

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Kristýna Neumannová	
Název práce	Identification and analysis of Czech equivalents of German compounds	
Rok odevzdání	2021	
Studijní program	Informatika	
Studijní obor	Obecná informatika Softwarové a datové inženýrství Programování a softwarové systémy *	
Autor posudku	RNDr. Daniel Zeman, Ph.D.	Vedoucí – Oponent *
Pracoviště	Ústav formální a aplikované lingvistiky	

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámec obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

K celé práci

lepší OK horší nevyhovuje

K celému zadání	lepší	OK	horší	nevyhovuje
Obtížnost zadání		×		
Splnění zadání		×		
Rozsah práce <i>... textová i implementační část, zohlednění náročnosti</i>		×		
Komentář				
Zajímavé téma i jeho zpracování. Z informatického hlediska je asi nejzajímavější sekce 6.2.5, kde jsou navrženy a použity různé metody, které se snaží (i s využitím existujících softwarových nástrojů pro zpracování přirozeného jazyka) rozložit český překlad německého složeného slova na komponenty, které půjde identifikovat s komponenty německé předlohy.				
K celkovému shrnutí práce se nabízí otázka (v práci nediskutovaná), zda autorka má nějakou představu i o opačném směru: kolik (pokud vůbec) českých složených slov v paralelním textu odpovídá německé frázi, která není složeným slovem?				

Posudek pokračuje na druhé straně.

* Nehodící se škrtněte.

Textová část práce

lepší OK horší nevyhovuje

		lepší	OK	horší	nevyhovuje
Formální úprava	<i>... jazyková úroveň, typografická úroveň, citace</i>		×		
Struktura textu	<i>... obsahuje kontext, cíle, analýzu, návrh, vyhodnocení</i>		×		
Analýza			×		
Vývojová dokumentace				?	
Uživatelská dokumentace				?	
Komentář Se strukturou textu, typografickou úpravou i jazykem práce jsem spokojen. Práce je psána srozumitelnou angličtinou, drobné chyby/překlepy v ní jsou, ale není jich mnoho. Práce je atypická (v rámci inženýrských BP) tím, že jejím cílem není jeden kus software, ale spíše analýza dat. Implementace navržené metody spočívá (pravděpodobně) v kombinaci řady drobných skriptů a několika existujících „velkých“ nástrojů jako FastAlign a UDPipe, které ale nejsou dílem autorky práce. Z tohoto pohledu je do jisté míry pochopitelné, že práce neobsahuje vývojovou ani uživatelskou dokumentaci. Přesto bych ale ocenil, kdyby k textu práce byly přiloženy jak dotyčné skripty, tak data, o kterých se v práci mluví (mám na mysli zejména výstupní data a ručně anotovaná data, nikoli vstupní korpusy). Pak by samozřejmě byla na místě i stručná dokumentace. Práce ale žádnou takovou přílohu nemá.					

Implementační část práce

lepší OK horší nevyhovuje

		lepší	OK	horší	nevyhovuje
Kvalita návrhu	<i>... architektura, struktury a algoritmy, použité technologie</i>		×		
Kvalita zpracování	<i>... jmenné konvence, formátování, komentáře, testování</i>		?		
Stabilita implementace			?		
Komentář Jak popisují výše, většina kritérií v této části není pro posuzovanou práci relevantní.					

Celkové hodnoceníVýborně – Velmi dobře – Dobře – ~~Dobře~~ – Neprospěl(a) ***Práci navrhuji na zvláštní ocenění**~~Ano~~ – Ne *Datum 15. 6. 2021

Podpis _____

* Nehodící se škrtněte.