



ARTIFICIAL INTELLIGENCE
LABORATORY · MONACO

FRONTIER
RESEARCH

PHYSICAL INTELLIGENCE · MARCH 2026

The Next AI Trade

How comparative cognition informs the road to AGI.

Leonardo Fabbri
Founder & Chief Executive

WE READ THE RESEARCH,
YOU GET THE INSIGHTS

EXECUTIVE SUMMARY

Evolution never optimised for conversation. It optimised for prediction — and prediction is where the value moves next.

A dragonfly intercepts prey on a curve it computes before the prey arrives. It does this in about fifty milliseconds — roughly three synapses of processing — and it succeeds about 95% of the time. No large model, no data centre, no training corpus. Whatever intelligence is doing in that animal, it is not describing the world. It is getting ahead of it.

In our December 2025 note, *From Words to World Models*, we argued that frontier AI systems can explain physics fluently and cannot predict it, and that the discriminating variable is architectural rather than one of scale.¹ This note takes the next step and asks the commercial question: if grounded prediction is the missing capability, where does the value accrue?

The biology sharpens the argument considerably. Mischiati and colleagues showed in *Nature* that dragonfly interception is driven by **internal forward and inverse models** of the animal's own body dynamics together with a model of prey motion — with vision reserved for correcting the unexpected.² Larval zebrafish, with roughly 100,000 neurons in total, combine prey position with prey velocity to project where the target will be; the velocity term alone improves hunting success by about 65% over a position-only strategy.³

Put beside the infant evidence from the December note, the pattern is hard to miss. Three systems separated by 500 million years of evolution, spanning five orders of magnitude in neuron count, all converge on the same architecture: a compact internal model, rolled forward, corrected by sensory error. None of them is running language.

The market implication is a rotation, and the first leg of it is already visible. Software-as-a-service

accounts for less than 0.5% of global GDP. The real economy — manufacturing, logistics, energy, construction, defence — is the other 99.5%, and AI has barely touched it.⁴ Meanwhile the enterprise software complex has de-rated hard: the iShares Expanded Tech-Software ETF fell roughly 17% over the first half of 2026 and about 26% from its October 2025 high, with forward multiples compressing from around 35× to 22× — the lowest since 2014.⁴

Capital is moving to the other side. Prometheus raised **\$12bn** at a ~\$41bn valuation in June 2026 to build an "artificial general engineer"; Anduril raised \$5bn at \$61bn; Figure's humanoids ran ten-hour shifts on a BMW line for five months, contributing to more than 30,000 vehicles.⁴ Morgan Stanley sizes the humanoid market at **\$5tn by 2050** across nearly a billion units.⁵

Our central caution is about implementation, not thesis. The trade is real and remarkably hard to own. The obvious public vehicle — the Global X Robotics & AI ETF — is 65 holdings and \$3.4bn, with its top weights in Keyence, Nvidia, ABB and Fanuc: a portfolio of twentieth-century factory automation, not of embodied intelligence.⁶ Every pure expression of this thesis is private, pre-revenue, and financed by venture and growth equity with no credit wrapper behind it.⁴ Being right about the direction and early on the instrument is the characteristic way this trade loses money.

CONTENTS

What is in this note

01	The Predator's Advantage: what a dragonfly actually computes	05
	Forward and inverse models, prey-centric steering, and the fifty-millisecond budget	
02	Convergent Evidence: from babies to dragonflies	08
	Three systems, five orders of magnitude of neurons, one architecture	
03	The Efficiency Case: prediction is cheap, description is expensive	10
	What the biology implies about the cost structure of physical intelligence	
04	The Trade: where the value moves	12
	The 99.5%, the software de-rating, and the capital already committed	
05	The Vehicle Problem: why this trade is hard to own	16
	What a robotics ETF actually holds, and the financing gap beneath the thesis	
06	The Cost Curve and the Adoption Path	19
	Unit forecasts, price compression, and where the honest disagreement sits	
07	Risks to the Trade	21
	Five ways to be right about the science and lose money anyway	
08	Conclusion and Watch List	23
	Dated signals we are tracking through 2028	
—	References	24
—	Methodology, Scope and Important Disclosures	25

EXHIBITS

- | | |
|--|--|
| 1 · The claim in one chart | 8 · Capital committed to physical AI |
| 2 · What a dragonfly computes | 9 · What a robotics ETF actually holds |
| 3 · Three predictive systems, compared | 10 · Three risk shapes, one thesis |
| 4 · What the December note established | 11 · The humanoid cost and unit curve |
| 5 · Prediction is cheap; description is not | 12 · What has to be true, and by when |
| 6 · The 99.5% the technology has not touched | 13 · Watch list to end-2028 |
| 7 · The software de-rating, measured | |

KEY JUDGMENTS

Five conclusions, and the evidence each rests on

01 Biology converges on prediction, at every scale it has been measured

A dragonfly runs forward and inverse models of its own body plus a model of prey motion. A zebrafish with 100,000 neurons combines position and velocity to project a target forward. An eleven-month-old expects solidity and continuity before it can speak. Five orders of magnitude of neural hardware, one architecture — and none of it linguistic.

02 The efficiency gap is the commercially interesting number

Fifty milliseconds and roughly three synapses buys a dragonfly a 95% capture rate. Our December note found intuitive physics emerging in a 115-million-parameter model trained on 128 hours of video, while systems thousands of times larger sat near chance. Both point the same way: grounded prediction is not what is expensive about AI. Description is.

03 The rotation has begun on the software side and is not yet visible on the other

Enterprise software de-rated roughly 17% in the first half of 2026 and 26% from its October 2025 high, with forward multiples at their lowest since 2014. The offsetting bid is going into private physical-AI platforms — \$12bn into Prometheus alone. One leg of the trade is priced in public markets; the other is almost entirely outside them.

04 The thesis is easier to state than to own

The main listed robotics vehicle holds \$3.4bn across 65 names, led by Keyence, Nvidia, ABB and Fanuc. That is high-quality factory automation and machine vision — the previous generation of the technology, not this one. An investor buying it is buying industrial capex cyclical with an AI label, and should know that is the exposure.

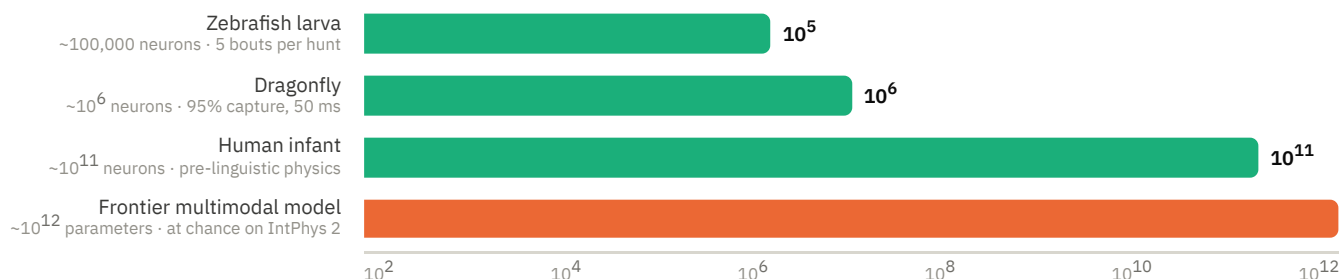
05 Timing risk sits in the financing structure, not in the science

Physical AI is funded by venture and growth equity, with late-stage private capital absorbing venture-stage technology risk and no credit wrapper for hardware-intensive deployment. Broad commercial deployment is not expected before 2027–2029. A thesis can be correct and still be three funding rounds early.

EXHIBIT 1

The claim in one chart: what prediction costs in nature

Neuron count, log scale. Each system solves a prediction problem under occlusion, motion and time pressure. The frontier model is included for scale, not as a like-for-like comparison.



Source: Bolton et al., *eLife* (2019); Sandia National Laboratories (2019); Mischiati et al., *Nature* (2015); AI Lab Monaco, *From Words to World Models* (2025). **Note:** parameters and neurons are not equivalent units and the bottom bar is shown for order-of-magnitude context only. Dragonfly neuron count is an order-of-magnitude estimate for odonates. The point of the chart is the ordering, not the arithmetic.

What a dragonfly actually computes

You cannot catch a moving target by aiming at it. By the time you arrive it has gone. Interception requires aiming at where the target *will be* — which means running a model of it, and of yourself, faster than the world updates.

This is a control problem before it is a biology problem, and it has a hard constraint at its centre: **delay**. Light hits the retina, signals propagate, muscles contract. By the time a sensory measurement has become a motor command, the measurement is stale. Any animal that hunts fast-moving prey must therefore act on a prediction, not on a perception.

In 2015 Matteo Mischiati and colleagues at Janelia published the cleanest available account of how one animal solves it. Working with *Plathemis lydia*, they tracked head and body orientation simultaneously in three dimensions, using retroreflective markers and high-speed videography at 1,000 frames per second — a resolution that finally made the internal computation visible from the outside.²

What they found was not a reflex. Dragonfly interception steering is driven by three distinct internal models operating together: a **forward model** of the animal's own body dynamics, predicting the consequences of its own motion; an **inverse model**, converting a desired flight state into the motor commands that achieve it; and a **model of prey motion**. Predictive rotations of the head

continuously track the prey's angular position, and the resulting head–body angle drives systematic body rotations that align the animal with the prey's flight path.

The sentence in that paper which matters most for our purposes is the one about vision: model-driven control accounts for the bulk of interception steering, while **vision is used for reactions to unexpected prey movements**.

Read that carefully. In the dragonfly, sensing is not the control loop. The internal model is the control loop, and sensing is the exception handler — the channel that fires when the model is wrong. This is precisely the inversion that separates a predictive architecture from a reactive one, and it is the same inversion we argued for on architectural grounds in the December note.¹

The performance this buys is the part that travels. A dragonfly's reaction time to manoeuvring prey is about **fifty milliseconds** — enough time for information to cross roughly three neurons. Capture rate is about **95%**.⁷ A human blink takes six times longer.

EXHIBIT 2

What a dragonfly computes, and where vision actually enters

The default path is model-driven. The sensory channel is a correction, not the control loop — the architectural point of this note.



The model carries the steering

Predictive head rotations track the prey's angular position continuously, so the animal is already oriented toward an interception course before any correction is needed.

Vision handles surprise

Sensory input is reserved for reactions to unexpected prey movement — the cases the model did not anticipate.

Source: Mischiati, M., Lin, H.-T., Herold, P., Imler, E., Olberg, R. & Leonardo, A., "Internal models direct dragonfly interception steering", *Nature* 517, 2015, pp. 333–338. AI Lab Monaco rendering.

The same solution, at a hundredth of the hardware

If the dragonfly result stood alone it would be a curiosity about one unusually good predator. It does not stand alone, and the corroborating case is more useful because the hardware is smaller and the computation has been characterised more explicitly.

Larval zebrafish hunt by a sequence of discrete swimming bouts, which makes the underlying strategy legible in a way continuous flight does not. Bolton and colleagues, working across Harvard and MIT, showed that the fish combines the prey's **position** — azimuth, altitude and computed radial distance — with its **velocity**, and projects the target forward in time in a manner analogous to Euler's method.³

The size of that velocity term is the finding. A model using position alone is materially worse than one using position and velocity together: adding the predictive component improves hunting success by roughly **65%**. The fish is not tracking. It is extrapolating.

The strategy is also elegantly cheap. Each bout reduces the angular error to the prey by about half — a halving rule that converges without requiring precise one-shot computation. The animal takes an average of five bouts per successful hunt, with bouts lasting around 176 milliseconds.

Two further details deserve attention from anyone thinking about robot policies. First, the authors found that adding *stochastic noise* to an otherwise deterministic algorithm improved capture speed by 14–25% — randomness as a feature of the search, not a defect in the controller. Second, the whole thing runs on a brain of roughly **100,000 neurons**.

Two animals, separated by several hundred million years of evolution and two orders of magnitude of neural hardware, converge on the same design: hold an internal model, roll it forward, correct on error. Neither of them is running anything that resembles language.

EXHIBIT 3

Three predictive systems, compared

The infant column summarises the evidence base of our December 2025 note; the two animal columns are the new material in this one.

	DRAGONFLY <i>Plathemis lydia</i>	ZEBRAFISH LARVA Prey capture	HUMAN INFANT 11 months, pre-linguistic
Neural scale	~10 ⁶ neurons; ~3 synapses in the interception path	~100,000 neurons in the entire brain	~10 ¹¹ neurons, mostly uncommitted
What is modelled	Own body dynamics (forward and inverse) plus prey motion	Prey position and velocity, projected forward	Object permanence, continuity, solidity, support
Time budget	~50 ms reaction to manoeuvring prey	~176 ms per bout; ~5 bouts per hunt	Not time-critical; measured by looking time
Role of sensing	Correction of unexpected prey movement only	Re-estimation between discrete bouts	Registers violation as surprise, then investigates
Measured performance	~95% capture rate	+65% success from the velocity term alone	Learns from the violated case, not the ordinary one
Shared architecture	A compact internal model, rolled forward under uncertainty, corrected by prediction error — acquired without instruction and without language		

Source: Mischianti et al., *Nature* (2015); Sandia National Laboratories (2019); Bolton, A. D. et al., “Elements of a stochastic 3D prediction engine in larval zebrafish prey capture”, *eLife* (2019); Stahl & Feigenson, *Science* (2015) via AI Lab Monaco (2025). Neuron counts for dragonfly and infant are order-of-magnitude estimates.

In the dragonfly, sensing is not the control loop. The model is the control loop, and sensing is the exception handler. Almost every deployed robot today has that relationship the other way round.

Why an insect result belongs in an investment note

The obvious objection is that dragonflies are not robots, that evolution is not engineering, and that biological analogies have a long history of misleading technologists. All three are fair, and the argument survives them for a narrow reason.

The dragonfly is not offered as a design template. It is offered as evidence about **where the difficulty lies**. If acting competently in a physical world required enormous representational capacity, no animal with three synapses of processing latency could do it at all. Something with a millionth of a frontier model's parameters achieves 95% reliability on a task — intercepting an evading target in three dimensions — that no robot can currently perform.

That tells us the expensive part of physical competence is not the physics. It is everything we have built *around* the physics: the perception stack that re-derives the world from scratch every frame,

the language layer that describes it, the planner that searches without a usable forward model.

There is also a specific engineering lesson, and it is the one that converts a biology finding into a capital-allocation view. Delay is not a solved problem in robotics; it is the reason demonstrations look fluid and deployments do not. A system that re-perceives before every action inherits the full sensing-to-actuation latency on every step. A system that runs a forward model pays that cost only when it is *wrong*.

This is exactly the architectural claim our December note made from the machine side, arrived at independently from the biological side. Two lines of evidence converging on the same conclusion is the strongest position a research argument can occupy, and it is why we are prepared to underwrite the direction of this trade even while being cautious about its timing.

SECTION 01 · WHAT WE TAKE FORWARD

- Interception is a prediction problem, and biology solves it with internal forward and inverse models rather than with faster sensing.
- In the best-characterised case, vision is the correction channel and the model is the control loop — the inverse of how most deployed robotic systems are built.
- The performance is achieved at trivial neural scale: ~50 ms, ~3 synapses, ~95% capture. Capacity is not the binding constraint on physical competence.
- The zebrafish result quantifies the value of the predictive term directly: +65% success from adding velocity to position.

A system that re-perceives before every action pays the full sensing-to-actuation delay on every step. A system that runs a forward model pays it only when it is wrong.

From babies to dragonflies

This note has a predecessor, and the two are designed to be read together. The December argument was about architecture and was made from machine evidence. This one is about capital and is made from biological evidence. They meet in the middle.

In *From Words to World Models* we set out a narrow, falsifiable claim: frontier language and multimodal systems describe physical constraints fluently and fail to predict them, and the variable that separates the systems which succeed from those which fail is the space in which prediction happens — not model scale.¹

The evidence was benchmark evidence. Models trained to predict masked video in learned representation space acquire object permanence, continuity and shape constancy; models that predict pixels or text sit within a few points of an untrained network on the same protocol. On IntPhys 2 most state-of-the-art systems score at chance against a near-perfect human baseline; on CausalVQA the gap to humans widens from 23 points overall to 38 on questions asking what happens next.

The efficiency result was the one we flagged as most likely to reorder priors: the capability appeared in a

115-million-parameter model trained on **128 hours** of unique video, with no physics supervision at all.

What the biology adds is independent confirmation from a completely different direction. The December argument had to reason from a controlled comparison between four model families. The dragonfly and zebrafish results are not comparisons between architectures — they are existence proofs about what the task actually requires. And they say the same thing: the predictive core is small, and it is not linguistic.

Two independent lines of evidence converging matters more than either alone. A benchmark result can be an artefact of the benchmark. An animal result can be a curiosity of that animal. Both pointing at internal forward models, at very different scales, through completely different methods, is a harder thing to explain away.

EXHIBIT 4

What the December note established, and what this one adds

The two notes are a pair: one establishes the architecture, the other prices it.

	<i>FROM WORDS TO WORLD MODELS</i> December 2025	<i>THE NEXT AI TRADE</i> March 2026 — this note
Question	Can frontier systems predict physics, or only describe it?	If prediction is the missing capability, where does the value accrue?
Evidence base	Benchmarks: IntPhys 2, GRASP, InfLevel, CausalVQA, NEWTON	Comparative neuroscience, plus market and capital-formation data
Central finding	Prediction in learned representation space acquires physical expectations; prediction in pixels or tokens does not	Biological systems solve the same problem with internal forward models at trivial scale — and the capital has begun to move
On scale	115M parameters and 128 hours of video suffice; far larger systems fail	~10 ⁵ –10 ⁶ neurons suffice in nature; ~50 ms and 3 synapses buy a 95% capture rate
On language	The right substrate for abstraction and instruction; the wrong one for grounded prediction	Absent entirely from every biological system examined — and they perform the task
Joint conclusion	The predictive core of physical intelligence is small, cheap and non-linguistic. What is expensive is the description layer we have built around it — and the market is priced the other way round	

Source: AI Lab Monaco, *From Words to World Models* (December 2025) and this note. Benchmark figures as cited in the December note; biological figures as cited in §01 above.

Where the industry itself has landed

It is worth noting that this is no longer a contrarian reading inside the field. It has become, over roughly eighteen months, something close to the consensus of the people building the systems — which is a different thing from being priced.

Fei-Fei Li's formulation is the one most often quoted, and it is the sentence that gave this note its starting point: evolution did not optimise for conversation. Her company, World Labs, raised \$1bn in February 2026 to build spatial intelligence. Yann LeCun left Meta and raised \$1.03bn in March 2026 for AMI Labs, explicitly on the world-model thesis.¹

Goldman Sachs, writing from the banking side rather than the research side, describes the same

shift in almost the same terms: after a decade defined by systems that recognise patterns and predict text, the frontier is moving toward models that understand how the world works — and notes that if world models scale *alongside* language models rather than replacing them, current infrastructure sizing may prove undersized.⁴

That last clause is worth pausing on. The prevailing capex forecasts — roughly \$7.6tn across compute, data centres and power between 2026 and 2031 — are calibrated to transformer-based language models. They do not price a second engine. If the thesis in these two notes is right, the demand side of the buildout has an unpriced call option attached to it.

SECTION 02 · WHAT WE TAKE FORWARD

Two independent evidence bases — controlled machine benchmarks and comparative neuroscience — converge on the same architecture: a **compact internal model, rolled forward, corrected by error**. Neither requires language and neither requires frontier scale.

The view is now broadly shared by the people building the systems and by at least one major bank writing about the capital. It is not yet reflected in how the public market is positioned, which is the subject of the rest of this note.

A note on what we are not claiming

We are not claiming that language models are overvalued, that they will be displaced, or that the infrastructure buildout is misconceived. The December note was explicit that the endpoint is a hybrid — a world model for grounded prediction with language as the interface layer — and nothing here revises that.

We are claiming something narrower: that the capability which unlocks the physical economy is architecturally distinct from the one currently being capitalised, that biology suggests it is cheap once you have it, and that the market has not yet built a clean way to own it. Sections 04 and 05 are about those last two points.

Prediction is cheap. Description is expensive.

Put the two efficiency results side by side and a commercial argument falls out of them. Roughly three synapses and fifty milliseconds buy a 95% capture rate in nature. A 115-million-parameter model trained on 128 hours of video acquires infant-like physical expectations. Neither is a frontier-scale system, and the frontier-scale systems fail the same tests.

The inference to draw is not that big models are useless. It is that the expensive part of what we have built is not the part that produces physical competence. That capability appears to be — in nature and in the one machine architecture that has demonstrated it — structurally compact.

The distinction lands directly on the economics. Goldman Sachs frames the split cleanly: training is the up-front capital event, while **inference is the operating cost**, running continuously across every interaction, and inference is where AI economics ultimately resolve.⁴ A predictive core small enough to run inside the machine changes the operating cost of an embodied system more than any improvement in the model describing what the machine is doing.

It also determines *where* the computation can live. Among the five factors Goldman identifies as

separating industrial AI winners from commodity players, one is edge deployment: most industrial AI must run inside the machine, on the embedded controller — not in a cloud data centre.⁴ A control loop that makes a round trip to a data centre before acting has reintroduced exactly the delay problem §01 described the dragonfly solving.

Two qualifications, because the argument is easy to overstate. **Cheap to run is not cheap to build:** V-JEPA 2 reached its action-conditioned capability at 1.2 billion parameters on over a million hours of video, so training may remain a frontier-scale capital event even where deployment is not. And **we are not making an energy claim** — a rigorous joules-per-decision comparison between an insect nervous system and a silicon inference stack does not exist in citable form. The claim is about parameter scale, synaptic depth and latency, all of which are measured.

EXHIBIT 5

Where the cost actually sits in a physical AI system

The layer that produces physical competence is the cheapest to run — and the one the market is least exposed to.

LAYER	WHAT IT DOES	COST CHARACTERISTICS	EVIDENCE ON SCALE REQUIRED
Predictive core	Carries state, rolls it forward under action, flags its own error	Small, latency-critical, must run at the edge	~10 ⁵ –10 ⁶ neurons in nature; 115M parameters, 128 hours of video in the machine case
Action-conditioning	Maps intentions onto the transition model of a specific body	Data-bound, not compute-bound — the sector's binding constraint	62 hours of unlabelled robot interaction data in the V-JEPA 2 disclosure
Perception stack	Reconstructs the scene from raw sensors, frame by frame	Continuous, heavy; re-derives what a good model already holds	Dominates current robot compute budgets — the layer biology mostly does without
Language interface	Goal interpretation, instruction, explanation, abstraction	Frontier-scale training; inference priced per token	Where essentially all current AI capital expenditure is concentrated
The asymmetry	The layer that determines whether a machine can act reliably in the world is the one requiring least capacity — and the one with almost no listed exposure		

Source: AI Lab Monaco analysis, drawing on Garrido et al. (2025) and Meta AI's V-JEPA 2 disclosure (2025) as cited in our December note, and Goldman Sachs, *Harnessing AI for the Real Economy* (2026) on the training/inference cost split and edge-deployment requirement.

The unpriced second engine

There is a version of this argument that reads as bearish on the infrastructure buildout, and we want to be clear that we do not hold it. The opposite case is at least as plausible and is better supported.

Current capital-expenditure forecasts — on the order of **\$7.6tn** across compute, data centres and power between 2026 and 2031 — are calibrated to transformer-based language models. Goldman's own framing is that this is a reasonable base case which assumes existing model architectures define tomorrow's compute demand, and that if world models scale alongside language models rather than replacing them, the total *may prove undersized*.⁴

There is a second demand channel that compounds it, and it is systematically under-modelled. As

autonomous systems proliferate, each deployed machine becomes a continuous generator of data that must be ingested, modelled and acted on. Physical deployment is not only a consumer of AI capability; it is a fast-growing source of the compute and the power the next phase will require — and that load sits largely outside projections calibrated to today's language-model usage.⁴

So the efficiency argument cuts in a specific direction. It says the *per-unit* cost of physical intelligence is lower than the per-unit cost of linguistic intelligence. It does not say the aggregate is smaller — because the unit count is the entire physical economy rather than a knowledge-worker seat count. Cheap per instance, enormous in aggregate, is how electrification behaved too.

SECTION 03 · WHAT WE TAKE FORWARD

The predictive core is small in every system where it has been measured, must run at the edge, and is data-bound rather than compute-bound. That is a different cost structure — and a different competitive structure — from the language layer.

This is not a bear case on infrastructure. If world models scale alongside language models, aggregate compute demand carries an **unpriced second engine** that current sizing does not include.

Cheap per instance, enormous in aggregate. That is not a contradiction — it is how electrification behaved, and it is the shape of the opportunity here.

Where the value moves

The whole of the software industry that AI has so far transformed is less than half a percent of global output. The argument for the rotation is arithmetic before it is anything else.

Software-as-a-service accounts for under 0.5% of global GDP. The technology sector as a whole is roughly 5%. The remaining **99.5%** — manufacturing, logistics, energy, construction, agriculture, defence, healthcare delivery — is the part AI has barely touched, and it is the part that requires a machine to do something in the world rather than describe it.⁴

That is the size of the prize, and it is worth being precise about why it has stayed out of reach. It is not that the industries are technologically unsophisticated. It is that acting in a physical environment demands exactly the capability §§01–

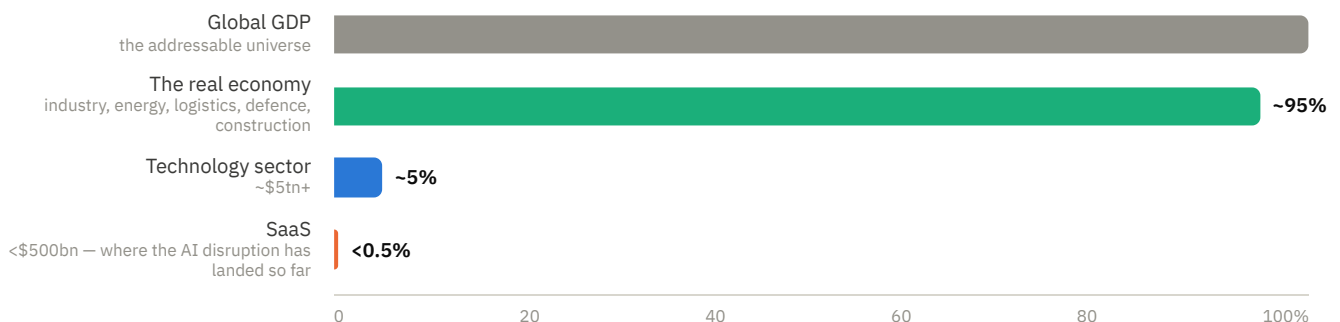
03 have been describing, and until recently nothing had it.

The evidence that this is changing is no longer only architectural. Figure's humanoids ran ten-hour daily shifts for more than five months on BMW's line in Spartanburg, contributing to the production of over 30,000 vehicles across roughly 1,250 operating hours. Siemens runs AI-driven predictive maintenance at its Amberg plant with a 40% reduction in production downtime.⁴ These are narrow deployments. They are also real ones, on real production lines, with real output attached.

EXHIBIT 6

The 99.5% the technology has not touched

Share of global GDP. The category AI has transformed so far is the narrow band at the bottom.



Source: Goldman Sachs, *Harnessing AI for the Real Economy* (2026), citing Precedence Research and IMF data; global GDP in excess of \$100tn. “The real economy” is shown as the residual after the technology sector and is indicative rather than a defined statistical aggregate.

Why the arithmetic only now has a buyer

That ratio has been true for a decade and did not constitute a trade. What has changed is that the demand side now has a hard constraint behind it rather than an efficiency argument in front of it.

The United States has roughly 13 million manufacturing workers, more job openings than available workers, and over one million unfilled materials-handling roles.⁴ That is not a productivity story in which a machine competes with a person on cost. It is one in which the work does not get done at all.

The same constraint recurs one layer down, in the buildout itself. Goldman estimates the US power and grid value chain alone needs more than **500,000 additional workers by 2030**, in roles requiring three to four years of training — against an apprenticeship pipeline running at 45,000 entrants a year where 65,000 would be needed.⁴

Labour scarcity is what converts a large addressable market into an urgent one. It is also why the first credible deployments have appeared in materials handling and warehousing rather than in the domestic settings that dominate the promotional videos.

The first leg is already priced

A rotation has two sides, and one of them has already happened in public markets — quickly, and with very little of the narrative attached to it that this note is making.

Enterprise software has de-rated hard. The iShares Expanded Tech-Software Sector ETF fell approximately 17% over the first half of 2026 and roughly 26% from its October 2025 high, with the sector's ten largest holdings shedding close to \$800bn of market capitalisation over the same span. Forward multiples compressed from about 35× in

late 2025 to about 22× — the lowest level since 2014.⁴

The proximate cause is not the argument in this note. It is that AI-generated code and agentic workflows are disrupting software itself, and the market cannot yet tell which incumbents will refund themselves. But the effect is the same: capital is leaving the application layer and moving toward the infrastructure beneath it — and, increasingly, toward the physical layer beyond it.

EXHIBIT 7

The software de-rating, measured

iShares Expanded Tech-Software Sector ETF (IGV) and its constituents, 2026 through June.

-17%

over the first half of **2026**; -26% from the October 2025 high

-\$800bn

of market capitalisation shed by the sector's **ten largest** holdings

22×

forward earnings, down from ~35× in late 2025 — the **lowest since 2014**

Source: Goldman Sachs, *Harnessing AI for the Real Economy* (2026). Figures as of June 2026 and therefore postdate this note's cover date; see the publication history on the disclosures page.

The second leg is happening off-exchange

The offsetting bid is real, large, and almost entirely private. Prometheus — an "artificial general engineer" for designing physical systems, co-founded by Jeff Bezos and Vik Bajaj — raised **\$12bn** in June 2026 at a roughly \$41bn valuation, among the largest early-stage investments on record. Anduril raised \$5bn at a \$61bn valuation in May 2026. Mobileye paid \$900m for Mentee Robotics in January 2026.⁴

Alongside them sit the world-model platforms we covered in December: AMI Labs at \$1.03bn, World Labs at \$1bn.¹ And alongside those, the data-supply layer: Galbot raised roughly \$362m in early 2026 specifically to address the robot-interaction-data bottleneck.⁴

Note what is absent from that list — a public-market instrument. Every name in it is private. That is the subject of §05, and in our view it is the single most important practical fact about this trade.

EXHIBIT 8

Capital committed to physical AI and adjacent platforms

Disclosed primary rounds and strategic acquisitions. Every entry is a private transaction.

COMPANY	DATE	ROUND	VALUATION	WHAT IT BUILDS
Prometheus	Jun 2026	\$12.0bn	~\$41bn	An “artificial general engineer” for designing physical systems, from jet engines to drug compounds. Co-founded by Jeff Bezos and Vik Bajaj
Anduril	May 2026	\$5.0bn	\$61bn	Autonomous defence systems across air, ground, maritime and counter-UAS, unified by a software backbone; partnered with OpenAI
AMI Labs	Mar 2026	\$1.03bn	\$3.5bn	World models on the JEPA thesis. Yann LeCun chairman; Nvidia, Samsung, Temasek and Toyota Ventures among backers
World Labs	Feb 2026	\$1.00bn	~\$5bn	Spatial intelligence; the Marble world model generates navigable 3D scenes from text or images
Mobileye / Mentee Robotics	Jan 2026	\$0.90bn	acquisition	Cross-layer consolidation — an autonomy platform buying a robotics stack, exactly what the full-stack thesis predicts
Galbot	Mar 2026	~\$0.36bn	n/d	Robot “data factories” manufacturing the interaction data that cannot be scraped
PhysicsX	Jun 2026	\$0.30bn	~\$2.4bn	AI-native physics simulation compressing aerospace, semiconductor and energy engineering from days to seconds
Disclosed total		~\$20.6bn		Across seven transactions in a seven-month window — none of it accessible in public markets

Source: Goldman Sachs, *Harnessing AI for the Real Economy* (2026); TechCrunch (March 2026); Reuters (February 2026); Caixin Global (March 2026). n/d = not disclosed. AI Lab Monaco compilation; figures as reported and not independently verified. Several transactions postdate this note's cover date — see publication history.

What has to happen for the rotation to complete

A de-rating on one side and private capital formation on the other is not yet a completed rotation. Three things have to occur, and only the first is under way.

Capability has to clear the reliability bar. Our December note put the warehouse-grade threshold at roughly 99% and found the best demonstrated generalist manipulation running 50–83% on precision tasks, with transfer to new objects and environments largely unmeasured.¹ Nothing in the biology changes that number; it only suggests which architecture is likeliest to reach it.

Unit economics have to become underwritable.

Goldman is explicit that funding is coalescing first where economics sit closest to existing frameworks — logistics automation, predictive maintenance, warehouse robotics — and forming unevenly or not at all elsewhere.⁴

A public instrument has to exist. Until the pure-play cohort lists, or the financing structures described in §05 mature, the trade is available to private capital and to almost nobody else.

Where the revenue actually is today

One correction to the popular version of this thesis is worth making plainly, because it is the difference between an investable view and a narrative one. Humanoids dominate the conversation and are not where industrial robotics generates revenue.

Millions of industrial robots, warehouse fleets and autonomous vehicles are in active use today. Next-generation fulfilment centres are being designed around ten-times-higher robotics density from the outset. Symbotix operates dark warehouses for Walmart. Amazon unveiled its latest Proteus autonomous mobile warehouse robot, responding to natural language, in June 2026. Caterpillar's autonomous mining fleet has hauled over 11 billion tons of material; Deere puts autonomy at the centre of its lineup.⁴

These are businesses with settled unit economics, existing customers and identifiable listed exposure. They are also, largely, *not* expressions of the architectural thesis in this note — they are conventional automation operating in structured environments, which is precisely why their economics work.

The honest framing is therefore a barbell rather than a single trade. Conventional automation is investable now and does not need world models to work. Embodied general-purpose systems need the architecture this note is about, and are not investable in public markets yet. Conflating the two is how the thesis gets sold, and how investors end up owning industrial capex cyclicality while believing they own embodied intelligence.

SECTION 04 · WHAT WE TAKE FORWARD

- The arithmetic is the argument: SaaS is under 0.5% of global GDP; the physical economy AI has not touched is the other 99.5%.
- The software leg of the rotation is priced — -17% in the first half of 2026, forward multiples at 2014 lows. The physical leg is forming privately, at roughly \$20bn of disclosed capital across seven transactions.
- Today's robotics revenue comes from conventional automation in structured environments, not from the architecture this note argues for. Do not confuse the two exposures.
- The rotation completes only when reliability clears ~99%, unit economics become underwritable, and a public instrument exists. One of the three is under way.

Why this trade is hard to own

Suppose you accept every argument in this note. The practical question is what you buy. The honest answer is that the obvious instrument gives you something quite different from what the thesis describes, and it is worth looking at the holdings rather than the label.

EXHIBIT 9

What a robotics ETF actually holds

Global X Robotics & Artificial Intelligence ETF (BOTZ), top 15 of 65 holdings, \$3.43bn in assets. Weights as of 14 August 2026.

HOLDING	WEIGHT	WHAT THE EXPOSURE ACTUALLY IS
Keyence	10.8%	Factory sensors and machine vision. Founded 1974
NVIDIA	9.4%	Semiconductors — the one holding with genuine physical-AI exposure, held by nearly every AI fund in existence
ABB	9.1%	Industrial automation and electrification. Founded 1883
Fanuc	7.8%	Industrial robot arms and CNC controls. Founded 1972
Intuitive Surgical	6.1%	Teleoperated surgical systems — a surgeon is in the loop, by design and by regulation
SMC	4.7%	Pneumatic components for factory automation
Inovance	3.8%	Chinese industrial automation and motion control
Daifuku	3.2%	Material-handling and conveyance systems
Aurora Innovation	2.5%	Autonomous trucking — one of the few holdings expressing the thesis directly
Cognex · Yaskawa Omron · RoboTechnik	8.6%	Machine vision, motion control and factory automation
Alphabet · AeroVironment	4.3%	Platform exposure and defence drones
Legacy factory automation and machine vision	~46%	Excellent businesses, and a bet on industrial capex cycles rather than on embodied intelligence. Exposure to the physical-AI cohort in Exhibit 8: zero — all of it is private

Source: stockanalysis.com holdings data for BOTZ as of 14 August 2026; AI Lab Monaco classification. The ETF is named as the category's most widely held vehicle, not as a recommendation for or against it — see the disclosures page. Weights sum to less than 100%; the fund holds 65 positions in total.

What no index can hold

The more revealing list is the one that cannot appear. Set the fund's holdings beside the companies actually building the capability this note describes and the two barely intersect.

Figure, 1X, Agility and Appttronik on the hardware side. Physical Intelligence and Skild on the policy side. World Labs, AMI Labs and the JEPa lineage on the world-model side. Prometheus on the

engineering side. Galbot on the data side. Not one of them is listed anywhere.^{1,4}

An index is a claim on what has already been financed in public markets. This category has been financed almost entirely outside them — which means the passive vehicle is structurally incapable of expressing the thesis, and will remain so until the cohort lists. That is not a criticism of the fund. It is a description of where we are in the cycle.

The gap between the thesis and the instrument

None of the companies in Exhibit 9 is a bad business. Keyence, Fanuc and ABB are among the highest-quality industrial franchises in the world. The point is narrower and more useful: they are the *previous* generation of this technology.

What they sell is automation for structured environments — a robot arm that repeats a known motion inside a fixture, a vision system that inspects a part in known lighting. That technology works precisely because the environment has been engineered so that prediction is unnecessary. It is the opposite of the capability this note is about.

An investor who buys the sector ETF on the strength of a physical-AI thesis is therefore buying industrial capital-expenditure cyclicalities, Japanese and Swiss equity exposure, and one large semiconductor position. Those may be perfectly good things to own. They are not the thesis, and the label suggests otherwise.

There is a second-order effect worth naming. Because the label attracts flows the constituents did not earn, thematic demand can bid up industrial franchises on a story their businesses do not express — which is a risk to the holder of the fund rather than a criticism of the companies in it.

The structural version of the problem is the one Goldman describes from the financing side. Digital infrastructure is investable through familiar instruments: investment-grade bonds, private credit, infrastructure debt, real estate trusts. Industrial AI software is accessible through public exposures and selective M&A. **Physical AI relies primarily on venture and growth equity**, carrying concentration and mismatch risk that institutions must weigh against the upside of underwriting the category early.⁴

EXHIBIT 10

Three risk shapes, one thesis

The same view expressed at three points on the stack produces three entirely different risk profiles. Only the third is a direct expression of it.

LAYER	HOW IT IS ACCESSED	WHAT YOU ARE ACTUALLY UNDERWRITING	RELATIONSHIP TO THIS THESIS
Enabling infrastructure	Listed semiconductors, power, grid, data centres; IG bonds and private credit	Aggregate compute and electricity demand growth	Indirect but real. Benefits whether the winning architecture is linguistic or predictive — and carries the unpriced second engine of \$03
Conventional automation	Listed industrials and the robotics ETF complex	Industrial capex cycles, Japanese and European industrial equity, labour substitution in structured environments	Weak. Works today, does not require world models, and does not benefit disproportionately if they arrive
Embodied general-purpose systems	Venture and growth equity only; no listed pure play, no credit wrapper	Whether the architecture reaches deployment reliability, on an unknown timetable	Direct. The only expression of the argument in this note — and the least accessible

Source: AI Lab Monaco analysis, incorporating the financing-divide framework in Goldman Sachs, *Harnessing AI for the Real Economy* (2026). Descriptive of instrument characteristics; not investment advice.

What would close the gap

The vehicle problem is not permanent. Three developments would each, in different ways, make the thesis ownable, and all three are visible in early form.

Listings. Unitree cleared a STAR Market IPO review in June 2026, and the 2025–26 window has already absorbed defence and space listings — Firefly, AEVEX, Hawkeye, York Space Systems, SpaceX — that would have been considered unfinanceable five years ago.⁴ A public physical-AI cohort is a question of when.

Structural credit. Goldman describes Robot-as-a-Service models extending the subscription template that enabled SaaS into physical AI procurement, shifting robotics from capital expenditure to

operating expenditure.⁴ That accounting change is what makes a fleet financeable, and it is the same move that made data centres financeable.

Cross-layer consolidation. Mobileye's \$900m acquisition of Mente Robotics in January 2026 is the pattern to watch: listed platforms buying private capability, which is how public investors get exposure before the pure plays list. Industrial acquirers have already spent more than \$110bn on software companies since 2020.⁴

Until at least one of those matures, an allocator's realistic options are the enabling infrastructure layer, a private allocation, or patience. We would rather say that plainly than manufacture a listed basket that does not honestly express the view.

SECTION 05 · WHAT WE TAKE FORWARD

The main listed robotics vehicle is roughly **46%** legacy factory automation and machine vision, with **zero** exposure to the physical-AI cohort — because that cohort is entirely private. Buying the label buys industrial capex cyclicality.

The gap closes through listings, Robot-as-a-Service credit structures, or listed platforms acquiring private capability. All three are visible in early form; none is mature. Until then the direct expression of this thesis is available to private capital and to very little else.

Being right about the direction and wrong about the instrument is the characteristic way a structural thesis loses money. This one is unusually exposed to that failure.

Unit forecasts, price compression, and where the disagreement sits

The published forecasts for this category differ by roughly a decade on timing and by an order of magnitude on units. That spread is the honest state of the analysis, and presenting a single number would misrepresent it.

EXHIBIT 11

The humanoid cost and unit curve, as forecast by two houses

Morgan Stanley and Goldman Sachs Global Investment Research. The disagreement is instructive and is not reconciled here.

	MORGAN STANLEY May 2025	GOLDMAN SACHS GIR October 2025 and 2026 updates
Headline size	~\$5tn market by 2050, including supply chain, repair and support	Units rather than market value: 20,000 in 2025 rising to 1.4 million by 2035
Unit forecast	Nearly 1 billion units by 2050 — ~930m industrial and commercial, ~80m household	1.4m by 2035; no 2050 figure published
Geography	China 302m units, United States 78m by 2050	Chinese developers advancing integrated stacks at materially lower cost
Unit cost path	~\$200k (2024) → ~\$150k (2028) → \$50k by 2050 in high-income markets; ~\$15k via Chinese supply chains	Cost falls on scale and full-stack control; no explicit price path published
Timing	Adoption relatively slow until the mid-2030s, accelerating in the late 2030s and 2040s	Broad commercial deployment not expected before 2027–2029
The spread	Roughly a decade of disagreement on when adoption inflects, from two houses looking at the same hardware. Treat any single-point forecast in this category with corresponding scepticism	

Source: Morgan Stanley, *Humanoid Robot Market Expected to Reach \$5 Trillion by 2050* (14 May 2025); Goldman Sachs Global Investment Research, *Humanoids and Profit Implications for Autos, Part II* (October 2025), and *Harnessing AI for the Real Economy* (2026).

The variable both forecasts turn on is Chinese full-stack cost

Strip away the modelling and the two houses disagree mainly about how fast one thing happens: cost compression through vertical integration, concentrated in a single country.

A cohort of Chinese developers — Unitree, UBTECH, Galbot, AgiBot, Spirit AI, XPENG and LimX — is advancing embodied platforms at materially lower cost, and Goldman's survey of fourteen of them found single-purpose systems moving toward integrated stacks that pair vision-language-action models with world models.⁹ Unitree cleared a STAR Market IPO review in June 2026.⁴

Morgan Stanley's own long-run path encodes the same dependency from the other end: its \$15,000 unit price for lower-income markets in 2050 is explicitly a Chinese-supply-chain figure, against \$50,000 in high-income markets.⁵

This makes the cost curve a geopolitical variable as much as an engineering one. Tariffs, export controls and procurement restrictions would not change what the technology can do; they would change what it costs and who can buy it — and, given that both forecasts assume that compression, they would move both timelines together.

Reading the cost curve honestly

Headline unit prices in this category are quoted loosely and compared carelessly, and the confusion runs in one direction — toward optimism. Three distinctions do most of the work.

Entry-level Chinese platforms are not industrial humanoids. The widely-cited five-figure price points come from lightweight research and developer units. Morgan Stanley's ~\$200,000 figure for 2024 refers to a deployable industrial platform in a high-income market. Both numbers can be accurate and they describe different products.

Hardware cost is not deployment cost. Goldman's assessment is that workflow integration lags both capital and hardware: embedding humanoids into

existing operations, training the workforce and reaching institutional unit economics all present real challenges.⁴ A cheaper robot that still needs a systems integrator, a safety case and a retrained workforce has not moved total cost of ownership as far as the sticker suggests.

The falling curve is real regardless. None of the above argues against the direction. Chinese full-stack producers are compressing cost genuinely, and Morgan Stanley's own path has prices roughly quartering by 2050 in high-income markets and falling by an order of magnitude through Chinese supply chains. The caution is about reading today's cheapest number as today's deployable cost.

Where we come out on timing

We think the Goldman window — broad commercial deployment in **2027–2029** — is the right frame for industrial and logistics settings with structured environments and identifiable labour shortfalls, and that Morgan Stanley's slower path is the better guide for general-purpose and household deployment.

The two are not really in conflict once separated that way. Structured industrial work needs less of the

capability this note is about; unstructured general-purpose work needs all of it. The former can deploy on today's architecture, the latter cannot.

Which returns the argument to where §04 left it. The near-term revenue is in conventional automation and does not need world models. The large prize needs them, and it arrives on the architecture's timetable rather than on the hardware's.

SECTION 06 · WHAT WE TAKE FORWARD

- Published forecasts differ by roughly a decade on timing and an order of magnitude on units. Any single-point number in this category is a choice of house, not a fact.
- Quoted unit prices mix developer platforms with deployable industrial systems, and hardware cost with total cost of ownership.
- We separate the timelines: 2027–2029 for structured industrial deployment, materially later for general-purpose systems — because the second needs an architecture the first does not.

Five ways to be right about the science and lose money anyway

01 The biology is suggestive, not transferable

Evolution optimises under constraints engineering does not share, and a dragonfly's three-synapse circuit is exquisitely specialised for one task in one body. Nothing guarantees the architecture generalises to a machine that must do many things. **What would show it:** world-model-based policies continuing to underperform conventional vision-language-action baselines on held-out objects and environments.

02 Structured automation captures the value instead

If it proves cheaper to engineer the environment than to build a machine that predicts it, the economics favour Fanuc and Keyence rather than the cohort in Exhibit 8. Most industrial value has gone that way for fifty years. **What would show it:** warehouse and factory operators continuing to hit their automation targets with fixtured systems, with humanoid pilots not converting to volume orders.

03 The private cohort is priced for an outcome it does not reach

\$12bn at \$41bn pre-revenue, \$61bn for a defence platform, \$3.5bn for a company with no product — these are valuations that require the category to work broadly and soon. Late-stage private capital is absorbing venture-stage technology risk, and there is no credit wrapper behind it.⁴ **What would show it:** down rounds or stalled raises across the cohort, or strategic investors — Nvidia, Samsung, Toyota — stepping back.

04 The energy and labour constraints bind before the technology does

Grid interconnection queues in key markets run eight to twelve years, and the US power value chain needs 500,000 additional workers by 2030 in roles that take three to four years to train for.⁴ Physical AI is a new load on a system that cannot serve its existing one. **What would show it:** deployment timelines slipping for siting and power reasons rather than capability reasons — the least discussed and most likely delay.

05 The trade is right and un-ownable for long enough to stop mattering

This is the risk we weight most heavily. A thesis available only through private capital, with no listed pure play and deployment economics five years out, can be entirely correct and still produce nothing for a public-market allocator for a decade. **What would show it:** the absence of the developments in §05 — no listings, no Robot-as-a-Service credit structures, no cross-layer acquisitions — through 2027.

OUR CONVICTION, STATED PLAINLY

We hold the *architectural* claim with reasonable confidence and the *direction* of the value shift with moderate confidence. We hold the *timing* with low confidence and the *instrument* question with none at all — there is currently no good public expression of this view, and we would rather say so than construct one. Sections 01–03 are the durable part of this note; 04–06 are a snapshot.

EXHIBIT 12

What has to be true, and by when

The thesis decomposed into testable conditions. Read the right-hand column as the falsification test for each.

CONDITION	BY WHEN	CURRENT STATE	WHAT WOULD FALSIFY IT
The architecture reaches deployment reliability	2027–2029	Best generalist manipulation 50–83% on precision tasks; threshold ~99%	No movement in generalist success rates on held-out objects through 2027
Action-conditioning stays data-thin	ongoing	62 hours of robot data sufficed in the V-JEPA 2 disclosure	Competitors reporting thousands of hours per new domain
Labour scarcity persists as the demand driver	through 2030	>1m unfilled US materials-handling roles; 500k power-sector shortfall by 2030	A labour-market loosening that removes the urgency premium
Unit economics become underwritable	2027–2029	Forming in logistics and warehousing; absent in general-purpose	Pilots not converting to volume orders at published price points
A public instrument emerges	2027+	Unitree cleared STAR Market review June 2026; no Western pure play listed	No listings and no Robot-as-a-Service credit structures through 2027

Source: AI Lab Monaco. Reliability figures from our December 2025 note, drawing on Epoch AI (2026); remaining rows from Goldman Sachs (2026) as cited throughout. “By when” dates are our estimates, not company or analyst guidance.

The asymmetry we think is mispriced

Stepping back from the individual risks, the shape of the opportunity is unusual and worth naming.

The downside case is well understood and well telegraphed: humanoids disappoint, the private cohort re-rates down, and structured automation keeps taking the industrial value. Anyone underwriting this category has heard that case.

The upside case has a feature that is much less discussed. If world models scale alongside language models rather than replacing them, then the compute and power buildout has a second demand

engine that none of the current sizing includes — and every deployed autonomous system becomes a continuous generator of data that must itself be processed.⁴

That means the enabling-infrastructure layer of Exhibit 10 may be the best risk-adjusted expression of this view available in public markets today: it participates in the upside without requiring any particular embodied company to succeed. It is an unsatisfying answer for anyone who wants to own the thesis directly. We think it is the accurate one.

What we are watching

The dragonfly is a rhetorical device and the argument does not depend on it. What it supplies is a hard existence proof: intercepting an evading target in three dimensions, at 95% reliability, is achievable with three synapses of processing depth. Whatever makes that possible is not scale, and it is not language.

Our December note reached the same conclusion from machine benchmarks. This one reaches it from comparative neuroscience and then asks what it is worth. The answer is that the capability unlocks the 99.5% of the economy that AI has not touched, that capital has begun to move accordingly, and that there is presently almost no honest way for a public-market investor to own it. Those three facts together are the note.

EXHIBIT 13

Watch list to end-2028

Dated or datable signals, ranked by how much each would move our view.

SIGNAL	HORIZON	WHY IT MATTERS
A Western physical-AI pure play lists	2027–28	The single development that would convert this from a private thesis into an ownable one. Unitree's STAR Market clearance is the precedent to watch
Generalist manipulation clears 95% on held-out objects	2027	The capability gate. Everything commercial in this note is downstream of it
A Robot-as-a-Service credit structure at scale	2027	Shifts robotics from capex to opex and makes fleets financeable — the move that made data centres investable
Listed platforms acquiring private robotics capability	ongoing	How public investors get exposure before the pure plays list. Mobileye/Mentee is the template
Capex forecasts revised to include world models	2026–28	Current sizing assumes transformer language models define compute demand. A revision would price the second engine
Humanoid shipments against the ~20,000 base	annual	The simplest reality check on both houses' forecasts, and on the 2027–2029 window
Grid interconnection and skilled-labour throughput	through 2030	The least discussed constraint. Power and people, not capability, may set the actual pace

Source: AI Lab Monaco. Horizons are our estimates, not company guidance.

Evolution never optimised for conversation. It optimised for getting to where the prey will be — and that is the capability the physical economy is waiting on.

REFERENCES

Sources

Every quantitative claim in this note is traceable to one of the sources below. Market and capital figures are as reported by the cited source and have not been independently verified. Several sources postdate this note's cover date; see the publication history on the following page.

1. Fabbri, L., *From Words to World Models: why babies beat ChatGPT at physics, and what it means for the next capability cycle*, Artificial Intelligence Laboratory Monaco, December 2025. All benchmark, model and world-model financing figures attributed to the December note are sourced there.
2. Mischiati, M., Lin, H.-T., Herold, P., Imler, E., Olberg, R. & Leonardo, A., "Internal models direct dragonfly interception steering", *Nature* 517, 2015, pp. 333–338. doi:10.1038/nature14045.
3. Bolton, A. D., Haesemeyer, M., Jordi, J., Schaehtle, U., Saad, F. A., Mansinghka, V. K., Tenenbaum, J. B. & Engert, F., "Elements of a stochastic 3D prediction engine in larval zebrafish prey capture", *eLife*, 2019. doi:10.7554/eLife.51975.
4. Goldman Sachs Investment Banking, *Harnessing AI for the Real Economy*, 2026. Source for the GDP composition, the software de-rating, the Figure/BMW and Siemens Amberg deployments, the Prometheus, Anduril, PhysicsX, Galbot and Mobileye transactions, the labour and grid constraints, the five factors framework, and the financing-divide analysis.
5. Morgan Stanley, "Humanoid Robot Market Expected to Reach \$5 Trillion by 2050", 14 May 2025.
6. Global X Robotics & Artificial Intelligence ETF (BOTZ) holdings, via stockanalysis.com, as of 14 August 2026.
7. Sandia National Laboratories, "What do dragonflies teach us about missile defense?", 24 July 2019 — source for the 95% capture rate, the ~50 ms reaction time and the three-neuron processing depth. See also Chance, F., "Interception from a Dragonfly Neural Network Model", *International Conference on Neuromorphic Systems*, 2020.
8. Goldman Sachs Global Investment Research, *Humanoids and Profit Implications for Autos, Part II*, October 2025 — source for the 20,000-to-1.4-million unit trajectory.
9. Goldman Sachs Global Investment Research, *China Humanoid Robot — Mid-Year Check-In*, 26 May 2026 — survey of fourteen Chinese robotics developers moving toward integrated stacks pairing vision-language-action models with world models.
10. Goldman Sachs Global Institute, "Tracking Trillions: The Assumptions Shaping the Scale of the AI Build-Out", April 2026 — source for the ~\$7.6tn 2026–2031 capital expenditure baseline and its stated assumptions.
11. Garrido, Q., Ballas, N., Assran, M., Bardes, A., Najman, L., Rabbat, M., Dupoux, E. & LeCun, Y., "Intuitive physics understanding emerges from self-supervised pretraining on natural videos", arXiv:2502.11831, February 2025.
12. Meta AI, *Introducing the V-JEPA 2 world model and new benchmarks for physical reasoning*, June 2025 — source for the 62-hour action-conditioning figure and the 65–80% zero-shot manipulation result.
13. Epoch AI, *Where Autonomy Works: Evaluating Robot Capabilities in 2026*, 2026 — source for the reliability threshold and generalist manipulation success rates cited from the December note.
14. Bordes, F., Garrido, Q., Kao, J. T., Williams, A., Rabbat, M. & Dupoux, E., *IntPhys 2: Benchmarking Intuitive Physics Understanding In Complex Synthetic Environments*, arXiv:2506.09849, June 2025.
15. Stahl, A. E. & Feigenson, L., "Observing the unexpected enhances infants' learning and exploration", *Science* 348(6230), 2015, pp. 91–94.
16. Fabbri, L., "The Next AI Trade? Check Dragonflies", Artificial Intelligence Laboratory Monaco, 17 March 2026 — the article this note develops.

METHODOLOGY, SCOPE AND IMPORTANT DISCLOSURES

How this note was made

PUBLICATION HISTORY

First published **17 March 2026** as “The Next AI Trade? Check Dragonflies”. This research edition was **updated in August 2026**; sources dated after March 2026 — including the Goldman Sachs real-economy report, several financing transactions and the ETF holdings in Exhibit 9 — were added in that update and are cited with their own dates. The cover carries the original publication date. Companion to *From Words to World Models*, December 2025.

SCOPE AND METHOD

All quantitative claims are drawn from primary sources — peer-reviewed papers, laboratory publications, bank research, and named market data — and each is cited. Where a figure is an order-of-magnitude estimate rather than a measured value, the exhibit footnote says so; neuron counts for the dragonfly and the human infant are of this kind. Where two reputable houses disagree materially, as on humanoid adoption timing, we present the spread rather than selecting between them. We have not independently verified company-reported funding figures, valuations, deployment statistics or fund holdings.

WHAT THIS NOTE IS NOT

This is a research note on the state of a technical field and its commercial context. **It is not investment research and it is not a recommendation to buy, sell or hold any security, fund or private interest.** AI Lab Monaco is not a registered investment adviser and nothing here should be read as investment, legal, accounting or tax advice. Named companies and funds — including the exchange-traded fund examined in Exhibit 9 — are discussed as illustrations of how the category is currently constituted, not as recommendations for or against them. The use of the word “trade” in this note's title reflects the framing of the underlying article and should not be read as a call to action. Any reader considering a financial decision in this category should take independent professional advice.

FORWARD-LOOKING STATEMENTS

Sections 04, 06 and 08 contain estimates, expectations and dated projections. These reflect our judgment, rest on assumptions we have tried to make explicit, and are subject to change without notice. Section 07 sets out the conditions under which we would consider the central argument falsified. Views expressed are those of the author and are as of the August 2026 revision.

ABOUT AI LAB MONACO

The Artificial Intelligence Laboratory Monaco is an independent research and advisory organisation based in Monaco. It publishes research on frontier artificial intelligence, advises decision-makers, and convenes the Artificial Intelligence Monaco conference. Its editorial principle is stated simply: we read the research, you get the insights.

HOW TO CITE THIS NOTE

Fabbri, L. (2026). *The Next AI Trade: what dragonflies, zebrafish and infants tell us about where the value in AI moves next*. Frontier Research. Artificial Intelligence Laboratory Monaco, March 2026; research edition revised August 2026.

CORRECTIONS

We correct errors of fact openly and date every revision. If you believe a figure, citation or characterisation in this note is wrong, please write to us — particularly if you are an author of work cited here.



Artificial Intelligence Laboratory Monaco

Frontier Research · March 2026
artificialintelligencemonaco.com

We read the research, **you get the insights.**

Also in this series — From Words to World Models: why babies beat ChatGPT at physics, December 2025.



Frontier Research • March 2026

artificialintelligencemonaco.com

© 2026 Artificial Intelligence Laboratory Monaco