Bürgerservice

Der Avatar als Bürgerberater

[04.11.2026] Wie muss ein Avatar gestaltet sein, um vom menschlichen Gegenüber positiv wahrgenommen zu werden? Mit dieser Frage beschäftigt sich die Frankfurt University of Applied Sciences. Zusammen mit einem Start-up werden zudem fotorealistische digitale Avatare entwickelt, die künftig etwa als Bürgerberater zum Einsatz kommen könnten.

Avatar ist nicht gleich Avatar. Vielmehr existiert eine Vielzahl von Erscheinungsformen – gemein ist ihnen allen jedoch, dass sie beim Betrachter eine Wirkung entfalten. Im Hollywood-Blockbuster „Avatar: Fire and Ash" erweckt die Schauspielerin Oona Chaplin die Anführerin des Mangkwan-Clans, Varang, durch Sprache, Mimik und Gestik zum Leben. Diese gelungene Verkörperung gibt zugleich Aufschluss darüber, wie Avatare konzipiert sein sollten: Eine unzureichende Umsetzung würde die Qualität der Darstellung spürbar mindern und den intendierten Effekt verfehlen. Es stellt sich die Frage, ob sich der Zusammenhang zwischen Gestaltungsqualität und Wirkung auch auf Avatare übertragen lässt, die in der öffentlichen Verwaltung zum Einsatz kommen.

Um diese Frage näher zu beleuchten, hat sich an der Frankfurt University of Applied Sciences (Frankfurt UAS) eine Forschungsgruppe etabliert, die sich der Gestaltung neuerer Konzepte der Human-Computer-Interaction in verschiedenen Bereichen widmet. Im Mittelpunkt stehen dabei sowohl die Entwicklung geeigneter Chatbots mittels Large Language Models (LLMs) als auch die Gestaltung von Avataren.

Der Begriff Avatar stammt ursprünglich aus dem Sanskrit und bedeutet die irdische Verkörperung einer Gottheit. Im Roman Snow Crash aus dem Jahr 1992 taucht der Begriff erstmals im digitalen Kontext auf und bezeichnet die grafische Repräsentation einer Person im virtuellen Raum, dem Metaverse – das kommt der heutigen Bedeutung, einer digitalen Stellvertreterfigur eines realen Nutzers, bereits sehr nahe. Im Gegensatz zum Avatar-Konzept der gleichnamigen Hollywood-Filmreihe, in denen die Na'vi-Figuren von Menschen gesteuert werden, handelt es sich bei den Avataren, die in der öffentlichen Verwaltung eingesetzt werden, um digitale – teilweise durch KI-Programme – gesteuerte Modelle. Das ist ein wesentlicher Unterschied, der sich beim derzeitigen State-of-the-Art negativ auswirkt, da Gestik und Mimik eines professionellen Schauspielers nur bedingt KI-generiert werden können.

Acht Grade des Realismus

Betrachtet man ausschließlich zweidimensionale Avatare in Verbindung mit Chatbots, lassen sich acht Kategorien unterscheiden, die sich in ihrem Grad an Menschenähnlichkeit und Realismus deutlich voneinander abgrenzen. Am einen Ende des Spektrums steht der symbolische Avatar: Er verzichtet weitgehend auf menschliche Merkmale und tritt stattdessen als Symbol, Icon oder abstrakte Figur in Erscheinung – etwa als Roboterkopf oder abstraktes Logo. Eine Stufe menschenähnlicher ist der Cartoon- oder Comic-Avatar, eine stilisierte menschliche oder tierische Figur, die bewusst unrealistisch gehalten ist, wie man es von klassischen Comicfiguren kennt.

Der anthropomorphe Avatar wiederum greift zwar eine menschliche Gestalt auf, bleibt dabei aber deutlich stilisiert, etwa in Form eines gezeichneten menschlichen Gesichts. Beim illustrativen Avatar steigt die gestalterische Qualität weiter: Er ist hochwertig gezeichnet oder gerendert, verzichtet jedoch nach wie vor

auf Fotorealismus, wie es für digitale Illustrationen typisch ist. Mit dem semirealistischen Avatar nähert sich die Darstellung dem menschlichen Vorbild weiter an – er verfügt über menschliche Proportionen und Gesichtszüge, bleibt jedoch erkennbar künstlich, etwa als digitaler „Porträt"-Avatar.

Der fotorealistische 2D-Avatar strebt bereits eine möglichst realistische Darstellung eines menschlichen Gesichts an, sei es als Foto oder als KI-generiertes Abbild. Die siebte Kategorie bildet dann der digitale Mensch oder Talking Head, der ein reales oder realitätsnahes menschliches Gesicht in Bewegung zeigt und dabei Sprache mit synchronisierter Mimik verbindet. Die achte und komplexeste Kategorie stellt schließlich der affektiv reagierende Avatar dar, der nicht nur optisch möglichst realistisch erscheint, sondern zusätzlich in der Lage ist, auf die Emotionen und Eingaben des Nutzers mit entsprechender Mimik und Reaktion einzugehen.

Wann wird ein Avatar positiv wahrgenommen?

In einer aktuellen empirischen Studie wurden vier Erscheinungsformen von Avataren und deren Einfluss auf den ersten Eindruck von Customer-Service-Chatbots untersucht – menschliche (also realistische oder fotorealistische) Avatare, karikierte Avatare (Comic- oder cartoonhafte Figuren), zoomorfe (tiergestaltige) Avatare sowie funktionale/symbolische Avatare. Die Studie gelangte zu dem Ergebnis, dass Anthropomorphismus für die positive Wahrnehmung eines Chatbots wichtiger zu sein scheint als visueller Realismus. Einfach ausgedrückt: Je mehr menschliche Eigenschaften ein Objekt besitzt oder zugeschrieben bekommt, desto positiver wird es wahrgenommen.

Für die Gestaltung von Avataren bedeutet das, das mindestens drei getrennte Dimensionen zu beachten sind: Wie realistisch sollte der Avatar aussehen (Virtual Realism)? Besitzt oder vermittelt er menschliche Eigenschaften (Anthropomorphism)? Wie menschlich verhält er sich und kommuniziert er (Behavioral Realism)?

Eine im Jahr 2025 veröffentlichte systematische Meta-Analyse teilt Avatare anhand ihres Realismusgrades in drei Klassen ein: den High Realism Avatar (HRA) mit hoher menschlicher Detailtreue, den Medium Realism Avatar (MRA) mit vereinfachtem, stilisiertem menschlichem Erscheinungsbild sowie den Low Realism Avatar (LRA), der Cartoon-, Sketch- oder Comic-Stil aufweist. Die Untersuchung wertete insgesamt 13 Studien mit knapp 2.300 Teilnehmenden aus. Der Studie zufolge wird ein Avatar umso attraktiver und vertrauenswürdiger wahrgenommen, je realistischer er gestaltet ist.

Hochschule untersucht drei Analyseebenen

An der Frankfurt UAS werden Avatare und deren Wirkung auf uns Menschen nun anhand unterschiedlicher Analyseebenen betrachtet. Erstens die kognitivistische Ebene im Sinne der Willing Suspension of Disbelief. Das Konzept geht auf Samuel Taylor Coleridge (1817) zurück und beschreibt die Bereitschaft, eine künstliche oder fiktionale Darstellung als glaubwürdig zu akzeptieren. Für digitale Medien bedeutet das, dass Nutzer einem künstlichen Gegenüber soziale Eigenschaften zuschreiben können, wenn dessen Darstellung und Verhalten hinreichend kohärent erscheinen.

Eine frühe und besonders eindrucksvolle Ausprägung dieses Phänomens ist der Eliza-Effekt. Das von Joseph Weizenbaum 1966 entwickelte Computerprogramm Eliza simulierte einen psychotherapeutischen Dialog, obwohl es lediglich auf der Erkennung bestimmter Schlüsselwörter und vorgegebenen Antwortmustern beruhte. Dennoch schrieben Nutzer dem System teilweise Verständnis und Empathie zu. Der Eliza-Effekt als wahrnehmungspsychologische Ebene zeigt, dass bereits einfache sprachliche

Interaktion ausreichen kann, um einem technischen System menschliche Eigenschaften zuzuschreiben.

Demgegenüber beschreibt die von dem japanischen Robotiker Masahiro Mori 1970 erstmals beschriebene Theorie des Uncanny Valley (Unheimliches Tal) die Grenze dieser Akzeptanz: Mit zunehmender menschlicher Ähnlichkeit steigt die Akzeptanz zunächst an, kann bei einer nahezu, aber nicht vollständig menschlichen Darstellung jedoch abrupt in Irritation und Ablehnung umschlagen, etwa, wenn das System eine unpassende emotionale Antwort gibt, starr starrt oder logisch versagt.

Eine dritte Ebene betrifft die physiologischen Reaktionen auf wahrgenommene soziale Signale. Die 1994 von Stephen Porges veröffentlichte Polyvagal-Theorie bietet hierfür einen theoretischen Bezugsrahmen, indem sie die Bedeutung von Signalen wie Stimme, Mimik und Körperhaltung für die Wahrnehmung von Sicherheit und die Regulation des autonomen Nervensystems hervorhebt.

Fotorealistischer Avatar für vertrauensvolle Interaktion

Aktuell werden an der Frankfurt University of Applied Sciences in Zusammenarbeit mit einem Start-up fotorealistische digitale Avatare entwickelt und über entsprechende Workflows mit Chatbots und generativen KI-Systemen verbunden. Neben Bereichen aus der Medizin und der Sicherheit werden auch Themen aus dem Bereich der öffentlichen Verwaltung bearbeitet. Ein Beispiel hierfür ist die Bürgerberaterin. Diese kann Anträge, etwa zur Kfz-Zulassung, schon vorab mit den Bürgerinnen und Bürgern besprechen und auf Vollständigkeit prüfen oder einfache Hinweise zu Fachfragen geben. Für die Verwaltung ergibt sich dadurch eine Effizienzsteigerung als positiver Effekt.

Im Gegensatz zu anderen Chatbots mit Avataren werden diejenigen der Frankfurt UAS mithilfe des Framework MetaHuman von Epic Games auf Basis der Unreal Engine programmiert. Bei einem MetaHuman besitzt das Gesicht ein sehr komplexes Facial Rig – mit solch einem technischen Steuerungssystem in einem 3D-Modell können beispielsweise verschiedene Gesichtspartien bewegt werden, um so Gesichtsausdrücke, Mundbewegungen und Emotionen zu animieren. Für eine natürliche Sprachinteraktion ist insbesondere die Lippenbewegung (Lip Sync) wichtig. Damit soll der Uncanny-Valley-Effekt vermieden und die Interaktion mit dem Avatar als vertrauensvoller wahrgenommen werden.

Die Entwicklung von fotorealistischen Avataren für die digitale Verwaltung steht noch am Anfang. Gemeinsam mit mehreren Kommunen werden entsprechende Konzepte derzeit praktisch erprobt, weiterentwickelt und hinsichtlich ihrer Einsatzmöglichkeiten untersucht. Dabei geht es nicht allein um die technische Integration von Avataren in Chatbots, sondern zunehmend um die Frage, wie soziale Präsenz, Vertrauen und Nutzerakzeptanz in der digitalen Interaktion gezielt gestaltet werden können. Mit der Weiterentwicklung generativer KI eröffnen sich dabei neue Möglichkeiten, Avatare nicht nur sprachlich, sondern auch durch Mimik, Gestik, Stimme und situationsabhängige Reaktionen an die jeweilige Interaktion anzupassen.

Die nächste Entwicklungsstufe im Blick

Damit rückt eine weitere Entwicklungsstufe in den Mittelpunkt: der affektive Avatar. Während herkömmliche Avatare in erster Linie Informationen vermitteln und ein möglichst menschliches Erscheinungsbild erzeugen, sollen affektive Avatare emotionale Zustände erkennen, berücksichtigen und angemessen darauf reagieren. Sie könnten beispielsweise Unsicherheit, Frustration oder Zufriedenheit des Nutzers aus dessen Sprache, Stimme oder Interaktionsverhalten ableiten und ihre Kommunikation entsprechend anpassen. Damit verschiebt sich der Schwerpunkt von der bloßen Vermenschlichung der

Darstellung hin zu einer emotional adaptiven Mensch-Maschine-Interaktion.

()

Stichwörter: Künstliche Intelligenz, Avatare, Bürgerservice, Frankfurt University of Applied Sciences