

The Data Motherlode: Why the Modern Gold Rush Is Happening in India

Wed, 02/25/2026 - 09:00

|
Blaine Fisher, Ph.D.

[View PDF](#)



In the summer of 1956, a handful of researchers gathered at Dartmouth College with a wild idea. They believed intelligence could be described precisely enough that a machine could simulate it. They coined the term "artificial intelligence," wrote

up a proposal, and spent a summer arguing about whether they could actually pull it off.

They could not, mostly. But the idea survived every winter that followed.

Seventy years later, the question has shifted entirely. It is no longer whether the conjecture is plausible. The question now is who will industrialize it at scale, control the supply chains, and train the operators who run the whole thing. That is a different kind of problem. It is less a scientific question and more a geographic and economic one, the kind that gets answered not in laboratories but in land deals, fiber contracts, and policy rooms.

Which is why, if you want to understand where AI is actually going, you have to look at what happened in New Delhi in February 2026.

The Rush Begins

Every gold rush needs a geography. In 1848, it was the Sierra Nevada foothills, specifically a stretch of the American River where James Marshall spotted something glinting in the water at Sutter's Mill. Within two years, California's non-native population had grown by a factor of ten. People did not come because they loved California. They came because that is where the gold was.

The India AI Impact Summit 2026, held from February 16 to 20 at Bharat Mandapam and Sushma Swaraj Bhawan in New Delhi, drew attendance in the hundreds of thousands and announced investment commitments on a scale normally reserved for power grids and ports. More than \$250 billion in commitments tied to AI infrastructure, data centers, and semiconductors were announced across the event's five days. The frontier labs, the chipmakers, the cloud providers, and the enterprise software companies all showed up. They did

not come because they love New Delhi in February, though February is actually a fine time to visit. They came because that is where the motherlode is.

The motherlode, in this rush, is data.

Why Data Is the New Gold

The Dartmouth proposal, drafted by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, included the claim that "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it." That sentence launched a discipline, and seventy years of argument, disappointment, and eventually genuine progress followed from it.

What the 1956 crew could not have anticipated is what intelligence at machine scale actually requires as raw material. It requires data, enormous quantities of it, diverse enough to capture the full range of human expression, behavior, error, and intent. English-language web data, the first great seam that labs mined, is already heavily worked. The easy surface deposits are gone. What remains requires more sophisticated extraction, and the richest untouched veins run through languages, regions, and interaction patterns that Western labs have historically underserved.

India sits on top of one of the largest and most linguistically diverse data deposits on the planet. The country has over a billion active internet users. It has hundreds of living languages and thousands of dialects. It has an enormous and growing population of voice-first users, people who interact with technology primarily through speech rather than text, because that is what works in their language and context. It has behavioral data flowing from agricultural apps, medical consultations, financial services reaching first-time banking customers, and educational platforms serving students who have never owned a laptop.

None of that data looks like the English-language web. Which means it fills gaps that current models handle poorly. Which means it is, in gold rush terms, a completely separate mountain range that nobody has properly staked yet.

The Assay Offices Show Up

In any gold rush, the prospectors arrive first. Then come the assay offices, the banks, the railroads, and the merchants selling picks and shovels. The infrastructure follows the ore.

On February 16, 2026, Anthropic opened a new office in Bengaluru, its first in India, framed around supporting a fast-growing market and expanding enterprise adoption. Anthropic has also emphasized work on Indic language performance and partnerships designed to improve model behavior in local contexts, including collaborations with Karya, an organization that pays rural Indian workers fair wages to generate and annotate training data in local languages. Google, Microsoft, and a roster of other labs and platforms have made similar moves, each staking a claim in a slightly different part of the territory.

The IndiaAI Mission has been publicly framed around scaling shared national compute capacity, with plans to onboard tens of thousands of GPUs into a national infrastructure that smaller players and researchers can access. That is the equivalent of building the assay office before the prospectors arrive so that when the ore comes out of the ground, there is somewhere to process it.

The logic running underneath all of this is straightforward. Better local language performance raises adoption. Higher adoption generates more high-quality local interactions. Those interactions feed back into training and evaluation. The cycle compounds, and whoever holds the infrastructure at each stage of that cycle holds

significant leverage over what gets built, what gets improved, and what the economics of Indian AI look like for the next decade.

The Claim Jumper Problem

The original California gold rush had a claim jumper problem. Independent prospectors would stake a territory, do the hard work of locating a productive seam, and then find that a larger operation with more capital and legal firepower had moved in and redrawn the boundaries.

India's government has been explicit about not wanting a repeat of that dynamic in the data economy. The framing around "data sovereignty" and "sovereign compute" at the summit was not accidental. It reflected a deliberate policy position: the data generated by Indian users should produce value that flows back to India, not simply upstream to servers in Virginia and Oregon.

That tension is real and unresolved. Frontier labs need Indian data to improve their models. India needs frontier labs to accelerate its own AI development. The negotiation happening between those two facts will shape the terms of the rush, who gets to mine what, who owns the claims, and who gets paid when the ore moves.

The One Thing You Cannot Import

You can import GPUs and finance data centers, but you cannot instantly produce a generation of people who understand how to build with AI, audit it, stress test it, and use it responsibly in messy real-world contexts. That pipeline takes years, and it starts earlier than most universities want to admit.

In the original gold rush, the people who got rich were not always the prospectors. The merchants, the engineers who built the water systems for hydraulic mining, the lawyers who understood land claims, and the people who figured out logistics and supply chains often did better than the people swinging picks. The same pattern tends to hold in technology transitions. The talent that understands how to operate and build within a new infrastructure often captures more durable value than the people who arrived earliest with the most capital.

India's grassroots AI education layer is, in this framing, the equivalent of training the engineers and the merchants, not just the prospectors. Programs that treat students as builders rather than consumers of technology are doing early-stage workforce development for an industry that does not fully exist yet in its mature form.

This summer, I am stepping into that layer personally. In July 2026, I will be in India to teach a week-long AI accelerator at Mayo College, in collaboration with **Big Red Education**, as part of their "Command Z: Future Tech Lab" program running July 6 to 11, 2026. The students I will be working with are high schoolers. They will be working adults in the early 2030s, which is roughly when the infrastructure being announced right now starts producing its full economic output.

The timing is not a coincidence. It is the point.

What the Map Looks Like Now

India is pushing from the top down, aligning sovereign compute with the global supply chain and negotiating the terms under which its data enters the global training ecosystem. Programs like the one I am joining push from the bottom up, building the human layer that determines whether infrastructure investments

actually convert into domestic capability.

When those two forces overlap, what emerges looks less like a curriculum and more like a national strategy taking shape at both ends simultaneously.

Infrastructure expands access, access generates adoption, adoption produces local data and local demand, and demand eventually funds the startups, deployments, and skills programs that grow the next round of talent. The cycle does not need a coordinator. It just needs enough entry points.

The 1849 prospectors who arrived in California were not thinking about statehood, railroad policy, or what the University of California system would eventually produce. They were thinking about the next pan, the next seam, the next claim. The larger arc assembled itself from millions of individual decisions made by people who were mostly just trying to find the gold.

That is probably how this goes too. The summit announcements, the office openings, the skills programs, and the teenagers in a classroom in Rajasthan learning to prompt and build and evaluate are all individual decisions that, taken together, start to look like something much larger.

Seventy years ago, Dartmouth's crew tried to bottle intelligence in math. They succeeded enough to change the world, and the world went off and built the machines to match the idea. Now those machines are being built at planetary scale, and the future will be written in power purchase agreements, chip diplomacy, and classrooms where teenagers learn how to steer a tool that can outpace their teachers.

The motherlode is in India. The rush is already underway. The question worth asking now is who gets to keep what they find.