

# Wir im Kiez – eine multimodale App zur gegenseitigen Nachbarschaftshilfe für Senioren

N. Reithinger<sup>1</sup>, A. Russ<sup>1</sup>, S. Schmeier<sup>1</sup>, A. Aue<sup>2</sup>, N. Beck<sup>2</sup>, and M. Gerlitschka<sup>2</sup>

<sup>1</sup>Deutsches Forschungszentrum für künstliche Intelligenz, Berlin, [norbert.reithinger, aaron.russ, sven.schmeier]@dfki.de

<sup>2</sup>Cocomore AG, Frankfurt, [andreas.aue, nico.beck, meike.gerlitschka]@cocomore.com

## Kurzfassung

In dem vom Bundesministerium für Bildung und Forschung geförderten Projekt „Wir im Kiez – Ich mach was für Dich, Barrierefreie Cloud für Senioren“ (Förderkennzeichen: 16SV6323) wurde eine multimodale App entwickelt, über die ältere Menschen miteinander in Kontakt treten können, um sich gegenseitig bei alltäglichen Problemen und Fragestellungen zu helfen. Der zentrale Gedanke dabei war die Realisierung einer virtuellen Nachbarschaftshilfe für Menschen, die keine oder nur noch wenig soziale Kontakte haben. Personen können Unterstützung im Kiez anbieten („Ich kann Gardinen aufhängen“, „Ich helfe beim Umgraben“,...), die in der Nachbarschaft nachgefragt werden, umgekehrt können Hilfesuche aufgegeben werden („Ich muss für 4 Tage ins Krankenhaus, wer kann meinen Wellensittich versorgen“). Die Hauptzielgruppe von „Wir im Kiez“ sind Senioren, die App wurde primär für Tablets entwickelt.

## 1 Aufgabenstellung

Die Gesellschaft in Deutschland verändert sich hinsichtlich der Altersstruktur und der Bevölkerungszahl. Entsprechend dieses demographischen Wandels nimmt der Anteil der Senioren an der Gesamtbevölkerung in Relation zu den Jüngeren bereits seit Jahrzehnten zu. Unter den Senioren gibt es auf der einen Seite viele Alleinlebende, die bestimmte Alltagssituation, wie z.B. Gartenarbeit, Hausputz oder Einkaufsgehen, nur noch schwierig oder gar nicht mehr alleine bewältigen können oder häufig einsam und alleine sind. Auf der anderen Seite gibt es viele Personen, die sich auch im Alter noch fit genug fühlen, um sich einzubringen und die ihre Energie für das Wohl der Gesellschaft einzusetzen möchten. Das Projekt „Wir im Kiez“ zielte darauf ab, die Interessen dieser beiden Gruppen auf innovative Art und Weise mittels einer sicheren, informationstechnischen Lösung zusammenzubringen. Da viele ältere Personen in wachsendem Ausmaß Smartphones und Tablets besitzen, stellten diese Geräteklassen eine ideale Plattform für das Vorhaben dar. Zur Verbindung der Interessen der beiden Gruppen auf einer cloud-basierten Plattform wurde eine virtuelle Nachbarschaftshilfe realisiert. Dort werden Hilfsangebote- und Gesuche in Bezug auf die individuelle, nachbarschaftliche Umgebung der Nutzer veröffentlicht und Menschen treten miteinander in Kontakt. Um sicherzustellen, dass die Zielgruppe die Plattform trotz unterschiedlicher, altersbedingter Einschränkungen sowie der teilweise geringen Erfahrung im Umgang mit technischen Geräten möglichst uneingeschränkt nutzen kann, mussten spezielle Anforderungen an die Usability erfüllt werden. Diese Anforderungen im Detail zu evaluieren war wesentlicher Bestandteil des Projektes. Das Sprachdialogsystem der Plattform ermöglicht die Interaktion per Sprache alternativ zur konventionellen Touch-Bedienung von Smartphones und Tablets. Dies

unterstützt insbesondere eine barrierefreie Bedienung der Plattform durch Menschen mit stark eingeschränkter Motorik oder Sehkraft.

## 2 Stand der Wissenschaft und Technik

Das Projekt „Vernetzte Nachbarschaft“ der Universität der Künste Berlin<sup>1</sup> verfolgt den Ansatz, ältere Personen aus der Nachbarschaft zusammenzubringen. Zusammen mit dem SeniorenComputerClub Berlin wird versucht, die Kommunikationslücke zu schließen und u.a. auch Legacy-Technologien wie Papierbrief mit einzubeziehen. Im Mittelpunkt dieses Projekts standen Designaspekte.

Im kommerziellen Umfeld gibt es einige Systeme, die Teilaspekte der geplanten Umsetzung von Wir im Kiez scheinbar realisieren. Das System der peopleAG<sup>2</sup> implementiert eine niederschwellige Vermittlung für leichte Helfertätigkeiten, wie wir sie in unserem Projekt vorgesehen haben. Das System Niri – dieses System ist in der Zwischenzeit aufgegeben worden – realisiert ein nachbarschaftsbasiertes soziales Netzwerk. Beide sind jedoch bezüglich ihrer Bedienbarkeit und Sicherheit nicht hinreichend für die speziellen Bedürfnisse der Zielgruppe von älteren Benutzer ausgelegt. Zudem mangelt es ihnen an modernen Merkmalen wie mobiler und sprachlicher Interaktion. Im Bereich der sozialen Vernetzung wurden im Projektzeitraum einige Plattformen realisiert, z.B. [www.nebenan.de](http://www.nebenan.de). Alle diese Plattformen folgen jedoch dem Standardansatz für soziale Netzwerke und untersuchen oder verwenden insbesondere keine fortgeschrittenen Interaktionstechnologien, wie dies in „Wir Im Kiez“ erfolgreich erforscht und getestet wur-

<sup>1</sup><http://www.design-research-lab.org/?projects=neighborhood-labs>

<sup>2</sup><https://www.peopleag.de/>

de. Zudem konnten wir für „Wir Im Kiez“ auf Erfahrungen aus einem der weltweit größten Forschungsprojekte im Bereich AAL zurückgreifen. In dem Projekt SmartSenior wurde das Themenfeld altersgerechte Benutzerschnittstellen von Styleguides für AAL-Anwendungen bis hin zu implementierten tabletauglichen Systemen bearbeitet [1, 4]. Vor dem Hintergrund dieser Arbeiten ist uns kein System in Forschung und Industrie bekannt, das die verschiedenen Aspekte gleichwertig abdeckt.

### 3 Aufbau und Realisierung der Plattform

Im Zentrum des Projektes stand eine Plattform, über die ältere Menschen miteinander in Kontakt treten können, um sich gegenseitig bei alltäglichen Problemen und Fragestellungen zu unterstützen. Dabei stand die Realisierung einer virtuellen Nachbarschaftshilfe für Senioren im Zentrum, die sich an vitale und hilfsbereite Senioren ebenso wendet, wie an solche mit altersbedingten Einschränkungen oder keinen bis wenigen sozialen Kontakten. Voraussetzung war dabei, eine möglichst barrierefreie Anwendung mit geringer Hemmschwelle zu schaffen, die sowohl von eingeschränkten wie auch motorisch und visuell nicht beeinträchtigten Menschen intuitiv und einfach bedient werden kann. Recherchen und ergänzende, alterssimulierende Untersuchungen – mittels Fehlsichtigkeitsbrille „Aging Glasses“ und Simulationshandschuh „Tremor-Explorer“ der Age Suite Germany GmbH – zeigten, dass zur Erfüllung dieser Anforderungen eine multimodale Anwendung realisiert werden muss, die nicht ausschließlich per Touch bedienbar ist. Einschränkungen des Sichtfeldes sowie motorische Störungen, die es nicht oder nur bedingt erlauben, touch-sensitive Elemente auf dem Bildschirm zu aktivieren, sprachen für eine alternative Bedienung mittels Sprachdialog sowie eine Funktion, die relevante Bildschirminhalte rezitiert. Damit das hohe Sicherheitsbedürfnis der Nutzer befriedigt und deren Skepsis gegenüber sozialen Plattformen reduziert werden konnten, musste zudem ein hoher Standard in Bezug auf den Datenschutz sowie den Schutz der Privatsphäre für den Erfolg der Anwendung definiert werden. Als Mindestvoraussetzungen galten daher:

- Ein rollenbasiertes Rechtssystem, welches (mindestens) zwischen nicht registrierten (öffentlich), registrierten (informell angemeldet) und verifizierten Nutzer (überprüfte Anmeldung) unterscheiden kann und entsprechend Zugriff auf sensible Daten und Anwendungsbereiche verweigert bzw. gewährt.
- Ein sicherer Prozess, bei dem die Daten, die bei der Registrierung eines Nutzers angegeben wurden durch eine dritte Stelle geprüft werden können, um einen Nutzer in den Status eines verifizierten Mitgliedes zu überführen. Die Prüfung kann beispielsweise durch einen Datenabgleich mit dem Personalausweis des zu verifizierenden Nutzers erfolgen.
- Eine klare Kennzeichnung des Status jedes Nutzers

innerhalb der Anwendung, sodass andere Nutzer zwischen verifizierten und nicht verifizierten Nutzern unterscheiden können.

Um eine zur aktiven Teilnahme anregenden und Mehrwert stiftende, virtuelle Nachbarschaftshilfe umsetzen zu können, wurden folgende Funktionen als Basis definiert:

- Aufgeben neuer Gesuche und Angebote via Tastatur- oder Spracheingabe.
- Durchsuchen von Gesuchen und Angeboten via Tastatur- oder Spracheingabe.
- Nachrichtensystem zur Kommunikation zwischen den Teilnehmern.
- Aktive Benachrichtigung, z.B. bei neuen Nachrichten oder möglicherweise interessanten, übereinstimmenden Anzeigen („Push-Meldungen“).
- Aussagekräftige Nutzerprofile mit Kontaktmöglichkeiten.
- Angebote zur Freizeitgestaltung, z.B. in Form eines Event-Kalenders.
- Sämtliche Nutzer und inhaltlichen Elemente werden mit einer Geoposition versehen, sodass die Bildung virtueller „Kieze“ möglich ist und Nutzer den Informationsfluss auf Ihre individuelle Nachbarschaft begrenzen können.

Um eine Grundlage für die Gestaltung aller geplanten Komponenten und Inhalte der Anwendung zu bilden, wurden Szenarien in gemeinsamen Workshops vom DFKI und Cocomore erarbeitet. Folgende wichtigste Ideen und Szenarien waren das Ergebnis:

- Austausch von Hilfssuchen und -angeboten in der Nachbarschaft („Ich kann Gardinen aufhängen“), möglicherweise auch im Tausch gegen andere Leistungen.
- Kontaktmöglichkeiten von bekannten und (bisher) unbekanntem Menschen, auch über größere Entfernungen hinweg (z. B. Tochter lebt 400km weit entfernt).
- Automatische Benachrichtigung über neue und relevante Hilfssuchen und -angebote und persönliche Nachrichten per Push-Dienst.
- Aktuelles aus der Nachbarschaft in einer Aktivitätsanzeige.
- Steuern und Diktieren der gesamten Anwendung per Sprache als Alternative zu Touch oder Tastatur/Maus.
- Digitales Mehrgenerationenhaus mit einem „Hausmeister“ oder Moderator als Vertrauensperson und Ansprechpartner bei Problemen mit der Anwendung.

In Zusammenarbeit mit der BAGSO<sup>3</sup> wurden in Workshops in Berlin, Treptow-Köpenick, die o. g. Szenarien iterativ vorgestellt und mit den Vertretern der verschiedenen

<sup>3</sup>Die Bundesarbeitsgemeinschaft der Senioren-Organisationen

Interessensgruppen diskutiert und abgeglichen. Dabei sind folgende, wertvolle Einsichten generiert worden (Nennung der Wichtigsten):

- Nicht nur Hilfsangebote und -gesuche sind für die Zielgruppe wichtig, sondern auch das Anbieten und Suchen von Freizeitaktivitäten und gemeinsamen Ausflügen.
- Die Diskussionsteilnehmer vermuten ein hohes Sicherheitsbedürfnis der Senioren – dieses sollte sich auch in der Anwendung widerspiegeln.
- Eine Registrierung der Nutzer vor Ort mit Unterstützung von Pflegediensten oder öffentlichen Stellen (z. B. Bürgerämter oder Wohnungsbaugesellschaften) wäre aufgrund der möglichen Sicherstellung der Identität und damit des Schutzes aller Nutzer wünschenswert. Auch würde dies die Hürde des Erstkontaktes mit der Anwendung reduzieren und eine Unterstützung bei dem Registrierungsprozess ermöglichen.
- Ein zuvor angedachtes Bewertungssystem für Helfer wurde verworfen, um keine negativen Bewertungen für freiwillige oder ehrenamtliche Helfer zuzulassen.
- Auch professionelle Hilfe ist gewünscht, um einerseits eine Finanzierung der Plattform zu erleichtern und andererseits bestimmte Dienste anzubieten, die private Personen nicht leisten können oder dürfen (Wo verläuft die Grenze zwischen Hilfe und professioneller Pflege, z. B. „Spritze setzen“?).
- Die Diskussionsteilnehmer sind sich einig, dass eine Unfall- oder Haftpflicht-Versicherung der Nutzer der Plattform ein besonders kritisches Thema ist. Dessen detaillierte Untersuchung war allerdings nicht Teil dieses Forschungsprojektes.

## 4 Zusammenarbeit mit der Zielgruppe und Evaluation

Zu Beginn des Projektes wurden innerhalb von Fokusgruppen und in teilstrukturierten Interviews (Altersgruppe zwischen 30 und 90 Jahren, Fokus auf 60-90) insgesamt 45 Fragebögen befüllt und ausgewertet, auf Basis derer Inhalte eine Haupt-Persona und zwei Neben-Personas entstanden. Diese Personas dienten von da an dem Team zur Entwicklung und Verifizierung der Gestaltung der Anwendung („Würde Margot Lammer die Funktion wirklich so benötigen und auch nutzen können?“). Zur Auswertung und Clusterung aller bis dato gesammelten Erkenntnisse wurde ein Affinitätsdiagramm gewählt, um auch in späteren Projektphasen schnell auf Änderungen aufgrund neuer Erkenntnisse aus den weiteren Usability-Tests und Interviews reagieren zu können.

Die entwickelten Konzepte wurden im gesamten Zeitraum der Entwicklung iterativ in Usability-Tests, Experten-Fokusgruppen und mit Endnutzern auf Ihre Gebrauchstauglichkeit und Akzeptanz verifiziert (siehe unten). Nach

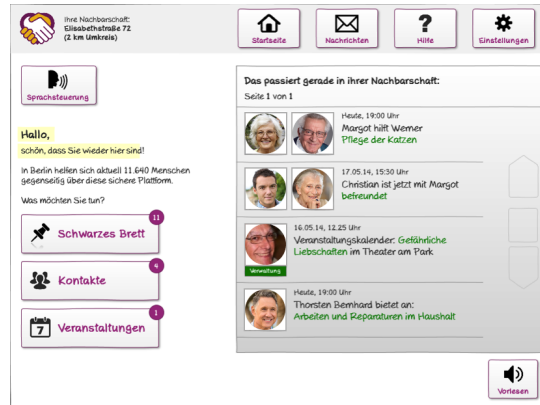


Abbildung 1 Klickbarer Prototyp

der Ausgestaltung der Wireframes für Teile der o. g. Bereiche wurde ein klickbarer Prototyp (Abbildung 1) erstellt und in einem ersten Usability-Test im September 2014 in Berlin unter kontrollierten Bedingungen mit 5 Probanden aus der Zielgruppe evaluiert – darunter auch ein Proband mit Schlaganfall und ein Proband mit Tremor. Die wichtigsten Erkenntnisse:

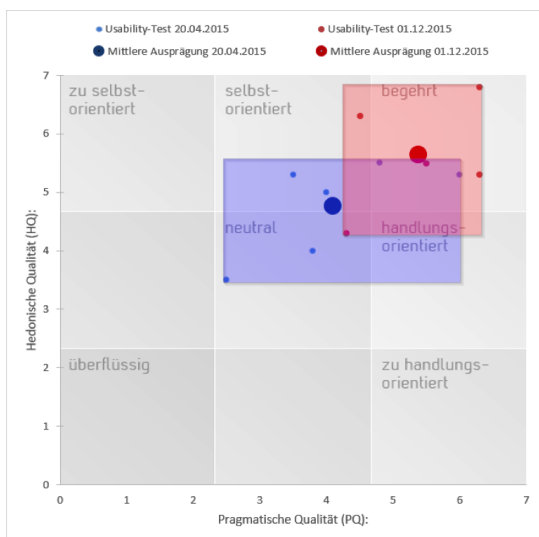
- Generell besitzt die Anwendung in diesem frühen Stadium bereits eine erstaunlich hohe Akzeptanz und lässt sich von den meisten Probanden intuitiv bedienen.
- Swipen und Wischen auf dem Tablet stellt für Probanden mit Tremor eine unangenehme Erfahrung dar.
- Bei den Eingabefelder-Dialogen der Spracheingabe war unklar, welcher Button welche Aktion auslöst.
- Manche Schaltflächen wurden optisch nicht wahrgenommen bzw. verstanden (z. B. Tab „Nachbarn“).
- Angebote oder Gesuche wurden nicht durchgesehen, bevor eine eigene Anzeige aufgegeben wurde.

Die identifizierten Probleme wurden nach Schweregraden eingeteilt, sortiert und durch Neukonzeptionen behoben. Diese sind in internen Tests erneut evaluiert worden. Auf Basis der aktualisierten und konzeptionell erweiterten Wireframes wurde anschließend ein erstes Layout (visuelles Design) erstellt. Dabei wurde ein Farbschema erarbeitet, welches der Orientierung der Nutzer dienen und dabei die interaktiven Elemente betonen sollte. Die Layouts wurden in einem interaktiven, funktionalen Prototyp implementiert und dann in einem zweiten Usability-Test im April 2015 in Saarbrücken mit 6 Probanden unter kontrollierten Bedingungen evaluiert. Zusätzlich wurde mittels AttrakDiff-Fragebogen [2] eine genauere Einschätzung der hedonischen und pragmatischen Qualität der Anwendung erfasst.

- Generell war die Anwendung und die Sprachsteuerung bereits in großen Teilen intuitiv bedienbar. Auch die AttrakDiff-Auswertung ergab ein überdurchschnittliches Ergebnis in der pragmatischen



**Abbildung 2** AttrakDiff-Auswertung des zweiten Usability-Tests



**Abbildung 3** AttrakDiff-Auswertung des dritten Usability-Tests

und hedonischen Qualität (Abbildung 2). Die mittlere Ausprägung war im Charakterbereich „selbst-orientiert“. Die Größe und Verortung des Konfidenz-Rechtecks lässt allerdings darauf schließen, dass das Produkt nicht eindeutig diesem Bereich zuordenbar ist und dass sich die Nutzer nicht ganz einig bei der Beurteilung waren.

- Klick- und Tap-Flächen waren zu klein, teilweise keine Reaktion auf die Eingabe.
- Schriftgröße und Abmessungen zu klein, Lesbarkeit suboptimal, Kontrast zu niedrig.
- Es wurde z. T. nicht immer erkannt, dass natürlichsprachliche Dialoge mit dem System möglich sind.

Im Rahmen eines Technology Experience Café (TEC) wurde im April 2015 in Saarbrücken eine weitere Fokusgruppenbefragung durchgeführt. TEC ist ein interkulturel-

les Begegnungsformat mit dem Ziel, den Dialog zwischen älteren Menschen und Produzenten von Technologie und Dienstleistung zu verbessern. Es wurde im Rahmen des EU-Projektes SiforAge entwickelt. Mit mehr als 20 Senioren und Seniorinnen wurden in kleinen Gruppen Demonstrationen an der entwickelten Plattform durchgeführt. Auswertungen der Fragebögen und Interviews zeigten eine sehr positive Bewertung des Prototyps. Speziell die Steuerung des Systems über natürliche Sprache, realisiert durch die speziell entwickelte hybride Sprachverstehens-Komponente, eingebettet in die Dialogplattform MMIR (Mobile Multimodal Interaction and Rendering), fand dort Gefallen.

Der dritte Usability-Test im November 2015 in Berlin mit 5 Probanden unter kontrollierten Bedingungen zeigte die Fortschritte bei der Ergonomie und Gestaltung. Die Probanden konnten den Prototyp nun ausnahmslos intuitiv und ohne vorherige Anleitung bedienen. Die AttrakDiff-Auswertung ergab nochmals eine deutliche Verbesserung der hedonischen und der pragmatischen Qualität (Abbildung 3). Die mittlere Ausprägung lässt darauf schließen, dass das Produkt in der vorliegenden Form „begehrt“ ist – ein ausgezeichnetes Ergebnis. Die Größe des Konfidenz-Rechtecks zeigt im Vergleich, dass sich die Nutzer etwas einiger bei der Beurteilung als im vorangegangenen Usability-Test waren (vgl. Abbildung 2).

## 5 Wissenschaftliche Herausforderungen und Ergebnisse

Die multimodale Dialogverarbeitung wird im MMIR-Framework per SCXML (W3C-Standard) definiert: multimodale Eingaben werden Ereignis-getrieben verarbeitet (Fusion) und gemäß dem aktuellen Anwendungszustand ausgewertet und auf die UI angewendet (Fission). Die Spracheingabe ist nach Aktivierung „always-on“, d.h. im Unterschied zu klassischen „push-to-talk“ muss nicht nach jedem Befehl erneut die Spracheingabe aktiviert werden, um neue Befehle einzugeben. Stattdessen werden nach Aktivierung die Eingaben kontinuierlich ausgewertet und erkannte Befehle ausgeführt, bis die Eingabe explizit beendet wird. Da diese Funktionalität standardmäßig nicht von den verwendeten Sprachbibliotheken (Nuance SpeechKit, Google Speech Recognition) unterstützt wurde, wurden diese dafür erweitert.

### 5.1 Technische Realisierung der Plattform

Die multimodale Dialogplattform wurde als Web-App bzw. hybride App realisiert. Dafür wurde das MMIR-Framework eingesetzt [3], welches als Open-Source entwickelt wird<sup>4</sup> und plattformunabhängige Lösungen mittels Webstandards ermöglicht. MMIR-basierte Anwendungen können als native Apps auf den üblichen mobilen Plattformen eingesetzt werden. Funktionalitäten, die nicht bzw. nicht zufriedenstellend mittels HTML5 realisiert werden können, werden als Web-Services und als native Plugins

<sup>4</sup><https://github.com/mmig/mmir>

realisiert: im Wesentlichen betrifft dies die Spracherkennung und -synthese. Basierend auf Erkenntnissen von vorangegangenen Projekten wurde ein Anforderungsdokument für die Einbindung der Inhalte aus der Internetplattform in die MMIR-basierte App erstellt. Die Kommunikation der Module erfolgt über eine REST API. Dadurch ergibt sich die Unabhängigkeit und Robustheit der Plattform als Ganzes und somit die Nachhaltigkeit für eine spätere Verwertung der Projektergebnisse.

## 5.2 Informationsextraktion (IE)

Die IE-Module für die Erkennung der eingesprochenen Texte bestehen aus drei kaskadierten Modulen, wobei die Mächtigkeit in den Resultaten der einzelnen Module wächst (die Robustheit allerdings sinkt). Die erste Schicht ist rein wortbasiert und ist von daher die flachste und robusteste IE-Einheit im System. Hier werden die Vorkommen von Schlagworten und deren linguistischen Varianten in den Texten gesucht. Sie dienen als Entscheidungsgrundlage für die weiteren Schritte in der Plattform. In der zweiten Schicht benutzen wir das auf die wesentlichen Texte in den jeweiligen Anwendungsszenarien trainierte Klassifikationssystem PCL [5]. Dieses auf maschinellem Lernen basierende Modul kann Anfragen des Benutzers an das System auf den Workflow der Plattform abbilden und dient insofern zur Verfeinerung der Informationen aus der ersten Schicht. Die dritte Schicht bildet ein deterministischer Dependenz-Parser (MD-Parser) [7]. Mit Hilfe der Ergebnisse, den Dependenzstrukturen, des Parsers werden – sofern ein erfolgreiches Parsing durchgeführt werden konnte – definierte Templates befüllt, die im Zusammenspiel mit der Dialogplattform detaillierte Rückfragen bei fehlenden Informationen ermöglichen.

Das Ergebnis der IE-Module besteht i.d.R. aus einer Liste von mehreren möglichen Befehlen sortiert nach Erkennungskonfidenz. Die Liste potentieller Befehle wird dann von der Dialogplattform anhand der Konfidenzen und mittels kontextabhängigen Information ausgewertet, z.B. ob ein Befehl im aktuellen Applikationszustand ausgeführt werden kann oder ob bzw. wie groß eine inhaltliche Übereinstimmung vom aktuell angezeigten Applikationszustand mit dem (parametrisierten) Befehl ist. Erkannte Befehle werden „optimistisch“ bewertet: wenn die Liste mit hinreichender Konfidenz auf einen (eindeutigen) Befehl reduziert werden kann, so wird dieser ausgeführt. Das Dialogsystem gibt zugleich Rückmeldung dazu, dass und welcher Befehl ausgeführt wurde, sowie die Möglichkeit den Befehl wieder rückgängig zu machen. Die „optimistische“ Befehlsausführung reduziert Nachfragen seitens des Systems bei mehrdeutig erkannten Befehlen (Klärungsdialoge) auf ein Minimum. Des Weiteren hat das Dialogsystem einen Modus für geführte Dialoge bei Dateneingaben per Sprache (etwa zum Einstellen eines Hilfsgesuches): Nutzern wird Schritt-für-Schritt per Sprachausgabe gesagt, was sie tun müssen bzw. Rückmeldung darüber gegeben, wenn eine Eingabe fehlerhaft oder nicht vollständig war.

## 5.3 Spracherkennung

Für die Erprobung der Sprachinteraktion wurde ein Demonstrator (für Android, iOS und als Chrome Extension) sowie mehrere Module für die Einbindung unterschiedlicher Spracherkennung realisiert: Nuance SpeechKit (Android, iOS) und Google Speech Recognition (Android, Chrome Extension). Die weitere Interpretation der Spracherkennung erfolgte innerhalb der IE-Module. So wurden in den Erkennungseinheiten von Google bzw. Nuance die häufigsten verwechselten Begriffe innerhalb der Domäne bzw. Subsprache zum einen über manuelle Regeln in der Datenbank aufgelöst (erste Schicht). Zu anderen wurden für das Training des Klassifikationssystems in der zweiten Schicht die unveränderten Transkriptionen der Spracherkennung verwendet. Dadurch konnte die Robustheit und Akkuratheit der Klassifikation insgesamt verbessert werden. Für die im finalen System 40 zu erkennenden Sprechakte konnte eine Akkuratheit von über 74% erreicht werden, die Akkuratheit ohne die Transkriptionen lag nur bei 67%<sup>5</sup>

## 5.4 Modellierung emotionales TTS (Sprachausgabe)

Für die Modellierung des emotionalen TTS haben wir die Open-Source-Sprachsynthese-Plattform MARY TTS [6] mit einer eigens trainierten Stimme versehen. Wir haben dazu die Stimme des Leiters unserer Partnerorganisation „Leben im Kiez“ in Treptow-Köpenick in vier Aufnahmesessions aufgezeichnet. Es war möglich, alle nötigen Syllables in ausreichender Qualität zu synthetisieren. Bzgl. des Sprachrhythmus sind folgende Herausforderungen entstanden: Professionelle Sprecher tragen mit einem Grundtakt vor, vergleichbar mit Musikstücken. Dadurch haben die in einer Aufnahme geäußerten Worte und Silben jeweils die gleiche Länge und sind somit sehr gut tokenisierbar. Während der synthetischen Generierung passen dann die Einzelsilben sehr gut aufeinander und die Sprache wirkt dadurch sehr natürlich. Bei ungeübten Sprechern gibt es diesen Takt nicht und somit braucht man ungleich mehr gesprochenen Text, um ähnliche Qualitäten zu erzielen. Oftmals wirken auch Worte unnatürlich gedehnt oder die Sprachrhythmik passt nicht optimal zu den gesprochenen Sätzen. Im Resultat haben wir innerhalb der Plattform nur an einigen Stellen auf diese synthetisierte Stimme, an den meisten Stellen jedoch auf professionelle Sprachpakete zurückgegriffen.

## 5.5 Benutzeradaption über Machine Learning

Über eine Machine Learning-Komponente wurde das Verhalten des Systems benutzeradaptiv gestaltet. Nach initialen Versuchen und Arbeit mit den Fokusgruppen, sowie aus Erfahrungen, die Microsoft mit benutzeradaptiven Benutzerschnittstellen in den Office Paketen 2007 berichtet<sup>6</sup>, haben wir bewusst darauf verzichtet, sichtbare Objekte aufgrund des Nutzerverhaltens zu verändern. Stattdessen haben wir die Adaption auf die Nachrichten angewendet, die

<sup>5</sup>Gemessen per 10-fold validation auf transkribierten Testmengen

<sup>6</sup><https://vimeo.com/3305642>

der Benutzer per Push-Dienst zugestellt bekommt. Hintergrund hierfür ist, dass diese Push-Meldungen mit einer hohen Präzision gefiltert werden sollten. Benutzer, die zum Beispiel selbst nach Hilfen für das Aufhängen von Gardinen suchen, sollen möglichst keine Anfragen dieser Art als Push-Meldung erhalten. Auf Basis der Ergebnisse aus den Fokusgruppen haben wir die identifizierten Hilfsangebote und -gesuche nach Ausschließlichkeit unterteilt: Hilfsangebote werden dann als ausschließlich klassifiziert, wenn der Suchende eine gleiche bzw. ähnliche Hilfe nicht leisten könnte. Beispiele für ausschließliche Gesuche sind: Aufhängen von Gardinen, Reparaturen, Tragen von schweren Gegenständen, usw. Beispiele für nicht-ausschließliche Aktionen sind: Einkaufen, Begleitung bei verschiedenen Aktivitäten, Besuch von Veranstaltungen, etc. Über Trainingscorpora mit Gesuchen, die in die jeweiligen Unterteilungen passen, wurde ein Dokumentenklassifikationsalgorithmus auf Basis des maschinellen Lernverfahren PCL trainiert. Einkommende Nachrichten werden gemäß obiger Unterteilung klassifiziert und je nach Ergebnis dem Benutzer angezeigt oder nicht.

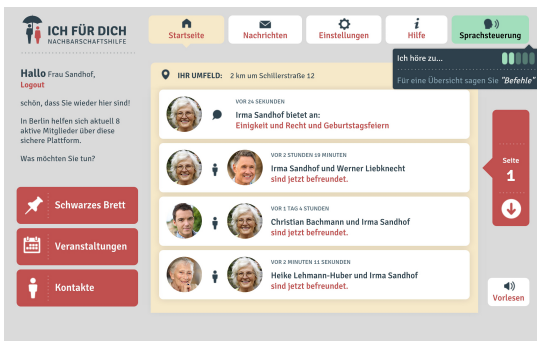


Abbildung 4 Startseite

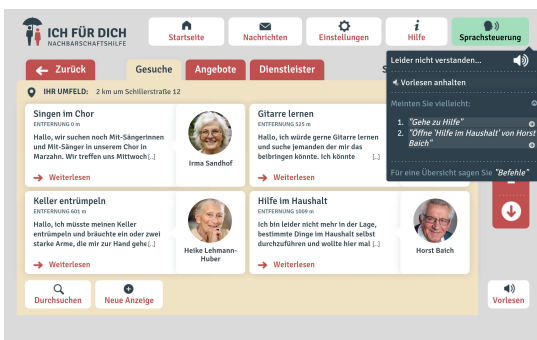


Abbildung 5 Detailansicht Kontaktdaten

## 5.6 Feldtest

Vom 15. September bis 8. Oktober 2015 fand im Rathaus St. Johann in Saarbrücken (Hauberrisser Saal) die interaktive Ausstellung „Zukunftsstadt“ statt. Dort wurde an ausgewählten Exponaten gezeigt, wie nachhaltige Stadtentwicklung funktioniert, wie man in Zukunft seine Wohnung für sich arbeiten lässt oder wie sich Nachbarschaften besser miteinander vernetzen können. Auf der Ausstellung konnten alle Besucher einen Prototyp der „Wir im Kiez“-App

testen. Das Feedback wurde gesammelt und floss in die Entwicklung mit ein. Die Anwendung erhielt eine durchweg positive Resonanz. Bei der Bedienung des Prototyps wurde oft beobachtet, dass er intuitiv bedient werden konnte, ohne dass die Anwendung zuvor erklärt werden musste. Abbildung 4 und Abbildung 5 zeigen exemplarische Screenshots der App.

## 6 Zusammenfassung

Mit dem Projekt „Wir im Kiez“ konnten wir erfolgreich eine Plattform für den Austausch zwischen Nachbarn erforschen und realisieren. Der maßgebliche Erfolgsfaktor war dabei die konsequente Einbeziehung von Benutzern aus der intendierten Zielgruppe bei der Auswahl der Funktionalitäten und bei dem Interaktionsdesign. Mit dem Einsatz aktueller Methoden der Künstlichen Intelligenz wie Sprachinteraktion, Informationsextraktion und maschinellem Lernen, können auch Benutzer mit der Plattform interagieren, die keine oder nur wenig Affinität zu technischen Geräten haben.

## 7 Literatur

- [1] John, M.; Klose, S.; Kock, G.; Jendreck, M.; Feichtinger, R.; Hennig, B.; Reithinger, N.; Kiselev, J.; Gövercin, M.; Kausch, S.; Polak, M.; Irmscher, B.: Smart-Senior's Interactive Trainer – Development of an interactive system for a home-based fall-prevention training for elderly people. In: Ambient Assisted Living – Advanced Technologies and Societal Change, Vol. 7, Pages 305-316, Springer 2012.
- [2] Hassenzahl, M.; Burmester, M.; Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität, Interaktion in Bewegung (pp. 187-196). Stuttgart, Leipzig: B. G. Teubner, 2003.
- [3] Russ, A.: MMIR Framework: Multimodal Mobile Interaction and Rendering, GI-Jahrestagung 2702-2713, 2013
- [4] Sven Schmeier: Ein mobiles Übersetzungssystem für den Einsatz bei Notfällen im Kontext von AAL 5. Deutscher AAL-Kongress mit Ausstellung, 24. - 25. Januar 2012, Berlin, Tagungsbeiträge, Berlin, Germany, VDE, 2012
- [5] Sven Schmeier: Exploratory Search on Mobile Devices, PhD-Thesis, 2014
- [6] Marc Schröder, J. Trouvain: The German Text-to-Speech Synthesis System MARY: A Tool for Research, Development and Teaching, International Journal of Speech Technology volume 6, Pages 365-377, 2003
- [7] Alexander Volokh, Günter Neumann: Task-oriented Dependency Parsing Evaluation Methodology, IEEE 13th International Conference on Information Reuse and Integration, Las Vegas, NV, USA, IEEE Systems, Man, and Cybernetics Society (SMC), 2012