

Talking (to) Things

**Eine Untersuchung von
Sprachassistenten und deren
Konversation-Aspektes im Kontext
von Inklusion und Lebenswelten**

Talking (to) Things, Adrian Demleitner
HKB BFH MA Design, Colloquium 3 - November 2021

Unter <https://ttt.things.care> ist die zum Projekt
gehörende Dokumentation einsehbar.

Abstrakt

Ich beschäftige mich in meiner Forschungsarbeit mit dem Potenzial von Sprachassistenten für Menschen mit Seh- oder Mobilitätsbeeinträchtigungen. Dabei interessiert mich inwiefern diese Technologien den Zugang zur Welt vermitteln und dabei assistieren können und wie die Interaktion gestaltet werden muss, damit der Fokus auf den realen Bedürfnissen dieser Menschen liegt.

Auf die Weiterentwicklung des gegenwärtigen Fokus dieses Forschungsvorhabens werde ich in der Präsentation eingehen.

In den letzten zehn Jahren haben sich Sprachassistenten wie Siri, Alexa und Google Home gesellschaftlich etabliert und gehören zu einem der stärksten wachsenden Bereiche der Consumer-Technologien. Sprachassistenten erweitern herkömmliche Sprachinterfaces, für das Diktat oder zur Sprachsteuerung, um den Aspekt der Konversation. Die weiterentwickelte Leistungsfähigkeit dieser Technologien sowie Innovationen bezüglich der Mensch-Maschine Interaktion werden vorerst im Bereich der Unterhaltung eingesetzt. Dabei bieten sich diese Potenziale für den Einsatz im Bereich von inklusiven Mensch-Maschinen-Interfaces an, allen voran für Menschen mit Sicht- oder Mobilitätsbeeinträchtigungen.

Stand der Forschung und eigener Beitrag

Sprachassistenten haben sich in den letzten zehn Jahren etabliert und sind damit auch ein beliebtes Forschungsobjekt geworden. Besonderes Augenmerk hatten bislang die der Technologie zugrundeliegenden Algorithmen und künstlichen Intelligenzen und deren Bias (Bajorek, 2019; Danielescu, 2020), die Ausgestaltung der Persona (Danielescu, 2020; Humphry & Chesher, 2020), Probleme rund um die Privatsphäre (Easwara Moorthy & Vu, 2014; Pal et al., 2020), die Postphänomenologie von solchen Assistenten (Hector & Hrcnal, 2020; Wiltse, 2020) sowie im kleineren auch Sprachassistenten als Hilfsgeräte für Menschen mit Beeinträchtigungen (Masina et al., 2020; Pradhan et al., 2018).

Begriffserklärung zur Postphänomenologie unter <https://tth.thingscare/research/glossary/postphenomenology>

Vordergründig liegen die Probleme in der profitgetriebenen Entwicklung und Verbreitung von Sprachassistenten durch die Marktführer Google, Amazon und weitere. Der Fokus auf die Unterhaltungsindustrie sowie das Sammeln von Daten (Daten-Extraktivismus) nehmen

Zumindest Amazon und Microsoft betreiben eigene Disability-Labs, respektive investieren in die Forschung für Menschen mit Beeinträchtigung. Jedoch lassen sich solche Ansätze nur schwer in den Endprodukten erkennen.

dabei einen besonderen Stellenwert ein. Die meisten meiner Teilnehmer:innen waren sich dieser Praktiken sehr wohl bewusst, sahen sich aber ausser Stande etwas an dieser Situation zu ändern. Ähnlich ergeht es Menschen mit Beeinträchtigung.

In den angeführten Untersuchungen wird der Sprachassistent vor allem aus technischer, sozialer oder legaler Sicht angegangen. Ich schlage in meinem Vorhaben einen Fokus auf die Postphänomenologie vor, welcher nahe an den Erfahrungen der Benutzer:innen vor allem die Mediation zwischen diesen und deren Lebenswelt untersucht wird. Mit dieser Öffnung hin zur User-Experience kann für Design relevante Fragestellungen nachgegangen werden.

Hypothese

Bisherige Sprachassistenten-Technologien sind vor allem ein Abhilfe, um Lebenswelten, Inhalte oder Prozesse, welche für der Norm zugehörige Menschen gestaltet wurde, Menschen mit Beeinträchtigungen zugänglich zu machen. Ein klassisches Computer-Terminal bedarf in der Regel Hände und Sehvermögen, um bedient werden zu können. Ein *Screenreader* oder eine Sprachsteuerung versuchen nun, als zweitklassige Schnittstelle an dieses Interface anzuknüpfen.

Sprachassistenten können innerhalb dieser Rahmung als ein alternative, *disability-first* Interfaces betrachtet werden, welche die gesprochene und gehörte Sprache als primären Modus haben. Dabei ist der Aspekt der Konversation von speziellem Interesse. Nebst der Möglichkeit unpräziser Inputs bringen Konversationen sowie die dem Sprachassistenten zugrunde liegenden Algorithmen weiterhin Dimensionen von Zeitlichkeit sowie Erinnerungen mit sich. Diese Aspekte ermöglichen einen niederschwelligeren Zugang als herkömmliche Technologien, indem sie weniger Expert:innen-Wissen Einlern-Zeit fordern sowie eine höhere Trainings- und Anpassungs-Fähigkeit mit sich bringen.

Begriffserklärung zu disability-first unter <https://tth.things.care/research/glossary/disability-first>

Forschungsfrage

Wie müssen Sprachassistenten, sowie die Interaktion mit diesen, gestaltet sein, damit diese Technologien für Menschen mit Sicht- oder Mobilitätsbeeinträchtigung zu einem disability-first Interface werden?

Untergeordnete oder weitergehende Fragen

Wie müssen Sprachassistenten beschaffen sein, damit Menschen mit Beeinträchtigungen mithilfe dieser ihre Lebenswelten aktiv mitgestalten können? Wie werden Machtverhältnisse, zwischen Benutzer:innen, Technologie und Gesellschaft über eine sprachgesteuerte Mensch-Maschinen-Schnittstelle verhandelt? Wie kann zum Beispiel innerhalb dieser Rahmung Privatsphäre zwischen Mensch und Maschine ausgehandelt werden? Wie können wir Sprachassistenten abseits von Unterhaltung und Daten-Extraktivismus weiterdenken?

Ziele, Daten und Methoden

Ziel des interdisziplinär angelegten Forschungsvorhabens ist es, die bislang unerforschten Potenziale und Problemstellen von Sprachassistenten zu ermitteln, mit Fokus auf eine Relevanz im Einsatz für Menschen, welche existentiell auf solche Interfaces und Schnittstellen angewiesen sind. Dabei besteht ein spezifisches Interesse an der Verhandlung von Privatsphäre innerhalb der Rahmung von Mensch-Maschine-Interaktion. Denkbar wäre zum Beispiel, die Dialoghaftigkeit dieser Technologie und deren Einsatzmöglichkeiten weiterzuerfolgen. Der materialisierte Output soll nebst der wissenschaftlichen Analyse und Publikation ein Leitfaden zur Aushandlung von Privatsphäre mit Sprachassistenten sein, welcher durch ausführbare Beispielen unterstützt wird. Auf diesem kann in der Praxis wie auch in weiteren Forschungsvorhaben aufgebaut werden.

Erweiterte Literatur-Review

Der Diskurs um assistierende Technologien wird schon lange geführt. Er soll hier auf seine Aspekte bezüglich Design und Sprachassistent-Technologien untersucht und aufbereitet werden.

Interviews und teilnehmende Beobachtung

Um die realen Bedürfnisse sowie die technischen Möglichkeiten möglichst gut in Betracht ziehen zu können, wird eine Interview-Reihe durchgeführt. Ich möchte mit jeweils vier Personen aus den folgenden zwei Gruppen sprechen:

- *Menschen mit Seh- oder Mobilitätsbeeinträchtigungen welche auf Sprachassistent-Technologien angewiesen sind*
- *Expert:innen welche im Bereich assistierender Technologien und Sprachassistent-Technologien aktiv sind, vorzugsweise Personen welche an den jeweiligen Disability-Labs der Marktführer arbeiten*

Die transkribierten Interviews und Erkenntnisse bilden die Basis für die partizipativen Workshops sowie das Prototyping.

Partizipative Workshops und Fokusgruppen

Die partizipativen Workshops und Fokusgruppen stehen in einem iterativen Prozess mit dem Prototyping zueinander und bilden das Herzstück des Forschungsvorhabens. Die Workshops werden in enger Zusammenarbeit mit Menschen mit Beeinträchtigungen abgehalten, um sicherzustellen, dass die Phasen des Prototyping die richtige Stossrichtung haben. Folgende Fragen sollen in den Workshops in Bezug auf die Bedürfnisse von Menschen mit Seh- oder Mobilitätsbeeinträchtigungen angegangen werden.

- *Was braucht es bezüglich Konversations-Design und muss die Interaktion zwischen Mensch und Assistent gestaltet sein?*
- *Was müssen Hardware, aber auch etwaige Peripherie-Geräte können?*
- *Wie müsste ein entsprechendes Internet of Things Ökosystem aussehen?*
- *Wie muss ein Sprachassistent in den Leben und Lebenswelten von Menschen mit Beeinträchtigungen eingebettet sein?*

Dem iterativen Vorgehen in kurzen Entwicklungsschritten liegen agile Organisations-Methoden zugrunde. (Ozkaynak et al., 2021) Die Fokusgruppen dienen, ähnlich der Sprint-Review und Retrospektive in Scrum, zur Evaluation, Diskussion und weiteren Planung nach Abschluss einer Prototyping-Phase. («Focus Groups», 2017)

Iteratives Prototyping

Durch die technische Natur des Projektes, sowie dem erfahrungsgemässen zusätzlichen Aufwand in technischen Umsetzungen, sollte zumindest für die Prototyping-Phasen ein:e zusätzliche:r Entwickler:in engagiert werden. Diese muss eine Nähe zum Thema, zur Arbeit mit Menschen mit Beeinträchtigungen mit sich bringen.

In dieser iterativen Herangehensweise wird das Prototyping verschiedene Umsetzungsgrade durchlaufen und aufeinander aufbauen, von Mockups und gespielten Interaktionen hin zu technisch umgesetzten Assistenten. (Matthews & Wensveen, 2015; Tripp & Bichelmeyer, 1990) Resultate aus den Prototyping-Phasen fliessen in die Folge-Workshops ein. Es sind vier Zyklen im Laufe eines Jahres geplant, wobei ein Zyklus einen Workshop, eine Prototyping-Phase und eine Fokusgruppe umfasst.

Erweiterte Dokumentation

Ich liste die Dokumentation bewusst unter Methoden auf, da sie gerne auf ein Müssen reduziert wird. Dokumentation ist aber, vor allem auch in der Entwicklung von technischen Produkten, ein unabdingbarer Prozess, welcher Reflexion, Zugänglichkeit und Transparenz schafft. Insofern soll die Dokumentation als Praxis gelebt Grundlagen für den Leitfaden als Output schaffen. Dabei werde ich auf den reflexiven Prozessen der thematischen Analyse aufbauen. (Braun et al., 2019)

Zeitplan

Phase 1 – Literaturstudium; 6 Monate

- *Verortung innerhalb eines Design-wissenschaftlichen Diskurses mit Analyse der Schnittmengen zu Disability Studies*
- *Analyse und theoretische Einordnung innerhalb relevanter und zeitgenössischer Diskurse innerhalb Mensch-Maschine Interaktion und User Experience Design*

Phase 2 – Qualitative Evaluation, Bedarfsanalyse; 6 Monate

- *Interviews sowie teilnehmende Beobachtung mit Menschen mit Beeinträchtigungen welche auf Sprachassistenten-Technologien angewiesen sind*
- *Experteninterviews mit Entwickler:innen und Gestalter:innen solcher Technologien; z. B. bei Google, Amazon oder Spezialeinheiten wie das Microsoft Disability-Lab*
- *Organisatorische Vorbereitungen für Entwicklungs-Jahr*

Phase 3 – Workshops, Fokusgruppe, Prototyping; 12 Monate

- *Konzeption, Durchführung und Auswertung von konkreten partizipativen und iterativen Workshops zur Frage der Gestaltung von Mensch-Maschinen Interaktion mit Sprachassistenten*
- *Iteratives Prototyping von Ansätzen, welche in den Workshops entstehen und überprüft werden*

Phase 4 – Auswertung und Finalisierung; 6 Monate

- *Zusammenführung, Strukturierung, Systematisierung und Auswertung der Erfahrungen und Erkenntnisse aller Phasen*
- *Aufarbeitung und Finalisierung des Prototyping*
- *Konzept und Entwicklung Leitfadens*

Phase 5 – Publikation; 6 Monate

- *Erstellung wissenschaftlicher Abschlussberichte*
- *Dissemination von Forschungsergebnissen durch Partizipation in wissenschaftlichen Konferenzen und Zeitschriftenartikel*
- *Veröffentlichung Leitfadens sowie Prototypen und Quellcodes*

Budget

Eigene Personalaufwände

Label	Beschreibung	Aufwände	Kosten
Salär 1. Jahr	nach Richtlinie doc.ch	-	47'040.-
Salär 1. Jahr	"	-	48'540.-
Salär 1. Jahr	"	-	50'040.-
			<u>145'620.-</u>

Externe Personalaufwände

Label	Beschreibung	Aufwände	Kosten
zus. Software Entwickler:in	40% Stellenprozent auf 8 Monate	75.-/h	5'040.-
Honorar Interviews	bei 8 Personen/1h	8h/75.-	600.-
Honorar Workshopteilnahme	4 Workshops/5 Personen	75.-/Person	1'500.-
Honorar Fokusgruppe	4 Fokusgruppen/5 Personen	50.-/Person	1'000.-
Korrektorate	Dissertation, Leitfaden	-	1'500.-
			<u>9'640.-</u>

Materialkosten, Weitere Aufwände

Label	Beschreibung	Aufwände	Kosten
Literaturrecherche	Fachliteratur, Zugriff Forschungsdokumente	-	250.-
Verbrauchs-Material	Workshopmaterial, Büro	-	500.-
Reisespesen	Konferenzen, Forschungs-Reisen	4 Reisen/500.-	2'000.-
Entwickler-Umgebung	Software Lizenzen, Hosting für Applikationen und Code	250.-/Jahr	750.-
Hardware	Arbeitsgeräte, Elektronische Bauteile, Rapid Prototyping, Anleihe von Geräten	-	8'000.-
Infrastruktur	Raum- und Mobiliarmieten für Workshops und Entwicklungsarbeiten	-	1'500.-
Buffer	Unvorgesehene Ausgaben	-	1'000.-
			<u>14'000.-</u>

Bibliographie

Bajorek, J. P. (2019, Mai 10). Voice Recognition Still Has Significant Race and Gender Biases. *Harvard Business Review*. <https://hbr.org/2019/05/voice-recognition-still-has-significant-race-and-gender-biases>

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic Analysis. In P. Liamputtong (Hrsg.), *Handbook of Research Methods in Health Social Sciences* (S. 843–860). Springer Singapore. https://doi.org/10.1007/978-981-10-5251-4_103

Danielescu, A. (2020). Eschewing Gender Stereotypes in Voice Assistants to Promote Inclusion. *Proceedings of the 2nd Conference on Conversational User Interfaces*, 1–3. <https://doi.org/10.1145/3405755.3406151>

Easwara Moorthy, A., & Vu, K.-P. (2014). Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space. *International Journal of Human-Computer Interaction*, 31, 307–335. <https://doi.org/10.1080/10447318.2014.986642>

Focus Groups. (2017). In M. Allen, *The SAGE Encyclopedia of Communication Research Methods*. SAGE Publications, Inc. <https://doi.org/10.4135/9781483381411.n208>

Hector, T. M., & Hrcnal, C. (2020). Intelligente Persönliche Assistenten im häuslichen Umfeld: Erkenntnisse aus einer linguistischen Pilotstudie zur Erhebung audiovisueller Interaktionsdaten. <https://doi.org/10.25819/UBSI/1013>

Humphry, J., & Chesher, C. (2020). Preparing for smart voice assistants: Cultural histories and media innovations. *New Media & Society*, 146144482092367. <https://doi.org/10.1177/1461444820923679>

Masina, F., Orso, V., Pluchino, P., Dainese, G., Volpato, S., Nelini, C., Mapelli, D., Spagnolli, A., & Gamberini, L. (2020). Investigating the Accessibility of Voice Assistants With Impaired Users: Mixed Methods Study. *Journal of Medical Internet Research*, 22(9), e18431. <https://doi.org/10.2196/18431>

Matthews, B., & Wensveen, S. (2015). Prototypes and prototyping in design research (S. 262–276). <https://doi.org/10.4324/9781315758466-25>

Ozkaynak, M., Sircar, C. M., Frye, O., & Valdez, R. S. (2021). A Systematic Review of Design Workshops for Health Information Technologies. *Informatics*, 8(2), 34. <https://doi.org/10.3390/informat-ics8020034>

Pal, D., Arpnikanondt, C., Razzaque, M. A., & Funilkul, S. (2020). To Trust or Not-Trust: Privacy Issues With Voice Assistants. *IT Professional*, 22, 46–53. <https://doi.org/10.1109/MITP.2019.2958914>

Pradhan, A., Mehta, K., & Findlater, L. (2018). «Accessibility Came by Accident»: Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (S. 1–13). Association for Computing Machinery. <https://doi.org/10.1145/3173574.3174033>

Tripp, S. D., & Bichelmeyer, B. (1990). Rapid prototyping: An alternative instructional design strategy. *Educational Technology Research and Development*, 38(1), 31–44. <https://doi.org/10.1007/BF02298246>

Wiltse, H. (Hrsg.). (2020). *Relating to Things: Design, Technology and the Artificial*. Bloomsbury Visual Arts. <https://doi.org/10.5040/9781350124288>