

# Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

**Student:** Bc. Stanislav Fífik  
**Oponent práce:** RNDr. Jakub Klímeck, Ph.D.  
**Název práce:** Recipes Detection and Recognition on the Web  
**Obor:** Webové a softwarové inženýrství (magisterský)

**Datum vytvoření:** 6. 5. 2015

<p><i>Hodnotící kritérium:</i></p> <p><b>1. Náročnost a další komentář k zadání</b></p>	<p><i>Způsob hodnocení - následující škálou 1 až 5:</i> 1=mimořádně náročné zadání, 2=náročnější zadání, <b>3=průměrně náročné zadání,</b> 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání</p>
<p><i>Popis kritéria:</i> 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.)</p> <p><i>Komentář:</i> Zadání hodnotím jako průměrně náročné</p>	
<p><i>Hodnotící kritérium:</i></p> <p><b>2. Splnění zadání</b></p>	<p><i>Způsob hodnocení - následující škálou 1 až 4:</i> 1=zadání splněno, <b>2=zadání splněno s menšími výhradami,</b> 3=zadání splněno s většími výhradami, 4=zadání nesplněno</p>
<p><i>Popis kritéria:</i> 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.</p> <p><i>Komentář:</i> Zadání uvádí požadavek na implementaci a testování různých přístupů ke kategorizaci textu a jeho částí. V práci se různé přístupy popisují, ale jejich komparativní testy na doméně receptů v práci nejsou. Použitý přístup je popsán přímočaře, neposuzuje různé varianty řešení.</p>	
<p><i>Hodnotící kritérium:</i></p> <p><b>3. Rozsah písemné zprávy</b></p>	<p><i>Způsob hodnocení - následující škálou 1 až 4:</i> 1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, <b>3=splňuje požadavky s většími výhradami,</b> 4=nesplňuje požadavky</p>
<p><i>Popis kritéria:</i> Porovnejte rozsah předložené písemné zprávy s požadovaným rozsahem, viz Směrnice děkana č. 9/2011, článek 3. Pro hodnocení ZP je také důležité, zda všechny části písemné zprávy jsou informačně bohaté a pro práci nezbytné. Text ZP by neměl obsahovat zbytečné části.</p> <p><i>Komentář:</i> Část věnovaná popisu NLP technik a technik pro programování v JavaScriptu je rozsáhlejší, část věnovaná samotné implementaci pak chudší. Zcela chybí programátorská dokumentace knihovny - Kapitola 5 je pouze textový popis toho, co knihovna umí za algoritmy a techniky. V práci je odkaz na GitHub <a href="https://github.com/Ficik/recipes">https://github.com/Ficik/recipes</a> kde by měla dokumentace být, ale tam jsem ji v době psaní posudku nenašel. Stejný nedostatek má i část věnovaná demo aplikaci, ta však nebyla součástí zadání, takže tento fakt nehodnotím. V části 5.10.4 se mluví o podpoře RDF a SPARQL, ale není vůbec jasné k čemu se tyto techniky používají, dále v práci se o nich již nemluví.</p>	
<p><i>Hodnotící kritérium:</i></p> <p><b>4. Věcná a logická úroveň práce</b></p>	<p><i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i> 50 (E)</p>
<p><i>Popis kritéria:</i> 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.</p> <p><i>Komentář:</i> V práci chybí strukturovaná programátorská dokumentace knihovny. Kapitola 4 má pouze jednu podkapitolu a zřejmě by šla sloučit s kapitolou 5. Kapitola 9 o testování obsahuje pouze stručný slovní popis testování, výsledky by to chtělo prezentovat systematictěji. 2 ze 4 obrázků v této kapitole nejsou z textu odkazovány.</p>	
<p><i>Hodnotící kritérium:</i></p> <p><b>5. Formální úroveň práce</b></p>	<p><i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i> 50 (E)</p>
<p><i>Popis kritéria:</i> Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 9/2011, článek 3.</p>	

### Komentář:

Jazyková úroveň angličtiny je nízká a práce se pak těžko čte. Typograficky je práce v pořádku, formální zápisy až na vzorečky v sekci 5.7. práce neobsahuje.

Hodnotící kritérium:

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

## 6. Práce se zdroji

10 (F)

### 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 uvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

### Komentář:

V práci se volně mísí popis toho, co student naprogramoval a toho, co převzal. Toto rozdělení bych čekal přehledně shrnuté např. ve formě tabulky.

Práce obsahuje 20 referencí v bibliografii.

Z textu ale vedou odkazy jen na 5 zdrojů: [7], [10], [12], [13], [18]

Zdroje [1] - [6], [8], [9], [11], [14], [15], [16], [17], [19], [20] jsou tedy v práci nepoužité a tedy zbytečné.

Hodnotící kritérium:

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

## 7. Hodnocení výsledků, publikační výstupy a ocenění

40 (F)

### 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:

Bookmarklet mi v Chrome pro Windows Version 42.0.2311.135 m (64-bit) nefungoval, nicméně nebyl součástí zadání, takže nehodnotím. Otestoval jsem tedy aplikaci na receptu [http://ona.idnes.cz/kruti-platky-na-tymianu-se-sladkym-hraskem-f26-/recepty.aspx?c=A120804\\_002056\\_recepty\\_jid](http://ona.idnes.cz/kruti-platky-na-tymianu-se-sladkym-hraskem-f26-/recepty.aspx?c=A120804_002056_recepty_jid) a výsledek obsahoval jisté chyby, například čerstvý mletý pepř byl naparsován jako "ml čerstvý etý pepř", kroky postupu byly očíslovány jinak a původní čísla v textu zůstala, některé ingredience byly vynechány a pak vypsány do prvního kroku receptu.

Při testování na receptu z sRecepty.cz <http://www.srecepty.cz/koktejl-s-citronovou-stavou>, na kterém byl přístup dle práce naučen a měl by tedy podávat dobré výsledky, obsahoval výsledek opět chyby, například místo 0,3l mléka bylo ve výsledku 3l mléka.

Výsledek také obsahuje chyby z hlediska formátu Microdata - každý krok postupu je uveden jako "recipeInstructions" přičemž tato vlastnost má být pro recept uvedena jednou a obsahovat všechny instrukce.

Také bych čekal pokus o naparsování dalších informací kromě postupu a ingrediencí, jako třeba dobu přípravy apod.

Vzhledem k tomu že se jedná o diplomovou práci bych čekal lepší výsledky, práce odpovídá spíše bakalářské práci.

Licenční podmínky převzatých částí kódu v práci nejsou řešeny.

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:

Výsledky práce nové poznatky nepřinášejí ale poskytují základ pro další zkoumání.

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 odřádkami).

### Otázky:

1. Které části stacku jste implementoval a které převzal?
2. V zadání je napsáno že váš přístup má být rozšiřitelný i mimo doménu receptů. Ověřoval jste si na nějaké jiné doméně že tomu tak je?
3. U kolika receptů jste ručně kontroloval správný výstup a jaká byla úspěšnost? Kapitola 9 je velmi stručná a obsahuje pouze výsledky automatizovaných testů.

Hodnotící kritérium:

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

## 10. Celkové hodnocení

45 (F)

### Popis kritéria:

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

*Text hodnocení:*

Práce působí uspěchaným dojmem, jazyková úroveň je nízká a velkým problémem je uvedení zdrojů které v práci nejsou odkazovány, čímž se počet použitých zdrojů dostal na 5 což je na diplomovou práci nedostatečné. Chybí mi programátorská dokumentace vyvinuté knihovny. Výsledek otestovaný na náhodně vybraných českých receptech je nepřesvědčivý - obsahuje chyby které by ani mnohem primitivnější přístup pomocí parsování regulárními výrazy neobsahoval. Část úvodu a popisu použité teorie zabírá větší část práce, samotný přínos je pak na diplomovou práci stručný, stejně jako celá práce celkově, odpovídá tak rozsahem i obsahem práci bakalářské.

Podpis oponenta práce: