

Die Welt – 05.11.2006

http://www.welt.de/webwelt/article92539/Wenn_der_Computer_spricht.html

Wenn der Computer spricht

PC-Programme erkennen die Stimmen menschlicher Anrufer. Gute Software bietet sogar von sich aus Produkte zum Kaufen an. Vielen deutschen Kunden ist das Gespräch mit Maschinen jedoch unangenehm.

Wenn Captain Kirk auf dem „Raumschiff Enterprise“ seinem Computer mündlich einen Befehl gibt, piepst der Rechner und zeigt damit, dass er verstanden hat. Auch im Science-Fiction-Film „2001: Odyssee im Weltraum“ sprechen Menschen mit Maschinen, manchmal sogar über Philosophie. Serie und Film entstanden Ende der 60er- und Anfang der 70er-Jahre. Mittlerweile haben Forscher die Filmideen in Software für Spracherkennung umgesetzt. Viele der aktuellen Systeme verstehen mehr als nur einzelne Ziffern oder Wörter, wenn jemand zum Beispiel seinen Kontostand abfragt oder Geld überweisen möchte. Neue Programme filtern Schlüsselwörter aus Sätzen heraus, lassen sich im Gespräch unterbrechen und erkennen Versprecher. Das gefällt nicht jedem Anwender.

Nach einer Untersuchung des Fraunhofer-Instituts und der Initiative Voice Business sehen drei Viertel aller Deutschen solche Sprachdialog-Systeme kritisch. Dabei sind sie schon seit Langem Wirklichkeit. „Viele Dinge, die vor zehn Jahren Forschungsgegenstand waren, sind heute kommerziell im Einsatz“, sagt Florian Gallwitz, Entwicklungsleiter bei Sympalog. Die Erlanger Firma erarbeitet Sprachanwendungen zum Beispiel für Autovermieter Sixt und Siemens.

Im vergangenen Jahr setzten Anbieter von Dialog-Systemen weltweit fast 350 Millionen US-Dollar um. Im Jahr 2010 sollen es nach einer Studie der Initiative Voice Business weltweit fast 1,1 Milliarden Dollar sein. 2005 investierten Telekommunikationsunternehmen 120 Millionen Dollar, um ihre Kundenkontakte zu automatisieren. Insgesamt gibt es nach Berechnungen der Initiative in Deutschland jeden Tag 38 Millionen Anrufe, bei denen Menschen mit einer Sprachsoftware der Servicestelle reden.

Für die Firmen rechnet sich der Einsatz automatisierter Dialogsysteme. T-Mobile, Deutschlands größter Mobilfunkanbieter, hat hierzulande etwa 30 Millionen Handykunden und muss für Abrechnungen oder Umstellungen monatlich mehr als fünf Millionen Fragen beantworten. Kein Callcenter der Welt könnte das bewältigen.

„Die Automatisierung von Prozessabläufen in Callcentern entwickelt sich zu einem entscheidenden Erfolgsfaktor für Unternehmen“, sagt Thiess Johannssen, Sprecher von D+S Europe. Die Firma betreibt in Deutschland acht Callcenter mit insgesamt 2000 Mitarbeitern. „Zwischen zehn und 15 Prozent der Anrufe könnten heute von einem Automaten erledigt werden“, schätzt Johannssen. Um Anfragen nach dem Kontostand zu bearbeiten, reiche eine Maschine. Die damit gewonnene Zeit könnten Mitarbeiter besser für intensive Kundenberatungen einsetzen.

Tatsächlich geht es darum, Kosten einzusparen. Zwischen 2,30 Euro und 2,50 Euro kostet T-Mobile ein Anruf an Personalkosten, wenn sich ein Callcenter-Mitarbeiter persönlich um eine Anfrage kümmert. Gerade einmal elf Cent sind es, wenn dies ein Dialogsystem übernimmt. Schon nach wenigen Monaten rechnet sich die Investition in ein Sprachdialog-System.

An solchen Programmen arbeiten Sprachwissenschaftler mit Neurobiologen und Technikern. Die Forscher untersuchen dazu den Verlauf der Tonhöhen, Dauer einzelner Laute, Sprechpausen, Dialekt oder Akzent, damit Maschinen Menschen verstehen können, wenn sie natürliche Sprache verwenden.

„Heutzutage wird Sprache am Computer bis auf die Klangelemente von

einzelnen Buchstaben heruntergebrochen. Das ermöglicht die Synthese natürlich klingender Sprachausgaben, die von einer menschlichen Stimme nicht mehr zu unterscheiden sind“, sagt Nuance-Sprecher Alexander Rosenwald. Die US-Firma Nuance entwickelt Software zur Spracherkennung, die andere Firmen für ihre Anwendungen nutzen können. „Die Spracherkennung haben wir im Griff, die Qualität der Erkennung ist sehr gut“, sagt Rosenwald. Dabei geht es nicht nur darum, Wörter und Sätze zu verstehen und entsprechend darauf zu reagieren.

T-Mobile zum Beispiel hat von etwa 2,5 Millionen Kunden Profile angelegt. Anhand der Telefonnummer erkennt der Computer den Anrufer und bietet ihm zuerst eine Wahlmöglichkeit an, die der Anrufer bei den Kontakten zuvor gewünscht hat. Das kann zum Beispiel eine Wettervorhersage für eine bestimmte Region sein. Langes Durchsprechen bis zum gewünschten Menüpunkt entfällt.

Ein interessantes System hat auch die Firma Semanticedge aus Berlin gezeigt, die schon dreimal den Voice Award gewonnen hat. Den Preis vergibt jährlich die Initiative Voice Business. In diesem Jahr zeichnete sie Semanticedge für einen Persönlichen Assistenten aus, den die Firma für DaimlerChrysler entwickelt hat. Das Sprachdialog-System kennt sich unter anderem im persönlichen Adressbuch seines Nutzers aus und hält Telefonnummern oder Adressen bereit.

Mittlerweile lassen auch Marktforschungsinstitute Umfragen vom Computer durchführen, Energieversorger nehmen Zählerstände und Versandhändler Bestellungen mit PC-Hilfe entgegen. Sprachautomaten sprechen Mahnungen aus, indem sie an fällige Zahlungen erinnern. Auch bieten Computer mittlerweile Produkte an. Rechner der Citibank zum Beispiel haben Kunden angerufen und Festgeld-Angebote unterbreitet. Bei Interesse des Kunden stellte der PC zu einem Mitarbeiter der Bank durch.

Ein weiteres Beispiel für moderne Sprachsoftware ist „Berti“. Mit dem automatisierten Ansagedienst zur Fußball-Bundesliga können Anrufer frei sprechen (siehe Interview links). Berti weiß zumindest, was seine Techniker ihm programmiert haben. Fragen zu Torschützen kann er nicht beantworten. „Das ist aber nur eine Frage der Wissensdatenbank“, sagt Florian Gallwitz von Sympalog. Würde sie entsprechend erweitert, könnte der Anrufer Berti auch nach dem Torschützen im vergangenen Spiel gegen Bayern München fragen und bekäme eine richtige Antwort.

Eine weitere Anwendung für Sprachsoftware ist Stimmenbiometrie. Die Software kann eine Person anhand ihrer Stimme identifizieren. Aus Sicherheitsgründen weisen viele Firmen ihre Mitarbeiter an, häufig das Passwort für das Firmennetzwerk zu wechseln. Dadurch vergessen es manche nach kurzer Zeit. Mit Stimmenerkennung ist das auszuschließen.

Stimmenbiometrie ist sogar noch zuverlässiger, als würden Fingerabdrücke verglichen. „Wir machten Versuche mit Elmar Brandt, der Gerhard Schröder wirklich perfekt imitieren kann“, sagt Hermann Geupel, Manager der Firma Voice Trust. Das System hat aber einen deutlichen Unterschied zwischen beiden Stimmen festgestellt.