Green Artificial Intelligence

An overview of the emerging field of green artificial intelligence and the development of a Green AI Index

Unsustainable Burden of AI

- ICT sector represents **~4%** of global electricity consumption in 2020.
- ICT sector represents **1.4%** of global GHG emissions in 2020.
- The ICT sector has increased its emissions by about **5%** from 2015.

*Source: IPCC Climate Change 2022: Mitigation of Climate Change

Green AI

Algorithm Optimization

Optimizing algorithms to reduce computational resource requirements and energy consumption, such as sparse training methods, quantization techniques, and low-precision arithmetic operations.

Data Center Optimization

Strategies to optimize data center operations, such as dynamic load management, efficient cooling systems, and use of renewable energy sources.

Hardware Optimization

Innovations in hardware design, including specialized AI accelerators, parallelization techniques, and edge computing to improve energy efficiency.

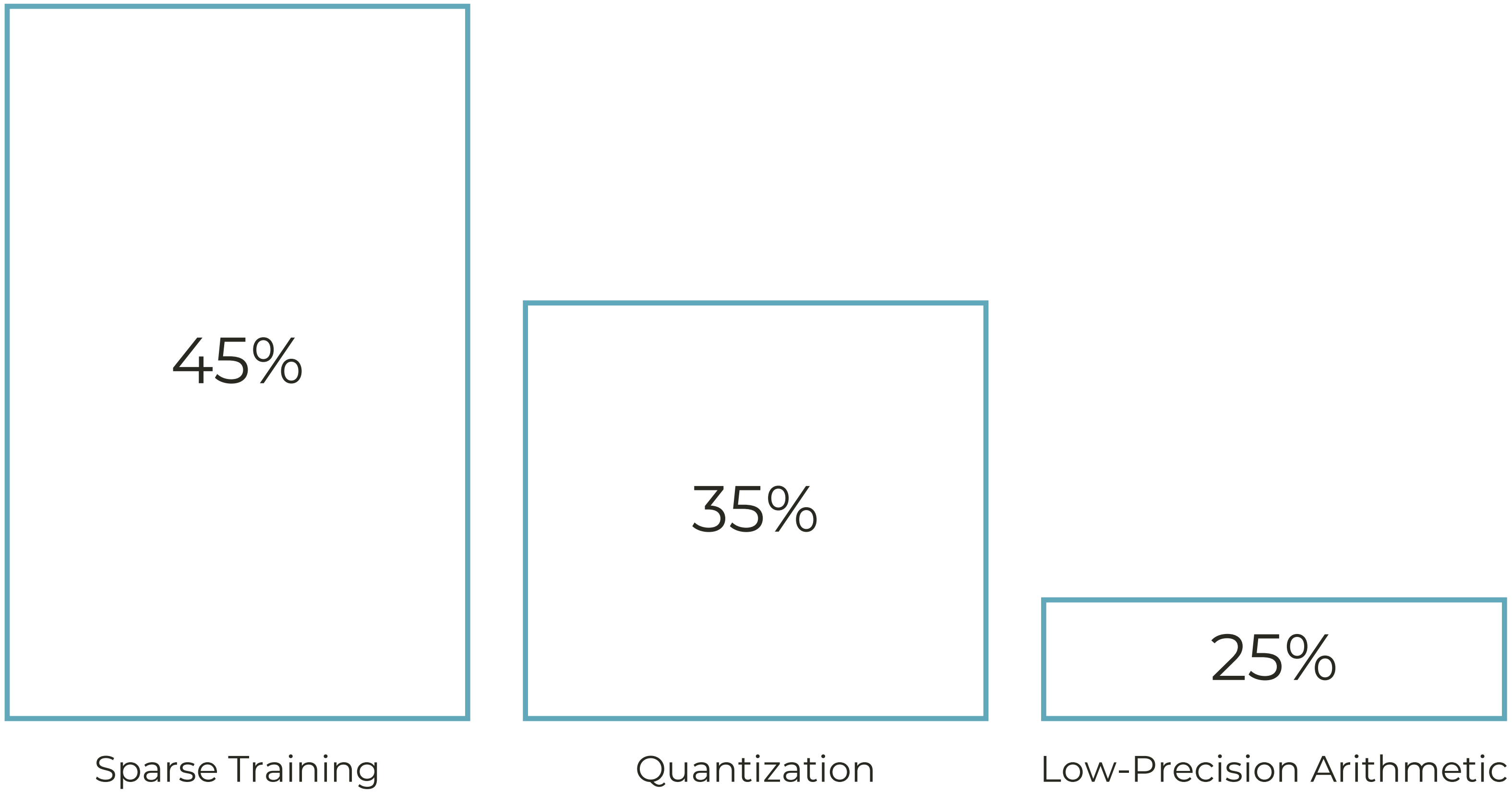
Energy Consumption Calculation Tools

Frameworks and tools to accurately measure and report the energy consumption and carbon footprint associated with AI systems, like CarbonTracker, CodeCarbon, and the 'Green Algorithms' project.

By implementing these green-in AI strategies, we can significantly reduce the environmental impact of AI systems and contribute to a more sustainable future.

Algorithm Optimization

Percent reduction in computational resource requirements and energy consumption



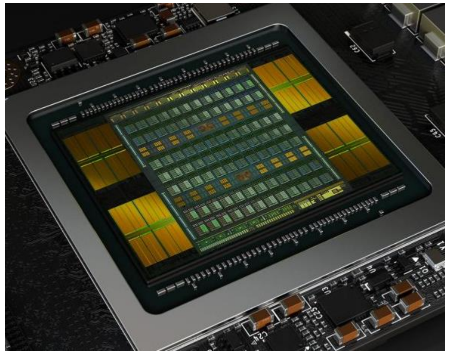
A review of green artificial intelligence: Towards a more sustainable future - Neurocomputing Journal, 2024

https://doi.org/10.1016/j.neucom.2024.128096

Green Algorithms

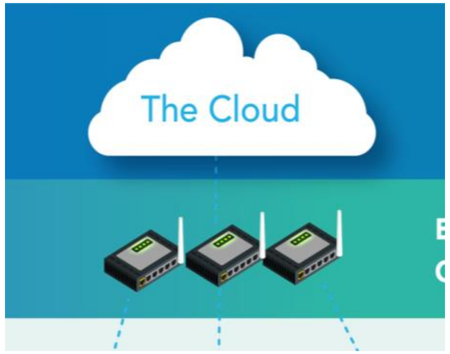
- Resource Utilization Optimization
- Power Consumption Minimization
- Computing Sustainability Promotion
- Reducing Environmental Impact

Hardware Optimization



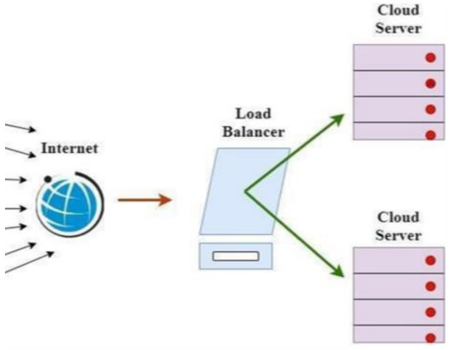
Tensor Processing Unit (TPU)

A specialized hardware chip designed by Google to accelerate deep learning workloads, providing significant performance and energy efficiency improvements over GPUs.



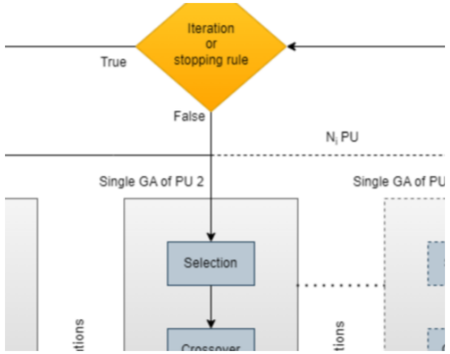
Edge Computing Device

Small, low-power computing devices that can perform data processing and analysis at the edge of the network, close to where data is generated, reducing the need for data transfer to the cloud.



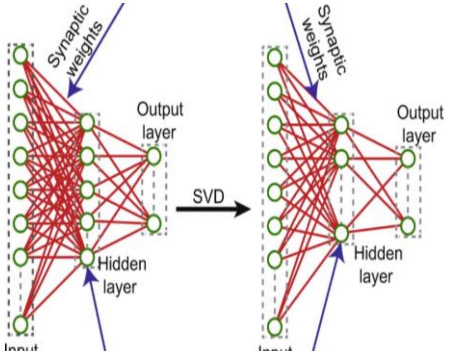
Cloud Data Center

Modern, purpose-built data centers that are highly optimized for energy efficiency, with advanced cooling and power distribution technologies, enabling significant reductions in energy consumption and emissions compared to traditional on-premises data centers.



Parallelization

Distributing computation across multiple processing cores to reduce the training time of algorithms, but with the caveat that increasing the number of cores does not always lead to proportional improvements in execution times and emissions reduction.



Hardware-Optimized Algorithms

Developing efficient and compact neural network architectures that can be executed more efficiently on specialized hardware like TPUs and edge devices, reducing computational resource requirements and energy consumption without compromising performance or accuracy.

Data Center Optimization

Leverage Renewable Energy

Power data centers with wind, solar, and hydroelectric power to reduce carbon footprint by 5-10x

Dynamic Load Management

Develop algorithms to dynamically manage server loads, cooling systems, and resource allocation for energy efficiency

Locational Optimization

Utilize cloud-based data centers optimized for energy savings compared to on-premises facilities

Waste Heat Capture

Capture and repurpose waste heat from data centers to heat nearby buildings or power other processes

AI-Powered Monitoring

Leverage AI-based monitoring systems to continuously optimize data center operations and energy usage

Energy Consumption Calculation Tools

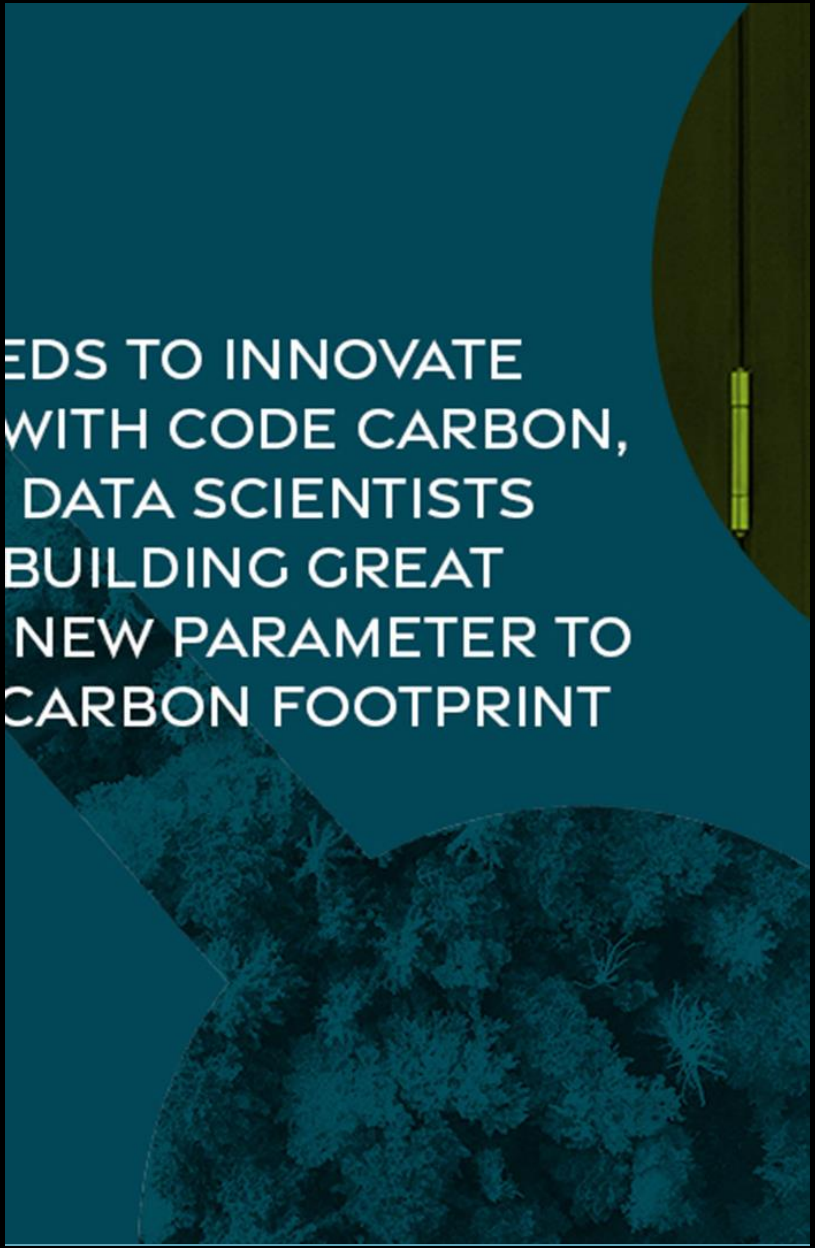
Exploring frameworks and tools to assess the environmental impact of machine learning systems and their carbon footprint.

CarbonTracker

CarbonTracker is an environmental monitoring tool designed to track and analyze GHG emissions, particularly carbon dioxide, at various sources and locations. It uses a combination of approaches, including atmospheric measurements and statistical modeling, to estimate and visualize real-time carbon emissions.

CodeCarbon

CodeCarbon is a specialized software tool that assists developers and organizations in tracking and managing the carbon emissions generated by their software and codebase. By integrating CodeCarbon into the development pipeline, users can monitor the energy consumption and environmental impact of their software applications, and receive actionable insights and recommendations to reduce their carbon footprint.



PowerTop
Optimizing Linux Power Efficiency

- Real-Time Power Monitoring

  Analyzes power consumption and management in real-time on Linux-based systems

- Power-Hungry Identification

  Identifies power-hungry processes, devices, and components to optimize energy efficiency

- Battery Life Extension

  Helps extend battery life on laptops and reduce power usage in servers

- Detailed Power Usage Reporting

  Displays detailed information on power usage to enable informed decisions

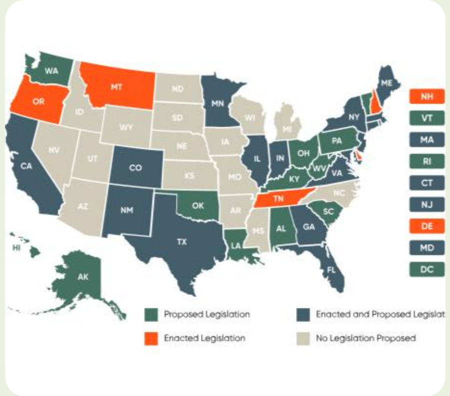
- Energy Conservation

  Enables users to fine-tune power-related settings and identify areas for improvement

Regulation for Green AI

Examining the role of regulations and sustainable practices in the development and deployment of AI systems

Regulations in the USA, China and EU

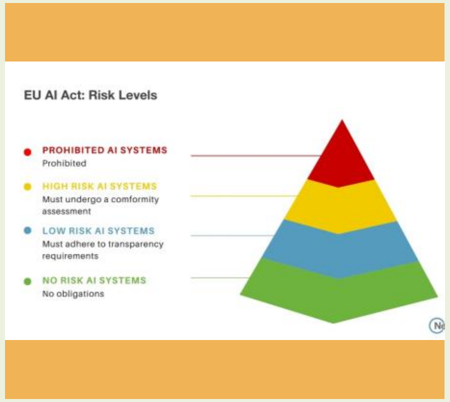


USA Regulations

The recent Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence in the USA briefly mentions an objective to promote the safe, responsible, and rights-affirming development and deployment of AI abroad, but does not enforce sustainability restrictions.

China Regulations

China's AI regulation, similar to the USA, does not enforce any sustainability restrictions on the development of AI algorithms.

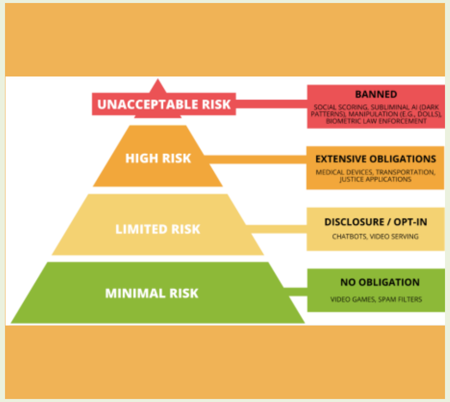


EU AI Act

In contrast, the EU's AI Act establishes obligations for AI systems based on their potential risks and impacts, aiming to safeguard fundamental rights, democracy, the rule of law, and environmental sustainability.

Reporting Energy Efficiency

The AI Act requires reporting energy efficiency as an additional obligation for high-impact general purpose AI models, particularly those posing systemic risk, such as generative AI.



Codes of Conduct

For non-high-risk AI systems, the AI Act encourages the development of codes of conduct that adhere to some or all the mandatory requirements applicable to high-risk systems, including minimizing the environmental impact of AI systems.


“By establishing guidelines and standards for energy efficiency in AI development, stakeholders can ensure that advances in AI not only drive innovation but also align with broader sustainability objectives.”

Incorporating Sustainability as a Regulatory Requirement

amazon

DeepMind

IBM

ANTHROP\C

NVIDIA

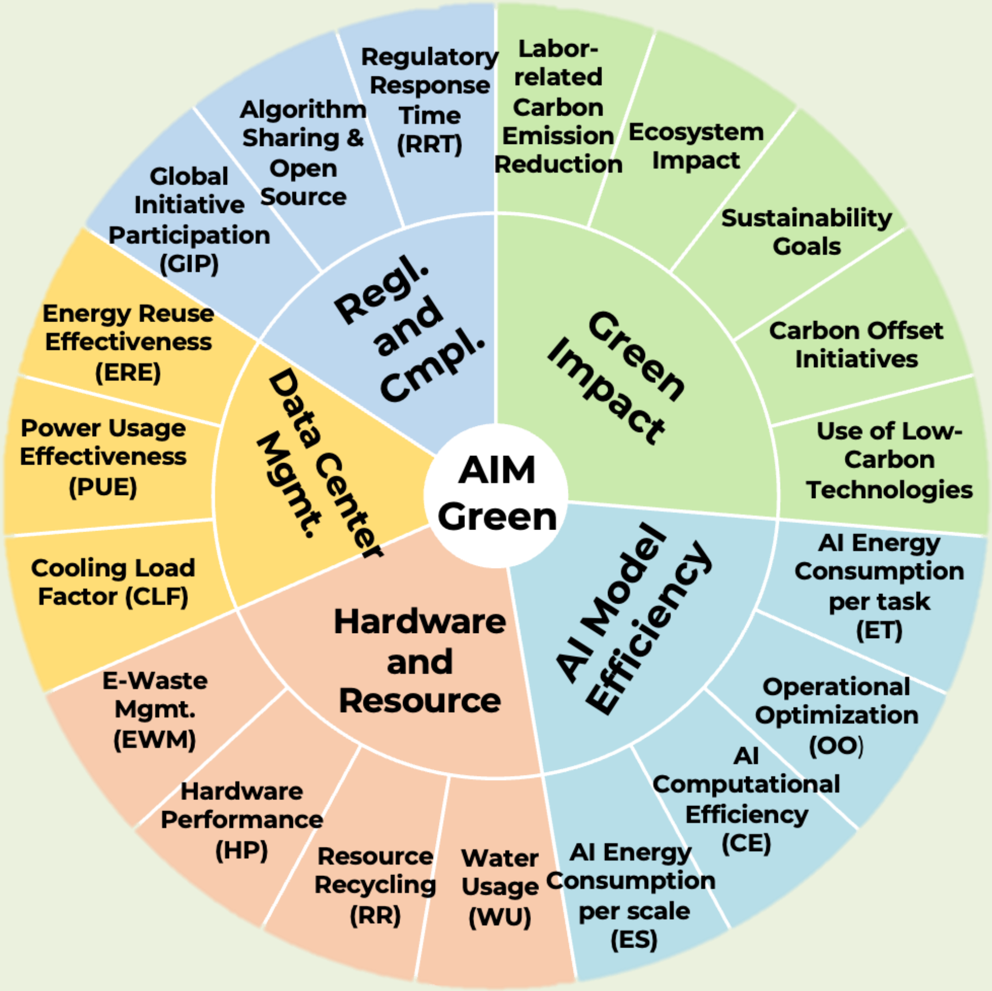
Meta

AIM Green Index

Measurement Dimensions

Five dimensions
19 key indicators

- AI Model Efficiency
- Hardware and Resource
- Data Center Management
- Regulation and Compliance
- Socio-economic and Green Impact



1 AI Model Efficiency 1/2

Tracks energy usage and carbon emission for AI model training and operation.

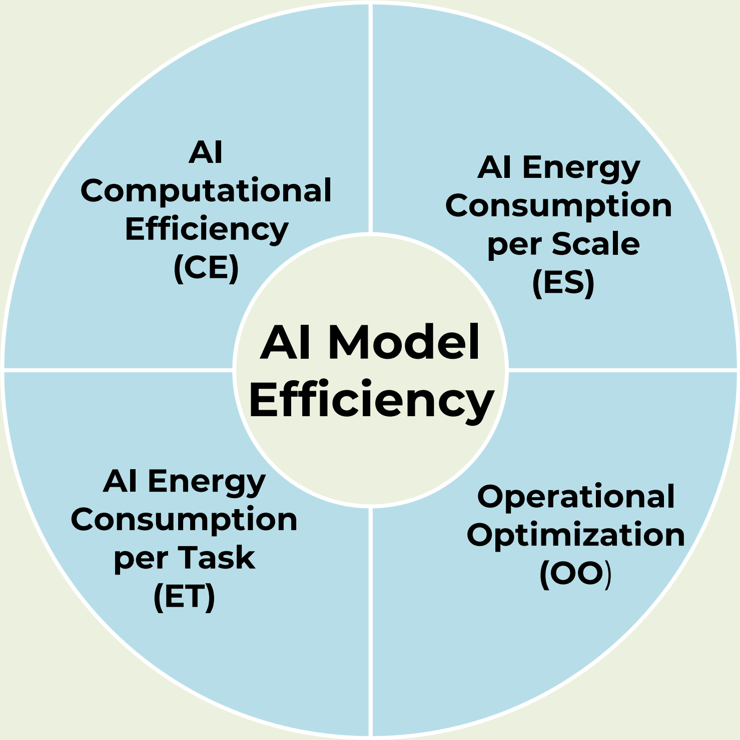
- AI Energy Consumption per Task (ET)

Energy usage for AI training and inference tasks.

- AI Energy Consumption per Scale (ES)

Total energy consumption to the business scale.

ES = the number of customer access per day/AI energy consumption per day



1 AI Model Efficiency 2/2

Tracks energy usage and carbon emission for AI model training and operation.

- AI Computational Efficiency (CE)

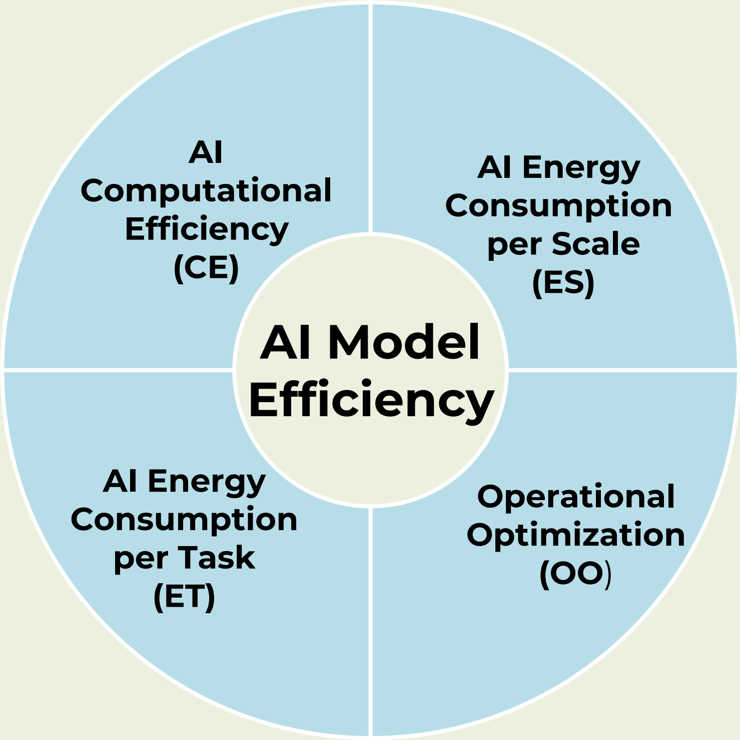
Efficiency of AI models.

CE = computing task complexity/ the number of floating-point operations (FPO) of the task

- Operational Optimization (OO)

Continuous improvement measures to optimize AI operations and reduce waste.

OO = the energy savings due to operational optimization/ total operating costs



2 Hardware and Resource 1/2

Tracks resource usage and recycling for AI hardwares.

- Hardware Performance (HP)

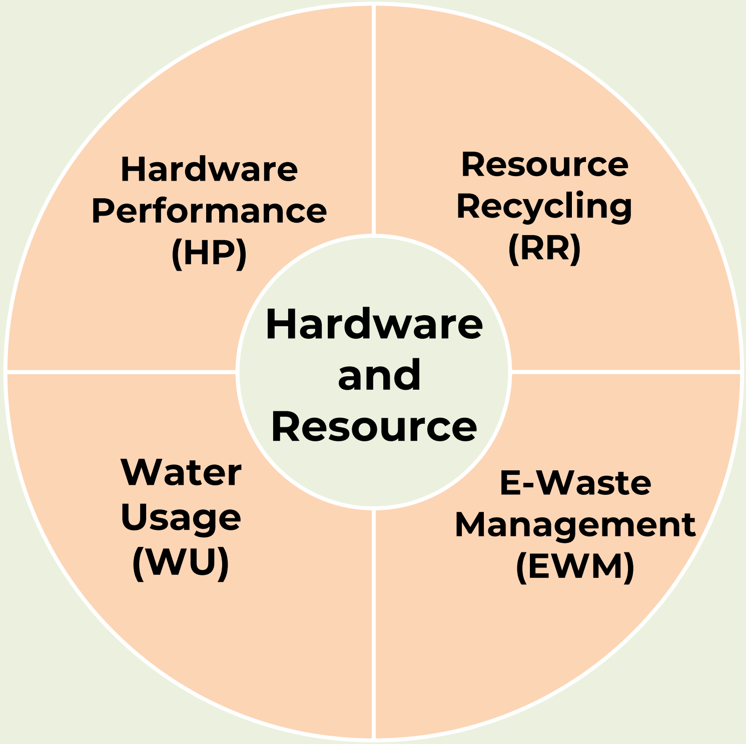
The advancement and efficiency of AI hardware.

HP = Operational performance / Energy consumption

- Resource Recycling (RR)

Rate of recycling and reuse of AI hardware and infrastructure.

RR = total value of recovered AI hardware and infrastructure / total value of all AI hardware and infrastructure



2 Hardware and Resource 2/2

Tracks resource usage and recycling for AI hardwares.

- Water Usage (WU)

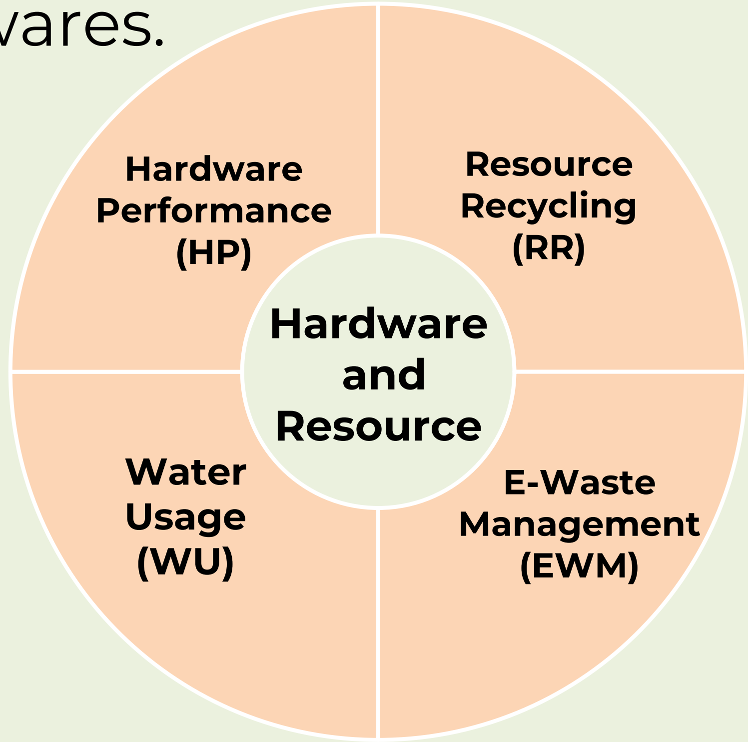
Water consumption for cooling and other processes, reflecting resource efficiency in AI operations.

WU = total water usage per day/ total AI task per day

- E-Waste Management (EWM)

Proper disposal and management of electronic waste, ensuring responsible handling of obsolete equipment.

EWM = ratio of proper disposal and management of electronic waste



3 Data Center Management

Measures the efficiency and sustainability of data centers.

- Power Usage Effectiveness (PUE)

Measures the ratio of total building energy usage to the energy used by IT equipment, indicating data center efficiency.

PUE = total data center energy usage / IT equipment energy usage

- Energy Reuse Effectiveness (ERE)

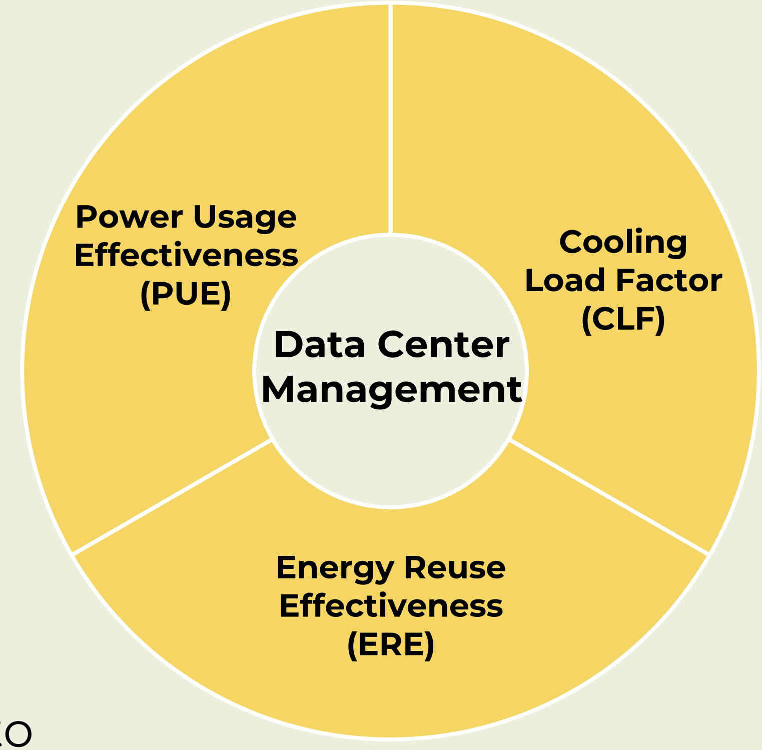
Measure the degree of energy reuse in a data center.

ERE = (Total energy − Recycled energy)/ IT equipment energy consumption

- Cooling Load Factor (CLF)

Indicates the ratio of energy consumption of cooling equipment to energy consumption of IT equipment.

CLF = Refrigeration equipment power consumption / IT equipment power consumption



4 Regulation and Compliance

Evaluates the extent to which companies adhere to regulatory standard.

- Global Initiative Participation (GIP):

Measure the company’s involvement in internationally recognized initiatives

GIP = Number of international initiatives in which the company participates / Total number of international initiatives in the relevant field

- Algorithm Sharing and Open Source (ASOS)

Measure the extent to which companies share and open source their AI algorithms and code bases.

ASOS = Number of algorithms shared publicly / Total number of algorithms

- Regulatory Response Time (RRT)

Measure how quickly the business responds to new regulations or policy changes.

RRT = Time from regulation release to enterprise adjustment /Total time before regulation implementation

5 Socio-economic and Green Impact 1/2

Assesses the environmental impact of AI in sector’s companies.

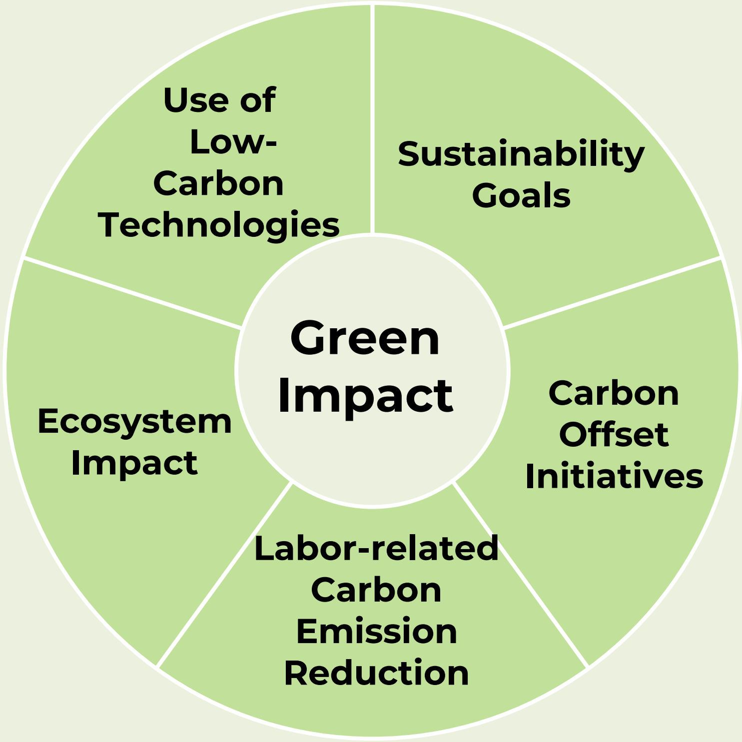
- Labor-related Carbon Emission Reduction (LCER)

The reduction of carbon emissions due to the use of AI models to replace traditional elements (labor, resources).

LCER= the total number of laborers being replaced by the AI Model(s) /the energy consumption of the AI model(s)

- Sustainability Goals

Alignment with international sustainability goals, showing a commitment to global environmental standards.



5 Socio-economic and Green Impact 2/2

Assesses the environmental impact of AI in sector's companies.

- Ecosystem Impact

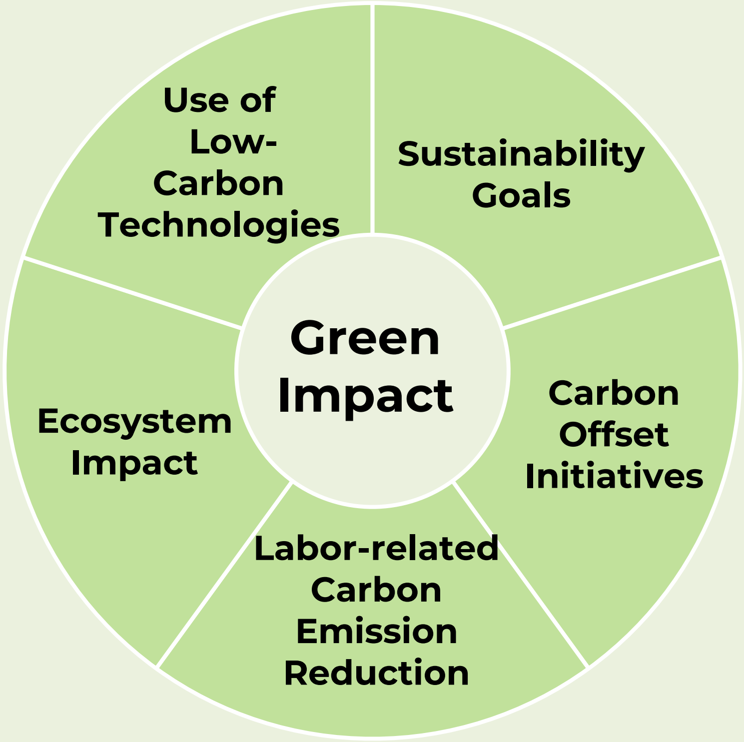
Measures to mitigate the impact on local ecosystems, reflecting a holistic view of environmental responsibility.

- Use of Low-Carbon Technologies

Adoption of technologies that reduce carbon emissions, indicating a preference for sustainable solutions.

- Carbon Offset Initiatives

Efforts to offset carbon emissions through projects and initiatives, demonstrating commitment to carbon neutrality.



UNIDO

BEIJING INSTITUTE OF TECHNOLOGY 北京理工大学 1940

SGG
School of Global Governance
国际组织创新学院

AIIM GLOBAL
Global Alliance on AI for Industry & Manufacturing

Green Artificial Intelligence

Pro. LIU Hao

Executive Dean, School of Global Governance

Chair, Space Safety Governance Committee, IAASS

Secretary General, JARUS