

Posudek oponenta diplomové práce

Jméno a příjmení studenta	Petra CHVOSTOVÁ
Název práce	Návrh aplikace pro automatizovanou extrakci atributů produktů webů
Název a typ studijního programu	Systémové inženýrství a informatika / Navazující
Fakulta / katedra	Ekonomická fakulta / KMI
Vedoucí práce	Beránek Ladislav, doc. Ing. CSc., MBA
Oponent	Mgr. Petr Kursá

Posudek závěrečné práce

1. Význam a náročnost tématu 2.0

2. Logická struktura práce 2.0

Poznámka: Logická struktura je uspokojivá, odpovídá vcelku standardu požadovanému pro tento typ práce

3. Naplnění cíle práce 2.5

Poznámka: Cílů práce bylo v zásadě dosaženo. V zadání bylo ale ještě návrh na kategorizaci dat a návrh, jak je dále zpracovávat. To zpracováno nebylo, ani v práci není nějaký návrh.

4. Metodický postup 3.0

Poznámka: V teoretické části mě chybí rozbor stávajících nástrojů nebo řešení, např. Scrapy. V praktické části jsou popsány skripty jazyka R. Avšak automatickou extrakci bych si představoval trochu jinak.

5. Hodnocení teoretického zobrazení a přínosu 3.0

Poznámka: Práce nemá velký teoretický ani praktický přínos. Může sloužit spíše jako ukázka skriptů jazyka R pro práci s webovými stránkami.

6. Práce s literaturou 2.5

Poznámka: V teoretické části mohlo být více odkazů. Prací na toto téma existuje celá řada.

7. Formální stránka 2.5

Poznámka: Formální stránku práce kazí způsob číslování (nečíslování) kapitol. Vyskytují se i drobné gramatické chyby

Závěr

Hodnocení práce (známka): **dobře**

Doporučuji práci k obhajobě: **ANO**

Otázky pro diskusi a poznámky

Kritické poznámky a celkový přínos, celková hodnota práce

Práce ukazuje použití jazyka R pro extrakci dat z webů. Bohužel se podle mého názoru nejedná o automatickou extrakci dat. Chybí dále návrh, co se získanými daty, alespoň v diskuzi. Není dostatečně popsána metodika, architektura řešení, není provedena dostatečná analýza. V teoretické části také chybí přehled možných nástrojů a řešení. Některé popisy jsou zbytečně

rozvláčné (např. jazyk html a další). Chybí návrh testů, vyhodnocení, porovnání. Přínosem práce je ukázka skriptů jazyka R, které mohou posloužit ve výuce.

Otázky nebo témata pro diskusi před komisí

1. Popište nějaká existující aplikace/řešení pro extrakci dat z webů.
2. Dalo by se zařadit do vašich skriptů automatické procházení webů (třeba předem definovaných)?

Datum: 05.10.2022

Podpis oponenta