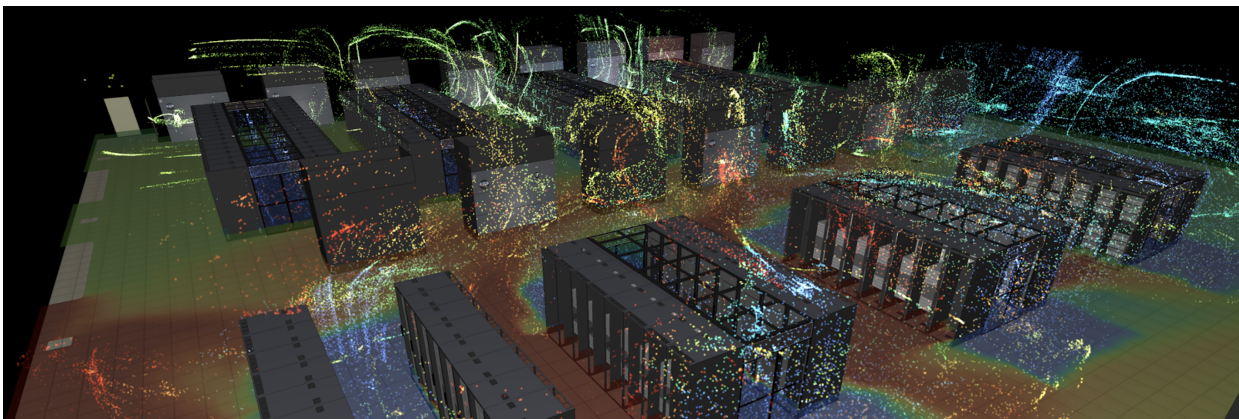# Analysis | Startups are working overtime to reduce the massive climate impact of AI

By **Tim Smith**, News editor
April 25, 2024

A datacentre optimisation simulation from Cadence

Depending on who you listen to, super-intelligent computer systems are either going to destroy humanity, or deliver a future of prosperity and abundance; unlock a new era of human artistry, or gut out entire creative industries.

The truth is probably somewhere in the messy middle, but one thing that pretty much everyone can agree on is that AI is using *a lot* of energy, with large generative models needing huge amounts of computing power (known as "compute" in the industry) to crunch vast quantities of data.

Last month a report by a group of climate change organisations estimated that AI could lead to an 80% increase in global carbon emissions, and experts are warning that the AI industry's huge power demands could threaten global energy supplies.
Luckily, a handful of plucky startups are working on ways to reduce the climate impact of the AI sector, making the technology more energy-efficient through hardware and software innovations, as well as strategic decisions on accessing green power.

The mountain shall go to Mohammed
OpenAI CEO Sam Altman is advocating for new nuclear energy capacity to quench AI's power thirst — but other operators are already finding smart ways to meet energy demands without increasing AI's carbon emissions.

London-based startup NexGen Cloud, which was founded in 2020, builds data centres — facilities that house the specialist chips for AI development — that are entirely powered by renewable energy.

They're not built in Britain, however. "We had to go further afield. We were like, 'Where can we go? We're not going to be able to do this in the UK…' We saw there was an abundance of power in places like Norway, Sweden and Iceland," says the company's CEO Chris Starkey.

"In some locations in Norway and Sweden they have really big [green energy] infrastructure but the towns are really small. So they were only using a very, very small amount of that power." NexGen Cloud's data centres are exclusively powered by hydroelectric power, meaning that they get a reliable supply of affordable renewable energy — and Starkey says the decision to go green has paid off for his business.

"Now I'd say one out of three conversations that we're having with customers is all about renewable and sustainability," he tells Sifted.

Efficiency savings
While NexGen Cloud has been able to scale up its own compute capacity with green power, Starkey emphasises that the surging demand for AI across the economy is likely to cause "big problems" with energy supply and demand in the coming years.

That means that it's not enough for data centre operators to hitch their wagon to green power providers — they also have to work overtime to make their technology more energy-efficient, particularly when it comes to the cooling systems they use to keep their chips running.
One entrepreneur working on that problem is Mark Seymour, founder of London-based Future Facilities Ltd, which was acquired by US multinational Cadence Design Systems in 2022. He tells Sifted that he's now using AI to help organisations building data centres design more energy-efficient systems, via something called "digital twin" technology.

Its product, which is being used by large cloud compute providers, lets people building data centres simulate the design and setup of their systems to make sure they can be cooled with the least amount of energy possible.

"Every time you change a server in the data centre, you're changing the component layout and therefore the thermal performance," he says, adding that — without simulating the design of those systems first — data centres often only run at 60-70% energy efficiency.
"With simulation, you can work out how to use that capacity and get much closer to running 100% efficiently."

Squeezing more juice
It's not only the design of data centres that can be made more energy-efficient. AI chips also leave a lot to be desired when it comes to smart use of power consumption.

This is partly because the processors that are used for AI training — graphics processing units (GPUs) — were originally designed for a different use-case: real-time rendering of computer-generated graphics in video games.

Startups like Oxford, UK-based Lumai are now developing new chips using photonics technology — which uses light rather than silicon to transmit computation information — to build processors specifically for AI development, which use less energy than a GPU.

"What we're creating is much more advanced computer speed using much lower energy than what GPUs can do at the moment," founder Andrew Weil told Sifted last year, when the company received a £1.1m grant from the UK government.

There are also gains to be made in the way AI chips hold data memory, according to Flavia Levi, early stage deeptech investor at Octopus Ventures, particularly when it comes to "inference" — the term for the computing power needed to run AI models after they've been developed.

"Traditional computing systems are not optimised for inference. They hold the memory and processor on separate chips, which means the processor must make calls to the memory when it requires anything stored there, which is time and energy consuming," she says, adding that a

number of startups including London-based Fractile are developing solutions to put the processor and memory on a single chip to reduce energy consumption.

Startups are working on software to complement these hardware improvements too. This week, Paris-based startup FlexAI came out of stealth mode, [announcing](#) a $30m seed round. It plans to build an operating system that helps AI developers design their systems in ways that get more performance with less energy usage.

Cofounder Brijesh Tripathi tells Sifted that, currently, AI chips waste about half of the energy that they consume.

He estimates that software optimisations can get GPUs running closer to 80% efficiency.

Stay cool
AI isn't entirely the bad guy, however. The technology [could *also* help](#) in the fight against climate change.

Last week Brussels-based think tank The Lisbon Council released a [report](#) supported by chip giant Nvidia, assessing how AI can contribute to a greener future. The research emphasised the need for greener technologies to build powerful systems, but also highlighted that AI could make other areas of the economy far more energy-efficient.

"AI is fundamental to integrating green energy into the grid and discovering new [more sustainable] materials," Lisbon Council director of research David Osimo tells Sifted.

"I don't want to make it hype-y — it's not that AI will solve all problems — but you just have to look at the holistic impact and have a balanced approach."

https://sifted.eu/articles/startup-ai-climate-impact-energy