

synthesia

Mit Synthesia ([synthesia.io](https://www.synthesia.io)) lässt sich Text in eine gesprochene Präsentation mit Video-Avatar erstellen. Die Qualität ist gut, Anwendungsbeispiele sind Marketing-Videos, Erklär-Filme, Firmen-Präsentationen.



Ähnlich wie bei RunwayML lassen sich mit **invideo.ai** ([invideo.io](https://www.invideo.io)) Video erzeugen. Der Funktionsumfang ist nicht ganz so groß wie bei RunwayML, jedoch haben sie beide den gleichen Slogan (-)



RunwayML (runwayml.com) ist ein online Videooditor und Video-Generator. Die Grundfunktionen können kostenlos nach Anmeldung getestet werden, darüber hinaus gibt es ab \$12/Monat unterschiedliche Tarife. Neben der Erzeugung von Videos durch Texteingabe lassen sich Videoeffekte (Vergrößern, Slow Motion, Effekte, Keying,...) mit exzellenter Bildqualität durchführen. Es lohnt sich definitiv ein Blick.

Hands down: OpenAI hat mit **Sora** (openai.com/sora) im Dezember 2024 die Welt zum Staunen gebracht, Google hat mit **Veo 2** (deepmind.google/technologies/veo/veo-2) das Ganze ein paar Tage später nochmals überboten. Beide Dienste sind zum Redaktionschluss noch nicht in Deutschland verfügbar, doch die sozialen Medien werden seitdem mit deren Videos überflutet. Dass die Film-, 3D- und VR-Branche ein Problem haben wird (oder sollte man sagen: ausgestorben [sind]?" - Jurassic Park), ist keine Frage mehr.

Video: Bearbeitung + Erstellung

Audio, Sprache, Musik

Das Feld von Audio, Sprache und Musik ist relativ jung. Die hier genannten sind teilweise erst seit 1-2 Jahren auf dem Markt, zeigen aber deutlich was bereits möglich ist und vor allem, wie schnell sich der Markt entwickelt:

Mit **Suno** (suno.com) aus den USA lässt sich Musik (inkl. Gesang) ausgehend von einer Texteingabe generieren. Nach der Anmeldung lassen sich 10 Songs kostenlos generieren. Es ist tatsächlich sehr spannend und das Ergebnis klingt wie ein durchproduzierter Song.

ElevenLabs (elevenlabs.io) aus den USA ist seit einiger Zeit eine der bekanntesten Text-To-Speech Plattformen. Im Vergleich zu herkömmlicher Sprach-Synthese überzeugen die Ergebnisse. Die eigene Erfahrung (Mitte 2024) hat jedoch auch gezeigt, dass es aktuell noch schwierig ist, die Tonalität der generierten Sprache so anzupassen, dass es für ein Projekt (bspw. Video-Trailer) wirklich brauchbar ist. Die "Regie-Führung" mit dem AI-Tool wird sich sicherlich bald verbessern. Auch das Klonen einer Stimme ist mit ElevenLabs möglich.



Lizzen (lizzen.de) ist ein junges StartUp aus Deutschland mit dessen Service sich Sprache aus Text generieren lässt.

Was ist mit Deepfakes? Auch hier ist heute noch mehr möglich als in der ersten Welle von Angst vor wenigen Jahren. Gute Anlaufstellen sind: [Reference.ai](https://www.reference.ai), [facemagic.ai](https://www.facemagic.ai), myheritage.de nicht mehr aktuell, dafür OpenSource: [deepfacerlab](https://deepfacerlab.com)

Face-Reverse Suche: Wer kennt es nicht: Das Foto einer fremden Person mit der Frage "Wer ist das?". Die Suchmaschinen haben bereits alle öffentlich zugänglichen Fotos von Personen (bspw. von Facebook, LinkedIn oder der Hochschul-Website) per Gesichtserkennung analysiert und geben diese Fotos entsprechend preis. Oder du möchtest einfach wissen, welche Bilder von dir im Internet sind? Viel Spaß mit [images.com](https://images.com oder facecheck.id) oder facecheck.id (Diese Tools wurden auch bei diesem Projekt zweiter Harvard-Studenten eingesetzt)



Bildgenerierung: Toll aber relativ aufwändig und einarbeitungsintensiv ist **ComfyUI** (comfy.org). Die Software ist OpenSource und "Node-Based" - dasheiß, einzelne KI-Elemente werden miteinander verbunden um am Ende ein Bild oder Video zu generieren. Es lässt sich über Pinokio installieren, die Models dazu müssen separat installiert werden.

Viele OpenSource KIs können kostenlos auf dem eigenen Computer ausprobiert werden. Die Nutzung ist jedoch nicht ganz so komfortabel wie bei den zuvor genannten Online-Anbietern. Die Installation der KIs wird über die App **Pinokio** (pinokio.computer) vereinfacht - quasi ein **App-Store für KIs**. Darüber lassen sich bspw. ComfyUI (s.u.) oder StableDiffusion installieren. Pinokio lädt aber auch zum entdecken und ausprobieren ein. Bei der Installation kann es helfen, etwas mehr Geld als die Durchschnittsperson zu sein. Bevor ihr es selbst installiert, könnt ihr es gerne im Labor der HfS "Ernst Busch" anschauen und ausprobieren.

Wie installiert man nun Stable Diffusion und ähnliche?!

Down the rabbit hole

Bildbearbeitung & -generierung ### Text-To-Image



[midjourney.com](https://www.midjourney.com)

Midjourney aus San Francisco ist ein *Generative Adversarial Network* (GAN) zur Erstellung von Bildern. Nutzbar ist es nur als bezahlter Dienst (ab \$10/Monat), die Ergebnisse gehören jedoch zu den Besten. Über den Messenger "Discord" können mittels des Kommandos `/imagine Bilder` erstellt werden.

Die Nutzung der Adobe-Produkte wie **Photoshop** ist generell kostenpflichtig, dafür wird eine KI zur Bilderstellung und Bearbeitung direkt mitgeliefert: Adobe **Firefly** - ~~Free~~ **Generative AI for creatives**



Canva aus Australien ist eine neue Generation von cloudbasierten Design-Tools, mit denen unterschiedlichstes erstellt werden kann: Präsentationen, Layouts, Flyer, Foto, Video... im Gegensatz zu den Adobe-Produkten wird man mit zahlreichen Design-Presets und -Vorschlägen unterstützt, so auch mit KI-basierten Bild- und Videobearbeitungs- und generierungs Tools. Teils kostenlos, für mehr Funktionen muss bezahlt werden.

Stable Diffusion von Stability AI (USA/Canada) kann direkt als bezahlte Version (ab \$9/Monat) bei stability.ai verwendet werden. Da die Firma den Quellcode zu ihrer KI im August 2022 öffentlich zugänglich gemacht hat, gibt es aber auch die Möglichkeit, Stable Diffusion lokal auf dem eigenen Rechner zu nutzen. Die Installation ist jedoch nicht ganz einfach und die Bildgenerierung ist deutlich langsamer. Viele Websiten und Programme nutzen zur Bildgenerierung jedoch Stable Diffusion. (siehe [Pinokio](https://www.pinokio.computer) auf der letzten Seite des Zines.)

Haftungsausschluss: Für die Inhalte der verlinkten Seiten ist stets der jeweilige Anbieter oder Betreiber der Seiten verantwortlich. Die verlinkten Seiten wurden zum Zeitpunkt der Verlinkung auf mögliche Rechtsversöße überprüft. Rechtswidrige Inhalte waren zum Zeitpunkt der Verlinkung nicht erkennbar.

Anastasia Lande
Annika Dura n Kandidator
Frederik Brizlmaier
Maria Gaiser
Paolo Frederico Attisi
Szeratina Schiesser
Yeganeh Shafie
&
Julian Jungel

"Ernst Busch", Berlin, erstellt.
Das Zine wurde gemeinsam mit Studierenden des Studiengangs *Spiel & Objekt* und dem Labor für Schauspiellkunst

labor@fts-berlin.de

Nichts funktioniert? Oder du vermisst etwas in unserem KI-Zine? Schreib uns, wir freuen uns über jeden Austausch!

Welcher Geschichte, die dir erzählt wird, glaubst du? Welchem Video, das du siehst, glaubst du?

bleibt spannend, was die Zukunft bringen wird. generierenden KIs wird dadurch immer einfacher und es hierzulande (Sora) erstellt werden. Auch Adobe, Canva, RunwayML etc. streben dies an. Die Nutzung von Mit einem (bezahlten) Zugang zu OpenAI können heute schon neben Text (ChatGPT) auch Bilder (Dall-E) und - bald auch alle Medien vereint.

Schlusswort: In diesem Zine unterteilen wir die Tools in drei Kategorien: 1. KI, die das korrekte, da bis etwa letztes Jahr die entsprechenden Rubriken wie Text, Bild, Video und Ton. 2. KI, die das korrekte, da bis etwa letztes Jahr die meisten Firmen singuläre Angebote hatten. Ziel der meisten Filme ist es, multimodale KIs anzubieten, die



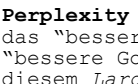
What everybody talks about: ChatGPT, Erstellung von Text, Übersetzung, Schreibhilfen



chatgpt.com

ChatGPT von OpenAI wurde Ende 2022 der Öffentlichkeit zugänglich gemacht und revolutionierte, wie wir arbeiten. Es ist eingeschränkt kostenlos - auch ohne login! - nutzbar.

Bei der bezahlten Version können hochgeladene Datenbanken und Medien analysiert und Bilder, Video und Audio generiert werden. Auch können eigene ChatBots erstellt werden und andere Modelle(*) verwendet werden.



perplexity.ai

Perplexity ist für viele das "bessere ChatGPT" und "bessere Google". Toll an diesem *Large Language Model* (LLM) ist, dass es die Quellen angibt, auf denen die Antwort basiert.



Perplexity ist eine relativ kleine Firma mit 100 Mitarbeitern im Vergleich zu OpenAI (das zu den "Big Playern" gehören) mit 1700 Angestellten. Beide Firmen haben ihren Hauptsitz in den USA.



[deepl.com/write](https://www.deepl.com/write)

DeepL kennen die meisten als Übersetzungs-Tool, es bietet aber mit **DeepL-Write** auch die Möglichkeit, das Schreiben von Texten zu unterstützen. DeepL kommt aus Deutschland mit Hauptsitz in Köln.

Biased: Alle drei genutzten KIs können gut mit **Sprachen der "westlichen Welt"** umgehen. Versucht man, bspw. auf Farsi zu übersetzen, wird kein zufriedenstellendes Ergebnis geliefert. Möchten wir Tools mit solchem Bias nutzen?

(*) *GPT-Modelle* sind die angelernten Trainingsmodelle und Algorithmen, mittels denen der Chatbot antwortet. Angelernt werden sie, indem Millionen Websites und Bücher durch die KI "gelesen" werden. GPTs (Generative Pretrained Transformer) gehören zur Gruppe der LLMs (Large Language Models)

Januar 2025



Pope Francis in puppy winter jacket -
Erstellt mit MidJourney, Unknown
author, 24.03. 2023

Die wichtigsten Apps und Services zur Erstellung und Bearbeitung von Medien mittels KI!

KÜNSTLICHE INTELLIGENZ

A QUICK GUIDE TO...

JANUAR
2025
EDITION

Wusstest du...

...dass der erste Chat-Bot "ELIZA" bereits 1965 von Joseph Weizenbaum programmiert wurde?

...dass die Idee zu "Künstlichen neuronalen Netzen" (worauf alle aktuellen KI-Entwicklung basieren) bereits 1957 von Frank Rosenblatt entwickelt wurde?

Was damals "künstliche Intelligenz" genannt wurde, sind aus heutiger Perspektive vergleichsweise einfache Algorithmen, die nicht (mehr) als 'intelligent' gelten. Die rasante Entwicklung von KI in den letzten Jahren ist den Computerspielen zu verdanken. Für die immer realistischeren Grafiken werden kontinuierlich bessere Grafikeinheiten ("GPUs") entwickelt. Zufälligerweise können diese nicht nur Millionen Pixel pro Sekunde auf den Bildschirm bringen, sondern auch die Millionen Neuronen eines Neuronalen Netzes berechnen.

Der Energieaufwand hierfür ist beachtlich: Eine normale Suchabfrage bei Google benötigte 0,3 Wh - eine Anfrage bei ChatGPT etwa zehnmal so viel, 3Wh(*). Das ist eine kleine LED-Glühbirne, die eine Stunde brennt. Anstelle eines Gesprächs mit ChatGPT, könnte man also auch das Licht zu Hause die Nacht brennen lassen.

Aufgrund der Kosten für die Energie und die dahinterliegenden Computer, sind nur wenige gute KI-Services kostenlos. Die, die (meist eingeschränkt) gratis nutzbar sind, wollen Appetit auf mehr machen - um mit Geld oder Daten bezahlt zu werden. Vertrauenswürdige KIs belassen die Daten auf unserem Computer und berechnen sie lokal, d.h. nicht auf einer Serverfarm, und wir behalten die Souveränität über unsere Daten. Denn auch ohne teure Grafikeinheiten zur Berechnung der Neuronalen Netze lassen sich immer mehr KIs auch auf normalen Laptops und Smartphones nutzen. Die Entwicklung läuft weiter und so können wir hoffen, dass der Energiekonsum auch bei den großen KIs irgendwann wieder geringer wird - denn aktuell planen Microsoft, Google & Amazon den (Wieder-)Aufbau von Atomreaktoren für ihre Datenzentren(**).

(*) <https://www.heise.de/en/news/ChatGPT-s-power-consumption-ten-times-more-than-Google-s-9852327.html>

(**) <https://www.nytimes.com/2024/10/16/business/energy-environment/amazon-google-microsoft-nuclear-energy.html>

Disclaimer: Die Zahl an KI-Tools zur Mediengestaltung wächst täglich und ist unüberschaubar. Unmöglich, ein Zine wie dieses auch nur annähernd Vollständigkeit zu nennen. Die genannten Produkte sind solche, die sich etwas stärker auf dem Markt ausgedehnt haben, möglichst einfach zu bedienen sind und auf die sich unserer Meinung nach ein Blick zu werfen lohnt.

Die Reise geht weiter...

Bei der Recherche fanden wir deutlich mehr Angebote für die Medienbearbeitung und Medienerstellung mit KI als wir selbst austesten, bewerten und im Zine abbilden konnten. Hier ist eine Liste weiterer Services, die auf den ersten Blick spannend aussahen:

MUSIC / SOUND / SPEECH

Music / Sound

- * Lyric jam live lyric generator (Free with login)
- * Aiva KI-Musikkomposition und -Arrangement (Free trial with login)
- * Amper Music Musikgenerierung durch KI (Free trial with login)
- * Boomy Eigene Musik durch KI erstellen (Free with login)
- * Endless Musik-Jams mit KI-Unterstützung (Free with login)

Text To Speech

- * lovo.ai (Free trial with login, KI-basiert)
- * DEEPai.org (Free trial without login, KI-basiert)
- * Natural Reader (Free trial with login, KI-basiert)
- * Speechify (Free trial with login, KI-gestützt)
- * Voicemaker.in (Free without login, KI-optimiert)

AVATARE / 3D / SPRECHENDE AVATARE

Fotorealistisch

- * DeepBrain AI Text-to-Video mit realistischen KI-Avataren (Free trial with login)
- * Movio KI-basierte Avatar-Videoerstellung (Free trial with login)
- * D-ID Realistische sprechende Avatare aus Text oder Audio (Free trial with login)

3D

- * Replika Therapeutischer Chatbot mit Avatar (Free trial with login)
- * DeepMotion Animate 3D 3D-Animation durch KI aus Videos (Free trial with login)
- * Avatarify KI-Avatare in Echtzeit basierend auf Deepfake-Technologie (Free without login)

AI Picture & Video

Generating Video

- * Filmcrux.com Generative Video AI (paid)
- * Move AI AI Motion Capture for creating 3D animation (free, paid upgrade)

Video Editing and Image/Video Search:

- * Papercup – An AI platform that artificially creates voiceovers and dubs for videos in different languages
- * Flim – An AI-powered fully searchable image database of stills from movies, TV series, music videos and commercials (free, paid upgrade)

Avatare

Avaturn

Erstellt photorealistische Avatare aus Fotos. Die Avatare sind 3D-animiert und bereit für den Einsatz in Spielen oder virtuellen Umgebungen.

- * Hochrealistische Avatare.
 - * Export in 3D-Umgebungen wie Unity oder Unreal Engine.
- (Basisversion kostenlos, kostenpflichtige Pro-Versionen verfügbar)

Ready Player Me

Diese Plattform wandelt ein Foto in einen 3D-Avatar um, der in Spielen, Apps oder virtuellen Welten verwendet werden kann.

- * Personalisierbar
- (Kostenlos)

Synthesia

Für professionelle Anwendungen. Wandelt Bilder in 3D-animierte Avatare um, die in Videos sprechen können.

- * Fokus auf realistische Animationen für professionelle Zwecke -> Unterstützung für Videos und Präsentationen.
- (Premium-Preisplan ab \$30/Monat.)

Replika

Ein AI-Companion, der personalisierte Avatare anbietet. 3D-Avatar individuell gestaltbar, aber mit Einschränkungen, visueller Gesprächspartner.

- * Fokus auf emotionale Interaktion und persönliche Gespräche.
 - * Anpassungsmöglichkeiten für den Avatar.
 - * Verfügbar als mobile App oder Web-Version.
- (Kostenlos mit optionalen Premium-Features)