

# Sprachassistenten zum Diktat

Von Christin Hohmeier

*Datenbanken verstehen nur das geschriebene Wort, mit Sprachdateien kann kein Big-Data-System, keine Callcenter-Anwendung und auch keine Predictive-Maintenance-App arbeiten. Damit Künstliche Intelligenz und Sprachsteuerung funktionieren, müssen die Daten konvertiert werden – von Sprache zu Schrift.*

Die Kommunikation mit einer sprachgesteuerten Big-Data-Anwendung erfolgt in beide Richtungen über Lautsprecher und Mikrofon – ein Techniker repariert eine Maschine und benötigt Informationen. Er wählt am Smartphone die Nummer seines Supports, ein Softwareroboter nimmt den Anruf entgegen.

Der übersetzt Sprache in Text, füttert mit diesem Ergebnis die Suchmaschinen und Datenbanken. Die spielen das Ergebnis zurück an den Roboter, der nun den Text in Sprache wandelt und wiederum über das Telefon erklärt, welche Klappe der Techniker an der Maschine öffnen soll. Der Bot beschreibt, mit welchem Werkzeug der menschliche Kollege welche Komponente austauscht und fragt, ob er weitere Information benötigt.

Auf ein „Ja“ schickt der digitale Assistent einen Bauplan der Maschine auf das Laptop. Hat der Assistent die Pläne verwechselt, kann der Mechaniker noch einmal nach dem richtigen Bauplan fragen. So spielen sich beide am Telefon die Bälle zu und die Maschine läuft.

Denn anscheinend ist das KI-System fehlerhaft programmiert und hat deshalb den falschen Bauplan geschickt. Ein Software-Entwickler kann das Problem betrachten und eine Lösung finden.

Auch die Intervalle, in denen eine Maschine gewartet oder repariert wird, kann eine wichtige Information sein. Hierfür erhält das KI-System alle transkribierten Dialoge zwischen Techniker und Software-Roboter. Es liest, vergleicht, zählt, lernt, wertet aus und berichtet – „alles im grünen Bereich“.

Und im nächsten Schritt sind der Phantasie, was ein Unternehmen mit den Protokollen anfangen kann, kaum mehr Grenzen gesetzt. Ein naheliegender Gedanke ist es, den Mensch-Maschinen-Dialog für Schulungen und Weiterbildungen zu nutzen. Dazu überarbeitet das Schulungsteam ganz einfach die Datei im Schreibprogramm und bekommt ein kurzes Handbuch für einen Wartungsprozess an einer ganz bestimmten Maschine.

Weitere Möglichkeiten sind Qualitätssicherung und der Nachweis darüber, dass ein Service zu einem Stichtag durchgeführt wurde.

## Digitale Assistenten assistieren beim Papierkram

Aber das Konzept lässt sich weiterdenken. Der Techniker kann den Sprachzugang zum Support-Bot nutzen, um über diesen Weg weitere Informationen zu diktieren. Der Verschleiß und der Ausfall der nächsten Komponente ist schon absehbar? Ein Eintrag in dem Kalender seiner Kollegen ist über die Sprachsteuerung ebenso möglich, wie die Bestellung der Komponente in der Einkaufsabteilung oder die Anfrage, ob das fragliche Teil auf Lager liegt.

Aus dem Bot wird ein Assistent, der dem Techniker den lästigen Papierkram vom Hals hält. Auf dem Weg zurück in die Firma und weiter in den Feierabend diktiert er dem digitalen Assistenten seine täglichen Einsatzberichte. Und noch vor der Tagesschau sind alle Dokumente fertig transkribiert und liegen in der Inbox seines Chefs und der Buchhaltung, und warten auf Freigabe, Rechnungsstellung für den Kunden und die endgültige Archivierung.

„Das KI-System erhält alle transkribierten Dialoge zwischen Mechaniker und Bot.“

## Protokolle der Supportanfragen

In der Natur dieses Dialogs liegt es, dass sich am Ende niemand mehr an dessen Inhalt erinnert. Aber tatsächlich kann das Gespräch Informationen enthalten, die so wichtig oder überraschend sind, dass es sinnvoll erscheint, alles mit Hilfe von Künstlicher Intelligenz noch einmal anzusehen und zu analysieren.

WERBEBEITRAG | INTERVIEW

## „Sprachanalyse profitiert von KI“



„Mit neuronalen Netzen und Deep Learning erreichen wir höchste Spracherkennungsraten“, heißt es bei EML European Media Laboratory GmbH, Heidelberg. Wir sprachen mit Dr. Volker Fischer (rechts) und Markus Klehr (links).

Herr Klehr, wofür setzen Sie Spracherkennung und -analyse ein? Unsere Software konvertiert gesprochene Sprache in Echtzeit in maschinenlesbaren Text und analysiert diesen Text dann. Die Technologie wird zum Beispiel in Callcentern eingesetzt, um eingehende Anrufe an den richtigen Ansprechpartner weiterzuleiten. Außerdem hilft sie,

Produktrends oder auch Krisen zu erkennen, wenn plötzlich bestimmte Stichwörter viel häufiger auftreten.

Dr. Fischer, welche Rolle spielt dabei die Künstliche Intelligenz (KI)? KI optimiert Spracherkennung und Sprachanalyse. Wir erreichen mit neuronalen Netzen und Deep Learning höchste Erkennungsraten in Echtzeit. Unsere Sprachmodelle, also die Fülle der statistischen Zusammenhänge, die wir zur Erkennung benötigen, lernen mit. Die Callcenter können Produktlisten, Dienstleistungen und Aktivitäten schnell und direkt anpassen. Und ihre Mitarbeiter nutzen die gewonnenen Daten, um kundenfreundlicher und effizienter zu kommunizieren.

Apropos Daten: Wie sieht es denn mit dem Datenschutz aus? Spracherkennung und Modellanpassung sind on-premise, also bei unseren Kunden. So halten wir – im Gegensatz zu manchen US-Anbietern – die strengen europäischen Datenschutzrichtlinien ein.

www.eml.org

Quelle: Bitkom, 2018

## Würden Sie für Bankgeschäfte wie Überweisungen eine Sprachsteuerung nutzen?

35%



Würde ich nutzen



61%

Würde ich nicht nutzen

4%



Weiß nicht/ keine Angabe