# The cloud is not in the sky: why the energy consumption cost of AI models is not equal and what this means for sustainability.

There has rarely been a type of technology in history as fast paced as Artificial Intelligence (AI), which now permeates every aspect of our daily lives. According to the UK Government's 2023 paper on 'A pro-innovation approach to AI regulation', AI is defined as "products and services that are 'adaptable' and 'autonomous'." However, the term AI is too broad to accurately describe the vast differences in technologies that it represents and is often conflated with machine learning, data processing and data analytics. Examples of uses of AI include:

- Expert systems – a piece of software which uses databases of knowledge to offer advice or make decisions.
- Natural language processing (NLP) – the application of computational techniques to the analysis and synthesis of natural language and speech.
- Speech recognition – the process of allowing software to identify and respond to the sounds produced in human speech.
- Machine vision – the ability of a computer to see the environment.

The use of AI causes an environmental impact. This is due to the data centres and network infrastructures that are needed to operate the AI models consuming large amounts of energy and water. Currently AI's energy use only represents around 2-3% of total global emissions. (Source: World Economic Forum, AI and Energy) However, this is likely to increase as adoption of AI increases to drive efficiency and productivity.

**Recommendations:**

Regulations should now look beyond the immediate risks of AI development to the much broader impact it has on the environment.

- The IET recommends a bronze, silver, gold standard for the approval of new data centres in the UK, based on a sustainability rating. The UK must become not only a global leader in innovation and technology, but a leader in sustainable innovation.

- Government should support green technologies by building sustainability into appropriate regulation of new and expanding technologies.

- Computing vs Storage - recognising that current storage models (data centres) are unsustainable without greener infrastructure.

## AI and sustainability

There is a unique juxtaposition between AI and sustainability, as AI is high consumer of energy, but also possesses huge potential to tackle climate change. AI can transform the energy efficiency of other carbon-intensive industries, such as modelling buildings to predict energy use (Source: European Cluster Collaboration Platform, Artificial Intelligence improves the energy efficiency of buildings).

Reports predict that the use of AI could help to mitigate 5-10% of global greenhouse gases (GHG) by 2030 (Source: Boston Consulting Group, How AI Can Speed Climate Action). This is where greater scrutiny is needed. The difference in energy consumption between a Large Language Model and a simple AI app is significant. AI should also not be used interchangeably to mean data centres, which are themselves, intensive and costly.

Cost cannot only be quantified in terms of financial return as the cost to the environment and the impact that climate change has on health must also be considered. One of the impacts we need to consider is the capacity of the UK grid. Achieving net-zero will require a comprehensive understanding of the most carbon efficient AI models, and the offsetting powers of technology. There needs to be serious thought on the placement of data centres where climate impact and resilience can be minimised with local access to green energy generation.

The IET proposes a ‘bronze, silver, gold’ standard for data centres: bronze for excessive energy consumption which is harmful to the environment, silver for less harmful, and gold for environmentally sustainable operations. This would emphasise the moral responsibility surrounding data centre energy consumption. Government should subsequently encourage the removal of "bronze" data centres (those not using green energy) in the UK. Similar to the country’s shift away from coal energy, this could support the push for sustainable technological progress. It would also be appropriate to tangibly recognise instances where developers have acted sustainably by adapting an existing product, for example an LLM, rather than retraining or starting from the beginning, which has led to a lower carbon footprint.

A clean, modern and decarbonised grid will be vital to move towards net zero. Companies like Microsoft and Google are already exploring alternative power options such as nuclear technologies to power Large Language Models (LLMs). Government should invest in other clean energy technologies such as offshore wind, which has a proven track record in the UK. The drive to decarbonise our energy system by 2050 offers the nation an opportunity to move from fuel-based solutions and become more energy independent.

**Technology trends**

The growing financial cost of Large Language Models (LLMs) is unsustainable, this is in part due to the extensive cost of the energy required to run them, making LLMs unprofitable. The UK must be cautious about investment in technologies that may or may not have reached their peak. There should be consideration surrounding whether the UK should invest in current technologies that are excelling so the tech’s security and durability is enhanced, invest in next generation technologies in an effort the lead in the field, or target and lead in niche topics such as: smaller machine learning or narrower AI approaches.

There are also ethical considerations of pursuing technology that is not net-productive. By focusing on carbon efficient AI and technologies the UK has the potential to become a leader in the global AI economy. The Government should disaggregate AI and data centres. It should be clarified that the costs of running hardware in data centres are significantly higher in both energy consumption and financial terms than solely storing data for AI processes. To limit harm, the environmental impact of data centres needs greater recognition across industry without being overlooked in favour of cost efficiency. There also needs to be a consideration on what type of AI to use, whether it is LLMs, traditional compute optimisation or search techniques as there is a considerable cost of training and re-training models to achieve this.