LINDAT DSpace at CLARIN.SI

Cyprian Laskowski

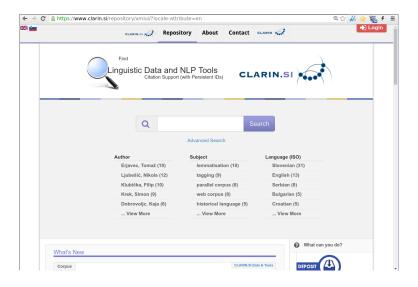
Jožef Stefan Institute

9 November 2016

Overview

- CLARIN.SI is small
- The local LINDAT/DSpace repository is the core of CLARIN.SI
- Now a CLARIN B centre
- Close to 50 submissions, all with substantial data, mostly in HLT
- Mostly from Slovenia and for Slovene
- No tools yet
- Example submissions: SNABI (speech database), Emoji Sentiment (lexicon), SlovParl (transcriptions)

Screenshot



Adaptations

- Local LINDAT DSpace installation in 2014
- Local branding (based on common-theme)
- Two-year "test" phase (but without "fake" submissions)
- Periodic updates/upgrades
- Local overlay management
- Test server
- Localisation (Slovene translation)
- Hiding of unused features (e.g., Google Analytics)

Features valued

- AAI
- OAI-PMH harvesting
- Overlays
- Common theme
- Shared submission
- Refbox
- Piwik
- Etc.

Features hoped for

- Simpler localisation
- Improved bibliography exports
- Integration with improved Kontext
- More harvesters (e.g, Thomson Reuters)
- Better statistics summaries
- Better admin documentation

Future of CLARIN.SI

- Stabilisation (minor local bug fixes, etc.)
- Better Kontext installation
- Improved monitoring
- Statistics analysis
- More complete localisation (and for DSpace?)
- Adding of CLARIN and custom licenses
- Move from sys to user support (documentation, webcasts, etc.)
- Providing web services and tools