



← JAN NEUMANN

ÜBER JAN NEUMANN

Jan Neumann leitet die Comcast Applied Artificial Intelligence Research Group mit Teammitgliedern in Washington DC, Philadelphia, Chicago, Denver und im Silicon Valley. Das Aufgabenspektrum des Teams umfasst Maschinelles Lernen, Deep-Learning, NLP und Computer-Vision sowie die Entwicklung neuer Algorithmen und Produktkonzepte. Die Aktivitäten verfolgen das Ziel, den Comcast-Kundenservice laufend zu verbessern, wie z.B. durch die X1-Sprachfernbedienung und Personalisierungsfunktionen, virtuelle Assistenten und Predictive-Intelligence im Kundenservice sowie intelligente Video- und Sensoranalysen. Bevor Jan Neumann 2009 zu Comcast kam, arbeitete er für Siemens Corporate Research an verschiedenen Projekten im Bereich Computer Vision, wie beispielsweise Fahrerassistenzsystemen und Videoüberwachung. Er hat über 20 Artikel in wissenschaftlichen Konferenzen und Zeitschriften veröffentlicht und ist ein gefragter Referent zum Thema Machine Learning und Data Science. Er promovierte in Informatik an der University of Maryland, College Park.

ÜBER COMCAST

Comcast Corporation (Nasdaq: CMCSA) ist ein globales Medien- und Technologieunternehmen mit drei Hauptgeschäftsbereichen: Comcast Cable, NBCUniversal und Sky. Unter der Marke Xfinity ist Comcast Cable einer der größten Video-, Hochgeschwindigkeits-Internet- und Telefonanbieter der Vereinigten Staaten für Privatkunden und ist auch im Firmenkundengeschäft aktiv. Darüber hinaus können Privatkunden bei Xfinity auch mobile Dienste sowie Sicherheits- und Automatisierungsdienstleistungen kaufen. NBCUniversal ist global tätig und betreibt zahlreiche Nachrichten-, Unterhaltungs- und Sportsender, die Rundfunknetze von NBC und Telemundo, Produktionsstudios, Fernsehstationen sowie Universal Pictures und Universal Parks und eigene Resorts. Sky ist eines der führenden europäischen Medien und Unterhaltungsunternehmen, das über seine Pay-TV-Dienste Kunden mit einer breiten Palette an Videoinhalten versorgt. Auch Telekommunikationsdienstleistungen wie Hochgeschwindigkeits-Internet-, Telefon- und Mobildienste für Privatanwender werden angeboten. Sky betreibt den Kanal Sky News sowie Sport- und Unterhaltungssender, produziert selbst Originalinhalte und verfügt über zahlreiche Exklusivrechte.

→ www.comcastcorporation.com



INTERVIEWER

Professor Christian Hildebrand führte das Interview im Juni 2019.

Sprechen statt Tippen: Sprachbasierte Fernbedienungen können mehr

Interview mit Jan Neumann, Senior Director, Applied AI,
Comcast Cable, Philadelphia, USA

Während viele Kunden immer noch zögern, sich in den eigenen vier Wänden Alexa, Cortona oder Siri anzuvertrauen, scheint ihnen die Sprachsteuerung ihrer Fernsehgeräte weniger Sorgen zu bereiten. Comcast hat im Jahr 2015 mit dem Angebot einer sprachbasierten Fernbedienung begonnen und diesen Service kontinuierlich erweitert. Die weite Welt des Home Entertainments scheint ein günstiges Umfeld für den Einsatz von Sprachsteuerung zu sein. Bei der Navigation durch die ständig zunehmenden Unterhaltungsoptionen auf TV-Geräten bietet sie Konsumenten zahlreiche Vorteile. Jan Neumann erklärt, wie Comcast es seinen Kunden ermöglicht, ein riesiges Unterhaltungsportfolio bequem auf persönlich relevante Inhalte zu reduzieren und im hart umkämpften Home Entertainment-Markt erfolgreich zu bleiben.



Christian Hildebrand ✕ *Der Aufstieg der Sprachsteuerung bzw. Voice Control ist heutzutage ein heißes Thema. Alle reden von Alexa, Google Home und Siri. Comcast hat nun eine sprachbasierte Fernbedienung eingeführt. Erfindet die sprachbasierte Steuerung das TV-Erlebnis neu?*

Jan Neumann ✕ Eines der Hauptprobleme der Fernsehkunden ist, dass immer mehr Inhalte verfügbar sind, während sich die Steuerungsgeräte nicht verändert haben. Es gibt nach wie vor die traditionellen Fernbedienungen mit Ziffern. In einer Welt, in der man nur zwischen 20 Fernsehsendern wechseln musste, von denen man sich die entsprechenden Zahlen merken konnte, hat das gut funktioniert. Heute hat man hingegen hunderttausende verschiedene Inhalte jederzeit zur Verfügung. In dieser Situation ist eine komplexere Steuerung, die einem hilft, das zu finden was man will, absolut entscheidend.

Bietet Voice nur ein anderes oder tatsächlich ein besseres Kundenerlebnis?

Die Sprachsteuerung ändert die Bedienung grundlegend und dreht sie zugunsten der Konsumenten. Traditionell diktiert das Steuerungsgerät dem Kunden, was er zu tun hat. Nun erlauben wir Kunden, sich selbst auszudrücken, und dann liegt es an uns, ihre Angaben zu interpretieren. Dieser Prozess stellt den Kunden vor die Plattform. Die Sprache baut Hierarchien in der Steuerung ab, sie ist die perfekte Abkürzung. Man sagt uns einfach, woran man interessiert ist, und dann sind wir am Zug, um das Gewünschte am Bildschirm zu präsentieren. Durch die Spracheingabe können wir die Menge der präsentierten Inhalte eingrenzen, sodass man dann wieder mit der herkömmlichen Fernbedienung navigieren kann. Das lästige und mühselige Herumtippen am Fernsehgerät wird überflüssig.

»

*Aus Nutzerfeedbacks wissen wir,
dass die sprachbasierte Fernbedienung
ein großer Treiber für Kundenzufriedenheit
und -bindung ist.*

«

Von anderen sprachbasierten Geräten wissen wir, dass es eine große Diskrepanz zwischen Kauf- und tatsächlichen Nutzungsraten gibt. Die Nutzung hinkt deutlich hinterher. Wissen Sie, ob Kunden Ihre sprachbasierte Fernbedienung tatsächlich nutzen?

Ja, sie ist eines der beliebtesten Produkte, die wir herausgebracht haben und läuft äußerst erfolgreich. Aus Nutzerfeedbacks wissen wir, dass die sprachbasierte Fernbedienung ein großer Treiber für Kundenzufriedenheit und -bindung ist. Sobald sich die Kunden daran gewöhnt haben, nutzen sie diese täglich und intensiv. Wir verarbeiten mehr als eine halbe Milliarde Sprachbefehle pro Monat und die Anzahl steigt laufend. Die sprachbasierte Fernbedienung ist sogar einer der Hauptgründe, warum Kunden unserer Plattform treu bleiben.

Ist Voice-Control nur eine neue Form der Kommunikation mit Ihren Kunden oder verändert sie auch das Verhalten grundlegend? Suchen Menschen mittels Sprachsteuerung anders?

Die Anwender beginnen mit einfachen Anwendungsfällen, wie z. B. der Suche nach „CNN“ oder „NBC“ oder bestimmten Nummern von Sendern. Sobald man ihnen Beispiele für weitere Suchmöglichkeiten gibt, werden ihre Abfragen komplexer. Wir verbessern ständig die Funktionalität. Zum Beispiel können Benutzer nach „Ergebnissen der aktuellen NBA-Playoffs“ fragen, und wir sind in der Lage, Statistiken zu liefern, zum Beispiel zum Duell Raptors gegen Warriors, oder auch Information darüber, wann die Spiele stattfinden. Im Grunde genommen müssen die Benutzer also nur noch beschreiben, was sie sehen wollen.

Könnten Sie uns ein weiteres Beispiel geben, wie ein Benutzer suchen könnte und was er oder sie dann bekommt?

Erst kürzlich haben wir ein neues Feature eingeführt. Benutzer können beschreiben, was in einer bestimmten Folge einer

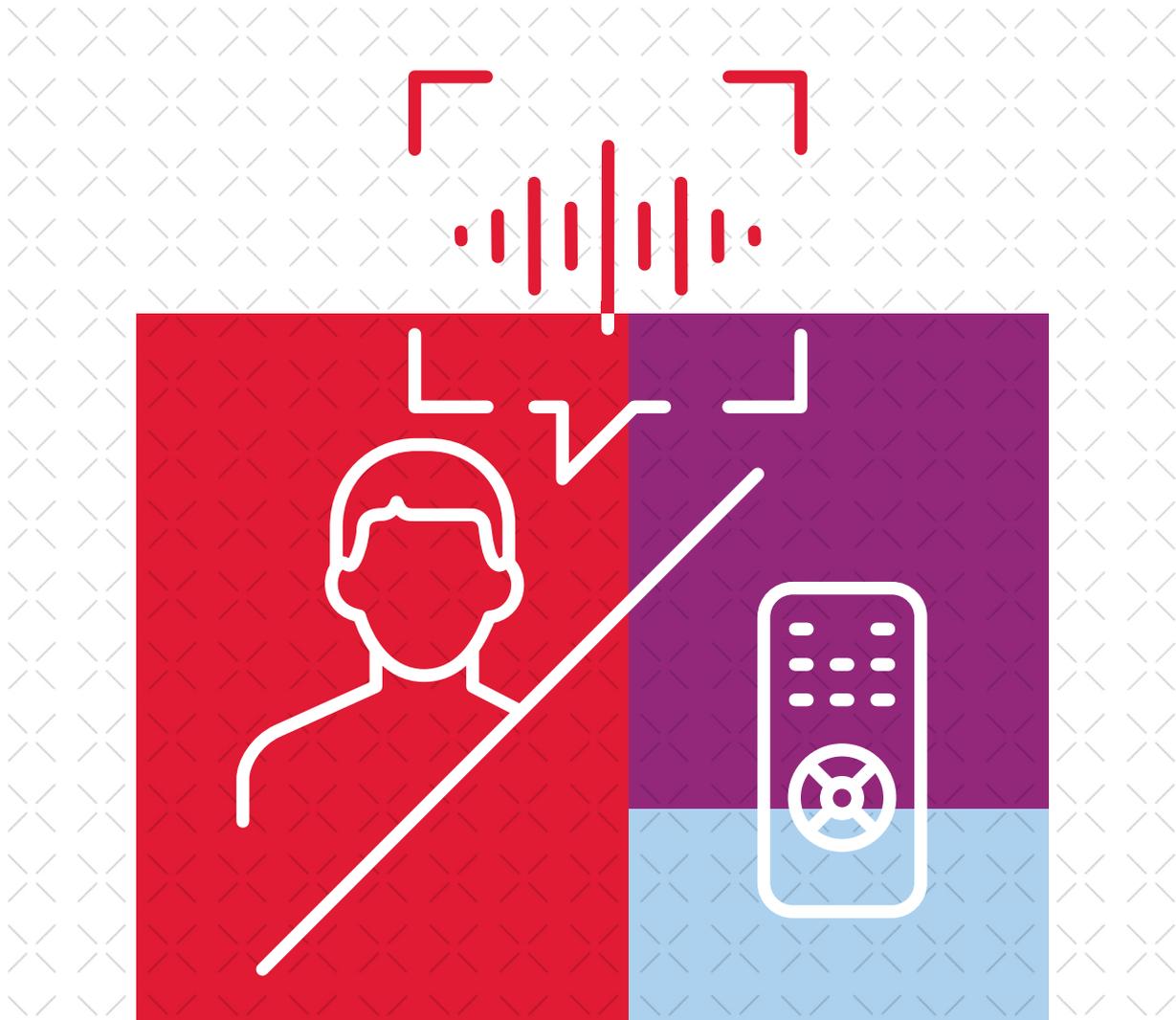
Serie passiert. Sie müssen sich nicht mehr an den Titel oder die Nummer der Folge erinnern. Sie müssen nur sagen: „Zeige mir die Friends-Episode mit Brad Pitt“, und sofort haben sie diese Episode auf dem Bildschirm.

Wie sprechen Kunden mit der sprachbasierten Fernbedienung? Ganz natürlich, wie mit einer anderen Person, oder versuchen sie es zunächst mit einzelnen Schlüsselwörtern?

Die Leute beginnen oft mit einfacheren, kommandoähnlichen Eingaben. Sie können aber praktisch alles ausprobieren – Versuch und Irrtum sozusagen – nicht nur für sie, sondern auch für uns. Wir geben unseren Kunden nicht vor, welche Befehle oder Wörter sie verwenden sollen. Die Nutzer haben die Möglichkeit und die Freiheit, genau das auszudrücken, was sie wollen. So können wir lernen, was beliebt oder interessant ist. Durch Zuhören erfahren wir, was Kunden wünschen und was das nächste Feature sein könnte, das wir liefern sollten.

Wie können Sie ein besseres Kundenerlebnis schaffen, wenn die Leute nicht alle Funktionen der Sprachsteuerung kennen? Müssen Sie Kunden aktiv ermutigen, Neues zu probieren und erklären, wie etwas funktioniert?

Um unsere Kunden zu informieren, nutzen wir zum Beispiel regelmäßig den Bildschirmschoner. Außerdem schalten wir viel Werbung, die potenzielle Anwendungsfälle beschreibt. Erst vor kurzem haben wir die Funktion „Sprachvorschläge“ eingeführt, die aktiv empfiehlt, wonach Verbraucher suchen könnten. Im Zuge der Einführung haben wir den Kunden einige Suchausdrücke und neue Suchmöglichkeiten beigebracht. Unsere Vorschläge sind immer kontextbezogen, sodass sich die Nutzer besser daran erinnern und sie beim nächsten Mal eher selbst verwenden, weil sie passend erscheinen. Dieses Feature ist so neu, dass wir noch keine messbaren Ergebnisse haben, aber wir selbst sind jedenfalls recht begeistert davon.

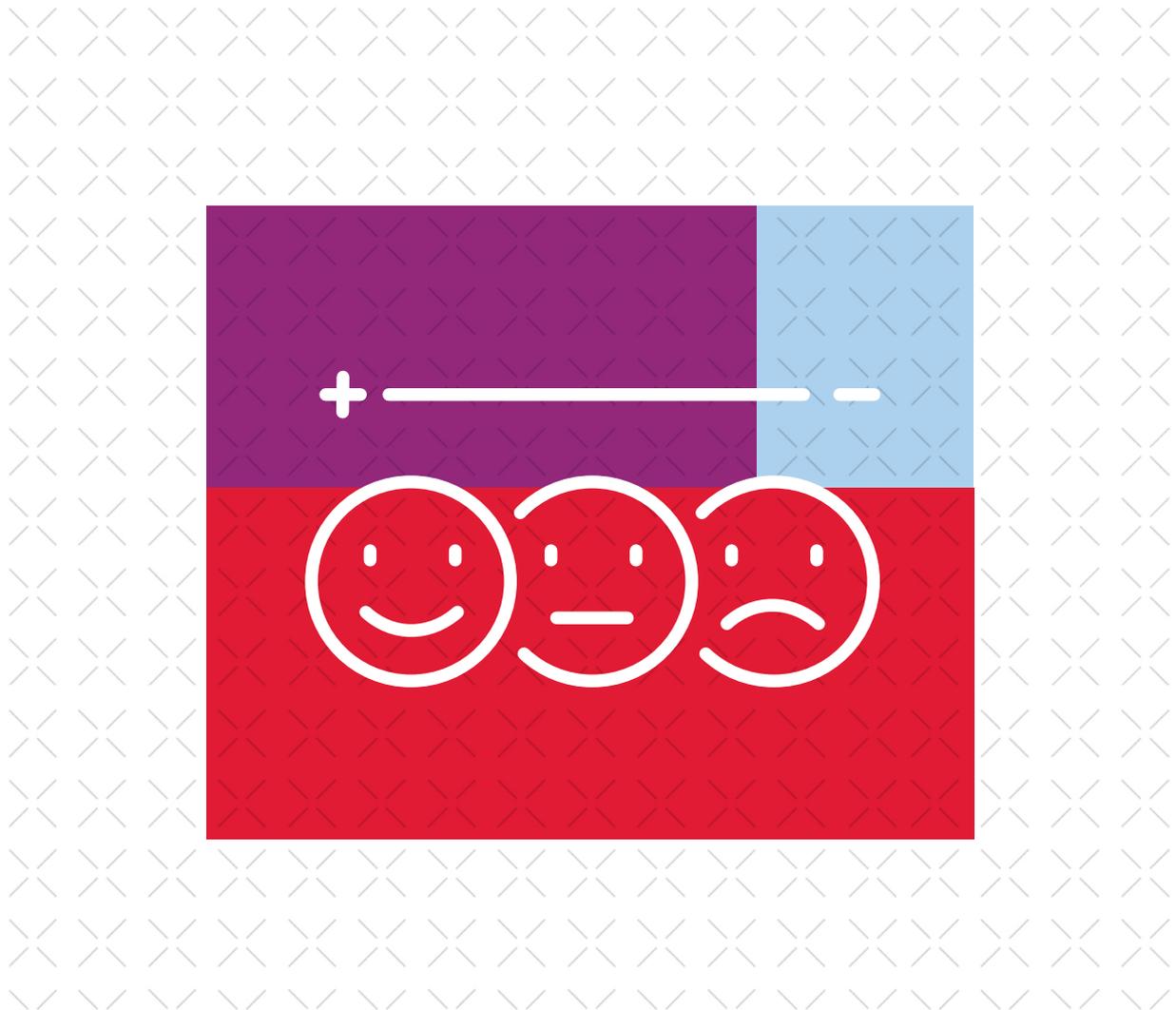


Absichten zu erkennen, ist oft ziemlich schwierig. Wie nutzt man Sprachverarbeitung, um das Verständnis zu verbessern und zu vermeiden, was Alexa-Anwender nur zu gut kennen – „Tut mir leid, das weiß ich nicht“?

Unser Vorteil ist, dass wir keinen Generalassistenten für alles anbieten. Unser Sprachservice konzentriert sich auf Unterhaltung, Heimsteuerung und Kundenservice. In jedem dieser Bereiche haben wir ein ziemlich gutes Verständnis davon, worum es geht und was getan werden kann. Der Bereich des Kundenservice ist etwas komplexer, da man längere Sätze verstehen und Chats bewältigen muss. Die eher kurzen Sprachbefehle in der Unterhaltung sind im Gegensatz dazu meist recht prägnant und eindeutig. Unser Servicespektrum ist weit weniger komplex, denn je mehr Dienstleistungen angeboten werden, desto schwieriger wird es, Anforderungen richtig zu interpretieren.

»
Das lästige und mühselige Herumtippen am Fernsehgerät wird überflüssig.

«



Eine weitere Herausforderung stellt sicher Ihr umfassendes Portfolio an ständig wechselnden Inhalten und laufend neuen Programmen dar. Wie managen und bewältigen Sie diese Dynamik und Vielfalt? Wie schaffen Sie es, Ihren Kunden maßgeschneiderte, zeitnahe Empfehlungen zu liefern?

Es ist wichtig, gute Metadaten zu generieren und zu pflegen. Man muss dazu fähig sein, automatisch zu erkennen, wer in einer Sendung auftritt, worum es in einem neuen Video inhaltlich geht, welche Kundensegmente interessiert sein könnten, usw. Es geht darum, in die Tiefe zu gehen, um Inhalte identifizieren zu können. Man muss vielschichtige Beschreibungen darüber haben, was zu erwarten ist, welche Emotionen angesprochen werden oder welcher Musikstil dominiert. Wir füllen Lücken in der Beschreibung mit Daten, die nicht automatisch bereitgestellt wurden und ermöglichen damit neue Formen der Interaktion.

Welche Lektionen haben Sie bei der Einführung der sprachbasierten Fernbedienung gelernt? Gibt es da Punkte, die für andere Branchen, die mit Sprachsteuerung experimentieren, relevant sind?

Vor allem im Home-Entertainment-Bereich hat man auch einen Bildschirm, den man zusätzlich nutzen sollte. Die Nutzung betrifft nicht nur die Stimme, sondern muss im Kontext von Sprache und Text gesamtheitlich betrachtet und konzipiert werden, und das ist eine große Herausforderung. Man kann sich beispielweise über die Sprachsteuerung eine Liste von aktuellen Komödien erstellen lassen, die Feinauswahl dann aber auf dem Bildschirm durchführen. Anstatt zu sagen „nächste Seite, nächste Seite oder zweite Zeile rechts oben“, ist in dieser Phase die Navigation über die Tastatur sinnvoller. Man muss die Stärken der beiden Steuerungsmöglichkeiten kombinieren.



Letztendlich geht es darum, wer am schnellsten von den Kunden lernen und am agilsten auf ihre Bedürfnisse eingehen kann.



Gehen wir nun zu allgemeineren Themen über. Wie hat die Arbeit mit KI und der neue Analysefokus das Innenleben von Comcast verändert? Was hat der Übergang von einer „System Administration Culture“ zu einer Kultur, bei der die kontinuierliche Verbesserung des Kundenerlebnisses im Vordergrund steht, im Unternehmen bewirkt?

Die größte Veränderung war, den Kunden in den Mittelpunkt zu stellen. Die Ausrichtung auf den Kunden und seine Zufriedenheit verändert den Fokus, die Art der Projekte und das gesamte Geschäftsmodell. Es ist äußerst motivierend zu sehen, welche Wirkungen wir mit unserem Technologiewechsel erzielen können. Anstatt Annahmen treffen zu müssen, was die Kunden mögen könnten, erhalten wir diese Informationen direkt aus unseren Technologien. Wir können nun viel schneller lernen, und das ist wirklich spannend.

Ironischerweise wurde Comcast vor ein paar Jahren wegen mangelnder Kundenorientierung kritisiert. Inzwischen aber scheinen Sie selbst durch die Nutzung von KI der Game-Changer zu sein und können ein besseres Kundenerlebnis bieten als andere. Das ist ziemlich beeindruckend.

Danke für das Kompliment. Die Neuausrichtung war eine unternehmensweite Initiative, die weit über Technologien und Software hinausreicht, und es ist wirklich spannend zu erleben, wie Kunden reagieren. Wir wissen, dass wir noch viel mehr tun müssen, aber als Technologe und KI-Wissenschaftler finde ich es besonders cool, diese Technologien anzuwenden, um die Erfahrungen unserer Kunden einfach und reibungslos zu gestalten.

Ihre Mitbewerber sind Player wie Netflix und andere globale Technologieunternehmen, die an Ihren Marktanteilen knabbern. Ist das nicht extrem herausfordernd oder gar beängstigend?

Als Ingenieur würde ich diese Situation eher als motivierend beschreiben. Dass so viele Unternehmen so viele spannende Dinge machen, inspiriert mich und mein Team enorm und

motiviert uns, Neues auszuprobieren und immer einen Schritt voraus zu sein. Ingenieure treten als Gruppe gern in einen Wettstreit mit anderen, und wenn wir sehen, wie jemand anderer etwas Cooles macht, ziehen wir daraus sowohl Inspiration als auch Motivation.

Hat derjenige einen Wettbewerbsvorteil, der den besten Algorithmus hat?

Algorithmen sind das Werkzeug, um bestimmte Ziele zu erreichen. Letztendlich geht es vielmehr darum, wer am schnellsten von den Kunden lernen und am agilsten auf ihre Bedürfnisse eingehen kann. Grundsätzlich ist das Geschäft für alle gleich, aber es gibt Möglichkeiten, den Feedback-Zyklus zu beschleunigen und durch Technologie laufend effizienter zu werden.

Werfen wir zum Schluss noch einen Blick auf Steuerungsmöglichkeiten der Zukunft. Was wird der nächste große Wurf, um Kunden besser zu bedienen? Vielleicht die Steuerung über Gedanken oder zusätzliche Sensoren im Wohnraum der Menschen?

Wie auch immer eine neue Steuerung aussehen mag, sie muss die Bedürfnisse der Konsumenten aktiv verstehen und antizipieren können, dabei helfen, Kundenwünsche zu erfüllen und klare Ziele zu erreichen. Das Sprechen ist eine natürliche Form der Kommunikation und derzeit ist sie ein sehr gutes Steuerungsmedium. Wir testen ebenfalls nonverbale Steuerungen und vielleicht können auch diese bei zukünftigen Bedienungselementen einbezogen werden. Eigentlich spielt das Medium aber gar keine so wichtige Rolle. Was auch immer es uns ermöglicht, die Bedürfnisse unserer Kunden effektiv und mit so wenig Reibungsverlusten wie möglich zu erfassen, wird sich letztendlich durchsetzen.

Was für ein Schlusswort! Danke für Ihre Zeit, Jan, und dafür, dass Sie uns durch das neue Tech-Universum bei Comcast begleitet haben!

