TECHNOLOGY

AI IS TAKING WATER FROM THE DESERT

New data centers are springing up every week. Can the Earth sustain them?

By Karen Hao

Illustration by Erik Carter

MARCH 1, 2024

SHARE AS GIFT SAVE

One scorching day this past September, I made the dangerous decision to try to circumnavigate some data centers. The ones I chose sit between a regional airport and some farm fields in Goodyear, Arizona, half an hour's drive west of downtown Phoenix. When my Uber pulled up beside the unmarked buildings, the temperature was 97 degrees Fahrenheit. The air crackled with a latent energy, and some kind of pulsating sound was emanating from the electric wires above my head, or maybe from the buildings themselves. With no shelter from the blinding sunlight, I began to lose my sense of what was real.

Microsoft announced its plans for this location, and two others not so far away, back in 2019—a week after the company revealed its initial $1 billion investment in OpenAI, the buzzy start-up that would later release ChatGPT. From that time on, OpenAI began to train its models exclusively on Microsoft's servers; any query for an OpenAI product would flow through Microsoft's cloud-computing network, Azure. In part to meet that demand, Microsoft has been adding data centers at a stupendous rate, spending more than $10 billion on cloud-computing capacity in every *quarter* of late. One semiconductor analyst called this "the largest infrastructure buildout that humanity has ever seen."

I'd traveled out to Arizona to see it for myself. The Goodyear site stretched along the road farther than my eyes could see. A black fence and tufts of desert plants lined its perimeter. I began to walk its length, clutching my phone and two bottles of water. According to city documents, Microsoft bought 279 acres for this location. For now, the plot holds two finished buildings, thick and squat, with vents and pipes visible along their sides. A third building is under construction, and seven more are on the way. Each will be decked out with rows of servers and computers that must be kept below a certain temperature. The complex has been designated partly for OpenAI's use, according to a person familiar with the plan. (Both Microsoft and OpenAI declined to comment on this assertion.) And Microsoft plans to absorb its excess heat with a steady flow of air and, as needed, evaporated drinking water. Use of the latter is projected to reach more than 50 million gallons every year.

That might be a burden in the best of times. As of 2023, it seemed absurd. Phoenix had just endured its hottest summer ever, with 55 days of temperatures above 110 degrees. The weather strained electrical grids and compounded the effects of the worst drought the region has faced in more than a millennium. The Colorado River, which provides drinking water and hydropower throughout the region, has been dwindling. Farmers have already had to fallow fields, and a community on the eastern outskirts of Phoenix went without tap water for most of the year.

After I'd walked for 20 minutes in the sun, my own water supply was nearly out. I'd managed to traverse only two sides of the data center, but it was time to call it quits. My face and neck were red, and my eyes had sprouted flecks of pink. If I wanted to continue my tour the next day, there were dozens of other facilities I could visit in the area, including those run by Apple, Amazon, Meta, and, soon, Google. Not too far from California, and with plenty of cheap land, Greater Phoenix is among the fastest-growing hubs in the U.S. for data centers.

The American Southwest has become the site of a collision between two civilization-defining trends. In this desert heat, the explosive growth of generative AI is pitched against a changing climate's treacherous extremes. The situation is already bad enough to worry residents in Goodyear, as several told me on my visit. And it's only going to get worse.

Microsoft, the biggest tech firm on the planet, has made ambitious plans to tackle climate change. In 2020, it pledged to be carbon-negative (removing more carbon than it emits each year) and water-positive (replenishing more clean water than it consumes) by the end of the decade. But the company also made an all-encompassing commitment to OpenAI, the most important maker of large-scale AI models. In so doing, it helped kick off a global race to build and deploy one of the world's most resource-intensive digital technologies.

Microsoft operates more than 300 data centers around the world, and in 2021 declared itself "on pace to build between 50 and 100 new datacenters each year for the foreseeable future." The dual task of laying down those new facilities and making them sustainable has fallen to Noelle Walsh, the head of Microsoft's data-center

division. Walsh told me the company is moving aggressively to reach its end-of-the-decade environmental benchmarks. It's partnering with power providers to bring more solar and wind energy online. It's investing in fusion companies and researching new battery designs. It's working across water-stressed regions, including Arizona, to put water back into local lakes and rivers. "Even with our surge in demand and surge in AI," she said, "our 2025 goals have stayed the same, and so have our 2030 goals."

But all of this work takes time—including years of consultation with local authorities, power providers, and communities, as well as risky, long-term bets on basic research. In comparison, the AI explosion happened overnight. On an earnings call last summer, Microsoft told investors that the Azure OpenAI Service was bringing in new customers at a rate of nearly 100 a day. Executives have gushed about the AI boom and suggested that the technology might even help *save* the climate. Internally, they're boasting about the wonders it has done for Microsoft's business. "Every now and then it's great to take a step back and marvel at just how far we've come in just one year," Eric Boyd, the corporate vice president of the AI platform, wrote in an August 2023 email to his division. In another, from last April, he wrote that the company would be "infusing AI" into all of its cloud solutions. Coca-Cola, Office Depot, Shell, and the Department of Defense were already among the customers with access to Azure's OpenAI features, including ChatGPT and GPT-4, he said across several emails last year.

Read: The internet's next great power suck

Public data hint at the potential toll of this approach. Researchers at UC Riverside estimated last year, for example, that global AI demand could cause data centers to suck up 1.1 trillion to 1.7 trillion gallons of fresh water by 2027. A separate study from a university in the Netherlands, this one peer-reviewed, found that AI servers' electricity demand could grow, over the same period, to be on the order of 100 terawatt hours per year, about as much as the entire annual consumption of Argentina or Sweden. Microsoft's own environmental reports show that, during the initial uptick in the AI platform's growth, the company's resource consumption was accelerating. In fiscal year 2022, the most recent year for which Microsoft has released data, the tech

giant's use of water and electricity grew by about a third; in absolute terms, it was the company's largest-ever reported increase, year to year.

A Microsoft spokesperson said the company will continue to invest in ways to mitigate the climate crisis, including those based on AI. He also noted that even as the use of data centers had exploded in recent years, their electricity consumption has grown far more slowly because of innovations in efficiency.

Microsoft has sought to "be a good neighbor," Walsh said. In Arizona, the company is funding projects for conserving and restoring water. But it acknowledged that most of these upstream efforts support Goodyear residents only indirectly. Many of the company's other efforts to address its environmental impacts also suffer from limitations: Carbon offsets and clean-energy power-purchase agreements may help Microsoft achieve carbon-negative and water-positive operations on paper, but they don't necessarily net out the effects on local communities, or anyone else for that matter. The power-purchase agreements, for example, give utility providers money up front to build more renewable-energy or carbon-free-energy capacity, but not necessarily on the grids that Microsoft uses. That means the company's data centers may still be running on fossil fuels and generating emissions, while clean energy is being underutilized somewhere else. "*Purchasing* clean energy is not the same as *physically consuming* clean energy," Microsoft wrote in its own 2023 white paper about decarbonizing the cloud. The Microsoft spokesperson said that the company's data centers hook into "large, interconnected electricity grids" that draw from thousands of power plants, including renewable ones supported by Microsoft's purchase agreements, but that it isn't possible to specify the exact source of the electricity consumed.

Read: Why Amazon's data centers are hidden in spy country

Walsh said her team seeks to go above and beyond what global sustainability standards require. She wants to make sure that, by the end of the decade, any renewable or carbon-free energy that Microsoft purchases is connected directly to the power grids that it runs on. The company is transitioning some data centers,

including those in Arizona, to designs that use less or no water, cooling themselves instead with giant fans. Her team is also working with steel and concrete suppliers to make the building materials for each data center more sustainable, and she'd like to see server components recycled whenever possible.

In the meantime, though, Microsoft has been reluctant to provide customers with specific details on the environmental impacts of their cloud-service needs. A series of internal memos, written from 2020 to 2022 by a group of employees from across the company, proposed giving Azure customers new tools to measure the carbon footprint of their cloud use, for AI and otherwise, according to four former employees who saw these documents, and who requested anonymity to safeguard their future job prospects. But these ideas were sidelined, they said. The group proposed, for example, to publish the energy and water efficiency of every Microsoft data center, so customers could make more informed choices about which ones to use, per two of the former employees who were there when some of the memos were presented to leadership. Microsoft executives expressed concern, they said, that some of the company's data centers would end up showing worse performance than Google's. (Google has been posting facility-specific energy-efficiency numbers for years.)

The Microsoft spokesperson said the company is constantly improving the sustainability of its data centers and committed to providing transparency to its customers. Indeed, Microsoft has released some tools to help customers estimate their carbon emissions based on regional averages. In late 2022, Walsh and her team put out average water- and energy-efficiency numbers for data centers in 28 regions across the globe, but not for specific facilities. (The Microsoft spokesperson said the information is provided at a regional and not facility level because customers can choose only among different regions.) A former employee told me that there was much more the company could do. "Microsoft is being lazy," he said.

On the fast track to dehydration, I called another Uber to bring me from the data center back to my hotel. Even that simple act, it occurred to me, required a data center not unlike the ones right next to me. A facility like this would have to connect the ping from my cellphone to Uber's driver network in order to find me a ride home.

Exactly how much power does this Goodyear data center use, and how much of it is renewable? Neither Microsoft nor the local utility company would say. As for water use, a records request to the city returned documents with all of the numbers redacted; a representative for the city said the numbers were "considered proprietary by Microsoft." But a report that Microsoft filed with the city council provides at least an estimate. An analysis commissioned by the company anticipates that when construction for the third building is complete, the complex will use about 56 million gallons of drinking water each year, equivalent to the amount used by 670 Goodyear families.

In other words, a campus of servers pumping out ChatGPT replies from the Arizona desert is not about to make anyone go thirsty. Barbara Chappell, Goodyear's water-services director, told me the city's relationship with Microsoft on the whole hasn't given her cause for concern. But the supply of water in the region *is* quite limited, and the more that's taken up by data centers, the less there is for, say, supplying tap water to new housing. "We're going to have to make tough choices in the near future to make sure our state is protected for future generations," Arizona Attorney General Kris Mayes told me. "Allowing one more data center to come to our state is an easy but stupid decision in a lot of cases. It's like the cotton candy of economic development."

For Chandler, a city some 40 miles away from Goodyear, in the wealthier East Valley area of Greater Phoenix, the benefits of more data-center investments weren't worth the water and energy costs, Micah Miranda, the city's economic-development director, told me. Chandler already hosts seven data-center complexes, the city said; in 2015 and 2022, the city passed two ordinances that effectively limited how many more could come in.

Read: The tech industry is producing a rising din

Goodyear sits in the less economically developed West Valley, and had more incentive to cut a deal. At a reveal event in 2019, Brian Janous, then Microsoft's general manager of energy and renewables, described the project as win-win. As of 2021, the company had provided more than $1 million for advancing STEM education, training workers, and revitalizing the environment, such as planting trees and cleaning up rivers, according to a press release. Microsoft was also well aware of Phoenix's environmental concerns, Janous said at the event, and fully intended for its facilities in Goodyear and just next door in El Mirage to be "among the most sustainably designed data centers in the world." Whenever the temperature outside was below 85 degrees, they would be cooled without water, he promised.

But temperatures in Goodyear clear 85 degrees on most days of the year. Furthermore, at the time the deal was struck, Microsoft had yet to face the unprecedented AI surge. City documents have since logged four amendments to Microsoft's data-center construction. "It's always changed. It's changed numerous times," Chappell said, later adding, "We can't keep up."

AI didn't create these problems, but it's certainly making them worse. According to Walsh, AI applications are among the most computationally intensive that Azure supports, which creates a need for more data centers overall while also upping the energy and cooling demands at each one. On top of that, *generative*-AI applications in particular can be orders of magnitude more energy-intensive than the predictive-AI applications that came before. "We still don't appreciate the energy needs of this technology," OpenAI CEO Sam Altman said in January at Davos. "There's no way to get there without a breakthrough." In the meantime, tensions over data centers' water use are cropping up not just in Arizona but also in Oregon, Uruguay, and England, among other places in the world.

As I wrapped up my trip to Arizona, preparations at Microsoft were in full swing to promote its own, very different narrative at the world's most important climate event of the year. In the run-up to the United Nations' Conference of Parties, held this past

fall in Dubai, Big Tech firms were proudly claiming that AI is not so much a source of new emissions as it is a way to lower them, by making batteries and buildings more efficient, for example, and reducing food waste. (This is mainly true of predictive-AI models, Sasha Luccioni, a researcher and climate lead at the AI firm Hugging Face, told me. The generative ones have yet to demonstrate significant environmental benefits that would justify their mounting footprint.) At the same time, Microsoft announced a new partnership with the UN that would use AI to help track global carbon emissions. "Simply put, you can't fix what you can't measure," the company's president, Brad Smith, said in a statement released by the company at the time.

In the background, though, Microsoft was withholding important measures and projections. Before the climate conference, the company sent around publicity materials for internal review. Some employees responded that they were concerned, according to screenshots of internal messages. The documents minimized the mounting energy costs of AI, they told the company's leaders; perhaps Microsoft should instead disclose internal estimates of how those costs might rise in years to come.

Walsh confirmed to me that her team did indeed keep those internal estimates—one for commercial-cloud growth, and one for AI. I asked her what they said. "I can't tell you that," she replied. By corporate standards, it was an ordinary response. Yet the circumstances have never been more extraordinary. When I asked the company how an accelerating use of natural resources might affect Microsoft's plan to erase its carbon footprint altogether by the end of the decade, the spokesman answered in broad terms: "We remain optimistic regarding our collective ability to decarbonize the global economy while continuing to grow and prosper as a global community." With more than 8,000 data centers whirring all around the world and venting heat, and many more on their way, that optimism may come off as nothing more than faith: Technology has gotten us into this predicament; perhaps technology will get us out of it.

ABOUT THE AUTHOR

Karen Hao

Follow

Karen Hao is a writer based in Hong Kong. She was formerly a foreign correspondent covering China's technology industry for the *Wall Street Journal* and a senior editor at *MIT Technology Review*.