

HAnS Epics

und Storys

Technische Hochschule Ingolstadt, Technische Hochschule Nürnberg, Hochschule Hof

07.07.2022

Übersicht

1. Notes from meeting 07-07-22
2. Epic 1: Audio Mining & Segmentation
3. Epic 2: Adaption & Anonymisierung
4. Epic 3: Test & Evaluation

Topics of interest

Mentioned during meeting on 07-07-22

- Thomas: OCR zur Extraktion von Text zur ASR Verbesserung
- Fabian: Speaker Diarization for Anonymization, Text summarization und Topic segmentation
- Christopher: ASR Verbesserung durch Lippenlesung
- Lia: Speaker Diarization, eigenes (E2E) ASR System

Epic 1: Audio Mining & Segmentation

Story 1: Segmentation & Alignment

Task 1: Segmentation

- Segmentierung des Audiosignals auf Satzebene mith. akustischer Features
- Satzsegmente besser geeignet für ASR und spätere Verarbeitung in HAnS

Task 2: Alignment

- Alignment von Audio und Transkript auf Wortebene mith. phonemischer Features
- Wichtig für Begriffssuche im Video und Adaption (Aussprache von Fachbegriffen)

AP 2.1.1 ASR Grundsystem und Anpassung an Lehrmaterial

AP 2.2.1 Audio Mining, zeitsynchroner Lattices & Konfidenzskalierung

Epic 1: Audio Mining & Segmentation

Story 2: Topic Segmentation

- Unterteilung der Vorlesungsvideos in Themenblöcke und Benennung der Subthemen
- Unterteilung der Vorlesung in Themenblöcke erleichtert Weiterverarbeitung im HAnS-Kontext und erhöht Nutzerfreundlichkeit
- Unsupervised & Supervised Ansätze zur Topic Segmentation und Topic Extraction

AP 2.2.4 Topic Modeling mit User-Interaktion

Epic 1: Audio Mining & Segmentation

Story 3: Audio Mining

- Detektieren von Störquellen/Soundproblemen (und anderen Sound Events) in der Audiosequenz erleichtert nachfolgende Bearbeitung in HAnS
- Nimmt Lehrenden Arbeit ab bei der Durchsicht der Aufzeichnung - Anreiz zur Verwendung von HAnS

AP 2.1.1 ASR Grundsystem und Anpassung an Lehrmaterial

AP 2.2.1 Audio Mining, zeitsynchroner Lattices & Konfidenzskalierung

Epic 2: Adaption & Anonymisierung

Story 1: Textbasierte Adaption

- Domain Adaptation der Language Models mit Transkripten und von OCR
- Lernen und Erkennen von Fachbegriffen für ASR (Lattice Rescoring), Topic Modeling, usw.
 - Generierung von domänenspezifischen Vokabularen
- Mehrsprachige Synonyme von Fachbegriffen sollten als solche erkannt werden
 - Auch evtl. als Schnittstelle zwischen deutschem und englischem Lehrmaterial
- Kontinuierliche Verbesserung nach dem Deployment (online learning) und Möglichkeit zur Ergänzung mit neuen Domains

AP 2.1.1 Anpassung an Lehrmaterial, **2.1.3** Domain-Adaptation mit OCR

AP 2.2.1 zeitsynchrone Lattices, **2.2.2** Erkennung & Adaption von Fachbegriffen, **2.2.3** Lernen von Fachausdrücken

Epic 2: Adaption & Anonymisierung

Story 2: Akustische Adaption

- Dynamische Sprecheradaption für das akustische Modell
 - Auch nach Deployment muss an neue Sprecher adaptiert werden
- Anpassung an verschiedene Räume und Aufnahmegeräte
- Konfidenzskalierung für die manuelle Korrektur möglicher Transkriptionsfehler
- Emphasis Detection zur Unterstützung der Fachbegriffserkennung und als Metainformation im Transkript
 - Kann in die UI integriert und von der Suchfunktion verwendet werden

AP 2.1.4 Sprecheradaption

AP 2.2.1 Audiomining & Konfidenzskalierung

Epic 2: Adaption & Anonymisierung

Story 3: Speaker Diarization und Anonymisierung

- Speaker Diarization zur Erkennung und Anonymisierung von Abschnitten, in denen ein Student spricht
- Hauptsprechererkennung für die Sprecheradaption
- TTS-System von Transkription anonymisierter Abschnitte für Accessibility
 - Könnte auch für den KI-Tutor verwendet werden

AP 2.1.5 Sprechererkennung/Anonymisierung der Sprache von Studenten

Epic 3: Test & Evaluation

Story 1: Aufnahmen

- Dozierende benutzen unterschiedliche Mikrofone in unterschiedlichen Räumen
- Spracherkennung & Sprecheradaption sollen trotzdem funktionieren
- Kontrollierte Aufnahmen in verschiedenen Umgebungen

AP 2.1.1 Grundsystem

AP 2.1.4 Sprecheradaption

Epic 3: Test & Evaluation

Story 2: Analyse

- Viele Hochschulen mit verschiedenen Fachbereichen benutzen HAnS
- Akzente & Fachbegriffe sollen erkannt werden
- Bias-Analyse — Wo braucht die Erkennung noch Verbesserung?

AP 2.1.1 Grundsystem

AP 2.1.4 Sprecheradaption

AP 2.2.2 Fachbegriffe