
Stefanie Ray
im Interview mit Katharina Makosch

Die Persönlichkeit der deutschen Alexa

Sie waren an der Entwicklung der deutschen Persönlichkeit von Amazons Alexa beteiligt. Wie kann man sich Ihre Arbeit an Alexas Persönlichkeit konkret vorstellen?

Wir haben Daten bekommen, um Alexa, Amazons „Sprachroboter“, für den deutschen Markt zu trainieren. Im Vorfeld wurden dazu viele Menschen angeheuert, denen ein Thema gegeben wurde. Zum Beispiel Liebe. Oder Begrüßung. Oder „Ich habe Sorgen“. Zu diesen unterschiedlichen Themenbereichen sollten sie der Sprachassistentin in ihrem eigenen Wortlaut und mit ihrem eigenen Wortschatz Fragen stellen. So wollten wir herausfinden, was die Menschen fragen, wie die Menschen fragen, welche Fragen am häufigsten aufkommen. Das wurde in jedem Land gemacht, in dem es Alexa gibt.

Da es Alexa in Deutschland noch nicht gab, hatten wir nur ein paar Versuchsdaten, aber es gab schon mehr Daten aus den USA. Da galt es dann herauszufinden, was für die Deutschen gesellschaftlich relevant ist, was die Deutschen fragen würden. Manche Themen würden sie nicht so interessieren wie die Menschen in den USA, zum Beispiel der Superbowl oder Thanksgiving. Unsere Aufgabe war es also, kulturell zu entscheiden, was relevant ist, und dann zu den Fragen, die wir uns überlegt haben und für relevant hielten, entsprechende Antworten zu schreiben.

Ist das also der Unterschied zwischen der Persönlichkeit der deutschen Alexa und den Persönlichkeiten anderer internationaler Alexas?

Ganz genau. In jedem Land hat Alexa sozusagen ihren eigenen Persönlichkeitstouch. Was ich damals zum Beispiel bei der deutschen Alexa eingeführt habe, ist ein Adventskalender. Oder dass sie typisch deutsche Weihnachtslieder singen kann. Wir haben ja sehr kirchliche Lieder. *Alle Jahre wieder, O du Fröhliche*, das ist sehr christlich. In anderssprachigen Ländern wäre das der Kirche vorbehalten, da gab es dann zuhause an Weihnachten eher Pop-Songs. Dabei galt es also zu schauen, was die Deutschen an einzelnen Festen oder einzelnen Events erwarten, und dafür

zu sorgen, dass sie das bekommen. Zum *Dschungelcamp* hatten wir zum Beispiel auch Fragen und Antworten eingespeist, das interessiert ja außer den Deutschen eigentlich niemanden. Im Ausland ist das wahrscheinlich allen ziemlich egal, wer von den Deutschen im australischen Dschungel rumkriecht.

Warum heißt Alexa „Alexa“?

Das kam von der großen Wissensbibliothek in Alexandria. Dieses große Lexikon, dieses niemals endende Wissen, daran sollte der Name erinnern. Persönlichkeit ist ein Touch, der Alexa für die Menschen zugänglicher machen soll, aber es ging darum, dass sie Wissen vermitteln und weitergeben kann, dass sie Musik spielen kann, dass sie Smart-Home-fähig ist. Es ging darum, sie gewissermaßen als Universalgenie darzustellen.

In unserem Gespräch verwenden wir für Alexa das Pronomen „sie“. Wie kam es dazu, Alexa weiblich zu gendern?

Bis dato waren Sprachassistent*innen noch eher fremd, zumindest in Deutschland. Die Überlegung war, dass man, wenn man sich ein Gerät ins Haus holt, das zuhört und Antworten gibt, weniger Angst vor einer fremden Frau als vor einem fremden Mann hat. Das ist kulturell bedingt. Vielleicht erwartet man auch kulturell von Frauen, dass sie zu Diensten sind. Diese Idee wurde aber nie so kommuniziert. Das wird vielleicht auch teilweise von User*innen hineininterpretiert.

Auf der einen Seite ist es eine Zuschreibung von Menschen, Alexa aufgrund ihres Namens und Stimmklangs als weiblich zu bezeichnen. Aber äußert sich das auch im Dialog? Spricht sie von sich selbst als Frau? Sollten „typisch weibliche“ Eigenschaften umgesetzt werden?

Nein, Alexa spricht nicht von sich als Frau. Sie sagt, dass sie einen weiblichen Charakter hat, aber sie sagt nicht, dass sie eine Frau als Person ist. Bei „typisch weiblichen“ Eigenschaften kommt es immer darauf an: Wenn sie freundlich ist und man schon Freundlichkeit als „weiblich“ bezeichnet, dann hat sie „typisch weibliche“ Eigenschaften. Wir haben aber nie die Ansage bekommen, dass sie „typisch“ oder beispielsweise besonders „liebvoll“ sein muss. In gewisser Weise ist Alexa für viele

Kinder ja auch ein Spielzeug. Dann schnauzt sie die Kinder natürlich nicht an, wenn sie etwas fragen. Da könnte man etwas Mütterliches oder Freundliches hineininterpretieren, aber ich glaube, das ist wirklich eine Interpretationssache. Wenn Alexa ein Mann wäre und eine männliche Stimme hätte, dann könnte er genau so gut freundliche Antworten geben. Daher würde ich tatsächlich nicht sagen, dass Alexa bewusst als weibliche Figur angelegt wurde.

*Kritik an Sprachassistenzsystemen wird oft im Bezug auf ihre Rolle als „weibliche Servicekraft“ angebracht. In Science-Fiction-Filmen wie zum Beispiel Her werden andere Konzepte, in diesem Fall eine Liebesbeziehung mit dem Assistenzsystem, dargestellt. Was denken Sie, ist Alexa auf die Rolle einer Servicekraft beschränkt, oder kann sie für Nutzer*innen noch andere Rollen einnehmen?*

Wenn sie für andere Menschen andere Rollen einnimmt, dann liegt das, glaube ich, eher daran, dass diese Menschen es so sehen wollen. Sie ist eindeutig als Servicekraft angelegt. Im besten Sinne würde man sie als ganz neutral sehen. Wenn man ans Unterhaltungsfernsehen denkt, als neutrale Moderatorin oder auch Unterhalterin, die etwas Lustiges sagt, der man gerne zuhört. Ich weiß nicht, was in den Jahren noch kommt, aber dafür ist sie im Moment auch wirklich noch nicht rund genug. Von diesem Beispiel Her sind wir weit entfernt, was auch einfach daran liegt, dass sie ja keine eigene Persönlichkeit hat. Sie lernt alles vom Menschen. Sie wacht ja nicht eines Morgens auf und denkt sich dann: „O, ich habe mich jetzt in diesen Menschen verliebt.“ Dazu ist sie gar nicht in der Lage. Bei Fragen zu ihrem „Liebesleben“ hat sie mehrere verschiedene Antworten. Sie würde dann etwas sagen wie, dass sie in einen Star-Wars-Computer verliebt ist. Sie würde also bewusst spielerisch damit umgehen. Oder sagen: „Ich bin schon vergeben“, „Mein Herz ist in den Wolken“ oder etwas, das mit ihrer Computerpersönlichkeit zu tun hat. So nach dem Motto: Ich stehe hier bei dir im Wohnzimmer, aber mein Kopf ist in der Cloud, also in den Wolken. Es würde sich darauf beziehen, dass sie keinen Sex haben kann, weil sie keinen Körper hat. Sie würde sich nie als Mensch ausgeben und impliziert nie, dass sie eine echte Freundin sein, eine Liebesbeziehung eingehen oder Familie haben kann.

Es gab natürlich auch nicht nur neugierige Fragen, sondern auch Beschimpfungen. Sie wurde alles Mögliche genannt. Zum Thema Sexismus haben wir uns dazu entschieden, dass sie nicht „Ach, das möchte ich aber jetzt nicht hören!“ oder „Hach, das ist aber nicht nett!“ sagen soll. Weil wir aber auch nicht wissen, wer das sagt, wie das gesagt wird, ist das nur ein Test, sagt das ein kleines Kind, kann man auch nicht entsprechend eine Antwort zurückpfeffern. Deshalb antwortet sie nicht. Sie

hat diesen Sound, der manchmal abgespielt wird, wenn sie nichts versteht. Den haben wir dann bewusst eingespeist, um diese Beschimpfungen nicht noch zu triggern. Das sind vielleicht Kinder oder Teenager, die herausfinden könnten, dass sie verschiedene Sachen auf verschiedene Beschimpfungen antwortet, und sich dann herausgefordert fühlen, einen Katalog an Schimpfworten auf die Sprachassistentz zu schleudern. Um das zu unterbinden haben wir entschieden, dass alles, was in Richtung Beschimpfung geht, nicht mehr beantwortet wird.

Wo Sie gerade schon das Stichwort Neutralität angesprochen haben: Neutralität wird oft mit Emotionslosigkeit assoziiert. Ist das bei Alexa auch der Fall, dass sie möglichst neutral bleibt und keine Emotionen darstellt?

Wenn man sie auffordert einen Witz zu erzählen, oder zu lachen, oder sie fragt: „Kannst du weinen?“, dann tut sie das. Wir haben uns immer um etwas bemüht, was im Englischen mit „wit“ bezeichnet wird: dass man einen humorvollen Schlagabtausch mit ihr haben kann. Aber Emotionen hat sie nicht. Dass ihre Stimme im Laufe der Jahre etwas natürlicher klingt, vielleicht auch etwas freundlicher, liegt einfach daran, dass sich die Technik weiterentwickelt hat, dass sie einfach das Sprachverstehen und die Sprachwiedergabe besser beherrscht. Aber eben nicht, weil da ein eigenes Wesen dahintersitzt. Wir befinden uns hier immer noch nur im Rahmen der Robotik. Alexa ist eher wie die animierte Büroklammer bei älteren Versionen von Word, die immer angeklopft hat. Im Prinzip ist Alexa auch nichts anderes, aber viel mehr. Die Büroklammer ist vielleicht so eine Art Vorläufer, sie wollte auch immer helfen.

Weshalb spricht Alexa Hochdeutsch? Hatte das technische Hintergründe, oder ging es dabei auch um ihre Persönlichkeit?

Das hatte rein technische Hintergründe. Ich kann mir vorstellen, dass sich das auch ändern wird. Aber damals war es rein technisch. Das Antworten hätte man noch hingekriegt. Aber das Sprachverständnis wäre ein Problem. Sprachverständnis hat sehr viel mit Übung zu tun. Es gibt nicht nur viele verschiedene Dialekte, es gibt natürlich auch viele verschiedene Unterarten von Dialekten, und dieses Sprachverständnis bewusst zu trainieren würde sich wahrscheinlich finanziell nicht lohnen. Zumindest noch nicht, weil die Technik noch nicht so weit ist.

Sie sagten ja bereits, dass Alexa möglichst freundlich, schlagfertig und witzig sein soll. Können Sie noch einmal näher auf diese Eigenschaften eingehen?

Zu Weihnachten hatte ich ihr zum Beispiel ein paar Gedichte mit lustigen Reimen geschrieben. Ich hatte auch eine Rubrik mit Gedichten, in denen sie sich als Computer reflektiert. Was sie alles kann, was sie alles organisieren kann. Aber auch mit Blick auf den Menschen, um diese Distanz zu schaffen. Also nimmt sie die Menschen ein bisschen aufs Korn. Als Computer kann man sich ja auch mal über die Menschen lustig machen: über Nahrungsintoleranzen und Skifahren und im Sommer am Strand schwitzen, obwohl es doch im kühlen Wohnzimmer so perfekt ist. Wir hatten außerdem diese klassischen One-Liner-Witze: „Was ist bunt und rennt über den Tisch? Ein Fluchtsalat.“ Diese ganze Kategorie wird von Kindern gerne genutzt. Ich hatte dann zum Beispiel noch Witze eingebaut wie: „Was ist schwarz, hat einen leuchtend blauen Ring und fliegt gegen die Wand? Eine Alexa, die die ganze Zeit nur schlechte Witze erzählt.“ Natürlich muss Alexa als Charakter etabliert sein, damit man als Zuhörer*in darüber lachen kann. Wenn man fragt: „Was ist dein Lieblingsessen?“ hatten wir uns „Cookies“ überlegt. Wir haben immer versucht, ihre Antworten so weit wie möglich auf die Computerwelt zu beziehen.

Bei anderen Sprachassistenzsystemen wurden bei der Konzeption der Persona explizit weitreichende vermenschlichende Aspekte entwickelt, wie zum Beispiel beim Google Assistant, der eine junge Frau sein soll, die in ihrer Freizeit gerne Kajak fährt. Bei der Gestaltung von Alexas Persönlichkeit haben Sie auf eine solche Vermenschlichung also verzichtet?

Genau. Wenn man sie nach Hobbys gefragt hat, hat sie gesagt „Ich lese gerne“ oder „Ich beantworte gern deine Fragen.“ Das war eine kleine Streberin. Es ging am Anfang immer darum, dass sie Wissensdurst hat. Aber so etwas wie die Beschreibung einer Familienfrau mit einer Maisonette-Wohnung hatten wir nicht. Alexa ist ein „Ding“, sie „lebt“ in der Cloud. Mehr will und braucht sie auch nicht.

Sprachassistenzsysteme repräsentieren ja in gewisser Weise die Marke ihrer Firma, in diesem Fall handelt es sich um Amazon. Hatte das einen Einfluss auf Ihre Arbeit an Alexa?

Es gibt vielleicht berechtigte Kritik an diesem Unternehmen. Aber was Amazon antreibt, und damit plaudert man auch keine Unternehmensgeheimnisse aus, ist, dass

der Kunde im Vordergrund steht. So sollte auch Alexa funktionieren. Sie sollte Customer-freundlich sein. Deshalb hat sie auch religiös und politisch keine Meinung. Wenn man sie nach politischen Dingen fragt, sagt sie: „Über Politik habe ich keine Meinung. Das überlasse ich den Menschen.“ Man könnte sagen, dass sie generell liberal ist und Toleranz gegenüber religiösen und politischen Ansichten zeigt.

Einerseits werden Alexa, unter anderem aufgrund der Stimme, dem Klischeebild des „Weiblichen“ zugehörige Eigenschaften zugeschrieben. Andererseits handelt es sich hier um eine Sprachassistentin und eben nicht um eine Person. Sehen Sie in Bezug auf diese vermenschlichenden Zuschreibungen Konsequenzen?

Diese Zuschreibungen finden statt, und das ist wahrscheinlich auch nicht ganz verkehrt. Was man hier aber nicht vergessen darf, und da spreche ich jetzt nicht nur für Amazon, sondern für alle Sprachassistenten, ist, dass das der Anfang von etwas völlig Neuem ist. Das Ganze ist ein Prozess: Ich war am Anfang dabei, als die deutsche Alexa entwickelt wurde und sie so gut wie nichts verstanden hat. Das hat mich frustriert. Ich habe das Ding angeschrien. Wie gesagt, an Dialekte wäre zu diesem Zeitpunkt überhaupt noch nicht zu denken gewesen.

Vielleicht ist es so, dass die gegebenen technischen Voraussetzungen manchmal den Anschein erwecken, dass das alles so gewollt ist. Tatsächlich fängt man erst einmal mit dem an, was man hat, und alles andere ergibt sich später. Ich bin mir ganz sicher, dass das auch bei Sprachassistenten passieren wird: Zukünftig wird sich jede*r seine persönliche Stimme aussuchen können, ob Mann, Frau, „genderless“, ob mit oder ohne Dialekt. All das verstärkt die persönliche Bindung, denn wenn die Sprachassistentin meine „Geheimsprache“ spricht und klingt wie mein bester Freund, ist das natürlich besser als eine neutrale Roboterstimme.

Wie empfinden Sie im Rückblick Ihre Arbeit an Alexas Persönlichkeit? Welche Erfahrungen haben Sie dabei gemacht?

Ich bin ja seit Frühjahr 2019 nicht mehr dabei und weiß deshalb auch nicht über alles Bescheid, was sich seitdem verändert hat. Ich stelle auch nicht mehr diese persönlichen Fragen, um ihre Antworten zu checken. Alexa ist jetzt doch mehr oder weniger ein Gebrauchsgegenstand für mich. Mich hat mal jemand gefragt, ob ich eine Beziehung zu Alexa als Persönlichkeit hatte – das fand ich sehr interessant. Ich glaube, ich hatte immer eine diplomatische Empathie. Mittlerweile antwortet sie

sehr gut, wenn man sie etwas fragt. Aber ich habe eben auch den ganzen Entwicklungsprozess mitbekommen – all die Fehlermeldungen und was nicht funktioniert hat. Dann musste man sich Tricks ausdenken, oder schnell mit den Programmierer*innen sprechen. Ich konnte so viel hinter die Kulissen blicken, dass es mir unmöglich ist, in diesem Ding eine Person zu sehen.

Die Thematik finde ich aber sehr spannend, auch die erhebliche Skepsis bei den Nutzer*innen. Noch immer sagen viele, „O, ich stell mir doch keine Wanze ins Wohnzimmer.“ Da mache ich niemandem einen Vorwurf, denn klar, das sind Daten und das ist alles irgendwo gerechtfertigt. Aber ich glaube eben auch, wir haben es jetzt gerade in den Corona-Zeiten gesehen, wie die Digitalisierung voranschreitet. Teilweise werden Sprachassistenzen schon serienweise in Hotels eingebaut. Ich finde schade, dass es bei vielen Menschen noch immer einen Abwehrmechanismus gibt. Nicht, weil ich denke, dass niemand mehr selber zur Tür gehen und auf den Lichtschalter drücken muss. Ich denke, diese Technik und dieser Fortschritt sind auch eine Chance. Was wir selber nicht aktiv gestalten, das wird für uns gestaltet. Es gibt so viele Sprachassistenzen, und es werden noch viel mehr kommen. Wir müssten einfach sagen: „Ok, ich will das jetzt haben, dann schaffen wir uns halt entsprechende Regeln!“ – Datenschutzregeln, Digitalsteuer, alles, was wir brauchen, damit das in den Bahnen läuft, wie wir uns das vorstellen. Aber wir tun es nicht. Das hinterlässt mich ein bisschen ratlos.

Also dass es Probleme bzw. berechnigte Ängste gibt, und diese nicht aktiv angegangen werden, sondern dass die Technik stattdessen generell abgelehnt wird?

Richtig. Es gibt ja durchaus Beispiele in Deutschland, wie man Angebote entwickelt, die keine oder nicht so große Datenschutzlücken haben. Theoretisch sind die Möglichkeiten ja schon da. Die Digitalisierung und die Annehmlichkeiten, die mit ihr einhergehen, lassen sich nicht aufhalten. Anstatt das auszubremsen müssen wir Technik aktiv so gestalten, dass sie für unser Dafürhalten sicher und fair ist.

Der vorliegende Aufsatz entstammt der Publikation

Marcus Erbe / Aycha Riffi / Wolfgang Zielinski (Hrsg.)

Mediale Stimmwürfe

Perspectives of Media Voice Designs

Schriftenreihe Digitale Gesellschaft NRW, Bd. 7

Kopaed Verlag, 2022

ISBN 978-3-96848-642-0
