



[Home](#) >> [Presse](#) >> [IHK WirtschaftsForum](#) >> [Januar 2002](#) >> [Sprachtechnologie](#)

Sprachtechnologie - mehr als Spracherkennung

Der Markt für Sprachtechnologie wird in den nächsten Jahren dynamisch wachsen. Allein im Bereich Sprachbearbeitungssoftware wird bis 2005 europaweit mit einem Umsatzanstieg um etwa 50 Prozent auf über 1,6 Milliarden Euro gerechnet. Dies ist das Ergebnis der Studie ¹Technologie-Trends: Sprachtechnologie¹, die von der DZ Bank erarbeitet wurde.

Sprachtechnologie wird mittelfristig in alle Bereiche des täglichen Lebens ihren Einzug halten. Neben dem technischen Fortschritt sind nach Einschätzung der DZ Bank vor allem die Kostenreduktion sowie die komfortable Bedienung der Endgeräte wesentliche Faktoren für eine zunehmende Akzeptanz der Sprachtechnologie. Für Unternehmen, die sich in dieser Branche erfolgreich positionieren möchten, bedeutet diese Entwicklung neben den erwarteten Umsatzchancen aber auch erhöhte Finanzierungs- und Investitionsanforderungen.

Unter Sprachtechnologie ("Human Language Technology", HLT) versteht man die Anwendung der Kenntnis von Sprache auf die Entwicklung von Computersystemen, die alle Formen menschlicher Sprache erkennen, verstehen, interpretieren und generieren können. Ziel ist es, den zunehmenden Informationsaustausch effizienter zu gestalten, neue Möglichkeiten der internationalen Kommunikation anzubieten sowie die Schnittstelle zwischen Mensch und Maschine zu optimieren. Vision der Entwickler ist dabei der weltweite, natürliche (sprachgesteuerte) Zugriff auf das gesamte digital gespeicherte Wissen - und das für jedermann, zu jeder Zeit und an jedem Ort. Mit ihrer Hilfe erfolgt ein Zugriff von Computersystemen auf gesprochene Sprache, geschriebene Texte oder Informationen in Datenbanken.

Die Anfänge der Sprachtechnologie reichen bis in die fünfziger Jahre zurück und haben ihren Ursprung in den USA. Ähnlich wie das Internet verdankt sie ihre Existenz dem Kalten Krieg. Parallel zu sprachwissenschaftlichen Fortschritten nahm sie mit der Entwicklung der Computertechnik seit den siebziger Jahren ihren Lauf. Erste kommerzielle Produkte enttäuschten und ernüchterten die euphorischen Erstanwender jedoch aufgrund ihrer noch unzureichenden Leistungsfähigkeit.

Dem Computer fehlen elementare Voraussetzungen, um mit dem Menschen kommunizieren zu können: Ihm fehlt das menschliche Vokabular, er beherrscht auch nicht die Grammatik. Besonders problematisch für ein echtes Verständnis ist ferner seine völlige Unkenntnis über sämtliche menschlichen Aspekte und Lebensbereiche. Für die Kommunikation mit dem Menschen steht ihm einzig der direkte Kanal "Stimme" zur Verfügung. Weite Bereiche der menschlichen Kommunikation sind aufgrund ihrer enormen Komplexität daher nach wie vor einer Lösung mit Hilfe des Computers verschlossen. Computerlinguistik macht aus heutiger Sicht deshalb nur dort Sinn, wo ein limitierter Wortschatz mit klar definierter Bedeutung genutzt wird. Maschinelle Anwendungen schneiden damit lediglich einen Teilaspekt aus dem "Gesamtkomplex Sprache" heraus. Trotz dieser Restriktionen gibt es eine Reihe interessanter Anwendungen, die sich in naher Zukunft rasch durchsetzen werden.

Am bekanntesten ist derzeit die reine Spracherkennung in so genannten Speech-to-Text-Anwendungen (Diktiersysteme). In einem ersten Schritt leistet das System dabei die eigentliche Spracherkennung; es

"übersetzt" sprachliche Information in Befehle. Bei der Erkennung gesprochener Sprache muss das System zunächst das akustische Signal in verarbeitbare digitale Information umsetzen. Anschließend erfolgt ein Vergleich mit bekannten (also gespeicherten) Sprachmustern. Dabei wird die Sprache nicht in einem tieferen Sinne "verstanden", sondern aufgrund von Wahrscheinlichkeitsmodellen (in der Regel Hidden-Markov-Modelle) einem Befehl zugeordnet. Ein Spracherkennungssystem besitzt einen großen Vorrat an derartigen Referenzmustern, die somit als "Schablone" für die akustisch aufgenommenen Wörter dienen.

Diese Systeme mussten in der Vergangenheit in einem langwierigen Prozess auf den jeweiligen Anwender trainiert werden. Dennoch blieben die Erkennungsraten meist mit 50 bis 70 Prozent unerwünscht niedrig. Heutige hochwertige Systeme erreichen Erkennungsraten von weit über 90 Prozent, sind oft sprecherunabhängig ausgelegt und erfordern nur noch minimale Trainingszeiten. Der erkannte Vokabularumfang hat sich ferner von anfänglich etwa 700 Wörtern innerhalb von 15 Jahren auf über 700 000 Wörter erhöht. Gleichzeitig sind die Preise für die Endgeräte drastisch gesunken.

Die bekannteste Anwendung für die Ausgabe natürlicher Sprache (Sprachsynthese) dürften Navigationssysteme sein. Die Sprachsynthese erfolgt anhand von ausgefüllten Schablonen durch das gleichzeitige Abspielen von zuvor aufgezeichneten Aufnahmen oder verkettenden Spracheinheiten (Phoneme, Worte). Generierte Sprache muss Aspekte wie Intensität, Dauer und Betonung berücksichtigen, um ein kontinuierliches, natürliches Resultat hervorzubringen.

Im Zusammenspiel von Spracherkennungs- und Sprachausgabeeinheiten arbeiten so genannte Dialogsysteme; Sprachcomputer sind wohl die bekannteste Form des Mensch-Maschine-Dialogs. Bekannte Anwendungsgebiete sind beispielsweise Börsen-, Kino- oder Bahninformationen. Diese für einen möglichen Gesprächsverlauf vorstrukturierten, einfachen Systeme werden zukünftig gemischt initiativ geführte Dialoge zulassen und damit bis zu einem gewissen Grad einen "echten" Dialog auf Basis von Metawissen der Maschine ermöglichen.

Größte Herausforderung bleibt auch bei diesen Systemen die Multilingualität. Übersetzungsprogramme können nach augenblicklichem Entwicklungsstand nur "Rohübersetzungen" von längeren Textpassagen liefern, die von Menschen mit der entsprechenden Sprachkompetenz noch nachgearbeitet werden müssen. Hier ist auch mittelfristig nur eine langsame Qualitätssteigerung zu erwarten. Weitere textbasierte Systeme sind Internetsuchmaschinen, Dokumentenmanagement-Systeme oder auch Systeme zur Bearbeitung von E-Mails (so genannte E-Mail Response Management Tools). Allen sprachtechnologischen Anwendungen gemeinsam ist das Ziel einer deutlichen Zeit- und Kostenersparnis sowie einer Reduktion von Komplexität.

Durch die Fortschritte in der Computertechnik, Sprachwissenschaft und Sprachtechnologie wurde ein gewisses Niveau bei HLT-Basistechnologien erreicht. Auf der CEBIT 2001 tauchten daher auch erstmals fertige Endlösungen für die Massenfertigung (beispielsweise Videorecorder mit Sprachsteuerung, sprachgesteuertes TV-Gerät, Handhelds) auf. Die Verbreitung sprachtechnologischer Lösungen wird sich nicht zuletzt aufgrund zunehmender Kundenakzeptanz, fortschreitender Globalisierung, ansteigender Informationsflut und einer Zunahme mobiler Endgeräte in den nächsten Jahren fortsetzen.

So prognostiziert die Marktforschungsgesellschaft Frost & Sullivan ein Wachstum des europäischen Marktvolumens für Sprachverarbeitungssoftware von 250 Millionen US-Dollar in 2000 auf über 1,6 Milliarden US-Dollar in 2005, was einem jährlichen Wachstum von etwa 50 Prozent entspricht. Damit ist der europäische Markt deutlich dynamischer als der US-amerikanische. Hiervon und vom Trend zu höherwertigen Anwendungen profitieren die europäischen Anbieter.

Nach den Anschlägen vom 11. September 2001 ist jedoch davon auszugehen, dass auch die HLT-Anwendungen unter der allgemeinen Kürzung der IT-Budgets in den Unternehmen leiden werden. Sprachtechnologie wird sich aber insbesondere in den Branchen durchsetzen, die ihren Kundenkontakt primär über das Internet beziehungsweise Telefon führen (zum Beispiel Call Center Industrie). Hier liefert sie neben qualitativen Verbesserungen auch einen direkt quantifizierbaren hohen Nutzen.

Der Einzelne wird sich mittelfristig den durch Sprachtechnologie herbeigeführten Veränderungen sowohl in seiner Arbeitswelt als auch im Privatbereich nicht mehr verschließen können. Die Unternehmen werden Routineaufgaben zukünftig aus Effizienzgründen vollautomatisch - oder bei komplexeren Aufgabenstellungen zumindest teilautomatisiert - von Sprachsystemen abwickeln lassen. Der Kontakt zwischen Unternehmen und Kunden wird durch die Verbreitung existierender Technologien wie auch durch neuartige Anwendungen in Verbindung mit Internet und mobilen Multimedia-Diensten geprägt sein.

Ralf Raebel
DZ Bank, Frankfurt

weitere Informationen: [Redaktion IHK WirtschaftsForum](#)

IHK WirtschaftsForum
Januar 2002

Bookmarks

Quelle: Studienreihe Technologie-Trends: Sprachtechnologie.
Herausgegeben von der DZ Bank, Frankfurt. Autoren: Ralf Raebel,
Dr. Jan Wessels.

Die 76-seitige Studie kann als PDF-Datei angefordert werden
bei [Karina Brands](#), DZ Bank

Für die Richtigkeit der in dieser Website enthaltenen Angaben können wir trotz sorgfältiger Prüfung keine Gewähr übernehmen.