

## I. IDENTIFIKAČNÍ ÚDAJE

<b>Název práce:</b>	<b>Využití konceptu BIG data v oblasti Geodézie a Kartografie</b>
<b>Jméno autora:</b>	<b>Martin Vajner</b>
<b>Typ práce:</b>	diplomová
<b>Fakulta/ústav:</b>	Fakulta stavební (FSv)
<b>Katedra/ústav:</b>	K155 - Katedra geomatiky
<b>Oponent práce:</b>	Ing. Bronislav Koska, Ph.D.
<b>Pracoviště oponenta práce:</b>	Valeo Detection Systems s.r.o.

## II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

<b>Zadání</b>	<b>mimořádně náročné</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání diplomové práce je velmi náročné. Zadání klade na studenta vysoké nároky na znalost programování (Python, Java) a na dobrou orientaci ve světě Big Data. Práce je zaměřena na praktické využití analytických nástrojů v čele s Apache Sparc, student tedy musí zvládnout ekosystém Apache Hadoop/Apache Sparc.	

<b>Splnění zadání</b>	<b>splněno s menšími výhradami</b>
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
„Cílem diplomové práce je analýza využití Big data v oblasti geodézie a kartografie.“ Zadání práce bylo splněno jenom částečně. Nejsou uvedeny oblasti z oboru geodézie a kartografie, kde se zkoumané přístupy a nástroje používají a kde by je bylo vhodné a smysluplné používat. Uvedený demonstrační příklad do této kategorie nespadá. „a práci s přicházejícími daty v reálném čase.“ – Práce s daty přicházejícími v reálném čase nebyla demonstrována.	

<b>Zvolený postup řešení</b>	<b>správný</b>
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Student zvolil správný postup – nejprve teoreticky analyzuje problematiku Big Data, následuje popis nástrojů Big Data. V této části se student věnuje popisu ekosystému Apache Hadoop a Apache Sparc. Nabyté znalosti se student pokouší aplikovat na několika příkladech z geodézie a kartografie.	

<b>Odborná úroveň</b>	<b>C - dobře</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Během čtení práce je evidentní, že student není erudovaným programátorem. Student rozumí nástrojům Big Data, bohužel nebyl schopen nástroje srozumitelně popsat a vysvětlit jejich hlavní výhody při použití. V některých kapitolách jde student do zbytečných detailů, následně věnuje méně prostoru klíčovým tématům týkající se Apache Sparc.  Histogram odchylek od průměrných hodnot na straně 59 neodpovídá žádnému známému rozdělení pravděpodobnosti a je zjevně chybný (četnost bodů v těsné blízkosti průměrné hodnoty je minimální a dále se zvyšuje). Vzhledem k tématu práce jde o okrajovou záležitost, ale technicky smýšlející člověk by si zcela neočekávaných výsledků měl všimnout.	

<b>Formální a jazyková úroveň, rozsah práce</b>	<b>C - dobře</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
V této oblasti je v práci celá řada nedostatků například: S. 8 - několikrát psáno o bilionech lidí. Billion je $10^{12}$ . Je pravděpodobné, že tento text vznikl kopírováním anglického textu a následně chybným překladem. Mnoho obrázků obsahuje anglický popis, stejně tak tabulka na straně 12, u které není uveden titulek ani zdroj!	

**Výběr zdrojů, korektnost citací**

**B - velmi dobře**

*Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.*

Student využil relevantní zdroje, stránky apache.org které jsou základem pro pochopení ekosystému Apache Hadoop a Apache Sparc. Oceňuji, že pro zvládnutí problematiky pro dimplovou práci, student využil systém Udemy.

**Další komentáře a hodnocení**

*Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.*

**III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE**

*Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.*

Výslednou známku nejvíce ovlivnila míra splnění zadání, kde student ne zcela demonstroval práci s daty přicházejícími v reálném čase. Dalším faktorem byla formální a jazyková úroveň.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

**Otázky:**

1. Jaké jsou výhody použití Apache Sparc v porovnání s Apache Hadoop MapReduce?
2. Vyzkoušel jste také jiné nástroje pro zpracování Big Data mimo rodinu Apache Hadoop a Apache Sparc?
3. Proč je v práci kombinováno několik programovacích Jazyků, jaký byl důvod?

Datum: 13.6.2022

Podpis: