

KI-gestützte Transkription von Audiodateien im polizeilichen Umfeld

Studiengang: MAS Digital Forensics & Cyber Investigation

Im polizeilichen Umfeld entstehen jährlich mehrere Terabyte an Audiodaten, die aufgrund ihrer hohen Sensibilität bislang überwiegend manuell ausgewertet werden müssen. Der Einsatz kommerzieller Cloud-basierter Tools ist aus Datenschutzgründen ausgeschlossen, während geeignete automatisierte Lösungen bisher fehlen. Ziel dieser Arbeit ist die prototypische Entwicklung und Evaluation eines solchen Transkriptionssystems.

Ausgangslage

Im polizeilichen Ermittlungsumfeld werden regelmässig grosse Mengen an Audiodaten sichergestellt. Die manuelle Auswertung dieser Daten ist zeitintensiv, personalaufwendig und fehleranfällig. Hinzu kommt, dass bestehende kommerzielle Transkriptionssysteme häufig Cloud-basiert sind und damit datenschutzrechtliche Risiken bergen. Zusätzlich sind Audiodaten häufig in Schweizer Dialekten, Fremdsprachen oder unter schwierigen akustischen Bedingungen vorhanden, was die Analyse weiter erschwert.

Zielsetzung

Ziel dieser Masterthesis ist die prototypische Entwicklung und Evaluation eines on-premise Transkriptionssystems auf Basis von Open-Source-Technologien. Das System soll Audiodaten automatisiert in Text überführen, eine effiziente Stichwortsuche ermöglichen und dadurch polizeiliche Ermittlungsprozesse deutlich vereinfachen und beschleunigen.

Methode

Die Arbeit kombiniert qualitative Interviews mit Ermittlerinnen und Ermittlern zur Erhebung praxisnaher Anforderungen mit einem designorientierten Entwicklungsprozess und der anschliessenden prototypischen Umsetzung eines Transkriptionssystems. Auf Basis der Interviewergebnisse sowie einer

systematischen Analyse bestehender Lösungen wurde ein Prototyp realisiert.

Resultat

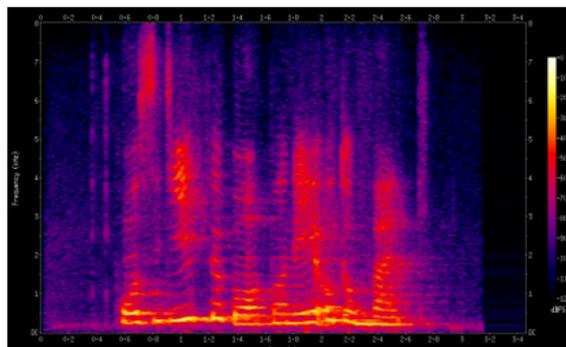
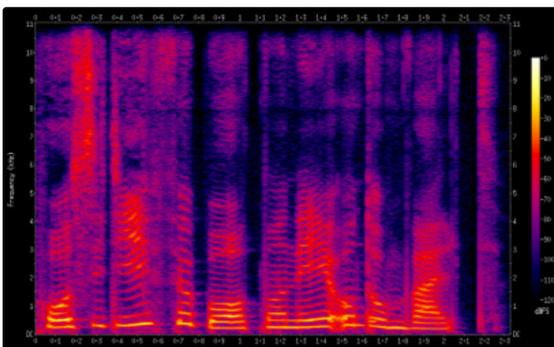
Das Resultat besteht aus einem vollständig entwickelten Prototyp, der das Open-Source-Modell Whisper nutzt, sowie einem Python-Skript, das die Schnittstelle zu einem forensischen Tool implementiert und dort integriert wurde. Für die Evaluation wurden praxisnahe Testszenarien erstellt, die Schweizer Dialekte, Fremdsprachen sowie Aufnahmen mit Hintergrundgeräuschen umfassten. Die Leistungsfähigkeit wurde anhand der Word Error Rate (WER) und durch qualitative Analysen bewertet. Auf dieser Basis konnte der Transkriptionsprozess erfolgreich automatisiert werden, sodass die aufbereiteten Daten den Ermittlerinnen und Ermittlern effizient bereitgestellt werden können.

Ausblick

Transkriptionen von Audiodateien aus Mobiltelefonen werden künftig mit dem entwickelten System durchgeführt. Das Transkriptionssystem soll dabei weiter ausgebaut und funktional erweitert werden. Mit der geplanten Überführung in den produktiven Einsatz entfällt der Prototypen-Status, und das System wird als vollwertiges Tool für die polizeiliche Praxis nutzbar.



Velko Filcev



Gegenüberstellung der Spektrogramme von WAV- und Opus derselben Datei. Links ist das unkomprimierte Signal mit grösserer Bandbreite und feineren Strukturen dargestellt, rechts die komprimierte Version