References and Links

This chapter compiles relevant references and provides links to projects where the JOS-SYN syntax has been developed and applied to Slovene texts.

Projects, in which the system has been developed or applied

JOS - Linguistic Annotation of Slovene: Methods and Resources

Communication in Slovene

Janes - Resources, Tools and Methods for the Research of Nonstandard Internet Slovene

Development of Slovene in a Digital Environment

Training corpora containing manually revised JOS-SYN tags

Arhar Holdt, Špela; Krek, Simon; Dobrovoljc, Kaja; Erjavec, Tomaž; Gantar, Polona; Čibej, Jaka; Pori, Eva; Terčon, Luka; Munda, Tina; Žitnik, Slavko; Robida, Nejc; Blagus, Neli; Može, Sara; Ledinek, Nina; Holz, Nanika; Zupan, Katja; Kuzman, Taja; Kavčič, Teja; Škrjanec, Iza; Marko, Dafne; Jezeršek, Lucija; Zajc, Anja, 2022, Training corpus SUK 1.0, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, http://hdl.handle.net/11356/1747.

The Q-CAT tool for manual annotation following the JOS-SYN system

Brank, Janez, 2022, Q-CAT Corpus Annotation Tool 1.4, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, http://hdl.handle.net/11356/1684.

References

Arhar Holdt, Špela; Fišer, Darja; Erjavec, Tomaž in Krek, Simon. Syntactic annotation of Slovene CMC: first steps. Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities, 27.–28. september 2016, Ljubljana, Slovenia, 2016, str. 3–6. http://nl.ijs.si/janes/cmc-corpora2016/proceedings/.

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

Erjavec, Tomaž. 2012. MULTEXT-East: morphosyntactic resources for Central and Eastern European languages. Language Resources and Evaluation, 46(1): 131–142.

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.

Toporišič, Jože. (2004): Slovenska slovnica. Maribor: Obzorja.

Revision #7
Created 2023-11-13 14:17:06 UTC by Tina Munda
Updated 2023-11-15 11:23:43 UTC by Tina Munda