

# **Der Energiehunger der Künstlichen Intelligenz: Eine globale Analyse und Prognose des Strombedarfs bis 2035**

## **1. Zusammenfassung für Entscheidungsträger**

Die rasante Entwicklung und globale Implementierung von Künstlicher Intelligenz (KI) markiert einen Wendepunkt für die Weltwirtschaft und stellt gleichzeitig eine fundamentale Herausforderung für das globale Energiesystem dar. Diese Analyse prognostiziert den zusätzlichen Strombedarf, der durch KI-Anwendungen in den nächsten zehn Jahren entsteht, und bewertet die strategischen Implikationen für Energieinfrastruktur, Umwelt und geopolitischen Wettbewerb.

Der globale Stromverbrauch von Rechenzentren, der Haupttreiber des Energiebedarfs von KI, wird sich von einem Niveau von etwa 415-460 Terawattstunden (TWh) in den Jahren 2022-2024 auf über 1.000 TWh bis zum Jahr 2030 mehr als verdoppeln.<sup>1</sup> Dieser prognostizierte Verbrauch im Jahr 2030 übersteigt den gesamten jährlichen Strombedarf Japans.<sup>2</sup> KI ist der entscheidende Katalysator für dieses Wachstum. Während KI-Anwendungen heute schätzungsweise 10-20 % des Stromverbrauchs von Rechenzentren ausmachen, wird ihr Anteil bis 2027 voraussichtlich auf 27 % ansteigen und den Großteil des zukünftigen Wachstums bestimmen.<sup>4</sup>

Geografisch konzentriert sich diese Energienachfrage auf drei Schlüsselregionen: Nordamerika (insbesondere die USA), China und Europa. Zusammen sind diese Regionen für über 85 % des aktuellen Verbrauchs von Rechenzentren verantwortlich.<sup>6</sup> Parallel dazu entwickelt sich Südostasien, angeführt von Malaysia, zu einem neuen Hyper-Wachstumszentrum, das durch strategische Investitionen und geografische Vorteile angetrieben wird.<sup>8</sup>

Gleichzeitig findet ein technologischer Wettlauf um Effizienz statt. Spezialisierte Chips und optimierte KI-Modelle führen zu einer drastischen Reduzierung des Energieverbrauchs pro einzelner Anfrage.<sup>10</sup> Dieser Fortschritt wird jedoch durch das exponentielle Wachstum der

KI-Nutzung und die zunehmende Komplexität der Modelle auf globaler Ebene überkompenziert. Dieses Phänomen, bekannt als Jevons-Paradoxon, führt dazu, dass der Gesamtenergieverbrauch trotz massiver Effizienzgewinne weiter stark ansteigt.

Die strategischen Konsequenzen sind tiefgreifend. Der KI-Boom erzwingt einen fundamentalen Wandel im Energiesektor und erfordert massive Investitionen in neue Erzeugungskapazitäten – insbesondere aus erneuerbaren Quellen, Erdgas und Kernkraft – sowie in die Modernisierung und den Ausbau der Netzinfrastruktur.<sup>5</sup> Der Zugang zu erschwinglicher, zuverlässiger und zunehmend kohlenstoffärmer Energie wird zu einem entscheidenden geopolitischen und wirtschaftlichen Wettbewerbsvorteil im 21. Jahrhundert. Für Entscheidungsträger in Politik und Wirtschaft ist das Verständnis dieser Dynamik unerlässlich, um die Weichen für eine nachhaltige und wettbewerbsfähige Zukunft im Zeitalter der KI zu stellen.

## 2. Der KI-Energie-Tsunami: Eine globale Prognose bis 2035

Die Einführung generativer KI hat eine Nachfragewelle nach Rechenleistung ausgelöst, die in ihrer Größenordnung beispiellos ist. Diese Nachfrage manifestiert sich direkt in einem rapide ansteigenden Stromverbrauch, der die globalen Energiemärkte nachhaltig verändern wird. Dieser Abschnitt quantifiziert das Ausmaß dieser Entwicklung, indem er verschiedene globale Prognosen analysiert und die treibenden Kräfte hinter dem Wachstum beleuchtet.

### 2.1 Baseline 2024: The Starting Point

Um die zukünftige Entwicklung zu verstehen, ist eine genaue Bestimmung des aktuellen Energieverbrauchs unerlässlich. Im Jahr 2022 belief sich der globale Stromverbrauch von Rechenzentren auf schätzungsweise 460 TWh.<sup>2</sup> Aktuellere Schätzungen für das Jahr 2024 bewegen sich in einer Spanne von 415

TWh<sup>6</sup> bis zu einer Prognose von 536

TWh für 2025.<sup>1</sup> Dieser Wert entspricht etwa 1.5 % bis 2 % des gesamten weltweiten Stromverbrauchs und ist damit bereits heute in einer ähnlichen Größenordnung wie der Energiebedarf der globalen Luftfahrtindustrie.<sup>1</sup>

Der Anteil, der direkt auf KI-Workloads zurückzuführen ist, ist derzeit noch begrenzt, wächst aber exponentiell. Schätzungen für die Jahre 2023 und 2024 verorten den KI-Anteil am Gesamtverbrauch der Rechenzentren bei 10 % bis 20 %.<sup>4</sup> Eine spezifischere Analyse von Goldman Sachs beziffert den Anteil für 2023 auf 14 %.<sup>5</sup> Diese Zahlen verdeutlichen eine entscheidende Dynamik: Während traditionelle IT-Anwendungen und Cloud-Computing nach wie vor den Großteil des Verbrauchs ausmachen, ist KI der mit Abstand wichtigste Wachstumstreiber, der die zukünftige Nachfragekurve dominiert.

## 2.2 The Growth Trajectory: Forecasts to 2030 and Beyond

Die Prognosen für die kommenden Jahre deuten auf eine dramatische Beschleunigung des Energieverbrauchs hin. Verschiedene renommierte Institutionen haben Modelle entwickelt, die zwar in den Details variieren, aber in der Tendenz ein klares Bild zeichnen: eine Verdopplung bis Verdreifachung des Bedarfs innerhalb des nächsten Jahrzehnts.

Die Internationale Energieagentur (IEA), eine der maßgeblichen Quellen auf diesem Gebiet, prognostiziert einen Anstieg des Gesamtverbrauchs von Rechenzentren auf über 1.000 TWh bereits bis zum Jahr 2026.<sup>2</sup> Für das Jahr 2030 wird ein Verbrauch von etwa 945 bis 1.050 TWh erwartet, was dem aktuellen Stromverbrauch Japans entspricht.<sup>6</sup> Im Basisszenario der IEA steigt dieser Wert bis 2035 weiter auf 1.200 bis 1.300 TWh an.<sup>16</sup>

Andere Analysten und Beratungsunternehmen kommen zu ähnlichen, teilweise sogar höheren Schätzungen:

- **Deloitte** prognostiziert für 2030 einen Verbrauch von etwa 1.065 TWh. Gleichzeitig wird eine erhebliche Unsicherheit betont: Sollten die erwarteten Effizienzsteigerungen nicht eintreten, könnte der Verbrauch auf über 1.300 TWh ansteigen; bei optimaler Effizienzentwicklung könnte er bei rund 1.000 TWh liegen.<sup>1</sup>
- **ABI Research** legt eine der aggressivsten Prognosen vor und erwartet bis 2030 einen Verbrauch von 1.479 TWh.<sup>13</sup>
- **Goldman Sachs** prognostiziert einen Anstieg der globalen Leistungsnachfrage von Rechenzentren um 165 % bis 2030 im Vergleich zum Niveau von 2023.<sup>5</sup>

Die Divergenz dieser Prognosen, die eine Spanne von über 500 TWh für das Jahr 2030 abdecken – ein Wert, der den gesamten Jahresstromverbrauch Deutschlands übersteigt<sup>19</sup> –, ist kein Zeichen von statistischer Ungenauigkeit. Vielmehr spiegelt sie eine fundamentale Unsicherheit über das zukünftige Zusammenspiel von drei kritischen Variablen wider: die

Geschwindigkeit der KI-Adaption in der Wirtschaft, die Rate der technologischen Effizienzverbesserungen und die zukünftige Komplexität von KI-Modellen. Die Zukunft des Energieverbrauchs ist keine einzelne, vorhersagbare Linie, sondern ein breiter Korridor von Möglichkeiten. Für politische und wirtschaftliche Entscheidungsträger bedeutet dies, dass Strategien nicht auf eine einzelne Zahl setzen dürfen, sondern robust gegenüber einer Reihe von Szenarien sein müssen, beispielsweise durch flexible Netzplanung und modulare Energieerzeugung.

Spezifisch für KI wird ein noch steileres Wachstum erwartet. Der Stromverbrauch von reinen KI-Rechenzentren allein soll sich laut Deloitte von 2022 bis 2026 verzehnfachen und 90 TWh erreichen.<sup>1</sup> Die IEA geht davon aus, dass sich der Bedarf von KI-optimierten Rechenzentren bis 2030 mehr als vervierfachen wird.<sup>15</sup>

Institution (Quelle)	Prognose 2026 (TWh)	Prognose 2030 (TWh)	Prognose 2035 (TWh)	Wesentliche Annahmen/Trei ber
<b>IEA</b>	>1,000	945 - 1,050	1,200 - 1,300	KI als Haupttreiber, Basis-Szenario mit moderaten Effizienzgewin nen. <sup>2</sup>
<b>Deloitte</b>	681	1,065 (Spanne: 1,000-1,300)	-	Prognose hängt stark von der Realisierung von Effizienzpotenz ialen ab. <sup>1</sup>
<b>ABI Research</b>	-	1,479	-	Hohe Annahmen zur Verbreitung von GenAI und Cloud-Comput ing. <sup>13</sup>
<b>BloombergNE</b>	-	-	1,200	Steigende Leistungsdicht

<b>F</b>				e durch KI-Modelle wie GPT-4. <sup>16</sup>
<b>Goldman Sachs</b>	-	Anstieg der Leistungsnach frage um 165 % (vs. 2023)	-	KI-Anteil am Strombedarf steigt von 14 % (2023) auf 27 % (2027). <sup>5</sup>

## 2.3 Primary Drivers of Demand

Mehrere miteinander verknüpfte Faktoren treiben diesen Energie-Tsunami an:

- **Verbreitung von Generativer KI:** Die Integration von Large Language Models (LLMs) und bildgenerierenden Modellen in alltägliche Anwendungen ist der stärkste Treiber. Eine einzelne KI-Anfrage, wie beispielsweise an ChatGPT, kann bis zu zehnmal mehr Energie verbrauchen als eine herkömmliche Google-Suche, was bei Milliarden von Anfragen täglich zu einem enormen Gesamtverbrauch führt.<sup>20</sup>
- **Steigende Modellkomplexität:** Der Wettlauf um leistungsfähigere KI-Modelle führt zu einer exponentiellen Zunahme des Rechenaufwands. Die für das Training von Spitzenmodellen erforderliche Rechenleistung verdoppelt sich derzeit etwa alle neun Monate.<sup>4</sup> Schätzungen zufolge verbrauchte das Training von GPT-4 40- bis 50-mal mehr Strom als das seines Vorgängers GPT-3.<sup>20</sup>
- **Leistungsdichte der Hardware:** Die Hardware, die KI antreibt, wird immer leistungsfähiger, aber auch energieintensiver. Die Leistungsaufnahme von spezialisierten Grafikprozessoren (GPUs) ist von 400 Watt im Jahr 2022 auf 700 Watt für generative KI-Modelle im Jahr 2023 gestiegen. Für die nächste Chip-Generation werden 1.200 Watt erwartet.<sup>1</sup> Diese Entwicklung erhöht die Leistungsdichte pro Server-Rack in Rechenzentren dramatisch und stellt extreme Anforderungen an Stromversorgung und Kühlung.<sup>1</sup>
- **Investitionsboom:** Angetrieben von der erwarteten Nachfrage, haben sich die globalen Investitionen in den Bau von Rechenzentren seit 2022 fast verdoppelt und erreichten im Jahr 2024 ein Volumen von 500 Milliarden US-Dollar.<sup>7</sup> Dieser massive Ausbau der Infrastruktur schafft die physische Grundlage für den prognostizierten Anstieg des Energieverbrauchs.

### **3. Regionale Hotspots: Analyse des Strombedarfs in Nordamerika, Europa und Asien**

Der globale Trend des steigenden Energiebedarfs durch KI manifestiert sich regional sehr unterschiedlich. Die Nachfrage konzentriert sich auf wenige, hoch entwickelte Märkte, die über die notwendige digitale Infrastruktur und das wirtschaftliche Ökosystem verfügen. Diese Konzentration führt jedoch zu erheblichem Druck auf die lokalen Stromnetze und macht die Energiepolitik zu einem strategischen Faktor im globalen Technologiewettbewerb.

#### **3.1 Nordamerika: Der unangefochtene Spitzenreiter**

Nordamerika, und insbesondere die Vereinigten Staaten, ist das Epizentrum der KI-Revolution und damit auch des damit verbundenen Energieverbrauchs. Im Jahr 2024 entfielen auf die USA allein 45 % des weltweiten Stromverbrauchs von Rechenzentren, was ihre dominante Stellung unterstreicht.<sup>6</sup>

Die Prognosen für die USA sind dramatisch. Der Anteil der Rechenzentren am nationalen Stromverbrauch wird sich voraussichtlich von heute etwa 4 % auf 7.5 % bis 12 % bis zum Jahr 2030 verdreifachen.<sup>16</sup> Eine Analyse von McKinsey prognostiziert einen Anstieg der Leistungsnachfrage von 25 Gigawatt (

GW) im Jahr 2024 auf über 80 GW im Jahr 2030. Dies entspricht einem zusätzlichen jährlichen Strombedarf von etwa 400 TWh – mehr als der gesamte aktuelle Stromverbrauch Frankreichs.<sup>19</sup>

Die wirtschaftlichen Auswirkungen dieser Verschiebung sind tiefgreifend. Bis 2030 wird der US-Wirtschaftssektor voraussichtlich mehr Strom für die Datenverarbeitung und KI verbrauchen als für die Herstellung aller traditionellen energieintensiven Güter wie Aluminium, Stahl, Zement und Chemikalien zusammen.<sup>15</sup> Dieser Wandel signalisiert die Entstehung einer neuen, digitalen Schwerindustrie.

Diese massive und schnelle Nachfragesteigerung stellt die Stromnetze vor immense Herausforderungen. Die Nachfrage konzentriert sich auf wenige Regionen mit bestehender Infrastruktur, wie den PJM-Interconnection-Markt im Osten der USA und den ERCOT-Markt in Texas.<sup>16</sup> Diese geografische Bündelung führt zu lokalen Überlastungen und erheblichen Verzögerungen bei Netzanschlüssen, was den weiteren Ausbau der KI-Infrastruktur zu verlangsamen droht.<sup>6</sup>

### **3.2 Europa: Wachstum trifft auf Regulierung und Netzengpässe**

Europa ist mit einem Anteil von 15 % am globalen Verbrauch im Jahr 2024 der drittgrößte Markt für Rechenzentren.<sup>6</sup> Auch hier wird ein starkes Wachstum erwartet, das jedoch durch eine andere Ausgangslage und spezifische Herausforderungen geprägt ist.

Die Prognosen deuten auf eine mehr als Verdopplung des Bedarfs bis 2030 hin. McKinsey erwartet einen Anstieg des Stromverbrauchs von heute etwa 62 TWh auf über 150 TWh bis 2030.<sup>26</sup> Die IEA prognostiziert ein Wachstum von über 45

TWh, was einer Steigerung von 70 % gegenüber dem heutigen Niveau entspricht.<sup>11</sup> Dieser Anstieg wird den Anteil der Rechenzentren am gesamten europäischen Stromverbrauch von heute ca. 2 % auf etwa 5 % im Jahr 2030 anheben.<sup>26</sup>

Im Gegensatz zu anderen Regionen wird erwartet, dass der zusätzliche Energiebedarf in Europa überwiegend aus kohlenstoffarmen Quellen gedeckt wird. Prognosen der IEA zufolge werden erneuerbare Energien und Kernkraft bis 2030 zusammen 85 % der Stromversorgung für europäische Rechenzentren ausmachen.<sup>28</sup>

Die größte Herausforderung für Europa liegt in der Netzinfrastruktur. Nach 15 Jahren stagnierender oder sogar sinkender Stromnachfrage führt der plötzliche Boom der Rechenzentren zu einem Nachfrageschock. Die Anzahl der Anträge auf Netzanschlüsse ist in den letzten Jahren exponentiell gestiegen, was die Verteilnetzbetreiber an ihre Kapazitätsgrenzen bringt und zu langen Wartezeiten führt.<sup>5</sup>

### **3.3 Asien-Pazifik: Chinas Dominanz und der Aufstieg Südostasiens**

Die Asien-Pazifik-Region ist ein Markt der Extreme, geprägt von der Dominanz Chinas und dem dynamischen Aufstieg neuer Wachstumszentren.

China ist mit einem Anteil von 25 % am globalen Verbrauch im Jahr 2024 der weltweit zweitgrößte Markt.<sup>6</sup> Das Land wird auch in Zukunft ein enormes Wachstum verzeichnen. Die IEA prognostiziert einen Anstieg des Stromverbrauchs von Rechenzentren um etwa 175

TWh bis 2030, was einer Steigerung von 170 % gegenüber dem heutigen Niveau entspricht.<sup>11</sup>

In den etablierten Märkten Japan und Südkorea wird ebenfalls ein starkes Wachstum von etwa

80 % (ca. 15 TWh) bis 2030 erwartet.<sup>11</sup>

Die bemerkenswerteste Entwicklung findet jedoch in Südostasien (ASEAN) statt. Die Region entwickelt sich schnell zu einem neuen globalen Hotspot für Rechenzentren. Die IEA prognostiziert hier eine Verdopplung des Bedarfs bis 2030.<sup>9</sup>

- **Malaysia** erlebt das mit Abstand schnellste Wachstum. Angetrieben durch staatliche Förderung, niedrigere Kosten und die Nähe zu Singapur, soll der Strombedarf der malaysischen Rechenzentren von 9 TWh im Jahr 2024 auf 68 TWh im Jahr 2030 ansteigen. Dies würde bedeuten, dass Rechenzentren im Jahr 2030 unglaubliche 30 % des gesamten nationalen Stromverbrauchs ausmachen würden.<sup>8</sup>
- **Indonesien und die Philippinen** verzeichnen ebenfalls ein explosives Wachstum. Es wird erwartet, dass sich die Emissionen aus dem Stromverbrauch der Rechenzentren in Indonesien vervierfachen und auf den Philippinen sogar um das 14-fache ansteigen werden, was die Herausforderungen für die Energiewende in diesen Ländern verdeutlicht.<sup>8</sup>

Diese regionale Analyse zeigt, dass KI-Infrastruktur nicht gleichmäßig verteilt wird. Sie konzentriert sich in Regionen, die eine Kombination aus verfügbarem Land, robusten Stromnetzen, günstigen regulatorischen Bedingungen und vor allem Zugang zu großen Mengen zuverlässiger und zunehmend kohlenstoffärmer Energie bieten. Dies führt zur Entstehung von "KI-Energie-Oasen" ("AI Power Havens"). Die Energiepolitik eines Landes wird somit zu einem direkten und entscheidenden Faktor für seine Wettbewerbsfähigkeit in der digitalen Weltwirtschaft. Nationen, die in der Lage sind, schnell neue Energieerzeugungskapazitäten aufzubauen und Netzanschlüsse zu rationalisieren, werden die Investitionen in die KI-Infrastruktur der Zukunft anziehen.

Region	Stromverbrauch 2024 (TWh)	Prognose Stromverbrauch 2030 (TWh)	Prognostizierte r Anteil am regionalen Gesamtstromverbrauch 2030 (%)	Wesentliche Treiber/Herausforderungen
<b>Nordamerika (USA)</b>	~187	~587	7.5 - 12	Dominanz der Tech-Giganten, Konzentration in Hubs (PJM, ERCOT), starker Netzdruck. <sup>6</sup>

<b>Europa (EU+UK)</b>	~62	~150	~5	Starke Regulierung, Fokus auf erneuerbare Energien, Netzengpässe nach langer Stagnation. <sup>11</sup>
<b>China</b>	~104	~279	n.a.	Staatlich gefördertes Wachstum, zweitgrößter Markt der Welt, hohe Abhängigkeit von Kohle. <sup>6</sup>
<b>Japan &amp; Südkorea</b>	~19	~34	n.a.	Etablierte Technologiemärkte, Fokus auf Kernkraft und Erneuerbare. <sup>11</sup>
<b>Südostasien (insb. Malaysia)</b>	~18	~123	bis zu 30 % (Malaysia)	Hyper-Wachstum, niedrigere Kosten, Herausforderungen bei der Dekarbonisierung. <sup>8</sup>

## 4. Das Effizienzparadoxon: Kann die Technologie die Nachfrage überholen?

Während die Prognosen ein Bild von unaufhaltsam steigendem Energieverbrauch zeichnen, findet parallel eine technologische Revolution statt, die auf maximale Effizienz abzielt. Auf den

ersten Blick scheint es widersprüchlich: Wie kann der Gesamtverbrauch explodieren, wenn jede einzelne KI-Operation immer weniger Energie benötigt? Die Antwort liegt im sogenannten Jevons-Paradoxon, einem ökonomischen Prinzip, das besagt, dass Effizienzsteigerungen oft zu einem höheren Gesamtverbrauch führen, weil sie die Nutzung einer Ressource billiger und zugänglicher machen.

## 4.1 Hardware-Evolution: Der Wettkampf um Performance-pro-Watt

Der Kern der Effizienzsteigerung liegt in der Entwicklung spezialisierter Hardware. Allgemeine Prozessoren (CPUs) sind für die massiv parallelen Berechnungen, die für KI erforderlich sind, ungeeignet und ineffizient. Der Durchbruch kam mit Grafikprozessoren (GPUs), aber die Entwicklung geht weiter zu hochspezialisierten Chips:

- **Tensor Processing Units (TPUs):** Von Google speziell für KI-Workloads entwickelte Chips (ASICs) sind von Grund auf auf Energieeffizienz optimiert. TPUs können eine 2- bis 3-fach bessere Performance-pro-Watt im Vergleich zu GPUs für bestimmte Aufgaben liefern und die Betriebskosten um 20-30 % senken.<sup>29</sup>
- **Generationensprünge:** Jede neue Chip-Generation bringt Quantensprünge in der Effizienz. Googles neueste TPU-Generation, "Trillium", ist 67 % energieeffizienter als ihr direkter Vorgänger.<sup>30</sup> Der Chiphersteller Nvidia verspricht für seine neuesten "Superchips" eine 30-fache Leistungssteigerung bei gleichzeitig 25-fach geringerem Energieverbrauch für spezifische generative KI-Dienste.<sup>22</sup> Neue Wettbewerber wie das Startup Positron zielen sogar auf eine 3- bis 6-fache Verbesserung der Energieeffizienz im Vergleich zu Nvidias kommender Chip-Generation ab.<sup>31</sup>

## 4.2 Software- und Modelloptimierung: Die unsichtbaren Einsparungen

Neben der Hardware sind Fortschritte bei Software und Algorithmen ein entscheidender Hebel zur Reduzierung des Energieverbrauchs. Durch Techniken der Modellkompression können riesige KI-Modelle auf eine handhabbare Größe reduziert werden, ohne signifikante Leistungseinbußen.

- **Quantisierung (Quantization):** Diese Technik reduziert die numerische Präzision der Modellparameter, z.B. von 32-Bit-Gleitkommazahlen (FP32) auf 8-Bit-Ganzzahlen (INT8). Dies kann die Modellgröße um bis zu 75 % reduzieren, den Speicherbedarf drastisch senken und die Inferenz (die Anwendung des Modells) erheblich beschleunigen.<sup>32</sup>
- **Pruning:** Hierbei werden redundante oder unwichtige Verbindungen (Neuronen) aus

einem neuronalen Netz entfernt. Dies ist vergleichbar mit dem Stutzen eines Baumes, um das Wachstum zu fördern. Mit dieser Methode kann die Modellgröße um das 10-fache reduziert werden, bei einem Genauigkeitsverlust von weniger als 1 %.<sup>32</sup>

- **Wissensdestillation (Knowledge Distillation):** Bei diesem Ansatz wird das "Wissen" eines großen, komplexen "Lehrer"-Modells auf ein kleineres, effizienteres "Schüler"-Modell übertragen. Ein prominentes Beispiel ist DistilBERT, das 97 % der sprachlichen Fähigkeiten des ursprünglichen BERT-Modells beibehält, dabei aber 40 % weniger Parameter benötigt und 60 % schneller ist.<sup>32</sup>

### 4.3 Der Nettoeffekt: Das Jevons-Paradoxon in der KI

Die Kombination dieser Hardware- und Software-Optimierungen führt dazu, dass der Energieverbrauch pro einzelner KI-Anfrage dramatisch sinkt. Google berichtet, dass der Energiebedarf für eine mediane Anfrage an sein Gemini-Modell innerhalb von nur 12 Monaten um den Faktor 33 gesunken ist.<sup>10</sup> Die von OpenAI-CEO Sam Altman genannte Zahl von 0.34 Wattstunden (

Wh) pro ChatGPT-Anfrage ist, obwohl in der Fachwelt noch diskutiert, ein Beleg für diesen Trend zu bemerkenswert niedrigen Pro-Anfrage-Kosten.<sup>33</sup>

Trotz dieser beeindruckenden Fortschritte auf der Mikroebene steigt der Gesamtverbrauch auf der Makroebene unaufhaltsam. Dies ist die zentrale Dynamik, die den Energiehunger der KI bestimmt. Die Effizienzgewinne machen KI-Anwendungen billiger, schneller und breiter verfügbar. Dies führt zu einer explosionsartigen Zunahme der Nutzung: Mehr Nutzer stellen mehr Anfragen, und KI wird in immer mehr Produkte und Dienstleistungen integriert. Gleichzeitig ermöglichen es die Effizienzgewinne den Entwicklern, noch größere und komplexere Modelle zu trainieren (z.B. für die Videoerzeugung), die den Energieverbrauch pro Anfrage wieder in die Höhe treiben.

Die Effizienzsteigerungen kompensieren das Wachstum also nicht, sondern sie ermöglichen es erst. Sich allein auf technologische Effizienz zu verlassen, um das Energieproblem der KI zu lösen, ist daher eine fehlerhafte Strategie. Politische und unternehmerische Ansätze müssen sich auch auf das Management des Nachfragewachstums konzentrieren, beispielsweise durch die Förderung kleinerer, aufgabenspezifischer Modelle und die Schaffung von Anreizen für einen nachhaltigen Einsatz von KI.

### 5. Die Zahlen im Kontext: Der Energie-Fußabdruck der

# KI im Vergleich

Um die abstrakten Zahlen von Terawattstunden greifbar zu machen, ist ein Vergleich mit bekannten Größen aus Wirtschaft und Alltag unerlässlich. Diese Vergleiche verdeutlichen die schiere Größenordnung des Energiebedarfs der KI und untermauern die These, dass wir Zeugen der Entstehung einer neuen globalen, energieintensiven Industrie sind.

## 5.1 Vergleich mit Volkswirtschaften

Der prognostizierte Stromverbrauch von Rechenzentren im Jahr 2030, der sich auf etwa 945 bis 1.050 TWh beläuft, stellt viele nationale Energiemarkte in den Schatten:

- Dieser Wert ist **größer als der gesamte jährliche Stromverbrauch Japans**, der 2023 bei rund 1.013 TWh lag.<sup>6</sup>
- Bereits bis 2026 könnte der Verbrauch von Rechenzentren dem von Volkswirtschaften wie **Deutschland** (ca. 506 TWh im Jahr 2023) oder Schweden entsprechen.<sup>4</sup>
- Der prognostizierte **Zuwachs** des Strombedarfs allein in China bis 2026 (ca. 1.400 TWh) ist mehr als die Hälfte des gesamten Jahresverbrauchs der Europäischen Union.<sup>2</sup>

## 5.2 Vergleich mit Industriezweigen

Traditionell gelten Industrien wie die Stahl- oder Zementherstellung als die größten Energieverbraucher. Der KI-Sektor ist im Begriff, in diese Liga aufzusteigen und sie in einigen Regionen sogar zu übertreffen.

- **Traditionelle Schwerindustrie:** In den USA wird der Sektor der Rechenzentren bis 2030 mehr Strom verbrauchen als die gesamte energieintensive Fertigungsindustrie (Aluminium, Stahl, Zement, Chemie) zusammen.<sup>15</sup> Die US-Stahlindustrie benötigt für ihre vollständige Dekarbonisierung bis 2050 schätzungsweise 174 TWh pro Jahr – ein Wert, den der jährliche Bedarf der US-Rechenzentren bereits um 2030 weit übersteigen wird.<sup>25</sup>
- **Kryptowährungen:** Das Bitcoin-Mining, oft als energieintensiv kritisiert, verbraucht jährlich etwa 160 TWh.<sup>36</sup> Der prognostizierte Verbrauch von Rechenzentren im Jahr 2030 wird diesen Wert um das Sechs- bis Siebenfache übertreffen.
- **Luftfahrtindustrie:** Wie bereits erwähnt, wird der heutige Stromverbrauch von

Rechenzentren (1-2 % des weltweiten Verbrauchs) oft mit dem der globalen Luftfahrtindustrie verglichen, was seine systemische Bedeutung unterstreicht.<sup>12</sup>

Diese Vergleiche signalisieren einen fundamentalen wirtschaftlichen Wandel. Die Tech-Industrie, lange als "saubere" digitale Wirtschaft wahrgenommen, entwickelt sich in Teilen zu einem Energieverbraucher im industriellen Maßstab. Ein einzelner großer Rechenzentrumscampus hat heute einen ähnlichen Leistungsbedarf wie eine Aluminiumhütte.<sup>17</sup> Diese Neuausrichtung hat weitreichende Konsequenzen für die Industriepolitik, die Umweltregulierung und die Anlagestrategien.

### 5.3 Vergleich im Alltag

Um die Dimensionen auf die Mikroebene herunterzubrechen, helfen Vergleiche mit alltäglichen Vorgängen:

- **Eine einzelne KI-Anfrage:** Eine typische Anfrage an ChatGPT verbraucht etwa 0.34 Wh.<sup>33</sup> Das ist eine winzige Menge Energie, vergleichbar mit dem Betrieb einer LED-Lampe für wenige Minuten<sup>37</sup> oder dem einmaligen Aufladen eines Smartphones.<sup>12</sup> Die schiere Masse von Milliarden Anfragen pro Tag führt jedoch zu einem erheblichen Gesamtverbrauch.
- **Ein KI-Rechenzentrum:** Ein typisches, auf KI spezialisiertes Rechenzentrum kann so viel Strom verbrauchen wie 100.000 Haushalte. Die größten derzeit im Bau befindlichen Anlagen werden das 20-fache davon benötigen und damit den Energiebedarf einer mittelgroßen Stadt erreichen.<sup>7</sup>

Entität	Jährlicher Stromverbrauch (TWh)	Vergleichsgröße	Quellen
<b>Prognose Rechenzentren 2030</b>	<b>~1,000</b>	Größer als der gesamte Stromverbrauch Japans.	6
<b>Japan (2023)</b>	<b>1,013</b>	Einer der größten Stromverbraucher der Welt.	19

<b>Deutschland (2023)</b>	506	Größte Volkswirtschaft Europas.	19
<b>Bitcoin-Netzwerk (jährlich)</b>	~160	Oft als Beispiel für hohen Energieverbrauch genannt.	36
<b>US-Stahlindustrie (Dekarbonisierungsbedarf 2050)</b>	174	Repräsentiert eine traditionelle Schwerindustrie.	35
<b>Globale Luftfahrtindustrie</b>	Vergleichbar	Oft als Benchmark für den Anteil am globalen Energieverbrauch genutzt.	12

## 6. Strategische Implikationen und Ausblick

Der durch KI angetriebene Anstieg des Strombedarfs ist mehr als eine technische Herausforderung; er ist ein strategischer Faktor, der die globale Energieinfrastruktur, die Umweltpolitik und die geopolitische Landschaft neu gestaltet. Die Fähigkeit, diesen Bedarf sicher, nachhaltig und kostengünstig zu decken, wird über die technologische und wirtschaftliche Führungsrolle im 21. Jahrhundert entscheiden.

### 6.1 Auswirkungen auf die Energieinfrastruktur

Die Deckung der prognostizierten Nachfrage erfordert einen massiven und schnellen Ausbau der gesamten Energiewertschöpfungskette.

- **Investitionsbedarf:** Die erforderlichen Investitionen sind enorm. Allein in Europa werden bis 2030 Investitionen in die Infrastruktur von Rechenzentren (ohne die Kosten für die Stromerzeugung) auf 250 Milliarden US-Dollar geschätzt.<sup>27</sup> In den USA könnten bis 2030

zusätzliche Netzinvestitionen in Höhe von 720 Milliarden US-Dollar erforderlich sein, um die Stabilität zu gewährleisten.<sup>5</sup>

- **Transformation des Energiemixes:** Der zusätzliche Strom wird aus einem diversifizierten Portfolio von Energiequellen stammen müssen. Die IEA prognostiziert, dass bis 2035 der Zuwachs hauptsächlich durch erneuerbare Energien (zusätzlich +480 TWh), Erdgas (+180 TWh) und Kernkraft (+190 TWh) gedeckt wird.<sup>17</sup> Insbesondere kleine modulare Reaktoren (Small Modular Reactors, SMRs) rücken in den Fokus der Tech-Industrie als Quelle für zuverlässige, kohlenstoffarme Grundlastenergie.<sup>3</sup>
- **Netzengpässe als Nadelöhr:** Die größte unmittelbare Herausforderung ist nicht die Erzeugungskapazität, sondern die Übertragungs- und Verteilungsinfrastruktur. Die IEA schätzt, dass sich fast 20 % der weltweit geplanten Rechenzentrumsprojekte aufgrund von Engpässen beim Netzanschluss verzögern könnten.<sup>6</sup> Die Wartezeiten für kritische Netzkomponenten wie Transformatoren haben sich in den letzten Jahren verdoppelt.

## 6.2 Ökologische und geopolitische Dimension

Der Energiehunger der KI hat weitreichende ökologische und geopolitische Konsequenzen.

- **CO2-Emissionen:** Obwohl der Stromverbrauch massiv ansteigt, bleiben die direkten CO2-Emissionen im globalen Kontext relativ begrenzt. Die IEA schätzt einen Anstieg der Emissionen aus dem Stromverbrauch von Rechenzentren von 220 Millionen Tonnen (Mt) CO2 im Jahr 2024 auf 300-320 Mt CO2 bis 2035. Dies entspricht weniger als 1.2 % der gesamten energiebedingten Emissionen.<sup>17</sup> Dennoch handelt es sich um eine der am schnellsten wachsenden Emissionsquellen im Energiesektor, was den Druck zur Dekarbonisierung der Stromerzeugung weiter erhöht.<sup>17</sup>
- **Wasserverbrauch:** Ein oft übersehener, aber kritischer Faktor ist der immense Wasserverbrauch, der sowohl für die direkte Kühlung der Rechenzentren als auch für die Stromerzeugung (insbesondere bei thermischen Kraftwerken) benötigt wird. Prognosen deuten auf eine mögliche Verzehnfachung des Wasserverbrauchs bis 2028 hin.<sup>38</sup> Verschärft wird dieses Problem dadurch, dass viele große Rechenzentrums-Hubs bereits heute in Regionen mit mittlerem bis hohem Wasserstress liegen.<sup>38</sup>
- **Geopolitik der Energie:** Der Zugang zu Energie wird zu einem zentralen Faktor für die digitale Souveränität. Der Wettbewerb zwischen den USA, China und Europa um die Vorherrschaft bei KI wird zunehmend auch ein Wettbewerb um den Zugang zu sicherer, sauberer und kostengünstiger Energie. Die Fähigkeit, die Energieinfrastruktur schnell und effizient auszubauen, wird zu einem entscheidenden nationalen Vorteil.

## 6.3 Die duale Rolle der KI: Problem und Lösung zugleich

Die vielleicht faszinierendste und strategisch wichtigste Eigenschaft der KI ist ihre duale Rolle als massiver Energieverbraucher und gleichzeitig als potenziell wirksamstes Werkzeug zur Optimierung des Energiesystems.

- **KI zur Steigerung der Energieeffizienz:** KI-Anwendungen können den Energieverbrauch in der gesamten Wirtschaft drastisch senken. Sie können den Stromverbrauch in Gebäuden optimieren (potenzielle globale Einsparungen von bis zu 300 TWh), industrielle Prozesse effizienter gestalten und die Auslastung bestehender Stromnetze verbessern, was die Notwendigkeit teurer Netzausbauten reduzieren kann (potenzielle Kapazitätsfreisetzung von bis zu 175 GW).<sup>6</sup>
- **Die Netto-Bilanz als offene Frage:** Die entscheidende Frage für die Zukunft ist, ob die durch KI ermöglichten Energieeinsparungen in der Gesamtwirtschaft ihren eigenen, direkten Energieverbrauch ausgleichen oder sogar übertreffen können. Die Antwort darauf ist noch offen und hängt von der Geschwindigkeit und dem Umfang der Implementierung ab. Eine Studie von PwC deutet darauf hin, dass der Nettoeffekt energie-neutral oder sogar leicht positiv sein könnte, wenn die Effizienzgewinne in der breiteren Wirtschaft realisiert werden.<sup>39</sup> Das Weltwirtschaftsforum schätzt, dass KI das Potenzial hat, die globalen Treibhausgasemissionen um 5-10 % zu senken.<sup>40</sup>

Diese Analyse zeigt zwei starke, gegenläufige Entwicklungen: KI als Klimarisiko aufgrund ihres Energiehunders und KI als Klimalösung aufgrund ihres Optimierungspotenzials. Der endgültige Ausgang ist nicht vorbestimmt. Er hängt von den strategischen Entscheidungen ab, die heute getroffen werden. Eine erfolgreiche Strategie muss zweigleisig sein: Sie muss den direkten ökologischen Fußabdruck der KI durch den Ausbau erneuerbarer Energien und Effizienzsteigerungen minimieren und gleichzeitig die Anwendung von KI zur Dekarbonisierung und Effizienzsteigerung in der gesamten Wirtschaft maximieren. Die Unternehmen und Nationen, die diese duale Strategie beherrschen – also KI nutzen, um die von KI geschaffenen Probleme zu lösen –, werden einen entscheidenden Wettbewerbs- und Umweltvorteil erlangen. Die "Kohlenstoffbilanz" der KI ist letztlich eine Gestaltungsaufgabe, kein unabwendbares Schicksal.

## Referenzen

1. As generative AI asks for more power, data centers seek more reliable, cleaner energy solutions - Deloitte, Zugriff am September 12, 2025, <https://www.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions/2025/genai-power-consumption-creates-need-for-more-sustainable-data-centers.html>
2. Executive summary – Electricity 2024 – Analysis - IEA, Zugriff am September 12, 2025, <https://www.iea.org/reports/electricity-2024/executive-summary>

3. AI surge to double data centre electricity demand by 2030: IEA | The Straits Times, Zugriff am September 12, 2025,  
<https://www.straitstimes.com/world/europe/ai-surge-to-double-data-centre-electricity-demand-by-2030-iea>
4. How AI Is Fueling a Boom in Data Centers and Energy Demand - Time Magazine, Zugriff am September 12, 2025,  
<https://time.com/6987773/ai-data-centers-energy-usage-climate-change/>
5. AI to drive 165% increase in data center power demand by 2030 | Goldman Sachs, Zugriff am September 12, 2025,  
<https://www.goldmansachs.com/insights/articles/ai-to-drive-165-increase-in-data-center-power-demand-by-2030>
6. Data center energy consumption will double by 2030: more than 450 TWh of additional renewable energy will be required to sustain its growth, Zugriff am September 12, 2025, <https://strategicenergy.eu/data-center/>
7. Executive summary – Energy and AI – Analysis - IEA, Zugriff am September 12, 2025, <https://www.iea.org/reports/energy-and-ai/executive-summary>
8. From AI to emissions: Aligning ASEAN's digital growth with energy transition goals - Ember, Zugriff am September 12, 2025,  
<https://ember-energy.org/app/uploads/2025/05/Report-From-AI-to-emissions-ASEAN.pdf>
9. ASEAN's data centres electricity demand keeps growing - Ember, Zugriff am September 12, 2025, <https://ember-energy.org/chapter/setting-the-scene/>
10. Measuring the environmental impact of AI inference | Google Cloud Blog, Zugriff am September 12, 2025,  
<https://cloud.google.com/blog/products/infrastructure/measuring-the-environmental-impact-of-ai-inference/>
11. Energy demand from AI – Energy and AI – Analysis - IEA, Zugriff am September 12, 2025, <https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai>
12. AI has high data center energy costs — but there are solutions | MIT ..., Zugriff am September 12, 2025,  
<https://mitsloan.mit.edu/ideas-made-to-matter/ai-has-high-data-center-energy-costs-there-are-solutions>
13. Data Center Energy Consumption Forecast, 2024-2030, Zugriff am September 12, 2025,  
<https://www.abiresearch.com/blog/data-center-energy-consumption-forecast>
14. www.goldmansachs.com, Zugriff am September 12, 2025,  
[https://www.goldmansachs.com/insights/articles/ai-to-drive-165-increase-in-data-center-power-demand-by-2030#:~:text=At%20present%2C%20Goldman%20Sachs%20Research,%2C%20and%20AI%20\(14%25\).](https://www.goldmansachs.com/insights/articles/ai-to-drive-165-increase-in-data-center-power-demand-by-2030#:~:text=At%20present%2C%20Goldman%20Sachs%20Research,%2C%20and%20AI%20(14%25).)
15. AI is set to drive surging electricity demand from data centres while offering the potential to transform how the energy sector works - News - IEA, Zugriff am September 12, 2025,  
<https://www.iea.org/news/ai-is-set-to-drive-surging-electricity-demand-from-data-centres-while-offering-the-potential-to-transform-how-the-energy-sector-works>

16. Power for AI: Easier Said Than Built | BloombergNEF, Zugriff am September 12, 2025,  
<https://about.bnef.com/insights/commodities/power-for-ai-easier-said-than-built/>
17. AI's Energy Hunger: Data Centers Set to Use Power Equal to Japan's by 2035, Zugriff am September 12, 2025,  
<https://carboncredits.com/ais-energy-hunger-data-centers-set-to-use-power-equal-to-japans-by-2035/>
18. How AI is Transforming Data Centers and Ramping Up Power Demand - Goldman Sachs, Zugriff am September 12, 2025,  
<https://www.goldmansachs.com/insights/articles/how-ai-is-transforming-data-centers-and-ramping-up-power-demand>
19. List of countries by electricity consumption - Wikipedia, Zugriff am September 12, 2025, [https://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_electricity\\_consumption](https://en.wikipedia.org/wiki/List_of_countries_by_electricity_consumption)
20. The AI Energy Crisis: A Looming Threat to Sustainability and Tech's Green Ambitions, Zugriff am September 12, 2025,  
<https://markets.financialcontent.com/stocks/article/marketminute-2025-9-11-the-ai-energy-crisis-a-looming-threat-to-sustainability-and-techs-green-ambitions>
21. AI and the Net-Zero Journey: Energy Demand, Emissions, and the Potential for Transition, Zugriff am September 12, 2025, <https://arxiv.org/html/2507.10750v1>
22. AI and energy: Will AI reduce emissions or increase power demand?, Zugriff am September 12, 2025,  
<https://www.weforum.org/stories/2024/07/generative-ai-energy-emissions/>
23. AI power: Expanding data center capacity to meet growing demand - McKinsey, Zugriff am September 12, 2025,  
<https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand>
24. 215 Data Center Stats (June-2025) - Brightlio, Zugriff am September 12, 2025,  
<https://brightlio.com/data-center-stats/>
25. How data centers and the energy sector can sate AI's hunger for power - McKinsey, Zugriff am September 12, 2025,  
<https://www.mckinsey.com/industries/private-capital/our-insights/how-data-centers-and-the-energy-sector-can-sate-ais-hunger-for-power>
26. The role of power in unlocking the European AI revolution - McKinsey, Zugriff am September 12, 2025,  
<https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/the-role-of-power-in-unlocking-the-european-ai-revolution>
27. Europe's data centre power demands plus other technology news to know, Zugriff am September 12, 2025,  
<https://www.weforum.org/stories/2024/11/europe-data-centre-plus-other-technology-news-to-know/>
28. Energy supply for AI – Energy and AI – Analysis - IEA, Zugriff am September 12, 2025, <https://www.iea.org/reports/energy-and-ai/energy-supply-for-ai>
29. Energy-Efficient GPU vs. TPU Allocation - Artech Digital, Zugriff am September 12,

2025,

<https://www.artech-digital.com/blog/energy-efficient-gpu-vs-tpu-allocation>

30. What's the difference between CPUs, GPUs and TPUs? - Google Blog, Zugriff am September 12, 2025,  
<https://blog.google/technology/ai/difference-cpu-gpu-tpu-trillium/>
31. Next-Gen AI Chips Developed to Cut Energy Use - The National CIO Review, Zugriff am September 12, 2025,  
<https://nationalcioreview.com/articles-insights/extra-bytes/next-gen-ai-chips-developed-to-cut-energy-use/>
32. How to Reduce AI Energy Costs with Model Optimization, Zugriff am September 12, 2025,  
<https://www.artech-digital.com/blog/how-to-reduce-ai-energy-costs-with-model-optimization>
33. www.devsustainability.com, Zugriff am September 12, 2025,  
<https://www.devsustainability.com/p/chatgpt-energy-usage-is-034-wh-per#:~:text=People%20are%20often%20curious%20about,one%20fifteenth%20of%20a%20teaspoon.>
34. Let's Analyze OpenAI's Claims About ChatGPT Energy Use - Towards Data Science, Zugriff am September 12, 2025,  
<https://towardsdatascience.com/lets-analyze-openais-claims-about-chatgpt-energy-use/>
35. CEBA Report: Cutting Steel Industry Emissions Will Require 174 Terawatt Hours of Electricity Annually by 2050 - Clean Energy Buyers Association, Zugriff am September 12, 2025,  
<https://cebuyers.org/blog/ceba-report-cutting-steel-industry-emissions-will-require-174-terawatt-hours-of-electricity-annually-by-2050/>
36. 61 Bitcoin Energy Consumption Statistics (2024), Zugriff am September 12, 2025,  
<https://buybitcoinworldwide.com/bitcoin-mining-statistics/>
37. How much energy does ChatGPT use? - Epoch AI, Zugriff am September 12, 2025, <https://epoch.ai/gradient-updates/how-much-energy-does-chatgpt-use>
38. AI boom may drain resources: Data centres' water use could hit 1,068 billion litres by 2028; Morgan Stanley report flags 11x rise, Zugriff am September 12, 2025,  
<https://timesofindia.indiatimes.com/business/international-business/ai-boom-may-drain-resources-data-centres-water-use-could-hit-1068-billion-litres-by-2028-morgan-stanley-report-flags-11x-rise/articleshow/123758252.cms>
39. AI Could Soon Offset Its Own Environmental Impact By Improving Energy Efficiency, Report Finds - TriplePundit, Zugriff am September 12, 2025,  
<https://triplepundit.com/2025/sustainable-ai-energy-use/>
40. Energy and AI: the power couple that could usher in a net-zero world, Zugriff am September 12, 2025,  
<https://www.weforum.org/stories/2025/01/energy-ai-net-zero/>