

# Reference in povezave

V tem razdelku so zbrane reference in povezave na projekte, v katerih se je označevalni sistem razvijal in uporabljal.

## **Projekti, na katerih se je označevalni sistem razvijal oz. uporabljal**

[MULTEXT-East - Multilingual corpora and text tools for Central and East European languages](#)

[JOS - Jezikoslovno označevanje slovenskega jezika: metode in viri](#)

[Sporazumevanje v slovenskem jeziku](#)

[Janes - Viri, orodja in metode za raziskovanje nestandardne spletne slovenščine](#)

[Razvoj slovenščine v digitalnem okolju](#)

## **Učni korpusi z ročno pregledanimi oznakami MULTEXT-East**

Arhar Holdt, Špela; et al., 2022, Training corpus SUK 1.0, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, <http://hdl.handle.net/11356/1747>.

Lenardič, Jakob; et al., 2022, CMC training corpus Janes-Tag 3.0, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, <http://hdl.handle.net/11356/1732>.

Krek, Simon; et al., 2019, Training corpus ssj500k 2.2, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, <http://hdl.handle.net/11356/1210>.

## **Reference**

Krek, Simon; Erjavec, Tomaž; Dobrovoljc, Kaja; Gantar, Polona; Arhar Holdt, Špela; Čibej, Jaka in Brank, Janez. The ssj500k training corpus for Slovene language processing. V: Fišer, D. in Erjavec, T. Jezikovne tehnologije in digitalna humanistika: zbornik konference: 24.-25. september 2020, Ljubljana, Slovenija. Ljubljana: Inštitut za novejšo zgodovino, 2020. Str. 24-33.  
[http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Krek-et-al\\_The-ssj500k-Training-Corpus-for-Slovene-Language-Processing.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Krek-et-al_The-ssj500k-Training-Corpus-for-Slovene-Language-Processing.pdf)     [\[PDF\]](#)

Erjavec, Tomaž. 2017. MULTEXT-East. V (Nancy Ide, James Pustejovsky, ur.): Handbook of linguistic annotation. pp. 441-462. Springer. DOI: [10.1007/978-94-024-0881-2\\_17](https://doi.org/10.1007/978-94-024-0881-2_17).

Erjavec, Tomaž. 2012. MULTEXT-East: morphosyntactic resources for Central and Eastern European languages. Language Resources and Evaluation, 46(1): 131-142. DOI: [10.1007/s10579-011-9174-8](https://doi.org/10.1007/s10579-011-9174-8)  
[\[PDF\]](#)

Erjavec, Tomaž; Fišer, Darja; Krek, Simon in Ledinek, Nina. 2010. The JOS Linguistically Tagged Corpus of Slovene. V: Proceedings of the Seventh conference on International Language Resources

and Evaluation (LREC'10), Valeta, Malta, Maj. European Language Resources Association (ELRA).

[http://www.lrec-conf.org/proceedings/lrec2010/pdf/139\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2010/pdf/139_Paper.pdf) [\[PDF\]](#)

---

Revision #16

Created 3 March 2023 10:56:45 by Tina Munda

Updated 5 December 2023 17:17:30 by Tina Munda