

Peer-to-Peer Kommunikation für Sprachübertragung in einem Praxisrufsystem

IP6 - Bachelor Thesis

2. November 2021

Studenten Joshua Villing

Fachbetreuer Daniel Jossen

Auftraggeber Daniel Jossen

Studiengang Informatik

Hochschule Hochschule für Technik

Management Summary

Lorem Ipsum

Inhaltsverzeichnis

1	Einleitung	1
2	Vorgehensweise	2
3	Anforderungen	4
4	Technologie Evaluation	6
4.1	Appentwicklung IOS	6
4.2	Text To Speech	7
4.3	Sprachübertragung	9
4.4	Full on VOIP mit Telefonnummern	10
5	Konzept	11
5.1	SystemArchitektur	11
5.2	Migration Benachrichtigungen	13
5.3	Integration Text To Speech	14
5.4	Integration Gegensprechanlage	15
5.5	Übersicht Erweiterung Praxisruf Cloud Service	16
6	Umsetzung	17
7	Schluss	18
	Literaturverzeichnis	19
	Abbildungsverzeichnis	21
A	Aufgabenstellung	22
B	Quellcode	23
C	Ehrlichkeitserklärung	24

1 Einleitung

”Ärzte und Zahnärzte haben den Anspruch in Ihren Praxen ein Rufsystem einzusetzen. Dieses Rufsystem ermöglicht, dass der behandelnde Arzt über einen Knopfdruck Hilfe anfordern oder Behandlungsmaterial bestellen kann. Zusätzlich bieten die meisten Rufsysteme die Möglichkeit eine Gegensprechfunktion zu integrieren. Eine durchgeführte Marktanalyse hat gezeigt, dass die meisten auf dem Markt kommerziell erhältlichen Rufsysteme auf proprietären Standards beruhen und ein veraltetes Bussystem oder analoge Funktechnologie zur Signalübermittlung einsetzen. Weiter können diese Systeme nicht in ein TCP/IP-Netzwerk integriert werden und über eine API extern angesteuert werden.

Im Rahmen dieser Arbeit soll ein Cloudbasiertes Praxisrufsystem entwickelt werden. Pro Behandlungszimmer wird ein Android oder IOS basiertes Tablet installiert.

Auf dieses Tablet kann die zu entwickelnde App installiert und betrieben werden. Die App deckt dabei die folgenden Ziele ab:

- Evaluation Frameworks für die Übertragung von Sprachinformationen (1:1 und 1:m)
- Erweiterung SW-Architektur für die Übertragung von Sprachdaten
- Definition und Implementierung Text-to-Speech Funktion
- Implementierung Sprachübertragung inklusive Gegensprechfunktion
- Durchführung von Funktions- und Performancetests

Die Hauptproblemstellung dieser Arbeit ist die sichere und effiziente Übertragung von Sprach- und Textmeldungen zwischen den einzelnen Tablets. Dabei soll es möglich sein, dass die App einen Unicast, Broadcast und Multicast Übertragung der Daten ermöglicht. Über eine offene Systemarchitektur müssen die Kommunikationsbuttons in der App frei konfiguriert und parametrisiert werden können.”¹[1]

¹Ausgangslage, Ziele und Problemstellung im Originaltext der Aufgabenstellung

2 Vorgehensweise

Projektplan



Abbildung 2.1: Projektplan

Lorem ipsum

Meilensteine

In der Anfangsphase des Projektes wurden folgende Meilensteine definiert:

Id	Beschreibung
M01	Initiale Anforderungsanalyse Die Anforderungen an das Projekt aus der Aufgabenstellungen sind in User Stories dokumentiert.
M02	Einarbeit und Setup IOS Umgebung Projektteilnehmende sind mit groben Konzepten der IOS Entwicklung vertraut. Die Entwicklungsumgebung ist bereit für die Umsetzung.
M03	Evaluation Technologien Die Evaluation der Technologien für Text To Speech und Gegensprechanlage (VOIP Kommunikation) ist abgeschlossen.
M04	Konzepte Die Konzepte für Systemarchitektur, Aufbau und Architektur der mobilen Applikation sowie Anpassungen an bestehenden Komponenten im System sind abgeschlossen.
M05	Migration bestehender Funktionalität Die Funktionen die im Mobile Client der Projektarbeit IP5 Cloudbasiertes Praxisrufsystem umgesetzt wurden, stehen in der neu entwickelten nativen IOS Applikation zur Verfügung.
M06	Umsetzung Text To Speech Alle Anforderungen zu der Text To Speech Funktion sind in der neu entwickelten nativen IOS Applikation umgesetzt.
M07	Umsetzung Gegensprechanlage 1:1 Alle Anforderungen für die 1:1 Kommunikation über die Funktion Gegensprechanlage sind in der neu entwickelten nativen IOS Applikation umgesetzt.
M08	Umsetzung Gegensprechanlage 1:n Alle Anforderungen für die 1:1 Kommunikation über die Funktion Gegensprechanlage sind in der neu entwickelten nativen IOS Applikation umgesetzt.
M09	Abnahme Die Abnahmetests wurden zusammen mit dem Kunden ausgeführt.

3 Anforderungen

Es gibt drei Rollen von Stakeholdern, welche Anforderungen an Praxisruf stellen. Die meisten Benutzer des Systems fallen in die Rolle Praxismitarbeitende. Diese verwenden die mobile Applikation von Praxisruf, um in der Praxis miteinander zu kommunizieren. Neben der Rolle der Praxismitarbeitenden, arbeitet auch die Rolle des Praxisverantwortlichen mit dem Praxisrufsystem. Diese Benutzergruppe ist dafür verantwortlich, Praxisruf für Praxismitarbeitende zu konfigurieren. Als dritte Rolle hat zudem der Auftraggeber ein Interesse daran, dass gewisse Rahmenbedingungen gesetzt und eingehalten werden. Siehe Projektbericht Cloudbasiertes Praxisrufsystem [2].

Im folgenden Kapitel werden die Anforderungen dokumentiert, die bei Projektstart ermittelt wurden. Die Anforderungen werden dabei aus fachlicher Sicht mit User Stories festgehalten. Jede User Story beschreibt ein konkretes Bedürfnis einer Stakeholdergruppe.

Praxismitarbeitende

Id	Anforderung
U01	Als Praxismitarbeiter/in möchte ich alle Funktionen aus der existierenden Applikation weiterhin verwenden können, damit mir diese weiterhin die Arbeit erleichtern. ²
U02	Als Praxismitarbeiter/in möchte ich, dass wichtige eingehende Benachrichtigungen vorgelesen werden, damit den Inhalt der Benachrichtigung kenne, ohne meine Aufmerksamkeit auf den Bildschirm zu richten.
U03	Als Praxismitarbeiter/in möchte ich, das Vorlesen von Benachrichtigungen deaktivieren können, damit ich bei der Arbeit nicht unnötig gestört werde.
U04	Als Praxismitarbeiter/in möchte ich, per Button eine Sprachverbindung zu einem anderen Praxiszimmer aufbauen können, damit ich mich mit einer anderen Person absprechen kann.
U05	Als Praxismitarbeiter/in möchte ich, per Button eine Sprachverbindung zu mehreren anderen Praxiszimmern aufbauen können damit, ich mich mit mehreren anderen Personen absprechen kann.
U06	Als Praxismitarbeiter/in möchte ich über geöffnete Sprachverbindungen in Echtzeit kommunizieren können damit es die Funktion einer Gegensprechanlage wirklich erfüllt.
U07	Als Praxismitarbeiter/in möchte ich nur Buttons für Sprachverbindungen sehen, die für mich relevant sind.
U08	Als Praxismitarbeiter/in möchte ich benachrichtigt werden, wenn ein anderes Zimmer eine Sprachverbindung öffnet, damit ich auf die Anfrage Antworten kann.
U09	Als Praxismitarbeiter/in möchte ich vergangene und verpasste Sprachverbindungen nachvollziehen können, damit ich mich zurückmelden kann.
U10	Als Praxismitarbeiter/in möchte ich, dass eingehende Sprachverbindungen aus anderen Praxiszimmern automatisch geöffnet werden damit ich meine Hände für besseres brauchen kann.
U11	Als Praxismitarbeiter/in möchte ich, direkte Sprachverbindungen aus anderen Praxiszimmern trennen können damit ich ein Gespräch beenden kann.
U12	Als Praxismitarbeiter/in möchte ich, aus Sprachverbindungen zu mehreren Praxiszimmern (Gruppenunterhaltungen) austreten können, damit ich nicht unnötig bei der Arbeit gestört werde.

Praxisadministrator

Id	Anforderung
U13	Als Praxisadministrator möchte ich konfigurieren können, welche Benachrichtigungen dem Praxismitarbeitenden vorgelesen werden, damit nur relevante Benachrichtigungen vorgelesen werden.
U14	Als Praxisadministrator möchte ich konfigurieren können, aus welchen Zimmern Sprachverbindungen zu welchen anderen Zimmern aufgebaut werden können, damit die Mitarbeitendend das System effizient bedienen können.
U15	Als Praxisadministrator möchte ich Benachrichtigungen, Clients und Benutzer wie zuvor konfigurieren können, damit ich das System weiterhin auf meine Praxis zuschneiden und bestehende Konfigurationen übernehmen kann.

Auftraggeber

Id	Anforderung
U16	Als Auftraggeber möchte ich die bestehende Betriebsinfrastruktur übernehmen, um von der bereits geleisteten Arbeit profitieren zu können.
U17	Als Auftraggeber möchte ichm, dass die bestehende Komopnenten des Systems wo immer möglich weiter verwendet werden, um von der bereits geleisteten Arbeit profitieren zu können.

4 Technologie Evaluation

Hier werden die Technologien und Frameworks mit denen das Projekt umgesetzt wird evaluiert.

4.1 Appentwicklung IOS

Übersicht

Bisherige Funktionalität muss implementiert werden können.

Firebase Anbindung muss funktionieren.

Audio abspielen muss funktionieren.

Push Benachrichtigungen müssen funktionieren.

Optionen

Swift ist standard

UIKit oder SwiftUI

SwiftUI ist neu und die Richtung in die Apple seit einiger Zeit pushed.

Es wird SwiftUI genommen, weil es der neue Standard ist.

Gewisse Hooks in den IOS Lifecycle sind mit rein SwiftUI noch nicht möglich.

Die entsprechenden Teil (AppDelegate) können auch in SwiftUI verwendet werden.

Kleiner POC wurde mit SwiftUI erstellt um zu verifizieren, dass Anforderungen von oben möglich sind.

Entscheidung:

Target IOS15

Swift 5

SwiftUI 3.0

4.2 Text To Speech

Praxisruf soll es neu Unterstützen, dass empfangene Benachrichtigungen vorgelesen werden. Um dies zu ermöglichen muss eine Technologie integriert werden, die es erlaubt den Textinhalt einer Benachrichtigung in Audio zu konvertieren welches abgespielt werden kann.

Grundsätzlich gibt es zwei Varianten wie dies erreicht werden kann. Die erste Option ist es, die Konvertierung auf dem Mobile Client selbst vorzunehmen. In diesem Fall ist die Konvertierung teil der Mobile App. Das IOS Core Framework bietet Packages welches dies erlauben.

Die zweite Option ist es, die Konvertierung an einen Cloud Service zu delegieren. Dazu muss ein externer Text To Speech Provider angebunden werden, der vom Praxisruf angesprochen werden kann.

4.2.1 Nativ

Vorteile

Einfach einzubinden

Garantierter nativer support

Nachteile

Starke bindung an Apple.

Wenn die Funktion künftig nicht mehr zur Verfügung steht, geht sie verloren

Mehr Verantwortung im Mobile Client selbst (weniger klar getrennte Verantwortung)

Funktion für Android Client kann damit nicht umgesetzt werden

Anbieter

Speech Synthesis von IOS[3]

4.2.2 Cloud

Vorteile

Unabhängig von Apple

Einfacher austauschbar (bei entsprechender implementation der Anbindung)

Klar getrennte Verantwortung

Funktion für Android Clients kann gleich Umgesetzt werden.

Nachteile

Zusätzliche Anbindung von Service

Mehrere Umsetzungsmöglichkeiten (Client spricht direkt mit T2S Cloud oder Cloud Service Spricht mit T2s Cloud)

Zusätzlicher Point of Failure / Komplexität

Anbieter

AWS Polly [4]

Unterstützt alles

Rest Infrastruktur ist bereits bei Amazon
Kunde wünscht einheitlich bei AWS wo möglich

4.2.3 Entscheidung

Anbindung per Cloud wird genommen. Die Implementierung kann dadurch komplizierter werden. Es ist aber sehr wichtig unabhängig zu bleiben. Apple Version bindet zu stark an Apple. Kann nicht für künftigen Android Client wiederverwendet werden. Zukunftssicherer weil der Konkrete T2S Service ausgewechselt werden kann. Es bestehen SDKs für Java und für Client. Damit ist die Anbindung direkt aus Client oder aus Cloud Service möglich. Welche Option implementiert wird, ist im Konzeptteil zu beschreiben.

4.3 Sprachübertragung

Praxisruf soll um die Funktion einer Gegensprechanlage erweitert werden. Um dies zu ermöglichen muss das System Sprachübertragung zwischen Mobile Clients in Echtzeit unterstützen. In diesem Kapitel werden die Technologien ermittelt, mit denen dies umgesetzt werden kann.

4.3.1 WebRTC

Mit WebRTC können Sie Ihrer Anwendung Echtzeit-Kommunikationsfunktionen hinzufügen, die auf einem offenen Standard basieren. Es unterstützt Video-, Sprach- und generische Daten, die zwischen Peers gesendet werden, sodass Entwickler leistungsstarke Sprach- und Videokommunikationslösungen erstellen können. Die Technologie ist in allen modernen Browsern sowie auf nativen Clients für alle wichtigen Plattformen verfügbar. Die Technologien hinter WebRTC sind als offener Webstandard implementiert und in allen gängigen Browsern als reguläre JavaScript-APIs verfügbar. Für native Clients wie Android- und iOS-Anwendungen steht eine Bibliothek mit derselben Funktionalität zur Verfügung.”[5]

Varianten

Externer Anbieter (twilio)

Vorteile:

Einfachere Integration auf Client Seite durch Vendor SDK

Auf Cloud Service Seite nur Configuration Domain betroffen, keine andere Erweiterung nötig

Alles andere kann auf Client Seite erledigt werden.

Nachteile:

Kompliziertere Integration mit Client -> Client muss korrekte Verbindungen anhand Button Konfiguration aufmachen

Verantwortung des Clients wächst. Aktuell hat er nur Verantwortung zum Empfangen, die Arbeit wird immer von andern gemacht

Self Hosted (WebRtc Server als Teil von cloud service implementieren)

Vorteile:

Vermittlung und Signaling kann vom Cloud Service übernommen werden.

Grössere Flexibilität, da Vermittlung in Cloud Service

Evtl. Synergie mit Rules Engine in Configuration Domain

Nachteile:

Kompliziertere Einbindung auf Client Seite.

4.3.2 Meeting Solution

AWS Chime, Teams oder Ähnliches als Basis

Nicht wirklich p2p?

Starke Vendor Bindung

4.4 Full on VOIP mit Telefonnummern

Not feasible

Nicht was wir brauchen, da nur vordefinierte calls zwischen clients gemäss konfiguration möglich sein sollen

Es ist nicht gefordert und nicht gewünscht dass irgendjemand per telefonnummer, den client erreichen kann.

4.4.1 Entscheidung

WebRTC self hosted.

Synergie Rules Engine

Dispatching in Verantwortung Cloud Service

SDK gut genug

Nicht vendor abhängig

5 Konzept

5.1 SystemArchitektur

Mit diesem Projekt wird Praxisruf um die Funktionen Text To Speech und Gegensprechanalge erweitert. Um dies zu ermöglichen, sind Erweiterungen an der Systemarchitektur nötig. Bestehende Module bleiben. Modularität wird erweitert. Bisher nur Package Trennung. Neu werden Gradle Module pro Domain gemacht. Immernoch in einem Service nachher. Aber eine Stufe näher daran, es in Microservices aufzutrennen. Übersichtlicher, einfacher erweiterbar. Trennung der Domänen garantiert.

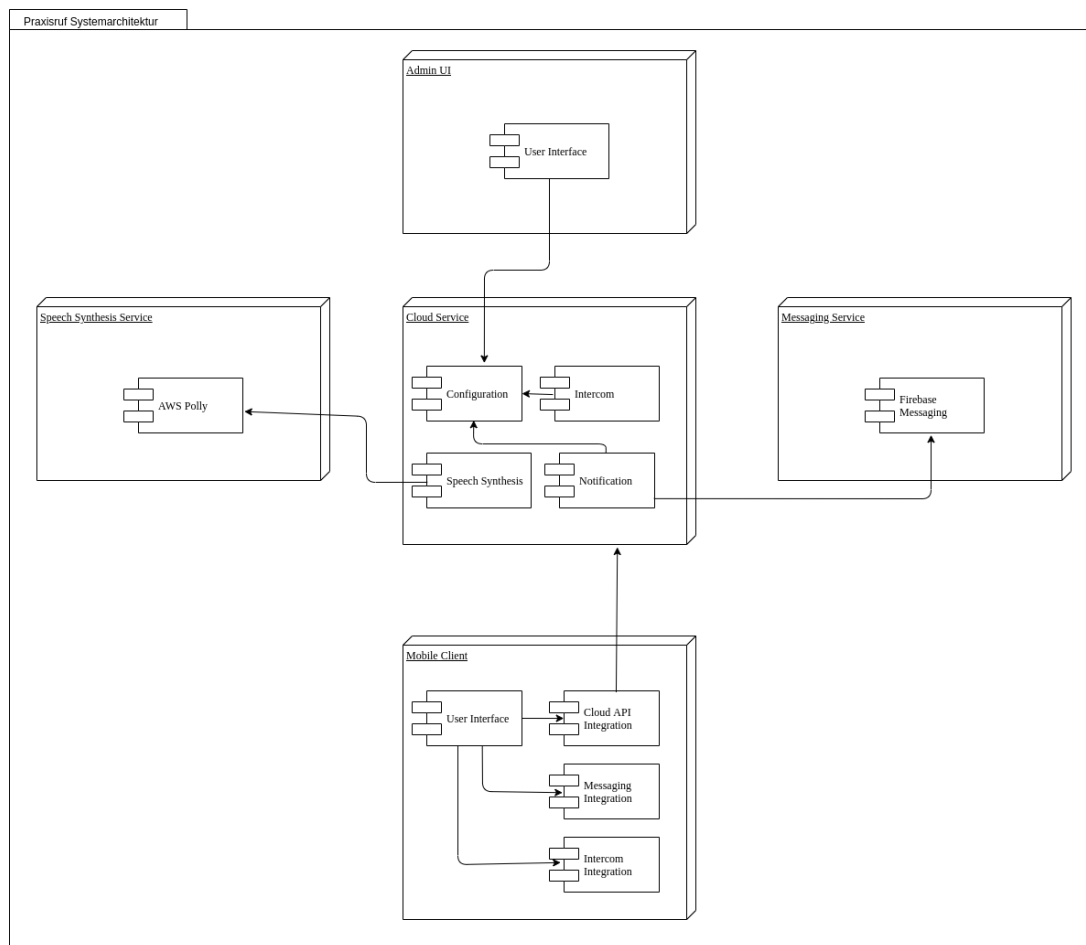


Abbildung 5.1: Systemarchitektur Praxisruf

Cloudservice

Auf Seite Cloud Service werden die Module Intercom und Speech Synthesis hinzugefügt. Intercom übernimmt das Signaling für WebRTC. Hat den Vorteil, dass künftig auch Web und Android Clients an denselben Signaling Service angebunden werden können. Vermittlung passiert anhand der vom Admin erfassten Konfiguration.

Speech Synthesis dient als einheitliche Schnittstelle zu einem externen Speech Synthesis Service. Dadurch kann auch wenn ein Android oder Web Client kommt, dieser genau gleich angebunden werden. Garantie, dass die Konfiguration und Funktionsweise dieselbe für alle Clients ist.

Mobile Client

Neu als nativer Client mit SwiftUI. Beinhaltet des Benutzer Interface. Sowie Komponenten zur Anbindung an Configuration, Notification, Speech Synthesis und Intercom. Details in Client Kapitel.

Admin UI

Das Admin UI dient weiterhin zur Administration der Configuration. Das Admin UI wird um Konfigurationsmöglichkeiten für Speech Synthesis und Gegensprechanalge werweitert.

Speech Synthesis Service

Als neuer externer Service AWS Polly angebunden. Dabei handelt es sich um die Speech Synthesis Funktion von Amazon Webservices.

Messaging Service

Der Messaging Service wird weiterhin zum Versenden von Benachrichtigungen verwendet. Am Messaging Service werden in diesem Projekt keine Änderungen vorgenommen.

5.2 Migration Benachrichtigungen

Mit IP5 wurde bereits ein Client umgesetzt. Dieser muss für IP6 migiert werden.

Architektur Anbindung Rest

Architektur Anbindung Firebase

5.3 Integration Text To Speech

Benutzeroberfläche

Erweiterung Inbox um T2S Icon.

Erweiterung Admin UI um Checkbox.

Zusätzlich Configuration Page mit preferences.

Konfiguration

Erweiterung NotificationType um ein boolean Flag für isTextToSpeech. Wenn Aktiviert, wird Benachrichtigung bei Empfang vorgelesen. Text To Speech kann auf Client Seite deaktiviert werden. Ist es deaktiviert, werden keine Benachrichtigungen vorgelesen. Audio Signal, dass Benachrichtigung empfangen wurde ertönt aber trotzdem. Keine zusätzlichen Endpoints am Cloud Service nötig.

Laufzeitsicht

Empfang und Versenden gleich wie bei IP5. Benachrichtigung enthält zudem neu Flag ob T2S gebraucht werden soll. Wenn ja, wird Vorlesen an T2S Service delegiert.

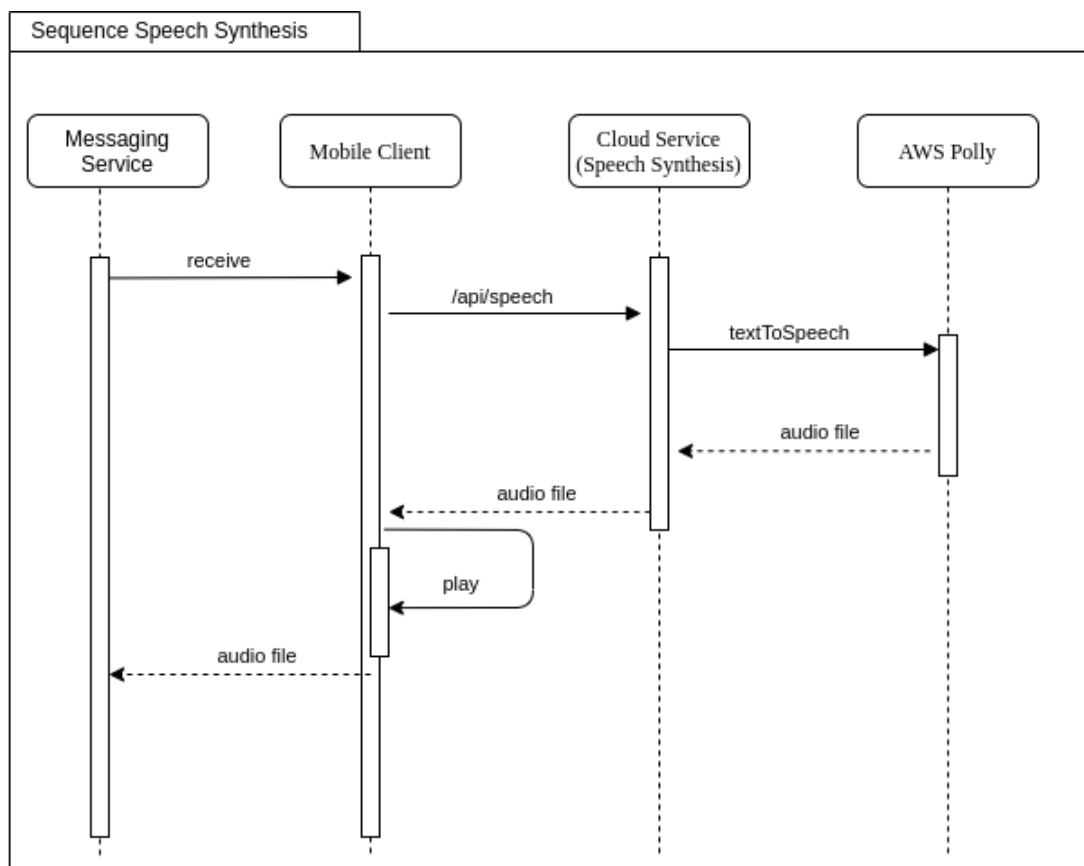


Abbildung 5.2: Ablauf Benachrichtigung empfangen

5.4 Integration Gegensprechanlage

Benutzeroberfläche

Button Screen wie IP5

Active Call Screen

Erweiterte Inbox

Neue Screens für Gegensprechanlage.

Erweiterung Admin UI für Konfiguration CallType

Konfiguration

Erweiterung Configuration Domain um CallType. Hat text property, dass als Anzeige auf dem Button dient. Hat Liste von Clients, die im Call angesprochen werden können.

Laufzeitsicht

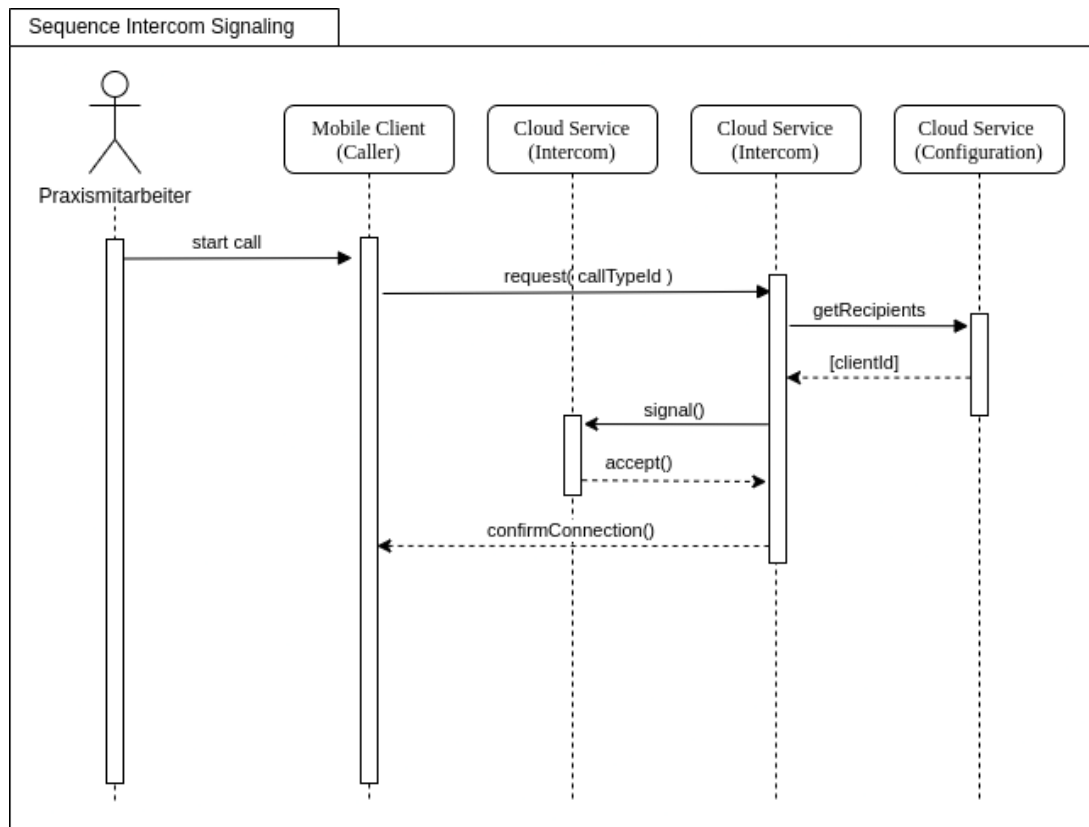


Abbildung 5.3: Ablauf Verbindungsaufbau Gegensprechanlage

5.5 Übersicht Erweiterung Praxisruf Cloud Service

;; Erweitertes ERD für Configuration Domain (Cloud Service) ::

6 Umsetzung

Lorem ipsum

7 Schluss

Lorem Ipsum

Literaturverzeichnis

- [1] D. Jossen, *21HS-IMVS38: Peer-to-Peer Kommunikation für Sprachübertragung in einem Praxisrufsystem*, 2021.
- [2] J. Villing, K. Zellweger, "Cloudbasiertes Praxisrufsystem," FHNW - Hochschule für Technik, Techn. Ber., 2021.
- [3] A. Inc. (). Speech Synthesis, Adresse: https://developer.apple.com/documentation/avfoundation/speech_synthesis.
- [4] I. Amazon Webservices. (). Amazon Polly, Adresse: <https://aws.amazon.com/polly/>.
- [5] Google Developers. (). WebRTC - Echtzeitkommunikation für das Web, Adresse: <https://webrtc.org/>.

Abbildungsverzeichnis

2.1	Projektplan	2
5.1	Systemarchitektur Praxisruf	11
5.2	Ablauf Benachrichtigung empfangen	14
5.3	Ablauf Verbindungsaufbau Gegensprechanlage	15
A.1	Aufgabenstellung	22

A Aufgabenstellung

21HS_IMVS38: Peer-to-Peer Kommunikation für Sprachübertragung in einem Praxisrufsystem

Betreuer: [Daniel Jossen](#)

Arbeitsumfang: P6 (360h pro Student)

Priorität 1

Priorität 2

Teamgrösse: Einzelarbeit

Sprachen: Deutsch

Ausgangslage

Ärzte und Zahnärzte haben den Anspruch in Ihren Praxen ein Rufsystem einzusetzen. Dieses Rufsystem ermöglicht, dass der behandelnde Arzt über einen Knopfdruck Hilfe anfordern oder Behandlungsmaterial bestellen kann. Zusätzlich bieten die meisten Rufsysteme die Möglichkeit eine Gegensprechfunktion zu integrieren. Ein durchgeführte Marktanalyse hat gezeigt, dass die meisten auf dem Markt kommerziell erhältlichen Rufsysteme auf proprietären Standards beruhen und ein veraltetes Bussystem oder analoge Funktechnologie zur Signalübermittlung einsetzen. Weiter können diese Systeme nicht in ein TCP/IP-Netzwerk integriert werden und über eine API extern angesteuert werden.



Ziel der Arbeit

Im Rahmen dieser Arbeit soll ein Cloudbasiertes Praxisrufsystem entwickelt werden. Pro Behandlungszimmer wird ein Android oder IOS basiertes Tablet installiert. Auf diese Tablet kann die zu entwickelnde App installiert und betrieben werden. Die App deckt dabei die folgenden Ziele ab:

- Evaluation Frameworks für die Übertragung von Sprachinformationen (1:1 und 1:m)
- Erweiterung SW-Architektur für die Übertragung von Sprachdaten
- Definition und Implementierung Text-to-Speech Funktion
- Implementierung Sprachübertragung inklusive Gegensprechfunktion
- Durchführung von Funktions- und Performancetests

Problemstellung

Die Hauptproblemstellung dieser Arbeit ist die sichere und effiziente Übertragung von Sprach- und Textmeldungen zwischen den einzelnen Tablets. Dabei soll es möglich sein, dass die App einen Unicast, Broadcast und Multicast Übertragung der Daten ermöglicht. Über eine offene Systemarchitektur müssen die Kommunikationsbuttons in der App frei konfiguriert und parametrisiert werden können.

Technologien/Fachliche Schwerpunkte/Referenzen

- Cloud Services (AWS)
- IOS App-Entwicklung (SWIFT)
- Sichere Übertragung von Sprach- und Textmeldungen

Bemerkung

Dieses Projekt ist für Joshua Villing reserviert.

Abbildung A.1: Aufgabenstellung

B Quellcode

Sämtlicher Quellcode der im Rahmen des Projektes entsteht, wurde mit Git verwaltet. Der Quellcode ist für Berechtigte unter github.com einsehbar². Berechtigungen können bei Joshua Villing angefordert werden.

²<https://github.com/users/jsvilling/projects/3>

C Ehrlichkeitserklärung

«Hiermit erkläre ich, die vorliegende Projektarbeit IP6 - Cloudbasiertes Praxisrufsystem selbständig und nur unter Benutzung der angegebenen Quellen verfasst zu haben. Die wörtlich oder inhaltlich aus den aufgeführten Quellen entnommenen Stellen sind in der Arbeit als Zitat bzw. Paraphrase kenntlich gemacht. Diese Projektarbeit ist noch nicht veröffentlicht worden. Sie ist somit weder anderen Interessierten zugänglich gemacht noch einer anderen Prüfungsbehörde vorgelegt worden.»

Name Joshua Villing
Ort Aarau
Datum 01.03.2022

Unterschrift