

Autor: MIRJAM HAUCK
Seite: 17

Ressort: Wirtschaft
Ausgabe: Hauptausgabe

¹ von PMG gewichtet 01/2023

² von PMG gewichtet 7/2022

Mediengattung: Tageszeitung
Auflage: 111.238 (gedruckt)¹ 130.027 (verkauft)¹
 131.981 (verbreitet)¹
Reichweite: 0,586 (in Mio.)²

Runter zur Titanic

Ein Informatiker taucht 4000 Meter tief zu dem Schiffswrack

München – In rund 4000 Meter Tiefe im Atlantik liegt das wohl berühmteste Wrack der Welt: die *Titanic*, gesunken auf ihrer Jungfernfahrt im April 1912. Seit der Entdeckung 1985 ist der untergegangene Stahlkoloss ein begehrtes Ziel von Forschungsexpeditionen. Allerdings nicht nur von Meeresbiologen oder Archäologen, wie man vermuten könnte – auch ein Informatiker tauchte dort kürzlich ab. **Alex Waibel, Informatikprofessor am Karlsruher Institut für Technologie (KIT) und der Carnegie Mellon Universität in Pittsburgh/Pennsylvania, fuhr mit einem Forschungs-U-Boot an den Meeresgrund, um herauszufinden, ob die von ihm entwickelte Sprach- und Videotechnologie auch unter Extrembedingungen funktioniert.**

Also in einer Unterwasserwelt mit Massen an Salzwasser, in der Kommunikation nur per Sonar, also mittels akustischer Signale, möglich ist.

Waibel forscht seit mehr als 30 Jahren zu Sprachtechnologie und künstlicher Intelligenz (KI), seine Vision ist es, dass Menschen ohne Sprach- und Kommunikationsbarrieren miteinander reden können. So entwickelte er mit anderen Forschenden zusammen in den frühen 1990er Jahren in Karlsruhe ein Übersetzungsprogramm, das erstmals gesprochene Sprache in Text umwandelte und diese ins Englische oder Japanische übersetzte. „Das war damals eine Sensation“, sagt Waibel. Allerdings verstand der Übersetzer seinerzeit nur 500 Wörter und eine festgelegte Syntax.

Das Projekt habe aber dazu geführt, sagt Waibel, dass man erkannt habe, dass ein leistungsfähiges Übersetzungsprogramm nicht so funktionieren könne, indem man einfach die Sprache codiere. Vielversprechender war vielmehr das Konzept eines neuronalen Netzwerks, das durch künstliche Intelligenz (KI) und maschinelles Lernen, also das Füttern mit immer mehr Daten, neue Fähigkeiten „lerne“. Doch es brauchte noch

einige Jahre, bis die Rechner so schnell wurden, dass sich diese Idee auch tatsächlich umsetzen ließ.

Mitte der 2000er Jahre stellte Waibel einen Simultanübersetzer vor, an dem er maßgeblich beteiligt war, 2009 eine Übersetzungs-App fürs Handy. Sie hieß Jibbiggo und wurde an Facebook verkauft. Waibel ging mit, kehrte aber nach zwei Jahren aber wieder an die Universität zurück: Facebook habe nicht so viel mit der App anzufangen gewusst, sagt er heute.

Das nächste Sprach-Start-up hieß 2015 Kites, schon dem Namen nach eine unverkennbare Ausgründung aus dem KIT. Zum Ziel setzten sich Alex Waibel und sein Mitgründer Sebastian Stücker mehrsprachige Übersetzungen in Echtzeit anzubieten. Mit weiteren Forschenden entwickelten sie ein Programm, das Texte aufzeichnen, sie transkribieren und sofort übersetzen kann. Die Software kam zunächst an der Uni und im EU-Parlament zum Einsatz. Dann kam die Pandemie und mit ihr der Einsatz in Videokonferenzen und digitalen Vorträgen. Und die US-Firma Zoom kaufte das deutsche Start-up 2021. Seitdem lassen sich auch in Zoom-Konferenzen – gegen Aufpreis – die Gespräche in Echtzeit übersetzen und als Untertitel anzeigen.

Ihre Sprachdaten bekommt solche Software aus öffentlich zugänglichen Quellen wie Filmen oder auch aus Vorträgen, bei denen die Sprecher der Weiterverwertung zugestimmt haben. Laut Waibel sind es 60 000 bis 100 000 Stunden Sprache, mit denen ein Programm gefüttert werden muss, damit es gute Ergebnisse liefert. Aber selbst dann funktioniert es nicht immer einwandfrei, zum Beispiel, „wenn Menschen mit starkem Akzent oder Dialekt sprechen“, sagt Waibel. Auch mit Abkürzungen oder Namen kommt die Software nicht immer zurecht: Zoom transkribiert Waibels Nachnamen im Gespräch zu „Wei-

del“ – sehr zu dessen Ärger.

Doch mit dem Zoom-Produkt hat Waibel nichts mehr zu tun, er arbeitet wieder an einem neuen Projekt, er nennt es „Face-Dubbing“. Dabei geht es darum, in Filmen und Videos automatisch in einer anderen Sprache zu synchronisieren, aber so, dass keine andere Stimme, sondern der Originalsprecher zu hören ist und vor allem, dass, was er in der anderen Sprache sagt, genau auf die Lippenbewegungen passt. „Wir wollen das Gesicht so manipulieren, dass es so aussieht, als würde die Person selbst in der anderen Sprache sprechen.“

Dafür sind einige Arbeitsschritte notwendig: Der Computer erkennt die Sprache, wandelt sie in Text um, übersetzt sie, synthetisiert sie in Sprache, die an die „echte“ angepasst werden muss. Und am Ende muss eine KI, die wiederum aus vielen Lippenbewegungen gelernt hat, das im Video umsetzen, dass es nicht roboterhaft oder künstlich aussieht. Dieses Zerlegen in einzelne Schritte war es auch, was Waibel zum Wrack der *Titanic* gelockt hat. Weil dort unten in der Tiefsee mit einem Schiff an der Oberfläche nur Kommunikation per Textnachrichten funktioniert, müssen Videonachrichten zerlegt und über der Wasseroberfläche wieder von einer KI zusammengesetzt werden.

Man müsse, um das zu testen, nicht unbedingt in 4000 Metern Tiefe abtauchen, gibt Waibel zu, aber das Ganze sei auch mehr als eine PR-Spielerei eines Abenteuer-suchenden Wissenschaftlers: „Es gibt genügend Stellen auf der Welt, wo die Bandbreite so schlecht ist, dass nur Text-Übertragung möglich ist“. Mit der neuen Technologie könnte in solchen Gegenden Video-Kommunikation möglich werden.

Noch funktioniert die Technologie nicht einwandfrei, Emotionen in der Stimme lassen sich noch nicht perfekt künstlich nachbauen. Und, sagt Waibel, man habe komplett die Mimik unterschätzt, wenn

ein Mensch nichts sagt. „Man kann das Gesicht nicht einfach einfrieren, das sieht komplett unnatürlich aus.“ Videos, in denen Prominenten und Politikern Wörter in den Mund gelegt werden, die sie nie gesagt haben. Aber in Karlsruhe weiß man um diese Gefahr oder bestimmte Klassifikatoren, die kleine Unreinheiten im Video entdecken.

Mit dieser Technologie lassen sich derzeit noch keine Videos erstellen, die an Deepfakes denken lassen, also an Karlsruhe weiß man um diese Gefahr und arbeitet an Authentifizierungsverfahren – sei es eine Art Wasserzeichen

MIRJAM HAUCK

Abbildung: Bei einem Tauchgang am Wrack der Titanic testete Alex Waibel seine Sprachtechnologie. Foto: Alex Waibel/KIT

Wörter: 867

Urheberinformation: DIZdigital: Alle Rechte vorbehalten – Süddeutsche Zeitung GmbH, München