



# Posudek oponenta závěrečné práce

**Student:** Jindřich Žák  
**Oponent práce:** doc. RNDr. Tomáš Skopal, Ph.D.  
**Název práce:** Souvislosti a doporučování obsahu na webu  
**Obor:** Webové a softwarové inženýrství

**Datum vytvoření:** 6. 6. 2019

<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení – následující škálou 1 až 4:</b>
<b>1. 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 dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.	
<b>Komentář:</b> Práce se zabývá doporučováním příspěvků v rámci webu FIT ČVUT. Autor analyzoval stávající stav systému (postavený na CMS Drupal), uživatelské potřeby a technologie. Poté navrhl řešení doporučování příspěvků na základě modelu personalizovaného PageRanku, které implementoval jako modul do Drupalu. V rámci implementace a s ohledem na doporučování pomocí PageRanku autor zvolil grafovou databázi Neo4j a dotazy v jazyce Cypher. Implementovaný modul byl otestován a odladěn na lokální instalaci fakultního webu.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>2. Písemná část práce</b>	<b>80 (B)</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. Dále 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. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.	
<b>Komentář:</b> Písemná část práce je zpracována pečlivě, dělení na kapitoly je standardní a text srozumitelný.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>3. Nepísemná část, přílohy</b>	<b>80 (B)</b>
<b>Popis kritéria:</b> Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů	
<b>Komentář:</b> Na příloženém USB je vedle souborů diplomové práce také adresář se zdrojovými kódy k doporučovacím modulu (konfigurační soubory a skripty). Rozsah kódu je poměrně malý.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>4. Hodnocení výsledků, jejich využitelnost</b>	<b>85 (B)</b>
<b>Popis kritéria:</b> Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.	

**Komentář:**

Student navrhl a implementoval modul pro doporučování příspěvků na fakulním webu FIT, což je jistě přínosem v procesu evoluce webu FIT. Na druhou stranu zvolený model personalizovaného PageRanku zohledňuje pouze grafové vztahy mezi příspěvků (články), což je informace již nativně poskytovaná samotnou strukturou webu. Model grafu sice uvažuje více typů hran, nicméně všechny tyto hrany se předpokládají jako manuálně zanesené operátorem (resp. adminem). Zcela chybí doporučování na základě porovnávání obsahu nebo dokonce kolaborativního filtrování (na základě podobnosti chování více uživatelů). Důvodem k tomuto konzervativnímu přístupu byla pravděpodobně snaha o zajištění maximální konzistence - doporučovat se budou hlavně příspěvky v topologicky blízkém okolí (potomci/rodiče stránky), což se ale dá implementovat už teď právě proto, že tato explicitní informace už v systému je. Jinými slovy, doporučovací neuvažuje žádnou novou hranu na základě např. fulltextové podobnosti, ačkoliv v databázovém modelu tato možnost je (hrana typu souvislost).

*Hodnotící kritérium:*

*Způsob hodnocení – nehodnotí se*

## 5. Otázky k obhajobě

*Popis kritéria:*

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

**Otázky:**

1. Proč se v doporučovacím modulu nezrealizovalo content-based anebo kolaborativní filtrování?
2. V sekci 3.3.3. Transformace grafu pro PageRank se píše, že z důvodu procházení grafu oběma směry se orientovaný graf dezorientoval (ke každé orientované hraně se přidala opačná). To ale zcela popírá princip PageRanku, který právě rozlišuje uzly, na které se hojně referuje od těch, na které se nereferuje. Prosím vysvětlete legitimitu této transformace.

*Hodnotící kritérium:*

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

## 6. Celkové hodnocení

85 (B)

*Popis kritéria:*

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.

**Text hodnocení:**

Student navrhl, implementoval a otestoval modul pro doporučování příspěvků na fakulním webu FIT ČVUT. Ačkoliv mám výše zmíněné výhrady, doporučuji práci k obhajobě.

Podpis oponenta práce: