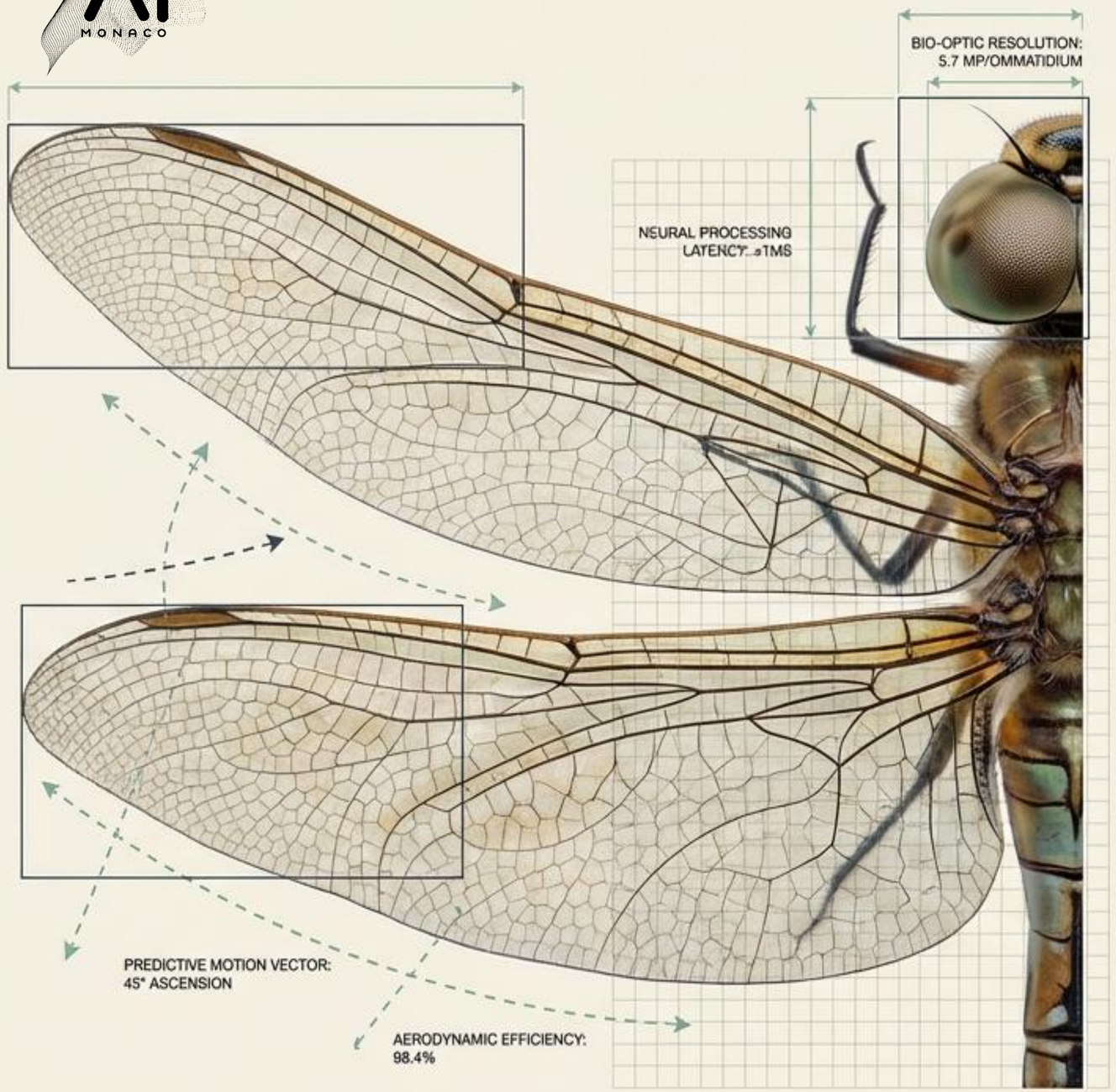
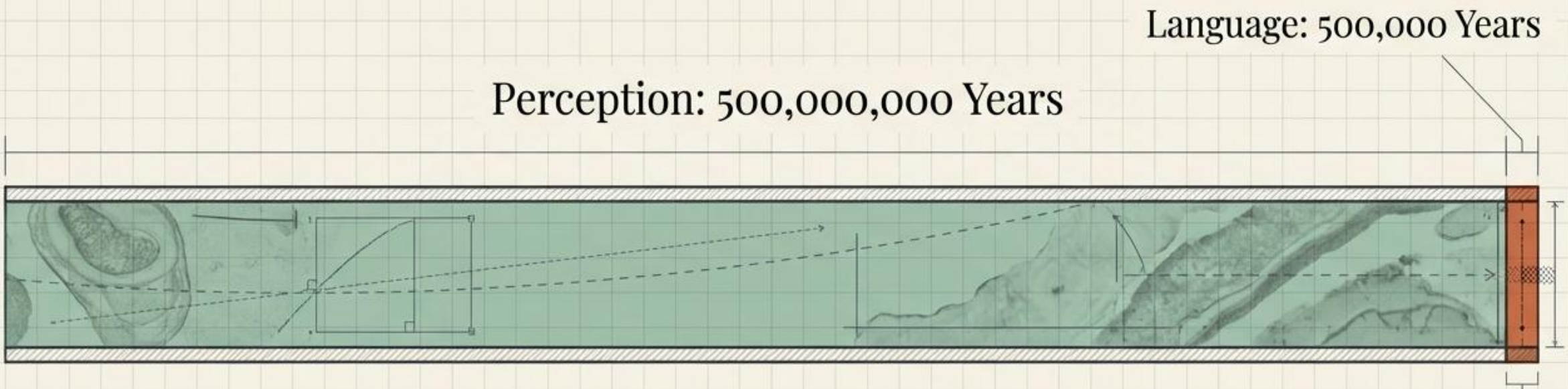
The Deepest Roots of Intelligence Are Physical

Why the next trillion-dollar AI trade relies on predictive world-modeling, not language.



Prepared by AI Monaco | Market Intelligence & Asset Allocation

The Time-Scale of Intelligence

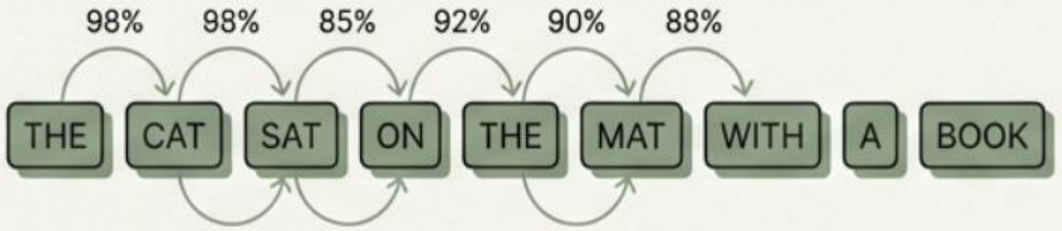


"Language is a half-million-year-old luxury.
Perception is a half-billion-year-old necessity."

— Fei-Fei Li, Professor of Computer Science, Stanford University

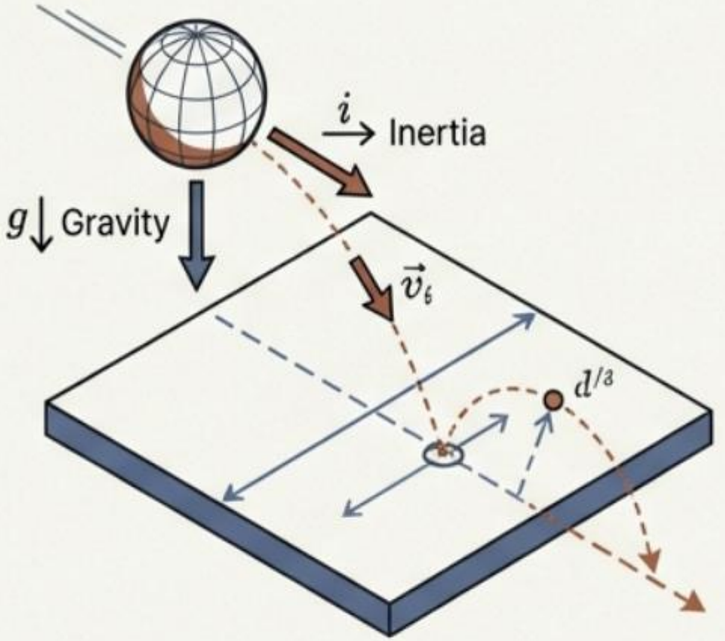
Biological systems evolved to survive in three-dimensional space—anticipating trajectories, forces, and delays long before they could speak.

Artificial intelligence is currently bifurcated



Mastery of Linguistic Symbols

- Large Language Models (LLMs) excel at statistical token prediction.
- Fundamentally detached from physical causality.

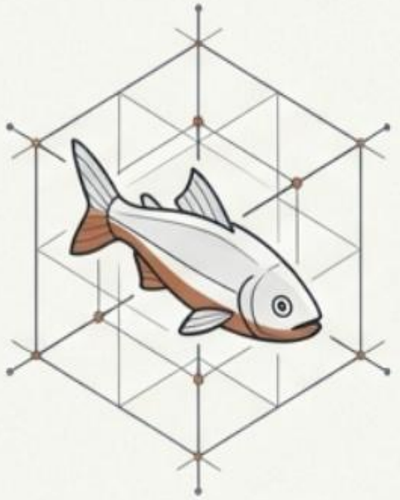


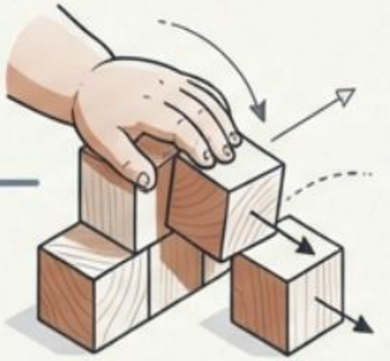
Mastery of Physical States

- Artificial General Intelligence (AGI) requires robust internal simulations of the material world.
- Predicting the 'next state' of a dynamic environment.

Intelligence emerged from navigating a three–dimensional medium

Evolution prioritized the ability to model environmental dynamics—gravity, inertia, and causality—as a prerequisite for survival.



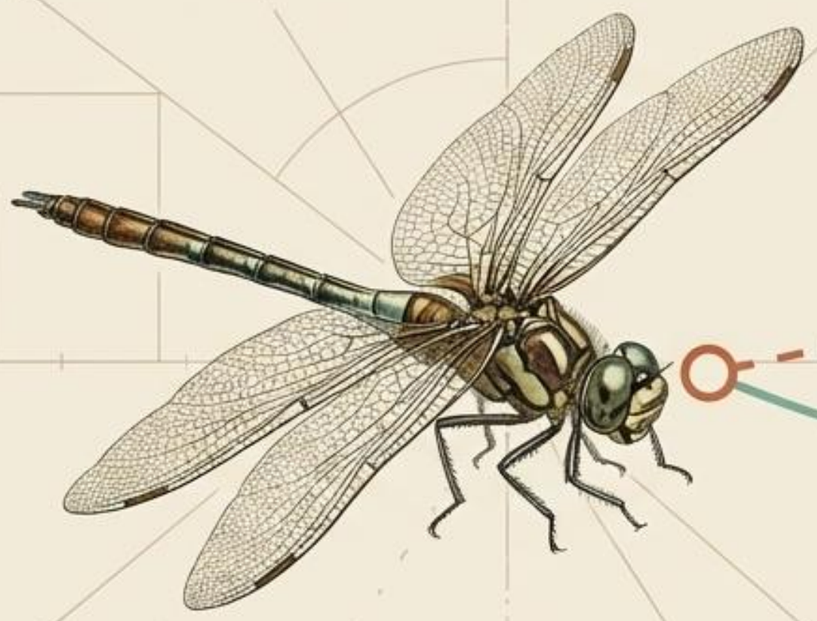


The World Model

- Internal simulation is a universal biological solution to environmental uncertainty.
- Allows agents to predict consequences of actions and trajectories of external agents.

Predictive Control in Apex Predators

95% Hunt Success Rate



Reactive Path: Neural delay guarantees failure.

Predictive Path: Estimating future states.

The Neural Delay: Visual stimulus to motor response creates a lag. Reaction alone is insufficient for survival.

The Solution: Intelligence emerges through estimating future states and updating movement.

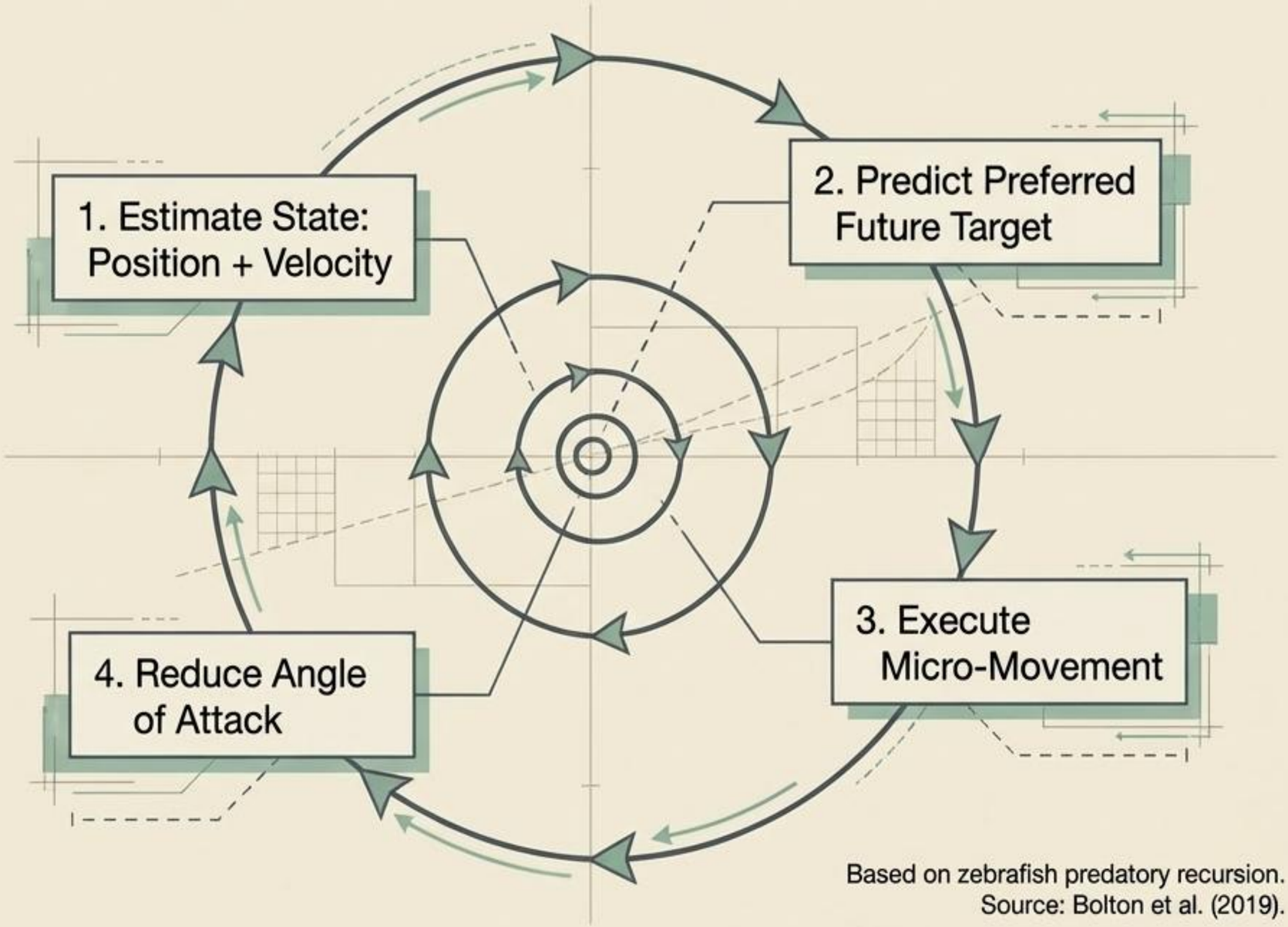
Source: Mischiati et al., *Nature* (2015).

The Architecture of Continuous Adaptation

Core Insight

Intelligence requires systems that track evolving states and adjust behavior based on anticipated futures.

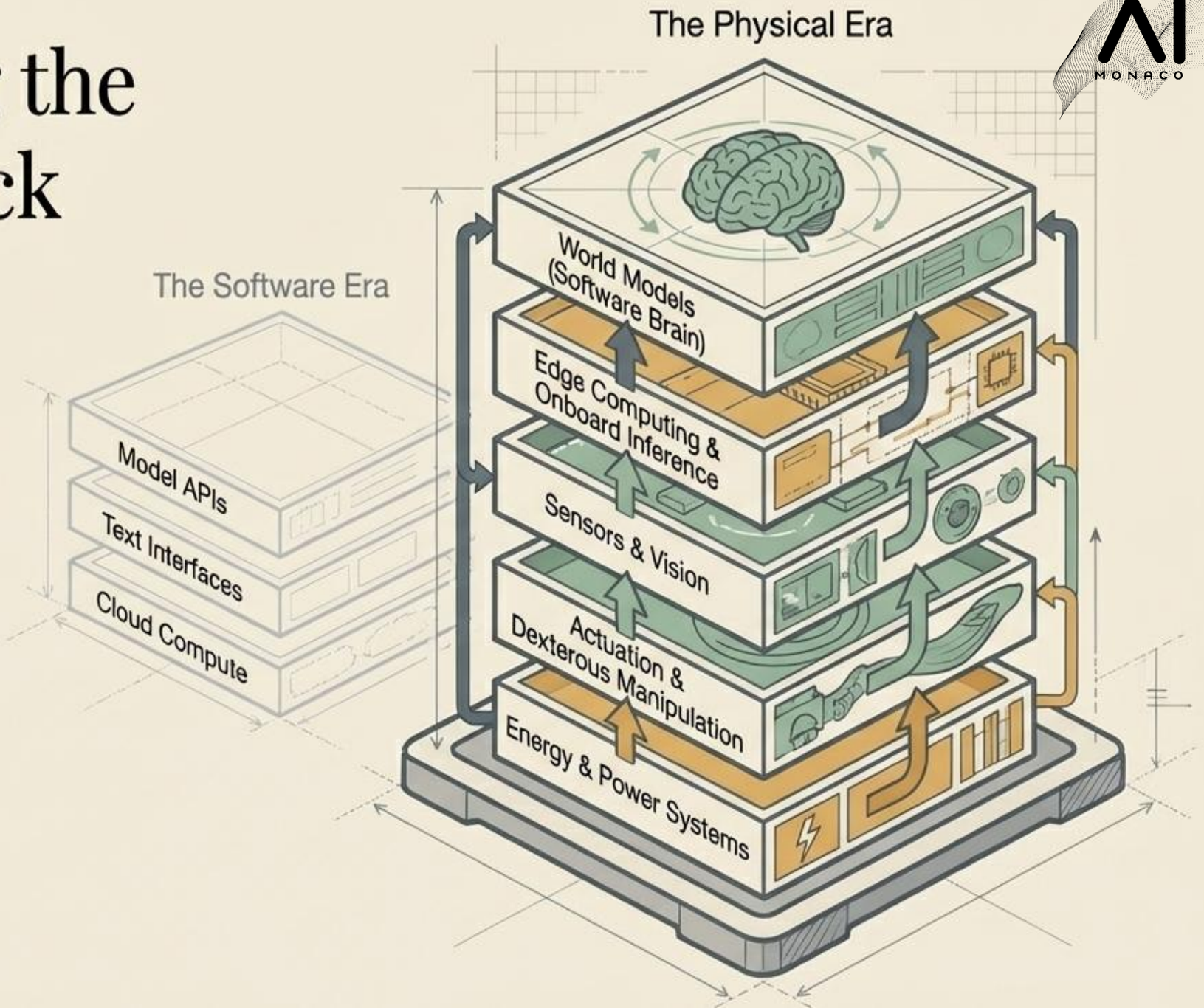
Prediction is not an accessory; it is central.



Based on zebrafish predatory recursion.
Source: Bolton et al. (2019).

AI MONACO

Deconstructing the Physical AI Stack



As software models commoditize, durable economic moats will belong to systems that successfully integrate software with physical sensing, planning, and real-world execution.

The radical transformation of humanoid unit economics

Battery Runtime

1.5 – 2 Hours

→ **8 – 12 Hours (Full Shift)**

Inference Latency

100+ ms

→ **18 – 20 ms**

Unit Cost (Est.)

$250k – $1M

→ **$15k – $30k**

Hand Dexterity

5 – 11 DoF

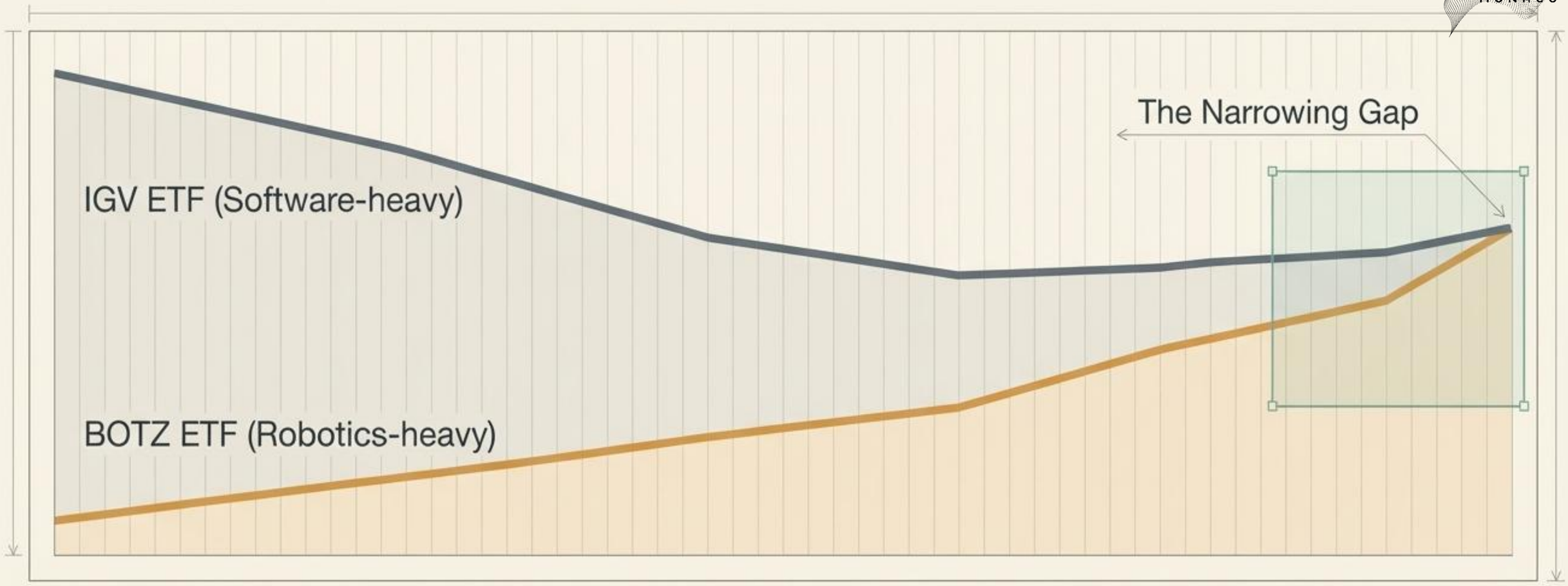
→ **22 DoF (Gen 3)**

Torque Density

Moderate →

High (Solid-State Gears)

The Repricing of AI's Next Frontier



The recent convergence between IGV and BOTZ is not a short-term anomaly. It is an early market signal of a deeper rotation toward the physical AI ecosystem.

The Multi-Trillion Dollar Rotation

Robotics Manufacturers

Sensor Providers

$5 Trillion

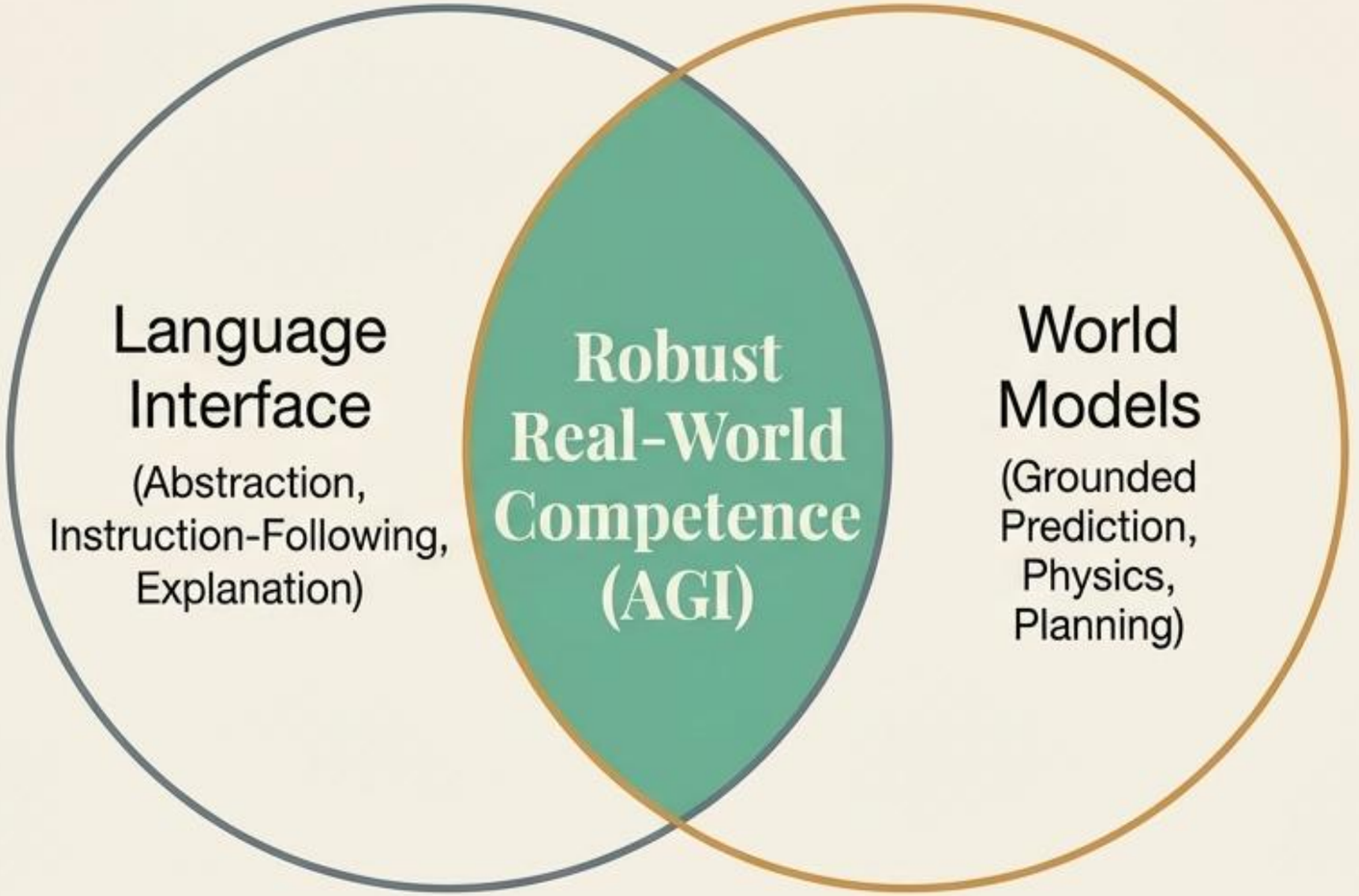
Morgan Stanley Research estimates the embodied AI market opportunity will reach ~US$5 Trillion by 2050.

Semiconductor Suppliers

Edge Inference Enablers

Value is migrating beyond pure software toward the critical infrastructure required for real-world intelligence.

The Ultimate Synthesis



"The most capable systems will look less like a 'bigger language model' and more like a fusion: world models for grounded prediction, with language as the interface layer."

Before intelligence becomes verbal, it is physical. True AI requires both.