

深耕智能演进、共筑 AI 新局——AIIA 与行业同频共进
Deepen Intelligent Evolution · Build a New AI
Landscape: AIIA Advances with the Industry

中国信息通信研究院人工智能研究所 曹峰

2026年3月

~2024

扩张

技术：技术范式确立

规模定律，范式确立

应用：试点探索落地，技术适配行业场景

制造业 AI 质检 政务智能问答 金融风控 医疗辅助诊断

2025

转折

技术：核心技术进化

基础模型 → 思考模型

聊天对话 → 执行任务

闭源主导 → 开源崛起

应用：规模化深度落地，全行业提质增效
知识库、BI、代码生成、文本生成

2026~

攻坚

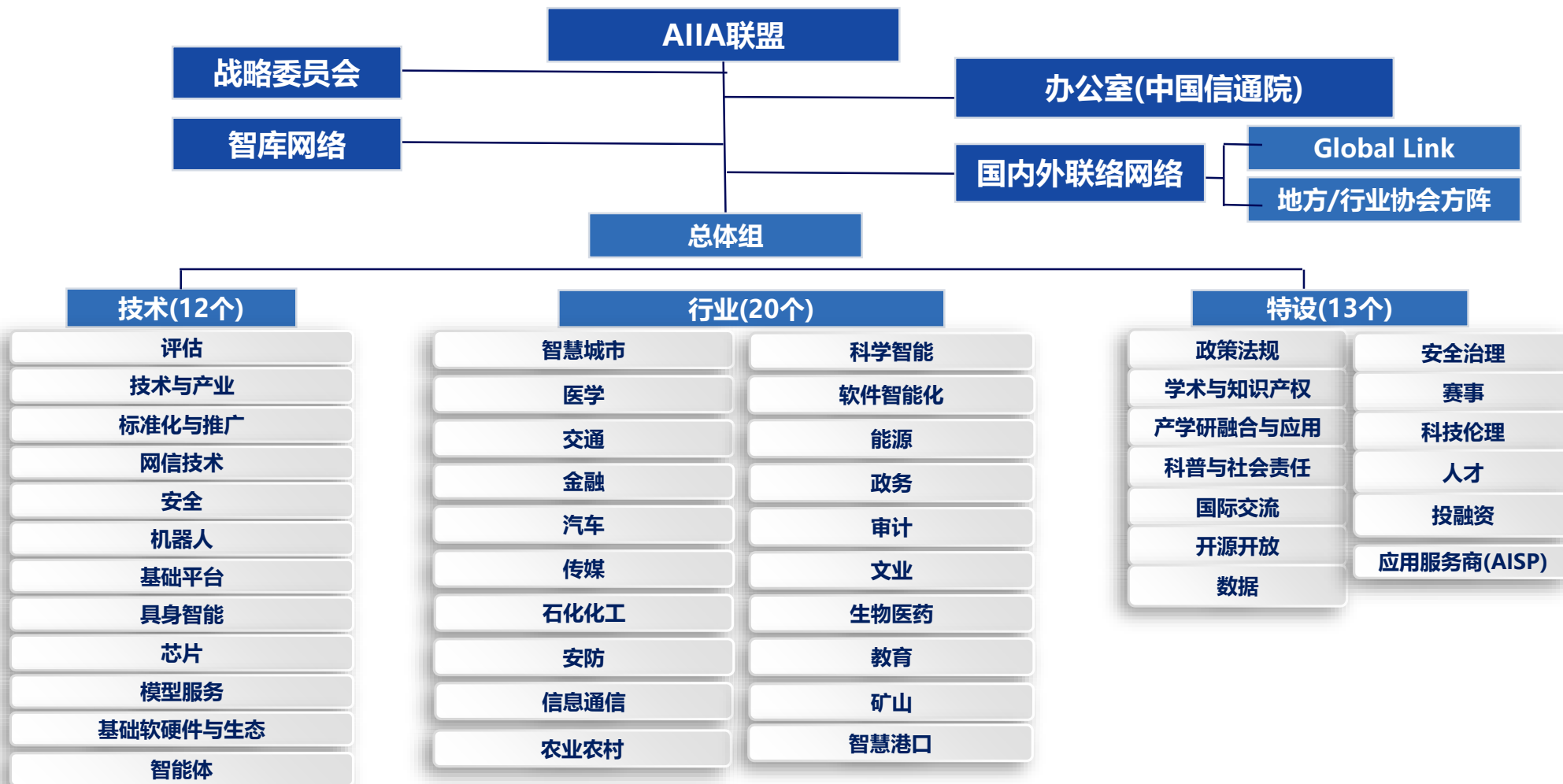
技术：继续扩张与架构创新

聚焦Agentic AI、智能体工程、Skills、AI原生等 前沿技术。

应用：智能体普及，生态迭代升级

自主决策智能体 智慧员工 科研 创新

➤ 中国人工智能产业发展联盟 (Artificial Intelligence Industry Alliance, 简称AIIA) 是在2017年由中国信息通信研究院等单位共同发起的行业合作机制。联盟目前已有1439家成员单位, 45个工作组持续推进。



定位

人工智能赋能数字经济发展的推动者、实践者

愿景

具有全球影响力的人工智能产业发展公共服务平台

任务

技术创新、应用推广、产业生态培育、国际合作

➤ AIIA联盟紧密围绕人工智能产业发展重点方向，组织联盟成员开展技术攻关、应用赋能、供需对接、标准制定和国际交流合作等工作，推动人工智能产业高质量发展。

技术创新与评测

- 建立多元评估体系，打造“**方升**”基准测试体系，构建AI数据集质量评估、具身智能、智能体等多类评估体系；
- 运行**人工智能智库网络**，开展趋势研究、技术研讨与政策支撑。

行业应用与生态培育

- 搭建**合作平台**，举办“供需对接、高水平赛事”；
- 建设**高价值场景库**，推动“人工智能+”行业落地；
- 落地**地方平台**，建设大模型公共服务平台，协助多地建数据测评中心。

安全治理与开放合作

- 推动《**人工智能安全承诺**》签署与验证，开展智能体安全标准预研；
- 组织参与国际活动如**WAIC、AI for Good全球峰会**，推动中国案例走向世界；
- 产业**协同攻关**，组织芯模联动、开源智算生态培育。



企业积极探索人工智能落地价值

落地价值

提升办公效率
知识数据洞察
生产工具赋能
管理模式革新

对内赋能

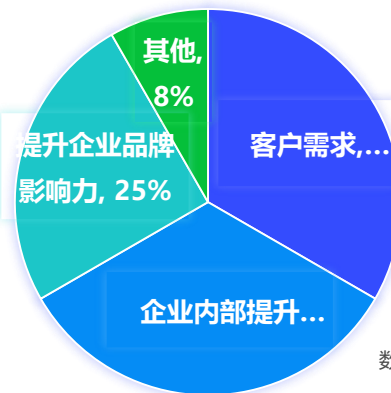


对外服务

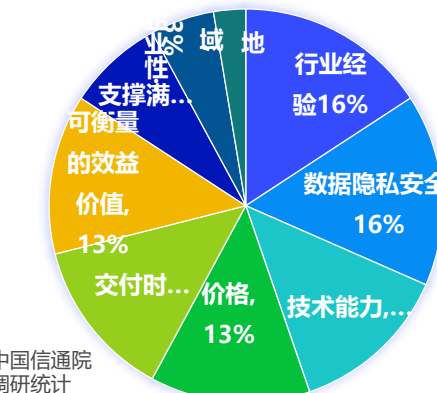
业务引导提效
产品服务升级
营销工具赋能
市场策略研判

使用大模型驱动力：满足应用需求和提升内部效率
影响采购因素：行业经验和数据隐私安全

企业使用大模型的驱动力



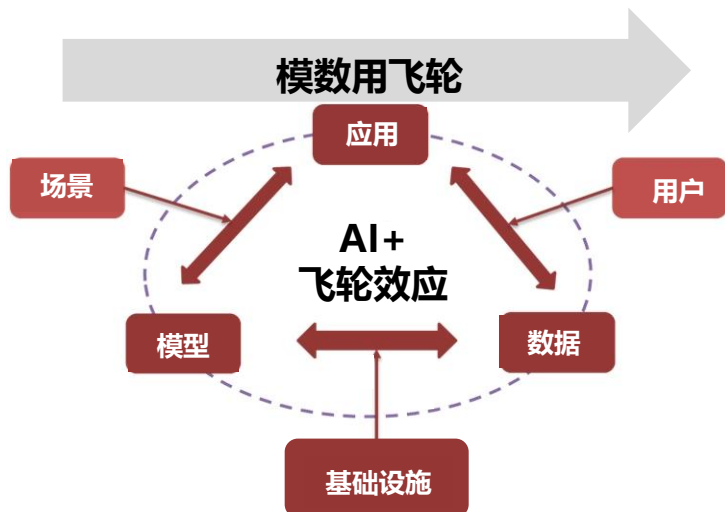
企业采购大模型的影响因素



数据来源：中国信通院
依据问卷调查统计

实现

模数用飞轮



2. 模型 -> 应用：推进模型与场景深度融合

3. 应用 -> 数据：通过应用深入挖掘高质量数据资源

支撑要素

资源向技术转化

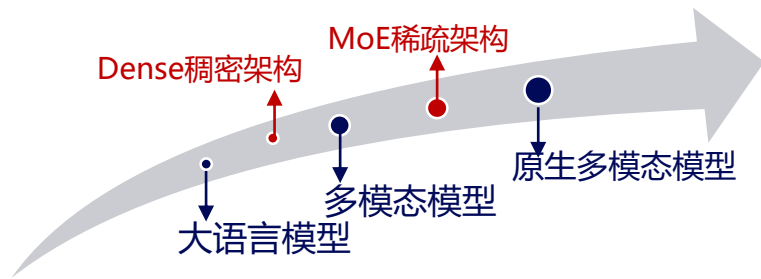
应用向资源转化

1. 数据 -> 模型：夯实资源底座，将数据资源优势转化为智能能力优势。

大模型规模化落地推动推理架构深度优化，实现“效果-性能-成本”多目标协同。

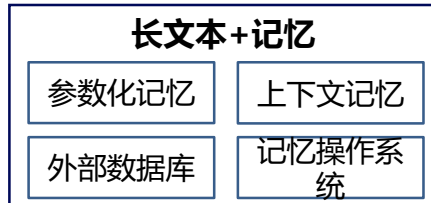
模型、应用、场景多维特性的适配挑战

① 模型：多模态、MoE成为主流



② 应用：长上下文、记忆能力的需求

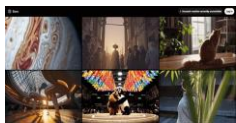
RAG、Agent服务形式不断发展，新需求爆发
→ 长上下文需求
→ 记忆需求



③ 场景：低时延、高吞吐、流量波动



低时延
智能客服、对话



高并发
数据集批量构造

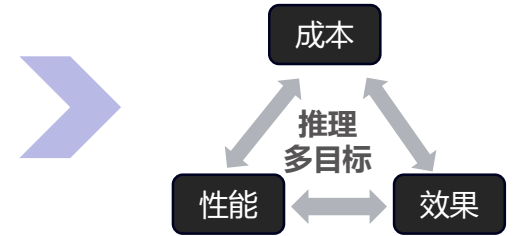


借助理推优化实现“效果-性能-成本”的多目标协同

算力与成本的博弈

现有困境：
如何在有限算力下提升性能、降低成本？

单纯堆叠硬件性能、数量
→ 加剧落地成本
→ 难以形成可持续的商业闭环



推理优化是核心抓手

单实例优化：请求高效计算

推理引擎、工具组件为载体，衔接硬件、软件与模型

模型压缩
计算优化
高性能存储

单步多次推理
批调度策略
并行策略

面向请求流量、算力资源较低的场景

多实例优化：请求与资源的分配

结合模型、资源情况、场景(长文本/时延/并发)特点等系统考虑

PD分离
AF分离
实例碎片管理

MoE集群
KV Cache池化
自适应调度策略

面向服务流量较大、算力资源较多场景

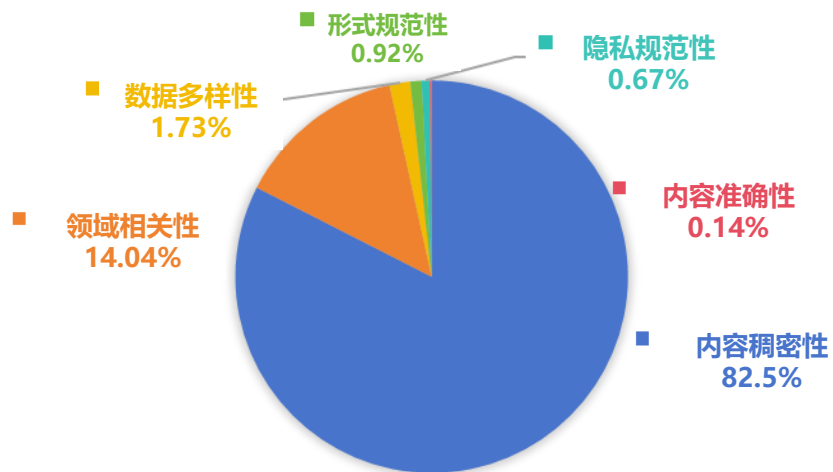
信通院相关工作

- ① 推理优化分级标准 → 优化技术指引；
- ② 推理测试床 → 水平量化定位；
- ③ 研究报告 → 系统级优化方案参考、优化路径指引。

数据集建设亟需方法论指导和技术创新，四类数据需求成为2026年重点

质量问题和建设方法论成为制约高质量数据集的关键

人工智能行业数据集建设质量问题占比统计



领域相关性

内容稠密性

数据多样性

形式规范性

从企业实践落地看，数据集建设重点从通用基础转向行业场景，数据集质量问题成为垂类模型落地和场景应用的瓶颈，三大问题制约企业推动高质量数据集建设。

目标定位模糊化

场景需求与数据集建设目标脱节，未能将数据工程目标与核心业务指标深度绑定

实施路径碎片化

数据采集到模型训练全链路缺乏系统性规划和设计，未形成体系化数据集构建和维护机制

技术底座薄弱化

缺乏适配行业特性的工具链，自动化程度低，人力依赖严重，工程落地效率受阻

未来AGI发展高度依赖于高知识密度和场景数据

① 世界模型数据需求

WorldScore理解和预测物理世界动态演化

② 具身智能数据需求

1% 真机 VS 99% 合成

③ 智能体数据需求

与测试环境深度耦合的交互数据



<https://os-world.github.io/>

④ 行业模型数据需求

- 医疗：诊断过程思维链数据
- 金融：风控和客户画像数据
- 交通：交通流和检测数据
- 工业：工艺知识、生产流程数据
- 能源：新能源场站、电力巡检数据

推动高质量数据集建设

① AI-READY 的企业数据工程

体系建设、开发维护、质量控制、资源运营、合规可信的全流程闭环

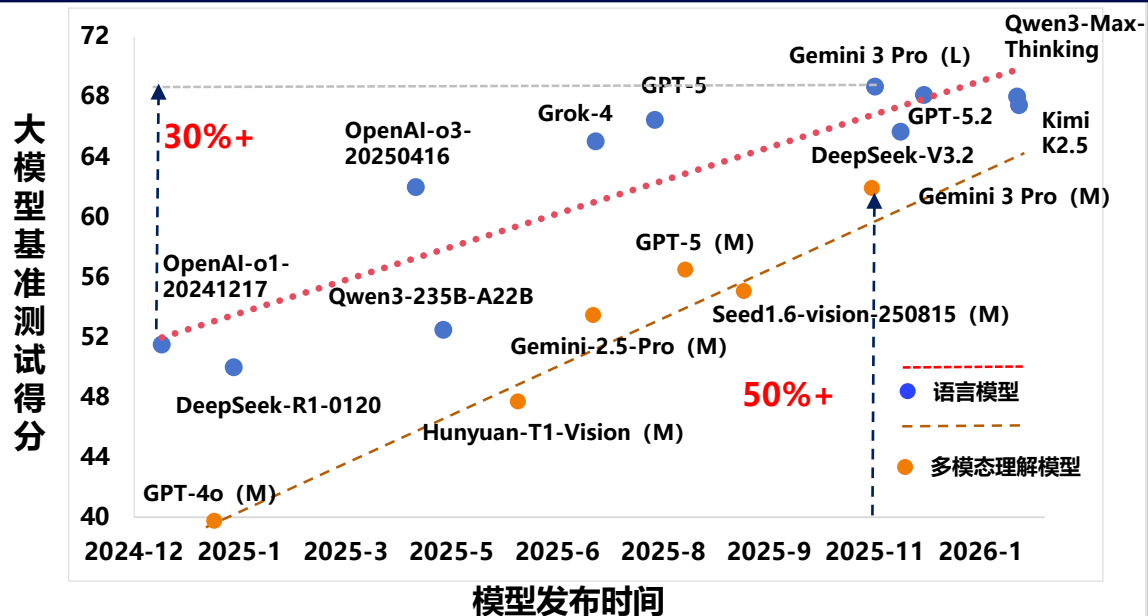
② 推动前沿数据技术攻关

鼓励企业联合高校、科研院所开展合成数据工业化生产、数据增强扩充、全链路质量控制、异构数据等技术突破

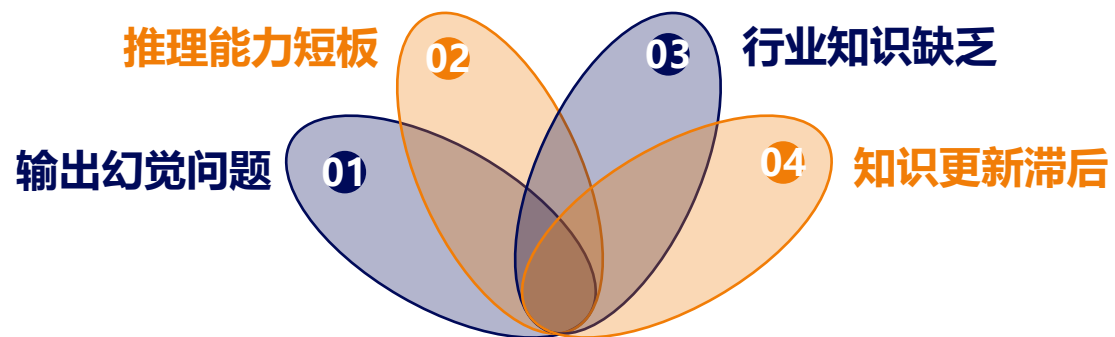
专精模型：大模型能力提升显著，行业应用问题仍需重点突破

➤ 2025年，全球语言和多模态大模型能力提升超过30%，行业应用仍存短板，测试是驱动模型行业应用的重要手段。

语言和多模态模型进步明显

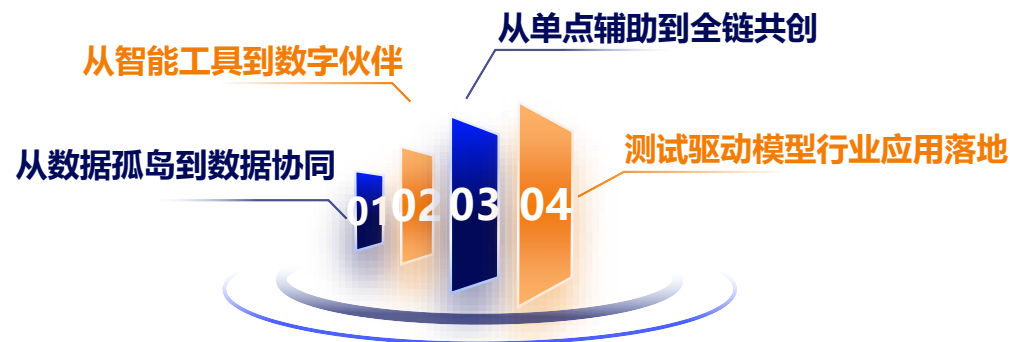


模型行业应用面临的问题及挑战



大模型基准测试体系助力模型应用落地

模型行业应用未来发展趋势



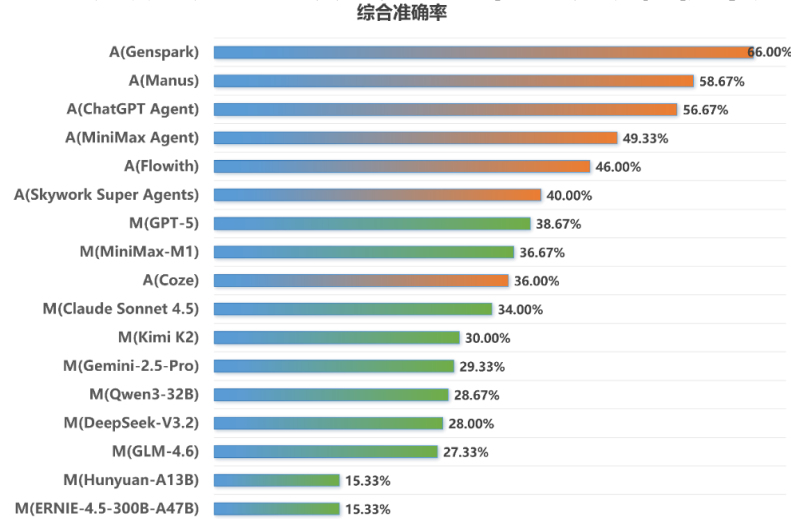
“方升”大模型基准测试体系

测试标准 **测试数据** **测试技术** **测试工具**

➤ 智能体是在大模型基础上的工程化增强，规模化应用仍需克服多重挑战。

智能体自主性不断提高，应用落地仍存在挑战

- 智能体可以获得比单一模型更好的性能表现，仍不及预期。



注：A代表智能体产品，M代表大模型；数据来源：中国信通院“方升”智能体基准测试，2026年2月

- 智能体技术能力和工程架构仍存在一定挑战。

技术能力方面

模型能力不足
复杂长程任务能力差
多智能体协同能力弱

工程架构方面

Skills、MCP、记忆等技术演进快
智能体工程架构复杂
记忆与知识管理瓶颈
工具调用效率与准确性仍有待提高

亟需从单点智能体应用向企业级智能体系统演进

- 企业级智能体架构推动单点智能体向复杂智能体系统演进

- 与企业原有IT、数据基础设施融合

- 与企业ERP、CRM等业务系统融合

- 与企业采购、研发、运维等 workflows 融合

- 中国信通院围绕智能体开展一系列研究工作。

标准研制

编制智能体标准体系

研制国际标准与行业标准

智能体测试体系

基础能力测试

通用任务测试

典型应用场景测试

智能体测试床

MCP协议研制测试

单智能体能力测试

多智能体协同测试

➤ 智能原生有望打破常规推动应用范式变革，从概念走向现实，亟需推动AI原生标准规范、测试能力建设，助力AI原生技术与产业生态圈建设。

智能原生重塑产品服务与企业组织模式

- 智能原生：以人工智能技术作为核心引擎，在技术、产品、商业模式和经营管理等方面突破常规、全面重构。

智能原生的特征

- 形成高效运转的“模型-场景-数据”闭环，支撑进化。



- 成就超级个体，实现超高人效，切入细分场景。

- 创业即盈利，敏捷迭代，持续进化，陡峭成长。

驾驭不确定性
关键：为智能原生提供包容性环境

从概念走向现实，AI原生产业实践面临挑战

高预期 vs. 低ROI

预期落差

定价困境

模型万能论 vs. 工程现实

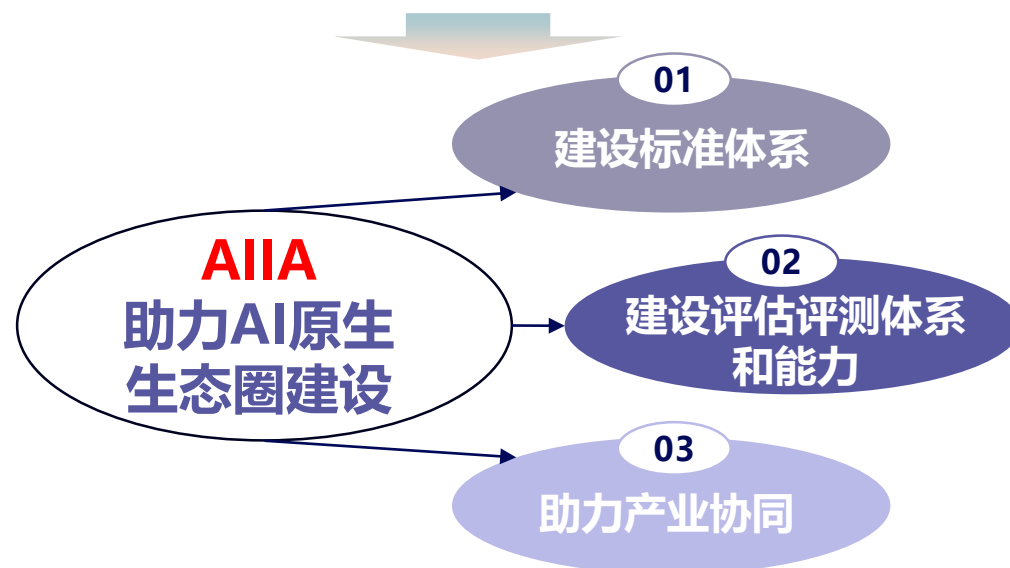
AI落地的“二八定律”
工程化能力缺失

数据驱动 vs. 数据贫瘠

数据资产化程度低
数据孤岛与流通壁垒

技术创新 vs. 安全伦理

Agent的不可预测性
自主智能对价值对齐需求迫切



向上突破——技术摸高

- **继续扩张：**更紧密软硬协同+新的模型架构
- **持续学习：**长期记忆能力为持续学习创造条件
- **世界模型：**智能体和具身智能进化的共性需求
- **经验时代：**在交互中强化学习的重要性凸显

向下扎根——应用落地

- **敏捷迭代：**适应新常态，在快速变化中获取价值
- **飞轮效应：**“场景-应用-数据-成效”需要闭环
- **长期规划：**形成体系化的企业智能化道路
- **生态成型：**建设以开源开放为特征的新型智算生态



感谢聆听！
THANKS

