



Ok Google...

The End of Search as we know it

Projektbericht zum Studienprojekt „Ok Google... The End of Search as we know it.
Sprachgesteuerte Suche im Test“

Projektdauer: 30.09.2015-03.02.2016

Projektleitung: Sebastian Sünkler (HAW Hamburg)
sebastian.suenkler@haw-hamburg.de

Friederike Kerkmann (HAW Hamburg)
friederike.kerkmann@haw-hamburg.de

Sonja Quirnbach (Deutsche Telekom AG, Darmstadt)
s.quirnbach@telekom.de

Projektteilnehmer: Alles, Ruben; Bertram, Jan; Bidar Nariman, Babak; Bilstein, Geraldine;
Camacho Galvez, Danilo Mirko; Danisch, Elena; Döring, Ingo; Hoppe,
Malte; Kraaz, Patrick; Machel, Marie; Oehlsen, Tyrone; Sakouhi, Amel;
Schultheiß, Sebastian; Tüysüz, Berna; Votteler, Lena



Inhaltsverzeichnis

Executive Summary	5
Einleitung	7
<i>Hintergrund.....</i>	<i>7</i>
<i>Problemstellung und Zielsetzung.....</i>	<i>7</i>
<i>Begriffsdefinition und Abgrenzung</i>	<i>8</i>
<i>Vorgehen</i>	<i>9</i>
<i>Aufbau der Arbeit</i>	<i>10</i>
State of the Art.....	11
<i>Übersicht: State of the Art</i>	<i>11</i>
<i>Google Now</i>	<i>14</i>
<i>Apple Siri.....</i>	<i>17</i>
<i>Microsoft Cortana.....</i>	<i>21</i>
<i>Amazon Fire OS.....</i>	<i>25</i>
Methodenbeschreibung	31
<i>Übersicht.....</i>	<i>31</i>
<i>Expertenbasierte Evaluation.....</i>	<i>33</i>
<i>Nutzerbasierte Evaluation</i>	<i>35</i>
Ergebnisse.....	37
<i>Übersicht: Auswertung Heuristische Evaluation und Szenarienbasierter Walkthrough.....</i>	<i>37</i>
<i>Google Now: Auswertung Heuristische Evaluation und Szenarienbasierter Walkthrough.....</i>	<i>44</i>
<i>Apple Siri: Auswertung Heuristische Evaluation und Szenarienbasierter Walkthrough</i>	<i>47</i>
<i>Microsoft Cortana: Auswertung Heuristische Evaluation und Szenarienbasierter Walkthrough</i>	<i>49</i>
<i>Amazon Fire OS: Auswertung Heuristische Evaluation und Szenarienbasierter Walkthrough</i>	<i>52</i>
<i>Auswertung Retrievaltest</i>	<i>54</i>

<i>Google Now: Auswertung Nutzertest</i>	<i>57</i>
<i>Apple Siri: Auswertung Nutzertest</i>	<i>59</i>
<i>Microsoft Cortana: Auswertung Nutzertest</i>	<i>62</i>
<i>Amazon Fire OS: Auswertung Nutzertest</i>	<i>66</i>
Empfehlungen für die Entwicklung der idealen Voice Web Search	69
<i>Erfolgskriterien</i>	<i>69</i>
<i>Prozessablauf und prototypische Umsetzung der idealen sprachgesteuerten Websuche</i>	<i>74</i>
Fazit	81
<i>Zusammenfassung</i>	<i>81</i>
<i>Kritische Würdigung</i>	<i>81</i>
<i>Weiterer Forschungsbedarf</i>	<i>82</i>
Quellenverzeichnis	83
Anlagen	89
<i>Anlage 1: Heuristik</i>	<i>89</i>
<i>Anlage 2: Personas</i>	<i>89</i>
<i>Anlage 3: Szenarien</i>	<i>89</i>
<i>Anlage 4: Fragebogen Vorabbefragung</i>	<i>89</i>
<i>Anlage 5: Testleitfaden</i>	<i>89</i>
<i>Anlage 6: Fragebogen Nachbefragung</i>	<i>89</i>
<i>Anlage 7: Suchprozess innerhalb einer optimalen Voice Search</i>	<i>89</i>

Executive Summary

Testgegenstand: Getestet wurde die sprachgesteuerte Websuche folgender Anbieter auf folgenden Geräten:

Voice Web Search System	Getestet auf
Google Now	LG Nexus 5
Apple Siri	iPhone 6
Microsoft Cortana	Microsoft Lumia 535
Amazon Fire OS	Amazon Fire Phone

Tabelle 1: Testsysteme und -geräte

Zielsetzung: Das Ziel der Untersuchung bestand darin, Stärken und Schwächen der betrachteten Systeme zu identifizieren und daraus Empfehlungen zur Gestaltung der optimalen Voice Web Search abzuleiten.

Methoden: Zur Umsetzung der genannten Ziele wurde ein Multi-Methoden-Ansatz gewählt. Zum Einsatz kamen eine heuristische Evaluation (mangels etablierter Heuristiken für den Bereich der mobilen sprachgesteuerten Websuche auf Basis einer eigenentwickelten Heuristik), ein szenarienbasierter Walkthrough (4 Personas, die jeweils vier Szenarien orientiert an den von Google entwickelten Micro-Moments Want-to-do, Want-to-know, Want-to-buy und Want-to-go durchliefen), ein Retrievaltest zur Relevanzbeurteilung sowie eine Nutzerstudie im freien Feld mit 25 Probanden zwischen 8 und 40 Jahren.

Drei Erfolgskriterien, die alle evaluierten Systeme umsetzten:

- Eindeutige Kennzeichnung der Anwendung als sprachgesteuerte Suche
- Spracherkennung funktioniert zuverlässig trotz lauter Umgebungsgeräusche
- Echtzeitfeedback bei der Verarbeitung von Suchanfragen

Drei Erfolgskriterien, die keines der evaluierten Systeme umsetzte:

- Interaktion der Suchergebnisse per Sprache möglich
- Ausgabe der aller Ergebnisse in akustischer Form
- Anpassung der Sprache der Ausgabe an die Sprache der Eingabe

Empfehlung: Für eine zuverlässige und anwenderorientierte sprachgesteuerte Websuche ist es notwendig, die suchspezifischen Anforderungen in die Sprachsteuerung zu integrieren. Sprachsteuerungssoftware ist technisch in der Lage, Sprache korrekt zu erkennen und zu verarbeiten, muss aber im Kontext der Websuche noch stärker auf die Nutzerbedürfnisse in Hinblick auf Sucheingabe, Verarbeitung der Suchanfragen, Ausgabe der Suchergebnisse und Interaktion mit diesen sowie auf die unterschiedlichen Nutzungskontexte bspw. definiert durch die vier Micro-Moments angepasst werden.

Einleitung

Hintergrund

Die Studierenden des Departments Information der Hochschule für Angewandte Wissenschaften (HAW) Hamburg nehmen im fünften Semester an einem interdisziplinären Studienprojekt ihrer Wahl teil. In diesen Projekten wird in der Regel eine Aufgabenstellung des Berufsalltags in enger Zusammenarbeit mit einem realen Auftraggeber aus der Praxis mit den Methoden des forschenden Lernens bearbeitet. Durch zeitkritisches und aufgabenbezogenes Arbeiten werden so die im späteren Berufsleben herrschenden Arbeitsbedingungen simuliert, kooperatives und lösungsorientierte Verhaltensweisen geschult sowie arbeitsteilige Organisationsformen und Projektmanagementmethoden erprobt.

Das Team zu dem im vorliegenden Bericht beschriebenen Studienprojekt „Ok Google... The End of Search as we know it“ im Wintersemester 2015/16 bestand aus 15 Studierenden der beiden am Department angebotenen BA-Studiengänge *Bibliotheks- und Informationsmanagement* und *Medien und Information*, zwei Projektleitern von Hochschuleseite sowie einer Projektleiterin von Seiten des Auftraggebers *Deutsche Telekom AG*.

Im Fokus des Projektes standen die Sprachsteuerungen der vier marktführenden Anbieter, deren Gebrauchstauglichkeit in Bezug auf die sprachgesteuerte Websuche im mobilen Kontext via Smartphones evaluiert wurde.

Problemstellung und Zielsetzung

„People communicate with each other by conversation, not by typing keywords“ (Singhal, 2013, o.S.), formuliert es Suchmaschinenbranchenführer Google und arbeitet seit geraumer Zeit verstärkt daran, auch seine Systeme durch Sprache steuerbar werden zu lassen. Google Now, der persönliche Assistent mit seiner Hotword-Aktivierung „Ok Google...“ ist inzwischen wortwörtlich in vieler Munde und auch in der Werbung zunehmend präsent. Er erinnert an Termine, berechnet die schnellste Route zu einem Ziel und sucht nach Informationen im Web – alles auf Zuruf. Ähnliches gilt für die Mitbewerber in Sachen Sprachsteuerungen: Cortana (Windows Phone), Siri (iOS) und die Sprachsteuerung des inzwischen eingestellten Amazon Fire Phone.

Wie aber funktioniert „googlen“, wenn der Nutzer seine Anfrage nicht wie bisher schreibt, sondern spricht? Wie zufriedenstellend, effektiv und effizient arbeiten die verschiedenen auf dem Markt befindlichen Systeme, wo liegen ihre jeweilige Stärken und Schwächen? Und nicht zuletzt – wie müsste die ideale Sprachsteuerung für Websuche gestaltet sein, um dem Nutzer

das optimale Sucherlebnis zu bieten?

Diesen Fragen ist das beschriebene Vorhaben nachgegangen. Erkenntnisse der Usability- und Suchmaschinenevaluierung flossen ebenso ein wie Aspekte der Prototypentwicklung; jeweils unter dem Einfluss mobiler Endgeräte. Damit nahm sich das Projekt einem innovativen, bislang wenig systematisch untersuchten Thema an.

Begriffsdefinition und Abgrenzung

In Wissenschaft und Praxis wird der Begriff der Sprachsteuerung allgemein als Befehlseingabe per Stimme an Geräte definiert und bildet damit eine weitere Zugangsmöglichkeit über die Benutzerschnittstellen zu Systemen und Anwendungen. Mögliche Anwendungsfelder von Sprachsteuerung sind bspw.:

- Sprachsteuerung als alternatives Eingabegerät zur Steuerung von Betriebssystemen, Fernsehern oder Spielkonsolen
- Speech-to-Text zur Umwandlung gesprochener Sprache in geschriebene Texte
- In-car-Systeme, z. B. zur Steuerung von Navigationssystemen in Pkws
- Sprachsteuerung in Operationssälen in Krankenhäusern
- Personal Assistant
- Voice Search (eine allgemeine sprachgesteuerte Suche, die sich auf Offline-Inhalte auf dem Gerät selbst sowie auf Online-Inhalte beziehen kann)
- Voice Web Search (sprachgesteuerte Suche, die sich auf Informationen im Web bezieht und in der Regel auf Websuchmaschinen basiert)

Die genannten Anwendungen lassen sich bei der praktischen Nutzung nicht immer trennscharf voneinander abgrenzen; häufig bildet die Suche den Zugang zu der eigentlichen Anwendung. Personal Assistants werden eingesetzt, um den Anwender bei Aufgabenerfüllung und Bedürfnisbefriedigung zu unterstützen, wobei ein Teil des Funktionsumfangs dafür auf direkten und indirekten Suchprozessen basiert. Somit stellt die Suche eine zentrale Anwendung für die Nutzung eines Personal Assistants dar. Ein Beispiel für die Nutzung der sprachgesteuerten Suche im Kontext des Personal Assistant ist die Aufforderung per Stimme einen Kontakt anzurufen. Dafür wird im ersten Schritte eine Suche in den gespeicherten Kontakten im Adressbuch auf dem Smartphone durchgeführt. Dieser Prozess gehört zu den Voice-Search-Anwendungen innerhalb der Sprachsteuerung.

Weiter lässt sich Voice Search von der Voice Web Search dadurch abgrenzen, da sich Voice Search allgemein sowohl auf Offline-Inhalte und Online-Inhalte bezieht, die nicht auf die

Websuche bezogen sein müssen. Voice Web Search hingegen bezieht sich ausschließlich auf die Inhalte aus Websuchmaschinen und ist damit unabhängig von den gespeicherten Inhalten auf dem Gerät. In der praktischen Anwendung bedeutet dies, dass ein Nutzer eine Suchanfrage einspricht, die durch Websuchmaschinen verarbeitet wird und die Ergebnisse dem Nutzer anschließend präsentiert werden.

Die Darstellung und Verarbeitung hängen dabei von der Integration der Websuchmaschinen in das Betriebssystem des Gerätes ab. So wird im einfachsten Fall die akustische Suchanfrage direkt in Text umgewandelt und an eine Suchmaschine weitergeleitet oder es findet eine Aufbereitung der Suchergebnisse bezogen auf das System statt.

In der Projektdurchführung wurde gezielt die Voice Web Search von vier Systemen evaluiert. Dabei wurden die Bereiche Personalisierung und Personal Assistant nicht untersucht, da das Ziel des Projekts darin bestand, festzustellen, wie gut die sprachgesteuerte Websuche auf den Systemen bisher funktioniert. Die Begriffe Voice Web Search und sprachgesteuerte Websuche werden in diesem Projektbericht synonym verwendet.

Vorgehen

Mittels Literaturrecherche wurde in einem ersten Schritt der gegenwärtige Entwicklungsstand der mobilen sprachgesteuerten Websuche aufgearbeitet und die vier zu testenden Systeme mit ihren Funktionalitäten und Besonderheiten beschrieben. Hierzu wurden vor allem populärwissenschaftliche Veröffentlichungen, vielfach Blogs, praxisorientierte Fachzeitschriften oder die Daten von Marktforschungsunternehmen herangezogen, da eine wissenschaftliche Auseinandersetzung mit dem Thema offenbar bislang kaum stattgefunden hat.

Im zweiten Schritt erfolgte die expertenbasierte Evaluation, in der die Studierenden als Experten fungierten. Hierbei kamen verschiedene Methoden (Heuristische Evaluation, Szenarienbasierter Walkthrough, Retrievaltest) zum Einsatz. Dieser Methodenblock diente zum einen der Identifikation konkreter Stärken und Schwächen, gleichzeitig aber auch dem intensiven Kennenlernen der Geräte durch die Studierenden. Kernstück der Untersuchung bildete der anschließende Nutzertest, in dem Testpersonen unter realen Bedingungen Suchaufgaben mit einem zur Verfügung gestellten Gerät und der dazugehörigen Sprachsteuerung bearbeiteten.

Die so gewonnenen Erkenntnisse wurden im Anschluss in einem gemeinsamen Workshop mit dem Auftraggeber diskutiert und im Hinblick auf die Gestaltung der idealen sprachgesteuerten Websuche konkretisiert und weiterentwickelt.

Aufbau der Arbeit

Im Anschluss an die Executive Summary für den eiligen Leser (Kapitel 0) und die Einleitung (Kapitel 1) beschreibt Kapitel 2 die zu testenden Systeme Google Now, Apple Siri, Microsoft Cortana und Amazon Fire OS mit ihrem jeweiligen Stand der Technik. Kapitel 3 widmet sich dem Studiendesign und beschreibt die angewandten Evaluierungsmethoden sowie ihr Zusammenwirken. Die Ergebnisse aus den einzelnen Untersuchungen werden in Kapitel 4 vorgestellt. Kapitel 5 konkretisiert diese mit Blick auf den Auftraggeber, skizziert den idealen Suchprozess und liefert prototypische Lösungsvorschläge für ausgewählte Probleme, die im Rahmen der Tests identifiziert wurden. Der Projektbericht endet mit dem Fazit, das neben Zusammenfassung und kritischer Würdigung der erfolgten Arbeit auch einen Ausblick mit künftigem Forschungsbedarf enthält.

State of the Art

Das folgende Kapitel stellt das Ergebnis der Literaturrecherche dar und bereitet den jeweiligen Stand der Forschung für alle vier Systeme auf. Die Quellenlage zeigt sich dabei schwierig – hauptsächlich populärwissenschaftliche Veröffentlichungen und Praktikermagazine, Blogs, Marktforschungsdaten und Selbstauskünfte der Anbieter ließen sich recherchieren. In der Wissenschaft scheint das Thema der sprachgesteuerten Websuche bislang kaum angekommen zu sein.

Übersicht: State of the Art

Systeme	Amazon Fire OS 	Apple Siri 	Google Now 	Microsoft Cortana 
Anbieter	Amazon.com, Inc.	Apple Inc.	Google Inc.	Microsoft
Verbreitung	schlechte Verkaufszahlen: Fire Phones im Wert von 83 Mio. Dollar auf Lager (2014) Veröffentlichungsdatum USA: 25.07.14 Veröffentlichungsdatum Deutschland: 30.09.14	Siri Inc. wurde von Apple Inc. aufgekauft Siri wurde 2010 für Apple-Produkte (iPhone, iPad und iPod) übernommen Verkaufszahlen steigen jedes Jahr 2015 wurden 78 Millionen Geräte verkauft	Sprachsuch-Funktion seit 2007 in den USA 2012 veröffentlichte Google den Assistenten Google Now, der eine weiterentwickelte Spracherkennungssoftware bietet - Google Voice Search	Vorstellung von Cortana im April 2014 Im Dezember 2014 kam die Alpha-Version in den europäischen Raum

Datenbasis	Spracheingaben werden in der Cloud gespeichert zur Verbesserung der Nutzererfahrung keine Informationen über Index oder andere Technik für Auswertung	Jeder Befehl bekommt eine Nummer und verbindet Sprachdatei und Nutzer die Aufbewahrung im Zusammenhang von Sprachdatei und Nutzer beträgt 6 Monate Audioschnipsel werden zwei Jahre gespeichert Sprachbefehle werden an Apple-Rechencenter geleitet Suchergebnisse kommen von Bing und Google	Google Index	Bing-Index man kann zu Google wechseln System legt Datenbank an - Personalisierung
Technik	Sprachbefehle werden aufgenommen und codiert codierte Befehle werden an die Server übermittelt Anfrage wird verarbeitet und eine Antwort zurückgesendet Algorithmus besteht aus einer Mischung aus Wahrscheinlichkeitsrechnung und Datenbankabfragen	Reine Spracherkennung Kontexterkennung Anbindung an entsprechende Systemfunktionen und Datenbanken	Spracherkennung und Verstehen von Suchanfragen Versucht eindeutige Ergebnisse zu liefern	Cortana funktioniert auf Smartphones und Tablets lernfähig

Funktionen	Starttaste wird gedrückt gehalten mit den Keywords „Suchen“ oder „im Internet suchen“ kann man Websuche starten diktieren, wonach gesucht wird Ergebnisse werden im Silk-Browser angezeigt	Zur Aktivierung Home-Button drücken und Frage einsprechen Ab IOS 8 Aktivierung ebenfalls per Sprachsteuerung „Hey Siri“ möglich lernfähig - Siri wird personalisiert	Voice Search per Sprache starten mit „Ok Google“ oder das abgebildete Mikrofon drücken Google interpretiert Suchanfrage und kann Fragen beantworten lernfähig	Symbol drücken und Frage einsprechen Ausgabe per Ergebnisliste
Dialogfähigkeit	keine Dialogfähigkeit	dialogfähig	dialogfähig	dialogfähig
Entwicklungen und Ausblick	Spekulationen über bessere Performance + Android für „Fire Phone 2“ in 2016 bisher keine offizielle Pressemeldung von Amazon	Einzigste Info: Siri soll künftig als Anrufbeantworter fungieren können	Standort- und wissensbezogene Fragen sollen verbunden werden Einkäufe sollen über Sprachsteuerung getätigt werden können	Cortana möchte auf die Windschutzscheibe von Autos Neben Android-Nutzern sollen künftig auch iOS-Nutzer Cortana verwenden können Cortana auf Xbox Soll in Browser „Spartan“ integriert werden

Tabelle 2: Übersicht State of the Art

Google Now



Anbieter

Google Inc. ist der Anbieter der Google Voice Search und ein US-amerikanisches Tochterunternehmen der Alphabet Inc., deren Hauptsitz im kalifornischen Ort Mountain View liegt. Google wurde durch Internetdienstleistungen, insbesondere durch die gleichnamige Suchmaschine „Google“ bekannt. Andere bekannte Produkte sind die Video-Plattform YouTube, der E-Mail-Anbieter Gmail und der Webbrowser Chrome.

Der Echtbetrieb der Suchmaschine startete 1996 unter dem Namen BackRub, seit 1998 agiert sie unter dem Namen Google. In Deutschland zeichnet sich der Marktführer Google Inc. mit 90% der Marktanteile aus und ist darüber hinaus auch in 4 von 5 Kontinenten Marktführer unter den Suchmaschinen. Jährlich werden 1,2 Billionen Suchanfragen weltweit an Google gestellt, 53 Milliarden Suchen gehen jährlich allein aus Deutschland ein.



Verbreitung/Historie

Die erste Sprachsuch-Funktion von Google wurde im Jahr 2007 ausschließlich in den USA veröffentlicht. Es war eine kostenlose Telefonauskunft namens GOOG-411, die ein lokales Branchenverzeichnis bot (Apretrica.com, 2015). Mittels Sprachbefehlen konnte man sich z.B. direkt mit dem gewünschten Geschäft verbinden, Restaurants in der Nähe durchsagen oder eine Wegbeschreibung per SMS zukommen lassen (Searchengineland.com, 2015a). Im Oktober 2010 stellte Google den Service von GOOG-411 ein (Googleblog.blogspot.de, 2015a). Vorher veröffentlichte Google allerdings eine Voice Search für Google Maps ausschließlich für Blackberry, die auf der Engine von GOOG-411 basiert (searchengineland.com, 2015b). Später in diesem Jahr wurde eine reine Voice Search für iPhone auf den Markt gebracht. 2009 folgte diese auch für Blackberry (Searchengineland.com, 2015c). Die ersten Versionen erschienen in englischer Sprache, bis 2010 folgten Sprachen für den chinesischen Markt sowie Deutsch, Französisch, Italienisch und Spanisch (Googlemobile.blogspot.de, 2015b).

Die Anwendung Voice Search wurde im August 2010 um weitere Befehle für die Sprachsteuerung erweitert. Zu den sog. „Voice Actions“ gehört das Anrufen von Kontakten und das Diktieren von Nachrichten. Die Google-Maps-Navigation ließ sich nun per Sprachbefehl aufrufen und einstellen. Auch den Musik-Player, den Browser und die Google-Suche können Android-Nutzer anhand des Features über die Stimme aufrufen und steuern, bis dato war die Sprachsteuerung nur als englische Version verfügbar (Searchengineland.com, 2015d).

Die Eingabe von Suchbegriffen via Mikrofon stellte Google 2011 als Browser-Erweiterung für Chrome bereit (Searchengineland.com, 2015e).

Am 9. Juli 2012 veröffentlichte Google seine Version eines intelligenten persönlichen Assistenten Google Now für Android-Systeme. Die Anwendung sollte Konkurrenzprodukten wie Apples Siri ebenbürtig sein. Google Now bietet eine weiterentwickelte Spracherkennungssoftware (Google Voice Search) zur Verarbeitung von natürlich gesprochener Sprache.



Datenbasis

Die Daten auf die die Google Voice Search zugreift, zieht die Suchmaschine aus dem eigenen Index. Die Ergebnisse der Google Voice Search können sowohl durch direkte Antworten sprachlich wiedergegeben als auch über die herkömmliche Web-Search-Ergebnisliste von Google angezeigt werden.



Technik

Der Nutzer muss seine Suchanfrage nicht mehr in das Smartphone eintippen, sondern kann diese über Spracheingabe einfügen. Die Ergebnisse der Anfrage werden vom Google-eigenen Index geliefert.

Suchanfragen sollen von der Sprachsteuerung als Frage erkannt und bearbeitet werden, z. B. werden bei der Eingabe, dies gilt auch für die Spracheingabe: „Eiffelturm“ und „Lage“, Dokumente geliefert, die beide Schlagworte enthalten. Das vermeintliche Suchziel, nämlich die Lage des Eiffelturms, könnte in diesen Dokumenten enthalten sein. Anders verhalten sich die Suchergebnisse bei der Suchanfrage: „Wo ist der Eiffelturm?“, hier versucht Google nach der Essenz (dem Suchziel) der Frage zu suchen, um passende Ergebnisse zu liefern. Das heißt, Google sucht nicht nach Dokumenten, die die eingegebenen Schlagworte (wo, ist, der, Eiffelturm) enthalten, sondern nach diesen, die die Frage beantworten. Google gleicht sich so der „alltäglichen Sprache“ an.

Die Ergebnisse der Voice Search sollen eine eindeutige Antwort liefern und nicht in der gewohnten Auflistung der Google-Ergebnisliste erscheinen. Die Antwort wird entweder in Form eines Knowledge Graphs und/oder gesprochen ausgegeben. Bei dieser Vorgehensweise, dem genauen Antworten auf gezielte Fragen, lernt das System stetig. Fragestellungen werden aufgenommen und verarbeitet, um künftig Anfragen noch besser zu erkennen. Namen und

Sprachmuster wie Dialekte, Akzente und Sprachfehler können durch diese Herangehensweise nach mehrmaliger Eingabe schneller und leichter vom System verarbeitet werden.



Funktionen

Die wesentlichen Funktionen der sprachgesteuerten Websuche beziehen sich vorwiegend auf die Spracherkennung und das Verstehen von Suchanfragen, um diese möglichst treffsicher mit Web-Search-Ergebnissen zu beantworten. „OK Google“ kategorisiert die Befehle in folgende Cluster: Navigation, Unterhaltung und Wissen.



Dialogfähigkeit

Die sprachgesteuerte Suche von Google kann sich bei erneuten Suchanfragen auf unmittelbar zuvor gestellte Suchanfragen beziehen, bspw.: „*Wie heißt der Präsident der Vereinigten Staaten?*“ → *Antwort* → „*Wie groß ist er?*“ → *Antwort*. Im Gegensatz zu Siri kann mit der Google Voice Search keine Unterhaltung geführt werden.



Entwicklungen und Ausblick

Für die Zukunft der Google Voice Search ist ein Feature geplant, welches standortbezogene Suchanfragen mit wissensbezogenen Suchanfragen verbinden soll. Der Nutzer soll also fragen können, wie hoch der Kirchturm der Kirche ist, vor der er in diesem Moment steht. Ein weiteres denkbares Informationsbedürfnis für den Nutzer der Websuche wäre die nach den Öffnungszeiten des Restaurants, in dem dieser gerade wartet.

Einige Fragen können bereits in der englischen Version beantwortet werden. Wann dieses Feature für die deutsche Version eingeführt wird, ist noch nicht bekannt (Androitpit.de, 2015). Zusätzlich will Google künftig auch Einkäufe über die Sprachsteuerung abwickeln lassen können (huffingtonpost.com, 2015).

Apple Siri



Anbieter

Apple Inc. ist ein kalifornisches Unternehmen mit Sitz in Cupertino, das Computer, Smartphones und Unterhaltungselektronik sowie Betriebssysteme und Anwendungssoftware entwickelt und vertreibt. Das Unternehmen wurde von Steve Jobs, Steve Wozniak und Ron Wayne 1976 gegründet.



Verbreitung/Historie

Die Sprachsteuerungssoftware Siri wurde von Apple 2010 übernommen und in die Apple-Geräte integriert. Siri Inc. war vorher ein eigenständiges Unternehmen und wurde mittlerweile von Apple aufgekauft.

Siri ist auf jedem iPhone, iPad und iPod integriert. Der Absatz vom iPhone mit der Sprachsteuerung steigt jedes Jahr; in 2015 wurden ca. 78 Millionen Geräte verkauft (Schimanke.com, 2016; Statista.com, 2007).

Siri ist auf folgenden Geräten verfügbar:

- iPhone 4S oder neuer
- iPad (3. Generation oder neuer)
- iPad mini
- iPod touch (5. Generation)



Datenbasis

Damit Apples Spracherkennung und Sprachsteuerung Siri Eingabebefehle ausführen kann, müssen diese via WLAN oder Datennetzwerk zu Apples Servern geschickt werden. Dort verarbeitet ein System dann den Inhalt und überträgt das Ergebnis zurück auf das jeweilige Gerät, z. B. iPhone oder iPad. Dieses Vorgehen ist erforderlich, da sonst eine riesige Datenbank auf jedem Mobilgerät gespeichert werden müsste, die Apple zudem fortwährend zu aktualisieren hätte.

Apple hat sich nun zu dieser Frage geäußert und angegeben, was bei Siri im Hintergrund abläuft. Der Sprachbefehl wird an Apples Rechencenter zur Analyse geschickt. Apple erzeugt dabei eine zufällige Nummer, die Sprachdatei und Nutzer miteinander verbindet. Diese Nummer

steht nicht in Verbindung mit Apple ID oder Mailadresse des Nutzers. Sechs Monate lang speichert Apple die Spracheingaben zusammen mit jener eindeutigen Nummer, anschließend wird sie aus der Datei gelöscht. Der gespeicherte Audioschnipsel hingegen verbleibt noch länger auf dem Server, nämlich insgesamt zwei Jahre lang. Apple verwendet diese Dateien, um Siri zu testen und den Dienst weiter zu verbessern (MacTechNews.de, 2016).

Apple Siri bezieht seine Daten von Bing und Google. Apple hat keine eigene Datenbank (Statista.com, 2016).



Technik

Siri ist auf drei Säulen aufgebaut:

- der reinen Spracherkennung (welche Wörter hat der Nutzer gesagt)
- der Kontexterkennung (geht es um eine Route oder einen Termin?)
- der Anbindung an entsprechende Systemfunktionen und Datenbanken.

Siri kann als persönlicher Assistent genutzt werden. Der digitale Assistent soll nun das Leben des Nutzers kennen, analysieren und aus den gewonnenen Erkenntnissen bessere Vorschläge machen. Ist etwa Stau auf dem Weg zu einem Termin, könnte Siri früher an die Abfahrt erinnern. Der Assistent kann nun auch Musik erkennen und Titel anzeigen oder Apps vorschlagen. Der Assistent Siri besteht im Prinzip aus 2 Komponenten – dem lokalen Teil auf dem iPhone (gewissermaßen dem GUI) und dem serverbasierten Teil (dem Apple-Server, auch bekannt als: guzzoni.apple.com). Sobald der Nutzer anfängt Siri etwas zu erzählen, verbindet sich Siri mit dem Apple-Siri-Server und schickt das Gesprochene dorthin (Focus.de, 2016). Wie bei Applidium beschrieben erfolgt dies via Speex Audio Codec (Sample Rate 16kHz). Soweit ersichtlich werden die oft angesprochenen „A5 Sprachchipmöglichkeiten“ z.Z. nicht genutzt (macwelt.de, 2016). Dies wiederum macht es auch möglich, Siri ohne Probleme auf anderen Devices laufen zu lassen.



Funktionen

Für die Aktivierung der Siri-Funktion muss einfach der Home-Button gedrückt und dabei die gewünschte Frage ausgesprochen werden. Wenn iOS 8 oder eine neuere Version benutzt wird, kann die Aktivierung durch die Sprachsteuerung mit „Hey, Siri“ erfolgen. Über das Headset/Kopfhörer kann Siri über die Mitteltaste aktiviert werden. Siri speichert die Stimme des Nutzers und reagiert dementsprechend nur beim Nutzer des Gerätes. Somit wird der Zugriff für

fremde Personen verhindert. Die Sprachfunktion Siri ist sofort ab der ersten Benutzung startklar und muss nicht vorher installiert oder programmiert werden (Computerbild.de, o. J.). Je öfter Siri verwendet wird, desto mehr Merkmale kann sich Siri über die Person merken. Je häufiger die Siri-Funktion genutzt wird, desto besser werden die Ergebnisse zu den Suchanfragen (Computerbild.de, o. J.). Die Antworten werden im Laufe der Zeit automatisch personalisiert (Datenbank), um besser auf die Bedürfnisse des Nutzers einzugehen und ihm relevantere Ergebnisse zu bieten.

Apple Siri kann insgesamt folgende Sprachen verstehen und anwenden (Computerbild.de, o. J.):

- USA (Englisch, Spanisch)
- Großbritannien (Englisch)
- Australien (Englisch)
- Frankreich (Französisch)
- Deutschland (Deutsch)
- Japan (Japanisch)
- Kanada (Kanadisches Französisch, Englisch)
- China (Mandarin)
- Hongkong (Kantonesisch)
- Italien (Italienisch)
- Korea (Koreanisch)
- Mexiko (Spanisch)
- Spanien (Spanisch)
- Schweiz (Französisch, Deutsch, Italienisch)
- Taiwan (Mandarin).



Dialogfähigkeit

Siri ist eine dialogfähige Funktion, die dem Nutzer hilft Fragen zu beantworten. Jedoch ist diese durchaus noch ausbaufähig und bedarf Verbesserungen. Nicht immer versteht Siri die Anweisungen, die der Nutzer der Sprachsteuerung stellt, doch kann sie sich auf das Eingesprochene beziehen und den Dialog einhalten.



Entwicklungen und Ausblick

Siri soll in Zukunft auch als Anrufbeantworter genutzt werden können. Der Sprachassistent soll nicht nur weiter verbessert werden, sondern in Zukunft noch weitere Funktionen bekommen. Jedoch befindet sich die Verbesserung momentan noch im internen Test. Daher kann Apple keine weiteren Informationen zu dem Thema geben. Business-Insider haben sich Siri als einen persönlichen Assistenten gewünscht. Doch auf diesen Wunsch kann Apple nicht eingehen, daher wird dieser Wunsch nicht umgesetzt.

IOS 9 – Die wichtigsten Neuerungen auf einen Blick: Der Akku soll ab sofort länger halten; laut Apple im Schnitt ca. 3 Stunden länger als bisher. Die Apple-Karte-App hat mit dem iOS 9 einen neuen Schalter, um die Routen im Nahverkehr besser anzeigen zu können. Für Berlin kann man mit dem neuesten iPhone sogar auf den Bahnverkehr auf der Apple-Karte zugreifen, bei anderen Städten klappt dieses aber leider noch nicht. Siri soll in Zukunft sprachlich besser optimiert werden und ca. 40% schneller reagieren.

Die Termine werden nun auch schneller geöffnet. Eine Meldung an Siri genügt und Siri greift ab sofort auf die eigenen Termine zu. Auch Internetseiten werden gespeichert und individuell angepasst. Der Nutzer braucht nur die Home-Taste gedrückt halten und hineinsprechen: „Erinnere mich an die Seite XY“. Man muss jetzt auch nicht mehr die eigenen Bilder durchsuchen, wenn man nach einem bestimmten Bild sucht. Siri ist ab sofort auch bei der Bildersuche behilflich. Am Design hat sich nur minimal etwas verändert.

Was Apple in Zukunft noch alles auf den Markt bringen wird, kann nicht vorausgesagt werden, da diese Daten geheim gehalten werden.

Microsoft Cortana



Anbieter

Der neue Sprachassistent von Microsoft heißt Cortana und tritt damit in die Fußstapfen von Siri und Google Now. Nachdem die Sprachsteuerung zunächst nur für Windows Phones konzipiert war, soll der Microsoft Assistent nach und nach auch auf fremden Systemen laufen. Darüber hinaus geht Microsoft eine Partnerschaft mit Cyanogen ein (ZDNet.de, 2015). Cyanogen OS ist eine alternative Android-Version.



Verbreitung/Historie

Cortana wurde im Rahmen der Microsoft-Konferenz Build in April 2014 durch Joe Belfiore präsentiert (Microsoft.com, 2014). Mit der Developer Preview von Windows Phone 8.1 konnten zunächst Nutzer in den USA die Beta-Version des Service testen. Das Windows Phone 8.1. Update (Juli 2014) ermöglichte die Nutzung der Beta-Version auch in China und Großbritannien sowie der Alpha-Version in Australien und Canada (vgl. ebd). Seit einem weiteren Update im Dezember 2014 war Cortana neben Frankreich, Italien und Spanien auch in Deutschland als Alpha-Version verfügbar (Windows.com, 2014). Im August 2015 startete Microsoft die öffentliche Beta-Version des Sprachassistenten auf Android-Geräten (Windows.com, 2015).

Cortana ist derzeit der einzige Assistent, der sowohl auf Smartphones als auch auf Tablets, Notebooks und PCs funktioniert. Seit Windows 10 ist der Sprachassistent auf Desktop- und Tablet-Betriebssystemen verfügbar. Zudem soll 2016 eine Version für die Xbox One erscheinen. Microsoft beschränkt seinen Service somit nicht nur auf das eigene System, sondern stellt es für Systeme der Konkurrenz kostenlos zur Verfügung. Während die Android-Version als Beta bereits seit Mitte 2015 verfügbar ist, arbeitet Microsoft noch immer an einer Version für iOS-Geräte. Ende 2015 sollen demnach auch iPhone-Nutzer Cortana als Alternative zu Siri nutzen können (ZDNet.de, 2015).



Datenbasis

Cortana benutzt grundsätzlich auf allen Endgeräten Bing als Suchmaschine (winfuture.de, 2015). So hat oftmals eine Frage oder ein Befehl nur eine Ergebnisliste der hauseigenen Maschine Bing zur Folge(giga.de, 2015). Es ist möglich, durch Nutzung von Google Chrome, über AddIns wie Bing2Google, auf Google zu wechseln (ZDNet.de, 2015). Cortana lernt den

Nutzer zudem, z. B. durch Analyse des Browserverlaufs, immer besser kennen. Dazu legt das System eine Datenbank an. Die Daten werden jedoch, anders als bei Google Now und Siri, direkt auf dem Gerät gespeichert. Cortana erfasst und nutzt außerdem Ortungsdaten, Spracheingaben, Kontakte, Informationen aus E-Mails und SMS, Kalenderdetails und andere Daten (Windowsphone.com, 2015).



Technik

Cortana ist auf den Betriebssysteme Windows 8.1, 10, Windows Phone 8.1 und Windows 10 Mobile Xbox One lauffähig. Microsoft plant ebenfalls Apps für Android und iOS zu entwickeln und zu testen, allerdings lokal beschränkt, sodass diese Testversionen im Google Play Store bisher noch nicht in Deutschland verfügbar waren. Laut Angabe im Google Play Store wäre Cortana ab Android 4.0 installierbar. Allerdings wurde bei Kompatibilitätsprüfung vor der Installation kein registriertes Gerät, worauf Android >4.0 installiert war, gefunden.



Funktionen

Cortana ist eine dialogorientierte, verstehende persönliche Assistentin, die den Nutzer bei verschiedenen Anforderungen unterstützt. Durch ihre Lernfähigkeit kann sie Themen vorausschauend behandeln und Empfehlungen aussprechen (Windowsphone.com, 2015).

Basisfunktion: Notizbuch: Dreh- und Angelpunkt der Kommunikation mit Cortana ist das Notizbuch. Hier hinterlegt der Nutzer seine persönlichen Interessen, favorisierte Themen, Vorlieben, Ruhezeiten, Adressen und Kontakte. Die Informationen können jederzeit verfeinert oder wieder gelöscht werden. Basierend auf diesen Informationen macht Cortana selbständig Vorschläge und bietet Inhalte an.

Intelligente Erinnerung: Durch den Einsatz der sog. Geofencing-Technik verknüpft Cortana Erinnerungen mit bestimmten Personen oder Orten und agiert dann proaktiv. So wird der Nutzer bspw. daran erinnert, bestimmte Lebensmittel einzukaufen, wenn er sich in der Nähe eines Supermarkts befindet. Ähnlich funktioniert die personenbezogene Erinnerung bei Kontakten.

Telefon- und Terminorganisation: Cortana ist in der Lage, die Kommunikation und die Termine des Nutzers zu organisieren. So kann sie beispielweise Anrufe starten, SMS oder Emails schreiben oder Smartphone und PC nach Kontakten, Musikdateien oder Kalendereinträgen durchsuchen. Außerdem macht sie Notizen und nimmt Kalendereinträge vor

oder ändert diese. Gibt es dabei Überschneidungen, informiert Cortana den Anwender und macht Vorschläge zur Lösung.

Echtzeit-Planung: Hat ein Nutzer z. B. zu einem Termineintrag eine Adresse hinterlegt, wird Cortana ihn rechtzeitig darauf hinweisen aufzubrechen, unter Berücksichtigung der aktuellen Verkehrslage und des aktuellen Standorts. Bei der Reiseplanung macht sie darauf aufmerksam, wenn es Zeit wird, zum Flughafen aufzubrechen, unter Berücksichtigung der aktuellen Verkehrslage. Außerdem zeigt sie aktuelle Informationen zum gebuchten Flug an und liefert Wetterinformationen zum Ankunftsort.

Information: Cortana nutzt die Bing-Suchmaschine zur Beantwortung von Fragen aller Art. Bestimmte Themen wie die zur Verkehrs- und Wetterlage sind bereits vorinstalliert. Außerdem hat sie Zugriff auf die aktuellen Entwicklungen der Finanzmärkte und stellt Neuigkeiten in vorgegebenen Nachrichtenkategorien zusammen. Zusätzlich erkennt Cortana sehenswerte Orte in der Umgebung selbstständig und gibt einen entsprechenden Hinweis.

Synchronisation und Schnittstellen: Cortana lässt sich mit Apps und anderen Anwendungen intelligent verknüpfen. So kann bspw. direkt eine Musik-App starten, wenn der Name eines Liedes oder einer Band genannt wird. Desktop-PC und Smartphone werden automatisch synchronisiert, sofern das gleiche Windows-Konto verwendet wird. Tabs, Lesezeichen, WLAN-Einstellungen oder Design-Konfiguration stehen so auf allen Geräten gleich zur Verfügung.

Einsatz in Unternehmen: Microsoft wird Cortana so maximal wie möglich in andere Produkte integrieren. Durch eine Kopplung mit Microsoft Dynamics CRM können so Meetings oder Termine angelegt, Kontakte, Kundenkonten und -listen gesucht oder Aktivitäten und Einträge erstellt werden. Auch eine Integration der digitalen Assistentin in die Office-Apps steht auf dem Plan.



Dialogfähigkeit

Neben der reinen Websuche, in der auch Siri und Google Now dem Entwicklungsstand entsprechend gute Performance erbringen, wird Cortana vor allem die Dialogfähigkeit in den Medien zu Gute gehalten. Hierbei ist der Sprachassistent in der Lage verschachtelte Dialoge zu erkennen und eine korrekte Antwort zu liefern. „Beispiel: Stellt man Cortana die Frage ‘Wie heißt der amtierende US-Präsident?’, erhält man als Antwort: Barack Obama. Fragt man anschließend: ‘Wie heißt seine Frau?’, erkennt Cortana, dass mit „seine“ der Bezug zum US-Präsidenten gemeint ist und antwortet korrekt: Michelle Obama.“ (computerbild.de, 2015).

Neben der englischen Sprache unterstützt Cortana in gleicher Weise auch die deutsche, somit kann Dialogfähigkeit auch bei deutschen Anfragen in vollem Umfang genutzt werden.



Entwicklungen und Ausblick

Seit 2015 gibt es Cortana nun auch in Deutschland und seitdem haben bereits viele verschiedene technische Fachmagazine über den neuen Sprachassistenten von Microsoft berichtet. Schaut man sich die Grundtonalität an, kann man von einer durchaus positiven Berichterstattung sprechen (heise.de, computerbild.de, 2015). Neben einer Terminverwaltung werden Fragen aller Art beantwortet, als Quelle für die Websuche setzt Microsoft auf die eigene Suchmaschine Bing.

Cortana verfolgt seit ihrer Erscheinung sehr ehrgeizige Ziele. So berichtet bspw. heise.de von „Windows in the car“: Damit möchte Microsoft Cortana die aus Windows 10 bekannte Helferin auf die Windschutzscheibe von PKWs bringen, die per Sprachbefehl z. B. Orte auf einer virtuellen Karte anzeigen kann. Ähnlich wie Apples CarPlay oder Googles Android Auto sollen Mobilgerät und Auto eng miteinander verzahnt werden (heise.de, 2015).

Neben Android sollen auch Smartphone-Nutzer von anderen Betriebssystemen in die Gelegenheit kommen, Cortana zu nutzen. Der Funktionsumfang soll allerdings geringer sein als auf Windows Phones (heise.de, 2015), weil keine tiefe Integration in fremde Systeme möglich ist.

XBOX-One-Besitzer können sich laut Microsoft freuen. Cortana soll es nämlich auch auf der XBOX geben. Geplant ist diese Erweiterung in Deutschland für das Jahr 2016. Laut der Winfuture (winfuture.de, 2015) soll die Sprachassistentin „per Sprache u. a. die Interaktion mit Freunden erleichtern. Man kann fragen, wer gerade online ist, eine Party starten und vieles mehr und das alles, ohne das Spiel unterbrechen zu müssen.“ Die Kombination aus der bisherigen XBOX-Spracheingabe und Cortana soll darüber hinaus auch kompliziertere Befehlseingaben ermöglichen.

Da sich Cortana unter anderem auch als Diktiergerät gut eigne, ist Microsoft auch daran interessiert, die hauseigene Sprachassistentin als Unterstützung in den kommenden Internet-Browser „Spartan“ zu integrieren (golem.de, 2015).

Amazon Fire OS



Anbieter

Der Anbieter des Fire Phones und somit dessen Sprachassistenten ist Amazon.com, welches ein börsennotiertes amerikanisches Unternehmen ist, das sich auf den Online-Versandhandel einer breit gefächerten Produktpalette spezialisiert hat (Wikipedia.org, 2015).



Verbreitung/Historie

Das Fire Phone begann wie viele gute Erfindungen zunächst auf dem Papier als Konzeptzeichnung mit seinen geplanten Funktionen. Um entscheiden zu können, wie das spätere Gerät auszusehen hat und welche Funktionen es haben soll, wurden mehr als drei Dutzend Mitarbeiter befragt. Das gesamte Telefon sollte an den Wünschen der Kunden angepasst werden. Es sollte das erste Smartphone mit vollständigem, ohne Hilfsmittel nutzbaren 3D-Effekt sein. Mit dieser Art von 3D sollte es möglich sein, alle Symbole, Icons, Spiele sowie vieles mehr am Blickwinkel des Auges und der Haltung des Geräts auszurichten und dadurch in 3D darzustellen. 2013 wurde dann alles, was bis dahin konzipiert worden war, verworfen und durch eine verschlankte, mehr auf ein günstiges Gerät ausgelegte Version ersetzt. Da der Hauptverantwortliche jedoch im Gegensatz zum Rest des Teams auf den 3D-Effekt bestand, hat man sich schlussendlich doch für einen höheren Preis mit besserer Hardware entschieden.

Der Plan war weiterhin, Amazon zu einer Lifestyle-Marke wie Apple zu machen (Carr, 2015). Dass das Unternehmen mit diesem Vorhaben jedoch nur wenig erfolgreich war, verdeutlichen die schlechten Verkaufszahlen. So hatte Amazon in 2014 noch immer Fire Phones im Wert von 83 Millionen Dollar auf Lager (Luckerson, 2014).

In den USA konnte man das Fire Phone ab dem 25. Juli 2014 erwerben (Winslow, 2014), in Deutschland ab dem 30. September desselben Jahres (DMEurope, 2014).

Die Recherche zeigt, dass Statista Amazon nicht als eigenständigen Smartphone-Hersteller sowie auch Fire OS nicht als eigenständiges Betriebssystem berücksichtigt, weshalb genauere Zahlen zur Verbreitung des Geräts und somit dessen Sprachsteuerung nicht ermittelt werden konnten.



Datenbasis

Spracheingaben, die über das Fire Phone getätigt werden, werden in der Amazon-Cloud gespeichert und im weiteren Verlauf der Nutzung dafür verwendet, die Nutzererfahrung zu verbessern (Amazon.de, 2015a, S. 39).

Ob Amazon einen Index oder andere Quellen als Grundlage für die Auswertung der Spracheingaben benutzt, ist nicht bekannt.



Technik

Die Sprachsteuerung des Fire Phones trägt keinen Namen und ist abzugrenzen von anderen sprachbasierten Amazon-Systemen wie bspw. „Alexa“. Leider existieren für die sprachgesteuerte Suche des Fire Phones keinerlei Quellen, die speziell auf die Technik dieses Gerätes eingehen.

Die Funktionsweise einer Sprachsteuerung funktioniert im Kern allerdings immer gleich, sodass davon ausgegangen werden kann, dass es auch beim Fire Phone zu keinen großen Abweichungen kommt.

Die Sprachbefehle werden aufgenommen, in ein spezielles komprimiertes Format codiert (welches, ist nicht bekannt) und dann an die jeweiligen Server hinter dem Dienst übermittelt. Die Server verarbeiten die Anfrage und senden eine passende Antwort zurück. Da die benötigte Rechenleistung, um diese Anfragen zu bearbeiten, sehr groß ist, muss man online sein, um mit den Servern kommunizieren zu können. „Je nach Algorithmus hinter den einzelnen Lösungen ist die Erkennung der Wörter mehr oder weniger genau“ (Go2android.de, 2014).

Die Algorithmen im Detail sind von Firma zu Firma natürlich ein wohlgehütetes Geheimnis, allerdings basieren alle auf den gleichen Grundlagen. „Es ist nämlich eine Mischung aus Wahrscheinlichkeitsrechnung und Datenbank-Abfragen, die dabei zum Einsatz kommt. Wirklich verstehen kann der Computer ‘Sprache’ nicht. Sprachmuster, die das System erkennt, werden mit denen bereits gespeicherter Muster verglichen“ (Martin-Jung, 2011).



Funktionen

„To search the Web, hold down the Home button, and then say, ‘Search’ or ‘Search the Web.’ Dictate your search term. The Fire launches the Silk browser, searches the Web, and displays the results as it does for any web search“ (Gralla, 2014, S. 84).

„1. Halten Sie die Startseite-Taste gedrückt und sagen Sie ‘Suchen’ oder ‘im Internet suchen’.

2. Diktieren Sie, nach was Sie suchen, wie zum Beispiel ‘Kinos’ oder ‘Museen’ und schauen Sie sich die Ergebnisse im Silk-Browser an“ (Amazon.de, 2015a, S. 38).

Die Richtung finden

- ‘Richtung für Tate Modern in London, UK finden’
- ‘Hotels in Berlin, Deutschland auf einer Karte anzeigen’
- ‘550 Terry Ave in Seattle, Washington finden’

Allgemeine Fragen

- ‘Wer ist der Präsident von Frankreich?’
- ‘Wie viele Gramm entspricht ein Teelöffel Mehl?’
- ‘Welche Lebensdauer hat ein Elefant?’ (Amazon.de, 2015b)”



Dialogfähigkeit

Die spärlich gesäten Nutzerberichte über die Sprachsuche des Fire Phones haben gezeigt, dass ein Dialog außerhalb der Standardkommandos, auf die das Gerät reagiert, nicht möglich ist. So werden bspw. Anfragen, die einen Dialog erwarten ließen, in eine Websuche eingegeben, anstatt direkt in ausformulierten Antworten an den Benutzer auszugeben, oder lediglich mit einer Fehlermeldung quittiert. Zudem werden die Anfragen nicht kontextuell eingeordnet, was zur Folge hat, dass für jede Suche ein neues Browserfenster mit einer neuen Anfrage geöffnet wird, anstatt die bestehenden Suchergebnisse mit einzubeziehen (Youtube.com, 2015).

Als beispielhaft für diesen Mangel kann folgende Feststellung nach der Eingabe einer Navigationsfrage gesehen werden: „And if you ask it for directions to the nearest gas station, it will return a query for ‘gas_station’ on the Yelp mobile site. Lots of room for improvement here“ (Sadauskas, 2014).



Entwicklungen und Ausblick

Nach den desaströsen Verkaufszahlen und der Einstellung des Verkaufs des Fire Phones zeigte sich Jeff Bezos, Gründer und Präsident Amazons, Ende des Jahres 2014 dennoch in seinem Glauben an das Fire Phone trotz Misserfolgs unerschüttert und versprach eine Weiterentwicklung des Geräts (Horizont, 2014). Daraufhin folgten seitens Amazons unbestätigte

Berichte über die Veröffentlichung sowie ebenso inoffizielle Meldungen darüber, dass die Entwicklung eines „Fire Phone 2“ vorerst beendet wurde (Brodersen, 2014).

Auch andere Quellen schätzen daraufhin die Bekanntgabe eines Erscheinungsdatums für eine zweite Version des Smartphones als eher unwahrscheinlich ein, obwohl der amerikanische Konzern verkraftbare finanzielle Einbußen hinnehmen musste (Haughton, 2015).

Laut eines Berichts des Wall Street Journals (Bensinger, 2015) wurden Dutzende Mitarbeiter des „Lab126“, das seit 2014 Hardware-Produkte für Amazon entwickelt haben soll – darunter vermutlich auch das Fire Phone, im Sommer 2015 entlassen. Angeblich führte diese Entlassungswelle auch zu der Einstellung von aktuellen Projekten, wie vertraute Personen, auf die sich das Wall Street Journal bezieht, berichtet haben sollen. Folglich kann dieser Umstand die Argumente weiter unterstützen, nach denen ein Erscheinen der zweiten Version des Fire Phones als unwahrscheinlich eingeschätzt wird.

Zuletzt wurde darüber spekuliert, ob das neue Fire Phone 2 womöglich im Jahr 2016 auf dem Markt erscheinen wird (Chip.de, 2015). Andere Quellen ziehen die Möglichkeit einer Veröffentlichung noch vor Beginn des neuen Jahres in Betracht. Das Fire Phone 2 soll nach diesen Spekulationen unter anderem über eine verbesserte Performance sowie größere Anteile von Google in Form von Android 6.0 Marshmallow als Betriebssystem auf den Geräten verfügen (Banks, 2015). Offizielle Pressemeldungen seitens Amazon bleiben jedoch weiterhin aus.



Gründe für den Misserfolg des Fire Phones sowie dessen Sprachsteuerung

Die Ursachen, warum das Fire Phone bisher zu keinem Erfolg wurde, sind vielfältig. So wird es als technisch mittelmäßiges Gerät bezeichnet, dessen hoher Preis angesichts seiner an Innovationen vergleichsweise armen Ausstattung als ungerechtfertigt erscheint (Android Advisor, 2014, S. 41). Dass es die wenigen Besonderheiten des Geräts, wie bspw. dessen 3D-Display, zudem nicht schafften, bei den potentiellen Käufern Begeisterung auszulösen, sorgte für schwierige Grundvoraussetzungen des Geräts bereits bei Markteinführung (Luckerson, 2014).

Das Geschäftsmodell des Fire Phones wird ebenso als problematisch erachtet. Ein Smartphone, das sich preislich zumindest bei Markteinführung im Bereich des iPhones bewegte, jedoch ähnlich wie das hingegen niedrigpreisige Kindle-Tablet stark auf Amazon-eigene Dienste und

Inhalte ausgelegt ist, wirkt auf mögliche Interessenten wenig attraktiv (Android Advisor, 2014, S. 44).

Zwar wurde der Anfangspreis von 199 US-Dollar bei einem Zweijahresvertrag nach Erkennen des Misserfolgs zügig auf 99 US-Cent reduziert, dies trug jedoch zu keiner spürbaren Kehrtwende bei (Luckerson, 2014). Auch in Deutschland hatte das Smartphone schon zu Beginn an einen schweren Stand, da das Gerät auf der Amazon-Seite selbst von den Kunden als schlecht beurteilt wurde. Erschwerend kam hinzu, dass für kurze Zeit fälschlicherweise ein Preis von 10 Euro angegeben worden ist. Bestellungen zu diesem Preis wurden storniert, was auf Twitter mit Häme quittiert wurde (Saal, 2014).

Im Vergleich mit Amazons Konkurrenten werden weitere Gründe für den Misserfolg des Fire Phones erkennbar. Amazon verfügt über einen verhältnismäßig kleinen App-Store, der in etwa ein Viertel so umfangreich ist wie der von Google. Des Weiteren ist der Einstieg Amazons in den Smartphone-Markt sehr spät gewählt. Apple- und Android-Geräte wurden bereits in zahlreichen Generationen vertrieben und stetig optimiert (Luckerson, 2014).

Die unter Funktionen und Dialogfähigkeit beschriebenen wenigen und im Vergleich zur Konkurrenz zudem schwachen Funktionen der Sprachsteuerung scheinen daher im Gesamtkontext nicht die ausschlaggebenden Gründe dafür zu sein, warum sowohl die Verkaufszahlen als auch die Quellenlage mit Bezug zum Amazon Fire Phone und dessen Sprachsteuerung überschaubar sind. Vielmehr sorgen die grundlegenden Schwächen und die dadurch begrenzte Verbreitung des Geräts wohl selbst dafür, dass die, wie unter 1. Anbieter genannte, exklusiv auf dem Fire Phone verfügbare Sprachsteuerung bisher nur wenig Beachtung finden konnte.

Methodenbeschreibung

Das folgende Kapitel stellt das Studiendesign vor und beschreibt die im Rahmen dieses Projektes angewandten Methoden. Eine Übersicht fasst alle Methoden zusammen; im Anschluss werden experten- und nutzerbasierte Methoden im Detail vorgestellt.

Übersicht

Usability-Evaluation als systematische und möglichst objektive Bewertung der Gebrauchstauglichkeit eines Gegenstandes oder einer Anwendung kann entweder aus Experten- oder Nutzerperspektive heraus erfolgen. Expertenbasierte Methoden greifen auf Usability-Experten zurück, die sich entweder in den Nutzer hineinversetzen und versuchen, so „durch die Brille des Nutzers zu schauen“ oder aber aus ihrer Erfahrung heraus Verstöße gegen etablierte Usability-Kriterien identifizieren können. Nutzerbasierte Methoden hingegen nehmen den realen Anwender in den Blick und erfahren durch diesen, wo Stärken und Schwächen des Testgegenstandes liegen. Im Rahmen des vorliegenden Projektes wurden beide Methodenansätze miteinander kombiniert.

	Usability	Ergebnisqualität
Expertenbasiert	Heuristische Evaluation	Retrievaltest
	Szenarienbasierter Walkthrough	
Nutzerbasiert	Vorabbefragung	Nutzertest Teilnehmende Beobachtung Retrospective Thinking Aloud
	Nutzertest	
	Teilnehmende Beobachtung	
	Retrospective Thinking Aloud	
	Nachbefragung	
	SUS-Fragebogen	

Tabelle 3: Forschungsdesign

Dabei orientieren wir uns an einem erweiterten Modell des Suchprozesses (in Anlehnung an Quirmbach 2013) und an den sog. Micro-Moments, die von Google in Bezug auf das Nutzungsverhalten mobiler Websuche identifiziert wurden (Google, 2015). Mobile Suchen sind im Gegensatz zur Suche an stationären Geräten praktisch von überall möglich und werden stärker von Spontaneität und Ungeduld der Nutzer getrieben.

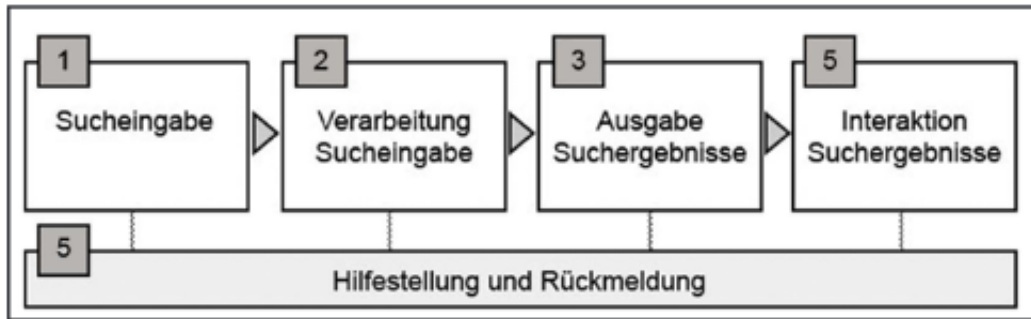


Abbildung 1: Modell Suchprozess (Quirnbach, 2013, S. 30)

Micro-Moment	Suchintention	Beispiel
WANT-TO-KNOW	informationsorientierte Suchanfragen	Ein Nutzer will sich über ein Produkt informieren, das er in diesem Moment in einer TV-Werbung sieht.
WANT-TO-GO	lokale Suche	Ein Nutzer sucht nach einem italienischen Restaurant in seiner Nähe.
WANT-TO-DO	Howto-Suche	Ein Nutzer sucht nach einem Rezept für einen Apfelkuchen.
WANT-TO-BUY	transaktionsorientierte Suche	Ein Nutzer befindet sich einem Geschäft und nutzt die Suche als Entscheidungshilfe beim Produktkauf.

Tabelle 4: Micro-Moments (Google, 2015)

Expertenbasierte Evaluation

Heuristische Evaluation

Bei einer heuristischen Evaluation beurteilen Experten die Gebrauchstauglichkeit des Testgegenstandes anhand von prüfbareren Usability-Prinzipien (sog. Heuristiken), die für einen bestimmten Anwendungsfall definiert sind (Nielsen, 1994).

Mangels etablierter Heuristiken für die sprachgesteuerte mobile Websuche wurde im Rahmen des Projektes eine eigene Heuristik entwickelt (s. Anlage 1). Basis dafür bildeten Harms et al., 2002; usability.gov, 2006; Sünkler, 2009 sowie Shitkova et al., 2015, die um eigene such- und/oder mobilspezifische Überlegungen ergänzt wurden. Die Bewertung der formulierten Prüfpunkte erfolgte binär mittels ja/nein und wurde durch die Studierenden vorgenommen, die bei der Projektdurchführung die Rollen der Usability-Experten einnahmen.

Anhand der Faktoren Problemhäufigkeit (In welcher Frequenz tritt das Problem auf?), Problemeinfluss (Inwieweit wird die Bewältigung einer Aufgabe durch das Problem beeinträchtigt?), Persistenz (Tritt das Problem zufällig oder regelmäßig auf?) und Markteinfluss (Wie könnte sich das Problem auf das Ansehen beim Nutzer auswirken?) wurden die identifizierten Usability-Probleme bewertet und entsprechend einer der folgenden Problemstufen zugeordnet:

Stufe	Bezeichnung	Priorität
-	Kein Problem	-
1	Kosmetisches Problem	Braucht nicht behoben zu werden, außer der Projektrahmen stellt genügend Zeit dafür zur Verfügung.
2	Geringfügiges Usability-Problem	Der Behebung sollte untergeordnete Priorität eingeräumt werden.
3	Bedeutendes Usability-Problem	Es ist wichtig, das Problem zu beheben.
4	Usability-Katastrophe	Es ist ein Muss, dieses Problem zu beheben, bevor das Produkt ausgeliefert wird.

Tabelle 5: Gewichtung von Usabilityproblemen (nach Nielsen, 1994)

In die Auswertung flossen die Ergebnisse ein, die die Evaluatoren als bedeutendes Usability-Problem (Stufe 3) oder als Usability-Katastrophe (Stufe 4) beurteilt hatten.

Szenarienbasierter Walkthrough

Walkthrough-Methoden gehören zu den expertenbasierten analytischen Usability-Inspektionsmethoden. Dabei werden im Gegensatz zur breitangelegten heuristischen Evaluation Personas und Szenarien erstellt, die dabei unterstützen, Usability-Probleme zu identifizieren. Experten versetzen sich in die Rolle der Persona und prüfen Systeme anhand der Szenarien (Sarodnick & Brau, 2006).

In dem Projekt wurden insgesamt vier Personas definiert, die verschiedene Archetypen typischer Suchmaschinennutzer mit verschiedenen kognitiven Fähigkeiten und Suchmaschinenerfahrungen repräsentieren (s. Anlage 2). Grundlage für die Szenarien bildeten die oben dargestellten Micro-Moments, um die Nutzerbedürfnisse speziell bei der mobilen Suche stärker berücksichtigen zu können (Szenarien s. Anlage 3). Die Durchführung der Walkthroughs wurde durch die Studierenden realisiert, die sich an den Aspekten aus der Heuristik orientierten und in den Ergebnissen sowohl identifizierte positive als auch negative Auffälligkeiten aufnahmen.

Retrievaltest

Eine gebrauchstaugliche Suchmaschine definiert sich neben der Qualität ihrer Gestaltung gleichermaßen auch durch die Qualität der ausgelieferten Ergebnisse. Ein System mit gebrauchstauglicher Oberfläche ist für den Nutzer dann erst in vollem Umfang benutzbar, wenn relevante Ergebnisse zu den Suchanfragen ausgeliefert werden (Lewandowski, 2011).

Retrievaltests dienen der Messung der Effektivität von Information-Retrieval-Systemen, beantworten also die Frage, wie gut die Ergebnisse zu einer bestimmten Suchanfrage passen. Als Standardkriterium für die Bewertung wird die Relevanz herangezogen (Petras, 2013).

Für eine erste Einschätzung der Ergebnisqualität der Voice-Web-Search-Angebote wurden in dem Projekt insgesamt 20 Suchanfragen, aufgeteilt nach den Micro-Moments (s. Tabelle 4) von zwei Studierenden anhand ihrer Relevanz bewertet. Die Bewertung erfolgte dabei binär (ja / nein) und auf einer aufsteigenden 5er-Skala.

Nutzerbasierte Evaluation

Die nutzerbasierte Evaluation fand in der ersten Hälfte des Dezembers 2015 statt; teilgenommen haben insgesamt 25 Probanden im Alter zwischen 8 und 40 Jahren. Die Stichprobe war als Convenience Sample konzipiert und damit nicht repräsentativ. Die Rekrutierung erfolgte im Umfeld der Projektteilnehmer_innen sowie über eine Ausschreibung auf einer studentischen Stellenplattform.

Alle ausgewählten Testteilnehmer_innen verfügten über Erfahrung im Umgang mit Smartphones, die wenigsten von ihnen jedoch brachten Vorerfahrung in der Nutzung der Sprachsteuerung und sprachgesteuerten Websuche mit.

Die Evaluation erfolgte unter realen Bedingungen im Umfeld der Hochschule, nicht im Labor. Die Testgeräte wurden den Probanden zur Verfügung gestellt. Der Nutzertest wurde umrahmt durch eine Vorabbefragung im Vorfeld und eine Nachbefragung im Anschluss.

Vorabbefragung

Mittels Vorabbefragung wurden die demographischen Daten, die Erfahrungen und Nutzungsgewohnheiten in Bezug auf Smartphones und Sprachsteuerung abgefragt sowie die Erwartungen der Probanden an Test und System ermittelt (Fragebogen s. Anlage 4).

Nutzertest

Der eigentliche Nutzertest wurde durch eine Vorstellung der Ziele und des Ablaufs der Studie und eine kurze Einführung in die Funktionsweise des Testgerätes eingeleitet. Anschließend bearbeitete jeder Proband zunächst eine Pretask, um sich mit dem Gerät vertraut zu machen. Eine Auswertung dieser Aufgabe erfolgte bewusst nicht.

Das Kernstück des Nutzertests bildeten vier Nutzungsszenarien, die jeweils einen Micro-Moment abbildeten (s. Tabelle 4) und jeweils aus mehreren Teilsuchschritten bestanden. Den Testabschluss bildete eine freie Suche, bei der jeder Nutzer eine beliebige Suche nach seinem persönlichen Interesse durchführen konnte (Testleitfaden s. Anlage 5).

Die Datenerhebung im Nutzertest erfolgte mittels teilnehmender Beobachtung und der Methode des Retrospective Thinking Aloud (dem Lauten Denken in der Rückschau an eine Aufgabe), um die Sprachsteuerung nicht versehentlich auszulösen oder durch paralleles Sprechen zu irritieren. Die Beobachtungen und Äußerungen wurden handschriftlich durch einen Protokollanten sowie durch Audioaufnahmen dokumentiert.

Nachbefragung und System Usability Scale (SUS)

In der abschließenden Nachbefragung wurden die Eindrücke und Gedanken zu dem getesteten System zum einen in freier Form (Fragebogen s. Anlage 6), zum anderen in standardisierter Form mittels SUS-Fragebogen abgefragt. Der Fragebogen (System Usability Scale, SUS) besteht aus insgesamt zehn Aussagen über die Usability eines Systems, wobei jeweils fünf Aussagen positiv und fünf negativ formuliert sind. Diese werden abwechselnd gestellt, um die Probanden auf diese Weise dazu anzuhalten, aktiv zu lesen und bewusste Antworten zu geben (Quirnbach 2013, S. 104). Für die Beantwortung der Items steht eine Bewertungsskala mit fünf Antwortmöglichkeiten zur Verfügung. Das Ergebnis ist dann der SUS-Score, eine Zahl zwischen 0 und 100, wobei 0 das schlechtmöglichste Ergebnis und damit die geringste Usability darstellt.

Für die Berechnung des SUS-Scores werden die Zahlenwerte der Bewertungsskala summiert und anschließend mit 2,5 multipliziert. Der ermittelte Wert kann nun als Prozentwert interpretiert werden:

- 100% entsprechen einem perfekten System ohne Usability-Probleme.
- Werte über 80% deuten auf eine gute bis exzellente Usability hin.
- Werte zwischen 60% und 80% sind als grenzwertig bis gut zu interpretieren.
- Werte unter 60% sind Hinweise auf erhebliche Usability-Probleme.

Ergebnisse

Das folgende Kapitel stellt die gewonnenen Ergebnisse dar und beschreibt – gegliedert nach Systemen und Methoden – die identifizierten Stärken und Schwächen. Den ersten Abschnitt bilden die Ergebnisse der expertenbasierten Methoden, den zweiten dann diejenigen aus dem Nutzertest.

Übersicht: Auswertung Heuristische Evaluation und Szenarienbasierter Walkthrough

In der nachfolgenden Darstellung werden die 'Pain Points' der heuristischen Evaluation aufgeführt, also diejenigen Usability-Probleme, die von den Evaluatoren als 'bedeutend' oder 'katastrophal' bewertet worden sind. Die Testergebnisse werden wie folgt eingestuft:

	bedeutendes Usability-Problem	Es ist wichtig, dieses Problem zu beheben, und sollte dadurch eine hohe Priorität erhalten.
	Usability-Katastrophe	Es ist ein Muss, dieses Problem zu beheben, bevor das Produkt ausgeliefert wird.

Zugang / Einstieg

1.1 Wartet das System mit dem Beginn der Suche, bis der Nutzer zu Ende gesprochen hat?		
System	Erläuterung Prüfkriterium	Gewichtung
Amazon Fire OS	ggf. notieren, wie lange die Pausen zwischen den gesprochenen Wörtern maximal sein können, bevor die Suche automatisch beginnt.	

Fazit zum Zugang / Einstieg: Die sprachgesteuerte Websuche innerhalb des Amazon Fire OS startet den Suchvorgang bereits dann, wenn die Sprechpausen zwischen den einzelnen Suchbegriffen lediglich eine Sekunde (oder mehr) beträgt. Das System wartet demnach nicht, bis der Nutzer zu Ende gesprochen hat.

Bedienung

2.1 Werden Sprachbefehle bezogen auf die Ergebnisliste angenommen?		
System	Erläuterung Prüfkriterium	Gewichtung
Amazon Fire OS	z. B. "Ok Google, öffne den ersten Treffer" o.Ä.	
Microsoft Cortana		
Google Now		

2.2 Kann man die akustische Rückmeldung akustisch beenden?		
System	Erläuterung Prüfkriterium	Gewichtung
Google Now	/	

Fazit zur Bedienung: Die sprachgesteuerte Websuche innerhalb des Amazon Fire OS sowie Google Now setzen keine Sprachbefehle um, die sich auf die Ergebnisliste beziehen. Bei Microsoft Cortana lässt sich zumindest teilweise der erste Treffer durch eine Spracheingabe öffnen, was jedoch nur für einige erstplatzierte und nicht für weitere Treffer gilt. Negativ aufgefallen ist des Weiteren bei Google Now, dass sich die akustische Rückmeldung nicht akustisch beenden lässt.

Suchanfrageverarbeitung

3.1 Werden (Rechtschreibfehler) Sprachfehler automatisch korrigiert?		
System	Erläuterung Prüfkriterium	Gewichtung
Apple Siri	/	

3.2 Wird die Suchanfrage bei lauten Hintergrundgeräuschen korrekt erkannt?		
System	Erläuterung Prüfkriterium	Gewichtung
Amazon Fire OS	/	
Microsoft Cortana	/	

3.3 Werden Dialekte korrekt verarbeitet?		
System	Erläuterung Prüfkriterium	Gewichtung
Microsoft Cortana	/	

3.4 Werden Sprachfehler korrekt verarbeitet?		
System	Erläuterung Prüfkriterium	Gewichtung
Amazon Fire OS	/	

3.5 Werden Operatoren korrekt verarbeitet?		
System	Erläuterung Prüfkriterium	Gewichtung
Google Now	/	

3.6 Werden Abkürzungen und Akronyme in der Suchanfrage aufgelöst?		
System	Erläuterung Prüfkriterium	Gewichtung
Amazon Fire OS	/	

Fazit zur Suchanfrageverarbeitung: Bezüglich der Verarbeitung der Suchanfragen weisen alle vier Systeme Schwächen auf. So hat Apple Siri Schwierigkeiten, sprachliche Fehler zu korrigieren, die sprachgesteuerte Websuche innerhalb des Amazon Fire OS sowie Microsoft Cortana werden durch Störfaktoren wie laute Hintergrundgeräusche beeinträchtigt. Für das System von Amazon stellen auch Lispeln, Abkürzungen sowie Akronyme Probleme dar, Microsoft Cortana erkennt hingegen französischsprachige Suchanfragen nur bruchstückhaft. Außerdem werden Boolesche Operatoren wie „UND“,

„ODER“ sowie „NICHT“ von Google Now nicht als solche erkannt und in der Suchanfrage verarbeitet.

Personalisierung

4.1 Wird die Sprache der Ausgabe der Sprache der Eingabe angepasst?		
System	Erläuterung Prüfkriterium	Gewichtung
Amazon Fire OS	Erhält der Nutzer eine englischsprachige Antwort, wenn er englischsprachig sucht?	
Apple Siri		
Google Now		

4.2 Unterstützt das System auch englischsprachige Suchanfragen?		
System	Erläuterung Prüfkriterium	Gewichtung
Amazon Fire OS	/	

4.3 Lassen sich Suchfilter durch Sprachsteuerung auswählen?		
System	Erläuterung Prüfkriterium	Gewichtung
Apple Siri	/	
Microsoft Cortana	Versuch nur Videos oder nur Bilder auszugeben.	

Fazit zur Personalisierung: Bis auf Microsoft Cortana setzen alle übrigen Systeme englischsprachige Suchanfragen nur unzureichend um. Das Amazon-System führt bei der Frage „Where do I find the Stadtpark?“ eine Bing-Suche anstelle einer Routenberechnung durch und stellt bei „Football clubs in Germany“ Hamburger Nachtclubs dar, Google Now liefert deutsche Sprachausgaben auf englische Fragestellungen. Eine Schwäche des Microsoft-Systems ist jedoch zu nennen: wie auch Apple Siri ist es nicht in der Lage, Filter innerhalb des Suchvorgangs durch Spracheingabe auszuwählen.

Barrierefreiheit

5.1 Verfügt das System über zusätzliche Funktionen der Barrierefreiheit?		
System	Erläuterung Prüfkriterium	Gewichtung
Amazon Fire OS	1. Kriterium: akustische Ausgabe der Suchergebnisse 2. Kriterium: Schriftvergrößerung	
Microsoft Cortana	/	

Fazit zur Barrierefreiheit: Bezüglich der Barrierefreiheit sind insbesondere die Systeme von Amazon und Microsoft zu nennen. In beiden Sprachsteuerungen ist eine Schriftvergrößerung nicht möglich. Eine akustische Ausgabe der Suchergebnisse erfolgt beim Amazon-System überhaupt nicht, bei Microsoft Cortana lediglich teilweise.

Ergebnispräsentation

6.1 Erfolgt die Antwort in akustischer Form?		
System	Erläuterung Prüfkriterium	Gewichtung
Microsoft Cortana	/	

Fazit zur Ergebnispräsentation: Die beim Abschnitt der Barrierefreiheit genannte Schwäche von Microsoft Cortana, die Suchergebnisse nicht immer auch akustisch auszugeben, ist gleichzeitig als Mangel der Ergebnispräsentation zu nennen.

Hilfestellung

7.1 Werden alternative Suchvorschläge unterbreitet („Meinten Sie ...“ oder verwandte Suchbegriffe) ?		
System	Erläuterung Prüfkriterium	Gewichtung
Apple Siri	/	
Microsoft Cortana	/	

7.2 Ist die Rückmeldung bei Nicht-Verstehen durch das System hilfreich ?		
System	Erläuterung Prüfkriterium	Gewichtung
Apple Siri	/	

7.3 Wird während der Suche Hilfe bspw. Durch einen Chat im Interface angeboten ?		
System	Erläuterung Prüfkriterium	Gewichtung
Google Now	/	

Fazit zur Hilfestellung: Bis auf die sprachgesteuerte Websuche des Amazon Fire OS offenbaren die übrigen drei Systeme gravierende Mängel bezüglich der Hilfestellungen. So unterbreiten weder Microsoft Cortana noch Apple Siri alternative Suchvorschläge, wie z.B. „Meinten Sie...?“, noch ist bei letzterem System die Rückmeldung verständlich, wenn die Spracheingabe nicht verstanden wurde. Des Weiteren wurde bei Google Now die fehlende Hilfe während des Suchvorgangs, z.B. durch einen Chat, als bedeutendes Usability-Problem erachtet.

Fazit Heuristik

Das nachfolgende Diagramm stellt die Ergebnisse aus der heuristischen Evaluation sämtlicher Fragestellungen auf allen getesteten Systemen in der Übersicht dar. Diese befinden sich in einer Reihenfolge. Google Now hat demnach vor Apple Siri, Microsoft Cortana und der sprachgesteuerten Websuche des Amazon Fire OS am stärksten abgeschnitten und bei 39 von insgesamt 46 Fragen die 'Gewichtung 0: kein Problem' erhalten (s. Abbildung 2).

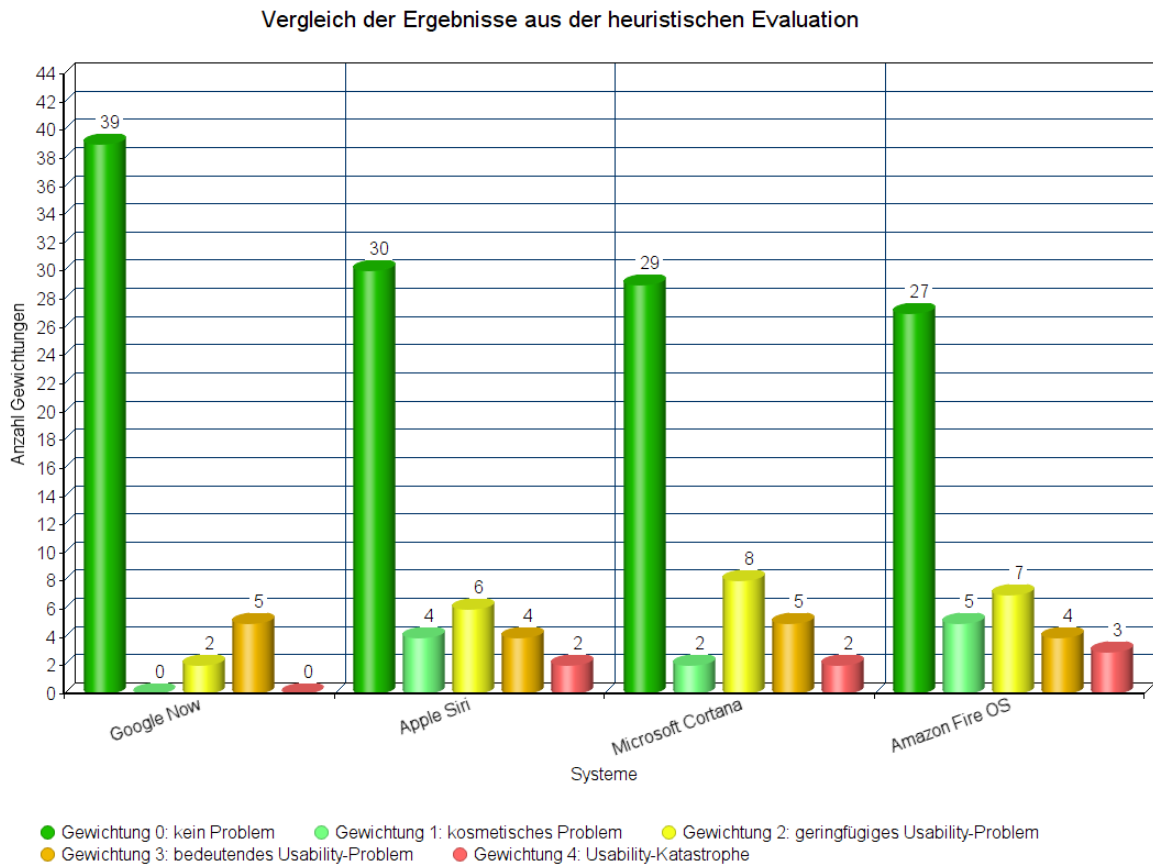


Abbildung 2: Vergleich der Ergebnisse aus der heuristischen Evaluation (eigene Darstellung, 2015)

Google Now: Auswertung

Heuristische Evaluation und Szenarienbasierter Walkthrough



Zugang/Einstieg

Positiv sollte man erwähnen, dass der Start reibungslos verlief und das Aufrufen der Suche mit dem Sprach-Shortcut „OK Google“ sehr einfach ist. Die Suche kann sowohl per Sprach-Shortcut als auch mit Hilfe eines Buttons begonnen werden. Zudem kennzeichnet ein akustisches Signal zu Beginn und zum Ende der Spracheingabe den jeweiligen Schritt der Suchanfragenstellung in dem sich der Nutzer gerade befindet.



Bedienung

Zunächst wirkt die allgemeine Bedienung der Google Voice Search sehr simpel, auch durch die sehr einfach gehaltene Gestaltung des Interface. Die Ergebnisse werden schnell und übersichtlich präsentiert, allerdings wurde bemängelt, dass die einzelnen Suchergebnisse nicht verbal geöffnet werden können, sondern per Hand ausgewählt werden müssen. Allgemein werden viele „Antworten“ noch nicht direkt als Ergebnis dargestellt (wie bspw. im Knowledge Graph), sondern befinden sich in Dokumenten, die erst von Hand angeklickt werden müssen.



Suchanfrageverarbeitung

Bis auf einige seltene Ausnahmen werden i.d.R. alle Suchanfragen richtig erkannt und verarbeitet. Eigenwörter oder Namen werden größtenteils direkt erkannt (falls dies nicht der Fall ist, ist das System aber lernfähig und erkennt diese beim zweiten oder dritten Mal als solche). Das System verarbeitet meist auch mehr als eine Frage und bezieht sich in Folgeanfragen auch auf zuvor gesuchte Begriffe und Suchwörter.



Personalisierung

Ist der Standort aktiviert, wird dieser verarbeitet und bspw. für Wegbeschreibungen genutzt. Auch für Want-to-know-Anfragen werden Ergebnisse auf die jeweilige Stadt, in der man sich befindet, zugeschnitten. Begriffe wie „Zuhause“ oder „Arbeitsplatz“ können definiert und später so von alleine erkannt und zugeordnet werden. Zudem werden Orte, die zuvor vom jeweiligen Google-Nutzer gesucht wurden, gespeichert und in darauffolgenden Suchanfragen vorgeschlagen bzw. in der Ergebnispräsentation (z.B. eine Landkarte) mit eingezeichnet.



Barrierefreiheit

Das System ist generell in der Lage Akzente, Dialekte oder Sprachfehler zu verstehen und korrekt zu verarbeiten. Durch die teilweise akustisch ausgegebenen Suchergebnisse müssen diese nicht erst gelesen werden. Allerdings ist das generelle Vorlesen der Suchergebnisse nicht möglich. Auch das Zoomen auf die Ergebnisse ist auf den ersten Blick nicht möglich. Schrifteinstellungen können allerdings im Menü über die „Einstellungen“ modifiziert werden.



Hilfestellung

Eine Hilfestellung ist lediglich durch das Erscheinen des Schriftzuges „Jetzt Sprechen“ gegeben. Allgemeine Hilfestellungen werden jedoch ebenfalls nur unter dem Menü unter „Hilfe“ bereitgestellt. Beim intuitiven Nutzen des Systems tauchten i.d.R. aber keine Probleme auf, da das System größtenteils selbsterklärend ist.



Ergebnispräsentation

Szenario 1: Want-to-buy: Bei den Tests fiel positiv auf, dass zu den gestellten Suchanfragen die passenden Ergebnisse ausgegeben wurden. Zusätzlich sollte man hervorheben, dass die verschiedenen Spracheingaben richtig erkannt wurden. Die Suchergebnisse wurden durch Shopping-Ergebnisse und Preisvergleichsseiten ordentlich und übersichtlich präsentiert.

Szenario 2: Want-to-go: Bei Suchanfragen nach einem bestimmten Ziel, erhält man eine genaue akustische Wegbeschreibung, die in wenigen Fällen wegen schlechter Betonung schwer zu verstehen ist. Ein positiver Aspekt ist die Wahl des Verkehrsmittels auf dessen Basis die Route berechnet werden soll. Ebenso ist es möglich eine Route mit den öffentlichen Verkehrsmitteln ausgeben zu lassen.

Bei ungenauen Anfragen, bspw. „Sehenswürdigkeiten in Hamburg“, erhält man eine Kartenansicht die, wie hier angefragt, Sehenswürdigkeiten im Raum Hamburg markiert. Zur besseren Übersicht wird ergänzend eine Liste mit Bildern der jeweiligen Sehenswürdigkeit angezeigt. Wählt man eine Sehenswürdigkeit aus, wird ein Knowledge Graph angezeigt, der die wichtigsten Informationen enthält, bspw. eine kurze Beschreibung, die Öffnungszeiten und/oder die Option zur Routenberechnung.

Szenario 3: Want-to-know: Bei Suchanfragen nach bekannten Persönlichkeiten oder Orten erhält man neben einer Sprachausgabe mit allen wichtigen Eckdaten einen Knowledge Graph, der alle wichtigen Informationen anzeigt. Google nutzt als Quelle Wikipedia.

In der Regel wird eine Auflistung aller relevanten Treffer angezeigt, in wenigen Fällen kommt es vor, dass die Ergebnisseite aus irrelevanten Treffern besteht.

Szenario 4: Want-to-do: Das Ende des Suchdurchlaufes wird via Sprachausgabe gekennzeichnet, z.B. „Hier sind die Einträge für ...“.

Wie in Szenario 2 ist es möglich eine Kartenansicht zu erhalten. Diese zeigt die Orte an wo man die gesuchte Aktivität besuchen kann. Wichtige Informationen oder die Routenberechnung zum gewünschten Ort können über den Knowledge Graph ausgewählt werden.

Apple Siri: Auswertung

Heuristische Evaluation und Szenarienbasierter Walkthrough



Zugang/Einstieg

Hier können wir positiv erwähnen, dass der Start ohne Probleme verlief und das Aufrufen der Suche mit dem Sprach-Shortcut „Hey Siri“ sehr gut funktioniert hat. Der Sprach-Shortcut funktioniert jedoch nur, wenn das Gerät an einer Stromquelle angeschlossen ist. Die Suche kann sowohl per Sprach-Shortcut als auch mit Hilfe eines Buttons begonnen werden. Zudem ertönt ein akustisches Signal zu Beginn und zum Ende der Spracheingabe.



Bedienung

Die allgemeine Bedienung von Siri ist sehr simpel, auch durch die sehr einfache Gestaltung des Interface. Die Ergebnisse werden schnell und übersichtlich präsentiert, allerdings wurde bemängelt, dass die einzelnen Suchergebnisse nicht verbal geöffnet werden können, sondern per Hand ausgewählt werden müssen. Allgemein werden viele „Antworten“ noch nicht direkt als Ergebnis dargestellt, sondern befinden sich in Dokumentenliste, die erst von Hand angeklickt werden müssen (sehr oft Wikipedia). Einmal öffneten sich die Siri-Einstellungen, als eine Suche begonnen werden sollte.



Suchanfrage/Verarbeitung

Meistens werden alle Suchanfragen richtig erkannt und verarbeitet. Akzente oder Dialekte werden meistens erkannt. Zusätzlich personalisiert Siri sich im Laufe der Zeit und lernt dazu. Das System verarbeitet meist auch mehr als eine Frage und bezieht sich in Folgeanfragen auch auf zuvor gesuchte Begriffe und Suchwörter.



Personalisierung

Der Ortungsdienst muss aktiviert sein, damit Siri die Ergebnisse entsprechend anpassen und verarbeiten kann. Siri gibt ein Feedback falls der Ortungsdienst nicht aktiviert ist und leitet den Nutzer auf die entsprechende Einstellung weiter. Die Ergebnisse werden in Form von Karten mit Wegbeschreibung, Entfernung etc. dargestellt. Siri kann nach einer Zeit und bei einer richtigen Einstellung den Nutzer mit einem gewünschten Namen ansprechen. Informationen werden gespeichert und eventuell für neue Suchen verwendet.



Barrierefreiheit

Siri ist meistens in der Lage Akzente, Dialekte oder Sprachfehler zu verstehen und korrekt zu verarbeiten. Allerdings ist das generelle Vorlesen der Suchergebnisse nicht möglich. Auch das Zoomen auf die Ergebnisse ist nicht möglich. Schrifteinstellungen können allerdings im Menü unter „Einstellungen“ angepasst werden. Siri stellt bei offenen oder nicht verstandenen Suchanfragen eine Gegenfrage, um ein passendes Ergebnis zu liefern.



Hilfestellung

Eine Hilfestellung ist bei Siri in Form von einem Fragezeichen-Symbol vorhanden, ebenso kann diese durch „Hilfe“ in die Sprachsuche aufgerufen werden. Siri öffnet bei zu langen Pausen zusätzlich eine Liste mit einigen Beispielfragen, die der Nutzer verwenden kann.



Ergebnispräsentation

Die Ergebnispräsentation hat bei Siri gut funktioniert. Uns ist dabei aufgefallen, dass die ortsbezogene Suche einwandfrei funktioniert hat. Hierbei werden die Routen aus der Apple-eigenen Karte angezeigt. Hierbei kann der Nutzer selbstständig auswählen, ob er zu Fuß oder mit dem Auto unterwegs ist. Die Verkehrslage lässt sich hierbei nicht anzeigen. Bei den Ergebnissen wird der Standort des Nutzers beachtet.

Bei einfachen Suchanfragen hat Siri neben der reinen webbasierten Ergebnisliste auch akustische Antworten geliefert, indem sie auch die Ergebnisse vorgelesen hat. Siri funktioniert auch gut bei komplexen und verschachtelten Fragen.

Man kann nicht per Sprachkommando auf die Ergebnisliste zugreifen, hier muss eine weitere Eingrenzung per Hand vorgenommen werden. Häufig waren die Treffer nicht immer relevant bzw. vermengt mit unwichtigen Ergebnissen.

Microsoft Cortana: Auswertung

Heuristische Evaluation und Szenarienbasierter Walkthrough



Zugang/Einstieg

Der Zugang zur Anwendung wird weitgehend als einfach und unkompliziert beschrieben. Es ist zumeist problemlos erkennbar wie die Anwendung gestartet und bedient wird.



Bedienung

Die Bedienung ist sehr simpel und das Gerät ist „mit einem Wisch“ bedienbar, jedoch ist keine Trefferauswahl per Sprachsteuerung möglich.



Suchanfrageverarbeitung

Bei einer sprachlich fehlerhaften Suchanfrage, z. B. „Wo isch die näschte Go-Cart-Bahn in Hamburg?“ oder „Wo kann i a billigs Smartphone kaufa?“ werden trotzdem passende Treffer angezeigt. Dabei werden die Suchanfragen synchron mitgeschrieben.

So versteht Cortana Anfragen nach Wegbeschreibungen sehr gut, kann sie jedoch nicht darstellen (z. B. bei „Fischmarkt, Verkehrsmittel“). Bei mehreren, verschiedenen Suchanfragen war es Cortana vereinzelt doch möglich, Routen für öffentliche Verkehrsmittel direkt zu berechnen.

Bei einer Anfrage wie „Bücher zu Napoleon“ oder „Bücher zu Napoleon in Bibliothek“ liefert Cortana jedoch keine relevanten Ergebnisse. Bei einer genaueren und spezifischeren Anfrage verweist Cortana auf die Hamburger Bücherhallen und deren Ausleihbedingungen, jedoch nicht direkt auf Bücher zu Napoleon. So muss jede Anfrage sehr genau und spezifisch formuliert werden, um von Cortana verstanden werden zu können.

Bei einer verschachtelten Anfrage wird „Hier“ oder „Dort“ nicht als der in der ersten Anfrage erwähnte Ort (z. B. Hamburg) verstanden. Z. B.: „Welche Touristen-Attraktionen gibt es in Hamburg?“ → „Gibt es dort auch Restaurants?“.



Personalisierung

Eine Personalisierung ist grundsätzlich möglich (z. B. durch die über große Zeiträume angesammelten Suchdaten oder durch das personalisierte Notizbuch, das auf persönliche Interessen und Bedürfnisse zugeschnittene Internetsuchen möglich macht). Jedoch ist es für den durchschnittlichen Einsteiger nicht möglich, innerhalb kurzer Zeit, einzusehen, ob und inwieweit die Personalisierung stattfindet.



Barrierefreiheit

Negativ anzumerken ist, dass keine genaue Wegbeschreibung per Spracheingabe möglich war. Microsoft Cortana zeigt hierbei eindeutige Probleme auf. Wörter, die nicht korrekt ausgesprochen werden (z. B. Buchstaben weggelassen, Umgangssprache), kann Cortana je nach Situation und Grad der Schwierigkeit unterschiedlich deuten und verstehen. Sprachlicher Dialekt wurde ohne Probleme ins Hochdeutsche übersetzt. Darüber hinaus kann das Zoomen von Schrift und Grafik deutlich positiv hervorgehoben werden.



Hilfestellung

Die Hilfestellung des Sprachassistenten bietet ebenfalls abhängig von Ziel und Erwartung von gar keiner Hilfe über ähnliche Anfragen bis hin zu konkreten Vorschlägen, welche Vokabeln verwendet werden sollten, um die Spracheingabe für Cortana verständlich zu machen. Dabei spielt die Komplexität der Anfrage eine große Rolle. Für einfache Anfragen wie die Kaufabsicht eines konkreten Smartphones werden auch ähnliche oder alternative Suchanfragen ausgegeben. Bei komplexen Anfragen fiel auf, dass für falsch verstandene Anfragen oder aber falschen Ergebnissen keine Hilfestellung von Cortana vorgeschlagen worden ist. Dennoch wollen wir die uns besonders positiv aufgefallene Cortana Hilfestellung „Sag...“ hervorheben, welche als Vorschlag ausgegeben wird, um die Spracheingabe verständlicher zu machen.



Ergebnispräsentation

Positiv anzumerken bei der Ergebnispräsentation ist die ortsbezogene Suche. Hierbei werden die aus der Bing-eigenen Kartenfunktion generierten Ortsergebnisse in den Suchergebnissen dem Nutzer präsentiert. In dem Kontext wird weiterhin auch die Kartennavigation inklusive Routenberechnung für PKWs bei entsprechender Anfrage korrekt aufgerufen. Auch der Verkehr

lässt sich in der Situation per Sprachbefehl problemlos anzeigen. In der Regel wird bei den Ergebnissen der Standort des Nutzers beachtet.

Bei simplen Anfragen hat Cortana neben der reinen webbasierten Ergebnisliste auch verbal ausgesprochene Antworten geliefert, indem die Ergebnisse vorgelesen wurden, auch wenn dies nur selten der Fall war. Komplexe bzw. verschachtelte Fragen wurden unzureichend bis gar nicht bearbeitet oder lieferten kein Ergebnis.

Neben den reinen Web-Ergebnissen werden dem Nutzer auch universelle Auswahlmöglichkeiten ausgegeben, so wie bei der Suche nach „Bücherhallen Hamburg“, wo eine Karte mit Standorten der Hamburger Bücherhallen angezeigt worden ist.

Je nach Situation und Erwartung ist Cortana in der Lage bei der Frage nach den Öffnungszeiten, diese in den Ergebnissen auszugeben, ohne dass der Nutzer eine weitere Interaktion ausführen muss, um an das Ziel zu kommen. Verbindungen mit den öffentlichen Verkehrsmitteln wurden ebenfalls situationsbedingt entweder angezeigt oder nicht. Dabei ist uns aufgefallen, dass Anfragen in dem Zusammenhang sehr spezifisch sein müssen, damit Cortana diese überhaupt versteht. Auf die Frage „Wie komm ich mit öffentlichen Verkehrsmitteln zum Hamburger Fischmarkt?“ wurden bspw. nur Routen mit dem PKW angeboten.

Eine Auswahl per Sprache aus der Ergebnisliste ist mit dem Sprachassistenten nicht möglich. Hier muss eine weitere Eingrenzung manuell vorgenommen werden. Die Tests haben auch ergeben, dass Kaufabsichten von Cortana nur mangelhaft verstanden wurden. Zum einen waren die Treffer nicht immer relevant bzw. vermengt mit unwichtigen Ergebnissen wie News-Artikeln, und zum anderen begannen die ersten relevanten Ergebnisse unter „ähnliche Anfragen“.

Amazon Fire OS: Auswertung

Heuristische Evaluation und Szenarienbasierter Walkthrough



Zugang/Einstieg

Der Zugang wird hier als neutral bis negativ bewertet. Man ist sich nicht sofort im Klaren darüber, welches Vokabular nun verwendet werden soll, um eine gezielte Websuche durchzuführen, da vorerst keine Hilfestellungen gegeben werden. Ansonsten funktioniert es simpel über einen langen Knopfdruck.



Bedienung

Die Bedienung selbst ist im Allgemeinen gut und simpel. Allerdings wird bemängelt, dass die Sprachsteuerung sich nicht per Sprache starten lässt. Wenn eine Websuche durchgeführt wird, kann man auch keine Folgeaktionen per Sprache befehlen.



Suchanfrageverarbeitung

Die Suchanfrageverarbeitung weist auch einige Mängel auf. So werden einzeln gesprochene Wörter einfach ausgelassen und nicht mit in die Suchanfrage übernommen. Bei Lautstärke werden Hintergrundgeräusche mit verarbeitet. Man muss zudem relativ schnell sprechen, da schon nach sehr kurzer Zeit die Anfrage abbricht und gesucht wird. Außerdem wäre es wünschenswert, wenn die Suchanfrage während des Sprechens mitgeschrieben wird. Beim Amazon Fire Phone allerdings wird die Suchanfrage am Ende für einen Sekundenbruchteil angezeigt und direkt verarbeitet.



Personalisierung

Eine Personalisierung in irgendeiner Form ist nicht möglich.



Barrierefreiheit

Bei Sprachfehlern reagiert das Amazon Fire Phone souverän. So wurde ein moderater schwäbischer Dialekt erkannt und ins Hochdeutsche umgesetzt. Auch mit Lispeln konnte die sprachgesteuerte Suche gut umgehen. Allerdings muss man laut sprechen, mit Flüstern bspw. bekommt man kaum positive Ergebnisse. Ob man das nun als barrierefrei bezeichnen kann, sei dahingestellt, allerdings ist der Eindruck in dieser Kategorie recht positiv.



Hilfestellung

Hilfestellungen sind eingeschränkt verfügbar. So kann man mit dem Fire Phone bspw. durch sein Smartphone navigieren und SMS und/oder Kontakte öffnen. Zu diesen Funktionen steht nach einer kurzen Wartezeit ein Vokabular als Vorlage zu Verfügung. Auch mit einer Wischbewegung von links nach rechts kann man das Vokabular öffnen. Allerdings hat das nichts mit der sprachgesteuerten Websuche zu tun, Beispiele oder Hilfestellungen hierzu sind nicht verfügbar.



Ergebnispräsentation

Die Ergebnispräsentation ist von der Suchmaschine Bing abhängig. So wird jede Suchanfrage, egal ob sie erkannt wird oder nicht, in die Suchmaschine übertragen und dort gesucht. Deshalb ist die Ergebnispräsentation keine Eigenschaft des Fire Phones. Die Bing-Ergebnispräsentation ist suchmaschinentypisch aufgebaut und mit Bildern und Videos angereichert. Die Infosnippets sind als gut zu bewerten. Die gefundenen Wörter werden fett hervorgehoben. Einzig eine Wegbeschreibung öffnet die Navigationsapp, die im Allgemeinen als positiv bewertet wurde, allerdings unter der Voraussetzung, dass das Ziel auch richtig gefunden wurde. Leider war der Navigationsdienst nie verfügbar, sodass keine geführte Wegbeschreibung gestartet werden konnte.

Die Sprachausgabe des Fire Phones ist als negativ zu bewerten. Bis auf „Kommt sofort“ oder „Das kann ich noch nicht“ gab es keine Antworten zu den Fragen. An einen Dialog ist gar nicht zu denken.

Auswertung Retrievaltest

Der Retrievaltest in dem Projekt wurde durchgeführt, um die Fähigkeiten der Systeme in Bezug auf die Qualität der Suchergebnisse zu evaluieren. Bei der Durchführung ergaben sich durch die Technik und Ergebnispräsentation der Anwendungen Einschränkungen, daher konnte die Methode des klassischen Retrievaltests nicht direkt übernommen werden. So waren den Juroren die Systeme bei der Bewertung bekannt. Daneben variierte die Darstellung der Ergebnisse stark zwischen den Systemen: Während bspw. ein System das Ergebnis akustisch wiedergab, wurde dieses bei einem anderen nur in der herkömmlichen Trefferdarstellung der Ergebnisliste angezeigt. Das führte dazu, dass eine typische Evaluierung einzelner Dokumente auf den angezeigten Trefferpositionen nicht durchführbar war. Es können daher auch keine Aussagen über die Retrievaleffektivität auf der Dokumentenebene gemacht werden.

Daher sind die aufgezeigten Ergebnisse des Retrievaltests nur als Tendenzen zu verstehen. Nichtsdestotrotz zeigen diese Tendenzen auf, bei welchen Micro-Moments (s. Tabelle 4) die Systeme die besten Ergebnisse liefern. Im Folgenden werden die zusammengefassten Ergebnisse des Retrievaltests dargestellt. Dabei wird zunächst das Verhältnis der beantworteten Suchanfragen zu der Gesamtanzahl der Suchaufgaben aufgezeigt. Anschließend folgt eine Darstellung der Ergebnisse aus den Skalenbewertungen zu Güte der Relevanz der Suchergebnisse.

Gelöste Suchaufgaben je System

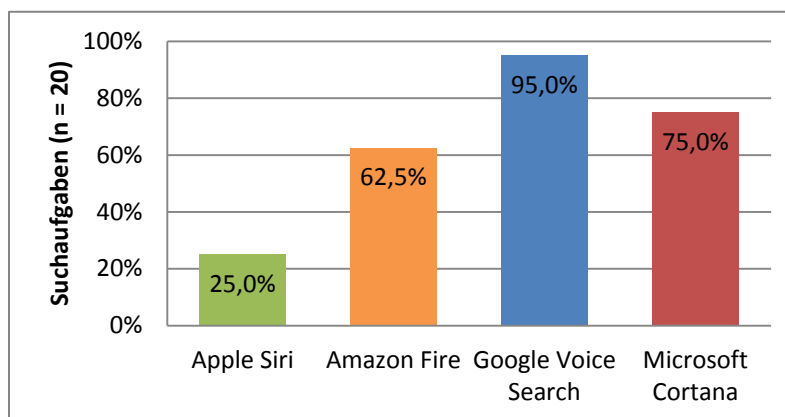


Abbildung 3: Verhältnis der gelösten Suchaufgaben zu der Gesamtanzahl der Suchaufgaben

Für die Beurteilung, ob ein System eine gestellte Aufgabe lösen konnte, bewerteten die Juroren die zurückgegebenen Suchergebnisse zu der Anfrage in binärer Form. Konnte das System ein

brauchbares Ergebnis liefern, wurde dies durch den Juror mit „Ja“ bewertet, wenn nicht, mit „Nein“.

Insgesamt konnte Google Voice Search die Juroren in dieser Fragestellung am meisten überzeugen. Die Zusammenfassung zeigt, dass Google Voice Search bei 95% der Aufgaben brauchbare Ergebnisse liefern konnte. Dahinter liegt Microsoft Cortana mit 75%, während Apple Siri nur bei 25% nutzbare Ergebnisse zeigte. Es zeigte sich in dem Test, dass Apple Siri zum einen bestimmte Spracheingaben als interne Befehle versteht (z. B. starten Anfragen, die mit „zeige“ beginnen, die Galerie auf dem iPhone) und zum anderen auch einige Anfragen nicht versteht und nur mit diesem Hinweis antwortet.

Ergebnisqualität nach System und Micro-Moment

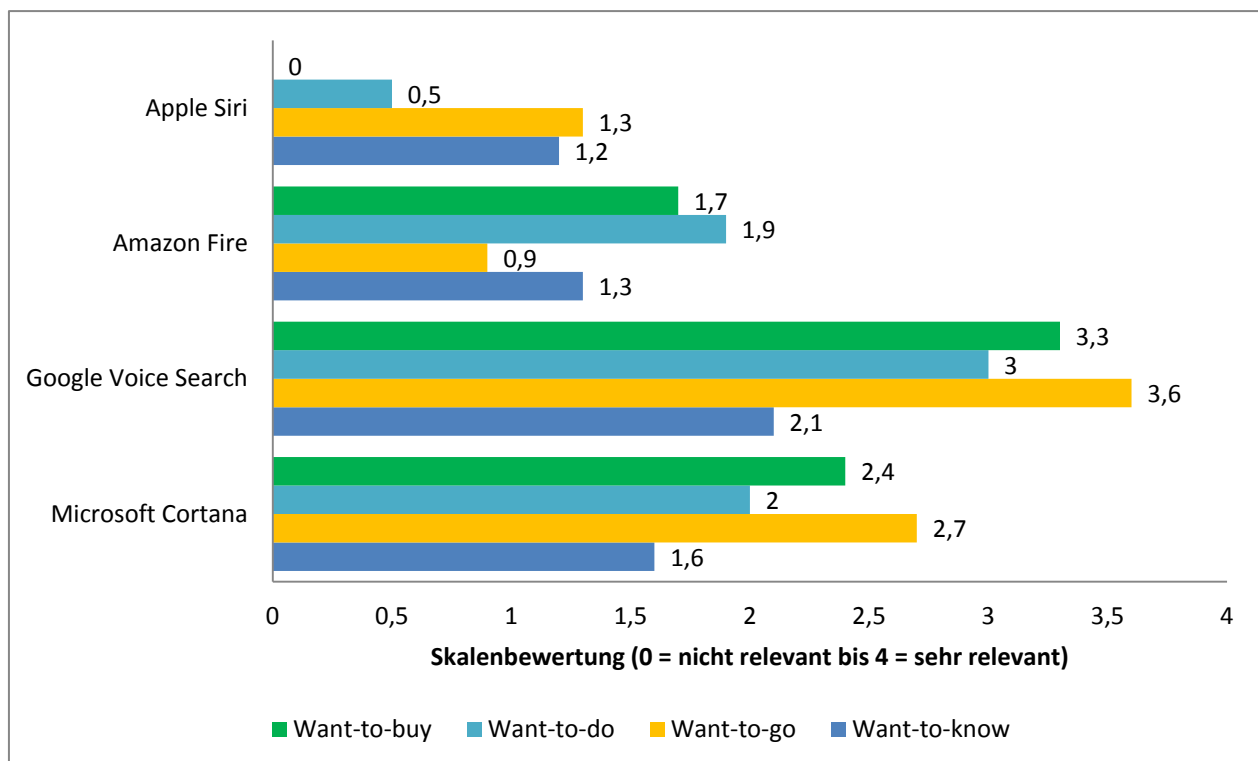


Abbildung 4: Durchschnittliche Skalenbewertung durch 2 Juroren zu den Micro-Moments

Bei der Auswertung mit der 5er-Skala und Aufteilung nach Micro-Moments (je 5 Suchaufgaben pro Micro-Moment) liegt Google Voice Search vor den anderen Systemen im Test. Das Verhältnis zu gelösten Aufgaben und zur Qualität der Ergebnisse spiegelt sich hier wieder. Apple Siri landet auch bei der Qualität auf dem letzten Platz. Mit Hilfe von Apple Siri konnte keine Aufgabe zu dem Micro-Moment Want-to-buy gelöst werden, daher gibt es dazu keine

Skalenbewertung in der Abbildung. Bei den Want-to-go-Aufgaben schneiden Google und Microsoft am besten ab. Beide Anbieter konnten im Test am häufigsten korrekte Routen und Landkarten anzeigen, während Amazon Fire OS dort keine brauchbaren Ergebnisse lieferte. Es fand zwar vereinzelt eine Umleitung auf einen Kartendienst statt, dieser zeigte aber keine nützlichen Informationen.

Fazit Retrievaltest

Die expertenbasierte Evaluation zur Ergebnisqualität zeigte Tendenzen dahingehend, dass Google mit Abstand die besten, d.h. relevantesten Suchergebnisse liefert. Bei der Qualitätsbewertung auf der Skala zeigte sich dabei, dass Google im Bereich der informationsorientierten Anfragen (Want-to-know) die schlechtesten Ergebnisse zeigte, bei den anderen Aufgabentypen aber im Verhältnis eine gleichbleibende Qualität lieferte. Der Verlierer im Retrievaltest ist der Dienst Siri von Apple. Die Ergebnisse konnten nicht überzeugen. Amazon Fire OS und Microsoft Cortana näherten sich in Bezug auf die Beantwortung der Suchaufgaben an. Unterschiede ergaben sich aber bei der Qualitätsbewertung. Dort ist Microsoft Cortana dem Dienst von Amazon bei allen Micro-Moments überlegen.

Google Now: Auswertung Nutzertest

N = 6



Zugang/Einstieg

Positiv sollte man erwähnen, dass die Probanden nach kurzer Einführung mit der Sprachsteuerung vertraut waren. Bei keinem der Probanden waren Startschwierigkeiten zu erkennen, ebenso war keine erneute Erklärung der Handhabung von Nöten.



Bedienung

Zu beobachten war, dass kein Proband mit der Bedienung Probleme hatte. Für die Probanden war mittels Anzeige und akustischen Signale immer klar erkennbar, an welchem Punkt der Suche sie sich befanden. Auffällig war, dass die meisten Probanden nach den ersten Fragen eher den Aktivierungsbutton anstatt den Voice-Shortcut „Ok Google“ nutzten. Bis auf eine junge Probandin, unter 16 Jahren, stellten die Probanden ihre Anfragen mittels Schlagwörtern. Es war zu erkennen, dass ausformulierte Suchanfragen relevantere Suchergebnisse ergaben als die reine Suche mit Schlagworten.



Suchanfrageverarbeitung

Bis auf einige Ausnahmen wurden die Suchanfragen richtig erkannt und verarbeitet. Trotz lauter Nebengeräusche bspw. durch Baulärm und hochfrequentierte Straßen, wurden die Suchanfragen korrekt erkannt. Wie oben bereits erwähnt, führten ausformulierte Suchanfragen zu relevanteren Ergebnissen. Dennoch sollte erwähnt werden, dass die Suche durch Schlagworteinsprache ebenso zum richtigen Ergebnis führte. Die Probanden waren mit der Qualität der Suchergebnisse sehr zufrieden.



Personalisierung

Der aktuelle Standort der Probanden wurde immer richtig erkannt und bei passenden Anfragen in die Suchanfragenverarbeitung einbezogen. Einige Probanden nutzten unbemerkt die Dialogfähigkeit von Google. Erst auf Nachfrage bemerkten sie, welche Funktion sie gerade nutzen.



Barrierefreiheit

Grafiken und Schrift der Ergebnispräsentation werden in ausreichender Größe dargestellt und können darüber hinaus vergrößert werden, ebenfalls der Kontrast ist angemessen. Nicht alle Ergebnisse können akustisch wiedergegeben werden. Die Ergebnisse, die von der Sprachausgabe unterstützt werden, werden in zufriedenstellender Lautstärke wiedergegeben.



Ergebnispräsentation

Überwiegend führte eines der ersten drei aufgelisteten Suchergebnisse zum Erfolg. Bei der Suche nach bekannten Orten oder Persönlichkeiten bspw. kam es vor, dass das Ergebnis zusätzlich über eine Sprachausgabe präsentiert wurde. Die Probanden waren von dieser Funktion positiv überrascht. Passend zur Suchanfrage wurden bspw. Bilder, Videos oder Google-Maps-Einträge angezeigt. Bei einigen Suchanfragen wurde ein Knowledge Graph angezeigt, welchen die Probanden als hilfreich empfanden. Es kam vor, dass die Zusatzoption der Routenbeschreibung von einigen Probanden nicht erkannt wurde. Dies führte dazu, dass Probanden mit dem Ergebnis unzufrieden waren.



Hilfestellung

Probanden nutzten bei Problemen nicht das Handbuch.



SUS-Score

Der durchschnittliche über alle Probanden ermittelte SUS-Score betrug 82% und weist damit auf eine gute Usability des Systems hin.

Apple Siri: Auswertung Nutzertest

N = 6 (+ ein 8jähriges Kind)



Zugang/Einstieg

Bei den Probanden handelte es sich um Smartphone-Nutzer, von denen nur zwei von sieben über ein iPhone verfügten. Insgesamt ist zu erwähnen, dass die sprachgesteuerte Websuche gar nicht bis vereinzelt genutzt worden ist. Vielmehr wurde Siri als Sprachassistent genutzt. Der 8-jährige Proband nutzte bspw. Siri auf seinem iPad zum Vergnügen, in dem er spaßige Fragen stellte („Siri, wo wohnst du? Wie alt bist du?“). Die restlichen Nutzer waren mit der Funktion vertraut, haben sie jedoch nie im Privaten genutzt. Trotz dessen fiel allen der Einstieg leicht.



Bedienung

Die Bedienung fiel allen Nutzern leicht, trotz wenigen Erfahrungen im Bereich der sprachgesteuerten Websuche. Die Bedienung ist nutzerfreundlich und leicht zu verstehen.



Suchanfrage/Verarbeitung

Die Suchanfrage wurde von den Nutzern vorerst in ganzen Sätzen frei ausgesprochen. Dadurch, dass Siri die Frage oftmals nicht verstanden hatte, stellten die Nutzer die Suchanfrage in kurzer Form nochmals. Bspw. wurde die Frage „Was hat Mark Zuckerberg studiert?“ von fast allen Nutzern auf „Studium Zuckerberg“ oder „Mark Zuckerberg Studium“ verändert. Im Laufe des Nutzertests wurden die Fragen immer häufiger in Stichwörtern gestellt.

Ebenfalls positiv aufgefallen war die Darstellung der Suchanfrage, die dem Nutzer half einen Überblick über seine Frage zu behalten. Durch das Mitschreiben waren die Nutzer besser in der Lage ihre Suchanfrage umzuformulieren.



Personalisierung

Mit Hilfe der Ortungsdienste wusste Siri, wo genau sich der Nutzer befand. Die Suchergebnisse zu den jeweiligen Orten (Museen) wurden in Form einer Karte angezeigt und zusätzlich die Entfernung, Öffnungszeiten und der Weg mit den öffentlichen Verkehrsmitteln.



Barrierefreiheit

Der Siri-Home-Button reagierte einwandfrei auf eine Suchanfrage: auf alle getesteten Stimmen, alle Akzente und alle Sprachfehler. Hintergrundgeräusche haben keinen negativen Einfluss auf Siri, so hinderte z.B. eine Suchanfrage an einer vielbefahrenen, mehrspurigen Straße Siri nicht daran, die Frage richtig zu verstehen und ein Ergebnis zu liefern. Längere Sprechpausen stellen jedoch eine Barriere da: Siri beendet die Suchanfrage nach mehreren Sekunden und gibt dann keine oder falsche Ergebnisse wieder. Die Sprachsteuerung gibt die Ergebnisse meist schriftlich wieder, jedoch kann die Schriftgröße nicht geändert werden (muss vor der sprachgesteuerten Suche vom Nutzer in den Einstellungen geändert werden). Darüber hinaus funktioniert Siri nicht ohne Internetverbindung, sie gibt allerdings dem Nutzer den Hinweis, dass eine Internetverbindung im möglichen Fall nicht besteht und macht ihn somit darauf aufmerksam (akustisch oder springt manchmal in die Einstellungen). Obwohl die meisten Probanden nicht im Besitz eines eigenen Apple-Geräts sind, waren die Erfahrungen und Umgang positiv.



Ergebnispräsentation

Der Großteil aller Testpersonen ist sehr zufrieden mit den Ergebnissen. Hierbei ist zu erwähnen, dass sowohl die Information in akustischer als auch geschriebener Form zufrieden stellend sind. Bei ortsbezogenen Fragen gibt Siri eine Wegbeschreibung bzw. eine Karte wieder. Dadurch bekommt der Nutzer den direkten Weg zu dem jeweiligen Ort visuell wieder. Darüber hinaus ist die optische Darstellung sehr ansprechend. Ergebnisse zu den aktuellen Aktienkursen werden bspw. in übersichtlichen Tabellen wiedergegeben. Ein kleiner Teil der Nutzer war mit Siris Leistung unzufrieden: Anfragen wurden nach mehreren Versuchen abgebrochen, da die Frage nicht richtig verstanden wurde. Dies sorgt beim Nutzer für Verwirrung, Unzufriedenheit und Frustration.



Hilfestellung

In manchen Fällen ist Siri fähig dem Nutzer mit einer Hilfestellung entgegen zu kommen. Auf die Frage „Wo ist in der Hamburger Meile ein Esprit-Shop?“ fragt Siri, nach welcher Art von Geschäft man suche. Im Anschluss wird das Ergebnis angezeigt. Dies geschieht jedoch nicht bei jeder Suchanfrage, sondern nur bei vereinzelt Situationen. Die Nutzer reagieren unterschiedlich: für manche scheint die Hilfestellung von Vorteil zu sein, für andere Nutzer ist dies verstörend und sie stellen eine neue Suchanfrage.



SUS-Score

Der durchschnittliche über alle Probanden ermittelte SUS-Score betrug 70,4% und weist damit auf eine grenzwertige Usability des Systems hin.

Microsoft Cortana: Auswertung Nutzertest

N = 6



Zugang/Einstieg

Der Startprozess des Sprachassistenten hat sich während des gesamten Nutzertests als einfach und unkompliziert erwiesen. Die Anwendung lässt sich schnell finden und mit wenigen Handgriffen starten. Die visuelle Darstellungen und akustische Wiedergabe hat die Probanden die Bereitschaft der Sprachsteuerung erkennen lassen, Nachfragen waren nicht notwendig. Der händische Start von Cortana wurde zu keinem Zeitpunkt negativ erwähnt, auch wenn an manchen Stellen eine Neuformulierung des Sprachbefehls aus Sicht der Nutzer notwendig war. Es wurde auch gar nicht versucht den Assistenten per Sprache zu starten.



Bedienung

Sowie der Startprozess bzw. der Zugang reibungslos verlaufen ist, so wurde auch die Bedienung der Sprachsteuerung von den Nutzern im Einzelnen nicht bemängelt. Das Gerät lässt sich „mit einem Wisch“ bedienen und wie bereits erwähnt ist das manuelle Starten auch aus Sicht der Bedienung problemlos möglich. Ein weiterer Aspekt ist das Feedback der Software, welches die Bedienung positiv beeinflusst hat. Es waren keine überhasteten Handgriffe oder negativen Äußerungen der Nutzer zu beobachten. Auch hat die Info darüber, in welchem aktuellen Status sich derjenige befindet, den Erwartungen entsprochen.

Auch wenn die Heuristik ergeben hat, dass über den ersten Sprachbefehl hinaus, Ergebnisse aus Trefferlisten nicht per Sprachbefehl abrufbar waren, so haben die Nutzer diese Möglichkeit der Navigation gar nicht in Erwägung gezogen. Aus den nicht aufgezeichneten Gesprächen, hatten die Probanden hinsichtlich dieser Thematik keine Erwartung an das System.

Besonders auffällig während des gesamten Nutzertests waren die Art und Weise der Sucheingabe. So haben alle Probanden analog zur bekannten händischen Suche via Desktop oder Mobil ihre Suchformulierung schlagwortartig ausgesprochen. Fragen wurden demnach nicht ausformuliert.



Suchanfrageverarbeitung

Einfache, kurze und simple Anfragen haben Cortana keine Schwierigkeiten bereitet. Aber auch längere Sprachbefehle wurden verarbeitet und zu keinem Zeitpunkt aufgrund der Länge abgebrochen. Diese Arten von Suchanfragen wurden stets zufriedenstellend verarbeitet. Durch kurze Sprachpausen wurde das System nicht dazu verleitet, die Suchanfrage vorzeitig beginnen zu lassen, als es eigentlich vom Nutzer gewünscht war.

Da Anfragen durch den Sprachassistenten immer mitgeschrieben werden, hatten die Nutzer bei Abkürzungen und englisch auszusprechenden Begriffen negative Anmerkungen hinsichtlich der Erkennung gemacht. So wurden weniger populäre Abkürzungen oder anglizistische Aussprachen von Cortana unzureichend verarbeitet bzw. nicht erkannt. Als Ergebnis wurden hier ähnlich klingende Treffer ausgeliefert. Die automatische Korrektur hingegen wurde positiv aufgefasst.

Darüber hinaus lieferte der Sprachassistent unregelmäßig und zu selten direkte Antworten. Dies wurde auf Nachfrage hin vom größten Teil der Probanden jedoch gewünscht, auch wenn die Erwartungen an das System hier nicht gegeben sind. Vordergründig wurde damit argumentiert, dass man der Sprachsteuerung dies nicht zutrauen würde, weil man es zu selten bis gar nicht nutzt und oder kennt. Die in den Fachzeitschriften und anderen Artikeln viel gelobte Dialogfähigkeit von Cortana wurde von den Probanden vollständig ignoriert bzw. schlichtweg nicht genutzt.

Cortana hat von Haus aus eine Routenfunktion implementiert. Diese wurde während der Tests bei entsprechenden Suchanfragen konsistent angezeigt. Jedoch empfanden die Probanden dies nicht hilfreich, weil die Funktion eingeschränkt ist auf PKW-Routenberechnungen. Daher wurden Anfragen nach öffentlichen Verbindungen nicht zufriedenstellend verarbeitet.



Personalisierung

Ähnlich wie bei den vorangegangenen Evaluierungen hat auch bei den Nutzertests keine wirkliche Personalisierung der Ergebnisse stattgefunden. Zum einen lag es daran, dass aufgrund der Testvoraussetzungen (mehrere Probanden) die Möglichkeit, persönliche Interessen und favorisierte Einstellungen des Nutzers zu hinterlegen, nicht genutzt werden konnte. Diese Angaben sind jedoch laut Hersteller zur Verfeinerung und Personalisierung der Ergebnisse notwendig. Zum anderen konnte aufgrund der ständig wechselnden Probanden und

somit des wiederholten Zurücksetzen des Systems auf Werkseinstellungen nicht die Lernfähigkeit der Anwendung in vollem Umfang genutzt werden.

Bei der ortsbezogenen Suche wurden individuelle Ergebnisse basierend auf dem Standort des Nutzers ausgegeben. Der Standort wurde in allen Fällen richtig lokalisiert und alle relevanten Ergebnisse wurden angezeigt. Dies wurde von allen Probanden als besonders positiv hervorgehoben.



Barrierefreiheit

Die Darstellung der Ergebnisse im Hinblick auf die Barrierefreiheit war grundsätzlich unproblematisch. Die verwendete Schrift, Schriftgröße und Farbgebung sowie die Darstellung aller weiteren Elemente wurde durchwegs als angenehm empfunden. Daher war die fehlende Möglichkeit, die Darstellung bspw. zu Gunsten einer besseren Lesbarkeit zu ändern, für die Probanden nicht relevant. Ähnlich verhielt es sich mit dem Aufrufen einzelner Ergebnisse mittels Sprachsteuerung. Diese Möglichkeit wurde von keinem der Probanden in Anspruch genommen. Die akustische Wiedergabe der Ergebnisse fand allerdings nur vereinzelt und in unregelmäßigen Abständen statt, was als irritierend empfunden wurde. Zudem war die Wiedergabe in lauter Umgebung zu leise und konnte kaum wahrgenommen werden. Insgesamt hätten sich die meisten Probanden durchgängig eine akustische Wiedergabe gewünscht, insofern diese gut verständlich wäre.



Ergebnispräsentation

Die Präsentation wurde generell als gut bewertet, besonders die Ergebnisdarstellung bei der ortsbezogenen Suche sowie die Karten- und Routenansicht von Cortana ist den Probanden positiv aufgefallen. Bei der Ergebnisliste der Bing-Suche war sowohl die Anzahl der angezeigten Ergebnisse als auch die Darstellung von verschiedenen Ergebnisformaten, wie bspw. Bildern und Videos, insgesamt zufriedenstellend. Bei der Suche nach öffentlichen Einrichtungen lieferte Cortana sporadisch eine eigene Ansicht, in der zwar einige relevante Informationen zusammengefasst waren, die jedoch von den Probanden als unübersichtlich und unvollständig empfunden wurde. So war bspw. die Angabe zu den Öffnungszeiten nur schwer zu finden. In diesem Zusammenhang wurde ebenfalls die variierende Ergebnisdarstellung, einerseits als systemeigene Übersichtsdarstellung und andererseits als gewöhnliche Ergebnislistenansicht, bemängelt. Dieses fiel den Probanden an mehreren Stellen negativ auf, so bspw. auch bei der Suche nach Aktienkursen, bei der es sowohl eine systemeigene Darstellung als auch eine

gewöhnliche Ergebnisliste gab. Dabei wurde der Ansatz, alle relevanten Informationen in einer zusammenfassenden Ansicht zu präsentieren, stets als sehr positiv und wünschenswert empfunden.



Hilfestellung

Ein Hinweis, der beim Starten der Anwendung erklärt, wann das System bereit ist und die Suche beginnen kann, lieferte den Probanden eine gute Orientierung für den Umgang mit der Anwendung. In Bezug auf die sprachgesteuerte Suche gab es seitens des Systems allerdings nur wenig bis keine Hilfe. Bei Anfragen, die aufgrund der Akustik nicht verarbeitet werden können, gab es einen entsprechenden Hinweis. Negativ fielen dagegen Fehlermeldungen auf, die lediglich mit einem Fehlercode versehen waren und wenig Hinweise zur Problembeseitigung lieferten.



SUS-Score

Der durchschnittliche über alle Probanden ermittelte SUS-Score betrug 65% und weist damit auf eine grenzwertige Usability des Systems hin.

Amazon Fire OS: Auswertung Nutzertest

N = 6



Zugang/Einstieg

Die Probanden hatten keinerlei Probleme, die Sprachsuche des Amazon Fire OS durch einen langen Knopfdruck zu starten, nachdem ihnen die Funktion einmalig vorgeführt wurde und empfanden den manuellen Einstieg als simpel.

Probleme gab es dagegen bei der Länge und Sprechgeschwindigkeit der Spracheingabe. So ist circa eine Sekunde Pause bereits problematisch, da Amazon Fire OS schon mit der Verarbeitung beginnt. Teilweise auch, obwohl die Probanden direkt mit der Ausformulierung der Suchanfrage fortfuhren. Dies war ebenso ein Problem bei Suchanfragen, die aus vollständig ausformulierten Sätzen bestanden. Auch kam es zur vorzeitigen Verarbeitung der Suchanfragen, obwohl die Nutzer die Spracheingabe nicht beendet hatten. Die Anfragen mussten daher zügig eingesprochen werden. Nach dieser Erfahrung überlegten einige Probanden zunächst einen Moment, um eine passende Formulierung zu finden, anstatt die Anfrage spontan einzusprechen.

Der Zugang wird hier als neutral bis negativ bewertet. Man ist sich nicht sofort im Klaren darüber, welches Vokabular nun verwendet werden soll, um eine gezielte Websuche durchzuführen, da vorerst keine Hilfestellungen gegeben werden. Ansonsten funktioniert der Einstieg simpel über einen langen Knopfdruck.



Bedienung

Zusätzlich zu den unter "Zugang/Einstieg" genannten Schwierigkeiten war es nicht möglich, die sich schon in der Verarbeitung befindlichen Suchanfragen zu beenden. Ein Abbruch der akustischen Rückmeldungen sowie die Auswahl von einzelnen Ergebnissen aus der Bing-Suche durch Sprachkommandos waren ebenso nicht durchführbar.



Suchanfrageverarbeitung

Positiv hervorzuheben ist die Verarbeitung der Suchanfragen: bei lauten Hintergrundgeräuschen wie z.B. einer vielbefahrenen Straße. So gut wie alle Suchanfragen wurden korrekt erkannt und verarbeitet. Deutschsprachige Suchanfragen wurden durchgehend korrekt verarbeitet.

Engelssprachige Personennamen dagegen wie „Mark Zuckerberg“ konnte das Amazon Fire Phone nicht korrekt verarbeiten. So wurde aus „Zuckerberg“ bspw. zu „Ackerberg“. Einfache englische und französische Wörter, die im Alltag gebräuchlich sind, hingegen wurden korrekt umgesetzt.



Personalisierung

Eine Personalisierung in irgendeiner Form ist nicht möglich. Filter innerhalb der Suche können nicht durch eine Spracheingabe ausgewählt werden. Allerdings wurde dies von den Probanden weder versucht noch bemängelt.



Barrierefreiheit

Die vom Amazon Fire Phone nicht unterstützten Funktionen zur Barrierefreiheit, wie die Möglichkeit zur Schriftvergrößerung, wurden in den Nutzertests nicht vermisst. Auch ist die grafische Darstellung der Ergebnisse im Sinne der Bild- und Schriftgröße sowie des Kontrasts nicht negativ bewertet worden. Die in der expertenbasierten Evaluation getesteten Szenarien, wie z.B. Dialekte und Sprachfehler, stellten in den Nutzertests ebenso keine Hürden dar. Dies kann jedoch auch durch die beschränkte Probandenanzahl begründet werden.



Ergebnispräsentation

Die Präsentation der Suchergebnisse wurde durch die Nutzer größtenteils als gelungen aufgefasst, dennoch nannten sie diverse Verbesserungsvorschläge. Bei der Suche nach Video-Tutorials und Börsenkursen wurde bspw. von mehreren Teilnehmern begrüßt, dass ein entsprechendes Youtube-Video bzw. der Börsenkurs direkt in die Ergebnisliste von Bing eingebettet worden ist. Jedoch wünschte sich ein Proband, das Video direkt präsentiert zu bekommen, ohne dafür das Suchergebnis extra auswählen zu müssen.

Ebenfalls positiv empfanden die Probanden, dass generell verschiedene Ergebnistypen (Bilder, Videos, Texte, Karten...) ausgegeben wurden. Allerdings wurde bei der Suche nach einem Shop in einem Einkaufszentrum eine Karte mit Routenberechnung geöffnet, wobei in diesem Fall ein Lageplan im Bildformat gewünscht war.

Nicht von den Nutzerinnen und Nutzern gefragt war hingegen die, in der expertenbasierten Evaluation bemängelte, fehlende Möglichkeit Antworten in akustischer Form auszugeben.



Hilfestellung

Das Fire Phone bietet eine sehr überschaubare Anzahl von Hilfestellungen. Dies ist in den Nutzertests jedoch kaum störend aufgefallen. Die Teilnehmerinnen und Teilnehmer unternahmen demnach nicht den Versuch, Hilfe in Textform (z.B. ein Handbuch) zu öffnen. Als störend empfanden die Probanden lediglich die Fehlermeldungen ("Das kann ich noch nicht"), die zum Teil willkürlich auftraten und somit nicht deutlich wurde, was die Fehlerquelle war und wie diese behoben werden könnte. Diese Tatsache unterstreicht die negative Nutzererfahrung und weist dadurch auf einen dringenden Handlungsbedarf hin.



SUS-Score

Der durchschnittliche SUS-Score über alle Probanden beträgt 70% und weist damit auf eine grenzwertige Usability des Systems hin.

Empfehlungen für die Entwicklung der idealen Voice Web Search

Erfolgskriterien

Die folgenden Tabellen zeigen die in den experten- und nutzerbasierten Evaluationen ermittelten Erfolgskriterien und ihre jeweilige Umsetzung bei den getesteten Geräten in der Übersicht. Diese Zusammenstellung bildet im Anschluss die Basis für die Konzeption des Idealprozesses der Voice Web Search sowie der Skizzen zur prototypischen Realisierung ausgewählter Anforderungen.

Zugang / Einstieg

Nr.	Kriterium		
1.1	Eindeutig als Sprachsteuerung erkennbar		

Bedienung

Nr.	Kriterium		
2.1	Lange Sprechpausen möglich (mehr als 2 Sekunden)		
2.2	Stichwortsuche führt zu brauchbaren Ergebnissen		
2.3	Interaktion mit Suchergebnissen per Sprache möglich		
2.4	Suchverarbeitung per Sprache beenden		
2.5	Unterstützung von Dialogen		

2.6	Echtzeitfeedback bei Suchanfrageverarbeitung	   	
2.7	Spracherkennung funktioniert zuverlässig trotz lauter Umgebungsgeräusche	   	

Suchanfrageverarbeitung

Nr.	Kriterium		
3.1	Akzente und Dialekte werden korrekt erkannt	   	
3.2	Korrekte Erkennung englischsprachiger Sucheingaben		  
3.3	Unbekanntere Abkürzungen können interpretiert werden		  
3.4	Korrekte Erkennung der Anfrage	   	
3.5	Autokorrektur wird unterstützt	   	

Personalisierung

Nr.	Kriterium		
4.1	Standort wird bei der Suchanfrage einbezogen		
4.2	Die Sprache der Ausgabe wird der Sprache der Eingabe angepasst		

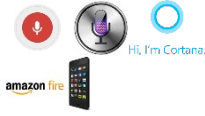
Barrierefreiheit

Nr.	Kriterium		
5.1	Grafiken bei den Ergebnissen werden in ausreichender Größe dargestellt		
5.2	Die Schrift wird bei den Ergebnissen in ausreichender Größe dargestellt		
5.3	Der Kontrast bei der Ergebnisdarstellung ist angemessen		
5.4	Alle Ergebnisse können akustisch wiedergegeben werden		
5.5	Schriftvergrößerung wird unterstützt		

Ergebnispräsentation

Nr.	Kriterium		
6.1	Sprachausgabe bei Faktenrecherchen	 	   Hi, I'm Cortana.
6.2	Akustisches Feedback zu den Ergebnissen nach der Suchanfrageverarbeitung		    Hi, I'm Cortana.
6.3	Sprachausgabe direkt zu den Suchergebnissen wird unterstützt		     Hi, I'm Cortana.
6.4	Zeigt passende Darstellungen der Ergebnisse zu den Micro-Moments		    Hi, I'm Cortana.
6.5	Die Quellen der Suchergebnisse sind immer erkennbar	     Hi, I'm Cortana.	
6.6	KWIC wird unterstützt	    Hi, I'm Cortana.	
6.7	Die Trefferbeschreibungen sind verständlich	     Hi, I'm Cortana.	
6.8	Die Ergebnisse sind auf dem aktuellen Stand	     Hi, I'm Cortana.	

Hilfestellung

Nr.	Kriterium		
7.1	Fehlermeldungen sind hilfreich		
7.2	Das System bietet aktive Unterstützung bei der Bedienung der sprachgesteuerten Suche		
7.3	Links zur Hilfe sind vorhanden und gut sichtbar		
7.4	Eine Hilfestellung in Textform wird angeboten		
7.5	Feedback und Hilfestellung bei Nulltreffer-Seiten und nicht verstandener Spracheingabe		

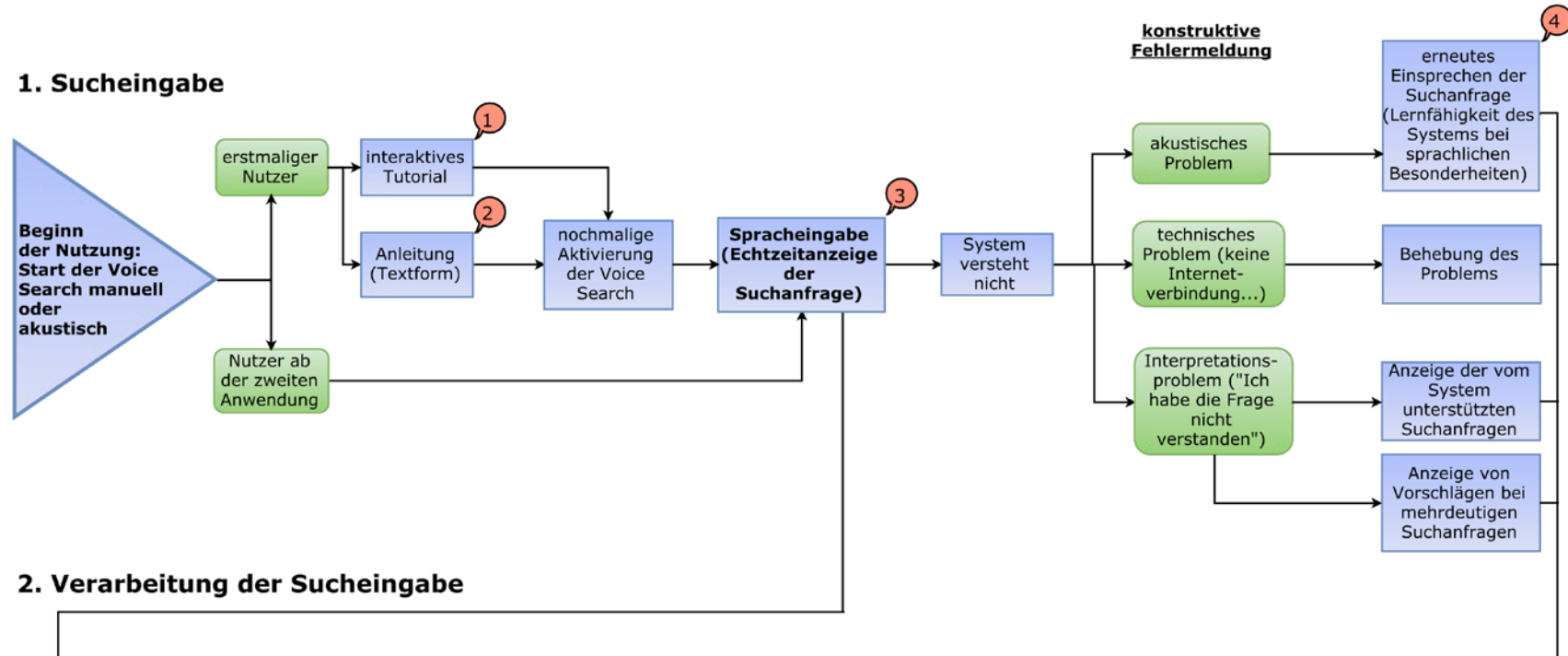
Prozessablauf und prototypische Umsetzung der idealen sprachgesteuerten Websuche

Der illustrierte Prozess und die Skizzen basieren auf den Ergebnissen der durchgeführten expertenbasierten Evaluationen und den Nutzertests und den daraus abgeleiteten ermittelten Empfehlungen für die Realisierung einer optimalen Voice Web Search. Es wird aufgezeigt, wie das System in den einzelnen Schritten im Suchprozess (Sucheingabe, Verarbeitung der Sucheingabe, Ausgabe der Suchergebnisse, Interaktion mit den Suchergebnissen) reagieren sollte und welche Form von Feedback und Hilfestellung wünschenswert wäre.

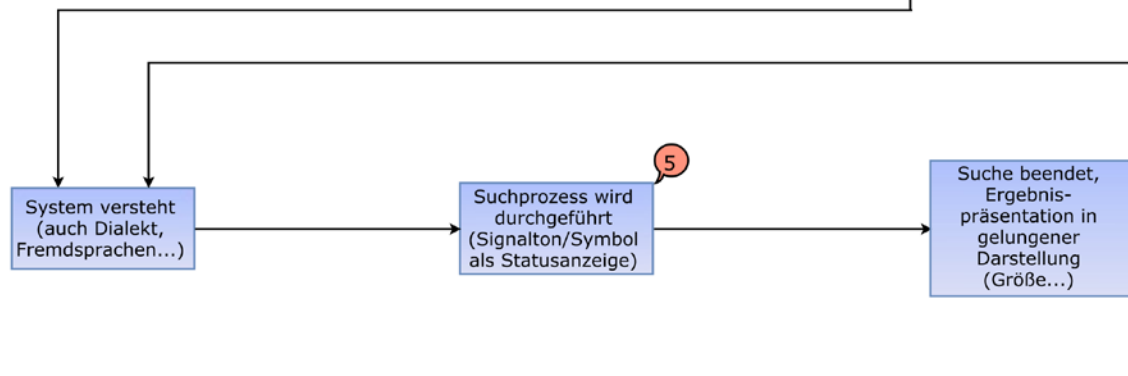
Ausgewählte Prozessschritte werden dann im Folgenden durch Skizzen möglicher Nutzerinterfaces visualisiert. Für eine vergrößerte Darstellung des Prozesses s. Anlage 7.

Suchprozess innerhalb einer optimalen Voice Search

1. Sucheingabe

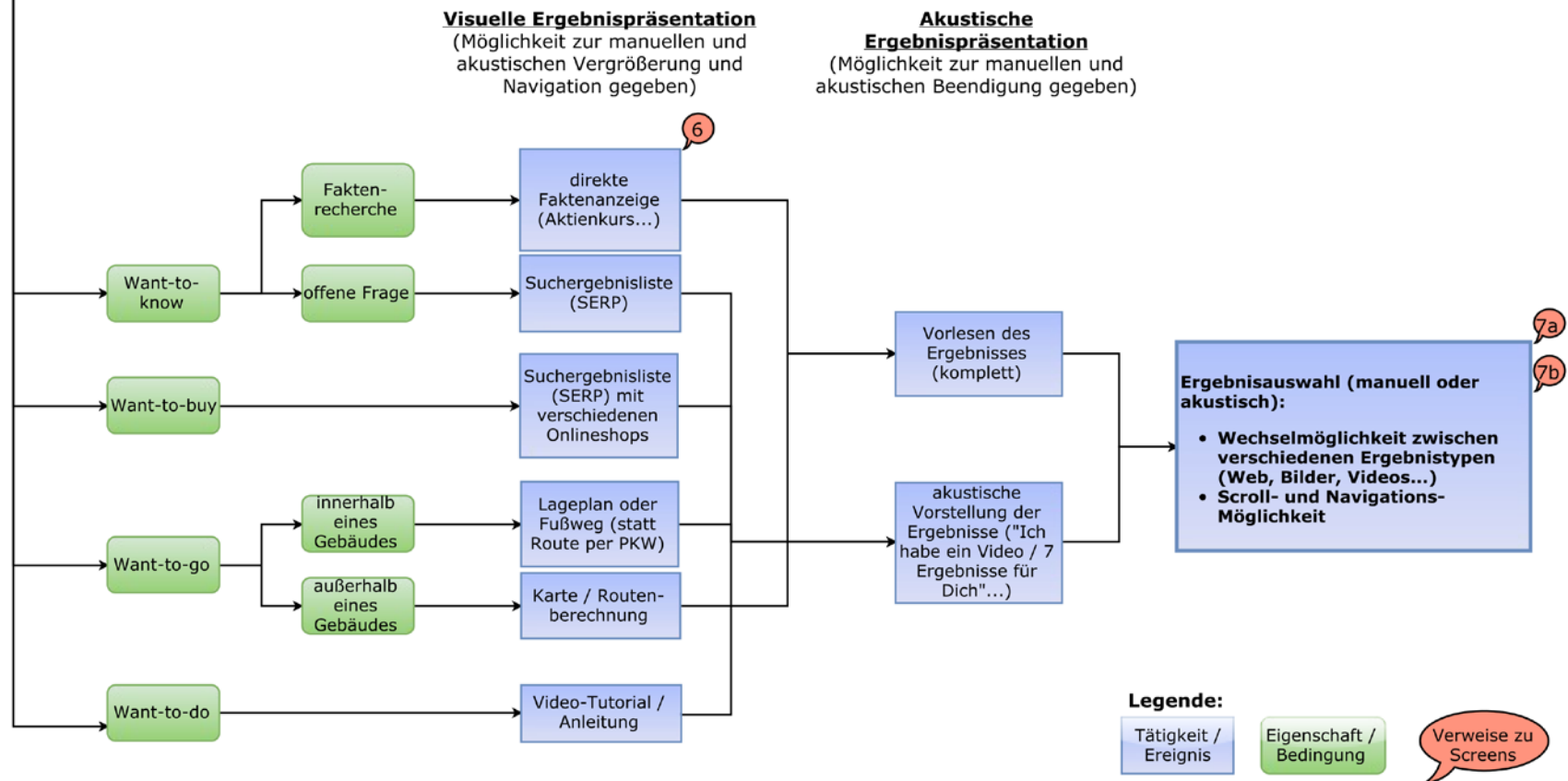


2. Verarbeitung der Sucheingabe



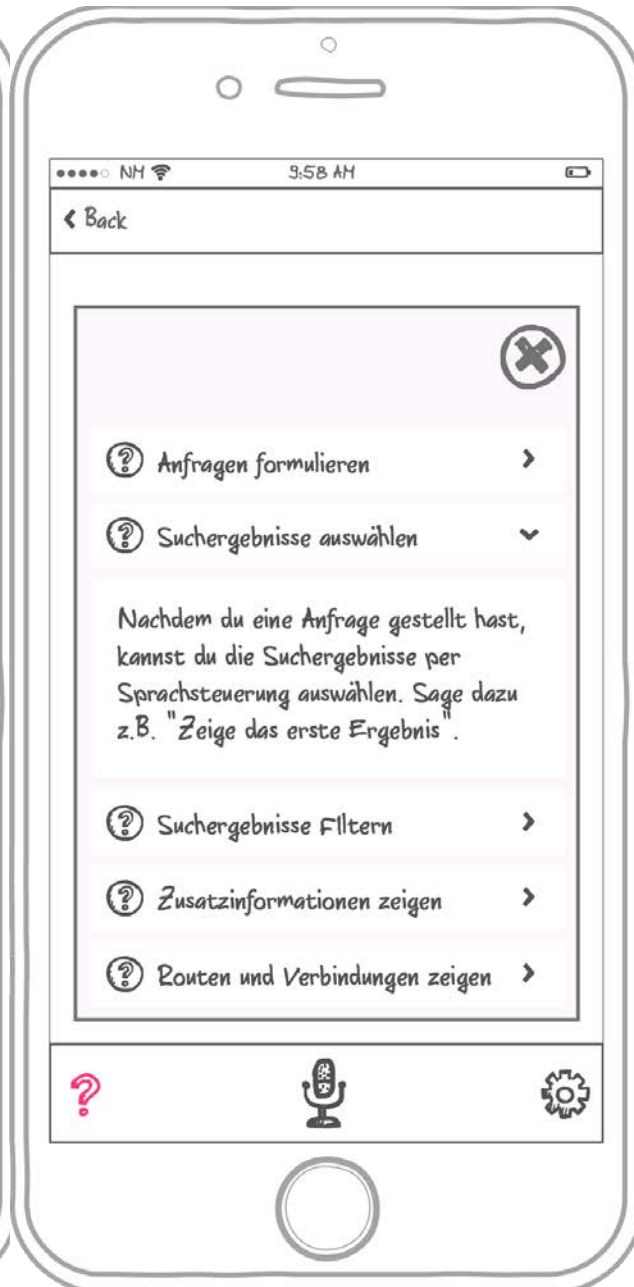
3. Ausgabe der Suchergebnisse

4. Interaktion mit den Suchergebnissen



1

2



3

4



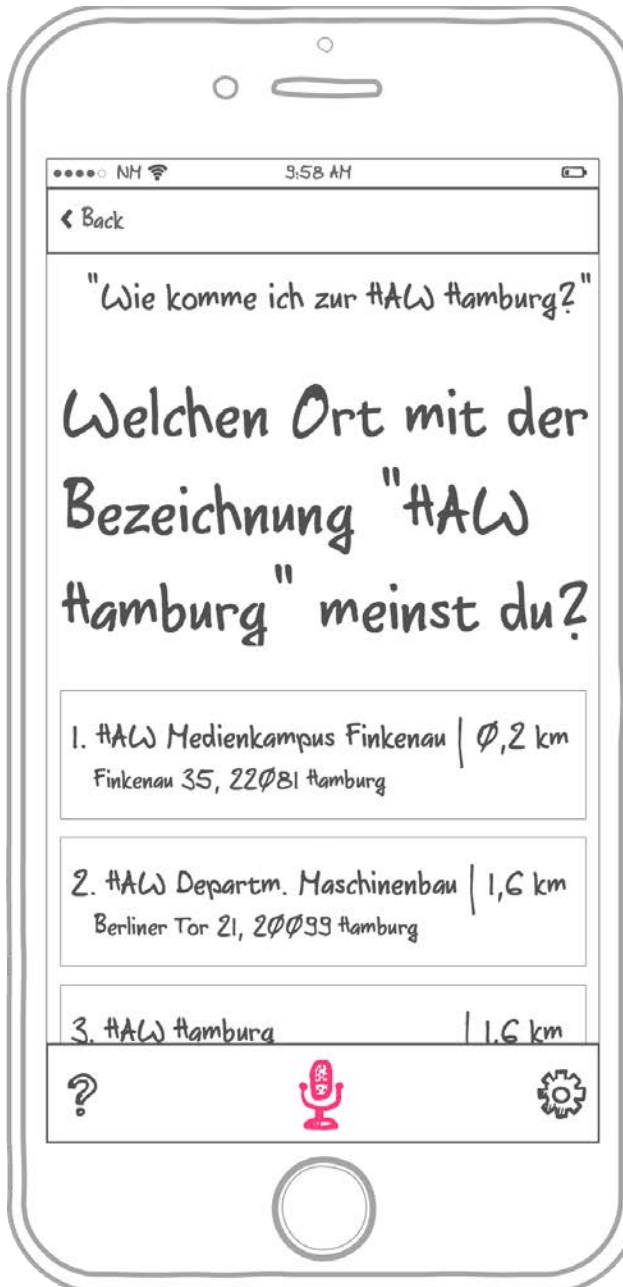
5

6



7a

7b



Fazit

Zusammenfassung

Der vorliegende Projektbericht beschreibt die notwendigen Handlungsempfehlungen für die Gestaltung einer nutzerfreundlichen Voice Web Search. Die Empfehlungen basieren auf einem multimethodischen Ansatz, bestehend aus heuristischer Evaluation, szenarienbasiertem Walkthrough, Retrievaltest und Nutzerstudie. Mit Hilfe dieser Methoden wurden die Stärken und Schwächen der markführenden Smartphone-Systeme in Bezug auf ihre sprachgesteuerten Websuchen (Amazon Fire OS, Apple Siri, Google Now, Microsoft Cortana) identifiziert.

Die durchgeführten Evaluationen zeigen, dass Google sich in Bezug auf Usability, User Experience und Ergebnisqualität von den anderen Systemen positiv abhebt und somit als Gesamtsieger bezeichnet werden kann. Daher wurden viele Funktionen und Gestaltungsaspekte Googles für die Konzeption einer nutzerfreundlichen sprachgesteuerten Websuche adaptiert. Bei den anderen Systemen zeigten sich mehr Schwächen in allen evaluierten Bereichen. So konnte Siri besonders bei der Ergebnisqualität nicht überzeugen. Amazon zeigte vereinzelt Schwächen in der Bedienung und Microsoft Cortana konnte bei der Hilfestellung nicht überzeugen, da nur Fehlercodes ohne direkten Bezug zu dem aufgetreten Problem angezeigt wurden.

Die Ergebnisse der Studien wurden genutzt, um daraus Erfolgskriterien abzuleiten, die wiederum für die Entwicklung einer eigenen Voice-Web-Search-Anwendung herangezogen wurden und in einem idealen Prozessablauf sowie dazugehöriger Visualisierung möglicher Nutzerinterfaces mündeten.

Kritische Würdigung

Das beschriebene Vorhaben betrat Neuland und hat somit einen pilothaften Charakter. Die Evaluierung von sprachgesteuerter mobiler Websuche ist bislang in der Wissenschaft nicht angekommen. Das Projektteam konnte daher bei der Methodenwahl auf wenig Etabliertes zurückgreifen. So wurden zwar gängige Methoden der Usability-Forschung und Suchmaschinenevaluierung für den Anwendungsfall der sprachgesteuerten mobilen Suche adaptiert; bedingt durch die hohe Aktualität des Themas basierten diese Adaptionen aber überwiegend auf eigenen Überlegungen, weniger auf wissenschaftlich bestätigten Ansätzen.

Ergänzend zu der Herausforderung der Methodenwahl zeigten sich Schwierigkeiten bei der genauen Abgrenzung des zu testenden Anwendungsfalls bzw. der sprachlichen Aufbereitung der Ergebnisse. So mussten sowohl die Begriffe „Sprachgesteuerte Suche“, „Sprachgesteuerte

Websuche“ sowie „Sprachsteuerung“ und „Personal Assistant“ als auch „Testsystem“ und „Testgerät“ differenziert werden, um – soweit möglich – ein einheitliches Begriffsverständnis aller Projektbeteiligten sicherzustellen. Dabei zeigte sich, dass eine trennscharfe Abgrenzung nicht in Fällen möglich war.

Weiterer Forschungsbedarf

In dem Projekt wurde die Voice Web Search von vier Systemen evaluiert. Dabei wurden die Bereiche Personalisierung und Personal Assistant nicht systematisch berücksichtigt, da das Ziel des Projekts darin bestand, festzustellen, wie gut die sprachgesteuerte Websuche auf den Systemen bisher funktioniert. Personalisierung, Zusammenspiel der Voice Web Search mit Personal-Assistent-Funktionalitäten sowie die Lernfähigkeit bieten daher sehr interessante Aspekte für weitere Forschungsarbeiten. So könnten im Bereich der Personalisierung datenschutzrelevante Aspekte aufgegriffen oder Langzeitstudien durchgeführt werden, die dabei zur Evaluierung der Lernfähigkeit (Anpassung an die Nutzerbedürfnisse, z. B. in Bezug auf Mundart und individuelle Suchvorlieben) der Systeme im natürlichen Umfeld dienen können. Zusätzlich ist es notwendig, die versuchsweise angewandten Methoden in weiteren Studien zur sprachgesteuerten mobilen Websuche zu validieren und zu verfeinern.

Quellenverzeichnis

- Amazon.de (2015a). Fire Phone User Guide. URL: <https://s3-us-west-2.amazonaws.com/customerdocumentation/Fire+Phone+User+Guide+DE.pdf>
- Amazon.de (2015b). Amazon.de Hilfe: Mit Ihrem Smartphone sprechen. URL: <http://www.amazon.de/gp/help/customer/display.html?nodeId=201555130>
- Android Advisor (2014). Amazon unveils its first smartphone. Android Advisor, 1(6), 33-37. URL: <http://212.50.14.233/Knowledge/Computing%20%26%20Games/Magazines/Android%20Advisor%20-%20Issue%2006%20-%202014.pdf> (Abruf am 05.11.15, Dokument auf dem Server nicht mehr verfügbar)
- Banks, C. (2015). Fire Phone 2 coming before the holidays? Geeksnap.com. URL: [http://www.geeksnack\(Banks, 2015\).com/2015/09/08/fire-phone-2-coming-before-the-holidays/](http://www.geeksnack(Banks,2015).com/2015/09/08/fire-phone-2-coming-before-the-holidays/)
- Beiersmann, Stefan (2015). Cyanogen-CEO verspricht „tiefe Cortana-Integration“ in Cyanogen OS. München: ZDNet.de. URL: <http://www.zdnet.de/88246322/cyanogen-ceo-verspricht-tiefe-cortana-integration-in-cyanogen-os/>
- Bensinger, G. (2015). Amazon Curtails Development of Consumer Devices. WSJ.com. URL: <http://www.wsj.com/articles/amazon-curtails-development-of-consumer-devices-1440632203>
- Bergert, Denise (2015). Connected Car: Microsoft will Cortana auf die Windschutzscheibe bringen. Hannover: heise.de. URL: <http://www.heise.de/newsticker/meldung/Connected-Car-Microsoft-will-Cortana-auf-die-Windschutzscheibe-bringen-2818993.html>
- Blum, C. (2015). iPhone: Siri als Anrufbeantworter - COMPUTER BILD. Computerbild.de. URL: <http://www.computerbild.de/artikel/cb-News-Handy-iPhone-Siri-Anrufbeantworter-12513929.html>
- Brodersen, B. (2014). Amazon wirft Pläne für Fire Phone 2 über den Haufen. Areamobile.de. URL: <http://www.areamobile.de/news/30179-bericht-amazon-wirft-plaene-fuer-fire-phone-2-ueber-den-haufen>
- Carr, A. (2015). The real story behind Jeff Bezos's Fire Phone debacle and what it means for Amazon's future. Fastcompany.com. URL: <http://www.fastcompany.com/3039887/under-fire>
- Chip.de (2015). Fire Phone Verkauf gestoppt: Ist das Fire Phone 2 bereits im Anmarsch? URL: <http://www.chip.de/news/Fire-Phone-Verkauf-gestoppt-Ist-das-Fire-Phone-2-bereits-im->

Anmarsch_81127926.html

Computerbild (o. J.). iOS 9: Die wichtigsten Neuerungen auf einen Blick - Bilder, Screenshots - COMPUTER BILD. Computerbild.de. URL: <http://www.computerbild.de/fotos/Die-ersten-Bilder-von-iOS9-11890860.html#1>

Czerulla, Hannes (2015). Microsoft integriert Sprachassistentin Cortana in Windows 10. Hannover: heise.de. URL: <http://www.heise.de/newsticker/meldung/Microsoft-integriert-Sprachassistentin-Cortana-in-Windows-10-2482133.html>

DMEurope (2014). Amazon cuts Fire phone to 99 cents, launches in Germany, UK. URL: <https://www.nexis.com/docview/getDocForCuiReq?oc=00240&lni=5D3K-BGM1-F0K1-N0HK&perma=true&csi=305444&secondRedirectIndicator=true>

Donath, Andreas (2015). Einstein soll Cortana intelligent machen. Berlin: golem.de. URL: <http://www.golem.de/news/microsoft-einstein-soll-cortana-intelligent-machen-1503-112957.html>

Dressler, Nadine Juliana (2015-07-17): Cortana für Android: jetzt gibt es Fotos und einen .APK-Leak. Berlin: WinFuture.de. URL: <http://winfuture.de/news,88067.html>

Focus.de (2015). Schlauere Siri und Multitasking - Das ist neu in iOS 9. FOCUS Online. URL: http://www.focus.de/digital/multimedia/technik-schlauere-siri-und-multitasking-das-ist-neu-in-ios-9_id_4954656.html

Go2android.de (2014). Sprachsteuerung: Was gibt es und wie funktioniert sie. URL: <http://www.go2android.de/sprachsteuerung-gibt-es-und-wie-funktioniert-sie/>

Google (2015). Micro-Moments Your Guide to Winning the Shift to Mobile. URL: <https://think.storage.googleapis.com/docs/micromoments-guide-to-winning-shift-to-mobile-download.pdf>

Gralla, P. (2014). Amazon Fire Phone: the missing manual (1. ed.). Sebastopol, CA [u.a.]:

O'Reilly Haughton, J. (2015). How Twitter chief's memo highlights personal responsibility at the top. Managers.org.uk. URL: <http://www.managers.org.uk/insights/news/2015/february/twitter-chiefs-memo-highlights-personal-responsibility-at-the-top>

Harms, I., Schweibenz, W. & Strobel, J (2002). Usability Evaluation von Web-Angeboten mit dem Web Usability Index. In Proceedings der 24. DGI-Online-Tagung 2002 - Content in Context. Frankfurt am Main 4.-6. Juni 2002 (S. 283-292). Frankfurt/Main.

- Horzont.net (2014). Firephone: Amazon-Chef glaubt weiter an geflopptes Smartphone. URL: <http://www.horizont.net/marketing/nachrichten/Firephone-Amazon-Chef-glaubt-weiter-an-geflopptes-Smartphone-131785>
- Lewandowski, D. (2011). Evaluierung von Suchmaschinen. In D. Lewandowski (Hrsg.), Handbuch Internet-Suchmaschinen 2: Neue Entwicklungen in der Web-Suche (S. 203–228). Heidelberg: Akademische Verlagsanstalt AKA.
- Luckerson, V. (2014). 4 Reasons Amazon's Fire Phone Was a Flop. Time.com. URL: <http://time.com/3536969/amazon-fire-phone-bust/>
- MacTechNews.de,. (2016). Apple speichert Siri-Spracheingaben zwei Jahre lang. URL: <http://www.mactechnews.de/news/article/Apple-speichert-Siri-Spracheingaben-zwei-Jahre-lang-155553.html>
- Maier, M. (2011). [IT] Allgemeine technische Informationen und mehr zu Siri und dem iPhone 4 - Ketchup-Mayo-Senf.de - Blog. Ketchup-mayo-senf.de. URL: <http://www.ketchup-mayo-senf.de/blog/-it-allgemeine-technische-informationen-und-mehr-zu-siri-und-dem-iphone-4>
- Martin-Jung, H. (2011). So funktioniert Siri. Sueddeutsche.de. URL: <http://www.sueddeutsche.de/digital/sprachsteuerung-am-computer-so-funktioniert-siri-1.1157482>
- Microsoft.com (2014). Microsoft News Center: Microsoft showcases latest updates to Windows, opportunities for developers. URL: <https://news.microsoft.com/2014/04/02/microsoft-showcases-latest-updates-to-windows-opportunities-for-developers/>
- Microsoft Corp. (2015,2016): Cortana - Android-Apps auf Google Play. Unterschleißheim: Microsoft Deutschland GmbH. URL: <https://play.google.com/store/apps/details?id=com.microsoft.cortana>
- Nielsen, J. (1994). Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Hrsg.), Usability Inspection Methods. John Wiley & Sons, New York, NY.
- Petras, V. (2013). Methoden für die Evaluation von Informationssystemen. In: Umlauf, Konrad (Hrsg.), Fühles-Ubach, Simone (Hrsg.), Seadle, Michael (Hrsg.): Handbuch Methoden der Bibliotheks- und Informationswissenschaft: Bibliotheks-, Benutzerforschung, Informationsanalyse, 2013, De Gruyter, Berlin.

- Porteck, Stefan (2015). Cortana künftig auch für Android und iOS. Hannover: heise.de. URL: <http://www.heise.de/newsticker/meldung/Cortana-kuenftig-auch-fuer-Android-und-iOS-2574978.html>
- Quirnbach, S. (2013). Suchmaschinen: User Experience, Usability und nutzerzentrierte Website-Gestaltung. Berlin: Springer.
- Saal, M. (2014). Kindle Unlimited: Amazon bringt E-Book-Flatrate für 10 Euro und muss Prügel fürs Fire Phone einstecken. Horizont.net. URL: <http://www.horizont.net/marketing/nachrichten/Kindle-Unlimited-Amazon-bringt-E-Book-Flatrate-und-muss-Pruegel-fuers-Fire-Phone-einstecken-130781>
- Sadauskas, A. (2014). Amazon Fire Phone first look: Gadget Watch. SmartCompany. URL: <http://www.smartcompany.com.au/technology/42943-amazon-fire-phone-first-look-gadget-watch.html>
- Sarodnick, F.; Brau, H. (2006). Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung. Bern: Huber.
- Schimanke, F. (2016). Flo's Weblog. Schimanke.com. URL: <http://schimanke.com/archives/3151-Erste-konkrete-Zahlen-zur-Nutzung-von-Siri.html>
- Schurwanz, T.; Schmidt, M.; Schmidt, F. et al (2015-12-10): Cortana: Microsoft veröffentlicht Sprachsteuerung für Android und iOS. Hamburg: COMPUTER BILD Digital GmbH. URL: <http://www.computerbild.de/artikel/cb-News-App-Check-Cortana-Android-iOS-11195534.html>
- Sebayang, Andreas (2015-11-05): Microsoft sucht Betatester für iOS-Version von Cortana. Berlin: Golem Media GmbH. URL: <http://www.golem.de/news/sprachassistent-microsoft-sucht-betatester-fuer-ios-version-von-cortana-1511-117311.html>; hrsg. am 5.11.2015 um 11:43
- Shitkova, M., Holler, J., Heide, T., Clever, N. & Becker, J (2015). Towards Usability Guidelines for Mobile Websites and Applications. Universität Münster. URL: <http://www.wi2015.uni-osnabrueck.de/Files/WI2015-D-14-00116.pdf>
- Singhal, A. (2013). Google Inside Search – The official Google Search blog: A multi-screen and conversational search experience. URL: <http://insidesearch.blogspot.de/2013/05/a-multi-screen-and-conversational.html>

- Statista (2016). Apple iOS 6: companies that power Siri | Statistic. URL:
<http://www.statista.com/statistics/235241/share-of-companies-that-power-siri-with-apples-ios-6/>
- Statista (2016). Apple iPhone - Absatz weltweit bis Q4 2015 | Statistik. URL:
<http://de.statista.com/statistik/daten/studie/12743/umfrage/absatz-von-apple-iphones-seit-dem-jahr-2007-nach-quartalen/>
- Sünkler, S. (2009). Vergleichende Analyse und szenariobasierte Usability Tests zur Evaluation der Gebrauchstauglichkeit neu gestarteter Suchmaschinen. Bachelorarbeit. HAW Hamburg. URL: http://edoc.sub.uni-hamburg.de/haw/volltexte/2009/891/pdf/Suenkler_Sebastian_090915.pdf
- Usability.gov (2006). Research-Based Web Design & Usability Guidelines. URL:
http://www.usability.gov/how-to-and-tools/resources/guidelines_book.pdf
- Wikipedia.org (2015). Amazon.com - Wikipedia. URL: <https://de.wikipedia.org/wiki/Amazon.com>
- Windows.com (2015). Windows Experience Blog: Cortana Now Available on Android in Public Beta. URL: <https://blogs.windows.com/windowsexperience/2015/08/24/cortana-now-available-on-android-in-public-beta/>
- Windows.com (2014). Windows Experience Blog: Cortana arrives in France, Italy, Germany, and Spain as a preview. URL:
<http://blogs.windows.com/windowsexperience/2014/12/05/cortana-arrives-in-france-italy-germany-and-spain-as-a-preview/>
- Windowsphone.com (2015). Hilfe und Anleitung: Cortana. URL:
<http://www.windowsphone.com/de-de/how-to/wp8/cortana/meet-cortana>
- Winfuture.de (2015). Xbox One: Großes Update bringt völlig neue Oberfläche und Cortana. URL: <http://winfuture.de/videos/Software/Xbox-One-Grosses-Update-bringt-voellig-neue-Oberflaeche-und-Cortana-14679.html>
- Winfuture.de (2015). Cortana soll E-Mails der Nutzer mitlesen um ihren Job zu machen. URL:
<http://winfuture.de/news,90740.html>
- Winslow, G. (2014). LaunchPAD: Fire Phone. Broadcastingcable.com. URL:
<http://search.ebscohost.com/login.aspx?direct=true&db=ufh&AN=96722955&lang=de&site=ehost-live>

Witold, Pryjda (2014-12-19): HERE: Microsoft übernimmt die Windows-Phone-App (Update).
Berlin: WinFuture.de. URL: <http://winfuture.de/news,85085.html>

Woods, P. (2016). Bestätigt: Siri nutzt Nuance-Technik. Macwelt web. URL:
<http://www.macwelt.de/news/Bestaetigt-Siri-nutzt-Nuance-Technik-7934468.html>

Youtube.com (2015). Comparing the amazon echo commercials to the crappy amazon fire
phone. URL: <https://www.youtube.com/watch?v=29-O2UjHPxA>

ZDNet.de (2015). Workspace: Cortana auf den Spuren von Siri und Google Now. URL:
<http://www.zdnet.de/88244128/cortana-auf-den-spuren-von-siri-und-google-now/>

ZDNet (2015). Microsoft integriert Kalender und E-Mail tiefer in Cortana. URL:
<http://www.zdnet.de/88257972/microsoft-integriert-kalender-und-e-mail-tiefer-in-cortana/>

Anlagen

Die folgenden Anlagen befinden sich auf der beiliegenden CD.

Anlage 1: Heuristik

Anlage 2: Personas

Anlage 3: Szenarien

Anlage 4: Fragebogen Vorabbefragung

Anlage 5: Testleitfaden

Anlage 6: Fragebogen Nachbefragung

Anlage 7: Suchprozess innerhalb einer optimalen Voice Search

