FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Posudek oponenta závěrečné práce

Student: Jakub Drahoš
Oponent práce: Ing. Marek Suchánek
Název práce: Webová aplikace pro online web scraping
Obor: Webové a softwarové inženýrství

Datum vytvoření: 28. 5. 2019

Hodnotící kritérium:	*Způsob hodnocení – následující škálou 1 až 4:*
1. Splnění zadání	***1=zadání splněno,*** *2=zadání splněno s menšími výhradami,* *3=zadání splněno s většími výhradami,* *4=zadání nesplněno*

Popis kritéria:
Posuďte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Komentář:
Zadání považuji za splněné i přes to, že aplikace není webová v typickém smyslu, ale jedná se o rozšíření webových prohlížečů (strukturou a technologiemi podobná webovým aplikacím - HTML, CSS, JS) - to je trochu nešťastné, ideální by bylo, kdyby s vedoucím práce přeformulovali zadání a název například pouze na "Aplikace pro online web scraping". Nicméně v práci je to vícekrát vysvětleno a zdůvodněno a je to i zcela pochopitelné. Jinak jsou všechny body naplněné a v práci dobře popsané.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
2. Písemná část práce	*80 (B)*

Popis kritéria:
Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Komentář:
Jednotlivé části práce na sebe logicky navazují dle tradičního cyklu vývoje SW (v souladu se zadáním práce) a obsahují relevantní informace. V práci jsou jen občasné typografické prohřešky, překlepy a gramatické chyby (velká/malá písmena v názvech, čárky, konzistence v použití stylů písma, apod.). Převzaté informace jsou řádně citovány dle citačních zvyklostí - jediný problém vidím v sekci 1.2, kde je celá strana bez citací a přitom obsahuje informace, které si autor "nevymyslel". V následujících podsekcích již jsou zdroje povětšinou správně a vhodně uvedeny.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
3. Nepísemná část, přílohy	*95 (A)*

Popis kritéria:
Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce
– opakovatelnost experimentů

Komentář:
Zdrojový kód rozšíření je dobře strukturovaný, přehledný a řádně otestovaný. Rozšíření do prohlížeče je funkční a snadno se s ním pracuje, není nutné studovat návod. Jediné co bych vytknul je absence souboru README, který by popisoval projekt a například spouštění testů či požadavky pro vývoj nebo běh.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
4. Hodnocení výsledků, jejich využitelnost	*100 (A)*

Popis kritéria:
Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Komentář:

Výsledkem je funkční rozšíření do prohlížeče Google Chrome (ale také Chromium a Opera), které umožňuje snadné získávání "tabulkových" dat z webových stránek pomocí intuitivního rozhraní. Rozšíření je velmi snadné na použití a přitom nabízí užitečné funkce na vysoké kvalitě. Zároveň je po technické stránce aplikace dobře navržena a implementována, což usnadní případný další rozvoj.

Hodnotící kritérium:	*Způsob hodnocení – nehodnotí se*
5. Otázky k obhajobě	

Popis kritéria:
Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

Otázky:

- Plánujete na projektu dále pracovat a otevřeně rozšíření distribuovat? Pokud ano, pod jakou licencí a proč?
- Bylo by možné (a případně jak) jednoduše rozšířit Váš Scraper o crawling například výčtem adres či označením typu odkazů k následování tak, aby bylo zachováno intuitivní rozhraní (viz sekce 4.4)? Jak byste případně postupoval?

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
6. Celkové hodnocení	*90 (A)*

Popis kritéria:
Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.

Text hodnocení:

Celkově práci na základě předchozích komentářů hodnotím jako výbornou. V rámci práce bylo vytvořeno kvalitní softwarové dílo ve formě rozšíření pro prohlížeč Google Chrome (i dalších obdobných), které umožňuje snadné a intuitivní získávání dat z webových stránek.

Podpis oponenta práce: