

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Borek Požár
Název práce Dialogový systém pro hlasové vytáčení
Rok odevzdání 2021
Studijní program Informatika
Studijní obor Obecná informatika

Autor posudku Mgr. Michal Novák, Ph.D. Oponent
Pracoviště Ústav formální a aplikované lingvistiky

K celé práci

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Obtížnost zadání		X		
Splnění zadání		X		
Rozsah práce <i>... textová i implementační část, zohlednění náročnosti</i>		X		
<p>Celkově práce splňuje zadání. Výjimkou je chybějící integrace skloňování českých jmen a vyhodnocení systému na testovací sadě.</p> <p>I když by případné zahrnutí skloňování mohlo pomoci u kratších jmen, implementovaný fuzzy matching ve většině případů postačuje.</p> <p>Co se týče testování aplikace, porovnávací komponenta zahrnuje unit testy, které ověřují funkčnost jednotlivých procedur na několika příkladech s použitím “hezkého” seznamu kontaktů, kde každý kontakt obsahuje jméno a příjmení. V praxi jsou ale seznamy kontaktů mnohem špinavější, u některých kontaktů bývá vyplněno jenom křestní jméno, přezdívkou nebo role (např. “máma”, “táta”). Vyhodnocování úspěšnosti na rozsáhlejší datové sadě, zejména během vývoje, by lépe odhalilo slabší místa porovnávací komponenty, které popisují níže. I když by tento postup vedl k vyšší kvalitě vyhledávání v kontaktech, u bakalářské práce neočekávám rigorózní metodiku provádění experimentů, a proto nesnižuji hodnocení u položky “Splnění zadání”.</p>				

Textová část práce

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Formální úprava <i>... jazyková úroveň, typografická úroveň, citace</i>	X			
Struktura textu <i>... kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>		X		
Analýza		X		
Vývojová dokumentace		X		
Uživatelská dokumentace		X		
<p>Práce je velmi dobře strukturována a její text se čte příjemně. Nad úrovní běžné bakalářské práce je citování většího množství související literatury.</p> <p>Vývojová dokumentace sice není centralizovaná, ale je k nalezení ve formě souborů README.md u každé z komponent.</p> <p>Uživatelská dokumentace je netradiční ve formě ukázkového videa, které demonstruje použití aplikace. To v tomto případě postačuje, jelikož použití je přímočaré. Očekával bych však jasnější popis toho, kam se odesílají údaje z telefonu, např. že se při spuštění aplikace odešlou jména ze seznamu všech kontaktů na servery IBM. Fakt, že spolu se jmény se neodesílají telefonní čísla, je zmíněn jenom v dotazníku pro uživatelské testování.</p>				

Implementační část práce

lepší OK horší nevyhovuje

Kvalita návrhu ... architektura, struktury a algoritmy, použité technologie			X	
Kvalita zpracování ... jmenné konvence, formátování, komentáře, testování	X			
Stabilita implementace		X		

Aplikace šikovně propojuje několik externích služeb (Google STT/TTS, IBM Watson Asistent) za použití jazyků Kotlin a Python. Kód je velmi dobře okomentován a porovnávací komponenta je ošetřena unit testy.

Z mého pohledu je největším vlastním přínosem porovnávací komponenta, která zabezpečuje vyhledání kontaktu v seznamu kontaktů. I když je vyhledávání sofistikované, v praxi mi uspokojivě fungovalo u kontaktů ve formě “jméno příjmení”, ale již hůř u kontaktů sestávajících z jednoho slova.

Jedním z důvodů je, že navržená metoda entity vyhledává na základě předem definovaného seznamu jmen a příjmení. Ačkoliv je dlouhý, na familiární jména, přezdívky nebo role nepostačuje. Na základě kontextu jsou entity identifikovány jenom ve specifickém případě, když už bylo sousedící slovo identifikováno jako entita.

Problémem je také absence přísnějších kritérií párování dotazu s kontaktem v případě, kdy systém nabízí výběr z několika dříve nalezených kontaktů. Např. po dotazu na jméno “Nelka” mě systém požádal o výběr z kontaktů “Peta ABC”, “Peta XYZ”, “Aneta suseda”, “Nelka” a “Tea EFG”. Následně se mi jej však nijak nepodařilo přesvědčit, aby vybral právě kontakt “Nelka”, který se dokonale shodoval s dotazem (rovněž ve výstupu STT komponenty). Tohle chování se opakovalo rovněž u jiného dotazu.

Uznávám, že jednou z příčin horší funkcionality na mém telefonu může být i to, že můj seznam kontaktů obsahuje z části slovenské jména a výrazy.

Celkové hodnocení Velmi dobře, spíše lepší
Práci navrhuji na zvláštní ocenění Ne

Datum: 23. srpna 2021

Podpis