

# Posudek oponenta diplomové práce

Autor/autorka práce: **Bc. Jan Černogurský**

Název práce: **Charakteristiky OSS projektů pro selektivní těžbu a analýzu jejich dat**

Obsah práce

Diplomová práce (DP) se zabývá metodikami vývoje software, projektovým řízením v oblasti vývoje sw, open-source softwarovými (OSS) projekty, získáváním projektových dat z relevantních zdrojů, nástrojem SPADe pro detekci procesních anti-vzorů, který je vyvíjený na KIV, a především analýzou a výběrem vhodných charakteristik OSS projektů s tím, že globálně sledovaným cílem je detekce procesních anti-vzorů v projektových datech. Součástí DP je také návrh, implementace a testování pomocné aplikace pro vyhledávání a získávání zdrojových projektových dat na platformě GitHub.

Kvalita řešení a dosažených výsledků

Úvodní tři kapitoly DP jsou popisné a textově zbytečně rozsáhlé (např. popis metodik vývoje software považuji za příliš vzdálený jádru práce). S ohledem na délku textu a způsob vyjadřování autora (více v hodnocení Formální úroveň) považuji tuto část práce za informačně chudou. Kapitola 4 již obsahuje jádro DP, tedy popis a analýzu charakteristik OSS projektů. Autor nejprve zkoumá pět nástrojů/zdrojů dat OSS projektů, z nichž záhy vybírá GitHub pro další úvahy. Autor zde píše, že prováděl důkladný průzkum nástrojů, nicméně z textu to není patrné. V kapitole pak následuje relativně podrobný popis jednotlivých charakteristik OSS projektů, u charakteristiky 4.7 Issues není zřejmé, jak moc souvisí podkapitoly 4.7.3 až 4.7.6 s tématem kapitoly 4.7. Celkově mi pak není jasné, jestli autor vybíral popsané OSS charakteristiky pouze dle struktury GitHubu, nebo hledal oporu i v literatuře či dalších zdrojích OSS projektů.

V kapitole 5 se znovu objevuje (bez porovnání s ostatními platformami) a rozšiřuje zdůvodnění výběru GitHubu jako zdroje dat OSS projektů. Poté je popsáno API GitHubu a vybrána podmnožina dále zpracovávaných charakteristik OSS projektů, a to s ohledem na náročnost získávání dat z GitHubu a jejich tzv. informačně-hodnotový přínos (str. 72). Autor se při tomto výběru odkazuje na kapitolu 4 a popis jednotlivých charakteristik OSS projektů v ní obsažených; nejsem však schopen z textu myšlenkově odvodit stejné závěry autora (tedy konečný výběr podmnožiny vhodných charakteristik OSS projektů).

Autor pak dále popisuje návrh pomocné aplikace pro vyhledávání a získávání projektových dat z GitHubu (tato část by si zasloužila samostatnou kapitolu). Tento popis neodděluje specifikaci požadavků na aplikaci a návrh samotný; nejsou také použity standardizované diagramy a notace. Kapitola Implementace pak míchá implementační detaily s úvahami nad výkonovými parametry dotazování do GitHubu a rozšiřitelností aplikace. Popis testování (kapitola 7) je velmi vágní. Samotná aplikace je funkční. I přes výhrady k podobě uživatelského rozhraní plní požadavky na ni kladené.

Formální úroveň

Celá DP má 113 stran; na úrovni nadpisů kapitol a podkapitol je (kromě částí popsaných výše) rozumně strukturovaná, text se však velmi špatně čte. DP je zejména v úvodních třech kapitolách psána (polo)marketingovým (nikoli technickým) stylem; organizace textu v jednotlivých

podkapitolách pak působí dojmem náhodného mixu (pravděpodobně přeložených a v různých obměnách podaných) obecných sdělení z oblasti softwarového inženýrství a přeložených (propagačních) materiálů využitých nástrojů a technik. Takový popis v češtině pak působí jako „výplňový“ a řada sdělení přestává dávat i smysl, např.

„...přičemž je třeba věnovat velkou pozornost nejen finančním rozhodnutím, ale také časovým rámcům, které jsou ústředními aspekty kulminujícími v míře úspěšnosti – což je zásadní faktor.“ (str. 23), „Pozorování také umožňuje uznání těm vývojářům, kteří nabízejí vstupy do projektu určením jejich významu pro něj, čímž je odpovídajícím způsobem odměňuje na základě úrovně jejich příspěvku (str. 36) či „Důslednou prací na uvedeném snažení zaručuje jeho účelnost a relevanci pro publikum, které jej často využívá“ (str. 46).

Netechnické (vágní, redundantní, povrchní, „vypravěčské“) vyjadřování provází celou práci, v částech, ve kterých autor již popisuje vlastní realizaci, je však trochu snesitelnější.

#### Práce s literaturou

Autor cituje celkem 28 zdrojů, tyto odpovídají předkládanému obsahu práce. Překvapuje mě však, že počet zdrojů, které se věnují jádru práce, tedy analýze a výběru charakteristik OSS projektů, je v tomto výčtu relativně malý.

#### Splnění zadání

Zadání bylo splněno (se značnými výhradami ke kvalitě textu DP).

#### Doplňující informace k práci

Na práci mám velmi ambivalentní pohled. Autor splnil body zadání, nepochybuji o tom, že se práci věnoval; tato pak zřejmě přispívá k efektivitě výběru projektů vhodných pro další zpracování nástrojem SPADe. Text DP a obecně způsob vyjadřování autora jsou však příliš vzdálené standardnímu techničtěji orientovanému dílu.

#### Dotazy k práci

1. Jakým způsobem jste sbíral a posléze vybíral vhodné charakteristiky OSS projektů? Byly poplatné struktuře GitHubu? Jak se liší/jsou shodné s odbornou literaturou?
2. Popis testování v kapitole 7 je velmi vágní. Píšete o potřebě objektivit testování, kritériích výběru testerů, zpětných vazbách a následných úpravách aplikace, to vše bez konkrétních údajů. Alespoň některé z těchto údajů prosím uveďte.

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