

## **Zápis**

z obhajoby disertační práce pana Mgr. Romana Chýly

konané dne: 27. 1. 2012

téma práce: „*Automatizované metody popisu struktury odborného textu a vztah některých prvků ke kvalitě textu*“

přítomní: Doc. PhDr. Rudolf Vlasák – předseda  
Prof. RNDr. Jan Rauch, CSc. – oponent  
Prof. PhDr. Oldřich Uličný, DrSc. – oponent  
Doc. RNDr. Jiří Ivánek, CSc.  
Doc. RNDr. Jiří Souček, DrSc.  
Prof. PhDr. Königová Marie, CSc.

omluveni: Doc. PhDr. Vladimír Smetáček – školitel  
PhDr. Eva Bratková, Ph.D.

Předseda komise Doc. PhDr. R. Vlasák zahájil obhajobu a představil přítomným kandidáta a jeho odborné zázemí.

Kandidát v prezentaci přítomné seznámil se svou disertační prací. Představil zejména hlavní cíle výzkumu a výsledky své práce. Cílem disertační práce bylo vytvořit nový softwarový nástroj pro univerzální sémantický kód (USK) – systém Séman, a ověřit jeho aplikovatelnost. Kandidát shrnul základní pojmy (univerzální sémantický kód, základní sémy, sémové ekvivalenty) a použité metody (morfologická analýza). Mezi hlavní výsledky práce patří srovnání vytvořeného systému Séman s extraktorem/systemem BibClassify (používá se pro klasifikaci dokumentů v CERNU), analýza problému správnosti překladu při automatické klasifikaci dokumentů (využita taxonomie Wordnet, shrnutí: změny v přesnosti se projeví ve zlepšení výsledků), vyhledávání kombinací sémů (dva syntetické testy, použit korpus umělého jazyka s 20 tis. dokumentů a korpus skutečného jazyka s 24 tis. dokumentů). Při evaluaci překladu se potvrdilo, že systém Séman je flexibilní a to i bez použití složitých algoritmů. Doplnil také, že systém je volně přístupný a jsou publikována použitá data.

Doc. Vlasák přečetl posudek omluveného školitele doc. Smetáčka. Vyzdvihuje samostatnost práce Mgr. Chýly, náročnost tématu, dosažení vytčených cílů práce. Systém lze použít pro sémantickou analýzu.

Oponent prof. Uličný předkládanou práci ocenil jako velmi kvalitní, přesahující požadavky a přínosnou oboru. Vyzdvihl interdisciplinaritu práce, která postihuje mnoho aspektů včetně lingvistických a představuje syntetizující, komplexní přístup k problematice. První otázka oponenta se týkala formalizovatelnosti některých používaných metod obsahové analýzy, druhá možnosti definice slova. Práci doporučuje k obhajobě.

Oponent prof. Rauch shrnul obsah práce a pět dílčích otázek v práci, na které se snažil autor odpovědět. K práci nemá podstatné komentáře a doporučuje ji obhajobě. Jeho připomínka se týkala toho, že by práci prospěl přehledný popis USK.

Kandidát na všechny otázky a připomínky oponentů odpověděl: Formalizovatelnost některých metod: existují nástroje pro kvalitativní obsahovou analýzu a teorie, které pomáhají

k objevování nových poznatků z obsahové analýzy textu. Jedná se o intelektuální práci, kterou lze těžko formalizovat, dalším problémem je neexistence jednotné teorie mysli, formalizování je celkem obtížná. Mgr. Chýla uvedl otázku do souvislosti se systémem Sémán a jeho lexikonem. Otázka definice slova: slovo vyjadřuje také postoj mluvčího; Sémán se neomezuje na slova, je omezen lexikální sémantikou, nedokáže se zpracovat vyšší struktura (věta, odstavec, text), souvisí s otázkou pochopení smyslu systémem. Kandidát se také vyjádřil k připomínce oponenta prof. Raucha o potřebě lepšího vysvětlení zásad použitého jazyka. Předseda komise i oponenti byly s odpověďmi kandidáta spokojeni.

*Diskuse:*

Doc. Souček připomněl, že významy v textu jsou závislé na kontextu (holistický přístup) a položil otázku, zda je možnost analyzovat kontext ve smyslu elementů. Kandidát odpověděl, že existuje problém paradigmatických vztahů – zatím není možné všechny směry zkoumat najednou (příliš komplexní a složité), je nutno vybrat doménu, která se zkoumá. Prof. Uličný doplnil souvislost s teorií mluvních aktů.

Prof. Konigová se zajímala o to, kdy doc. V. Smetáček začal s myšlenkou umělého jazyka a systému Sémán a zda byl v této oblasti podán nějaký patent. Mgr. Chýla odpověděl, že původní návrh pochází cca z let 1979-80, patent nebyl podán, nyní je tento nástroj dostupný zdarma všem. Prof. Konigová upozornila na zajímavost teorie myšlenkových map, která vyjadřuje nelineárnost lidského myšlení (přístup založený na asociacích). Tento princip podle kandidáta nelze aplikovat v systému Sémán.

Doc. Ivánek se zajímal o metody testování použité v disertaci, např. Fisherův test – R. Chýla se řídil počítačovou lingvistikou.

Následovala další diskuse, kterou předseda komise shrnul a uzavřel takto: práce se jeví jako přínosná pro vývoj škály přístupů k obsahové analýze a možnosti přesnějšího porovnávání obsahu textů a překladů.

Členové komise přistoupili k tajnému hlasování.

Vyhlášení výsledků tajného hlasování:

Komise navrhla udělit titul doktor (Ph.D.)

V Praze 27. 1. 2012 zapsala Mgr. Věra Pilecká

Podpis předsedy komise: