

New features in SkE

Vojtěch Kovář, Vít Suchomel



`{vojtech.kovar, vit.suchomel}@sketchengine.co.uk`

5th Sketch Engine Workshop
Bolzano, July 14, 2014

New features in Bonito

- Filtering candidate definitions (experimental)
- Corpus info page
- “Auto” part of speech in word sketch and thesaurus
- Show only 1 hit per document in concordance
- Filter out overlaps in concordance
- Reciprocal bilingual manual search
- Shortening long references
- * and ? wildcards in simple query
- Printer-friendly layout + icon for printing
- Short URL service <http://ske.li>

New features in Corpus Architect

- Featured corpora
- Corpus list filtering
- Local admin user space quota allocation
- User-friendly “Corpus should be recompiled” message
- Automatic conversion of user corpora to a new tagset
- Convert packed documents to a single plaintext
- Term extraction for user corpora

New processors of user data installed

- Arabic Stanford segmenter and POS tagger
- Estonian TreeTagger
- Finnish TreeTagger
- Hungarian Hunpos
- Polish WCRFT
- Swedish Hunpos