

Möchten Sie Bananen? Dann sagen Sie Bananen!

Eines ist wohl ein-eindeutig: Wenn ich den Anrufer befrage, egal ob Kunde oder Interessent und meinetwegen durch eine Sprachmaschine: „Möchten Sie von unserem automatischen Kundendialogsystem Gebrauch machen oder möchten Sie zu einem Berater durchgestellt werden?“ wählen über 95% den menschlichen Berater!



Voice compass: ...um sich in der Folge über die Inkompetenz des Callcenter-Agenten zu beschweren. Ob über die Fach- oder Entscheidungskompetenz – das sei mal dahin gestellt.

Dennis Schottler im teilweise menschlichen Dialog mit Detlev Artelt, dem Herausgeber und Autor der Fachbuchreihe ‚voice compass‘ zum Thema: „Was können Sprachanwendungen heute leisten?“ Zweiundsechzig E-Mails, 4 Telefonate über mehrere Stunden und 1 Treffen auf einer Messe ergaben folgenden zusammengefassten ‚Dialog‘:

Schottler: Unsere Stimme oder im Englischen voice lässt uns sprechen. Die Sprache ist eine der wichtigsten Kommunikationsformen, die wir Menschen besitzen. Warum quälen wir uns in unserem Alltag dennoch mit Tasten und Ziffernblöcken, um mit Maschinen zu kommunizieren?

Artelt: Ihre Frage lässt mich unweigerlich etwas schmunzeln, denn sie stammt aus dem Vorwort meines Buches ‚voice compass‘. Ich sitze hier, das Headset am Ohr, das Diktiersystem aktiviert und spreche immer noch mit diesem Schmunzeln im Gesicht eine Antwort in meinen PC. Was soll ich sagen, für nur wenige 100€ kann heute jeder diese Technik ganz einfach nutzen. Meine Freunde, egal ob Lehrer, Berater oder Handwerker schauen mir gelegentlich dabei über die Schulter und sind

alle sehr schnell von der Einfachheit, wie auch von der Effizienz, begeistert. Wir alle sind mit der Tastatur aufgewachsen, haben es mit 10-Fingerschreiben sehr schwer und müssen trotzdem erst einmal mit eigenen Augen sehen, dass Spracherkennung funktioniert. Für mich hat sich dieses Thema bereits seit langem erledigt, da ich den Vorteil der Technologie für meine Bequemlichkeit nutze. Um z.B. eine Adresse in meine Navigation zu sprechen, eine Person mit meinem iPhone anzurufen oder, wie jetzt, um einen Text zu diktieren.

Schottler: Nun gut, Sie digitalisieren Ihr Gesprochenes mit Hilfe eines Spracherkennungssystems. OK! Was aber, wenn derartige Systeme genutzt werden, um den Dialog zu ersetzen? Kann Kunden damit geholfen werden? Kann man damit Support leisten? Kann man damit Kunden binden oder gar gewinnen? Laufen Unternehmen hier nicht Gefahr eher abzuschrecken statt zu begeistern?

Artelt: Ehrlich gesagt – es kommt darauf an. Sinnvoller Einsatz, sowie professionelles Design bei der Implementierung sind maßgeblich für die Akzeptanz des Nutzers. Ein völlig freier, natürlicher Dialog mit der Maschine ist auch heute nicht möglich, oder besser gesagt, wirtschaftlich sinnvoll, nicht in jedem Fall

möglich. Ein Dialog mit einer Frage und einer Antwort in einem ganz bestimmten Fachgebiet klappt wunderbar, aber ein philosophisches Gespräch mit dem Computer im Stil eines Dialoges von ‚Captain Kirk auf dem Raumschiff Enterprise‘ ist nach wie vor Fiktion. Es lassen sich aber ganz bestimmte Teilbereiche für die Gewinnung von Neukunden mit der Maschi-

**Ein völlig freier,
natürlicher Dialog mit der
Maschine ist auch heute
nicht möglich**

ne abwickeln. Eine der beliebtesten Lösungen ist ein Gewinnspiel mit einer automatischen Adresserfassung. Die Gewinnfrage, ob Apfel oder Birne, lässt sich einfach durch die Maschine abwickeln. Das machen Sprachcomputer bereits seit 15 Jahren! Die Adresserfassung ist schon etwas komplexer und lässt sich auch heute nur zu 80 bis 90% vollautomatisch abwickeln. Der Rest muss manuell erfasst werden. In diesem Bereich gibt es Dienstleister, die bereits die gesamte Technologie zur vollautomatischen Erfassung deutscher Adressen aufgebaut haben und Unternehmen diesen Service für unter 1€ je Adresse anbieten.

Oder denken Sie mal an das Kontoфон: Hier können Sie Ihre Bank mitten in der Nacht ▶

anrufen, um Ihren aktuellen Kontostand abzufragen, Überweisungen zu senden oder um neue Schecks zu bestellen. Funktioniert einwandfrei und wird rege genutzt.

Schottler: Überzeugt mich nicht wirklich! Wie dynamisch kann eine Dialogmaschine sein? Wenn jemand ein Callcenter sucht stellen wir als CallCenter Scout wesentliche Fragen zur Analyse des Vorhabens. Jede Antwort des Suchenden generiert die nächste Frage mit weiteren, vielfältigen Antwortoptionen. Leichteste Mathematik zeigt auf, dass bei nur vier Antwortoptionen aus zehn Fragen das Schneeballsystem einsetzt. Soll heißen: Es müssten mehrere Tausend Folgefragen und –Antworten vordefiniert werden. Ob dieser Aufwand in einem gesunden Verhältnis steht, ist fraglich.

Artelt: Das ist ein sehr schönes Beispiel, denn es zeigt sehr deutlich, wie dumm die Maschine eigentlich ist. All die notwendige Intelligenz, um selber Entscheidungen zu treffen, müssen der Maschine im Vorfeld beigebracht werden. Vier Antwortoptionen bei zehn Fragen ist für eine Spracherkennung überhaupt kein Problem. Der Aufwand in diesem Beispiel besteht darin, die Fragen in einer IT-Struktur abzubilden. Das geht am besten mit einer Datenbank oder einer XML-Datei, die als Grundlage für die Dialogsteuerung verwendet wird. Dieser Aufwand lohnt sich nur, wenn genügend viele Anrufer ein System nutzen, um damit einen Preisvorteil gegenüber einer manuellen Bearbeitung zu generieren. Soll heißen, dass ist absolut machbar, nur vermutlich für 40 Anrufe je Tag zu teuer. Ab wann sich das lohnt, lässt sich für uns als Berater schnell errechnen, wenn weitere Details zur Verfügung stehen. Sie können solche Dienste heute bereits bei einigen Hosting Anbietern erstellen lassen und zahlen dann neben der einmaligen Dienstleistung für die Generierung des Fragenkataloges im Weiteren nur noch für die Benutzung des Systems je Anruf. Das kann sich dann wiederum sehr schnell rechnen, da Sie keine teure Technik anschaf-

fen und betreuen müssen, sondern nur die Kosten für die Erstellung haben und danach viel günstiger als im Callcenter durch die Maschine arbeiten lassen.

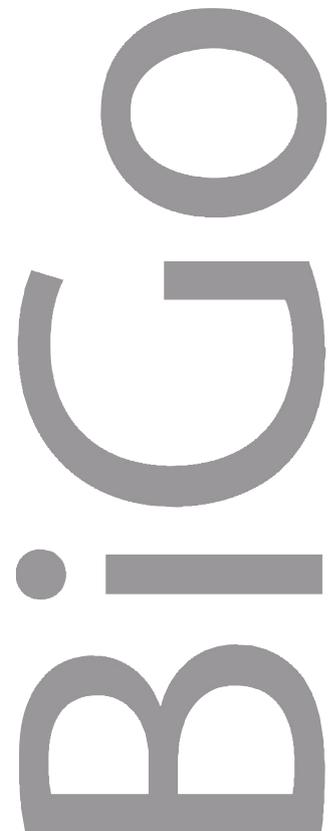
Schottler: Keiner scheut sich vor werthaltigen Techniken. Trotzdem bin ich der Meinung, dass der Konsument entscheidet, ob sich derartiges Führen von Dialogen durchsetzen wird. Neben dem Informationsaustausch soll Dialog ja auch Spaß machen und Türen öffnen, die vorher nicht absehbar waren.

Artelt: Türen öffnen! Da fällt mir was ein. Sie können heute ins Kino gehen und an der Kasse Ihre Handynummer hinterlassen. Zu Beginn des Filmes erhalten Sie den Hinweis: „Bitte lassen Sie Ihr Mobilfunkgerät während der ersten 10 Minuten der Vorstellung eingeschaltet.“ Der Film beginnt. Eine sichtlich nervöse, ängstliche Frau auf der Flucht – sie rennt durch die Hallen eines Fabrikgebäudes um ihr Leben, bis sie vor einer verschlossenen Tür steht mit dem Hinweisschild – „Nicht Öffnen, Explosionsgefahr!“. Die Spannung steigt an. Soll Sie die Tür öffnen, um dem Verfolger zu entkommen oder läuft sie zurück, ihrem Verfolger womöglich direkt in die Arme. Sie zückt ihr Handy und tippt eine Nummer. Im Kinosaal klingelt ein Handy. Die Frau aus dem Film fragt den Zuschauer: „Was soll ich tun?“ Jetzt entscheidet der angerufene Zuschauer über den Fortgang des Filmes durch seine Antwort.

Eine sichtlich nervöse, ängstliche Frau auf der Flucht

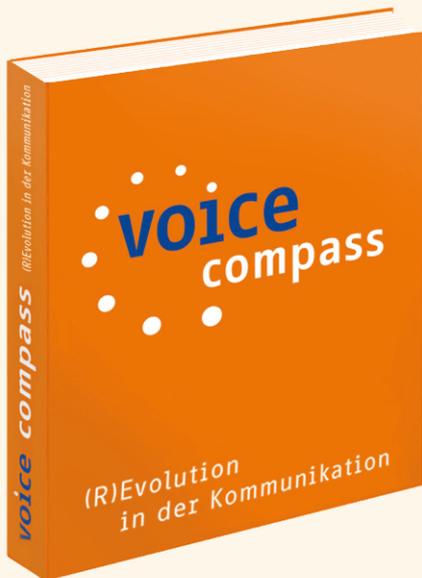
Schottler: Kracher! Das klingt wirklich spannend. Sprache in Nullen und Einsen zu wandeln und daraus Folgeprozesse einzuleiten funktioniert. Ja, das glaube ich! Wie steht es denn mit dem wirklichen Mehrwert der Stimme, die uns Aufschluss darüber gibt, ob das Gegenüber unsicher oder zielgenau agiert? Bleiben Emotionen auf der Strecke? Emotionen die dem Angerufenen die Möglichkeit geben, auf sein Gegenüber derart einzuwirken, dass er sich seiner wirklich annehmen kann.

Artelt: Sie werden es nicht glauben, aber das Erkennen von Emotionen ist in einem einfa-



Sie wollen weiter in die Materie einsteigen?

Der CallCenter Scout empfiehlt das Fachbuch:



Jetzt bestellen und sparen (inkl. SCOUT-Bonus)
Einfach auf der CallCenter Scout Seite des voice
compass bestellen:

<http://ccscout.voice-compass.com>

chen Rahmen heute schon möglich. Dazu sind in einigen Callcentern Sprachcomputer per Konferenz als stille Zuhörer in die Gespräche geschaltet. Diese erkennen Eskalationssituationen, wie z.B. einen auf einmal deutlich lauter sprechenden Kunden, oder sind in der Lage, das Alter der Stimme zu schätzen, um bei Minderjährigen einen Kauf eines bestimmten Produkts zu verhindern. Als Zusatzinformation auf dem Bildschirm des Agenten kann es teilweise sehr sinnvoll sein und helfen. Andere Systeme erkennen wie schnell oder wie sicher ein Benutzer mit einer Sprachanwendung umgeht und passen die weiteren Interaktionen daran an. So werden die Fragen bei wiederkehrenden Benutzern kürzer gefasst und einem schnell agierenden Benutzer werden Abkürzungen angeboten. Das beschleunigt die Abarbeitung, erfreut den ‚Power-User‘ und spart

Kosten. Reine Emotionen, die durch nur kleinste Veränderungen der Stimme von einem geschulten Ohr erkannt werden, gehen der Maschine eher verloren, denn die Maschine soll den Menschen ja nicht ersetzen sondern ihm nur die Arbeit leichter machen. So finden auch immer mehr Systeme ihren Einsatz, bei denen der Agent den Anrufer sozusagen ‚aufwärmt‘, um dann eine neutrale und immer gleich freundliche Befragung durch die Maschine abzuwickeln. Zu diesem Mensch-Maschine-Mix haben wir im voice compass ein ausführliches Kapitel.

Schottler: Wie hoch ist die Fehler- und Abbruchquote bei einem simplen Bestellprozess mit Ihrer Technik und welche Kosten entstehen dabei im Vergleich zu einem gut geschulten Callcenter-Agenten?

Artelt: Die meisten Fehler passieren nicht in der eigentlichen Spracherkennung, sondern in der Verarbeitung bzw. in der Vorbereitung des automatischen Bestellsystems. Wenn das Dialogdesign des Bestellprozesses gut gemacht ist, nämlich kundenfreundlich und zur Zielgruppe und dem Produkt passend, dann ist die Fehlerquote sehr gering (unter 5%). Nur dies bedeutet Aufwand und vor allen Dingen: Viel Nachdenken bevor das System gebaut wird.

Ich berate Unternehmen seit Anfang der 90er Jahre zum Einsatz von Sprachautomation und bekomme heute immer noch die gleichen Fragen und Ideen präsentiert, wie damals. Oftmals sind die Idee und das Ziel Kosten zu senken. Wenn dann klar wird, dass für eine erfolgreiche Sprachanwendung viel Vorarbeit geleistet werden muss, entscheiden sich einige Kunden doch für einfache und leider oftmals unfreundliche Systeme, die dann hohe Fehlerquoten haben. Dass sich die Automation auf der Kostenseite lohnt, haben uns nahezu alle Banken und Versicherungen vorgemacht, denn diese Branchen haben alle schon vor vielen Jahren Lösungen in Betrieb genommen. Immer wiederkehrende Prozesse, wie Nummerneingabe per Sprache oder Tasten und die Ausgabe von Informationen per Sprache, lassen sich wunderbar und preisgünstig automatisieren. Dadurch hat der Kunde schnell und egal zu welcher Uhrzeit eine Information und kann bei weiteren Fra-

gen ausführlicher mit dem Berater sprechen, wenn es nicht gerade nachts um 3:00 Uhr ist. Es liegt auf der Hand, dass ein solch automatisierter Prozess günstiger ist. Wie hoch die genaue Ersparnis ist, lässt sich pauschal nicht sagen, denn hier wirken verschiedene Faktoren, wie Anzahl der Anrufe, Komplexität des Bestellprozess und Anbindung an das Bezahlsystem bzw. die weitere IT, hinein.

Schottler: Ich habe meine Freundin bei der Buchung einer Reise am Telefon kennengelernt. Kann ich mich künftig via IVR verlieben?

Artelt: Naja, wenn Sie Ihre Freundin durch einen Sprachcomputer ersetzen möchten vielleicht, aber dies würde Einschränkungen in anderen Lebenslagen mit sich bringen. Ich bevorzuge ein Abendessen mit meiner Frau. Sie besteht darauf, dass ich mein Diktiersystem im Büro lasse.

Schottler: Ehrlich? Sie haben sich doch längst das App zum voice filing auf Ihr iPhone geladen...

Artelt: Stimmt ;-) und auch eine Spracherkennung, aber diese Apps ruhen, wenn ich mich entspanne, denn es gibt auch ein Leben ohne diese Technologien! ■



Detlev Artelt

Geschäftsführer aixvox GmbH
Aachen, Deutschland

IVR

Interactive Voice Response (IVR) - Ein System, das die Verbindung zu einem Computersystem per Telefon ermöglicht