

# KI in der Film- und Videoproduktion

Der praktische Einstieg

» Hier geht's  
direkt  
zum Buch

# DAS VORWORT

# Einführung

Liam Winter nahm den letzten Schluck Espresso und blickte über die Elbe. Eigentlich hatte er keine Zeit – aber die Strandperle war sein Ruhepol in einem vollen Arbeitstag. »Liam? Du bist es wirklich?« Er sah auf. Nora Kohlmann stand da, überrascht und lächelnd. »Nora! Wahnsinn, wie lange ist das her? Komm, setz dich!« Sie sprachen über alte Zeiten, wer wohin verschwunden war, alte Uni-Legenden.

Dann kam die Frage, die sich früher ständig gestellt hatte. »Und? Schauspieler geworden?« »In gewisser Weise.« »Film oder Theater?« »Film.« »Mit großem Set, Kamera, Licht und Action?« »Nicht ganz. Wenn du zehn Minuten Zeit hast – unser Büro ist gleich ums Eck.«

Sie betraten ein Industrie-Loft mit rohen Backsteinwänden, hohen Decken und bunten Graffiti. An drei Schreibtischen arbeiteten junge Leute konzentriert an ihren Laptops. Kein Set. Kein Scheinwerfer. Keine Kamera.

»Hier macht ihr Filme?«, fragte Nora skeptisch. »Pass auf.« Liam ging zur kleinen Küche, griff sich einen Kochlöffel und stürmte mitten ins Loft. »Ich bin Aedric von Valmor, Wächter der Flammenfeste!«, rief er dramatisch und schwang die imaginäre Waffe durch die Luft. Ben, ein Kollege, filmte ihn dabei kommentarlos mit dem Smartphone. »Wow«, sagte Nora trocken. »Also wenn das hier die Zukunft der Filmproduktion ist, dann bin ich Wonder Woman.«

Ben grinste und setzte sich wieder an seinen Laptop. »Gib mir zwei Minuten. Ich lass den Kochlöffel gleich mal zum Schwert werden.« Er begann, in ein Textfeld zu tippen: Medieval battlefield, Aedric of Valmor, cinematic lighting, smoke, epic

camera sweep, intense sword fight... Liam reichte Nora eine schlichte schwarze Brille. »Setz die mal auf.« Sie zögerte. »Ich seh' da jetzt hoffentlich keine Werbung für Ritterküchen?« »Nein. Du siehst einfach nur die Zukunft.«

Nora setzte die Brille auf. Zunächst sah sie noch den Raum – leicht überlagert von Menü-Elementen. Dann wechselte das Bild. Plötzlich stand Liam nicht mehr im Loft, sondern in voller Rüstung auf einem rauch verhangenen Schlachtfeld. Um ihn herum tobte ein epischer Kampf. Schlachtrufe, Explosionen, Pferde, Schwerter. Und er mittendrin – kein Schauspiel, sondern ein filmreifer Auftritt.

»Oh mein Gott«, flüsterte Nora. »Ziemlich echt, oder?« Sie nahm die Brille ab. »Okay, Liam. Ich geb's zu. Du hattest recht. Das ist total irre.« »Und das ist erst der Anfang. Netflix testet gerade ein neues Format.« »Und ich könnte da theoretisch auch ein Teil davon sein?« »Nicht nur theoretisch.«

Am Abend saß Nora auf dem Sofa. Auf ihrem Netflix-Startbildschirm blinkte ein neues Banner: »Shadow Protocol – Jetzt personalisiert verfügbar.« Aus Neugier hatte sie sich angemeldet, ein paar Selfies hochgeladen, ein Genre gewählt. Mehr war nicht nötig gewesen. Sie klickte auf »Play«. Und erstarrte.

Sie sah sich selbst. Als Agentin in einem futuristischen Thriller. Hochwertige Kameraarbeit, starke Dialoge – und mittendrin sie. Nicht billig animiert, sondern verblüffend lebensecht. Kein Deep Fake. Kein Filter. Sondern eine individuell generierte Hauptrolle in einem echten Netflix-Film. Nora griff zum Handy und schrieb an Liam: »So wie's aussieht, trage ich ab jetzt ein rotes Cape und ein goldenes Lasso. ;-)«

## Die Magie der Bilder – und der Beginn einer neuen Ära

1993 sprachen meine Freunde nur noch über einen neuen Film, den man unbedingt gesehen haben musste. Ein Film, der Dinosaurier so lebensecht auf die Leinwand brachte, wie man es zuvor noch nie gesehen hatte: *Jurassic Park*. Und da mich Filme schon immer begeistert haben, vor allem, wenn sie neue technische Tricks und visuelle Innovationen boten, war ich sofort dabei.

Ich saß also im Kino, John Williams' majestätischer Soundtrack begann, und auf einmal stand da ein echter Dinosaurier vor mir. Zumindest fühlte es sich so an. Diese Szene hat mich damals vollkommen umgehauen. Und offenbar ging es

nicht nur mir so: Der Film spielte weltweit über eine Milliarde Dollar ein – ein Rekord für die damalige Zeit.

Immer wieder in der Geschichte des Kinos führen technologische Durchbrüche zu neuen Filmideen.

- Als 1933 *King Kong* auf die Leinwand kam, staunten die Zuschauer über die damals revolutionäre Stop-Motion-Animation.
- 1940 begeisterte *The Thief of Baghdad* mit den ersten Bluescreen-Effekten.
- Und *Star Wars* veränderte 1977 die Branche mit einer der ersten Motion-Control-Kameraaufnahmen.
- 2002 erlebten wir mit *Der Herr der Ringe: Die zwei Türme* Gollum als ersten glaubwürdigen CGI-Charakter mit Motion-Capture-Technologie.

Und heute? Heute stehen wir erneut am Beginn einer Revolution. Ihr Name: *Künstliche Intelligenz*. Doch diesmal ist etwas anders. Unsere erste Begegnung mit dieser Technik findet nicht im Kinosaal statt, sondern auf TikTok und Instagram. Millionen einfacher, teils amateurhaft produzierter KI-Clips fluten unsere Feeds. Viele empfinden sie eher als lästig – oft, weil sie nicht wissen, was dahintersteckt, oder weil es immer schwieriger wird, echtes Bildmaterial von KI-generierten Fakes zu unterscheiden.

Und dennoch dürfen wir nicht vergessen: Das ist erst der Anfang. Die Frage ist nicht mehr, *ob* KI unsere Film- und Videoproduktion verändern wird, sondern *wie schnell* und *wie tiefgreifend*. Wird es bald möglich sein, dass Einzelpersonen komplette Kinofilme mit dem Laptop im eigenen Zimmer produzieren – als moderne »Bedroom-Producer« der Filmwelt?

Dabei zeigt die KI längst auch im großen Hollywood-Kino, wozu sie fähig ist. Im Film *Here* von Regisseur Robert Zemeckis zum Beispiel werden Tom Hanks und Robin Wright mit Hilfe von KI über sechs Jahrzehnte hinweg gealtert und verjüngt – direkt am Set, in Echtzeit. Die Technologie dahinter heißt *Metaphysic Live*. Zwei Kameras liefen parallel: eine zeigte das rohe Bild, die andere die KI-bearbeitete Version mit jüngerer oder älterer Gesichtsstruktur – live und ohne monatelange Postproduktion. Tom Hanks selbst sagte über diese neue Technik: »Was früher ein halbes Jahr Arbeit bedeutete, passiert jetzt in einer Nanosekunde.«

Das ist nicht nur ein Meilenstein für digitale Effekte – es verändert die Produktionsweise selbst. Visuelle Magie entsteht dort, wo früher noch Komparsen,

Maskenbildner und ganze VFX-Teams wochenlang arbeiteten. Rob Bredow, Visual-Effects-Meister bei Industrial Light & Magic, sagte in einem TED-Talk: »Wir glauben nicht, dass bald Einzelne in dunklen Räumen alleine ganze Filme machen werden. Wir brauchen das Know-how unserer Künstler, Techniker, Lichtgestalter und Animator\*innen. Diese Vielfalt schafft Magie.«

Ich teile diese Meinung – aber nur teilweise. Ja, auch in Zukunft werden aufwendige Spielfilme weiterhin von großen Teams produziert, in denen jeder seine oder ihre Spezialaufgabe übernimmt. Die Zusammenarbeit vieler Expert\*innen erzeugt jene Tiefe, Qualität und narrative Raffinesse, die große Kinoproduktionen auszeichnen. Und natürlich wird künstliche Intelligenz in diesem Kontext eine unterstützende Rolle spielen, Prozesse effizienter machen, Kosten senken und kreative Möglichkeiten erweitern.

Aber ich muss Rob auch widersprechen. Ich bin überzeugt, dass wir gerade dabei sind, einen völlig neuen Markt entstehen zu sehen: Filme, die von Einzelpersonen produziert werden – unterstützt durch KI, in kürzester Zeit, mit beeindruckender Qualität. Diese Filme werden möglicherweise nicht für das Kino gedacht sein, vielleicht schaffen sie es auch nicht auf Netflix oder Amazon Prime. Doch auf Plattformen wie YouTube oder neuen Streaming-Diensten werden sie ihr Publikum finden – und das nicht zu knapp. Millionen Aufrufe, große Reichweiten und ein nachhaltiges Einkommen sind realistische Szenarien für diese neue Generation von Solo-Filmemacher\*innen.

Der amerikanische Unternehmer und Medien Visionär Tom Bilyeu – Mitgründer von Impact Theory und Quest Nutrition – formulierte es vor Kurzem in einem viel beachteten Video drastisch: »Hollywood ist tot.« Er sieht in den neuen KI-Technologien einen so tiefgreifenden Wandel, dass traditionelle Filmproduktionen in ihrer heutigen Form kaum überleben werden. Stattdessen prognostiziert er eine Ära, in der Einzelpersonen mit KI-Tools beeindruckende Inhalte erschaffen – ganz ohne großes Team.

Und dieser Umbruch ist nicht mehr Zukunftsmusik. Er hat bereits begonnen – mitten in Hollywood und bei den größten Streaming-Diensten. Netflix etwa hat kürzlich bei der argentinischen Serie *El Eternauta* erstmals KI eingesetzt, um eine spektakuläre Szene – den Einsturz eines Gebäudes in Buenos Aires – zu realisieren. Das Ergebnis: zehnmal schneller und deutlich günstiger als mit herkömmlichen Methoden. Co-CEO Ted Sarandos nannte KI »die Chance, Content nicht nur günstiger, sondern auch besser zu machen«.

Und es bleibt nicht beim Spielfilm: Auch in der Werbebranche hat die KI bereits Fuß gefasst. So produzierte etwa das Unternehmen Kalshi einen surrealen Werbespot für die NBA Finals – in weniger als 48 Stunden und mit einem Budget von nur rund 2.000 Euro. Der Macher PJ Accetturo ist stolz auf genau das: hohe kreative Wirkung bei minimalem Aufwand – ermöglicht durch Tools wie Google Gemini und Veo 3.

Auch das konservative Kreditkartenunternehmen Coign ließ seinen TV-Spot komplett mit KI erstellen – inklusive realistischer Darsteller, die in Wahrheit nie vor einer Kamera standen. Diese Spots liefen landesweit auf Plattformen wie Fox News oder YouTube – bei einem Bruchteil der Kosten klassischer Kampagnen.

Wenn also sogar Streaming-Giganten und große Marken mit Milliarden Budgets diesen Weg gehen, ist klar: KI ist nicht nur ein Trend, sondern längst Produktionsrealität.

Was bedeutet das für dich und mich? Ganz einfach: Wenn Netflix heute KI einsetzt, um Kosten und Zeit zu sparen, wird dieselbe Technologie schon morgen für alle verfügbar sein. Genau darauf bereitet dich dieses Buch vor.

## **Am Puls der KI-Revolution – Was dich in diesem Buch erwartet und warum jetzt der richtige Zeitpunkt ist**

Die letzten Monate haben im Bereich »KI und Film« eine ungeahnte Dynamik entfesselt. Quasi über Nacht scheint sich eine völlig neue Welt aufzutun, in der künstliche Intelligenz unsere Art des Filmmachens revolutioniert. Was vor Kurzem noch futuristisch klang, ist plötzlich greifbare Realität: KI-Tools generieren auf Knopfdruck beeindruckende Bilder und Videosequenzen, die noch vor einem Jahr unmöglich schienen. Sicher, noch mögen viele dieser KI-Clips auf TikTok oder Instagram als Spielerei erscheinen – manche nervig, andere faszinierend. Doch wir dürfen nicht vergessen: Das ist erst der Anfang. Die Geschwindigkeit, mit der sich KI-Technologien zur Videoproduktion entwickeln, ist atemberaubend, und sie nimmt weiter zu.

In diesem Buch erfährst du, was das konkret bedeutet. Es erwartet dich ein praxisnaher Streifzug durch die neuesten KI-Werkzeuge und -Techniken, die die

Film- und Videoproduktion bereits jetzt beeinflussen. Dabei ist mir eines besonders wichtig: Authentizität durch eigene Erfahrung. Schon in meinem letzten Buch »Künstliche Intelligenz in der Musik- und Audioproduktion« habe ich den Ansatz verfolgt, jede relevante Software und Plattform selbst auszuprobieren. Genauso bin ich auch diesmal vorgegangen. Ich habe jedes KI-Tool, das in diesem Buch vorkommt, eigenhändig getestet und für meinen YouTube-Kanal Schritt-für-Schritt-Tutorials erstellt.

In vielen Kapiteln findest du am Ende eines Themas einen QR-Code, der dich direkt zu einem passenden YouTube-Tutorial von mir führt. Dort siehst du das im Kapitel behandelte Tool in der Praxis: Benutzeroberfläche, Arbeitsablauf, Tipps und typische Stolpersteine. So kannst du das Gelesene sofort ansehen und umsetzen – eine ideale Ergänzung zu den Texten in diesem Buch. Denn Theorie ist wichtig – aber erst in der Praxis zeigt sich, was wirklich funktioniert.

Um nicht nur abstrakt über Möglichkeiten zu schreiben, habe ich mich selbst ans Werk gemacht und mehrere Praxisprojekte umgesetzt. Dazu gehören zum Beispiel Musikvideos zu eigenen Songs, die ich unter meinem Namen veröffentlicht habe, aber auch kleinere Videoprojekte für Instagram. Vielleicht bist du ja sogar auf eines meiner KI-Rätsel gestoßen: In kurzen Clips habe ich deutsche Sprichwörter visualisiert, und die Zuschauer konnten raten, welches Sprichwort gemeint ist. Zugegeben, das berühmte Beispiel »Ein Elefant im Porzellanladen« war leicht zu erkennen – ein Elefant trampelt durch ein Geschäft und zerdeppert dabei alles Geschirr. Aber einige der weniger bekannten Sprichwörter waren schon kniffliger zu erraten. Mir ging es dabei nicht nur um den Rätselspaß, sondern vor allem darum, Hands-on-Erfahrung zu sammeln. Solche Experimente zeigen nämlich sehr gut, wo die Technik heute steht.



Der Elefant im Porzellanladen

Die Ergebnisse waren oft verblüffend. Der gesamte Prozess – vom Ausdenken eines Szenarios über den Prompt (eine textbasierte Beschreibung für die KI, meist mithilfe von ChatGPT formuliert) bis hin zum fertigen Clip – dauerte manchmal nur ein bis zwei Minuten. Und das Resultat sah auf Anhieb beeindruckend aus. Man könnte tatsächlich glauben, ein echter Elefant sei mit einer echten Kamera in einem echten Porzellanladen gefilmt worden.

Doch zugleich fiel auf: Irgendeine Kleinigkeit stimmte fast immer nicht. Ein Beispiel: In einem meiner Clips ließ ich einen Sportler in seinem Wohnzimmer Anlauf nehmen, um wie bei einem Wettkampf mit dem Kopf voran durch die Wand zu rennen. Die KI generierte daraus eine spektakuläre Szene – bis auf ein Detail: Der Athlet blieb kurz vor der Wand stehen, wackelte nur seltsam mit dem Kopf, und im nächsten Moment explodierte die Wand von allein. Action? Ja. Logik? Nicht ganz. Den Prompt so lange nachzujustieren, bis die Szene stimmig war, erwies sich als überraschend langwieriger Prozess.

Ähnliche Grenzen zeigten sich bei anderen Ideen. Für ein weiteres Experiment wollte ich einen Löwen erzeugen, der auf ein kleines Kind in einem Käfig zuläuft, während außen mehrere Leute durch eine Glasscheibe zusehen. Doch hier spielte die KI gar nicht erst mit: Statt eines Videos bekam ich nur eine Fehlermeldung – mein Vorhaben verstieß wohl gegen die Inhaltsrichtlinien.

Diese Beispiele führen vor Augen, dass die Möglichkeiten zwar schon beeindruckend, aber längst noch nicht perfekt und grenzenlos sind. Die derzeit verfügbaren KI-Tools liefern verblüffende Ergebnisse, sind jedoch (noch) keine vollständig zuverlässigen Profi-Werkzeuge. Trotzdem bin ich überzeugt: Gerade jetzt ist es für alle, die in der Film- und Videoproduktion tätig sind, entscheidend, sich mit diesen Technologien vertraut zu machen. Früh dran sein lautet die Devise. Wer heute experimentiert, lernt, wo KI bereits Kosten senken und Zeit sparen kann – oder welche Aufnahmen plötzlich möglich werden, die früher undenkbar oder unbezahlbar gewesen wären. Natürlich wird es noch eine Weile dauern, bis wir wirklich komplette Kinofilme rein am Computer generieren können, ohne ein menschliches Filmteam und echte Schauspieler. Doch die Weichen sind gestellt. Es ist keine Frage mehr, ob KI unsere Branche verändert, sondern wie tiefgreifend und wie schnell.

Dieses Buch gibt dir einen umfassenden Überblick über den aktuellen Stand der Dinge. Du wirst unter anderem lernen:

- Mit reinen Textbeschreibungen Videos erzeugen – also nur durch Eingabe von Worten einen kompletten Clip erstellen.
- KI-generierte Bilder animieren – statische von KI erschaffene Grafiken in Bewegung versetzen.
- Virtuelle Darsteller sprechen oder singen lassen – Videos mit künstlich erzeugter Sprache oder Gesang vertonen.
- Visuelle Effekte per KI erschaffen – und erfahren, wie diese neuen Möglichkeiten mit klassischen Tools wie Schnittprogrammen zusammenspielen.
- Stimmen klonen und verfremden – echte Stimmen nachahmen oder vorhandene Sprachaufnahmen in eine andere Stimme umwandeln.
- Geräusche und Musik generieren – per KI Soundeffekte und Musikstücke komponieren und Audiomaterial optimieren.

All das erkläre ich anhand konkreter Tools und Beispiele, die ich selbst ausprobiert habe. Mein Ziel ist es, dir einen möglichst breiten – aber dennoch kritischen – Überblick zu verschaffen. Nicht alles, was machbar ist, liefert schon auf Knopfdruck professionelle Ergebnisse. Doch wenn du weißt, was heute schon geht und wo noch Grenzen sind, bist du einen entscheidenden Schritt voraus.

Aus eigener Erfahrung kann ich sagen: Der einzige Weg zu wirklich guten Ergebnissen führt über Neugier und Ausprobieren. Je mehr Tools du kennengelernt und selbst getestet hast, desto gezielter kannst du sie einsetzen. Dieses Buch möchte dir genau dabei helfen – als Kompass in der faszinierenden, manchmal verwirrenden Welt der KI-gestützten Film- und Videoproduktion. Ich lade dich ein, gemeinsam mit mir diese neue Welt zu erkunden.

## Über den Autor – Vom Ö3-Jingle zur KI-Regie

Als Kind hatte ich zwei große Träume: Musikproduzent und Filmregisseur zu werden. Mit sieben gründete ich mit meinem Freund Peter meine erste Band, die Flying Potatoes. Parallel drehte ich mit Freunden kleine Filme – zehnmünütige Clips mit Drehbuch, Regie und Szenenwechseln, für mich schon damals »echte« Filme.

Den Traum vom Musikproduzenten habe ich verwirklicht. Den Traum vom Regieführen lebt heute meine Tochter, die Film studiert hat und inzwischen beim ORF

in Wien arbeitet. Das hat eine schöne Logik: Mein Vater wollte Musiker werden, konnte es nie – seinen Traum habe ich gelebt, und meine Tochter lebt meinen.

Gemeinsam mit Peter gründete ich später die Firma *Foster Kent*. Mit 22 produzierten wir unsere ersten Produktionen für Hitradio Ö3 – damals der einzige österreichische Sender mit Popformat. Unsere Musik lag als durchgängiges Sounddesign unter dem gesamten Wetter und unter den kompletten Verkehrsmeldungen; sie trug diese Sendestrecken, statt nur in kurzen Jingles aufzublitzeln. Kurz darauf gewannen wir eine große Ausschreibung des ORF-Fernsehens. Mit dem Wechsel zu ORF1 und ORF2 lieferten wir die komplette musikalische Verpackung – von Magazinen über das Kinderprogramm bis zur »Zeit im Bild«, dem österreichischen Pendant zur Tagesschau. Aus diesen Anfängen ist eine internationale Erfolgsgeschichte gewachsen: Unsere Musik ertönt heute unter anderem bei RTL (z. B. vor »RTL Aktuell«), auf 1010 WINS, einer der erfolgreichsten Nachrichten-Stationen der USA, und beim landesweiten US-Nachrichten-Netzwerk News Nation – einem schnell wachsenden Sender, der sich als Alternative zu CNN positioniert.

Vor zehn Jahren startete ich meinen YouTube-Channel *Thomas Foster Musikproduktion*, in dem ich die Praxis moderner Musikproduktion erkläre – von Logic Pro über Ableton Live bis hin zu vielen weiteren Tools. Später habe ich den Kanal in *Thomas Foster Sound & Vision* umbenannt: Für dieses Buch begann ich, mich intensiv auch mit Film zu beschäftigen, erste Tutorials zu Final Cut zu produzieren und meine Themen vom reinen Sound hin zu Sound und Bild zu erweitern. Aus dieser Arbeit entstanden Bücher. Mein drittes, »Künstliche Intelligenz in der Musik- und Audioproduktion«, wurde mehrfacher Amazon-Bestseller. Bei der Recherche merkte ich, dass es im deutschsprachigen Raum noch kein wirklich fundiertes Werk zum Thema »Künstliche Intelligenz in der Film- und Videoproduktion« gab. Also habe ich monatelang die besten YouTube-Tutorials systematisch ausgewertet, transkribiert und geordnet – weil dort Wissen aktuell, offen und mit großer Leidenschaft geteilt wird.

Parallel habe ich begonnen, Musik unter meinem eigenen Namen zu veröffentlichen – samt Videos, die ich selbst produziere. Innerhalb von zwölf Monaten wuchs meine Hörerschaft von 250 auf über 100.000 monatliche Hörer. In jüngster Zeit arbeite ich in Bild und Ton intensiv mit KI: Sie ermöglicht mir, Regie zu führen, Storyboards zu entwickeln und Ideen ohne klassische Produktionsgrenzen in bewegte Bilder zu übersetzen.

Als Mitgründer von *Mugent* entwickle ich außerdem Software und Sounds für Produzenten. Unser *Mugent Player* – ein Plugin (VST/AU) – liefert hochwertige Instrumente und wurde zuletzt um einen vielseitigen Arpeggiator erweitert. Dahinter steht mein Anspruch, Musizieren und moderne Technik sinnvoll zu verbinden.

Selten hat Technik für mich so viel Magie entfaltet wie heute. Wenn dieses Buch einen Bruchteil dieser Energie weitergeben kann, dann hat es seinen Zweck erfüllt. Viel Freude beim Lesen.