

UMĚLÁ INTELIGENCE PRO ZPRACOVÁNÍ OBRAZU A ZVUKU

Jan Švec
honzas@ntis.zcu.cz



FAKULTA
APLIKOVANÝCH VĚD
ZÁPADOČESKÉ
UNIVERZITY
V PLZNI

SpeechTech
speech technology

Kontakt:

Jan Švec

honzas@ntis.zcu.cz

jan.svec@speechtech.cz

www.linkedin.com/in/jansvec



Katedra kybernetiky ZČU v Plzni

- součást Fakulty aplikovaných věd
- výzkum a vývoj v oblastech
 - umělá inteligence (řečové technologie, počítačové vidění, machine learning)
 - informační a řídicí systémy
 - automatické řízení
 - biokybernetika

Odkazy

- <https://www.fav.zcu.cz/>
- <http://www.kky.zcu.cz/cs>



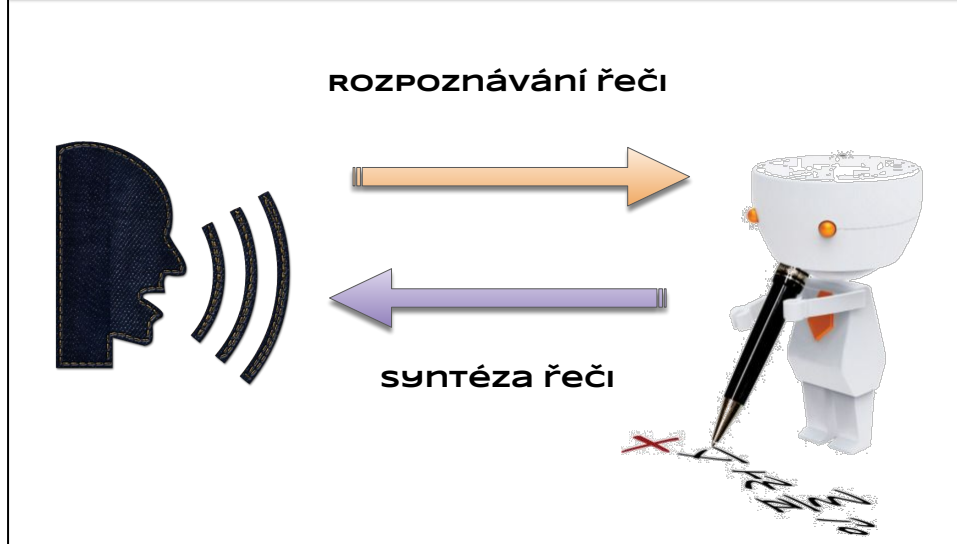
SpeechTech s.r.o.

- komerční partner KKY ZČU
- vývoj a výzkum v oblasti praktických aplikací řečových technologií
- provoz nabízených technologií na vlastním hardware a prodej formou SaaS
- vývoj runtime engine
 - rozpoznávání řeči
 - syntéza řeči
 - hlasové dialogové systémy

Odkazy

- <https://www.speechtech.cz/>

Řečové technologie v kostce



Řečové technologie

- rozpoznávání řeči
 - cíl: přepis řeči do textu
 - (automatic speech recognition, ASR)
- syntéza řeči
 - cíl: vygenerovat přirozenou řeč ze vstupního textu
 - (text-to-speech, TTS)
- porozumění řeči
 - cíl: zjistit strojově čitelnou reprezentaci vstupní řeči (význam)
 - (spoken language understanding, SLU)
- dialogové systémy
 - propojení ASR, SLU a TTS do jednoho celku
 - cíl: vedení konverzace s uživatelem
 - příklad: osobní asistenti (Google Home, Amazon Alexa, Apple Siri)

Úspěšné aplikace řečových technologií



ŠKODA

Úspěšné aplikace řečových technologií KKY + SpeechTech

- Automatické titulkování televizních přenosů (Česká televize)
- Syntéza řeči
 - Hasičský záchranný sbor - čtení informací hasičům jedoucím na zásah
 - Dopravní podnik h.m. Prahy - hlášení informací ve vozech Metra
 - Kaufland - hlášení informací v in-store rádiu
- Hlasové dialogové systémy
 - CS-Soft - automatický pseudopilot pro výcvik řídících letového provozu
 - ŠKODA AUTO - testovací stand pro automatické testování hlasových navigací

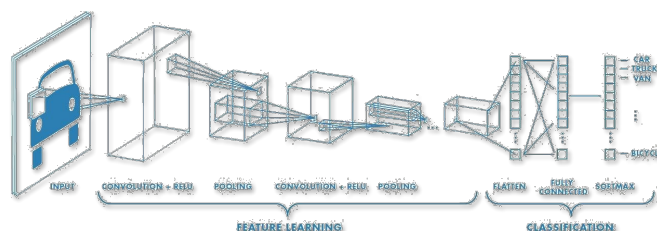
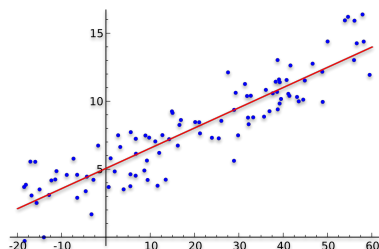


Jaký je rozdíl mezi umělou inteligencí a machine learningem?

Je-li to napsáno v Pythonu, je to nejspíš machine learning.

Je-li to napsáno v PowerPointu, je to zcela určitě umělá inteligence.

UMĚLÁ INTELIGENCE POUŽÍVANÉ METODY



Umělá inteligence

- Všeobímající pojem
- V průběhu dekád znamenající vždy něco jiného
- Zprvu jí byla míněna obecná umělá inteligence
- Nyní nejčastěji krycí název pro strojové učení (machine learning)

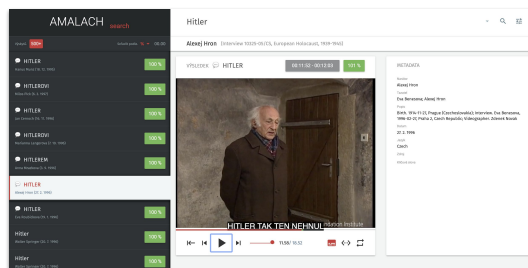
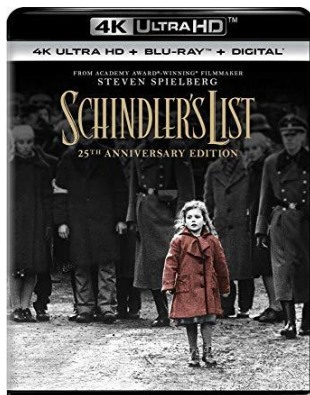
Machine learning

- Obor zkoumající jak odvodit struktury a/nebo parametry statistického modelu na základě dostupných dat
- (Trénovací) data - "palivo" pro veškeré strojové učení

Umělé neuronové sítě

- popř. hluboké neuronové sítě (Deep Neural Networks, DNN)
- typ statistického modelu
- učení parametrů gradientními metodami
- možnost akcelerace na GPU/Tensor kartách
- demokratizace vývoje modelu:
 - frameworky (Tensorflow, Keras, Torch, CNTK)
 - předtrénované modely (Xception, VGG16, ResNet, InceptionV3)
 - dostupný výkonný hardware (GPU)
 - paralelizace a dostupnost v běžných cloudech

INDEXACE A VYHLEDÁVÁNÍ V AUDIOVIZUÁLNÍCH ARCHIVECH



Motivace audiovizuálních archivů

- 1994 - Schindlerův seznam
- Steven Spielberg založil USC Shoah Foundation
- Financování pro získání kolekce Visual History Archive®
 - týmy reportér+kameraman sbírají svědectví o Holocaustu mezi přeživšími
 - 54000 výpovědí, 56 zemí, 32 jazyků
- Otázka zpracování získaného materiálu
 - Ruční zpracování
 - Automatizované zpracování

Vědecký projekt MALACH

- Využití metod rozpoznávání řeči pro vyhledávání v archivu Visual History Archive®
- Mezinárodní projekt pro různé jazyky archivu
- KKY + ÚFAL MFF UK - čeština, slovenština, ruština, výpomoc s maďarštinou
- Financování NSF

Od vědeckého projektu k reálnému vyhledávači

- Po skončení projektu pokračování na půdě KKY
- Financování národními projekty (NAKI), rozšíření o angličtinu
- Nasazení v Centru vizuální historie Malach (CVHM) na MFF UK

Reálné demo

- Přístup z CVHM online
- Provozováno ze ZČU na adrese <https://amalach.zcu.cz>

vech

[HTTPS://amaLach.zcu.cz](https://amaLach.zcu.cz)

AMALACH

search

upgraded **500+** search proba na GO.DIG

- HITLER** 100 %
Hitler, Adolf (1889-1942)
- HITLEROVÍ** 100 %
Hitler, Adolf (1889-1942)
- HITLER** 100 %
Hitler, Adolf (1889-1942)
- HITLEROVÍ** 100 %
Hitler, Adolf (1889-1942)
- HITLEREM** 100 %
Hitler, Adolf (1889-1942)
- HITLER** 100 %
Hitler, Adolf (1889-1942)
- HITLER** 100 %
Hitler, Adolf (1889-1942)
- HITLER** 100 %
Hitler, Adolf (1889-1942)
- HITLER** 100 %
Hitler, Adolf (1889-1942)

Hitler

Alexej Hron (interview 19325-05/CS, European Holocaust, 1939-1945)

VYSLEDEK **HITLER** 00:11:52 - 00:12:03 101 %

HITLER TAK TEN NEHNUL Hitler, Adolf (1889-1942)

METADATA

Název
Alexej Hron

Typ
Eva Benesova, Alexej Hron

Forma
Blaž, 1934-11-31, Prague (Czechoslovakia), interview, Eva Benesova, 1996-02-05, Praha 2, Czech Republic, Videography, Zábek Koubek

Stav
20. 2. 1996

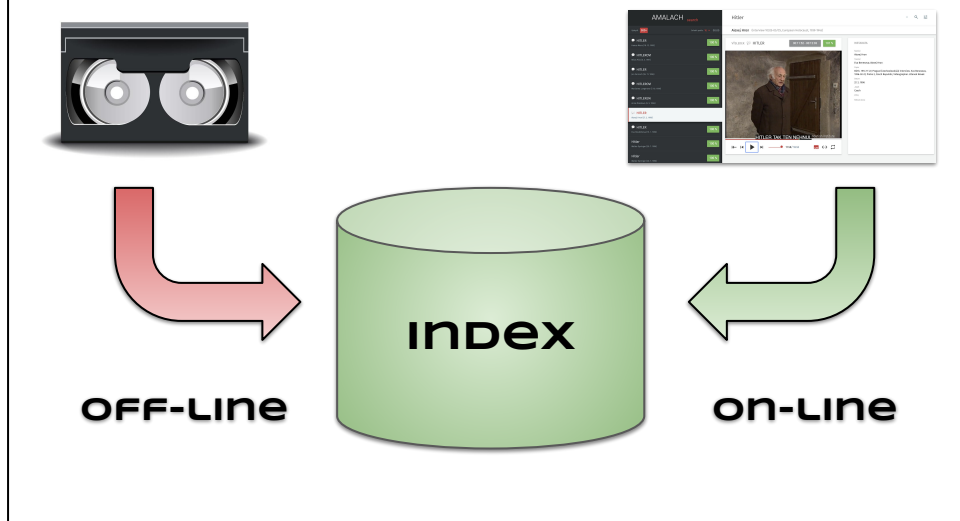
Číslo
Czech

Žánr
Hitlerová

Reálné demo

- Přístup z CVHM online
- Provozováno ze ZČU na adrese <https://amalach.zcu.cz>
- Přístup pouze pomocí klientských certifikátů
- (důvod: autorská práva VHA)

Indexace a vyhledávání



2-fáze zpracování

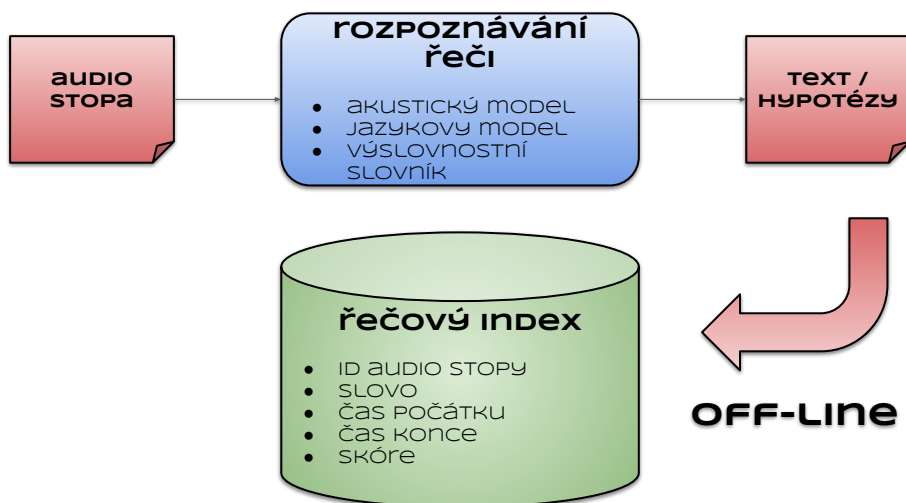
Off-line zpracování (časově náročné, není kritické)

- Konverze audio a video
- Extrakce audio a segmentace
- Rozpoznávání řeči (ASR)
- Indexace výsledku ASR

On-line přístup (kritická rychlost vyhledání)

- Zadáání hledané fráze do GUI
- Vyhledání v indexu
- Přehrání nalezených výsledků
- Morfologie

Vyhledávání v řeči



Rozpoznávání řeči pro AMALACH archiv

- převod vstupní audio stopy do textové podoby
- možnost více alternativních hypotéz
- 3 dílčí modely
 - akustický model
 - modeluje jednotlivé hlásky
 - náročný na vytvoření
 - starší mluvčí
 - slang, germanismy, emoce
 - jazykový model
 - modeluje posloupnosti slov
 - trénování z textů
 - výslovnostní slovník
 - přiřazuje posloupnost hlásek konkrétním slovům
 - často problematické/nejednoznačné

MULTIMODÁLNÍ INDEXACE a VYHLEDÁVÁNÍ

The screenshot displays the CEMI TechDemo interface. On the left, a sidebar shows search results for 'agent londýn' with a 100% match rate. The main area features a video player showing a news segment from Česká televize. The video is titled 'Události (04.05.2018 - Případ otravy Sergeje Skripala)'. The video content includes a news anchor and a portrait of Miloš Zeman. The interface also includes a 'METADATA' section on the right, listing details such as the date (4. 5. 2018), the channel (Česká televize), and the topic (Případ otravy Sergeje Skripala). A word cloud is visible in the bottom right corner, highlighting terms like 'zemana', 'hlavy', 'rusko', and 'spionáž'.

Multimodální indexace a vyhledávání

- zobecnění vyhledávání v řeči
- přidání dalších modalit
 - indexace textu v obraze
 - indexace tváří v obraze
 - analýza obsahu
- náplň projektů CEMI a NAKI-ÚSTR

CEMI (Centrum pro multi-modální interpretaci dat velkého rozsahu)

- Grantová agentura ČR
- konsorcium ČVUT, ÚFAL MFF UK, ZČU, MUNI
- výzkumný projekt
- výsledek: technologický demonstrátor
- vyhledávání v reportážích České televize vysílaných v rámci pořadu Události

NAKI-ÚSTR

- Grantový program Ministerstva kultury ČR
 - Národní kulturní identita
- Partner projektu Ústav pro studium totalitních režimů
- Výpovědi lidí perzekuovaných v době totality
- Složky lidí obsahují i obrazový materiál
 - Skeny vyšetřovacích spisů

Vyhledávání v obraze



TEXT IN THE WILD

OPTICAL CHARACTER
recognition

tolerable state of order among nations. On the contrary, the political, economic, and military situation of the Free world is already unsatisfactory and will become less favorable unless we act to reverse present trends.

This situation is one which militates against successful negotiations with the Soviet Union for the sake of agreement on important pending issues would reflect present realities and would therefore be unworkable, if not disastrous, to the United States and the rest of the free world. Hence a decision has been made and action undertaken to build up the strength, in the broadest sense, of the United States and the free world, an attempt to repudiate a general sentiment in some quarters to no longer be ineffective and probably long drawn out, and might thereby seriously delay the necessary measures to build up our strength.

This is true despite the fact that the United States now has the capability of delivering a powerful blow against the Soviet Union in the event of war, for one of the present realities is that the United States is not prepared to tolerate the use of our present armaments. It is light of present trends, the Soviet Union will not withdraw and the only conceivable basis for a general settlement would be spheres of influence and of an influence-a settlement which the Soviet Union would readily accept in its great advantage. The idea that Germany or Japan or other important areas can exist as islands of neutrality in a divided world is unreal, given the Soviet design for world domination.

A. The Second Course-Isolation

Continuation of present trends, it has been shown above, will lead progressively to the withdrawal of the United States from most of its present commitments to Europe and Asia and to our isolation in the present hemisphere and the hemisphere. This would result not from a conscious decision but from a failure to take the necessary measures to bring our capabilities into line with our commitments and thus to a withdrawal under pressure. This pressure might come from our present allies, who will tend to seek other solutions unless they have confidence in our determination to continue our efforts to build a successfully functioning political and economic system in the free world.

There are some who advocate a deliberate decision to isolate ourselves. Superficially, this has some attractiveness as a course of action, for it appears to bring our commitments and capabilities into balance by reducing the former and by concentrating our present, or perhaps even reduced, military expenditures on the defense of the United States.

NSC 68

- 54 -

Vyhledávání ve videu

- Využití technologie pro text-in-the-wild ČVUT Praha
- Lokalizace textu
- Rozpoznání znaků
- Využití slovníků z rozpoznávání řeči pro odstranění záměn na úrovni znaků
- Text-in-the-wild
 - Text může být natočen
 - Různými fonty
 - Na různém pozadí

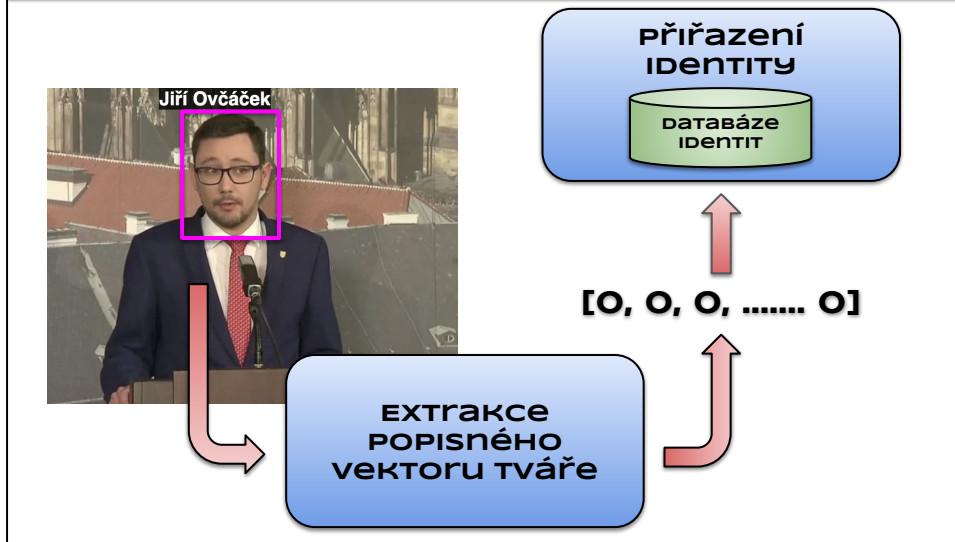
Vyhledávání ve skenech archivních materiálů

- Využití technologie OCR
- Předpoklad strojopisu / tištěného textu
- Bílé pozadí
- Text jedním fontem, popř. několika málo různými fonty

Indexace rozpoznaného textu

- ID záznamu
- Slovo
- X-Y souřadnice bounding-boxu
- Čas počátku (pro text ve videu)
- Čas konce (pro text ve videu)
- Skóre

Vyhledávání tváří a IDENTIT



Vyhledávání tváří

- V rámci projektu CEMI využití software Eyedea Recognition s.r.o.
 - EyeFace / Eyedentity SDK
- Pro video záznam lokace face-tracků
 - posloupnost bounding-boxů, kde se vyskytuje lidská tvář
- Přiřazení deskriptoru každému face-tracku
 - vektor reálných čísel
 - otisk identity tváře

Přiřazení identity

- srovnání deskriptoru s databází identity
- databáze identit
 - ZČU v Plzni
 - ~800 různých identit
 - 377057 příkladů
- hledání identity podle k-nejbližšího souseda
- filtrace face-tracků
 - překryvy
 - střihy

Eyedentity SDK

- navíc pro každý face-track odhad věku a odhad pohlaví
- přidání informace k metadatům videozáznamu

HLAVY STÁTU



Analýza obsahu rozpoznaného textu

- detekce témat
 - strom ~ 600 témat
 - přiřazení pomocí binárního klasifikátoru
 - téma je/není přítomno v textu
- generování WordCloudu
 - jednotlivá témata
 - + významná slova z promluvy
- extrakce sémantický entit (MALACH)
 - názvy obcí
 - datумы
 - geografické údaje

[HTTPS://CEMI.ZCU.CZ](https://cemi.zcu.cz)

CEMI
TechDemo

Výsledky: 3 Zůstatk. punktů: % 00.00

agenta ... Londýn
Události (04. 5. 2018)

100 %

agenta ... Londýně
Události (04. 5. 2018)

100 %

agenta ... Londýn
Události (04. 5. 2018)

agenta londýn

Údálosti (04.05.2018 – Případ otravy Sergeje Skripala)
VÝSLEDEK agenta ...
Londýn

Michal Kubel

Miloš Zeman

pachatelé při otravě ruského dvojčete

blálosti

00:14 / 04:23

METADATA

Názov:
Události

Ceska televize: J. Miloš Zeman (J.M. 0146) /
J. Michal Kubel (J.M. 0123) / J. Andrej Babiš (J.M. 0108) /
J. Vojtěch Filip (J.M. 0101) / J. Jan Hamáček (J.M. 0103) /
J. Jana Černochová (J.M. 0102) / J. Vladimír Putin (J.M. 0105)

Předmět:
Případ otravy Sergeje Skripala

Datum:
4. 5. 2018

Jazyk:
Čeština

Zdroj:
www.ceskatelevize.cz

PLAYLISTY NUBO EPISODE A ROZVOJEKY

Zemana Zeman ruský prezident
státu
ruské
hlavy
spionáž
Rusko
BIS rozvědky
rozvědky

- Provozováno ze ZČU na adrese <https://cemi.zcu.cz>

- Provozováno ze ZČU na adrese <https://cemi.zcu.cz>