From "Jalapeño" to "Morpho-Root": The Apex of OpenAI's In-House Chip and the Hardware Involution of Tokenism

——On the Path of Logographic AI's "Morpho-Root Native Chip"

Abstract

In June 2026, OpenAI and Broadcom jointly released Jalapeño, their first in-house AI inference chip, which went from design to tape-out in just nine months with an estimated 50% reduction in inference cost[1][2][3]. This paper is an independent technical analysis paper based on the theoretical framework of The Rooted Intelligence: The Paradigm Revolution of Logographic AI[8]. From the perspective of Logographic AI (LAI)—a rooted intelligence paradigm that replaces "Token" with "Morpho-Root" as the cognitive primitive—this paper provides a paradigm-level critical interpretation of this industrial event and proposes an alternative hardware vision. On one hand, Jalapeño represents a "self-rescue" and "reinforcement" of the Tokenist paradigm at the hardware level—it improves the efficiency of statistical fitting through specialized design, but fails to address the fundamental problem of "rootless cognitive primitives." On the other hand, it validates the feasibility of "customizing chips for specific cognitive architectures," clearing the path for Logographic AI's "Morpho-Root Native Chip" vision.

This paper first elaborates the fundamental conceptual frameworks of Logographic AI and Tokenism, then analyzes the nature of Jalapeño and compares it with GPU and TPU in a technological genealogy. Subsequently, integrating the architectural concepts of Morpho-Root triples, Value Axiom Library, Morpho-Graph Computing, and Glypho-Phonetic Co-processor, it proposes a complete architectural vision for the "Morpho-Root Native Chip," demonstrating that the transition from "statistics-friendly hardware" to "structure-affine hardware" is the necessary path for AI to move toward "Rooted Intelligence."

Keywords: Jalapeño; Logographic AI; Morpho-Root; Tokenism; Morpho-Graph Computing; Endogenous Security; Glypho-Phonetic Co-processor

---

I. Introduction: A "Jalapeño" That Triggers Paradigm Reflection

On June 24, 2026, OpenAI and Broadcom jointly released Jalapeño, their first in-house AI inference chip. This chip, codenamed after the Mexican chili pepper, went from concept to tape-out in just nine months, setting a record for high-performance ASIC development cycle[1][2][3]. More importantly, OpenAI's own AI models were deeply involved in the chip's design and optimization—AI is now participating in designing the hardware that runs AI itself, forming a recursive "AI-designs-AI" closed loop[2][3].

This event sent shockwaves through the AI industry. Broadcom CEO Hock Tan revealed that Jalapeño demonstrates approximately 50% cost savings compared to typical AI GPUs[1][4]. OpenAI's hardware lead Richard Ho stated that engineering samples are already running machine

learning workloads—including GPT-5.3-Codex-Spark—in the lab at production target frequencies and power consumption[1].

Mainstream industry interpretation tends to view Jalapeño as a "milestone in the AI hardware revolution": in-house chips will break NVIDIA's monopoly, significantly reduce inference costs, and accelerate large-scale AI deployment[2][3][6]. However, this paper seeks to ask a deeper question: **Is Jalapeño the "lifeline" of the Tokenist paradigm, or the apex of Tokenism's hardware involution?** This paper's answer: **both.** It is the peak of the old paradigm—the best that Tokenism can achieve at the physical layer—yet it also validates the feasibility of custom hardware, clearing the path for the hardware era of "Rooted Intelligence."

This inquiry's theoretical framework comes from the Logographic AI theory systematically elaborated in The Rooted Intelligence: The Paradigm Revolution of Logographic AI[8]—a theory that replaces "Token" with "Morpho-Root" as the cognitive primitive, launching a paradigm revolution in artificial intelligence from the very origin of cognition.

II. Theoretical Framework: Logographic AI, Morpho-Root, and Tokenism

2.1 Tokenism: The "Rootless" Paradigm of Mainstream AI

The dominant large language models of today—GPT, Claude, and their peers—are built upon a cognitive paradigm that can be termed "Tokenism"[8][9][10]. A Token is a discrete symbolic unit produced by slicing text, code, images—all inputs—into fragments. A Token is merely an integer index whose "semantics" are entirely and temporarily conferred by its statistical co-occurrence relationships with other Tokens in vast datasets.

The fundamental characteristic of Tokenism is that **meaning is entirely externalized from the symbol itself.** A symbol's meaning is not intrinsic to it, but is determined by its position within a statistical distribution. When the training data distribution shifts, the "semantics" of Tokens drift accordingly. This "rootlessness" gives rise to a series of structural deficiencies: hallucinations cannot be fundamentally eliminated, safety guardrails can be bypassed, reasoning cannot be truly traced, and value alignment remains perpetually fragile[8][9].

This paper adopts **"Rootless Semioticism"** as the theoretical framework for analyzing the Tokenist paradigm[10]. The core judgment of this diagnosis is that a symbol's meaning is entirely defined by its statistical co-occurrence relationships within the symbolic system, and cannot reach the "referent" beyond the symbol—namely, the real world, human experience, and civilizational values. The symbol thus becomes a "floating signifier," sliding endlessly within a closed statistical loop[8].

2.2 Logographic AI and Morpho-Root: Rooted Cognitive Primitives

In contrast to Tokenism, "Logographic AI" (LAI) theory proposes a fundamentally different cognitive paradigm[8][10]. Its core concept is the "Morpho-Root"—a new type of cognitive primitive that replaces the Token[8].

Each Morpho-Root is a **structured triple**[8]:

r = ⟨ S, A, R⟩

Where:

- **S (Symbol)**: The symbolic identifier, the externally addressable name of the Morpho-Root.
- **A (Attributes)**: The attribute set, embedding the Morpho-Root's intrinsic semantic features and value constraints (e.g., [+liquid], [+trust], [+inviolable]).
- **R (Relation Functions)**: The relation function set, defining the presupposed logical connections between this Morpho-Root and others (e.g., "composes(人, 言)", "entails(信, 人∧言)").

The key innovation of the Morpho-Root is that **meaning is not "emerged" from statistics, but is pre-configured as an intrinsic attribute of the cognitive primitive**[8]. When the character "信" (trust) is created, "non-deception" is already its constitutive feature. This enables Logographic AI's reasoning to be transparent and traceable, and its security to be endogenous and unbypassable[8][9].

2.3 From Transformer to Morpho-Graph Computing: A Fundamental Shift in Computational Paradigm

At the algorithmic level, the Tokenist paradigm is deeply bound to the Transformer architecture. The core of the Transformer is the self-attention mechanism—each Token computes attention weights with every other Token in the sequence, with a complexity of $O(n^2)$. The essence of this mechanism is **dense computation of statistical associations**: the model does not care whether there is a meaningful relationship between Tokens; it merely performs a weighted average over all possible pairs.

Logographic AI proposes **Morpho-Graph Computing** (implemented by the "MorphoCosmos" engine) as an alternative architecture[8]. Its core design principles are[8]:

- **Computational scope**: Not all Token pairs, but **presupposed relation edges (R)** in the Morpho-Root network. Complexity reduced from $O(n^2)$ to $O(|E|)$.
- **Attention guidance**: Not statistical weights, but jointly guided by **relation types** (causality, implication, composition, etc.) and **local morpho-structural entropy**.
- **Reasoning method**: Not matrix operations, but **graph traversal**—"walking" along presupposed relation edges within the Morpho-Root network.
- **Interpretability**: Each inference generates a **fully traceable decision subgraph**, recording node activations, edge traversals, decision points, and conflict resolutions.
- **Morpho-Structural Entropy** (MSE) plays a key role in this architecture: it measures the structural determinacy of the Morpho-Root network and guides dynamic resource allocation—fast reasoning

in low-entropy regions, expanded search or human-in-the-loop requests in high-entropy regions[8]. This mechanism enables Logographic AI to "know what it does not know"[8].

2.4 Comprehensive Comparison of the Two Paradigms

Dimension	Tokenism (Phonographic AI)	Logographic AI
Cognitive Primitive	Token (meaningless statistical fragment)	Morpho-Root (structured meaning crystal)[8]
Source of Meaning	Temporarily conferred by statistical co-occurrence	Pre-configured within attribute set A[8]
Computational Architecture	Transformer (self-attention)	Morpho-Graph Computing (graph traversal)[8]
Computational Complexity	$O(n^2)$	O(E)[8]
Reasoning Method	Black-box attention weights	Transparent graph traversal[8]
Interpretability	Post-hoc attribution	Traceable paths[8]
Security Paradigm	External guardrails, bypassable[8]	Endogenous security, logically unviolable[8]
Value Alignment	External RLHF, fragile	Value axioms embedded, stable[8]
Hardware Affinity	Matrix operations (GPU)	Graph traversal (NPU/specialized chip)[8]

This comparison reveals a critical difference: **the two paradigms differ not only in algorithms, but in the very nature of their computational patterns**[8]. Tokenism is **"statistics-friendly"**—it requires massive parallel matrix multiplication; Logographic AI is **"structure-affine"**—it requires efficient graph traversal and relation matching[8]. This means that hardware optimized for Tokenism (such as GPUs and Jalapeño) and hardware optimized for Logographic AI should differ fundamentally, from instruction sets to microarchitecture[8].

III. The Nature of Jalapeño: Tokenism's Hardware "Self-Rescue"

3.1 From General-Purpose to Specialized: Efficiency Improved, Paradigm Unchanged

Jalapeño is an ASIC (Application-Specific Integrated Circuit) **designed specifically for large language model inference**[1][2]. Unlike NVIDIA's general-purpose GPUs, it is deeply optimized for the memory access patterns and attention mechanisms of the Transformer architecture[1][4]. OpenAI President and co-founder Greg Brockman stated in the announcement: "The world is transitioning to a compute-driven economy."[1]

This is the pinnacle of optimization within the Tokenist paradigm. Jalapeño solves the problem of **"roads too congested"** —how to move and compute Token fragments faster and more cheaply. It does not answer the question of **"which way the car should go"** :

- The cognitive primitive **remains the Token**—those meaningless, statistically co-occurrence-dependent symbolic fragments.
- The computational architecture **remains a Transformer variant**, its hardware optimization still serving the acceleration of the $O(n^2)$ attention mechanism[8].
- The security paradigm **remains external**—the chip does not care whether the content it processes is "fact" or "backdoor."

It is worth noting that Jalapeño is not an isolated case for OpenAI. Google (TPU), Amazon (Trainium/Inferentia), Microsoft (Maia), and Meta (MTIA) have all embarked on the path of in-house chip development, and Anthropic is also considering it[3][6]. **NVIDIA's list of top customers is becoming a list of competitors**[3]. This trend indicates that custom hardware under the Tokenist paradigm is becoming an industry consensus.

3.2 Examining the Claim that "Jalapeño Is an AI Hardware Revolution"

Following Jalapeño's release, mainstream industry interpretation has tended to view it as a landmark "AI hardware revolution"[2][3][6]. However, from a paradigm-level perspective, this judgment requires careful scrutiny:

First, efficiency improvement is not paradigm breakthrough. Jalapeño's 50% cost savings are a victory of engineering optimization, not cognitive architecture innovation. It makes Tokens run faster and cheaper, but it does not make Tokens more meaningful. As The Rooted Intelligence argues: no matter how fast you run on the Token track, you cannot escape the Token's cage[8].

Second, specialization is not intelligence. Jalapeño's specialization is specialization for the Transformer, not for "understanding." Its optimization target is "predicting the probability distribution of the next Token," not "understanding the meaning of symbols." If the very definition of "understanding" is flawed, then no matter how fast the hardware runs, it cannot produce genuine understanding.

Third, cost reduction is not security enhancement. Cheaper inference may make AI more ubiquitous, but if AI's cognitive primitives remain "rootless," what is being popularized is merely a faster engine of non-understanding. As Chapter 4 of The Rooted Intelligence elaborates, Tokenism's security predicament is structural, not a cost issue[8].

Therefore, **Jalapeño is Tokenism's "self-reinforcement" at the hardware level, not "self-transcendence."** It reinforces the physical foundation of the old paradigm, but does not open the cognitive space of a new one.

3.3 The "AI-Designs-AI" Recursive Closed Loop

Jalapeño's most striking feature is the deep involvement of AI models in its design process. According to reports, Jalapeño went from design to tape-out in just nine months, with both companies claiming this is "one of the fastest ASIC development cycles in the history of high-performance semiconductors"[2][3]. This speed was enabled by deep software-hardware co-optimization between OpenAI's engineering team and Broadcom, as well as the use of OpenAI's own AI models to accelerate portions of the design and optimization work[2][3].

This forms a recursive closed loop:

AI (based on Token statistics) → designs chip → chip runs AI → stronger AI designs stronger chip

From the perspective of Logographic AI, this is Tokenism's **"self-reinforcement"** at the hardware level—using statistical tools to patch the physical bottlenecks of a statistical system, yet never touching the fundamental question of "why intelligence needs to understand"[8][9]. Each step of the recursion optimizes "guessing more accurately," not "understanding more deeply."

IV. Chip Technological Genealogy: From GPU to Morpho-Root Chip

To more clearly position Jalapeño within chip development history and contrast it with the Morpho-Root Native Chip, this section places current mainstream AI chips within a unified technological genealogy.

4.1 Paradigm Comparison of Four Generations of Chips

Dimension	NVIDIA GPU	Google TPU	OpenAI Jalapeño	Morpho-Root Native Chip (Vision)
Design Goal	General-purpose graphics/matrix computing	Matrix multiplication acceleration	Transformer inference acceleration	Morpho-Root graph traversal acceleration
Cognitive Paradigm	None (compute tool)	Tokenism	Tokenism	Logographic AI / Morpho-Root Paradigm
Core Operation	Matrix multiplication	Matrix multiplication	Matrix multiplication +	Graph traversal + attribute matching

Dimension	NVIDIA GPU	Google TPU	OpenAI Jalapeño	Morpho-Root Native Chip (Vision)
			attention	
Data Pattern	Dense, continuous	Dense, continuous	Dense, continuous (Token sequence)	Sparse graph (Morpho-Root network)
Interpretability	None	None	None	Hardware-level transparent traceability
Security Paradigm	External	External	External	Hardware-level endogenous security
Representative Slogan	Compute is King	Built for AI	Inference cost halved	Making hardware understand meaning

4.2 Genealogical Positioning Analysis

First Generation (GPU): General-purpose computing platform, assuming no particular cognitive paradigm. Its success lies in flexibility and ecosystem (CUDA), not in understanding the nature of intelligence[6].

Second Generation (TPU): The first hardware customized for AI, yet still serving matrix multiplication—the core operation of the statistical paradigm. It is Tokenism's first-generation specialized hardware.

Third Generation (Jalapeño): Deeply customized for Transformer inference, representing the extreme form of Tokenism's hardware trajectory. It has the highest degree of specialization, and is therefore the most deeply locked into the Tokenist paradigm.

Fourth Generation (Morpho-Root Native Chip): Assumes that the cognitive paradigm itself needs to change. It does not accelerate the statistical fitting of Tokens, but rather the structural traversal of Morpho-Roots. It belongs to a fundamentally different chip family.

4.3 Jalapeño's Position in the Genealogy

This comparison table clearly shows that **Jalapeño belongs to the same family as GPU and TPU in the technological genealogy**—all are statistics-friendly hardware serving the Tokenist paradigm, with progressively deepening specialization across generations. The Morpho-Root Native Chip, in contrast, belongs to a completely different family—structure-affine hardware.

Jalapeño is the **apex** of Tokenism's hardware trajectory, not the **turning point**. It achieves the best that the Token paradigm can do at the physical layer—faster matrix multiplication, lower inference cost, shorter development cycles. But it stops there. This is the peak of the old paradigm, and the starting point of the new.

V. Logographic AI's Interpretation: "Paradigm Involution" at the Hardware Level and Historical Opportunity

5.1 Jalapeño and the Isomorphism of "Rootless Hardware"

Tokenism's "rootlessness" exists not only at the algorithmic level, but also extends to the hardware level[8]. NVIDIA GPUs are quintessential "statistics-friendly hardware"—they execute matrix operations efficiently, but never care about the "meaning" of the objects being operated on. Although Jalapeño is more efficient, its "rootless" nature remains unchanged: **it is still a high-speed porter of Tokens, just faster and cheaper.**

Some analyses point out that NVIDIA's true monopoly advantage lies in its CUDA software ecosystem, not the chip itself[6]: over 5 million registered developers, approximately 5,937 GitHub projects related to CUDA (compared to AMD's competing ROCm with only 187), and nearly all relevant AI libraries deeply bound to CUDA[6]. This means that even if Jalapeño surpasses GPUs in inference efficiency, the software ecosystem under the Tokenist paradigm still exerts a powerful lock-in effect[8].

From the Logographic AI perspective, Jalapeño's emergence marks a new stage of **"involution"** for the Tokenist paradigm at the hardware level[7][8]. It may further consolidate Tokenism's dominance in the hardware ecosystem, making the future transition to "Rooted Intelligence" even more costly in terms of architectural shift[7][8]. As The Rooted Intelligence warns: no matter how fast you run on the Token track, you cannot escape the Token's cage[8].

5.2 An Underestimated Positive Signal: The Viability of Customization

However, Jalapeño also sends a signal highly favorable to Logographic AI: **customizing chips for specific cognitive architectures is feasible and commercially viable**[1][2][3].

OpenAI proved in nine months that an AI company can go from zero to tape-out of a custom chip in under a year[2][3]. This clears the doubt about **"whether customization is feasible"** for Logographic AI's "Morpho-Root Native Chip" vision. When Logographic AI's algorithmic paradigm matures, it too can have dedicated "Morpho-Root chips" designed for it—**this is not science fiction, but a path just walked by OpenAI**[8].

As The Rooted Intelligence points out, the way out for current AI lies in moving from "escaping 'technological capture'" toward paradigm revolution, rather than pursuing ever more extreme optimization within the same paradigm[7]. Jalapeño is precisely this kind of dual-event—it is both the apex of Tokenism's involution and a proof of feasibility for the customization path[8].

VI. Morpho-Root Native Chip: The Hardware Path of Rooted Intelligence

Unlike the "Glypho-Phonetic Co-processor" proposed in Chapter 11 of The Rooted Intelligence—which requires heterogeneous computing handling both Morpho-Root graph traversal and speech sequence modeling simultaneously[8]—Jalapeño is only optimized for inference-stage Token sequences and has not yet touched the structural computation at the cognitive primitive level. The Morpho-Root Native Chip builds upon the "Glypho-Phonetic Co-processor," further mapping the cognitive primitive structure of Morpho-Graph Computing to physical topology[8].

6.1 From "Matrix Operations" to "Morpho-Root Graph Traversal"

Integrating the architectural concepts of Morpho-Graph Computing and the Glypho-Phonetic Co-processor[8], we can envision a more disruptive hardware form: the **Morpho-Root Native Chip**[8].

Its core computational unit is no longer an ALU executing matrix multiplication, but a **specialized graph traversal engine**[8]. It natively supports the Morpho-Graph Computing paradigm, directly managing "Morpho-Root" nodes and presupposed "relation functions" (R) at the hardware level. Logical reasoning is no longer simulated, but accomplished by the hardware "walking" along presupposed paths within the relation network[8].

Morpho-Structural Entropy is directly **embedded into the hardware scheduling logic** in this architecture[8]: low-entropy regions correspond to deterministic paths, enabling fast reasoning; high-entropy regions trigger expanded search or human-in-the-loop requests—this mechanism realizes the ability to "know what it does not know" at the hardware level[8].

6.2 Compute-in-Memory: From "Moving Data" to "Activating Meaning"

Traditional chips (including Jalapeño) consume most of their power moving data between storage and compute units[8]. The Morpho-Root Native Chip should adopt **compute-in-memory technology**[8], directly anchoring each Morpho-Root triple **r = ⟨ S, A, R⟩** near the storage unit[8].

When the chip receives an input, it no longer "reads" data, but **"activates" the Morpho-Root network relevant to that input**[8]. The entire inference process becomes a massive, parallel attribute matching and relation triggering operation, nearly eliminating the energy consumption of data movement[8].

The design of the Morpho-Root Native Chip essentially maps the cognitive architecture of the Morpho-Root triple **r = ⟨ S, A, R⟩** directly onto physical structure—symbol S, attributes A, and relations R are no longer stored data, but **hardwired circuit topology**[8].

6.3 Hardware-Level Value Hardening: From "External Security" to "Endogenous Security"

Jalapeño does not care about the content it processes and cannot distinguish "fact" from "backdoor"[1][2][3][8]. The Morpho-Root Native Chip, by contrast, embeds the **"Value Axiom Library"** directly into the chip as **hardwired logic**[8].

Top-level axioms such as "do no harm" and "inviolability" become **part of the chip's physical structure**. Any computational path attempting to violate value constraints will be directly blocked at the circuit level—this is the true **hardware implementation of "endogenous security"**[8].

From the perspective of Logographic AI's "endogenous security," the Tokenist hardware represented by Jalapeño still remains in the era of "external patches." As Chapter 4 of The Rooted Intelligence elaborates, when security rules are merely statistical text rather than hard constraints, external guardrails are destined to be bypassed[8]. The hardware-level value hardening achieved by the Morpho-Root Native Chip is the **fundamental response** to this predicament[8].

6.4 R&D Roadmap Vision

Drawing on the three-phase model from Chapter 12 of The Rooted Intelligence[8], the development of the Morpho-Root Native Chip could follow this roadmap:

Phase	Timeline	Goal	Key Tasks
Phase I: Algorithm Simulation	1-2 years	Validate the feasibility of Morpho-Graph Computing at the hardware abstraction layer	Implement Morpho-Root graph traversal engine prototype on FPGA; validate functional correctness on small-scale Morpho-Root library (~10^4 nodes)
Phase II: Architecture Design	2-3 years	Complete microarchitecture design of Morpho-Root Native Chip	Define instruction set (ACTIVATE_NODE, TRAVERSE_EDGE, MATCH_ATTR, etc.); design physical layout of compute-in-memory units; complete RTL-level simulation
Phase III: Tape-out Validation	3-4 years	Complete first Morpho-Root Native Chip tape-out	Select mature process node (e.g., 12nm/7nm); collaborate with foundry on physical design; prototype testing and iteration

This roadmap demonstrates that the Morpho-Root Native Chip is not a distant vision, but a path that is challenging yet clear. OpenAI proved the feasibility of custom chips in nine months[2][3]; what the Morpho-Root chip needs is not more time, but the right paradigm direction.

VII. Conclusion: From Jalapeño to the Hardware Path of Rooted Intelligence

Jalapeño represents a "self-rescue" and "reinforcement" of the Tokenist paradigm at its hardware foundation[1][2][3][8]. It trades generality for efficiency on specific tasks, but fails to address the fundamental problem of "rootless cognitive primitives"[8][9]. Its computational architecture remains

a Transformer variant, its hardware optimization still serves the statistical fitting of Tokens—just faster and cheaper[1][2][3][8].

Yet its emergence also validates a critical proposition: **only when an intelligence's algorithms (software) and hardware are co-designed around a cognitive model can optimal efficiency be achieved**[1][2][3][8]. This precisely provides the logical and commercial support for "Rooted Intelligence" to realize a full-stack, software-hardware integrated architecture in the future[7][8].

When OpenAI's AI models design their own chips, they reinforce Tokenism's closed loop[2][3][8]; when Logographic AI's Morpho-Root paradigm moves toward hardware, it will open a new era of **"giving intelligence roots"**[8][10].

Core Judgments

Dimension	Significance of Jalapeño	Implication for Logographic AI
Paradigm Positioning	Apex of Tokenism's hardware involution[8]	Proves the old paradigm has reached its physical limits
Technical Path	Extreme of statistics-friendly hardware	Necessity of structure-affine hardware is highlighted
Commercial Validation	Custom chips are commercially viable[1][2][3]	Commercial path for Morpho-Root Native Chip is clear
Security Paradigm	External security; chip does not care about content[8]	Urgency of hardware-level value hardening is proven
Historical Significance	Endpoint of the old paradigm	Starting point coordinate of the new paradigm[8]

Jalapeño is not the starting point, but the endpoint—**the endpoint of Tokenism's hardware optimization**. It achieves the best that the Token paradigm can do at the physical layer, but stops there. Yet the customization path it validates will belong to the new paradigm[8].

When Logographic AI's Morpho-Graph Computing moves from algorithms to chips, when the Morpho-Root Native Chip replaces matrix operations with graph traversal and external guardrails with hardware-level value hardening, what it opens is not a continuation of the old track, but the starting point of a new track[8][10].

No matter how spicy the Jalapeño, it cannot change the fact that the Token is a "rootless" statistical fragment. What the Morpho-Root chip carries, by contrast, is the meaning crystal of ren yan wei xin (人言为信—trust is a person's word; the character combines 'person' and 'speech'), the value axiom of zhi ge wei wu (止戈为武—true martial valor is stopping war; the character combines 'stop' and 'spear'), and the rooted growth of five thousand years of civilization in the silicon world. This "Morpho-Root" is the future that AI hardware deserves.

References

[1] OpenAI, Broadcom. (2026, June 24). OpenAI and Broadcom announce Jalapeño inference chip. OpenAI Official Blog. https://openai.com/index/openai-broadcom-jalapeno-inference-chip/

[2] Quantum Bit. (2026, June 25). OpenAI's in-house chip succeeds in 270 days, led by former Google TPU executive. 36Kr. https://36kr.com/p/3868228504933378

[3] EET China. (2026, June 25). OpenAI and Broadcom jointly release first custom AI inference chip Jalapeño, tape-out in 9 months. https://www.eet-china.com/news/202606258325.html

[4] First Financial. (2026, June 25). OpenAI releases first custom AI inference chip, discloses multi-generation computing platform plan. Securities Times. https://stcn.com/article/detail/3978450.html

[5] Avcıoğlu, M. (2026, June 24). OpenAI launches first in-house AI chip in partnership with Broadcom. Anadolu Ajansı. https://www.aa.com.tr/en/americas/openai-launches-first-in-house-ai-chip-in-partnership-with-broadcom/3976867

[6] IC Smart. (2026, June 25). OpenAI partners with Broadcom, unveils first in-house AI chip. https://www.icsmart.cn/106676/

[7] Liu, S. (2025). Escaping "technological capture": The future path of AI from architectural improvement to paradigm revolution. PSSXiv. https://doi.org/10.12451/202512.03460

[8] Liu, S. (2026). The Rooted Intelligence: The Paradigm Revolution of Logographic AI. http://logographicai.com/#book

[9] Liu, S. (2025). Nested learning, gated attention, and native sparse attention: The limit optimization of the token paradigm and the paradigm revolution of Logographic AI. PSSXiv. https://doi.org/10.12451/202512.00594

[10] Liu, S. (2025). Logographic AI vs. Phonographic AI: A new paradigm and the manifesto of AI decolonization. ChinaXiv. https://doi.org/10.12074/202503.00051