

@iBusiness

Dossier

JAHRGANG 12 | NOVEMBER 2017 | AUSGABE 4
THEMEN-BACKGROUNDER DER IBUSINESS-REDAKTION

Alexa: Commerce- Strategien für den Mittelstand



sponsored by

magnalister
...boost your Online Shop

Die Shop-Schnittstelle zu Amazon, eBay & Co

Erhältlich für Shopware, Magento,
xt:Commerce und viele mehr ...



magnalister ist optimiert für
amazonbusiness

@iBusiness

Dossier

JAHRGANG 12 | NOVEMBER 2017 | AUSGABE 4
THEMEN-BACKGROUND DER IBUSINESS-REDAKTION

Alexa: Commerce- Strategien für den Mittelstand



sponsored by

magnalister
...boost your Online Shop

Die Shop-Schnittstelle zu Amazon, eBay & Co

Erhältlich für Shopware, Magento,
xt:Commerce und viele mehr ...



magnalister ist optimiert für
amazonbusiness



Dominik Grollmann ist Redakteur des iBusiness-Redaktionsteams

Agenten-Commerce: So lernen Marken zu denken wie Alexa

Sprachassistenten gelten als das nächste große Ding im Online-Business. Doch bislang gibt es keine funktionierenden Konzepte, um die Geräte kommerziell zu nutzen. Der Beitrag erklärt, wie funktionierende User-Interfaces aussehen müssen.



Der Erfolg von Sprachassistenten kam überraschend. Zwar hat sich die Technologie mit Siri und Co. bereits angekündigt. Doch erst Amazons smarte Echo-Lautsprecher haben den Durchbruch geschafft - und die hatte niemand auf dem Radar. Inzwischen gilt in der Digitalbranche als ausgemacht, dass Sprachassistenten - neben AR und VR vielleicht - das nächste große Ding sein werden.

Allerdings: Auch wenn es kaum einer zugehen mag - wohin die Reise genau gehen wird, weiß immer noch niemand genau zu sagen. Für die Alexa-Plattform - deren System recht offen gestrickt ist - gibt es zwar inzwischen tausende „Skills“ genannte Anwendungen. Die meisten davon haben jedoch einen äußerst eingeschränkten Nutzwert. Wetterbericht und TV-Programm gehören im Reigen der zahlreichen Fuzkissen- und Witze-Skills bereits zu den Highlights.

Die bisherigen Einsatzmöglichkeiten geben nur einen eingeschränkten Hinweis auf die zukünftige Entwicklung. Genaue Nutzungsstatistiken sind nicht verfügbar. Bei Befragungen stellt sich bisher heraus, dass die Kunden zwar viele Funktionen ausprobieren, sich aber schnell auf eine übersichtliche Anzahl von Anwendungen beschränken. Das Abspielen von Musik und die Steuerung des Smart Homes stehen ganz oben auf den Listen. Allenfalls lässt sich festhalten, dass das Interface „Spracheingabe“ auf einen fruchtbaren Boden fällt und von den Nutzern - vor allem in den eigenen vier Wänden - akzeptiert wird.

Bislang mangelt es vor allem an Business-relevanten Szenarien. Allenfalls Amazon selbst verdient bislang mit dem Verkauf seiner Speaker etwas Geld. Obwohl es für Alexa bereits ein offenes App-Ökosystem in Form des Skill-Stores gibt, hat sich bislang kaum eine Anwendung gefunden,

► Das Beispiel zeigt: Beim Entwickeln sprachgesteuerter Skills muss komplett neu gedacht werden. Einerseits sind Assistenzfunktionen gefordert. Andererseits muss klassische Visualisierung durch Sprachkonzepte abgelöst werden.



mit der sich das App-Ökosystem monetarisieren ließe. Dabei lägen die wirtschaftlich sinnvollen Einsatzzwecke im Marketing, E-Commerce sowie Service und Support auf der Hand.

Woran es bei Alexa & Co hapert

Zunächst: Entwickler müssen komplett neu denken. Sprachassistenten dürfen nicht überschätzt werden. Die Dinger wirken schlau (weil sie ein Gespräch simulieren und manchmal überraschende, charmante und erheiternde Antworten liefern), tatsächlich sind sie es aber nicht.

Der einzig wirklich „intelligente“ Teil, den Alexa bietet, ist die automatische Spracherkennung (ASR, Automatic Speech Recognition) und das Verständnis natürlicher Sprache (NLU, Natural Language Understanding). Im

direkten Anschluss daran endet Alexas Intelligenz sehr schnell. Bei den meisten Anwendungen beginnt im Anschluss ein sehr schematischer Teil der Kommunikation.

Die Grenzen von Alexa lassen sich plastisch darstellen:

- Die Automatic Speech Recognition übernimmt die Rolle einer Tastatur, die aus einem gesprochenen Kommando („Alexa, wie wird das Wetter am Wochenende?“) einen Text erzeugt.
- Diese Zeichenkette wird an das Natural Language Understanding weitergereicht und dort in Intention und Parameter zerlegt. „Wie wird das Wetter“ bestimmt die Intention („Abfrage Wetterbericht“); die Angabe „am Wochenende“ wird in ein konkretes Datum übersetzt, die fehlende Ortsangabe wird durch den hinterlegten Standort des Geräts ersetzt.
- Anschließend werden diese Informationen in der durch die API definierte Form an den Wetterserver weitergereicht.
- Dieser startet eine entsprechende Datenbankabfrage und liefert einen aus vorgefertigten Bausteinen zusammengestellten Text an den Assistenten zurück, den die dort eingebaute Text-to-Speech-Engine in gesprochene Sprache umwandelt.
- Alternativ kann auch eine Audio-Datei oder eine Stream als Antwort erfolgen. Der Assistent spielt dann diese ab.

Manche Intentionen - wie die Abfrage des Wetterberichts - werden von Alexa an Services weitergeleitet, die Amazon selbst bereitstellt. Nur wenn ein Skill installiert ist, von denen es inzwischen mehr als 15.000 gibt, werden Anfragen auch an externe Server geleitet. Die Phrase „Alexa, frage WetterOnline wie das Wetter am Wochenende wird“ würde Alexa dann anweisen, Intention und Parameter an die Serveradresse des Drittanbieters zu

übermitteln und die Antwort von dort zu erwarten.

Alexa: Wie Tastatur und Display - nur völlig anders

Die künstliche Intelligenz hat in Sachen Spracherkennung und -Verständnis zweifellos gewaltige Fortschritte gemacht. Trotzdem: Sie funktionieren am besten mit einfach strukturierten Sätzen, die eine klare Intention erkennen lassen. Komplexe Anfragen oder implizite Informationen verstehen die Systeme gar nicht erst.

Und selbst wenn die Spracheingabe funktioniert, macht das noch keinen „intelligenten“ Assistenten aus. Denn für die passende Antwort ist der Sprachassistent in den wenigsten Fällen zuständig. Auch wenn der Nutzer den Eindruck hat, mit Alexa zu reden, stammt die Antwort nicht unbedingt von ihr. Alexa ist wie Tastatur und Display. Schlau muss der dahinter liegenden Service schon selbst sein.

Was das anbelangt, ist bislang nicht viel geboten. Alexa scheitert selbst an Aufgaben, für die lediglich bereits vorliegende Nutzerinformationen verknüpft werden müssten. Auf der Amazon-Website kann etwa dankenswerter Weise eingestellt werden, welches genaue Automodell der Kunde fährt - doch Alexa ist nicht in der Lage, diese Information zu berücksichtigen, wenn der Kunde neue Scheibenwischer bestellt. Stattdessen empfiehlt der Sprachassistent einen Wischer Typ „Bosch Dreimilliarden dreihundertsiebenundneunzigmillionen einhundertachtzehntausend neunhundertneundsiebzig Aerotwin“. Der natürlich nicht passt.

Noch weniger kann der Assistent intelligente Vorschläge aus der Einkaufshistorie ableiten. Wer Ersatzzahnbürsten bestellt, darf nicht erwarten, dass Alexa berücksichtigt, welche elektrische Zahnbürste man zuvor gekauft hat. Immerhin: Hat man einmal einen bestimmten Bürstentyp bestellt, schlägt Alexa diesen auch bei zukünftigen Käufen vor. Verbrauchsartikel lassen sich so unkompliziert nachbestellen. ▶

Das Beispiel zeigt: Beim Entwickeln sprachgesteuerter Skills muss komplett neu gedacht werden. Einerseits sind Assistenzfunktionen gefordert, denen nicht alles haarklein erklärt werden muss. Sie sollten statt dessen sinnvolle Vorschläge aus den vorhandenen Informationen ableiten können. Andererseits muss klassische Visualisierung durch Sprachkonzepte abgelöst werden. Diese Voice-User-Interfaces (VUI) müssen die speziellen Anforderungen einer Sprachein- und -ausgabe berücksichtigen. Die Informationsmenge muss reduziert sein und gegebenenfalls schrittweise abgefragt werden. Suchergebnisse müssen durch Filter vorselektiert und strukturiert werden. Bei der Gestaltung der Kommunikation mit Alexa gibt es zwei grundsätzliche Ansätze:

- Zum einen das „One-Shot“-Modell, durch welches eine einfache, schnelle Art der Kommunikation mit den Skills realisiert wird. „One Shots“ eignen sich für einfache Fragen oder Befehle („Wie wird das Wetter heute?“) oder „Spiele SWR3-Radio“.
- Für komplexere Anfragen eignet sich ein mehrstufiges Dialogmodell. Dabei

werden Daten des Nutzers abgefragt, gespeichert, zurückgegeben, verändert oder gelöscht. Die eindeutige Identifikation des Nutzers bei Alexa erfolgt über die Amazon.UserProfil_ID.

Für eine Fahrplanauskunft wählt der Nutzer beispielsweise zuerst den Service („Öffne Fahrplanauskunft“) und teilt dem System schrittweise den Start- und Zielbahnhof sowie die gewünschte Ankunfts- oder Abfahrtszeit mit. Anschließend fragt das System, ob die Ergebnisse nach günstigsten oder schnellsten Verbindungen sortiert werden sollen und nennt erst dann eine passende Verbindung mit Start- und Ankunftszeit, Anzahl der Umsteigebahnhöfe und Fahrpreis. Anschließend fragt das System, ob diese Verbindung passend ist, oder ob noch eine frühere oder spätere, eine schnellere oder eine günstigere Verbindung gesucht werden soll. Zuletzt kann das System - sofern der Kunde eingeloggt ist - die Verbindungsdetails per Mail schicken oder gleich ein Ticket kaufen und einen Sitzplatz reservieren.

Ebenso komplex würde ein funktionierender Einkaufsdialog aussehen. Nach der

Scheibenwischerbestellung würde sich Alexa vergewissern, ob das hinterlegte Automodell das richtige ist. Anschließend danach fragen, ob die Suchergebnisse nach Preis, Beliebtheit oder Bewertung sortiert werden sollen oder ob eine bestimmte Marke präferiert wird. Erst dann nennt Alexa das erste Suchergebnis (ohne dabei die Artikelnummer vorzulesen!) und bietet die Bestellung an.

Daten neu strukturiert

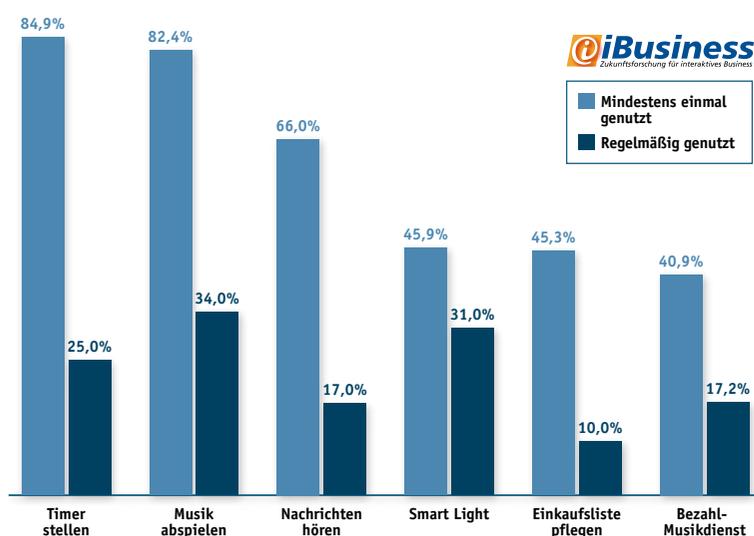
Die Beispiele zeigen, warum die meisten Alexa-Skills bislang One-Shots sind und sich kaum für komplexe Szenarien eignen. Obwohl die formalen Anforderungen an eine Alexa-Skill gering sind und sie sich schnell entwickeln lassen, sind für die meisten nützlichen Anwendungen grundlegende Vorarbeiten notwendig. Einfach nur den ersten Treffer einer Suchergebnisseite an die Text-to-Speech-Maschine zu parsen, kann langfristig kein Konzept für ein Voice-User-Interface sein.

In vielen Fällen muss nicht nur der Prozess der Informationsaufbereitung und -Abfrage neu durchdacht werden. Selbst die vorhandenen Datensätze müssen zunächst für eine Sprachausgabe passend aufbereitet werden, wie der Scheibenwischer „Bosch 3397118979 Aerotwin“ zeigt. Vor allem müssen die Daten mittels mehrstufiger Filter selektierbar sein, die vom VUI durch Nachfragen oder mittels bereits vorhandener Nutzerinformationen gesetzt werden.

Immerhin können Entwickler in der Regel auf bereits bestehende Schnittstellen wie eigene Datenbanken oder Server zugreifen und müssen deswegen selten komplett neue Datenstrukturen aufzubauen.

Und: Die Arbeit lohnt sich. Denn entsprechende Systeme lassen sich nicht nur über Smartspeaker ansprechen - von denen Prognosen zufolge bis 2020 in den USA rund 138 Millionen installiert sein sollen - sondern auch von Sprachassistenten in Smart Homes und textbasierten Chatbots. Der Bedarf an smarten Informationssystemen wird also steigen. ■

Diese Skills werden von Alexa-Nutzern ausprobiert



Quelle: Experian; Grafik: HighText Verlag

Die Shopping-Funktionen von Smart Speaker werden nur von wenigen dauerhaft genutzt



Dominik Grollmann ist Redakteur des iBusiness-Redaktionsteams



Foto: Alexas Fotos / CCO Public Domain

Arbeitswelt 2.0: Wie Siri und Alexa die Callcenter- Branche erschüttern

Millionen von Arbeitsplätzen werden in Zukunft von smarten Software-Robotern erledigt, sagen Analysten. Die ersten Vorboten werden schon jetzt in Callcentern spürbar. Dort könnten Alexa und Siri schon bald menschliche Mitarbeiter vollständig ablösen.

Agenten, KI und Robotic Process Automation (RPA) werden die Arbeitswelt in Zukunft komplett durcheinanderwirbeln. Besonders Dienstleistungen, bei denen Menschen viele Routinetätigkeiten ausführen, sind betroffen. Kassierer, Versicherungsagenten, Finanzberater, Buchhalter – und natürlich Callcenter-Agenten.

Gerade in letzterem Bereich scheint die Entwicklung besonders fortgeschritten. Mit Alexa, Siri und Co. stehen virtuelle Assistenten bereit, die schon heute erschöpfende Auskünfte geben können. Es scheint nur eine Frage der Zeit, bis sie auch Callcenter-Agenten ersetzen.

„In den nächsten zwei Jahren wird sich die Entwicklung beschleunigen“ meint daher auch Manfred Stockmann, Präsident des deutschen Callcenterverbands CCV und Inhaber der Change-Management-Beratung CMBS. *„Bislang werden komplexe Sprachassistenzsysteme noch eher zögerlich eingesetzt – sowohl bei Verbrauchern wie in Unternehmen. Das wird sich nun deutlich beschleunigen.“* Alexa, Siri und Co. ebnen den Weg.

Die Herausforderungen

Dem aktuellen Hype zum Trotz darf nicht übersehen werden, dass die Systeme noch ganz am Anfang stehen. Licht und Radio einschalten, Wetterbericht oder Verkehrsinformationen abrufen – das funktioniert meist tadellos. Aber schon wer ein Taxi oder eine Pizza bestellen will, kann die Tücken der Sprachsteuerung erleben.

Damit die Sprachassistenten zum wirklich intelligenten Helfer werden, müsste die KI die Pfade der Nutzer analysieren können und daraus kybernetische Schlussfolgerungen ziehen. Was meint der Nutzer, wenn er von einem „großen Auto“ spricht? Und selbst dann würde noch immer eine herausragende Eigenschaft fehlen: Die Systeme können zwar inzwischen schon Emotionen erkennen – aber noch nicht adäquat darauf reagieren. Vorgefertigte Textbausteine helfen nicht, wenn ein Mensch verärgert eine Hotline anruft.

„Solange ein Mensch lösungsorientiert vorgeht, können Sprachassistenten prinzipiell funktionieren. Wenn Menschen aber ▶

► Die Callcenter-Branche macht derzeit vor, wie Bestellungen per Sprachassistenten funktionieren können – und wo die Grenzen liegen.

emotionale Ansprache suchen – etwa weil sie verärgert sind – muss auch ein Mensch am Apparat sein“, erklärt Stockmann. Dabei ist Ärger nicht das einzige Gefühl, das eine Rolle spielt. Gar nicht so selten rufen Menschen bei der Bestellhotline eines Händlers an, weil sie Kontakt suchen.

Welche Auswirkung Sprachassistenten haben

Auch wenn sich die technischen Voraussetzungen verbessern, ist der Einsatz von sprachgeführten Systemen also keineswegs so zwangsläufig, wie es zunächst scheint. Die Gretchenfrage lautet: Wie viele Geschäftsvorfälle lassen sich in einem Callcenter schon jetzt automatisieren? Und: Sind Hintergrundprozesse schon soweit, dass sie sich leicht automatisieren lassen?

In der jetzigen Situation ist daher ein maßvolles Vorgehen angeraten. Folgende Voraussetzungen sollten Sie prüfen:

1. Automatisierung: Erlauben die Prozesse im Callcenter bereits eine Automatisierung?

2. Anwendungsumgebung: Ist ein Sprachassistenzsystem sinnvoll? Die Zielgruppe oder die Dienstleistung können dagegen sprechen. Dagegen sprechen beispielsweise

- Kunden, die auf persönlichen Kontakt Wert legen.
- Händler, die einen intensiven Kundenkontakt pflegen wollen.

3. Emotionaler Faktor: Steht eine emotionale oder eine lösungsorientierte Kundenansprache im Vordergrund? Selbst wenn alle vorhergehenden Prüfungen positiv ausfallen, verbietet sich

der Einsatz von Maschinen in emotional aufgeladenen Fällen. Eine Beschwerdhotline kann nicht von Maschinen besetzt sein – kein Mensch würde sich ernst genommen fühlen.

Anders sieht es bei standardisierten, technischen Abläufen aus. Wer seine Bankverbindung oder Adresse ändern will, eine Reise umbuchen oder ein Taxi bestellen will, kommt oft mit einer Maschine klar.

Rationalisierung statt Automatisierung

In allen anderen Fällen bieten sich hybride Lösungen an. Sogenannte IVR-Systeme (Interactive Voice Response) helfen, Anrufer vorzuselektieren, das genaue Anliegen abzufragen oder Vorklärungen vorzunehmen (etwa: Abfrage der Kunden- oder Versicherungsnummer). Solche Systeme sind keine neue Erfindung, Sprachassistenten, die natürliche Sprache verstehen, können sie aber attraktiver machen.

Noch deutlich lassen sich die Kosten reduzieren, wenn ein paralleler Ansatz gewählt wird, in dem zwar einerseits ein vollständig automatisierter Prozess angeboten wird, der Nutzer aber jederzeit auf einen menschlichen Mitarbeiter zurückspringen kann. Allerdings muss bei diesem Ansatz auch eine hinreichende Aussicht darauf bestehen, dass ein bestimmter Anteil der Kunden das automatisierte Verfahren auch abschließen wird. Greift dagegen die überwiegende Mehrheit auf die menschliche Unterstützung zurück, wurde lediglich ein zusätzlicher Kostenblock aufgebaut – und der Kundenfrust gesteigert. In diesen Fällen ist es günstiger, die Mitarbeiter mittels automatisierter Vorselektion und Vorklärung zu entlasten.

Automatisierung hinterlässt Spuren

Auf diese Weise werden in drei Jahren rund 160.000 der derzeit 500.000 Arbeitsplätze in Callcentern direkt oder indirekt beeinflusst sein, schätzt Stockmann. Das bedeutet aber nicht, dass damit automatisch ein radikaler Stellenabbau verbunden ist. Zum einen, weil derzeit ohnehin etwa 20.000 Stellen nicht besetzt werden können. Diese werden als erstes von der Automatisierungswelle aufgefangen. Zum zweiten werden vor allem Routinetätigkeiten durch Maschinen erledigt. Etwa die erwähnten Vorklärungsfragen oder bei der Prozessabwicklung im Nachklapp des Gesprächs. Dadurch werden Kapazitäten für den wesentlichen Teil des Gesprächs frei.

Mittels künstlicher Intelligenz sind noch genauere Selektionen möglich. Kunden, die kürzlich ein Handy gekauft haben, rufen bei der Hotline meist wegen Vertragsfragen an, nach dem Rechnungsversand sind es eher Tariffragen. Ist die Anrufererkennung mit den richtigen Daten verknüpft, können bereits im Vorfeld solche Schlüsse gezogen werden.

Grenzen der Automatisierung

Menschen werden aber in absehbarer Zeit nicht durch Sprachassistenten ersetzt werden können. „Das ist nicht nur eine technologische, sondern auch eine soziale Frage“ sagt Stockmann. Selbst wenn durch Siri und Co. die Akzeptanz für Sprachassistenten steigt, werden soziale Faktoren einen flächendeckenden Einsatz verhindern. „Die Frage lautet: Fühlen sich Menschen von Maschinen ernst genommen“, meint Stockmann und liefert die Antwort gleich mit. „Auf emotionaler Ebene wird das nicht funktionieren.“ ■

iBusiness Dossier –
Themenbackground der iBusiness-Redaktion
ISSN 1861-8197

Jahresbezug: Der Bezug von iBusiness Dossier ist in der iBusiness-Premium-Mitgliedschaft enthalten.

Jahresbeitrag: 390,- Euro, Preise inkl. Porto, zzgl. MwSt. iBusiness Dossier dient ausschließlich der Hintergrundinformation für iBusiness-Premium-Mitglieder.

Freiweisungen sind aus grundsätzlichen Überlegungen nicht möglich. Nachdruck und Vervielfältigung in jeder Form, auch auszugsweise nur mit schriftlicher Genehmigung.

Unberechtigter Nachdruck oder Vervielfältigung ziehen Schadensersatzforderungen nach sich.

Mitteilung gemäß § 8, Artikel 3 des Bayerischen Pressegesetzes

Inhaber des HighText-Verlag Graf und Treplin OHG sind zu jeweils 50 Prozent Joachim Graf (verantwortlich für Redaktion und Anzeigen) und Daniel Treplin (beide Journalisten, München);

Handelsregister München HR A 72216

Anschrift Verlag und Redaktion

HighText Verlag
Wilhelm-Riehl-Str. 13
80687 München
Tel.: (089) 57 83 87-0
Fax: (089) 57 83 87 99
E-Mail: info@ibusiness.de
URL: http://www.ibusiness.de

Redaktion@ibusiness.de
Joachim Graf, Dominik Gröllmann,
Sebastian Halm, Susan Rönisch, Daniel Treplin

Buchhaltung@ibusiness.de
Daniela Seitz

Anzeigen@ibusiness.de
Nicos Fucidis, Benjamin Karg, Dagmar Kraux
Disposition: Anna Strommer

Leserservice@ibusiness.de
und Vertrieb@ibusiness.de
Bianka Matthes (bma@ibusiness.de)

Technik@ibusiness.de
Sven Noherr

Grafik@ibusiness.de
Gertrud Graf, Monika Herbig

Druck: Holzmann Druck GmbH & Co. KG,
Gewerbestraße 2,
86825 Bad Wörishofen

Titelfoto: PublicDomainPictures/Pixabay

Das komplette Dossier auch
als klickbares PDF:
www.ibusiness.de/dossier

 iBusiness





Dominik Grollmann ist Redakteur des iBusiness-Redaktionsteams

Finde die Nische! Das wahre kommerzielle Potenzial von Alexa und Co.

Foto: HighText Verlag

Auch in diesem Jahr werden wohl wieder Tausende von intelligenten Lautsprechern verkauft. Doch trotz der riesigen Gerätebasis tut sich die Branche extrem schwer, einen kommerziellen Nutzen für die Geräte zu finden. Welche kommerziellen Szenarien funktionieren.

Digitale Sprachassistenten und intelligente Lautsprecher sind einer der wichtigsten Technologietrends des Jahres – vielleicht neben künstlicher Intelligenz, Augmented und Virtual Reality. Eine wahre Flut an Geräten ist 2017 auf den Markt gekommen (oder zumindest angekündigt worden). Schätzungen zufolge steht in den USA bereits in jedem zehnten Haushalt ein Alexa-Gerät. In Deutschland haben nach einer Statista-Umfrage bereits 28 Millionen Menschen Erfahrung mit Siri & Co. gesammelt. Und smarte Lautsprecher dürften auch 2018 zum Renner werden.

Angesichts dieser Aussichten kann es nicht verwundern, dass sich Analysten weltweit mit der Frage auseinandersetzen, wie sich diese Entwicklung für Werbung und Onlinehandel nutzen lässt. Der Werbeexperte Unruly erklärte auf der Digitalmesse Dmexco 2017 die Geräte und die damit verbundenen Einkaufsszenarien gar zum Messetrend überhaupt.

Die wichtigste Zukunftsaufgabe für Marken und Händler ist es, einen Zugang zu den smarten Devices der Konsumenten zu finden, erklärte Unruly-Futurist Simon

Gosling den Messebesuchern gebetsmühlenartig. Seine Szenarien: Wer sich in Zukunft einen Wagen kaufen will, wird die Innenausstattung in einer perfekten VR-Simulation konfigurieren – natürlich im eigenen Wohnzimmer und auf Zuruf an den smarten Agenten. Wem morgens die Rasierklingen ausgehen, wird noch im Bad per Sprachassistent neue bestellen.

Und wenn eine neue Klingenmarke auf den Markt kommt, wird der digitale Helfer freundlich darauf hinweisen und fragen, ob man nicht einmal diese ausprobieren möchte. *„Bislang müssen Marken und Händler versuchen, in das Relevant Set der Verbraucher vorzudringen“*, erklärt Gosling. *„In Zukunft wird es darum gehen, das Relevant Set der digitalen Assistenten zu erreichen.“*

Nicht zum ersten Mal springt die Branche so euphorisiert auf einen Trend auf. Schließlich kann Technologie das Leben verändern – wie Internet, GPS oder Smartphones beweisen. Oder spurlos vorüber gehen – wie es iBeacons, Dash-Buttons, QR-Codes oder der berühmte Internet-Kühlschrank ►

► Am Erfolg von Sprachassistenten ist kaum zu zweifeln – kommerziell ist das Potenzial aber schwer zu heben.

zeigten. Welchen der beiden Wege Einkaufsvarianten mit digitalen Assistenten nehmen werden ist bislang kaum abzusehen. Klar ist nur, dass nicht nur der britische Werbespezialist Unruly auf die Macht der Computerstimmen setzt.

Allein für Alexa wurden inzwischen mehr als 15.000 Skills entwickelt. Bei den meisten davon ist zwar kein Monetarisierungsziel erkennbar. Es gibt aber auch einige rein kommerzielle Ansätze. In Deutschland versuchen sich etwa die Supermarktketten Real und Hit sowie der Online-Marktplatz Allyouneed daran.

Taugt Alexa als Shopping-Assistent?

Dass sich die Entwicklung bislang fast ausschließlich auf Alexa konzentriert, ist kein Wunder. Der Amazon-Assistent ist für kommerzielle Anbieter am vielversprechendsten. Alexa ist der einzige Assistent, der mittels Schnittstelle um beliebige Anwendungen („Skills“) erweitert werden kann. Auf diese Weise erhält das System neue Fähigkeiten und kann auf Zuruf beispielsweise Sonderangebote des Supermarktes nebenan abspulen („Alexa, frage Real nach den Angeboten des

Tages“). Andere Assistenten lassen sich nicht so gezielt mit Inhalten und Aktionen füttern und würden auf die entsprechende Anfrage lediglich das Ergebnis einer Websuche präsentieren.

Doch unter allen verfügbaren Skills scheint ausgerechnet die Rubrik „Shopping“ eine der am wenigsten attraktivsten zu sein: Von den mehr als 2.500 in Deutschland erhältlichen Erweiterungen stammen gerade einmal 19 aus dieser Rubrik. Natürlich: Das kann auch daran liegen, dass eine der Grundfertigkeiten von Alexa darin besteht, Bestellungen bei Amazon auszuführen. Es ist also gar nicht nötig, einen Skill zu aktivieren, um mit Alexa einzukaufen. Außerdem sagt die Zahl der verfügbaren Skills auch nichts über die aktivierte Basis aus.

Amazon selbst veröffentlicht keine Statistiken über die Nutzungsraten einzelner Anwendungen. Noch weniger gibt es Zahlen zu den mit Alexa durchgeführten Einkäufen oder gar den daraus resultierenden Umsätzen. Aus Nutzerbefragungen lässt sich allerdings ein recht detailliertes Bild gewinnen. Zuletzt hat Pricewaterhouse Coopers im Juni 2017 eine bevölkerungsrepräsentative Umfrage unter

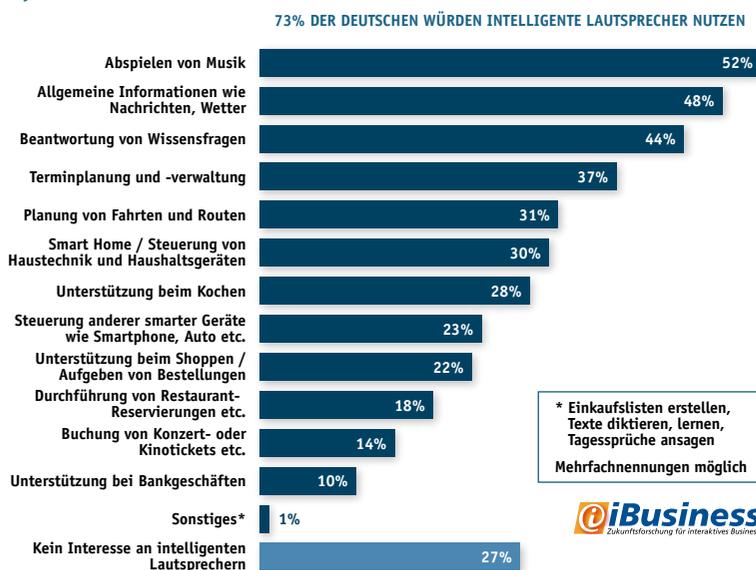
deutschen Bundesbürgern ab 18 Jahren zu ihrer Einstellung gegenüber digitalen Assistenten durchgeführt. Die Ergebnisse in der Zusammenfassung:

- **Hohe Bekanntheit, konkretes Interesse, geringe Verbreitung:** Die Bekanntheit von Sprachassistenten ist in Deutschland sehr hoch. Amazons Alexa ist mit 78 Prozent der bekannteste Vertreter. Die Verbreitung ist aber gering. Nur fünf Prozent besitzen ein Gerät mit Alexa. Das größte Interesse herrscht am Smart Speaker Google Home, immerhin jeder fünfte Deutsche gibt an, sich dafür erwärmen zu können.
- **Große Akzeptanz trotz Datenschutzbedenken:** Fast drei Viertel der Befragten können sich generell vorstellen, intelligente Lautsprecher zu nutzen. Die größten Bedenken herrschen hinsichtlich der Sicherheit. Die Gefahr von Datendiebstahl sowie die Manipulation oder der Missbrauch durch Hacker erscheinen 70 Prozent der Befragten „sehr groß“ oder „eher groß“. 69 Prozent fürchten, für Amazon, Apple oder Google zum gläsernen Kunden zu werden.

- **Anwender wollen Entertainment und Wissen:** Gefragt ist vor allem ein hochwertiger Ersatz für das Radio. Die meisten Nutzer wollen mit den Geräten Musik abspielen und auf Abruf Nachrichten und Wetterberichte hören. Es folgen allgemeine Wissensfragen sowie persönliche Assistenzfunktionen (Termin-, Routenplanung). Knapp ein Drittel wollen ihr Smart Home steuern. Konsumaufgaben (etwa Bestellungen, das Erstellen von Einkaufslisten oder Tischreservierungen, Ticketbuchungen) liegen weit abgeschlagen auf den hinteren Rängen (siehe Chart Seite 8).

Andere Statistiken kommen zu ähnlichen Ergebnissen. Im März 2017 fragte das Marktforschungsunternehmen Norstat, für welche Zwecke die Deutschen digitale Sprachassistenten nutzen würden, wenn sie denn welche hätten. Die Top-3-Einsatzgebiete: Suchmaschinen befragen, Terminerinnerungen und

Interesse: Diese Anwendungen wünschen sich Anwender von Sprachassistenten



Quelle: comScore; Grafik: HighText Verlag

Wettervorhersagen abhören. Ganz am Ende der Liste: „Taxi bestellen“ und „Waren bestellen“ mit je gerade einmal 14 Prozent positiven Nennungen.

An dem Interesse ändert sich auch nichts, wenn die Anwender die entsprechenden Funktionen ausprobieren konnten. Der Marktforscher Comscore hat im ersten Quartal 2017 amerikanische Besitzer von vernetzten Lautsprechern nach ihren Nutzungsgewohnheiten befragt. Ergebnis: 60 Prozent stellen allgemeine Fragen, 57 Prozent hören den Wetterbericht, 54 Prozent Musik. Und E-Commerce? Die Funktionen „Produkte bestellen“ (11 Prozent) und „Lebensmittel/Essen bestellen“ (8 Prozent) landeten auf den hintersten Rängen 11 und 12.

Bleibt festzuhalten: Die meisten Anwender wollen keine Produkte per Sprachassistenten bestellen und die wenigsten Anwender tun dies in der Praxis. Aber vielleicht bedingen sich die Ergebnisse gegenseitig? Ist es nur eine Frage der Gewöhnung und die Nutzer würden eifrig Sprach-Shopper, wenn sie die Funktion nur einmal probieren würden?

Für diese These gibt es keinen Beleg – im Gegenteil: Im Mai 2016 wollte Experian von amerikanischen Alexa-Nutzern wissen, welche Funktionen sie bereits ausprobiert haben und welche sie anschließend regelmäßig nutzen. Ergebnis: Erwartungsgemäß haben sich vergleichsweise wenige (45,3 Prozent) für die Shopping-Funktion interessiert. Noch viel weniger blieben ihr aber treu. Nur zehn Prozent kauften anschließend „regelmäßig“ mit Alexa ein – das schlechteste Ergebnis in der gesamten Umfrage. Selbst Funktionen, die anfangs von viel weniger Nutzern ausprobiert wurden, kamen schließlich auf deutlich mehr regelmäßige Anwender.

Shopping-Hype nur herbeigeredet

So eindrucksvoll manche Präsentation auf der Dmexco war, so ernüchternd fällt das Fazit nach Auswertung der statistischen Parameter aus.

- **Anfang des Hype-Cycles:** Smarte Lautsprecher und die enthaltenen Assistenten haben eine enorme Bekanntheit, das Interesse daran ist hoch und die Verbreitung gering. Sie stehen damit am Anfang eines klassischen Hype-Cycles und die Absatzprognosen für die kommenden Monate dürften rosig sein. Typisch ist aber auch, dass auf die Anfangseuphorie eine Ernüchterung folgt, bevor sich die Anwendung im Alltag verfestigt. Diese Phase steht intelligenten Lautsprechern noch bevor.

- **Kein Shopping-Hype:** Für ECommerce-Szenarien gilt diese Entwicklung dagegen nicht. Ein Hype ist gar nicht erst zu erkennen. Es gibt weder eine anfängliche Euphorie, noch Nutzungsszenarien, die sich im Alltag herausbilden. Im Gegenteil: Anwender, die die Funktion einmal ausprobiert haben, verwenden sie meist anschließend nicht mehr.

Diese Einschränkungen gelten nicht nur für ECommerce-Anwendungen. Durchweg alle Funktionen, die mit dem Einkauf oder Transaktionen zu tun haben, erhalten schlechte Noten.

- Pflegen von Einkaufslisten,
- Essensbestellung (z.B. Pizzadienst),
- Tickets für Konzert, Kino und Theater,

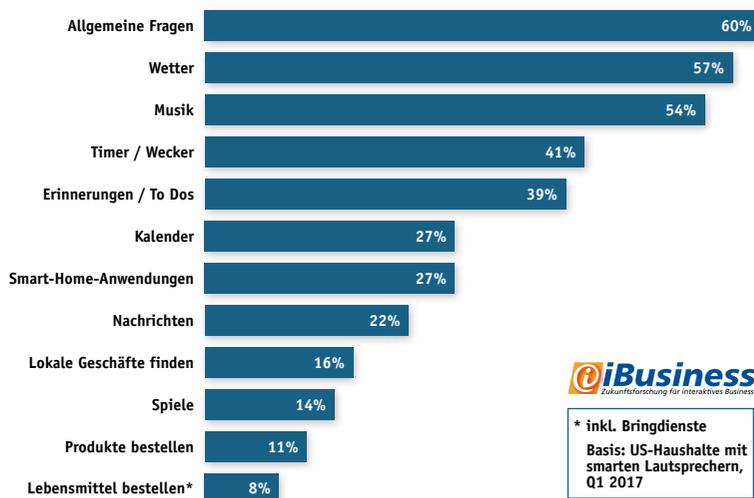
- Taxiruf
- und selbst die Reservierung in einem Restaurant.

Chance liegt in der Nische

Alexa und Co. werden so schnell keine Shopping-Revolution auslösen. Völlig vernachlässigen sollten Händler die Systeme aber trotzdem nicht. Selbst wenn nur jeder zehnte Nutzer seinen Sprachassistenten regelmäßig zum Einkauf verwendet, kann dies eine interessante Zielgruppe werden. Zudem ist es möglich, dass zwar insgesamt nur ein kleiner Teil der Verkäufe über Assistenten abgewickelt wird, in bestimmten Situationen oder Produktkategorien jedoch der Großteil. Anders gesagt: Es kann sich eine lukrative Nische auf tun.

Zunächst muss sich jeder Anbieter bewusst sein, dass er einen juristischen Graubereich betritt. „Wir sehen bei Alexa-Bestellungen die gleiche Problematik wie beim Dash-Button“, meint Carsten Fölisch, Jurist des Shopzertifizierers Trusted Shops. Den Kunden stellen diese Einschränkungen allerdings nicht schlechter. Denn dieser hat im Zweifel nicht nur ein Widerrufsrecht, sondern kann sich sogar darauf berufen, dass gar kein Vertrag zustande gekommen ist. ▶

Nutzung: Diese Sprachassistenten-Anwendungen werden tatsächlich genutzt



Quelle: comScore; Grafik: HighText Verlag

iBusiness
Zukunftsforschung für interaktives Business

* inkl. Bringdienste
Basis: US-Haushalte mit smarten Lautsprechern, Q1 2017

Schwieriger ist die Situation dagegen für den Händler. „Wir glauben, dass die gesetzlichen Informationspflichten im Fernabsatz und im E-Commerce weder beim Dash-Button noch bei Käufen über Alexa eingehalten werden“, sagt Föhlisch.

Juristisch stellen Transaktionen über Sprachassistenten noch immer ein Wagnis dar, wenngleich ein kalkulierbares. Weniger Einfluss kann der Händler dagegen auf die Einkaufssituation nehmen. Vier Szenarien lassen sich unterscheiden – mit unterschiedlichen Erfolgsaussichten für den Händler.

ECommerce-Szenarien für digitale Assistenten:

1. Komplexe Produkte

Die meisten, selbst vermeintlich einfache ECommerce-Einkäufe sind für einen Sprachassistenten ungeeignet. Die genaue Auswahl erfolgt zumeist am Bildschirm und Bestellwerte liegen meist über zehn Euro. Selbst bei vermeintlich einfachen Dingen wie einem USB-Stick sind am Ende noch zahlreiche Parameter (Preis, Marke, Kapazität, Größe, Standard) zu klären. Die Hürden:

- Ein mehrstufiges Voice-User-Interface (VUI), das Rückfragen beherrscht, wäre nötig. Kein Sprachassistent beherrscht dies derzeit zufriedenstellend.
- Eine echte Assistentenfunktion („Kaufe eine neue Autobatterie“) oder gar Stilberatung scheitert an mangelndem Wissen und Fähigkeiten. Die KI hinter der Sprachverarbeitung ist wenig leistungsfähig.
- Selbst wenn diese technischen Hürden gelöst wären, müsste der Nutzer darauf vertrauen, dass der Sprachassistent eine Wahl im Sinne des Benutzers trifft.
- Funktionieren kann das Szenario nur, wenn Produkt und Preis klar definiert sind. Zum Beispiel bei einem Musik-Album oder einem Buch.

Bewertung: Weitgehend ungeeignet.

2. Convenience-Produkte

Am besten könnte Assi-Commerce bei Produkten funktionieren, die gerne beim Lebenspartner bestellt werden („Schatz, bringst du noch Milch mit“). Das Produkt ist einfach, ein Fehlkauf praktisch ausgeschlossen. Der Partner weiß schon, was gemeint ist (fettarme Milch). Die Hürden:

- Die meisten Produkte sind niedrigpreisig und das Sortiment für E-Commerce nicht geeignet.
- Die Präsenz des stationären Einzelhandels ist in diesen Segmenten enorm, die Konkurrenz groß.
- Digitalen Assistenten fehlt Wissen („Kaufe Batterien ... die kleinen, für die Fernbedienung.“) Am ehesten lassen sich Produkte, die regelmäßig verbraucht werden, nachordern. Spontankäufe sind ausgeschlossen.
- Entscheidend ist technologischer Fortschritt. Der Assistent muss mehr Wissen anhäufen, um nützlich zu sein. Welche Glühbirne passt in die Wohnzimmerlampe? Welcher Scheibenwischer ans Auto?
- Fraglich bleibt, ob Kunden dieses Wissen weitergeben möchten. Datenschutzbedenken zählen schließlich zu den größten Befürchtungen.

Bewertung: Vor allem betriebswirtschaftlich herausfordernd.

3. Kaufentscheidungsprozess

Marken und Händler können die Geräte nutzen, um sich im Kaufentscheidungsprozess zu positionieren. Rezeptideen vom Supermarkt sind eine naheliegende Idee. Die zunehmende Vernetzung mehrerer Geräte (Smartphone, Laptop, TV, Smart Speaker) erlaubt es, digitale Assistenten auch in komplexere Kaufentscheidungen einzubinden. Die Aussichten:

- Derzeit ist der Einfluss von Sprachassistenten auf Kaufentscheidungen eher gering. Das VUI ist ungenügend.

- Am ehesten sind einfache Suchanfragen für lokale Ladengeschäfte denkbar („Wo ist der nächste Schlüsseldienst?“). Allerdings ist das Potenzial begrenzt.

- In Zukunft könnten Sprachassistenten eine wichtigere Rolle übernehmen. Sie könnten zentrales Ein- und Ausgabegerät einer cloudbasierten Intelligenz werden. Alexa, Siri und Cortana halten die lose vernetzten Geräte (Entertainment-System im Auto, Smartphone, PC, TV-Set und Smart Home) zusammen.

- Auf Kommando („Zeige meine Suchergebnisse auf dem Fernseher“) würden die vernetzten Geräte mit den gespeicherten Inhalten beschickt.

Das Szenario ist zukunftsweisend, der Zeithorizont allerdings noch relativ groß. Im Vorteil sind Hersteller, die über eine breite Basis an Geräten und Diensten verfügen. Allen voran wäre hier Google, gefolgt von Apple, zu nennen.

Bewertung: Langfristiges Disruptionspotenzial.

4. Digitale Services

Kurzfristig sind dagegen digitale Produkte am ehesten für Assistenten-Commerce geeignet. Die Aussichten:

- An erster Stelle sind Audio-Inhalte und Abos von Streaming-Diensten denkbar.

- Mit den Skills hat Amazon gezeigt, dass auch andere Erweiterungen praktikabel sind.

- Noch fehlt ein kommerziell attraktives Ökosystem („Skill-Store“). Konkurrenten wie Apple oder Google könnten Bewegung in den Markt bringen.

- Die kommerziell attraktivsten Dienste (Medien, Streaming Audio) werden von den Inhabern der Ökosysteme selbst angeboten. Die Eintrittsbarrieren sind entsprechend hoch.

Bewertung: Hohes Disruptionspotenzial, ebenso hohe Eintrittsbarrieren. ■



Dominik Grollmann ist Redakteur des iBusiness-Redaktionsteams

Agenten-Commerce: So lernen Marken zu denken wie Alexa

Sprachassistenten gelten als das nächste große Ding im Online-Business. Doch bislang gibt es keine funktionierenden Konzepte, um die Geräte kommerziell zu nutzen. Der Beitrag erklärt, wie funktionierende User-Interfaces aussehen müssen.



Der Erfolg von Sprachassistenten kam überraschend. Zwar hat sich die Technologie mit Siri und Co. bereits angekündigt. Doch erst Amazons smarte Echo-Lautsprecher haben den Durchbruch geschafft - und die hatte niemand auf dem Radar. Inzwischen gilt in der Digitalbranche als ausgemacht, dass Sprachassistenten - neben AR und VR vielleicht - das nächste große Ding sein werden.

Allerdings: Auch wenn es kaum einer zugehen mag - wohin die Reise genau gehen wird, weiß immer noch niemand genau zu sagen. Für die Alexa-Plattform - deren System recht offen gestrickt ist - gibt es zwar inzwischen tausende „Skills“ genannte Anwendungen. Die meisten davon haben jedoch einen äußerst eingeschränkten Nutzwert. Wetterbericht und TV-Programm gehören im Reigen der zahlreichen Fuzkissen- und Witze-Skills bereits zu den Highlights.

Die bisherigen Einsatzmöglichkeiten geben nur einen eingeschränkten Hinweis auf die zukünftige Entwicklung. Genaue Nutzungsstatistiken sind nicht verfügbar. Bei Befragungen stellt sich bisher heraus, dass die Kunden zwar viele Funktionen ausprobieren, sich aber schnell auf eine übersichtliche Anzahl von Anwendungen beschränken. Das Abspielen von Musik und die Steuerung des Smart Homes stehen ganz oben auf den Listen. Allenfalls lässt sich festhalten, dass das Interface „Spracheingabe“ auf einen fruchtbaren Boden fällt und von den Nutzern - vor allem in den eigenen vier Wänden - akzeptiert wird.

Bislang mangelt es vor allem an Business-relevanten Szenarien. Allenfalls Amazon selbst verdient bislang mit dem Verkauf seiner Speaker etwas Geld. Obwohl es für Alexa bereits ein offenes App-Ökosystem in Form des Skill-Stores gibt, hat sich bislang kaum eine Anwendung gefunden,

► Das Beispiel zeigt: Beim Entwickeln sprachgesteuerter Skills muss komplett neu gedacht werden. Einerseits sind Assistenzfunktionen gefordert. Andererseits muss klassische Visualisierung durch Sprachkonzepte abgelöst werden.



mit der sich das App-Ökosystem monetarisieren ließe. Dabei lägen die wirtschaftlich sinnvollen Einsatzzwecke im Marketing, E-Commerce sowie Service und Support auf der Hand.

Woran es bei Alexa & Co hapert

Zunächst: Entwickler müssen komplett neu denken. Sprachassistenten dürfen nicht überschätzt werden. Die Dinger wirken schlau (weil sie ein Gespräch simulieren und manchmal überraschende, charmante und erheiternde Antworten liefern), tatsächlich sind sie es aber nicht.

Der einzig wirklich „intelligente“ Teil, den Alexa bietet, ist die automatische Spracherkennung (ASR, Automatic Speech Recognition) und das Verständnis natürlicher Sprache (NLU, Natural Language Understanding). Im

direkten Anschluss daran endet Alexas Intelligenz sehr schnell. Bei den meisten Anwendungen beginnt im Anschluss ein sehr schematischer Teil der Kommunikation.

Die Grenzen von Alexa lassen sich plastisch darstellen:

- Die Automatic Speech Recognition übernimmt die Rolle einer Tastatur, die aus einem gesprochenen Kommando („Alexa, wie wird das Wetter am Wochenende?“) einen Text erzeugt.
- Diese Zeichenkette wird an das Natural Language Understanding weitergereicht und dort in Intention und Parameter zerlegt. „Wie wird das Wetter“ bestimmt die Intention („Abfrage Wetterbericht“); die Angabe „am Wochenende“ wird in ein konkretes Datum übersetzt, die fehlende Ortsangabe wird durch den hinterlegten Standort des Geräts ersetzt.
- Anschließend werden diese Informationen in der durch die API definierte Form an den Wetterserver weitergereicht.
- Dieser startet eine entsprechende Datenbankabfrage und liefert einen aus vorgefertigten Bausteinen zusammengestellten Text an den Assistenten zurück, den die dort eingebaute Text-to-Speech-Engine in gesprochene Sprache umwandelt.
- Alternativ kann auch eine Audio-Datei oder eine Stream als Antwort erfolgen. Der Assistent spielt dann diese ab.

Manche Intentionen - wie die Abfrage des Wetterberichts - werden von Alexa an Services weitergeleitet, die Amazon selbst bereitstellt. Nur wenn ein Skill installiert ist, von denen es inzwischen mehr als 15.000 gibt, werden Anfragen auch an externe Server geleitet. Die Phrase „Alexa, frage WetterOnline wie das Wetter am Wochenende wird“ würde Alexa dann anweisen, Intention und Parameter an die Serveradresse des Drittanbieters zu

übermitteln und die Antwort von dort zu erwarten.

Alexa: Wie Tastatur und Display - nur völlig anders

Die künstliche Intelligenz hat in Sachen Spracherkennung und -Verständnis zweifellos gewaltige Fortschritte gemacht. Trotzdem: Sie funktionieren am besten mit einfach strukturierten Sätzen, die eine klare Intention erkennen lassen. Komplexe Anfragen oder implizite Informationen verstehen die Systeme gar nicht erst.

Und selbst wenn die Spracheingabe funktioniert, macht das noch keinen „intelligenten“ Assistenten aus. Denn für die passende Antwort ist der Sprachassistent in den wenigsten Fällen zuständig. Auch wenn der Nutzer den Eindruck hat, mit Alexa zu reden, stammt die Antwort nicht unbedingt von ihr. Alexa ist wie Tastatur und Display. Schlau muss der dahinter liegenden Service schon selbst sein.

Was das anbelangt, ist bislang nicht viel geboten. Alexa scheitert selbst an Aufgaben, für die lediglich bereits vorliegende Nutzerinformationen verknüpft werden müssten. Auf der Amazon-Website kann etwa dankenswerter Weise eingestellt werden, welches genaue Automodell der Kunde fährt - doch Alexa ist nicht in der Lage, diese Information zu berücksichtigen, wenn der Kunde neue Scheibenwischer bestellt. Stattdessen empfiehlt der Sprachassistent einen Wischer Typ „Bosch Dreimilliarden dreihundertsiebenundneunzigmillionen einhundertachtzehntausend neunhundertneundsiebzig Aerotwin“. Der natürlich nicht passt.

Noch weniger kann der Assistent intelligente Vorschläge aus der Einkaufshistorie ableiten. Wer Ersatzzahnbürsten bestellt, darf nicht erwarten, dass Alexa berücksichtigt, welche elektrische Zahnbürste man zuvor gekauft hat. Immerhin: Hat man einmal einen bestimmten Bürstentyp bestellt, schlägt Alexa diesen auch bei zukünftigen Käufen vor. Verbrauchsartikel lassen sich so unkompliziert nachbestellen. ►

Das Beispiel zeigt: Beim Entwickeln sprachgesteuerter Skills muss komplett neu gedacht werden. Einerseits sind Assistenzfunktionen gefordert, denen nicht alles haarklein erklärt werden muss. Sie sollten statt dessen sinnvolle Vorschläge aus den vorhandenen Informationen ableiten können. Andererseits muss klassische Visualisierung durch Sprachkonzepte abgelöst werden. Diese Voice-User-Interfaces (VUI) müssen die speziellen Anforderungen einer Sprachein- und -ausgabe berücksichtigen. Die Informationsmenge muss reduziert sein und gegebenenfalls schrittweise abgefragt werden. Suchergebnisse müssen durch Filter vorselektiert und strukturiert werden. Bei der Gestaltung der Kommunikation mit Alexa gibt es zwei grundsätzliche Ansätze:

- Zum einen das „One-Shot“-Modell, durch welches eine einfache, schnelle Art der Kommunikation mit den Skills realisiert wird. „One Shots“ eignen sich für einfache Fragen oder Befehle („Wie wird das Wetter heute?“) oder „Spiele SWR3-Radio“.
- Für komplexere Anfragen eignet sich ein mehrstufiges Dialogmodell. Dabei

werden Daten des Nutzers abgefragt, gespeichert, zurückgegeben, verändert oder gelöscht. Die eindeutige Identifikation des Nutzers bei Alexa erfolgt über die Amazon.UserProfil_ID.

Für eine Fahrplanauskunft wählt der Nutzer beispielsweise zuerst den Service („Öffne Fahrplanauskunft“) und teilt dem System schrittweise den Start- und Zielbahnhof sowie die gewünschte Ankunfts- oder Abfahrtszeit mit. Anschließend fragt das System, ob die Ergebnisse nach günstigsten oder schnellsten Verbindungen sortiert werden sollen und nennt erst dann eine passende Verbindung mit Start- und Ankunftszeit, Anzahl der Umsteigebahnhöfe und Fahrpreis. Anschließend fragt das System, ob diese Verbindung passend ist, oder ob noch eine frühere oder spätere, eine schnellere oder eine günstigere Verbindung gesucht werden soll. Zuletzt kann das System - sofern der Kunde eingeloggt ist - die Verbindungsdetails per Mail schicken oder gleich ein Ticket kaufen und einen Sitzplatz reservieren.

Ebenso komplex würde ein funktionierender Einkaufsdialog aussehen. Nach der

Scheibenwischerbestellung würde sich Alexa vergewissern, ob das hinterlegte Automodell das richtige ist. Anschließend danach fragen, ob die Suchergebnisse nach Preis, Beliebtheit oder Bewertung sortiert werden sollen oder ob eine bestimmte Marke präferiert wird. Erst dann nennt Alexa das erste Suchergebnis (ohne dabei die Artikelnummer vorzulesen!) und bietet die Bestellung an.

Daten neu strukturiert

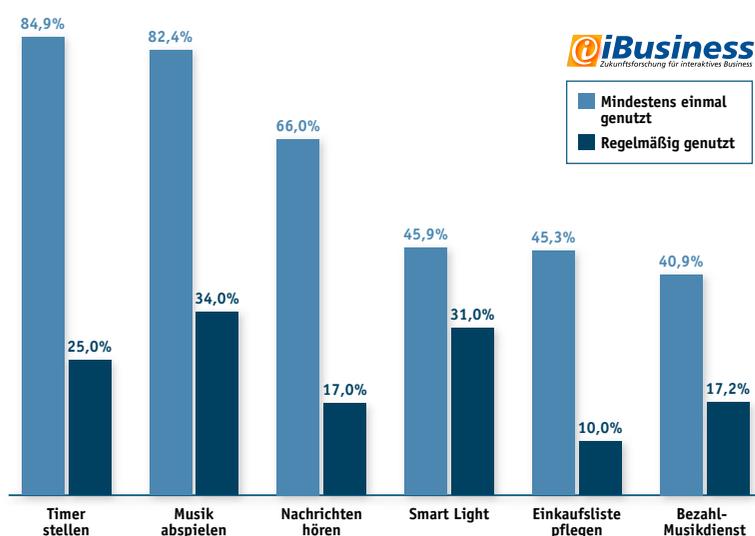
Die Beispiele zeigen, warum die meisten Alexa-Skills bislang One-Shots sind und sich kaum für komplexe Szenarien eignen. Obwohl die formalen Anforderungen an eine Alexa-Skill gering sind und sie sich schnell entwickeln lassen, sind für die meisten nützlichen Anwendungen grundlegende Vorarbeiten notwendig. Einfach nur den ersten Treffer einer Suchergebnisseite an die Text-to-Speech-Maschine zu parsen, kann langfristig kein Konzept für ein Voice-User-Interface sein.

In vielen Fällen muss nicht nur der Prozess der Informationsaufbereitung und -Abfrage neu durchdacht werden. Selbst die vorhandenen Datensätze müssen zunächst für eine Sprachausgabe passend aufbereitet werden, wie der Scheibenwischer „Bosch 3397118979 Aerotwin“ zeigt. Vor allem müssen die Daten mittels mehrstufiger Filter selektierbar sein, die vom VUI durch Nachfragen oder mittels bereits vorhandener Nutzerinformationen gesetzt werden.

Immerhin können Entwickler in der Regel auf bereits bestehende Schnittstellen wie eigene Datenbanken oder Server zugreifen und müssen deswegen selten komplett neue Datenstrukturen aufzubauen.

Und: Die Arbeit lohnt sich. Denn entsprechende Systeme lassen sich nicht nur über Smartspeaker ansprechen - von denen Prognosen zufolge bis 2020 in den USA rund 138 Millionen installiert sein sollen - sondern auch von Sprachassistenten in Smart Homes und textbasierten Chatbots. Der Bedarf an smarten Informationssystemen wird also steigen. ■

Diese Skills werden von Alexa-Nutzern ausprobiert



Quelle: Experian; Grafik: HighText Verlag

Die Shopping-Funktionen von Smart Speaker werden nur von wenigen dauerhaft genutzt



Dominik Grollmann ist Redakteur des iBusiness-Redaktionsteams



Foto: Alexas Fotos / CCO Public Domain

Arbeitswelt 2.0: Wie Siri und Alexa die Callcenter- Branche erschüttern

Millionen von Arbeitsplätzen werden in Zukunft von smarten Software-Robotern erledigt, sagen Analysten. Die ersten Vorboten werden schon jetzt in Callcentern spürbar. Dort könnten Alexa und Siri schon bald menschliche Mitarbeiter vollständig ablösen.

Agenten, KI und Robotic Process Automation (RPA) werden die Arbeitswelt in Zukunft komplett durcheinanderwirbeln. Besonders Dienstleistungen, bei denen Menschen viele Routinetätigkeiten ausführen, sind betroffen. Kassierer, Versicherungsagenten, Finanzberater, Buchhalter – und natürlich Callcenter-Agenten.

Gerade in letzterem Bereich scheint die Entwicklung besonders fortgeschritten. Mit Alexa, Siri und Co. stehen virtuelle Assistenten bereit, die schon heute erschöpfende Auskünfte geben können. Es scheint nur eine Frage der Zeit, bis sie auch Callcenter-Agenten ersetzen.

„In den nächsten zwei Jahren wird sich die Entwicklung beschleunigen“ meint daher auch Manfred Stockmann, Präsident des deutschen Callcenterverbands CCV und Inhaber der Change-Management-Beratung CMBS. *„Bislang werden komplexe Sprachassistenzsysteme noch eher zögerlich eingesetzt – sowohl bei Verbrauchern wie in Unternehmen. Das wird sich nun deutlich beschleunigen.“* Alexa, Siri und Co. ebnen den Weg.

Die Herausforderungen

Dem aktuellen Hype zum Trotz darf nicht übersehen werden, dass die Systeme noch ganz am Anfang stehen. Licht und Radio einschalten, Wetterbericht oder Verkehrsinformationen abrufen – das funktioniert meist tadellos. Aber schon wer ein Taxi oder eine Pizza bestellen will, kann die Tücken der Sprachsteuerung erleben.

Damit die Sprachassistenten zum wirklich intelligenten Helfer werden, müsste die KI die Pfade der Nutzer analysieren können und daraus kybernetische Schlussfolgerungen ziehen. Was meint der Nutzer, wenn er von einem „großen Auto“ spricht? Und selbst dann würde noch immer eine herausragende Eigenschaft fehlen: Die Systeme können zwar inzwischen schon Emotionen erkennen – aber noch nicht adäquat darauf reagieren. Vorgefertigte Textbausteine helfen nicht, wenn ein Mensch verärgert eine Hotline anruft.

„Solange ein Mensch lösungsorientiert vorgeht, können Sprachassistenten prinzipiell funktionieren. Wenn Menschen aber ▶

► Die Callcenter-Branche macht derzeit vor, wie Bestellungen per Sprachassistenten funktionieren können – und wo die Grenzen liegen.

emotionale Ansprache suchen – etwa weil sie verärgert sind – muss auch ein Mensch am Apparat sein“, erklärt Stockmann. Dabei ist Ärger nicht das einzige Gefühl, das eine Rolle spielt. Gar nicht so selten rufen Menschen bei der Bestellhotline eines Händlers an, weil sie Kontakt suchen.

Welche Auswirkung Sprachassistenten haben

Auch wenn sich die technischen Voraussetzungen verbessern, ist der Einsatz von sprachgeführten Systemen also keineswegs so zwangsläufig, wie es zunächst scheint. Die Gretchenfrage lautet: Wie viele Geschäftsvorfälle lassen sich in einem Callcenter schon jetzt automatisieren? Und: Sind Hintergrundprozesse schon soweit, dass sie sich leicht automatisieren lassen?

In der jetzigen Situation ist daher ein maßvolles Vorgehen angeraten. Folgende Voraussetzungen sollten Sie prüfen:

1. Automatisierung: Erlauben die Prozesse im Callcenter bereits eine Automatisierung?

2. Anwendungsumgebung: Ist ein Sprachassistenzsystem sinnvoll? Die Zielgruppe oder die Dienstleistung können dagegen sprechen. Dagegen sprechen beispielsweise

- Kunden, die auf persönlichen Kontakt Wert legen.
- Händler, die einen intensiven Kundenkontakt pflegen wollen.

3. Emotionaler Faktor: Steht eine emotionale oder eine lösungsorientierte Kundenansprache im Vordergrund? Selbst wenn alle vorhergehenden Prüfungen positiv ausfallen, verbietet sich

der Einsatz von Maschinen in emotional aufgeladenen Fällen. Eine Beschwerdeline kann nicht von Maschinen besetzt sein – kein Mensch würde sich ernst genommen fühlen.

Anders sieht es bei standardisierten, technischen Abläufen aus. Wer seine Bankverbindung oder Adresse ändern will, eine Reise umbuchen oder ein Taxi bestellen will, kommt oft mit einer Maschine klar.

Rationalisierung statt Automatisierung

In allen anderen Fällen bieten sich hybride Lösungen an. Sogenannte IVR-Systeme (Interactive Voice Response) helfen, Anrufer vorzuselektieren, das genaue Anliegen abzufragen oder Vorklärungen vorzunehmen (etwa: Abfrage der Kunden- oder Versicherungsnummer). Solche Systeme sind keine neue Erfindung, Sprachassistenten, die natürliche Sprache verstehen, können sie aber attraktiver machen.

Noch deutlich lassen sich die Kosten reduzieren, wenn ein paralleler Ansatz gewählt wird, in dem zwar einerseits ein vollständig automatisierter Prozess angeboten wird, der Nutzer aber jederzeit auf einen menschlichen Mitarbeiter zurückspringen kann. Allerdings muss bei diesem Ansatz auch eine hinreichende Aussicht darauf bestehen, dass ein bestimmter Anteil der Kunden das automatisierte Verfahren auch abschließen wird. Greift dagegen die überwiegende Mehrheit auf die menschliche Unterstützung zurück, wurde lediglich ein zusätzlicher Kostenblock aufgebaut – und der Kundenfrust gesteigert. In diesen Fällen ist es günstiger, die Mitarbeiter mittels automatisierter Vorselektion und Vorklärung zu entlasten.

Automatisierung hinterlässt Spuren

Auf diese Weise werden in drei Jahren rund 160.000 der derzeit 500.000 Arbeitsplätze in Callcentern direkt oder indirekt beeinflusst sein, schätzt Stockmann. Das bedeutet aber nicht, dass damit automatisch ein radikaler Stellenabbau verbunden ist. Zum einen, weil derzeit ohnehin etwa 20.000 Stellen nicht besetzt werden können. Diese werden als erstes von der Automatisierungswelle aufgefangen. Zum zweiten werden vor allem Routinetätigkeiten durch Maschinen erledigt. Etwa die erwähnten Vorklärungsfragen oder bei der Prozessabwicklung im Nachklapp des Gesprächs. Dadurch werden Kapazitäten für den wesentlichen Teil des Gesprächs frei.

Mittels künstlicher Intelligenz sind noch genauere Selektionen möglich. Kunden, die kürzlich ein Handy gekauft haben, rufen bei der Hotline meist wegen Vertragsfragen an, nach dem Rechnungsversand sind es eher Tariffragen. Ist die Anrufererkennung mit den richtigen Daten verknüpft, können bereits im Vorfeld solche Schlüsse gezogen werden.

Grenzen der Automatisierung

Menschen werden aber in absehbarer Zeit nicht durch Sprachassistenten ersetzt werden können. „Das ist nicht nur eine technologische, sondern auch eine soziale Frage“ sagt Stockmann. Selbst wenn durch Siri und Co. die Akzeptanz für Sprachassistenten steigt, werden soziale Faktoren einen flächendeckenden Einsatz verhindern. „Die Frage lautet: Fühlen sich Menschen von Maschinen ernst genommen“, meint Stockmann und liefert die Antwort gleich mit. „Auf emotionaler Ebene wird das nicht funktionieren.“ ■

iBusiness Dossier – Themenbackground der iBusiness-Redaktion ISSN 1861-8197

Jahresbezug: Der Bezug von iBusiness Dossier ist in der iBusiness-Premium-Mitgliedschaft enthalten.

Jahresbeitrag: 390,- Euro, Preise inkl. Porto, zzgl. MwSt. iBusiness Dossier dient ausschließlich der Hintergrundinformation für iBusiness-Premium-Mitglieder.

Freiweisungen sind aus grundsätzlichen Überlegungen nicht möglich. Nachdruck und Vervielfältigung in jeder Form, auch auszugsweise nur mit schriftlicher Genehmigung.

Unberechtigter Nachdruck oder Vervielfältigung ziehen Schadensersatzforderungen nach sich.

Mitteilung gemäß § 8, Artikel 3 des Bayerischen Pressegesetzes

Inhaber des HighText-Verlag Graf und Treplin OHG sind zu jeweils 50 Prozent Joachim Graf (verantwortlich für Redaktion und Anzeigen) und Daniel Treplin (beide Journalisten, München);

Handelsregister München HR A 72216

Anschrift Verlag und Redaktion

HighText Verlag
Wilhelm-Riehl-Str. 13
80687 München
Tel.: (089) 57 83 87-0
Fax: (089) 57 83 87 99
E-Mail: info@ibusiness.de
URL: http://www.ibusiness.de

Redaktion@ibusiness.de
Joachim Graf, Dominik Gröllmann,
Sebastian Halm, Susan Rönisch, Daniel Treplin

Buchhaltung@ibusiness.de
Daniela Seitz

Anzeigen@ibusiness.de
Nicos Fucidis, Benjamin Karg, Dagmar Kraux
Disposition: Anna Strommer

Leserservice@ibusiness.de
und Vertrieb@ibusiness.de
Bianka Matthes (bma@ibusiness.de)

Technik@ibusiness.de
Sven Noherr

Grafik@ibusiness.de
Gertrud Graf, Monika Herbig

Druck: Holzmann Druck GmbH & Co. KG,
Gewerbestraße 2,
86825 Bad Wörishofen

Titelfoto: PublicDomainPictures/Pixabay

Das komplette Dossier auch
als klickbares PDF:
www.ibusiness.de/dossier

 iBusiness





Dominik Grollmann ist Redakteur des iBusiness-Redaktionsteams

Finde die Nische! Das wahre kommerzielle Potenzial von Alexa und Co.

Foto: HighText Verlag

Auch in diesem Jahr werden wohl wieder Tausende von intelligenten Lautsprechern verkauft. Doch trotz der riesigen Gerätebasis tut sich die Branche extrem schwer, einen kommerziellen Nutzen für die Geräte zu finden. Welche kommerziellen Szenarien funktionieren.

Digitale Sprachassistenten und intelligente Lautsprecher sind einer der wichtigsten Technologietrends des Jahres – vielleicht neben künstlicher Intelligenz, Augmented und Virtual Reality. Eine wahre Flut an Geräten ist 2017 auf den Markt gekommen (oder zumindest angekündigt worden). Schätzungen zufolge steht in den USA bereits in jedem zehnten Haushalt ein Alexa-Gerät. In Deutschland haben nach einer Statista-Umfrage bereits 28 Millionen Menschen Erfahrung mit Siri & Co. gesammelt. Und smarte Lautsprecher dürften auch 2018 zum Renner werden.

Angesichts dieser Aussichten kann es nicht verwundern, dass sich Analysten weltweit mit der Frage auseinandersetzen, wie sich diese Entwicklung für Werbung und Onlinehandel nutzen lässt. Der Werbeexperte Unruly erklärte auf der Digitalmesse Dmexco 2017 die Geräte und die damit verbundenen Einkaufsszenarien gar zum Messetrend überhaupt.

Die wichtigste Zukunftsaufgabe für Marken und Händler ist es, einen Zugang zu den smarten Devices der Konsumenten zu finden, erklärte Unruly-Futurist Simon

Gosling den Messebesuchern gebetsmühlenartig. Seine Szenarien: Wer sich in Zukunft einen Wagen kaufen will, wird die Innenausstattung in einer perfekten VR-Simulation konfigurieren – natürlich im eigenen Wohnzimmer und auf Zuruf an den smarten Agenten. Wem morgens die Rasierklingen ausgehen, wird noch im Bad per Sprachassistent neue bestellen.

Und wenn eine neue Klingenmarke auf den Markt kommt, wird der digitale Helfer freundlich darauf hinweisen und fragen, ob man nicht einmal diese ausprobieren möchte. *„Bislang müssen Marken und Händler versuchen, in das Relevant Set der Verbraucher vorzudringen“*, erklärt Gosling. *„In Zukunft wird es darum gehen, das Relevant Set der digitalen Assistenten zu erreichen.“*

Nicht zum ersten Mal springt die Branche so euphorisiert auf einen Trend auf. Schließlich kann Technologie das Leben verändern – wie Internet, GPS oder Smartphones beweisen. Oder spurlos vorüber gehen – wie es iBeacons, Dash-Buttons, QR-Codes oder der berühmte Internet-Kühlschrank ►

► Am Erfolg von Sprachassistenten ist kaum zu zweifeln – kommerziell ist das Potenzial aber schwer zu heben.

zeigten. Welchen der beiden Wege Einkaufszenarien mit digitalen Assistenten nehmen werden ist bislang kaum abzusehen. Klar ist nur, dass nicht nur der britische Werbespezialist Unruly auf die Macht der Computerstimmen setzt.

Allein für Alexa wurden inzwischen mehr als 15.000 Skills entwickelt. Bei den meisten davon ist zwar kein Monetarisierungsziel erkennbar. Es gibt aber auch einige rein kommerzielle Ansätze. In Deutschland versuchen sich etwa die Supermarktketten Real und Hit sowie der Online-Marktplatz Allyouneed daran.

Taugt Alexa als Shopping-Assistent?

Dass sich die Entwicklung bislang fast ausschließlich auf Alexa konzentriert, ist kein Wunder. Der Amazon-Assistent ist für kommerzielle Anbieter am vielversprechendsten. Alexa ist der einzige Assistent, der mittels Schnittstelle um beliebige Anwendungen („Skills“) erweitert werden kann. Auf diese Weise erhält das System neue Fähigkeiten und kann auf Zuruf beispielsweise Sonderangebote des Supermarktes nebenan abspulen („Alexa, frage Real nach den Angeboten des

Tages“). Andere Assistenten lassen sich nicht so gezielt mit Inhalten und Aktionen füttern und würden auf die entsprechende Anfrage lediglich das Ergebnis einer Websuche präsentieren.

Doch unter allen verfügbaren Skills scheint ausgerechnet die Rubrik „Shopping“ eine der am wenigsten attraktivsten zu sein: Von den mehr als 2.500 in Deutschland erhältlichen Erweiterungen stammen gerade einmal 19 aus dieser Rubrik. Natürlich: Das kann auch daran liegen, dass eine der Grundfertigkeiten von Alexa darin besteht, Bestellungen bei Amazon auszuführen. Es ist also gar nicht nötig, einen Skill zu aktivieren, um mit Alexa einzukaufen. Außerdem sagt die Zahl der verfügbaren Skills auch nichts über die aktivierte Basis aus.

Amazon selbst veröffentlicht keine Statistiken über die Nutzungsraten einzelner Anwendungen. Noch weniger gibt es Zahlen zu den mit Alexa durchgeführten Einkäufen oder gar den daraus resultierenden Umsätzen. Aus Nutzerbefragungen lässt sich allerdings ein recht detailliertes Bild gewinnen. Zuletzt hat Pricewaterhouse Coopers im Juni 2017 eine bevölkerungsrepräsentative Umfrage unter

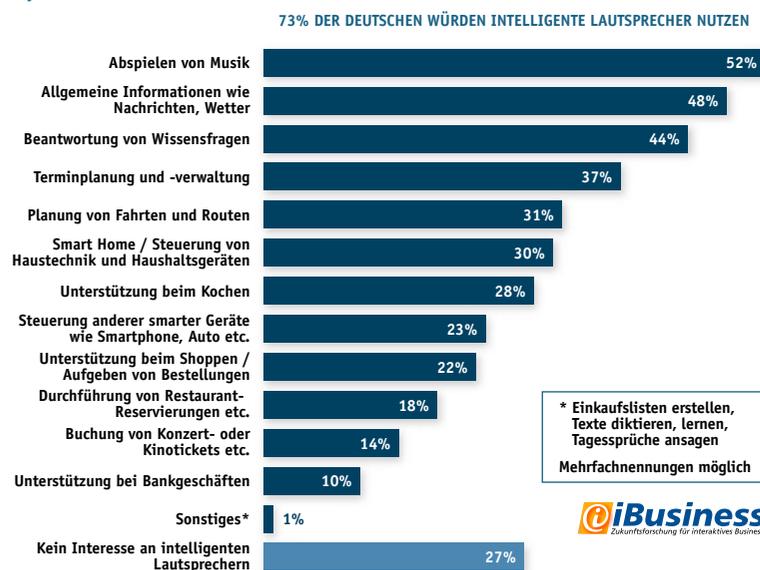
deutschen Bundesbürgern ab 18 Jahren zu ihrer Einstellung gegenüber digitalen Assistenten durchgeführt. Die Ergebnisse in der Zusammenfassung:

- **Hohe Bekanntheit, konkretes Interesse, geringe Verbreitung:** Die Bekanntheit von Sprachassistenten ist in Deutschland sehr hoch. Amazons Alexa ist mit 78 Prozent der bekannteste Vertreter. Die Verbreitung ist aber gering. Nur fünf Prozent besitzen ein Gerät mit Alexa. Das größte Interesse herrscht am Smart Speaker Google Home, immerhin jeder fünfte Deutsche gibt an, sich dafür erwärmen zu können.
- **Große Akzeptanz trotz Datenschutzbedenken:** Fast drei Viertel der Befragten können sich generell vorstellen, intelligente Lautsprecher zu nutzen. Die größten Bedenken herrschen hinsichtlich der Sicherheit. Die Gefahr von Datendiebstahl sowie die Manipulation oder der Missbrauch durch Hacker erscheinen 70 Prozent der Befragten „sehr groß“ oder „eher groß“. 69 Prozent fürchten, für Amazon, Apple oder Google zum gläsernen Kunden zu werden.

- **Anwender wollen Entertainment und Wissen:** Gefragt ist vor allem ein hochwertiger Ersatz für das Radio. Die meisten Nutzer wollen mit den Geräten Musik abspielen und auf Abruf Nachrichten und Wetterberichte hören. Es folgen allgemeine Wissensfragen sowie persönliche Assistenzfunktionen (Termin-, Routenplanung). Knapp ein Drittel wollen ihr Smart Home steuern. Konsumaufgaben (etwa Bestellungen, das Erstellen von Einkaufslisten oder Tischreservierungen, Ticketbuchungen) liegen weit abgeschlagen auf den hinteren Rängen (siehe Chart Seite 8).

Andere Statistiken kommen zu ähnlichen Ergebnissen. Im März 2017 fragte das Marktforschungsunternehmen Norstat, für welche Zwecke die Deutschen digitale Sprachassistenten nutzen würden, wenn sie denn welche hätten. Die Top-3-Einsatzgebiete: Suchmaschinen befragen, Terminerinnerungen und

Interesse: Diese Anwendungen wünschen sich Anwender von Sprachassistenten



Quelle: comScore; Grafik: HighText Verlag

Wettervorhersagen abhören. Ganz am Ende der Liste: „Taxi bestellen“ und „Waren bestellen“ mit je gerade einmal 14 Prozent positiven Nennungen.

An dem Interesse ändert sich auch nichts, wenn die Anwender die entsprechenden Funktionen ausprobieren konnten. Der Marktforscher Comscore hat im ersten Quartal 2017 amerikanische Besitzer von vernetzten Lautsprechern nach ihren Nutzungsgewohnheiten befragt. Ergebnis: 60 Prozent stellen allgemeine Fragen, 57 Prozent hören den Wetterbericht, 54 Prozent Musik. Und E-Commerce? Die Funktionen „Produkte bestellen“ (11 Prozent) und „Lebensmittel/Essen bestellen“ (8 Prozent) landeten auf den hintersten Rängen 11 und 12.

Bleibt festzuhalten: Die meisten Anwender wollen keine Produkte per Sprachassistenten bestellen und die wenigsten Anwender tun dies in der Praxis. Aber vielleicht bedingen sich die Ergebnisse gegenseitig? Ist es nur eine Frage der Gewöhnung und die Nutzer würden eifrig Sprach-Shopper, wenn sie die Funktion nur einmal probieren würden?

Für diese These gibt es keinen Beleg – im Gegenteil: Im Mai 2016 wollte Experian von amerikanischen Alexa-Nutzern wissen, welche Funktionen sie bereits ausprobiert haben und welche sie anschließend regelmäßig nutzen. Ergebnis: Erwartungsgemäß haben sich vergleichsweise wenige (45,3 Prozent) für die Shopping-Funktion interessiert. Noch viel weniger blieben ihr aber treu. Nur zehn Prozent kauften anschließend „regelmäßig“ mit Alexa ein – das schlechteste Ergebnis in der gesamten Umfrage. Selbst Funktionen, die anfangs von viel weniger Nutzern ausprobiert wurden, kamen schließlich auf deutlich mehr regelmäßige Anwender.

Shopping-Hype nur herbeigeredet

So eindrucksvoll manche Präsentation auf der Dmexco war, so ernüchternd fällt das Fazit nach Auswertung der statistischen Parameter aus.

- **Anfang des Hype-Cycles:** Smarte Lautsprecher und die enthaltenen Assistenten haben eine enorme Bekanntheit, das Interesse daran ist hoch und die Verbreitung gering. Sie stehen damit am Anfang eines klassischen Hype-Cycles und die Absatzprognosen für die kommenden Monate dürften rosiger sein. Typisch ist aber auch, dass auf die Anfangseuphorie eine Ernüchterung folgt, bevor sich die Anwendung im Alltag verfestigt. Diese Phase steht intelligenten Lautsprechern noch bevor.

- **Kein Shopping-Hype:** Für ECommerce-Szenarien gilt diese Entwicklung dagegen nicht. Ein Hype ist gar nicht erst zu erkennen. Es gibt weder eine anfängliche Euphorie, noch Nutzungsszenarien, die sich im Alltag herausbilden. Im Gegenteil: Anwender, die die Funktion einmal ausprobiert haben, verwenden sie meist anschließend nicht mehr.

Diese Einschränkungen gelten nicht nur für ECommerce-Anwendungen. Durchweg alle Funktionen, die mit dem Einkauf oder Transaktionen zu tun haben, erhalten schlechte Noten.

- Pflegen von Einkaufslisten,
- Essensbestellung (z.B. Pizzadienst),
- Tickets für Konzert, Kino und Theater,

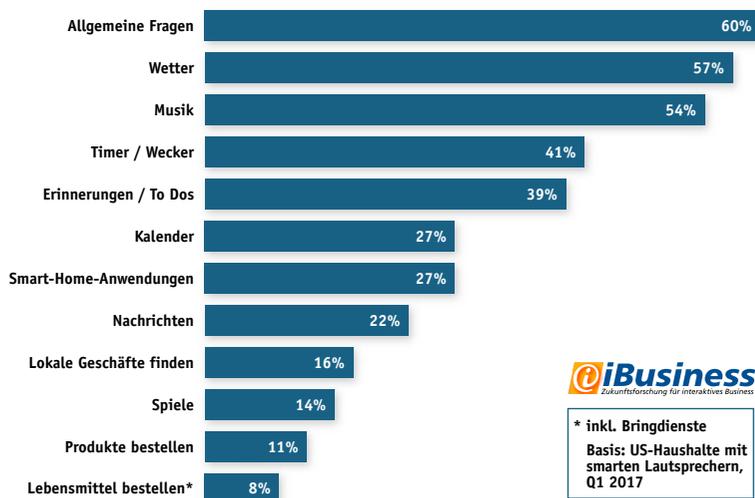
- Taxiruf
- und selbst die Reservierung in einem Restaurant.

Chance liegt in der Nische

Alexa und Co. werden so schnell keine Shopping-Revolution auslösen. Völlig vernachlässigen sollten Händler die Systeme aber trotzdem nicht. Selbst wenn nur jeder zehnte Nutzer seinen Sprachassistenten regelmäßig zum Einkauf verwendet, kann dies eine interessante Zielgruppe werden. Zudem ist es möglich, dass zwar insgesamt nur ein kleiner Teil der Verkäufe über Assistenten abgewickelt wird, in bestimmten Situationen oder Produktkategorien jedoch der Großteil. Anders gesagt: Es kann sich eine lukrative Nische auf tun.

Zunächst muss sich jeder Anbieter bewusst sein, dass er einen juristischen Graubereich betritt. „Wir sehen bei Alexa-Bestellungen die gleiche Problematik wie beim Dash-Button“, meint Carsten Fölisch, Jurist des Shopzertifizierers Trusted Shops. Den Kunden stellen diese Einschränkungen allerdings nicht schlechter. Denn dieser hat im Zweifel nicht nur ein Widerrufsrecht, sondern kann sich sogar darauf berufen, dass gar kein Vertrag zustande gekommen ist. ▶

Nutzung: Diese Sprachassistenten-Anwendungen werden tatsächlich genutzt



Quelle: comScore; Grafik: HighText Verlag

iBusiness
Zukunftsforschung für interaktives Business

* inkl. Bringdienste
Basis: US-Haushalte mit smarten Lautsprechern, Q1 2017

Schwieriger ist die Situation dagegen für den Händler. „Wir glauben, dass die gesetzlichen Informationspflichten im Fernabsatz und im E-Commerce weder beim Dash-Button noch bei Käufen über Alexa eingehalten werden“, sagt Föhlisch.

Juristisch stellen Transaktionen über Sprachassistenten noch immer ein Wagnis dar, wenngleich ein kalkulierbares. Weniger Einfluss kann der Händler dagegen auf die Einkaufssituation nehmen. Vier Szenarien lassen sich unterscheiden – mit unterschiedlichen Erfolgsaussichten für den Händler.

ECommerce-Szenarien für digitale Assistenten:

1. Komplexe Produkte

Die meisten, selbst vermeintlich einfache ECommerce-Einkäufe sind für einen Sprachassistenten ungeeignet. Die genaue Auswahl erfolgt zumeist am Bildschirm und Bestellwerte liegen meist über zehn Euro. Selbst bei vermeintlich einfachen Dingen wie einem USB-Stick sind am Ende noch zahlreiche Parameter (Preis, Marke, Kapazität, Größe, Standard) zu klären. Die Hürden:

- Ein mehrstufiges Voice-User-Interface (VUI), das Rückfragen beherrscht, wäre nötig. Kein Sprachassistent beherrscht dies derzeit zufriedenstellend.
- Eine echte Assistentenfunktion („Kaufe eine neue Autobatterie“) oder gar Stilberatung scheitert an mangelndem Wissen und Fähigkeiten. Die KI hinter der Sprachverarbeitung ist wenig leistungsfähig.
- Selbst wenn diese technischen Hürden gelöst wären, müsste der Nutzer darauf vertrauen, dass der Sprachassistent eine Wahl im Sinne des Benutzers trifft.
- Funktionieren kann das Szenario nur, wenn Produkt und Preis klar definiert sind. Zum Beispiel bei einem Musik-Album oder einem Buch.

Bewertung: Weitgehend ungeeignet.

2. Convenience-Produkte

Am besten könnte Assi-Commerce bei Produkten funktionieren, die gerne beim Lebenspartner bestellt werden („Schatz, bringst du noch Milch mit“). Das Produkt ist einfach, ein Fehlkauf praktisch ausgeschlossen. Der Partner weiß schon, was gemeint ist (fettarme Milch). Die Hürden:

- Die meisten Produkte sind niedrigpreisig und das Sortiment für E-Commerce nicht geeignet.
- Die Präsenz des stationären Einzelhandels ist in diesen Segmenten enorm, die Konkurrenz groß.
- Digitalen Assistenten fehlt Wissen („Kaufe Batterien ... die kleinen, für die Fernbedienung.“) Am ehesten lassen sich Produkte, die regelmäßig verbraucht werden, nachordern. Spontankäufe sind ausgeschlossen.
- Entscheidend ist technologischer Fortschritt. Der Assistent muss mehr Wissen anhäufen, um nützlich zu sein. Welche Glühbirne passt in die Wohnzimmerlampe? Welcher Scheibenwischer ans Auto?
- Fraglich bleibt, ob Kunden dieses Wissen weitergeben möchten. Datenschutzbedenken zählen schließlich zu den größten Befürchtungen.

Bewertung: Vor allem betriebswirtschaftlich herausfordernd.

3. Kaufentscheidungsprozess

Marken und Händler können die Geräte nutzen, um sich im Kaufentscheidungsprozess zu positionieren. Rezeptideen vom Supermarkt sind eine naheliegende Idee. Die zunehmende Vernetzung mehrerer Geräte (Smartphone, Laptop, TV, Smart Speaker) erlaubt es, digitale Assistenten auch in komplexere Kaufentscheidungen einzubinden. Die Aussichten:

- Derzeit ist der Einfluss von Sprachassistenten auf Kaufentscheidungen eher gering. Das VUI ist ungenügend.

- Am ehesten sind einfache Suchanfragen für lokale Ladengeschäfte denkbar („Wo ist der nächste Schlüsseldienst?“). Allerdings ist das Potenzial begrenzt.

- In Zukunft könnten Sprachassistenten eine wichtigere Rolle übernehmen. Sie könnten zentrales Ein- und Ausgabegerät einer cloudbasierten Intelligenz werden. Alexa, Siri und Cortana halten die lose vernetzten Geräte (Entertainment-System im Auto, Smartphone, PC, TV-Set und Smart Home) zusammen.

- Auf Kommando („Zeige meine Suchergebnisse auf dem Fernseher“) würden die vernetzten Geräte mit den gespeicherten Inhalten beschickt.

Das Szenario ist zukunftsweisend, der Zeithorizont allerdings noch relativ groß. Im Vorteil sind Hersteller, die über eine breite Basis an Geräten und Diensten verfügen. Allen voran wäre hier Google, gefolgt von Apple, zu nennen.

Bewertung: Langfristiges Disruptionspotenzial.

4. Digitale Services

Kurzfristig sind dagegen digitale Produkte am ehesten für Assistenten-Commerce geeignet. Die Aussichten:

- An erster Stelle sind Audio-Inhalte und Abos von Streaming-Diensten denkbar.

- Mit den Skills hat Amazon gezeigt, dass auch andere Erweiterungen praktikabel sind.

- Noch fehlt ein kommerziell attraktives Ökosystem („Skill-Store“). Konkurrenten wie Apple oder Google könnten Bewegung in den Markt bringen.

- Die kommerziell attraktivsten Dienste (Medien, Streaming Audio) werden von den Inhabern der Ökosysteme selbst angeboten. Die Eintrittsbarrieren sind entsprechend hoch.

Bewertung: Hohes Disruptionspotenzial, ebenso hohe Eintrittsbarrieren. ■