

Posudek oponenta bakalářské práce

Student: Matuš Adam
Téma: Služba pro archivaci webového obsahu (id 21833)
Oponent: Polčák Libor, Ing., Ph.D., UIFS FIT VUT

- 1. Náročnost zadání** **jednoduché zadání**
Samotné zadání nedává konkrétní požadavky na strukturu požadovaného API pro práci s frameworkem Lemmiwinks. Student následně vytvořil pouhé dva koncové body, každý implementoval v samostatné aplikaci.
- 2. Splnění požadavků zadání** **zadání splněno s vážnými výhradami**
Práce však nepopisuje novější verze protokolu HTTP: HTTP/2 a připravované HTTP/3. Vzhledem k tomu, že tyto detaily vytvořené aplikace neovlivňují, nepokládám tento nedostatek za důležitý.
Podle mého názoru mělo v bodě 3 vzniknout API, které dává uživateli lepší kontrolu nad procesem stahování, poskytuje autentizaci a autorizaci uživatelů a umožňuje stáhnout výsledky archivace.
- 3. Rozsah technické zprávy** **splňuje pouze minimální požadavky**
- 4. Prezentací úroveň předložené práce** **65 b. (D)**
Smysl rozdělení na stahovací a archivační službu není vysvětlen pochopitelně. Z textu práce mně nebylo zřejmé, co jednotlivé služby vykonávají a jaká je návaznost na Lemmiwinks. Např. na str. 15 se píše: "*služba pomoci analýzy získaných zdrojů zjistí, které další zdroje je potřeba stáhnout. V případě úspěchu stahování všech potřebných zdrojů se ze získaných dat vytvoří archiv ve formátu MAFF.*" Veškerou tuto činnost vykonává Lemmiwinks a analýza zdrojů a stahování potřebných zdrojů není předmětem této práce.
Některé části práce, jako např. popis protokolu HTTP jsou napsané pěkně.
- 5. Formální úprava technické zprávy** **75 b. (C)**
Práce obsahuje drobné typografické prohřešky jako chybějící mezery před citacemi, obrázky v anglickém jazyce, i když existuje česká terminologie (např. obr. 2.1), a nedodržení okrajů (např. str. 22). Strany 28-33 jsou potištěny jen částečně.
- 6. Práce s literaturou** **80 b. (B)**
Práce obsahuje podstatné zdroje. U zdrojů č. 1, 5, 9, 13 a 17 jsou jména některých autorů nahrazena jen jejich počátečními písmeny.
- 7. Realizační výstup** **50 b. (E)**
Vytvořené API je velice jednoduché a neumožňuje ani autentizaci a tím pádem se každý, kdo práci nasadí bez dodatečných omezení vystavuje riziku sehrání role prostředníka pro potenciálního útočníka.
Zdrojové kódy nejsou komentované a není je možné otestovat automatickými testy. Práce neobsahuje návod pro instalaci. Autorovi práce trvala instalace u mě v kanceláři několik desítek minut.
V archivu je odevzdán adresář lemmiwinks, jehož autorem je vedoucí práce, práce studenta a vedoucího práce není odlišena.
- 8. Využitelnost výsledků**
Práce je natolik jednoduchá, že její přínos je jen velmi malý. Největší přínos vidím v sekci 6.3, která obsahuje seznam části skrytých stránek v síti Tor.
- 9. Otázky k obhajobě**
 - Proč rozlišujete archivační a stahovací službu? Jaký kód se mezi službami sdílí a jaký by se potenciálně mohl sdílet?
 - Je možné nějakým způsobem měnit parametry archivace, jako preferovaný jazyk, parametry Toru (např. výstupní uzel v určité zemi pro získání konkrétní jazykové mutace), či *User Agent String*, který pro síť Tor není nastavený na Torbrowser, čímž může být archivační služba nápadná?
- 10. Souhrnné hodnocení** **55 b. dostatečně (E)**
Pan Matuš splnil požadavky kladené na vypracování bakalářské práce. Splnění bodu 3 je však nejjednodušší možné. Pro praktické nasazení by práce vyžadovala větší zamyšlení nad poskytovaným API. Výsledky práce nejsou jednoduše nasaditelné, protože chybí návod pro instalaci vytvořených programů. Navrhuji hodnotit práci stupněm E.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 24. května 2019

.....
podpis