



Die Slavistik in einer nachhaltigen digitalen Forschungsinfrastruktur

Roland Meyer (Deutscher Slavistenverband)
Institut für Slawistik, Humboldt-Universität zu Berlin
roland.meyer@hu-berlin.de

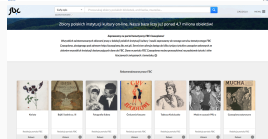
Wissenschaftsgeleitete Forschungsinfrastrukturen, Berlin, 15.02.2018

Digitale Forschungsdaten in der Slavistik

- **Textkorpora** (tiefenannotiert, multimedial, historisch, parallel, Feldforschung)



- **Digitale Bibliotheken** mit Datenmassiven in Bildform, selten Volltext



- **Webportale** zur (multimedialen) Präsentation von Einzelthemen

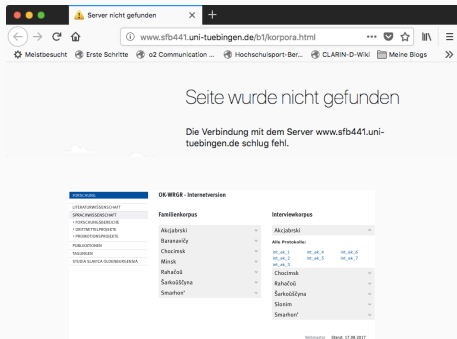


- (vereinzelt) **Software**, z.B. Parameterdateien für Tagger, R-Skripts
- **Forschungsprimärdaten** aus empirischen Studien – z.B. Befragungsdaten, psycholinguistische Daten, ...
- **Online-Publikationen**

Warum braucht die Slavistik eine digitale Infrastruktur?

- **Nachhaltigkeit** der Speicherung und des Zugangs zu Forschungsdaten

- **Vereinheitlichung** des Zugangs zu Forschungsdaten



- **Übergreifende Katalogisierung**, auch für Online-Ressourcen
- **Praktikable Standardlösungen** für datenintensive, dezentrale Projekte
- **Internationaler Anschluss** an Leuchtturmprojekte in Mittel-/Osteuropa (z.B. Nationalkorpora)

Nutzen für Lehre und Weiterbildung:

- **Einfacherer Zugang** zu Daten für studentische und interne wie externe Nutzer (z.B. Lehrkräfte)
- Bessere **Dokumentation**, technische **Synergien**
- Mehr Konzentration auf das fachlich Wesentliche: Anwendung, Auswertung, Analyse

Nutzen für Projektbegutachtung und Antragstellung:

- **Standardprozeduren**, weniger Trickkisten und Hacks, mehr Chancengleichheit
- Vereinfachung des Technikkapitels in Anträgen, Absicherung durch technologisches **Expertenwissen**
- zugleich **Freiraum** für **dezentrale Innovationen**, Vermeidung von Metadaten-Lastigkeit und Überregulierung

Korpora:

- UTF-8, historische Fonts, Mehrebenenannotation, Bild-/Audio-/Videodaten, Parallelkorpora, Metadaten
- Bedienfreundlichkeit für Administratoren *und* Nutzer, konfigurierbare Abfrageschnittstellen, APIs
- Annotationssoftware: **große, strukturierte Tagsets**; Lexikonmanagement; Transkriptionsunterstützung

Digitale Bibliotheken und Repositorien:

- Einbindungsmöglichkeit in eigene Webseiten, APIs
- **Mehrsprachige** Zugänge und **sprachübergreifende Ontologien**
- Rechtemanagement, Zugangsbeschränkungen, Zitierbarkeit
- Katalogisierung von Software und Forschungsprimärdaten
- Kompatibilität mit **Standards in Osteuropa** (z.B. Clarin-PL/-CZ)
- praktikable Austauschformate