

## Panelová diskusia 2

# Textové a jazykové zdroje na Slovensku: kde a čo máme, kde a čo hľadáme

**M. Šimková (SNK JÚLŠ SAV)**  
**L. Hluchý (ÚI SAV), J. Juhár (KEMR TUKE)**  
**D. Kluvanec (DGT EK)**



- **Verejne dostupné textové a jazykové zdroje  
SNK JÚLŠ SAV**
- **Dostupné zdroje ďalších pracovísk**
  - **ÚI SAV, TUKE, DGT EK**



- **Pozitíva vs. negatíva/problémy**
- **Potreba a užitočnosť d'alších zdrojov**
- **Zdroje potenciálnych zdrojov**



## textové zdroje – materiál, jazykové zdroje – databázy

- hlavný korpus súčasných písaných textov (publicistika, beletria, náučné texty)
- hovorený korpus, nárečový korpus, historický korpus
- paralelné korpusy vrátane dokumentov EÚ (en, cz, fr, de, hu, bg; ru, la)
- špecializované korpusy: právne, ekonomické texty
- špecializované databázy: terminologická, morfológická, kolokačná; prekladové frázy; slovníky

<http://korpus.sk>



**ŠIMKOVÁ, Mária — GARABÍK, Radovan — GAJDOŠOVÁ, Katarína — LACLAVÍK, Michal — ONDREJOVIČ, Slavomír — JUHÁR, Jozef — GENČI, Ján — FURDÍK, Karol — IVORÍKOVÁ, Helena — IVANECKÝ, Jozef:**  
**Slovenský jazyk v digitálnom veku – The Slovak Language in the Digital Age.**  
Berlin – New York: Springer 2012.

**<http://www.meta-net.eu/whitepapers/volumes/slovak>**



## negatíva/problémy

- **získavanie textov – autorský zákon**
- **zákony na ochranu osobnosti a os. údajov**
- **slabšia finančná a technická podpora**
- **ne/možná spolupráca s komerčnou sférou**



## pozitíva

- korpusy a databázy bezplatne online dostupné na vyhľadávanie jazykových informácií
- aktuálny materiál reálneho používania jazyka (frázy, ekvivalenty, synonymá, neologizmy ...)
- spracované údaje v slovníkoch a databázach
- bežné (nie len vedecko-výskumné) využitie



**plán: korpus textov štátnej a verejnej správy**



**terminológia štátnej a verejnej správy**





## **Ak by sa budoval slovenský korpus textov štátnej a verejnej správy,**

- koho možno/treba osloviť**
- aké a od koho na to treba súhlasy, zmluvy**
- sú informácie na webe automaticky voľné?**
- výročné správy, charakteristiky sekcií, odborov, oddelení, projektov, najčastejšie otázky (FAQ) a pod.**
- kde sa texty budú primárne zhromažďovať**
- formáty textov ...**