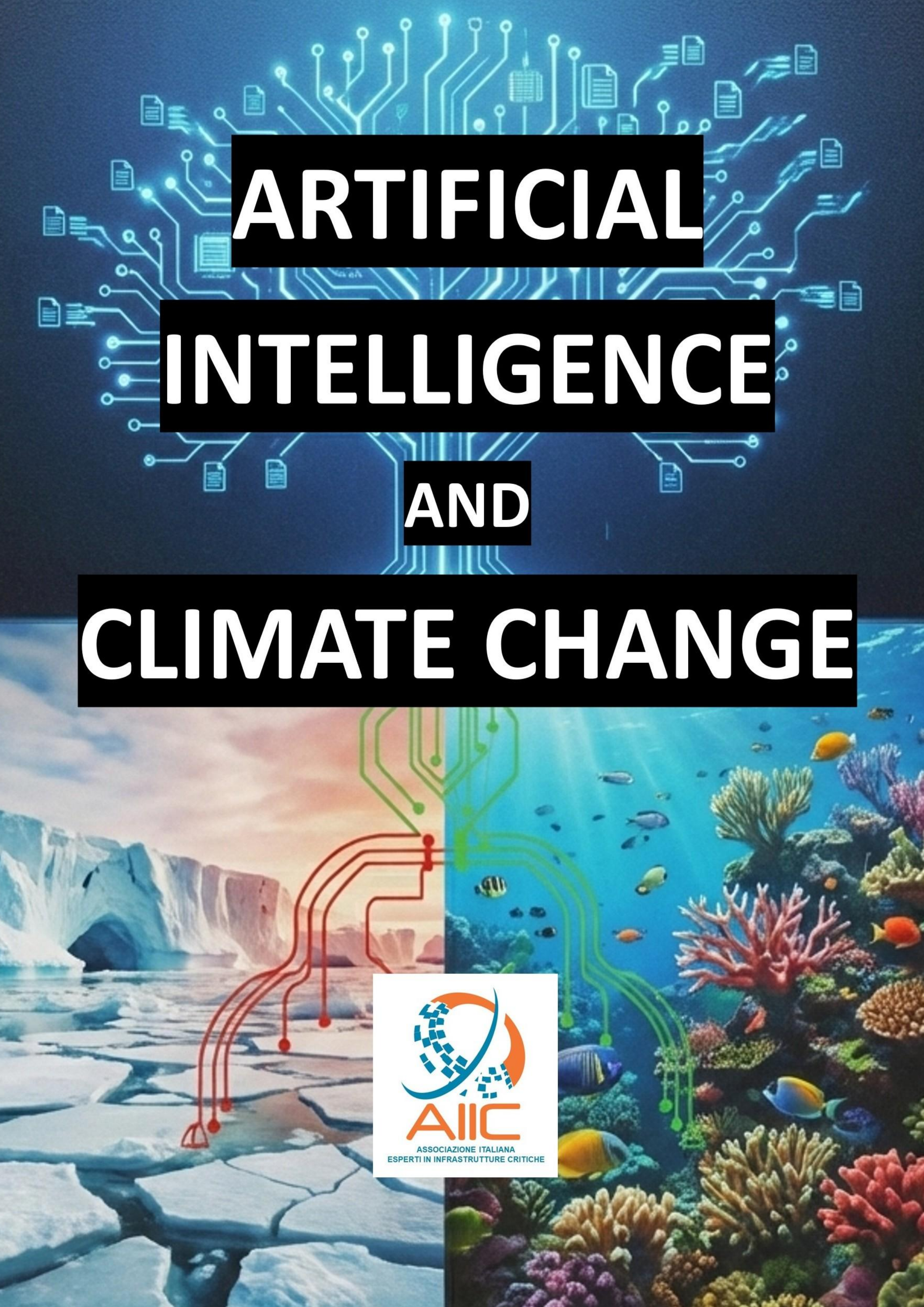
ARTIFICIAL INTELLIGENCE AND CLIMATE CHANGE



AIIC
ASSOCIAZIONE ITALIANA
ESPERTI IN INFRASTRUTTURE CRITICHE

Artificial Intelligence and Climate Change

ASSOCIAZIONE ITALIANA
ESPERTI IN INFRASTRUTTURE CRITICHE

Published by AIIC

February 2026

This document is the result of a joint project coordinated by Sandro Bologna and carried out with the contribution of: Silvano Bari, Glauco Bertocchi, Sandro Bologna, Gabriele Balzano, Luigi Carrozzi, Raffaella D'Alessandro, Tommaso Diddi, Elenio Dursi, Adriana Peduto, Beatrice Rosa, Alberto Stefanini, Cristina Turconi, Lorenzo Vandoni, Maria Beatrice Versaci.

The authors would like to thank the Italian Association of Critical Infrastructure Experts (AIIC) for its support and encouragement.

This version of the report represents the state of the art at the date of publication.

SUMMARY

1 INTRODUCTION *(Sandro Bologna)*

Artificial Intelligence (AI) is rapidly advancing frontiers in climate science, providing new insights into our understanding of the climate system and helping to transformation climate science into actionable information. This Report will explore innovative ways AI can enhance and accelerate climate action with a focus on decision-making and adaptation measures to foster resilience to climate change impacts. The Report provides a forum for interdisciplinary, cross-sectoral dialogue to facilitate cross-sector engagement, identify applications where AI can inform climate action at speed and scale, and consider how AI's broader societal impacts affect approaches to addressing climate change.

Artificial Intelligence is a powerful tool that has the potential to revolutionize the approach to many climate-related actions, from optimizing energy and water consumption to forecasting extreme weather events. As we leverage and expand the applications of AI technologies, it is essential to ensure that these models are accurate and reliable and to find ways to mitigate the climate impact of AI-based data center energy and water use. That is normally named **AI's environmental footprint**.

The Report seeks to foster ongoing discussions, shared learning, and simply coordination around emerging issues related to AI and climate change, including: how AI can combat climate change; the environmental impact of AI itself; and strategies for mitigating the impacts of AI energy consumption and climate effects.

There's a lot of talk about Artificial Intelligence applied to meteorology and climatology. However, while the benefits in meteorology are already visible and the initial results are encouraging, in climatology the situation is more complex. The possibilities are there, but it's not yet entirely clear how to best exploit them and with what guarantees.

Growing investment in AI is fueling a broader debate about sustainability, from the enormous demand for electricity to power data centers to the lifecycle emissions of the hardware needed to enable this technology. Globally, companies and governments are preparing to significantly increase energy consumption and raw material extraction to meet demand. On the other hand, technological evolution and innovations in AI optimization could lead to lower emissions than previously expected.

To understand how to map, measure, and mitigate the impacts of AI as they relate to data center electricity usage remain to be investigated, together with how metrics can reflect the trade-offs involved in AI's impacts on the economy and society, both locally and globally. Possible outcomes include describing the proportion of current data center electricity use and emissions related to AI; identifying a range of scenarios for potential increased electricity use and emissions; identifying options to mitigate potential increased electricity use and emissions; and catalyzing conversations among stakeholders on new approaches to understanding and meeting these challenges.

In addition to energy demands and related impacts on the grid, the rapid expansion of AI data centers can increase air pollution, affect water use and quality, and contribute to material use and environmental impacts from information technology equipment manufacturing and e-waste at the local, national, and global scales. Data center expansion can affect economies and communities through impacts on energy consumption, the environment, and workforces.

Running the AI chips to support society's demand for AI applications requires a great deal of energy, and a large portion of that energy demand comes from technologies used to keep information technology (IT) equipment cool. As the energy density of each server rack increases, this challenge only grows more acute. Cooling technologies are identified as an area for efficiencies and innovations to alleviate the load on the grid and impacts on communities.

AI has an urgent need for strong policies to combat data centers' power demands, which unfairly threaten climate goals and raise consumer electricity rates. As examples of the negative impacts AI data centers can have on local communities is the water consumed. In addition to these environmental impacts data center expansion may create serious economic tensions.

The emergence of AI has both raised concerns that AI-fuelled data center growth might fuel climate change and also raised expectations that AI applications in the energy sector could help reduce emissions by unlocking new optimisations and efficiencies. As over 100 countries – and the European Union – have targets to reach net zero emissions between 2030 and 2070, it is pertinent to explore what AI's impact on emissions could potentially be.

The ways in which Artificial Intelligence technologies are used will influence the future impact of AI data centers on electricity use and computing resource demands. AI presents opportunities to transform society and drive clean energy, but these benefits could also come at great costs. Utilities are arguing that they can only meet today's increased demand with an increase in natural gas investment, forgoing decarbonization strategies. Do we use this moment to move fast toward a decarbonized future … or do we use this moment as an opportunity to backtrack into natural gas?

This Report aims to address the above-mentioned issues, both:

- by exploring ways AI can be used to provide new insights into our understanding of the climate system;
- by providing new insights into strategies for mitigating the impacts of AI energy and water consumption with consequent climate effects, normally referred as **AI's environmental footprint**.

Fig. 1.1 – AI and Climate Change
(source: Internet)

2 AI AND CLIMATE CHANGE: DEFINITIONS AND APPLICATIONS

(Sandro Bologna, Alberto Stefanini)

2.1 Conceptual and Terminological Framework *(Alberto Stefanini)*

Understanding the interaction between Artificial Intelligence (AI) and climate change requires a shared conceptual and terminological basis. The two fields evolve rapidly and often use partly overlapping vocabularies. This section provides a concise glossary of the key terms used throughout the Report. Definitions are intentionally short, non-technical and aligned with standard scientific usage.

The purpose of this glossary is to ensure consistency across chapters and avoid conceptual ambiguity, particularly between meteorology, weather forecasting and climatology, an issue previously highlighted within the Working Group. These distinctions become essential when analysing the dual role of AI in:

1. strengthening climate science and
2. supporting the management of infrastructure impacts due to climate-driven stresses.

This glossary constitutes the common reference baseline for all subsequent chapters (3–9).

Glossary of Key Terms[1]

Artificial Intelligence (AI): technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy (IBM, 2024).

AI model: An artificial intelligence (AI) model in computer science is defined as software code that performs a function with the capability to achieve human-level accuracy in certain cognitive tasks, ranging from simple classification and regression to complex probabilistic decision making (Science Direct, 2026)

Carbon Intensity: a measure of how clean our electricity is. It refers to how many grams of carbon dioxide (CO_2) are released to produce a kilowatt hour (kWh) of electricity (National Grid (UK), 2022).

Data Center: a facility used to house computer systems and associated components, such as telecommunications and storage systems (Wikipedia, 2026).

Deep Learning (DL): Deep learning is a form of machine learning that enables computers to learn from experience and understand the world in terms of a hierarchy of concepts (Goodfellow, Bengio, & Courville, 2016)

---

[1] The definitions provided in this glossary are derived from internationally recognized standards and authoritative institutional or scientific sources. They are included to ensure terminological consistency across the report and do not constitute original normative definitions by AIIC.

Demand Response: balancing the demand on power grids by encouraging customers to shift electricity demand to times when electricity is more plentiful or other demand is lower, typically through prices or monetary incentives (IEA, the International Energy Agency, 2023).

Electric Grid (Power Grid): an interconnected network for the transmission and distribution of electricity from producers to consumers (UCTE, 2004).

Grid Flexibility: The ability of a power system to respond to variability and uncertainty in supply and demand (IEA, 2019).

Large Language Model (LLM): a neural network trained on very large text corpora to perform natural language understanding and generation tasks (Bommasani, Hudson, Aseli, & Russmann, 2021).

Load Forecasting: The process of predicting future electricity demand over different time horizons to ensure grid reliability and efficiency (ENTSO-E, 2021).

Machine Learning (ML): a branch of artificial intelligence focused on algorithms that improve performance through experience and data (Mitchell, 1997).

Power Usage Effectiveness (PUE): the ratio between total data-center energy consumption and the energy consumed by IT equipment (The Green Grid Assoc., 2012).

Retrieval-Augmented Generation (RAG): an AI architecture combining information retrieval systems with generative models to produce grounded and verifiable outputs (Lewis, et al., 2020).

State-Space Models (SSM): a class of models representing systems through latent states evolving over time, recently adapted for efficient sequence modelling in A (Gu, Goel, & Ré, 2022).

Water Usage Effectiveness (WUE): a metric expressing the amount of water used by a data center relative to its IT energy consumption (Azevedo, Belady, & Pouchet, 2011).

Workload Shifting: the practice of relocating computational tasks across time or geographic locations to optimise energy use, cost or carbon intensity (IEA, 2023).

Sources

Azevedo, D., Belady, C., & Pouchet, J. (2011, March 1). *WP#35 - Water Usage Effectiveness (WUE): A Green Grid Data Center Sustainability Metric.* Tratto da The Green Grid: https://archive.thegreengrid.org/en/resources/library-and-tools/238-WP

Bommasani, R., Hudson, D. A., Aseli, E., & Russmann, R. B. (2021). *On the Opportunities and Risks of Foundation Models.* Ithaca, NY: Cornell Univ.

ENTSO-E. (2021). *WEATHER PROCESS AND ENERGY PROGNOSIS Implementatin Guide.* Brussels: European Network of Transmission System Operators for Electricity. Tratto da https://eepublicdownloads.entsoe.eu/clean-documents/EDI/Library/cim_based/Weather_Process_Energy_Prognosis_IG_v1.3.pdf

Goodfellow, J., Bengio, Y., & Courville, A. (2016). *Deep Learning* . Cambridge, MA: MIT Press.

Gu, A., Goel, K., & Ré, C. (2022, Aug. 5). *Efficiently Modeling Long Sequences with Structured State Spaces.* Tratto da Cornell University: https://arxiv.org/abs/2111.00396

IBM. (2024). *What is artificial intelligence (AI)?* Tratto da IBM: https://www.ibm.com/think/topics/artificial-intelligence

IEA. (2019). *Status of Power System Transformation 2019.* Brussels: IEA - the International Energy Agency. Tratto da https://www.iea.org/reports/status-of-power-system-transformation-2019

IEA. (2023). *Electricity Grids and Secure Energy Transitions.* Paris: IEA - International Energy Agency. Tratto da https://www.iea.org/reports/electricity-grids-and-secure-energy-transitions

IEA, the International Energy Agency. (2023). *Demand Response.* Tratto da IEA: https://www.iea.org/energy-system/energy-efficiency-and-demand/demand-response

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., . . . Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *34th Conference on Neural Information Processing Systems (NeurIPS 2020),* (p. 1-16). Vancouver, Canada: Facebook AI Research; University College London; New York University. Tratto da https://proceedings.neurips.cc/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf

Mitchell, T. (1997). *Machine Learning.* Pittsburgh (PA): McGraw-Hill.

National Grid (UK). (2022, May 17). *What is carbon intensity?* Tratto da National Grid: https://www.nationalgrid.com/stories/energy-explained/what-is-carbon-intensity

Science Direct. (2026). *Artificial Intelligence Model.* Tratto da Science Direct: https://www.sciencedirect.com/topics/computer-science/artificial-intelligence-model

The Green Grid Assoc. (2012). *PUE: A Comprehensive Examination of the Metric. https://datacenters.lbl.gov.* The Green Grid. Tratto da https://datacenters.lbl.gov/sites/default/files/WP49-PUE%20A%20Comprehensive%20Examination%20of%20the%20Metric_v6.pdf

UCTE. (2004). *Operation Handbook – Glossary.* Brussels: UCTE - union for the co-ordination of transmission of electricity. Tratto da https://www.google.com/search?q=ucte&rlz=1C1CHBD_itIT1179IT1189&oq=UCTE&gs_lcrp=EgZjaHJvbWUqBwgAEAAYgAQyBwgAEAAYgAQyBwgBEAAYgAQyDQgCEC4YrwEYxwEYgAQyBwgDEAAYgAQyCQgEEAAYChiABDIHCAUQABiABDIHCAYQABiABDIHCAcQABiABDIJCAgQABgKGIAEMgkICRAAGAoYgATSAQg1MjE5ajBqN6g

Wikipedia. (2026, Jan. 10). *Data center*. Tratto da Wikipedia, the free encyclopedia: https://en.wikipedia.org/wiki/Data_center

2.2 What We Mean by Climate Change *(Sandro Bologna)*

Climate change refers to significant, long-term alterations in the Earth's climate system—encompassing temperature, precipitation, atmospheric circulation, and extreme events—driven primarily in the modern era by anthropogenic emissions of greenhouse gases (GHGs). According to the Intergovernmental Panel on Climate Change (IPCC), global surface temperature has increased by approximately **1.1–1.2°C** since 1850–1900 (IPCC, 2021)[2]. This trend parallels the unprecedented rise in atmospheric CO_2 concentrations from **~280 ppm** pre-industrially to **over 420 ppm** today, levels not observed in the past 2 million years.

Radiative forcing analyses indicate that human activities—especially fossil fuel combustion and land-use change—have contributed over **95%** of the net positive forcing since 1750 (IPCC, 2021). These changes manifest in multiple Earth system components. The cryosphere has experienced accelerated

---

[2] IPCC. (2021). *Climate Change 2021: The Physical Science Basis*. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

mass loss, with Greenland and Antarctica shedding hundreds of gigatons of ice annually. Global mean sea level has risen by approximately **20 cm** since 1900, with rates tripling in recent decades. Oceans have absorbed over **90%** of excess heat and undergone measurable acidification, with surface pH dropping by about 0.1 units since pre-industrial times.

Climate change also alters the frequency and intensity of extreme weather events. Heatwaves have become significantly more likely and more severe, with attribution studies demonstrating that many would have been extremely unlikely without human influence. Heavy precipitation events have strengthened due to increased atmospheric moisture capacity—about **7% more water vapor per °C** of warming.

In summary, climate change is a measurable, data-supported shift in global and regional climate behavior, driven primarily by human emissions and characterized by rising temperatures, altered hydrological cycles, sea-level rise, cryospheric decline, and more frequent extreme events.

AIIC has already covered the topic of Climate Change and its impact on the Resilience of Critical Infrastructure in the year 2023 (AIIC, 2023)[3].

2.3 AI for what? *(Sandro Bologna)*

Artificial intelligence is often touted as a crucial lever for the ecological transition. And rightly so: thanks to its ability to optimize electricity grids, industrial processes, and logistics, AI can reduce waste and inefficiencies along the entire value chain. At the same time, however, each generative model requires highly energy-intensive physical infrastructure, powered by large quantities of electricity and water for cooling.

The paradox is here: the same technology that promises to make production systems more sustainable, at the same time produces new climate-altering emissions if it is not embedded in a clear, measurable, and transparent **E**nvironmental, **S**ocial, **G**overnance (ESG) strategy. This is what emerges from the different scenarios developed by different studies worldwide, among others **ClimateSeed**[4], a startup specializing in decarbonization software and consulting, which has highlighted the "dark side" of AI based on Italian and international data.

AI is by no means "immaterial." Every digital interaction with a model like ChatGPT or Gemini involves passing through data centers working at full capacity, supported by specialized chips like GPUs and TPUs, and cooling systems that consume significant water resources.

Studies mentioned by ClimateSeed indicate that some queries can consume up to 0.43 Watt-hours (Wh), an amount of energy comparable to keeping an LED light on for almost a minute. For longer or more complex requests, consumption can rise to 2–4 Watt-hours per single interaction. These seemingly small numbers become significant when multiplied by millions of daily uses around the world.

This energy demand fits into an already critical situation: according to projections from the **I**nternational **E**nergy **A**gency (IEA), global data center electricity consumption could rise from 460 Tera-watt-hours (TWh) in 2022 to over 1,000 TWh by 2026, more than doubling in just four years

---

[3] AIIC. (2023). Resilienza delle Infrastrutture Critiche e Cambiamenti Climatici (in italian) and Proceedings of CRITIS 2024, Springer LNCS 15549.

[4] ClimateSeed, https://climateseed.com/

and driven by the spread of generative AI. The TWh is the unit of measurement of energy that corresponds to one thousand billion Watt-hours, or 10^{12} Watt-hours.

It's no surprise, then, that major digital platforms are seeing a significant increase in their climate footprint. Sustainability reports show that, since 2020, Microsoft has seen a 29% increase in overall emissions, while Google recorded a 48% increase in 2023 compared to 2019. Even more significant is the data on indirect emissions, for Microsoft, Amazon, and Meta, there is a 150% increase in emissions between 2020 and 2023: *Greening Digital Companies: Monitoring emissions and climate commitments,* Report 2025[5]. These emissions include all emissions along the value chain arising, for example, from digital services, external infrastructure, suppliers, and product use. In other words, the long-term impact of the digital revolution and AI is impacting the entire ecosystem, not just the operational boundaries of big tech. Also if AI impact has the greater value.

If we narrow our focus to the Italian context, in 2023, overall electricity consumption reached 509.7 GWh, with Lombardy, Lazio, Emilia-Romagna, and Piedmont alone accounting for almost 85% of the total. These figures reflect an increasingly central digital infrastructure for the productive fabric, but they raise pressing questions about how to ensure this growth is compatible with national and European climate objectives.

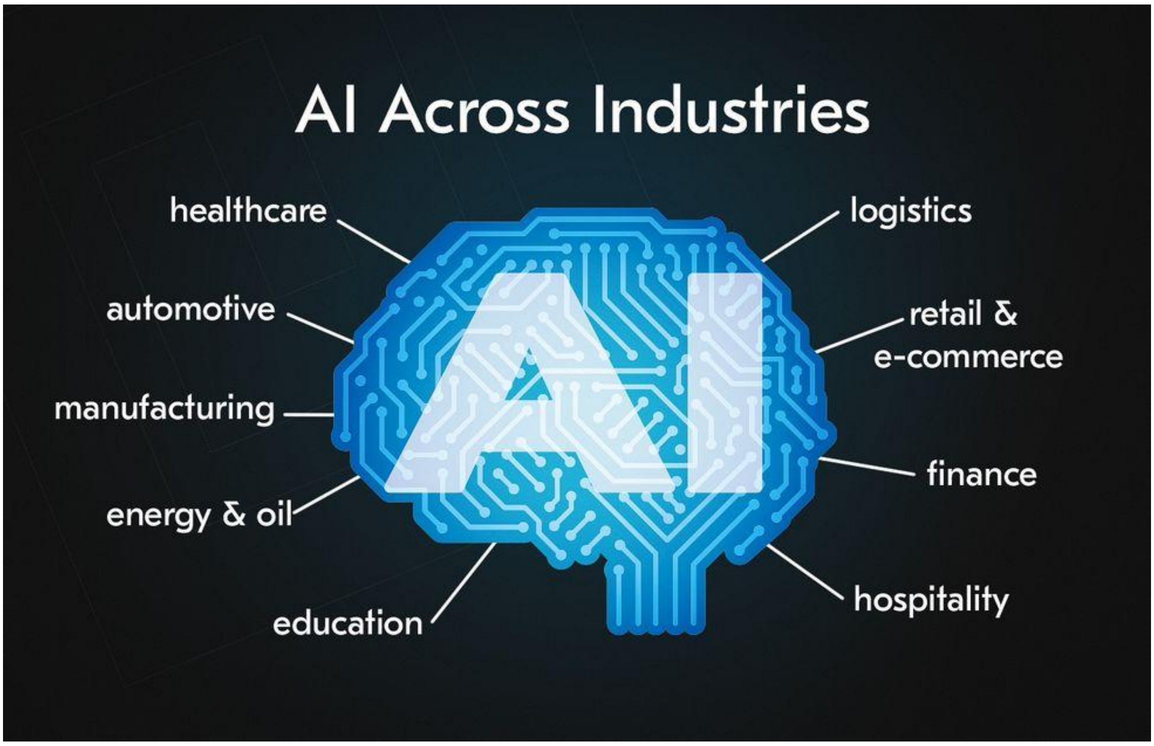


Fig.2.1 Nine possible AI use in major industries
(source Internet: Acropolium, March 17, 2025: https://acropolium.com/blog/ai-use-cases-in-major-industries-elevate-your-business-with-disruptive-technology/)

---

[5] Greening Digital Companies: Monitoring emissions and climate commitments, Report 2025, https://www.itu.int/en/ITU-D/Environment/Pages/Publications/GDC-25.aspx.

2.4 Machine Learning *(Sandro Bologna)*

Machine learning (ML), a central branch of artificial intelligence, comprises computational methods that enable algorithms to learn from data and generalize beyond explicitly programmed instructions. Through approaches such as supervised, unsupervised, reinforcement, and deep learning, ML systems extract complex patterns from large, multidimensional datasets and use them to generate robust predictions. In climate science, ML has become an essential predictive tool, with a particularly significant application in **forecasting extreme weather events**, where ML models improve the precision and timeliness of predictions for phenomena such as storms, droughts, and heatwaves. As a mitigation tool aimed at reducing the human footprint on the climate, ML also plays a critical role in **improving the energy efficiency of buildings and industrial processes**, using real-time data to optimize heating, cooling, lighting, and manufacturing operations in ways that substantially lower greenhouse gas emissions. Together, these two application domains illustrate ML's dual importance in anticipating climate-related risks and actively reducing the drivers of climate change.

ML, like any technology, does not always make the world a better place—but it can. In the fight against climate change, ML has significant contributions to offer across domain areas. ML can enable automatic monitoring through remote sensing (e.g., by pinpointing deforestation, gathering data on buildings, and assessing damage after disasters). It can accelerate the process of scientific discovery (e.g., by suggesting new materials for batteries, construction, and carbon capture). ML can optimize systems to improve efficiency (e.g., by consolidating freight, designing carbon markets, and reducing food waste). And it can accelerate computationally expensive physical simulations through hybrid modeling (e.g., climate models and energy scheduling models). Applying ML to tackle climate change has the potential both to benefit society and to advance the field of ML. Meaningful action on climate problems requires dialogue with fields within and outside computer science and can lead to interdisciplinary methodological innovations, such as improved physics-constrained ML techniques.

In recent years, **machine learning (ML)** has been recognized as a broadly powerful tool for technological progress. Despite the growth of movements applying ML and **artificial intelligence (AI)** to problems of societal and global good, there remains the need for a concerted effort to identify how these tools may best be applied to tackle climate change. Many ML practitioners wish to act, but are uncertain how. On the other side, many fields have begun actively seeking input from the ML community, see **Tackling Climate Change with Machine Learning (2019)**[6]

Technology is not in itself enough to solve climate change, nor is it a replacement for other aspects of climate action such as policy. Many technological tools useful in addressing climate change have been available for years but have yet to be adopted at scale by society. While we hope that ML will be useful in accelerating effective strategies for climate action, humanity also must decide to act. ML is not a silver bullet.

The nature of climate-relevant data poses challenges and opportunities. For many of the applications we identify, data can be proprietary or include sensitive personal information.Where datasets exist, they may not be organized with a specific task in mind, unlike typical ML benchmarks that have a clear objective. Datasets may include information from heterogeneous sources, which must be integrated using domain knowledge. Moreover, the available data may not be representative of global use cases. For example, forecasts of electricity demand based on a dataset from the US will not necessarily generalize to India, where patterns of demand may be different. Tools from transfer

---

[6] Tackling Climate Change with Machine Learning (2019), https://www.researchgate.net/publication/366788501_Tackling_Climate_Change_with_Machine_Learning

learning and domain adaptation will likely prove essential in low-data settings. For some tasks, it may also be feasible to augment learning with carefully simulated data. Of course, the best option if possible is always more real data; we strongly encourage public and private entities to release datasets and to solicit involvement from the ML community. For those looking to use ML to help tackle climate change, we strongly recommend to visit the **Climate Change AI initiative**[7].

2.5 Large Language Models *(Sandro Bologna)*

The climate crisis, exacerbated by fossil fuel burning, deforestation, and industrial processes, poses a grave global threat. Its impacts range from rising sea levels to intensified weather events and biodiversity loss. Addressing this crisis is urgent, prompting widespread efforts to reduce greenhouse gas emissions and adopt more sustainable practices. In this context, large language models (LLMs) like GPT-5 can play a vital role in raising awareness and educating the public about the climate emergency. LLMs have the potential to reach a global audience and provide accurate, up-to-date information on the causes and consequences of the climate crisis. They can also engage in discussions with users, answering questions and addressing concerns related to climate change. Once for all, we recommend to have a look to **Climate Change from Large Language Model (2024)**[8]**.**

Extracting climate crisis knowledge from LLMs is a nontrivial task due to limited interpretability. The approach suggested in the paper aims to improve understanding and evaluation of this knowledge, enabling a more human-interpretable assessment of their capabilities. The parameter knowledge in the text is symbolized through elaborately designed prompts.

The second challenge is evaluating knowledge related to the climate crisis. Prior studies have primarily relied on perplexity to assess generated content, but this approach falls short in accurately capturing knowledge from a human cognitive perspective. Certain research efforts have resorted to human evaluation, an approach that can be both costly and time-consuming. Other studies have attempted to utilize classifiers to grade answers, yet these methods prove inadequate for accurately evaluating knowledge pertinent to the climate crisis. To address this issue, the paper proposes a method to automatically evaluate the knowledge of LLMs related to the climate crisis by evaluating the quality of questions and answers.

---

[7] Climate Change AI initiative, https://www.climatechange.ai/

[8] Climate Change from Large Language Model (2024), https://arxiv.org/pdf/2312.11985

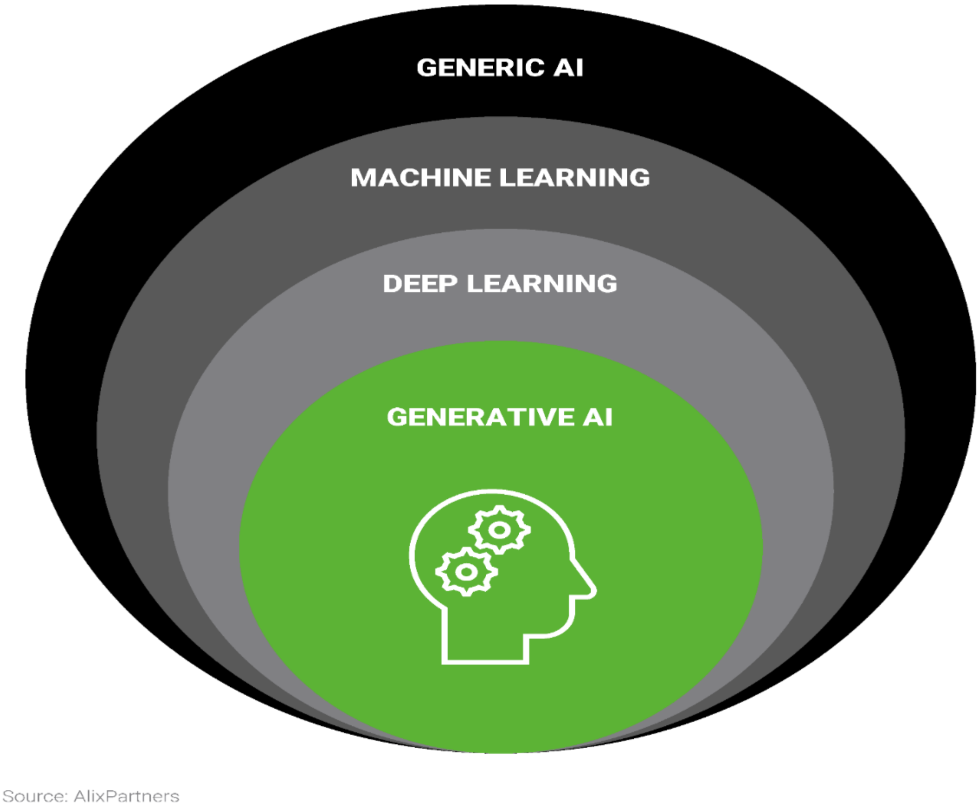


Fig.2.2 Different AI models
(source Internet: AlixPartners - AI, Machine Learning and LLMs, February, 2025: https://www.alixpartners.com/insights/102jysq/investigating-decision-making-algorithms-part-3-ai-machine-learning-and-llms/)

3 AI TECHNOLOGY EVOLUTION AND CLIMATE CHANGE *(Luigi Carrozzi)*

3.1 Foreword

The relation between AI and Climate Change has a systemic nature. That is, there are several intertwined factors within a complex system that include not only technological but also economic and social factors. Having the objective of taking an overall picture of the relations existing between the development of AI technology and the relevant impact that this may have on climate variations would be the case, as it possible, to outline the relevant factors that may influence the above-mentioned relation.

According to this approach, in the following are introduced what we may consider some of the most relevant factors underlying this relation and how they may impact climate change. Typically, the correlation between the growth of AI and climate change is related to the huge amount of energy (electricity) necessary to power AI systems. And energy consumption is a primary driver of climate change, being a source of carbon dioxide emissions and hence contributing to global carbon intensity growth. Related to the growth of AI systems there are also other important impacts on environmental sustainability like water consumption used to cool the huge data centers running AI software, and the land consumption required to place such computing infrastructures. In the following we will focus only on the impact of climate change due to energy consumption.

3.2 The factors of AI growth impacting on energy consumption and climate change

As above mentioned, we may identify different major technical and non-technical factors of AI growth and how these factors impact energy consumption. Some of the major factors characterizing the AI evolution that shape the impact on climate change due to energy consumption may be identified in the following.

1. Growth of processing power given by hardware components and architectures dedicated to computational tasks of AI (Computational power)
2. Development of new AI optimized techniques reaching higher level of algorithmic efficiency
3. Data availability
4. AI market growth
5. Private and public sector investments in AI developments
6. Governments' involvements on research and developments of AI as a strategic asset
7. Rules and legislative frameworks governing the development and the use of AI

All these factors contribute differently to the growth of AI. And the resulting mode and characteristics of AI growth impacts also differently on energy consumption and consequently on climate change.

Here follows an analysis of those factors and how they contribute to rise climate change effects due to rise of energy consumption. This analysis has the objective to present an *exercise* trying to carry out a more in-depth, systemic analysis on how artificial intelligence evolution may impact on energy consumption.

These results are presented discussing how each factor contribute (positively or negatively) to the rise of energy consumption and are based on the best knowledge of the current state of the art and trends as gathered by the author.

3.3 Analysis of Factors' influence on energy consumption

IEA (the International Energy Agency)[9] finds that the global electricity consumption for data centers is projected to double to reach around 945 TWh by 2030 and from 2024 to 2030, and it will grow by around 15% per year, more than four times faster than the growth of total electricity consumption from all other sectors. In the following we will try to identify the way in which the above-mentioned factors will influence such energy consumption.

Growth of processing power given by hardware components and architectures dedicated to computational tasks of AI (Computational power)

From the hardware side, the following should be considered:

- the evolution of processing units power such as CPU, GPU (Graphical Processing Units) and ASIC (Application-Specific Integrated Circuit) are devoted to support the computational power required by the unending growth of AI machine learning workloads (in particular, about ASIC it's worth of note that as set of electronic circuits integrated on a single chip and designed for specific applications rather than for general-purpose use, they may perform some specific functions faster than general purpose CPUs and GPUs and have also the characteristics to be energy-efficient);
- the large, distributed data centers where the processing units may be connected as a single machine;
- EDGE and local AI computing, that are contributing to a more rational use of energy;
- the game changer of the forthcoming quantum computing and optical computing.

***Driving Factor*:** Greater is the processing power the higher is the energy consumption
***Balancing Factor**:* new efficient technologies may enable less energy consumption
***Counterbalancing Factor**:* If hardware rise performance at an incredibly high pace, on the other side more the computing power is available, the greater the use that AI algorithms will make of it, regardless of the levels of efficiency that can be achieved. (see Wirths's law[10], also known as "*What Intel giveth, Microsoft taketh away*").

Development of new AI optimized techniques reaching higher level of algorithmic efficiency

Algorithmic efficiency in the context of Artificial Intelligence and Machine earning refers to the evaluation of how effectively an algorithm utilizes computational resources to solve a problem.[11]
It is often measured in terms of time complexity (how the execution time increases with input size) and space complexity (how much memory the algorithm needs during execution).

---

[9] https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai
[10] https://people.inf.ethz.ch/wirth/Articles/LeanSoftware.pdf
[11] https://www.opentrain.ai/glossary/algorithmic-efficiency

Efficient algorithms make possible to handle larger datasets, execute faster, require less computational power and make them more suitable for real-world applications. The search for algorithmic efficiency is related to balancing speed and computational resource consumption.

A key factor is also to enable the scalability and feasibility of ML models, especially when dealing with big data and complex computations. For example, the "*k-Nearest Neighbors algorithm*" (k-NN) is a machine learning algorithm used for both classification and regression known for its simplicity in classification tasks. However, its efficiency can be significantly impacted by the size of the dataset. There are specific techniques to optimize the efficiency of this algorithm such as other deep learning algorithms, that can reduce the model size and computational requirements.
In general, it's worth of note that the training phase of AI models is highly energy intensive and interacting with an LLM could likely cost 10 times more than a standard keyword search[12]

***Driving Factor*: **Greater is the algorithmic efficiency the lower is the energy consumption
Balancing Factor: new efficient algorithms lead the way to profit of the resources made available to perform more complex tasks which inevitably leads to greater energy consumption

Data availability

The issue of the availability of data to train AI systems is crucial. Better say, "data scarcity"[13] is the biggest challenge faced by Artificial Intelligence when the lack of data may become a significant impediment to train machines to real-world processes. This problem is typically managed using *Synthetic Data* or machine learning techniques such as *Transfer Learning* or *Few-Shot Learning*. *Transfer Learning* uses pre-trained models from a machine learning task to improve performance and generalizability across a related task; *Few-Shot Learning* is a machine learning framework, typically used to train models for classification tasks when proper training data is scarce, where an AI model learns through a very small number of labeled examples.
As far as it concerns the relation between the amount of data used by an AI system and energy consumption may be synthetized as follows: more data typically implies larger models and more computing resources during training operations with higher level of energy consumption.

***Driving Factor*: **Greater is the amount of data available the larger is the energy consumption.
Balancing factors: Techniques aimed to face data scarcity drive to fewer computing resources during training operations and hence lower level of energy consumption.

AI market growth

The Artificial Intelligence market is expected to have a significant growth until 2030. A UN Trade and Development (UNCTAD) report projects that the global AI market rises from $189 billion in 2023 to $4.8 trillion by 2033[14].
The increasing adoption of AI across industries as pushed by the evolution of AI algorithms, the development of computing infrastructure and the growing investment in AI research and development, drive the growth. AI is progressively becoming an integral part of business operations and consumer-facing applications[15].
***Driving Factor*: **The market demand (industry and consumers) drive consistently the overall growth of AI systems supported implying consequently higher level of energy consumption

---

[12] Alex de Vries"*The growing energy footprint of artificial intelligence*" https://doi.org/10.1016/j.joule.2023.09.004
[13] Hemn Barzan Abdalla, Yulia Kumar, Jose Marchena, Stephany Guzman, Ardalan Awlla, Mehdi Gheisari, Maryam Cheraghy, The Future of Artificial Intelligence in the Face of Data Scarcity, Computers, Materials and Continua, Volume 84, Issue 1, 2025,
[14] https://unctad.org/publication/technology-and-innovation-report-2025
[15] https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide

Here follows a picture representing the AI market growth compared with other technologies from a report of the UN Trade and Development organization.

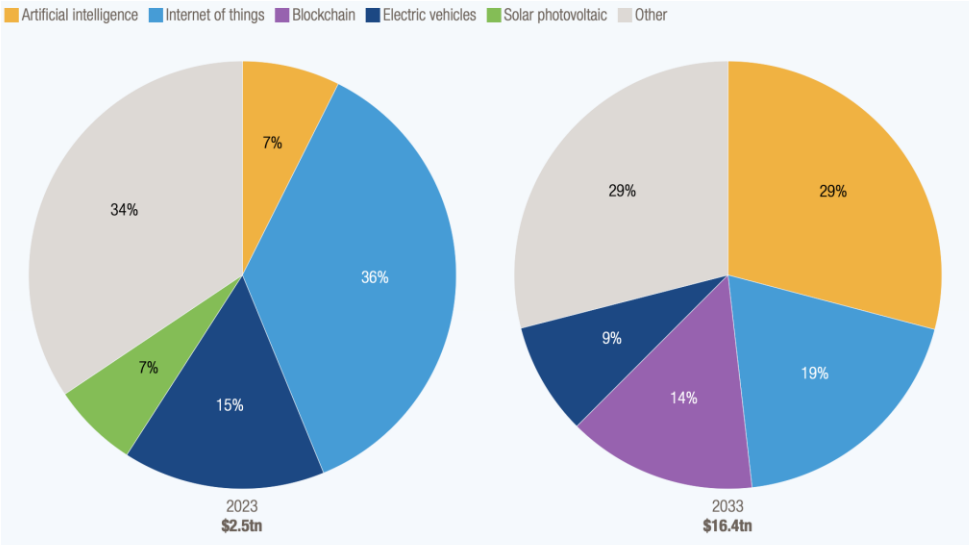


Fig. 3.1 Estimated frontier technology market size and share of selected technologies, 2023 and 2033
(Source: *UN Trade and Development UNCTAD)*[16]

Private and public sector investments on AI developments

According to the Stanford University's "2025 AI Index Report"[17] private investments in artificial intelligence rose to 109.1 billion dollars, almost 12 billion dollars to 9.3 billion dollars in China and 24 billion dollars to 4.5 billion dollars of UK. Artificial intelligence generated a significant growth, with 33.9 billion dollars in global private investment, with an increase of 18.7% until 2023.
It goes with that the high level of private investments pushes the value chain of AI growth also through Venture Capital firms that are consistently financing the new digital world shaped by AI technology. Public investment in AI is growing as well. The European Commission announced recently an investment of €1.98 billion to advance the digital transition significantly financing also programs involving AI projects. The overall growth of investment inevitably pushes development and use of AI that are direct factors of energy consumption. Here follows a picture describing the amount of private investment by geographic area in 2024.

Driving Factor: Greater is the amount of investment in AI the larger is the energy consumption.

Balancing factors: The enthusiasm in AI technology may have a contraction due to a paved "stock market bubble".

---

[16] https://unctad.org/news/ai-market-projected-hit-48-trillion-2033-emerging-dominant-frontier-technology
[17] https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf

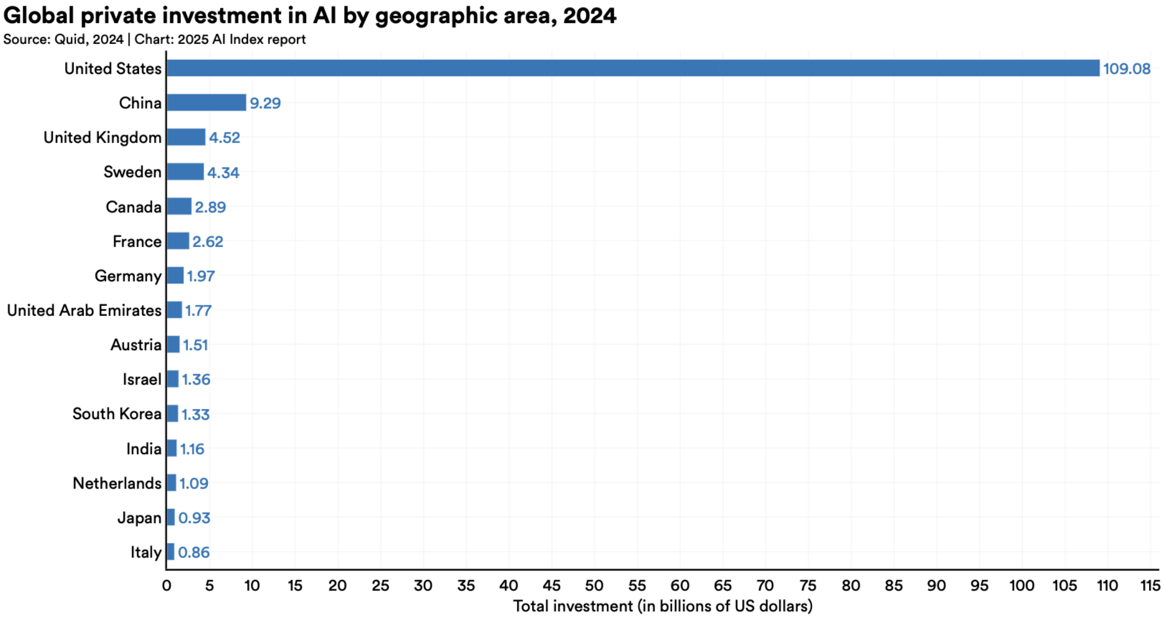


Fig. 3.2 - Global private investment in AI by geographic area, 2024
(Source: Artificial Intelligence Index Report 2025 – Stanford University – Human centered Artificial Intelligence)

Governments' involvements on research and developments of AI as a strategic asset

Governments often consider AI as a strategic asset also in terms of strategic, geopolitical, digital sovereignty power, pushing the development of AI systems in all industry sectors, including dual-use application.

Driving Factor: Greater is the importance of Governments in the strategic use of AI the higher is the development and use of AI systems with the unavoidable impacts on energy consumption.
Reinforcing factor: Competitiveness among States may give the way to a reinforcing loop leading to further overall energy consumption.

Rules and legislative frameworks governing the development and the use of AI

The development of Legal frameworks, Policy and Standards governing the development and use of AI system may be considered as a constrains on AI technology developments. But good rules may also enable trust in pushing trustworthiness in a so powerful but also risky technology.

Driving Factor: Greater is the development and the enforcing of rules for providers and deployers of AI system, the higher are the level of constraints in development and use of AI and consequently, lower the overall energy consumption.
Balancing Factor: The higher level of trustworthiness given by Legal frameworks, Policy and Standards may facilitate higher level of market confidence on quality and trustworthiness of AI technology supporting further AI development and hence overall energy consumption

4 AI AND ENVIRONMENTAL IMPACT: REGULATIONS AND STANDARDS *(Raffaella D'Alessandro, Adriana Peduto, Beatrice Rosa, Cristina Turconi)*

In order to understand how to properly address the development and adoption of AI in relation to climate change, let's look at how regulations and standards can support us. In this chapter we identify what the applicable European and national regulation say on this matter and depict the state of the art of the applicable standards.

4.1 European Regulations (*Adriana Peduto, Beatrice Rosa)*

Regulation (EU) 2024/1689 (the AI Act) undoubtedly represents one of the first comprehensive attempts to regulate Artificial Intelligence (AI) in a systemic manner, integrating requirements relating to safety, fundamental rights and risk governance, while addressing—albeit only at a nascent and still distinctly embryonic stage—the environmental and energy-sustainability dimension.
Be that as it may, while the AI Act seems to address this dimension only in a limited manner, the European Union's engagement with the relationship between AI, environmental sustainability and the climate transition has been far more extensive. The establishment of expert groups and dedicated funding instruments underscores the EU's commitment to fostering a sustainable digital transformation where AI serves as a powerful accelerator for achieving climate neutrality by 2050, distinguishing its comprehensive and integrated strategy from more fragmented global efforts. forming part of a broader strategic trajectory initiated with the European Green Deal in 2019[18] and subsequently developed through a substantial body of policy documents, institutional analyses and regulatory initiatives. Within this framework, the European Parliament's study The Role of Artificial Intelligence in the European Green Deal (May 2021)[19] constitutes a crucial conceptual turning point, as it clearly articulated the dual nature of AI, that has been later regarded also in the EU AI Act, namely AI as a technology with significant potential to enable sustainability, yet simultaneously a source of autonomous environmental and energy-related risks.

Keeping in mind this perspective, the key point of the EU strategy regarding the AI-Climate change relationship has always been very straightforward: without targeted mandatory regulatory intervention, AI risks undermining the very climate-related objectives it is expected to support. Hence, it is essential to develop rules and methodologies able to properly assess the environmental impact of AI systems, govern their systemic effects and steer technological innovation towards objectives consistent with climate neutrality.
Despite this well-acknowledged need, starting from the EU Green Deal back in 2019 and culminating with the current AI Act, these recommendations have always been explicitly framed as mere policy and soft-law guidance instruments, rather than as immediately binding legal obligations.

As a matter of fact, the current AI Act[20] actually proposes a very soft approach regarding the impact of AI usage in environmental terms, an approach that is surely innovation-driven, but marked by a

---

[18] The European Green Deal Text is available here: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0640.

[19] European Parliament, Directorate-General for Internal Policies of the Union and Öko-Institut e.V, The role of artificial 0ointelligence in the European Green Deal, European Parliament, 2021, https://data.europa.eu/doi/10.2861/882830.

[20] The AI Act full text is available here: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

striking lack of immediate binding force. Rather than imposing concrete obligations, enforceable limits, or sanction-backed requirements, the AI Act consistently defers environmental governance to future developments, Member States' initiatives, technical standardisation processes, and, above all, voluntary instruments.

Taken together, Recitals 4, 27, 121, 142, 165 and 174 of the AI Act articulate an ostensibly coherent vision in which the environmental and energy impacts of AI are not treated as ancillary issues, but as structurally relevant components of the European Union's AI policy. Through these provisions, the EU legislator acknowledges three interrelated dimensions that actually shape the entire EU strategy regarding the relationship between AI and Climate Change: first, the potential environmental benefits of AI alongside its growing energy and resource costs; second, the incorporation of sustainability into the ethical and operational principles of trustworthy AI in line with the EU Green Deal well-known strategy; and third, the reliance on dynamic and progressive regulatory mechanisms capable of evolving alongside technological development.

This regulatory posture is already apparent in Recital 4, which frames AI as a potentially transformative tool for energy efficiency, environmental monitoring, and climate change mitigation. While this Recital clearly situates AI within the objectives of the European Green Deal, it does so at the level of political orientation rather than legal command. The underlying assumption is that AI can support sustainability, provided it is designed and deployed responsibly; however, no immediate duty is imposed on providers or deployers to ensure that such responsibility is actually realised.[21]

Recital 27 further reinforces this approach by incorporating the principle of social and environmental well-being drawn from the 2019 Ethics Guidelines for Trustworthy AI. Sustainability thus becomes an interpretative benchmark and a normative reference point, but not a directly enforceable obligation. Its primary function is to inform the development of future codes of conduct and to guide the application of the Regulation, rather than to generate autonomous legal duties in the present.[22]

The centrality of voluntary governance instruments becomes even more evident in Recitals 121 and 165. While Recital 121 situates codes of good practice within a broader culture of transparency and accountability, Recital 165 explicitly promotes the adoption of voluntary codes of conduct addressing environmental sustainability across the AI value chain, including for systems that do not qualify as high-risk.[23]
This choice reflects a deliberate legislative strategy: rather than mandating environmental requirements, the Union encourages operators to self-regulate, thereby extending responsibility in principle while avoiding binding constraints in practice.
To date, however, these codes remain largely underdeveloped, weakly institutionalised, and scarcely discussed outside specialised policy circles.

Recital 142 shifts attention to the role of Member States, encouraging them to promote research and innovation in AI systems capable of delivering positive environmental outcomes. Once again, sustainability is framed as a criterion for funding priorities, research agendas, and interdisciplinary collaboration, rather than as a mandatory condition for market access or system deployment. Environmental considerations thus become an objective to be pursued through national policies and incentives, not through uniform and enforceable EU-level obligations.[24]

---

[21] See Recital 4 of the AI Act, available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.
[22] See Recital 27 of the AI Act, available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.
[23] See Recitals 121 and 165 of the AI Act, available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.
[24] See Recital 142 of the AI Act, available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

This pattern culminates in Recital 174, which encapsulates the adaptive logic of the AI Act. The Commission is tasked with monitoring, over time, the development of energy-efficient technical standards, the energy consumption of general-purpose AI models, and the effectiveness of voluntary codes of conduct, with the possibility of proposing additional measures in the future if necessary. Crucially, however, this provision confirms that the current regulatory framework is provisional: environmental and energy impacts are acknowledged as potentially systemic risks, yet their concrete regulation is postponed, conditional, and contingent upon future assessments.[25]

The same logic permeates the binding Provisions and Annexes of the Regulation.

Article 40(2) links energy efficiency and resource use to future harmonised standards for AI high-risk systems, but the operative obligations depend entirely on standardisation processes that have yet to be completed.[26] Similarly, Article 59 recognises environmental protection and energy sustainability as public interest objectives justifying broader data usage within regulatory sandboxes, thereby facilitating experimentation rather than constraining environmental impact.[27] Then, Article 95 and Article 112(7) further entrench reliance on voluntary codes of conduct, whose effectiveness is to be evaluated only years after the Regulation's entry into force.[28]

Even the Annexes, while introducing novel transparency requirements, such as the obligation to document estimated energy consumption (Annex XI) and the relevance of computational intensity for assessing systemic risk (Annex XIII), stop short of establishing any binding thresholds, reduction targets, or legal consequences. Measurement and disclosure are required, but reduction and accountability remain optional.

In sum, the AI Act unmistakably recognises the environmental and energy implications of AI and embeds them within a broader strategic narrative aligned with the European Green Deal. However, this recognition does not translate into immediate, enforceable obligations yet. Environmental sustainability is largely deferred to future regulatory action, to Member State discretion, and to voluntary codes of conduct that, at present, remain peripheral and weakly operationalised.

The result is a framework in which sustainability is declared, monitored, and encouraged - but not yet legally governed.[29]

4.2 National Regulations (*Cristina Turconi*)

Recent legislative developments, with specific reference to Law No. 132 of September 23[30], 2025, outline a regulatory framework in which "sustainability" is elevated to the status of a cornerstone principle, yet it currently lacks an immediate operational implementation.

---

[25] See Recital 174 of the AI Act, available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

[26] See Article 40(2) of the AI Act, available at:https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

[27] See Article 59 of the AI Act, available at:https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

[28] See Article 95 and 112(7) of the AI Act, available at:https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

[29] For a further analysis see Nicolas Alder, Kai Ebert, Ralf Herbrich, and Philipp Hacker, AI, Climate, and Transparency: Operationalizing and Improving the AI Act, arXiv preprint arXiv:2409.07471v1 (2024), https://arxiv.org/abs/2409.07471; Thomas Le Goff (2025): Environmental law's principles applied to artificial intelligence: a path towards regulation, International Review of Law, Computers & Technology, DOI: 10.1080/13600869.2025.2506166.

[30] Law No. 132 of September 23, 2025 ("Provisions and delegations to the Government concerning artificial intelligence"): This constitutes the cornerstone regulation for the digital sector. Article 3 elevates sustainability to a general principle, while Article 24 grants a legislative delegation to the Government to ensure the alignment of the national legal system with Regulation (EU) 2024/1689 (AI Act) within twelve months.

Article 3 of the aforementioned Law includes sustainability among the "General Principles" that must guide the research, development, and application of AI systems; however, it does not define the technical parameters necessary for its measurement.

The statute is characterized as "framework legislation", systematically delegating the actualization of its objectives to subsequent legislative and administrative acts.
In particular:

a) Delegation of Powers to the Government: Article 24 confers a twelve-month delegation to the Government for the adoption of legislative decrees aimed at aligning the national legal system with Regulation (EU) 2024/1689 (AI Act).

b) Absence of Implementing Decrees: Multiple sectors, such as health research and data management, are currently contingent upon the issuance of future ministerial decrees (e.g., from the Ministry of Health) to define simplified processing methods and security safeguards.

Consequently, in the current state environmental protection within the technological-digital sector remains confined to a "programmatic imperative" pending the exercise of delegated powers required to establish the enforcement framework and specific execution measures.

This legal framework functions like a building permit for a skyscraper that lists the safety standards (the principles) but lacks the structural blueprints (the implementing decrees). Without the blueprints, the builders know the safety goals they must reach, but they don't yet have the technical instructions to actually start the construction.

Insofar as Data Centers are concerned, the current legal framework is primarily established by the "Guidelines for Environmental Assessment Procedures for Data Centers" (MASE, 2024).
Although this regulatory instrument specifies procedural obligations and thermal thresholds (50 MWt and 150 MWt) for the activation of Environmental Impact Assessment and Integrated Environmental Authorization (AIA/VIA), its nature as "Guidelines" confirms its function as interpretative guidance rather than a self-sufficient, detailed technical regulation.

Indeed, an analysis of the sources reveals the following:

I. Discrepancy between Principle and Practice: The Guidelines "translate" regulatory principles into requirements, yet the actual operational implementation of circular economy criteria—such as waste heat recovery—is still frequently perceived as a strategic option or a design requirement to be negotiated with local authorities.
It has not yet been crystallized as a mandatory obligation within technical implementing decrees providing detailed, enforceable technical specifications.

II. Subsidiary Role of International Standards: In the absence of exhaustive national technical legislation that operationally defines "energy sustainability," operators are compelled to rely on "private international frameworks" (such as ANSI/TIA-942 or Uptime Institute) to fill the existing regulatory vacuum and provide verified efficiency metrics.
The adoption of such standards is not a mere technical preference but a strategic necessity to "de-risk" authorization processes,. This occurs because national norms define the overarching objectives without providing the specific operational instruments required to achieve them.

The current regulatory framework for data centers is like a GPS that shows the final destination (the sustainability goals) but lacks the detailed street maps (the implementing decrees) to get there. To avoid getting lost or stalled, operators are forced to buy private maps from third parties (international standards) to navigate the actual terrain of the authorization process.

Further Primary Legislation and Emergency Decrees:
- Decree-Law No. 153 of October 17, 2024 (commonly known as the "Environment DL"): An emergency legislative measure that introduced substantial amendments to the Consolidated Environmental Act. Its primary objectives are to promote the circular economy and streamline environmental assessment procedures, resulting in a direct impact on the mandatory heat recovery obligations for data centers.
- Legislative Decree No. 152 of April 3, 2006 ("Environmental Regulations" or the "Consolidated Environmental Act" – TUA): This is the general regulatory framework governing Environmental Impact Assessment (EIA/VIA) and Integrated Environmental Authorization (IEA/AIA). It has been recently amended by Decree-Law 153/2024 to integrate new sustainability requirements.
- Law No. 132 of June 28, 2016: The founding statute of the National Network System for Environmental Protection (SNPA). It governs the coordination between the Higher Institute for Environmental Protection and Research (ISPRA) and the Regional Environmental Protection Agencies (ARPA), which are the technical bodies responsible for the monitoring and enforcement of environmental requirements and prescriptions imposed on data centers.

4.3 Standards *(Raffaella D'Alessandro)*

The environmental impact of Artificial Intelligence (AI) is a growing concern, and attention is focused on creating **international standards and frameworks** to measure, manage, and mitigate this impact. Standardization is important in ensuring that AI is developed and deployed in an environmentally sustainable manner. Standards provide guidelines and benchmarks for energy efficiency, hardware optimization, and sustainable data management practices. By adhering to standardized guidelines, AI developers and operators can ensure that their systems meet specific environmental performance criteria, such as energy consumption and GHG[31] emissions.
Here are the main relevant international standards with a detailed focus on how each addresses the environmental impact of AI.

ISO/IEC Standards
Standards from the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) are crucial for AI responsible management, which includes environmental considerations.

ISO/IEC 42001: Artificial Intelligence Management System (AIMS)
This is the first international standard for an AI Management System. It establish controls for the responsible use of AI. It emphasizes the need for operational controls over the AI lifecycle (development, deployment, operation). These controls can be specifically applied to measure and manage the energy and resource consumption of the AI infrastructure (data centers, cloud resources, hardware lifespan).

ISO/IEC 42005: AI System Impact Assessment

---

[31] GHG (Greenhouse Gas) emissions are gases like carbon dioxide (\(CO_{2}\)), methane (\(CH_{4}\)), and nitrous oxide (\(N_{2}O\)) released from human activities, primarily burning fossil fuels, that trap heat in Earth's atmosphere, intensifying the greenhouse effect and causing global warming and climate change.

This standard provides practical guidance on how to conduct an AI System Impact Assessment (AIA). Environmental impact is an explicit domain of this assessment. Organizations must identify environmental risks such as high energy consumption (training and inference), water usage (for data center cooling), and e-waste generation (short lifespan of specialized AI hardware like GPUs).

ISO 14040/14044: Life Cycle Assessment (LCA)

LCA is the globally recognized methodology for assessing environmental impacts associated with all stages of a product or service's life. This is the most direct standard for environmental accounting. It provides the framework for calculating the AI model’s total environmental footprint, covering: 1. Embodied Carbon: The carbon emitted during the manufacturing of hardware (servers, chips, networking gear). 2. Operational Energy: The energy and resultant Carbon Dioxide (CO2) emissions from training and running the model (inference). 3. Water Consumption: The water used for cooling data centers (a significant factor).

CEN/CENELEC standards

The JTC21 of CEN/CENELEC shall produce standardization deliverables in the field of Artificial Intelligence (AI) and related use of data, as well as provide guidance to other technical committees concerned with Artificial Intelligence. The JTC shall produce standardization deliverables to address European market and societal needs and to underpin primarily EU legislation, policies, principles, and values.

At the moment there is one AI Environmental Impact standard published:

CEN/CLC/TR 18145:2025 Environmentally sustainable Artificial Intelligence

This document establish a framework for quantification of environmental impact of AI and its long-term sustainability and encourage AI developers and users to improve efficiency of AI use. It also provides a summary of the state of the art of AI technology for direct control and optimisation of energy use in energy systems.

Work in progress for the development of new AI environmental impact standards are:

prEN 18287 Sustainable Artificial Intelligence – Guidelines and metrics for the environmental impact of artificial intelligence systems and services

This document describes the principles and framework for environmental impact measurement of artificial intelligence systems and services and provides guidelines for impact reduction throughout its lifecycle. It includes: - A framework for defining the environmental impact of artificial intelligence - A harmonized calculation method for assessing the environmental impact of artificial intelligence systems and services - Reporting guidelines - Best practices for reducing the environmental impact of AI systems and services throughout their lifecycle.

prEN 18228 AI Risk Management

This document specifies requirements on risk management for AI systems and provides clear and actionable guidance on how risk can be addressed and mitigated throughout the entire lifecycle of the AI system. Risks covered include both risks to health and safety and risks to fundamental rights which can arise from AI systems, with impact for individuals, organisations and society.

In summary, the trend in international standardization is to shift from general ethical principles to **mandated, measurable requirements** concerning the environmental impact of AI. Despite the benefits, there are gaps in the current standards that need to be addressed. For example, there is a lack of comprehensive metrics for measuring the energy efficiency of AI systems, and guidelines for lifecycle assessment are not yet fully developed. Addressing these gaps is essential for realizing the full potential of standardization in promoting sustainable AI.

4.4 Concluding evaluations *(Raffaella D'Alessandro, Adriana Peduto, Cristina Turconi, Beatrice Rosa)*

Ultimately, the European and national regulatory framework recognises the environmental and energy implications of AI and places them within the strategic trajectory of the European Green Deal. Yet this political recognition has not translated into binding obligations: sustainability is largely deferred to future regulation, Member State discretion, and voluntary codes that remain marginal and weakly operationalised. International standardisation is moving from high-level ethical principles toward measurable requirements on environmental impact, but important gaps persist — notably the absence of consolidated energy-efficiency metrics and mature lifecycle assessment (LCA) guidelines — which limit the ability of standards to drive a genuine sustainability transition at the moment.

The cautious regulatory stance appears deliberate: policymakers seek to protect fundamental rights and environmental goals without erecting rigid boundaries that could stifle experimentation and European competitiveness. In light of this, s**oft-law instruments** — voluntary codes, guidance, and emerging standards — would therefore be employed as a pragmatic balancing tool, enabling innovation to proceed while shaping expectations and practices. Nevertheless, to avoid soft law becoming mere rhetoric, however, clear **criteria** must be set for when and how to escalate to binding measures.

Evidence shows that AI can materially support climate mitigation and adaptation but also creates energy, governance, and equity risks that demand specific rules and metrics. **Policy priorities** should be sequenced and assigned: **(1) mandatory energy-use and LCA reporting** for AI systems in critical sectors**; (2) EU-wide metrics** for kWh per inference and tCO_2e per model; (3) **environmental criteria in public procurement;** (4) funding for low-power AI research and pilot projects; (5) coordinated guidance aligning the AI Act with climate policy. Only by aligning strategy, standards, and regulation — and by using soft law as a time-bound, conditional stepping stone toward hard law where thresholds are met — can AI sustainability evolve from a declared principle into a genuinely governed dimension that supports Europe's climate objectives while preserving space for innovation.

References

- Environmental law's principles applied to artificial intelligence: a path towards regulation
- The EU AI Act: Insights from the Green AI Committee | GSF
- The role of Artificial Intelligence in the European Green Deal
- EU Commission aims to regulate environmental impact of AI through delegated act | Euronews
- https://green-forum.ec.europa.eu/green-business/environmental-footprint-methods_en
- https://commission.europa.eu/energy-climate-change-environment_en
- Green and intelligent: the role of AI in the climate transition | npj Climate Action
- https://va.mite.gov.it/it-IT/DatiEStrumenti/MetadatoRisorsaCondivisione/768769f3-3f9c-4080-b81c-1d8801fbd401
- https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legge:2024-10-17;153
- Ecological footprints, carbon emissions, and energy transitions: the impact of artificial intelligence (AI) | Humanities and Social Sciences Communications
- AI and the Environment - International Standards for AI and the Environment SG5 - EMF, environment, climate action, sustainable digitalization, and circular economy
- https://standards.cencenelec.eu/ords/f?p=205:22:::::FSP_ORG_ID,FSP_LANG_ID:2916257,25&cs=114251C6C0B684FBBC069923513BF6348

5 AI AND CLIMATE CHANGE: UNDERSTANDING THE ENERGY-WATER AI NEXUS *(Glauco Bertocchi, Tommaso Diddi, Beatrice Versaci)*

The main objective of this chapter is to analyze the energy-water nexus in AI "production." The authors are aware that there are other aspects to consider, such as the impact of the energy used in AI on the production and distribution of electricity and how AI can contribute to optimizing these processes. The latter aspect is specifically addressed in Chapter 7, "AI and Climate Change: AI for Energy System Optimization." This chapter will refer to these and other parts of the report that specifically address complementary aspects.

Artificial intelligence (AI) is based, in a very concise and simplified way, on three pillars:

1. Algorithms that enable artificial intelligence systems to correctly perform all the required operations.
2. Technology declined into: a) basic hardware technology, chips, which enable algorithms to be executed efficiently and effectively; b) the possibility of creating data centers that aggregate hundreds of thousands (and more) servers to enable the processing and provision of AI services to all users who request them.
3. Availability of large amounts of data (Big Data) necessary for training AI models; availability means the possibility of processing and storing both the original data and the results of the processing phases.

This is because current AI is a composite universe in which there are systems that require progressively increasing processing capacity (computing speed, memory size, and transfer speed between components).

Data centers were initially created to meet the demands of cloud computing and are still mainly used for this purpose. Over the past few years, the possibility for anyone with an Internet connection, to use generative AI systems has seen an explosion in demand for processing capacity and, therefore, the need for "specialized" data centers with technology optimized for parallel computing (e.g. GPUs) and high-bandwidth memory. Data centers for AI production will be discussed in paragraph 5.3.

However, the entire AI system requires two essential components: the energy to run the servers and an efficient system for dissipating the heat produced by the processing systems. Given the current state of technology, all equipment used in AI utilizes energy in the form of electricity, and therefore, throughout the remainder of this chapter, the terms "energy" and "electricity" will be used interchangeably.

In the following of this chapter, we use the term" water" to indicate cooling even though heat dissipation can be achieved with different means (e.g. air, gases, fluids, etc.). We are aware that the environmental "footprint" of these means can be quite different, and we will indicate it when relevant. It is therefore clear that the use of AI has a cost in terms of environmental impact and is therefore not a priori neutral with respect to climate change.

Consequently, AI depends on both electricity (energy) and cooling (water), while energy and water are already interconnected through power generation and treatment.

- AI increases energy demand, which can mean more water is needed for cooling and electricity production.

- AI increases water demand (directly for cooling, indirectly via energy production).
- This makes AI a new stress factor in the traditional energy–water nexus.

5.1 Energy and water demand from AI

A simplified scheme, based on current (2025-26) usage of AI, is the following

> a) Domestic or personal use, via connections to cloud systems, with a limited amount of processing that can be carried out locally using "home" computers with specialized processors (GPUs, NPs, etc.). This use does not substantially change local energy demand, as the devices have similar energy requirements to previous generations of devices. Nowadays, almost all "domestic" use of AI involves cloud-based systems accessible via the Internet.
>
> b) Business use, which can take place either through connections to cloud systems or by on-premises devices, and, of course, in a mixed manner. This use can change local energy demand in all cases where it is decided to use specialized on-premises equipment, since AI system processing is, as mentioned, energy-intensive.

In both cases, it should be noted that the use of AI requires the provision of services through a cloud structure capable of meeting user demands.

From the above considerations, it can be deduced that today, the use of AI, and particularly access to LLMMs and similar technologies for so-called generative AI, cannot be separated from the existence of large data centers capable of performing the necessary processing.

It is therefore important to identify, albeit in a simplified way, how the energy of data centers is used and what the resulting cooling requirements are.

Among the various types of AI, those that require greater computing power due to the algorithms used are those that use LLMMs with Transformer or similar capabilities. These are the ones most required and offered to both the private and business sectors. The main characteristics of these types of AI, with a particular focus on their impact on data centers, will be illustrated in following paragraphs.

To identify the areas of greatest energy and cooling demand, it may be useful to break them down into basic factors, even if simplified. This operation will use an approximation sufficient to identify the main factors and will allow details to be "masked," such as the comparative efficiency between different algorithms or the use of a specific type of hardware.

To break down energy and cooling demands into factors, we can proceed as follows: attribute energy demand primarily to intangible processing activities, i.e., software execution, while cooling demand is a consequence of the hardware technology used and the time the systems are turned on.

The energy demand of AI can be schematically divided into the two main phases of current AI technologies, increased by the energy required to execute the functions necessary for the data center to operate.

- **a)** the training of AI models;
- **b)** the usage or 'inference' phase, when users use AI applications;
- **c)** the execution of complementary software necessary for the operation of the data center (operating systems, networking, etc.).

The first two phases are those that are most dependent on the algorithms used. Energy demand obviously also depends on the technology and efficiency of the hardware that executes the algorithms. The latter factor concerns energy efficiency and technology and will be briefly considered below with efficiency (algorithms, hardware, etc.) in AI operations. Additional information on these subjects can be found in the already cited chapter 7 and in the first part of chapter 3.

Cooling demand can be divided into two components:

1. The first relates to the direct needs of the data center, which depend on:
   - **a)** the hardware technology used and the aggregation characteristics (density) of the circuits
   - **b)** the power-up time of the hardware systems
   - **c)** the type of technology used in cooling circuits.

2. The second relates to indirect use resulting from the production of the total energy consumed by the data center.

The energy demands arising from AI "production," especially generative AI, have been described in several reports that have helped to shed light on certain aspects of the problem.

As a broad point of view in AI production can be useful the approach followed by ITU in a [32]study. The report examines how international standards can guide the ICT industry toward environmentally sustainable AI development. While AI offers transformative potential for addressing global challenges like climate change, it poses significant environmental concerns including high energy consumption in data centers, greenhouse gas emissions from computational demands, and increasing e-waste.
The document emphasizes understanding AI's lifecycle—from problem identification through data collection, model design, training, evaluation, deployment, and inference—to identify stages where energy use and emissions are highest. Model training and inference are particularly resource-intensive phases requiring targeted mitigation efforts.
International standards play a critical role in ensuring AI's environmental benefits outweigh its footprint. The ITU-T SG5 has developed foundational standards for environmental efficiency across AI systems, data centers, and e-waste management. These standards provide frameworks to measure, manage, and reduce AI's environmental impacts at product and network levels.

As a specific point of view of the AI "production" a Google report.[33].evaluate the energy and water consumption used to answer a prompt for Gemini (Google's AI System).
The authors propose and implement a comprehensive methodology to measure energy usage, carbon emissions, and water consumption of AI serving workloads at scale.
The study examines Google's AI infrastructure serving the Gemini AI assistant, accounting for the complete AI serving stack including AI accelerator power, host system energy, idle capacity, and data center overhead. Key findings reveal that the median Gemini Apps text prompt consumes just 0.24Wh of energy—significantly lower than many public estimates.
The research demonstrates that Google's software efficiency improvements and clean energy procurement have achieved a 33x reduction in energy consumption and a 44x reduction in carbon footprint for the median text prompt over one year. To provide context, the median prompt uses less energy than watching nine seconds of television and consumes water equivalent to five drops (0.26 mL).
It is important to note that Google's study focuses entirely on consumption during the inference phase and does not consider other phases of the AI lifecycle that are very energy-intensive, such as training.

---

[32] https://www.itu.int/dms_pub/itu-t/opb/env/T-ENV-ENV-2024-1-PDF-E.pdf

[33] https://arxiv.org/pdf/2508.15734

Despite these limitations, which are well known to the authors of the report, the work is significant because it provides a "real" assessment, even if difficult to verify.

Another paper[34] examines the environmental impacts of GenAI training on a specific GPU (Nvidia A100) This study presents the first comprehensive multi-criteria life cycle assessment (LCA) of AI training, addressing a significant gap in environmental impact research. While current assessments focus primarily on operational carbon emissions using secondary data, this research examines 16 environmental impact categories using detailed primary data from the Nvidia A100 SXM 40GB GPU. This multi-criteria analysis broadens the Sustainable AI discourse beyond operational carbon emissions, challenging existing sustainability narratives and emphasizing the need for comprehensive policy frameworks that address AI's full spectrum of environmental impacts.

The second pillar of AI is represented by algorithms. Among the most widely used in generative AI systems are Transformers[35] (the reference is the foundation paper of this technique), which are based on parallel computing but require a huge number of operations. To speed up Transformer operations, research is exploring various alternatives that reduce the number of operations, especially for long sequences (hence the difficulty of "examining" texts consisting of hundreds of pages). Among the most promising methods is the use of models (SSM) that employ more efficient algorithms to select the data to be "examined."[36]

In summary the characteristic of AI production, given the current state of algorithms and theoretical studies, is the use of hardware in "very high density" configurations and with computational speeds that require ever-larger data centers.
Without these facilities and those currently being planned (see paragraph 5.3), it does not seem possible, at the current state of the art, to provide AI services on demand and on a large scale with acceptable response times and costs.

5.2 Energy and water supply for AI

The previous paragraph highlighted how the infrastructure for AI operations using current techniques (software and hardware) is particularly significant in terms of processing capacity and therefore requires significant amounts of energy and cooling. This paragraph will examine possible energy sources and their classification in terms of renewability and climate impact. The same breakdown will also be made for cooling, considering mainly water consumption.
The required level of energy is such that it can only be met by connecting to an electrical grid with adequate capacity or by on-site production using dedicated power plants.

Energy from a grid can be generated by

- Renewable sources (solar, wind, hydro).
- Non-renewable sources (coal, natural gas, petrol, nuclear).

However, the division into renewable and non-renewable sources must be supplemented by another parameter, namely climate impact. In very broad terms, and limiting ourselves to the operation of these sources, we can say that:

- Low-impact sources (solar, wind, hydro, nuclear)
- High-impact sources (coal, natural gas, oil)

---

[34] https://arxiv.org/abs/2509.00093
[35] https://arxiv.org/abs/1706.03762
[36] https://arxiv.org/abs/2312.00752

This division does not consider the climate and environmental impact of the construction of these types of energy sources. For example, a hydroelectric plant can have a considerable environmental impact during construction, resulting in changes to the local microclimate.
A more detailed consideration of energy use and distribution can be found in chapter 7 of this report. Sections 7.1 and 7.2 highlight the current and future impact of data centers on the electricity grid.

As regards cooling, it is considered that this is mainly produced using water and therefore the origin of this water will be considered. In the case of water, however, it should be noted that it is a source whose availability is limited, even though it is a renewable source. Furthermore, its availability over time is subject to various factors, including climate change, which could significantly alter the amount of water, either increasing or decreasing it. The reduction of the quantity of water available in a "region" can produce, in the medium-long term, a desertification process that will feed more climate change.

Water used by AI may come from:

- Surface water (rivers, lakes, reservoirs).
- Groundwater (aquifers, wells).
- Municipal supply (city water systems).

This water is used in:

- Cooling systems in data centers (direct).
- Power plants that generate electricity for AI (indirect).

It is also important to note that, especially for high-performance hardware circuits and servers, cooling is adopting technologies that use gases other than air and liquids other than water. This change reduces water consumption but still involves the problems of disposing of or reusing the heat produced.

It is important to consider that the construction of data centers with dedicated power plants on the same site increases water consumption and could therefore have a greater impact. A nuclear plant of the 4th generation can reduce the impact of water consumption, but the usage of the heat produced remains.[37]
On the other hand, connection to an electricity grid "relocates" and increases water consumption in places other than data centers but could be less efficient due to energy transmission losses.

In areas with water stress, AI's cooling needs can compete with agriculture, industry, and households. This competition could create social tensions that could also compromise even the possibility of building or operating the data center.

5.3 The Impact of Data Centers

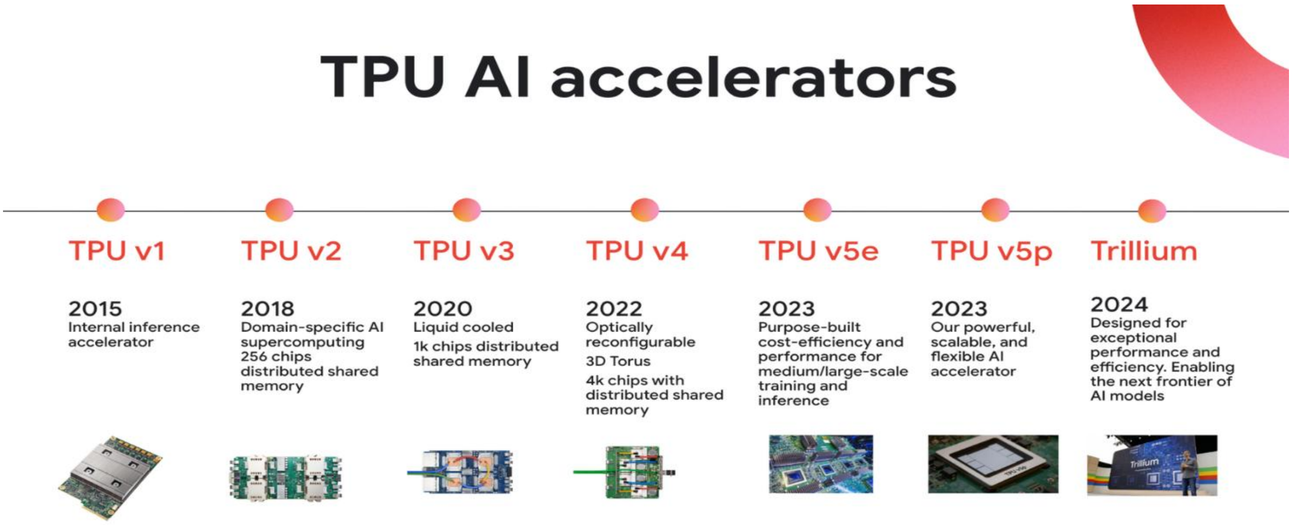
Data centers constitute the foundational infrastructure underpinning contemporary digital services and, increasingly, Artificial Intelligence applications.

A data center is a purpose-built facility designed to provide secure, reliable, and continuous support for digital services, including cloud computing, large-scale artificial intelligence (AI) operations, and other critical information systems. A modern data center constitutes an integrated socio-technical

---

[37] https://www.world-nuclear-news.org/Articles/Pilot-Natrium-plant-to-be-built-in-Wyoming

system in which computing, storage, networking, energy, cooling, and security infrastructures operate in concert to maintain high levels of performance, availability, and resilience. At its core, the information technology infrastructure comprises servers, storage arrays, and networking devices. Servers execute computations, ranging from conventional data processing to highly demanding AI workloads that rely on specialized hardware such as graphics processing units (GPUs) and tensor processing units (TPUs).
The history of ten years of AI TPU chips made by Google is synthesized in figure 5.1



TPU AI accelerators

TPU v1
2015
Internal inference accelerator

TPU v2
2018
Domain-specific AI supercomputing
256 chips distributed shared memory

TPU v3
2020
Liquid cooled
1k chips distributed shared memory

TPU v4
2022
Optically reconfigurable
3D Torus
4k chips with distributed shared memory

TPU v5e
2023
Purpose-built cost-efficiency and performance for medium/large-scale training and inference

TPU v5p
2023
Our powerful, scalable, and flexible AI accelerator

Trillium
2024
Designed for exceptional performance and efficiency. Enabling the next frontier of AI models

Fig. 5.1 Google's AI-specialized-chips-TPU-history-
(source: https://cloud.google.com/transform/ai-specialized-chips-tpu-history-gen-ai)

Storage systems, whether implemented as direct-attached, network-attached, or distributed architectures, provide reliable, high-speed access to data while ensuring redundancy and fault tolerance. Networking components, including routers, switches, firewalls, and load balancers, orchestrate data flows both within the facility and across broader networks, maintaining low latency and high throughput essential for real-time applications.

The effective operation of these IT systems depends critically on the underlying energy and cooling infrastructures. Data centers consume substantial amounts of electricity, which is supplied both by the public grid and by backup systems such as uninterruptible power supplies (UPS) and diesel or gas generators that ensure continuity during outages. Redundant power distribution schemes, often aligned with Tier III or Tier IV[38] standards, are implemented to mitigate the risk of service interruption. Thermal management is equally critical; the high-density computing environment generates considerable heat, which must be dissipated to maintain optimal operating conditions. Cooling solutions range from traditional air-conditioning units to advanced liquid immersion systems, and increasingly leverage free-air cooling when climate conditions permit. Water usage, typically associated with cooling systems, can be substantial, and operators are exploring approaches such as reclaimed water circuits and direct-to-chip liquid cooling to reduce the environmental footprint.

A data center's functionality is further enhanced by sophisticated management and security systems. Data Center Infrastructure Management (DCIM) platforms monitor and optimize energy consumption, thermal conditions, and IT workload performance. Physical security measures,

---

[38] Data center Tier classification is an industry standard system (by the *Uptime Institute*) that categorizes data centers into four levels (Tier I–IV) based on their infrastructure redundancy, fault tolerance and expected uptime, with higher tiers indicating more robust power/cooling paths and greater availability guarantees. https://journal.uptimeinstitute.com/explaining-uptime-institutes-tier-classification-system/

including restricted access, biometric authentication, surveillance systems, and fire suppression mechanisms, protect the facility from unauthorized access and environmental hazards. Cybersecurity frameworks, encompassing firewalls, encryption, and network segmentation, ensure the integrity and confidentiality of the data processed within the facility.

The operational cycle of a data center begins when a user or application submits a request, which is routed through networking devices to the appropriate server cluster. The servers execute the computation or AI model, access necessary storage resources, and generate results that are subsequently transmitted back to the user or distributed across networks. Throughout this cycle, monitoring systems continuously assess energy use, thermal conditions, and performance metrics, ensuring reliability and enabling dynamic optimization.

Modern data centers vary in scale and function. Enterprise data centers serve individual organizations and are often managed internally, whereas colocation facilities host the equipment of multiple clients, providing shared infrastructure for efficiency. Hyperscale data centers, operated by leading cloud and AI providers, represent the apex of scale and complexity, often comprising tens of thousands of servers and consuming power in the multi-gigawatt range. These facilities exemplify the convergence of computational demand, energy infrastructure, and environmental considerations, particularly as AI workloads become increasingly dominant.

In recent years, the scale and density of data centers have expanded considerably, driven by the exponential growth of cloud computing, the proliferation of connected devices, and the computational intensity of AI workloads[39]

Figure 5.2 gives an idea of the size of today's data centers.

Fig. 5.2 Daytime view of the new Albany data center campus in central Ohio
(source: https://datacenters.google/discover-more/photo-gallery)

---

[39] International Energy Agency (IEA). (2024). Electricity 2024: Analysis and Forecast to 2026. Paris: OECD/IEA. https://iea.blob.core.windows.net/assets/18f3ed24-4b26-4c83-a3d2-8a1be51c8cc8/Electricity2024-Analysisandforecastto2026.pdf ; https://www.science.org/doi/10.1126/science.aba3758 ;

Energy demand and AI-related consumption

Global electricity consumption by data centers, cryptocurrencies, and artificial intelligence was estimated at approximately 460 TWh in 2023, corresponding to around 1.7–2% of total global electricity demand[40] . Although the sector has achieved notable efficiency gains through improved hardware, virtualization, and advanced cooling systems, the recent diffusion of AI models, particularly large language models (LLMs), has altered this trajectory. AI workloads are characterized by high computational density and the necessity for specialized hardware, such as GPUs and TPUs, which exhibit significantly higher energy intensity compared to conventional servers.

The share of total data center electricity attributable to AI is currently estimated between 10% and 15%, but this figure is expected to rise substantially as generative AI becomes pervasive in both consumer and industrial applications. Scenario analyses suggest that, under high-adoption pathways, AI-related demand could account for 30–40% of total data center energy use by 2030[41].

Training large-scale AI models represents the most energy-intensive phase of the AI lifecycle. Recent empirical assessments indicate that a single training cycle of a frontier LLM may consume hundreds of megawatt-hours (MWh), a quantity equivalent to the annual electricity consumption of several hundred European households[42]. By contrast, inference activities, though less energy-demanding per transaction, occur continuously at a massive scale, implying that aggregate energy demand for inference may soon surpass that for training.

Water consumption and thermal management

The operational sustainability of data centers cannot be understood without considering their water footprint, primarily linked to cooling requirements. Thermal management systems, which maintain the operational stability of processors and other components, typically rely on either evaporative cooling or chilled-water systems, both of which entail significant water withdrawals. Depending on the technology and climatic context, data centers consume between 0.1 and 5 liters of water per kWh of electricity used. A medium-sized facility may use up to approximately 416 million liters of water per year, while large-scale data centers can consume more than 6.8 billion liters annually, volumes comparable to the water demand of entire urban communities[43].

AI-driven workloads exacerbate these demands by increasing thermal density and, consequently, the cooling load. The industry's response has included the adoption of liquid immersion cooling, direct-to-chip liquid cooling, and the use of non-potable or reclaimed water sources, as well as the strategic siting of facilities in regions with favorable climatic or hydrological conditions.

Spatial distribution and strategic concentration

The geography of hyperscale data centers reflects the interaction of technological, environmental, and geopolitical factors. While the number of data center facilities is widely distributed across countries, hyperscale capacity is far more concentrated, as it depends on the size and power density of individual sites. As of 2025, approximately 40% of global hyperscale capacity is located in the United States, despite the country hosting a broader mix of facility types; this concentration is particularly evident in Northern Virginia, Oregon, and Texas, where robust grid connectivity and favorable regulatory frameworks prevail. Europe accounts for roughly 25% of global hyperscale

---

[40] IEA, 2024

[41] International Energy Agency (IEA). (2025) In focus: Data centers – an energy-hungry challenge https://energy.ec.europa.eu/news/focus-data-centers-energy-hungry-challenge-2025-11-17_en#:~:text=According%20to%20the%20International%20Energy,towards%20115%20TWh%20by%202030 .

[42] Patterson, D., Gonzalez, J., Le, Q. et al. (2021). The Carbon Footprint of Machine Learning Training. Google Research White Paper. **https://arxiv.org/pdf/2204.05149.**

[43] https://www.eesi.org/articles/view/data-centers-and-water-consumption

capacity, even though it comprises a large number of smaller and colocation facilities; hyperscale clusters are primarily found in Ireland, the Netherlands, Sweden, and Finland, regions benefiting from abundant renewable energy and cooler climates that reduce energy and water demands. The Asia–Pacific region, which hosts a substantial share of the world's data centers by count, represents approximately 20% of global hyperscale capacity, with major hubs in Singapore, Japan, and India, where expansion is increasingly constrained by high energy costs and water availability[44].

The location of AI-oriented data centers is particularly strategic, as operators seek to balance latency requirements, energy availability, and climate resilience. The recent trend toward co-locating data centers near renewable generation sites or within regions offering free-air cooling potential exemplifies the sector's adaptive strategies to energy-water constraints.

A particularly notable development illustrating the scale of AI-driven infrastructure investment is the OpenAI–Oracle–SoftBank "Stargate"[45] initiative, announced in 2025. OpenAI plans to invest approximately US$400 billion to develop five new U.S. data center sites across Texas, New Mexico, and Ohio, adding a combined 7 GW of capacity, comparable to the power demand of several medium-sized cities. The project, part of a broader commitment of US$500 billion in AI infrastructure, represents one of the largest coordinated data center expansions ever undertaken. Its primary objective is to support large-scale AI computation for OpenAI services, including ChatGPT, which now serves an estimated 700 million weekly users. The new facilities, such as those in Shackelford County (Texas) and Doña Ana County (New Mexico), exemplify the emerging pattern of AI infrastructure clustering around renewable-rich, low-latency regions of the continental United States. The initiative underlines how AI infrastructure is becoming a macro-scale energy actor. Concentrating multi-gigawatt data centers in select U.S. regions will have measurable implications for regional electricity grids, water basins, and climate adaptation policies, demanding integrated planning approaches that bridge digital and environmental governance.

Environmental and policy implications

The rapid growth of data center infrastructure raises critical questions for energy and environmental policy. Without continued efficiency improvements and a systemic transition to renewable energy, the expansion of AI-driven computing risks offsetting global decarbonization gains. Accordingly, regulatory frameworks in the European Union and other jurisdictions increasingly mandate energy performance disclosures, Power Usage Effectiveness (PUE) and Water Usage Effectiveness (WUE) reporting, and the integration of waste heat recovery and renewable sourcing strategies into data center design and operation[46].

The convergence of AI, energy, and water infrastructures thus presents a paradigmatic case of the energy–water–digital nexus, where efficiency and sustainability cannot be treated in isolation. Ensuring the long-term viability of AI systems requires coordinated governance approaches that integrate technological innovation, resource management, and policy coherence.

---

[44] Synergy Research Group. (2025) The World's Total Data Center Capacity is Shifting Rapidly to Hyperscale Operators https://www.srgresearch.com/articles/the-worlds-total-data-center-capacity-is-shifting-rapidly-to-hyperscale-operators ; Visual Capitalist. (2025) Visualizing All of the World's Data Centers in 2025. https://www.visualcapitalist.com/visualizing-all-of-the-worlds-data-centers-in-2025/

[45] South China Morning Post. (2025). OpenAI expands Stargate with 5 new data center sites across US. https://www.scmp.com/tech/big-tech/article/3326619/openai-expands-stargate-five-new-data-center-sites-across-us?module=top_story&pgtype=section

[46] European Commission. (2024). EU Code of Conduct for Data Centers (Energy Efficiency) – 2024 Update. Brussels. https://joint-research-center.ec.europa.eu/jrc-news-and-updates/eu-code-conduct-data-centers-towards-more-innovative-sustainable-and-secure-data-center-facilities-2023-09-05_en

5.4 Additional References

- European Commission. (2024). 2024 *Best Practice Guidelines for the EU Code of Conduct on Data Center Energy Efficiency* https://e3p.jrc.ec.europa.eu/en/publications/2024-best-practice-guidelines-eu-code-conduct-data-center-energy-efficiency
- Masanet, E., Shehabi, A., Lei, N., Smith, S., & Koomey, J. (2020). *Recalibrating global data center energy-use estimates*. Science, 367(6481), 984–986. https://www.science.org/doi/10.1126/science.aba3758
- Lawrence Berkeley National Laboratory — *2024 United States Data Center Energy Usage Report*: https://climate.mit.edu/sites/default/files/2025-02/lbnl-2024-united-states-data-center-energy-usage-report.pdf

6 AI AND CLIMATE CHANGE: AI FOR PREDICTING CLIMATE CHANGE *(Alberto Stefanini, Elenio Dursi)*

6.1 Introduction – AI Techniques in Climate Modelling

Artificial intelligence (AI) is increasingly used to analyse environmental data and support predictive models relevant to the understanding of climate dynamics. It is important to distinguish clearly between *meteorology* and *climatology*: meteorology focuses on short-term atmospheric phenomena (weather), whereas climatology studies long-term trends, aggregated behaviours, and structural shifts in the climate system. This distinction guides the appropriate use of AI techniques in climate-related analyses and ensures coherence with the objectives of the present Report, centerd on climate change rather than short-term forecasting.

Recent developments in AI have enabled the integration of vast and heterogeneous datasets—such as satellite observations, oceanographic measurements and land-surface records—into coherent analytical frameworks. While AI-enhanced weather forecasting tools remain valuable in other domains, the focus here is on AI methods that support *climate modelling*: long-term projections, multi-decadal variability, and the identification of emerging climate risks.

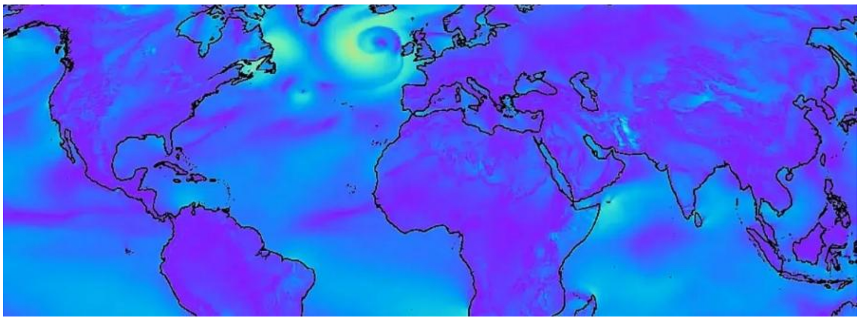
Climate modelling relies on General Circulation Models (GCMs) and Earth System Models (ESMs), which require substantial computational resources. AI techniques are increasingly used to accelerate specific components of these models, such as parameterisation of sub-grid processes, anomaly detection in multi-decadal datasets, and the downscaling of coarse-resolution simulations. These approaches complement rather than replace physical models, allowing more efficient experimentation and more detailed assessments of long-term climate trajectories.

Hybrid approaches combining physics-based models with machine learning have shown promising results in improving projections of temperature trends, ocean heat content, cryospheric dynamics, and the frequency of extreme events under different emission scenarios. AI also plays a growing role in analysing paleoclimate records, assessing regional climate impacts, and supporting integrated assessment models (IAMs) that link socio-economic pathways with climate outcomes.

Overall, AI enhances climate science by strengthening long-term predictive frameworks, improving the robustness of global and regional climate projections, and contributing to a deeper understanding of climate-change-related risks.

6.1.1 Machine Learning for Forecasting

Recent breakthroughs have demonstrated that neural networks can achieve or even surpass the skill of traditional numerical models in medium-range forecasting. A key example is **DeepMind's GraphCast**, developed with the European Center for Medium-Range Weather Forecasts (ECMWF).



It employs a **graph neural-network architecture** to represent the Earth as a mesh of interlinked nodes, learning spatiotemporal dependencies from forty years of ERA5 reanalysis data. The model produces global weather forecasts up to ten days ahead with accuracy comparable to the ECMWF's operational IFS system while running thousands of times faster (DeepMind, 2023).

Similar innovations include **NVIDIA's FourCastNet**, which uses adaptive Fourier neural operators to model global atmospheric dynamics (Pathak et al., 2022), and **Huawei's Pangu-Weather**, which achieves high accuracy with a 3D transformer architecture. These systems demonstrate that AI can learn deterministic and probabilistic relationships governing climate variables without explicit equations, offering both speed and energy savings.

However, the actual improvement of AI in operational meteorological forecasting remains limited. Efficient forecasting tools for renewable generation already exist and are routinely used by TSOs; the contribution of AI is currently marginal due to data-quality constraints and the chaotic nature of atmospheric dynamics.

6.1.2 Data Assimilation and Surrogate Modelling

Classical models depend heavily on **data assimilation**, the process of merging observations with model outputs to obtain an optimal estimate of the current state of the atmosphere. AI can enhance this process through **recurrent and ensemble-learning techniques** that identify correlations among heterogeneous data sources, filling gaps and detecting sensor anomalies. Machine-learning surrogates can emulate specific components of GCMs—such as radiative transfer or cloud microphysics—at a fraction of the computational cost, enabling larger ensembles and longer time horizons.

In this sense, AI serves as an *accelerator* of climate simulation, allowing researchers to explore multiple emission pathways or sensitivity analyses that would otherwise be prohibitively expensive. Hybrid models, sometimes termed **physics-informed neural networks (PINNs)**, embed conservation laws directly within the learning process, ensuring physical plausibility while maintaining flexibility.

6.1.3 Downscaling and Bias Correction

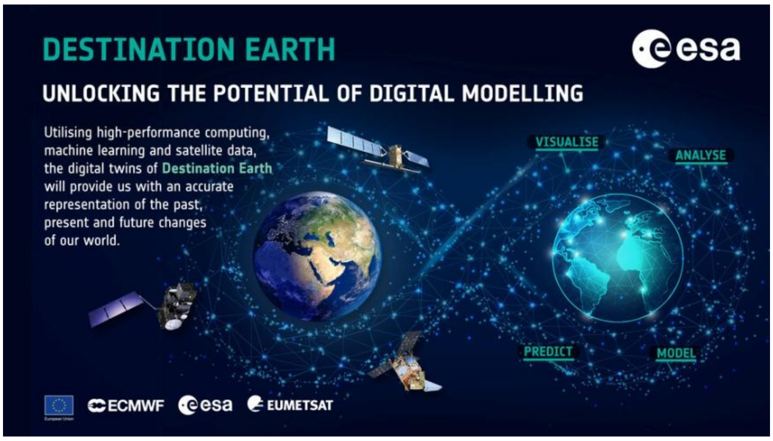
Perhaps the most practical contribution of AI lies in **statistical and dynamical downscaling**, i.e. translating coarse global projections into local and regional climate information required for adaptation planning. Convolutional neural networks and generative adversarial networks (GANs) can reconstruct fine-scale temperature, precipitation, and wind fields using topographic and land-use data. Other models perform **bias correction**, aligning simulated time series with local observational records.

In Italy, early examples include ENEA's AI-based regional projections for Mediterranean drought patterns and the Politecnico di Torino's applications of machine learning to alpine snowpack forecasting. These studies show that AI can complement limited station data and improve the granularity of risk maps for municipalities and energy operators.

6.2 European and Italian Initiatives

Europe is currently leading the integration of AI and climate modelling through large-scale programmes that combine supercomputing, data infrastructures, and open science principles.

Destination Earth (DestinE), launched by the European Commission and implemented by ESA, EUMETSAT and ECMWF, aims to create *digital twins of the Earth* capable of simulating climate and environmental processes with unprecedented precision (European Commission, 2023). AI components are central to this vision: they accelerate numerical solvers, manage uncertainty ensembles, and allow interactive exploration of "what-if" scenarios on exascale computing platforms (EuroHPC JU, 2024).



The **Copernicus Climate Change Service (C3S)**, operated by ECMWF, provides open reanalysis datasets and climate indicators that underpin many AI training workflows. Its sub-programme *AI4Weather* encourages the use of machine learning for bias correction and post-processing of forecasts (ECMWF, 2024).

Italy participates in several related efforts. **ENEA** coordinates research on national climate modelling and adaptation (ENEA, 2023). **RSE** studies the coupling between climate projections and energy-system resilience, particularly in relation to hydropower and transmission-grid vulnerability. The **CNR-ISAC** and universities of Bologna, Turin, and Pisa contribute to regional downscaling and climate-impact assessments. Collectively, these initiatives situate Italy within the European research fabric connecting AI, HPC, and climate services.

6.3 Open Issues and Integration with Climatology

Recent debate within the Group has emphasised that the discussion should move beyond Large Language Models (LLMs) toward emerging Small Language Models (SLMs) and Retrieval-Augmented Generation (RAG). These approaches can deliver targeted, energy-efficient inference, opening the way to agentic AI systems with a smaller environmental footprint.

Furthermore, the critical factor for a renewable-based energy system is not only generation but storage. The intermittency of solar and wind requires scalable accumulation solutions—from fast lithium batteries and pumped-hydro plants to long-term hydrogen storage. The challenge of the energy transition is therefore less scientific than logistical and economic: how rapidly and widely these "energy reservoirs" can be deployed.

Marino Sforna (Terna) shows that the assumption that AI will make weather forecasts increasingly accurate is misleading: highly efficient forecasting models for renewable-energy production have existed for years and are already used daily by TSOs such as Terna; AI adds little or no value in their operational context. The present limitations are essentially twofold: (1) strong dependence on the quality and availability of training data and historical reanalyses, and (2) the intrinsic boundaries imposed by atmospheric chaos—the "butterfly effect"—which constrain long-term predictability (TERNA, 2025).

Therefore, despite its potential, the application of AI to climate prediction faces several open challenges.

1. **Scientific validation.** AI models often outperform traditional ones on specific benchmarks but remain "black boxes". Without explicit physical constraints, their long-term stability and extrapolation under non-observed conditions are uncertain. Climatologists are essential to define metrics and protocols for validation and uncertainty quantification.

2. **Data limitations.** High-resolution observations are unevenly distributed, especially over oceans and the Global South. AI's dependence on data quality may therefore amplify regional biases. Combining remote-sensing data with physical constraints offers a partial solution.

3. **Interpretability and trust.** Decision-makers require transparent indicators of model reliability. Efforts are underway to create **explainable AI (XAI)** methods that reveal which variables drive predictions and to align them with established climate diagnostics (e.g. ENSO indices).

4. **Integration with existing infrastructures.** AI must coexist with long-standing operational workflows at meteorological agencies and climate centers. The emerging

standardisation under ISO/IEC 42001 (Artificial Intelligence Management Systems) and the EU AI Act may help formalise documentation and risk management procedures.

5. **Energy and carbon cost.** Training large AI models for climate prediction is itself energy intensive. A balanced assessment must consider whether computational efficiency gains offset these costs — an issue linking Chapter 3 directly with Chapter 4 on AI's role within the energy system.

Ultimately, AI should be viewed not as a replacement for physical modelling but as an additional layer of intelligence capable of accelerating, refining, and democratising climate knowledge. A collaboration between AI researchers and climate scientists — possibly facilitated by the GdL AIIC through contacts in ENEA, RSE or CNR-ISAC — would ensure that the resulting framework remains scientifically credible and operationally valuable. In recent years, the discussion has shifted from large-scale language models (LLMs), which remain the core of generative AI, toward smaller, more energy-efficient architectures such as Small Language Models (SMLs) and so-called Agentic AI. These approaches aim to preserve reasoning capability while drastically reducing computational demand, data movement, and cooling energy. They represent a pragmatic path toward sustainable AI in climate-related applications.

6.4 References (APA 7th Edition)

DeepMind. (2023). *GraphCast: Learning skillful medium-range global weather forecasting.* **Nature, 626**, 533–539.

ECMWF. (2024). *AI4Weather & Copernicus Climate Change Service (C3S).* Retrieved from https://www.ecmwf.int/

ENEA. (2023). *Progetti per la modellazione climatica nazionale.* Rome: ENEA. Retrieved from https://www.enea.it/

European Commission. (2023). *Destination Earth (DestinE) Implementation Plan.* Brussels: Directorate-General for DG CNECT.

EuroHPC Joint Undertaking. (2024). *High-Performance Computing Initiatives for AI and Climate Integration.* Retrieved from https://eurohpc-ju.europa.eu/

Pathak, J., Subramanian, S., et al. (2022). *FourCastNet: Global medium-range weather forecasting using adaptive Fourier neural operators.* **arXiv:** 2202.11214.

RSE. (2023). *Studi sul collegamento tra modelli climatici e resilienza dei sistemi energetici.* Milan: Ricerca sul Sistema Energetico S.p.A.

TERNA. (2025). *Data center e futuro: la sfida energetica che accende l'Italia.*
Tratto da Lightbox: https://lightbox.terna.it/it/insight/data-center-rete-trasmissione

7 AI AND CLIMATE CHANGE: AI FOR ENERGY SYSTEM OPTIMISATION *(Alberto Stefanini, Elenio Dursi)*[47]

Artificial-intelligence techniques can enhance the operation of modern power systems in three traditional domains of grid control: **forecasting and predictive control**, **real-time optimisation and grid balancing**, and **asset management and lifecycle efficiency**. The underlying principles are classical, but AI offers new ways to combine large data flows and make faster, adaptive decisions. However, public evidence on the magnitude of efficiency gains remains limited, as most applications are recent and commercial. Available EU demonstration projects (for example ENEA, 2022; ETIP-SNET, 2024; https://www.etip-snet.eu/) mainly report methodological advances and partial field results rather than peer-reviewed quantitative assessments.

Forecasting and Predictive Control

Electric power systems must maintain a continuous balance between generation and demand. Forecasting is therefore one of their core functions. Machine-learning algorithms can merge multiple data sources — meteorological forecasts, historical load curves, distributed-generation output, and even mobility or market-price information — to predict electricity consumption and renewable production over time horizons ranging from minutes to several days.

Such forecasts feed **predictive-control systems**, which act before an imbalance occurs. For instance, when a drop in solar radiation is expected, the controller can pre-charge storage units or schedule flexible industrial loads to postpone operation. Unlike traditional rule-based automation, predictive control continuously recalculates its decisions as new data arrive, creating a feedback loop between the physical grid and its digital twin. In transmission and distribution system operators (TSO/DSO) environments, this approach reduces reliance on costly spinning reserves and improves integration of variable renewable energy sources.

European field trials — for example within ENEA’s “AI for Energy Efficiency” initiative (https://www.enea.it/it/Ricerca_sviluppo/energia/efficienza-energetica) — show how short-term solar forecasting can lower curtailment and ramping stress on conventional plants. Still, the robustness of these results under diverse conditions remains to be verified through independent studies.

Real-Time Optimisation and Grid Balancing

While forecasts provide foresight, real-time optimisation ensures second-by-second stability. Here AI augments classical control by analysing massive streams of sensor data — phasor-measurement units (PMUs), smart meters, and substation telemetry — to detect emerging instabilities before they propagate.

A central concept is the **digital twin**, a high-fidelity virtual replica of the grid that mirrors voltages, currents, and component status. The twin allows operators or algorithms to simulate hundreds of possible corrective actions within milliseconds. Reinforcement-learning agents then choose the most effective response based on the simulated outcomes, gradually improving their policy through experience.

---

[47] The authors are indebted to Marino Sforna for his comments and his contributions to the text of this chapter.

For example, if a sudden voltage rise occurs due to a surge of rooftop-PV generation, the AI controller can test alternative inverter-set-point adjustments in the digital twin and implement the one that restores stability with minimal loss. In pilot projects under the ETIP-SNET umbrella and the EU-funded INTERPLAN project, such AI-assisted balancing has reduced renewable curtailment by 10–20 % and cut grid losses by several percentage points. These results, though promising, come mainly from demonstration settings rather than fully commercial systems.

Asset Management and Lifecycle Efficiency

Beyond daily operations, AI supports long-term asset management. Modern substations and transformers are instrumented with sensors that measure temperature, vibration, oil chemistry and electrical discharges. These data feed predictive-maintenance models that estimate the **remaining useful life (RUL)** of each asset. The models, often based on ensemble learning or recurrent neural networks, forecast degradation trends and alert operators when the probability of failure exceeds a threshold.

Integrating such predictions with **risk-based maintenance planning** enables utilities to allocate resources efficiently — repairing components that are both critical and near end-of-life while postponing low-risk interventions. When combined with **Lifecycle Assessment (LCA)** (ISO 14040/44, https://www.iso.org/standard/37456.html) and **Lifecycle Circularity Index (LCI)** metrics, these tools align technical reliability with sustainability goals. Decisions about refurbishing versus replacing equipment can then consider not only cost and downtime but also embodied emissions and material recovery potential.

AI also aids supply-chain optimisation: by analysing procurement data, it can identify suppliers with lower carbon footprints or higher recyclability scores, extending the principles of circular economy to the entire infrastructure lifecycle.

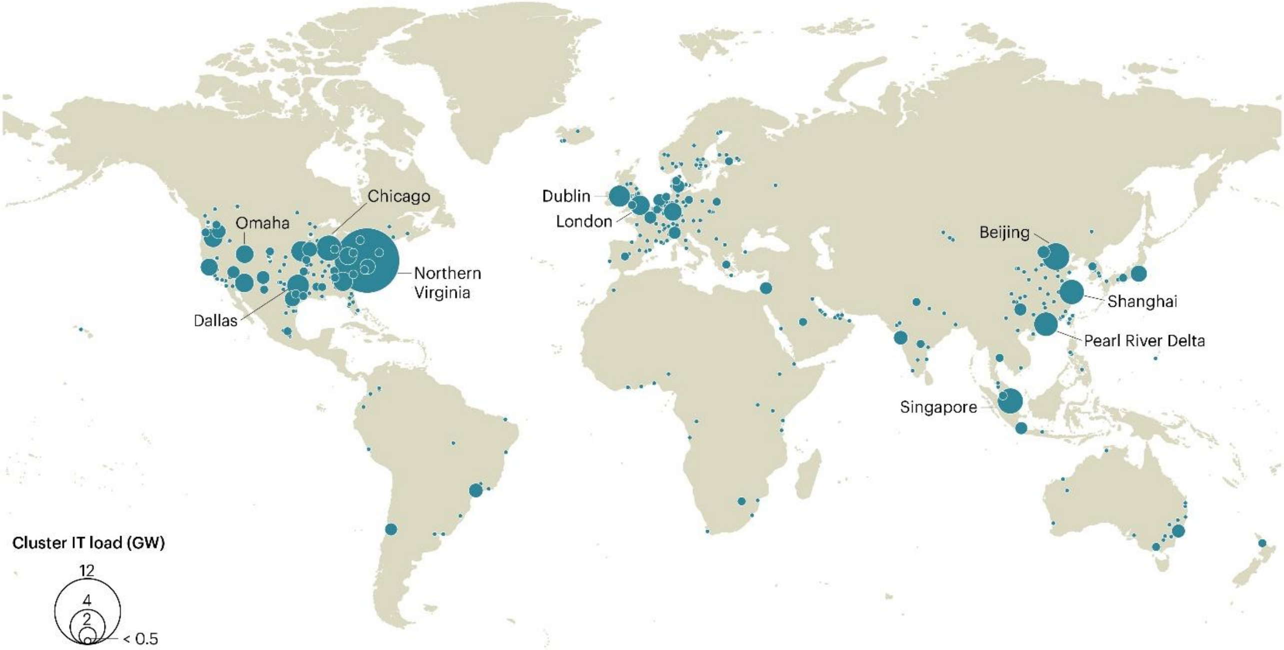
Critical Reflection

Despite numerous pilot projects, large-scale quantitative evidence **confirming that AI consistently reduces emissions or operational costs remains scarce.** Most current deployments are proprietary, and results are communicated through industrial reports rather than peer-reviewed studies. A persistent limitation of renewable sources lies in their intermittency: solar and wind generation cannot be dispatched on demand. In this sense, energy storage represents the “missing link” that transforms intermittent sources into a reliable backbone of the grid. Solutions range from high-capacity batteries for daily management to traditional pumped-hydro systems and emerging hydrogen storage technologies. The challenge is now less technological than logistical, economic, and political, determining how rapidly and pervasively such storage infrastructures can be deployed. Future work should therefore focus on establishing transparent benchmarks and open datasets to evaluate AI’s real contribution to grid efficiency and decarbonisation.

Nevertheless, the examples above illustrate the potential of combining classical control engineering with AI-enabled analytics: a convergence that may gradually transform power systems from reactive infrastructures into predictive, self-optimising ecosystems.

7.1 AI Data Centers and the Electric Grid: Current Impact and Long-Term Outlook (2025–2035)

The rapid expansion of data centers driven by artificial intelligence (AI) workloads is reshaping the European and global electricity landscape. AI inference and training tasks have sharply increased computational intensity and therefore energy demand. The International Energy Agency (IEA) estimates that total data-center electricity consumption rose to around **460 TWh in 2024**, with projections between **700 and 1,000 TWh by 2030**, depending on the trajectory of AI adoption and efficiency gains. This corresponds to roughly **2–3 % of global electricity demand** today and up to **5 % by 2030**, a share comparable to that of medium-sized industrial nations.



Concentration of Data Centers worldwide.
Retrieved from IEA (2023). *Artificial Intelligence and Energy.* Paris: International Energy Agency.

Current Incidence (2025)

In 2025, AI-related computing accounts for about **one fifth** of total data-center load worldwide. Concentrations are highest in Ireland, the Netherlands, Germany, northern Italy, and the Alpine corridor, where abundant connectivity and mild climates favour hyperscale facilities. Most installations remain fully grid-dependent, with less than **15 %** drawing significant on-site renewable generation. Existing self-production usually comes from rooftop solar, small hydro, or power-purchase agreements (PPAs) with nearby renewable plants rather than true energy autonomy.

Medium-Term Outlook (2028)

Over the next three years, the combination of strong AI demand and decarbonisation targets will intensify stress on regional transmission systems. By 2028, data-center consumption in Europe could reach **160–180 TWh**, equivalent to Spain's entire electricity use for industry. Grid operators such as **ENTSO-E** forecast local congestion in high-density clusters around Dublin, Frankfurt, Milan, Zurich, and Vienna. At the same time, the share of *prosumer* data centers—those integrating on-site renewables and storage—is expected to rise to about **30 %**, largely through hybrid solar-battery systems and waste-heat recovery schemes.

Long-Term Outlook (2035)

Assuming continued AI diffusion and partial electrification of ancillary cooling systems, global demand could exceed **1,300 TWh by 2035**, even after efficiency improvements. In this horizon, data centers become both heavy consumers and potential stabilisers of the grid, operating as **flexible prosumers** that can modulate consumption or inject stored energy. Yet this transformation implies substantial network reinforcement: IEA modelling foresees the need for **150–200 GW of new transmission capacity** across the OECD to accommodate the additional load. For the Alpine and Northern Italian regions, the challenge will be balancing new renewable generation with limited space for high-voltage corridors and growing social opposition to new lines.

Critical Issues

Three main risks emerge across timeframes:

1. **Short-term (2025):** voltage instability and local congestion around data-center clusters.
2. **Medium-term (2028):** competition with electrified industry and households for limited renewable supply.
3. **Long-term (2035):** the environmental footprint of grid extensions themselves, as new corridors and substations may encroach on protected ecosystems or landscapes. Thus, the solution to AI's energy hunger—building more lines—could in turn raise environmental and social resistance, complicating the very transition it aims to sustain.

The Role of Quantum and Emerging Technologies

Quantum computing, photonic processors, and specialised neuromorphic chips may reduce energy per computation by one or two orders of magnitude, but deployment at scale before 2035 remains speculative. For the foreseeable future, most AI workloads will still rely on conventional GPUs and large-scale data-center infrastructure, keeping the demand curve upward.

Future Trajectories and Possible Saturation

While mainstream forecasts (IEA 2025; Ember 2025) anticipate sustained growth through the decade, several structural limits may slow expansion after 2030: diminishing marginal utility of larger models, public disaffection, stricter carbon-accounting frameworks, and the rise of **Small Language Models** and **edge inference** reducing reliance on hyperscale training. As efficiency and regulation tighten, AI-related electricity demand is expected to follow an **S-shaped trajectory**—rapid growth through 2030, then gradual stabilisation by 2035. The challenge for policymakers and grid operators will be managing this transition so that today's digital acceleration does not become tomorrow's systemic vulnerability.

7.2 Outlook: AI Data Centers, Grid Evolution, and Long-Term Sustainability (2025–2035)

The global expansion of data centers driven by artificial intelligence is reshaping the demand profile of modern electricity systems. As of 2025, total data-center consumption approaches **460 TWh**, roughly **2–3% of global electricity use**, and could rise to **1,000–1,300 TWh by 2035**, depending on technological and policy trajectories. This increase reflects the combined effects of AI training workloads, constant inference demand, and the electrification of cooling and support systems.

The European grid, particularly in Ireland, Germany, northern Italy, and the Alpine corridor, is already experiencing localized congestion linked to data-center clustering. While regulatory frameworks encourage efficiency and renewable sourcing, less than **15%** of current facilities operate with meaningful on-site generation or storage capacity.

By **2028**, AI-related consumption in Europe could reach **160–180 TWh**, equivalent to Spain's total industrial use. Grid operators anticipate new stress points at key interconnection nodes. However, the rise of hybrid *prosumer* architectures—centers equipped with solar, battery, or hydro assets—is likely to reach **30%** penetration, reducing net dependency on the grid during peak hours.

Looking further to **2035**, the global expansion of digital infrastructure could require **150–200 GW of additional transmission capacity** in advanced economies. This investment wave will itself present environmental and societal challenges, as new lines, substations, and cooling-water systems interact with fragile ecosystems and densely populated areas. The paradox is clear: the digital transition intended to optimise sustainability risks generating new forms of spatial and ecological pressure.

Emerging Technologies and Possible Turning Points

Next-generation computing paradigms—**quantum**, **photonic**, or **neuromorphic** processors—promise to cut energy per computation by one or two orders of magnitude. Yet their commercial diffusion before 2035 remains limited. Consequently, most AI workloads will still rely on conventional GPUs and large-scale data-center clusters for the next decade.

Meanwhile, architectural evolution within AI itself may alleviate energy intensity. **Small Language Models (SLMs)**, **retrieval-augmented generation (RAG)**, and **edge-based inference** distribute intelligence closer to devices, reducing dependence on hyperscale training. This shift, together with more stringent carbon accounting, could flatten the growth curve of energy demand beyond 2030.

Expected Trajectory and Policy Implications

Current projections outline an **S-shaped trajectory**: rapid acceleration through the late 2020s, followed by gradual stabilisation around 2035 as saturation, efficiency, and regulatory constraints converge. The "ceiling" of AI-driven electricity demand is therefore not unlimited, but contingent upon policy choices, innovation in energy management, and social acceptance.

For policymakers and system operators, three actions are crucial:

1. **Incentivise prosumer models** combining data-center reliability with local renewable generation.

2. **Accelerate permitting for grid reinforcements** while integrating environmental safeguards.
3. **Promote energy-aware AI architectures** through standards and public procurement requirements.

In conclusion, the interplay between AI and energy infrastructures is neither linear nor deterministic. The digital and ecological transitions can reinforce each other—but only if efficiency and governance advance as rapidly as computation itself. The challenge for the next decade will be to ensure that the intelligence we build does not outgrow the energy systems that sustain it.

7.3 References (APA 7th Edition)

Ajanovic, A., & Haas, R. (2024). *Digitalisation and artificial intelligence for climate-neutral energy systems.* **Energy Policy, 183**, 113517.

Ember (2025, June). *Grids for Data Centers in Europe*. London: Ember, retrieved from: https://ember-energy.org/app/uploads/2025/06/Grids-for-data-centers-in-Europe.pdf

ENEA (2022). *Artificial Intelligence for Energy Efficiency: Applications and Guidelines.* Rome: ENEA. Retrieved from https://www.enea.it/it/Ricerca_sviluppo/energia/efficienza-energetica

ENTSO-E (2024). *Grid Map Downloads: interconnected network of ENTSO-E*. European Network of Transmission System Operators for Electricity. Retrieved from https://www.entsoe.eu/data/map/downloads/

ETIP-SNET. (2024). *AI for Energy System Integration.* European Technology and Innovation Platform for Smart Networks for Energy Transition. Retrieved from https://www.etip-snet.eu/

European Commission. (2023). *EU Taxonomy for Sustainable Activities: Technical Screening Criteria for Data Centers and ICT.* Brussels.

IEA (2023). *Artificial Intelligence and Energy.* Paris: International Energy Agency. Retrieved from https://www.iea.org/reports/energy-and-ai/understanding-the-energy-ai-nexus

IEA (2025). International Energy Agency (IEA). *Energy and AI: Energy demand from AI*. Paris: IEA, Retrieved from: https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

ISO. (2006). *Environmental management — Life cycle assessment — Principles and framework (ISO 14040/44).* Geneva: International Organization for Standardization.

ISO. (2023). *ISO 42001 — Artificial Intelligence Management System Requirements.* Geneva: International Organization for Standardization.

Rolph, J., Smith, K., & Zhao, Y. (2024). Sustainable computing architectures for AI. *Nature Electronics, 7*(3), 251–264.

Terna S.p.A. (2023). *Piano di Sviluppo della Rete di Trasmissione Nazionale 2023.* Rome: Terna.

ETSI. (2021). *EN 305 174-1 V2.1.1: Environmental Engineering (EE); Measurement Methodology for Energy Efficiency of Data Centers.* Sophia Antipolis: European Telecommunications Standards Institute.

European Parliament and Council. (2024). *Artificial Intelligence Act (Proposal COM/2021/206 final).* Brussels. Retrieved from https://artificialintelligenceact.eu/

Quantum Flagship. (2024). *European Quantum Technology Flagship Programme.* Retrieved from https://qt.eu/

EuroHPC Joint Undertaking. (2024). *High-Performance Computing Initiatives for Quantum and AI Integration.* Retrieved from https://eurohpc-ju.europa.eu/

8 ARTIFICIAL INTELLIGENCE AND CLIMATE CHANGE – IMPLICATIONS FOR THE HEALTHCARE SECTOR *(Silvano Bari)*

This chapter examines how Artificial Intelligence (AI) applied to the healthcare sector can have an indirect yet significant impact on climate change, an impact that can be both positive and negative at the same time.
The considerations to be presented underline the importance of optimizing the use and management of such technologies to reduce the overall environmental footprint of healthcare activities, particularly those related to hospitals.

8.1 Healthcare Sector and Pollution: Why Hospitals Consume So Much Energy

The relevance of the healthcare sector in the global climate crisis is significant: as evidenced by numerous studies, the healthcare sector is responsible for approximately **4-5% of global greenhouse gas (GHG) emissions**.
The OECD report, "Decarbonising Health Systems Across OECD Countries"[48] estimates an average of 4.4% of total emissions in member countries is attributable to hospitals, outpatient clinics, pharmaceuticals, devices, and health services. **Hospitals account for 30% of these emissions**, largely due to complex logistics and the high resource intensity required for continuous operations, including transportation and supply chain activities.

Hospitals, in particular, are major sources of emissions because they must operate continuously, 24 hours a day, all year round, while adhering to strict hygiene, sanitation, and environmental standards: their priority is to guarantee safety, hygiene, and constant comfort for patients, which results in enormous energy consumption. In addition to consuming large amounts of energy for internal uses, they also make massive use of single-use materials, medical transport, pharmaceuticals, and have very complex supply chains.

The largest share of hospital energy consumption (between 40% and 75%) is linked to **Heating, Ventilation, and Air Conditioning (HVAC) systems** because thermal and humidity control (thermo-hygrometric comfort) for patients and staff must be continuously guaranteed, along with high indoor air quality (IAQ) year-round on a 24/7 basis. Furthermore, critical areas (operating theaters, intensive care units, infectious disease wards, laboratories) require sophisticated ventilation management, control over temperature, humidity, and air quality, highly efficient air filtration with high hourly air change rates and the maintenance of differential atmospheric pressures (positive/negative) for contamination control—all leading to significant energy consumption.

Hospitals operate non-stop, which implies the continuous use of **artificial lighting** in every environment, day and night.

Adding to this is the use of **energy-intensive diagnostic and therapeutic equipment**—such Magnetic Resonance Imaging (MRI), Computed Tomography (CT) scanners, and radiotherapy

---

[48] https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/09/decarbonising-health-systems-across-oecd-countries_73dfae4d/5ac2b24b-en.pdf

systems—which not only require high power for operation but also dedicated auxiliary cooling systems for thermal stability.

The energy consumption and CO_2 emissions of an average-sized radiology model generally range between approximately 4 and 10 tons per single diagnosis. This varies depending on the type and efficiency of the equipment, operating mode, usage duration, and energy source (use of renewable energy or not). These estimates primarily refer to imaging technologies such as MRI, CT, X-rays and ultrasound, considering the entire life cycle, from raw material procurement to production, use and disposal.

For instance, an Australian study[49] estimated that the CO_2 emissions associated with a single MRI scan are equivalent to approximately 145 km driven by car in terms of CO_2 equivalent emissions[50].

The requirement for **cold and hot water** in hospitals is enormous and is necessary for constant sanitization, hygiene, and disinfection.

The sterilization of instruments and linens, often based on high-temperature steam, entails a further and significant demand for thermal energy.

Moreover, absolute **operational continuity** mandates the adoption of redundant energy systems — such as emergency generators — that must always be ready for activation in the event of a failure or blackout. Maintaining these backup systems also involves additional consumption and costs.

To direct energy consumption are added the **indirect emissions** related to the **supply chain**: the production, transport, and disposal of single-use materials, pharmaceuticals and various devices. The total energy consumption associated with these activities is very high and represents a challenge for environmental sustainability, especially because the environmental impact throughout the entire life cycle must be considered. For example, the management of single-use materials (gloves, gowns, and other sterile devices, often made of complex, difficult-to-recycle plastics), including their production and disposal, contributes significantly to hospital waste and energy consumption.

Furthermore, the impact of CO_2 emissions from **medical transport and the travel of doctors and patients** represents a significant component. In Italy, transport contributes 28.3% of total national greenhouse gas emissions, which includes health-related travel, with a preponderance of road transport. There are no precise data for health transport alone, but the healthcare sector as a whole has a relevant weight[51].

An example reported by studies in Canada indicates that an in-person patient visit can contribute approximately 13 kg of CO_2 emissions, based on a model that also considers the economic and environmental costs of the journey[52].

The strong energy impact of hospitals is therefore an inevitable consequence of their mission: ensuring safe care, sterile environments, and absolute operational continuity. These needs, combined

---

[49] McAlister et al, The carbon footprint of hospital diagnostic imaging in Australia, Lancet Regional Health Western Pacific 2022, in https://www.thelancet.com/journals/lanwpc/article/PIIS2666-6065(22)00074-8/fulltext

[50] Specifically, regarding MRI scanners, they are high-power devices that consume a lot of energy: a single scanner can emit orders of magnitude of tens of tons of CO_2 per year (an estimate can be around 50–60 tons of CO_2/year, which consists of 14.6 kg CO_2 per scan for a usage of 4000 scans per year. This data depends on the type of machine, energy efficiency and clinical use. However, the use in inference mode (diagnosis, image segmentation) has a negligible impact (~0.5 g per scan). See: Daniel Truhn et al, The ecological footprint of medical AI, in https://pmc.ncbi.nlm.nih.gov/articles/PMC10853292/

[51] Source: https://indicatoriambientali.isprambiente.it/it/trasporti/emissioni-di-gas-serra-dai-trasporti

[52] Estimating Patient and Family Costs and CO_2 Emissions for Telehealth and In-Person Health Care Appointments in British Columbia, Canada: Geospatial Mixed Methods Study, https://www.sciencedirect.com/org/science/article/pii/S1438887125002456

with often obsolete infrastructure and a large quantity of technological equipment, make hospitals among the most energy-demanding buildings.
These data indicate the importance of optimizing equipment usage and employing technologies to reduce the overall environmental footprint of healthcare activities.

In this context, Artificial Intelligence applied to the healthcare sector can have an indirect but significant impact on climate change, an impact that can be simultaneously both positive and negative.

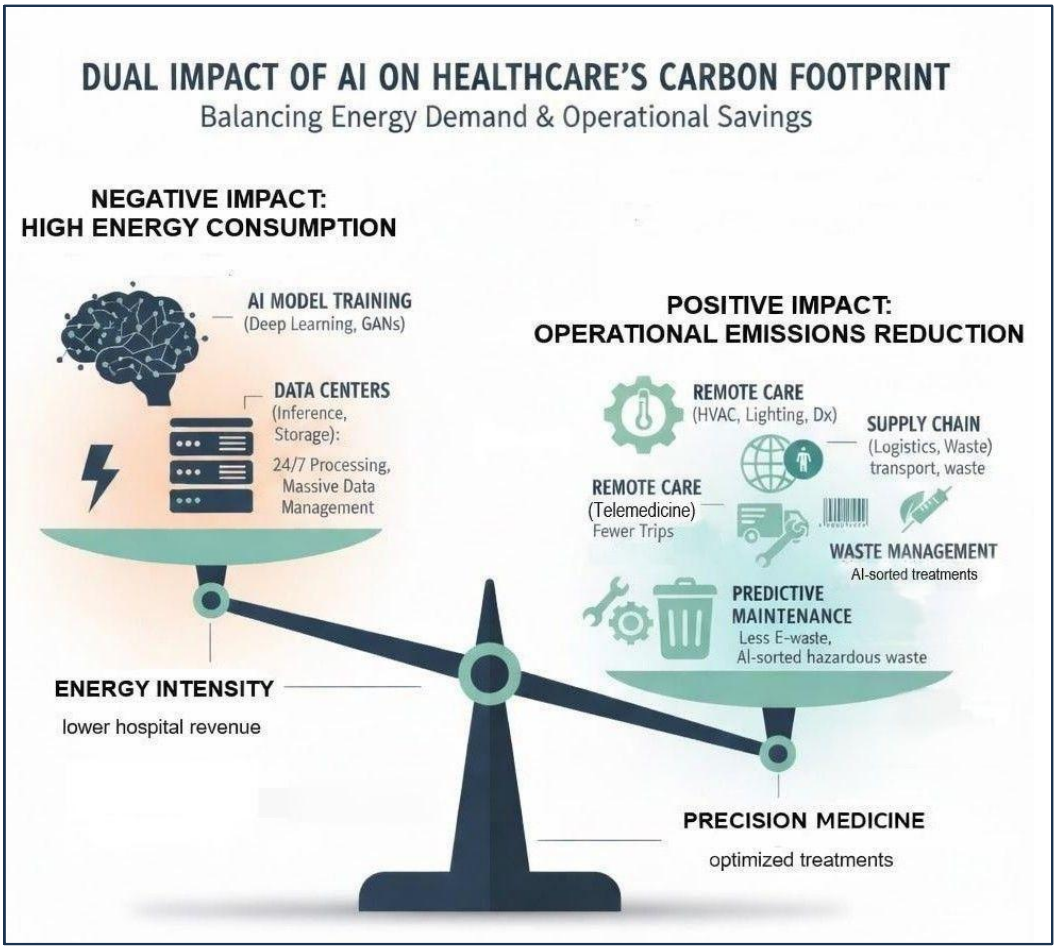


Fig. 8.1 Dual impact of Artificial Intelligence on healthcare carbon footprint.
(Source: Author's elaboration, 2026)

8.2 Healthcare Sector and Artificial Intelligence: Negative Climate Impact

While AI offers advantages such as improved resource management, more efficient diagnostics, and predictive equipment maintenance, it possesses a significant environmental impact, particularly during certain phases of its development and utilization. This impact stems from the energy required to train complex models and manage large amounts of data, thereby contributing to the sector's overall carbon emissions.

Some studies have quantified the **energy consumption of artificial intelligence**, even if not specifically in the healthcare domain, highlighting a significant impact.
A study conducted by the University of Massachusetts Amherst indicates that training a single large AI model can emit up to 284 tons of CO_2, comparing this figure to the environmental impact of other activities, such as transatlantic flights or the use of automobiles over their entire life cycle (e.g. nearly five times the lifetime emissions of the average American car including manufacture of the car itself)[53].

Currently, there are no precise and uniform values specific to the healthcare sector, but this data is particularly relevant for AI applications in health, such as **radiological diagnosis**, **disease prediction**, and **automated triage management**, all of which require intense data processing and, consequently, high energy consumption.

During the **training phase** (the most energy-intensive stage), Deep Learning models, such as the neural networks used for analyzing medical images (CT, MRI) or drug discovery, demand enormous computational power. A large AI model can require days or weeks of intensive computation on infrastructure involving thousands of GPUs (Graphics Processing Units) and Data Centers that consume notable energy, which can translate into a heavy carbon footprint.

Generative models such as Generative Adversarial Networks (GAN) are revolutionizing medical imaging, with numerous applications used to generate synthetic medical images especially in the brain field[54].
While there are still no precise and universal estimates for each specific GAN model in healthcare, some current data show that these deep learning systems can require significant amounts of energy to power the models, with an exponential increase in energy needs over time: it can be said that their energy consumption can reach tens or hundreds of MWh per training cycle, with proportional CO_2 emissions[55].

In the healthcare sector, energy consumption is not limited to the training of AI models but also includes the **processing, storage, and continuous management of large data volumes** during daily use (the so-called **Inference phase**), which necessitates energy-intensive data center infrastructure. The daily use of AI models by users (e.g., a decision support system aiding a physician in diagnosis) is less intensive per single operation, but the total consumption increases with the number of users and frequency of use. Generative AI models used in medicine to summarize clinical records or assist in research are experiencing rapidly growing energy consumption.

In summary, in the hospital context, the overall energy consumption of facilities is already high (approximately 5% of the sector's revenue), and the increasing use of AI escalates its relevance, making energy efficiency strategies necessary.

However, AI can contribute to reducing waste and optimizing resources, such as in the predictive maintenance of equipment and the rational management of medical supplies, leading to indirect environmental benefits.

---

[53] https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/

[54] In detail, GANs are widely used in brain imaging with the synthetic generation of MRI data, used both to expand training datasets and for diagnostic applications such as segmentation of brain tumors and translation of images from one modality (e.g. CT) to another (e.g. MRI). See: Generative Adversarial Networks for Data Augmentation, in https://arxiv.org/pdf/2306.02019

[55] Rongguang Wang et al, Applications of generative adversarial networks in neuroimaging and clinical neuroscience, in https://www.sciencedirect.com/science/article/pii/S1053811923000472

8.3 Healthcare Sector and Artificial Intelligence: How Health AI Can Reduce Climate Impact

While consuming energy to operate, Artificial Intelligence also holds the potential to significantly reduce the overall energy consumption of the healthcare system. AI can mitigate climate impact through a series of measures that reduce CO_2 emissions, not by directly intervening on greenhouse gases, but by optimizing processes, avoiding travel, reducing waste, and improving efficiency.

What are the main measures:

Artificial Intelligence for Energy Efficiency in Hospital Facilities

AI offers targeted solutions for hospital and pharmaceutical facilities, transforming structures into "efficient hospitals" through significant reductions in energy consumption.

System Optimization and Core Functionalities
AI analyzes real-time data to optimize the use of utility systems, which on average account for approximately 75% of a hospital's energy consumption.
These systems automatically optimize the operation of key components (heating, ventilation, lighting, air conditioning). They apply Deep Learning and Predictive Control models, enabling advanced system control and the automatic application of optimal adjustments without the need for operator intervention.
The primary functionalities of AI include:

1. ***Dynamic HVAC Regulation***: AI is utilized to forecast heating, ventilation, and cooling (HVAC) requirements. These predictions are based on real-time data such as space occupancy, weather forecasts, and the building's thermal inertia, allowing for dynamic system regulation.
2. ***Ancillary Optimization***: AI can also be deployed to optimize lighting and predict operating room usage, further reducing energy waste and operational costs.
3. ***Predictive Maintenance***: The systems continuously monitor equipment performance to identify faults or inefficiencies in advance, thereby preventing energy waste caused by malfunctions.

Proven Results and Transfer Potential
A specific case study involved the application of AI to HVAC systems at a multinational pharmaceutical company located in Los Angeles, with the goal of minimizing the environmental impact of its CO_2 emissions[56].
The installed solution optimized the systems' operation across a large campus, resulting in a **reduction of 95 metric tons of CO_2 annually**. This translated into a **16% reduction in energy consumption**, primarily attributed to the HVAC systems, which are the largest consumers in large facilities.
This solution is easily transferable as a savings potential to a typical hospital setting. Hospitals exhibit consumption profiles that are highly similar and complex compared to pharmaceutical facilities, with

---

[56] Source: https://brainboxai.com/hubfs/Case%20Studies/BBAI_Case%20Study_Pharmaceutical.pdf

extreme demands for climate control and ventilation, confirming the high potential for efficiency gains.

In Italy, an energy efficiency project conducted in a large public hospital in Lombardy involved the optimization of HVAC systems through the use of IoT (Internet of Things) sensors and Artificial Intelligence (Machine Learning) algorithms, achieving an overall energy saving of 8% by intervening on the control logic of the systems, during the 2020-2021 heating season.
The results of the study indicate that the adoption of these advanced control strategies can lead to energy savings of up to 45% on the energy consumed by the HVAC systems subject to the intervention. Based on this reduction in energy consumption (kWh not used), the corresponding reduction in CO_2 emissions can be estimated.[57]

Another documented example concerns a hospital in Singapore that focused specifically on optimizing air conditioning (AC), through a system of multi-variable sensors and predictive controls based on Machine Learning.
The AI learned room occupancy patterns and air quality metrics to proactively adjust the air conditioning system. In this way, the hospital achieved more than 20% overall energy savings on AC consumption, while improving comfort.
An energy saving of 20% in an energy-intensive system such as a hospital translates into tens or hundreds of tons of CO_2 avoided per year[58].

Telemedicine and Remote Diagnostics

AI revolutionizes telemedicine and remote diagnostics by transforming how health data is acquired, analyzed, and utilized. It does not just connect patients and doctors but automates, personalizes, and enhances the care pathway and, ultimately, reduces the need for physical transport (fewer ambulances, patient travel). For example, during COVID, the surge in telemedicine eliminated millions of trips, indirectly reducing emissions.

A systematic review conducted in 2021 on 48 studies (68 million consultations) calculated a saving of approximately 692,000 tonnes of CO_2 thanks to the substitution of travel for medical appointments[59].

A UCLA study estimated that telemedicine in 2023 reduced monthly CO_2 emissions equivalent to 61,000–130,000 automobiles, saving up to 47.6 million kg of CO_2 per month[60].

---

[57] R. Frassanito et al, How IoT and Artificial Intelligence can improve energy efficiency in hospitals - a North Italian case study, in https://www.researchgate.net/publication/359088037_How_IoT_and_Artificial_Intelligence_can_improve_energy_efficiency_in_hospitals_-_a_North_Italian_case_study; see also https://www.e3s-conferences.org/articles/e3sconf/pdf/2022/10/e3sconf_52ndaicarr2022_02001.pdf
Currently, this White Paper is the most specific document for a methodological overview on the application of AI for energy efficiency in Italian hospitals.

[58] https://www.sciencedirect.com/science/article/abs/pii/S0306261921013362. See also: https://www.researchgate.net/publication/355821018_Machine-learning-based_model_predictive_control_with_instantaneous_linearization_-_A_case_study_on_an_air-conditioning_and_mechanical_ventilation_system

[59] S. Rodler et al, The Impact of Telemedicine in Reducing the Carbon Footprint in Health Care: A Systematic Review and Cumulative Analysis of 68 Million Clinical Consultations, in https://pubmed.ncbi.nlm.nih.gov/38036339/

[60] Telemedicine had an impact on carbon emissions equivalent to reducing up to 130,000 car trips each month in 2023 - Internal Medicine | UCLA Health, in https://www.uclahealth.org/news/release/telemedicine-had-impact-carbon-emissions-equivalent-reducing

A US hospital, Stanford Health Care (SHC), reduced emissions from in-person visits by 36% between 2019 and 2021, moving to 0.02-0.04 kg CO_2 per virtual visit (versus 20 kg for an in-person visit)[61]. This reduction is represented in the figure below which compares the trend of visits with the trend of total greenhouse gas emissions, showing how the increase in telemedicine has made it possible to reduce the overall environmental impact despite a greater number of appointments.

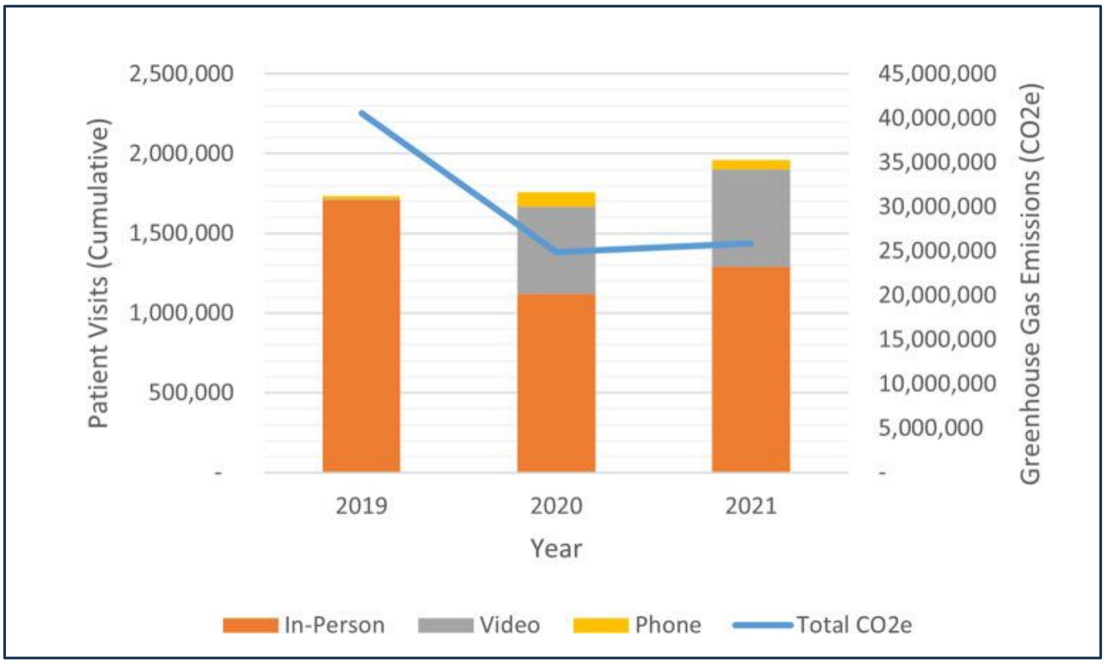


Fig. 8.2 Number of Patient Visits and Cumulative Estimated GHG Emissions.
(Source: reprinted from Cassandra L Thiel et al, 2023)

In fact, the research shows that total visits increased by 13% (from 1,733,020 in 2019 to 1,961,768 in 2021) but, despite this increase, total GHG emissions decreased by 36%, from about 40,600 tons of CO_2e in 2019 to 25,900 tons in 2021. This result is due to the massive growth of telemedicine, supported by Artificial Intelligence systems, which has replaced many face-to-face visits. Emissions decrease even if visits increase for the following reasons:

1. an *in-person visit* produces an average *of 20 kg of* CO_2e (mainly due to patient travel);
2. a *telephone visit* produces only *0.02 kg of* CO_2e.
3. a *video visit* produces *0.04 kg of* CO_2e.

So, when many visits are converted to virtual-visits, the total emissions drop dramatically.

Supply Chain and Pharmaceutical Management

AI can be used to improve the logistics and inventory management of pharmaceuticals and materials (such as single-use devices), taking into account consumption and expiration dates. This way, the need for production and transport—which are major sources of emissions in the healthcare supply chain—is reduced, decreasing the manufacturing of materials and the waste of drugs and supplies, and consequently, CO_2 emissions as well.

Predictive Maintenance

---

[61] Cassandra L Thiel et al, Telemedicine and the environment: life cycle environmental emissions from in-person and virtual clinic visits, in https://pmc.ncbi.nlm.nih.gov/articles/PMC10169113/

Smart hospitals that prevent malfunctions in medical machinery thanks to AI avoid premature replacements and reduce electronic waste (e-waste).

Precision Medicine

AI can enhance diagnostic accuracy and personalize treatment plans, reducing the need for ineffective tests or treatments and thereby decreasing resource consumption and the energy associated with the cycle of care.

Another area where AI has a direct and quantifiable impact is the use of anesthetic agents, which are potent greenhouse gases. A Clinical Institute in Italy uses an AI solution to optimize the use of low-flow anesthesia: anesthesiologists are supported in the automatic and precise control of gas administration. This not only reduces the costs of anesthetic agents, but also their greenhouse gas emissions (CO2 equivalent) into the environment[62].

Healthcare Waste Management

AI and computer vision can improve the recycling and management of infectious and hazardous waste, which otherwise requires high-energy incineration. Predictive algorithms based on historical data and smart sensors aid in the planning of collection and disposal, forecasting peaks and optimizing routes and intervention times. AI-equipped robots can automate the separation of hazardous waste at treatment facilities, reducing operator exposure to risky materials and increasing the precision in recycling recoverable materials, such as metals or toxic electronic components. This reduces operating costs, consumption, and CO_2 emissions compared to traditional methods[63].

8.4 Conclusions

The integration of Artificial Intelligence in the healthcare sector represents a paradigm shift with profound environmental implications. While the computational intensity of AI models - particularly during the training and inference phases - introduces a new source of energy demand and carbon emissions, the evidence suggests that its systemic benefits can far outweigh these costs.

Through the optimization of energy-intensive HVAC systems, the radical reduction of travel-related emissions via AI-enhanced telemedicine, and the refinement of pharmaceutical supply chains, AI acts as a powerful catalyst for healthcare decarbonization. However, the transition toward "Green AI" in medicine requires a conscious approach: healthcare facilities must prioritize energy-efficient algorithms and sustainable data management to ensure that technological progress does not come at the expense of environmental health. Ultimately, AI should not be viewed merely as a clinical tool, but as a strategic asset for building a more resilient and sustainable global health system.

---

[62] https://www.gehealthcare.it/-/jssmedia/gehc/uk/files/products/carestation-insights/new-carestation-insights_white-paper_porta-sole_italy_jb23747xx.pdf?rev=-1

[63] https://www.beataladifferenziata.it/it/lintelligenza-artificiale-nella-gestione-dei-rifiuti-innovazione-e-sostenibilita

9 EDGE COMPUTING AND CLIMATE CHANGE - AN INDUSTRIAL PROJECT *(Lorenzo Vandoni, Gabriele Balzano)*

The rapid expansion of digital services, artificial intelligence, and ultra-low-latency applications is putting increasing pressure on conventional hyperscale data centers. Centralized architectures are energy-intensive, water-dependent, and often disconnected from the local contexts in which data are generated and used. At the same time, regulatory requirements on data sovereignty, security, and auditability are becoming stricter, particularly in Europe.

This project proposes an industrial alternative based on distributed edge computing. Instead of concentrating computation and storage in large, remote facilities, the model relies on small, modular edge server rooms deployed close to users, networks, and energy sources. Each edge node operates as a compact "server farm", powered primarily by renewable energy and integrated into the local energy system. Waste heat generated by computing equipment is recovered and reused for local purposes, such as aquaculture, greenhouses, or district heating, turning digital infrastructure into an active component of the local circular economy.

The edge infrastructure hosts latency-critical and data-sensitive applications. These include 5G services requiring real-time responsiveness, such as autonomous or assisted driving, as well as Immutable Data Vaults (IDVs) built on distributed technologies such as IPFS, enabling secure, auditable, and resilient data storage. Artificial intelligence workloads are deployed directly within the edge farm, allowing data to be processed locally and avoiding the transfer of sensitive information to centralized cloud platforms.

By combining edge computing, renewable energy, heat recovery, distributed data management, and local AI, the project defines a new class of industrial digital infrastructure. The objective is not only to reduce environmental impact, but also to increase resilience, regulatory compliance, and technological sovereignty, while enabling advanced digital services where they are most effective: at the edge.

The project is promoted by **Hal Service**, a technology-oriented company focused on the design and integration of advanced digital solutions for industrial and enterprise environments. Hal Service operates at the intersection of IT infrastructure, automation, and applied innovation, with particular attention to emerging technologies, sustainability, and regulatory compliance. Through this initiative, the company aims to extend its role from system integration to the development of proprietary industrial platforms, positioning edge computing as a strategic asset for resilient, low-impact, and locally integrated digital services.

9.1 Edge Computing

Edge computing is a distributed computing paradigm in which data processing, storage, and AI inference are performed close to where data are generated and consumed, rather than in centralized hyperscale data centers. The primary objectives are reduced latency, improved reliability, data sovereignty, and tighter integration with local energy and industrial systems.

From a technical perspective, an edge computing facility is built as a modular micro–data center. The core components include:

- a small cluster of servers equipped with CPUs and, where required, GPUs or AI accelerators;
- high-speed local storage, typically NVMe-based, with redundancy and replication across nodes;
- networking equipment supporting fiber backhaul and 5G connectivity, with deterministic latency and traffic prioritization;
- containerized and virtualized software stacks (e.g. Kubernetes-based orchestration) enabling dynamic workload placement and remote management;
- monitoring and control systems for power, thermal conditions, and cybersecurity.

Unlike traditional data centers, the edge node is dimensioned for local demand. Typical installations range from a few tens to a few hundreds of kilowatts of IT load, allowing air or liquid-assisted cooling without the complexity of hyperscale infrastructure. This scale makes it feasible to colocate the server room with renewable energy sources and heat-recovery systems.

Cost structures reflect this modularity. Capital expenditure is dominated by IT hardware, power electronics, cooling, and grid connection. For an industrial-grade edge node, initial investment can be one order of magnitude lower than a conventional data center, while operating costs benefit from reduced energy consumption, local renewable generation, and the absence of large-scale cooling and water systems. The modular design also allows incremental expansion, aligning investment with real demand.

The project foresees the construction of two pilot edge facilities. The first prototype will be deployed in an industrial warehouse acquired specifically by Hal Service. This site will host the edge server room together with a Fab Lab dedicated to prototyping, testing, and skills development, and a photovoltaic plant integrated into the building. The co-location of digital infrastructure, renewable generation, and manufacturing-oriented spaces is intended to demonstrate an integrated industrial model. The second prototype will be installed in proximity to a hydroelectric power plant. In this configuration, the edge facility can directly leverage continuous renewable generation, test tight coupling between computing workloads and energy availability, and validate operation in a context where grid stability and resilience are critical. Together, the two prototypes will serve as reference implementations for scalable, energy-integrated edge computing in industrial environments.

9.2 Energy Aspects: Renewable Supply and Heat Reuse

The energy model of the proposed edge infrastructure is based on the direct coupling between computing workloads and local renewable energy sources. Unlike hyperscale data centers, which rely on large and often distant power plants, the edge facilities are designed to operate as energy-aware industrial systems, integrated into their surrounding environment.

Power supply is primarily provided by on-site or nearby renewable generation. In the industrial warehouse prototype, this is achieved through a photovoltaic plant installed on the building and adjacent areas, combined with inverters and local energy management systems. In the hydroelectric prototype, the edge node is connected directly to a renewable source with continuous and predictable output. In both cases, the objective is to maximize self-consumption, reduce dependency on the grid, and align computational loads with real-time energy availability through intelligent scheduling.

From a technical standpoint, the edge facilities are equipped with advanced power electronics, uninterruptible power supplies, and monitoring systems that allow fine-grained control of energy flows. Non-critical workloads, such as batch AI training or data synchronization, can be shifted in time to periods of high renewable production, while latency-critical services remain continuously available. This approach reduces peak demand, lowers operational costs, and minimizes the carbon intensity of computing activities.

A central element of the energy strategy is the reuse of waste heat generated by servers. Even at small scale, edge computing systems produce a stable and valuable thermal output. Instead of dissipating this heat through conventional cooling, the project integrates heat recovery systems that capture low- to medium-temperature heat and make it available for local use.

Recovered heat can support applications such as aquaculture systems, greenhouses, space heating for industrial or office areas, or pre-heating processes in nearby facilities. In the warehouse-based prototype, heat reuse can directly supply the Fab Lab and adjacent spaces. In the hydroelectric context, thermal integration can complement existing infrastructure or serve nearby users. This transforms the edge facility into a digital–energy hybrid asset, where electricity is converted not only into computation, but also into useful thermal energy.

By combining renewable power, energy-aware workload management, and systematic heat reuse, the project demonstrates how digital infrastructure can evolve from an energy consumer into an active participant in local energy ecosystems, improving overall efficiency and environmental performance.

9.3 Applications: AI, Data Integrity, and Ultra-Low Latency Services

The edge infrastructure is designed to host applications that benefit directly from local computation, data proximity, and deterministic performance. The focus is on use cases where latency, data sovereignty, and energy efficiency are critical, and where centralized cloud platforms introduce technical, regulatory, or economic constraints.

A primary application domain is artificial intelligence, with particular emphasis on the local hosting of Large Language Models (LLMs) and other AI inference and training workloads. By deploying AI models directly within the edge farm, sensitive industrial, operational, or personal data can be processed without being transferred to external cloud environments. This approach reduces exposure risks, simplifies compliance with data protection regulations, and enables real-time AI services tightly coupled with local processes. Typical use cases include industrial decision support, predictive maintenance, natural-language interfaces for operators, and AI-assisted control systems. The availability of GPUs or dedicated accelerators at the edge allows scalable AI performance while maintaining full control over data and models.

Immutable Data Vaults (IDVs) represent a second core application layer. Built on distributed storage technologies such as IPFS, IDVs provide resilient, verifiable, and tamper-evident data storage. Data and model artifacts stored in the vault are cryptographically addressed and replicated across nodes, ensuring integrity and long-term availability. This architecture is particularly suited for regulatory compliance, traceability of AI models and datasets, and secure data sharing among multiple stakeholders. The combination of edge computing and IPFS-based IDVs enables decentralized yet coordinated data management, independent from large centralized data centers.

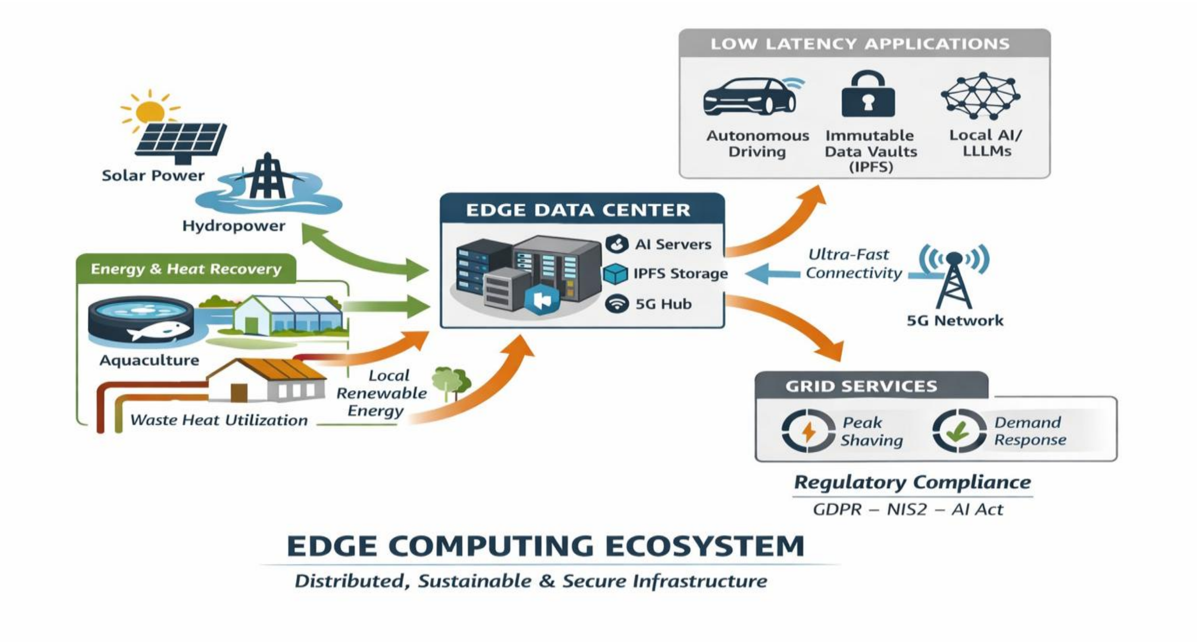
The infrastructure also supports 5G-enabled services requiring ultra-low latency and high reliability. By placing computing resources close to the radio access network, the edge nodes can host applications such as autonomous or assisted driving, smart mobility, industrial automation, and real-time video analytics. In these scenarios, millisecond-level latency is essential for safety and performance, and cannot be guaranteed by remote cloud infrastructures. The edge facility acts as a local processing hub, enabling fast decision loops and resilience even in the presence of network disruptions.

Together, AI at the edge, IPFS-based IDVs, and 5G low-latency services form a coherent application ecosystem. The edge infrastructure becomes a programmable industrial platform, capable of supporting advanced digital services while preserving control over data, energy use, and operational continuity.

9.4 Industrial Perspective and Climate Change

The proposed industrial project addresses climate change by rethinking the structure and governance of digital infrastructure. Rather than scaling through centralized hyperscale data centers, it adopts a distributed edge model in which computing capacity grows in proportion to local demand and available renewable energy. This reduces grid congestion, transmission losses, and systemic risk, while enabling compliance with European requirements on data sovereignty, security, and auditability. Immutability and local data processing become intrinsic properties of the infrastructure, simplifying regulatory alignment without adding procedural overhead.

From an industrial and economic perspective, the model is designed to be scalable and replicable. Modular edge nodes can be deployed as a territorial network, supporting SMEs with shared yet sovereign digital infrastructure and lowering barriers to advanced AI and data services. Energy-aware workload management and heat reuse improve overall efficiency, while the ability to provide flexibility to the electrical system (e.g. peak shaving and load shifting) aligns digital growth with climate and energy policy objectives. In this configuration, digital infrastructure evolves from a climate stress factor into an enabling asset for a low-carbon, resilient, and regulation-ready industrial ecosystem.



10 CONCLUSIONS *(Sandro Bologna, Alberto Stefanini)*

The work carried out by the AIIC Working Group on Artificial Intelligence and Climate Change reflects the intrinsic complexity and multidimensionality of the subject. Artificial Intelligence is not a single technology, nor does it provide a single answer. Rather, it comprises a broad family of methods—statistical, computational and increasingly hybridised with physics-based models—that can support understanding, forecasting, mitigation and adaptation in the context of global climate change. Across the sections of this Report, several domains have emerged in which AI is already delivering tangible contributions.

First, AI strengthens climate science by enhancing the analysis of long-term environmental datasets and improving specific components of climate modelling. Machine learning does not replace Earth-system models but can increase spatial and temporal resolution, reduce computational costs, and accelerate scenario exploration, provided that physical interpretability is preserved.

Second, AI plays a growing role in managing the impacts of climate change on critical infrastructures. The energy sector, in particular, is experiencing a rapid rise in computational demand driven by data centers, digital platforms and the widespread adoption of generative AI. This introduces new challenges for electricity grids, water usage and heat management, but also opens opportunities for optimisation, demand response and the coordination of distributed renewable resources. AI can support operators in sizing networks, forecasting stress conditions and orchestrating flexibility in real time.

Third, AI is increasingly used in domains where climate change directly interacts with societal needs: healthcare, civil protection, agriculture and early-warning systems. The examples discussed in this Report show that AI can support decision-making, enhance monitoring capabilities and improve preparedness during extreme events, reducing exposure and vulnerability.

Fourth, the pilot applications explored—such as the reuse of existing industrial infrastructures for data-center operations and the proposed Immutable Data Vaults—demonstrate that the convergence of AI and critical-infrastructure management is already triggering concrete innovation paths. These early experiences are partial but promising, and they highlight the value of controlled experimentation in real-world settings.

Despite these advances, significant challenges remain. The rapid evolution of AI raises concerns regarding energy consumption, water use, supply-chain constraints, ecosystem impacts, cybersecurity vulnerabilities and the difficulty of integrating highly dynamic digital systems into infrastructures traditionally designed for stability and long-term planning. Moreover, the expertise required spans climate science, engineering, ICT, environmental management, data governance and regulatory frameworks. No single organisation—or Working Group—can cover all dimensions exhaustively; broader interdisciplinary collaboration will therefore be essential.

Looking ahead, two lines of work appear particularly crucial.

Strengthening the link between AI methods and domain knowledge in climate science, ensuring that novel approaches remain physically grounded, interpretable and trustworthy.

Broadening the scope of analysis across critical infrastructures—including water systems, transport networks and urban environments—which will be deeply affected by climate shifts and where AI-enabled tools have potential yet remain underexplored.

It is also important to keep expectations realistic. According to the International Energy Agency (IEA), the emissions reductions achievable through the widespread adoption of existing AI applications could amount to roughly 5% of energy-related emissions by 2035—a significant contribution, yet insufficient on its own to meet global climate targets. Barriers to adoption, governance, interoperability and operational integration will need to be overcome to unlock even these gains. AI can be a meaningful tool for emissions reduction, but it is not a silver bullet and does not remove the need for proactive policy and investment.
Finally, as data-center expansion continues, further work will be needed to explore AI-based solutions for water-efficiency optimisation, in line with emerging metrics such as Water Usage Effectiveness (WUE). This topic sits at the crossroads of digitalisation and environmental stewardship and deserves specific attention in future editions of the Report.

In conclusion, AI represents both an opportunity and a responsibility. If adequately understood, governed and integrated, it can contribute to more resilient infrastructures, more efficient resource management and more informed climate strategies. The Working Group hopes that this Report will serve as a constructive reference for future initiatives, and as a starting point for deeper cooperation among researchers, practitioners and institutions engaged in addressing one of the defining global challenges of our time.

The Working Group recognizes and thanks all the previous AIIC Working Groups for the work already done, producing a set of AIIC Reports about the issue of Critical Infrastructure Resilience. All the Reports can be downloaded from AIIC web site.

AUTHORS

Silvano Bari

Graduated in Statistical and Demographic Sciences and Master's Degree in Computer Law, former Head of Security and Privacy of Alitalia - Italian Airlines. Today he is a professor of "Risk Assessment" at the "Campus Bio-medico di Roma" University and vice president of AIIC (Italian Association of Critical Infrastructure Experts). He is CISM certified.

Glauco Bertocchi

Degree in Physics at the University of Rome "la Sapienza" More than 40 years of experience in IT and security acquired within universities and national institutions. Active in research in the field of protection and resilience of critical infrastructure. Vice-President of the ISACA Rome chapter, member of the AIIC Board of Directors.

Sandro Bologna

Graduated in Physics at the Sapienza University of Rome. The main research activity concerns Critical Infrastructure Resilient to Disasters of different nature, with particular attention to the modeling, simulation and analysis of vulnerabilities and interdependencies, as well as the use of different technologies in order to increase their security.

Luigi Carrozzi

(CGEIT, CRISC, Auditor L.A. ISMS, ISMS Senior Manager, Privacy Officer) Graduated in Statistical Sciences, he has over 35 years of experience in ICT Governance, Risk Management and Compliance at leading private and public organizations.

Raffaella D'Alessandro

Degree in Economics. 40 years of experience in ICT, of which 35 years in Cybersecurity, Data Protection, Management Systems, Standards and Digital Law Compliance. She worked at Olivetti, Arthur Andersen, Ernst & Young and IBM. Member of the Board of Directors of AIIC, Founding Member and Secretary of TOPForGrowth,

Elenio Dursi

IT project manager and Scientific Committee Board Member at CLUSIT. Manages cyber security and the architecture of advanced solutions for intranet and internet infrastructures.

Tommaso Diddi

Graduated in Information Engineering with a curriculum in telecommunications. He has gained experience in the audio and electronics sector. He holds the role of Junior Software Developer at Hermes Bay s.r.l.

Beatrice Rosa

Graduated with honors from LUISS Guido Carli, she has studied copyright, data protection and innovation law, with attention to the impact of AI, new technologies and sustainability in the media and copyright sectors. She is currently a trainee lawyer at the E-lex Law Firm, combining legal skills and innovative vision.

Adriana Peduto

Lawyer, partner of the E-lex firm, he deals with privacy and new technology law, with many years of experience in copyright and consumer law. ISO/IEC 27001:2022. Lead Auditor, she is Data Protection Officer (DPO) of public institutions and private companies. He deals with ethics & privacy compliance profiles in EU research projects.

Alberto Stefanini

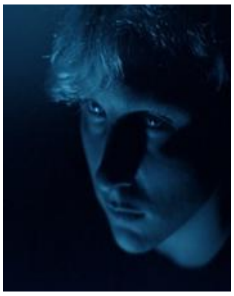
Degrees in Electronic Engineering and Classical Studies. He has extensive experience in national and EU research centers, where he has mainly worked on collaborative proposals and projects in automation, AI, energy, and critical infrastructure security. His current focus include AI applications, the economics of cybersecurity and digital editions.

Cristina Turconi

Legal Counsel with over thirty years of experience gained as an in-house lawyer at a leading Italian group operating in the airport sector.
Specialized in airport law and aviation, with a solid expertise in contracts and an in-depth knowledge of technical and sector regulations. Active member of AIIC and TOPForGrowth.

Maria Beatrice Versaci

She holds a master's degree in Oriental Languages and Civilizations (Arabic) from La Sapienza University in Rome, then specialized in Strategic Protection of the Country System (Cyber Intelligence, Big Data and Security of Critical Infrastructures) at the Italian Society for International Organization (SIOI).



Gabriele Balzano

Senior Research & Innovation Manager at Hal Service Spa, where he leads research initiatives, technological innovation and integration of new ICT solutions. He has experience in business development and project management in digital and telecommunications services.

Lorenzo Vandoni
IT professional with over 35 years of experience in software development and project management: skills in C++, .NET and management applications. He has held roles as project manager, R&D director, PMO and Quality Assurance at Emisfera, an Italian software house. He currently works as a freelancer, collaborating mainly with Hal Service Spa.

AIIC
ASSOCIAZIONE ITALIANA
ESPERTI IN INFRASTRUTTURE CRITICHE

febbraio 2026