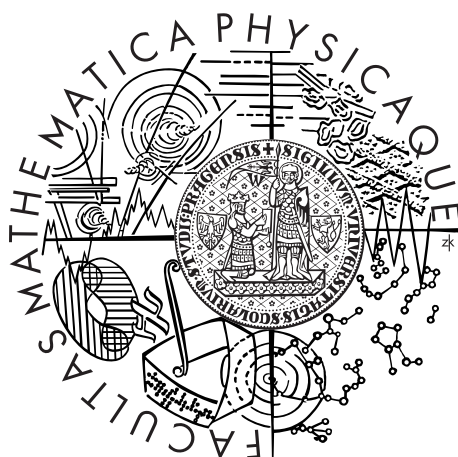Charles University in Prague

Faculty of Mathematics and Physics

MASTER THESIS

Ondřej Klejch

Development of a cloud platform for automatic speech recognition

Institute of Formal and Applied Linguistics

Supervisor of the master thesis:	Mgr. Ing. Filip Jurčíček Ph.D.
Study programme:	Informatics
Specialization:	Theoretical Computer Science

Prague 2015

Dedication.

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In .... date ....   signature of the author

Název práce: Development of a cloud platform for automatic speech recognition

Autor: Ondřej Klejch

Katedra: Ústav formální a aplikované lingvistiky

Vedoucí diplomové práce: Mgr. Ing. Filip Jurčíček Ph.D., Ústav formální a aplikované lingvistiky

Abstrakt:

Klíčová slova:

Title: Development of a cloud platform for automatic speech recognition

Author: Ondřej Klejch

Department: Institute of Formal and Applied Linguistics

Supervisor: Mgr. Ing. Filip Jurčíček Ph.D., Institute of Formal and Applied Linguistics

Abstract:

Keywords:

Contents

Introduction

The most natural form of people's communication is speech. In order to be able to talk with a computer, it is crucial to have a good Automatic Speech Recognition (ASR) system. On one hand, there are several open-source speech recognition toolkits that can be used but deployment of such toolkits requires substantial knowledge and therefore for common software developers it's not easy to use them. On the other hand, there are a few webservices that provide speech recognition as a service, yet these webservices don't solve all problems - either they are paid, closed-source or they are not customizable. The first goal of the thesis is to develop a cloud platform for ASR that is easy to use both from user's and maintainer's point of view.

Although accuracy of ASR systems is improving, these systems are still far from perfect. One of the reasons is that accuracy of ASR systems relies heavily on the amount of the training data and there isn't enough publicly available transcribed speech data. By providing free ASR webservice it is possible to collect vast amount of recordings. Consequently, the second goal of the thesis is to create an annotation interface so that recordings obtained by CloudASR platform can be annotated and given back to the community.

In the following text there will be described development and deployment of CloudASR platform and of its annotation interface. Chapter 1 introduces ... In Chapter 2 architecture of CloudASR is described. Annotation interface and theory related to obtaining of human transcriptions is presented in Chapter 3. Finally, Chapter 5 concludes this thesis. User manual and programmer manual can be found in the Attachments.

1. Title of the first chapter

1.1 Title of the first subchapter of the first chapter

1.2 Title of the second subchapter of the first chapter

2. Title of the second chapter

2.1 Title of the first subchapter of the second chapter

2.2 Title of the second subchapter of the second chapter

Conclusion

Future work

Bibliography

List of Tables

List of Abbreviations

Attachments