

## Posudok habilitačnej práce

Téma: **Evalvácia strojového prekladu**

Autor: **RNDr. Daša Munková, PhD.**

### Oponent:

prof. RNDr. Peter Mikulecký, PhD.

Fakulta informatiky a managementu

Univerzita Hradec Králové

Česká republika

Tento oponentský posudok habilitačnej práce vznikol na základe poverenia dekana FF UKF v Nitre prof. PaedDr. Bernarda Garaja, CSc. zo dňa 23. 1. 2014. K dispozícii bol viazaný výtláčok habilitačnej práce RNDr. Daši Munkovej, PhD. „*Evalvácia strojového prekladu*“ v rozsahu 188 strán vrátane príloh.

### 1. Náročnosť spracovania riešenej problematiky práce

Strojový preklad predstavuje už celý rad rokov mimoriadne náročnú, o to viac však komerčne i používateľsky zaujímavú oblasť výskumu, na ktorú sa zameriava množstvo odborníkov pôsobiacich v napohľad rozdielnych výskumných oblastiach, ako je informatika, umelá inteligencia, matematická štatistika, Lingvistika, translológia, atď. Je preto treba vysoko oceniť odvahu habilitantky, že sa do tejto vysoko interdisciplinárnej (až transdisciplinárnej) oblasti výskumu pustila a dosiahla v nej pozoruhodné pôvodné výsledky. Konštatujem, že predložená práca je spracovaná na primerane vysokej odbornej úrovni, dokumentujúcej zrelosť a erudíciu habilitantky.

### 2. Posúdenie úrovne spracovania práce

Predložená habilitačná práca má 8 číslovaných kapitol, dve kapitoly nečíslované (Úvod a Diskusia a záver), rozsiahly zoznam použitých zdrojov a 8 číslovaných príloh. Okrem toho obsahuje zoznamy obrázkov, tabuliek a skratiek. Za úvodom, v ktorom vysvetľuje svoju motiváciu a okolnosti napísania práce, najmä definuje šesť cieľov, ktoré v práci hodlá splniť. Potom nasledujú tri kapitoly úvodné, (kap. 1 až 3), ktoré čitateľa uvádzajú do problematiky strojového prekladu, do problematiky štatistického strojového prekladu a do problematiky jeho evalvácie. Tieto kapitoly zaberajú cca tretinu celého textu a dajú sa považovať za veľmi dobrý a fundovaný prehľad súčasného stavu poznania v danej oblasti. K týmto kapitolám môžeme ďalej priradiť kapitoly 4 a 5, vysvetľujúce prístupy a metódy evalvácie strojového prekladu. Všetkých 5 prvých kapitol teda tvorí akúsi teoretickú časť habilitácie, veľmi dobrým a fundovaným spôsobom pripravujúcu pôdu pre prezentáciu vlastných vedeckých výsledkov, tvoriacich náplň kapitol 6 až 8.

Bolo prípadne možné uvažovať o väčšej hutnosti a stručnosti prvých 5 kapitol (prípadne prvých troch kapitol pomocou ich prípadného zlúčenia do jednej, úvodnej), lebo predsa len mám pocit, že tieto v podstate prípravné kapitoly, aj keď uznávam potrebnosť ich začlenenia, zaberajú veľmi podstatnú časť celej habilitačnej práce, možno i na úkor rozsiahlejšej prezentácie výsledkov výskumu. Napriek tomu vysoko hodnotím úroveň tejto teoretickej časti habilitácie, kde habilitantka úspešne zhrnula všetko, čo je v oblasti strojového prekladu zásadné a preukázala vynikajúcu orientáciu v problematike.

Vlastné výsledky sú, ako už bolo spomenuté, predmetom kapitol 6 až 8. Rozsahom ide o nejakých 36 strán, čo sa mi zdá byť pomerne málo. Výsledky tam prezentované si zaslúžili väčší priestor. Problematika evalvácie strojového prekladu a najmä problematika automatickej evalvácie patrí medzi zložité oblasti, ktoré nie sú až takým častým objektom výskumu. O to viac je treba oceniť habilitantku, že sa na túto oblasť zamerala. Zamerala sa na analýzu spoľahlivosti automatickej evalvácie strojového prekladu (kap. 6), na komparáciu strojového a posteditovaného prekladu mierami automatickej evalvácie v závislosti od cieľového jazyka (anglického či nemeckého, kap. 7) a na komparáciu strojového, posteditovaného a humánneho prekladu mierami automatickej evalvácie (kap. 8). Výsledky sú zaujímavé, vhodne vizualizované početnými grafmi. Chýba mi však vysvetlenie, prečo bol v kap. 6 na skúmanie zvolený text s 37 vetami (prečo nie iný počet? Alebo to nie je dôležité? Stačí 37 viet? Ak áno, prečo?) a podobne v kapitole 8 text pozostávajúci zo 14 viet (tie isté otázky, ako predtým). Okrem toho je v kap. 6 spomenuté, že na evalváciu strojového prekladu habilitantka použila vlastný nástroj, ale zdá sa, že chýbajú nejaké informácie, aký nástroj to bol. Prosím toto pri obhajobe spresniť. Výskumná vzorka posudzovateľov (kap. 6) pozostávala z 80 posudzovateľov, z toho 50 na jazyk anglický a 30 na nemčinu. Tento počet považujem za primeraný, dalo by sa však viac diskutovať o profesijnom či vzdelanostnom zložení tejto vzorky. Je uvedené, že ju tvorili študenti FF UKF i „native speakers“, ale chýba informácia o pomere študentov ku „native speakers“, a pod.

V práci mi ešte vadí rozpor medzi formuláciou cieľov práce na str. 15, kde je formulovaných 6 vedeckých cieľov, ktoré sa snaží habilitantka v práci dosiahnuť a naplniť, a konštatovaním na str. 151, že predkladaná práca sleduje dva hlavné ciele, ktoré potom v príslušnej záverečnej kapitole (str. 151 a nasl.) rozoberá. Totiž v úvode práce pri formulácii cieľov nie je reč o nijakých dvoch hlavných cieľoch, pričom u tých, ktoré sú na str. 15 formulované, mám miestami pochybnosti, či sa dajú považovať za ciele. Napr. formulácia „Analyzovať jazykovú komplexnosť hypotézy (...) v závislosti od referenčného prekladu“ v podstate nijaký cieľ nedefinuje, analyzovať sa dá čokoľvek akokoľvek. Dôležité je, čo sa tu chcelo dosiahnuť, objaviť, ako by táto analýza mala prispieť k naplneniu naozajstných cieľov práce. Tie dva ciele, ktoré sú rozoberané v záverečnej kapitole od str. 151 však v podstate vychádzajú z viac menej účelného zlúčenia cieľov zo str. 15 do dvoch skupín:

Prvý cieľ, formulovaný ako „interdisciplinárny výskum, ktorého výsledkom je analýza, návrh a implementácia algoritmov automatickej evalvácie strojového prekladu založenej na metrikách ako je presnosť (precision), pokrytie (recall), f-miera (f-measure), WER, PER, BLEU-1, BLEU-2, BLEU-3, BLEU-4, TER a CDER“ v sebe zahrňuje zrejme prvé tri pôvodne definované ciele. Druhý cieľ je potom formulovaný ako „evalvácia strojového prekladu, konkrétne evalvácia vygenerovaného prekladu on-line systémom štatistického strojového prekladu“ a očívidne viac – menej pokrýva ďalšie tri ciele zo strany 15. Ak zhodnotíme naplnenie takto preformulovaných cieľov, môžeme konštatovať, že v práci sú celkom iste na veľmi dobrej úrovni splnené a že habilitantka pri ich naplňaní dosiahla celý rad zaujímavých a pôvodných výsledkov.

Inak má habilitačná práca veľmi dobrú úroveň a dobre sa číta. Úroveň jej spracovania považujem za mimoriadne kvalitnú.

### **3. Prínos a úroveň výsledkov práce**

Výsledky, dosiahnuté v práci habilitantka zhrnula do samostatnej kapitoly bez čísla, nazvanej *Diskusia a záver* (od str. 151). Možno by bolo lepšie kapitolu explicitne nazvať nejakou ako

*Dosiahnuté výsledky* a pripojiť potom samostatný stručný záver, ale i takto sa čitateľ dostane k súhrnným výsledkom, spolu s istým kritickým pohľadom na výsledky zo strany habilitantky. Habilitantka tu stručne popisuje a analyzuje tri prípadové štúdie, ktoré použila pri evalvácií textu vygenerovaného systémom štatistického strojového prekladu.

Prvá štúdia bola zameraná na preklad populárno-náučného textu zo slovenského do anglického jazyka. Tu bol použitý text pozostávajúci z 37 viet vrátane nadpisu; celkove text obsahoval 658 slov. Otázky, ktoré by bolo vhodné v tejto súvislosti vyjasniť, som už uviedol v predchádzajúcej časti posudku. Habilitantka tu podrobne analyzuje chybovosť prekladu uvádza svoje názory na príčiny chýb u jednotlivých viet. S jej argumentáciou pravdepodobne možno súhlasiť, ale tento aspekt výsledkov práce nie je celkom v rozsahu mojich odborných kompetencií. Za veľmi hodnotné však v tejto časti považujem výsledky dosiahnuté pri vyhodnocovaní metrík automatickej evalvácie prekladu, za signifikantný výsledok považujem zistenie významných rozdielov v hodnotách metrík automatickej evalvácie PE a MT.

Druhá štúdia bola zameraná na preklad vzorky textu slovenského jazyka do nemeckého jazyka, na jeho posteditovanie, ďalej následnú komparáciu zistených rozdielov, a konečne komparáciu kvality anglického a nemeckého strojového prekladu a posteditovaného prekladu. Tu bol síce dosiahnutý rad zaujímavých zistení, ale samotné zaradenie komparácie kvality anglického a nemeckého prekladu do habilitačnej práce nie je vzhľadom na rozdiely v mohutnosti jednotlivých textových databáz úplne zdôvodnené. Z dosiahnutých výsledkov je pozoruhodné zistenie najvýznamnejších rozdielov medzi posteditovaným a strojovým prekladom prostredníctvom metrík automatickej evalvácie v závislosti od referenčného prekladu.

Tretia štúdia pozostávala z overenia efektívnosti používania systému štatistického strojového prekladu z hľadiska posteditovania oproti ľudskej prekladateľskej aktivite. Ide o zaujímavý experiment, ale počet študentov, ktorí sa v tomto prípade zúčastnili experimentu (7 + 7) sa mi však zdá byť nízky, prosím o vysvetlenie kritérií, podľa akých bol tento počet stanovený a konkrétni jednotlivci vyberaní.

Celkove sú výsledky zhrnuté v závere kapitoly *Diskusia a záver*. Asi nie je celkom patričné tento súhrn degradovať na niekoľko faktických poznámok, na druhej strane týchto 5 poznámok skutočne realisticky a v kondenzovanom tvare zhrnulo to najdôležitejšie, do čoho zaujímavý a podnetný výskum habilitantky vyústil.

Prínos a úroveň výsledkov, dosiahnutých v habilitačnej práci hodnotím pozitívne a výsledky považujem za vítaný a dôležitý príspevok k ďalšiemu rozvoju odboru.

#### **4. Posúdenie formálnej stránky práce**

Predložená habilitačná práca má pomerne veľký rozsah, 188 strán vysoko odborného textu nie je pre oponenta celkom ľahké dôkladne preštudovať. Práca je ale napísaná veľmi zrozumiteľne, pritom fundovane, zrelým odborným jazykom, štylisticky je veľmi kvalitná. Štruktúra práce je logická, možno sa dali niektoré kapitoly (napr. 1 a 3) zhrnúť do jednej úvodnej, ale to je vec prístupu ku spracovaniu problematiky. Ciele práce sú formulované v úvodnej kapitole, porozumeniu textu napomáha rad obrázkov (zhruba 68, ak sa nemýlim) veľmi vhodne ilustrujúcich úvahy habilitantky. Osem príloh práce vhodne dopĺňa hlavný text. Práca obsahuje 232 položiek zoznamu použitej literatúry, čo veľmi reprezentatívne dokumentuje autorkin rozhľad v skúmanej problematike. Z uvedeného zoznamu sú štyri práce v autorstve či spoluautorstve habilitantky, čo je samozrejme príjemné zistenie. Citácie použitých zdrojov sú v súlade s predpísanými normami a zvyklosťami.

**5. Explicitne vyjadrený návrh na udelenie vedecko-pedagogického titulu docent**

Vzhľadom na celkové posúdenie predloženej habilitačnej práce a jej vyššie uvedené prínosy, ako aj vzhľadom na to, že RNDr. Daša Munková, PhD. sa svojou publikačnou a vedecko-výskumnou činnosťou už dlhší čas radí k odborníkom v danej oblasti navrhujem, aby jej bol po úspešnom habilitačnom konaní v študijnom odbore **2.1.35 prekladateľstvo a tlmočníctvo** udelený vedecko-pedagogický titul **docent**.

V Hradci Králové, 14. 3. 2014



Peter Mikulecký