



## POSUDEK VEDOUCÍHO BAKALÁŘSKÉ PRÁCE

**Jméno studenta:** Vlasta Matějková

**Název práce:** Nástroje pro přípravu českého textu na strojové zpracování

**Autor posudku:** Mgr. Jiří Haviger, Ph.D.

**Cíl práce:** Cílem práce je popsat specifika českého textu vzhledem k angličtině, získat přehled dostupných nástrojů k preprocessingu a otestovat je na reálných textech.

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)					
	A	B	C	D	E	F
Práce svým zaměřením odpovídá studovanému oboru	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vymezení cíle a jeho naplnění	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### Vyjádření k výsledku anti-plagiátorské kontroly

Celková podobnost 0%.

### Dílní připomínky a náměty:

Autorka v práci porovnávala několik nástrojů – balíčků v jazyce Python na různé části preprocessingu českého textu, a to na dvou dostupných textech – Saturnin a Povídky o pejskovi a kočičce.

Struktura práce je velmi dobrá, čtenář se rychle orientuje v tom, která část popisuje aktuální stav poznání, která data se použila, které metody byly použity i kde najde výsledky použitých metod.

K popisu současného stavu poznání by bylo vhodné mít více literatury, jinak připomínky nemám. Ohledně dalších částí jen drobnosti, např. ve shrnutí výsledků zdůraznit, kterého nástroje se daný odstavec týká.

Osobně mohu pochválit i přiměřený rozsah, protože práce neobsahuje zbytečné odstavce a čtenář se rychle dostane k hlavním cílům.

**Celkové posouzení práce a zdůvodnění výsledné známky:**

Práce splňuje požadavky kladené na bakalářskou práci.

**Otázky k obhajobě:**

V tabulce 7 uvádíte v prvním řádku počet slov bez uvedení nástroje, kterým jste to měřila. Uveďte prosím rozdíl mezi pojmy token, slovo a lemma a specifikujte, co jste měla v tabulce 7 na mysli.

**Práci doporučuji k obhajobě.**

**Navržená výsledná známka: A**

**V Hradci Králové, dne 14. května 2024**

---

**podpis**