

**Stuttgarter Zeitung 13.12.2006**

## **Mit Maschinen auf Du und Du - Spracherkennung im Alltag**

Ob Werbeanrufe, die von Computern getätigt werden, oder die Bedienung des Navigationssystems im Auto: Sprachdialogsysteme sind nicht mehr wegzudenken

Bald können Menschen den Computer per Sprache steuern. Und Ansagedienste können schon heute weit mehr als nur Ziffern herunterbeten. Dahinter steckt hochkomplizierte Spracherkennung. Ein Blick hinter die Kulissen und in die Zukunft.

Von Oliver Klempert

Keine Frage: es hat Charme, wenn Captain Kirk auf dem Raumschiff Enterprise seinem Computer mündlich einen Befehl gibt, dieser piepsend Verständnis signalisiert und den Befehl ausführt. Auch im Science-Fiction-Film "2001: Odyssee im Weltraum" sprechen Menschen mit Computern - dort ist die Maschine sogar in der Lage, eine philosophische Unterhaltung zu führen. Serie und Film entstanden Ende der 60er- und Anfang der 70er-Jahre. Aus der Vorstellung, sich eines Tages mit Maschinen unterhalten zu können, ist nun längst Wirklichkeit geworden.

Allerdings nehmen die Deutschen dies bisher kaum wahr. Drei von vier Befragten stehen einer Untersuchung zufolge, die das Fraunhofer-Institut gemeinsam mit der deutschen Initiative Voice Business durchgeführt hat, so genannten Sprachdialogsystemen kritisch gegenüber. Obwohl die Systeme längst mehr verstehen als nur einzelne Ziffern oder Wörter. Stand der Technik ist, dass die Dialogsysteme Schlüsselwörter aus Sätzen herausfiltern, sie lassen sich unterbrechen und erkennen sogar Versprecher.

Damit dies möglich ist, arbeiten Linguisten mit Neurobiologen und Technikern an immer besseren Dialogsystemen. Intonation, Lautdauer, Sprechpausen, Dialekt oder Akzent sind Punkte, die beim so genannten Dialogdesign beachtet werden müssen. Das Zauberwort heißt "natürliche Spracherkennung". Diese hat die Kommunikation zwischen Mensch und Maschine in den vergangenen Jahren deutlich weitergebracht.

"Heutzutage wird Sprache am Computer bis auf die Klanglaute von einzelnen Buchstaben heruntergebrochen. Das ermöglicht die Synthese natürlich klingender Sprachausgaben, die von einer menschlichen Stimme nicht mehr zu unterscheiden sind", sagt Alexander Rosenwald von Nuance. Das amerikanische Unternehmen hat eine Spracherkennungssoftware entwickelt, auf der viele Anwendungen anderer Firmen aufbauen. "Die Spracherkennung haben wir im Griff, die Qualität der Erkennung ist sehr gut."

Das hat in den vergangenen Jahren mehrfach die Firma **Semanticedge** aus Berlin gezeigt. Den so genannten Voice Award, der jedes Jahr von der Initiative Voice Business vergeben wird, hat das Unternehmen dreimal gewonnen. In diesem Jahr gewannen die Berliner mit dem "Persönlichen Assistenten", den sie für Daimler-Chrysler entwickelten. Dies ist ein Sprachdialogsystem, das sich im persönlichen Adressbuch des Nutzers auskennt. Vergessene Telefonnummern oder Adressen gehören der Vergangenheit an.

Auch sonst sind den Anwendungen keine Grenzen gesetzt: Marktforschungsinstitute lassen Umfragen durch Computer erstellen, Energieversorger nehmen automatisch Zählerstände und Versandhändler Bestellungen entgegen. Auch im Mahnwesen können die Sprachautomaten

eingesetzt werden, indem sie zum Beispiel an fällige Zahlungen erinnern. Auch unterbreiten Computer mittlerweile Angebote: So ließ die Citibank unlängst ihre Rechner bei ihren Kunden anrufen und ein Festgeldangebot unterbreiten. Bei Interesse konnten sich die Bankkunden dann gleich zu einem Mitarbeiter der Bank durchstellen lassen. "Viele Dinge, die vor zehn Jahren noch Gegenstand der Forschung waren, sind heute kommerziell im Einsatz", sagt Florian Gallwitz, Entwicklungsleiter bei Sympalog, einer Erlanger Firma, die Sprachanwendungen zum Beispiel für den Autovermieter Sixt und für die Firma Siemens erarbeitet hat.

Mit "Berti", einem automatisierten Ansagedienst rund um die erste Bundesliga, zeigt Sympalog, was mit modernen Sprachanwendungen möglich ist. Denn mit "Berti" können Nutzer völlig frei sprechen - auch wenn er nicht alles weiß. Fragen zu Torschützen kennt er nicht. "Das ist aber nur eine Frage der Wissensdatenbank", sagt Gallwitz. Würde sie erweitert werden, könnte der Nutzer Berti ebenfalls fragen: "Wer hat letzte Woche das Tor gegen Bayern München geschossen?" Gallwitz ist überzeugt: "Technisch sind wir schon sehr weit. Allerdings spiegeln dies viele Systeme, die derzeit noch am Markt sind, nicht wider."

Ebenfalls längst Realität ist auch die Stimmenbiometrie. Im Gegensatz zum Fingerabdruck ist sie viel sicherer, denn in ihrer Komplexität und Dynamik sind Stimme und Sprache eines jeden Menschen einmalig. Sie werden durch die Größe der Mundhöhle, des Kiefers, der Zunge und des Gaumens bestimmt. "Wir machten Versuche mit Elmar Brandt, der Gerhard Schröder perfekt imitieren kann", sagt Hermann Geupel von der Firma Voice Trust aus München, "aber unser System stellte einen deutlichen Unterschied zwischen beiden Stimmen fest." Einsatzmöglichkeit: viele Firmen verschärfen seit einigen Jahren die Passwortregeln. Das führt dazu, dass die Mitarbeiter häufiger ihr Passwort vergessen. Mit Stimmenerkennung können diese schnell und unkompliziert wiederhergestellt werden. "Das spart viel Zeit", sagt Geupel. Die Allianz und die Hypo-Vereinsbank nutzen das System seit einigen Jahren.

Doch neue Technik soll noch mehr können. Ein Beispiel: "Einmal Terminator 3 bitte." Höflich fragt der zehnjährige Junge beim automatischen Ansage- und Kartendienst nach einer Karte für den Actionstreifen im Kino nach. Doch der Computer verweigert den weiteren Dienst: "Ich denke, dafür bist du noch zu jung. Wie wäre es stattdessen mit Pipi Langstrumpf?" fragt die Maschine zurück. Was wie eine Zukunftsvision klingt, soll vom nächsten Jahr an bei der Telekom möglich sein - Alterserkennung am Telefon. Ein solches Projekt wird derzeit bei T-Systems gemeinsam mit der TU Berlin entwickelt. Fragt eine ältere Person etwa das Kinoprogramm ab, so werden zunächst Klassiker aufgezählt, sind Kinder am Telefon, gibt es als erste Option Kinderfilme. Außerdem werden ältere Menschen eher an einen Mitarbeiter im Call-Center durchgestellt, jüngere müssen länger mit dem Automaten sprechen. "Mit intelligenten Gesprächsführungen wollen wir den Komfort und die Nutzerfreundlichkeit steigern", sagt der Projektleiter Florian Metzke, der an der Berliner Universität für die Telekom das Projekt leitet.

Die neue Interaktion wird nach Ansicht vieler Forscher in den kommenden Jahren eine Flut neuer Möglichkeiten ergeben. Die Kaffeemaschine könnte schon bald fragen, ob es Espresso oder Cappuccino sein soll. Auch wird per Spracheingabe das Umschalten des TV-Programms möglich oder Videotext vorgelesen. Selbst Lernspielzeuge für Kinder sind bereits angedacht: Kinder können Fragen stellen oder Anweisungen geben, worauf das Spielgerät mit einer Aktion reagiert.

Auch beim klassischen PC werde sich einiges ändern. So beginne demnächst die Post-PC-Ära, sagt Wolfgang **Wahlster**, Leiter des deutschen Forschungszentrums für künstliche Intelligenz in Saarbrücken. "Mobile Geräte, die ohne Maus und Tastatur auskommen müssen, werden alle menschlichen Sinne nutzen, um mit dem Menschen zu kommunizieren." Deutschland stünde

bei dieser Entwicklung weit vorn. "Wir haben rund 120 kleine und mittelständische Firmen in Deutschland, die sich damit auseinander setzen." Sprachliche Regelungen im Auto können sogar die Sicherheit erhöhen. Per Zuruf wird künftig die Lautstärke des Radios oder die Heizung reguliert. Die Hände können am Steuer bleiben. An einem System, das darüber hinausgeht, arbeitet derzeit das Fraunhofer-Institut für Rechnerarchitektur und Softwaretechnik. Unter dem Begriff "Smart Web-Kfz" soll es schon bald für Autofahrer möglich sein, während der Fahrt Informationen aus dem Internet abzufragen, Beispiel: "Wo finde ich die nächste Pizzeria?" Die im Internet gefundenen Antworten werden in ganzen gesprochenen Sätzen ausgegeben.

Weil eine direkte Kommunikation mit dem Internet per Sprache nicht möglich ist, behilft sich die neue Kommunikation hinter dem Steuer digitaler Radiosignale. Sobald das Auto den "Smart Web"-Empfangsbereich eines Radiosenders erreicht, erscheint auf einem Touchscreen ein Dialogsymbol mit Menüpunkten, die zur Verfügung stehen. Dann kann der Fahrer per Spracheingabe sich die nächste Tankstelle, die Bundesligaergebnisse oder Staumeldungen ausgeben lassen.

Ein weiteres Projekt, das zu den Olympischen Spielen 2008 in Peking einsatzbereit sein soll, heißt "Compass 2008". Erstmals soll ein Handy mit Übersetzungssystem für gesprochene Sprache auf den Markt kommen. Wenn man etwa im Restaurant sitzt, spricht man die Bestellung auf Deutsch ins Handy, das dann alles ins Chinesische übersetzt. Aus dem Lautsprecher gibt eine elektronische Stimme die Bestellung an den Keller weiter.

Navigation im Auto wird künftig per Sprachbefehl gesteuert. Foto dpa

1144183, STZ , 13.12.06; Words: 1217, NO: STZ-20061213-10-16