



# Posudek oponenta závěrečné práce

**Oponent práce:** prof. Dr. Ing. Petr Kroha, CSc.  
**Student:** Lukáš Rynt  
**Název práce:** Dvojjazyčné vyhledávání v dokumentech  
**Obor / specializace:** Znalostní inženýrství  
**Vytvořeno dne:** 28. května 2022

## Hodnotící kritéria

### 1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Práce je velmi pěkná. Student prokázal, že chápe problémy, které řeší. Rozsah je nadprůměrný - 54 stran.

### 2. Písemná část práce

90/100 (A)

Práce je výborně napsaná, použité zdroje (34) jsou správně citovány.

K obsahu mám některé připomínky, které lze označit jako drobné.

1. Str. 1 - "zpracování informací (anglicky information retrieval)" - Information retrieval se zabývá metodami, které naleznou dokumenty, ve kterých se vyskytují zadaní klíčová slova. Nejedná se tedy o zpracování informací, nicméně metody vyvinuté pro information retrieval se používají v mnoha dalších oborech souvisejících s computational linguistics, text mining a machine translation.

2. Str. 1 - Využití právnických textů EU pro trénování dvojjazyčného vyhledávání je odvážná myšlenka. Sice se nezabývá srovnáváním právních systémů, ale vím, že např. britské právo se opírá o rozhodnutí soudů v minulosti na rozdíl od německého i českého práva. Británie sice už není členem EU, ale ten problém se může objevovat i v jiných právních systémech. Lze si představit i situaci, kdy některé entity se vyskytují v právních systémech některých členských zemí, ale v jiných nikoli. Na druhé straně je mi známo, že firma Linguee (DeepL) také použila EU předpisy.

2. Str. 14 - "... dotaz nebude vyhovovat všem dotazům" - ???

### 3. Nepísemná část, přílohy

80 /100 (B)

Autor implementoval navržený systém z poloviny v jazyce Ruby a z poloviny v jazyce Python. Elegantnější by bylo provést implementaci jen v Pythonu. Odůvodnil to tím, že Ruby dobře ovládá, ale Ruby nemá potřebné linguistické a AI moduly.

Práce neobsahuje diagramy popisující model jako základ implementace. Model MVC je zmíněn, ale diagram tříd není uveden. Třídy jsou zmíněny v textu.

Práce neobsahuje popis testování.

### 4. Hodnocení výsledků, jejich využitelnost

90 /100 (A)

Prototyp funguje, ale na jeho nasazení je příliš brzy. Problematika je velmi složitá a bude vyžadovat podrobnější testování a analýzu výsledků, než je v této bakalářské práci uvedeno.

## Celkové hodnocení

90 /100 (A)

Dvojjazyčné vyhledávání, do kterého se autor pustil, je složitý, otevřený problém. Překladaelé často překládají text podle jeho smyslu, ale nikoli podle jednotlivých slov.

Souvisí to se sémantikou, fantazií a uměním překladu.

Příklady:

Název britského románu *Changing places* byl do češtiny přeložen jako *Hostující profesoři*.

Název britského románu *Small world* byl do němčiny přeložen jako *Schnitzeljagd*.

K tomu je třeba dodat, že autor německého překladu tím nemyslí druh řízku, jak by se třeba znalec německých jídelních lístků mohl domnívat,

ale dětskou hru, kdy jedna parta dětí umísťuje v lese na větvičky papírové fáborky (*Papierschnitzel*) a druhá skupina dětí tu první podle těch fáborků "stopuje - honí" - (*Jagd* je německy hon).

Takže pokud by se někdo pokusil čistě na základě německého názvu najít anglický originál, tak se mu to asi nepovede. Německý překladael tím názvem ale přesně popsal, o co v románu jde.

## Otázky k obhajobě

Nemám otázky

## **Instrukce**

### **Splnění zadání**

Posudte, 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, posudte 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.

### **Písemná část práce**

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, 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. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, 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.

### **Nepísemná část, přílohy**

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.

### **Hodnocení výsledků, jejich využitelnost**

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.

### **Celkové hodnocení**

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.