

Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

## BAKALÁŘSKÁ PRÁCE



Jan Milota

## Vývoj hlasově ovládaných webových her pomocí CloudASR

Ústav formální a aplikované lingvistiky

Vedoucí bakalářské práce: Mgr. Ing. Filip Jurčíček, Ph.D.

Studijní program: Informatika

Studijní obor: Obecná Informatika

Praha 2015

Děkuji panu doktorovi Jurčíčkovi za profesionální vedení této práce a za jeho neutuchající víru v mou schopnost samostatného vývoje.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V ..... dne .....

Podpis autora

Název práce: Vývoj hlasově ovládaných webových her pomocí CloudASR

Autor: Jan Milota

Katedra: Ústav formální a aplikované lingvistiky

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

Abstrakt: Cílem práce je navrhnout a vyvinout software pro výuku jazyků hrou za použití webových technologií a čerstvě vznikající CloudASR knihovny. Běžný uživatel provozuje interakci se svým prohlížečem skoro výhradně prostřednictvím myši a klávesnice. Díky softwaru, který tato práce reprezentuje, má nyní uživatel příležitost zabřednout do někdy ne úplně populární výuky jazyka i za pomoci svého hlasu. Což nabízí zmíněné výuce netušené možnosti, obzvláště stran uživatelské interaktivity. Důraz byl kladen na uživatelskou přívětivost, grafickou fidelitu a na kompetitivní aspekt výuky, využívajíc Facebookovou integraci a bodové hodnotící žebříčky.

Klíčová slova: automatické rozpoznávání řeči, ASR, hry, web, HTML5, Javascript

Title: Development of speech enabled web games using CloudASR

Author: Jan Milota

Department: Institute of Formal and Applied Linguistics

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

Abstract: The main goal of this thesis is to design and implement a piece of software for playful language learning, using web technologies and the fresh CloudASR library. A common user interacts with their web browser almost exclusively using a mouse and keyboard. Thanks to the software this thesis represents the user has an opportunity to delve into sometimes unpopular language learning process using his natural voice. This fact presents new and exciting possibilities, mainly regarding user interactivity. A lot of stress has been put to user friendliness, graphical fidelity and to the competitive aspect of language education, exploiting Facebook integration and point-scoring leader boards.

Keywords: automatic speech recognition, ASR, games, web, HTML5, Javascript

# Obsah

<b>Úvod</b>	<b>2</b>
<b>1 e-Learning jako způsob předávání informací</b>	<b>3</b>
1.1 Kapitola . . . . .	3
1.2 Kapitola . . . . .	3
<b>2 Automatické rozpoznávání řeči</b>	<b>4</b>
2.1 Kapitola . . . . .	4
2.2 Kapitola . . . . .	4
<b>3 Použité technologie</b>	<b>5</b>
3.1 Kapitola . . . . .	5
3.2 Kapitola . . . . .	5
<b>4 Implementace</b>	<b>6</b>
4.1 Kapitola . . . . .	6
4.2 Kapitola . . . . .	6
4.2.1 Ukázka L <sup>A</sup> T <sub>E</sub> Xu . . . . .	7
<b>Závěr</b>	<b>8</b>
<b>Seznam použité literatury</b>	<b>9</b>
<b>Seznam tabulek</b>	<b>10</b>
<b>Seznam použitých zkratek</b>	<b>11</b>
<b>Přílohy</b>	<b>12</b>

# Úvod

Základním komunikačním kanálem mezilidské interakce je mluvená řeč. Ačkoli trend v poslední době spíše směřuje k separaci této komunikace virtuálními médii, je mluvené slovo i nadále nenahraditelnou součástí socializace a výměny informací. Proč tedy nepřenést schopnost porozumět lidské mluvě i na stroje, jakožto právě na ona separující virtuální média? Současná technologie je k tomu jistě dostačující, výpočetní výkon lidstva roste po zběsilé křivce a technologické know-how po ještě zběsilejší.

Aby tento cíl mohl být dosažen, je kritické mít dobrý ASR (Automatic Speech Recognition) systém. Těchto se pohybuje v éteru mnoho, volně dosažitelných, rozličné kvality a spolehlivosti. Pro specifické využití v této práci byl zvolen systém CloudASR<sup>1</sup>, jehož autorem je kolega Ondřej Klejch.

TrogASR (Translating Online Game using Automatic Speech Recognition) je, jak možná název napovídá, online hra, zaměřující se na překlad slov a frází z jazyka do jazyka za využití automatizovaného rozpoznávání řeči.

Cílem této práce je návrh a implementace graficky přitažlivé hry pro webové prohlížeče, která poskytne uživatelům další, neotřelou, možnost, jak si vyzkoušet a zdokonalit své jazykové schopnosti. Tato hra pak bude sloužit jako kompetitivitou hnaný motor pro sběr verifikovaných dat pro použitou CloudASR knihovnu.

Velký důraz je kladen na grafickou stránku věci; uživatelsky neatraktivní hra nemá valnou naději na šíření mezi uživateli samotnými samovolně. Navíc se na ošklivou aplikaci nikdo nebude dívat po delší časový úsek, než je jedno sezení, a už vůbec nikdo se k takovéto aplikaci nebude vracet. Proto bylo veškeré stylování a malování a navrhování layoutů a přechodů prováděno s výraznou pílí.

Dále je důraz kladen i na zmíněnou kompetitivitu. Každý má touhu vidět své jméno někde vysoko na žebříčku, zanecháváje své přátele na tomto žebříčku daleko pod sebou. Proto byla zapojena i integrace s Facebookovým API, jež přináší kýženého výsledku poměrně nenásilnou formou. Na „Like“ a „Login“ tlačítka jsou uživatelé zvyklí a nebojí se jej extenzivně používat.

Následující text popisuje a dokumentuje vývoj této aplikace a myšlenkové pochody při něm stojící. První kapitola analyzuje e-learning jako způsob předávání informací. Druhá kapitola zmiňuje ASR teorii, popisuje některé postupy a základní paradigmaty. Třetí kapitola seznámí čtenáře s využitými technologiemi a frameworky pro vývoj aplikace trogASR. Kapitola čtvrtá nabídne čtenáři vhled do řešení a implementace aplikace trogASR, do použitých návrhových vzorů a do problematiky s nimi spojené. A konečně, kapitola pátá shrne celou práci.

---

<sup>1</sup><https://github.com/UFAL-DSG/cloud-asr/>

# 1. e-Learning jako způsob předávání informací

1.1 Kapitola

1.2 Kapitola

## **2. Automatické rozpoznávání řeči**

**2.1 Kapitola**

**2.2 Kapitola**



## **3. Použité technologie**

**3.1 Kapitola**

**3.2 Kapitola**

## 4. Implementace

4.1 Kapitola

4.2 Kapitola



Obrázek 4.1: Logo MFF UK

### 4.2.1 Ukázka $\text{\LaTeX}$ u

V této krátké části ukážeme použití několika základních konstrukcí  $\text{\LaTeX}$ u, které by se vám mohly při psaní práce hodit.

Třeba odrážky:

- Logo Matfyzu vidíme na obrázku 4.1.
- Tato subsekce má číslo 4.2.1.
- Odkaz na literaturu [1].

Druhy pomlček: červeno-černý (krátká), strana 16–22 (střední), 45 – 44 (minus), a toto je — jak se asi dalo čekat — vložená věta ohraňčená dlouhými pomlčkami. (Všimněte si, že jsme za **a** napsali vlnovku místo mezery: to aby se tam nemohl rozdělit řádek.)

„České uvozovky.“

**Definice.** *Strom je souvislý graf bez kružnic.*

**Věta 1.** *Tato věta neplatí.*

*Důkaz.* Neplatné věty nemají důkaz. □

# Závěr

# Seznam použité literatury

- [1] LAMPORT, Leslie. *ΛT<sub>E</sub>X: A Document Preparation System*. 2. vydání. Massachusetts: Addison Wesley, 1994. ISBN 0-201-52983-1.

# Seznam tabulek

# Seznam použitých zkratek

ASR            Automatic Speech Recognition - automatické rozpoznávání řeči

TrogASR    Translating Online Game using Automatic Speech Recognition

# Přílohy