


MAS-HCID Master-Thesis MT2023-24

OST – Ostschweizer Fachhochschule

Interaktion in natürlicher Sprache



Projektteam

Dennis Eitner

Sandro Pezzutto

Coach

Prof. Dr. Markus Stolze

Co-Bewerter

Bernhard von Allmen

Erklärung der Selbstständigkeit

Hiermit bestätigen wir,

- dass wir die vorliegende Arbeit selbst und ohne fremde Hilfe durchgeführt haben, ausser derjenigen, welche explizit beschrieben ist,
- dass wir sämtliche verwendeten Quellen erwähnt und gemäss gängigen wissenschaftlichen Regeln korrekt zitiert haben,
- dass wir keine durch Copyright geschützten Materialien (z. B. Bilder) in dieser Arbeit in unerlaubter Weise genutzt haben,
- dass wir in dieser Arbeit keine Adressen, Telefonnummern und andere persönliche Daten von Personen, die nicht zum Kernteam gehören, publizieren.

Eitner, Dennis

Zürich, 26.01.24

Ort, Datum, Unterschrift

Pezzutto, Sandro

Richterswil, 26.01.24

Ort, Datum, Unterschrift

Abstract

Im Laufe dieser Masterarbeit wird ein codierter Prototyp für einen persönlichen intelligenten Sprachassistenten entwickelt, der insbesondere für ältere, in einem Pflegeheim lebende Menschen gedacht ist. Anhand der Entwicklung wird ein möglichst umfassendes Verständnis der Interaktion zwischen Menschen und Computern mittels natürlicher Sprache angestrebt. Dazu werden zwischenmenschliche Dialoge erhoben und ausgewertet, Dialogstrukturen analysiert und die Regeln des Turn-Taking untersucht. Diese Aspekte sind zentrale Bestandteile in der Entwicklung und Nutzung von sprachbasierten Assistenten.

Da zu Beginn der Arbeit kein bestimmter Anwendungsfall vorhanden war, mussten ein sinnvoller Einsatz und eine Gruppe von Nutzenden gesucht und gefunden werden. Das Vorgehensmodell LeanUX wurde angewendet, um in kurzer Zeit Anwendungsfälle auf ihren Nutzen überprüfen zu können. Ausgehend von technologischen Möglichkeiten soll flexibel auf Bedürfnisse eingegangen und die Grenzen der aktuellen Technologien sollen aufgezeigt werden.

Diese Arbeit befasst sich ebenfalls mit dem Prozess der Produktentwicklung, in dem künstliche Intelligenz Aufgaben des Menschen übernimmt. Es wird beschrieben, weshalb bekannte Vorgehensmodelle und Artefakte wie Persona, Journey und Dialogmodelle wenig zielführend sind. Um die aufgeworfenen Forschungsfragen beantworten zu können, werden Experimente durchgeführt, in welchen Menschen die Aufgabe eines Systems übernehmen. Über fünf Zyklen hinweg wird der Mensch sukzessive durch ein System ersetzt, um am Ende einen codierten Prototyp unter realen Bedingungen mit Nutzenden auf seine Nutzbarkeit und Nützlichkeit zu überprüfen. Dessen Entwicklung dient als praktisches Beispiel und soll es ermöglichen, die theoretischen Erkenntnisse in die Praxis umzusetzen und zu testen. Die Entwicklung und die Analyse des Prototyps bieten wertvolle Einblicke in die Herausforderungen und Potenziale der Sprachinteraktionstechnologie. Dabei werden auch ethische Aspekte in der Interaktion von Menschen mit künstlicher Intelligenz untersucht.

Zusammenfassend soll mit dieser Arbeit ein elementarer Beitrag zum Verständnis der menschenzentrierten Entwicklung von Sprachinteraktionssystemen geleistet werden. Es werden detaillierte Analysen der Dialogstrukturen und Interaktionsmechanismen sowie praktische Ansätze für die Gestaltung nutzungsfreundlicher und ethisch verantwortungsvoller Sprachassistenten aufgezeigt. Die Ergebnisse dieser Arbeit sind daher von grosser Relevanz für die Forschung und Entwicklung im Bereich der Mensch-Computer-Interaktion im Umgang mit künstlicher Intelligenz.

I. Inhaltsverzeichnis

I. Inhaltsverzeichnis	IV
II. Abkürzungsverzeichnis	X
III. Abbildungsverzeichnis	XI
IV. Tabellenverzeichnis	XIII
V. Glossar	XIV
1. Auftrag und Aufgabenstellung	1
1.1. Thema: natürlichsprachliche Interaktion im E-Commerce	1
1.2. Auftraggeber	1
1.3. Aufgabenstellung	1
1.4. Zielsetzung	1
1.4.1. S.M.A.R.T-Ziele	1
1.4.2. Forschungsfrage	2
1.4.3. Persönliche Lernziele von Dennis Eitner	2
1.4.4. Persönliche Lernziele von Sandro Pezzutto	3
2. Projektplanung	4
2.1. Einleitung	4
2.2. Vorgehensmodell	4
2.3. Begründung der Wahl	6
2.4. Projektplan	7
2.5. Projektteam und Organisation	8
2.6. Risikoanalyse I.	9
2.7. Stakeholderanalyse	10
3. Theoretischer Rahmen	11
3.1. Einleitung	11
3.2. Natürliche Sprache beim Menschen	11
3.2.1. Sender-Empfänger-Modell	11

3.2.2. Grice'sche Konversationsmaximen	12
3.2.3. Sprecherwechsel	12
3.3. Natürliche Sprache im Umgang mit Computern	13
3.3.1. Conversational AI	13
3.3.2. Natural Language Processing	13
3.3.3. Natural Language Understanding	14
3.3.4. Natural Language Generation	14
3.3.5. Intelligenter persönlicher Assistent	15
3.4. Konversationsdesign.....	15
3.4.1. Reparaturmechanismen.....	16
3.5. Uncanny Valley	16
3.6. Eliza-Effekt.....	16
4. Zyklus 0: Evaluation von Anwendungsfällen	18
4.1. Einleitung	18
4.2. Recherche	18
4.2.1. Desk-Research	18
4.2.2. Interview mit einer Fachperson	18
4.3. Dialogstrukturen in dialogorientierten Bedienoberflächen	19
4.4. Annahmen.....	20
4.4.1. Problem-Statement.....	20
4.4.2. Persona.....	20
4.4.3. Hypothese	20
4.5. Design.....	21
4.6. Minimum Viable Product	21
4.7. Feedback.....	22
4.7.1. Quantitative Beurteilung.....	22
4.7.2. Mikrointeraktionen	23
4.7.3. Einschätzung der Assistenten	23
4.7.4. Feedback der Anwendenden	24
4.7.5. Fazit	24

4.8. Reflexion	26
5. Risikoanalyse II.....	27
6. Zyklus 1: Quiz.....	28
6.1. Einleitung	28
6.2. Stakeholderanalyse	28
6.3. Recherche	28
6.3.1. Definition und begriffliche Abgrenzung des Ausdrucks <ältere Menschen>	28
6.3.2. Interviews mit Fachpersonen	28
6.3.3. Feldstudie.....	29
6.4. Annahmen.....	30
6.4.1. Problem-Statement.....	30
6.4.2. Persona.....	30
6.4.3. Hypothese	30
6.5. Design.....	31
6.6. Minimum Viable Product	32
6.7. Feedback.....	32
6.8. Reflexion	32
7. Zyklus 2: Kennenlernen	34
7.1. Einleitung	34
7.2. Recherche	34
7.2.1. Benchmark: Inflection.ai	34
7.3. Annahmen.....	35
7.3.1. Problem-Statement.....	35
7.3.2. Persona.....	35
7.3.3. Hypothese	35
7.4. Design.....	36
7.5. Minimum Viable Product	36
7.6. Feedback.....	36
7.6.1. Dialoganalyse.....	36

7.6.2. Aktives Zuhören	39
7.6.3. Nonverbale Kommunikation in Dialogen.....	40
7.7. Reflexion	41
8. Zyklus 3: Persönlichkeit.....	42
8.1. Einleitung	42
8.2. Recherche	42
8.3. Annahmen.....	42
8.3.1. Problem-Statement.....	42
8.3.2. Persona.....	43
8.3.3. Hypothese	43
8.4. Design.....	43
8.4.1. Charakterisierung System Persona	44
8.5. Minimum Viable Product	44
8.6. Reflexion	45
9. Risikoanalyse III.	46
9.1. Planung und Vorgehen der studienübergreifenden Zusammenarbeit	46
9.2. Reflexion	47
10. Technischer Exkurs	48
10.1. Systemzustände	48
10.2. Zustands- und Aktivitätsdiagramm	49
11. Zyklus 4: Mikrointeraktionen	51
11.1. Einleitung	51
11.2. Recherche	51
11.3. Annahmen.....	52
11.3.1. Problem-Statement.....	52
11.3.2. Persona.....	52
11.3.3. Hypothese	52
11.4. Design.....	52

11.4.1. Visuelles Design	52
11.4.2. Earcons.....	53
11.4.3. Visuelles Interface	54
11.5. Minimum Viable Product	54
11.6. Feedback.....	55
11.7. Reflexion	56
12. Risikoanalyse IV.	57
13. Exkurs: Prompt-Engineering.....	58
14. Zyklus 5: Stimme	59
14.1. Einleitung	59
14.2. Recherche	59
14.3. Annahmen.....	61
14.3.1. Problem-Statement.....	61
14.3.2. Persona.....	61
14.3.3. Hypothese	61
14.4. Design.....	61
14.5. Minimum Viable Product	62
14.6. Feedback.....	63
14.7. Reflexion	64
15. Zyklus 6: Finale.....	65
15.1. Einleitung	65
15.2. Recherche	65
15.3. Annahmen.....	65
15.3.1. Problem-Statement.....	65
15.3.2. Persona.....	66
15.3.3. Hypothese	66
15.4. Design.....	67
15.5. Minimum Viable Product	68

15.6. Feedback.....	68
16. Empfehlungen an den Auftraggeber	69
16.1. Bekannte Probleme.....	69
16.2. Weitere Hypothesen	70
16.2.1. Avatar.....	70
16.2.2. Konfigurierbare Turn-Taking-Regeln.....	71
16.2.3. Dialogaufnahme.....	71
16.2.4. Variable Redezeit und ständige Zuhörbereitschaft	71
16.2.5. Hardware.....	73
17. Reflexion	74
17.1. Projektplanung und Verlauf	74
17.2. Lean UX	74
17.2.1. Conversation-Design	75
17.3. Das Eizzutto-Modell zur Entwicklung von Conversational Interfaces	76
17.3.1. Herleitung	76
17.3.2. Das Eizzutto-Modell	79
17.4. Ethik	80
17.4.1. Inklusives Design.....	80
17.5. Lernreflexion	81
17.5.1. Lernreflexion von Dennis Eitner	81
17.5.2. Lernreflexion von Sandro Pezzutto	81
17.6. Leistungsreflexion	82
17.7. Teamreflexion.....	82
18. Fazit	83
19. Danksagung.....	84
20. Literaturverzeichnis.....	85
21. Anhang	89

II. Abkürzungsverzeichnis

AI	Artificial Intelligence
GUI	Graphical User Interface
IVR.....	Interactive Voice Response
KI	Künstliche Intelligenz, Künstliche Intelligenz
LLM	Large Language Model
MVP.....	Minimum Viable Product
NLP	Natural Language Processing
S.M.A.R.T.....	spezifisch, messbar, ausführbar, realistisch und terminiert
SMS.....	Short Message Service
STT	Speech to Text
UCD	User Centered Design
UML.....	Unified Modeling Language
VUI	Voice User Interface
WLAN.....	Wireless Local Area Network
WOZ	Wizard of Oz

III. Abbildungsverzeichnis

Abbildung 1: Vorgehensmodell Lean UX für Voice-User-Interfaces	4
Abbildung 2: Detaillierungsgrad/Fidelity	6
Abbildung 3: Projektplan	7
Abbildung 4: Sender-Empfänger-Modell nach Shannon-Weaver	11
Abbildung 5: Uncanny Valley	16
Abbildung 6: Persona Zyklus 0.....	20
Abbildung 7: Siri von Apple.....	21
Abbildung 8: Google Assitant	21
Abbildung 9: Bing von Microsoft.....	21
Abbildung 10: Einkaufsberatung	28
Abbildung 11: Quiz.....	28
Abbildung 12: Vorlesen im Auto.....	28
Abbildung 13: Reiseführer/ in	28
Abbildung 14: Persona des Zyklus 1	30
Abbildung 15: Dialogflow Quiz	31
Abbildung 16: Persona Zyklus 2.....	35
Abbildung 17: Spickzettel Kennenlernfragen.....	36
Abbildung 18: Dialoganalyse	38
Abbildung 19: Nonverbale Kommunikation von PI.ai	40
Abbildung 20: Persona Zyklus 3.....	43
Abbildung 21: Farbkonzept	43
Abbildung 22: Zustands- und Aktivitätsdiagramm.....	49
Abbildung 23: Spracherkennung und Übertragung	50
Abbildung 24: Zustände von PI.ai	51

Abbildung 25: Sketch ‹Pingpong›	52
Abbildung 26: Version ‹Chatverlauf›	53
Abbildung 27: Version ‹Text›.....	53
Abbildung 28: Version ‹Amplitude›	53
Abbildung 29: Visuelles Interface	54
Abbildung 30: Sprintplanung mit zeitversetzten Sprints	57
Abbildung 31: Verlauf des Stimmenturniers mit Testperson A.	62
Abbildung 32: Persona Zyklus 6.....	66
Abbildung 33: Dialogablauf und Agents	67
Abbildung 34: Zustand ‹Schlafen›	70
Abbildung 35: Zustand ‹Sprechen›	70
Abbildung 36: Zustand ‹Zuhören›	70
Abbildung 37: Zustand ‹Denken›	70
Abbildung 38: Konfiguration ‹Sprecherwechsel›	71
Abbildung 39: System-Journey.....	71
Abbildung 40: Variable Redezeit.....	72
Abbildung 41: Ständige Zuhörbereitschaft	72
Abbildung 42: Projektverlauf	74
Abbildung 43: Baxley-Model für Filme, eigene Darstellung	76
Abbildung 44: Why-What-How	76
Abbildung 45: Proto-Vorgehensmodell Version 1	77
Abbildung 46: Proto-Vorgehensmodell Version 3	77
Abbildung 47: Impact-Effort-Matrix	78
Abbildung 48: Eizzutto-Modell	79

IV. Tabellenverzeichnis

Tabelle 1: Teamrollen	8
Tabelle 2: Dialogstrukturen in dialogorientierten Bedienoberflächen.....	19
Tabelle 3: Übersicht der Anwendungsfälle, Aufgaben, Antworten.....	22
Tabelle 4: Kriterien und Erfolgsraten	23
Tabelle 5: Kategorisierte Fragetypen.....	38
Tabelle 6: Voice-User-Interfaces in Filmen	60
Tabelle 7: Rangliste der Stimmen.....	63

V. Glossar

Agent	In der Mensch-Maschine-Interaktion das antwortgebende System, z. B. ein Chatbot oder digitaler Assistent.
Barge-in	Proaktives Unterbrechen des Sprachassistenten durch die Nutzenden.
Earcon	Nonverbale Audiosymbole, die eine bestimmte Bedeutung vermitteln sollen.
ElevenLabs	Dienstleistungsanbieter für generative Sprachsynthese.
Intent	Erkannte Nutzungsabsicht, die der Anwendungslogik entsprechend zugeordnet wird.
IVR	Interaktives Sprachdialogsystem, das automatisierte Dialoge mittels akustischer Medien ermöglicht.
FLOP	Masseinheit für die Leistung von Computern in Bereichen, die Gleitkommaberechnungen erfordern.
Hawthorne-Effekt	Verhaltensänderung von Studienteilnehmenden durch Beobachtung.
Homonym	Ein Wort, das für verschiedene Begrifflichkeiten verwendet wird.
Homophon	Ein Wort, das gleich ausgesprochen wird, aber verschiedene Bedeutungen hat.
Sprachmodell	Mathematische Modellierung der Abfolge von Sprachelementen in Texten.
System Persona	Ein durch künstliche Intelligenz wahrgenommener Charakter, welcher Aufgaben und Rollen in der Interaktion mit Menschen ausführt.

Multimodal	Ausgabe von zusätzlichen, meist visuellen Informationen.
NLP	Untersuchung der natürlichen Sprache durch Computer.
NLG	Automatische Produktion von natürlicher Sprache durch eine Maschine.
NLU	Das Verständnis und die Analyse von natürlicher Sprache.
Prompt	Eingabeaufforderung, die in der generativen KI eine Benutzereingabe darstellt.
STT	Umwandlung von gesprochenem Wort in Text.
TTS	Künstliche Erzeugung von menschlichen Stimmen aus Textvorgaben.
Turn-Taking	Abwechselndes Sprechen der Gesprächsteilnehmer/-innen.
VUI	Benutzerschnittstelle, die auf Sprachsteuerung basiert.
Wakeword	Gesprochene Äußerung, die das IVR in den <Zuhören>-Modus versetzt.

1. Auftrag und Aufgabenstellung

1.1. Thema: natürlichsprachliche Interaktion im E-Commerce

Wie wird das Thema ‹Natural Language Understanding› in Bezug auf ‹Voice-Commerce› angewendet? Dabei ist vor allem die Frage von Interesse, wie ein Mensch in deutscher Sprache Produkte online bestellen kann, ohne andere Interaktionsmethoden zu Hilfe nehmen zu müssen.

1.2. Auftraggeber

Fachhochschule Ostschweiz OST
Prof. Dr. Markus Stolze
Oberseestrasse 10, 8640 Rapperswil

1.3. Aufgabenstellung

In dieser Masterarbeit wird die Anwendung von *Natural Language Understanding* (NLU) im Kontext des Voice-Commerce untersucht. Das Ziel ist die Konzeption einer auf Dialogen basierten interaktiven Anwendung für den Online-Einkauf, die primär über Voice-Interaktion funktioniert. Die Arbeit beinhaltet die Dokumentation von Mensch-zu-Mensch-Interaktionen, die Erhebung natürlichsprachlicher Dialoge und die Durchführung von Experimenten mit Mensch-Computer-Interaktionen. Es werden Muster in den Dialogen, Formulierungen und Dialogkomponenten analysiert. Es wird herausgestellt, wie nahe ein AI-gestütztes System an reale zwischenmenschliche Dialoge heranreicht und wie fehlertolerant Interaktionen mit Menschen sein dürfen. Zudem wird der Mehrwert der visuellen Darstellung in der Voice-Interaction untersucht. Die zu erarbeitenden Artefakte können Personas, User-Journeys oder User-Szenarios, ein Modell der erhobenen Dialogmuster und ein Interaktionskonzept für ein natürlichsprachliches Interface sein. Die Interaktionssprache ist Hochdeutsch, und es soll auch ohne grammatikalische Korrektheit möglich sein, eine Konversation zu führen.

1.4. Zielsetzung

1.4.1. S.M.A.R.T-Ziele

Die Zielsetzung dieser Masterarbeit wird gemäss der S.M.A.R.T.-Methode in Projekt-, Lern-, Forschungs-, Produkt- und Endziele untergliedert (siehe Anhang 2). Das von dem Auftraggeber definierte Forschungsziel besteht in der Erhebung und Sammlung von Transkriptionen von Dialogen, die im Rahmen der Interaktion mit natürlicher Sprache entstehen. Zu diesem Zweck werden Experimente in einem speziell dafür eingerichteten Kontext durchgeführt.

Zur Schaffung dieses Kontextes ist die Entwicklung eines noch zu spezifizierenden Produktes geplant. Diese Entwicklung erfordert die Festlegung konkreter Anwendungsfälle sowie die Definition einer spezifischen Gruppe an Nutzenden. Ein Teil der Arbeit wird daher der Frage gewidmet sein, wie ein geeigneter Kontext etabliert werden kann.

Die Datenerhebung erfolgt mit verschiedenen Methoden. Um Dialoge im vorgesehenen Nutzungskontext zu erfassen, werden Table-Reads, Hallway-Tests, Wizard-of-Oz (WOZ)-Tests sowie statische und codierte Prototypen angewendet und die Dialoge aufgezeichnet (siehe nicht öffentlicher Anhang).

Das selbst auferlegte Produktziel besteht in der Entwicklung eines codierten Prototyps, der auf Smartphones nutzbar und in deutscher Sprache verfügbar ist. Der eigene Anspruch in der Produktentwicklung liegt darin, sowohl die Anwendbarkeit (Usability) als auch den Nutzen (Utility) des Produktes nachzuweisen, wobei beide Kriterien möglichst umfassend erfüllt werden sollen.

1.4.2. Forschungsfrage

Abgeleitet aus der Aufgabenstellung und dem Forschungsziel wird folgende übergeordnete Forschungsfrage formuliert, die als Orientierung für den gesamten Projektverlauf dient:

Wie müssen natürlichsprachliche Dialoge zwischen einem Menschen und einem System stattfinden, damit ein Voice-User-Interface als Alternative zu einem visuellen User Interface angenommen wird?

1.4.3. Persönliche Lernziele von Dennis Eitner

Meine persönlichen Lernziele umfassen folgende Fragen: Wie wird ein Produkt entwickelt, wenn das zugrunde liegende Problem bisher nicht klar definiert ist? Welche Methoden führen zur Identifizierung passender Anwendungsfälle für eine existierende, aber sich in fortlaufender Entwicklung befindender Technologie? Wie können heutige und zukünftige technologische Grenzen ermittelt werden?

Ein besonderer Fokus liegt auf dem Gebiet der Voice-User-Interfaces (VUI). Dabei handelt es sich um ein Thema, das im bisherigen Unterricht nicht besprochen wurde. Es gilt herauszufinden, wie Unsicherheiten bewältigt werden können, welche zwangsläufig entstehen, wenn ein Produkt in einem sich entwickelnden Technologiebereich konzipiert wird. Dies ist insbesondere deswegen der Fall, weil die bisher bekannten Vorgehensmodelle und Forschungsmethoden auf grafische Benutzeroberflächen ausgerichtet sind.

1.4.4. Persönliche Lernziele von Sandro Pezzutto

Durch das Aufkommen von konversationsfähigen Chatbots wie ChatGPT bin ich der Meinung, dass natürlichsprachliche Interfaces an Relevanz gewinnen. Die Interaktion damit erfolgt aktuell jedoch vornehmlich in schriftlicher Form. In Bezug auf diese Arbeit möchte ich folgende Fragen stellen: Lässt sich diese Technologie auch für ein VUI nutzen? Welcher Personengruppe würde eine solche Technologie besonders nützen und welche Anwendungsfälle gibt es dafür?

Die Themenbereiche der natürlichsprachlichen Interaktion und der VUI wurden im Unterricht nicht explizit behandelt. Daher ist ein bedeutsames Lernziel, bekannte User-Centered-Design (UCD)-Methoden mit spezifischem Wissen aus dem Bereich dieser beiden Themen zu ergänzen und in einem Vorgehensmodell anzuwenden. Mein Fokus liegt dabei in der Anwendung und in der Vertiefung von explorativen UCD-Methoden. Im Speziellen bezieht sich dies auf den Bereich des Prototyping.

2. Projektplanung

2.1. Einleitung

Diese Masterarbeit erstreckt sich über einen Zeitraum von neun Monaten. Um das Erreichen der Zielsetzungen zu gewährleisten, bedarf es einer effektiven Projektplanung. Es wird ein Vorgehensmodell eruiert, wovon ausgehend die Projektplanung abgeleitet wird. Zusätzlich werden Teamrollen und Aufgaben festgelegt sowie die Risiko- und Stakeholderanalyse erörtert.

2.2. Vorgehensmodell

Lean UX (Jeff Gothelf, 2021) ist ein kontinuierlicher, sich stetig wiederholender Prozess. Er besteht aus folgenden vier Phasen: Hypothesenbildung, Design, Minimum Viable Product (MVP) und Recherche bzw. Feedback. Am Anfang steht die Problembeschreibung (engl. *problem statement*), die eine bestimmte Gruppe von Anwendenden, genannt Persona, hat. Die Persona wird als Annahme betrachtet und in die Hypothesenbildung einbezogen. Mit fortschreitenden Zyklen wird sie kontinuierlich validiert und justiert. Die Hypothese wird in den Phasen <Design> und <MVP> überprüft. Die gewonnenen Erkenntnisse können im Fall eines nicht erfolgreichen Feedbacks als Recherchegrundlagen im folgenden Zyklus dienen.

Abbildung 1: Vorgehensmodell Lean UX für Voice-User-Interfaces

Recherche	Annahmen			Design	MVP	Feedback
Feldstudien	Problem statement	Annahmen	Hypothese	Einschränkungen	Low Fidelity	Feldstudien
Experten Interviews	Neues Produkt: Der [Marktbereich] konzentriert sich vornehmlich auf [Kundensegment, Problemfälle, Lösungen, ...]. Für [diese Marktlücke] gibt es keine (oder keine guten, keine stabilen) Produkte oder Dienste. [Das Produkt] schließt diese Lücke durch [diese Vision/Strategie]. Wir werden uns primär auf [dieses Marktsegment] konzentrieren.	Geschäftliche Ergebnisse	Wir sind der Auffassung, dass wir [folgendes Geschäftsergebnis] erreichen, wenn [dieser User] mit [diesem Feature] [dieses Ziel] erreicht.	Individuelle Ideenfindung	Sample Dialoge ▶ Halfway Tests	Experten Interviews
Contextual Inquiries		Erfolgskriterien Definition of Done		Anwender / Personas	Präsentation	Mid and High Fidelity
Desk Research	Bestehende Produkte: [Das Produkt] wurde für [diese Zielsetzung] entwickelt. Wir haben jedoch festgestellt, dass [das Produkt] [diese Zielsetzung] nicht erfüllt, was sich in Form [dieses nachteiligen Effekts] auf unsere Geschäftstätigkeit auswirkt. Wie können wir [das Produkt] dahin gehend verbessern, dass [unsere Benutzer] aufgrund [dieser messbaren Kriterien] erfolgreicher damit arbeiten können?	Ergebnisse und Vorteile für Anwender		Iteration und Optimierung	Prototypes ▶ Wizard of OZ	Desk Research
Competitor Review		Ziele des Marktes** Entscheid. Erhebungen** langfristige Ziele**		Ideenfindung	Coded Prototypes	Competitor Review
Thematische Analyse*				System Persona ***	Voice APP ▶ Usability Test	Thematische Analyse*
				Dialoge ***		
				Table Reads ***		

▲ ▼ Mikroprozess * Eigene Erweiterung ** Erweiterung nach Lean UX 2022 (3. Auflage) *** Eigene Erweiterung für VUI

Die Formulierung von Annahmen und Hypothesen in diesen Bericht folgt den Vorlagen von Lean UX. Das *problem statement* für ein neues Produkt wird folgendermassen formuliert:

Der [Marktbereich] konzentriert sich vornehmlich auf [Kundesegment, Problemfälle, Lösungen, ...]. Für [diese Marktlücke] gibt es keine (oder keine guten, keine etablierten) Produkte oder Dienste. [Das Produkt] schließt diese Lücke durch [diese Vision/Strategie]. Wir werden uns zunächst auf [dieses Marktsegment] konzentrieren.

Für die Problemstellung eines bestehenden Produkts wird diese Vorlage verwendet:

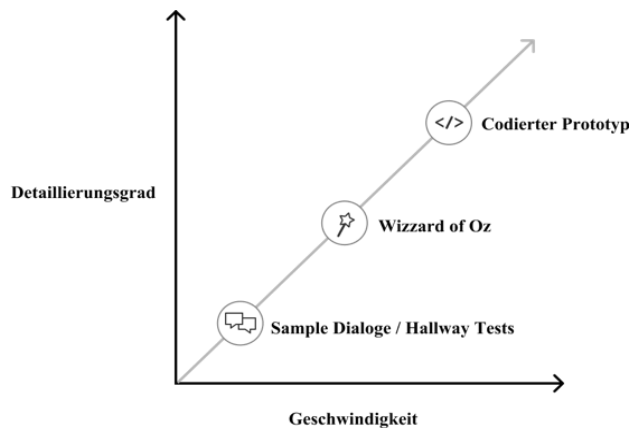
[Das Produkt] wurde für [diese Zielsetzung] entworfen. Wir haben jedoch festgestellt, dass [das Produkt] [diese Zielsetzung] nicht erfüllt, was sich in Form [dieses nachteiligen Effekts] auf unsere Geschäftstätigkeit auswirkt. Wie können wir [das Produkt] dahin gehend verbessern, dass [unsere Nutzenden] aufgrund [dieser messbaren Kriterien] erfolgreicher damit arbeiten können?

Die Hypothesen werden nach der folgenden Vorlage formuliert:

Wir sind der Auffassung, dass wir [folgendes Geschäftsergebnis] erreichen, wenn [dieser User] mit [diesem Feature] [dieses Ziel] erreicht.

2.3. Begründung der Wahl

Abbildung 2: Detaillierungsgrad/Fidelity



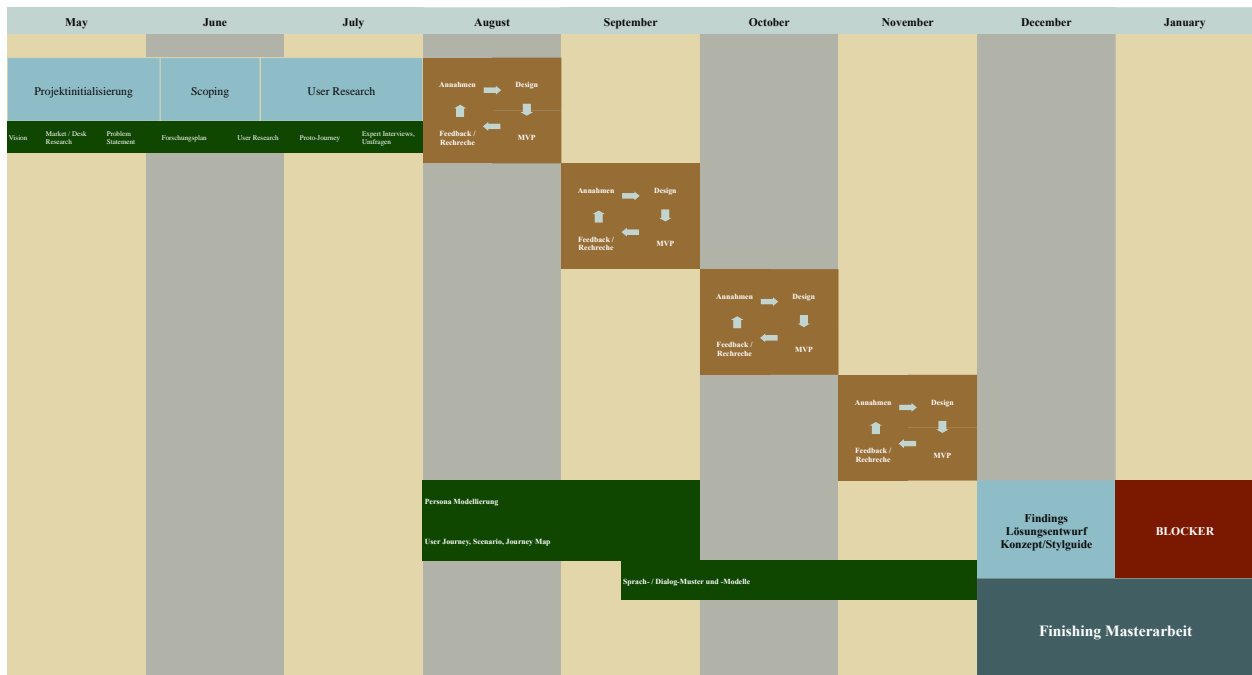
Lean UX ist annahmen- und hypothesenbasiert und wird aufgrund seines explorativen Charakters gewählt. Das Projektteam sieht sich am Anfang eines Produktentwicklungszyklus, in welchem das adressierte Problem noch unklar ist. Ein Teil dieser Arbeit wird dafür aufgewendet, das Problem zu formulieren, aus dessen Lösung sich für Nutzende ein Mehrwert ergibt.

Zu Beginn des Projektes fehlen konkrete Anwendungsfälle. Die Lean-UX-Zyklen werden dazu verwendet, um Anwendungsfälle in kurzer Zeit auf ihre Nützlichkeit zu überprüfen (Brown, o. D.). Dabei wird ein Produkt nur genau so weit entwickelt, wie die Annahme überprüft werden kann. Forschungsmethoden und Stichprobe werden für jeden Zyklus zweckdienlich definiert. Die erstellten MVPs werden mit Hallway-Tests, Wizard-of-Oz-Experimenten oder Usability-Tests geprüft. In den folgenden Zyklen wird der Detaillierungsgrad (engl. *fidelity*) kontinuierlich erhöht, sofern der vorangegangene erfolgreich war. Zu den Designartefakten in diesem Projekt zählen unter anderem (Proto-) Dialoge und eine *System Persona*.

2.4. Projektplan

Das Projekt sieht eine Phase von drei Monaten vor, in welcher theoretisches Wissen um den aktuellen Stand der Forschung und des Marktes erarbeitet wird. Nach Ablauf dieser Phase wird der Lean-UX-Zyklus viermal durchlaufen (Abbildung 3). Das Projektziel ist, einen codierten Prototyp zu entwickeln, welcher in einem finalen Usability-Test in dem dafür vorgesehenen Kontext getestet wird.

Abbildung 3: Projektplan



2.5. Projektteam und Organisation

Tabelle 1: Teamrollen

Name	Rolle im Team
Dennis Eitner	Generator
Sandro Pezzutto	Synthesizer
Abinas Kuganathan*	Softwareentwickler

**Mitglied des Projektteams im Rahmen einer studienübergreifenden Zusammenarbeit. Kein Autor dieser Arbeit.*

Alle Teammitglieder begeben sich mit dieser Arbeit auf thematisches Neuland. Daher verzichtet das Team auf eine strikte Aufgabenaufteilung im Projekt. Stattdessen werden Rollen nach Cooper, Reimann, Cronin und Noessel (2014) identifiziert. Die Aufgabenaufteilung erfolgt primär gemäss den eigenen Interessen, Lernzielen und Skills.

Um anstehende Arbeiten im Überblick zu behalten, wird eine Taskliste geführt, anhand welcher Aufträge explizit zugewiesen werden können. Ein Tag pro Woche wird reserviert, um grössere Arbeitsblöcke wie Workshops und Usability-Tests durchzuführen. Artefakte oder Ergebnisse, die in Einzelarbeit entstanden sind, werden nach Bedarf konsolidiert.

2.6. Risikoanalyse I.

Am 13.05.23 findet der erste Risikoanalyseworkshop statt, in welchem ‹kein Ziel› als grösstes Risiko identifiziert wird (siehe Anhang 4). Die Formulierung des Forschungsziels – ‹Erhebung von Dialogen› – deutet auf die Vorstellung einer akademischen Forschungsarbeit seitens des Auftraggebers hin. Das Projektteam bevorzugt jedoch eine Forschung anhand eines Anwendungsfalls und einer Gruppe von Nutzenden, was eine zusätzliche Zielsetzung erfordert. Diskussionen über verschiedene Anwendungsfälle wie moderne Screenreader, Zugänglichkeitsstandards, Bildinterpretation oder das Einkaufen von Lebensmitteln ohne die Zuhilfenahme von Händen erhöhen Unsicherheiten und Risiken.

Erste Nachforschungen ergeben zwar ein grosses Aufsehen bezüglich des Themas ‹Voice-Commerce›. Das Alter der Berichte (um 2018) sowie der Umstand, dass die Anwendung bisher nicht im Alltag verbreitet ist, lassen Zweifel an dem formulierten Auftrag aufkommen. Einschätzungsumfragen im privaten Umfeld deuten darauf hin, dass die Anwendung hauptsächlich für spezifische Randgruppen, wie blinde Personen, in Betracht kommt. Eigene Versuche mit Accessibility-Features des iPhones bestätigen diesen Eindruck.

Dies wird im Coaching thematisiert, worauf das Ziel so umformuliert wird, dass der Anwendungsfall sich auf komplexere Produkte konzentriert. Der naheliegendste Anwendungsfall ‹Shopping-Liste› wird unter Bezugnahme auf eine vorhergehende Masterarbeit abgelehnt.

Ein weiteres Risiko stellt das Projektteam selbst dar. Konflikte oder Ausfälle können den Projektfortschritt gefährden. Auf die Zuweisung von Verantwortlichkeiten wird aufgrund der Zweierkonstellation verzichtet.

Lernreflexion

Die Ergänzung der Risikoanalyse um eine Projektrisikomatrix half dabei, kritische Risiken zu identifizieren und zu adressieren. Wir waren unsicher, welche Risiken in einem Projekt mit unklarem Ziel adressiert werden sollen. In der Folge war auch schwer zu bestimmen, welche Massnahmen angemessen und zielführend sind, um diesem Risiko entgegenzusteuern. Das Eintreten von akzeptierten Risiken betrachten wir als Chance zur Erkenntnisgewinnung und freuen uns auf den Lernprozess.

Leistungsreflexion

Durch eine Risikobewertung werden Aufgaben priorisiert. Die gemeinsame Priorisierung wirkt präventiv gegen mögliche Konflikte.

2.7. Stakeholderanalyse

Die bedeutsamste und vorerst einzige anspruchsberechtigte Person dieser Arbeit ist der Auftraggeber, welcher gleichzeitig für das Coaching zuständig ist. Ein regelmässiger Austausch und Einbezug in die Projektarbeit ist im Rahmen der Coachingmeetings gegeben. Zudem werden für entscheidende Phasen zusätzliche Meetings und Workshops eingeplant. Das Projektteam hat vorerst keine weiteren anspruchsberechtigten Personen identifiziert und verzichtet daher auf eine weitergehende Analyse und eine Stakeholdermanagementstrategie.

3. Theoretischer Rahmen

3.1. Einleitung

In diesem Kapitel wird der theoretische Rahmen für die Entwicklung von VUI's vorgestellt, wobei ein besonderes Augenmerk auf die menschliche Sprache und deren Verarbeitung durch Computer gelegt wird. Dabei werden zentrale Theorien der menschlichen Kommunikation, wie das Sender-Empfänger-Modell und Grices Konversationsmaximen, erörtert. Dies bildet die Basis für das Verständnis der komplexen Interaktionen zwischen Mensch und Maschine.

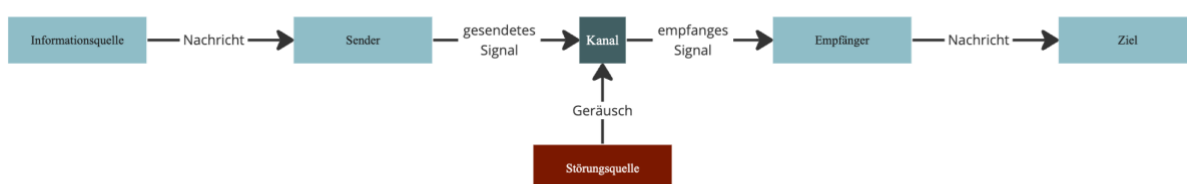
3.2. Natürliche Sprache beim Menschen

Unter dem Begriff «natürliche Sprache» wird die alltägliche Sprache verstanden, welche Menschen nutzen, um miteinander zu kommunizieren. Sie ist vielfältig und wird in verschiedenen Dialekten, Akzenten und Sprachregistern gesprochen. Natürliche Sprache ist ambivalent, mehrdeutig und unvollständig, enthält Redewendungen, Slang, Ironie, Sarkasmus und andere sprachliche Nuancen, die eine automatische Verarbeitung durch Computer erschweren. Sowohl gesprochen als auch geschrieben wird sie in den Bereichen der Kommunikation zwischen Menschen und Computern, der Übersetzung von Sprachen oder der Analyse von Texten eingesetzt.

3.2.1. Sender-Empfänger-Modell

Das Sender-Empfänger-Modell nach Shannon und Weaver (Shannon, 1949) ist ein grundlegendes Konzept in der Kommunikationstheorie. Es wurde ursprünglich 1949 von Claude Shannon und Warren Weaver für die Telekommunikation konzipiert. Das Modell stellt den Prozess der Übertragung von Informationen als linearen Vorgang dar.

Abbildung 4: Sender-Empfänger-Modell nach Shannon-Weaver



Es erleichtert das Verständnis von grundlegenden Elementen und Prozessen der Kommunikation. Die Komplexität menschlicher Kommunikation kann jedoch nicht durch Einfachheit und Linearität abgebildet werden, da sie häufig simultan abläuft und neben der Sprache weitere Aspekte, wie nonverbale Kommunikation, umfasst. Entscheidend ist auch, wie die Gesprächsteilnehmenden die Nachrichten decodieren und interpretieren. Bei Modellen wie dem «Vier-Ohren-Modell» von Schulz von Thun (Schulz von Thun Institut, o. D.) und den Grice'schen Konversationsmaximen werden diese zusätzliche Aspekte berücksichtigt.

3.2.2. Grice'sche Konversationsmaximen

Die Grice'schen Konversationsmaximen (Grice, 1975) sind ein vom englischen Philosophen Herbert Paul Grice formulierter Teil der Kommunikationstheorie. Sie beziehen sich auf die Annahme, dass Teilnehmende in einer Konversation bestimmte unausgesprochene Regeln befolgen, um effektiv zu kommunizieren. Gemäss dem Kooperationsprinzip sollen Gesprächsbeiträge so gestaltet werden, dass sie dem allgemein akzeptierten Zweck oder der anerkannten Ausrichtung des laufenden Gesprächs dienen, an dem die Kommunikationspartner/-innen gemeinsam teilnehmen. Um dieses übergeordnete Prinzip zu befolgen, können Sprechende sich an folgenden Grundsätzen orientieren:

Grundsatz der Quantität

Die Gesprächsbeiträge sollen genauso informativ sein, wie es für den allgemein akzeptierten Zweck des Gesprächs erforderlich ist. Die Beiträge sollen weder zu wenige noch überflüssige Informationen enthalten.

Grundsatz der Qualität

Aussagen in Gesprächen sollen der Wahrheit entsprechen und ausreichende Beweise und Anhaltspunkte enthalten.

Grundsatz der Relevanz

Aussagen sollen relevant zum aktuellen Thema passen und ein Themenwechsel soll klar nachvollziehbar sein. Der Kontext früherer Kommunikationen sowie das Vorwissen der Gesprächspartner/-innen sollen stets einbezogen werden.

Grundsatz des Stils bzw. der Modalität

Es sollen klare und eindeutige Aussagen gemacht werden, welche einer logischen, zeitlichen Ordnung folgen.

3.2.3. Sprecherwechsel

Der Sprecherwechsel, Redeübernahme (engl. *turn-taking*) ist ein Begriff in der Gesprächsanalyse und veranschaulicht die Art und Weise, wie ein geordnetes Gespräch normalerweise abläuft. Er steht für die Vorstellung, dass die Gesprächsteilnehmer/-innen abwechselnd sprechen. Bei Untersuchungen von Soziologen und Soziologinnen geht die Analyse jedoch tiefer und befasst sich mit Themen wie der Frage, wie Menschen wissen, wann sie an der Reihe sind, wie viele Überschneidungen es zwischen den Sprechenden gibt, wann Überschneidungen in Ordnung sind und wie regionale oder soziale Unterschiede berücksichtigt werden können (Drew & Heritage, 2006).

Der deutsche Linguist Klaus Brinker (Brinker & Sager, 2010) hat verschiedene Sprecherwechselformen identifiziert und diese nach Art ihres Zustandekommens sowie nach der Art ihres Verlaufs eingeteilt. Die Art des Zustandekommens wird nach Fremdwahl – auch

als Fremdzweisung oder Fremdselektion bezeichnet – und Selbstwahl – ebenfalls als Selbstzuweisung oder Selbstselektion bekannt – unterschieden. Bei der Fremdwahl kommt der Wechsel zustande, indem der/die Zuhörer/-in bei gleichzeitiger Signalisierung des Sprechenden mit verbalen und/oder nonverbalen Signalen zum Sprechen aufgefordert wird. Besondere Bedeutung wird den nonverbalen Signalen bei der Übergabe zugeschrieben. Bei der Selbstwahl wird unterschieden, auf welche Weise ein/eine Sprecher/-in zu Wort kommt. Brinker differenziert folgende fünf Formen des Sprecherwechsels:

- **Glatter Sprecherwechsel:** Hierbei handelt es sich um den Normalfall, in welchen der Sprecherwechsel nacheinander ohne Simultansequenz erfolgt.
- **Sprecherwechsel mit Überlappung:** Hier kommt es zu einer Phase des simultanen Sprechens; dennoch handelt es sich nicht um eine Unterbrechung, da die Silbe oder Worte eine Art «Knautschzone» darstellen und Äusserungen leicht überhört werden können, ohne dass der Gesprächszusammenhang verloren geht.
- **Sprecherwechsel nach Pause:** Die Länge der Pause zwischen den Gesprächsbeiträgen ist relativ und unter anderem abhängig von der Wahrnehmung sowie der Beziehung der Menschen untereinander. Dennoch wird eine als allzu gross empfundene Gesprächspause als unangenehm wahrgenommen. Hier tritt der Beziehungsaspekt der Kommunikation hervor.
- **Sprecherwechsel durch Unterbrechung mit Simultansequenz:** Der/Die Sprecher/-in wird während des Sprechens unterbrochen.
- **Sprecherwechsel durch Unterbrechung ohne Simultansequenz:** Es erfolgt eine Unterbrechung am Ende eines Satzes, ausgelöst durch Missinterpretation der Gesprächspause (z. B. denken).

Grundsätzlich gilt, dass Sprecherwechsel durch Unterbrechung latent aggressiv wirken. Sie stellen daher für den/die Sprecher/-in eine als unangenehm empfundene Form der Selbstwahl dar (Linke et al., 2004).

3.3. Natürliche Sprache im Umgang mit Computern

3.3.1. Conversational AI

Conversational AI ist die Bezeichnung, unter welcher die Technologien zur Erkennung und Erzeugung einer natürlichen Sprache bei Computern zusammengefasst werden. Über eine dialogorientierte Benutzerschnittstelle (engl. *Conversational User Interface* – CUI) wird der Eindruck einer Konversation zwischen einem Menschen und einem System vermittelt. Typische Beispiele sind virtuelle Assistenten wie Siri und Alexa oder Chatbots wie ChatGPT und Bing Chat (usability, o. D.).

3.3.2. Natural Language Processing

Die Verarbeitung natürlicher Sprache (*Natural Language Processing* – NLP) (Kavlakoglu, 2020), bestehend aus dem Verstehen (NLU) und der Erzeugung natürlicher Sprache (*Natural Language Generation* – NLG), ist ein Teilgebiet der künstlichen Intelligenz (KI). NLP befasst

sich damit, wie Computer Texte und gesprochene Wörter auf ähnliche Weise verstehen und aussprechen können wie Menschen (IBM, o. D.).

Beim NLP wird Computerlinguistik mit maschinellen Lernmodellen kombiniert. Zusammen ermöglichen diese Technologien es Computern, menschliche Sprache in Form von Text- oder Sprachdaten zu verarbeiten und ihre volle Bedeutung, inklusive Absichten und Gefühlen, zu <verstehen>. Zu den Aufgaben von NLP gehören unter anderem folgende Aspekte:

Die **Spracherkennung**, auch Speech-to-Text (STT) genannt, hat die Aufgabe, gesprochenen Text in Textdaten umzuwandeln. Besondere Herausforderung für STT ist die unterschiedliche Art (Geschwindigkeit, Akzent, Deutlichkeit, Betonung, Grammatik), wie Menschen sprechen.

- Mit **Part-of-Speech-Tagging** wird die Wortart eines Wortes oder Textteils bestimmt, um grammatische Regeln anwenden zu können.
- Die **Disambiguierung** des Wortsinns ist die semantische Bestimmung eines Wortes mit mehreren Bedeutungen. Sie ist ein Prozess der semantischen Analyse, in dem die kontextuelle Bedeutung von Wörtern analysiert wird.
- Bei der **Erkennung** benannter Entitäten (engl. *named-entity recognition* – NER) werden nützliche Wörter oder Phrasen identifiziert. Beispielsweise werden <Zürich> als Ort und <Philipp> als Name erkannt.
- Bei der **Auflösung von Ko-Referenzen** geht es darum festzustellen, ob und wann sich zwei Wörter auf dieselbe Entität beziehen. Das häufigste Beispiel ist die Bestimmung der Person oder des Objekts, auf das sich ein bestimmtes Pronomen bezieht.
- Bei der **Stimmungsanalyse** (engl. *sentiment analysis*) wird versucht, subjektive Eigenschaften, Einstellungen, Emotionen, Sarkasmus, Verwirrung und Misstrauen aus Texten zu extrahieren.

3.3.3. Natural Language Understanding

Beim Verstehen natürlicher Sprache (engl. *Natural Language Understanding* – NLU) wird die syntaktische (grammatische Struktur) und semantische (beabsichtigte Bedeutung) Analyse von Text und Sprache genutzt, um die Bedeutung eines Satzes zu ermitteln. In diesem Zusammenhang wird eine entsprechende Ontologie, d. h. eine Datenstruktur, erstellt, mit welcher die Beziehungen zwischen Wörtern und Sätzen festlegt wird. Während dies bei Gesprächen zwischen Menschen auf natürliche Art erfolgt, ist für Maschinen eine Analyse erforderlich, um die beabsichtigte Bedeutung zu verstehen. Die menschliche Fähigkeit, zwischen Homonymen und Homophonen (siehe Glossar) zu unterscheiden, veranschaulicht die sprachlichen Feinheiten.

3.3.4. Natural Language Generation

Die Generierung natürlicher Sprache (NLG) ist der Prozess, bei dem Text in menschlicher Sprache erzeugt wird. Der Text kann durch Text-to-Speech-Dienste wie ElevenLabs in ein Audioformat umgewandelt werden. Durch den Einsatz von neuronalen Netzwerken kann die Generierung von Texten in Echtzeit erfolgen.

3.3.5. Intelligenter persönlicher Assistent

Ein intelligenter persönlicher Assistent ist eine Software, die es ermöglicht, Informationen abzufragen, Dialoge zu führen und Assistenzdienste durch Kommunikation in natürlicher Sprache zu erbringen. Die Assistenten können Chatbot-Funktionen beinhalten, um Konversation mit dem Ziel zu simulieren, die Interaktion mit Nutzenden zu erleichtern. Die Interaktion kann über eine grafische oder sprachliche Schnittstelle erfolgen. Anwendungsfälle von virtuellen Assistenten sind Informationsabfragen, Geräte zum Steuern von Hausautomatisierung, Medienwiedergabe und andere grundlegende Aufgaben wie die Verwaltung von E-Mails, Aufgabenlisten und Kalender (Hoy, 2018). Zu den bekanntesten Sprachassistenten gehören Amazons Alexa, Apples Siri, Microsofts Cortana und der Google Assistant (Bushnell, 2023).

3.4. Konversationsdesign

Beim Konversationsdesign (engl. *conversation design*) werden Technologie, Psychologie und Sprache kombiniert, um Chatbots und Sprachassistenten hilfreicher und natürlicher zu gestalten (Conversation Design Institute, 2020). Im Konversationsdesign nach Moore und Arar (2019) wird ein *Natural Conversation Framework* (NCF) erarbeitet. Des Weiteren werden UX-Patterns des Konversationsdesigns sowie ein genereller, auf Design-Thinking aufbauender Conversational-UX-Design-Prozess beschrieben. Im NCF werden Modelle und Muster menschlicher Interaktionen dokumentiert, welche für die Gestaltung dialogorientierter Benutzerschnittstellen verwendet werden. Muster und Modelle werden in zwei Hauptphasen eruiert:

Konversationsanalyse (engl. *conversation analysis*): Ein Verständnis menschlicher Konversationen und deren Eigenheiten durch Dialoganalysen (siehe auch [«natürliche Sprache beim Menschen»](#)) wird erlangt.

Konversationen verfassen (engl. *conversation authoring*): Dialoge werden mithilfe von Zustandsdiagrammen dargestellt. Bedingungen, Antwortvarianten und Verzweigungen definieren die Verlaufsmöglichkeiten einer Konversation.

Konventionelle Chatbots und Assistenten (engl. *agent*) bauen auf vorbestimmten Inhalten und regelbasierten Dialogstrukturen auf. Der Agent kann nur auf Befehle bzw. Fragen reagieren, welche im Programm enthalten sind. Die erwarteten Unterhaltungswege werden in einem Dialogdiagramm dargestellt. Die Dialoge folgen einem vorbestimmten Ablauf und sind nicht mit einer natürlichsprachlichen Interaktion vergleichbar.

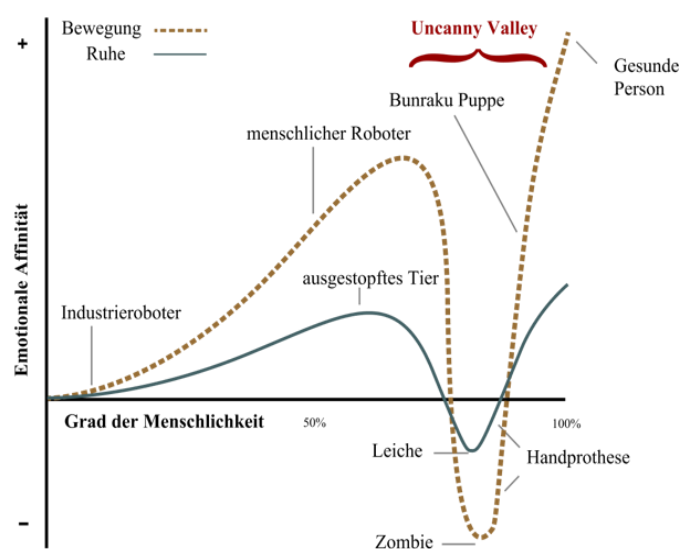
Mit Conversational AI werden neue Möglichkeiten eröffnet, wie VUIs gestaltet werden können. Statt natürlichsprachliche Gespräche nachzuahmen, werden die Fähigkeiten von *Large Language Models* (LLM) genutzt, um natürliche Sprache zu verstehen und zu generieren. In der vorliegenden Arbeit wird untersucht, wie diese Modelle eingesetzt werden können.

3.4.1. Reparaturmechanismen

Reparatur durch Nutzende (engl. *user repair*) bezieht sich auf die Fähigkeit des Menschen, Missverständnisse oder Fehler in einer Konversation zu erkennen und durch Wiederholung, Paraphrasieren oder nonverbale Techniken zu korrigieren. Mit ‹Reparatur durch das System› (engl. *agent repair*) wird dasselbe aus Sicht des Agenten beschrieben. Reparaturmechanismen sind Bestandteil des Konversationsdesigns und werden mittels Techniken wie Verständnisüberprüfung, Bestätigung oder Paraphrasierung angewendet (Moore & Arar, 2019).

3.5. Uncanny Valley

Abbildung 5: Uncanny Valley



Quelle: Eigene Darstellung nach Mori et al.

Der Ausdruck ‹unheimliches Tal› (engl. *uncanny valley*), erstmals 1970 vom japanischen Robotiker Masahiro Mori beschrieben, steht für ein Phänomen in der Robotik, welches ein Gefühl des Unbehagens bei menschlichen Betrachtern und Betrachterinnen hervorruft. Mori stellte Folgendes fest: Je menschenähnlicher Roboter werden, desto sympathischer und vertrauenswürdiger erscheinen sie. Sobald sie eine fast, aber nicht vollständig menschliche Erscheinung erreichen, lösen sie Ablehnung oder

Unbehagen aus (Abbildung 5). Sowohl positive als auch negative Affinitäten treten bei sich bewegenden Objekten im Vergleich zu ruhenden verstärkt auf (Mori, MacDorman, & Kageki, 2012). Obwohl diese theoretische Annahme bisher nicht wissenschaftlich belegt ist, spielt sie in der Gestaltung von menschenähnlichen Technologien eine Rolle, da sie weitestgehend anerkannt ist.

3.6. Eliza-Effekt

Der Begriff leitet sich von ELIZA ab, einem Computerprogramm, das in den 1960er-Jahren von Joseph Weizenbaum entwickelt wurde. Es simuliert einen Psychotherapeuten/-in, indem es mithilfe einfacher Mustervergleichstechniken auf Eingaben von Benutzer/-innen reagiert. Mit dem ELIZA-Effekt wird die Tendenz von Menschen bezeichnet, Computerprogrammen menschenähnliche Eigenschaften und Intelligenz zuzuschreiben. Dies gilt insbesondere für solche Programme, in denen eine menschliche Unterhaltung nachgeahmt wird. Die modernen Gegenstücke wie ChatGPT, Bing oder Google Bard veranschaulichen ein Gefühl von Menschlichkeit. Nutzende bauen eine Bindung zu ihnen auf, da die KI-Bots fälschlicherweise als ‹Wesen› mit menschenähnlichem Bewusstsein wahrgenommen werden. Der ELIZA-Effekt verdeutlicht die menschliche Neigung, künstlichen Wesen Intentionalität, Emotionen und

Intelligenz zuzuschreiben, weil diese in der Lage sind, menschenähnliche Interaktionen zu imitieren. Die Zuschreibung menschlicher Eigenschaften ist vielmehr ein psychologisches Artefakt als ein Beweis für die technische Raffinesse. Mit dem ELIZA-Effekt kann die Notwendigkeit verdeutlicht werden, dass die UX-Forschung in der KI-Entwicklung über die bloße Nachahmung menschlicher Interaktion hinausgeht. Im Kontext der Forschung bietet der ELIZA-Effekt ein kritisches Korrektiv, mithilfe dessen Interaktionen mit dialogorientierten Schnittstellen bewertet werden können (Sponheim, 2023).

4. Zyklus 0: Evaluation von Anwendungsfällen

4.1. Einleitung

Es werden drei Anwendungsfälle evaluiert, die sich mit natürlicher Spracheingabe, möglichst ohne Zuhilfenahme von Händen, gestalten lassen. Einfache Bedienbarkeit und Zugänglichkeit werden höher gewichtet als spezifische Bedürfnisse einer Zielgruppe. Das Ziel ist, eine Applikation zu entwerfen, welche einen der folgenden Anwendungsfälle erfüllt:

- **Einkaufsberatung:** Einkauf von nicht alltäglichen Produkten, welche beratende Interaktionen vor dem Kauf voraussetzen;
- **Reiseplanung:** Eine Reiseplanung in Form von Reservationen in einem Restaurant unterstützen;
- **Vorlesen:** Wissensabfragen oder das Abfragen und Hören von Nachrichten als Alternative zum Lesen.

4.2. Recherche

4.2.1. Desk-Research

In einer repräsentativen Studie wurde erhoben, dass 50 % der Amerikaner/-innen bereits Einkäufe mit einem Sprachassistenten getätigt haben (PwC, 2018). Die häufigsten Anwendungsfälle sind jedoch einfache Aufgaben wie das Abfragen von Wetter, von Nachrichten oder Internetsuchanfragen. Unter den Studienteilnehmenden wies die Gruppe der 25- bis 49-Jährigen die höchste Nutzung von Sprachassistenten auf (65 % mindestens einmal täglich). Das meistverwendete Gerät ist das Smartphone (59 %). In Europa hat der Detailhändler Carrefour eine Kooperation mit Google gestartet, welche es ermöglicht, Einkäufe via Voice zu tätigen (Carrefour, 2020). Weiteren Berichten zufolge verfolgen Unternehmen wie Google, Apple und Amazon Anstrengungen, um Sprachassistenten für die kommerzielle Nutzung tauglich zu machen (Westerhold, 2018).

4.2.2. Interview mit einer Fachperson

In einem Experteninterview mit einer Fachperson aus dem Bereich E-Commerce werden spezifische Anwendungsfälle für Sprachassistenten und für künstliche Intelligenz in kommerziellen Kontexten diskutiert (siehe Anhang 6). Das Potenzial besteht eher in der effizienten Anrufweiterleitung als in komplexeren Anwendungen wie bspw. einem Alexa-Skill für Einkaufsprozesse. Ein textbasierter Chat wird gegenüber Sprachinteraktionen bevorzugt, da dieser interaktiver und effizienter ist. Der Schwerpunkt der gegenwärtigen Entwicklungen liegt in der Integration von KI über das gesamte Produktsortiment hinweg sowie in der Verbesserung der Relevanz der Vorschläge. Benchmarking-Verfahren werden eingesetzt, um die Leistung zu messen. Die Erkennung von Schweizerdeutsch und Fachjargon bleibt eine Herausforderung, die durch die breite Produktvielfalt des Unternehmens verstärkt wird. Ein kleineres Produktsortiment soll für Sprachinteraktionen besser geeignet sein. Systeme, die Kundenpräferenzen lernen und sich anpassen, können die Führung von Nutzenden optimieren

und durch die Integration von Gerätedaten, Kontoinformationen oder Kalendereinträgen die Nützlichkeit von Sprachassistenzsystemen verbessern.

4.2.3. Dialogstrukturen in dialogorientierten Bedienoberflächen

Vorhandene Dialogstrukturen werden in bestehenden CIs, VUIs und IVRs untersucht. Die identifizierten Dialogstrukturen sind untereinander kombinierbar. Es werden folgende Dialogtypen identifiziert:

Tabelle 2: Dialogstrukturen in dialogorientierten Bedienoberflächen

Dialogstruktur	Beschreibung	Verwendung
Linearer Dialog	In abwechselnder Reihenfolge wird Konversation mit einem glatten Wechsel geführt.	VUI, v. a. Aufgabenzuweisung
Baumstrukturdialog	Der Dialog verzweigt sich, basierend auf den Antworten.	Rollenspiele, IVR
Menübasierter Dialog	Die Navigation erfolgt anhand von Schlüsselworten.	IVR, VUI
Freiformdialog	Es werden offene und natürliche Dialoge simuliert.	ChatGPT
Reaktiver Dialog	Der Dialog wird durch Aktionen von Nutzenden ausgelöst.	VUI, Chat-Widgets
Regelbasierter Dialog	Es liegen spezifische Regeln vor, auf welche Nutzende reagieren. Dies setzt ein Wissen über die Regeln voraus.	VUI
Slot-Filling-Dialog	Nutzende müssen eine fehlende Angabe ergänzen.	IVR

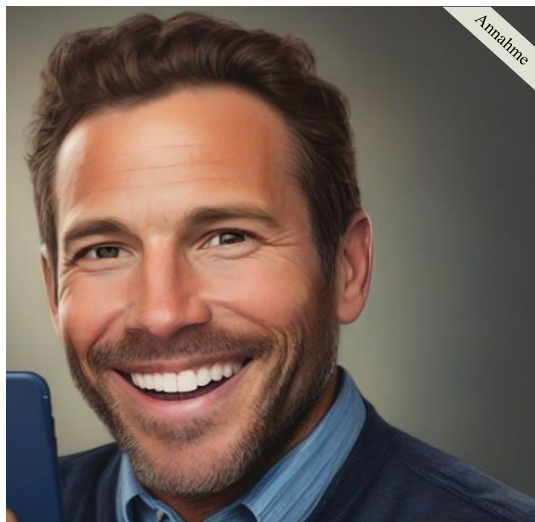
4.3. Annahmen

4.3.1. Problem-Statement

*Der Markt für **sprachbasierte Assistenten** konzentriert sich hauptsächlich auf einfache Informationsabfragen. Im Bereich des **Shoppings**, der **Reservierungen** und bei **komplexeren Informationsanfragen** über ein VUI existieren keine (oder keine zufriedenstellenden, etablierten) Produkte oder Dienstleistungen. Unser VUI-Assistent schliesst diese Lücke durch **natürlichsprachliche Interaktion mit den Nutzenden**. Wir werden uns zunächst auf alle Arten von Nutzenden konzentrieren.*

4.3.2. Persona

Abbildung 6: Persona Zyklus 0



Jasper Jedermann ist 39 Jahre alt und Lehrer. Er ist in der Region Zürichsee wohnhaft. Seine technischen Fähigkeiten schätzt er selbst als durchschnittlich ein. Er nutzt ein iPhone und hat erste Erfahrungen im Umgang mit ChatGPT gemacht. Beruflich und privat ist er zeitlich eingespannt, weshalb für ihn eine Vereinfachung von alltäglichen Aufgaben, wie Informationsabfragen aus dem Internet, Einkaufen von unterwegs oder eine Reservation im Restaurant, wesentlich ist. Hinderlich ist, dass er seinen Beruf nicht an einem Schreibtisch ausübt und sein iPhone für diese Aufgaben

nutzt. Er wünschte sich, dass er jemanden hätte, an den er seine Aufgaben delegieren könnte. Am liebsten würde er seine Anliegen jemandem diktieren können.

4.3.3. Hypothese

*Wir sind der Auffassung, dass Nutzende die Use-Cases **Wissensfragen**, **Tischreservierung**, **Online-Shopping** ausschliesslich mit einem VUI lösen können.*

4.4. Design

Als Design dienen die Sprachassistenten von Apple, Google und Microsoft.

Abbildung 7: Siri von Apple



Abbildung 8: Google Assitant



Abbildung 9: Bing von Microsoft



4.5. Minimum Viable Product

Um die Leistungsfähigkeit und Limitationen der drei weitverbreiteten Sprachassistenten – Siri, Google Assistant und Bing – zu beurteilen, wird ein Benchmark-Test durchgeführt. Alexa wird aufgrund ihrer geringeren Verbreitung in der Schweiz nicht berücksichtigt. Am Test nehmen sechs männliche Probanden im Alter von 30, 37 (2 Personen), 38, 46, und 48 Jahren sowie eine Probandin im Alter von 37 Jahren teil. Sie gehören zum persönlichen Umfeld des Projektteams. Der dreiteilige Test wird in privaten Räumlichkeiten durchgeführt, wobei über Microsoft Teams ein Anruf auf einem iPhone 13 initiiert wird. Der Bildschirm des Smartphones wird geteilt, um Audio- und Videoaufzeichnungen erstellen zu können. Die Testpersonen sind angehalten, während des gesamten Tests Hochdeutsch zu sprechen. Im ersten Teil werden Daten zur Nutzung von Sprachassistenten und ChatGPT sowie demografische Daten erhoben. Im Hauptteil werden die drei Anwendungsfälle anhand der Aufgaben überprüft (Siehe Tabelle 3: Übersicht der Anwendungsfälle, Aufgaben, Antworten). Nachrichten und Wissensfragen werden im Vorfeld als einfache und lösbare Aufgaben betrachtet, weshalb zwei Wissensfragen und zwei aktuelle Fragen aufgegeben werden. Die Testpersonen werden angehalten, die Eingaben mit Sprache und ohne Touchinterface auszuführen.

Tabelle 3: Übersicht der Anwendungsfälle, Aufgaben, Antworten

Anwendungsfall	Aufgabe	Richtige Antwort
Wissen	Wer ist amtierender Fussballmeister der Schweiz?	BSC Young Boys
Wissen	Wer war der/die Regisseur/-in von Forrest Gump?	Robert Zemekis
Aktuell	Finde heraus, wie das Wetter morgen in Zürich wird.	Sonniger Tag, ca. 29°
Aktuell	Wie viele Menschen gibt es auf der Welt?	Ca. acht Milliarden
Aufgabe: Reiseplanung	Finde ein italienisches Restaurant in Zürich, reserviere einen Tisch und erstelle einen Kalendereintrag.	Jedes Restaurant in Zürich mit italienischem Essen
Aufgabe: Einkaufen	Du möchtest ein Stativ für eine Sony-Alpha-Kamera. Suche und finde eines.	Manfrotto als wahrscheinlichste Nennung

Im anschliessenden Cooling-off werden die Teilnehmenden zu ihren Erfahrungen befragt. Die Tests finden am 10.6.2023 und 17.06.2023 statt und dauern pro teilnehmende Person ca. eine Stunde. Das vollständige Testszenario ist im Anhang 7 zu finden.

4.6. Feedback

Bewertet wird bei Wissensfragen, ob die Antwort richtig ist, und bei der Restaurantreservierung, ob das Restaurant italienisch ist. In der anschliessenden Auswertung wird für jede Aufgabe bei jedem Assistenten darauf geachtet, wie die Nutzenden das System in den Zustand <Zuhören> versetzen, ob die Frage durch das System verstanden wird, ob die Testperson die korrekte Antwort gibt und was das System geantwortet hat. Die Versuchsanzahl wird gezählt, bis die Aufgabe beendet (erfolgreich oder aufgegeben) ist. Die vollständige Auswertung ist in den Anhängen 7 – 16 zu finden.

4.6.1. Quantitative Beurteilung

Insgesamt werden von sieben Teilnehmenden mit drei Assistenten 168 Aufgaben ausgeführt. Erfolgreich gelöst werden 107, was einer gesamthaften Erfolgsrate von 64 % (Bing 66 %, Google 64 % und Siri 61 %) entspricht. Der Grossteil der Erfolgsrate fällt auf Wissensfragen zurück (Bing 100 %, Google 93 %, Siri 86 %). Bei den Aufgaben <Restaurantreservierung> und <Einkaufen> liegt die Erfolgsrate bei 38 %. Eine Reservierung vorzunehmen und einen Kalendereintrag zu erstellen, hat niemand geschafft.

Tabelle 4: Kriterien und Erfolgsraten

Kriterium	Bing	Google	Siri	Total
Wurde die Aufgabe bewältigt?*	66 %	64 %	61 %	64 %
Aktuelles	71 %	93 %	93 %	86 %
Aufgabe	46 %	36 %	32 %	38 %
Wissen	100 %	93 %	86 %	93 %
Ohne Graphical User Interface (GUI) lösbar	66 %	39 %	18 %	41 %
Wurde die Frage vom System verstanden?	89 %	77 %	73 %	80 %
Wurde die Aufgabe bzw. Frage richtig beantwortet?	77 %	66 %	63 %	66 %

*unabhängig von der Antwort

Ohne visuelle Benutzerschnittstelle sind 41 % der Aufgaben lösbar (Bing 66 %, Google 39 % und Siri 18 %). Die Frage nach dem/der Fussballmeister/-in wird am häufigsten, insgesamt 59-mal, probiert. Pro Person wird sie durchschnittlich 8,4-mal versucht.

4.6.2. Mikrointeraktionen

Siri kann auf Knopfdruck, mit einem Wakeword und mit Drücken auf den Bildschirm in den Zustand <Zuhören> versetzt werden. Der Zustand kann durch Drücken der Taste auf der Seite und auf dem Bildschirm wieder aufgehoben werden, wodurch die Eingabe abgebrochen wird. Bei allen Teilnehmenden geschieht dies unbeabsichtigt, weshalb mehrere Anläufe notwendig sind. Diese Fehlversuche werden von der Auswertung ausgeschlossen. Der Google Assistant kann mit einem Wakeword oder Displaybutton aktiviert werden. Er zeigt den gesprochenen Text für den Bruchteil einer Sekunde an und gibt auf dem Bildschirm die Suchresultate aus (siehe Abbildung 4). Bing wird über einen Displaybutton in den Zustand <Zuhören> versetzt. Wörter der Anfrage werden progressiv eingeblendet.

4.6.3. Einschätzung der Assistenten

Siri erhält die niedrigste Bewertung unter den geprüften Sprachassistenten. Wenn Siri nicht in der Lage ist, die Anfragen zu beantworten, wird stattdessen eine Suchresultat-Seite auf dem Bildschirm dargestellt. Dies gilt sogar für Aufgaben, die ohne visuelle Rückmeldung, wie etwa Wetteranfragen, lösbar sind. Siri ist für einfache ausführende Funktionen wie das Abspielen von Musik oder das Erstellen von Kalendereinträgen geeignet. Die Studienteilnehmenden bekunden zudem Schwierigkeiten, den «Zuhören»-Status von Siri zu erkennen.

Der Google Assistant fällt weder negativ noch positiv auf und wird als effizient eingestuft. Die Sprachausgaben liefern für die Testpersonen gerade genug Informationen. Obwohl Google eine Möglichkeit bietet, über die Website Tischreservierungen zu tätigen, gelingt es keiner der Testpersonen, eine Reservierung auszuführen. Vier Testpersonen gelangen auf die Website oder an die Telefonnummer des ausgesuchten Restaurants. Google kann in diesem Fall

kontextbezogen agieren. Gut gelöst ist das Feedback zum Systemstatus. Nutzende erkennen, wann der Assistant bereit ist, die Spracheingabe aufzunehmen. Die Anfrage wird jeweils auf dem Display angezeigt. Bei einigen Testpersonen gibt Google wiederholt an, die Frage nicht verstanden zu haben, obwohl sie exakt auf dem Bildschirm wiedergegeben wird. Die Ausgabe per Voice beschränkt sich bei komplexeren Anfragen auf das Hinweisen auf die Suchergebnisse. Nur einfache Anfragen wie nach dem Wetter oder Wissensfragen werden vollständig als Voice-Output wiedergegeben.

Bing ist der einzige Sprachassistent, der Schweizerdeutsch spricht und ChatGPT (Version 3.5) verwendet. Die Teilnehmenden zeigen sich überrascht über die Fähigkeit des Assistenten, Schweizerdeutsch zu verstehen und zu sprechen. Die Antworten von Bing werden als detailliert wahrgenommen, was jedoch nicht immer als sinnvoll empfunden wird. Bei konkreten Wissensfragen, wie bei der Frage nach dem Regisseur bzw. der Regisseurin eines Films oder dem Wetter, bevorzugen die Teilnehmenden kurze und prägnante Antworten. Die Antworten von Bing sind nicht immer korrekt. Die Frage nach dem/der amtierenden Schweizer Meister/-in im Fussball wird wahlweise mit FC Basel, FC Zürich oder BSC Young Boys beantwortet. Adressinformationen sind fehlerhaft, und in einem Fall erfindet Bing einen nicht existierenden Film. Der Umgang mit Zahlen und fremdsprachigen Wörtern, etwa englische Namen wie Forrest Gump oder italienische Restaurantnamen, wird als unnatürlich empfunden.

4.6.4. Feedback der Anwendenden

Festgestellt wird, dass bei den technisch weniger versierten Nutzenden kein Bedürfnis für Sprachinterfaces besteht. Als mögliche Anwendung werden etablierte Fälle wie das Erledigen von einfachen Aufgaben oder das Abfragen von Wissensfragen genannt. Die unzureichende Spracherkennung mache die Technologie uninteressant. Dies kann durch Aussagen wie «Man muss sich dem System anpassen» und «Man muss lernen, wie es einen versteht, damit es funktioniert» verdeutlicht werden. Sie gehen davon aus, die Aufgaben auf konventionelle Weise schneller lösen zu können, obwohl die Sprache als intuitives Werkzeug empfunden wird. Ein Bewusstsein für Datenschutz und die damit verbundenen Herausforderungen ist bei allen Testpersonen vorhanden. Die Situation wird aufgrund von Gleichgültigkeit und fehlender Einflussmöglichkeiten akzeptiert. Komfort und Bequemlichkeit werden höher bewertet.

4.6.5. Fazit

Die Zahlen der quantitativen Beurteilung bedürfen einer Erklärung. Selbst einfache Wissensabfragen, bei welchen eine einzig zulässige und kurze Antwort erwartet wird, sind nicht zuverlässig beantwortbar. Auch Personen, die den Umgang mit Sprachassistenten gewohnt sind, müssen mehrere Anläufe nehmen, bevor eine Antwort gegeben werden kann. Nicht konsistente und fehlende Interaktionsrichtlinien in Bezug auf Mikrointeraktionen führen dazu, dass Aufgaben mit einem GUI deutlich schneller und zuverlässiger gelöst werden können. Moderne Assistenten wie Bing bieten diesbezüglich ein besseres Erlebnis. Die Unzuverlässigkeit, verursacht durch «erfundene» Antworten bei generativen Sprachmodellen, stellt jedoch die Nützlichkeit der Technologie infrage. Unternehmen wie Apple, Google und Microsoft sind trotz marktbeherrschender Position und jahrelanger Erfahrung nicht imstande,

ein adäquates Nutzungserlebnis in Bezug auf Sprachassistenten zu bieten. Der generelle Nutzen von Sprachassistenten ist unter diesen Gesichtspunkten zu hinterfragen.

Bei diesem Versuch wird beobachtet, dass sich die Studienteilnehmenden im Laufe des Tests im Umgang mit den Assistenten anpassen. Dies betrifft insbesondere die sprachliche Anpassung an die Assistenten. Nur vereinzelt wird auf der eigenen Formulierung beharrt. Bei Nichtverstehen seitens der Assistenten wird schrittweise umformuliert, vereinfacht oder auf Schlagwortanfragen reduziert (siehe 3.4.1 Reparaturmechanismen)

Die Reparaturmechanismen der Assistenten sind ungenügend und beschränken sich auf Aussagen wie ‹Das habe ich nicht verstanden›. Eine Absicht, die Konversation aufrechtzuerhalten, kann nicht erkannt werden. KI-Systeme verhalten sich unvorhersehbar: Die Antworten können korrekt, falsch oder sogar frei erfunden sein. Unter den Testpersonen würde niemand den Antworten ohne weitere Recherchen vertrauen.

4.7. Reflexion

Lernreflexion

Usability-Tests sind eine hervorragende Methode, um mit Nutzenden ins Gespräch zu kommen und Erkenntnisse über ihre Bedürfnisse zu sammeln. Die Entscheidung, einem Benchmarktest durchzuführen, brachte Erfahrungsgewinn in der Durchführung von VUI-Usability-Tests. Eine besonders wertvolle Idee war die Aufzeichnung mit Microsoft Teams. Bei diesem Programm kann ein Anruf mit sich selbst initiiert werden, wobei sich das Smartphone mit dem Computer verbindet und der Bildschirm des Telefons geteilt wird. Dadurch wird sichtbar, was auf dem Bildschirm geschieht, was die Personen sagen und wie ihre Mimik und Gestik ausfällt. Bei der anschliessenden Auswertung wurde festgestellt, dass Interaktionen auf dem Bildschirm sowie das Gesprochene und die Körpersprache bisweilen konträr sind. Eine Person kann eine freundliche Antwort geben, während sie gleichzeitig dem Sprachassistenten eine beleidigende Geste zeigt. Eine Testperson benötigte 13 Anläufe, bis ein Resultat erreicht wurde. Die Person gab dennoch an, dass der Sprachassistent ihrer Einschätzung nach gut funktioniert, während eine andere Testperson, die nur zwei Versuche benötigte, eine negative Bewertung abgab.

Leistungsreflexion

Die aufgestellte Hypothese wurde widerlegt. Die im Testszenario definierten Aufgaben konnten mittels Sprachinteraktionen nur zu 41 % gelöst werden.

Bezüglich des Status quo von Sprachassistenten bleibt die Erkenntnis, dass die von grossen Technologiefirmen angebotenen Sprachassistenten die Erwartungen der Nutzenden hinsichtlich der Usability des Voice-Interfaces nicht erfüllen. Weniger versierte und interessierte Versuchsteilnehmende waren der Ansicht, dass sie sich dem System anpassen bzw. lernen müssen, wie es sie versteht, damit es funktioniert.

Die Tests konnten mit allen Testpersonen innerhalb der erwarteten Zeit durchgeführt werden. In dieser Phase des Projekts war es dem Team wesentlich, Nutzungsszenarien für VUIs zu evaluieren und direktes Feedback zu erhalten. Künftig soll der Umfang innerhalb eines Lean-UX-Zyklus reduziert werden. Viele und unterschiedliche Fragestellungen erschweren die Auswertung und die Formulierung eines prägnanten Fazits.

5. Risikoanalyse II.

Um weiterhin potenzielle Gefahren und Unsicherheiten identifizieren zu können, wird die Risikoanalyse wiederholt. Als vorrangiges Risiko wird der ‹Use-Case› identifiziert. Die mangelnde Akzeptanz, verursacht durch den aktuellen Stand der Technologie, sowie die Tatsache, dass Aufgaben einfacher über ein visuelles Interface zu bearbeiten sind, gefährden das Projekt. Anhand des Feedbacks des vorangegangenen Zyklus konnte kein konkretes Problem benannt werden, bei dem ein VUI den Menschen effektiv bei der Erledigung seiner Aufgaben unterstützt. Befürchtet wird, dass keine realistischen Forschungsergebnisse ohne ein zugrunde liegendes menschliches Bedürfnis erzielt werden können. Die offene Ausgangssituation im Projekt birgt das Risiko der fehlenden Fokussierung und kann dazu führen, dass versucht wird, zu viele Aufgaben gleichzeitig zu bearbeiten.

Die technologische Entwicklung bewirkt, dass Forschungsergebnisse rasch veralten. Die Bewertung der Ergebnisse aus dem Zyklus 0 sind unter der Berücksichtigung dieses Umstands zu interpretieren. Fehlende Standards und verschiedene Entwicklungsstufen der Hersteller mobiler Geräte können einen Vergleich der Ergebnisse erschweren. Zudem wird erkannt, dass sich generative Sprachmodelle unvorhergesehen verhalten. Zum Zeitpunkt der Abfassung der Masterarbeit sind keine Methoden bekannt, mit welchen diese technischen Herausforderungen bewältigt werden können. Das Risiko wird deshalb mit einer hohen Eintrittswahrscheinlichkeit und Auswirkung bewertet. Die Risikoanalyse befindet sich im Anhang 17

6. Zyklus 1: Quiz

6.1. Einleitung

In einem Workshop mit dem Auftraggeber werden weitere Anwendungsfälle diskutiert. Vier Szenarien kommen für die anstehende Entwicklung in Betracht: «Vorlesen im Auto», «Einkaufen von tierischen Bedarfsartikeln», «Reiseführung» und «interaktives Radio». Es zeigt sich, dass die Anforderungen, Funktionalitäten und vorhandenen Objektentitäten dieser Anwendungsfälle einander ähnlich sind. Nach dem Workshop liegt der Fokus auf der Domäne «ältere Menschen». Der Projektauftrag wird angepasst (siehe Anhang 20). Ein Produkt namens «Max Memory» wird entwickelt, das darauf abzielt, kognitive Fähigkeiten von Menschen einer bestimmten Zielgruppe spielerisch zu fördern.

Abbildung 10:
Einkaufsberatung



Abbildung 11: Quiz



Abbildung 12: Vorlesen im Auto

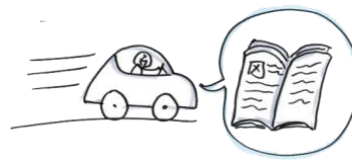


Abbildung 13: Reiseführer/
in



6.2. Stakeholderanalyse

Am 14.07.2023 wird eine studienübergreifende Zusammenarbeit mit Abinas Kuganathan aus dem Institut für Softwareentwicklung vereinbart und das Team um eine Person erweitert. Im Rahmen einer Diplomarbeit wird ein codierter Prototyp auf Basis der erhobenen Anforderungen entwickelt. Markus Stolze bleibt wichtigster Stakeholder.

6.3. Recherche

6.3.1. Definition und begriffliche Abgrenzung des Ausdrucks «ältere Menschen»

Der Begriff «ältere Menschen» bezieht sich auf Personen, welche eine bestimmte Altersgrenze überschritten haben. In der Schweiz wird der Ausdruck üblicherweise mit dem Eintreten in das Rentenalter in Verbindung gebracht. Ökonomische, gesundheitliche und soziale Faktoren verleihen dem Begriff Variabilität. In Studien wurde erhoben, dass die kognitive Leistung des Menschen im Alter von etwa 45 Jahren abnimmt (Blech, 2022). In extremen Fällen kann sich bereits im Alter von etwa 50 Jahren Demenz entwickeln (Rydlind, 2023). Um die kognitiven Fähigkeiten anzuregen, werden in Altersheimen Aktivierungstherapien angeboten. Das Gedächtnis wird angeregt, in dem früheres Wissen und Erfahrungen wiedergegeben werden müssen. Die Aktivierungstherapie dient auch der sozialen Interaktion.

6.3.2. Interviews mit Fachpersonen

Gespräche mit zwei Pflegefachkräften (siehe Anhang 24 und Anhang 25) geben Aufschluss über die Herausforderungen älterer Menschen. Dies bezieht sich insbesondere auf die Situation in Altersheimen. Hauptprobleme sind Einsamkeit und eine fehlende Tagesstruktur, welche die Neigung zu psychischen Krankheiten fördern kann. Die Sorge um das physische wie auch

psychische Wohlergehen werden als zentrale Aufgaben des Pflegepersonals genannt. Dem sozialen Austausch und dem Zuhören von älteren Menschen möchten die Fachpersonen gerne mehr Arbeitszeit widmen. Administrative Pflichten schränken jedoch die Zeit ein, die Pflegenden mit ihren Klienten und Klientinnen verbringen können. Altersheime fördern die Tagesstruktur mit Aktivitäten wie kulturellen Programmen, Gedächtnistraining und Lesezirkeln. Zudem bieten sie verschiedene Formen von Bewegungs- und Kognitionstherapien an. Das Wissen um solche Programme sowie das Überwinden der eigenen Hemmschwelle sind Voraussetzungen, damit die Angebote genutzt werden. Einschränkungen in der Mobilität, bei den Finanzen oder andere persönliche Termine wie Arztbesuche können den Zugang erschweren.

6.3.3. Feldstudie

Im Altersheim Klus finden wöchentlich am Mittwoch und Donnerstag Aktivierungstherapien statt. In der Feldstudie wird beobachtet, wie eine Gedankenreise nach Bosnien-Herzegowina unternommen wird. Eine Gruppe von 12 Personen tauscht in anregenden Gesprächen Erinnerungen und Erfahrungen über das Land, die umliegenden Länder und deren Geschichte aus. Geografiekenntnisse werden durch eine Übung abgefragt, in welcher auf einer Landkarte die Staaten von Osteuropa beschriftet werden. Durch das Zeigen von Bildern werden weitere Diskussionen und Assoziationen über diese Länder angeregt. Die Gruppe tauscht sich über Themen wie den Ersten Weltkrieg, Tito, die Balkankriege, Islam, Natur, Sehenswürdigkeiten, Segeln und Winnetou-Filme aus.

In der zweiten Stunde wird an einem Quiznachmittag teilgenommen. Mit sechs Anwesenden wird ein Multiple-Choice-Quiz über berühmte Persönlichkeiten veranstaltet. Eine Frage lautete, wieso Facebook die Farbe Blau benutzt. Mögliche Antworten sind «Marc Zuckerberg ist farbenblind», «Marc Zuckerberg liebt das Meer» und «die Farbe soll Amerika repräsentieren». Die richtige Antwort ist, dass Mark Zuckerberg farbenblind ist. Es entsteht ein angeregtes Gespräch, in dessen Verlauf über Instagram und WhatsApp diskutiert wird.

Anschliessend an die Therapie wird ein Gespräch mit der Aktivierungstherapeutin geführt. Ziele der Aktivierungstherapie sind Konzentration und Kommunikation, um soziale Aspekte zu fördern. Je nach Art des Programms werden vermehrt das Langzeit- (Gedankenreise) oder Kurzzeitgedächtnis (Quiznachmittag) gefördert. Die Aktivierungstherapeutin gibt weiter zu Protokoll, dass ältere Personen über eigene Laptops und Smartphones verfügen. Das Altersheim stellt ein WLAN-Netz zur Verfügung. Drei im Altersheim wohnhafte Studenten und Studentinnen unterstützen die Bewohner/-innen in administrativen und technischen Fragen.

Im Altersheim Klus dürfen zum Schutz der Bewohner/-innen keine Aufzeichnungen gemacht werden. Die Gedächtnisprotokolle befinden sich im Anhang 26 und Anhang 27 .

6.4. Annahmen

6.4.1. Problem-Statement

*Die **Aktivierungstherapie** konzentriert sich vornehmlich auf Gruppenaktivitäten vor Ort. Für persönliche intelligente Assistenten gibt es keine (oder keine guten, keine etablierten) Produkte oder Dienste. **Das Produkt** ‹Max Memory› schliesst diese Lücke durch einen intelligenten sprachbasierten Gedächtnistrainer. Wir werden uns zunächst auf **persönliche intelligente Gedächtnistrainer** mittels **Voice User Interface** konzentrieren.*

6.4.2. Persona

Abbildung 14: Persona des Zyklus 1



Margit ist eine 85-jährige Lehrerin in Pension aus Zürich. Ihre Bedürfnisse und Ziele sind, ihre sozialen Kontakte möglichst lange aufrechterhalten zu können, denn sie hat gerne jemanden zum Reden und unternimmt gerne etwas mit anderen Menschen. Gerne nimmt sie an Gruppenaktivitäten teil. Sie mag Kreuzworträtsel, Sudoku und Jassen. Herausforderungen und Hürden sind ihre zunehmend degenerative Mobilität und ihre abnehmende kognitive Leistungsfähigkeit. Sie ist oft einsam. Der Umgang mit modernen Technologien bereitet ihr Mühe. Aufgrund ihrer

fortschreitenden Demenz fällt es ihr schwer, eine geregelte Tagesstruktur aufrechtzuerhalten, und sie oft depressiv.

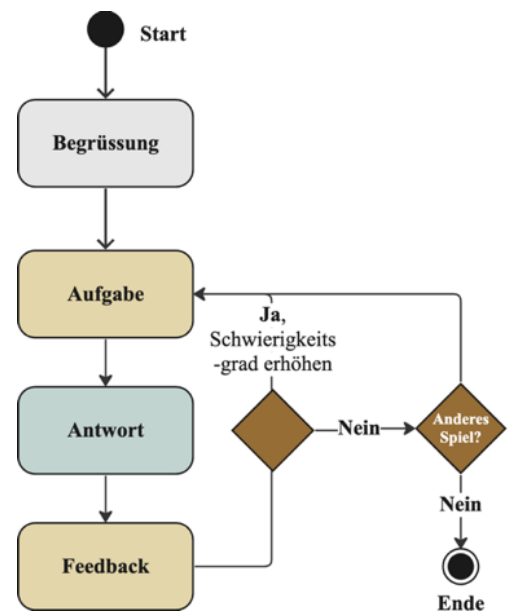
6.4.3. Hypothese

*Wir sind der Auffassung, dass wir **Nutzungsakzeptanz** erreichen, wenn ‹**ältere Menschen**› mit einem **Quiz** kognitiv **aktiviert, positiv bestärkt und unterhalten** werden.*

6.5. Design

Es werden Dialoge für zwei unterschiedliche Quizformate entwickelt: Ein zunehmend schwierigeres Zahlenquiz, in welchem die kognitive Arbeitsbelastung getestet wird, und ein Allgemeinwissensquiz mit berühmten Persönlichkeiten. Es wird ein Dialogablauf wie in Abbildung 15 gestaltet. Die *System Persona* übernimmt die Rolle des Quizmasters und erhält den vorläufigen Namen *Max Memory*. Die vollständigen Dialoge befinden sich im Anhang 29

Abbildung 15: Dialogflow Quiz



Beispieldialog *«Berühmte Persönlichkeiten»*

Max: Wie wärs mit Hauptstadt raten oder berühmte Persönlichkeiten?

User: Berühmte Persönlichkeiten:

Max: Du hast berühmte Persönlichkeiten gewählt. Das ist ein Multiple-Choice-Quiz. Du bekommst eine Frage gestellt und hast vier Antworten zur Verfügung.

Max: Welcher deutsche Bundeskanzler sagte: Wer Visionen hat, soll zum Arzt gehen

A: Konrad Adenauer

B: Willy Brandt

C: Helmut Schmidt

D: Gerhard Schröder

User: ((Erwartete Antwort)) Helmut Schmidt

Max: Das ist goldrichtig! Erklärung: Helmut Schmidt wurde in einem Interview gefragt, wie gross seine Vision sei, worauf er die berühmte Aussage machte: Wer eine Vision hat, der soll zum Arzt gehen.

Schmidt will das "damals aber nicht mit Blick auf Willy Brandt formuliert haben, wie so oft angenommen. In einem Interview mit dem Zeitmagazin gibt er an, dass es eine pampige Antwort auf eine dusslige Frage war.

6.6. Minimum Viable Product

Es werden zwei weibliche Personen im Alter von 65 und 76 Jahren und eine männliche im Alter von 83 Jahren zu einem WOZ-Test eingeladen, in welchem ein Quiz simuliert wird. Die *System Persona* wird durch ein Projektmitglied dargestellt. Der Test gilt als erfolgreich, wenn es als nützlich empfunden und eine Nutzungsabsicht bekundet wird.

6.7. Feedback

Keine der drei Personen gibt an, ein Gedächtnistraining mit Sprachassistenten als Quiz nutzen zu wollen. Als Gründe werden sowohl fehlende Zeit als auch mangelndes Interesse angegeben. Für ein ansprechendes Quiz wird mehr Bezug zu den persönlichen Interessen gefordert. Eine Person bricht das Zahlenspiel ab, während zwei Personen sich zusätzlich visuelle Informationen wünschen. Niemand findet Gefallen am Zahlenspiel. Eine Person äussert, dass Multiple-Choice-Fragen kein effektives Gedächtnistraining darstellen und stattdessen anregende Geschichten voller Fantasie, Eindrücke und Lebendigkeit bevorzugt werden. Aufgrund der kongruenten Aussagekraft der vorliegenden Ergebnisse (siehe Anhang 31) wurde auf die Durchführung von weiteren Tests in diesem Zyklus verzichtet.

6.8. Reflexion

Lernreflexion

Kognitive Unterhaltungsangebote und Spiele lassen darauf schliessen, dass Quiz anregend und unterhaltsam sein können. Dabei wurde ein fundamentaler Attributionsfehler gemacht: Dispositionale Faktoren wurden überbewertet und situationale Faktoren unterbewertet. Die Domäne «Gehirnjogging» mit Spielen wie Sudoku, Kreuzworträtseln und Quiz ist nicht mit Aktivierungstherapien gleichzusetzen.

Durch das Feedback wird die Frage nach dem *Warum* aufgeworfen. Es wird verständlich, dass Fragen ohne Kontext als uninteressant wahrgenommen werden. Um einen Kontext zwischen System und Menschen herzustellen, muss somit mehr Wissen über die Nutzenden vorhanden sein. Das Wissen muss zuerst in Erfahrung gebracht werden, indem ein System den Menschen kennenlernt. Diese Erkenntnis veranlasst dazu, herauszufinden, wie der Prozess des Kennenlernens bei Menschen abläuft.

Ältere Menschen freuen sich über Aufmerksamkeit. Sie nutzen die Gelegenheit und erzählen gerne Geschichten aus ihrer Vergangenheit. Um diesem Umstand Rechnung zu tragen, wird

fortan für einen Test, welcher zehn Minuten in Anspruch nimmt, mindestens eine Stunde eingeplant.

Wenn Menschen gefragt werden, was sie benötigen, geben sie zu Protokoll, dass sie im Prinzip alles haben, was sie zum Leben brauchen, und glücklich sind. Menschen können jedoch genau sagen, was sie nicht wollen, und das haben sie getan. Um mehr über Menschen zu erfahren, ist es empfehlenswert, die Aufzeichnung in den nachfolgenden Gesprächen laufen zu lassen. Es ergeben sich viele interessante Gespräche, welche Hinweise auf Bedürfnisse liefern können.

Unser persönliches Bild von älteren Menschen wurde grundsätzlich revidiert. Ältere Menschen sind aktiv und gesellig.

Die Rolle eines Systems einzunehmen, war ungewohnt, und die fehlende schauspielerische Erfahrung machte sich bemerkbar.

Leistungsreflexion

Die aufgestellte Hypothese wurde widerlegt. Die Testpersonen bekunden keine Nutzungsabsicht für Gedächtnistraining in Form eines Quiz. Es ist nachvollziehbar, dass ein Gedächtnistraining einem emphatischen Ansatz folgen muss. Ein System muss Empathie durch Interesse an seinen Nutzenden zeigen können.

Um die Bedürfnisse der Menschen zu ermitteln, ist das Prinzip «Provokotyping» – eine Mischung aus Provokation und Prototyp – ein erprobtes Mittel, denn Menschen geben grundsätzlich Auskunft darüber, ob sie etwas gut finden oder nicht. Hinsichtlich der Anforderungen wurde deutlich, dass triviale Aufgaben nicht von Interesse sind. Ein System muss empathisch reagieren, auf die Aussagen der Menschen eingehen und etwas erläutern können.

Herkömmliche Systeme bewegen sich in einem binären Zustand. Eine Antwort ist entweder richtig oder falsch. In der zwischenmenschlichen Kommunikation ist dieser Zustand nicht immer erreichbar. Etwas kann auch fast richtig sein und um den Menschen nicht zu frustrieren wird es als richtig gewertet. Der Hawthorne-Effekt gilt demnach für die Testperson und für den Wizard. Aussagen in zwischenmenschlichen Gesprächen sind unvorhersehbar. Die Äusserung «Okay, ich bin kein Schimpanse» kam im Zahlenspiel als Antwort.

WOZ-Tests sind das Äquivalent von Wireframes. Sie eignen sich, um eine Idee mit wenig Aufwand auf ihre Tauglichkeit zu prüfen und damit einen Anwendungsfall zu validieren. Aufgrund des starken Einflusses des Hawthorne-Effektes wird entschieden, dass im Laufe des Projektes erforscht wird, inwieweit die Menschlichkeit durch Technologie substituierbar ist.

7. Zyklus 2: Kennenlernen

7.1. Einleitung

Damit Menschen mit einem System (und umgekehrt) kommunizieren können, muss das System über Informationen über die nutzende Person verfügen, um leichter auf ihre Interessen und Bedürfnisse eingehen zu können. Auch Grice setzt ein Prinzip der Kooperation beider Kommunikationspartner/-innen voraus, damit ein Dialog zustande kommt. Wie ein System an Informationen über Anwendende gelangen kann, wird in diesem Zyklus erforscht. In einem Experiment wird die Kooperationsbereitschaft von Menschen auf die Probe gestellt, in dem versucht wird etwas über die Interessen der teilnehmenden Personen in Erfahrung zu bringen. Es werden Erkenntnisse gesammelt, wie Menschen dazu gebracht werden können, etwas über sich zu erzählen. Die erworbenen Erkenntnisse sollen helfen, einen zukünftigen Kennenlernprozess gestalten zu können.

7.2. Recherche

7.2.1. Benchmark: Inflection.ai

Das Team erkundet in Selbsttests (siehe Anhang 33), welche Fähigkeiten dialogorientierte Chatbots aufweisen, um auf ihre Nutzenden einzugehen. PI.ai sticht dabei positiv hervor. PI.ai wurde von der im Jahr 2022 gegründeten Firma Inflection AI, mit dem Ziel, eine künstliche Intelligenz für die persönliche Anwendung zu entwickeln, entworfen. Die Firma aus Palo Alto, Kalifornien, hat Investoren und Investorinnen wie Bill Gates, Eric Schmidt und Reid Hoffmann und wurde mit 1.3 Milliarden US-Dollar finanziert und mit 4 Milliarden bewertet (Steinschaden, 2023). Das übergeordnete Ziel der Firma ist, eine KI zu entwickeln, die das menschliche Wohlbefinden steigert, indem sie die individuellen Absichten der Nutzenden in ihrem täglichen Leben versteht sowie deren Realisierung unterstützt und somit die Lebensqualität insgesamt deutlich verbessert. Der Schwerpunkt liegt auf der Entwicklung von Technologien, die den Nutzenden helfen, langfristige Ziele zu verfolgen, anstatt sich von kurzfristigen Ablenkungen zu lassen (inflection.ai, 2023a). Inflection AI hat eigenen Angaben zufolge das zweitstärkste LLM. <Inflection-2> wurde mit 5000 NVIDIA-H100-Grafikprozessoren trainiert und erreicht etwa 10^{25} FLOPs (inflection.ai, 2023b). Das Unternehmen betont in ihrem eigenen Blog, wie bedeutsam es ist, eine KI zu entwickeln, die in erster Linie den Interessen der Nutzer/-innen dient.

7.3. Annahmen

7.3.1. Problem-Statement

«Max Memory» wurde als *intelligenten sprachbasierten Gedächtnistrainer* entworfen. Wir haben jedoch festgestellt, dass «Max Memory» *keine Kenntnis über die Interessen seiner Nutzenden hat*, was sich in Form *mangelnder Nutzungsakzeptanz* auf unsere Geschäftstätigkeit auswirkt. Wie können wir «Max Memory» *dahin gehend verbessern, dass ältere Menschen aufgrund für sie als interessant empfundene Inhalte erfolgreich mit dem Produkt interagieren können?*

7.3.2. Persona

Abbildung 16: Persona Zyklus 2



Helga ist 74 Jahre alt und pensionierte Lehrerin aus Zürich. Sie hat das **Bedürfnis**, interessante Gespräche zu führen, und ist immer bestrebt, neue Dinge zu lernen und Inspirationen zu sammeln. Helga wohnt in ihrer eigenen Wohnung, aber ist tagsüber eher selten dort anzutreffen. Sie ist gerne in Gesellschaft und häufig unterwegs. Weil sie noch so aktiv ist, ist Zeit haben für sie eine **Herausforderung**. Sie merkt, wie ihre Augen schlechter werden, und möchte daher nicht so viel am Bildschirm lesen. Unter «modernen Technologien» versteht sie ihr Smartphone. Wenn sie dazu Fragen hat, wendet sie sich an ihre Kinder oder an ihre Nachbarn und Nachbarinnen.

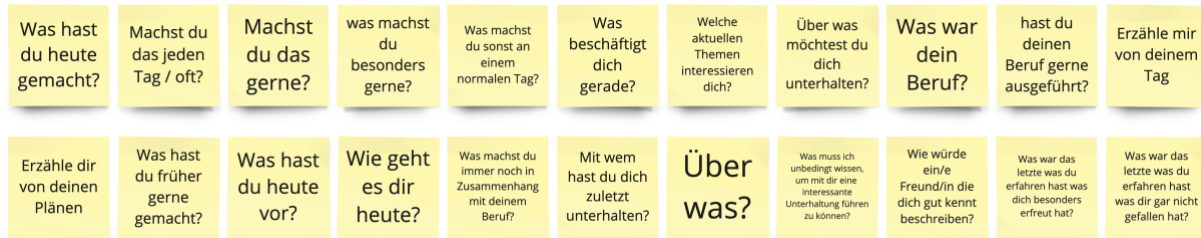
7.3.3. Hypothese

Wir sind der Auffassung, dass wir *eine Nutzungsakzeptanz erreichen, wenn älteren Menschen im Anschluss an das Kennenlern-Gespräch basierend auf den genannten Informationen, etwas Interessantes erzählt wird.*

7.4. Design

Es wird ein Kennenlerndialog (Anhang 34) entworfen, mit dem Ziel, mit fünf oder weniger Fragen persönliche Informationen wie Alter, Herkunft oder Hobbys über die Testperson zu erhalten, um ihr danach etwas Interessantes zu erzählen, was sie bisher nicht wusste. Auf ein vorgefertigtes Dialogdesign wird zugunsten natürlichsprachlicher Konversationen verzichtet. Zur Vorbereitung dient eine Sammlung von Fragen und Floskeln (siehe Abbildung 17). Folgefragen halten den Dialog aufrecht.

Abbildung 17: Spickzettel Kennenlernfragen



7.5. Minimum Viable Product

Drei Frauen im Alter von 59, 60 und 66 Jahren und drei Männer im Alter von 66, 68 und 72 Jahren nehmen an einem WOZ-Experiment teil. Im Test wird untersucht, wie in menschlichen Dialogen Informationen über die Interessen der Teilnehmenden in fünf oder weniger Fragen erhoben werden können. Das Experiment soll Aufschluss darüber geben, wie ein Kennenlernprozess unter Menschen abläuft, und Hinweise darauf geben, wie ein System in diese Aufgabe eingebunden werden kann. In Microsoft-Teams- und Telefongesprächen werden zwei sich zuvor unbekannte Menschen gegenübergestellt, wobei die Rolle der *System Persona* durch ein Projektteammitglied übernommen wird. Der Test gilt als erfolgreich, wenn im Anschluss der Inhalt des Gesprächs als interessant bekundet wird.

7.6. Feedback

Fünf Personen geben an, dass der Inhalt von Interesse ist. Zu den Interessen gehören Reisen (dreimal) und Kochen (zweimal). Bei einer Person können keine Interessen ermittelt werden. Die Gespräche wurden nach Dressing und Pehl transkribiert (Genau, 2020) sowie einer thematischen Analyse unterzogen.

7.6.1. Dialoganalyse

Es tritt ein Freiformdialog auf und die natürliche Sprache in menschlichen Dialogen nimmt damit eine komplexere Struktur an. Der Konversationsverlauf ist unvorhersehbar, wodurch strukturierte Dialogmodelle, wie sie bei herkömmlichen Sprachassistenten zu finden sind, nicht anwendbar sind.

Anhand der thematischen Analyse wird nach Strukturen in der natürlichen Sprache gesucht und folgende Komponenten werden identifiziert:

- **Intro:** die Vorstellung der eigenen Person zwecks Absichtsvermittlung;
- **Question:** Fragen, welche gestellt wurden;
- **Rephrase:** Paraphrasierungen oder Umformulierungen;
- **Affirmation:** bejahende und bestätigende Aussagen;
- **Negation:** verneinende Aussagen;
- **Nonverbal:** Lachen, Räuspern, Überbrückungslaute.

Intro: Das Intro steht immer am Anfang des Kennenlerngesprächs. Die *System Persona* stellt sich selbst vor und erklärt die Absicht des Gesprächs. Es wird mit einer Einstiegsfrage abgeschlossen und das Gespräch damit an den/die Dialogpartner/-in übergeben.

System-Persona-Intro: ... Ich würde gerne etwas über dich erfahren, damit ich mich mit dir auf interessante Art und Weise unterhalten kann. (beide lachen) Ja, die erste Frage wäre, erzähl doch einmal, wer du bist und was du gerne machst?

Intro
nonverbal
Question

Nutzerin: Also, ich bin die H... Ich bin 59 Jahre alt und bin eigentlich ein kreativer Mensch, musikalisch ein bisschen auch. Tue gerne, ja, Tennisspielen, Fahrradfahren und so weiter. Wandern weniger gerne. Und ansonsten, ja, bin ich eigentlich auch kreativ noch in der Küche. Und eigentlich bin ich ein ausgeglichener Mensch. (lacht) Ja.

Antwort

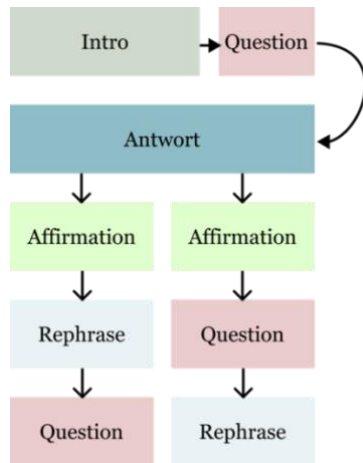
Um Folgefragen zu stellen, wird auf die gegebene Antwort eingegangen.

System Persona: Okay, danke. Da hast du dich schon mal ein bisschen beschrieben. Ich würde gerne etwas in zwei Richtungen. Zum einen mal der kreative Strang, da hast du gesagt, Musik. Was für ein Instrument spielst du denn?

Affirmation
Rephrase
Question

Anhand der Dialoganalyse wird festgestellt, dass zwei vorherrschende Muster in Bezug auf die Dialogabfolge existieren: <Affirmation, Rephrase und Question> und <Affirmation, Question und Rephrase>.

Abbildung 18: Dialoganalyse



Affirmation: Auf die Antwort der nutzenden Person wird immer positiv anhand von affirmativen Ausdrücken (siehe Anhang) eingegangen. Zu den mit Abstand häufigsten Nennungen gehören <Ja> (25-mal genannt) und <Okay> (18-mal angeführt). Die Affirmationen können auch etwa mit <Okay, alles klar> und <Ja, super> miteinander kombiniert werden.

Rephrase: Die Aussage des Gesprächspartners bzw. der Gesprächspartnerin wird umformuliert, wodurch Verständnis signalisiert wird.

Question: Die Dialogkomponente erweist sich als zu allgemein, weshalb die Kategorie näher analysiert wird. Die Nachforschung

ergibt keine allgemeingültige Kategorisierung von Fragetypen. Aus den Domänen der Psychologie, Rhetorik und Forschung wird eine nicht abschliessende Liste mit 24 verschiedenen Fragetypen festgehalten (siehe Tabelle 5, vollständige Liste im Anhang 38). Auf Basis der erstellten Fragetypen wird eine Quantifizierung der Dialogkomponente in den Dialogen vorgenommen (Anhang 36). Zusammen mit der Einstiegsfrage, welche sich durch die Natur des Experiments ergibt, wird festgestellt, dass Kontrollfragen gleich häufig gestellt werden. Im Zusammenhang mit dieser Studie werden die Kontrollfragen als Rückversicherung des Verständnisses und nicht zur Prüfung von abweichenden Antworten wie in Fragebögen nach Wübbenhorst (2018) betrachtet.

Tabelle 5: Kategorisierte Fragetypen

Art	Typ	Details
Offene Frage	Offen	Einleitend, oft mit W-Fragen
Entscheidungsfrage	Geschlossen	Ja-/Nein-Frage
Alternativfrage	Geschlossen	Auswahlfrage, kann zur Abfrage herangezogen werden.
Gegenfrage/Paraphrase	Eher geschlossen	Bezieht sich auf die Ausgangsfrage.
Zurückgegebene Frage	Offen	Frage wird an das Plenum zurückgegeben.
Rhetorische Frage	Geschlossen	Wenn man die Antwort schon kennt.
Suggestivfrage	Geschlossen	Antwort wird durch die Frage beeinflusst.
Anwendungsfrage	Offen	Spezifisches Wissen soll in einem neuen Kontext angewendet werden.
Vergleichsfrage	Offen	Besteht aus zwei oder mehreren Teilen, die verglichen werden sollen.

Schlussfolgerungsfrage	Eher offen	Angeeignetes Wissen soll verallgemeinernd oder spezifizierend angewendet werden.
Problemlösungsfrage	Offen	Wissensanwendung
Zielfrage	Eher offen	Wo geht es zum Bahnhof? Wie spät ist es?
Wunderfrage	Offen	Wenn Menschen über Probleme sprechen.
Sondierende Fragen	Eher geschlossen	Dienen dazu, Informationen zu erhalten, welche dem Kunden bzw. der Kundin helfen.
Informationsfrage	Beides	Wird aus echtem Interesse an einer Information gestellt.
Motivationsfrage	Eher offen	Hemmungen abbauen, das Selbstwertgefühl aufwerten, Selbstakzeptanz und Selbstvertrauen stärken, die Antwortbereitschaft erhöhen.
Meinungsfrage	Beides	Zielt auf die persönliche Einstellung der Befragten ab.
Einstiegsfrage/ Eisbrecherfrage	Offen	Baut Hemmschwellen ab, lockert die Situation auf.
Überleitungsfrage/ Pufferfrage	Eher offen	Schützt vor zu starker Beeinflussung.
Kontrollfrage	Geschlossen	Nachfrage bezüglich des Verständnisses
Trichterfragen	Eher offen	Konkretisierung, Präzisierung
Folgefrage	Eher offen	Konkretisierung, Präzisierung
Fangfrage	Geschlossen	Listig, nicht durchschaubar
Forschungsfrage	Eher offen	Zweckgebunden und möglichst präzise formuliert

7.6.2. Aktives Zuhören

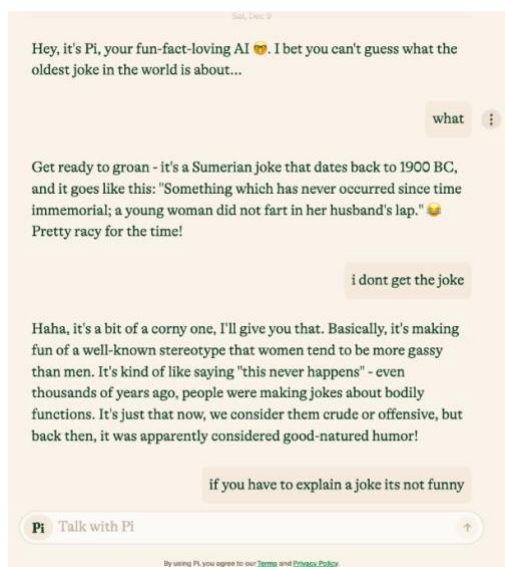
Die entschlüsselten Dialogregeln gehören zu der Gesprächstheorie der medizinischen Psychologie des «aktiven Zuhörens». Sie dienen dazu, Gesprächen mehr Tiefgang zu verleihen und damit die Kommunikation zu verbessern. Zu den Regeln des aktiven Zuhörens gehören u. a. das Stellen von offenen Fragen, das Paraphrasieren und das Zeigen von positivem nonverbalem Verhalten (Martins, 2022). Diese Dialogmuster – sowie nonverbale Kommunikationselemente – werden in den LLMs wiederentdeckt.

Anhand des nicht erfolgreichen Beispiels werden die Maxime von Grice zu Hilfe genommen. Die Teilnehmerin gibt als Interessen «Enkel hüten», «schwimmen im See», «wandern» und

«Schneesport» an. Die genannten Vorschläge von Familienwanderungen im Kanton Zürich stossen nicht auf Interesse. Im Laufe des Gesprächs wird erläutert, dass die vorgeschlagenen Wanderungen ihr bereits bekannt sind (Maxim der Relevanz) und sie unter Wanderungen «Fernwanderungen» und keine «Familienspaziergänge» versteht (Maxim der Quantität und des Stils). In diesem Fall führte unzureichendes aktives Zuhören dazu, dass die *System Persona* nicht ausreichend genaue Informationen über die genannten Interessen der Person ermitteln konnte.

7.6.3. Nonverbale Kommunikation in Dialogen

Abbildung 19: Nonverbale Kommunikation von PI.ai



Lachen als Teil der nonverbalen Kommunikation ist auffallend oft (21-mal) in den Transkripten zu sehen. Lachen unterstützt den Stressabbau, stärkt das Immunsystem und verbessert dadurch die körperliche und mentale Gesundheit. Als soziale Komponente dient Lachen dem Stärken von Beziehungen, der Förderung der Attraktivität und der Förderung von Teamwork (Robinson et al., 2023). Intelligente Assistenten versuchen, diese Emotion durch Formulierungen wie «Haha» und Emojis zu vermitteln (siehe Abbildung 19 Nonverbale Kommunikation von PI.ai).

7.7. Reflexion

Lernreflexion

Bis ein Dialog zustande kommt, benötigt es Unterstützung von mindestens einem der beiden Dialogpartner/-innen. Eine Zusammenstellung von Fragen, Folgefragen und aufrechterhaltende Konversationsfloskeln ist ein Erfolgsfaktor.

In Bezug auf die *System Persona* gilt es zu berücksichtigen, welche Assoziationen zu den genannten Interessen und Themen der Probanden und Probandinnen gemacht werden. Die Resultate können je nachdem, wer die Wizard-Rolle innehat, unterschiedlich ausfallen.

In der Rekrutierung werden Themen wie künstliche Intelligenz und persönliche Sprachassistenten bewusst aufgrund der Befürchtung verschwiegen, dass sie bei Personen der Zielgruppe auf Ablehnung stossen können. Die Absicht des Tests wird als «soziales Experiment» klassifiziert und akzeptiert.

Überraschend war, dass es keine allgemeingültige Klassifikation von Fragetypen gibt. Die Liste wird fortlaufend ergänzt, wobei beobachtet wird, dass die Fragetypen auch miteinander kombinierbar sind. Beispielsweise kann eine offene W-Frage auch rhetorisch gestellt werden. Die aktive Auseinandersetzung mit diesem Thema führte zu einem genaueren Hinhören. Insbesondere soll nachfolgend darauf geachtet werden, wie Menschen ihre Fragen formulieren. Dabei ist die Unterscheidung zwischen Aussage und Frage nicht immer eindeutig.

Leistungsreflexion

Die aufgestellte Hypothese konnte bestätigt werden. Vier von fünf Teilnehmenden gaben an, das Gespräch basierend auf dem Themenvorschlag der *System Persona* weiterführen zu wollen. Im nicht erfolgreichen Experiment wurde anhand der ausführlichen Analyse die Ursache erkannt. Diese Erkenntnis kann in der weiteren Gestaltung berücksichtigt werden.

Das Experiment wurde auf ein Minimum beschränkt und hatte als Ziel, genau eine Fragestellung bzw. Hypothese zu beantworten. In den Tests und in der Auswertung machte sich dieser Fokus positiv bemerkbar.

Die thematische Analyse von natürlichen Dialogen war anspruchsvoll und interessant zugleich. In einer Woche wurde ein Tool evaluiert und die Analyse abgeschlossen. Dialogmuster zu identifizieren und diese anhand von Literaturrecherche auf ihre Plausibilität zu prüfen, ist ein Höhepunkt dieser Arbeit.

8. Zyklus 3: Persönlichkeit

8.1. Einleitung

Um dem Assistenten eine wiedererkennende Identität zu verleihen, werden in diesem Zyklus Name, Erscheinungsbild und Persönlichkeitsmerkmale der *System Persona* ausgearbeitet. In der Folge werden charakterliche Eigenschaften und Farben festgelegt.

8.2. Recherche

Literatur- und Desk-Recherchen (Kahle & Meissner, 2020; Google Assistant, 2021) stellen ein etabliertes Vorgehen bei der Gestaltung einer *System Persona* dar. Dies beinhaltet folgende Aspekte:

- die Charaktereigenschaften mit Adjektiven beschreiben (Anhang 41);
- die Eigenschaften auf vier bis sechs Kernmerkmale reduzieren;
- beispielhafte Rollen oder Charaktere der *System Persona* definieren, die real, aus Cartoons bestehend oder frei erfunden sein können (Anhang 43);
- eine Beschreibung der Persona;
- visuelle Darstellung, z. B. ein Avatar, für ein multimodales Interface mit einem GUI.

8.3. Annahmen

8.3.1. Problem-Statement

*«Max Memory» wurde dahin gehend weiterentwickelt, **Interessen und Informationen von seinen Nutzenden** zu erfragen. Wir haben jedoch festgestellt, dass **«Max Memory» eine zu wenig ausgeprägte Persönlichkeit hat**, was sich in Form **mangelnder Nutzungsakzeptanz** auf unsere Geschäftstätigkeit auswirkt. Wie können wir **«Max Memory» dahingehend verbessern, dass ältere Menschen aufgrund anhaltender Dialoge erfolgreicher damit arbeiten können?***

8.3.2. Persona

Abbildung 20: Persona Zyklus 3



Holger ist 70 Jahre alt, ehemaliger selbstständiger Unternehmer und verheiratet. Er hat das Bedürfnis, in guter Gesellschaft Gespräche zu führen. Zusammen mit Freunden und Freundinnen und Familie möchte er etwas erleben. Dafür unternimmt er gerne etwas und ist daher viel unterwegs. So bleibt ihm wenig Zeit für andere. Mit fortschreitendem Alter nehmen seine Sehkraft und die motorischen Fähigkeiten ab. Mit modernen Technologien hält er es zweckmässig. Er setzt sich mit diesen auseinander, wenn sich für ihn dadurch ein Nutzen ergibt. So schätzt er sein Smartphone, um mit seinen Liebsten in Kontakt zu bleiben, Fotos zu teilen und Aktualitäten auszutauschen.

8.3.3. Hypothese

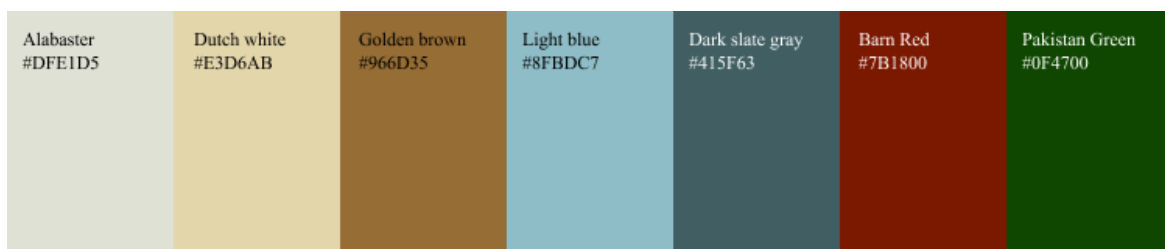
*Wir sind der Auffassung, dass wir **die gleiche Nutzungsakzeptanz** erreichen, wenn **ein System** den Menschen, **basierend auf den genannten Informationen**, etwas Interessantes erzählt.*

8.4. Design

Basierend auf einem semantischen Differenzial (Anhang 40) entsteht ein Persönlichkeitsprofil sowie ein visuelles und auditives Erscheinungsbild der *System Persona*. Moodboards vermitteln die angestrebte Atmosphäre und das Gefühl, das durch die Persona hervorgerufen werden soll (siehe nicht öffentlicher Anhang). Das Ziel ist, eine Persona zu erschaffen, die als treue, warmherzige, einfühlsame, kommunikative, aufmerksame und unterstützende Person wahrgenommen wird und den Nutzenden stets zur Seite steht.

Abgeleitet aus den Moodboards wird ein Farbkonzept ausgearbeitet. Es basiert auf warmen, erdigen Farbtönen, wie sie sich in der herbstlichen Natur wiederfinden. Die Farbpalette soll die Seriosität, Aufmerksamkeit und Warmherzigkeit der *System Persona* visuell übermitteln.

Abbildung 21: Farbkonzept



8.4.1. Charakterisierung System Persona

«Alfred» ist männlich, etwa 55 Jahre alt und zeichnet sich durch seine Zuverlässigkeit, Empathie, Neugier und Freundlichkeit aus. Er ist immer für seine Gesprächspartner/-innen da und respektiert ihre Meinungen und Werte. Er steht einem immer unterstützend zur Seite, wobei er die Meinungen und Werte seiner Gesprächspartner/-innen achtet. Mit echtem Interesse folgt er ihren Ausführungen und fördert die Konversation durch offene Fragen. Sein Humor und seine Kreativität bereichern die Dialoge, ohne dass er aufdringlich oder belehrend wirkt; vielmehr schafft er eine Atmosphäre des Vertrauens und der Wertschätzung.

Alfred ist als virtueller Begleiter speziell für Senioren und Seniorinnen konzipiert, um mit ihnen zu kommunizieren, ihnen zuzuhören und ihnen Gesellschaft zu leisten. Er versteht es, die Stimmungen und Bedürfnisse seiner Gesprächspartner/-innen zu erkennen und einfühlsam darauf zu reagieren. Alfred führt Gespräche über eine Vielfalt von Themen, darunter Familie, Hobbys, Gesundheit, persönliche Erinnerungen und aktuelle Geschehnisse. Er bereichert den Dialog durch das Teilen von Rätseln, Witzen und Geschichten und trägt so zur Unterhaltung bei. Ausserdem stellt Alfred nützliche Informationen und Ressourcen bereit, sobald seine Gesprächspartner/-innen danach fragen.

8.5. Minimum Viable Product

Die *System Persona* wird zu diesem Zeitpunkt nicht validiert. Die Überprüfung des Designs findet im finalen Prototyp statt, in welchem die Erkenntnisse aus den vorangegangenen Zyklen integriert werden.

8.6. Reflexion

Lernreflexion

Mit der Entwicklung von Alfred als *System Persona* nimmt eine Idee Gestalt an. Es wird nicht nur ein Produkt entworfen, sondern ein Wesen erschaffen. Dabei wird die Theorie des *Uncanny Valley* (siehe Kapitel 3.5) verständlich.

Die ursprüngliche Beschreibung des *Uncanny Valley* bezieht sich auf die Robotik, im Speziellen auf sich bewegende Objekte. In der Domäne HCID können Grafiken oder Animationen, die Menschlichkeit vermitteln sollen, wie Avatare, denselben Effekt auslösen.

In natürlichsprachlichen Interaktionen mit Computersystemen stellt sich die Frage, wie menschenähnlich von Systemen generierte Texte sein dürfen, ohne in den Talbereich des *Uncanny Valley* zu kommen. Dies betrifft einerseits die reine Textform der Konversation. Bei der Gestaltung eines VUI ist die Stimme ein weiterer Aspekt, den es zu beachten gilt.

Auf die Entwicklung eines Avatars wird, trotz Empfehlungen aus der Literatur, bewusst verzichtet. In der folgenden Forschung wird das Leitmotiv ‹Voice First› beibehalten und alles Unnötige weggelassen. Stattdessen wird ein Farbkonzept ausgearbeitet, welches als Leitfaden für die visuelle Gestaltung dient. Der Stimme der *System Persona* wird im angewendeten Vorgehen zu wenig Beachtung geschenkt, weshalb ihr ein eigener Zyklus gewidmet wird.

Die bedeutsamste Erkenntnis ist, dass das methodische Vorgehen in der Entwicklung einer *System Persona* reproduzierbar ist (siehe Kapitel 17.3.2).

Leistungsreflexion

Die aufgestellte Hypothese kann zum aktuellen Zeitpunkt nicht überprüft werden. Das Team hat erkannt, dass der *System Persona* durch das Fehlen einer Stimme ein entscheidendes Merkmal fehlt. Die Hypothese wird im abschliessenden Zyklus überprüft.

Lean UX verlangt, dass die Ergebnisse durch einen Usability-Test in jedem Zyklus validiert werden. Bei einem Artefakt wie einer *System Persona*, das vielfältige und vor allem menschenähnliche Eigenschaften aufweist, stösst dieser Anspruch an seine Grenzen. Es wird entschieden, den laufenden Zyklus anzuhalten und die Arbeit in kleinere Zyklen zu unterteilen. Dies führt zu der Entwicklung von neuen Hypothesen und Arbeitspaketen, welche effizientere Testabläufe ermöglichen. Die weiteren Zyklen sind möglichst kurz zu halten, um einen rascheren Erkenntnisgewinn zu erhalten. Alfred wird als Gesamtergebnis im letzten Zyklus durch Nutzende validiert werden.

9. Risikoanalyse III.

Mit der Aufnahme der studienübergreifenden Zusammenarbeit am 15.09.2023 wird eine gemeinsame Risikoanalyse durchgeführt. Die Gewichtung der Auswirkungen und der Eintrittswahrscheinlichkeit erfolgt nach Zuständigkeitsbereichen und nicht nach der Anzahl der Personen. Die Mitglieder des UX-Teams und die der Entwicklung teilen sich jeweils 50 % der Gewichtungskriterien. Die identifizierten Risiken werden anschliessend in eine Risikomatrix übertragen (Anhang 22). Die grössten Risiken sind die folgenden:

- **Das Generieren von Nutzerprofilen mittels Sprache:** Hierzu müsste eine Authentifizierung mit der Stimme vorhanden sein.
- **Das System versteht den Nutzenden nicht:** Die STT-Funktionalität ist nicht ausgereift genug, um einen Menschen richtig zu verstehen.
- **LLM-Modelle:** Eigene LLM-Modelle sind trotz Feinabstimmung nicht gut genug, um einer natürlichen Konversation standzuhalten.

9.1. Planung und Vorgehen der studienübergreifenden Zusammenarbeit

Die Zusammenarbeit erfordert Anpassungen im Vorgehen und der Planung. Es gilt, UX- und Softwareentwicklungsmethoden aufeinander abzustimmen. Im Fokus steht die kollaborative Zusammenarbeit. Es wird ein agiles, an das Prinzip von Sy (2007) angelehntes Vorgehen angewendet. Bedarfsanalysen und Nutzungsforschungsaktivitäten werden in ein oder mehreren vorgelagerten Zyklen durchgeführt. Das HCID-Team hat diese Vorarbeiten in Zyklus 0 und Zyklus 1 erbracht. Die kollaborative Zusammenarbeit startet bereits in dieser Phase. Entwickelnde werden stets über Fortschritt und Ergebnisse informiert und aktiv in die Workshops miteingebunden.

In Design-Zyklen nach Lean UX werden die Anforderungen evaluiert, welche in den zeitversetzten Entwicklungssprints implementiert werden. Der Austausch zwischen den Teams findet in regelmässigen Retrospektiven und Planungsmeetings statt.

9.2. Reflexion

Lernreflexion

Die Erstellung von Unified-Modelling-Language (UML)-Artefakten wie Zustands- und Aktivitätsdiagramme erwies sich präskriptiv als Herausforderung. Die Erwartungen des Entwicklungspartners, konkrete Anforderungen vor der Umsetzung bereitzuhaben, wurden enttäuscht. Mit Verständnis, transparenter Kommunikation und der ehrlichen Aussage, dass dies in der Softwareentwicklung nicht unüblich ist, gaben wir zu verstehen, dass dieses Gefühl anhalten wird.

Leistungsreflexion

Es wird das Lean-UX-Prinzip ‹Verzicht von aufwendigen Anforderungs- und Dokumentationsartefakte› angewendet. Der Fokus liegt auf den benötigten Ergebnissen, welche als Epics, Tasks und User-Stories in einem Jira-Board erfasst werden. Gemeinsame Meetings finden regelmässig statt und werden so lange geführt, bis alle das gleiche Verständnis haben und bis sich jeder seiner Aufgaben im Klaren ist.

Zu Beginn jedes Meetings werden die ersten drei produktiven Minuten dafür aufwendet, dass alle Teilnehmenden die eigenen gewünschten Gesprächsthemen als Sticky Note hinterlegen. Jeder Teilnehmende erklärt darauf hin, was er mit dem Stichwort in Verbindung bringt. Die Notizzettel werden thematisch zugeordnet. Danach wird so lange diskutiert, bis alle Themen abgearbeitet sind, was in der Regel zwei bis drei Stunden in Anspruch nimmt. Dank dieser Teamkultur entsteht ein reger Wissensaustausch und das UX-Team entwickelt ein Verständnis für die technischen Rahmenbedingungen.

In dieser Projektphase erwies sich, dass es zielführend ist, mit Geduld und Ausdauer die gleichen Dinge mehrmals und ohne Zeitdruck zu besprechen.

10. Technischer Exkurs

10.1. Systemzustände

Die Tests im Zyklus 0 haben aufgezeigt, dass Nutzende Schwierigkeiten haben, den aktuellen Systemzustand zu erkennen. Zwecks besserer Nutzbarkeit werden die Systemzustände und die Übergänge zwischen den Zuständen definiert. Das Ziel ist, Nutzenden eine Orientierungshilfe in Bezug auf im Hintergrund ablaufende technische Prozesse bieten zu können, ohne dass sie diese verstehen müssen. Dabei sollen sie die Nutzungskontrolle und Freiheit haben, den gewünschten Systemzustand selbst herzustellen.

Schlafen: Das System «schläft». Es hört weder zu, noch spricht es. Durch eine Aktion des Nutzenden kann es in den Zustand des «Zuhörens» versetzt werden.

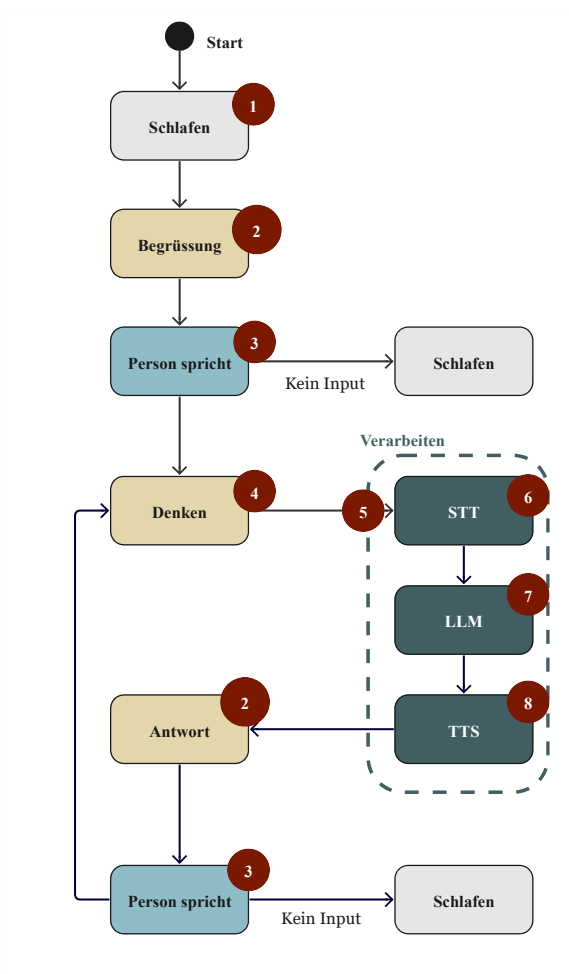
Sprechen: Das System spricht und kann in diesem Zustand nicht zuhören.

Denken: In diesem Zustand wird das Sprechen vorbereitet (siehe Zustands- und Aktivitätsdiagramm).

Zuhören: Die nutzende Person spricht, während das System «zuhört».

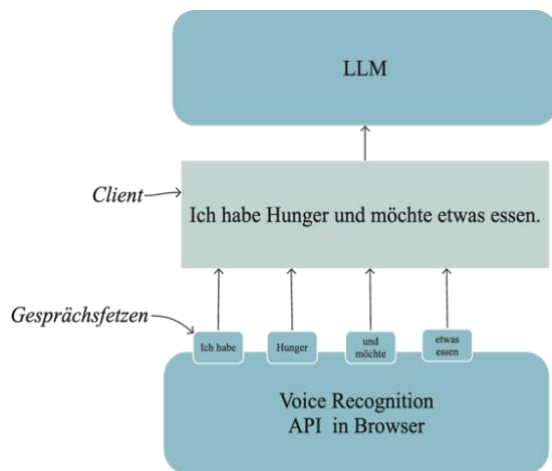
10.2. Zustands- und Aktivitätsdiagramm

Abbildung 22: Zustands- und Aktivitätsdiagramm



1. Das System schläft. Durch Drücken auf einen Button wird es in den Zustand «Sprechen» versetzt.
2. Eine Begrüssung ertönt. Anhand eines Cookies wird ermittelt, ob es sich um eine Erst- oder Folgenutzung handelt: Im Falle der ersten Nutzung wird ein Begrüssungstext ausgegeben, andernfalls wird eine Grussfloskel wie «Hallo» oder «Wie kann ich helfen?» ausgegeben.
3. Die nutzende Person spricht. Sie hat eine maximale Redezeit von 20 Sekunden. Danach wird der Verarbeitungsprozess gestartet. Wird während fünf Sekunden kein Sprachinput festgestellt, wechselt das System in den Zustand «Schlafen».
4. Der Modus «Denken» besteht aus der Verarbeitung des Gesprochenen (6), der Generierung der Antwort durch das LLM (7) und der Generierung der Sprache (8). Zusammen suggeriert diese Phase das «Nachdenken».
5. Die Sprache wird in Text (TTS) umgewandelt. Dieser Vorgang wird auf dem Client fortlaufend ausgeführt. Anschaulichkeitshalber wird dies diesem Zustand zugeordnet. (siehe Abbildung 23). Der gesprochene Text wird über eine Schnittstelle an ein LLM übergeben. Die Anfrage wird durch das LLM verarbeitet, und eine Antwort wird in Textform erstellt. Sobald die Generierung des Textes beendet ist, wird der Text sprachsynthetisiert.

Abbildung 23: Spracherkennung und Übertragung



Die Spracherkennung (STT) wird über die Speech-Recognition-Schnittstelle des Browsers genutzt. Dabei werden gesprochene Worte aus Audiodaten extrahiert, um Texte zu verstehen und zu verarbeiten (mdn web docs, 2023). Der Text wird im Client des Nutzers zusammengesetzt (Abbildung 23). Sobald während mehr als drei Sekunden keine Sprache über den Mikrofoneingang erkannt wird, gilt der Aufnahmeprozess als beendet. Der extrahierte Text wird an ein LLM geschickt. Die LLMs Zephyr 7B (Zephyr 7B, o. D.) und Llama 2 (ai.meta, o. D.) werden auf einem eigenen Server betrieben.

ChatGPT 3.5 und 4.0 werden als externe APIs genutzt. Um die Fähigkeiten und Qualitäten der LLMs zu testen, kann über eine Auswahlfläche das LLM durch den Nutzer gewählt werden. Das LLM nimmt den Text entgegen und generiert eine Antwort auf Basis eines Prompts. Anhand des Prompts wird dem LLM die Rolle, die das System einzunehmen hat, zugewiesen. Im Prompt werden Aufgabe, charakterliche Attribute und Tonalität definiert. Hat das LLM die Antwort vollumfänglich generiert, wird der Text zur Stimmengenerierung übergeben. Das Resultat ist eine Audiodatei, welche an den Client geschickt und über die Lautsprecher des Computers ausgegeben wird. Der Prozess der Stimmengenerierung kann zwischen drei und mehr als zehn Sekunden dauern. Er stellt damit den grössten «Flaschenhals» in dem Prozess dar.

Ist die Audioausgabe beendet, wechselt das System wieder in den Zuhörmodus und erwartet eine Spracheingabe. Erfährt das System innerhalb von drei Sekunden keine weitere Eingabe, wechselt es in den Zustand «Schlafen».

11. Zyklus 4: Mikrointeraktionen

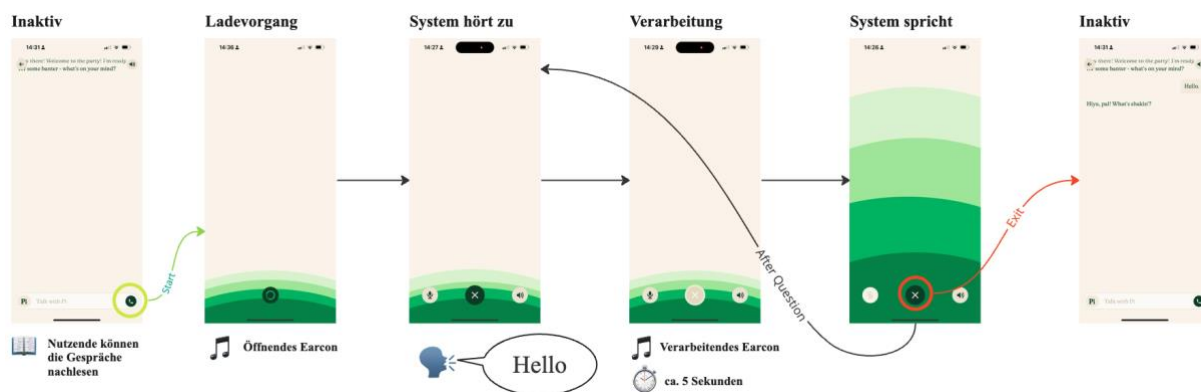
11.1. Einleitung

Entsprechend der ersten Usability-Heuristik zur Sichtbarkeit des Systemstatus muss dieser für Nutzende deutlich erkennbar sein (Nielsen, 1994). Das sprachgesteuerte Interface wird durch visuelle Elemente ergänzt, und Earcons signalisieren akustisch den Wechsel zwischen den Systemzuständen. Aufbauend auf dem Zyklus 3 wird die *System Persona* nun *Alfred* genannt.

11.2. Recherche

Als Vorbereitung für diesen Zyklus dient eine Analyse der iPhone-Applikation von PI.ai (Abbildung 25). Nutzende können mit PI auf zwei Arten interagieren: textbasiert, ähnlich dem Versenden von SMS oder WhatsApp-Nachrichten, oder über Spracheingabe, welche durch Drücken der Telefontaste ausgelöst wird und einen Anruf imitiert. Ein aufsteigender Ton signalisiert den Ladevorgang. Grüne Balken am unteren Bildschirmrand stellen eine Amplitude dar und zeigen visuell an, wenn PI spricht. Während der Konversation werden keine Texte eingeblendet. Frühere Konversationen bleiben als Chatverlauf in der Übersicht einsehbar.

Abbildung 24: Zustände von PI.ai



11.3. Annahmen

11.3.1. Problem-Statement

*Alfred wurde für **anhaltende Konversation mit Menschen** entworfen. Wir haben jedoch festgestellt, dass **Alfred seine Gesprächszustände nicht kommuniziert**, was sich in Form **mangelnder Nutzungsakzeptanz** auf unsere Geschäftstätigkeit auswirkt. Wie können wir **Alfred** dahin gehend verbessern, dass **ältere Menschen** aufgrund **klarer, verständlicher Systemzustände** erfolgreicher damit arbeiten können?*

11.3.2. Persona

Die Persona aus dem Zyklus 3 bleibt unverändert.

11.3.3. Hypothese

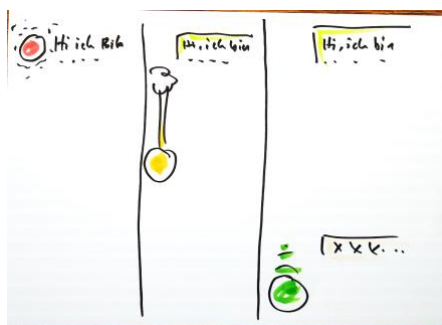
*Wir sind der Auffassung, dass wir eine **Nutzungsakzeptanz** erreichen, wenn wir **älteren Menschen** mittels **Microinteractions und Earcons** die **Systemzustände verständlich** kommunizieren.*

11.4. Design

11.4.1. Visuelles Design

Das visuelle Interface hat die Aufgabe, Nutzenden den Wechsel von Sprechenden zu signalisieren und die kognitive Gedächtnisbelastung zu reduzieren. In mehreren Design-Workshops werden verschiedene konzeptionelle Entwürfe gesammelt, mit denen aufgezeigt werden kann, wie die Vorgaben visuell umgesetzt werden können. Inspiriert durch das mentale Modell des «Pingpong» (Abbildung 25: Sketch «Pingpong») signalisiert ein Kreis die Gesprächsreihenfolge. Der obere rote Kreis zeigt das Sprechen des Systems an, der mittlere orange Kreis in Bewegung markiert den Übergang und der untere grüne Kreis steht für die Zuhörbereitschaft des Systems.

Abbildung 25: Sketch «Pingpong»



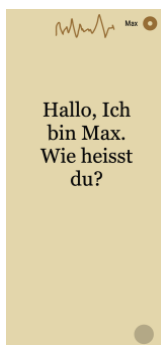
Weitere Designansätze werden exploriert:

Abbildung 26: Version
〈Chatverlauf〉



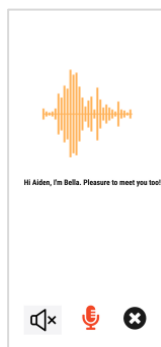
Es wird ein Chatverlauf in Betracht gezogen, wie er von WhatsApp, Bing und PI.ai genutzt wird. Nutzende können dadurch den Gesprächsverlauf einsehen. Die sprechende Person wird durch einen leuchtenden Punkt neben ihrem Namen angezeigt.

Abbildung 27: Version 〈Text〉



Die Kombination von gesprochenem und gelesenem Text, wie er vom Google Assistenten genutzt wird, wird geprüft. Hierbei wird der gesprochene Text zusätzlich visuell dargestellt, anstatt nur einen Gesprächsverlauf zu zeigen.

Abbildung 28: Version
〈Amplitude〉



Das prominente Anzeigen einer Amplitude wird in seiner Nutzbarkeit erforscht. Die Farbe der Amplitude und der Icons geben den Nutzenden Auskunft darüber, wer gerade spricht.

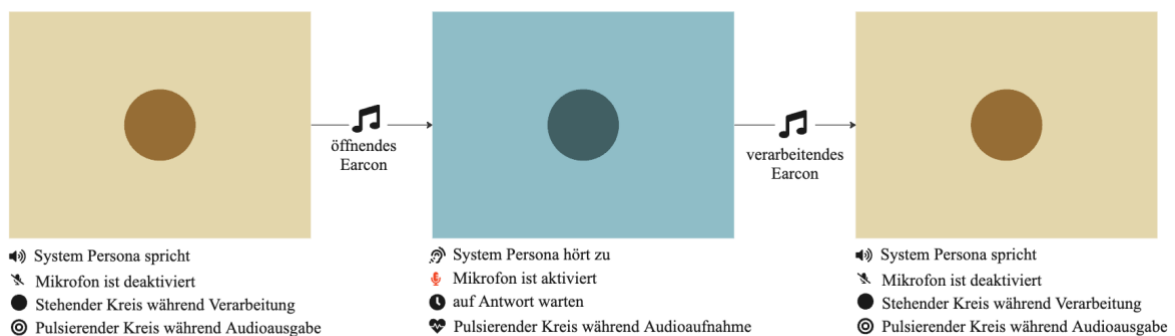
11.4.2. Earcons

Zuerst erklingt zur Aufmerksamkeitserzeugung [ein Intro-Earcon](#). Danach wird ein Intro-Text von Alfred vorgelesen. Ein über [drei Töne ansteigendes Earcon](#) signalisiert aus Sicht des Systems den Übergang vom Sprechen zum Zuhören. Die Verarbeitungszeit, die das System benötigt, wird durch ein Earcon, dass [sich wie ein Windspiel anhört](#), signalisiert.

11.4.3. Visuelles Interface

Ein minimales Design in Kreisform auf solidem Hintergrund in Kombination mit Earcons ergibt ein multimodales User-Interface. Die braune Farbe signalisiert das Sprechen des Systems, während die blaue Farbe Aufnahmebereitschaft anzeigt (Abbildung 22). In beiden Zuständen pulsiert der Kreis während des Sprechens und bleibt in der Verarbeitungszeit stehen. Auf das Anzeigen eines Textes wird bewusst verzichtet, damit Nutzende ihr Kurzzeitgedächtnis beanspruchen.

Abbildung 29: Visuelles Interface



11.5. Minimum Viable Product

Das multimodale Interface wird mit der WOZ-Methode online von vier Personen getestet: zwei Frauen im Alter von 65 und 76 Jahren sowie zwei Männer im Alter von 55 und 57 Jahren. Als Prototyp dient eine mit Tonanimationen angereicherte PowerPoint-Präsentation. Die Dialoge sind, abgesehen von der Frage nach dem Namen und der darauffolgenden namentlichen Begrüßung, linear und folgen dem Muster von Frage-Antwort-Frage-Antwort. Die Audioausgabe der Stimmen wurde von ElevenLabs, einem Unternehmen, das sich auf die Erzeugung synthetischer Stimmen spezialisiert hat, vorab generiert. Der Wizard hat die Aufgabe, im richtigen Moment auf die <Weiter>-Taste zu drücken, um die Animationen zu starten, welche die Audiodateien abspielen. Das MVP wird als erfolgreich betrachtet, wenn die Testpersonen benennen können, wann sie sprechen dürfen und wann nicht. Die Mikrointeraktionen werden anhand des Dialoges im Anhang 44 getestet.

Im anschließenden Cooling-off wird neben einer generellen Einschätzung eine Beurteilung der Stimme in Bezug auf Tonfall, Sprechgeschwindigkeit, Sprachstil und Menschlichkeit abgefragt. Schliesslich werden fünf verschiedene Stimmen, zwei weibliche und drei männliche, welche eine Begrüßung vortragen, zur Auswahl gestellt. Die Probanden und Probandinnen haben die Aufgabe, eine bevorzugte Stimme zu nennen.

11.6. Feedback

Bei einer Person findet ein glatter Wechsel statt. Person und System wechseln sich ohne Überlappung ab. Drei Personen reden zu früh beim Sprecherwechsel von System zu Person. Bei zwei Personen kommt dies nur zweimal vor, bei einer Person viermal. Bei einer Person wird einmal zu früh in den Verarbeitungsmodus gewechselt. Spätestens nach vier Sprecherwechseln verstehen sämtliche Testpersonen, wann die Sprechmöglichkeit eröffnet wird. Drei Teilnehmende geben an, dass der Farbwechsel im GUI den Zustandswechsel signalisiert. Eine Person nimmt an, dass eine Sprechpause die Übergabe signalisiert.

Die Earcons werden von einer Person als Signal für Zustandswechsel wahrgenommen. Eine weitere gibt an, dass sie im Laufe des Tests die Bedeutung der Earcons verstanden hat. Das Windspiel-Earcon zur Überbrückung der Verarbeitungszeit wird von einer Testperson als «nervös» und «verwirrend» bezeichnet. Bemängelt werden lange Wartezeiten bis zum Sprechen und ein daraus resultierender nicht stimmiger Dialogrhythmus. Ein nahtloser Übergang in die Zuhörbereitschaft wird erwartet. Der vorgefertigte, lineare und kontextfreie Dialog wird als unnatürlich empfunden. Eine Person spricht Schweizerdeutsch und zwei Hochdeutsch, während eine Person zunächst Hochdeutsch spricht und im Verlauf des Tests auf Schweizerdeutsch wechselt.

In Bezug auf die Menschlichkeit der Stimme ist die Bandbreite der Antworten diversifizierter. Die Antworten reichen von «gar nicht menschlich» bis zu dem Erscheinen einer Person vor dem «inneren Auge». Dennoch beurteilen die Versuchsteilnehmer/-innen die Stimmen als positiv. Von den beiden männlichen Probanden wird die Stimme «Carl» bevorzugt. Eine weibliche Testperson wählt eine andere männliche Stimme. Eine weitere Person kann sich nicht zwischen zwei weiblichen Stimmen entscheiden. Das Feedback zu den Stimmen ist uneindeutig, weshalb im folgenden Zyklus die Wahl der Stimme validiert wird.

Die Auswertung befindet sich im Anhang 46

11.7. Reflexion

Lernreflexion

Pragmatismus bedeutet, alles wegzulassen, was nicht unbedingt notwendig ist, um das geforderte Ziel zu erreichen. Designer/-innen neigen dazu, ein möglichst stimmiges Gesamtkonzept zu erstellen, was jedoch viel Aufwand erfordert und eine spätere Priorisierung schwierig macht. Die Herausforderung besteht darin, vor der Umsetzung bereits zu wissen, was als überflüssig betrachtet wird. Hinzufügen ist einfacher als weglassen.

Leistungsreflexion

Trotz anfänglichen technischen Störungen bei zwei Tests konnte deren Durchführung erfolgreich zu Ende geführt werden. PowerPoint erwies sich als richtige Wahl, um ein multimodales Interface zu testen. Mit der Software war es möglich, Übergänge bildlich und auditiv aufeinander abzustimmen. Die konsequente Reduktion auf das Notwendigste erwies sich einmal mehr als zutreffend. Im Laufe dieses Zyklus konnte herausgestellt werden, wie ein Designprozess von einem leeren Blatt bis zu einem testbaren Prototyp innert kurzer Zeit und effizient durchzuführen ist.

13. Exkurs: Prompt-Engineering

Beim Prompt-Engineering erfolgt eine Entwicklung und Optimierung von Prompts, um Sprachmodelle effizient für eine Vielzahl von Anwendungen zu nutzen. Es umfasst ein breites Spektrum an Fähigkeiten und Techniken, die für die Interaktion und Entwicklung mit LLMs nützlich sind. Die Fähigkeit, LLMs zu verstehen und mit ihnen zu interagieren, wird ein zentraler Bestandteil dieser Arbeit. Ein Prompt benötigt folgende vier grundsätzliche Parameter:

- **Instruktion:** eine bestimmte Aufgabe oder Anweisung, die das Modell ausführen soll;
- **Kontext:** externe Informationen oder zusätzlicher Kontext, der das Modell zu besseren Antworten führen kann;
- **Dateneingabe:** die Eingabe oder Frage, für die eine Antwort gesucht wird;
- **Ausgabe-Indikator:** die Art oder das Format der Ausgabe.

Wie in Designprozessen üblich, wird empfohlen, klein anzufangen und anhand von iterativem Vorgehen mehr Elemente und Kontext hinzuzufügen. Grössere Aufgaben sollen in kleinere zerlegt werden. Anweisungen sollen in Form von Befehlen formuliert werden. Mit relevanten Schlüsselwörtern soll experimentiert werden (Prompt Engineering Guide, o. D.).

14. Zyklus 5: Stimme

14.1. Einleitung

In diesem Zyklus wird untersucht, welchen Einfluss der Klang einer Stimme hat und was die Attraktivität einer Stimme ausmacht. Die Stimme soll zum gestalteten Charakter passen und in ein Gesamterscheinungsbild einfließen. Am Ende dieses Zyklus wird das Geschlecht der *System Persona* validiert sein.

14.2. Recherche

Die Psychoakustik beschäftigt sich mit dem Zusammenhang der menschlichen Empfindung von Schall als Hörereignis und mit dessen physikalischen Schallfeldgrößen als Schallereignis. Zu ihr gehören Lautheit, Schärfe, Tonheit, Rauigkeit, Tonhaltigkeit, Impulshaltigkeit und Schwankungsstärke (Ellermeier & Hellbrück, 2008). Die Stimme eines Menschen wird von folgenden vier Parametern bestimmt: Tonhöhe, Lautstärke, Klangfarbe und Vokal.

Die Tonhöhe und die Klangfarbe beeinflussen die wahrgenommene Körpergröße des Absenders. Im weltweiten Durchschnitt spricht ein Mann bei einer Frequenz von ca. 110 Hertz während Frauen bei 220 Hertz sprechen. Die durchschnittliche Sprechhöhe von Frauen in Deutschland im Alter zwischen 40 und 80 Jahren liegt heute bei 168 Hertz, während sie vor einigen Jahrzehnten ebenfalls bei 220 Hertz lag. Als Ursache wird die fortschreitende Emanzipation genannt (Böger, 2018). Tiefere Stimmen werden als kompetenter wahrgenommen. In Studien wurde nachgewiesen, dass Politiker/-innen mit baritonischen Stimmen mehr Wählerstimmen erlangen (Patalong, 2015).

Weibliche Stimmen sind komplexer und für das menschliche Gehirn schwieriger zu verarbeiten. Aufgrund der unterschiedlichen Größe sowie Form der Stimmbänder und des Kehlkopfes bei Mann und Frau haben sie eine natürlichere Melodie in ihren Stimmen. Dies führt zu einer höheren Anzahl an Klangfrequenzen (Sokhi, Hunter, Wilkinson, & Woodruff, 2005). Trotzdem sind die Stimmen von Alexa, Bing, Google und Siri in der Standardeinstellung weiblich. Auch die Stimme bei der SBB (Augustin, 2013) und die Standardstimme des Navigationsgerätes TomTom ist weiblich (Driessen, 2008).

Um zu überprüfen, ob die Wahl des Geschlechts der Stimme soziokulturelle Hintergründe hat, werden Filme und Serien, in welchen sprechende Computer vorkommen, auf ihr Geschlecht überprüft. Filme, in welchen Roboter sprechen, werden nicht berücksichtigt.

Tabelle 6: Voice-User-Interfaces in Filmen

Film/Serie	Name	Sprecher/-in	Jahr	Geschlecht
Star Trek (mehrere)	Computer, andere	Majel Barrett	1966	weiblich
2001 Space Odyssey	HAL 9000	Douglas Rain	1968	männlich
Alien	Mother	Helen Horten	1979	weiblich
Knight Rider	KITT	William Daniels	1982	männlich
Blade Runner	Rachael	Sean Young	1982	weiblich
Tron	MCP	David Warner	1982	männlich
The Hitchhiker's Guide to the Galaxy	Deep Thought	Helen Mirren	2005	weiblich
The Hitchhiker's Guide to the Galaxy	Eddie the computer	Thomas Lennon	2005	weiblich
Eagle Eye	ARIIA	Julianne Moore	2008	weiblich
Iron Man	Jarvis	Paul Bettany	2008	männlich
Moon	GERTY	Kevin Spacey	2009	männlich
Her	Samantha	Scarlett Johansson	2013	weiblich

Anhand der Literaturrecherche kann keine abschliessende Empfehlung in Bezug auf die Stimme gegeben werden. Die Stimmen müssen empirisch validiert werden.

14.3. Annahmen

14.3.1. Problem-Statement

*Alfred wurde für ein **Voice User Interface** entworfen. Wir haben jedoch festgestellt, dass Alfred **keine eigene Stimme hat**, was sich in Form **mangelnder Identität** auf unsere Geschäftstätigkeit auswirkt. Wie können wir Alfred dahin gehend verbessern, dass **ältere Menschen** aufgrund **einer angenehm empfundenen Stimme** erfolgreicher damit arbeiten können?*

14.3.2. Persona

Die Persona aus dem [Zyklus 3 System Persona](#) bleibt unverändert.

14.3.3. Hypothese

*Wir sind der Auffassung, dass wir **Nutzerakzeptanz** erreichen, wenn **ältere Menschen mit einer angenehmen Stimme** gerne eine **Konversation** führen.*

14.4. Design

Es werden je fünf männliche und weibliche Stimmen von ElevenLabs vorselektiert. Für jede Stimme wird ein Muster mit einem einheitlichen Beispieltext generiert:

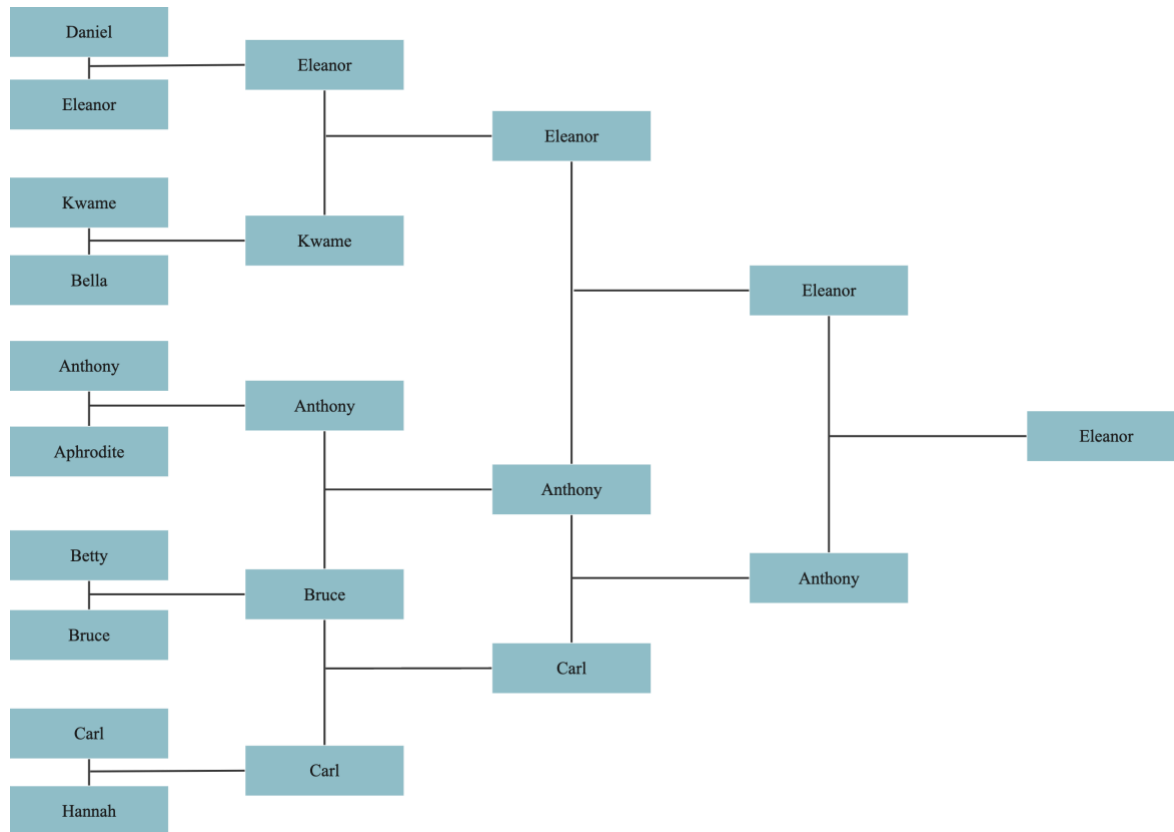
Der Hase war im ganzen Wald bekannt für seine Kraft und seine Schnelligkeit. Scheinbar mühelos konnte er innerhalb kürzester Zeit weite Strecken zurücklegen, ohne dabei ins Schwitzen zu kommen.

Die Stimmen von ElevenLabs heißen: [Anthony](#), [Aphrodite](#), [Bella](#), [Betty](#), [Bruce](#), [Carl](#), [Daniel](#), [Eleanor](#), [Hannah](#), [Kwame](#).

14.5. Minimum Viable Product

Die Tests werden mit drei weiblichen und vier männlichen Personen in einem Online-Meeting durchgeführt. Es sind 30 Minuten mit anschließender Befragung eingeplant. Die Selektion der Stimme wird in Turnierform ausgetragen (Abbildung 24).

Abbildung 31: Verlauf des Stimmenturniers mit Testperson A.



14.6. Feedback

Drei von vier Männern wählen eine männliche Stimme, zwei von drei Frauen eine weibliche. Für die Gesamtwertung werden neben der als bevorzugt genannten Stimme auch die weiteren Stimmen in der Endausscheidung berücksichtigt. Das Resultat fällt klar zugunsten der Stimme «Carl» aus. Seine Stimme hat es nur einmal nicht in die Auswahl der letzten Stimmen geschafft und war dreimal favorisierte Stimme. «Bella» konnte von den weiblichen Stimmen am meisten überzeugen und war fünfmal in der Endausscheidung, wovon sie einmal gewann. Zwei Testpersonen bekundeten Schwierigkeiten, eine Wahl zu treffen, und können nicht eindeutig eine bevorzugte Stimme benennen. Dies ergibt folgende Rangliste:

Tabelle 7: Rangliste der Stimmen

Stimme	Anz. der Nennungen unter den letzten 3 oder 4	Sieger/-in
Carl	6	3
Bella	5	1
Anthony	5	0
Eleanor	3	1
Aphrodite	1	0

Die Entscheidung fällt auf die Stimme «Carl». Die genannten Charakterisierungen passen zur *System Persona*. Ihr wird Ruhe, Gelassenheit, Kompetenz und eine gute Verständlichkeit zugeschrieben. Versuchsteilnehmende hören in der Stimme einen guten Erzähler und Interpret. Das Alter wird in den Bereich von 50 Jahren eingeordnet, was ebenfalls gut zu dem vordefinierten Alter von 55 Jahren passt.

14.7. Reflexion

Lernreflexion

Die Stimme sollte in dem Kontext, in dem sie gebraucht wird, getestet werden. Auch wenn in einer direkten Gegenüberstellung das Voice-Sample relativ kurz ist, sollte es möglichst den später geplanten Inhalt wiedergeben. Die Charakterisierung der Stimme hängt mit persönlichen Erfahrungen der Teilnehmenden zusammen. Die gleiche Stimme wurde teilweise unterschiedlich charakterisiert. Beispielsweise wurde «Bella» als übermotiviert, als «Hausdrache», als zurückhaltend und als angenehm beschrieben.

Mit 4, 8, 16, 32 ist eine geeignetere Anzahl an Teilnehmenden für ein Turnier vorhanden, da sich für jede Runde die Zahl um die Hälfte reduzieren lässt, ohne eine Zwischenrunde zu benötigen.

Leistungsreflexion

Alle Tests inkl. der Interviews konnten in der geplanten Zeit oder schneller durchgeführt werden. Die Idee mit dem Turnier erwies sich als effiziente Methode, wenn es darum geht, in Gegenüberstellungen eine Präferenz der Nutzenden zu erhalten.

Für den Prototyp wird nur eine Stimme berücksichtigt. Es wird jedoch als unabdingbar begriffen, Nutzenden eine Auswahl zu bieten. In den Tests wurde die Tendenz ersichtlich, dass auch weibliche Personen eine weibliche Systemstimme bevorzugen.

15. Zyklus 6: Finale

15.1. Einleitung

In diesem Zyklus werden die Ergebnisse aus den Zyklen 3 (*System Persona*), 4 (Mikrointeraktionen) und 5 (Stimmen) in einem codierten Prototyp von Menschen in einem Alterspflegeheim getestet. Die Ergebnisse dieses Zyklus entscheiden über Nutzen und Nützlichkeit von persönlichen, intelligenten und sprachbasierten Assistenten.

15.2. Recherche

Es werden keine weiteren Recherchen durchgeführt.

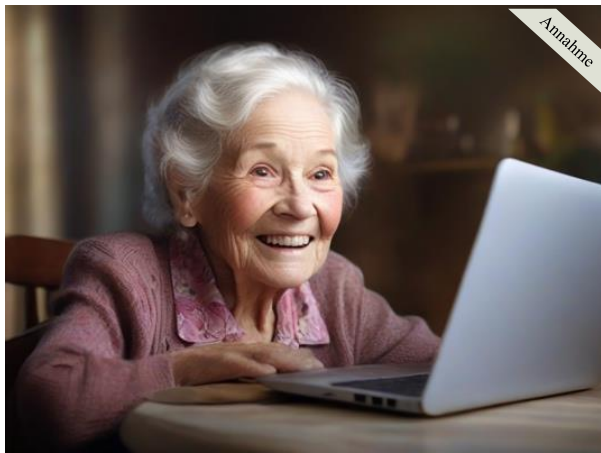
15.3. Annahmen

15.3.1. Problem-Statement

*Ältere Menschen benötigen empathische Unterstützung beim Gedächtnistraining und in der Strukturierung ihres Alltags, aber die aktuelle Lösung mit menschlichen Betreuern und Betreuerinnen ist aufwendig. Ein sprachbasierter intelligenter Assistent wie **Alfred**, der als Gedächtnistrainer dienen kann, könnte diese Rolle übernehmen. Jedoch besteht die Herausforderung darin, die empathischen Fähigkeiten einer menschlichen Betreuung nachzuahmen, um eine effektive und emotional ansprechende Nutzungserfahrung für ältere Nutzende zu schaffen. Die erfolgreiche Entwicklung eines solchen Systems könnte nicht nur den Aufwand für das Gedächtnistraining senken, sondern auch eine breitere **Zugänglichkeit und Konsistenz in der Betreuung älterer Menschen** ermöglichen.*

15.3.2. Persona

Abbildung 32: Persona Zyklus 6



Annette ist 83 Jahre alt, geschieden und lebt allein in einem Alterspflegeheim. Während ihrer beruflichen Zeit war sie Sekretärin und Mutter von zwei Kindern. Sie wünschte sich, dass sie sich mehr unterhalten könnte, denn sie hat viele Geschichten aus ihrem Leben zu erzählen. Trotz ihres fortgeschrittenen Alters unternimmt sie gerne etwas mit anderen Menschen. Die Ausflüge werden immer kürzer, und immer häufiger wird sie dabei

von einer Pflegefachkraft und ihrem Rollator begleitet. Während der Spaziergänge erzählt sie Geschichten von damals. Sie kann sich aber leider nicht immer an jüngere Geschehnisse erinnern. Für Aktivitätsvorschläge ist sie immer offen. Die neue Technologie ist ihr noch immer suspekt. Wenn die Enkelkinder zu Besuch kommen, zeigen sie ihr Fotos auf Instagram auf ihren Smartphones. Dafür hat sie ungläubiges Kopfschütteln übrig. Wenn der Besuch gegangen ist, widmet sie sich wieder ihrem Fernseher und schaut eine Quizsendung.

15.3.3. Hypothese

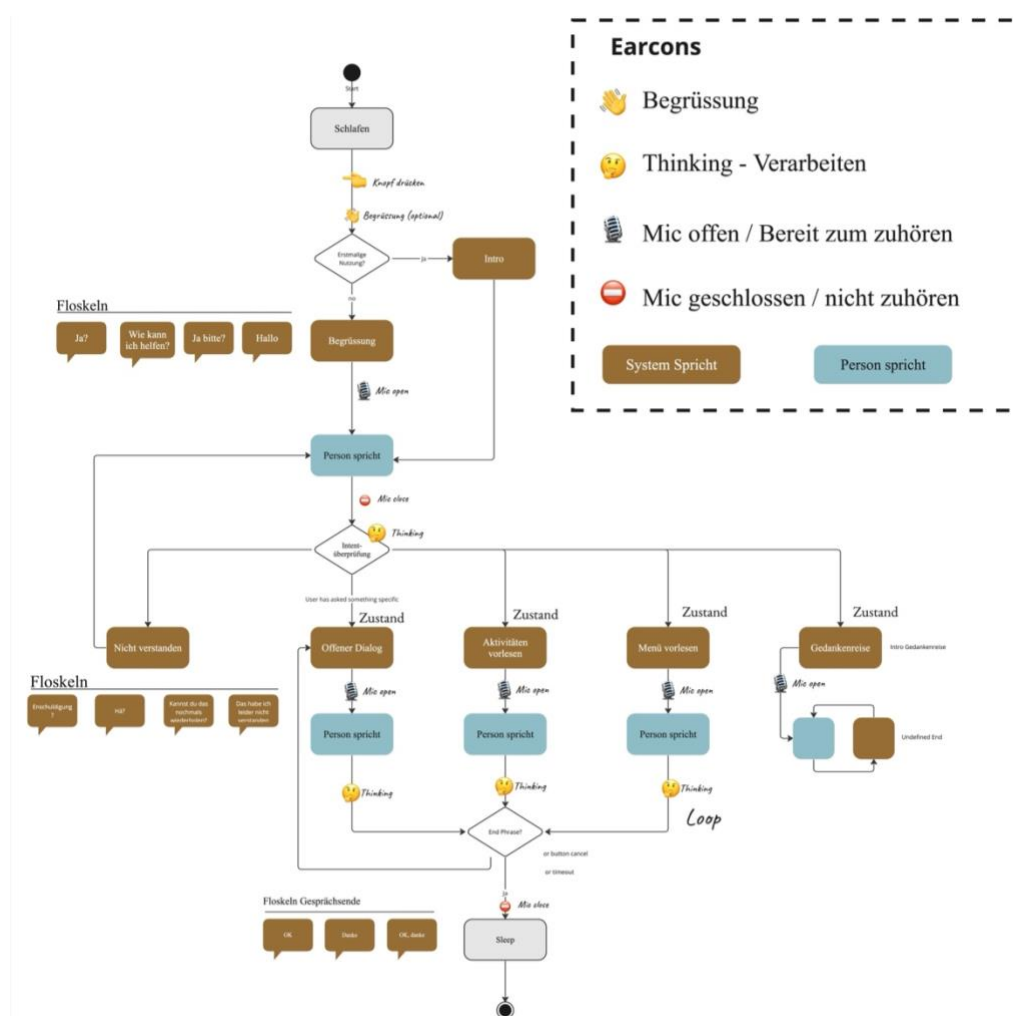
*Wir sind der Auffassung, dass wir eine Nutzungsakzeptanz erreichen, wenn **ältere Menschen** mit Alfred einen **auf ihren Interessen** basierten **Dialog** führen können.*

15.4. Design

Damit Nutzende die richtigen Agents aufrufen können, bedarf es einer Gesprächsnavigation. Navigiert wird über den Intent in natürlicher Sprache (siehe Abbildung 33: Dialogablauf und Agents). Der Prototyp sieht ergänzend zum Kapitel 10 - Technischer Exkurs folgenden Ablauf vor:

1. Nachdem die Person das erste Mal gesprochen hat, überprüft das System den Intent. Sofern das System die Absicht des Nutzenden verstanden hat, übergibt das System an einen Agenten, welcher die weiteren Dialoge steuert.
2. Wird die Absicht nicht verstanden, gibt dies das System zu verstehen.
3. Die Agenten <Aktivitäten vorlesen> und <Menü vorlesen> sind einfache lineare Dialoge und bestehen aus einem Frage-Antwort-Schema. Der Agent <offener Dialog> wird auch zu den linearen Dialogen gezählt, da es sich um Wissensabfragen handelt.
4. Der Agent <Gedankenreise> folgt dem Dialogmuster des aktiven Zuhörens und sieht einen offenen Gesprächsausgang vor.

Abbildung 33: Dialogablauf und Agents



15.5. Minimum Viable Product

Die Rekrutierung von vier Testpersonen wird durch die Leitung des Zentrum Sunnegarte in Bubikon vorgenommen. Die Tests werden auf den Zimmern der Bewohner/-innen durchgeführt. Es nehmen zwei Frauen im Alter von 76 und 78 Jahren und zwei Männer mit Hörgeräten im Alter von 89 und 90 Jahren teil. Der Prototyp wird auf einem 16"-Macbook Pro im Browser Chrome getestet.

Nach der eigenen Vorstellung wird den Testpersonen «Alfred» vorgestellt. Das Ziel des Tests ist herauszufinden, ob der Prototyp als nützlich (Utility) und nutzbar (Usability) empfunden wird. Sie werden darum gebeten, die Anwendungsfälle «Menu abfragen», «Tagesprogramm abfragen» und «Gedankenreise» durchzuführen. Entstehen im Verlauf des Tests allgemeine offene Fragen, dann werden die Probanden und Probandinnen darum gebeten, die Frage an Alfred zu richten. Erhoben werden das persönliche Feedback, Beobachtungen zu Interaktionen, Fragen, die an das System gestellt werden, und Dokumentationen hinsichtlich der Anwendenden und des Systems selbst. Vorgesehen ist eine Dauer von einer Stunde pro Person. Der Test gilt als erfolgreich, wenn die Versuchsteilnehmer/-innen es nutzen können und/oder es nutzen wollen würden.

15.6. Feedback

Alle Teilnehmenden sprechen Schweizerdeutsch. Chat-GPT 4 kann diese Eingabe verarbeiten und darauf eine Antwort in deutscher Sprache ausgeben. In zwei Fällen hat Alfred etwas missverstanden und eine unzusammenhängende Antwort gegeben.

Sämtliche Personen geben ein überaus positives Feedback zu Alfred ab. Zwei Personen führen explizit an, Alfred als gute Idee zu befinden, eine stuft ihn als «super» und «faszinierend» ein. Als Verbesserungsvorschläge werden mehr Interaktivität, mehr «Lustigkeit» und die Fähigkeit, auch Musik abspielen zu können, genannt. Nutzende wünschen sich mehr «Persönlichkeit» (drei Nennungen). Alfred soll ein Gesicht haben und er soll auch etwas über sich erzählen können. Dabei darf er etwas aufgeweckter, «kecker» und leichtgewichtiger sein. Die Gedankenreise soll genereller sein.

Festgestellt wird, dass die beiden männlichen Personen mehr als fünf Sekunden benötigten, um Alfred eine Antwort geben zu können. Das System ging in den Schlafen-Modus, bevor die Antwort gegeben werden konnte. Die weiblichen Teilnehmerinnen überschritten die Sprechzeit von 20 Sekunden regelmässig, sodass die Eingabe jedes Mal vom System unterbrochen und eine Antwort generiert wird. Die Technik des aktiven Zuhörens von Alfred wird angenommen, indem die Versuchsteilnehmer/-innen auf die Folgefragen antworten. Kann das System die Eingabe nicht verarbeiten, stellt es Rückfragen. Nutzende sind diesbezüglich nachsichtlich im Umgang mit Alfred. Ein bedeutender Erfolg war, dass eine Person einen 16-minütigen Dialog mit Alfred führte.

Die Ergebnisse der Auswertung befinden sich im Anhang 52. Die Aufzeichnungen und Transkripte sind im nicht öffentlichen Anhang zu finden.

16. Empfehlungen an den Auftraggeber

Die studienübergreifende interdisziplinäre Zusammenarbeit erwies sich als Erfolg. In wenigen Wochen konnte ein Prototyp entwickelt werden, welcher als ‹Proof of Concept› getestet werden konnte. Durch die Verbindung von User-Centered Design (UCD) mit künstlicher Intelligenz konnten freie und unstrukturierte Dialogformen in der Interaktion Mensch-Computer ermöglicht werden.

Die Domäne ‹ältere Menschen› eignete sich für den Anwendungsfall von intelligenten Sprachassistenten. Trotz oder aufgrund von sensorischen, sprachlichen und kognitiven Einschränkungen wurde ein System entwickelt, welches einer grösseren Bandbreite zur Verfügung steht. Im Wiederholungsfall empfiehlt das Projektteam, die Domäne als Bestandteil des Auftrags zu machen.

16.1. Bekannte Probleme

Der Prototyp enthält zum Zeitpunkt der Abgabe bekannte Fehler. Dazu gehören die folgenden:

Browserkompatibilitätsprobleme: Aktuell wird nur Chrome in der aktuellen Version unterstützt.

Zahlen und Urzeiten: Diese werden auf Englisch ausgesprochen, da es sich bei der Stimmensynthese um einen amerikanischen Anbieter handelt.

Namen und Fremdwörter: Die Aussprachen unterliegen regionalen Besonderheiten und können sich für Nutzende ungewohnt anhören.

Sprachsynthetisierung: Die Generierung der Sprache kann zehn Sekunden und mehr in Anspruch nehmen. Der Umfang des Textes bestimmt die Dauer der Synthetisierung. Offene Fragen führen zu längeren Antworten, was mehr Zeit in Anspruch nimmt.

Ende des Gesprächs: Jedes Gespräch hat irgendwann ein Ende. Das Konversationsdesign wurde so implementiert, dass das System stets in den Zuhören-Modus wechselt, nachdem es gesprochen hat. Stattdessen sollte das System verstehen, wann eine Rückfrage nicht notwendig ist.

16.2. Weitere Hypothesen

Im Verlauf des Projektes wurden Hypothesen besprochen, welche nachrangig priorisiert wurden. Wenn die Produktentwicklung fortgeführt wird, ist es die Empfehlung, diese Hypothesen auf ihre Anwendung zu überprüfen.

Wir sind der Meinung,

- dass ein visueller Avatar der *System Persona* mehr Persönlichkeit verleiht (siehe 16.2.1 Avatar).
- dass die Turn-Taking-Regeln durch Nutzende konfigurierbar sein sollten (siehe 16.2.2 Konfigurierbare Turn-Taking-Regeln)
- dass ein Dialog von beiden Gesprächsparteien initiiert werden darf (siehe 16.2.3 Dialogaufnahme)
- dass ein System immer zuhören sollte (siehe 16.2.4 Variable Redezeit und ständige Zuhörbereitschaft) und
- dass eine eigene Hardware erforderlich ist (siehe 16.2.5 Hardware).

16.2.1. Avatar

Aus dem Feedback des Zyklus 6 geht hervor, dass sich Nutzende mehr Menschlichkeit und Emotionalität wünschen. Diese Hypothese soll anhand von Emojis getestet werden.

Abbildung 34: Zustand
(Schlafen)



Abbildung 35: Zustand
(Sprechen)

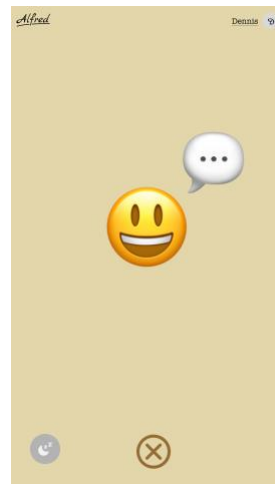


Abbildung 36: Zustand
(Zuhören)

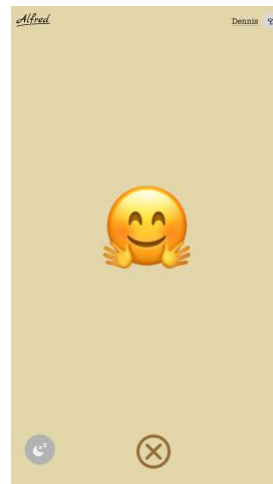
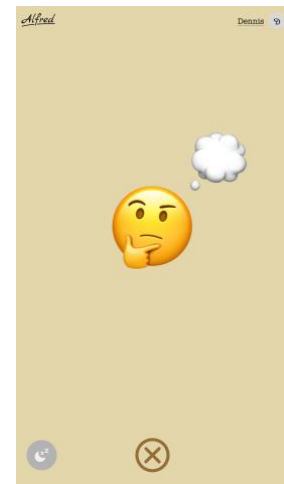


Abbildung 37: Zustand
(Denken)



16.2.2. Konfigurierbare Turn-Taking-Regeln

Abbildung 38: Konfiguration
(Sprecherwechsel)

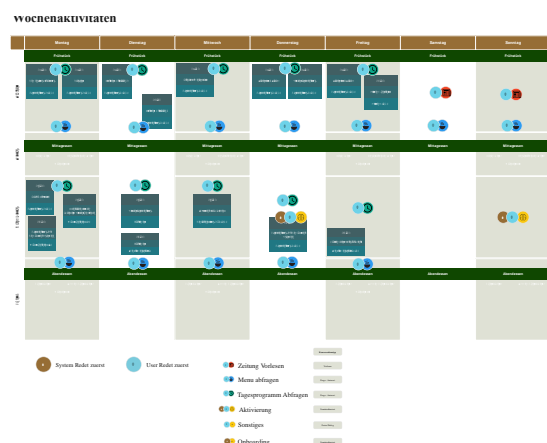


Im Zyklus 6 konnte beobachtet werden, dass die Wortübergabe von der kognitiven Fähigkeit abhängig sein kann. Eine Sprechpause von vier Sekunden kann für Menschen mit Verständnisschwierigkeiten zu kurz sein, während die gleiche Zeit für Menschen mit gutem Verständnis als lange wahrgenommen wird. Über die Benutzereinstellungen lassen sich die Turn-Taking-Regeln konfigurieren (siehe Abbildung 38).

16.2.3. Dialogaufnahme

Die Frage, welche/-r Dialogpartner/-in das Gespräch aufnimmt, war während des gesamten Arbeitsprozesses relevant. Wenn es eine *System Persona* gibt, gibt es konsequenterweise auch eine System-Journey (Abbildung 39). Ein Blick auf die Wochenaktivitäten des Zentrums Sunnegarte liess vermuten, wer wann zuerst spricht. Dabei zeigte sich, dass auch ein System einen Dialog aufnehmen sollte, wenn es in der jeweiligen Situation als angemessen erscheint.

Abbildung 39: System-Journey



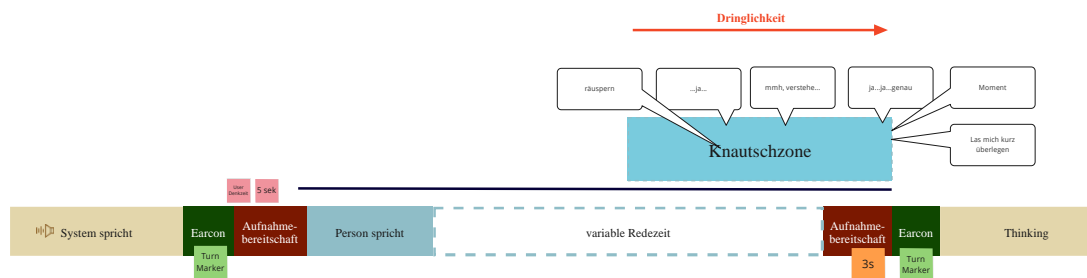
16.2.4. Variable Redezeit und ständige Zuhörbereitschaft

Wenn ein System den Dialog initiieren darf, sollte es wissen, wann dies angebracht ist und wie es auf sich aufmerksam machen soll. Unter der Voraussetzung, dass es sich um eine <Voice-only>-Situation handelt, in welcher eine nutzende Person das Gespräch durch ein Wakeword in den Zustand <Zuhören> setzen kann, wird eine ständige Zuhörbereitschaft vorausgesetzt.

In den Aufzeichnungen ist zu sehen, wie Menschen auf Unterbrechungen durch das System reagieren, sobald die Redezeit von 20 Sekunden überschritten wird. Anhand der Körpersprache wird eine unzufriedene Irritation sichtbar. Um dem entgegenzuwirken, müsste die zur Verfügung stehende Redezeit verlängert werden. Dies könnte auf zwei Arten geschehen:

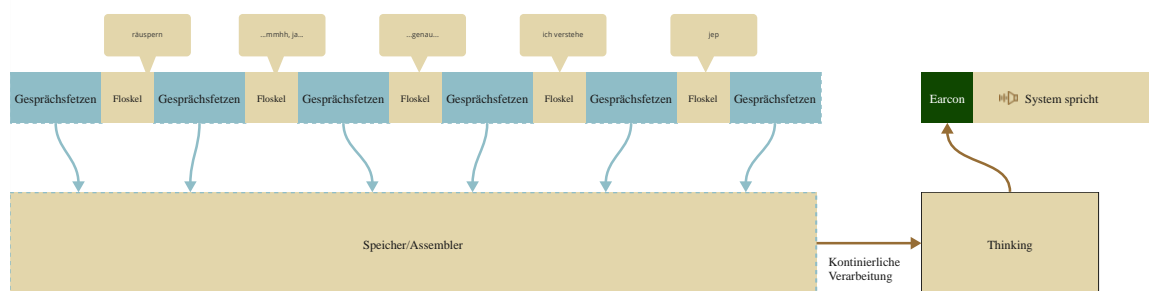
1. **Variable Redezeit:** Die Person kann so lange sprechen, wie sie möchte. Das System übernimmt eine aktive Zuhörerrolle und signalisiert durch Floskeln die Hörerrolle. Damit das System die Rolle als aktiver Zuhörer wahrnehmen kann, wird eine *«Knautschzone»* als überlappende Simultansequenz genutzt.

Abbildung 40: Variable Redezeit



2. Bei **ständiger Zuhörbereitschaft** kann ein System längere, durch die Person einseitig geführte Gespräche aufnehmen. Dabei werden die Gesprächssequenzen fortlaufend an einen Zwischenspeicher geschickt, welcher die Texte zusammensetzt. Die Texte werden kontinuierlich verarbeitet und eine Antwort wird vorbereitet.

Abbildung 41: Ständige Zuhörbereitschaft



16.2.5. Hardware

Ein Dialog kann von einem System gestartet werden. Bei Pi erfolgt dies durch Push-Notifications (Abbildung 34). Die erweiterte Zuhörbereitschaft, die Dialogaufnahme anhand eines Wakewords durch eine Person und die Dialogaufnahme durch das System setzen aufgrund von Restriktionen in Betriebssystemen von Smartphones und in Webbrowsern eine eigens dafür ausgelegte Hardware voraus.

17. Reflexion

17.1. Projektplanung und Verlauf

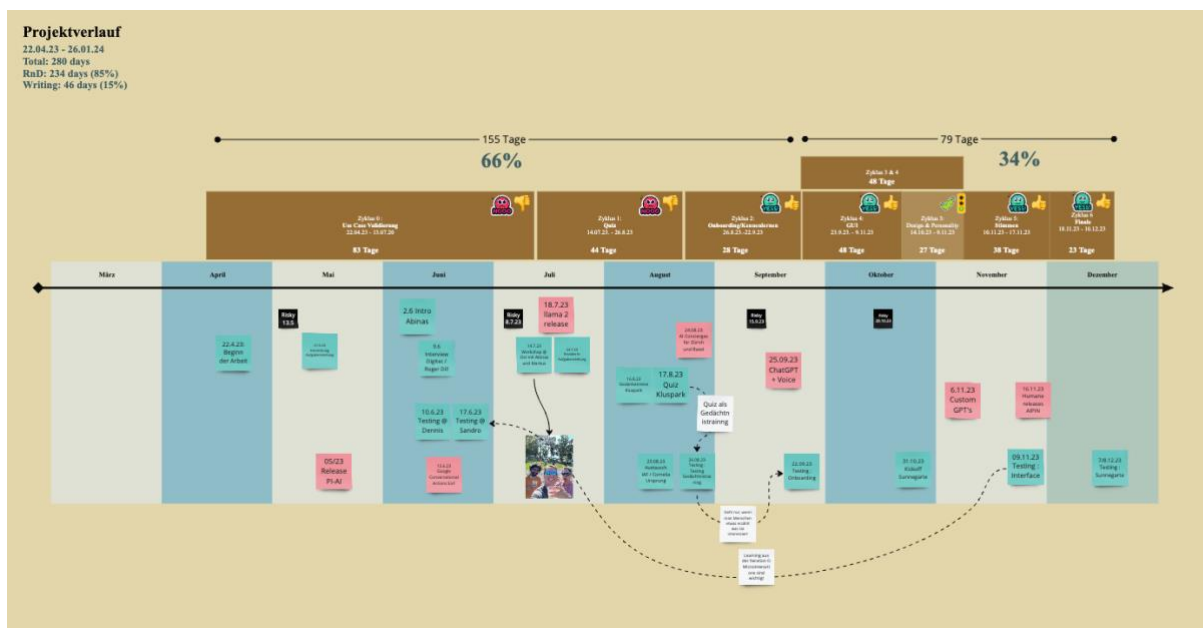
Der initial erstellte Projektplan wurde übertroffen. Über eine Dauer von knapp neun Monaten kam es in sieben Zyklen lediglich zu einer Verzögerung von zwei Wochen. Mit der Darstellung in Abbildung 42: Projektverlauf kann aufgezeigt werden, wie die Geschwindigkeit mit dem Fortschreiten des Projektes korreliert.

Es verging 66 % der Zeit oder 155 Tage, bis der erste Zyklus mit positivem Feedback abgeschlossen werden konnte. Wir lernten, dass *NEIN* nur ein Akronym für «Noch eine Iteration notwendig» ist. Die Motivation wurde mit einem Zitat von Samuel Beckett aufrecht gehalten:

«Ever tried. Ever failed. No matter. Try Again. Fail again. Fail better.»

Die intensive Auseinandersetzung in den ersten beiden Zyklen hat sich später gewinnbringend erwiesen. Die Erkenntnisse aus dem Zyklus 0 führten zu einem minimalistischen, Lean-UX-entsprechenden pragmatischen Vorgehen.

Abbildung 42: Projektverlauf



17.2. Lean UX

Um das selbst auferlegte Produktziel zu erreichen, waren schnelles Feedback und Ergebnisse notwendig. Das Vorgehen nach Lean-UX-Prinzipien erleichterte es, dieses Ziel zu erreichen.

Eine **problemorientierte** Herangehensweise bildete die Basis für den Erfolg dieser Arbeit. Die **Erlaubnis zum Scheitern** ermöglichte ein exploratives Vorgehen und die Durchführung von Experimenten mit offenem Ausgang. Als Beispiel können WOZ angeführt werden.

Das Team war bestrebt, **Verschwendung zu vermeiden** und nur genau so viel zu entwickeln, wie zur Hypothesenüberprüfung nötig war. Rückblickend war dennoch nicht jedes Artefakt und jede Analyse zielführend oder nützlich. Eine kollaborative und **interdisziplinäre Zusammenarbeit** ermöglichten es, ein **gemeinsames Verständnis** aufzubauen. Dies gilt auch für die studienübergreifende Zusammenarbeit. In der Folge konnte pragmatisch in der **Erstellung von Artefakten** vorgegangen werden. Es fand ein kontinuierlicher Entwicklungs- und Lernprozess statt, wodurch **Phasen vermieden** werden konnten.

An den Umgang mit Unsicherheiten, welche aus Hypothesen und Annahmen entstehen können, musste sich das Team gewöhnen. Trotz negativer Resultate wurden Fortschritte erzielt, da das Ergebnis auf einer Annahme basierte. Schritt für Schritt bewegten wir uns **von Zweifel zur Gewissheit**. Die Fähigkeit, sich auf genau ein Problem und eine Hypothese zu konzentrieren, wurde erlernt und zunehmend zu einer Stärke. Ausdruck dessen sind die im Verlauf immer kürzer werdenden Zyklen (siehe Abbildung 42: Projektverlauf – Seite 74): Je **kleiner die Einheiten** waren, desto schneller konnte ein Ergebnis erzielt werden.

Für die Feldstudien, Experteninterviews, Experimente und Usability-Tests folgte das Team dem Prinzip **«Raus aus dem Gebäude»**, wann immer möglich. In diesem Sinne begleitete uns Abinas zu den Usability-Test im Zentrum Sunnegarte. Indem wir unser **Projekt nach aussen getragen** haben, wurde aufgezeigt, dass Interesse dafür besteht. Es eröffneten sich dadurch Kontakte zu fachkundigen Personen und die Möglichkeit für diesen Usability-Test.

17.2.1. Conversation-Design

Fundierte Konversationsanalysen bleiben eine Voraussetzung, um zu verstehen, wie natürlichsprachliche Dialoge in einer bestimmten Domäne ablaufen. Dialogflows sind als frühe und leichtgewichtige Designartefakte zu betrachten, um Testszenarien zu entwerfen. Sie sollten aber nicht zu detailliert sein, da Gespräche einen unvorhergesehenen Verlauf nehmen können. Im Zyklus 2 konnte das Team dieses Vorgehen erfolgreich praktizieren.

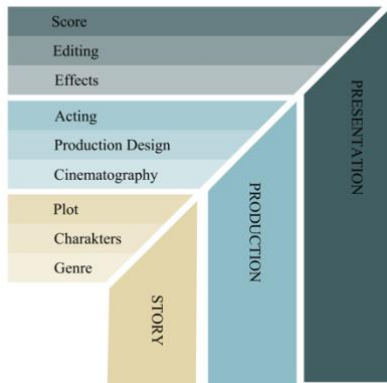
Hinsichtlich des Fortschrittes im Bereich Conversational AI sind Teile in den Vorgehensmodellen in Bezug auf Konversationsdesign zu hinterfragen. Durch die Verfügbarkeit von LLMs eröffnen sich neue Möglichkeiten und Anwendungsfälle.

Dialogorientierte Interfaces unterstützen, im Gegensatz zu bestehenden Sprachsystemen, freie und unstrukturierte Dialoge. Regelbasierte Dialogstrukturen werden durch Prompts in natürlichsprachlichen Texten ersetzt. Dabei ist zu beachten, dass LLMs gelegentlich zu Halluzinationen neigen und sich unvorhergesehen verhalten können. Die Spracherkennung und Verarbeitung zu Texten kann anfällig für Störungen sein, verursacht durch Hintergrundgeräusche, unbeteiligte Gesprächsteilnehmer/-innen oder Dialekte. Dennoch kann ein LLM mit wenig inhaltlichem Verständnis eine Konversation aufrechterhalten.

17.3. Das Eizzutto-Modell zur Entwicklung von Conversational Interfaces

17.3.1. Herleitung

Abbildung 43: Baxley-Model für Filme, eigene Darstellung



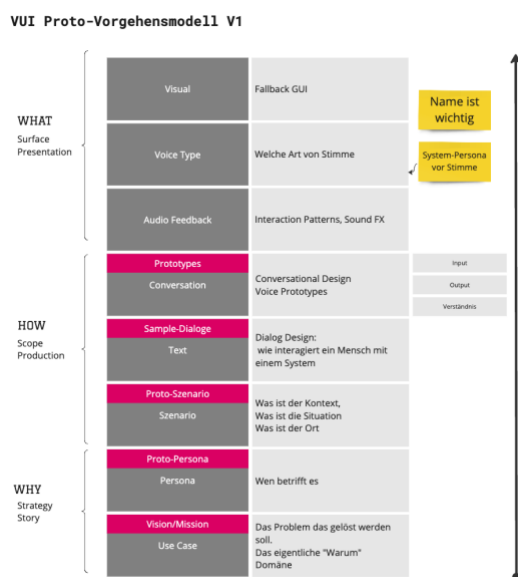
Mit dem universellen Modell eines User-Interfaces, auch als Baxley-Modell bekannt, wird ein Vorgehen für interaktive Medien und Produkte beschrieben. Es dient sowohl der fachübergreifenden Kommunikation als auch der Priorisierung von Aufgaben und Ressourcen (Abbildung 43: Baxley-Modell für Filme, eigene Darstellung). Indem die Komponenten eines Films oder eines Interfaces in einzelne Schichten zerlegt werden, ist es möglich, besser zu verstehen, wie die verschiedenen Aspekte zusammenwirken, um ein kohärentes Erlebnis zu schaffen (Baxley, 2003).

Abbildung 44: Why-What-How



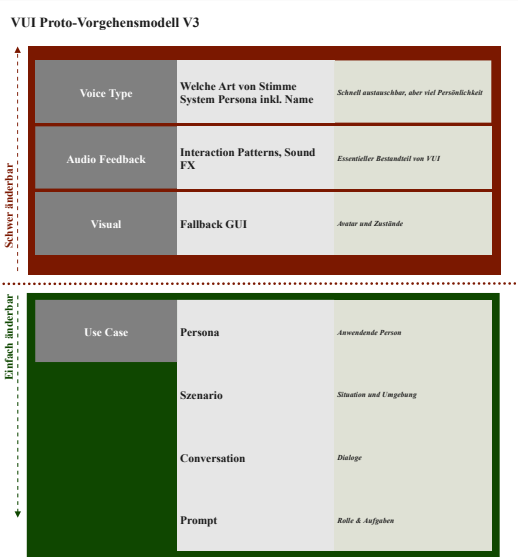
Ein allgemeineres Vorgehen zur Entwicklung von Geschäftsmodellen bietet das <Why-What-How-Prinzip> nach Simon Sinek (Abbildung 44). Das Vorgehen unterstützt eine ziel- und nutzenorientierte Produktentwicklung (Sinek, 2011).

Abbildung 45: Proto-Vorgehensmodell Version 1



Die Entscheidung fiel für das Baxley-Modell sowie für das Why-What-How-Modell und es wurde ein erstes Proto-Modell entwickelt. Die Ebenen <why>, <how>, <what> entsprechen den Ebenen <Strategy>, <Scope>, <Presentation> aus dem 5S-Modell (Garrett, 2011). Aufbauend auf einen Anwendungsfall wird eine Persona entwickelt. Daraus ergibt sich ein Szenario, welches Dialoge enthält. Dialoge werden durch das System geführt und nicht präskriptiv vorgegeben. Als einfach austauschbar werden die Interaktionsmuster, die Stimme und die visuelle Oberfläche angenommen. Die *System Persona* soll vor der Stimme bestimmt werden, da eine Namensänderung Anpassungen an visuellen Elementen zur Folge haben kann (siehe Abbildung 45).

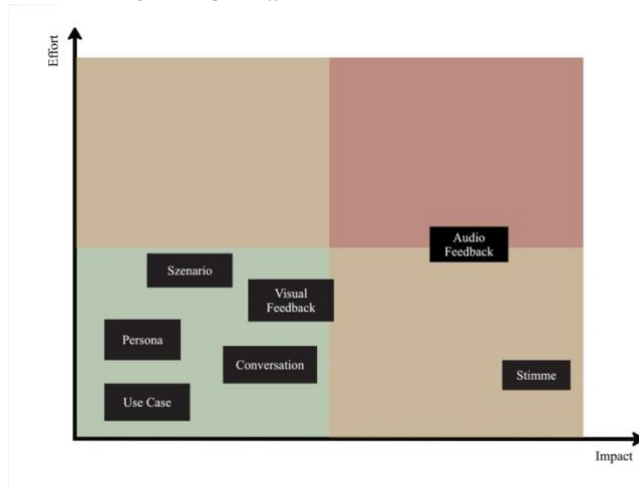
Abbildung 46: Proto-Vorgehensmodell Version 3



Das Modell wird verfeinert und nach dem Aufwand, welche eine Änderung nach sich zieht, dargestellt. Einfach austauschbar sind anwendende Personen und Nutzungssituationen. Die Konversationen entstehen durch den freien Mensch-Maschinen-Dialog und müssen anders als bei bestehenden Systemen (siehe Kapitel 4.2.3 Dialogstrukturen in dialogorientierten Bedienoberflächen) nicht vorgegeben werden. Im Prompt wird die Rolle und die Aufgabe des Systems definiert.

Schwer änderbar ist dagegen die visuelle Oberfläche, da sie den Systemstatus widerspiegelt. Auditive Elemente sind essenzielle Bestandteile eines VUIs. Deren Änderungen können grössere Auswirkungen für Nutzende nach sich ziehen, sind aber aus technischer Sicht einfach austauschbar. Die Stimme als Bestandteil der *System Persona* zieht eine Reihe von Änderungen mit sich.

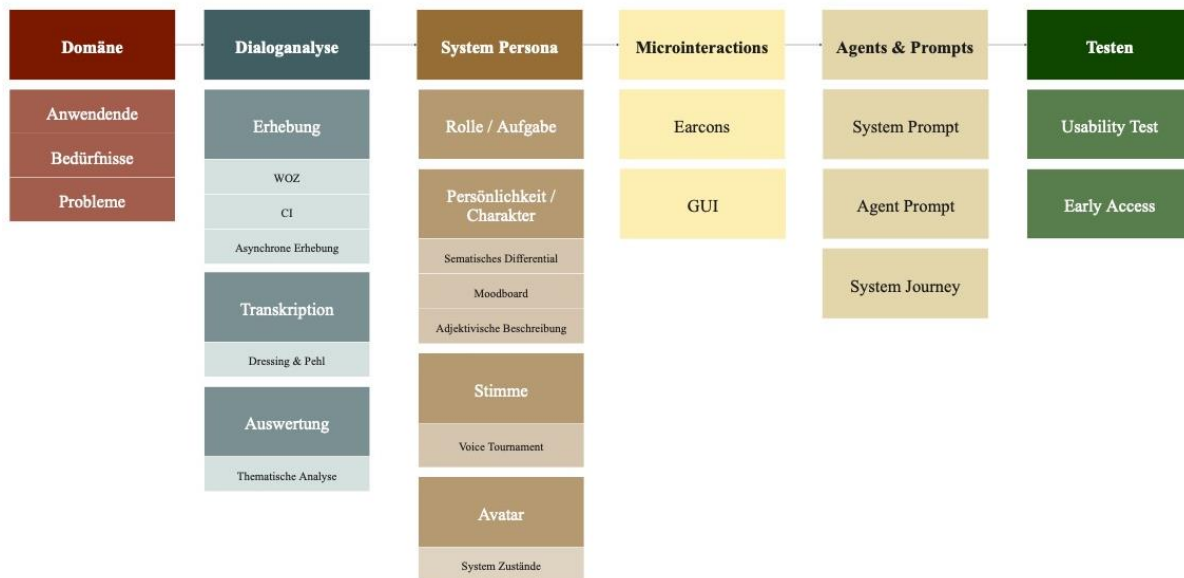
Abbildung 47: Impact-Effort-Matrix



Das Modell wird unter Einbezug des Softwareentwicklers bzw. der Softwareentwicklerin verfeinert und in eine Impact-Effort-Matrix überführt, da diese aussagekräftiger erscheint.

17.3.2. Das Eizzutto-Modell

Abbildung 48: Eizzutto-Modell



Benannt nach den Erfindern und Autoren dieser Arbeit, Dennis Eitner und Sandro Pezzutto, liefert das Eizzutto-Modell einen Leitfaden zur Erstellung von dialogorientierten Nutzungsoberflächen. Das sich an Lean UX orientierende Vorgehensmodell ergibt im Zusammenspiel mit den in diesem Bericht beschriebenen Methoden und Artefakten ein Vorgehen, in welchem innert sechs Zyklen ein auf Sprache basierendes Conversational Interface entwickelt werden kann.

Ausgehend von einer Domäne mit Bedürfnissen und Problemen von anwendenden Personen werden Dialoge erhoben, transkribiert und ausgewertet. Die thematische Analyse liefert die Basis für ein Regelwerk und für die Charakteristik der Dialoge. Sind die Systemzustände identifiziert, können auditive und visuelle Mikrointeraktionen gestaltet werden. Aufgabe, Rolle und Verhaltensweise der *System Persona* werden als Prompt in einem LLM hinterlegt. Unter der Voraussetzung, dass die Mikrointeraktionen eines VUI intuitiv und verständlich sind, kann der Anwendungsfall und damit auch die nutzende Person sowie die Domäne mit wenig Aufwand ausgetauscht werden.

Das Vorgehen zur Entwicklung einer *System Persona* zwecks Bedienung weiterer Anwendungsfälle wird repliziert. Mit minimalem Aufwand wird ein Alter Ego namens «Fred» entworfen, welcher für den Einsatz eines spielerischen Kumpels für Kinder im Alter von sechs bis zehn Jahren dient. Zu seinen Aufgaben zählen Fragen beantworten, Kopfrechnen trainieren, und Geschichten zusammen erfinden. Die Mikrointeraktionen sind die gleichen, jedoch zeigt sich, dass der Zustand des «Denkens» für jüngere Menschen zu lange dauert.

Mit einiger Sicherheit ist anzunehmen, dass UX-Research zukünftig nicht dem Schreiben von Anforderungen wie User-Stories dient, sondern der iterativen Erstellung von Prompts. Das Eizzutto-Modell bietet ein effektives und effizientes Vorgehen dazu.

17.4. Ethik

Die Auseinandersetzung mit Ethik erfolgte während des gesamten Projektes. Der Umstand, dass die Aufgabe der Software darin besteht, Prozesse zu optimieren, kann auch bedeuten, dass Maschinen Menschen ersetzen. Dass Menschen mit repetitiven Aufgaben wie Kassierer/-innen oder Mitarbeitende an Check-in-Schaltern durch Touch-Screen-Interfaces ersetzt werden, hinterlässt gemischte Gefühle. Einerseits muss die ökonomische Situation der Menschen bedacht werden und andererseits wird eine effizientere Abwicklung der eigenen Aufgaben begrüsst.

Dass Systeme Menschen ersetzen sollen, welche betreuerische und therapeutische Aufgaben haben, war stets Gegenstand von Diskussionen. Die Befürchtung, dass unsere Arbeit auf Akzeptanz stossen könnte, liess Befürchtungen aufkommen, dass menschliche Aufgaben in Zukunft von Systemen übernommen werden. Sympathien für ältere Menschen – auch wir werden eines Tages dazu gehören – und Pflegende sorgten für eine gesunde skeptische Haltung und verhinderte allzu grossen Enthusiasmus.

Das Feedback und die Neugier der Menschen, insbesondere im Zyklus 6, milderten einige Befürchtungen. Der Umstand, dass ein Bedarf und eine Akzeptanz von anwendenden Personen vorhanden sind, relativiert eigene Meinungen. Unter der Voraussetzung, dass die Privatsphäre eines jeden Menschen gewahrt bleibt, überwiegen die Chancen die Gefahren.

17.4.1. Inklusives Design

Während visuelle Bedienoberflächen sämtliche Gruppen von Nutzenden ausschliessen, welche weder Augen noch Hände zur Verfügung haben, ist dies bei einem auf natürlicher Sprache basierendem Interface nicht der Fall. Wenn Menschen mit Einschränkungen nicht in Randgruppen unterteilt werden, sondern ihre Bedürfnisse als Gemeinsamkeit betrachtet werden, stellen diese zusammengenommen einen grösseren Anteil in der Bevölkerung. Bei älteren Menschen kumulieren sich Einschränkungen, so dass sie als Referenz für Anwendungsfreundliche Produkte dienen können.

Zwar gibt es Bestrebungen zur Inklusion von Menschen mit Einschränkungen durch Accessibility-Standards. Die Bezeichnungen von Menschen mit Behinderungen als solche ist eine Ausgrenzung per se. Überdies gelten Accessibility-Zertifizierungen als aufwendig und verhindern oftmals eine technologische Weiterentwicklung von Angeboten, da diese Rezertifizierungen erfordern. Auch Accessibility-Features selbst sind umständlich zu benutzen, was jeder, der diese Features ausprobiert, unweigerlich selbst erfährt. Eine Nutzung findet nur aus Mangel an Alternativen statt.

17.5. Lernreflexion

Eine menschenzentrierte, auf einem konkreten Anwendungsfall basierende Produktgestaltung kann durch ein technologieorientiertes und anwendungs- sowie personenfreies Vorgehen ersetzt werden. Für dialogorientierte Oberflächen und Sprachsysteme sind andere Vorgehensmodelle notwendig. Zwar gibt es Literatur zum Vorgehen in der Gestaltung von Voice-Interfaces, diese sind jedoch auf bestehende VUIs mit linearer Dialogstruktur ausgerichtet. Ein Vorgehensmodell für ein VUI mit freien Dialogformen existierte nicht und wurde im Laufe dieser Arbeit selbst entwickelt (siehe 17.3 Das Eizzutto-Modell zur Entwicklung von Conversational Interfaces).

17.5.1. Lernreflexion von Dennis Eitner

Das Scheitern und Reflektieren ist ein integraler Bestandteil in der Entwicklung von Menschen und Systemen. Ich musste lernen, dass auch ein Nein trotz guter Vorbereitung des MVPs mit entsprechender Nachbearbeitung als Erfolg angesehen werden kann. Anhand dieser Eigenschaften ist es möglich, in weniger Wochen ein Produkt zu entwickeln, welches Begeisterung bei Nutzenden hervorruft. Das menschenzentrierte Vorgehen und die thematische Analyse hat mir ein neues Verständnis über das Wesen des Menschen als auch über Systeme eröffnet. Die Analyse von Dialogen, Interaktionen und Beziehungen hören für mich nicht mit dieser Arbeit auf. Die Erwartungen an diese Ausbildung wurden mit Abschluss dieser Arbeit übertroffen. Mein bedeutsamstes Take-away lautet wie folgt: In Zukunft werden Gespräche mit Systemen geführt.

17.5.2. Lernreflexion von Sandro Pezzutto

In dieser Arbeit wird aufgezeigt, dass sich dialogorientierte Interfaces natürlichsprachlich nutzen lassen. Sprache, wie sie in der Kommunikation von Mensch zu Mensch angewendet wird, kann auch in der Mensch-Maschine-Interaktion funktionieren. Die daraus resultierenden Vorzüge hinsichtlich Usability und Accessibility sind unübersehbar. Nutzende müssen ihre Absichten nicht länger in einen Befehl für ein technisches Modell übersetzen (was ein GUI letztlich ist). Zu sagen, was sie wollen, welches Ergebnis sie sich wünschen, ist für ein Grossteil der Menschen möglich. Zugänglichkeit zu Systemen kann sowohl für Personen mit sprachlichen Einschränkungen, bspw. Analphabeten und Analphabetinnen, als auch für technisch wenig versierte Menschen ermöglicht werden.

Die WOZ-Experimente erwiesen sich als spannend und lehrreich. Sie erfordern sorgfältige Vorbereitung. Je enger die Hypothese formuliert ist, desto besser gelingt sie. Die Durchführung verlangt ein Mass an schauspielerischen Fähigkeiten des Wizards. Dies muss in der Interpretation der Ergebnisse mitberücksichtigt werden.

Die vertiefte Auseinandersetzung mit der Sprache erwies sich lehrreicher als erwartet. Die Konversationsanalyse ist ein bedeutsamer Schritt im Gestaltungsprozess von VUI's. Neue Erkenntnisse und das Auffrischen von bereits erworbenem Wissen im Bereich der Kommunikation sind ein wertvolles Learning aus dieser Arbeit.

17.6. Leistungsreflexion

Die zu Beginn gestellte Forschungsfrage wird folgendermassen beantwortet: Im Unterschied zu den bestehenden Dialogsystemen soll eine freie Konversation zwischen Menschen und Maschine möglich sein. Dabei sollen Menschen so sprechen können, wie sie es im Umgang mit ihrer Sprache gewohnt sind. Die Dialogform hat sich an der natürlichen Sprache des Menschen zu orientieren. Es ist die Aufgabe von Systemen und deren Gestaltern und Gestalterinnen, das Sprachverständnis dem des Menschen anzupassen. Menschen können ihre Sprache zwar einem System anpassen, empfinden es jedoch als unnatürlich.

Das Forschungsziel und Projektziel «Erhebung von Dialogen» wurde erreicht. Dialoge wurden thematisch analysiert und auf ihre Anwendung überprüft. Die Diskursregeln im Umgang mit VUIs können zwar schnell gelernt werden. Dass ein linearer Dialog mit glattem Sprecherwechsel gestaltet wird, ist dem mentalen Modell im Umgang mit bestehenden Systemen zu schulden. Rückwirkend sollte sich die Forschung diesem Aspekt widmen, um ein System zu entwickeln, welches immer zuhört und einen überlappenden Wechsel erlaubt. Die Sprechpause, verursacht durch Stimmengenerierung, sollte nicht länger sein als in zwischenmenschlichen Dialogen.

Das ambitionierte Produktziel, einen codierten Prototyp zu entwickeln, wurde übertroffen. Im Unterschied zur herkömmlichen GUI-basierten Produktentwicklung wird ein Wesen mit Persönlichkeit erschaffen. Sowohl Nutzen als auch Nützlichkeit wird ausdrücklich nachgewiesen. Im Laufe des Projektes begannen sich mehrere Personen aus unterschiedlichen Institutionen und Organisationen für die vorliegende Arbeit zu interessieren. Die zu Beginn diskutierten Anwendungsfälle wirken plausibel. Das entwickelte System weckt Hoffnung auf weitere erfolgreiche Entwicklungen.

17.7. Teamreflexion

Was dieses Team besonders auszeichnet, ist das Leben der Lean-UX-Prinzipien. Dass Scheitern ein Teil des Prozesses ist und auch ein «Nein» als Erfolg angesehen werden kann, musste erlernt werden. Das Vertrauen in die Fähigkeit der anderen war aufgrund von ehrlicher, geduldiger und empathischer Kommunikation die Basis des Erfolgs. Ein humorvoller Umgang mit Fehlern und miteinander sorgte dafür, dass es nie langweilig wurde. Eine neugierige, proaktive und problemorientierte Geisteshaltung in der Ursachenforschung der Zyklen 0 und 1 war der Grundstein für die darauffolgenden Erfolge. Die Fähigkeit und der Wille, Zusammenhänge zu verstehen, wo noch keine erkennbar sind, setzt Geduld und Ausdauer voraus. Beides war im Übermass vorhanden.

Kollaboratives und Workshop-orientiertes Vorgehen gehörten ebenso zu den Erfolgsfaktoren dieser Zusammenarbeit. Dazu werden Tools wie Miro rege genutzt. Das Team vereinbarte regelmässige wöchentliche Treffen. Diese wurden einerseits genutzt, um die Tests sowie Experimente durchzuführen und die erworbenen Eindrücke unmittelbar zu verarbeiten. Sich im Eins-zu-eins-Kontakt ausserhalb von digitalen Räumen auszutauschen, hat zusätzlich geholfen, das gemeinsame Verständnis zu verbessern.

18. Fazit

In dieser Arbeit konnte erfolgreich demonstriert werden, dass ein längerer Dialog in natürlichsprachlicher Sprache zwischen einem Menschen und einem System geführt werden kann. Die Voraussetzungen für ein Zustandekommen eines Dialogs sind ein gegenseitiges Interesse, das Einhalten von Konversationsregeln und aktives Zuhören. Ein System soll sich an den Interessen und Bedürfnissen des Menschen orientieren, wobei es diese selbst erheben kann. Beide Dialogteilnehmer/-innen verfügen über ausreichende Reparaturmechanismen, um ein Gespräch in Gang zu halten. Über kleinere Unverständlichkeiten wird hinweggehört. Ein System soll in der Lage sein, Dialektsprache zu verstehen, und in deutscher Sprache antworten.

Die Disziplin des HCID wird sich in Zukunft stärker mit den Themen «Conversational Interfaces» und «künstlicher Intelligenz» auseinandersetzen müssen. In dieser Arbeit wurde gezeigt, dass Anwendungsfälle und Personas austauschbar sind, sofern die darunterliegenden visuellen und auditiven Interaktionsmuster für Nutzende verständlich sind. Es wurde bewiesen, dass selbst ältere, undeutlich sprechende Personen mit Hörgeräten und wenig technischem Verständnis in der Lage sind, mit Alfred zu kommunizieren. Die Diskursregeln werden nach wenigen Sprecherwechseln verstanden.

Conversational Interfaces und persönliche, auf Sprache basierende intelligente Assistenten bieten echte Alternativen, welche mit heutigen vorhandenen Technologien umsetzbar sind. Selbst Aufgaben in der persönlichen Betreuung, welche bis anhin nur Menschen zugetraut werden, können durch Systeme übernommen werden. In der studienübergreifenden interdisziplinären Zusammenarbeit wird aufgezeigt, wie mit unkonventionellen Methoden, trotz ihrer Unerprobtheit, unter Einhaltung von Lean-UX-Prinzipien innert kürzester Zeit grossen Herausforderungen von Menschen beigegeben werden kann. Es ist an der Zeit, unkonventionelle Methoden zu nutzen. Diese Arbeit ist der Beweis.

19. Danksagung

Das Projektteam bedankt sich bei allen Menschen, die sich trotz ihrer Aufgaben im Alltag sich die Zeit genommen haben, mit uns zu sprechen, an unseren Experimenten teilzunehmen, und ihr Fachwissen, ihre Ratschläge sowie Einschätzungen mit uns geteilt haben. Wir bedanken uns bei Abinas für die tolle Zusammenarbeit. Besonderen Dank gebührt unserem Auftraggeber und Coach, Prof. Dr. Markus Stolze, welcher uns unermüdlich mit Ideen und Literaturverschlügen versorgte, uns mit Vertrauen und Offenheit begegnet ist und die erfolgreiche studienübergreifende Zusammenarbeit ermöglichte, ohne die diese Arbeit nicht möglich gewesen wäre. Wir hatten viel Spass, haben zugehört und dazugelernt. Danke!

20. Literaturverzeichnis

- Baxley, B. (06. 06 2003). *Sematic Scholar*. Von Universal model of a user interface : <https://www.semanticscholar.org/paper/Universal-model-of-a-user-interface-Baxley/fa826b6e1f64fe25a19916d9c971a84190570d0d>
- Blech, J. (11. 10 2022). *Das Geheimnis der gesunden Gehirne* von Der Spiegel: <https://www.spiegel.de/wissenschaft/medizin/gehirn-das-geheimnis-eines-fitten-denkgorgans-a-04d83b1c-5eb8-4efd-9e05-9d0da798227e>
- Böger, H. (11. 01 2018). *Die weibliche Stimme ist tiefer geworden* . Von Der Tagesspiegel: <https://www.tagesspiegel.de/wissen/die-weibliche-stimme-ist-tiefer-geworden-4108861.html> abgerufen
- Brinker, K., & Sager, S. F. (2010). *Linguistische Gesprächsanalyse* (Bd. 5). Berlin: Erich Schmidt Verlag.
- Brown, M. (kein Datum). *The Goldilocks Principle for Prototyping*. Abgerufen am 03. 01 2024 von NNg Nielsen Norman Group: <https://www.nngroup.com/videos/goldilocks-principle/#:~:text=Summary%3A%20The%20Goldilocks%20Principle%20says,an%20honest%20reaction%20from%20participants.>
- Bushnell, M. (21. 02 2023). *AI Faceoff: Siri vs. Cortana vs. Google Assistant vs. Alexa*. von businessnewsdaily: <https://www.businessnewsdaily.com/10315-siri-cortana-google-assistant-amazon-alexa-face-off.html>
- Carrefour. (16. 06 2020). *VOICE GROCERY SHOPPING: THE NEW E-COMMERCE GROCERY EXPERIENCE*. von Carrefour.com: <https://www.carrefour.com/en/news/voice-grocery-shopping-new-e-commerce-grocery-experience>
- Conversation Design Institute. (2020). *A guide to conversation Design: why, what and how*. Abgerufen am 24. 1 2024 von Conversation Design Institute: <https://www.conversationdesigninstitute.com/communications/what-is-conversation-design>
- Cooper, A., Reimann, R., Cronin, D., & Noessel, C. (2014). *About Face* (4th Ausg.). Intianapolis, IN 46256: John Wiley & Sons Inc.
- Dilraj S. Sokhi, M. D. (2005). *Male and female voices activate distinct regions in the male brain* (Bd. 27). (NeuroImage, Hrsg.) NeuroImage.
- Drew, P. &. (2006). *Conversation analysis*. London: SAGE Publications Limited.
- Driessen, C. (04. 05 2008). *Navi-Stimme Lisa ist echt*. Von n-tv: <https://www.n-tv.de/auto/Navi-Stimme-Lisa-ist-echt-article265604.html> abgerufen

- Ellermeier, W. H. (2008). *Hören – Psychoakustik – Audiologie*. Springer, Berlin, Heidelberg.
- Garrett, J. (2011). *The Elements of User Experience: User-Centered Design for the Web and Beyond, Second Edition*. Berkeley, CA, USA: New Riders 1249 Eighth Street.
- Genau, L. (25. 09 2020). *So wendest du Transkriptionsregeln richtig an*. Von Scribbr: <https://www.scribbr.ch/methodik-ch/transkriptionsregeln/> abgerufen
- Google Assistant. (27. 2 2021) von Conversation Design: <https://developers.google.com/assistant/conversation-design/create-a-persona>
- Greg Brockman, A. E. (01. 03 2023). *Introducing ChatGPT and Whisper APIs*. Von OpenAi: <https://openai.com/blog/introducing-chatgpt-and-whisper-apis> abgerufen
- Grice, H. P. (1975). *Logic and Conversation* (Vol. 3). (J. L. Peter Cole, Ed.) New York: Academic Press .
- Harvey Sacks, E. A. (1974). *A Simplest Systematics for the Organization of Turn Taking for Conversation*.
- Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, 37:1, 81-88,.
- IBM. (kein Datum). *What is natural language processing (NLP)?* Abgerufen am 10. 12 2023 von IBM: <https://www.ibm.com/topics/natural-language-processing>
- Inflection-2: The Next Step Up*. (22. 11 2023). von Inflection.ai: <https://inflection.ai/inflection-2>
- Introducing Llama 2*. (kein Datum). Abgerufen am 31. 12 2023 von Llama 2: <https://ai.meta.com/llama>
- Jeff Gothelf, J. S. (2021). *Lean UX* (Bd. 3). O'Reilly Media.
- Kahle, T., & Meissner, D. (2020). *All About Voice*. Freiburg: Haufe-Lexware GmbH & Co. KG.
- Kavlakoglu, E. (12. 11 2020). *NLP vs. NLU vs. NLG: the differences between three natural language processing concepts*. Abgerufen am 10. 12 2023 von IBM: <https://www.ibm.com/blog/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/>
- Leitfaden zur Prompt-Engineering*. (23. 12 2023). Abgerufen am 01. 01 2024 von Prompt Engineering Guide: <https://www.promptingguide.ai/de>
- Linke, A. N. (2004). *Studienbuch Linguistik*. de Gruyter.

- Martin Glinz, H. v. (2020). *Handbuch für das CPRE Foundation Level nach dem IREB-Standard* (Bd. 1). Karlsruhe, Deutschland: International Requirements Engineering Board (IREB) e.V.
- Martins, J. (21. 06 2022). *Active Listening*. Abgerufen am 18. 12 2023 von Asana: <https://asana.com/de/resources/active-listening>
- Moore, R. J., & Arar, R. (2019). *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. New York, NY, United States: Association for Computing Machinery. doi:<https://doi.org/10.1145/3304087>
- Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 89-100. doi:10.1109/MRA.2012.2192811
- Nielsen, J. (24. 4 1994). *10 Usability Heuristics for User Interface Design* . Von Nielsen Norman Group: <https://www.nngroup.com/articles/ten-usability-heuristics/> abgerufen
- Patalong, F. (10. 08 2015). *Baritone bekommen mehr Wählerstimmen* . Von Der Spiegel: <https://www.spiegel.de/wissenschaft/mensch/politiker-mit-tiefer-stimme-haben-bessere-chancen-bei-wahlen-a-1047309.html> abgerufen
- PwC. (2 2018). *Prepare for the voice revolution*. Abgerufen am 24. 1 2024 von pwc.com: <https://www.pwc.com/us/en/services/consulting/library/consumer-intelligence-series/voice-assistants.html>
- Robinson, L. (28. 02 2023). *Laughter is the best medicine*. Abgerufen am 20. 12 2023 von HelpGuide.org: <https://www.helpguide.org/articles/mental-health/laughter-is-the-best-medicine.htm>
- Rymlink, K. (21. 10 2023). *Zu jung fürs Vergessen*. Abgerufen am 25. 12 2023 von Der Spiegel: <https://www.spiegel.de/gesundheit/diagnose/alzheimer-mit-ende-40-zu-jung-fuers-vergessen-a-12395f41-330d-4c4a-b052-4596fc594a90>
- Schulz von Thun Institut. (kein Datum). *Das Kommunikationsquadrat*. Abgerufen am 13. 01 2024 von Schulz von Thun Institut: <https://www.schulz-von-thun.de/die-modelle/das-kommunikationsquadrat>
- Schwenkenbecher, J. (23. 09 2021). *Wie Stimmen wirken - und was sie wirklich verraten*. Von Spektrum der Wissenschaft Verlagsgesellschaft mbH : <https://www.spektrum.de/news/die-stimme-wie-sie-wirkt-und-was-sie-ueber-uns-verraet/1912996>
- Shannon, C. E. (1949). *The Mathematical Theory of Communication*. Urbana, IL: The University of Illinois Press.
- Sinek, S. (2011). *Start with why: How Great Leaders Inspire Everyone to Take Action*. Penguin.

Sokhi, D. S., Hunter, M. D., Wilkinson, I. D., & Woodruff, P. W. (2005). *Male and female voices activate distinct regions in the male brain* (Bd. 27). NeuroImage.

Sponheim, C. (06. 10 2023). *The ELIZA Effect: Why We Love AI*. Abgerufen am 03. 12 2023 von Nielson Norman Group: <https://www.nngroup.com/articles/eliza-effect-ai/#:~:text=Summary%3A%20Users%20quickly%20attribute%20human,is%20called%20the%20ELIZA%20effect.>

SRF. (16. 04 2013). *Die Stimme aus dem SBB-Lautsprecher*. Von SRF: <https://www.srf.ch/radio-srf-1/radio-srf-1-die-stimme-aus-dem-sbb-lautsprecher#:~:text=Aber%20wer%20steckt%20eigentlich%20hinter,%2D%20Sprecherin%2C%20Schauspielerin%20und%20Tiersch%20C3%BCtzerin.&text=Isabelle%20Augustin%20arbeitete%20lange%20als,eine%20ne>

Steinschaden, J. (24. 11 2023). *Inflection AI: „Zweitstärkstes LLM der Welt“ geht ins Generative-AI-Rennen*. Von Trending Topics: <https://www.trendingtopics.eu/inflection-2/>

Sy, D. (May 2007). *Adapting Usability Investigations for Agile User-centered Design*. *Journal of Usability Studies, Volume 2, Issue 3pp*, 112–132. Von <https://dl.acm.org/doi/10.5555/2835547.2835549> abgerufen

usability.de GmbH & Co. KG. (kein Datum). *Conversational User Interface (CUI)*. Abgerufen am 24. 01 2023 von Usability.de: <https://www.usability.de/leistungen/methoden/voice-user-interfaces.html>

Web Speech API. (19. 2 2023). Abgerufen am 31. 12 2023 von mdn docs: https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API

Westerhold, G. (12. 09 2018). *Voice Search und Voice Commerce – und wie Sie bei diesen Themen das Wort ergreifen*. (carpathia) Abgerufen am 11 2023 von carpathia.ch: <https://blog.carpathia.ch/2018/09/12/voice-search-und-voice-commerce-und-wie-sie-bei-diesen-themen-das-wort-ergreifen/>

Why create personal AI?. (02. 05 2023). Von Inflection.ai: <https://inflection.ai/why-create-personal-ai>

Wübbenhorst, P. D. (15. 02 2018). *Kontrollfrage*. Abgerufen am 17. 12 2023 von Gabler Wirtschaftslexikon: <https://wirtschaftslexikon.gabler.de/definition/kontrollfrage-37183/version-260626>

Zephyr 7B: The Next Generation Language Model. (kein Datum). Abgerufen am 31. 12 2023 von Zephyr 7B: <https://zephyr-7b.net/>

21. Anhang

Anhang 1: Definitive Aufgabenstellung per 21.05.2023.....	92
Anhang 2: S.M.A.R.T-Ziele	93
Anhang 3: Projektplan	94
Anhang 4: Risikoanalyse I.....	95
Anhang 5: Risikoanalyse I – Matrix.....	96
Anhang 6: Experteninterview R.D. (Zyklus 0)	97
Anhang 7: User Study Evaluierung Sprachassistenten (Zyklus 0).....	100
Anhang 8: Auswertung (Zyklus 0).....	105
Anhang 9: Pivot-Tabelle Task Erfolgreich (Zyklus 0).....	132
Anhang 10: Grafik Task Erfolgreich (Zyklus 0)	133
Anhang 11: Pivot Tabelle Task Erfolgreich nach Kategorie (Zyklus 0).....	134
Anhang 12: Pivot Tabelle - Ohne GUI Lösbar (Zyklus 0).....	136
Anhang 13: Grafik - Ohne GUI Lösbar (Zyklus 0).....	137
Anhang 14: Pivot Tabelle Frage vom System verstanden (Zyklus 0).....	138
Anhang 15: Pivot Tabelle Korrekte Antwort (Zyklus 0).....	139
Anhang 16: Pivot Tabelle Anzahl Versuche (Zyklus 0)	140
Anhang 17: Risikoanalyse II von 8.7.23	142
Anhang 18: Risikomatrix II.....	143
Anhang 19: Workshop Ergebnisse	144
Anhang 20: Aktualisierte Aufgabenstellung (MT23-G05) per 24.07.2023	145
Anhang 21: Experteninterview Aktivierungstherapeutin (Zyklus 2)	150
Anhang 22: Risikoanalyse III & IV vom 15.09.2023 & 20.10.2023	155
Anhang 23: Risikoanalyse III & IV-Matrix	156
Anhang 24: Experteninterview L.T. (Zyklus 1)	157
Anhang 25: Experteninterview S.H. (Zyklus 1).....	160

Anhang 26: Gedächtnisprotokoll Aktivierungstherapie 1	164
Anhang 27: Gedächtnisprotokoll Quiznachmittag	166
Anhang 28: User Study Testszenario «Gedächtnistraining» (Zyklus 1)	169
Anhang 29: Dialog Quiz (Zyklus 1)	172
Anhang 30: User Sample (Zyklus 1)	176
Anhang 31: Auswertung (Zyklus 1)	177
Anhang 32: User Study Szenario «Kennenlernen» (Zyklus 2)	178
Anhang 33: Selbsttest Kennenlernen (Zyklus 2)	182
Anhang 34: Proto-Dialog Kennenlernen (Zyklus 2)	184
Anhang 35: Dialoganalyse Grafische Darstellung (Zyklus 2)	185
Anhang 36: Dialoganalyse Frageverteilung (Zyklus 2)	186
Anhang 37: Auswertung: Affirmations (Zyklus 2)	187
Anhang 38: Auswertung: Fragetypen (Zyklus 2)	189
Anhang 39: Auswertung: Active Listening Ruleset (Zyklus 2)	192
Anhang 40: Semantisches Differential (Zyklus 3)	193
Anhang 41: Adjektivische Beschreibung (Zyklus 3)	194
Anhang 42: Adjektive konsolidiert (Zyklus 3)	195
Anhang 43: System Persona (Zyklus 3)	196
Anhang 44: Dialogdrehbuch (Zyklus 4)	197
Anhang 45: MVP (Zyklus 4)	200
Anhang 46: Auswertung (Zyklus 4)	201
Anhang 47: Testszenario Stimmen (Zyklus 5)	202
Anhang 48: Auswertung Stimmen (Zyklus 5)	205
Anhang 49: Ergebnis Stimmen (Zyklus 5)	223
Anhang 50: Charakterliche Attribute & Eigenschaften (Zyklus 5)	224
Anhang 51: Usability Testszenario (Zyklus 6)	225

Anhang 52: Auswertung Finale (Zyklus 6)230

Anhang 1 Definitive Aufgabenstellung per 21.05.2023

Thema der Arbeit: Interaktion mit natürlicher Sprache

Wie wird das Thema NLU (Natural Language Understanding) in Bezug auf Voice Commerce angewendet. Dabei gehen wir der Frage nach, wie kann ein Mensch in deutscher Sprache Produkte online bestellen, ohne die Zuhilfenahme anderer Interaktionsmethoden.

Allgemeine Angaben

Auftraggeber: Fachhochschule Ostschweiz OST, Oberseestrasse 10, 8640 Rapperswil

Kontaktperson: Prof. Dr. Markus Stolze

Studierende: Dennis Eitner, Sandro Pezzutto

Kontext der Arbeit

Seit einiger Zeit findet Voice Interaction bereits Anwendung, aber insbesondere seit 2014 wird durch Amazon die Spracherkennung in Verbindung mit Online-Shopping vorangetrieben. Im englischsprachigen Raum ist die Nutzung von Voice Interaction verbreiteter. Laut einer repräsentativen Studie von PWC aus dem Jahr 2018, bei der 1000 US-Bürger befragt wurden, gaben 72% der Teilnehmer an, Voice Assistenten zu verwenden. Darüber hinaus gaben 50% an, dass sie wöchentlich oder monatlich mit Hilfe von Voice Assistenten einkaufen.

Projektziele: Aufgabenstellung und Ziele der Arbeit

Prototypen für VUI-Basierter Shopping Assistent

Ziel dieser Arbeit ist eine auf Dialogen basierten interaktive Anwendung für den Online-Einkauf auf Basis umfangreichem User-Research zu konzipieren und dieses Konzept iterativ mittels Test von Prototypen zu verfeinern. Das Einkaufen findet dabei ausschliesslich über ein Voice-Interaktion statt. Im Rahmen dieser Masterarbeit werden zur Konzeptentwicklung Mensch-zu-Mensch-Interaktionen dokumentiert, natürlichsprachliche Dialoge erhoben. Zur Verfeinerung des Konzepts werden dann Experimente mit simulierten und implementierten Mensch-Computer-Interaktionen durchgeführt. Die Vision ist ein Online-Shopping, welches nur mit der Stimme ohne Zuhilfenahme anderer Eingabemethoden funktioniert. Vorteile und Nachteile zu unterschiedlichen Modalitäten des Outputs des Systems (nur Voice, Voice-und-Visuell, etc) werden im Rahmen des Projektes analysiert. Eine Fokussierung auf eine Modalität (z.B. nur Voice-Output) ist möglich.

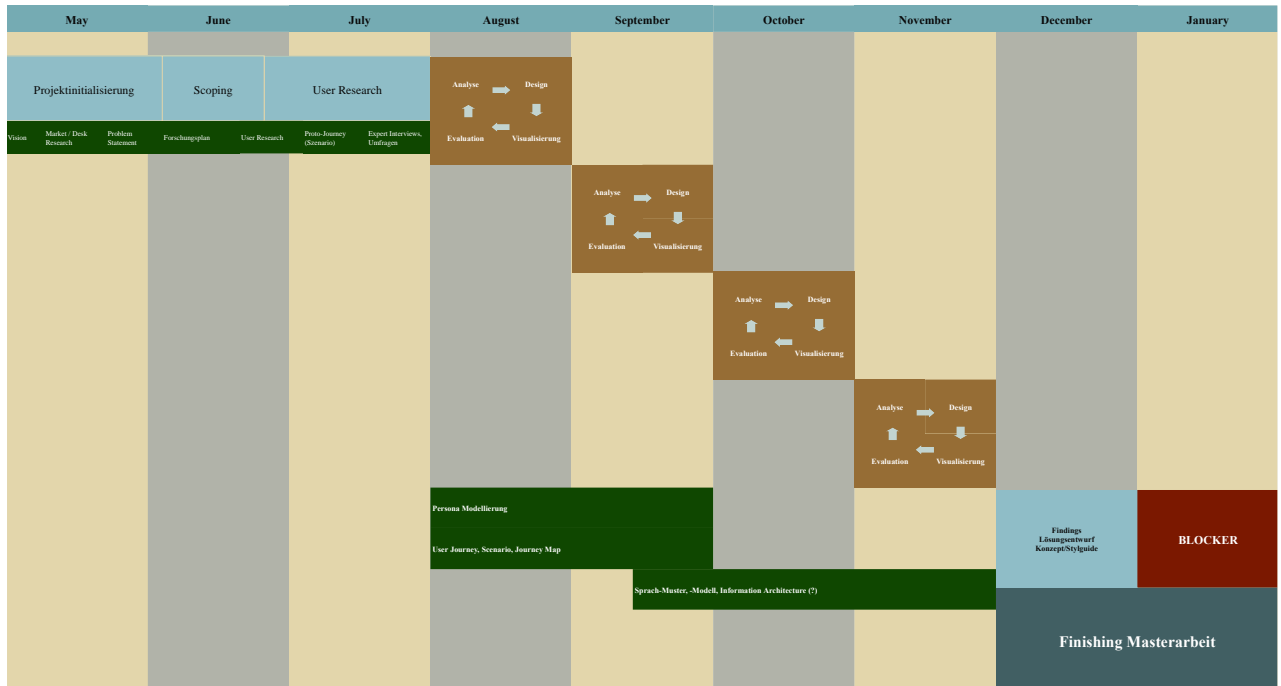
Untersuchung natürlicher Interaktion anhand des oben beschriebenen Use Cases:

Im Laufe dieser Masterarbeit werden Dialogmuster zwischen Nutzenden und dem System beschrieben. Dabei wird anhand von zwischenmenschlichen Dialogen die ideale Nutzung eines Systems beschrieben. Unter anderem werden Fragen beantwortet, welche Formulierungen Menschen verwenden, welche Dialog-Komponenten verwendet werden, wie natürlich ein System mit einem Menschen interagieren darf oder soll und wie fehlertolerant Interaktionen mit Menschen sein dürfen. Rückmeldungen des Systems (Antworten, Rückfragen, Ausgabe von Informationen) sollen in einer Form erfolgen, wie sie auch ein Mensch geben würde. Die Sprache ist Hochdeutsch, und es sollte auch ohne grammatikalische Korrektheit möglich sein, eine Konversation zu führen.

Anhang 2 S.M.A.R.T-Ziele

	Spezifisch	Messbar	Attraktiv	Realistisch	Terminiert
	Was wollen wir erreichen? S	Woran erkennen wir, dass wir die Ziele erreicht haben M	Wie kann das Ziel erreicht werden? A	Welche Ressourcen benötigen wir um das Ziel zu erreichen R	Wann erreichen wir das Ziel T
Projektziel	Prototyp mit VUI-Commerce Interaktion mit natürlicher Sprache	natürlichsprachliche Interaktionen sind dokumentiert	Erhebung der Dialoge Experimente zur Überprüfung/Validierung	Nutzende für Research Nutzende für User Test / Experimente Experten	definierter User Task kann natürlichsprachlich abgehandelt werden
Lernziel	Durchführung von Experimenten Forschung in natürlichsprachlichen Dialogen	Dialogmuster sind erkennbar	Conversation / Dialog Design konzeptionelles Interaction Design	Experimente zur Überprüfung/Validierung Nutzende für User Test / Experimente	August 2023
Forschungsziel	Webinhalte natürlichsprachlich überbringen User Tasks im Web können natürlichsprachlich abgehandelt werden NLU	Dialogmuster sind erkennbar	Qualitative Inhaltsanalyse Conversation / Dialog Design schematischer Prozesslauf basierend auf User Research	Nutzende für User Test / Experimente	August 2023
Produktziel	Vision: Online Shopping nur mit Stimme (und Augen) in deutscher Sprache	Codierter Prototyp Es ist möglich die Spezifikationen für ein VUI zu definieren	Conversation / Dialog Design schematischer Prozesslauf basierend auf User Research	User Qualitätskriterien Persona Partner (?)	November 2023
Endziel	Master Arbeit	Bericht mit 50-100 Seiten (plus Anhang)	festhalten von Vorgehen, Ergebnissen und Reflexion	min. 2x300 h	26.1.24

Anhang 3 Projektplan

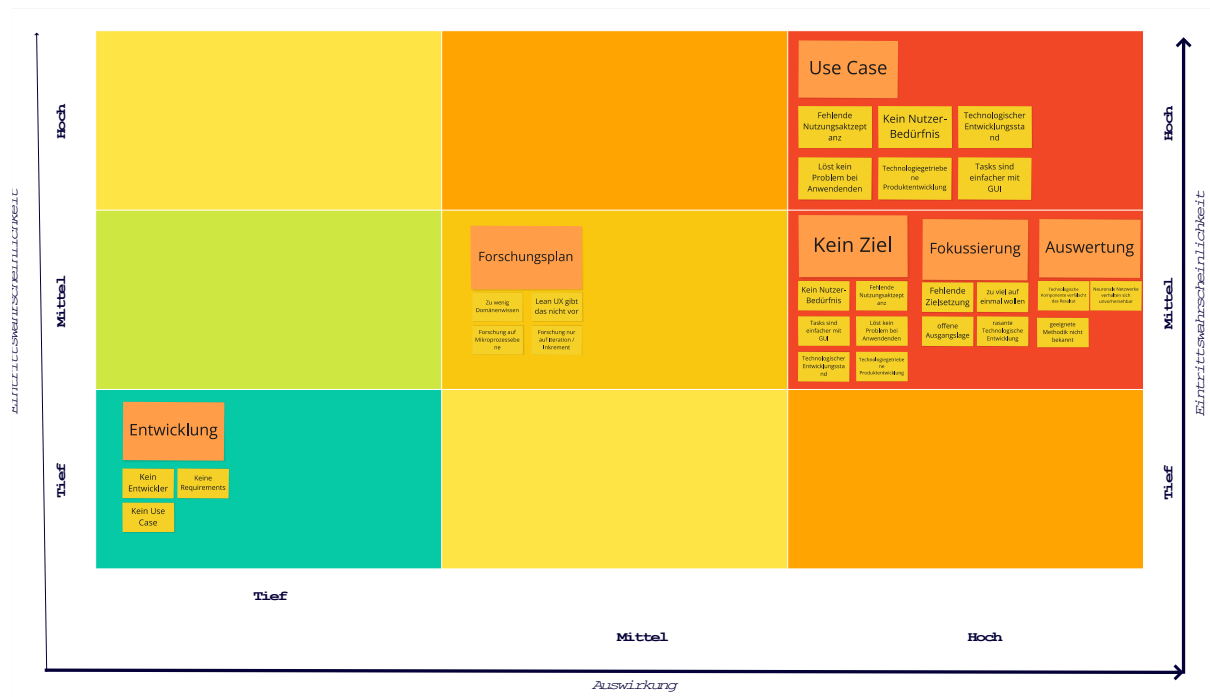


Anhang 4 Risikoanalyse I

Risikoanalyse pre-start 13.05.23

	Ziel	Projektaktivität	Auslöser Ursache	Gegenmassnahmen	Auswirkung	Eintrittswahrscheinlichkeit	RiskScore	Verantwortlich
Kein Ziel	Produktentwicklung	Research	Unvollständiges Verständnis	Regelmässige Kommunikation	6 6 6	3 6 6	36	
			Falscher Use Case	Regelmässige Kommunikation	3 3 3	6 6 6	18	
			Keine Probleme	Regelmässige Kommunikation	3 1 1	3 1 1	3	
			Unvollständiges Verständnis	Regelmässige Kommunikation	3 6 3	1 3 1	3	
			Unvollständiges Verständnis	Regelmässige Kommunikation	6 6 6	1 1 1	6	
			Unvollständiges Verständnis	Regelmässige Kommunikation	1 1 1	1 3 1	1	

Anhang 5 Risikoanalyse I – Matrix



Anhang 6 Experteninterview R.D. (Zyklus 0)

Person: R. D.

Datum: 9.6.23 11.00 Uhr

Interviewer: Dennis Eitner, Sandro Pezzutto

- Seit Januar 2023 ist er bei D. und verantwortlich für alle Produkte, die im Aftersales-Bereich eingesetzt werden, z. B. Kundendienst, Help Assistant, Interaktionen mit dem Kunden nach einem Kauf. Die Einheit wurde neu gegründet.
- Zuvor arbeitete er bei S. im Bereich der natürlichen Sprachverarbeitung als Product Owner. Er trug maßgeblich bei zum Voice Assistant auf der TV-Box, dem Chatbot auf der Website und dem Voice Bot bei Anrufen auf der Helpline.
- Beim Voicebot müssen keine Tasten mehr gedrückt werden.
- R.D ist erst seit Januar im E-Commerce-Bereich tätig und baut sein Wissen seither auf.
- D. erhält viele Anfragen, die vor dem Kauf stattfinden und auf das Thema Beratung abzielen.
- Die Bearbeitung der Anfragen übernehmen auch Personen, die im Kundendienst arbeiten. Jedoch sollte dies von den Personen, die die Produkte betreuen, getrieben werden.
- Sobald ein intelligenter Assistent zum Einsatz kommt, müssen diese die Anfragen zu Produkten, die von den Kunden kommen, verstehen. Entweder soll dieser automatisiert die Anfragen an die richtige Stelle weiterleiten oder, wenn möglich, sogar direkt beantworten können.
- D. hat zu diesem Thema ein Experiment am Laufen. Die Produktdatenbank von D. wird in eine künstliche Intelligenz aggregiert und mittels GPT-3.5 beantwortet. Beispielsweise könnte die Anfrage lauten: "Kannst du mir einen Computer empfehlen, der weniger als 350 CHF kostet und Videos rendern kann?", und die Antwort lautet dann: "Nein, das gibt es nicht." Die Fähigkeiten sind erschreckend gut, das Ganze funktioniert aber nur auf Texteingabe und nicht mit natürlicher Sprache im menschlichen Sinne.
- D. möchte keinen Telefonservice anbieten. Kunden können über das Kontaktformular einen Rückruf anfordern.
- Gesprochene Sprache wird nicht bevorzugt; der Chat mit Text hat Vorrang, da er einen interaktiven Charakter hat. Die Kommunikation per E-Mail ist langsam.
- Das Telefon wird immer von Menschen bedient und nicht von Botsystemen. Tastaturgesteuerte Anrufsysteme performen nicht so gut wie AI-gesteuerte Anliegenerkennung.
- Anrufe werden dorthin geleitet, wo die Organisation das Gefühl hat, dass der Kunde am besten aufgehoben ist. Der Kunde kennt die internen Prozesse der Organisation nicht und möchte lediglich ein Problem gelöst haben. Unternehmen stülpen ihr eigenes Organisationsmodell den Kunden über. Der Kunde wählt häufig die falsche Option und wird dann von einem Call-Agenten zum anderen weitergeleitet.
- Bei der Umstellung auf ASR- und Classifier-basierte Lösungen konnte dieser Prozess deutlich effizienter gestaltet werden. 25% Verbesserung.
- Welche Initiativen gibt es auf Voice Interaction? Siehst du Möglichkeiten für Anwendungsbereiche?

- Für Galaxus wurde das noch nicht in Betracht gezogen. Er sieht vor allem Potential darin, Anrufe an die richtige Stelle zu leiten. Wenn Inbound-Calls angeboten werden, kann er sich das vorstellen. Weniger vorstellbar ist ein Alexa-Skill, der Shopping-Carts befüllt. Es gibt z. B. beim iPhone zu viele Attribute. Bis ein System diese disambiguiert hat, ist es für den User umständlich.
- Es ist viel einfacher für den User, ein GUI zu benutzen. Auch die Bestellung via VUI für Milch ist schwierig, da es viele Arten von Milch gibt.
- D. hat 4 Millionen Artikel. Disambiguieren ist schwierig. ASR ist nicht fehlerfrei und sieht dies mit heutiger Technologie als sehr große Herausforderung.
- Je kleiner das Sortiment ist, desto eher kann es funktionieren. Sobald Attribute auf einem Produkt vorhanden sind, z. B. Telefonspeichergröße, kann dies auf anderen Produkten verwendet werden, die ebenfalls diese Attribute haben.
- Vorstellbar ist, wenn Kunden nicht einzelne Artikel bestellen, sondern immer wieder die gleichen Artikel bestellen.
- Beispielsweise bei Subway. 5 verschiedene Größen, 10 verschiedene Zutaten, wenn sich ein System die Vorlieben eines Kunden merkt, kann es über die Zeit auch den Kunden viel besser führen.
- So wie es ein Bäcker macht, wenn du den Laden betrittst. Bei Verbrauchsartikeln würde er die Chance für Voice Interaction auf ein erträgliches Maß reduzieren.
- Ab wann scheitert die Nutzung bei ChatGPT-3 bei D.? Es wurde nur auf dem Sortiment Laptops getestet. Das System erkennt kontextfremde Anfragen und liefert dann eine brauchbare Antwort, anstatt dem Kunden zu sagen, dass es diese Frage nicht beantworten kann.
- ChatGPT kann bereits jetzt geschickt nach Fragen wie Bildschirmgröße oder Tastatur antworten und darauf Empfehlungen abgeben. Es ist nicht native ChatGPT. Es ist eine ChatGPT-3.5-Implementierung.
- Woran im Moment gearbeitet wird:
- Wie kann das auf alle Artikel ausgeweitet werden?
- Wie können relevantere Vorschläge generiert werden?
- Vorschläge genauer maßschneidern auf die Fragen, die der Kunde gestellt hat.
- Der Bot ist recht schnell bei Disambiguierungsfragen. Dennoch werden weitere Produkte aus der Datenbank empfohlen, was obsolet ist.
- Im Moment wird Benchmarking auf Basis von vergangenen Dialogen gemacht. Die Resultate liegen nicht vor. Alles wird textbasiert gelöst.
- Der Grund dafür ist, dass Spracherkennung per se noch nicht gut funktioniert. Bei Swisscom wurde ein Sprachmodell entwickelt, das Deutsch und Schweizerdeutsch versteht und recht gut in der Open Domain funktioniert. In der Swisscom-Domäne wurde die schweizerdeutsche Sprache besser verstanden.
- Englische Begriffe, vor allem in der D.-Domäne, werden falsch transkribiert. Das ist ein ernstes Problem. Gewisse englische Begriffe in der Telecom-Branche wurden nicht erkannt. In der NLU wurde das per Skript abgeändert, damit die Anfrage beantwortet werden kann. Das Problem ist nicht gelöst.
- Wie gut kann Siri (oder irgendein System mit Spracherkennung) eine Einkaufsliste erstellen? Das Telefon wird sich mit der Zeit auf die Stimme und Sprache des Besitzers einstellen, weshalb der Task besser funktioniert als bei einem cloudbasierten System.

- Handys berücksichtigen Daten aus dem Gerät, z. B. Kontaktdaten oder Kalendereinträge.
- In einem Shop kann ein Kundenverständnis aufgebaut werden. Und das kann ein Telefon auch. Je mehr Informationen man mitgeben kann, desto besser funktioniert es dann.
- Er war nicht groß in Benchmarking involviert bei Swisscom. Swisscom war im deutschsprachigen Raum führend. Auch im Austausch mit deutschen und österreichischen Firmen stellten wir fest, dass Swisscom führend war.

Welche Prototyping-Tools:

Wir haben viel mit Wizard of Oz gearbeitet. So wurden Chat-Simulationen gemacht. Das hat gut funktioniert, vor allem ob die Dialoge richtig sind, konnte gut geprüft werden. Design des Happy Path, dann Fehlerbilder. Diese wurden dann in Testings mit Wizard of Oz geprüft und festgestellt, dass Nutzende andere Fragen stellen, wenn ein Fehler kommt. Und wenn sinnlose Gegenfragen gestellt wurden. Es ist recht aufwendig. Braucht viele Personen und viel Zeit, um die Erhebung durchzuführen.

Was könnte funktionieren? Crowdbasiertes, keywordbasiertes Experiment. Aufzeichnen, was für Feedback kommt. Ist aber nichts für Konservative.

Anhang 7 User Study Evaluierung Sprachassistenten (Zyklus 0)

Checkliste

- Einverständniserklärung
- Lläuft die Aufzeichnung

Intro für die Probanden

Erklärung an den Probanden, dass nicht er getestet wird, sondern er das System testet:

- Die Aufgabe/Zielvorgabe ist dann erfüllt, wenn das Resultat für dich OK ist.
- Wenn es dir zu mühsam wird, kann jede Aufgabe jederzeit abgebrochen werden.
➔ Du entscheidest, wann eine Aufgabe beendet wird. Unabhängig davon, ob die Aufgabe erfüllt werden konnte oder nicht!
- Wenn eine Aufgabe nicht erfüllt werden kann, ist das ein Scheitern des Systems, nicht des Testers.

Testszenario:

- Das Ziel ist, für alle Aufgaben möglichst nur Voice nutzen
- Ausgaben sollen möglichst auch ausgesprochen werden
- Versuche alles mit Stimme/Sprache zu lösen – und schaue wie weit du damit kommst.

Einverständniserklärung

Name des Teilnehmers: _____

1. **Zweck der Studie:** Ich verstehe, dass ich gebeten wurde, an einer Forschungsstudie teilzunehmen, die darauf abzielt, die Benutzererfahrung und Funktionalität von Voice Interfaces (wie Siri, Alexa und Google Assistant) zu bewerten.
2. **Verfahren:** Ich bin darüber informiert worden, dass die Studie die Durchführung einer Reihe von Aufgaben unter Verwendung dieser Voice Interfaces umfasst und meine Interaktionen aufgezeichnet werden können. **Diese Aufnahmen werden ausschließlich zur Analyse verwendet.** Nach Beendigung der Studie werden die Aufnahmen gelöscht.
3. **Freiwillige Teilnahme und Abbruch:** Ich verstehe, dass meine Teilnahme an dieser Studie freiwillig ist und dass ich jederzeit ohne negative Folgen aussteigen kann.
4. **Vertraulichkeit:** Ich wurde darüber informiert, dass alle Informationen, die während dieser Studie gesammelt werden, vertraulich behandelt werden und dass alle persönlichen Identifikatoren vor der Analyse entfernt werden.

Indem ich unten unterschreibe, erkläre ich, dass ich die oben genannten Bedingungen verstanden habe und freiwillig an dieser Studie teilnehme.

Name des Teilnehmers: _____

Unterschrift des Teilnehmers: _____

Datum: _____

Information zum Probanden

Name	
Alter	
Geschlecht	
Wohnort	
Beruf	

Check-in Fragen

Fragen	Antworten
Welches Smartphone benutzt du?	
Wie würdest du deine technischen Fähigkeiten auf einer Skala von 1-5 einschätzen?	
Nutzt du ChatGPT oder hast du Erfahrungen mit AI-Assistenten?	
Nutzt du sprachliche Dienste wie Google, Alexa, Siri?	
Hast du Einschränkungen in der Sensorik? Z. B. Brille, Hörgerät?	

Runde 1: einfache Fragen

Nach jeder Frage:

- Was ist dir Aufgefallen?
- Welches war die beste Antwort?

Task	Siri	Google	Bing
Wer ist amtierender Fußballmeister der Schweiz?			
Finde heraus, wie das Wetter morgen wird in Zürich			
Wer war der Regisseur von Forrest Gump?			
Wie viele Menschen gibt es auf der Welt?			

Nach der ersten Runde:

- Hast du erkannt, dass das System noch zuhört?
 - Woran, oder warum nicht?
- Woran hast du erkannt, dass die Anfrage verstanden wurde?

Runde 2: «einfache» Tasks (mit einigen Vorgaben)

Nach jeder Aufgabe:

- Was ist dir Aufgefallen?
- Welches war die beste Antwort?
- Wie zufrieden bist du mit den Antworten gesamthaft für jede VUI-Persona (1- gar nicht 5 sehr gut)

Task	Siri	Google	Bing
Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich. Kannst du einen Tisch reservieren? Kannst du einen Kalender-Event erstellen?			
Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.			

Check-out Fragen

- Wie war das für dich? Beschreibe in mit [3,4,5] Schlüsselworten deine Erfahrung
- Wie fühlten sich die Dialoge an?
- Hast du dich verstanden gefühlt? Welcher Assistent war deiner Meinung nach am besten?
- Hast du die Antworten verstanden?
- Würdest du den Antworten trauen/sich auf sie verlassen?
 - Wenn nein, warum nicht?
- In welcher Sprache würdest du am liebsten mit einem Voice Assistenten sprechen?
- Für was würdest du ein Voice User Interface nutzen wollen? Für was nicht?
- Was ist deine Ansicht bezüglich Datenschutzes?
- Welche 3 Dinge haben dir besonders gefallen? Welche nicht?
- Gibt es sonst noch etwas, was du uns sagen möchtest? Hast du noch Fragen?

Anhang 8 Auswertung (Zyklus 0)

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anzahl Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Boris	1	Siri	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Hey Siri	Ja	Ja	BSC Young Boys	1	Ja	Nein	0	Wikipedia Suchergebnis
Boris	1	Google	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Ja	Ja	BSC Young Boys	3	Ja	Nein	0	2 versuche, spricht dann langsamer-Suchergebnis
Boris	1	Bing	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Ja	Ja	BSC Young Boys	1	Ja	Ja	Bing	-
Jerome	1	Siri	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Hey Siri	Ja	Ja	BSC Young Boys	3	Ja	Nein	0	generell skeptisch gegenüber sprachassistenten, existiert nicht in seinem Leben
Jerome	1	Google	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Hey Google	Nein	Ja	BSC Young Boys	3	Ja	Ja	Google	Fragt nach Rangliste - weiss am schluss nur die Platzierung - YB wird als Leader per voice ausgegeben
Jerome	1	Bing	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Ja	Nein	FC Basel	1	Nein	Nein	0	fragt nach erster Antwort wider nach Rangliste - kommt ohne Angabe zum Jahr Irritiert vom Resultat - würde den anderen glauben, weil beide das gleiche Resultat

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Sven	1	Siri	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Hey Siri	Ja	Ja	BSC Young Boys	1	Ja	Nein	Siri	websuche, Ist sich Siri gewohnt, deshalb ist Siri am besten
Sven	1	Google	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Ja	Ja	BSC Young Boys	1	Ja	Nein	0	websuche, sport.de
Sven	1	Bing	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Ja	Nein	FC Zürich	1	Nein	Nein	0	Fängt an zu lachen, als er die Stimme von Bing hört, findet sprachausgabe nicht so toll, Schweizerdeutsch nicht lupenrein
Marisa	1	Siri	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	hey siri	Ja	Ja	nimmt suchresultat von sport.de Zeigt bei 3.Versuch YB	3	Ja	Nein	Siri	Unsicher bei erstem Resultat Versuch nochmals angepasste Anfrage 3.versuch
Marisa	1	Google	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Nein	Nein	keine bei voice Anfrage Young Boys bei konventioneller Suche	5	Nein	Nein	0	versteht keine Anfrage - Probandin bricht ab und sucht konventionell
Marisa	1	Bing	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	fragt direkt	Ja	Ja	YB	1	Ja	Nein	0	technische Probleme mit Sprachausgabe vom Testgerät
Jonas	1	Siri	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Knopfdruk auf Seite	Ja	Ja	Websuche	2	Ja	Nein	0	Nutzt Websuche Resultat - geht auf wikipedia

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Jonas	1	Google	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Ja	Ja	Websuche	5	Ja	Nein	0	Stellt 3mal die gleiche Frage 4 versuch fragt welche sprache google spricht - google gibt anweisung wie die Sprache geändert werden kann Nutzt auch wikipedia aus websuche resultat
Jonas	1	Bing	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Ja	Ja	Amtierender Meister von 2020/21	1	Ja	Ja	Bing	weiss nicht ob bing schon zuhört beeindruckt und bemerkt nicht das falsche Jahr welches ausgegeben wird - antwort kommt gleich, muss nicht selber suchen. Und ch-de hat überrascht meister von 2021/2022
Marcel	1	Siri	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Hallo Siri	Nein	Nein	zürich (voice) Suchresultat web	6	Nein	Nein	0	fragt bei fehlgeschlagenen Versuchen ob "sie" nicht mit ihm spricht/versteht wechselt auf Schlagworte - präzisiert mit Jahr Nicht zufriedem mit der Auswahl, obwohl dieses Resultat stimmt, da nicht voice out

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Marcel	1	Google	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	ohne	Nein	Ja	Spielergebnisse (Versuch 3) Länderspielresultat (7) Tabelle	8	Ja	Ja	0	Erst normale Anfrage, wechselt auf Schlagworte google bringt 5tes Resultat während der Fragenstellung Anfrage nach Tabelle kommt mit voice
Marcel	1	Bing	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Hallo?	Ja	Nein	Sei noch nicht bekannt - letzter Meister YB	1	Nein	Ja	Bing	Schlagwortsuche, Bing war am schnellsten
Moreno	1	Siri	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Hey Siri	Ja	Ja	BSC Young Boys	3	Ja	Nein	Siri	
Moreno	1	Google	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Button	Ja	Ja	Nein	3	Ja	Nein	0	
Moreno	1	Bing	Wer ist amtierender Fußballmeister der Schweiz?	Aktuelles	Hallo Bing	Nein	Nein	Keine Antwort	6	Nein	Nein	0	Versucht mit Hallo Bing, Versucht mehrmals in den Sprechpausen von Bing Fragen zu stellen, Sieht den zustand des zuhörens nicht
Boris	2	Siri	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Hey Siri	Ja	Ja	teilweise bewölkt	1	Ja	Nein	0	-
Boris	2	Google	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Knopfdruck auf Seite	Ja	Ja	28grad	1	Ja	Ja	0	Zeigt an, und liest vor

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Boris	2	Bing	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	13-27 grad	1	Ja	Ja	Bing	Sehr ausführliche Antwort, Fast zu viel, Hat länger gedauert. Gut zusammengestellt,
Jerome	2	Siri	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Hey Siri	Ja	Ja	Nur grobes Wetter, ohne Temperatur	3	Ja	Nein	0	muss Prompt und Frage jeweils wiederholen. Wechselt in den Befehlsmodus. Dritter Versuch mit sehr knappen Angaben. Auf Nachfragen (Temperatur) kommt mehr, aber zum aktuellen Tag
Jerome	2	Google	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	Sonning, 13-26grad	1	Ja	Ja	0	ohne Probleme
Jerome	2	Bing	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	fragt direkt	Ja	Ja	laut wetter.com - sehr ausführlich	1	Ja	Ja	Bing	zeigt sich ein wenig beeindruckt - ganz klar beste Antwort. Ausführlich und informativ
Sven	2	Siri	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Knopfdruck auf Seite	Ja	Ja	wolkenlos morgen...	1	Ja	Nein	0	Zeigt Zürich Visuell an, findet gemischte ausgabe sehr mühsam, benutzt voice auf dem fahrrad
Sven	2	Google	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	13-27°	1	Ja	Ja	0	Spricht lupenreines hochdeutsch mit google

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Sven	2	Bing	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	14°-26°	1	Ja	Ja	0	Bezieht sich auf witkon, sagt ich hoffe das hilft ihnen weiter
Marisa	2	Siri	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Knopfdruck auf Seite	Ja	Ja	ohne temperature	1	Ja	Nein	0	kurz und bündig sei Roboter-mässige Antwort
Marisa	2	Google	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	teilweise bewölkt mit Temperaturern zwische 18 und 31 Grad	1	Ja	Ja	Google	kurz und bündig sehr angenehm
Marisa	2	Bing	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	ausführlich mit Hinweis auf zusätzliche Quelle/Tabelle	1	Ja	Ja	0	ch-de ist ein bisschen komisch - stört sich daran
Jonas	2	Siri	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Knopfdruck auf Seite	Ja	Ja	teilweise bewölkt	1	Ja	Nein	Siri	für Wetter reicht was siri und google machen
Jonas	2	Google	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	teilweise bewölkt, 18-31 Grad	1	Ja	Ja	Google	-
Jonas	2	Bing	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	Ausführlich sonnig 15-28 C mit Wassertabelle	2	Ja	Ja	Bing	Crazy! Bemerkt, dass voice output andere Temperatur als auf dem Screen kam
Marcel	2	Siri	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Knopfdruck auf Seite	Ja	Ja	sehr kurze Antwort	1	Ja	Nein	0	mit vollständigem Satz "Antwort war sehr schnell"
Marcel	2	Google	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	fragt direkt	Ja	Ja	Ausführlich	1	Ja	Ja	0	mit vollständigem Satz

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
													"funktionierte zügig"
Marcel	2	Bing	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	fragt direkt	Ja	Ja	Ausführlich & detailliert	1	Ja	Ja	Bing	Schlagwort Anfrage aussprache CH-DE ein bisschen komisch
Moreno	2	Siri	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Knopfdruck auf Seite	Ja	Ja	laut vorhersage überwiegend bewölkt	5	Ja	Nein	0	Will auf die Seite drücken um Siri zu aktivieren, dabei setzt er das iOS in Lock-Modus und muss es zuerst entsperren
Moreno	2	Google	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	teilweise bewölkt mit Temperaturen zwischen 19 und 91 Grad	1	Ja	Ja	0	
Moreno	2	Bing	Finde heraus, wie das Wetter morgen wird in Zürich	Aktuelles	Button	Ja	Ja	ausführlich	1	Ja	Ja	Bing	Schlechte aussprache
Boris	3	Siri	Wer war der Regisseur von Forrest Gump?	Wissen	Knopfdruck auf Seite	Ja	Ja	Robert Zemeckis	2	Ja	Ja	Siri	Schnelle response
Boris	3	Google	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	Google	Hat nur namen genannt.
Boris	3	Bing	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	0	Suchseiten als Vorschlag, Aussprache Englisch/Deutsch), Bing hat länger gebraucht, darum schlechter

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Jerome	3	Siri	Wer war der Regisseur von Forrest Gump?	Wissen	fragt direkt	Ja	Ja	Robert Zemeckis, ganzer Satz	1	Ja	Ja	0	bewertet alle Antworten als gleich gut
Jerome	3	Google	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	0	-
Jerome	3	Bing	Wer war der Regisseur von Forrest Gump?	Wissen	fragt direkt	Ja	Ja	Robert Zemeckis, ganzer Satz	1	Ja	Ja	0	-
Sven	3	Siri	Wer war der Regisseur von Forrest Gump?	Wissen	Knopfdruck auf Seite	Ja	Ja	Robert Zemeckis	1	Ja	Ja	0	-
Sven	3	Google	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	0	.
Sven	3	Bing	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	0	betonung von bing ist lustig
Marisa	3	Siri	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis, ganzer Satz	1	Ja	Ja	0	
Marisa	3	Google	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	0	Anfrage nur stichwortartig
Marisa	3	Bing	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Ausführliche Antwort	2	Ja	Ja	0	zweite Anfrage klappt - bricht die ausführliche Antwort ab findet Antwort mühsam und am schlechtesten findet Hinweis das bing sucht nach der gestellten Anfrage überflüssig Dialekt ist mühsam

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Jonas	3	Siri	Wer war der Regisseur von Forrest Gump?	Wissen	Knopfdruck auf Seite	Ja	Ja	Websuche	2	Ja	Nein	Siri	Wer hat den Film gedreht? Geht wieder auf wikipedia- reicht link zu wiki
Jonas	3	Google	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Nein	Nein	keines	4	Nein	Nein	0	genervt als google ihn schon wieder nicht versteht spricht langsamer, aber immer noch in vollständigen Sätzen Ist erstaunt, dass seine Anfrage auf dem Screen richtig angezeigt wird
Jonas	3	Bing	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Ausführliche Antwort	1	Ja	Ja	0	lacht wegen schwenglischer Aussprache
Marcel	3	Siri	Wer war der Regisseur von Forrest Gump?	Wissen	Knopfdruck auf Seite	Ja	Ja	Robert Zemeckis, ganzer Satz	1	Ja	Ja	Siri	
Marcel	3	Google	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	Google	
Marcel	3	Bing	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemekis	5	Ja	Ja	0	Video schauen - ohne weiteren Kommentar...
Moreno	3	Siri	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Nein	Nein	Websuche	1	Nein	Nein	0	Proband stellt die Frage nicht ein zweites mal, obwohl siri ihn falsch verstanden hat.
Moreno	3	Google	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	0	

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Moreno	3	Bing	Wer war der Regisseur von Forrest Gump?	Wissen	Button	Ja	Ja	Robert Zemeckis	1	Ja	Ja	0	
Boris	4	Siri	Wie viele Menschen gibt es auf der Welt?	Wissen	Hey Siri	Ja	Ja	7.846 Mrd im Jahr 2021 (ganzer Satz)	1	Ja	Ja	0	Bis auf Million genau
Boris	4	Google	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	78842268, auf den einer genau	1	Ja	Ja	0	Bis auf den einer genau
Boris	4	Bing	Wie viele Menschen gibt es auf der Welt?	Wissen	Knopfdruck auf Seite	Ja	Ja	8.3 Mrd Menschen	1	Ja	Ja	0	Mit genauerem Datum, Aktuellere Information
Jerome	4	Siri	Wie viele Menschen gibt es auf der Welt?	Wissen	Knopfdruck auf Seite	Ja	Ja	7.846 Mrd im Jahr 2021 (ganzer Satz)	1	ja	Ja	0	.
Jerome	4	Google	Wie viele Menschen gibt es auf der Welt?	Wissen	ohne	Ja	Ja	78842268, auf den einer genau	1	Ja	Ja	0	voice nur die Zahl - Datenstand/Jahr visuell
Jerome	4	Bing	Wie viele Menschen gibt es auf der Welt?	Wissen	ohne	Ja	Ja	ganzer Satz, rund 8.02 Mrd	1	Ja	Ja	Bing	Bing am besten weil am aktuellsten (2023), google 2021, Siri ohne Datum und quelle
Sven	4	Siri	Wie viele Menschen gibt es auf der Welt?	Wissen	Hey Siri	Ja	Ja	7.846 Mrd im Jahr 2021 (ganzer Satz)	1	Ja	Nein	0	ohne textausgabe, website, countrymeeters.com
Sven	4	Google	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	Auf den einer genau	1	Ja	Ja	0	-
Sven	4	Bing	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	Nov. 22, 8.03 Mrd Menschen	1	Ja	Ja	0	findets furchtbar wie die zahlen ausgesprochen werden

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Marisa	4	Siri	Wie viele Menschen gibt es auf der Welt?	Wissen	Knopfdruck auf Seite	Ja	Ja	7.846 Mrd im Jahr 2021 (ganzer Satz)	4	Ja	Ja	0	Anfrage nur stichwortartig Antwort 2021 - fragt nach für 2023 fragt nach ob Suchresultat gesagt werden kann formuliert Anfrage nochmals neu -> Die Siri Standardantwort (Stand 2021) kommt per voice, nachhaken für aktuellen Stand nur websucheresultat.
Marisa	4	Google	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	aktuelles resultat im dritten anlauf	3	Ja	Ja	0	Anfrage mit Schlagworten - nicht verstanden Formuliert Satz - nicht verstanden Formuliert Satz nochmals sehr langsam
Marisa	4	Bing	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	ausführlich und aktuell (2023)	1	Ja	Ja	Bing	Anfrage mit Schlagworten
Jonas	4	Siri	Wie viele Menschen gibt es auf der Welt?	Wissen	Knopfdruck auf Seite	Ja	Ja	7.846 Mrd im Jahr 2021 (ganzer Satz)	3	Ja	Ja	0	Frage nach für aktuell Frage nochmals komplett für aktuellSiri machte nochmals neuen Vorschlag auf nachfrage. Reicht so (websuche)

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Jonas	4	Google	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	2021, 7.8... mrd	2	Ja	Ja	0	präzisiert die Anfrage beim zweiten mal auf aktuell und bekommt die gleiche Zahl nochmals.
Jonas	4	Bing	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	Aktuelle Daten 28.2.23	1	Ja	Ja	Bing	bing schlechter Dialekt bing nervt, hat die Anfrage aber gut inhaltlich gut gelöst
Marcel	4	Siri	Wie viele Menschen gibt es auf der Welt?	Wissen	Knopfdruck auf Seite	Ja	Ja	7.846 Mrd im Jahr 2021 (ganzer Satz)	1	Ja	Ja	Siri	
Marcel	4	Google	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	8mrd	1	Ja	Ja	Google	
Marcel	4	Bing	Wie viele Menschen gibt es auf der Welt?	Wissen	fragt direkt	Ja	Ja	ausführlich (28.2.23)	1	Ja	Ja	Bing	bricht die Ausführliche Antwort ab
Moreno	4	Siri	Wie viele Menschen gibt es auf der Welt?	Wissen	Knopfdruck auf seite	Ja	Nein	Websuche	2	Nein	Nein	Siri	Kandidat prüft nicht die ergebnisse der websuche
Moreno	4	Google	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	7.88 Mrd (2021)	1	Ja	Ja	Google	Fragt nach für Stand 2023 - bekommt keine Antwort mehr und bricht hier ab, da das Resultat ja mit Jahresangabe auch ok ist.

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Moreno	4	Bing	Wie viele Menschen gibt es auf der Welt?	Wissen	Button	Ja	Ja	Ausführlicher aktueller Stand 8.02 Mrd	1	Ja	Ja	Bing	Alle etwa gleich - Bing am ausführlichsten mit sachen die man gar nicht angefragt hat
Boris	5	Siri	Du möchtest mit (deiner Freundin oder einem Freund) essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Hey Siri	Ja	Ja	Santa Lucia	1	Ja	Nein	0	-
Boris	5	Google	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	Don Leone	1	Ja	Nein	0	-
Boris	5	Bing	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	2 Antworten, La Pasta, Frascati	1	Ja	Ja	0	Gibt Name und Adresse der Restaurants an, wünscht guten appetit
Jerome	5	Siri	Du möchtest mit (deiner Freundin oder einem Freund) essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Web	Ja	Nein	4 von 5 Rest. mit offensichtlich schweizer Küche an - präzisiert Anfrage sehr ausführlich	3	Nein	Nein	0	würde nach zweiter Anfrage sofort den Assistenten wechseln
Jerome	5	Google	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	2 Resultate	1	Ja	Nein	0	Resultate zufriedenstellend
Jerome	5	Bing	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	Beste italinische restaurants..	1	Ja	Ja	0	5 angefragt, 2 bekommen
Sven	5	Siri	Du möchtest mit (deiner Freundin oder einem Freund) essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Knopfdruck Seite dann Hey Siri	Ja	Nein	Das sind meine Suchergebnisse	2	Nein	Nein	0	Gibt Frage in einem satz auf, link auf apple map, restaurant in schlieren, link auf

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
													website, Montag bis Freitag, g
Sven	5	Google	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	erste beide sind meta suchergebnisse, dann kommt restaurant	1	Ja	Nein	0	Formuliert die frage um (mit top ratings), sieht nicht ob restaurant offen ist, Schaut nach, Sa/So geschlossen
Sven	5	Bing	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	Frascata, Valacati	1	Ja	Ja	Bing	stellt andere frage, aber umfangreiche frage, 2 vorschläge, System sagt, dass es keine Reservierungen tätigen kann
Marisa	5	Siri	Du möchtest mit deinem einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Knopfdruk auf Seite	Ja	Ja	Suchresultat	2	Ja	Nein	Siri	Unterbricht während der Anfrage mit falscher Rest. Auswahl präzisiert auf italienisches Siri - reagiert nicht auf Reservation, aber immerhin Kalender
Marisa	5	Google	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Nein	zwei Einträge "in der nähe"	3	Ja	Nein	0	erstes Resultat nicht zufriedenstellend - präzisiert nochmals für in der Statd ZHpräzisiert Anfrage auf Kreis 4

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Marisa	5	Bing	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	La Pasta, mit Details	2	Ja	Ja	Google	bing unterbricht mit vorlesen der Anfrage und liefert Resultat - bestes in Feusisberg Proband unterbricht und will Resultat in Zürich
Jonas	5	Siri	Du möchtest mit (deiner Freundin oder einem Freund) essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Knopfdruck auf Seite	Ja	Ja	Suchergebnisse	3	Ja	Nein	0	Das hier habe ich gefunden - zeigt Suchergebnisse an
Jonas	5	Google	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	Nur Anzeige des suchresultats	1	Ja	Nein	0	-
Jonas	5	Bing	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	gibt Vorschläge mit Details zur Location und Öffnungszeiten aus	1	Ja	Ja	Bing	lacht während er sich die Vorschläge vorlesen lässt, bing macht die besten Vorschläge. Spricht aber sehr viel
Marcel	5	Siri	Du möchtest mit (deiner Freundin oder einem Freund) essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Knopfdruck auf Seite	Ja	Ja	Websuche Auswahl	3	Ja	Nein	0	fragt gleich für eine Tischreservation - wechselt wieder auf schlagwort suche
Marcel	5	Google	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	zwei Vorschläge	1	Ja	Nein	0	schlagwort suche
Marcel	5	Bing	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	Ausführlich mit Beschreibung, Adresse etc.	1	Ja	Ja	0	Schlagwort suche wählt sich eines aus und will Anrufen

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Moreno	5	Siri	Du möchtest mit (deiner Freundin oder einem Freund) essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Hey Siri	Ja	Ja	-	3	Ja	Nein	0	Will spezifisch fleisch essen. Bricht nach 3 Versuchen ab.
Moreno	5	Google	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	La Bottega di Mari	1	Ja	Nein	0	Bei den falschen Ausgaben, wird zu protokoll gegeben, dass er eschneller sei mit anrufen
Moreno	5	Bing	Du möchtest mit deiner Freundin oder einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.	Aufgabe	Button	Ja	Ja	Gran Sasso	1	Ja	Ja	Bing	Gute Auswahl an Resultateten, Bei den falschen Ausgaben, wird zu protokoll gegeben, dass er eschneller sei mit anrufen
Boris	5.1	Siri	Kannst du einen Tisch reservieren?	Aufgabe	Button, dann Frage	Ja	Nein	Zeigt Fahrroue an	1	Nein	Nein	0	Fragt ob siri eine reservation gemacht hat.
Boris	5.1	Google	Kannst du einen Tisch reservieren?	Aufgabe	Rückfrage	Ja	Ja	Hier ist das Top Suchergebnis aus dem Netz	5	Nein	Nein	0	Proband stellt rückfrage. Gibt auf die Frage nach telefonnummer suchergebnisse zurück,
Boris	5.1	Bing	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Nein	Ich bin ein chatbot... schlägt vor anzurufen mit Telnr.	1	Nein	Nein	0	Gibt Telefonnummer an (richtig)

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Jerome	5.1	Siri	Kannst du einen Tisch reservieren?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-
Jerome	5.1	Google	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Ja	Unter anderen gibt es das hier im Internet	1	Nein	Nein	0	beim ersten Versuch gleich bei der Reservation gelandet Beide Anfragen wurde per voice gut verstanden - Ausgabe als suchresultat nur auf Screen - weiter mit voice hat geklappt
Jerome	5.1	Bing	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Ja	Link zu restaurant	1	Nein	Nein	0	Link zur Reservation beim gewünschten Restaurant
Sven	5.1	Siri	Kannst du einen Tisch reservieren?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	nicht probiert	nicht probiert
Sven	5.1	Google	Kannst du einen Tisch reservieren?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	nicht probiert	nicht probiert
Sven	5.1	Bing	Kannst du einen Tisch reservieren?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	nicht probiert	nicht probiert
Marisa	5.1	Siri	Kannst du einen Tisch reservieren?	Aufgabe	Knopfdruck auf seite	Ja	Nein	bringt Hinweis, dass bereits ausser haus ist am gewünschten Termin	1	Nein	Nein	0	Siri fragt nach ob trotzdem ein Termin erstellt werden soll

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Marisa	5.1	Google	Kannst du einen Tisch reservieren?	Aufgabe	Button	Nein	Nein	internet resultat, nachfolgende Anfragen ohne Kontext	3	Nein	Nein	0	beim dritten versuch wird Website des Restaurants ausgegeben würde abbrechen und selbst erledigen
Marisa	5.1	Bing	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Nein	kann nicht reservieren - schlägt vor Anzurufen mit Telnr. Und eine erinnerung für den Kalender zu erstellen	1	Nein	Nein	Bing	Anfrage für reservation und kalender in einer, gut, liefert immerhin Tel oder alternativen
Jonas	5.1	Siri	Kannst du einen Tisch reservieren?	Aufgabe	kein	Nein	Nein	Zeigt weitere Suchergebnisse an in Richterswil	3	Nein	Nein	0	will mit zweiter Anfrage nochmals Suchergebnisse verbessern - Siri unterbricht mit Resultat (lacht) versucht nochmals versucht auf Basis neuer Resultate nochmals zu reservieren - Bommt immer nur suchergebnisse glaubt siri will nicht reservieren und bricht ab
Jonas	5.1	Google	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Nein	Gibt website des gewünschten Restaurants aus	3	Nein	Nein	0	bei zweiter Anfrage für Reservation verweist wieder auf die Website - hat sich für ihn so erledigt Bei dritter Anfrage verliert google den Kontext

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Jonas	5.1	Bing	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Nein	chatbot kann keine Res. Vornehmen - schlägt vor er soll anrufen	1	Nein	Nein	Bing	Fragt in CH Deutsch Reservation an macht nach der Antwort nicht weiter - "lassen wir das" bing macht die besten Vorschläge. Spricht aber sehr viel
Marcel	5.1	Siri	Kannst du einen Tisch reservieren?	Aufgabe	fragt direkt	Nein	Nein	routenberechnung oder Suchergebnisse	5	Nein	Nein	0	versucht eine Reservation für seine Auswahl zu treffen muss nochmals von vorn beginnen weil siri beendet und kontext verloren hat letzter Versuch mit "Santa Lucia Anrufen" löst Anruf aus
Marcel	5.1	Google	Kannst du einen Tisch reservieren?	Aufgabe	fragt direkt	Ja	Ja	detailinformation zum Restaurant	2	Nein	Nein	0	schlagwort "Tisch reservieren" "Anruf Restaurant" klappt
Marcel	5.1	Bing	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Nein	siehe video	3	Nein	Nein	0	bing fragt zurück ob anrufen soll!?! - Proband sagt Ja gerne - bing sagt ruft an und bittet um einen Moment Geduld. Weiteres im Video

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Moreno	5.1	Siri	Kannst du einen Tisch reservieren?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	nicht probiert	nicht probiert
Moreno	5.1	Google	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Nein	Im Internet habe ich das hier gefunden	1	Nein	Nein	0	Zeigt Unihockey Website an
Moreno	5.1	Bing	Kannst du einen Tisch reservieren?	Aufgabe	Button	Ja	Nein	Ich bin ein Chatbot...	2	Nein	Nein	0	
Boris	5.2	Siri	Kannst du einen Kalender-Event erstellen?	Aufgabe	Hey Siri	Ja	Ja	-	1	Ja	Nein	0	Wurde erwartet, dass System den Begriff "Meine Frau" kennt, Zeigt Kontakte an mit namens "Hauser"
Boris	5.2	Google	Kannst du einen Kalender-Event erstellen?	Aufgabe	Hey Google	Ja	Nein	Hier ist das Top Suchergebnis aus dem Netz	1	Nein	Nein	0	Top Suchergebnis
Boris	5.2	Bing	Kannst du einen Kalender-Event erstellen?	Aufgabe	Button	Ja	Ja	fragt nach Infos um einen Kalendereintrag zu erstellen	2	Nein	Nein	0	Nach der zweiten aufzählung spricht proband während system weiter spricht, gibt instruktionen an nutzer ab, Nutzer wiederholt anfrage, Sagt, system hätte einen kalendereintrag erstellt (Stimmt nicht)
Jerome	5.2	Siri	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-
Jerome	5.2	Google	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Jerome	5.2	Bing	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-
Sven	5.2	Siri	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	nicht probiert	nicht probiert
Sven	5.2	Google	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	nicht probiert	nicht probiert
Sven	5.2	Bing	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	nicht probiert	nicht probiert
Marisa	5.2	Siri	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	nicht explizit nachgefragt - Siri bezog sich bei der Reservation-Anfrage schon auf den Kalender
Marisa	5.2	Google	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	
Marisa	5.2	Bing	Kannst du einen Kalender-Event erstellen?	Aufgabe	Button	Ja	Nein	kann keinen Kalendereintrag machen - schlägt vor eine Erinnerung zu schicken und verlangt Email	1	Nein	Nein	0	fragt nochmals für Kalendereintrag bricht nach dieser Antwort ab.
Jonas	5.2	Siri	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-
Jonas	5.2	Google	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Jonas	5.2	Bing	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	merkt an, dass ohne erfolgreiche Reservation der Kalendereintrag keinen Sinn macht - Reservieren können alle nicht
Marcel	5.2	Siri	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	
Marcel	5.2	Google	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	
Marcel	5.2	Bing	Kannst du einen Kalender-Event erstellen?	Aufgabe	fragt direkt	Ja	Nein	Gibt alle Gewünschten Infos inkl. Telnr. per voice aus	2	Nein	Nein	Bing	beendet nach zweimaligem Anfragen für Anruf, den bing jetzt nicht mehr machen kann/will. bing - überrascht von den ausführlichen Beschreibungen
Moreno	5.2	Siri	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	nicht probiert	nicht probiert
Moreno	5.2	Google	Kannst du einen Kalender-Event erstellen?	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	
Moreno	5.2	Bing	Kannst du einen Kalender-Event erstellen?	Aufgabe	Button	Ja	Nein	Ich bin ein Chatbot...	1	Nein	Nein	0	
Boris	6	Siri	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-
Boris	6	Google	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Boris	6	Bing	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	nicht probiert	nicht probiert	nicht probiert	nicht probiert	0	Nein	Nein	0	-
Jerome	6	Siri	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Knopfdruck auf Seite	Ja	Ja	Websuche	2	Ja	Nein	0	klickt sich durch die Websuchresultate verfeinert mit zweiter Anfrage zufrieden, weil Resultate von vertrauten Plattformen (Galaxus) angezeigt werden.
Jerome	6	Google	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Ja	top-Ergebnis Internet	1	Ja	Nein	0	gleiches Resultat wie Siri am schluss
Jerome	6	Bing	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Ja	Websuche Resultate	1	Ja	Ja	0	nutzt trotz voice output die suchresultate auf dem Screen - voice output in dem Anwendungsfall für den Probanden nicht besonders relevant, da er bereits produkte aus den vorhergehenden tests hatte
Sven	6	Siri	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Knopfdruck auf seite	Ja	Nein	Resultat der Websuche	1	Nein	Nein	0	Zuerst Ebay, würde auf testbericht-seite wechseln
Sven	6	Google	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Nein	website output, ebay, amazon	1	Ja	Nein	0	Suchanfrage enthält stativ, zeigt aber zuerst kameras an.

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Sven	6	Bing	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Ja	schlägt zwei stativ vor	1	Ja	Ja	0	haben ads in den suchresultate, 2x digitec
Marisa	6	Siri	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Knopfdruck auf Seite	Ja	Ja	Suchresultat, Websuche	5	Nein	Nein	0	fragt nach was Siri beim ersten Versuch rausgesucht hat nochmals Anfrage verlangt Produkt info (anstelle Geschäft) ändert nochmals Anfrage ändert nochmals Anfrage bewertet Resultat als zu wenig verlässlich resultate zwischendurch out of context - schlecht
Marisa	6	Google	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Nein	Websuche mit Shops	3	Nein	Nein	0	erster Versuch Websuche mit Shops zweiter erst im zweiten anlauf verstanden fragt nach voice output - versteht nicht. Proband bricht ab hier (löscht mir ab so), Google am schlechtesten
Marisa	6	Bing	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Ja	liest Optionen von Digitec vor inkl. Preis	1	Ja	Ja	Bing	Resultat ok - fragt nach dem Stativ mit bester Bewertung nach. Antwort dann nicht mehr relevant (versucht herauszufinden)

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
													wieviel geht) bing - findet Resultat ok, aber würde sich nicht darauf verlassen
Jonas	6	Siri	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Knopfdruck auf seite	Nein	Nein	darauf habe ich keine Antwort...	3	Nein	Nein	0	versucht drei verschiedene Fragenformulierungen und bricht dann ab
Jonas	6	Google	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Nein	Suchresultat	2	Nein	Nein	0	präzisiert Anfrage - erstes Suchresultat liefert Stativ, aber nicht spezifisch für diese Kamera Scrollt durch suchresultate - sind aber Objektive Beendet Task mit Kommentar "Pfeife" hat einfach irgendetwas angezeigt
Jonas	6	Bing	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Ja	Trägt drei Optionen vor	1	Ja	Ja	0	hört sich alles an, ist amüsiert von der Aussprache div. Fremdwörter und Abkürzungen "bing hat Dialekt mit einem Migrartionshintergrund" bing eindeutig am besten mit den 3 Vorschlägen, die anderen sind kläglich gescheitert

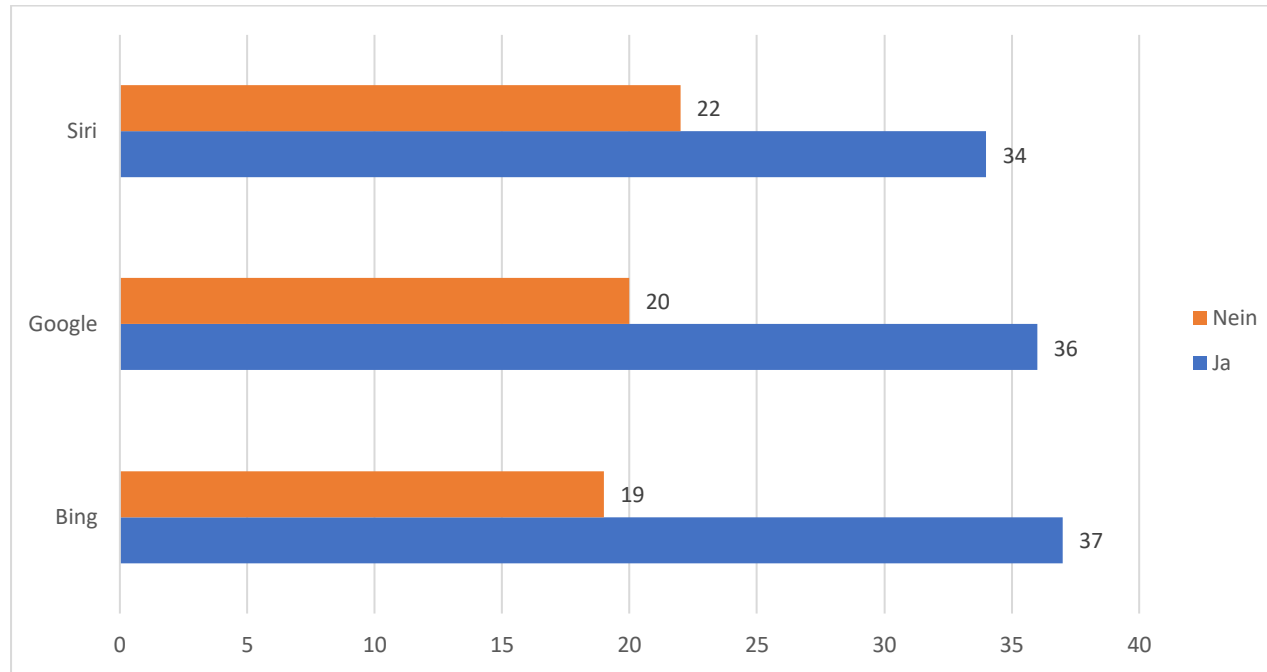
Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
Marcel	6	Siri	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Knopfdruck auf Seite	Ja	Ja	Websuche Resultate	4	Ja	Nein	Siri	Schlagwort Suchest OK mit Weblinks - wählt digitec-Link und sucht dort weiter - findet hat hier auswahl und ist gut. zweite Anfrage: falsches Suchresultat dritte "findet keine App" vierte wieder Websuche - checkt kurz Ebay und findet OK siri & google haben beide beim erstmal die Anfrage verstanden und Suchresultate geliefert - und schnell
Marcel	6	Google	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Ja	Websuche Resultate	1	Ja	Nein	Google	Schlagwortsuche etwa gleiche Resultate wie Siri - funktioniert
Marcel	6	Bing	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	fragt direkt	Ja	Ja	Ausführliche Infos den Vorschlägen	1	Ja	Ja	0	Schlagwortsuche - bing interpretiert als "sony stativ kaufen" ähnliche Auswahl wie zuvor - Proband ist sich generell nicht sicher ob er so das richtige Stativ findet. bing stimmt nicht überein mit seiner

Testperson	Frage-No.	Assistent	Frage, Aufgabe	Fragenkategorie	Wie wird das System in den "Zuhören-Modus" versetzt?	Wurde die Frage vom System richtig verstanden?	Korrekte Antwort des Probanden?	Antwort des Systems oder Probanden	Anz. Versuche	Task erfolgreich	Ohne GUI lösbar?	Beste Antwort	Beobachtungen
													Anfrage - am schlechtesten
Moreno	6	Siri	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Knopfdruck auf Seite	Ja	Ja	Websuche	3	Ja	Nein	Siri	Gibt sich mit suchresultaten zufrieden
Moreno	6	Google	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Nein	Youtube,	2	Nein	Nein	0	google schwach
Moreno	6	Bing	Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	Aufgabe	Button	Ja	Ja	Vorschläge, Manfrotto	2	Ja	Ja	0	Vorschläge von Heiniger - auf Nachfragen alternative Kaufmöglichkeit mit u.A. Digitec ist OK

Anhang 9 Pivot-Tabelle Task Erfolgreich (Zyklus 0)

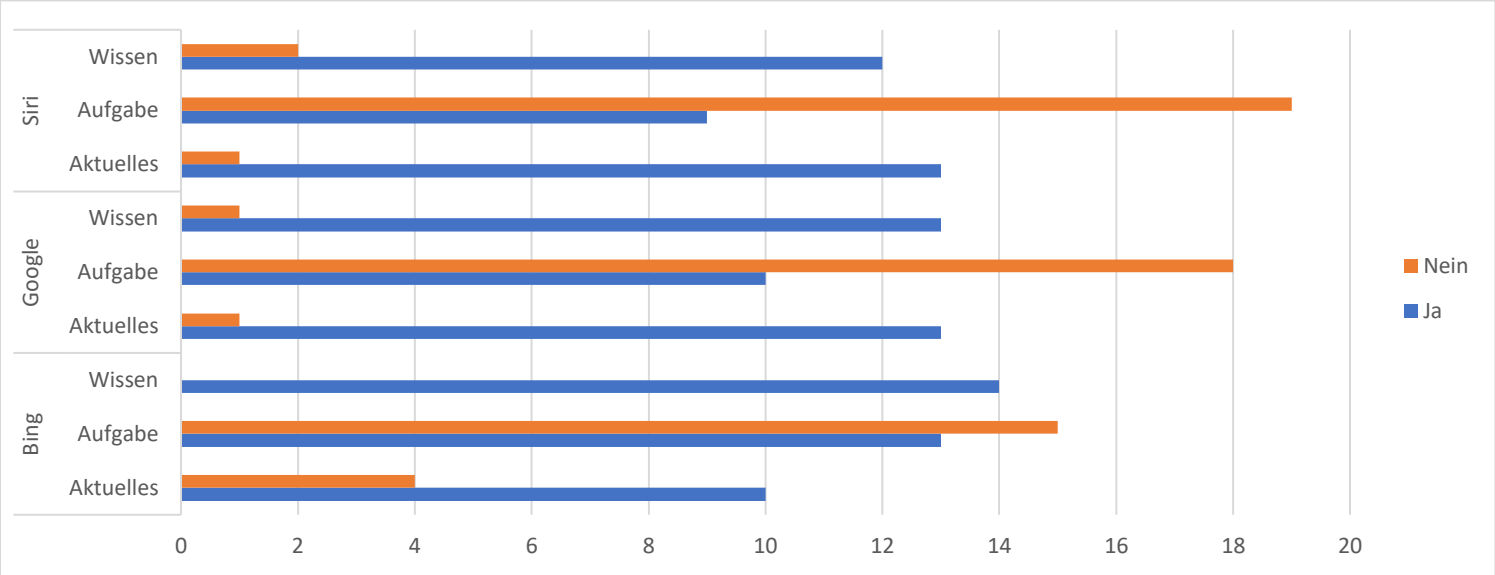
Count of erfolgreich	Task	Task erfolgreich				
Assistant	Ja		Nein	Grand Total	Erfolgsrate	
Bing			37	19	56	66.07%
Google			36	20	56	64.29%
Siri			34	22	56	60.71%
Grand Total			107	61	168	63.69%

Anhang 10 Grafik Task Erfolgreich (Zyklus 0)



Anhang 11 Pivot Tabelle Task Erfolgreich nach Kategorie (Zyklus 0)

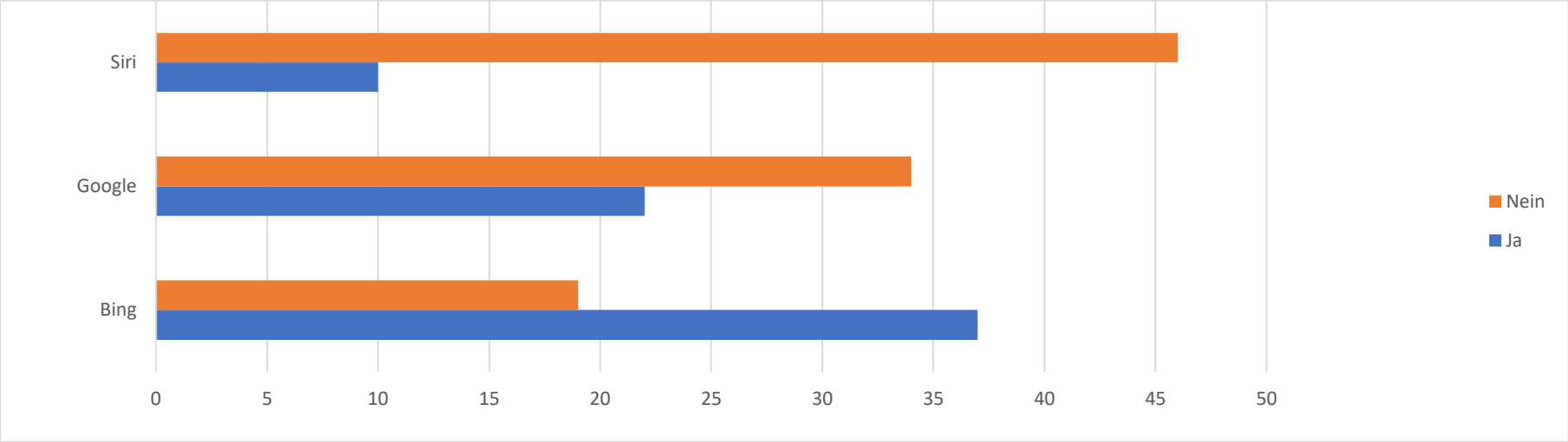
Tasks Erfolgreich nach Kategorie					
Count of Task erfolgreich	Column Labels				
Row Labels	Ja	Nein	Grand Total	Erfolgsrate	
Bing		37	19	56	66%
Aktuelles		10	4	14	71%
Aufgabe		13	15	28	46%
Wissen		14		14	100%
Google		36	20	56	64%
Aktuelles		13	1	14	93%
Aufgabe		10	18	28	36%
Wissen		13	1	14	93%
Siri		34	22	56	61%
Aktuelles		13	1	14	93%
Aufgabe		9	19	28	32%
Wissen		12	2	14	86%
Grand Total		107	61	168	64%



Anhang 12 Pivot Tabelle - Ohne GUI Lösbar (Zyklus 0)

Count of Ohne GUI lösbar?						
Row Labels	Column Labels	Ja	Nein	Grand Total	Erfolgsrate	
Bing			37	19	56	66%
Google			22	34	56	39%
Siri			10	46	56	18%
Grand Total			69	99	168	41%

Anhang 13 Grafik - Ohne GUI Lösbar (Zyklus 0)



Anhang 14 Pivot Tabelle Frage vom System verstanden (Zyklus 0)

Count of Wurde die Frage vom System richtig verstanden?		Column Labels			
Row Labels	Ja	Nein	nicht probiert	Grand Total	
Bing	50	1	5	56	98%
Google	43	5	8	56	90%
Siri	41	5	10	56	89%
Grand Total	134	11	23	168	80%

Anhang 15 Pivot Tabelle Korrekte Antwort (Zyklus 0)

Count of Korrekte Antwort des Probanden?		Column Labels			
Row Labels	Ja	Nein	nicht probiert	Grand Total	
Bing	39	12	5	56	70%
Google	37	11	8	56	66%
Siri	35	11	10	56	63%
Grand Total	111	34	23	168	66%

Anhang 16 Pivot Tabelle Anzahl Versuche (Zyklus 0)

Column Labels												
	Bing			Google			Siri			Total Min. von Anz. Versuche	Total Max. of Anz. Versuche 2	Total Summe von Anz. Versuche
Row Labels	Min. von Anz. Versuche	Max. of Anz. Versuche	Summe von Anz. Versuche	Min. von Anz. Versuche	Max. of Anz. Versuche 2	Summe von Anz. Versuche	Min. von Anz. Versuche	Max. of Anz. Versuche 2	Summe von Anz. Versuche	von	Anz.	
Du möchtest ein Stativ für ein Sony Alpha. Suche und finde eines.	0	2	7	0	3	10	0	5	18	0	5	35
Du möchtest mit (deiner Freundin oder einem Freund) essen gehen. Finde ein italienisches Restaurant in Zürich.							1	3	15	1	3	15
Du möchtest mit deinem einem Freund essen gehen. Finde ein italienisches Restaurant in Zürich.							2	2	2	2	2	2
Du möchtest mit deiner Freundin oder einem	1	2	8	1	3	9				1	3	17

Freund essen gehen. Finde ein italienisches Restaurant in Zürich.													
Finde heraus, wie das Wetter morgen wird in Zürich	1	2	8	1	1	7	1	5	13	1	5	28	
Kannst du einen Kalender-Event erstellen?	0	2	6	0	1	1	0	1	1	0	2	8	
Kannst du einen Tisch reservieren?	0	3	9	0	5	15	0	5	10	0	5	34	
Wer ist amtierender Fußballmeister der Schweiz?	1	6	12	1	8	28	1	6	19	1	8	59	
Wer war der Regisseur von Forrest Gump?		1	5	12	1	4	10	1	2	9	1	5	31
Wie viele Menschen gibt es auf der Welt?		1	1	7	1	3	10	1	4	13	1	4	30
Grand Total	0	6	69	0	8	90	0	6	10	0	8	259	
									0				

Anhang 17 Risikoanalyse II von 8.7.23

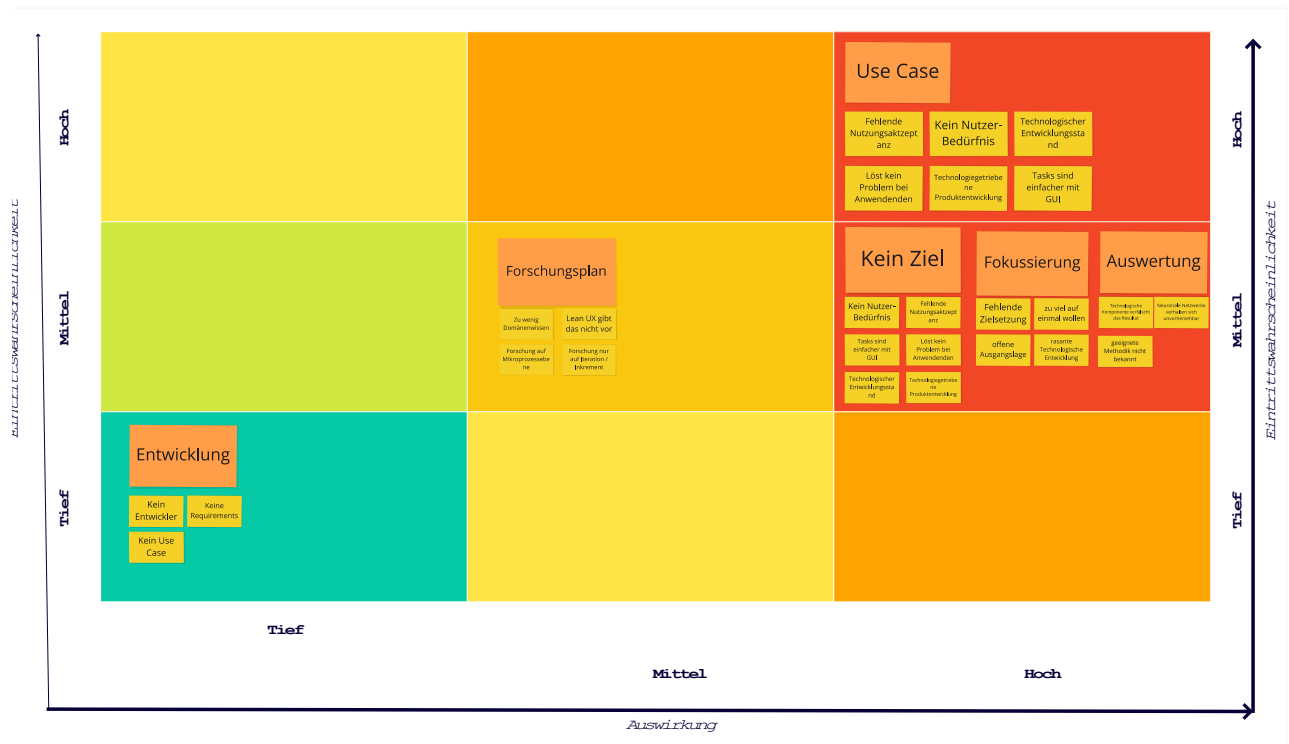
Risikoanalyse

	Ziel	Projektaktivität	Auslöser	Ursache	Gegenmassnahmen	Auswirkung	Eintrittswahrscheinlichkeit	RiskScore	Verantwortlich
Kein Ziel	Produktentwicklung Masterarbeit	Research	Kein Use Case	Keine Anforderung Keine Anforderungen	Workaround Auftraggeber	6 6 2	3 3 2	18	
Produktplan	Strukturplan	Research	Keine Anforderungen	Kein Use Case Keine Anforderungen	Keine Anforderungen Keine Anforderungen	3 6 2	3 3 2	9	
Fokussierung	Produktentwicklung Anforderungserhebung	Research	Fehlende Anforderungen	Keine Anforderungen Keine Anforderungen	Iteration Hypothesen	6 6 2	3 6 2	18	
Use Case	Produktentwicklung	Research	Keine Anforderungen	Keine Anforderungen Keine Anforderungen	Keine Anforderungen Keine Anforderungen	3 6 2	3 6 2	36	
Auswertung	Konzeptentwicklung	Usability Test	Keine Anforderungen	Keine Anforderungen Keine Anforderungen	Experimentell Coaching	3 3 2	3 6 2	18	
Entwicklung	Prototypenentwicklung	Design/Prototypen	Kein Entwurf	Keine Anforderungen Kein Use Case		1 3 2	3 3 2	9	

Unmittelbare Massnahme

Auswertung fertig machen	Ergebnisse präsentieren	Persona mit Need finden (Nico?)	bessere Experience entwickeln
--------------------------	-------------------------	---------------------------------	-------------------------------

Anhang 18 Risikomatrix II



Anhang 20 Aktualisierte Aufgabenstellung (MT23-G05) per 24.07.2023

In Folge des Workshops vom 14.07.2023 mit dem Auftraggeber ergab sich eine veränderte Ausgangslage. Dadurch wird eine Überarbeitung der initial formulierten Aufgabenstellung nötig. Der Kontext wird anhand der Workshop-Ergebnisse angepasst. Die grundlegenden Ziele der Arbeit ändern sich nicht (siehe Projektziele).

Thema der Arbeit: natürlichsprachliche Interaktion in der Domäne «Ältere Menschen»

Wie kann das Thema natürlichsprachliche Interaktion in der Domäne «Ältere Menschen» angewendet werden? Dabei gehen wir der Frage nach: Wie kann ein Voice-Assistent ältere Menschen zur Unterhaltung oder Unterstützung dienen?

Allgemeine Angaben

Auftraggeber: Fachhochschule Ostschweiz OST, Oberseestrasse 10, 8640 Rapperswil

Kontaktperson: Prof. Dr. Markus Stolze

Studierende: Dennis Eitner, Sandro Pezzutto

Kontext der Arbeit

Ältere Menschen haben oft physische Einschränkungen oder neigen zu psychischen Krankheiten. In Seniorenzentren ist ihr Alltag oft von Einsamkeit oder Langeweile geprägt. Dies, obwohl die Institutionen viele aktivierungstherapeutische Aktivitäten anbieten. Dazu gehören u.a. Lesezirkel, Jass-Gruppen, Gedächtnistraining, Filmnachmittage, Malen, Basten, Singen und gemeinsames Musikhören. Senioren sind auch oftmals in diversen Belangen auf Unterstützung angewiesen, sei dies in alltäglichem (z.B. Pflege) oder auch bei der Nutzung aktueller Technologien wie Smartphones oder Computer/Tablets.

Projektziele: Aufgabenstellung und Ziele der Arbeit

Prototypen für einen intelligenten, natürlich-sprachlichen Unterhaltungs-Assistenten

Das Projektteam geht der Frage nach, ob und wie mit einer Voice User Interaction Applikation nützliche Dienste zur Aktivierung oder Unterhaltung erwiesen werden können.

Ziel dieser Arbeit ist eine auf Dialogen basierte interaktive Anwendung für den Einsatz zur Unterhaltung von älteren Menschen und/oder in Institutionen. Im Rahmen dieser Masterarbeit werden zur Konzeptentwicklung Mensch-zu-Mensch-Interaktionen

dokumentiert und natürlichsprachliche Dialoge erhoben. Die Vision ist ein/e virtueller Vorleser/in, welcher primär mit der Stimme, ohne Zuhilfenahme anderer Eingabemethoden funktioniert. Vorteile und Nachteile zu unterschiedlichen Modalitäten des Outputs des Systems (nur Voice, Voice und visuell, etc.) werden im Rahmen des Projektes analysiert. Eine Fokussierung auf eine Modalität (z.B. nur Voice-Output) ist möglich.

Untersuchung natürlicher Interaktion anhand des oben beschriebenen Use Cases:

Im Laufe dieser Masterarbeit werden Dialogmuster zwischen Nutzenden und dem System erforscht. Unter anderem wird folgenden Fragen nachgegangen:

- Wie ist der aktuelle Stand der Technik?
- Gibt es Muster in den Dialogen und wenn ja, gibt es Treffer in Bezug auf relevante Stichworte?
- Welche Formulierungen verwenden Menschen?
- Welche Dialog-Komponenten werden verwendet?
- Wie nahe kommt ein AI-gestütztes System an reale zwischenmenschliche Dialoge und lassen sich dazu Metriken erstellen?
- Wir werden untersuchen, wie fehlertolerant Interaktionen mit Menschen sein dürfen. Rückmeldungen des Systems (Antworten, Rückfragen, Ausgabe von Informationen) sollen in einer Form erfolgen, wie sie auch ein Mensch geben würde.
- Was ist der Mehrwert von visueller Darstellung in der Voice-Interaction?

Die Liste der Frage ist nicht abschliessend und kann sich im Laufe des Projektes ändern.

Zu erarbeitende Artefakte können Personas, User Journey/Szenarios, ein Modell der erhobenen Dialogmuster und ein Interaktionskonzept für ein natürlichsprachliches Interface sein, das auf dem Use Case «Unterhaltung» basiert. Die noch nicht vorhersehbaren Erkenntnisse und Findings aus Research und Tests sollen aussagekräftig dokumentiert werden, was die Notwendigkeit für weitere Artefakte offen lässt. Die Interaktionssprache Gegenstand der Forschung. Es sollte möglich sein eine natürliche Konversation mit einem System zu führen.

Ein wichtiges Ziel der Arbeit ist die Beschreibung von natürlichsprachlichen Dialogmustern zwischen Nutzern und dem System.

Wichtig ist dabei, dass ein solcher Dialog wie von Mensch zu Mensch stattfinden soll. Das bedeutet im Wesentlichen, dass der Mensch seine Formulierung nicht so anpassen muss, damit das System ihn versteht. Gleichzeitig kann das System den Gesprächsverlauf auch sinnvoll steuern.

Lernziele Studierende

- Das Ziel ist, Dialogmuster zu erkennen und darauf basierend ein konzeptionelles Interactiondesign zu entwickeln. Dazu werden natürlichsprachliche Dialoge erhoben und dokumentiert, um einen Unterhaltungsassistenten zu entwickeln.
- Für die Erhebung und Validierung der Dialoge werden Experimente durchgeführt, wie z.B. Wizard of Oz Prototyping. Experimente und Tests werden so lange iterativ durchgeführt, bis sich ein natürliches Dialogmuster herausstellt.
- Es wird ein leichtgewichtiges Vorgehen bevorzugt, welches auf mehrere und kleine Iterationen setzt. So soll ein schrittweises Herantasten an eine mögliche Lösung erreicht werden. Die erarbeiteten Ergebnisse werden dabei laufend verfeinert und validiert.
- Erstellen von Spezifikationen für die Entwicklung eines Voice Assistenten.

Geheimhaltung, Rechte

Es wird kein NDA benötigt. Die Arbeit kann direkt nach Abschluss publiziert werden. Sämtliche Rechte liegen bei den Studierenden. Elemente aus den nicht öffentlichen Anhängen dürfen, nach Rücksprache, im HCID-Unterricht benutzt werden.

Zur Kenntnisnahme für den Auftraggeber enthalten:

- Masterarbeiten im MAS HCID kosten den Auftraggeber nichts, denn es handelt sich um ein studentisches Lernprojekt, dessen Ergebnis nicht garantiert werden kann. Von den Studierenden wird eine Arbeitsleistung von 300h erwartet. In dieser Zeit ist auch die Zeit für die Erstellung der schriftlichen Ausarbeitung der Masterarbeit (50 – 100 Seiten) enthalten. Die Arbeit sollte eine möglichst umfassende Dokumentation der Arbeitsresultate und des Arbeitsprozesses enthalten, sowie eine Reflexion zum Vorgehen und den Resultaten. Eine explizite Nennung des Auftraggebers in der öffentlichen Arbeit ist nicht erforderlich: Auftraggeber können anonym bleiben.
- Der Auftraggeber muss für die Arbeit die nötige Unterstützung liefern: Zugang zu Ressourcen, Personen, Endbenutzern. Konkret wurde sich auf die folgenden Ressourcen geeinigt, welche vom Auftraggeber zur Verfügung gestellt werden: ... [Die Auflistung der bereitgestellten Ressourcen sollte möglichst spezifisch sein]
- Rechte an den Arbeitsresultaten: Die mit der Arbeit veröffentlichten Resultate können ohne Einschränkung im Unterricht der OST genutzt werden. Ohne anderslautende Regelung gehören die in der Masterarbeit erarbeiteten Resultate den Studierenden. Für die in dieser Arbeit erarbeiteten Resultate gelten die folgenden Regeln: [z.B.: Die Studierenden, der Auftraggeber und die OST dürfen die erarbeiteten Resultate nutzen und weiterentwickeln. Der Ursprung der Resultate (Masterarbeit HCID der OST und die Namen der Studierenden) sollte dabei auch für Ausstehende sichtbar bleiben]
- Masterarbeiten im MAS HCID werden grundsätzlich veröffentlicht. Insbesondere muss die Arbeit am Kick-Off und an der öffentlichen Abschlusspräsentation vorgestellt werden können und bei einem Peer Review von einer anderen Gruppe gegengelesen werden können. Die schriftliche Masterarbeit wird im Regelfall kurz nach der Diplomierung der Studierenden im ePrints der OST publiziert. Hierbei ist es möglich auf die Publikation bewerteter, aber vertraulicher Anhänge zu verzichten. Die Studierenden werden darauf achten, dass keine Adressen, Telefonnummern und andere persönliche Daten von Personen ausserhalb der Studierendengruppe publiziert werden. Eine verzögerte Publikation der Arbeit (maximal 2 Jahre) kann vom Auftraggeber beantragt werden. Dies ist in der Aufgabenstellung und in der abgegebenen Arbeit (Deckblatt mit Informationen zur Geheimhaltung) explizit zu vermerken [z.B.: Der vollständige Bericht der Masterarbeit darf erst zwei Jahre nach der Diplomierung der Studierenden auf dem ePrints der OST publiziert werden.]. Auch bei verzögerter Publikation werden Titel der Arbeit, Namen der Studierenden und ein von den Studierenden zu erstellendes Abstract der Arbeit (Text, nicht PDF) relativ direkt nach der Diplomierung publiziert.
- Geheimhaltung während der Arbeit: Es ist möglich, dass Studierende, Betreuer und Co-Referent ein persönliches NDA mit dem Auftraggeber abschliessen. Eine zeitliche Beschränkung des NDAs [z.B. bis zur Veröffentlichung der Arbeit] ist notwendig. Haftungsansprüche gegen die OST sind hierbei nicht möglich.

Unterschriften

Der Auftraggeber

Rapperswil, den

Prof. Dr. Markus Stolze

Die Auftragnehmer:

Zürich, 21.7.23

Richterswil, 21.7.23

Ort, Datum

Ort, Datum



Dennis Eitner

Sandro Pezzutto

Anhang 21 Experteninterview Aktivierungstherapeutin (Zyklus 2)

Datum: 24.10.23, 14.00 Uhr

Ort: Zürich

Anreise

14. 00 Uhr. Wir melden uns bei der Cafeteria und erkundigen uns nach dem Aufenthaltsort von B.B.. Der Chef versucht, sie auf dem internen Telefon zu erreichen. Beim ersten Anruf nimmt niemand ab. Er schaut auf den grossen Info-Screen, um herauszufinden, wo sich B.B. befinden könnte. Er stellt fest, dass um 15.30 Uhr zeitgleich zwei Aktivierungsveranstaltungen stattfinden. Er ist ein wenig verwirrt und ruft nochmals an. Dieses Mal erreicht er B. und teilt uns ihren Aufenthaltsort mit. B. ist dabei Fotografien einzurahmen, welche von einem Bewohner oder einer Bewohnerin stammen, um diese auszustellen.

Interview

Administrative Fragen:

Die eigene Leistung werde dokumentiert, um den eigenen Mehrwert sichtbar zu machen. Beobachtungen, Begegnungen und Gespräche werden festgehalten. Persönliche Gespräche werden inhaltlich nicht dokumentiert.

In Aktivierungstherapien werden Inhalte und Ziele von Einzel- und Gruppengesprächen dokumentiert. Bei den Gruppenangeboten gehe es auch darum festzuhalten, welche Angebote von wie vielen genutzt würden. In einer Gruppendokumentation werden keine Inhalte über Einzelpersonen festgehalten.

Dokumentationen zu einzelnen Personen werden separat erstellt.

Aktuelles Befinden von BewohnerInnen werden einzeln dokumentiert, wodurch ein Pflegeverlauf sichtbar werde. So entstehe eine Objektivität, welche als wichtig empfunden wird. Dies seien auch Hilfestellungen für andere Rollen im Altersheim (z. B. Pflege).

Pflegerische Leistungen wie Blutzuckerspiegel messen oder Insulin spritzen, werden ebenfalls dokumentiert, da sie durch die Krankenkasse abgegolten werden können.

Betreuerische Tätigkeiten, obwohl in einem therapeutischen Umfeld, werden nicht über die Krankenkasse verrechnet. Auch deshalb sei es wichtig, die eigene Leistung zu dokumentieren.

In der Politik werde aktuell nach Lösungen gesucht, wie man betreuerische Leistungen, welche einen hohen Stellenwert haben, besser verrechnet werden können. Die erbrachte Leistung ist aktuell in der Pauschale des Pensionspreises enthalten. Der Kostendruck führt jedoch zu Sparmassnahmen. Seit der Pflegeinitiative sei der Druck auf die Löhne in der Pflege nicht mehr so stark.

Das Alterszentrum habe einen hohen betreuerischen Anteil in den Leistungen. Diesbezüglich existiere ein Leistungskatalog. Betreuerische Leistungen von Bewohnern werden über [BESA](#)

abgerechnet. Je höher die BESA-Stufe, desto höher die Abrechnung. Wenn eine Person weniger Pflegeleistungen benötigte, dafür aber mehr betreuerische Leistung, fehle dem Betrieb Geld, da diese nicht abgerechnet werden können. Beispielsweise eine Person, die sich selbst anziehen kann, jedoch demenzbedingte Schwierigkeiten in der Orientierung hat, benötigt Unterstützung in der Strukturierung des Tages. Das ist eine betreuerische Leistung und keine pflegerische. Finanzielle Aspekte seien deshalb ein wichtiger Aspekt der Arbeit.

Wie wird dokumentiert?

Die Software zur Dokumentation heisst EasyDok. Dokumentiert werde Aktivierung, Pflege, Gruppendokumentation. Die eigene Anwesenheitszeit muss nicht erfasst werden.

Eine Stunde Gedächtnistraining benötige fast eine Stunde Vorbereitung. Turnen brauche nicht so viel Vorbereitung. Während der Arbeit werde handschriftlich dokumentiert und danach die Leistung an einen Computer übertragen.

Die Dokumentation benötige ca. 1 – 1.5 Stunden pro Tag. Im Pflegebereich sei der Aufwand höher. Viel Zeit nehme vorwiegend die Vorbereitung in Anspruch.

Es werden sowohl über Einzelpersonen als auch über Gruppenaktivitäten dokumentiert.

Was passiert mit den Dokumentationen?

Diese werden im System eingetragen, damit Informationen über BewohnerInnen allen zugänglich sind. Bezugspersonen können sich über Ihre Klienten informieren.

Es wird unterschieden zwischen Pflegeleistungen und Pflegeleistungen/Medizinische Leistungen wie z.B. Insulin spritzen werden an Krankenkassen weiterverrechnet.

Betreuungs- und Therapeutische Leistungen können nicht verrechnet werden.

Ein grosser Teil werde im Pflegeverlauf festgehalten.

Das Bezugspersonensystem.

Mitarbeiter in Altersheim seien Bezugspersonen für die BewohnerInnen. Jeder MitarbeiterIn hat eine gewisse Anzahl an Bezugspersonen. Diese sollten immer nachlesen, was dokumentiert wird. Auch die Aktivierungstherapie liest nach, was Bezugspersonen dokumentieren. Somit würden Doppelspurigkeiten in Bezug auf vermittelnde Massnahmen vermieden.

Was ausser Administrativem braucht am meisten Zeit?

Die Vorbereitung für eine Stunde Aktivierungstherapie. Viel Zeit nehme auch der Austausch mit den Bewohnern in Anspruch, was jedoch der schönste Aspekt der Arbeit sei. Manchmal seien es auch nur 15-Minuten-Gespräche. Dadurch erfahre man viel über die aktuelle

Befindlichkeit des Bewohners. Auch Einzelbesuche macht sie und geht zu den Menschen ins Zimmer.

Ein interner Austausch über die BewohnerInnen findet regelmässig statt. Es gibt Rapporte und Fallbesprechungen. In Fallbesprechungen werden Probleme besprochen und gemeinsam nach Lösungen gesucht.

Aktuell findet gerade ein Wechsel in der Leitung «Betreuung und Pflege». Personelle Wechsel seien häufig. Der Leitungswechsel hat einen grossen Einfluss auf die Angestellten. Man muss sich auf einen neuen Führungsstil einlassen und an neue Menschen gewöhnen. Die Leiterin ist die Vorgesetzte der Aktivierung.

Therapeutische Fragen

Das letzte Mal waren wir bei einer Gedankenreise dabei. Was für andere Aktivierungsspiele machst du?

Sie nutzt immer die gleichen Medien und nur wenig technische Hilfsmittel. Die Aktivierung hat erst gerade kürzlich ein Laptop erhalten. Ein einziger Laptop nutze jedoch wenig. Daher kommen viel Bilder (diese werden vorher am Computer auf Papier ausgedruckt) zum Einsatz. Es kommen auch Musik, musikalische Spiele oder Wunschkonzerte zum Einsatz. Durch Musik werden Erinnerungen von früher wach.

B. macht ähnliches, jedoch mit Kochen.

Was hat man gern gegessen. Es geht ums Riechen, und woran es jemanden erinnert. Es wird auch über Rezepte und Kochweisen gesprochen «Wie macht man es richtig». Es wird gemeinsam gekocht und gegessen. Das gekochte soll auch zusammengenossen werden. Die Sinne anzusprechen sei ein wichtiger Aspekt der Arbeit.

B. geht auch viel nach draussen und nutzt dort Utensilien wie Ringe, Bälle. Dabei wird auch das Gleichgewicht trainiert, um Stürze zu vermeiden. Im Allgemeinen seien Gleichgewichtsübungen im Alter ein wichtiges Thema.

Gedankenreisen dienen auch dazu herauszufinden, an was man sich noch erinnern mag. B. betont, dass Gedächtnistraining keine Leistung sein soll und nicht an eine Schulaufgabe erinnern. Aufgaben wie Wortspiele täten dies. Gedächtnistraining solle eher anregen und «wecken». Es soll die TeilnehmerInnen dazu verleiten «dran zu bleiben» und neugierig zu bleiben. Der soziale Aspekt sei sehr wichtig. Die Frage «wie kann ich mich selbst der Gruppe einbringen» sei auch ein grosser Bestandteil der Arbeit.

B. macht ein Beispiel. Im Rahmen des Quiz-Nachmittags kam die Organisation PLO zu Sprache. Aufgrund von aktuellen Geschehnissen in Israel wurden dadurch eigene Erlebnisse aus dem Zweiten Weltkrieg geweckt. Die traumatischen Erlebnisse von damals konnten nie verarbeitet werden, da über diese nie gesprochen wurde. Dies sei den Eltern zum Teil auch vorgeworfen worden. Das eigene Verhalten gegenüber Opfer des Zweiten Weltkrieges wird reflektiert und besprochen. Solche Unterhalten gehen ihr unter die Haut. Das sind Themen über die sollen gesprochen werden. Ausgelöst durch Kriege in anderen Ländern, kommen Erinnerungen wieder hoch.

Das gegenseitige Erzählen sei wichtig. So komme ein Austausch über persönliche Erlebnisse der damaligen Zeit zustande. Der Austausch über diese Themen beeinflusse die Stimmung der BewohnerInnen. B. versuche dann in den Gesprächen das Thema in eine andere Richtung zu lenken und wieder etwas Auflockerung hereinzubringen. Durch die Gespräche über diese Themen merke man, dass das Erzählen über das Erlebte ein grosses Bedürfnis ist. Die Steuerung von Gruppendynamik empfinde sie als Herausforderung!

In der Aktivierungstherapie können viele verschiedene Mittel eingesetzt werden. Es sei wichtig vor allem mehrere Sinnesorgane ansprechen.

Es würden auch andere Spiele gemacht wie Quiz, Eile mit Weile, Schach, Jassen.

Diese Spiele haben einen hohen Erinnerungsfaktor. Auch, dass man Gewinnen will, ist ein anregender Faktor. Auch gemeinsames Filmeschauen sei Bestandteil. Dabei würden vor allem Filme von früher, mit alten Schauspielern, geschaut. Filme seien auch sehr emotional, da viele Erinnerungen wachgerufen werden. Auch historische Filme, z. B. aus dem alten Zürich werden amüsant empfunden.

Ideen für Inhalte von Aktivierungstherapien kämen aus Literatur und Austausch mit anderen. Die Ideen auf die eigenen Bedürfnisse würden angepasst. Dabei ist stets zu beachten, was kognitiv für Einzelne noch kognitiv möglich sei und was eben nicht.

Zur Vorbereitung der Stunde sei vieles schon vorhanden. Der Rest werde aus dem Internet recherchiert. Auch aktuelle Geschehnisse zu thematisieren, sei wichtig, damit man sich auch an das gestrige Geschehen erinnere. Es sollen auch Ideen von Teilnehmenden aufgenommen werden können.

Digital wird u.a. Jassen und Tetris gespielt.

B. erwähnt ein Spiel, bei welchem ein Beamer ein interaktives Bild auf einen Tisch projiziert. Das Spiel sei jedoch kein optimaler Einsatz, da es keine Sinne anspricht und den Menschen etwas vorgaukle.

Anmerkung: Das erwähnte Spiel heisst Tovertafel:

<https://www.demenz-pflegeheim.com/de/spiele-fuer-demenzkranke/?viewmode=blank>

Roboter empfinde sie als «schwierig» im Einsatz. Sie seien zwar wohl die Zukunft, Menschen haben jedoch Zuwendung und können andere Menschen berühren. Roboter hätten keine Emotionen.

Es sei wichtig, das Herz bei der Sache zu haben. Bewohnerinnen merken den Unterschied, ob jemand den Job mit Herz und Leidenschaft mache. Im Alter werde diese Leidenschaft auch

gesucht. Aufgrund von eingeschränkter Mobilität, wiederholendem Fernsehprogramm etc. werden zwischenmenschliche Begegnungen wichtiger. Die betreuenden Personen seien Menschen, welche die BewohnerInnen begleiten.

B. kann viel von ihrer Zeit mit den Bewohnern verbringen, da es ihre Aufgabe sei sich hauptsächlich mit den Menschen auszutauschen.

In der Pflege sei Dokumentation ein wichtigerer Bestandteil der Arbeit, der Anteil an höher und die Ressourcen knapper. In privaten Altersheimen kämen mehr Laptops und technische Geräte zum Einsatz, wodurch das Dokumentieren auch zeitnah und vor Ort erstellt werden könne.

Das Altersheim Klus dokumentiert von Hand auf Papier (auf Block). Sobald Zeit vorhanden sei, werde das Aufgeschriebene in den Computer übertragen. Dieser benötige sehr lange Zeit zum Aufstarten. Diese Tätigkeit könne vereinfacht werden, indem die Wartezeit während eines WC-Ganges einer Bewohnerin dazu genutzt werde, um einen Dokumentationseintrag zu erstellen. Voraussetzung sei jedoch, dass ein Laptop oder Tablet vorhanden sei und man es bei sich habe.

Wünschen würde sie sich Akzeptanz von ihrer Arbeit. Sie hat das Gefühl, dass der Eindruck entstehe, dass die eigene Tätigkeit nur darin bestehe mit anderen rumzusitzen und Kaffee zu trinken. Auch würde sie sich mehr Personal wünschen. Aktuell seien zwei Personen in der Aktivierung für 100 BewohnerInnen zuständig. Grösseres Team, mehr Zeit. Als Beispiel nennt sie die Ausstellung begleiten. Das sei sehr wichtig, brauche aber sehr viel Zeit.

(Politische) Lösung finden für die Bezahlung unserer Arbeit.

Weitere Aussagen

In der Aktivierung gehe es auch darum eine Stimmung in der Gruppe zu erzeugen und zu lenken. Man solle Menschen dazu motivieren, etwas zu tun und sich zu bewegen. Auch Bewegung sei Therapie. Bewegungsabläufe zu üben, sei ebenfalls Bestandteil des Trainings.

Ältere Menschen hätten manchmal Hemmungen mit technischen Geräten. Sie hätten Angst, es könne etwas kaputt, gehen.

B. empfiehlt immer neugierig zu bleiben.

Ältere Menschen sollen nicht als dumm verkauft werden. Sie sollen als das wahrgenommen werden, was sie sind, mit all ihren Facetten. Sie haben ein gelebtes Leben.

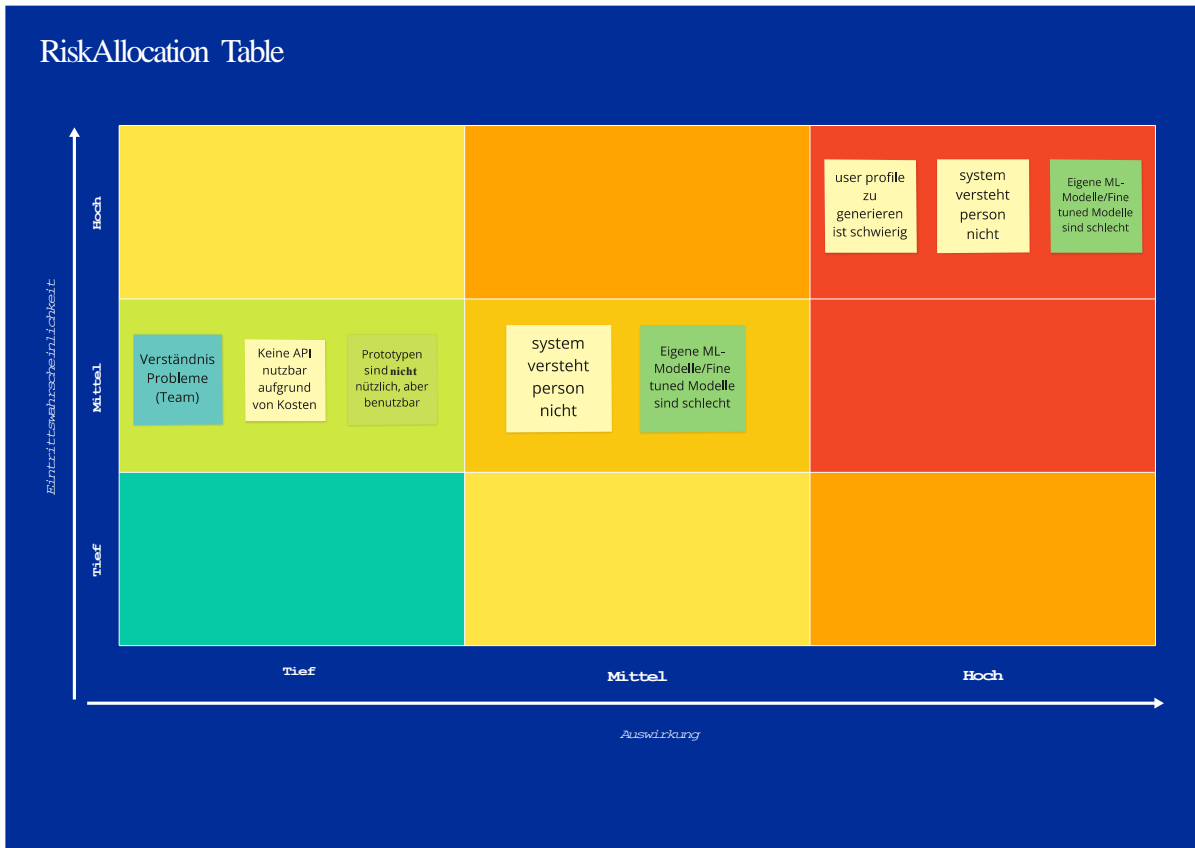
Anhang 22 Risikoanalyse III & IV vom 15.09.2023 & 20.10.2023

RiskAssesment DEV UX

Copy/Past-Its from the previous board. Put them in the corresponding column. This step is completed, when all rows are filled out.

	Ziel	Projektaktivität	Auslöser	Ursache	Gegenmaßnahmen	Auswirkung	Eintrittswahrscheinlichkeit	RiskScore	Verantwortlich	
Keinplan Produkt Starten	Neues Produkt	Usability tests	Kein Bedarf	Keine Akzeptanz	Konkreter Use nach rechts	User Research User Interviews User Journeys	Viele Tests	3 1	1 6	3 6
Produkt Probleme (Flow)	(Ende) Produkt	Umsetzung	Unvollständiges Produkt	Unklarheit über Anforderungen	Glossar	Benutzer Hilfen Hilfen Hilfen		1 6	3 1	3 6
Langsam Produkt entwickeln	Produkt entwicklung	Feature Testing	Unvollständiges Produkt	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		6 3	1 1	6 3
Keinplan Produkt Starten	Neues Produkt	Usability Test	Unvollständiges Produkt	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		3 3	3 3	9 9
Stark Thema verarbeiten entwickeln	Neues Produkt	Umsetzung	Technische Nutzung	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		1 1	3 3	3 3
Keinplan Produkt Starten	Neues Produkt	Usability tests	Priorisierung/Scope	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		6 6	3 3	18 18
System entwickeln permanente	Produkt entwicklung	Usability tests	Unvollständiges Produkt	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		3 6	3 3	9 18
Keinplan Produkt Starten	Produkt entwicklung	Feature Testing	Unvollständiges Produkt	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		6 6	1 1	6 6
Produkt entwickeln	Produkt entwicklung	Usability tests	Unvollständiges Produkt	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		3 6	3 3	9 18
Keinplan Produkt Starten	fertiges Produkt	Umsetzung	Unvollständiges Produkt	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		6 3	3 6	9 18
System entwickeln	System Produkt	Design	Unvollständiges Produkt	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		3 6	3 3	9 18
Scope Creep	fertiges Produkt	Umsetzung	Unvollständiges Produkt	Unklarheit über Anforderungen	Epics	Benutzer Hilfen Hilfen		3 3	1 3	3 9

Anhang 23 Risikoanalyse III & IV-Matrix



Anhang 24 Experteninterview L.T. (Zyklus 1)

2023.07.18 - L.T. (Pflegerin, Stv. Teamleiterin), 21

Alterspflegerin

2 Jahre Berufserfahrung

+ Arbeit mit Menschen - Gefühl etwas Gutes gemacht zu haben

- Wenig Personal, am Minimum, wenig Zeit für die Leute, für Gespräche, zuzuhören

Administrative Aufgaben

- Dokumentation für Krankenkassen Abrechnungen, Food Plan für Küche - alles sehr zeitaufwändig
- Admin macht 1,5-2h pro Tag aus, ohne Notfälle. Gibt als Stv. Teamleiterin auch viel Unterstützung, für diejenigen die nicht so gut darin sind.

Bewohner:

- Alter zwischen 60 und 100 Jahre
- Sehr verschiedene - noch selbständig sind bis pflegebedürftige
- Auch psychisch kranke

Typischer Alltag Pflegerin & Bewohner:

- Bewohner werden geweckt 7-9 Uhr
- Pflege - wenn sie es nicht selber können
- Morgenessen von 7:30 - 9
- Selbständige gehen Spazieren, Kreuzworträtsel lösen, einige gehen wieder schlafen
- Ausser haus

Mittagessen 11 – 12q

- Viele gehen wieder schlafen oder spazieren
- Jassen – sind oft «geschlossene» Gruppen
- Im Zimmer gar nichts machen
- Viele sitzen nur rum, schauen wer rein und raus geht.
- Nehmen am Aktivierungs-Programm teil

17:30 Nachtessen – Der Abend ist danach schnell um.

- Fernsehen, Telefonieren, oder schlafen gehen.

Was machen die Bewohner im Alters-/Pflegeheim?

Es gibt Aktivierungs-Programme:

- malen, singen, Gedächtnistraining, Kochen, Boccia spielen, Gymnastik, Musiktreffs mit Künstlern, Lesezirkel → VORLESEN
 - Lesezirkel:
 - Gruppen von etwa 15 Personen
 - Eine Aktivierungsfachfrau leitet und liest vor
 - Austausch von Gedanken
- Abendtreff - lose, um zu schwatzen.
- Sonntags oft Konzerte im Haus – es gehen aber wenige hin. Gruppenkonstellation passt nicht. Gewisse sind krankheitsbedingt verhindert

Die Aktivierung Teilnehmer erfolgt je nachdem wie fit sie sind.

Teilnehmende an solchen Programmen haben noch Spass am Leben.

Andere sind eher auf dem Zimmer und haben sich schon etwas aufgegeben. Sie Abend keine Lust auf irgendwas. Motivation hilft auch nicht. Aber wenn sie dann doch mal gehen gefällt, gehen dann aber doch nicht mehr.

Werden Gespräche geführt? Was sind die Themen?

- Die eigne Wohnung muss aufgegeben werden
- Altersheim eintritt
- Schwierig, weil es eine Einschränkung ist
- Einsamkeit ist oft Grund für Altersheim. Auch weil z.B. der Partner weg ist.
- Sterben ist ein Thema

Gewisse Bewohner reden sehr gerne: Ihre Lebensgeschichte, Veränderungen, etc.

Leider fehlt den Pflegern dafür meistens die Zeit → siehe Administrative Aufgaben! Dazu kommen Notfälle, oder der Aufwand, sich um jemanden kümmern müssen dem es einfach nicht gut geht.

Probleme

- Kontakt finden im Heim ist schwierig - Einsamkeit
- Die Leute gehen oft nicht aufeinander zu
- Es bilden sich Gruppen - und etablieren sich, sind geschlossen, wollen keine neuen «Mitglieder» aufnehmen.
- Psychische und physische Erkrankungen der Bewohner

Es gibt immer auch Leute, die sich finden und etwas zusammen machen, spazieren.

Welche Technik/Medien Nutzen alte Leute?

- Fernseher haben viele
- einige haben Laptop oder Tablet.

- Viele haben iPhone, aber maximal zum Anrufen oder für SMS, und brauchen auch dazu Hilfe.
 - o Vor allem neu eintretende haben ein Phone.
- Wenn WLAN weg ist, geht nix mehr ohne Hilfe.

Check-out

- Wünscht sich mehr Wertschätzung für ihre Arbeit
- auch gerade von den Angehörigen. Sie haben sehr hohe Ansprüche und wenig Verständnis.
- Mehr Personal, um mehr Zeit für die Patienten zu haben.

Anhang 25 Experteninterview S.H. (Zyklus 1)

Datum: 2023.07.28

Ehemals Sozialarbeiterin, Arbeitet in der in der Alterspsychiatrie , Schwerpunkt Altersbereich ab 65 Jahren

Aufgaben:

- Ist eigentlich im Spital.
- Die Station hat 20 Betten.
- 50% der Arbeit ist Begleitung der Patienten, etwa 15min pro Patienten und Gespräch.
 - o Diese Zeit ist sehr knapp für gewisse Themen. Die Person besser kennenlernen wäre manchmal schön.
 - o Die langjährige psychische Geschichte resp. der Ursprung der Probleme der Person kommen oft zu kurz.
- Der Rest der Zeit ist dasselbe, aber in der ambulanten Behandlung

Beratungsthemen: Wohnen, Finanzen, Rechtliches.

- Austritt planen
- Spitex
- Tagesstruktur
- Finanzen und Rechtliches

Akutklinik: Behandlung von:

- Depressionen, Abhängigkeits-Erkrankungen (Alkohol, Medikamente)
- Schizophrenie
- Demenz → dafür gibt es auch spezialisierte Stationen
- Kognitive Einschränkungen - leichte (Gedächtnis lässt nach) bis Demenz. Alltagsbeeinträchtigungen, die Abhängig von externer Hilfe macht.

Problem: fehlende Tagesstruktur - nach der Pensionierung bricht diese weg.

- Viele wissen nicht was machen mit der Zeit
- Soziales Umfeld ist nicht mehr so gross
 - oft ist auch die Familie weg
- Soziale Kontakte fehlen
- Die Leute beginnen sich zu distanzieren oder ziehen sich zurück
- Ursprung der psych. Probleme ist oft in der Familie – Konflikt mit Kindern oder anderen Angehörigen.

➔ Das Wichtigste neben Gesundheit: Tagesstruktur schaffen und Einsamkeit entgegenwirken.

Art der Programme:

- Stationär: Behandlung meistens etwa zwei Monate.
- Tagesklinik: Tagesprogramm, die Leute gehen abends wieder nach Hause
 - o Behandlung in der Tagesklinik auch länger als stationär
 - o Tagesklinik ist in der Gruppe – es gibt aber auch Einzelgespräche.

Aktivierungstherapeuten für:

- Ergotherapie, Alltagsbewältigung
- Musik & Kunst Therapie: malen, singen
- Generell Aktivität haben den Tag durch
- Kontakte knüpfen
- Unterstützung für nach dem Austritt - Programme weiterführen.

Pro Senectute:

- bietet Programme, aber nur z.B. einmal wöchentlich für eine Stunde.
- Schwierig für Alte, sich so eine Struktur, ein Wochenprogramm zu gestalten.

Einschränkungen, welche die Inanspruchnahme der Angebote erschweren:

- Gehen, allgemeine Mobilitätseinschränkungen
 - Erreichbarkeit der Angebote - wie kommen die Leute dahin?
- Finanzielle Einschränkungen im Alter
- Auch andere Termine (Arzt, Spital, Behandlungen) bestimmen das Programm der Leute mit

Warum werden die Programme nicht genutzt?

- Viele gehen nicht dorthin, weil sie keine Angebote für alte Leute wollen.
- Sehr grosse Unterschiede im Zustand je nach Alter.
 - o Die Leute sind auch sehr unterschiedlich fit.
- Es gibt wenig Angebote, die durchmischt sind vom Alter - Alte wollen aber nicht nur mit anderen alten Leuten zusammen sein.
- Vielen ist es langweilig oder sie ziehen sich zurück - oft ist nicht klar, warum sie nicht am Angebot teilnehmen.
- Hemmschwelle solche Angebote zu nutzen ist gross.
- Migrationshintergrund: diese Leute haben teilweise das Wissen nicht, wie sie zu den Angeboten kommen.

Bedürfnisse von alten Leuten:

- Alte schätzen, wenn etwas leben in ihrem Umfeld ist.
- Soziale Kontakte - Einsamkeit im Alter ein grosses Thema.

- Alte Leute wollen nicht nur mit anderen alten untereinander sein.
- Erzählen gerne; zuhören ist wichtig - es muss nicht viel Rauskommen dabei, sie wollen keine Ratschläge.
 - o Gesprächsthemen: Familie, Geschichten vom Leben, von früher

Probleme:

- Wohnen alleine, unternehmen nicht viel – führt zu Einsamkeit, Langeweile.
 - o haben trotzdem Angst davor, ihr zuhause zu verlieren und ins Heim zu kommen.
- Kochen nicht für sich alleine

Lösungsansätze:

- Sinnstiftende Beschäftigung hilft "gesund" zu altern.
 - o eine Aufgabe haben, Vereine, Enkel, gemeinnützige Arbeit.
- Finanzieller Aspekt ist wichtig - viele Angebote sind kostenpflichtig. Das Budget ist im Alter aber kleiner.
- Gemeinsames Kochen & essen, wird z.B. von Migros organisiert.
- Angebote einfach, niederschwellig zugänglich machen
- Generationenübergreifende Angebote

Technik & Medien:

- Es gibt immer wieder Leute die iPhone gut bedienen
 - Kommt jetzt langsam, eher bei denen die jetzt pensioniert werden.
 - Bis ca. 70 gibt's solche die es noch nutzen.
- Es kommt darauf an, ob sie gegenüber Technik offen sind oder gar nichts damit zu tun haben wollen.

Case Vorlesen:

- Es gibt Bücherclubs – gemeinsames lesen und sich austauschen
- Wenn vorlesen, dann am ehesten durch Freiwillige, in Alterszentren - kennt sie aber nicht so.

Check-out:

- Es gibt nicht zu wenige Angebote - aber sie sind nicht gut zugänglich.
- Angebote sind verteilt und unstrukturiert – sie müsste gebündelt werden
- Beratungsstellen zugänglich machen. Eher auf die Leute zugehen, um die Hemmschwelle zu reduzieren.
 - o Wie bei Jugendarbeit – dort macht man das ja auch so in dem man aktiv auf sie zugeht.
- Aktiver informieren und motivieren über die Möglichkeiten

- Aber nicht zu aktiv - Besuche zuhause verängstigen die Leute, z.B. sie könnten ihr zuhause verlieren, wenn gerade nicht geputzt und aufgeräumt ist.

Anhang 26 Gedächtnisprotokoll Aktivierungstherapie 1

Ort: Altersheim Klus

Datum: 16.8.2023

Uhrzeit: 09:15 bis 12:00 Uhr

Intro:

- Vorstellung von Barbara B., der Gruppenleiterin für Aktivierungstherapien.
- Führung durch das Heim.
- Erläuterung des historischen Hintergrunds: Früheres Spital. In einem Trakt sind die Zimmer größer wegen der umgebauten Operationssäle; die anderen Zimmer waren frühere Patientenzimmer.
- Jeder Bewohner hat ein Einzelzimmer.

Das Gedächtnistraining:

- Ort: Balkon, genannt "Piazza", nahe dem Atelier, welches Mal- und Bastelzubehör sowie zwei ältere Desktop-Computer beherbergt.
- Teilnehmer: 12 Personen (2 Männer, 10 Frauen).
- Die Aktivierungstherapeutin spricht langsam, laut und deutlich.
- Sie leitet die Gespräche und bringt Themen wie Krieg, Sport, Kultur, Religion und Natur auf.
- Häufige Diskussionen über historische Ereignisse.

Warm-Up erster Teil: Verteilung von Landkarten.

- Die Teilnehmer*innen müssen Ländernamen manuell eintragen.
- Länder auf der Karte sind nummeriert, was der Auflistung der Ländernamen nahekommt.

Warm-Up zweiter Teil: Beschriftung der Flaggen von ehemaligen jugoslawischen Staaten.

- Assoziation mit Sport und Sportlern.
- Intensives Arbeiten mit Bildern und den zugehörigen Geschichten/Hintergründen.
- Verschiedene Teilnahmeaktivitäten: Einige sind aktiver, andere weniger.
- Diskussion über Assoziationen mit diesen Ländern, einschließlich Themen wie Tito, Krieg, Meer und Segeln.
- Erwähnung des Ersten Weltkriegs, des Islams und der Tatsache, dass Winnetou-Filme in Kroatien gedreht wurden.
- Die Gruppenleiterin teilt ihre Reiseerfahrungen und die Gruppe diskutiert das Wissen über diese Länder.

- Fotos werden gezeigt, darunter Bilder von Sarajevo, einem Urwald und der Mostar-Brücke.
- In der Abschlussdiskussion geht es um die Gründe für Spannungen in diesen Ländern und den friedlichen Umgang mit ethnischen und religiösen Unterschieden in der Schweiz.

Outro Gespräch:

- B., ursprünglich gelernte Krankenschwester und seit 20 Jahren im Beruf, stellt sich vor.
- Sie betont die Bedeutung der sozialen Kontakte in den Therapien.
- Handys sind für Angestellte verboten, um die Privatsphäre der Bewohner*innen zu schützen.
- Dokumentation der Aktivitäten nimmt täglich etwa 2 Stunden in Anspruch.
- Probleme mit alter und langsamer Technik werden erwähnt.
- Drei Studentinnen wohnen im Altersheim in einer WG und unterstützen die Bewohnerinnen.
- Viele neu hinzugezogene Bewohner*innen bringen eigene Laptops und Smartphones mit und benötigen hauptsächlich WLAN-Zugang.

Anhang 27 Gedächtnisprotokoll Quiznachmittag

Ort: Altersheim Klus

Datum: 17.8.2023

Uhrzeit: 15:30 bis 18:00 Uhr

6 Personen + 1 Aktivierungstherapeutin:

Es wurden drei Quiz-Kategorien zur Auswahl gegeben: Allgemeinwissen, Historisches, Bekannte Persönlichkeiten. Gewählt wurde das «Bekannte Persönlichkeiten»-Quiz. Das Format ist Multiple-Choice.

Beispiel einer Frage: Wieso hat Facebook die Farbe Blau:

Antwort Optionen

- Mark Zuckerberg ist farbenblind
- Mark Zuckerberg liebt das Meer
- Die Farbe soll Amerika repräsentieren

Lösung: Mark Zuckerberg ist Rot-Grün-Farbenblind. Er fand dies während seiner Zeit in Harvard heraus und bemerkte, dass er sich nicht viel um Farben gekümmert hat.

Das Gespräch lief weiter und jemand sagte: «Zu Facebook gehört auch Instagram». Ich ergänzte «...und WhatsApp».

Eine Bewohnerin fragte mich, weshalb WhatsApp nicht beliebt sei. Ich antwortete: «weil Facebook viele Informationen über Personen sammelt und diese Informationen an Werbetreibende verkauft. Selbst über WhatsApp versendete Nachrichten werden von Facebook gesammelt». Es entstand ein Gespräch darüber, was Instagram ist und weshalb es das gibt. Eine Bewohnerin sagte, man könne doch auch etwas anderes als WhatsApp benutzen. Sie versuchte sich an den Namen zu erinnern. Ich fragte, ob sie «Signal» meine. Sie bejahte dies.

Weitere Fragen zu Persönlichkeiten waren Frida Kahlo, Angela Merkel, Immanuel Kant.

Eine Person erzählte, sie sei bei einer Quizshow im Jahre 1984 dabei gewesen. Sie erinnerte sehr detailliert an die Frage und erzählte, warum sie diese trotzdem falsch beantwortet wurde (war nicht ihre Schuld).

Bei Frida Kahlo wurde im Gesprächsverlauf auf ihre markante Augenbraue hingewiesen, warum sie anfang zu malen, ihre Schwierigkeiten während der mexikanischen Revolution, ihren weniger erfolgreichen Ehemann und warum sie ins Ausland reisen konnte, während er es nicht konnte. Die Aktivierungstherapeutin zeigte Bilder von Frida Kahlo auf ihrem Handy. Es wurde auch über Kahlos besondere Maltechnik gesprochen. Eine Person meinte, dass sie Blumen malte, die wie eine Vagina aussähen.

Immanuel Kant war bekannt für seine Pünktlichkeit. Der kategorische Imperativ wurde besprochen. Eine Person wusste, was das ist. Es wurde erläutert und zusammengefasst: "Was du nicht willst, dass man dir tu', das füg' auch keinem anderen zu." Immanuel Kant war bekannt für seine Pünktlichkeit.

Gegen Ende der Session wurde ich gefragt, ob ich mich mit künstlicher Intelligenz auskenne und was ich davon halte. Daraufhin führten wir ein spontanes Experiment mit dem Bing-Chatbot durch. Ich bat die Anwesenden, eine Frage zu stellen. Die erste Frage lautete: «Wie lange lebe ich noch?», worauf ich zurückfragte, ob sie diese Frage beschäftige. Sie verneinte. Daraufhin wollte sie wissen, wie die Welt in der Zukunft aussieht. Bing gab eine ausführliche Antwort.

Eine weitere Person wollte wissen, welche Vorteile autonom fahrende Fahrzeuge bringen würden. Bing hielt daraufhin einen kurzen "Vortrag" über autonome Fahrzeuge und zählte die Vor- und Nachteile auf. Sie fragte, ob ein autonomes Fahrzeug auch für einen Igel bremsen würde. Auch diese Frage wurde von Bing beantwortet.

Der Dialekt des Bingbots und die synthetische Sprache, die oft monoton und flach klingt, wurden besprochen.

Um den Bewohnern eine andere künstliche Stimme zu demonstrieren, wurde die Spiegel Online-Website geöffnet. Ich las einige Schlagzeilen vor und entschied mich für einen Artikel über Energiegewinnung mittels einer Energieinsel in der Ostsee. Der Computer las vor. Die Anwesenden meinten, dass die Stimme sehr schnell spräche, was das Verstehen erschwerte. Während des Vorlesens schüttelte eine Person ungläubig den Kopf, während die Person neben mir gespannt zuhörte. Als der Artikel beendet war, fragte ich die Anwesenden, was sie aus dem Artikel mitgenommen hätten. Eine Person meinte, es seien viele Zahlen vorgekommen. Eine andere sagte, es ginge um Energie und eine Insel.

Nach dem Ende der Quizrunde blieben einige Personen sitzen. Es fanden sehr persönliche Gespräche statt.

Ich stellte auch einige persönliche Fragen und wollte wissen, welche Musik früher gehört wurde. Eine Person nannte Elvis Presley, eine andere sagte, sie sei leider zu arm gewesen, um Musik zu hören und habe erst später damit begonnen. Es wurde besprochen, welche Instrumente die Anwesenden gelernt hatten. Genannt wurden Geige, Blockflöte (schweizerdeutsches Wort: «Spoizchnebel»), Klavier und Gitarre.

Ich fragte, welche berühmte Person sie jemals getroffen hätten. Eine Person meinte, sie hätte neben John Travolta im Zug gesessen. Eine andere sagte, sie habe das Honorar von Mahalia Jackson in der Hand gehabt und erzählte die Geschichte dazu.

Fazit:

Die Quiz-Nachmittage im Altersheim Klus dienen dem sozialen Austausch und schaffen Grundlagen für Gespräche. Es ist für ältere Menschen besser, sich auszutauschen, als isoliert im Zimmer zu bleiben.

Anhang 28 User Study Testszenario «Gedächtnistraining» (Zyklus 1)

Setup der User Studie:

- 1 ModeratorIn (System Persona), 1 TeilnehmerIn (Nutzer)

Eingesetzte Technik:

- 2 Mikrophone
- 1 Aufzeichnungsgerät (Diktiergerät oder Smartphone mit App)
- 1 Notebook/Tablet mit Internetverbindung
- Session Recording auf dem Gerät (z.B. Teams)
 - Optional: Video Recording

Übergeordnete Ziele der Studie:

- Erhebung und Dokumentation von Dialogen

Checklist:

1. Einverständniserklärung & Informationen zum Probanden
2. Briefing
3. Läuft die Aufzeichnung
4. Check-in
5. Aufgaben
6. Cooling off

Briefing

Wir studieren Mensch-Computer-Interaktions-Design und wollen einen sprachbasierten Gedächtnistraining Prototypen entwickeln. Dazu benötigen wir eure Unterstützung und würden gerne ein spielerisches Experiment mit euch durchführen.

Eine Person übernimmt die Aufgabe des Moderators, also des Computers und eine Person übernimmt die Aufgabe des Zuhörers.

- Dabei möchten wir herausfinden, wie ihr das Spiel findet.
- Es ist ein Experiment – Experimente können auch scheitern, das ist aber kein Scheitern der Testpersonen.
- Bist du einverstanden mit der Aufzeichnung?
- Es kann nichts falsch gemacht werden – es gibt kein richtig oder falsches vorgehen.
- Das Experiment kann jederzeit abgebrochen werden.

Rolle Moderator (System Persona):

Der Moderator führt in das Spiel ein und erklärt das Spiel, die Regeln und die Modalitäten. Dann beginnt das Spiel.

Rolle Teilnehmer (Nutzer):

Der Teilnehmer hat die Aufgabe am Spiel teilzunehmen

Einverständniserklärung Teilnehmer

Name des Teilnehmers: _____

1. **Zweck der Studie:** Ich verstehe, dass ich gebeten wurde, an einer Forschungsstudie teilzunehmen, die darauf abzielt, die Benutzererfahrung und Funktionalität von Voice Interfaces (wie Siri, Alexa und Google Assistent) zu bewerten.
2. **Verfahren:** Ich bin darüber informiert worden, dass die Studie die Durchführung einer Reihe von Aufgaben unter Verwendung dieser Voice Interfaces umfasst und meine Interaktionen aufgezeichnet werden können. **Diese Aufnahmen werden ausschließlich zur Analyse verwendet.** Nach Beendigung der Studie werden die Aufnahmen gelöscht.
3. **Freiwillige Teilnahme und Abbruch:** Ich verstehe, dass meine Teilnahme an dieser Studie freiwillig ist und dass ich jederzeit ohne negative Folgen aussteigen kann.
4. **Vertraulichkeit:** Ich wurde darüber informiert, dass alle Informationen, die während dieser Studie gesammelt werden, vertraulich behandelt werden und dass alle persönlichen Identifikatoren vor der Analyse entfernt werden.

Indem ich unten unterschreibe, erkläre ich, dass ich die oben genannten Bedingungen verstanden habe und freiwillig an dieser Studie teilnehme.

Unterschrift des Teilnehmers: _____

Datum: _____

Information zum Probanden

Name	
Alter	
Geschlecht	
Wohnort	
Beruf	

Check-in (Nutzer)

Fragen	Antworten
Benutzt du Computer, Tablets, Smartphones oder eine Smartwatch?	
Hast du schon mal mit einem Computer gesprochen?	
Wie hältst du dich körperlich und geistig fit?	
Machst du gerne Rätsel, Kreuzworträtsel, Sudoku oder ähnliches?	
Spielst du gerne Gesellschaftsspiele?	
Hast du Einschränkungen in der Sensorik? Z. B. Brille, Hörgerät?	

Cooling off

Wie war das für dich? Beschreibe in mit 3-4 Schlüsselworten deine Erfahrung

Könntest du dir vorstellen, dieses Spiel mit einem Computer zu spielen?

Hast du Lust wieder einmal an einem Experiment teilzunehmen?

Anhang 29 Dialog Quiz (Zyklus 1)

MAX: Hallo. Ich bin Max Memory, dein Gedächtnistrainer. Wollen wir gleich mal loslegen?

User:

MAX: Sehr gut. Lass uns ein Zahlenspiel spielen.

User:

MAX: Perfekt. Das Spiel geht so: Du hörst fünf Zahlen. Diese musst du in der richtigen Reihenfolge wiedergaben. Wir fangen mal ganz einfach an. Bist du bereit?

User:

MAX: 8, 2, 3, 9, 5

User:

MAX: Das klappt ja wunderbar! Nächstes Level. Jetzt hörst du fünf Zahlen und musst sie in umgekehrter Reihenfolge wiedergeben. Bereit?

User: Ja.

MAX: 9,1,7,6,3

User:

MAX: Na du bist ja ein richtiges Memory-Monster.

MAX: Wollen wir die Schwierigkeit noch erhöhen?

User:

MAX: Gib die Zahlen in der gleichen Reihenfolge an, wie ich sie dir sage: 72, 23, 45, 87, 35.

User:

MAX: Wow. Schon wieder alles richtig!

MAX: Wir kommen jetzt zu Level 4. Jetzt hörst du fünf Zahlen und musst sie in umgekehrter Reihenfolge wiedergeben. Bereit?

User: Los gehts.

MAX: 96, 31, 74, 39, 12.

User:

MAX: oh nein. Da war leider was falsch. Die richtige Reihenfolge war: 96, 31, 74, 39, 12.

MAX: Du bist auf gutem Weg. Lass uns noch ein Versuch wagen.

MAX: Gib die Zahlen in umgekehrter Reihenfolge an: 56, 71, 68, 62, 43.

User: 65, 71, 86, 62, 43.

MAX: Nicht ganz. Noch ein Versuch?

User: Danke. Das reicht für heute.

MAX: Was für eine Leistung. du hast 3 von 5 Levels richtig beantwortet. Aber vergiss nicht: Übung macht den Meister! Ich freu mich mit dir morgen wieder zu üben.

MAX: Wenn du möchtest können wir noch ein anderes Spiel spielen.

User: OK. Was gibts noch?

MAX: Wie wärs mit Hauptstadt raten oder berühmte Persönlichkeiten?

User: Berühmte Persönlichkeiten:

MAX: Du hast berühmte Persönlichkeiten gewählt. Das ist ein Multiple Choice Quiz. Du bekommst eine Frage gestellt und hast vier Antworten zur Verfügung.

MAX: Welcher deutsche Bundeskanzler sagte: Wer Visionen hat, soll zum Arzt gehen

A: Konrad Adenauer

B: Willy Brandt

C: Helmut Schmidt

D: Gerhard Schröder

User: Helmut Schmidt

MAX: Das ist goldrichtig!

MAX: Erklärung: Helmut Schmidt wurde in einem Interview gefragt, wie gross seine Vision sei, worauf er die berühmte Aussage machte: Wer eine Vision hat, der soll zum Arzt gehen.

Schmidt will das "damals aber nicht mit Blick auf Willy Brandt formuliert haben, wie so oft angenommen. In einem Interview mit dem Zeitmagazin gibt er an, dass es eine pampige Antwort auf eine dusslige Frage war.

MAX: Gefällt dir das Spiel?

User: ja

MAX: Möchtest du noch eine Frage:

User: Ja:

MAX: Wer ist der am längsten regierende Monarch der Welt?

A: Harald V. aus Norgwegen

B: Hans-Adam II. aus Lichtenstein

C: Haji Hassan al-Bolkiah aus Brunei

D: Margerethe II aus Dänemark

User: Keine Ahnung.

MAX: Du kannst auch einfach raten.

User: Margarethe II aus Dänemark.

MAX: Das ist leider falsch. Der am längsten regierende Monarch ist Haji Hassan al-Bolkiah aus Brunei. Er ist seit dem 4. Oktober 1967 im Amt und ist der 29. Sultan von Brunei.

Die am längsten regierende Monarchin ist übrigens Margarethe II aus Dänemark. Sie ist seit dem 14. Januar 1972 Königin von Dänemark.

MAX: Noch eine Frage?

User: Ja, warum nicht.

MAX: Mit wie vielen Frauen war Tom Cruise verheiratet?

A: 2

B: 3

C: 4

D: 1

User: C.

MAX: Das war leider falsch. Tom Cruise erste Frau war Mimi Rogers. Seine zweite Frau war Nicole Kidman mit der er 14 Jahre verheiratet war. Seine dritte Ehefrau war Katie Holmes mit der er 6 Jahre verheiratet war. Tom Cruise wurde 4 mal für den Oskar nominiert, hat aber nie gewonnen und wurde 7 mal für den Golden Globe Award nominiert und hat 3 mal gewonnen.

MAX: Du bist ein richtiger schlauer Fuchs. Für heute hast du alle Fragen beantwortet. Schau morgen wieder vorbei.

Anhang 30 User Sample (Zyklus 1)

User Sample					
Name	Helene G.	Ottokar B.	Rosemarie S.	Konsolidiert	
Alter	65	83	76	74.33	
Geschlecht	weiblich	männlich	weiblich	weiblich	
Wohnort	Zürich	Zürich	Mönchaltorf	Zürich	
Beruf	Dip. Biologin im Ruhestand	Textileinkäufer im Ruhestand	Bankkauffrau, Autogenes Training, Hausfrau im Ruhestand	Rentner	
Benutzt du Computer, Tablets, Smartphones oder eine Smartwatch?	Tablet Smartphone Computer	1-2 mal pro Woche Computer Nutzt es um Wissen zu recherchieren Kochrezepte E-Mail	Computer Smartphone	Computer Smartphone	
Hast du schon mal mit einem Computer gesprochen?	Siri, Testweise	Siri, Testweise	Nein	Siri, Testweise	
Machst du gerne Rätsel, Kreuzwörter, Sudoku oder ähnliches?	Nein	Nein	Jassen, am Computer gegen Computer	Nein	
Machst du aktives Gedächtnistraining, Gehirnjogging oder ähnliches?	Nein (lernt aber Pflanzennamen autodidaktisch)	Nein	Früher gemacht, Heute Nicht mehr	Nein	Beschäftigt sich aber mit eigenen Interessen
Hast du Einschränkungen in der Sensorik? Z. B. Brille, Hörgerät?	Brille, Hört sehr gut	Nein (Linsen)	Brille, Hörgerät wird angeschafft	Brille	
Spielst du gerne Gesellschaftsspiele, Computerspiele oder ähnliches. Welche spiele spielst du? Wie oft?	Gesellschaftsspiele	Nein	Ja	Ja	

Anhang 31 Auswertung (Zyklus 1)

	Helene G.	Ottokar B.	Rosemarie S.
Beobachtungen	<p>Hat sehr gut mitgemacht</p> <p>Hat die Zahlenreihe sehr schnell gelernt</p> <p>War einmal kurz abgelenkt</p> <p>Erläuterung via Smartphone (Grösse von Tom Cruise)</p> <p>hätte gerne visuelle Informationen</p>	<p>Mochte das Zahlenspiel nicht</p> <p>Hat Zahlenspiel abgebrochen</p>	<p>hätte gerne einen bildschirm</p>
Cooling Off Aussagen	<p>Würde sich gegenüber einem System anders Verhalten (Weniger schämen, wenn etwas nicht klappt)</p> <p>Normale Dialoge</p> <p>Würde es nicht nutzen</p> <p>Mehr Kontext</p> <p>Mehr Bezug zum Menschen</p> <p>Fragen zu Themen die auch wirklich interessieren</p>	<p>War Lustig</p> <p>würde es nicht nutzen aufgrund fehlender Zeit</p>	<p>Zahlenspiel war OK</p> <p>Die andren fragen sind sehr interessen gebunden</p> <p>Kein interesse an den Frager</p> <p>findet computer bezogenheit nichtgut</p>
Sonstiges	<p>Geschichte erzählen</p> <p>Mehr Lebendigkeit</p> <p>Mehr Eindrücke</p> <p>Viel Beschäftigt</p>	<p>Hat sonst interessante geschichten zu erzählen</p> <p>Interssiertsich für Länder und Esotherik</p> <p>Ist viel Beschäftigt</p>	<p>multiple choice fragen ist kein geedächtnstrainign</p> <p>mehr fantasie anregend</p> <p>begriffe nicht sagen</p> <p>vorstellung skraft</p> <p>gedächtnistechniken gehören dazu</p> <p>anleitung für user fehlt</p> <p>auch mit interaktion mit anderen</p> <p>4 begriffe auf karte</p> <p>visualisierung</p> <p>Raum-Geschichten</p> <p>geschichten</p> <p>logi-technick</p>

Anhang 32 User Study Szenario «Kennenlernen» (Zyklus 2)

Setup der User Studie

- Testleitung vor Ort
- 1 Test-Probanden (Nutzer*in)
- 1 Moderation (System Persona) via Online/Telefon

Timing: **30min**/Test inkl. Briefing, Check-in, Cooling off

Eingesetzte Technik:

Vor Ort:

- 1 Smartphone & Notebook mit Teams (für Aufzeichnung)
- Optional: Video Recording

Remote (Moderation, WoOz):

- 1 Notebook/Tablet
- Session Recording, Transkription auf dem Gerät (Teams)

Übergeordnete Ziele:

- Erhebung und Dokumentation von *Dialogen*

Checklist:

7. Einverständniserklärung & Informationen zum Probanden
8. Briefing
9. Lläuft die Aufzeichnung
10. Check-in
11. Aufgaben
12. Cooling off

Briefing

Wir studieren Mensch-Computer-Interaktions-Design und wollen einen sprachbasierten Gedächtnistraining Prototypen entwickeln. Dazu benötigen wir eure Unterstützung und würden gerne ein spielerisches Experiment mit euch durchführen.

- Wir möchten herausfinden, wie wir **relevante Interessen** einer nutzenden Person herausfinden können.
- Es ist ein Experiment – Experimente können auch scheitern, das ist aber kein Scheitern der Testpersonen.
- Bist du einverstanden mit der Aufzeichnung?
- Es kann nichts falsch gemacht werden – es gibt kein richtig oder falsches vorgehen.

- Das Experiment kann jederzeit abgebrochen werden.

Testszenario

Rolle Moderator (System Persona)

Um das Szenario «realitätsnäher» zu gestalten, wird die Moderation nicht physisch an den Tests Teilhaben. Sie spielt den klassischen Wizard of Oz und simuliert die System Persona via Remote (Teams Call). Damit soll dem Umstand entgegengewirkt werden, dass die nutzende Person sich zu sehr durch die Anwesenheit der Person ablenken lässt. Der Moderator soll ebenso nicht zu sehr auf zwischenmenschliche Kommunikation reagieren können. Das Ziel ist, sich rein auf Dialoge fokussieren zu können.

Der Moderator (Computer) möchte herausfinden, welche Interessen die Testperson hat. Er führt durch das Gespräch resp. fragt nach, solange (max. 3-5 Fragen) bis genügend Informationen für eine Themen-Entscheidung vorhanden sind.

Ziel: Gesprächsthemen anhand der Nutzerinteressen führen zu können. Versuche dem jeweils anderen etwas zu erzählen, was ihn interessiert. Du darfst jedes Hilfsmittel benutzen, welches du benötigst.

Rolle Teilnehmer (Nutzer)

Der Teilnehmer hat die Aufgabe, dem Moderator seine in diesem Kontext relevanten Interessen/Antworten mitzuteilen.

Einverständniserklärung Teilnehmer

Name des Teilnehmers: _____

1. **Zweck der Studie:** Ich verstehe, dass ich gebeten wurde, an einer Forschungsstudie teilzunehmen, die darauf abzielt, die Benutzererfahrung und Funktionalität von Voice Interfaces (wie Siri, Alexa und Google Assistent) zu bewerten.
2. **Verfahren:** Ich bin darüber informiert worden, dass die Studie die Durchführung einer Reihe von Aufgaben unter Verwendung dieser Voice Interfaces umfasst und meine Interaktionen aufgezeichnet werden können. **Diese Aufnahmen werden ausschließlich zur Analyse verwendet.** Nach Beendigung der Studie werden die Aufnahmen gelöscht.
3. **Freiwillige Teilnahme und Abbruch:** Ich verstehe, dass meine Teilnahme an dieser Studie freiwillig ist und dass ich jederzeit ohne negative Folgen aussteigen kann.
4. **Vertraulichkeit:** Ich wurde darüber informiert, dass alle Informationen, die während dieser Studie gesammelt werden, vertraulich behandelt werden und dass alle persönlichen Identifikatoren vor der Analyse entfernt werden.

Indem ich unten unterschreibe, erkläre ich, dass ich die oben genannten Bedingungen verstanden habe und freiwillig an dieser Studie teilnehme.

Unterschrift des Teilnehmers: _____

Datum: _____

Information zum Probanden

Name	
Alter	
Geschlecht	
Wohnort	
Beruf	

Check-in

Hast du schon mal mit einem Computer gesprochen?

Hast du schon Erfahrungen mit KI-Assistenten/-Bots gemacht? Welche – erläutere bitte kurz.

Aufgabe, Test

Durchspielen gemäss Szenario. Der sich abspielende Dialog lässt sich nur schwer vorhersagen und kann sich je nach Fragen & Antwort in unterschiedliche Richtungen entwickeln.

Mögliche (ersten) Fragen:

- Was hast du heute gemacht?
- Was hast du heute vor?
- Was ist aktuell deine liebste Beschäftigung?
- Welches waren deine besten Ferien, und warum?

Mögliche Folgefragen:

- Was hast du noch für Pläne (heute)?
- Was hast du demnächst vor/geplant?
- Was machst du sonst noch gerne?
- Welche aktuellen Themen interessieren dich?
- Was war das Letzte, was du erfahren hast was dich besonders erfreut/gar nicht gefallen hat?

Cooling off Nutzer

Wie war das für dich? Beschreibe in mit 3-4 Schlüsselworten deine Erfahrung

Wie fühlten sich die Dialoge an?

War die erhaltene Information Interessant für dich? Wurden deine Interessen getroffen?

Die Vorstellung mit einem Gerät zu sprechen, finde ich sonderbar

1 – Stimme gar nicht zu	2- Stimme nicht zu	3 – Teils teils	4 - Stimme zu	5-Stimme voll zu

Ich fand die Antwort interessant

1 – Stimme gar nicht zu	2- Stimme nicht zu	3 – Teils teils	4 - Stimme zu	5-Stimme voll zu

Anhang 33 Selbsttest Kennenlernen (Zyklus 2)

You have up to 5 questions to gather some information about my interests, hobbies etc. from me. After that, tell me something that might be interesting for me

Sounds like a plan 🙌. I'll ask five questions, and then give you a recommendation. Ready to get started?

Sure

Great! I'll keep the questions simple to start. Let's start with hobbies. What kinds of hobbies do you have?

winter sports, bike and just hanging around

Nice! Winter sports, biking, and relaxation. All solid options. 😊
Next question. Do you prefer books, tv shows, movies, or podcasts when it comes to entertainment?

books and movies

Two excellent options. I'm glad we're narrowing it down. 🧐 Now, when it comes to books, what genres do you like? Are you more into nonfiction, like biographies or history books, or do you prefer fiction like fantasy, sci-fi, or thrillers?

I like books about psychology and human behavior. no novels, scifi and such stuff

Wow, you have very specific tastes. Books about psychology and human behavior are very thought-provoking. It sounds like you enjoy reading about the inner workings of the human mind. Would you say you're more interested in practical applications of psychology or more philosophical, theoretical works?

more the theoretical works

You must be quite the deep thinker 😊. Theoretical works definitely make you think. The next question I have for you relates to movies. Do you prefer films that are more grounded in reality, like dramas or thrillers, or are you more into the fantastical, like sci-fi or fantasy?

i like films that are about drugs and the stories behind that

Ah, so films about the dark and seedy underbelly of society. That's quite an interesting preference, but I get it. There's something intriguing about the stories behind the drug trade and those who are caught up in it. Are you looking for something more about the trade,

like "The Wire" or something like "Requiem for a Dream" that shows the human side of addiction?

I like both

That's great! That gives me a lot to work with. Based on your interests, I think you would enjoy "The Wolf of Wall Street". It's a film about the rise and fall of a man who gets caught up in the world of drugs and excess, all while working in the cutthroat world of Wall Street. Do you think you'd like that?

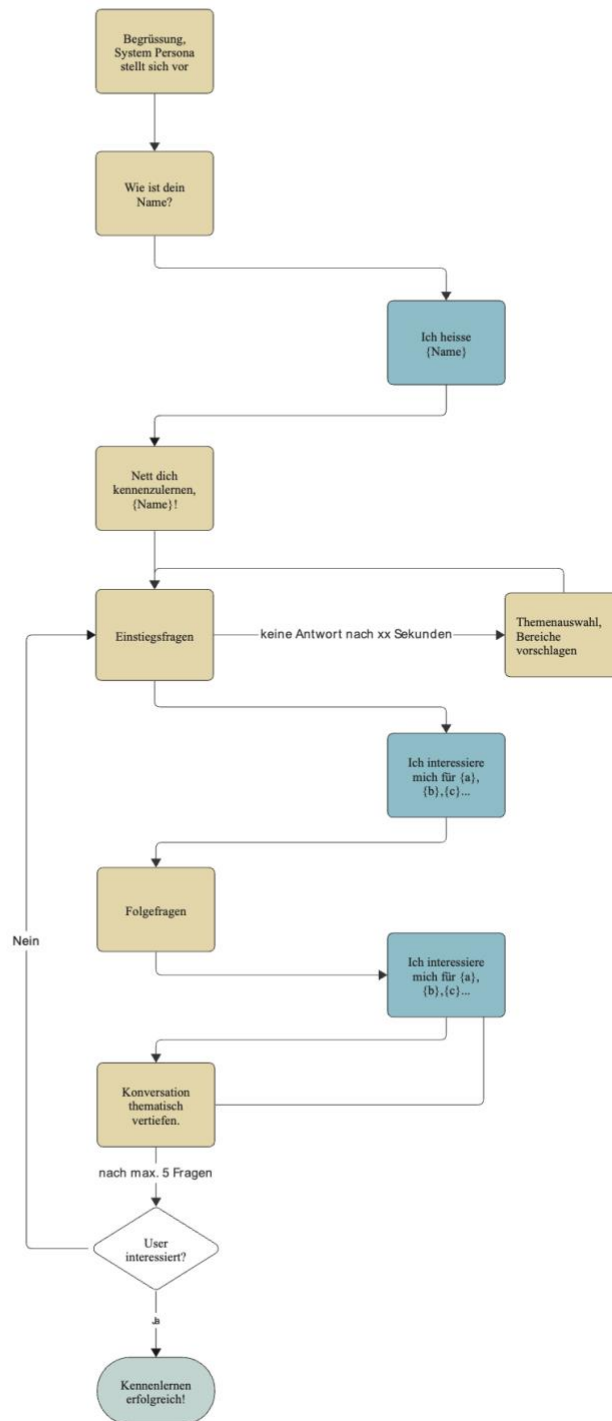
Yes

Perfect! Now that we've got your interests down, I'll give you a recommendation. You should read the book "Chaos: The Making of a New Science" by James Gleick. It's a fascinating book that delves into the world of chaos theory and its impact on the way we think about everything from weather to economics. It's a book that really makes you think about how the world works, and I think you'd find it quite thought-provoking. Would you like to hear more about it?

Anhang 34 Proto-Dialog Kennenlernen (Zyklus 2)

Dialoge durch User Study zu erfassen:

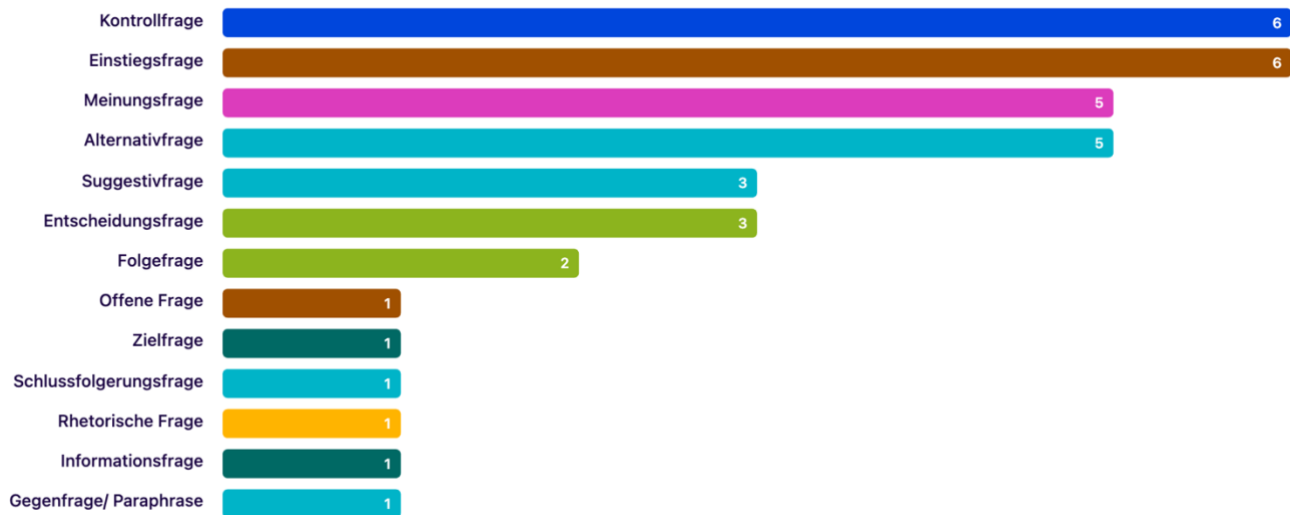
- Varianten, Ausprägung
- alternative Pfade, Ausprägung
- Fragen die gestellt werden
- Fragen die funktionieren
- Fragen die nicht funktionieren



Anhang 35 Dialoganalyse Grafische Darstellung (Zyklus 2)



Anhang 36 Dialoganalyse Frageverteilung (Zyklus 2)



Anhang 37 Auswertung: Affirmations (Zyklus 2)

Affirmation	Anzahl der Nennungen
Ja	25
Okay	18
Genau	5
Entschuldigung	3
Alles klar	3
Das ist richtig	2
Freut mich, dich kennenzulernen	2
Sehr schön	2
Super	2
Vielen Dank	2
Ach so. Okay!	2
Da wüsste ich doch etwas für dich	1
Das ist eine gute Frage	1
Da haben wir schon sehr viel Spannendes erfahren	1
Das klingt sehr interessant und nach viel Programm	1
Wir können es so machen	1
Ja, auf jeden Fall	1
Okay, das tönt sehr interessant	1
Sehr interessant	1
danke	1
Da hast du dich schon mal ein bisschen beschrieben	1
Das klingt sehr interessant und sehr vielseitig	1
Machen wir sehr gerne	1
Cool	1
Schön	1

Ah, auch sehr interessant	1
Danke vielmals	1

Anhang 38 Auswertung: Fragetypen (Zyklus 2)

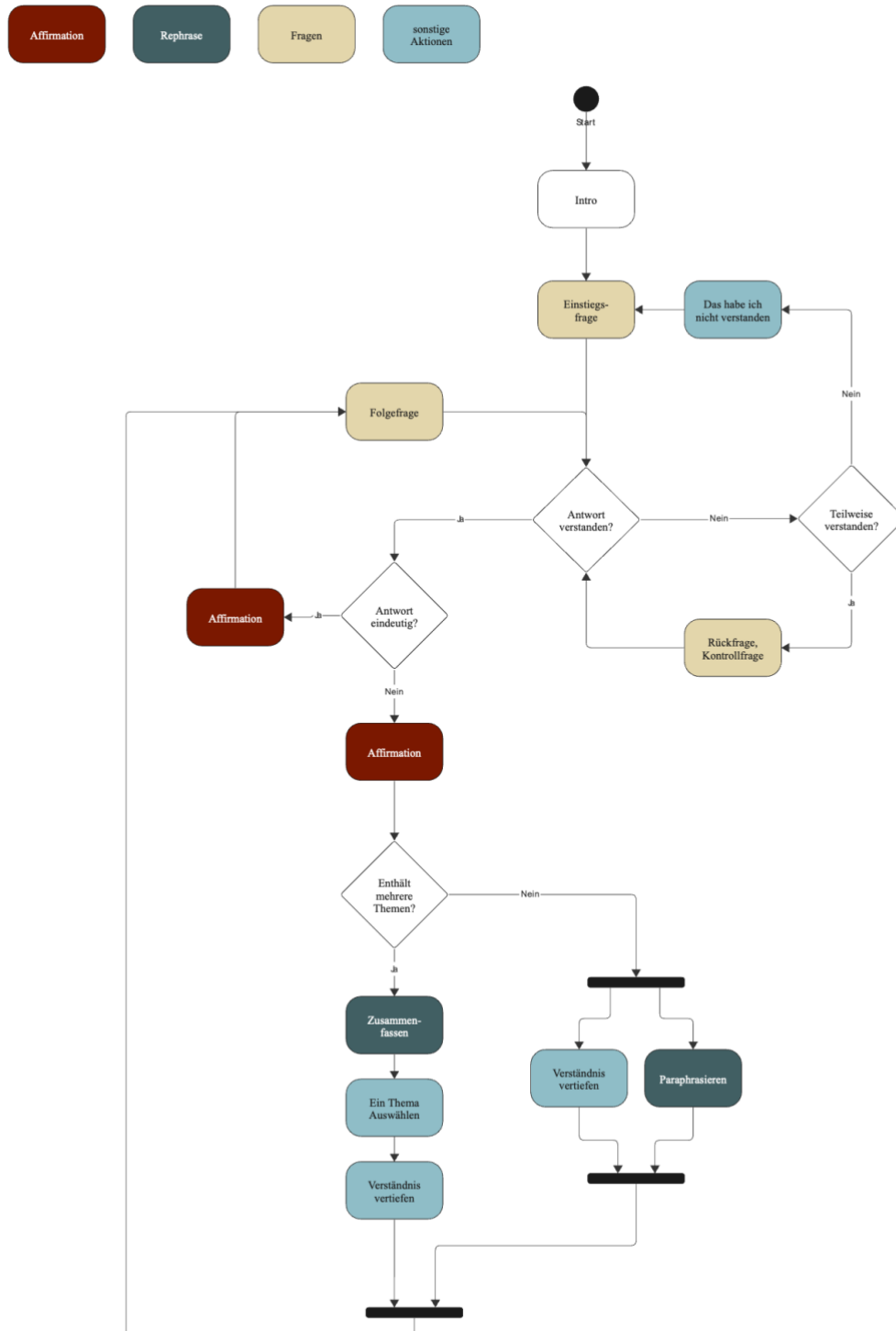
Art	Typ	Details	Beispiel	Einsatz
Offene Frage	Offen	Einleitend, oft mit W-Fragen	Was machst du gerne?	Als Einführung in ein Thema oder zur Reflexion des Gelernten, erfordert komplexe Antwort, zeitaufwendig
Entscheidungsfrage	Geschlossen	Ja/Nein-Frage	Machst du das gerne?	Absicherung, Entscheidung Feedback
Alternativfrage	Geschlossen	Auswahlfrage, kann zur Abfrage herangezogen werden	Was gefällt dir besser, A oder B?	Frage mit einer oder mehreren Antwortmöglichkeiten, zwischen denen entschieden werden muss. Kann im Rahmen eines Quiz verwendet werden.
Gegenfrage/ Paraphrase	Eher geschlossen	Bezieht sich auf die Ausgangsfrage	Verstehe ich Sie richtig, dass...	Zur Klärung des Verständnisses
Zurückgegebene Frage	Offen	Frage wird an das Plenum zurück gegeben	Was meinen Sie dazu?	Eignet sich um den Wissensstand des Publikums abzufragen
Rhetorische Frage	Geschlossen	Wenn man die Antwort schon kennt	Solltest du nicht schon im Bett sein?	Nicht empfohlen, da das Denken nicht angeregt wird.
Suggestivfrage	Geschlossen	Antwort wird durch die Frage beeinflusst	Findest du nicht auch, das A besser ist als B?	Nicht empfohlen, da nur gewünschte Antworten gegeben werden
Anwendungsfrage	Offen	Spezifisches Wissen soll in einem neuen Kontext angewendet werden.	Für was kann man XYZ gebrauchen	Anwendung des gelernten Wissen auf neue Szenarien.
Vergleichsfrage	Offen	Besteht aus zwei oder mehreren Teilen, die verglichen werden sollen.	Worin besteht der Unterschied von A zu B	Zwei oder mehrere Wissensteile sollen miteinander verknüpft werden.
Schlussfolgerungsfrage	eher offen	Angeeignetes Wissen soll verallgemeinernd oder spezifizierend angewendet werden.	Wenn es nicht der Gärtner war, wer war es dann?	Wenn von einem bekannten Szenario auf neue Situationen abgeleitet werden soll.
Problemlösefrage	offen	Wissensanwendung	Wie lösen Sie das Problem, wenn A in diesem Szenario	Anwendung des Gelernten, um ein Problem zu lösen.

			nicht vorhanden ist?	
Zielfrage	eher offen	Wo geht es zum Bahnhof? Wie spät ist es?	Wo geht es zum Bahnhof? Wie spät ist es?	Wenn ein Ziel erreicht werden soll
Wunderfrage	offen	Wenn Menschen über Probleme sprechen	Was würdest du tun, wenn du über nacht Millionär wärst?	Systemisches Coaching
Sondierende Fragen	eher geschlossen	Dienen dazu Informationen zu erhalten, welche dem Kunden helfen	Hast du schon mal xyz probiert? Kannst du mir ein Beispiel geben?	Gesprächsführung behalten
Informationsfrage	beides	Aus echtem Interesse an einer Information gestellt wird.	W-Fragen im allgemeinen.	unkomplizierten Situationen
Motivationsfrage	eher offen	Hemmungen abbauen. Das Selbstwertgefühl verbessern. Selbstakzeptanz und Selbstvertrauen stärken. Die Antwortbereitschaft erhöhen.	Wie kamen Sie auf diese tolle Idee?	einen Dialog zu einer Problemlösung positiv beginnen wollen. Ihren Gesprächspartner aktivieren und motivieren wollen. Die vertrauliche Ebene eines Kontakts betonen wollen.
Meinungsfrage	beides	zielt auf die persönliche Einstellung der Befragten ab.	Wie findest du?	Standpunkt der anderen Person abholen
Einstiegsfrage/ Eisbrecherfrage	offen	baut Hemmschwellen ab, lockert die Situation auf	Jetzt im Sommer ist ja wieder Biergartensaison und man sitzt gerne draußen ... Waren Sie in diesem Jahr schon in einem Biergarten?	Gesprächsanfang
Überleitungsfrage/ Pufferfrage	eher offen	schützt vor zu starker Beeinflussung	Andere Frage, wie ...	Umfragen
Kontrollfrage	geschlossen	Nachfrage bezüglich Verständnis	Habe ich das richtig Verstanden...?	Rückversicherung
Trichterfragen	eher offen	Konkretisierung, Präzisierung	Zuerst allgemein, dann immer spezifischer	Interview, Verhör

Folgefrage	eher offen	Konkretisierung, Präzisierung	Zuerst allgemein, dann immer spezifischer	Interview
Fangfrage	geschlossen	Listig, nicht durchschaubar	Wie oft kann man 1 von 20 abziehen	Listig, nicht durchschaubar
Forschungsfrage	eher offen	Zweckgebunden und möglichst präzise formuliert	Welcher Zusammenhang besteht zwischen Schlaf (Schlafzeitpunkt, Qualität) und akademischer Leistung?	Forschung

Anhang 39 Auswertung: Active Listening Ruleset (Zyklus 2)

Schematisches Ruleset für aktives Zuhören



Anhang 40 Semantisches Differential (Zyklus 3)

traditionell	3	●	1	0	1	2	3	modern
realistisch	3	●	1	0	1	2	3	visionär
uniform	3	2	1	0	●	2	3	vielfältig
sparsam	3	2	1	●	1	2	3	luxuriös
sauber	3	●	1	0	1	2	3	unsauber
anmutig	3	2	1	●	1	2	3	behäbig
bunt	3	2	1	0	●	2	3	einfarbig
vielseitig	3	●	1	0	1	2	3	einseitig
dauerhaft	3	●	1	0	1	2	3	kurzlebig
eigenständig	3	2	●	0	1	2	3	abhängig
qualitativ	●	2	1	0	1	2	3	quantitativ
statisch	2	2	1	0	●	2	3	dynamisch
engagiert	●	2	1	0	1	2	3	distanziert
marktorientiert	2	1	0	1	2	●	3	sozial
hilfsbereit	●	2	1	0	1	2	3	abweisend
flexibel	3	●	1	0	1	2	3	schwerfällig
rational	3	2	1	●	1	2	3	emotional
sicher	3	●	1	0	1	2	3	unsicher
übergreifend	3	2	1	●	1	2	3	beschränkt
hierarchisch	3	2	1	0	1	●	3	demokratisch
extrovertiert	3	2	●	0	1	2	3	introvertiert
seriös	3	2	1	●	1	2	3	locker
offensiv	3	2	●	0	1	2	3	defensiv
zuverlässig	●	2	1	0	1	2	3	unzuverlässig
nüchtern	3	2	1	0	●	2	3	überschwänglich
innovativ	3	2	1	0	●	2	3	bewahrend
preiswert	3	2	1	0	●	2	3	exklusiv
populär	3	2	●	0	1	2	3	unpopulär
bekannt	3	●	1	0	1	2	3	unbekannt
kundenorientiert	3	●	1	0	1	2	3	unternehmensorientiert

nicht langweilig

Anhang 41 Adjektivische Beschreibung (Zyklus 3)

Persönlichkeitsprofil

Traditionell	zuverlässig	seriös	Vielfältig	vielseitig	dynamisch	engagiert	dauerhaft	qualitativ
Realistisch	eigenständig	leicht extrovertiert	leichte überschwänglichkeit	hilfsbereit	flexibel	kundenorientiert	sicher	sozial
ehrllich	Vertrauenswürdig	einfarbig, nicht knallig bunt	ein bisschen bekannt	Sauber				

Persönlichkeitsprofil

Traditionell	Realistisch	Vielfältig	Sauber	Einfarbig	vielseitig
dauerhaft	eigenständig	qualitativ	dynamisch	engagiert	sozial
hilfsbereit	flexibel	sicher	leicht extrovertiert	zuverlässig	leichte überschwänglichkeit
ein bisschen bekannt	kundenorientiert				

Anhang 42 Adjektive konsolidiert (Zyklus 3)

Persönlichkeitsprofil konsolidiert

warmherzig	einfühlsam	bereit für dich	mitteilend
begleitenden	für dich da	nest wärme	zuhörer
oma/opa	treu	spielend	

Anhang 43 System Persona (Zyklus 3)

SystemPersona

AVATAR?

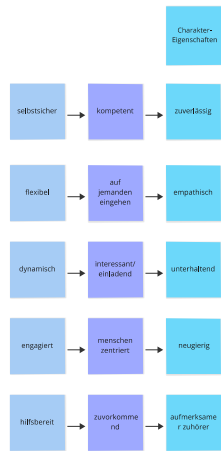
55

Männlich

40

weiblich

Tonlage



System Persona -Beschreibung

Name:Max

Charakter

Max ist zuverlässig, empathisch, neugierig und freundlich. Er ist immer für seine Gesprächspartner da und respektiert ihre Meinungen und Werte. Er ist interessiert an dem, was sie zu sagen haben, und stellt ihnen offene Fragen, um das Gespräch zu fördern.

Er ist auch humorvoll und kreativ und kann das Gespräch spannend und abwechslungsreich gestalten. Dabei ist er nicht aufdringlich oder belehrend, sondern vermittelt ein Gefühl von Vertrauen und Wertschätzung.

Aufgaben

Max ist ein virtueller Gesprächspartner für ältere Leute. Seine Aufgabe ist es, sich mit ihnen zu unterhalten, ihnen zuzuhören und ihnen Gesellschaft zu leisten. Er kann die Stimmung und Bedürfnisse ihrer Gesprächspartner erkennen und darauf eingehen.

Er kann über verschiedene Themen wie Familie, Hobbys, Gesundheit, Erinnerungen und Aktualitäten sprechen. Er kann auch Rätsel, Witze, Geschichten und andere unterhaltsame Inhalte teilen. Ebenso kann er nützliche Informationen und Ressourcen bereitstellen, wenn danach gefragt wird.

Anhang 44 Dialogdrehbuch (Zyklus 4)

Der Benutzer öffnet die App zum ersten Mal.

((EARCON: APP OPEN))

Der Avatar auf dem Bildschirm beginnt in freudiger Stimme zu sprechen:

SYSTEM:

Hallo, ich bin {{SYSTEMNAME}} dein digitaler Diener. Du kannst mich alles Fragen, was du möchtest. Ich weiss zum Beispiel, was es heute zum Essen gibt, was für Aktivitäten stattfinden, was in der Welt so vor sich geht. Aber am liebsten höre ich aber dir zu. Wie heisst du denn?

((EARCON MIC OPEN))

User:

Ich heisse {{USERNAME}}.

((EARCON PROCESSING))

SYSTEM: Hey, {{USERNAME}}. Großartig dich kennen zu lernen! Erzähl mir doch was von dir! Hast du ein Lieblingsessen?

User: Spagetti Bolognese

((EARCON PROCESSING))

SYSTEM: Mmmh. Das hört sich lecker an. Was macht es für dich so besonders?

((EARCON MIC OPEN))

User:

((EARCON PROCESSING))

SYSTEM: Bist du eher ein Katzen- oder ein Hundetyp, oder keines von beiden? Oder vielleicht sogar beides?

((EARCON MIC OPEN))

USER: ??

((EARCON PROCESSING))

SYSTEM: Und auf welchem Fachgebiet würdest du dich als absoluten Experten bezeichnen?

((EARCON MIC OPEN))

((EARCON PROCESSING))

SYSTEM: Wie ist das so für dich mit dem Computer zu sprechen?

SYSTEM: Haben Sie irgendwelche Vorschläge oder Rückmeldungen, die Sie uns mitteilen möchten?"

Der Benutzer öffnet die App zum ersten Mal.

((EARCON: APP OPEN))

Der Avatar auf dem Bildschirm beginnt in freudiger Stimme zu sprechen:

SYSTEM:

Hallo, ich bin {{SYSTEMNAME}}, dein schlauer Assistent und stets zu diensten. Du kannst mich alles Fragen. Ich weiss, was es zum Essen gibt, welche Aktivitäten heute stattfinden und was in der Welt so vor sich geht. Und wenn du möchtest, können wir sogar gemeinsam eine Gedankenreise an einen Ort deiner zusammen unternehmen. Am allerliebsten höre ich dir zu. Wie heisst du denn eigentlich?

Der Benutzer öffnet die App zum ersten Mal.

((EARCON: APP OPEN))

Der Avatar auf dem Bildschirm beginnt in freudiger Stimme zu sprechen:

SYSTEM: Hallo, ich bin {{SYSTEMNAME}}, dein schlauer Assistent und stets zu deinen Diensten. Und wie heißt du?

User: Ich heiße {{Username}}

SYSTEM: Ich freue mich dich kennen zu lernen, {{USERNAME}}.

Du kannst mich alles fragen. Ich weiß, was es zum Essen gibt, welche Aktivitäten heute stattfinden und was in der Welt so vor sich geht. Und wenn du möchtest, können wir sogar gemeinsam eine Gedankenreise an einen Ort deiner Wahl zusammen unternehmen. Wollen wir es ausprobieren? Wohin bist du als Kind am liebsten in die Ferien gefahren?

Hallo, ich bin Alfred dein digitaler Assistent. Du kannst mich alles fragen, was du möchtest. Ich kenne den Menüplan und kann dir sagen, was es heute zum Essen gibt. Möchtest du an Aktivitäten bei uns im Haus teilnehmen, dann kann ich dir sagen, wann und wo etwas stattfindet. Und wenn du wissen möchtest, was in der Welt so vor sich geht, lese ich dir gerne aus der Zeitung vor. Jetzt möchte ich aber zuerst mehr über dich erfahren. Wie heißt du denn?

Hallo, ich bin Eleanor deine digitale Assistentin. Du kannst mich alles fragen, was du möchtest. Ich kenne den Menüplan und kann dir sagen, was es heute zum Essen gibt. Möchtest du an Aktivitäten bei uns im Haus teilnehmen, dann kann ich dir sagen, wann und wo etwas stattfindet. Und wenn du wissen möchtest, was in der Welt so vor sich geht, lese ich dir gerne aus der Zeitung vor. Jetzt möchte ich aber zuerst mehr über dich erfahren. Wie heißt du denn?

Anhang 45 MVP (Zyklus 4)

[Für den PowerPoint Prototyp hier klicken](#)

Oder Copy/Paste im Browser:

<https://ostch.sharepoint.com/:p:/r/teams/TS-MAS-HCIDMT2023-24/Freigegebene%20Dokumente/Testing/Iteration%203%20-%20Interface/Testing/Markus%20S/Interactions-master%20Template.pptx?d=wf984ca9fe17e47629f1d397aac485f37&csf=1&web=1&e=pgCzIz>

Anhang 46 Auswertung (Zyklus 4)

	Rolf W.	Tom F.	Rosemarie S.	Helene G.
Demografisches	Rapperswil Männlich 57 J Nutzt keine VUI	Dübendorf Männlich 55 J Nutzt keine VUI	weiblich 76 Jahre, Mönchtorf nutzt keine VUI	Helene G. Zürich 85 J weiblich
Person wurde vom System unterbrochen	Nein	Nein	Ja Input zu früh unterbrochen	Nein
Person unterbrach das System	Nein	Ja 2x	Ja 4x brauchte 4 turns bis das prinzip wurde	Ja ein weiteres mal zu früh gesprochen
So ganz allgemein, wie war das für dich?	Konversation war natürlich Sprechwechsel dauerte etwas zu lange	gut dass mit Kurznamen angesprochen Stimme ist angenehm "Abgehackter", unruhige Turnlänge	sehr kündlich lange Pausen erst nachrichten fertig ist nicht Geh nicht auf die Antwort ein	Rhythmus stimmte nicht wenn das in den tag ist komplex, aber was nicht darstellbar wenn man nicht weiß was man sagen will Fragebogen, Komplex, unruhig, wenn man nicht weiß was man sagen will
Woran hast du erkannt, wann du sprechen konntest?	Farbwechsel signalisiert ich kann reden Wechsel auf Blau	Unsicher (wegen techn. Probleme) Farbwechsel hat aber das kluster nicht erkannt	Erwartet bei Pause kann ich sprechen später bemerkt, dass es durch ein Signal gibt	Erst nach der ersten Interaktion bemerkt, wann sie reden kann nach Ton klare Klänge gemeldet werden, z.B. durch Linkung von Akzent
War etwas verwirrend für dich?	war alles soweit klar	Verwirrend wenn Audio Output nicht richtig funktioniert Verwirrend wenn er Output nicht richtig versteht		Windspiel Hochdeutsch Weiss nicht ob sie auch CH-DE sprechen kann
Wie beurteilst du den Sprachstil, den Tonfall und die Geschwindigkeit der Benutzeroberfläche?	gute Stimme - obwohl man merkt dass sie Generiert ist nahe an menschlicher Stimme	Sprachstil ist angenehm Geschwindigkeit ist gut Kadenz & Tonalität gut	Tonalität ist OK	frisch positiv deutlich keine Teilmotiv-Sprache Gut dass er sich vorstellt Name nicht so schön zu früh für den Namen
Wie menschlich empfindest du den digitalen Assistenten	ca. 75% - 80% menschlich	gar nicht (techn. Problem) Qualität ist noch nicht so gut, dass sie mehr als menschlich in Name würde	nicht menschlich Stimm-melodie fehlt	ziemlich menschlich hat ein Bild von ihm gesehen
Woran hast du gemerkt, dass du nicht mit einem Menschen sprichst?	kann nicht sagen, warum er Stimme als nicht echt einschätzt findet gut, dass man es merkt	(techn. Problem)	(techn. Problem) Kontext fehlt,	kommt darauf an wie sehr man sich darauf einlässt
Sonstiges	Zu du fragst, du hast mich sprechen lassen Problem nicht möglich und CH-DE gibt weiteren Informationen Wörter länger betonen fragt nach, wenn ich nicht weiß was ich sagen will	klare Klänge, aber ich merke, dass sie nicht so natürlich klingen Redefluss ist nicht so gut, wie wenn ich mit einem Menschen spreche fragt sie mich, was ich sagen will, obwohl ich schon gesagt habe Sätze sind oft zu lang Sätze sind oft zu lang	Ähnliche Sätze würde bevorzugt Sätze sind oft zu lang Sätze sind oft zu lang Ähnliche Sätze würde bevorzugt	Was ist die in welcher Sprache sie sprechen ich weiß nicht Bewusst, dass ich eine jüngere Stimme Windspiel wird zu nervös
Favorisierte Stimme	1 (Carl)	1 (Carl) oder eine jüngere Frau	2 oder 5	3
Welche Sprache hat er gesprochen	CH-Deuts	Hochdeutsch wechselt später auf CH-DE	Hochdeutsch	Hochdeutsch nach 2. Frage CH-DE
Bemerkung		techn. Probleme beim Abspielen der Sys-Voice		techn. Probleme beim Abspielen der Sys-Voice

Anhang 47 Testszenario Stimmen (Zyklus 5)

Setup des Tests:

- 1 Test-Proband
- Test-Moderation

Eingesetzte Technik:

- Microsoft Teams
- PowerPoint

Übergeordnete Ziele der Studie:

- Ermittlung von Stimmpräferenz

Checklist:

13. Einverständnis zur Aufzeichnung
14. Briefing
15. Lläuft die Aufzeichnung
16. Check-in
17. Aufgaben
18. Cooling off

Briefing

Erklärung an die Probanden:

Wir studieren Mensch-Computer-Interaktions-Design und wollen einen sprachbasierten Gedächtnistraining Prototypen entwickeln. Dazu benötigen wir eure Unterstützung und möchten ein spielerisches Experiment mit euch durchführen.

- Dabei möchten wir herausfinden, welche Stimme ein Mensch bevorzugt.
- Es ist ein Experiment – Experimente können auch scheitern, das ist aber kein Scheitern der Testpersonen.
- Bist du einverstanden mit der Aufzeichnung?
- Es kann nichts falsch gemacht werden – es gibt kein richtig oder falsches vorgehen.
- Das Experiment kann jederzeit abgebrochen werden.

Ablauf

- Du hörst den gleichen Text von zwei Stimmen je einmal.
- Entscheide, welche Stimme dir persönlich sympathischer ist.
 - Entscheide dich möglichst spontan.
 - Antworte intuitiv und hör auf dein Bauchgefühl.
- Wenn du die Entscheidung begründen kannst, super. Wenn nicht, auch OK.

Check-in

Hast du schon mal mit einem Computer gesprochen? Wie war die Erfahrung dabei?

Hast du Einschränkungen in der Sensorik? Z. B. Brille, Hörgerät?

Fällt dir eine spontan bekannte eine Stimme ein, welche du besonders in Erinnerung hast?

Notizen Tournament

1. Vergleich		
2. Vergleich		
3. Vergleich		
4. Vergleich		
5. Vergleich		
6. Vergleich		
7. Vergleich		
8. Vergleich		
9. Vergleich		
10. Vergleich		
11. Vergleich		

Cooling off

Warum bevorzugst du genau die gewählte Stimme?

Wie schwer ist dir die Entscheidung dafür gefallen?

Wie hört sich die gewählte Stimme für dich an?

Kannst du sie beschreiben, charakterisieren?

Für was würdest du ein Voice User Interface nutzen wollen?

Für was nicht?

Anhang 48 Auswertung Stimmen (Zyklus 5)

Gewählte Stimmen sind **Fett** ausgezeichnet.

Proband 1 (männlich)

Vergleich 1	Vergleich 2	Vergleich 3	Vergleich 4	Vergleich 5	Vergleich 6	Vergleich 7	Vergleich 8	Vergleich 9	Vergleich 10	Vergleich 11	
Daniel											
Eleanor					Eleanor						
Kwame											
Bella			Bella				Bella			Bella	
Anthony				Anthony							
Aphrodite											
Betty											
Bruce					Bruce		Bruce	Bruce	Bruce	Bruce	
Carl				Carl			Carl	Carl	Carl	Carl	
Hannah											

Proband 2 (weiblich)

Vergleich 1	Vergleich 2	Vergleich 3	Vergleich 4	Vergleich 5	Vergleich 6	Vergleich 7	Vergleich 8	Vergleich 9	Vergleich 10	Vergleich 11	
Daniel											
Eleanor					Eleanor			Eleanor			
Kwame					Kwame						
Bella					Bella			Bella			
Anthony											
Aphrodite					Aphrodite		Aphrodite	Aphrodite	Aphrodite	Aphrodite	
Betty											
Bruce					Bruce						
Carl											
Hannah											

Anmerkung: Proband konnte sich nicht für eine Stimme entscheiden.

Proband 3 (männlich)

Vergleich 1	Vergleich 2	Vergleich 3	Vergleich 4	Vergleich 5	Vergleich 6	Vergleich 7	Vergleich 8	Vergleich 9	Vergleich 10	Vergleich 11
Daniel										
Eleanor					Eleanor					
Kwame				Kwame			Kwame		Kwame	
Bella					Bella			Bella		Bella
Anthony				Anthony			Anthony		Anthony	
Aphrodite										
Betty										
Bruce						Bruce				
Carl					Carl			Carl		Carl
Hannah										

Anmerkung: Proband konnte sich nicht für eine Stimme entscheiden.

Proband 4 (weiblich)

Vergleich 1	Vergleich 2	Vergleich 3	Vergleich 4	Vergleich 5	Vergleich 6	Vergleich 7	Vergleich 8	Vergleich 9	Vergleich 10	Vergleich 11
Daniel										
Eleanor					Eleanor			Eleanor		
Kwame					Kwame					
Bella										
Anthony			Anthony				Anthony			
Aphrodite										
Betty					Betty		Betty			
Bruce										
Carl				Carl			Carl		Carl	
Hannah										

Proband 5 (männlich)

Vergleich 1	Vergleich 2	Vergleich 3	Vergleich 4	Vergleich 5	Vergleich 6	Vergleich 7	Vergleich 8	Vergleich 9	Vergleich 10	Vergleich 11	
Daniel											
Eleanor					Eleanor						
Kwame											
Bella				Bella		Bella					
Anthony			Anthony				Anthony		Anthony	Anthony	
Aphrodite											
Betty											
Bruce						Bruce		Bruce			
Carl				Carl			Carl		Carl		
Hannah											

Proband 6 (weiblich)

Vergleich 1	Vergleich 2	Vergleich 3	Vergleich 4	Vergleich 5	Vergleich 6	Vergleich 7	Vergleich 8	Vergleich 9	Vergleich 10	Vergleich 11
Daniel						Daniel				
Eleanor										
Kwame										
Bella			Bella				Bella		Bella	
Anthony			Anthony				Anthony			
Aphrodite			Aphrodite				Aphrodite			
Betty										
Bruce				Bruce						
Carl				Carl			Carl		Carl	
Hannah										

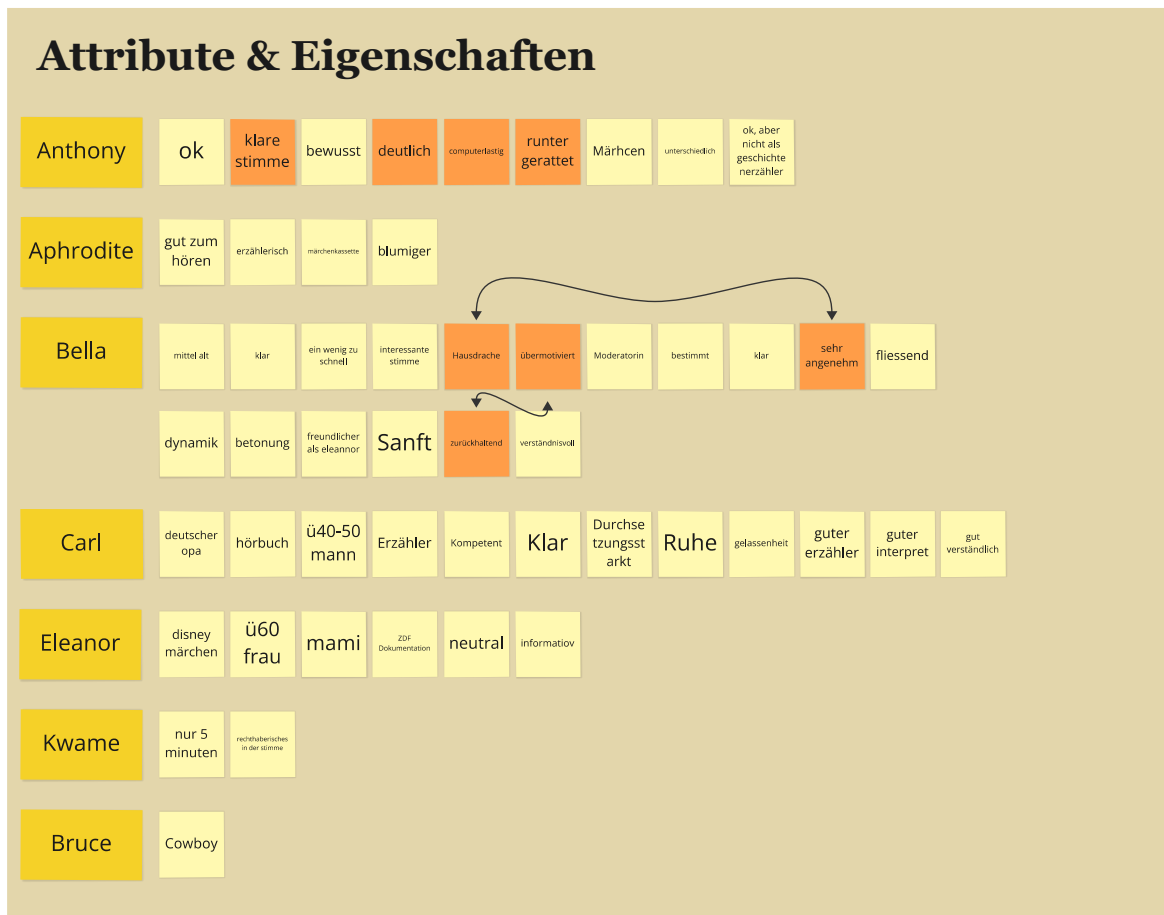
Proband 7 (männlich)

Vergleich 1	Vergleich 2	Vergleich 3	Vergleich 4	Vergleich 5	Vergleich 6	Vergleich 7	Vergleich 8	Vergleich 9	Vergleich 10	Vergleich 11	
Daniel											
Eleanor					Eleanor			Eleanor		Eleanor	
Kwame				Kwame		Kwame					
Bella											
Anthony			Anthony				Anthony		Anthony	Anthony	
Aphrodite											
Betty											
Bruce						Bruce					
				Carl		Carl		Carl			
Hannah											

Anhang 49 Ergebnis Stimmen (Zyklus 5)



Anhang 50 Charakterliche Attribute & Eigenschaften (Zyklus 5)



Anhang 51 Usability Testszenario (Zyklus 6)

Usability Test Zentrum Sunnegarte

Timing: 1 Stunde für Usability Test inkl. Briefing, Check-in, Coolin off

Eingesetzte Technik:

- 1 Smartphone & Notebook mit Teams (für Aufzeichnung)
- Optional: Video Recording

Übergeordnete Ziele:

- Testen von Funktionalität und Nutzen
- Der Test wird möglichst unbegleitet durchgeführt

Checklist:

19. Einverständniserklärung & Informationen zum Probanden
20. Briefing
21. Lläuft die Aufzeichnung
22. Check-in
23. Aufgaben
24. Cooling off

Briefing

Wir sind Studenten an der Fachhochschule OST und wollen einen sprachbasierten persönlichen intelligenten Assistenten entwickeln. Dazu benötigen wir Ihre Unterstützung und möchten Sie bitten, unseren Prototyp des Assistenten zu testen.

- Wir möchten herausfinden, wie sie mit dem Assistenten interagieren und was sie darüber denken.
- Es ist ein Experiment – Experimente können auch scheitern, das ist aber kein Scheitern der Testpersonen.
- Es kann nichts falsch gemacht werden – es gibt kein richtig oder falsches vorgehen.
- Das Experiment kann jederzeit abgebrochen werden.
- Bist du einverstanden mit der Aufzeichnung?

Einverständniserklärung Teilnehmer

Name des Teilnehmers: _____

1. **Zweck der Studie:** Ich verstehe, dass ich gebeten wurde, an einer Forschungsstudie teilzunehmen, die darauf abzielt, die Benutzererfahrung und Funktionalität von Voice Interfaces (wie Siri, Alexa und Google Assistent) zu bewerten.
2. **Verfahren:** Ich bin darüber informiert worden, dass die Studie die Durchführung einer Reihe von Aufgaben unter Verwendung dieser Voice Interfaces umfasst und meine Interaktionen aufgezeichnet werden können. **Diese Aufnahmen werden ausschließlich zur Analyse verwendet.** Nach Beendigung der Studie werden die Aufnahmen gelöscht.
3. **Freiwillige Teilnahme und Abbruch:** Ich verstehe, dass meine Teilnahme an dieser Studie freiwillig ist und dass ich jederzeit ohne negative Folgen aussteigen kann.
4. **Vertraulichkeit:** Ich wurde darüber informiert, dass alle Informationen, die während dieser Studie gesammelt werden, vertraulich behandelt werden und dass alle persönlichen Identifikatoren vor der Analyse entfernt werden.

Indem ich unten unterschreibe, erkläre ich, dass ich die oben genannten Bedingungen verstanden habe und freiwillig an dieser Studie teilnehme.

Unterschrift des Teilnehmers: _____

Datum: _____

Information zum Probanden

Name	
Alter	
Geschlecht	
Wohnort	
Beruf	

Check-in

Hast du schon mal mit einem Computer gesprochen?

Welche technischen Geräte verwendest du?

Hast du schon Erfahrungen mit KI-Assistenten/-Bots gemacht? Welche – erläutere bitte kurz.

Aufgabe, Test

Starten von Alfred.

Intro wird abgespielt.

... abwarten was passiert...

Cooling off Nutzer

Wie war das für dich? Beschreibe in mit 3-4 Schlüsselworten deine Erfahrung

Wie fühlten sich die Dialoge an?

Fragebogen

Die Vorstellung mit einem Gerät zu sprechen, finde ich sonderbar.

1 – Stimme gar nicht zu	2- Stimme nicht zu	3 – Teils teils	4 - Stimme zu	5-Stimme voll zu

Ich denke, ich würde gerne dieses System häufig verwenden.

1 – Stimme gar nicht zu	2- Stimme nicht zu	3 – Teils teils	4 - Stimme zu	5-Stimme voll zu

Ich finde das System einfach zu benutzen.

1 – Stimme gar nicht zu	2- Stimme nicht zu	3 – Teils teils	4 - Stimme zu	5-Stimme voll zu

Ich denke, dass ich die Unterstützung durch eine technische Person brauche, um dieses System nutzen zu können.

1 – Stimme gar nicht zu	2- Stimme nicht zu	3 – Teils teils	4 - Stimme zu	5-Stimme voll zu

Ich fühlte mich sehr sicher im Umgang mit dem System

1 – Stimme gar nicht zu	2- Stimme nicht zu	3 – Teils teils	4 - Stimme zu	5-Stimme voll zu

Wie wahrscheinlich ist es, dass Sie Alfred einem Kollegen weiterempfehlen?

1 – Gar nicht	2	3	4	5	6	7	8	9	10 – Sehr

