

# Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

<b>Autor práce</b>	Tadeáš Tomiška	
<b>Název práce</b>	Vývoj mobilní aplikace a generátoru otázek pro hru Desítka	
<b>Rok odevzdání</b>	2023	
<b>Studijní program</b>	Informatika	
<b>Studijní obor</b>	Informatika se specializací Programování a vývoj software	
<b>Autor posudku</b>	Mgr. Rudolf Rosa, Ph.D.	Oponent
<b>Pracoviště</b>	Ústav formální a aplikované lingvistiky	

## K celé práci

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Obtížnost zadání		X	X	
Splnění zadání			X	
Rozsah práce <i>... textová i implementační část, zohlednění náročnosti</i>				X
<p>Zadání je spíše jednodušší (a nedostatečné v otázce seznamu odborné literatury), navíc práce byla zadána týden před odevzdáním, tedy v době, kdy pravděpodobně byla práce prakticky hotová. Je proto s podivem, že ani tak se studentovi nepodařilo zadání zcela splnit – automatické generování otázek je dořešeno pouze částečně, kontrola obtížnosti otázek není řešena vůbec, a manuální potvrzení správnosti a relevantnosti otázky uživatelem není implementováno v kódu ale spočívá v manuální editaci vygenerovaného souboru s otázkami.</p> <p>Celkový rozsah a úroveň práce jsou nedostatečné, práce odpovídá spíše zápočtovému programu či prototypu bakalářské práce (na úrovni kvalitnějšího ročníkového projektu), mnoho věcí je řešeno jen základním způsobem bez podrobnějšího zkoumání a doladění do kvalitní podoby, žádná z částí práce není nějak výrazně náročná či zajímavá. Text je velmi stručný a mnoho věcí není dostatečně popsáno a odůvodněno, i tak zároveň text obsahuje mnoho nepřesností, chyb a nekonzistencí, a to jak vnitřních, tak vůči odevzdané aplikaci (například screenshoty odpovídají odevzdanému programu jen částečně).</p> <p>Ačkoliv má práce i komplexnější nedostatky, mnoho problémů lze snadno odhalit a odstranit řádným testováním a kontrolou práce v relativně krátkém čase. Práce tedy nese známky kvapného dokončení bez řádné konzultace s vedoucím.</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	

Textová část práce je nedostatečná. Přestože je velmi stručná (25 stran textu), tak obsahuje mnohé nepřesnosti a nekonzistence. Zároveň dostatečně nepopisuje a nevysvětluje mnohé části řešení.

Text popisuje velmi stručně zvolené řešení, jen příležitostně diskutuje proč bylo zvoleno právě toto řešení a jaké má výhody a nevýhody a jaké jsou další možnosti – diskutují se méně podstatné otázky (jako volba programovacího jazyka), ale nediskutuje se například metoda extrakce dat, filtrování dat, generování otázek... To vše vzbuzuje pocit, že student si rozmyslel pouze základní věci, ale na rozmyšlení těch pokročilejších částí už neměl čas, takže prostě zvolil první řešení, které ho napadlo, ani se nad ním příliš nezamyslel a ani jej příliš neotestoval a neodladil a jen ho použil a stručně popsal. Dojem z textu je v tomto smyslu konzistentní s dojmem z vlastní aplikace.

Některé části nejsou popsány téměř vůbec, například jak se parsují a extrahují data z Wikipedie, jak se z nich generují otázky a odpovědi, jak je implementováno GUI. Práce také zcela postrádá aspoň elementární vyhodnocení kvality řešení, tj. nějakou analýzu, do jaké míry se daří extrahovat vhodné otázky a odpovědi; toto by šlo dobře dělat kvantitativně, například na vzorku 20 extrahovaných otázek vyhodnotit několik kritérií jako smysluplnost otázky, správnost odpovědi, vhodnost otázky pro účely hry, a podobně. Práce ani neobsahuje reálné ukázky extrahovaných dat a generovaných otázek a odpovědí (otázky ve screenshotech zjevně neodpovídají skutečným otázkám v přiložených JSON souborech). Jen jednou větou se zmiňuje i generování otázek na seřazení, toto by si zasloužila výrazně více textu (jak se na toto extrahují data, jak se to generuje, jak se to pak používá v GUI...)

Nepřesnosti a překlepy například hovoří o tom, že pro komunikaci se využívá wifi (to je nesmysl, využívá se prostě síťová komunikace, není zde nic specifického ohledně wifi), že otázky se generují pomocí webových stránek (otázky se generují pomocí aplikace, s využitím webových stránek jako zdroje), hovoří se o využití NLP jako protiváhy k využití strojového učení (NLP dnes typicky využívá strojové učení, navíc práce prakticky žádné NLP nevyužívá, jedná se prostě a pravidlové zpracování), hovoří se o využití dat z Wikipedie namísto Wikidat (ale užívané infoboxy jsou často generované právě z Wikidat, tj. místo parsování Wikipedie by šlo využít Wikidata přímo). Zdůvodnění některých rozhodnutí je také nevhodné (například argumentace velikostí aplikace jako důvodu pro uložení otázek na serveru, když samotných otázek je poměrně malé množství). Obecně by bylo vhodné v práci méně fabulovat na základě vlastních dojmů a více se opírat o odborné zdroje (práce de facto žádné odborné zdroje necituje, odkazuje pouze na 7 webových stránek).

## Implementační část práce

lepší    OK    horší    nevyhovuje

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

Řešení je velmi základní, neobsahuje prakticky žádné zajímavé prvky. Implementace GUI klientské aplikace se dle screenshotů zdá být dobré, ale v práci není vůbec popsáno.

Serverová část má celkem rozumně navrženou strukturu a fungování, ale mám výhrady k souborovým formátům a komunikačnímu protokolu – autor kombinuje různé standardní postupy a formáty (JSON, CSV) s vlastními postupy formáty (např. vlastní komunikační protokol přímo přes Sockets); vhodnější by bylo použít jeden standardní postup a formát (např. JSON) a toho se držet, takto se nepodařilo v práci dostatečně popsat použité formáty a protokoly a navíc se zdá že hrozí problémy např. s neošetřenými speciálními znaky (což standardní postupy typicky řeší inherentně). Zvolený způsob komunikace také snadno umožňuje podvádět, neboť klient společně s otázkou dostává i správnou odpověď (toto se v práci ani nijak nezmiňuje, autor si tento nedostatek tedy zřejmě neuvědomuje).

Část generování otázek je z větší části neexistující. Otázky a odpovědi se těží z Wikipedie a základním způsobem filtrují, ale dále se již nezpracovávají a předkládají se uživateli v syrové podobě, či ještě hůře s odstraněnou diakritikou a mezerami. Data v podobě např. “Česká republika – hlavní město – Praha” se tak transformují do otázky “hlavni\_mesto”, což není právě uživatelsky přívětivé. V práci se toto ani nijak dále nediskutuje, a ze zadání i abstraktu i screenshotů se celou dobu zdá, že se zde provádí nějaké chytřejší generování otázek (mluví se dokonce o NLP, ale žádné znaky NLP jsem v práci nenašel). Přitom by stačil na první řešení i poměrně jednoduchý postup, např. prefixovat otázku “Co je”, tedy bychom například získali otázku “Co je hlavní město?”, “Co je rok vzniku?”... Toto by pak už dále šlo přirozeně vylepšovat pomocí základních NLP postupů, například určením kategorie odpovědi a rozrůzněním otázek (věc, člověk, vlastnost – co je, kdo je, jaké je), prací s jednotným/množným číslem, a podobně. Doporučit také lze vycházet z nějaké existující odborné literatury z oblasti Question Answering a generování datasetů pro tuto úlohu; práce z žádné odborné literatury nevychází. Vzhledem k tomu, že dle zadání mělo zjevně jít o podstatnou část práce, která by jí dodala na zajímavosti a novosti, je praktická absence této části závažným nedostatkem.

Není také například nijak chytře řešena kontrola správnosti odpovědi na otázku, vyžaduje se prostě přesná stringová shoda včetně diakritiky (a to i u exotických znaků s diakritikou mimo možnosti české klávesnice). Nijak se ani nevyužívá například fakt, že některé názvy mají více možných variant, což používaná Wikipedie přímo zachycuje a umožňuje vytěžít; Wikipedie je zde však používána bez hlubšího pochopení její struktury a fungování. Dokonce ani nejsou v odpovědích odstraňovány upřesnitka v závorkách, takže například u otázky na autory uměleckých děl je socha Davida uvedena ne jako “David”, ale tak jak se jmenuje článek na Wikipedii, tedy “David (Michelangelo)”, čímž se stává správné zodpovězení autorství směšných. Toto vše ukazuje na zcela nedostatečné testování a ladění aplikace, neboť mnohé tyto problémy by bylo skutečně triviální nalézt a odstranit.

Mobilní klientskou aplikaci se mi nepodařilo spustit, je totiž extrémně uživatelsky nepřívětivá. Spuštění mobilní aplikace vyžaduje úpravu zdrojového kódu (viz pokyny na straně 19 v práci) a instalaci a konfiguraci několika nástrojů. Přitom by jistě bylo možné konfiguraci provádět přímo v aplikaci a ne úpravou zdrojáku, a tedy by bylo možné odevzdat už zkompilevanou aplikaci připravenou k nahrání do mobilního telefonu standardním způsobem.

Serverovou aplikaci se mi po určitých obtížích podařilo spustit (práce např. neuvádí, jaké nástroje a v jaké verzi jsou k tomu potřeba), pokusil jsem se s ní komunikovat pomocí telnetu (kvůli neschopnosti spustit plnou klientskou aplikaci) dle popsaného komunikačního protokolu, ale popis protokolu v práci je nedostatečný a tedy se mi nepodařilo tímto způsobem vlastní hru otestovat.

Autor uvádí, že se nepodařilo serverovou část spustit na veřejně přístupném serveru, neuvádí však, zda se o to alespoň pokusil. Věřím, že ve spolupráci s vedoucím by neměl být problém spustit server na nějakém veřejně přístupném počítači na MFF, nejde (v rozporu s tím, co autor tvrdí) o nijak hardwarově náročnou aplikaci.

---

<b>Celkové hodnocení</b>	Neprospěl(a)
<b>Práci navrhuji na zvláštní ocenění</b>	Ne

Datum

Podpis