IT und KI.

Durch die Steigerung von Energieeffizienz und Ressourcenmanagement Nachhaltigkeitsziele erreichen

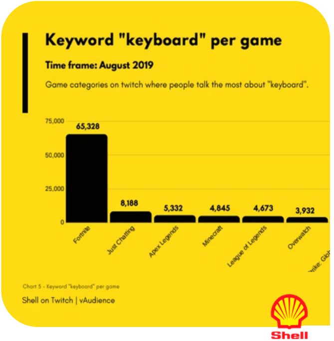
18.09.2024

Forum Energie 2024

Dr. Toni Wagner

- 1977 in Bamberg geboren
- 1986 erster Programmierspaß mit dem C64
- 1997-2002: Studium Molekular-Biologie
- 2002-2007: PhD Signal-Quantification in StemCells
- 2006-2011: Group-Leader Signalquantifikation & Stammzellforschung Universität Würzburg
- 2010-2014: myLinkCloud (IT-Startup) Founder&CEO
- 2011-2016: Co-Founder & CTO meinUnterricht.de
- seit 2016: Founder & CEO vAudience GmbH
- seit 2016: KI-Projekte (mostly machine-learning)
- seit 2022: vA Fokus auf KI-Automation & Multi-Agenten-Systeme

Seit 2016 Beispielprojekte mit KI (Machine Learning)



TEXT (~2017-2019)

Nutzung von über 1Mrd Chat-Nachrichten pro Monat um Persona-, Sentiment und Themen-Analysen auf Basis für Kunden durchzuführen

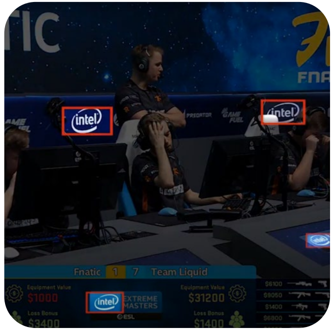
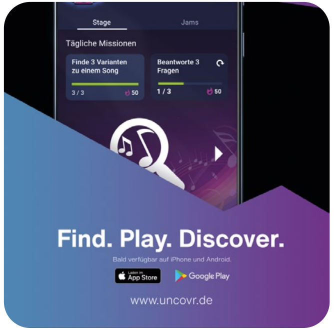


BILD (~2018-2019)

Quantifizierung der Reichweite und "Sichtbarkeit" von Objekten (z.B: Logos) in Live-Videos



AUDIO (2020-2021)

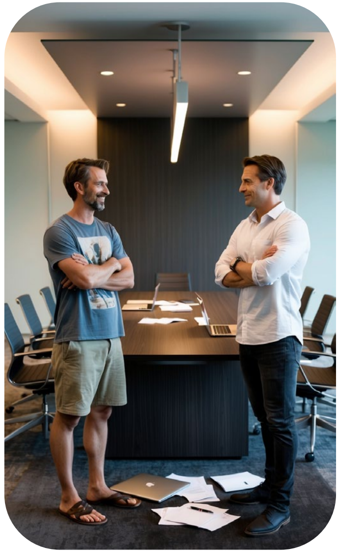
Konzeption und Umsetzung eines gamifizierten Labelling-Systems und Training eines Modells für Audio-Erkennung

MASCHINENBAU (2024)

Auswertung von Sensordaten in Fertigungsprozessen um Event-Vorhersage zu ermöglichen

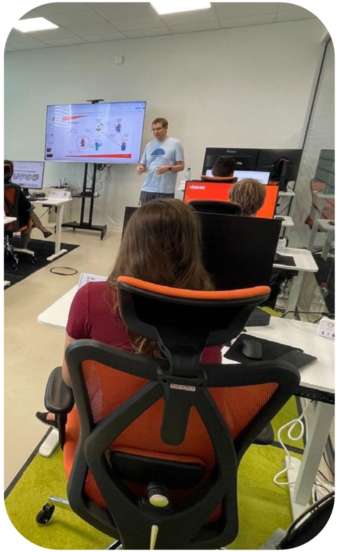
vAudience Customer Journey

1



Erstgespräch

2



KI-Kurs

3

Workshop/ Prozess-begleitung

4

Handlungs-empfehlung

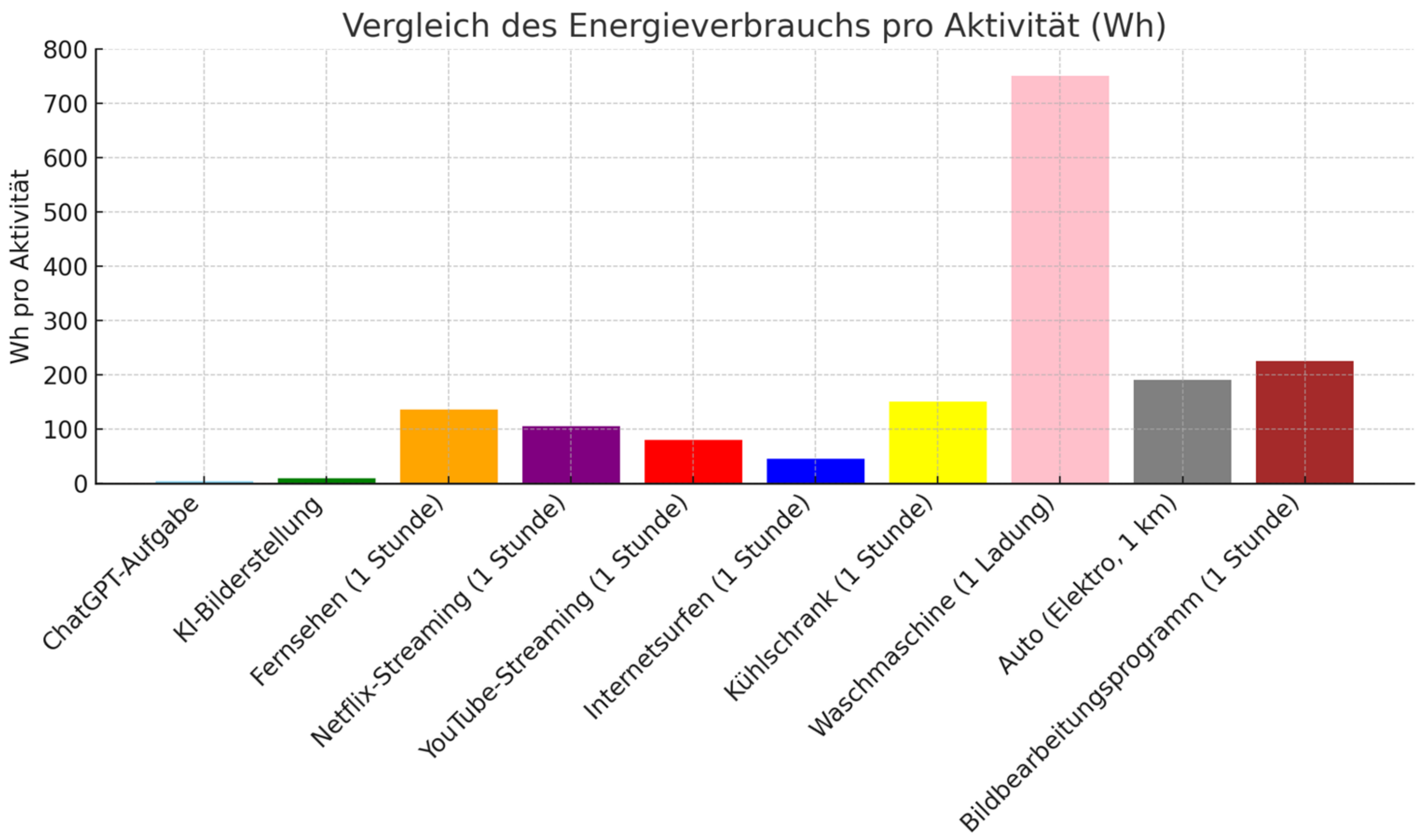
5

Produkt-Umsetzung

Intelligenz braucht Energie

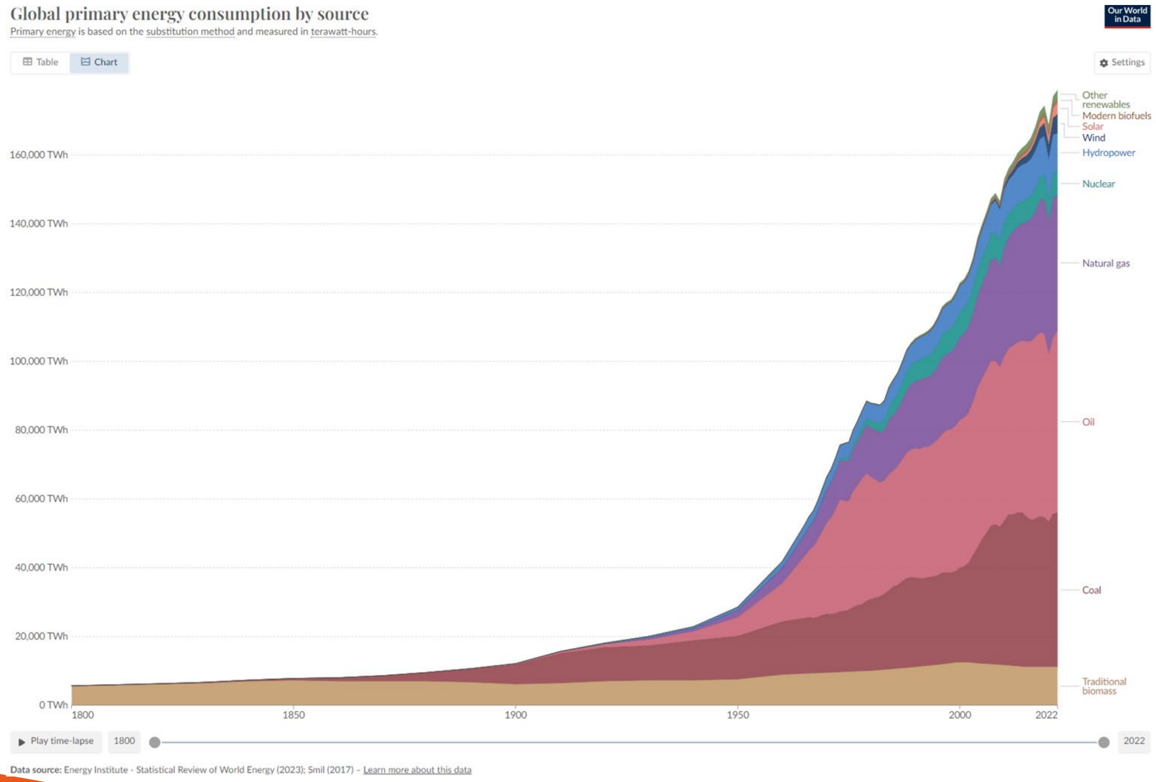
Spezies	Energieverbrauch des Gehirns (% des Grundumsatzes)	Energieverbrauch des Gehirns (kWh/Tag)	Neuronen im Gehirn
Mensch	20-25%	~0.34	~86 Mrd
Ratte	3-5%	~0.0025	~200 Mio
Maus	3-5%	~0.0004	~71 Mio
Katze	5-10%	~0.015	~250 Mio
Hund	5-10%	~0.028	~530 Mio
Schimpanse	8-10%	~0.120	~28 Mrd
Delfin	5-15%	~0.245	~5.8 Mrd
Elefant	0.5-0.7%	~0.4	~257 Mrd

Energieverbrauch im Alltag

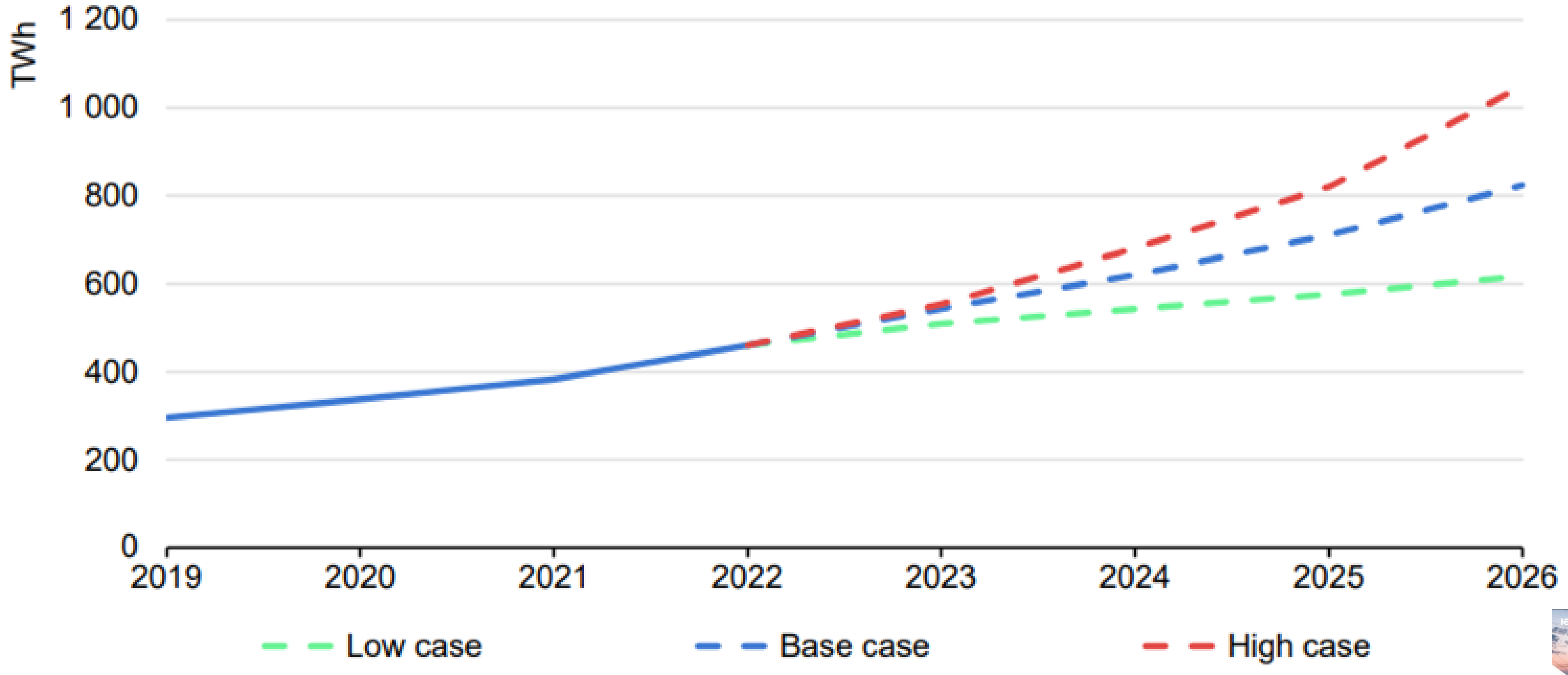


KI braucht (mehr) Energie

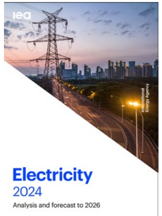
Weltweiter Energieverbrauch



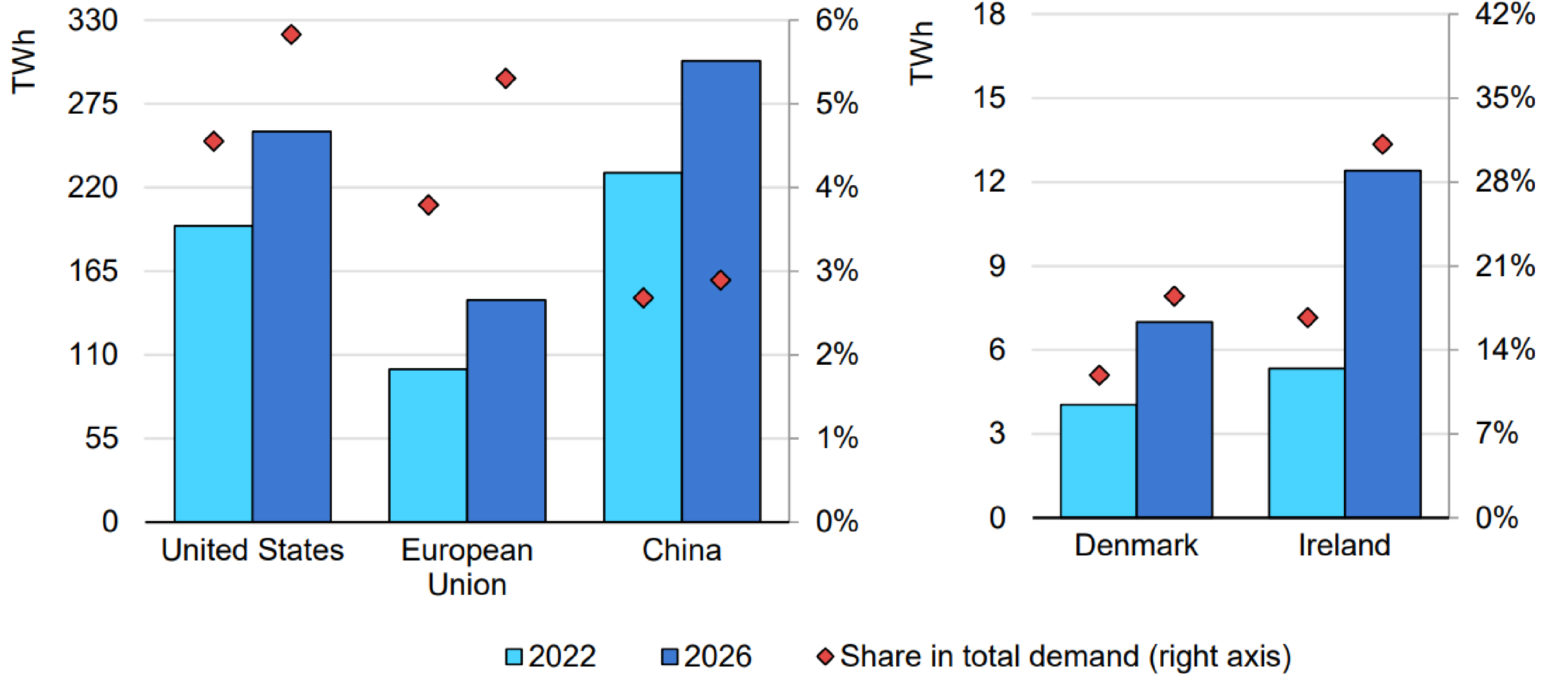
Weltweiter Energieverbrauch durch Datenzentren 2019-2026



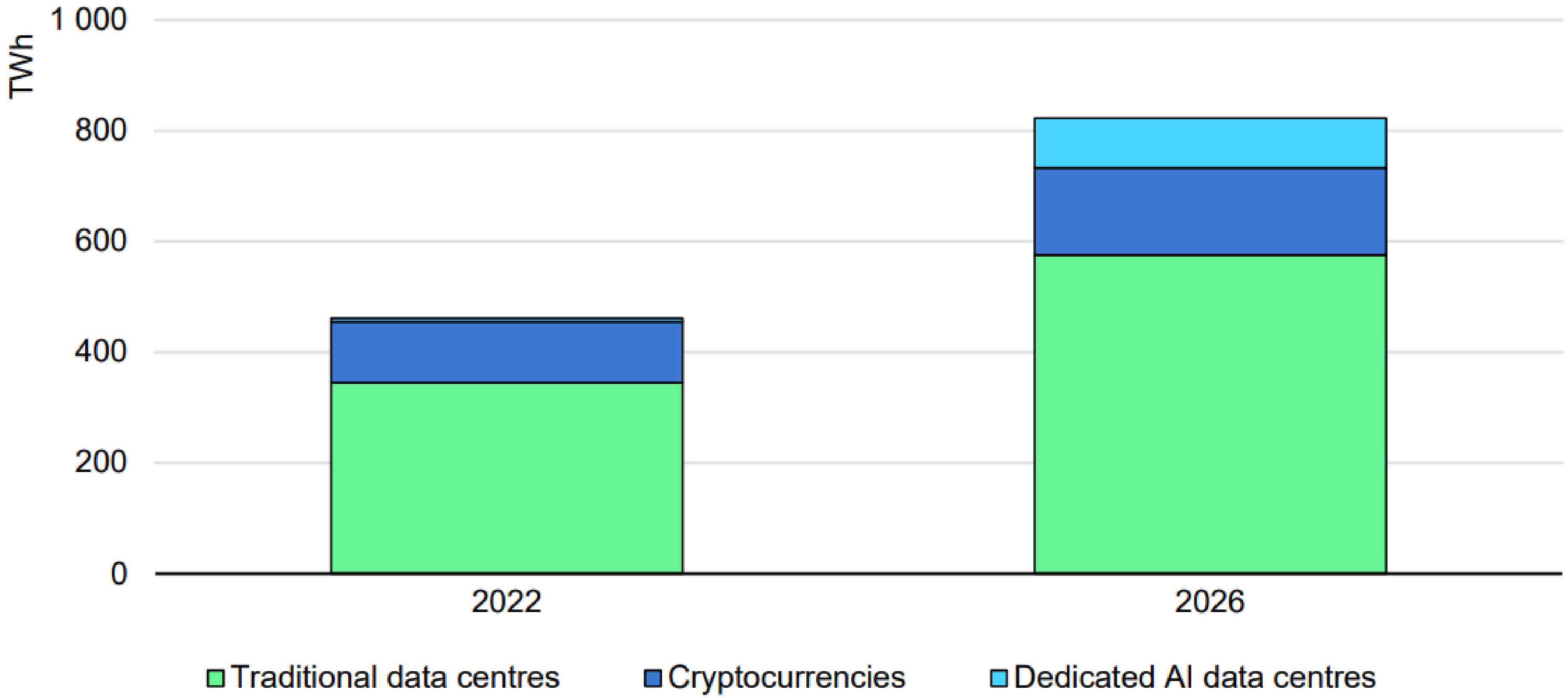
8Mrd. menschliche Gehirne brauchen etwa 1460TWh pro Jahr



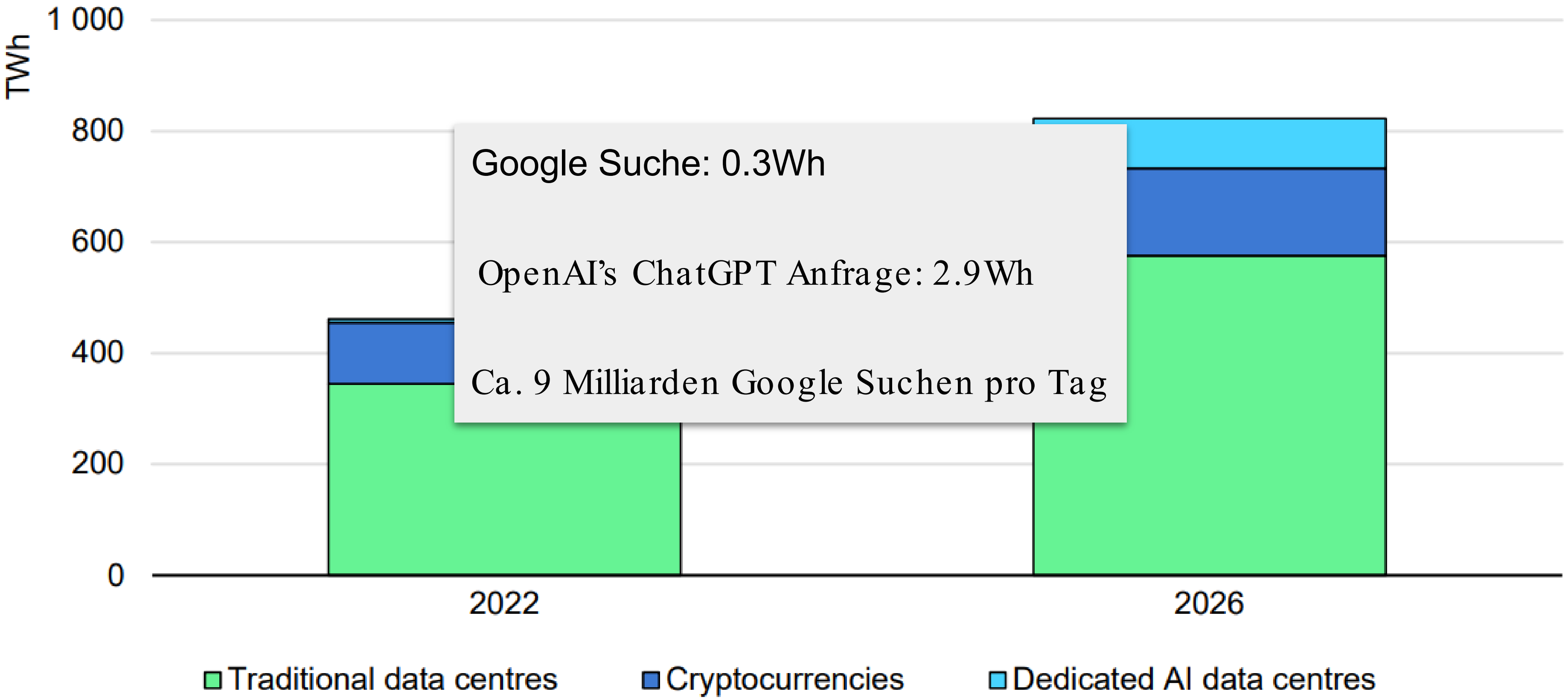
Estimated data centre electricity consumption and its share in total electricity demand in selected regions in 2022 and 2026



Estimated electricity demand from traditional data centres, dedicated AI data centres and cryptocurrencies, 2022 and 2026, base case



Estimated electricity demand from traditional data centres, dedicated AI data centres and cryptocurrencies, 2022 and 2026, base case



Google Suche: 0.3Wh

OpenAI's ChatGPT Anfrage: 2.9Wh

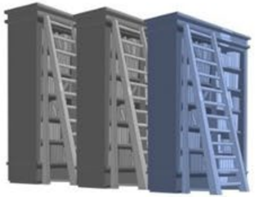
Ca. 9 Milliarden Google Suchen pro Tag

Energieverbrauch für KI-Training und Inferenz

Das OpenAI GPT-4 Modell

TRAININGSDATEN
für 13 Billionen* Tokens

650 km
lange Reihe von Bücherregalen



100.000 Tokens pro Buch
100 Bücher pro Regal
2 Regale pro Meter

RECHENZEIT
für 2,15 e25 FLOPs

7 Mio. Jahre
auf mittelgroßem Laptop (100 GFLOPs)

100 GLOPs pro Sekunde

MODELLGRÖSSE
für 1,8 Billionen* Parameter

30.000
fußballfeldgroße Excel-Tabellen

1x1 cm pro Excel-Zelle
100x60 m Feldgröße

Model-Training Kosten

Year	Model Name	Model Creators/Contributors	Training Cost (USD) Inflation-adjusted
2017	Transformer	Google	$930
2018	BERT-Large	Google	$3,288
2019	RoBERTa Large	Meta	$160,018
2020	GPT-3 175B (davinci)	OpenAI	$4,324,883
2021	Megatron-Turing NLG 530B	Microsoft/NVIDIA	$6,405,653
2022	LaMDA	Google	$1,319,586
2022	PaLM (540B)	Google	$12,389,056
2023	GPT-4	OpenAI	$78,352,034
2023	Llama 2 70B	Meta	$3,931,897
2023	Gemini Ultra	Google	$191,400,000

https://www.visualcapitalist.com/training-costs-of-ai-models-over-time/

Model-Training Ausstoß

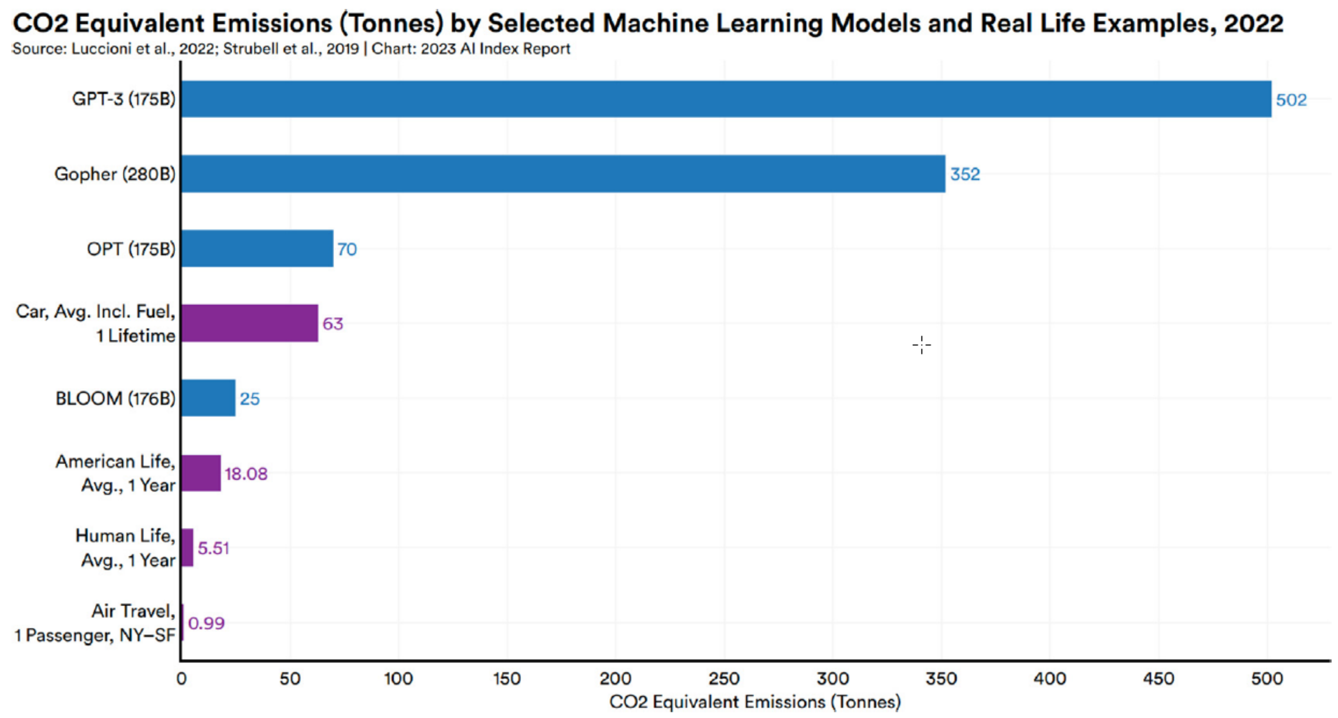
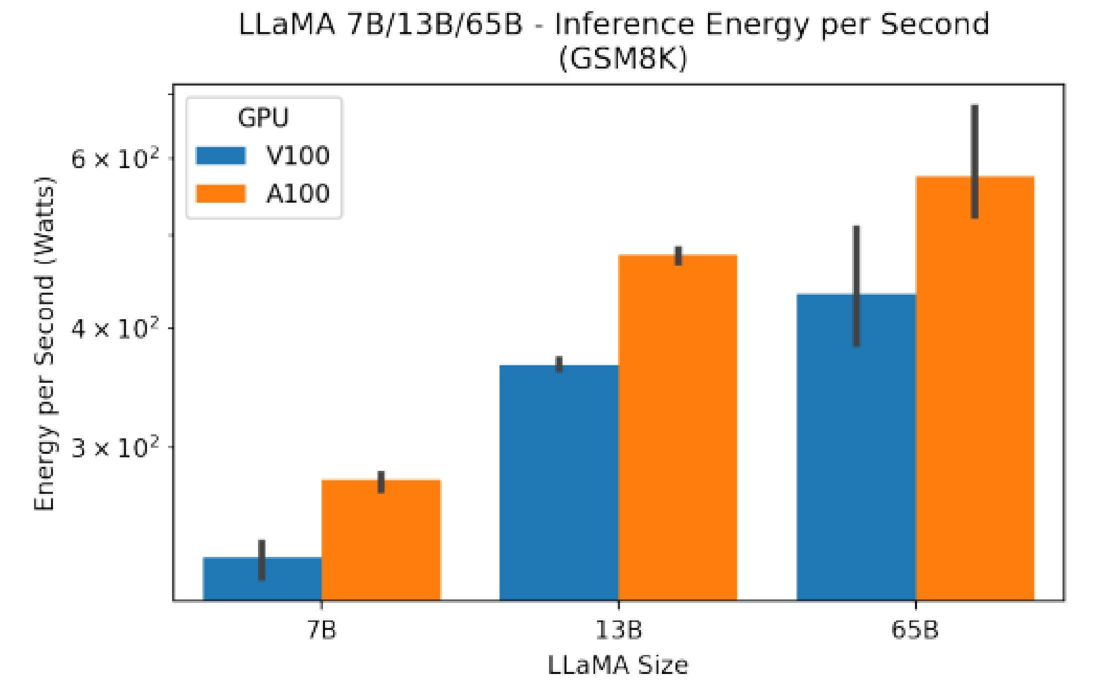


Fig. 1. CO_2 equivalent emissions for training ML models (blue) and of real-life cases (violet). In brackets, the billions of parameters adjusted for each model.

https://www.sciencedirect.com/science/article/pii/S0925231224008671

From Words to Watts: Benchmarking the Energy Costs of Large Language Model Inference

Siddharth Samsi*§, Dan Zhao†, Joseph McDonald*, Baolin Li‡, Adam Michaleas*,
Michael Jones*, William Bergeron*, Jeremy Kepner*, Devesh Tiwari‡, Vijay Gadepally*
* MIT, † NYU, ‡ Northeastern University



A100 65B needs 600W/s, which is about 14,4 kWh per day.

Remember, your brain needs about 0,5 kWh per day!

Energy consumption for training of
- GPT-3 ~1287 MWh
- HF BLOOM ~433MWh

In 2023, OpenAI consumed about 564MWh per DAY for ChatGPT's ops

Kyle Corbitt
@corbtt

Spoke to a Microsoft engineer on the GPT-6 training cluster project. He kvetched about the pain they're having provisioning infiniband-class links between GPUs in different regions.

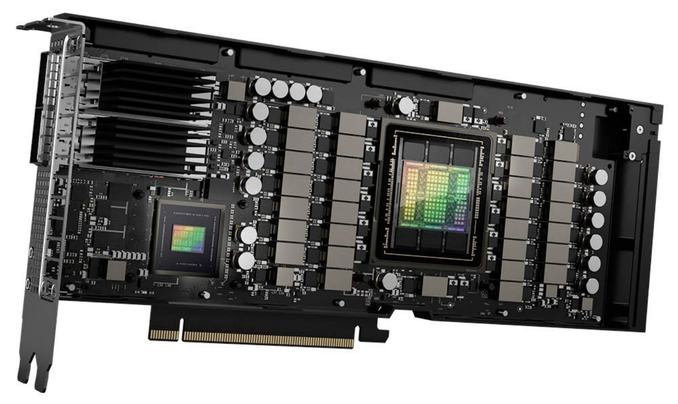
Me: "why not just colocate the cluster in one region?"
Him: "Oh yeah we tried that first. We can't put more than 100K H100s in a single state without bringing down the power grid." 🤯

Last edited 11:38 PM · Mar 25, 2024 · **1.8M** Views

225 1.1K 5.9K 2K

Post your reply Reply

Kyle Corbitt @corbtt · Mar 26
All the AI haters in my replies don't realize what GPT-6+ will do for humanity's standard of living.



This is Nvidia's H100 GPU. It has a peak power consumption of ~700W. At a 61% annual utilization, it is equivalent to the power consumption of the average American household occupant (based on 2.51 people/household).

Nvidia's estimated sales of H100 GPUs is 1.5-2mil H100 GPUs in 2024.

Nachhaltigkeit & Herausforderungen durch KI

Herausforderungen für die IT -Nachhaltigkeit durch KI

Direkter Anstieg des Energieverbrauchs und der CO2-Emissionen durch erhöhten Bedarf an Rechenleistung und Speicherkapazität für das Training und die Nutzung generativer KI-Modelle in Rechenzentren.

Indirekte Zunahme des Energieverbrauchs durch wachsende Nutzung und Verbreitung von KI-Anwendungen, die auf rechenintensiven generativen Modellen basieren.

Herausforderungen für die IT -Nachhaltigkeit durch KI

Erhöhter Bedarf an seltenen Rohstoffen und Materialien für die Produktion leistungsfähiger Hardware, die für generative KI benötigt wird, was indirekte Umweltauswirkungen mit sich bringt.

Herausforderungen für die IT-Nachhaltigkeit durch KI

Schwierigkeiten bei der Entwicklung und Umsetzung von Standards und Regulierungen für die nachhaltige Nutzung generativer KI aufgrund der rasanten technologischen Entwicklung.

Anti-electricity cartoon from ~1900

Nachhaltigkeitsziele durch KI erreichen

Grüne Künstliche Intelligenz

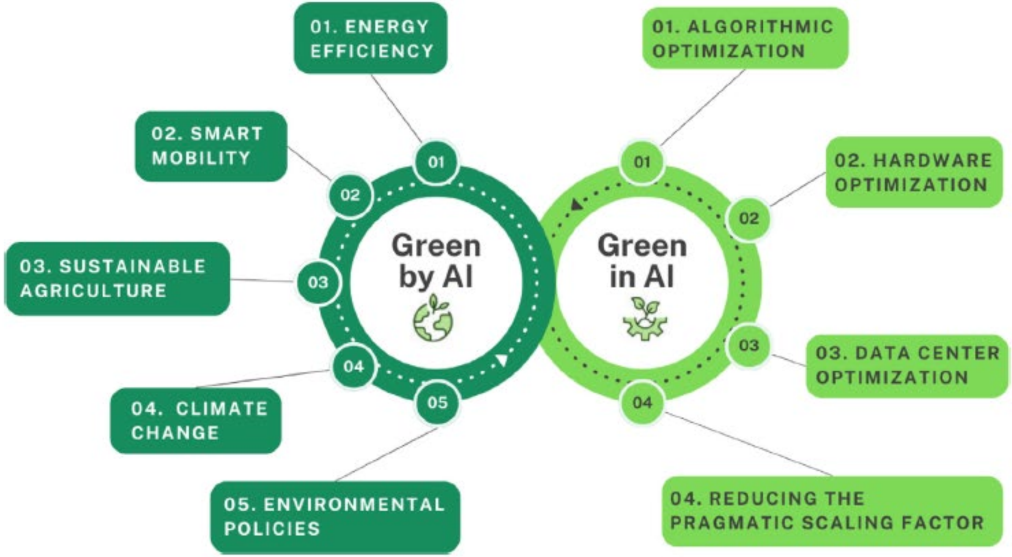
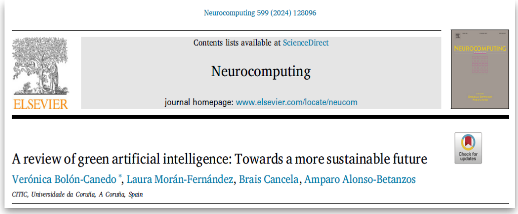


Fig. 3. Overview of green-in vs green-by algorithms.



Neurocomputing 599 (2024) 128096

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A review of green artificial intelligence: Towards a more sustainable future

Verónica Bolón-Canedo*, Laura Morán-Fernández, Brais Cancela, Amparo Alonso-Betanzos

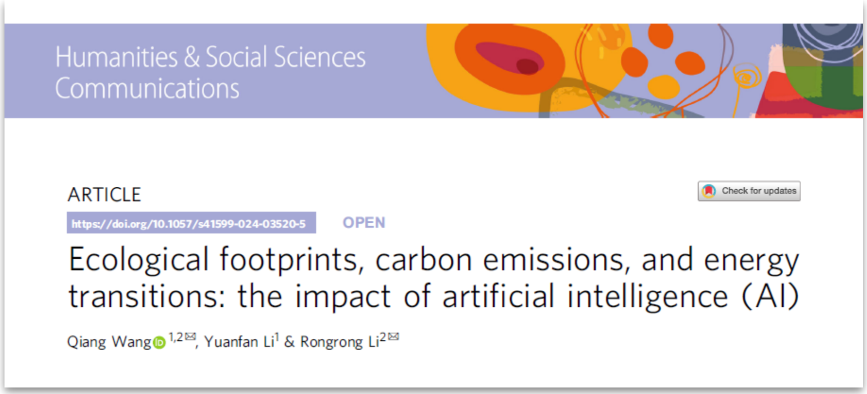
CITIC, Universidade da Coruña, A Coruña, Spain

https://www.sciencedirect.com/science/article/pii/S0925231224008671

Der Einfluss von Künstlicher Intelligenz

Every 1% increase in the level of AI development corresponds to a 0.0025% increase in energy transition, a 0.0018% decrease in ecological footprint, and a 0.0013% decrease in carbon emissions.

In other words, AI has the strongest effect in promoting energy transition, followed by the reduction effect on ecological footprint, and then the impact of AI on carbon emissions.


Humanities & Social Sciences Communications

ARTICLE

Check for updates

https://doi.org/10.1057/s41599-024-03520-5 OPEN

Ecological footprints, carbon emissions, and energy transitions: the impact of artificial intelligence (AI)

Qiang Wang[1,2], Yuanfan Li[1] & Rongrong Li[2]

https://www.nature.com/articles/s41599-024-03520-5

Handlungsempfehlungen

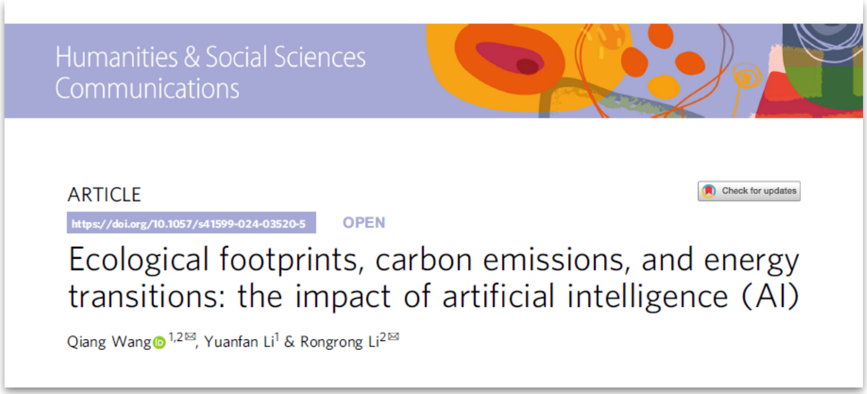
Countries should increase investment in AI R&D, propose AI development strategies and establish long-term development mechanisms

Countries with high industrial proportions should make full use of the advantages of their own industrial systems to develop renewable energy with the help of AI

Trade openness plays an important role in exerting the positive effects of AI

Setting binding long-term goals for AI development is of great significance to all countries

A higher level of energy transition can help AI better reduce its environmental footprint


Humanities & Social Sciences Communications

ARTICLE

Check for updates

https://doi.org/10.1057/s41599-024-03520-5 OPEN

Ecological footprints, carbon emissions, and energy transitions: the impact of artificial intelligence (AI)

Qiang Wang[1,2], Yuanfan Li[1] & Rongrong Li[2]

https://www.nature.com/articles/s41599 -024-03520-5

ELSEVIER

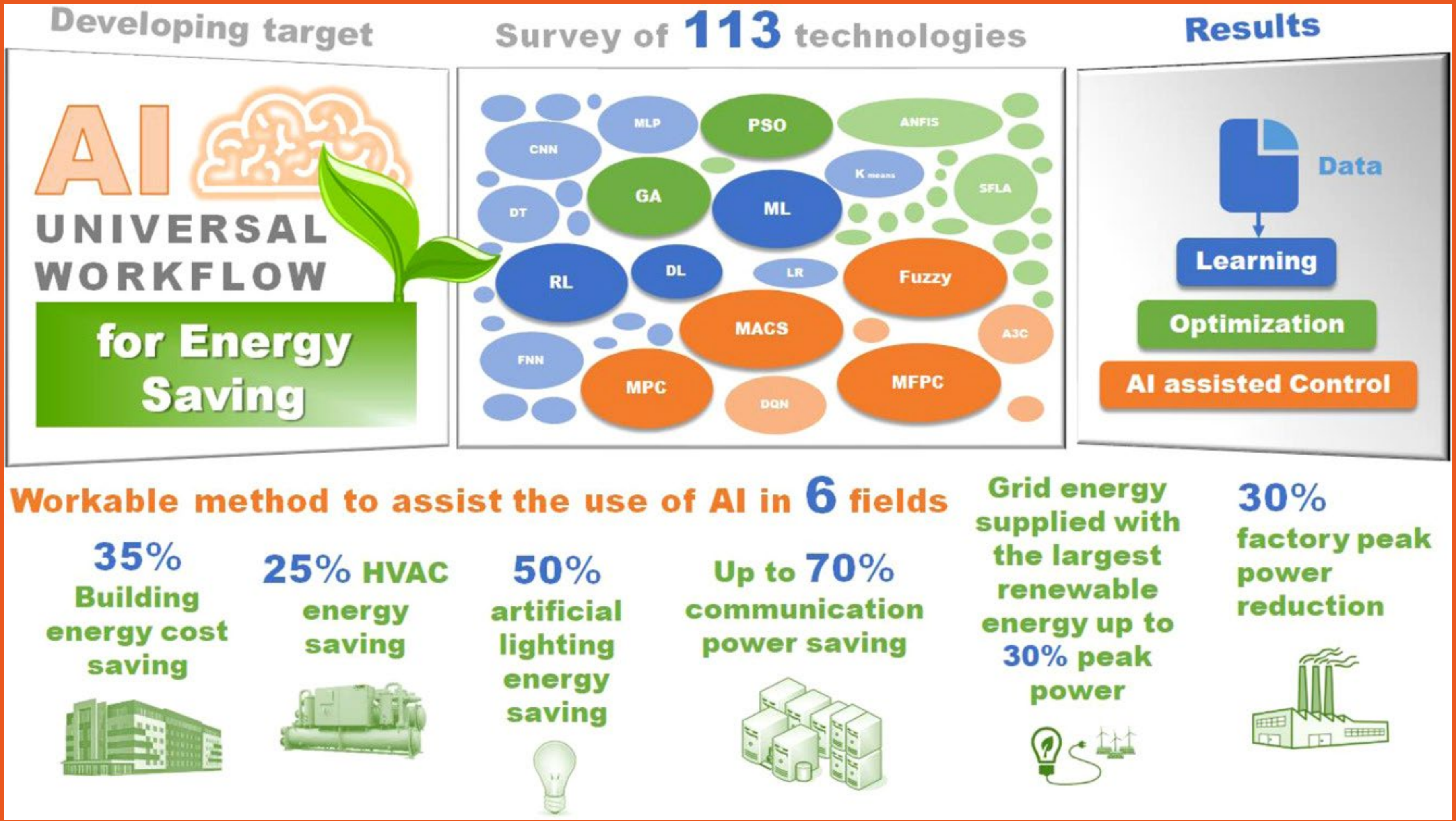
Energy Reports
Volume 8, November 2022, Pages 1602-1633

Research paper

Universal workflow of artificial intelligence for energy saving

Da-sheng Lee [a], Yan-Tang Chen [a], Shih-Lung Chao [b]

Source: https://www.sciencedirect.com/science/article/pii/S2352484721015055



Source: https://www.sciencedirect.com/science/article/pii/S2352484721015055

Sustainable Development Goals der UN

13 CLIMATE ACTION

13.1 Strengthen resilience and adaptive capacity to climate-related hazards and natural disasters in all countries

13.2 Integrate climate change measures into national policies, strategies and planning

13.3 Improve education, awareness-raising and human and institutional capacity on climate change mitigation, adaptation, impact reduction and early warning

13.A Implement the commitment undertaken by developed-country parties to the United Nations Framework Convention on Climate Change to a goal of mobilizing jointly $100 billion annually by 2020 from all sources to address the needs of developing countries in the context of meaningful mitigation actions and transparency on implementation and fully operationalize the Green Climate Fund through its capitalization as soon as possible

13.B Promote mechanisms for raising capacity for effective climate change-related planning and management in least developed countries and small island developing States, including focusing on women, youth and local and marginalized communities

*Acknowledging that the United Nations Framework Convention on Climate Change is the primary international, intergovernmental forum for negotiating the global response to climate change.

14 LIFE BELOW WATER

14.1 By 2025, prevent and significantly reduce marine pollution of all kinds, in particular from land-based activities, including marine debris and nutrient pollution

14.2 By 2020, sustainably manage and protect marine and coastal ecosystems to avoid significant adverse impacts, including by strengthening their resilience, and take action for their restoration in order to achieve healthy and productive oceans

14.3 Minimize and address the impacts of ocean acidification, including through enhanced scientific cooperation at all levels

14.4 By 2020, effectively regulate harvesting and end overfishing, illegal, unreported and unregulated fishing and destructive fishing practices and implement science-based management plans, in order to restore fish stocks in the shortest time feasible, at least to levels that can produce maximum sustainable yield as determined by their biological characteristics

14.5 By 2020, conserve at least 10 per cent of coastal and marine areas, consistent with national and international law and based on the best available scientific information

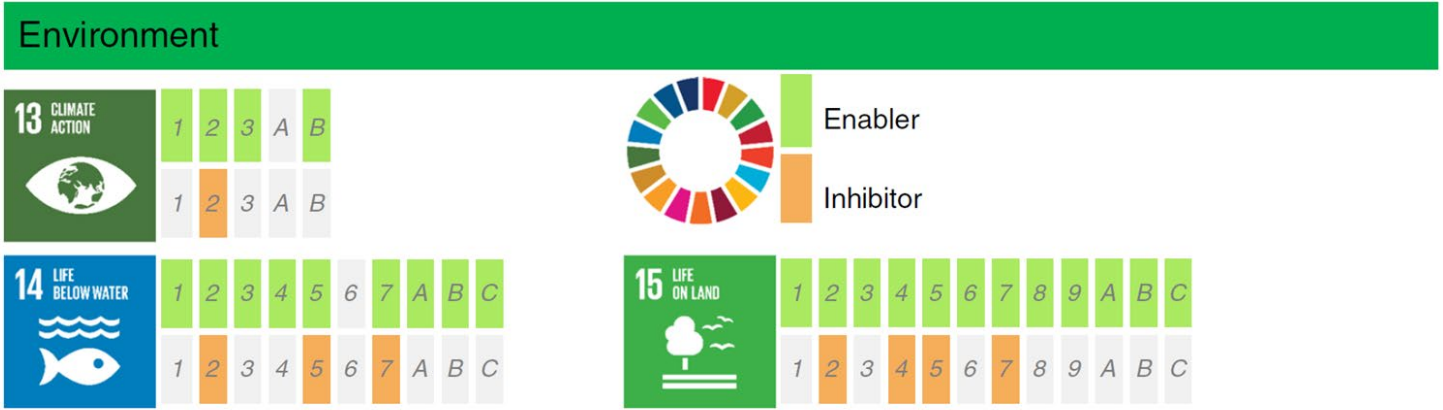
14.6 By 2020, prohibit certain forms of fisheries subsidies which contribute to overcapacity and overfishing, eliminate subsidies that contribute to illegal, unreported and unregulated fishing and refrain from introducing new such subsidies, recognizing that appropriate and effective special and differential treatment for developing and least developed countries should be an integral part of the World Trade Organization fisheries subsidies negotiation

14.7 By 2030, increase the economic benefits to Small Island developing States and least developed countries from the sustainable use of marine resources, including through sustainable management of fisheries, aquaculture and tourism

14.A Increase scientific knowledge, develop research capacity and transfer marine technology, taking into account the Intergovernmental Oceanographic Commission Criteria and Guidelines on the Transfer of Marine Technology, in order to improve ocean health and to enhance the contribution of marine biodiversity to the development of developing countries, in particular small island developing States and least developed countries

14.B Provide access for small-scale artisanal fishers to marine resources and markets

Sustainable Development Goals der UN




nature COMMUNICATIONS

PERSPECTIVE

https://doi.org/10.1038/s41467-019-14108-y OPEN

The role of artificial intelligence in achieving the Sustainable Development Goals

Ricardo Vinuesa[1*], Hossein Azizpour[2], Iolanda Leite[2], Madeline Balaam[3], Virginia Dignum[4], Sami Domisch[5], Anna Felländer[6], Simone Daniela Langhans[7,8], Max Tegmark[9] & Francesco Fuso Nerini[10*]

https://www.nature.com/articles/s41467 -019-14108-y

Sustainable Development Goals der UN

Analyzing large-scale interconnected databases to develop joint actions aimed at preserving the environment.

AI advances will support the understanding of climate change and the modeling of its possible impacts.

AI will support low-carbon energy systems with high integration of renewable energy and energy efficiency.

Algorithms for automatic identification of possible oil spills.

Neural networks and objective-oriented techniques can be used to improve the classification of vegetation cover types based on satellite images.

nature COMMUNICATIONS

PERSPECTIVE

https://doi.org/10.1038/s41467-019-14108-y OPEN

The role of artificial intelligence in achieving the Sustainable Development Goals

Ricardo Vinuesa[1*], Hossein Azizpour[2], Iolanda Leite[2], Madeline Balaam[3], Virginia Dignum[4], Sami Domisch[5], Anna Felländer[6], Simone Daniela Langhans[7,8], Max Tegmark[9] & Francesco Fuso Nerini[10*]

https://www.nature.com/articles/s41467-019-14108-y

Beispiele im Bereich Energieeffizienz-Optimierung

Präzise Bedarfsvorhersagen

KI kann eine Vielzahl verschiedener Kennzahlen in Vorhersagen einbeziehen

- Historischer Bedarf
- Wetter
- Auslastung

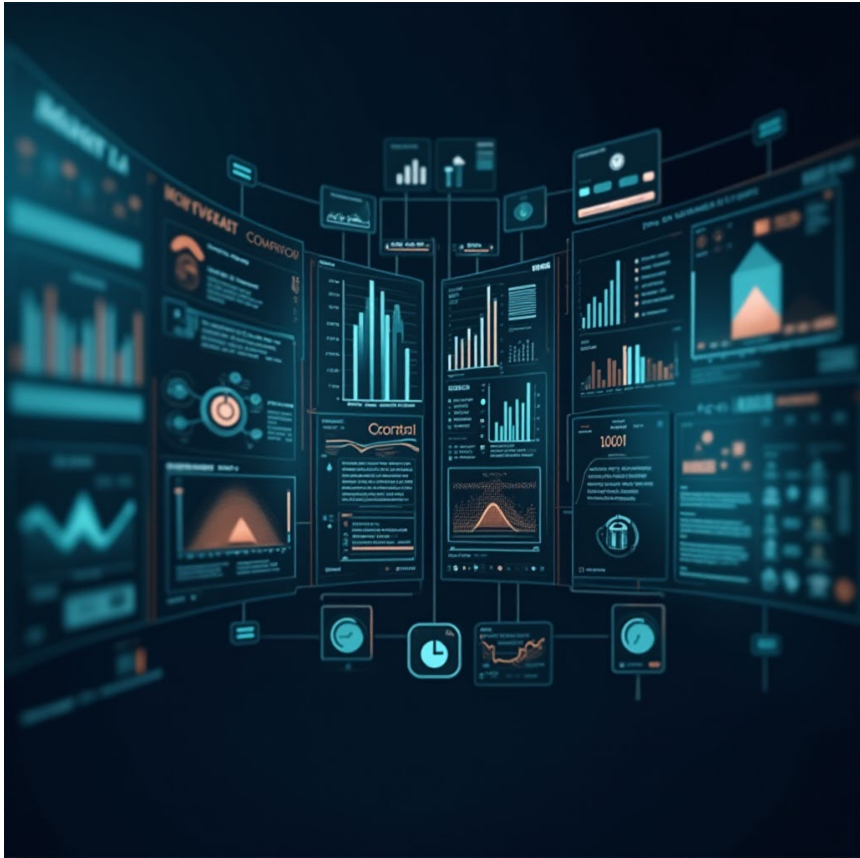
https://fepbl.com/index.php/estj/article/view/1015

Echtzeit Monitoring und Kontrolle

Ermöglicht Reaktion auf sich ändernde Gegebenheiten.

- Auslastung
- Wetter
- Preis

Erhöht Effizienz und hilft dabei, mit Stoßzeiten besser umzugehen.



https://fepbl.com/index.php/estj/article/view/1015

Erneuerbare Energien einbinden

KI kann erneuerbare Energiegewinnung vorhersagen und deren Verwendung maximieren.

“Demand Response Programs”, wodurch Verbraucher ihren Bedarf an Angebot-Nachfrage-Signale anpassen können.

https://fepbl.com/index.php/estj/article/view/1015

Datensicherheit

Verbrauchsdaten können sehr sensitiv sein.

Robuste Cybersecurity ist essentiell, aber auch anspruchsvoll.

https://fepbl.com/index.php/estj/article/view/1015

Regulierungen

Verstehen, Anwenden und Einhalten von Regulierungen/Vorschriften fügt ein weiteres Level an Komplexität hinzu.

https://fepbl.com/index.php/estj/article/view/1015

Neue Kompetenzen

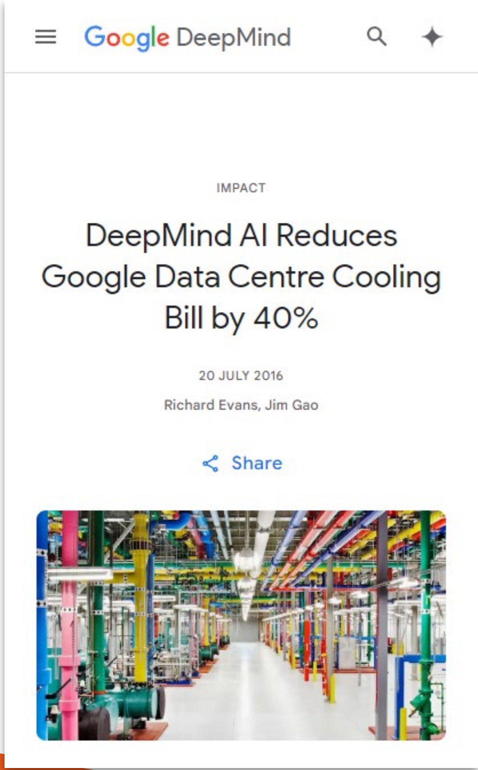
Implementierung und Verwendung von KI Lösungen bedarf spezielles Wissen in Bereichen

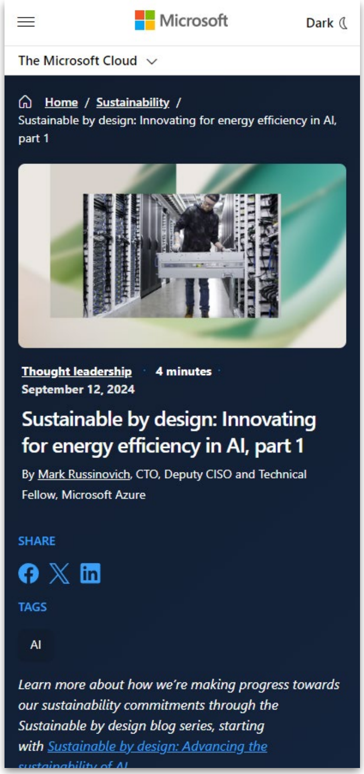
- KI Algorithmus
- Datenanalyse
- Energiesysteme

Integration in existierende Systeme kann komplex werden.

https://fepbl.com/index.php/estj/article/view/1015

Heute im Einsatz





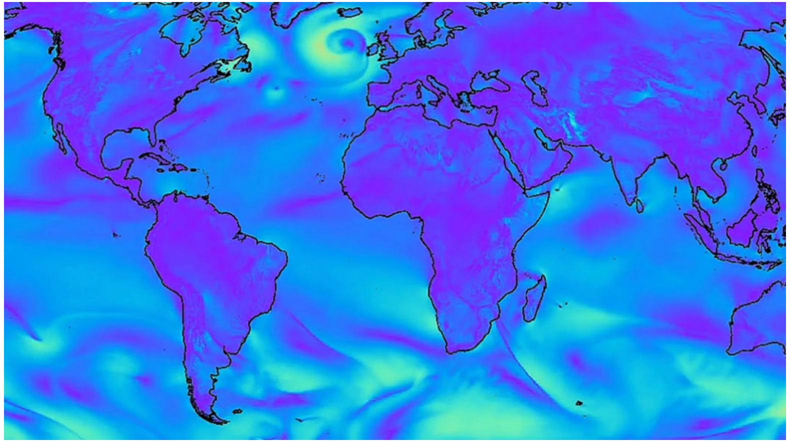
https://fepbl.com/index.php/estj/article/view/1015

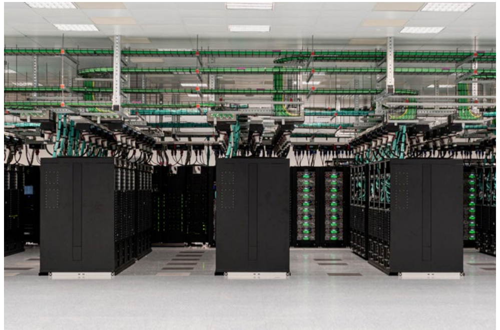
Entdeckung ungeahnter Möglichkeiten durch KI

GraphCast

- schlägt die aktuell besten aktiven Modelle in über 99% aller Wetter-Variablen in über 90% aller getesteten (1.300) Regionen

- läuft auf Standard-PC und generiert die Vorhersagen innerhalb weniger Sekunden



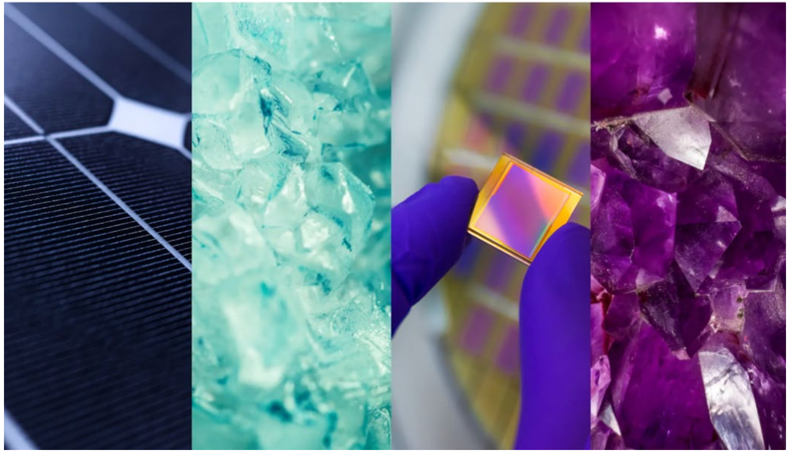


Source: https://deepmind.google/discover/blog/graphcast-ai-model-for-faster-and-more-accurate-global-weather-forecasting/

VAUDIENCE.AI

GNoME

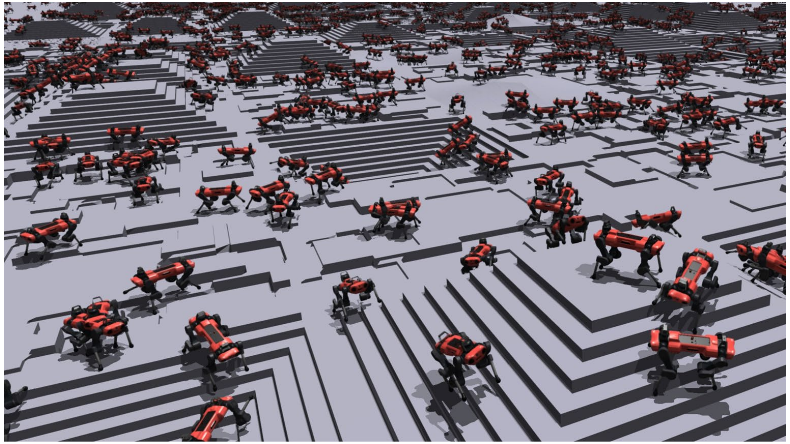
- identifiziert > 2,2 Millionen neue Kristallstrukturen
- 380.000 als stabile neue Materialien vorhergesagt
- faszinierenden Eigenschaften, die neue Technologien ermöglichen werden
- >70% erfolgreich bestätigt
- Entspricht etwa 800 Jahre Materialforschung



Source: https://deepmind.google/discover/blog/millions-of-new-materials-discovered-with-deep-learning/

NVIDIA Isaac Sim

- trainiert Roboter virtuell
- spart sehr viel Zeit und Geld
- entdeckt aktiv neue Ansätze



VAUDIENCE.AI

Fragen? Panik? Diskussion?

VAUDIENCE.AI

DANKE!

Wir hoffen, es war spannend!

Quellen

References:

- Raichle, M. E., & Gusnard, D. A. (2002). Appraising the brain's energy budget. PNAS, 99(16), 10237-10239.
- Herculano-Houzel, S. (2016). The Human Advantage: A New Understanding of How Our Brain Became Remarkable. MIT Press.
- Mink, J. W., et al. (1981). Energy consumption of cerebral cortex and subcortical nuclei of the rat. Journal of Neuroscience, 1(3), 245-252.
- Herculano-Houzel, S. (2011). Scaling of brain metabolism with a fixed energy budget per neuron: Implications for neuronal activity, plasticity and evolution. PLoS ONE, 6(3), e17514.
- Roth, G., & Dicke, U. (2005). Evolution of the brain and intelligence. Trends in Cognitive Sciences, 9(5), 250-257.
- Jardim-Messeder, D., et al. (2017). Dogs have the most neurons, though not the largest brain: Trade-off between body mass and number of neurons in the cerebral cortex of large carnivoran species. Frontiers in Neuroanatomy, 11, 118.
- Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. Frontiers in Human Neuroscience, 3, 31.
- Azevedo, F. A., et al. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. Journal of Comparative Neurology, 513(5), 532-541.
- Marino, L. (2002). Convergence of complex cognitive abilities in cetaceans and primates. Brain, Behavior and Evolution, 59(1-2), 21-32.
- Ridgway, S. H., et al. (2017). Dolphin brain size and smarts. Frontiers in Neuroanatomy, 11, 16.
- Herculano-Houzel, S., et al. (2014). The elephant brain in numbers. Frontiers in Neuroanatomy, 8, 46.
- Shulman, R. G., et al. (2004). Energetic basis of brain activity: implications for neuroimaging. Trends in Neurosciences, 27(8), 489-495.