

---

# Das NVIDIA Omniverse und Audio2Face

Marvin Clauß

Hochschule der Medien, Stuttgart, Deutschland

Wintersemester 2021/22, Aktuelle Themen, Professor Dr. Koch

Künstliche Intelligenz, Aktueller Stand und ihre Auswirkungen auf Ihre Zukunft

---

23. März 2022

**D**ie Technologiebranche fokussiert sich in den vergangenen Jahren immer mehr auf den Aufbau und Ausbau einer massentauglichen Schnittstelle für KI-Entwicklung und Anwendung. Durch die Veröffentlichung des Omniverse ist NVIDIA diesem Ziel nun einen beachtlichen Schritt näher gekommen. Dieses Whitepaper befasst sich mit dem Aufbau des Omniverse, gibt Aufschluss über den starken Fokus auf KI-Unterstützung und bietet einen Überblick über USDs, OptiX, DLSS und die RTX-Technologie, auf der NVIDIAs Umsetzung eines Metaverse basiert. Zusätzlich wird das Deep-Learning-basierte Gesichtsanimation-Programm Audio2Face genauer beleuchtet und der Aufbau dessen neuronaler Netze erklärt. Abschließend wird ein kurzer Überblick zu aktuellen und potenziellen Anwendungsfeldern der Software gegeben.

## 1 Das Omniverse

Bereits seit einigen Jahren findet KI eine immer breitere Anwendung innerhalb einer großflächigen Auswahl an Industriebranchen. Besonders im Bereich des Deep Learnings, Machine Learnings und des Natural Language Processings werden dabei immer weiter Fortschritte erzielt, welche die KI-Forschung vorantreiben und den Einsatzbereich KI-basierter Systeme stetig erweitern. (Schoppmeier, 2021) Mit der Ankündigung des Omniverse im Frühjahr 2021 legt auch NVIDIA als großer Grafikkartenhersteller das eindeutige Ziel fest, die Forschung im Bereich der KI-basierten und unterstützten Systeme maßgeblich vorantreiben zu wollen. Dabei setzt der Konzern vor allem auf eine starke Partnerschaft mit einer Vielzahl marktführen-

der Unternehmen und will mit dem Omniverse eine Art Metaverse für die Entwicklung, Forschung und das Arbeiten mit KI, sowie KI-gestützten Systemen schaffen. (Johnston, 2021; CNET Highlights, 2021)

### 1.1 Was ist das Omniverse

Laut NVIDIA handelt es sich bei Omniverse um eine einfach erweiterbare Open Collaboration Plattform, welche neben ihrer Funktion als HUB für Plug-Ins und Erweiterungen sowie Microservices auch als Launcher für NVIDIA-eigene Anwendungen dient. Omniverse nimmt die Rolle eines Knotenpunkts für Programme zum Prototypen, Visualisieren, Animieren und für Echtzeit-Testing ein. Dabei spielen die hauseigenen RTX-Grafikkarten sowie die damit verbundenen Möglichkeiten des Echtzeit-Raytracings und -Pathtracings eine zentrale Rolle bei der Ermöglichung einer virtuellen Zusammenarbeit in einem cloudbasierten Serversystem. NVIDIAs RTX-Server sollen Daten mehrerer Nutzer gleichzeitig streamen und dabei live bearbeitbar machen. Das gesamte System stützt sich dabei auf das von Pixar entwickelte USD-Dateiformat, welches für einen einheitlichen Workflow und eine gemeinsame, optimierte Arbeitspipeline sorgen soll. (NVIDIA, 2022c; Bayer, 2021; NVIDIA Studio, 2022)

Um den Arbeitsansatz mit Omniverse konkurrenzfähig zu machen, zielt NVIDIA auf die Kollaboration mit einem möglichst breiten Spektrum an Partnerfirmen und etablierten Pipelines ab. Zum Anfang des Jahres 2022 befinden sich so bereits Firmen wie Adobe, Autodesk, Maxon, Turbosquid, wacom, Epic oder Unity unter den Projektpartnern und bieten variierende Grade an Einbindung in das Omniverse an. (NVIDIA, 2022c)

Basierend auf einer Ausrichtung auf ein vormerklich Business-orientiertes Klientel stößt die Plattform zum aktuellen Zeitpunkt vor allem in der Architekturbranche sowie der Automotive- und Film- beziehungsweise Games-Industrie auf eine positive Resonanz und verzeichnet bereits kurze Zeit nach Veröffentlichung des Omniverse-Launchers einen regen Zustrom an Interessenten. Die Möglichkeit der Beschleunigung von Renderings- und Simulationszeiten sowie der auf KI abgestimmte Workflow innerhalb der Omniverse-Apps bietet hier in vielen Bereichen eine potenzielle Zeitersparnis und Optimierung und eröffnet Firmen den Zugang zu leistungsstarker Technik, mit der eine Umsetzung von Projekten in Echtzeit ermöglicht wird. Auch im Bereich der Forschung lässt sich ein vermehrtes Interesse an der Plattform erkennen. Hier bildet vor allem die Möglichkeit der RTX-Server-gestützten KI-Forschung einen treibenden Faktor. Durch den öffentlichen und im privaten Rahmen kostenlosen Zugang zum Omniverse stößt die Plattform auch bei Künstlern und Hobby-Enthusiasten auf viel Anklang, da durch RTX-Server und die KI-Optimierung eine leistungsstarke Umgebung geboten wird, welche die meisten Heimnetzwerke bei weitem übersteigt. (NVIDIA, 2022c)

## 1.2 USDs

Ein großer und auch treibender Faktor bei der Umsetzung des Omniverse ist eine universelle Nutzung des von Pixar entwickelten USD-Fileformats (Universal Scene Description), welches in vieler Hinsicht maßgeblich zur Vereinheitlichung der Arbeits-Pipelines beiträgt. Der Aufbau des USD-Format lässt sich mit einer Programmiersprache oder Game-Engine vergleichen und dient ähnlich wie diese als zentraler Ankerpunkt für Modeling, Shading, Animation, Lighting, FX und Rendering. Anders als bei herkömmlichen Pipelines bietet das USD-Format allerdings einen allgemeingültigen Dateityp, der von allen Anwendungen eingelesen und gemeinsam genutzt werden kann. Dies erzielt den Vorteil, dass Dateien beim Vorschreiten in der Pipeline nicht konvertiert werden müssen, was Arbeitszeit und Verwaltung einspart.

Durch den Multilayer-Aufbau des USD-Formats ist zusätzlich ein gleichzeitiges Arbeiten an einer Scene möglich, ohne dabei Merge-Konflikte hervorzurufen. Dank nativer Multi-GPU-Unterstützung und integrierem Data-Streaming ist außerdem ein Arbeiten in Echtzeit möglich, was den generellen Arbeitsfluss deutlich beschleunigt. (NVIDIA, 2022d; Pixar Animation Studios, 2021)

## 1.3 Omniverse und KI

Da sich das Omniverse zu großen Teilen auf KI-basierten Systemen stützt und als Arbeitsgrundlage Echtzeit-Raytracing und Pathtracing verwendet, ist die Umsetzung der Plattform erst mit dem Aufkom-

men der RTX-Grafikkarten und den damit verbundenen Tensor-Cores möglich geworden. (NVIDIA, 2022c) Bei diesen handelt es sich um speziell auf neuronale Netze ausgelegte GPU-Kerne, welche in den CUDA-Einheiten der Grafikkarten verbaut sind. Die Hauptaufgabe der Tensor-Cores besteht dabei aus Matrix-Multiplikationen, welche mit einem 16 Bit Float Standard berechnet werden und eine Präzision von bis zu 32 Bit aufweisen. Reduziert man diese, können die Kerne auch mit 8 beziehungsweise 4 Bit Int rechnen, was den Prozess bei Deep-Learning-Operationen massiv beschleunigt. RTX-Grafikkarten mit Tensor-Cores erreichen dabei eine im Durchschnitt 8-fach gesteigerte Geschwindigkeit bei Rechenoperationen gegenüber Grafikkarten ohne Tensor-Cores. (NVIDIA Developer, 2019)

Besonders die NVIDIA-eigenen Anwendungen profitieren von dieser nativen RTX-Unterstützung beim Deep-Learning. Machinima nutzt Deep-Learning für das Webcam-basierte Aufzeichnen von Animationen und Live-Mapping auf In-Engine Characters. Bei Audio2Face kommt die Technik bei der Synthese von Audiodateien zur Erstellung von Gesichtsanimationen zu Einsatz. (NVIDIA, 2022c) Generell profitieren alle Omniverse-Apps vom auf Deep-Learning basierenden Antialiasing-System DLSS (Deep Learning Super Sampling). Dieses nutzt neuronale Netze, um Ergebnisbilder beim Echtzeit-Raytracing hochzuskalieren und den Detailgrad des Ausgabebildes zu erhöhen. Dabei werden mehrere Samples eines Pixels betrachtet, in Subpixel zerlegt und die Farbwerte auf Basis der KI-Trainingserfahrung zugewiesen. Da viele Objekte bei einer entfernten Darstellung auf dem Bildschirm kleiner als ein Pixel sind, kann so mithilfe der KI ein deutlich realitätsnäheres und schärferes Bild erzeugt werden als bei algorithmenbasierten Verfahren. (NVIDIA, 2022a)

Ein weiterer Bereich, der von KI übernommen wird, ist das Denoising (Entrauschen) der Renderergebnisse. Diese Aufgabe übernimmt bei den Omniverse-Apps das von NVIDIA entwickelte OptiX-System. Als klassisch trainiertes System kann OptiX auf ein umfangreiches Training mit zufälligen Noisepatterns zurückgreifen, durch welches sich die KI Wissen über Licht und Schatten, Tiefenschärfe, Reflexionen und andere definierende Faktoren aneignen konnte. Mit diesem Wissen geht der Denoiser über die Renderergebnisse und befüllt dabei fehlende Informationen mit passenden Werten. Ein solches Denoising-Verfahren ist bei Echtzeit-Raytracing unerlässlich, da auf Grund des Zeitmangels nur eine begrenzte Anzahl an Raytracing-Samples pro Bild erreicht werden kann. Ohne Denoising würde bei der Bildausgabe nur ein stark rauschendes Ergebnis ankommen. Der Vorteil, den OptiX und KI-basiertes Denoising hier bieten, ist, dass im Gegensatz zu algorithmenbasiertem Denoising eine adaptive Filterkerngröße angewandt werden kann, wodurch ein weniger stark verschwommenes Ergebnis erzielt wird. (NVIDIA Developer, 2022)

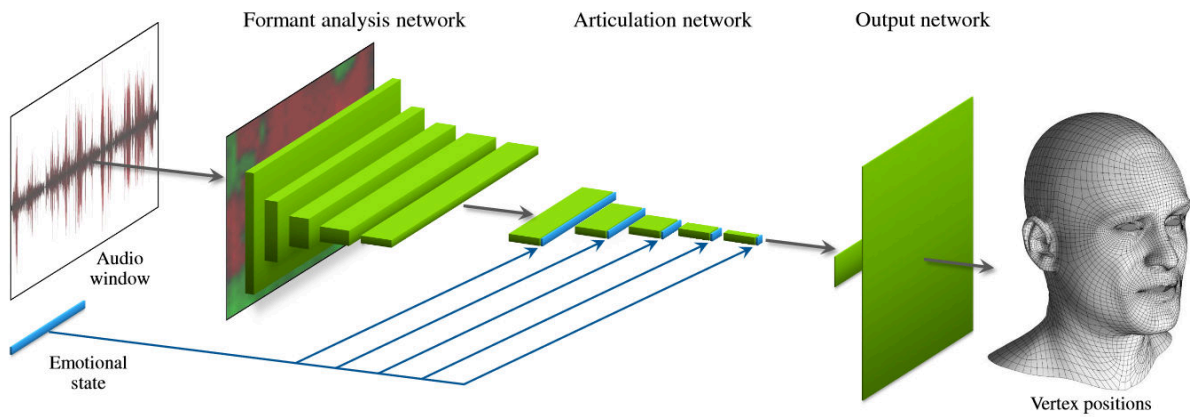


Abbildung 1: Aufbau der neuronalen Netze von Audio2Face. (Karras u. a., 2017)

## 2 Audio2Face

### 2.1 Was ist Audio2Face

Wie der Name dieser Omniverse-App vermuten lässt, handelt es sich hierbei um eine Anwendung zur Synthese von Audiodaten zur Gesichtsanimation. Das Ziel der App ist die Vereinfachung und Optimierung von Gesichtsanimations-Pipelines mithilfe einer Echtzeit-Animationssteuerung auf Basis von Audio-Input. Wie alle Omniverse-Apps basiert auch Audio2Face auf dem USD-File-Standard und ermöglicht ein Arbeiten in Echtzeit mit Raytracing-Qualität.

Als Input für die Anwendung dienen sowohl Audiodateien im Waveform-Format als auch live eingesprochene Tonaufnahmen. Diese werden mithilfe einer KI in Echtzeit auf ein 3D-Mesh gemappt und entsprechend der Audioquelle animiert. Audio2Face stellt hierbei einen Vorreiter auf dem Gebiet der Gesichtsanimationen dar, da bisherige Ansätze zur Deep-Learning-basierten Gesichtsanimation fast ausschließlich mit Videoerkennung als Referenz zur Animation arbeiten. Die synthetisierten Dateien können direkt in der Anwendung gebaked werden und anschließend als Animationsdateien für eine Verwendung in beispielsweise der Unreal Engine exportiert werden. (NVIDIA, 2022b)

### 2.2 Die KI hinter Audio2Face

Die grundlegende Funktionsweise der KI hinter Audio2Face entspricht einem angepassten Deep-Learning-Netzwerk. (NVIDIA, 2022b) Das bedeutet, die KI wird mit ausgewählten Daten trainiert, um gewünschte Muster zu erkennen und sich entsprechende Verhaltensweisen anzueignen. (Luber und Litzel, 2017)

Im Fall der Audio2Face-KI dient als Ausgangsdatei für die Eingabeschicht des neuronalen Netzes ein Audio-Input im Waveform-Format. Dieser Input wird zu einer 16 kHz Mono-Audiospur konvertiert

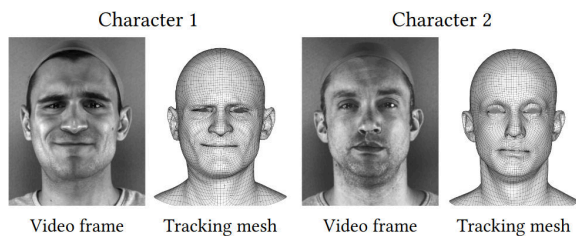
und in ein 2D-Raster unterteilt, bestehend aus dem sogenannten „Linear Filter“ und dem „Excitation Signal“. Die Informationen über Tonhöhe und Klangfarbe des Sprechers, welche dem „Excitation Signal“ entnommen werden können, ergeben kombiniert mit der Verhaltensweise des „Linear Filters“ ein klares Bild der Mund-, Zungen- und Rachenbewegungen des Sprechers, welches der KI als Ausgangspunkt für die Animation dient. (Karras u. a., 2017)

Um die Daten sinnvoll verarbeiten zu können, wird der Audio-Input zunächst in jeweils 520 ms lange „Snapshots“ unterteilt. Diese setzen sich aus dem aktuellen Aufnahmezeitpunkt und jeweils 260 ms vorheriger und folgender Aufnahme zusammen. Diese „Snapshots“ werden erneut in 64 sogenannte „Frames“ zerteilt. Da jeder „Frame“ aus 16 ms Audio beziehungsweise 256 Samples besteht, entsteht so ein Abstand der Frames von 8 ms oder 128 Samples. Das bedeutet, die Daten werden mit einem doppelten Overlap verarbeitet. Der fertige Frame wird in eine 64x32 Matrix übertragen und der KI als Eingabewert übergeben. Hier ist anzumerken, dass eine Matrix mit einer Datendichte von 64x12 Werten durchaus für die Mimikerkennung ausreicht, bei einer Dichte von 64x32 aber genügend Redundanz erzeugt wird, um die Mikromimik des Sprechers bestimmen zu können. (Karras u. a., 2017)

Bei der Verarbeitung der Audiodaten werden im Gesamten drei neuronale Netze durchlaufen, bevor ein Animationsergebnis erzielt wird. Im Format Analysis Network, welches als initiale Eingabeebene dient, werden die Audio-Daten aufgeschlüsselt und korrelierenden Werten zugeordnet. Im darauffolgenden Articulation Network werden die Daten mit zusätzlich eingespeisten Stimmungs- und Emotionsdaten verrechnet und anschließend dem Output Network übergeben. Dieses verarbeitet die entstandenen Daten und übergibt die Ausgabewerte den 5022 Kontroll-Vertices des Animationsmeshs. Bei diesem 3D-Mesh handelt es sich um das von

NVIDIA vordefinierte „Digital Mark“, an welchem die Audio2Face-KI trainiert wurde. (NVIDIA, 2022b; Karras u. a., 2017)

Als Trainingsdaten für die neuronalen Netze dienen mehrere Datensätze mit 3-5 Minuten hochqualitativer Audioaufnahme. Das Sprachmaterial ist auf Englisch und beinhaltet je ein bis drei Sätze (Pangramme) mehrerer Sprecher, welche mit wechselnden Schlüsselementen wiederholt werden. Der Fokus liegt dabei auf einer möglichst großen Varianz an Phonemen innerhalb der Sätze. Zusätzlich wird sogenanntes „In-Character-Material“ verwendet. Dabei handelt es sich um Referenzaufnahmen bereits existierender Gesichtsanimationen und den zugehörigen Sprachaufnahmen. Die KI soll auf diese Weise trainieren, eine bestimmte, vorgegebene Emotion beizubehalten, auch wenn der Voice Actor einen „Gesichtsausdrucks“ hat und kurz von der erwünschten Mimik abweicht. (Karras u. a., 2017)



**Abbildung 2:** Vergleichsbild zwischen Gesichtsausdrücken der Sprecher und dem Ergebnis der Audio2Face-KI. (Karras u. a., 2017)

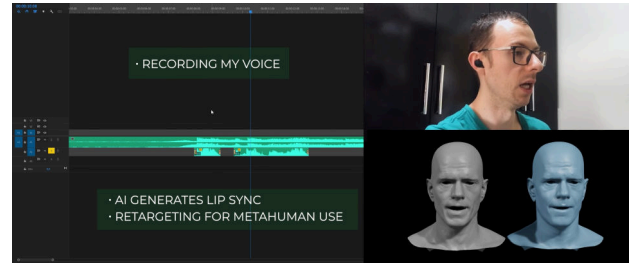
Die Pangramme und das „In-Character-Material“ ergeben zu gleichen Anteilen zwei Drittel der Trainingsdaten. Das verbleibende Drittel der „Audioframes“ bleibt zum Validieren der Trainingsergebnisse reserviert. Als Kontrolldaten für das Training dienen Videoaufzeichnungen der Sprecher beziehungsweise der Charaktere, welche mit einem videobasierten Vergleichsprogramm analysiert und ausgewertet werden. Der gesamte Trainingszyklus der Netzwerke entspricht dabei 500 Epochen. (Karras u. a., 2017)

## 2.3 Anwendungsbereiche

Da sich Audio2Face aktuell (Q1 2022) noch in der Entwicklung befindet und kein voller Feature-Umfang verfügbar ist, bleibt eine großflächige Einbindung in etablierte Workflows und Studio-Pipelines bisher aus. Zwar wächst das Interesse in der Film-Industrie an der App stetig, zum Beispiel als Hilfe bei der Gesichtsanimation von Hintergrundcharakteren, findet aber nur begrenzt Anwendung. (NVIDIA, 2022b)

Anders sieht es bei Indie-Produktionen und privaten Content-Creators aus. Hier kommt Audio2Face besonders in Kombination mit Epics MetaHuman und der

Unreal Engine 4 zum Einsatz und bedeutet für viele Projekte einen deutlich reduzierten Arbeitsaufwand und gesteigerte Arbeitsgeschwindigkeit. Auch einige YouTuber nutzen Audio2Face bereits für ihre Videoprojekte. (Anderson Rohr, 2021; Lee, 2021; Epic Games, 2022)



**Abbildung 3:** Anderson Rohr beim Aufnehmen der Animationsdaten seines Musikvideos. (Anderson Rohr, 2021)

Eine deutlich weitläufigere Anwendung findet die App bereits in der Games-Industrie. Hier ist besonders das Spiel „Mount & Blade 2: Bannerlord“ von TaleWorlds Entertainment zu nennen, bei dem sämtliche Gesichtsanimationen der computergenerierten Charaktere von Audio2Face übernommen werden. Auch andere Studios, darunter der Entwickler Flying Wild Hog (Shadow Warrior 3), Offworld Industries (SQUAD) und Piranha Games (Mechwarrior 5) nutzen bereits die Software. (NVIDIA, 2022b)

Ein weiteres Anwendungsgebiet ergibt sich aus dem Mobile-App- und Web-Bereich, wo das KI-Gesichtsanimationsverfahren bei der Umsetzung von Avataren für Chatbots eine Verwendung findet. (Tech Demo Team und NVIDIA, 2020)

## 3 Ein Blick in die Zukunft

Da Audio2Face bereits in limitierter Anwendung einen Einzug in die Animationsfilm-Sparte gefunden hat und dort sehr positiv aufgenommen wird, ist davon auszugehen, dass sich die Technik in den kommenden Jahren auf einem breiteren Markt durchsetzen kann. Besonders als Methode für die Gesichtsanimation von Hintergrundcharakteren bietet sich das Programm als zeitsparende und kostengünstigere Alternative zur händischen Animation oder einem Motion-Tracking der Sprecher an.

Auch in der Gaming-Branche ist das große Potenzial der Software erkennbar. Besonders in Kombination mit immer einfacher zugänglichen hochqualitativen Character-Meshes und einer nativen Einbindung in Game-Engines bietet Audio2Face hier einen Zugang zu hochwertigen Ergebnissen bei der automatisierten Gesichtsanimation, welche qualitativ von einem professionellen Motion-Tracking-System kaum unterscheidbar sind.

Zudem bietet sich Audio2Face als passende automatisierte Lösung für ein Umfeld an, in dem die Gesichtsanimation in Echtzeit generiert werden muss. Somit eignet es sich beispielsweise für robotische KI und Chatbots. In Collaborative Virtual Environments und dem Metaverse kann es das Lip-Syncing der virtuellen Avatare mit ihren Echtwelt-Pendants durchführen.

## 4 Fazit

NVIDIA treibt die Entwicklung von massentauglicher KI-Software mit dem Omniverse maßgeblich voran. Besonders durch den starken Fokus auf Echtzeit-Raytracing und -Pathtracing sowie eine Implementierung des USD-Fileformats bietet die Plattform einen gelungenen Lösungsansatz zur Vereinheitlichung von Arbeitspipelines und Workflows. Als Entwicklungs- und Arbeitsumgebung mit Fokus auf KI-gestützte Programme bietet das Omniverse eine ideale Grundlage für eine weitere Verbreitung von KI-Technologie auf dem Endkonsumentenmarkt.

Gerade Audio2Face mit seiner innovativen auf Audio-Input basierenden Deep-Learning-Technologie, zeigt sich als valide Alternative zu herkömmlichen Motion-Tracking-Verfahren und bietet Nutzern einen kostengünstigen und zeiteffizienten Lösungsansatz zur Automatisierung von Gesichtsanimationen. Besonders für Indie-Developer und -Künstler bietet Audio2Face in Kombination mit zum Beispiel Epics Unreal Engine 4 eine neue Möglichkeit zur Generierung von hochwertigen Animationsinhalten. Aber auch im professionellen Umfeld bringt Audio2Face das Potenzial mit sich, auf lange Sicht ein fester Bestandteil automatisierter Animationspipelines zu werden.

Schlussendlich gliedern sich das Omniverse und Audio2Face sehr gut in das immer weiter wachsende Feld der KI-basierten Anwendungen ein und werden auch in Zukunft weiter an Relevanz gewinnen.

## Quellenangaben

Anderson Rohr (Okt. 2021). *Audio2Face to Metahuman: Unreal Engine + NVIDIA Omniverse & RTXGI*. Video. Zuletzt aufgerufen am 14. März 2022. URL: [https://www.youtube.com/watch?v=fo-tmkdgZF8&ab\\_channel=AndersonRohr](https://www.youtube.com/watch?v=fo-tmkdgZF8&ab_channel=AndersonRohr).

Bayer, Martin (Nov. 2021). *Der digitale Zwilling der Erde*. Website. Zuletzt aufgerufen am 08. März 2022. URL: <https://www.computerwoche.de/a/der-digitale-zwilling-der-erde,3552174>.

CNET Highlights (Nov. 2021). *Nvidia's Jensen Huang reveals vision for the Metaverse! (with Omniverse)*. Video. Zuletzt aufgerufen am 08. März 2022. URL: [https://www.youtube.com/watch?v=RPfItXmIRG4&ab\\_channel=CNETHighlights](https://www.youtube.com/watch?v=RPfItXmIRG4&ab_channel=CNETHighlights).

Epic Games (2022). *High-fidelity digital humans made easy*. Website. Zuletzt aufgerufen am 21. März 2022. URL: <https://www.unrealengine.com/en-US/metahuman-creator>.

Johnston, Kasia (Apr. 2021). *NVIDIA Launches Omniverse Design Collaboration and Simulation Platform for Enterprises*. Website. Zuletzt aufgerufen am 08. März 2022. URL: <https://nvidianews.nvidia.com/news/nvidia-launches-omniverse-design-collaboration-and-simulation-platform-for-enterprises>.

Karras, Tero u. a. (2017). „Audio-driven facial animation by joint end-to-end learning of pose and emotion“. In: *ACM Transactions on Graphics* 36.4, S. 1–12. ISSN: 0730-0301. DOI: 10.1145/3072959.3073658.

Lee, Angie (Dez. 2021). *Meet the Omniverse: Omniverse Creator Uses AI to Make Scenes With Singing Digital Humans*. Website. Zuletzt aufgerufen am 21. März 2022. URL: <https://blogs.nvidia.com/blog/2021/12/16/omniverse-creator-anderson-rohr/>.

Luber, Dipl.-Ing. (FH) Stefan und Nico Litzel (Apr. 2017). *Was ist Deep Learning?* Website. Zuletzt aufgerufen am 14. März 2022. URL: <https://www.bigdata-insider.de/was-ist-deep-learning-a-603129/>.

NVIDIA (2022a). *NVIDIA DLSS*. Website. Zuletzt aufgerufen am 08. März 2022. URL: <https://www.nvidia.com/de-de/geforce/technologies/dlss/>.

– (2022b). *NVIDIA Omniverse Audio2Face-App*. Website. Zuletzt aufgerufen am 14. März 2022. URL: <https://www.nvidia.com/de-de/omniverse/apps/audio2face/>.

– (2022c). *NVIDIA OMNIVERSE Creating and Connecting Virtual Worlds*. Website. Zuletzt aufgerufen am 08. März 2022. URL: <https://www.nvidia.com/en-us/omniverse/>.

– (2022d). *What Is Universal Scene Description?* Website. Zuletzt aufgerufen am 08. März 2022. URL: <https://developer.nvidia.com/usd>.

NVIDIA Developer (2019). *Tensor Cores in a Nutshell*. Video. Zuletzt aufgerufen am 08. März 2022. URL: [https://www.youtube.com/watch?v=yyR0ZoCeB08&ab\\_channel=NVIDIADeveloper](https://www.youtube.com/watch?v=yyR0ZoCeB08&ab_channel=NVIDIADeveloper).

– (2022). *NVIDIA OptiX™ Ray Tracing Engine*. Website. Zuletzt aufgerufen am 08. März 2022. URL: <https://developer.nvidia.com/rtx/ray-tracing/optix>.

NVIDIA Studio (Jan. 2022). *What is NVIDIA Omniverse?* Video. Zuletzt aufgerufen am 08. März 2022. URL: [https://www.youtube.com/watch?v=dvdB-ndYJBM&ab\\_channel=NVIDIAStudio](https://www.youtube.com/watch?v=dvdB-ndYJBM&ab_channel=NVIDIAStudio).

Pixar Animation Studios (2021). *Universal Scene Description*. Zuletzt aufgerufen am 08. März 2022. Pixar Animation Studios. URL: <https://graphics.pixar.com/usd/release/index.html>.

Schoppmeier, Luisa (Sep. 2021). *KI in der Industrie: Auf dem Weg zur smarten Unternehmensführung*. Website. Zuletzt aufgerufen am 08. März

Das NVIDIA Omniverse  
und Audio2Face

2022. URL: <https://www.d-velop.de/blog/branchenprozesse/ki-in-der-industrie/>.

Tech Demo Team und NVIDIA (Okt. 2020). *Intelligent End-to-End AI Chatbot with Audio-Driven Facial Animation*. Website. Zuletzt aufgerufen am 21. März 2022. URL: <https://developer.nvidia.com/techdemos/video/d2s39-vid>.