Spectra Industries

White Paper

Sofiane Croisier

February 2026

PRIVATE AND CONFIDENTIAL - NOT TO BE SHARED

Brief

The pace of intelligence is now quicker than our ability to make sense of it.

2024 saw conversational AI.

2025 was agentic AI.

2026 will be physical AI.

As intelligence becomes explosive and embodied, progress is no longer limited by algorithms only, but by our ability to feed it with knowledge of the real word in which it will live.

We are building the Matrix for Physical AI, a new infrastructure to solve today’s main bottleneck for the training and safe adoption of autonomous and robotic systems: data scarcity.

Problem

"The ChatGPT moment for physical AI is here" said Jensen Huang in early 2026.

But it is slow.

Because data is missing.

Training robots' brains at scale is either a bankruptcy exercise (collecting real-world data can cost billions of dollars), or technologically impossible today (traditional simulators have perfect math but Tetris-like visuals and GenAI has Hollywood-grade content but remains a slot machine that randomly hallucinates physics).

This has created a tall data scarcity wall that prevents makers of robotics, drones and industrial automation from moving forward in training and deploying safe and accurate physical AI in our daily lives and across industries.

Physical AI does not need a "ChatGPT moment".

What it needs is an ImageNet moment.

Solution

Our technology can be summarized as a simulation infrastructure for rapid AI policy learning, evaluation and transfer.

The goal is to drastically accelerate the time-to-market and usefulness of robotic intelligence (as measured by sim-to-real performance). That is, we want to collapse physical AI's training cost & time and boost its adoption.

Our Neuro-Kinematic Engine (NKE), a world-generating interface that can create scenarios that both generate high visual realism (neuro-symbolism: the "skin") *and* respect the laws of motion physics (kinematics: the "skeleton"), will drive our one-stop-shop data platform and push the frontier of physical AI.

To the best of our knowledge, this has never been achieved at scale in the market.

Our NKE will manufacture reality into multiple parallel world scenarios to provide a hyper-fast "training arena" for robotic AIs.

To train a robot to fold a shirt or land on Mars, we will generate not just the visuals of these scenarios but also their constitutive physical equations and edge-case variations (1% humidity, 10% humidity, 25% humidity, etc.).

By creating custom and cost-effective "synthetic nervous systems" that can be downloaded into robots, we will aim to eventually provide them with 1,000 days of real-world experience in ~1 hour.

Here is the link to the demo of our product vision:

https://spectra-industries.netlify.app/

Why now

The timing is optimal because access to simulation will be a natural chokepoint within the next 5-10 years as demand for autonomy across driving, manufacturing, delivery and everyday household tasks exceeds the supply of reliable training data.

For example, to statistically validate safety in autonomous driving, engineers must empirically test for extremely rare events (edge cases). Real-world testing alone demands billions of kilometres driven (only Tesla can afford to do it today).

As humanity technologically expands, data will remain physically inaccessible in some use cases (e.g. autonomous deployment of solar-powered data centres in space, rover explorations on the Moon and Mars).

The way we want to solve this massive painpoint is to simulate and download in robots' ML policies "long tail" edge cases with up to a ~1,000x cheaper time and cost curve.

We will also build a special *Chaos* feature to inject synthetic Perlin noise such as tremor, lag, hesitation and recalibration to make our platform-trained robots statistically robust against real-world chaos.

As physical AI scales, real-world data becomes the fixed resource. Simulation is the only elastic one.

When training and deployment velocity is essential, access to high-fidelity simulation becomes the rate limiter. Companies that control that chokepoint will become infrastructure.

Market potential

We estimate the expanded TAM (autonomy training for industrial and household robots, drones, automotive, aerospace, CAD-based construction, AI-powered factories, digital twins for smart city planning, etc.) to be around ~$100B and growing at a ~20-30% CAGR rate.

We believe that a defensible synthetic data & validation infrastructure for physical AI could capture 5% of that market over 5 years, creating a $5B ARR company.

Even at a conservative half of today's ~25x average multiple for AI startups, this could pave the path towards a decacorn valuation.

But first, our focus is to pressure-test and validate product-market fit not with LOIs but with real demand signals: working with 1-3 robotics startups to develop a robust proof-of-concept.

This will enable us to calibrate & deploy the engine and back-propagate real market learnings into our product roadmap.

Differentiation

Three key moats will make our business defensible:

1. Foundational technology moat

We are building a quantum-ready, physics-valid and photorealistic simulation IP purpose-built for physical AI training & deployment and architected for high-throughput.

This vertically integrated stack is deeply technical and difficult to replicate. It will be protected not only by aggressive anti-reverse-engineering NDAs, but also by an automatic prompt optimizer, quantum processing offramps, accumulated physics solver tuning and sim-to-real know-how that compounds over time.

As enterprise robotic adoption increases, we will further harden our technical moat by researching and testing cutting-edge technologies:

- *New RL methods such as "Reinforcement Learning by Scenario Tuning" (RLST)* to fine-tune robotic behaviours by "domain specialization" for specific, scalable use cases. For example, teaching a solar-powered delivery drone to find the optimal sunlit spot and angle every 2 hours to recharge its battery throughout the day.

- *Self-recursive calibration pipeline* to continuously learn and integrate residual corrections that cannot be easily simulated as priors. For example, if "air friction" cannot be pre-computed in the simulated environment, our AI will perform micro-experiments and generalize its findings back into the policy and simulator.

- *Acoustic simulation* to further collapse the sim-to-real gap by amplifying robots' contact awareness (e.g. grasp slip) and failure detection (e.g. motor strain). Today, robotic software focuses on vision and kinematics. Hearing is an underserved channel for continuous, omnidirectional bandwidth. There will likely be a first-mover advantage to the platform that exploits that gap early and builds the capability natively in its stack.

2. Data and ecosystem moat

Reusable, sector-specific environment assets (e.g. training packs for logistics, manufacturing, mobility, government) will encode task distributions, failure modes, and physics regimes and create a compounding data flywheel.

As customers train, validate, and deploy within these environments on our platform, they will continuously create an anonymized shared empirical prior that improves training efficiency and policy generalization for the entire industry, helping it grow its size over time.

In the long run, like CUDA did for GPU optimization, we will become the main language and system of record for physical AI workflows, making switching costs extremely painful.

3. Adoption and workflow moat
An Apple-style, user-friendly UX for complex simulation workflows dramatically lowers friction for robotics and ML teams. It will provide a clear performance snapshot of how their new robot performs.

By abstracting away complexity behind an intuitive, opinionated interface, we aim to shorten time-to-value of robotic hardware and embed ourselves deeply into daily engineering loops.

This makes our infrastructure hard to replace not because it is complex-to-copy for customers, but because removing it from their workflows would materially damage their P&L.

Techno-business model

We aim to build a healthy operation from the start, not a spray-and-pray compute spending business.

The platform will be priced as core infrastructure, not a dev tool. Pricing will have two main components:

1) *Base license priced per customer size* (~$50k–$200k for startups, ~$500k–$2M+ for enterprises) that unlocks the simulation engine including core physics, APIs and the validation dashboard.

2) *Usage-based sim-hours* priced by fidelity tier (~$2–$5 per blended sim-hour, with weighted-average costs of ~$0.50–$1 including GPU/CPU, orchestration and storage costs), driving a ~70% gross margin target. This should create a revenue scaling flywheel as demand for physical AI matures and customers train more robots on our platform.

We will pre-emptively protect blended margins by:

- Selling validated environments (not compute hours) which will be cheaper to customers and enable us to package and redeploy sectorized environments, cache core physics artifacts and negate customer-specific compute paths by default.

- Including lock-in factors in enterprise contracts, such as explicit fidelity tier pricing, minimum sim-hour commits, burst pricing above commits and/or at peak usage hours.

- Making price elastic to the number of hardware units being trained on our simulation platform after a certain threshold (e.g. additional 5% pricing > 1,000 trained and validated units).

- Turning the highest photoreal rendering off by default (customers pay blended and we incur weighted-average costs).

- Approaching sectors with a land-and-expand mindset through clear upsell pathways, such as 4K rendering mode, sensor features, monthly autonomous validation reports to keep robots' brains "fresh", enterprise options (e.g. on-prem deployment, air-gapped, premium SLAs).

Future vision

We see our robotics training platform as an "orchestrating pipeline" play that will be our beachhead entry into a much bigger market that does not exist yet and that we want to help create.

Over time, we want to build a foundational simulation technology at the frontier that can define and anchor two new categories: reality-grade simulative AI (for firefighters' training, smart city planning or space exploration) and next-gen entertainment ("prompt what you want to experience" games and movies).

LLM benchmarks have become useless as they keep spawning and shifting the goalpost of intelligence. Simulative intelligence is the new category we want to shape and will be the last irrefutable evidence that AGI has arrived. That is because we will "feel" this intelligence, not just read it.

To facilitate this long-term vision, we plan to build in our very first product iteration the quantum primitives to offload the meshes and matrices of our procedural world generation to quantum computers (e.g. IBM, Rigetti, etc.) when are sufficiently mature.

This quantum readiness will eventually make our infrastructure exponentially faster than incumbent competitors like Isaac Sim, MuJoCo or Unreal Engine, giving us a significant headstart and displacing market power in the next technological race.

By helping to make tomorrow’s autonomous hardware safer and smarter with more speed, our technology aims to power global GDP and return human time so we can live richer lives.