

# Gemeinsamer Schlussbericht zu dem Projekt „Einsatz mobiler KI-Sprach- technologien zur Qualitätssteigerung in der medizinischen Versorgung“ – HYKIST

Vorgelegt von Fraunhofer FOKUS, RWTH Aachen, Triaphon gGmbH, AppTek GmbH

Version	1.0
Laufzeit des Vorhabens	01.10.2020-31.12.2023
Förderkennzeichen BMG	ZMVI1-2520DAT04A
Zuwendungsempfänger	Fraunhofer FOKUS <sup>1</sup> , RWTH Aachen <sup>2</sup> , Triaphon gGmbH <sup>3</sup> , AppTek GmbH <sup>4</sup>
Projektleiter	Fraunhofer FOKUS
Berichtszeitraum	01.10.2020 - 31.12.2023
Erstellungsdatum	13.06.2024

Autoren: Dr. Michael John<sup>1</sup>, Thilo Ernst<sup>1</sup>, PD Dr. Ralf Schlüter<sup>2</sup>, Christoph Lüscher<sup>2</sup>,  
Tina Raissi<sup>2</sup>, Dr. med. Korbinian Fischer<sup>2,3</sup>, Dr. Volker Steinbiß<sup>4</sup>, Dr. Evgeny  
Matusov<sup>4</sup>

Gefördert durch:



aufgrund eines Beschlusses  
des Deutschen Bundestages

# Inhaltsverzeichnis

1	Zusammenfassung .....	4
2	Einleitung .....	4
3	Erhebungs- und Auswertungsmethodik.....	5
4	Detaillierte Darstellung der geleisteten Arbeit und Ergebnisse .....	10
4.1	AP 1: Anforderungsanalyse und Datenerfassung .....	10
4.1.1	Aktivitäten und Ergebnisse Fraunhofer FOKUS .....	10
4.1.2	Aktivitäten und Ergebnisse RWTH Aachen .....	11
4.1.3	Aktivitäten und Ergebnisse Triaphon gGmbH.....	14
4.1.4	Aktivitäten und Ergebnisse AppTek GmbH .....	16
4.2	AP 2: Systeme zur automatischen Spracherkennung .....	18
4.2.1	Aktivitäten und Ergebnisse RWTH Aachen .....	18
4.3	AP 3: Systeme zur maschinellen Übersetzung .....	21
4.3.1	Aktivitäten und Ergebnisse AppTek GmbH .....	21
4.3.2	Aktivitäten und Ergebnisse Fraunhofer FOKUS .....	25
4.3.3	Aktivitäten und Ergebnisse Triaphon gGmbH.....	29
4.4	AP 5: Integration zum Proof of Concept der Gesamtlösung .....	31
4.4.1	Aktivitäten und Ergebnisse Fraunhofer FOKUS .....	31
4.4.2	Aktivitäten und Ergebnisse Triaphon.....	37
4.4.3	Aktivitäten und Ergebnisse AppTek.....	39
4.5	AP 6: Praxistest, Evaluation und Projektabschluss .....	41
4.5.1	Aktivitäten und Ergebnisse Fraunhofer FOKUS .....	42
4.5.2	Aktivitäten und Ergebnisse RWTH Aachen .....	43

4.5.3	Aktivitäten und Ergebnisse Triaphon gGmbH.....	45
4.5.4	Aktivitäten und Ergebnisse AppTek GmbH .....	49
4.6	AP 7: Übergreifende Aktivitäten .....	51
	Aktivitäten und Ergebnisse Fraunhofer FOKUS .....	51
4.6.1	Aktivitäten und Ergebnisse RWTH Aachen .....	53
4.6.2	Aktivitäten und Ergebnisse Triaphon gGmbH.....	55
4.6.3	Aktivitäten und Ergebnisse AppTek GmbH .....	56
5	Erreichte Ergebnisse.....	59
6	Gender Mainstreaming Aspekte.....	61
7	Diskussion der Ergebnisse, Gesamtbeurteilung .....	61
8	Verbreitung und Öffentlichkeitsarbeit der Projektergebnisse .....	69
9	Verwertung der Projektergebnisse .....	69
10	Publikationsverzeichnis.....	72

# 1 Zusammenfassung

Im Projekt HYKIST wurde mit Technologien der Künstlichen Intelligenz ein echtzeitbasiertes Übersetzungsassistenzsystem in der medizinischen Basis- und Notfallkommunikation entwickelt. HYKIST assistiert dolmetschenden Sprachmittlerinnen und Sprachmittlern durch eine KI-unterstützte Kommunikation, um so komplexe medizinische Sachverhalte und Fachtermini besser während des Arzt-Patienten-Gesprächs übersetzen zu können. Hierfür wurden Technologien der automatischen Spracherkennung und maschinellen Übersetzung mit einem Dialogsystem zur Erstanamnese gekoppelt und in eine Telekonferenzplattform zur Unterstützung der Sprachmittlerinnen und Sprachmittler integriert. Zuerst wurden in dem Projekt umfangreiche Dialoge für die Sprache Arabisch sowie für Vietnamesisch erhoben, die den Grundstock für die Algorithmen- und Anwendungsentwicklung bilden. Innerhalb der Aufstockung HYKIST-PLUS wurden die im HYKIST-Projekt etablierten Verfahren dann auch auf die Bearbeitung des Ukrainischen erweitert. Bereits projektbegleitend erfolgten erste technische Tests hinsichtlich Akkuratheit und Qualität der automatisierten Übersetzungen. Das zu entwickelnde Gesamtsystem wurde abschließend in einem Pilottest im Sanaklinikum Berlin-Lichtenberg für den Bereich der Notfallaufnahmen und Erstanamnese in Akutsituationen erprobt und in einer abschließenden klinischen Studie hinsichtlich Nutzerakzeptanz evaluiert. Das Projekt wurde zum 01.10.2020 bewilligt und zum 31.12.2023 erfolgreich abgeschlossen.

## 2 Einleitung

In Deutschland leben 20,8 Mio. Menschen mit Migrationshintergrund (Statistisches Bundesamt, Mikrozensus 2018), davon haben laut Literatur 10-30% geringe Deutschkenntnisse (Brenne et al. 2015). Mangelnde oder fehlerbehaftete Kommunikation in der medizinischen Versorgung führt zu Unter- und Fehlversorgung, geringerer Versorgungsqualität, erhöhter Rate an Behandlungsfehlern und ineffektiveren Präventionsmaßnahmen für die Patienten und Patientinnen sowie zu Unzufriedenheit beim medizinischen Personal (Lampert et al. 2016, Pette et al. 2004). Von dieser Unter-/ und Fehlversorgung aufgrund von Sprachbarrieren sind Frauen und Kinder in besonderem Maße betroffen (Schouler-Ocak et al. 2017, Ullrich et al. 2016). In Studien wird die zunehmende Gewalt gegen medizinisches Personal in den Notaufnahmen u.a. auf Verständigungsprobleme zurückgeführt (GINA-Studie, „Krankenhausbarometer“ des Deutschen Krankenhausinstituts). Nur bedarfsgerecht jederzeit verfügbare, aber auch hochqualitative Dolmetsch-Dienstleistungen können dieses Problem lösen. Aktuelle Sprachmittler-Plattformen zur telefonischen Echtzeit-Vermittlung qualifizierter Sprachmittler und Sprachmittlerinnen konnten hier bereits erhebliche qualitative Verbesserungen bewirken. Allerdings müssen Sprachmittler

und Sprachmittlerinnen – u.a. auch aufgrund zu wenig verfügbaren universitär ausgebildeten Dolmetscher und Dolmetscherinnen - mit geschulten Sprachmittlern und Sprachmittlerinnen, für ihren Einsatz trainiert und fortlaufend weitergebildet werden. Hier können automatische Spracherkennungs- und maschinelle Übersetzungssysteme einen entscheidenden Beitrag für ein besseres Verständnis von Sprachmittlerinnen, Ärzten und Patienten liefern.

In den letzten Jahren gab es infolge theoretischer Durchbrüche in der Künstlichen Intelligenz, insbesondere auf dem Gebiet des Deep Learnings, sehr substanzielle Fortschritte auch bei der Entwicklung automatisierter Sprachverarbeitungs- und Übersetzungswerkzeuge und entsprechender Online-Dienste. Allerdings zeigt die aktuelle Studienlage auch immer noch eine hohe Fehleranfälligkeit von aktuell vorhandenen Übersetzungstools (wie bspw. Google Translate o.Ä.) auf, insbesondere in den derzeit in der medizinischen Versorgung in Deutschland benötigten Sprachen. Das Sprachproblem betrifft auch erste Systeme, die versuchen, durch KI die menschliche Anamnese (oder sogar weitere Aspekte der Triage) zu ersetzen. Abgesehen von anderen Schwächen bewirkt primär das Fehlen einer multilingualen Sprachschnittstelle, dass solche Systeme für nicht-deutschsprachige Patienten nicht nutzbar sind.

Aufgrund dieser Problemstellungen wurde im Projekt eine Kopplung der bereits existierenden Triaphon-Plattform zur Vermittlung menschlicher Dolmetsch-Leistung mit einer KI-basierten Übersetzungslösung untersucht. Diese Kopplung erfolgte bidirektional: Einerseits werden die menschlichen Dolmetscher durch maschinell erzeugte Übersetzungsvorschläge unterstützt, andererseits werden die bei der Nutzung von Triaphon durch Dolmetscher und Dolmetscherinnen auf laufender Basis manuell erzeugten Übersetzungsbeispiele der weiteren Verbesserung der Abdeckung und Qualität der KI-Übersetzungslösung zugänglich gemacht. Analog zu im Gebiet des (asynchronen) Übersetzens feststellbaren Trends wird also keine Vollautomatisierung angestrebt, sondern eine Symbiose zwischen maschineller Vor- und Detailarbeit und menschlich vorgenommener Auswahl, Integration und Qualitätssicherung. Flankierend wurde zur Kommunikationseinleitung und zur Erhebung initialer Anamnese-Informationen die Integration mit einer ebenfalls KI-basierten Sprachassistentenlösung untersucht und ebenfalls für die Nutzung durch nicht-deutschsprachige Patienten und Patientinnen optimiert.

### **3 Erhebungs- und Auswertungsmethodik**

Das Gesamtziel des Projekts war eine optimierte medizinische Versorgung von nicht-deutschsprachigen Patienten und Patientinnen durch Verbesserung der Kommunikation mittels eines hybriden Gesamtsystems, das menschliche Dolmetsch-Leistung mit einer KI-

basierten Übersetzungslösung kombiniert sowie KI-basierte Dialogassistentenfunktionen anbietet.

Konkret wurden dafür die folgenden Teilziele verfolgt:

- **Projektziel 1:** Hauptziel des Projekts war die Entwicklung einer hybriden (Mensch-KI) Dolmetsch-Lösung als Proof of Concept, welche Chancen und Fallstricke von Mensch und Maschine möglichst ergänzt bzw. aufhebt.
- **Projektziel 2:** Weiterhin wurde eine Integration dieser „Fremdsprachschnittstelle“ mit einem ebenfalls KI-basierten Anamnese-Assistenten verfolgt, womit als weiterer Proof of Concept eine integrierte Lösung für die Akutanamnese nicht-deutschsprachiger Patienten und Patientinnen entsteht. Das Vorhaben sah eine Partizipation folgender **Zielgruppen** vor: Patienten und Patientinnen mit Sprachbarriere, gesamtes medizinisches Personal (Pflegepersonal, Ärzte, Therapeuten, MTAs etc.) und bilinguale Sprachmittler.

Die im Arbeits- und Zeitplan definierten Meilensteine dokumentieren die zu erzielenden Ergebnisse des HYKIST-Projekts:

- M1 (nach 12 Monaten): Ein externer Evaluationspartner ist identifiziert und die Verhandlungen zu dessen Integration in HYKIST wurden erfolgreich abgeschlossen.
- M2 (nach 12 Monaten): Die initiale Integration der zentralen Komponenten des Systems wurde erfolgreich abgeschlossen.
- M3 (nach 23 Monaten): Die endgültige Integration, grundlegende Tests und eine erste praxisorientierte Evaluierung des integrierten Systems sind abgeschlossen.
- M4: (nach 34 Monaten): Die durch den externen Evaluationspartner durchzuführende klinische Studie ist abgeschlossen.

Zur Erreichung der Projektziele wurde im Projekt ein Projektmanagement eingerichtet, das die enge Verzahnung von Software-Entwicklungsprozessen und klinischen Prozessen forderte und fortwährend überprüfte. Für die Softwareentwicklung kamen agile Entwicklungsprozesse zur Anwendung. Durch den nutzerorientierten Entwicklungsprozess konnte frühzeitig auf Aspekte der Barrierefreiheit eingegangen werden. Frühzeitig wurden die entwickelten Prototypen auf Usability und Akzeptanz hin überprüft. Die Testergebnisse flossen zur Optimierung der Nutzerinteraktion in den technischen Entwicklungsprozess ein. Die Entwicklungsziele für die automatische Spracherkennung und maschinelle Übersetzung wurden fortwährend anhand eines Evaluationsdatensatzes überprüft und optimiert.

Das Studiendesign für die klinische Erprobung des HYKIST-Prototypen adressierte v.a. die Machbarkeit des entwickelten Systems. Hier wurden einerseits die Frameworks (Craig et al. 2008) des Medical Research Councils zur Entwicklung und Evaluation komplexer Interventionen sowie das Throughput-Modell der Versorgungsforschung nach Pfaff (Pfaff 2003, Schrappe et al. 2016) herangezogen. Die Ziele der Feasibilitystudie entsprachen denen in aktuellen Frameworks (Eldridge et al. 2016, Ukoumunne et al. 2015) dargestellten. Im Fokus standen hierbei Aspekte der Rekrutierung und Retention von Teilnehmenden, der Randomisierung, der Verblindung, der Festlegung und Bestimmung von Versorgungsrelevanten Ergebnisparametern für eine spätere Prüfung der Effektivität und Kosten-Effektivität sowie der Frage nach der Belastung der Beteiligten durch die Studiendurchführung und die Evaluation. Geplant war die Einführung des Systems in 3 kooperierenden Notaufnahmen für einen Zeitraum von drei Monaten. Die Feasibilitystudie wurde mit einer Stichprobe von 100-150 Anrufen geplant. Das Studiendesign wurde in Kooperation mit dem Institut für Pflegewissenschaften der Universität Köln als einem externen Evaluationspartner erarbeitet.

Die folgenden wissenschaftlichen Fragestellungen standen dabei im Fokus der Evaluation:

1. Wie sind die Usability und Akzeptanz des entwickelten Systems zu bewerten und welches Optimierungspotenzial besteht für das System? Stellt die echtzeitbasierte Übersetzung und Assistenz der Übersetzungsvorschläge eine Verbesserung für die Nutzer dar?
2. Inwieweit führen einzelne teilautomatisierte Dialoge des Erstanamneseassistenten, die auf der Systemlogik basieren zu einer höheren Akzeptanz sowohl bei den Patienten als auch dem medizinischen Fachpersonal?
3. Ändert sich das Nutzungsverhalten der Anwender über den Zeitraum der Systemnutzung und führt der Prozess der andauernden Feedbackgebung bzw. Bewertung der Übersetzungsvorschläge zu einer merkbaren Qualitätssteigerung im System?

Für die Auswertung der Studie wurden zum einen bereits erprobte Fragebögen zur Benutzerfreundlichkeit von digitalen Anwendungen (u.a. Telehealth Usability Questionnaire, System usability scale, etc.) eingesetzt, zum anderen wurden spezifische Fragebögen zur Anwendung des Systems entwickelt. Hierbei waren zur subjektiven Nutzenbewertung des Systems auch freie Antwortmöglichkeiten vorgesehen. Ergänzend wurden kurze leitfadengestützte Interviews durchgeführt. Die empirischen Daten wurden deskriptiv statistisch bzw. nach Methoden der qualitativen Sozialforschung (Inhaltsanalyse) ausgewertet. Die qualitative Sozialforschung ermöglicht mit Beobachtungs- oder Interviewtechniken möglichst nah im Lebensalltag die Nutzung digitaler Medien zu erfassen und zu

## HYKIST-Schlussbericht

dokumentieren. Weiterhin wurden die systemgenerierten Daten hinsichtlich der tatsächlich erfolgten Interaktion von Patienten, Sprachmittlern und Ärzten wie auch der Systemnutzung insgesamt deskriptiv ausgewertet.

Das Projekt wurde gemäß dem im Projektplan abgestimmten Projektplan durchgeführt:

Zeit- und Meilensteinplan		Jahr1				Jahr2				Jahr3							
		I	II	III	IV	I	II	III	IV	I	II	III	IV				
<b>1. Anforderungsanalyse und Datenerfassung</b>		Lead: Triaphon															
1.1	Klärung der Rahmenbedingungen inklusive Datenschutz und Ethikvotum	2,9	2,4														
1.2	Studiendesign sowie Festlegung von Bewertungsmetriken und Anforderungen an die Datenerhebung	5,15	4,4														
1.3	Datenerhebung und Annotierung	3,3	3,7	6	5,7	3,4	3,4	1,35	1,35								
<b>2. Systeme zur automatischen Spracherkennung</b>		Lead: RWTH															
2.1	Basissysteme automatische Spracherkennung Basissprachen	6	6	1,5	1,5	0,7	0,7										
2.2	Adaption automatische Spracherkennung an Zieldomäne			4,5	4,5	2	2	1,2	1,2	0,8	0,8	0,5	0,3				
2.3	Anbindung an maschinelle Übersetzung					1,75	2,25	2,25	2	2,5	1,5	0,95	0,4				
<b>3. Systeme zur maschinellen Übersetzung</b>		Lead: AppTek															
3.1	Basissysteme maschinelle Übersetzung Basissprachpaare	4	3,5	1	0,5												
3.2	Adaption maschinelle Übersetzung an Zieldomäne			2,5	2,5	1	0,8	0,4	0,4	0,1	0,1	0,1	0,1				
3.3	Systeme zur Simultanübersetzung					2,5	2,5	2,5	2,5	2,5	1,2	0,9	0,9				
<b>4. KI-basierter Erstanamnese-Sprachassistent</b>		Lead: FOKUS															
4.1	Domänenspezifische Vorgaben für Dialogschemata	0,5	1,5														
4.2	Prototypentwicklung einsprachig, Integration mit Spracherkennungssystem			2	1												
4.3	Prototypentwicklung mehrsprachig, Integration mit Übersetzungssystem			1	1	1											
4.4	Optimierung, Integration regelbasierter/statistische KI					1	1	1	1	1							
<b>5. Integration zum Proof of Concept der Gesamtlösung</b>		Lead: FOKUS															
5.1	User Interface, Usability		1,8	1,85	1,85	1,75											
5.2	Schnittstellen(API) Spracherkennung, Übersetzungssystem, Sprachassistent, Triaphon-Plattform			1	1	0,5	0,75										
5.3	Middleware und Anwendungslogik			1	1	1	1	1	1	1							
5.4	Systemintegration und Test					1,25	1,25	1,35	1,35	1,85	1,85	1,1					
<b>6. Praxistest, Evaluation und Projektabschluss</b>		Lead: RWTH															
6.1	Praxisorientierte Evaluation nach Erstintegration							4,25	4,25								
6.2	Klinische Studie inkl. Praxistest des entwickelten System-Prototypen										8	5,5					
6.3	Datenauswertung und Evaluation anhand der Ergebnisse und Daten der klinischen Studie											4,25	4				
6.4	Lessons learned, Formulierung der nächsten Schritte												4,75				
<b>7. Übergreifende Aktivitäten</b>		Lead: FOKUS															
7.1	Projektkoordination	2,05	1,6	1,1	1,2	1,1	1,1	1,2	1,1	3,6	1,35	1,85	2,05				
7.2	Dokumentation, Veröffentlichungen, Workshops, Standardisierung			0,75	0,85	0,85	1	0,35	0,35	1,1	0,85	0,95	1,05				
<b>Meilensteine</b>																	
M1	Klinischer Partner (N.N.) in das Konsortium integriert																
M2	Erstintegration abgeschlossen																
M3	Praxisorientierte Evaluation abgeschlossen																
M4	Klinische Studie abgeschlossen																
Gesamt-PM pro Quartal		23,9	24,9	22,2	22,6	18,8	18,8	16,9	16,5	14,5	15,7	16,1	13,6				
Gesamt-PM pro Kalenderjahr		23,9	88,5				66,6			45,3							

**2020:**

Im Jahr 2020 erfolgten im Rahmen von AP 1 erste Absprachen der Projektpartner zu den Anforderungen, Nutzungsszenarien und den rechtlichen Rahmenbedingungen. AppTek führte einen besonders zügigen Projektstart durch und setzte bereits im ersten Projektquartal Basissysteme zur Übersetzung auf, basierend zunächst auf textuellen Daten außerhalb des HYKIST-Projekts.

**2021**

In 2021 wurde im Rahmen von AP 1 die Anforderungsanalyse abgeschlossen, das Datenschutzkonzept finalisiert und entsprechende Datenschutzverträge unter den Partnern geschlossen. Ebenso wurden mehrere Stunden Trialoge zu Entwicklungszwecke aufgenommen und transkribiert. In AP 2 wurden die akustischen Modelle auf Basis der in AP 1 transkribierten Sprachdaten trainiert, evaluiert und fortlaufend optimiert. In AP 3 die Übersetzungsmodelle für die medizinische Domäne adaptiert und eine Erkennung und farbliche Markierung zentraler Begriffe umgesetzt. Ebenso wurde an dem Problem der automatischen Erkennung eines Sprecherwechsels (sog. Code-Switching) in der

Telefonkonferenz gearbeitet. In AP 4 wurden die Anforderungen an den Erstanamneseassistenten finalisiert und erste Dialoge inkl. der Benutzeroberflächen prototypisch implementiert. Im Rahmen des AP 5 wurde die HYKIST-App konzipiert und ebenfalls prototypisch implementiert, damit die automatischen Übersetzungsergebnisse den Dolmetschern angezeigt werden. Unter Berücksichtigung der Entwicklungsfortschritte der Partner wurde das Planungsdokument zur Koordination der HYKIST-Prototypen gepflegt. In AP 7 erfolgten projektbegleitende Aktivitäten sowie die Berichterstattung für die Monate 1-3 im Jahr 2020.

## **2022**

Im Jahr 2022 wurden die einzelnen Komponenten des HYKIST-Gesamtsystems weiterentwickelt, sodass über den Berichtszeitraum mehrere begleitende Anwendertestungen zur Überprüfung der Stabilität und Qualität des Systems erfolgen konnten. Die aus den Testungen resultierenden technischen und Nutzeranforderungen wurden sukzessive durch die Projektpartner implementiert, aufgetretene technische Probleme behoben, sodass für den Pilotbetrieb von einem zuverlässigen Betrieb der verteilten Komponenten und Datenflüsse ausgegangen werden kann. Dennoch musste der Meilenstein 3 in das Jahr 2023 verschoben werden, da sich die Entwicklungsarbeiten insbesondere im Sprachassistentenbackend aufgrund ihrer hohen Komplexität bei der Erkennung von arabischen Dialekten verzögerten und abschließende Testungen des Gesamtsystems erst in 2023 erfolgen können. In AP 2 und AP 3 wurden die akustischen und Übersetzungsmodelle auf Basis der in AP 1 transkribierten Sprachdaten für die Erkennung arabischer Dialekte optimiert. In AP 4 wurden die Programmlogik und die Bedienoberflächen des Erstanamneseassistenten entwickelt. In AP 5 wurde die HYKIST-App finalisiert und die Entwicklung und Testung des Gesamtsystems koordiniert. In AP 6 erfolgte die Konzeption des Studiendesigns und die Erstellung der notwendigen Dokumente zur Vorbereitung des Pilotbetriebs, in AP 7 die Berichterstattung und Teilnahme an Vernetzungsveranstaltungen sowie die Öffentlichkeitsarbeit für das Projekt. Ebenso erfolgten im Berichtszeitraum die Erweiterung des HYKIST-Systems um die ukrainische Sprache.

## **2023**

Aufgrund von notwendigen Optimierungen im Sprachassistenten-Backend (insbesondere an der Spracherkennung und maschinellen Übersetzung) sowie der umfangreichen Vorbereitung der Dokumente für die Ethikkommission erfolgte der Start der klinischen Phase für das HYKIST-Gesamtsystem mit einer Verzögerung von 3 Monaten im März 2023. Demzufolge wurde eine kostenneutrale Verlängerung des Projektes bis 31.12.2024 beantragt und bewilligt. Das HYKIST-Gesamtsystem wurde zum 16.03. im Sana-Klinikum in Berlin-Lichtenberg installiert und für die klinische Phase in Betrieb genommen. Aufgrund hoher Fehlerraten bei der

Spracherkennung und maschinellen Übersetzung erfolgten auch in 2023 parallel zu der klinischen Studie Optimierungsarbeiten in AP 2 und AP 3. Hierzu gehörten die Adaption vortrainierter akustischer Modelle, die Optimierung Suchraum/Dekodierer in Bezug auf Streamingmodus, die Reduktion der Latenz sowie ein verbessertes Glossarmanagement für die maschinelle Übersetzung. In AP 4 wurde auf Basis der Rückmeldungen aus der Pilotierung ein erweiterter Loggingmechanismus implementiert und Optimierungen am Bedieninterface durchgeführt. Ebenso wurden in AP 4 und AP 5 die Interfaces für die Anzeige von ukrainischen Sprach- und Übersetzungsstreams erweitert. Ebenso erfolgte in AP 5 der Pilotbetrieb des HYKIST-Systems. In AP 6 wurde die klinische Studie durchgeführt und anschließend die erhobenen Daten zur Usability, Akzeptanz und klinischen Machbarkeit ausgewertet. Das Projekt wurde am 31.12.2023 erfolgreich abgeschlossen.

## **4 Detaillierte Darstellung der geleisteten Arbeit und Ergebnisse**

Im Folgenden werden die detaillierten Arbeiten der Projektpartner nach Arbeitspaketen strukturiert. Die geleisteten Arbeiten werden jeweils in Jahresscheiben für die Projektlaufzeit 2020 – 2023 dargestellt.

### **4.1 AP 1: Anforderungsanalyse und Datenerfassung**

Dieses Arbeitspaket dient der Anforderungsanalyse für das zu entwickelnde Gesamtsystem. Ebenso werden die Methodiken für die durchzuführenden technischen und klinischen Studien definiert und die Beispieldatensätze erhoben.

#### **4.1.1 Aktivitäten und Ergebnisse Fraunhofer FOKUS**

##### **2020**

Fraunhofer FOKUS hat die Partner über die Wichtigkeit der datenschutzrechtlichen Aspekte des Projektes informiert und federführend die ersten Zuarbeiten zu dem Datenschutzkonzept koordiniert. Zur Klärung der Datenflüsse erstellte FOKUS einen ersten Entwurf der Gesamtsystemarchitektur als Diskussionsgrundlage. Gemeinsam mit den Partnern wurde das Dokument für die Nutzungs- und Anwendungskontexte des Gesamtsystems erstellt.

##### **2021**

Fraunhofer FOKUS hat federführend das Datenschutzkonzept koordiniert, das sämtliche Datenflüsse und Verarbeitungstätigkeiten im HYKIST-Gesamtsystem beschreibt. Zur Klärung der Datenflüsse finalisierte FOKUS die architektonische Beschreibung der Gesamtsystemarchitektur. Für die Analyse der Sicherheitsrisiken analysierte FOKUS die dezentrale Serverinfrastruktur der Partner und erstellte eine Datenschutzfolgeabschätzung (DSFA).

Ebenso koordinierte FOKUS den Vertragsabschluss einer gemeinsamen Verantwortung nach Art. 26 DSGVO, die die Verantwortung für die einzelnen Systemkomponente, Transportwege und Datensenzen offenlegt. Zur Überprüfung der Nutzungs- und Anwenderszenarien wurden erste Skizzen der Oberflächen für den Erstanamnese-assistenten (EAA) und die HYKIST-App erstellt und mit den Anwendern diskutiert.

## **2022**

Im Jahr 2022 erfolgten lediglich kleine Anpassungen an dem HYKIST-Datenschutzkonzept, da für die klinische Studie in AP 6 zusätzliche Daten erhoben werden sollen.

## **2023**

In AP 1 erfolgten im Jahr 2023 keine weiteren Arbeiten.

### **Ergebnisse 2020**

- Entwurf und Dokumentenversionen des Datenschutzkonzeptes
- Grafiken der Gesamtsystemarchitektur und Datenflüsse (siehe auch 5.5)
- Dokument zu Anwendungs- und Nutzungsszenarien des HYKIST-Systems

### **Ergebnisse 2021**

- Final abgestimmtes Datenschutzkonzept
- Finale Gesamtsystemarchitektur und Datenflüsse
- Analytierte Sicherheitsrisiken in der Datenschutzfolgeabschätzung (DSFA)
- Vereinbarung über eine gemeinsame Verantwortung nach Art. 26 DSGVO
- Powerpoint-Skizzen der Bedienoberflächen (Erstanamneseassistent und HYKIST-App)

### **Ergebnisse 2022**

- Anpassung des Datenschutzkonzeptes an die Anforderungen der klinischen Studie

#### **4.1.2 Aktivitäten und Ergebnisse RWTH Aachen**

##### **2020**

Die Anforderungsanalyse und das Gesamtsystemdesign wurden zusammen mit den Partnern erarbeitet. Die Transkription der im Projekt erhobenen Daten wurde vorbereitet, Transkriptionsrichtlinien erarbeitet und entsprechendes muttersprachliches studentisches Personal für die jeweiligen Sprachen angeworben. Die Bewertungsmetriken für die automatische Spracherkennung wurden spezifiziert und Darstellungsmöglichkeiten für die

erkannten Trialoge zusammengetragen. Weiterhin wurden Beiträge zum Datenschutzkonzept geleistet.

## **2021**

Die Anforderungsanalyse und das Gesamtsystemdesign wurden zusammen mit den Partnern erarbeitet. Die Transkription der im Projekt erhobenen Daten wurde vorbereitet, Transkriptionsrichtlinien erarbeitet und entsprechendes muttersprachliches studentisches Personal für die jeweiligen Sprachen eingeworben. Die Aufnahmebedingungen wurden in Abstimmung mit Triaphon dem realen Szenario angepasst. Das studentische Personal transkribierte die erhobenen Sprachdaten. Die Bewertungsmetriken für die automatische Spracherkennung wurden spezifiziert und Darstellungsmöglichkeiten für die erkannten Trialoge zusammengetragen. Weiterhin wurden Beiträge zum Datenschutzkonzept geleistet.

## **2022**

Die Aktivitäten zur Transkription aller Aufnahmen sind für Deutsch, Vietnamesisch und Arabisch abgeschlossen. Es wurden ungefähr 3 Stunden unsegmentierte deutsch-vietnamesische Audiodaten segmentiert und transkribiert.

Für das ukrainische System wurde muttersprachliches Personal eingestellt und mit deren Hilfe Dialoge und Trialoge im Umfang von sechs Stunden aufgenommen und transkribiert. Die Aufnahmen wurden in sieben Sitzungen, unter Beteiligung von 49 Freiwilligen (ukrainische Geflüchtete) und fünf Mitarbeitern in Aachen und Düren durchgeführt. Die ukrainischen Transkriptionsrichtlinien wurden von den Transkriptionsrichtlinien der anderen Sprachen adaptiert.

## **2023**

Die Aufnahmetätigkeiten für die ukrainische Sprache wurden erfolgreich abgeschlossen. Die Gesamtanzahl der beteiligten Freiwilligen belief sich auf 83 Personen. Zusätzlich haben die Mitarbeiter der RWTH aktiv an den Aufnahmen teilgenommen. Insgesamt wurden 13 Sitzungen sowohl in Aachen als auch in Düren durchgeführt, was zu 121 Aufnahmen mit einer Gesamtdauer von 10 Stunden unsegmentiert und 6,8 Stunden segmentiert führte.

## **Ergebnisse 2020**

- Klärung der Rahmenbedingungen
- Erläuterung des Spracherkennungssystems für die Projektpartner
- Vorbereitung der Transkription von Sprachdaten
- Suche nach weiteren Sprachdaten für das Training der Spracherkennungssysteme
- Beiträge zur Anforderungsanalyse und zum Systemdesign

- Definition der Bewertungsmetriken
- Beiträge zum Datenschutzkonzept

### **Ergebnisse 2021**

- Klärung der Rahmenbedingungen
- Erläuterung des Spracherkennungssystems für die Projektpartner
- Vorbereitung der Transkription von Sprachdaten
- Suche nach weiteren Sprachdaten für das Training der Spracherkennungssysteme
- Beiträge zur Anforderungsanalyse und zum Systemdesign
- Definition der Bewertungsmetriken
- Beiträge zum Datenschutzkonzept
- Transkription von Sprachdaten
- Bis zum 31.12.2021 wurden folgende Datenmengen transkribiert:
  - o Arabisch: 7,4 Stunden, 44,2k lfd. Wörter (111 Audiodateien),
  - o Deutsch: 11,0 Stunden, 111k lfd. Wörter (163 Audiodateien) und
  - o Vietnamesisch: 6,2 Stunden, 82,3 k lfd. Wörter (103 Audiodateien).

### **Ergebnisse 2022**

- Segmentierung und Transkription von drei Stunden unsegmentierten deutsch-vietnamesischen Audiodaten
- 83 Aufnahmen [Dialoge (UK), Trialoge (UK-DE)], 6:18h Rohaufnahmen davon 4:04h segmentiert, transkribiert und verifiziert
- 7 Sitzungen, unter Beteiligung von 49 Freiwilligen (ukrainische Flüchtlinge) und 5 Mitarbeitern in Aachen und Düren (Dauer der Aufnahmesessions ca. ~24 Stunden).

### **Ergebnisse 2023**

- Segmentierung, Transkription, und Verifizierung von zwei Stunden unsegmentierten ukrainischen Audiodaten
- 38 Aufnahmen [Dialoge (UK), Trialoge (UK-DE)], 3.82h Rohaufnahmen davon 2:16h segmentiert, transkribiert und verifiziert
- Sitzungen, unter Beteiligung von 34 Freiwilligen (inklusive ukrainische Flüchtlinge) und 5 Mitarbeitern in Aachen und Düren (Dauer der Aufnahmesessions ca. ~15 Stunden).

### 4.1.3 Aktivitäten und Ergebnisse Triaphon gGmbH

#### **2020**

Triaphon hat die technischen (zweite Telefonanlage) sowie die organisatorischen Voraussetzungen (Rekrutierung der beteiligten Personen, Einwilligungserklärungen, Datenschutzaspekte, Terminvereinbarungen) geschaffen, um simulierte gedolmetschte Arzt-Patienten-Gespräche aufzuzeichnen. Mittels Cloud-Telefonie werden drei Konferenzteilnehmer (Arzt-Patient-Sprachmittler) zusammengeführt und aufgezeichnet (wav-Datei). Im AP 1 wurde in enger Zusammenarbeit mit der RWTH Aachen University die Audiodaten geteilt, um diese dann durch die RWTH Aachen auswerten zu lassen (Transkription etc.). Mittels enger Feedbackschleifen wurde dieser Prozess optimiert und Feinheiten der Aufnahmetechnik und des Aufnahmesettings angepasst.

#### **2021**

Triaphon hat nach Klärung der Rahmenbedingungen und Schaffung der technischen Voraussetzungen simulierte gedolmetschte Arzt-Patienten-Gespräche (arabisch/deutsch und vietnamesisch/deutsch) aufgezeichnet. Die Gespräche wurden anhand vorher angefertigter Skripte bzw. skizzierten Gesprächssituationen geführt. Die Dialoge wurden mit verschiedenen Sprechern in allen beteiligten Rollen (Arzt, Patient, Sprachmittler) aufgenommen. Hier wurde auch auf einzelne Dialekte (insbesondere bei Vietnamesisch) geachtet und zudem Besonderheiten wie „das Einsprechen mit Maske“ berücksichtigt. Die Aufnahmen hatten eine durchschnittliche Aufnahmedauer pro Dialog von etwa 10 Minuten.

#### **2022**

Im Zuge von HYKIST-Plus hat Triaphon den neuen Sprachpool „Ukrainisch“ aufgebaut. Hierzu wurden insgesamt 46 Sprachmittler rekrutiert, geschult und im weiteren Betrieb betreut. Sie wurden angeleitet medizinische Gespräche zu dolmetschen und konnten über die „Dolmetsch-Nothilfe: Ukrainisch“ bis Ende 2022 über 3.500 Gespräche am Telefon ad hoc dolmetschen.

#### **2023**

Triaphon hat seinen kostenlosen-Dolmetsch-Service für Ukrainisch bis zum 31.01.23 erfolgreich fortgeführt. Zudem hat Triaphon die Aufnahme-Tätigkeit für HYKIST-Plus der RWTH Aachen unterstützt (technisch, Material etc.).

### **Ergebnisse 2020**

- Separate Hykist-Triaphon-Telefonanlage zur Aufnahme konfiguriert (Server und Testsystem)
- Klärung der datenschutzrelevanten Fragestellungen

## HYKIST-Schlussbericht

- Einholung der Einwilligungserklärungen aller Beteiligter
- Verfassen der ersten Trialoge für die Aufnahmen
- Rekrutierung und Koordinierung von teilnehmenden Personen
- Bis 31.12.2020 erfolgten 34 Aufnahmen, welche 5 Stunden und 28 Minuten Aufnahmezeit ergaben; davon anteilig die Hälfte in deutscher Sprache, ein Viertel auf Arabisch und ein weiteres Viertel auf Vietnamesisch

**Ergebnisse 2021**

- Verfassen von diversen Trialogen bzw. Vorgabe von Gesprächsszenarien für die Sprecher
- Rekrutierung und Koordinierung von teilnehmenden Personen
- Aufnahme von über 300 simulierten Arzt-Patienten-Gesprächen in Arabisch-Deutsch oder Vietnamesisch-Deutsch
- Rohmaterial an Audiodaten von insgesamt über 30 Stunden für je Arabisch und Vietnamesisch. Für Deutsch ergab sich ungefähr der doppelte Anteil.
- Bis Ende November 2021 konnte die Datenerhebung abgeschlossen werden.

**Ergebnisse 2022**

- Aufbau, Schulung und Betreuung des neuen Sprachpools „Ukrainisch“ inklusive Vor-Ort-Schulung (mit Translationswissenschaftlerin) sowie Einzelschulungen per Telefon (HYKIST-Plus)
- Verdolmetschung von über 3.500 medizinischen Gesprächen im Jahr 2022
- Registrierung von über 600 medizinischen Einrichtungen bei "Dolmetsch-Nothilfe: Ukrainisch"; Anrufe zu über 90% von niedergelassenen oder ehrenamtlich tätigen Ärzten

**Ergebnisse 2023**

- Anfertigen und Pflege von Vokabellisten für Ukrainisch und Russisch
- Erfolgreicher Projekt-Abschluss der kostenlosen Dolmetsch-Hilfe Ukrainisch zum 01.02.2023

#### 4.1.4 Aktivitäten und Ergebnisse AppTek GmbH

##### **2020**

AppTek hat intensiv mit den anderen Partnern zusammengewirkt und an der Anforderungsanalyse und dem Systemdesign gearbeitet. Die Bewertungsmetriken für MT wurden spezifiziert und die Möglichkeiten der Interaktion des MT-Systems mit der Sprachmittlerin (Liste von Übersetzungen etc.) durchdacht.

##### **2021**

AppTek hat die von Triaphon transkribierten Trialoge verwertet, indem diese Daten als so genannte "Seed"-Daten genutzt wurden, um ähnliche deutsch-arabische und deutsch-vietnamesische parallele Textdaten zu finden. Diese Daten wurden im 5.3 für die Adaption der Übersetzungssysteme verwendet. Die von Triaphon zur Verfügung gestellten Wortlisten und Glossare wurden direkt ins Training der Systeme integriert.

Ferner wurden die transkribierten Trialoge verwendet, um zu testen, wie die automatische Erkennung funktioniert, welche Sprache (z.B. Deutsch oder Arabisch) gerade auf demselben Audio-Kanal gesprochen wurde.

Was die Datenspeicherung angeht, so wurden Maßnahmen gemäß dem Datenschutz-konzept getroffen, dass die Daten sicher gespeichert und abgerufen werden können. Dies gilt sowohl für Audiodaten, als auch für Transkriptionen, Übersetzungen und unterstützendes Material wie Sprachmittler-spezifischen Glossare.

##### **2022**

AppTek hat weiterhin die von Triaphon transkribierten Trialoge verwertet, indem diese Daten als so genannte "Seed"-Daten genutzt wurden, um ähnliche deutsch-arabische und deutsch-vietnamesische parallele Textdaten zu finden. Für HYKIST-Plus wurden die Ukrainisch-Deutsche Trialoge auf gleiche Weise verwendet. Ferner wurden die transkribierten Trialoge verwendet, um zu testen, wie die automatische Erkennung funktioniert, welche Sprache (z.B. Deutsch oder Arabisch) gerade auf demselben Audio-Kanal gesprochen wurde.

##### **2023**

AppTek hat für die Übersetzungssysteme Deutsch<->Vietnamesisch und Deutsch<->Ukrainisch künstliche Textdaten erzeugt, in dem GPT4 gefragt wurde, Dialoge zwischen einem deutschen Arzt/einer deutschen Ärztin und einem Patienten oder Patientin zu erzeugen. Dabei wurde in der Anweisung erwähnt, dass die Patienten aus dem Ausland kommen und der komplette Dialog auch ins Vietnamesische bzw. Ukrainische automatisch übersetzt werden soll. Die Anweisung beinhaltete auch ein Gesprächsthema (ein bestimmtes

gesundheitliches Leiden). Die so erzeugten parallelen Daten beinhalteten mehr als 500 kurze Dialoge. Die Inhalte der Dialoge waren den Inhalten der echten Dialoge ähnlich, so dass diese Daten danach als Adaptionen- bzw. Trainingsdaten für die Verbesserung der Übersetzungssysteme verwendet wurden. Ein Teil der Daten wurde beiseitegelegt, damit sie als zusätzliche Testdaten verwendet werden konnten. Die Übersetzungsqualität von GPT4 wurde in diesen Dialogen als gut bewertet, sodass die Daten verwendet werden konnten. Es wurde darauf geachtet, dass in diesen künstlichen Dialogen keine Personeninformationen enthalten sind, weder in den Anweisungen an GPT4 noch in den Antworten des Systems.

Diese künstlichen Daten wurden von AppTek an RWTH Aachen weitergeleitet, damit sie auch für die bessere Sprachmodellierung im automatischen Spracherkennungssystem genutzt werden können.

### **Ergebnisse 2020**

- Klärung der Rahmenbedingungen
- Erläuterung der für die Partner relevanten Aspekte des Übersetzungsmoduls
- Beiträge zur Anforderungsanalyse und zum Systemdesign, dabei auch Anforderungen hinsichtlich personalisierter Unterstützung (Vokabelliste) und Auswirkungen auf das Datenschutzkonzept
- Bewertungsmetriken
- Anteil am Datenschutzkonzept

### **Ergebnisse 2021**

- Erfolgreiche Verwendung der simulierten Dialoge und deren Transkriptionen
- Erfolgreiche Verwendung anderer Textdaten, die von Triaphon bereitgestellt wurden (medizinische Begriffe, andere Daten aus dem Kontext der Sprachmittlung in Krankenhäusern)
- Technische Implementierung des Datenschutzkonzepts

### **Ergebnisse 2022**

- Erfolgreiche Verwendung der simulierten Dialoge und deren Transkriptionen

### **Ergebnisse 2023**

- Erfolgreiche Konzipierung und Verwendung der künstlichen Dialoge, die von einem großen Sprachmodell (GPT4) erstellt wurden

## 4.2 AP 2: Systeme zur automatischen Spracherkennung

In AP 2 werden die projektspezifischen Anforderungen an die automatische Spracherkennung implementiert.

### 4.2.1 Aktivitäten und Ergebnisse RWTH Aachen

#### **2020:**

Die RWTH hat transkribierte telefonsprachliche Audiodaten für das Training der Spracherkennungssysteme für die Sprachen Deutsch, Arabisch und Vietnamesisch sowie weitere Wissensquellen (insb. Aussprachelexika) zusammengetragen und vorverarbeitet. Trainingskonfigurationen wurden erstellt und implementiert und das Training der akustischen Modelle erster Basissysteme wurde für alle drei im Projekt vorgesehenen Sprachen (Deutsch, Arabisch, Vietnamesisch) gestartet.

#### **2021:**

Die akustischen Modelle wurden auf den erhobenen Sprachdaten evaluiert. Die akustischen Modelle der Basissysteme wurden verbessert und optimiert, indem die Fehler der Systeme analysiert wurden und Lösungen entwickelt wurden. Während der Fehleranalyse wurde festgestellt, dass in den Sprachdaten Wörter auftreten, die nicht der Originalsprache entstammen. Für die Basissysteme könnte diese Fehlerquelle ein ernsthaftes Problem darstellen. Daher wurden erste Überlegungen und Ideen für ein System ausgearbeitet, dass diese spezifischen Worte, die nicht aus der Originalsprache stammen, berücksichtigt.

Das arabische Spracherkennungssystem verwendet ein Conformer-basiertes akustisches Modell mit Wortfehlerrate (WER) von 38,9%. Das Sprachmodell hat eine Perplexität von 1500. Für Deutsch und Vietnamesisch wurden mit BLSTM-basierten akustischen Modellen WER-Werte von 24,7% bzw. 34,2% erreicht. Die jeweilige Perplexität des Sprachmodells beträgt 71 bzw. 81 erzielten Wortfehlerraten (WER) aufgelistet.

#### **2022:**

Die Sprachmodelle der automatischen Spracherkennungssysteme für die Sprachen Deutsch, Vietnamesisch und Arabisch wurden mit medizinspezifischen Begriffen erweitert und adaptiert. Neue akustische Modelle auf Basis der Transformer Architektur wurden für das Deutsche und Vietnamesische entwickelt und in das Produktionssystem integriert. Die Suchraumparameter und die Segmentierung wurden für den Anwendungsfall unter Streaming-Bedingungen optimiert.

**HYKIST+:** Ein erstes Spracherkennungssystem mit akustischer Modellierung auf Basis bidirektionaler LSTM rekurrenter neuronaler Netze wurde in das Produktionssystem überführt

und befindet sich im Einsatz. Desweiteren steht auch bereits ein akustisches Modell für das Ukrainische basierend auf der Conformer Architektur zur Verfügung.

## **2023**

Entwicklung eines akustischen Modells für das Vietnamesische basierend auf der Conformer-Architektur. Die Suchraumparameter wurden für den Anwendungsfall unter Streaming-Bedingungen weiter optimiert.

Für das Arabische wurde das NLLB-Meta-Modell verwendet, um von MSA (Modern Standard Arabic) in den Dialekt zu übersetzen, und anschließend wurden diese Daten genutzt, um ein Sprachmodell zu trainieren. Die MSA-Daten enthielten u.a. Texte aus dem medizinischen Bereich.

**HYKIST+:** Die endgültige Version der transkribierten Daten wurde in Dev-, Test- und Adaptationssets unterteilt. Unter Verwendung der Dev-Daten wurden die Decodierparameter für das auf der Conformer-Architektur basierende Modell optimiert. Mit den Adaptationsdaten wurde ein neues Sprachmodell trainiert. Für das abschließende Sprachmodell und das Decoding-Lexikon wurde eine Übermenge des Baseline-Sprachmodells und Lexikons berücksichtigt. Das Lexikon wurde um Wörter erweitert, die nicht im Grundvokabular enthalten waren.

## **Ergebnisse 2020**

- Zusammenstellung Korpora für das Training akustischer Modelle (Deutsch, Arabisch, Vietnamesisch)
- Zusammenstellung von Aussprachelexika für alle beteiligten Sprachen
- Erarbeitung von Trainings- und Evaluationskonfigurationen
- Beginn des Trainings erster Basissysteme für deutsche, arabische und vietnamesische Telefonsprache
- Erforschung geeigneter Spracherkennungsarchitekturen

## **Ergebnisse 2021**

- Zusammenstellung von Korpora für das Training akustischer Modelle (Deutsch, Arabisch, Vietnamesisch)
- Zusammenstellung von Aussprachelexika für alle beteiligten Sprachen
- Erarbeitung von Trainings- und Evaluationskonfigurationen
- Beginn des Trainings erster Basissysteme für deutsche, arabische und vietnamesische Telefonsprache

- Erforschung geeigneter Spracherkennungsarchitekturen
- Entwicklung, Training und Evaluierung von akustischen Modellen
- Integration der akustischen Modelle ins AppTek System

### Ergebnisse 2022

- Transformer-basierte akustische Modelle für Deutsch und Vietnamesisch
- Adaption der Sprachmodelle an medizinische Domäne/medizinische Termini
- Adaption vor-trainierter akustischer Modelle für Deutsch, Vietnamesisch und Arabisch
- Optimierung Suchraum/Dekodierer in Bezug auf Streamingmodus
- Reduktion der Latenz ohne signifikanten Einfluss auf die Fehlerraten
- Die erreichten Wortfehlerraten (WER) der Spracherkennungssysteme im Streamingmodus betragen für Deutsch 20.0%, Vietnamesisch 37.7% sowie Arabisch 41.1%, wobei die Zeichenfehlerrate (CER) für das Arabische 20.7% beträgt.
- Das Vokabular besteht aus 202K (Deutsch), 11k (Vietnamesisch), und 370k (Arabisch) Wörtern. Die *Out-of-Vocabulary* (OOV) Raten von den Vokabularien nicht abgedeckter Wörter der Testdaten betragen 1.6% (Deutsch), 0.2%(Vietnamesisch) sowie 1.0% (Arabisch).
- Die Perplexität der Sprachmodelle beträgt 78 (Deutsch), 69 (Vietnamesisch) sowie 972 (Arabisch).
- **HYKIST+:**
  - o Zusammenstellung von Korpora für das Training akustischer Modelle für die ukrainische Spracherkennung
  - o Entwicklung, Training und Evaluierung von akustischen Modellen für das ukrainische System, die eine Wortfehlerrate von 39.8% liefern. Das Vokabular besteht aus 141K Wörtern, die OOV-Rate beträgt 7.7%. \Das Sprachmodell hat ein Perplexität 669.

\*Zeichenfehlerrate (CER) Arabisch

### Ergebnisse 2023

- Erstellung eines akustischen Modells für das Vietnamesische unter Verwendung der Conformer-Architektur. Die Wortfehlerrate (WER) wurde mit dieser Architektur auf bis zu 29.0% reduziert
- Die Dialektzuordnung im Fall des arabischen Systems führte zu einer Verbesserung von 40.4% auf 38.9% WER auf Testdaten für die Offline-Spracherkennung (ASR).
- Optimierung Suchraum/Dekodierer in Bezug auf Streamingmodus
- **HYKIST+:**

- Aufteilung in Entwicklungs-, Test- und Adaptationsdaten.
- Evaluierung von Conformer-basierten akustischen Modellen für das ukrainische System.
- Vereinheitlichung des Vokabulars für akustische und Sprachmodelle.
- Entwicklung eines Sprachmodells unter Verwendung der Adaptationsdaten.
- Evaluation aktueller Optimierungen (Wortfehlerraten auf HYKIST Testdaten)
- Erzielte Wortfehlerrate (WER): 24.4%, Vokabulargröße: 291K Wörter, Rate unbekannter Wörter in den Testdaten (OOV): 3.2, Sprachmodellperplexität: 265.

### 4.3 AP 3: Systeme zur maschinellen Übersetzung

In AP 3 werden die projektspezifischen Anforderungen an die maschinelle Übersetzung implementiert.

#### 4.3.1 Aktivitäten und Ergebnisse AppTek GmbH

##### 2020:

Es waren MT-Systeme (Systeme zur maschinellen Übersetzung, *Machine Translation*) für die Übersetzungsrichtungen Deutsch → Arabisch, Arabisch → Deutsch, Deutsch → Vietnamesisch und Vietnamesisch → Deutsch zu trainieren, die neben einer grundlegenden Performanz im Allgemeinen insbesondere für das Patientengespräch in der Notaufnahme geeignet sind. Die Herausforderung liegt hierbei in der Gesprächsdomäne: MT-Systeme werden üblicherweise mit einer großen Menge von übersetzten Texten trainiert, die hier weder vorhanden sind noch es in großer Menge sein werden, und das dadurch gestellte technische Problem gilt es zu lösen. AppTek hat im Zeitraum m01 – m03 initiale MT-Systeme trainiert und domänenspezifische Daten akquiriert. Dabei wurden auch schon erste Adaptionen der MT an die Zieldomäne durchgeführt. Es sei darauf verwiesen, dass das MT-System auf eine Verwendung in einer Pipeline *automatische Spracherkennung* → *MT* hin von Grund auf angelegt ist und dass sich die Übersetzung gesprochener Sprache im Vergleich zur Übersetzung geschriebener Texte durch Besonderheiten auszeichnet. Insbesondere haben wir für die Sprachen Deutsch, Vietnamesisch und Arabisch unsere Textnormalisierungssoftware geschrieben bzw. aktualisiert. Es handelt sich u.A. um die Konvertierung von Zahlenfolgen, Datumsangaben, Angaben von Medikamentenmengen und Messungsergebnissen von der geschriebenen Form mit Ziffern in die gesprochene Form.

Zusätzlich sind die Ergebnisse der Forschung zur simultanen Übersetzung gesprochener Sprache in das Projekt geflossen. Es geht um die Anwendung, d.h. wenn eine partielle

Übersetzung schon geliefert wird, während der Sprecher / die Sprecherin den Satz noch nicht zu Ende gesprochen hat. Diese Anwendung ist besonders dann wichtig, wenn bestimmte Phrasen aus dem medizinischen Bereich den Sprachmittlern als Übersetzungen sehr schnell vorgeschlagen werden können, damit es keine lange Pause zwischen Äußerungen der Patientin / des Patienten und der mündlichen Übersetzung entsteht.

**2021:**

In diesem Zeitraum wurden von AppTek MT-Systeme (Systeme zur maschinellen Übersetzung, *Machine Translation*) für die Übersetzungsrichtungen Deutsch → Arabisch, Arabisch → Deutsch, Deutsch → Vietnamesisch und Vietnamesisch → Deutsch trainiert und auf die medizinische Domäne adaptiert. Dabei wurde die Struktur der MT-Modelle so angepasst, dass eine Markierung von wichtigen Begriffen in der Quellsprache (anhand von Glossaren) dazu führt, dass die entsprechenden Begriffe in der Übersetzung des Satzes in der Zielsprache ebenfalls markiert werden. Diese Funktionalität ist wichtig, um in der HYKIST-App die wichtigsten Begriffe und deren Übersetzungen farblich hervorzuheben bzw. nur diese einzublenden. Außerdem wurden aus den existierenden, öffentlich verfügbaren parallelen Daten solche Daten für die Adaption der Systeme ausgewählt, die am ähnlichsten zu den im AP1 gesammelten simulierten Gesprächen sind. Dies geschah mit mehreren automatischen Verfahren der textbasierten Datenselektion, inklusive Verfahren, die eine semantische Repräsentation der Wörter und Sätze mit Hilfe von neuronalen Netzen unterstützen.

Die Schnittstelle zum HYKIST-Backend wurde erweitert, um Sprachmittler-spezifische Glossare zu unterstützen. Die Abgleichung der Glossar-Einträge mit einem gesprochenen Satz wurde implementiert und getestet. Es wurde auch sichergestellt, dass die Glossar-Übersetzungen vom MT-System übernommen werden.

Viel Zeit wurde von AppTek investiert, um das Problem zu lösen, die Identität der Quellsprache zu bestimmen. Da sowohl die Patienten als auch die Ärzte über denselben Telefonkanal sprechen, muss in Echtzeit, nach jeder Äußerung, automatisch entschieden werden, welche Sprache (Deutsch oder Fremdsprache) gesprochen wurde, um die in dieser Sprache erkannte Wortfolge an das entsprechende Übersetzungssystem zu senden. Als erste Lösung dieses Problems wurde vorgeschlagen, zwei Erkennungssysteme (für Deutsch und Arabisch bzw. Vietnamesisch) parallel laufen zu lassen und dann segmentweise die richtige Sprache zu bestimmen. Dies wurde umgesetzt; die richtige Sprache bestimmt man durch Mittelung aller Konfidenzwerte, die ein ASR-System für jedes Wort ausgibt. Es wurde viel mit simulierten Dialogen experimentiert, um eine Normalisierung der Konfidenzwerte auszuwählen, die zu einer höheren Präzision der Bestimmung der richtigen Sprache führt.

**2022:**

In diesem Zeitraum wurden die AppTek-MT Systeme (Systeme zur maschinellen Übersetzung) weiter verbessert, sowohl durch Verwendung von mehr Daten, die für den medizinischen Kontext passend sind, als auch durch eine komplexere Architektur der Modelle. Die Funktionalität der Übernahme der Übersetzungen wichtiger Begriffe wurde erheblich verbessert durch besseres Matching der Begriffe, auch wenn sie in einer anderen morphologischen Form (z.B. plural vs. Singular) im Eingabesatz zu finden waren. Die Systeme wurden auch so trainiert, dass sie mit Code-Switching (Verwendung deutscher Begriffe in der Fremdsprache) klarkommen und trotzdem eine richtige Übersetzung liefern, die diese Begriffe enthält.

Basierend auf Ergebnissen der Anwendertestungen wurden weitere Verbesserungen vorgenommen. Insbesondere wurde die Komponente verbessert, die die gesprochenen Zahlen und Zahlenfolgen, Datumsangaben usw. in die schriftliche Form mit Ziffern konvertiert.

Für HYKIST+ wurden die Basis-Übersetzungssysteme für Deutsch<->Ukrainisch, Russisch auf den Trainingsdaten optimiert, die einen Bezug zu Gesprächsinhalten haben, die den Bereich Gesundheit betreffen. Dazu wurden aus den existierenden, öffentlich verfügbaren parallelen Daten solche Daten für die Adaption der Systeme ausgewählt, die am ähnlichsten zu den im AP1 gesammelten simulierten Gesprächen sind. Dies geschah mit mehreren automatischen Verfahren der textbasierten Datenselektion, inklusive Verfahren, die eine semantische Repräsentation der Wörter und Sätze mit Hilfe von neuronalen Netzen unterstützen. Die Daten wurden im Training so balanciert, dass eine gute Übersetzungsqualität sowohl für Ukrainisch als auch Russisch erreicht wird.

Die automatische segmentweise Bestimmung der jeweils gesprochenen Sprache (von Patienten/Patientin oder von einem Arzt/einer Ärztin) wurde ebenfalls verbessert durch eine verbesserte Normalisierung und Kalibrierung der Konfidenzwerte der zwei parallellaufenden Spracherkennungssysteme für Deutsch und die Fremdsprache.

**2023**

Der Verlauf der internen App-Testungen im November 2022 und Februar 2023 wurde von AppTek in Bezug auf die Qualität der maschinellen Übersetzung ausgewertet. Es wurden kleine Mängel in der Textdarstellung durch Postprocessing-Regeln verbessert. Die Qualität der Übersetzung ins Deutsche wurde als hinreichend gut bewertet, vorausgesetzt dass es keine Spracherkennungsfehler in der Eingabe gibt. Die Qualität der Übersetzung ins Arabische wurde ebenfalls als gut bewertet. Die Qualität der Übersetzung ins Vietnamesische wurde als nicht hinreichend gut bewertet, wegen der hohen Mehrdeutigkeit der Woreinheiten in der vietnamesischen Sprache. Diese Einheiten entsprechen eher Silben statt Wörtern. Es wurde ein neues Training für dieses Sprachpaar durchgeführt, bei welchem die Option eingesetzt

wurde, mehrere häufig zusammen auftretende Einheiten zu einer Einheit (Token) zusammenzuführen. Zusätzlich wurden mehr Domäne-relevanten parallelen Trainingsdaten verwendet, inklusive die im AP 4.1 erzeugten künstlichen Dialoge zu medizinischen Themen. Das alles hat zur signifikanten Verbesserung der Übersetzungsqualität ins Vietnamesische geführt.

Für HYKIST+ wurde das MT-System für die Übersetzung ins Ukrainische und Russische ebenfalls verbessert durch ein neues Training mit Hinzunahme der künstlich erzeugten Dialoge als parallele Daten. Ferner wurde die Glossar-Funktion der HYKIST+ Übersetzungssysteme verbessert.

### **Ergebnisse 2020**

- Basissysteme zur maschinellen Übersetzung (s.o.)

### **Ergebnisse 2021**

- Systeme zur maschinellen Übersetzung, die für die HYKIST-Domäne (medizinischer Kontext) adaptiert sind und das Hervorheben der wichtigen Begriffe unterstützen
- Unterstützen von Glossaren und deren Verwendung in MT-Systemen
- Eine Softwarekomponente in der Schnittstelle zwischen Erkennung und Übersetzung, die die gesprochene Sprache automatisch erkennt (s.o.)

### **Ergebnisse 2022:**

- Verbesserte (neu trainierte) Systeme zur maschinellen Übersetzung, die für die HYKIST-Domäne (medizinischer Kontext) adaptiert sind, das Hervorheben der wichtigen Begriffe unterstützen, und mit Code-Switching umgehen können.
- Verbesserte Suche für Glossar Matches bei Verwendung des Glossars
- Verbesserte segmentweise Erkennung der gesprochenen Sprache
- HYKIST+: Adaptierte Systeme zur maschinellen Übersetzung, die gute Übersetzungsqualität in der Domäne der gesundheitsrelevanten Gespräche aufweisen und gleichermaßen gut von Ukrainisch und Russisch nach Deutsch übersetzen können und umgekehrt.

### **Ergebnisse 2023**

- Verbessertes (neu trainiertes) automatisches Übersetzungssystem "Deutsch nach Vietnamesisch"
- Kleine Verbesserungen bei den anderen laufenden Übersetzungssystemen
- Erweiterung der Glossare, Verbesserung der Glossarfunktionen
- Verbesserte Übersetzungssysteme für HYKIST+ (Deutsch <-> Ukrainisch, Russisch)

- Evaluierung der Übersetzungsqualität durch Sprachmittler-Feedback und Anwender-Feedback (manuelle Evaluierung auf echten Dialogen)

#### **AP 4: KI-basierter Erstanamnese-Assistent**

In diesem AP wird ein KI-basierter Sprachassistent für das Anwendungsszenario der Kontaktaufnahme und Erstanamnese entwickelt.

##### 4.3.2 Aktivitäten und Ergebnisse Fraunhofer FOKUS

###### **2020:**

Fraunhofer FOKUS hat mit den Partnern erste Anforderungen an den KI-basierten Erstanamnese-Sprachassistenten (EAA) erarbeitet und abgestimmt.

###### **2021:**

Fraunhofer FOKUS hat mit Triaphon die Anforderungen an den KI-basierten Erstanamnese-Sprachassistenten (EAA) finalisiert und mit den Partnern RWTH Aachen und AppTek die technischen Schnittstellen zu der Übersetzungsassistenz definiert. Hierzu gehörte das Interaktions- und Webdesign für die EAA WebApp basierend auf dem Anwendungsszenario und dem Dialogschema. In Zusammenarbeit mit Triaphon wurde ein grafisch-interaktiver Body Selector implementiert, der den Nutzern eine intuitive Erfassung der Schmerzen (u.a. Körperregion und Schmerzintensität) entwickelt. Für das Deployment des EAA auf einem Kiosk-Terminal erfolgten Vorarbeiten zur Paketierung der EAA-WebApp als native Electron-Anwendung. Um neue Anforderungen effizient zu verwalten und die Entwicklungsarbeiten zu strukturieren, wurde ein Issue-Tracking-System für EAA und EAAS aufgebaut. Begleitend zu den Entwicklungsarbeiten erfolgten ausführliche Einzeltests der Komponenten. Für die Testung des Zusammenspiels des EAA mit den anderen Systemkomponenten im HYKIST-Gesamtsystem wurden Integrationstests der Implementierungen von EAA und EAAS zusammen mit den kollaborierenden Komponenten HYKIST-Backend und Sprachassistenten-Backend durchgeführt. Zur Optimierung der Testabläufe wurde eine automatisierte Testinfrastruktur aufgebaut. Für den Betrieb im Rahmen der abschließenden Evaluation wurde für den EAA und den EAAS ein Cloud-Betriebskonzept sowie eines CI/CD-Konzept entworfen, um erforderliche Änderungen auch zu Laufzeit zu ermöglichen.

###### **2022:**

Die Entwicklung des EAA wurde intensiv fortgeführt. Durch Integration der vom Partner AppTek bereitgestellten Sprachservices (Spracherkennung ASR und Sprachsynthese TTS) wurden Sprachschnittstellen für alle HYKIST-Zielsprachen realisiert und mittels Aufbau auf bei

FOKUS entwickelte generische Komponenten die Sprachassistenten-Funktionalität des EAA sprachübergreifend in einer initialen Version bereitgestellt. Die EAA-Nutzerschnittstellen wurden gemäß der im Projekt verwendeten agilen Entwicklungsmethodik mehrfach überarbeitet und aufgrund neu klar gewordener Praxisanforderungen (z.B. aus Diskussionen mit dem Praxispartner Triaphon, aus projektinternen Testungen und aus Ergebnissen des Anwendertest-Workshops) optimiert.

Alle bisherigen Entwicklungen der verschiedenen Subkomponenten des EAA wurden zu einem ersten für die Anwendertestung geeigneten Gesamt-Release integriert und initial auf die vom neuen assoziierten Partner HERTZ Systemtechnik bereitgestellte Kioskterminal-Hardware portiert.

Im Rahmen von **HYKIST+** erfuhr der EAA eine erste Anpassung für Ukrainisch, indem sowohl seine Nutzerschnittstellen für die ukrainische Sprache lokalisiert wurden als auch die von den HYKIST-Partnern bereitgestellten Ukrainisch-Sprachverarbeitungs-komponenten integriert wurden.

### **2023:**

Die Entwicklungsarbeiten am EAA wurde nochmals intensiviert mit dem Ziel, einen leistungsfähigen und stabilen Entwicklungsstand zum Beginn der klinischen Evaluation zu erhalten. Eine erhebliche Anzahl durch den Praxispartner Triaphon in 2023 (und z.T. erst kurz vor Beginn der Evaluation) festgestellte spezifische Anforderungen hierfür wurde agil und in enger Abstimmung mit dem gesamten Konsortium zeitnah umgesetzt. Als Beispiel sei eine Erweiterung der Nutzerschnittstelle genannt, um auch den Fall zu berücksichtigen, dass Nutzer nicht selbst Patienten sind, sondern ein Kind zur Behandlung bringen: In diesem Fall erfolgt die Online-Zustimmung zum Datenschutz ja durch den Elternteil als juristische Vertretung, und es ist eine weitere Namenseingabe erforderlich. Auch die Erstanamnese-Fragenliste und Details z.B. zur Darstellung von Schmerzgraden wurden überarbeitet, und eine neuer Fragebogen am Sitzungsende, mit dem Patient Feedback übermitteln können, integriert. Nach festgestellten Problemen mit der Qualität der Spracherkennung wurde die Präsentation der Sprachassistentenfunktionalität im EAA überarbeitet. Es wurde ein Backup-Dialog realisiert, der im Falle eines Druckerausfalls im EAA-Terminal trotzdem die Fortführung des Prozesses, also den nachfolgenden Aufbau einer HYKIST-Sprachmittlerkonferenz, sichert.

Für den produktiven Betrieb in der klinischen Evaluationsphase wurde auch der Entwicklungsprozess noch stärker professionalisiert. Im Entwicklungszyklus wurden automatische End-to-End-Tests verankert, die bei jedem Deployment automatisch ausgeführt werden und so dafür sorgen, dass keine fehlerhaften Softwareversionen in den produktiven

Betrieb gelangen können. Zur Unterstützung der Weiterentwicklung parallel zum Produktivbetrieb wurde eine Staging-Infrastruktur aufgebaut und genutzt. Auftretende Fehler wurden in einem Ticketsystem verwaltet und bis zur vollständigen Bearbeitung verfolgt. Für verschiedene im Klinikeinsatz durch suboptimale Rahmenbedingungen (z.B. unzuverlässige WLAN-Anbindung der EAA-Terminals) aufgetretene Probleme wurden kurzfristig wirksame Maßnahmen realisiert.

Im Rahmen von **HYKIST+** wurden die EAA-Anpassungen für Ukrainisch finalisiert. Im Rahmen von **HYKIST-Future** wurde für den HYKIST-Erstanamneseassistenten eine zusätzliche Internationalisierung mit der Fremdsprache Englisch realisiert. Diese erleichtert die Durchführung von Demos, wenn im Publikum keine Sprachkenntnisse der primären Projekt-Fremdsprachen vorhanden sind. Ebenfalls wurden die hierfür nötigen erforderlichen Erweiterungen und Anpassungen im Sprachassistentz-Backend vorgenommen, das nunmehr bei Fraunhofer FOKUS gehostet wird.

### **Ergebnisse 2020**

- Initiales Anwendungsszenario für den Erstanamnese-Assistenten
- Grobentwurf von Struktur und prinzipiellen Inhalten des Dialogschemas
- Grundsätzliche Architektur (EAA + EAA-Service)
- Basisanforderungen für die Schnittstellen des EAA mit anderen Komponenten des HYKIST-Systems

### **Ergebnisse 2021**

- Implementierte EAA-WebApp
- Implementierte Schnittstelle zum EAA-Service
- Implementierter erster Prototyp einer EAA-Zusatzkomponente zur grafisch-interaktiven Auswahl der von Symptomen betroffenen Körperregionen,
- Implementierter EAA-Service mit Schnittstellen zu anderen Komponenten des HYKIST-Systems (u.a. zum Sprachassistentz-Backend)
- Automatisierte Testinfrastruktur für die Bildschirm-Interaktion mit dem EAA, Spezifikation erster, diese Testinfrastruktur nutzender automatisierter Tests
- Deployment zur Umsetzung des Testbetriebs des EAA-Services und des EAA, inklusive IT-Infrastruktur-Bereitstellung durch FOKUS sowie Entwurf und Umsetzung eines vorläufigen Sicherheitskonzepts für den EAA

### **Ergebnisse 2022**

- Initiale Implementierung des EAA-Sprachassistenten

## HYKIST-Schlussbericht

- Integration aller entwickelten EAA-Subkomponenten zu anwendertestfähigem Release
- Initiale Portierung des EAA auf die Kioskterminal-Hardware von HERTZ Systemtechnik
- Agile Weiterentwicklung der Nutzerschnittstellen des EAA entsprechend iterierter Anforderungen aus der Praxis und Erfahrungen aus internen Tests und dem Anwendertest-Workshop
- **HYKIST+:**
- Initiale Anpassung des EAA für Ukrainisch (Ukrainisch-Lokalisierung der EAA-Nutzerschnittstellen, EAA-Integration der Ukrainisch-Sprachverarbeitungs-komponenten)

**Ergebnisse 2023**

- Erweiterung der FOKUS-Infrastruktur um eine Staging-Infrastruktur zur Absicherung eines stabilen Produktivbetriebs des EAA bei paralleler Weiterentwicklung
- Finale Portierung des EAA auf die Kioskterminal-Hardware von HERTZ Systemtechnik, inkl. Druckeransteuerung
- Vielfache agile Weiterentwicklungen an EAA-Front- und Backendssoftware gemäß Anforderungen der klinischen Evaluierung
- Integration geeigneterer Onscreen-Tastaturkomponenten für Fremdsprachen
- Umsetzung eines zusätzlichen (im Normalbetrieb versteckten) Debug-Modus im EAA zur Verfolgung von im Praxiseinsatz aufgetretenen Problemen
- Integration von durch Muttersprachler geprüften finalen Versionen der internationalisierten EAA-Textinhalte
- Überarbeitung der Logging-Funktionalitäten zwecks Datenerhebung in der Evaluierungsphase
- Realisierung von end-to-end Tests und *Continuous Integration*-basierte Integration in den Entwicklungsprozess
- Optimierungen an EAA-Komponenten zu Beginn der klinischen Evaluation, u.a. Realisierung zusätzlicher Resilienz-funktionalität zur Mitigation instabiler WLAN-Anbindung des EAA in der Klinik
- Vorbereitung und Installation eines EAA-Terminals für die HYKIST-Präsentation auf der BMG-Konferenz „Data4health“

**HYKIST+:**

- Finale Anpassung des EAA für Ukrainisch (Ukrainisch-Lokalisierung der EAA-Nutzerschnittstellen)

- **HYKIST-FUTURE:**
- EAA mit neuer, zusätzlicher Fremdsprache Englisch inkl. eigenes hosting im Fraunhofer FOKUS-Rechenzentrum

#### 4.3.3 Aktivitäten und Ergebnisse Triaphon gGmbH

##### **2020:**

Triaphon hat mit seiner medizinischen Expertise die grobe, inhaltliche Ausrichtung des Erstanamnese-Assistenten erarbeitet. Welche Fragen sind essentiell für eine schnelle Erstanamnese in der Notaufnahme?

##### **2021:**

Triaphon hat die Inhalte des EAA (Fragen und Antwortoptionen) erarbeitet und zusammen mit FOKUS die graphische Umsetzung erarbeitet. Zusammen mit dem Gesamtkonsortium wurde der Use-Cases des EAA (integriert in den patient-journey) erarbeitet und den sich hieraus ergebenden Anforderungen für den Praxiseinsatz bzw. das Einsatzszenario (bsp. Terminal am Anmeldebereich der ZNAs, Ausgabe von Dokumenten zur Unterschrift (Einwilligungserklärung), Mitgabe des PIN an Patienten mit Hinweisen für den Arzt (Zusammenfassung der Kontextdaten, erfolgte Einwilligung und Rufnummer plus „PIN“ für Call)

##### **2022**

Triaphon hat im Projektverlauf stets auf eine Anpassung des Gesamtsystems sowie ihrer Teilsysteme im Hinblick auf die Anforderungen des klinischen Settings (wenig Zeit, große Lautstärke, einfachste Handhabung) hingewirkt mit dem Ziel einen möglichst niederschweligen Zugang für das gesamte medizinische Personal zu erreichen (Berücksichtigung der Bedürfe von Pflegekräften und ärztlichen Personal). Zudem wurden die Inhalte der Fragen des EAA in weiteren Feedback-Schleifen mit Experten stetig überarbeitet, einzelne Dimensionen ersetzt oder ergänzt sowie Formulierungen angepasst. Die Inhalte des EAA wurden durch Triaphon-Sprachmittler in die Zielsprachen Vietnamesisch und Arabisch übersetzt.

##### **2023**

Triaphon hat den patient-journey insbesondere in Detailfragen und bis zur finalen Umsetzung in der klinischen Studie stets mit den Partnern ausgearbeitet und hier auf eine niederschwellige und praxisnahe Umsetzung hingewirkt, inklusive Einbindung in das konkrete Szenario vor Ort in den Notaufnahmen.

Triaphon hat die inhaltliche Ausgestaltung des Dialogschemas (Anzahl und Dimensionen der Anamnese-Fragen) bis zuletzt in weiteren Feedback-Schleifen auch mit der Testklinik abgestimmt. Es galt hier die Balance zwischen einer gewünschten umfassenden Informationslage und der technischen Machbarkeit (insbesondere im Hinblick auf Usability und Nutzerakzeptanz) zu finden.

Im Zuge der klinischen Studie hat Triaphon Lotsen zur Unterstützung für die Patienten in der Bedienung des EAA in der Testklinik installiert (Rekrutierung, Schulung, Betreuung, strukturiertes Sammeln von Feedback). Die Lotsen haben das Gesamtsystem und insbesondere den EAA regelmäßig getestet, das Feedback an Triaphon und bei relevanten technischen Störungen direkt an FOKUS gemeldet (via von Triaphon moderierten Messenger-Chatgruppe und Zoom-Calls). Triaphon hat nach Beendigung der klinischen Studie eine Übersicht über die aufgetretenen Störungen am EAA und die erfolgten Updates durch FOKUS verfasst und dem Gesamtprojekt zur Verfügung gestellt. Triaphon hat die Universität Köln bei der Durchführung und Planung von simulierten Anrufen unterstützt, inklusive Vor-Ort-Einführungstermin in die Bedienbarkeit des EAA inklusive Test-Durchläufe.

### **Ergebnisse 2020**

- Mitarbeit an den ersten Entwürfen der zu erfragenden Informationen im EAA.

### **Ergebnisse 2021**

- Ausgestaltung der erhobenen Dimensionen und Formulierung der einzelnen Fragen inklusive der Antwortoptionen
- Zusammen mit FOKUS Entwicklung der graphischen Darstellung (Bspw. Abfrage von Schmerzlokalisierung und Intensität)
- Probelauf bezüglich Usability, Verständlichkeit und inhaltliche Korrektheit (Feedbackschleifen)
- Übersetzung der EAA-Inhalte auf Vietnamesisch
- Definieren und Erarbeiten der Schnittstelle EAA zu bestehendem Triaphon-System (PIN-Eingabe etc.) in Zusammenarbeit mit der Priorist GmbH

### **Ergebnisse 2022**

- Anpassung und Überarbeitung der Inhalte des EAA (Dimensionen, Antwort-Optionen und Fragen-Baum)

- Zusammen mit FOKUS Entwicklung von Lösung von Usability-Schwierigkeiten in der Bedienbarkeit des EAA, inkl. Feedback der Sprachmittler sowie Erörterung sprach- und kulturspezifischer Besonderheiten
- Übersetzung der EAA-Inhalte in die Zielsprachen Arabisch und Vietnamesisch

### **Ergebnisse 2023**

- Übersetzungen von EAA-Inhalten in den Sprachen Arabisch und Vietnamesisch durch Triaphon-Sprachmittler.
- Abstimmung der finalen Fassung der EAA-Inhalte mit der ärztlichen und pflegerischen Leitung der Testklinik

## **4.4 AP 5: Integration zum Proof of Concept der Gesamtlösung**

In diesem Arbeitspaket erfolgt die Integration der einzelnen Komponenten zu einem Gesamtsystem. Zum AP gehören der Entwurf und die Umsetzung der Architektur des HYKIST-Gesamtsystems, die Definition aller Schnittstellen zwischen den Subsystemen und Komponenten sowie deren Kollaborationen, die Realisierung der erforderlichen Middleware, aber auch die Entwicklung der UIs für den Erstanamnese-Sprachassistenten und die Sprachmittler-WebApp. Ebenfalls werden die auf kontinuierlicher Basis erfolgenden Komponentenintegrations- und -testaktivitäten im Projekt durch dieses Arbeitspaket abgedeckt.

### **4.4.1 Aktivitäten und Ergebnisse Fraunhofer FOKUS**

#### **2020:**

Projektkoordinator FOKUS hat, parallel und mit engem Bezug zum in intensiver Kooperation entstehenden Datenschutzkonzept, substanzielle Abstimmungen zur Gesamtarchitektur des HYKIST-Systems vorangetrieben.

#### **2021:**

FOKUS hat als Projektkoordinator die Gesamtarchitektur des HYKIST-Systems definiert und dokumentiert. Hierfür wurde im Rahmen der AG Architektur und Implementierung ein Dokument zur Definition der Schnittstellen erarbeitet. Für die Integration der bestehenden Triaphon-Plattform wurde eine Spezifikation erarbeitet. Um die Implementierungsarbeiten im Projekt aufeinander abzustimmen wurde ein Planungsdokument angelegt, das als Entwicklungsroadmap verschiedene Ausbaustufen definiert und mit internen Meilensteinen hinterlegt. Darüber hinaus hat FOKUS ein Interfacedesign für die Bedienoberflächen konzipiert und die einzelnen Interaktionsmöglichkeiten in verschiedenen Mockups zunächst definiert und nachfolgend gemäß den vereinbarten Entwicklungsstufen implementiert. Im Rahmen der

regelmäßig stattfindenden AG Architektur und Implementierung hat Fraunhofer FOKUS den Arbeitsfortschritt der einzelnen Ausbaustufen sowie die Einhaltung der Termine und Absprachen koordiniert. Für die Integration der verteilt ausgeführten HYKIST-Dienste wurde die Integrationsschicht (Middleware) im HYKIST-Backend entwickelt, das die einzelnen Datenströme, Abläufe und Funktionsaufrufe im HYKIST-Gesamtsystem koordiniert. Als zentrale Nutzerschnittstelle für die Ausgabe der maschinellen Übersetzungsergebnisse wurde die HYKIST-App konzipiert und implementiert.

## **2022**

FOKUS hat als Projektkoordinator innerhalb der regelmäßigen Treffen der AG Architektur und Implementierung die Entwicklung des Gesamtsystems inklusive der verteilten Programmabläufe, Datenflüsse und Bedieninterfaces koordiniert und bis zu den avisierten Teststellungen vorangetrieben. Für die Echtzeit basierte Verarbeitung der Gesprächsdaten über das HYKIST-Backend wurde die Koordinationsschicht (Middleware) der Dienste inklusive des EAA-Backendservices implementiert und durch Deployment in der FOKUS-Infrastruktur in den laufenden Testbetrieb überführt. Auf dieser Basis wurde die Kollaboration der Middleware mit den anderen HYKIST-Komponenten (Triaphon-Plattform, Sprachassistenten-Backend, HYKIST-App und Erstanamneseassistent) im Rahmen des angestrebten verteilten HYKIST-Gesamtsystems erfolgreich getestet und eingefahren. Über den Unterauftrag an die Firma priorist GmbH hat FOKUS die Triaphon-Plattform technisch über die zu schaffenden Schnittstellen in das HYKIST-Gesamtsystem eingebunden und war gemeinsam mit den Partnern Triaphon und priorist GmbH an deren Konzeption und Abstimmung beteiligt. In Middleware und Infrastruktur wurden erste Maßnahmen zum Monitoring und Logging realisiert, um die Analyse des Systemverhalten während Tests (z.B. des Anwendertest-Workshops) zu unterstützen. Auf die Rückmeldungen der Nutzer aus den Testungen (siehe AP 6) wurde die HYKIST-App zur Ausgabe der Übersetzungsergebnisse jeweils an die neuen Anforderungen angepasst.

Parallel zu den Integrationsarbeiten für die Sprachen Arabisch und Vietnamesisch erfolgten **für HYKIST+** die Erweiterungen des HYKIST-Systems für das Ukrainische. Hierfür wurden die entsprechenden Schnittstellen, Datenflüsse und Bedieninterfaces im HYKIST-Backend für die Audioverarbeitung zwischen dem Triaphon-Backend (Twilio-Datenstrom) und dem AppTek-Backend analog zu den bereits existierenden Strukturen für die neue Zielsprache angepasst und die Bedieninterfaces der HYKIST-App sowie EAA (siehe AP 4) um die ukrainische Sprache erweitert.

## **2023**

Für die Durchführung der klinischen Studie wurde das Gesamtsystem an die infrastrukturellen Bedingungen im Sanaklinikum Berlin und der Universität Köln angepasst und für eine

möglichst fehlerfreien Pilotbetrieb mehreren Ende-zu-Ende-Tests unterzogen. Nach den Testungen wurden 2 Gesamtsysteme an das Sanaklinikum Berlin und die Universität Köln ausgeliefert. Neben der Generierung dynamischer PINs wurde das HYKIST-Backend um die Verarbeitung statischer Patienten-PINs erweitert, die zur Ermittlung der angeforderten Sprache an das Triaphon-Backend zurück übermittelt werden müssen. Zur erweiterten Auswertung des Nutzungsverhaltens und zur Fehlersuche wurde außerdem im HYKIST-Backend ein neuer Endpunkt zur Aufzeichnung der Logging-Daten aus dem EAA hinzugefügt. Für die schnelle Fehlerbehebung und Rückmeldungen von den Testnutzern und Lotsinnen hat sich FOKUS an dem Messengerdienst von Triaphon beteiligt und technischen Support geleistet. Gemeldete Fehlermeldungen aus den Pilotbetrieben wurden schnellstmöglich behoben.

In der HYKIST-App wurden auf Wunsch der Partner letzte Änderungen bei der Visualisierung des Live-streams implementiert. So wird jetzt im Stream jeweils mit dem Symbol grün/gelb angezeigt, ob der Satz final erkannt/übersetzt wurde oder nicht. Hierfür wurde das Protokoll zwischen HYKIST-Backend und App angepasst, sodass nachträgliche Änderungen (ältere Einträge im Stream) im Frontend angezeigt werden können, wenn Erkenner/Übersetzer bessere/andere Ergebnisse liefern.

Die EAA-Kioskterminals wurde auf Hardware- und Systemebene für den öffentlichen Einsatz in der klinischen Evaluierungsphase vorbereitet, und es wurden eine Installations- und Betriebsdokumentation des EAA-Terminals sowie eine Bedienungsanleitung für die HYKIST-App erstellt.

Für **HYKIST-Future** wurde die HYKIST-App angepasst und für die geänderten Systemkomponenten das Triaphon-Login deaktiviert. Ebenso wurden einzelne Darstellungen in der App wie z.B. verschiedene Sprecheransichten deaktiviert, da für die Prototyperstellung zuerst nur von der Darstellung des Übersetzungsstreams ausgegangen wird.

Ebenso wurde basierend auf den HYKIST-Ergebnissen und dem bei FOKUS entwickelten Sprachassistenten-Framework FISA (FOKUS Intelligent Speech Assistant) neue Komponenten implementiert, die eine Anbindung der nachgeordneten Services bzw. Open-Source-Komponenten für die einzelnen Sprachassistenten-Basisfunktionalitäten basierend auf übersichtlichen, in Python definierten, und auf dem offenen W3C-Protokollstandard Websockets beruhenden Interfaces herstellen. Konkret für die Spracherkennung/ASR sowie für die Sprachidentifikation/LI wurde beispielhaft die Open-Source-Lösung fasterwhisper eingebunden und getestet. Darüber hinaus wurde auch für die Sprachassistenten-Basisfunktionalität Sprachsynthese/TTS eine Anbindung der nachgeordneten Services bzw. Open-Source-Komponenten basierend auf einem in Python definierten, und auf dem offenen W3C-Protokollstandard Websockets beruhenden Interface realisiert; konkret wurde

beispielhaft die Open-Source-Lösung CoquiTTS angebunden. Für die Maschinelle Übersetzung wurde die Open-Source-Lösung Facebook M2M100 eingebunden und getestet.

Für den Aufbau einer generischen Schnittstelle zur Anbindung existierender VOIP-Open Source-Dienste wurden zuerst unterschiedliche Anbieter (u.a. Asterisk PBX, LinPhone (Flexisip), OpenSIPS) recherchiert und evaluiert, inwieweit sie die Anforderungen eines ausleitbaren Audiostreams erfüllen. Hinsichtlich des zu verwendenden Protokolls wurden http, websockets und SSL präferiert. Im Kontext von HYKIST-Future wurde die VoIP-Technologie des Anbieters Asterisk aufgesetzt und in die HYKIST-Plattform integriert, da Asterisk ein eigenes Modul anbietet, über welches der Audiostream gegabelt und somit parallel zur VoIP-Konferenz verarbeitet werden kann. Um den Proof of concept nachzuweisen sind bislang Anrufe nur aus dem internen Telefonnetz von FOKUS möglich. Da unter den analysierten Open source VOIP-Anbietern bislang nur Asterisk die Funktionalität eines ausleitbaren Audiostreams besitzt wurde es im Projektverlauf nicht weiter notwendig ein Adapterkonzept zu implementieren.

Für die Weiterleitung der Sprachdaten vom VOIP-Server an das Sprachassistentz-Backend über die HYKIST-Middleware mussten die bestehenden HYKIST-Protokolle angepasst werden (vorher gRPC, jetzt websockets). Ebenso musste die Kodierung des Streams angepasst werden, aus der VOIP-Verbindung die Audio-Rohdaten für die aktuell eingesetzten Dienste der Sprachassistentz in einen Base64-String konvertiert werden müssen. Diese Aufgabe wird durch die HYKIST-Middleware umgesetzt. Weitere Adapter sind in dem aktuellen Aufbau nicht notwendig. Ebenso wurden Schemaanpassungen der zurückgelieferten JSON-Response aus der Sprachassistentz für die Visualisierung in der App notwendig. Auch Anpassungen an den Authentifizierungsmechanismen wurden notwendig. Hierfür wurde der Keycloak-Dienst in das Gesamtsystem eingebunden, um die Anrufe den Nutzern der App zuordnen können.

Unter Nutzung der oben beschriebenen Komponenten und offenen Interfaces für nachgeordnete Sprachservices wurde ein Integrationsproxy realisiert, der die Ersetzbarkeit der seinerzeit im HYKIST-Projekt genutzten proprietären Sprachservices mittels (über die genannten offenen Schnittstellen angebundenen) Open-Source-Sprachkomponenten praktisch demonstriert und belegt.

Begleitend zu den Anpassungs- und Erweiterungsarbeiten fanden für HYKIST-Future die notwendigen Funktions- und Integrationstest statt, um einen fehlerfreien Ablauf und Aufruf der verteilten Komponenten des Gesamtsystems zu garantieren. Der Nachweis über ein funktionierendes Gesamtsystem wurde in einer abschließenden Teststellung am FOKUS erbracht, bei dem die gesamte Datenstrecke vom Aufbau eines VOIP-Calls, über die HYKIST-Middleware zur Sprachassistentz und zurück bis in die HYKIST-App getestet wurde.

**Ergebnisse 2020** (dokumentiert im Datenschutzkonzept, siehe 5.1):

- Erste abgestimmte Darstellung der HYKIST-Gesamtarchitektur (Basis für den nachfolgenden detaillierten Architekturentwurf)
- Identifikation der erforderlichen Datenflüsse (Grundlage für die in der weiteren Architektureraarbeitung vorgesehene Schnittstellendefinition)

**Ergebnisse 2021:**

- Finale HYKIST-Gesamtarchitektur
- Dokument zu HYKIST-Architektur und Schnittstellen
- Planungsdokument HYKIST-Prototypen
- Protokolle der AG Architektur und Implementierung
- Designvorschläge für die HYKIST-Oberflächen
- Erste Prototypen des Erstanamneseassistenten (EAA) und der HYKIST-App
- Aufbau der Integrationsschicht (Middleware) im HYKIST-Backend
- Erfolgreiche erste Integrationstests in diversen Kombinationen der in Entwicklung befindlichen HYKIST-Komponenten, Beleg der Funktionsfähigkeit aller Schnittstellen der HYKIST-Komponenten sowie der grundsätzlichen Funktionsfähigkeit der Ende-zu-Ende-Integration des Gesamtsystems.

**Ergebnisse 2022:**

- Planungsdokumente der HYKIST-Prototypen
- Protokolle der AG Architektur und Implementierung
- Implementierte und integrierte HYKIST-Anwendungen inkl. der Nutzeroberflächen der HYKIST-App und des Erstanamneseassistenten (EAA, siehe Kap. 4) im HYKIST-Gesamtsystem
- Implementierte, lauffähige Integrationsschicht (Middleware) im HYKIST-Backend inkl. Schnittstellen zu Triaphon-Plattform und zum Sprachassistenten-Backend von RWTH und AppTek
- Konzipierte und abgestimmte Schnittstellen zu der Triaphon-Plattform über den vergebenen Unterauftrag an die priorist GmbH
- Deployment der Backend-Komponenten in der FOKUS-Infrastruktur für laufenden Testbetrieb

- Implementierte Schnittstellen und Bedieninterfaces für die ukrainische Sprache für das HYKIST-Gesamtsystem

### **Ergebnisse 2023**

- Erweitertes HYKIST-Protokoll und entsprechende Visualisierungen in der HYKIST-App (erste/finale Übersetzungen)
- Erweitertes HYKIST-Backend um die Verarbeitung statischer PINs
- Ende-zu-Ende-Test für die Systeminstallationen im Sanaklinikum Berlin und der Universität Köln
- Hardware-Anpassungen der EAA-Terminals (Mikrofon-Integration, Gestaltung im HYKIST-Design)
- Erarbeitung von Installations- und Betriebsdokumentation für HYKIST-App und EAA
- 2 ausgelieferte Gesamtsysteme (Sana Klinikum Lichtenberg, Universität Köln)
- **HYKIST-FUTURE:**
- Angepasstes Login der HYKIST-App
- Angepasste Anzeige der übersetzten Gespräche in der App
- Definition der o.g. offenen Schnittstellen für Spracherkennung (ASR), Sprachsynthese (TTS) und Maschinelle Übersetzung (MT)
- Beiträge zum HYKIST-Future-Architekturdokument (ASR, TTS, MT, VOIP)
- Realisierung und Test der o.g. ASR-, TTS- und MT-Komponenten durch beispielhafte Anbindung geeigneter Open-Source-Lösungen
- Recherche unterschiedlicher Open Source VOIP-Anbieter
- Beiträge zum HYKIST-Future-Architekturdokument
- Integrierte Asterisk VOIP-Technologie
- Implementierte Schnittstellen zum HYKIST-Backend
- Angepasstes HYKIST-Streaming-Protokoll
- Angepasste Schnittstellen der HYKIST-Middleware zum VOIP-Dienst, zur Sprachassistenz und zur HYKIST-App
- Realisierung und Test des Integrationsproxys
- Funktions- und Integrationstests
- Lauffähiges HYKIST-Future Gesamtsystem

#### 4.4.2 Aktivitäten und Ergebnisse Triaphon

**2020:**

keine

**2021:**

Zusammen mit dem Gesamtkonsortium hat Triaphon an der Gesamtarchitektur gearbeitet. Insbesondere bei der Entwicklung der HYKIST-App (für die Sprachmittler) hat Triaphon Zuarbeiten geleistet.

**2022:**

Zusammen mit dem Gesamtkonsortium hat Triaphon an der Gesamtarchitektur gearbeitet und insbesondere bei der Konzeption der HYKIST-APP sowie des EAA-Anforderungen und Lösungsideen eingebracht. Die technische Implementierung des HYKIST-Systems in die laufende Triaphon-App wurden durch den IT-Dienstleister priorist GmbH (im Unterauftrag von Fraunhofer FOKUS) erfolgreich umgesetzt und in den Triaphon-Code integriert

In Feedback-Schleifen wurde Verbesserungsvorschläge durch Sprachmittler und des Triaphon-Teams bezüglich Inhalte und der Usability in die weitere Entwicklung aufgenommen. Im Zuge von HYKIST+ hat Triaphon das Triaphon-HYKIST-Backend um die Sprache Ukrainisch ergänzt.

**2023**

Triaphon hat bis zum Praxisstart in der Konzeption, der allgemeinen Gesamtarchitektur und in der speziellen Umsetzung in der Testklinik auf eine möglichst niederschwellige Bedienbarkeit und Praxistauglichkeit des Gesamtsystems und seiner Teilsysteme (EAA und Hykist-APP) hingewirkt. Triaphon hat die konkrete Implementierung des Gesamtsystems an der Testklinik vorbereitet und die Kommunikation zwischen und mit den relevanten Akteuren übernommen (Klinikleitung, ärztliche und pflegerische Leitung der ZNA, Lotsen, Sprachmittler, Projektpartner, externe Evaluationspartner). In Zusammenarbeit mit FOKUS wurden notwendige Hilfsmittel ausgearbeitet (bsp. Infozettel für Patienten und Personal, Zettel für das Vorgehen bei der Verwendung von Notfall-PINs etc.).

Bei der Entwicklung der HYKIST-App (für die Sprachmittler) hat Triaphon in 2023 weitere Zuarbeiten geleistet. Zusammen mit den Projektpartnern hat Triaphon mit Sprachmittler an Live-Praxis-Tests teilgenommen und diese mit vorbereitet. Nach den ersten Live-Tests und den ersten erfolgten Einsätzen in der klinischen Studie durch die HYKIST-Sprachmittler wurde konkretes Feedback in Chatgruppen und einzelnen Zoom-Meetings eingeholt. Auf dieser Grundlage ausgearbeitete Verbesserungsvorschläge bezüglich des Gesamtsystems und insbesondere zur HYKIST-App wurden umgehend an FOKUS weitergereicht.

Im Zuge von HYKIST+ hat Triaphon das Triaphon-HYKIST-Backend um die Sprache Ukrainisch erweitert.

### **Ergebnisse 2020:**

keine

### **Ergebnisse 2021:**

- Die relevanten Inhalte und Erfordernisse für die HYKIST-App wurden zusammen mit den zukünftigen Anwenderinnen (Sprachmittler von Triaphon) erarbeitet.
- Die Funktionsweise und Usability wurde zusammen mit FOKUS entwickelt und auf den spezifischen Use-Case abgestimmt: Telefondolmetschen (Schnelligkeit, kognitiv fordernde Tätigkeit) per Smartphone (kleines Display).
- Die zunächst erarbeiteten Inhalte wurden in Feedback-Schleifen durch die Sprachmittler stetig geprüft und um Verbesserungsvorschläge oder add-ons erweitert. Beispielhaft sei zu nennen: Die Darstellung der Stream-Funktion (Ausgangssprache vs. Zielsprache); Menüoptionen und Usability-Anpassungen (Swipe-Funktion) sowie Details zur Glossarfunktion mit Übersetzungsvorschläge für einzeln Fachbegriffe.
- Inhalt und Terminologie wurden teils aus der bestehenden Triaphon-App übernommen und an die HYKIST-App angepasst (bspw. „Statistik“).
- Schnittstellen zwischen der HYKIST-App und der Triaphon-App wurden definiert und durch die Priorist GmbH in Abstimmung mit den Projektpartnern programmiert (bspw. PIN-Eingabe, Auth-Verfahren (redirect), Glossarfunktion, Statistik)
- In Abstimmung im Konsortium und des Evaluationspartners wurden die Evaluationsfragen zur Bewertung der APP nach jedem Call durch die Sprachmittler erarbeitet.

### **Ergebnisse 2022:**

- Fertigstellung der HYKIST-Triaphon-Einheit mit simpler Anwendung: Aufbau einer HYKIST-Telefonkonferenz über Anwahl einer Servicenummer und Eingabe der Patient-PIN über die Telefontastatur (von EAA generiert)
- Programmieren der Schnittstellen und Endpunkte der Triaphon-App durch die priorist GmbH (im Unterauftrag von Fraunhofer FOKUS) inklusive Einlogg- /Registrierprozess (Redirect, Übermittlung relevanter Daten an middleware wie Sprachmittler-ID, Aufbau Telefonkonferenz, Versenden des Stream Endpunkts etc.)
- Vorbereitung und Durchführung des ersten Live-Praxis-Tests in den Räumen von Triaphon unter Teilnahme von Sprachmittler mit Feedback zu notwendigen

Anpassungen bezüglich Gesamtsystem, Inhalt, Usability und Bewertung (EAA und HYKIST-APP)

- Input bezüglich Optimierung der HYKIST-App bei Anrufannahme, Login, Ansichten und Content (Triaphon-Glossare als Datengrundlage für die Termini-Erkennung)
- Abstimmung der Inhalte und Bezeichnungen zwischen Triaphon-App und HYKIST-App („Meine Zeiten“, „Statistik“, das medizinische Glossar der Zielsprachen etc.)

### **Ergebnisse 2023**

- Planung der konkreten Implementierung des Gesamtsystems an der Testklinik
- Erarbeiten von Informationsmaterial (Aushänge, Infozettel etc.) für Patienten und medizinisches Personal der Testklinik in Zusammenarbeit mit FOKUS.
- Fortlaufendes Einholen von Feedback zur HYKIST-App von den Sprachmittlern durch engen Kontakt via Messenger-Chatgruppe sowie regelmäßige Videocalls
- Sammlung und Strukturierung von Usability-Aspekten, technischen Störungen und Optimierungsvorschlägen zur HYKIST-App aus dem Feedback der Sprachmittler
- HYKIST+: Das Triaphon-HYKIST-Backend wurde um die Sprache Ukrainisch ergänzt

#### 4.4.3 Aktivitäten und Ergebnisse AppTek

##### **2020:**

keine

##### **2021:**

AppTek hat einen wesentlichen Teil der Gesamtarchitektur implementiert. Dazu wurde eine gRPC-Schnittstelle entworfen, mit Projektpartnern intensiv besprochen und implementiert. Diese Schnittstelle akzeptiert Anfragen vom HYKIST-Backend in Form eines Textes oder auch eines Audio-Streams und liefert fast zeitsynchron die Transkription der gesprochenen Äußerungen sowie deren Übersetzung in die richtige Sprache. Die Schnittstelle, die zunächst nur die einfache Transkription und Übersetzung unterstützte, wurde schrittweise erweitert um die Unterstützung der Erkennung, welche Sprache gesprochen wurde, sowie um Glossar-Unterstützung (Hochladen, Speichern, Löschen und Verwenden von Sprachmittler-spezifischen Glossare) und Hervorheben der wichtigen Begriffe in der Quell- und Zielsprache (siehe Abschnitt 5.3.1 für Details).

Die Schnittstelle wurde mehrmals von anderen Projektpartnern getestet, eventuelle kleine Fehler wurden behoben.

**2022:**

AppTek hat seine gRPC-Schnittstelle fürs Sprachassistentz-Backend erheblich verbessert. Es wurde eine separate Komponente fürs Glossar-Management (Hochladen, Löschen, Anwenden) implementiert, assoziiert mit Sprachmittler-ID. Die Schnittstelle wurde auf den Servern von AWS in Deutschland (Frankfurt) aufgebaut und umfangreich getestet. Es wurde sichergestellt, dass zusätzliche Übersetzungs-Modell-Instanzen bei einer hohen Last innerhalb weniger Sekunden gestartet werden können. Das Sprachassistentz-Backend wurde kontinuierlich von AppTeks Ingenieuren überwacht, bei etwaigen technischen Problemen bei den Anwendertestungen wurde Support-Arbeit geleistet. Es wurde sichergestellt, dass alle Ereignisse bei der Verwendung des Backends sicher protokolliert werden können. Damit wurde eine datenschutzkonforme Analyse der geloggtten Daten ermöglicht, womit die Verbesserung der Erkennungs- und Übersetzungsqualität erzielt werden kann.

Zusätzlich wurden AppTek's Systeme zur inversen Textnormalisierung für Deutsch, Arabisch, Vietnamesisch installiert, um insbesondere bei der Verwendung der HYKIST-Spracherkennungssysteme für die KI-basierte Erstanamneseassistentz die gesprochenen Zahlen und Zahlenfolgen (Alter, Datums- und Mengenangaben) korrekt und mit Berücksichtigung des Kontexts in einen korrekt schriftlich formatierten Ausdruck mit Ziffern zu umwandeln.

**2023**

AppTek hat weiterhin das Sprachassistentz-Backend betrieben und verbessert. Es wurden mehrere Spracherkennung-Modell-Updates durchgeführt, in Verbindung mit einer Neugewichtung der Konfidenzmaße der beiden Spracherkennungssysteme (Deutsch und Fremdsprache), die bei den Updates notwendig war. Es wurde auch die inverse Textnormalisierung für Vietnamesische als ein verbessertes Modell installiert und betrieben. Ferner wurde das System für die Übersetzung von Deutsch nach Vietnamesisch neu eingespielt und weitere Übersetzungsmodell-Updates durchgeführt.

AppTek hat auch dafür gesorgt, dass in den internen Testungen sowie bei den Anwender-Testungen alle Systeme einwandfrei und ohne Unterbrechung liefen. Das gesprochene und die Transkriptionen und Übersetzungen wurden vollständig geloggt. Dadurch war es möglich, die simulierten Dialoge direkt über das laufende System/App zu transkribieren, damit sie später fürs Training der Spracherkennungskomponenten und anderen Modellen genutzt werden können.

**Ergebnisse 2020:**

keine

**Ergebnisse 2021:**

- Eine lauffähige, ausgiebig getestete Schnittstelle zwischen HYKIST-Backend und Sprachassistentz-Backend, mit Unterstützung von Glossaren und weiteren HYKIST-spezifischen Funktionen

**Ergebnisse 2022:**

- Verbesserte Schnittstelle zwischen HYKIST-Backend und Sprachassistentz-Backend, mit Unterstützung von Glossaren sowie eines sprachspezifischen, verbesserten Glossar-Matchings
- Installation und Updates von Spracherkennungs-, Übersetzungsmodellen, sowie Modellen zur inversen Textnormalisierung.
- Installation des Sprachassistentz-Backends auf sicheren Servern in Deutschland
- Ständige Überwachung und Nutzer-Support des laufenden Systems
- Integration der neuen Versionen der Spracherkennungsmodelle von RWTH in die Infrastruktur von AppTek, mit anschließender Neubestimmung der Gewichte der Konfidenzmaße für die segmentweise Erkennung der gesprochenen Sprache

**Ergebnisse 2023**

- Updates von Spracherkennungs-, Übersetzungsmodellen, sowie Modellen zur inversen Textnormalisierung.
- Ständige Überwachung und Nutzer-Support der laufenden Systeme
- Integration der neuen Versionen der Spracherkennungsmodelle von RWTH in die Infrastruktur von AppTek, mit anschließender Neubestimmung der Gewichte der Konfidenzmaße für die segmentweise Erkennung der gesprochenen Sprache
- Unterstützung beim Loggen der simulierten Dialoge, deren Analyse und Extraktion der Audio- und Textdaten aus den Logs.

**4.5 AP 6: Praxistest, Evaluation und Projektabschluss**

In diesem Arbeitspaket wurde die integrierte Gesamtlösung zu zwei Zeitpunkten auf Systemperformance, Usability und Nutzerakzeptanz evaluiert. Die abschließende Evaluation der Gesamtsystemlösung erfolgte im Rahmen einer klinischen Studie anhand des in AP1 entwickelten Studiendesigns sowie standardisierter Erhebungsinstrumente zur

Technologieakzeptanz und Usability. Die Daten beider Evaluationen wurden ausgewertet und in einem Evaluationsbericht zusammengeführt. Abschließend wurden die Lessons learned sowie nächste Schritte zur Verwertung der Projektergebnisse erarbeitet.

#### 4.5.1 Aktivitäten und Ergebnisse Fraunhofer FOKUS

##### **2020:**

Keine Aktivitäten

##### **2021:**

Keine Aktivitäten

##### **2022:**

Im Rahmen der projektbegleitenden Evaluation einzelner Entwicklungsstände wurden zwei Anwendertestungen gemeinsam mit den Partnern vorbereitet und durchgeführt (30./31.05. bei und organisiert durch Triaphon 07.11.2022 bei und organisiert durch FOKUS) sowie die daraus resultierenden Optimierungsansätze des Gesamtsystems sowie einzelner Systemkomponenten dokumentiert. Im Rahmen der AG Evaluation hat FOKUS regelmäßig an den Absprachen und Vorbereitungen für die abschließende Pilotphase teilgenommen und sich an der Erstellung des Studienprotokolls beteiligt sowie Fragestellungen zur Usability-Evaluation der Bedienoberflächen eingebracht. Die Erstellung und Einreichung des Ethikvotums und des Studienprotokolls erfolgte über einen Unterauftrag von FOKUS durch die Universität zu Köln (Institut für Pflegewissenschaft).

##### **2023**

Für die Erstellung des abschließenden Ethikvotums lieferte FOKUS als Gesamtkoordinator die notwendigen Informationen zu dem HYKIST-System und nahm an einer Sitzung der Ethikkommission des Berliner Ärzteverbandes teil. Während der klinischen Studie leistete FOKUS erhebliche Unterstützungsarbeiten bei Installation, Einführung und Betrieb der EAA-Terminals in den beteiligten Kliniken (Sana-Klinikum Lichtenberg und Uniklinik Köln) für die 2 gelieferten Gesamtsysteme. Ebenso nahm FOKUS an den Arbeitstreffen der AG Evaluation teil und beteiligte sich an Diskussionen zur Erhöhung der Teilnehmerzahl. Nach Abschluss der klinischen Studie wurden die für den EAA und das HYKIST-Gesamtsystem relevanten Evaluationsfragen in einer Excell-Tabelle aufbereitet, in ihrer Qualität gesichert und ausgewertet. Die Datenauswertungen zur Usability und Akzeptanz wurden auf dem Projektabschlusstreffen vorgestellt und in Bezug zu den Datenauswertungen der Universität Köln gesetzt. Gemeinsam mit den Projektpartnern wurde an den Lessons learned und der Formulierung von über das Projekt hinausgehenden Anwendungsperspektiven z.B. in der psychiatrischen Versorgung Geflüchteter und im ambulanten Sektor gearbeitet.

**Ergebnisse 2022:**

- Erfolgreich durchgeführte Anwendertests in diversen Kombinationen der in Entwicklung befindlichen HYKIST-Komponenten zum Beleg der Funktionsfähigkeit aller Schnittstellen der HYKIST-Komponenten sowie der grundsätzlichen Funktionsfähigkeit der Ende-zu-Ende-Integration des Gesamtsystems.
- Ablauf- und Ergebnisprotokolle der Anwendertestungen und Verbesserungsvorschläge
- Konzept zur Usability-Evaluation der Bedienoberflächen von HYKIST-App und EAA
- Beiträge zum Ethikvotum und Studienprotokoll für die klinische Phase
- Konzeption und Implementierung des digitalen Einwilligungsprozesses im EAA

**Ergebnisse 2023**

- Installation / Betrieb / Support des EAA in der Evaluierungsphase inkl. Unterstützung vor Ort aktiver „HYKIST-Lotsen“
- Systemanpassungen am EAA-Terminal zur Absicherung des öffentlichen Betriebs während der klinischen Evaluation (Sperrung des Zugriffs auf Linux-Desktop etc.)
- Konzeption und Implementierung des digitalen Einwilligungsprozess der Patienten zur Studienteilnahme via EAA-Terminal
- Technischer Support und Fehlerbehebungen für die 2 Gesamtsysteme im Pilotbetrieb
- Evaluationsdaten und Datenauswertungen für die 2 Subsysteme EAA und das HYKIST-Gesamtsystem
- Formulierung von Lessons learned und weitergehenden Anwendungsperspektiven für das HYKIST-Gesamtsystem
- Evaluationsbericht der Universität Köln

**4.5.2 Aktivitäten und Ergebnisse RWTH Aachen****2020:**

Keine Aktivitäten

**2021:**

Keine Aktivitäten

**2022:**

RWTH Aachen hat zusammen mit Triaphon die AG Evaluation koordiniert und geleitet. In der AG Evaluation wurden in Zusammenarbeit mit der Pflegewissenschaft der Universität Köln das Studiendesign für den Praxistest finalisiert, das Studienprotokoll angefertigt, der Ethikantrag erarbeitet und sich auf die Evaluationsinstrumente festgelegt.

**2023**

Aktivitäten zu den Anwendertests, der Leitung der AG Evaluation sowie der Vorbereitung und Durchführung der klinischen Studie erfolgten im Rahmen einer Abordnung an den Dienort Triaphon gGmbH, Berlin und werden als Teil der Aktivitäten von Triaphon unten beschrieben.

Darüber hinaus wurden Gespräche von Patienten in arabischer und vietnamesischer Sprache aus den simulierten Anrufen segmentiert, transkribiert und verifiziert. Die derzeit im Einsatz befindlichen Streaming-Systeme zur automatischen Spracherkennung wurden auf diesen Daten evaluiert.

**Ergebnisse 2022:**

- Zusammen mit Triaphon: Leitung AG Evaluation und Koordination zwischen Gesamtkonsortium und Uni Köln (Evaluation) mit Finalisierung des Studienprotokolls, des Studiendesign und Festlegung der Evaluationsinstrumente
- Zusammen mit Triaphon: Kommunikation (Onboarding) mit Praxispartnern für Praxistest
- Beteiligung an Durchführung und Auswertung der Anwendertests

**Ergebnisse 2023**

- Insgesamt wurden 26 simulierte vietnamesische Aufnahmen mit 4611 laufenden Wörtern und manuell transkribiert und ebenso vom im HYKIST-System eingesetzten Streaming-Spracherkennungssystem erkannt. Die Wortfehlerrate beträgt 31,5%, diese liegt in einem ähnlichen Bereich wie die Ergebnisse auf den HYKIST-Testdaten. Die Out-of-Vocabulary (OOV) Rate beträgt 0.17%.
- Ein ähnliches Verfahren wurde für die arabische Sprache angewandt, die zu evaluierenden simulierten Patientengespräche umfassten hier 27 Aufnahmen mit insgesamt 1955 laufenden Wörtern. Die Wortfehlerrate betrug 39,2%, also auch hier in einem ähnlichen Bereich wie auf den HYKIST Testdaten für diese Sprache, die Out-of-Vocabulary (OOV) Rate betrug 5.2%.

#### 4.5.3 Aktivitäten und Ergebnisse Triaphon gGmbH

**2020:**

Keine Aktivitäten

**2021:**

Keine Aktivitäten

**2022:**

Triaphon hat den ersten Anwendertest in den Räumen der Triaphon gGmbH in Berlin organisiert, gemeinsam mit den Partnern vorbereitet und durchgeführt (30.05./31.05.). Die sich ergebenden Entwicklungsstände wurden dokumentiert inkl. Beschreibung von noch zu optimierenden Komponenten des Gesamtsystems sowie der Teilsysteme. Triaphon hat zusammen mit der RWTH Aachen die AG Evaluation koordiniert, die regelmäßigen Treffen vorbereitet und dokumentiert und die Zusammenarbeit des Konsortiums mit dem Evaluationsinstitut (Pflegewissenschaft der Universität Köln) abgestimmt. In der AG Evaluation wurden in Zusammenarbeit mit der Pflegewissenschaft der Universität Köln das Studiendesign für den Praxistest finalisiert, das Studienprotokoll angefertigt, der Ethikantrag erarbeitet und sich auf die Evaluationsinstrumente festgelegt.

**2023**

Triaphon hat die vorbereitenden Anwendertests zusammen mit den Partnern geplant und organisiert. Bei den Live-Tests waren stets Sprachmittler der Sprachen Arabisch und Vietnamesisch anwesend, welche das Gesamtsystem getestet, bewertet und Optimierungsideen eingebracht haben.

Triaphon hat zusammen mit der RWTH Aachen die AG Evaluation geleitet, die regelmäßigen Treffen vorbereitet und protokolliert und die Zusammenarbeit des Konsortiums mit dem externen Evaluationsinstitut (Pflegewissenschaft der Universität Köln) abgestimmt. In der AG Evaluation wurden in Zusammenarbeit mit der Pflegewissenschaft der Universität Köln das Studiendesign sowie das Studienprotokoll für die klinische Studie finalisiert, der Ethikantrag erarbeitet und die Evaluationsinstrumente festgelegt sowie die Rekrutierung koordiniert. Zusammen mit der RWTH Aachen hat Triaphon die Arbeit des Evaluationspartners (Uni Köln) aktiv unterstützt und Teilergebnisse geleistet. So hat Triaphon die Sitzung der Ethikkommission der Berliner Ärztekammer zusammen mit der Universität Köln vorbereitet, absolviert und nachbereitet, inklusive notwendiger Anpassungen im Studiendesign bzw. Studienprotokoll.

Triaphon hat die klinische Studie gemeinsam mit RWTH federführend vorbereitet und durchgeführt. Hierzu übernahm Triaphon die primäre Kommunikation mit allen beteiligten Akteuren, wie Sana Klinik Lichtenberg (Klinikleitung, ärztlicher Dienst, pflegerischer Dienst,

etc.), den Sprachmittlern, den Lotsen, den Projektpartnern sowie der Uni Köln und den Technikern von FOKUS (zu relevanten Inhalten des EAA und der HYKIST-App). Triaphon war primärer Ansprechpartner für alle beteiligten Akteure der klinischen Studie. Bei unerwartet zu geringer Rekrutierungszahl bzw. zu geringer Systemnutzung hat Triaphon zusammen mit der RWTH Aachen, dem externen Evaluationspartner und dem Gesamtkonsortium Lösungen entwickelt und diese in mehreren Eskalationsschritten umgehend umgesetzt: Etablierung von Lotsen, Aufwandsentschädigung pro Anruf für die Testklinik, Durchführen von simulierten Anrufen, dritter Standort für simulierte Gespräche beim externen Evaluationspartner). Im Anschluss erfolgte eine Auswertung der Ergebnisse inklusive einer Übersicht der lessons learned.

### **Ergebnisse 2022:**

- Zusammen mit der RWTH Aachen: Leitung AG Evaluation und Koordination zwischen Gesamtkonsortium und Uni Köln (Evaluation) mit Finalisierung des
- Planung, Durchführung und Auswertung von Testläufen
- Studienprotokolls, des Studiendesign und Festlegung der Evaluationsinstrumente
- Erfolgreiche Durchführung des ersten Anwendertests mit Sprachmittlern in Berlin (30./31.05.2022) mit erster Praxiserfahrung und Ergebnissen für die weitere Projektentwicklung
- Zusammen mit den Partnern erfolgreiche Durchführung von weiteren Anwendertests mit Sprachmittlern Erkennen von technischen Hürden oder Fehlern und Erarbeiten von Lösungen in der Bedienbarkeit des Gesamtsystems sowie der Teilsystems (EAA-Terminal und HYKIST-App) vor Studienstart
- Erarbeiten und Finalisieren des Studienprotokolls, der Evaluationsinstrumente sowie des Ethikantrags (federführend durch Pflegewissenschaft Uni Köln)
- Kommunikation (Onboarding) mit Praxispartnern für Praxistest, in enger Absprache mit der RWTH Aachen
- Koordination zwischen dem Gesamtkonsortium und der Pflegewissenschaft Uni Köln
- Mitwirkung bei der Gestaltung des digitalen Einwilligungsprozess der Patienten zur Studienteilnahme via EAA-Terminal
- Onboarding der Kliniken als Praxispartner für den Praxistest
- Akquise und erste Informationsveranstaltungen für die HYKIST-Sprachmittlern der Sprachen Vietnamesisch und Arabisch

## Ergebnisse 2023

- Koordination und Übernahme der primären Kommunikation mit allen relevanten Akteuren der klinischen Studie, insbesondere mit der Testklinik und dem externen Evaluationspartner
- Finales Onboarding der Kliniken als Praxispartner
- Anpassung der Planungen nach kurzfristiger Absage der zentralen Notaufnahme der Universität Köln zusammen mit den Partnern und enger Abstimmung mit RWTH Aachen
- Vorbereitung, Planung, Durchführung, Betreuung und Steuerung der klinischen Studie (inkl. Follow-ups mit allen Beteiligten wie Klinikpersonal, Klinikleitung, Lotsen, Sprachmittlern, Projektpartnern, Uniklinik Köln etc.)
- Initiierung, Rekrutierung, Schulung und Betreuung von Lotsen in der Testklinik, die regelmäßige Systemtests (EAA) durchgeführt haben, Fehler und Usability-Aspekte gemeldet (in Messenger-Chatgruppe und Videocalls) und notwendige Neustarts durchgeführt haben.
- Einrichten und Betreuen von eigens eingerichteten Messenger-Chatgruppen (allgemeines Feedback zum Gesamtsystem, spezifisches technisches Feedback zum EAA)
- Erstellen einer tabellarischen Übersicht über aufgetretene Störungen, die erfolgten Updates und zu Usability-Aspekten des EAA
- Vor-Ort-Einführung in die EAA-Bedienbarkeit sowie in die Durchführung von simulierten Triaphon-Gesprächen über das Hykist-Gesamtsystem an der Universität Köln (Pflegewissenschaft)
- Vermittlung der notwendigen Informationen an die Testklinik; Koordination der Implementierung vor Ort (Treffen mit Leitungsebene der Klinik, Planung der Einführungsschulungen für das Personal etc.)
- Akquise und Informationsveranstaltungen für HYKIST-Sprachmittlern der Sprachen Vietnamesisch und Arabisch
- Enge Betreuung der HYKIST-Sprachmittlern durch das Triaphon-Team (via Messenger-Chatgruppen, Einzelgespräche, Übungscalls und Videocalls zum Austausch) inklusive Einführungsschulungen mit Vermittlung eines Verständnisses für das Gesamtsystem sowie für den Aufbau der klinischen Studie und insbesondere Einführung in die Bedienung der HYKIST-App
- Sicherstellung einer lückenlosen Schichtabdeckung der HYKIST-Sprachmittler für eine gelingende (ungeplante) Anrufannahme während der Studienlaufzeit
- Kommunikation mit ärztlicher und pflegerischer Leitung sowie Klinikleitung zur Vorbereitung, Durchführung (Kick-off am 23.03.2023) und Betreuung während der gesamten Laufzeit

## HYKIST-Schlussbericht

- Durchführung von insgesamt sechs Schulungen des Personals (Ärzeschaft und Pflegekräfte) vor Ort (zu Beginn und im Verlauf) sowie Follow-Up-Calls mit pflegerischer und ärztlicher Leitung (der Notaufnahmen)
- Betreuung der Klinik bei insgesamt 50 HYKIST-Anrufen aus den zentralen Notaufnahmen (Erwachsene und Kinder)
- Ständige Systemüberwachung und Dauer-Support sowie Ansprechpartner für das HYKIST-Gesamtsystem für das Klinikpersonal vor Ort, die Sprachmittler und die Lotsen
- Kontrolle jedes Anrufs und strukturiertes Erfassen von Metadaten der Anrufe
- Erstellen einer Übersichtstabelle aller Anrufe während der Studienlaufzeit mit den relevanten Metadaten inklusive der simulierten Dialoge
- Organisation, Durchführung und Dokumentation von etwa 60 simulierten Anrufen mit muttersprachlichen simulierenden Patienten und medizinischen Personal vor Ort (n= 60)
- Enge Betreuung der Lotsen während der Einsätze vor Ort (Messenger-Chatgruppe, Follow-up-Calls), Erfassung von Feedback sowie Koordinierung der Abrechnung
- Koordination und Kommunikation zwischen Lotsen und den Projektpartnern bezüglich Feedbacks zu technischen Problemen und Usability-Aspekten des EAA
- Bei Netzwerkproblemen des EAA auf LAN-Anschluss des Terminals in Testklinik hingewirkt
- Betreuung, Anleitung und Begleitung des Aufbaus eines zusätzlichen EAA-Standorts an Uni Köln für weitere simulierte Anrufe mit Evaluation
- **Initiieren und Durchführen von Maßnahmen zur Erhöhung der Rekrutierungsrate während der Studienlaufzeit:**
- Koordination, Analyse und Nachjustieren bei zu geringer Rekrutierungsrate von Patienten in der klinischen Studie zusammen mit dem externen Evaluationspartner und der RWTH Aachen
- Absprachen und Entwicklung von Lösungen zur Usability im engen Austausch mit pflegerischer und ärztlicher Leitung vor Ort sowie den Projektpartnern und dem Projektträger
- Gespräche vor Ort mit einzelnen Betroffenen und Verantwortlichkeit eines Ansprechpartners pro Team (Ärzeschaft und Pflegekräfte Erwachsenen und Kinder-ZNA)
- Rundmails an das Personal mit Informationen über HYKIST, die Funktionsweise und wichtige Anpassungen bzw. Optimierungen (Aufwandsentschädigung, Lotsen, Updates des EAA etc.)

- Installation von Lotsen vor Ort (ab Juni 2023) mit Einführungsschulungen vor Ort und enger Betreuung über Messenger-Chatgruppen, Telefon und Follow-up-Calls (Protokolle)
- Vor-Ort-Präsenz von Team und Projektpartnern inklusive Gespräche mit dem medizinischen Personal, der Klinikleitung, den Lotsen etc.
- Etablierung einer Aufwandsentschädigung pro Anruf für die Klinik inkl. Kommunikation (ab Juni 2023)
- Vorbereitung (Anfertigen von klinischen Fällen pro Call) und Organisation von simulierten Anrufen vor Ort mit Abstimmung, Rekrutierung, Koordinierung und Einweisung des Klinikpersonal und der simulierenden Patienten
- Konzeption, Planung und Durchführung von simulierten Anrufen vor Ort unter Einbezug von Klinikpersonal, simulierenden Patienten und teilweise von Lotsen in den Notaufnahmen der Testklinik
- Supervision und Unterstützung bei Durchführung simulierter Anrufe an Uni Köln
- **Evaluation während und nach der klinischen Studie:**
- Koordination und Unterstützung der Universität Köln in der Organisation der Evaluation (Instrumente, Zeitpunkte etc.) in der Klinik (Fragebögen, Fokusgruppen etc.)
- Strukturierte Erfassung und Dokumentation von technischen Fehlern und Usability-Aspekten der HYKIST-App (Feedback der Sprachmittler) sowie des EAA (Lotsen und Patienten) mit Vermerk auf die verschiedenen Update-Versionen mit den jeweils vorgenommenen Anpassungen
- Triaphon-Projektabschluss mit (internem) Triaphon-Workshop und Formulierung der lessons learned
- Durchführung eines HYKIST-Workshops für das Triaphon-Team in Berlin als Projektabschluss zur Auswertung der HYKIST-Ergebnisse und Zusammentragen der lessons learned

#### 4.5.4 Aktivitäten und Ergebnisse AppTek GmbH

##### **2020:**

Keine Aktivitäten

##### **2021:**

Keine Aktivitäten

**2022:**

AppTek war aktiv an den Anwendertests beteiligt, AppTek's Forscher haben die Tests direkt verfolgt und etwaige technische Schwierigkeiten und Verbesserungswünsche beim Weiterentwickeln der Systemkomponenten berücksichtigt. AppTek's Ingenieure haben dafür gesorgt, dass die Anwendertestung reibungslos funktioniert. Außerdem wurden Testclients zur Verfügung gestellt, um das bei den Anwendertests aufgenommene Audio nochmals an das Sprachassistenten-Backend zu schicken, um den Einfluss der geänderten Parameter auf die Latenz und andere Eigenschaften des Systems zu testen.

**2023**

In Folge der internen Testungen hat AppTek analysiert, wie man das Zusammenspiel der Spracherkennung und Übersetzung verbessern kann. Es wurde von AppTek vorgeschlagen, die automatische Übersetzung erst dann den Sprachmittlern in der App anzuzeigen, wenn die partielle Erkennung ein stabiles Segment ausgibt, das sich nicht mehr ändert. Ferner wurden verschiedene Möglichkeiten der akustischen Segmentierung, basierend auf Pausenlänge, erprobt und deren Auswirkung auf die Qualität der Identifizierung der gesprochenen Sprache und der maschinellen Übersetzung der Segmente analysiert. Es wurde letztendlich eine Konfiguration der Parameter ermittelt und verwendet, bei der es nur zu einer möglichst minimalen Zeitverzögerung kam und dennoch die Kontextlänge für die automatische Übersetzung lang genug war, um eine korrekte Übersetzung des gesprochenen Satzes zu ermöglichen.

**Ergebnisse 2022:**

- Erfolgreich durchgeführte Anwendertests, die zur Stabilität des Sprachassistenten-Backends beigetragen haben
- Ablauf- und Ergebnisprotokolle der Anwendertestungen und Verbesserungsvorschläge in Bezug auf maschinelle Übersetzungssysteme und deren Zusammenspiel mit anderen Komponenten der Gesamtlösung
- Umfangreiches Testen mit aufgenommenen Audio-Streams aus den Anwendertestungen, um eine Verbesserung bei bestimmten aufgefallenen Problemen durch verbesserte Modelle und Parameter festzustellen

**Ergebnisse 2023**

- Verbesserte Schnittstelle zwischen Spracherkennung und Übersetzung (Optimierung der Segmentierung)
- Optimierung der Geschwindigkeit und Verkürzung der Latenz des Gesamtsystems
- Weitere erfolgreich durchgeführte Anwendertestungen

## **4.6 AP 7: Übergreifende Aktivitäten**

In diesem AP ist zunächst die Steuerung der Projektprozesse sowie die fachliche und terminliche Koordination der Arbeitspakete verortet. Weitere Gegenstände des AP sind administrative Aufgaben wie die Durchführung von Workshops, Projekttreffen, Beiratssitzungen, das Erstellen und Abstimmen von Projektberichten mit den Projektpartnern, dem Auftraggeber (BMG), dem Projektträger VDIVDE-IT und den Anwendern (Kliniken, Ärzte, Patienten) sowie die gesamte Öffentlichkeitsarbeit des Projekts.

### **Aktivitäten und Ergebnisse Fraunhofer FOKUS**

#### **2020:**

Projektkoordinator hat Fraunhofer FOKUS die entsprechenden Verträge mit den Partnern abgestimmt und zur Unterzeichnung gebracht sowie die Projektinfrastruktur aufgesetzt. Für die Steuerung der Projektarbeiten wurden die Verantwortlichkeiten verteilt und ein Berichtswesen eingeführt. Ebenso wurde ein Projektblatt erstellt und eine öffentlich zugängliche Projektwebseite angelegt.

#### **2021:**

Als Projektkoordinator hat Fraunhofer FOKUS die entsprechenden Verträge (Vertrag zur gemeinsamen Verantwortung, Unteraufträge an die Priorist GmbH und die Universität zu Köln) zur Unterzeichnung gebracht sowie die Projektinfrastruktur gepflegt. Für die Steuerung der Projektarbeiten wurden die Verantwortlichkeiten verteilt und ein Berichtswesen gepflegt. Für die Öffentlichkeitsarbeiten wurde ein Logo entwickelt. Darüber hinaus hat Fraunhofer FOKUS die regelmäßig stattfindenden Statustreffen vor- und nachbereitet sowie die Einhaltung der Absprachen koordiniert. Für den Projektträger wurde ein erster Bericht erstellt, der den Zeitraum der Monate 1-3 dokumentiert. Fraunhofer FOKUS hat weiterhin zur Vorbereitung der Testphase Kontakte zu Herstellern von Kiosk-Terminal-Hardware etabliert.

#### **2022:**

Als Projektkoordinator hat Fraunhofer FOKUS die Projekt-internen Prozesse koordiniert und die Arbeiten der einzelnen AGs und Projektpartner aufeinander abgestimmt. Darüber hinaus hat Fraunhofer FOKUS die regelmäßig stattfindenden Statustreffen vorbereitet und protokolliert. Fraunhofer FOKUS hat weiterhin zur Vorbereitung der Testphase die Fa. HERTZ Systemtechnik als Hersteller von Kiosk-Terminal-Hardware über einen Assoziierungsvertrag in das Projekt eingebunden. Für die Öffentlichkeitsarbeit wurde ein Projektblatt mit den Partnern abgestimmt und eine Internetpräsenz eingerichtet. Ebenso wurde im Rahmen des

AP 7 der Aufstockungsantrag für die Erweiterung des HYKIST-Systems für das Ukrainische vorbereitet und eingereicht.

## **2023**

Fraunhofer FOKUS hat die regelmäßig stattfindenden Statustreffen vorbereitet und protokolliert sowie über das zentrale Projektmanagement darauf hingewirkt, dass zum Beginn der klinischen Studie die technischen Arbeiten weitgehend abgeschlossen wurden. Ebenso hat Fraunhofer FOKUS das HYKIST-Gesamtsystem auf der Data4Health-Conference des BMG präsentiert. Ebenso hat Fraunhofer FOKUS die Beiträge für den Projektabschlussbericht koordiniert. Für die Umsetzung der Lessons learned aus dem HYKIST-Projekt konnte Fraunhofer FOKUS noch einen weiteren Aufstockungsantrag stellen.

## **Ergebnisse 2020**

- Kooperationsvertrag
- Auswahl, Konfiguration und Inbetriebnahme projektspezifischer Dokumentenablage- und Kollaborationstools (basierend auf den Plattformen *Microsoft Teams* und *OwnCloud*)
- Projektinterne Organisationsstruktur und Berichtswesen (fortlaufendes Protokoll)
- Konfiguration und Inbetriebnahme Projekt- und Arbeitspaket-spezifischer Mailverteiler
- HYKIST-Projektblatt (Entwurf)

## **Ergebnisse 2021:**

- Unterbeauftragung der Universität zu Köln als Evaluationspartner
- Projektbericht für die Monate 1-3
- Design des HYKIST-Logos
- Protokolle der Statustreffen
- Präsentation zum Meilensteintreffen am 30.09.2021
- Exceltabelle für das Projektmonitoring (übermittelt an PT)

## **Ergebnisse 2022:**

- Projektbericht für die Monate 4-15
- Projektblatt und Internetpräsenz des Projektes
- Protokolle der 3-wöchigen Statustreffen
- Präsentation zum BMG-Vernetzungstreffen am 27./28.01. und 15./16.11.2022

- Kooperationsvertrag mit HERTZ Systemtechnik
- Aufstockungsantrag zur Erweiterung des HYKIST-Systems um die ukrainische Sprache (Ergebnisse siehe andere Arbeitspakete)

### **Ergebnisse 2023**

- Projektbericht für die Monate 16-27
- Protokolle der 3-wöchigen Statustreffen
- Eingereichter Aufstockungstrag HYKIST-Future
- Abschlussbericht für das HYKIST-Projekt

#### 4.6.1 Aktivitäten und Ergebnisse RWTH Aachen

##### **2020:**

RWTH Aachen hat zu einer guten Kommunikation und Vernetzung mit den anderen Projektpartnern beigetragen. Dies wurde erreicht durch Teilnahme an regelmäßigen Projektmeetings und durch direkte Kommunikation und Abstimmung, insbesondere mit Triaphon und AppTek.

##### **2021:**

RWTH Aachen hat zu einer guten Kommunikation und Vernetzung mit den anderen Projektpartnern beigetragen. Dies wurde erreicht durch Teilnahme an regelmäßigen Projektmeetings und durch direkte Kommunikation und Abstimmung, insbesondere mit Triaphon und AppTek.

##### **2022:**

Teilnahme an regelmäßigen Projektmeetings und direkte Kommunikation insbesondere mit Triaphon und AppTek, für die Entwicklung und Test im Streamingmodus der automatischen Spracherkennungssysteme, sowie Aufnahmen für das ukrainische System.

##### **2023**

Teilnahme an regelmäßigen Projekttreffen sowie bilaterale Kommunikation, insbesondere mit AppTek im Hinblick auf die ASR-Systeme für die finale Evaluation und den Abschluss des ukrainischen Systems sowie mit Triaphon bezüglich der Umsetzung und Begleitung der klinischen Studie. Organisation im Zuge der Durchführung und Begleitung der klinischen Studie.

**Ergebnisse 2020**

- Koordination mit Triaphon zur Abstimmung der Audiodatenerhebung und -transkription
- Abstimmung mit Apptek bezüglich transkribierter Audiodaten für das Training akustischer Modelle
- Regelmäßige Onlinemeetings zur Projektkoordination
- Regelmäßige Onlinemeetings zur Datenerhebung
- Abstimmung im Gesamtprojekt zum Kooperationsvertrag
- Koordination der RWTH-spezifischen Aufgaben

**Ergebnisse 2021:**

- Koordination mit Triaphon zur Abstimmung der Audiodatenerhebung und -transkription
- Abstimmung mit Apptek bezüglich transkribierter Audiodaten für das Training akustischer Modelle
- Regelmäßige Onlinemeetings zur Projektkoordination
- Regelmäßige Onlinemeetings zur Datenerhebung
- Abstimmung im Gesamtprojekt zum Kooperationsvertrag
- Koordination der RWTH-spezifischen Aufgaben
- Beitrag zu den Testläufen
- Beitrag zu den initialen Planungen zur Studie 1

**Ergebnisse 2022:**

- Koordination der RWTH-spezifischen Aufgaben
- Regelmäßige Meetings für die Verarbeitung der Transkriptionsrichtlinien für das ukrainische System
- Beiträge zu den Testläufen

**Ergebnisse 2023**

- Koordination der RWTH-spezifischen Aufgaben
- Teilnahme an den regelmäßigen Status-Telkos
- Regelmäßige Treffen zu Evaluation und Praxistest
- Gemeinsam mit Triaphon Koordination der klinischen Studie
- Beiträge zu Berichten
- Übersetzung begleitender Dokumente für die Durchführung der klinischen Studie

- Organisation zum begleitenden Einsatz von Lotsen in der klinischen Studie

#### 4.6.2 Aktivitäten und Ergebnisse Triaphon gGmbH

##### **2020:**

In enger Abstimmung mit dem Vorhabenskoordinator FOKUS und allen anderen Partnern bestimmt Triaphon basierend auf seiner klinischen Praxiserfahrung in den anwendungsbezogenen Kernfragen die Richtung des Projekts. Zudem wurde die Kontaktaufnahme mit einem Evaluationspartner für die folgende Evaluationsstudie initiiert.

##### **2021:**

In enger Abstimmung mit dem Vorhabenskoordinator FOKUS und allen anderen Partnern bestimmt Triaphon basierend auf seiner klinischen Praxiserfahrung in den anwendungsbezogenen Kernfragen die Richtung des Projekts und die relevanten Inhalte der Kernmodule wie EAA und HYKIST-App. Zudem konnte u.a. durch die Aktivitäten von Triaphon ein Evaluationspartner für die folgende Evaluationsstudie gewonnen werden.

##### **2022**

Triaphon hat mit den Projektpartnern Schwierigkeiten und mögliche Lösungswege diskutiert, insbesondere im Hinblick auf die besondere Anwendersituation in einer zentralen Notaufnahme (EAA) bzw. im herausfordernden Einsatzbereich eines konsekutiven Telefondolmetsch-Settings (HYKIST-APP). Hier hat Triaphon aufgrund seiner Expertise stets auf eine möglichst praxisorientierte und niederschwellige Projektausrichtung versucht hinzuwirken.

##### **2023**

Zusammen mit den Projektpartnern hat Triaphon mit seiner Praxiserfahrung in den anwendungsbezogenen Kernfragen die grundlegende Ausrichtung des Projekts mitbestimmt und relevante Inhalte für Teilsysteme (HYKIST-APP und EAA) erarbeitet. Im Projektverlauf hat Triaphon an den Workshops, Meetings und BMG-Terminen teilgenommen und diese zusammen mit den Partnern vorbereitet. In der Verantwortung über die praktische Umsetzung der klinischen Studie erarbeitete Triaphon zusammen mit RWTH Aachen, FOKUS und der Universität Köln Lösungswege bei auftretenden Herausforderungen (technisch, organisatorisch, kommunikativ etc.) im Studienverlauf.

##### **Ergebnisse 2020**

- Onboarding Evaluationspartner für Evaluationsstudie
- Koordination Aufnahmen (AP1), insbesondere mit RWTH Aachen University

- Koordination der Teilarbeiten von Triaphon

### **Ergebnisse 2021**

- Erfolgreiches Onboarding Evaluationspartner für Evaluationsstudie
- Koordinierung der AG Evaluation, insbesondere mit RWTH Aachen University, FOKUS und Universität Köln
- Koordination der Teilarbeiten von Triaphon
- Teilnahme am BMG-Vernetzungstreffen

### **Ergebnisse 2022:**

- Mitwirkung an Projektberichten bzw. Teilnahme und Vorbereitung von BMG-Workshops (Präsentationsinhalte)
- Abstimmung mit Partnern zu einzelnen Arbeitspaketen und Inhalten
- Beitrag zu Testläufen sowie Übersetzungsleistungen durch Triaphon-Sprachmittler bzw. Triaphon-Team

### **Ergebnisse 2023**

- Abstimmung mit den Partnern zu einzelnen Arbeitspaketen und Inhalten
- Kollaboratives Erarbeiten von Dokumenten, Berichten, Workshops mit regelmäßigem Austausch in den Status-Telkos
- Vorbereitung und Teilnahme an den Workshops, Vernetzungstreffen und dem Abschlusstreffen in Aachen
- Vorbereitung und Teilnahme mit Präsentation des Projekts zusammen mit FOKUS auf der Data4Health-Conference
- Mitwirkung an Berichten inklusive dieses Schlussberichts
- Schriftliche Übersetzungsleistungen von Dokumenten durch Triaphon-Sprachmittler bzw. durch das Triaphon-Team

#### **4.6.3 Aktivitäten und Ergebnisse AppTek GmbH**

##### **2020:**

AppTek war, ebenso wie die anderen Projektbeteiligten, um eine intensive Vernetzung des Projekts gleich zum Projektstart bemüht und hat u.a. die kollaborative Bearbeitung von Projektdokumenten vorangetrieben. Die Projektpartner, wenngleich durch die Pandemie noch nie zugleich an einem Ort, sind durch regelmäßige (im Jahr 2020 wöchentliche) Videokonferenzen gut miteinander bekannt und die Kooperation funktioniert gut, auch dann, wenn zusätzlich zum Kernteam einzelne Spezialisten hinzugezogen werden. AppTek als kleines

Unternehmen ohne ausgeprägte administrative Infrastruktur war insbesondere in vertraglichen Fragen sehr flexibel und kompromissbereit und konnte dadurch zur Geschwindigkeit der entsprechenden Abläufe beitragen.

**2021:**

AppTek hat aktiv an Projektbesprechungen teilgenommen und war maßgeblich daran beteiligt, die Schnittstellen zwischen HYKIST-Backend und Sprachassistentz-Backend in gemeinsamen Gesprächen mit Projektpartnern zu definieren. Vorschläge von AppTek zur Definition der Prototypen, Abfolge der Testungen, inklusive initialer technischer Testungen der HYKIST-App (so genannte "Dry Run" Testungen) wurden sehr positiv von anderen Projektpartnern aufgenommen. AppTek war auch erheblich daran beteiligt, die Evaluierung mit Sprachmittlern zu bestimmen, und hat am Fragenkatalog mitgearbeitet, um die Fragen an die Sprachmittler so zu formulieren, dass die Antworten darauf so gut wie möglich dazu dienen, im weiteren Verlauf des Projekts die Systeme zur automatischen Spracherkennung und Übersetzung basierend auf diesem Feedback zu verbessern.

AppTek hatte sich sehr eng mit Partner RWTH abgesprochen, um gemeinsam z.B. die relevanten Daten für die Sprache Arabisch auszuwählen. Außerdem haben AppTek und RWTH zusammen daran gearbeitet, dass die ASR-Systeme von RWTH erfolgreich in die Schnittstelle des Sprachassistentz-Backends, entwickelt von AppTek, integriert werden konnten.

**2022:**

AppTek hat aktiv an Projektbesprechungen teilgenommen und war maßgeblich daran beteiligt, die Schnittstellen zwischen HYKIST-Backend und Sprachassistentz-Backend in gemeinsamen Gesprächen mit Projektpartnern zu verbessern. Vorschläge von AppTek zum Ablauf der Anwendertests wurden sehr positiv von anderen Projektpartnern aufgenommen.

AppTek hatte sich sehr eng mit Partner RWTH abgesprochen, um gemeinsam z.B. die Schnittstelle zwischen Spracherkennung und maschinellen Übersetzung zur verbessern. Außerdem haben AppTek und RWTH weiterhin zusammen daran gearbeitet, dass die ASR-Systeme von RWTH erfolgreich in die Schnittstelle des Sprachassistentz-Backends, entwickelt von AppTek, integriert werden konnten.

**2023**

AppTek hat aktiv an internen Testungen und Projektbesprechungen teilgenommen und war unmittelbar daran beteiligt, die Schnittstellen zwischen den Modellen innerhalb des Sprachassistentz-Backends und auch zum HYKIST-Backend zu verbessern, damit ein stabiler Betrieb des Gesamtsystems und eine niedrige Latenz realisiert werden kann. AppTek hat weiterhin besonders eng mit Partner RWTH gearbeitet, um gemeinsam die Schnittstelle

zwischen Spracherkennung und maschinellen Übersetzung zu verbessern. AppTek hat auch am Fragenkatalog für die Anwender-Testungen mitgearbeitet, und hat dafür gesorgt, dass Fragen zur Nützlichkeit und Qualität der beiden Komponenten gestellt wurden, damit sie auch getrennt und nicht nur gemeinsam bewertet werden können.

**Ergebnisse 2020:**

- Fachliche Absprache mit Partner RWTH zur Annotation der gesammelten Sprachdaten
- Einführung kollaborativ bearbeiteter Dokumente für verbesserte Effizienz
- Teilnahme an einer internationalen Systemevaluierung, die in eine Veröffentlichung münden wird (siehe Kap. 7)
- Koordination der AppTek-spezifischen Teilarbeiten

**Ergebnisse 2021:**

- Integration der Spracherkennungsmodelle von RWTH in die Infrastruktur von AppTek
- Mitarbeit mit Projektpartnern an einem Fragenkatalog für die Evaluierung mit Sprachmittlern
- Kollaborative Bearbeitung mehrerer Dokumente, inklusive Planung der Prototypen, Datenschutzkonzept, usw. Aktive Teilnahme an Projekttreffen, inklusive einer wichtigen Rolle im HYKIST-Architekturboard.

**Ergebnisse 2022:**

- Kollaborative Bearbeitung mehrere Dokumente, inklusive Planung der Anwendertests. Aktive Teilnahme an Projekttreffen, inklusive einer wichtigen Rolle im HYKIST-Architekturboard.

**Ergebnisse 2023**

- Integration von Modell-Updates in die Infrastruktur des Sprachassistenten-Backends
- Mehrere Vorschläge zur Verbesserung des Gesamtsystems, die zum großen Teil von den anderen Partnern gut bewertet und anschließend implementiert wurden.
- Aktive Teilnahme an internen Testungen und Projekttreffen
- Aktive Beteiligung an der Erstellung des Fragenkatalogs für die Anwender-Testungen

## 5 Erreichte Ergebnisse

Im Zeitraum **01.10.2020-31.12.2020** wurden die folgenden Zwischenergebnisse erzielt:

- Etablierung einer Projektmanagementstruktur mit internem Berichtswesen
- Aufbau der Projektinfrastruktur (Server, Anwendungsprogramme) für die projektinterne Kommunikation und Dokumentenablage
- Erste simulierte Aufnahmen von Dialogen über eine Projektinstallation des Triaphonsystems
- Kontaktaufnahme zu Evaluationspartner (Uniklinik Köln)
- Initiierung der Klärung der relevantesten Datenschutzfragen (inklusive Beginn der Erarbeitung des DS-Konzepts und der DSFA)
- Erste Überlegungen und Konkretisierungen zur System-Architektur (sowohl Erstanamnese-Assistent als auch Übersetzungsassistent)
- Basissysteme zur maschinellen Übersetzung für die Basissprachpaare Arabisch-Deutsch und Vietnamesisch-Deutsch; dazu erste Arbeiten mit öffentlich verfügbarem Textmaterial zur Adaption an die Zieldomäne

Im Berichtszeitraum **01.01.2021-31.12.2021** wurden die folgenden Zwischenergebnisse erzielt:

- Klärung der relevantesten Datenschutzfragen (u.a. Erarbeitung des DS-Konzepts, der Datenschutzfolgeabschätzung, Abschluss einer gemeinsamen Verantwortung nach DSGVO Art. 26)
- Finalisierung der System-Architektur (sowohl Erstanamnese-Assistent als auch Übersetzungsassistent)
- Erste prototypisch implementierte Systemkomponenten und Benutzeroberflächen (u.a. Erstanamneseassistent, HYKIST-App, HYKIST-Backend, etc.)
- Unterbeauftragung der Universität zu Köln und Integration in das Projekt als Evaluationspartner zu M1
- Erfolgreiche Absolvierung des Meilensteins 1

Im Berichtszeitraum **01.01.2022-31.12.2022** wurden die folgenden wichtigsten Zwischenergebnisse erzielt:

- Erster Prototyp des integrierten Gesamtsystems mit Datenfüßen über alle verteilten Systemkomponenten (Triaphon-Plattform, HYKIST-Backend, HYKIST-EAA, Sprachassistentz-Backend, HYKIST, App)

## HYKIST-Schlussbericht

- Implementierte Benutzeroberflächen für HYKIST-App und HYKIST-Erstanamneseassistent (EAA), initiale Portierung des EAA auf die Kioskterminal-Hardware
- Implementierte Applikationslogiken und Datenflüsse im HYKIST-Backend für HYKIST-App, HYKIST-EAA und Sprachassistenten-Backend
- Optimierung der Spracherkennungssysteme für die Sprachen Deutsch, Vietnamesisch und Arabisch durch die Adaption vortrainierter akustischer Modelle
- Optimierung der Suchparameter optimiert zur schnellen und möglichst fehlerfreie Spracherkennung
- Aufnahme, Segmentierung und Transkription ukrainischer Audiodaten für die Evaluation
- Entwicklung und Integration des ukrainischen Spracherkennungssystems und Integration in das HYKIST-System als separater Entwicklungsbranch
- Erfolgreicher Anwendertest-Workshop mit Vorbereitung und ausführlicher Testung des HYKIST-Gesamtsystems
- Aufbau, Qualifizierung und Betreuung des neuen Sprachenpools „Ukrainisch“ (HYKIST-Plus)
- Verbesserte Übersetzungsmodelle Deutsch ↔ Arabisch, Deutsch ↔ Vietnamesisch, mit einer verbesserten Glossarunterstützung und Markierung wichtiger Begriffe
- Neu adaptierte Übersetzungsmodelle HYKIST+ (Deutsch ↔ Ukrainisch, Russisch)
- Assoziierungsvereinbarung mit Industriepartner HERTZ Systemtechnik GmbH zur Bereitstellung robuster Kioskterminal-Hardware zwecks Betriebs des Erstanamneseassistenten (EAA) im Rahmen der klinischen Studie

Im Berichtszeitraum 01.01.2023-31.12.2023 wurden die folgenden wichtigsten Ergebnisse erzielt:

- Finalisierung des Gesamtsystems und Vorbereitung des HYKIST-Systems für die klinische Studie
- Erteilung des Ethikvotums durch die Ärztekammer Berlin
- Durchführung der Machbarkeitsstudie im Sanaklinikum Berlin-Lichtenberg und an der Universität Köln
- Evaluation des Gesamtsystems in den Monaten März bis September 2023
- Usability- und Akzeptanzevaluation des Erstanamnese-Assistenten und der HYKIST-App
- Technische Testung des ukrainisch-sprachigen Prototypen

## 6 Gender Mainstreaming Aspekte

Aspekte des Gender Mainstreaming finden seit Beginn des Projektes Berücksichtigung z.B. auch bei der Publikation des Projektblattes. Bei den Evaluierungen wurde darauf geachtet, die in der Informatik immer noch häufig anzutreffende Überrepräsentation männlicher Sprecher zu vermeiden.

## 7 Diskussion der Ergebnisse, Gesamtbeurteilung

### Gesamtbeurteilung Fraunhofer FOKUS

Fraunhofer FOKUS hat alle im Projektantrag beschriebenen Ziele erreicht. Fraunhofer FOKUS hat mit dem HYKIST-Projekt seine Kompetenzen auf den Gebieten „Echtzeitbasierte Sprachverarbeitung“ und „Entwicklung und Management sicherer, verteilter Medizininformatiksysteme“ weiter ausgebaut.

Im Rahmen der Gesamtprojektleitung in AP 7 hat FOKUS den Projektplan mit dem jeweiligen Arbeitsstand im Projekt abgeglichen und Entscheidungsprozesse im Konsortium moderiert und dokumentiert. Hierzu gehörten auch die Abstimmungen mit dem BMG und dem Projektträger, die Öffentlichkeitsarbeit und die Koordinierung des Berichtswesens. Für die technische Projektleitung in AP 5 hat FOKUS eine Entwicklungsroadmap mit den Partnern aufgestellt, die Zuarbeiten zu den jeweiligen Ausbaustufen kontrolliert und die technischen Integrationstestungen geleitet. Hierzu gehörte auch die Gesamtsystemdokumentation inklusive der Schnittstellenbeschreibungen sowie die Vertragsverhandlungen zur Einbindung der Triaphon-Software. In AP 1 moderierte FOKUS den Anforderungsprozess bis hin zur technischen Konzeption des HYKIST-Gesamtsystems und koordinierte federführend die vertragliche Ausgestaltung des Datenschutzkonzept inklusive Datenschutzfolgeabschätzung. Die technischen Entwicklungsziele in AP 4 und AP 5 wurden ebenfalls erreicht. Hierzu zählten die Entwicklung des multilingualen Erstanamneseassistenten (EAA) mit Frontend und in die HYKIST-Gesamtarchitektur integrierten Backend, die HYKIST-Middleware als Koordinierungsschicht der dezentralen Sprachassistentendienste inklusive der Schnittstellen zu den technischen Komponenten der Projektpartner. Zur Unterstützung der Sprachmittler wurde die einfach zu bedienende HYKIST-App entwickelt. Insgesamt entstand durch die Arbeiten von FOKUS im Projektverlauf ein stabiler Prototyp für die anschließende klinische Studie in AP 6. Im Rahmen der Evaluationsarbeiten untersuchte FOKUS insbesondere die entwickelten Bedienoberflächen und zentralen technischen Systemfunktionalitäten auf ihre Akzeptanz und Usability und wertete dafür die Rückmeldungen aus der klinischen Studie aus.

Einige Punkte im Projektverlauf dürfen rückblickend als herausfordernd beschrieben werden: So dauerten die Abstimmungen zum Datenschutz wesentlich länger als ursprünglich geplant,

da in dem HYKIST-System Audiodaten aus Arzt-Patienten-Gesprächen verarbeitet wurden. Hierfür wurde ein komplexes Pseudonymisierungsverfahren entwickelt, um im Falle eines Löschbegehrens die Rückverfolgbarkeit der an sich anonym verarbeiteten Audiodaten zu gewährleisten. Als Gesamtkoordinator und Verantwortlicher für das Gesamtsystem war für FOKUS insbesondere der Echtzeitbetrieb des HYKIST-Systems in den Notfallaufnahmen eine weitere Herausforderung. Mit mehreren Optimierungen am System und dem notwendigen Nutzersupport während der Pilotphase konnte aber schließlich die avisierte Zahl an Probanden erreicht werden.

Trotz der sicherlich als schwierig zu bezeichnenden Rahmenbedingungen bei der Evaluation (siehe dazu das Gesamtfazit weiter unten) kann aus Sicht von Fraunhofer FOKUS ein positives Fazit für das Projekt gezogen werden: Das Projekt wurde trotz der hohen technischen Anforderungen mit geringem zeitlichem Verzug von 3 Monaten abgeschlossen. Die personelle Kontinuität in der Zusammenarbeit bei den Partnern war hierfür ausschlaggebend. Die Performance des Gesamtsystem von Datenquelle, über das Sprachassistentzbackend bis hin zu Benutzerschnittelle der HYKIST-App darf als sehr gut beschrieben werden. Die technische Systemintegration der einzelnen Komponenten verlief aufgrund der guten Zusammenarbeit der technischen Partner weitgehend reibungslos, Probleme wurden frühzeitig durch die begleitenden Integrationstestungen identifiziert und behoben. Schließlich wurden aufgrund der Testung des HYKIST-Prototypen unter klinischen Alltagsbedingungen die notwendigen Anschluss- und Optimierungsarbeiten erkannt und in Ansätzen bereits während der Projektlaufzeit Schritte zu deren Umsetzung begonnen wurden.

### **Gesamtbeurteilung RWTH Aachen**

Die RWTH Aachen sieht die Projektziele mit der Entwicklung und Optimierung automatischer Spracherkennungssysteme und deren domänenspezifischer Auslegung für die Zielsprachen Arabisch, Vietnamesisch, Ukrainisch und Deutsch unter den Bedingungen von Telefonsprache in wechselnden akustischen Umgebungen mit angemessenen Fehlerraten weitestgehend erfüllt. Hierfür wurden Adaption-, Entwicklungs- und Evaluationsdaten für alle drei Dialog-Sprachpaare Arabisch-Deutsch, Vietnamesisch-Deutsch sowie später auch Ukrainisch-Deutsch im geplanten Umfang erhoben, transkribiert und für die Adaption an die Domäne zweisprachiger Anamnesegespräche und die Optimierung und Evaluation der entwickelten Spracherkennungssysteme eingesetzt.

Die größte Herausforderung zum Training der Spracherkennungsmodelle in HYKIST stellte die fehlende Verfügbarkeit domänenspezifischer Sprachdaten dar. Reale Anwendungsdaten konnten aus Datenschutz-rechtlichen Gründen nicht eingesetzt werden und wurden durch im Projekt generierte Daten ersetzt, die die Situation des Arzt/Ärztin-Patient/in-Sprachmittler/in

Dialogs in Zusammenarbeit mit Triaphon simulierten. Der Aufwand für die manuelle Transkription dieser Sprachdaten für Adaption, Entwicklung und Evaluation der Spracherkennungsmodelle war aufgrund der anspruchsvollen Aufnahmebedingungen dieser Domäne erhöht und die Rekrutierung von Hilfskräften für die Transkription durch die Corona-Pandemie erschwert. Dennoch konnten für jede der Zielsprachen Daten in der Größenordnung des beabsichtigten Umfangs gesammelt und annotiert werden.

Schon während der Sammlung und Annotation domänenspezifischer Daten konnten vom Partner AppTek bereitgestellte Telefon-Sprachdaten für die Entwicklung erster Spracherkennungsmodellen für die Zielsprachen des Projekts verwandt werden. Zugleich wurden aktuelle Entwicklungen zum Stand der Technik in der neuronalen Modellierung von Spracherkennungssystemen aufgegriffen, die zu entsprechenden Performanzverbesserungen führten und für die HYKIST-Spracherkennungsmodelle nutzbar gemacht wurden. Desweiteren wurden Weiterentwicklungen in der Architektur von Spracherkennungssystemen bzgl. Integration von Sprachmodellen, Glättung phonetischer Kontexte, sowie dem effizienten Training von Grund auf sowie des diskriminativen Sequenz-basierten Trainings verfolgt. Ebenso wurden aktuelle Entwicklungen bzgl. sehr großer, auf erheblichen Mengen multilingualer Sprachdaten unüberwacht vortrainierter Foundation-Modelle erfolgreich zur verbesserten und effizienten Adaption an die HYKIST-Domäne genutzt. Die Entwicklung der Spracherkennungssysteme sowie deren Adaption an die in HYKIST gesammelten domänenspezifischen Daten wurde auf der ITG-Tagung Sprachkommunikation 2023 in Aachen vorgestellt.

Die in HYKIST entwickelten Spracherkennungssysteme und im Laufe des Projekts erzielte Verbesserungen der zugrundeliegenden Modelle wurden im Rahmen der gemeinsam genutzten Plattformen RWTH-ASR für die Spracherkennungssysteme und RETURNN für die entsprechende neuronale Modellierung mit dem Projektpartner AppTek abgestimmt und zur Integration in die für HYKIST eingesetzten Sprachübersetzungssysteme übergeben. Die Spracherkennungskomponenten konnte in der klinischen Studie eingesetzt sowie deren Performanz im Rahmen der Transkription von simulierten Dialogen erfolgreich evaluiert werden.

### **Gesamtbeurteilung Triaphon**

Triaphon hat in enger Zusammenarbeit mit allen beteiligten Partnern zum großen Teil die Projektziele von HYKIST erfüllt, und sämtliche Meilensteine wurden erfolgreich erreicht. Im Rahmen des Projekts wurde eine technische Lösung entwickelt und im klinischen Umfeld evaluiert, um die Sprachmittler zu unterstützen. Um das Gesamtziel einer optimierten medizinischen Versorgung zu bewerten, wäre eine umfassendere Evaluationsstudie erforderlich.

Im Teilprojekt HYKIST+ wurde in kürzester Zeit ein neuer Sprachmittler-Pool für die ukrainische Sprache aufgebaut. Im Rahmen des Projekts "Dolmetsch-Nothilfe: Ukrainisch" wurde dieser Pool medizinischen Einrichtungen unkompliziert und kostenfrei zur Verfügung gestellt. Dadurch konnte ein Beitrag zur medizinischen Versorgung von Geflüchteten in Deutschland geleistet werden, da eine Verständigung und somit eine medizinische Behandlung von ukrainischen Geflüchteten in Arztpraxen und Kliniken ermöglicht wurde. Nach Abschluss der Projektlaufzeit wurde der aufgebaute ukrainische Sprachmittler-Pool in den regulären Betrieb von Triaphon überführt und konnte somit von den Triaphon-Kunden auch nach dem Ende der "Dolmetsch-Nothilfe: Ukrainisch" ohne Unterbrechung genutzt werden.

Die Durchführung der Pilotstudie brachte große Herausforderungen mit sich, die es zu bewältigen galt. Beispielsweise setzt eine erfolgreiche Nutzung des Systems (EAA) die Akzeptanz und Anwendung von drei verschiedenen Anwendergruppen voraus: den Patienten, dem pflegerischen und ärztlichen Dienst. Besonders herausfordernd gestaltete sich die Situation in der zentralen Notaufnahme (ZNA), wo der EAA nur im "Sonderfall" (nur bei vorliegender Sprachbarriere) Anwendung finden sollte und keinen standardisierten Ablauf darstellte.

Um diesen Herausforderungen zu begegnen, wurden verschiedene Maßnahmen ergriffen. Eine klare Darstellung des Mehrwerts für die einzelnen Anwendergruppen erfolgte durch Schulungen, den Lotsen-Dienst, Informationsmaterial und Rundmails. Alle Akteure wurden frühzeitig in den Entwicklungsprozess eingebunden, um Transparenz über Verbesserungen und Updates des EAAs sicherzustellen. Vor-Ort-Präsenz und enger Austausch mit ausgewählten Ansprechpartnern der Abteilungen waren ebenfalls essentiell.

Das HYKIST-Projekt zeigt, dass die erfolgreiche Einführung neuer Systeme eine umfassende Strategie erfordert, die auf enger Zusammenarbeit mit den betroffenen Berufsgruppen, benutzerfreundlicher Gestaltung und kontinuierlicher Unterstützung vor Ort basiert. Als Handlungsempfehlungen für zukünftige KI-Projekte und Evaluationsstudien kann Triaphon aus den Projekterfahrungen folgende Punkte ableiten:

- Einbindung der betroffenen Anwender in den Entwicklungsprozess von Anfang an.
- Gestaltung eines User-Journey, der möglichst wenige Berufsgruppen für eine erfolgreiche Systemnutzung adressiert.
- Implementierung in bestehende Abläufe und technische Infrastruktur.
- Etablierung eines offensichtlichen Mehrwerts der Systemnutzung für alle beteiligten Nutzergruppen.
- Präsenz vor Ort mit regelmäßiger Kommunikation über Projektstand und Ziele.

- Aktive Einholung von Kritik, Ideen und Lösungsvorschlägen der Nutzer und deren Integration in die Entwicklung.
- Kritische Prüfung des Prototyps vor dem Einsatz in der Klinik auf Nutzerakzeptanz.
- Möglicherweise gestuftes Testsetting: Erster Anwender-Test in einem weniger arbeitsverdichteten Umfeld; nach Auswertung und ggf. Optimierungen des Systems dann Tests in Arbeitsumfeld mit hoher Arbeitsverdichtung (wie Zentrale Notaufnahmen)
- Etablierung standardisierter Abläufe, beispielsweise eine digitale Erstanamnese-Erhebung für alle Patienten in ihrer Muttersprache.

### **Gesamtbeurteilung AppTek**

AppTek hat die im Projektantrag gestellten Ziele im Wesentlichen erreicht – es wurden domänenspezifische Sprachübersetzungssysteme von akzeptabler Qualität trainiert und im Laufe des Projekts verbessert. Deren der Aufgabe angemessene Qualität wurden in den Anwender-Testungen und weiteren internen Testungen evaluiert. Die Funktionen der Glossar-basierten Übersetzung und Markierung wichtiger Begriffe und deren Übersetzungen auf Wort- und Phrasenebene wurden ebenfalls erfolgreich implementiert und validiert.

Deutlich mehr Aufwand als zunächst anvisiert kostete der Aufbau des Sprachassistenten-Backends, inklusive der Definition der Schnittstellen und Testungen. Dabei spielte ein Fakt eine wesentliche Rolle: der Arzt / die Ärztin und der Patient / die Patientin sprechen auf demselben Telefonkanal. Diese Besonderheit wurde bei der Antragstellung nicht als wichtig beachtet, musste aber für die Realisierung der Gesamtlösung auf jeden Fall berücksichtigt werden, in dem eine automatische Erkennung stattfindet, wer gerade welche Sprache auf dem Kanal spricht. AppTek hat als Lösung des Problems ein System vorgeschlagen, in welchem zwei Spracherkennungssysteme (für Deutsch und für die entsprechende Fremdsprache) gleichzeitig laufen und die Erkennung, welche Sprache gerade gesprochen wurde, segmentweise erfolgt, basierend auf Konfidenzmaßen, die die beiden Spracherkennungssysteme für jedes erkannte Wort liefern. Diese Lösung wurde in den Testungen nachgebessert und letztendlich erfolgreich in der Praxis eingesetzt, mit einer geringen Fehlerrate (Fälle, wo z.B. eine kurze Äußerung auf Deutsch als eine ähnlich lautende Äußerung auf Arabisch als Arabisch inkorrekt erkannt und übersetzt wird). Dennoch hat uns diese Lösung davon abgehalten, tiefer an simultanen Übersetzungsalgorithmen zu arbeiten, wie in UAP 3.3 ursprünglich geplant. Eine solche simultane Übersetzung erlaubt es, Teile eines Satzes bereits zu übersetzen, bevor der Satz überhaupt zu Ende gesprochen wurde. Es hat sich aber in den Testungen gezeigt, dass die Erkennung für kurze Segmente nicht stabil genug ist (die

Worthypothesen ändern sich mit mehr Kontext), und daher auch die Erkennung, welche Sprache gerade gesprochen wurde, nicht verlässlich genug ist. Von daher haben sich die Aktivitäten von AppTek in UAP 3.3 vor allem darauf konzentriert, eine niedrige Latenz bei gleichzeitig hinreichendem Kontext für die Übersetzung zu ermöglichen. Die Idee, Teilübersetzungen auszugeben, bevor ein Satz zu Ende ist, wurde verworfen. Stattdessen wurde zusammen mit Partner RWTH daran gearbeitet, die Segment- bzw. Satzlänge so zu optimieren, dass auch bei kürzeren Pausen ein Segmentende generiert wird. Die so erstellten Segmente waren normalerweise lang genug, um für die Übersetzung genug Kontext bereitzustellen; gleichzeitig waren sie kurz genug, um die Verzögerung beim Anzeigen der Übersetzung eines solchen Segments in der App gering zu halten.

### **Zusammenfassende Bewertung aus Sicht aller Partner**

Auch wenn die klinische Phase in den Entwicklungsprojekten für neue digitale Angebote immer eine Herausforderung für das Konsortium darstellt, so waren im HYKIST-Projekt die Anforderungen und Erwartungen an das System für den Einsatz in der Notaufnahme besonders hoch.

An erster Stelle ist hier die echtzeitbasierte Verarbeitung der Audiostreams von der Datenquelle bis hin zur Ergebnispräsentation der finalen Übersetzungen in der HYKIST-App inklusive der erforderlichen Kanaltrennung zu nennen. Die Erkennungsrate für die 3 in HYKIST bearbeiteten Fremdsprachen Arabisch, Ukrainisch und Vietnamesisch betrug zum Ende der Entwicklungsphase ca. 75%, sodass aufgrund der noch zu optimierenden Sprachmodelle es im Pilotbetrieb auch zu Fehlerkennungen kam, die Auswirkungen auf den Ablauf hatten. Dennoch war die Qualität der Spracherkennung insbesondere für Deutsch im Vergleich zum internationalen Standard auf einem hohen Niveau. Auch für Arabisch und Vietnamesisch konnten im Laufe des Projekts signifikante Verbesserungen der Erkennungsqualität erzielt werden, indem aktuellste Entwicklungen in die eingesetzten neuronalen Architekturen eingingen.

Die Herausforderung für die maschinelle Übersetzung im HYKIST-Projekt war Systeme zu trainieren, die direkt zwischen den relevanten Sprachen übersetzen ohne die Verwendung einer Pivot-Sprache. Ein weiteres Problem für die maschinelle Übersetzung waren die gelegentlichen Spracherkennungsfehler, die logischerweise negativ die Qualität der Übersetzung beeinflussten. Die im Projekt aufgebaute sukzessive Glossarfunktion zur Forcierung korrekt erkannter Fachbegriffe half dabei, eine deutlich bessere Übersetzungsqualität zu erreichen. Die automatische Übersetzung zwischen Vietnamesisch und Deutsch war dennoch zum Teil fehlerbehaftet wegen der Knappheit der zur Verfügung stehenden Trainingsdaten und der Besonderheiten der vietnamesischen Sprache, die viel

Mehrdeutigkeit bei den verwendeten Wörtern bzw. Silben zulässt. Da es sich im HYKIST-Projekt um zu erkennende und zu übersetzende Spontansprache (oftmals in einem Dialekt) handelt, sehen wir weiterhin Verbesserungsbedarf in den Bereichen automatische Spracherkennung und maschinelle Übersetzung. Insbesondere Sprachdialekte waren in den Trainingsdaten des Spracherkennungssystems deutlich unterrepräsentiert.

Als weitere Herausforderungen für das Projekt sind die hohen Benutzeranforderungen an ein optimales multilinguales Interaktionsdesign sowohl für das EAA-Interface wie auch HYKIST-App zu sehen. Sie erforderten einen intensiven Diskussions- und Abstimmungsprozess im Projekt. Herausfordernd für das EAA-Interface war, neben der mehrere Iterationen erfordernden medizinisch-fachlichen Abstimmung des Anamnesefragens, den Einwilligungsprozess zur Studienteilnahme rein digital abzubilden und den Patienten für die Anamnesefragen ein multimodales und bediensicheres Sprachinterface bereitzustellen. Hier mussten auch in der klinischen Phase aufgrund der Praxiserfahrungen und der Rückmeldungen der Studienteilnehmer noch spontane Änderungen am Programmablauf und einzelnen Interfaceelementen erfolgen. Z.B. wurde erst kurz nach Beginn des klinischen Pilotbetriebs ein geeigneter Weg umgesetzt, um auch die Situation korrekt abzubilden, dass die Dateneingabe von EAA von Eltern als gesetzlichen Stellvertretern eines zur Behandlung erschienenen Kindes vorgenommen wird. Ebenfalls musste nach den Erfahrungen von Probandinnen mit Problemen bei der Spracheingabe (welche sowohl durch die Leistung der ASR-Modelle als auch durch den suboptimalen Aufstellort des EQQ-Terminals bedingt waren) die Alternative besser herausgestellt werden, dass eine Eingabe auch über den Touchscreen erfolgen konnte.

Unabhängig von diesen technischen Herausforderungen hat sich auch der reine Umfang der von medizinisch-fachlicher Seite definierten Anamnesefragen als problematisch erwiesen. Viele Probanden waren nach Abschluss dieses langen Fragendialoges nicht mehr bereit, noch einen weiteren Fragebogen zu ihrer Nutzungserfahrung mit dem EAA zu beantworten. Während der Pilotphase wurden deshalb weitere Logging- und Debugging-Funktionalitäten sowohl im EAA als auch im HYKIST-Backend implementiert, um aufgekommene Probleme besser eingrenzen und beheben zu können und ein feingranulareres Bild der praktischen Nutzung des Systems zu erhalten. Als eine technische Herausforderung erwies sich auch die Netzwerkperformance in den klinischen Einrichtungen: Nachdem z.B. eines der in einer Notaufnahme installierten EAA-Terminals aus baulichen Gründen nur eine Netzwerkverbindung per WLAN-erhalten konnte (kein Ethernet-Kabelanschluss möglich) und dieser drahtlose Netzwerkzugang immer wieder kurzzeitige Ausfälle zeigte, wurden am EAA mehrfache Verbesserungen vorgenommen, nach welchen kurzzeitige Netzwerkausfälle zum einen keine Prozessabbrüche mehr nach sich ziehen konnten, zum anderen die Probanden in

solchen Situationen nicht mehr mit für sie unverständlichen Fehlermeldungen konfrontiert wurden.

Die Schwierigkeit für die HYKIST-App bestand darin, die Funktionen eines komplexen Übersetzungssystems mit Alternativvorschlägen auf einer App-Oberfläche so zu komprimieren, dass sie auch noch parallel zu der eigentlichen Gesprächssituation verwendet werden konnte. Aufgrund der zeitlichen Begrenzung für die Entwicklung der Sprachassistenten konnten für die klinische Phase nicht alle Ideen umgesetzt werden, da sie noch eine tiefere inhaltliche Analyse des eigentlichen Gesprächskontextes vorausgesetzt hätten. Die viel diskutierten Ansätze zur Vereinfachung des Interfaces wie z.B. einzelne ein- und ausblendbare Sichten auf die übersetzten Daten, die Markierung vorübergehender und finaler Übersetzungsergebnisse sowie die farbliche Hervorhebung von Fach- und Glossarbegriﬀen wurden jeweils in den Ausbaustufen umgesetzt, sie basierten aber ausschließlich auf dem linearen Übersetzungsstream aus dem Sprachassistentenbackend, der entsprechend im Frontend gefiltert wurde. Gewünscht wurde darüber hinaus von den Sprachmittlern immer wieder eine inhaltliche Präsentation der wichtigsten Begriffe im Gespräch und die prägnante Hervorhebung zentraler Daten wie Angaben zur Medikation, zur Diagnose oder zu weiteren Behandlungsschritten: Diese Funktionen, die HYKIST-App quasi als intelligente Gedächtnisstütze für Sprachmittler umzusetzen, konnte aufgrund des gewählten Ansatzes für das Sprachassistentenbackend im Projekt leider nicht mehr umgesetzt werden.

Auch der Systembetrieb unter realen Bedingungen in der Notaufnahme eines Krankenhauses für Maximalversorgung ist als große Herausforderung zu bezeichnen. Nachdem die hohen Erwartungen des medizinischen Personals aufgrund der technischen Einschränkungen des Forschungsprototypen nicht erfüllt wurden, mussten im Projekt Lotsen eingesetzt und ein Erwartungsmanagement zu betreiben und die Anwender in der Nutzung zu unterstützen. Die gegebene Zeitknappheit des medizinischen Personals bei gleichermaßen gesundheitskritischen Versorgungsprozessen führte schnell dazu, dass das System nicht aktiv für den Versorgungsprozess angenommen wurde. Die naht- und reibungslose Integration des Systems direkt in den Prozess der Notaufnahme als ein Projektziel konnte somit leider nicht vollends erreicht werden. Zur Verbesserung des Systemverhaltens wurden deshalb noch während der Pilotphase Anstrengungen zur Optimierung des Gesamtsystems unternommen. Diese betrafen zum einen das Zeitverhalten des Systems; hierzu wurden durch Messungen und Protokollierungen an verschiedenen Stellen der Prozessierungskette die Ursachen für erhöhte Latenzen identifiziert und abgestellt. Zum anderen ging es darum, die Resilienz des Systems zu steigern und insbesondere auch seine Stabilität unter widrigen Bedingungen zu sichern. Auch die im Konsortium getroﬀenen Maßnahmen zur Erhöhung der Teilnehmerzahl waren von mäßigem Erfolg, sodass zum Schluss für die Evaluationsarbeiten zum Teil auf simulierte Dialoge zurückgegriffen werden musste. Zusammen mit den simulierten Dialogen

konnten aber dennoch abschließend 170 Trialoge mit dem HYKIST-System durchgeführt werden (111 HYKIST-Trialoge im Sanaklinikum, davon 48 echte Trialoge und 63 simulierte, sowie 59 simulierte HYKIST-Trialoge in der Universität zu Köln).

In zukünftigen Arbeiten könnte durch die Optimierung der Spracherkennung und maschinellen Übersetzung die Fehlerquote auf ca. 10% reduziert werden. Ebenso sollte für zukünftige Arbeiten vermehrt auf die Erkennung wichtiger Gesprächsinhalte und deren leicht erfassbare Darstellung für die Sprachmittler fokussiert werden. Das Gesamtkonzept eines Assistenzsystems für die Sprachmittlung wird seitens der Anwender wie auch Projektpartner weiterhin als erfolgsversprechend angesehen.

## 8 Verbreitung und Öffentlichkeitsarbeit der Projektergebnisse

Folgende Meldung zu dem Projekt wurde von Fraunhofer FOKUS publiziert:

[https://www.fokus.fraunhofer.de/de/fokus/news/hykist\\_2022\\_12](https://www.fokus.fraunhofer.de/de/fokus/news/hykist_2022_12)

Das Projekt ist ebenso auf den Webseiten von Fraunhofer FOKUS präsent:

<https://www.innovationszentrum-telehealth.de/go/hykist>

<https://www.fokus.fraunhofer.de/de/dps/projekte/hykist>

Ebenso beteiligte sich das HYKIST-Konsortium an den Vernetzungstreffen des BMGs.

Der HYKIST-Demonstrator wurde auf der Data4health-Konferenz des BMG am 20./21.06.2023 präsentiert.

Darüber hinaus erfolgte mehrere Publikationen der aktuellen Projektergebnisse auf nationalen und internationalen Tagungen (siehe dazu Kap. 10 Publikationsverzeichnis)

## 9 Verwertung der Projektergebnisse

### Verwertung der Projektergebnisse Fraunhofer FOKUS

Fraunhofer FOKUS arbeitet nach dem HYKIST-Projekt und unter Aufbau auf dessen Resultate weiter daran, die Nutzung von Sprachassistenzsystemen in weiteren Anwendungsdomänen (wie öffentliche Verwaltung und produzierende Industrie) voranzutreiben und dafür tragende Verfahren und Technologien zu entwickeln. Eine wichtige Rolle spielt dabei das von FOKUS unter Aufbau auf Erfahrungen aus früheren Sprachassistenz-Projekten, aber auch HYKIST entwickelte FISA-Framework (FOKUS Intelligent Speech Assistant), welches mit geringem Aufwand ermöglicht, existierende Applikationen minimal-invasiv mit Sprachassistenzfunktionalität zu erweitern. FISA unterstützt technologieoffen und anbieterübergreifend eine leichte Integration verschiedener sprachbezogener Basisdienste

(Spracherkennung, Sprachsynthese und maschinelle Übersetzung). Somit ist eine Nutzung sowohl entsprechender kommerzieller Services als auch geeigneter Open-Source-Lösungen möglich, bei Bedarf auch kombiniert.

FOKUS ist Gründungsmitglied der vom Stab des Vorstands der Fraunhofer-Gesellschaft ins Leben gerufenen dedizierten Task Force „KI-basierte Sprachassistenten“, die die Kompetenzen, die Erfahrungen und das Partnernetzwerk mehrerer Fraunhofer-Institute zusammenbringt, um die Marktchancen dieser vielversprechenden Technologie, konkreter zu verfolgen und mittels angestrebter gemeinsamer Projektaktivitäten zu nutzen. Eine wesentliche Zielrichtung für diese Vorhaben wird es sein, die aus Vorgängerprojekten verfügbaren Erfahrungen und Technologien mit dem Einsatz leistungsfähiger generativer KI-Verfahren und großer Sprachmodelle (large language models, LLMs) zu kombinieren, um eine den durch Systeme wie ChatGPT massiv gestiegenen Nutzererwartungen an konversationale Systeme adäquat entgegenzukommen.

In den Anwendungsdomänen Gesundheit und Öffentliche Verwaltung spielen Ziele wie Datenschutz und Datensouveränität eine besondere Rolle – daher konzentriert sich FOKUS in naher Zukunft besonders auf die Erprobung von Open-Source-LLMs, welche einen on-premises-Betrieb erlauben und somit Unabhängigkeit von den Cloud-Lösungen der derzeitigen Marktführer ermöglichen, und baut derzeit ein Testbed zur Evaluierung und anwendungsbezogenen Anpassung dieser Art von Modellen auf.

Die im HYKIST-Projekt entwickelten Komponenten von Fraunhofer FOKUS können für eine Vielfalt weiterer Anwendungsfälle nutzbar gemacht werden, z.B. in der Katastrophenmedizin, oder für multilinguale Erste-Hilfe-Informationen im Alltag. Ebenso bieten sich Anschlussprojekte z.B. im Rahmen des Innovationsfonds auf dem Gebiet der echtzeitbasierten Sprachverarbeitung in der medizinischen Versorgung an. Die Lessons learned können publiziert und in Folgeprojekten Technologiepartnern in der Medizinbranche vermittelt werden. Darüber hinaus wurde ein Folgeprojekt zur psychiatrisch-psychosozialen Unterstützung von Geflüchteten bereits als Skizze formuliert.

### **Verwertung der Projektergebnisse RWTH Aachen**

Entwicklungen zu Spracherkennungssystemen der RWTH Aachen geschehen auf Basis der kontinuierlich gepflegten Toolkits RWTH ASR (RASR) sowie dem RWTH Extensible Training framework for Universal Recurrent Neural Networks (RETURNN) und sind auf diesem Weg, für Forschungszwecke (frei) und für eine wirtschaftliche Verwertung verfügbar. Die wirtschaftliche Verwertung geschieht u.a. innerhalb einer Kooperation mit dem Projektpartner AppTek.

Wissenschaftlich brachten die Entwicklungen der Spracherkennungssysteme in HYKIST vielfältige Erfahrungen, die auch schon in Folgearbeiten eingegangen sind und auf

internationalen Benchmarks Ergebnisse auf Höhe des aktuellen Stands der Technik zeigen. Die in HYKIST erzielten Ergebnisse zur automatischen Spracherkennung wurden in 11 wissenschaftlichen Veröffentlichungen auf unterschiedlichen Konferenzen im Bereich der automatischen Verarbeitung natürlicher Sprache vorgestellt und stehen der weltweiten Forschung somit zur Verfügung. So konnten u.a. neue neuronale Modellarchitekturen (insb. Conformer) erfolgreich zur Verbesserung der Performanz der Spracherkennung eingesetzt werden, deren Arbeiten [ZX22a, ZX22b] seit Veröffentlichung in 2022 schon signifikant zitiert wurden. Ebenso wurden große vortrainierte Modelle zur Adaption an Aufgaben mit geringen Adaptiondatenmengen untersucht [VL21, VS23], ein mittlerweile sehr aktiver Zweig der Forschung, nicht nur im Bereich der automatischen Spracherkennung. Die Erfahrungen zur Entwicklung der Spracherkennungssysteme in HYKIST werden in [LZ23] zusammengefasst.

### **Verwertung der Projektergebnisse Triaphon**

Der Sprachmittler-Pool für Ukrainisch, den Triaphon im Rahmen von Hykist+ aufgebaut hat, ist nun Teil des regulären Betriebs von Triaphon. Dadurch steht neben der Sprache „Russisch“ nun auch "Ukrainisch" als neunte Sprache für die Kunden von Triaphon (Kliniken, Arztpraxen, medizinische Beratungsstellen) zur Verfügung.

Im Verlauf des Projekts konnte Triaphon zahlreiche Erfahrungen und Erkenntnisse sammeln, insbesondere über den spezifischen Unterstützungsbedarf von Sprachmittlern und Sprachmittlerinnen. Diese Erkenntnisse werden in die zukünftige Weiterentwicklung der Serviceleistungen von Triaphon einfließen, insbesondere im Bereich der Sprachmittlung

Durch das HYKIST-Projekt hat Triaphon eine realistische Einschätzung der Chancen und Grenzen der aktuellen technischen Möglichkeiten zur Unterstützung von Sprachmittlern und Sprachmittlerinnen gewonnen. Es hat sich gezeigt, dass die Verdolmetschung eines medizinischen Gesprächs in Notaufnahmen derzeit am besten von Personen durchgeführt werden kann, die beide Sprachen sehr gut beherrschen und mit beiden Kulturen vertraut sind. Bei dieser kognitiv und emotional sehr anspruchsvollen Tätigkeit können die Sprachmittler und Sprachmittlerinnen durch technische Hilfsmittel durchaus unterstützt werden.

### **Verwertung der Projektergebnisse AppTek**

AppTek wird das im Rahmen des Projekts aufgebaute Sprachassistenten-Backend weiter kommerziell nutzen können für Echtzeit-Spracherkennung und Sprachübersetzung sowohl in den Telefonie-Anwendungen als auch z.B. bei automatischer Verschriftung und Übersetzung von Nachrichtensendungen und Talk-Shows.

Die Möglichkeit, bestimmte Begriffe zu markieren und diese Markierung auf die automatischen Übersetzungen dieser Begriffe zu übertragen, die speziell für HYKIST bei AppTek entwickelt wurde, erlaubt uns, auch in anderen Anwendungen Markierungen wie HTML/XML Tags auf

Wort- und Phrasenebene in der maschinellen Übersetzung beizubehalten. Dies ist oft in verschiedenen Kontexten von unseren Kunden gewünscht, und die allermeisten Anbieter von kommerziellen MT-Systemen bieten eine solche Lösung nicht an, da sie sich in neuronalen Übersetzungsmodellen nur schwer und mit viel Kreativität implementieren lässt.

Wissenschaftlich interessant für AppTek war auch, dass künstlich (durch ein GPT-basiertes Modell) erzeugten domänenspezifischen parallelen bilingualen Daten enorm hilfreich sein können, die Übersetzungsqualität in einer bestimmten, sogar sehr eng definierten Domäne deutlich zu verbessern, wie das bei HYKIST für medizinische Gespräche zwischen medizinischem Personal und Patienten und Patientinnen der Fall war. Wir werden diesen Ansatz in weiteren Experimenten und auch auf anderen Sprachpaaren tiefer untersuchen. Ebenfalls wichtig waren für uns die Erkenntnisse, wie lang Äußerungen sein müssen, um die gesprochene Sprache mit einer guten Sicherheit automatisch zu identifizieren. Diese Erkenntnisse spielen eine wichtige Rolle bei unseren weiteren Forschungen im Bereich der multilingualen Spracherkennung und Übersetzung.

## 10 Publikationsverzeichnis

[ZG21] M. Zeineldeen, A. Glushko, W. Michel, A. Zeyer, R. Schlüter, H. Ney: "Investigating Methods to Improve Language Model Integration for Attention-based Encoder-Decoder ASR Models," *Interspeech*, Brno, Czech Republic, Aug./Sept. 2021, arXiv:2104.05379.

[VL21] P. Vieting, C. Lüscher, W. Michel, R. Schlüter, H. Ney: "On Architectures and Training for Raw Waveform Feature Extraction in ASR," *IEEE Automatic Speech Recognition and Understanding (ASRU) Workshop*, Cartagena, Colombia, Dec. 2021, arXiv:2104.04298.

[RB22] T. Raissi, E. Beck, R. Schlüter, H. Ney: "Improving Factored Hybrid HMM Acoustic Modeling without State Tying," *IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, May 2022, arXiv:2201.09692.

[ZX22a] M. Zeineldeen, J. Xu, C. Lüscher, R. Schlüter, H. Ney: "Conformer-based Hybrid ASR System for Switchboard Dataset," *IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, May 2022, arXiv:2111.03442.

[ZX22b] M. Zeineldeen, J. Xu, C. Lüscher, R. Schlüter, H. Ney: "Improving the Training Recipe for a Robust Conformer-based Hybrid Model," *Interspeech*, Incheon, South Korea, Sep. 2022, arXiv:2206.12955.

- [YG22] Z. Yang, Y. Gao, A. Gerstenberger, J. Jiang, R. Schlüter, H. Ney: "Self-Normalized Importance Sampling for Neural Language Modeling," *Interspeech*, Incheon, South Korea, Sep. 2022, arXiv:2111.06310.
- [RZ23] T. Raissi, W. Zhou, S. Berger, R. Schlüter, H. Ney: "HMM vs. CTC for ASR: Comparison Based on Full-Sum Training from Scratch," *IEEE Spoken Language Technology Workshop (SLT)*, Doha, Qatar, January 2023, arXiv:2210.09951.
- [YZ23] Z. Yang, W. Zhou, R. Schlüter, H. Ney: "Lattice-Free Sequence Discriminative Training for Phoneme-based Neural Transducers," *IEEE Intern. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, June 2023, arXiv:2212.04325.
- [RL23] T. Raissi, C. Lüscher, M. Gunz, R. Schlüter, H. Ney: "Competitive and Resource Efficient Factored Hybrid HMM Systems are Simpler Than You Think," *Interspeech*, Dublin, Ireland, Aug. 2023, arXiv:2306.09517.
- [LZ23] C. Lüscher, M. Zeineldeen, Z. Yang, P. Vieting, K. Le-Duc, W. Wang, R. Schlüter, H. Ney: "Development of Hybrid ASR Systems for Low Resource Medical Domain Conversational Telephone Speech," *ITG Conference on Speech Communication*, Aachen, Germany, Sep. 2023, arXiv:2210.13397.
- [VS23] P. Vieting, R. Schlüter, H. Ney: "Comparative Analysis of the wav2vec 2.0 Feature Extractor," *ITG Conference on Speech Communication*, Aachen, Germany, Sep. 2023, arXiv:2308.04286.
- [BW23] P. Bahar, P. Wilken, J. Iranzo-Sánchez, M. Di Gangi, E. Matusov, Z. Tüske: "Speech Translation with Style: AppTek's Submissions to the IWSLT Subtitling and Formality Tracks in 2023". *Proceedings of the 20th International Conference on Spoken Language Translation (IWSLT 2023)*, pages 251–260, Toronto, Canada, Jul. 2023.
- [WM22] P. Wilken and E. Matusov: "AppTek's Submission to the IWSLT 2022 Isometric Spoken Language Translation Task". *Proceedings of the 19th International Conference on Spoken Language Translation (IWSLT 2022)*, pages 369–378, Dublin, Ireland, May 2022.
- [BW21] P. Bahar, P. Wilken, M. A. Di Gangi, E. Matusov: "Without Further Ado: Direct and Simultaneous Speech Translation by AppTek in 2021". *Proceedings of the 18th International Conference on Spoken Language Translation (IWSLT 2021)*, pages 52–63, Bangkok, Thailand (online), Nov. 2021.
- [MW20] E. Matusov, P. Wilken, C. Herold: "Flexible Customization of a Single Neural Machine Translation System with Multi-dimensional Metadata Inputs". *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas (Volume 2: User Track)*, pages 204–216, virtual, Oct. 2020.