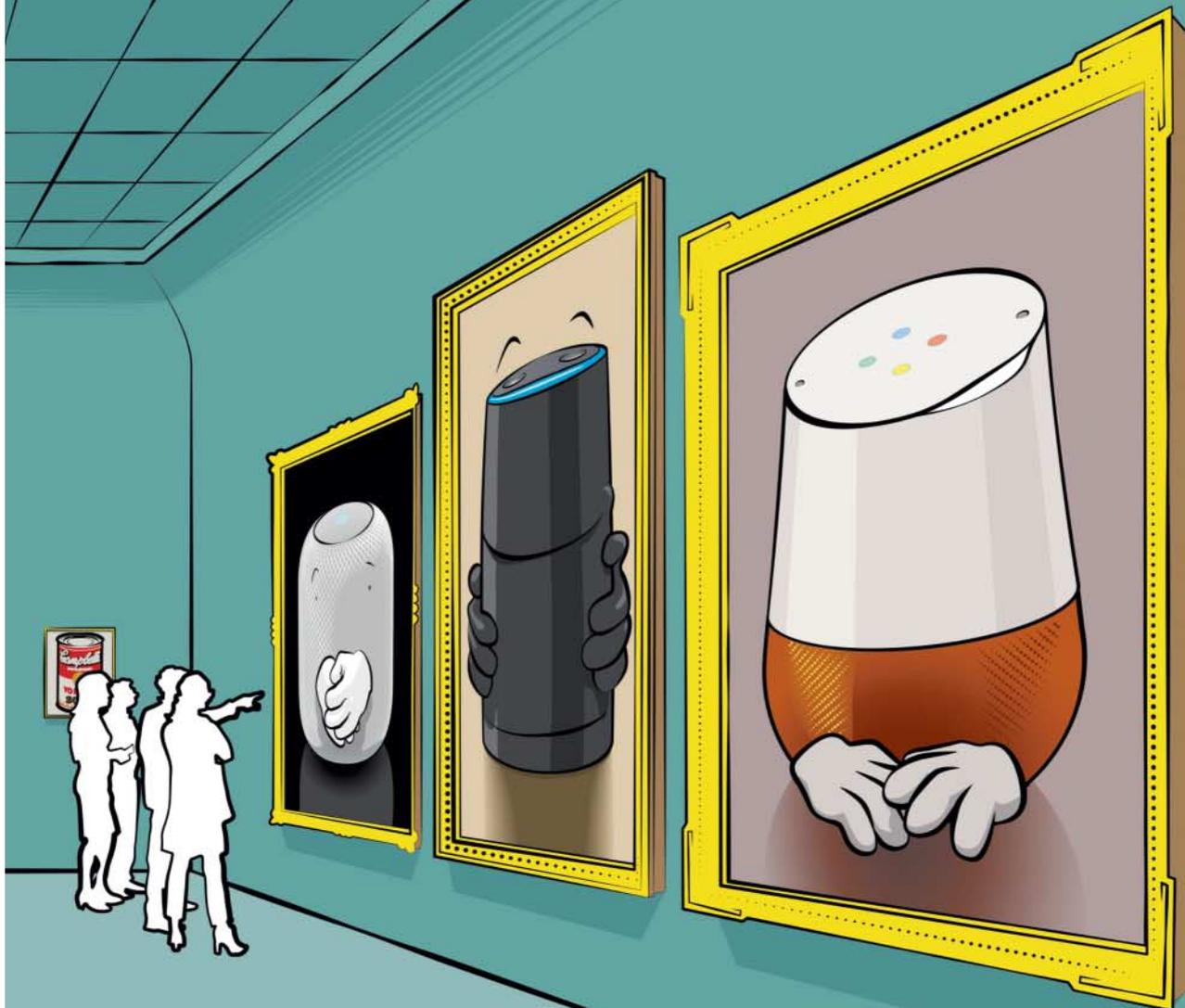


Sie sind da!

Sprachassistenten durchdringen den Alltag



Was Assistenten leisten	Seite 60
Test: Google Assistant, Alexa & Co.	Seite 66
Der Assistent der Zukunft	Seite 72
Assistenten im Datenschutz-Check	Seite 74

Bild: Rudolf A. Blaha

Ob man die Sprachassistenten für unverzichtbare Helferlein hält oder für den absoluten Daten-GAU – fast jeder ist schon mal mit Alexa & Co in Kontakt gekommen. Deshalb ist es wichtig, die Funktionsweise der Assistenten zu verstehen, denn sie sind weder reine Spielerei noch Alleskönner oder Datenspione.

Von **Sven Hansen, Nico Jurran**
und **Stefan Porteck**

Der Faszination einer menschelnden Maschine kann sich kaum jemand entziehen. Mit Amazons Echo stand in vielen Haushalten erstmals eine Sprechdose auf dem Tisch, die auf Zuruf Nachrichten, Informationen und Musik liefert, Witze erzählt und Geräte im Haushalt steuert. Auch wenn die erste Begeisterung verfliegen ist, sind die Sprachassistenten Teil des Alltags geworden: Laut einer aktuellen Studie nutzt heute rund ein Drittel aller Deutschen Alexa & Co., vor einem Jahr war es nur ein Fünftel.

Außer dem Sprachassistenten des Versandriesen Amazon sind in deutschen Haushalten der Google Assistant und Apple Siri anzutreffen. Siri nimmt eine Sonderstellung ein: Nicht nur, dass der Assistent – obwohl älter als Alexa – erst sehr spät über den „HomePod“ genannten vernetzten Lautsprecher einen festen Platz in Wohnräumen erhielt. Apple kündigte bislang auch keine weitere Variante des Geräts oder gar ein Integrationsprogramm für Dritthersteller an. Und während Amazon und Google sowohl Entwickler-Kits als auch die nötigen Begleit-Apps unter Android und iOS bereitstellen, lässt sich Siri nur auf dem HomePod und nur von iOS-Nutzern einrichten und konfigurieren.

Abseits der drei großen Assistenzsysteme versuchen sich immer wieder einzelne Gerätehersteller an einer eigenen Sprachsteuerung – beispielsweise Microsoft mit Cortana am PC, Samsung mit Bixby auf Smartphones und LG mit dem „Voice Mate“ auf seinen aktuellen Smart-TVs. Über letzteren lassen sich nicht nur TV-Sender nach Namen ansteuern und die Lautstärke verändern, er durchsucht beispielsweise auch YouTube nach Titeln und das TV-Angebot nach Genres. Auch die Autohersteller versuchen es mit eigen-

nen Assistenten: Bei den Premiummarken Mercedes und BMW kann man sich in ausgewählten Fahrzeugen mit MBux oder dem Personal Digital Assistant unterhalten.

Wie geht das?

Wenn sie in einem smarten Speaker verpackt sind, arbeiten Sprachassistentensysteme nach demselben Muster. Sie horchen über im Lautsprecher integrierte Fernfeldmikrofone permanent auf ein Aktivierungswort (englisch Hotword) wie „Alexa“ oder „Hey Google“. Dabei fließt noch nichts in die Cloud. Erst wenn sie ihr Wort erkennen, zeichnen sie den nachfolgenden Satz auf und schicken ihn verschlüsselt an die Cloud des jeweiligen Anbieters. Manche Assistenten in Geräten wie dem Fire-TV-Medienplayer und dem angesprochenen LG-Fernseher springen nur an, wenn der Nutzer eine bestimmte Taste auf der Fernbedienung mit integriertem Mikrofon drückt.

In der Cloud wird die Sprachaufzeichnung in Text umgewandelt und nach Kommandos durchsucht. Handelt es sich



Mit dem Amazon Echo (links) und Google Home begann der Sprachassistenten-Boom.

um eine Frage, suchen die Server die passende Antwort heraus, wandeln sie über ein Text-To-Speech-System in Sprache und schicken diese Audiodaten an den vernetzten Lautsprecher des Nutzers zurück. Geht es dem Anwender hingegen darum, mit seinem Sprachbefehl ein Smart-Home-Gerät zu schalten, übermitteln Amazon, Google & Co. das Kommando an den passenden Cloud-Dienst.

Die Intelligenz der Sprachassistenten steckt somit komplett in der Cloud und nicht im Gerät auf der heimischen Kommode. Das hat für die Anbieter den Vorteil, dass sich die Hardware auf Nutzerseite praktisch auf eine Kombination aus Lautsprecher, Fernfeldmikrofonen und Internetzugang beschränkt. Die lässt sich preisgünstig in vernetzte Lautsprecher integrieren – aber auch in Smart-TVs, Soundbars, Tablets und sogar in smarte Thermostate. Einen Überblick, in welchen Geräten mittlerweile Assistenten direkt zu finden sind, finden Sie auf Seite 64. Das Cloud-Konzept wirft allerdings immer wieder eine Reihe datenschutz- und verbraucherrechtlicher Fragen auf, die der Artikel ab Seite 74 beleuchtet.

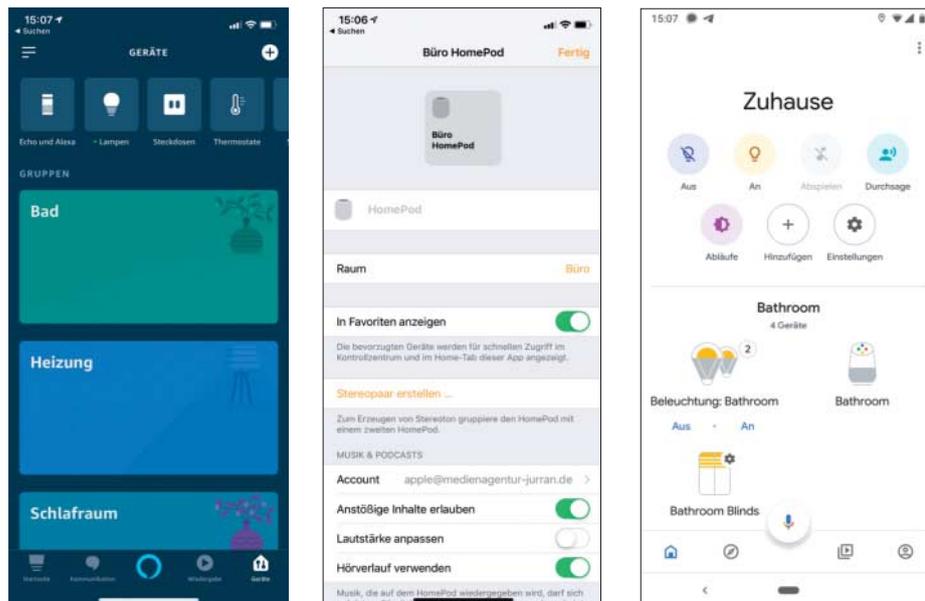
Sprachsynthese

Eine gelungene Sprachausgabe ist bei den Assistenzsystemen mindestens so wichtig wie die Erfassung und korrekte Auswertung der vom Nutzer gesprochenen Befehle. Schließlich soll der Anwender die Antworten auf Anhieb verstehen und nicht erst rätseln müssen, was die Assistentin gesagt hat. Eine Herausforderung stellen dabei fremdsprachige Begriffe und Eigennamen dar.

Aber auch die Stimme, die Sprechgeschwindigkeit und die Betonung sind wichtige Faktoren: Auf Dauer ist es anstrengend, einer Sprachassistentin zuzuhören, die sehr schnell oder sehr langsam spricht oder sprichwörtlich ohne Punkt und Komma vor sich hinredet.

Um zum gewünschten Ergebnis zu kommen, nutzen die Entwickler verschiedene Ansätze. Im einfachsten Fall lässt man eine reale Person Sätze komplett oder in größeren Abschnitten einsprechen und spielt diese Audiodateien dem Nutzer später passend vor. Dies bietet sich beispielsweise an, wenn der Nutzer nach einem Witz oder einer Geschichte fragt, wo es oft auf Pausen und die richtige Betonung ankommt.

Alternativ können die Anbieter Texte mit sogenannten SSML-Tags (Speech Syn-



Die Sprachassistenten von Amazon, Apple und Google werden über Smartphone-Apps eingerichtet. Hierüber laufen gewöhnlich auch alle weiteren Einstellungen.

thesis Markup Language) versehen, mit denen sich Pausen, Formatierung von Datum und Uhrzeit sowie andere Befehle zur Aussprache hinzufügen lassen. Auch hier muss zu irgendeinem Zeitpunkt ein Mensch eingreifen und die nötigen Metadaten hinterlegen.

In den übrigen Fällen hört der Nutzer eine automatisierte Sprachsynthese, die Texte Wort für Wort abarbeitet – etwa, wenn der Assistent nicht auf einen eigenen Datenbestand zurückgreifen kann und stattdessen einen Wikipedia-Eintrag vorliest. Den Unterschied merkt man sofort, da hier oft eine unrunde Satzmelodie zustande kommt. Auch die buchstabenweise Aussprache von Abkürzungen wie ü.NN für „über Normal Null“ macht das Zuhören anstrengend.

Funktionsumfang

Das Repertoire der Sprachassistenten ist schon heute riesig. Doch es gibt weder „die eine“ Killerapplikation noch „den einen“ Nutzer.

Die Anbieter selbst haben zum Marktstart der Smart-Speaker gerne die integrierten Wissensdatenbanken beworben. Das klang fast schon so, als würden Schulkinder bald Alexa ihre Hausaufgaben machen lassen. Das klappt vielleicht beim Übersetzen einzelner Vokabeln, aber tatsächlich lesen die Sprachassistenten bei Wissensfragen aktuell gerade mal Kurzzusammenfassungen aus Wikipedia und anderen Online-Quellen vor.

Ganz vorne steht für viele Nutzer von Alexa & Co. die Steuerung der Medienwiedergabe – angefangen von Internetradio über Spotify-Musik bis hin zu Videofilmen, die man sich auf vernetzten Lautsprechern mit integriertem Touch-Bildschirm, sogenannten Smart-Speakern, zeigen lassen kann. Praktisch ist dabei, dass Alexa, Google Assistant und Siri auch die Multiroom-Wiedergabe von Musik beherrschen.

Auch im Smart-Home haben die Assistenten ihren festen Platz. Für viele Nutzer dürften sie in Kombination mit smarten Leuchtmitteln – etwa von Philips, Innr und Ikea – sogar den Erstkontakt mit dem Thema herstellen. Doch nicht nur Lampen kann man per Sprache kontrollieren: Mittlerweile lassen sich Komponenten aus nahezu allen Bereichen ansprechen, darunter Schalter und Dimmer, Thermostate und Rolläden. Möglich machen dies bei Amazon und Google „Skills“ und „Actions“ genannte Software-Erweiterungen, über die Hersteller ihre IoT-Geräte an die Systeme andocken können. Apple brät hier wieder eine Extrawurst und unterstützt nur Komponenten, die das Unternehmen vorher für die eigene Smart-Home-Plattform HomeKit zertifiziert hat.

In manchem Haushalt kommen Alexa & Co. in den Smart-Speakern zudem als sprachgesteuerte Timer und Wecker zum Einsatz. Glaubt man Amazon, ist das Kochen nach Schritt-für-Schritt-Anleitung am Smart-Display eine

der beliebtesten Funktionen. Und natürlich kann man die Sprachassistenten mit klassischen Assistenzaufgaben betreuen – etwa dem Führen von Kalendern und Einkaufslisten.

Die Assistenten unterstützen auch die zwischenmenschliche Kommunikation und initiieren auf Zuruf über ihre Smart-Speaker beziehungsweise -Displays kostenfreie (Video)-Gespräche mit Freunden, Kollegen oder Verwandten über das Internet. In großen Wohnungen und Häusern können die smarten Lautsprecher schließlich die Aufgabe des Haustelefons übernehmen. Statt laut ins Treppenhaus zu brüllen, dass das Essen auf dem Tisch steht, gibt man beim Assistenten von Amazon oder Google eine Sprachnachricht als Rundruf an einen oder alle Smart-Speaker im Haushalt auf.

Dass praktisch alle Anbieter zumindest einen Großteil der genannten Felder mit ihren Assistenzsystemen abdecken, bedeutet nicht, dass man blind zu einer Lösung greifen sollte. Für den Assistenten-Test ab Seite 66 haben wir uns angeschaut, an welcher Stelle welcher Assistent seine Stärken und seine Schwächen hat – und wo alle versagen.

Wieso nicht offline?

Mancher Anwender wünscht sich, dass der Assistent gar keine Daten ins Internet schickt, sondern alles komplett auf dem Gerät läuft. Passende Projekte gibt es, auch im kommerziellen Bereich – etwa vom amerikanisch-französischem Startup Snips. Wie eine solche Lösung im privaten Bereich aussehen kann, hatten wir zudem in [1] am Beispiel des Open-Source-Projekts Jasper gezeigt.

Auch Amazon und Google denken in diese Richtung, wie unser Artikel über die Zukunft der Sprachassistentensysteme auf Seite 72 zeigt. Solche Offline-Lösungen fallen allerdings meist in die Kategorie der reinen Sprachsteuerung – schon, weil für aktuelle Informationen wie Nachrichten oder Wettervorhersagen ein Internetzugang nötig ist. Vor allem wäre so ein Ansatz zumindest mit aktuell günstigen Speakern, wie den Echos und Home Minis, nicht zu machen, denn die Offline-Hardware muss deutlich leistungsfähiger sein. (nij@ct.de) **ct**

Literatur

[1] Hilko Holweg, Der eigene Butler, Digitaler Assistent mit Offline-Spracherkennung im Eigenbau, c't 2/2016, S. 14

Anzeige

Wo Assistenten zu finden sind ...



Smart-Speaker und Soundbars

Vernetzte Lautsprecher mit integriertem Sprachassistenten gibt es heute in allen möglichen Formen, Farben und Größen von Amazon, Apple und Google. Soundspezialist Sonos bietet als einziger Produkte, auf denen man wahlweise Alexa oder Google Assistant nutzen kann.

Wer vorrangig seinen Fernseher mit gutem Sound ausstatten will, für den eignen sich Soundbars mit integriertem Google Assistant, die es unter anderem von JBL, LG und Teufel gibt. Wer den Platz unter dem Fernseher Alexa geben möchte, greift zur „Command Bar“ von Polk Audio. Bei der Sonos Beam hat man wieder die freie Wahl zwischen beiden Assistenten.

Smartphones und Tablets

Fast alle Assistenten waren vor ihrem Auftritt in Smart-Speakern zunächst auf Smartphones und oft auch auf Tablets verfügbar – und sind es immer noch.

Um die Sprachassistenten auf einem Mobilgerät zu nutzen, muss man diese allerdings meist in die Hand nehmen und entsperren. Apple und Samsung adressieren dieses Problem, in dem die Geräte nun auch ständig aufs Hotword horchen können. Amazon bietet bei seinen Tablets Fire HD 7 und 10 die Funktion „Alexa Hands-free“. Hier lässt sich Alexa bei aktiver WLAN-Verbindung auch im Standby ohne Knopfdruck ansprechen.



Fernseher und Medienplayer

Auf TV sind Assistenten eine wertvolle Hilfe – wenn sie etwa auf Zuruf den Sender wechseln oder Videos bei YouTube, Netflix & Co. heraussuchen. Philips und Sony nutzen Google Assistant, der als Teil von Android-TV nachgeliefert wurde.

LGs Fernseher bieten einen eigenen Assistenten, der auch einfache Fragen beantwortet, etwa nach dem Wetter. Dennoch spendierte ihnen der Hersteller 2018 zusätzlich den Google Assistant und in diesem Jahr noch Alexa. Samsung hält auf seinen TVs mit „Bixby“ dagegen.

Als Nachrüstlösung für alle TV gibt es Amazons Fire-TV mit integrierter Alexa und das Apple TV (4K) mit Siri. Auf Nvidias Shield TV kam der Google Assistant per Android-TV-Update.

Smartwatches

Einen Assistenten kann man heutzutage auch einfach am Handgelenk tragen. So steht Apples Siri bei der Apple Watch dank integriertem Mobilfunkmodem auch ohne Smartphone unterwegs Rede und Antwort.

Google hat seinen Assistant in sein Wear OS integriert, das auf Uhren unter anderem von Huawei, LG und sehr vielen Modellen von Fossil und deren Tochtermarken zu finden ist. Etliche der neuen Uhren verfügen auch über Lautsprecher, weshalb man bei ihnen nicht auf die Sprachausgabe verzichten muss.

Amazon lag hier lange zurück, bekommt Mitte September mit der Fitbit Versa 2 aber ein erstes Domizil am Handgelenk.



Bild: Fitbit

Smart-Displays

Smarte Displays sind eine recht neue Produktkategorie. Es handelt sich um Netzwerklautsprecher, die mit einem Touchscreen ausgestattet sind. Auf ihren Displays zeigen sie Zusatzinformationen zu den Sprachantworten des Assistenten an, geben aber beispielsweise auch Videos wieder, ermöglichen Videotelefonate und manche haben einen Browser.

Amazon hat zwei Modellreihen im Sortiment: „Echo Show“ mit eckigem Display in den Größen 5,5, 7 und 10 Zoll und „Echo Spot“ mit rundem 2,5-Zoll-Display (siehe Bild). Googles Angebot beschränkt sich hierzulande noch auf den „Google Nest Hub“ mit 7 Zoll großem Bildschirm. In englischsprachigen Ländern ist ab September der „Nest Home Hub Max“ mit 10"-Display erhältlich.



Im Auto

Eigentlich ist das Auto der perfekte Einsatzort für Assistenten – schließlich soll man während der Fahrt tunlichst die Finger vom Smartphone lassen und sich aufs Verkehrsgeschehen konzentrieren.

Die Verbreitung der Sprachassistenten ist hier bislang jedoch dürrig: Audi hat Alexa Auto in seinem e-Tron integriert, bei BMW soll sie per Update auf die älteren Infotainmentsysteme mit iDrive 5 / 6 kommen. Beim aktuellen BMW OS 7.0 setzt BMW hingegen auf seinen eigenen Sprachassistenten und folgt damit Mercedes, die mit MBux (siehe Bild) ebenfalls ein eigenes System nutzen.

