DARIAH-CZ
Digital Research Infrastructure for the Arts and Humanities
Czech node of the DARIAH ERIC - Digital Research Infrastructure for the Arts and Humanities

Jan Hajič
Institute of Formal and Applied Linguistics
Computer Science School
Faculty of Mathematics and Physics
Charles University, Prague
Czech Republic
DARIAH-CZ

• “Digital humanities”
  – Enable, support digitally-enabled research

• Users
  – Researchers in Arts and Humanities

• Technology
  – Repository (resources), Services, Applications

• Current status
  – Successfully passed Ris’ “Evaluation 2017”
  – Proposal to join LINDAT/CLARIN: LINDAT/CLARIAH-CZ
Main Goals for 2019-2022

• Becoming stable part of DARIAH ERIC
  – Membership of Czech Republic
  – Integrated with LINDAT/CLARIN by Jan. 1, 2020
  – Operational and fully connected by Jan. 1, 2021

• Construction phase (2019-2020), following preparations (2017-8)
  – Collect data, assess quality, completeness, licencing, determine main repository(ies)
  – Prioritize, convert data / metadata

• Operational phase (2021-)
  – Repositories running and connected to DARIAH central hub
  – Services
    • integrated existing services
    • new services: search in multimodal data, statistics, visualization
Timeline of DARIAH-CZ w/relation to LINDAT/CLARIN

- LINDAT/CLARIN est. 2010, operational since Jan. 1, 2014
- CLARIN ERIC est. 2012 (Utrecht, NL)
- DARIAH ERIC est. 2014 (Paris, FR)
Institutions

• Charles University, Prague
  – Faculty of Mathematics and Physics (CS School)
  – Faculty of Arts (multiple departments)
• Univ. of West Bohemia, Pilsen (Applied Sciences)
• Masaryk University, Brno
  – Faculty of Informatics, Faculty of Arts (multiple departments)
• Academy of Sciences of the Czech Republic
  – Library of the Academy of Sciences
  – Institute of the Czech Language
  – Institute of Philosophy
  – Institute of History
• National Library, Prague
• Moravian Library, Brno
• National Film Archive, Prague
• National Gallery, Prague

24.4.2018
Research Fields

- Linguistics, computational linguistics
- History, historical bibliography
- Study of culture
- History of art
- Philosophy
- Film and film history including new audiovisual media
- Visual arts
- Musicology and history of music
- Ethnology
- Archaeology/egyptology
- ... and interdisciplinary areas of the above
The Workplan

• Current situation
  – Existing collections – at all institutions
  – Varying degree of
    • Digitization, metadata description, access and availability

• Construction phase
  – “inventory” (18 (sub)projects identified)
    • E.g. MEMORI (IP CAS, CAS Library), Archive Jan Patocka, …
  – Metadata unification (DARIAH compatible)
  – Technology development for access
Workplan scheme

User applications – search, Statistics, visualization, ...
Thank you!