Background Guide

The Endpoint of AI: Photovoltaics and Solar Energy

In the contemporary global context, digitalization and green transitions are becoming the twin engines driving a new round of technological, energy, and industrial revolutions, continually propelling global technological innovation and governance transformation. The report from the 20th National Congress of the Communist Party of China emphasized the acceleration of digital economy development and its deep integration with the real economy, alongside a hastened transition towards greener methods of development. The “dual carbon” goals and “Digital China” initiatives stand as China’s core strategies for high-quality development pathways, each distinct yet inherently interconnected. On one hand, the global impacts of climate change present significant challenges to human development, prompting major countries and regions to accelerate their steps towards carbon neutrality. On the other hand, represented by information technology, the new wave of technological and industrial revolutions is advancing rapidly, intensifying the competition for key core technologies and causing turbulence in digital industry chains, as countries actively enhance their digital governance systems and foster new drivers for digital development.

Recent technological breakthroughs in large language models herald a disruptive era for AI’s impact on human life. In 2023, the generative artificial intelligence ChatGPT became a focal topic in the tech community and a leading trend in future AI technology. Released publicly on November 30, 2022, ChatGPT rapidly attracted approximately 100 million active users within two months, making it the fastest-growing consumer application in history. In May 2023, NVIDIA, known for its production of “superchips” optimized for AI applications, saw its market value surge to one trillion dollars.

Despite AI’s promising role in sustainability, the accelerated digital transformation is driving greater energy demand and increasing carbon emissions. The massive energy
consumption during large model training and usage, along with the supporting infrastructures like data centers and 5G stations, continues to grow, raising societal concerns. The rapid development of artificial intelligence and digital technologies results in significant energy consumption, drawing increasing research attention to the substantial tension between AI technology development and the environment (Vinuesa et al., 2020). Jensen Huang, founder of NVIDIA, publicly stated, “The endpoint of AI is photovoltaics and energy storage. Don’t just think about computing power—if only considering computation, we would need the energy of 14 Earths.” Without major advancements in the energy sector, AI development cannot surpass theoretical and technological limitations. Sam Altman, father of ChatGPT, noted: In the future, computational power and energy will become two key “currencies” that are interconvertible.

As debates over artificial intelligence and energy intensify, a consensus is emerging in the tech community: the limit of AI is energy, and the limit of computing power lies in electricity, including photovoltaics, energy storage, and nuclear fusion. A single training of GPT-3 emits the equivalent CO2 of the average human’s annual carbon emissions 91.1 times over, equivalent to 1976 round-trip flights from Beijing to Shanghai. The electricity consumed by one GPT-3 training could power a typical Chinese family of four for 340 years. Regular usage emissions far exceed those of training large models.

For instance, the emissions from using a model like ChatGPT can exceed its training emissions in just a few weeks. Generating images is one of the most carbon-intensive tasks in artificial intelligence. For example, producing 1000 images with a generative AI model might release a significant amount of CO2, equivalent to the emissions produced by driving a car for 6.6 kilometers. Considering these generative AI models are used millions or even billions of times daily, their negative environmental impact is alarming. According to the International Energy Agency's (IEA) latest power report in 2024, global data centers, cryptocurrencies, and AI electricity consumption amounted
to 460 terawatt-hours in 2022, accounting for 2% of global electricity usage, with projections increasing to 620-1050 TWh by 2026. This increase is equivalent to the annual electricity consumption of Sweden at its minimum and Germany at its maximum.