

Posudek oponenta diplomové práce

Autor práce: **Martin Hospaska**

Název práce: **Systém porovnání studijních programů českých univerzit**

Obsah práce

Autor se zabývá vysoce aktuální problematikou automatického porovnávání obsahu infromatických předmětů nabízených vysokými školami v ČR. Za tímto účelem navrhl a vytvořil systém, který stahuje informace o předmětech z webových stránek deseti českých univerzit, ukládá je do datového úložiště, umožňuje v nich vyhledávat podle vybraných kritérií, porovnávat a ve svém konečném důsledku doporučuje uživateli univerzitu nabízející nejvíce relevantních předmětů. Zamýšlenou cílovou skupinou uživatelů vzniklého systému jsou studenti posledních ročníků středních škol.

Kvalita řešení a dosažených výsledků

Kvalitu řešení a dosažených výsledků považuji za velmi dobrou. Student se rozhodl jít cestou vytvoření vlastního webového robota zkombinovaného s knihovnamy a moduly třetích stran. Důsledkem tohoto přístupu je menší rozsah vytvořeného zdrojového kódu, ale o to větší nároky jsou kladeny na sladění a integraci všech komponent. Výsledná webová aplikace byla otestována množinou uživatelů, ale není bohužel uvedeno jak velkou (s. 79). Nevýhody zvoleného přístupu se podle mého názoru projeví především při budoucím rozšiřování systému také na jiné univerzity a obory, což je ostatně také již naznačeno na s. 80 a 81, kde se pro přidání další univerzity počítá s časem kolem šesti hodin a důkladnou znalostí zdrojového kódu. Není tedy zcela jasné, jak by se systém vypořádal s měnící se strukturou webových stránek desítek, stovek, či tisíců vysokých škol.

Formální úroveň

Formální úroveň diplomové práce (DP) je rovněž velmi dobrá. Text práce je logicky vhodně členěný a autorův projev je na většině z 80 hlavních stránek kultivovaný a bez většího množství chyb a překlepů. Rušivě působí jen některé anglicismy, které přitom zavedený český ekvivalent mají, např. *crawler* (webový pavouk nebo robot), *crawling* (stahování nebo procházení) a s tím související podivné výrazy typu „crawlerem vyselektován“ (s. 91), „předmět je vylistován“ (s. 95) či „terminováním“ (s. 102). Mnohé podrobnosti z implementační kapitoly 8 by se hodily již spíše do přílohy a v příloze A (s. 90 – 97) mi poněkud chybějí náhledy obrazovek jednotlivých webových stránek sledovaných univerzit.

Práce s literaturou

Práce s literaturou je odpovídající a zahrnuje 35 citovaných pramenů, z nichž jen pět jsou tištěné publikace. To lze ale u takto technologicky náročného tématu pochopit.

Splnění zadání

Zadání DP bylo splněno, i když budoucí rozšíření vytvořené aplikace o další české nebo zahraniční univerzity a obory bude zřejmě ztíženo omezeními vynucenými současnou architekturou celého systému.

Doplňující informace k práci

Největším nedostatkem této DP je podle mého názoru neflexibilní vlastní webový pavouk s natvrdo zabudovanou strukturou stahovaných stránek, což má za následek nefunkčnost programu (anebo nutný zásah do jeho zdrojového kódu) při jakékoliv změně sledovaných webů. Tyto změny přitom evidentně nejsou systémem nijak monitorovány, aby se na ně dalo reagovat ještě před stahováním dat. Pokud už se autor rozhodl pro vlastní webový pavouka, měl se alespoň pokusit do něj vložit nějakou umělou inteligenci, která by z něj nedělala jen „hrubou sílu“, ale usnadnila i jeho rozšíření na další univerzity. V opačném případě měl autor raději sáhnout k nějakému z existujících řešení, jejichž přehled (s. 11 – 13) ale zcela jistě není dostačující.

Dotazy k práci

Jak systém zjistí, že se struktura webových stránek změnila? (s. 45)

Jak je Kibana integrována do celé aplikace? (s. 53 - 56)

Kolik uživatelů testovalo webovou aplikaci? (s. 79)

Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.

V Plzni dne 6. června 2017



doc. Ing. Dalibor Fiala, Ph.D.

**SOUHLASÍ
S ORIGINÁLEM**



Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky

①