

Automatische Transkription von Interviews

Die Transkription gesprochener Konversationen ist ein Anwendungsfall, der in unzähligen Gebieten wie der Politik, der Medizin oder der Geschäftswelt vorhanden ist. Seien es Meetings, Interviews im Mediumfeld oder Bewerbungsgespräche; die meisten Unternehmen berufen sich auf menschliche Transkription, was nicht nur ein kosten-, sondern auch ein zeitaufwändiger Prozess ist. Mehr und mehr zeigt sich jedoch die automatisierte Audio-Transkription als realistische Alternative. Die Auswahl an Programmen, die ohne technische Kenntnisse auf der Seite des Benutzers zur Audio-Transkription verwendet werden können, ist jedoch klein und viele wichtige Funktionen, wie Sprecher- oder Pausenerkennung, werden nicht oder nur minimal unterstützt.

Ziel dieser Arbeit war es, basierend auf den Erkenntnissen und Ergebnissen aus einer vorangegangenen Projektarbeit, ein System für die automatische Transkription von Interviews zu entwickeln. Es sollen Audio-Dateien der Anwendung übergeben werden können und das Transkript zurückerhalten werden. Anschliessend soll der Benutzer in einem integrierten Editor Text- sowie Spracherkennungsfehler korrigieren können. Weiter soll das Transkript im Editor durch Hervorheben einzelner Wörter mit der Audio-Datei synchronisiert werden und es sollen dem Benutzer Möglichkeiten zum Export des Transkriptes bereitgestellt werden.

Die Software wurde mittels eines agilen Vorgehensmodells, unter der Verwendung von Elementen aus Scrum und Kanban, in einwöchigen Sprints entwickelt. Jede Woche wurden die im vorgehenden Sprint erarbeiteten Resultate präsentiert, wonach die verbleibenden Funktionen diskutiert und neu priorisiert wurden. Anhand dessen wurde bestimmt, welche User Stories als nächstes in Angriff zu nehmen waren.

Das bereits bestehende «Proof of Concept» wurde weitgehend überarbeitet und stark verbessert. Die Software «InterScriber» beinhaltet alle wichtigen Grundfunktionen, damit das Arbeiten an einem transkribierten Interview möglich ist. Als alternative Transkriptions-Schnittstelle zur Google-Cloud-API wurde ausserdem die IBM-Cloud-API eingebunden, und Deutsch wurde als unterstützte Sprache hinzugefügt. Während die Funktionalität zum Transkribieren von Interviews nun gegeben ist, bleiben jedoch noch einige Funktionen zu implementieren, um aus dem Prototyp ein marktaugliches Produkt zu machen.



Diplomierende
Nicolas Elvis Eckhart
Mathias Richard Marxer

Dozent
Mark Cieliebak



InterScriber unterstützt englische und deutsche Interview-Aufnahmen.



Mit dem integrierten Editor können entstandene Fehler im Transkript korrigiert werden.