

Speech to Text for Medical Record Applications

In dieser Arbeit sollten für den Industriepartner Noser Engineering AG potentielle Einsatzmöglichkeiten von Smart Speakern im Medizinalbereich geprüft werden. Es sollten mögliche Einsatzszenarien gefunden, die Smart Speaker auf ihr Potential und ihre Risiken hin evaluiert, Sprachbibliotheken eingebunden, und den so entstandenen Text analysiert und strukturiert abgespeichert werden. Die Aufgabenstellung stellte sich in dieser Form als zu umfangreich heraus. Diese Arbeit fokussiert sich deshalb auf die Evaluation der Smart Speaker. Die drei marktführenden Smart Speaker (Amazon Echo, Microsoft Invoke, Apple Home Pod) wurden analysiert. Für ein Einsatzszenario wurde mit einem Amazon Echo Dot ein Prototyp programmiert. Mangels verfügbarer Sprachbibliotheken wurde ein eigener Prototyp zur Erkennung von medizinischen Fachwörtern erstellt.

Zur Evaluation der Einsatzszenarien wurden Interviews mit Fachpersonen im Gesundheitsbereich durchgeführt und der technische Aufwand der einzelnen Szenarien bewertet. Für die Spracherkennung wurden zum einen die öffentlich verfügbaren Services mit Beispieltextrn getestet, andererseits wurde mit zwei verschiedenen Technologien je eine mögliche Lösung programmiert: Ein Ansatz verwendet ein Neuronales Netz basierend auf Tensor Flow und dem Google Speech Commands Dataset, der andere Ansatz arbeitet mit Hidden Markov Models und einer statistischen Analyse des Sprachkorpus mithilfe der Kaldi Linguistikbibliothek.

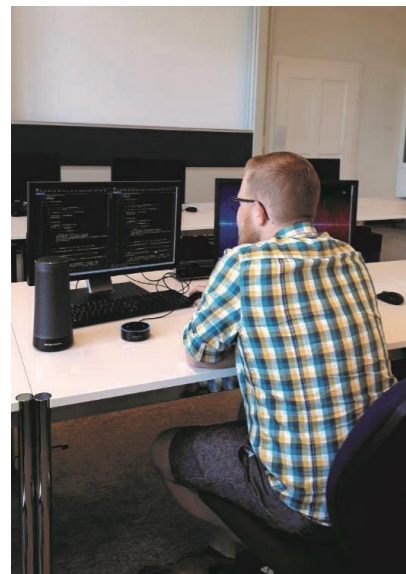
Es konnte festgestellt werden, dass Berufsleute im Gesundheitswesen ein grosses Interesse an Spracherkennung haben. Im Speziellen deshalb, weil ihre Arbeit mit einem grossen administrativen Aufwand einhergeht. Die Bewertung der Nützlichkeit der einzelnen Szenarien variierte aber stark. Bevor ein kommerzielles Produkt entwickelt wird, müssten die Nutzerbedürfnisse noch stärker abgeklärt werden. Die Evaluation der Smart Speaker hat ergeben, dass die Systeme sehr abgeschlossen und sich die Spracherkennung auf kurze Sätze beschränkt. Die Möglichkeiten für Drittentwickler mit diesen Geräten sind eingeschränkt. Ein Anbinden einer externen Sprachbibliothek an einen Smart Speaker ist so nicht möglich. Die beiden Spracherkennungs-Prototypen zeigten aber, dass es durchaus möglich wäre, medizinische Fachbegriffe zuverlässig zu erkennen.

Trotz den Einschränkungen gibt es erfolgsversprechende Einsatzszenarien für Smart Speaker in Arztpraxen und Spitälern.



Diplomierende
Sascha Bergmann
Roman Meier
Simon Raphael Walter

Dozent
Phillipp Ackermann



Programmier- und Testarbeiten an den Smart Speakern Amazon Echo Dot und Harman Kardon Invoke.