UDC 004.8:504

223. THE ENVIRONMENTAL COST OF ARTIFICIAL INTELLIGENCE AND LARGE-SCALE COMPUTING

Danchenko O.A., student of the gr. 521901, Omelyanchuk A.V., student of the gr. 521701

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Yeliseyeva V.A. - Senior Lecturer, Master of Arts

Annotation. The rapid development of artificial intelligence and cloud computing has significantly increased demand for computational resources and data centre infrastructure. This study examines the environmental impact of large-scale computing, focusing on energy consumption, carbon emissions, water usage, and electronic waste. It analyses statistical evidence to estimate the ecological costs of training and operating machine learning models. Particular attention is given to growing computational demands and global infrastructure expansion. The paper also discusses strategies for improving sustainability, including algorithm optimisation, renewable energy use, and advances in data centre design.

Keywords. artificial intelligence, machine learning, data centres, energy consumption, carbon emissions, large-scale computing, sustainability, green computing.

During the last decade artificial intelligence has become one of the most rapidly developing technological fields. Machine learning systems are now widely applied in medical diagnostics, logistics optimisation, natural language processing, financial forecasting, industrial automation, and scientific research. These technologies are capable of analysing vast quantities of information and identifying patterns that are difficult for humans to detect. As a result, artificial intelligence is increasingly considered a fundamental component of modern digital infrastructure. However, the rapid expansion of artificial intelligence capabilities has required the development of increasingly powerful computing systems. Large technology companies and research institutions invest substantial financial resources in the construction of new data centres and high-performance computing facilities. These infrastructures allow complex machine learning models to be trained and deployed at a global scale. Although the technological advantages of artificial intelligence are widely recognised, its environmental consequences have only recently become a subject of systematic scientific investigation. Artificial intelligence relies on large-scale computational processes that require substantial energy resources. Training advanced neural networks involves processing extremely large datasets through millions or even billions of mathematical operations. This process requires specialised hardware operating continuously for long periods of time. Consequently, the environmental cost of artificial intelligence is increasingly associated with energy consumption, greenhouse gas emissions, water usage for cooling systems, and the generation of electronic waste. For this reason, many researchers emphasise the importance of evaluating the sustainability of artificial intelligence development in order to minimise its environmental impact.

One of the most significant environmental factors associated with artificial intelligence is electricity consumption. Training modern machine learning models requires specialised processors such as graphics processing units (GPUs) and tensor processing units (TPUs). These processors are designed to perform large numbers of parallel mathematical calculations and therefore require substantial electrical power. In large data centres thousands of such processors operate simultaneously in order to accelerate complex computational tasks. Global demand for cloud computing services has increased considerably as artificial intelligence technologies become integrated into commercial platforms, digital assistants, recommendation systems, and automated decision-making tools. According to the International Energy Agency, global data centres consumed approximately 460 terawatt-hours of electricity in 2022, which represents around two percent of global electricity demand [1]. This level of consumption is comparable to the annual electricity usage of entire industrialised countries. Although improvements in hardware efficiency have partially compensated for the growing demand for digital services, the total energy consumption of global computing infrastructure continues to increase.

The training phase of artificial intelligence models is particularly energy intensive. During this stage algorithms repeatedly process large datasets in order to optimise the parameters of neural networks. Researchers analysing the environmental impact of machine learning estimated that training a large natural language processing model can generate approximately 284 tons of carbon dioxide equivalents depending on the carbon intensity of electricity production [2]. Such emissions are comparable to the lifetime carbon footprint of several passenger vehicles. Furthermore, the size and complexity of machine learning models have increased dramatically during the past decade.

Contemporary large language models often contain hundreds of billions of parameters, which significantly increases computational requirements and energy consumption. Another environmental issue associated with artificial intelligence involves the physical infrastructure of data centres. Artificial intelligence systems require large facilities that contain thousands of high-performance servers and network systems. These devices generate considerable amounts of heat during operation because electrical energy is partially converted into thermal energy. In order to maintain safe operating temperatures and prevent hardware damage, data centres employ complex cooling systems. Cooling infrastructure represents a substantial portion of the total energy consumption of computing facilities. Studies indicate that cooling processes may account for up to forty percent of the total electricity usage of a typical data centre [3]. Various cooling technologies are currently employed, including air-based cooling systems, chilled water circulation, and advanced liquid cooling solutions. Improvements in cooling efficiency can therefore significantly reduce the environmental impact of computing infrastructure. In addition to electricity consumption, data centres frequently require large volumes of water for cooling operations. Many facilities rely on evaporative cooling technologies in which water absorbs heat generated by electronic equipment. This process allows thermal energy to be dissipated efficiently but also results in significant water consumption.

Research published in **the journal** Nature Climate Change indicates that the training of large artificial intelligence models can indirectly require hundreds of thousands of liters of freshwater depending on the cooling technology and climatic conditions of the data centre location [4]. In regions experiencing water scarcity, such demand may contribute to environmental stress on local water resources. The environmental consequences of artificial intelligence are also related to the production and disposal of specialised computing hardware. Machine learning systems rely heavily on high-performance processors designed specifically for parallel computation. Technological progress in semiconductor engineering occurs rapidly, and new generations of processors frequently provide significantly higher performance than previous versions. As a result, computing equipment becomes technologically obsolete relatively quickly. Large technology companies periodically replace outdated servers with more efficient hardware in order to maintain competitive performance levels. This process contributes to the global problem of electronic waste. According to the Global E-Waste Monitor produced by the United Nations University, more than 53 million tonnes of electronic waste were generated worldwide in 2019, and the volume of such waste continues to grow annually [5]. Although artificial intelligence infrastructure represents only a portion of global electronic waste, the rapid expansion of computing capacity increases the rate at which electronic components are discarded. Despite these environmental challenges, researchers and technology companies have proposed several strategies aimed at reducing the ecological footprint of artificial intelligence systems, including algorithm optimisation and the use of renewable energy in data centres. One important direction involves improving the efficiency of machine learning algorithms. Techniques such as model pruning, parameter sharing, and knowledge distillation allow neural networks to be simplified while preserving their functional capabilities. By reducing the number of parameters and computational operations required for training, such methods can significantly decrease electricity consumption. Experimental research demonstrates that optimised machine learning models can require several times less computational power while maintaining similar levels of accuracy [6]. In addition to algorithmic improvements, researchers also investigate methods for developing specialised hardware architectures designed specifically for energy-efficient artificial intelligence computation. The development of more efficient processors may reduce the environmental impact of future artificial intelligence systems. Another important strategy involves the use of renewable energy sources in data centre operations. Several major technology companies have invested heavily in renewable electricity generation in order to reduce the carbon footprint of their computing infrastructure. Google reported that its global data centre operations have matched one hundred percent of electricity consumption with renewable energy purchases since 2017 [7]. Similar initiatives have been announced by other companies operating large cloud computing platforms. Renewable energy sources such as wind and solar power can significantly reduce greenhouse gas emissions associated with artificial intelligence workloads. However, the integration of renewable energy into computing infrastructure also presents technical challenges. Renewable energy production is often intermittent and depends on weather conditions. Consequently, reliable energy storage systems and flexible grid management strategies are required in order to ensure stable electricity supply for data centres. Another important area of research focuses on improving the architectural design of data centres. Modern facilities increasingly employ advanced cooling technologies such as liquid cooling systems, immersion cooling methods, and heat recovery mechanisms. These technologies allow thermal energy generated by computing equipment to be dissipated more efficiently than traditional air cooling solutions.

Artificial intelligence technologies themselves can also contribute to improving data centre efficiency. Machine learning algorithms can analyse operational data from computing facilities and automatically adjust parameters such as server workload distribution, cooling system intensity, and power management settings. Such optimisation methods allow energy consumption to be reduced without compromising computational performance. The efficiency of data centres is commonly evaluated using the Power Usage Effectiveness metric, which represents the ratio between total facility energy consumption and the amount of electricity used directly by computing equipment. The structure of this metric and the distribution of energy consumption in data centre infrastructure are illustrated in Figure 2.

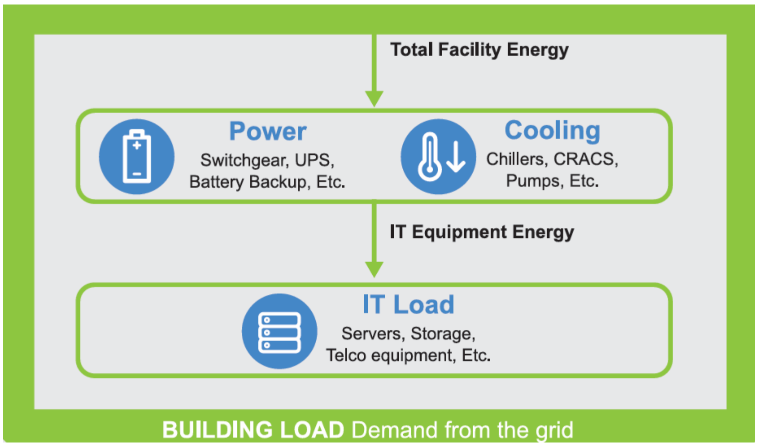

Figure 2 – Power Usage Effectiveness (PUE)

Early data centres constructed in the 2000s often had PUE values close to 2.0, indicating that a large portion of consumed electricity was required for cooling and other supporting infrastructure. Modern facilities have achieved significantly lower PUE values, often approaching 1.5 or even lower in highly optimised environments [1]. Continued improvements in infrastructure design may therefore reduce the environmental impact of large-scale computing. Government policies and regulatory initiatives may also influence the sustainability of artificial intelligence technologies. Several countries have introduced energy efficiency standards for digital infrastructure and have encouraged the adoption of renewable energy in data centre construction. Researchers also suggest the introduction of transparency requirements that would obligate organisations to report the computational resources and carbon emissions associated with training artificial intelligence models [2]. Such transparency may encourage developers to design more efficient algorithms and discourage unnecessary large-scale computational experiments.

The environmental cost of artificial intelligence and large-scale computing represents a complex and multidimensional challenge. Artificial intelligence technologies rely on extensive computing infrastructure that requires substantial electricity consumption, sophisticated cooling systems, large quantities of water, and specialised electronic hardware. Empirical studies demonstrate that the training of large machine learning models can generate significant greenhouse gas emissions and consume considerable natural resources. At the same time, technological progress provides several opportunities for reducing the environmental impact of artificial intelligence. Improvements in algorithmic efficiency, the development of energy-efficient hardware, the integration of renewable energy sources, and the optimisation of data centre architecture may significantly decrease the ecological footprint of large-scale computing systems. Furthermore, artificial intelligence itself can be applied to improve energy efficiency and support environmental monitoring. It can therefore be stated that the environmental sustainability of artificial intelligence will depend on the balance between technological growth and responsible resource management. Continued interdisciplinary research involving computer scientists, environmental researchers, and energy engineers will be essential for ensuring that artificial intelligence development remains compatible with the principles of sustainable technological progress.

References:

1. *International Energy Agency.* Data Centres and Data Transmission Networks – Analysis. – *Paris : International Energy Agency, 2023. – URL: https://www.iea.org/energy-system/digitalisation/data-centres-and-data-transmission-networks (date of access: 12.03.2026).*

2. *Strubell, E. Energy and policy considerations for deep learning in NLP / E. Strubell, A. Ganesh, A. McCallum //* Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. – *Florence : ACL, 2019. – P. 3645–3650. – URL: https://aclanthology.org/P19-1355/ (date of access: 12.03.2026).*

3. *Shehabi, A.* United States Data Centre Energy Usage Report */ A. Shehabi, S. Smith, D. Sartor et al. – Berkeley : Lawrence Berkeley National Laboratory, 2016. – URL: https://eta-publications.lbl.gov/sites/default/files/lbnl-1005775_v2.pdf (date of access: 12.03.2026).*

4. *Li, P. The carbon and water footprint of large AI models / P. Li, J. Yang, M. Islam //* Nature Climate Change. *– 2022. – Vol. 12. – URL: https://www.nature.com/articles/s41558-022-01385-5 (date of access: 12.03.2026).*

5. *United Nations University.* The Global E-Waste Monitor. – *Bonn : United Nations University, 2020. – URL: https://ewastemonitor.info/ (date of access: 12.03.2026).*

6. *Patterson, D. Carbon emissions and large neural network training / D. Patterson, J. Gonzalez, Q. Le et al. –* arXiv preprint, *2021. – URL: https://arxiv.org/abs/2104.10350 (date of access: 12.03.2026).*

7. *Google.* Environmental Report and Carbon-Free Energy Progress. – *Mountain View : Google LLC, 2022. – URL: https://www.google.com/environment/ (date of access: 12.03.2026).*

8. Growth in data centre energy consumption *[diagram]. – Paris : International Energy Agency, 2023. – 1 fig. – URL: https://www.iea.org/energy-system/digitalisation/data-centres-and-data-transmission-networks (date of access: 12.03.2026).*

9. Power Usage Effectiveness (PUE) *[infographic]. – Portland : The Green Grid, 2017. – 1 fig. – URL: https://www.thegreengrid.org/system/files/store/TGGInfographicPUEvF.pdf (date of access: 12.03.2026).*