

# Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

**Student:** Jana Ernekerová  
**Oponent práce:** Mgr. Rudolf Bohumil Blažek, Ph.D.  
**Název práce:** Webová demonstrace základních statistických výpočtů s využitím matematického software R a SAGE  
**Obor:** Web a multimédia (bakalářský)

**Datum vytvoření:** 6. 6. 2015

<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 5:</b>
<b>1. Náročnost a další komentář k zadání</b>	<b>1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání</b>
<b>Popis kritéria:</b> Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)	
<b>Komentář:</b> Zadání považuji za průměrně náročné, avšak při přísnější interpretaci by zadání mohlo být považováno za náročnější. Jedná se o hloubku porovnání efektivity různých řešení pro propojení webových aplikací s výpočetními a algebraickými systémy.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 4:</b>
<b>2. Splnění zadání</b>	<b>1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</b>
<b>Popis kritéria:</b> Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.	
<b>Komentář:</b> Zadání bylo splněno, snad jen testování vytvořených softwarových aplikací by mělo být podrobnější. Implementační část bohužel nezahrnuje instalaci serverových částí použitých řešení. Proto nebylo možné podrobně porovnat náročnost integrace obou řešení a měření jejich výpočetního výkonu. Použitím existujících cloudových řešení se porovnání omezilo pouze na náročnost integrace ve vyvinuté webové aplikaci, což je spíše porovnání náročnosti použití. Zadané porovnání efektivity by v případě nasazení serverové části na jednom počítači mohlo zahrnout i porovnání výpočetního výkonu (benchmarking), přidaných zpoždění a náročnosti na výpočetní zdroje. Takovéto podrobné porovnání není možné v případě různých serverů pro různá řešení. Tato výhrada ale souvisí s rozdílnou interpretací zadání. Práce proto formálně splňuje požadavky zadání.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 4:</b>
<b>3. Rozsah písemné zprávy</b>	<b>1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky</b>
<b>Popis kritéria:</b> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.	
<b>Komentář:</b> Práce je rozsahem vyhovuje požadavkům. Kapitoly 1 a 2 by mohly být spojeny do jedné kapitoly, kapitola 4 by měla být rozsáhlejší. Na druhou stranu, kapitola 3 s přehledem statistických pojmů a metod je možná zbytečně podrobná. Nepřináší nic nového a nevysvětluje dostatečně jasně, proč jsou tyto pojmy předkládány. V úvodu kapitoly by mělo být jasněji řečeno, které metody budou demonstrovány ve vytvořené aplikaci, a teprve pak přistoupit k jejich diskusi s tím, že by mohla být stručnější.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>4. Věcná a logická úroveň práce</b>	<b>90 (A)</b>
<b>Popis kritéria:</b> Posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře.	

#### Komentář:

Z věcného hlediska práce obsahuje některé nepřesnosti, nikoliv však závažné chyby. Např. na str. 9, ř. 2, autorka píše: ""Po každém volání funkce, OpenCPU ukončí (případně zabije) R proces..."". Toto však není správně, jedná se o chybný překlad z dokumentace [24]. R proces není nutně ukončen, někdy je pouze ""cleaned"", tedy jsou vymazány informace související s právě ukončeným zpracováním požadavku. Ukončování a opětné spouštění R procesu by mohlo mít negativní vliv na výkon cloudového nasazení OpenCPU. Podobně, na str. 13. v 5. řádce sekce 1.4.2 se uvádí, že cloudové nasazení je možné na Ubuntu, ale ve skutečnosti podporovaných Linuxových odnoží je více, včetně Fedora, SUSE, Debian a CentOS. Podobných nepřesností je v práci několik.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

### 5. Formální úroveň práce

85 (B)

Popis kritéria:

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 12/2014, článek 3.

#### Komentář:

Anglický abstrakt by zasloužil větší pozornost, případně kontrolu rodilým mluvčím anglického jazyka. Problémem jsou především použití členů a neobvyklá slovní spojení, např. "the integration", nebo "calculating algebraic systems" místo "computer algebra systems". Zbytek práce je z formálního hlediska na vysoké úrovni, obsahuje příliš mnoho překlepů (např. "doktoranského" na str. 8, ř. 1 sekce 1.3.2, "pojektu", posl. řádek stejného odstavce, či "elemntárních" na str. 25, odst. 2, ř. 3). Formálním problémem je odkazování na "skripta pana Jiřího Anděla" a "pana Jiřího Pavlíka" (str. 25, posl. věta 1. odst.). Vhodnější by bylo použít titulu "profesor", příp. žádného titulu, jako je tomu u anglické literatury ve stejné větě. Občas najdeme i drobné chyby ve vzorcích, např. str. 25, 3. řádek odspodu, kde v definici míry Q chybí závorky u míry P, nebo str. 26, kde v definici F(x) chybí specifikace, která x uvažujeme. Těchto chyb je více, jedná se ale o méně závažná přehlédnutí. Fonty symbolů P a E pro pravděpodobnost a střední hodnotu mají špatnou velikost.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

### 6. Práce se zdroji

90 (A)

Popis kritéria:

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. 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.

#### Komentář:

Práce se zdroji je na vysoké úrovni a odpovídá vědecké publikaci, až na několik výjimek, kde jsou přímo v textu použity odkazy URL (např. odkaz na dokumentaci OpenCPU na str. 13, sekce 1.4.2, ř. 7).

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

### 7. Hodnocení výsledků, publikační výstupy a ocenění

90 (A)

Popis kritéria:

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

#### Komentář:

Výsledkem práce je analýza různých webových rozhraní pro výpočetní systém R, popis nástrojů a služeb OpenCPU a Sage Cell Server, popis použitých statistických metod a webová aplikace, která umožňuje nasazení těchto statistických metod na webu. Kapitola o statistických metodách je zbytečně podrobná, ale kvalitně zpracovaná, až na nepřesnosti a chyby v některých vzorcích. Nedostatkem je, že chybí porovnání výpočetního výkonu, zpoždění a systémové náročnosti studovaných řešení (benchmarking a overhead). V případě bakalářské práce jsou tyto výsledky ale adekvátní.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

### 8. Komentář o využitelnosti výsledků

Popis kritéria:

Uvedte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uvedte možnosti využití výsledků ZP v praxi.

#### Komentář:

Porovnání studovaných řešení je prezentováno z pohledu uživatele a programátora webových aplikací. Bohužel neposkytuje informaci, který systém a za jakých podmínek je výkonnější, a jaké mají nároky na výpočetní zdroje. Výsledky práce lze tedy použít jako návod pro použití těchto řešení, nikoliv jako vodítko k rozhodnutí, které řešení je vhodnější k nasazení za různých podmínek.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

### 9. Otázky k obhajobě

Popis kritéria:

Uvedte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

*Otázky:*

1. Proč nebyly serverové části nasazeny lokálně? V tom případě by bylo možné provést přímé porovnání výpočetního výkonu a náročnosti na zdroje obou řešení a analyzovat i zpoždění způsobené integrací webového aplikačního serveru s použitými systémy pro počítačovou algebru.

2. Vysvětlete prosím podrobněji, proč nebylo možné využít výpočetní systém R pomocí služby Sage Cell Server. Vysvětlení na str. 43, sekce 4.2.1, ř. 7 naznačuje, že by bylo možné tento problém obejít v případě nasazení služby Sage Cell Server na lokálním počítači. Pak by se mohlo provést porovnání výkonu nasazení R s oběma řešeními (viz otázka 1).

*Hodnotící kritérium:*

*Způsob hodnocení - bodové hodnocení 0 až 100 bodů  
(známka A až F):*

**10. Celkové hodnocení**

90 (A)

*Popis kritéria:*

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nemusí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

*Text hodnocení:*

Práce poskytuje popis uvažovaných technologií a návod na jejich použití, včetně porovnání jejich předností z pohledu programátora webových aplikací. Část zadání porovnávací služby "z hlediska praktické použitelnosti, efektivity a jednoduchosti integrace" je dle mého názoru příliš stručná, protože nezahrnuje porovnání nároků na výpočetní zdroje serverů ani porovnání výpočetního výkonu z pohledu zpoždění způsobeného integrací webových technologií s výpočetními systémy. Výsledky jsou ale akceptovatelné pro bakalářskou ZP.

Podpis oponenta práce: