

TS Belgrade

Exercises – cascades for NER

Preparatory steps

1. Make sure you have installed the latest version of Unitex.
2. Install all language packages, English is mandatory.
3. Please chose the private user directory (all languages will be there). Recommendation is to use the folder „*TS_Belgrade*“. Unitex will ask you to set private user directory during the first launch. Subfolders of all languages will be there.
4. Check if the subfolder *English* contains files *Alphabet.txt* and *Alphabet_sort.txt*.
5. Subfolder *Dela* should contains the dictionary *dela-en-public(.bin/.inf)*. Please add two additional dictionaries *CasENAmbiguites-. (dic/inf)* and *CasEnIstexDico.(dic/inf)* from the *TS_Belgrade.rar* package from the folder with the same name (*Dela*).
6. Please copy the text *ENG18900_Doyle_noHeader_xml.txt* from the *Corpus* subfolder from *TS_Belgrade.rar* package into *Corpus* subfolder of your private user directory. It is the text from the ELTeC collection in English with header ommited.
7. Please copy files *TS_analysis.csc* and *TS_synthesis.csc* from subfolder *Cassys* from *TS_Belgrade.rar* package in the *Cassys* subfolder of your user directory.
8. Please copy all graphs from *TS_Belgrade.rar\TS_Belgrade\Graphs\repository* into the same subfolder (*Graphs\repository\TS_Belgrade*) of your user directory.