



Mittelstand-Digital
Zentrum
Schleswig-Holstein



Praxisprojekt

Sprechen statt Tippen

KI-gestütztes Ausfüllen von Formularen

Das Team des Mittelstand-Digital Zentrum Schleswig-Holstein konzipiert mit der Univelop GmbH im Rahmen einer Machbarkeitsanalyse Ansätze zur Spracheingabe in Formulare, um die Anwenderfreundlichkeit von Apps zu erhöhen.

Smartphones und Tablets überzeugen als Multifunktionsgeräte in Unternehmen nicht nur durch E-Mail-, Kalender- oder Kontaktfunktionen, sondern auch durch den Einsatz individueller Unternehmens-Apps, die bei der täglichen Arbeit helfen. Dabei sind einfache und schnelle Bedienungsmöglichkeiten für die Anwender Voraussetzung für die Akzeptanz und die erfolgreiche Nutzung der App. Was liegt also näher, als sich darüber Gedanken zu machen, ob man anstelle des Eintippens von Inhalten in Formulare den Nutzern nicht die Möglichkeit an die Hand gibt, Inhalte in Formulare einzusprechen?

Sprechen statt Tippen – einfach und schnell unterwegs.

Genau diesen Ansatz möchte Univelop mit Unterstützung des Mittelstand-Digital Zentrum Schleswig-Holstein (MDZ-SH) auf den Prüfstand stellen. Das Unternehmen bietet einen No-Code-App-Baukasten, mit dem Unternehmen Ihre Geschäftsprozesse eigenständig entsprechend ihrer individuellen Prozesse digitalisieren können. Bausteine wie Textfelder, Drop-Down-Auswahlfelder, Formeln und vieles mehr können in beliebiger Reihenfolge zusammengestellt werden und werden anschließend als App zur Verfügung gestellt.

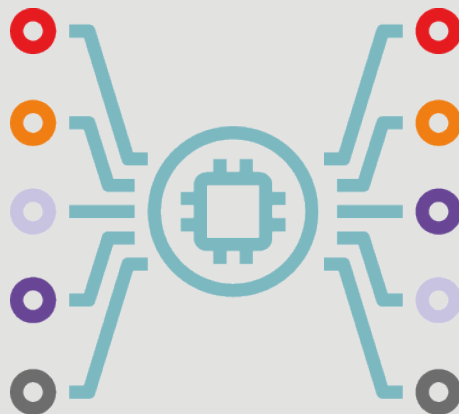
Die Herausforderung

Im Idealfall können Benutzer die Formulare zukünftig sprachgesteuert ausfüllen. Dazu soll die natürliche Sprache vom Benutzer aufgenommen und direkt in das passende Formularfeld eingefügt werden. Im Rahmen des Projektes erarbeitete das Team des MDZ-SH mit Univelop zwei Ansätze:

1. Szenario 1: Der Nutzer wählt das Formularfeld aus und spricht dann den Text ein
Diese Möglichkeit unterstützt die Bedienerfreundlichkeit bereits erheblich. Allerdings muss der Anwender trotzdem noch die Felder anklicken und sich durch das Eingabeformular hangeln.
2. Szenario 2: Der eingesprochene Text wird automatisch den richtigen Formularfeldern zugeordnet.
Ein Wunschscenario in Sachen Bedienerfreundlichkeit: Formulareingabe starten und nur noch sprechen. Allerdings ist das Erkennen des richtigen Formularfeldes nicht einfach, denn die App muss eigenständig den Sinn und die Bedeutung der Spracheingabe erkennen und diese den Formularfeldern zuordnen. Erschwert wird diese Zuordnung dadurch, dass es keine Standardformulare gibt. Schließlich werden diese individuell von den Nutzern entworfen.



Spracheingabe
in APP



Erkennen der
Formularfelder
und Inhalte



Zuordnen und
ausfüllen

Eingesprochener Text soll mittels Künstlicher Intelligenz analysiert und den Formularfeldern automatisch zugeordnet werden.

Das Vorgehen

Im weiteren Projektverlauf wurden ausschließlich Lösungsansätze für Szenario 2 entwickelt. Dabei wurden unterschiedliche Möglichkeiten beleuchtet, wie zum einen das gesprochene Wort den Formularfeldern zugeordnet werden kann und zum anderen, wie der Inhalt korrekt eingetragen werden kann.

Für die Sprachaufnahme und Umwandlung in Text wird auf die Spracherkennungsfunktion des Android bzw. Apple Smartphones zurückgegriffen. Anschließend muss der eingesprochene Text daraufhin analysiert werden, welche Teile welchem Formularfeld zuzuordnen sind. Damit ein Formularfeld als solches erkannt wird, werden alle Felder mit einem eindeutigen Namen (=TAG) versehen. Bei der Sprachaufnahme wird der Benutzer die Inhalte in der Reihenfolge „TAG – Formularfeldinhalt“ einsprechen. Und das natürlich in beliebiger Feldreihenfolge und auch für mehrere Felder hintereinander in einer Sprachaufnahme. TAGs müssen dann als TAGs identifiziert werden und wenn der eingesprochene Text zu einem TAG passt, wird er in das entsprechende Formularfeld eingetragen. Wie passen vom System beurteilt wird, war eine der Kernfragestellungen der ImpulsWerft.

Dazu wurden mehrere Metriken und Verfahren daraufhin analysiert, wie gut sie sich für die hier beschriebene Aufgabenstellung eignen.

Der Projektpartner



Univelop GmbH



11 Mitarbeitende



Software-Entwicklung

Univelop bietet Digitalisierungslösungen für Prozesse, beispielsweise im Außendienst, in der Produktion, im Vertrieb oder in der Geschäftsführung. Zugrunde liege eine Bausteinsystem, mit dem sich in wenigen Schritten individuelle Prozesse abbilden lassen.

Die Lösung

Wie gut ein Wort zu einem TAG passt kann über unterschiedliche Metriken berechnet werden. Hier kann man zwischen syntaktischen und semantischen Metriken unterscheiden. Die syntaktischen Metriken vergleichen das erkannte Wort mit den TAGs Zeichen für Zeichen und ergeben eine niedrigere Distanz, je ähnlicher sich die beiden Worte sind. Semantische Metriken hingegen ignorieren die Schreibweise und setzen dafür Wort und TAG in eine semantische Beziehung. So wären zum Beispiel die Worte "Kugelschreiber" und "Bleistift" deutlich näher miteinander verwandt als mit "Kugellager", obwohl eine syntaktische Metrik ein anderes Ergebnis liefern würde. Unterschiedliche Metriken können in Kombination angewandt werden, um sich die jeweiligen Stärken der Metrik zu Nutze zu machen

Betrachtet wurden

1. Edit Distance: Hier wird quantifiziert, wie unterschiedliche zwei Wörter sind
2. Phonetische Distanz: Vergleichende Worte werden in Lautschrift überführt und dann Stück für Stück verglichen
3. Word2Vec: Modell aus dem Maschinellen Lernen, über das Wörter in ihrer Bedeutung verglichen werden können
4. Semantische Distanz im WordNet: WordNets sind Sammlungen aller Worte einer Sprache, über die Worte verglichen werden können

Univelop hat während des Projektes unterschiedliche Techniken und Metriken für die automatisierte Textverarbeitung kennengelernt und konnte so die eigene Projektidee weiterentwickeln.

Ihr Ansprechpartner

Mittelstand-Digital
Zentrum Schleswig-Holstein
Christoph Linse
linse@inb.uni-luebeck.de
www.digitalzentrum-sh.de