

SPRACHERKENNUNG ERLEICHTERT DIE TÄGLICHE ARBEIT

MARTIN MÜLLER, LEITER DER 4VOICE-SOFTWARE-ENTWICKLUNG, IM GESPRÄCH ÜBER SPRACHERKENNUNG UND DIE NEUE VOICE4XY VERSION 2.

Der Diplom-Informatiker Martin Müller leitet die Software-Entwicklungsabteilung der 4voice AG. Seit 2001 beschäftigt er sich mit Spracherkennung und nutzt sein Know-How seit 2003 als Chefentwickler für die 4voice AG. Zu seinen Aufgaben gehören die Weiterentwicklung aller Produkte der 4voice AG sowie die Spezifikation und Implementierung von Schnittstellen zu Fremdsystemen, aber auch Tests zu neuen Funktionalitäten und Einsatzmöglichkeiten von Spracherkennung.

Welche Schwierigkeiten oder Anforderungen stellen sich Ihrer Meinung nach heute an moderne Spracherkennung?

Die größte Herausforderung an moderne Spracherkennung stellt nach wie vor der Anwender dar. Dass Spracherkennung an sich funktioniert, ist seit längerem bewiesen, aber leider wird häufig übersehen, dass Spracherkennung nicht als Selbstzweck betrieben wird, sondern sich die Anwender ihre tägliche Arbeit damit erleichtern wollen. Dieses Tagesgeschäft besteht aber häufig nicht darin, einen Computer zu bedienen, so dass die Anwender einen möglichst einfachen und integrierten Einsatz von Spracherkennung benötigen. Die funktionierende Spracherkennung ist dabei schlicht als Grundvoraussetzung zu sehen, die Integration in die bestehende IT-Landschaft und den bestehenden Workflow ist das primäre Ziel, damit die Anwender mit der Gesamtlösung zufrieden sind.

Was ist der Hintergrund für die neue Version 2?

Uns sind die vorgedruckten CD-Aufkleber für Version 1 ausgegangen (lacht). Für Version 1.x wurden einige Erweiterungen und neue Funktionalitäten entwickelt, die zunächst nur kundenspezifisch waren. Etliche dieser Erweiterungen wurden nun in den regulären Leistungsumfang der Produkte integriert. Zudem sollten die nach und nach entstandenen Komponenten der Version 1.x auf eine gemeinsame Basis gestellt und konsolidiert werden.

Was ist das Besondere an der neuen Version?

Die vielleicht auffallendste Änderung ist die Verschmelzung von reinem digitalem Diktatmanagement und serverbasierter Umsetzung. Wo bei Version 1 noch zwei getrennte Produkte (voice4xy dictate und voice4xy server) vorlagen und sich der Anwender im Prinzip bei der Installation für den einen oder anderen Weg entscheiden musste, gibt es mit Version 2 einen fließenden Übergang. Für jeden einzelnen Sprecher kann somit festgelegt werden, ob er mit Spracherkennung arbeiten möchte oder lediglich digitales Diktatmanagement mit manueller Transkription nutzen will.

Daneben sind etliche Ergänzungen dazugekommen, wie beispielsweise die Möglichkeit, jedem Diktat einen gesprochenen Bearbeitungshinweis oder beliebige andere Dateien als Anhang mitzugeben, oder die Möglichkeit automatischer Updateverteilung bei Großinstallationen.

Welche Vorteile bietet v4m V2 gegenüber anderen Spracherkennungssystemen?

Ein großer Vorteil ist sicherlich die Möglichkeit der „sanften“ Einführung von Spracherkennung. Je nach Wunsch und Kenntnisstand des Benutzers bieten wir mit voice4medicine Lösungen, die am einen Ende der Skala komplett ohne Spracherkennung arbeiten und

„nur“ digitales Diktatmanagement bieten, über die serverbasierte Umsetzung oder die Korrektur durch eine Schreibkraft, bis hin zu voice4medicine local, wo der Sprecher selbst die kompletten



Funktionalitäten der Spracherkennung an seinem eigenen Rechner nutzt. Dennoch verfügen alle Komponenten über dasselbe Look&Feel. Ein Benutzer, der bspw. von digitalem Diktatmanagement zu serverbasierter Umsetzung wechseln möchte, muss sich nicht umstellen.

Wie sieht die Zukunft der Sprachverarbeitung aus?

Nachdem die Verwendung analoger Bänder mehr und mehr ausstirbt erwarte ich, dass der Einsatz digitaler Medien auch immer natürlicher erfolgt. Genauso selbstverständlich wie heutzutage Tastatur und Maus wird in Zukunft wohl auch ein Mikrofon zur Standardausstattung von PCs gehören, zusammen mit Spracherkennung, die ein zusätzliches Eingabemedium darstellt. Auch die Verbreitung auf mobilen Geräten dürfte zunehmen, so dass die Grenzen zwischen den Arbeitsweisen mehr und mehr verschwimmen. Für die jeweils gerade anstehende Aufgabe wird einfach das effizienteste Arbeitsmittel gewählt.