



In immer mehr Alltagsgegenständen wird Sprachverarbeitung eingesetzt. Dabei können Gegenstände gleichzeitig Sender und Empfänger sein. Sprachverarbeitung kann aktuell zwar schon vielfältig eingesetzt werden, beispielsweise für Übersetzung und Recherche, im Automobil und in der Produktion. Allerdings steht ihr Gebrauch noch am Anfang. Dieses stand bei dem digitalen Format „Alles spricht. Assistenzsysteme mit Sprachverarbeitung“ im Rahmen der Veranstaltungsreihe ZIRPzoom – Perspektiven der Wirtschaft in Rheinland-Pfalz im Mittelpunkt unter Leitung von Staatssekretärin Heike Raab, Bevollmächtigte des Landes Rheinland-Pfalz beim Bund und für Europa, Medien und Digitales. Prof. Dr. Sven Pagel, Hochschule Mainz, Dr. Aljoscha Burchardt, DFKI, sowie Frank Jordan, Stellantis N.V. (vorher Opel Automobile GmbH), erörterten die Sicht der Wissenschaft, die Systematik hinter den Anwendungen und den sinnvollen Einsatz von Sprachverarbeitungssoftware sowie daraus resultierende neue Geschäftsmodelle.

Neue Plattformen als Gatekeeper

Im Medienstaatsvertrag, seit November 2020 in Kraft, wurde das deutsche Medienrecht auf die aktuelle Medienlandschaft angepasst. Staatssekretärin Heike Raab, Bevollmächtigte des Landes Rheinland-Pfalz beim Bund und für Europa, Medien und Digitales, hat maßgeblich an seiner Ausarbeitung mitgewirkt und erläuterte die ihm zugrundeliegenden Ansprüche: „Das Modell des klassischen Rundfunks ist längst diversen öffentlich und privaten Produktions- und Sendemodellen gewichen. Der Medienbegriff hat sich erweitert. Mittlerweile können technische Anwendungen sowohl Empfänger als auch Sender von Informationen sein, die mittels Sensorik, QR-Codes, Face-ID und KI generiert werden“, so Staatssekretärin Heike Raab. Neue Plattformen seien zu den Gatekeepern der digitalen Welt geworden, indem sie den Zugang zu Inhalten regulieren. Die Beliebtheit von Geräten mit Sprachsteuerung und -ausgabe liege an ihrer niederschweligen und barrierefreien Nutzung. Solche

Anwendungen könnten das Leben von Menschen erleichtern. Hierbei seien Spracherkennungsmodulle im rundfunkrechtlichen Sinne keine Medien und dennoch stelle ihre Basis die Nutzung von Medien dar. Ein besonders erfolgreiches Beispiel für die Verbesserung von Medienanwendungen mittels KI sei der EU Council Presidency Translator, der unter Mitarbeit des Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI) im Rahmen der deutschen EU-Ratspräsidentschaft 2020 entwickelt wurde. Staatssekretärin Heike Raab führte aus:

„Es war uns besonders wichtig, die Vielfalt, die Meinungsfreiheit und die Transparenz bei Intermediären wie Smart Speakern sicherzustellen. Die Orientierung an Reichweitenstärke sowie nutzer- und verbraucherfreundliche Einstellungen müssen gewährleistet werden, damit journalistische und redaktionelle Inhalte gefunden werden können und die Logik der Auswahl transparent bleibt.“

Zudem biete die Entwicklung und der Einsatz von Anwendungen mit Sprachverarbeitung ein erhebliches Potenzial für unternehmerische Geschäftsmodelle in Rheinland-Pfalz, so die Staatssekretärin.

Nutzer fragt, System antwortet

Prof. Dr. Sven Pagel, Professor für Wirtschaftsinformatik und Medienmanagement an der Hochschule Mainz, lieferte Hintergrundinformationen zum Einsatz von Künstlicher Intelligenz (KI) im Bereich der Sprachverarbeitung. „Sie kann mehrere automatisierte Aktivitäten eigenständig verknüpfen und eine Chance für den Einsatz von Sprachverarbeitung sein“, so Prof. Dr. Sven Pagel. Die Sprachverarbeitung biete sehr viele Chancen, doch würden bisherige Technologien nicht zeitnah abgelöst werden, auf Grund der Nutzungsgewohnheiten der Menschen und dem noch auszuschöpfenden Entwicklungspotenzial im Bereich Spracherkennung und -verständnis der Anwendungen:

„Eingabe und Ausgabe müssen reibungslos funktionieren, um eine intuitive Nutzung zu



„Unser Ziel ist ein geregelter Fluss von Daten, der Kreativität ermöglicht und gleichzeitig Diskriminierung und Einschränkungen von Freiheiten verhindert.“

Staatssekretärin Heike Raab, Bevollmächtigte des Landes Rheinland-Pfalz beim Bund und für Europa, Medien und Digitales



„Sprachverarbeitung kann auf allen Ebenen von Kommunikation eingesetzt werden: Von der direkten Individualkommunikation bis hin zur Massenkommunikation. Es handelt es sich jeweils um eine einzelne Aktivität, die mit KI umgesetzt wird.“

Prof. Dr. Sven Pagel, Professor für Wirtschaftsinformatik und Medienmanagement, Hochschule Mainz

ermöglichen. Nutzer:innen sollen irgendeine Frage stellen können, das System übersetzt das Audiosignal, gleicht die Informationen mit Datenbanken ab und formuliert mittels Sprachausgabe eine Antwort“.

Als Anwendungsbeispiele stellte Prof. Pagel die Unternehmensgründung LEVOOBA Kids vor, die Smart Speaker mit Frage-Antwort-Spielen für Bildung zuhause nutzt sowie den automatisierten Untertitelungsprozess von AI4mediadata für ZDF digital.

Individuelle Sprachsysteme durch pain points identifizieren

Dass der Nutzen von Sprache umgebungs- und situationsabhängig ist, betonte Dr. Aljoscha Burchardt, Forschungsbereichsmanager am Language Technology Lab des DFKI Projektbüro Berlin:

„In bestimmten Umgebungen ist gesprochene Sprach geradezu prädestiniert. Beim Autofahren oder bei der Anamnese beim Zahnarzt. In einer lauten Umgebung wie eine Werkshalle ergibt es keinen Sinn, mit Sprachbefehlen und -ausgabe zu arbeiten“.

Sprachverarbeitung kann auf zwei unterschiedlichen Mechanismen beruhen: datenbasiert oder wissensbasiert. „Das maschinelle Übersetzen basiert auf dem Training mit schon bestehenden Datensätzen. Neuronale Netze werten beispielsweise Übersetzungsdaten aus und nutzen diese, um neue Texte zu übersetzen“, veranschaulichte er. Im Gegensatz hierzu würden smarte Assistenten regelbasiert über Vorgaben programmiert. Unternehmen, die Sprachtechnologie einführen möchten, rät Dr. Burchardt sogenannte pain points, also wirkliche Bedürfnisse, zu identifizieren, bei denen Sprachverarbeitung eine Lösung ist. Der nächste Schritt ist die Suche nach Daten, die für das Trainieren eines Systems und/ oder die Identifikation von Regelwerken genutzt werden könnten. Als ein positives Beispiel verwies Dr. Burchardt auf die Nutzung von aufgezeichneten Kundendialogen eines Computerherstellers für die Entwicklung eines automatisierten Help Desks. Nach Auswertung der Daten konnte die Anwendung so trainiert werden, dass sie Anfragen eigenständig beantworten kann.

Durch KI-Sprachsteuerung mehr Fahrspaß und -sicherheit

Das KI der enabler für gute Sprachsteuerung sei, betonte Frank Jordan, Leiter der Abteilung Forschung & Innovation bei der Stellantis N.V.: „Erst durch den Einsatz von Künstlicher Intelligenz ist es uns möglich, alle Potenziale der Sprachtechnologie auszuschöpfen und sie auch für Kund:innen attraktiv zu machen.“ Erwartung der Kunden ist, dass das Produkt Freude bereitet und die Handhabung erleichtert. Um die Nutzung des Autos intuitiver und individueller zu machen, sei das Ziel nicht nur Dialekte zu verstehen, sondern auch die Fahrererkennung anhand der Stimme. Allerdings: Nur weil bestimmte Anwendungen technologisch möglich seien, empfinde man die Sprachausgabe nicht automatisch als Bereicherung. Auch visuelle Signale als Bestätigung von Fragen oder Anweisungen können bevorzugt sein: „Zu viel Sprachausgabe wird eher als störend empfunden. Es gehe daher mehr um die Sprachverständlichkeit der Kunden und Kundinnen.“ Der Schwerpunkt der Entwicklung von sprachgesteuerten Anwendungen liege zudem auf dem Aspekt der Sicherheit:

„Wir wollen das Potenzial der Sprachsteuerung nutzen, um die Fahrsicherheit zu maximieren. Autofahrer:innen sollen Navis und andere Anwendungen bedienen können und gleichzeitig die Hände am Lenkrad und den Blick auf der Straße haben.“

In der Diskussion ging es vor allem um den zukünftig möglichen Einsatz „sprechender“ Assistenzsysteme in Produktion, Verwaltung und Alltag. Staatssekretärin Heike Raab betonte die Notwendigkeit, technische Möglichkeiten mit persönlichen Wünschen und übergeordneten Werten wie informationelle Selbstbestimmung und Pluralismus in Einklang zu bringen. Spracherkennung ist aus Sicht von Frank Jordan besonders für KMU interessant, beispielsweise in der Produktion. Prof. Dr. Sven Pagel unterstrich die Relevanz der digitalen Bildung: „Nur jemand der fundiert informiert ist, kann aktiv Entscheidungen für oder gegen die Nutzung bestimmter Technologie treffen.“ Dr. Aljoscha Burchardt wies darauf hin, dass Unternehmen zunächst Ressourcen für die Identifikation von Potenzialen und die Entwicklung von Anwendungen bereitstellen müssen.



„Die Nutzung von Sprachanwendungen wie voice search hat erheblich mit den Mediennutzungssituationen und -gewohnheiten zu tun. Nur weil Anwendungen zur Verfügung stehen, heißt es noch lange nicht, dass Nutzer:innen die technologischen Möglichkeiten für sich als hilfreich empfinden.“

Dr. Aljoscha Burchardt,
Forschungsbereichsmanager am
Language Technology Lab des DFKI
Projektbüro Berlin



„Voice control wird auch in Deutschland zum Standard werden. Daher werden wir in Zukunft immer mehr auf Sprache im Auto setzen und so neue Anwendungsmöglichkeiten schaffen.“

Frank Jordan,
Leiter der Abteilung Forschung &
Innovation bei der Stellantis N.V.