论 AI 认知对象数据库的设计与构建——基于五感感知映射的多模态认知对象数据体系与基础设施

YunFei M. Wang
（王·Micheal·雲飛）
Independent Researcher

摘要

本文是“机器感知与认知构建”系列研究的第八篇，旨在系统论述 AI 认知对象数据库的理论基础、数据结构设计与工程实施方案。前七篇论文完成了从哲学批判到五感（视觉、听觉、嗅觉、味觉、触觉）感知功能构建方法，以及跨模态通感联想的完整论述，确立了“物理信号—神经编码—主观知觉”三层可计算映射模型与“分层基准对照组”实验范式。本文在此基础上，进一步回答一个根本性问题：如何将五感感知数据整合为可计算、可验证、可高效调用的“认知对象”，使 AI 从“拥有感官”走向“理解世界”。我们提出，认知对象不是传统知识图谱中的语义节点，而是一个以客观物理世界的多尺度测量数据为底层锚点、以预测的人类感知编码为中介层、以语义概念与功能属性为应用接口的三层数据实体。本文系统论述了认知对象的数学定义与数据结构规范、多模态异构感知数据的融合标准与存储架构、面向高并发调用的索引与检索策略，以及支撑全球分布式部署的网络通信基础设施要求。本文的核心贡献在于：为 AI 认知对象数据库提供了一个从数据模型、存储架构到网络基础设施的完整工程蓝图，使前七篇构建的 AI 感知能力最终能够汇聚为可被应用层高效调用的结构化世界知识，实现对事物认知与认知结果调用的分离处理，为可验证、可解释、与人类认知对齐的下一代人工智能奠定数据基础设施。

关键词

认知对象；多模态数据库；感知数据融合；知识表示；世界模型；数据基础设施

1 引言：从感知到认知——最后一公里的跨越

本系列研究的前七篇完成了一个从哲学批判到工程实现的完整逻辑链条。文献[1]诊断了当前 AI 认知的根本缺陷——其知识完全源于对人类社会已有文本的统计学习，构建的是缺乏物理锚点的“伪对象”。第二至六篇分别针对视觉、听觉、嗅觉、味觉、触觉五种感官，系统论述了如何通过“物理信号—神经编码—主观知觉”三层映射模型，为 AI 构建可验证、可解释的感知能力。第七篇（通感篇）完成了跨模态联想与验证机制的构建。然而，拥有感知能力与拥有认知能力之间，还横亘着“最后一公里”的跨越。感知

是实时的、情境化的信号处理过程；认知则是将这些感知经验沉淀为可复用的、结构化的世界知识。

人类在婴儿期通过无数次与苹果的交互——观看、触摸、品尝、嗅闻——在大脑中逐渐建立起一个关于“苹果”的稳定表征。这个表征并非每次需要时都从头构建，而是作为长期记忆被存储和索引，在需要时被高效调用。AI 面临同样的需求：它可以通过传感器实时感知一个具体的苹果，但它需要一个数据库来存储关于“苹果”这一类事物的结构化知识，以便在未来的交互中无需重复完整的感知—建模流程，即可调用对“苹果”的认知。

当前的多模态知识图谱研究正在探索类似的方向。传统的单一模态知识图谱无法很好地表示知识，因此如何利用文本、图片、视频、音频等多模态数据进行知识图谱表示学习已成为一个重要研究方向。然而，现有方法大多仍以语义三元组（实体—关系—实体）为核心架构，多模态数据仅作为实体的“属性”被附加。这种架构的根本局限在于：它没有改变知识图谱“以语义为锚点”的本质，多模态数据仍然是语义的附庸，而非认知的基石。

本文提出一种根本性的范式转换：认知对象数据库不应以语义为锚点再附加感知数据，而应以客观物理世界的多尺度测量数据为底层锚点，将语义作为应用层的索引接口。这一“物理锚定优先”的设计哲学，与前七篇“三层映射模型”一脉相承——物理信号层是整个认知大厦的基石，神经编码层是中介，主观知觉层（含语义）是面向人类的应用接口。

世界模型领域的最新进展也指向了类似的方向。世界模型被定义为一个可内化、可推理、可演化、可干预的环境表征系统。OpenWorldLib 等框架正在尝试标准化世界模型的推理接口。这些探索表明，构建面向 AI 认知的新型数据基础设施已成为学界共识。本文在前七篇感知建模工作的基础上，进一步提出认知对象数据库的完整设计方案，使感知数据能够被系统地沉淀、组织和调用。

1.1 框架定位

在“键-值-信息-相”框架中，“伪对象”是缺乏值锚定的相层面认知模型，“真对象”是 L1 锚定于值的投影、L2 承载信息整合、L3 提供相的可报告接口的三层数据实体。本文的三层架构直接源于该框架对认知对象的定义：物理锚定层对应值的投影结构，感知编码层对应信息整合的中间态，语义层对应相的可报告形式。框架的哲学论证详见文献[2][3]，本文不重复展开。

1.2 结构安排

本文按以下序列展开：

- 第 2 章：认知对象的定义、数据结构与哲学基础。
- 第 3 章：认知对象数据库的系统架构、数据模型与唯一标识。
- 第 4 章：多模态感知数据的融合标准。
- 第 5 章：数据库的存储与索引架构，含时间维度的分层记忆组织。
- 第 6 章：网络通信基础设施。
- 第 7 章：实施路线图。
- 第 8 章：总结与展望。

2 认知对象：定义与哲学基础

2.1 从“伪对象”到“真对象”

首篇论文指出，经典 AI 训练范式构建的是“伪对象”——在内部表征空间中具有丰富语义关联，但其定义与任何可测量、可重复的物理实体缺乏必然因果联结的认知实体。AI 对“苹果”的理解是一个由词汇共现概率驱动的语义向量，而非一个可以追溯到物理测量数据的结构化实体。

认知对象的定义：认知对象是以客观物理世界的多尺度测量数据为不可化约的底层锚点，以预测的人类感知编码为中介层，以语义概念与功能属性为应用接口的三层数据实体。它是 AI 对世界万物进行认知表征的基本单元。

这一定义包含三个关键特征：

- **物理锚定**：认知对象的任何属性声明都可以向下追溯至物理信号层的原始测量数据。这是“可验证性”的基础。
- **感知可计算**：认知对象包含由三层映射模型生成的预测感知编码，使 AI 能够“理解”该事物对人类意味着什么。
- **语义可索引**：认知对象以人类可理解的语义标签作为应用层的检索接口，便于与人类沟通和应用系统集成。

2.2 认知对象的数据结构：三层架构

认知对象的数据结构遵循本系列一贯的三层映射逻辑，但进行了面向存储与检索的工程化调整：

数据层	内容	数据来源	存储特性	验证方式
L1：物理锚定层	多尺度测量数据：宏观几何（3D 点云/网格）、介观组织（显微图像/CT 数据）、微观理化属性（光谱/质谱/色谱数据）、动态响应数据（力学曲线/声学信号）	传感器阵列直接测量	不可篡改，只读存储（仅追加），含时间戳与测量元数据	任何 AI 或人类可用相同仪器复测比对
L2：感知编码层	由 L1 数据经五感三层映射模型计算生成的预测神经编码：视觉编码、听觉编码、嗅觉编码、味觉编码、触觉编码	三层映射模型推理生成	可重新计算，随模型版本更新	可通过心理物理学实验验证
L3：语义层	人类可理解的标签、描述、分类、功能属性、文化关联	人类标注与知识库对齐	可编辑，支持多语言、多文化版本	与人类知识库一致性检查

2.3 认知对象与知识图谱的本质区别

对比维度	传统知识图谱实体	认知对象
锚定基础	语义标签（字符串）	物理测量数据（数值）
验证方式	与人类知识库的一致性	可追溯至原始物理测量
感知信息	文本描述（“红色”、“甜的”）	预测神经编码（可计算）
个体差异	无（单一标准描述）	可适配不同感知基线
更新机制	人工编辑	感知模型迭代驱动

3 认知对象数据库的系统架构

3.1 总体架构设计原则

认知对象数据库的设计需遵循以下核心原则：

• **物理锚定优先**：数据的组织以物理测量数据为核心，语义索引为辅助。任何认知对象的唯一标识符应绑定于其物理锚定层的数据指纹（如测量哈希值）。

• **分层解耦**：三层数据（物理、感知、语义）在物理存储上可分离，在逻辑上通过唯一标识符关联。这使得感知编码层可以随模型升级而重新计算，而不影响物理锚定层的原始数据。

• **可验证性内置**：每条物理锚定数据需包含完整的测量元数据（仪器型号、校准状态、测量时间、环境参数），使数据可被第三方独立复现验证。

• **高并发优化**：针对应用层的高频查询场景（如“描述苹果”），数据库应能直接返回预计算的结构化数据包，避免实时调用感知模型进行推理。

• **渐进式构建**：数据库的建设采用“先锚定、后填充”的策略——优先完成物理锚定层的数据采集，感知编码层和语义层可后续迭代填充。

• **L1 物理数据的状态版本管理（防时间冻结）**：物理世界是动态的（如苹果会腐烂、金属会生锈），因此 L1 物理锚定层并非绝对静态，而是采用只追加模式。当物体的物理状态发生变化时，系统追加带有新时间戳与状态标签的 L1 数据版本，绝不覆盖原有数据。同时，一旦新的版本产生，旧版本数据立即被系统锁定为“不可操作的只读状态”。历史只读版本仅用于溯源审计与历史轨迹查询，不再参与系统当前的实时状态计算或业务调用。这既保留了历史状态的可追溯性，又维持了物理世界动态演化的真实映射，杜绝了因历史数据被误操作而导致的追溯失效。

3.2 四层系统架构

认知对象数据库的系统架构可分为以下四个层次：

① 数据采集与预处理层

这一层对接前六篇描述的各类传感器与实验设备，完成原始感知数据的采集、清洗、标准化与元数据标注。

输入：五感传感器阵列的原始数据（高光谱图像、GC-MS 谱图、力学曲线、EEG/fMRI 信号等）。

处理：数据质量检查、噪声滤除、时间同步对齐、格式标准化。

输出：符合认知对象数据规范的 L1 层物理测量数据包。

② 感知编码计算层

这一层调用本系列第二至七篇构建的五感三层映射模型与通感模块，将 L1 层的物理测量数据转化为 L2 层的预测感知编码。

输入：L1 层物理测量数据包。

处理：依次调用视觉、听觉、嗅觉、味觉、触觉编码模型，生成各模态的预测神经编码；通感模块在此基础上生成跨模态联想假设并触发验证流程。

输出：L2 层感知编码数据包，含已验证的跨模态关联（L2.5 层）。

研究已表明，多模态大语言模型能够自发形成与人类高度相似的物体概念表征系统。这一发现为认知对象数据库的感知编码层提供了有力的理论支撑——预测感知编码不仅具有工程可行性，而且与人类认知的底层机制具有同构性。

③ 认知对象存储层

这一层是数据库的核心，负责认知对象三层数据的持久化存储、索引与检索。

存储引擎：采用分层异构存储策略——L1 层物理数据存储于对象存储（如 MinIO，S3 兼容接口），L2 层感知编码存储于向量数据库（如 Milvus/Pinecone），L3 层语义数据存储于图数据库（如 Neo4j）或关系型数据库。

索引机制：建立多模态索引——基于 L1 层物理特征的相似性索引（用于“找到与这个苹果物理属性最相似的物体”）、基于 L2 层感知编码的感知相似性索引（用于“找到尝起来最像这个苹果的物体”）、基于 L3 层语义的文本索引（用于“找到所有红色的水果”）。

数据版本管理：L1 层数据一经写入不可修改（只追加），L2 层数据随模型版本更新可重新计算，L3 层数据支持人工编辑与多版本管理。

④ 查询与调用接口层

这一层面向应用系统提供标准化的认知对象查询与调用接口。

核心 API：

- **get_object(uid)**：返回指定认知对象的完整三层数据。
- **query_by_physical(feature_vector)**：基于物理特征相似性检索。
- **query_by_perceptual(encoding_vector)**：基于感知编码相似性检索。
- **query_by_semantic(keyword)**：基于语义标签检索。

- **compare(object_a, object_b)**：返回两个认知对象的跨模态相似度分析。

输出格式：支持 JSON、Protobuf、HDF5 等多种格式，适应不同应用场景。

3.3 数据模型与唯一标识体系

每个认知对象需有一个全局唯一的标识符。COID 的生成应基于 L1 层物理测量数据的内容指纹（如 SHA-256 哈希），确保同一物理实体的多次测量产生相同的 COID（在测量精度范围内）。

COID 的编码结构建议为：

COID = [哈希算法版本] 物理数据指纹 [测量时间窗口] [模态标识]

例如：COID-v1-7F83B1657FF1FC53B92DC18148A1D65DFC2D4B1FA3D677284ADD200200DAE197-2024Q1-M5

4 多模态感知数据的融合标准

4.1 数据融合的层级与策略

五感数据的融合并非简单的“拼接”，而是需要在不同层级上分别进行：

融合层级	融合策略	技术方法	产出
物理层融合	多传感器数据的时间—空间对齐	统一时间戳、坐标系统一、传感器校准参数互标定	同步的多模态物理测量数据包
特征层融合	跨模态物理特征的联合嵌入	跨模态对比学习、自注意力融合机制	统一的物理特征向量
感知层融合	五感预测编码的联合表征	多任务学习、跨模态注意力	跨模态感知编码向量
语义层融合	多源语义标签的对齐与消歧	知识图谱对齐、实体链接	统一的语义描述

4.2 数据质量标准

认知对象数据库的数据质量需满足以下标准：

• **可复现性**：L1 层数据需包含足够的测量元数据，使第三方能够在相同或等效条件下复现测量。

• **可追溯性**：每条数据需记录完整的采集链路（传感器→采集设备→预处理算法→存储）。

• **精度分级**：根据测量精度将数据分为不同等级（如“科研级”、“工程级”、“消费级”），应用层可根据需求选择。

• **时效性标注**：物理属性可能随时间变化（如水果的成熟度），需标注测量时间与数据有效期。结合 L1 层的版本管理，系统应能追踪同一 COID 在不同时间戳下的状态演化。

5 数据库的存储与索引架构

5.1 结构维度：分层异构存储策略

认知对象数据库的数据量极为庞大。以“苹果”为例，单个认知对象的 L1 层数据可能包含：高光谱图像（GB 级）、3D 点云（百 MB 级）、GC-MS 谱图（百 MB 级）、力学曲线（MB 级）。若覆盖 10 万种常见物体，总数据量可达 PB 级。因此，必须采用分层异构的存储策略。

数据层	数据类型	推荐存储方案	容量规划	访问模式
L1 物理锚定层	大文件（图像、点云、谱图）	对象存储（MinIO/S3）+冷热分层	PB 级	低频写入、中频读取
L2 感知编码层	高维向量（各模态编码）	向量数据库（Milvus/Pinecone）	TB 级	高频写入、高频读取
L3 语义与应用层	结构化数据（标签、关系）	图数据库（Neo4j）+关系型数据库	GB—TB 级	高频读写

多模态数据管理与分析的研究正在快速发展，涵盖多模态数据统一建模、存储与索引、查询与挖掘等全生命周期问题。面向云边端协同的多模态数据建模技术也已提出了包含六类基元的数据模型架构。这些技术进展为认知对象数据库的存储层实现提供了可借鉴的工程基础。

5.2 时间维度：分层记忆与递归状态的迁移

上述三层结构是认知对象的“结构维度”，回答的是“一个认知对象由哪些数据层构成”。但认知对象并非静态实体。苹果会成熟、腐烂、被储藏、被加工；金属会氧化；生物会生长、衰老、死亡。数据库因此还需要一个“时间维度”，回答的是“认知对象如何随时间变化、如何追踪其状态演化”。这一维度借鉴第二套方案中长程智能体的分层记忆结构[1]:

时间层	对应内容（以苹果为例）	更新策略
本底键	物种特征、遗传属性、物理常数、固有代谢规律	极稳定，跨越对象全生命周期，仅当对象被重新归类时更新
经验键	生长阶段、成熟度、储藏记录、加工工序、运输路径	阶段性更新，可追溯至关键节点，支持坍缩链式回溯
活跃键	实时温度、当前水分含量、当前成熟状态、当前损伤程度	频繁变化，每个查询点都可能不同

递归状态管理负责维护“这个认知对象现在处于什么状态”，并在状态发生关键分化时生成快照——如同长程智能体为分支宇宙保存坍缩时刻的状态变量。这一机制使数据库不仅能回答“苹果是什么”，还能回答“这个苹果现在怎么样、经历过什么、下一步可能变成什么”。

结构与时间两个维度的结合，才构成完整的认知对象：前者保证可追溯性，后者保证可演化性。没有结构维度，时间维度的状态变化缺乏可验证的物理锚点；没有时间维度，结构维度只是一个静态标本，无法支撑需要长期状态维护的智能体任务。

5.3 索引策略

认知对象数据库需支持三种类型的查询，对应三种索引策略:

① 物理特征索引（用于“找相似的物理对象”）

对 L1 层数据提取全局物理特征（如光谱特征向量、形状描述子），建立高维向量索引（如 HNSW、IVF），支持基于欧氏距离或余弦相似度的 k-NN 查询。

② 感知编码索引（用于“找感知相似的物体”）

对 L2 层感知编码向量建立独立的向量索引，支持跨模态感知相似性查询（如“找视觉

上像苹果但触觉上像橙子的物体”）。可采用多模态知识图谱表示学习方法，将不同模态的实体嵌入到统一的低维空间中。

③ 语义索引（用于“找符合语义描述的对象”）

基于 L3 层语义数据建立全文索引与图索引，支持 SPARQL-like 的结构化查询，支持自然语言查询（通过 LLM 将自然语言转译为查询语句）。

5.4 数据更新与版本管理

认知对象数据库的更新面临一个根本性挑战：物理世界的测量数据是不断积累的，而感知编码模型是持续迭代的。

L1 层：采用只追加模式。新的测量数据追加存储，已有数据不可修改。通过 COID 的内容指纹机制，相同实体的多次测量可产生不同的状态版本（如“新鲜苹果”与“腐烂苹果”属于同一 COID 下的不同状态时间戳）。

L2 层：采用可重新计算模式。当五感映射模型升级时，可批量重新计算所有认知对象的感知编码层。旧版本编码可保留作为历史参照。

L3 层：采用人工审核+众包模式。语义标签的更新需经过质量控制流程。

6 网络通信基础设施

6.1 分布式部署架构

认知对象数据库的全球部署需应对以下挑战：数据量巨大（PB—EB 级）、全球低延迟访问（<100ms）、高频并发调用（每秒百万级查询）。建议采用三层分布式架构：

层级	部署位置	存储内容	访问延迟	容量
边缘缓存层	用户就近的边缘节点	高频访问认知对象的 L3 语义层+L2 感知编码摘要	<10ms	GB—TB 级
区域中心层	各大洲数据中心	区域常用认知对象的完整三层数据	<50ms	PB 级
全局归档层	全球主数据中心	全部认知对象的完整数据（冷存储）	<200ms	EB 级

6.2 通信协议与数据格式

认知对象数据库的通信协议需满足以下要求：

- **高效传输**：采用二进制序列化格式（如 Protobuf、FlatBuffers）替代 JSON，减少传输体积。
- **流式传输**：对于大文件（如 3D 点云），支持分块传输与渐进式加载。
- **增量更新**：支持数据的增量同步，避免全量传输。
- **断点续传**：支持大文件传输的中断恢复。

6.3 安全与访问控制

认知对象数据库涉及大量真实世界的物理测量数据，可能包含敏感信息（如地理位置、个人物品特征）。安全架构需包含：

- **数据分级**：将认知对象按敏感度分级（公开、受限、机密）。
- **访问控制**：基于角色的访问控制+基于属性的访问控制。
- **隐私保护**：对涉及个人的认知对象数据进行匿名化处理。
- **审计日志**：记录所有数据访问与修改操作。

7 实施路线图

7.1 阶段划分

阶段	周期	核心任务	里程碑目标
第一阶段：数据标准制定	0—12 个月	制定认知对象数据规范（含 L1/L2/L3 三层的数据格式、元数据标准、质量分级）；设计通感联想暂存区与验证结果的数据接口	发布认知对象数据标准 v1.0
第二阶段：原型系统构建	12—36 个月	以“苹果”等 10—20 种常见物体为样本，完成三层数据的完整采集与存储；搭建四层系统架构的原型	首个跨模态认知对象完成；原型系统通过功能验证
第三阶段：规模扩展与感知	36—60 个月	扩展至万级认知对象；集成五感三层映射模型与通感模块；完成 L2	万级认知对象库上线；感知编码随模

模型迭代		层感知编码的批量重算机制	型版本可迭代
第四阶段：全球部署与生态开放	60—84个月	完成三层分布式部署；开放标准化API；建立数据质量审核与社区贡献机制	百万级认知对象库；API接口对外开放

7.2 关键里程碑

认知对象数据标准 v1.0：完成 L1 层物理数据的格式规范、L2 层感知编码的接口规范、L3 层语义数据的标注规范，以及通感关联（L2.5 层）的存储格式。

首个跨模态认知对象：完成“苹果”的完整认知对象——包含五感物理测量数据、五感预测感知编码、完整语义描述、时间维度的状态记录。

百万级认知对象库：覆盖常见自然物（动植物、矿物）、人造物（工具、食品、日用品）、抽象概念（颜色、形状、情感）。

8 总结与展望

8.1 本系列论文的理论闭环

本文的完成标志着本系列八篇论文形成了完整的理论闭环：

论文	**核心贡献**	**与前篇关系**
第一篇	诊断问题：经典 AI 构建“伪对象”的认知局限	—
第二至六篇	构建方案：五感的三层映射模型与工程实现	为认知对象提供感知数据来源
第七篇	跨模态整合：通感联想与验证机制	为认知对象提供跨模态关联
第八篇（本文）	整合方案：认知对象数据库的数据模型、存储架构与基础设施	将感知数据与通感关联沉淀为可调用的认知对象

从哲学批判到感知建模，再到跨模态整合与认知整合，本系列完成了一个从“AI 如何感知世界”到“AI 如何理解世界”的完整论述。

8.2 与显式世界模型及数据优先本体研究的对话

DaoQL（2026）提出将 LLM 视为推理引擎、把确定性知识移入显式的多模态数据库，以实现可审计的反事实推理。这与本文的物理锚定优先原则高度同构：两者都拒绝将世界知识全部隐式编码在模型参数中。本文的差异在于，它的 L2 层不是一般的多模态嵌入，而是预测的人类感知编码——知识不仅是真的，还是可体验的；它服务于与人类感知对齐的 AI，而不是纯粹的形式推理。

Puffin-World（2026）在单一框架中同时建模力（重力场）、几何（深度）与外观（RGB），实现感知、仿真、生成与重建的统一。第三套方案的三层生成与 Puffin-World 的原生 3D 世界状态在结构约束优先上一致，但第三套方案的贡献在于把结构约束按矢量逻辑、投影逻辑、位图逻辑三层显式分离。本文的认知对象数据库可以作为 Puffin-World 这类世界模型的外部显式知识层：世界模型负责生成与仿真，认知对象数据库负责存储可追溯的 L1 锚点与可调用的 L3 语义。

具身大脑与显式世界模型趋势（2026）。当前具身智能的共识是多模态感知输入+大模型/世界模型推理内核+高层规划与底层动作输出。本系列的位置是：它为多模态感知输入提供了从哲学批判到工程实验的完整工程范式，为世界模型提供了显式知识锚点（认知对象数据库），但它本身不替代具身控制系统，而是作为感知与知识层嵌入具身大脑。

在具身智能中，认知对象数据库必须作为受访问控制的显式知识层存在。每个机器人只能基于自身传感器与合法通信事件构建世界模型，任何试图越过感知接口（键）直接读取全局数据库的行为，都属于系统级违规，必须通过架构层面的白名单机制予以杜绝。

8.3 认知对象数据库的远景价值

认知对象数据库的建成将产生深远影响：

- **为 AI 提供可验证的认知基础**：AI 对“苹果”的理解不再是一个黑箱向量，而是一个可追溯到物理测量的结构化数据实体。这是根治“幻觉”问题的根本方案。

- **实现认知与推理的分离**：应用层无需在每次推理时都调用完整的感知模型，而是直接查询预计算的认知对象数据包，大幅降低算力开销。

- **支撑跨模态推理与创造**：当认知对象具备统一的多层数据结构，AI 可进行跨物种、跨物质的深度类比——这正是人类创造性思维的核心机制。

- **推动具身智能的落地**：类人形机器人可以实时查询认知对象数据库，获取对物体的多模态认知，而无需在本地存储全部数据。

8.4 结语

本系列八篇论文从对经典 AI 训练范式的深刻批判出发，逐步构建了一个从机器感知到机器认知的完整理论体系与实践蓝图。我们提出了“物理信号—神经编码—主观知觉”三层映射作为构建 AI 感官功能的通用框架，设计了“分层基准对照组”与“单变量纵深实验”的严谨实验范式，规划了五感实验室建设的完整工程方案，构建了通感联想的生成与验证机制，最终以认知对象数据库完成了从感知到认知的闭环。

这一体系的核心精神可以概括为一句话：AI 的认知必须扎根于物理世界，而非悬浮于语言符号之上。唯有如此，AI 才能真正从“处理符号的工具”进化为“理解世界的智能体”。

参考文献

[1] YunFei M. Wang 从长上下文到分层状态：面向长程智能体的设计框架[PP/OL]. AiraXiv, 2026. eprint: 2611.0018
[2] YunFei M. Wang 论经典 AI 训练方式认知局限的产生原因与破局思路——迈向基于多尺度机器感知与认知对象数据库的下一代人工智能[PP/OL]. AiraXiv, 2026. eprint: 2611.0009
[3] YunFei M. Wang 数学有效性的哲学解释：基于计算模拟视角的“键-值-信息-相”四层认知模式[预印本]. PhilPapers, 2026. 链接: https://philpapers.org/rec/WANPOT-8
[4] YunFei M. Wang AI 视觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0010
[5] YunFei M. Wang AI 听觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0011
[6] YunFei M. Wang AI 嗅觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0012
[7] YunFei M. Wang AI 味觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0013
[8] YunFei M. Wang AI 触觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0014
[9] YunFei M. Wang 论 AI 通感（机器联觉）功能的可行性构建方法——基于五感感知体系与认知对象数据库的跨模态联想与验证机制[PP/OL]. AiraXiv, 2026. eprint: 2611.0015

AIGC 使用声明

本文作者提出了论文的核心观点、理论框架、方案设计思路、全部数据模型与架构逻辑。在文本撰写过程中，作者使用了 DeepSeek AI 模型辅助进行文本结构的组织、部分章节的起草与修辞润色。作者对 AI 生成的所有内容进行了全面的审阅、批判性修改

与最终定稿，并对论文的全部学术内容、观点及结论负最终责任。