

«Die Bedenken sollten uns nicht abhalten»

Was die Stimme über uns verrät, berichtet Jarek Krajewski im Interview

Mittels Stimmanalyse identifizieren Unternehmen wie die Swisscom ihre Kunden. Ein enormes neues Geschäftsfeld tut sich gerade auf – doch mit welchen Folgen? Wir haben uns mit dem Wirtschaftspsychologen Jarek Krajewski unterhalten, Gründer und CEO des Instituts für experimentelle Psychophysiologie (IXP) in Düsseldorf.

Sandro Küng

Ein sprachgesteuerter Billettautomat ist doch eine praktische Sache: «Einmal von Zürich nach Bern retour, bitte.» Und schon wird das Ticket ausgedruckt. Oder auch nicht. Im Gehäuse könnte eine Software am Werk sein, die Stimmen analysiert und auf dem Display anzeigt «Sie haben Grippe. Um die anderen Reisenden vor einer Ansteckung zu schützen, bekommen Sie leider kein Billett. Gute Besserung!» Noch gibt es keine solchen Automaten.

Laut Jarek Krajewski wäre es heute technisch kein Problem, anhand der Stimme eine Krankheit zu diagnostizieren. Auch könne die Stimme Aufschluss geben über Charakter und Gefühlslage der Sprechenden.

den. Wie und was unser Körper über uns verrät, sagt uns der Wissenschaftler im Telefoninterview.

Herr Krajewski, wo wird die Stimmanalyse heute bereits eingesetzt?

In erster Linie in der Wissenschaft. Einige Start-ups haben aber schon Anwendungen entwickelt, z.B. im telemedizinischen Bereich und in der Personalrekrutierung und bieten diese am Markt an.

Können Sie uns ein Beispiel nennen?

Mindestens zwei deutsche Firmen arbeiten bereits mit Personaldienstleistern an Stimmanalyse-Tools. Die Personalbranche interessiert sich vor allem für die sogenannten Big Five, also darum, wie erfinderrisch, nachlässig, gesellig, mitfühlend oder neuro-

tisch ein Mensch ist. Die Firma Precire bietet in Deutschland zum Beispiel eine stimmanalytische Personalauswahl-Unterstützung, indem sie eine linguistische und eine phonetische Analyse macht.

Was heisst das konkret?

Linguistisch bedeutet, es wird untersucht, *was* gesagt wird – und phonetisch betrifft die Stimme, also *wie* etwas gesagt wird. Das alles ruft im wissenschaftlichen Betrieb derzeit noch Skepsis hervor. Ich bin zu 80 Prozent Wissenschaftler und zu 20 Prozent Unternehmer. Als Wissenschaftler bin ich bei Precire skeptisch, weil das, was dort gemacht wird, noch ziemlich unscharf ist. Die Substanz nach innen ist noch sehr

ausbaufähig, und das Auftreten des Unternehmens im Markt ist zudem plakativ.

Wie sieht denn die Akzeptanz am Markt aus?

Es gibt Substanzprojekte und Showprojekte. Diejenigen, die für Showprojekte empfänglich sind, machen eher mit.

Was verstehen Sie unter einem Showprojekt?

Diejenigen Unternehmen, die gerne etwas KI-basiertes (KI=Künstliche Intelligenz, Anm. d. Red.) machen möchten, das derzeit in aller Munde ist, die springen bereits auf den Zug auf. Der Flaschenhals liegt in der Qualität der Lerndaten.

Von welchen Faktoren hängt diese Qualität ab?

Unterschiedliche Störeinflüsse wie Hintergrundgeräusche, unterschiedliche Sprecher, unterschiedliche Dialekte. Die Datenbasis ist das Hauptnadelöhr. Eine Depression kann heute mit einer Treffsicherheit von 80 Prozent erkannt werden, weil hier bereits viele Daten vorliegen, nämlich von etwa 800 Sprechern.

Bei der Personaldiagnostik profitieren doch alle davon, wenn der Geeignete am geeigneten Platz sitzt.

Wo sonst ist eine breite Datenbasis vorhanden?

Vor allem bei den Big Five (ein Modell der Persönlichkeitspsychologie, auch Fünf-Faktoren-Modell genannt; Anm. d. Redaktion) sind viele Daten öffentlich. Auch Müdigkeit kann gut erkannt werden, da es sich um ein klares physiologisches Phänomen handelt. Immer, wenn grosse Unterschiede in den Basisemotionen erkennbar sind wie Ärger, Freude, Überraschung, Trauer, sind genügend Daten vorhanden.

Parkinson soll mittels Stimmanalyse zwei Jahre früher erkennbar sein als beim Arzt, war in den Medien zu lesen – was ist da dran?

Da wäre ich skeptisch. Wenn ein Produkt entwickelt, aber noch nicht von aussen verifiziert wurde, sind Zweifel berechtigt. Es wäre dann plausibel, wenn ein Vergleichswert da ist, also wenn sich die Stimme von einer an Parkinson erkrankten Person durch den Tremor zittriger anhört als vor der Erkrankung.

Wo sehen Sie die ethischen Grenzen?

Ich selbst bin von meiner Persönlichkeit her kein Bedenkenträger. Es gibt in Deutschland viele, die technologiekritisch sind und etwas über das Ziel hinaus-



schliessen. Viele Themen lassen sich gut lösen. Bei der Personaldiagnostik profitieren doch alle davon, wenn der Geeignete am geeigneten Platz sitzt. Beim autonomen Fahren wird immer wieder die Frage aufgeworfen, ob sich Autos in Konfliktfällen ethisch korrekt verhalten können. Diese Bedenken sollten uns nicht abhalten, weiter zu forschen. Sie sind für mich keine prinzipiellen Hemmnisse, sondern interessante Fragen. Im Gegenteil sollten wir mit einer Wucht in KI und stimmbasierte Analysen investieren. Wir haben momentan in den deutschsprachigen Ländern einen guten wissenschaftlichen Stand. Und es eröffnen sich enorme Chancen, insbesondere auch für den Wissensstandort Deutschland. Wir sind auch in Kontakt mit einem internationalen Anbieter, der über Video Personalassessments anbietet und zwar in Russland, wo die KI-basierte Stimmanalyse wesentlich positiver aufgenommen wird.

Wo liegt das Scharnier zum Unternehmer bei Ihnen?

Es gibt viele bürokratische Hemmnisse für die Wissenschaft. Deshalb haben wir eine GmbH gegründet, wo wir gerade Projekte zur Schmerz-, Depressions- oder Herkunftsland-Erkennung umsetzen. Dabei arbeiten wir mit einer Vielzahl von Partnern aus Industrie, Forschung und Behörden zusammen.

Wie realistisch ist das eingangs erwähnte Beispiel mit dem Fahrkartenautomaten?

Ich glaube, das ist sehr realistisch, diese Omni-präsenz führt dazu, dass Stimmen analysiert werden. Ich glaube, die öffentliche Akzeptanz ist dann schnell da, wenn die Analyse eine hohe Genauigkeit hat.

Wie schnell konkret?

Ich denke, in zehn Jahren hat sich die smarte Technologie im Smart-Home-Bereich durchgesetzt. Es sieht zumindest momentan danach aus, zumal sich die grossen Player auch auf die Sprachanalyse stürzen. Bereits in drei bis fünf Jahren werden die Produkte ziemlich gut sein. Amazon hat jüngst ein neues Patent für Alexa angemeldet, dem hauseigenen digitalen Assistenten. Da geht man schon in die Richtung. Die fangen mit einfachen Phänomenen an. Laut dem neuen Patent soll Alexa über das Mikrofon auch weitere Zusatzinformationen des Nutzers sammeln und analysieren können, die auf die Gesundheit des Anwenders schliessen. Anschliessend sollen dem Kunden passende Inhalte angeboten werden. Wird beispielsweise ein Husten über das Mikro wahrgenommen oder lassen sich aus der Spracheingabe vergleichbare Erkenntnisse über etwaige Erkrankungen des Nutzers gewinnen, soll der digitale Assistent von Amazon sich aktiv einschalten und entsprechende Medikamente

bzw. Arzneimittel zum Kauf – in unserem Beispiel etwa Hustensirup – anbieten.

Wird es hier nicht datenschutzrechtlich heikel?

Im Moment sind die Datenschutzbedingungen in Deutschland noch sehr eng. Es wird eher noch mit Bonussystemen operiert. Wer gesund ist und die App nutzt, wird belohnt.

Wie viel lässt sich tatsächlich von der Wirklichkeit durch die Stimme spiegeln?

Gemütszustände, Verfassungen, Emotionen, kurz: alles, was ein Mensch bei einem guten Bekannten oder einem Freund, den er gut kennt, anhand der Stimme erkennt, werden Systeme mit guten Lerndaten auch erkennen können. Alles, was darüber hinausgeht, ist eher unrealistisch. Google Assistant fragt übrigens manchmal nach, wenn für die Software etwas unklar ist. So kann die Genauigkeit gesteigert werden.

Wie lässt sich die Genauigkeit der Sprachanalyse wissenschaftlich begründen und nachweisen?

Bei Müdigkeit ist es noch relativ einfach, zum Beispiel über EEG-Systeme oder die Erkennung von Mikro-Schlafsymptomen. Geht es um die Persönlichkeit, dann verifizieren wir mit einer Kombination der Big Five, einem Selbsturteil oder je nach Motivationslage kann das in Übereinstimmung mit der Aussensicht geschehen.

Wie könnte die Stimmanalyse im Coaching eingesetzt werden?

Ich könnte mir vorstellen, dass man Lernmodule oder Übungsmodule entwickelt, indem man das System damit füttert, was ein guter Coach auch kann. So könnten Lernübungsplattformen entwickelt werden mit einem virtuellen Coach.



Prof. Dr. Jarek Krajewski ist internationaler Experte im Bereich Digitales Betriebliches Gesundheitsmanagement und Affective Computing. Im Bereich der digitalen Sprachsignalforschung ist er Initiator grosser internationaler Forschungsverbände. Er forscht zudem in zahlreichen öffentlich- und industriegeförderten Drittmittelprojekten.
j.krajewski@ixp-duesseldorf.de