

Computational Literary Studies: Data Landscape Review and Inventory

Ingo Börner (ingo.boerner@uni-potsdam.de) / Vera Maria Charvat (VeraMaria.Charvat@oeaw.ac.at) / Matej Ďurčo (Matej.Durco@oeaw.ac.at) /
Michał Mrugalski (michal.mrugalski@hu-berlin.de) / Carolin Odebrecht (carolin.odebrecht@hu-berlin.de)

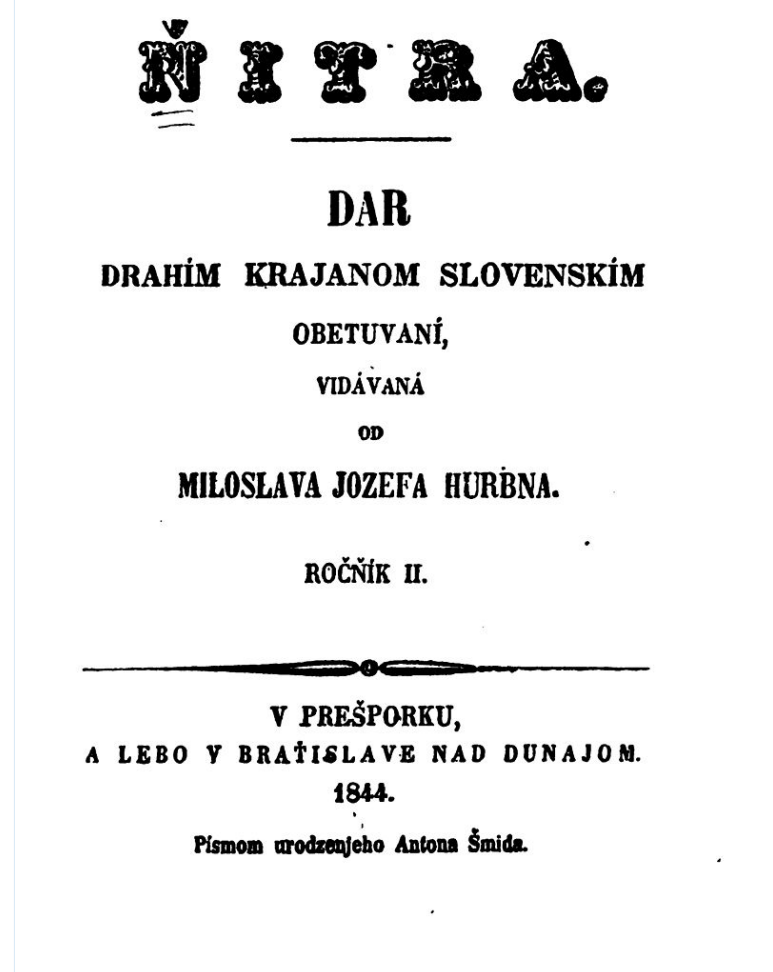
DATEN- Landschaft heute

Embarras

DE RICHESSES



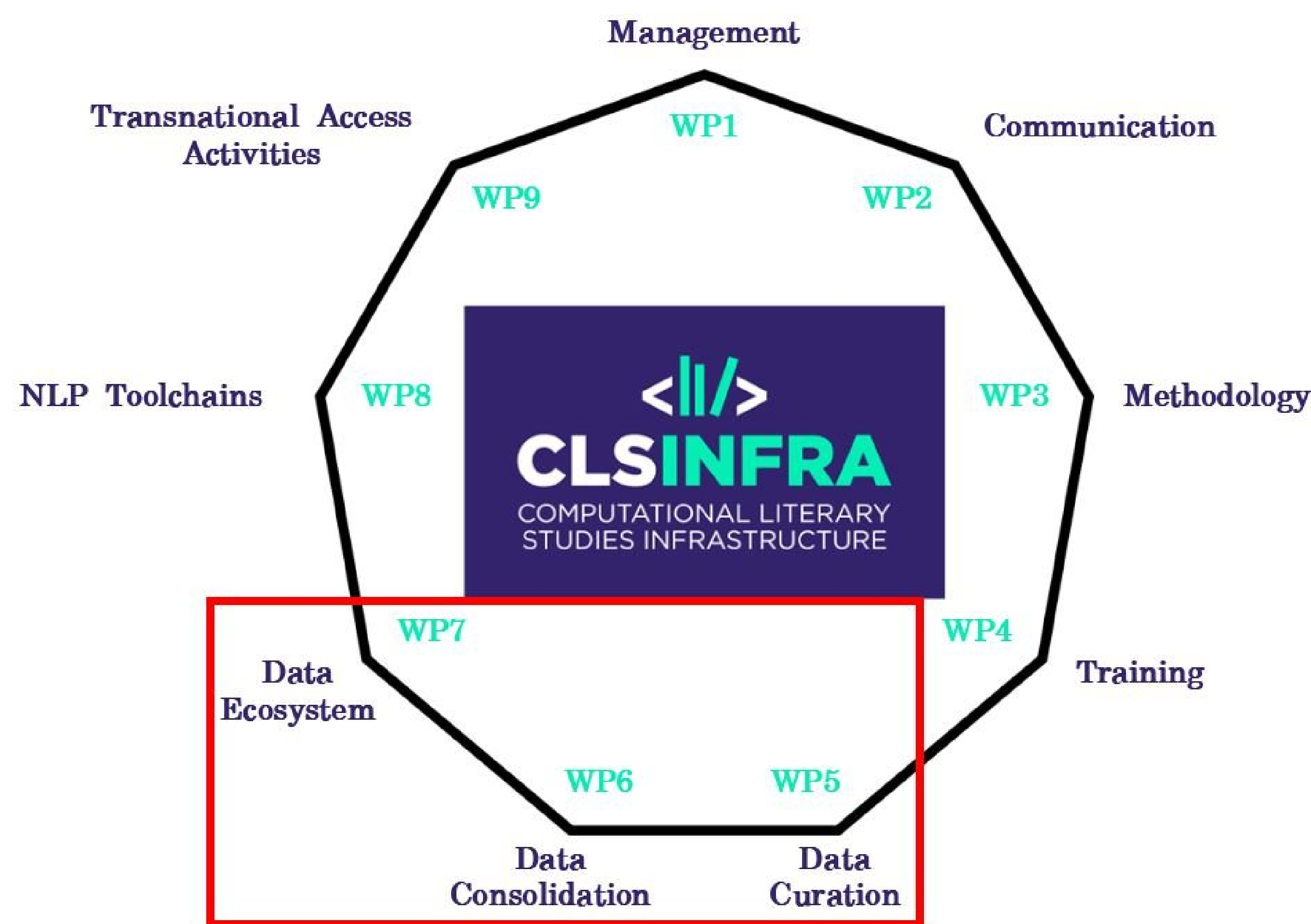
인간은 태어날 때부터
우며 그 존엄과 권리에
동등하다. 인간은 천부적으로
양심을 부여받았으며 서로
의 정신으로 행동하여야



Zwei Usecases
→ slowakischer Roman : Fragmentarizität von Daten
Beliebigkeit von Metadaten : Haiku ←

IASC-Projekt CLS INFRA

Einheitlicher, einfacher, europaweiter Zugang zu Daten und Tools für die CLS-Community.



Fokus auf "Data Curation and Consolidation" mit Perspektive auf das Ökosystem der "Programmable Corpora"
Rekonfiguration von Zugangsparadigmen für literarische Daten nach den FAIR-Prinzipien (findable, accessible, interoperable, reusable)

D5.1 Review report documenting the state of literary data

Eher konzeptionelle als technische Zugangsbedingungen zu Textdaten

Leitfragen

Welche Beschreibungsmerkmale sind für die Daten als Kollektion im Sinne einer eigenen epistemischen Einheit wesentlich? in Bezug auf die literarischen Vorlagen und deren Aufbereitungen relevant?

Typologie

literarischer Textsammlungen entsprechend

- Aufbau, Textauswahl, Zweck
- Formate, Beschreibungen, Auszeichnungen
- Zugriffsmöglichkeiten

D6.1 Inventory of existing data sources and formats supported/required

Online-Katalog literarischer Korpora

Modell

- Basierend auf Metamodell für Korpusmetadaten (MKM)
- Basisentitäten Korpus, Dokument und Annotation
- Kompatibel mit CIDOC CRM, FRBR
- Verweise zu externen Referenzressourcen (GND, WikiData, GeoNames, DDC, ...)
- Dissemination über Aggregatoren wie CLARIN VLO, Europeana oder OpenAIRE
- Entsprechende Mappings zu Dublin Core, CMDI, EDM

Technologie

ResearchSpace + TripleStore

