

Mgr. Rudolf Rosa, Ph.D.

+420 776 577 617

rosa@ufal.mff.cuni.cz

<http://ufal.mff.cuni.cz/rudolf-rosa>

Vybrané výzkumné projekty

- 2019–nyní **THEaiTRE**, automatické generování scénářů divadelních her, hlavní řešitel.
- 2018–2020 **LSD**, analýza lingvistických struktur v hlubokých neuronových sítích.
- 2013–nyní **HamleDT/Universal Dependencies**, mnohojazyčné závislostní korpusy.
- 2015–2019 **SloNLP**, Slovenskočeský NLP Workshop, hlavní organizátor a předseda PC.
- 2014–2018 **croSSSynt**, modelování závislostní syntaxe napříč jazyky [6].
- 2013–2016 **TectoMT**, hloubkově syntaktický strojový překlad.
- 2011–2015 **Depfix**, automatická post-editace výstupů statistického strojového překladu [4].

Výuka

- 2014–nyní **Technologie pro NLP**, NPFL092, Bash, Git, Python, XML, JSON.
- 2017–nyní **NLP na výpočetním clusteru**, NPFL118, použití naší sítě a SGE.
- 2018–nyní **Mnohojazyčné NLP**, NPFL120, mezijazyčný tagging a parsing, překlad.
- 2019–nyní **Programování I.**, NPRG030, úvod do programování, Python.
- 2018–nyní **Vedení studentů**, 10× bakalářský, 1× magisterský, 1× pre-PhD stáž.

Ocenění

- 2018 **GA UK kulatý stůl**, nejlepší projekty grantové agentury, projekt croSSSynt.
- 3× **ÚFAL best paper award**, nejlepší články ústavu, 2013 [5], 2014 [7], 2017 [8].

Práce

- 2011–nyní **akademický vědecký pracovník (VP2A)**, Univerzita Karlova, Česko.
Matematicko-fyzikální fakulta, Ústav formální a aplikované lingvistiky
- 2015–2016 **softwarový inženýr (šestiměsíční stáž)**, Google, Curych, Švýcarsko.
automatické určování anafory pro hlasové vyhledávání
- 2004–2014 **vývojář; učitel angličtiny**, Jazyková škola Rosa, Kladno, Česko.
vývoj administračního systému a webu; výuka dětí a dospělých

Vzdělání

Univerzita Karlova, Praha, Česko.

Matematicko-fyzikální fakulta, Ústav formální a aplikované lingvistiky

- 2013–2018 **doktorské studium**, matematická lingvistika.
Odhalování struktury vět přirozeného jazyka pomocí částečně řízených metod [3]
- 2007–2013 **bakalářské a magisterské studium**, informatika, matematická lingvistika.
Automatická post-editace výstupů frázového strojového překladu [2]

Programátorské dovednosti

plynně Python, Bash, Subversion, vim, L^AT_EX
dobře C++, Java, Perl, Linux, Git, SGE

Jazykové dovednosti a osvědčení

úroveň C2 angličtina (CPE)
úroveň B1 francouzština (DELF), němčina
úroveň A2 španělština, italština
úroveň A1 český znakový jazyk

Vybrané publikace

shrnutí **40 recenzovaných publikací**, 756 citací, h-index 15 (podle Google Scholar).

- [1] David Mareček a Rudolf Rosa. „From Balustrades to Pierre Vinken: Looking for Syntax in Transformer Self-Attentions“. In: *Proc of BlackboxNLP*. Stroudsburg, PA, USA: ACL, 2019, s. 263–275.
- [2] Rudolf Rosa. „Automatic Post-editing of Phrase-based Machine Translation Outputs“. Diplomová práce. Praha, Česko: Univerzita Karlova, Matematicko-fyzikální fakulta, 2013.
- [3] Rudolf Rosa. „Discovering the Structure of Natural Language Sentences by Semi-supervised Methods“. Dis. Praha, Czechia: Charles University, Faculty of Mathematics a Physics, 2018.
- [4] Rudolf Rosa, David Mareček a Ondřej Dušek. „DEPFI_X: A System for Automatic Correction of Czech MT Outputs“. In: *Proc of WMT*. Montreal, Kanada: ACL, 2012, s. 362–368.
- [5] Rudolf Rosa, David Mareček a Aleš Tamchyna. „Deepfix: Statistical Post-editing of Statistical Machine Translation Using Deep Syntactic Analysis“. In: *Proc of ACL Student Research Workshop*. Sofie, Bulharsko: ACL, 2013, s. 172–179.
- [6] Rudolf Rosa a Zdeněk Žabokrtský. „ KL_{cpo}^3 – a Language Similarity Measure for Delexicalized Parser Transfer“. In: *Proc of ACL-IJCNLP, Short Papers*. Stroudsburg, PA, USA: ACL, 2015, s. 243–249.
- [7] Rudolf Rosa et al. „HamleDT 2.0: Thirty Dependency Treebanks Stanfordized“. In: *Proc of LREC*. Ed. Nicoletta Calzolari et al. Reykjavík, Island: ELRA, 2014, s. 2334–2341.
- [8] Rudolf Rosa et al. „Slavic Forest, Norwegian Wood“. In: *Proc of VarDial4*. Ed. Preslav Nakov et al. Stroudsburg, PA, USA: ACL, 2017, s. 210–219.