



# **Voice Search**

Voraussetzungen und Handlungsempfehlungen  
für Unternehmen

Thomas Fröse  
Monika Engelen

*IBAL Schriftenreihe: Contributions to Business Administration and Leadership  
Jahrgang 2022, Band 2*

## Impressum

TH Köln  
Gustav-Heinemann-Ufer 54  
50968 Köln  
[www.th-koeln.de](http://www.th-koeln.de)

**Zitation:** Fröse, Thomas; Engelen, Monika (2022): Voice Search - Voraussetzungen und Handlungsempfehlungen für Unternehmen. TH Köln: Institute for Business Administration and Leadership (Schriftenreihe: Contributions to Business Administration and Leadership, 2022, 4).

Abbildung Titelblatt: Brett Sayles von Pexels

Stand: 06/2022



# Inhalt

<b>Inhalt</b>	<b>3</b>
<b>Vorwort</b>	<b>4</b>
<b>I Einleitung</b>	<b>5</b>
<b>II Voice Marketing</b>	<b>6</b>
1 Was ist Voice Marketing?	6
1.1 Position von Voice Search im Voice Marketing	6
2 Was ist Voice Search?	8
2.1 Nutzung von Voice Search	8
2.2 Speech-to-Text	8
2.3 Ergebnisse einer Voice Search	12
2.4 Einordnung von Voice Search in den Kaufprozess	14
2.5 Maßnahmen zur Optimierung der Voice Search (VSEO)	16
<b>III Handlungsempfehlungen für Unternehmen</b>	<b>20</b>
1 Vorgehen zur Ableitung von Handlungsempfehlungen	20
2 Handlungsempfehlungen	21
2.1 Ziele und Ambitionen	21
2.2 Grundlagen	22
2.3 Spezifische Ansätze für die Voice Search Optimierung	23
<b>IV Fazit und Ausblick</b>	<b>26</b>
<b>Literaturverzeichnis</b>	<b>27</b>

# Vorwort

Das vorliegende Paper beschäftigt sich mit Maßnahmen für eine erfolgreiche Auffindbarkeit von Unternehmen bei Sprachsuchen – dem sogenannten Voice Search.

Es basiert auf der Bachelorarbeit von Herrn Thomas Fröse, der von November 2021 bis Februar 2022 von Prof. Dr. Monika Engelen und Prof. Dr. Thomas Münster betreut wurde. Neben der Einordnung, Definition und Aufarbeitung des aktuellen Stands von Voice Search wurden Methoden und Techniken für eine erfolgreiche Auffindbarkeit von Unternehmen mittels Voice Search analysiert. Dafür wurde eine qualitative Inhaltsanalyse von Empfehlungen von Marketingagenturen durchgeführt.

Dieses Paper soll Erkenntnisse und Handlungsempfehlung zur erfolgreichen Auffindbarkeit von Unternehmen mittels Voice Search aufzeigen. Die Techniken, Methoden und Algorithmen von Voice Search entwickeln sich kontinuierlich weiter. Daher kann hier nur der Stand Anfang 2022 dargestellt werden.

# I Einleitung

„Hey, Google! Warum sollte ich mein Unternehmen auf Voice Search vorbereiten?“

Die Nutzung von Voice Search mit Hilfe virtueller Assistenten auf den verschiedenen Geräten steigt stetig an. 2011 launchte Apple den ersten virtuellen Assistenten Siri. Ein halbes Jahr später gab Google den Nutzern die Möglichkeit, die Sprache in Schrift umzuwandeln, um anschließend durch einen Klick die Suche starten zu können. Heutzutage werden nicht nur Informationen durch Sprachsuchen ausgetauscht, sondern auch bereits mit dem eigenen Smartphone oder Smart Speaker rein sprachgesteuerte Käufe getätigt.

Für die Sucheingabe oder Käufe benötigen bestehende und potenzielle Kunden künftig nur eins: die Sprache. Daher ist es für Unternehmen relevant, den aktuellen Stand zum Thema Voice Search zu kennen und sich darauf vorzubereiten, mittels Sprachsuche gefunden zu werden.

# II Voice Marketing

## 1 Was ist Voice Marketing?

Bei Voice Marketing handelt es sich um Marketingstrategien und -taktiken, in denen hauptsächlich die gesprochene Sprache zur Kommunikation mit den Kunden verwendet wird.

Marketing generell bedeutet, laut Homburg (2020): „[...] in unternehmensinterner Hinsicht die Schaffung der Voraussetzungen im Unternehmen für die effektive und effiziente Durchführung dieser marktbezogenen Aktivitäten. Dies schließt insbesondere die Führung des gesamten Unternehmens nach der Leitidee der Marktorientierung ein. Sowohl die externen als auch die internen Ansatzpunkte des Marketings zielen auf eine im Sinne der Unternehmensziele optimale Gestaltung von Kundenbeziehungen ab.“

Voice Marketing findet hauptsächlich im und über das Internet statt, deshalb ist das Voice Marketing ein Teilbereich des Online-Marketings, auch Internet-Marketing oder Digital-Marketing genannt.

### 1.1 Position von Voice Search im Voice Marketing

Die Einordnung von Voice Search in den Bereich Voice Marketing veranschaulicht die Abbildung 1 nach Kreuzter und Seyed Vousoghi (2020). Voice Search basiert auf der Corporate oder Brand Language und dem Sound Branding. Es kann weiter führen zu Voice Commerce und Voice Integration in Produkten. Der Fokus dieses Papers liegt allerdings auf dem Voice Search und dessen Grundlagen.



Abbildung 1: Herausforderungen des Voice Marketings; Eigene Darstellung nach Kreuzter und Seyed Vousoghi (2020), S. 14

### **Corporate oder Brand Language**

Eine gute Gestaltung von Kundenbeziehung kann beim Voice Marketing nur mit einer guten Unternehmenssprache erzielt werden. Daher ist es wichtig, erstmals die Corporate oder Brand

Language als Unternehmen festzulegen<sup>1</sup>. Um als Unternehmen festlegen zu können, wie das eigene Unternehmen „sprechen“ soll, muss man seine bisherige Kommunikation mit den Kunden analysieren. Dabei sollte sich das Unternehmen mit der Sprache, den Unterschieden von gesprochener und geschriebener Kundenkommunikation und den technischen Bedingungen von Sprachassistenten auseinandersetzen<sup>2</sup>.

Die Tonalität der Unternehmen ist für das Voice Marketing von großer Bedeutung und bildet für alle weiteren Bereiche des Voice Marketings den Grundbaustein der Pyramide. Geht ein Unternehmen seriös mit den Kunden um, so ist die Kommunikationsart eher sprachlich distanziert. Z. B. sind Banken oder Fluggesellschaften oftmals Unternehmen, die ihre Kunden förmlich ansprechen, wobei bei Einzelhändlern die Kunden häufiger geduzt werden. Hier wird versucht durch Sprache keine Distanz aufzubauen, es wird vielmehr versucht, Einigkeit zu erzeugen, um so ein „Wir-Gefühl“ zu schaffen. Nach der persönlichen Ansprache ist es für Unternehmen wichtig, die richtigen Keywords (Schlüsselwörter) zu definieren. Dabei sollte darauf geachtet werden, dass der Kunde diese Keywords möglichst eindeutig mit dem jeweiligen Unternehmen in Verbindung setzen kann. Abschließend sollte zum Corporate bzw. Brand Language die Textverständlichkeit gegeben sein. Dabei sollte die Grammatik der ausgewählten Sprache berücksichtigt werden und gleichzeitig auch der Unterschied, der zwischen gesprochener und geschriebener Sprache besteht.<sup>3</sup>

### **Sound Branding**

Doch wie klingt das jeweilige Unternehmen oder die Marke? Damit beschäftigt sich das Sound Branding. Beim Sound Branding geht es um den Klang oder um die Stimme des Unternehmens oder der Marke. Benutzt das Unternehmen eine bestimmte Melodie oder gibt es eine Stimme, die für das Unternehmen spricht?<sup>4</sup> Statistiken zu Sprachassistenten sagen aus, dass 34 % der Sprachassistentenbenutzer eher die weibliche Stimme, 21 % definitiv die weibliche Stimme und 29 % der Nutzer bevorzugen weder die weibliche noch männliche Stimme<sup>5</sup>. Auch Slogans, die mit einer bestimmten Melodie in Verbindung gesetzt werden, können als Sound-Branding verstanden werden. Sei es McDonalds: „Ich liebe es“ oder Haribos: „Haribo macht Kinder froh und Erwachsene ebenso“. Mit beiden Slogans geben sich Haribo und McDonalds eine Stimme. Der Klang und die Stimme bleibt den Kunden in Erinnerung und führt zu einem erfolgreichen Sound Branding.

### **Voice Search**

Steht fest, wie das Unternehmen spricht und klingt, muss das Unternehmen sich auffindbar und sichtbar für die Sprachsuche zu machen. Dieser Bereich wird Voice Search genannt und ist die wichtigste Stufe in der Pyramide. Denn wenn Unternehmen bei Sprachsuchen nicht gefunden werden können, finden weitere Voice Marketing Prozesse beim Kunden nicht statt. D. h., es führt nicht zum Kauf von Produkten oder Dienstleistungen mittels Voice Search oder einer Kundenbindung.

<sup>1</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 13

<sup>2</sup> Vgl. Hörner (2019), S. 66

<sup>3</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 13

<sup>4</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 13

<sup>5</sup> Vgl. Statista (2019), <https://de.statista.com/prognosen/984086/bevorzugung-maennlicher-oder-weiblicher-stimmen-bei-virtuellen-assistenten>, (10.12.21)

## 2 Was ist Voice Search?

Der Begriff Voice Search stammt aus dem Englischen und bedeutet Sprachsuche. Die Sprachsuche unterscheidet sich von der traditionellen Suche im Internet, da sie rein sprachgesteuert erfolgt und keine manuelle Eingabe im Suchfeld der Suchmaschinen (Google, Bing etc.) notwendig ist. Die sprachgesteuerte Eingabe bietet Benutzern den Nutzen, ohne das Gerät anfassen zu müssen, ihre Stimme zu verwenden, um einen Befehl oder eine Frage zu stellen und innerhalb weniger Sekunden eine Antwort zu erhalten. Bei Voice Search kommunizieren die Nutzer mit Smart Speaker, Smartphones oder anderen analogen Geräten.<sup>6</sup>

Google ist die bekannteste und meistgenutzte Suchmaschine im deutschsprachigen Raum. Deshalb ist es besonders relevant, die Auffindbarkeit des eigenen Unternehmens auf Google zu erhöhen. Der Fokus dieses Papers liegt deshalb nur auf Google als Suchmaschine. Die Empfehlungen lassen sich aber meist ähnlich auf andere Suchmaschinen übertragen.

### 2.1 Nutzung von Voice Search

Bei der Nutzung von Voice Search entsteht eine Konversation zwischen Benutzer und Voice Assistenten. Das Aktivieren der Sprachassistenten ist meist durch spezifische Voice Commands (z. B. „Hey, Google...“ oder „Alexa...“) möglich. Es entsteht ein nahezu reales Gespräch, bei dem allerdings nicht immer Standardhochdeutsch gesprochen wird. Die Nutzer von Voice Search nutzen häufig Füllwörter und sprechen mit Akzent. Generell werden Nutzer von Voice Search dazu gezwungen, Fragen zu stellen, anstatt gezielt Schlüsselwörter anzugeben wie bei der traditionellen Suche.

Dadurch ist die intensive Nutzung der W-Fragen unausweichlich. Wenn jemand nach einem Bäcker in der Nähe sucht, stellt sich diese Person grundsätzlich die Frage: „Wo befindet sich der nächste Bäcker?“. Eine umgangssprachliche Voice Search Frage könnte auch sein: „Was für Alternativen gibt es für WhatsApp?“.

Werden die ganzen Fragen in Bereiche gefiltert, so können Voice Search Suchanfragen in mehrere unterschiedliche Suchkategorien gegliedert werden.

Mögliche Beispielfragen nach Kreutzer und Seyed Vousoghi (2020)<sup>7</sup>:

- Lokales: Wo ist das nächste Restaurant?
- Shopping: Welche Kühlschränke sind am besten?
- Navigation: Wie komme ich mit dem Zug nach Köln?
- Information: Was läuft heute Mittag im Fernsehen?
- Befehl: Stelle ein Wecker für morgen um 7 Uhr.

### 2.2 Speech-to-Text

Sprache ist die natürlichste, effizienteste und die bevorzugte Art der Kommunikation unter Menschen. Daher wird davon ausgegangen, dass Menschen lieber sprechen als tippen. Studien haben sogar bewiesen, dass es viermal schneller ist, in die Suchmaschine zu sprechen, als die Tastatur zu nutzen. Um das Gesprochene anschließend fehlerfrei in einen Text umwandeln zu können, werden Spracherkennungssysteme und Sprachumwandlungssysteme benötigt. Die für

<sup>6</sup> Vgl. Kabel (2020), S. 126

<sup>7</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 8



Voice Search relevantesten Methoden sind die Automatic Speech Recognition, das Natural Language Processing, das Understanding und die Generation.

### Automatic Speech Recognition

Automatic Speech Recognition (ASR) (deutsch: automatisches Spracherkennungssystem) ermöglicht es einem Computer, die Audiodatei, welche durch ein Mikrofon aufgenommen wurde, in den entsprechenden Text umzuwandeln. Das passiert vorzugsweise in der Schrift der gesprochenen Sprache. Ein gutes ASR sollte in der Lage sein, die Eingabe zu erkennen, die gesprochenen Wörter zu unterscheiden und anschließend die identifizierten Wörter als Eingabe für eine andere Maschine zu verwenden, damit weitere Aktionen ausgeführt werden können. Es gibt mehrere Spracherkennungssysteme, jedoch ist die Grundstruktur der Spracherkennungssoftware nahezu immer gleich.<sup>8</sup>

Generell kann ASR in vier Module unterteilt werden:

- ein Klangverarbeitungsmodul,
- ein Merkmalsextraktionsmodul,
- ein Klassifizierungsmodul und
- ein Sprachmodul,

wie in Abbildung 2 gezeigt.

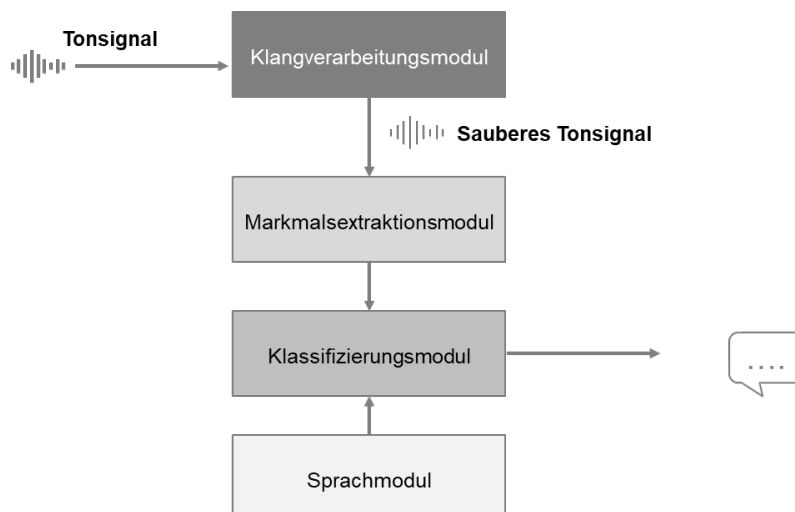


Abbildung 2: Aufbau Automatic Speech Recognition (ASR); Eigene Abbildung nach Arora, (2012), S. 5

Das einem ASR zugeführte Tonsignal wird mit einem Mikrofon erfasst. Innerhalb des Tonsignals treten i.d.R. auch Rauschen oder andere Störgeräusche auf. Das Ziel der Audiovorverarbeitung ist die Reduzierung des Signal-Rausch-Verhältnisses. Es gibt verschiedene Filter und Verfahren, die auf ein Tonsignal angewandt werden können, um das zugehörige Rauschen zu reduzieren. Framing, Normalisierung und Endpunkterkennung sind einige der häufig verwendeten Methoden, um das Rauschen in einem Signal zu reduzieren. Die Vorverarbeitungsmethoden variieren auch basierend auf dem Algorithmus, der für die Merkmalsextraktion verwendet wird. Bestimmte Merkmalsextraktionsalgorithmen erfordern die Anwendung einer bestimmten Art

<sup>8</sup> Vgl. Malik et al. (2020), S. 9412

von Vorverarbeitungsverfahren auf ihr Eingangssignal, damit ein bereinigtes Tonsignal an das Merkmalextraktionsmodul weitergeleitet wird.<sup>9</sup>

Nach der Vorverarbeitung wird das reine Sprachsignal durch das Merkmalextraktionsmodul geleitet. Die Leistung und Effizienz des Klassifizierungsmodells hängen stark von den extrahierten Merkmalen ab. Es gibt verschiedene Methoden, Merkmale aus Sprachsignalen zu extrahieren. Merkmale sind normalerweise die vordefinierte Anzahl von Koeffizienten oder Werten, die durch Anwenden verschiedener Verfahren auf das Eingangssprachsignal erhalten werden. Das Merkmalextraktionsmodul sollte gegenüber verschiedenen Faktoren, wie bspw. Rauschen und Echoeffekten, robust sein.

Das dritte Modul ist das Klassifizierungsmodul. Dieses wird verwendet, um den entsprechenden Text vorherzusagen. Klassifizierungsmodule verwenden dafür die aus der vorherigen Stufe extrahierten Merkmale. Das am häufigsten verwendete Modell basiert auf dem generativen Ansatz. Der generative Ansatz verwendet eine datensatzbasierte gebildete Wahrscheinlichkeitsverteilung um den gesuchten Text zu erkennen. In einem anderen Ansatz berechnet ein parametrisches Modell unter Verwendung eines Trainingssatzes von Eingabevektoren und ihren entsprechenden Ausgabevektoren. Dieser Ansatz wird als diskriminativer Ansatz bezeichnet.<sup>10</sup>

Das Sprachmodell ist das letzte Modul der ASR. Es besteht aus den Regeln und Semantiken einer Sprache. Sprachmodelle sind notwendig, um das vom Klassifizierer vorhergesagte Phonem zu erkennen. Anschließend werden Wörter oder Sätze, unter Verwendung aller vorhergesagten Phoneme, aus dem gesendeten Tonsignal gebildet. Bei Phonemen handelt es sich um kleinste Bedeutungsunterschiede. Z.B. kann das Wort „Bein“ vom Merkmalextraktionsmodul auch als „Pein“ verstanden werden. Wird also vom Klassifizierer ein Wort falsch erkannt, ist es die Aufgabe vom Sprachmodell, die Wörter unter Berücksichtigung der Phoneme richtig zu identifizieren.<sup>11</sup>

Die meisten modernen ASR sind so konzipiert, dass sie auch ohne Sprachmodelle funktionieren können. Diese ASR können Wörter und Sätze, die in der gegebenen Eingabe gesprochen werden, vorhersagen. Ihre Effizienz kann jedoch durch die Verwendung eines Sprachmodells deutlich erhöht werden.

### **Natural Language Processing, Understanding und Generation**

Damit die Sprachassistenten die natürliche, menschliche Sprache erkennen und verarbeiten können, sind Spracherkennungstechnologien notwendig. Eine davon ist das Natural Language Processing (NLP). Das NLP ist eine Computerlinguistik, welche dem Bereich der KI u.a. Machine Learning zuzuordnen ist. NLP beschäftigt sich mit der Verarbeitung der natürlichen menschlichen Sprache. Dabei werden zusammenhängende Informationen aus Texten extrahiert und identifiziert.<sup>12</sup>

Nachdem der extrahierte Text mit dem NLP-Verfahren analysiert wurde, geht es im nachfolgenden Schritt darum, die Absicht des Gesprochenen zu erkennen. Dafür dient das Natural Language Understanding (NLU). NLU ist zuständig für die Informationsbeschaffung und Identifizierung der Absicht (engl. Intent), nachdem die im Gesprochenen erkannte Information in die richtigen Themengebiete unterteilt wurde.<sup>13</sup>

<sup>9</sup> Vgl. Malik et al. (2020), S. 9426

<sup>10</sup> Vgl. Arora, (2012), S. 6

<sup>11</sup> Vgl. Kabel, (2020), S. 52

<sup>12</sup> Vgl. Buxmann und Schmidt, (2021), S.155

<sup>13</sup> Vgl. Kabel, (2020), S. 46

Mit Hilfe anderer, dem Themengebieten zugehörigen, Informationen kann eine Künstliche Intelligenz der Natural Language-Verfahren die Nutzerabsicht besser identifizieren und dadurch den genauen Befehl oder die Frage der Nutzer erkennen. Das Gesprochene wird im NLP, sowohl als auch im NLU, immer mit Hilfe der ASR in Absichten (Intents) umgewandelt. Daraufhin können Machine Learning-Algorithmen diese auswerten und für den weiteren Prozess verwenden.

Die NLP- und NLU-Methode sind beides Verfahren, die zur Spracherkennungstechnologie gehören, anders als beim Natural Language Generation (NLG). NLG dient dazu, eine natürliche Sprache, u.a. den vom Sprachassistenten zu sprechendem Text, zu generieren. Das ist notwendig, da der Mensch mit einem Gerät kommuniziert und die Sprachassistenten die natürliche Sprache des Menschen nicht immer nachvollziehen können. Für die Sprachassistenten ist NLG wichtig, um der Sprache des Menschen zu ähneln und dem Nutzer das Gefühl zu geben, dass sie mit einer realen Person reden.

Bei der ASR hingegen wird Sprache zum Text, wobei die Sprachgenerierungstechnologie NLG die Aufgabe hat, den Text des Tonsignals (Suchanfrage) zu generieren. Hier wird der richtige Inhalt des Textes bestimmt, um mit Hilfe von Dialog-Management-Systemen die dazugehörige natürliche Sprache zu finden<sup>14</sup>.

<sup>14</sup> Vgl. Kabel, (2020). S. 8

## 2.3 Ergebnisse einer Voice Search

Eine Search Engine Result Page (SERP), zu Deutsch Suchergebnisseite, ist die Ergebnisseite einer Suche mit der Suchmaschine Google. Abbildung 3 ist ein Beispiel für eine SERP zum Suchbegriff „Wie hoch ist die Zugspitze?“.

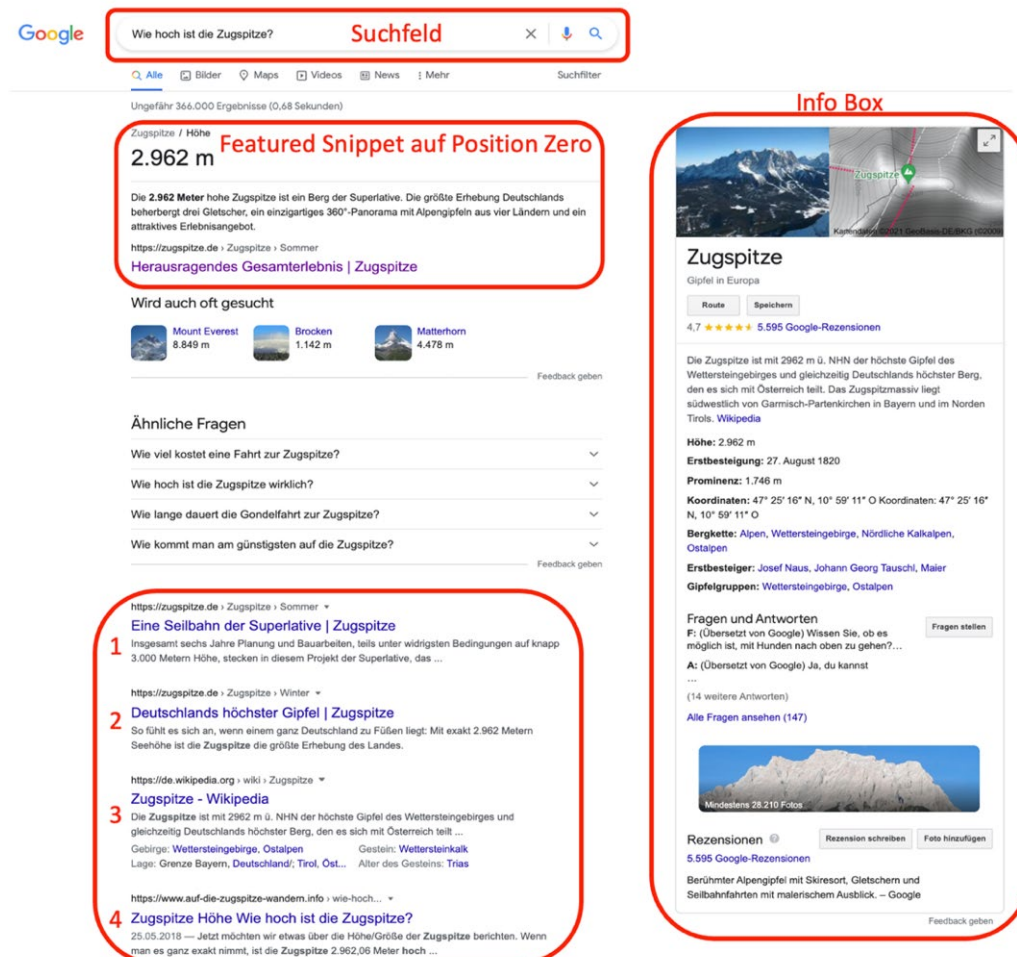


Abbildung 3: Screenshot der Google Sprachsuche mit eigenen Hervorhebungen, Vgl. Google (2021), [shorturl.at/bnJTW](https://shorturl.at/bnJTW), (29.11.2021)

Rich Snippets sind kurze Auszüge aus einer Webseite, die auf Googles SERPs dargestellt werden. Durch besonders herausgestellte und an erster Stelle genannte Snippets, die sogenannten Featured Snippets, sollen Benutzern eine Antwort auf die Suchanfrage zu erhalten, ohne die Quellwebseite besuchen zu müssen. Featured Snippets stellen präzise Antworten dar und sind meistens im Einsatz bei Verwendung informativer Suchanfragen.<sup>15</sup>

Ein Featured Snippet ist an oberster Stelle auf der Position Zero oder rechts in Form einer Info-Box platziert. Das Aussehen der Info-Box ändert sich je nach Suchanfrage. Die Info-Box soll einen Überblick über den gesuchten Inhalt geben. Bspw. werden in Abbildung 3 die wichtigsten Daten und Bilder Höhe der Zugspitze gefiltert dargestellt.

Featured Snippets bestehen nicht nur aus Textinhalten, sondern auch aus Bildern, Listen und Tabellen. Die dort dargestellten Informationen ruft Google aus verschiedenen Quellen ab. Vor

<sup>15</sup> Vgl. Kreutzer (2021), S. 370

allein die Internet-Enzyklopädien Wikipedia und Wikidata werden oft als Quellen benutzt.<sup>16</sup> Googles eigener Algorithmus Hummingbird filtert alle Suchergebnisse der Suchanfrage und vergleicht diese miteinander. Hummingbird sucht das passende Suchergebnis aus und ordnet es bei richtiger Übereinstimmung dem Intent zu, um im Nachhinein ein Featured Snippet erzeugen zu können.<sup>17</sup>

Featured Snippets spielen dabei eine wichtige Rolle für die Sprachsuche, da von den Suchenden meistens nur eine präzise Antwort erwartet wird und Voice Search die Antwort des Featured Snippet wiedergibt<sup>18</sup>.

Die Antwort der Suchanfrage wird unter dem Suchfeld angezeigt. Die oberste, meistbegehrte Stelle wird auch Position Zero, Position 0 oder auf Deutsch Nullposition genannt. Wenn die richtige Antwort daraufhin direkt erscheint (Vgl. Abb. 3), dann handelt es sich um ein Zero-Click-Search. Dieser Suchvorgang des Zero-Click-Search wird im Deutschen auch als Null-Klick-Suche bezeichnet. Das Ergebnis der Suchanfrage ist sofort bekannt und erhält die beste Platzierung, die Position Zero, vor allen anderen Suchergebnissen.<sup>19</sup>

Unter der Position Zero werden oft andere ähnliche Suchanfragen dargestellt. Unter „Wird auch oft gesucht“ werden vergleichbare Vorschläge der Suche angezeigt. Darunter befinden sich außerdem „Ähnliche Fragen“, die der vorher getätigten Suchanfrage ähneln und den suchenden Benutzer hinzukommend interessieren könnten. Dies macht Google, um Anreize zu weiteren Themen oder Fragestellungen zu setzen. Danach werden die anderen Suchergebnisse angezeigt. Die übereinstimmenden Suchergebnisse werden in der Reihenfolge von oben nach unten angezeigt. Oben befinden sich i.d.R. die zutreffendsten Suchergebnisse.

Der Google-Knowledge-Graph ermöglicht die Informationssortierung und -zusammenhänge. Ein Knowledge-Graph, auch als semantisches Netzwerk bekannt, ist ein Netzwerk von Entitäten (Personen, Unternehmen, Produkte oder Dienstleistungen) und veranschaulicht die Beziehung zwischen ihnen. Diese Informationen werden in einer Graphikdatenbank gespeichert und als Graphikstruktur visualisiert. In diesem Knowledge-Graph Googles werden diese Entitäten eindeutig identifiziert.<sup>20</sup>

Wie bei einer Wissensdatenbank werden alle Ergebnisse analysiert und mit einer Vielzahl von gewonnenen Informationen aus diversen Datenquellen verglichen. Daraufhin landen die analysierten Daten in der Info-Box, dem Featured Snippet, sowohl als auch bei „Wird auch oft gesucht“ oder bei „Ähnliche Fragen“, falls hier vom Google-Knowledge-Graph Ähnlichkeiten gefunden wurden. Google erstellt diese Elemente der SERP automatisiert, d. h. man kann es als Unternehmen nicht direkt beeinflussen, ob und wie eigene Inhalte zu Featured Snippets werden.

Voice Search und Voice Assistants bedienen sich an diesen Funktionen des Featured Snippets und der Zero-Klick-Search. Deshalb können die Sprachassistenten die richtige Antwort möglichst schnell abrufen und ermöglichen für den Benutzer damit eine bequeme Sprachsuche mit Voice Assistants, egal ob Smart Speaker oder Smartphone.

Für Unternehmen bedeutet dies, dass es wichtig ist, für die gewünschten Keywords und Fragen als Featured Snippet und möglichst weit oben auf der SERP aufzutauchen. Dazu sollten die Antworten der mit Hilfe von Voice Search gestellten Suchanfragen nicht länger als 40 Wörter

<sup>16</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 27

<sup>17</sup> Vgl. Lin et al. (2014), S. 52

<sup>18</sup> Vgl. Kreutzer (2021), S. 370

<sup>19</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 28

<sup>20</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 28

sein und dürfen eine ausgesprochene Zeit von 20 bis 30 Sek. nicht überschreiten, denn die Suchenden erwarten eine kurze und präzise Antwort. Webseitenbetreiber sollten die gewünschte Antwort direkt nach der Frage auf der Webseite positionieren, damit der Algorithmus einer Suchmaschine (z.B. bei Googles Hummingbird) die Antworten passend zuteilt. Das erhöht die Wahrscheinlichkeit, als Featured Snippet zu erscheinen.<sup>21</sup>

## 2.4 Einordnung von Voice Search in den Kaufprozess

Unternehmen sollten die Entscheidungsfindung im Kaufprozess der Kunden verstehen, um erfolgreiches Voice Marketing zu betreiben zu können. Der Kaufprozess kann als Customer Journey weiter mit Informationsquellen und -bedürfnissen der potenziellen Kunden detailliert werden. Mit Hilfe des Customer Journey können Unternehmen die „Reise“ der Kunden nachvollziehen, um effizient und effektiv Maßnahmen für eine gute Kundenbeziehung zu gestalten.<sup>22</sup>

Kotler et al. (2017) stellen im Fünf-Phasen-Modell mögliche Prozessschritte, die Konsumenten/Kunden durchlaufen, dar (Abbildung 4).

### Fünf-Phasen-Modell des Kaufprozess

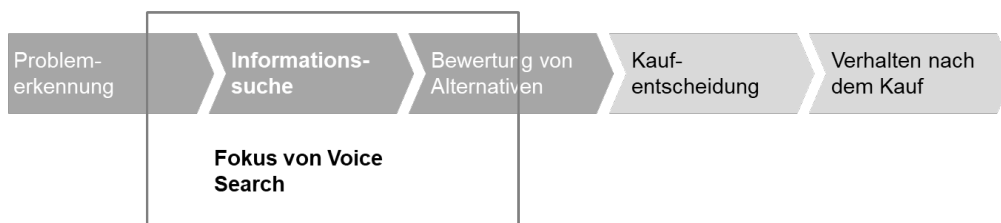


Abbildung 4: Fünf-Phasen-Modell des Kaufprozess; eigene Darstellung nach Kotler et al. (2017), S. 219

### Problemerkennung

Die erste Phase im Kaufprozess ist die Problemerkennung. In dieser Phase wird einem potenziellen Kunden bewusst, dass er oder sie ein bestimmtes Produkt oder eine bestimmte Dienstleistung zur Bewältigung seines Problems benötigt oder Bedarf an etwas Neuem besteht. Das Problembewusstsein kann entweder durch intern oder extern ausgelöste Reize entstehen. Ein intern auslösender Reiz könnte bspw. sein, dass die Person Hunger verspürt und somit die Neugier oder auch Bedarf nach Nahrung besteht oder dass sie ein bestimmtes Gerät für die Verbesserung seines Hobbys benötigt.<sup>23</sup>

Teilt man dem Google Assistant mit, dass man Durst hat, antwortet der Sprachassistent, dass man dringend etwas trinken sollte, und gibt die Möglichkeiten an, ein Café in der Nähe aufzusuchen. Wenn potenziellen Kunden die mögliche Lösung durch ein Produkt oder Unternehmen schon bekannt ist, dann wird er oder sie eher konkret nach dem Produkt- oder Unternehmensnamen suchen.

### Informationssuche

Nachdem ein Problem oder Bedürfnis erkannt wurde, begibt sich der Nutzer zur nächsten Phase, der Informationssuche. Nach Hörner (2019) können diese vier Fragearten, in der Kommunikation mit Sprachassistenten, zur Informationssuche verhelfen<sup>24</sup>:

<sup>21</sup> Vgl. Hörner (2019), S. 262

<sup>22</sup> Vgl. Homburg (2020), S. 10

<sup>23</sup> Vgl. Hörner (2019), S. 69

<sup>24</sup> Vgl. Hörner (2019), S. 72

- Anfrage direkt an Sprachassistenten z.B. „Ok Google, wo oder bei wem kann ich Fitnesskleidung kaufen?“
- Anfrage an Voice-App des eigenen Unternehmens z.B. „Alexa, frage Adidas nach Fitnesskleidung“
- Anfrage an Voice-App eines Dritten z.B. „Hey Google, rede mit Personaltrainern über geeignete Fitnesskleidung“
- Aktives Informationsangebot in einer Voice-App z. B. Während der Konversation wird passendes Produkt oder Händler platziert.

Bei der direkten Anfrage an einem Sprachassistenten ist es für Unternehmen schwieriger, auf der SERP an erste Stelle zu stehen, es sei denn der Kunde fragt spezifisch nach dem eigenen Unternehmensnamen. Daher sollte Voice Marketing nicht nur einzeln betrachtet werden, sondern auch bei traditionellen und anderen digitalen Kommunikationskampagnen beachtet werden, wie potenzielle Kunden die dort präsentierten Informationen sprachlich suchen würden.

Die Informationssuche ist die wichtigste Phase für ein erfolgreiches Voice Marketing im Bereich des Voice Search. In dieser Phase durchlaufen die Suchenden meist zwei Suchstufen. In der ersten Stufe wird generell nach einer Lösung gesucht. Dabei ist die Aufmerksamkeit des Suchenden sehr hoch und der Suchende ist neuen Produktinformationen gegenüber besonders offen. Anschließend erfolgt meist die zweite Suchstufe, in der Suchende aktiv Informationen zu einer davor identifizierten Lösung/Produkt/Unternehmen suchen.<sup>25</sup>

Während der Informationssuche sollte der Suchende eine passende Antwort bekommen haben. Das passende Unternehmen sollte direkt gefunden werden und der Sprachassistent sollte die passende Antwort geben. Hierbei ist es von hoher Bedeutung, dass der Sprachassistent die Suchintention der Benutzer versteht und die passende Antwort bietet.

### **Bewertung von Alternativen**

Hat ein Suchender seine Informationssuche abgeschlossen, sortiert und bewertet er die Informationen des Suchergebnisses. Bei Smart Speaker ist der dritte Schritt in Kotlers Kaufprozess (Bewertung von Alternativen) weniger präsent, da dem Kunden kaum Alternativen zur Verfügung gestellt werden, sondern nur eine Antwort gegeben wird.

Auf Smartphones, PCs und anderen Geräten zeigen virtuelle Assistenten bei Suchanfragen mehrere Suchergebnisse auf. Da bei der Sprachausgabe der Suchergebnisse nur eine Antwort wiedergegeben wird, können, wenn das gewünschte Unternehmen gefunden wurde, die restlichen Suchergebnisse für den Suchenden irrelevant sein. Dies bestätigt die Notwendigkeit, möglichst als erstes Ergebnis in den Suchergebnissen angezeigt bzw. vorgelesen zu werden.

### **Kaufentscheidung und das Verhalten nach dem Kauf**

Nach einer erfolgreichen Informationssuche und -bewertung kann ein Kauf oder eine Nutzung der gefundenen Lösung erfolgen. Momentan ist der Anteil der rein durch Spracheingaben getätigten Käufe eher gering, dies kann sich allerdings in Zukunft ändern. Da der Fokus dieses Praxis Papers auf dem Thema Suche liegt, wird der Bereich Kauf und Kundenbindung nicht weiter vertieft.

<sup>25</sup> Vgl. Kotler et al. (2017), S. 219

## 2.5 Maßnahmen zur Optimierung der Voice Search (VSEO)

Die Autoren Alpar et al. (2015) beschreiben: „Suchmaschinenoptimierung (zu Englisch: Search Engine Optimization - SEO) ist eine Marketing-Fachdisziplin, die alles Wissen, alle Fähigkeiten und Techniken umfasst, mit denen das zu optimierende Ziel in den organischen Ergebnissen von Suchmaschinen durch höhere Platzierungen bei unterschiedlichen Suchbegriffen besser auffindbar gemacht wird.“<sup>26</sup>

Diese Definition von SEO kann auf alle Arten von Suchen angewandt werden. Die Maßnahmen des Voice SEO setzen meist in der Phase der Informationssuche im Kaufprozess an. Im Voice SEO gibt es mehrere Ansätze, um auf Suchmaschinen erfolgreich gefunden zu werden. Nach Alpar et al. (2015) sind folgende Aspekte die wichtigsten für ein erfolgreiches Voice SEO<sup>27</sup>:

- Onpage-Optimierung
- Offpage-Optimierung
- Keyword-Strategie und -Recherche
- Content-Optimierung

Diese Aspekte beruhen alle auf generellen Ansätzen zur Optimierung der Auffindbarkeit bei Suche und werden speziell für Sprachsuchen ergänzt.

### **Keyword-Strategie und Recherche**

Bei der Keyword-Strategie geht es um die Suchbegriffe und -anfragen, die Suchenden eingeben. Hier müssen Unternehmen alle relevanten Schlüsselwörter zum eigenen Themenbereich herausfinden und analysieren. Es sollten Keywords definiert werden, die möglichst oft mit dem Unternehmen assoziiert werden. Die Herausforderung bei Voice Search liegt eher auf längeren Suchanfragen (Long-Tail-Keywords).

Bei Voice Search werden mehrere, in Verbindung stehende, Suchbegriffe als textbasierte Suchanfrage verwendet, daher ist eine Spezialisierung auf Long-Tail-Keywords notwendig<sup>28</sup>. Short-Tail-Keywords sind einzelne Schlüsselwörter, bestehend aus ein bis zwei Wörtern. Long-Tail-Keywords (deutsch: Langer-Schwanz-Schlüsselwörter) bestehen aus drei bis fünf Wörtern, gelegentlich auch mehr. Da diese Keywords spezifischer als generische Einzelbegriffe sind, erzielen Long-Tail-Keywords geringere Suchvolumen. Da aber jeder Long-Tail-Suchbegriff eine konkrete Intention des Suchenden aufweist, besteht eine höhere Handlungsbereitschaft und höhere Chance auf Konvertierung zu einer Handlung (siehe Abbildung 5).

<sup>26</sup> Alpar et al. (2015), S. 1

<sup>27</sup> Vgl. Alpar et al. (2015)

<sup>28</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 54



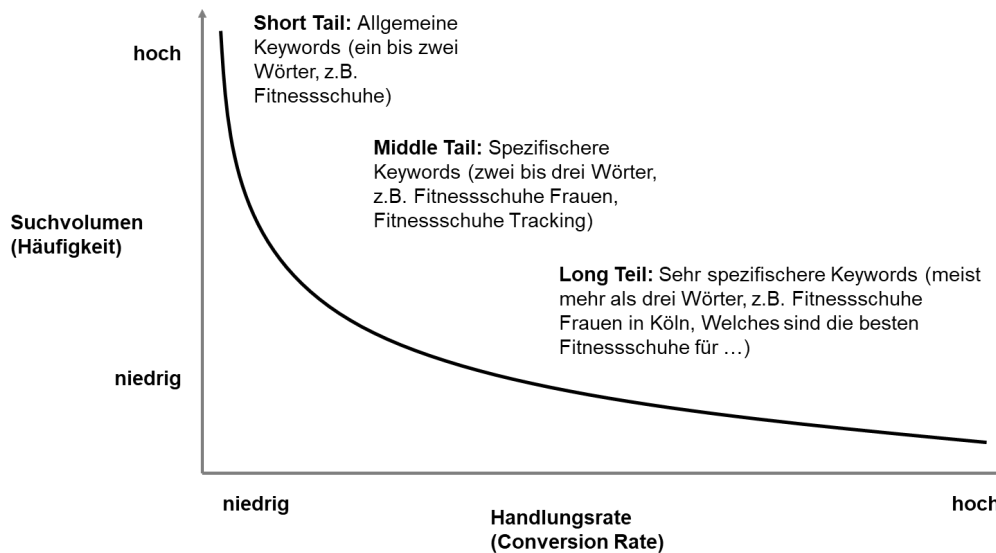


Abbildung 5: Keywords im Short, Middle und Long-Tail; Eigene Abbildung nach Kreuzter und Seyed Vousoghi (2020), S. 55

Unternehmen müssen sich darauf vorbereiten, dass die potenziellen Kunden bei Voice Search mit ihrer Sprache suchen. D. h. sie geben keine genauen Schlüsselwörter in das Suchfeld ein, wie z.B. „Fitnesskleidung“, sondern fragen ganz intuitiv „Wo oder bei wem finde ich Fitnesskleidung?“. Zu beachten ist ebenfalls, dass Keywords in Verbindung mit anderen Wörtern eine andere Bedeutung besitzen können (Intents).

Es gibt zahlreiche Tools zur Unterstützung der Keyword-Definition für das eigene Unternehmen. Beispielsweise können mit dem Google AdWords Keyword-Planner Keyword-Strategie und Recherche betrieben werden. Mit dem Tool lassen sich neue Keywords entdecken und hinzufügen, Keyword-Ergebnisse verwalten und optimieren, Suchvolumen bestimmen und Prognosen erstellen.

Unter anderem ist es wichtig, die Wettbewerber zu betrachten. Wie nutzten andere Unternehmen Keywords und welche Kombinationen mit anderen Suchbegriffen wurden bisher gewählt? Es sollte eine Datenerhebung gestartet werden, um einen Überblick der bisher gewählten Keywords zu erlangen. Nach Betrachtung aller Kriterien und gesammelten Daten soll eine Keyword Strategie entwickelt werden, welche ständig aktualisiert werden sollte.

### Onpage-Optimierung

Onpage-Optimierung bedeutet, dass Unternehmen ihre eigene Webseite für Voice SEO optimieren müssen, um ein möglichst hohes Ranking auf Suchmaschinen zu erzielen<sup>29</sup>. Man muss die eigene Informationsarchitektur an sprachbasierende Inhalte anzupassen, um von dem Google Algorithmus bei Sprachsuchen angezeigt zu werden. Webseitenbetreiber sollten eine möglichst flache Seitenstruktur erstellen<sup>30</sup>. Die Auffindbarkeit und Struktur wird klarer und den Suchmaschinen wird eine bessere Zugänglichkeit gewährt.

Beim Crawling arbeiten Crawler (softwaregesteuerte Suchroboter der Suchmaschinen<sup>31</sup>) regelmäßig das Internet durch, d. h. sie analysieren die Inhalte der Webseiten und indexieren diese

<sup>29</sup> Vgl. Onlinemarketing-Praxis (2021), <https://www.onlinemarketing-praxis.de/glossar/on-page-optimierung/>, (29.12.2021)

<sup>30</sup> Vgl. Alpar et al. (2015), S. 184

<sup>31</sup> Vgl. Gründerszene Lexikon (2019), <https://www.businessinsider.de/gruenderszene/lexikon/begriffe/crawler/>, (29.12.2021)

anschließend in der Google Knowledge Datenbank. In diesem Prozess entscheiden die Algorithmen der Suchmaschinen, welche Informationen für eine Antwort auf Suchanfragen geeignet sind.

Es können Inhalte für die Crawler hervorgehoben werden. Das geschieht auf programmierbarer Ebene mit Hilfe von robots.txt. Robots.txt ist eine Textdatei, in der Webseitenbetreiber hinterlegen können, welche Inhalte von den Suchmaschinen gelesen werden dürfen oder nicht<sup>32</sup>. Es gibt noch andere Methoden, die zur Hervorhebung oder Vernachlässigung bestimmter Inhalte führen. Mit diesen markiert man Inhalte auf der Webseite und verbessert die Chancen auf ein erfolgreiches Ranking nach dem Crawlen. Diese Markierungen werden im SEO-Fachjargon zum Gebiet der strukturierten Daten gezählt.

Die Betrachtung, Strukturierung und Markierung auf einer Website für Crawler bezeichnet man als Crawling- und Indexierungsmanagement. Nach Alper et al. (2015) ist die Berücksichtigung des Crawling- und Indexierungsmanagements aus drei Gründen für ein besseres Ranking notwendig<sup>33</sup>:

- Informationsübermittlung dank Markups schaffen eine bessere Transparenz der Webseite für Suchmaschinen
- Doppelte und unnötige Inhalte sollte man vorher ausschließen
- Bessere Steuerung des Index-Budget (Anzahl der indexierten Inhalte)

Schema.org bietet eine Möglichkeit die Texte auf den Webseiten zu markieren. Das erfolgt nach Webstandard und zeigt dem Crawler den Speakable Content (sprechbaren Inhalt) an<sup>34</sup>. Diese Texte werden mit Hilfe von Speakable Markups oder auch Semantic Markups (sprechbare oder semantische Markierungen genannt) markiert, um anschließend Structured Data (strukturierte Daten) zu erhalten.

Das empfohlene Format für Markups bei Google ist JSON-LD<sup>35</sup>. Dank Speakable Markups kann der Webseiteneigentümer (Unternehmen) seinen Content für eine Umwandlung in Sprache (Text-to-Voice) zur Verfügung stellen<sup>36</sup>. Diese Maßnahmen führen dazu, dass die Informationen der Webseiten im Google-Knowledge-Graph mit aufgenommen werden. Eine vorherige On-Page-Optimierung des Speakable Contents ist die beste Voraussetzung für Speakable Markups.

Für die Überprüfung der erstellten strukturierten Daten eignet sich bspw. Googles Markup Tool. Googles Schema Markup Validator gibt Webseiteneigentümern die Möglichkeit, alle getätigten Markierungen zu validieren und zu überprüfen. Hierfür wird zunächst die URL (Adresse der Webseite) eingegeben, um die Testung der strukturierten Daten zu starten.<sup>37</sup>

## **Offpage-Optimierung**

Bei der Offpage-Optimierung werden die Maßnahmen für eine bessere Auffindbarkeit der Webseite außerhalb der eigenen Webseite getätigt. Dieser Ansatz ist auch ein wichtiger Bestandteil

<sup>32</sup> Vgl. Lindner (2021), <https://www.seo-kueche.de/lexikon/robots-txt/>, (29.12.2021)

<sup>33</sup> Vgl. Alpar et al. (2015), S. 194

<sup>34</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 43

<sup>35</sup> Vgl. Google Search Central (2022), <https://developers.google.com/search/docs/advanced/structured-data/sd-policies>, (05.01.2022)

<sup>36</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 43

<sup>37</sup> Vgl. Schema.org (2022), <https://validator.schema.org/>, (05.01.2022)

einer generellen SEO. Eine der häufigsten Offpage-Optimierungsmethode ist der Backlinkaufbau. Backlinkaufbau funktioniert, indem ein extern eingehender Link zur Zielseite (Webseite des Unternehmens) führt.

Für Voice Search in Google ist eine Verlinkung zu Google My Business relevant. In den Unternehmensprofilen können wichtige Unternehmensinformationen bspw. Kontaktdaten, Fotos und Logos hinzugefügt werden.<sup>38</sup> Anhand der Google My Business Daten stellt Google Unternehmensprofile und -informationen bei Suchergebnissen in einer Info-Box (mit Öffnungszeiten, Telefonnummer oder sonstige Daten und Bilder zum Unternehmen) dar. Anschließend werden Benutzer zur Navigationsapplikation Google Maps weitergeleitet. Diese Verlinkungen sind besonders für lokale und transaktionale Suchen relevant.<sup>39</sup>

### **Content-Optimierung**

Der Begriff Content beschreibt die Inhalte, die man als Unternehmen schafft, veröffentlicht und nutzt. Welche Inhalte für die potenziellen Kunden relevant und interessant sind, sollte daher im ersten Schritt analysiert werden. Dabei kann eine Definition von Zielgruppen und die Erstellung von Profilen typischer Kunden (Personas) unterstützen. Die in den Nutzerprofilen aufgelisteten Informationen (demographische Merkmale, Interessen, Probleme und das Online-Kaufverhalten) verhelfen dabei, einen besseren Inhalt zu erstellen. Mit dem an die Zielgruppe angepassten Content besitzen Unternehmen wiederum die Möglichkeit, den Content in der Keyword-Strategie zu berücksichtigen.<sup>40</sup>

Content-Audits können helfen, den bisher veröffentlichten Content zu evaluieren und zu managen. Hier können Unternehmen ein individuelles Bewertungsschema formulieren, um bestenfalls die Erstellung von gleichen Contents zu vermeiden und künftig besseren Content zu erstellen, indem Erfolgsfaktoren und Risiken des bisher veröffentlichten Contents gesichtet werden.<sup>41</sup>

In der Optimierung des Contents müssen Unternehmen ausgewählte Keywords der Keyword-Recherche berücksichtigen. Diese hängen eng mit dem zu veröffentlichten Content zusammen. Nach Alpar et al. (2015) gibt es grundsätzlich drei Content-Arten<sup>42</sup>:

- Originärer Content: Content-Herstellung im eigenen Unternehmen
- User Generated Content: wurde von Nutzern/Kunden/Geschäftspartnern produziert, Erwähnung von Content-Hersteller möglich
- Curated Content: Nutzung und Zusammenfassung von Inhalten Dritter

Man muss sich als Unternehmen überlegen, zu welchen Anteilen man den Content selbst herstellen, von Kunden schaffen oder von Dritten beziehen möchte. Möchte ein Unternehmen seine Produkte oder Dienstleistungen für eine Problembewältigung präsentieren, empfiehlt sich auf bisherige Fragen und Probleme der Kunden und Nutzer in dem jeweiligen Themengebiet einzugehen. Die Probleme und Fragen sollten dann beantwortet werden.

Die Texte und Inhalte von Unternehmen auf deren Webseiten müssen für ein erfolgreiches Voice Search sprechbar und der natürlichen Sprache gerecht sein, damit nach Wiedergabe des Inhalts durch Sprachassistenten eine menschenähnliche natürliche Sprache ermöglicht wird.

<sup>38</sup> Vgl. Google My Business (2021), <https://backlinko.com/voice-search-seo-study>, (22.12.2021)

<sup>39</sup> Vgl. Kreuzer und Seyed Vousoghi (2020), S. 64

<sup>40</sup> Vgl. Alpar et al. (2015), S. 332

<sup>41</sup> Vgl. Alpar et al. (2015), S. 332

<sup>42</sup> Vgl. Alpar et al. (2015), S. 333

Dieser sprechbare Inhalt wird im Voice-SEO Speakable oder Conversational Content genannt.<sup>43</sup>

## III Handlungsempfehlungen für Unternehmen

Google ist mit ca. 83 % aller Suchanfragen die meistgenutzte Suchmaschine weltweit<sup>44</sup>. Auf Mobilgeräten wird Google sogar in 97% aller Suchanfragen genutzt<sup>45</sup>. Auch der virtuelle Assistent Googles führt die Liste aller virtuellen Assistenten an. Deshalb werden in diesem Kapitel Handlungsempfehlungen nur für die Voice SEO bei Google analysiert.

Bei Voice Search mithilfe virtueller Assistenten entsteht für Unternehmen folgende Herausforderung: Der Suchende muss mit seiner Suchanfrage direkt auf das Unternehmen stoßen. Doch was sollte ein Unternehmen tun, um möglichst oft und von den richtigen Personen in Sprachsuchen gefunden zu werden?

### 1 Vorgehen zur Ableitung von Handlungsempfehlungen

Für die Ableitung von übergreifenden Empfehlungen wurde eine qualitative Inhaltsanalyse nach Mayring (2021) erstellt. Die Experten in diesem Fachgebiet sind meist kommerzielle SEO-Agenturen. Viele davon sprechen im Internet Empfehlungen und Tipps aus, um mit der eigenen Webseite besser zu ranken und die Auffindbarkeit im Internet zu steigern. Es wurden zehn Fach- und Blogbeiträge und zwei YouTube-Videos ausgewählt. In diesen Quellen wurden alle Handlungsempfehlungen betrachtet, die zu einer erfolgreichen Auffindbarkeit von Unternehmen mittels Voice Search führen sollten. Diese Empfehlungen wurden analysiert und strukturiert zusammengefasst.

<sup>43</sup> Vgl. Kreutzer und Seyed Vousoghi (2020), S. 42

<sup>44</sup> Vgl. StatCounter (2021), <https://gs.statcounter.com/search-engine-market-share/desktop/germany#monthly-202110-202110-bar>, (06.12.2021)

<sup>45</sup> Vgl. StatCounter (2021), <https://gs.statcounter.com/search-engine-market-share/mobile/germany#monthly-202110-202110-bar> (06.12.2021)

## 2 Handlungsempfehlungen

Aus der qualitativen Inhaltsanalyse lassen sich drei Handlungsbereiche und -empfehlungen ableiten (siehe Abbildung 6).

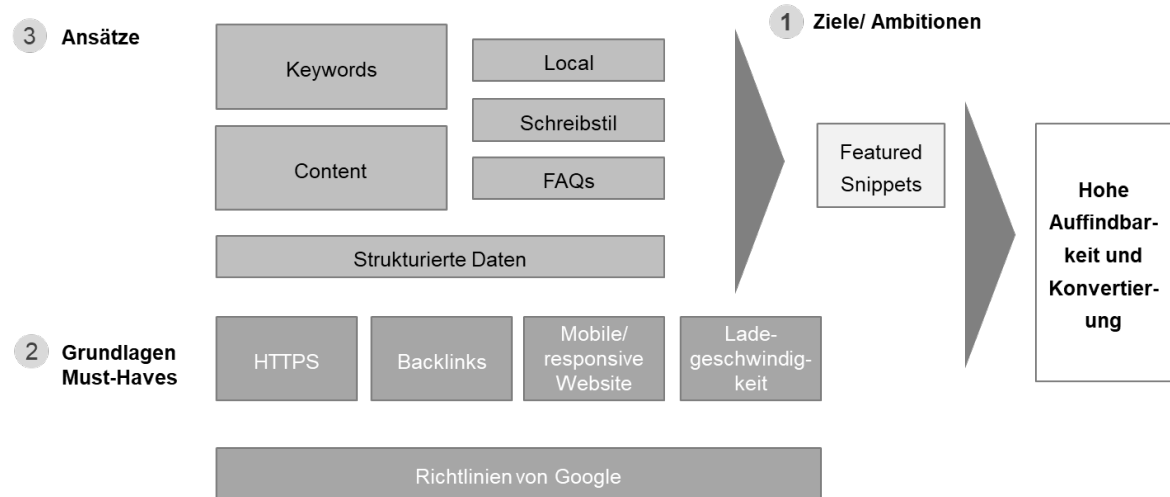


Abbildung 6: Übersicht an Handlungsempfehlungen für Voice Search; eigene Darstellung

Die Agenturempfehlung können grob in drei Themen gruppiert werden:

1. Ziel und Ambitionen, die man mit Voice Search verfolgen kann und sollte
2. Grundlagen, die erfüllt sein müssen (und so meist auch für schriftliche Suchen gelten)
3. Spezifische Ansätze für die Voice Search Optimierung

### 2.1 Ziele und Ambitionen

Grundsätzlich bezweckt man mit Search Voice Optimierung für bestimmte Suchbegriffe in Sprachsuchen gefunden zu werden, und eine Handlung, wie einen Kauf oder eine Nutzung, zu erreichen. Neben der Platzierung im oberen Bereich der Suchergebnisse ist es bei Sprachsuchen noch relevanter als Featured Snippet angezeigt zu werden.

Ein Featured Snippet ist der Auszug aus einer Website, der direkt auf der Suchergebnisseite angezeigt wird. Da Featured Snippets an allererster Stelle angezeigt werden und Sprachwiedergaben von Suchen meist auf diesen Inhalt zurückgreifen, sollte die Anzeige als Featured Snippet eines der Ziele der Optimierungsmaßnahmen sein.

Dazu kann man als Unternehmen aktuelle Featured Snippets bisheriger Suchanfragen zu sich holen, um diese anschließend zu evaluieren. Generell wird für eine Text, den man als Featured Snippet vorsieht, eine Wörteranzahl von 20 bis 30 Sek. Sprechzeit empfohlen. Um für einen Suchbegriff zu einem Featured Snippet zu werden, sollte man den Content nach klaren Fragen der Nutzer strukturieren, möglichst weit oben auf der eigenen Website unterbringen und in ca. 250 bis maximal 370 Zeichen eine klare und einfache Antwort liefern.

Das Featured Snippet ist ein relativ neues Element der SERP und die Anforderungen und Kriterien ändern sich noch. Daher empfiehlt es sich, in die jeweils neuen Erkenntnisse zu Features Snippets einzulesen.

## 2.2 Grundlagen

### **Richtlinien**

Die Richtlinien und Vorgaben der Suchmaschinen einzuhalten, ist natürlich die Grundlage für die Suchmaschinenoptimierung. Für Google sind dies die Search Speed Guidelines. Des Weiteren ist es sehr sinnvoll, sich an die Vorgaben für strukturierte Daten (siehe Kapitel 1.5.2) auch zu halten und die entsprechenden Markups zu platzieren.

### **Backlinks**

Die Erstellung von Backlinks ist eine Standardmethode und gehört zum klassischen SEO. Backlinks werden nicht zwingend benötigt, um durch Voice Search erfolgreich gefunden zu werden. Sie beeinflussen aber den generellen Suchalgorithmus und wirken sich daher auch direkt auf die Auffindbarkeit bei Sprachsuchen aus.

### **Umstellung von HTTP auf HTTPS**

Das Hypertext Transfer Protocol (HTTP) ist ein Protokollprogramm, das es ermöglicht über das Internet zu kommunizieren<sup>46</sup>. Das Kürzel „http“ steht am Anfang der Webseitenadresse und ist die Grundvoraussetzung für jede Webseite. Bei HTTPS (Hypertext Transfer Protocol Secure) wird auf Webseiten eine sichere Verschlüsselungs- und Authentifizierungsmethode genutzt, um den Datenverkehr zwischen Browser und Web-server vor unberechtigten Dritten zu schützen.

Google legt großen Wert auf die Verwendung der starken HTTPS-Verschlüsselung, um für seine Nutzer eine sichere Verbindung gewährleisten zu können. Aus diesem Grund wurde HTTPS als Ranking-Faktor für die oberen Platzierungen auf einer SERP Googles Anfang August 2014 festgesetzt.<sup>47</sup>

### **Mobile und responsive Webseite**

Eine mobile Webseite ist eine Webseite, die auf einem Smartphone oder Tablet angezeigt wird. Diese unterscheiden sich in der Größe, in der Form und dem Aufbau zur klassischen Desktop-Version. Mobile Webseiten umfassen i. d. R. kleinere Datenmengen als normale Webseiten und sind in der Form des jeweiligen Bildschirms des Smartphone-Gerätes abhängig. Mobile Webseiten sind so konzipiert, dass sie auf dem, im Vergleich zu Desktops, kleinen Bildschirmen möglichst viel Inhalt besitzen. Das wird möglich, indem Webseitenhersteller von responsive Design Gebrauch machen. Responsive Design bedeutet, dass Unternehmen anstelle von zwei Webseiten (Desktop, Smartphones oder Tablets) eine Webseite für beide Gerätetypen erstellen. Dadurch entsteht dieselbe Webseite auf beiden Gerätetypen, welche nur formabhängig vom Gerätetyp ist, jedoch der Inhalt und die Art des Aufbaus komplett identisch bleiben.<sup>48</sup>

Websites sollten heute grundsätzlich responsiv aufgebaut werden, da Endnutzer von verschiedenen Geräten auf Websites zugreifen. Sprachsuchen werden sehr oft von Tablets oder Smartphones aus getätigt, daher ist hierfür eine mobil-optimierte Ansicht notwendig.

Accelerated Mobile Pages (AMP) ist eine spezielle HTML-Variante, um mobile Webseiten nutzerfreundlich zu gestalten. Des Weiteren können mit AMP-Codier-Sprache programmierte mobile Webseiten die Ladegeschwindigkeit der Webseite beschleunigen. Die AMP-Codier-Sprache

<sup>46</sup> Vgl. Gourley und Totty (2002), S. 3

<sup>47</sup> Vgl. Bahajji und Illyes (2014), <https://developers.google.com/search/blog/2014/08/https-as-ranking-signal>, (20.12.21)

<sup>48</sup> Vgl. Kahle und Meißner (2020), S. 138

wird besonders mobilen Webseiten empfohlen, welche viele statische Inhalte wie Blogspots oder Produktzielseiten anbieten.<sup>49</sup>

### Ladegeschwindigkeit der Webseite

Mit der Ladegeschwindigkeit der Webseite ist die Ladezeit der Webseite in Sekunden gemeint. Die durchschnittliche Voice Search-Ergebnisseite lädt in 4,6 Sekunden<sup>50</sup>. Dies kann nur so schnell passieren, wenn sich möglichst wenige Dateien auf der Webseite befinden und die Abrufbarkeit der Webseite von Servern schnell erfolgt<sup>51</sup>. Bereits im Juli 2018 verkündigte Google ein Speed Update, welches festlegt, dass die Page Speed (Seitengeschwindigkeit) für die mobile Suche nun auch als Rankingfaktor gewertet wird<sup>52</sup>.

Folgende Handlungsempfehlungen wurden von verschiedenen Agenturen ausgesprochen:

- Ladegeschwindigkeit der eigenen Webseite nicht länger 1,5 Sekunden
- Ladegeschwindigkeit auf Mobilgeräten durch AMP in weniger als einer Sekunde
- Größe der Bilder und Videos auf Webseiten komprimieren

## 2.3 Spezifische Ansätze für die Voice Search Optimierung

### Keywords

Die Keyword-Optimierung ist die wichtige Voraussetzung für eine erfolgreiche Auffindbarkeit mittels Voice Search. Es gibt folgende Methoden in diesem Bereich:

- Keyword-Analyse, um das eigene Unternehmen in der eigener Nische genau positionieren zu können
- Mittels Keyword-Analyse, Keywords mit dem höchsten Ranking auswählen
- Konzentration auf Long-Tail-Keywords bestehend aus Frage-Keywords (W-Wörter), Füllwörter, Befehlswörter, lokale Suchwörter

Die Auswahl der Keywords ist der Startpunkt der SEO. Um zu den richtigen Begriffen (die zu einer hohen Konvertierung führen) gefunden zu werden, müssen Unternehmen alle für das eigene Unternehmen geeignete Keywords analysieren, bewerten und diejenigen auswählen auf die man fokussieren möchte. Die Keywordauswahl ist ein experimenteller und dynamischer Prozess, da Keywords kontinuierlich neu analysiert und selbst optimiert werden sollten.

Am Beginn der Keywordauswahl steht das Verständnis der Zielgruppen. Dafür gibt es zahlreiche Ansätze. Die folgenden Methoden können hilfreich sein:

- Durch Persona der möglichen Nutzer die Kundenpersönlichkeiten, deren Informationsbedürfnisse und -medien besser verstehen
- Kommunikationskanäle der Zielgruppe sichten (Social Media, Blogposts, FAQ und sogar Kommunikation mit Vertriebsteam des Unternehmens)
- Absichten (Intent) der Suchanfragen mit einer Nutzer-Intent-Analyse versuchen zu verstehen

<sup>49</sup> Vgl. Google Search Central (2022), <https://developers.google.com/search/docs/beginner/seo-starter-guide#mobile>, (05.01.2022)

<sup>50</sup> Vgl. Dean (2018), <https://backlinko.com/voice-search-seo-study>, (22.12.21)

<sup>51</sup> Vgl. Hörner (2019), S. 263

<sup>52</sup> Vgl. Google Search Central (2018), <https://developers.google.com/search/blog/2018/01/using-page-speed-in-mobile-search>, (4.01.2022)

Diese Methoden können dazu verhelfen, auf künftige Suchanfragen vorbereitet zu sein. Durch eine erfolgreiche Auseinandersetzung mit der Zielgruppe, können Keywords leichter analysiert werden, und durch Sichtung aller Kommunikationskanäle kann festgestellt werden, wie der Kunde spricht. Doch nicht nur einzelne Keywords können aufgegriffen werden, sondern auch ganze Fragen, welche bspw. bei den FAQ's der eigenen Webseite beantwortet können.

## **Content**

Eine erfolgreiche Auffindbarkeit setzt voraus, dass die passende Antwort auf die Suchanfrage vorhanden ist. Unternehmen müssen den Content auf die richtigen Keywords (siehe oben) im richtigen Format (siehe strukturierte Daten und MarkUps) bereitstellen. In den Handlungsempfehlungen werden folgende Methoden zur erfolgreichen Content-Generierung und Auffindbarkeit erwähnt:

- Kurzer und präziser Content, vorgelesene Antwort max. 20-30 Sek. Sprechzeit
- Konzeption von Conversational Content – also Antwort/Frage Formate
- Content in kleineren Abschnitten mit präzisen Überschriften
- An Zielgruppen angepasster, leichter und verständlicher Content
- Content anpassen auf Intent der Long-Tail-Keywords
- Content-Generierung durch Sichtung des Bereiches „Ähnliche Fragen“ auf Google
- Sichtung des Contents der SERP von traditionellen Suchanfragen

Da die Suchanfragen über Sprache ablaufen, werden viele Keywords bei der Sucheingabe benutzt. Die Herausforderung besteht darin, die richtige Intention der Suchanfragen zu verstehen. Zusätzlich kommt hinzu, dass Suchanfragen dieselbe Intention besitzen können, jedoch andere Keywords benutzt werden. Deshalb ist für die Optimierung des Contents eine gute Keyword-Recherche von sehr hoher Relevanz.

## **Local Content**

Mit Local SEO soll eine erfolgreiche Auffindbarkeit bei lokalen Suchanfragen gewährt werden. Befindet sich ein Nutzer des Voice Search in der Nähe des Unternehmens und sucht nach einem zum Unternehmen passendem Produkt oder geeignete Dienstleistung, so entsteht für das Unternehmen die Möglichkeit erfolgreich durch Voice Search gefunden zu werden

Bei den Handlungsempfehlungen des Local SEO steht der Aufbau eines Unternehmensprofil bei Google My Business im Vordergrund. Zu den bisher genannten Optimierungsansätzen des Voice Marketings gilt zusätzlich:

- Aktualität des Unternehmensprofils sicherstellen
- Ortsbezogene Informationen auf der Unternehmerwebseite einbauen
- Ortsbezogene Informationen auf der Webseite mit strukturierten Daten markieren, um bei Google „in meiner Nähe“ angezeigt zu werden

## **Schreibstil**

Der Schreibstil des Contents ist wichtig, da diese Inhalte in natürlicher Sprache wiedergegeben werden sollten. Für eine erfolgreiche Voice Search ist das Ziel, die Konversation zwischen virtuellem Assistenten und Nutzer einleiten und führen zu können, überaus relevant. Deshalb ist ein natürlich klingender Schreibstil die beste Voraussetzung, um eine möglichst natürliche Konversation mit den Nutzern aufzubauen. Einige der Handlungsempfehlungen waren:

- Natürlichen Sprachfluss, z. B. ganze Sätze und Fragen, verwenden.
- Verständlichkeitsniveau für alle Nutzer einhalten oder an Zielgruppe anpassen.



- Komplexe Sprachstrukturen (z.B. Schachtelsätze) oder komplizierte Wörter meiden.
- Leseverständlichkeit für Sprachverarbeitungssysteme (NLP und BERT) sicherstellen.
- Konversationeller Schreibstil (z.B. durch Frage/Antwort-Formate) verwenden.

Insgesamt wurde der Schreibstil von mehreren Empfehlungsgebern als wichtiger Bestandteil für ein erfolgreiches Voice Search ausgesprochen.

### **Strukturierte Daten**

Die Informationen müssen so strukturiert sein, dass diese von dem Google-Crawler sinnvoll erfasst und indexiert werden können. Hier wurden zu den im Voice Marketing genannten Methoden, folgende Techniken ausgesprochen:

- Mit Markierungen und Beschriftungen Inhalten (z.B. Bildern) Kontext verleihen
- Speakable Markup nur für Text sinnvoll, Markierung wichtiger Absätze max. 2-3 Sätze
- Speakable Markup für Unternehmensdaten in FAQ, Produktinfos und Ratings
- Speakable Markup für Speakable Content von hoher Bedeutung
- Format JSON-LD, für schnelleres Crawlen

Strukturierte Daten sind für eine erfolgreiche Auffindbarkeit im Voice Search sehr wichtig. Besonderes Augenmerkmal wird auf das Speakable Markup gesetzt. Diese Markierung ist unerlässlich, wenn sprechbare Inhalte im Featured Snippet angezeigt werden sollen.

### **FAQ**

Mit Hilfe von Frequently Asked Questions (FAQ) können Unternehmen oft gestellte Fragen beantworten. Der Ort der Zusammenstellung der FAQ ist den Webseitenbetreibern selbst überlassen. Für Voice Search bietet sich die Nutzung von FAQ besonders an, da hier Fragen formuliert sind, welche der menschlichen gesprochenen Sprache identisch sind. Die Beantwortung der Fragen sollten kurz und präzise sein. Damit eignen sich diese Inhalte auf Websites besonders für Features Snippets und die Wiedergabe in Suchanfragen.<sup>53</sup> FAQs können aus bestehenden Kundenanfragen generiert und aktualisiert werden. Die Inhalte, der Schreibstil und die Struktur der Daten sollte sich natürlich an den Empfehlungen der vorherigen Kapitel orientieren.

<sup>53</sup> Vgl. Google Search Central (2021)

## IV Fazit und Ausblick

Bereits ca. 27 % aller Smartphone-Nutzer nutzen Voice Search. Ein Großteil davon nutzt die Suchmaschine Google, wenn sie etwas suchen oder recherchieren<sup>54</sup>. Auch die Anzahl und Nutzung von sprachgesteuerten Assistenten nimmt zu. Die Sprachverarbeitungs- und Spracherkennungstechnologien werden kontinuierlich verbessert, so dass die Sprache des Menschen zukünftig noch besser von Maschinen zu verstehen ist.

Folglich werden auch die Suchen, die durch Sprache erfolgen zunehmen. Für Unternehmen ist es jetzt schon wichtig und wird es noch wichtiger werden, auch bei diesen Suchen erfolgreich gefunden zu werden.

Aufbauend auf einer allgemeinen Search Engine Optimierung für Suchen per Texteingabe können Unternehmen sich auch für Spracheingaben optimieren. Folgende grundlegende Maßnahmen und Methoden eignen sich speziell für die Voice Search Optimierung:

- Die Anzeige als Featured Snippet für relevante Suchanfragen als Ziel setzen und regelmäßig überprüfen wie Featured Snippets inhaltlich und technisch von Google erstellt werden.
- Keyword-Recherche, -Analyse und Auswahl mit Blick auf Spracheingaben der relevanten Zielgruppen (entlang des Customer Journeys und von W-Fragen) durchführen.
- Content für mögliche Sprachein- und -ausgaben optimieren durch
- Leseverständlichkeit für Kunden als auch Spracherkennungstechnologien (Voice Search-geeigneter Schreibstil) und strukturierte Daten mit Speakable Markups.
- Local SEO ausbauen (ortsbezogene Informationen auf Webseiten erstellen und Anlegung eines Unternehmensprofil auf Google My Business).

Aufgrund der sich dynamisch ändernden Richtlinien und Kriterien im Bereich Voice Search und von Google sind dies generelle Empfehlungen gültig zum Stand Anfang 2022.

<sup>54</sup> Vgl. Global Web Index (2018)

# Literaturverzeichnis

Alpar et al. (2015) SEO - Strategie, Taktik und Technik. Wiesbaden: Springer Fachmedien Wiesbaden.

Apple (2021) iOS - CarPlay. <https://www.apple.com/de/ios/carplay/> (Zugegriffen am: 7.12.2021).

Apple (2021) Siri. <https://www.apple.com/de/siri/> (Zugegriffen am: 6.12.2021).

Arora (2012) Automatic Speech Recognition: A Review. International Journal of Computer Applications.

Buxmann und Schmidt (2021) Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg. Berlin, Heidelberg: Springer Berlin Heidelberg.

Cloudflare (2022) Cloudflare CDN - das Beste seiner Art. <https://www.cloudflare.com/de-de/cdn/> (Zugegriffen am: 5.01.2022).

Dean (2018) Voice Search SEO Study: Results From 10k Voice Searches. <https://backlinko.com/voice-search-seo-study> (Zugegriffen am: 22.12.2021).

Devlin et al. (2019) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. <http://arxiv.org/abs/1810.04805> (Zugegriffen am: 5.01.2022).

Google Cloud (2022) Cloud CDN – Übersicht. <https://cloud.google.com/cdn/docs/overview?hl=de> (Zugegriffen am: 5.01.2022).

Google Developers (2021) Häufig gestellte Fragen mit strukturierten Daten versehen, Search Central. <https://developers.google.com/search/docs/advanced/structured-data/faqpage?hl=de> (Zugegriffen am: 22.12.2021).

Google Developers (2021) HTTPS als Ranking-Signal | Blog von Google Search Central. <https://developers.google.com/search/blog/2014/08/https-as-ranking-signal?hl=de> (Zugegriffen am: 20.12.2021).

Google Developers (2021) Seitengeschwindigkeit als Rankingfaktor für die mobile Suche. <https://developers.google.com/search/blog/2018/01/using-page-speed-in-mobile-search> (Zugegriffen am: 4.01.2022).

Google Developers (2021) Startleitfaden zur Suchmaschinenoptimierung (SEO): Grundlagen <https://developers.google.com/search/docs/beginner/seo-starter-guide?hl=de> (Zugegriffen am: 5.01.2022).

Google Developers (2022) Allgemeine Richtlinien für strukturierte Daten, Google Search Central. <https://developers.google.com/search/docs/advanced/structured-data/sd-policies?hl=de> (Zugegriffen am: 5.01.2022).

Google My Business (2021) Google My Business – Unternehmensprofil verwalten. [https://www.google.com/intl/de\\_de/business/](https://www.google.com/intl/de_de/business/) (Zugegriffen am: 22.12.2021).

Google Suche (2021) Wie hoch ist die Zugspitze.

Google Support (2021) Informationen zur Search Console - Search Console-Hilfe. <https://support.google.com/webmasters/answer/9128668?hl=de> (Zugegriffen am: 5.01.2022).

Gourley und Totty (2002) HTTP: the definitive guide. 1st ed. Beijing ; Sebastopol, CA: O'Reilly.

Homburg (2020) Marketingmanagement: Strategie - Instrumente - Umsetzung - Unternehmensführung. Wiesbaden: Springer Fachmedien Wiesbaden.

Hörner (2019) Marketing mit Sprachassistenten: So setzen Sie Alexa, Google Assistant & Co strategisch erfolgreich ein. Wiesbaden: Springer Fachmedien Wiesbaden.

Kabel (2020) Dialog zwischen Mensch und Maschine: Conversational User Interfaces, intelligente Assistenten und Voice-Systeme. Wiesbaden: Springer Fachmedien Wiesbaden.

Kahle und Meißner (2020) All about Voice: So gestalten und vermarkten Sie Produkte oder Dienstleistungen auf digitalen Sprachassistenten. 1. Auflage. Freiburg: Haufe Group.

Kotler et al. (2017) Marketing-Management: Konzepte - Instrumente - Unternehmensfallstudien. 15. aktualisierte Auflage. Hallbergmoos, Deutschland.

Kreutzer (2021) Kundendialog online und offline: Das große 1x1 der Kundenakquisition, Kundenbindung und Kundenrückgewinnung. Wiesbaden: Springer Fachmedien Wiesbaden.

Kreutzer und Seyed Vousoghi (2020) Voice-Marketing: Der Siegeszug der digitalen Assistenten. Wiesbaden: Springer Fachmedien Wiesbaden.

Lin et al. (2014) How Google's New Algorithm, Hummingbird, Promotes Content and Inbound Marketing. American Journal of Industrial and Business Management.

Malik et. al (2020) Automatic speech recognition: a survey. Springer Science+Business Media.

Mayring und Gläser-Zikuda (2008) Die Praxis der qualitativen Inhaltsanalyse. 2. Auflage. Weinheim Basel: Beltz Verlag (Pädagogik).

Onlinemarketing-Praxis (2021) On-Page-Optimierung - Definition. <https://www.onlinemarketing-praxis.de/glossar/on-page-optimierung> (Zugegriffen am: 29.12.2021).

Poushneh (2021) Humanizing voice assistant: The impact of voice assistant personality on consumers attitudes and behaviors. Journal of Retailing and Consumer Services.

Schema Markup Validator (2021) <https://validator.schema.org/> (Zugegriffen am: 5.01.2022).

SPLENDID Research (2019) <https://de.statista.com/statistik/daten/studie/1031358/umfrage/umfrage-zu-bekanntheit-und-nutzung-verschiedener-sprachassistenten-in-deutschland/> (Zugegriffen am: 7.12.2021).

Stäbe (2021) Webseite und Online-Shop auf HTTPS umstellen. <https://www.onlinemarketing-praxis.de/webdesign-webentwicklung/webseite-und-online-shop-auf-https-umstellen> (Zugegriffen am: 20.12.2021).

StatCounter (2021) Mobile Search Engine Market Share Germany. <https://gs.statcounter.com/search-engine-market-share/mobile/germany> (Zugegriffen am: 6.12.2021).

Statista (2019) Virtuelle Assistenten - Männliche oder weibliche Stimme in Deutschland 2019. <https://de.statista.com/prognosen/984086/bevorzugung-maennlicher-oder-weiblicher-stimmen-bei-virtuellen-assistenten> (Zugegriffen am: 10.12.2021).

Think with Google (2021) Voice search mobile use statistics. <https://www.thinkwithgoogle.com/marketing-strategies/search/voice-search-mobile-use-statistics/> (Zugegriffen am: 29.11.2021).

yext.de (2021) Wie Sprachsuche alles verändert. <https://www.yext.de/ressourcen/ver%C3%B6ffentlichungen/wie-sprachsuche-alles-verandert/> (Zugegriffen am: 21.01.2022).