UC SANTA BARBARA



April 15, 2025 Michaela Galarza

To power AI, data centers need more and more energy

In the race to dominate artificial intelligence, technology companies are fueling an unprecedented expansion of data centers. These digital infrastructures power everything from search engines to cloud computing and, increasingly, artificial intelligence (AI) applications. But as AI's demand for computational power skyrockets, so too does its energy footprint.

UC Santa Barbara <u>Professor Eric Masanet</u>, a leading expert in industrial sustainability and the Mellichamp Chair in Sustainability Science for Emerging Technologies at UCSB's Bren School of Environmental Science & Management, has spent nearly two decades studying the evolving <u>energy dynamics of data centers</u>. In a recent conversation, he provided key insights into the challenges and solutions surrounding data center sustainability in the age of AI.

Masanet's journey into the data center space was unexpected. With a background in engineering and a passion for sustainability, he initially worked in product design before pivoting to research. His expertise in industrial energy efficiency led him to co-lead a Congressional study on data center energy use in 2006, which sparked his interest in the field.

"Energy became a critical part of my research because decarbonizing manufacturing was my main goal," he said. "You can't improve industrial sustainability without considering energy sources and efficiency."

The three eras of data center growth

According to Masanet, the evolution of data centers can be divided into three distinct eras: early growth, the cloud computing boom and the AI acceleration that defines our current landscape. The early growth phase, spanning the 1990s and 2010, was driven by the rapid expansion of internet infrastructure, including the proliferation of computers, internet services and smartphones. This period was marked by a lack of concern for energy efficiency.

The second phase, from 2010 to 2018, saw the rise of cloud computing, which led to the consolidation of data centers into large-scale, centralized facilities operated by major technology companies. This shift significantly improved efficiency by replacing smaller, inefficient centers with hyperscale data centers — massive facilities that house thousands of servers and are designed to process large amounts of data with optimized energy use.

Today's AI acceleration era is characterized by an explosion in energy demand. Aldriven applications require specialized, power-intensive hardware, and the rapid deployment of AI technologies is pushing data center energy consumption to unprecedented levels.

Why AI is driving an energy surge

Al workloads require specialized graphics processing units (GPUs) that consume significantly more electricity than conventional servers. Al model training, particularly for large-scale models like ChatGPT, involves processing vast amounts of data and billions of parameters, making it an energy-intensive process. Even after training, running the models to generate responses, known as the inference stage, requires extensive amounts of energy.

"Al servers use up to 10 times the power of a standard server, and companies are deploying them at an unprecedented scale," Masanet noted. "The combination of high power needs and rapid expansion is what's straining the grid."

The sustainability challenge

The increasing electricity demand from AI data centers poses a significant challenge to sustainability goals. Masanet points to Northern Virginia as a cautionary tale, where the region's concentration of data centers has forced utilities to keep fossil fuel plants online to meet demand. While major tech companies pledge to power data centers with renewable energy, the reality is that the expansion is outpacing the deployment of clean energy sources.

"Some companies are looking at nuclear power or geothermal solutions, but these are not yet widely available," he said. "Renewable energy simply isn't scaling fast enough to match Al's growth."

Potential solutions for Al's energy footprint

Despite these challenges, there are several promising strategies to mitigate Al's energy consumption:

- Geographic optimization: Placing AI data centers in regions with abundant renewable energy, such as Iceland or the Pacific Northwest, could reduce reliance on fossil fuels.
- Workload shifting: Cloud providers can <u>shift AI tasks to locations</u> where renewable energy is most available at any given time.
- Algorithmic efficiency: Al models can be trained using fewer data points and optimized software to reduce energy consumption.
- Cooling innovations: Advanced liquid cooling systems can improve efficiency by reducing the energy required to cool AI servers.

However, these solutions come with limitations. "Even with efficiency gains, Al's energy footprint is still expected to grow," Masanet warned. "The key is ensuring that this growth aligns with sustainable energy deployment rather than exacerbating fossil fuel dependence."

Looking ahead: Will AI energy demand level off?

While AI energy demand is surging now, Masanet believes it won't continue indefinitely. "Some business models will collapse, and AI data centers won't always run at full capacity," he said. "Efficiency improvements will happen, but for now, rapid expansion is the reality."

The challenge remains: Can the tech industry align Al's rapid rise with sustainability goals? As data centers become the backbone of the digital economy, their energy consumption will be a critical issue in the fight against climate change.

Masanet's work highlights the urgent need for innovation, transparency and policy intervention to ensure that AI's benefits outweigh its environmental costs. The future of sustainable computing depends on it.

Tags Artificial Intelligence

Media Contact **Debra Herrick** Associate Editorial Director (805) 893-2191 <u>debraherrick@ucsb.edu</u>

About UC Santa Barbara

The University of California, Santa Barbara is a leading research institution that also provides a comprehensive liberal arts learning experience. Our academic community of faculty, students, and staff is characterized by a culture of interdisciplinary collaboration that is responsive to the needs of our multicultural and global society. All of this takes place within a living and learning environment like no other, as we draw inspiration from the beauty and resources of our extraordinary location at the edge of the Pacific Ocean.