Martin Popel

- **Treex** NLP framework (+ Treex::Web by Michal Sedlák)
  - more language-independent code for tecto-analysis (and synthesis) exploiting Interset (by Dan Zeman)
- **TectoMT** machine translation (PhD thesis on transfer)
  - experiments with Vowpal Wabbit and structured prediction
- **HamleDT** 30+ treebanks (ACL14: stanfordization, LRE paper)
- **PBML** (next deadline: January 15th 2015) LaTeX volunteer?
- **Technical reports** (2014 deadline: December 1st)
- **Teaching**
  - autumn: Modern Methods in CL I (“Reading group”)
  - spring: Modern Methods in CL II (for PhD and staff, Deep NN ?)

Language Data Resources (with ZŽ, me: significance)
Český jazyk neboli čeština je západoslovanský jazyk, nejvíce příbuzný se slovenštinou, poté polštinou a lužickou srbskou.
8 partners: (FCUL, DFKI, CUNI, IICT-BAS, UBER, UPV/EHU, UG, HF)

Build MT system(s) for 14 language pairs:
English ↔ Basque, Bulgarian, Czech, Dutch, German, Portuguese, Spanish

Extrinsic evaluation: IT-related question answering
(user questions translated to English → IR → answers translated back)

TectoMT considered for MT Pilot 1 for:
Czech, Dutch, Portuguese, Spanish

Depfix considered for: Basque (and Czech via Chimera)

ÚFAL leads WP2 ”Deep MT” (55 PMs)

14 PMs also on WP5 ”Lexical Semantics: Linking and resolving” (Linked Open Data)