

European Media
Laboratory GmbH



ZUR SPRACHE GEBRACHT!

Automatische Transkription

Callcenter-Analyse

Automatische Übersetzung

Sprachsteuerung im Auto

Echtzeit

Neuronale Netzwerke

Maschinelles Lernen

**„BEST-PRACTICE“-BEISPIELE FÜR DEN EINSATZ
VON EML-SPRACHTECHNOLOGIE**

TELEFONIE

Branche: Telefonanlagen / Kundenkommunikation
Partner: ADDIX Software GmbH, Kiel



Mit verschrifteten Sprachnachrichten zu mehr Effizienz im Unternehmen

Wie eine Telefonanlagenfunktion ein Unternehmen veränderte

Es ist 17.00 Uhr am Freitag, die Mitarbeiter eines Sanitätsversandfachhandels aus Berlin machen Feiertag. Der Anrufbeantworter wird eingeschaltet, und die Mitarbeiter gehen ins Wochenende. Am darauffolgenden Montag die Szene, die sich tagtäglich wiederholt: Ein Mitarbeiter sitzt mit Zettel und Stift am Anrufbeantworter und hört diesen ab. Notiert die Bestellwünsche der Kunden und verteilt anschließend die Zettel an die Kollegen im Vertrieb. Je nachdem wie ausgelastet der Vertrieb ist, kann es einen halben Tag dauern, bis die Zettel am richtigen Schreibtisch angekommen sind. Es wird anschließend der Inhalt erfasst und die notwendigen Bestellungen im System für die jeweiligen Kunden hinterlegt. Ein langwieriger und fehlerbehafteter Weg, der Personalressourcen bindet.

Der Sanitätsfachhandel ist seit einigen Jahren Kunde der ADDIX in Kiel. Bei einem Workshop zur Optimierung des Callflows innerhalb des Unternehmens wurde fast per Zufall bekannt, wie das Unternehmen bisher mit den in die Astimax Telefonanlage integrierten Voicemailboxen umging. Um den Prozess des Abhörens von Voicemailboxen zu beschleunigen, schlugen die Kieler dem Kunden vor, die eingehenden Anrufe auf dem Anrufbeantworter automatisch zu verschriften.

Warum? Lesen ist schneller als Hören: Ungefähr sechs Sekunden dauert es, eine Sprachnachricht von einer Minute zu lesen – ein klarer Vorteil für das Unternehmen. Seit der Einführung der verschrifteten Mailboxen werden die darauf gesprochenen Bestellungen automatisch verschriftet und per Mausclick

an den zuständigen Mitarbeiter als E-Mail weitergeleitet. Dieser muss dann die notwendigen Angaben für den Bestellprozess nur noch aus der E-Mail herauskopieren, statt vorher mühsam abzutippen. Die telefonische Bestellung des Kunden kann inklusive Voicefile im System archiviert werden.

Die Verschriftung der Mailboxen der Astimax Telefonanlagen funktioniert dank der Zusammenarbeit von ADDIX und der EML European Media Laboratory GmbH. Das Heidelberger Unternehmen verschlüsselt die Sprachdaten und verarbeitet sie ausschließlich auf Servern am Standort Deutschland - damit befolgt es die strengen deutschen Datenschutzgesetze. Die hohen Erkennungsraten sorgen für eine reibungslose Verschriftung. Spezielle Fachtermini, die das Unternehmen häufig nutzt, wurden im System durch das Astimax Team ergänzt. Der Wortschatz kann so auch an aktuelle Änderungen wie beispielsweise neue Produktnamen angepasst werden.

„Mit der Verschriftung von Voicemailboxen wird den Anwendern ein realer Mehrwert mit messbaren Einsparpotenzialen geboten. Gerade solche Beispiele machen deutlich, welchen Beitrag die Verschriftungstechnologie zur Digitalisierung leisten kann“, so EML-Forschungs- und Entwicklungsleiter Dr. Siegfried Kunzmann. Die Verschriftung ist auch optionaler Bestandteil der Astimax UC Lösung, die 2016 als Produkt auf den Markt gekommen ist.

ADDIX Software GmbH, Kiel

Die ADDIX Software GmbH hat sich als Softwareentwicklungsfirma auf die Entwicklung und Herstellung von Telekommunikationsprodukten spezialisiert. Die von ADDIX entwickelte professionelle Voice over IP Telefonanlage Astimax wird seit 2005 international erfolgreich eingesetzt und bietet mit ihren zahlreichen Funktionen Unternehmen eine maximale Flexibilität.

Inzwischen ergänzt der Astimax Konferenzserver mit WebRTC Integration das Angebot. Mit der Astimax VCC bietet ADDIX als erstes Unternehmen eine professionelle WebRTC Integration, die Unternehmen weitere Kommunikationswege erschließt.

Björn Schwarze, Geschäftsführender Gesellschafter ADDIX Software GmbH

„Die EML Spracherkennung nutzen wir für die Verschriftung von Sprachmailboxen, um Nachrichten schnell und diskret als E-Mail zu lesen und für die Dokumentation von Telefonkonferenzen. Da unsere Zielgruppe Businesskunden sind, steht die Zuverlässigkeit der Spracherkennung bei uns im Fokus.“



SPRACHSTEUERUNG

Branche: Automotive / Sprachsteuerung
Partner: Mobilcenter Zawatzky, Meckesheim



Barrierefrei mobil

Wie Sprachtechnologie Menschen mit Behinderung das Autofahren erleichtert

Wenn Josef Fleischmann in seinem Auto über die Autobahn fährt, genießt er wie viele junge Leute die Freiheit, mit seinem fahrbaren Untersatz fast unbegrenzt mobil zu sein. Doch der 23-jährige Student aus Stuttgart fährt anders als seine Kommilitonen: Er steuert sein Fahrzeug mit einem Joystick, denn er sitzt aufgrund einer angeborenen Muskeldystrophie im Elektrorollstuhl. Für seine Mobilität sorgt ein Auto, das der Fahrzeugumbauer Mobilcenter Zawatzky den Bedürfnissen Fleischmanns angepasst hat. Er gibt Gas, bremst und lenkt mit einem speziellen, von Zawatzky entwickelten Joystick. Die Sekundärfunktionen, also Blinker, Hupe, Scheibenwischer und Licht, steuert der angehende Wirtschaftsingenieur mit Sprachbefehlen. Dabei nutzt er die EML-Spracherkennungstechnologie. Die beiden Technologiepartner EML und Zawatzky haben vor einigen Jahren dieses System für barrierefreies Fahren entwickelt. Nach dem erfolgreichen Projektabschluss im Zentralen Innovationsprogramm Mittelstand (ZIM) des Bundeswirtschaftsministeriums vertritt Zawatzky die gemeinsame Entwicklung unter dem Namen „CenterVoice“ in Deutschland und einigen europäischen Ländern.

Bei der Entwicklung ganz oben stand die Sicherheit. „Uns war klar: Die Spracherkennung und die Übertragung auf die Funktionen muss vor allem schnell funktionieren“, erläutert EML-Forschungs- und Entwicklungsleiter Dr. Siegfried Kunzmann. Mit einem relativ kleinen Wortschatz von 200 bis 300 Begriffen und der optimalen Position des Mikrofons wurde die Erkennungsrate von nahezu 100 Prozent erreicht. Zusätzlich bauten die EML-Sprachtechnologen auch noch sogenannte „Komfortfunktionen“ ein, wie die Steuerung der Navigation oder die



Möglichkeit, Nachrichten zu diktieren. „Ich nutze die Sprachsteuerung schon fast zwei Jahre und komme im Alltag damit sehr gut zurecht“, zeigt sich Josef Fleischmann zufrieden. Der angehende Wirtschaftsingenieur studiert an der Dualen Hochschule Baden-Württemberg in Stuttgart und ist Werkstudent bei einem internationalen Telekommunikationsausrüster. Er pendelt mit seinem Auto zwischen Universität und Arbeitsstelle, fährt aber auch längere Strecken, zum Beispiel wenn er seine Eltern in Nürnberg besucht. „Das Autofahren mit Sprachsteuerung ist deutlich komfortabler, weil ich mich ganz auf das Fahren konzentrieren kann. Es ist auch sicherer, weil ich nur eine Taste betätigen muss, im Gegensatz zu einer Fernbedienung, auf die ich dann schauen muss, wenn ich zum Beispiel die Nebelschlussleuchte einschalten will, die ich nur selten brauche.“ Dank moderner Technologie ist Josef Fleischmann barrierefrei mobil – im Studium und in der Freizeit.

Mobilcenter Zawatzky GmbH

Mechatronische Systeme ermöglichen es Menschen mit Behinderung, ein Fahrzeug selbst zu steuern. Wer jedoch den Blinker setzen oder die Klimaanlage einstellen will, musste je nach Art der Behinderung auf kostenintensive Zusatzanfertigungen zurückgreifen. Für solche sogenannten Sekundär- und Komfortfunktionen ist die Sprache als Steuerungsinstrument ideal. Das EML und Mobilcenter Zawatzky GmbH, Meckesheim / Baden, entwickelten dafür ein System, das inzwischen in mehreren Sprachen auf dem Markt ist.

Andreas Zawatzky, Geschäftsführer, Mobilcenter Zawatzky GmbH

„Seit über 50 Jahren bauen wir Fahrzeuge für Menschen mit Behinderungen. Zur sicheren Bedienung von Blinker, Hupe, Wischer, Radio oder Fensterheber fehlte uns noch eine Sprachsteuerung. Durch die Zusammenarbeit mit dem EML konnten wir dieses Ziel erreichen.“



CALL-CENTER

Branche: Kundendialog, Call-Center

Interview mit Dennis Schottler, Geschäftsführer der voiXen GmbH, Berlin

Ihr Unternehmen bietet eine Webapplikation zur Archivierung, Analyse und Suche von Sprachdateien an. In welchen Bereichen agieren Ihre Kunden und welche Motive haben sie, die Lösungen von voiXen zu nutzen?

Der Großteil unserer Kunden stammt aus der Branche der Callcenter-Dienstleister, die täglich tausende Telefonate aus verschiedenen Gründen aufzeichnen. Hier sind die Motive ganz klar in der Qualitätssicherung zu finden. Ohne den Aufwand für manuelle, stichprobenartige Qualitätsprüfung einzelner Voicefiles findet voiXen schulungswürdige Gespräche genau dann, wenn sie stattfinden, und ermöglicht flexibles und effektives Coaching. Aber auch im Bereich Medien / Printmedien unterstützen wir mit unserer Lösung Kunden. Stellen Sie sich beispielsweise die vielen Interviews vor, die für Zeitschriften oder fürs Radio gemacht werden. VoiXen-Kunden sind hier in der Lage, jede Aufzeichnung aufzufinden, in der beispielsweise ein beliebiges Stichwort oder ein bestimmter Name genannt wurde.

„Ohne Speech2Text geht in Sachen ‘Big Data’ künftig gar nichts mehr.“

Sie sind seit über 20 Jahren im Call-Center-Bereich tätig und bieten speziell für diese Branche maßgeschneiderte Lösungen an. Welchen Mehrwert hat aus Ihrer Sicht der Einsatz von Sprachtechnologie und automatischer Transkription im Call-Center?

Mit einem Wort lautet meine Antwort: „Automatisierung!“. Der Kundendialog kann und wird künftig nicht mehr mit Niedriglöhnern exklusiv am Telefon stattfinden. Ich glaube an asynchrone Kommunikation. Böse Zungen behaupten, ich predige das Revival des Anrufbeantworters, und irgendwie ist da auch was dran. Nur werden Servicesuchende neben Selfservice-Modulen und Selbsthilfe-Motivation durch das permanent verfügbare WWW künftig nicht irgendwo anrufen, um ihre Serviceanfrage zu platzieren. Der Frust über Inkompetenz der Callcenter-Agenten (nicht nur fachlich – auch in Bezug auf Entscheidungskompetenz) ist zu groß geworden. Künftig wird man Sprachnachrichten auf Websites oder sogar auf an das Internet angeschlossenen Geräten hinterlassen. Und dann kommen die Transkription und die echte Business Intelligence zum Zuge, in Form einer Weiterleitung (Routing) der Anfrage zum Sachbearbeiter, der wirklich helfen und lösen kann. Das gesprochene Wort in Text gewandelt, editiert mit Informationen zur betroffenen Hardware (Alter, Garantie, Kaufort, nächstgelegener Servicetechniker etc.), wird die Kundenkommunikation in den nächsten fünf Jahren grundlegend verändern. Fragen Sie

mal 14-Jährige, wie die heute mit Serviceanfragen umgehen, dann bekommen Sie ein Bild davon, was Serviceanbieter künftig erwartet. Ohne Speech2Text geht in Sachen „Big Data“ künftig gar nichts mehr. Wer da nicht mitspielt, wird vom Markt verschwinden, und zwar lautlos.

Das EML ist „Technical Partner“ von voiXen. Welche EML-Produkte und Services nutzen Sie?

Wir verwenden den „EML Transcription Server“, der eingehende Anrufe sofort vollautomatisch und sprecherunabhängig in maschinell durchsuchbaren Text umwandelt. Außerdem nutzen wir den „EML Language Model Workplace“, mit dem wir die Spracherkennung an die jeweilige Anwendungsdomäne unserer Kunden anpassen können.

Wie sind Ihre bisherigen Erfahrungen damit? Wodurch unterscheiden sich die EML-Lösungen von denen anderer Anbieter?

Das EML entwickelt seine Produkte und Lösungen aus der Forschung heraus. Das macht EML zu einem Partner. Denn im Gegensatz zu anderen Anbietern will das EML nicht einfach nur von voiXen den Marktzugang erhalten, um seine Produkte zu verkaufen. Und die Forschung zahlt sich aus: Durch den Einsatz von „Deep Learning“-Methoden und nicht zuletzt durch ihre große Erfahrung mit automatischer Spracherkennung haben die EML-Technologen eine Erkennungsrate erreicht, die derzeit auf dem Markt nicht zu toppen ist.

Sie glauben an die Zukunft der „asynchronen Kommunikation“ im Call-Center, also: Voicemail statt Warteschleife. Welche Vorteile bietet hier die Sprachtechnologie?

Der Vorteil liegt auf der Hand: Der Anrufer kann sein Problem oder sein Anliegen sofort loswerden, ohne die nervtötende Odyssee durch ein IVR oder das Festsitzen in einer Warteschleife. Einfach losreden! Durch die Sprachtechnologie können die Voicemails nach Schlüsselworten (Keywords), Auftragsnummern, Namen etc. durchsucht und dem richtigen Bearbeiter oder Projekt zugeordnet werden. Dieser kümmert sich dann um das Anliegen. Callcenter geben viel zu viel Geld für Erreichbarkeit aus. Darunter leidet die FCR-Quote, die First Contact Resolution. Der Anrufbeantworter mit Spracherkennung kann mehr zum besseren Preis.

„Callcenter geben viel zu viel Geld für Erreichbarkeit aus. Der Anrufbeantworter mit Spracherkennung kann mehr zum besseren Preis.“

Welche Bereiche werden Ihrer Meinung nach in Zukunft vom Einsatz der automatischen Transkription gesprochener Sprache profitieren?


Jeder, der Audioinhalte (gesprochen - nicht gesungen) publiziert, kann mit Speech2Text Inhalte findbar machen und genau sagen, an welcher Stelle welche Werbeeinblendung sinnvoll ist. Den Rest überlassen wir dann Google, Facebook und Co. Oder: Durchsuchen Sie mal YouTube nach allen Interviews von Frau Merkel, in denen sie das Wort „Datenschutz“ sagt, und springen Sie direkt zu dieser Stelle im Video. Geht heute nicht. Aber morgen!

EML European Media Laboratory GmbH wurde von SAP-Mitbegründer Klaus Tschira als privates IT-Unternehmen gegründet. Das EML entwickelt Software und Technologien in der automatischen Sprachverarbeitung. Schwerpunkt ist die automatische Umwandlung von Sprache in Text (Transkription), die in Telefonie (Speech Analytics, Voice Mail) und mobilen Anwendungen (Voice Texting, Voice Search) eingesetzt wird.

www.eml.org

 @EMLVoice

 EMLEuropeanMediaLaboratory

 EMLVoiceMessaging

Kontakt:

EML European Media
Laboratory GmbH
Mathematikern
Berliner Str. 45
69120 Heidelberg

Forschung und Entwicklung:

Dr. Siegfried Kunzmann
kunzmann@eml.org
Tel. +49-6221-533-206
Fax +49-6221-533-282

Kommunikation:

Dr. Peter Saueressig
saueressig@eml.org
Tel. +49-6221-533-245
Fax +49-6221-533-282

© EML 2016



“Die Zeichen der Zeit stehen auf **Sprache!** Nach der neuesten **Bitkom-Studie** vom Februar 2016 bedienen in Deutschland mehr als die Hälfte aller Smartphone-Nutzer ihre Geräte per Stimme. Wir verschriften gesprochene Sprache automatisch in Echtzeit. Dabei arbeiten wir mit **neuronalen Netzen** und **Deep Learning**. Lesen ist schneller als Hören: Für eine Sprachnachricht von 79 Sekunden benötigt man nur 7 Sekunden, um sie zu lesen. Im Geschäftsalltag bedeutet das einen Unterschied, der sich rechnen kann. Unser **Know-how** stellen wir verschiedenen Branchen zur Verfügung – vom Call-Center bis zur Sprachsteuerung für barrierefreies Autofahren.“

Dr. Siegfried Kunzmann,
EML-Forschungs- und Entwicklungsleiter