

Andre Alpar, Claudius Konopka

Beyond Touch – intelligente Assistenten: Teil III: Wie der perfekte Assistent der Zukunft aussieht

In den vorangegangenen zwei Artikeln unserer dreiteiligen Serie „Beyond Touch: Intelligente Assistenten“ haben wir uns die Entwicklung und den Status quo der neuen Plattformen um den Bereich Voice angeschaut. Während wir bei Adaption und Nutzung intelligenter Assistenten längst über die Anfangsphase hinaus sind, fragen wir uns, was die Zukunft bringt. Die fortlaufende Verbesserung der Technologie integriert Interaktionen per Sprache immer weiter in unseren Alltag. Fakt ist: Was heute noch relativ neu für viele Nutzer ist, wird der Standard von morgen sein. Die Ansprüche sind hoch und die Entwicklung schnell. Doch wie können die Sprachassistenten der Zukunft aussehen? Wie interagieren Sie mit uns und was erwarten wir von ihnen?

Von eindimensional zu interaktiv

Die Interaktion mit intelligenten Assistenten wie Alexa, Google Assistant oder auch Siri geht heute noch recht kompliziert vonstatten. Von der nahtlosen Integration in die häusliche Umgebung ist die Entwicklung derzeit noch ein paar Schritte entfernt. So müssen wir vor der Inbetriebnahme erst zahlreiche Accounts verbinden und die Spracheingabe ist sicher noch verbesserungswürdig. Viele Brüche in der Nutzung und Integration in unseren Haushalt und unser Leben erschweren derzeit eine 100 % verfügbare Nutzung der Assistenzsysteme. Die Integration der Voice-Technologie in unseren Alltag wird aber zunehmend wichtiger. Auch außerhalb unseres Heims, auf Reisen im PKW oder bei Aufenthalt in Hotels soll die Verfügbarkeit ohne Frustramente oder Wiederholungen gegeben sein. Als Schnittstelle zwischen den Systemen und digitalen Hilfsmitteln nimmt die Bedeutung der Sprachassistenten immer weiter zu.

Dabei sind die Bedürfnisse der Menschen durchaus verschieden. Nicht für jeden ist die Nutzung einer sprachgesteuerten Anwendungspalette denkbar oder komfortabel. Denken wir an Menschen mit Beeinträchtigungen oder Behinderungen, für die ein Zugang nicht immer möglich

DER AUTOR



Andre Alpar ist seit über 20 Jahren unternehmerisch im Bereich Online-Marketing tätig. Heutzutage ist er als Investor, Fachautor, Keynote Speaker und Podcaster (#askOMR) aktiv.

DER AUTOR



Claudius Konopka ist Gründer und Geschäftsführer der Beyto GmbH, eine auf Voice und Conversational AI spezialisierten Beratungsagentur. Mit Beyto berät er Unternehmen zum Thema Voice und virtuelle Assistenten.



ist. Oder an Situationen, in denen eine Texteingabe sinnvoller wäre oder der jeweiligen App hilft, die gestellte Aufgabe bestmöglich zu erfüllen. Wir sehen also: Was für den einen eine neue Zukunft ist, kann für den anderen ein Hindernis bedeuten.

Für die Zukunft der intelligenten Assistenten bedeutet dies die Entwicklung angepasster und optimierter Modelle. Anstatt eindimensional und isoliert zu funktionieren, werden der Dialog und die persönliche Beziehung zwischen dem Nutzer und dem Sprachassistenten wichtiger. Das Gefühl, dass der Sprachassistent jederzeit verfügbar ist und als eine Stimme für viele Anwendungen fungiert, erleichtert es dem Nutzer, die Maschine hinter der Voice-Technologie quasi zu vergessen. Die verschiedenen Nutzungsräume müssen nahtlos ineinandergreifen und es möglich machen, dort anzusetzen, wo der Nutzer aufgehört hat. Die Interaktion des Assistenten mit dem Nutzer beispielsweise durch das Vorschlagen weiterführender Themen bringt die Technologie weg von einer reinen Reaktion und hin zu einer aktiven Hilfestellung.

Was bedeutet dies für die neue Generation intelligenter Assistenten?

Aus den oben genannten Bedingungen lassen sich folgende Anforderungen an die zukünftigen Sprachassistenten ableiten:

1. Omnipräsenz
2. Proaktivität
3. Menschlichkeit

Im Film „Her“ aus dem Jahre 2013 verliebt sich der Hauptdarsteller in sein lernfähiges Computer-Betriebssystem Samantha. Die beiden führen tiefe Gespräche, die künstliche Intelligenz reagiert auf seine Aussagen und kombiniert verschiedene Datenquellen, um

Rückschlüsse auf seinen Gemütszustand und seine Wünsche zu ziehen. Sie führt ihn zu ihm unbekanntem Orten, die sie aufgrund seiner Vorlieben ausgewählt hat, um ihn zu überraschen, und sendet proaktiv von seinem E-Mail-Account aus von ihm geschriebene Briefe an einen Buchverlag, der diese daraufhin veröffentlicht. Diese Vision ist gar nicht so weit entfernt.

All diese Dinge sind mit den derzeitigen eher reaktiven, eindimensionalen Systemen nicht möglich. Es braucht ein interaktives, lernfähiges und hochvernetztes System – die digitalen Assistenten der Zukunft.

1. Omnipräsenz

Alle voicebasierten Assistenten, die wir heutzutage kennen und nutzen, werden mit Lautsprechern in Verbindung gebracht. Amazons Alexa und Google Home kommen in einem stylischen Zylinder daher, Microsofts Cortana und Apples Siri sind an das jeweilige Gerät gebunden. Zukünftige Generationen werden sich in vielen Devices und Geräten verbergen, die wir nicht mehr als externe Peripherie wahrnehmen werden. Eingebaut in Kühlschränke, unser Auto, unsere Armbanduhr und sogar in Brillen wird eine übergangslose Nutzung gang und gäbe sein.

Als neuer Vermittler wird der digitale Assistent der Zukunft außerdem eine Multimodalität sicherstellen. Er wird nicht nur plattformbasierte Erweiterungen nutzen, die auf den jeweiligen Hersteller wie Amazon oder Google zugeschnitten sind, sondern Angebote und Features von verschiedenen Plattformen zusammenführen.

Im Bereich Multimedia werden so Medien aus dem gesamten Unterhaltungssystem von TV, Apps, Videospiele, Video-on-Demand und VR-Angebote zusammengeführt, um sicherzustellen, dass wir das finden, was wir suchen.

2. Engagiert und proaktiv

Sprachgesteuerte Geräte entwickeln sich immer weiter und werden intelligenter. Dabei müssen Sie eine immer höhere Anzahl an miteinander verbundenen Aufgaben übernehmen.

Laut der Beyto-Smart-Speaker-Studie 2019 sind derzeit die häufigsten Themen bei der Nutzung intelligenter Assistenten Smart Home, Musik und Audio, Wetter, Nachrichten, Produktivität oder auch Dienstprogramme (Studie: <http://einfach.st/beyto>). Sprachassistenten übernehmen in Anfängen schon heute die Aufgabe, verschiedene Informationen zu kombinieren. So empfehlen sie beispielsweise dem Nutzer den idealen Weg unter Beachtung des Wetters und des Verkehrs zu einem bestimmten Ort. Dabei müssen die Parameter jedoch vorher durch den Nutzer selbst definiert werden, um ein sinnvolles Ergebnis zu erhalten. Doch das wird in Zukunft nicht mehr ausreichen.

Die Fähigkeit, hochkomplexe Fragen zu verarbeiten, wird die Sprachassistenten der Zukunft auszeichnen. Während heute noch eine definierte Aktion erstellt werden muss, werden die Systeme von morgen in der Lage sein, vielschichtige Fragen wie: „*Wohin kann ich morgen gehen und wie komme ich am besten dorthin?*“, frei zu beantworten.

Auch die Fähigkeit, neue Features und Funktionen zu entdecken, ist in Anfängen bereits in einigen Assistenten implementiert, hält für die Zukunft aber noch viel Potenzial bereit. So wird beispielsweise bei der Frage nach einem Beautytipp automatisch der „My Beauty Chat“ von L'Oréal geöffnet und interagiert mit dem Nutzer.

Doch gehen wir noch einen Schritt weiter. Die Sprachassistenten von morgen werden in der Lage sein, eine Abfolge von Handlungen in unserem Namen vorzunehmen. So können sie bei Eingang von Rechnungen automatisch die eigenen Bank- und Kartenkonten



prüfen und proaktiv und termingerecht den Betrag überweisen. Darüber hinaus können sie Beträge zwischen den Konten verschieben, um einen größtmöglichen Nutzen aus beispielsweise Zinsschwankungen zu ziehen. Im Anlagebereich bieten sich darüber hinaus noch mehr Möglichkeiten. Denken wir an eine erweiterte Anlageberatung, wie sie digitale Anlageassistenten bereits heute in Teilen vornehmen. Die Comdirect-Bank hat als erste deutsche Bank einen eigenen Alexa-Skill herausgebracht, der es Depotinhabern ermöglicht, Kurse und Entwicklungen im Auge zu behalten und zu reagieren. In Zukunft können uns unsere Assistenten an unsere Lebenssituation, Risikofaktoren und das Alter angepasste Produkte empfehlen und proaktiv das Depot verwalten.

3. Menschlichkeit

Die Nutzung eines Sprachassistenten muss sich für den Menschen natürlich anfühlen. Vorbei sind die Zeiten steifer Roboterstimmen, die blechern die Sätze herunterrattern. Auffällig dabei ist, dass die meisten Stimmen der Sprachassistenten weiblich sind. Obwohl der

Nutzer auch eine männliche Stimme auswählen kann, sind die Standardstimmen weiblicher Natur. Das liegt daran, dass die weibliche Stimme laut Studien als angenehmer und entspannender wahrgenommen wird (<http://einfach.st/mitchell>). Je nach Land werden die Stimmen den gesellschaftlichen und charakterlichen Eigenheiten angepasst. Diese Befähigung der Technologie, den Nutzer emotional zu erreichen, wird die zukünftigen Assistenten ausmachen. Unbelebte Objekte werden zu Helfern, von denen menschenähnliche Qualitäten und eine Zwei-Wege-Kommunikation erwartet werden. Erfüllen die Sprachassistenten nicht die Anforderungen an Zuhören und Verstehen, werden die Chancen, die eine emotionale Ansprache der Nutzer bietet, vertan.

Für eine personalisierte Erfahrung wird sich das Lernverhalten des eigenen Assistenten aus der Vielzahl an Informationen zusammensetzen, die er sowohl aus unserer Interaktion bezieht als auch aus der Menge der Gesamtdaten aller anderen Nutzer. Das Wissen über die Menschen in ihrer Gesamtheit wird mit dem, was er über einen bestimmten

Benutzer weiß, verknüpft. Aber auch das maschinelle Lernen aller anderen Assistenten wird herangezogen, um eine maßgeschneiderte persönliche Erfahrung für den einzelnen Nutzer zu gewährleisten.

Wir sehen: Trotz des Anspruches an eine menschenähnliche Interaktion wird die Ähnlichkeit mit dem Menschen wohl bei den zukünftigen Sprachassistenten weniger wichtig sein als die Eignung, den Menschen zu befähigen. Der perfekte Assistent soll die Menschen unterstützen, sie weiterführen und ihnen helfen, sich selbst zu verbessern.

Use Cases und Ausblick

Die Nutzung der intelligenten Assistenten steckt erst in den Kinderschuhen, aber die Visionen und das Potenzial sind gigantisch. Dies bestätigt sich, wenn wir uns einige Entwicklungen oder auch Patente von Google, Amazon und Co. betrachten.

» **Amazon** hat ein Patent zur Bestimmung des Gesundheitszustandes des Nutzers beantragt und erhalten. Hierbei sollen Krankheiten und Gemütszustände anhand der Stimmuster und des Bestellverlaufs im Internet



bestimmt werden. Der Assistent schlägt daraufhin passende Medikamente vor, die er auch gleich online bestellt.

- » **Microsoft** arbeitet an einem neuartigen Flüstermodus für das haus-eigene Cortana. Anders als der bereits bestehende Modus bei Alexa kann das System Wörter aus dem beim Sprechen erzeugten Luftstrom extrapolieren.
- » **Google** arbeitet an einer Voice-Technologie, die das gesamte Heim eines Nutzers mit Kameras scannt und in Kombination mit dem Browserverlauf von Google Chrome auf bestimmte Vorlieben zugeschnittene Produkte und Aktivitäten vorschlägt. Diese sind dann auch direkt über Spracheingabe bestellbar.
- » Ebenfalls von **Google** kommt das Patent für vernetzte Smart-Home-Geräte, die automatisch den Modegeschmack, das geschätzte Einkommen sowie Geschlecht und Alter der Nutzer erkennen. Darauf aufbauend werden Vorschläge für beispielsweise Einkauf und Entertainment gemacht. Dieses System findet heute schon bei der

Datenerhebung im Internet und bei Entertainmentanbietern wie Netflix Verwendung, kann aber in Zukunft auf den gesamten heimischen Bereich ausgeweitet werden.

- » **Apple** ist eine Partnerschaft mit dem Versicherer United Healthcare eingegangen, der kostenlose Apple-Uhren verteilt hat, um die Mitarbeiter von Unternehmen anzuhalten, bestimmte Fitnessziele zu erreichen. Die Daten werden hierbei von dem Versicherungsunternehmen ausgewertet. Diejenigen, die sich sportlich betätigen, können demnach Vorteile erhalten, oder diejenigen abgestraft werden, die dies nicht tun. Ein Weiterdenken dieses Anwendungsfalles wäre die Erhebung von Ernährungsdaten aus Smart-Home-Geräten aus der Küche, um Essgewohnheiten zu ermitteln und Tarife entsprechend anzupassen.

Doch die Anwendung heimischer Sprachassistenten in Verbindung mit Smart-Home-Geräten geht noch einen Schritt weiter. Ein weiteres Anwendungsszenario zeichnet dieses Patent: Daten aus installierten Sensoren und

Kameras sollen eingesetzt werden, um das Verhalten der Kinder in Abwesenheit der Eltern zu überwachen. Eltern können ein Gerät programmieren, das erfasst, ob eine „unerwünschte Sprache“ benutzt wird, was im Internet oder TV geschaut wird und ob bestimmte Tabuzonen des Hauses benutzt werden. Dabei könnten LED-Lichter zur Warnung der Kinder eingesetzt werden, wenn diese beispielsweise den Alkoholschrank öffnen.

Fazit

Was muss der perfekte intelligente Assistent der Zukunft bieten? Das hängt ganz von der Art der Anwendung ab. Allem gemein ist jedoch, dass der ideale Sprachassistent intelligent und lernfähig sein muss. Er muss Daten in Echtzeit verarbeiten und kombinieren und den Nutzer in seinen Eigenheiten dort abholen, wo dieser sich befindet. Fakt ist, dass es nur noch eine Frage der Zeit ist, bis die Entwicklung und Vernetzung in jedem Bereich unseres Lebens präsent ist. Bestimmende Faktoren sind nur das Tempo der technischen Entwicklung und die Akzeptanz der Nutzer. ¶