

# PROJECT IMMERSIVE DEMOCRACY

Part of the European Metaverse Research Network

## Hassrede im Metaverse

Esen Küçüktütüncü & Dani Shanley | Juli 2023

### ÜBER DIE AUTORINNEN

**Esen Küçüktütüncü** ist Dokumentarfilmerin, bildende Künstlerin und Forscherin in den Bereichen VR (Virtuelle Realität) und Klinische Psychologie. Sie schloss ihr Masterstudium in kognitiven Systemen und interaktiven Medien an der Universität Pompeu Fabra ab und ist derzeit Doktorandin an der Universität Barcelona im Event Lab unter der Leitung von Mel Slater. Ihre Doktorarbeit konzentriert sich auf die Entwicklung gemeinsamer virtueller Umgebungen und virtueller Agenten, die sich mit Konfliktlösung und sozialer Harmonie befassen.

**Danielle Shanley** ist Postdoktorandin an der Fakultät für Kunst- und Sozialwissenschaften der Universität Maastricht in den Niederlanden. Derzeit erforscht sie die ethischen Implikationen immersiver Umgebungen (die fortschrittliche Technologien wie virtuelle Realität und künstliche Intelligenz kombinieren). Shanleys Fachwissen liegt hauptsächlich im Bereich der Wissenschafts- und Technologiestudien (STS) und der Technologiephilosophie – mit besonderem Schwerpunkt auf reflexiven, partizipativen Designmethoden (oder verantwortungsvollen Innovationen) wie Social Labs und Value Sensitive Design (VSD).

Hassrede bezeichnet nach der Definition der Vereinten Nationen (UN) jede Form von Kommunikation, Gesten oder Verhaltensweisen, die zu Gewalt, Diskriminierung, Feindseligkeit oder vorurteilsbasierten Handlungen gegen Einzelpersonen oder Gruppen auf der Grundlage von Merkmalen wie Rasse, ethnischer Zugehörigkeit, Religion, Geschlecht, sexueller Orientierung oder anderen Eigenschaften aufrufen. Sie umfasst Äußerungen, die Einzelpersonen oder Gemeinschaften erniedrigen, entmenschlichen oder stereotypisieren, schädliche Stereotypen aufrechterhalten und Intoleranz fördern (UN o.J.).

Obwohl diese Definition klar und prägnant erscheint, ist die Definition von Hassreden nicht so leicht. Jeder Versuch, Hassreden zu definieren, muss ein Gleichgewicht zwischen der Wahrung der Meinungsfreiheit – einem grundlegenden Menschenrecht – und dem Schutz von Einzelpersonen und Gruppen vor Schaden anstreben. Dies wird noch dadurch erschwert, dass verschiedene Personen ein- und dieselbe Äußerung unterschiedlich wahrnehmen können, und was eine Person als beleidigend oder schädlich ansieht, kann eine andere Person als legitime Meinungsäußerung betrachten. Die subjektive Natur der Beurteilung, ob bestimmte Äußerungen die Grenze zur Hassrede überschreiten, macht es schwierig, allgemeingültige Definitionen festzulegen.

Trotz dieser Definitionsschwierigkeiten ist es von entscheidender Bedeutung, sich mit dem Thema zu befassen, insbesondere in Online-Umgebungen wie dem Metaverse. Wie wir aus der frühen Forschung zu virtuellen Welten gelernt haben, kann das, was in diesen Welten geschieht, die Werte der Menschen in der realen Welt beeinflussen – und umgekehrt. So argumentierte etwa Christine Ward Gailey in den frühen 1990er-Jahren, dass Videospiele die vorherrschenden kulturellen Werte in der Gesellschaft widerspiegeln und Verhaltensweisen verstärken und fördern, die mit der vorherrschenden Ideologie übereinstimmen (Gailey 1993). Kommerziell erfolgreiche Spiele wiederholen und verstärken demnach oft die Werte und Aktivitäten, die mit den vorherrschenden gesellschaftlichen Normen verbunden sind.

## VORURTEILE IN VIRTUELLEN WELTEN

Daher ist es vielleicht nicht überraschend, dass die Darstellung verschiedener sozialer Gruppen in den virtuellen Welten von Videospiele tendenziell bestehende Vorurteile widerspiegelt, die Einzelpersonen in der realen Welt hegen, einschließlich Sexismus und Rassismus. So ist beispielsweise die geschlechtsspezifische Stereotypisierung des Aussehens von Videospielecharakteren, bei der Frauen oft als dünn mit großen Brüsten und betonten sexuellen Merkmalen dargestellt werden (ein Trend, der durch die Figur Lara Croft in der „Tomb Raider“-Reihe verkörpert wurde) und Männer oft mit einem muskulösen Körperbau und aggressivem Verhalten dargestellt werden, gut dokumentiert. Psychologische Studien unter Jugendlichen und Studierenden haben gezeigt, dass diese stereotypen Darstellungen zu einer Desensibilisierung gegenüber Sexismus in der realen Welt führen können, sodass Sexismus weniger schockierend erscheint und schädliche Überzeugungen wie „Vergewaltigungsmythen“ aufrechterhalten werden, die den Opfern sexueller Übergriffe die Schuld geben (Breuer et al. 2015).

Sexismus in Videospiele ist natürlich nicht nur ein Problem, wenn es um den Inhalt geht. Auch das Nutzungsverhalten, insbesondere in Multiplayer-Online-Spielen, ist oft höchst problematisch. So werden beispielsweise weibliche Spielerinnen oft angegriffen und belästigt und müssen ihre eigenen Bewältigungsmechanismen und Strategien entwickeln, um sich vor unerwünschtem Verhalten zu schützen.

Auch rassistische Darstellungen sind seit Anfang der 1990er-Jahre ständig ein Thema in Videospiele. Spiele wie Duke Nukem 3D und Shadow Warrior sind ein Beispiel dafür, wie verletzend rassistische Stereotype die Erzählungen von Videospiele geprägt haben. Die Handlung von „Duke Nukem 3D“ zum Beispiel dreht sich um die eugenische Panik vor einer Rassenvermischung zwischen eindringenden Außerirdischen und weißen Frauen in einem zukünftigen Los Angeles, das als mono-ethnisch dargestellt wird. Die Hauptfigur begibt sich auf die Mission, die außerirdischen Invasoren aufzuhalten, um die genetische Reinheit der menschlichen Spezies zu bewahren. Im Spiel „Shadow Warrior“ hingegen werden mit der stereotypen und generischen „asiatischen“ Identität des Protagonisten eine Reihe von Fähigkeiten verbunden, die als biologisch bedingt dargestellt werden, und welche so die Vorstellungen über die vermeintliche mangelnde Männlichkeit des Protagonisten aufrechterhalten (Weise 2021).

## SEXISTISCHE UND RASSISTISCHE VORSTELLUNGEN HABEN AUSWIRKUNGEN AUF DIE REALE WELT

Die in Videospiele dargestellten sexistischen und rassistischen Ideen und Darstellungen können über die virtuelle Welt hinausgehen und Auswirkungen auf die reale Welt haben. Psychologische Forschungen deuten darauf hin, dass das Spielen von gewalttätigen Videospiele Ethnozentrismus verstärken und eine erhöhte Aggression auslösen kann, wenn die spielenden Personen jemandem begegnen, der anders ist als sie selbst (Ewoldsen et al. 2012). Obwohl Aktivistinnen wie Anita Sarkeesian diese Themen in der Spieleindustrie aufgreifen,<sup>1</sup> haben sich Hass und diskriminierende Einstellungen in Videospiele im Laufe der Zeit normalisiert. Das wahrscheinlich auch die Gestaltung und Entwicklung anderer virtueller Welten beeinflussen wird. Im Jahr 2022 wurde in den Medien immer wieder auf das „Groping-Problem“ im Metaverse (unangemessene sexuelle Berührungen der Avatare) hingewiesen, was zeigt, wie wichtig es ist, die Probleme anzusprechen und zu hinterfragen, um inklusive und sichere Online-Umgebungen zu schaffen.

## HASSREDE IN SOZIALER VR

Auf sozialen VR-Plattformen wie VRChat, AltSpace und Meta Horizons nimmt Hassrede oft ähnliche Formen an wie in den traditionellen sozialen Medien. Nutzende können diskriminierendes oder beleidigendes Verhalten an den Tag legen, das sich gegen Einzelpersonen oder Gruppen aufgrund ihrer Eigenschaften oder Identitäten richtet.

---

<sup>1</sup> Feminist Frequency (2013). Damsel in Distress: Part 1 - Tropes vs Women in Video Games. [https://youtu.be/X6p5AZp7r\\_Q](https://youtu.be/X6p5AZp7r_Q) (Abgerufen am 27. Juni 2023).

Einige Beispiele hierfür sind:

- **rassistische oder abfällige Bemerkungen:** Nutzende können sich verbal rassistisch äußern, sich an rassistischer Stereotypisierung beteiligen oder diskriminierende Kommentare auf der Grundlage von race oder ethnischer Zugehörigkeit einer Person machen.
- **homophobes oder transphobes Verhalten:** Hassrede kann sich als verbale Belästigung, Mobbing oder Ausgrenzung äußern und sich gegen Personen aufgrund ihrer sexuellen Orientierung oder Geschlechtsidentität richten.
- **religiöse oder kulturelle Intoleranz:** Nutzende können Hassrede äußern, indem sie diskriminierende Einstellungen oder beleidigende Bemerkungen gegenüber bestimmten Religionen oder kulturellen Gruppen machen.
- **Cybermobbing und Belästigung:** Ähnlich wie herkömmliche soziale Medien können soziale VR-Plattformen zu Räumen für gezielte Belästigung werden, in denen Menschen Online-Mobbing, Drohungen oder beleidigendem Verhalten ausgesetzt sind.

Es lassen sich zwar wichtige Lehren aus der Art und Weise ziehen, wie diese Formen des Verhaltens in sozialen Medien stattfinden, gleichwohl gibt es eine Reihe wichtiger Unterschiede zwischen Hassrede in sozialen Medien und Hassrede in sozialen VR-Plattformen, insbesondere angesichts ihrer immersiven Natur.

Soziale VR-Plattformen bringen zusätzliche Elemente in die Nutzungserfahrungen ein, welche die Art und den Umfang von Hassrede, Mobbing und Diskriminierung beeinflussen können. Diese Plattformen bieten Nutzenden die Möglichkeit, Avatare zu verkörpern und sich in immersivere Interaktionen zu begeben. Infolgedessen können neue oder andere Verhaltensweisen entstehen.

Einige Beispiele hierfür sind:

- **nonverbale Äußerungen:** Hassrede kann über die verbale Kommunikation hinausgehen. Nutzende können Avatare nutzen, um beleidigende oder diskriminierende Gesten, Handlungen oder visuelle Darstellungen vorzunehmen, was die Wirkung von Hassrede verstärken kann.
- **räumliche Nähe und Präsenz:** In VR-Umgebungen können die Nutzenden physisch navigieren und mit anderen in unmittelbarer Nähe interagieren. Diese physische Präsenz kann die emotionale Wirkung von Hassreden verstärken und zu einem erhöhten Gefühl der Belästigung oder Diskriminierung führen.
- **immersive Erfahrungen und Anonymität:** Die immersive Natur von VR kann ein verstärktes Gefühl von Anonymität und Enthemmung vermitteln, was im Vergleich zu traditionellen Social-Media-Plattformen zu extremerem oder beleidigendem Verhalten führen kann.

## ABSCHWÄCHUNG VON HASSREDE IN SOZIALER VR

Um die negativen Auswirkungen von Hassrede in sozialen VR-Plattformen abzuschwächen, ist es für Plattformanbieter wichtig, eine proaktive Moderation in den Vordergrund zu stellen, klare Community-Standards festzulegen und die Selbstbestimmung der Nutzenden durch Meldetools und Bildungsinitiativen zu fördern. Gemeinsame Bemühungen von Plattformentwicklung, Nutzenden und relevanten Interessenvertretungen können dazu beitragen, sichere und einladende virtuelle Umgebungen zu schaffen, in denen Nutzende sich frei äußern können, ohne Angst vor Belästigung oder Diskriminierung zu haben.

In der neueren technologie-ethischen Literatur wird der Wert eines deliberativen Engagements für die Gestaltung der technologischen Entwicklung betont.<sup>2</sup> Diese Arbeiten stützen sich weitgehend – sowohl implizit als auch explizit – auf Konzepte der „deep democracy“ (Buhmann and Fieseler 2021, 101475), indem sie das epistemische Potenzial offener Beteiligungsprozesse hervorheben. Ganz allgemein wird in dieser Literatur vorgeschlagen, dass Innovator\*innen als proaktive Teilnehmende einer breiteren öffentlichen Debatte bzw. eines Diskurses zu verantwortungsvollen Innovationsprozessen beitragen können. Im

---

<sup>2</sup> Siehe zum Beispiel eine Reihe von Artikeln, die im *Journal of Responsible Innovation* und im *Journal of Responsible Technology* veröffentlicht wurden.

Wesentlichen geht es also darum, das Potenzial der verschiedenen Formen des Engagements zu nutzen, um optimale Lösungen zu finden.

Um das Jahr 2010 herum wurde der Begriff „Verantwortungsvolle Innovation“ („Responsible Innovation“) zu einem beliebten Begriff, um über Fragen der Verantwortung zu sprechen – sowohl in der Wissenschaft als auch bei politischen Entscheidungsträger:innen. Gemeint ist damit, Forschung und Innovation so zu organisieren, dass ihre Auswirkungen sicher, gerecht und auf die gesellschaftlichen Bedürfnisse abgestimmt sind.

Neue und aufkommende Technologien, wie KI und VR, werden unsere Zukunft auf beeindruckende Weise verändern. Wie wir bereits sehen können, wird dies dazu führen, dass wir uns mit neuen Fragen zu Risiken, Ethik, Gerechtigkeit und Gleichheit auseinandersetzen müssen. Verantwortungsvolle Innovation liefert uns die Konzepte und Praktiken, die erforderlich sind, um diese Fragen anzugehen, und hilft uns, über Dinge wie Hype, Umfang, Macht und Einbeziehung in Forschung und Innovation nachzudenken.

In Anlehnung an die Arbeiten von Buhmann und Fieseler kann verantwortungsvolle Innovation als ein Konzept betrachtet werden, das drei Hauptdimensionen umfasst (und die sich häufig im öffentlichen Diskurs über neue und aufkommende Technologien widerspiegeln). Erstens die *Verantwortung, Schaden zu vermeiden*, was sich zum Beispiel auf Risikomanagementansätze bezieht, mit denen potenziell schädliche Folgen kontrolliert werden sollen. Zweitens die *Verantwortung, Gutes zu tun*, die sich auf die Verbesserung der Lebensbedingungen bezieht, wie sie in den Zielen für nachhaltige Entwicklung festgelegt sind. Und schließlich die *Governance-Verantwortung*, die sich auf die Verantwortung für die Schaffung und Unterstützung globaler Kontrollstrukturen bezieht, die die beiden erstgenannten Verantwortlichkeiten unterstützen können (Ebd.).

Beispiele für Instrumente und Konzepte, die unter das Dach der verantwortungsvollen Innovation fallen, sind:

- **wertorientiertes Design:** ein Rahmen zur Erforschung der Werte von Interessengruppen, um diese Werte dann durch iterative konzeptionelle, empirische und technische Untersuchungen in operationelle Designkriterien zu übersetzen. Wertorientiertes Design stellt Fragen wie: Welche Werte sollen in die Gestaltung einbezogen werden? Wie lassen sich diese Werte in den Entwurfsprozess einbringen? Wie können Entscheidungen und Kompromisse zwischen widersprüchlichen Werten getroffen werden? Wie lässt sich überprüfen, ob das entworfene System die beabsichtigten Werte verkörpert?
- **Szenario-Planungsworkshops:** Narrative oder Szenarien sind im Wesentlichen hypothetische Abfolgen von Ereignissen, die konstruiert werden, um die Aufmerksamkeit auf kausale Prozesse und Entscheidungspunkte zu lenken. In diesem Sinne kann die Entwicklung von Szenarien zum Lernen und zur Beratung genutzt werden, um Entscheidungsprozesse zu schaffen, die auf dem Einbezug und der Interaktion verschiedener Interessengruppen beruhen.
- **Visionen-Karten:** Die Visionen-Karten sind eine Kombination aus wertorientiertem Design und Szenarienplanung. Sie basieren auf Visionierungskriterien, die das Bewusstsein für langfristige und systemische Fragen im Design schärfen sollen. Die Karten geben Anregungen zum Durchdenken verschiedener Implikationen und Wertespannungen und können in Workshops oder Teamsitzungen eingesetzt werden, um Diskussionen und Reflexionen anzustoßen.

Letztlich können wir durch diese Ansätze zur Technologieentwicklung versuchen, potenzielle Schäden wie Hassrede und Vorurteile zu konfrontieren und abzumildern, bevor sie sich festsetzen.

## DISKUSSION

Im Hinblick auf soziale VR-Anwendungen zeichnen sich bereits mehrere wichtige Probleme ab, die eher früher als später angegangen werden müssen. So stellt beispielsweise die Verwaltung des Verhaltens von Nutzenden auf Online-Plattformen aufgrund des fragmentierten Bereichs und der vielen beteiligten Systeme eine große Herausforderung dar. Jede Plattform kann ihre eigenen Regeln und Richtlinien haben, was zu widersprüchlichen Werten und Normen führt. Außerdem wird die Durchsetzung von Altersbeschränkungen schwierig, da die Nutzenden ein falsches Alter angeben oder die Beschränkung einfach umgehen können. Ein weiteres Problem besteht darin, dass die Geschwindigkeit der Entwicklungen im digitalen Raum oft die Fähigkeiten übersteigt, wirksame Moderationsmaßnahmen umzusetzen. Es ist

auch schwierig, branchenweite Standards gesetzlich zu verankern, sodass es schwierig ist, einheitliche Leitlinien für die Regulierung von Inhalten aufzustellen. Das Testen dieser Plattformen unter realen Bedingungen ist ebenfalls kompliziert, da die virtuelle Umgebung noch relativ unbekanntes Terrain ist. In Anbetracht dieser Faktoren müssen wir sicherstellen, dass virtuelle Welten verantwortungsvoll entwickelt werden. Das erfordert eine kontinuierliche Anpassung und Zusammenarbeit zwischen den Entwickler\*innen der Plattformen, den Benutzenden und den Regulierungsbehörden.

Bisher haben die Bemühungen zur Bekämpfung von Hassrede vor allem darin bestanden, den Dialog zu fördern, die Bildung zu unterstützen und die Medienkompetenz zu fördern, um das Verständnis, die Empathie und den Respekt zwischen Einzelpersonen und Gemeinschaften zu verbessern. Im Falle der Entwicklung des Metaverse wird die Zusammenarbeit zwischen Regierungen, zivilgesellschaftlichen Organisationen und Technologieunternehmen auch eine entscheidende Rolle bei der Entwicklung von Richtlinien, Maßnahmen und Werkzeugen spielen, um Hassrede wirksam zu bekämpfen und gleichzeitig unterschiedliche Perspektiven zu respektieren und kulturell sensibel zu agieren.

Es ist wichtig anzuerkennen, dass soziale VR-Plattformen zwar zweifellos Möglichkeiten zur Sozialisation und Kreativität bieten, dass sie aber auch vor großen Herausforderungen stehen, wenn es darum geht, gegen Hassrede, Mobbing und Diskriminierung vorzugehen. Wie erörtert, sind die Moderation von Inhalten und die Durchsetzung von Richtlinien aufgrund der dynamischen und immersiven Natur von VR-Umgebungen komplexe Aufgaben. Die Nutzung von Ideen und Konzepten aus dem Bereich der verantwortungsvollen Innovation, die Einführung von Meldemechanismen, die Aufklärung der Nutzenden und die Entwicklung von Gemeinschaftsrichtlinien können und sollten jedoch alle eine Rolle bei der Förderung integrativer und respektvoller virtueller Gemeinschaften spielen.

## LITERATURVERZEICHNIS

Breuer, J., Kowert, R., Festl, R., & Quandt, T. (2015). Sexist games= sexist gamers? A longitudinal study on the relationship between video game use and sexist attitudes. *Cyberpsychology, Behavior, and Social Networking*, 18(4), 197-202.

Buhmann, A. & Fieseler, C., (2021). Towards a deliberative framework for responsible innovation in artificial intelligence. *Technology in Society*, 64, 101475.

Ewoldsen, D. R., Eno, C. A., Okdie, B. M., Velez, J. A., Guadagno, R. E., & DeCoster, J. (2012). Effect of playing violent video games cooperatively or competitively on subsequent cooperative behavior. *Cyberpsychology, Behavior, and Social Networking*, 15(5), 277-280.

Gailey, C. W. (1993). Mediated messages: Gender, class, and cosmos in home video games. *Journal of popular culture*, 27(1), 81.

UN (o. D.). What is hate speech? <https://www.un.org/en/hate-speech/understanding-hate-speech/what-is-hate-speech> (Abgerufen am 27. Juni 2023).

Weise, Matthew (2021). The Hidden, Destructive Legacy of 'Duke Nukem'. <https://www.vice.com/en/article/pkdvxb/the-hidden-destructive-legacy-of-duke-nukem> (Abgerufen am 27. Juni 2023).