

**Univerzita Karlova v Praze**  
**Matematicko-fyzikální fakulta**  
**ZÁZNAM O PRŮBĚHU OBHAJOBY**  
**DISERTAČNÍ PRÁCE**

**Název práce:** Automatické propojování lexikografických zdrojů a korpusových dat

**Jazyk práce:** čeština

**Jméno studentky:** RNDr. Eduard Bejček

**Studijní program:** Informatika

**Studijní obor:** 4I3 Matematická lingvistika

**Školitel:** doc. RNDr. Markéta Lopatková, Ph.D. (ÚFAL MFF UK)

**Oponenti:** RNDr. Aleš Horák, Ph.D. (FI MU Brno)  
doc. Ing. Zdeněk Žabokrtský, Ph.D. (předseda, ÚFAL MFF UK)

**Členové komise:** prof. RNDr. Jan Hajič, Dr. (předseda, ÚFAL MFF UK)  
prof. PhDr. Jarmila Panevová, DrSc. (místopředsedkyně, ÚFAL MFF UK)  
prof. PhDr. Eva Hajičová, DrSc. (ÚFAL MFF UK)  
doc. Ing. Zdeněk Žabokrtský, Ph.D. (předseda, ÚFAL MFF UK)  
Mgr. Barbora Vidová Hladká, Ph.D. (ÚFAL MFF UK)  
Mgr. Jan Štěpánek, Ph.D. (Barclays Capital Services Limited)  
Ing. Alexandr Rosen, PhD. (FF UK)

**Datum obhajoby:** 22. září 2015

**Průběh obhajoby:** Obhajoba proběhla v úterý 22. 9. 2015 od 13:00 hodin v budově MFF UK v místnosti S1, Malostranské nám. 25, Praha 1.  
Přítomní členové: viz prezenční listina  
Hosté: viz prezenční listina

Celá obhajoba probíhala v češtině. Předsedající, prof. RNDr. Jan Hajič, Dr., přivítal přítomné, představil členy komise a konstatoval, že komise je usnášeníschopná. Předsedající dále stručně představil uchazeče a zkontroloval úplnost dokladů vyžadovaných k obhajobě. Poté konstatoval, že byly splněny všechny podmínky pro konání obhajoby a že nedošly žádné písemné připomínky. Přítomní dostali možnost nahlédnout do výtisků předkládané disertační práce, do seznamu publikovaných prací, citací a životopisu uchazeče.

Školitelka, doc. RNDr. Markéta Lopatková, Ph.D., vyjádřila své stanovisko k průběhu doktorského studia uchazeče a seznámila přítomné se svým hodnocením obhajované disertační práce.

Uchazeč dostal slovo a představil svou práci. Presentaci rozdělil na dvě základní části: vzájemné propojování valenčních slovníků a propojování slovníků s korpusy. V první části podrobněji představil automatickou heuristickou metodu na propojování slovníků

---

Pokyny pro předsedy nebo místopředsedy komisi:

Práce v elektronické podobě musí být studentem vložena do SIS. Formulář vyplňte ve všech bodech v elektronické podobě. V bodě Členové komise se uvedou všichni členové komise a za jejich jména se uvede „(přítomen)“ nebo „(nepřítomen)“. Předseda nebo místopředseda komise je jejím členem. V bodě Průběh obhajoby by měly být uvedeny alespoň čtyři věty vystihující průběh obhajoby. Po vyplnění formuláře ho vytiskněte, dole formulář ještě vlastnoručně podepište a přiložte k zápisu o státní závěrečné zkoušce. Současně vložte formulář v elektronické podobě (bez vlastnoručního podpisu) do SIS.

VALLEX a PDT-Vallex a výsledek tohoto propojení, který je k dispozici na webu, demonstroval na ukázce. Stručně zmínil i propojování různých verzí slovníku VALLEX a mezijazyčné propojování slovníku VALLEX a anglického slovníku FrameNet. V druhé části se věnoval propojování slovníku víceslovných lexií s Pražským závislostním korpusem a korpusem SYN2000, který je součástí Českého národního korpusu. Na závěr uchazeč shrnul výsledky a přínos své disertační práce.

Předsedající představil oponenty, RNDr. Aleše Horáka, Ph.D. a doc. Ing. Zdeňka Žabokrtského, Ph.D., a požádal je o přednesení posudků. Dr. Horák vyzdvihl komplexnost a precizní zpracování předložené práce. Přínosy viděl jak v kvalitním rozboru automatické metody na propojení lingvistických zdrojů, tak i v rozboru výstupů a evaluaci navržené metody. Jako odborník z prostředí mimo ÚFAL, pracoviště uchazeče, ocenil v práci podrobné představení vzniku a vývoje jednotlivých lingvistických zdrojů. Jako hlavní nedostatek zmínil absenci generalizace použitých metod, s čím souvisí rovněž opomenutí tzv. Linked Open Data, jakožto univerzální metodiky na propojování nejen lingvistických zdrojů. Představení posudku uzavřel doporučením práce k obhajobě a následného udělení titulu Ph.D.

Doc. Žabokrtský ve svém posudku zdůraznil, že zaměření práce je velmi aktuální. Uvítal technickou diskuzi nad hledáním společného formátu pro propojované slovníky, ocenil rovněž množství ilustrativních příkladů z korpusů a způsob navržení a vyhodnocení experimentů. Naopak ho překvapilo spojení dvou podle něj poměrně vzdálených témat: valence a víceslovných lexií. Rovněž by při řešení úloh uvítal aplikaci metod strojového učení. Nicméně svůj posudek uzavřel doporučením práce k obhajobě.

Uchazeč se detailně vyjádřil ke všem poznámkám a otázkám oponentů. Oba oponenti vyjádřili svou spokojenost s odpověďmi uchazeče.

Předsedající následně otevřel veřejnou rozpravu. Prof. Panevovou zajímalo, jestli se ještě vyvíjí slovník PDT-Vallex. Uchazeč zmínil korpus mluveného jazyka PDTSC a paralelní korpus PCEDT, při jejích tvorbě se slovník doplňoval. Školitelka doc. Lopatková zdůraznila, že automatická metoda kterou uchazeč navrhl je replikovatelná, a tudíž je možné pomoci jí propojit i takto rozšířené slovníky. Dr. Rosen se zeptal, proč se při propojování slovníku a korpusu uchazeč věnoval právě slovníku víceslovných jednotek. Rovněž ho zajímalo, jak těžké je propojit VALLEX a anglický slovník FrameNet. Na oba dotazy uchazeč odpověděl ke spokojenosti tazatele. Na heuristickou povahu navržené metody reagoval Dr. Bojar otázkou, jestli je žádoucí při propojování zdrojů mluvit o míře spolehlivosti jednotlivých propojení. Uchazeč odpověděl, že ve své práci tuto míru zohlednil a do slovníku zveřejněném na webu zahrnul jenom ta nejspolehlivější propojení.

Následovala neveřejná diskuse členů komise a tajné hlasování.

Komise schválila udělení titulu Ph.D.

**Počet publikací: 20**

**Výsledek hlasování:**

Počet členů s právem hlasovacím: 7

Počet přítomných členů: 7

Odevzdáno hlasů kladných: 7

Odevzdáno hlasů neplatných: 0

Odevzdáno hlasů záporných: 0

**Výsledek obhajoby:**  prospěl/a  neprospěl/a

**Předseda nebo místopředseda komise:**

---

Pokyny pro předsedy nebo místopředsedy komisi:

Práce v elektronické podobě musí být studentem vložena do SIS. Formulář vyplňte ve všech bodech v elektronické podobě. V bodě Členové komise se uvedou všichni členové komise a za jejich jména se uvede „(přítomen)“ nebo „(nepřítomen)“. Předseda nebo místopředseda komise je jejím členem. V bodě Průběh obhajoby by měly být uvedeny alespoň čtyři věty vystihující průběh obhajoby. Po vyplnění formuláře ho vytiskněte, dole formulář ještě vlastnoručně podepište a přiložte k zápisu o státní závěrečné zkoušce. Současně vložte formulář v elektronické podobě (bez vlastnoručního podpisu) do SIS.