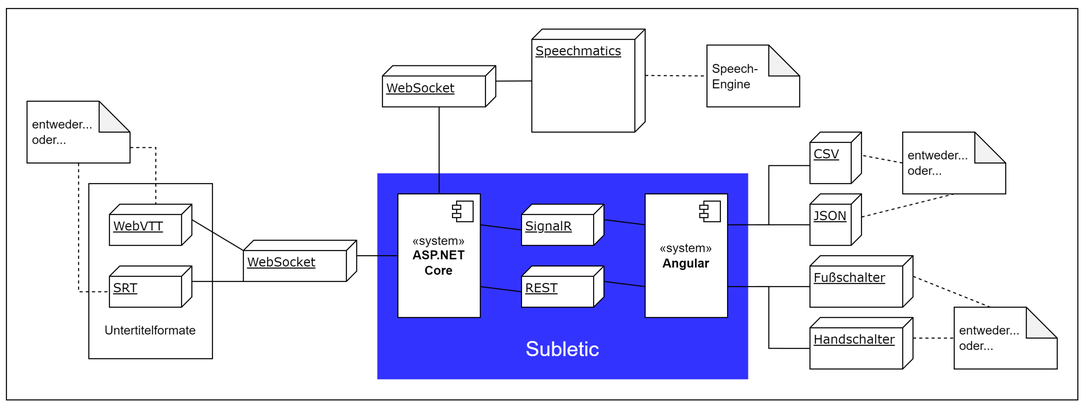
HTWK
30 Jahre
Hochschule für Technik,
Wirtschaft und Kultur Leipzig

Projekt O

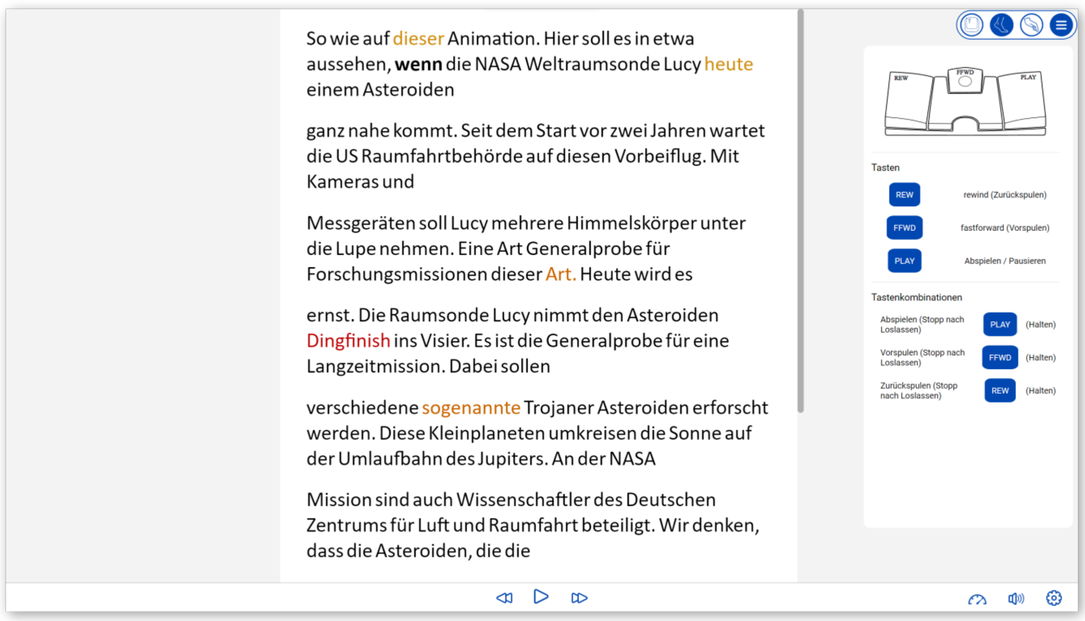
Subletic: Live Stream Editor zur Korrektur von Untertiteln

Mit der von unseren Entwicklern erdachten Software, möchten wir ein Werkzeug schaffen, welches das Korrigieren von Live-Untertitelung, so einfach wie möglich macht. Dabei liest unsere Software einen beliebigen Videostream ein, erstellt mittels einer Speech-Engine eine Untertitelung und bietet die generierten Untertitel dem Anwender zu Korrektur an. Dieser prüft den Untertitel und korrigiert diesen, falls nötig, sodass der Untertitel an den Endkunden ausgeliefert werden kann. Das Endergebnis? Ein perfekt untertitelter Videostream, intuitiv bedienbar, DSGVO-konform und unter Beachtung der Barrierefreiheit.



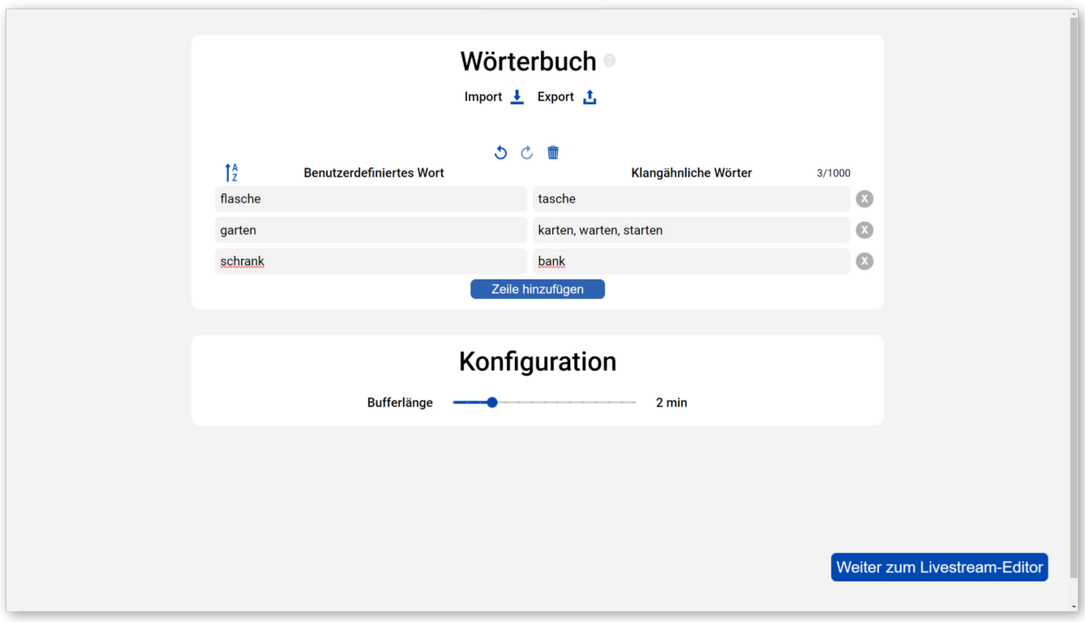
Zielstellung

Live-Untertitelung existiert schon heute, wie die automatische Untertitelung von YouTube zeigt. Diese ist jedoch noch fehleranfällig und eignet sich daher nicht für Live-Streams, bei denen ein hoher Wert auf Korrektheit gelegt wird, wie es unter anderem bei Landtagsdebatten der Fall ist. Dabei besitzen diese Technologien das Potenzial, Berufsgruppen wie Stenografen zu entlasten, und lässt Menschen, die auf Untertitel angewiesen sind, von besserer Untertitelung profitieren. Unser Ziel ist es, die bereits heute existierende Software für automatische Untertitelung, um Echtzeit Korrektur zu erweitern. Dabei sollen die Bedürfnisse von Stenografen unter Beachtung der DSGVO-Konformität umgesetzt werden. Somit kann unsere Software überall da eingesetzt werden, wo es auf fehlerlose Untertitel und die Einhaltung des Datenschutzes ankommt.



Funktionalitäten

Zum Release unserer Software konnten alle Kernfeatures realisiert werden. Wir bieten einen abonnierbaren WebSocket-Service an, mit dem ein Audio-Stream transkribiert wird. Außerdem haben wir eine Web-Anwendung entwickelt, mit dem die generierten Transkriptionen korrigiert werden können. Diese werden in einer Liste von sogenannten Sprechblasen dargestellt und können editiert werden. Dabei unterstützt unser Tool die korrigierende Person indem Wörter hervorgehoben werden, die vermutlich falsch erkannt wurden. Parallel wird der Audio-Stream für den Anwender abgespielt und ihm visuell angezeigt, welches Wort gerade gesagt wird. Um sich zu vergewissern ob ein Wort richtig erkannt wurde, kann der Nutzer vor und zurück springen und dabei auch die Wiedergabegeschwindigkeit anpassen. Eingabegeräte wie Hand und Fußschalter werden unterstützt, um die Bedienung zu erleichtern. Wörter die häufig falsch erkannt werden, wie zum Beispiel Eigennamen oder Fachbegriffe, können in einem Wörterbuch gespeichert werden und sorgen so für eine verbesserte Worterkennung.



Qualitätssicherung

Da unsere Software von jungen Entwicklern programmiert wird, haben wir von Anfang an auf Code-Reviews, Unit-Tests und Linting in der CI-Pipeline gesetzt, um die Qualität unseres Codes zu gewährleisten. Erfahrende Teammitglieder unterstützen die jungen Entwickler durch Code-Reviews und geben ihr Wissen weiter. So haben wir eine durchschnittliche Methoden-Komplexität von 2 und eine Klassenkopplung von 16 erreicht. Die durchschnittlichen Zeilen pro Methode beträgt 6 und die von Klassen 125. Die Line-Coverage liegt bei 66% im Backend und 75% im Frontend. Um die User-Experience zu verbessern, sind unsere Frontend-Entwickler angehalten die 60:30:10-Faustregel anzuwenden. Diese hilft uns ein schlichtes und stimmiges Design durchzusetzen und den Blick des Anwenders durch Akzentfarben zu lenken.

.NET
WebSocket
FFMPEG
GitLab

Wir bedanken uns bei allen Beteiligten:

GRUNDIG Business Systems
i.V. Philipp Platis

Benedikt Beigang
Finn Romeis
Chantal Bley
Pascal Dittes

Luca Noack
Luca Franke
Christoph Neidahl
Amine Jegani

Softwareprojekt II +
Projektmanagementpraktikum II
Wintersemester 23/24
DOZ: Prof. Dr. Karsten Weicker, Tobias Höppner