



# Mit Chatbots und Voice zum flexiblen, Kontext-unterstützenden Arbeiten in der digitalen Welt

Christian Ochsenkühn, OPITZ CONSULTING Deutschland GmbH

Dieser Artikel widmet sich einigen grundlegenden Fragen, die sich Unternehmen stellen, wenn sie im aktuellen Hype um Chatbots und/oder Sprachassistenten vorne mit dabei sein wollen: Wie funktioniert ein Chatbot? Warum sollten wir auf Chatbots setzen? Was bringt das unseren Nutzern?

Auf der Hauskonferenz „inspire|IT“ hat der Autor in diesem Jahr zum ersten Mal „Alexa Skill“ vorgestellt. Dazu wurde ein Tisch vorbereitet, auf dem ein Amazon Echo, ein Laptop und ein Poster mit einem sogenannten „Wimmelbild“ auf die Besucher warteten. Letzteres ist ein Bild, auf dem es vor Details nur so wimmelt; daher der Name. In diesem Fall zeigte es die unterschiedlichsten Projekte, die das Unternehmen des Autors bereits im Digitalisierungs-

kontext durchgeführt hat (siehe Abbildung 1). An dem Tisch konnten die Besucher nun von Alexa etwas zu diesen Projekten erfahren, etwa indem sie nach der Aufforderung „Erzähle mir etwas zum Flughafen“ anschließend die Frage stellten: „Welche Technik steckt dahinter?“. Alexa merkte sich in diesem Fall, dass wir zuvor über das Flughafen-Projekt gesprochen hatten und nannte die zugehörigen Informationen zu einigen technischen Details.

## Wie ist man innovativ?

Nach einiger Zeit kam ein älterer Herr im Anzug auf den Autor zu und ließ sich die verschiedenen Sprachbefehle erklären. Interessiert fragte er: „Wie hat das alles angefangen?“ Eine gute Frage. Der Autor wollte direkt loslegen und erklären, dass er einen ersten Prototyp in Python entwickelt hatte und später – als ihn ein Kollege bei der Entwicklung unterstützte



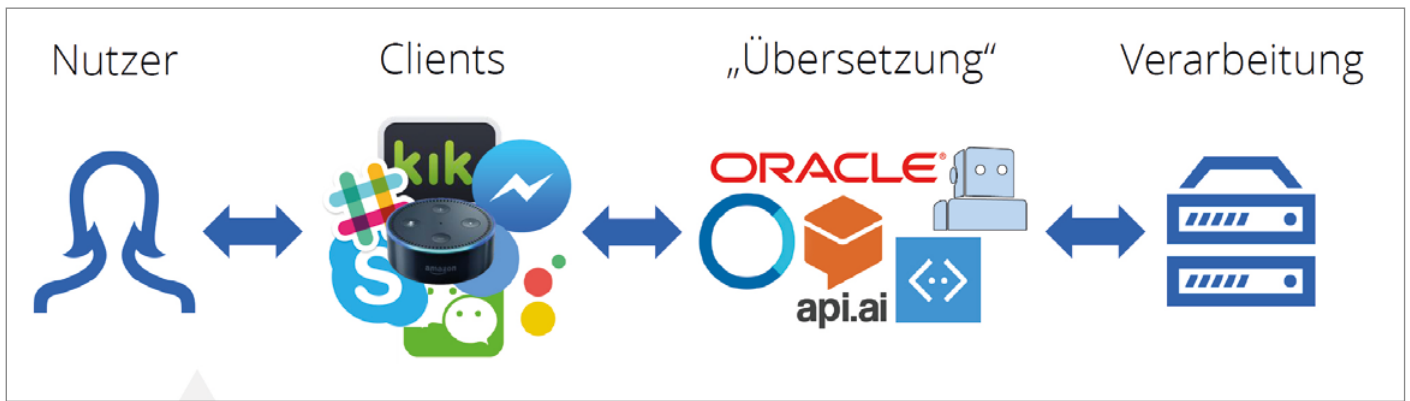


Abbildung 2: Die grundlegende Funktionsweise von Chatbots und Spracheingaben

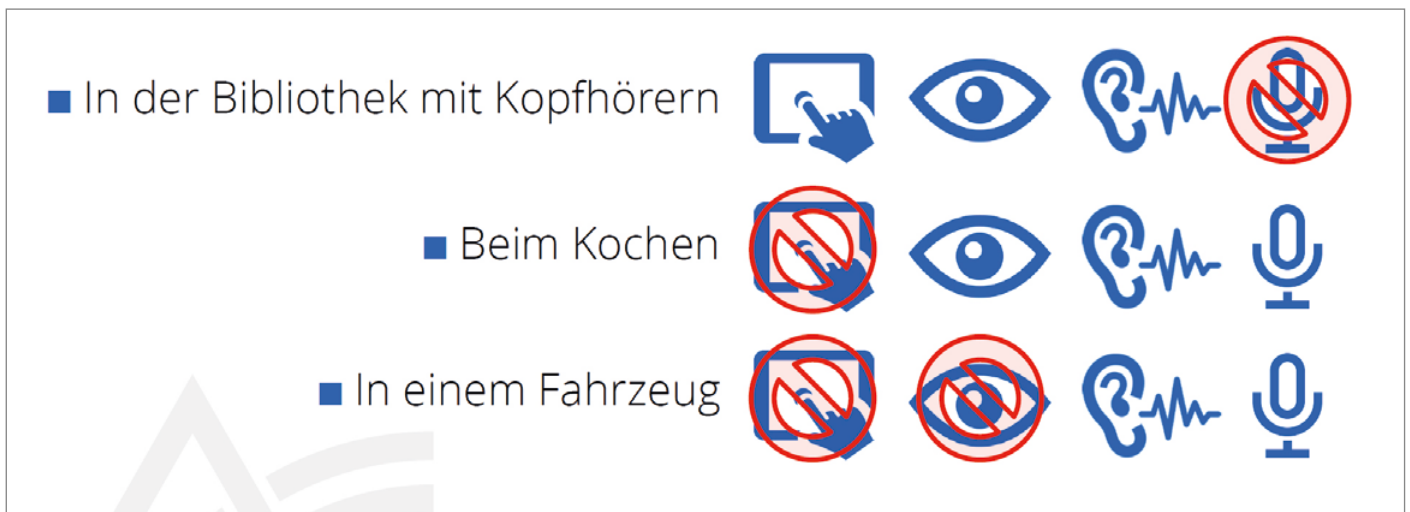


Abbildung 3: Je nach Kontext sind unterschiedliche Interaktionen (wie Spracheingabe) sinnvoll

Daraus sollen sie lernen und auch ähnliche Sätze dem richtigen Intent zuordnen.

Erst jetzt kommt von dort eine Anfrage beim Code des Chatbots an. Dies geschieht über eine Web-Schnittstelle. Man kann den Code also grundsätzlich dort hosten, wo man möchte. In der eingehenden Anfrage erhalten wir aber nicht etwa den gesprochenen oder geschriebenen Satz, sondern den zuvor erkannten Intent, also die Aktion, die ausgeführt werden soll. Zusätzlich gibt es noch einige Informationen aus dem Request sowie einen Status, eine User-ID oder bestimmte Worte aus dem initialen Satz, den man weiterverarbeiten will. Eventuell möchten wir zum Beispiel eine Zahl, ein Datum, einen Ort oder ein anderes bestimmtes Wort aus dem Satz verwenden, um eine bestimmte Aktion auszulösen.

Auf diesen Intent kann man nun reagieren und eine passende Antwort kreieren. Was unsere Aktion genau auslöst, hängt vollkommen vom Anwendungsfall ab. Der Bot kann direkt eine ganz einfache

Antwort zurückgeben, eine Audio-Datei abspielen, einen weiteren (Web-)Service ansprechen, der wiederum weiterführende Informationen bereitstellt, oder gar eine zusätzliche künstliche Intelligenz nutzen, um eine dynamisch generierte und wirklich intelligente Antwort liefern zu können. Egal welche Aktion wir auslösen, letztendlich wird dazu eine Antwort generiert, die wieder im Client und somit beim Nutzer landet. Der Kreis schließt sich.

### Sind Bots die neuen Apps?

Jetzt, wo wir wissen, wie Chatbots grundsätzlich funktionieren, stellen sich neue Fragen: „Brauchen wir überhaupt einen Bot?“ oder „Sollen wir ausschließlich Bots einsetzen, um unseren Kunden Mehrwert zu bieten?“ Die Antwort lautet: „Ja und nein“. Die Interaktion mit menschlicher Sprache bietet eine gute Möglichkeit, das Angebot schnell um einen weiteren Ser-

vice beziehungsweise ein weiteres Produkt zu erweitern, indem man bereits vorhandene Daten und Schnittstellen im Unternehmen nutzt und diese wiederverwendet. Jedoch ist nicht jeder Service dafür geeignet, um zum Beispiel per Sprache bedient zu werden. Vielmehr ist der Kontext wichtig, damit ein Chatbot einen Mehrwert zu bewährten Benutzeroberflächen schaffen kann.

Das klingt erst einmal plausibel, aber was sagt uns das? Bots sind nicht die neuen Apps, sondern eine On-Top-Erweiterung, die im richtigen Kontext eine zusätzliche und komfortablere Möglichkeit bietet, mit einem digitalen Produkt oder Service zu interagieren.

Haben wir bisher via Auge (Display), Händen (Tastatur, Touch) oder Ohren (Sound) interagiert, ist das Sprechen beziehungsweise Schreiben menschlicher Sprache der nächste logische Schritt in dieser Reihe. Die folgenden „Placeonas“ (Idee von [1]: Place + Persona) bieten einen Einblick in das Zusammenspiel von

Kontext und in diesem Fall Spracheingabe (siehe Abbildung 3).

Man stelle sich vor, man wäre in einer Bibliothek und hätte Kopfhörer auf (1). Dort kann man sein Smartphone mit Touch-Gesten bedienen, auf dem Display etwas sehen und durch die Kopfhörer auch auf Sprachausgabe hören. Laut seinem Gerät sollte man an diesem Ort aber eher nicht sprechen (2). Anders ist es, wenn man in einem Fahrzeug sitzt. Dort sollte man die Augen auf den Straßenverkehr richten und auch seine Hände am Lenkrad lassen, statt ein Touch-Gerät zu bedienen. Hören und vor allem das Sprechen von Befehlen ist hier absolut sinnvoll (3). Ähnlich ist es in der Küche, wo man vielleicht vom Teig verklebte Hände hat und daher lieber per Sprachbefehl einen Timer startet oder laut nach dem nächsten Schritt des Rezepts fragt.

Diese einfachen Beispiele lassen bereits erahnen, wohin die Reise geht. Für jeden Kontext gibt es das passende Mittel zur Interaktion. Lange Texte in sein Smartphone eingeben? Nein, danke. Den Text hingegen sprechen und in ein digitales Schriftstück umwandeln lassen, klingt wiederum sehr angenehm. Vorlesen lassen möchte man sich einen langen Text, ein Diagramm oder eine Tabelle nicht. Hier ergibt eine grafische Oberfläche am meisten Sinn, am besten auf einem größeren Screen als auf dem Smartphone. Das Ziel, dem uns Chatbots und vor allem die Spracheingabe einen großen Schritt näherbringen, muss somit sein, ein flexibles und Geräte-übergreifendes Arbeiten zu ermöglichen.

## Flexibles, Kontextunterstützendes Arbeiten in einer digitalen Welt

Die Interaktion mit einem Chatbot oder einem Sprachassistenten kann entweder so ablaufen, dass man einen Befehl absetzt und dieser direkt einen Prozess anstößt, der letzten Endes eine Lösung und eine Antwort bietet. Häufiger wird es aber vorkommen, dass der Befehl Teil eines größeren und Geräte-übergreifenden Ganzen ist. Der folgende fiktive Teil eines Arbeitstags soll diesen Mehrwert verdeutlichen: Ein Geschäftsmann sitzt morgens im Hotel beim Frühstück und bekommt eine E-Mail auf seine Smartwatch.

Da dort eingestellt ist, dass nur Mails mit einer hohen Priorität ankommen sollen, öffnet er sie, scrollt kurz durch den Text und verschafft sich einen Überblick.

Antworten wird er nicht sofort. Erstens wäre das auf der Smartwatch viel zu mühsam, sein Smartphone liegt noch oben im Hotelzimmer, und zweitens möchte er zuerst sein Frühstück genießen. Später fährt er mit dem Auto vom Hotel in sein Büro vor Ort und spricht währenddessen mit dem Chatbot: „Nimm das Excel-Sheet aus der heutigen E-Mail von Tobias und erstelle ein Diagramm nach Template drei.“

Der Bot erstellt das Diagramm anhand eines vorher definierten Templates und legt es in der Cloud für ihn ab. Wenn er dann im Büro ankommt, sind seine Informationen bereits erstellt. Er analysiert die Tabellen aus der E-Mail dort auf einem großen Bildschirm, um den besten Überblick zu haben. Anschließend macht er sich zu Fuß auf den Weg zum Kunden, wo er dieses Wissen und das Diagramm brauchen wird.

Unterwegs schickt er jenes Diagramm noch schnell per Smartphone an einen Kollegen, der später zum Kundentermin hinzustoßen wird und sich so vorher einen Überblick verschaffen kann. Zum Verschicken nutzt er den unternehmensinternen Chat, weil er seinen Kollegen darüber am schnellsten erreicht.

Dieses Beispiel zeigt eindrucksvoll, wie Chatbots eine weitere Lücke schließen, sodass man wirklich flexibel und vor allem passend zum jeweiligen Kontext mit digitalen Produkten und Services agieren kann.

## Fazit

Bots sind also nicht die neuen Apps, sondern eine logische Erweiterung der bisherigen Interaktionsmöglichkeiten mit digitalen Produkten und Services. Wo anfangs die Interaktion nur per Kommandozeile stattfand, folgte später die Maus. Sie ersetzte die reinen Textbefehle allerdings nicht (komplett). Ebenso wie heute Sprachassistenten hatte es die Maus bei ihrer Einführung nicht leicht und wurde von vielen nur als nettes Gimmick gesehen. So schrieb die PC Week im April 1984: „Mice are nice ideas, but of dubious value for business users.“ Wie das heute aussieht, wissen wir alle.

Ähnlich erging es dem Smartphone, als es mit seinen Touch-Gesten Einzug hielt. Apps, die ein digitales Feuerzeug simulierten – indem sich die Flamme passend zur Smartphone-Neigung bewegte – wirkten zunächst ja auch wirklich nur wie eine Spielerei. Doch zeigten sie schon eindrucksvoll, wohin die Reise gehen sollte. Heute nutzen wir Smartphones massiv im Alltag wie auch im Businesskontext. Verdrängt haben sie Laptops, Maus und Tastatur aber dennoch nicht, auch wenn wir diese mittlerweile deutlich weniger nutzen. Ebenso wird es mit Chatbots und Spracheingabe kommen. Sie werden sowohl Apps als auch grafische Oberflächen nicht ersetzen, sondern diesen Mix aus Interaktionsmöglichkeiten erweitern.

Der Autor hat große Hoffnung, dass Chatbots und Spracheingabe uns einen Schritt näher an eine digitale Welt bringen, in der das, was wir nicht gerne tun, automatisiert und das, was wir gerne tun, digital angereichert („augmented“) wird [2]. Dieser Satz verdeutlicht genau das, was Digitalisierung für ihn bedeutet und weshalb er auf die weitere digitale Zukunft mit freudiger Erwartung blickt.

## Weitere Informationen

- [1] D. Traynor, What voice UI is good for: <https://blog.intercom.com/benefits-of-voice-ui/>
- [2] U. O'Broin. Chatbots in the Enterprise – Not Just All Talk: <https://chatbotsmagazine.com/chatbots-in-the-enterprise-not-just-talk-de385ec6f2>



Christian Ochsenkühn  
christian.ochsenkuehn@opitz-consulting.com