



# Institute of Formal and Applied Linguistics (ÚFAL) Projects

Jan Hajič



# Projects in 2019



- LINDAT
  - Research Infrastructure (JH, MSMT), not research in itself
  - LINDAT/CLARIN: 2016-2019 plus OP VVV, P. Stranak, J. Misutka
  - LINDAT/CLARIAH-CZ: 2019-2022, OP VVV II submitted, B. Vidová
    - Combines membership in EU CLARIN and DARIAH networks
      - DARIAH: Digital humanities and arts
    - 11 institutions in the Czech combined network
    - UK, MU, Academy of Sciences, National and Moravian Libraries, NG, NFA
  - <http://lindat.cz> (new web: temporarily at <https://clariah.lindat.cz/>)
  - New people
    - Maarten Janssen [janssen@ufal.mff.cuni.cz](mailto:janssen@ufal.mff.cuni.cz) (Spain), TEITOK adaptation
    - ?Wolfgang Sauer (Germany) – repository development
  - Universal Dependencies project: under LINDAT
  - Center for Visual History Malach: under LINDAT (& Herzl Center)



# Projects in 2019



- ELG European Language Grid, EU H2020 Call 29a
  - „Marketplace“ for mainly commercial use of Language Technologies
  - Partners: **DFKI**, UK, Sheffield, ELRA, commercial partners
  - Pilot projects: organized by us (2M EUR), L. Kačena, J. Hamrlová
  - Some UFAL tools provided – UDPipe, maybe MT (M. Straka)
    - UDPipe integrated in the first round, shown at META-FORUM in Brussels (Oct. 2019)
- SSHOC, Infrastructural project
  - Part of CLARIN ERIC participation
  - UFAL/LINDAT provides basic language tools for IR, MT for Social Surveys
    - D. Variš, J. Arnold
- Both projects 2019-2021 (SSHOC until mid-2022)



# Continuing Projects



- VIADAT
  - Virtual assistant for Oral History collections
  - NFA, Inst. for Contemporary History (Academy of Sciences)
  - 0.75 FTE free for finishing the software
- Oral History Archives for 20th century history
  - Coordinated by Pilsen university (UWB)
  - Cooperation with ÚSTR (Institute for the study of totalitarian regimes)



# Shared Tasks

- CoNLL Shared Task on dependency parsing
  - Run in 2017 and 2018 (D. Zeman et al.)
    - 2018: M. Straka wins EPE (extrinsic evaluation) with UDPipe 2
- CoNLL Shared Task on Meaning Representation Parsing (MRP)
  - Run in 2019 (maybe again 2020, Cz added)
  - Stephan Oepen & colleagues



# To be submitted (H2020)



- Call 48/49, Deadline Nov 13, 2019
- HumaineAI
  - Coordinated by DFKI (Paul Lukowicz)
  - NLP small part (maybe with VUT Brno)
- Security, NLP, Speech, ... (Ondrej Bojar)
  - Czech Technical University (M. Pechoucek)
- AI4EU
  - Unclear participation
- Human Language Project
  - Not submitted (?)
- FET Proactive
  - Interaction – Ondrej Dusek