

## Künstliche Intelligenz und die Zukunft virtueller Existenzen

### Gesellschaftliche und ethische Perspektiven

*Was bedeutet es, wenn Verstorbene durch Künstliche Intelligenz (KI) für ihre Hinterbliebenen virtuell präsent bleiben? Matthias Meitzler hat in einer qualitativen Studie zu einer Veränderung der Sterbe- und Trauerkultur durch solche Angebote und deren möglichen gesellschaftlichen Folgen geforscht – ein Thema, das auch die Trauerpastoral zutiefst betrifft.*

#### KI-basierte Formen des digitalen Weiterlebens

Digitale Technologien prägen zunehmend das gesellschaftliche Alltagsleben – und auch für den Umgang mit dem Sterben, mit Tod und Trauer spielen sie eine immer größere Rolle. Beispielsweise können Verstorbene schon seit längerer Zeit auf virtuellen Friedhöfen, Gedenkseiten oder zurückgelassenen Social-Media-Profilen digital betrauert und erinnert werden. Durch die jüngsten Entwicklungssprünge im Bereich der Künstlichen Intelligenz haben die Möglichkeiten der postmortalen Online-Präsenz nun eine weitere Dimension erreicht. Auf einem unter dem Label der *Digital Afterlife Industry* (DAI) firmierenden Markt werden KI-Systeme angeboten, die das Kommunikationsverhalten Verstorbener – zum Teil mitsamt ihrer früheren optischen Erscheinung und dem (vermeintlich) unverwechselbaren Klang ihrer Stimme – nachahmen und mit den hinterbliebenen Nutzer:innen interagieren können.

Diesbezüglich gibt es verschiedene Anwendungsformen. Während Chatbots darauf ausgelegt sind, textbasierte Konversationen zu führen, äußern sich Avatare darüber hinaus auch optisch in Form eines virtuellen Körpers. In beiden Fällen wird auf größere Datenmengen (z. B. E-Mails, Messenger-Nachrichten, Postings, Foto-, Video- oder Audioaufnahmen) zugegriffen, die das analoge Original zu Lebzeiten produziert hat. Wie genau diese Daten verarbeitet werden, hängt wiederum vom konkreten Programm ab. Vergleichsweise einfache Applikationen bedienen sich zwar selektiv aus dem Ursprungsmaterial, dieses wird den Nutzenden jedoch unverändert ausgegeben. So sendet z. B. eine Smartphone-App Hinterbliebenen zu bestimmten Anlässen Text-, Sprach- oder Videobotschaften der Verstorbene. In Großbritannien wurde im Jahr 2022 während einer Trauerfeier die digitale Repräsentation einer kürzlich verstorbenen älteren Dame auf einem großen Bildschirm eingeblendet und stand den versammelten Gästen Rede und Antwort. Hierzu wurden einzelne Sequenzen aus längeren Interviews, welche die Frau kurze Zeit vor ihrem Tod gegeben hatte, von einer Software passend zu der jeweils gestellten Frage abgespielt. Überdies existieren auch solche KI-Anwendungen, die die gespeicherten Informationen nicht lediglich unverändert abrufen, sondern daraus neuen Output generieren. Nutzer:innen haben dann die Möglichkeit, in einen Dialog mit dem digitalen ‚Double‘ zu treten, ihm entweder vor einem Display oder etwa unter Verwendung einer Datenbrille in einer dreidimensionalen virtuellen Realität zu ‚begegnen‘. Aus Südkorea stammt das Beispiel einer TV-Show, in der die Zusammenkunft einer Frau mit der 3D-Simulation ihrer verstorbenen siebenjährigen Tochter gezeigt wurde. Bei den technisch erzeugten Kommunikationsinhalten handelt es sich folglich um Sätze, die von der repräsentierten Person zwar nie so gesagt wurden, die sie aber auf diese oder ähnliche Weise sagen *könnte*, wäre sie noch am Leben.

Ein so konzipiertes digitales Weiterleben umfasst sowohl jene Konstellationen, in denen Individuen zu Lebzeiten eine spezifische Verwendung ihrer Daten für die Zeit nach dem eigenen Tod arrangieren (oder explizit untersagen), als auch die Perspektive der Hinterbliebenen, die daran interessiert sind, die Präsenz von Verstorbene in dieser Form zu bewahren bzw. zu erfahren. All das ist nicht allein für private Kontexte relevant, etwa beim Tod von nahestehenden Familienangehörigen, sondern u. a. auch innerhalb der Populärkultur, wenn verstorbene Schauspieler:innen oder Musiker:innen ihr postmortales Comeback als Avatar feiern, oder in der historisch-politischen Bildung, vor allem bei der digitalen Repräsentation von Zeitzeug:innen, mit denen z. B. Museumsbesucher:innen interagieren können. Auch wenn die meisten Dienste der DAI noch nicht sehr weit verbreitet sind und die Produktion realistischer Avatare von Verstorbene einen größeren technischen wie finanziellen Aufwand bedeutet, ist künftig von einer vermehrten Nutzung auszugehen.

So sehr sie die kollektive Fantasie von der Überwindung der Endlichkeit bedienen, bringen die genannten Angebote einige noch weitgehend ungeklärte Probleme mit sich und werfen verschiedene ethische, rechtliche und sicherheitstechnische Fragen auf. Diese standen im Mittelpunkt eines 2022 bis 2024 an der Universität Tübingen in Kooperation mit dem Fraunhofer SIT (Darmstadt) durchgeführten und vom BMBF geförderten



Matthias Meitzler ist wissenschaftlicher Mitarbeiter am Internationalen Zentrum für Ethik in den Wissenschaften (IZEW) an der Eberhard-Karls-Universität Tübingen. Als Soziologe befasst er sich seit längerer Zeit mit unterschiedlichen gesellschaftlichen Phänomenen rund um Sterben, Tod und Trauer.

**Forschungsprojektes.** Was bedeuten die aktuellen und künftigen Möglichkeiten der virtuellen Existenz nach dem Tod für das gesellschaftliche Zusammenleben im Allgemeinen und für künftige Trauer- und Erinnerungskulturen im Besonderen? Kann der Avatar eines/ einer geliebten Verstorbenen die persönliche Verlustbewältigung bzw. den Blick auf den eigenen Tod erleichtern – oder wird dies gerade hierdurch zusätzlich erschwert? Wer trifft die finale Auswahl über die zu verwendenden Daten und wem obliegt die Gestaltungshoheit über die digitale Identität nach dem Lebensende? Welche Wünsche und Sorgen haben Menschen diesbezüglich? Welche Risiken der Manipulation bzw. des Missbrauchs bestehen und wie können die Rechte Betroffener durchgesetzt werden? Um angesichts dieser und weiterer Forschungsfragen tiefere Einsichten in die jeweiligen Zusammenhänge zu gewinnen, bedurfte es der interdisziplinären Kooperation von Rechtswissenschaften, Informatik, Theologie, Soziologie, Medienwissenschaften und Ethik.

Das Projekt verfolgte einen partizipativen Ansatz im Lichte der qualitativen Sozialforschung. Mittels Analysen bestehender DAI-Dienste und unter Einbeziehung der Expertisen, Erfahrungen und Einstellungen unterschiedlicher Akteure wurde eine Kartierung und erste Exploration des Forschungsfeldes vorgenommen. Dabei kamen sowohl Privatpersonen als auch Stakeholder aus unterschiedlichen Tätigkeitsfeldern insbesondere der Sterbe- und Trauerkultur (z. B. Bestattungswesen, konfessionelle Seelsorge, Sterbe- und Trauerbegleitung) im Rahmen von Fokusgruppendifkussionen und Einzelinterviews zu Wort. Auf diese Weise konnten verschiedene gesellschaftliche Perspektiven auf das Thema erschlossen und ein umfassendes Verständnis seiner vielfältigen Aspekte und Herausforderungen gewonnen werden. Unter der Annahme, dass es sich dabei um einen Fiktions- und Projektionsraum handelt, dessen weitere Entfaltung noch nicht klar abzusehen ist, wurden außerdem verschiedene fiktive Narrationen betrachtet, die das digitale Weiterleben behandeln. So findet sich in entsprechenden Filmen oder Serien üblicherweise die dominante Vorstellung, dass die Avatare der Toten nicht nur suggestiv, sondern auch manipulativ auf das Leben der Hinterbliebenen einwirken und dabei ein kaum mehr zu kontrollierendes ‚Eigenleben‘ entfalten.

Die Forschungsarbeit knüpft an vorausgegangene Untersuchungen zum Wandel der Sterbe-, Trauer- und Erinnerungskultur sowie zu Fragen der Datensicherheit, des Datenschutzes, der informationellen Selbstbestimmung und des digitalen Nachlasses an. Sie zielt darauf ab, die aktuellen Angebote der DAI zu reflektieren, zukünftige Entwicklungen zu antizipieren, Regulierungsbedarf zu kennzeichnen und konkrete Handlungsoptionen für Politik und Gesellschaft zu erarbeiten. Ein zentrales Anliegen besteht folglich darin, für die Thematik der digitalen Fortexistenz mittels KI innerhalb einer breiteren Öffentlichkeit zu sensibilisieren und einen empirisch informierten sowie ethisch begründeten Impuls für die gesellschaftliche Diskussion über geeignete Rahmenbedingungen zu geben.

#### **Leitlinien für die Entwicklung und Nutzung KI-basierter Avatare von Verstorbenen**

Nachfolgend sollen einige Handlungsoptionen benannt werden, die sich aus den Ergebnissen der Studie ableiten lassen. Von herausragender Bedeutung ist dabei der Schutz von Usern sowie die Wahrung der Rechte von Verstorbenen und Angehörigen. Um potenzielle Nutzungsrisiken zu minimieren, ist ein kritisches Technikverständnis der Bevölkerung zu fördern. Hierfür braucht es zuverlässige Anlaufstellen und einen niedrigschwelligen Zugang zu umfangreichen zielgruppenorientierten Informationen. In diesem Zusammenhang erweist sich *digital literacy* als ein Schlüsselbegriff, der sich nicht nur auf das Wissen über technische Funktionsweisen und konkrete Handhabungen bezieht, sondern auch und vor allem auf die Fähigkeit, betreffende Anwendungen hinsichtlich ihrer Auswirkungen auf persönlicher, gesellschaftlicher und ethischer Ebene zu reflektieren und verantwortungsbewusst mit ihnen umzugehen. *Digital literacy* ist indes nicht allein aufseiten der Nutzenden gefragt; auch weitere Akteure werden sich in Zukunft mit dem Thema des digitalen Weiterlebens näher befassen müssen, um zuverlässige Empfehlungen aussprechen, beratend tätig sein und generelle Orientierung geben zu können. Dazu zählen vor allem solche Personen, die von Berufs wegen mit Fragen rund um das Lebensende befasst sind und an die sich potenzielle Nutzer:innen solcher digitalen Angebote wenden könnten. Aber auch deren Entwickler:innen sind angesprochen, denn sie stehen vor der Herausforderung, Produkte zu kreieren, die einerseits kommerziell erfolgreich sind und andererseits ethischen Standards gerecht werden. Hierfür bedarf es nicht zuletzt eines umfassenden Wissens über Trauerprozesse und -bedürfnisse sowie über deren jeweilige psychologische, soziale, kultur- bzw. religionspezifische Implikationen.

Durch die weiterhin zu erwartenden Fortschritte im Bereich der KI dürfte es künftig noch schwieriger werden, zwischen einer echten Person und ihrer digitalen Simulation zu unterscheiden. Weil dieser bereits für den allgemeinen Umgang mit sogenannten *deepfakes* bekannte Umstand gleichermaßen auf Avatare von Verstorbenen zutrifft, wären auch hier Kennzeichnungspflichten für KI-generierte Inhalte relevant, wie sie bereits in der [EU-Verordnung](#) „zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz“ fixiert sind. Nicht minder bedeutsam ist die Notwendigkeit der Erfüllung datenschutzrechtlicher Pflichten, die auch die Daten von Verstorbenen mit einschließen. Ebenso ist der postmortale Persönlichkeitsschutz im Hinblick auf Anwendungen der DAI noch lückenhaft und bedarf einiger juristischer Nachjustierungen. Dazu gehören etwa konkrete Nachlassregelungen der repräsentierten Person bezüglich des Umgangs mit ihren Daten nach dem Tod. Wenn Menschen eine solche postmortale Nutzung ihrer persönlichen Daten ablehnen, dann benötigen sie Möglichkeiten, diesen Willen rechtlich verbindlich kundzutun. Ebenso könnte

es hilfreich sein, spezifische Vorkehrungen zu treffen, um zu vermeiden, dass bestimmte Hinterbliebene unerwartet und unfreiwillig mit dem Avatar eines verstorbenen Familienmitglieds konfrontiert werden. Auch aus technischer Sicht braucht es klare Richtlinien für die Erstellung, Verwendung und Löschung von Avataren. Dies umfasst u. a. Maßnahmen hinsichtlich der zu gewährenden bzw. einzuschränkenden Autonomie solcher KI-Systeme und des Personenkreises, mit dem sie interagieren dürfen. Eine häufig angesprochene Befürchtung besteht darin, dass der Avatar Dinge mitteilen könnte, die seine Interaktionspartner:innen irritieren oder gar beleidigen. Solchen ungewollten Output gilt es durch hohe sicherheitstechnische Standards zu verhindern.

Im Unterschied zu anderen Formen des medialen Totengedenkens wie beispielsweise Fotos und Videos entfalten Avatare von Verstorbenen durch die Ausgabe neuer KI-generierter Kommunikationsinhalte eine noch größere Suggestivkraft. Damit können sie einen nicht zu unterschätzenden Einfluss auf das emotionale Wohlbefinden von Hinterbliebenen nehmen. In diesem Zusammenhang ist insbesondere zu bedenken, dass es sich bei den Diensteanbietern um kommerzielle Unternehmen handelt, deren Interessen sich nicht zwangsläufig mit den Trauerbedürfnissen der Nutzenden decken müssen. In einem derart sensiblen Bereich wie dem Lebensende erscheinen beispielsweise Produktplatzierungen oder andere Werbekommunikationen, die die meist vulnerable Trauersituation ausnutzen, höchst problematisch.

Eine weitere ethische Überlegung betrifft die Angemessenheit der Darstellung eines Avatars. Seine Entstehung basiert – wie bei allen anderen medialen Artefakten – auf einer spezifischen Auswahl, Kombination und Ausblendung bestimmter Elemente. Die zugrundeliegende Datenbasis repräsentiert somit lediglich einen Ausschnitt aus der früheren Lebenswelt des/der Verstorbenen. Andere Persönlichkeitseigenschaften und Kommunikationsweisen erhalten demgegenüber keine Berücksichtigung in den Trainingsdaten und schlagen sich darum auch nicht im Auftreten des Avatars nieder. Vor diesem Hintergrund ließe sich über die Freiheitsgrade bei der Gestaltung solcher digitaler Repräsentationen diskutieren: Sollten ihre Auftraggeber:innen – seien es die Verstorbenen zu Lebzeiten oder deren Hinterbliebene – diesbezüglich völlige Entscheidungshoheit genießen und Avatare über sämtliche Themen sprechen und dabei sämtliche Meinungen vertreten dürfen? Oder besteht auch hier die Notwendigkeit von *content moderation*, indem bestimmte als bedenklich empfundene (z. B. rassistische, sexistische oder sonst wie diskriminierende bzw. demokratiefeindliche) Inhalte *per se* ausgeschlossen werden? Und soll das digitale Weiterleben – ähnlich wie bei Social Media – einen Raum zur *Selbstoptimierung* bereitstellen? Unliebsame Eigenschaften (z. B. in körperlicher Hinsicht oder auch bestimmte Persönlichkeitsfacetten betreffend) ließen sich dann mithilfe der technischen Möglichkeiten digital ‚retuschieren‘ bzw. von vornherein ausblenden. Doch wer würde hierbei die finale Entscheidung treffen? Die verstorbene Person, die frühzeitig entsprechende Festlegungen getroffen hat? Ihre Angehörigen, die den Dienst nutzen und dabei eigene Bedürfnisse entwickeln? Oder letztendlich das Unternehmen, welches das Tool bereitstellt und damit ein ökonomisches Kalkül verfolgt?

#### **Gegenwärtige und zukünftige Relevanzen für Politik und Gesellschaft**

Die in der Studie befragten Teilnehmenden beurteilen das digitale Weiterleben durch KI überwiegend skeptisch. Besonders häufig wird dabei die Sorge geäußert, dass das simulierte Zwiegespräch mit einem Computerprogramm einen negativen Einfluss auf den Trauerprozess haben könnte. Gelungene Verlustbewältigung, so die dominante Auffassung, setze voraus, dass man das Nicht-mehr-da-Sein des geliebten Menschen realisiere und akzeptiere – und sei dies noch so schmerzhaft. Eine KI-Anwendung, die den Eindruck erwecke, die verstorbene Person sei gar nicht tot, sondern weiterhin verfügbar, würde die notwendige Akzeptanz allerdings erheblich erschweren bzw. verunmöglichen. Trauernde könnten hierdurch in einer Art ‚Scheinwelt‘ versinken, die das Risiko der Vereinsamung erhöhe und ein Suchtpotenzial besäße, weshalb man nur noch mit großen Mühen aus ihr herausfinden könne.

Wenngleich die vorgebrachten Einwände durchaus ihre Berechtigung haben und für die weitere Reflexion berücksichtigt wurden, mangelt es angesichts der relativ geringen Verbreitung von DAI-Anwendungen aktuell noch an belastbaren empirischen Erkenntnissen über die tatsächlichen trauerpsychologischen Wirkungen. Um dies näher einordnen zu können, müsste u. a. zwischen unterschiedlichen Anwendungsszenarien und damit verbundenen Parametern differenziert werden. Dazu gehören etwa das erreichte Lebensalter der Verstorbenen, die Todesumstände (plötzlicher Herzstillstand, Unfall, Suizid oder ein vorausgegangener längerer Krankheitsverlauf), die empfundene Beziehungsqualität, der Zeitpunkt und die Dauer der Inanspruchnahme des Dienstes sowie die konkreten Erwartungen der Nutzenden. Und weil Trauer keinen statischen Zustand beschreibt, sondern einer gewissen Dynamik unterliegt, ändern sich die auf den erlebten Verlust bezogenen Bedürfnisse im Laufe der Zeit. Obwohl sich Nutzer:innen für gewöhnlich darüber im Klaren sein dürften, dass sie nicht mit der verstorbenen Person, sondern mit ihrer virtuellen Nachbildung, also mit einem nicht bewusstseinsfähigen technischen System kommunizieren (und das ‚Weiterleben‘ darum eher in einem symbolisch-metaphorischen Sinne zu verstehen ist), kann dieses Wissen durchaus von Emotionen überlagert sein. Indem bestimmte Erwartungen an ein lebendiges menschliches Gegenüber in die Technologie hineinprojiziert werden, nimmt der Avatar für seine Anwender:innen eine Bedeutung an, die über den Status einer ‚unpersönlichen‘ Maschine hinausreicht.

Die Entwicklungen im Umfeld der DAI sind nicht lediglich Ausdruck eines technologischen Wandels, sondern auch der Transformation von Trauer- und Erinnerungskulturen – und damit ein Thema von hoher gesellschaftlicher und politischer Bedeutung. Neben einigen ethischen Fragen zu einem verantwortungsvollen Umgang und der Schaffung adäquater rechtlicher Rahmenbedingungen werden hierdurch traditionelle Vorstellungen von Identität, Lebendigkeit, sozialer Präsenz und Endlichkeit herausgefordert. Diesbezügliche Grenzüberschreitungen erscheinen umso bemerkenswerter in einer Gesellschaft, die zwar als säkularisiert beschrieben wird, in der gleichwohl weiterhin Transzendenzsehnsüchte zu beobachten sind. Auf einer abstrakteren Ebene ließe sich wiederum diskutieren, was das Wesen einer Person ausmacht, inwiefern es sich in den Mengen hinterlassener Daten überhaupt adäquat abbilden lässt – und ob die Anerkennung der eigenen Endlichkeit ein Schlüssel für ein gutes Leben ist oder gerade die Sehnsucht nach der Überwindung dieser Grenze ein menschliches Kerncharakteristikum darstellt.

Obschon sich das Thema des digitalen Weiterlebens mittels KI insbesondere innerhalb der journalistischen Medienberichterstattung derzeit einer auffallend großen Popularität erfreut, befinden sich die besagten Technologien noch in einem vergleichsweise frühen Stadium und ihr Potenzial ist längst nicht ausgeschöpft. Sollte es künftig möglich sein, anhand von immer weniger Daten und zu immer geringeren Kosten Avatare zu kreieren, die sich ihren analogen Vorbildern in sprachlicher wie optischer Hinsicht zunehmend annähern, dann wird der Markt der DAI weiter wachsen und neue Geschäftsmodelle hervorbringen. Auch der stetige demografische Wandel könnte hierfür ein begünstigender Faktor sein: Wenn die Zahl der sogenannten *digital natives* weiter steigt und betreffende Personen in ein Lebensalter kommen, in dem sie vermehrt mit Tod, Trauer und Verlust konfrontiert werden, dürfte dies zu einer noch stärkeren Verbreitung und gesellschaftlichen Akzeptanz von digitalen Umgangsweisen führen. Und vielleicht ist es auch schlichtweg eine Frage der kulturellen Gewöhnung daran, dass sich Menschen nicht mehr nur face-to-face, sondern via Webcam am Bildschirm begegnen, dabei auch mit künstlichen virtuellen Personen interagieren (inklusive *realer* emotionaler Folgen) und Lebensrückblicke online zu finden sind. Angesichts dieser und weiterer Normalitätsverschiebungen könnten spätere Generationen die Frage nach der Berechtigung und Brauchbarkeit von Avataren des digitalen Weiterlebens möglicherweise anders beurteilen, als dies die allermeisten Menschen aktuell tun, wenn es um die Suche nach einer adäquaten Antwort auf einen Verlust geht.

Bei alledem besteht jedoch die Gefahr, dass zentrale ethische und rechtliche Fragen übersehen werden, wenn besagte Entwicklungen bzw. die mit ihnen einhergehenden Risiken nicht frühzeitig erkannt und im Sinne des Gemeinwohls moderiert werden. Der Einfluss von KI auf das gesellschaftliche Leben im Allgemeinen und auf Trauer- bzw. Gedenkkulturen im Besonderen, der Bedarf an juristischen Regelungen zum Schutz postmortaler Persönlichkeitsrechte und die Herausforderung, ethische Standards in einem sich schnell entwickelnden kommerziellen Markt zu wahren, zeigen die Dringlichkeit einer breiten wissenschaftlichen und gesellschaftlichen Auseinandersetzung. Gerade für solch ein noch wenig reguliertes Aktionsfeld erscheint es umso notwendiger, einen geeigneten Raum zu schaffen, der Menschen einen informierten, reflektierten und selbstbestimmten Umgang mit Techniken des digitalen Weiterlebens ermöglicht.