DCVC

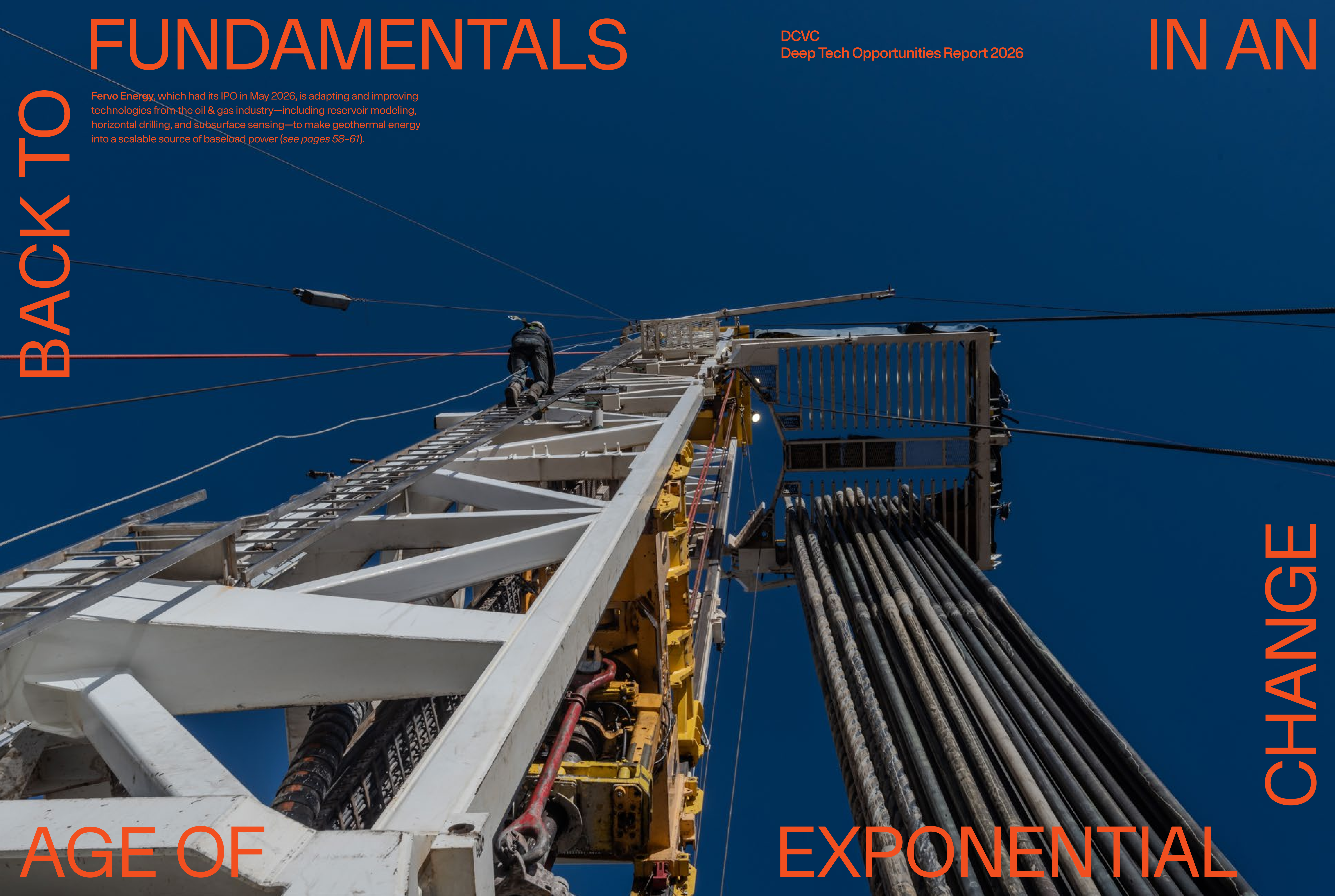
004

Deep Tech Opportunities Report

2026

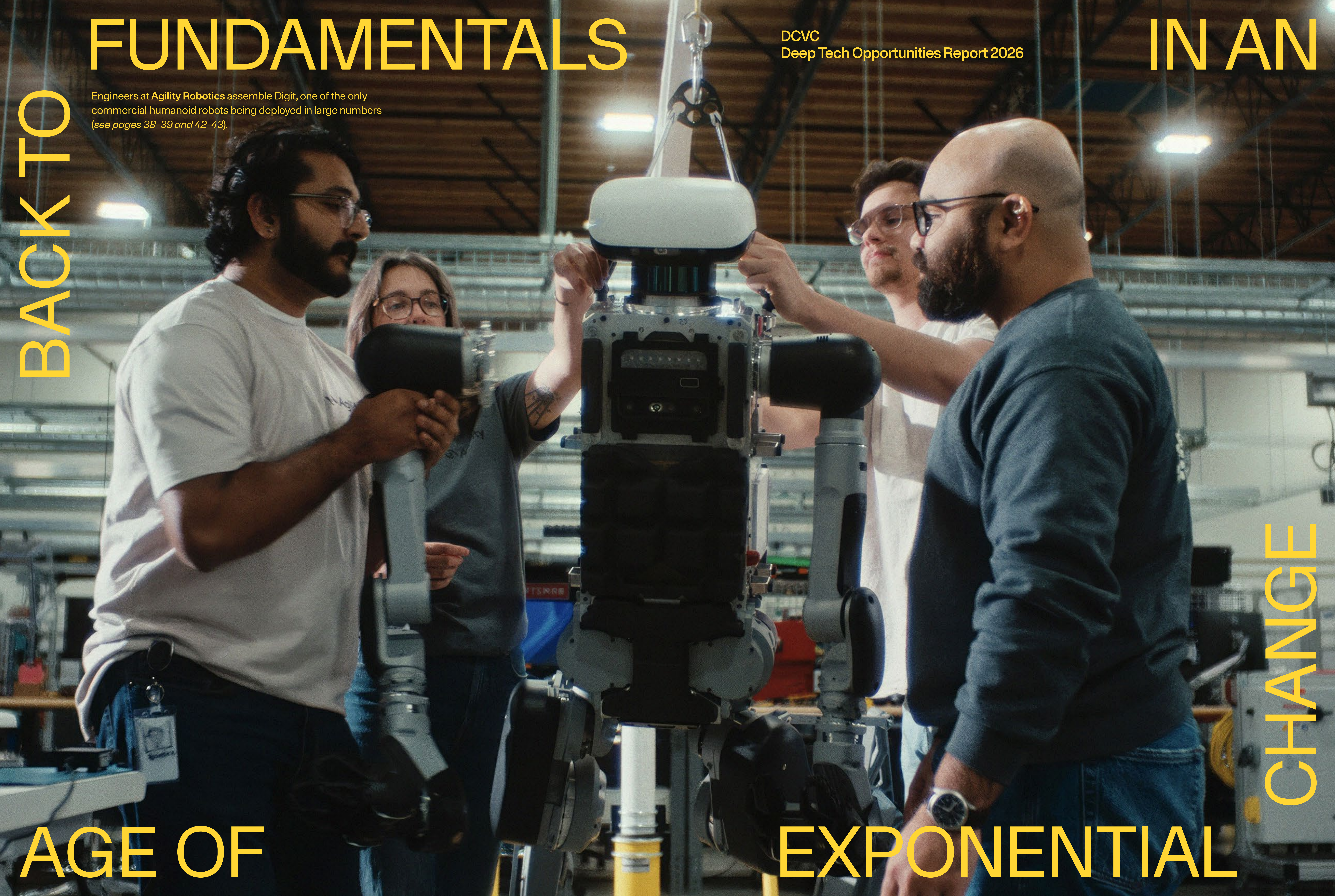
BACK TO FUNDAMENTALS

BACK TO FUNDAMENTALS IN AN AGE OF EXPONENTIAL CHANGE



Fervo Energy, which had its IPO in May 2026, is adapting and improving technologies from the oil & gas industry—including reservoir modeling, horizontal drilling, and subsurface sensing—to make geothermal energy into a scalable source of baseload power (*see pages 58–61*).

BACK TO FUNDAMENTALS IN AN AGE OF EXPONENTIAL CHANGE



Engineers at **Agility Robotics** assemble Digit, one of the only commercial humanoid robots being deployed in large numbers (*see pages 38–39 and 42–43*).

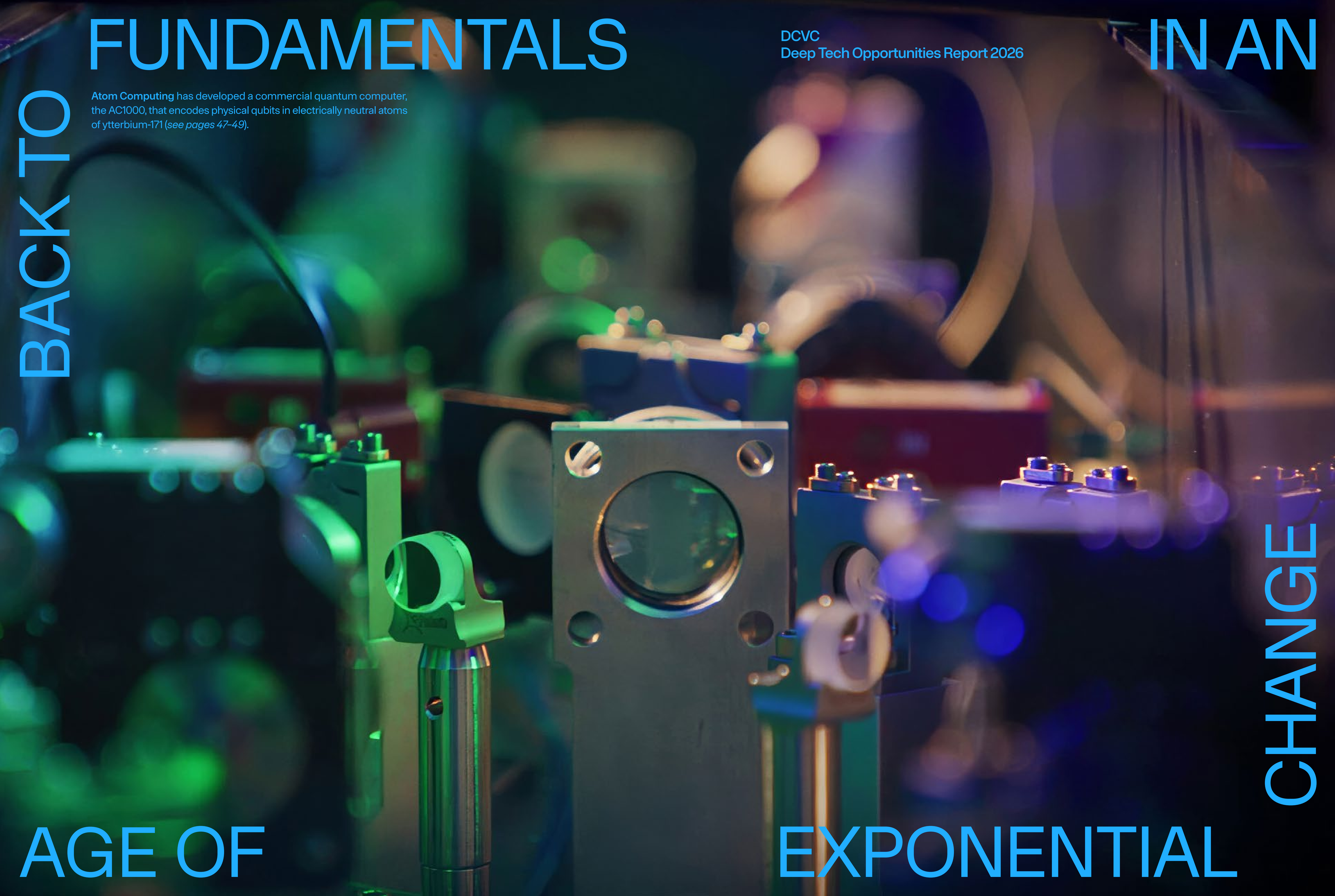
BACK TO FUNDAMENTALS IN AN AGE OF EXPONENTIAL CHANGE

Halter builds solar-powered smart collars for cattle that replace physical fencing. The collars use sound and vibration to guide animals away from or into digitally designated areas—allowing ranchers to manage more cattle on the same land with better weight gain from precision grazing.

BACK TO FUNDAMENTALS IN AN AGE OF EXPONENTIAL CHANGE

Rocket Lab has delivered 260 satellites to orbit with its two-stage Electron rocket (*see page 131*).

BACK TO FUNDAMENTALS IN AN AGE OF EXPONENTIAL CHANGE



Atom Computing has developed a commercial quantum computer, the AC1000, that encodes physical qubits in electrically neutral atoms of ytterbium-171 (*see pages 47–49*).

The real economy of 2026—the environment where all entrepreneurs, innovators, and investors must function—is marked by persistent unpredictability. Thanks to a handful of compounding forces operating at national and planetary scale, it's becoming harder to plan for growth, not easier:

01

Hyperscale computing and AI, plus the electrification of industry and transportation, are driving unsustainable competition for energy.

Power, water, land, permitting, and capital constraints add months or years to construction projects and make it difficult to scale domestic production quickly when needed.

02

Natural sources of fresh water are being exhausted.

Regional water shortages—intensified by climate volatility—affect industry, agriculture, and cities.

03

The physical infrastructure is aging and brittle.

Power grids, water systems, and transportation networks in the United States and many other developed nations were built for a different era, and are vulnerable to cascading systems-level failures and external attacks.

04

Supply chains for critical minerals and materials are fragile.

Markets are opaque and overly concentrated; geopolitical rivals control key inputs for semiconductors, batteries, vehicles, and other products.

05

Runaway costs in healthcare, medicine, and insurance are inflicting enormous pain on households.

This is especially true in the United States and other nations with aging populations and a high chronic disease burden.

06

Labor shortages in skilled technical and industrial roles limit U.S. competitiveness.

We aren't grappling with AI- or robotics-induced mass unemployment (yet), but with its opposite.

07

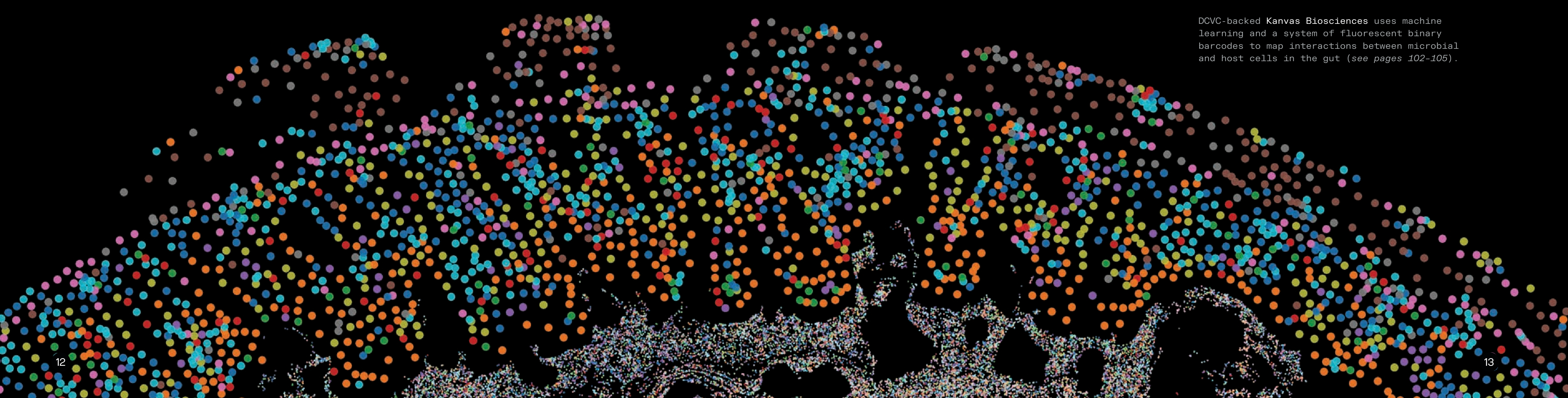
Climate instability is generating more frequent and more costly disasters.

Storms, floods, heat waves, and wildfires are magnifying infrastructure repair costs and insurance losses.

08

The geopolitical picture remains unsettled.

The United States faces global and regional adversaries who are experimenting with new forms of warfare, including autonomous and attritable systems and influence operations.



DCVC-backed **Kanvas Biosciences** uses machine learning and a system of fluorescent binary barcodes to map interactions between microbial and host cells in the gut (*see pages 102–105*).

To compress the point: The risks and uncertainties that have always made planning a challenge are only growing today as powerful technologies collide with physical and institutional limits. For example, if supplies of AI chips keep growing at current rates, the electricity required to operate and cool those chips will exceed the power capacity of the entire state of California by 2028, according to RAND. Trajectories like this are obviously untenable. Exponentially improving technological capabilities *cannot be allowed* to translate into exponentially increasing energy requirements, costs, material demands, and project timelines—if they do, the industrial economy will collapse under its own weight.

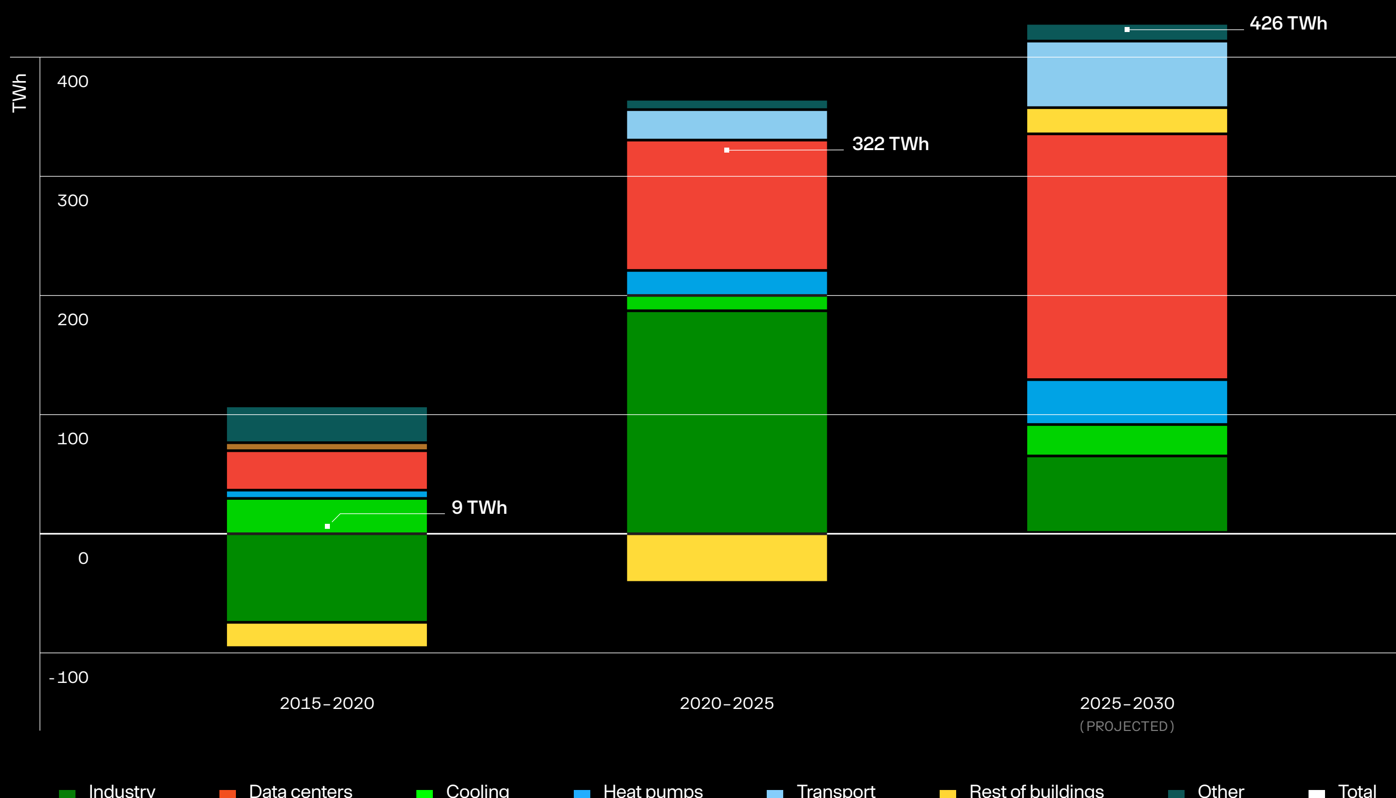
But the way we see things here at DCVC, there's still one critical lever that can help bend escalating cost and schedule curves downward, begin to restore predictability, and, as we like to say around here, "cancel the apocalypse."

That lever is **fundamental deep tech innovation.**

Figure 0.0.1

Electricity demand growth by sector in the United States.
Between 2015 and 2020 demand decreased in some sectors and increased in others.
Between 2020 and 2025, industry and data centers sparked renewed demand growth.
After 2025, data centers began to outstrip all other sources of demand.

Source: IEA.



1

Our quest as investors—as we've explained in previous editions of this report—is to find companies using computation to engineer solutions to trillion-dollar problems that older technologies haven't fixed. But not just *any* solutions. The ideas that interest us most have three special qualities:

Most importantly, we look for innovations that pay for themselves rather than adding to ever-rising costs. This means they don't require permanent subsidies; they don't shift costs elsewhere in the system; and they don't look good at pilot scale only to grow unaffordable as they scale up. They reduce the need for blind experimentation through sensing, simulation, and prediction. They're built around computation and rapid learning curves that make physical processes faster with iteration, so their unit economics improve with deployment. And the gains they achieve are measurable, defensible, and compounding.

For an example, look at advanced geothermal company **Fervo Energy,** which we've backed since 2022, and which raised $2.2 billion in an initial public offering on the Nasdaq exchange this spring. By applying cutting-edge drilling techniques from the oil and gas industry, real-time downhole sensor data relayed by fiber optic cables, and hands-on experience in the field, Fervo has been able to dramatically lower the time and cost required to drill new wells. By 2028 the company expects to have 500 megawatts of firm clean power on the grid—as much as a typical combined-cycle natural gas power plant—and over time its plan is to outcompete natural gas on a cost-per-megawatt basis.

Or look at **Relation Therapeutics,** a TechBio standout where we first invested in 2021. The company's "lab-in-the-loop" approach lowers the cost of drug development by shaving months to years off the usual discovery timelines for new drugs. Relation collects enormous amounts of gene and protein data from healthy and diseased human tissues; trains AI models on that data so they'll recognize the genes and proteins most likely to be involved in disease; uses predictions from those models to modify cells for the next round of experiments; and so on, until it identifies wholly new molecular targets for interrupting disease processes and molecules that can act on them. Pharma giant GSK, among others, has validated Relation's approach through partnership agreements worth tens of millions of dollars in upfront cash and billions of dollars in milestone-based payments.

2

As we search for ideas that will help retool old industries and give rise to new ones, we also focus on those that will be relatively invariant to political and regulatory change. To be venture-investable, in our view, a new technology needs a regulation- and ideology-agnostic value proposition that will survive each successive administration. When technologies win on economics first, adoption can be driven by lower cost, higher reliability, and better performance, not mandates or subsidies.

Consider **Mainspring Energy,** which builds innovative flex-fuel linear generators that can be placed on-site near factories or data centers to provide primary or backup power, which replaces dirty, expensive, inflexible diesel generators. Advanced power electronics let these parking-space-sized generators switch between natural gas, propane, biogas, hydrogen, or ammonia with no modification while compensating automatically for different energy densities. They can be deployed under existing permitting standards, without grand infrastructure overhauls—Prologis, for example, turned to Mainspring to power EV chargers for electric trucks at its Los Angeles and Long Beach port facilities, completing the installation in just nine months. With wait times for new grid connections stretching to five to eight years, it'll take exactly this kind of innovation in behind-the-meter power sources to support the "electrification of everything" we highlighted in our 2023 report and the swift reshoring of manufacturing we talked about last year.

Similar logic applies to **Radiant Industries,** which makes 1-megawatt advanced nuclear fission reactors perfect for powering military bases, replacing backup diesel generators, and powering off-grid communities. In 2026 the company expects to reach the landmark achievement of 'turning on' the United States' first-ever advanced fission nuclear reactor, and by 2034, Radiant expects to have up to 50 reactors in the field. Energy resilience like this matters under any geopolitical scenario. It doesn't presuppose consensus across party lines; it strengthens domestic capacity; and it provides direct benefits to manufacturers and municipalities, creating local champions.

3

A third quality that unites almost all the companies we back is that they use advanced forms of computation, especially machine learning, to make natural or mechanical systems more programmable and engineerable and, as a result, more efficient.

Consider **Sabanto, AgZen,** and **Verdant,** three agtech companies backed by DCVC Bio. They're all confronting the reality that farming still involves enormous amounts of waste, whether of human labor time or chemical inputs like fertilizer and pesticides. Sabanto builds kits farmers can use to make their existing tractors self-driving. Verdant and AgZen make sensing systems that can adjust spray output for weather conditions and deliver pesticides more precisely, down to the level of individual weeds. In complementary ways, these companies are using imaging, AI, and robotics to make farms behave more like computational systems—so that farmers can tend their fields with fewer people and spend less money on crop protection.

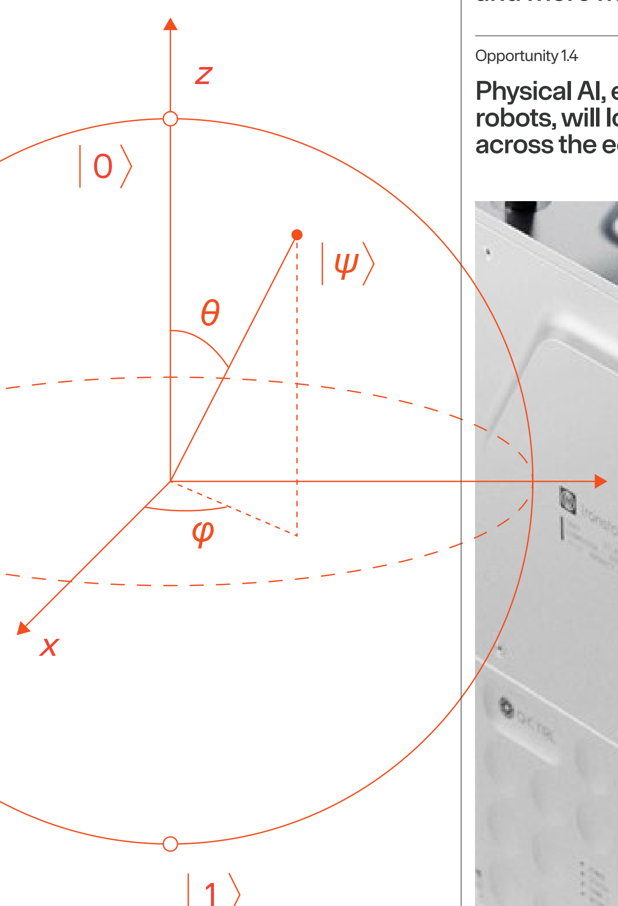
Identifying and supporting the entrepreneurs and companies that best embody these three qualities is our job. The story plays out differently across our areas of specialization—computing, energy, water, healthcare, drug discovery, agriculture, space, and defense.

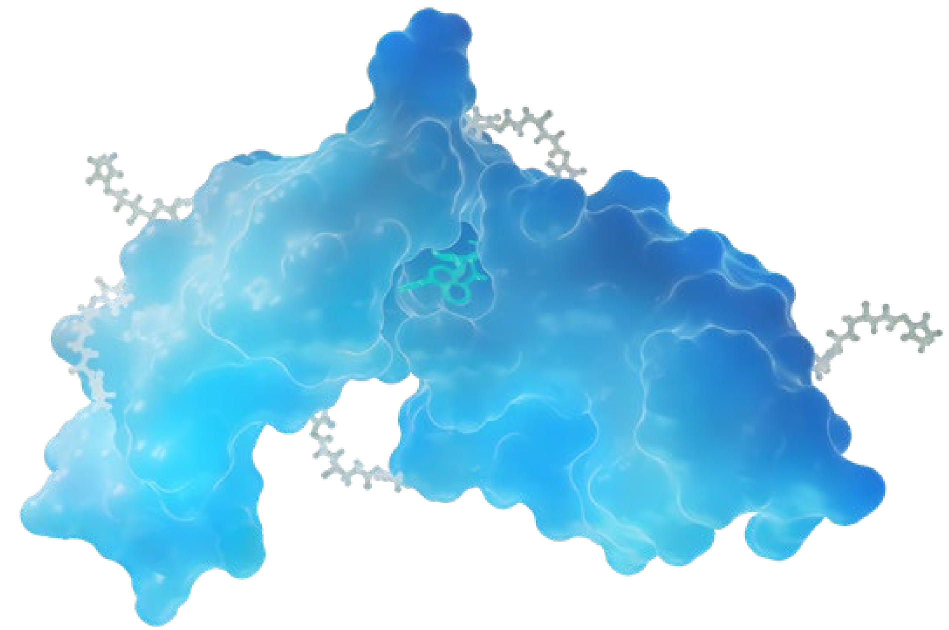
And different elements move to the foreground as market conditions evolve. But DCVC's fundamental premise hasn't changed since our founding in 2010. We see advanced computing as a universal solvent that allows the ingredients of innovation to mix—lowering the cost and risk of experimentation, revealing ground truths faster, smoothing out costly technical wrinkles, and making chaotic systems more predictable and engineerable.

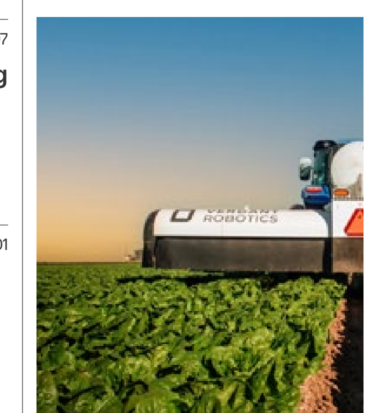
Each year the Deep Tech Opportunities Report gives us the chance to describe what new bets we're placing, how old bets are paying off, and how we're adapting our thinking in each investing area. Here are a few of the high-level trends you'll notice in the report; many of our investments respond directly to these realities.

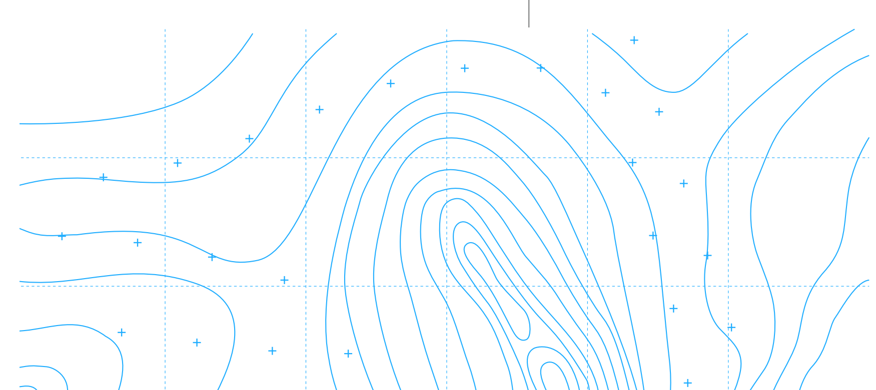
- In classical computing, there's an urgent need for fundamental improvements in speed and efficiency at all levels, from chips to data centers to machine learning architectures to physical AI.
- In quantum computing, we are at the edge of perfecting new hardware modalities, error correction mechanisms, and sensing technologies that will make massive computational gains a reality for business and industry.
- In the energy system, deep tech can provide a way out of the dilemma posed by exponentially rising power demand; aging, outdated physical infrastructure; and an increasingly volatile climate.
- In energy and climate tech, advances in classic deep tech disciplines such as power electronics or subsurface drilling techniques can help us get structural risks under control and improve the resilience of communities and industries.
- In industry, we can achieve our reindustrialization, reshoring, and decarbonization goals if we find ways to rationalize logistics, build smarter factories, make low-carbon steel, cement, and fuels, and secure predictable, sovereign supplies of water and critical minerals.
- In the life sciences and agriculture, we can speed up drug discovery, lower failure rates, reduce healthcare costs, cure more human diseases, and make food systems healthier and more efficient by introducing bigger and better computational models, trained on more high-quality data.
- In defense, deep tech innovators can deploy new forms of autonomy to transform the battlefield and deploy countermeasures against hostile actors.

If there's a master trend here, it's that better technology reduces risk and uncertainty and makes the world a more predictable place. And we believe the best place to look for fixes and improvements is deep inside a system rather than on its surface. Not just better software but more efficient chips. Not just greener hydrocarbon fuels but carbon-neutral, drop-in replacements for those fuels. Not just clean, renewable energy but energy that's so cheap that businesses must switch away from hydrocarbons out of economic self-interest. Not just more drugs but drugs that are safer in a predictable way and developed at far lower cost and in less time. Those are the kinds of ideas venture-backed, computation-first deep tech companies can uniquely explore and exploit. We hope you'll enjoy this year's look at how our portfolio companies are doing this, and how we see them driving the evolution of their industries in the near future.


z
|0⟩
|ψ⟩
θ
φ
x
|1⟩







A few words on the investing climate.

Anticipated IPOs, new AI architectures, and geopolitical reorganization create countless openings for deep tech innovation.

THE GLOBAL ECONOMY is like an ever-shifting weather system affected by short-term storms, seasonal El Niño-like patterns, and long-term weather and climate trends. And like good meteorologists, we constantly monitor the financial, market, and regulatory conditions creating tailwinds or headwinds for our portfolio companies.

As always, we'd like to open the Deep Tech Opportunities Report with a data-driven survey of some of the weather we're tracking. Our observations this year center on the reopening of the IPO window, the growing likelihood of large liquidity events in AI and space, the continued scramble to lower the cost of computation and energy for an AI-heavy economy, and the rising strategic pressure to rebuild supply chains closer to home.

Biotech's IPO window has begun to reopen.

Historically, 35 to 45 biotech companies went public each year. But a multiyear post-pandemic IPO drought dragged into 2025, when there were only nine meaningful biotech IPOs ("meaningful" in the sense that the offerings raised at least $75 million). In the second quarter of 2025—for the first time in a decade—there were no biotech IPOs at all, due in part to higher interest rates, regulatory uncertainty, and budget uncertainty at the Food and Drug Administration and related government agencies. But the market began to thaw in the fourth quarter, and in the first quarter of 2026, we saw half a dozen meaningful biotech IPOs, including for Aktis, Veradermics, Eikon, Spyglass, Agomab, and Generate. In February alone, companies raised $1.4 billion—a sign that generalist investors are coming back into the market. Barring a major macroeconomic shock, the market could produce 25 to 30 IPOs in 2026. The upshot is that venture investors in biotech can begin to cash out their positions in this generation of biopharma companies and return funds to their LPs, allowing the innovation cycle to continue.

The prospect of mega-IPOs in AI and space could reset venture liquidity for years.

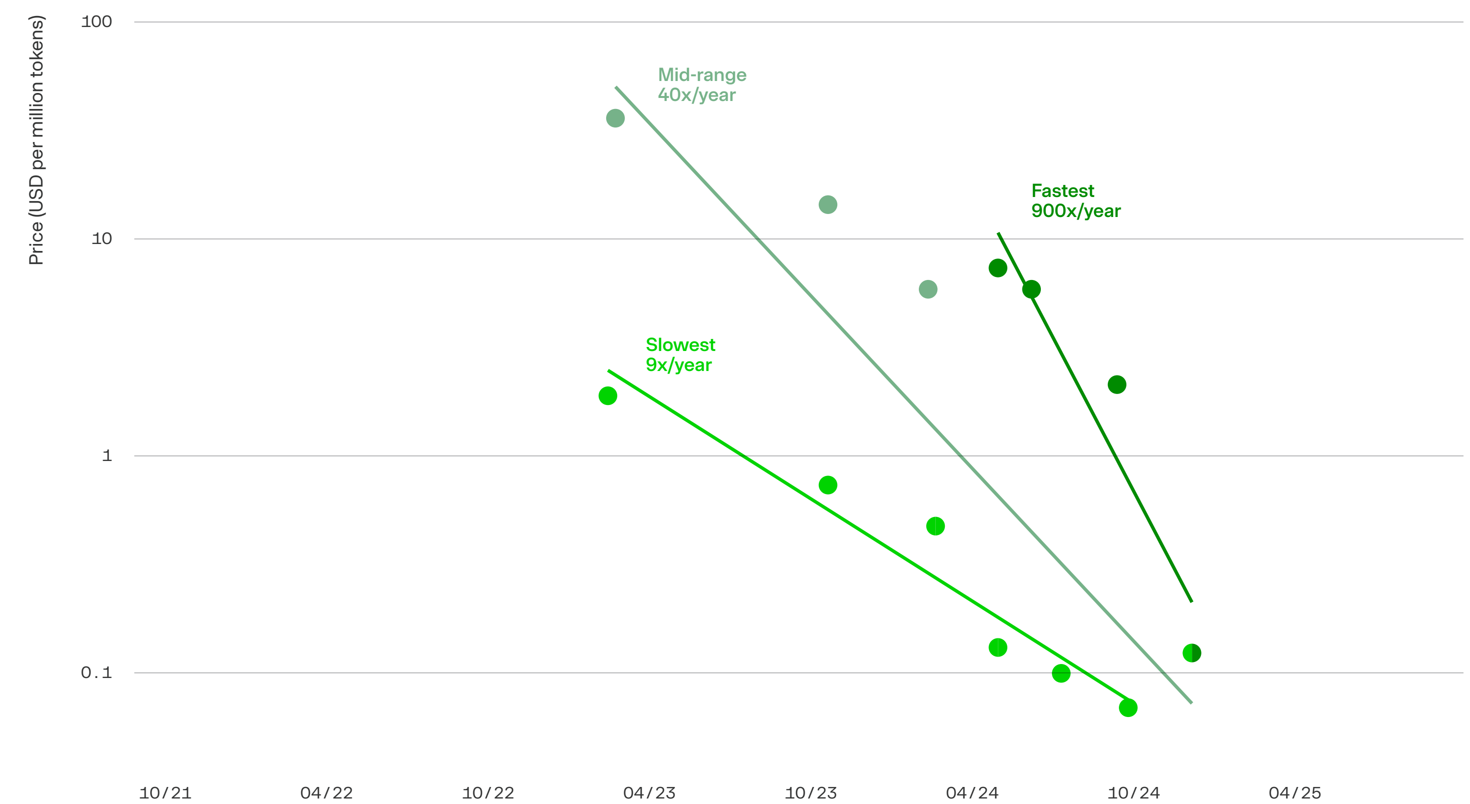
In May SpaceX filed the prospectus for its initial public offering, targeting a valuation of $1.75 trillion. As this report went to press (on May 22), Open AI was not far behind SpaceX, and was reported to be preparing to file confidentially for an IPO within weeks. The LLM maker was valued at $730 billion in its most recent private funding round. And Anthropic is on a similar trajectory; its most recent financing valued the company at $900 billion, with revenue doubling to $10.9 billion in the second quarter, according to reports in the *Wall Street Journal*. Even the credible prospect of IPOs on this scale is already reshaping the venture environment, as private-market valuations get marked up in anticipation of future liquidity. Such listings could release enormous amounts of venture LP capital for new investments; leave all three companies in better position to acquire newer, smaller firms; and free up talent to found the next generation of AI and space companies. But they could also suck up public capital and temporarily suppress the market for other tech IPOs. We'll be watching all of these developments closely. And we're also monitoring a tide of energy-company IPOs that could lift many boats in the industries devoted to building physical things.

Figure 0.1.1 Biotech IPOs, by year and disease category; 2026 figures are current as of May 15. Source: Biopharma Dive.

FOCUS	2018	2019	2020	2021	2022	2023	2024	2025	2026
Cancer	20	18	38	49	7	7	7	1	2
Immune Diseases	4	4	7	7	1	3	6	3	3
CNS Disorders	3	6	6	16	5	1	4	2	1
Rare Diseases	10	3	12	12	1	0	2	1	1
Infectious Diseases	5	1	4	6	3	1	0	0	0
Other	12	15	12	14	5	7	5	4	4

Figure 0.1.2 The price of performing an inference on an LLM has decreased by 9x to 900x per year, depending on the task, but application companies are offsetting these gains by performing far more inferences.

Source: Epoch AI, Artificial Analysis.



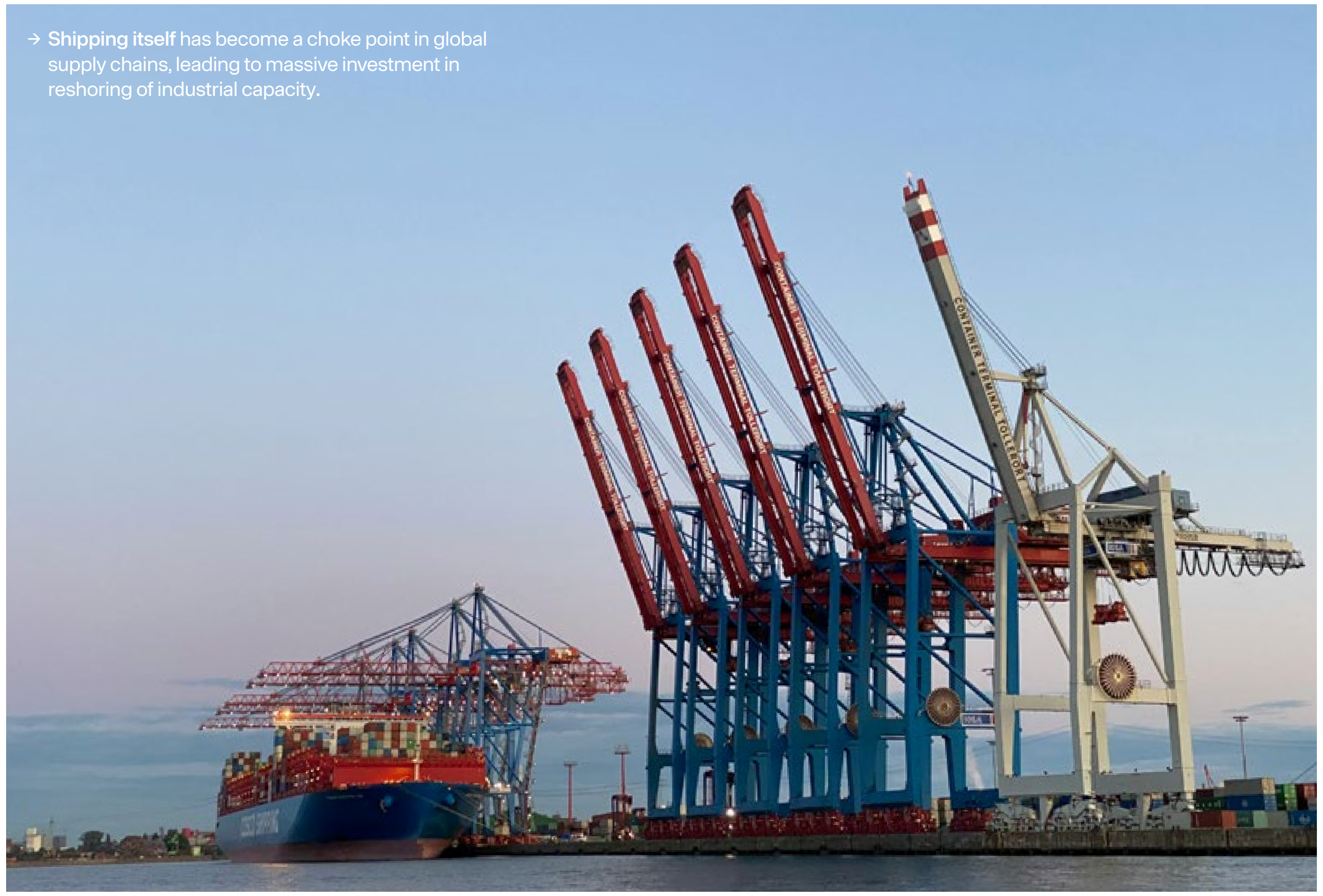
As AI shows up in more applications across the business world, there's a looming cost-structure problem.

The cost of inference from frontier AI models is falling rapidly (see Figure 0.1.2) thanks to model improvements, hardware advances, and other techniques—but not fast enough to offset exploding inference demand as AI gets integrated into more products. For application-layer companies whose primary expense is model inference—think coding assistants like Cursor, search engines like Perplexity, or productivity tools like Grammarly—this widening gap will squeeze margins and create significant headwinds. But it is also creating powerful tailwinds and investment opportunities where companies are developing more efficient models and compute infrastructure that can reduce the cost of inference by orders of magnitude. For more, see Opportunity 1.2 and Opportunity 1.3.

The age of monolithic AI architectures is already giving way to something more complex.

The early generative AI boom (roughly 2020 to 2022) was driven by a simple hypothesis: make one model large enough, train it on enough data, build a data center with enough standardized chips to run it, and it will be able to perform almost every cognitive task. This made system design relatively straightforward. The problem, as we just noted, is that frontier models are extremely expensive to run. In practice, most tasks do not require a frontier model at all. A more efficient architecture is beginning to emerge—

(↪ Continued on page 22)



→ **Shipping itself** has become a choke point in global supply chains, leading to massive investment in reshoring of industrial capacity.

one that distributes work across multiple agents, multiple models, and multiple types of hardware. Powerful GPUs may handle large reasoning tasks, while smaller models run on cheaper inference chips, edge accelerators handle real-time loads, and old-fashioned CPUs coordinate all of their workflows. This shift is creating a market for a new class of heterogeneous compute companies building model-routing platforms, AI orchestration frameworks, composable multi-model agents, and cross-chip scheduling software to determine which model is best suited to the job, where it runs, and how tasks are distributed. As Adam Smith observed in *The Wealth of Nations* 250 years ago, productivity tends to increase when tasks become specialized. AI systems appear to be evolving in the same direction. (Again, see Opportunities 1.2 and 1.3.)

Global decoupling means that supply chains are gradually being restructured along geopolitical lines.

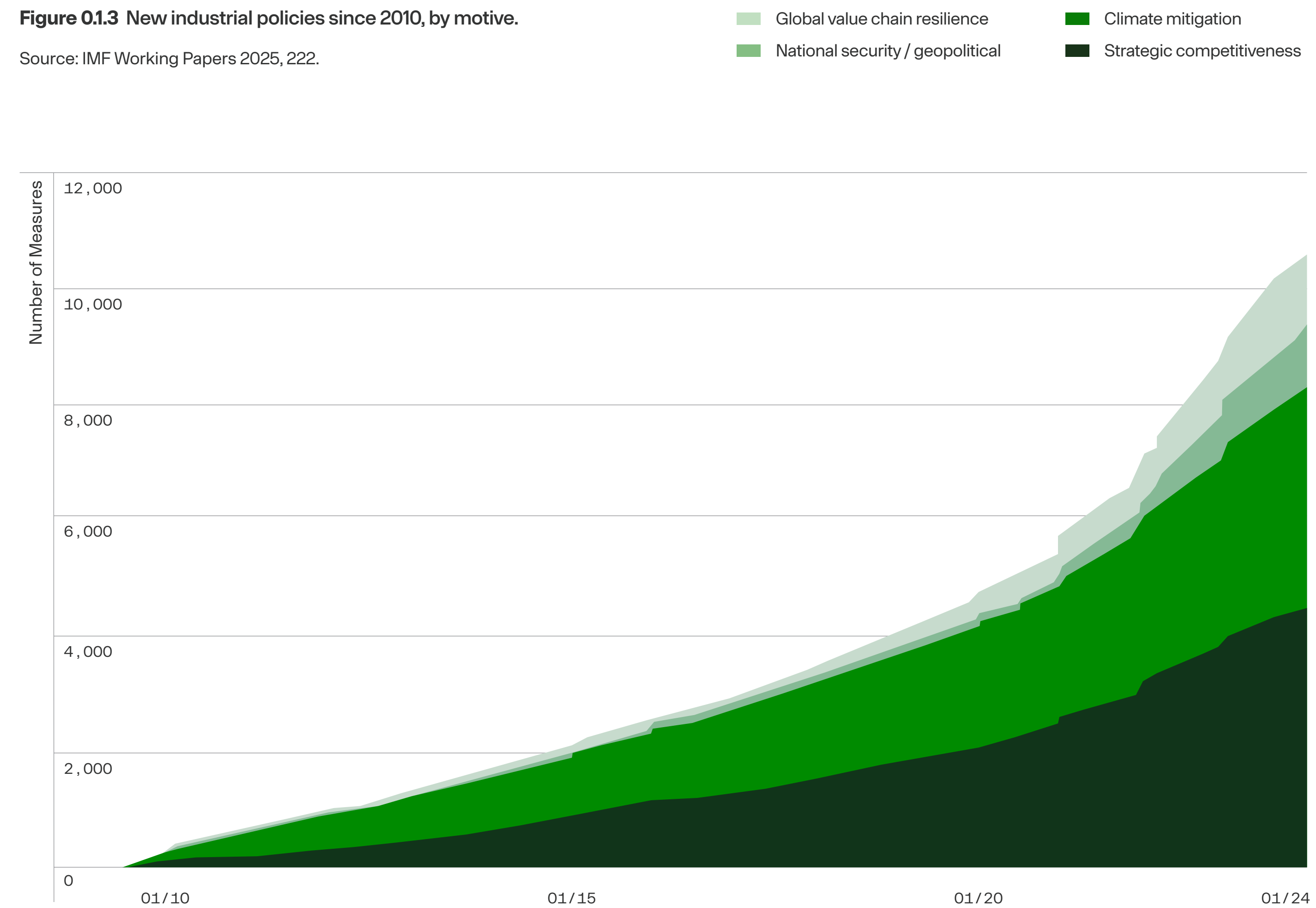
For decades, venture capital mostly ignored industrial sectors, because everyone assumed globalization had already optimized them for cost efficiency and scale. But the pandemic, rising U.S.-China strategic competition, and national security concerns around critical technologies and materials have revealed intolerable imbalances and choke points in global production and distribution networks. As a result, each year the U.S. is now spending hundreds of billions of dollars rebuilding industrial capacity and adding hundreds of thousands of manufacturing jobs through reshoring and foreign direct investment. (Manufacturing construction spending has more than doubled since 2021, driven largely by semiconductor, battery, and clean-energy factories.) Because labor costs are higher in the U.S., smart reshoring requires AI-driven quality control, robotics, and other forms of automation and advanced manufacturing. It also means developing alternative material supply chains for rare earth elements, batteries, and semiconductors, as well as large amounts of firm baseload power. All of this is stimulating a wave of investment in the deep tech innovations required to manufacture goods, process materials, and produce energy closer to home. (We talked about all these trends at length in the 2025 Deep Tech Opportunities Report; see also Chapters 3 and 4 of this volume.)

Industrial policy is returning worldwide, although in many different forms.

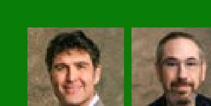
After decades in which governments largely relied on markets to allocate industrial investment, major economies are now actively subsidizing strategic sectors such as semiconductors, batteries, energy, and critical minerals. The IMF observes that new industrial subsidies shot upward globally after 2020, in response to the same concerns over supply chain resilience, national security, and geopolitical competition we just mentioned. In the United States, this takes the form of a patchwork of federal and state programs rather than a single national industrial strategy. Many of the programs launched during the previous administration—notably the CHIPS and Science Act, the Inflation Reduction Act, and Department of Energy (DOE) loan programs—remain in place even as they're being reinterpreted, renegotiated, or partially scaled back. Meanwhile, states have instituted big incentive packages to attract industrial projects, including New York's $10 billion Green CHIPS program, incentives in Georgia for EV and battery factories, and incentives in Arizona for semiconductor fabs. And these incentives are working: In 2025 companies invested $278 billion in the manufacture and deployment of clean energy, clean vehicles, building electrification, and carbon management technology, up 5 percent from 2024, according to MIT and Rhodium Group. That investment wave inevitably expands the market for deep tech companies in advanced manufacturing, materials, and energy.

Figure 0.1.3 New industrial policies since 2010, by motive.

Source: IMF Working Papers 2025, 222.



↳ Deep tech entrepreneurs play a stabilizing role in an unstable and rapidly changing world where the economic weather seems increasingly stormy. Rather than fixating on lofty dreams about techno-utopian futures, they roll up their sleeves and ask the practical questions: *How's it going to work?* and *Who's going to pay for it?* Our team of investors is hyperfocused on finding and funding the best of these innovators, and it's our continuing privilege to help them do their work. ■



Zachary Bogue and Matt Ocko
Co-Founders and Managing Partners, DCVC

CHAPTER 1

We must reinvent classical computing—and physical AI—for an era of rapid scale-up.

As the AI revolution drives new demand for computing capacity—and for the energy needed to power data centers—we need ways to make this revolution affordable, which means doing everything much more cost-effectively. Better underlying technologies in areas like chip manufacturing, data center management, internet architecture, and robotics will bring greater speed, power efficiency, and utilization to the whole IT infrastructure, making it more sustainable to scale up. We care about this problem because computing is the flywheel that powers almost every other form of innovation—including what goes on at most of our portfolio companies. Solving it is a prerequisite for continued economic growth and national competitiveness.

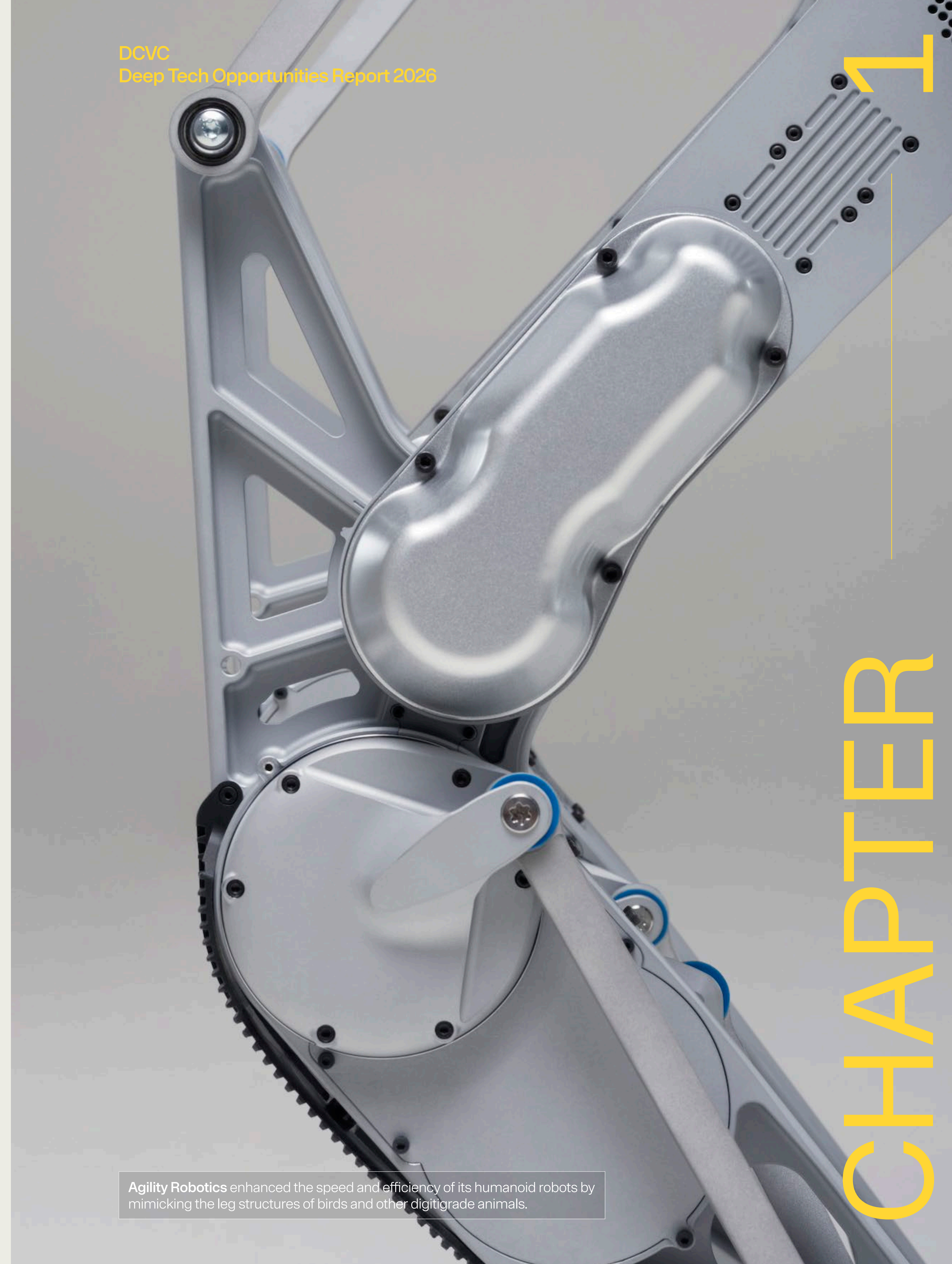
OPPORTUNITIES

COMPANIES

Agility Robotics, Fulfil, Mythic, Proprio, Remedy Robotics, SF Compute, Slip Robotics, TypeSafe, Unconventional

VOICES

Zachary Bogue, Alan Cohen, James Hardiman, Ali Tamaseb



Agility Robotics enhanced the speed and efficiency of its humanoid robots by mimicking the leg structures of birds and other digitigrade animals.

Next-gen lithographic techniques and smarter chip architectures will help rein in AI's power requirements.

There's nothing broken, exactly, about the current regime in semiconductor manufacturing. It has brought us unexpected and unprecedented benefits—putting a supercomputer into every pocket, hundreds to thousands of microchips into every vehicle, and AI-powered features into practically every software application. But this success rests on standard chip design and manufacturing paradigms that are becoming increasingly fragile, costly, and difficult to scale. Unlocking new efficiencies will depend on rethinking every layer of the hardware stack: how chips are made, how they are physically organized, how computation is distributed, and how efficiency itself is conceptualized and represented. "The infrastructure layer of the AI stack is becoming interesting and investable," says DCVC Co-Founder and Managing Partner Zachary Bogue.

Manufacturing the CPUs and GPUs that power every mobile device or data center starts with the optical lithography machines companies use to build their chips. This equipment is among the most intricate and expensive machinery ever made. Transitioning from one generation of lithographic techniques to the next, to achieve greater feature density and extend Moore's Law-style device scaling, is always an engineering tour de force.

Just reaching the current plateau—using extreme ultraviolet or EUV wavelengths to etch features at the 7-nanometer scale and below—demanded sustained research over roughly two decades and many billions of dollars of cumulative R&D investment before production-ready tools emerged. Today's commercial EUV lithography systems are manufactured by only one company, ASML, with optics from Carl Zeiss and lasers from Trumpf, and they can cost $200 million to $400 million each. The whole point of these machines is to create smaller and smaller features on silicon substrates. (Smaller transistors mean more of them can be crammed onto each chip; since the electrons travel shorter distances they're more efficient as well.)

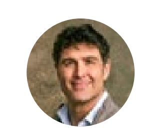
In traditional lithography, the only way to make features smaller is to use either shorter wavelengths of light or brighter light, or both. This light shines through and/or reflects off a patterned, stencil-like photomask, exposing a photoresist, which allows the patterned regions to be selectively removed so that the pattern can be etched into the underlying silicon. But as the EUV approach pushes toward smaller feature sizes (into the low-single-digit nanometer regime), it's hitting fundamental limits. Defects in photoresist chemistry and noise from individual photons constrain the precision and resolution of lines and other features.

One feasible way to keep going smaller is to switch to a different lithographic regime altogether. What if the photomask were no longer a passive stencil? Future nanoscale photomask architecture could sculpt and reshape the electromagnetic field from visible light (not EUV), concentrating energy into localized pockets on the photoresist far smaller than the light's wavelength. If the mask were to become an active optical device—not just a shadow- or reflection-caster—it might be possible to achieve features as small as 2–4 nm, without the high-energy lasers or vacuum conditions required in EUV. "Foundational tech for heavy compute has a number of layers, starting all the way down with the semiconductor layer," says DCVC General Partner Ali Tamaseb. "New mask methods would enable better and finer lithography."

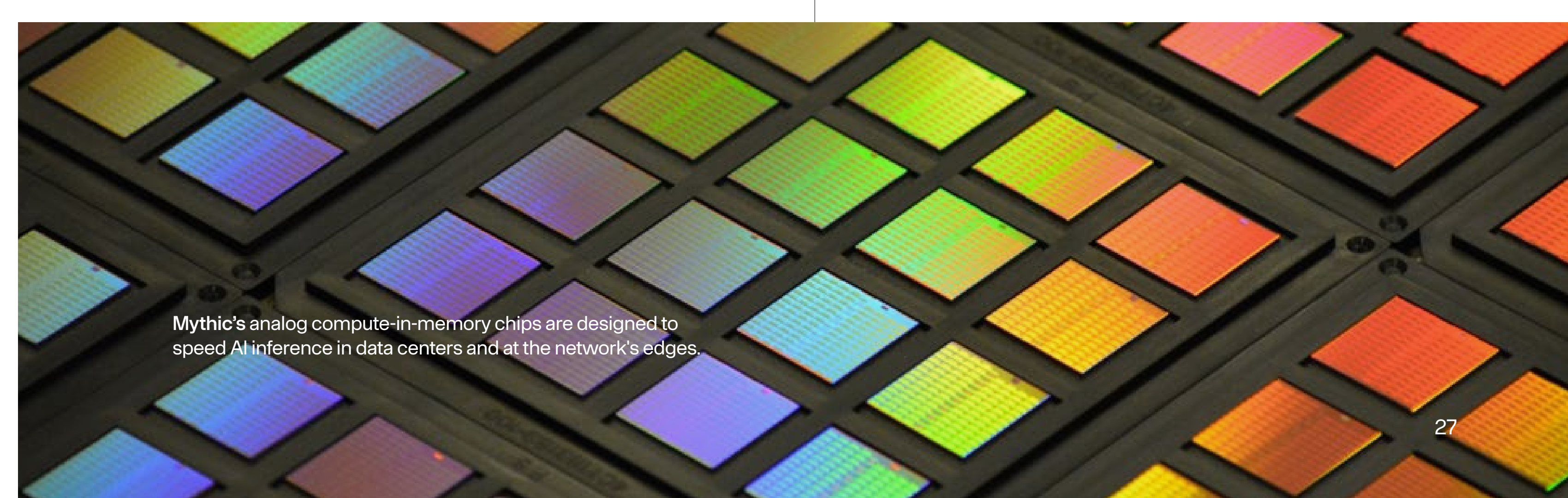
The next way to create chips capable of more operations per second, and at higher energy efficiency, is to think multi-dimensionally. "We either need to go smaller, or we need to build up, like a city with multiple levels," Tamaseb explains.

As an example, chip designers are exploring stacking integrated circuits and building blocks called chiplets to create dense multi-layer designs. These can shorten interconnect distances and allow logic, memory, and other components to be interspersed. The overall idea is to extract more performance and energy efficiency by rethinking how chips are physically organized, not just how finely they're etched. Advanced chipmakers like TSMC, Intel, and Samsung are active here, as are newer companies such as Ayar Labs, which specializes in fast optical interconnects for chiplets.

> "The infrastructure layer of the AI stack is becoming interesting and investable."



Zachary Bogue
Co-Founder and Managing Partner, DCVC

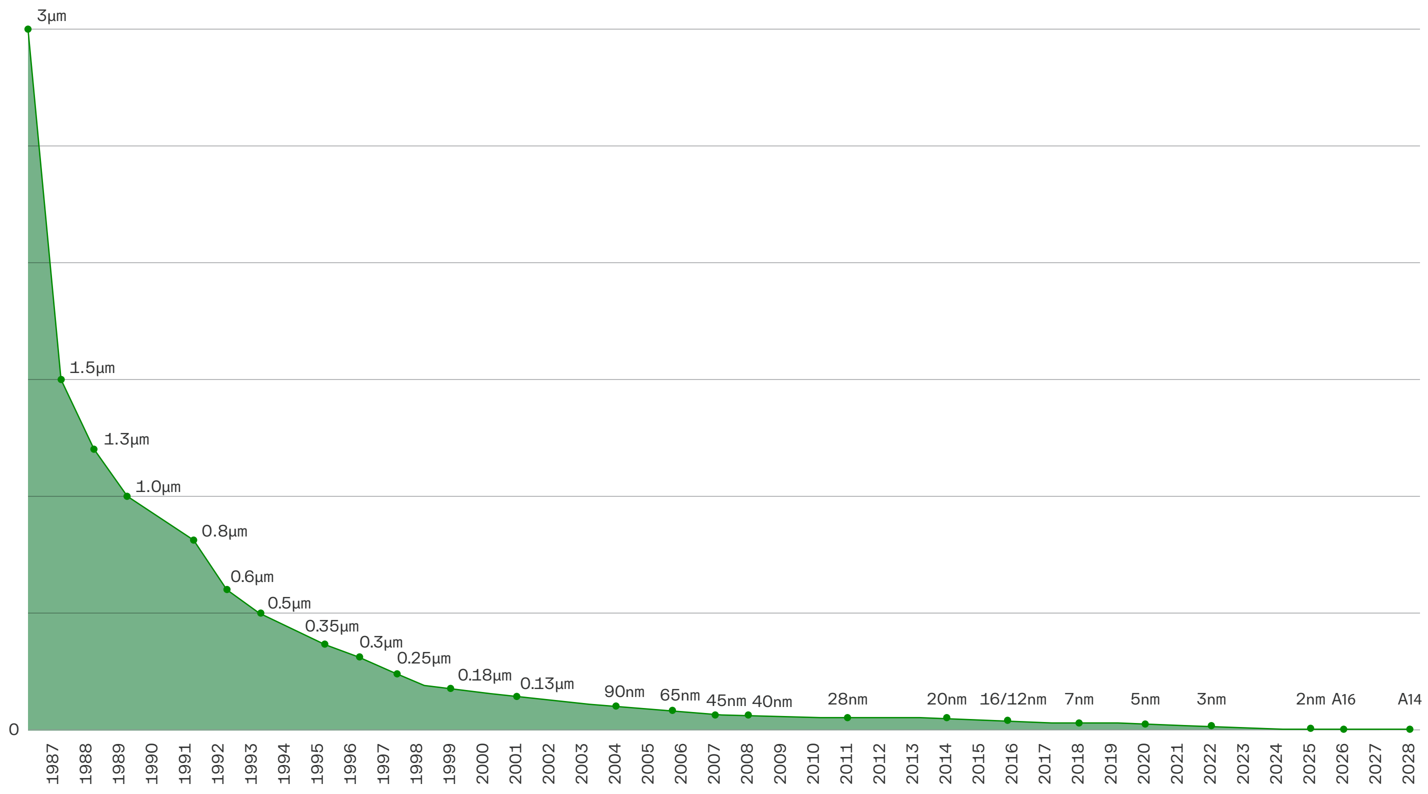


Mythic's analog compute-in-memory chips are designed to speed AI inference in data centers and at the network's edges.

Figure 1.1.1

Since 1987, the size of the smallest features on transistors manufactured by leading chipmaker TSMC has shrunk from 3 µm to 2 nm.

Source: TSMC.



When processor speed and efficiency are paramount, there's also room to bet on more disruptive approaches. One of them is to change the way computers handle the basic math behind machine learning. And it turns out that running that math on analog rather than digital hardware is dramatically more energy-efficient, as our portfolio company Mythic is demonstrating.

Modern ML hardware performs one dominant class of calculations over and over: matrix multiplication and accumulation, where input activations are multiplied by an array of neural network weights. In a digital AI chip that means fetching the weights from memory, fetching the activations as well, then multiplying and adding up the answers in logic and reading out the results. The problem is that in energy terms, all this fetching of data is orders of magnitude more expensive than the actual arithmetic. "We need better ways to reduce the cost of data movement," says Tamaseb. "Mythic has solved it in one way by bringing together compute and memory in the same place."

Mythic's analog chips are, in fact, memory arrays, wired up so that circuit physics performs the math. A Mythic chip is an array of non-volatile analog flash memory elements connected in rows and columns and operated as tunable resistors. Neural network weights are represented in the array by setting the flash cells at different levels of conductance. Input activations are applied along the rows as voltages, and at each element, the voltage drives a current whose magnitude is set by the conductance—the physical equivalent of a multiplication. These currents flow along a shared column, where readout circuitry captures the final sums. Far less data moves at all; the array itself carries out a massively parallel computation by virtue of geometry, Ohm's Law, and Kirchhoff's current law.

Mythic's M1076 compute-in-memory chip can deliver AI throughput comparable to an edge GPU for 1/10th of the power, and the company says its next-generation analog processing units are targeted to be up to 100 times more energy-efficient than conventional digital GPUs. That means tasks like AI inference that used to be confined to data centers can increasingly happen at the network's edges, where computing meets the environment—think drones, smart city and smart home devices, and especially robots. "We're going to see robotics actually take off because edge compute is becoming a thing," Tamaseb predicts. "It used to be that you couldn't run all of this compute on these robots—you had to send it to the cloud, and you had latency problems. Now we are able to actually use top models, big AI models, on the edge, running directly on these robots with low latency. That's going to be a big unlock."

Another potentially transformative edge computing application: brain-computer interfaces. Here we're watching companies such as Infera Neuro, a spinoff from École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland. They're developing tiny, low-power implantable AI chips that could read brain signals and translate them into text, or identify symptoms of epilepsy or movement disorders. One prototype has just 1/25th the area of Elon Musk's Neuralink device—and unlike Neuralink, it carries out processing inside the brain, with no need for power-hungry telemetry. It's one example of what's possible when integrated circuits are small, dense, and frugal enough to take AI out of the data center and into the real world.

One final way to engineer the next big leap in computing efficiency could be to emulate the brain itself. The irony of current-day AI is that while neural network architectures loosely mimic the operation of biological neurons, they require far more power to operate; the human brain runs on only 20 joules of energy per second, or 20 watts, while a modern GPU like NVIDIA's H100 chip needs 350-700 watts under load. Unconventional, a DCVC-backed company that launched last year, wants to change that—and maybe even head off the coming bottleneck in global energy supplies for AI—through brain-inspired or "neuromorphic" computing.

The company was co-founded by Naveen Rao, a computational neuroscientist who previously founded DCVC-backed Nervana Systems (acquired by Intel) and

↑ **New types of AI chips** must be manufacturable at scale. Here, a pick-and-place robot extracts silicon dies from a diced wafer.

MosaicML (acquired by Databricks). The company aims to design a new silicon-based "substrate for intelligence" that requires far less power. Rao and his co-founders say the company will, like Mythic, exploit the analog properties of silicon, running neural networks "on the physics directly" rather than trying to simulate biological system digitally.

It may only be by mimicking evolution, and learning its lessons, that we can create more powerful computers that don't suck up more electricity than we have to give them. "The creatures that live inside their energy envelope are the ones that tend to survive for the longest period," says DCVC General Partner Alan Cohen. "The point is, the constraints of power are forcing us to go back and look at other models."

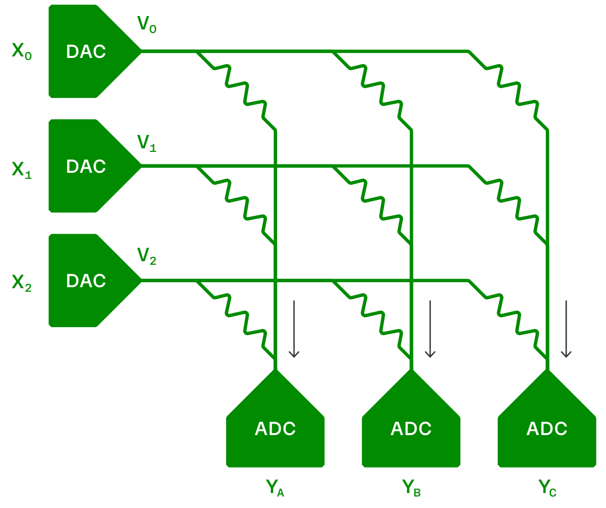
> "The constraints of power are forcing us to go back and look at other models."

Alan Cohen
General Partner, DCVC

Figure 1.1.2

How Mythic performs parallel computation using geometry and physics
Mythic's compute-in-memory arrays speed up neural network matrix operations, where input vectors are multiplied by weight matrices. Tunable resistors at the intersections represent network weights; inputs are supplied as voltages, and outputs are collected as currents.

Source: Mythic.



We must invent new forms of efficiency at the data center and cloud levels.

Not all computation can move to the edges of the network. Huge data centers with shareable compute resources offer users a level of accessibility and affordability that would be impossible to deliver any other way. Until recently, cloud users operated in a regime where the supply of computing was relatively abundant—but actual GPU hours were, paradoxically, quite illiquid. Today, growing power needs, construction delays, and capital requirements are putting a ceiling on the supply of raw compute power, even as there's more demand for flexible access to that power. We're investing in companies that tackle those problems at the level of the market and the level of the models themselves.

Spending on data center construction in the United States exceeds $2 billion per month.

Financing a large GPU cluster or an AI-heavy data center is remarkably similar to financing a new clean energy plant. The first job, even before construction begins, is to line up long-term "offtake" agreements with customers who agree to buy the resources at a certain price for a certain length of time. But this is how companies at the application layer (who make up most of a data center's customers) end up locked into inflexible, long-term contracts. "Amazon Web Services might give you a big discount, but they will force you into a three-year contract and a minimum spend with them, because they need to know whether they have enough customers to build the next data center or not," says DCVC's Tamaseb.

But what if a cloud user such as a SaaS company doesn't end up selling enough monthly subscriptions to cover its contract? They're still on the hook for the unused capacity. The San Francisco Compute Company, where we first invested in 2025, believes that the market for GPU hours should be more flexible and that cloud customers shouldn't have to bear the entire risk of fixed-price, long-term contracts.

SF Compute's platform works like this: the company itself buys large chunks of time on clusters comprising thousands of GPUs. It uses its own software to virtualize these resources, then sells contracts to its own customers—but with the twist that these customers can exit the contract at any time and sell the unused resources on SF Compute's market for GPU time.

> "Amazon Web Services might give you a big discount, but they will force you into a three-year contract and a minimum spend with them, because they need to know whether they have enough customers to build the next data center or not."

Ali Tamaseb
General Partner, DCVC

SF Compute is conceptually similar to a commodities exchange: “It connects GPU providers to GPU hunters on an hourly basis,” as Tamaseb explains. That lets hunters work within a budget. “You can code it into your buy that you only want to execute this training task if the price of a GPU goes below $2.50 per hour per node. And then maybe three days later at 3:00 am, SF Compute finds that capacity and actually runs your code. So you can also hedge against the future.” (Interestingly, DCVC backs another company, ElectronX, that provides a nearly identical market-making service in the realm of electricity contracts. “ElectronX and SF Compute are basically the same company, but they’re working in two different domains,” Tamaseb says.)

To Tamaseb, the attraction of a platform that financializes GPU hours is that it opens up and smooths out access to a resource—distributed computing—that’s becoming so homogeneous and utility-like that it should work like plugging into a wall socket. Another bonus: SF Compute gives smaller companies a practical way to tap into the power of the latest AI models. “A two-person startup in a San Francisco Victorian can’t realistically sign a 5-year take or pay contract on a $100 million supercomputer,” the company says. “But they may be able to buy the month of liquidity” that someone else wants to sell back.

There’s another way to control the cost of scarce cloud resources: it’s to make them go further. Today’s vast wave of data center construction—which will cost $3 trillion globally over the next five years, according to Moody’s—is premised on the assumption that scale always wins in AI: that making frontier models like ChatGPT, Gemini, and Claude even bigger will automatically make them smarter and more capable. But what if it won’t? What if, in fact, general models are wasteful overkill for many of the automated services we’ll want in the future?

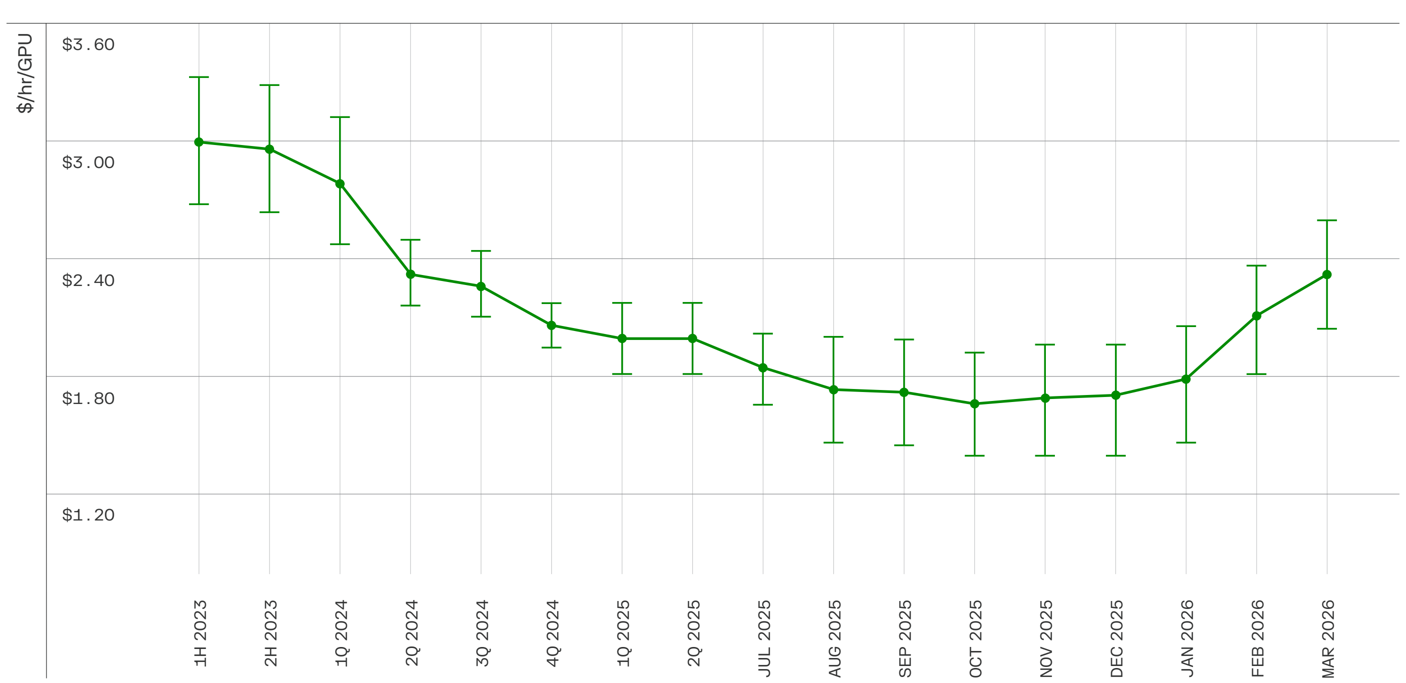
Those are the questions the founders of TypeSafe are asking. The San Francisco company, which we’ve backed since 2025, is making two major bets: First, that AI’s biggest benefits will come from automating more tasks done by humans, including online tasks. Second, that the path to economically viable AI automation runs not through ever-larger general models ceaselessly chatting away with humans, but through networks of smaller, so-called “composable” AI components that can be reliably selected, combined, and deployed by other software.

Figure 1.2.1

GPU rental pricing trends

In 2026, AI usage is driving a run on GPU hours at the hyperscalers, forcing the price of a 1-year contract upward.

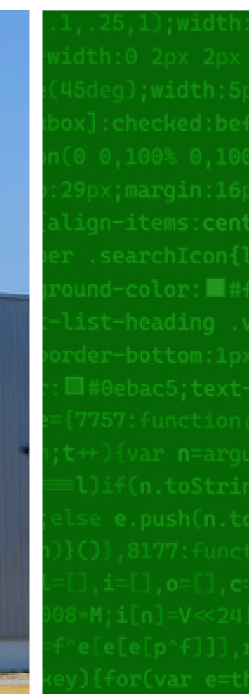
Source: SemiAnalysis.



> Projecting the trends of energy consumption forward for these data centers, it’s exponential growth, and it’s hugely destabilizing. The infrastructure layer of the AI stack might make those growth curves sub-exponential.”

Zachary Bogue
Co-Founder and Managing Partner, DCVC

↓ In 2025 data centers consumed roughly 5 percent of U.S. electrical generation.



Composable software modules are like LEGO bricks: they’re built to connect predictably. But they’re also designed to be machine-readable and machine-callable, with structured inputs and outputs. On an internet populated with composable systems reasoning over structured, typed data (not just natural language), all of the tasks that go into, say, planning a vacation could be more efficiently delegated and scheduled, avoiding expensive AI inference calls whenever possible. That means more work per watt, fewer GPU hours burned, and better utilization of cloud resources.

“The number one rule in machine learning is that you get what you optimize for,” says DCVC General Partner James Hardiman. “So what if, instead of training these models to understand human preferences, we optimized them for machine-to-machine communication? That’s what will allow us to use these tools properly for agentic workflows in the new AI economy. And once we fine-tune these models for that stuff, they’ll be way more performant. What TypeSafe is proposing is a refreshing take on the coming AI wave, not just a regurgitation of what everyone else is saying.”

The next leap in AI won't come from larger models, but from architectures that make intelligence cheaper and more machine-native.

Today's leading large language models are optimized for working directly with humans. With each new generation, these models grow larger and more expensive to build and run. But it's far from clear whether AI companies can unlock durable new industries or deliver promised productivity as long as most systems still operate in chatbot mode.

We think large language models are on their way to becoming infrastructure, not the product itself. Increasingly, they'll be used by other software rather than by humans directly. The next breakthroughs will come one layer down, at the level of architecture: how models are specialized, trained, and orchestrated—and how they operate inside an internet that has been retrofitted for machine-native workflows.

It's worth being precise about what today's frontier LLMs actually do. Models from Google, Anthropic, and OpenAI are trained to interact with us using natural language. But at a mechanistic level they're not doing anything remotely similar to human cognition. They're large probabilistic systems that use a vast set of neural network weights, established during their training phases based on multimodal data from numerous sources, to predict the next most likely token (word or part of a word) in response to a prompt.

Before the models are released, engineers use a process called reinforcement learning with human feedback (RLHF) to fine-tune them to be helpful to humans. They're optimized for *conversational fluency* and *alignment with human needs*. And it all works remarkably well—except for two fundamental problems.

The first is that the cost of training and running these models tends to scale linearly (or worse) with their fluency and capabilities. The longtime ideal in Silicon Valley—and a quality prized by investors—was that new software technologies would grow *cheaper* as they scaled. Frontier AI is doing the opposite. OpenAI's GPT-4 was widely reported to have a trillion or more parameters and required 13 trillion tokens to train; CEO Sam Altman said publicly that the process cost more than $100 million. Training its successor GPT-5, by contrast, required two full training runs over more than a year, each costing more than $500 million, according to reporting in the *Wall Street Journal*. Scientists at research firm Epoch AI found that LLM training costs have gone up an average of 240 percent per year since 2016, with most of the spending going to AI chips, server components, staff costs, and energy. And Anthropic CEO Dario Amodei has projected that by 2027, training LLMs could cost $10 billion to $100 billion per run. Unless the profits from AI begin to scale exponentially as well, it's hard to see how this level of spending growth will be sustainable.

However, there are ways to grapple with the training-cost problem without abandoning the LLM paradigm. When DCVC-backed Databricks acquired MosaicML ↳

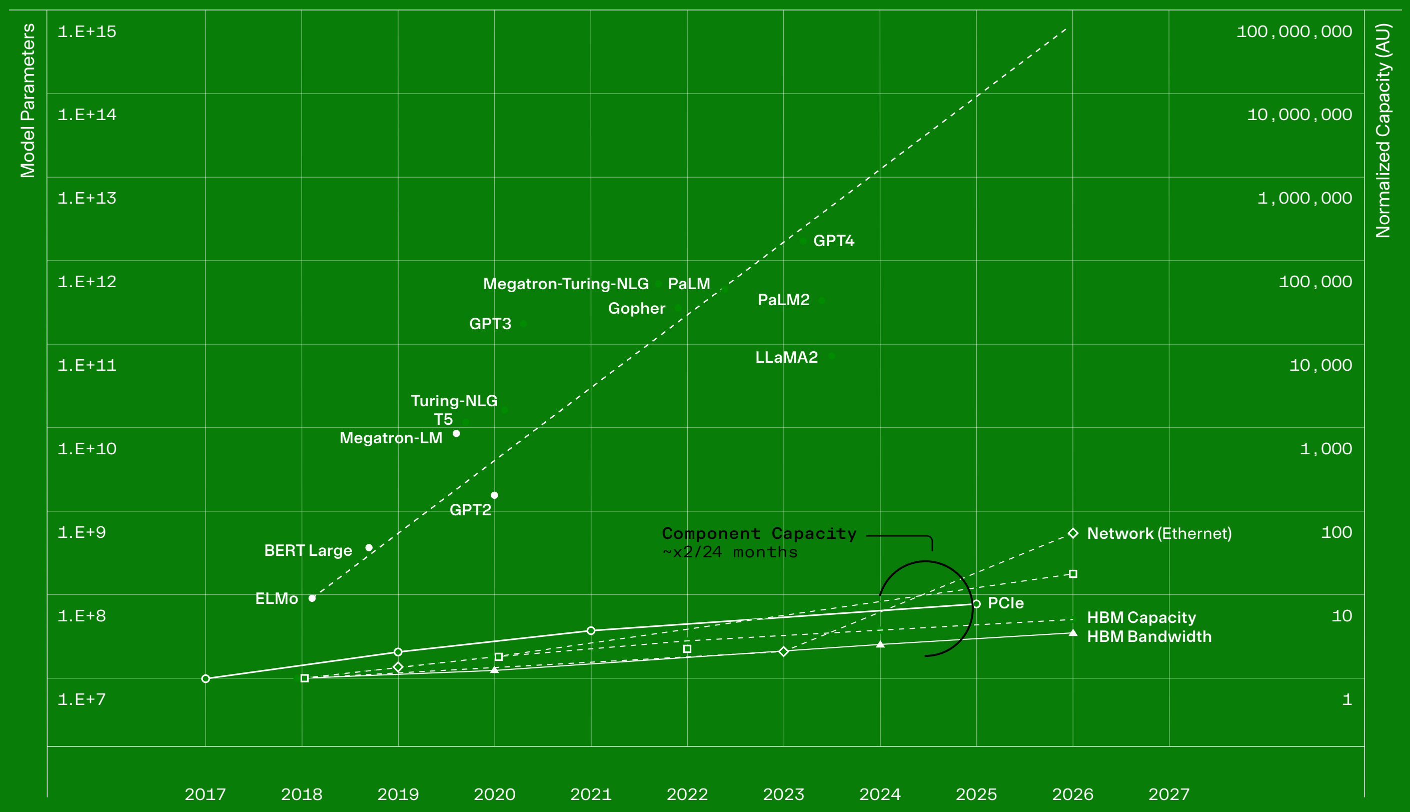
↓ **Naveen Rao,** right, CEO and Co-Founder of MosaicML and Unconventional, argues that the next generation of AI chips should mimic the human brain (*see page 29*).

Figure 1.3.1

Growth of training data for a collection of LLMs

The number of parameters used to train leading AI models is doubling roughly every 4 months—and is increasing far faster than system parameters such as network and peripheral (PCIe) speed and high-bandwidth memory (HBM) capacity and bandwidth.

Source: Open Compute Project.



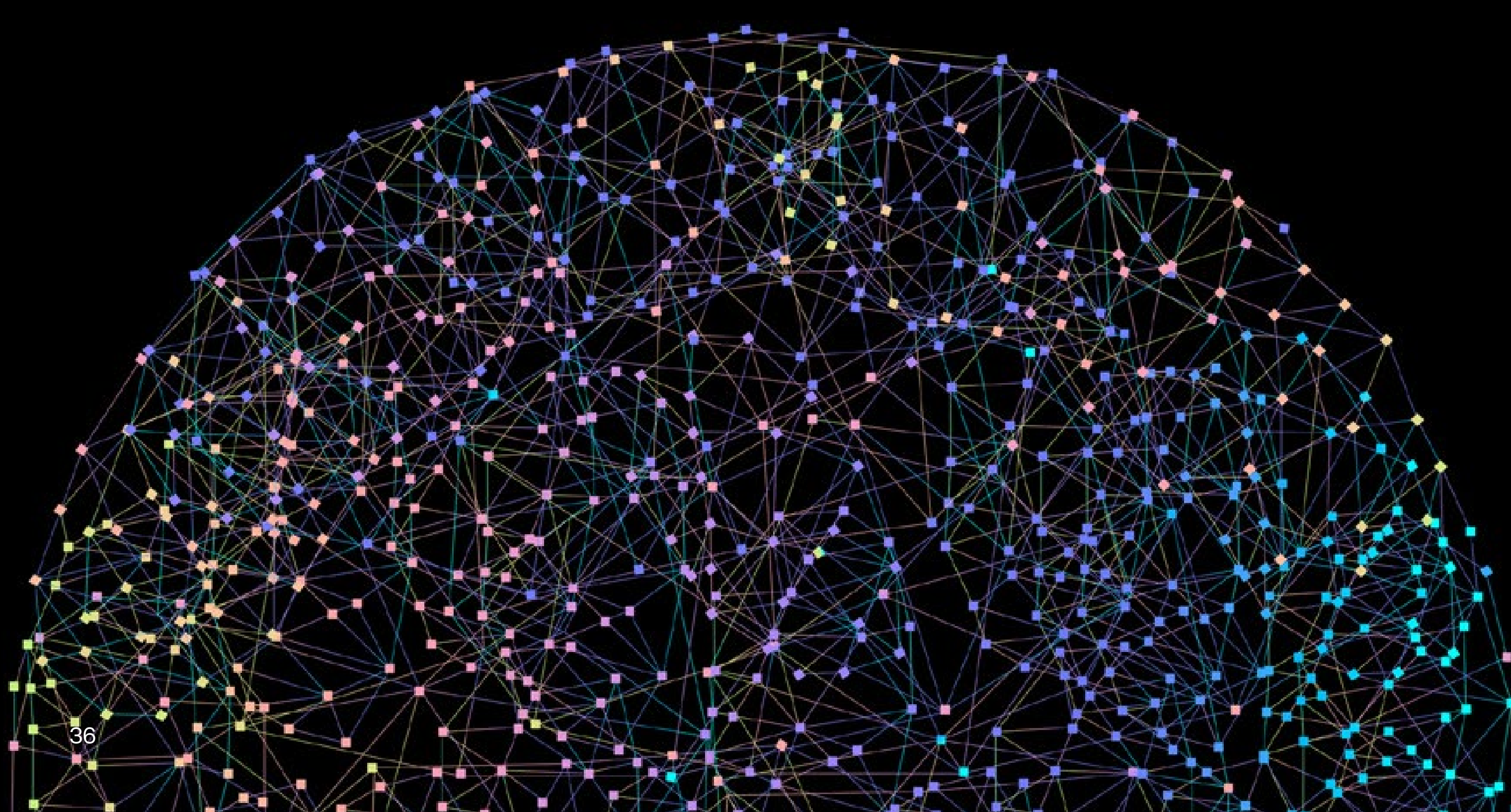
in 2023, it brought in a team of engineers working to drastically lower the costs to enterprises of training their own machine learning and generative AI applications. Using Databricks' Mosaic AI Model Training, customers can now tap into a managed system that's been optimized for training LLMs, allowing them to build useful models with fewer GPU hours. That helps democratize large-scale model training beyond the frontier labs.

We're also impressed by a company called Ceramic that's trying to make training cheaper by changing the fundamental training regime for transformer-based AI models. Transformer architectures rely on self-attention, which means that every word (or token) in a sequence compares itself to every other word. The number of pairwise comparisons goes up as the square of the sequence length, so most training pipelines chunk text into very short windows of 2,000 to 4,000 words. This approach is inefficient. It takes a long time, and many self-attention passes, before a model gets signals about relationships or structure that stretch across many chunks.

Ceramic's training stack has an advanced mathematical foundation that makes it feasible to train models using much longer sequences or "context windows" of up to 96,000 tokens (roughly equivalent to tens of thousands of words) at once. The company also reorders training data so that batches are aligned by topic. Longer context matters because it helps models reason and remember across more complex information. Just as important, it speeds up training by surfacing high-level structure with fewer passes. (Each step may be more expensive, but fewer steps are needed overall.) Ceramic says its platform can train models 2.5x faster than open-source training stacks. "This is just one widget, but it's going to be part of the re-architecting of the internet for machines," DCVC's Hardiman says.

One more way to scale up AI affordably and advance faster toward artificial general intelligence and even superintelligence may be to mimic the open-ended processes of Darwinian and cultural evolution that gave rise to human intelligence. That's the strategy at Recursive Superintelligence, a company we backed starting this spring. Their team includes researchers who have developed open-ended algorithms and self-improving coding agents for OpenAI, Deep Mind, Google Brain, and Meta, and they aim to develop AI that recursively improves itself—gradually replacing all the hand-designed methods in machine learning with AI-driven processes.

But lowering the cost of training AI is only half of the problem. Another big limitation of today's leading LLMs is that they're arguably mis-optimized for the next era of value creation in computing. A truly helpful AI would be capable of taking care of real-world tasks automatically—which would mean interacting not just with humans, or with online services designed for humans, but with other software entities directly. "AIs shouldn't be scrolling through Kayak to book a flight for you," says Hardiman. "That is an abomination. When you take a model that was trained for human preferences and you apply it to other things like agentic workflows, you can imagine why there's a mismatch."



One solution here is to design smaller, less wasteful, more task-focused AI models. This is the kind of work TypeSafe is engaged in, as we noted in Opportunity 1.2. They're designing AI systems to be natively composable (available to be mixed and matched by other software) and to respect the rigid software types that show computer systems what they're allowed to do (a "type" in programming terms might be Integer, String, Date, UserID, or ShippingAddress).

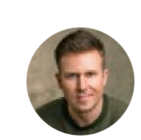
In a world of composable AI, cognitive tasks would not be assigned to general LLMs—which are costly to run and never come up with the same answer twice—but would instead be distributed to agents that each do one thing well, with clearly specified inputs and outputs. As TypeSafe puts it, "Automation demands systems that can consistently choose what to do and when to do it"—think scheduling, monitoring, or transaction processing. It's a more efficient approach to automation, prizing structure, predictability, and correctness over creativity, synthesis, or ambiguity—and reserving expensive AI inference for the few moments when it's actually needed. Early projects like OpenClaw, an open-source AI assistant that runs on a home computer and communicates via messaging apps like WhatsApp, Telegram, Discord, Slack, and Teams, are already showing how this might play out in the real world (with consequences that are both impressive and worrying).

But there's one major speed bump on the way to a future made up of composable AIs, and that's the current-day internet, which was designed largely for humans, not agents. For the biggest payoffs in a world of agentic workflow, we need a new infrastructure for the web's "second user," the machines themselves. Parallel is one of the companies working on this larger vision. It's building web tools such as an agent-friendly search engine that can act as a facilitating layer between AI and existing internet databases. In agent-directed search, for example, there's no inherent reason to limit the inputs to keywords or limit the outputs to URLs. Parallel Search turns an agent's keywords into a declarative, semantic objective. It then hands back results and excerpts that are ranked and compressed so that only the most relevant tokens go into the agent's context window. Parallel also builds its own web agents, such as a Task API that helps the company's customers make complex research tasks more easily repeatable. The overall idea is to get the web ready for the day when there are far more AI users than human users. As the company puts it in its manifesto, "The choice is binary: we build the web for its second user, or it fractures beyond repair."

"There's search for humans, and there needs to be search for agents," says Hardiman. "It's just one widget—it's just one part of the re-architecting of the internet. But we're going need a Parallel for everything." Taken together, such efforts point toward an AI ecosystem that does less brute-force thinking and more structured work—spreading intelligence across systems rather than concentrating it in ever-larger, ever-more-expensive models.

"

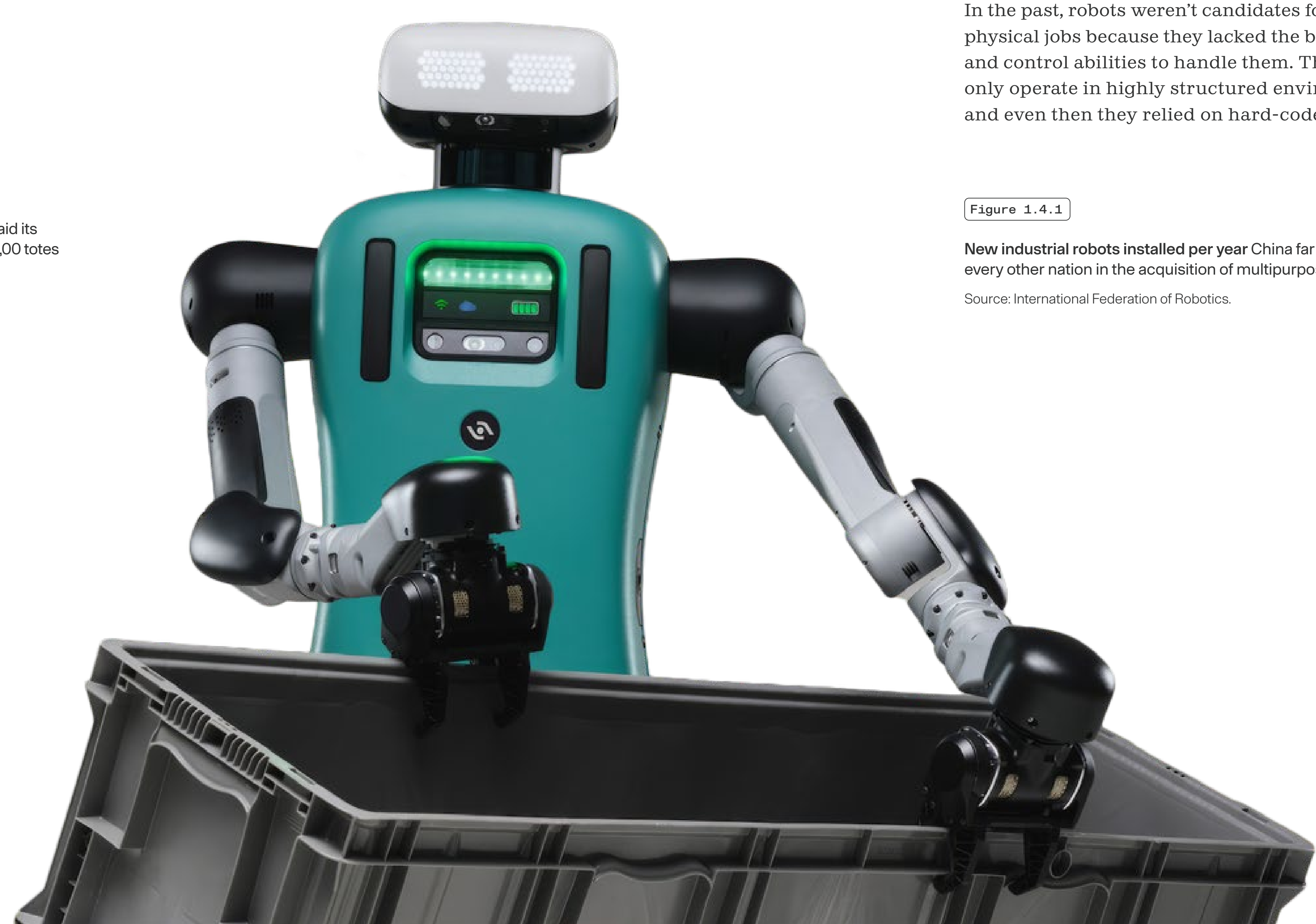
There's search for humans, and there needs to be search for agents. It's just one part of the re-architecting of the internet."



James Hardiman
General Partner, DCVC

Physical AI, embodied in robots, will lower costs across the economy.

There aren't enough human laborers to do all the demanding and physical jobs that go into maintaining the existing U.S. economy, let alone an ideal future economy marked by the reshoring of more manufacturing sectors. Yet, so far, robotic systems haven't had the dynamic sense-and-adjust capabilities these jobs require. As companies develop better robotic sensing and actuation hardware and smarter control software, there's a huge opportunity to fill labor gaps and perform many tasks faster and more safely.

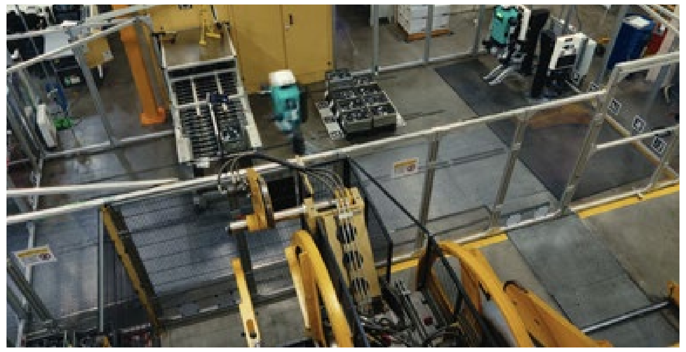


→ In November 2025 **Agility Robotics** said its Digit robot had moved more than 100,00 totes at a GXO Logistics facility in Georgia.

Roofing. Painting. Carpentry. Electrical work. Plumbing. Fruit picking. Groundskeeping. Hotel housekeeping. Building maintenance. All of these activities take place in unstructured environments, and they require dexterous manipulation of irregular objects and continuous adjustment to changing conditions. That's why they've long been resistant to automation.

Yet that's beginning to change—as we're already seeing in the field of logistics and material handling. Warehouses used to be cited as hard-to-automate environments, since they have an ever-changing inventory of objects of varying size and weight, which often must be picked from cluttered bins. But what was considered impossible two decades ago is now routine, as Amazon and other companies deploy multiple varieties of robots that speed up fulfillment, from simple drive units that move shelving around to sophisticated pick-and-place robot arms. (Amazon said in 2025 that it had deployed 1 million fulfillment robots globally; it employs roughly 1.5 million human warehouse workers.)

In the past, robots weren't candidates for many physical jobs because they lacked the basic perceptual and control abilities to handle them. They could only operate in highly structured environments, and even then they relied on hard-coded rules and pre-programmed motion trajectories. But the same machine learning techniques that transformed software are now transforming hardware. Advances in visual models allow robots to identify irregular objects and make sense of cluttered scenes. Reinforcement learning and imitation learning mean we don't have to tell robots exactly how to move, but can instead let them learn control policies that adjust to shifting conditions. Emerging "vision-language-action" models trained on internet-scale data allow robots to integrate camera images with natural-language instructions to generate their own actions.

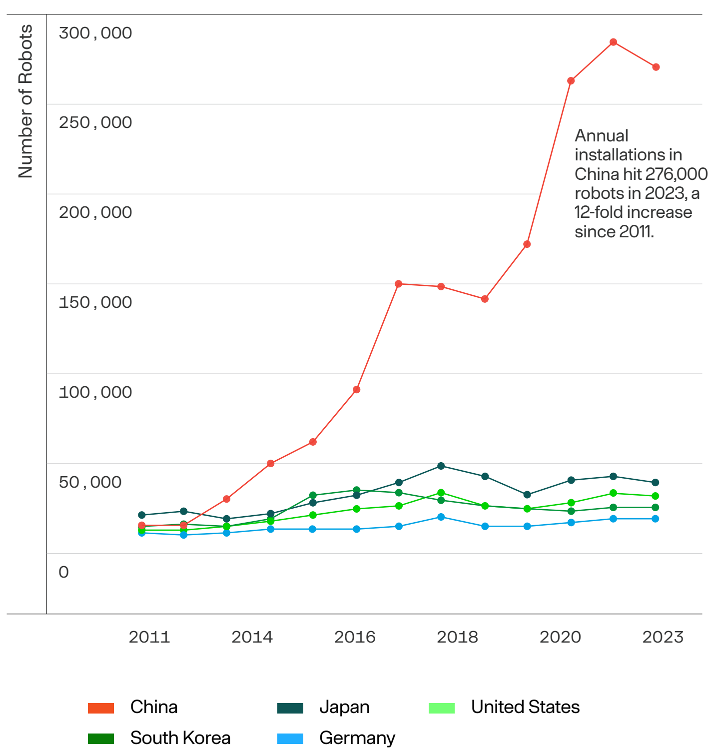


↑ **Agility Robotics'** Digit humanoid robots help keep parts moving through a factory.

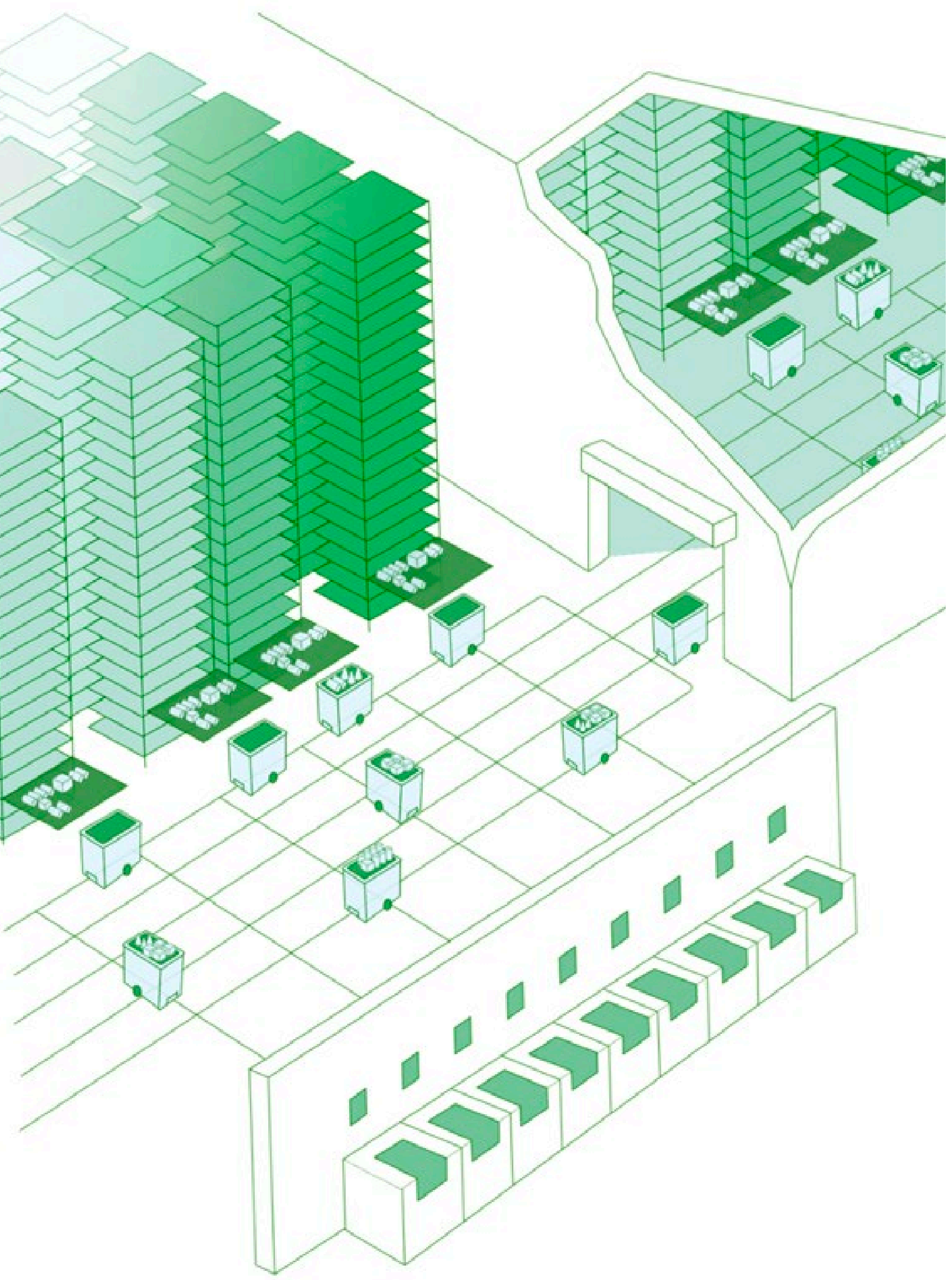
Figure 1.4.1

New industrial robots installed per year China far outstrips every other nation in the acquisition of multipurpose robots.

Source: International Federation of Robotics.



In short, robots are becoming less "robotic" and more like flexible agents who can respond to their environments. That doesn't mean they'll be piloting fighter jets or performing solo neurosurgery anytime soon—but it's already clear that robots will show up in a variety of simpler domains that offer a modicum of predictability and structure and a lower penalty for error. We see this happening at three tiers of complexity.



↖ **Fulfil Robotics'** ShopBots collect grocery items from vertical storage trays and deliver them to pickup portals.

①

First off, robots can help with repetitive jobs that involve low cognitive load, where they're supplementing human muscle and hand-eye coordination rather than problem-solving.

We've invested in several companies in this category. Fulfil Robotics, for example, makes squat, rolling robots called ShopBots that automate the multitude of pick-and-pack actions that go into preparing an online grocery order for delivery. "Fulfil can replicate an entire supermarket, with gantries holding the food and little R2-D2-like robots picking and packing," says DCVC General Partner Alan Cohen. "They know you do the dry goods first, then you go to the refrigerated case. They know you put the heavy stuff in the bottom of the bag. And the timing is so much faster that they can take an operation that is often unprofitable and turn it into something profitable."

Slip Robotics is automating another familiar environment: the warehouse loading dock. Traditionally, workers driving forklifts move goods into or out of semitrailers pallet by pallet. But that system results in forklift congestion around the dock, not to mention accident risk. Slip has two solutions: SlipBot, an automated pallet on wheels that carries loads into trailer and stays with the freight, and SlipLift, which shuttles heavier freight loads in and then scoots back out—like an Amazon drive unit, but for moving a whole truckload of goods at a time.

Slip's robots align to the trailer opening and navigate without beacons, markers, or other aids; onboard sensors and software allow them to monitor their surroundings and stop if an obstacle appears. A warehouse still needs human forklift operators to place goods on the SlipBots or SlipLifts manually, but when robots take over the most hazardous and delicate part—trailer loading—it can speed up operations dramatically. One Slip customer, automotive component manufacturer Valeo, was able to reduce loading time from 30+ minutes per trailer to just 5 minutes.

②

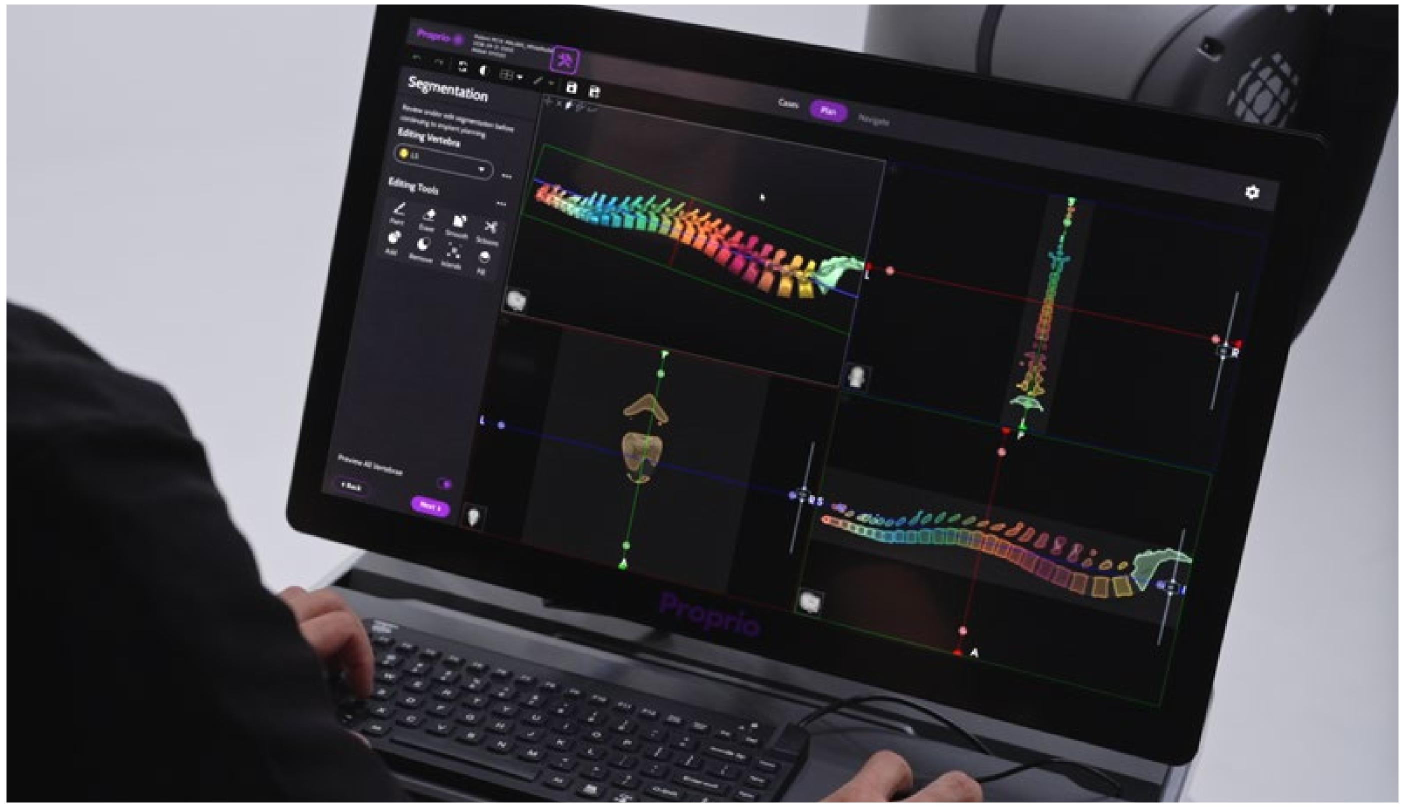
At the next tier of complexity, robots aren't replacing humans but helping to sharpen their skills.

There are many work contexts, such as surgery, that still call for human-level perceptual intelligence, motor intelligence, and task intelligence. But physical AI can magnify these abilities, as companies like Remedy Robotics are showing. Remedy uses machine learning to help endovascular surgeons guide catheters through the human body—for example, to conduct an angiogram or reach a life-threatening clot in the brain. The AI aligns live 2D X-ray images from an operating-room fluoroscope with 3D images from a preoperative CT or MRI scan. That helps surgeons stay oriented as they push the springy, elastic catheters around tight corners, all the way from the groin to the brain.

↗ **Slip Robotics'** SlipBot and SlipLift autonomous loading robots reduce semitrailer loading and unloading times by 6x to 10x, while increasing safety.

Combined with a robotic system for manipulating the catheter, the Remedy technology will make it possible for surgeons to carry out both routine and emergency procedures remotely, helping patients in far-flung hospitals that may lack a trained endovascular surgeon. Last year, a team at St. Michael's Hospital in Toronto used Remedy's robot and imaging system to perform history's first fully remote cerebral angiograms. "Surgical robotics has been such a slow-to-innovate and hardware-focused field," says Remedy co-founder and CEO David Bell. "The machine learning component is really what makes this possible."

→ **Remedy Robotics** helps surgeons perform remote endovascular interventions.



↑ **Proprio** uses advanced sensors and AI to build 3D models that help surgeons plan and execute spinal surgery, significantly improving outcomes for patients.

Spinal surgery is another discipline that requires enormous dexterity, experience, and stamina. In many procedures, surgeons need to place pedicle screws to anchor implants that stabilize the vertebrae, and that's historically been a freehand technique, with lots of blood and tissue in the way and only the occasional fluoroscopic snapshot as a guide. As a result, outcomes are variable and surgical revision rates are high.

Proprio, where we first invested in 2020, places depth-sensitive lightfield cameras over the surgical area. Like Remedy, it uses machine-learning models to align the 3D data with preoperative maps and show surgeons in real time whether they've placed implants in the optimal positions and angles. Cohen calls Proprio's AI "an ever-watchful assistant, performing continuous measurements and analysis that would be impossible for a human to do in real time."

③

In the most complex environments, such as existing workplaces built for humans, robots will need to be not just autonomous and precise, but versatile; breadth will be just as important as brilliance. This is where general-purpose humanoid robots come into the picture.

DCVC-backed Agility Robotics makes one of the only commercially deployed humanoid robots being manufactured in large numbers—it's called Digit. In February the Canadian wing of Toyota—a multinational famous for manufacturing excellence—signed an agreement with Agility to deploy the robot in manufacturing, supply chain, and logistics roles. Our intuition is that a humanoid robot becomes economically viable the moment it can perform enough distinct tasks to amortize its capital cost across many such jobs. "Would I pay $20,000 a year for a robot that can only do the laundry? Probably not," says DCVC's Hardiman. "But if it did the laundry, folded it, put it away, did the dishes, picked up after itself, and pulled the weeds in the backyard, then yeah, I would definitely pay $20,000 a year. Suddenly, if it can do all that, the market is enormous."

One of Digit's selling points is that it's multi-purpose and multi-skilled. Inside a factory, warehouse, or fulfillment center, it can move totes to or from autonomous mobile drive units, load and unload conveyors and flow racks, load and unload pallets, or stack, unstack, or nest tote boxes—all in the same spaces where humans are working, without special barriers or protections. For humans, this kind of work is obviously repetitive and monotonous, and when Digit takes over, workers can be redeployed to higher-value tasks. Agility says industrial customers who buy fleets of Digits can recoup their investment in as little as two years.

No matter how quickly businesses adopt robots for their loading docks, farm fields, operating rooms, or warehouses, the broader trend is unmistakable: Intelligence is leaving the data center and entering the physical world. The American industrial renaissance we described in our 2025 report will depend not just on building more factories, but on embedding more intelligence within them—so that humans can move into more stimulating, higher-paying roles. The companies that win in robotics won't simply build machines; they'll build data loops, collecting and refining proprietary datasets that make physical AI models more capable with every deployment. And if we pair smarter robots with the reshoring of critical components—from sensors, actuators, and batteries to edge computing and control software—this can all happen here in the United States.

"

Would I pay $20,000 a year for a robot that can only do the laundry? Probably not, but if it did the laundry, folded it, put it away, did the dishes, picked up after itself, and pulled the weeds in the backyard, then yeah, I would definitely pay $20,000 a year. Suddenly, if it can do all that, the market is enormous."

James Hardiman
General Partner, DCVC

CHAPTER 2

Quantum computing is evolving from a lab experiment into an industrial stack.

“Since backing the very first commercial quantum computer company in 2003, I’ve been optimistic about the field’s massive potential impact on human progress and overall social capability, but pessimistic about its commercial time frames,” says DCVC Co-Founder and Managing Partner Matt Ocko. “The DCVC team has continued to be a leader in quantum computing and the overall quantum technology space, making judicious investments over the last 20 years in a significant majority of its leading and commercially successful companies. But now we are seeing a compression of both the timeframe and the risk along the path to business-relevant quantum computing.”

Specifically, we are seeing:

- New qubit modalities with higher factory yields, longer coherence times, and higher fidelity
- Practical and powerful error correction codes
- Benchmarking frameworks that show which modalities are becoming credible engineering ventures
- Industrial applications with real customers and near-term revenue
- Long-horizon upside in areas like drug discovery and energy.

Individual physical qubits are still too noisy and fragile to support meaningful computation. But we’re learning how to manufacture them cheaply enough to make redundancy affordable, and how to yoke them together into logical qubits where math can smooth out the noise. That’s the “silicon moment” for quantum computing—and it’s the difference between a physics experiment in a university clean room and an industrial platform that can function as part of a national computing strategy.

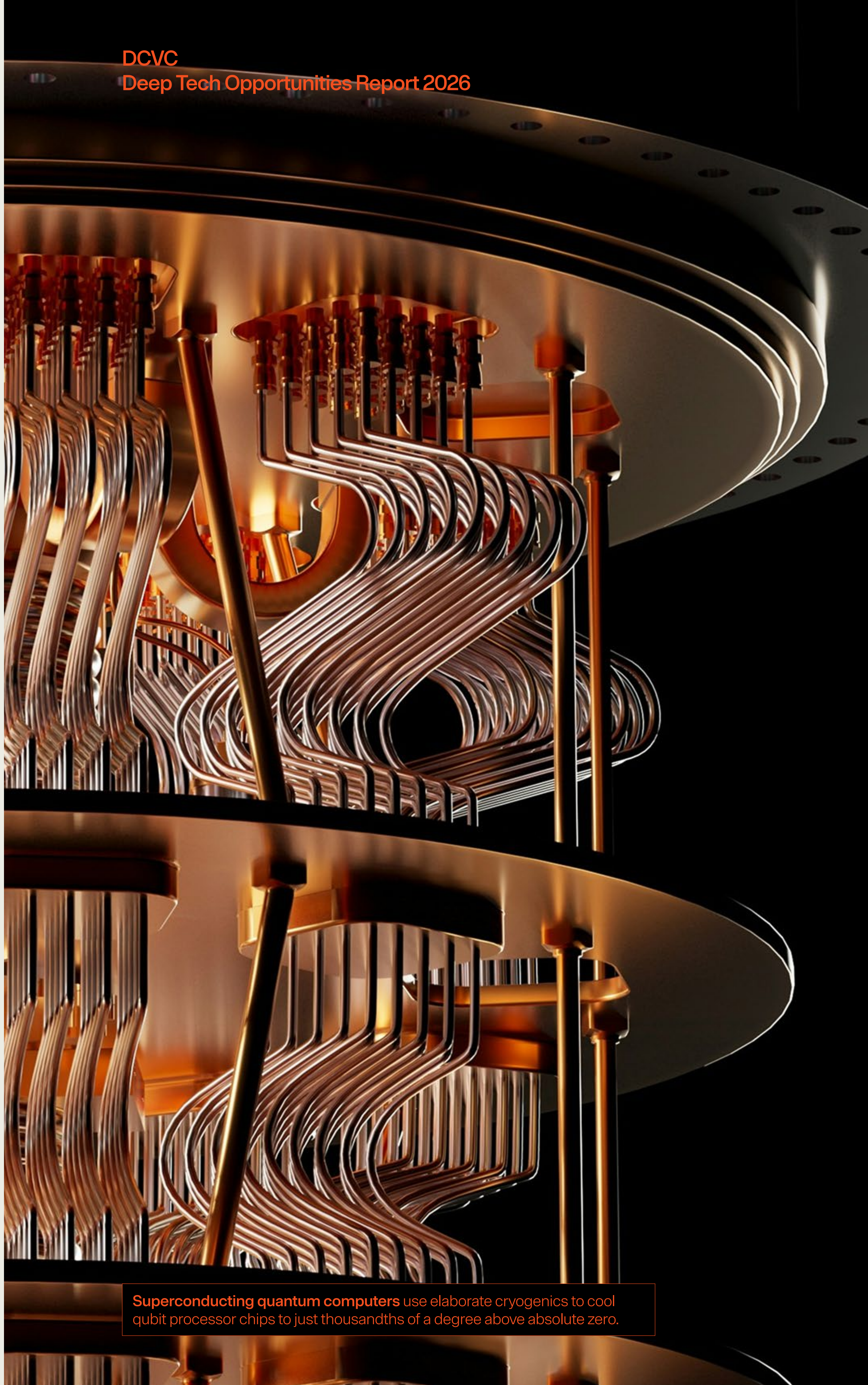
OPPORTUNITIES

COMPANIES

Atom Computing, Iceberg Quantum, Logiqal, Mesa Quantum, Q-CTRL, Quantum Motion, Rigetti Computing

VOICES

James Hardiman, Dr. Prineha Narang, Matt Ocko



Superconducting quantum computers use elaborate cryogenics to cool qubit processor chips to just thousandths of a degree above absolute zero.

CHAPTER 2

The effort to build reliable hardware to hold qubits is not a winner-take-all race.

Quantum hardware has finally crossed the line from “perpetual research project” to having a credible 7-to-10-year runway to commercialization, and that changes the capital calculus. Across modalities, qubit coherence time, fidelity, and fabrication consistency are ticking up, error rates are moving down, and government benchmarks have emerged to make overall progress measurable. Now that qubits are beginning to look more like manufacturable components—noisy ones, but good enough to assemble into reliable logical machines—there are many opportunities for deep tech venture investors to help.

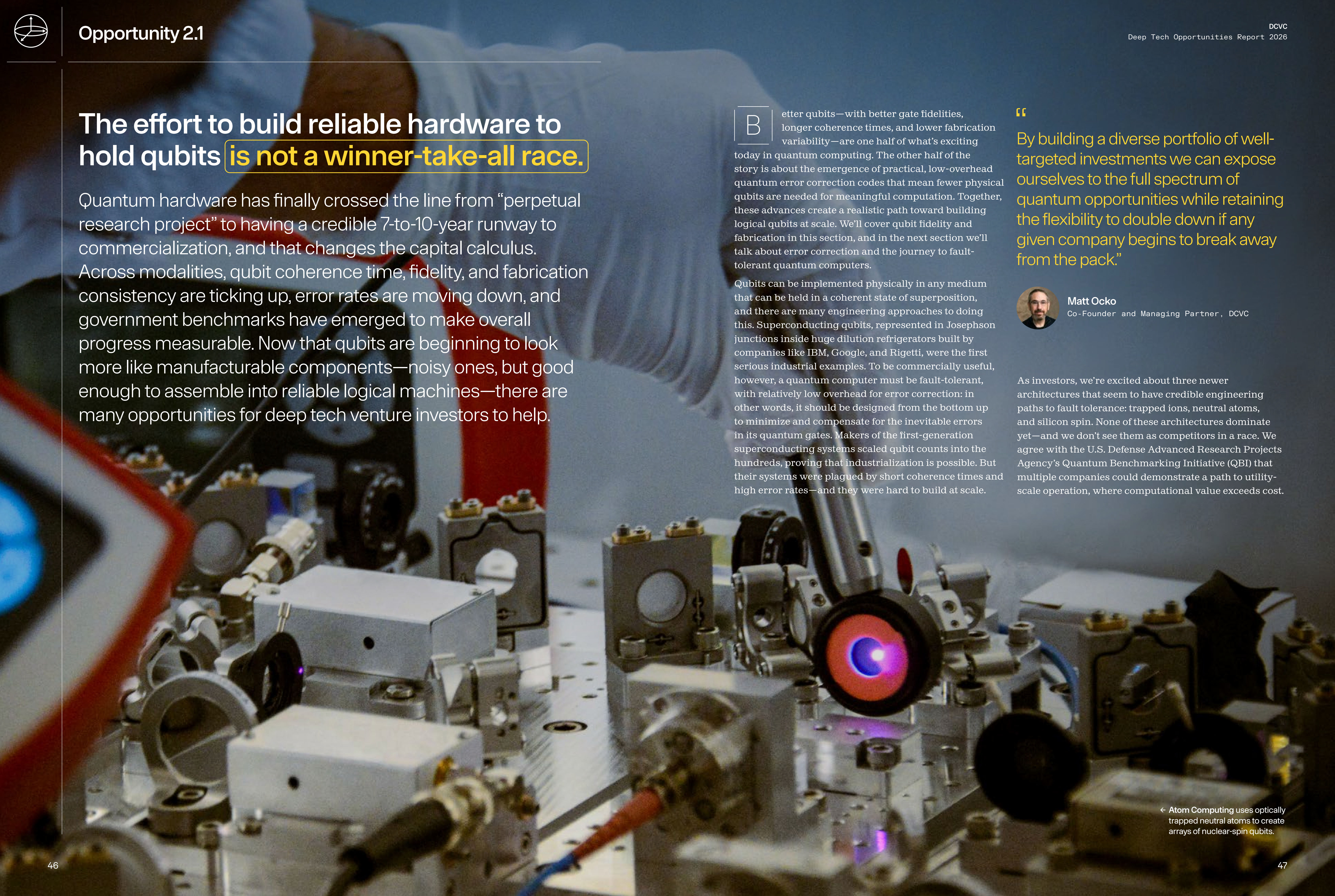
Better qubits—with better gate fidelities, longer coherence times, and lower fabrication variability—are one half of what’s exciting today in quantum computing. The other half of the story is about the emergence of practical, low-overhead quantum error correction codes that mean fewer physical qubits are needed for meaningful computation. Together, these advances create a realistic path toward building logical qubits at scale. We’ll cover qubit fidelity and fabrication in this section, and in the next section we’ll talk about error correction and the journey to fault-tolerant quantum computers.

Qubits can be implemented physically in any medium that can be held in a coherent state of superposition, and there are many engineering approaches to doing this. Superconducting qubits, represented in Josephson junctions inside huge dilution refrigerators built by companies like IBM, Google, and Rigetti, were the first serious industrial examples. To be commercially useful, however, a quantum computer must be fault-tolerant, with relatively low overhead for error correction: in other words, it should be designed from the bottom up to minimize and compensate for the inevitable errors in its quantum gates. Makers of the first-generation superconducting systems scaled qubit counts into the hundreds, proving that industrialization is possible. But their systems were plagued by short coherence times and high error rates—and they were hard to build at scale.

> “By building a diverse portfolio of well-targeted investments we can expose ourselves to the full spectrum of quantum opportunities while retaining the flexibility to double down if any given company begins to break away from the pack.”

Matt Ocko
Co-Founder and Managing Partner, DCVC

As investors, we’re excited about three newer architectures that seem to have credible engineering paths to fault tolerance: trapped ions, neutral atoms, and silicon spin. None of these architectures dominate yet—and we don’t see them as competitors in a race. We agree with the U.S. Defense Advanced Research Projects Agency’s Quantum Benchmarking Initiative (QBI) that multiple companies could demonstrate a path to utility-scale operation, where computational value exceeds cost.



← **Atom Computing** uses optically trapped neutral atoms to create arrays of nuclear-spin qubits.

Dr. Prineha Narang is an Operating Partner at DCVC, Howard Reiss Chair in Physical Sciences at UCLA, and head of a lab working at the intersection of quantum physics, quantum photonics, and quantum information science. She had close interactions with the DARPA program managers overseeing QBI. She notes that multiple companies and qubit technology approaches—including trapped ions, neutral atoms, silicon spin, superconducting processors, and photonic systems—have been selected to advance to "Stage B" of the initiative, where they've been judged to have an R&D plan capable of realizing a utility-scale quantum computer by 2033. "The program was not intended to pick winners and losers—just to say who has a credible path to the fault-tolerance goalpost," Narang emphasizes.

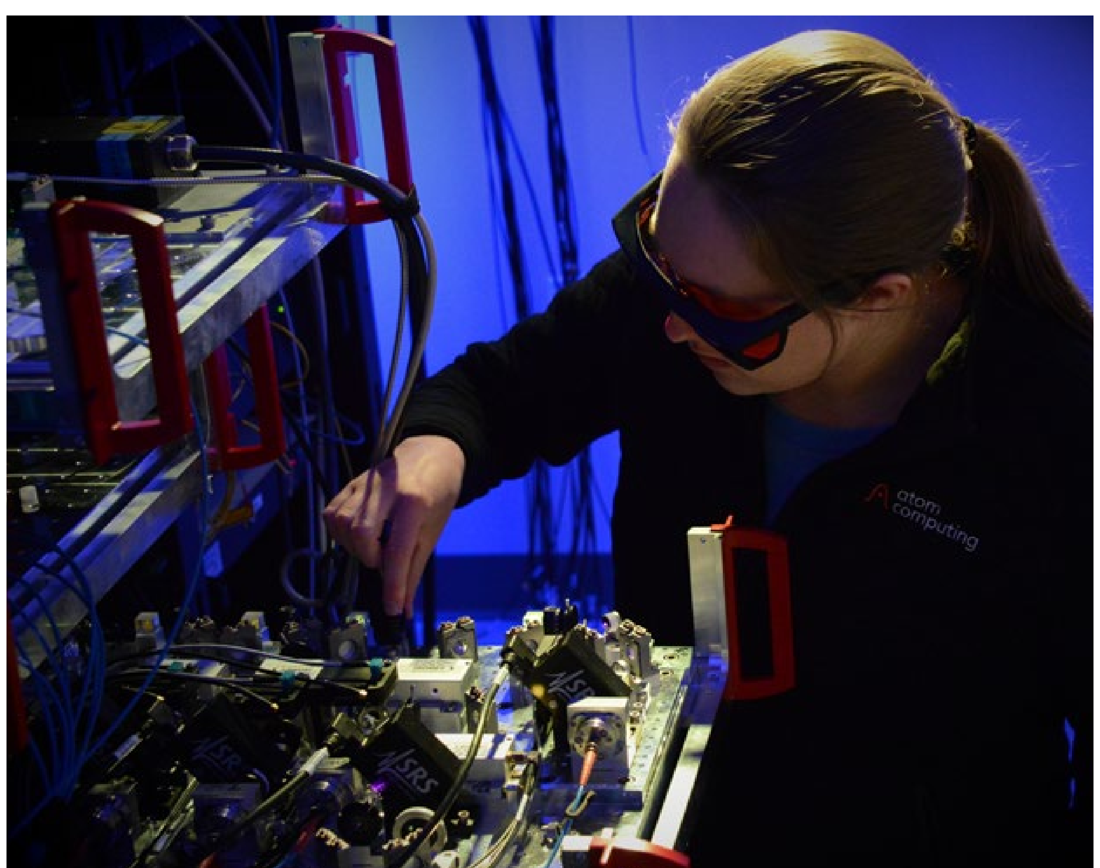
Trapped-ion systems: Fidelity first

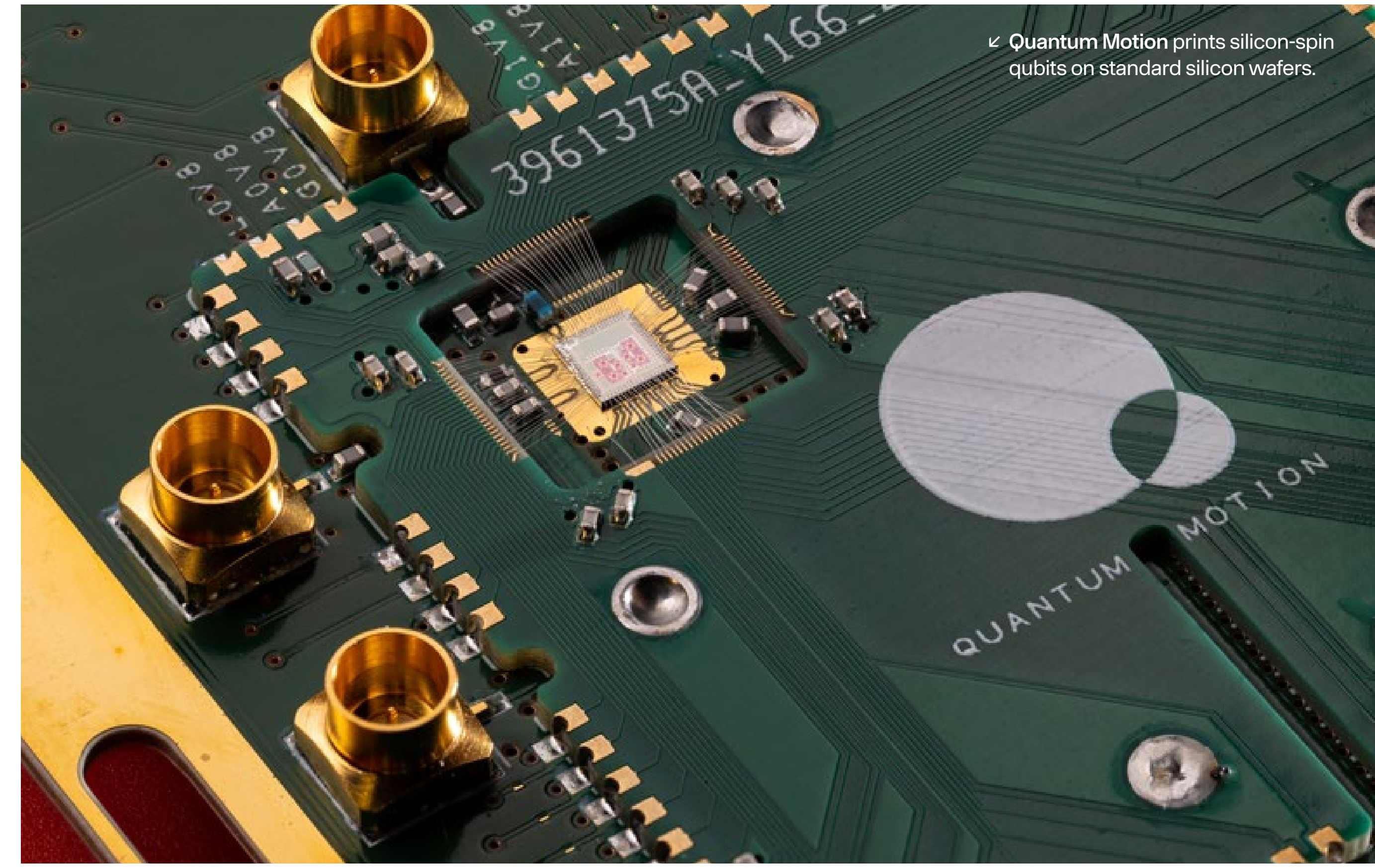
We know what "good qubits" look like: they're controllable, with long coherence times and high fidelity. In the trapped-ion approach, qubits are encoded in the internal atomic states of charged atoms confined in an electromagnetic field. Because they're suspended in an ultra-high vacuum, where they're shielded from environmental noise and material defects, these qubits have coherence times measured in seconds or minutes, as opposed to microseconds or milliseconds. And because all the qubits are physically identical, and controlled by precisely tuned microwave or laser interactions, there's little variability or material disorder that could throw off two-qubit quantum logic gate operations.

UK-based Universal Quantum is one of the leaders in this technology. The company is systematically solving the problems with laser stability and optical alignment that have been the main pain points as trapped-ion systems have scaled up. The company uses a combination of global microwave signals and locally generated fields to move and address individual qubits and reduce laser complexity. While the ions themselves are laser-cooled to microkelvin temperatures, the apparatus runs at a relatively balmy 70 K (-203 C), meaning it doesn't need to be placed inside a dilution refrigerator. Just as important, Universal Quantum has found a way to fit its qubit arrays into microfabricated modules on silicon, which it integrates with classical control chips and digital-analog converters to help shuttle information into and out of the qubits.

The company already has a customer: it's building a 100-qubit system for the German Aerospace Center (DLR) that could eventually be scaled to millions of qubits, thanks to its repeatable modules. The more physical qubits are available, and the lower their error rate, the more logical qubits can be assembled and the lower the probability of logical qubit failure. That sounds to us like the beginnings of a classic industrial scaling curve.

↘ Control electronics and software are key to performing error correction and forming logical qubits on **Atom Computing's** quantum computing platform.





↙ **Quantum Motion** prints silicon-spin qubits on standard silicon wafers.

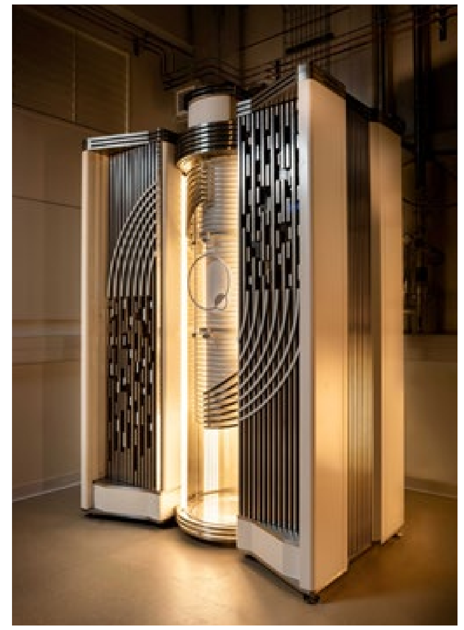
Neutral-atom systems: Reconfigurability first

Architecture and layout can be just as important as device physics. In neutral-atom quantum systems, qubits are encoded in the atomic energy levels or nuclear spin of electrically neutral atoms of alkali or alkaline-earth elements such as rubidium or ytterbium. When excited to high-energy "Rydberg states," neutral atoms can become entangled with other nearby Rydberg atoms, allowing the formation of quantum logic gates. But here's the key: because they're trapped by individually controllable optical tweezers (laser beams), neutral atoms can be dynamically shuffled to new positions, which means they can be reconfigured on the fly to specify which atoms form gates. The kind of connectivity between gates that's baked in with lithographically fabricated silicon systems thus becomes software-defined in these arrays. This means, among other things, that circuit layouts can be optimized for specific problems, and that lost qubits can be detected and reloaded.

We're intrigued by and/or already invested in three companies building neutral-atom quantum computers. One is Atom Computing, where we first invested in 2018, and doubled down in 2026. The Boulder, Colo.-based company offers a commercially available quantum computer called AC1000 that ships with 1,200 physical qubits that can be combined into 50 or more logical qubits. The physical qubits are encoded in atoms of ytterbium-171, whose high fidelity and long coherence times make it easier to combine them into reliable logical qubits. In 2024 Atom formed a partnership with Microsoft to strive for what the tech giant calls "Level 2" quantum computing, where logical qubit error rates are low enough to complete complex computations. Atom is just one of just 11 companies that have graduated to QBI's Stage B. It will also benefit from a $100 million equity investment from the U.S. Department of Commerce, part of a package of nine investments in utility-scale quantum computing announced in May that materially involves four DCVC portfolio companies.

The other neutral-atom quantum builder we're backing is Logiqal. (Yes, all of the C's in the quantum industry are qollapsing into Q's.) Logiqal, like Atom, is using ytterbium for its qubits; it's focused on building scalable systems from manufacturable sub-modules controlled by silicon optoelectronics.

Recent studies have shown that cryptographically relevant instances of Shor's algorithm for finding the prime factors of an integer could be executed on a neutral-atom quantum computer with as few as ~10,000 qubits. "For a long time, neutral-atom was yet another technology that was a lab curiosity," says Narang. "People were doing Nobel Prize-type experiments, but nobody was really thinking of it as how quantum computing would be implemented using these systems. But now it's a technology where—between Atom Computing, Logiqal, and others—there's just insane amounts of traction, and it seems like they'll be very good at solving all kinds of encryption problems."



Quantum Motion combined its Quantum Processing Unit, manufactured using standard silicon CMOS chip fabrication processes, with UI and control elements to build a data-center-friendly full-stack quantum computer.

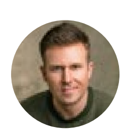
Silicon spin systems: Manufacturability first

One of the biggest signs that quantum computing is becoming real is that it has entered the semiconductor supply chain. In silicon-spin quantum systems, a qubit is represented in the spin of a quantum dot—a single electron confined in all three spatial dimensions. The nanoscale potential wells that trap these electrons can be defined lithographically. That means silicon-spin qubits can be fabricated using machines and processes found throughout the global fab ecosystem; can theoretically be created in extreme numbers, at extreme density; and can easily be integrated with classical control electronics.

"With quantum dots I don't have a real estate problem on a chip because I can get millions of them on there," Narang says. "Moving toward that required some insane engineering challenges to be solved. But that technical risk is retired now, so that's why there are three silicon-spin qubit companies that have moved to the next stage of DARPA's QBI."

"

> The most advanced silicon-spin companies have the promise of being able to leapfrog the superconducting, trapped-ion, and photonic qubit companies."



James Hardiman
General Partner, DCVC

One of them is U.K.-based Quantum Motion, a company we backed starting this year. It starts with standard 300-millimeter silicon wafers and uses the standard chip fabrication process to print quantum processing units (QPUs) with all the compute, readout, and control elements needed to perform useful operations. One of the more insane challenges the company had to solve was that the qubit layer in silicon QPUs operates at very low temperature, just tens of thousandths of a degree Kelvin, while the cryogenic control electronics naturally operate at temperatures above 1–4 Kelvin and the classical control systems operate at room temperature—so much higher that the leaking heat can make it difficult to reliability initialize qubits.

Quantum Motion found ways to keep the temperature gradient from becoming a showstopping constraint, principally by working with pairs of qubits and using a technique called "algorithmic cooling" to transfer entropy from one to the other (in effect, using logic operations to create higher-purity qubits than the temperature would normally allow).

For the dense wiring and control electronics needed to make these initialization protocols feasible, Quantum Motion also depended on advanced fabrication and packaging techniques, an area where silicon foundries such as GlobalFoundries and Imec are making huge strides.

In 2025 Quantum Motion delivered what it calls the world's first full-stack silicon spin quantum computer to the U.K.'s National Quantum Computing Centre. It has three server racks positioned next to a dilution refrigerator housing the quantum chips, and could fit nicely into any data center. The company's goal is to scale up to millions of qubits per machine, with qubit fidelities high enough to support the latest error correction protocols. Narang calls Quantum Motion "a very exciting company." And because there's already a huge global industry built around manufacturing silicon chips, "the most advanced silicon-spin companies have the promise of being able to leapfrog the superconducting, trapped-ion, and photonic qubit companies," says DCVC General Partner James Hardiman.

Again, though, there likely won't be a single path to fault-tolerant quantum computing. By investing in a range companies, "We're not trying to say 'we don't know which type of qubit will win, so let's do a bunch of them,'" Narang comments. "We're saying, multiple qubits could win, and they might also address different types of problems."

Figure 2.1.1

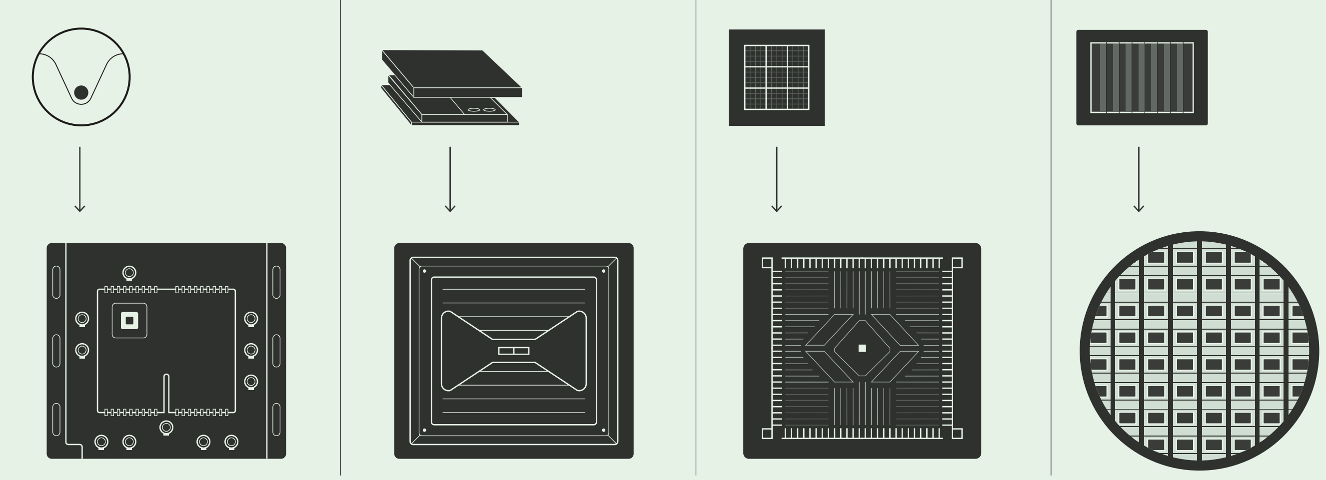
Quantum Modalities: Different approaches to embodying qubits in hardware require unique combinations of technologies.

Neutral atom: Encodes quantum information in the energy states of individual neutral atoms such as rubidium or ytterbium, held in place by optical tweezers.

Trapped ion: Suspends a single charged atom in electromagnetic fields, using tuned lasers or microwave pulses to encode information.

Silicon spin: Represents qubits in the "spin up" or "spin down" of a single electron or atomic nucleus confined within a silicon chip.

Photonic: Encodes qubits in the polarization or paths of individual particles of light, allowing the quantum states to move through optical components.

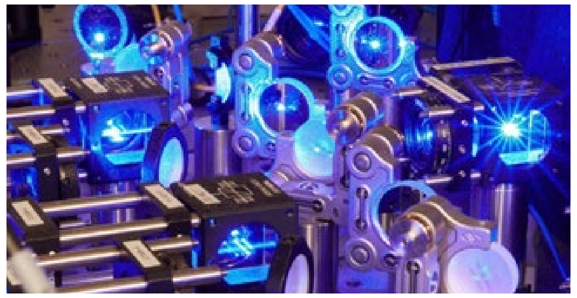


Advanced error detection and correction are critical elements in the quantum revolution.

The first generation of quantum startups tried to win mostly by building better qubits. The next generation—the group of companies we've already been talking about—is building systems designed from the outset to support scalable error correction, and that's what will finally make qubits usable in the fault-tolerant quantum computers of the near future.

To review: physical qubits are the finicky objects that hold a 1 and a 0 in superposition inside a quantum computer. Logical qubits are error-corrected qubits where quantum information is encoded across many entangled physical qubits and continuously stabilized through measurement. When the fidelity of physical qubits was low, and when error detection and correction codes were relatively primitive, it could take thousands—sometimes tens of thousands—of physical qubits to make just one reliable logical qubit. For two decades, then, the bottleneck to fault-tolerant quantum computing was whether we could even make decent physical qubits.

And today we're getting there: as we explained in the previous section, physical qubits are gaining fidelity and becoming more manufacturable, to the point that some amount of redundancy is affordable. But the big question is how much redundancy is needed, and how that redundancy is organized: will a fault-tolerant quantum computer require millions of physical qubits, or merely hundreds of thousands? And that's where better error-correcting schemes enter the picture. If physical qubits are good enough—below the error threshold required by a given scheme—then increasing the number of physical qubits participating in that scheme causes logical error rates to fall exponentially. But not all schemes use redundancy equally efficiently; the decisive question is how much reliability you get per additional qubit.



↑ **Iceberg Quantum's** Pinnacle architecture could achieve fault tolerance with 1/10 as many physical qubits.

Error correction schemes called "surface codes" were the first practical bridge from fragile qubits to scalable machines. They arrange physical qubits in a 2D grid (a surface) and assume that each qubit interacts only with its nearest neighbors. Quantum information is stored in data qubits, and neighboring ancillary qubits detect errors by repeatedly measuring parity relationships (seeing whether pairs of qubits agree or disagree, without revealing their individual states). This scheme is conceptually simple and it matches real hardware layouts, but it is qubit-hungry: making the quantum circuit more reliable requires widening the grid, and the number of physical qubits grows with the square of that width.

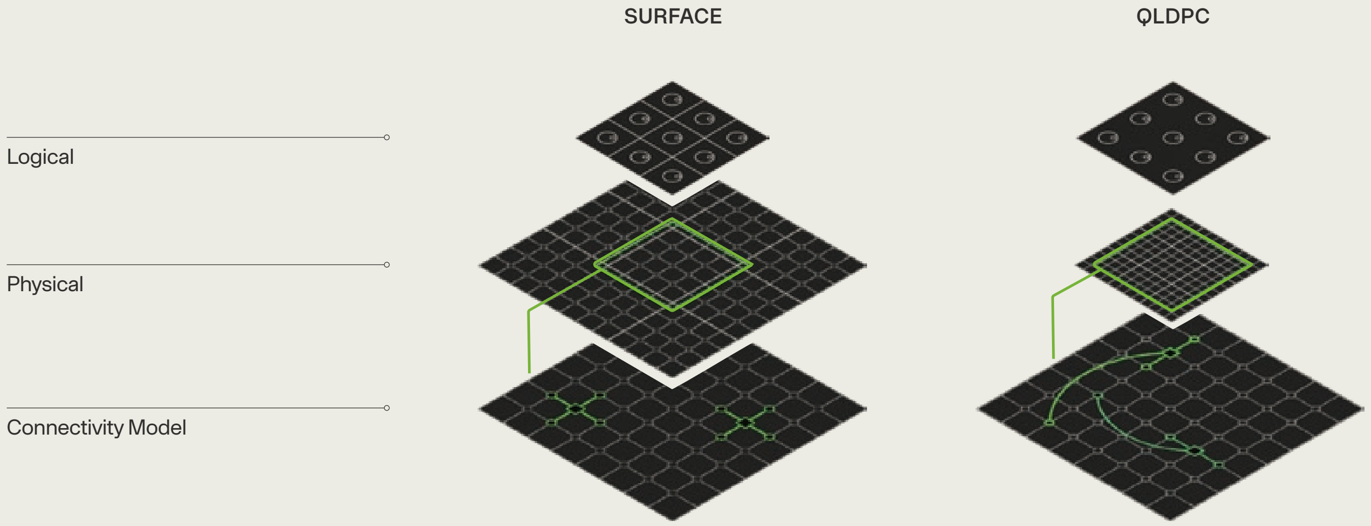
Thanks to recent advances, there's a more economical alternative: Quantum Low-Density Parity Check codes, or QLDPC, inspired by classic LDPC codes used in wireless communications protocols like Wi-Fi and 5G.

Figure 2.2.1

Advanced Error Correction

Surface codes restrict parity checks to nearest-neighbor qubits. QLDPC codes allow non-local parity checks, packing more logical qubits into fewer physical qubits.

Source: Quantum Zeitgeist.



In these schemes, parity checks occur across a sparse, cleverly connected graph. Logical errors can be suppressed without requiring a 2D grid of physical qubits that balloons in area. Whereas surface codes were designed for hardware with fragile qubits and limited connectivity, QLDPC codes assume fabrication improvements and higher connectivity, and use advances in sparse graph design to extract more logical protection per physical qubit.

That may sound like an incremental shift, but it could determine whether the field becomes economically viable within a venture time horizon. "With surface codes, the overhead of physical to logical qubits was always 100 to 1, or 1,000 to 1, or even higher, and now it's coming down to single-digits-to-1," says DCVC's Narang. "Of course, the devil is always in the details. But there's a lot the field has been able to accomplish, which makes companies focused entirely on quantum architecture and error correction valuable to us."

Recently we invested in Iceberg Quantum, an Australian company creating state-of-the-art QLDPC codes that are already accelerating the work of quantum hardware companies such as Diraq (silicon spin qubits), IonQ (trapped-ion qubits), and PsiQuantum (photonic qubits). Narang calls Iceberg Quantum "the ARM of quantum computing"—a company that doesn't build the hardware itself, but provides the architectural layer that multiple hardware vendors can then adapt. Iceberg's work "gives you a blueprint of which qubit you're going to move, which you're going to keep stationary, what pairs you're going to address," she explains. "It's like putting together a seating chart for a wedding that takes into account everyone's social proclivities—but where the constraints are what your qubits are capable of doing and what number of operations you can accomplish before decoherence."

In early 2026 the company published a paper showing that, in theory, a quantum computer using its new Pinnacle QLDPC architecture could factor a 2048-bit RSA key—the backbone of much of today's public-key cryptography—using just 100,000 physical qubits, compared to the 1 million or more researchers thought would be required using surface codes. That finding is a warning about the vulnerability of classical cryptography schemes. But even more, it's a signal that the physical scale required for fault-tolerant quantum computing may be compressing far faster than expected.

"

> [The Iceberg architecture] is like putting together a seating chart for a wedding that takes into account everyone's social proclivities—but where the constraints are what your qubits are capable of doing and what number of operations you can accomplish before decoherence."

Dr. Prineha Narang
Operating Partner, DCVC

We don't need to wait for fault-tolerant quantum computing to achieve "quantum advantage."

We know that we can manufacture qubits using multiple modalities, and we know that we can make them reliable using practical, low-overhead quantum error correction. In other words, the path to full fault-tolerant quantum computing is now clearer than ever. But can we create value even before we get there? Again, the answer is yes. The first scalable market will be quantum-enhanced sensing and measurement.

The Global Positioning System is an orbital metronome providing timing information that underpins modern travel, logistics, finance, telecom, and defense. It's a magnificent achievement, but it's not perfect. Outside of specialized applications such as surveying, its accuracy is limited to about 30 centimeters. GPS signals degrade at the extreme latitudes of the Arctic and Antarctic because satellite inclination is optimized for mid-latitudes. And signals can be unintentionally degraded by natural phenomena such as solar flares, or intentionally jammed or faked in active conflict zones and contested shipping lanes, such as those in the Middle East.

But today, an alternative is emerging: quantum-assured positioning, navigation, and timing (PNT) powered by quantum sensors and advanced software. One of the leaders in this technology is Q-CTRL, where we first invested in 2018. The company got its start in the field of quantum error suppression—not purely mathematical error correction in the style of Iceberg Quantum, but software that pinpoints environmental noise causing decoherence and optimizes control pulses to counteract it. Along the way, Q-CTRL got so good at stabilizing quantum systems against noise and other disturbances that it realized it could build its own hardware powered by quantum control.

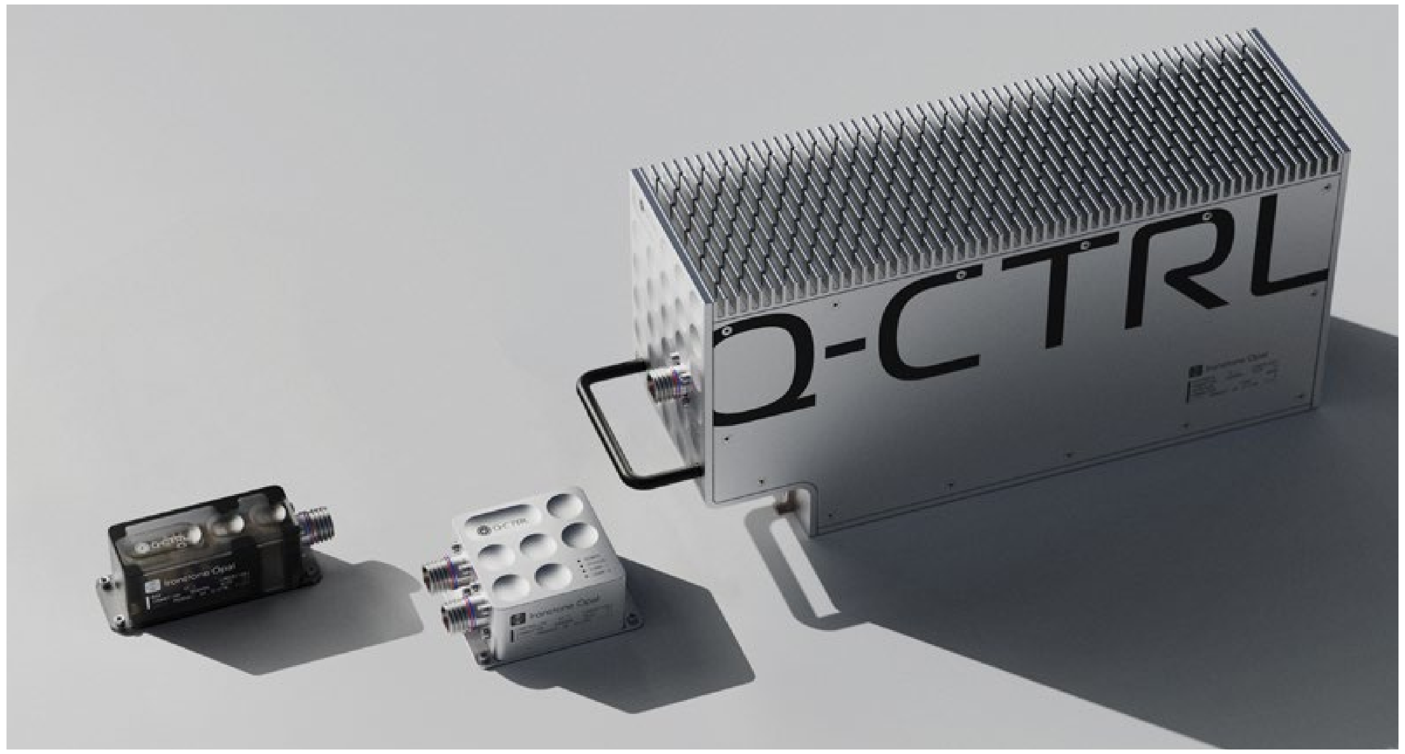
One of the things that can disturb a quantum system is a tiny shift in position or velocity. Treat that vulnerability as an asset, and suddenly you've got a quantum sensor that can detect very weak acceleration—and the core of a new kind of wayfinder. Together with Lockheed Martin, Q-CTRL has a contract with the U.S. Department of Defense Innovation Unit to build quantum inertial navigation systems that are immune to jamming and spoofing and will help warfighters determine their position, speed, and orientation even in GPS-denied environments.

Separately, Q-CTRL built a navigation system called Ironstone Opal that senses tiny changes in magnetic and gravitational fields, as well as acceleration. When data from Ironstone Opal's sensors is matched against fine-grained geomagnetic and gravimetric maps of the Earth, it's possible to dramatically reduce positioning error compared to standalone GPS. (In a 700-kilometer flight in Australia last year, the Q-CTRL system showed 111 times greater positioning accuracy than a high-end legacy GPS device.) *Time* Magazine named Ironstone Opal one of the best inventions of 2025, and the company is working with Airbus to test it on large aircraft. Q-CTRL's progress suggests that quantum precision, once confined to laboratories, can now be engineered into field-deployable systems—and sold.

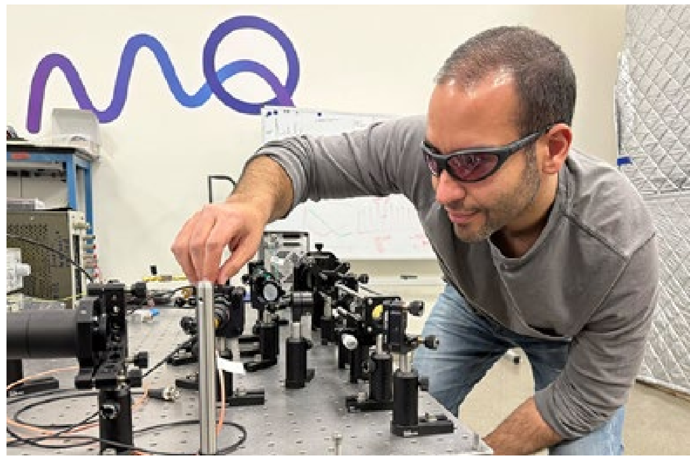
Quantum sensing doesn't require thousands of qubits or full logical stacks. It monetizes quantum decoherence directly—through defense and aerospace budgets first, and then the larger civilian market. In fact, it's an industry complete with an emerging supply chain. Mesa Quantum, where we invested in 2025, makes a specialized type of component for quantum sensors called vapor cells. They're small glass containers filled with clouds of vaporized alkaline metal atoms, whose absorption of laser light changes if their internal quantum states are shifted by magnetic fields, acceleration, or other disturbances. They enable magnetometers, gravimeters, accelerometers, gyroscopes, and atomic clocks with exceptionally low SWaP—size, weight, and power.

"In this case, low-SWaP means poker-chip-size or blueberry-size," Narang says. "You can imagine many of these sensors on drones or your favorite flying platform. So we're very excited about this investment in Mesa Quantum."

↘ **Q-CTRL's** Ironstone Opal quantum navigation system is suitable for drones, airliners, and autonomous cars.



> GPS-denied environments are everyone's fear in current conflicts. So you want navigation equipment that is resilient to electronic warfare—but the last thing you want is for everything to get heavier and heavier. [With Mesa Quantum] we can reduce the SWaP of these platforms and fly many more of them, and fly them cheaper."
>
> **Dr. Prineha Narang**
> Operating Partner, DCVC



↑ **Mesa Quantum's** vapor cells are at the heart of quantum sensors used for positioning, navigation, and precision time-keeping.

CHAPTER 3

In energy and climate, deep tech innovation is our best shot at limiting the cost of growth.

Today, the United States energy system is facing unprecedented constraints. Power demand is rising dramatically for the first time in a generation, driven by data centers, industrial reshoring, and electrification. At the same time, the physical infrastructure that delivers power is aging; interconnection queues and permitting cycles are slowing the addition of generation and the improvement of transmission; and climate volatility is increasing the frequency and cost of disruptions. We can debate policy responses, which are indeed potent when enabled, but electrons don't care. Unless the system's resilience improves, it will buckle under compounding stress.

In this environment, leveraging practical solutions is the only sustainable way forward. We can apply advances in disciplines such as subsurface engineering, nuclear science, power electronics, and intelligent control systems to reimagine energy production and management at their roots. And we can use better data to understand and manage the growing structural risks from climate change. This is how we improve reliability, predictability, and affordability even in the face of accelerating electricity demand.

"Climate change is real, but the solutions need to pay for themselves," says DCVC Co-Founder and Managing Partner Zachary Bogue. Deep tech will be the mechanism by which resilience is transformed from a slogan into a declining cost curve—and nowhere is this more operative than in the energy and climate-tech sectors.

OPPORTUNITIES

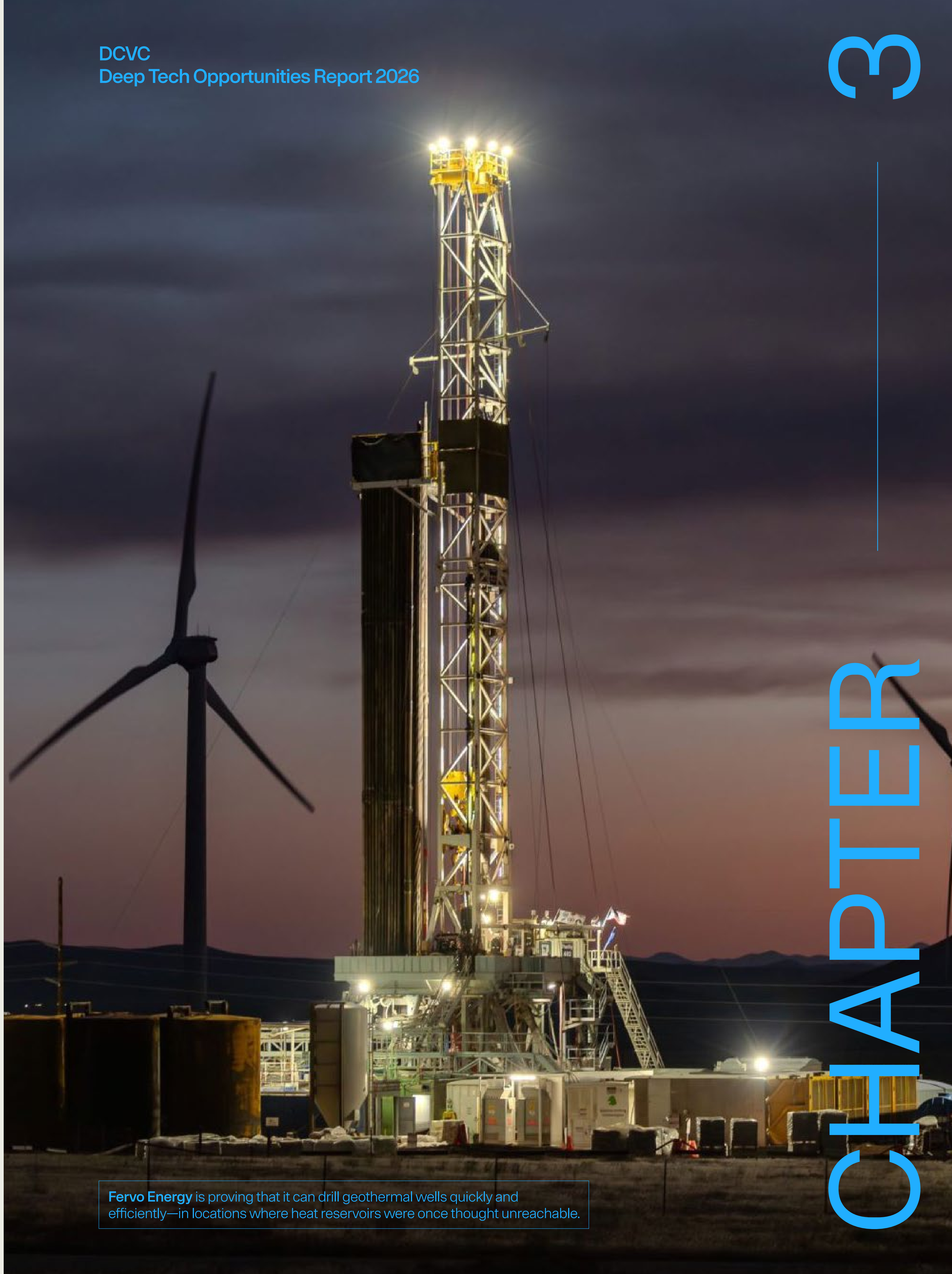
COMPANIES

Fervo Energy, Mainspring Energy, Lunar Energy, Oklo, Pacific Fusion, Radiant Industries, Rainmaker, Zap Energy

VOICES

Zachary Bogue, Dr. Rachel Slaybaugh, Milo Werner

CHAPTER 3



Fervo Energy is proving that it can drill geothermal wells quickly and efficiently—in locations where heat reservoirs were once thought unreachable.

New approaches are unlocking geothermal power in places that were previously thought uneconomic.

What the U.S. grid needs most urgently is more clean firm power: predictable, 24/7 electricity that doesn't vary depending on weather, time of day, or volatile fuel supplies, and that doesn't add more greenhouse gases to the atmosphere. Advances in reservoir modeling, horizontal drilling, and subsurface sensing—many adapted from oil & gas—are transforming geothermal power from a niche resource into a scalable source of clean baseload power.

Enhanced geothermal systems (EGS) have crossed from technical plausibility into infrastructure-scale deployment. The key to efficient geothermal energy production is reliably getting sufficient hot water out of the ground to make power. Traditional geothermal wells could be built only in rare geologic formations where naturally occurring water was flowing through hot, permeable rock. In EGS, by contrast, pump water flows down an injection well, travels horizontally through fractures in the rock, absorbs heat, rises back up to the surface through a production well, and passes heat to a turbine, which converts it into electricity. Fervo Energy's addition of horizontal drilling and hydraulic fracturing means the well lengths are much longer and the fracture surface area is engineered—creating enough hot surface area to ensure significant power production.

Exponential learning curves—where performance improves and cost declines in a virtuous cycle with cumulative deployment—are core to the economics of deep tech innovation, especially in climate tech. Fervo Energy, which we've backed since 2022, is riding such a curve. Between 2022 and 2025, the company reduced drilling times by approximately 75 percent and reduced per-foot drilling costs by about 70 percent. The two wells at Fervo's first "Project Red" site in Nevada took between 65 and 75 days to drill, at more than $800 per foot; more recent wells at its Cape Station site in Utah have reached their total depth in under 20 days, for $200-$300 per foot. CEO Tim Latimer has described the company's strategy as one of repetition rather than reinvention. "For us, going 100 times bigger means doing exactly what we've already done, but just 100 more units of it," he told Volts podcast host David Roberts. "This is the whole key of the repeatability of making this work."

← **Fervo Energy's** plans call for Cape Station in Utah (shown here under construction) to reach 100MW of operating capacity by 2027.

Just as important, Fervo says Project Red in Nevada has not experienced the kind of thermal decline that often plagued traditional geothermal projects. Now that it has substantially derisked its technology, Fervo plans to take a modular, standardized approach to scaling up: it will build 50-megawatt power facilities it calls GeoBlocks and collect them into multi-gigawatt energy hubs called GeoClusters, spreading across nearly 600,000 acres of land where it has already obtained geothermal leases. In other words, Fervo's growth is unlikely to resemble the slow, site-specific trajectories historically seen in the geothermal sector—it could look more like the learning curves followed in the past by the semiconductor, solar, and battery industries.

Publicly announced milestones suggest that EGS is already moving beyond pilot scale. At Cape Station, Fervo expects to deliver approximately 100 megawatts to the grid by early 2027, with an additional 400 megawatts to follow. As of the end of March 2026, the company had 658 megawatts of power purchase agreements in place, as well as an agreement with Google to develop up to 3 gigawatts of capacity to power current and planned data centers. Fervo raised about $2.2 billion in an oversubscribed IPO this May described by *Fortune* as "the biggest clean energy IPO ever."

The National Renewable Energy Laboratory estimated in 2023 that the United States harbors 230 gigawatts of potential geothermal energy at depths of less than 4 kilometers—a slice of the Earth's crust that is accessible with current-day drilling technology. (U.S. utilities had about 1,230 gigawatts of total capacity in 2024.) "If we only ever produce a few hundred gigawatts of power in the United States, only get to the point where we're 20 percent to 30 percent of U.S. electricity and that's it, it's probably still worth doing," Latimer told Volts. "But I still think that as we get to economies of scale and come down the cost learning curve, we're going to be able to go to even deeper resources and make that work." And that could make geothermal one critical element in a future of abundance and resilience.

↑ **Fervo** Co-Founder and CEO Tim Latimer and Co-Founder and CTO Jack Norbeck rang the Nasdaq opening bell in celebration of the company's initial public offering in May.

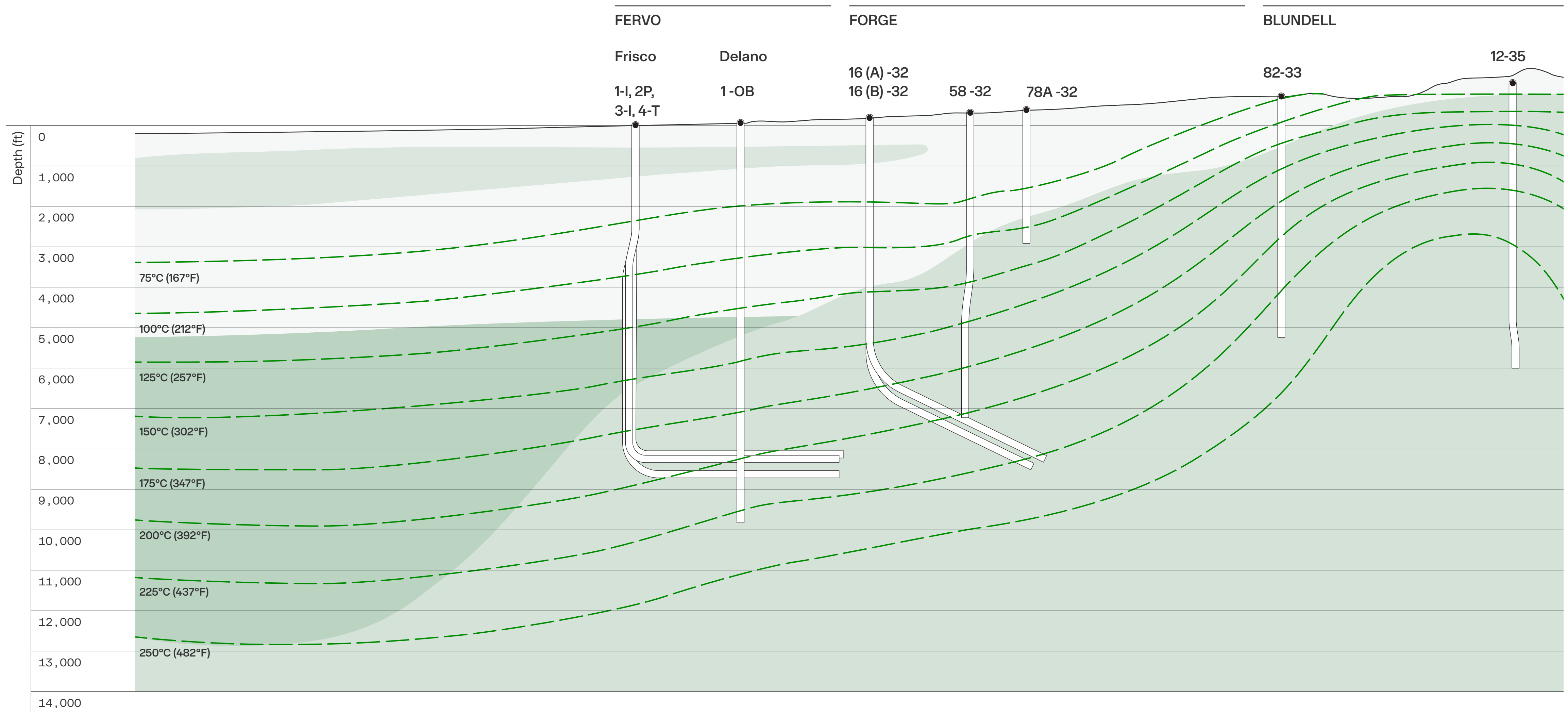


Figure 3.1.1

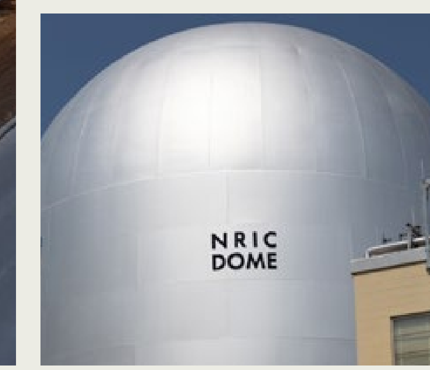
Fervo's Cape Station wellfield taps heat reservoirs in the granitic basement formations (bottom green layer) of western Utah. It's adjacent to the DOE FORGE project and PacifiCorp's Blundell geothermal plant, but Fervo is drilling to hotter layers at 8,000 to 9,000 feet of vertical depth.

Source: Jack Hunter Norbeck, Christian Gradl, and Timothy Latimer, "Deployment of Enhanced Geothermal Systems technology leads to rapid cost reductions and performance improvements," EarthArXiv, 2024.

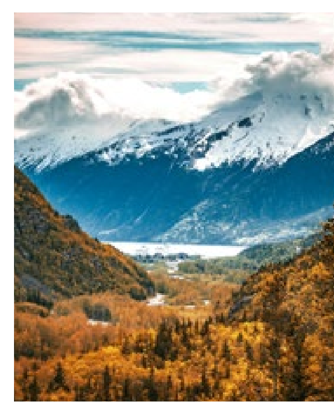
Micro nuclear reactors will bring firm power to locations the grid cannot reach.

In many remote or high-risk environments, the weakest link in energy supply isn't generation technology, it's fuel logistics. Micro nuclear reactors—designed from the outset for transportability, rapid deployment, and sealed operation—offer one solution.

Radiant's 1-MW microreactors (in the fenced area of this rendering) can be placed outside data centers or other facilities.



Radiant will test the Kaleidos reactor inside Idaho National Laboratory's DOME facility.



Transporting diesel fuel to remote locations is difficult and expensive. And yet there are hundreds of communities and settlements in the United States—primarily in Alaska—that are so far off the grid that they're forced to depend almost entirely on diesel generators, at a cost-per-kilowatt-hour that's much higher than the national average. In the winter months, cold, darkness, and harsh arctic weather conditions make fuel delivery especially unreliable. Numerous Defense Department forward- and remote-operating bases face similar challenges, with added risk for the military personnel who deliver the fuel.

Small factory-built fission reactors are an intriguing alternative. They provide firm, continuous electricity; they are virtually drag-and-drop, requiring minimal on-site setup; and the optimal designs operate for years without refueling.

In 2022 we invested in Radiant Industries, a company in El Segundo, Calif., building a 1-megawatt microreactor called Kaleidos that's scheduled to undergo its first fueled test this summer. The technology is wholly unlike the utility-scale reactors familiar since the 1950s, with their giant containment structures, elaborate water-based cooling systems, and tall cooling towers. The entire Kaleidos system—reactor core, heat exchanger, generators, electronics, and all—fits inside a semitrailer. Inside the core are kernels of uranium, carbon, and oxygen encased in layers of carbon and ceramic—a durable, robust combination known as TRISO (TRI-structural ISO-tropic particle). Helium gas carries heat from the reactor to the power generation loop, and an air-cooling system ejects the heat to the atmosphere, with no water required.

Kaleidos' small footprint makes it deliverable to any location by road, rail, or air. Its sealed, self-contained architecture and passive safety systems mean it needs minimal operations or support staff. The reactor has a 20-year lifetime and only needs to be refueled every three to five years. And modeling suggests that microreactors can be cheaper than diesel generators in locations where delivered fuel prices exceed $1.50 per liter, about $5.70 per gallon. The conflict in Iran has already driven U.S. average pump prices above that level in some states. And transport logistics and structural dynamics in some regions drive prices much higher. That's when alternatives like microreactors make economic sense.

We admire Radiant in part because they're so focused on execution. "For the last five years, since I started working with the company, they've been planning to turn on their demonstration reactor in 2026, and they've stayed on schedule and on budget," says DCVC General Partner Dr. Rachel Slaybaugh, who leads DCVC's nuclear investments and created the nuclear fission program for the DOE's ARPA-E. Radiant will be the first company to test their reactor inside the Idaho National Laboratory's DOME facility—where they can take advantage of the lab's infrastructure—as well as the first advanced reactor to do a full power demonstration.

Also this year, Radiant will break ground on a factory in Oak Ridge, Tenn., where it plans to build as many as 50 reactors per year. Takers will include the Pentagon's Defense Innovation Unit and the Air Force, which have signed a deal to buy microreactors for military bases. Equinix, the world's largest operator of data centers and colocation centers, has put down a deposit on 20 reactors.

Interest in nuclear power is spiking in the wake of executive actions that created the DOE's Reactor Pilot Program, a streamlined process for testing and early commercial deployment. Ten companies have been selected for the pilot program, including Radiant and DCVC-backed Oklo (see sidebar). As regulatory timelines compress and capital flows into the sector, Slaybaugh emphasizes the importance of rigorous engineering and governance. "There are a few serious companies, and there's a cacophony of other people in the space that are not as responsible," Slaybaugh says. Radiant's board includes former Nuclear Regulatory Commission leadership, and the company has engaged experienced external reviewers to scrutinize its design and safety case. In an industry where credibility matters as much as innovation, that preparation may prove decisive.

Diesel generators aren't just sources of carbon and particulates—they make energy supply into a recurring logistics problem. A sealed microreactor provides multi-year predictability in locations the grid hasn't yet reached or will never reach. We see advanced, small-scale fission as part of the larger shift to a less volatile, more resilient energy infrastructure.

↓ **Radiant** microreactors are modular, scalable to any location's power needs.

> "For the last five years, since I started working with [Radiant], they've been planning to turn on their demonstration reactor in 2026, and they've stayed on schedule and on budget."

Dr. Rachel Slaybaugh
General Partner, DCVC

OKLO: A PROGRESS REPORT

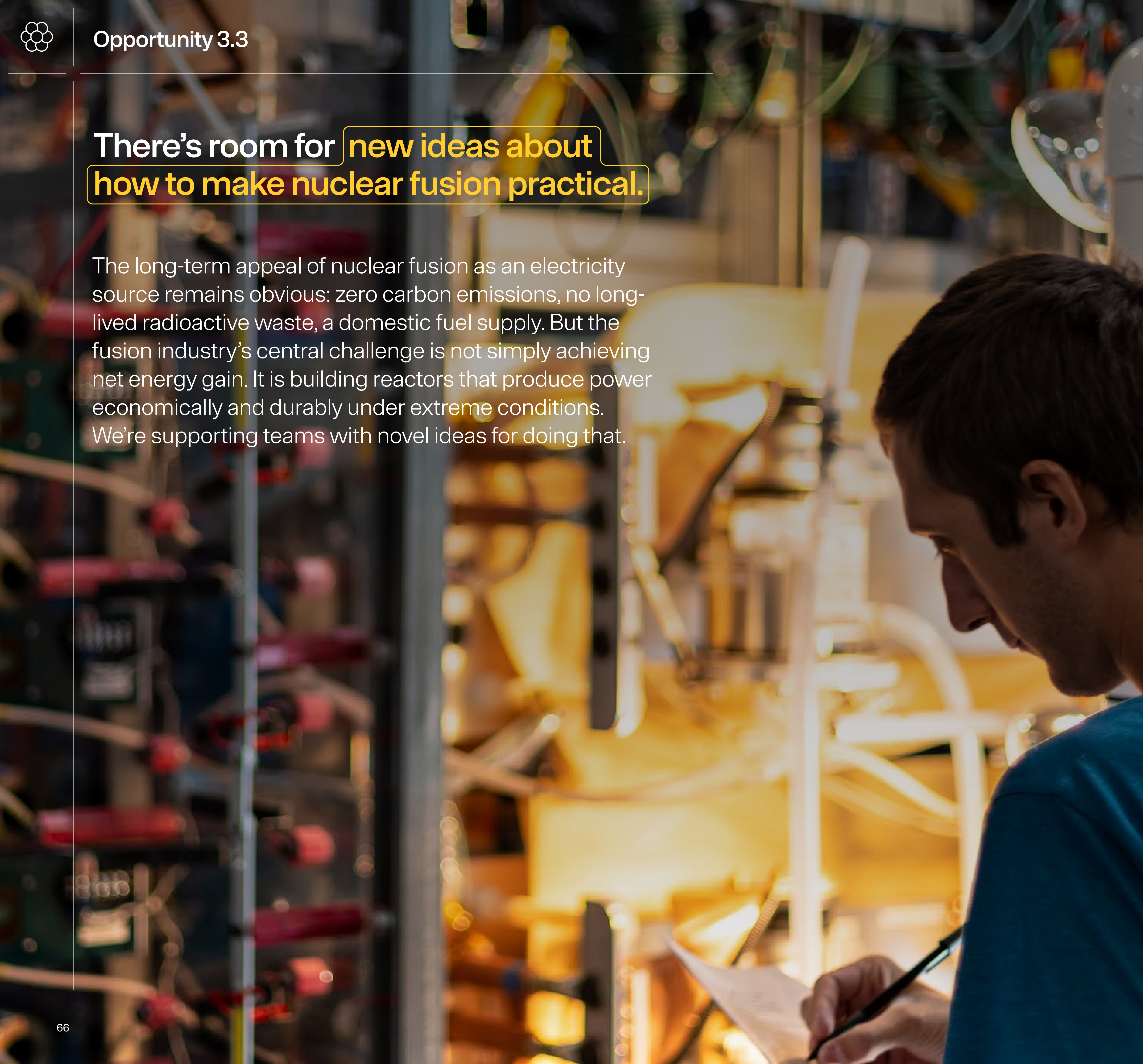
↑ In September 2025, **Oklo** broke ground on its first Aurora powerhouse at Idaho National Laboratory.

One of DCVC's first and most successful investments in next-generation nuclear fission energy was in Oklo, which went public in 2023 [NYSE: OKLO]. Oklo's business spans three critical areas, all enabled by its fast reactor technology: producing reliable baseload power from microreactors; fabricating and recycling fuel for its own reactors and others'; and supplying high-value isotopes domestically. Bolstered by a series of federal regulatory reforms, executive orders, tax credits, and financing tools meant to accelerate nuclear projects, the company has moved from product development phase into true project deployment.

→ The formula starts with power production. The company's Aurora liquid-metal-cooled fast reactor can run on a variety of fuels, including nuclear waste from other reactors. In 2025 Oklo broke ground for its first 75-megawatt Aurora powerhouse at Idaho National Laboratory, where it's targeting operations for late 2027 to early 2028. Thanks to a signed prepayment agreement with Meta, Oklo will also build a powerhouse campus in Ohio supplying up to 1.2 gigawatts of power capacity for future Meta-owned facilities. Oklo ultimately plans to supply firm power to industrial and military customers as well as data centers.

→ The supply of nuclear fuel is a growth-limiting step in the nuclear industry. At its advanced nuclear fuel recycling and manufacturing facility in Oak Ridge, Tennessee, Oklo is building capacity to convert enriched uranium and nuclear materials into reactor-ready fuel for Oklo reactors and third-party reactors. It will also recover uranium for reuse in its own reactors, as well as transuranic elements for use in other advanced reactors.

→ Oklo's isotopes business—formed partly through the 2025 acquisition of Atomic Alchemy, another former DCVC portfolio company—will use specially designed reactors and facilities to source and package high-value isotopes for healthcare, industrial, space, and defense, with the goal of reducing reliance on foreign supply chains. Construction is nearly complete on a test isotope production reactor in Texas.



There's room for new ideas about how to make nuclear fusion practical.

The long-term appeal of nuclear fusion as an electricity source remains obvious: zero carbon emissions, no long-lived radioactive waste, a domestic fuel supply. But the fusion industry's central challenge is not simply achieving net energy gain. It is building reactors that produce power economically and durably under extreme conditions. We're supporting teams with novel ideas for doing that.

In 2022 the DOE briefly achieved Q>1, net energy gain at the target level, from nuclear fusion at Lawrence Livermore National Laboratory's laser-driven National Ignition Facility (NIF). Multiple commercial fusion players now aim to demonstrate Q>1 within this decade. The most advanced companies are pursuing magnetic confinement systems designed to hold superheated plasma in a steady state until deuterium and tritium atoms fuse, releasing enormous amounts of energy, partly in the form of fast-moving neutrons. Those neutrons heat a blanket surrounding the core, which in turn heats a working fluid that drives a turbine to produce power.

The problem is that the neutrons also bombard structural materials, such as the vessel wall and the first-wall components surrounding the plasma. Over time, this leaves them weakened and radioactive. "Any material that is exposed to the plasma basically gets shredded and has to be replaced pretty frequently," says DCVC's Slaybaugh. "And that maintenance has to be done with radiation-hardened robots. And so that's expensive."

In other words, the economics of fusion will depend partly on how cheaply reactors can be repaired after neutron damage. That means there's an opening for companies with new architectures that will optimize for OpEx and maintainability. And that's the niche Pacific Fusion is exploring.

← **Zap Energy** is a DCVC-backed company that aims to achieve fusion using the Z-pinch effect, compressing deuterium-tritium gas to extremely high density.

↑ **Pacific Fusion's** demonstration system will be the only pulsed-power inertial fusion facility of its scale, and will be capable of producing some of the most intense bursts of energy on Earth.

The company was co-founded in 2023 by Will Regan, Keith LeChien, Eric Lander, and Carrie von Muench. Regan is a physicist who spent several years at ARPA-E studying the technical and economic challenges facing today's commercial fusion approaches, then worked for the Google X Moonshot Factory, as did von Muench. LeChien is a former member of the technical staff at Lawrence Livermore. Lander has led enormous, technically challenging international research efforts, including the Human Genome Project. Their reactor design is built around pulsed magnetic fusion, which is conceptually very different from steady-state magnetic confinement and more similar to pulsed-power systems such as Sandia National Laboratories' Z Machine. "There's a lot of existing science they can build on," Slaybaugh says.

Pacific Fusion's design starts with hundreds of pulser modules, each ringed by capacitors that release electrical energy in coordinated pulses. The energy travels through pulse tubes that converge inside the fusion chamber on two electrodes, which drive current through a tiny fusion target. The target electromagnetically compresses until fusion occurs, releasing energy. The chamber is surrounded by water shielding to limit neutron flux through the other components.

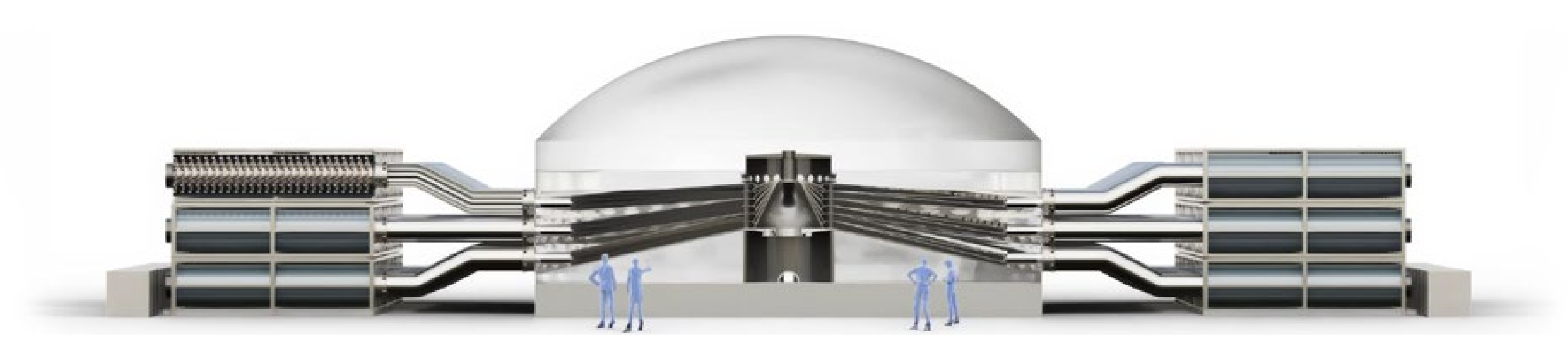
"If you don't have as many components that are near the plasma, you don't have as much stuff that gets irradiated, so it's lower-cost," Slaybaugh explains. "And you're not using lasers—you're using electrical pulses, where the components are also lower-cost."

Reflecting the high capital costs and long timelines in nuclear R&D, Pacific Fusion raised more than $900 million in initial funding. While all the funding was committed upfront, it will be unlocked in milestone-based chunks to ensure accountability. The company is headquartered in Fremont, Calif., but it will build its first research and manufacturing campus in Albuquerque, not far from Sandia. The goal there will be to demonstrate net facility gain—more fusion energy out than all the energy stored in the system—by 2030.

Pacific Fusion is not alone in pursuing innovative architectures. We invested in another alternative fusion company, Zap Energy, in 2022. Zap's approach is to initiate fusion by squeezing a filament of deuterium-tritium gas within the powerful magnetic field that forms as a current flows along a tube. That phenomenon is called the Z-pinch effect; another phenomenon, sheared flow, breaks the pinched plasma into layers to keep it from dissipating and fizzling out.

Figure 3.3.1

Pacific Fusion is betting on a technology called pulsed magnetic inertial fusion. Pulser modules made from thousands of identical parts focus energy on centimeter-scale fuel containers in the fusion chamber.



> You don't have as many components that are near the plasma [in Pacific Fusion's design], so you don't have as much stuff that gets irradiated, so it's lower-cost."

Dr. Rachel Slaybaugh
General Partner, DCVC

↑ **Zap Energy**'s style of fusion requires no magnets, cryogenics, or high-powered lasers. It works by running a powerful current through a tube of plasma stabilized by a phenomenon called sheared flow.

In fusion, a reactor's performance depends on the product of plasma density, temperature, and confinement time—a measure known as the Lawson criterion. In its latest experiments, Zap has created microsecond-long "shots" that pressurize plasma to 1.6 gigapascals, or 16,000 times atmospheric pressure at sea level. That is orders of magnitude higher than the plasma pressure sustained inside conventional tokamak-style reactors. But tokamaks operate over far longer confinement times. Zap's strategy is to achieve extreme pressure and density for a very short time. That means they can use simpler equipment that may require less capital (since it doesn't depend on superconducting magnets or high-powered lasers) and could be easier to maintain under neutron exposure. Zap says its reactors will also be small and modular, potentially scaling up or down to serve a single facility or a whole city.

We see fusion as a field where architectural diversity matters—and as with the effort to create stable qubits for quantum computing (see Opportunity 2.1), we think there could be more than one winner. Yet few serious observers expect fusion to contribute meaningful grid power before the late 2030s or early 2040s. To become part of our infrastructure, fusion technology must follow a downward cost curve driven by manufacturing learning and the practical economics of maintenance and component replacement. The fusion investments we're making today are bets that new reactor architectures will in fact align physics with economics.

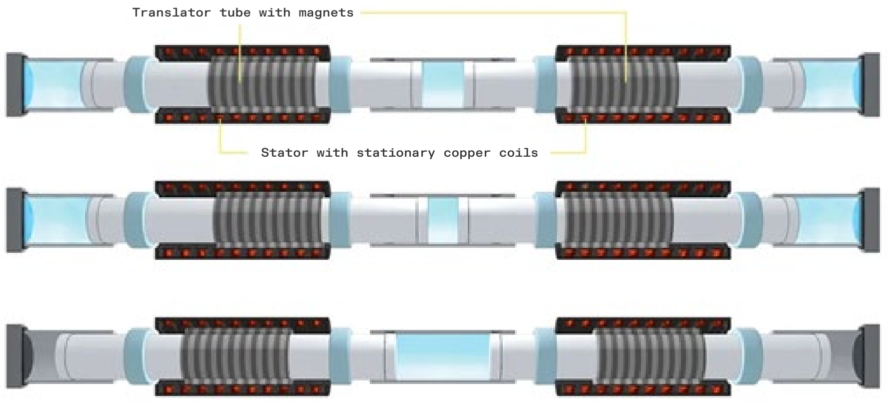
Big power users are turning to behind-the-meter power capacity and power management.

The North American electrical grid (which is actually a patchwork of five separate grids) was an indispensable creation that bolstered economic growth for decades. But it's built around concepts of centralized control and governance that reflect 20th-century limitations. Today, thanks to more sophisticated power electronics and flexible generating capacity, the grid is shifting from synchronized bulk power to modular, digitally controlled sources, shared across smaller geographies—which will make the system more resilient by design.

When Westinghouse won the "war of the currents" around 1890—prevailing in its argument that an alternating current system could supply large areas more efficiently than Edison's direct current system—it created constraints we're still grappling with today. With AC, power can be generated centrally, "stepped up" in voltage by transformers for long-distance transmission, then "stepped down" for use in homes and businesses. But as an AC grid grows, every generator feeding into it must stay precisely synchronized with all the others at an alternating frequency of 60 hertz. On top of that, because grids weren't designed with storage in mind, demand and supply must match in real time—meaning every watt of power produced must be consumed almost instantly.

Over time, these requirements—together with burgeoning bureaucracy, misaligned incentives, and conflicting vested interests—led to today's rigid and inflexible grid. There's a years-long "interconnection queue" for proposed new power plants and storage systems that need government and utility approval to hook into the grid; the projects on this waiting list amount to more than twice the nation's installed generating capacity (though many will never be built). But even after approval, long permitting times and regional planning processes mean that it can take 5-15 years to build high-voltage transmission lines to get power where it's needed—a timeline that complicates planning for new facilities such as factories or data centers. (One utility told Google it would take 12 years just to study building transmission lines to a proposed data center, a company official recently disclosed.)

↗ In **Mainspring Energy's** linear generator, fuel is compressed until a flameless reaction drives tubes outward; this motion is converted into electricity.



The fundamental challenge is that the grid was engineered for stability across large geographies, not for rapid capacity expansion at the network's edges. But today technologies are emerging that allow users to provision power locally—decoupling generation from geography. We've invested in several companies that use computation to coordinate power flows and make dynamic growth feasible. What's happening, ultimately, is that electricity is becoming programmable. Electrical hardware that used to be single-function is now more reconfigurable. Synchronization that used to be mechanical is now controlled through electronics. Switching and coordination are going digital. Together, these advances will help us reengineer the grid for adaptability.

Look at Mainspring Energy, where we invested in 2025. The company builds small, 250-kilowatt linear generators, each about the size of a parking space, that extract energy by compressing fuel between a pair of air-cushioned pistons, avoiding high-temperature combustion and its associated emissions. The power electronics in the generators allow them to switch between any gaseous fuel—natural gas, propane, biogas, hydrogen, ammonia—and compensate automatically for different energy densities. And unlike conventional turbines, which must spin at precisely 3,600 revolutions per minute to match the grid's 60 Hz frequency, Mainspring's generators create an electrical waveform digitally. That lets them operate in parallel with the grid, or form an islanded microgrid when needed. In industry language, they can "load follow"—ramping output up or down rapidly in response to real-time demand.

> "You could have one Mainspring generator at a supermarket, or you can have 100 of them at a data center."

Milo Werner
General Partner, DCVC

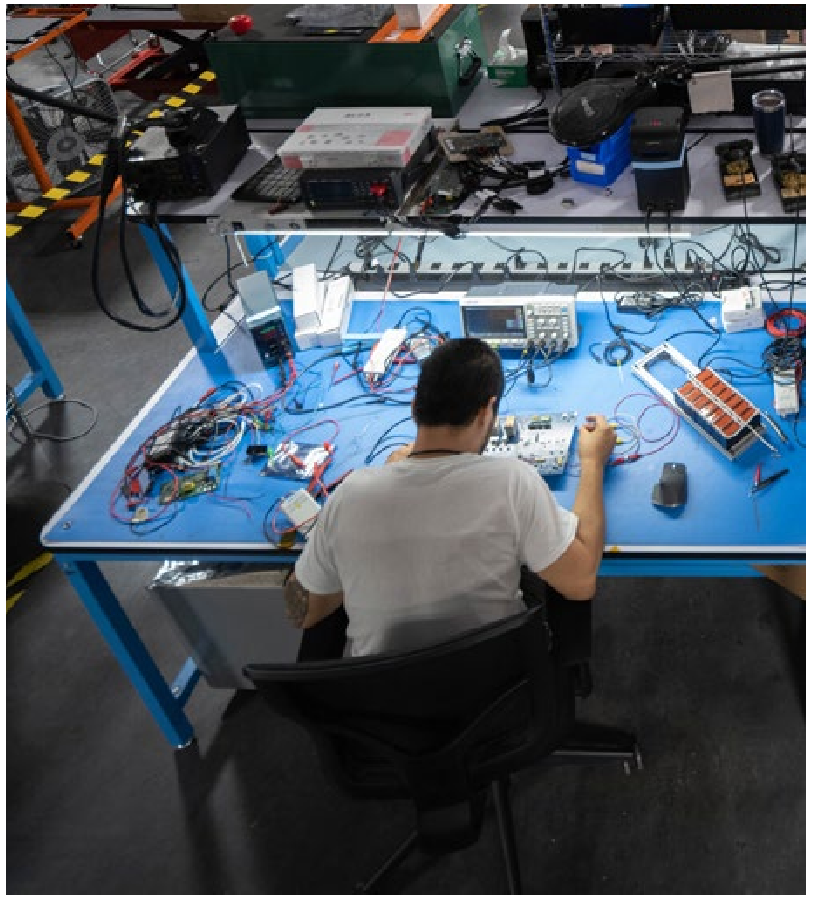
They're also modular, which means they're scalable via replication. "You could have one Mainspring generator at a supermarket, or you can have 100 of them at a data center," says DCVC General Partner Milo Werner. Already, Mainspring has hundreds of megawatts of generators in the field or in advanced development, supporting applications and facilities ranging from municipal generation (helping cities grow despite long interconnection queues) to conventional power plants (helping utilities meet peak demand) to dairy digesters (where the generators are often powered by dairy biogas). "Now that we have the capacity to create digital waveforms at kilowatt scale, it's really transforming the power industry," Werner says.

Better power electronics mean it's also possible to turn individual homes into managed energy systems. Lunar Energy, where we invested this year, offers a system of modular storage batteries and control devices designed to make solar-equipped homes more independent from the grid—and more affordable to run.

Lunar's batteries come in modular blocks that are charged by a home's rooftop photovoltaic panels, together with an inverter that converts power for household use. A nearby Bridge device monitors energy consumption—displaying the data via a smartphone app—and also monitors the grid, switching to battery and solar power in the event of an outage. In locations where homeowners are eligible for utilities' virtual power plant programs, software in the Bridge shows when a home is producing enough extra energy that it can send some back to the grid.

"It creates digital twin of your home and says, this is how we're going to run it," Werner explains. 'If there's going to be a power outage, or if there is a power outage, it can turn off your pool pump or turn down your HVAC system. And then it's always monitoring the price of electricity and deciding when to charge the battery and when to take the house totally off the grid"—and also when to sell power back to the grid. "It's the brain behind your home power consumption."

Mainspring and Lunar represent the industrial and residential halves of a single transformation: electricity customers want more flexibility and reliability than utilities can provide and are building their own capacity systems behind the meter. That was too difficult and expensive using legacy electrical equipment, but now advanced solid-state switches, transformers, filters, and management software make it feasible to imagine a more distributed grid where both generation and power management happen at the edges. Under that architecture, users take on part of the burden of reliability. But in exchange, they gain the ability to add capacity incrementally on their own timelines, rather than waiting for centralized upgrades. Resilience becomes a distributed feature of the system itself.



↑ **Lunar Energy's** system includes modular batteries (at left, above) and a Bridge device that connects customers' homes to the grid.

← **Lunar** develops, tests, and assembles every part of its system in Mountain View, California.

RAINMAKER: WE ARE RELEARNING HOW TO MAKE IT RAIN

↑ **Rainmaker** engineers prepare for a remote field test.

Managing for water scarcity and restoring water abundance must be part of any plan to make regional and local economies more resilient against a warming climate. There's plenty of water in the sky—and today deep tech is making an old technology, cloud seeding, measurable, scalable, and investable.

In 2025, two-thirds of the western U.S. was in drought, and the dry winter of 2025–2026 left Colorado's snowpack at its lowest point in decades. Mandatory cutbacks along the Colorado River are now hitting farms, ranches, and tribal communities hardest, and groundwater reserves are vanishing even faster than surface supplies. This change isn't cyclical—it's structural, and adaptation will have to come in many forms.

American researchers discovered the principles of cloud seeding in the 1940s: dispersing nucleating particles into clouds primes them to release precipitation that would otherwise stay aloft. At least nine western states already fund programs today. But the technology has barely advanced in decades, and crucially, operators have never been able to prove how much rain or snow they actually produced—which has kept private capital out.

Rainmaker, a California company where we invested in 2025, is changing that. Its "glaciogenic cloud seeding system" uses ground-based X-band radar and satellite data to identify clouds carrying pockets of supercooled liquid water, then deploys purpose-built drone fleets to disperse nucleation particles precisely where they will generate the most precipitation. Radar and satellite data feed a proprietary quantitative precipitation model that measures how much rain or snow fell in the seeded zone versus surrounding areas—turning an unmeasurable service into a billable one.

Operating in five western states as well as internationally, Rainmaker is helping customers build snowpack, recharge aquifers, refill reservoirs, stabilize hydropower streamflow, and deliver supplemental rainfall to farmland. Unlike speculative geoengineering, its operations are local, modest, and short-lived: adaptation at the human scale, buying time while longer-term transitions unfold.

CHAPTER 4

Industrial leadership will be won by redesigning manufacturing and materials production from first principles.

In industry after industry, the old ways no longer make sense.

- → Why build a factory that produces a single product when we can build one that makes many?
- → Why extract hydrocarbons from the ground when we can synthesize them from air, water, and waste streams?
- → Why treat wastewater as a liability when it contains clean water, minerals, and usable energy?
- → Why depend on fragile global supply chains when smarter chemistry and modular systems can produce critical minerals closer to home?

Across manufacturing, materials, and water infrastructure, the common thread is a return to fundamentals: physics, chemistry, biology, and smarter control systems. When these tools are used to break through the real constraints on industrial operations, they make it possible to build more flexible factories, more resilient supply chains, and less wasteful resource systems.

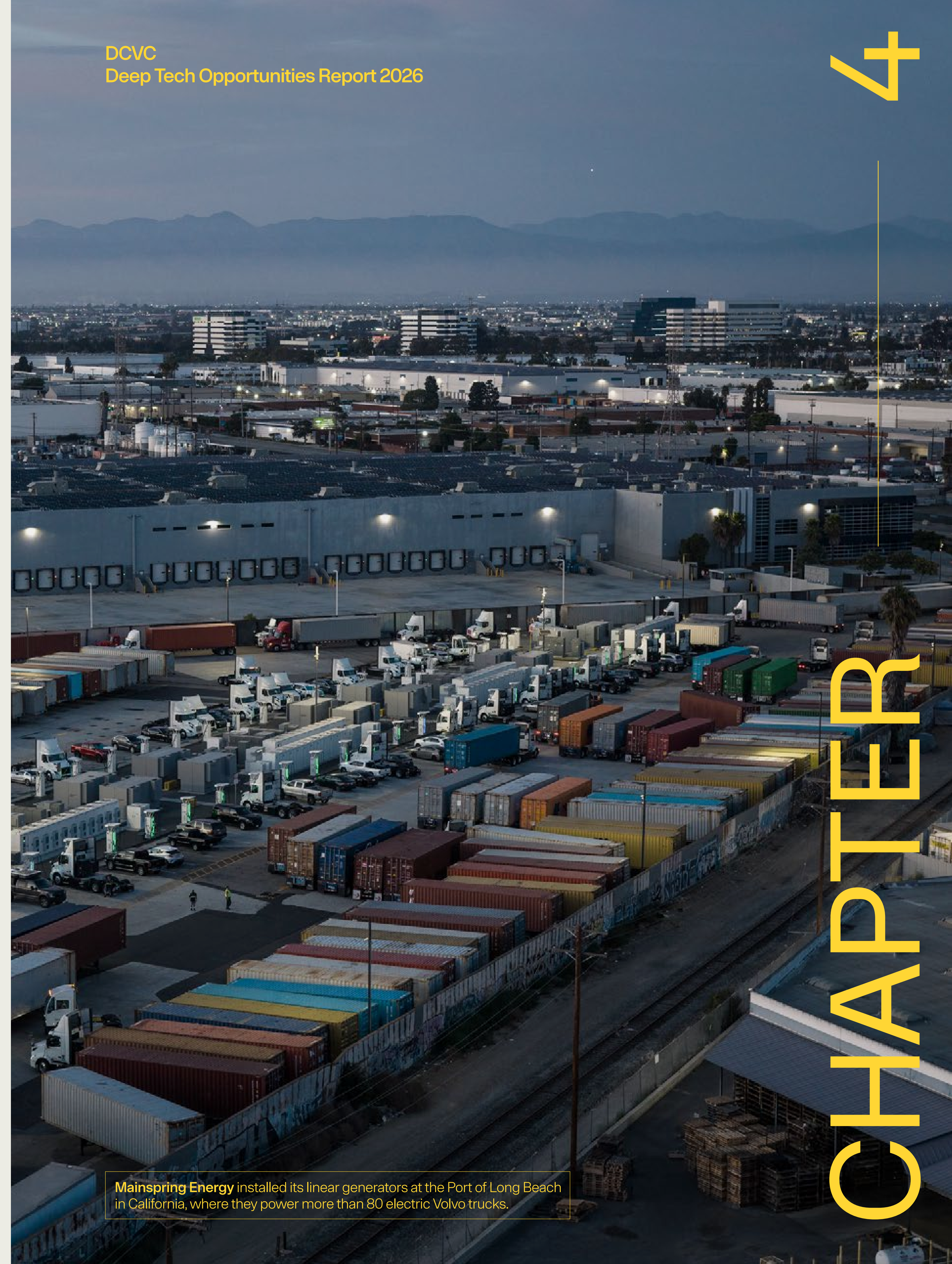
OPPORTUNITIES

COMPANIES

Agility Robotics, Alta Resource Technologies, Aquafortus, Brimstone, Circularity Fuels, Fulfil Robotics, Lumafield, Mainspring Energy, Pivot Bio, Slip Robotics, Tidal Metals, Twelve, Unspun

VOICES

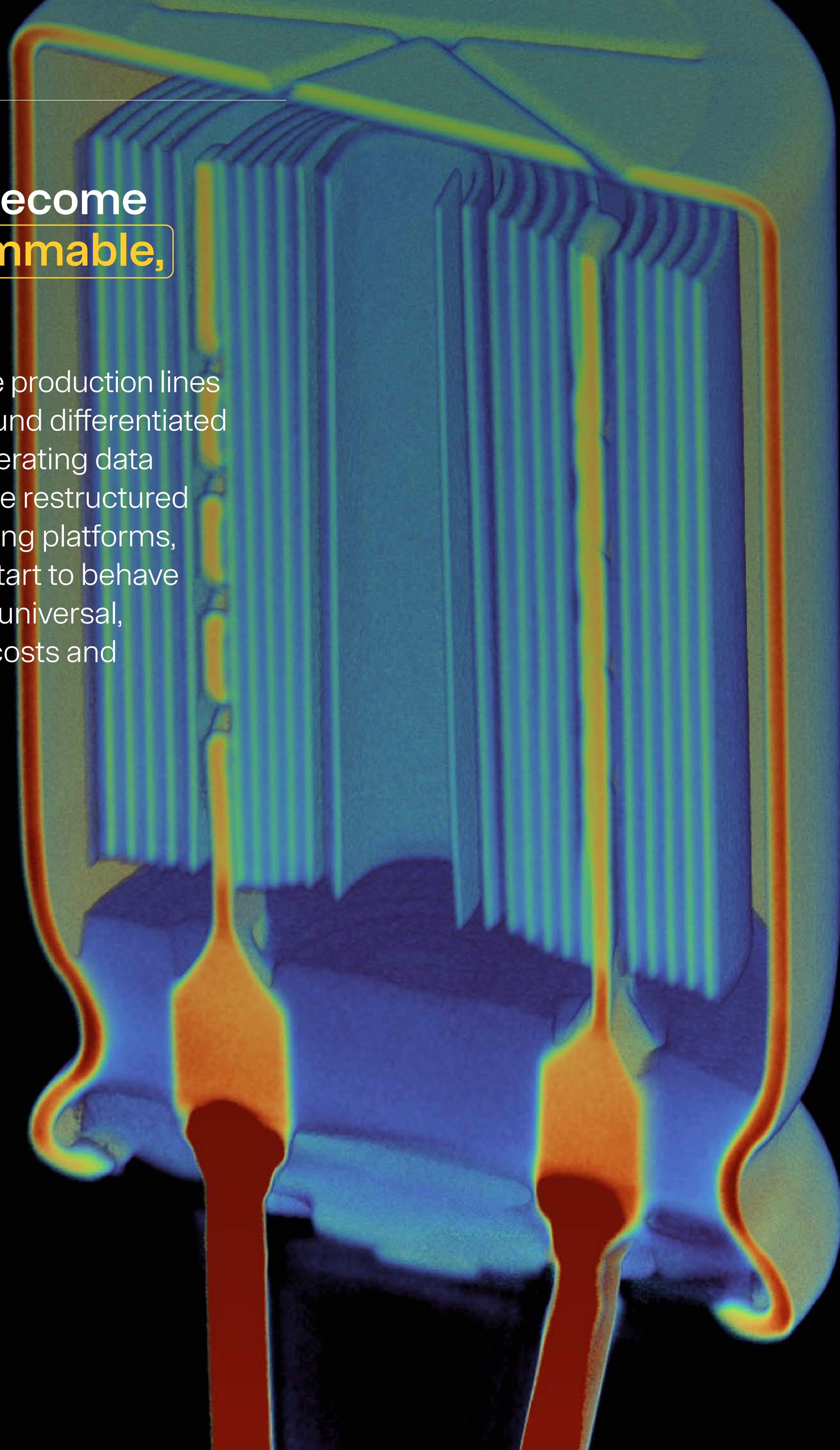
Dr. Stephen Beaton, Earl Jones, Kristen Rocca, Milo Werner



Mainspring Energy installed its linear generators at the Port of Long Beach in California, where they power more than 80 electric Volvo trucks.

Industrial production will become more modular and programmable, and therefore cheaper.

The age of single-purpose, capital-intensive production lines is giving way to modular systems built around differentiated hardware, embedded software, and accelerating data flywheels. As factories and warehouses are restructured around robotics, configurable manufacturing platforms, and distributed power systems, they will start to behave less like fixed infrastructure and more like universal, programmable machines—driving down costs and increasing flexibility and resilience.



↗ **Lumafield** builds industrial X-ray CT scanners and software that use AI to spot cracks, voids, and other imperfections in mass-manufactured parts.

In 2025, Anduril said it would spend nearly $1 billion building Arsenal-1, a "software-defined manufacturing" platform for mass-producing autonomous systems, in Ohio. That same year, First Solar opened a $1.1 billion solar panel factory in Louisiana using AI to detect defects, and Hyundai opened its $7.6 billion "Metaplant America" in Georgia, where hundreds of robots move and assemble parts while sensors supply data to digital twins that can catch quality issues and suggest fixes. When several of the world's largest new factories are built or announced with explicit commitments to AI, robotics, and highly automated tooling, that's no longer a pilot project—it's a structural shift in how industrial production works.

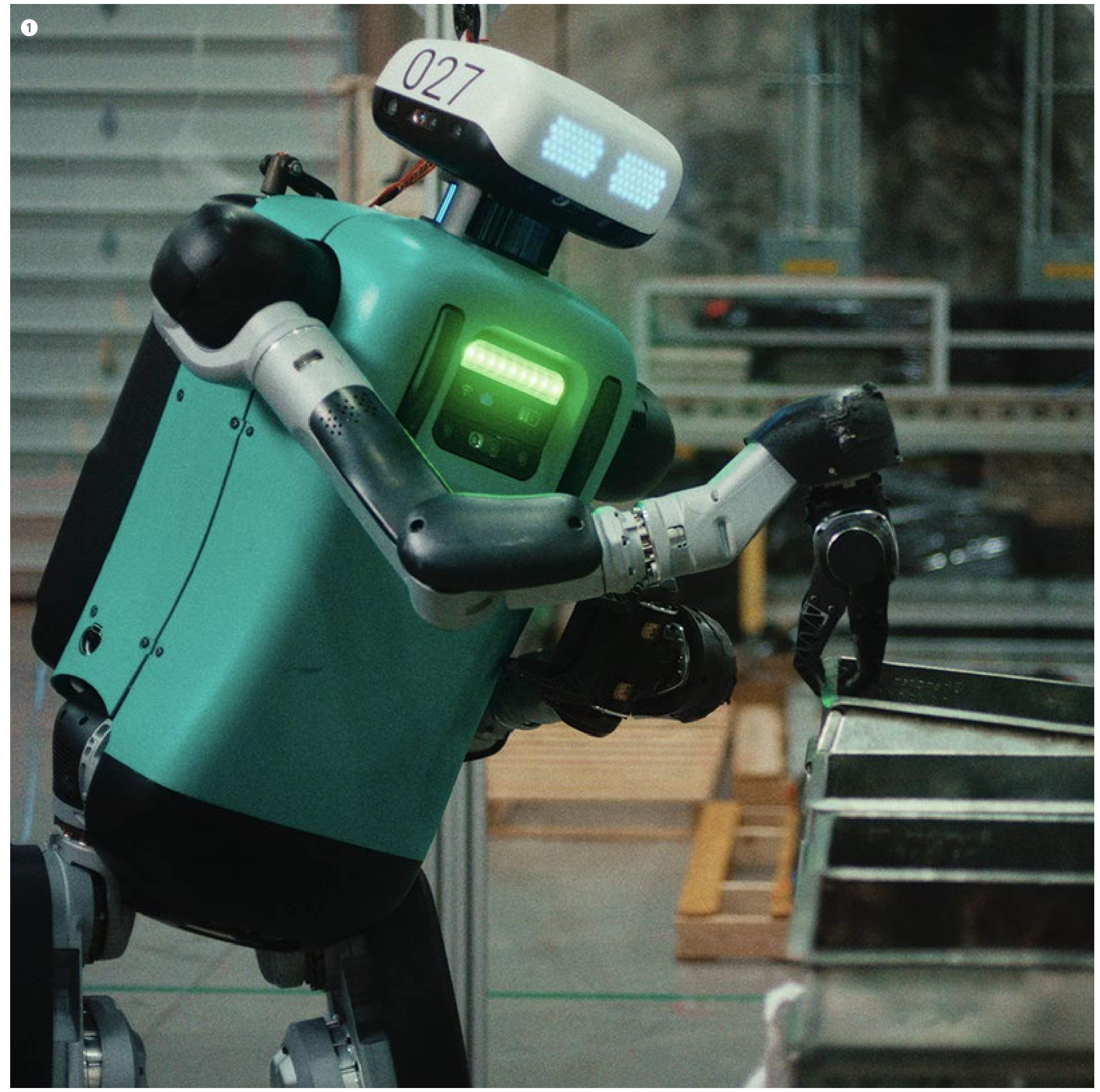
But software-defined manufacturing—the combination of responsive, repurposable production technology and software-optimized operations—doesn't have to cost billions. Dozens of startups are tackling narrower problems: detecting product defects, improving robot dexterity, or diagnosing performance drift. We see a coherent, fundable ecosystem trend emerging, away from monolithic lines dedicated to a single SKU, and toward distributed robotic cells that can be reconfigured, repurposed, and continuously improved through software.

"We make so many things that are kind of the same but just a little different," observes DCVC General Partner Milo Werner, who led new product introduction at Tesla prior to joining DCVC. "Think of battery packs. They're all about the same size. They all have cells and collector plates. They all come in an enclosure. There's no reason you couldn't run them down the same production line, with smart, automated arms making everything."

"

> The days of pure software innovation are over. In deep tech, the only way to create sustained value with large moats is through differentiated software built on differentiated hardware."

Milo Werner
General Partner, DCVC



❶ **Agility Robotics** offers a simulation and learning framework that lets customers teach Digit new capabilities.

❷ **Lumafield's** X-ray CT scanners help manufacturers detect and correct problems before they mar big batches of products.

❸ **Slip Robotics** designed its autonomous cargo pallets to fit with existing loading dock infrastructure, but use it more efficiently.

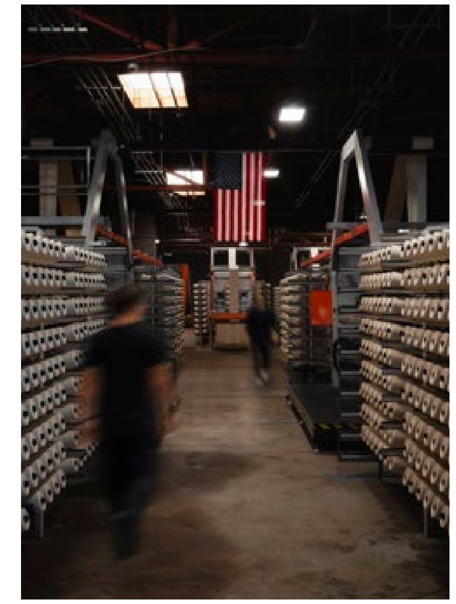
The point is not that software is eating factories. We talked earlier about physical AI and the rise of vision-language-action models that make robots more capable. But these models will likely converge in terms of their abilities, much as large language models have. Durable advantage will belong to companies that deeply understand physical operations—and pair that knowledge with flexible hardware, smarter sensing, and proprietary control or analytics software. "The days of pure software innovation are over," Werner says. "In deep tech, the only way to create sustained value with large moats is through differentiated software built on differentiated hardware."

Agility Robotics and Slip Robotics both fit that definition (see Opportunity 1.5). Agility's humanoid robots are designed to work alongside humans, in spaces built for humans, without requiring costly retrofits. Slip built its self-driving cargo pallets on top of a deep understanding of the loading dock—the universal interface between the warehouse and the road transportation system.

Lumafield, which helps manufacturers look inside their products non-destructively using industrial X-ray CT scanners and proprietary visualization tools, offers another example. Industrial CT scanning isn't new, but it long relied on slow, massive, expensive hardware using outdated software. Lumafield reengineered the stack, combining physics insight with modern software to make high-resolution inspection dramatically faster and more accessible. Its Triton scanner captures hundreds of 2D X-ray projections in fractions of a second, and its Voyager software converts that data into 3D volumes that reveal cracks, voids, leaks, and tolerance deviations—feeding insights back into production and preventing costly recalls. The breakthrough came not from inventing a new imaging modality, but from knowing which improvements would matter most to product developers and high-volume manufacturers.

Unspun fits the model as well. The company, where we invested starting in 2024, makes an automated circular loom called Vega that creates 3D textiles. Unlike conventional cut-and-sew lines, Vega routes thousands of individual yarns into seamless shapes on demand, eliminating cutting waste. Because the process is software-guided, a single machine can be reprogrammed to produce different styles without changing tooling.

↑ **Unspun** built a showroom at its Oakland microfactory where it can display apparel woven by its reprogrammable circular looms.



↑ **Unspun** can switch out the yarns feeding into its Vega loom to create new garments.

"

> Climate aside, [Unspun] makes a lot of sense for business. You reduce dead stock on your books, cut labor costs, and avoid shipping from Asia. There's a lot of cost savings that is baked into rolling out this technology. And you're saving a ton on emissions."

Kristen Rocca
Principal, DCVC

The company has built a microfactory in Oakland, Calif., where it's producing pants. Installed in other future locations in the U.S., Vega could shorten the apparel supply chain. Sustainability is also part of the appeal—the carbon footprint of an Unspun garment is smaller than that of a traditional import—but the primary advantage is economic. Reprogrammable production reduces lead times, cuts inventory risk, minimizes forecasting errors, and reduces unsold stock. "Climate aside, this makes a lot of sense for business," says DCVC Principal Kristen Rocca. "You reduce dead stock on your books, cut labor costs, and avoid shipping from Asia. There's a lot of cost savings that is baked into rolling out this technology. And you're saving a ton on emissions."

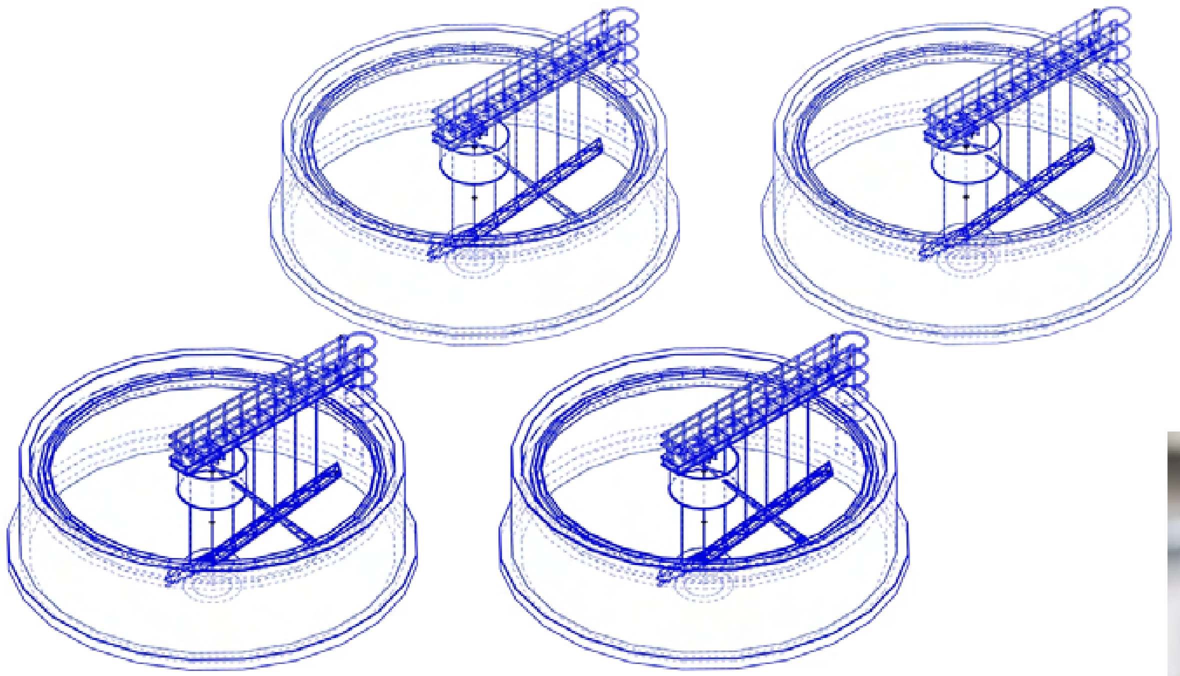
Building differentiated, reconfigurable, software-defined hardware is one way to create a successful deep tech product; another is to focus on modularity. Mainspring Energy does all of the above. In Opportunity 3.4, we explained how the software in their linear generators can reset the machines to use whichever gaseous fuel is most readily available. In addition the generators are small, each producing about 250 kilowatts of power. Customers can start with a few units and scale incrementally. In February the Utah Municipal Energy Agency chose Mainspring to supply 48 megawatts of new dispatchable generating capacity for Provo and the other cities it serves; that will mean collecting nearly 200 generators at a greenfield site in central Utah.

What ties all these companies together is a return to fundamentals: Understanding how factories work, how goods move, how product defects arise, how garments are woven, how electrons flow—and then redesigning those systems from the ground up with modular hardware and smarter control layers. Clever code is not enough. Industrial leadership will belong to those who pair deep knowledge of physics and on-the-ground operations with reconfigurable hardware and software that improves with every deployment.

Precision chemistry will reshape how we produce key industrial inputs.

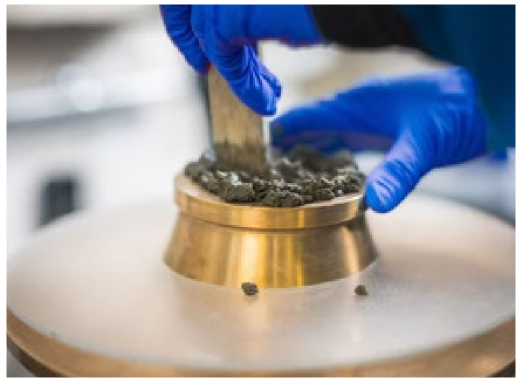
Programmable factories are only part of the story. An equally important transformation is happening upstream, in the chemistry that produces fuels, critical materials, and the materials that underpin modern infrastructure.

Brimstone is using chemistry to turn abundant rocks into building materials and light metals.



← **Brimstone** adds chemicals to calcium silicate feed rocks in large mixing and separation tanks.

↓ The process produces Portland cement, smelter-grade alumina, and supplementary cementitious materials (SCM).



For the last two centuries, we've relied on heat from fossil fuels to crack rocks and crude oil, refine metals, and rearrange molecules—while largely treating the resulting CO_2 emissions as an unavoidable externality. But increasingly, those processes are being redesigned around electrons, catalysts, and precise chemical control systems—opening the door to cleaner, more efficient materials production that can be carried out domestically.

Brimstone, where we first invested in 2020, is one leader in this movement. The company is applying the art of co-production—the hallmark of world-changing industries from oil refining to fertilizer manufacturing—to critical minerals and commodities, with calcium-bearing silicate rocks such as basalt or gabbro as the basic feedstock.

The company obtains these abundant rocks from simple quarry operations. They're crushed to the desired particle size, then run through a hydrometallurgical process to extract calcium, aluminum compounds, and residual silicates. Additional steps allow Brimstone to create:

1. Ordinary Portland cement, or OPC. The traditional method for making OPC is to break down limestone ($CaCO_3$) into quicklime (CaO) and CO_2—lots of CO_2. For every 1,000 kilograms of traditional cement produced, between 650 and 920 kilograms of CO_2 escape into the air, comprising 5 to 8 percent of anthropogenic CO_2 emissions. But by starting with calcium silicate rocks instead of limestone, Brimstone can extract the calcium compounds needed to make quicklime with no inherent CO_2 release.

2. Smelter-grade alumina, the feedstock for aluminum, a critical mineral. Historically, alumina has been refined from the ore bauxite, which is mined predominantly in Australia, China, Brazil, Guinea, and India. There's only one operating bauxite refinery in the United States, in Louisiana, which leaves the nation deeply dependent on imported feedstocks for aluminum. Given that this strong, light, conductive metal is indispensable for automobiles, airplanes, military craft, and high voltage transmission lines, among many other products, establishing a new domestic source of alumina is a matter of national security.

3 SCM, or supplementary cementitious materials. These are added to concrete to improve strength and durability. They traditionally came from the slag from steel production and fly ash produced by coal power plants, but with the reduction in coal generation, supplies have dwindled.

And those are just the company's first three products. Brimstone is developing additional processes to refine calcium silicate rocks into:

4 Steel, an alloy of iron and carbon that, along with cement, undergirds almost all modern construction. Traditional steelmaking in blast furnaces emits roughly 2 kilograms of greenhouse gases for every kilogram of steel produced. Brimstone will produce the metal from the iron found in its feed rocks, with lower carbon emissions.

5 Magnesium, the lightest structural metal and another critical mineral. It's needed for the aluminum alloys used in cars, planes, electronics, and defense systems, and provides a lightweighting advantage for applications that are mass-sensitive, such as EVs, drones, and humanoid robots. Currently, production of magnesium is heavily concentrated in China, which controls 85 to 95 percent of the global supply.

6 Titanium, yet another critical mineral, whose strength and corrosion resistance make it important for aerospace components, medical implants, and high-performance industrial equipment. Since the 2020 closing of a mine in Nevada, the U.S. has zero production of the metal's precursor; the leading producers are China, Mozambique, South Africa, and Australia.

← **Brimstone** aims to become a domestic supplier of smelter-grade alumina, a feedstock for aluminum. A modern passenger jet is approximately 80 percent aluminum by weight.

↑ In 2023 **Twelve** broke ground on its AirPlant facility for making eFuels and feedstocks with its power-to-liquid technology.

In other words, the company is building not just an alternative cement process, but an entire rock refinery. And if Brimstone demonstrates how rethinking the chemistry of rocks can reshape the materials economy, companies such as Twelve and Circularity Fuels are applying similar ideas to fuels—reprogramming how we produce hydrocarbons themselves.

We've been backing Twelve since its seed round in 2018. The company uses novel catalysts and renewable electricity to split biogenic CO_2 and H_2O, creating syngas (a mixture of CO and H_2) and O_2. Syngas is a feedstock for many valuable products, including sustainable aviation fuel (SAF) and chemicals such as E-Naphtha, a building block for plastics. With help from $200 million in Series C financing and $400 million in project equity from TPG Rise Climate, the company is building its first commercial-scale facility, called AirPlant, in the central Washington city of Moses Lake, where it expects to make 50,000 gallons per year of E-Jet SAF and 11,000 gallons of E-Naphtha. Airlines have been rushing to partner with, and invest in, Twelve to smooth the way for compliance with requirements such as the European Union's ReFuelEU initiative, which mandates that they use increasing amounts of SAF over time (a 6 percent blend by 2030, and a 70 percent blend by 2050).

But carbon transformation—converting CO_2 from emissions into critical materials that used to come from petroleum—isn't just about emissions reductions; "it's a business imperative that builds resilience, competitiveness, and long-term profitability," in the words of Twelve CEO Nicholas Flanders. The argument is that when industries use CO_2-derived fuels, plastics, and other chemicals as drop-in replacements for fossil-derived products, it can lower their exposure to volatile petroleum markets, enable modular, on-shore manufacturing, and help their customers respond to sustainability mandates.

Circularity Fuels is taking a complementary approach to carbon transformation. The company starts with biogas from dairy farms. It uses power from solar and wind farms to electrify a microchannel reactor, called Ouro, where proprietary catalysts split CO_2 and CH_4 from the biogas into syngas. Using the century-old Fischer-Tropsch process, Circularity converts the syngas into liquid hydrocarbons, which are then refined into SAF.

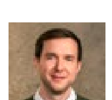
↑ **Circularity Fuels'** Ouro reactor is the size of a coffee thermos, allowing modular construction of syngas plants at biogas resource sites.

One Ouro reactor—essentially, a modified version of an automotive catalytic converter—can produce enough syngas to make two barrels of SAF per day, at 1/100 the cost of the steam or steam-methane reforming process traditionally used to make natural gas into syngas. The company says that if its biogas-to-SAF technology were deployed at all U.S. farms, landfills, and wastewater treatment plants, they could produce 42 million gallons of SAF per day, enough to replace 70 percent of the nation's supply of fossil-based jet fuel.

"Farmers in the U.S. and around the world are sitting on an untapped goldmine," says Entrepreneur in Residence Dr. Stephen Beaton, Founder and CEO of Circularity. "We're giving them the ability to turn waste into a profitable product that airlines desperately need."

Farmers can reduce synthetic fertilizer costs using Pivot Bio's microbial crop additives, which convert atmospheric nitrogen into natural ammonia fertilizer.

Farmers in the U.S. and around the world are sitting on an untapped goldmine. We're giving them the ability to turn waste into a profitable product that airlines desperately need."



Dr. Stephen Beaton
Entrepreneur in Residence at DCVC and Founder and CEO of Circularity

Speaking of farmers, there's a longstanding member of the DCVC portfolio that's using engineered microbes to produce a chemical essential for crops: nitrogen. Most of the nitrogen farmers add to their fields comes in the form of synthetic fertilizer—which suffers from enormous price volatility, and also happens to generate 5 percent of global greenhouse gas emissions. Pivot Bio, which we helped launch in 2014, sells crop additives containing microbes that take up residence in the root systems of corn, wheat, and small grain crops and turn atmospheric nitrogen into ammonia that plants can use to grow. Farmers using Pivot's PROVEN 40 product can maximize crop yields while using up to 40 pounds per acre less synthetic fertilizer. After the conflict in Iran shut down shipping lanes and caused fertilizer prices to soar, Pivot introduced a price assurance program that lets farmers lock in a consistent price for its products for the 2026, 2027, and 2028 growing seasons.

The theme here is that future fuels and other essential materials won't come just from mining or fossil sources, but from electrochemical factories and biology as well. One of those is magnesium, which, as we noted above, is critical for the defense, aerospace, automative, and other industries. Almost all magnesium is mined and refined in China—and when authorities there cut magnesium production to save energy in 2022, the result was an immediate, global price spike and supply shortages. But Tidal Metals, which we've backed since 2023, has created a breakthrough technology for accessing an essentially inexhaustible magnesium supply: the world's oceans.

Dow Chemical harvested magnesium from seawater for over 60 years, but its process could not compete with low-cost Chinese labor and energy and China's weak regulatory environment. Tidal has reimagined the problem and can now pull magnesium from seawater at a globally leading price point, with zero environmental harm. The company precipitates a hydrated magnesium chloride salt from seawater and industrial brines using a thermally driven separation system called a temperature-swing adsorption vapor pump. It dehydrates this salt using a patented, proprietary process and then electrolyzes the anhydrous magnesium chloride into magnesium metal and a valuable chlorine co-product.

Another approach to mineral recovery comes from DCVC portfolio company Alta Resource Technologies, which describes itself as a "precision mining" company. It doesn't take anything out of the ground: instead, it separates critical minerals from many novel feedstocks using customized metal-binding proteins. One protein might bind selectively to neodymium in e-waste; another might bind to dysprosium. To capture the targeted metals, Alta passes the bound proteins through a standard column-chromatography-style separation system that isolates the desired elements. Unlike traditional mineral separation, Alta's technology avoids large volumes of toxic solvents and can operate on a wide range of feedstocks—including e-waste, ores, brines, mine tailings, and industrial wastewater. The U.S. is home to an abundance of such waste streams, meaning Alta offers yet another path to supply-chain resilience.

What ties these companies together is a simple idea: the next generation of materials supply chains will be designed not around combustion and cracking, but around precision chemistry. By controlling how molecules are separated and reassembled—using catalysts, electrons, modular reactors, engineered organisms, and novel feedstocks—the innovators we back today are creating cleaner, more resilient supplies of the fuels, materials, and minerals modern economies depend on. The collective opportunity here is significant.

Wastewater treatment can unlock new supplies of clean water and critical materials.

Globally, 360 trillion liters of wastewater are produced every year, much of it containing dissolved minerals, nutrients, and energy-rich organic compounds. We rarely think of this waste as a resource—more often, it's treated as a liability to be disposed of. But new chemical and biological tools are making it possible to convert the growing flows of dirty water from industry, agriculture, and cities into productive industrial inputs. That process begins with separating and concentrating the valuable substances found in industrial wastewater, desalination brines, and produced water.

The most important energy asset in the United States is the Permian Basin, which produces nearly 50 percent of the nation's oil and 20 percent of its natural gas. But the basin has a big problem—water. For every barrel of oil extracted, wells yield an additional 3 to 13 barrels of so-called produced water, totaling nearly 24 million barrels every day. Disposing of this water is becoming an acute problem that could threaten the long-term viability of the basin.

Produced water is salty—often three to five times more salty than seawater. The practice in the industry has been to re-inject this water back into the ground in saltwater disposal wells. Over time, this has led to several negative outcomes, including pressurizing the formation and generating earthquakes. As a result, the Railroad Commission of Texas, which regulates the state's oil and gas, mining, and pipeline industries, is tightening the permitting process for new disposal wells.

"The net effect is that we're running out of places to go put produced water," says DCVC Operating Partner Earl Jones. "This means that the marginal cost of new produced water disposal is going up. The math is simple: if disposal costs go to $2 per barrel, and we're producing five barrels of water for each barrel of oil, that's a $10 per barrel equivalent. This is an existential threat and the industry has recognized that it has to do something about produced water."

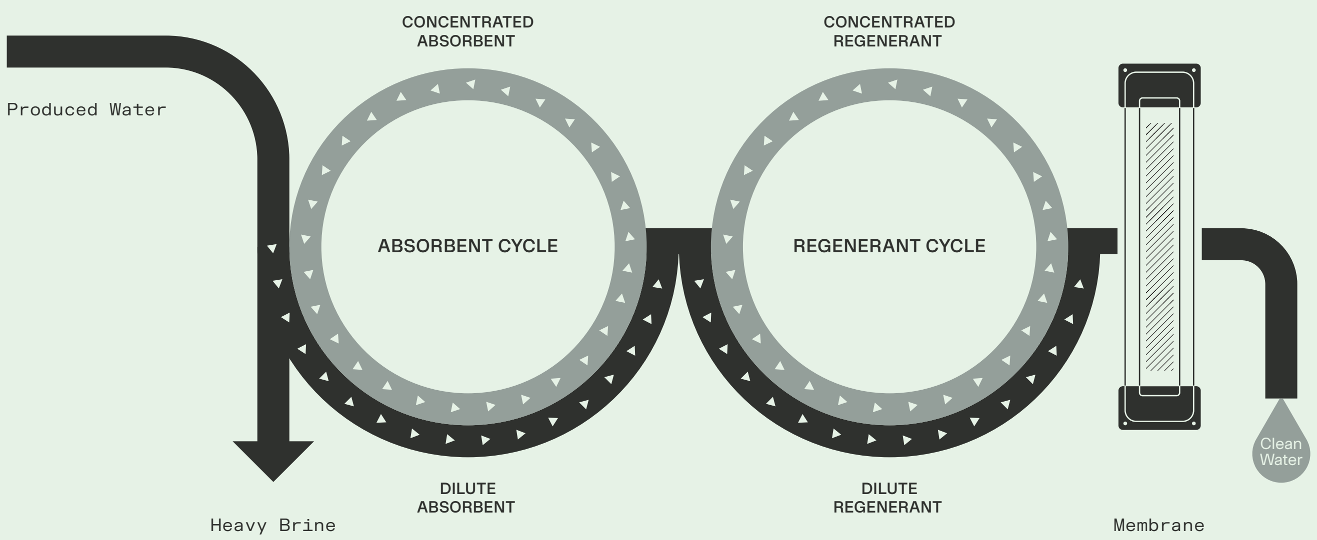
← Much of the brackish, saline water extracted at oil and gas wells ends up in evaporation ponds that dot the Permian Basin.

↓ **Aquafortus** built one of its first demonstration desalination plants for produced water in Colorado City, Texas.

Figure 4.3.1

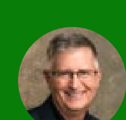
In Aquafortus' ABX process, an absorbent soaks up water from brine and rejects everything else. The water is then transferred to a regenerant, which travels through a membrane where clean water is recovered. Both the absorbent and regenerant are fully recovered and re-used.

Source: Aquafortus.



"

The joke is that the entire periodic table resides in produced water. There are valuable resources like lithium, iodine, gallium, magnesium, and ammonia, all of which can be recovered more economically at higher concentration when you're working with concentrated brine."



Earl Jones
Operating Partner, DCVC

Enter Aquafortus, a company we've backed since 2021. It has developed an inexpensive, scalable solution for desalinating heavy brines that halves the amount of produced water that must be re-injected underground, while simultaneously recovering fresh water and a range of valuable materials.

The high salinity and large volumes of produced water are a challenge for existing treatment technologies like membrane filtration and thermal evaporation, both technically and economically. Aquafortus' approach, called ABX, is to desalinate brines using a chemical mixing process called solvent extraction. It begins by mixing brine with a novel chemical absorbent, which acts like a liquid sponge, pulling water out of the brine while rejecting salts and other chemicals. The now dilute absorbent is then mixed with a second chemical called a regenerant, which releases the absorbent for re-use. In the final step of the process, the dilute regenerant is concentrated using membranes and also recovered for re-use, leaving high-purity water for beneficial use.

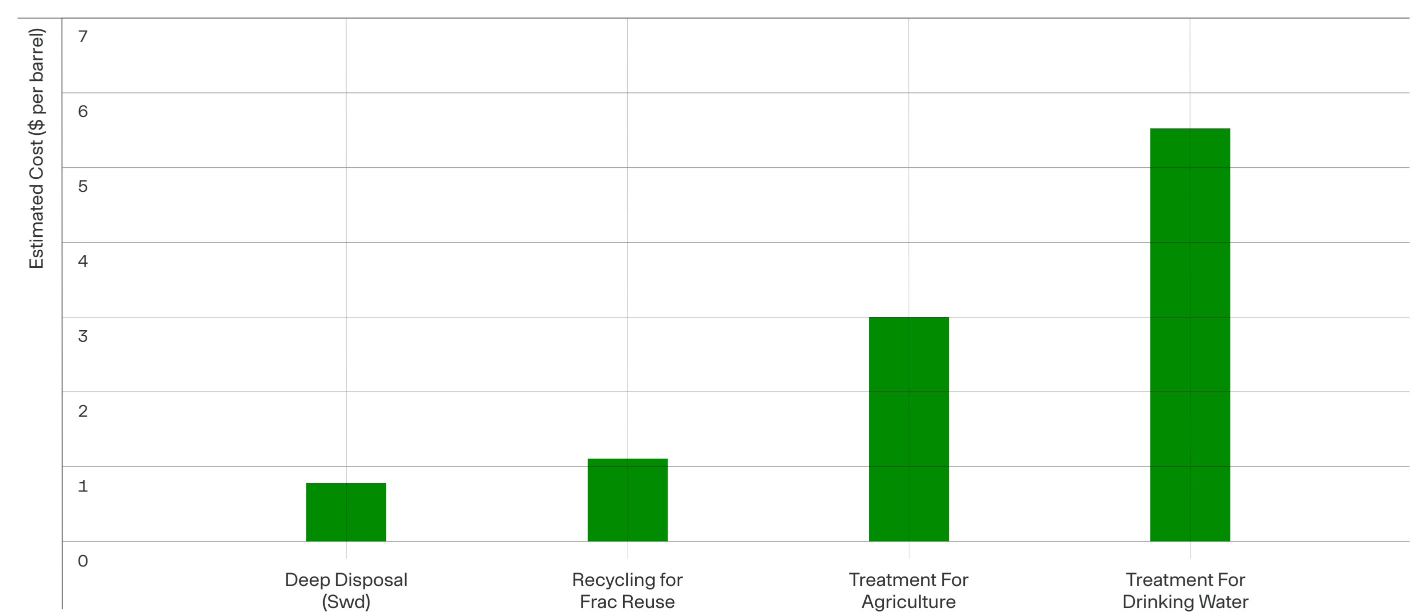
Texas and New Mexico, the epicenter of the produced water problem, are also seen as ideal locations for new factories and data centers—both water-intensive applications. But population growth and prolonged drought have dramatically depleted conventional freshwater resources. Desalinating produced water could solve this twin problem, by generating over 100 billion gallons per year of clean water while at the same time reducing disposal volumes. At its showcase facility in Colorado City, Texas, Aquafortus has already desalinated tens of thousands of barrels of produced water, and the company has a large pipeline of projects across the oil and gas, mining, and manufacturing industries.

In a very real sense, Aquafortus is not a desalination company, but a resource recovery platform. Performing solvent extraction on industrial brines can recover not only clean water but many critical metals and minerals. "The joke is that the entire periodic table resides in produced water," Jones says. "There are valuable resources like lithium, iodine, gallium, magnesium, and ammonia, all of which can be recovered more economically at higher concentration when you're working with concentrated brine." Put all that together, he says, and "you unlock an unbelievable economic story."

Figure 4.3.2

The cost of managing produced water in Texas (2025 estimates)
About 85 percent of produced water in the Permian Basin is injected into saltwater disposal wells. SWD remains cheaper than recycling or treatment—but the cost gap is closing.

Source: Texas A&M University.



CHAPTER 5

Computing-heavy, data-rich strategies are transforming drug discovery.

Biology and human health have historically been too complex to model accurately and too expensive to experiment on exhaustively. Drug discovery and medical innovation therefore relied on slow, costly, high-risk trial and error. Today those constraints are breaking down. Biology and medicine can increasingly be understood through closed-loop computation and experimentation.

WINNING TECHBIO COMPANIES DO THREE THINGS REPEATEDLY, IN A CLOSED LOOP:

1. Model biology
2. Generate new biological data
3. Improve the relevant models.

Out of that cycle come predictions that clarify biological networks, reduce the risks inherent in drug discovery, and speed the diagnosis of disease. None of this is an accident: it's happening today because of the converging revolutions in abundant omics data, foundation models in AI, and cheap automated lab infrastructure. (For more about the latter see the 2025 Deep Tech Opportunities Report, pages 82–83.) Together, these technologies are making biology more computable, drug modalities more designable, and medical care more measurable.

We believe that 20 years from now, or even sooner, most or all drug and healthcare companies will also be software companies. We've already seen how the most powerful deep tech businesses engineer breakthroughs in cost and efficiency by integrating computation with physical systems in energy (Chapter 3) and infrastructure (Chapter 4). In TechBio, computation is meeting living systems—including the human body.

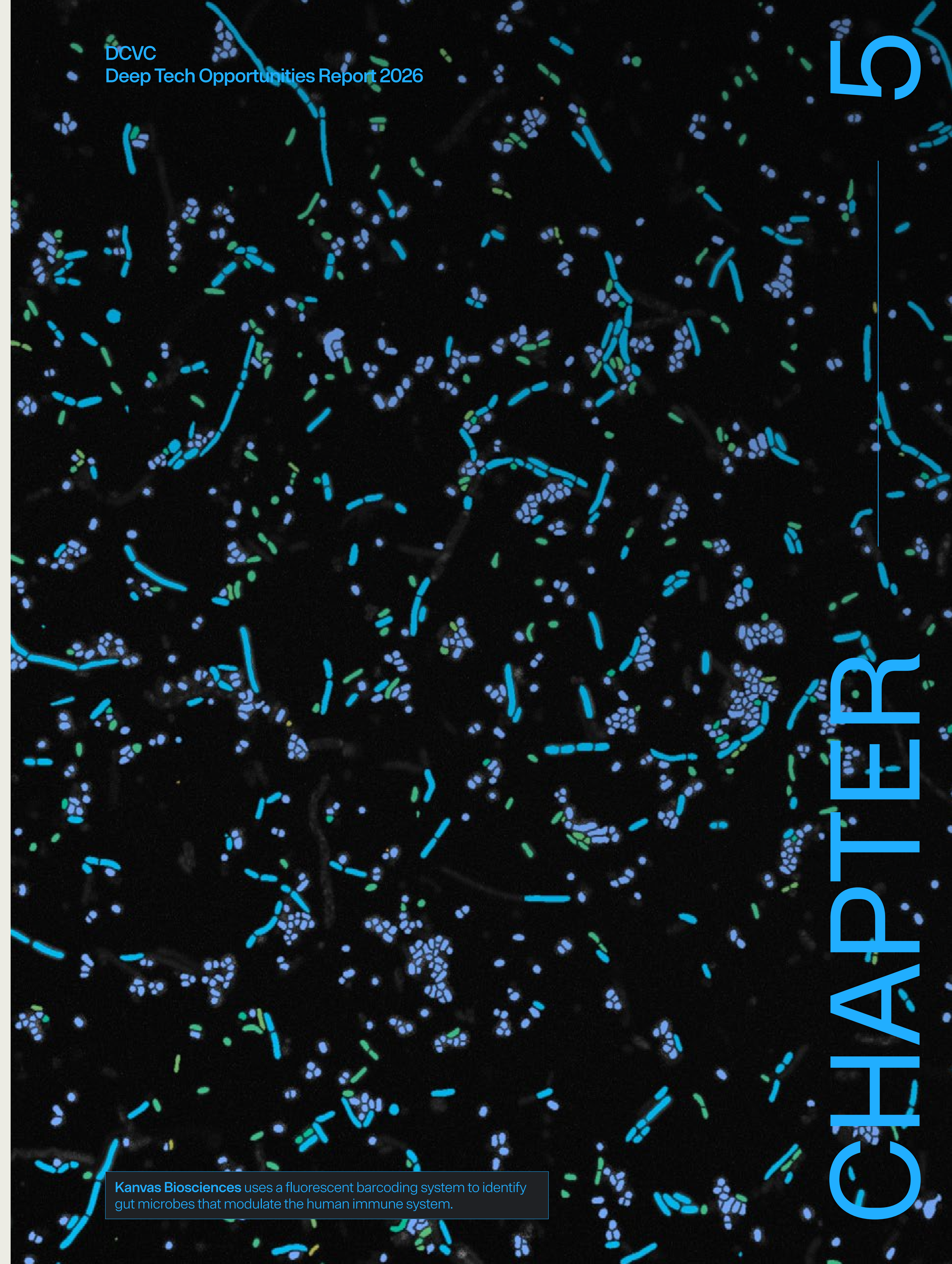
OPPORTUNITIES

COMPANIES

Auron Therapeutics, Kanvas Biosciences, Noetik, Plexium, Proxima, Recursion, Relation

VOICES

Jason Pontin, James Hardiman, Dr. Emily Park



Kanvas Biosciences uses a fluorescent barcoding system to identify gut microbes that modulate the human immune system.

State-of-the-art computing models and proprietary data generation are opening up new frontiers for TechBio companies.

It's possible to discover new cures for untreated diseases from first principles by integrating single-cell analysis, genomics, and machine learning in a virtuous cycle now widely known as the "lab-in-the-loop" model. That's the discovery approach at almost all of the TechBio companies we back. Their foundation AI models and the biology atlases they're building set the stage for early pharma partnerships worth billions of dollars, because they can solve problems big biopharma can't solve by itself.

Noetik uses powerful AI to analyze tumor core slices for telltale patterns of morphology and protein expression.

There's something a little magical about combining large-scale biological modeling with automated experimentation. When companies do this, they can stop screening thousands of molecules blindly. Instead they can continuously generate hypotheses about how genes and proteins interact to shape the destinies of cells, test those hypotheses in the lab, and feed the results back into their models. Keeping the lab in the loop is key, because it ensures that computational predictions are grounded in real biological data. The result is a discovery process that becomes faster and more informative with every experiment.

One of the earliest and most influential examples of this model is Utah-based Recursion (Nasdaq: RXRX). In 2016 DCVC made its initial investment in the company, which was one of the first to operationalize the lab-in-the-loop paradigm in drug discovery, in the service of an audacious vision to build virtual cell models that can be used to predict and explain the effects of drug interventions and generate testable hypotheses about novel treatments.

To realize that vision, Recursion, which had its IPO in 2021, worked with NVIDIA to build BioHive-2, the largest and fastest supercomputing cluster owned by a biopharma company. It's built to run Recursion OS, an automated platform that orchestrates the company's entire drug discovery process. That spans every step from collecting phenomic, genomic, and patient data to running the virtual-cell model to validating predictions in silico, designing new molecules, testing them in the lab, and recruiting patients for clinical trials.

One of the seven drug candidates in Recursion's pipeline, REC-1245, targets an RNA-binding protein involved in cellular DNA damage response. By degrading that protein, the drug could selectively kill genomically unstable tumor cells. Using Recursion OS, the company was able to move from initial identification of REC-1245 to clinical trials in just 18 months, compared to the industry average of 42 months. That kind of speed-up gives drug developers more shots on goal—and decreases their reliance on any single asset.

But while Recursion is the canonical example of a drug discovery company built around a lab-in-the-loop cycle, a new generation of TechBio companies is extending this approach even further. Relation, a London-based company that we've backed since its seed round in 2021, has perfected a version of the model that starts with single cells derived from patient tissue—for example, the osteoblasts of people with osteoporosis. It builds proprietary atlases of omics data on these cells—data that informs AI-generated maps of disease biology.

This, in turn, yields predictions about the kinds of perturbations that might affect the course of a disease. Relation then tests those perturbations in the wet lab, uses the resulting data to refine the maps, and so on, around and around.

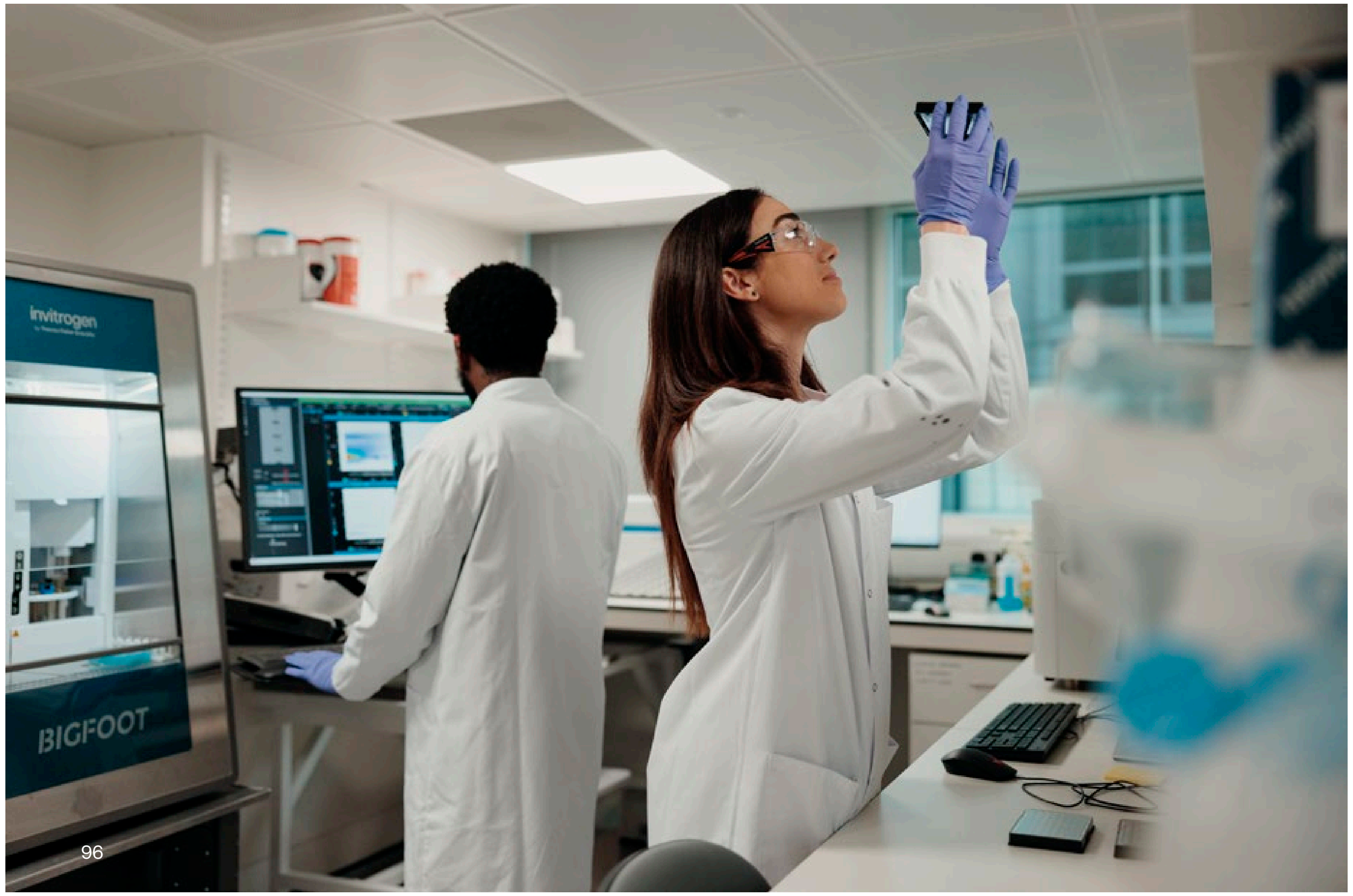
The company is advancing its own drug candidates against bone disease and has struck blockbuster deals and research collaborations with drug-industry titans. GSK, for example, has agreed to pay $108 million up front and after early milestones to treat fibrotic disease and osteoarthritis—plus up to $200 million per target for preclinical, development, commercial, and sales milestones, along with tiered royalties on net product sales. Meanwhile, Novartis will pay $55 million up front for research on atopic diseases, plus up to $1.7 billion in downstream milestone payments tied to specific drugs, as well as royalties.

Such deals are attractive for Relation and its investors because they provide proof of concept while shifting much of the financing burden away from equity fundraising. And they make sense for big pharma because they lead to a deeper understanding of disease biology while efficiently generating far more drug candidates than large pharmaceutical organizations are equipped to develop on their own.

“Here’s a nuance that is often missed: greater productivity means lower risk at scale,” says DCVC General Partner Jason Pontin. “You increase the rate of innovation because when you fail, you know you’ve failed more quickly. Relation can run multiple experiments at once, because it’s cheap, and they can find a novel drug target in weeks. Then they can generate novel chemistries and put them into cells and see if they work, remarkably rapidly.”

Another DCVC-backed company, Noetik, is pushing the lab-in-the-loop paradigm in a different direction: using large foundation models to understand spatial patterns of gene expression, protein signaling, and cellular structure in tumors, and then predict how patients will respond to new cancer therapies before those drugs ever reach large clinical trials.

↓ **Relation** A lab-in-the-loop process integrates single-cell analysis, genomics, and AI to find new drug targets.



“

Here’s a nuance that is often missed: greater productivity means lower risk at scale. You increase the rate of innovation because when you fail, you know you’ve failed more quickly. Relation can run multiple experiments at once, because it’s cheap, and they can find a novel drug target in weeks. Then they can generate novel chemistries and put them into cells and see if they work, remarkably rapidly.”

Jason Pontin
General Partner, DCVC

The work starts in Noetik’s wet lab, where it collects slices of tumors from thousands of cancer patients to create histology slides. These images become the training data for OCTO-VC, Noetik’s multimodal foundation model. Using a self-supervised approach similar to that used in large language models, OCTO-VC breaks down each image into tokens representing small patches of cells. By masking and reconstructing these patches during training, the model learns how the local neighborhood of each cell—nearby immune cells, fibroblasts, blood vessels, and tumor cells—influences its biological behavior.

Once trained on billions of such tokens, OCTO-VC can begin to simulate unknown biology and answer clinically important questions, such as: Which changes in gene expression might make tumor cells more susceptible to attack by the immune system? Which patients are most likely to respond to immunotherapy?

In January, GSK and Noetik unveiled a five-year, $50+ million agreement under which GSK will pay up-front capital fees as well as annual subscription fees to access OCTO-VC in the development of drugs for non-small-cell lung cancer and colorectal cancer. “As far as we know, it’s the first example of a substantial pure-play AI licensing deal in TechBio,” says DCVC General Partner James Hardiman. “That makes it a kind of existence proof for this new business model.”

The deeper lesson from companies like Recursion, Relation, and Noetik is that once biology enters a closed computational loop, discovery accelerates and the economics of drug development begin to change. That shift benefits both the big pharma companies lining up to work with these TechBio specialists—and ultimately the patients waiting for better therapies.

AI-designed molecular glues could turn previously undruggable proteins into targets— and replace many injectable biologics with simple oral drugs.

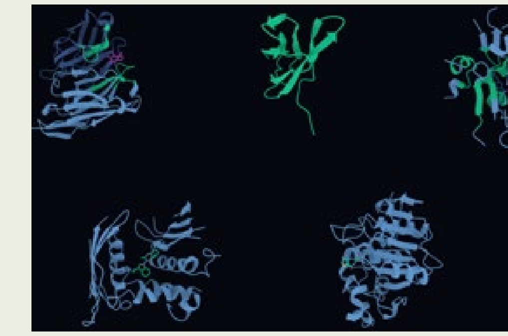
Once we can systematically map disease biology and run lab-in-the-loop discovery cycles, we can begin programming entirely new drug mechanisms. One of the most promising frontiers is protein-protein interactions.

Many effective drugs focus on proteins as isolated targets, altering the course of disease by blocking or activating their function. But inside living cells, proteins rarely act alone—instead they assemble into dynamic complexes with other proteins, transmitting signals that control processes such as immune responses, cell growth, and metabolism. A new generation of TechBio companies is beginning to exploit this deeper level of biology by designing molecules that can bring two proteins together—or force them apart—to trigger novel cellular behaviors.

Mastering protein-protein interactions is a huge therapeutic opportunity. A number of molecular glue molecules, such as Revolution Medicines' daraxonrasib, are already advancing in clinical trials. (That molecule works by slipping between two proteins called cyclophilin A and RAS to block cancer signaling.) A related drug category called PROTACs —for proteolysis targeting chimeras, pairs of ligands that bring target proteins into contact with ligases that break them down—is also booming, with dozens of drugs now in clinical trials. The limitation with both kinds of drugs is that there's been no systematic way to design them. They tend to be discovered serendipitously, or through laborious trial and error.

Early this year we invested in Proxima, a company using a series of frontier AI models called Neo to change that reality. Proxima starts by measuring how proteins interact in the real world. In the lab, it assembles vast numbers of proteins, enzymes, and transcription factors—currently covering 70 percent of the human genome—and digests them down to the peptide level to see where and how these protein fragments naturally link up. Using a technique called cross-linking mass spectrometry (XLMS), it captures snapshots of these diverse interactions at the interatomic level. These measurements of molecular structure become Neo's training data, giving it the ability to understand larger molecules and carry out generative design tasks. Starting from the amino-acid sequences of two proteins, for example, Neo can find pockets on both proteins where small molecules might bind, and generate a hypothetical structure for a ligand that will program their interactions.

That ability—to simultaneously simulate how proteins co-fold and design new ligands that precisely control them—is what makes Neo so powerful. "Once you can do that, you can do three things," explains DCVC General Partner Jason Pontin. "You can create so-called inducers, which bring two proteins into proximity. Or you can make blockers, which prevent protein-protein interactions. Or you can make modulators, to modify interactions. So you're doing what biologic drugs have historically done, but you're doing it with small-molecule chemistry inside the cell."



↑ **Proxima's** Neo model can design proteins, peptides, or small molecules from scratch.

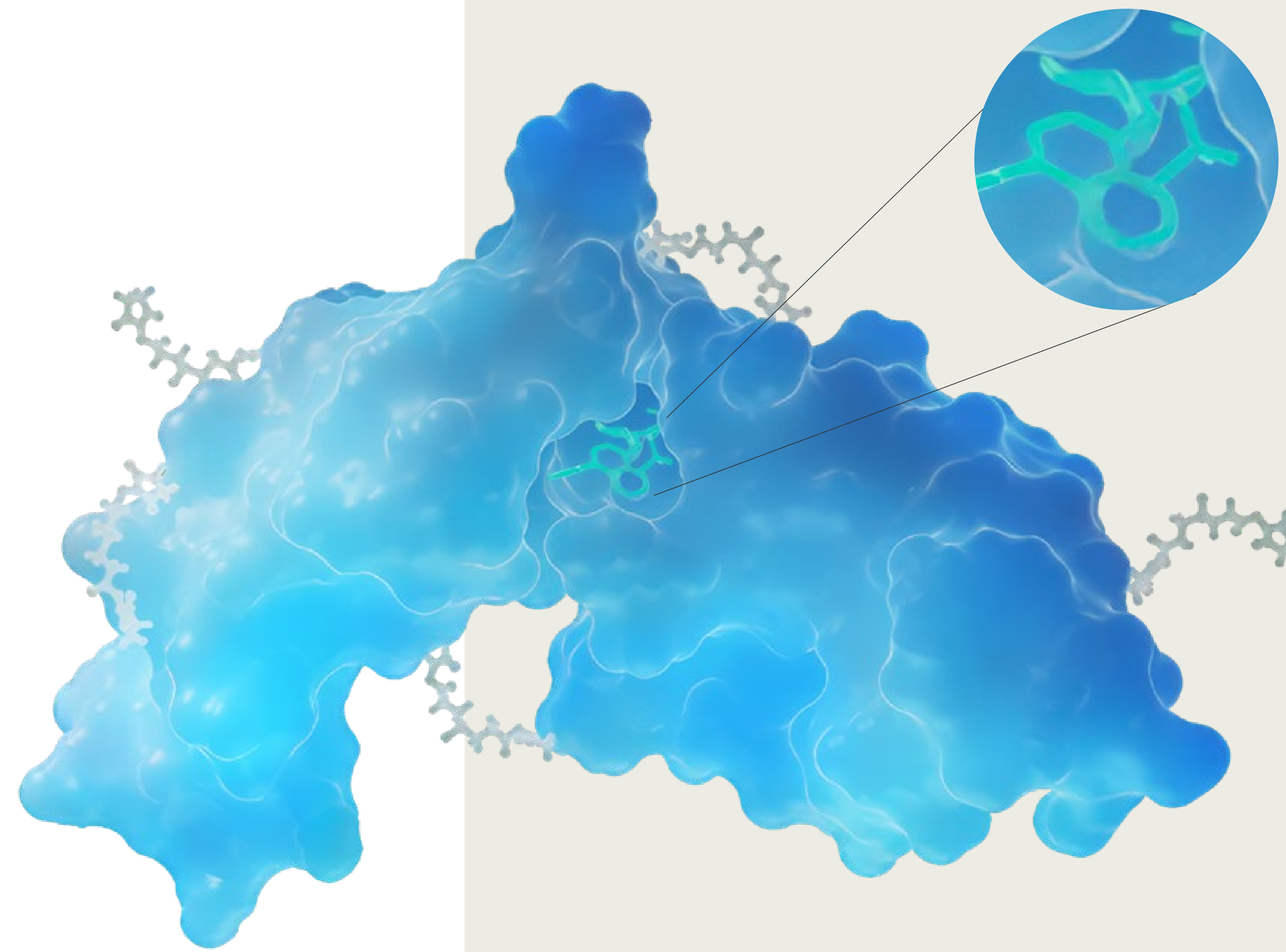
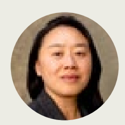


Figure 5.2.1

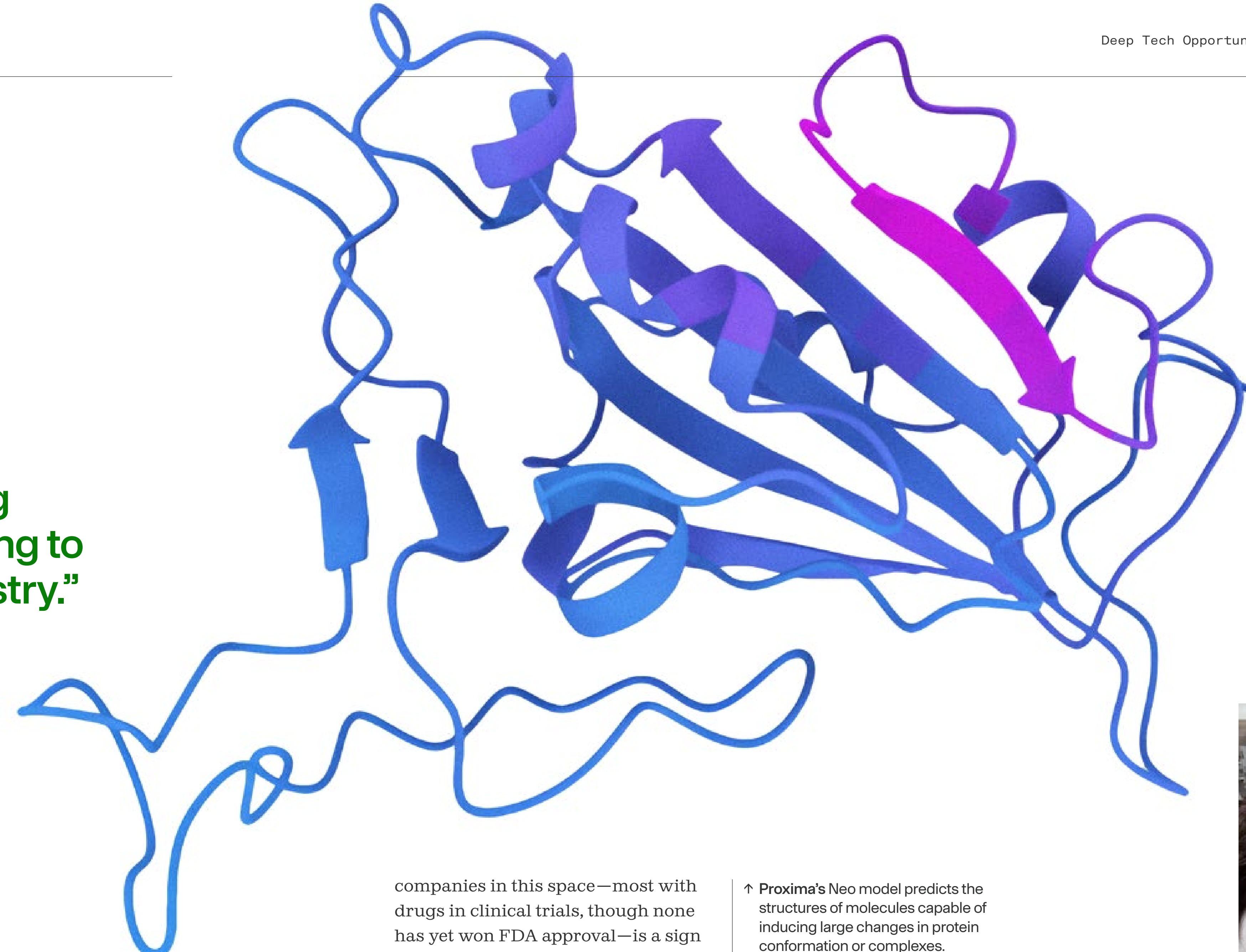
AI-designed molecular glues
A molecular glue is a small ligand molecule (green) that helps proteins fit together, unlocking new functions. Proxima uses a vast proprietary atlas of protein interaction data to train AI to design new ligands.

↑ **Auron** looks for protein degrader drugs that could knock tumor cells out of their proliferative, plastic states and inhibit their growth.

> “When the first oral protein degrader drug gets approved, it’s going to unlock an entire industry.”



Dr. Emily Park
Principal, DCVC Bio



↑ **Proxima’s** Neo model predicts the structures of molecules capable of inducing large changes in protein conformation or complexes.

→ **Auron** has identified protein degraders that could treat several forms of cancer as well as inflammatory disease.

Already, Proxima is putting this technology to work in partnerships with some of the largest companies in biopharma. Collaborations with Johnson & Johnson, Bristol Myers Squibb, and Blueprint Medicines—now part of Sanofi—focus on identifying new proximity-modulating drug candidates for cancer and other complex diseases. If the resulting therapies reach the market, the upfront payments, research funding, and milestone-based rewards from these deals alone could add up to billions of dollars. On top of that, the company is growing its team of researchers, physicians, and programmers, who aim to build an internal pipeline of protein-modifying drugs.

But Proxima has plenty of company in its search for proximity-based therapeutics. Plexium and Auron Therapeutics, two companies backed by DCVC Bio (see Chapter 6), show how TechBio and traditional biotech sometimes interweave. Auron has two AI-designed drugs approaching clinical trials—against acute myeloid leukemia, myelodysplastic syndrome, small-cell lung cancer, and prostate cancer—that are both protein degraders, selectively targeting pathogenic proteins for elimination. The company’s drug AUTX703, for example, targets enzymes called lysine acetyltransferases that push cells into a more proliferative, plastic state. Plexium, meanwhile, is in Phase I trials with two small-molecule drugs, a degrader and a molecular glue. The degrader targets a protein called SMARCA2 that some cancer cells depend on to help regulate gene expression, while the glue targets a transcription factor called IKZF2 that helps tumors evade immune cells.

And Plexium and Auron themselves have competition from degrader companies like Arvinas, Prelude Therapeutics, Foghorn Therapeutics, and Monte Rosa Therapeutics, not to mention all the companies developing PROTACS. But to Dr. Emily Park, a Principal at DCVC Bio, the profusion of companies in this space—most with drugs in clinical trials, though none has yet won FDA approval—is a sign that a new generation of drugs against cancer, immunological disorders, and other undertreated diseases is about to emerge.

“These drugs do something similar to antibodies, but you can make oral versions,” Park says. “That’s why I think it’s really exciting.” She points out that Dupixent, an injectable or intravenously delivered antibody drug made by Regeneron and Sanofi for the treatment of atopic dermatitis and some types of asthma, works by blocking interleukin-4 and interleukin-13 signaling, which are involved in inflammatory responses. “The thesis here is that using a degrader drug, a PROTAC, you could make an oral version, and just bring over a ligase to degrade the target,” Park says.

The market pull for cheaper oral equivalents of biologic drugs is enormous: without insurance, Dupixent costs around $1,500 per monthly or semimonthly dose, indefinitely, and in 2025 the drug earned Regeneron and Sanofi $17.8 billion. “When the first oral protein degrader drug gets approved, it’s going to unlock an entire industry,” Park predicts.

It’s easy to see how oral versions of injectable biologics could lower the cost and complexity of existing treatments for millions of patients. But what’s enabling that shift is just as exciting. On top of that, companies like Proxima, Auron, and Plexium are combining AI with lab experimentation to come up with small-molecule drugs that have no biologic precedent—and that will dramatically expand the range of treatable diseases.

New imaging, AI, and manufacturing tools are finally making the microbiome druggable.

For years, the microbiome has been one of medicine's most tantalizing but uncrossable frontiers. Researchers have long known that the trillions of microbes living on and inside the body influence immunity, metabolism, and development. But until recently they lacked a reliable way to see how particular microbes interacted with host cells, identify which strains actually mattered, or turn those insights into manufacturable drugs. As a result, most microbiome therapies were built on guesswork. Kanvas Biosciences represents a different approach: Modeling and manufacturing the microbiome in a way that may transform cancer treatment (for starters) and global health.

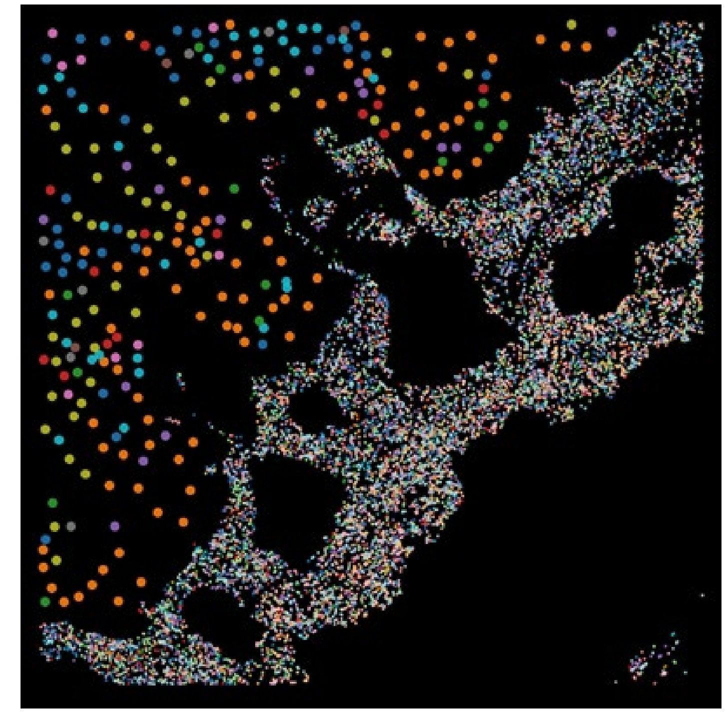
Microbial cells in the gut and on the skin make up only 1 to 3 percent of human body mass, but there are roughly 40 trillion of them—outnumbering the body's 30 trillion human cells—and they collectively carry over a hundred times more genetic information. In an important sense, humans are not individuals but ecosystems. So when our microbiome is out of balance, we are too.

Yet physicians aiming to modify the microbiome have often been stumbling in the dark. Early microbiome therapies relied on crude fecal microbiota transplants (FMTs), transferring entire microbial communities without understanding which organisms were responsible for the therapeutic benefit. FMT is now recognized as an effective treatment for recurrent *Clostridioides difficile* infections, but it is by no means a scalable or precision therapy. Even when promising strains were identified, manufacturing reliable "live biotherapeutic products" proved difficult, because microbial communities are complex, variable, and hard to standardize.

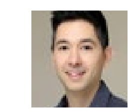
Kanvas, which we've backed since 2021, was founded on the premise that seeing the microbiome and its functions clearly is the first step to making it engineerable. Making that work required several platform innovations developed over the course of half a decade:

→ Spatial labeling and decoding. Kanvas' scientific leaders developed a way to label RNA from microbial and host cells with fluorescent binary barcodes, then use spectral imaging and machine learning to untangle the multiplexed signals and generate detailed 3D maps of host-microbiome interactions.

↓ **Kanvas Biosciences**' labeling system pinpoints specific RNA transcripts from host and microbial cells, offering a window into their interactions.



> "The data coming out of this microscope will allow us to chart out the first true host-microbiome atlas."



Dr. Matt Cheng
Co-Founder and CEO of Kanvas

→ Industrial-scale imaging. Later the company built a dedicated lightsheet microscope capable of scanning thin slices of labeled biopsy tissue at unprecedented speed, allowing Kanvas to expand its data collection efforts to industrial scale. "The data coming out of this microscope will allow us to chart out the first true host-microbiome atlas," says Dr. Matt Cheng, Co-Founder and CEO of Kanvas. That atlas allows Kanvas to identify which microbial strains are most active in the guts of healthy human "superdonors"—and which genes they're expressing.

→ Precision microbiome manufacturing. Finally, Kanvas built an advanced anaerobic fermentation facility in South San Francisco, giving it the ability to manufacture live biotherapeutic products (LBPs)—complex cocktails of live microorganisms derived from superdonor strains—consistently and at scale.

The platform is already yielding candidate LBPs that could benefit thousands of cancer patients. A class of drugs known as immune checkpoint inhibitors (ICIs) can produce near-miraculous results in patients with melanoma, non-small-cell lung cancer, renal cell carcinoma, and other cancers. Merck's market-leading ICI, Keytruda, is the world's top-selling drug, with revenue of $31.7 billion in 2025. Yet only 20 to 40 percent of patients respond to these therapies.

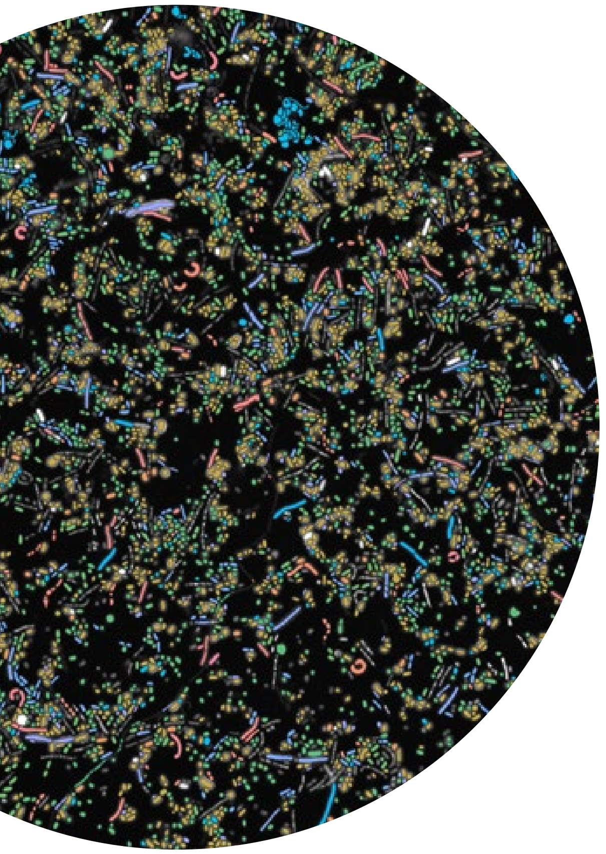
Researchers have known for several years that patients who respond well to ICIs tend to have higher baseline microbiome diversity, along with higher abundance of specific microbes such as *Akkermansia muciniphila*. Early studies combining ICIs with FMTs have shown positive promising results, suggesting that the microbiome plays an important role in determining who benefits from immunotherapy.

Kanvas built on those insights by studying the microbiomes of high-responding "superdonors" and designing an oral LBP, called KAN-001, that boosted the effects of ICIs against non-small-cell lung cancer in preclinical studies. In collaboration with the MD Anderson Cancer Center, Kanvas plans to move the therapy into Phase I clinical trials this year. (A separate drug, KAN-004, is designed to protect patients from colitis, a common side effect of checkpoint inhibitor therapy, and allow them to stay on the therapy longer. Kanvas recruited KAN-004's first clinical trial patients in April.)

“What Kanvas and MD Anderson have discovered is that you may be able to double the number of patients who respond to immune checkpoint inhibitors, if you modulate the response through the microbiome,” says DCVC General Partner Jason Pontin. “If it works, Kanvas will have both done a mitzvah and created a blockbuster drug.”

But Kanvas isn’t just a cancer company. Once you know how to engineer the microbiome, you can begin to tackle disorders of the gut itself—potentially including a condition called environmental enteric dysfunction (EED) that is a persistent health problem for low- and middle-income people in the developing world. Familiarly known as tropical sprue, or, in the past, as dysentery, EED is thought to be caused by fecal-oral contamination from unsanitary water and food. It leads to inflammation of the gastrointestinal tract, a lowered ability to absorb nutrients, impaired immune responses, stunted growth, and long-term developmental deficits. Some 150 million children are at risk worldwide. In pregnant mothers, EED is a major contributor to undernutrition, low infant birthweight, and higher neonatal health risk.

↓ **Kanvas** can use its visualization system to measure the abundance of different microbial strains in its bioreactors.



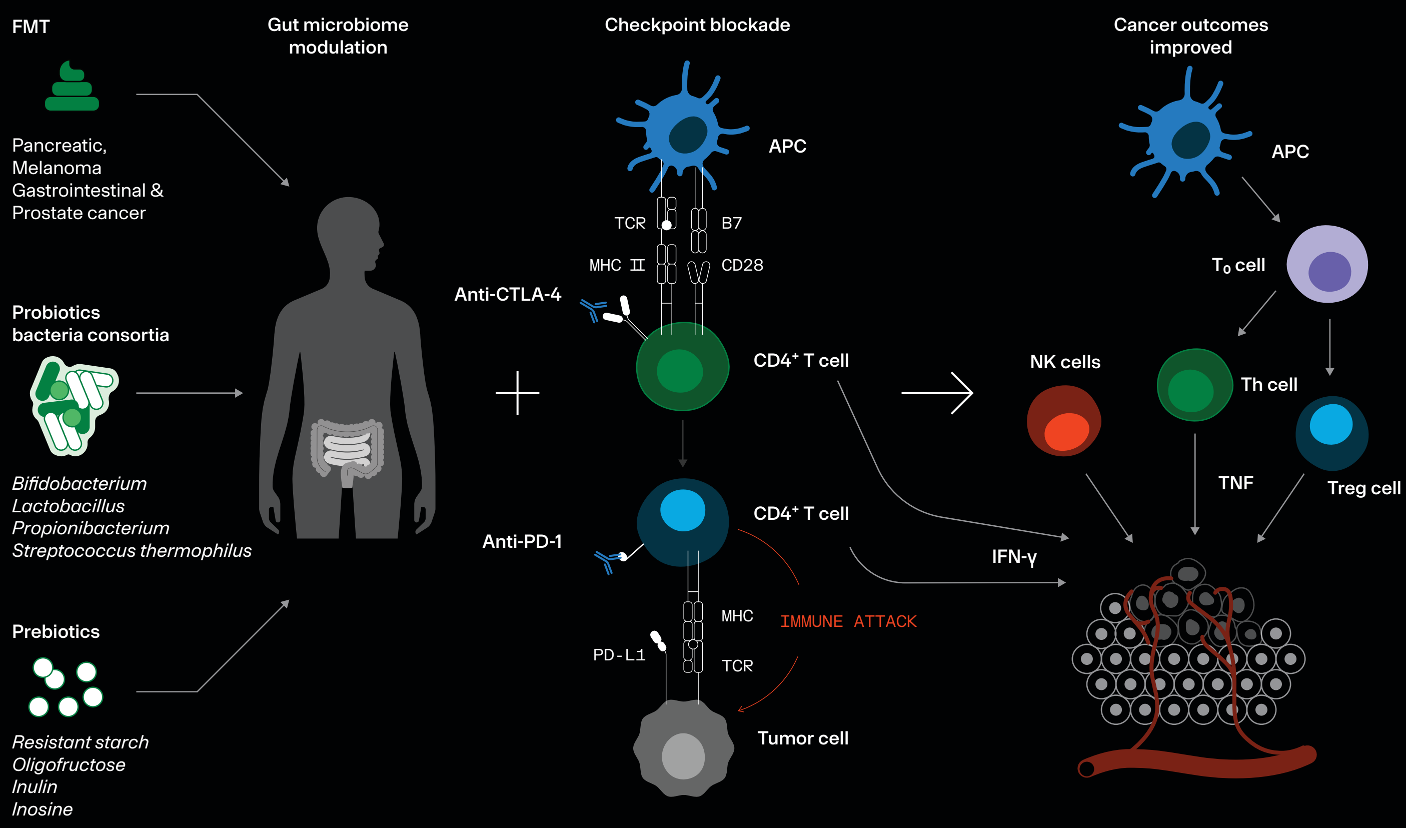
> What Kanvas and MD Anderson have discovered is that you may be able to double the number of patients who respond to immune checkpoint inhibitors, if you modulate the response through the microbiome. If it works, Kanvas will have both done a mitzvah and created a blockbuster drug.”

Jason Pontin
General Partner, DCVC

Figure 5.3.1

Modulating the gut microbiome to improve cancer immunotherapy
Introducing tailored bacterial consortia into the gut alongside checkpoint inhibitors can reprogram the tumor microenvironment and make it easier for immune cells to do their work.

Source: eBioMedicine.



The Gates Foundation has long invested in programs aimed at improving access to clean water and reducing the burden of infectious diseases in the world’s poorest regions. Now it is also targeting EED through a direct investment in Kanvas. This spring the foundation contributed to Kanvas’ $50 million Series A fundraising round (co-led by DCVC), helping support the development of a fully synthetic LBP designed to treat and prevent EED in pregnant women. The strategy is to use global datasets and biological samples to design complex microbial mixtures—with strains tailored to specific regions—that can restore gut health in both women and children.

“If you could treat this—for me personally, and for the infectious disease doctors at the Gates Foundation, and for Matt Cheng, who’s an infectious disease doctor—this would be your legacy,” Pontin says. “This is what you put on your damn tombstone.”

The cancer and global-health programs at Kanvas both emerge from the company’s engineering mindset and its methodical quest to observe, model, and manufacture complex microbial ecosystems. The era when the microbiome was a black box—amenable only to imprecise interventions—is finally ending.

CHAPTER 6

DCVC Bio invests in modalities that could unlock novel classes of medicines or food technologies.

The team here at DCVC Bio, the firm's dedicated life sciences fund, doesn't bet on specific molecules or drug candidates. We bet on scientific leverage points—biological or engineering insights that can create new paradigms for drug discovery, new ways of targeting undertreated conditions, new platforms for delivering therapies, or new approaches to growing food.

In essence, we're looking for new control knobs for biology and agriculture. And to intelligently adjust those knobs, we assemble and invest in the most talented science and engineering teams we can find. In this year's report, we'll talk about DCVC Bio companies looking to

- **control where biologic drugs can reach**, through improved antibody targeting
- **control how we harness the human immune system's own ingenuity**, by studying antibodies in elite responders
- **control where therapies are manufactured**, by enabling in vivo CAR-T cell engineering
- **control the chemistry applied to crops**, through real-time sensing and precision spraying.

Biology is chaotic. Traditional biotechnology companies try to find molecules that nudge the mess in the right direction. DCVC Bio companies try to rewrite the rules that govern it.

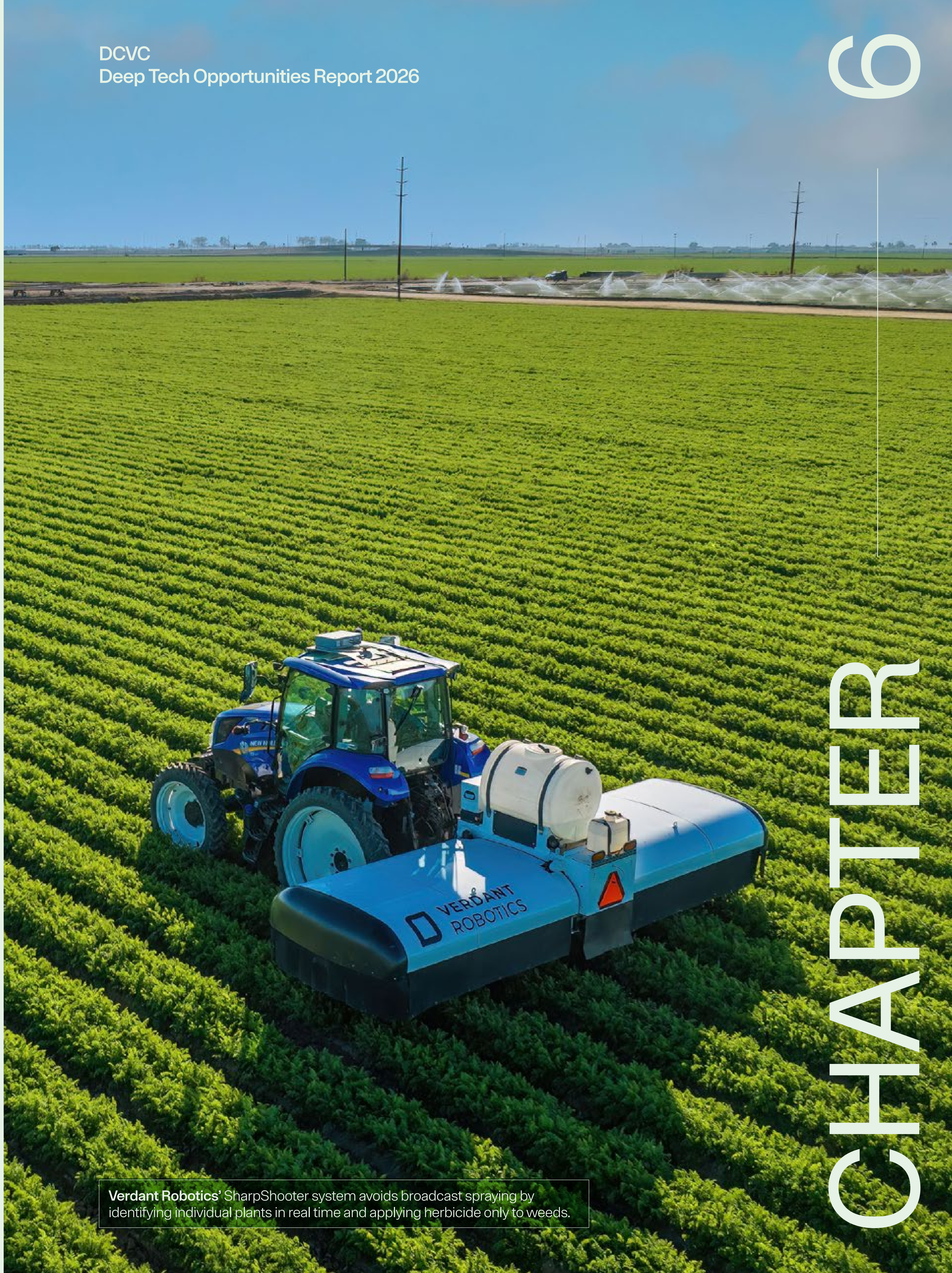
OPPORTUNITIES

COMPANIES

AgZen, Alchemab, Sabanto, Solu Therapeutics, Umoja Biopharma, Verdant Robotics

VOICES

Dr. John Hamer, Dr. Kiersten Stead, Dr. Justin Kern, Dr. Emily Park



Verdant Robotics' SharpShooter system avoids broadcast spraying by identifying individual plants in real time and applying herbicide only to weeds.

CHAPTER 6

The central challenge for antibody therapies is shifting from discovery to targeting and delivery.

Antibody-based drugs have become one of the most successful therapeutic modalities in modern medicine, helping to treat everything from breast and lung cancer to rheumatoid arthritis. But despite their clinical success, antibodies have still faced a fundamental limitation: they can only reach a fraction of the targets inside the human body that physicians would like to hit. Today, we're seeing researchers explore a much wider array of strategies for expanding where antibodies can go and what they can do.

Antibody drugs bring in annual revenues of hundreds of billions of dollars in a large swath of therapeutic areas, including cancer and autoimmune and inflammatory diseases, but their main limitation is that they don't readily enter cells or penetrate certain tissues. Many companies are working on variations that could unlock biological targets antibodies can't reach. Solu Therapeutics, where DCVC Bio invested in 2023, has a powerful idea for overcoming these longstanding constraints: creating "bifunctional" small molecules called Cytotoxicity Targeting Chimeras (CyTACs) and Therapeutic Index Control Targeting Chimeras (TicTACs) that bind to cell surface target receptors on one end and to engineered antibodies on the other. That gives the antibodies a toehold against cellular targets that were previously off limits—enabling them to kill tumor cells or eliminate pathogenic immune cells. In effect, Solu's technique gives these CyTACs and TicTACs the pharmacology of an antibody.

"What I find interesting about Solu is that they've taken two modalities that people are familiar with, small molecules and antibody-drug conjugates, but they've combined them in a very clever way to turn it into a new technology—a better version of both," says Dr. Emily Park, a Principal at DCVC Bio. Already, Solu is in Phase I clinical trials testing a drug called STX-0712 in patients with relapsed or chemotherapy-resistant chronic myelomonocytic leukemia (CMML), a rare but deadly type of blood cancer that leads to an overproduction of white blood cells called monocytes. Malignant monocytes express high levels of a G-protein-coupled receptor called CCR2 on their surfaces, and Solu developed CyTACs that potently bind to CCR2, allowing their partner antibodies to destroy the cells. (In primate studies, the therapy eliminated 95 percent of CCR2-positive monocytes in the blood.)

Solu's TicTACs have a particularly useful property: they have very long half-lives in the body, on the order of weeks. That's why the company is also investigating their usefulness as long-acting treatments for pain. Here's how it would work: Nociceptors—the specialized sensory neurons that carry pain signals—have sodium ion channels such as NaV1.7 and NaV1.8 embedded in their membranes. When these neurons detect injury, the sodium channels open up, and sodium ions rush into the cell, generating a pain signal that travels along the nerve to the spinal cord and brain. By designing a TicTAC that binds to a sodium ion channel on one end and an antibody on the other, Solu could selectively block these channels and reduce chronic pain.

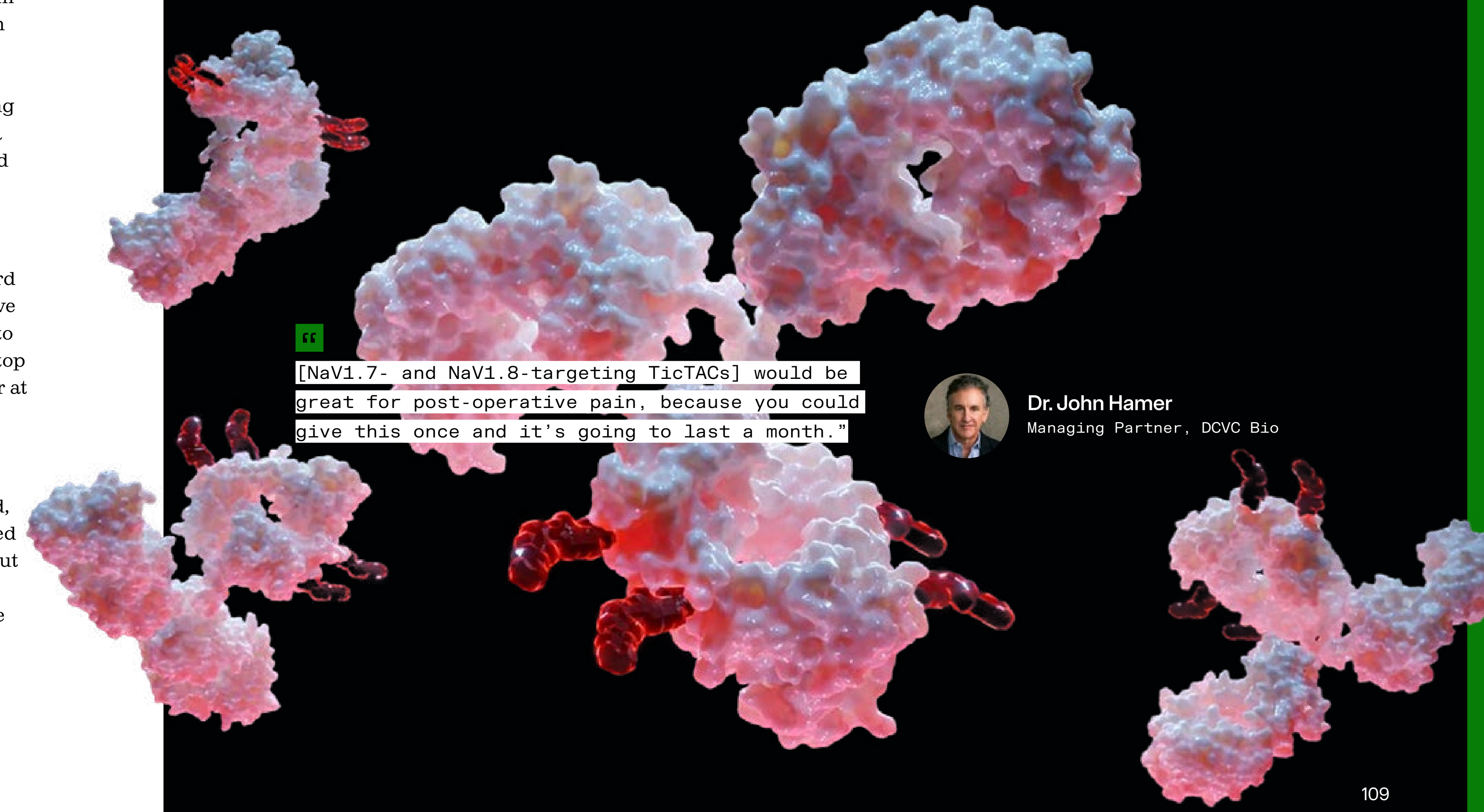
Solu is in preclinical studies with one TicTAC that targets NaV1.7, another that targets NaV1.8, and a third that inhibits both. "It would be great for post-operative pain, because you could give this once and it's going to last a month, and you don't have to give pain med on top of pain med," says Dr. John Hamer, Managing Partner at DCVC Bio.

The old bottleneck in antibody drug development was finding effective antibody molecules in the first place, but the pace of discovery has vastly accelerated, thanks in part AI-driven companies like DCVC-backed Alchemab (see Opportunity 6.2). The story at Solu is about how the emphasis is now shifting from discovery to delivery. And another DCVC portfolio company, Fable Therapeutics, is providing a third critical ingredient: Since the mid-2010s, the team of leading researchers and biotech entrepreneurs that founded Fable in 2021 has been creating generative models, graph neural networks, graph transformers, LLMs, diffusion models, Boltzmann generators, and other AI tools for protein design. Today the company deploys two complementary models: one focused on the 3-D structures of antibodies and the antigens they bind to; the other focused on the amino-acid sequences of these proteins, which behave like words in a large language model. Together, the models can rapidly generate novel, developable designs for binders to fit any given target structure—designs that can then be tested in the lab. That last part, developability, is important because it helps teams prioritize the antibodies that look the most manufacturable and stable, lowering downstream execution risk.

"Fable is really on the bleeding edge of antibody design," says Hamer. "Say you want an antibody that's going to go after this particular target. They can design it from scratch so that it hits it in just the right way, substantially better than antibodies made in the conventional way." Using its platform, Fable is building a pipeline of biologics to go after metabolic diseases, starting with obesity and all its complications. The company believes it can use its understanding of protein design to make antibodies and peptides that could improve on GLP-1 agonists like Ozempic and Zepbound, helping patients improve their basal metabolic rate and insulin sensitivity without the muscle loss associated with the GLP-1 drugs.

Stepping back, the real shift here is that computation is turning antibody therapeutics from a search problem to an engineering problem—where design, targeting, and delivery can be optimized using the same AI tools transforming many areas of biology and medicine. As our ability to read and engineer the immune system improves, researchers are learning not just how to design antibodies, but also how to discover protective ones in human populations—and ultimately how to reprogram immune cells themselves. In the next section, we'll see how those tools are being used to study "superhero" individuals and apply their biology to help others overcome or resist disease.

↓ **Solu Therapeutics**
Proprietary CyTAC molecules link engineered antibodies to receptors that are overexpressed on the surfaces of cancer cells.



"[NaV1.7- and NaV1.8-targeting TicTACs] would be great for post-operative pain, because you could give this once and it's going to last a month."

Dr. John Hamer
Managing Partner, DCVC Bio

Studying elite responders can reveal protective antibodies that help the rest of us fight disease.

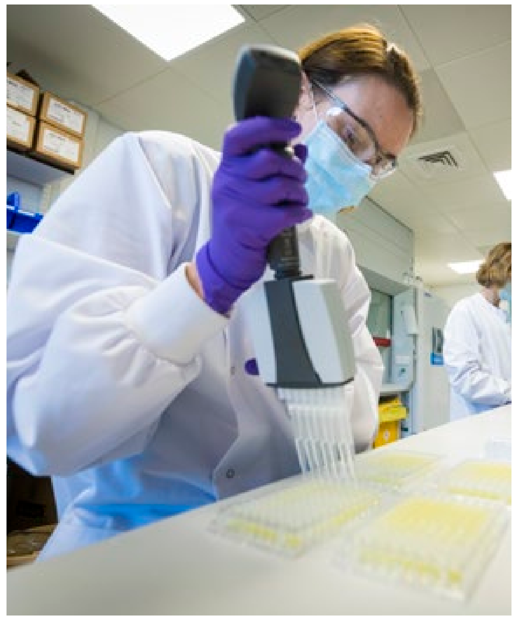
When most people hear the term auto-antibody, they think of autoimmune diseases like lupus or rheumatoid arthritis. But researchers have discovered that some auto-antibodies may actually protect us from disease—neutralizing toxic proteins, dampening inflammation, or eliminating diseased cells. One DCVC Bio–backed company, Alchemab, has become expert at finding these protective antibodies by studying individuals who should be sick but somehow remain healthy.

The human immune system is a gigantic search engine that is constantly running millions of antibody experiments. When antibody-producing B cells form, their DNA is randomly stitched together in segments that create millions of unique templates for recognizing antigens. On top of that, B cells that encounter pathogens rapidly mutate their own genes, in a kind of accelerated Darwinian evolution called affinity maturation, allowing the immune system to iteratively improve the antibodies it produces.

As a result, a single person can generate billions of distinct antibodies, and every individual's repertoire is largely unique. Over time, the immune system continues to evolve as it responds to infections, vaccinations, aging, and the microbiome. In some individuals, this process may produce antibodies that do more than fight infections. For example, a cancer survivor or a slow-progressing Alzheimer's patient may have antibodies that clear malignant cells or neutralize toxic protein tangles.

This diversity may sound impossibly chaotic: every immune system exploring a different region of antibody space. But researchers have discovered that when many people's immune systems confront the same biological challenge—say, a virus like SARS-CoV-2—they often converge on similar antibody solutions. The insight at Cambridge, U.K.-based Alchemab, where DCVC Bio invested in 2019, is that modern sequencing and AI allow scientists to scan the immune repertoires of many individuals and identify these convergences—signals that nature may have discovered a protective solution. In effect, the company is crowdsourcing drug discovery from millions of human immune systems, each conducting its own combinatorial search.

To do this, Alchemab sequences millions of antibodies from individuals who are resistant to, or have recovered from, disease. It uses machine learning to cluster these antibodies into families, looking for those that appear



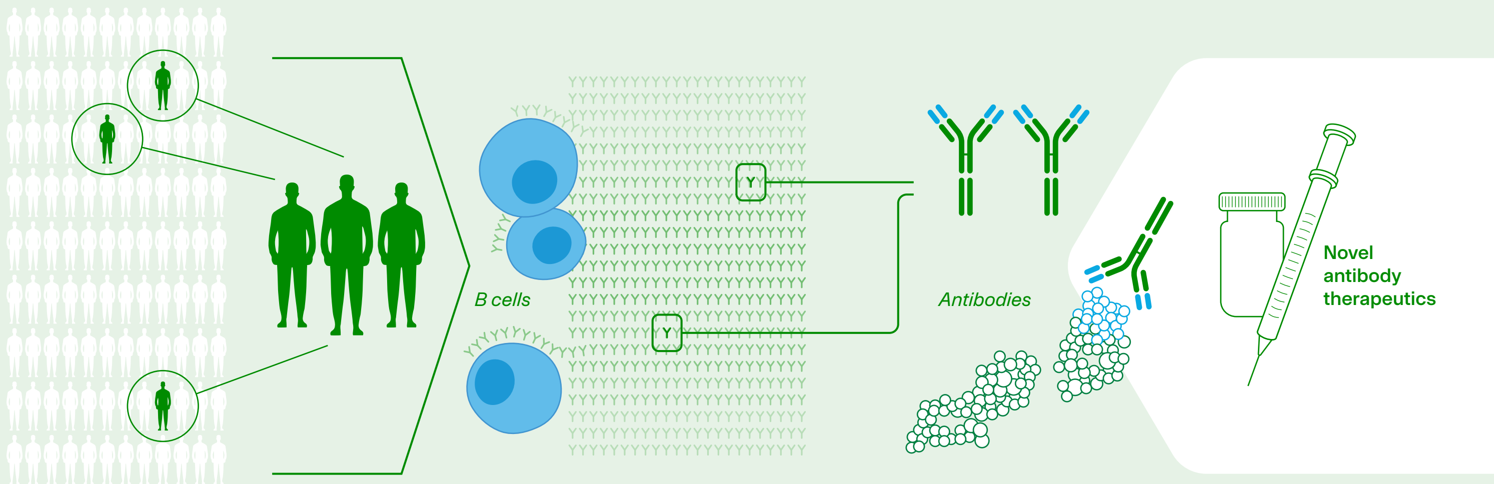
↑ **Alchemab** takes advantage of the human immune system's own combinatorial search capability.

Figure 6.2.1

▸ Alchemab begins by identifying resilient individuals whose naturally occurring antibodies allowed them to resist or recover from disease.

▸ The company sequences B cells from these resilient individuals to see which protective antibodies they have in common.

▸ Alchemab looks for the binding targets of these antibodies and develops therapeutic candidates that replicate the protective effect.



repeatedly across different immune systems confronting the same disease. Then it identifies the binding targets of those antibodies, teases apart their protective effects, and develops therapeutic molecules that replicate those effects.

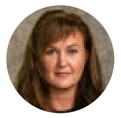
It's almost the inverse of traditional drug-discovery logic. Rather than treating the immune system as an enemy to combat, Alchemab treats it as a vast library of biological solutions that evolution has already tested. "Instead of looking at sick people and asking how you can make them well, you look at the unusually well people and say, 'How is this happening?'" says DCVC Bio General Partner Dr. Kiersten Stead.

At least four drugs developed this way are in the pipeline at Alchemab, including one in Phase I clinical trials for the treatment of amyotrophic lateral sclerosis and other neurodegenerative conditions. The drug, ATLX-1282, has been licensed to Eli Lilly and Company in a deal worth up to $415 million in upfront, milestone, and royalty payments.

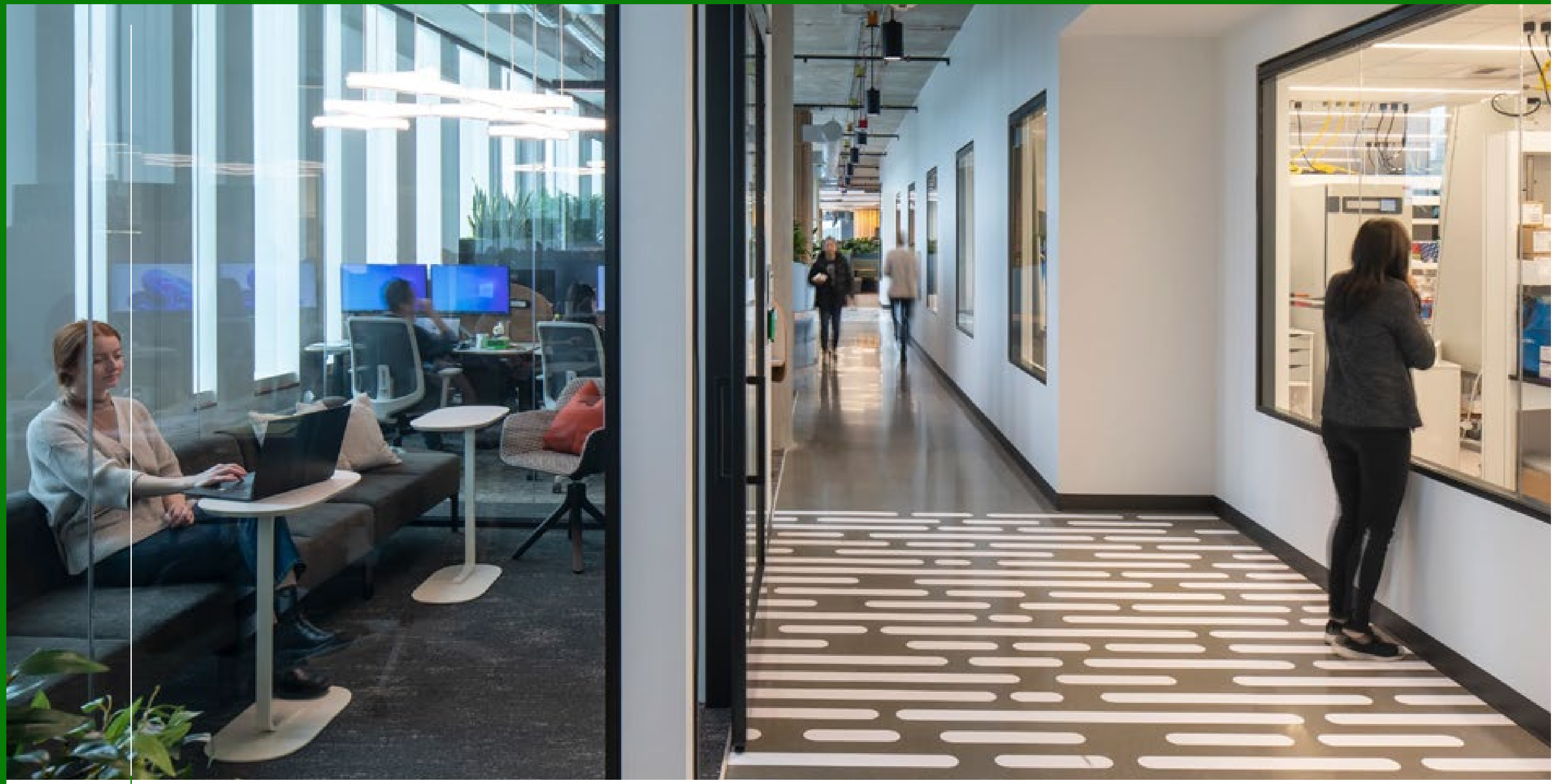
Alchemab created ATLX-1282 by identifying a common antibody sequence in a cohort of individuals carrying mutations that normally lead to frontotemporal dementia but who stayed healthy into old age. Further research showed that the antibody plays a role in neuroprotection across multiple neurodegenerative conditions.

Because Alchemab's approach begins with protective antibodies rather than known disease targets, it can reveal entirely new biological pathways involved in disease. "They've built a huge database of this kind of information, and they've discovered some extremely novel targets," says Hamer. "And not only novel targets, but antibodies that no one would have predicted would bind in the way they're binding." Modern sequencing and AI are, in effect, turning the immune system into a new discovery engine—one companies like Alchemab can harness to find signals in the chaos.

> "Instead of looking at sick people and asking how you can make them well, you look at the unusually well people and say, 'How is this happening?'"



Dr. Kiersten Stead
Managing Partner, DCVC Bio



↑ Umoja Biopharma
The company delivers the genes for cancer-attacking antigens directly to human T cells in the body—meaning doctors no longer have to extract these cells for reprogramming in the lab.

In vivo CAR-T therapies could make one of medicine's most powerful treatments far easier to deploy.

Antibody therapeutics like those being developed by Solu, Fable, and Alchemab harness the immune system indirectly. CAR-T therapy (for chimeric antigen receptor T-cell therapy) takes a more hands-on approach: it reprograms a patient's immune cells to recognize and attack cancer on their own. But while CAR-T therapies have produced deep remissions in patients with otherwise untreatable blood cancers, their cost and logistical complexity has limited how widely they can be used. Overcoming the bottlenecks will require several adjacent innovations at once—which is why DCVC's bet in this space is Umoja Biopharma.

Since the first CAR-T therapies were approved in 2017, more than 40,000 patients worldwide have received them, with remission rates above 80 percent in some studies—a startling result given that nearly all of these patients had relapsed after multiple traditional therapies, and often had expected survival times under one year. "CAR-Ts are amazing—they're the closest to curing cancer that we've ever gotten," says DCVC Bio Principal Dr. Emily Park.

It's frustrating, then, that CAR-T is so daunting to deploy. The process involves extracting a patient's T cells (a process called leukapheresis), genetically engineering them to allow them to recognize specific antigens on cancer cells, growing the modified T cells in bioreactors, and reinfusing the cells into the body. Expanding the reprogrammed T cells outside the body can take three to four weeks; meanwhile, patients must undergo chemotherapy to deplete their existing immune cells. The total price tag can exceed $400,000 per infusion, not even counting the cost of treating the dangerous immune reactions the therapy can provoke.

> "CAR-Ts are amazing—they're the closest to curing cancer that we've ever gotten."

Dr. Emily Park
Principal, DCVC Bio

Many patients never receive CAR-T despite being eligible—because of the high cost, the long expansion time, deterioration during the waiting period, and a lack of specialized treatment centers able to deliver the therapy. The core problem is that today's CAR-T therapies are painfully artisanal and personalized. Every population of genetically modified immune cells is a batch of one, requiring a dedicated manufacturing slot, complex cold-chain shipping, and the like.

Umoja has come up with ways to skip nearly every step in the traditional CAR-T process. They've developed a lentiviral vector that delivers the CAR gene directly to T cells in vivo, so that the body itself performs the manufacturing step. No leukapheresis, no cell expansion, no personalized manufacturing queue, and little or no need for lymphodepletion. When CAR-T looks more like a conventional biologic therapy—with a standardized product, stored and shipped like an injectable drug, and administered at a hospital—the treatment begins to look dramatically more scalable.

Umoja doesn't have this industry to itself—in fact, the field is crowded with in vivo CAR-T startups, including some that have already exited (in 2025 AstraZeneca acquired EsoBiotec for $1 billion and Kite acquired Interius BioTherapeutics for $350 million; this year Lilly bought Orna Therapeutics for $2.4 billion). But Umoja's early start mattered.

For one thing, the company has had time to develop its lentiviral platform, called VivoVec. Modeled after the human immunodeficiency virus, the particles can carry any CAR payload, and are engineered not just to deliver transgenes but to prime their target T cells to attack cancer cells. The genetic payload can also include a system called RACR, for rapamycin-activated cytokine receptor, intended to boost T-cell expansion inside the body and reduce or eliminate the need for lymphodepletion. Many competing in-body CAR-T companies use lipid nanoparticles loaded with messenger RNA as their vehicles for reengineering T cells, but Umoja argues that lentiviruses have greater single-dose persistence and more closely mimic traditional CAR-T expansion. "A lot of these lipid nanoparticles have to be dosed again and again and again," Park says.

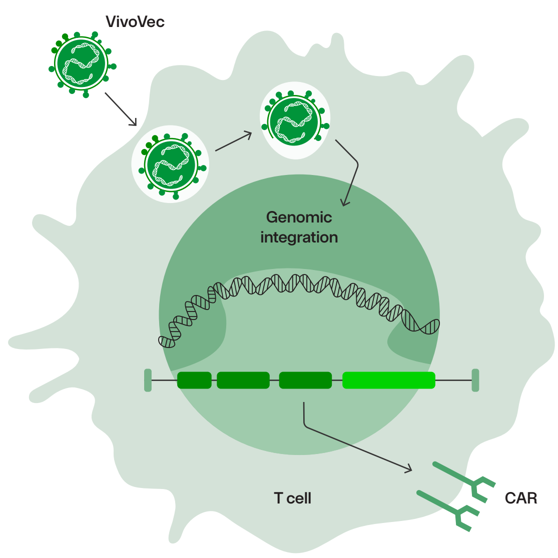
In addition, Umoja is the only in vivo CAR-T company to invest in its own lentiviral manufacturing facility—an unusual move in a field where most early-stage companies rely on contract manufacturers. And crucially, it's the first to receive clearance from the FDA to begin Phase I clinical trials. In late 2024 the company started in-human trials of a lentiviral vector called UB-VV111 in combination with rapamycin for the treatment of relapsed or chemotherapy-resistant large B-cell lymphoma and chronic lymphocytic leukemia. And in 2025 it began trials of a separate construct, UB-VV400 plus rapamycin, for the treatment of aggressive lymphoma. AbbVie has agreed to pay up to $1.44 billion for the rights to UB-VV111 and several related CAR-T programs.

What creates Umoja's technical moat is not just its first-mover status: it's a combination of the company's elegant lentiviral vector, a built-in expansion control system via RACR, and in-house manufacturing. Together, these attributes make the company far more than another me-too oncology startup—it's a true deep tech systems play. "In vivo CAR-T has been the hot topic this year, and some people feel like they need to hop on the boat," Park says. "But Umoja is in the clinic now, so it's perfect timing."

What ties all these approaches together is the growing realization that the immune system is not just a defense mechanism—it can also be an engineering platform. With the right tools, its molecules, cells, and evolutionary search processes can be made to produce new medicines for cancer and other devastating diseases.

Figure 6.3.1

In vivo CAR-T cell therapy: Umoja's VivoVec platform, a modified lentivirus, carries chimeric antigen receptor (CAR) RNA into the patient's T cells, which manufacture and express the receptors on their own.



Autonomy and AI are making farms programmable—and helping farmers optimize labor, equipment, and chemical use.

When autonomy is paired with machine learning technology that can measure and adjust field operations in real time, it can drastically reduce farmers' spending on their biggest inputs.

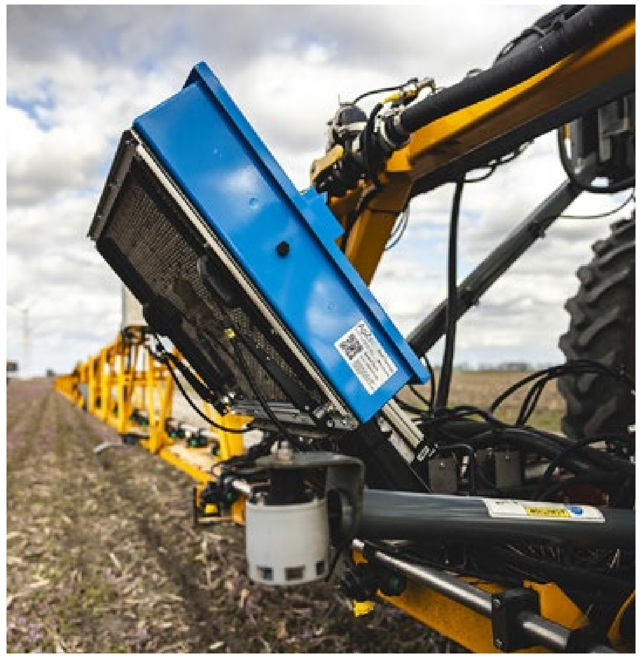
Sabanto's retrofit kits make standard John Deere and Kubota tractors fully autonomous.

Three pressures are forcing farmers to consider adopting more autonomous technology: labor scarcity, fertilizer and chemistry costs, and the capital inefficiency of farm equipment, which sits idle for much of the year. Autonomy and machine learning offer a way to ease all of these constraints simultaneously.

One company tackling the capital and labor problems is Sabanto, where DCVC Bio first invested in 2019. The company builds autonomous retrofits for conventional farm equipment, so farmers don't have to buy entirely new machines to access the latest in autonomy. The system combines GPS guidance, sensors, onboard computing, and safety systems that allow a tractor to operate without a driver. The autonomy kit can be installed on widely used tractors from all major manufacturers, allowing them to plant, spray, or till fields while the farmer supervises remotely.

The appeal is economic. Farmers often feel pressure to purchase ever larger machines to manage expanding acreage during narrow planting and harvesting windows. Sabanto's retrofit approach lets farmers automate tractors they already own—or purchase smaller, less expensive machines and operate several at once. Instead of a $500,000 flagship tractor running a few weeks per year, multiple autonomous tractors can work in parallel for longer hours, improving equipment utilization and reducing labor requirements. "You can't always get people to sit on tractors all day in these remote areas," comments DCVC General Partner Alan Cohen. "If all you're doing is going up one row and coming back down the next row, that's solving a need."

One Montana farmer, Justin Yirsa, faced exactly this decision. Yirsa needed to replace aging equipment to run his 10,000-acre wheat farm. One option was to spend about $1.5 million on a new Case IH Steiger tractor and a Bourgault seeder. Instead Yirsa used two Sabanto kits to retrofit an idle 2013 John Deere 5115M utility tractor as well as a 2008 John Deere 1590 No-Till Drill, used for seed and fertilizer delivery. His total cost, including the Sabanto systems, was $490,000. In trials, the retrofitted tractors seeded 8 acres per hour and burned half as much fuel as the Steiger and Bourgault would have.



← **AgZen** The company's RealCoverage system helps farmers measure and control how much pesticide spray reaches the leaves of their crops.

With the savings Yirsa purchased more land—an appreciating asset, unlike a massive depreciating tractor. In effect, autonomy allowed Yirsa to rethink the economics of his farm rather than scaling up equipment.

One of the main uses of large farm vehicles is pesticide spraying, and this is another area where computation can bring efficiency gains. American farmers spend at least $14 billion per year on herbicides, insecticides, fungicides, and fumigants, according to the EPA. Yet only 1 percent of it actually reaches pests. A lot of this chemistry simply misses the leaves of crop plants due to spray drift or evaporation. And because many plant leaves are hydrophobic—repelling water and other liquids—much of the pesticide that does land on leaves simply bounces off, rolls off, or evaporates, finding its way into the environment. Farmers historically have had almost no way to measure or respond to any of these effects in real time. They fall back on guesswork, and they typically overcompensate for uncertainty by vastly overapplying chemicals.

AgZen, an MIT spinoff company we funded in 2025, is solving the measurement problem. The company builds a camera system called RealCoverage that farmers can bolt on to their existing foliar sprayers to examine leaves before and after spraying to see how much chemical actually sticks to plants. This data, together with other variables such as weather, feeds into an AI model that can suggest retention-improving adjustments to sprayer speed, droplet size, boom height, or surfactant use.

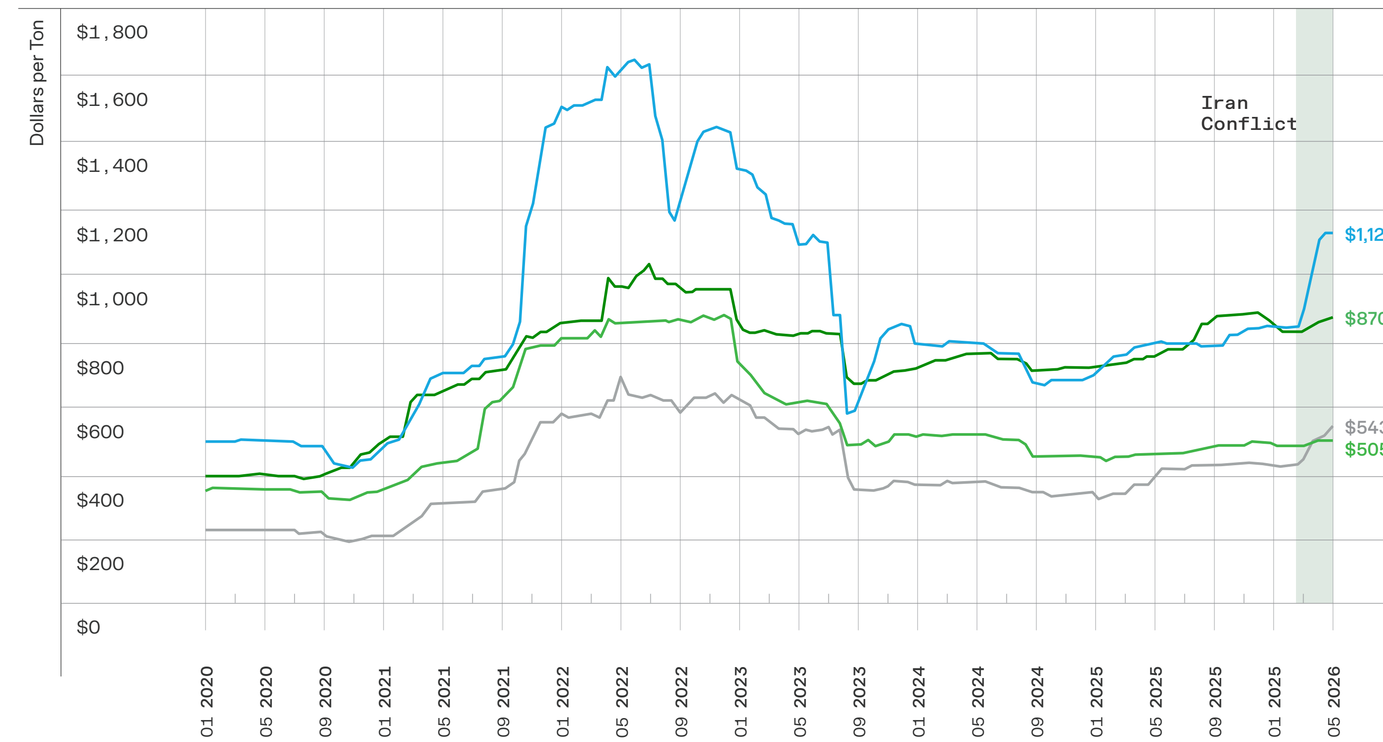
"We can literally cut the spray rate in half and get the same level of pest control just by optimizing those parameters, which farmers would normally just set heuristically or based on label recommendations about the highest rate permissible by regulation," says DCVC Bio Partner Dr. Justin Kern. After one trial season using RealCoverage, the Ernest & Julio Gallo Winery reported a 30 percent savings on chemicals. For large farms—the ones that spend hundreds of thousands of dollars per season on crop protection—this kind of efficiency gain can add up fast.

AgZen is making traditional spraying systems far more efficient. DCVC Bio–backed Verdant Robotics is taking a more radical approach: replacing broadcast spraying altogether. Cameras on the company's SharpShooter sprayer distinguish between crops and weeds in real time. Then aimable nozzles pivot to deliver chemistry exactly where it's needed—herbicides can be targeted to hit weeds as small as 2 millimeters in diameter. The company says that its technology can slash chemical use by up to 99 percent, while also saving on labor costs, since the system can eliminate hand weeding. "What used to take 30 people all day can now be done with one machine and one operator," one California grower of romaine and broccolini told *Western Grower & Shipper* magazine.

Figure 6.4.1

Nitrogen fertilizer price volatility in Illinois: Prices increased early in the Ukraine conflict in 2022, gradually eased, then shot up again sharply with the onset of hostilities in Iran and the closing of the Strait of Hormuz.

Source: University of Illinois Farmdoc Daily.



← **Verdant Robotics** In a study of spinach crops, Verdant's Aim & Apply technology produced 67 percent higher yields by keeping herbicide away from crop plants.

As we noted above (*see page 87*), DCVC-backed Pivot Bio is also helping farmers save on fertilizer, by manufacturing a microbial crop additive that helps plants fix their own nitrogen from the air. By reducing chemical inputs, making better use of existing equipment, and replacing manual labor, technologies like those at Pivot, Sabanto, AgZen, and Verdant are helping farms stay profitable in the midst of fuel- and chemical-cost volatility and growing labor constraints.

In effect, autonomy allows farm equipment to behave more like software—bringing a new level of flexibility to farming operations. "It's not a one-to-one replacement of human beings with autonomous systems," Kern says. "It's a change in how people deploy their time on a farm. Autonomy lets you do things that you couldn't do with labor."

CHAPTER 7

Venture-backed companies are building a computational defense and space infrastructure to meet 21st-century needs and protect against 21st-century threats.

The aerospace and defense industries have long been dominated by a small set of prime contractors building exquisite hardware systems—fighter jets, submarines, aircraft carriers, missile systems—that are very expensive and take decades to design and deploy. Those systems remain essential. But the strategic environment in which they operate is changing.

Today's security threats are smaller, faster-moving, and often software-driven. A $1,000 drone can threaten a military base or a multibillion-dollar Navy ship. Online influence campaigns can shape political outcomes across continents. Meanwhile the physical infrastructure that supports the global digital economy—from power grids to data centers to satellites—faces growing risks from both adversaries and nature itself.

These shifts have opened the door to a new generation of defense and aerospace companies that can outperform the traditional defense industrial base by focusing on software, sensing, autonomy, and data integration. Rather than replacing the primes, these companies aim to complement them—filling critical gaps in areas such as counter-drone defense, space logistics and communications, and the emerging battlefield of the information environment.

"When you're facing autocrats and terrorists who can manufacture weapons cheaply, you have to be able to defend against them," says DCVC Co-Founder and Managing Partner Matt Ocko. "We need ways of turning the cost aspects of war planning and manufacturing back against our enemies."

In an industry where change is traditionally slow and expensive—far too slow to keep up with today's emerging threats—the space, defense, and cybersecurity companies we back are bringing a dose of deep tech computational leverage and efficiency to this problem.

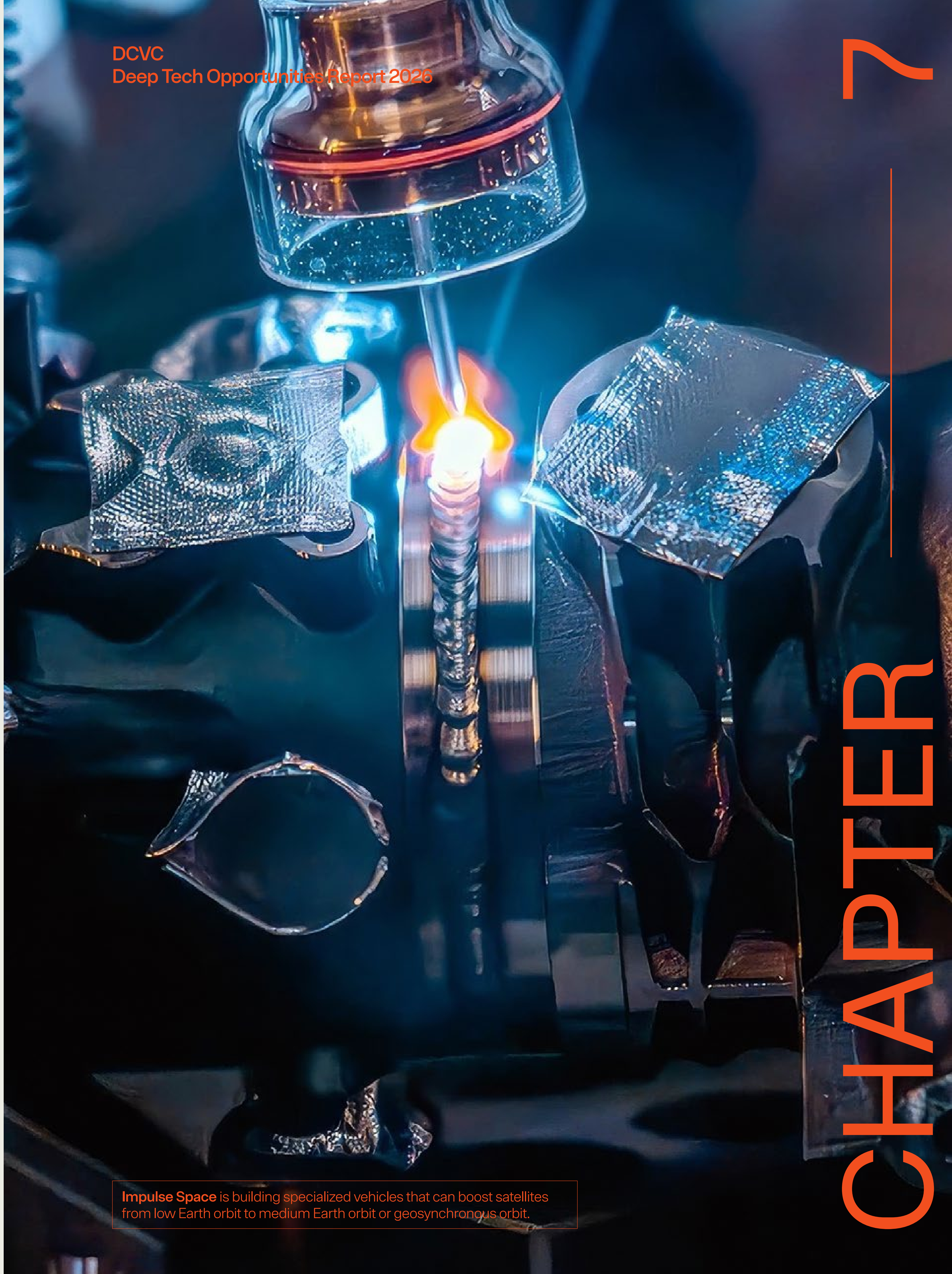
OPPORTUNITIES

COMPANIES

Fortem Technologies, Impulse Space, Reality Defender, Rocket Lab

VOICES

Matt Ocko, Matt O'Connell



Impulse Space is building specialized vehicles that can boost satellites from low Earth orbit to medium Earth orbit or geosynchronous orbit.

Cheap drones are changing the character of conflict—and computation-heavy startups are responding to the challenge.

If there's a field where Silicon Valley's old motto "move fast and break things" makes literal sense, it's counter-drone systems. Defending against drones is fundamentally a computational problem: detecting small aircraft in cluttered airspace, classifying them correctly, and coordinating interceptors in time to stop them. Engineering that kind of real-time sensing and decision-making is exactly where venture-backed, software-driven aerospace companies can move faster than traditional defense contractors.

Fortem Technologies' DroneHunter can act alone or in groups to protect restricted areas from rogue drones.

In military planning, the cost-per-kill—the cost of destroying an incoming threat—should ideally be lower than or comparable to the cost of the weapon being stopped. During the Cold War, this balance held: expensive missiles were used to intercept other expensive missiles or aircraft. But modern drone warfare has broken the equation. The typical defense against a $500 commercial quadcopter is a surface-to-air missile costing $100,000 or more. One effective defense against a $20,000 Iranian Shahed-136 drone is a $4 million Patriot missile. Cheap, mass-produced, expendable, hard-to-detect drones that fly low and slow allow bad actors—from major powers to insurgent groups—to impose enormous costs on more technologically advanced militaries.

The goal of companies building new counter-drone systems is to flip the cost equation back in the defender's favor: instead of firing expensive missiles, the defender launches another relatively cheap interceptor drone. Sometimes these drones are designed to be expendable, quickly executing kinetic or munitions attacks, as in the case of Anduril's Anvil drone series. In the case of Fortem Technologies, whose DroneHunter can capture and bring down enemy drones, they're built to be reusable.

We first invested in Fortem a decade ago, and now that small drones are being recognized as urgent battlefield threats, they've grown into one of the leaders of a large new defense market. The company's SkyDome suite integrates 3D radar detection and tracking, a pan-tilt-zoom camera system, AI-based threat classification software, and the DroneHunter itself, a hexacopter that fires nets to capture both fixed-wing drones and other rotary-wing drones. The DroneHunter's capture rate for fixed-wing drones is more than 85 percent, and its capture rate for rotary targets is above 95 percent. (The company also makes an explosive kamikaze drone for situations where the user wants the enemy drone destroyed.)

“

America has been behind the eight ball with respect to autonomy in warfare, but now we are dramatically increasing spending on drones. We salute the administration for reacting so promptly.”

Matt O’Connell
Operating Partner, DCVC

↑ DroneHunter F700 The autonomous, radar-guided drone snares enemy drones with its onboard NetGuns.

The wars in Ukraine and Iran, attacks on shipping in the Red Sea, drone overflights of military bases and airports, and drone proliferation globally have awakened Western militaries and elected officials to the need for such technology, resulting in major contracts and deployments for counter-drone companies. “America has been behind the eight ball with respect to autonomy in warfare, but now we are dramatically increasing spending on drones,” says Matt O’Connell, an Operating Partner at DCVC and chairman of Fortem. “We salute the administration for reacting so promptly.”

After Fortem dramatized SkyDome’s ability to stop swarms of incoming drones in a “5-vs-5” intercept demonstration—neutralizing five hostile drones autonomously, with no collateral damage—the Utah-based company won a three-year, $18 million contract from the U.S. Army to deploy DroneHunters and field support services to protect Army installations and operations. That came on top of a dozen orders last fall from U.S. allies in Europe and the Middle East, as well as deployments in Ukraine and along the U.S. southern border.

Last November, Fortem became the first company admitted to the U.S. Army’s new Global Tactical Edge Acquisition Directorate (G-TEAD) Marketplace, which is designed to get advanced technology to operational units faster. “G-TEAD gives commanders a way to procure proven counter-UAS capabilities like DroneHunter and SkyDome in weeks rather than years—which means they can field what they need at the pace the threat is evolving,” Fortem CEO John Gruen says. “This is transformation for our customers and for the mission of saving lives.”

Counter-drone technology is increasingly important outside military contexts as well. The 2026 FIFA World Cup—with 48 teams competing across the U.S., Mexico, and Canada—is expected to attract six million spectators. In February Fortem revealed that the Department of Homeland Security had placed a multimillion-dollar order for interceptors to shield the 11 U.S. soccer venues. Fortem’s net-capture technology is ideal for this kind of civilian infrastructure protection, since it aims to “remove hostile drones without showering debris on bystanders below,” as the news organization Breaking Defense puts it.

“Long term, the civilian opportunity for Fortem is bigger than the military one,” O’Connell says. “Every baseball field, every football field, every soccer stadium, every utility, every nuclear plant are areas where Fortem can help.” In January the company delivered DroneHunter 5.0, which carries enhanced onboard cameras and computing power for defense against coordinated swarms. The 5.0 system and the 5-vs-5 demo show DroneHunter’s capabilities not just as an interceptor, but as part of an overall airspace defense system.

The counter-drone category spans a huge spectrum of threats, from $400 first-person-view quadcopters to the delta-winged Shahed-136, which is estimated to cost $20,000 to $50,000. A system called Merops from Perennial Autonomy has reportedly logged over 1,000 Shahed interceptions in Ukraine, and the Iranian-built drones are also thought to be vulnerable to ramming interceptors like Anduril’s Anvil. But no single defense system can handle all types of unmanned aerial systems. Fortem’s sweet spot is low-collateral-damage urban airspace defense—and its reusability means customers can begin to reverse the economics of asymmetric conflict.

“The drones that we’re all investing in are much smaller now. The downside of that is that there’s less of a barrier to entry, and any moron in a garage can build one,” says O’Connell. “Bad people are doing more bad things. And that means people who have been focused on drones as attack vehicles need to invest just as much in drones as defense.” The much-discussed “democratization of air power” may have handed sophisticated aerial surveillance and attack capabilities to insurgent militias, terrorist organizations, criminal cartels, and pariah states—but fortunately, Fortem is democratizing them right back.

Wars are won on logistics, and new computational systems will determine how quickly the next generation of defense technologies can be deployed.

In modern warfare, technological breakthroughs don't matter unless planners can swiftly integrate them into operational systems. A new generation of defense and space companies is focusing on the infrastructure that determines how fast new capabilities can get to where they're needed—whether that means coordinating fleets of autonomous systems, moving spacecraft between orbits, or integrating sensors, data, and weapons into operational networks.

"Defense amateurs talk about 'tip of the spear' weapons," says DCVC Co-Founder and Managing Partner Matt Ocko. "Defense professionals quietly plan logistics."

There's a reason, in other words, that people remember the dictum attributed to Napoleon that "an army marches on its stomach." The ability to move people, equipment, and supplies faster and more reliably than an adversary has always been a decisive advantage in war.

One new challenge, however, is that the planet's surface, its oceans, and even space are becoming increasingly transparent to surveillance systems. That visibility makes traditional logistics—moving ships, aircraft, and cargo through contested environments—more vulnerable than ever. The bottlenecks are therefore shifting from innovation and production to speed, planning, and rapid deployment. And that shift is creating opportunities for companies using more advanced software, AI, and autonomous systems to move materiel and capabilities where they're needed far more quickly than traditional approaches allow.

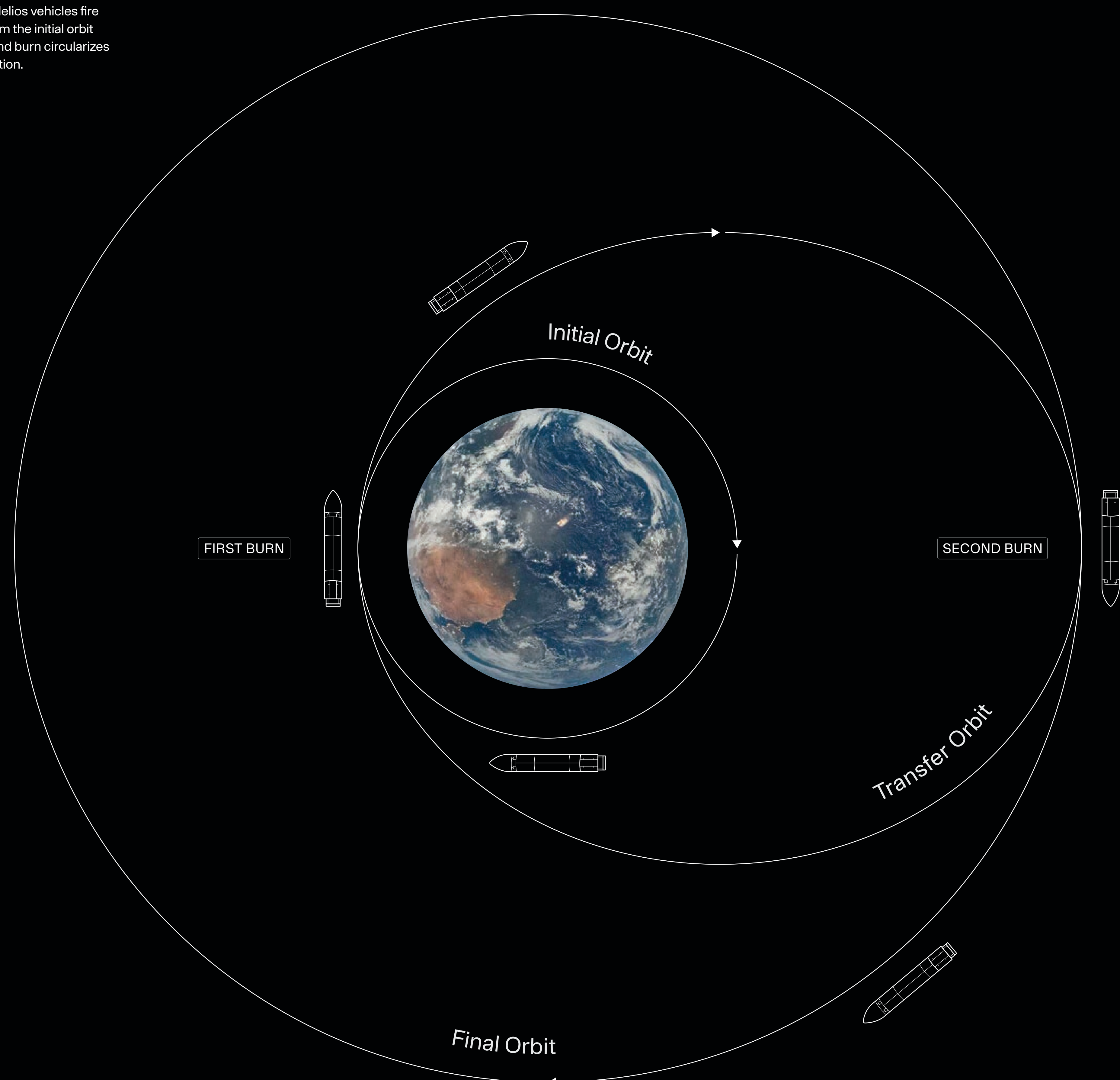
One of the clearest examples is orbital logistics. Launch providers such as SpaceX have dramatically lowered the cost of getting payloads into low Earth orbit (LEO). But placing satellites precisely where they need to operate—and moving them when circumstances change—remains a complex and expensive problem.

Impulse Space, part of our portfolio since 2024, is focused on solving that challenge. The company builds orbital transfer vehicles designed to move payloads from LEO to medium Earth orbit (MEO) or geosynchronous orbit (GEO) quickly and efficiently. Its Mira spacecraft, which first flew in 2023, can maneuver payloads of up to 300 kilograms and is built for rendezvous and proximity operations (RPO). Meanwhile its Helios booster is designed to ride a larger rocket such as a SpaceX Falcon 9 to LEO and then rapidly lift more than five tons of payload to MEO,

↑ **Impulse Space** engineers assemble the company's Helios kick stage vehicle. Paired with a medium-lift launch vehicle such as a SpaceX Falcon 9 or Rocket Lab Neutron, Helios will deliver large payloads to MEO, GEO, or even the Moon.

Figure 7.2.1

Impulse Space's Mira and Helios vehicles fire once to boost a payload from the initial orbit into a transfer orbit. A second burn circularizes the orbit at the final destination.



GEO, or even higher. Helios is set for its first flight this year, and the company plans to use it to send landers and cargo to the Moon in 2028.

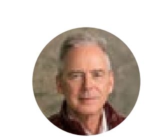
In December, Impulse completed a fully automated rendezvous between two Miras in LEO, guided by software from Starfish Space that brought one craft alongside the other solely using optical navigation from onboard cameras (in contrast with systems that historically relied on multiple complex sensors for RPO). And in a mission planned for this year, Impulse will use Mira to boost a fuel depot built by Orbit Fab to GEO, where it will refuel a Space Force test satellite called Tetra-5.

In an era when satellite networks underpin everything from communications to navigation and surveillance, the ability to refuel, reposition, or replace spacecraft could become an increasingly important form of space logistics. "There's significant government interest in what they're building," DCVC Operating Partner Matt O'Connell noted in last year's Deep Tech Opportunities Report. "If a government satellite were to stop functioning, being able to get a replacement satellite up quickly and inexpensively has a lot of appeal."

But moving hardware is only half of the logistics challenge. The other half is planning: determining where assets should go, when they should arrive, and how they should be redeployed as battlefield conditions change. That kind of dynamic planning has historically been slow and largely manual—but companies are beginning to apply AI to the problem.

"

> If a government satellite were to stop functioning, being able to get a replacement satellite up quickly and inexpensively has a lot of appeal."



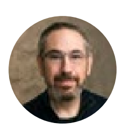
Matt O'Connell
Operating Partner, DCVC

→ **Smack Technologies** aims to strengthen battlefield decision-making by fusing, analyzing, and reasoning over vast streams of multimodal data.

DCVC's Ocko says logistics planning inside the Pentagon often still relies on manually updated spreadsheets and whiteboards—and it can take hours or days to recompute when battlefield conditions change. At best, general staff use separate tools operating on different time horizons—months, days, or the literal present moment—but those tools are siloed and rarely communicate with each other. In effect, planners have to figure it out as they go. "As we look at modern peer-level conflict, that's just not going to work," says Andy Markoff, a former Marine special operations officer who worked on operations and strategy at Palantir before founding Smack Technologies.

Smack is building AI models for defense planners that are designed from the outset to bridge the different time horizons of military planning and help commanders make better decisions. One of the company's core planning platforms, Omega, operates across three distinct time horizons. At the top layer, it correlates orders and directives with battlespace data to form a concept of operations or "con op." One layer down, it establishes a repeating execution cycle for a given campaign—usually on the order of hours or days—and assembles battlespace data into a synchronized matrix of maps and timelines that help commanders see what's happening, or what's about to happen, and refine the larger con op as needed. At the most granular layer, execution mode, the system shows the con op playing out and assesses how the adversary is reacting, whether the campaign is on-plan or off-plan, and how deviations from the plan would impact logistics, supply chains, and resource management.

> "
>
> Defense amateurs talk about 'tip of the spear' weapons. Defense professionals quietly plan logistics."



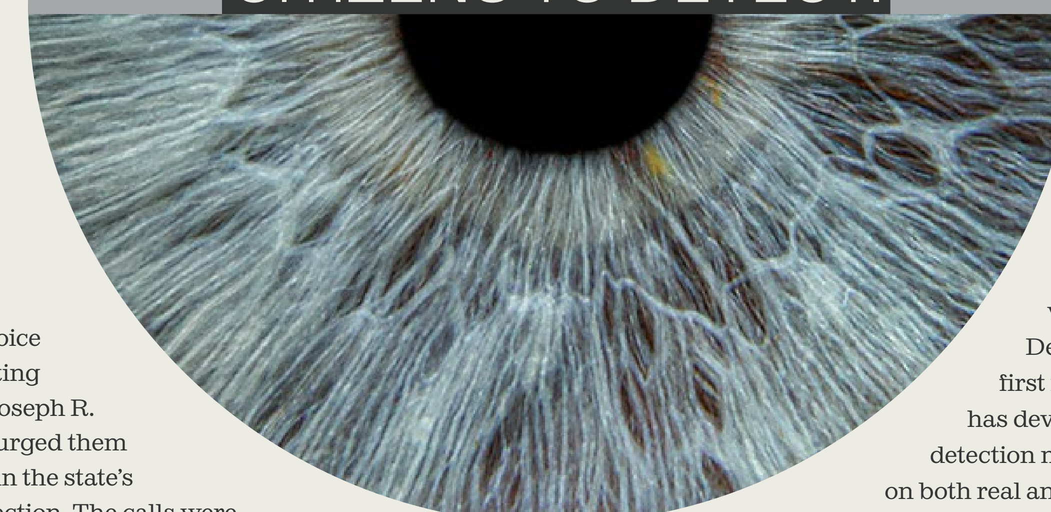
Matt Ocko
Co-Founder and Managing Partner, DCVC

This March Smack announced that it has raised $32 million in seed and Series A funding and secured contracts with multiple branches of the U.S. armed forces, including the Joint Fires Network and the Marine Corps Warfighting Lab. Markoff says Omega is the first in a series of "reasoning models rooted in physics that can make complex time-space calculations with precious resources in seconds." As warfare becomes faster and more data-driven, the ability to plan and replan operations at that time scale may prove decisive. Companies applying AI to logistics and operational planning are attempting to give U.S. and allied commanders exactly that advantage.

INFOSPACE DEFENSE

REALITY DEFENDER

THE RISE OF GENERATIVE AI IS MAKING DISINFORMATION CAMPAIGNS EASIER TO CONDUCT AND HARDER FOR CITIZENS TO DETECT.

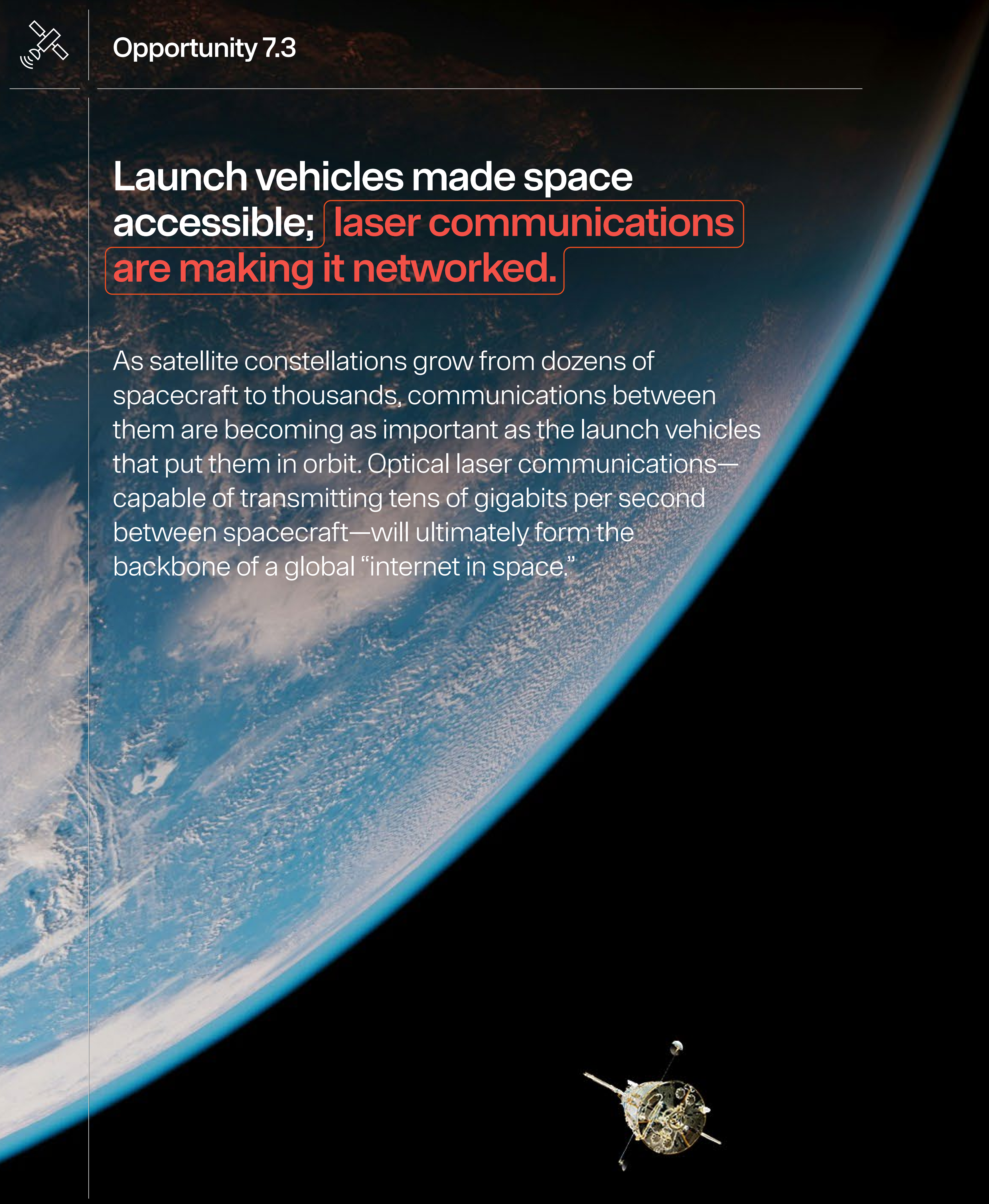


In January 2024, voters in New Hampshire received robocalls featuring a synthetic voice impersonating President Joseph R. Biden that urged them not to vote in the state's primary election. The calls were generated using commercial voice-cloning tools; the FCC proposed a $6 million penalty for the spoofer. AI-generated audio and video has also been used to impersonate Secretary of State Marco Rubio, U.K. Prime Minister Keir Starmer (then Labour party leader), Ukrainian president Volodymyr Zelenskyy, and other leaders. Meanwhile, pro-Russia actors have used software such as OpenAI's Sora 2 text-to-video generator to make dozens of viral clips purportedly illustrating desperation or low morale among Ukrainian troops. Investigators at U.S. Cyber Command have described sophisticated cross-platform influence campaigns using cloned websites, synthetic media, and bot amplification to push pro-Kremlin narratives and erode support for Ukraine in the West.

When citizens are unable to tell the difference between legitimate and manipulated information, it can "decay their mental resilience," lower their trust in media, and foster "us vs. them" polarization that ultimately leads to the marginalization and exclusion of whole populations, NATO's Allied Command Transformation has warned.

That puts a premium on tools that can detect and flag synthetic or manipulated audio and video. Reality Defender, where we first invested in 2023, has developed deepfake detection models—trained on both real and manipulated media—that can screen both recorded and live-streamed media for artifacts and inconsistencies introduced by generative models. One case study unfolded in Taiwan, where actors believed to be affiliated with the Chinese Communist Party used AI-generated content to influence the January 2024 presidential election. Audio clips surfaced during the fall 2023 campaign in which presidential nominee Ko Wen-Je purportedly defamed his opponent, Lai Ching-te. Taiwan's Ministry of Justice Investigations Bureau used Reality Defender to show that the recording was fake and warn the public.

To DCVC Managing Partner Matt Ocko, companies like Reality Defender are the venture industry's best play against foes who aim to exploit the viral power of social media platforms. "Social media is intrinsically cognitively destructive," Ocko says. "It doesn't matter what magic military hardware and software you have if your society can be inexpensively turned into the last days of Rome by enemies from within and without. So the idea of cognitive defense, of infospace defense, is a very interesting one."



Launch vehicles made space accessible; laser communications are making it networked.

As satellite constellations grow from dozens of spacecraft to thousands, communications between them are becoming as important as the launch vehicles that put them in orbit. Optical laser communications—capable of transmitting tens of gigabits per second between spacecraft—will ultimately form the backbone of a global "internet in space."

↑ **Rocket Lab's** Electron is one of the commercial vehicles helping to pepper the sky with satellites, creating a need for better inter-satellite networking.

The launch business to low Earth orbit (LEO) is booming thanks to large public and private investments in reusable rockets. (In this category, we're a proud backer of Rocket Lab [Nasdaq: RKLB], whose Electron rocket has carried 260 satellites to orbit.) Elon Musk's Falcon 9 rocket at SpaceX has revolutionized the launch business—bringing the internal cost of reaching LEO below $2,000 per kilogram and enabling the company to reach profitability through its Starlink internet division. Remote-sensing constellations from Planet and Capella Space (both backed by DCVC) as well as ICEYE, Umbra, and others are generating high-resolution, 24/7 coverage of Earth's surface that is invaluable for intelligence gathering, climate monitoring, and disaster response.

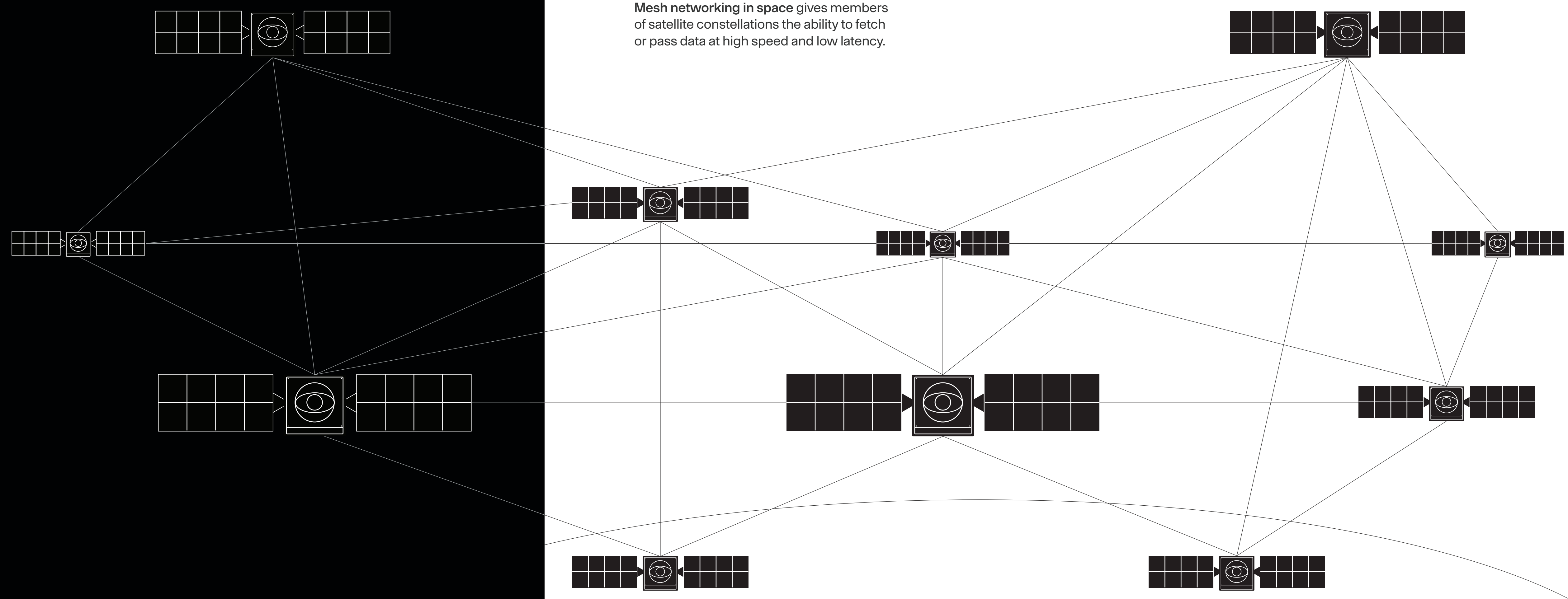
But the communications infrastructure linking all these satellites to the ground and to one another is still relatively primitive. As constellations grow, inter-satellite networking is the new bottleneck. Optical laser communications are one emerging solution—and we think it's a field where venture-backed companies will have an impact. "I think launch is overinvested at this point. Earth observation is probably overinvested too," says Matt O'Connell, an Operating Partner at DCVC. "But comms is worth looking at. I am interested in how we increase the efficiency and security of communications."

Optical inter-satellite links, or OISLs, are far from speculative—SpaceX has already demonstrated the power of the approach. Its latest Starlink satellites use OISL to route data from spacecraft to spacecraft before returning to Earth, creating a space-based mesh network capable of sharing traffic around the planet. Each satellite carries three lasers capable of transmitting tens of gigabits per second. Company engineers said in early 2024 that their network of more than 9,000 installed laser links carried 42 petabytes of data per day; the number is likely higher today. (The company also sells mini-lasers to other satellite manufacturers, providing up to 25 Gbps at distances up to 2,500 miles.) Amazon, meanwhile,

↑ **Rocket Lab's** Electron is the second most frequently launched U.S. rocket.

Figure 7.3.1

Mesh networking in space gives members of satellite constellations the ability to fetch or pass data at high speed and low latency.



is building its own constellation of broadband internet satellites, called Leo, and has already launched hundreds of satellites with laser crosslinks operating at 100 Gbps.

Researchers, meanwhile, are extending optical laser communications far beyond LEO. In 2023 NASA's Deep Space Optical Communications demonstration established an ultra-high-definition, 267 Mbps video link with the Psyche spacecraft over 19 million miles away. And earlier this year the European Space Agency used laser communications to maintain a 2.6-Gbps link between an aircraft in flight and a satellite in geosynchronous orbit—a major engineering feat, given that both targets were moving, and the Alphasat TDP-1 satellite was orbiting 22,300 miles above the Earth. The Space Development Agency, the Pentagon's rapid procurement office for space technology, has also demonstrated two-way laser communications between a satellite and an aircraft in flight—a capability that will be integral to the agency's Proliferated Warfighter Space Architecture, a planned constellation of thousands of small military satellites in LEO.

The commercial opportunities are broad. Laser links require high-precision optical ground terminals that can transmit through clouds and atmospheric turbulence, as well as global ground-station networks. Future constellations will require standardized, mass-produced laser terminals and optical switches with built-in routing protocols. Though companies such as Mynaric and Skyloom have begun to explore this market, there isn't yet a "Cisco for space networks." These are the fields O'Connell says he's scouting. "Is laser comms to LEO and MEO cost-effective?" he asks. "Can it be faster and more affordable? I don't know yet."

The connections these networks support aren't just human-to-human—increasingly, they'll be machine-to-machine. Researchers and entrepreneurs are exploring whether future data centers could operate in space, where they'd have access to limitless solar power—in fact, DCVC-backed Planet is building two prototype satellites for Google as part of its Project Suncatcher moonshot, testing whether it would be effective to scale AI in space on networks of orbiting, laser-connected TPUs. Ground-to-satellite lasers might also evolve into an ideal medium for quantum key distribution (QKD) for eavesdropper-proof digital communications, which we mentioned in section 2.4. The quantum computing and networking company IonQ bought DCVC-backed Capella Space in 2025 as a step in the development of its own space-based QKD network.

Space is gradually acquiring the same networked architecture that transformed computing on Earth in the early internet era. Launch vehicles made space accessible—but networking technologies will determine what we can actually do there, and how the billions of us still here on Earth will benefit.

Throughout this report we've been arguing that deep tech innovation is our most effective tool for restoring predictability in a world experiencing exponential change. And the deepest opportunities, we believe, don't merely digitize the surfaces of our old systems: they reach down into those systems' innards to make them more legible, controllable, reliable, and affordable.

→ In April 2026 astronauts aboard **NASA's Artemis II** lunar flyby mission captured this image of our lush, delicate planet about to pass behind the Moon.

That, to us, is the real through line of deep tech in 2026: reducing uncertainty by turning natural and physical systems into programmable ones. Not in the sense that nature becomes code, but in the more powerful sense that once a system can be measured with precision, modeled with fidelity, and acted upon in tight feedback loops, its behavior becomes more designable. A geothermal reservoir, a tumor microenvironment, a drone-filled airspace, a wastewater stream: each begins as something noisy and variable, but becomes investable when new tools make performance more repeatable.

In a time of exponential change, that's what matters most. Earth doesn't need more technologies that pile complexity onto brittle systems, demand ever-escalating spending, or rely on permanent subsidies or favorable politics. It needs ways to reconceive essential systems so that they're faster to build, cheaper to run, and harder to break. The companies we back operate in a volatile world—but they are redesigning the fundamentals of computing, energy, industry, biology, and security to make that world more predictable and engineerable.

Thank you to the following DCVC Partners, Principals, and Associates

Zachary Bogue
Alan Cohen
Steve Crossan
Dr. Vandon Duong
Dr. John Hamer
James Hardiman
Clay Hutmacher
Earl Jones
Dr. Justin Kern
Dr. Josué López
Dr. Linda Maxwell
Dr. Prineha Narang
Matt Ocko
Matt O'Connell
Dr. Emily Park
Jason Pontin
Kristen Rocca
Jack Sauls
Dr. Rachel Slaybaugh
Dr. Kiersten Stead
Ali Tamaseb
Richard Waines
Milo Werner
Cara Yesko

Editing Nate Nickerson, Jason Pontin
Writing Wade Roush
Data, Graphics, Fact Checking, and Copy Editing Matt Mahoney
Photo Research Hailey Hiss
Design One Design Company, Chicago, IL

Typefaces FK Grotesk Mono, Kleukens-Egyptienne, Onsite
First Published June 2026

Index of Companies (DCVC portfolio companies are shown in bold type)

DCVC

Deep Tech Opportunities Report

2026

270 University Avenue
Palo Alto, CA 94301

500 2nd Street, Floor 2
San Francisco, CA 94107

dcvc.com

IN AN AGE OF EXPONENTIAL CHANGE