski slovar sodobne slovenščine

**HVALA ZA POZORNOST**  
**ČUJEČNOST**  
**ZANIMANJE**  
**ZBRANOST**