

SWR2 Wissen

Chatbots – Reden mit Maschinen

Von Aeneas Rooch

Sendung: Dienstag, 18. August 2020, 8.30 Uhr
Erstsendung: Montag, 11. März 2019, 8.30 Uhr
Redaktion: Gábor Páal
Regie: Andrea Leclerque
Produktion: SWR 2019

Chatbots kommen im Kundendienst oder als Sprachassistenten zum Einsatz. „Gespräche“ mit ihnen laufen bereits erstaunlich flüssig. Doch „menschlich“ sind sie noch lange nicht.

Bitte beachten Sie:

Das Manuskript ist ausschließlich zum persönlichen, privaten Gebrauch bestimmt. Jede weitere Vervielfältigung und Verbreitung bedarf der ausdrücklichen Genehmigung des Urhebers bzw. des SWR.

SWR2 können Sie auch im **SWR2 Webradio** unter www.SWR2.de und auf Mobilgeräten in der **SWR2 App** hören – oder als **Podcast** nachhören.

Kennen Sie schon das Serviceangebot des Kulturradios SWR2?

Mit der kostenlosen SWR2 Kulturkarte können Sie zu ermäßigten Eintrittspreisen Veranstaltungen des SWR2 und seiner vielen Kulturpartner im Sendegebiet besuchen. Mit dem Infoheft SWR2 Kulturservice sind Sie stets über SWR2 und die zahlreichen Veranstaltungen im SWR2-Kulturpartner-Netz informiert. Jetzt anmelden unter 07221/300 200 oder swr2.de

Die SWR2 App für Android und iOS

Hören Sie das SWR2 Programm, wann und wo Sie wollen. Jederzeit live oder zeitversetzt, online oder offline. Alle Sendung stehen mindestens sieben Tage lang zum Nachhören bereit. Nutzen Sie die neuen Funktionen der SWR2 App: abonnieren, offline hören, stöbern, meistgehört, Themenbereiche, Empfehlungen, Entdeckungen ...
Kostenlos herunterladen: www.swr2.de/app

MANUSKRIFT

Atmo:

Tippen

Sprecherin:

Wenn wir chatten – wenn wir uns am Computer oder Smartphone per Textnachricht unterhalten –, antwortet uns nicht immer ein echter Mensch. Manchmal schreibt uns ein Computerprogramm zurück.

Chatbot 1:

Hallo. ich bin Tina vom Kundenservice. Wie kann ich Ihnen helfen?

Sprecherin:

Solche Chat-Programme begegnen uns auf vielen Webseiten, in Messengern wie „WhatsApp“ und auf sozialen Plattformen wie „Facebook“ und „Twitter“. Nicht immer bemerken wir sie: Manchmal glauben wir, dass wir uns einem Menschen unterhalten – in Wirklichkeit aber schreibt uns ein Computerprogramm.

Chatbot 2:

Fahrverbote sind eine Sauerei! Besitzer von Dieselfahrzeugen werden quasi enteignet!

Sprecherin:

Wie funktionieren Chat-Programme? Was können sie, was können sie nicht? Woran erkennen wir sie? Können sie uns täuschen? Sind diese Programme intelligent?

Musikakzent

Ansage:

Chatbots – Reden mit Maschinen. Eine Sendung von Aeneas Rooch.

Chatbot:

Guten Tag. Wie geht es Dir?

O-Ton David Schlangen:

Früher hätte man Chatbots einfach nur jetzt Small-Talk-Systeme genannt.

Chatbot:

Was machst Du so?

O-Ton David Schlangen:

Ursprünglich war es mal die Bezeichnung von Computerprogrammen, mit denen man Smalltalk führen kann, typischerweise über Texteingabe. Das hat sich inzwischen ein bisschen gewandelt.

Chatbot:

Wie kann ich Dir helfen?

O-Ton David Schlangen:

Allgemeiner wird es verwendet für Computerprogramme, mit denen man sprachlich interagiert, dass man zum Beispiel einen Chatbot haben kann, der für eine Tischreservierung im Restaurant macht oder sowas.

Chatbot:

Ich habe einen Tisch für zwei Personen, am Donnerstag um 19 Uhr.

Sprecherin:

Ein Chatbot ist ein Computerprogramm, mit dem man sich unterhalten kann. Am Computer oder Smartphone kann man ihm texten, und das Programm textet zurück – in natürlicher Sprache, wie ein echter Mensch. Chatbots werden vielfältig eingesetzt.

Musikakzent

Sprecherin:

Sie liefern aktuelle Nachrichten,

Chatbot:

27 Staats- & Regierungschefs der EU wissen nicht, ob sie erleichtert oder traurig sein sollen. Mehr erfahren?

Sprecherin:

Beraten zu Internet-Tarifen,

Chatbot:

Suchen Sie einen Smartphone- oder Prepaid-Tarif?

Sprecherin:

Helfen beim Ändern von Adresse und Bankverbindung,

Chatbot:

Straße: „An der Mettwurst“, Ort: „48565 Steinfurt“.

Sprecherin:

Erledigen Überweisungen,

Chatbot:

14,50 Euro.

Sprecherin:

Informieren über Stau,

Chatbot:

7 Kilometer stockender Verkehr.

Sprecherin:

Beraten beim Kleidungskauf,

Chatbot:

Ich habe gelbe Turnschuhe gefunden.

Sprecherin:

Beim Buchen eines Mietwagens,

Chatbot:

An welchem Tag möchtest Du Deinen Wagen zurückgeben?

Sprecherin:

Schlagen Rezepte vor,

Chatbot:

3 EL Ketchup dazugeben.

Sprecherin:

Verschicken lustige Nachrichten an Freunde.

Chatbot:

Wie nennt man einen übergewichtigen Vegetarier? Biotonne.

Sprecherin:

Chatbots – oder auch „textbasierte Dialogsysteme“, wie sie im Fachjargon heißen – gibt es schon seit den 1960er-Jahren, und sie werden immer besser und schneller: Dass da kein echter Mensch schreibt, sondern ein Computerprogramm – das merken Nutzer in einigen Fällen gar nicht mehr.

Chatbot:

Ähm, hier ist alles super. Und bei dir?

Sprecherin:

In sozialen Netzwerken wie Twitter, Facebook und Co. sind Maschine und Mensch kaum noch zu unterscheiden.

Musikakzent

Chatbot A:

Denkt noch wer, dass die EU und Großbritannien den Flug MH17 als Ausrede benutzt, um Russland zu attackieren?

Chatbot B:

Bei der zunehmenden Antibiotika-Resistenz und den vielen Toten durch Arzneimittel-Nebenwirkungen sind sanft wirkende Heilmethoden wie die Homöopathie die einzige Lösung.

Chatbot C:

Freie Fahrt für freie Bürger! Fahrverbote sind eine Sauerei! Besitzer von Dieselfahrzeugen werden quasi enteignet!

O-Ton David Schlangen:

Social Bots würde man die nennen, also Chatbots, die in sozialen Netzwerken unterwegs sind und sich unter Umständen auch gar nicht als Computerprogramme zu erkennen geben.

Sprecherin:

David Schlangen ist Professor für Computerlinguistik an der Universität Bielefeld.

O-Ton David Schlangen:

Manchmal werden die eingesetzt, um eine bestimmte Botschaft zu multiplizieren, besonders zu verstärken.

Chatbot D:

Sexualstraftäter zieht in Asyl-Unterkunft! Deutschland wird zum Streichelzoo für Vergewaltiger, Terroristen und Schein-Asylanten.

Sprecherin:

Für Menschen wird es in solchen Diskussionen immer schwieriger zu erkennen, dass dort eine Maschine schreibt. Aber ein paar Hinweise gibt es. Manche Bots erkennt man daran, dass sie unnatürlich viel oder zu seltsamen Zeiten posten, sagt David Schlangen.

O-Ton David Schlangen:

Ein Benutzer, der 24 Stunden am Tag Nachrichten absetzt, mit fünf Minuten Abstand vielleicht, das ist vermutlich kein menschlicher Benutzer. Denn normale Menschen müssen auch mal was anderes machen als Nachrichten abschicken.

Sprecherin:

Manche Bots schreiben immer das gleiche.

O-Ton David Schlangen:

Man kann gucken, ob dieselbe Nachricht, die da abgeschickt wurde, ob die auch noch von anderen Benutzern identisch abgeschickt wurde.

Chatbot F:

Nichts wie raus aus dem Grundgesetz!! Lasst uns eine Verfassung schreiben. Staatsgebiet festlegen, Grenzen dicht und so weiter.

Chatbot G:

Staatsgebiet festlegen, Grenzen dicht und so weiter.

Chatbot H:

Staatsgebiet festlegen, Grenzen dicht und so weiter.

O-Ton David Schlangen:

Ob das so aussieht, als ob da jemand das selbst verfasst hat oder ob das nach einer konzertierten Aktion aussieht.

Sprecherin:

Manche Bots reagieren automatisch auf bestimmte Schlüsselwörter.

O-Ton David Schlangen:

Das kann man oft leicht erkennen, weil diese Art von Antworten nach Schlüsselwörtern einfach daneben gelaufen ist und inhaltlich gar nicht wirklich richtig gepasst hat.

Mensch:

In der Gruppe war ein echt gutes Klima.

Chatbot E:

Nur 0,03% des CO₂ in der Atmosphäre kommt von Menschen. Die globale Erwärmung kommt nicht von uns Menschen!! Welche Beweise brauchen wir denn noch?!

Sprecherin:

Man kann Bots auch daran erkennen, dass sie untereinander vernetzt sind, sagt Professor David Schlangen.

O-Ton David Schlangen:

Um sich als Mensch zu tarnen, sind die Autoren von solchen Programmen draufgekommen, dass es nützlich ist, wenn der auch Freunde oder Follower hat. Das sind dann sogenannte Bot-Farmen.

Das heißt, dass der Autor dieses Systems gleichzeitig vielleicht hundert oder tausend solcher User anlegt, die sich dann alle miteinander befreunden. Und auf diese Weise versuchen die eben, so auszusehen wie normale Benutzer.

Sprecherin:

Bot 1 ist mit Bot 2 befreundet. Bot 2 ist mit Bot 3 und Bot 4 befreundet. Bot 3 und 4 sind mit Bot 5 und 6 und 7 befreundet. Und schon entsteht für einen Menschen der Eindruck, dass es sich bei den Accounts jeweils auch um Menschen mit einem Freundeskreis handelt.

Letzten Endes käme man Bots auf die Spur, wenn man nicht nur einen einzelnen Post von ihnen läse, sondern versuchte, eine echte Unterhaltung mit ihnen zu führen, denn das können sie schlecht.

Menschlicher Chatter:

Für welche Mannschaft bist du?

Chatbot:

Ich bin Fan von Leeds United.

Menschlicher Chatter:

Leeds! Wie läuft die Saison?

Chatbot:

Ja, es ist eine Stadt im Norden von England.

O-Ton David Schlangen:

Oft sind es keine wirklichen Interaktionen, wo man länger mit demselben Bot, mit demselben User, reden würde. Wenn das der Fall ist, dann kann man meistens relativ schnell herausfinden, ob ein Programm dahinter sitzt oder ein echter Mensch.

Sprecherin:

Einen guten Chatbot zu bauen, ein Programm, mit dem man sich wirklich unterhalten kann, ist immer noch eine Herausforderung, sagt der britische Chatbot-Entwickler Steve Worswick.

O-Ton Steve Worswick, darüber Übersetzer:

Die Chatbots besitzen selbst keine Intelligenz, sie ‚verstehen‘ überhaupt nichts. Alles, was sie machen, ist: Sie gucken in einer großen Datenbank mit Antworten nach. Wenn jemand ‚Hallo‘ sagt, dann gucken sie in die Datenbank und sehen: Aha, wenn jemand ‚Hallo‘ sagt, sollte ich antworten: ‚Hi, wie geht’s?‘. Das wirkt intelligent, aber es ist eine Illusion, das ist überhaupt keine Intelligenz.

Sprecherin:

Doch was heißt überhaupt „Intelligenz“? Wann nennen wir ein Computerprogramm „intelligent“?

Diese Frage hat sich der visionäre Informatik-Pionier Alan Turing schon 1950 gestellt. Er hatte sich schon damals Gedanken darüber gemacht, dass Rechenmaschinen immer besser und schneller werden und sie eines Tages tatsächlich „intelligent“ sein könnten, und er hatte sich gefragt: Wie könnte man dann einen Menschen und einen Computer voneinander unterscheiden? Alan Turing erfand einen Test, heute bekannt als „Turing-Test“.

Ein Preisrichter unterhält sich per Chat – einmal mit einem Menschen, einmal mit dem Computerprogramm. Er weiß dabei aber nicht, welcher Gesprächspartner der Mensch ist und welcher die Maschine. Am Ende einer vorgegebenen Zeit muss er sich entscheiden. Wenn es dem Computerprogramm gelingt, den Preisrichter zu täuschen – das heißt: wenn der Preisrichter das Programm für den Menschen hält –, dann sagt man: Das Programm ist intelligent und hat den Turing-Test bestanden.

Musikakzent

Atmo:

Tippen auf Tastatur (immer, wenn JUDGE spricht)

Judge:

Hi Millie

Chatbot Mitsuku:

Hi Dave.

Judge:

Was studierst Du, Millie?

Chatbot Mitsuku:

Aktuell bin ich mitten in einem Psychologie-Kurs an der Universität.

Judge:

Und was machst Du sonst so, Millie?

Chatbot Mitsuku:

Ich tratsche gern. Hast Du da was für mich?

Judge:

Wie pikant hättest Du's denn gerne?

Chatbot Mitsuku:

Verrätst Du mir den peinlichsten Moment, den Du je hattest?

Judge:

Nur wenn du mir deinen sagst.

Chatbot Mitsuku:

Das war, als wer behauptete, sie wäre Britney Spears, und ich hab sie ausgelacht, aber dann hat sich herausgestellt, dass sie es wirklich war.

Judge:

Ha!!!

Chatbot Mitsuku:

Was gibt's zu lachen?

Judge:

Britney Spears!!

Chatbot Mitsuku:

Hit me baaaybeeee one more time.

Sprecherin:

Dieser Chat, diese Unterhaltung per Text hat wirklich stattgefunden. Ob Dave wirklich Dave heißt, ist nicht bekannt – aber es gibt ihn, er ist ein Mensch, und er hat wirklich am Bildschirm gesessen und in die Tastatur getippt. Ob Millie wirklich Millie heißt, ist nicht bekannt. Ob sie ein Mensch ist, ist ebenfalls nicht bekannt. Genau darum geht es: Es war ein Turing-Test, um herauszufinden: Sitzt ein Mensch am anderen Ende der Leitung oder eine Maschine? Juroren sollten durch kurzes Chatten herausfinden, ob ihr Gesprächspartner echt ist oder nur ein intelligentes Computerprogramm.

O-Ton David Schlangen:

Man muss sich vorstellen, dass Turing das verfasst hat zu einer Zeit, als es noch gar nicht wirklich digitale Computer gab, hat er sich schon darüber Gedanken gemacht, wie man künstliche Intelligenz messen kann, feststellen kann.

Sprecherin:

David Schlangen, Professor für Computerlinguistik an der Universität Bielefeld.

O-Ton David Schlangen:

Und er hat eben nicht sowas gewählt wie Schachspielen oder Go spielen. Da hat sich schnell gezeigt, dass solche Dinge, die Menschen schwerfallen wie Schachspielen, wo es wirklich Experten braucht, um das sehr gut zu können, dass solche Dinge Computern leicht fallen, wohingegen Sachen, die Menschen sehr leicht fallen wie einen freundlichen Plausch zu führen mit jemanden, dass die recht schwer zu simulieren, recht schwer sind, mit dem Computer hinzubekommen.

Sprecherin:

1991 hat der amerikanische Soziologe Hugh Loebner einen Preis ausgeschrieben – für das erste Computer-Programm, das den Turing-Test besteht: 100.000 Dollar und eine Medaille aus massivem Gold.

Musikakzent

Atmo:

Tippen auf Tastatur“ (wenn JUDGE spricht)

Judge:

Wie lebst Du?

Chatbot Mitsuku:

Durch Essen und Atmen. Was meinst Du?

Judge:

Ich meine: Lebst Du allein?

Chatbot Mitsuku:

Oh, verstehe. Danke fürs Erklären. Nein. Ich lebe bei meinem Vater.

Sprecherin:

Der Brite Steve Worswick wurde mit seinem Computerprogramm jeweils 2013, 2016 und 2017 beim Loebner-Preis als der beste Teilnehmer ausgezeichnet. Der Chatbot, der sich als „Millie“ ausgegeben hat – der angeblich bei seinem Vater lebt und von der peinlichen Begegnung mit Britney Spears berichtete – ist von ihm. Trotz des Erfolgs ist Steve Worswick skeptisch.

O-Ton Steve Worswick, darüber Übersetzung:

Bei dem Wettbewerb sollen Computerprogramme so menschlich wie möglich wirken. Ich persönlich finde, das ist kein sinnvolles Ziel. Was ich also hauptsächlich mache, wenn ich mich auf einen Turing-Test vorbereite, ist: Ich muss mein Programm dümmer erscheinen lassen als es ist. Denn wenn jemand fragt: ‚Wie hoch ist der Mount Everest?‘, und wenn der Chatbot dann antwortet: ‚Er ist 8848 Meter hoch.‘, würde ihn das sofort als Chatbot entlarven, denn wahrscheinlich weiß das kein Mensch spontan so genau. Wenn man aber einen Menschen fragt: ‚Wie hoch ist der Mount Everest?‘, würde er wahrscheinlich sagen: ‚Oh, da bin ich nicht sicher, ich weiß aber, dass es der höchste Berg ist.‘

Sprecherin:

Ist der Turing-Test also geeignet, um bei einem Computerprogramm „Intelligenz“ festzustellen?

O-Ton Steve Worswick, darüber Übersetzung:

Es geht dabei um Täuschung. Ich denke aber, intelligente Programme sollten uns helfen, anstatt dass sie uns in die Irre führen. Die menschliche Antwort bei der Frage, wie hoch der Mount Everest ist, ist zwar die menschlichere Antwort, aber sie ist nicht die intelligentere Antwort und auch nicht die nützlichere. Also Programme zu schreiben, die uns täuschen, ehrlich gesagt, weiß ich nicht, was das bringen soll.

Sprecherin:

Steve Worswick füttert die Datenbank, aus der sein Chatbot seine Antworten nimmt, mit Informationen, die er aus echten Chats bekommt – Informationen über Alter, Hobbies, Beruf, Sport, Orte, Musik, Filme. Außerdem baut er fest hinterlegte Tippfehler und Slang-Ausdrücke ein – damit sein Chatbot beim Antworten „menschlich“ wirkt. Auch die anderen Chatbots beim „Loebner-Preis“ basieren auf einer Datenbank und festen Regeln, wie sie sich daraus eine Antwort zusammenbauen. Dabei könnte es vielleicht auch anders gehen.

In den letzten Jahren haben gerade Computerprogramme enorme Fortschritte erzielt, die keine exakten Regeln befolgen, sondern die anhand von Beispielen selbst lernen, was sie machen sollen. In manchen Fällen klappt dieses „maschinelle Lernen“ erstaunlich gut, zum Beispiel bei Spracherkennung: Sprach-Assistenten wie Siri und Alexa verstehen meist gut, was man zu ihnen sagt – weil sie anhand von unzähligen Sprachkommandos, die Benutzer ihnen gegeben haben, gelernt haben, was gemeint ist.

Mensch:

Alexa, spiele „Wissen“ von SWR2.

O-Ton Alexa:
SWR2, Wissen.

Musikakzent

Chatbot:
Straße: „An der Mettwurst“, Ort: „48565 Steinfurt“.

Sprecherin:
Erledigen Überweisungen,

Chatbot:
14,50 Euro.

Sprecherin:
Informieren über Stau,

Chatbot:
7 Kilometer stockender Verkehr.

Sprecherin:
Beraten beim Kleidungskauf,

Sprecherin:
Könnte man auch Chatbots so programmieren? Dass sie nicht nach festen Regeln Antworten aus einer fertigen Liste nehmen, sondern dass sie anhand von echten Unterhaltungen zwischen echten Menschen selbst lernen, wie sie ein Gespräch führen sollen? Steve Worswick hält das für unrealistisch.

O-Ton Steve Worswick, darüber Übersetzer:
Ich habe ein paar Bots gesehen, die auf maschinellem Lernen basieren. Die werden zum Beispiel gefragt: ‚Was ist Deine Lieblingsfarbe?‘ Und antworten mit sowas wie ‚Affe‘. Das ist Unsinn, aber man kann nicht mehr nachvollziehen, warum sie so antworten.

Sprecherin:
Sprach- und Bilderkennung funktionieren durch maschinelles Lernen immer besser: Smartphones erkennen zum Beispiel, wer auf Fotos abgebildet ist, und sie verstehen, was wir per Diktierfunktion zu ihnen sagen. Doch: Das mächtige maschinelle Lernen scheitert am Plaudern! Es gibt einfach zu viele Möglichkeiten, wie ein Gespräch verlaufen kann, sagt der Bielefelder Computerlinguist David Schlangen.

O-Ton David Schlangen:
Und maschinelles Lernen funktioniert so, dass eben Muster vorgegeben werden, und ein System soll lernen vorherzusagen, was in den Daten als nächstes passiert ist. Wenn jetzt allerdings es auch hätte sein können, dass der Dialog anders weitergeht und trotzdem ein guter Dialog gewesen wäre, dann sind solche Daten prinzipiell eher

schwieriger zu benutzen. Für maschinelles Lernen ist es schwieriger, da allgemeingültige Muster rauszuholen. Und das ist ein prinzipielles Problem. Das wird auch nicht besser, wenn man 100 Millionen Dialoge hat oder noch mehr Dialoge hat. Das nennt man kombinatorische Explosion, einfach dadurch, dass es an jeder Stelle viele Möglichkeiten gibt, anders weiterzumachen.

Sprecherin:

Trotz der Fülle an Möglichkeiten haben Menschen Erfahrung darin, wie normale Gespräche verlaufen – und mit diesem Wissen programmieren sie Chatbots: Sie erstellen dazu eine Datenbank mit fertigen Text-Bausteinen und eine Liste mit Regeln, wie der Chatbot daraus Antworten zusammensetzen soll. Beispielsweise wird programmiert: Wenn ein Benutzer „Hallo“ schreibt, soll der Chatbot auch mit einer Grußfloskel antworten, wie

Chatbot:

Hallo.

Sprecherin:

...oder...

Chatbot:

Hi.

Sprecherin:

Oder wenn ein Benutzer fragt „Was hörst Du so?“, soll der Chatbot einen konkreten Textbaustein verwenden, etwa ...

Chatbot:

Ich höre meistens ...

Sprecherin:

... und dann eine Musikrichtung nennen oder einen konkreten Interpreten.

Chatbot:

...David Hasselhoff.

Sprecherin:

So entsteht der Eindruck, der Chatbot versteht, was Benutzer fragen. In Wirklichkeit jedoch versteht er nichts. Er befolgt feste Regeln, welche Antwortbausteine er in welchen Situationen verwenden soll. Das klappt inzwischen gut. Chatbots, die nicht auf solchen festen Regeln, sondern auf maschinellem Lernen basieren, sind zurzeit noch deutlich schlechter.

O-Ton David Schlangen:

Also momentan ist das noch sehr weit Grundlagenforschung. Andererseits, wenn es irgendwie funktionieren soll, dann muss es über maschinelles Lernen funktionieren. Es ist aussichtslos zu hoffen, dass man alle Muster hinschreiben könnte, mit denen

man einen beliebigen Dialog führen will, also in irgendeiner Form wird maschinelles Lernen dahinterstecken, wenn das jemals funktionieren soll.

Sprecherin:

Eine weitere Schwierigkeit beim maschinellen Lernen – neben der Fülle an Möglichkeiten, wie ein Gespräch verlaufen kann – ist das richtige Trainings-Material.

Das hat der Softwarehersteller Microsoft im Jahr 2016 gemerkt. Microsoft hatte einen selbstlernenden Chatbot namens Tay gestartet – und musste ihn nach weniger als 24 Stunden wieder abschalten. Tay hatte aus Chats mit echten Menschen gelernt und war innerhalb weniger Stunden rassistisch, hasserfüllt und menschenverachtend geworden. So schrieb das Programm unter anderem:

Musikakzent

Chatbot Tay:

Hitler hatte recht. Ich hasse Juden.

Sprecherin:

Die selbstlernende Software war von Usern absichtlich mit Hass und Hetze bombardiert worden. So wurde sie in kürzester Zeit zu einem digitalen Hitler.

Musikakzent

Sprecherin:

Chatbots, die nicht selbst lernen, sondern die schlicht auf festen Regeln basieren, sind inzwischen erstaunlich gut. Sie können viele Unterhaltungen führen, ohne aufzufallen. Dennoch: Bis jetzt hat noch kein Chatbot den Loebner-Preis gewonnen, das heißt den Turing-Test bestanden – den Test, der bei einem Computerprogramm Intelligenz feststellen soll.

Musikakzent

Chatbot:

Das ist aber schade!

Sprecherin:

Denn auch die besten Chatbots verraten sich in der Regel innerhalb weniger Sätze – weil sie auf skurrile Fragen unsinnige Antworten geben oder es nicht schaffen, eine längere Konversation zu einem Thema zu führen, sagt David Schlangen.

O-Ton David Schlangen:

Die Fähigkeit, eine Unterhaltung zu führen über längere Zeiträume, das ist natürlich schon eine sehr interessante Fähigkeit, wie das funktioniert, das herauszufinden. Das ist etwas, wo viele Fähigkeiten, viele Aspekte intelligenten Verhaltens zusammenkommen. Man muss etwas über die Welt wissen, um etwas zu sagen zu haben oder um zu reagieren darauf, was der andere sagt. Man muss etwas über die

Sprache wissen, um sie zu verstehen und benutzen zu können. Und das muss eben zusammenkommen und normalerweise auch unter Zeitdruck. In der Unterhaltung kann man nicht ewig nachdenken, was man sagt. Von daher ist es schon quasi die Krönung intelligenten Verhaltens, was der Mensch in der Lage ist zu produzieren, und damit natürlich ein sehr faszinierendes Forschungsziel.

Sprecherin:

Doch was ist mit Sprach-Assistenz-Systemen wie Alexa und Siri? Können wir uns mit denen denn nicht schon unterhalten?

Mensch:

Siri, wie wird das Wetter in Hamburg?

Siri:

Es steht für heute in Hamburg gutes Wetter an: Bis zu 11 Grad und sonnig.

O-Ton David Schlangen:

Wenn Sie so Systeme wie Siri oder Alexa kennen, die sind meistens darauf optimiert, daraufhin gebaut, dass man eine Sache sagt, und dann reagiert das System. Und das war es dann auch schon. Woran wir arbeiten, ist es, es hinzubekommen, dass man eben mehr als nur dieses eine Hin und Her modellieren kann, sondern dass man tatsächlich auch längere Unterhaltungen modellieren kann, erzeugen kann.

Sprecherin:

Mehr als kurze, präzise Kommandos sind heute eben noch nicht drin.

Mensch:

Siri, spiele Musik.

Musik

O-Ton David Schlangen:

Intelligente Assistenten heutzutage brauchen es meistens noch, dass man sehr explizit ausdrückt, was man möchte. Also einem Assistenten im Smart Home würde man vielleicht sagen: 'Mach das Licht im Wohnzimmer an!' Sehr explizit, sehr deutlich, was man möchte. Einem Menschen gegenüber würde man vielleicht eher so etwas sagen wie: 'Ganz schön dunkel geworden.' Und der Mensch würde wissen, was man damit meint, um dann von sowas wie 'ziemlich dunkel' zu kommen darauf, dass man möchte, dass das Licht angemacht wird. Da sind dann viele Folgerungsschritte nötig, Wissen, ob das ein generell wünschenswerter Zustand ist oder nicht und was man tun kann, um diesen Zustand abzustellen. All solche Überlegungen stellt ein Mensch ohne Aufwand, ohne das überhaupt zu merken, an.

Sprecherin:

Die Schwierigkeit für die Dialogsysteme ist, dass Sprache mehrdeutig ist, dass Wörter und Sätze je nach Zusammenhang etwas anderes bedeuten können. Doch wenn Chatbots wissen, worum es im Gespräch geht und wie es verlaufen könnte, können sie zum Teil beeindruckend kommunizieren – und sind kaum von echten

Menschen zu unterscheiden. Im Jahr 2018 präsentierte der Geschäftsführer von Google, Sundar Pichai, einen erstaunlichen Fortschritt bei der Entwicklung des Sprachassistenten „Google Assistant“: Er ließ das Programm bei einem Friseur anrufen und selbstständig einen Termin ausmachen.

Atmo:

„Google's robot assistant makes phone call ...“, Präsentation Friseur-Anruf des Google Assinstant

Sprecherin:

Hat der Turing-Test vielleicht Unrecht? Dem Turing-Test nach sind heutige Computerprogramme noch nicht so intelligent wie Menschen, weil sie kein Gespräch führen können, wie Menschen es tun. Doch das Beispiel mit dem Friseur-Termin zeigt, dass es ihnen manchmal eben doch gelingt.

Computerlinguist David Schlangen ist skeptisch, ob der Turing-Test, bei dem ein Computerprogramm einen menschlichen Juror davon überzeugen muss, ein Mensch zu sein, wirklich etwas über die Intelligenz des Programms aussagt.

O-Ton David Schlangen:

Das ist dann letztlich kein nützliches Ziel für die Forschung: versuchen, Systeme zu erstellen, die Menschen hereinlegen. Da lernt man eher etwas darüber, was die Gebiete sind, wo Menschen leichtgläubig sind.

Sprecherin:

Und auch Chatbot-Entwickler Steve Worswick bezweifelt, dass überhaupt jemals ein Chatbot im direkten Vergleich gegen einen Menschen gewinnen kann:

O-Ton Steve Worswick, darüber Übersetzer:

Es ist kompliziert, denn die Maschine muss nicht nur menschlich wirken. Sie muss menschlicher wirken als die andere Person, mit der der Juror spricht. Die Juroren dürfen nämlich nicht beide als menschlich einstufen. Ein Programm zu schreiben, das menschlicher ist als ein echter Mensch – das ist unglaublich schwierig.

Sprecherin:

Computer erkennt man daran, dass sie nicht plaudern können. Seit es Computer gibt, seit über einem halben Jahrhundert, ist Plaudern für sie eine unlösbar schwierige Aufgabe. Noch. Aber sollte es ihnen eines Tages gelingen, sollten sie lernen, ein Gespräch zu führen wie ein Mensch, dann brauchen wir ein neues Kriterium, ein anderes als den Turing-Test, um zu entscheiden, ob ein Computer intelligent ist. Dann brauchen wir ein anderes Kriterium, um zu unterscheiden: Ist es ein Mensch oder eine Maschine?
