

# Posudek oponenta bakalářské práce

Autor/autorka práce: **Jan Vandlíček**

Název práce: **Analýza profilu tvůrčí činnosti FAV v období 2018-2023**

Obsah práce

Autor nejprve poskytuje čtenáři základní přehled o tvůrčí činnosti na FAV na základě výročních zpráv a přidává kontextuální informaci o pozici ZČU v mezinárodních žebříčcích pro hodnocení univerzit. V další části jsou představeny Datový sklad ZČU a způsob evidence výsledků tvůrčí činnosti. Následuje kapitola věnovaná nástroji Oracle APEX používaným v ekosystému Datového skladu ZČU. V kapitole 5 jsou definovány metriky, které budou použity v analýze, a popsána implementace dashboardu v prostředí Oracle APEX, který má za cíl odpovědět na 4 autorem stanovené otázky. Zároveň jsou zde i diskutovány výsledky. V závěrečné kapitole autor nabízí kromě souhrnu i sadu doporučení pro budoucí směřování analytik ohledně tvůrčí činnosti.

Kapitoly se mi zdají obsahově nevyvážené. Student popisuje např. teorii datových skladů a tržišť, i když na jádro práce nemá toto vliv. Naopak diskuze výsledků, popis struktury a obsahu dostupných dat, či popis možnosti úprav vizualizací v APEX by si dle mě zasloužily více pozornosti.

Kvalita řešení a dosažených výsledků

Autor v rámci práce sestavil tabulku se 72 kandidátními indikátory k posuzování tvůrčí činnosti, z nichž malou podmnožinu použil pro svou analýzu. Je škoda, že kromě určení priority (důležitosti) indikátorů není v tabulce i informace, kde se autor s indikátorem setkal, jakým způsobem byl jinými autory definován, či co brání jeho vyhodnocení pro FAV/ZČU. Takto vypracovaný přehled by mohl dobře posloužit pro případné další analytické práce. Autorem není vysvětleno, proč jsou v tabulce některé indikátory obarveny červeně a některé jsou přeškrtnuté.

V kapitole 5.2.1 autor dobře poukazuje na problém s rozpočítáváním publikací, na kterých se podílelo více sekcí FAV, mezi jednotlivé sekce a nabízí řešení, které by zachovalo správné celkové součty publikací v grafech. Toto řešení ale nakonec z nějakého důvodu nepoužívá a výsledný graf je tedy zkreslený.

Tvar grafu zobrazující počet článků s afilací FAV v databázi Scopus v jednotlivých letech (obrázek 5.1, str. 34) se autor snaží vysvětlit pandemií COVID19, ale ponechává bez komentáře, že absolutní hodnoty vynesené v grafu jsou (3, 2, 2, 3, 1, 1), což na první pohled ukazuje na problém v datech nebo zpracování. Kontrolou skrze databázi Scopus lze vidět, že v příslušných letech je evidováno řádově více položek. Očekával bych, že tato neshoda bude v textu diskutována.

Ve zpracování 5.2.3. autor neprovádí filtraci na období let 2018 – 2023 jako v ostatních případech, což v kontextu této práce považuji za nevhodné. Zároveň příslušný graf není sestaven intuitivně, když osa X obsahující počty autorů začíná na hodnotě -5 a kvartily jsou vynášeny jako čísla 1, 2, 3, 4 na reálné ose. Překrývání ploch jednotlivých bublin také není přehledné. Z textu není zřejmé, zda je možné v prostředí APEX vizualizaci lépe stylizovat a upravit, nebo jestli je autor limitován pouze jedním základním způsobem vykreslení.

V podkapitole 5.2.4. autor uvádí, že je pro něj podstatné porovnání počtu publikací v Q1 a Q2 proti Q3 a Q4. Tato hodnota ale není spočtena ani prezentována v grafu.

#### Formální úroveň

V textu nejsou zásadní typografické chyby a množství překlepů je zanedbatelné.

V textu vycházejícím z výročních zpráv FAV (2.1.1) je citována řada čísel pro jednotlivé roky a je obtížné si na základě toho udělat představu o vývoji. Ocenil bych v této části přidání grafického souhrnu čísel – tabulkou, případně grafem. Zároveň je nesystematické, že u roku 2022 autor uvádí 7 různých získaných ocenění s vazbou na FAV, ale v ostatních letech není uvedeno žádné.

#### Práce s literaturou

Použité zdroje jsou pro práci relevantní. V textu zabývajícím se mezinárodními žebříčky univerzit chybí odkaz na datový zdroj, resp. tabulku s žebříčkem

#### Splnění zadání

Zadání bylo splněno s výhradami.

Bod 2 by bylo vhodné detailněji rozebrat. Z mého pohledu je pro kontext práce důležité uvést, jaké atributy jsou v příslušném pohledu datového skladu vystaveny, vysvětlit jejich význam, či kriticky zhodnotit kvalitu dat v dostupném datasetu.

Bod „*Seznamte se s nástrojem Oracle APEX pro vizualizaci dat*“ není dostatečně řešen. Z textu není zřejmé, jak lze vizualizace modifikovat, zda je možné definovat vlastní typ grafu, nebo jestli je možné v dashboardech mít interaktivitu mezi grafy (cross-filter) běžnou v dashboardovacích nástrojích. Uživatelské prostředí je představeno pouze obrázkem 4.1, bez dalšího vysvětlení v textu.

#### Dotazy k práci

- Jak složité by bylo v prostředí APEX implementovat vlastní typ grafu, například violin plot zmiňovaný v kapitole 5.2.3?
- Co způsobilo, že součet publikací za jednotlivé roky na obrázku 5.1 str. 34 se pohybuje v rozsahu 1-3?
- Jedním z indikátorů v přiložené tabulce je „Počet sdílení publikací na sociálních sítích“. Dokážete naznačit, jakým způsobem by se získávala podkladová data pro vyčíslení tohoto indikátoru?

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