Wissenschaftliche Nutzung der korpuslinguistischen Infrastruktur LAUDATIO

Carolin Odebrecht, Dennis Zielke, Thomas Krause, Benjamin Weißenfels, Malte Belz, Tino Schernikau, Vivian Voigt
Humboldt-Universität zu Berlin
www.laudatio-repository.org

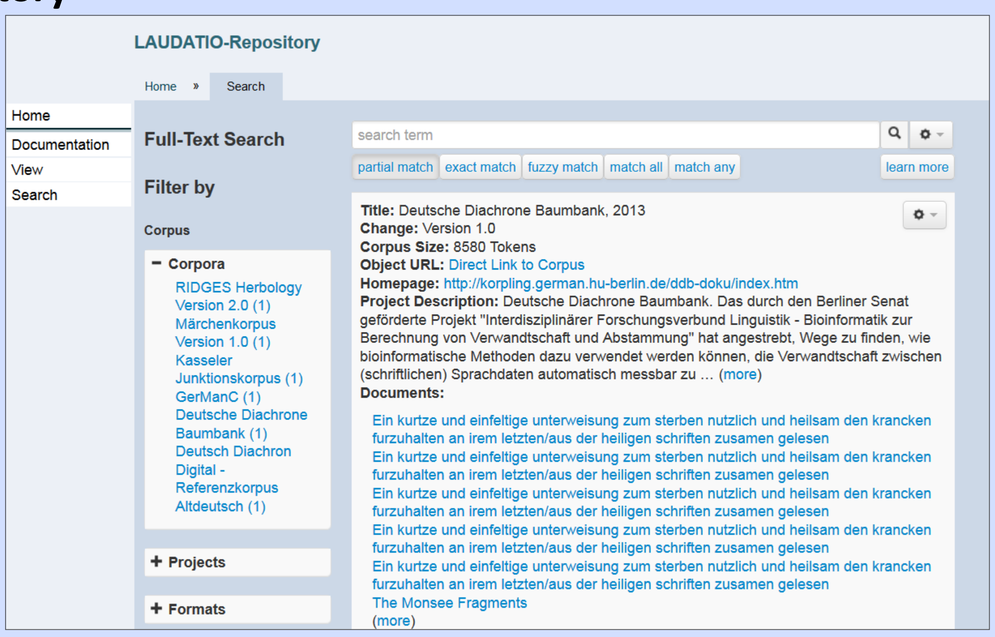
36. Jahrestagung der DGfS 2014 Marburg

Wie werden linguistische Forschungsdaten für eine kollaborative Nutzung aufbereitet?

Nutzung von	durch
Forschungsdaten	Finden, Durchsuchen, Herunterladen, Wiederverwenden, Hochladen
Software Modelle Schnittstellen	Anwenden, Einrichten, Anpassen

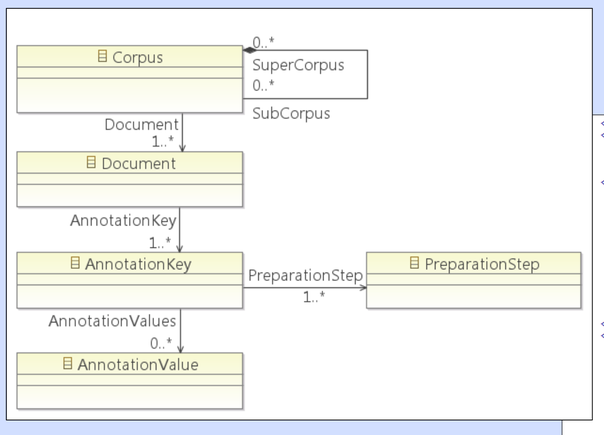
Open Access[7] LAUDATIO-Infrastruktur

Facetten- und Freitextsuche von historischen Korpora im LAUDATIO-Repository[9]

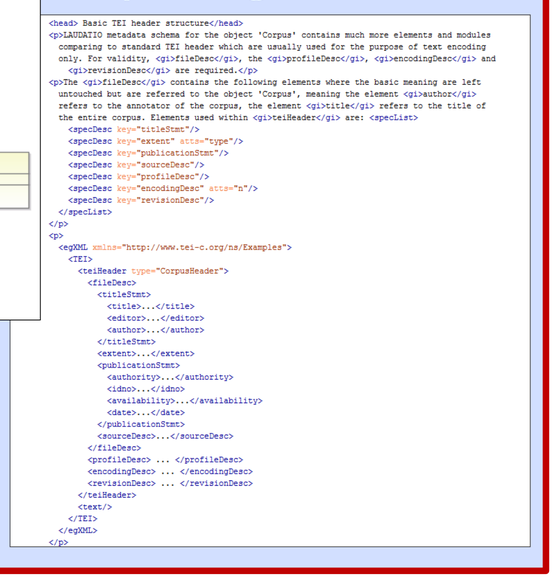


Strukturierte und einheitliche Anzeige im LAUDATIO-Repository

Metamodell von Forschungsdaten[3,4]



Metadaten in TEI XML Anpassung mit ODD[1,2]



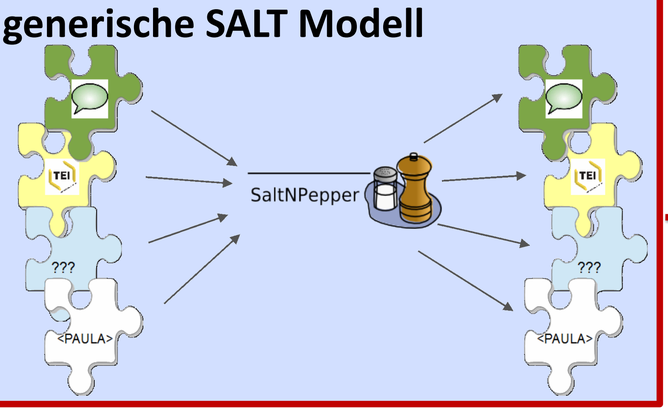
TEI < Text Encoding Initiative >

Einheitliches Metadatenformat für das Repository

Analyse und Visualisierung der Korpusdaten aus dem Repository in ANNIS[5]

Konvertierung verschiedener Formate durch SaltNPepper[6]

Import, Export und Manipulation über das generische SALT Modell



Aufbereitung der Daten für ANNIS durch SaltNPepper

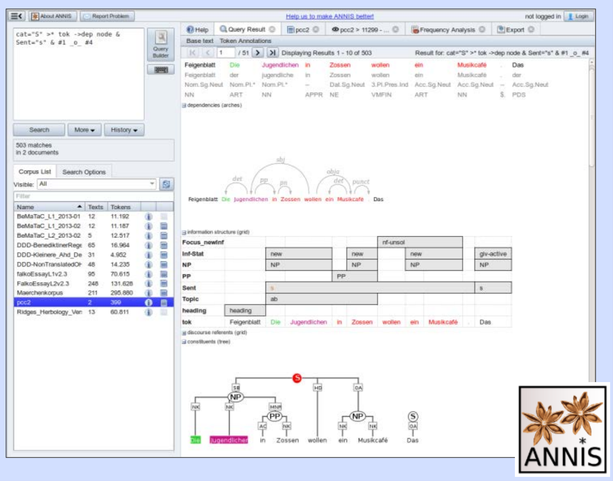
Unterstützung[8] von:

Mehrebenen-Korpora

Multiple Segmentierung

Korpusspezifische Visualisierungen

Multimodale Korpora



Ausblick

- Erweiterung auf historische Textkorpora aus verschiedenen Geisteswissenschaften
- Einbindung von SaltNPepper in der Repository Weboberfläche
- Vernetzung der Inhalte mit anderen Datenbanken (Linked Open Data und Interoperabilität mit anderen Metadatenformaten)

Referenzen [1] Burnard, L., Bauman, S. (Hg.) (2008) *TEI P5: Guidelines for Electronic Text Encoding and Interchange. Oxford.* [2] Burnard, L., Rahtz, S. (2004) *RelaxNG with Son of ODD. Extreme Markup Languages.* [3] Odebrecht, C. (erscheint) Modeling Linguistic Research Data for a Repository for Historical Corpora. *Digital Humanities 2014 Conference.* 8.7.-12.7.2014, Lausanne. [4] Odebrecht, C., Krause, T. (2013) Metadata in an Infrastructure for Historical Corpora. *SFB 732 Incremental Specification in Context. Kolloquium.* 20.6.2013, Stuttgart. [5] Zeldes, A., Ritz, J., Lüdeling, A., Chiarcos, C. (2009) ANNIS: A search tool for multi-layer annotated corpora. In *Proceedings of Corpus Linguistics 2009.* Liverpool, UK. [6] Zipser, F., Romary, L. (2010) A model oriented approach to the mapping of annotation formats using standards. In *Workshop on Language Resource and Language Technology Standards, LREC 2010.* La Valette, Malta. | [7] Alle Forschungsdaten verfügbar unter Creative Commons-Lizenzen https://creativecommons.org/licenses/ [8] Dokumentation ANNIS http://www.sfb632.uni-potsdam.de/annis/documentation.html [9] Technische Dokumentation des LAUDATIO Repository http://rtd.cms.hu-berlin.de/docs/laudatio-repository-documentation/