

Supervisor's Opinion of Final Thesis

Název práce	Design and Implementation of an End-to-End Speech Assistant
Jméno autora	Felix Staudigl
Typ práce	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Thesis reviewer	Sascha Schade Dipl.-Ing. Dipl.-Ing. (RWTH Aachen University) Head of Software / Robotacist Robotise GmbH Germany

Assignment (very challenging)

The student has discussed state-of-the-art approaches for a newly developed end-to-end speech assistant in a very specific use case with a very demanding requirement of full offline-capability and implemented an excellent solution.

Satisfaction of assignment (overfulfilled [on basis of shifted focus])

Literature research was very extensive as it provided a good basis for a solid and state-of-the-art implementation. Carefully considering newest research results raised in importance during presenting intermediate results. The design of the system was taken very seriously as the code is running in production and facing real customers in a very critical business environment. The implementation was done accordingly to the latest software engineering principles, including modern microservices and continuous integration and continuous delivery pipelines. The required confidentiality of the work the lack of access to real world test samples (due to data protection law, et cetera) made the task of comparison and evaluation difficult to impossible to achieve. Evaluation is an ongoing process and was also partially removed from the thesis by company's request of confidentiality. The work presented to the company has proven to have a great business value.

Method of conception (A fulfilled with very minor objections)

The student chose the Design Science Research Methodology (DSRM) which is very appropriate for this thesis and was carefully followed. The requirements for the end-to-end speech assistant, and its implications to the technical possibilities, were intensively discussed and set accordingly in strong collaboration with the business development department. The offline requirement is a clear requirement from the use case which made access to other solutions (Google, Amazon, IBM) impossible. The review of the literature, the design and the implementation were done in due time and met the agreed deadlines. The part of comparison and evaluation was conducted but stripped due to confidentiality requirements.

Technical level (A+ outstanding)

As research literature shows, speech processing and synthesis are very challenging tasks. Gaining an in-depth understanding of broad technologies in short time and presenting results in a concise manner was done outstanding well. The approach of a system consisting of microservices proved to be very successful and was a challenging task. The fully adhered software development principles (test-driven development, agile sprints, feature branches, merging) are on a very professional level and brought software development to a next level

at Robotise GmbH. The developed system outperforms in the given use-case any other systems.

Formal and language level, scope of thesis (A excellent)

The work presented is well structured and uses appropriate graphs to illustrate key principles. Language is very professional and concise.

Selection of sources, citation correctness (A excellent)

Extensive literature was researched, clustered and referenced appropriately. All relevant resources were used. References were given correctly and credit was given when work was incorporated into the thesis and work.

Additional commentary and evaluation

The extensive research made it possible to develop a solution for the company which outperforms all other possible solutions. The developed work has so far proved to be easily to maintain and has a high quality and set new standards at the company. As the work has become a crucial part of proprietary IP for the company, confidentiality reasons limited the amount of work which was possible to be presented in the report.

The work and thesis present an outstanding result to the company. I evaluate the thesis with grade **A+ outstanding**.

Dipl.-Ing. Dipl.-Ing. Sascha Schade

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Design and Implementation of an End-to-End Speech Assistant
Jméno autora:	Felix Staudigl
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra řídicí techniky
Oponent práce:	Ing. Jiří Spilka Ph.D.
Pracoviště oponenta práce:	CIIRC, ČVUT v Praze

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
<p>The thesis provides a comprehensive description of an end-to-end speech assistant implemented for a service robot use case. Several objectives for the assistant are defined: security, privacy, scalability, offline, user experience, and domain specific. The solution is composed of three building blocks: automatic speech recognition, intent recognition, and text to speech. The thesis provides a general overview of available frameworks and algorithms for each block. The selection of frameworks/algorithms is well reasoned and documented. The whole platform is assembled from the available frameworks, where each building block is dockerized. I consider the thesis assignment as challenging as it requires an understanding of the broad discipline (from speech recognition to speech synthesis) and needs to connect together different frameworks. Further, it is implemented using modern microservices pattern and utilize continuous integration and delivery making the development and deployment smooth and easy.</p>	

Splnění zadání	splněno s menšími výhradami
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
<p>The state of the art review, design, and implementation of the voice assistant were clearly completed. The last task was to define test cases, evaluation metric, and compare the implemented system with existing solutions. However, this task was only partially accomplished. The implemented system was not thoroughly evaluated (no test cases, no performance metric). Further, the comparison to Amazon Alexa is very brief. This is unfortunate since up to this task everything was described thoroughly and systematically.</p>	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
<p>The methodology is in my opinion correct. The objectives are clearly defined and reasoned. Only, the offline requirement is perhaps too strict and limits the number of available algorithms that can be considered. The selection of methods is great and thorough. Finally, the design using microservices architecture, dockerization of each service, and continuous integration and delivery is perfect.</p>	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
<p>The technical level of the thesis is excellent. The work uses state of the art algorithms/frameworks and unifies them into a single platform. Although, for the selected algorithms/frameworks, I would welcome a more detailed and technical description.</p>	

Formální a jazyková úroveň, rozsah práce	A - výborně
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
<p>The thesis is written in very nice English. It is easy to read and contains only a few typos.</p>	

Výběr zdrojů, korektnost citací

A - výborně

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

The references are, in my opinion, correct. From the plethora of available algorithms and frameworks, the author has selected relevant papers. I'm convinced that citation ethics was not violated.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Since the platform was developed for a company I was not able to review actual implementation due to confidentiality. I had asked for the access but was not able to gain it because of the short period between thesis submission and review. In that respect, the review is based on report only. The code quality/maintainability/documentation is not reviewed.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

The thesis focuses on the development of a voice assistant platform. It thoroughly describes existing algorithms and frameworks and assembles them together. The description of design and implementation is excellent. I firmly believe that it has required a large amount of work to understand the whole pipeline, select relevant algorithms, and connect them together. However, the evaluation of the quality and accuracy of the proposed system should be more detailed and comprehensive. Having no access to the implementation and without the detailed evaluation, it is difficult for me to judge the quality of the implemented system.

In that context, I evaluate the thesis with classification grade **C - dobře**.

Questions for discussion:

1. It is not clear how difficult is to write and deploy a dialogue? Is it sufficient to provide question/answer pairs?
2. How was the whole system tested?
3. One of the requirement is scalability of the proposed solution. The system will be deployed on a single robot where only one user can interact with the robot at the same time. What is meant by scalability under these conditions?

Datum: 2.6.2019

Podpis: