

ITG Workshop

Sprachassistenten

Anwendungen, Implikationen, Entwicklungen

Magdeburg, 3. März, 2020

Herausgeber

Ingo Siegert

Sebastian Möller

Impressum

DOI: 10.25673/32572

Das Copyright der Beiträge liegt bei der Otto-von-Guericke-Universität Magdeburg und den oben genannten Herausgebern. Darüber hinaus dürfen die Autoren Kopien der Veröffentlichung auf ihren eigenen Webseiten bereitstellen. Die Beiträge werden weiterhin auch online auf der Website des Workshop veröffentlicht.

Herausgeber: Otto-von-Guericke-Universität Magdeburg

Inhaltsverzeichnis

Vorwort	5
Programm	7
Kurzfassungen der Beiträge	9
<i>Reinhold Häb-Umbach</i> Siri und Alexa: Verstehen wir uns?	9
<i>Serpil Tas, Rene Arnold</i> Nutzung von Sprachassistenten in Deutschland	11
<i>Benjamin Weiss</i> Unterwegs mit digitalen Sprachassistenzsystemen – ein Vergleich verschiedener Konzepte zur Mobilität	13
<i>Steffen Werner</i> „Hey Mercedes!“ — Ein Einblick in die Entwicklung des MBUX Voice Assistenten	15
<i>Timo Baumann</i> Inkrementelle Verarbeitung als Hebel für responsive Dialogführung . . .	17
<i>Hans-Günter Hirsch, Alexander Micheel, Jan Stähler, Michael Gref, Janik Göbel</i> Smart-Home Steuerung mit einem hausinternen Sprachassistenzsystem	19
<i>Delphine Reinhardt</i> Sprachassistenten: Unsichtbare und unerwartete Gäste am Tisch? . . .	21
<i>Stefan Hillmann</i> Usability von Sprachassistenten im Smart-Home Bereich	23
<i>Felix Burkhardt</i> Dialogsysteme mit Gefühlen: Wie, Warum und Wohin?	25
<i>Jan Nehring</i> Chatbots mit Persönlichkeit	27
<i>Julia Krüger & Ingo Siegert</i> “das ist schon gruselig so dieses Belauschtwerden” — Subjektives Er- leben von Interaktionen mit Sprachassistenzsystemen zum Zwecke der Individualisierung	29
<i>Jessica Szczuka, Stefan Kopp, Nicole Krämer, Lina Varonina</i> Alexa, kannst Du mich verstehen? Empirische Langzeituntersuchung zur Interaktion zwischen Sprachassistenten und Kindern	31
<i>Andreas Nautsch, Massimiliano Todisco, Jose Patino and Nicholas Evans</i> Audio Security & Privacy	35
<i>Marcus Petersen, Martin Suffa, Oliver Neumann, Tino Köhler, Florian Mar- quardt</i> Sprachbasierte Self-Services auf Basis etablierter und kostengünstiger Cloud-Dienste am Beispiel einer vollautomatisierten Zählerstandserfassung	37

Adarsh Kuzhipathalil, Anto Thomas, Keerthana Chand, Frank Ortmeier, Ingo Siegert

Intelligent LSF-Answering System - An Alexa Skill 39

Vorwort

Der ITG Workshop „Sprachassistenten – Anwendungen, Implikationen, Entwicklungen“ fand am 3. März 2020 in Magdeburg statt.

Die Verbreitung und Funktionalität digitaler Sprachassistentensysteme, wie Apple’s Siri, Amazon’s Alexa oder der Google Assistant, nimmt immer weiter zu. Die Attraktivität dieser Geräte ist in ihrer einfachen Bedienung begründet, sie gestatten es uns Internetrecherchen, Online-Bestellungen, Raumüberwachungen und andere Smart-Home-Dienste nur durch Zuruf durchzuführen.

Der Nutzerin oder dem Nutzer sind die Implikationen der sprachbasierten Interaktion jedoch noch nicht immer klar, insbesondere, da heutige Sprachassistenten teilweise nur bessere Fernbedienungen sind und das Gerät nur über einen Teil der Funktionen sprachliches Feedback gibt. Zukünftig sollen diese jedoch nicht nur einfache Befehle verarbeiten, sondern auch eine natürliche und reibungslose Interaktion ermöglichen. Hierzu ist neben technischen Verbesserungen der Spracherkennung auch eine verbesserten Sprachverständnis sowie intelligentere Dialogführung nötig. Weiterhin gehören neben neuen technischen Lösungen aber auch rechtliche Aspekte, die sich durch die Verbreitung von Sprachassistenten und der strengeren Datenschutzregelungen ergeben. Wann und was mitgehört wird ist häufig nicht ersichtlich und auch was mit den Sprachdaten geschieht weiß der Nutzer zumeist nicht. Dies kann zu Akzeptanzproblemen führen.

Auf dem Workshop wurden vielfältige und interdisziplinäre Beiträge präsentiert, die vor allem folgende Fragestellungen angesprochen haben:

- Wie zuverlässig arbeiten moderne Sprachassistenten, und welche Aspekte der wahrgenommenen Qualität und des Nutzererlebens werden adressiert?
- Wie können Sprachassistenten anzeigen, wann Spracheingaben getätigt werden, und was davon erkannt wird?
- Wie kann die Dialogkomplexität für Sprachassistenten weiter gesteigert werden?
- Wie können Sprachassistenten ihre Funktionalitäten aufzeigen, inklusive möglicher Implikationen?
- Wie können Spracheingaben unterschiedlicher Nutzerinnen getrennt analysiert werden?
- Welche weiteren Informationen lassen sich aus diesen Sprachdaten über die Nutzerin bzw. den Nutzer gewinnen? Mit welcher Zuverlässigkeit?
- Wo werden diese Informationen in welcher Form gespeichert? Für wen sind sie zugänglich?
- Wie können Sprachinteraktionen gestaltet werden, um besser die Privatsphäre ihrer Nutzer zu respektieren?
- Welche Implikationen ergeben sich für die Akzeptanz von Sprachassistenten?

Durch die gute Mischung von Beitragenden sowohl aus der Hochschullandschaft als auch aus der Industrie konnten die verschiedensten Aspekte anwendungsnah diskutiert werden.

Die Organisatoren des Workshops danken dem ITG Fachbereich „Dienste und Anwendungen“ und der VDE Bereichsgruppe Magdeburg für die finanzielle Unterstützung,

sowie allen Autoren und Teilnehmern für ihre aktive Teilnahme, die diesen Workshop erst zu einem erfolgreichen Event gemacht habt.

Magdeburg, Februar 2020

Sebastian Möller, Ingo Siegert

Programm

- 12:30 - 12:45 Registrierung und Anreise
- 12:45 - 13:00 Begrüßung und Einführung
- 13:00 - 15:00 Session 1: Praxisbeispiele
- Siri und Alexa: Verstehen wir uns? *Reinhold Hüb-Umbach (Uni Paderborn)*
- Nutzung von Sprachassistenten in Deutschland. *Serpil Tas (Wissenschaftliches Institut für Infrastruktur und Kommunikationsdienste)*
- Unterwegs mit digitalen Sprachassistentensystemen – ein Vergleich verschiedener Konzepte zur Mobilität. *Benjamin Weiss (Carmeq)*
- „Hey Mercedes!“ — Ein Einblick in die Entwicklung des MBUX Voice Assistenten. *Steffen Werner (Daimler AG)*
- Inkrementelle Verarbeitung als Hebel für responsive Dialogführung. *Timo Baumann (Uni Hamburg)*
- Smart-Home Steuerung mit einem hausinternen Sprachassistentensystem. *Hans-Günter Hirsch, Alexander Micheel, Jan Stähler, Michael Gref, Janik Göbel (Hochschule Niederrhein)*
- 15:00 - 15:30 Kaffeepause
- 15:30 - 17:30 Session 2: Nutzererleben
- Sprachassistenten: Unsichtbare und unerwartete Gäste am Tisch? *Delphine Reinhardt (Uni Göttingen)*
- Usability von Sprachassistenten im Smart-Home Bereich. *Stefan Hillmann (TU Berlin)*
- Dialogsysteme mit Gefühlen: Wie, Warum und Wohin? *Felix Burkhardt (Audeering GmbH)*
- Chatbots mit Persönlichkeit. *Jan Nehring (DFKI Berlin)*
- “das ist schon gruselig so dieses Belauschtwerden — Subjektives Erleben von Interaktionen mit Sprachassistentensystemen zum Zwecke der Individualisierung. *Julia Krüger Ingo Siegert (Uni Magdeburg)*
- Alexa, kannst Du mich verstehen? Empirische Langzeituntersuchung zur Interaktion zwischen Sprachassistenten und Kindern. *Jessica Szczuka, Stefan Kopp, Nicole Krämer, Lina Varonina (Universität Duisburg-Essen, Universität Bielefeld)*
- 17:30 - 18:00 Podiumsdiskussion
- ab 18:00 Posterpräsentation und getTogether

Kurzfassungen der Beiträge

Siri und Alexa: Verstehen wir uns?

Reinhold Hüb-Umbach
Universität Paderborn

Nach Jahrzehnten der Forschung ist die automatische Spracherkennung schließlich in Alltagsprodukten angekommen, und fast jeder von uns hat schon Erfahrungen mit Siri, Alexa oder anderen Sprachassistenten gesammelt. Warum hat das eigentlich so lange gedauert und warum war es so schwierig, dem Computer das Hören und Verstehen beizubringen? Der Vortrag gibt einen Überblick über die aktuelle Technologie digitaler Sprachassistenten und diskutiert auch zukünftige Chancen und Risiken der Sprachtechnologie.

Reinhold Hüb-Umbach ist Professor für Nachrichtentechnik an der Universität Paderborn. Nach dem Studium der Elektrotechnik und der Promotion an der RWTH Aachen war er Postdoc am IM Almaden Research Center in San Jose, Kalifornien. Von 1990 bis 2001 arbeitete er als wissenschaftlicher Mitarbeiter, seit 1997 als „Senior Scientist“, am Philips Forschungslaboratorium Aachen. Im Jahr 2000 war er Visiting Scientist am Philips Naturkundig Laboratorium in Eindhoven, und seit dem 1.4.2001 ist er Professor an der Universität Paderborn. Er ist Koautor von mehr als 200 Veröffentlichungen, darunter das Buch „Robust Automatic Speech Recognition - a Bridge to Practical Applications“ (Academic Press, 2015) und ein Artikel zu digitalen Sprachassistenten, der im Nov. 2019 im IEEE Signal Processing Magazine veröffentlicht wurde. Er ist sowohl IEEE Fellow (IEEE: Institute of Electrical and Electronics Engineers) als auch ISCA Fellow (ISCA: International Speech Communication Association). Seine Forschungsinteressen sind die statistische Signalverarbeitung und maschinelle Lernverfahren, mit Anwendungen im Bereich der Sprachsignalverbesserung, der automatischen Spracherkennung und des unüberwachten Lernens aus Sprach- und Audiosignalen.

Nutzung von Sprachassistenten in Deutschland

Serpil Tas, Rene Arnold

*Wissenschaftliches Institut für Infrastruktur und
Kommunikationsdienste*

Ob Amazons Alexa, Apples Siri, Googles Assistant, Microsofts Cortana oder Samsungs Bixby: Sprachassistenten werden immer populärer. Laut der Befragung von 3.184 Konsumenten in Deutschland, die für diesen Diskussionsbeitrag durchgeführt wurde, nutzten Ende des Jahres 2018 schon 26% zumindest einen solchen Sprachassistenten. Die Adoptionsrate ähnelt damit derjenigen von Smartphones etwa fünf Jahre nach Einführung der wesentlichen ersten Modelle. Der entscheidende Unterschied zwischen den beiden Technologien ist, dass die meisten Konsumenten in Deutschland keinen Cent zusätzlich ausgeben und nicht einmal eine zusätzliche App herunterladen müssen, um einen Sprachassistenten zu nutzen. Rund 85% haben schon zumindest ein Endgerät zuhause, das einen Sprachassistenten vorinstalliert hat. Oft handelt es sich dabei um das Smartphone oder den Laptop. Bei etwa 11% der Befragten sind es spezielle sogenannte Smart Speaker. Mit dieser großen Verbreitung von Endgeräten mit vorinstallierten integrierten Sprachassistenten ist ein sprunghafter Anstieg der Adoptionsrate möglich. Wahrscheinlich ist ein solcher Anstieg nach den Ergebnissen der Befragung in naher Zukunft jedoch nicht. Die überwiegende Mehrheit der Nicht-Nutzer hat nicht die Absicht, im nächsten Jahr einen der fünf populärsten Sprachassistenten zu verwenden. Die derzeitigen Nutzungsmuster deuten ebenso nicht auf einen großen Einfluss der Assistenten hin. Zumeist werden sie eher selten und für eine recht eingeschränkte Auswahl von ca. 2 bis 3 Funktionen genutzt. Die am häufigsten genutzte Funktion ist das Erfragen einfacher Informationen, die aus dem Internet abgerufen werden, wie beispielsweise das Wetter, Sportergebnisse oder die nächste Tankstelle. Alexa, der Sprachassistent von Amazon, wird im Durchschnitt deutlich häufiger genutzt als die anderen Sprachassistenten. Dabei verwendet der durchschnittliche Alexa-Nutzer rund eine Funktion mehr als die Nutzer der anderen Systeme. Besonders beliebt ist das Abspielen von Musik, sei es über den Amazon-eigenen Dienst oder einen Drittanbieter.

Serpil Tas Frau Serpil Tas ist als Senior Economist in der Abteilung „Märkte und Perspektiven“ am WIK Wissenschaftliches Institut für Infrastruktur und Kommunikationsdienste (WIK) tätig. Dort arbeitet sie an Projekten und Studien im Bereich der Telekommunikations- und Internetmärkte. Zu ihren thematischen Schwerpunkten gehören vor allem die Analyse der Nachfrageseite von neuen B2C- und B2B-Märkten für Telekommunikationsdienste und neue internetbasierte Anwendungen sowie die Durchführung von Geschäftsmodell- und Branchenanalysen. Frau Tas studierte Volkswirtschaftslehre an der Philipps-Universität in Marburg sowie an der Heinrich-Heine-Universität in Düsseldorf.

Rene Arnold leitet seit Juni 2013 die Abteilung „Märkte und Perspektiven“ der WIK-Consult. Von 2010 bis 2013 war er Referent bei der IW Consult, einer hundertprozentigen Tochter des Instituts der deutschen Wirtschaft Köln (IW Köln). Dort baute er den Bereich IKT- und Internet-Studien auf und leitete zahlreiche Projekte aus den Bereichen Branchen- und Regionalanalysen sowie Innovationspolitik. Zu den thematischen

Schwerpunkten gehörten die ökonomischen Effekte von Internet-basierten Anwendungen, die Digitalisierung von Geschäftsmodellen innerhalb der deutschen Wirtschaft, die Bedeutung der Metall- und Elektroindustrie, regionale Entwicklungsstrategien für den ländlichen Raum, sowie die Ausrichtung der europäischen Innovationspolitik. Studium der Betriebswirtschaftslehre in Heilbronn, Abschluss 2007 (Diplombetriebswirt (FH)) mit den Schwerpunkten auf Marketing und Marktforschung. Es folgte eine Promotion im Bereich Konsumentenverhalten an der University of Edinburgh.

Unterwegs mit digitalen Sprachassistenzsystemen – ein Vergleich verschiedener Konzepte zur Mobilität

Benjamin Weiss
Carmeq

Mit dem aktuellen wirtschaftlichen Erfolg digitaler Sprachassistenzsysteme haben sich bereits zwei unterschiedliche Ansätze etabliert, die sich insbesondere über ihre Produkttyp der Hardware differenzieren lassen: Stationäre SmartHome-Systeme („intelligente Lautsprecher“) stehen Smartphones („persönliche Assistenten“) gegenüber. Dabei können einzelne Smart Speakers keine mobilen Nutzer assistieren, sondern nur als Verbund versuchen, eine reibungslose User Experience zu ermöglichen. Für mobile Szenarien ergeben sich je nach Ansatz unterschiedliche Anforderungen und Potentiale, die ich für die Aspekte der Sprachinteraktion, Kommunikationstechnik und User Experience systematisiere.

Benjamin Weiss studierte Kommunikationsforschung/-wissenschaft, Erziehungswissenschaft und Skandinavistik in Bonn, Trondheim und Berlin. Danach promovierte er an der Humboldt-Universität zu Berlin in germanistischer Linguistik zu sprechtempoabhängigen Aussprachevariationen. 2007 begann er am Quality and Usability Lab der Deutschen Telekom Innovation Laboratories / TU Berlin seine Forschung in den Themenbereichen multimodaler Sprachdialogsysteme, Sprachübertragungsqualität und stimmliche Sympathie. Benjamin Weiss schloss 2019 seine Habilitation zur Qualität natürlicher Dialogen und sprachbasierter, multimodaler Mensch-Computer-Interaktion ab. Er war Senior Researcher bei der audEERING GmbH und arbeitet nun im Bereich automotive conversational interfaces bei der Carmeq GmbH und lehrt als Privatdozent an der TU Berlin.

„Hey Mercedes!“ — Ein Einblick in die Entwicklung des MBUX Voice Assistenten

Steffen Werner

Daimler AG

Mercedes-Benz ist bekannt als Erfinder des Automobils. Aber auch im Bereich der fahrzeugbasierten Sprachdialogsysteme wurde Pionierarbeit geleistet. Was vor 50 Jahren in der Forschung begann und 1996 das erste Mal in einem Serienfahrzeug verbaut wurde, mündete im MBUX Voice Assistenten, der 2018 das erste Mal dem Kunden zur Verfügung gestellt wurde. Der Vortrag gibt einen Einblick in die Entwicklung des MBUX Sprachassistenten, stellt die größten Herausforderungen und deren Lösungen vor und zeigt, wie die Kundenakzeptanz – als das wichtigste Gütemaß von Sprachassistenten – bereits im Entwicklungsprozess adressiert werden kann.

Steffen Werner machte bei seinem Studium der Elektrotechnik an der Technischen Universität Dresden erste Kontakte mit der digitalen Sprachsignalverarbeitung. Nach einem Intermezzo an der Gerhard-Mercator-Universität Duisburg-Essen, absolvierte er seine Promotion ebenfalls an der TUD. Seine Arbeit befasste sich mit Modellierung von Aussprachevarianten in Sprachmodellen für die Verwendung in Sprachsynthese- und erkenntnisystemen. Er wechselte 2007 zur Daimler AG in Sindelfingen und entwickelte hier Sprachdialogsysteme für Mercedes-Benz Fahrzeuge. Sein größtes Projekt war dabei die komplette Neu-Entwicklung des Dialogsystems LINGUATRONC für die Mercedes-Benz User Experience (MBUX) welches 2018 in der A-Klasse erstmals zum Einsatz kam. Ende 2018 wechselte Steffen zur Daimler AG gehörigen nordamerikanischen LKW Entwicklung von Freightliner nach Portland, Oregon. Dort ist er seither für die HMI und Telematik Hard- und Software Entwicklung der zukünftigen elektrisch angetriebenen LKWs zuständig.

Inkrementelle Verarbeitung als Hebel für responsive Dialogführung

Timo Baumann
Universität Hamburg

Menschen im Dialog verarbeiten Sprache schritthaltend „just-in-time“ wobei die notwendigen Verstehens-, Verarbeitungs-, und Generierungsprozesse nebenläufig ablaufen und dennoch flexibel miteinander interagieren. Im Unterschied dazu nutzen typische Dialogsysteme eine einfache Verarbeitungspipeline, die nur eine Ping-Pong-artige Interaktion ermöglicht: entweder zuhören oder sprechen – und zwischendurch muss der Nutzer warten während die Maschine denkt. Der Vortrag präsentiert eine Architektur welche konversationalen Agenten eine flexiblere Interaktion ermöglicht, sowie diverse Anwendungsfälle in denen auf der Architektur basierende Systeme bereits deutliche Mehrwerte zeigen konnten.

Timo Baumann forscht und lehrt als Postdoctoral Researcher in der Language Technology Group am Fachbereich Informatik der Universität Hamburg. Zuvor arbeitete er zwei Jahre am Language Technologies Institute der Carnegie Mellon University in Pittsburgh an einem Industrieprojekt zu konversationellen virtuellen Agenten und promovierte bis 2013 an der Universität Bielefeld zu schritthaltender Sprachdialogverarbeitung. Seine Forschungsinteressen liegen im Bereich gesprochensprachlicher Mensch-Maschine- und Mensch-Mensch-Interaktion, der Erforschung der gesprochenen Sprache mit einem Fokus auf Prosodie, und den Digital Humanities. Timo fährt bei jedem Hamburger Wetter mit dem Fahrrad und engagiert sich in einer Energiegenossenschaft.

Smart-Home Steuerung mit einem hausinternen Sprachassistenzsystem

*Hans-Günter Hirsch, Alexander Micheel, Jan Stähler, Michael Gref,
Janik Göbel
Hochschule Niederrhein*

Es wird ein Sprachassistenzsystem vorgestellt, mit dem im Gegensatz zu den Systemen von z.B. Amazon, Apple oder Google die Spracherkennung und -steuerung lokal in einem häuslichen oder firmeninternen Umfeld stattfindet. Das Sprachsignal wird nicht zu einem externen Server übertragen, so dass die Privatsphäre der Benutzer gewahrt bleibt. Dies bedingt allerdings beim Einsatz in einem Privathaushalt oder einer Firma die Notwendigkeit einer kostengünstigen Realisierung des Servers. Es wird ein Assistenzsystem vorgestellt, das aus kleinen kompakten Client Systemen besteht, die beispielsweise in handelsübliche Unterputzdosen oder in Schaltsteckdosengehäusen integriert sind. Die Systeme übernehmen im Gegensatz zu den Client Systemen kommerzieller Anbieter nicht nur die Aufnahme und Wiedergabe der Sprache, sondern auch die Dialoggestaltung und das unmittelbare Aussenden von Steuerungsbefehlen. Dies besitzt neben dem Aspekt, dass man keinen zusätzlichen Server für die Dialoggestaltung benötigt, den Vorteil, recht einfach einen beispielsweise für bestimmte Räume oder Clients eingerichteten Dialog integrieren zu können. Zur sprachlichen Aktivierung des Dialogs wird ein Algorithmus zur Erkennung eines gesprochenen Schlüsselworts vorgestellt, der die begrenzten Rechenressourcen der Client Systeme berücksichtigt. Die Kommunikation der Clients mit dem Erkennungsserver und einem eventuell vorhandenen Smart Home Server findet über WLAN oder LAN statt. Es werden drei Realisierungen des Erkennungsservers vorgestellt, a) die eigene Realisierung eines GMM-HMM basierten Erkenners auf einem Raspberry PI, b) die eigene Realisierung eines DNN-HMM basierten Erkenners auf einem NVidia Board und c) die Implementierung eines auf Kaldi beruhenden Erkenners auf einem kostengünstigen Rechner. Mit den unterschiedlichen Systemen können Ansprüche abgedeckt werden, die von der Erkennung einiger Kommandowörter bis zur Erkennung großer Vokabularien reichen. Eine Demonstration verschiedener Client- und Serverkomponenten im Kontext der Anwendung zur Smart Home Steuerung wären im Rahmen des Workshops möglich.

Hans-Günter Hirsch ist seit 2001 Professor an der Hochschule Niederrhein und seit 2010 Leiter des Instituts für Mustererkennung, seit 1982 ist er im Arbeitsgebiet der Sprachverarbeitung/Spracherkennung tätig.

Sprachassistenten: Unsichtbare und unerwartete Gäste am Tisch?

Delphine Reinhardt
Universität Göttingen

Sprachassistenten haben unsere Welt bereits erobert. Alexa ist schon bei manchen eingezogen. Fast jedes Handy ist mit Siri oder dem Google Assistent ausgestattet. Durch ihre Sprachsteuerung bieten diese neuartigen Assistenten neue Möglichkeiten, um Informationen online zu finden oder Aufgaben zu erfüllen. Besonders für Personen, die eine Tastatur nicht benutzen können, bieten sie eine interessante Schnittschnelle, um mit der digitalen Welt zu interagieren. In unserer Tasche, an unserer Seite auf dem Tisch oder im Wohnzimmer begleiten sie uns in unserem Alltag. Sie bleiben meist unsichtbar, außer wenn sie sich unerwartet und ungefragt in einem Gespräch zu Wort melden. Nach dem ersten Überraschungsmoment können sich die Teilnehmer des Gesprächs ausgespäht fühlen. In meinem Vortrag werde ich deswegen mögliche Risiken für die Privatsphäre der Besitzer von Sprachassistenten sowie Dritten erläutern. Diese Analysis wird sich nicht nur auf der technischen Ebene bewegen sondern wird auch die Sicht der Nutzer widerspiegeln. Darüber hinaus werde ich mögliche Ansätze, um den Schutz der Privatsphäre dieser Personen zu gewährleisten, betrachten und Förderungen für die Zukunft vorstellen.

Delphine Reinhardt ist W3-Professorin und leitet seit 2018 die Forschungsgruppe „Computersicherheit und Privatheit“ an der Georg-August-Universität in Göttingen. In 2019 wurde sie als eine der 10 weltweiten „Rising Stars in Networking and Communications“ durch N2Women ernannt und wurde mit dem Johann-Philipp-Reis-Preis für herausragende innovative Veröffentlichungen ausgezeichnet. Zwischen 2014 und 2018 war sie Junior-Professorin an der Rheinischen Friedrich-Wilhelms-Universität in Bonn und leitete auch eine Gruppe bei Fraunhofer FKIE. Sie absolvierte ihre Promotion mit Auszeichnung an der Technischen Universität Darmstadt in 2013. Ihre Dissertation wurde mehrmals für herausragende akademische Leistungen ausgezeichnet. Sie erwarb einen doppelten Abschluss in Elektrotechnik an der TU Darmstadt und der Ecole Nationale Supérieure de l’Electronique et ses Applications in Frankreich in 2009.

Usability von Sprachassistenten im Smart-Home Bereich

Stefan Hillmann
TU Berlin

Die Akzeptanz eines Sprachassistenten/Conversational Agent (CA) durch den Nutzer hängt wesentlich von der Usability des CA ab. Die drei häufigsten Szenarien zur Nutzung von CA sind die Interaktionen mit einem Smart Phone/Tablet, einem Smart-Home und mit dem Auto. Der Vortrag gibt anhand des Smart-Home Szenarios eine Einführung in Faktoren welche die Usability von CA beeinflussen, als auch Methoden zur Usability-Evaluation von Sprachassistenten. Es werden die für eine Usability-Evaluation wesentlichen Interaktionsparameter erläutert, sowie passende Fragebögen und deren Einsatz in empirischen Untersuchungen vorgestellt. Da die Interaktion mit CAs heutzutage selten rein sprachbasiert ist, sondern gewöhnlich weitere Modalitäten unterstützend verwendet werden, wird auch auf multimodale Interaktionen eingegangen.

Stefan Hillmann forscht seit 2010 zur Usability-Evaluation von Sprachdialogsystemen und multimodalen Dialogsystemen, davor studierte er Informatik (Dipl.-Inf.) und arbeitete als Softwareentwickler. Seit 2017 ist er Post-Doc am Quality and Usability Lab der TU Berlin und koordiniert die Arbeitsgruppe Speech and Dialog. Im Rahmen seiner Dissertation beschäftigte er sich insbesondere mit Nutzersimulationen und deren Verwendung zur automatischen Usability-Evaluation. Im Rahmen von 2 BMBF-geförderten Projekten (Smart Senior und Usability @ Universal Home Control Interfaces) beschäftigte er sich mit der Konzeption und Evaluation sprachbasierter Mensch-Maschine-Interaktion im Smart-Home.

Sein aktuelles Forschungsinteresse betrifft die Simulation von Nutzerverhalten mittels maschinellen Lernens als auch das Trainieren von Dialogmanagern auf der Basis von Reinforcement Learning.

Seit 2019 ist er Gastwissenschaftler des Forschungsbereiches Speech and Language Technology des Deutschen Forschungsinstitut für Künstliche Intelligenz.

Dialogsysteme mit Gefühlen: Wie, Warum und Wohin?

Felix Burkhardt
Audeering GmbH

Erklärungsversuche für emotionales Verhalten sind so alt wie die Menschheit und auch der Zusammenhang mit intelligentem Verhalten wird seit längerem untersucht. Seit es Maschinen gibt wird auch hier der Umgang mit emotionalem Verhalten betrachtet, allerdings sind moderne Dialogsysteme wie z.B. Alexa oder Siri eher selten dazu in der Lage, Ironie zu erkennen oder auf Ungeduld zu reagieren. Der Vortrag gibt eine Übersicht über psychologische Modelle, technische Umsetzung, Anwendungen und ethische Aspekte von emotionaler Verarbeitung in der Mensch-Maschine Interaktion.

Felix Burkhardt widmet sich mittlerweile seit mehr als 20 Jahren dem Gebiet der Text- und Sprachverarbeitung, speziell für intelligente Sprachanalyse und Ermittlung von Emotionen aus der menschlichen Stimme. Als Research Director bei dem KI-Innovationsführer im Bereich intelligenter Sprachanalyse *audEERING* leitet Burkhardt ein Team von zehn Forschern und verantwortet acht Projekte, die auf Bundes- und EU-Ebene gefördert werden. Im Fokus steht dabei die AI-basierte Erkennung von Emotionen und akustische Szenenanalyse in unterschiedlichsten Themenfeldern wie beispielsweise Depressionserkennung bei Heranwachsenden, Roboterinteraktion bei Kindern mit Autismus sowie Angsttherapien mithilfe von Virtual Reality. Den Großteil seiner Fachkompetenz entwickelte Felix Burkhardt in 18 Jahren als Technologieexperte innerhalb der Deutschen Telekom. Darüber hinaus war er als Experte für Sprachsynthese an der Technischen Universität Berlin beschäftigt. Im Rahmen seiner Doktorarbeit widmete er sich der Simulation emotionaler Sprache durch Maschinen. Noch heute ist er nebenberuflich als Dozent in dem Bereich der auditiven Mensch-Maschine-Interaktion tätig.

Chatbots mit Persönlichkeit

Jan Nehring
DFKI Berlin

User Experience wird nicht allein von Effizienz bestimmt. Freude und Faszination sind ebenfalls wichtige Faktoren. Chatbots sollen darum mehr sein als nur nützliche Helfer. Präsentiert werden Erfahrungen aus der technischen Implementierung von Chatbots kombiniert mit Erfahrungen aus dem Seminar meta:dating an der Universität der Künste in Berlin. In diesem Seminar entwickeln Studierende der Gesellschaft- und Medienkommunikation Chatbots. Vorgestellt werden Methoden, um Chatbots mit Persönlichkeit zu entwickeln.

Jan Nehring Jan Nehring arbeitet als wissenschaftlicher Mitarbeiter am Deutschen Forschungszentrum für künstliche Intelligenz in Berlin. Dort beschäftigt er sich mit Natural Language Processing, Sprachtechnologien und Chatbots. Im Bereich Chatbots betreut er vor Allem Industriepartner bei der Umsetzung ihrer Projekte. Außerdem arbeitet er als Lehrer im Bereich Chatbots an der Technischen Universität Berlin und an der Universität der Künste Berlin. Vor seiner Tätigkeit am DFKI arbeitete Jan Nehring als wissenschaftlicher Mitarbeiter an der Technischen Universität Berlin und als freiberuflicher Softwareentwickler im Web Bereich.

“das ist schon gruselig so dieses Belauschtwerden” —
Subjektives Erleben von Interaktionen mit
Sprachassistenzsystemen zum Zwecke der Individualisierung

Julia Krüger & Ingo Siegert

*Universitätsklinik für Psychosomatische Medizin und Psychotherapie,
Mobile Dialogsysteme, Institut für Informations- und
Kommunikationstechnik, Universität Magdeburg*

Sprachassistenzsysteme werden in Beruf und Alltag immer bedeutsamer. Dem Anspruch moderner HCI folgend, sollen sie individuumszentriert sein. Dies setzt voraus, dass relevante Informationen über den/die Nutzer/in vorliegen. Sind diese weder messbar, noch beobachtbar, ist es notwendig, diese direkt in der Interaktion zu erfragen. Doch wie geht es Nutzer/innen, die von Sprachassistenzsystemen nach persönlichen und privaten Informationen gefragt werden? Im 2011/12 aufgenommenen LAST MINUTE Korpus (LMC) wurde dies mithilfe von Nutzerinterviews untersucht. Es wurde deutlich, dass das Erleben des Systems – verstanden als Resultat nutzerseitiger Zuschreibungen von Charakteristika, Fähigkeiten, Zielen etc. – vielgestaltig ist und Zusammenhänge zum nutzerseitigen Selbsterleben in der Interaktion, insb. auch im subjektiv reflektierten Selbstoffenbarungsgrad, bestehen. Das im LMC mittels Wizard-of-Oz-Technik simulierte technische System war hierbei durch eine artifizielle männliche Stimme sowie eine Kombination aus Bildschirm, Kameras und Mikrofonen repräsentiert. In der vorliegenden Arbeit wurde nun untersucht, inwiefern sich das subjektive Erleben von Nutzer/innen verändert, wenn sie einen weitgehend identischen Dialog durchlaufen, das System dabei jedoch ein ALEXA Skill mit einer menschlichen, männlichen Stimme, repräsentiert durch ein heutiges ALEXA Echo von Amazon, darstellt. Hierfür wurde eine Kurzversion des im LMC angewendeten Interviewleitfadens genutzt. Erste Ergebnisse der qualitativen Inhaltsanalyse dieser Interviewdaten werden präsentiert und mit denen des LMC verglichen. Sie deuten u.a. darauf hin, dass ALEXA überraschenderweise klarer als technisches Gegenüber erlebt wird als das System im LMC und eine ausgeprägtere Skepsis bezüglich der Weiterverarbeitung der preisgegebenen Daten herrscht. Die Ergebnisse werden unter Bezug auf die Limitationen der Untersuchung diskutiert.

Julia Krüger Diplom-Psychologin, ist sowohl wissenschaftlich als auch psychotherapeutisch an der Universitätsklinik für Psychosomatische Medizin und Psychotherapie der Otto-von-Guericke-Universität Magdeburg tätig. In ihrer Forschung beschäftigt sie sich vornehmlich mit subjektiven Zuschreibungen von Nutzer/innen gegenüber intelligenten technischen Systemen und verwendet hierbei Methoden der qualitativen Sozialforschung.

Ingo Siegert ist seit November 2018 Juniorprofessor für Mobile Dialogsysteme an der Fakultät für Elektrotechnik und Informationstechnik der Uni Magdeburg. In seiner Forschung beschäftigt er sich mit intelligenten Dialogführungsstrategien unter Ausnutzung prosodischer Sprachmerkmale für Alexa Siri und Co sowie der Sprachsignalanalyse in mobilen Umgebungen. Hierbei nutzt er vielfältige Methoden des maschinellen Lernens und der Signalverarbeitung.

Alexa, kannst Du mich verstehen? Empirische Langzeituntersuchung zur Interaktion zwischen Sprachassistenten und Kindern

*Jessica Szczuka, Stefan Kopp, Nicole Krämer, Lina Varonina
Universität Duisburg-Essen, Universität Bielefeld*

Sprachassistenten halten mehr und mehr Einzug in die Lebensräume der Besitzer und somit letztendlich auch in den der Familienmitglieder. Dies ermöglicht, dass auch Kinder Zugang zu den nicht-menschlichen Kommunikationssystemen haben. Erste Untersuchungen zur Interaktion zwischen Kindern und Sprachassistenten zeigen, dass Kinder (in Abhängigkeit deren Entwicklungsstufe) naive Techniken nutzen, um herauszufinden, wie Alexa, Siri und Co einzuschätzen sind. Anders als Erwachsene, die bereits über ein mentales Modell, also über eine Vorstellung über das System verfügen, nutzen Kinder Fragen wie „[Alexa], what is your favorite color“, „Can you open doors?“ oder erkundigen sich, ob der Assistent essen möchte, um dessen (soziale) Identität und Intelligenz einzuordnen. Daraus ergeben sich einerseits Fragen dazu, wie gut eine Interaktion zwischen Sprachassistenten und Kindern überhaupt funktionieren kann, da die stark domänen-spezifischen Sprachsysteme nur bedingt auf solche Fragen eingerichtet sind. Außerdem ist aus psychologischer Sicht zu analysieren, welchen Einfluss solche Interaktionen auf das mentale Modell, das Kinder sich vom Wesen der Sprachassistenten machen, haben. Ein weiterer wichtiger Bereich ist die Betrachtung der Beziehungsbildung. Hier ist bekannt, dass Kinder durchaus innige Beziehungen zu nicht-menschlichen Entitäten unterhalten. Forschung zum Prozess des Anthropomorphismus zeigt, dass die Fähigkeit, Lebendigkeit in Dingen zu sehen, mit zunehmendem Alter abnimmt, da sie durch Reflexionsvorgänge gehemmt wird. Vor diesem Hintergrund bleibt zu prüfen, wie stark Kinder Beziehungen ausbilden und inwieweit sie Sprachassistenten als „Freunde“ empfinden. Darüber hinaus kann erwartet werden, dass – wie bei Erwachsenen – auch unwillkürliche soziale Reaktionen im Sinne der Media Equation Theorie auftreten, da das Sprachassistentensystem natürlichsprachlich interaktiv interagiert und die Rolle einer weiblichen Helferin (sucht z.B. Informationen, führt bestimmte Befehle aus) ausfüllt. Die vereinzelt Studien, welche bisher durchgeführt worden sind, haben gemeinsam, dass diese sich auf sehr kurzfristige Interaktionen im Labor beziehen und daher nur kurzfristige Reaktionen, anstatt die Entwicklung von Interaktionen unter realweltlichen Bedingungen betrachten (z.B. Tätigen Kinder zu Beginn andere Anfragen als nach einigen Wochen der Nutzung? Unter welchen Bedingungen verlieren Sie das Interesse an der Interaktion?). Ebenfalls offen ist, inwieweit und in welchem Ausmaß Kinder Beziehungen zu Sprachassistenten entwickeln. Vor diesem Hintergrund präsentieren wir eine Langzeituntersuchung, die im Rahmen des interdisziplinären Projektes IMPACT (Implications of conversing with intelligent machines in everyday life for people's beliefs about algorithms, their communication behavior and their relationship building) mit Beteiligung der Disziplinen Informatik, Psychologie, Rechtswissenschaft und Ethik durchgeführt wird. Insgesamt 20 Familien bekommen für 5 Wochen einen Sprachassistenten zur Verfügung gestellt, um zu untersuchen, wie sich Interaktionen zwischen Kindern und Sprachassistenten entwickeln. Dabei werden Kinder und Eltern in regelmäßigen Abständen befragt sowie bei der Interaktion beobachtet. Zusätzlich werden Logfiles der Interaktionen berücksichtigt. Hierbei wird ein besonderer Fokus auf

das mentale Model der Kinder in Hinblick auf den Sprachassistenten gelegt und wie sich positive aber auch gescheiterte Anfragen auf das Verständnis der Funktionen sowie die Beziehungsbildung auswirken. Vor dem Hintergrund rechtswissenschaftlicher Überlegungen zur digitalen Selbstbestimmung wird ebenfalls erfragt, inwieweit ein Wissen darüber besteht, dass Daten gesammelt und gespeichert werden. Erste Analysen können auf dem Workshop vorgestellt werden.

Jessica Szczuka arbeitet als Postdoktorandin am Lehrstuhl für Sozialpsychologie: Medien und Kommunikation der Universität Duisburg-Essen. Hier hat sie zuvor den Bachelor und Master of Science in dem durch Psychologie und Informatik geprägten Studienfach angewandte Kognitions- und Medienwissenschaft erlangt. Dr. Szczuka hat sich innerhalb ihrer Promotion mit sexuellen Aspekten der Mensch-Roboter-Interaktion beschäftigt und arbeitet auch weiterhin im Forschungsbereich der digitalisierten Sexualität bzw. der Kommunikation mit artifiziellen Interaktionspartnern. Innerhalb des IMPACT Projekts engagiert sich Dr. Szczuka in der Koordination des Projekts und konzipiert bzw. leitet die durchzuführenden empirischen Untersuchungen.

Stefan Kopp ist Professor für Informatik an der Technischen Fakultät der Universität Bielefeld, Leiter der Gruppe Soziale Kognitive Systeme und Hauptforscher des Kompetenzzentrums „Cognitive Interaction Technology“ (CITEC). Er ist anerkannt als internationaler Experte auf dem Gebiet der Gesprächsagenten, der multimodalen Mensch-Maschine-Interaktion und der kognitiven Modellierung sozialer Interaktion. Prof. Dr.-Ing. Kopp koordinierte bereits zahlreiche Projekte zur Mensch-Maschine-Interaktion mit intelligenten Systemen, wie z.B. sozialen Robotern, mit virtuellen Menschen oder Avataren sowie mit multimodalen Dialogsystemen. Seine Expertise ist hochgradig interdisziplinär und verbindet kognitionswissenschaftliche sowie psychologische Theorien mit empirischen Untersuchungen des menschlichen Kommunikationsverhaltens und der Entwicklung von KI-Techniken zur Realisierung sozial intelligenter Systeme.

Nicole Krämer leitet den Lehrstuhl für Sozialpsychologie: Medien und Kommunikation an der Universität Duisburg-Essen. Mit 20 Jahren Erfahrung ist sie eine ausgewiesene Expertin auf dem Gebiet der Mensch-Computer-Interaktion, vor allem im Bereich der Interaktion mit intelligenten, virtuellen Agenten und Robotern. Sie leitete bereits vielzählige, besonders ebenfalls interdisziplinäre Forschungsprojekte. Unter Verwendung einer großen Anzahl von unterschiedlichen, empirischen Forschungsmethoden fokussieren ihre Arbeiten sozialpsychologische Effekte der Interaktion mit Maschinen und analysieren damit, wie Menschen auf intelligente, künstliche Entitäten in Abhängigkeit von deren Fähigkeiten sowie Verhalten reagieren.

Lina Varonina ist wissenschaftliche Mitarbeiterin ist wissenschaftliche Mitarbeiterin in der AG Social Cognitive Systems an der Universität Bielefeld. Sie hat Wirtschaftsinformatik im Bachelor an der Hochschule Hannover und Intelligente Systeme im Master an der Universität Bielefeld studiert. Sie war 2012 bis 2013 Tutorin für Mathematik an der Hochschule Hannover und zweimal Praktikantin beim Fraunhofer-Institut für Windenergiesysteme (IWES), in den Jahren 2013 und 2015, wo sie auch ihre Bachelorarbeit über die Architektur der Software für Simulationen der Instandhaltungsmaßnahmen für

Offshore- Windparks geschrieben hat. Ihre Forschungsinteressen liegen im Bereich von intelligenten Dialogsystemen, insbesondere wenn es um die Erhöhung der Lebensqualität und Autonomie mithilfe solcher Systeme geht.

Audio Security & Privacy

*Andreas Nautsch, Massimiliano Todisco, Jose Patino and Nicholas
Evans*
EURECOM

Smart speakers, virtual assistants and interactive voice response systems successfully integrated as a part of our daily life. Intuitively, we use speech as a medium to seamlessly impart and exchange information for a plethora of services. Speech applications are across all domains from personal assistants, smart home and call centers to online banking, health care as well as forensic sciences.

Repurposing speech data is a potential threat to privacy infringement for speech being a very rich medium. Voice data, however, cannot only be exploited to characterize individuals but also to subvert system security: countermeasures of voice biometric systems are in high demand for detecting spoofing attacks on voice biometrics (e.g., payment services rely on biometric identity proofs). An overview of the audio security and privacy lab at EURECOM is provided: results of co-organizing the biannual ASVspooft challenges (attack countermeasures); collaborations with legal and cryptography experts, and co-lead activities of the newly formed ISCA special interest group on security and privacy in speech communication.

Security. The ASVspooft editions investigated different settings: speech synthesis attacks (2015), replay attacks in the wild (2017), and voice conversion and replay attacks (2019). For its 2019 edition, the primary evaluation changed from the equal error rate to a decision risk placing a tandem emphasis on the costs resulting from decisions misled by either the voice biometric and the countermeasure subsystem. Investigating on different environments, countermeasures operated well in settings known for poor voice biometric performance.

Privacy. The emerging field is highly interdisciplinary (speech, legal, cryptography). A common understanding of technology and legal experts is necessary for carrying out data protection impact assessments, meaningful. A mutual understanding of methodology is necessary for harmonizing used terminology. When providing safeguards, complexity analyses demand computational approximations which lead to poor estimates of decision risks, in turn. A discourse in navigation.

Andreas Nautsch is with the Audio Security and Privacy research group (EURECOM). He received the doctorate from TU Darmstadt in 2019, B.Sc. and M.Sc. degrees from Hochschule Darmstadt in 2012 and 2014, respectively. From 2014 to 2018, he was with the biometrics group within the German National Research Center for Applied Cybersecurity. He served as an expert delegate to ISO/IEC and as project editor of the ISO/IEC 19794-13:2018 standard. He is a co-organizer of the ASVspooft 2019 evaluation, the related special sessions at Interspeech 2019 and ASRU 2019, and is a Guest Editor of the related CSL special issue. Andreas is a co-initiator and secretary of the ISCA Special Interest Group on Security Privacy in Speech Communication.

Sprachbasierte Self-Services auf Basis etablierter und kostengünstiger Cloud-Dienste am Beispiel einer vollautomatisierten Zählerstandserfassung

Marcus Petersen, Martin Suffa, Oliver Neumann, Tino Köhler, Florian Marquardt
regiocom SE

Die regiocom SE realisiert im Geschäftsfeld Business Process Outsourcing (BPO) komplette Geschäftsprozesse für Mandanten aus dem Energieversorgungsbereich. Für diese Mandanten ist das Entgegennehmen, Validieren und Verbuchen von Zählerständen essentieller Bestandteil des Geschäfts und zugleich ein wesentlicher Mengentreiber im Kundenkontakt. Die Entgegennahme der Zählerstände erfolgt über diverse Kanäle. In der Regel sind dies Brief, Mail, Fax, Webportal, App und Telefon. Das Telefon stellt dabei den kostenintensivsten Eingangskanal dar, sodass hier eine weitgehende Automatisierung angestrebt wird. In der Vergangenheit wurden hierfür Telefonautomaten auf Basis von Tastatureingaben (DTMF-basiert) eingerichtet, die Zählerstände vollautomatisiert verarbeiten können. Der Einsatz solcher Automaten steigert Kundenzufriedenheit und Servicequalität, indem die Kunden den Zählerstand ohne Wartezeit und rund-um-die-Uhr selbst übermitteln können.

Der Telefonautomat auf Basis von Tastatureingaben entspricht nicht mehr dem heutigen Stand der Technik. Sprachbasierte Dienste wie Siri, Cortana und Bixby halten Einzug in unseren Alltag und erleichtern zum Beispiel das Abrufen von Informationen oder das Steuern von (Haus-)Technik. Durch den Einsatz dieser Dienste in immer mehr Geräten, wie dem Auto oder universell-einsetzbarer Hardware (z.B. die Echo-Serie von Amazon) nimmt die Durchdringungstiefe stetig weiter zu. Dadurch eröffnen sich völlig neue Eingangskanäle, die moderne Unternehmen zukünftig anbieten müssen. Für unsere Mandanten ergibt sich aus dieser Entwicklung ein Spannungsfeld: Auf der einen Seite erwarten die Kunden denselben technischen Standard, den sie aus ihrem privaten Umfeld gewohnt sind; auf der anderen Seite sehen sich die Firmen einem enormen Kostendruck am Markt gegenübergestellt.

Vor diesem Hintergrund haben wir eine sprachbasierte Zählerstandserfassung auf Basis etablierter und kostengünstiger Dienste aus der Cloud entwickelt, die zukünftig über diverse Sprachkanäle angeboten werden soll. Der Cloud-Telefonie-Spezialist Tenios (www.tenios.de) nimmt in der ersten Ausbaustufe die eingehenden Telefongespräche entgegen und leitet die transkribierten Texte (Speech-to-Text) an Google Dialogflow (dialogflow.com) weiter. In Dialogflow wurde der Gesprächs- resp. Programmablauf mit der notwendigen Logik umgesetzt. Hier erfolgen Intenterkennung, Backendintegration und Antwortgenerierung. Die Antworten werden textbasiert an Tenios zur Ausgabe (Text-to-Speech) zurückgegeben. Durch den Einsatz etablierter Cloud-Dienste wird die Anforderung nach einer hochverfügbaren und skalierbaren Lösung erfüllt. Die Pay-per-Use-Modelle der Dienste reduzieren die Investitionskosten und ermöglichen eine flexible Preisgestaltung. Das System ist Multiprozessfähig. Die Zählerstandserfassung dient als Demonstrationsbeispiel.

Marcus Petersen verantwortet seit April 2019 als Product Owner das Team Sprachassistentensysteme der regiocom SE. In dieser Funktion dient er internen und externen Sta-

keholdern als Ansprechpartner für alle Themen und Projekte rund um das Erkennen, Verstehen und Verarbeiten natürlicher Sprache. Bei der regiocom SE engagiert er sich seit 2015 in unterschiedlichen Positionen. So hat er in dieser Zeit bspw. das Thema Robotics / Robotic Process Automation (RPA) mitaufgebaut und als festen Bestandteil im Firmenverbund etabliert. Neben der Softwareentwicklung bilden seine Aktivitäten im Bereich Software-, System- und Unternehmensarchitektur einen weiteren Schwerpunkt seiner Tätigkeiten. Nach dem Informatikstudium an der Technischen Universität Clausthal in Clausthal-Zellerfeld, arbeitete er knapp 4 12 Jahre als wissenschaftlicher Mitarbeiter in der Fachgruppe Produktentstehung am Heinz Nixdorf Institut der Universität Paderborn.

Intelligent LSF-Answering System - An Alexa Skill

*Adarsh Kuzhipathalil, Anto Thomas, Keerthana Chand, Frank Ortmeier,
Ingo Siegert
Uni Magdeburg*

Currently in the market, there are many voice assistants performing an intended action by the user, furthermore voice platforms allow developers to create their own speech based applications. The purpose of the presented project was to develop a new Alexa Skill as a Voice User Interface (VUI) to the LSF web portal of the Otto-von-Guericke University, Magdeburg. LSF is a student information system, accessible via web, but hardly optimized for mobile devices. It provides information about teaching, study courses, and the structure of the organizational units and university members.

By developing a VUI for LSF we aim to improve the accessibility, and getting the information in an efficient manner – on the example of course information for the study course Digital Engineering.

To implement the whole concept, we used the well known Amazon Alexa platform. This platform is designed in such a way that one can use the whole potential of it through its application programming interface (API). The LSF skill is implemented on the Alexa Platform with the help of other cloud based Web Services from Amazon.

The developed skill was tested with 20 users in terms of accuracy, robustness, and usability. The functionality test was followed by an attractiveness test (ATTRAKDIFF). The test showed that the VUI worked well with all of its designed functions, even though some particular utterances from users wasn't recognised by the system, mainly due to the accent of non-native English speakers. The overall success rate is 90%. The reason for minor failures are also studied as part of the test.

Adarsh Kuzhipathalil, Anto Thomas, Keerthana Chand are students of the Digital Engineering study course at the Faculty of Computer Science at Otto von Guericke University Magdeburg. The presented work was conducted as an interdisciplinary team project under the supervision of Prof. Ortmeier and Jun.-Prof. Siegert and part of their study course.