

**Computerlinguistisches Arbeiten: Protokoll zum Repetitorium der Computerlinguistik.  
Computerlinguistik in Anwendungen der BSB-IT**

Die Bayerische Staatsbibliothek ist eine Landesbibliothek des Freistaates Bayern. Im Bestand sind über 10 Millionen Bände. Die Entwicklungsabteilung des BSB arbeitet mit Metadaten, Digitalen Objekten und digitaler Außenwelt. Das Ziel ist, drei Aufgaben für den Netzbenuer zu erledigen: Finden, Auswählen, Zugang zu dem Buch erhalten. Die folgenden allgemeinen Anwendungen wurden präsentiert: Bibliothekskatalog (OPACplus), Fachportale, Spezialkataloge, OpenData des Bayerischen Digitalen Repositorium, LinkedOpenData. Datenprocessing enthält drei Aufgaben: Daten abholen, aufbereiten, weitergeben.

Daten-Aufbereiten enthält Umwandlung, Anreicherung, Data Mining (NER), Normalisierung.

Als nächstes wurde eine Architektur des Suchprojekt Apache Solr präsentiert.

Zum Schluss wurde die Zusammenfassung der computerlinguistischen Anwendungen erwähnt: Information Retrieval mit Solr, NER mit Stanford NER.

Als zukünftige Aufgaben sind vorgesehen:

- Optical Music Recognition
- Automatische inhaltliche Erschließung
- Disambiguierung und Verlinkung erkannter Entitäten
- Verlinkung zusammenhängender Titeldaten