

— Posudek oponenta diplomové práce —

Název: History and forecast weather and wave data processing
and visualization framework
Student: Bc. Lubomír Bucek
Vedoucí: Doc. Ing. Jiří Cajthaml, Ph.D.
Fakulta: FSv ČVUT v Praze
Katedra: Katedra geomatiky
Oponent: Ing. Martin Landa, Ph.D.
Pracoviště oponenta: Katedra geomatiky, FSv ČVUT v Praze

Zadání. (*Hodnocení náročnosti zadání závěrečné práce.*)

Zadání práce považuji za náročnější, vyžadující komplexní řešení problému. Pohybuje se od rešerše dostupných datových zdrojů a technologií až po implementaci vybraných komponent do produkčního prostředí. Drobná výtka směřuje k názvu práce, který působí šroubovaně — je přeplněn spojkami „and“.

Splnění zadání. (*Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena.*)

Výsledek implementační části lze rozdělit na části dvě. V první řadě se jedná o sadu skriptů pro stahování, zpracování a uložení historických dat o počasí a vlnách, a dále návrh přístupového rozhraní (tzv. REST API). Následuje sada skriptů pro stahování a zpracování rastrových souborů ve formátu GeoTIFF včetně jejich konfigurace v prostředí mapového serveru GeoServer. Podstatné pro hodnocení je úspěšné nasazení výsledků práce v reálném produkčním prostředí. Z tohoto pohledu lze zadání považovat za splněné.

Zvolený postup řešení. (*Posuďte, zda student zvolil správný postup nebo metody řešení.*)

Postup řešení je v podstatě pouze konstatován. Není dán téměř žádný prostor k diskuzi. Rešeršní část práce v podstatě neexistuje. V práci jsou popsány použité technologie, alternativy jsou zmíněny okrajově. Chybí jejich kritické zhodnocení. Sám autor v kapitole 2 zmiňuje, že volba databázového úložiště je pro návrh systému zásadní. Bohužel tímto konstatováním to končí. Podobně je tomu v kapitole 3. Autor kromě relevantních OGC webových standardů pouze uvádí použité technologie GeoServer a knihovnu GDAL. Tímto text degraduje z práce vědecké na technický popis. Náznaky rešerše se objevují až v kapitole 6 „Further remarks“.

Odborná úroveň. (*Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů. Posuďte též schopnost studenta vnímat řešenou problematiku v širších souvislostech a aplikovat inženýrský přístup při řešení.*)

Odbornou úroveň práce považuji za mírně nadprůměrnou. Student si musel osvojit znalost široké škály technologií od databázových a vyhledávacích úložišť typu ElasticSearch či MongoDB přes nastavení mapového serveru až po návrh sady skriptů pro stahování a zpracování zájmových dat v programovacím jazyce Python, a to navíc v produkčním prostředí. Úspěšné zvládnutí svědčí o schopnosti studenta zasadit řešenou problematiku do širších souvislostí.

Na několika místech se autor dopouští drobných terminologických přehmatů a nejasností. Například v kapitole 4 mluví o „API“, pravděpodobně se jedná ve skutečnosti o tzv. „REST API“. Další důležité pojmy jako např. NoSQL jsou uvedeny jakoby mimochodem a nejsou ani nijak vysvětleny.

Formální a jazyková úroveň, srozumitelnost práce. (*Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku práce a její celkovou srozumitelnost.*)

Práce je napsána v anglickém jazyce, což je vzhledem k jejímu zacílení přirozené. Po stránce typografické a jazykové nemám vážnějších námitek. Horší je to s její srozumitelností. Některé pojmy nejsou dostatečně vysvětleny. Na čtenáře je kladena jistá dávka odborné erudice anebo ochoty si chybějící útržky dohledat z dalších zdrojů. Jako příklad mohu uvést první větu kapitoly 2.1, ve které je zmíněno řešení Apache Lucene, které ale není nijak dále popsáno.

Výběr zdrojů, korektnost citací. (*Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Posuďte výběr pramenů. Ověřte, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi.*)

Seznam uvedených pramenů je poměrně bohatý, i když obsahuje celou řadu online zdrojů. V textu jsou zdroje řádně citovány. K porušení citační etiky dle mého názoru nedošlo, text působí autenticky. Po typografické stránce dopadl seznam literatury nejhůře. Navíc neodpovídá citačním zvyklostem.

Celkové hodnocení, otázky k obhajobě, návrh klasifikace.

Jako oponent práce mám následující dotazy:

1. V kapitole 4 zmiňujete, že Python skripty spouští externí programy (modul `subprocesses`) a jejich výsledky poté zpracovávají. Tento postup nesvědčí o dobrém návrhu, i když může být relativně běžný i v produkčním prostředí. Není ale v práci nijak argumentován. Pokud mají knihovny dostupné Python API, mělo by být použito. Jasná demonstrace je uvedena v kapitole 5, kde používáte konzolové nástroje knihovny GDAL namísto existujícího široce používaného GDAL Python API. Prosím o další objasnění.
2. V případě spojení GRIB souborů používáte unixový nástroj `cat`. I když je tento nástroj velmi mocný, zkoušel jste najít rozumnou Python alternativu?
3. Jsou opravdu pomocné skripty pro Bash nutné pro výslednou implementaci? Jejich nutnost není nijak argumentována, pouze jaksi konstatována. Nicméně v kapitole 6 se objevuje náznak argumentace. Můžete ji mírně rozvést?
4. Celé řešení je koncipováno jako sada skriptů poplatná procedurálnímu přístupu. Neuvážoval jste o objektově orientovaném kódu? Ten se obecně lépe udržuje, je čitelnější a většinou i snadněji rozšiřitelný. To ale samozřejmě závisí na kvalitě jeho návrhu.
5. V kapitole 5 se objevuje v souvislosti s GeoServerem zmínka o databázovém úložišti PostgreSQL. Na konci této kapitoly je dokonce uvedeno, že jde o doporučené řešení.

K čemu GeoServer databázi PostgreSQL konkrétně používá? Náznak je uveden na konci kapitoly 6.

6. V textu je bez jakékoli další rešerše uváděn mapový server GeoServer. Jaké vás vedly důvody pro jeho nasazení?

Úroveň práce snižuje předložený text, který ne zcela odpovídá požadavkům kladeným na diplomovou práci. Namísto vědecky orientované, kritické práce je předložen text, který má spíše povahu technické dokumentace. Výjimkou je z tohoto pohledu kapitola 6. Na druhou stranu je nutné ocenit implementační výsledek práce především z pohledu jeho komplexnosti a snahy dosažení jeho produkčního nasazení. Přes výše uvedené výhrady hodnotím předloženou závěrečnou práci klasifikačním stupněm

– **A (výborně)** –

V Praze dne 1. února 2019

.....
Ing. Martin Landa, Ph.D.
Fakulta stavební, ČVUT v Praze