

KI generierte Filme und die Umwelt

Eine Recherche

Stand: 7/2026

Vor Kurzem habe ich in einem Seminar ein kleines Quiz veranstaltet: „Erkennen Sie KI?“. Die Ergebnisse waren mehr oder wie ich sie erwartet hatte. Einige Bilder wurden richtig identifiziert, andere nicht. Worauf ich nicht vorbereitet war, war jedoch die Fragen einer Studierenden: „Haben Sie denn die Bilder extra für uns generieren lassen – wissen Sie denn nicht, wieviel Wasser da verbraucht wird?“ Nein, ich wusste es nicht. Und da wir als Rixfilm unter anderem auch als Green Consultants unterwegs sind, hat mich das zugegebenermaßen ziemlich gewurmt.

Die Ergebnisse meiner Recherchen habe ich hier zusammengestellt. Und um es gleich zu sagen, sie sind mit Vorsicht zu genießen. Ich habe mich redlich bemüht und trotzdem wenig gefunden.

Es gilt daher: Nichts genaues weiß man nicht.

- 1 Der CO₂e-Fußabdruck von KI-generierten Videos
- 2 Die Antwort auf die Ausgangsfrage
- 3 Ein Vergleich mit der realen Welt
- 4 Mein Fazit oder warum die Frage nach dem Ressourcenverbrauch trotzdem relevant ist.

Bevor es losgeht der folgende Disclaimer: Meine Recherche hat nicht den Anspruch einer wissenschaftlichen Analyse. Methodisch vergleiche ich vermutlich immer wieder Äpfel mit Birnen. Trotzdem wollte ich herausfinden, was ich Menschen antworten kann, die die Umweltverträglichkeit von KI generell infrage stellen. Haben sie recht oder liegen sie falsch?

1 Der CO₂e-Fußabdruck von KI-generierten Filmen

Als naiver Ingenieur denkt man, dass sich der Energieverbrauch von Rechnern, die ein KI-Video generieren, einfach messen lassen müsste. Doch so banal scheint es nicht zu sein. Offensichtlich beeinflussen zu viele Parameter die Ergebnisse, sodass keine klaren Aussagen möglich sind.

Trotzdem gibt es aktuelle Studien und Reports, aus denen sich einige Informationen ableiten lassen. Links und Credits dazu finden sich am Ende meines Berichts.

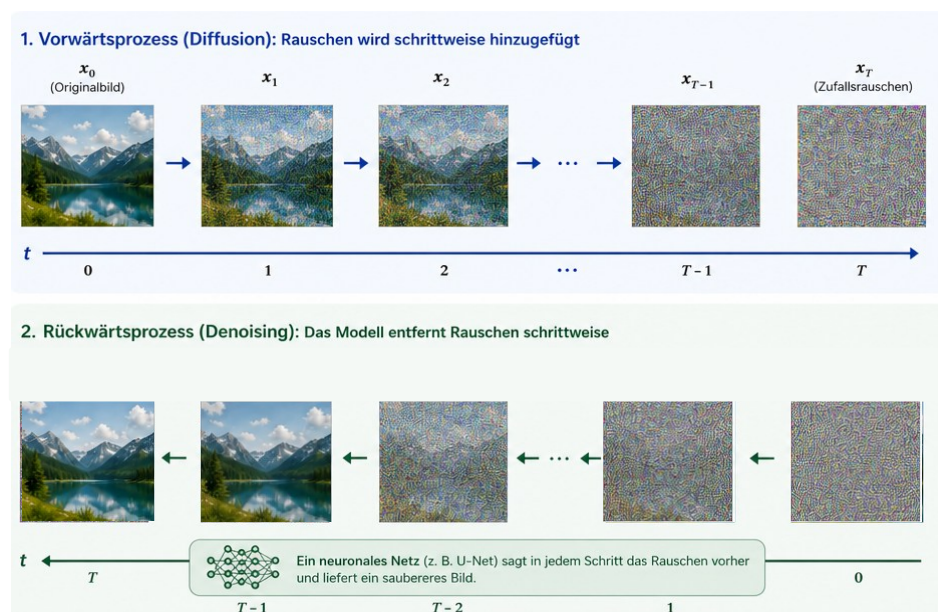
Das Wichtigste vorab: Der Ressourcenverbrauch und damit die Umwelt- und Klimabelastung steigt exponentiell von Text-to-Text über Text-to-Image zu Text-to-Video.

KI-generierte Filme sind „Stromfresser“, und verbrauchen etwa 2.000-mal so viel Energie wie ein klassischer Text, generiert von einem LLM (Delavande et al., 2025).

Der Energieverbrauch

Die genauen Werte variieren in den Studien, bewegen sich jedoch in ähnlichen Größenordnungen. Zentrale Parameter sind hierbei Länge des Clips, Auflösung, Bildrate, Anzahl der Denoising-Schritte sowie das verwendete KI-Modell.

Die Abhängigkeit des Energieverbrauchs von Länge, Auflösung und Bildrate ist etwa quadratisch (Delavande et al., 2025). Zumindest bei Länge und Bildrate verblüfft das zunächst, weil wir hier aus der „normalen“ Videowelt eher eine lineare Abhängigkeit sowie wie beim Denoising erwarten.



Denoising ist der Prozess, der sich im Hintergrund der Bildgenerierung per KI abspielt und in dem das Bild Schritt für Schritt aus einer völlig verrauschten Ausgangsbild entsteht. Wie viele Schritte dazu in den aktuellen Modellen notwendig sind, blieb nach meinen Recherchen unklar. Nur bei einigen Open-Source-Modellen, wie Wan 2.1-14B von Alibaba kann sie direkt beeinflusst werden. Wan 2.1-14B kann auf entsprechend leistungsstarken Rechnern auch lokal ausgeführt werden und ist deshalb bei Forschern sehr beliebt.

Verbrauchsangaben

Zu diesen Forschern zählen auch Nidhal Jegham und seine Mitstreiter. Sie kommen experimentell (bei lokalen Modellen) und durch Modellierungen auf folgende Zahlen für ein achtsekündiges Video in 720p-Auflösung (Jegham et al. 2026).

Modell	Energieverbrauch Mittelwert
Veo 3 (geschätzt)	30,8 Wh
Seedance-1 (geschätzt)	80 Wh
Runway Gen 4.5 (geschätzt)	322 Wh
Wan 2.2-14B (gemessen)	382,7 Wh
Sora 2 pro (geschätzt)	418,5 Wh

Die großen Unterschiede erstaunen und lassen sich, wenn ich die Autoren richtig verstehe, unter anderem durch die Rechnerarchitektur in den Rechenzentren erklären. Hier scheint Google den anderen Anbietern im Moment einen guten Schritt voraus zu sein.

Der Strommix

Vom Energieverbrauch zum CO₂e-Fußabdruck ist es theoretisch nur ein kleiner Schritt, sofern man weiß, wie viel CO₂ pro kWh erzeugtem Strom emittiert werden. Die landestypischen Werte sind hier sehr unterschiedlich.

Land	Fußabdruck Strommix in g CO ₂ e/kWh
Deutschland	330
Europa	284
USA	384
China	526

Alle Zahlen: <https://ember-energy.org/>

Da wir in der Regel nicht genau wissen, wo genau unsere Daten mit welchem Anteil an erneuerbaren Energien verarbeitet werden, ist der Strommix ein weiterer Unsicherheitsfaktor in meiner Berechnung.

Meine Kalkulation

Kalkuliert habe ich einmal für einen typischen deutschen TV-Primetime-Film wie den „Tatort“ und einmal für eine aufwendige und hochwertige internationale High-Budget-Kinoproduktion. Dabei habe ich jeweils das Modell mit der niedrigsten Emission (Veo) und das mit der höchsten (Sora 2 Pro) verwendet, um die Spannbreite möglicher Ergebnisse herauszufinden.

Folgende Annahmen habe ich getroffen:

	Kinoproduktion	TV Movie
Länge	120 Minuten	90 Minuten
Auflösung	4K	2K
Shots pro Film	2400	1000
Iterationen	100	50
Standort Rechenzentrum	USA	Europa

Die Iterationen entsprechen dem Drehverhältnis bei normalen Filmen. Also, wie oft ein Shot neu generiert werden muss, bis die Regie zufrieden ist. Die KIs schätzen 50 bis 100 Iterationen als durchaus realistisch ein, wenn die Qualität gut sein soll. Ebenso basiert die Annahme der Anzahl der Shots pro Film auf Erfahrungswerten.

Unter diesen Vorgaben ergeben sich folgende CO₂e-Fußabdrücke:

	Kinoproduktion	TV Movie
Veo 3	3,5 t CO ₂ e	0,4 t CO ₂ e
Sora 2 pro	49 t CO ₂ e	6 t CO ₂ e

Wie schon gesagt, ist das Ergebnis mit Vorsicht zu genießen, da die Berechnungen auf Schätzungen und Pauschalwerten beruhen. Grundsätzlich sollten die Größenordnungen jedoch stimmen.

Außerdem beziehen sich die Zahlen lediglich auf die Anwendung und Generierung (Inference) der Modelle. Der einmalige Energieaufwand für das vorgelagerte Training durch die Hersteller ist hierin nicht enthalten, verteilt sich aber bei verbreiteten Modellen auf hunderte Millionen Bilder und Videos.

2 Die Antwort auf die Ausgangsfrage

Eine Antwort auf die Ausgangsfrage nach dem Wasserverbrauch liefern die bisherigen Daten nicht – bestenfalls indirekt. Klar ist, dass ein Großteil der elektrischen Energie eines Rechenzentrums in Wärme umgewandelt wird und im Normalfall Wasser als ideales Kühlmittel eingesetzt wird. Das macht jedoch nur 15% bis 30% des Wasserverbrauchs aus. Der Rest wird indirekt bei der Stromerzeugung z. B. durch Kernkraftwerke oder fossile Kraftwerke verbraucht (Gage & Malarkey, 2025).

Google selbst nennt für den Wasserverbrauch einen Wert von 0,26 ml pro Text-to-Text-Prompt (Elsworth et al., 2025). In Bezug auf Text-to-Video konkreter (und dennoch schwammig) sind die Angaben des United Nations University Institute for Water, Environment and Health in seiner Studie Environmental Cost of AI's Energy Use (Aczel et al. 2026):

„Der mit der Stromerzeugung verbundene Wasser-Fußabdruck beträgt bei einem einzelnen Bild etwa zwei Esslöffel (29 ml), steigt bei einem komplexen Video jedoch auf 4,1 Liter“.

Doch was bedeutet komplex? Wie lange ist das Video? Auf Anhieb konnte ich dazu nichts finden.

3 Ein Vergleich mit der realen Welt

Für den Vergleich habe ich verschiedene Quellen genutzt:

- Der BFI Albert Report „A screen new deal“ weist für eine „Tentpole Production“ (mit einem Budget ab 70 Mio. USD) eine Treibhausgasemission von 2.840 t CO₂e aus (albert et al., 2026).
- Für deutsche Kinofilme schätzt der Berufsverband der Green Consultants Film&TV den Wert auf 150 t CO₂e.
- Ein Abschlussbericht von Zieglerfilm für einen nachhaltig produzierten „Tatort“ kommt auf ca. 80 t CO₂e (Schega, 2015)

Überraschend erscheint zunächst vor allem der gewaltige Unterschied bei den Kinofilmen. Stellt man die Emissionen jedoch in Relation zu den Produktionsbudgets, wird deutlich, dass der Unterschied gar nicht so groß ist. Grob gesagt liegt der CO₂e-Fußabdruck pro 1 Million EUR Produktionskosten zwischen 30 und 50 Tonnen.

Und was sagt der Vergleich zur KI-Produktion? Selbst wenn hier der schlechteste Fall betrachtet wird, scheinen KI-Produktionen im Vorteil. Einkalkuliert habe ich bereits CO₂e-Emissionen in Höhe von 3 t CO₂e für die benötigten Büroflächen (6 Monate), die Bereitstellung von Geschirr etc. und die Postproduktion (3 Monate) – berechnet mit dem CO₂-Rechner der MfG.

	Kinoproduktion		TV Movie	
	Best Case	Worst Case	Best Case	Worst Case
KI-Film	3,5 t CO ₂ e	49 t CO ₂ e	3,4 t CO ₂ e	9 t CO ₂ e
Realfilm	150 t CO ₂ e	2860 t CO ₂ e	80 t CO ₂ e	150 t CO ₂ e

4 Mein Fazit

Zuallererst: All das ist eine Rechenspielerie. Reine Theorie. Mit einer ziemlich hohen Wahrscheinlichkeit, dass reale Zahlen völlig anders aussehen. Zuviel basiert auf Schätzwerten und Annahmen, zu wenig auf konkreten Messungen.

Dennoch werde ich in Zukunft Studierenden relativ guten Gewissens eine Antwort darauf liefern können, inwieweit die Generierung von KI-Bildern und KI-Videos ökologisch verantwortbar ist. **Sie lautet: Als Alternative zu „echten“ Fotos und Filmen aus ökologischer Sicht, absolut ja.** Es spricht vieles dafür, dass KI-generierte Filmproduktionen einen deutlich geringeren CO₂-Fußabdruck verursachen als konventionelle Filmproduktionen ähnlicher visueller Komplexität.

Warum die ökologischen Folgen von KI-generierten Bilder und Videos dennoch problematisch ist, liegt in der schieren Menge an KI-Videos, die generiert werden. Derzeit – und in meinen Augen auch noch auf einige Jahre hinaus – vor allem für die diversen Social-Media-Kanäle. KI-Videos scheinen sich hier geradezu explosionsartig zu vermehren. Die United Nations University geht davon aus, dass in naher Zukunft 25 Millionen KI-generierte Clips am Tag generiert werden (Aczel et al. 2026). Wenn jeder Clip nur 8 Sekunden lang ist, entspricht das einem Volumen von über 400 „Tatorten“ pro Tag (bei 90 Minuten Länge, 50 Iterationen, 2K Auflösung) ...

Solche Rebound-Effekte durch die neue Technologie wird es auch in der eigentlichen Filmproduktion geben. Viele Produzierende und Regieleute sind sich darin einig, dass KI ein Türöffner für aufwendigere Effekte und Szenarien ist, die sonst gar nicht finanzierbar wären. KI ersetzt also nicht die klassische Produktion, sondern erweitert sie, sodass damit tendenziell auch deren ökologischen Fußabdruck steigt.

Best Practices

Was können wir tun, um die KI-Videoproduktion so umweltverträglich wie möglich zu gestalten? Hier einige Tipps:

- Effiziente Modelle wählen: Nicht für jeden Entwurf das größte Modell (z. B. Sora) nutzen, sondern für Layouts/Storyboards leichtere Modelle (z. B. Veo). Tendenziell sind neuere Modelle effizienter als ältere.
- Iterationen reduzieren durch präzise Prompts oder Image-to-Video-Referenzen
- Anbieter nutzen, die Nachhaltigkeitsberichte zum Erneuerbaren-Energien-Anteil ihrer Rechenzentren vorlegen und regionale oder grüne Rechenzentren betreiben.

Zum Weiterlesen

Zum Weiterlesen empfehle ich:

- Den [Albert-Report „a screen new deal“](#), der sich als „route map of sustainable film production“ sieht
- Den [Albert-Report „Accelerate 2025“](#), der die CO₂e-Emissionen in der UK-Film- und TV-Industrie analysiert
- Die Nachhaltigkeitsberichte von [Google](#), [Microsoft](#) und [Alibaba](#) (Andere Anbieter halten sich hier sehr bedeckt)

Über den Autor

Frank Becher arbeitet als Konzeptionist für edukative Medien und Filmproducer in Berlin. Mit über 25 Jahren Erfahrung und mehr als 350 realisierten Projekten gehört er zu den Pionieren des interaktiven und multimedialen Storytellings in Deutschland. Er hat Lehraufträge für Filmproduktion an der Macromedia Hochschule und der SRH University.

Über Rixfilm

Rixfilm wurde im Mai 2009 von Frank Becher und Angela Poschet in Berlin gegründet. Von Beginn an arbeiten wir in zwei Geschäftsfeldern - E-Learning und Animationsfilm.

Gemeinsam ist beiden Geschäftsfeldern eines: Wir möchten, dass die Projekte unserer Kundinnen und Kunden erfolgreich sind. Unser Schwerpunkt ist die Beratung der Kundinnen und Kunden sowie die umfassende Betreuung und Umsetzung der Projekte. Von der initialen Idee bis zum fertigen Produkt. Hinzu gekommen ist 2024 unser Angebot „Green Animation“. Hier unterstützen wir Filmproduktionen bei der ökologisch nachhaltigen Transformation Ihrer Tätigkeiten.

So erreichen Sie uns:

Rixfilm Ltd. & Co. KG
Bouchéstr. 39
12435 Berlin
Tel. 0179-3916691

www.rixfilm.de

Quellen

Aczel, M., Chamanara, S., Matin, M., Farsi, A., Marwala, T., & Madani, K. (2026). *Environmental cost of AI's energy use: Carbon, water and land footprints*. United Nations University Institute for Water, Environment and Health (UNU-INWEH). <https://doi.org/10.53328/INR26RMA002>

albert, British Film Institute, & Arup. (2026). *A screen new deal: A route map to sustainable film production*. British Film Institute.

Bundesverband Green Film & TV Consultants Deutschland e. V. (n. d.). *Green Producing*. Abgerufen am 29. Juli 2026 von <https://bvgcd.de/green-producing/#tv>

Delavande, J., Pierrard, R., & Luccioni, S. (2025). Video Killed the Energy Budget: Characterizing the Latency and Power Regimes of Open Text-to-Video Models. *arXiv preprint arXiv:2509.19222*.

Ember. (o. D.). *Electricity data explorer*. Abgerufen am 27. Juli 2026 von https://ember-energy.org/data/electricity-data-explorer/?data=co2_intensity&fuel=total&entity=Europe

Elsworth, C., Huang, K., Patterson, D., Schneider, I., Sedivy, R. A., Goodman, S., Townsend, B., Ranganathan, P., Dean, J., Vahdat, A., Gomes, B., & Manyika, J. (2025). *Measuring the environmental impact of delivering AI at Google scale* (arXiv:2508.15734). arXiv. <https://doi.org/10.48550/arXiv.2508.15734>

Gage, G., & Malarkey, B. (2025, 12. November). *Water demands of the data economy*. Roland Berger. <https://www.rolandberger.com/en/Insights/Publications/Water-demands-of-the-data-economy.html>

Jegham, N., Gamazaychikov, B., & Luccioni, S. (2026). Lights, Camera, Carbon: Architectural Scaling Laws for Video Generation Energy Consumption. *arXiv preprint arXiv:2607.04553*.

Schega, C. (2015). *Green-Shooting-Ergebnisbericht – Tatort: Fünf Minuten Himmel*. MFG Medien- und Filmgesellschaft Baden-Württemberg; Zieglerfilm Baden-Baden; Südwestrundfunk. https://greenshooting.mfg.de/files/02_MFG_Filmfoerderung/PDF/tatort_green_shooting_ergebnisbericht.pdf

Berechnung der CO₂-Emissionszahlen

	Kinoproduktion	TV Movie
D _F (Duration / Länge des Films) in Minuten	120	90
R (Resolution / Auflösung)	4K (2160p)	2K (1080p)
S (Shots pro Film)	2400	1000
I (Iterationen)	100	50
E _{mix} (Strommix-Emissionen) in kg CO ₂ e/kWh	0,386 (USA)	0,286 (Europa)
E _{S1} (Emission pro 8s-Shot 720p) Veo 3 in kWh	0,308	0,308
E _{S2} (Emission pro 8s-Shot 720p) Sora 2 pro in kWh	0,4185	0,4185

Angenommene Abhängigkeiten:

- Die Zahl der Iteration wirkt sich linear auf die Gesamtemission aus.
- Die Dauer der Shots und die Auflösung wirken sich quadratisch auf die Gesamtemission aus

Schritt 1: Berechnung der durchschnittlichen Shotlänge D_s in Sekunden:

$$D_s = \frac{(D_F * 60)}{S}$$

Schritt 2: Berechnung Gesamtemission:

$$E = \left(\frac{D_s}{8}\right)^2 * \left(\frac{R}{720}\right)^2 * I * E_{mix} * E_S * S$$

Beispielrechnung für Kino / Sora 2 pro / USA

$$E = \left(\frac{3}{8}\right)^2 * \left(\frac{2160}{720}\right)^2 * 100 * 0,386 * 0,4185 * 2400 \approx 49.000 \text{ kg}$$