

KI-Avatare im Alltag: Wie digitale Assistenten Akzeptanz finden

Ausgesprochen digital. Der Podcast für digitale Trends.

Intro

[00:00:02.120] - Steffen Wenzel

Ausgesprochen digital. Der Podcast für digitale Trends. Sind Sie sicher, dass dies eine Anmoderation von einer lebenden Person ist, oder könnte auch hier ein Avatar sprechen? Jetzt könnten Sie sagen: Na ja, die leicht hessische Einfärbung, das bekommt wohl noch kein Avatar hin. Aber vielleicht doch. Keine Sorge. Hier spricht ihr mittlerweile KI gestählter Moderator von Ausgesprochen digital in der 51. Folge dieses wunderbaren Podcasts, zu der ich Sie herzlich begrüße und in der wir uns mit dem Thema KI Avatare beschäftigen wollen. Martin Wunderwald ist heute zu Gast bei uns. Hallo Martin.

[00:00:44.810] - Martin Wunderwald

Hallo Steffen, Grüß dich!

[00:00:46.520] - Steffen Wenzel

Ja, Martin, In deiner täglichen Arbeit als Portfolio Lead für AI und Cognitive Services übersetzt du unter anderem Kundenherausforderungen in passende KI-Lösungen. Eine mögliche Lösung können KI-Avatare sein, über die wir heute sprechen wollen. Was denkst du, könnte ein Avatar die Moderation für einen Podcast übernehmen? Ich frage natürlich nur für einen Freund.

[00:01:09.020] - Martin Wunderwald

Ja, also in jedem Fall. Ich glaube, das ist sogar eine der leichtesten Übungen, die ein Avatar übernehmen kann. Wenn man überlegt, dass es inzwischen ja schon sozusagen generierte Podcasts aus Texten und Dokumenten gibt. Ja.

[00:01:22.310] - Steffen Wenzel

Okay, das beruhigt mich jetzt nicht besonders. Aber gibt es da Filter? Nicht so ein bisschen die menschliche Komponente oder der Dialekt?

[00:01:29.780] - Martin Wunderwald

Stimmt, das macht es natürlich sympathisch. Also die menschliche Komponente ist immer wieder das, was den Vorsprung vor der KI in jeder Disziplin behalten wird. Und es wird immer Leute geben, die das vor sozusagen automatischen oder roboähnlichen Stimmen oder ja Avataren bevorzugen werden.

Definition und Abgrenzung von KI-Avataren

[00:01:49.010] - Steffen Wenzel

Lass uns zu Beginn mal darüber reden, was Avatare überhaupt sind und wie man sie definieren kann. Also kannst du uns mal helfen?

[00:01:56.120] - Martin Wunderwald

Ja, ich versuche es auch mal! Also für mich ist es sozusagen etwas Menschenähnliches bis hin zu hyperrealistischen Abbildungen von Menschen, die sich in einem virtuellen Raum oder Endgerät

bewegen und letzten Endes Menschen in ihrer Sprache und mit Inhalten sozusagen simulieren. Ja, das ist für mich so die Abgrenzung.

[00:02:21.300] - Steffen Wenzel

Und was wäre jetzt so der Unterschied zu einem Voicebot oder zu einem Chatbot?

[00:02:26.520] - Martin Wunderwald

Genau. Also den Avatar, den kann man sich sozusagen vorstellen als insofern personifiziert ist dann vielleicht nicht das richtige Wort, aber einer personifizierten äußereren Erscheinung, die eben mehr als lediglich einen Sprachkanal hat, sondern eben auch ein Antlitz, die die Lippen bewegt, die gegebenenfalls sogar Arme und Körper bewegt. Egal, ob es jetzt sozusagen in dem virtuellen 3D Raum oder als 2D ist und dem gegenüber suggeriert, dass sie menschenähnlich ist. Also so die Vorstufe zum Roboter, sage ich mal, der ja in der physischen Welt augenscheinlich roboterähnlich ist. Aber ähnliches kann man eben sozusagen als Vorstufe betrachten, dass der Avatar dies in einem virtuellen oder in einem zweidimensionalen Raum ist.

Technologische Entwicklungen von KI-Avataren

[00:03:13.650] - Steffen Wenzel

Avatare gibt es ja jetzt schon eine Zeit lang. Jetzt mit künstlicher Intelligenz hat sich natürlich einiges noch mal verändert. Auf welcher Technologie beruhen denn eigentlich Avatare und welche Entwicklung hat denn da in den letzten Jahren stattgefunden, dass das jetzt so einen enormen Sprung getan hat?

[00:03:30.210] - Martin Wunderwald

Eine gute Frage. Also Avatare würde ich sagen. Sie waren schon immer da, aber es gab schon immer den Wunsch, sozusagen insbesondere im Servicebereich solche Schnittstellen zu digitalisieren, die zu echten Kundenberatern oder zu IT Service Desk oder bei einer Krankenkasse zum Service. Quasi die Schnittstellen waren, wo man einfach Fragen in natürlicher Sprache in ein Chatfenster eingeben konnte, weil die Menge der Informationen einfach so groß war, dass man sie nicht sich nicht nichts durchlesen will, sondern einfach einen Dialog mit einer Person oder eben einem Chat führen wollte. Und insofern kam der Chat wahrscheinlich als erstes. Ja und? Und sozusagen das virtuelle Gegenüber, was in Vergangenheit sehr stark regelbasiert beantwortet wurde. Also soll heißen, im Hintergrund gab es quasi Mechanismen, die anhand der Frage und Stichwörtern oder bestimmten Schlagworten innerhalb der Frage erkannt haben, was die passende Antwort sein konnte und diese dann in der Datenbank gesucht haben und angezeigt haben, so dass die einfachste Form von so einem Support oder von einem Chat. Und jetzt mit den Sprachmodellen kommt es natürlich auf eine völlig neue Ebene.

[00:04:43.440] - Martin Wunderwald

Das heißt also, egal wie komplex und wie verschachtelt sich eine Person, ein Mensch ausdrückt, in so einem Chat wird das Sprachmodell im Hintergrund das Anliegen verstehen und kann es viel besser sozusagen zu den Inhalten, die im Hintergrund bereitgehalten werden, mappen und dann eben die passende Antwort geben in einer wahrscheinlich höheren Qualität als aus einer Datenbank vorgefertigte Antworten.

[00:05:06.750] - Steffen Wenzel

Es gibt ja verschiedene Sprachmodelle. Gibt es da welche, die besonders geeignet sind, um sie bei Avataren einzusetzen?

[00:05:14.190] - Martin Wunderwald

Naja, dadurch, dass sozusagen jetzt erst diese natürliche Sprachlichkeit in diesen Sprachmodellen stattfindet, kommt überhaupt das Bedürfnis zu sagen, man sollte jetzt auch ein äußeres, also irgendwie

ein Erscheinungsbild dieser Person geben. Ja, wenn man zurückdenkt, was die ersten Avatare waren, sind das ja eher sozusagen so ich will nicht sagen doch vielleicht Maskottchen von Krankenkassen. Daran muss ich immer denken. Ja, also oder Versicherungen haben sowas oft genutzt, die sozusagen eine Alternative Kunden Kommunikationsschnittstelle gesucht haben zum klassischen Durchsuchen der Website nach Informationen. Und die wollten sie dann also einfach ein Äußeres geben. Die waren aber eher comichaft und sehr statisch, auch in ihrem Verhalten. Also im Zweifel haben die sich gar nicht bewegt und waren einfach auf der Webseite über dem Chatschlitz. Und inzwischen dadurch, dass es eben sehr natürlich sprachlich ist, kommen jetzt Animationen also dazu. Also man versucht die Avatare menschlicher aussehen zu lassen, immer näher, sozusagen an die reale Person heran, mit dem Ziel, dass man dann sich sozusagen als Kunde oder als Benutzer dieses Avatars viel stärker verbunden fühlt. Zu dem, mit dem man mit der Person, mit der virtuellen, mit der man spricht.

Akzeptanz und Vorbehalte gegenüber KI-Avataren

[00:06:33.060] - Steffen Wenzel

Es gibt ja verschiedene Studien, unter anderem von Capgemini. Die sagen zum Beispiel, dass es eine große Akzeptanz auch für Avatare gibt. Auf der anderen Seite wird auch oft noch der persönliche Assistent vorgezogen. Woran liegt das deiner Meinung nach?

[00:06:49.020] - Martin Wunderwald

Also ich hätte auch gerne einen persönlichen Assistenten. Leider ist das bis jetzt, hat sich das noch nicht eingestellt. Nee, Also ich verstehe natürlich diese Frage. Es gibt eine Reihe Beweggründe, warum Leute sozusagen, die den Kontakt mit interaktiven Medien oder interaktiven Avataren ein Stück weit scheuen. Also sei es quasi die Berührungsangst mit dem Nichtmenschlichen, Ja und dem Roboterähnlichen. Oder ist es einfach generell die Ablehnung dieser Art der Kommunikation zu einem bestimmten Thema? Und ich glaube, es wird immer Personen oder Menschen geben, die nur den physischen Kontakt mit Personen bevorzugen, wenn man einfach auch historisch schaut, wie einfach Verbindung, wie Geschäfte gemacht wurden. Das ist immer zwischen Personen per Handschlag und und dieses Vertrauen, was man auch aufbaut in zwischenmenschlichen Beziehungen, also zwischen dir und mir, wie wir uns hier auch virtuell sehen. Das wäre wahrscheinlich noch größer, auch wenn es schon groß ist, wenn wir uns physisch gegenübersetzen an einem Tisch und wir wirklich die gesamte Mimik und Gestik wahrnehmen können und darauf letzten Endes unsere Kommunikation aufbauen können. Also insofern sind die Vorbehalte durchaus berechtigt.

[00:08:12.180] - Martin Wunderwald

Aber die Vorteile sind eben nicht von der Hand zu weisen. Und die Potenziale, insbesondere in der Automatisierung von Aufgaben, die wo sehr viele Menschen gebunden sind, eben zum Beispiel solchen Call Centern, dort ist das große Potenzial, was die Forschung vorantreibt.

[00:08:30.690] - Steffen Wenzel

Ich glaube, da gibt es natürlich auch eine große Angst hinsichtlich, dass man gar nicht mehr weiß, ob man mit einem Avatar spricht. Ich sage mal, so was wie Deepfakes haben wir ja auch. Auf einer anderen Ebene ist es da wichtig, auch weiterhin oder auch in Zukunft Transparenz auch zu zeigen und nicht zu sagen na ja, also wir wollen so gut sein, dass man den Unterschied zu einem Menschen gar nicht mehr feststellt. Und es ist doch umso besser, wenn der Kunde beispielsweise jetzt in einem Servicebereich gar nicht mehr merkt, dass da ein Avatar mit ihm redet.

[00:09:02.850] - Martin Wunderwald

Ja, das ist wirklich. Also es ist ja auch ein Forschungsfeld, sozusagen. Also inwieweit diese Akzeptanz der Avatare oder das Aussehen, ob es quasi besser ist, ist sehr menschlich zu gestalten und hyperrealistisch, also fotografisch echt. Wo wir dann in genau das von dir benannte in diese Deepfake-Bereiche hineinkommen oder ob man augenscheinlich eine gewisse Künstlichkeit in dem Äußeren lässt, um

einfach nicht dieses Unwohlsein zu erzeugen. Bei einer Person. Also man man ist dort noch auf der Suche nach dem Besten, sozusagen nach der nach der besten Verkörperlichung. Der Trend, den ich jetzt sehe bei vielen Anbietern eben aus dieser wahrscheinlich ein Stück weit aus diesem Wow Effekt, dass man Menschen sehr ähnlich nachbauen kann. Getrieben ist doch dieses hyperrealistische, ja, dass man wirklich versucht, in einem zweidimensionalen Video Personen so realistisch abzubilden, dass es kaum ein Unterschied erkennbar ist. Und dann ist wirklich die Frage wie ist dann die Akzeptanz, ja der Kern an der Stelle oder das, was man natürlich erreichen will, ist je höher die quasi die Ähnlichkeit zu einem echten Menschen, desto eher ist die Hypothese, desto stärker das Vertrauen.

[00:10:23.580] - Martin Wunderwald

Also dieses innerliche Commitment in so einem Gespräch, was ja dann auch das ist, was letzten Endes zu einem Geschäftsabschluss oder zu einer Qualität in irgendeiner Form führt. Oder in einem Wohlempfinden und einem positiven Gesprächsverlauf. Das versucht man natürlich zu maximieren. Es gibt aber in gewissen Risikobereich, die nennt man in dem Fall Uncanny Valley, also so ein unsicheres Tal. Und der kommt nämlich genau dann, wenn versucht wird, den Avatar so echt wie möglich zu gestalten. Das aber nicht gelingt vollständig. Ich mache ein Beispiel. Wir haben eine Person, die als Avatar vor einem steht, vielleicht auch lebensgroß oder im Web. Und sie scheint sich aber seltsam zu bewegen und hat das Gesicht, hat Gestik und Mimik, aber die Arme hängen schlaff herunter. Also es gibt gar keine Interaktion mit dem Körper und das ist einfach dann etwas, wo man ins Zweifeln kommt und anfängt überhaupt darüber nachzudenken als Gegenüber und was als sehr, was einfach Unsicherheit hervorruft. Und in diesem Uncanny Valley möchte man sich eigentlich nicht befinden in so einer Kommunikation. Deswegen ist es entweder die wirklich hyperrealistische und sehr qualitative Darstellung solcher Personen oder Avatare oder die augenscheinlich künstliche, aber trotzdem qualitativ hochwertige Darstellung, wo man gerade versucht, hinzusteuer. Ja.

Einsatzmöglichkeiten von KI-Avataren

[00:11:47.700] - Steffen Wenzel

Ich glaube, die Akzeptanz hängt natürlich auch noch mal von dem Einsatzbereich ab. Also wo treffe ich auf einen Avatar? Was habe ich gerade für ein Anliegen? Kannst du uns da mal Beispiele bringen, wo Avatare heute schon eingesetzt werden und wo sie zukünftig eingesetzt werden?

[00:12:03.160] - Martin Wunderwald

Ja, also ich hatte es ja schon mehrfach genannt, also insbesondere in den Serviceschnittstellen, also wo es da wo aktuell letzten Endes Telefongespräche hinführen oder Chatverläufe, dass man dort die Informationsmenge im Hintergrund maximieren möchte, also einfach sehr viel Wissen vorhalten kann im Hintergrund. Egal ob das jetzt eine Krankenkasse, eine Versicherung ist, die über bestimmte Verträge, die man hat, über Fallkonstellationen, über Vertragstexte usw. einfach komplett informiert ist. Dieses Wissen, das in einer Person, in einer Künstlichen zu vereinen, ist einer der größten Punkte, warum man Avatare auch einsetzt. Und das heißt also immer, wenn es sozusagen so viel ist, dass ein einzelner Mensch nicht wissen kann oder sich aus vielen Systemen bedienen muss, um das Wissen zu generieren. Insbesondere im Callcenter wird man auf solche Avatare zurückgreifen. Das sind eben klassisch die ist eigentlich die Ablösung perspektivisch von Callcentern und vielleicht auch eine Bewegung weg vom Callcenter im Sinne von Telefon hin zu, dass dieser Kanal eben mit einem Avatar auf einem Mobiltelefon dauerhaft zur Verfügung steht. Man kann sich das so vorstellen Egal welche Applikationen, noch so komplex, Webseite App von dem Hersteller könnte perspektivisch durch einen Avatar unterstützt werden, der all das, was man eben über dieses Mensch Computer Interface, also Tastatur, Maus wie auch immer bedient, abgelöst wird und also durch natürliche Sprache und dann können im Hintergrund kann der Avatar eben Aktionen durchführen, Tickets erstellen, also Informationen geben. Ja, wie auch immer, die Anwendungsfälle sind dabei so breit und sehr branchenspezifisch, dass es schwer ist zu sagen, wo dieses Thema als erstes auf uns trifft. Einmal aus dieser gerade genannten Automatisierung, aus dem Automatisierungswunsch der Unternehmen getrieben, dass man einfach Personal abbauen will. In solchen Call Centern aber viel mehr auch noch getrieben oder mindestens so stark durch eine hohe Service Qualität, die man haben will beispielsweise. Ich bewege mich auf einem Bahnhof und möchte jetzt eben wissen, ob mein

Anschlusszug rechtzeitig ist und ob es vielleicht sogar der Zug ist, der hier gerade an dem ich vorbei laufe auf Gleis sieben, wo ein Zielort eingegeben steht. Und dann im besten Fall fotografiere ich diesen Zug und frage dann mein Avatar Ist das der Zug, der mich dorthin bringt, wo ich hin möchte?

[00:14:42.280] - Martin Wunderwald

Und kannst du mir ein Ticket bestellen? Und dann habe ich das eben in zwei Minuten gelöst und muss nicht erst zum Ticketautomaten, muss nicht zu an den Schalter gehen und mich bedienen lassen. Das sind dann solche Szenarien, die durchaus realistisch werden können.

Judy: Ein Beispiel für KI-Avatare

[00:14:56.240] - Steffen Wenzel

Da sind wir wieder bei deinem persönlichen Assistenten oder deiner persönlichen Assistentin, die du dir wünschst. Und. Aber es gibt ja auch noch andere Einsatzmöglichkeiten. Ihr habt beispielsweise Judy bei der MMS. Das ist auch ein Avatar, den ihr gerade jetzt bei der Digital X auch präsentiert habt. Was kann Judy und was macht sie so besonders?

[00:15:15.860] - Martin Wunderwald

Was kann Judy? Also Judy spezifisch ist ein 3D Avatar, da können wir dann noch mal tiefer abtauchen, wenn du möchtest in die technologischen Grundlagen von solchen Avataren, wenn es um die Gestaltung dieser geht. Also sie ist ein 3D Avatar, den wir individuell erstellt haben, nach einer Vorlage der Deutschen Telekom für spezifisch diese Messe für die Digital X. Sie läuft, sage ich mal oder läuft, laufen tut sie ,nicht. Sie wird gerendert in 3D, quasi auf einem lokalen Computer und steht quasi den Messebesuchern Frage und Antwort zu Wann sind welche Vorträge geplant? Wo ist mein Workshop zum Thema XY? Wo finde ich etwas? Also insbesondere auch wie komme ich von A nach B? Also eine klassische, eigentlich so eine Art Informationsperson. Ja, man kann es auch Wellcome Manager oder Concierge nennen. Da sind der Fantasie auch keine Grenzen gesetzt, wie man diese, diese virtuellen Charaktere dann auch benennt. Und tatsächlich stellt man also diese Frage in dem Fall in den Telefonhörer, typisch Telekom, hinein und bekommt dann ad hoc eine Antwort. Und diese Antwort wird quasi gespeist aus einem System mit eben all den Informationen zu dieser Messe.

[00:16:33.320] - Martin Wunderwald

Und in Echtzeit generiert. Ja, und über. Über die über den Avatar letzten Endes ausgegeben, der dann zu dem, was er was sie in dem Fall spricht. Die Lippen eben auch authentisch bewegt und damit ein sehr interessantes Gefühl auslöst Ja als, als Kunde oder als Besucher.

Content Creation durch KI-Avatare

[00:16:51.850] - Steffen Wenzel

Neben dem Servicebereich gibt es aber auch noch andere. Zum Beispiel auch den Content Creation Bereich, habe ich mir noch gemerkt. Also das heißt Avatare selbst Inhalte generieren. Was hältst du davon? Wie weit ist da die Technologie?

[00:17:06.610] - Martin Wunderwald

Wie meinst du das? Dass die Avatare selbst.

[00:17:09.790] - Steffen Wenzel

Ja, beispielsweise, dass du ja ein Avatar baust, der quasi Social Media Kanäle für dich bespielt und dementsprechend du da gar keine Redakteure oder Redakteurinnen mehr benötigst.

[00:17:23.350] - Martin Wunderwald

Ja, genau. Also hier ist es ja dann sozusagen nicht ein interaktiver Avatar, der quasi in so einer Realsituation Frage und Antwort steht, sondern eher vorproduziertes Videomaterial. Das heißt also, man nutzt generative KI Modelle, um einfach menschenähnliche oder tatsächlich auch oft Menschen zu generieren, die über bestimmte Sachen eben berichten. Ja, ich kenne es. Insbesondere in China ist das ein großes Thema in den sozialen Kanälen, dass also dort einfach im Akkord solche Bewertungen von Produkten beispielsweise durch so eine virtuelle Person durchgeführt werden und Rezensionen vorgelesen werden. Man sieht da auch viele Entwicklungen. Insbesondere bei YouTube ist es mir jetzt auch mal übergegangen, dass wenn man ein Video anschaut, was eigentlich natürlich sprachlich Englisch ist, es sofort simultan übersetzt wird auf die jeweilige Landessprache, in der ich mich befinde und sozusagen drübergelegt wird über den Content des Videos. Das heißt also, es ist inzwischen so weit, dass man nicht nur vorprozessieren kann, Sprache oder Simultanübersetzung, sondern auch die sozusagen in Echtzeit stattfinden lassen kann. Auch dafür hat die MMS im Übrigen ein Angebot, was transkribby AI heißt.

KI-Avatar Beratung von der Telekom MMS

[00:18:40.730] - Steffen Wenzel

Ja, kommen wir mal zu euch und euren Aktivitäten und Tätigkeiten. Ich habe ja eben gesagt, du berätst also Unternehmen dabei, wenn sie mit Avataren unter anderem arbeiten wollen. Welche Einsatzmöglichkeiten es dort gibt? Wie kann ich mir das vorstellen? Also kommt ein Unternehmen zu euch und sagt Wir würden da gerne was machen. Wir haben beispielsweise die und die Bereiche, wo wir uns das vorstellen. Wie gehst du dann vor?

[00:19:03.040] - Martin Wunderwald

Ja, also ähnlich wie du sagst. Nur meistens sind die Unternehmen natürlich über diesen physischen Kontakt zu dem Avatar gekommen. Also gerade die Digital X war so ein Anlaufpunkt. Wo also, wenn so ein Avatar in 3D hochauflösend in Körpergröße vor einem steht, mit dem man sprechen kann, schon ein Wow Effekt auslöst. Und insbesondere das so als kleiner Funfact, ist es eben eine Management Etage, die sich gern abbilden lässt in so einem virtuellen Raum. Wahrscheinlich auch in der Hoffnung, dass dann die Aufgaben gelöst werden durch den Avatar. Und so sieht man das, dass das oft eben Führungskräfte nachgebildet werden, die dann auf. Ja, wie auch immer, Hauptversammlungen, Tagungen als Avatar auftreten. Das haben wir in der Telekom einige Male habe ich das gesehen, aber auch bei einer Rewe Digital habe ich gesehen, dass da Personen sich in einem virtuellen Hologramm als Avatar nachbilden lassen. Also scheint doch ein starker Wunsch da zu sein, den Menschen nachzubilden und nicht nur quasi künstliche neue Menschen zu schaffen oder neue Personen.

[00:20:11.410] - Steffen Wenzel

Das ist ja spannend. Entspringt das einer Eitelkeit dieser Person oder ist das eher so, weil die Akzeptanz dann auch größer wird bei dem Rezipienten?

[00:20:21.010] - Martin Wunderwald

Eine gute Frage. Also ich bin eigentlich der Meinung, dass diese hyperrealistischen Avatare zwar diesen Wow-Effekt haben, dass sie aber in der breiten Masse nicht so eine gute Akzeptanz finden, weil sie natürlich insbesondere dann auch nur eine Person widerspiegeln. Und gerade wenn man jetzt über Neutralität der Geschlechter zum Beispiel nachdenkt, warum sollte eine Firma nur einen männlichen Avatar haben oder nur einen weiblichen? Warum sollte der schwarze Haare haben und weiße Haut? Ja, also all diese Fragen auch da die Neutralität zu wahren, drängt eigentlich mehr in die Richtung, einen, einen oder sogar mehrere Avatare zu erstellen, die also dort Gleichberechtigung aufgreifen und wegziehen, ein Stück weit, von den von den Abbildungen von realen Personen.

Avatar Anbieter

[00:21:12.710] - Steffen Wenzel

Vielleicht sollte man dann Tiere in der Zukunft nehmen, die sprechen können. Das ist dann vielleicht ein bisschen neutraler. Aber lass uns doch mal ein bisschen in die Technologie reinschauen und in die Anbieter, die es von Avataren jetzt mittlerweile gibt. Da gibt es ja eine Menge. Kannst du da ein paar nennen? Und kannst du uns auch mal sagen, wie die sich unterscheiden und wem du was dann auch empfehlen würdest?

[00:21:32.870] - Martin Wunderwald

Ja. Also es gibt eine Reihe von ja oder sagen wir drei, vier große Anwendungsgebiete. Das eine hat man schon gesagt, ist die Content Creation. Also wenn ich Videos generieren will mit künstlichen Avataren, die über Inhalte sprechen. Da gibt es einen großen Markt, auch eine Reihe von Anbietern. Das ist meistens quasi zweidimensional, das heißt fotorealistisch. Man lädt im besten Fall ein Bild von sich hoch, gibt den Text und es wird sozusagen ein Video generiert, was diesen Text spricht mit dem Avatar. Den kann man sich herunterladen, einbinden, veröffentlichen. Das ist so der eine Teil. Also ja, will sagen wie Sand am Meer Anbieter gibt. Einer davon ist beispielsweise HeyGen. Aber auch go Ava ist ein Unternehmen der Tech Start up Programms der Deutschen Telekom. Was in diese Richtung geht. Und dann gibt es eine weitere Gruppe der Avatare, die man immer unterschiedlich nennt. Meistens den Begriff interaktive Avatare habe ich gehört. Also wo es wirklich um die in Echtzeit Kommunikation mit einem physischen Gegenüber geht. Also wo wo quasi in Echtzeit gerendert werden muss oder oder die Bilder generiert werden müssen vom Avatar auf Basis dessen, was gerade gesprochen wird.

[00:22:48.980] - Martin Wunderwald

Und auch technologisch im Hintergrund ein Backend sein muss, was die Daten vorhält, die und das sind jetzt die wesentlichen Schritte von so einer interaktiven Avatar-Lösung. Das heißt also, es muss ja die Frage, mit der ich spreche, erst mal verstanden werden und in Text übersetzt. Ja, die wird sozusagen dann an ein Sprachmodell gestellt, was vorbereitet ist für diese Art Konversation. Also gepromptet ist ja da so ein bisschen der Fachbegriff und dann wird eine Antwort generiert vom Sprachmodell und dieses wird wieder in Sprache synthetisiert, welche dann wieder dazu führen muss, dass passend zu dem kleinen Audio Snippet der Mund und die Gestik des Avatar sich bewegt vom Avatar. Ja, und da gibt es eben so eine grobe Unterscheidung zwischen 2D und 3D, also Ansätzen, also dass man sagt, man baut quasi mit generativer KI eine 2D hyperrealistische Person, die aussieht wie ein Video. Oder man geht sogar in den 3D Raum und animiert eben Charaktere, wie wir sie aus Computerspielen beispielsweise kennen. Das sind dann die zugrundeliegenden 3D Engines wie Unreal und Unity zu nennen. Animiert in diesen virtuellen 3D räumen die Personen und die Gesichter, so wie eben der Text gesprochen wird und zeigt diese dann entsprechend dem Kunden, was dann einen größeren Freiheitsgrad hat, dass man diese Avatare eben auch interagieren lassen kann mit nicht physischen Objekten, mit Kaffeemaschinen oder wie auch immer die auch in diesem 3D Raum stehen können. Jetzt hatte ich gerade ein interessantes Thema von einem Museum, was also eine virtuelle Welt in der Vergangenheit nachbauen möchte in so einem Metaversum und dann eben auch möchte, dass diese Avatare durch diese 3D Welt in der Vergangenheit durchlaufen und dann aber auch sprechen können mit dem Besucher des Museums und auf Basis der Biografien der jeweiligen Avatare oder jeweiligen Personen, die sie dann in diesem 3D und dem historischen Raum darstellen, eben antworten. Und dann kann man eben mit jemandem aus der Vergangenheit sprechen, wenn man so will.

[00:25:04.250] - Steffen Wenzel

Ja. Okay, das klingt ja jetzt ziemlich komplex, aber vielleicht geht es ja auch ein bisschen schneller, als man da vielleicht vermuten kann. Was denkst du denn, wie lange braucht man dafür, um jetzt eine ganz simple Version erstmal dann auch starten zu können mit so einem Avatar-Generator? Wenn ich jetzt als Unternehmen zu euch komme und sage so, das hätten wir gerne jetzt für unseren Servicebereich oder was auch immer, hast du da ungefähr so eine Zeiteinschätzung?

[00:25:30.020] - Martin Wunderwald

Ja, also man kann sozusagen auch Avatare von der Stange nehmen, also sagen, ich nehme Avatare, die vorgefertigt sind und möchte die einsetzen. Mit dem kleinen Nachteil, dass die unter Umständen bei anderen Kunden eben auch wieder auftreten. Aber man kann auch sagen wer bringt eine Vorlage mit und wir gestalten diese Avatare individuell, was dann einfach so einem Projekt einfach noch mal einen Zeitraum von zwei Wochen verschieben lässt, bis man einen sehr guten, für seine Firma passenden Avatar erstellt hat, der alle Bedingungen erfüllt, wie man sich das vorstellt. Das ist insbesondere wichtig, wenn es größere Firmen sind, weil ja auch so ein Avatar sozusagen die Marke repräsentiert und ein Gesicht nach außen ist, was im besten Fall einheitlich ist und eben einigen Kriterien des Brands, wie man so schön sagt, entsprechen müsste. Und da lohnt es sich zu investieren und individuell etwas zu erstellen. Aber es geht auch von der Stange.

[00:26:28.950] - Steffen Wenzel

Und wahrscheinlich ist es ja wie immer, dass man damit anfängt und dann natürlich weiterentwickelt und dann natürlich auch wahrscheinlich ein bisschen progressivere Formen auch dann sich noch vorstellen kann, wenn man erst mal ein bisschen auf den Geschmack gekommen ist. So ist es.

Zukunftsansicht

[00:26:42.170] - Steffen Wenzel

Lass uns doch mal ein bisschen in die Zukunft blicken, Martin. Wohin geht die Reise? Das ist natürlich immer so zum Ende hin kommend bei einem Podcast eine typische Frage, aber gerade natürlich bei so einer Technologiefrage ganz wichtig. Wie schätzt du dort die weitere Entwicklung ein?

[00:26:58.700] - Martin Wunderwald

Ja, also ich denke, dass diese Anfangshürden bald genommen werden und wir auch diese den Hyperrealismus bei Avataren sehen werden. Also insbesondere hat Microsoft da ja schon vor einiger Zeit fertiges Modell mit Vasa-1 vorgelegt, als Paper, es aber bewusst noch nicht implementiert. Ja, eben aus dieser Deepfake Assoziation und den Gefahren, die sich damit ergeben. Aber ich denke, das wird was sein, was bald kommt, dass also diese Hyperrealismus erreicht wird und gleichzeitig einfach die Kosten für die Erstellung eines solchen Avatars einfach deutlich geringer werden. Das ist so das eine, was ich denke und ich habe auch noch eine große steile Hypothese, dass die, dass diese Assistenten, wie wir sie kennen, Siri, Alexa, dass auch die von den insbesondere von den Herstellern der Endgeräte, also den großen Apple, Google usw. letzten Endes ein Stück weit kommen werden und vielleicht auch ein Aussehen bekommen werden und dieser Sprachkanal dann eben zentral über das Telefon gesteuert wird. Und auch die Ein- und Ausgabe und die Übersetzung in Sprache oder in Text. Und letzten Endes nicht mehr in den jeweiligen Applikationen diese Avatare stattfinden müssen, sondern dass die sozusagen zentraler agieren werden als wirklich ein Assistent für mein tägliches Leben. Ja, der aber dann viel mehr Aufgaben übernehmen kann und sozusagen einzelne Apps jetzt zum Beispiel dann eben die Schnittstellen bieten müssen, um eine überliegende KI sage ich jetzt mal oder so ein Avatar Assistenten sozusagen zu bedienen, Ja.

[00:28:42.330] - Steffen Wenzel

Und das, was wir bei Star Wars schon erlebt haben mit dem Hologramm. Also, dass ich dann auf einen Knopf drücke an meinem Telefon und das dann quasi Siri, Alexa oder wer auch immer dann erscheint, wird das auch passieren?

[00:28:55.830] - Martin Wunderwald

Tja. Macht es das einfacher? Das ist die Frage. Ich glaube, wir werden noch viel sehen. Wenn wir jetzt gerade die Bewegung bei den Hologrammen und bei den AR/VR Brillen sehen. Ich glaube, die werden schnell kleiner werden. Ja, und vielleicht gerade weil wir beide auch Brillenträger sind, sieht man irgendwann gar nicht mehr, dass eigentlich dieses Glas eben auch unser Bildschirm ist in so eine digitale

Welt. Und wir brauchen dann keine Handys mehr, sondern wir können anders über Sprache oder Fingergesten steuern. Also ich glaube, das sind der Fantasie keine Grenzen gesetzt. Ich bin gespannt, was sich durchsetzt. Noch glaube ich fest daran, dass die jahrelange Forschung der Mensch-Computer-Interaktion noch nachhält, obwohl man sieht, dass gerade die Jugend natürlich, also ich meine WhatsApp ist so ein schönes Beispiel, dass man dass es eben Leute gibt, die Sprachnachrichten schicken und andere Telefonieren lieber. Und es gibt welche, die tippen lieber und nutzen ganz viele Emoticons. Also das heißt, jeder hat da seine Präferenzen und es wird einfach ein weiterer Kanal sein.

Persönlicher Einblick von Martin Wunderwald

[00:29:55.800] - Steffen Wenzel

Du hast es eben auch schon angesprochen, dass es natürlich manchmal auch darum geht, die Geschichte vielleicht noch mal lebendig zu machen mit Avataren. Deswegen meine Abschlussfrage an dich: Wenn du dir eine historische Persönlichkeit aussuchen könntest, die du als Avatar zum Leben erwecken könntest, wer wäre das und warum?

[00:30:16.380] - Martin Wunderwald

Das ist wirklich eine gute Frage und ich mag es auch total, da in den Vergangenheitswelten zu schwelgen. Sei es jetzt irgendwie eine Game of Thrones Serie, in die man eintaucht oder was anderes. Aber ich würde tatsächlich mal mit so einem Ritter aus dem Mittelalter reden, mit so Ritter Lancelot oder König Arthur. Und mal fragen: Wie war es denn wirklich? Ja, genau. Also das würde mich schon interessieren.

[00:30:45.210] - Steffen Wenzel

Hoffentlich sagt er dann nicht, dass es gar nicht so spannend war. Vielen Dank, Martin, dass du heute hier zu Gast warst. Ich fand das sehr spannend, dass du uns einen Einblick gegeben hast in die Welt der KI Avatare. Und hoffentlich wiederholen wir das bald, damit wir dann auch wieder in der Zukunft von dir hören, was alles so passiert ist und passieren wird. Danke schön. Martin.

[00:31:04.470] - Martin Wunderwald

Ich danke. Tschüss.

[00:31:06.510] - Steffen Wenzel

Ja. Herzlichen Dank, dass Sie heute bei uns hier zu Gast waren in diesem Podcast. Natürlich auch an Sie. Wir wünschen Ihnen ein frohes Fest. Das ist der letzte Podcast in diesem Jahr und ich kann Ihnen versprechen, im neuen Jahr geht es weiter und noch nicht mit einem KI-Avatar, sondern wieder mit mir und darüber hinaus finden Sie alle weiteren Infos zu diesem spannenden Thema wie immer in den Shownotes. Bis dahin alles Gute