

Projekt: Smart Speaker

## Einsatz von Sprachassistenzsystemen in der Wertschöpfung von KMU des Maschinen- und Anlagenbaus

### Vorgehensmodell zur Identifikation nutzenstiftender Einsatzszenarien für Sprachassistenzsysteme für KMU



Während Sprachassistenzsysteme im privaten Kontext weit verbreitet sind, existieren im industriellen Bereich nur vereinzelte Anwendungsbeispiele. Im Rahmen des Forschungsprojekts ‚Smart Speaker‘ wird genau dieser Mangel adressiert. Ziel des Projekts ist die Entwicklung eines Vorgehensmodells, welches Industrieunternehmen zum Einsatz von Sprachassistenzsystemen befähigen soll. Das IGF-Vorhaben 20983 N der Forschungsvereinigung FIR e. V. an der RWTH Aachen wird über die AiF im Rahmen des Programms zur Förderung der industriellen Gemeinschaftsforschung (IGF) vom Bundesministerium für Wirtschaft und Energie (BMWi) aufgrund eines Beschlusses des Deutschen Bundestages gefördert.

**S**prachassistenzsysteme – auch als Smart Speaker bezeichnet – finden sich in immer mehr Privathaushalten. Hierzu zählen Assistenten wie Google Home (Google Assistant), Apple HomePod (Siri) oder Amazon Echo (Alexa). Ein Überblick über die Funktionsweise von Sprachassistenzsystemen, die nachfolgend beschrieben werden, bietet Bild 1.

Der Nutzer kommuniziert über gesprochene Sprache mit einem Voice-User-Interface (1). Diese Hardware-Schnittstelle kann ein Smartphone, Kopfhörer, Mikrofon oder ähnliches sein (2). Sie erfasst die gesprochene Sprache und leitet sie als Input an eine Recheneinheit weiter. Dort werden nacheinander verschiedene Software-

Blöcke gestartet. Zunächst wird die Sprachnachricht – welche aus akustischen Schwingungen bestehend vorliegt – mithilfe eines „Speech-to-Text-Algorithmus“ (STT) in geschriebene Wörter umgewandelt (3). Im nächsten Schritt werden aus dem daraus entstandenen Text durch einen NLP- bzw. NLU-Algorithmus die für die Erkennung der Absicht des Nutzers notwendigen Informationen gewonnen (4). Diese können aus bestimmten Signalwörtern bestehen, sich aber auch auf den Kontext oder Formulierungen beziehen. Nachdem die Intention des Nutzers feststeht, werden im Dialogue-Management entsprechende Aktionen ausgeführt (5). Hierfür kann etwa über APIs auf weitere Software zugegriffen

(beispielsweise Datenbanken) oder eine Nachfrage an den Nutzer gestellt werden, falls Informationen fehlen. Nachfragen oder auch eine fertige Antwort an den Nutzer werden zunächst als geschriebene Wörter übermittelt. Diese können wahlweise direkt zum Lesen an den Nutzer übermittelt werden, über einen „Text-to-Speech-Algorithmus“ (TTS) in gesprochene Sprache – akustische Schwingungen – umgewandelt oder anderweitig verarbeitet werden (6). Der Nutzer erhält die Antwort dementsprechend über ein Display oder das Voice-User-Interface.<sup>1</sup>

<sup>1</sup> S. LI ET AL. 2016, S. 10; YU U. DENG 2015, S. 101 f.; SINGH ET AL. 2019, S. 73; TAYLOR 2009, S. 40f.

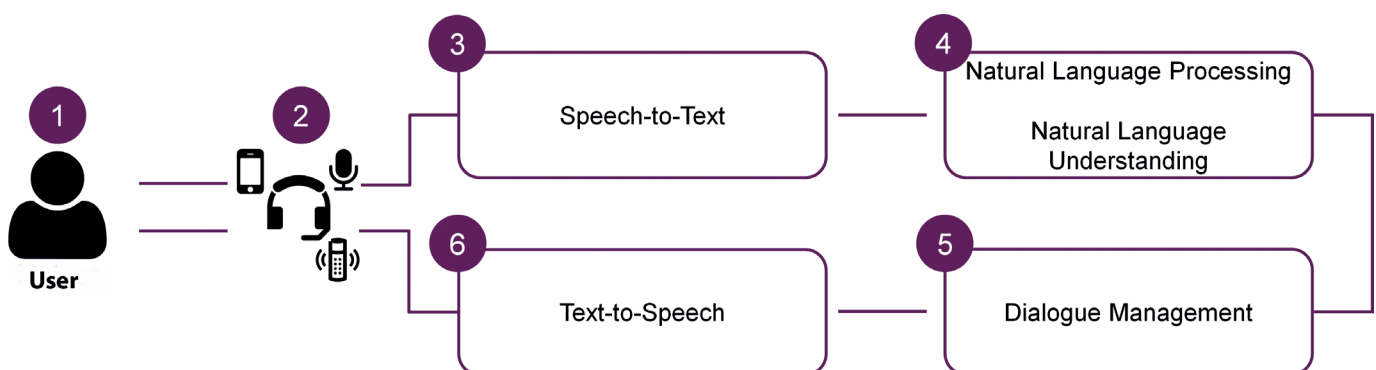


Bild 1: Funktionsweise von Sprachassistenzsystemen (eigene Darstellung)

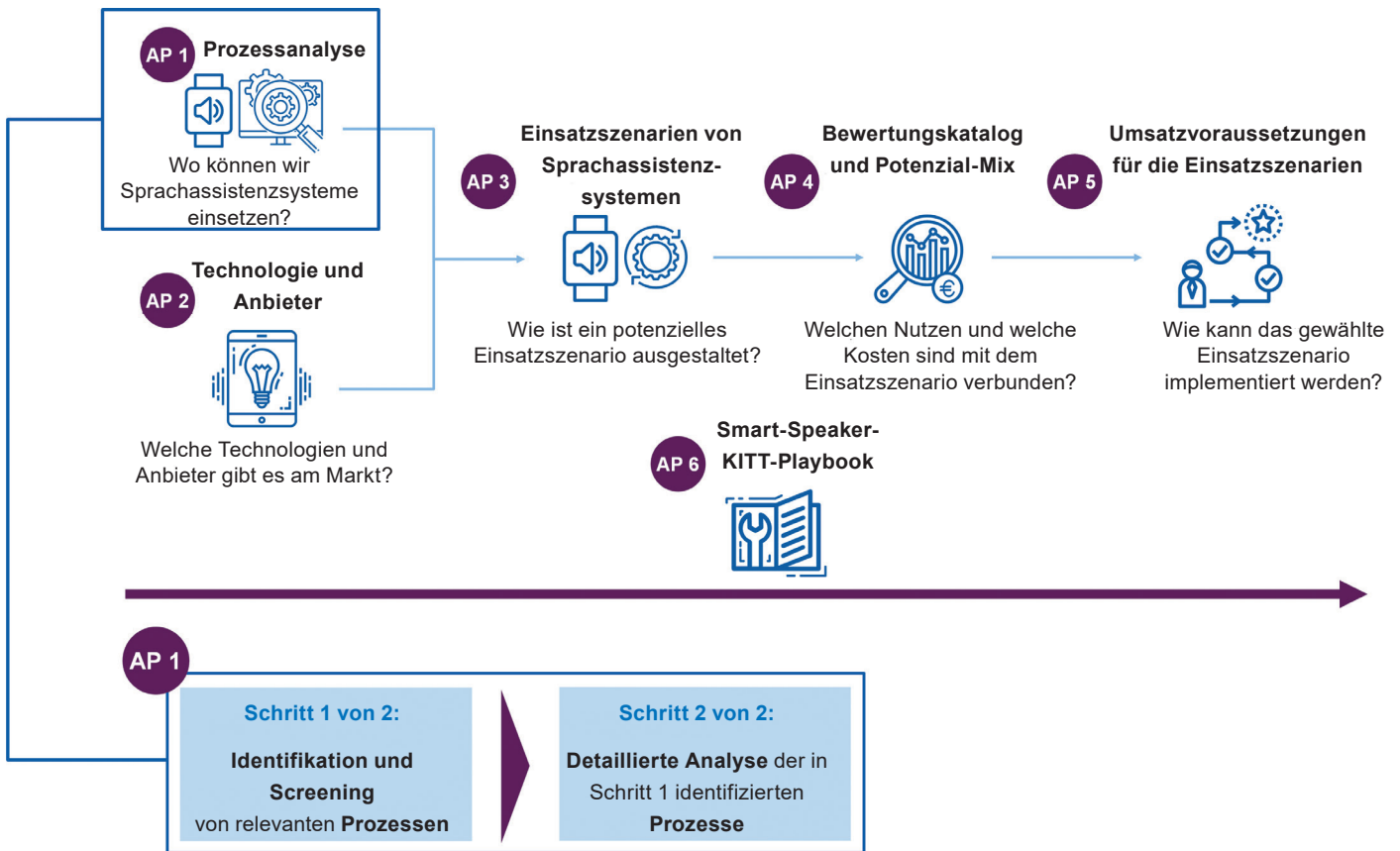


Bild 2: Vorgehensmodell zur Einführung von Sprachassistenzsystemen (eigene Darstellung)

Im privaten Bereich kommen Sprachassistenzsysteme beispielsweise im Automobil zur Sprachsteuerung des Navigationssystems oder von Multimedia-Anwendungen zum Einsatz. Im Gegensatz zum privaten Kontext existieren im industriellen Bereich aktuell bisher nur wenige Anwendungen. Bekannte Anwendungsbeispiele beschränken sich auf die Pick-by-Voice-Technologie in der Lagerlogistik oder die Verwendung von Chat-Bots im Kundenservice.<sup>2</sup>

Ziel des Forschungsprojekts ‚Smart Speaker‘ ist es, die beschriebene Forschungslücke zu schließen und industrielle Unternehmen zu befähigen, Sprachassistenzsysteme einzusetzen, um so ihre Unternehmensprozesse schneller, stabiler und effizienter gestalten zu können. Hierfür wird im Projekt

‚Smart Speaker‘ ein Vorgehensmodell zur Einführung von Sprachassistenzsystemen entwickelt und parallel mit einem projektbegleitenden Ausschuss erprobt und validiert. Das Vorgehensmodell umfasst sechs Schritte und ist in Bild 2 dargestellt.

Das erste Arbeitspaket dient dazu, die Unternehmensprozesse zu identifizieren und zu analysieren, die sich für den Einsatz von Sprachassistenzsystemen eignen. Hierfür werden zwei Unterschritte durchgeführt: Zunächst erfolgt im ersten Unterschritt ein grobes Screening möglicher Prozesse. Mithilfe von Prozesscharakteristika, die als Eignungsmaß für Einsatzmöglichkeiten von Sprachassistenzsystem dienen, werden aus einer Vielzahl möglicher Prozesse diejenigen identifiziert, die das höchste Potenzial bieten. Diese werden im zweiten Unterschritt

im Rahmen einer Ist-Prozess-Analyse detailliert betrachtet.

Gleichzeitig werden im zweiten Arbeitspaket die bestehenden Technologien untersucht. Hierbei werden kommerziell verfügbare Smart-Speaker-Lösungen und -Produkte betrachtet. Ziel des Arbeitspakets ist eine Marktübersicht der Technologien und entsprechender Anbieter. Das dritte Arbeitspaket ist eine Synthese der ersten beiden Arbeitspakete. Die im ersten Schritt analysierten Ist-Prozesse und die im zweiten Schritt eruierten Technologien werden gemeinsam mit den Unternehmen des projektbegleitenden Ausschusses zu Einsatzszenarien für Sprachassistenzsysteme zusammengeführt. Im vierten Schritt werden die Einsatzszenarien einer Kosten-Nutzen-Analyse unterzogen. Ziel ist die

<sup>2</sup> S. KREUTZER U. SIRRENBURG 2019, S. 149; KRUSE BRANDÃO U. WOLFRAM 2018, 273 f. UND S. 297 – 299; HÖRNER 2019, S. 43 UND S. 158 ff.

Entwicklung einer Bewertungsmethodik zur Priorisierung von Einsatzszenarien mit hohem Potenzial. Darauf aufbauend werden im fünften Schritt die Umsetzungsvoraussetzungen für die Einsatzszenarien betrachtet. Parallel werden in Arbeitspaket 6 die Ergebnisse der zuvor beschriebenen Arbeitspakete sukzessive gesammelt, konsolidiert und interessierten Unternehmen in Form eines visuell ansprechenden Playbooks zugänglich gemacht.

## Literatur

HÖRNER, T.: *Marketing mit Sprachassistenten. So setzen Sie Alexa, Google Assistant & Co strategisch erfolgreich ein.* Springer Gabler, Wiesbaden 2019.

KREUTZER, R. T.; SIRRENBURG, M.: *Künstliche Intelligenz verstehen: Grundlagen – Use-Cases – unternehmerische KI-Journey.* Springer, Wiesbaden [u. a.] 2019.

KRUSE BRANDÃO, T.; WOLFRAM, G.: *Digital Connection. Die bessere Customer Journey mit smarten Technologien – Strategie und Praxisbeispiele.* Springer, Wiesbaden [u. a.] 2018.

LI, J.; DENG, L.; HAEB-UMBACH, R.; GONG, Y.: *Robust automatic speech recognition. A bridge to practical*

*applications.* Academic Press (an imprint of Elsevier), Waltham (MA) 2016.

SINGH, A.; RAMASUBRAMANIAN, K.; SHIVAM, S.: *Building an Enterprise Chatbot: Work with Protected Enterprise Data Using Open Source Frameworks.* APRESS, New York 2019.

TAYLOR, P.: *Text to speech synthesis.* Prentice Hall, Upper Saddle River (NJ) 2009.

YU, D.; DENG, L.: *Automatic Speech Recognition. Speech a deep learning approach.* Springer, London [u. a.] 2015.

## Ansprechpartner:



Thies Bach, M.Sc.  
FIR e. V. an der RWTH Aachen  
Wissenschaftlicher Mitarbeiter  
Bereich Produktionsmanagement  
Tel.: +49 241 47705-405  
E-Mail: Thies.Bach@fir.rwth-aachen.de



Eren Cirit, M.Sc.  
FIR e. V. an der RWTH Aachen  
Wissenschaftlicher Mitarbeiter  
Bereich Produktionsmanagement  
Tel.: +49 241 47705-418  
E-Mail: Eren.Cirit@fir.rwth-aachen.de

Gefördert durch:



aufgrund eines Beschlusses  
des Deutschen Bundestages



**Projekttitel:** Smart Speaker

**Forschungs-/Projektträger:** Bundesministerium für Wirtschaft und Energie (BMWi);  
Arbeitsgemeinschaft industrieller Forschungsvereinigungen "Otto von Guericke" e. V. (AiF)

**Förderkennzeichen:** 20983 N

**Projektpartner:** IPRI International Performance Research Institute gemeinnützige GmbH

**Internet:** smart-speaker.info