



Beyond Touch – Voice Commerce 2030:

Wie Voice-assisted Interfaces den Handel
in Europa revolutionieren werden

Einleitung	04
Voice-assisted Interfaces im Überblick	08
Überblick und relevante Dimensionen	10
Die Technologie hinter den Voice Assistants	16
Einblicke in den Voice-Assistant-Markt	22
Forschungsbedarf: Die Auswirkungen von Voice-assisted Interfaces auf den Handel	29
Die Zukunft von Voice-assisted Interfaces	30
Vier mögliche Szenarien zur Zukunft von Voice Commerce	34
Auswirkungen der Szenarien auf Kunden, Händler und Hersteller	44
Kritische Treiber und Barrieren für den Erfolg von Technologiefirmen	50
Die Champions des Voice-Marktes	60
Handlungsempfehlungen – Voice-assisted Interfaces als Chance nutzen	64
Conclusio	70
Ansprechpartner	76

Einleitung

Herausforderung Voice-assisted Interfaces – die nächste Revolution für den Handel nach Touch.

Die Digitalisierung hat längst zu tiefgreifenden Veränderungen im Handel geführt. Nach der durch den Touchscreen veränderten Interaktion steht nun mit dem sprachgesteuerten Einkauf die nächste große Revolution bevor. Schon heute dringen intelligente Sprachassistenten wie Amazons Alexa, Google Assistant oder Siri von Apple in immer mehr Lebensbereiche ein, zu Hause, im Auto oder bei der Arbeit. Dabei werden Voice-assisted Interfaces nicht nur für den Abruf von Informationen, zur Gerätesteuerung oder zur Unterhaltung genutzt, sondern vor allem in den USA auch zunehmend für den sprachgesteuerten Einkauf. Zwar steckt Voice Commerce in Europa noch in den Anfängen, doch hat dieser neue Trend auch hier eindeutig das Potenzial, den Handel ein weiteres Mal zu revolutionieren.





Sind Voice-assisted Interfaces die nächste Revolution nach Touch, die den Handel grundlegend verändern wird? Dieser Frage sind die Vertreter führender Handelsunternehmen aus Versand, Supermarkt und Discount, Google und wir bei Monitor Deloitte im Rahmen dieser Studie gemeinsam nachgegangen. Eins können wir direkt vorwegnehmen: Wenn die Rahmenbedingungen stimmen, wird das sprachgesteuerte Einkaufen einen großen Einfluss auf das Kaufverhalten der Verbraucher haben und weitreichende Veränderungen für Hersteller, Händler und Verbraucher mit sich bringen.

Experten aus Handel und Industrie sind sich einig, dass Voice Assistants und sprachgesteuertes Einkaufen einen tiefgreifenden Einfluss auf das bisherige Wirtschaftsleben haben und die Marktteilnehmer unter Druck setzen werden. „Das ist der nächste große Sprung. Wir glauben, dass diese Technologie viele Teile des Handels grundlegend verändern wird“, sagt Jan Kunath, stellvertretender Vorstandsvorsitzender von Rewe.

Insbesondere die Dominanz der großen Technologiekonzerne wie Amazon, Google oder Apple kann gravierende Auswirkungen auf Händler und Hersteller haben. „Sprachassistenten entscheiden, welches Produkt von welchem Händler bestellt werden soll. Langfristig werden also die Selektions- und Marketingprozesse von Sprachassistenten einen wichtigen Faktor für den Erfolg und Misserfolg von Produkten und Händlern darstellen“, erklärt Stephan Tromp, stellvertretender HDE-Hauptgeschäftsführer.

Doch wie genau werden Voice Assistants das Einkaufen in Europa im Jahr 2030 beeinflussen? Dieser Frage geht die vorliegende Studie „Beyond Touch“ auf den Grund. Mithilfe von Artificial-Intelligence-Analysen und der Mitarbeit zahlreicher Experten wirft Monitor Deloitte einen Blick in die Zukunft. Die Studie beleuchtet die wichtigsten Chancen und Risiken des Voice Commerce und gibt erfolgskritische Handlungsempfehlungen, wie sich Hersteller und Händler optimal auf die bevorstehenden Veränderungen vorbereiten können. Denn eines ist schon heute sicher: Voice-assisted Interfaces werden das Shopping Erlebnis der Zukunft maßgeblich verändern und sowohl Hersteller als auch Händler sollten sich intensiv mit diesem Thema beschäftigen.

Doch wie bei jeder neuen Technologie sind auch bei Voice-assisted Interfaces weitreichende Prognosen von großen Unsicherheiten geprägt. Denn wohin genau sich der Markt entwickeln wird, hängt von dem Einfluss verschiedenster Treiber und Barrieren ab. Daraus ergeben sich folgende zentrale Fragen:

Wird Voice lediglich ein weiterer Interaktionskanal unter vielen bleiben oder entwickelt sich daraus mehr? Können deutsche Marktteilnehmer den Vorsprung amerikanischer Konzerne überhaupt noch aufholen? Werden die Konsumenten bereit

sein, ihre Privatsphäre zugunsten von mehr Bequemlichkeit aufzugeben? Und wie stark wird der Staat regulatorisch eingreifen?

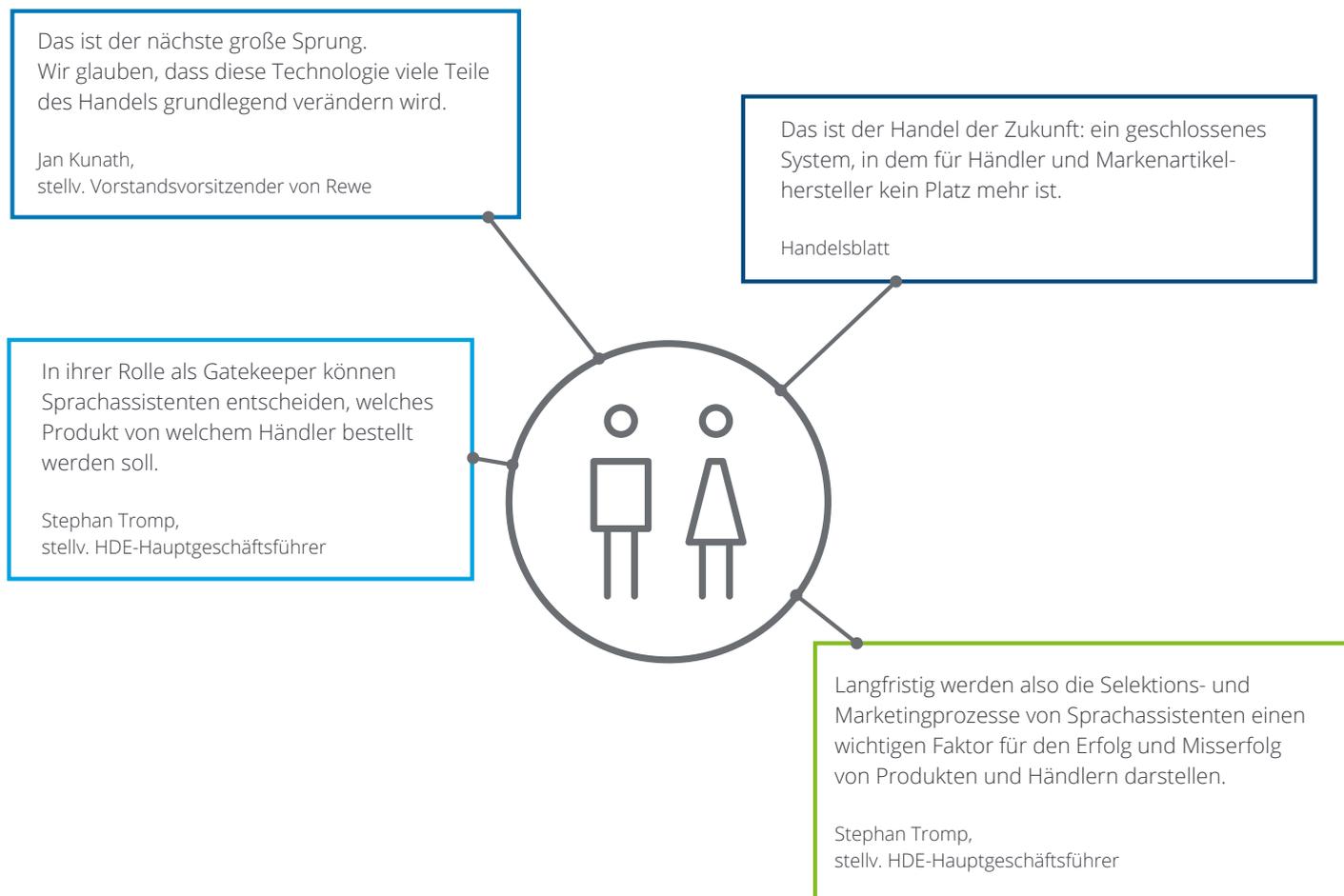
Auf Basis der wichtigsten Treiber und Barrieren des Voice Commerce hat Monitor Deloitte vier extreme, aber wahrscheinliche Zukunftsszenarien entwickelt: World of Aggregators, Old Europe, The new Internet und Age of heterogenic Alliances. Daraus leiten sich verschiedene Konsequenzen für Händler, Hersteller und Kunden ab.

Entscheidend für den Erfolg von Voice Assistants werden jedoch vor allem Koope-

rationen zwischen den Marktteilnehmern und eine Offenheit der Systeme sowie die Integration mit anderen Technologien sein.

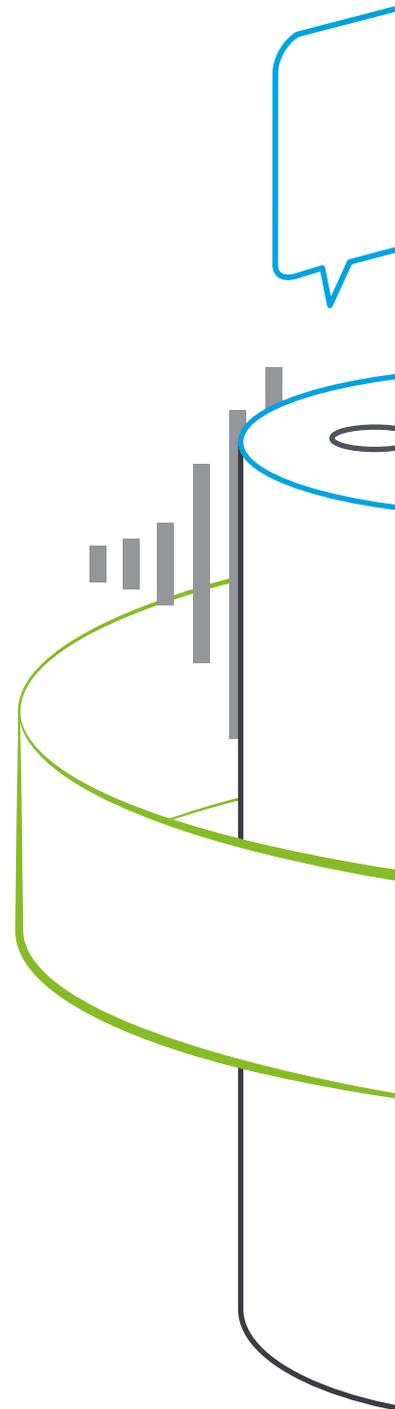
Europa steht im Voice-Commerce-Markt erst am Anfang und lässt den Akteuren noch Freiheiten bei der Entwicklung der Chancen. Deswegen sind alle Marktteilnehmer gefordert, jetzt individuelle Strategien zu entwickeln. Die Studie gibt wichtige Handlungsempfehlungen für Händler und Konsumgüterhersteller, um die Herausforderungen der Zukunft zu meistern und die Potenziale zu heben.

Abb. 1 – Expertenmeinungen zum Impact von Voice Assistants auf den Handel



Voice-assisted Interfaces im Überblick

Voice-assisted Interfaces sind intelligente Schnittstellen, die auf Sprachbefehl den Nutzer bei Aufgaben unterstützen. Durch die Weiterentwicklung von Artificial Intelligence (A.I.) wurden aus einfachen Assistenten intelligente Systeme, welche komplexe Aufgabenstellungen lösen können und stetig dazulernen. Dabei wird die menschliche Komponente übernommen und es ist eine flexible Kommunikation über verschiedene Medien möglich. Im engeren Sinne und im allgemeinen Sprachgebrauch spricht man in diesem Zusammenhang oft von Voice Assistants, geprägt durch die hohe Präsenz von Amazons Alexa, Google Assistant und Apples Siri. Der Begriff Voice Assistant soll neben der Kurzform Voice auch im Folgenden verwendet werden.





Überblick und relevante Dimensionen

Voice Assistants im engeren Sinne sind persönliche, intelligente und digitale Assistenten, die mithilfe von gesprochenen Anweisungen bestimmte Aktionen ausführen. Sie nutzen Spracherkennung, Sprachanalyse, -synthese und Artificial Intelligence, um die Anwender bei verschiedensten Aufgaben zu unterstützen. Voice Assistants ermöglichen dem Nutzer, mit Geräten über natürliche Sprache zu kommunizieren, ohne dabei einen Browser zu öffnen, die Hände einzusetzen oder eine Frage einzutippen.

Die Anwendungsgebiete von Voice Assistants sind vielfältig und reichen von Abruf von Informationen und Gerätesteuerung über Unterhaltung bis hin zum Shopping. Dazu zählen etwa Musik hören, Wiedergabe von Informationen, persönliche Assistenz, Steuerung von Smart-Home-Geräten, aber auch zunehmend die Möglichkeit, einzukaufen.

Durch Cloud Computing lassen sich die mit dem Internet verbundenen Systeme in verschiedenste Geräte integrieren, wie zum Beispiel in Smart Speaker, Smart Phones oder auch in Infotainmentsysteme von Fahrzeugen. Genutzt werden Voice Assistants vor allem zu Hause oder im Auto, aber auch unterwegs und bei der Arbeit.

Die bekanntesten Voice Assistants sind Alexa von Amazon, der Google Assistant und Siri von Apple. Diese werden unter anderem auf den zugehörigen Smart Speakers wie Echo (Amazon), Google Home oder auf dem HomePod (Apple) verwendet. Allein 2017 betrug der globale Smart Speaker Umsatz 4,4 Mrd. USD. Führend in diesem Segment ist Amazon mit einem Marktanteil von 62 Prozent, gefolgt von Google mit von 25 Prozent. Während die USA weltweit Vorreiter bei der Nutzung und den Anwendungsmöglichkeiten von Voice Assistance sind, steht der europäische Markt erst noch am Anfang.

Abb. 2 – Zusammenfassung Voice Assistants

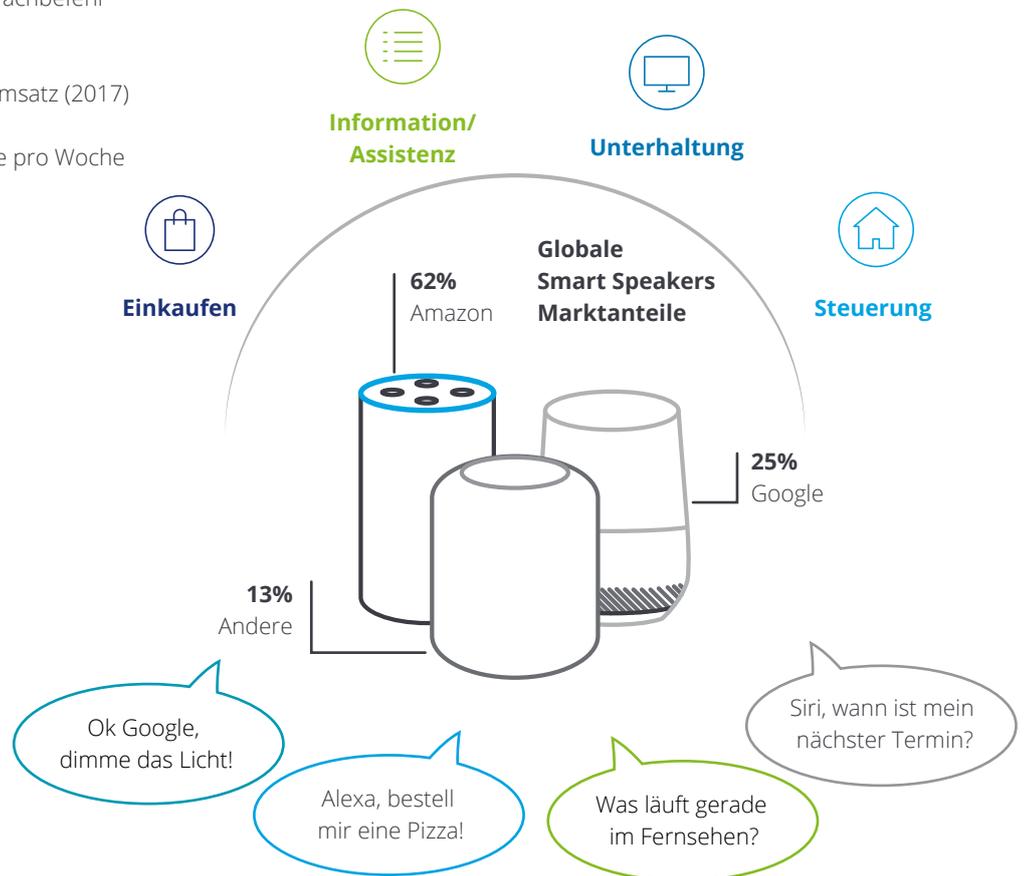
Voice Assistant

ist ein persönlicher Assistent, der auf Sprachbefehl Aufgaben des Alltags übernimmt.

4,4 Mrd. USD globaler Smart Speakers Umsatz (2017)

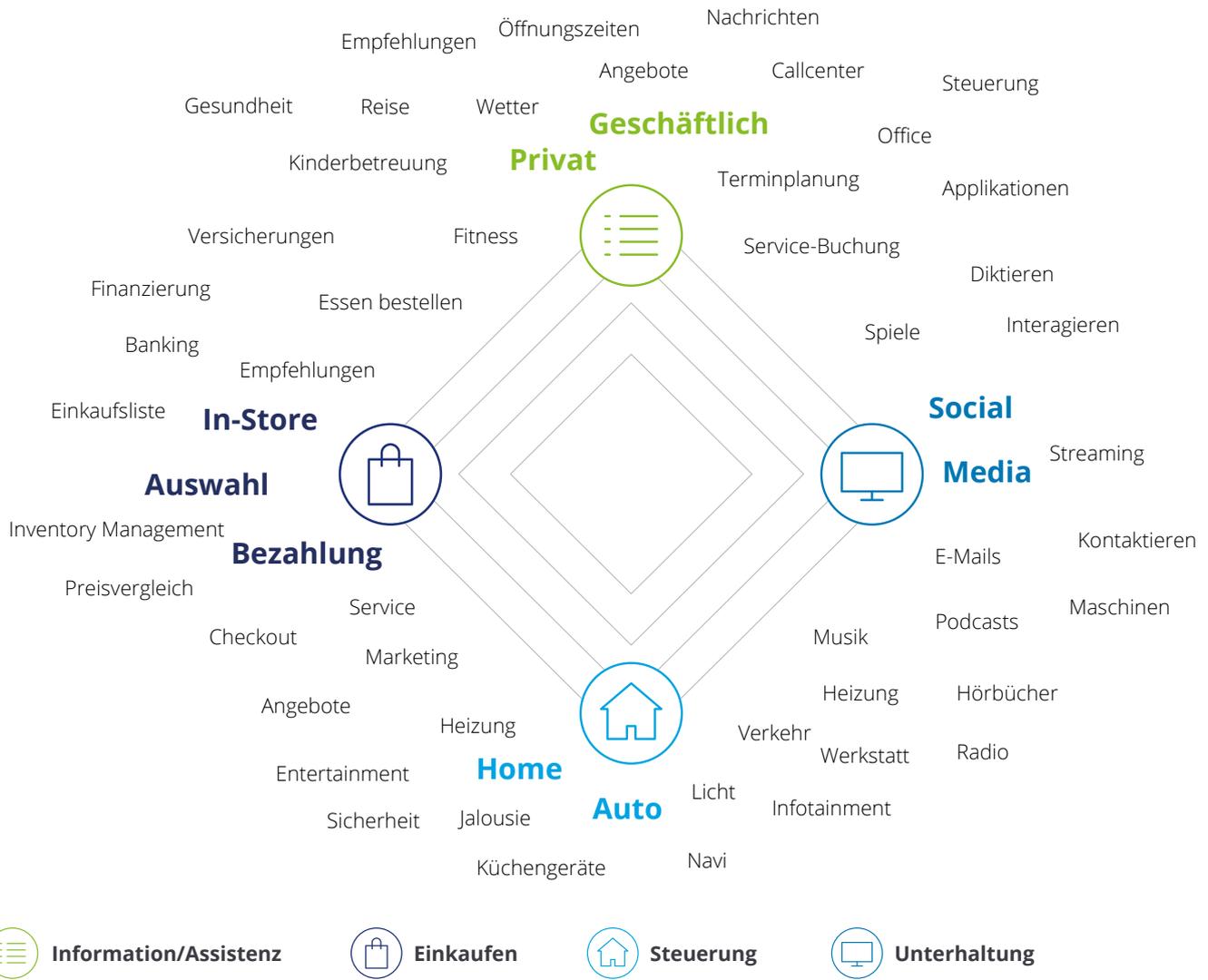
2 Mrd. Anfragen durch Nutzer über Voice pro Woche

Ort der Nutzung ...



Über 30 Mio. Smart Speaker wurden 2017 verkauft, um Nutzer mit Informationen, Unterhaltung, Steuerungs- und Einkaufsmöglichkeiten zu versorgen.

Abb. 3 – Reichweite des Voice-Assistent-Ökosystems



Quelle: Monitor Deloitte 2018.

Das Voice-Ökosystem

Die Reichweite von digitalen Sprachassistenten ist enorm und dringt in immer mehr Bereiche von Gesellschaft und Wirtschaft vor. Das Potenzial beschränkt sich bei Weitem nicht nur auf den privaten Einsatz, sondern betrifft auch zunehmend berufliche und geschäftliche Segmente. Mit rasanter Geschwindigkeit werden immer neue Themengebiete erschlossen. Das heißt, künftig werden Voice-assisted Interfaces für alle Branchen und Player relevant. Um wettbewerbsfähig zu bleiben, wird sich kaum ein Unternehmen diesem Megatrend entziehen können.

Bereits heute ist das Angebotsspektrum äußerst vielfältig. Privat werden Sprachassistenten häufig verwendet, um allgemeine Informationen aus dem Internet abzurufen, wie etwa den Wetterbericht oder Nachrichten. Aber auch die Steuerung von Geräten zu Hause, wie etwa der Heizung,

des Fernsehers, von Jalousien, Licht oder Küchengeräten erfolgt zunehmend über Voice Assistants. Im Auto können Navigationsgerät, Infotainment und Verkehrsinformationen mit Sprachbefehlen gesteuert werden. Und in der Arbeitswelt bieten Office-Applikationen, Callcenter oder Maschinensteuerung weitere intelligente Einsatzmöglichkeiten für Voice-assisted Interfaces. Auch das Einkaufen wird durch die neue Technologie revolutioniert.

Unternehmen, Hersteller und Händler, deren Business-Umfeld nahe der Kernfunktionen liegt, sind besonders stark betroffen. Sie müssen Strategien entwickeln, wie sie ihre Produkte und Services am besten in den Markt für Voice-assisted Interfaces integrieren können. Hier gibt es die Möglichkeit, eigene Dienste für die Sprachassistenten der Big Player zu entwickeln und auf deren Plattformen anzubieten.

Das Beispiel Google zeigt, wie große Technologiekonzerne die Vision eines weitreichenden Voice-Ökosystems, dem sich Unternehmen mit eigenen Anwendungen anschließen können, bereits umsetzen. Um Firmen bei der Entwicklung eigener Actions gezielt zu unterstützen, bietet Google sogenannte vertikal spezifische Lösungen in vier Themen-Bereichen an: Home Control Actions, Productivity Actions, Game & Fun Actions sowie Shopping Actions. Der bisherige Schwerpunkt von Google liegt hier vor allem auf Informations- und Assistentenfunktionen, weitere Bereiche werden jedoch sukzessive erschlossen. Insgesamt bedient das Marktangebot der großen Player jedoch schon jetzt substantielle Teile des Ökosystems.

Abb. 4 – Google Verticals



Quelle: Monitor Deloitte 2018.

basierend auf Google, Grafik zu finden unter <https://developers.google.com/actions/>

Der Voice-Assistent-Markt

In den letzten Jahren hat eine Vielzahl von Anbietern eigene Voice Assistants entwickelt. Allerdings konnten sich nur wenige Systeme großer Konzerne durchsetzen, die enorme Summen in die neue Technologie investiert und einen hohen Forschungsaufwand in diesen Zukunftsbereich gesteckt haben.

Den ersten Platz auf dem Markt für Voice Assistants in den USA belegt ganz klar der Konzern Apple mit seinem System Siri, wie aktuelle Zahlen vom Mai 2018 zeigen. Dabei nimmt der Hightech-Gigant aus Cupertino einen Marktanteil von 46 Prozent ein. Mit großem Abstand folgt Google auf Platz 2.

Die Tochter des Alphabet-Konzerns kommt mit Google Assistant auf einen Marktanteil von 29 Prozent. An dritter Stelle steht Amazons Alexa (Marktanteil: 13 %), gefolgt von Samsungs Bixby (6 %) und Cortana von Microsoft (5 %).

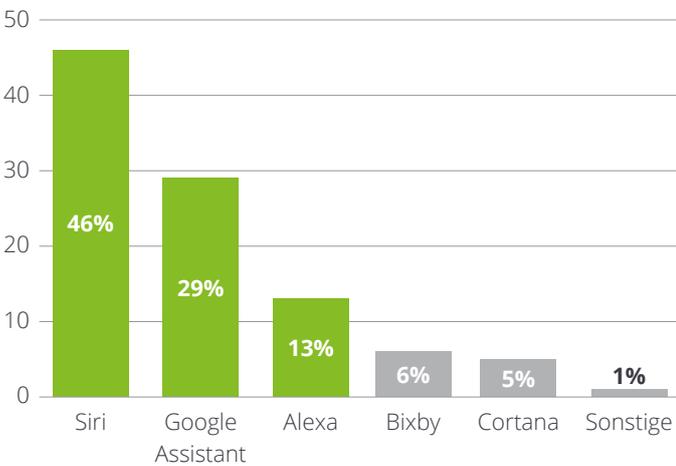
Die geringen Marktanteile der letzten beiden Konzerne haben hausgemachte Gründe. So ist Samsungs Bixby bisher nur auf Englisch und in wenigen asiatischen Sprachen verfügbar. Und Microsofts Cortana wurde anfangs nur auf den eigenen und wenig nachgefragten Windows Phones eingesetzt. Alle sonstigen Anbieter kommen zusammen lediglich auf einen marginalen Marktanteil von 1 Prozent.

Schaut man sich die Verkaufszahlen der Smart Speakers der letzten beiden Jahre an, fällt jedoch auf, dass Amazon und Google den Markt für das Zuhause als First Mover dominieren. Apple, Alibaba und weitere Player sind klar als Nachzügler zu erkennen.

Die Zahlen zeigen ein eindeutiges Bild: Durch die überwältigende Dominanz der großen Player haben kleinere Anbieter kaum noch realistische Chancen, sich auf dem Voice-Assistent-Markt zu etablieren.

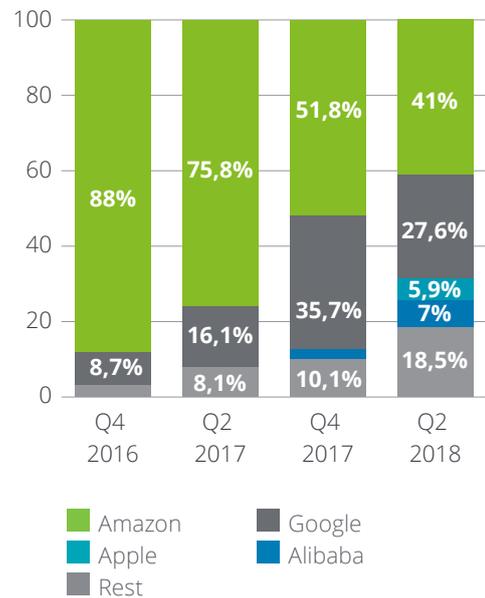
Abb. 5 – Marktanteile der größten Teilnehmer

Marktanteile Sprachgestützte Assistenzfunktionen (USA 2018)



Quelle: Statista.

Smart Speaker Absatz (Global)



Die Technologie hinter den Voice Assistants

Software-Komponenten

Zweifelsohne erfordert der technische Einsatz von Voice Assistants ein intelligentes Zusammenspiel von Software und Hardware. Die Software-Komponenten bei Voice-assisted Interfaces basieren auf vier wichtigen Kerntechnologien:

- Automated Speech Recognition (ASR)
- Natural Language Processing (NLP)
- Dialog Manager (DM)
- Text-to-Speech (TTS)

Bei der Nutzung von Voice Assistants durch den Anwender wird zunächst die Spracheingabe mithilfe von Automated

Speech Recognition (ASR) in Textform umgewandelt. Dann ordnet die Komponente Natural Language Processing (NLP) einzelnen Textfragmenten eine Bedeutung zu. Im Anschluss führt der Dialog Manager (DM) die notwendigen Schritte zur Beantwortung der Anfrage durch. Schließlich ermöglicht Text-to-Speech (TTS) die akustische Sprachausgabe der Antwort.

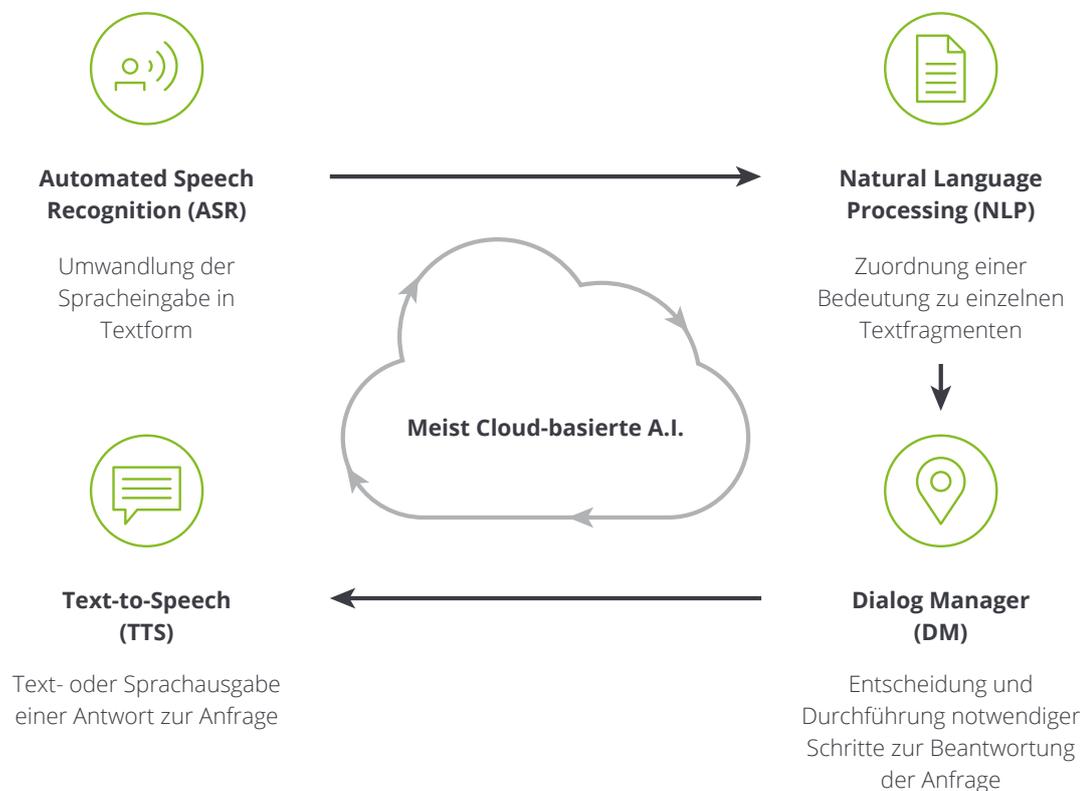
Zu den besonderen Herausforderungen dieser Technologien zählen neben dem richtigen akustischen und phonetischen Verstehen des gesprochenen Textes vor allem auch die kontextbezogene Interpretation der Suchwörter sowie das Erkennen der Absicht hinter der Suche.

Doch die konstante Weiterentwicklung dieser vier Kerntechnologien führt zu einer kontinuierlichen Verbesserung und

Nutzensteigerung. So wird die Voice Recognition durch technologische Fortschritte der zugrunde liegenden A.I. sukzessive optimiert. Außerdem ermöglicht der Einsatz immer leistungsstärkerer Prozessoren, aufwendige statistische A.I.-Modelle zu verarbeiten und eine große Menge strukturierter Daten effizient und rasch zu analysieren. Gleichzeitig werden verstärkt Metadaten im sprachlichen Kontext genutzt. Zudem ist im Bereich TTS eine zunehmend natürliche Sprachausgabe in immer mehr Sprachen möglich.

All diese Fortschritte in den einzelnen Software-Komponenten verbessern also fortlaufend die natürliche Interaktion der Nutzer mit den Voice Assistants. Mit dem wachsenden Mehrwert der Sprachassistenten beschleunigt sich auch deren exponentielle Verbreitung.

Abb. 6 – Übersicht der Software-Komponenten (meist Cloud-basiert)



Tatsächliche Aufteilung offline/online ist variabel und hängt vom zugrunde liegenden System ab.

Quellen: Google Cloud; Amazon Alexa Developer.

Software-Hardware-Schnittstelle

Für das reibungslose Zusammenspiel von Software und Hardware bei Voice Assistants ist auch die dazwischenliegende, beide Komponenten verbindende Schnittstelle von großer Bedeutung. Beim Voice Commerce liegt diese Schnittstelle in der Regel in der Cloud.

Cloud-basierte Systeme sind mit dem Internet verbunden und haben den Vorteil der unmittelbaren Verfügbarkeit von Neuerungen und Updates. Auch können auf diese Weise große Datenmengen rasch verarbeitet werden.

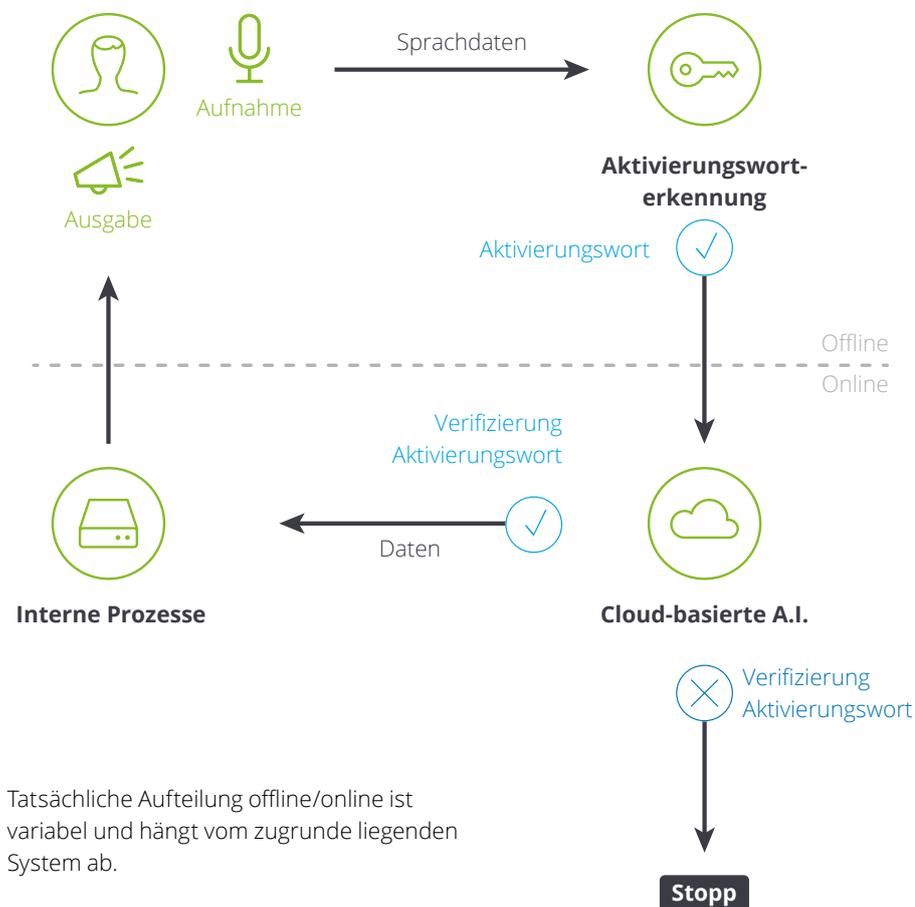
Ein Blick auf den typischen Ablauf einer Suchanfrage über voice-integrated Interfaces macht dies deutlich: Der Nutzer

aktiviert die Online-Verbindung durch ein Aktivierungswort und die Daten werden zur weiteren Verifizierung des Aktivierungswortes übertragen. Bei fehlender Verifizierung bricht der Prozess schon im Offline-Modus ab. Bei erfolgreicher Verifizierung startet die Cloud-basierte Weiterverarbeitung der Daten und es erfolgt die gewünschte Ausgabe.

Daneben gibt es auch Voice Assistants, die eine lokal verortete A.I. besitzen, bei der sämtliche Prozesse offline verlaufen. Für den Voice Commerce sind diese Systeme allerdings nicht relevant, da sie keine aktuellen Informationen wie Preise oder Verfügbarkeiten von Produkten aus dem Internet abrufen können.

So hängt auch der Erfolg Cloud-Basierter Smart Speaker für den Voice Commerce sehr stark von der Breitbandverfügbarkeit sowie den technischen Entwicklungen der A.I. ab. Außerdem müssen bei derartigen Systemen höhere Risiken im Hinblick auf Datenschutz und Sicherheit berücksichtigt werden. Da die Intelligenz des Systems Cloud-Basiert ist, ist die Hardware nicht der Hauptdifferenzierungsfaktor von Voice Assistants.

Abb. 7 – Hardware-Software-Schnittstelle



Tatsächliche Aufteilung offline/online ist variabel und hängt vom zugrunde liegenden System ab.

Quellen: Amazon Alexa Developer; Apple Machine Learning Journal.

Möglichkeiten der Voice-Integration für Händler und Hersteller

Für Händler und Hersteller gibt es verschiedene Möglichkeiten, die Voice Assistants für die Interaktion mit dem Kunden zu nutzen, die sich im Grad ihrer Integration und den verfügbaren Funktionen unterscheiden.

• Webbasierte Integration

Chatbots sind textbasierte Dialogsysteme, die die Kommunikation mit dem System über natürliche Sprache ermöglichen. Sie werden häufig in Online-Shops oder Support-Foren eingesetzt, um Standardfragen zu beantworten, aber auch, um markierte Texte vorlesen zu lassen.

• App-basierte Integration

Auch hier werden Chatbot-Funktionen eingesetzt, die eine Spracheingabe und -ausgabe ermöglichen. Händler und Hersteller können mit eigenen Apps in den Betriebssystemen verschiedener Anbieter wie etwa iOS von Apple oder Android von Google vertreten sein. Um hier die Bandbreite der Funktionen nutzen zu können, sind die Apps für die Sprachsteuerung kompatibel zu gestalten.

• Skill-/Action-basierte Integration für Smart Speakers

Die besten Chancen, den Voice Commerce-Kanal zu erschließen, liegen in einer Partnerschaft mit den etablierten großen Konzernen im Bereich der Smart Speaker. Hier gibt es für Hersteller und Händler die Möglichkeit, eigene Dienste für die Sprachassistenten der Big Player zu entwickeln und auf deren Plattformen anzubieten. Diese Anwendungen, die dem Prinzip nach mit Apps auf Smartphones vergleichbar sind, heißen im Smart-Speaker-Umfeld bei Amazon Skills und bei Google Actions. Dagegen bietet Apple bisher noch keine Möglichkeit für Drittanbieter, Siri-Apps zu programmieren.

Abb. 8 – Möglichkeiten der Voice-Integration

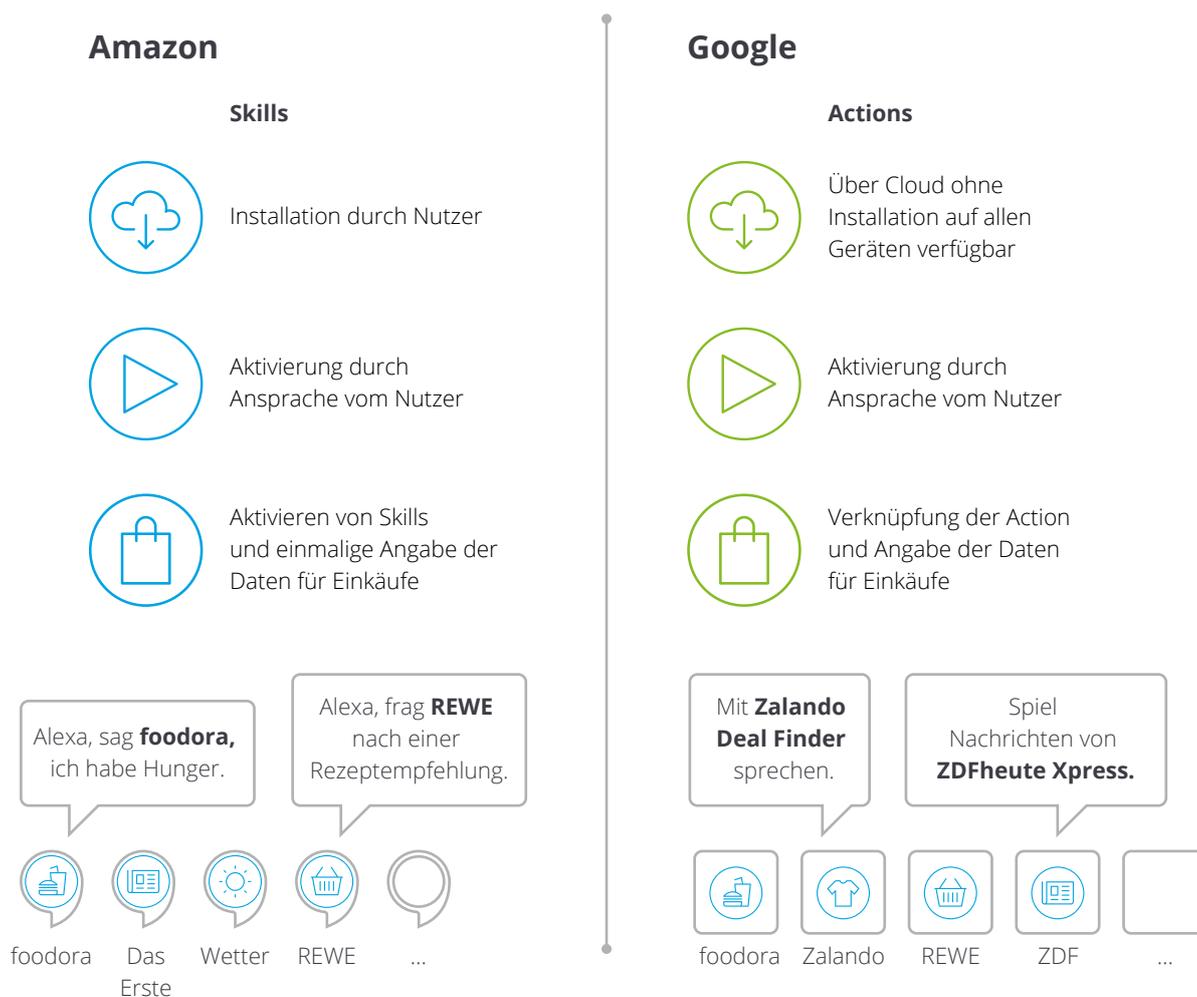


Quelle: Monitor Deloitte 2018.

Skills und Actions besitzen App-Charakter und sind auf ähnliche Weise zu bedienen: Im Voice-Umfeld kann sie der Nutzer jedoch sprachlich durch ein bestimmtes Aktivierungswort starten. Während bei Amazon ein aktives Freischalten notwendig ist, ist das bei Google nicht der Fall. Derzeit arbeiten alle Player aktiv daran, ihre Systeme technisch weiterzuentwickeln. Eine wichtige Rolle kommt dabei der sogenannten „Implicit Invocation“ zu. Hier wird versucht, implizite Anfragen der Nutzer direkt über den Konversationskontext einer Action zuzuweisen.

Da die Plattformen verschiedener Anbieter nicht untereinander kompatibel sind, müssen Firmen für jede Plattform eine eigene Anwendung entwickeln. Für den erfolgreichen Einsatz von Skills oder Actions ist es für Hersteller und Händler entscheidend, das Nutzungsverhalten ihrer Kunden zu verstehen und auf den richtigen Plattformen präsent zu sein. Für die technische Umsetzung steht eine große Auswahl potenzieller Partner im Bereich Voice Assistant Provider zur Verfügung, die die Firmen dabei unterstützen können.

Abb. 9 – Alexa Skills und Google Actions



Quelle: Monitor Deloitte 2018.

Voice-assisted Hardware

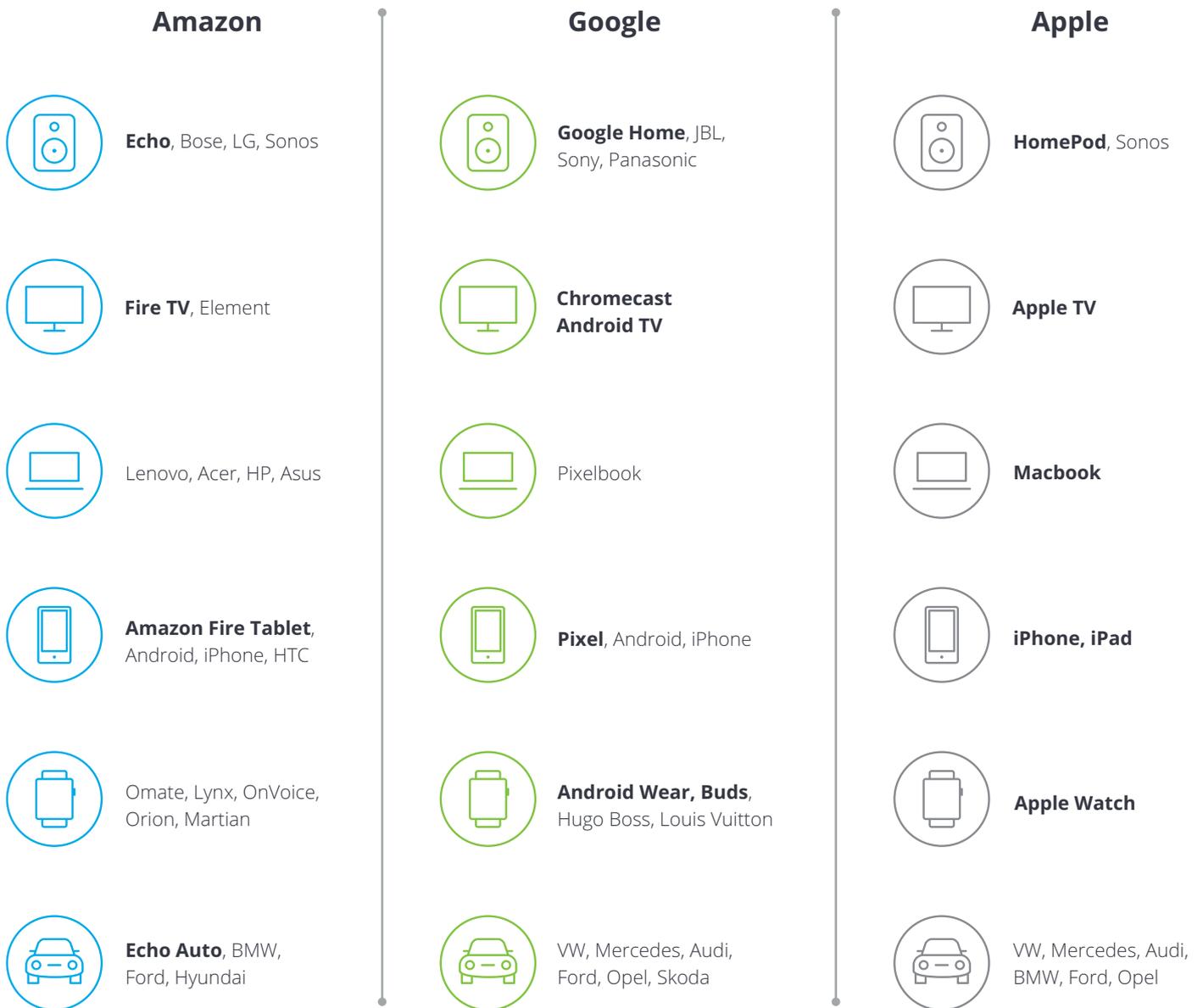
Neben den Voice Assistants bieten die großen Tech-Firmen auch die passenden eigenen Hardware-Geräte in verschiedenen Kategorien an, von Smart Speakers über Smartphones bis hin zu Smart TVs. Vor allem Amazon, Google und Apple haben ihre Voice Assistants in eine Vielzahl verschiedener Hardware-Geräte integriert. Dabei verfolgt Amazon eine offene Lizenzpolitik, sodass auch viele andere Hardware-Hersteller Alexa nutzen können. So ist Alexa nicht nur auf Amazon-Geräten wie Smart Speaker Echo, Fire TV oder Fire Tablet verfügbar, sondern auch auf den entsprechenden Produkten von Bose, Element, Lenovo oder vieler weiterer Firmen.

Auch Google Assistant ist neben der firmeneigenen Hardware wie Google Home, Chromecast oder Pixel auf vielen Geräten anderer Hersteller verfügbar, darunter Sony und Panasonic. Dagegen verfolgt Apple eine restriktive Lizenzpolitik und macht Siri in der Regel nicht für die Hardware anderer Hersteller nutzbar.

Die offene Lizenzpolitik von Amazon und Google hat Vorteile. Sie führt zu einer größeren Reichweite und Nutzung ihrer Voice Assistants über Hardware von Drittanbietern. Deswegen sind Alexa und Google Assistant auf bestem Weg, sich als die dominierenden hardwareübergreifenden Standards zu etablieren.

Die großen Technologiefirmen bieten neben der Software der Voice Assistants auch passende Hardware in allen Kategorien an. Hierbei ist der Hardware gegenüber der Software jedoch eine geringere Wichtigkeit beizumessen.

Abb. 10 – Hardware-Angebot der Tech-Firmen (Auszug)



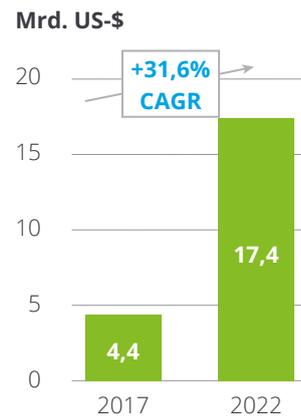
Quelle: Monitor Deloitte 2018.

Einblicke in den Voice-Assistent-Markt Die globale Marktentwicklung bei Smart Speakers

Der Einsatz von Voice Assistants ist eng verbunden mit der Hardware des Smart Speakers. Dabei handelt es sich um intelligente Lautsprecher, die über das Internet mit der Cloud verbunden sind. Über eingebaute Mikrofone können sie Sprachbefehle empfangen und verarbeiten sowie die Funktionen eines intelligenten persönlichen Assistenten integrieren. Die bekanntesten Beispiele dafür sind Amazon Echo, Google Home oder der HomePod von Apple.

Wie volatil dieses Hightech-Segment ist, zeigt ein Blick auf die Verkaufszahlen der wichtigsten Anbieter, die zum Teil einer Achterbahnfahrt gleichen. So konnte Marktführer Amazon die Zahl der weltweit verkauften Echos vom 3. auf das 4. Quartal 2017 auf 10 Millionen Stück nahezu verdoppeln. Dieser starke Umsatzsprung resultiert vor allem aus dem Weihnachtsgeschäft. Doch bereits im ersten Quartal 2018 ging der Absatz gleich wieder auf 4 Millionen Stück zurück. Auch Google, die Nummer zwei im Markt, weist bei dem Verkauf seiner Smart Speakers eine ähnliche Entwicklung auf.

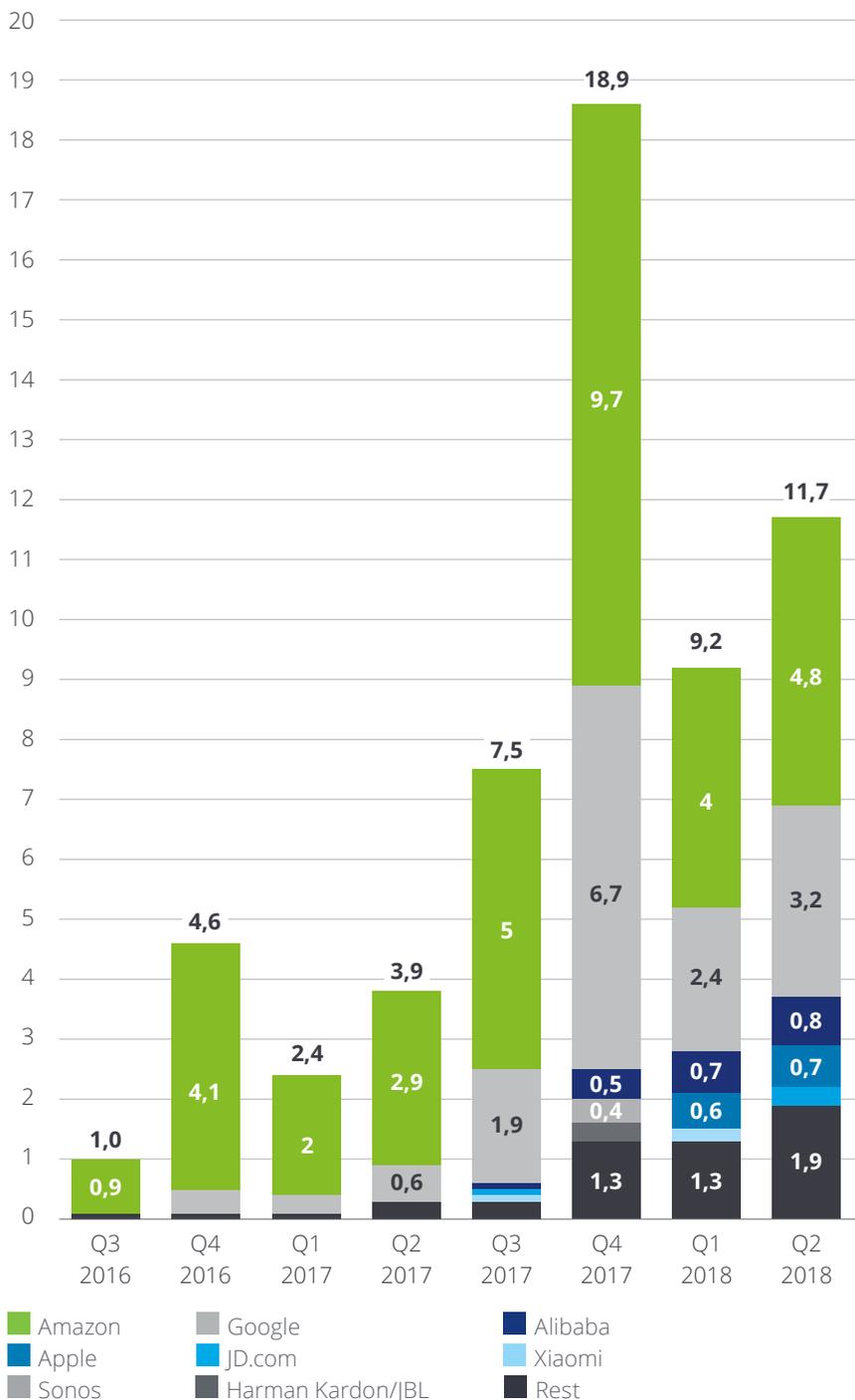
Abb. 11 – Globaler Smart-Speaker-Umsatz 2017 und 2022 in Milliarden US-\$



Quelle: International Data Corporation.

Als Marktführer im Bereich der Smart Speaker können Amazon und Google insbesondere von der Kombination aus Hardware und den entsprechenden Datenanalysefähigkeiten profitieren. Dagegen spielt Apple mit seinem HomePod bisher kaum eine Rolle im Markt und legt in seiner aktuellen Vermarktungsstrategie den Fokus eher auf andere Produkte wie iPhones, MacBooks und iPads. Gleichzeitig drängen neue asiatische Player wie Alibaba und Xiaomi auf den Markt, die bisher vor allem in Asien wachsen. So konnte Alibaba ab Q4 2017 drei Quartale in Folge jeweils 1 Million Smart Speakers absetzen. Dass die Asiaten früher oder später auch nach Europa expandieren dürften, ist sehr wahrscheinlich.

Abb. 13 – Weltweit verkaufte Smart Speakers (in Mio. Stück)



Was bedeutet diese Entwicklung nun für Hersteller und Händler? Um nicht den Anschluss zu verlieren, müssen sie sich so rasch wie möglich mit diesem Zukunftsthema auseinandersetzen. Dabei ist gut zu überlegen, in welche Technologie sie investieren und mit welchem Player sie langfristige Partnerschaften eingehen möchten. In jedem Fall lohnt es sich, die Entwicklung in den beiden größten Märkten USA und China im Auge zu behalten oder gar selbst dort aktiv zu werden. Eine naheliegende Möglichkeit wäre, zunächst etwa eine Kooperation mit einem der beiden Marktführer Amazon und Google zu starten. Später kann es sinnvoll sein, auch bei anderen Playern einzusteigen, um Nischen zu besetzen.

Quellen: Statista – Global smart speaker unit shipment 2016–2018, by vendor.

Anbietertypologien

Um den vielfältigen und teils unübersichtlichen Markt für voice-integrated Interfaces klarer zu strukturieren, ist es sinnvoll, die Player den wichtigsten Anwendungsgebieten zuzuordnen: Information, Unterhaltung, Steuerung und Einkaufen. So lassen sich die Anbieter basierend auf ihrem ureigenen Geschäftsmodell in vier verschiedene Archetypen einteilen.

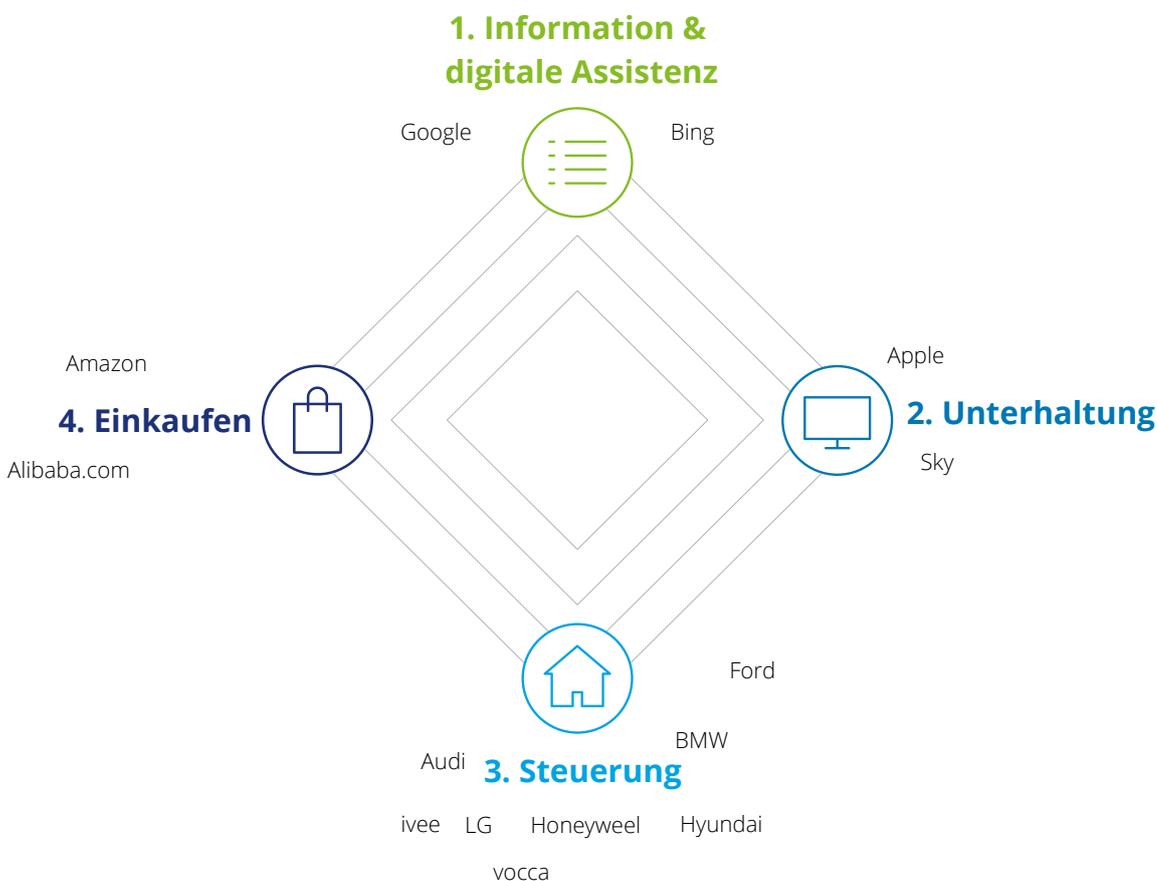
Der Archetyp Information und digitale Assistenz ist gekennzeichnet durch informationsgetriebene Serviceleistungen, die auf einer überlegenen Crawler-Technologie basieren. Hier spielen Suchmaschinen wie Google oder Bing eine wichtige Rolle.

Im Bereich Unterhaltung geht es um Hardware und Software-Lösungen mit starker Multi-Device-Interaktion, bei denen Usability und Design eine wichtige Rolle spielen. In diesem Archetyp nehmen Unternehmen wie Apple, Samsung oder Sky eine führende Position ein.

Bei der Dimension Steuerung stehen vor allem Smart-Home-Applikationen etwa von Isee oder Vocca im Vordergrund, die auf einer Kombination aus eigener Hardware und Smart Voice Features basieren. Diesem Archetyp sind jedoch auch Automobilhersteller wie Audi, Ford oder BMW zuzuordnen, die sprachgesteuertes Infotainment oder Navigationsfunktionen im Fahrzeug integrieren.

Der Archetyp Einkaufen bündelt Dienstleistungen rund um den sprachgesteuerten digitalen Retail, der durch ein vielfältiges Sortiment und unkomplizierte Distribution gekennzeichnet ist. Wichtige Player sind hier Amazon sowie in China Alibaba. Innerhalb des Marktes treffen verschiedene Anbieter aufeinander, die versuchen, ihre Kompetenzen Gewinn bringend einzusetzen. Um sich als Händler oder Hersteller richtig aufzustellen, ist es ratsam, die entsprechenden Entwicklungen genau zu betrachten. Zudem sollten Händler überlegen, wie sie sich selbst in den Markt einbringen und einzelne Marktteilnehmer im Rahmen eines Partnering-Prozesses unterstützen können.

Abb. 14 – Exemplarische Marktübersicht



Quelle: Monitor Deloitte 2018.

Globale Marktabdeckung

Auch die globale Marktabdeckung mit Voice Assistants ist schon weit vorangeschritten. Für die Verbreitung kommt es unter anderem darauf an, in welchen Sprachen die Dienste zur Verfügung stehen.

Von allen Playern weist Amazon aktuell die größte regionale Abdeckung mit Voice Assistants auf. Allerdings beschränkt sich der Konzern aus Seattle bisher noch

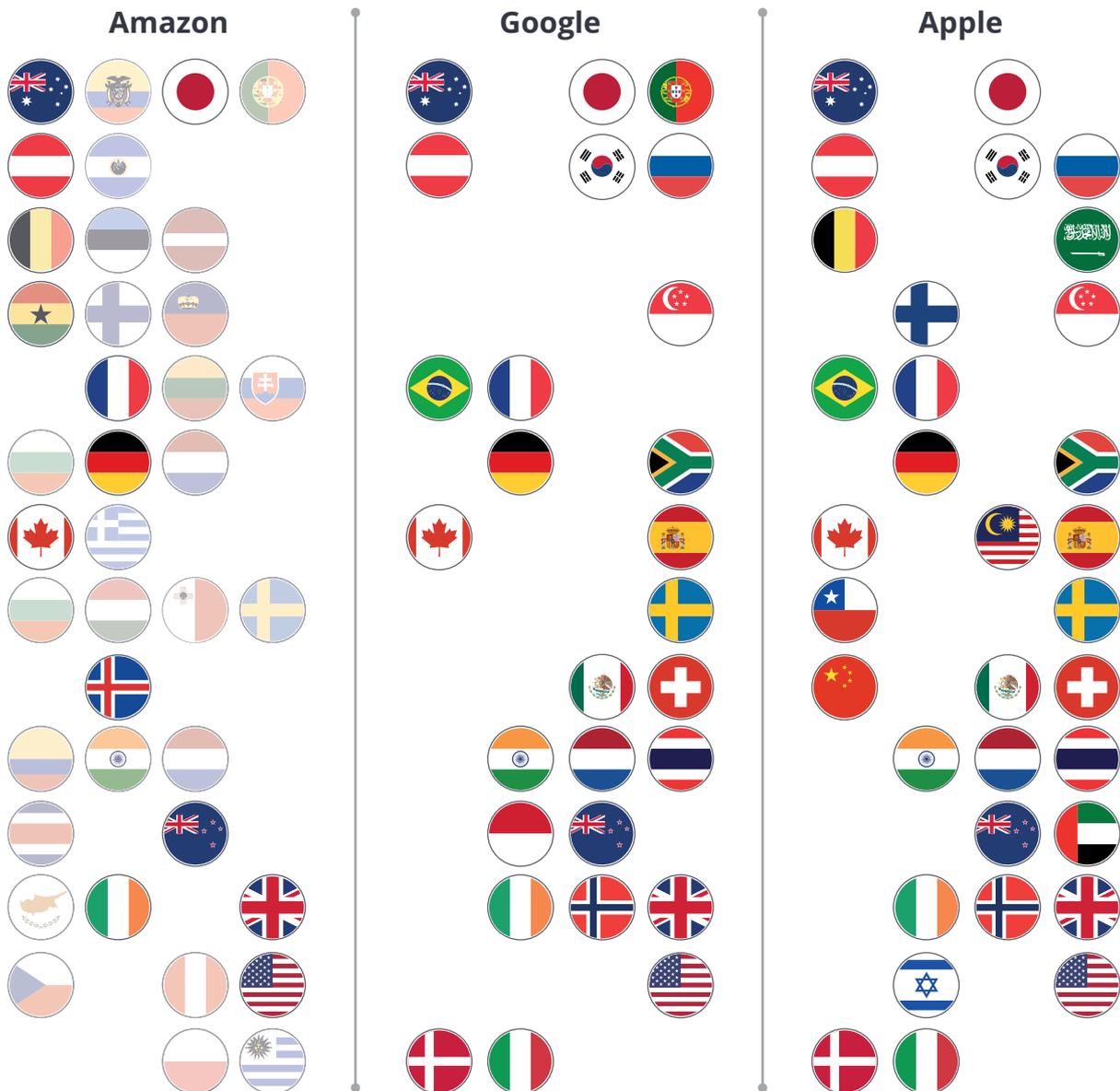
auf wenige Sprachen und verwendet international primär die englische Version, was viele Funktionen einschränkt. Damit dringt Amazon aggressiv in Märkte ein, strebt eine schnelle Marktabdeckung mit einheitlichen Produkten an und setzt auf eine Pfadabhängigkeit.

Dagegen breitet sich Google langsamer, dafür aber spezifischer in den Märkten aus und bietet mehr Sprachen an. Heute ist die

Alphabet-Tochter bereits in 27 Nationen in der jeweiligen Landessprache vertreten, bis Jahresende sollen weitere Sprachen und Anpassungen hinzukommen.

Die breiteste Sprachpalette bietet Apple an. Der Konzern legt großen Wert auf Qualität und Individualisierung und lässt Siri in über 30 Ländern in der jeweiligen Muttersprache mit den Nutzern kommunizieren.

Abb. 15 – Globale Marktabdeckung mit Voice Assistants (Auswahl an Ländern, Stand Mitte 2018)



Ausgegraut = nur in englischer Sprache.

Quelle: Monitor Deloitte 2018.

Voice Commerce hat ein riesiges bislang ungenutztes Potenzial, das kurz davor steht, freigesetzt zu werden und E-Commerce sowie stationären Handel zu revolutionieren.

Voice Commerce in Deutschland

Während Voice Commerce in den USA bereits weit verbreitet ist, steckt es in Deutschland noch in den Kinderschuhen. Die Amerikaner haben sprachgesteuerte Einkäufe etwa über den Google Assistant bei Walmart und Target längst in ihr Alltagsleben integriert.

Dagegen wird Voice Shopping hierzulande nur sehr vereinzelt angeboten und Voice Assistants werden vorrangig zu Informationszwecken genutzt. Insbesondere das Einkaufsangebot für Drittanbieter ist noch sehr eingeschränkt. Ausnahmen sind Lieferservices oder der Deal des Tages. Was jedoch ganz klar noch in Deutschland fehlt, ist ein breites, attraktives Shopping-Angebot.

Dennoch hat Voice Commerce in Deutschland ein riesiges, bislang ungenutztes Potenzial, das kurz davor steht, freigesetzt zu werden. Doch erst wenn sich attraktive Shopping-Möglichkeiten über den Voice-Kanal entwickeln, wird die Technologie wirklich interessant für Anbieter und Nutzer. Deswegen wird der sprachgesteuerte Einkauf früher oder später auch hierzulande sowohl den E-Commerce als auch den stationären Handel revolutionieren.

Abb. 16 – Voice-Commerce-Angebot in Deutschland

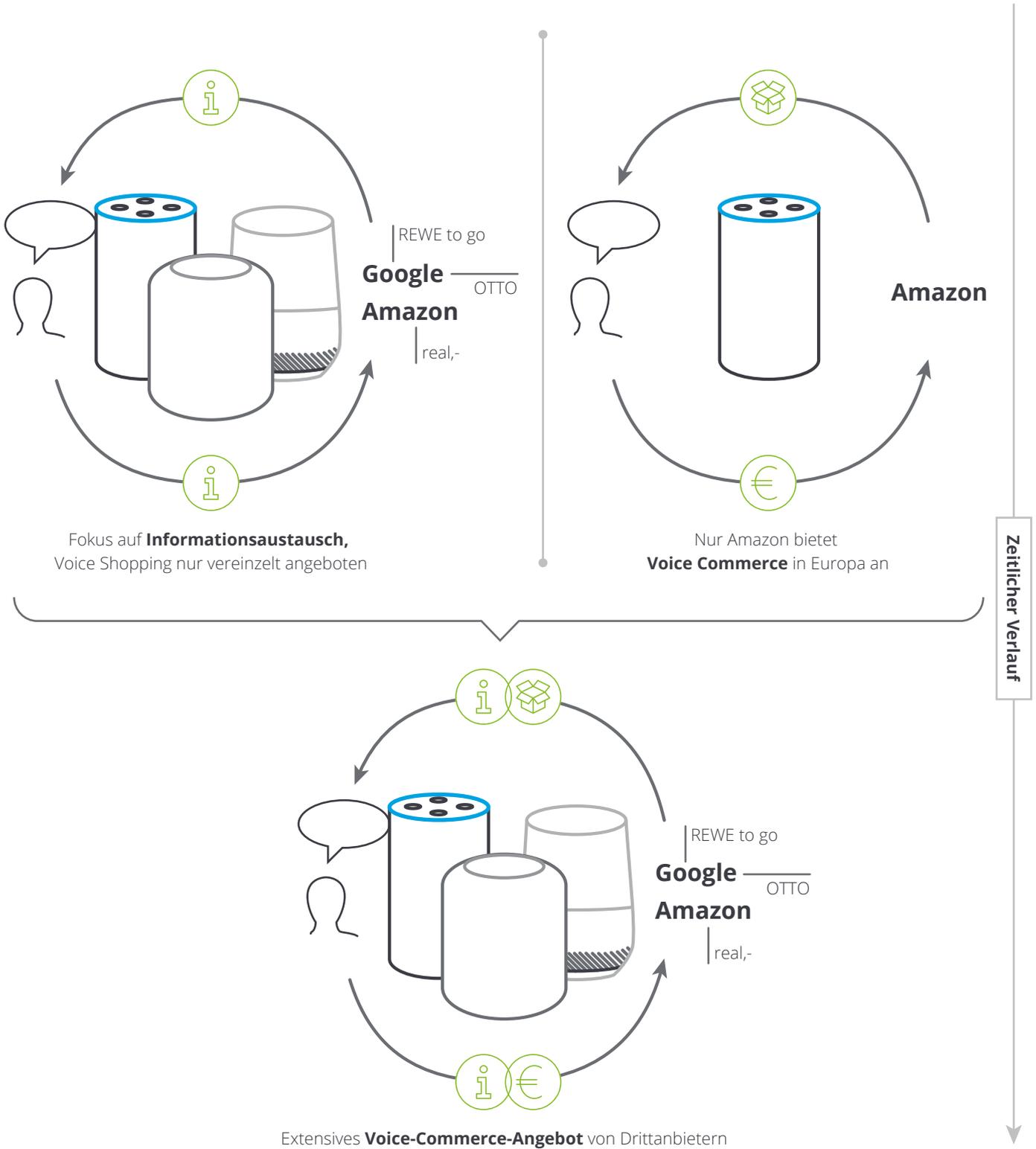
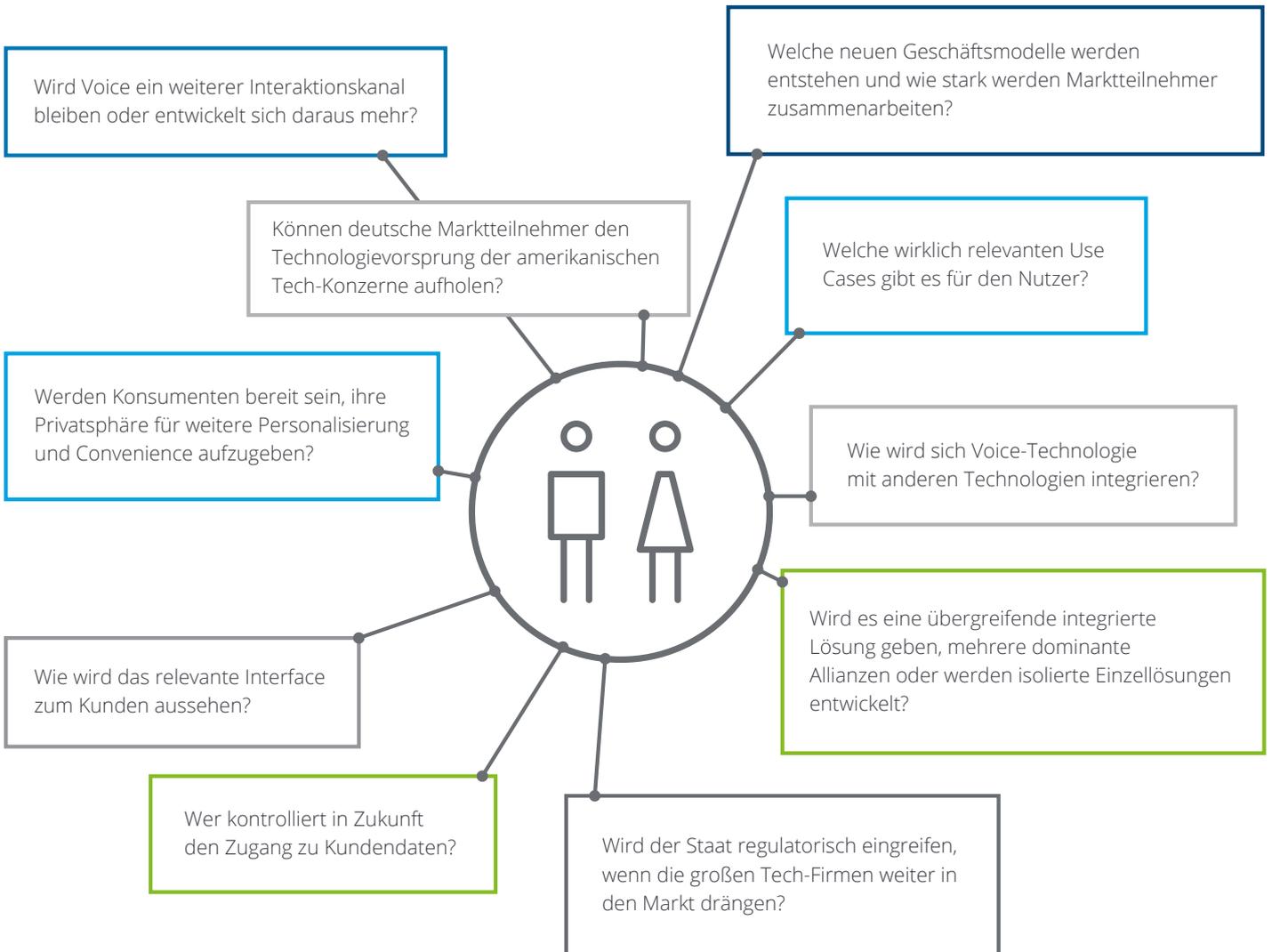


Abb. 17 – Fragen zur Zukunft von Voice Commerce



Quelle: Monitor Deloitte 2018.

Forschungsbedarf: Die Auswirkungen von Voice-assisted Interfaces auf den Handel

Die bisherige Bestandsaufnahme macht deutlich, dass die Bedeutung von Voice Commerce zunehmend größer wird. Allein der globale Smart-Speaker-Umsatz wächst scheinbar unaufhaltsam und wird Prognosen zufolge bis 2022 auf 17,4 Mrd. US-Dollar ansteigen. Darüber hinaus werden Voice Interfaces in immer mehr Geräte integriert. Schon heute durchdringen Voice Assistants als ständige Begleiter fast alle Bereiche des täglichen Lebens.

Auch kristallisiert sich bereits heraus, dass eine geringe Anzahl an Playern den Großteil des Marktes für sich behaupten wird. Vor allem amerikanische Tech-Firmen dominieren den Markt und expandieren aggressiv. Gleichzeitig drängen immer neue Player, etwa aus Asien, auf den Markt.

Doch die Zukunft ist geprägt von einer Vielzahl an Unsicherheiten. Deswegen besteht ein hoher Forschungsbedarf über die Zukunft des Voice Commerce, um Händlern und Herstellern konkrete Handlungsempfehlungen für ihre künftige Geschäftsstrategie in diesem wichtigen Sektor zu geben.

Trotz der bisherigen eindeutigen Entwicklung von Voice-Technologien und Markt bleiben viele Fragen ungeklärt.

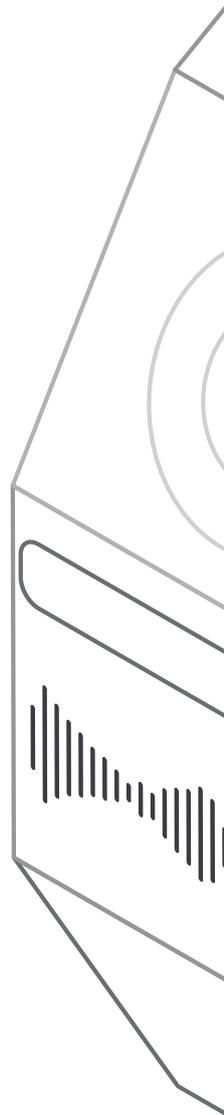
- Welchen Einfluss haben Voice-assisted Interfaces auf die Zukunft des Einkaufens?
- Welche Konsequenzen ergeben sich daraus für Händler, Hersteller und Konsumenten?
- Welche strategischen Fragestellungen sollten sich einzelne Marktteilnehmer stellen, um sich optimal aufzustellen?

Diesen Fragen geht Monitor Deloitte in den folgenden Kapiteln dieser Studie auf den Grund. Mithilfe von Artificial Intelligence und der Zusammenarbeit mit Experten aus Technologie und Handel wurden vier Szenarien zur Zukunft von Voice Assistants entwickelt. Und daraus wichtige Implikationen für die jeweiligen Marktteilnehmer abgeleitet.

Die Zukunft von Voice-assisted Interfaces

Entwicklung von Zukunftsszenarien mit dem Dynamic-Strategy-Ansatz

Um die künftigen Entwicklungen des Voice Commerce und die damit verbundenen Auswirkungen auf Hersteller und Händler besser verstehen zu können, hat Monitor Deloitte in der vorliegenden Studie einen detaillierten Blick in die Zukunft gewagt. Sie gibt Antworten auf die zentrale Frage: Wie werden Voice Assistants das Einkaufen in Europa im Jahre 2030 beeinflussen? Als zentrales Element der Studie wurden mithilfe von Szenariodesign vier Zukunftsszenarien entwickelt und daraus konkrete Handlungsempfehlungen für Händler und Hersteller abgeleitet. Die Studie ist das Ergebnis einer Reihe von Workshops, an denen Experten aus Handel und Technologie, Branchenexperten von Monitor Deloitte sowie erfahrene Szenariodesign-Praktiker aus dem „Center for the Long View“ beteiligt waren.



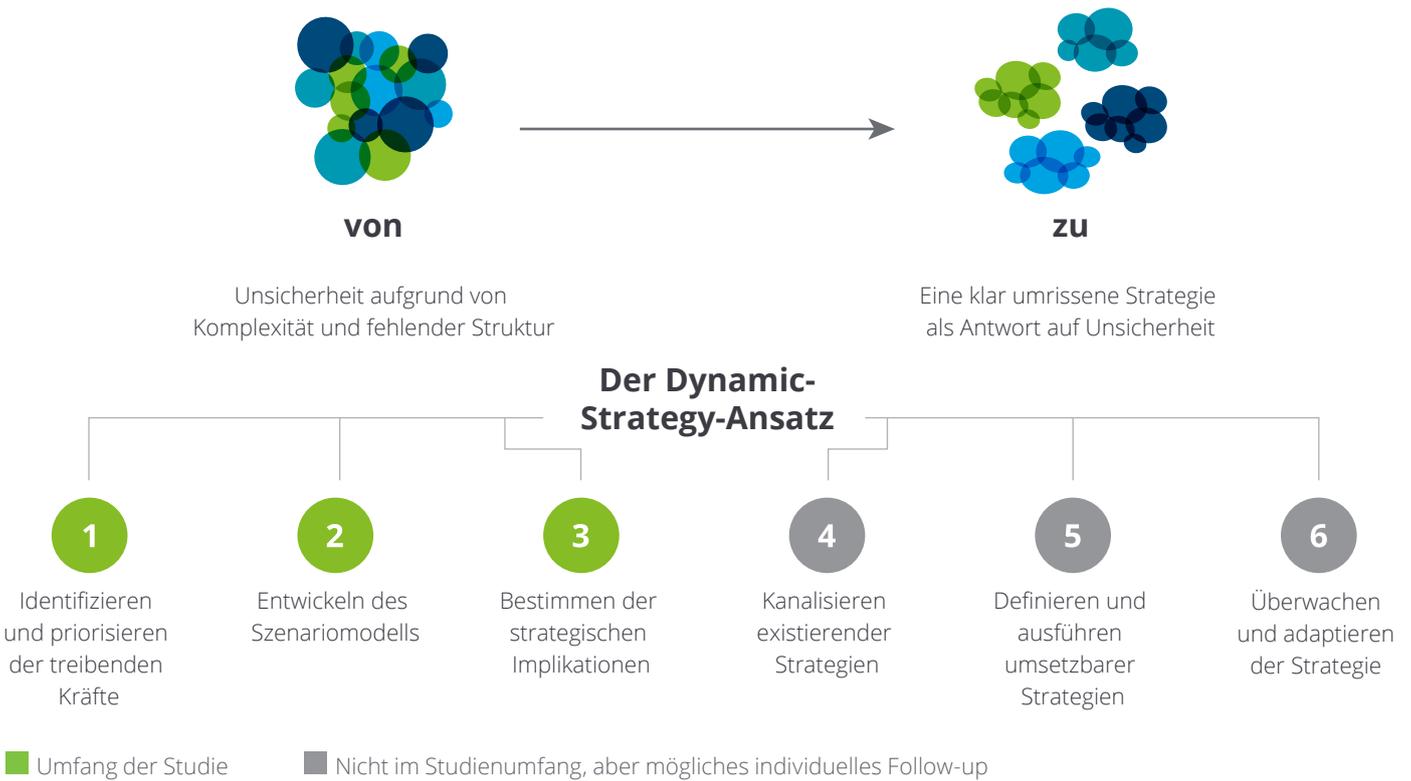


Die Methodik dieser Studie basiert auf dem von Deloitte entwickelten sechsstufigen Dynamic-Strategy-Ansatz und wendet die wissenschaftlichen Leitprinzipien der Objektivität, Zuverlässigkeit und Validität an. Er beschreibt den Weg von einem Zustand der Unsicherheit, der von Komplexität und fehlender Struktur geprägt ist, hin zu einer klar umrissenen Strategie. In der vorliegenden Studie über die Zukunft des Voice Commerce liegt der Fokus auf den ersten drei Schritten dieses Ansatzes: identifizieren und priorisieren der treibenden Kräfte, entwickeln des Szenariomodells und bestimmen der strategischen Implikationen. Die letzten drei Schritte finden sich nicht in der Studie wieder, da sie sehr

individuell auf einzelne Unternehmen anzupassen sind und auf den Ergebnissen der Studie aufbauen. Dabei geht es um das Kanalisieren existierender Strategien, das Definieren und Ausführen umsetzbarer Strategien sowie schließlich um das Überwachen und Adaptieren der Strategie – und das alles jeweils auf ein konkretes Unternehmen bezogen. Dieses Angebot nehmen Unternehmen gerne wahr und definieren ihre individuelle Strategie gemeinsam mit Monitor Deloitte und dem Center for the Long View, um sich erfolgreich für eine dynamische Zukunft aufzustellen. Sollten auch Sie Interesse daran haben, so wenden Sie sich gerne an unsere Ansprechpartner. (S. 76)

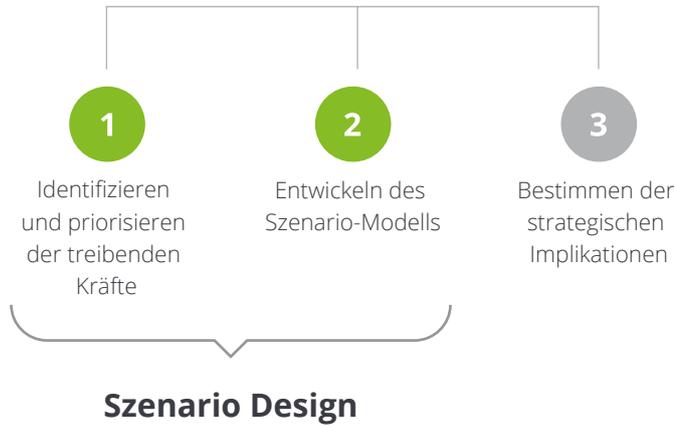
Weitere Informationen zur verwendeten Methodik des Szenariodesigns finden Sie auf der Seite des Center for the Long View unter <https://www2.deloitte.com/de/de/pages/strategy/topics/center-for-the-long-view.html>.

Abb. 18 – Dynamic Strategy Ansatz des Center for the Long View



Quelle: Monitor Deloitte 2018.

Abb. 19 – Für Studie gewähltes Vorgehen



Google



Experten aus
führenden LEH,
Discountern &
Versandhandel

Deloitte.

Quelle: Monitor Deloitte 2018.

Vier mögliche Szenarien der Zukunft von Voice Commerce

Zur Entwicklung des Szenariorahmens wurden die beiden ausgewählten kritischen Unsicherheiten – also der mögliche Grad der Kooperation zwischen den Marktteilnehmern und die damit verbundene Offenheit des Systems sowie das potenzielle Niveau der technologischen Integration – auf zwei orthogonale Achsen übertragen. Die Endpunkte der ersten Achse zeigen die möglichen Gegensätze, dass die Integration mit anderen Technologien entweder niedriger oder höher als erwartet ausfällt. Die Endpunkte der zweiten Achse markieren die unterschiedlichen Pole, ob es zu einer geringen Kollaboration der Marktteilnehmer mit geschlossenen Systemen kommt oder ob es kollaborative, offene Systemen geben wird. Die dadurch entstehenden Felder ergeben vier sehr unterschiedliche, aber plausible Szenarien: World of Aggregators, Old Europe, The new Internet sowie Age of heterogenic Alliances. Schließlich werden die vier verschiedenen Szenarien zur Zukunft des Voice Commerce mit all ihren Besonderheiten und Charakteristika ausführlich definiert. Dabei spielen die zuvor untersuchten Treiber eine wichtige Rolle.

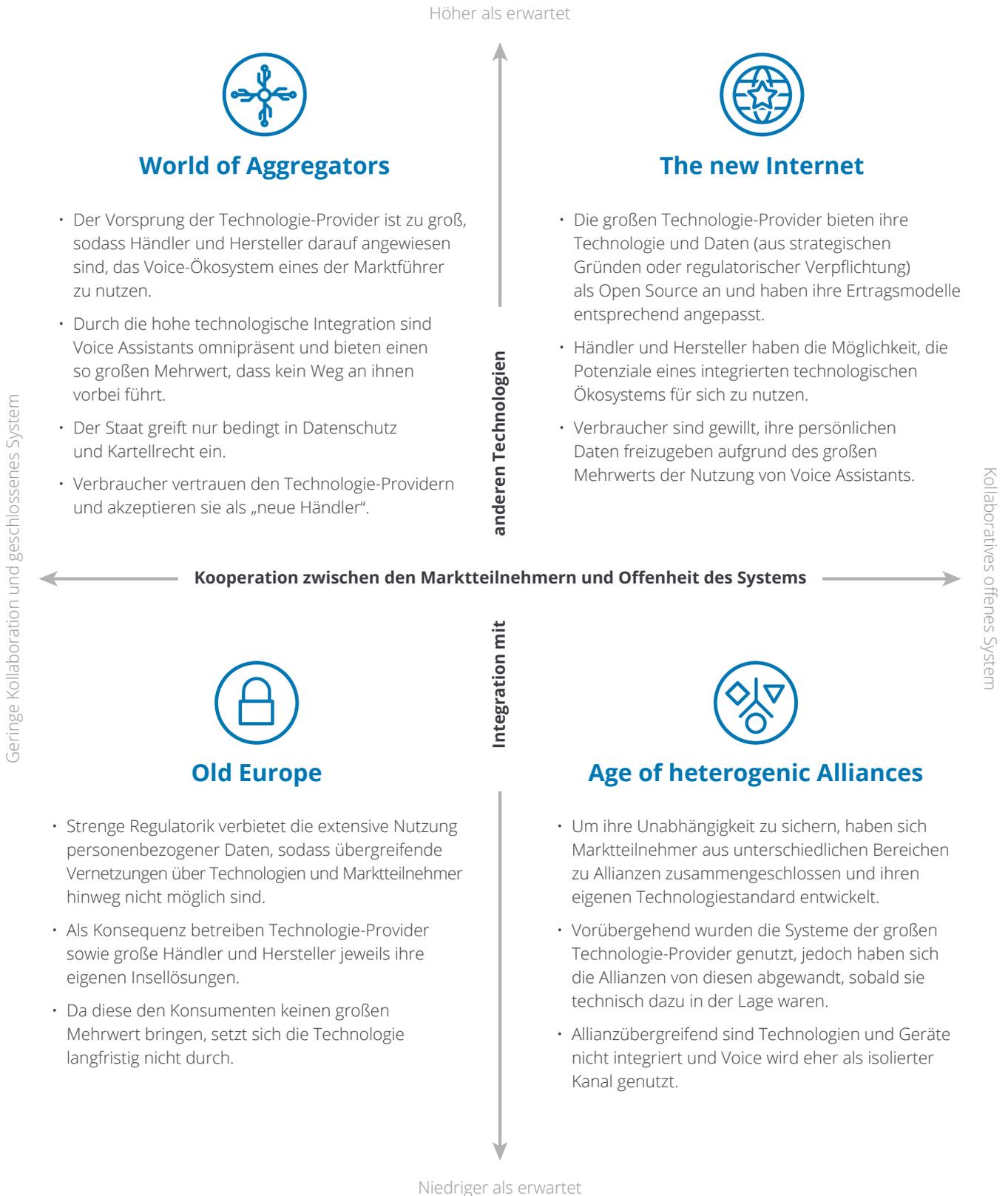
Hier noch eine wichtige Anmerkung zum richtigen Verständnis des Dargestellten: Ziel der Studie ist es nicht, mit den Szenarien vier verschiedene Geschichten zu erzählen und dann zu hoffen, dass eine davon wahr sein wird. Vielmehr wird davon ausgegangen, dass die reale Zukunft nicht genau einem der vier Szenarien entsprechen wird, sondern dass sie Elemente aller vier Szenarien enthalten wird. Die Methodik zielt darauf ab, die Ecken der plausiblen Zukunft festzulegen. Diese sind Extrempunkte und stellen die äußeren Grenzen des Plausiblen dar, sodass die Szenarien eine überzeichnete Qualität aufweisen.

Das Modell hat viele Vorteile: In den einzelnen Szenarien wird nicht nur die zukünftige Voice-Commerce-Landschaft genau beschrieben. Darüber hinaus werden auch die Auswirkungen der Szenarien auf verschiedene Arten von Marktteilnehmern untersucht und daraus Konsequenzen und konkrete Handlungsempfehlungen abgeleitet.

Abb. 20 – Definition der Szenarien



Abb. 21 – Eintrittsvoraussetzungen der Szenarien



Quelle: Monitor Deloitte 2018.

Eintrittsvoraussetzungen der Szenarien

Die wichtigsten Voraussetzungen für das Eintreten eines bestimmten Szenarios liegen vor allem in der Entwicklung der kritischen Unsicherheiten. Doch darüber hinaus gibt es noch eine Reihe weiterer Faktoren, die hier eine entscheidende Rolle spielen.

• The new Internet

Das Eintreten dieses Szenarios setzt voraus, dass die großen Technologie-Provider ihre Technologien und Daten – aus strategischen Gründen oder regulatorischer Verpflichtung – als Open Source anbieten und ihre Ertragsmodelle entsprechend anpassen. So haben Händler und Hersteller die Möglichkeit, die Potenziale eines integrierten technologischen Ökosystems für sich zu nutzen. Außerdem müssen die Verbraucher auch dazu bereit sein, ihre persönlichen Daten freizugeben, und gleichzeitig einen großen Mehrwert in der Nutzung von Voice Assistants sehen.

• Age of heterogenic Alliances

Dieses Szenario bedingt, dass sich die Marktteilnehmer aus unterschiedlichen Bereichen tatsächlich zu Allianzen zusammenschließen und ihre eigenen Technologiestandards entwickeln, um ihre Unabhängigkeit zu sichern. Vorübergehend wurden die Systeme der großen Technologie-Provider genutzt, jedoch haben sich die Allianzen von diesen abgewandt, sobald sie technisch dazu in der Lage waren. Allianzübergreifend sind Technologien und Geräte nicht integriert und Voice wird eher als isolierter Kanal genutzt.

• Old Europe

Eine wichtige Voraussetzung für diese mögliche Realität ist, dass eine strenge Regulatorik die extensive Nutzung personenbezogener Daten verbietet, sodass übergreifende Vernetzungen über Technologien und Marktteilnehmer hinweg nicht möglich sind. Als Konsequenz betreiben Technologie-Provider, große Händler und Hersteller jeweils ihre eigenen Insellösungen. Da diese den Konsumenten keinen großen Mehrwert bieten, wird sich die Technologie dieser Insellösungen langfristig nicht durchsetzen.

• World of Aggregators

Hier ist entscheidend, dass der Vorsprung der Technologie-Provider wirklich so groß ist, dass Händler und Hersteller darauf angewiesen sind, das Voice-Ökosystem einer der Marktführer zu nutzen. Zudem sollten die Voice Assistants durch die hohe technologische Integration omnipräsent sein und müssen einen so großen Mehrwert bieten, dass kein Weg mehr an ihnen vorbei führt. Auch darf hier der Staat nur bedingt in Datenschutz und Kartellrecht eingreifen. Schließlich müssen die Verbraucher den Technologie-Providern auch vertrauen und sie als „neue Händler“ akzeptieren.



The new Internet

Im Szenario „The new Internet“ werden Voice-assisted Interfaces einen maßgeblichen Einfluss auf das Einkaufsverhalten und die Marktdynamik haben.

In diesem Szenario hat sich Voice als Plattform mit einheitlichen Technologiestandards und offenen Schnittstellen etabliert, die alle Marktteilnehmer frei nutzen können. Das „neue Internet“ ist sowohl in unterschiedlichste Geräte – vom Smart Speaker bis zum Kühlschrank – als auch in verschiedene andere Technologien – wie Bilderkennung oder Sensorik – vollständig integriert.

Hier arbeiten alle Marktteilnehmer in einem offenen System sprachgestützter Schnittstellen zusammen. Auch die Art des Einkaufens hat sich völlig verändert. Denn die Verbindung von datengesteuerten Analysen mit intelligenten Geräten macht ein prädiktives und automatisiertes Einkaufen möglich.

Die wichtigsten Auswirkungen auf Handel, Hersteller und Verbraucher

- Die Relevanz des **Handels** hat in diesem Szenario abgenommen und großvolumige Einkäufe werden über Direktverkäufe beim Hersteller getätigt. Gleichzeitig sind neue Logistikdienstleister entstanden, die sich 24/7 um die Lieferung der Waren direkt zum Kunden kümmern. Erfolgreiche Händler konzentrieren sich vor allem auf den Verkauf von High-Engagement-Produkten. Sie haben ihre Filialen mit modernster Technologie aus dem Voice-Universum ausgestattet und die Kanäle Voice, Online und Offline komplett in ihre Geschäftsstrategie integriert.
- Erfolgreiche **Hersteller** haben frühzeitig erkannt, dass Investitionen in Marken unerlässlich sind, um die Kunden erreichen zu können und mit ihnen ein direktes Gespräch über sprachgestützte Schnittstellen zu führen. Darüber hinaus konzentrieren die Hersteller ihre Marketingbudgets auf Affiliate-Programme mit Voice-Assistant-Anbietern und incentivieren diese für die bevorzugte Bewerbung ihrer Produkte.
- **Verbraucher** können hier die Vorteile sprachgestützter Schnittstellen voll ausschöpfen, da die vollständige technologische Integration und die Zusammenarbeit mit den Marktteilnehmern eine höchstmögliche Transparenz der Kundendaten ermöglichen. Den Verbrauchern wird durch prädiktive Angebote und automatisierte Bestellungen der volle Shopping-Komfort geboten. Außerdem übernehmen Voice Assistants weitgehend den Auswahlprozess der Produkte: Sie filtern die verfügbaren Optionen etwa gemäß den Präferenzen des Kunden vor, bevor sie diesem konkrete Vorschläge unterbreiten. Dafür müssen die Verbraucher allerdings als „gläserne Kunden“ ihre persönlichen Daten freigeben.

Voice verschmilzt mit anderen Technologien zu einer Plattform mit offenen Schnittstellen, die den Nutzer in allen Lebenslagen umgibt und unterstützt.



Age of heterogenic Alliances

Im Szenario „Age of heterogenic Alliances“ schließen sich Händler und Hersteller zusammen, um sich erfolgreich gegenüber Technologie-Providern zu behaupten.

Ein wichtiges Kennzeichen dieses Szenarios sind heterogene Allianzen großer Händler und Hersteller. Diese verwenden jeweils eigens entwickelte Sprachtechnologien, die auf verschiedenen Standards beruhen und nicht allianzübergreifend kombinierbar sind. Hier gibt es also keine technologie- und geräteübergreifenden Lösungen.

In der Welt der heterogenen Allianzen arbeiten die Marktteilnehmer nur innerhalb ihres eigenen Verbundes sehr kooperativ zusammen. Denn hier sind mehrere unterschiedliche Allianzen zwischen Händlern und Herstellern entstanden, die jeweils eigene technologische Standards für ihre eigene sprachgestützte Schnittstelle verwenden. Das bedeutet: Offene Systeme und einen aktiven Datenaustausch gibt es nur innerhalb einer Allianz. Dagegen ist ein Datenaustausch über Allianzen hinweg aufgrund technologischer Hürden nicht möglich.

Die wichtigsten Auswirkungen auf Handel, Hersteller und Verbraucher

- Angetrieben von der Konkurrenz aus den USA haben sich in diesem Szenario erfolgreiche **Einzelhändler und Hersteller** zusammengeschlossen. Idealerweise ergänzen diese sich optimal in ihrem Produktportfolio und weisen nur einen geringen Überschneidungsgrad auf. Eine nahtlose Customer Journey und ein komplettes Produktportfolio sind hier die Schlüssel, um Kunden für eine bestimmte Allianz gewinnen zu können. Um sich zu differenzieren, investieren die Allianzen in Service-Bundling- und Loyalty-Programme sowie in Voice-Technologien entlang der Wertschöpfungskette. Ein kritischer Erfolgsfaktor ist hier der Aufbau von Barrieren, die verhindern sollen, dass Kunden von einer Allianz zur anderen wechseln.
- Die **Verbraucher** haben die Qual der Wahl und müssen sich für eine bestimmte Allianz entscheiden. Dadurch verpassen sie jedoch die Customer Experience oder mögliche Preisvorteile einer anderen Allianz. Alternativ können sich Verbraucher für mehrere Allianzen anmelden, müssen sich so jedoch mit mehreren Payment-Modellen und Anmeldeprozessen auseinandersetzen. Folglich verhindert die fehlende technologische Integration, dass Kunden beim Wechsel zwischen Allianzen eine nahtlose Customer Journey erleben können. Um von den Vorteilen einer Allianz profitieren zu können, müssen Kunden dieser eine vollständige Transparenz über ihre persönlichen Daten geben.

Heterogene Allianzen großer Händler und Hersteller entwickeln eigene voice-assisted Interfaces. Diese beruhen auf verschiedenen Standards und sind nicht allianzübergreifend kombinierbar.



Old Europe

Im Szenario „Old Europe“ verhindern **Datenschutz und eine starke Marktregulierung, dass sich ein integrierter Voice Commerce durchsetzen kann.**

Hier sind durch eine Verschärfung von Kartellrecht und Datenschutz keine übergreifenden Voice-Lösungen möglich. Der Markt für Voice Assistants besteht aus nicht integrierten Silo-Lösungen einzelner Teilnehmer, die den Verbrauchern nur einen marginalen Mehrwert bieten.

In dieser wettbewerbsintensiven Welt haben die Marktteilnehmer isolierte Voice-Commerce-Lösungen entwickelt, die weder mit anderen Technologien kompatibel noch geräteübergreifend integrierbar sind. Die stärksten Treiber dieser Realität sind strenge rechtliche Datenschutzvorschriften und eine starke Marktregulierung, die Monopole vermeiden soll.

Die wichtigsten Auswirkungen auf Handel, Hersteller und Verbraucher

- Erfolgreiche **Einzelhändler** haben in diesem Szenario frühzeitig begonnen, stark in ihre eigenen Voice Commerce-Lösungen zu investieren. Um die richtige Technologie zu entwickeln und sich im Markt durchzusetzen, mussten sie jedoch hohe Aufwände und Kosten bewältigen. Aufgrund einer fehlenden technologischen Integration und Zusammenarbeit zwischen den Marktteilnehmern müssen sich die Händler auf mehrdimensionale Kundenbeziehungen konzentrieren, um Customer Experience, Convenience und Preise zu optimieren.
- Die großen **Hersteller** haben ebenfalls jeweils eigene Voice Assistants im Einsatz, um direkte Vertriebskanäle aufzubauen und kontinuierlich mit ihren Endverbrauchern zu interagieren. Diese vertikale Integration stellt für die Hersteller jedoch eine komplexe Herausforderung dar. Deswegen konnten sich nur diejenigen durchsetzen und ihre Positionierung im Markt stärken, die ein breites Produktportfolio besitzen und stark in Technologie und Markenaufbau investiert haben.
- Die **Verbraucher** werden hier mit einer komplexen Welt von separaten Voice Assistants konfrontiert, die ihnen nur einen begrenzten Nutzen bieten. So sind personalisierte Einkäufe zwar möglich, allerdings nur in begrenztem Umfang, da kein Datenaustausch zwischen den Marktteilnehmern stattfindet. Zudem verwendet der typische Kunde nicht alle verfügbaren Voice Assistants, sondern konzentriert sich auf die für ihn relevantesten.

Dieses Szenario wird weitere gravierende Auswirkungen auf den Markt haben. Aufgrund hoher Komplexität und Wechselkosten wird sich mittelfristig je Segment nur ein präferierter Anbieter durchsetzen können. Infolgedessen werden auf lange Sicht nur einige wenige große Player den Markt dominieren, sodass die Versuche, durch eine starke Marktregulierung Monopole zu verhindern, letztendlich scheitern.

Der Markt für Voice Assistants besteht aus nicht integrierten Silo-Lösungen einzelner Marktteilnehmer und liefert Verbrauchern nur marginalen Mehrwert.



World of Aggregators

Im Szenario „World of Aggregators“ sind Verbraucher und Anbieter stark abhängig von wenigen großen Tech-Firmen, die den Marktzugang als Gatekeeper kontrollieren.

Dieses Szenario kennzeichnet geschlossene, technologisch integrierte Systeme, die als Oligopole von den großen Technologiefirmen kontrolliert werden. Dabei agieren die Big Player als Gatekeeper zwischen Käufern und Verkäufern und haben somit einen maßgeblichen Einfluss auf Angebot und Nachfrage.

In dieser Welt haben sich die großen Technologie-Provider durchgesetzt und das Einkaufen durch ihre geschlossenen Systeme von sprachgestützten Schnittstellen neu definiert. Dabei sind zwar Spracherkennungstechnologien stark integriert, die Zusammenarbeit zwischen den Marktteilnehmern ist jedoch schwach. Erfolgreiche Tech-Firmen haben die Voice-Commerce-Standards gesetzt und sind zum Gatekeeper der Shopping-Welt geworden. Big Data Analytics bietet diesen Plattform-Aggregatoren eine vollständige Kundentransparenz. Sie nutzen diese, um über datenbasiertes Marketing und prädiktive Serviceangebote eine größere Kundenbasis zu gewinnen.

Die wichtigsten Auswirkungen auf Handel, Hersteller und Verbraucher

- Der **Handel** musste hier erhebliche Entscheidungskompetenzen und die Verantwortlichkeit für kritische Fragestellungen, wie den Umgang mit Daten und IT-Sicherheit, an die Plattformanbieter abtreten. Einige Aggregatoren haben damit begonnen, ihre vertikale Reichweite auszubauen und somit die neuen dominierenden Händler zu werden.
- Die **Hersteller** konzentrieren ihre Marketing- und Vertriebsaktivitäten auf die Technologie-Aggregatoren, um ihre Weiterempfehlungsraten bei Voice Assistants zu optimieren. Gleichzeitig üben der verschärfte Wettbewerb im Markt und die hohe Transparenz durch die Plattformen einen hohen Druck auf Preise und Margen aus.
- Eine Vielzahl von Waren – von Lebensmitteln bis hin zu Autos – ist für die **Verbraucher** leicht zugänglich, effizient gebündelt und auf einer einzigen Plattform verfügbar. Die vollständig integrierten Technologien erhöhen zwar den Komfort der Endkunden, schränken aber den Datenschutz ein. Bei der Produktauswahl nehmen die Verbraucher eine immer passivere Rolle ein, da das Einkaufsverhalten durch Analytics-getriebene und individualisiert optimierte Angebote stark verzerrt wird.

Die großen Technologiefirmen agieren über ihre geschlossenen, aber technisch integrierten Lösungen als Gatekeeper zwischen Anbietern und Verbrauchern.

Die vier Szenarien aus Sicht der Verbraucher

Die vier beschriebenen Szenarien für die Zukunft des Voice Commerce sind in ihren Ausprägungen sehr unterschiedlich. Das heißt, jedes einzelne Szenario hat für die Konsumenten andere Konsequenzen.

Der Blick auf ein konkretes Fallbeispiel macht das deutlich. Sehen wir uns den Tagesablauf von Juliette an. Sie wohnt in München im Jahre 2030 und hat je nach Szenario ein komplett unterschiedliches Einkaufserlebnis. Dies beginnt beim Kauf der Geräte, mit denen Voice Commerce möglich ist, und der Anmeldung. Je nach Szenario stößt Juliette entweder sofort auf Hindernisse, oder sie erfährt eine einfache und bequeme Shopping Experience.

Die folgende Übersicht zeigt, wie einfach oder kompliziert das Leben für Juliette in den vier Szenarien ist, wenn sie Voice Assistants für verschiedene typische Aufgaben einsetzt. Was passiert etwa, wenn sie am Morgen keine Kaffeekapseln mehr vorrätig hat? Was soll sie für den Tag anziehen und kann sie dabei auch auf neue Produkte zurückgreifen? Und zuletzt: Wie kann sie über Voice Assistants ihren Tag planen? Egal, ob ein kaputtes Auto in die Werkstatt muss, es Termine zu koordinieren gibt oder sie einfach nur ihre Smart-Home-Geräte steuern möchte?

Aus Verbrauchersicht bietet das Szenario „Old Europe“ den geringsten und das Szenario „The new Internet“ den größten Mehrwert. So ist Juliette bei „Old Europe“ mit ihren unzähligen Accounts, Voice Assistants und Zahlungsmethoden überfordert und muss alles einzeln genehmigen. Zudem muss sie alle Einkäufe und Aufgaben des Alltags mit verschiedenen Systemen und Geräten verwalten. Und individualisierte Vorschläge stehen ihr nicht zur Verfügung.

Ganz anders sieht es für Juliette im Szenario „The new Internet“ aus. Hier kann sie mit einem einzigen Account ganz bequem auf alle Voice-integrierten Geräte zugreifen und alle Aufgaben über ein einziges System steuern. Außerdem machen prädiktive Bestellungen das Leben leichter und Juliette kann sich über individuell auf ihre Bedürfnisse zugeschnittene Empfehlungen freuen.

Abb. 22 – Die Zukunft von Voice Commerce – Sicht der Verbraucher

	Gerät kaufen und anmelden	Kaffeekapseln kaufen	Outfit auswählen/ bestellen	Tag planen
 <p>The new Internet „Eine Plattform für alles“</p>	<ul style="list-style-type: none"> • Ein Account • Alle Voice-integrierten und vernetzten Geräte möglich, egal welchen Anbieters • Vernetzung aller Geräte 	 <ul style="list-style-type: none"> • Prädiktive Bestellungen • Oder einfach mit Sprachbefehl • System kennt passende Marke und bevorzugte Sorte • Schnelle Lieferung 	 <ul style="list-style-type: none"> • Vorschläge basierend auf Bestand, Geschmack und Wetter • Individualisiertes Shopping 	<ul style="list-style-type: none"> • Alles über ein System steuern • Automatisierter Service • Abgestimmte Tagesplanung
 <p>Old Europe „Kaum Mehrwert“</p>	<ul style="list-style-type: none"> • Unzählige Accounts • Alles einzeln genehmigen • Verschiedene Zahlungsmethoden • Viele einzelne Voice Assistants nötig 	<ul style="list-style-type: none"> • Kauf auf Sprachbefehl an zugehöriges Gerät • Richtiges Gerät muss angesprochen werden 	<ul style="list-style-type: none"> • Vorschlag aus Bestand • Neuvorschläge nur von bereits gekauften Marken • Keine individualisierten Vorschläge 	<ul style="list-style-type: none"> • Aufgaben des Alltags sind einzeln über verschiedene Systeme und Geräte zu verwalten
 <p>World of heterogenic Alliances „Wer die Wahl hat, hat die Qual!“</p>	<ul style="list-style-type: none"> • Ein bis mehrere Accounts • Zur Allianz gehörige Geräte, nur von spezifischen Anbietern • Vernetzung aller Geräte innerhalb der Allianz 	<ul style="list-style-type: none"> • Kauf auf Sprachbefehl an zur Allianz gehörendes Gerät • System kennt nur zur Allianz gehörende Marken • Fehlerhafte Bestellung möglich 	<ul style="list-style-type: none"> • Vorschläge basierend auf Bestand • Individualisiertes Shopping innerhalb der Allianz-Marken 	<ul style="list-style-type: none"> • Aufgaben des Alltags über jeweilige Allianz(en) steuern • Keine allianzübergreifende Synchronisation
 <p>World of Aggregators „Der neue Handel“</p>	<ul style="list-style-type: none"> • Ein Account • Voice Assistants des Aggregators und seiner Partner möglich 	 <ul style="list-style-type: none"> • Prädiktive Bestellungen • Oder einfach mit Sprachbefehl • System bestellt passende Eigenmarke (wenn möglich) • Schnelle Lieferung 	 <ul style="list-style-type: none"> • Vorschläge basierend auf Bestand, Geschmack, aber verzerrt durch Vorselektion • Individualisiertes Shopping und schnelle Lieferung 	<ul style="list-style-type: none"> • Aufgaben des Alltags über ein System steuern • Abgestimmte Tagesplanung

 Produkt entspricht genau den Kundenwünschen

 Lieferung genau zum richtigen Zeitpunkt

Quelle: Monitor Deloitte 2018.

Auswirkungen der Szenarien auf Verbraucher, Händler und Hersteller

Implikationen für Verbraucher

Wie wirken sich die unterschiedlichen Szenarien nun konkret auf Kunden, Hersteller und Händler aus? Welche speziellen Vor- und Nachteile weisen sie für die unterschiedlichen Marktteilnehmer auf?

Betrachtet man die Implikationen für die Kunden, so bieten Szenarien mit hoher technologischer Integration einen größeren Mehrwert für den Verbraucher als andere. Denn hier können die Nutzer von durchgängigen Customer Journeys mit individualisierten Angeboten profitieren.

Die Implikationen, die die vier Szenarien auf die Kunden haben, lassen sich am besten anhand der folgenden vier kundenrelevanten Kategorien bewerten: Convenience, Auswahl, Shopping Experience und Freigabe persönlicher Daten. Dabei steht „Convenience“ für die Einfachheit der Nutzung, „Auswahl“ beschreibt die Breite des verfügbaren Produktportfolios. Im Bereich „Shopping Experience“ werden Themen wie die Möglichkeit der Individualisierung oder eine nahtlose Customer Journey bewertet. Die letzte Kategorie beschäftigt sich mit der Freigabe persönlicher Daten.

Den größten Mehrwert für die Kunden bieten die Szenarien „The new Internet“ und „World of Aggregators“. Hier können Verbraucher die Vorteile des Voice Commerce voll und ganz für sich nutzen. Allerdings müssen sie dafür ihre persönlichen Daten freigeben. In der „World of Aggregators“ werden die Kunden jedoch bei der Auswahl von Produkten stärker von den Tech-Firmen beeinflusst als bei „The new Internet“.

Nicht so reibungslos verläuft der Voice Commerce dagegen in den verbleibenden beiden Szenarien. So bietet „Old Europe“ dem Verbraucher keinen großen Mehrwert und ist mit viel administrativem Aufwand verbunden. Auch die „World of heterogenic Alliances“ bietet kein attraktives Einkaufserlebnis. Hier müssen die Kunden multiple Voice Assistants verwenden, um das gesamte Leistungs- und Produktportfolio nutzen zu können.

Abb. 23 – Auswirkung der Szenarien | Implikationen für Verbraucher

Verbraucher	 The new Internet	 World of heterogenic Alliances	 Old Europe	 World of Aggregators
Convenience	<p>5</p> <ul style="list-style-type: none"> • Alles auf einer Plattform • Über Geräte und Technologien integrierte Voice-Systeme 	<p>2</p> <ul style="list-style-type: none"> • Keine technische Integration über Allianzen hinweg • Mehrere Geräte notwendig 	<p>1</p> <ul style="list-style-type: none"> • Keine technische Integration • Viele einzelne Silo-Lösungen mangels übergreifender Lösung 	<p>5</p> <ul style="list-style-type: none"> • Integrierte Voice-Systeme • Alles auf einer Plattform
Auswahl	<p>5</p> <ul style="list-style-type: none"> • Breite Auswahl, die auf Basis der persönlichen Präferenzen optimiert wurde 	<p>2</p> <ul style="list-style-type: none"> • Verbraucher beschränken sich auf das Produktportfolio der ausgewählten Allianz(en) 	<p>1</p> <ul style="list-style-type: none"> • Verbraucher können nur die Produkte des Händlers/Herstellers bestellen, dessen Voice Assistant sie nutzen 	<p>4</p> <ul style="list-style-type: none"> • Breite prädiktive Auswahl • Auswahl durch Tech-Firmen vorgefiltert und beeinflusst
Shopping Experience	<p>5</p> <ul style="list-style-type: none"> • Hohe Individualisierung der Shopping Experience, die auf die Bedürfnisse des Verbrauchers abgestimmt ist 	<p>2</p> <ul style="list-style-type: none"> • Keine nahtlose Customer Journey • Lästiger Wechsel zwischen Allianzen 	<p>2</p> <ul style="list-style-type: none"> • Keine durchgehende Customer Journey • Viel Administration aus Datenschutzgründen 	<p>5</p> <ul style="list-style-type: none"> • Hohe Individualisierung der Shopping Experience, die auf die Bedürfnisse des Verbrauchers abgestimmt ist
Freigabe persönlicher Daten	<p>1</p> <ul style="list-style-type: none"> • Der gläserne Kunde ist Realität • Freigabe der persönlichen Daten zur Nutzung von Voice Commerce 	<p>2</p> <ul style="list-style-type: none"> • Daten werden nur an ausgewählte Allianzen weitergegeben 	<p>4</p> <ul style="list-style-type: none"> • Personengebundene Daten dürfen nur sehr begrenzt genutzt und gespeichert werden 	<p>1</p> <ul style="list-style-type: none"> • Der gläserne Kunde ist Realität • Tech-Firmen haben alle Daten und nutzen diese effektiv
Fazit	<p>5</p> <ul style="list-style-type: none"> • Verbraucher können Vorteile von Voice Commerce voll nutzen, müssen dafür ihre persönlichen Daten freigeben 	<p>2</p> <ul style="list-style-type: none"> • Verbraucher müssen multiple Voice Assistants nutzen, um das gesamte Leistungs-/Produktportfolio zu nutzen 	<p>1</p> <ul style="list-style-type: none"> • Voice bietet keinen großen Mehrwert und ist mit viel administrativem Aufwand verbunden 	<p>4</p> <ul style="list-style-type: none"> • Verbraucher können Vorteile von Voice Commerce nutzen, werden jedoch stark von Tech-Firmen beeinflusst, die alle Daten haben

1 sehr negative Auswirkungen
4 positive Auswirkungen

2 eher negative Auswirkungen
5 sehr positive Auswirkungen

3 neutral / kaum Auswirkungen

Implikationen für Händler und Hersteller

Jedes einzelne Szenario weist unterschiedliche Implikationen für Hersteller und Händler auf. Diese können sich zudem noch – je nach Industrie und Positionierung des jeweiligen Marktteilnehmers – weiter verändern.

The New Internet

Hier handelt es sich um eines der besten Szenarien für Hersteller. Der zugängliche Direktvertrieb bietet für sie die Chance, die Händler zu umgehen und damit höhere Margen zu erzielen. Das ist insbesondere für große Hersteller von Vorteil, da sie Skaleneffekte nutzen können. Vor allem im Non-Food-Bereich können die Produzenten vom direkten Endkundenkontakt profitieren. Zudem haben A-Brands durch ihre bestehende Markenpräsenz Vorteile.

Dagegen ist dieses Szenario für Händler – vor allem im Food Bereich – eher negativ zu bewerten. Hier ist eine Kanalverschiebung hin zum teuren Versandhandel zu erwarten und der Markt wird von einem enormen Preisdruck geprägt sein. Deswegen wird die Fulfillment-Kompetenz immer mehr zum

wichtigen Wettbewerbsfaktor. Erfolgsscheidend werden auch die Kundenbindung sowie die analytischen Fähigkeiten der Händler sein, um zum Beispiel A.I.-generierte Warenkörbe zu erstellen, die den Kunden ein einfaches und komfortables Einkaufen ermöglichen. Im Bereich der Non-Food-Händler wird die Bedeutung von Customer Experience und Kuratation als weitere wichtige Wettbewerbsfaktoren zunehmen.

World of heterogenic Alliances

Dieses Szenario hält für Hersteller die größten Schwierigkeiten bereit. Hier wird sich die Auswahl einer starken Allianz als überlebenswichtig erweisen, was jedoch aufgrund eines volatilen und undurchsichtigen Marktes schwierig ist. Wer versucht, dieses Problem mit der Präsenz in meh-

Abb. 24 – Auswirkungen der Szenarien | Implikationen für Hersteller

Hersteller	The new Internet	World of heterogenic Alliances	Old Europe	World of Aggregators
 Food	<ul style="list-style-type: none"> • Direktvertrieb einfacher durch geringere Macht stationärer Händler und bessere Marge • Insbesondere für große Hersteller vorteilhaft durch Skaleneffekte/Sortimentsbreite • A-Brands im Vorteil <p>4</p>	<ul style="list-style-type: none"> • Auswahl der stärksten Allianz überlebenswichtig, was aufgrund eines volatilen und undurchsichtigen Marktes schwierig ist • Teilnahme an mehreren Allianzen führt zu erhöhten Kosten und Aufwand <p>1</p>	<ul style="list-style-type: none"> • Voice Commerce wird nur für breite Produktportfolios genutzt, daher eher Händler-Apps, insbesondere im niedrigpreisigen Low-Involvement-Bereich <p>2</p>	<ul style="list-style-type: none"> • Direktvertrieb einfacher möglich • Kanalverschiebung zu Tech-Providern, d.h., mühsam aufgebaute Allianzen mit Händlern bröckeln • Marketingbudgetverschiebung zu Recommender Marketing • Besseres Targeting (kostet) <p>4</p>
 Non-Food	<ul style="list-style-type: none"> • Direktvertrieb einfacher durch geringere Macht stationärer Händler und bessere Marge • Insbesondere für große Hersteller vorteilhaft durch Skaleneffekte/Sortimentsbreite • A-Brands im Vorteil <p>4</p>	<ul style="list-style-type: none"> • Teilnahme an mehreren Allianzen führt zu erhöhten Kosten und Aufwand • Geringe Relevanz für High-Involvement-Produkte durch fehlenden Cross-Media-Einsatz <p>2</p>	<ul style="list-style-type: none"> • Voice Commerce wird nur für breite Produktportfolios genutzt, daher eher Händler-Apps, insbesondere im niedrigpreisigen Low-Involvement-Bereich <p>2</p>	<ul style="list-style-type: none"> • Direktere Kundeninteraktion • Kanalverschiebung von Händlern zu Tech-Providern, d.h., mühsam aufgebaute Allianzen mit Händlern bröckeln • Starkes Branding schützt <p>4</p>

1 sehr negative Auswirkungen **2** eher negative Auswirkungen **3** neutral / kaum Auswirkungen
4 positive Auswirkungen **5** sehr positive Auswirkungen

renen Allianzen zu umgehen, muss hohe Kosten und einen enormen Aufwand in Kauf nehmen. Im Non-Food-Bereich ist durch fehlende Cross-Media-Einsätze eine geringe Relevanz für High-Involvement-Produkte zu erwarten.

So schlecht dieses Szenario auch für die Hersteller sein mag, so positiv stellt es sich vor allem für große Händler dar. Sie haben die Chance, das Kernelement einer Allianz zu werden, und können von einer Verschiebung der Marktanteile profitieren.

Old Europe

Auch dieses Szenario birgt für die Hersteller große Herausforderungen. Hier wird der Voice Commerce nur für breite Produktportfolios genutzt. Deswegen werden in „Old Europe“ eher Händler-Anwendungen

im Vordergrund stehen, insbesondere im niedrigpreisigen Low-Involvement-Bereich.

Für Händler ist dieses rückständige Szenario sogar als positiv zu bewerten. Der Voice Commerce bietet ihnen hier die Möglichkeit, ihre Eigenmarken stärker zu platzieren, was höhere Margen mit sich bringt.

World of Aggregators

Für Hersteller ist dies ein positives Szenario. Ähnlich wie bei „The New Internet“ ist für sie auch hier der Direktvertrieb wesentlich einfacher möglich. Allerdings beginnen die mühsam aufgebauten Beziehungen zu Händlern durch eine Kanalverschiebung hin zu Tech-Providern immer mehr abzubrockeln. Durch eine rechtzeitige Marketingbudgetverschiebung in Richtung Recommender Marketing kön-

nen die Hersteller jedoch langfristig sehr positive Effekte erzielen. Denn auf diese Weise ist ein zwar kostenintensives, aber dennoch effizienteres und besseres Targeting möglich. Zudem schützt eine starke Marke vor allem im Non-Food-Bereich vor Umsatzeinbrüchen.

Für Händler ist die World of Aggregators dagegen das Szenario mit den größten Herausforderungen. Schließlich verlieren sie hier die wertvolle digitale Kundenschnittstelle an die Tech-Provider. Zudem wird ein stark erhöhter Preisdruck zu geringeren Margen führen. Um langfristig erfolgreich und profitabel im Markt zu bleiben, sind für Händler neue Geschäftsmodelle notwendig.

Abb. 25 – Auswirkungen der Szenarien | Implikationen für Händler

Händler	The new Internet	World of heterogenic Alliances	Old Europe	World of Aggregators
 Food	2 <ul style="list-style-type: none"> Starke Kanalverschiebung zu teurem Versandhandel Erhöhter Preisdruck Fulfillment-Kompetenzen als Wettbewerbsfaktor Hohe Anforderungen an Analytik, u.a. A.I.-Baskets 	4 <ul style="list-style-type: none"> Große Händler können Kernelement einer Allianz sein Profitieren durch Allianz (Partizipation an Verschiebung von Marktanteilen) 	4 <ul style="list-style-type: none"> Voice Commerce bietet die Möglichkeit, Eigenmarken stärker zu platzieren und dadurch die Marge zu erhöhen 	1 <ul style="list-style-type: none"> Händler verlieren digitale Kundenschnittstelle an Tech-Provider Erhöhter Preisdruck führt zu geringeren Margen Neue Geschäftsmodelle notwendig
 Non-Food	4 <ul style="list-style-type: none"> Multimodaler Technologieeinsatz stärkt Versandhandel Erhöhter Preisdruck Kundenbindung wichtig CX und Kuratation als Wettbewerbsfaktor 	4 <ul style="list-style-type: none"> Große Händler können Kernelement einer Allianz sein Profitieren durch Allianz (Partizipation an Verschiebung von Marktanteilen) 	4 <ul style="list-style-type: none"> Voice Commerce bietet die Möglichkeit, Eigenmarken stärker zu platzieren und dadurch die Marge zu erhöhen 	1 <ul style="list-style-type: none"> Händler verlieren digitale Kundenschnittstelle an Tech-Provider Erhöhter Preisdruck führt zu geringeren Margen Neue Geschäftsmodelle notwendig



Entwicklung von Voice- und E-Commerce

Die Integration von Voice mit anderen Technologien führt nicht nur zu höheren Voice-Commerce-Anteilen, sondern auch zu einem starken Wachstum von E-Commerce gegenüber dem stationären Handel.

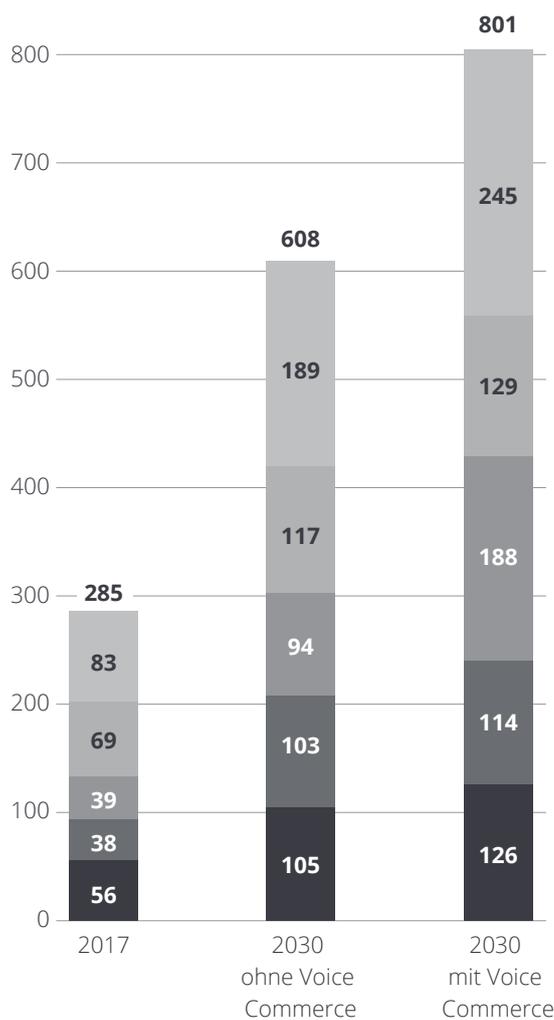
Der künftige Erfolg von Voice Commerce hängt auch stark von der Voice-Integration in andere Technologien ab. Je stärker der Grad der Integration, desto höher werden die Umsätze sein. Dabei erweist sich Voice als Wachstumstreiber für den E-Commerce. Die Hauptgründe dafür sind, dass die Integration von Voice in andere Technologien,

wie etwa die Bilderkennung, ein besseres Einkaufserlebnis für den Kunden schafft. Das erhöht nicht nur die Voice-Commerce-Anteile, sondern fördert auch generell das Wachstum von E-Commerce gegenüber dem stationären Handel. Dagegen werden nicht integrierte Voice-Systeme als isolierte Insellösungen nur wenig wettbewerbsfähig gegenüber dem stationären Handel und anderen E-Commerce-Systemen sein.

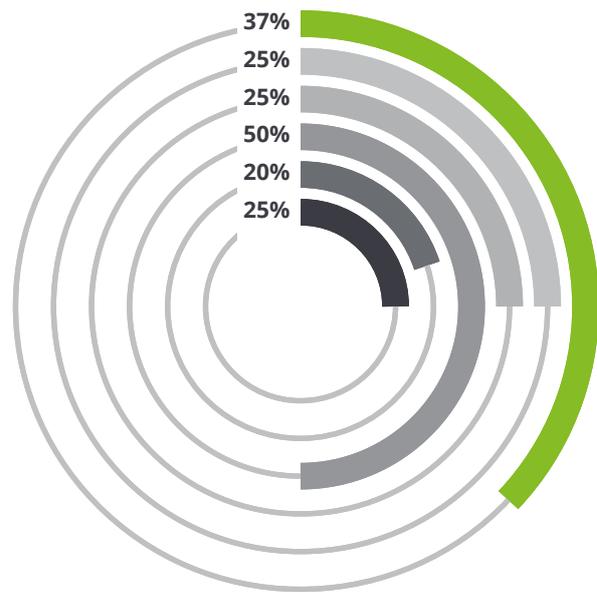
Integrierte Voice-Systeme: Von 2017 bis 2030 wächst der E-Commerce angetrieben durch Voice Commerce um 181 Prozent.

Experten schätzen, dass bei der Integration von Voice mit anderen Technologien der E-Commerce-Umsatz von 2017 bis 2030 um 181 Prozent wachsen wird. Dabei entfallen auf Voice Commerce sogar 37 Prozent der E-Commerce-Umsätze. Untersucht man einzelne Produktgruppen, so wird der Voice-Commerce-Anteil im Jahr 2030 bei Lebensmitteln und Körperpflege mit 50 Prozent am größten sein, gefolgt von Mode, Electronics und Medien sowie Spielzeug, Hobby, Heimwerken, die alle einen Voice-Umsatzanteil von jeweils 25 Prozent für sich verbuchen können.

Abb. 26 - Marktübersicht Europa (technologisch integriert, Umsatz in Mrd. €)



Anteil Voice Commerce am E-Commerce Umsatz



	Mode	Elektronik und Medien	Lebensmittel und Körperpflege	Möbel und Haushaltsgeräte	Spielzeug, Hobby, Heimwerken	Gesamt
CAGR 2017-2030 ohne Voice	6,0%	6,5%	4,2%	7,1%	8,0%	2,9%
CAGR 2017-2030 mit Voice	8,3%	8,7%	4,9%	12,9%	8,8%	6,4%

- Mode
- Elektronik und Medien
- Lebensmittel und Körperpflege
- Möbel und Haushaltsgeräte
- Spielzeug, Hobby, Heimwerken
- Gesamt

Quellen: Statista (E-Com Markt bis 2025); Monitor Deloitte Expertenschätzung (E-Com Markt 2025-2030, Voice Impact)

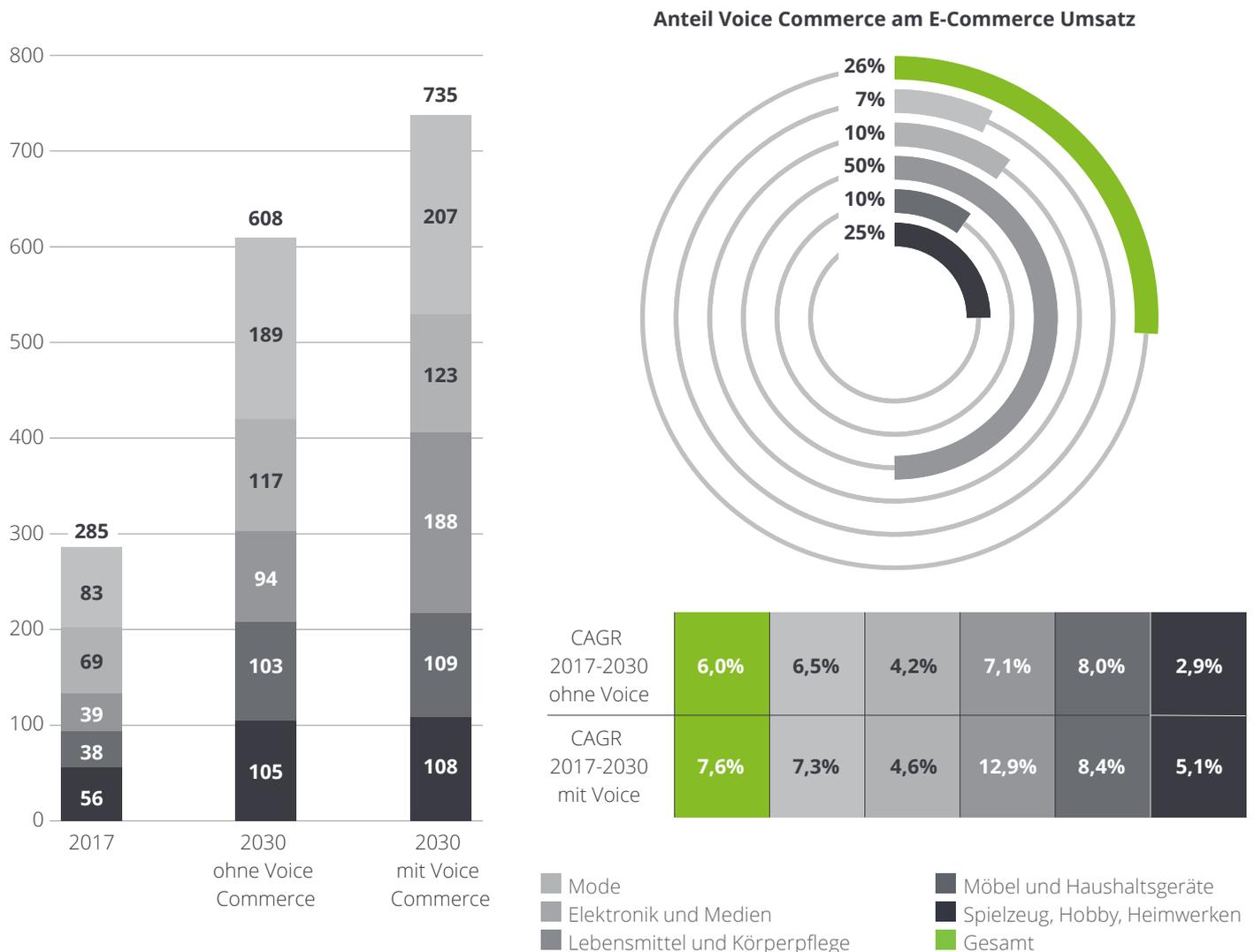


**Nicht integrierte Voice-Systeme:
Von 2017 bis 2030 wächst der E-Commerce angetrieben durch Voice Commerce um 158 Prozent**

Wird Voice Commerce dagegen als nicht integrierte Insellösung eingesetzt, dann wird der E-Commerce-Umsatz von 2017 bis 2030 lediglich um 158 Prozent zunehmen. Auch der Voice-Commerce-Marktanteil im Jahr 2030 wird hier mit 26 Prozent wesentlich niedriger ausfallen als bei integrierten Lösungen.

Das bedeutet, dass nicht integrierte Voice-Systeme nur wenig wettbewerbsfähig gegenüber dem stationären Handel und anderen E-Commerce-Systemen sein werden. Sie treiben den E-Commerce-Markt nur bedingt und führen zu geringeren Voice-Commerce-Umsätzen.

Abb. 27 – Marktübersicht Europa (technologisch integriert, Umsatz in Mrd. €)



Quellen: Statista (E-Com Markt bis 2025); Monitor Deloitte Expertenschätzung (E-Com Markt 2025-2030, Voice Impact)

Kritische Treiber und Barrieren für den Erfolg von Technologiefirmen

Für den Erfolg von Technologiefirmen im Bereich Voice Assistants spielen mehrere kritische Faktoren eine wichtige Rolle, die als Treiber oder Barrieren wirken können.

Im Folgenden werden acht wesentliche Faktoren genauer beleuchtet: das Henne-Ei-Problem, die Pfadabhängigkeit, die technische Integrationsfähigkeit, Technologievorsprung, Kooperationen, Datensicherheit, Regulatorik sowie die öffentliche Meinung.

Abb. 28 – Erfolgsfaktoren



Quelle: Monitor Deloitte 2018.



Henne-Ei-Problem
Ohne Anwendungen keine User – ohne User kein Anreiz für Dritte, Anwendungen zu programmieren

Das Henne-Ei-Problem beschreibt die schwierige Situation, dass externe Anbieter keine Anwendungen (Skills/Actions/Apps) für eine bestimmte Plattform programmieren, wenn diese nicht von einer relevanten Masse an Kunden genutzt wird. Nutzer verschreiben sich wiederum keiner Plattform, die keinen Mehrwert aufgrund eines zu geringen Angebotes an attraktiven Anwendungen bietet. Diesen Teufelskreis gilt es zu durchbrechen.

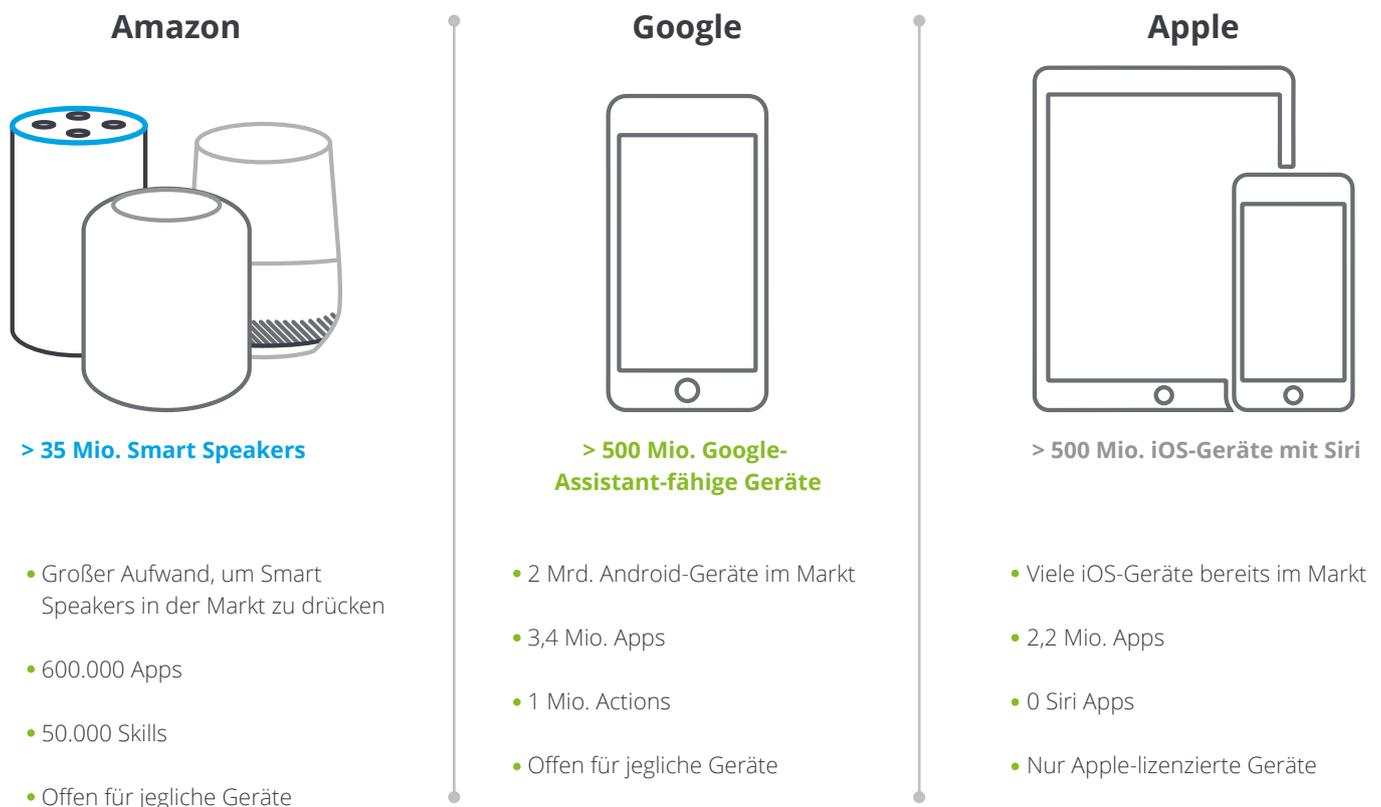
Google und Apple haben hier klar einen Vorsprung vor anderen Technologiefirmen, da deren sprachassistentenfähigen Geräte global bereits von einer breiten Kundenanzahl genutzt werden. Allerdings beschränkt Apple den Einsatz seines Voice Assistant Siri auf die eigenen Geräte. Dadurch wird die Anzahl der von Dritten entwickelten Anwendungen für Siri stark limitiert.

Deswegen gehen Amazon und Google im Markt für Voice Assistants deutlich vor Apple in Führung. Denn beide Player ermöglichen Drittfirmen, ihre Software zu nutzen, und gewinnen dadurch eine zusätzliche, von deren eigener Hardware

unabhängige Nutzerbasis. Zudem baut Amazon seine Kundenzahl durch den Verkauf von Smart Speakers zu geringen Preisen und mithilfe eines intensiven Marketings massiv aus.

Führende Technologieanbieter lösen das Henne-Ei-Problem also vor allem dadurch, dass sie auf ihre bereits existierende große Nutzerbasis bauen oder Hardware aggressiv in den Markt drücken.

Abb. 29 – Vergleich der Voice Assistant Nutzerbasis führender Technologieanbieter



Quelle: Monitor Deloitte 2018.



Pfadabhängigkeit

Wer als Erster eine Technologie als Standard etabliert, kann einen Lock-in-Effekt erreichen und den gesamten Markt für sich gewinnen.

Wenn sich bestimmte Technologiesysteme in einem Markt etablieren und als Standard durchsetzen, sodass sich dort neue Produkte, die auf anderen Technologiesystemen beruhen, nur schwer verbreiten können, wird von einer Pfadabhängigkeit gesprochen.

Etablierte Systeme und Technologiestandards stellen für Newcomer hohe Markteintrittsbarrieren dar. Denn um Kunden für sich zu gewinnen, müssen sie diese zu einem Systemwechsel bewegen. Doch der ist für Verbraucher mit Kosten und Aufwand verbunden, sei es für einen neuen Anschluss, die Inbetriebnahme oder die Einarbeitung in das neue System. Diese hohen Wechselkosten führen daher zu einem Lock-in-Effekt, der einen Wechsel erschwert und den Kunden an den bisherigen Anbieter bindet.

Im Rahmen dieser Wechselkosten bei der Einführung neuer Technologien ist der Effekt der gelernten Verhaltensänderung besonders relevant. Etabliert sich eine Technologie als Standard, wird ihre Handhabung für den Anwender intuitiv. Bei Neuentwicklungen oder anderen Anbietern mit unterschiedlicher Handhabung muss dieser Lernprozess erstmal durchlaufen werden, was für Nutzer mit Aufwand verbunden ist.

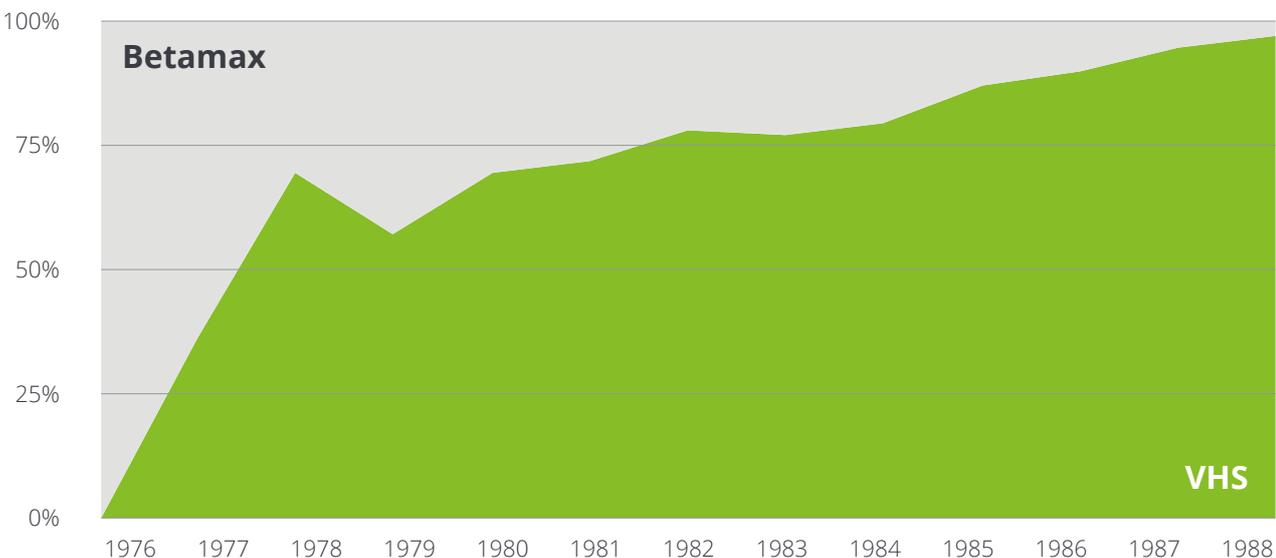
Eines der bekanntesten Beispiele zur Pfadabhängigkeit ist der Formatkrieg der Videorekordersysteme. Auf diesem Markt hat sich das Videokassettensystem VHS gegenüber dem qualitativ überlegenen System Betamax von Sony durchgesetzt. Die Gründe dafür lagen in besseren Allianzen und einer als höher wahrgenommenen Kompatibilität.

Aus diesem Beispiel lassen sich einige generelle und entscheidende Erfolgsfaktoren für standardsetzende Systeme ableiten. Zum einen wird der Standard nicht zwangsläufig vom First Mover gesetzt und zum anderen spielt auch die Qualität nicht immer die entscheidende Rolle. Vielmehr

kommt es auch auf andere Variablen wie die Befriedigung der Kundenbedürfnisse an. Zudem setzen Kunden eher auf langfristig funktionierende Systeme und legen je nach Ausrichtung auch Wert auf Dinge wie z.B. Preis, Kooperationen, Kompatibilitäten und Marketing. Auf dem Weg zu einer nachhaltigen Pfadabhängigkeit sind vor allem auch die Kooperationen entscheidend. Durch ein großes und starkes Netzwerk kann ein Lock-in-Effekt der Kunden gefördert werden.

Auf den Voice Assistant Markt übertragen, sind viele Parallelen zwischen VHS und Amazon zu erkennen. Auch der Tech-Konzern aus Seattle scheint eindeutig eine Pfadabhängigkeit anzustreben, indem er auf ein offenes System und eine große Anzahl von Anwendungsmöglichkeiten setzt. Außerdem betreibt Amazon eine aggressive Preispolitik und Marktpenetration. Haben Verbraucher die Anwendung von Alexa gelernt und ihr Verhalten darauf umgestellt, wird es für andere Player sehr schwer sein, am Markt eine relevante Rolle zu spielen.

Abb. 30 – Formatkrieg der Videorekorder



Quelle: Park, S. (2004). Quantitative analysis of network externalities in competing technologies: The VCR case, Review of Economics and Statistics, 86(4), 937-945.



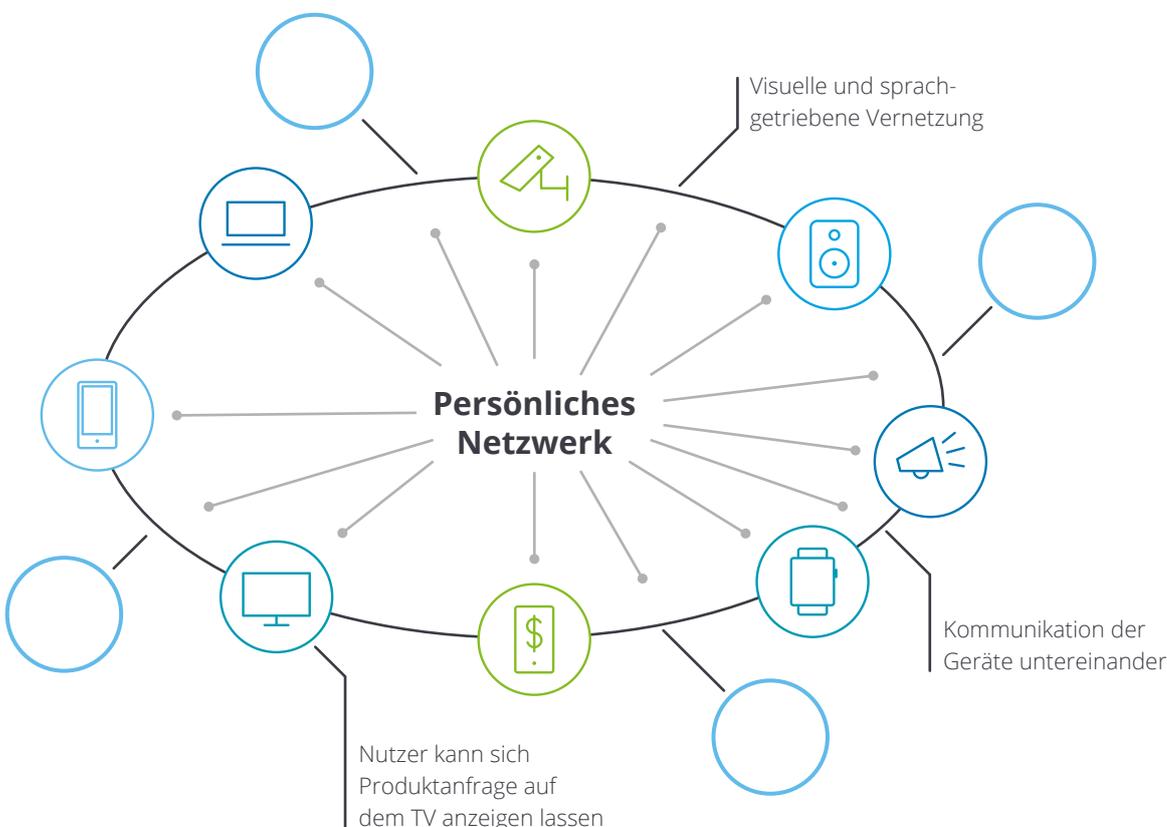
Technische Integrationsfähigkeit
Die Voice-Software muss einfach in externe Geräte/Systeme zu integrieren sein. Auch die Erstellung von Anwendungen durch Dritte muss unkompliziert möglich sein.

Die Integrationsfähigkeit von Voice in andere Technologien und Geräte ist für den Erfolg von Voice Commerce entscheidend. Denn dadurch erreicht die User Experience von Voice Assistants eine neue Dimension und der Mehrwert wird exponentiell erhöht. Die Voice-Integration kann in verschiedensten Geräten und über unterschiedliche Kanäle erfolgen. Beispiele dafür sind etwa Smartphones/Smart Speakers, Smart TV, Wearables, Tablets, PCs und viele weitere.

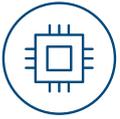
Zudem ist auch eine visuell getriebene Integration möglich, bei der Sprachsysteme mit Displayfunktionen ausgestattet werden. Dies macht eine zusätzliche Steuerung über Gestik möglich. Erfolgsentscheidend für die Sicherheit des Voice Commerce ist eine Integration zahlungs- und identitätsbestätigender Funktionen durch Biometrie, wie etwa Voice- und Gesichtserkennung.

Generell ist die Integrations- und Kompatibilitätsfähigkeit einzelner Systeme ein wichtiger kritischer Faktor, der den Gesamterfolg des Voice Commerce maßgeblich beeinflusst.

Abb. 31 – Überblick kanal- und systemübergreifender Systeme



Quelle: Monitor Deloitte 2018.



Technologievorsprung

Der Technologievorsprung ist ein entscheidender Wettbewerbsvorteil im Voice Commerce. Hier kommt es vor allem auf Spracherkennung, Artificial Intelligence und die Natürlichkeit der Konversation an.

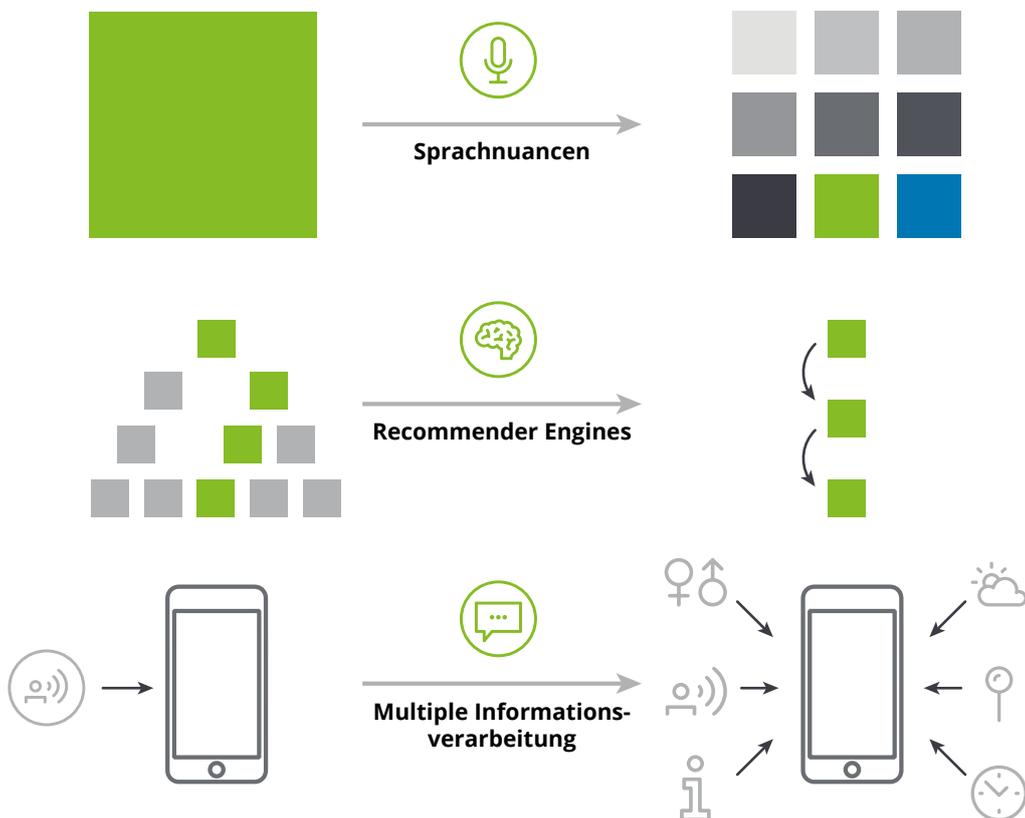
Der technologische Fortschritt ist ein wichtiger Treiber, der die User Experience verbessert und die Voice-Absatzkanäle attraktiver macht. Wer hier die Nase vorn hat, steigert ganz klar seine Attraktivität bei Kunden. Wer dagegen auf technologisch veraltete Systeme setzt, verliert an Wettbewerbsfähigkeit.

Dabei spielt die kontinuierliche Weiterentwicklung der Spracherkennung eine entscheidende Rolle. Wichtige Aspekte sind die Aufnahmequalität, die Fähigkeit der Software, Nuancen der Stimme zu erken-

nen, sowie die stimmabhängige Erkennung des Sprechers. Hinzu kommt die Natürlichkeit der Konversation. Dabei sollte eine situationsbedingte Informationssammlung kontextspezifische Konversationen möglich machen. Im Zentrum der Voice-Technologie steht die Artificial Intelligence. Hier sind die üblichen Logikbäume wegen der hohen Anzahl von Möglichkeiten keine Option. Der Fokus im Voice Commerce muss stattdessen klar auf Recommender Engines liegen.

All diese Features sind wichtige Differenzierungsmerkmale der Tech-Firmen im Wettbewerb um die Kunden. Wer technologisch bestens aufgestellt und stets auf dem neuesten Stand ist, erhöht die User Experience und gewinnt an Attraktivität. Natürlich steigert die Bedeutung des technischen Fortschritts für den Voice Commerce auch die Abhängigkeit von den großen Tech-Firmen, die auf diesen Gebieten führend sind.

Abb. 32 - Überblick Fortschritt Spracherkennung



Quelle: Monitor Deloitte 2018.



Kooperationen

Um im Wettbewerb das größte und beste Netzwerk aufzubauen, sind Kooperationen essenziell, hier am Beispiel Automobilindustrie dargestellt.

Für die großen Tech-Firmen spielen Kooperationen eine wichtige Rolle. Sie verfolgen damit das Ziel, die Nutzerbasis und die Anzahl integrierter Systeme weiter auszubauen. Ein Beispiel dafür ist das Bestreben der Tech-Firmen, im Automobilbereich Assistenzsysteme zur Verfügung zu stellen.

Die Systeme CarPlay von Apple und Android Auto von Google sind bereits bei allen namhaften Automobilherstellern platziert und ermöglichen so die Integration von Smartphones und Voice Assistants. Dagegen baut Amazon Allianzen zur Systemintegration auf und setzt mit Echo Auto auf die einfache Nachrüstung dieser Technologie durch den Nutzer. Dies ist in den Fahrzeu-

gen aller Hersteller und für jedes Baujahr möglich.

Der Einzug von Voice Assistants im Auto steigert den Mehrwert für die Verbraucher und fördert perspektivisch die Nutzung von Voice Commerce. Allerdings kommt es hier auch zu einer Erhöhung der Pfadabhängigkeit.

Beim Aufbau von Kooperationen ist es essenziell, die richtigen Geschäftsmodelle zu definieren und klare Governance-Strukturen aufzusetzen, damit alle relevanten Anspruchsgruppen optimal an den Vorteilen der Zusammenarbeit partizipieren. Die Governance muss hierbei einen rechtlichen und faktischen Ordnungsrahmen für die Leitung und Überwachung der Zusammenarbeit gewährleisten, wohingegen das Geschäftsmodell insbesondere die Monetarisierung der Leistungen definieren muss.

Abb. 33 – Kooperationsübersicht am Beispiel Automobil



Quelle: Monitor Deloitte 2018.



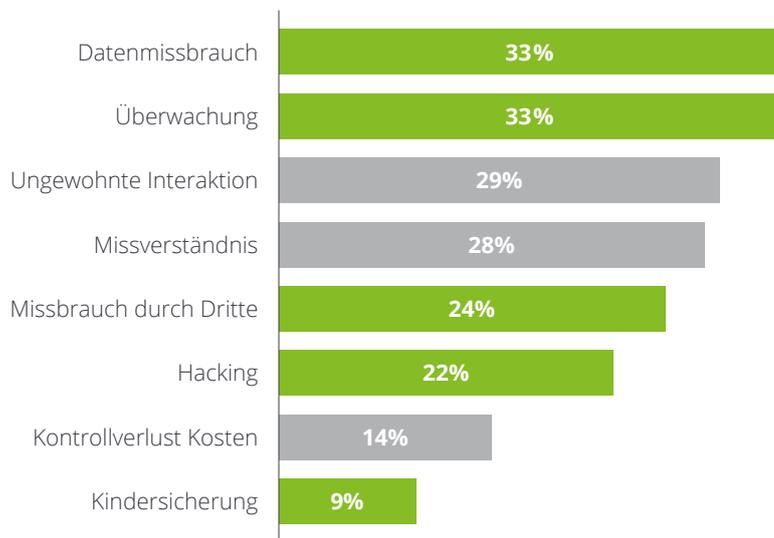
Datensicherheit

Die Gewährleistung der Datensicherheit ist auch bei Voice Commerce ein entscheidender Faktor.

Die Datensicherheit ist in fast allen Umfragen das Hauptbedenken der Verbraucher bei der Nutzung von Voice Assistants. Die Befürchtungen der Nutzer reichen von Datenmissbrauch über das Überwachen und Abhören durch andere bis hin zum

Hacking der smarten Anwendung über Sprachbefehle. Google hat die Thematik erkannt und informiert sehr ausführlich über die Sicherheit, Nutzung und Speicherung von Daten. Dagegen fährt Amazon in diesem Bereich eine eher zurückhaltende Kommunikationsstrategie. Für den Erfolg von Voice Assistants sind die technische Gewährleistung der Datensicherheit sowie ein aktives Aufklären der Nutzer zwingend notwendig.

Abb. 34 – Bedenken der deutschen Nutzer



Quelle: Bundesverband Digitale Wirtschaft (BVDW) e.V., Digital Trends: Umfrage zu Sprachassistenten, November 2017.



Regulatorik

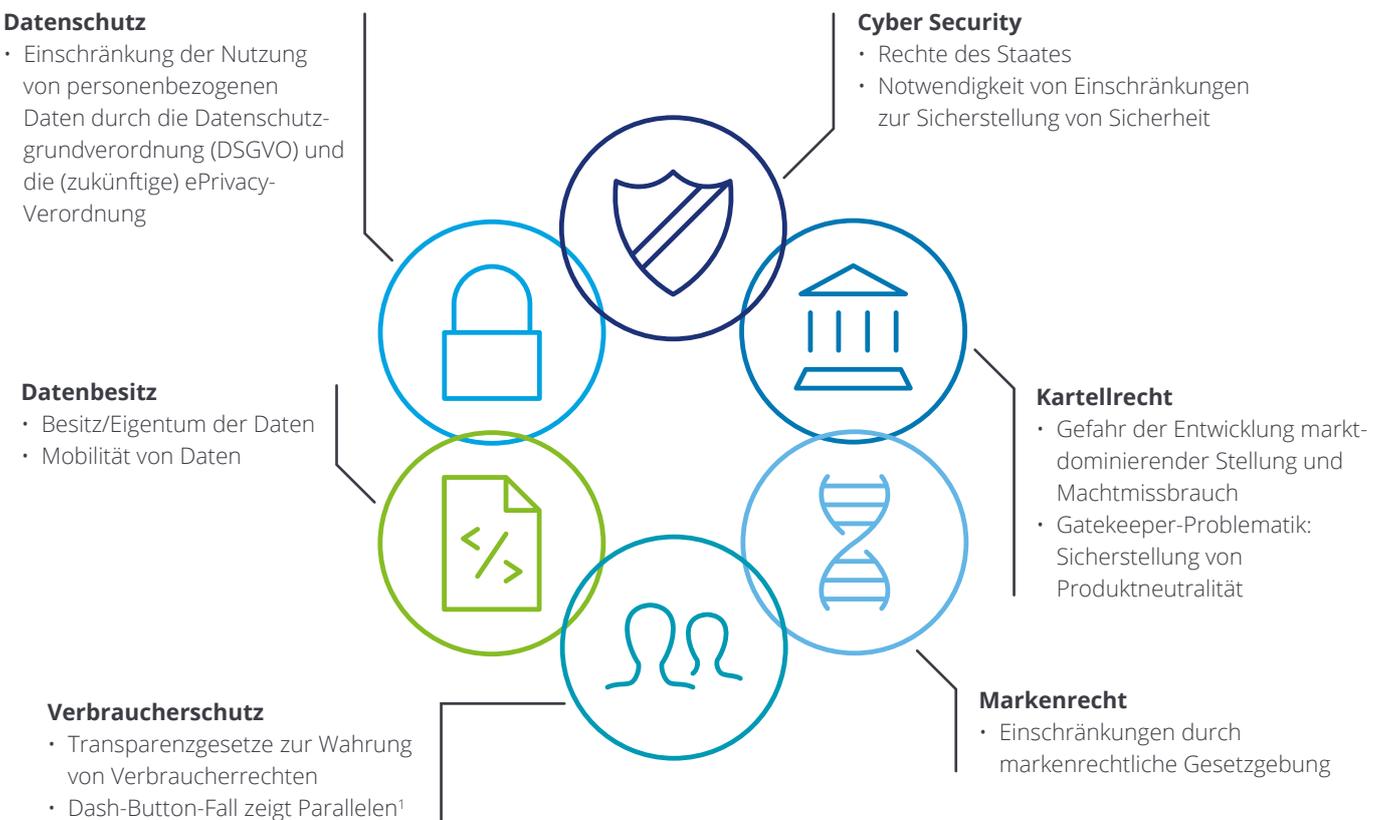
Eine Verschärfung der gesetzlichen Regelungen kann zu Risiken für den Voice Commerce führen und den Nutzen von Voice Assistants stark einschränken.

Die Veränderung gesetzlicher Rahmenbedingungen ist stets möglich und kann die weitere Entwicklung von Voice Commerce stark beeinflussen. Das betrifft zahlreiche Bereiche. Da Systeme, die mit dem Internet verbunden sind, immer anfälliger für Cyber-Kriminalität werden, könnte der Staat etwa weitere rechtliche Maßnahmen zur Stärkung der IT-Sicherheit ergreifen. Zudem steigt die Bedeutung von Datenschutz und dem Schutz der Privatsphäre stetig an. Hier

wurde durch die Datenschutzgrundverordnung (DSGVO) bereits die Nutzung von personenbezogenen Daten eingeschränkt und künftig ist die ePrivacy-Verordnung zu beachten.

Auch in Zukunft muss mit weiteren staatlichen Eingriffen gerechnet werden, etwa beim Verbraucherschutz, im Kartellrecht, Markenrecht oder Datenrecht. All diese Faktoren können teils auch negative Auswirkungen auf den Markt für Voice Assistants und den Voice Commerce haben. Deswegen sollte die Unsicherheit über den zukünftigen Stand der Regulatorik für jeden Marktteilnehmer ein wichtiger Teil der Strategieplanung sein.

Abb. 35 – Fragestellungen im Bereich Regulatorik



¹ In einem Urteil des Landgerichts München wurde der Dash-Button bemängelt, da der Verbraucher nicht darüber informiert wird, einen zahlungspflichtigen Vorgang zu tätigen, Amazon die Möglichkeit besitzt, eine Produktalternative bzw. zu einem anderen Preis zu liefern, sowie Vertragsinformationspflichten nicht erfüllt werden. (ECLI:DE:LGMUEN1:2018:0301.120730.17.0A)



Öffentliche Meinung

Die öffentliche Meinung ist für die Akzeptanz neuer Technologien von großer Bedeutung. Auch im Bereich Voice Commerce kommt es darauf an, Vertrauen zu stärken und Ängste abzubauen.

Die öffentliche Meinung zu Voice Assistants ist bisher noch maßgeblich von Zweifeln geprägt. Wie bereits angesprochen, stellen Sicherheitsbedenken die größte Hürde für Verbraucher dar, um Voice Assistants – vor allem auch beim Einkaufen – zu nutzen. Dabei sind insbesondere die Tech-Konzerne gefordert, Ängste abzubauen und das Vertrauen der Verbraucher in Voice Assistants zu stärken.

Schließlich läuft beim Voice Commerce die direkte Interaktion mit dem Nutzer über Big Player wie Amazon, Google und Apple, die als Gatekeeper fungieren. Hersteller und Händler stehen dagegen im Hintergrund und sind von den Systemanbietern abhängig.

Eine wichtige Voraussetzung für den künftigen Erfolg von Voice Assistants ist es, die öffentliche Meinung ins Positive zu drehen. Das kann nur mit einer stärkeren Aufklärung der Verbraucher über die Sicherheit von Voice Assistants gelingen. Dabei sind in erster Linie die großen Tech-Konzerne gefordert, aber auch die Händler und Hersteller müssen ihren Beitrag dazu leisten.

Abb. 36 – Änderung der öffentlichen Meinung



Quelle: Monitor Deloitte 2018.

Erfolg der großen Technologiekonzerne

Die Betrachtung der Treiber und Barrieren für den Erfolg von Voice Assistants hat gezeigt, dass die großen amerikanischen Technologiekonzerne einiges richtig machen, um ihre Marktführerschaft zu festigen und weiter auszubauen. Das liefert starke Argumente für das Eintreten des Szenarios 4: „World of Aggregators“. In dieser Welt setzen sich die großen Technologie-Konzerne durch und kontrollieren ihre technologisch integrierten Systeme als Oligopole. Dabei agieren die Big Player als Gatekeeper zwischen Käufern und Verkäufern.

Alexa als Standard etablieren und damit Anbieter und Nutzer binden. Amazon und Google integrieren gleichermaßen ihre Voice-Software über diverse Geräte und Technologien und werden so dem Treiber der technologischen Integrationsfähigkeit gerecht. Auch bezüglich des Technologievorsprungs sind Amazon, Google und Apple dem Wettbewerb weit voraus, insbesondere mit ihrer A.I.-Software. Und last but not least kooperieren Amazon, Google und Apple mit einer Vielzahl von Marktteilnehmern in diversen Industrien, vom Handel bis zum Automobilsektor.

Ein kurzer Rückblick auf den Umgang der einzelnen Tech-Konzerne mit Treibern und Barrieren macht dies deutlich. So dominiert Google den Markt mit der größten Nutzerbasis und umgeht damit das Henne-Ei-Problem. Passend zur Pfadabhängigkeit will Amazon durch eine hohe Marktpenetration

Damit spricht für die hohe Eintrittswahrscheinlichkeit der „World of Aggregators“ nicht nur das bisherige starke Marktwachstum der großen Tech-Firmen im Voice Markt. Sondern auch die Tatsache, dass die Big Player mit den Treibern und Barrieren dieser Technologie richtig umgehen können.

Abb. 37 – Erfolgsfaktoren und Best Practices



Quelle: Monitor Deloitte 2018.

Die Champions des Voice-Marktes Marktentwicklung

Wie sieht nun die konkrete Positionierung der Voice-Champions aus? Welche Strategien verfolgen Google, Amazon, Apple und Co. genau? Das klare Ziel dieser Anbieter ist es, eine überlegende Komplettlösung innerhalb der vier Dimensionen Information und digitale Assistenz, Unterhaltung, Steuerung und Einkaufen anzubieten. Dabei spielt sich der Kampf um Marktmacht hauptsächlich innerhalb des Dreiecks Information, Unterhaltung und Einkaufen ab, während das Thema Steuerung lediglich ein Hygienefaktor ist. Bei jedem Big Player sind jeweils die eigenen Kernfähigkeiten Ausgangspunkt für die Erweiterung des Angebots. So hat Google seine ursprüngliche Kernkompetenz eindeutig im Bereich Information und digitale Assistenz. Der

sprachgesteuerte Einkauf mit Google Home ist aktuell nur in den USA möglich. Das Angebot umfasst etwa eine Kooperation mit Walmart sowie die Google Express App, einen Marktplatz mit verschiedenen Händlern (Shop-in-Shop). Seit dem Kauf von YouTube ist Google auch in der Dimension Unterhaltung präsent.

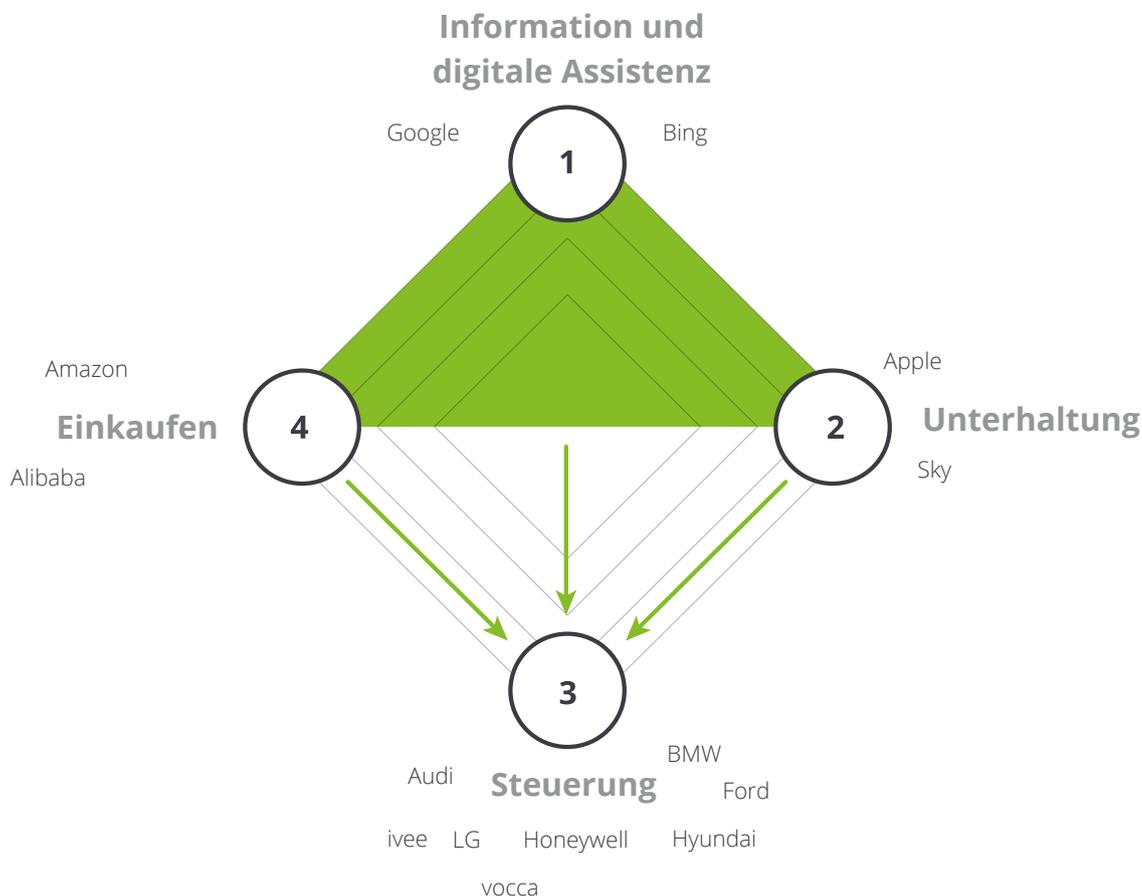
Dagegen setzt Apple vor allem auf Unterhaltung und Shopping. In den App Stores können auch Software und Apps anderer Anbieter gekauft werden. Der Sprachassistent Siri kann neben allgemeinen Infos wie dem Wetter auch Termine vereinbaren.

Amazon ist klassischerweise im E-Commerce-Sektor zu Hause. Darüber hinaus besetzt der Konzern mit Amazon Prime auch die Dimension Unterhaltung. Und der Smart Speaker Echo bietet zahlreiche Funktionalitäten.

Ganz anders sieht die Strategie der Steuerungsexperten aus. Sie streben nicht nach Höherem, sondern integrieren lediglich bestehende Voice-Lösungen der Big Player in ihre eigenen Geräte. So kooperieren zum Beispiel Honeywell beim Thermostat oder LG beim Smart TV jeweils mit Alexa und Google Assistant.

Alles in allem deutet sich eine starke Marktkonzentration auf die dominierenden Player Google, Apple und Amazon an. Gleichzeitig nimmt der Wettbewerb zwischen den großen Anbietern erheblich zu. Sie versuchen, die eigenen Lösungen in alle vier Bereiche auszuweiten. Im Folgenden werden die Entwicklung, Stärken und Positionierung der großen Technologiekonzerne im Voice-Markt beschrieben.

Abb. 38 – Voice Assistant Marktübersicht



Quelle: Monitor Deloitte 2018.

Amazon

Amazon führt als First Mover ganz klar den Voice-Markt an. Der Konzern setzt erfolgreich auf ein offenes System mit vielen Kooperationen und verfolgt eine aggressive Platzierung der eigenen Hardware.

Abb. 39 – Marktübersicht und zu beobachtende Ausbreitung – Amazon

Beispielhafte Kooperationen



Bing



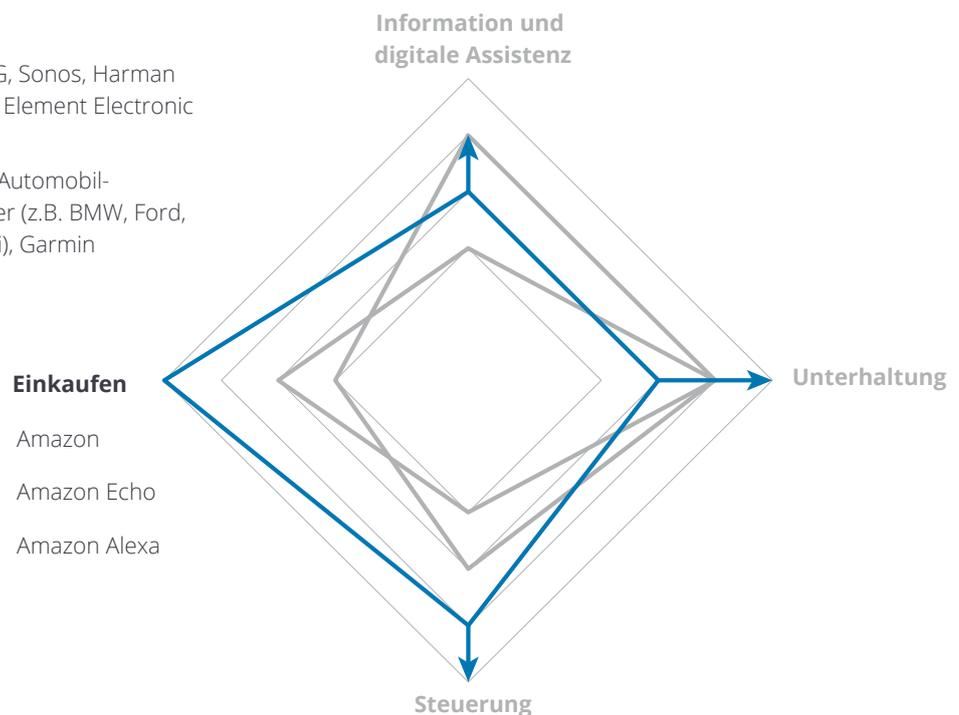
Bose, LG, Sonos, Harman Kardon, Element Electronic



REWE



GE, LG, Automobilhersteller (z.B. BMW, Ford, Hyundai), Garmin



Quelle: Monitor Deloitte 2018.

2013 hat Amazon das Start-up Ivona übernommen und die Sprachsteuerung Alexa entwickelt. Schon früh hat der Konzern aus Seattle Alexa Voice Services (AVS) auch Dritten zugänglich gemacht. Der Smart Speaker Amazon Echo kam zunächst im Juni 2015 in den USA auf den Markt. In Europa folgten 2016 United Kingdom, Deutschland und Frankreich.

Seine Kernkompetenz „Einkaufen“ stärkt Amazon besonders durch ein offenes System und unzählige Kooperationen, zum Beispiel mit Rewe. Eine wichtige Rolle spielt

auch der eigene Marketplace. Gleichzeitig dringt der Big Player auch massiv in alle anderen drei Dimensionen vor.

Zudem liegt der Fokus hier auf dem schnellen Verbreiten der eigenen Hardware aus der Amazon-Echo-Serie mithilfe von aggressiven Angeboten und Kooperationen mit anderen Hardware-Anbietern, die primär aus dem Unterhaltungs- und Steuerungsbereich kommen.

Die Kernkompetenzen von Amazon liegen natürlich im Bereich Einkaufen, aber auch

in dem breiten Unterhaltungsangebot von Amazon Prime. Zu Amazons Kooperationspartnern im Unterhaltungsbereich zählen etwa Bose, LG, Sonos, Harman Kardon und viele weitere. Für den Zugriff auf Informationen nutzt Amazon primär Wikipedia und Bing. Auch die Steuerungsmöglichkeiten hat der Konzern in den letzten Jahren stark optimiert. Kompatible Produkte gibt es zum Beispiel von GE und LG. Aber auch Auto-mobillösungen sind verfügbar und etwa bei BMW, Ford und Hyundai im Einsatz.

Google

Google ist die Nr. 2 im Voice-Markt. Der Erfolg des Konzerns aus Mountain View basiert auf einem stark Daten- und Algorithmus-getriebenen Nutzerverständnis, dem Unterhaltungswert von YouTube und ersten Retail-Lösungen.

Abb. 40 - Marktübersicht und zu beobachtende Ausbreitung - Google

Beispielhafte Kooperationen



Smartwatch-Kooperationen mit Lifestyle Brands (z.B. Hugo Boss, Guess, TAG Heuer, Montblanc, Polar)



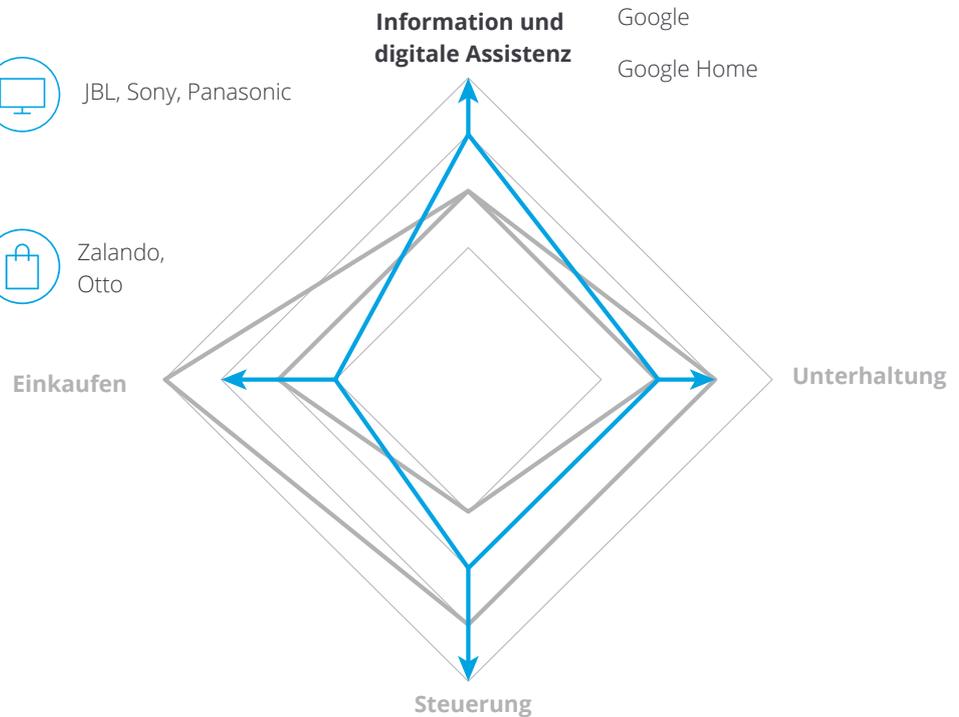
JBL, Sony, Panasonic



Philips, TP-Link, Nest, Android Auto (z.B. VW, Mercedes, Audi, Ford)



Zalando, Otto



Quelle: Monitor Deloitte 2018.

Die ursprüngliche Kernkompetenz von Google liegt eindeutig im Informationsbereich. Der entscheidende Schritt in den Voice-Markt fand im November 2016 statt. Damals wurde der Google Home Smart Speaker zusammen mit dem Google Assistant erstmals in den Vereinigten Staaten auf den Markt gebracht. Bereits einen Monat später erlaubte Google Drittanbietern, ihre Produkte oder Dienste in den Google Assistant über die Plattform Dialogflow (vorher API.ai) zu integrieren. Das zeigt, dass Google ebenfalls auf ein offenes System mit zahlreichen Kooperationen setzt.

Auch bei Android Auto kooperiert Google mit einer Vielzahl von namhaften Partnern aus der Automobilindustrie. Hier wurde eine Technologie entwickelt, die es möglich macht, verschiedene Funktionen von Android-Smartphones mit dem Infotainmentsystem in Kraftfahrzeugen zu nutzen. Auffällig sind die unzähligen Smart-Watch-Kooperationen mit Lifestyle-Marken wie Hugo Boss, Montblanc und Guess. Auch den Bereich Einkaufen baut der Konzern immer weiter aus. So gibt es Google Shopping als Shop-in-Shop-Lösung für Drittanbieter sowie das bisher nur in den USA verfügbare Angebot Google Express. Zudem bestehen Kooperationen mit Zalando und Otto.

In der Dimension Unterhaltung bietet Google Produkte wie Chromecast, YouTube und den Zugriff auf verschiedene Streamingdienste an. Kooperiert wird hier etwa mit JBL, Sony oder Panasonic. Auch im Bereich der Steuerung ist der kalifornische Tech-Gigant gut vertreten und arbeitet mit Partnern wie Philips, Nest oder TP-Link zusammen. Hier stehen sogar Lösungen mit Touchpad und Videoerweiterung zur Verfügung. Um eine Vielzahl von Kunden zu erreichen, setzt Google nicht nur auf die gezielte Vermarktung der eigenen Smart Speakers, sondern nutzt auch die Basis der bereits etablierten Android-Geräte.

Apple

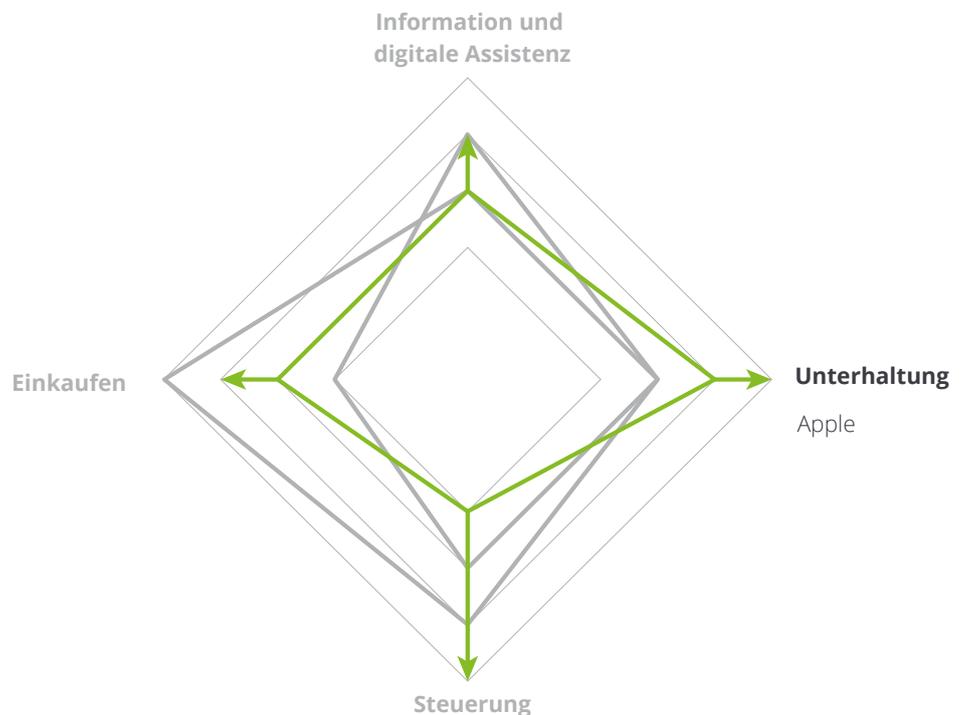
Apple ist die Nr. 3 der Big Player im Voice-Markt. Der Konzern aus Cupertino punktet mit dem Voice Assistant Siri, der zuerst auf dem iPhone zum Einsatz kam, und mit einer Vielzahl an weiteren gut vernetzten Geräten. Zudem verfügt Apple über eine starke Kompetenz in den Bereichen Unterhaltung und User Experience.

Abb. 41 – Marktübersicht und zu beobachtende Ausbreitung – Apple

Beispielhafte Kooperationen



Philips, Honeywell, IKEA, GE, CarPlay (z.B. BMW, Audi, Mercedes, Porsche)



Quelle: Monitor Deloitte 2018.

Der Smart Speaker HomePod ist seit Februar 2018 in den USA, Australien und UK erhältlich. Im Juni 2018 erfolgte der deutsche Verkaufsstart. Im Unterschied zu Amazon und Google verfolgt Apple jedoch die Strategie eines geschlossenen Systems mit einer sehr restriktiven Lizenzpolitik.

Bei der Voice-Hardware konzentriert sich Apple bisher primär auf die Nutzung der eigenen bereits etablierten Geräte wie iPhone und iPad sowie auf den HomePod. Dabei geht der Konzern nur sehr wenige Kooperationen mit Technologiepartnern

ein, was zu einer eher geringen Marktdurchdringung führt. Eine Ausnahme bildet Apple CarPlay, eine weitverbreitete Schnittstelle zur einfachen Benutzung des iPhones samt Siri in Fahrzeugen. Hier gibt es Kooperationen mit Automobilherstellern wie zum Beispiel BMW, Audi, Mercedes und Porsche.

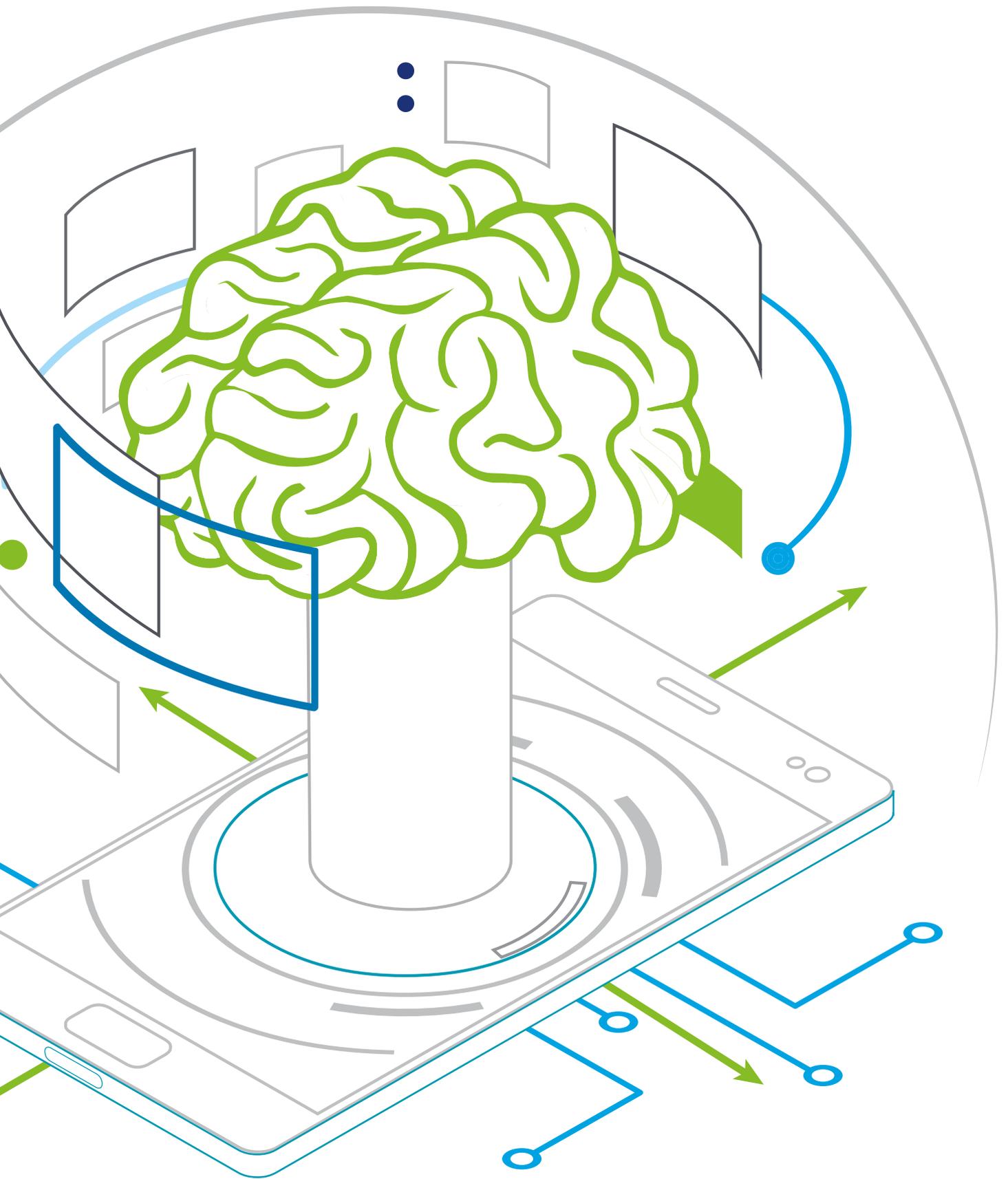
Die Kernkompetenz von Apple liegt eindeutig in der Unterhaltung. Das Hauptprodukt ist hier Apple Entertainment, das über iTunes angeboten wird. In den anderen Bereichen hat der Tech-Konzern dagegen

nur wenig Kompetenzen aufgebaut. So zählt der Informationsbereich nicht zu den Kernkompetenzen von Apple und der Voice Assistant Siri ist im Vergleich zur Konkurrenz eher schwach einzustufen. Einkaufen lassen sich Softwarekomponenten im iTunes-Sortiment. Die Steuerung von Smart-Home-Geräten und Ähnlichem ist über Siri möglich. Hier gibt es kompatible Produkte von Philips, Honeywell, IKEA oder auch GE.

Handlungsempfehlungen – Voice-assisted Interfaces als Chance nutzen

„Where to play and how to win“ mit Voice Commerce – die vorliegende Studie hat die neuesten Entwicklungen im Markt und in der Technologie von Voice Assistants untersucht. Auf dieser Basis wurden für die Zukunft des Voice Commerce die wichtigsten Treiber und Barrieren identifiziert, vier wegweisende Szenarien entwickelt und ihre die Auswirkungen auf Händler, Hersteller und Kunden beleuchtet. Um das Potenzial von Voice im eigenen Unternehmen vollumfassend zu nutzen und aus dem zusätzlichen Kundenkontaktpunkt wirklich einen funktionierenden Absatzkanal zu machen, wird nun die richtige, auf die individuelle Situation abgestimmte Strategie benötigt.





Die „Cascade of Choices“-Methodologie ist optimal geeignet, um die richtigen Schlüsse aus dieser Studie für das eigene Unternehmen zu ziehen.

Die Grundlage für eine zielführende Strategie, die auf unternehmerische Unsicherheiten vorbereitet ist, ist die Analyse der individuellen Risiken und Chancen. Dazu sollte zunächst geklärt werden, wie relevant Voice-assisted Interfaces für das Unternehmen überhaupt sind. Mittels dediziertem Impact Assessment kann der Einfluss von Voice-Technologien auf bestehende Geschäftsfelder und Kundenstrukturen für ein Unternehmen individuell analysiert werden.

Dafür müssen Unternehmen zunächst ihren eigenen Standpunkt bestimmen, um die genauen Auswirkungen der Szenarien zu identifizieren und entsprechend der eigenen Stärken und Schwächen eine passende Strategie für sich abzuleiten. Wichtig zu beachten ist, dass sich die Potenziale und Risiken der Szenarien je nach Marktteilnehmer stark unterscheiden. Dafür sind verschiedene Differenzierungskriterien ausschlaggebend.

Abb. 42 – Differenzierungskriterien

Rolle im Markt	Produkte	Marke	Sortiment	USP
Hersteller	Food	A-Marke	Monomärke	Preis
Händler	Fashion	B/C-Marke	Multimärke	Convenience
Logistik	Elektronik		Eigenmarken	Sortiment
Tech-Provider	Andere		Vollsortiment	Qualität
Politik/Verband	...		Discount	Experience
...		

Quelle: Monitor Deloitte 2018.

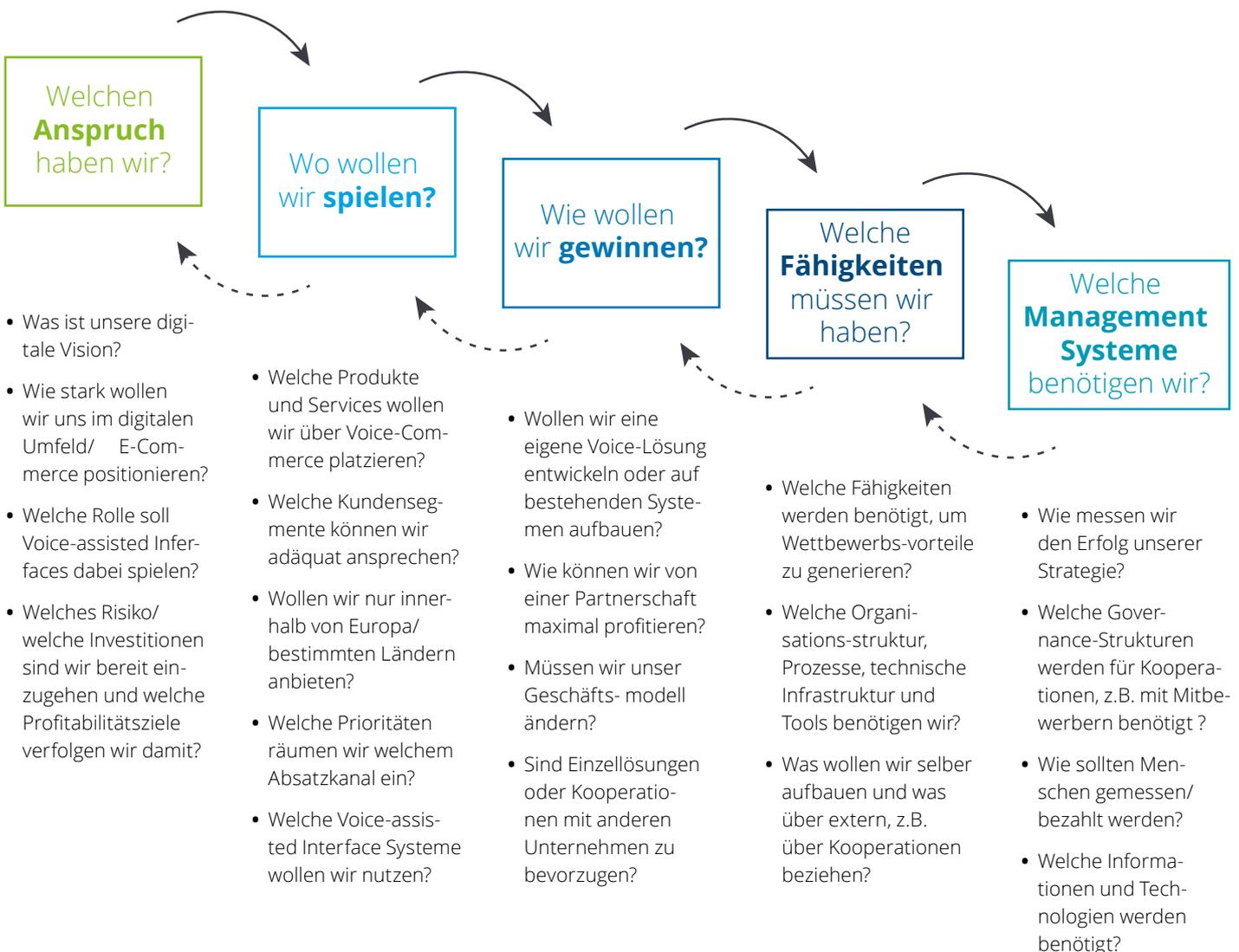
Basierend auf der Bewertung der Ausgangssituation aus dem Voice Impact Assessment können Unternehmen einen Weg entlang der Cascade of Choices entwickeln, der flexibel genug ist, um sich immer wieder ändernden Bedingungen anzupassen.

Wo und wann stellt man sich dem Wettbewerb (Where to Play) und wie tut man das erfolgreich in den Geschäftsbereichen, in denen man aktiv ist (How to Win)? Folgende nach Themen geordneten Fragen spielen dabei für jedes Unternehmen eine wichtige Rolle.

Die entwickelte Strategie muss die Basis bilden, um fundierte Entscheidungen treffen zu können, und demnach eine Vielzahl unterschiedlicher Fragen beantworten.

Untenstehend sind nach der „Cascade of Choices“-Methodologie die fünf großen Themenblöcke und eine Auswahl an zugehörigen Fragen dargestellt.

Abb. 43 – „Cascade of Choices“-Methodologie



Goals and Aspirations

Welche digitale Vision hat das Unternehmen oder der Geschäftsbereich und welche Rolle spielen dabei Voice-assisted Interfaces? Welche finanziellen Ziele sind gesteckt, welcher Wert soll geschaffen werden und wie sollen Voice-assisted Interfaces helfen, diese zu erreichen? Was ist das gewünschte Risikoprofil im direkten Zusammenhang mit den Profitabilitätszielen?

Where to Play

In welchen Geschäftsfeldern des Unternehmens sollen Voice-assisted Interfaces genutzt werden? Welche Zielkunden sollen mit welchen Voice-Produkten oder Dienstleistungen angesprochen werden? In welchen Regionen soll man sich dem Wettbewerb stellen? In welchen Teilen der Wertschöpfungskette und in welchen Prozessen kommt der Einsatz von Voice-assisted Interfaces infrage?

How to Win

Welche spezifischen Wertversprechen richten sich an die Zielkunden? Woher rührt der Wettbewerbsvorteil? Wie will das Unternehmen mit Voice-assisted Produkten oder Dienstleistungen Gewinne generieren? Welche Use Cases sind erfolgsversprechend? Sind Einzellösungen oder Kooperationen mit anderen Unternehmen zu bevorzugen?

How to Configure

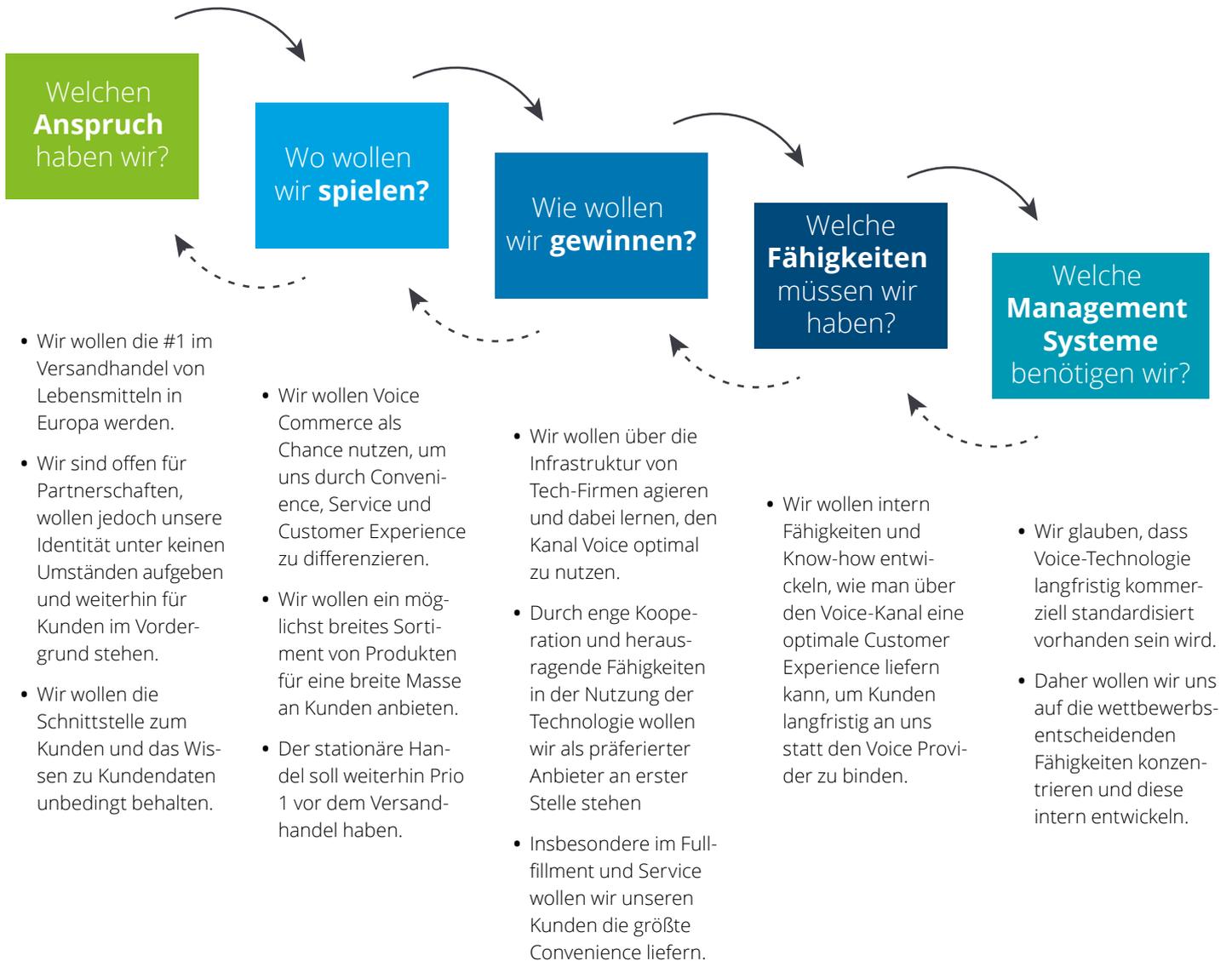
Wie sind die Aktivitäten zu strukturieren, mit denen die Strategie umgesetzt werden soll? Welche Kompetenzen im Bereich der Voice-assisted Interfaces hat das Unternehmen bereits und welche müssen aufgebaut, ausgebaut oder verinnerlicht werden? Wie müssen sich Organisation und Governance verändern, um die strategischen Ziele wirksam zu verfolgen? In welchem Umfang ist ein Wandel oder eine Transformation erforderlich und wie sind diese voranzutreiben? Welche Möglichkeiten gibt es, Partnerschaften zu gestalten, um die erforderlichen Fähigkeiten für die verschiedenen Geschäftsfelder zu gewährleisten?

Implementierung

Welche Maßnahmen, Initiativen und Investitionen sollte das Unternehmen mit hoher Priorität angehen? Welche konkreten Schritte sind zu verschiedenen Zeitpunkten notwendig, um Fortschritte in Richtung einer erfolgreichen Umsetzung der Strategie zu erzielen – etwa um die Performance zu verbessern, Geschäftsfelder neu zu ordnen oder zu erweitern, Wertversprechen oder Wettbewerbsvorteil auszubauen, erforderliche Kompetenzen zu schaffen und die Organisation weiterzuentwickeln? Wie sind Partnermanagement und Governance aufzusetzen, insbesondere, wenn einer der Kooperationspartner ein direkter Wettbewerber ist?

Als konkretes Beispiel wurden hier einige der Fragen für einen Lebensmitteleinzelhandel-Vollsortimenter im Szenario „Age of Aggregator“ beantwortet. Dieser Case ist äußerst interessant, da der Voice Commerce hier nicht nur eine Bedrohung darstellt, sondern auch ein großes Potenzial bietet, den Online-Kanal zu beleben.

Abb. 44 – Cascade of Choices am Beispiel eines Lebensmitteleinzelhandel-Vollsortimenters



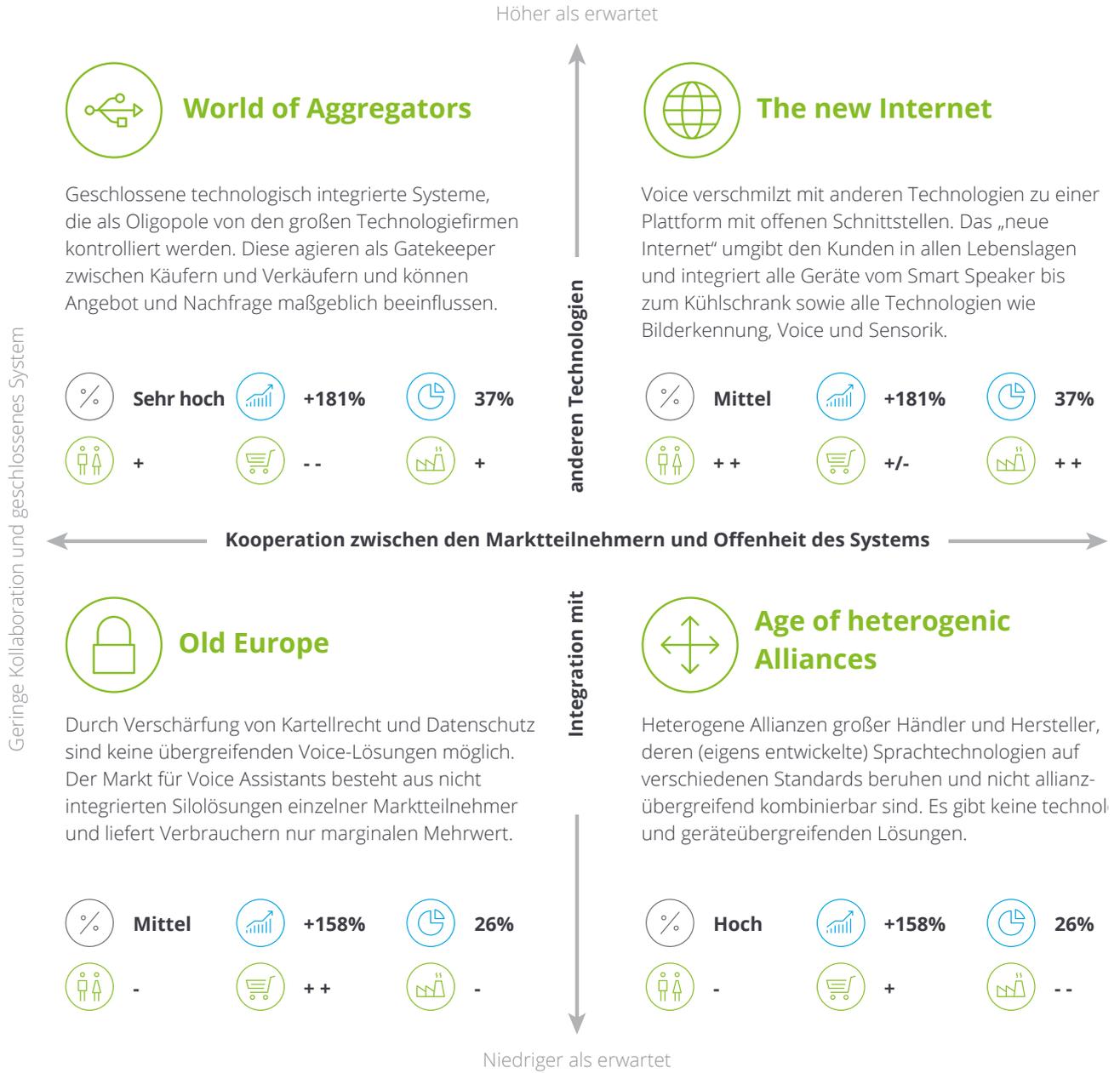
Conclusio

Die Szenarien zur Zukunft von Voice-assisted Interfaces und ihren Auswirkungen

Die „Beyond Touch“-Studie hat einen detaillierten Blick in die Zukunft des Voice Commerce geworfen: die wichtigsten Treiber und Barrieren definiert, vier extreme, aber gleichzeitig wahrscheinliche Szenarien entwickelt, die Auswirkungen auf Händler, Hersteller und Kunden analysiert sowie wichtige Handlungsempfehlungen für alle Marktteilnehmer daraus abgeleitet.

“Jedes der vier Szenarien hat individuelle Auswirkungen auf die Zukunft von Voice-Commerce in Europa und birgt unterschiedliche Vor- und Nachteile für Verbraucher, Hersteller und Händler”

Abb. 45 – Szenarien zur Zukunft von Voice-assisted Interfaces und ihren Auswirkungen



Quelle: Monitor Deloitte 2018.

Jedes Szenario birgt individuelle Vor- und Nachteile für die unterschiedlichen Akteure. Ideal für Verbraucher und Hersteller wäre das Szenario „The new Internet“, in dessen Zentrum ein komplett offenes und vernetztes System steht. Aber auch Händler könnten von dieser möglichen Realität profitieren, wenn sie den richtigen strategischen Ansatz finden und es ihnen gelingt, die Verbraucher mit Kundenbindungsprogrammen und durch Customer Experience zu überzeugen.

Aus heutiger Sicht ist am wahrscheinlichsten, dass im Jahr 2030 das Szenario „World of Aggregators“ Wirklichkeit wird. Bereits jetzt sind die beiden dominierenden Tech-Provider Amazon und Google auf bestem Wege zu einem Oligopol, in dem sie als Plattform-Aggregatoren ihre geschlossenen Systeme kontrollieren. Sie sind damit auch wichtige Gatekeeper, die den Zugang zum Kunden bestimmen.

Mit großer Wahrscheinlichkeit wird sich die „World of Aggregators“ für Hersteller als Chance und für den Handel als Bedrohung erweisen. Denn hier können die Hersteller von der Möglichkeit des Direktvertriebs profitieren. Vorausgesetzt, sie richten ihre Vertriebsaktivitäten auf die Technologie-Aggregatoren aus, um ihre Weiterempfehlungsraten durch Voice Assistants zu optimieren. Die Händler dagegen werden ihre bisherige Funktion als Bindeglied zwischen Kunde und Hersteller an die großen Tech-Provider verlieren.

Allerdings bemühen sich Händler bereits aktiv um die Entwicklung eigener Voice-Lösungen. Um sich in der Welt der Sprachassistenten einen Platz zu sichern, besteht sogar teilweise die Bereitschaft mit Mitbewerbern zu kooperieren und es sind erste Partnerschaften am Markt erkennbar. Daher ist das Szenario Age of heterogenic Alliances auch eine realistische, alternative Zukunftsvision für 2030.

Unabhängig von den einzelnen Szenarien wird es durch Voice Commerce auf jeden Fall zu einer starken Kanalverschiebung weg vom stationären Handel hin zum Onlinevertrieb kommen. Treiber dafür ist der klare Convenience-Vorteil, den sprachgesteuertes Einkaufen für den Kunden liefert. 2030 könnte Voice Commerce sogar einen Anteil von 30 % der E-Commerce-Umsätze für sich verbuchen.

Szenarienübergreifend sollten alle Marktakteure stets das Risiko staatlicher Eingriffe und Gesetzesänderungen im Blick haben. Schließlich können mögliche gesetzliche Verschärfungen des Kartellrechts, des Datenschutzes oder anderer Bereiche jederzeit große Auswirkungen auf den Voice Commerce haben.

Abb. 46 – Kernaussagen der „Beyond Touch“-Studie



Quelle: Monitor Deloitte 2018.

Nun sind die Hersteller und Händler selbst gefordert, auf Basis der Studienergebnisse eine für sie individuell passende Strategie zu entwickeln. Die Notwendigkeit zu Handeln ist heute schon gegeben. Schließlich lässt der starke Voice-Assistent-Markt in den USA auch in Europa einen großen Impact dieser Technologien erwarten. Europa steht erst am Anfang und lässt noch gewisse Freiheiten zur Entwicklung des Marktes zu.

Abb. 47 – Schritte zur dynamischen Strategie



Quelle: Monitor Deloitte 2018.

Ansprechpartner



Egbert Wege

Partner Monitor Deloitte
Strategy Lead
Tel: +49 (0)40 32080 4596
ewege@deloitte.de



Andreas Bauer

Partner Monitor Deloitte
Sector Lead Consumer Business
Tel: +49 (0)89 29036 7779
andbauer@deloitte.de



Nicolai Andersen

Partner Deloitte Garage
Head of Innovation
Tel: +49 (0)40 3208 04837
nicandersen@deloitte.de



Florian Klein

Director Monitor Deloitte
Head of Scenario Planning
Tel: +49 (0)69 97137 386
fklein@deloitte.de



Tim Hübner

Senior Consultant Monitor Deloitte
Consumer Business
Tel: +49 (0)30 25468 5564
thuebner@deloitte.de



Jerome Honerkamp

Manager Monitor Deloitte
Consumer Business
Tel: +49 (0)69 97137 572
jhonerkamp@deloitte.de



Maximilian Schulze-Frölich

Manager Monitor Deloitte
Tel: +49 (0)89 29036 4417
mschulzefroelich@deloitte.de



Jan-Niklas Keltsch

Manager Deloitte
AI-Hub Lead
Tel: +49 (0)40 32080 4240
jkeltsch@deloitte.de



Johann-Maximilian Bohle

Senior Consultant Monitor Deloitte
Tel: +49 (0)69 97137 418
jbohle@deloitte.de



Julia Weiss

Consultant Monitor Deloitte
Tel: +49 (0)89 29036 6561
jweiss@deloitte.de



Tobias Sahrer

Consultant Monitor Deloitte
Tel: +49 (0)69 97137 320
tsahrer@deloitte.de

Monitor **Deloitte.**

Diese Veröffentlichung enthält ausschließlich allgemeine Informationen, die nicht geeignet sind, den besonderen Umständen des Einzelfalls gerecht zu werden, und ist nicht dazu bestimmt, Grundlage für wirtschaftliche oder sonstige Entscheidungen zu sein. Weder die Deloitte Consulting GmbH noch Deloitte Touche Tohmatsu Limited, noch ihre Mitgliedsunternehmen oder deren verbundene Unternehmen (insgesamt das „Deloitte Netzwerk“) erbringen mittels dieser Veröffentlichung professionelle Beratungs- oder Dienstleistungen. Keines der Mitgliedsunternehmen des Deloitte Netzwerks ist verantwortlich für Verluste jedweder Art, die irgendetwas im Vertrauen auf diese Veröffentlichung erlitten hat.

Deloitte bezieht sich auf Deloitte Touche Tohmatsu Limited („DTTL“), eine „private company limited by guarantee“ (Gesellschaft mit beschränkter Haftung nach britischem Recht), ihr Netzwerk von Mitgliedsunternehmen und ihre verbundenen Unternehmen. DTTL und jedes ihrer Mitgliedsunternehmen sind rechtlich selbstständig und unabhängig. DTTL (auch „Deloitte Global“ genannt) erbringt selbst keine Leistungen gegenüber Mandanten. Eine detailliertere Beschreibung von DTTL und ihren Mitgliedsunternehmen finden Sie auf www.deloitte.com/de/UeberUns.

Deloitte erbringt Dienstleistungen in den Bereichen Wirtschaftsprüfung, Risk Advisory, Steuerberatung, Financial Advisory und Consulting für Unternehmen und Institutionen aus allen Wirtschaftszweigen; Rechtsberatung wird in Deutschland von Deloitte Legal erbracht. Mit einem weltweiten Netzwerk von Mitgliedsgesellschaften in mehr als 150 Ländern verbindet Deloitte herausragende Kompetenz mit erstklassigen Leistungen und unterstützt Kunden bei der Lösung ihrer komplexen unternehmerischen Herausforderungen. Making an impact that matters – für rund 286.000 Mitarbeiter von Deloitte ist dies gemeinsames Leitbild und individueller Anspruch zugleich.