

Aneignung und Gebrauchstauglichkeit von Sprachassistenten im Anwendungsbereich Küche

Dominik Pins¹, Alexander Boden¹, Gunnar Stevens² und Britta Hofmann¹

¹ Fraunhofer-Institut für Angewandte Informationstechnik FIT, Sankt Augustin, Germany
{vorname.nachname}@fit.fraunhofer.de

² Universität Siegen, Germany
gunnar.stevens@uni-siegen.de

Abstract. Die Küche ist neben dem Wohnzimmer ein Lebensmittelpunkt in vielen Haushalten. Hier wird nicht nur gekocht, sondern auch gelebt und gewirtschaftet. Sie bietet daher viele Potenziale für Smart-Home-Technologien, insbesondere der Verwendung von Sprachassistenten, die eine Reihe von Aufgaben im Bereich Küche unterstützen können. Ziel unserer Forschung ist es, unter Berücksichtigung der Aspekte nutzerorientierte Gestaltung, Anthropomorphismus der Interaktion und Zuverlässigkeit der bereitgestellten Information zu ermitteln, wie eine erfolgreiche Interaktion zwischen Menschen und Sprachassistenten gelingen kann. Die Herausforderung dabei ist unter anderem, dass Anwendungen ineinandergreifen und eine komplexe und vor allem flexible Form der Interaktion mittels Sprache erfordern. Um das Forschungsfeld besser zu verstehen, haben wir in einem ersten Schritt zehn Haushalte über ihre Nutzung und Zufriedenheit mit Sprachassistenten befragt. Erste Ergebnisse dieser explorativen Studie stellen wir im Rahmen dieses Beitrags vor.

Keywords: Sprachassistenten, Ernährung, Verbraucherinformatik, Mensch-Computer-Interaktion, Usability & UX Design,

1 Einleitung

Die Küche ist nicht bloß ein Ort, an dem Mahlzeiten zubereitet werden. Hier fallen viele weitere Aufgaben und Tätigkeiten zusammen, die meistens unabhängig voneinander betrachtet werden, doch im Rahmen des Komplexes „Wohnen und Wirtschaften“ zusammenhängen (Johannes-Hornschuh 2010). Neben dem Kochen für sich allein oder gemeinsam mit anderen ist die Küche auch ein wichtiger Ort für soziales Miteinander, sei es auf der Party oder beim Kochen mit FreundInnen. Die Küche ist demnach kein rein rationaler Hauswirtschaftsraum, sondern dient auch als Erlebnisraum. Sie ist aber auch ein Wissens- und Informationsraum, in dem Medien wie (Familien-)Terminplaner, Kochbücher, Werbeprospekte, Einkaufslisten, aber auch Postkarten, Bilder und Kunstwerke der Kinder oft eine prominente Rolle spielen. Wohntrends zeigen zudem immer häufiger offene Wohn-Kochbereiche, sodass die Küche immer mehr in den Mittelpunkt des häuslichen Lebens rückt (Johannes-Hornschuh 2010).

Zusätzlich zu analogen Medien halten derzeit vermehrt digitale Werkzeuge Einzug in der Küche. Neben dem Smartphone spielen dabei Sprachassistenten eine zunehmende Rolle, die nicht nur zum Musikhören oder als Smart-Home-Steuerung dienen, sondern auch als Organizer zum Kochen und Einkaufen (Hyland et al. 2018). So bieten sie etwa Funktionen für die Verwaltung von Einkaufslisten, können Rezepte vorlesen, oder einfache Aufgaben wie das Anlegen von Erinnerungen und zeitgesteuerten Alarmen übernehmen. Sprachassistenten sind dabei besonders durch die Möglichkeit der handlosen Sprachsteuerung für die Nutzung in der Küche attraktiv. Für die Verbraucherinformatik stellen sie ein interessantes Forschungsthema im Bereich digital vermittelter Formen des Konsums dar, was wir im Rahmen unserer zukünftigen Forschung näher betrachten möchten.

2 Forschungsthema

Die Küche spielt eine zentrale Rolle im Praxiskomplex rund um das Thema Ernährung (Warde 2016), ist sie doch der Ort, an dem Lebensmittel gelagert, verarbeitet und (teilweise) entsorgt werden. Sprachassistenten können dabei nicht nur als Unterstützer beim Kochen fungieren. Durch ihre Rolle als Alltagsassistenten transzendieren sie vielmehr in gewisser Weise die Grenzen zwischen der Küche und weiteren zentralen Orten wie dem Wohn- bzw. Esszimmer bzw. dem (Super-)Markt als Stätten für Praktiken rund um den Verzehr und den Erwerb von Lebensmitteln. So dienen Sprachassistenten potentiell als Ratgeber zur Beantwortung von Fragen zu Gesundheit und Ernährung, oder als Einkaufsberater mit aktuellen Angeboten und Preisen von Produkten. Sie stellen zudem eine direkte Schnittstelle zu Rezeptdatenbanken und dem Online-Handel dar.

In der Praxis stellen solche Aufgaben jedoch aufgrund ihrer Komplexität eine erhebliche Herausforderung in der Umsetzung dar, weil die zugrundeliegende Technik der Spracherkennung noch viele Einschränkungen aufweist. Mittels Cloud-Anbindung können die Geräte zwar mittlerweile sprachliche Befehle recht zuverlässig erkennen,

die Illusion eines intelligenten Gesprächspartners zerbricht jedoch schnell an den eingeschränkten Möglichkeiten „echten“ sprachlichen Verstehens. Zudem sind die verfügbaren Skills für Sprachassistenten in ihrer Funktionalität derzeit noch sehr begrenzt, oder leiden an Problemen der Bedienfreundlichkeit, insbesondere durch die mangelhafte Selbstbeschreibungsfähigkeit reiner Sprach-Interfaces.

Da die Technik bezüglich Implementierung und Verbreitung derzeit noch in den Kinderschuhen steckt, zielt unsere Forschung darauf, anhand des Anwendungsbereichs Küche herauszustellen, wie eine erfolgreiche und zufriedenstellende Mensch-Sprachassistent-Interaktion unter den Aspekten a) nutzerorientierte Gestaltung, b) Anthropomorphismus der Interaktion und c) Zuverlässigkeit der bereitgestellten Informationen und Wissen gelingen kann.

Unter den ersten Aspekt fällt vor allem das Prüfen der Gebrauchstauglichkeit von Sprachinterfaces im Bereich der Küche in Hinblick auf ihre effiziente Umsetzung vor dem Hintergrund der Dialogprinzipien (gemäß DIN EN ISO 9241-110, DAkkS 2010). Dies erfordert vor allem eine umfassende empirische Untersuchung des Nutzungskontexts sowie die Identifikation von Aufgaben, Erfordernissen und Nutzungsanforderungen, die der Sprachassistent in der Küche erfüllen kann oder soll. Dabei sollen Voraussetzungen für eine zufriedenstellende und zielführende Konversation insbesondere in komplexeren Interaktionen ermittelt werden, die beispielsweise Situationen aufdecken, die vom Sprachassistenten Proaktivität verlangen.

Der zweite Aspekt umfasst vor allem die Erwartungen der NutzerInnen an das (Sozial-)Verhalten des Sprachassistenten sowie dessen Rolle im Bereich der Wohnung bzw. der Mobilität (Seymor 2018, Moore R.J 2017). Da es sich bei Sprachassistenten um anthropomorphe Interfaces handelt, ist zu fragen, wie sich der Charakter des Gerätes oder des ausgeführten Skills – beispielsweise Stimme, Umgangsformen und Proaktivität der Situation – auf die Nutzungserfahrung der AnwenderInnen auswirkt, und wie dies beispielsweise mit dem Aspekt Vertrauensbildung in die dargestellten Informationen zusammenhängt.

Der dritte Aspekt, knüpft schließlich an die Vertrauensbildung an und nimmt die Zuverlässigkeit der Informationen und des Kanals, über den diese Informationen bereitgestellt werden, in den Blick. Dabei soll untersucht werden, in welchen Kontexten in der Küche ein Informationsbedarf besteht und in welchem Umfang die Informationen durch Sprachinterfaces bereitgestellt werden können. Dabei sind auch Medienkombinationen mit visuellen Kanälen interessant und es stellt sich die Frage, wie eine Interaktion mit einem Sprachassistenten gestaltet werden kann, in der komplexere Informationen kommuniziert werden müssen, die eine sehr kontextbezogene und situationsgerechte Kommunikation erfordern (Képuska et al. 2018) – insbesondere wenn mehrere Aktivitäten parallel thematisiert werden sollen.

3 Erste Ergebnisse

Im Rahmen einer explorativen Studie mit zehn NutzerInnen von Sprachassistenten (Amazon Echo und Google Home Assistant) haben wir alltägliche Nutzungssituationen

identifiziert, die sich durch einen längeren Gebrauch des Sprachassistenten abzeichnen. Dazu wurden die ProbandInnen im Rahmen von ca. zweistündigen Interviews zu ihren Nutzungsgewohnheiten befragt und Einblicke in die Interaktionshistorie (d. h. den Verlauf der vom Sprachassistenten ausgeführten Aktionen) vorgenommen.

Die Interviews zeigen, dass neben dem Wohnzimmer die Küche der häufigste Nutzungsbereich ist, gefolgt von Badezimmer und Schlafzimmer. Unsere Forschung bestätigt vorhergehende Studien insoweit sich häufig eine Ernüchterung über die Fähigkeiten des Sprachassistenten breitmacht, und die hohen Erwartungen an die Fähigkeiten und das Verstehen zumeist enttäuscht wurden (Moore R.K. 2017, Bentley 2018, Cho 2018). Dennoch sind die befragten NutzerInnen mit der Interaktion insgesamt zufrieden, sobald sie die Sprachsteuerung und die möglichen Befehle und Ausdrucksweisen verstanden haben. Demnach müssen die Fragen und Anweisungen möglichst klar, präzise und kurz ausgesprochen werden. Sie berichteten zudem von einem immer besseren Verstehen der Befehle durch den Sprachassistenten, was sie vor allem auf regelmäßige Updates der Hersteller zurückführen.

Probleme bereiten aus den ersten Erkenntnissen der Studie vor allem die Steuerung von Skills, da diese umständlicher durch notwendige Nennung des Namens angesteuert werden müssen. Auch die fehlende Freiheit in der Wortverwendung (z. B. Variation zwischen Licht, Lampe und Leuchte) wird teilweise bemängelt. Der größte Benefit ist nach Ansicht unserer NutzerInnen die Steuerung und Informationsgewinnung durch reine Sprache, sodass bis auf das Sprechen keine physischen Aktionen notwendig sind (z. B. der Gebrauch von Händen), was besonders beim Kochen vorteilhaft sein kann.

Die Nutzung eines Sprachassistenten im Bereich Küche konzentriert sich bei den befragten NutzerInnen fast ausschließlich auf das Setzen eines Timers und das Musikhören im Hintergrund während des Kochens / Essens. Teilweise nutzen Haushalte über den Sprachassistenten eine gemeinsame Einkaufsliste. Größte Barriere für den Einsatz beim Kochen ist die fehlende Visualität, da der Nutzer bzw. die Nutzerin gerne zwischendurch mal auf das Rezept gucken möchte oder zumindest bildhaft sehen möchte, wie das Resultat aussieht oder aussehen soll. Der momentan in Deutschland einzig verfügbare über den Sprachassistenten steuerbare Bildschirm ist der Amazon Echo Show, der dem Großteil der befragten NutzerInnen jedoch zu teuer ist beziehungsweise in seinem Nutzungspotenzial gegenüber einem Laptop, Tablet oder Smartphone noch zu uninteressant erscheint.

4 Diskussion

Unsere Ergebnisse greifen bereits die oben genannten Bewertungsaspekte auf und geben Anlass, weitere Forschungen in diese Richtung zu betreiben. Denn auch wenn von einer steten Verbesserung des Verstehens von Befehlen berichtet wird, ist die Interaktion noch sehr beschränkt und auf kurze, geschlossen Sequenzen reduziert. Weiterhin sind komplexere Aufgaben, wie das Kochen nach Rezept, ineffizient ausführbar, da insbesondere die Steuerbarkeit und Flexibilität in der Konversation fehlt oder unzureichend ist. Hinzu kommt aber auch, dass NutzerInnen nicht wissen, welche Skills ein Sprachassistent unterstützt, die in der Küche bei der Aufgabenerledigung

unterstützen können. Auch die anthropomorphe Wahrnehmung der Interaktion beschränkt sich aufgrund der häufigen Frage-Antwort- bzw. Befehl-Ausführung-Interaktion auf ein Minimum. Dennoch berichteten die NutzerInnen, relativ "normal" mit dem Sprachassistenten zu sprechen, und vereinzelt achten sie auch darauf Umgangsformen wie „Bitte“ und „Danke“ zu verwenden. Die NutzerInnen bemängelten im Hinblick auf den dritten Aspekt, dass Alexa bis vor einiger Zeit häufiger auf Fragen keine Antwort wusste und deshalb die Probanden in dem Fall auf andere Medien zurückgriffen. Auch vereinzelt Versuche, Rezepte zu kochen, scheiterten aufgrund der fehlenden Flexibilität in der Anforderung für die in der Situation notwendigen benötigten Informationen.

Komplexere Aufgaben, aber auch ineinandergreifende Aktionen erfordern schlussfolgernd eine Mensch-Sprachassistent-Interaktion, die erstens eine effektive, effiziente und zufriedenstellende Aufgabenerledigung ermöglicht, dem Nutzer zweitens ein Gefühl von Freiheit und Vertrauen in der Interaktion innerhalb seines Kontextes gibt und drittens Informationen so bereitstellt, dass der Nutzer bzw. die Nutzerin sie versteht und darauf basierend Nachfragen und anknüpfende Aktionen möglich sind. So bedingt die Vielschichtigkeit im Praxiskomplex um die Ernährung einen flexiblen und individuellen Informationsbedarf und stellt damit auch die Sprachassistenten vor die Herausforderung, nutzerspezifische Informationen zu filtern, bereitzustellen und darauf eingehen zu können, wenn beispielsweise Alternativen gefordert werden. Der Bereich Küche verlangt mithin auch die Möglichkeit des Multi-Taskings, wenn etwa das „Obst der Woche“ erfragt wird, man nach einem passenden Rezept suchen möchte, das Obst und einige Zutaten auf die Einkaufsliste setzen und dann noch seine Nährwerte erfragen möchte, ohne jedes Mal eine neue Anwendung starten zu müssen oder wenn eine (Kalender-)Erinnerung dazwischen spricht, von vorne beginnen zu müssen.

5 Ausblick

Um das Potenzial und die Anforderungen an einen Sprachassistenten im Anwendungsbereich Küche besser zu erfassen, planen wir weiter, mithilfe von Beobachtungen und Kontextinterviews Nutzungspraktiken im Kontext „Küche“ genauer zu untersuchen. Hierzu soll auch auf Basis tiefergehender Analysen etwa von Interaktionsverläufen untersucht werden, wie der Sprachassistent in der Küche insgesamt genutzt wird, welche Interaktionen stattfinden, wo sie gescheitert sind und ob sich bestimmte Interaktionsabläufe oder -strukturen identifizieren lassen, aus denen sich proaktive Handlungen seitens des Sprachassistenten ableiten lassen. Zudem soll die Interaktion genauer betrachtet werden: Wie wirken sich Aspekte wie Stimme, Umgangsformen und Proaktivität auf die Wahrnehmung des und Zufriedenheit mit dem Sprachassistenten in den Kontexten aus. Außerdem soll untersucht werden, ob oder wie der Sprachassistent und die damit verknüpften Anwendungen in außerhäuslichen Situationen eingesetzt werden, und inwieweit sich Interaktionen an anderen, mit der Ernährung zusammenhängenden Interaktionsorten wie dem Supermarkt unterscheiden. Diese Fragen wollen wir im Rahmen unserer weiteren Forschung bearbeiten.

References

1. Bentley, F., LuVogt, C., Silverman, M., Wirasinghe, R., White, B., Lottridge, D.: Understanding the Long-Term Use of Smart Speaker Assistants. *Proc ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3 Article 91 (2018)
2. Cho, Janghee: Mental Models and Home Virtual Assistants. *CHI'18 Extended Abstracts*, Montreal, QC, Canada (2018)
3. Deutsche Akkreditierungsstelle (DAkKS): Leitfaden Usability. https://www.dakks.de/sites/default/files/71_sd_2_007_leitfaden_usability_1.3_0.pdf, Version 1.3 (2010).
4. Hyland, L., Crabtree, A., Fischer, J. et al.: What do you want for dinner? – need anticipation and the design of proactive technologies for the home. In: *Computer Supported Cooperative Work*. Volume 27, Issue 3–6, pp 917–946 (2018)
5. Johannes-Hornschuh, M.: Die moderne Küche als sozialer Raum. *Triga, Gründau-Rothenbergen* (2010)
6. Kěpuska, V., Bohouta, G.: Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). *Conference: IEEE CCWC 2018, The 8th IEEE Annual Computing and Communication Workshop and Conference At: University of Nevada, Las Vegas, USA* (2018)
7. Moore, R. J., Arar, R., Ren, G.J., Szymanski, M.H.: Conversational UX Design. *CHI'17 Extended Abstracts*, Denver, CO, USA (2017)
8. Moore, R. K: Is Spoken Language All-or-Nothing? Implications for Future Speech-Based Human-Machine Interaction. In: Jokinen K., Wilcock G. (eds) *Dialogues with Social Robots*. *Lecture Notes in Electrical Engineering*, vol 427. Springer, Singapore (2017)
9. Muhle, F.: "Are you human?" Plädoyer für eine kommunikationstheoretische Fundierung interpretativer Forschung an den Grenzen des Sozialen. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 17(1) (2016)
10. Seymor, W.: How loyal is your Alexa? Imagining a Respectful Smart Assistant. *CHI'18 Extended Abstracts*, Montreal, QC, Canada (2018)
11. Warde, A: *The practice of eating*. Polity press, London (2016).