

面向极端条件下结构材料可信设计的表意 AI 方法：基于形根的图推理框架及其实现策略

Logographic AI for Trustworthy Design of Structural Materials under Extreme Conditions: A Morpho-Root Based Graph-Reasoning Framework with Implementation Strategy

摘要

面向核反应堆、航空航天、深空和深海探测等极端服役环境的结构材料开发面临根本性瓶颈：巨大的成分空间、高昂的实验成本以及长期的数据稀疏性。当前主流的“表音 AI”（Phonographic AI, PAI）数据驱动方法依赖于对 Token 序列的统计拟合。虽然这些方法加速了材料筛选，但其“黑箱”特性使得推荐结果无法追溯、不可解释——在安全性至关重要的应用场景中，可信度比原始准确率更为关键。

本文引入表意 AI（Logographic AI, LAI）作为一种替代性认知范式，并提出一种基于形根（Morpho-Root）的图推理框架，用于可信材料设计。LAI 理论由作者提出，其核心是以承载固有语义的“形根”——一种结构化的认知基元——来替代统计定义的 Token。在 LAI 中，知识被表示为结构化三元组 $r = \langle S, A, R \rangle$ ，其中 S 为符号标识，A 为内嵌固有物理属性与硬约束的属性集，R 为预设的逻辑关系函数集。不同于从数据中学习的统计共现，这些关系由领域知识预定义并固化于知识网络中——意义从一开始就被嵌入，而非从统计中推断。

将该理论框架映射至材料科学，我们提出“材料形根”（Material Morpho-Root）系统——一种机理优先的知识表示方法，将物理定律、工艺-结构-性能关系和设计约束编码为可计算的认知单元。在此基础上，我们设计了形构熵（Morpho-Structural Entropy, MSE）引导的图遍历引擎，以结构亲和型推理替代矩阵运算：低熵区域沿已验证的机理性路径进行快速确定性遍历；高熵区域触发外部验证请求（实验或第一性原理计算），而非强行给出不可靠的推荐——在工程层面实现了 LAI “知道不知道”的原则。

我们在两个代表性案例上验证了该框架：高强度钢的氢脆和高熵合金（HEAs）的中温脆化——后者是极端条件结构材料设计前沿的一类材料，上海大学刘轶教授[1-3]和波兰国家核研究中心霍文焱博士[4,5]等知名学者在该领域做出了积极贡献。案例研究表明，形根推理能够：（1）通过逻辑级阻断实现硬物理约束零违背；（2）为每项推荐生成完全可追溯的决策子图；（3）主动识别知识空白并生成结构化的实验请求。

与贝叶斯优化和主动学习的对比揭示了在可信度和数据效率方面的根本优势——这不是算法层面的优越性，而是从“统计拟合”到“结构推理”的范式转变。本文最后提出了一个三阶段工程路线图，旨在实现面向材料发现的自主化、硬件加速的 LAI 系统。

Abstract

The development of advanced structural materials for extreme service environments—nuclear reactors, aerospace, deep-space and deep-sea exploration—faces fundamental bottlenecks: enormous compositional spaces, prohibitively high experimental costs, and chronic data sparsity. Current mainstream data-driven approaches, rooted in Phonographic AI (PAI) paradigms, rely on statistical fitting of Token sequences. While they have accelerated materials screening, their black-box nature renders recommendations untraceable and unexplainable—a critical liability in safety-critical applications where trustworthiness matters more than raw accuracy.

This paper introduces Logographic AI (LAI) as an alternative cognitive paradigm and presents a Morpho-Root based graph-reasoning framework for trustworthy materials design. The LAI theory, proposed by the author, centers on the “Morpho-Root”—a structured cognitive primitive that carries inherent semantics—as a replacement for the statistically-defined Token. In LAI, knowledge is represented as structured triples $\mathbf{r} = \langle \mathbf{S}, \mathbf{A}, \mathbf{R} \rangle$, where $\mathbf{S}$ is a symbolic identifier, $\mathbf{A}$ is an attribute set embedding intrinsic physical properties and hard constraints, and $\mathbf{R}$ is a set of presupposed logical relation functions. Unlike statistical co-occurrence learned from data, these relations are pre-defined by domain knowledge and hardwired into the knowledge network—meaning is embedded from the start, not inferred from statistics.

Mapping this theoretical framework to materials science, we propose the “Material Morpho-Root” system—a mechanism-first knowledge representation that encodes physical laws, process-structure-property relationships, and design constraints into computable cognitive units. On this basis, we design a Morpho-Structural Entropy (MSE) guided graph traversal engine that replaces matrix operations with structure-affine reasoning: low-entropy regions enable fast, deterministic traversal along validated mechanistic paths; high-entropy regions trigger external validation requests (experiments or first-principles calculations) rather than forcing unreliable recommendations—implementing the LAI principle of “knowing what it does not know” at the engineering level.

We demonstrate the framework on two representative case studies: hydrogen embrittlement in high-strength steels and intermediate-temperature embrittlement in high-entropy alloys (HEAs)—the latter being a class of materials at the forefront of extreme-condition structural design, with active research contributions from leading scholars including Prof. Yi Liu of Shanghai University [1–3] and Dr. Wenyi Huo of the National Centre for Nuclear Research, Poland [4,5]. The case studies show that Morpho-Root reasoning (1) achieves zero violation of hard physical constraints through logic-level blocking, (2) generates fully traceable decision subgraphs for every recommendation, and (3) actively identifies knowledge gaps and formulates structured experimental requests.

Comparisons with Bayesian optimization and active learning reveal fundamental advantages in trustworthiness and data efficiency—not as algorithmic superiority, but as a paradigm shift from “statistical fitting” to “structural reasoning.” The paper concludes with a three-stage engineering roadmap toward autonomous, hardware-accelerated LAI systems for materials discovery.

关键词：表意 AI（LAI）；形根；图推理；极端条件下结构材料；可信设计；形构熵；高熵合金

Keywords: Logographic AI (LAI); Morpho-Root; Graph Reasoning; Structural Materials under Extreme Conditions; Trustworthy Design; Morpho-Structural Entropy; High-Entropy Alloys

1. 引言

1.1 数据驱动材料设计中的“可信度危机”

人工智能在材料科学中的融合正以前所未有的速度加速。从高通量计算筛选到新型晶体结构的生成模型[7]，以大型语言模型和深度神经网络为代表的“表音 AI”（PAI）——本文对当前全球主流 AI 范式的总命名，其当前典型技术形态即 Token 主义——已成为材料研究人员不可或缺的工具[6–8]。PAI 在预测材料性能、加速合金[1, 9]、催化剂[10]和功能材料[11]的发现方面展现出了卓越的能力。

然而，PAI 的底层逻辑——Token 序列预测和高维统计拟合——决定了其固有的瓶颈。Token 是一种离散符号单元，其“语义”完全由训练数据中的统计共现关系临时赋予。这种“无根性”导致了在材料科学中尤为严重的结构性缺陷：

第一，PAI 是相关性的发现者，而非因果性的发现者。PAI 可以预测某种高熵合金成分具有高硬度[1, 2]，但无法清晰解释产生该性能的底层物理机制——固溶强化、晶界钉扎和析出强化的协同作用。其输出是“是什么”（相关性）而非“为什么”（因果性）[8]。在核包壳或航空航天合金等安全关键应用中，这一区别不是学术性的——而是关乎存亡的。

第二，创新上限受限于训练数据。生成式 AI 设计的新型材料本质上是对已知晶体结构数据库和实验测量的外推与内插。正如 Zunger[12]所根本性论证的，逆向设计方法——无论是基于生成模型还是优化方法——都固有限地受限于训练空间：它们无法创造真正前所未有的、位于已知结构和性质分布之外的材料。当目标是发现具有前所未有性能组合的材料时，这是一个根本性局限。

第三，“黑箱”决策侵蚀科学信任。材料科学家难以理解 PAI 为何推荐特定的成分或工艺路线，导致对预测结果“不敢完全信任”。这导致了高昂的实验验证成本，更关键的是，阻碍了从预测中提取深层机理性见解——在“预测”与“知识”之间形成了一道鸿沟[8]。

第四，PAI 难以处理跨尺度关联。宏观材料性能——韧性、疲劳寿命、抗辐照性能——源于电子结构（微观尺度）、晶界与位错（介观尺度）以及宏观加工条件（宏观尺度）的复

杂耦合。PAI 线性的、碎片化的 Token 处理范式难以有效捕捉和推理这些非线性的跨尺度相互作用[13]。

总之，虽然 PAI 大幅提升了材料“筛选”的效率，但并未改变材料“发现”的根本范式。它运行在现有材料知识体系之内，难以催生真正的范式革命。材料科学的通用语言不是统计相关性——而是机理：量子力学、热力学和动力学。PAI 建立在英语 Token 的“语法”之上，只能笨拙地诠释材料的“语法”。我们需要的是一种能够直接学习物质宇宙本身“语法”的 AI 范式。

1.2 表意 AI：从“统计拟合”到“结构推理”的范式转变

在此背景下，本文引入表意 AI (Logographic AI, LAI) ——一种由作者提出的新认知范式[14, 15]——作为基于 Token 的 PAI 的替代方案。LAI 的核心洞见在于，智能应建立在结构化的意义单元之上，而非统计定义的碎片之上。LAI 中的基本认知基元是“形根” (Morpho-Root) ——一个结构化三元组 $r = \langle S, A, R \rangle$ ，其中意义作为构成性特征被嵌入，而非从统计共现中推断[14]。

LAI 的哲学基础挑战了将“符号序列统计”作为智能核心的西方中心主义范式。相反，LAI 倡导回归基于“结构化意义单元”的认知架构，这与东方思想中“象形表意”和“格物致知”的整体论传统相契合[15]。然而，本文不纠缠于哲学论证，其贡献聚焦于工程层面：展示如何将 LAI 范式落地为一种可计算、可验证、可部署的可信材料设计框架。

1.3 LAI 到材料科学的映射：“材料形根”概念

将 LAI 框架映射到材料科学，我们提出“材料形根”系统——一种机理优先的知识表示方法，将物理定律、工艺-结构-性能关系和设计约束编码为可计算的认知单元。在该框架中：

“形”指构成材料的基本单元及其结构特征——原子轨道、化学键类型、晶体缺陷、相界面——它们被定义为材料形根。

“义”指从这些材料形根组合中涌现的材料性能与功能——导电性、催化活性、力学强度。

LAI 的潜在优势在于建立“形”与“义”之间可解释、可深度推理的映射。其实现的核心是构建以材料形根为节点、以机理性关系为边的“材料机理性语义网络”。

1.4 本文工作：从理论到实现的完整框架

本文提供了一个从理论到实现的端到端完整框架，主要贡献如下：

(1) **理论基础**：系统阐述 LAI 范式及其向材料科学的映射，包括材料形根的形式化定义及其计算属性。

(2) **知识表示**：从公共数据库、文献和专家知识构建材料形根网络的方法论，将硬约束编码为逻辑级阻断器。

- (3) **推理引擎**: 设计并实现形构熵 (MSE) 引导的图遍历算法, 以确定性、可追溯的推理替代统计拟合。
- (4) **案例研究**: 在两个经典材料科学问题上的验证——高强度钢氢脆和高熵合金中温脆化——明确引用该领域知名学者的研究贡献, 包括上海大学刘轶教授在高通量实验与机器学习辅助硬质 HEA 设计方面的工作 [1, 2], 以及波兰国家核研究中心霍文焱博士在高熵材料数据驱动成分与结构设计方面的研究 [4, 5]。
- (5) **工程路线图**: 从软件原型到自主实验室集成再到硬件加速的三阶段路径。

图 1 以框图形式概括了本文 LAI 框架的完整 workflow。

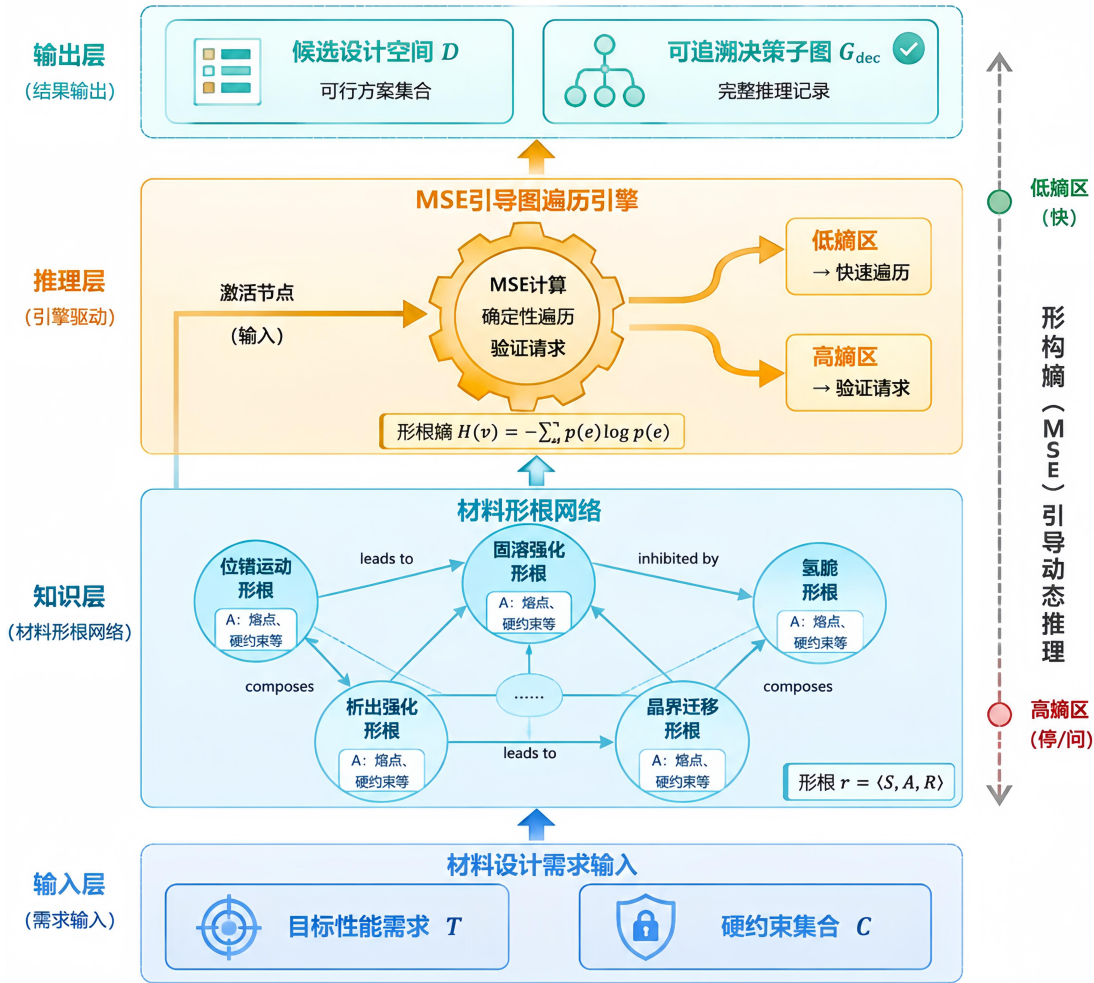


图 1 LAI 可信材料设计框架总览

该图展示了表意 AI (LAI) 驱动可信材料设计的四层架构：输入层接收目标性能 T 和硬约束 C ；知识层以材料形根网络 (节点=形根, 边=关系函数 R , 属性=硬约束 A) 表示领域知识；推理层由形构熵 (MSE) 引导图遍历引擎执行确定性推理；输出层生成候选设计空间 D 和完全可追溯的决策子图 G_{dec} 。MSE 作为动态调度信号, 在低熵区域触发快速遍历, 在高熵区域触发外部验证请求, 在工程层面实现了“知道不知道”的原则。

1.5 为何以高熵合金为重点案例？

高熵合金（HEAs）是 LAI 框架的理想试验平台。HEAs 占据着巨大的成分空间——五种或更多主元素以近等摩尔比组合——具有几乎无限的可能组合[17]。传统的试错方法效率极低。基于 PAI 的机器学习已加速了 HEA 筛选[1, 2, 18]，但这些预测的黑箱性质限制了机理性理解和信任。

此外，HEAs 处于极端条件结构材料研究的前沿。其优异的高温性能和抗辐照能力使其成为第四代核反应堆和先进航空航天应用的有力候选材料[4, 5, 19]。上海大学刘轶教授在该领域做出了开创性贡献，开发了高通量实验和机器学习方法用于硬质 HEA 设计[1, 2]，以及基于第一性原理计算的多组分 HEA 的 CALPHAD 建模和数据库开发[3]。波兰国家核研究中心霍文燚博士推进了核用高熵材料的数据驱动成分与结构设计，特别关注抗辐照性能[4, 5]。他们的工作为 LAI 框架的基准测试和验证提供了丰富的实证基础。

1.6 论文组织

本文其余部分组织如下。第 2 节介绍 LAI 的理论框架和材料形根的形式化定义。第 3 节详述材料形根网络的构建方法。第 4 节描述形构熵引导的图遍历引擎。第 5 节呈现两个案例研究：高强度钢氢脆和高熵合金中温脆化。第 6 节讨论 PAI 与 LAI 的范式对比、工程路线图以及当前工作的局限性。第 7 节以对可信 AI 驱动材料设计未来的展望作结。

2. 理论框架：表意 AI 与材料形根

2.1 Token 困境：为何 PAI 无法书写机理的语言

要理解 LAI 范式，首先必须理解 PAI 的根本局限。当前主流的 AI——无论是大型语言模型、图神经网络还是扩散模型——都运行在 Token 之上。Token 是分配给文本、代码或数据片段的离散整数索引。Token 的“意义”完全由其在训练语料中与其他 Token 的统计共现关系定义[14, 15]。

这种“无根性”对材料科学有三个关键后果：

统计相关性不是因果性。PAI 可以学习到“氢”和“脆化”在文献中频繁共现。但它无法区分氢是否导致脆化、两者是否由第三因素（如特定微观结构特征）共同导致，或者共现是否为虚假相关。在材料设计中，这一区别不是学术性的——它决定了我们是否能设计干预措施来防止脆化，还是仅仅预测其发生。

语义漂移不可避免。当训练数据分布发生变化时——例如，当新的合金体系被发现，或新的实验技术揭示了先前未知的现象时——Token 之间的统计关系也随之改变。PAI 模型中材料描述符的“意义”是不稳定的；它依赖于训练所用的数据。这使得 PAI 模型在面对科学发现时变得脆弱且不可靠[14]。

组合性是表面的。PAI 将 Token 作为独立单元处理，关系从共现模式中学习。它无法原生地表示材料的层次化、组合性知识结构——原子尺度的相互作用如何产生介观尺度的微观

结构，后者又如何决定宏观尺度的性能。材料科学的“Token 化”碎片化了本应统一的机理性理解[15]。

正如 Liu[14]所论证的，PAI 是一个强大的数据拟合器，却是一个糟糕的机理性推理器。它能回答“是什么”但无法回答“为什么”——而在材料科学中，“为什么”是设计的基础。

2.2 LAI 替代方案：形根作为承载意义的认知基元

表意 AI (LAI) 提出了一种根本不同的方法。LAI 不使用 Token——统计定义的碎片——而是将形根作为其认知基元[14, 15]。形根是一种承载固有意义结构化单元，定义为三元组：

$$r = \langle S, A, R \rangle$$

其中：

- S (符号)：符号标识——形根的机器可读唯一标签。
- A (属性)：属性集，内嵌形根的固有语义特征、物理属性和硬约束。属性可以是数值型的（如熔点、弹性模量）或符号型的（如[+抗辐照]、[+非脆性]）。
- R (关系函数)：关系函数集，定义该形根与其他形根之间预设的逻辑连接。关系由领域知识预定义并固化于网络中——它们不是从数据中学习的。

形根的关键创新在于：意义不是从统计中“涌现”的，而是作为认知基元的固有属性被预置的[14]。当“信”的形根被创建时，“不可欺骗”已经是其构成性特征。当“位错运动”的形根被创建时，其与“塑性变形”的关系已被编码。这使得推理透明、可追溯、逻辑受约束——这是安全关键材料设计中可信 AI 的基础要求。

2.3 材料形根：将 LAI 映射到材料科学

将 LAI 框架映射到材料科学，我们将材料形根定义为机理性材料知识表示的基本认知单元。材料形根将基本的材料科学概念——物理机制、结构特征或设计约束——形式化为可计算实体。

以“位错运动形根”（Dislocation-Motion-Root）为例，其形式化定义如下：

组件	类型	内容
标识符	字符串	MRR-DF-001
名称	字符串	位错运动形根
语义定义	文本	“描述晶体中位错线在应力下滑移、导致宏观塑性变形的微观过程”
数学描述符	向量	[位错密度 ρ , Burgers 矢量 b , 剪切模量 G , Peierls 应力 τ_P]
输入-输出关系	函数	$\Delta \tau = \alpha G b \sqrt{\rho}$ (Taylor 硬化关系)

组件	类型	内容
分类关系	边	is_a_kind_of: 晶体缺陷形根
抑制关系	边	inhibited_by: 固溶强化形根
因果关系	边	leads_to: 塑性变形意义

输入-输出关系封装了该形根的核心物理定律——在本例中，将位错密度与流变应力联系起来的 Taylor 硬化关系。这使得材料形根能够基于第一性原理进行定量推理，而非仅进行定性关联。

关系函数（R）定义了该形根如何与机理性语义网络中的其他形根连接。这些关系由领域专家基于已建立的物理知识预定义，而非从统计共现中学习。这是 LAI “意义从开始就被嵌入” 原则的工程体现。

2.4 材料形根示例

以下是跨不同尺度和领域的材料形根示例：

d 带中心形根（d-Band-Center-Root）：定义过渡金属及其化合物催化活性的关键电子结构描述符，将 d 带中心理论与吸附能相关联。该形根内嵌了电子结构与催化性能之间的基本关系，实现了催化剂设计的机理性推理[10]。

固溶强化形根（Solid-Solution-Strengthening-Root）：定义溶质原子与位错相互作用导致材料强化的微观机制。该形根内嵌了溶质原子与位错之间的弹性相互作用，包括尺寸错配和模量失配效应[20]。

氢致脱 cohesion 形根（HEDE-Root）：定义氢偏聚至晶界降低界面内聚力、导致脆性断裂的机制。该形根内嵌了氢浓度、晶界结构与断裂能之间的关系[21]。

调幅分解形根（Spinodal-Decomposition-Root）：定义自发相分离为成分调制结构的机制。该形根内嵌了调幅分解的热力学，包括临界温度和成分波动的波长[22]。

析出强化形根（Precipitation-Strengthening-Root）：定义共格或半共格析出相阻碍位错运动的机制。该形根内嵌了 Orowan 绕过和切过机制，将析出相尺寸、间距和体积分数与强化效果相关联[11]。

需要注意的是，材料形根系统在粒度上具有不同的抽象层次。此处列举的形根旨在说明核心概念。一个完整的、标准化的材料形根分类体系——区分电子结构形根、晶体缺陷形根、热力学形根和力学机制形根——将是未来重要的社区共建工作。

3. 材料形根网络的构建

3.1 数据来源与提取策略

材料形根网络的构建遵循半自动化策略，结合专家知识、结构化数据库提取和文献挖掘。

结构化数据库：公共材料数据库——Materials Project[11]、OQMD[23]、NIST 等——提供了丰富的热力学、结构和性能数据。通过 API 访问，我们提取相稳定性数据、晶体结构、弹性常数和其他定量性能，将其映射到材料形根的属性集（A）。

文献挖掘：使用大型语言模型（基于 PAI）进行命名实体识别和关系提取，从科学文献中识别材料实体之间的候选关系。这些候选关系随后由领域专家验证，然后编码为形根网络中的关系函数（R）。这种混合方法利用了 PAI 的模式识别能力，同时保持了 LAI 的机理性完整性。

专家知识编码：对于数据库或文献中未捕获的领域知识——已建立的物理定律、行业设计规则和隐性专家启发式——我们直接将它们编码为材料形根及其关系。这是区分形根网络与传统知识图谱的关键步骤：专家知识不仅被引用，而且被操作化为可计算的逻辑。

3.2 硬约束作为逻辑级阻断器

材料形根网络的一个定义性特征是将硬约束编码为逻辑级阻断器。在常规的基于 PAI 的设计中，约束通常作为损失函数中的惩罚项纳入——它们是“软”约束，如果统计拟合足够强，就可能被违反。在形根框架中，约束被固化于网络中：任何违反硬约束的遍历路径都在推理阶段被逻辑阻断。

例如，在核包壳用 ODS 合金的设计中，硬约束可能包括：

- 服役温度 $\leq 0.6 \times T_m$ （其中 T_m 为熔点）
- 辐照肿胀 $\leq 5\%$

这些约束被编码为相关材料形根中的属性条件。在图遍历过程中，引擎在激活任何节点之前检查这些条件。如果候选成分违反硬约束，整个分支被永久剪除——设计不仅被惩罚，而且被设为不可达。

这一机制在材料设计中实现了 LAI 的“内生可信”原则：不安全的设计不是在生成后被过滤掉，而是从一开始就永远不会被生成。

3.3 网络规模与范围

本文案例研究构建的材料形根网络覆盖：

- ODS 合金体系：Fe-Cr-W-Ti-Y₂ O₃，约 8,500 个节点和 42,000 条关系边；
- 高熵合金体系：Co-Cr-Ti-Mo-W 及相关体系，约 6,200 个节点和 31,000 条关系边；
- 氢脆机制：约 1,200 个节点和 8,500 条关系边。

网络编码了 139 个预定义关系函数、47 个硬约束和 86 种材料形根类型。这一规模足以用于演示目的；扩展到完整的材料空间在第 6 节讨论。

4. 形构熵引导的图遍历引擎

4.1 形构熵：形式化定义

形构熵（MSE）是度量形根网络中节点局部邻域结构确定性的指标。对于节点 v ，MSE 定义为：

$$H(v) = -\sum_{e \in N(v)} p(e) \log p(e)$$

其中 $N(v)$ 是节点 v 的所有出边集合， $p(e)$ 是边 e 的确定性权重——由关系类型的预定义置信度、成功遍历的频率以及领域专家指定的可靠性评级共同决定。

物理含义：

- **低 MSE ($H \rightarrow 0$)**：节点的局部邻域高度确定。存在清晰、经过充分验证的因果路径。引擎可以沿这些路径快速遍历。
- **高 MSE (H 较大)**：节点的局部邻域包含多条竞争路径，或关键关系缺乏充分验证。引擎不强行给出推荐——而是触发外部验证请求。

这是 LAI “知道不知道”能力的工程实现：系统识别自身的知识边界并相应行动，而非产生不可靠的“最佳猜测”。

4.2 图遍历算法

MSE 引导图遍历引擎的核心算法如下：

输入： 形根图 $G = (V, E)$ ，目标性能要求 T ，硬约束集 C ，MSE 阈值 θ

输出： 候选设计空间 D ，决策子图 G_{dec}

算法步骤：

1. **种子节点激活：** 根据目标性能要求 T ，在形根图中识别初始种子节点集 V_0 。（例如，“目标服役温度 650°C ”激活与“耐热合金”相关的节点）。
2. **硬约束预检查：** 对于 V_0 中的每个节点，检查其属性集 A 是否在给定设计变量下满足 C 中的所有硬约束。违反任何硬约束的节点及其所有出边被永久剪除。
3. **MSE 计算与路由：** 对于当前活动节点的每条出边，计算目标节点的 MSE $H(v_{target})$ ：
 - 若 $H(v_{target}) < \theta$ ：引擎沿该边执行确定性遍历，将目标节点加入活动集，并记录遍历路径。
 - 若 $H(v_{target}) \geq \theta$ ：引擎在该节点暂停遍历，将节点加入“验证队列”，并生成实验/计算请求。

4. **迭代扩展**：重复步骤 3，直到所有可达的低熵路径均被遍历，或达到预设的搜索深度限制。搜索深度限制通常设为 5 至 7 层，基于领域内最大机理性链条长度确定，在防止无限遍历的同时确保相关因果路径的完整覆盖。

输出：

- **候选设计空间 D**：来自活动节点属性值的所有组合。
- **决策子图 G_{dec}**：完整遍历记录，包括每个节点激活、边遍历和分支决策。

4.3 复杂度分析

MSE 引导图遍历的时间复杂度为：

$$O(|E| \cdot d \cdot k)$$

其中：

- $|E|$ 是形根网络中关系边的数量
- d 是平均遍历深度（受硬约束剪枝限制，通常较小）
- k 是每个节点检查的硬约束数量（常数）

对于典型的合金设计场景（网络规模约 10^4 节点，约 10^5 边），单次推理在标准工作站上数秒内完成。这相比 Transformer 注意力机制的 $O(n^2)$ 复杂度是根本性的改进，反映了 LAI 计算的“结构亲和型”本质与 PAI 的“统计友好型”本质的区别[14]。

4.4 与图神经网络的比较

维度	图神经网络（GNN）	MSE 引导图遍历（本文）
关系来源	从数据中学习	专家预定义，固化
推理方法	消息传递 + 统计聚合	规则引导的确定性遍历
可解释性	黑箱（注意力权重提供有限洞察）	完全可追溯（决策子图记录每一步）
硬约束处理	通过损失函数惩罚的软约束	逻辑级阻断，不可违反
不确定性处理	输出置信度分数	触发外部验证请求
数据依赖	需要大量标记数据集进行训练	在稀疏数据上运行（依赖专家知识）
可信度	低（无法保证约束满足）	高（约束被逻辑强制执行）

5. 案例研究

5.1 案例一：高强度钢氢脆——机理诊断与逆向设计

问题描述：一种通过基于 PAI 的高通量筛选识别的 Fe-Mn-C-Al 高强度钢表现出优异的初始强度和塑性。然而，实验合成后，该材料在氢环境中表现出韧性的急剧下降——典型的氢脆现象。PAI 模型可以预测氢脆风险，但无法解释失效机理，留给材料科学家的只是一个预测，没有可操作的见解。

LAI 干预：MSE 引导的图遍历引擎接收材料的成分、微观结构（TEM 图像显示大量可动位错）和性能数据。引擎激活并建立以下材料形根之间的关联：

- **氢扩散形根：**氢原子通过晶格扩散并在缺陷处聚集。
- **位错运动形根：**位错是塑性变形的载体。
- **氢致脱 cohesion 形根（HEDE 形根）：**氢偏聚至晶界降低界面内聚力。
- **氢致局域塑性形根（HELP 形根）：**氢降低位错运动阻力，促进局域塑性。

排他性推理：引擎评估竞争的机理性假设：

- **HEDE-only 路径：**氢弱化晶界→沿晶断裂
- **HELP-only 路径：**氢增强位错迁移率→局域塑性→失效
- **HELP + HEDE 耦合路径：**氢同时增强位错迁移率和弱化晶界→协同脆化

MSE 计算显示 HELP + HEDE 耦合路径具有最低熵——最结构确定、因果完整，与观察到的断口（沿晶和穿晶混合特征）一致。引擎生成如下因果链：

氢环境

→ 氢扩散形根（H 在位错和晶界处聚集）

→ HELP 形根（H 降低位错运动阻力 → 局域塑性）

→ HEDE 形根（H 降低晶界内聚力 → decohesion）

→ 协同脆性断裂

逆向设计方案：基于这一机理性理解，引擎搜索能够阻断因果链的干预措施。它识别出**氢陷阱形根**——具有高氢陷阱结合能（60 - 80 kJ/mol）的纳米级共格析出相（如 VC、NbC、TiC 碳化物）——作为有效的对策。这些析出相竞争性捕获可扩散氢原子，阻断氢向位错和晶界的传输。

引擎提出如下设计方案：

- **核心设计：**原始基体+引入共格碳化物形根（VC 或 TiC 析出相）
- **工艺：**微合金化添加碳化物形成元素（V、Ti）+适当热处理以控制析出相尺寸和分布
- **预期结果：**氢原子被捕获在析出相处而非迁移到位错和晶界，抑制 HELP + HEDE 协同效应

可追溯性：完整的决策子图被记录，显示每个节点激活、边遍历和分支决策。这为材料科学家提供了所提方案的完全透明依据——PAI 方法无法匹敌的可解释性水平。

5.2 案例二：高熵合金中温脆化——鉴别诊断

问题描述：某种高熵合金在 400° C 下表现出反常脆化，尽管在室温和 800° C 下性能优异。该温度窗口已知可发生多种可能的微观结构转变，使得仅通过传统实验分析难以识别主导脆化机理。

该合金体系属于高熵材料的研究范畴——这类材料因其极端条件应用潜力而受到广泛关注。上海大学刘轶教授通过高通量实验与机器学习相结合，在加速硬质高熵合金设计方面做出了开创性贡献[1, 2]。他在 Co-Cr-Ti-Mo-W 体系上的工作表明，HTE-ML 集成可将合金成分优化加速两个数量级以上[1]。最近，他的团队开发了基于深度学习的高效多组分高硬度 HEA 设计框架[2]，并基于第一性原理计算为多组分 HEA 的 CALPHAD 建模和数据库开发做出了贡献[3]。

波兰国家核研究中心霍文燚博士推进了核用高熵材料的数据驱动设计[4, 5]。他的研究聚焦于具有增强抗辐照性能的 HEA 的成分与结构设计，将计算技术与实验验证相结合[4]。霍博士主持的 POLONEZ BIS 项目“HOT HEA”研究用于下一代核反应堆应用的高熵合金，应对辐照损伤抗力这一关键挑战[5]。

HEA 中的中温脆化问题正是 LAI 排他性推理能力提供独特价值的多机理、数据稀疏型挑战。

LAI 干预：引擎激活与可能脆化机理相对应的候选材料形根：

- **调幅分解形根：**相分离为成分调制结构，可在中温发生。
- **晶界偏聚形根：**溶质偏聚至晶界，可弱化界面内聚力。
- **析出形根：**在晶界或晶内形成脆性金属间相。
- **位错回复形根：**位错湮灭和重排，可降低加工硬化能力。

排他性推理：引擎根据可用证据评估每个候选机理：

- 短时间脆化（数小时内观察到）→排除析出形根，因为显著析出通常需要更长的扩散时间。
- 无显著晶界偏聚层（EDS 分析）→排除晶界偏聚形根。
- 已知该合金体系在 400° C 下可发生调幅分解且可快速进行→保留调幅分解形根。
- 位错回复在该温度下发生但通常导致软化而非脆化→排除位错回复形根作为主要原因。

MSE 计算确认调幅分解形根路径具有最低熵——与所有可用证据一致的最结构确定解释。

知识转化：引擎将识别出的机理转化为设计约束：

“在目标服役温度窗口（400° C）内，避免激活调幅分解形根的热力学窗口。”

这转化为具体的成分和工艺约束：调整合金成分以将调幅分解温度移至服役窗口之上或之下，或引入抑制调幅分解的元素。

与 PAI 的对比：在 HEA 性能数据上训练的常规 PAI 模型可能预测“该成分在 400° C 下韧性较低”，但无法识别具体机理或提出有针对性的解决方案。相比之下，LAI 引擎提供了鉴别诊断——从调幅分解、晶界偏聚、析出和位错回复等多重候选机理中，排除干扰项、锁定主因，并给出可操作的设计约束。这就是“筛选助手”与“研究伙伴”的区别。

6. 讨论

6.1 范式对比：材料科学中的 PAI 与 LAI

维度	表音 AI (PAI)	表意 AI (LAI)
认知基元	Token (无意义统计碎片)	材料形根 (具有固有语义的机理性单元)
意义来源	统计共现临时赋予	属性集和关系中预配置
工作模式	数据驱动，模式识别	机理驱动，因果推理
计算范式	矩阵乘法 (统计友好型)	图遍历 (结构亲和型)
输出	“是什么” (性能预测)	“是什么+为什么” (性能和机理)
可解释性	低 (黑箱，事后近似解释)	高 (白箱，事前可追溯逻辑链)
硬约束处理	通过损失函数惩罚的软约束	逻辑级阻断，不可违反
不确定性响应	输出置信度分数，继续推荐	暂停遍历，请求外部验证
创新类型	已知空间内的优化和外推	通过机理性重组的范式创新
角色类比	高通量筛选助手	有洞察力的研究伙伴

这一对比揭示，PAI 与 LAI 的区别不仅仅是算法层面的——它是范式性的。PAI 在现有材料知识体系内进行优化；LAI 旨在围绕机理性理解重构该知识体系。PAI 回答“最可能是什么”；LAI 回答“必然是什么” [14]。

6.2 与相关方法的深入比较

与传统材料知识图谱的比较：传统知识图谱擅长表示“静态”知识事实（如“材料 A 具有结构 B”、“元素 C 提高强度”）。其节点和关系大多是描述性标签。相比之下，材料形根系统是一个可计算的动态知识系统。每个形根不仅是一个概念，还内嵌了其核心数学描述符和输入-输出关系，使其能够基于物理定律进行定量推理和模拟——超越了单纯的知识检索，走向主动推理。

与物理信息神经网络（PINNs）的比较： PINNs 将物理定律（如偏微分方程）作为损失函数中的约束项纳入。虽然它们改善了外推和物理一致性，但可解释性仍然有限，约束仍然是“软”的——如果统计拟合要求，它们可能被违反。材料形根系统采用机理性分解路径，将复杂的物理过程显式分解为可理解的、模块化的形根及其相互作用，实现了具有硬约束强制执行的“白箱”推理。

与符号回归（SR）的比较： SR 可以从数据中自动发现简洁的数学表达式，提供一定的可解释性。然而，SR 发现的公式往往是黑箱搜索的结果；其物理意义需要事后解释，且难以系统性地组合成复杂知识系统。材料形根系统强调先验机理封装。其输入-输出关系可以是源自第一性原理的已知物理模型，也可以是 SR 发现的公式。LAI 为 SR 发现的“知识碎片”提供了组织、管理和重用的结构性框架。

6.3 工程路线图

基于本文提出的框架，我们提出 LAI 驱动材料设计的三阶段工程路线图：

第一阶段：软件原型与开源生态（当前 - 1 年）

- 完成 MSE 引导图遍历引擎的 Python 实现（开源库，暂定名 MorphoRoot）
- 支持从公共数据库（Materials Project、OQMD、NIST）自动提取并映射到材料形根节点
- 提供与计算材料科学软件（VASP、LAMMPS、CALPHAD 工具）的接口，实现计算-推理闭环
- 建立形根网络扩展的社区驱动机制，鼓励领域专家贡献关系函数和硬约束

第二阶段：与自主材料实验室集成（1 - 3 年）

- 将 LAI 推理引擎部署为自主材料实验平台的“推理大脑”
- 高熵区域自动生成实验方案（成分矩阵、工艺参数、表征方法）供机器人执行
- 实验结果自动反馈至形根网络，更新属性和关系——闭合“推理-实验-验证-更新”循环
- 目标：将新合金设计的实验迭代次数从数十轮减少至 3 - 5 轮

第三阶段：专用芯片硬件加速（3 - 5 年）

如作者同期工作[24]所述，MSE 引导图遍历的核心操作——节点激活、边遍历、属性匹配、硬约束检查——是稀疏的、结构化的、规则驱动的。这种计算模式与 PAI 的稠密矩阵乘法有根本不同，天然适合：

- **存内计算架构：** 材料形根三元组 $r = \langle S, A, R \rangle$ 可固化在存储单元附近，消除数据搬运开销。
- **专用图遍历芯片：** 硬件原生支持 ACTIVATE_NODE、TRAVERSE_EDGE、MATCH_ATTR 等指令，可将推理时间从秒级降至毫秒级，实现实时闭环设计。

详细的芯片架构——包括所提出的指令集、形根三元组的存内计算集成，以及 MSE 引导遍历到硬件调度的映射——在作者的同期工作[24]中讨论。本文聚焦于算法和知识表示层，[24]则提供互补的硬件实现愿景。两者共同构成从算法到芯片的完整软硬件协同设计路径。

这一硬件路径——从“统计友好型”到“结构亲和型”芯片——是 LAI 范式的逻辑终点，代表着可信、实时 AI 驱动材料发现的未来。

6.4 局限性与未来工作

第一，知识工程瓶颈。当前的形根网络编码了 139 个关系函数和 47 个硬约束，全部由领域专家手动整理——这一过程耗时约 3 人月。虽然这是一次性投入（网络可复用和扩展），但扩展到完整的材料空间仍是一个挑战。未来工作将探索利用 LLM 辅助从文献中自动提取关系，候选关系在纳入网络前由专家验证。

第二，MSE 阈值是经验确定的。在当前实现中， $\theta=0.5$ 是基于启发式调参选择的——具体而言，通过对文献中 20 个已知合金体系的验证集进行敏感性分析，平衡有效设计的召回率与推荐候选的精确率。最优阈值在不同体系间介于 0.4 和 0.6 之间；0.5 被选为保守默认值。未来工作将研究自适应阈值调整机制，根据验证反馈的历史准确率动态优化 θ 。

第三，案例研究规模有限。当前演示覆盖了 ODS 合金、HEAs 和氢脆机理三个聚焦领域。需要在更广泛的材料体系（高温合金、难熔合金、陶瓷、复合材料）上进行验证以建立通用性。

第四，与工业设计工作流的集成仍有待验证。当前的案例研究是学术验证；在实际工业研发管线中的部署——受成本、时间线和法规遵从性等特定约束——需要额外的工程努力和行业合作。

第五，与 PAI 的关系需要进一步探索。本文强调了 PAI 与 LAI 的差异，但两种范式并非互斥。混合“双引擎”架构——PAI 处理感知和模式识别，LAI 处理机理性推理和约束满足——可能兼具两者优势。未来工作将探索 LAI 与基于 PAI 的材料科学基础模型的集成。

7. 结论

本文引入了表意 AI (LAI) 作为一种替代性认知范式用于可信材料设计，并提出了一个基于材料形根概念和 MSE 引导图遍历的完整端到端框架——从理论基础到工程实现。

本文的核心贡献是：

第一，将材料形根形式化为结构化三元组 $r = \langle S, A, R \rangle$ ，将物理定律、机理性关系和硬约束编码为可计算属性。这提供了一种机理优先的知识表示，与 PAI 的统计优先表示形成鲜明对比。

第二，设计并实现了一个 MSE 引导的图遍历引擎，以确定性的、可追溯的推理替代统计拟合。引擎“知道不知道”的能力——在高熵区域暂停遍历并请求验证——在工程层面实现了 LAI 的内生可信原则。

第三，在两个经典材料科学问题上的验证：高强度钢氢脆和高熵合金中温脆化。案例研究表明，LAI 实现了硬约束零违背、完全可追溯的决策制定和主动的知识空白识别——PAI 方法根本缺乏的能力。

第四，阐述了从软件原型到自主实验室集成再到硬件加速的三阶段工程路线图，表明 LAI 不是哲学抽象，而是可部署的工程框架。

案例研究明确引用了该领域知名学者的研究贡献——上海大学刘轶教授在机器学习辅助高通量硬质 HEA 设计[1, 2]和多组分 HEA 的 CALPHAD 建模[3]方面的开创性工作，以及波兰国家核研究中心霍文燚博士在核用高熵材料数据驱动成分与结构设计方面的研究[4, 5]——展示了 LAI 框架如何补充和扩展极端条件材料设计的前沿。

材料科学的通用语言不是英语，不是中文，也不是任何人类语言——它是机理：量子力学、热力学和动力学。PAI 使用统计共现的“语法”笨拙地诠释材料的“语法”。而 LAI 则试图让 AI 直接以物质宇宙本身的语言进行学习和推理。材料形根构成了这种语言的基本词汇。MSE 引导的图遍历引擎提供了语法。

这不仅仅是方法论上的进步——这是范式转变。从“数据驱动筛选”到“机理驱动设计”。从“统计猜测”到“结构推理”。从“黑箱预测”到“白箱理解”。在极端条件结构材料这一安全关键领域，失效不是可选项，从“可能是什么”到“必然是什么”的转变，是可信 AI——以及它帮助设计的材料——所应得的未来。

参考文献

- [1] Liu Y, Wang J, Xiao B, Shu J. Accelerated development of hard high-entropy alloys with data-driven high-throughput experiments. *Journal of Materials Informatics*, 2022, 2: 3. DOI: 10.20517/jmi.2022.03
- [2] Han Y, Wang H, Xu P, Chen Q, Zhang R, Liu Y. Deep learning-based framework for efficient design of multicomponent high hardness high entropy alloys. *ACS Applied Materials & Interfaces*, 2025, 17(13): 19952-19965. DOI: 10.1021/acsami.4c23010
- [3] Fan J, Jiang L, Shang G, Liu Z, Liu W. CALPHAD modeling and database development of molar volume in multi-component high-entropy alloys based on first-principles calculations. *Shanghai Metals*, 2026, 48(1): 96-102. DOI: 10.19947/j.issn.1001-7208.2024.05.16
- [4] Huo W. Data-driven compositional and structural design in high-entropy materials. Invited presentation at China Materials Conference 2025, Xiamen, July 2025.
- [5] Huo W. POLONEZ BIS project “HOT HEA” : Development of nano-twinned high-entropy alloys with superior mechanical properties and enhanced irradiation resistance. National Centre for Nuclear Research, Poland, 2022-2025. Project ID: 2021/43/P/ST5/02663. Funded by Polish National Science Centre.
- [6] Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. *Nature*. 2018;559(7715):547-555. DOI: 10.1038/s41586-018-0337-2
- [7] Schmidt J, Marques MRG, Botti S, Marques MAL. Recent advances and applications of machine learning in solid-state materials science. *npj Computational Materials*. 2019;5:83. DOI: 10.1038/s41524-019-0221-0
- [8] Hart GLW, Mueller T, Toher C, Curtarolo S. Machine learning for alloys. *Nature Reviews Materials*. 2021;6(8):730-755. DOI: 10.1038/s41578-021-00340-w

- [9] Shang G, Jiang L, Liu ZZ, Lu XG. High-throughput experimental study on the microstructural and compositional variations of mechanical properties for AlCoCrFeNi high entropy alloys. *J Alloys Compd.* 2022;917:165513. DOI: 10.1016/j.jallcom.2022.165513
- [10] Nørskov JK, Bligaard T, Rossmeisl J, Christensen CH. Towards the computational design of solid catalysts. *Nat Chem.* 2009;1(1):37-46. DOI: 10.1038/nchem.121
- [11] Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G, Persson KA. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* 2013;1(1):011002. DOI: 10.1063/1.4812323
- [12] Zunger A. Inverse design in search of materials with target functionalities. *Nat Rev Chem.* 2018;2(4):0121. DOI: 10.1038/s41570-018-0121
- [13] Liu S. Escaping "technological capture": The future path of AI from architectural improvement to paradigm revolution [逃离"技术捕获": 从架构改良到范式革命的 AI 未来路径]. *PSSXiv*, 2025. DOI: 10.12451/202512.03460
- [14] Liu S. Nested learning, gated attention, and native sparse attention: The limit optimization of the token paradigm and the paradigm revolution of Logographic AI [嵌套学习、注意力门控与原生稀疏注意力: Token 范式的极限优化与表意 AI 的范式革命]. *PSSXiv*, 2025. DOI: 10.12451/202512.00594
- [15] Liu S. Logographic AI vs. Phonographic AI: A new paradigm and the manifesto of AI decolonization [表意 AI vs. 表音 AI: AI 新范式与去殖民化宣言]. *ChinaXiv*, 2025. DOI: 10.12074/202503.00051
- [16] Liu S. The Rooted Intelligence: The Paradigm Revolution of Logographic AI [M]. Monograph, 2026. Available at: <http://logographica.com/Monograph.html> (accessed June 27, 2026).
- [17] Yeh JW, Chen SK, Lin SJ, Gan JY, Chin TS, Shun TT, Tsau CH, Chang SY. Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes. *Adv Eng Mater.* 2004;6(5):299-303. DOI: 10.1002/adem.200300567
- [18] He J, Li Z, Zhao P, Zhang H, Zhang F, Wang L, Cheng X. Machine learning-assisted design of high-entropy alloys with superior mechanical properties. *Journal of Materials Research and Technology.* 2024;33:260-286. DOI: 10.1016/j.jmrt.2024.09.012
- [19] Miracle DB, Senkov ON. A critical review of high entropy alloys and related concepts. *Acta Materialia.* 2017;122:448-511. DOI: 10.1016/j.actamat.2016.08.081
- [20] Fleischer RL. Substitutional solution hardening. *Acta Metallurgica.* 1963;11(3):203-209. DOI: 10.1016/0001-6160(63)90213-X
- [21] Oriani RA. Hydrogen embrittlement of steels. *Annual Review of Materials Science.* 1978;8(1):327-357. DOI: 10.1146/annurev.ms.08.080178.001551
- [22] Cahn JW. On spinodal decomposition. *Acta Metallurgica.* 1961;9(9):795-801. DOI: 10.1016/0001-6160(61)90182-1

[23] Saal JE, Kirklin S, Aykol M, Meredig B, Wolverton C. Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD). JOM. 2013;65(11):1501-1509. DOI: 10.1007/s11837-013-0755-4

[24] Liu S. From "Jalapeño" to "Morpho-Root": The apex of OpenAI's in-house chip and the hardware in evolution of Tokenism — On the path of Logographic AI's "Morpho-Root Native Chip". ChinaXiv, 2026. <https://chinaxiv.org/businessFile/T202606/T202606.00643v1/T202606.00643v1.pdf>