

Hodnocení vedoucího diplomové práce

Autor/autorka práce: **Bc. Josef Bozděch**

Název práce: **Možnosti analytického rozšíření úložiště Data Lakehouse**

Aktivita studenta

Aktivita studenta byla příkladná. Nesnadné téma, které navazovalo na předchozí diplomovou práci Lukáše Moučka úspěšně obhájenou v akademickém roce 2023/2024, mělo řadu úskalí. Diplomant inovoval realizovaný koncept datového úložiště heterogenních dat a snažil se o uspokojivé řešení zásadního problému Data Lakehouse – vytvoření datové vrstvy umožňující analytické procesy.

Diplomant po celou dobu úzce spolupracoval s vedoucím práce, na stanovené schůzky vždy docházel připraven.

Téma kvalifikační práce koresponduje se současnými trendy analytického zpracování velkých heterogenních dat. V této oblasti je známa platforma Databricks založená na technologii Lakehouse a komerční řešení Azure Databricks. Koncept datového úložiště Data Lakehouse vychází z koncepce Databricks. V průběhu implementace původního návrhu došlo k celé řadě podstatných úprav, navržené úložiště je originální.

Kvalita řešení odpovídá složitosti navrhovaného systému a postupnému získávání zkušeností se zpracováním specifické oblasti dat. Problém, který nebyl včas identifikován a do jisté míry ovlivnil kvalitu řešení, spočíval v modifikované struktuře a způsobu extrakce dat ze zdrojového systému. Poměrně značnou dobu se student orientoval ve stávající aplikaci. Posléze pokračoval v implementaci pomocí frameworku SpringBoot, dynamickou práci s webovou aplikací a zlepšení interakce s uživatelem realizoval aplikací knihovny Preact.js a Chart.js.

Aktuální řešení Data Lakehouse není konečné, bude nutno dále vyvíjet.

Menší výhrady lze mít ke stylistické a formální stránce diplomové práce, které se diplomantovi nepodařilo zcela odstranit.

Pro plánované nasazení v rámci analytického zpracování klinických dat jsou získané zkušenosti a výsledky přínosné. Pro vědeckou práci v oblasti medicínské informatiky je vývoj datového úložiště typu Data Lakehouse důležité, stávající systém MRE poskytuje výborné prostředí pro bezpečné, jednoduše modifikovatelné uložení dat a získávání dat, systém není přizpůsoben pro analytické zpracování.

Zadání diplomové práce bylo splněno.

Vzhledem k obtížnosti tématu jak po stránce teoretické, tak praktické, a aktivnímu přístupu diplomanta navrhuji hodnocení známkou **výborně** a práci doporučuji k obhajobě.