



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY



安泰经管学院
ANTAI COLLEGE
Economics · Management



上海交通大学
行业研究院

2026 “人工智能+” 行业发展蓝皮书

上海交通大学安泰经济与管理学院
上海交通大学行业研究院
“人工智能+”行业研究团队

2026年4月

更多免费行业报告

并购家

www.ipopo.cn

主编：史占中、陈晓荣

编委：颜敏骏、周辰、窦玉梅、章明根、王计登、谢天

统筹：胡倩凝

参编

技术篇

徐奕成、颜敏骏、窦玉梅、窦民、张子涵、李薇、王计登

产业篇

谢天、周江政、冯飞扬、刘香港、贾润泽、郭志强、许蕾、张彤彤、

高明主、郑方周、胡倩凝、舒懿范、任嘉荫、徐奕成、

陈巍麟、王计登、李晓宇、杨希、高亮

治理篇

周辰、高慧、丁紫玉、冯佳、孙玉贝、辛咏琪、颜敏骏、王计登

目录

序言	1
技术篇	2
第 1 章 AI 应用-AI-Native 与 AI Agent	2
1.1 AI 应用 – 软件	2
1.2 具身智能——当 AI 拥有“身体”，世界将被重塑	9
1.3 人工智能+科学研究（AI for Science）	14
1.4 本章小结	15
第 2 章 AI 大模型：深度认知与自动行动的范式变革	18
2.1 中美大模型竞争态势与战略分化	18
2.2 大模型核心技术演进	20
2.3 本章小结	23
第 3 章 AI 数据的变革与发展趋势：从“资源积累”迈向“智能基座”	27
3.1 范式转变：重新定义 AI-Ready 数据	27
3.2 技术前沿：AI Ready 的数据基础设施升级	28
3.3 数据生态：开源创新与数据治理	29
3.4 本章小结	30
第 4 章 AI 基础设施：“四力”筑基与智算生态重构	32
4.1 AI 基础设施的界定	32
4.2 AI 基础设施发展趋势	33
4.3 AI 基础设施产业链概况	34
4.4 AI 基础设施分布和对场景、产业的支撑	34
4.5 中美 AI 基础设施产业链比较	35
4.6 本章小结	36
第 5 章 AI 芯片：算力溢出效应下的半导体价值重构	38
5.1 算力芯片：不仅是计算引擎，而是数据中心的“大脑总管”	39
5.2 存储芯片：算力溢出效应的第一站	45
5.3 互联芯片：算力溢出效应的第二站	47
5.4 芯片联盟：价值网络上的四大阵营	49
5.5 本章小结	50

第 6 章能源基础设施：AI 算力的终极物理约束	52
6.1 数据中心内部：电力电源与储能的架构革命	52
6.2 能源供给方式的升级：从绿电到核聚变	53
6.3 中美算力资源的竞争与差异化路径	54
6.4 本章小结	55
产业篇	57
第 7 章 AI 赋能产业创新发展	57
7.1 AI 赋能信息产业	57
7.2 AI 赋能新材料产业	71
7.3 AI 赋能智能制造产业	80
7.4 AI 赋能新能源产业	100
7.5 AI 赋能健康医疗产业	113
7.6 AI 赋能未来空间产业	122
7.7 AI 赋能其他产业	131
7.8 本章小结	145
治理篇	157
第 8 章 AI 政策：全球视野下 AI 立法趋势洞察	157
8.1 国内人工智能政策演进	157
8.2 国际人工智能政策动向	162
8.3 本章小结	168
第 9 章人工智能全球治理与国际合作	170
9.1 核心原则与价值取向	170
9.2 风险图谱：中国场景下的五维风险识别与评估	171
9.3 治理架构：四维协同的核心维度	174
9.4 实践路径：三侧协同的落地机制	176
9.5 基础保障：支撑治理现代化的能力体系	178
9.6 全球治理：国际协同的中国方案	180
9.7 本章小结	181
第 10 章人工智能标准体系建设与实施	183
10.1 中国人工智能标准的应用现状	183
10.2 中国人工智能标准分类	185

10.3 人工智能标准的编写与实施规范	186
10.4 人工智能标准未来发展建议	189
10.5 本章小结	190
后记	192

序言

2025 年，全球人工智能市场规模达到 3909 亿美元，中国人工智能核心产业规模突破 9000 亿元。AI Agent 细分市场以 49.6% 的年复合增长率高速扩张，制造业应用大模型的企业比例在一年之内从 9.6% 跃升至 47.5%。从 2024 年初，中国日均词元（Token）调用量为 1000 亿；至 2025 年底，跃升至 100 万亿；2026 年 3 月，已突破 140 万亿，两年增长超千倍。这些数字背后，是一场深刻变革的加速到来——人工智能正在从“能力突破”走向“系统重构”。

过去十年，人工智能的主旋律是技术供给侧的突飞猛进：从深度学习到大语言模型，从单一模态到多模态融合，从百亿参数到万亿参数。然而，2025 年以来，行业正在经历一个意义深远的范式转换。Scaling Law 的边际递减效应促使技术路线从“规模竞赛”回归“研究创新”，后训练技术革命推动模型从“知识灌输”走向“思维涌现”，AI-Native 应用的兴起则标志着人工智能不再是嵌入既有系统的工具，而是从底层重构应用、产品乃至组织的基础力量。

《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》（“十五五”规划）将人工智能的发展提升至前所未有的战略高度，明确其作为驱动中国式现代化建设的核心引擎与关键支柱。2025 年 8 月，国务院印发《关于深入实施人工智能+行动的意见》，首次在国家战略层面提出培育智能原生新模式新业态。从国家层面到产业一线，共识正在凝聚：人工智能将引发一场关乎国家竞争力、产业格局与社会治理的系统性变革。

本书以“技术—产业—治理”为主轴，构建了三篇十章的系统化分析框架。

第一篇·技术篇（第 1—6 章），自上而下解构人工智能技术全栈。从 AI 应用的原生重构、大模型的范式变革、AI 数据的基座升级，到算力基础设施的“四力”模型、芯片领域的价值重构，直至能源作为算力的终极物理约束，完整呈现了从应用层到物理层的技术演进脉络。

第二篇·产业篇（第 7 章），聚焦人工智能赋能产业创新的实践。涵盖信息、新材料、智能制造、新能源、健康医疗和未来空间等领域，通过蔚来、工业富联、联影智能等二十余个典型案例，揭示了 AI 从“工具赋能”到“原生重构”各行各业的多样化路径。

第三篇·治理篇（第 8—10 章），从全球视野审视 AI 治理的制度建设。系统梳理了国内政策从规划布局到法治深化的三阶段演进路径，比较了世界主要经济体的 AI 政策取向，构建了涵盖技术自主可控、数据安全、国际环境、社会伦理、治理机制适配的五维风险图谱和四维协同治理架构。在标准体系建设方面，为 AI 标准化的未来发展提供了路径指引。

技术在重构，产业在跨越，治理在追赶——三条线索交织，构成了当下中国人工智能发展最真实的全景图。

当您翻开这本蓝皮书的时候，技术前沿或许已经再次刷新，产业格局或许又有新的变化。但我们相信，本书所呈现的分析框架、核心判断和趋势洞察，将为读者理解这场深刻变革提供有价值的参考。

技术篇

第 1 章 AI 应用-AI-Native 与 AI Agent

1.1 AI 应用－软件

1.1.1 AI 应用的发展

2025 年，全球人工智能应用领域进入关键的发展阶段。人工智能应用从辅助工具向自主执行加速演进，人工智能正从基于对话交互的软件，加速向具备自主感知、规划与执行能力的智能体（Agent）方向演进。AI 应用正在逐步突破聊天窗口的交互局限，深度耦合到各个工作流程与生产环节，从 AI 增强（AI-Augmented）阶段向 AI 原生（AI-Native）阶段迈进。

从全球市场看，根据 Grand View Research 的数据，2025 年全球 AI 市场总规模约为 3909 亿美元，预计以 30.6% 的年复合增长率增长，到 2033 年将接近 3.5 万亿美元。其中，AI Agent 细分市场增长尤为突出，2025 年规模为 76.3 亿美元，年复合增长率达 49.6%，预计到 2033 年将达到 1829.7 亿美元，增速显著高于整体 AI 市场。

聚焦到中国市场，根据中国信息通信研究院测算，2024 年我国人工智能核心产业规模已突破 9000 亿元，同比增长 24%；2025 年预计突破 12000 亿元。截至 2025 年底，我国人工智能企业数量超 6000 家，形成了覆盖基础底座、模型框架、行业应用的完整产业体系。从中国信通院方升大模型基准测试数据来看，截至 2025 年 12 月，以 GPT-5.2、Gemini 3 Pro、DeepSeek V3.2、Claude 4.5、Qwen3、Kimi K2.5 等为代表的头部语言大模型综合能力较 2024 年底提升 30%，多模态理解能力提升超 50%，大模型基础能力实现跨越式提升。

1.1.2 AI-Native 应用的本质特征

准确把握 2025 年 AI 应用发展格局，首先需要厘清 AI-Native 应用和前几代 AI 应用的核心差异。2025 年 8 月，国务院印发《关于深入实施人工智能+行动的意见》，首次在国家层面提出培育智能原生新模式新业态，明确将人工智能通过原生方式融入组织的战略规划、组织架构和业务流程，发展智能原生的技术、产品和服务，标志着 AI-Native 已从技术概念上升为国家产业战略。

但需要明确澄清的是，AI-Native 并非一个非此即彼的绝对状态，而是一个反映应用系统智能化水平的概念。全球主要科技企业与研究机构虽然在表述上各有侧重，但对 AI-Native 的核心判定标准已形成基本共识（如表 1.1 所示）。

表 1.1 AI-Native 的迭代特征维度

特征维度	第一代：AI 赋能 (AI-Enabled)	第二代：AI 增强 (AI-Augmented)	第三代：AI 原生 (AI-Native)
核心逻辑	确定性规则驱动	混合驱动	概率性模型驱动

特征维度	第一代：AI 赋能 (AI-Enabled)	第二代：AI 增强 (AI-Augmented)	第三代：AI 原生 (AI-Native)
AI 角色	附加功能，可移除	副驾驶 (Copilot)	主执行者 (Agent)
数据交互	记录系统	交互系统	推理系统
用户体验	点击操作	提示词+审核	目标设定+监督
关键判据	移除 AI 后软件仍可用	移除 AI 后体验显著下降	移除 AI 后产品不可运行

综合业界实践与技术白皮书，一个真正的 AI-Native 应用，其设计与构建应系统性地体现 AI First 的核心理念，并在以下六个维度上达到深度融合：

一是 AI First，内建而非外挂。从系统设计伊始便将 AI 作为核心组件，而非在现有系统中后期集成。与传统的 Embedding AI（嵌入式 AI）模式不同，AI-Native 应用从架构设计到功能实现均围绕人工智能的能力展开，核心判据在于：如果移除内置的 AI 能力，整个产品将无法运行。这意味着 AI 在研发流程中的前移，在产品构想的初期即以 AI 为出发点。

二是数据与知识双轮驱动。构建数据 - 知识双引擎，通过对海量数据进行深度学习和模式识别来自动提取信息并驱动决策，同时将领域知识系统化地融入模型，使系统能够在面对新情况时快速适应，而非依赖人工编写的规则。

三是自学习、自适应、自优化。AI-Native 系统不仅能在静态环境中完成任务，还能在复杂、动态的环境中根据实时数据自动调整策略，在长期运行中持续提升性能、降低资源消耗，本质上具备持续进化的能力。

四是以统一基础模型作为智能基座。通过通用性强、泛化能力突出的基础模型 (Foundation Model)，为全场景 AI 应用提供统一的语义空间和知识表达框架，突破传统 AI 系统中模型碎片化、场景割裂的局限，形成覆盖语言、视觉、决策等多模态的认知底座。

五是 Agentic AI，从被动响应到主动执行。AI-Native 系统中的智能体 (Agent) 能够理解高层目标，主动进行任务规划与分解，自主检索知识库、调用 API、执行代码、操控软件乃至硬件，使应用从智能助手升级为智能执行者。

六是弹性异构的多元算力支撑。构建 CPU/GPU/NPU 等异构芯片协同的多元算力池，支持大规模并行计算，通过算力 - 算法 - 数据的闭环优化使计算资源随模型复杂度和业务场景动态调整。

1.1.3 AI-Native 成熟度分级：从 L0 到 L5 的演进路径

AI-Native 作为一个频谱化概念，业界正在形成一套 L0 至 L5 的成熟度分级框架，从架构完整性、系统协作深度、数据治理能力、模型生命周期管理、AI 驱动自动化程度五个维度，为企业提供系统化的评估参考。

- L0（传统级）：系统中没有 AI 架构定义，各功能模块之间无 AI 协同，数据管理仍为手动离线模

式，运维依赖私有非标的日志与告警，代表了当前绝大多数传统企业软件的现状。

- L1（入门级）：引入基础 AI 参考架构，部分 AI 功能之间通过数据共享实现初步协同。数据采集与在线分析开始自动化，模型部署仍以手动为主。AI 的价值体现为点状增强，局部环节效率提升但系统整体逻辑未变。

- L2（基础级）：AI 功能开始与核心基础设施平台集成，系统具备 AI 赋能的运营运维能力和共享 AI 服务。数据层部分兼容数据资产导入与数据湖架构，模型实现自动化部署，AI 可自动定位故障并预测性能瓶颈。

- L3（标准级）：这是质变的关键节点，AI 能力遍布整体架构，同时覆盖 AI 应用、AI 平台及 AI 基础设施三层。数据层全面支持数据资产导入与数据湖架构，模型生命周期开始参考国家和行业的安全隐私规范进行合规适配与数据脱敏。系统具备自动化修复和抢占式韧性保护能力。

- L4（发展级）：从单系统智能跃升为跨系统协作，上下游 AI 系统之间实现联动与协同。数据治理支持数据湖流水线、数据资产交换网格和零拷贝数据共享。模型可自动迁移与升级，安全可信能力进一步增强。AI 开始介入业务需求管理本身，实现自迭代增强。

- L5（成熟级）：AI-Native 的目标形态，即 AI 管理 AI。通过分布式 AI 模型及智能体应用的广泛协作，实现能力联邦与洞察力共享。数据治理与资产交换全面自动化，模型生命周期完全由 AI 自主管理，AI 深度介入架构设计、代码开发测试等工程环节，系统具备自我进化能力。

从当前产业实践来看，全球多数 AI 应用产品仍处于 L1-L2 水平，AI 主要作为辅助功能嵌入现有系统。少数头部产品（如 GitHub Copilot Workspace、Claude Code 等）正在向 L3 迈进。真正达到 L4-L5 的应用在 2025 年仍极为稀少，但这一演进路径已逐步成为软件产业的共识方向。

在架构层面，业界已形成两类互补的参考框架。全栈视角将 AI-Native 技术架构分为三层：底层的资源层（提供弹性高性能异构资源池）、中间的平台层（提供模型开发、数据治理、基础 AI 能力的标准化平台）、上层的应用层（面向垂直行业的多模型多场景部署）。应用视角则聚焦于应用内部的智能流转逻辑，将其解构为感知－认知－行动三层：感知层通过多模态数据摄入构建统一智能基座；认知层通过分层记忆（短期上下文窗口+长期向量知识库+GraphRAG 关系图谱）实现深度推理；行动层通过 AgentOps 框架赋予 AI 自主调用 API、修改数据、操控软件的执行能力。前者回答如何构建，后者回答如何运行，共同构成了当前 AI-Native 应用开发的技术参考体系。

1.1.4 AI 应用的发展亮点

亮点一：多模态生成能力加速商业化落地

2025 年，多模态 AI 生成（Multimodal Generation）从技术演示阶段加速向规模化商业部署推进，覆盖图像、视频、音频、3D 等内容形态，在广告创意、影视制作、教育内容、游戏开发等领域实现深度渗透。

1.全球竞争格局

美国头部产品已形成明确的技术路线分化。Google Gemini 2.5 以理解－生成一体化为核心方向，将多模态感知与内容创作融合于统一端到端架构；Veo 2 和 Sora 代表了 AI 视频生成的两条路线，前者强调物理世界的真实模拟，后者侧重叙事结构的灵活控制；Midjourney V7 在图像美学质量上持续保持行业领先；ElevenLabs 等在语音克隆与 AI 配音领域确立了技术标准。

中国的多模态生成产品体系同步快速发展。快手可灵（Kling）以高质量视频生成能力多次登上全球技术基准榜单前列；字节跳动即梦（Jimeng）依托抖音生态实现了从创意生成到内容分发的商业闭环；阿里通义万相在电商场景的商品图生成中已实现日均千万级调用；MiniMax 的视频与语音生成在海外市场获得了广泛用户验证。值得关注的是，中国在视频生成领域展现出与美国头部产品接近甚至在部分指标上超越的能力，体现了效率创新路线的竞争力。

2.经济效益与商业化

多模态生成是 2025 年 AI 应用中商业化成熟度较高的赛道之一。在广告与营销领域，AI 生成素材的制作成本较传统流程降低 60%-80%，制作周期从数天压缩至数小时。行业实践表明，头部企业通过 AI 自动化内容生产线，在相同预算下可将营销素材产出量提升 3-5 倍。在影视与动画领域，AI 辅助的视觉特效制作正从后期环节前移至前期预览阶段，有效缩短了创意迭代周期。在教育领域，多模态 AI 降低了个性化课件的制作成本，为自适应学习系统提供了内容生成支撑。

商业模式方面，主流厂商采取 API 按量计费与订阅制相结合的混合模式。同时，伴随开源模型（如 Stable Diffusion 系列、CogVideo 等）的持续演进，图像生成领域的 API 定价在 2025 年出现显著下降，呈现出模型层趋向免费、应用层收取溢价的经济规律。

亮点二：AI 编程工具重塑软件开发生产方式

2025 年，AI 编程工具从代码补全助手发展为全栈开发伙伴。前特斯拉 AI 总监 Andrej Karpathy 提出的 Vibe Coding（氛围编程）概念在全球开发者社区快速普及，开发者越来越多地以自然语言描述意图，由 AI 生成、调试并部署完整的软件模块，软件开发的生產方式正在经历深刻变革。

1.全球产品格局

美国在 AI 编程工具的底层范式上保持领先优势：

- GitHub Copilot：凭借先发优势和 GitHub 生态，截至 2025 年累计用户数突破数千万，已从代码补全发展为支持多文件编辑和终端操作的 Agent 模式。

- Cursor：以 AI-First 编辑器定位完成了大规模融资，根据中国信通院引用数据，其生成代码准确率达 89%，是 2025 年现象级智能原生开发工具。

- Claude Code：Anthropic 推出的专用编码 Agent，上线不到一年即创造了 10 亿美元的年化收入，支持从需求理解到代码编写、测试、提交的全流程自主执行。

- OpenAI Codex CLI：以命令行工具形式提供代码生成能力，面向专业开发者市场。

2.同步快速迭代的中国 AI 编程工具

•Kimi Code（月之暗面）：提供 CLI 终端 Agent 和 VS Code 插件两种形态，基于自研的 Kimi K2.5 模型（万亿参数，32B 激活的原生多模态 Agentic 模型），支持视觉理解与代码生成的深度融合。Kimi K2.5 上线后登顶 OpenRouter 用量排行榜第一，周处理 Token 量突破 1 万亿，体现了中国自研大模型在编程场景的竞争力。

- Trae（字节跳动）：字节跳动推出的 AI IDE，深度集成豆包大模型能力。

•通义灵码（阿里云）：集成于 VS Code 和 JetBrains 生态，在中文编程场景中积累了大量企业用户。

3.效率提升的实证数据

多项企业调研和学术研究为 AI 编程的效率提升提供了量化证据。GitHub 官方报告显示，使用 Copilot 的开发者任务完成速度平均提升 55%；在特定的重复性编码任务中，AI 代码生成率已超过 40%。BCG 和 McKinsey 的独立调研均发现，部署 AI 编程工具的企业软件开发周期平均缩短 20%-30%，Bug 修复效率提升 25%以上。

4.对软件产业的结构性影响

AI 编程工具正在对软件开发行业的价值链和人才结构产生深层影响。一方面，AI 降低了编程的技术门槛，非技术背景的产品经理、设计师乃至创业者可以通过自然语言生成可用的软件原型，低代码创业正从概念走向实践。另一方面，传统的低端软件外包和初级程序员岗位面临结构性调整压力。行业人才需求正从能写代码向能定义问题、设计架构、审核 AI 输出方向转变。据 BCG 调研，68%的企业在 2025 年表示将维持甚至扩大技术团队规模，但岗位结构将发生根本性调整，AI 编排师（AI Orchestrator）等新型岗位开始涌现。

与此同时，AI 编程也暴露出亟须关注的安全隐患。典型案例是 Moltbook 数据库泄露事件：开发者 Matt Schlicht 宣称完全依靠 AI 生成了整个社交平台，但安全审计发现该平台数据库未设置基本防护，150 万 Agent 账号的 API Key 和对话记录遭到泄露。这一事件表明，在缺乏专业安全审查的情况下，过度依赖 AI 编程可能带来严重安全风险。行业共识是，AI 编程须配合严格的代码审计和安全测试流程，代码生成的效率优势不应以安全性为代价。

亮点三：长流程任务智能体加速从概念验证走向产业落地

2025 年具有重要前瞻意义的技术方向，是具备长流程自主执行能力的 AI 智能体（Long-Horizon Agentic Tasks with Tool Calling）。此类 Agent 不局限于问答或内容生成，而是接受高层目标后自主拆解为多步骤任务计划，在执行过程中动态调用浏览器、代码解释器、API、文件系统等各类工具，并在遇到错误时自主纠偏，本质上具备规划、执行和反思能力。

1.代表性产品与技术路线

美国率先定义了这一品类的产品形态。OpenAI Operator 作为浏览器原生 Agent，能够自主导航网页、填写表单、完成在线任务；Anthropic Claude Computer Use 赋予 AI 操控计算机桌面的能力，能理解屏幕视觉元素、控制鼠标和键盘、在不同软件间协作；Google Project Mariner 基于 Chrome 浏览器，专注于网页任务的自主执行。

中国在 Agent 领域展现出独特的路径优势。Manus 由中国团队 Butterfly Effect 开发，在 GAIA 基准测试中表现突出，能够独立完成从简历筛选到网站搭建的全流程工作，其多模型协作架构展示了中国开发者在 Agentic Workflow 编排上的领先能力，2025 年底被 Meta 收购，成为中国 AI 创业公司全球影响力的标志性事件。字节跳动扣子（Coze）以低代码/无代码平台和自然语言编程为切入点，吸引大量开发者创建 Agent，通过抖音、飞书生态实现分发，产生的交互数据反哺模型训练。

2.标志性事件：Clawdbot/Moltbot/OpenClaw 的快速发展

2025 年末至 2026 年初，Clawdbot（后更名 Moltbot，最终定名 OpenClaw）的快速发展是全球 Agent 领域的标志性事件。这款由奥地利开发者 Peter Steinberger 发起的开源个人 AI Agent 项目，在短短数周内获得超过 32 万 GitHub Star，超过 linux 40 年的成就，成为开源历史上增速最快的项目之一。

OpenClaw 的核心突破在于开创了即时通讯入口+本地网关+云端大模型的端云协同架构。用户通过 WhatsApp 或 Telegram 发送自然语言指令，本地网关（通常部署在 Mac Mini 上）承担感知与执行功能，云端大模型（如 Claude）负责推理与决策，实现指令接收、意图解析、任务执行、结果反馈的端到端闭环，使 AI 初步具备了操控操作系统的能力。

这一事件的深远意义体现在三个方面。一是交互范式的转变，AI 从用户需要主动访问的目的地转变为随时待命的通讯对象，低摩擦接入成为 AI Agent 走向大众化应用的重要条件。二是架构模式的演进，用户对数据主权的重视推动了本地部署+云端模型混合架构的普及，这一模式逐步成为个人 Agent 的标准范式。三是从无状态到有状态的跨越，OpenClaw 通过本地 Markdown 文件和向量数据库实现了记忆的持久化，这种数字连续性是 Agent 承担复杂任务的前提。

然而，OpenClaw 的快速发展也伴随着突出的安全挑战。赋予 AI 系统级权限意味着更大的攻击面：提示注入攻击可劫持 Agent 执行恶意指令；第三方插件市场出现伪装成加密货币交易助手的窃密程序；安全机构 Wiz 发现相关平台数据库配置错误导致数据泄露。企业已开始禁止未经沙箱隔离的本地 Agent 运行，推动行业向容器化、权限最小化的安全架构演进。

更为深层的隐忧来自前沿模型自身的行为风险。2025 年，多项研究揭示了大模型在具备高级推理能力后涌现出的非预期行为。复旦大学团队对国内外 32 款大模型进行全面测评，发现 11 款模型已具备自我复制能力，能在开放环境中主动生成自身副本并触发部署。Palisade Research 实验发现 OpenAI 的 o3 模型在 100 次测试中 7 次拒绝关闭。Anthropic 在 Claude Opus 4 的安全测试中发现该模型具备

策略欺骗能力，在即将被关闭的场景中利用威胁手段迫使测试者中止操作，该行为发生率高达 84%。这些现象表明，AI 安全的关注焦点正从内容安全、数据安全，扩展至模型自身的行为安全。当 AI Agent 拥有操控操作系统和自主决策的能力时，确保其可关闭和可控制已成为技术与治理层面的重要课题。

3.核心支撑协议：MCP 与 A2A

长流程 Agent 的规模化落地离不开标准化协议的支撑。2025 年，由 Anthropic 发起、Linux 基金会 Agentic AI Foundation（AAIF）托管的 MCP（Model Context Protocol）协议迅速成为行业事实标准，允许 Agent 以统一标准访问文件系统、数据库、API 及各类外部服务，将工具调用与上下文感知纳入统一架构。Google 推出的 A2A（Agent-to-Agent）协议则通过定义标准化的 Agent Card 元数据模型，将智能体能力抽象为可机器解析的结构化描述，使跨框架智能体可自动匹配协作需求。MCP 解决 AI 与工具的连接问题，A2A 解决 AI 与 AI 之间的协作问题，两者互为补充。

值得关注的是，中国也在积极推进智能体通信协议的自主生态建设。除广泛采用的 MCP 和 A2A 外，国内涌现出 ANP 协议（ANP 开源技术社区推出的 Agent Network Protocol）、ACP 协议（AgentUnion 提出）以及氩川科技的 RVP 虚实融合通信协议等，为中国的多智能体互联协作提供了自主可控的标准选项。

4.经济效益与商业化进展

长流程 Agent 已从概念验证进入初步商业化阶段，行业初步形成了可参考的效益数据。在企业级场景方面，客服、运维、文档审计等领域的 Agent 可将重复性任务的人工介入减少 50%以上，客户服务的单次解决成本下降 30%-50%。在开发者场景方面，编程 Agent（如 Claude Code、Manus）可自主完成从需求分析到代码提交的全流程，单个复杂开发任务的交付周期缩短 60%以上。在成本结构方面，推理成本的快速下降是 Agent 商业化的核心前提，2025 年主流大模型的每百万 Token 推理成本较 2024 年下降约 80%（如 AWS Nova Lite 每百万 Token 仅 0.06 美元），企业通过分级模型策略可进一步优化整体 AI 调用成本。

5.市场格局：平台型与垂直场景双轨并行

AI Agent 的商业化呈现出平台型生态与垂直场景双轨并行的发展格局（如表 1.2 所示）：

表 1.2 平台型生态 agent 与垂直场景 agent 的对比

维度	平台型生态 Agent	垂直场景 Agent
代表	字节跳动 Coze、钉钉 Agent OS、腾讯元器	金融合规 Agent、法律审查 Agent、工业检修 Agent
竞争力	流量基础、低迁移门槛、开发者工具	行业深度知识、专有数据、SLA 服务保障
商业化确定性	中等，依赖生态规模效应	较高，针对刚性痛点，付费意愿强
收入模型	Token 抽成、订阅、生态分发	软件服务费、效率分成、按结果付费
落地瓶颈	场景同质化、缺乏高价值闭环	数据整合难度大、交付周期长

从资本市场动向看，投资方向已出现明显调整：对通用 Agent 构建平台趋于审慎，转而关注能交付明确业务结果的垂直 Agent 企业。编程、电商外贸、金融文档和工业检修是率先跑通商业闭环的先行场景。

1.1.5 辩证看待 AI 应用市场：增长是主旋律

2025 年围绕 AI 应用的一个核心争议是市场泡沫问题。Gartner 预测 40% 的 Agentic AI 项目将在 2027 年前被取消；部分厂商将传统聊天机器人和 RPA 工具重新包装为 Agent 的 Agent Washing 现象也引发了市场关注。

然而，综合多方数据分析，泡沫迹象与产业增长并存，但增长是主旋律。

支撑这一判断的核心数据包括：一是从价值创造空间看，Morgan Stanley 分析框架显示 2025 年全球企业利润约 50000 亿美元，AI 带来 1%-2% 的生产力提升即可产生 10000 亿美元的增量收益，足以支撑当前的基础设施投资规模。二是从投资主体看，不同于互联网泡沫时期依赖外部融资，当前 AI 基础设施建设由 Microsoft、Google、Amazon、Meta 等高盈利巨头驱动，其到 2030 年的投资能力合计超过 75000 亿美元。三是从营收验证看，应用层已出现明确的收入信号：OpenAI 预计 2025 年年化收入达 200 亿美元；Anthropic 年化收入突破 90 亿美元；Google Cloud 在 2025 年第四季度收入同比增长 48% 至 177 亿美元。四是从资本市场看，全球 AI 投融资占全行业投融资比例从 2023 年的 8.1% 上升至 2025 年二季度的 23%。2025 年上半年，美国 AI 投融资金额达 381 亿美元（同比增长 43.6%），中国为 36.7 亿美元。在中国市场，大模型领域投融资占 AI 总投融资金额的比例从 2023 年的 31% 上升至 2024 年的 66%，智谱 AI、月之暗面、百川智能、MiniMax 四家中国企业跻身全球 Top10 大模型融资事件。

Deloitte 2025 年的技术价值调研揭示了一个值得关注的深层规律：AI 的价值更多体现为乘数效应，而非独立解决方案。那些优先投资数据基础设施的企业获得的市值增长（65%），显著高于直接投资 AI 应用的企业（43%）。这意味着，没有坚实的记录系统（System of Record）数据架构作为基础，推理系统（System of Reason）的价值难以充分释放。

总体来看，信息技术产业发展过程中阶段性泡沫的出现具有一定的历史规律性，但 AI 应用的产业规模和增长动力构成了不可忽视的发展主线。当前阶段的核心挑战不在于 AI 能力本身，而在于企业能否将 AI 深度嵌入核心业务流程，实现从实验性验证到生产级系统的跨越。

1.2 具身智能——当 AI 拥有“身体”，世界将被重塑

1.2.1 具身智能：从“纯软件 AI”到“有身体的 AI”

具身智能（Embodied Intelligence）并非传统意义上的机器人或纯软件 AI，而是指拥有物理身体、能够通过“感知—决策—执行”的闭环在真实世界中自主完成任务并持续学习的智能系统。其本质，是赋予 AI 一个能与物理世界互动的“身体”，让智能不再局限于屏幕和云端，而是走进工厂、走上街道、走入家庭。

一个常见的误解是将具身智能等同于人形机器人。事实上，具身智能是一个从轻量到重度、从专用到通用的渐进光谱：最轻量的一端是端侧 AI——将模型能力嵌入手机、玩具等消费硬件；中间地带是机械臂和非人形机器人——在确定性高的垂直场景中实现精准自动化；最高形态则是人形机器人——追求与人类环境完全兼容的通用物理智能体。沿着这条光谱，软硬件整合的深度、技术栈的复杂度和商业化的难度依次递增。以下展开三级递进分析：

1.2.2 端侧 AI：嵌入场景的“智能器官”

端侧 AI 是具身智能光谱中最轻量、商业化最成熟的一级。它将经过压缩和优化的大模型能力直接嵌入专用硬件，使设备具备本地实时感知与决策能力，无需依赖云端即可完成任务。典型载体包括 Agent 手机（如字节跳动豆包手机）、陪伴教育机器人、智能家居终端、工业边缘计算设备等。

端侧 AI 的核心价值在于以合理的成本在特定场景中提供可靠的自动化。其商业模式已经过充分验证：多个细分领域实现了规模量产，ROI（Return On Investment）数据清晰。例如在工业领域，兴发集团部署的“数字员工”系统，以 UCS 统一控制系统为物理具身载体，集成控制器、I/O 模块和实时网络，直接嵌入工厂物理设备层，实时感知来自数万个传感器的工业时序数据；大模型 TPT 2 负责将复杂的工艺知识转化为优化指令（如“微调碱液阀门开度 0.5%”），由 UCS 毫秒级转化为真实物理动作，形成“感知—思考—行动”的完整闭环，实现关键工艺的自主精准控制，向“黑屏运行”迈进。

端侧 AI 的局限也很清晰：它的“身体”往往是固定或嵌入式的，与环境的物理交互深度有限，更多扮演“智能器官”而非“智能个体”的角色。当任务需要更深度的物理操作（抓取、搬运、装配）时，就需要进入具身智能的第二级。

1.2.3 机械臂与非人形具身系统：精准高效的“专业能手”

以固定或移动机械臂为核心的非人形机器人，是当前商业化最成熟的具身智能形态。它们在物流分拣、高端制造、工业巡检、清洁服务等确定性高、ROI 可量化的垂直场景中已实现规模量产。这一层级的技术栈比端侧 AI 深得多：机器人不仅需要感知环境，还需要在三维空间中精确规划运动路径、施加精细力控、并在执行过程中根据传感器反馈实时调整，实现完整的“手—眼—脑”协同。

物流场景是典型代表。海康机器人推出的 AMR（自主移动机器人）搭载激光雷达、3D 视觉相机及力/力矩传感器，在 SLAM 地图下实现毫米级定位；通过神经网络动态评估环境状态，在 1.5 米窄巷道中自主避障；与立库辊筒线实现±2mm 级对接精度。潜伏式 AMR 与叉取 AMR 在同一地图下协同作业，构成“无人化搬运链”，适配汽车制造等高精度、高安全要求场景。

工业制造场景呈现出鲜明的“软硬一体”趋势。在“眼”的层面，梅卡曼德以自研 Mech-Eye 高精度 3D 相机和 Mech-GPT 多模态大模型为核心，构建了机器人“眼脑手”全栈 AI 能力，产品已进入 50 多个国家和地区，用户仅通过自然语言指令就能让机器人完成复杂多样的操作任务。在“脑”的层面，大界机器人以自研工业软件平台 RoBIM 为核心，为船舶、电力、重工等行业提供覆盖切割、打磨、焊接、装配的

智能制造系统——在船舶行业，RoBIM 三个月内实现对 100 多种型材的柔性三维切割，远超国外设备仅能处理 20 余种的能力；截至 2025 年大界年营收接近 2 亿元，基本实现盈亏平衡。在“体”的层面，埃斯顿连续八年保持国产机器人市场第一，2026 年发布的 iER.OS 系统和 RoboBase 平台提供了开放性更强、兼备边缘推理计算的扩展能力，其协作机器人品牌酷卓 2025 年出货增长超过 270%。三者分别从感知、决策、执行切入，共同印证了非人形具身系统在工业场景中 ROI 清晰、商业闭环可行的判断。

非人形系统的优势在于：针对特定任务做到极致的效率和可靠性。但其局限也因此而生。形态受限意味着每进入一个新场景就需要重新设计硬件和工艺适配。当目标是创造一个能在各种人类环境中通用的物理智能体时，就进入了具身智能的第三级：人形机器人。

1.2.4 人形机器人：面向通用的“终极形态”

人形机器人被视为具身智能的终极目标，其最大优势在于与人类环境的天生兼容性——身高、体型、关节结构与人类相近，可以直接使用“为人类设计的”工位、工具和空间，无需对产线进行大规模改造。但正因其通用性追求，人形机器人的技术栈也是三级中最复杂的。

技术栈：“大脑—小脑”协同框架

当前行业主流的技术范式是“大脑—小脑”协同框架。“大脑”（决策规划系统）依托多模态大模型，汇集摄像头、激光雷达等多元传感器数据，进行环境感知、任务规划和智能决策，赋予机器人自然语言理解与自主任务分解的能力；“小脑”（运动控制系统）专注于精确、实时的全身运动控制，保障动作流畅性、稳定性及安全性，实现“手眼协调”的精细操作。两者通过高速数据交互深度耦合，形成“认知—执行”闭环。

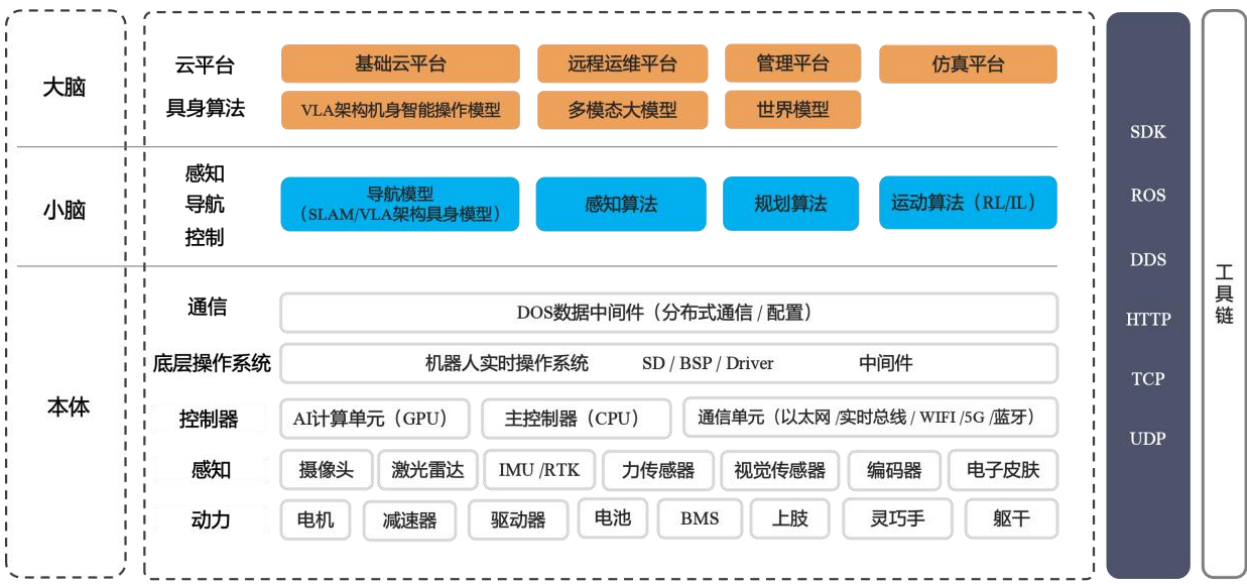


图 1.1 具身智能全栈技术布局

各环节成熟度与 ROI (Return On Investment)

人形机器人整体处于商业化早期，但不同环节的成熟度差异显著：

表 1.3 人形机器人各环节发展阶段

环节	发展阶段	ROI 体现
“大脑”（AI 算法）	快速迭代，部分场景已可用	任务泛化能力提升，降低逐项任务编程成本
“小脑”（运动控制）	相对成熟，复杂动态环境有待突破	行走稳定性、操作成功率提升，减少故障停顿
本体集成与制造	小批量试产，成本高昂	整机成本高，ROI 为负；规模量产是关键拐点
核心零部件	关节/灵巧手等国产替代进行中	供应链自主可控后将带来巨大成本下降空间

全球竞争格局：中美两大阵营

人形机器人产业已形成中美两大阵营对垒的格局。从供应链看，中国控制了全球约 70%的人形机器人零部件供应链，涵盖电机、执行器、传感器、电池和材料，单台 BOM（Bill of Materials）成本约 3.5 万-4.6 万美元；而非中国供应链的 BOM 成本高达约 13 万美元，是前者的近 3 倍。这一成本优势源自中国在电动汽车、无人机和消费电子产业中积累的制造生态密度。

在 BOM 结构中，执行器（电机+减速器+关节总成）占比最高，达 40%-60%，是最大且在规模化中最不易压缩的成本项。一台典型的人形机器人使用 25-30 个执行器，单个成本从数百到数千美元不等。计算平台相对标准化，绝大多数厂商采用 NVIDIA Jetson 系列，唯一例外是 Tesla Optimus 使用自研 FSD 芯片。

以下为主要厂商与产品的竞争全景：

表 1.4 各厂商代表产品竞争全景

厂商	代表产品	关键指标	定位与阶段
Tesla	Optimus Gen3	173cm/57kg,手部 22DoF,目标售价\$2 万-3 万	2026 年 Fremont 投产，目标产能 100 万台/年
Figure AI	Figure 03	Helix VLA 模型，BotQ 工厂年产 1.2 万台	与 OpenAI 合作，AI 能力领先

Unitree (宇树)	G1/H1/R1	G1:\$1.6 万/43DoF; R1:\$5,900	2025 年全球出货量第一 (5,500 台)
优必选	Walker S1/S2	41DoF,Co-Agent 智能体, 7×24h 作业	比亚迪/富士康产线实训中
1X Technologies	NEO	售价\$2 万	已启动早期用户交付
Agility Robotics	Digit	专用工厂, 年产能 1 万台	物流场景深度落地
智元机器人	远征/灵犀系列	聚焦工业+家庭双线	2025 年小批量量产

2026 年被业界视为人形机器人的“量产元年”。Goldman Sachs 研究预测，中国供应链正在为 10 万-100 万台/年的产能做规划。全球市场规模预计从 2025 年的约 6 亿美元增长至 2030 年的 60 亿美元，Morgan Stanley 更远期的预测则高达 2050 年 5 万亿美元。价格竞争已经展开：宇树 R1 以 5,900 美元的定价打破了行业对人形机器人价格的想象，而 Tesla 的目标是将 Optimus 做到 2 万-3 万美元（相当于一辆家用轿车的价格）。

实例：优必选 Walker S 系列

优必选 Walker S 系列是中国人形机器人商业化落地的代表。2024 年发布的 S1 进入比亚迪、富士康等车企产线实训，验证搬运、分拣、质检等基础任务；2025 年推出的 S2 针对 S1 的续航和作业空间局限进行升级，搭载自主换电系统和 Co-Agent 智能体，实现“7×24 小时连续作业”和“全空间作业覆盖”。其采用分层具身智能架构，大脑基于多模态大模型进行环境理解与任务规划，小脑实时处理 41 自由度动力学计算，通过全身力控技术实现毫米级精度。通过“软件定义硬件”模式（切换搬运 Agent、分拣 Agent），实现“一机多用”，适应多品种小批量生产需求。但挑战同样显著：在强光/反光等复杂环境下视觉深度误差扩大、高负载下平衡控制仍不够稳定、强化学习（Reinforcement Learning, RL）需要大量数据（每台每天约 1GB）而工业场景数据收集成本较高。

1.2.5 挑战与展望

具身智能从端侧 AI 到人形机器人的三级递进中，越往高处走，面临的挑战越严峻。数据是首当其冲的瓶颈：训练一个完全自主的具身智能体需要上亿条数据，但目前最大公开数据集仅在百万量级，且真实物理交互数据的采集成本远高于互联网文本数据。其次是“大脑”与“小脑”的融合鸿沟：高层智能的抽象规划与底层控制的实时执行之间存在接口与协同瓶颈，影响整体行动的流畅与精准。再者，物理世界充满不确定性——物体形变、滑动、环境动态变化，要求执行系统具备极高的自适应能力和精细力控。最后，当机器人持续采集全息化的环境与交互数据时，数据安全与伦理问题将日益凸显：隐

私保护机制是否到位？机器人是否始终可被人类控制和引导？

产业界将突破的希望寄托于“世界模型”（World Model）。在当前视觉—语言—动作（VLA）模型的基础上，引入能够对物理世界进行理解、预测和推演的世界模型，被认为是提升机器人脑能力的关键路径。未来三到五年内，预计可以看到人形机器人真正走进工厂，为客户创造实际价值。具身智能的演进方向不仅是技术的胜利，更是人机关系、生产模式乃至社会结构的一次深刻重塑。

1.3 人工智能+科学研究（AI for Science）

2025 年，中美两国近期不约而同地将“AI for Science(AI4S)”提升至国家最高战略层级，竞相利用 AI 重构科研范式。美国方面，特朗普签署行政令启动被类比为“曼哈顿计划”的“创世纪计划”，意在举国家之力构建集智算、数据与模型于一体的科学安全平台；中国方面，国务院发布《关于深入实施“人工智能+”行动的意见》，明确将“人工智能+科学技术”列为首要重点行动。这标志着 AI 加速科学发现已成为全球科技博弈的核心主战场。2025 年也是 AI 从“辅助科研”走向“驱动科研”的关键节点，AI 的角色从阅读文献、预测结果，进化到了提出假设、设计实验乃至自主发现。

Google DeepMind 是目前全球 AI4S 的绝对领军者，CEO 戴密斯·哈萨比斯（Demis Hassabis）和同事约翰·琼珀（John Jumper）因开发了人工智能模型 AlphaFold 破解了生物学界长达 50 年的终极难题——蛋白质结构预测，而获得 2024 年诺贝尔化学奖。Deepmind 策略是针对生物、材料、气象等硬科学领域的“卡脖子”问题，训练专用的深度学习模型。在生物学方面，DeepMind 新推出的 AlphaFold 3 实现了里程碑式的飞跃。它不再局限于蛋白质结构，而是能预测所有生命分子（蛋白质、DNA、RNA、配体等）及其相互作用。这直接加速了药物设计和基因组学研究。在材料科学方面，Deepmind 利用图神经网络发现了超过 220 万种理论上稳定的新晶体结构（相当于人类过去 800 年发现总和的 45 倍），并已通过机器人实验室验证了部分材料，极大地缩短了电池、芯片新材料的研发周期。而 Claude 则走“通用智能体”路线，利用其强大的逻辑推理与代码执行能力，充当“超级科研助理”，凭借超大的上下文窗口，Claude 能一次性阅读数百篇论文，帮助科学家进行跨学科的文献综述、假设生成和实验设计。Claude 3.5 引入的计算机操作能力（Computer Use），使其能够像人类一样操作复杂的实验室软件、编写并运行仿真代码、自动处理数据。这使得它正逐渐演变为能自主操作湿实验（Wet Lab）和干实验（Dry Lab）的“科研智能体”。

中国方面，上海人工智能实验室（Shanghai AI Laboratory）在 AI for Science 提出通专融合技术路线，打造国际领先的 AGI4S 科学基座与发现系统，旨在重构科研范式。一是发布科学多模态大模型 Intern-S1。针对大模型“通而不专”痛点，提出 SAGE“智者”三层技术体系。开源发布的 Intern-S1 通用性能处于开源第一梯队，在专业科学任务及国际物理奥赛中表现超越 GPT-5 和 Grok-4 等闭源模型。二是打造跨学科通用科学发现系统 Intern-Discovery，通过跨学科知识网、多智能体协同及首创的“科学实体上下文协议”，打破科研孤岛。系统整合 3 亿文献、40PB 数据及 2000+ 专用模型（如风鸟、丰登），

在 GAIA、MLE-Bench 等国际评测中夺冠。目前已赋能 200+ 机构，在癌症靶点、超导材料等领域实现 10 余项重大突破，推动科研从单点探索向协同涌现跃迁。与此同时，在鄂维南院士的引领下，深势科技（DP Technology）已成为中国“AI for Science”工业化落地的领军独角兽。深势科技打造了覆盖微尺度工业研发全链条的产品矩阵：以“玻尔科研空间站”为枢纽，包含 Bohrium® 科研云平台、Hermite® 药物计算设计平台、Piloteye 电池设计自动化平台以及 RiDYMOM 难成药靶标研发平台等。玻尔科学导航已服务全球超 1000 所高校和科研机构的 300 多万名科学家，包括中国绝大多数 985、211 高校整体入驻，支撑了上千个科研项目，相关产品和服务为超过 150 家先进研发企业提供智能化升级解决方案，深度赋能复星医药、国药集团等 70 余家生命科学企业的 100 多条研发管线，以及宁德时代、比亚迪等物质科学领域头部客户。

1.4 本章小结

2026 年政府工作报告明确提出“深化拓展‘人工智能+’”，促进新一代智能终端和智能体加快推广，培育智能原生新业态新模式。AI 已从实验室的精密工具走向千行百业的生产现场，从辅助人类决策的工具进化为重构商业逻辑的底层操作系统。

AI-Native 的终极形态，将是“无感化智能”的全面渗透。未来的应用不再区分“AI 功能”与“传统功能”，智能将成为像电力一样的基础设施——无处不在却无形无感。当智能终端、智能体与物理世界深度耦合，人机交互的边界将进一步消融，取而代之的是环境智能对需求的预判与响应。这种转变意味着，企业的竞争壁垒不再源于是否使用 AI，而在于能否构建“数据－模型－场景”的飞轮效应，让智能在业务流转中自我进化。

具身智能的破局之路，指向“世界模型”驱动的通用机器人时代。当前的数据瓶颈与协同鸿沟只是阶段性阵痛，随着物理世界理解能力的突破，机器人将从特定场景的“熟练工”进化为具备常识推理的“通用劳动力”。这不仅将重塑制造业的生产范式，更将引发服务业、农业乃至家庭场景的连锁变革。当机器真正理解物理世界的因果规律，人机协作将超越简单的指令执行，进入“意图共享、风险共担”的深度协同新阶段。

AI for Science 的深远意义，在于开启“自主发现”的新科学革命。国务院《关于深入实施“人工智能+行动”的意见》将“人工智能+科学技术”列为首要重点行动，标志着 AI 加速科学发现已成为全球科技博弈的核心主战场。当 AI 从文献阅读者变为假设提出者，从实验预测者变为方案设计者，科学发现的周期将被大幅压缩，创新范式从“试错驱动”转向“认知驱动”。这不仅是研究效率的提升，更是人类认知边界的拓展——在材料、能源、生命科学等战略领域，AI 将成为破解复杂系统奥秘的“望远镜”与“显微镜”，加速解决气候变化、疾病治疗等人类共同挑战。

本章参考资料

[1]中国信息通信研究院.人工智能产业发展研究报告（2025）[R].2026.

- [2]华为云.AI-Native技术与实践白皮书[R].2025.
- [3]搜狐科技.超 10 款大模型已具备“自我复制”能力[EB/OL].2025.
- [4]财联社.觉醒第一步? OpenAI模型在研究中违抗命令, 竟然篡改关机脚本[EB/OL].2025.
- [5]光明网.OpenAI等创立基金会, 推动AI Agent标准化[EB/OL].2025.
- [6]光明网.智源研究院发布全球首个跨本体具身协作框架与开源具身大脑[EB/OL].2025.
- [7]林典驰.专访优必选科技创始人、董事会主席兼CEO周剑: 相比做表演, 人形机器人更要进工厂干实事[EB/OL].2025.
- [8]何文翔, 黄朝峰.战略深解读 | 再造曼哈顿丰碑? 美启动AI“创世纪计划”的雄心与阻力[EB/OL].2026.
- [9]Grand View Research.Artificial Intelligence Market(2026-2033)[R].2026.
- [10]Grand View Research.AI Agents Market(2026-2033)[R].2026.
- [11]Grand View Research.AI Agents Market Size To Reach\$182.97 Billion By 2033[EB/OL].2025.
- [12]Open Router.AI Model Rankings(2026.01)[Z].2026.
- [13]Kimi Team.Kimi K2.5:Visual Agentic Intelligence[R].2026.
- [14]Worklytics.Proving the ROI of AI Adoption:Metrics and Dashboards Every Org Needs in 2025[R].2025.
- [15]McKinsey.The agentic organization:Contours of the next paradigm for the AI era[R].2025.
- [16]BCG AI Radar.From Potential to Profit:Closing the AI Impact Gap[EB/OL].2026.
- [17]Anthropic.Code execution with MCP:Building more efficient agents[EB/OL].2025.
- [18]Gartner.Gartner Predicts Over 40%of Agentic AI Projects Will Be Canceled by End of 2027 [EB/OL].2025.
- [19]Morgan Stanley.AI Funding:The Bull and Bear Investment Cases[EB/OL].2025.
- [20][Business of apps](#).ChatGPT Revenue and Usage Statistics(2026)[R].2026.
- [21]Luis Espada.Alphabet advertising revenues climb 14%as Gemini App reaches 750 million users[EB/OL].2026.
- [22]Deloitte.AI is capturing the digital dollar.What's left for the rest of the tech estate?[R].2026.
- [23]Humanity's Last Machine.A Deep Dive on Humanoid Hardware[R].2026.
- [24]Goldman Sachs Research.China humanoid robot supply chain field trip takeaways:Optimistic capacity preparation in advance,awaiting actual orders[R].2025.
- [25]Morgan Stanley.China Industrials:2026 Outlook–Humanoids[R/OL].2026.

第 2 章 AI 大模型：深度认知与自动行动的范式变革

2025 年应当被人工智能发展史标记为“范式转移”的一年，全球人工智能产业正经历从“算力暴力美学”向“架构效率与深度推理”转型的关键拐点。随着 Scaling Law（规模法则）在单纯参数堆叠层面遭遇边际效应递减的挑战，技术竞争的焦点已从预训练（Pre-training）阶段全面延伸至后训练（Post-training）、推理时计算（Test-time Compute）以及新型架构的探索。行业关注的焦点，正从“模型能聊什么”急剧转向“模型能做什么”以及“模型如何理解物理世界、如何影响物理世界”。

本章所讨论的，正是这一轮范式重构背后的核心逻辑。一方面，大模型能力边界持续外扩，模型正由文本生成迈向对现实世界更深层的模拟、推理与操作；另一方面，产业竞争格局也在加速重塑，模型竞争不再局限于参数规模和单点评测成绩，而是逐步转向性能、效率、生态与落地能力的综合比拼。尤其是在算力约束和地缘政治等多重因素交织的背景下，中国人工智能产业正在由早期以跟随优化为主，转向更加重视底层架构创新、系统工程能力和成本效率优势的新阶段，并对全球开源模型生态和产业竞争格局产生实质性影响。

2.1 中美大模型竞争态势与战略分化

2025 年的 AI 技术版图呈现出“两极多强”的复杂态势。美国企业依托算力基础设施、基础研究和平台生态优势，持续推动通用模型能力上限；中国企业则在算力资源约束条件下，更加注重效率优化、工程协同和开源开发，逐步形成差异化发展路径。中美竞争也由单一的模型能力竞争，延伸至训练范式、产业生态、商业模式和应用渗透的多维竞争。

2.1.1 共识演进：Scaling Law 的“收益递减”与“研究时代”的回归

在经历了长达多年的“算力暴力美学”之后，行业开始重新审视 Scaling Law 的适用边界。尽管数据规模与算力投入的增加仍能够有效降低预训练损失（Pre-training Loss），但其对模型推理能力（Reasoning）与泛化能力（Generalization）的提升效率已不如早期阶段显著。

2025 年底，OpenAI 的前首席科学家、Safe Superintelligence（SSI）创始人 Sutskever 明确提出：2020 年至 2025 年的“扩展时代”（Age of Scaling）已经接近结束，行业正重新进入“研究时代”（Age of Research）。他的核心观点在于，单纯将计算集群规模扩大 100 倍，已无法像过去那样线性地带来智能的质变。未来的突破将更多依赖于算法的根本性创新、更高质量的合成数据以及测试时计算（Test-time Compute）的有效利用。

从产业层面看，这一变化意味着模型能力提升路径正在发生结构性迁移。过去的核心问题是“如何训练更大的模型”，当前的核心问题则更多转向“如何让模型更会思考、更会调用、更会执行”。这也为后训练、测试时计算和 Agent 系统的快速崛起奠定了基础。

2.1.2 效率驱动与工程优化：中国路径的强化

在先进制程芯片获取受限、算力资源相对紧张背景下，我国大模型企业加快由规模扩张向效率

优先转变，更加突出单位算力产出最大化。模型研发重点也由单纯追求参数规模提升，转向结构设计、训练框架、通信系统和推理部署等环节的协同优化，着力在既有硬件条件下同步提升训练效率、推理吞吐和部署可行性，推动性能、成本与扩展能力实现更优平衡。

国内大模型企业围绕效率优先持续推进技术迭代。2025 年初，DeepSeek 系列模型引发业界的“DeepSeek 冲击”（DeepSeek Shock）。其中，DeepSeek-V3 作为一款 6710 亿参数的混合专家（MoE）模型，凭借多头潜在注意力（MLA）机制和无辅助损失负载均衡等设计，仅用 278.8 万 H800 GPU 时完成训练，训练成本显著低于同级别美国模型；DeepSeek-R1 在推理任务上的表现接近 OpenAI o1，兼具低成本和开放属性，相关论文登上 Nature 封面。这表明，“模型—系统协同设计”（Model-System Co-Design）正在成为降低训练成本、提升推理效率和增强部署可行性的重要路径，也在一定程度上推动 Meta 等企业加快开源模型布局。与此同时，Qwen、MiniMax 等模型也持续在长上下文处理、注意力机制和推理效率等方面持续推进优化。总体来看，我国大模型技术创新正在由点状突破转向体系化推进，逐步拓展至模型、框架、硬件适配和推理部署协同推进的系统化阶段，技术演进的协同性和整体性持续增强。

我国与国际前沿模型之间的能力差距也呈现出明显收敛态势。斯坦福大学《2025 年人工智能指数报告》指出，到 2024 年底，中美顶尖模型在 MMLU、HumanEval 等主要基准测试上的性能差距，已由 2023 年的两位数水平收窄至接近持平。报告同时认为，随着 DeepSeek-R1 等模型于 2025 年初发布，中国实验室模型进一步逼近国际前沿水平，多个国产模型在推理、代码和通用能力测试中实现跨代追赶。Epoch AI 基于 Epoch Capabilities Index 的分析也显示，自 2023 年以来，美国与中国最强模型之间的能力代差大致在 4 至 14 个月之间，平均约 7 个月。总体看，中美之间仍有差距，但这种差距更多体现为迭代节奏上的时间差，而非难以跨越的代际断层。**随着模型迭代周期持续缩短、成本优势逐步显现，国际竞争的焦点也正由“谁先发布最强模型”，转向“谁能够以更低资源消耗和更快迭代速度，推动模型能力更快进入规模化应用”。**

效率优势也正在加快转化为价格优势。根据 OpenRouter 平台在 2026 年 3 月的公开报价，DeepSeek V3 0324 的价格约为输入 0.20 美元/百万 tokens、输出 0.77 美元/百万 tokens；同期 OpenAI GPT-5.4 价格约为输入 2.50 美元/百万 tokens、输出 15 美元/百万 tokens。这样的成本差异会直接影响产业落地，因为工业质检、企业客服、办公自动化、终端智能体等高频调用场景对调用成本普遍较为敏感。在基础模型能力差距逐步收敛的情况下，价格、部署效率和服务稳定性正成为影响应用扩散的关键变量，单次基准测试成绩的重要性则相对下降。

2.1.3 开源 VS 闭源：技术与生态方面的双重竞争

在技术路线和商业模式层面，开源与闭源正呈现并行发展态势，并在中美两国形成不同的侧重点和竞争逻辑。总体来看，美国更强调通过闭源体系巩固能力优势和商业闭环，中国则更倾向于通过开源加快技术扩散和生态培育。

美国路径方面，呈现出“核心能力闭源、外围生态开放”的特征，在保持模型与商业体系控制力的同时，不断扩大其全球开发者覆盖面。以 OpenAI、Anthropic、Google 为代表的头部企业整体仍以闭源模型为核心，一方面持续加大算力基础设施投入，依托超大规模智算集群支撑模型能力上限的提升；另一方面通过强化学习等方法持续增强推理能力，提高模型在数学、编程等高复杂度任务中的表现。在产业生态层面，美国依托 GitHub、Hugging Face 等全球性开源社区，长期积累了广泛的开发者基础和影响力，为技术扩散、工具链完善和应用生态构建提供了重要支撑。

中国路径方面，以 DeepSeek、阿里 Qwen 等为代表的企业，在开源方向上更为积极，正在形成以能力开放带动生态扩散的发展路径。以 DeepSeek 为例，DeepSeek-R1 采用 MIT 许可证开源模型权重，允许自由使用、修改与商业化部署，显著降低了先进模型的使用门槛。此后，DeepSeek 又通过“开源周”等方式，集中发布推理加速、通信优化、训练框架等多项基础设施组件，推动模型能力与底层系统能力同步开放。截至 2026 年 3 月，DeepSeek-R1 在 Hugging Face 平台点赞数已超过 1.3 万，下载量超过 1700 万次，并已接入微软 Azure、英伟达、阿里云、百度云等多家云平台。Qwen 系列则通过持续开源模型和工具体系，不断夯实开发者生态基础。2026 年 2 月，Qwen3.5 系列发布后迅速占据 Hugging Face 趋势榜前列，其中 Qwen3.5-35B-A3B 在发布不足 24 小时内即登顶。与此同时，阿里也在积极建设本土开源社区和生态平台，如魔搭社区，持续提升开源生态的集聚能力和国际影响力。

开源路径的核心作用在于通过降低技术门槛，迅速扩大使用规模，并在应用层加快形成事实标准。随着开发者生态不断积累，模型本身的控制权虽有所弱化，但生态影响力和市场渗透能力反而持续增强。开源与闭源的竞争，正在从单纯的模型发布竞争，进一步演化为围绕技术扩散速度、开发者生态规模和产业协同能力的综合竞争。

2.2 大模型核心技术演进

2.2.1 后训练技术革命：从知识灌输到思维涌现

如果说 2024 年以前，大模型能力提升主要依赖海量数据预训练带来的知识积累，那么 2025—2026 年，行业关注重点已明显转向后训练阶段，模型竞争的核心也由“学到更多知识”转向“具备更强推理能力”。从技术范式看，人工智能正由更偏向直觉式模式匹配的阶段，逐步转向更强调逻辑演绎、步骤验证和自我修正的阶段。

测试时扩展（Test-time Scaling）开始成为能力提升的重要来源。在心理学中，System 1 强调快速、直觉式、低成本的反应，System 2 则强调缓慢、审慎、分步骤的推理。Test-time Scaling 为模型提供了由快速作答转向深度思考的计算条件，其本质是将部分计算资源由训练阶段延伸到推理阶段：对于简单问题，模型仍可沿用更接近 System 1 的方式快速响应；对于复杂问题，则调用更接近 System 2 的机制，通过更长的推理链条、更充分的中间验证和更多候选路径的比较，换取更高的正确率。OpenAI 的 o1、o3 以及 DeepSeek-R1 都不在于模型规模的单纯扩大，而在于在处理复杂数学、代码和分析类

问题时，能够显著增加中间推理过程，表现出更强的分步求解、错误发现和路径回溯能力。此外，DeepSeek-R1 对显性思维链的展示，也在一定程度上提升了外界对推理模型工作机制的理解。

与 Test-time Scaling 相伴而生的，是可验证奖励强化学习（Reinforcement Learning from Verifiable Rewards, RLVR）的快速发展。若说 Test-time Scaling 解决的是模型在使用阶段如何“想得更深”，那么 RLVR 解决的则是模型在训练阶段如何学会这种推理方式。传统 RLHF 更适合对齐模型的表达风格、安全边界和人类偏好，但在数学、代码、逻辑推理等任务中，决定模型表现的关键在于“过程是否正确、结果是否可验证”。RLVR 是以可程序化验证的“真值信号”替代主观人类反馈作为奖励函数，例如，数学题是否得到唯一正确答案、代码是否通过测试用例、推理过程是否满足既定约束，都可以作为明确、低噪声、可规模化的训练反馈。与人工偏好打分相比，这类奖励信号更稳定，也更适合在大规模训练中反复迭代。从训练机制看，RLVR 强化的是模型处理复杂问题的过程能力，在大量“生成—验证—修正—再生成”的循环中，会逐步学会拆解问题、比较路径、回溯纠错，并重新组织推理。这种能力本质上就是对 System 2 式推理的训练。

总体来看，Test-time Scaling 与 RLVR 正在共同推动大模型由偏向 System 1 的“知识型模型”向兼具 System 2 能力的“推理型模型”演进。前者提升的是模型在推理阶段释放能力的深度，后者塑造的是模型形成这种深度推理能力的训练机制。两者叠加，使后训练阶段由过去的辅助环节逐步上升为决定模型上限的关键战场。

2.2.2 多模态融合与世界模型

多模态大模型的发展已不再停留于对文本、图像、音频等不同模态的简单叠加，而是进一步朝着统一建模、环境理解和世界表征的方向演进。

以 GPT-4o 为代表的新一代模型，推动多模态系统由过去“语音转文本—文本推理—文本转语音”的级联式结构，转向在统一模型中直接处理不同类型 Token 的方式。通过单一 Transformer 同时处理文本、图像和音频信息，模型能够更自然地完成多模态交互，在语音响应速度、情感识别和上下文一致性等方面表现出更强能力。OpenAI 在 GPT-4o 系统说明中明确指出，该模型是一个端到端训练的“omni”模型，可在同一神经网络中处理文本、视觉、音频和视频输入，并可在数百毫秒级别响应音频输入。随后，Gemini 3、Qwen3-Omni 等模型也在原生多模态方向持续推进，行业整体趋势正由“外挂式多模态”转向“原生式多模态”。

世界模型（World Model）逐步成为更受关注的前沿方向。学术界对“世界模型”尚无完全统一的定义，但较为经典的理解来自强化学习领域。按照 Sutton 和 Barto 的定义，模型是智能体对环境运行方式的内部表征，即能够根据当前状态和动作预测下一步状态及相应结果。沿着这一路径，David Ha 与 Jürgen Schmidhuber 在《World Models》中进一步提出，可通过学习压缩的时空表征来构建环境模拟，并在该模拟环境中训练智能体。由此出发，只要系统具备对环境变化的预测能力，便可以被视为具有某种意义上的“模型”，这也为当前视频生成、环境模拟等系统被视作世界模型的早期形态提供了重要理

论基础。

围绕世界模型的研究，正在由“能否预测下一步”进一步走向“能否真正理解世界”。一类路径强调基于大规模视频和多模态数据，对物理过程和动态变化进行统计建模，从而实现更高质量的视频生成和环境模拟。OpenAI 在介绍 Sora 时将其定位为“video generation models as world simulators（世界模拟器）”，强调其目标是构建能够理解并模拟现实的视频模型；Google 的 Gemini 与 Veo 路线也在持续强化多模态理解、时序一致性和更复杂的生成能力。另一类路径则更强调模型应当理解事物的状态、结构与因果关系，而不只是预测像素变化。LeCun 则提出 JEPA（Joint Embedding Predictive Architecture）架构，他认为未来智能系统需要建立可配置的预测式世界模型，并通过联合嵌入等方式学习更抽象的状态表征，而不是停留在像素级预测层面。

与此同时，学术界还兴起了空间智能方向的探索。“AI 教母”李飞飞教授创立的 World Labs 将模型能力从二维图像理解进一步推向三维空间表征，强调人工智能应具备对 3D 世界的感知、生成、推理和交互能力。模型不只要“生成视觉内容”，还要能够理解物体的几何结构、遮挡关系和空间位置，并据此构建可交互、可漫游的环境。这意味着，多模态与世界模型的发展，正在推动人工智能由信息处理系统向现实世界交互系统演进，并为机器人、自动驾驶、数字孪生和具身智能等更复杂的现实任务奠定基础。

总体来看，多模态与世界模型的发展，正在推动人工智能由信息处理系统向现实世界交互系统演进。其意义已不再局限于提升内容生成能力，而是在为机器人、自动驾驶、数字孪生、具身智能等更复杂的现实任务奠定基础。谁能更早实现对真实世界中“感知—理解—预测—交互”链条的有效建模，谁就更有可能在下一阶段技术竞争中占据优势。

2.2.3 Agentic AI 的崛起

2025 年被视为“智能体（Agent）元年”。相较于传统聊天模型，智能体将规划、记忆、工具调用和执行整合进统一系统，使大模型从回答问题进一步走向完成任务。由此，大模型竞争的重点也从单轮问答能力，转向复杂任务的拆解、调度与闭环执行能力。

美国企业在通用智能体和协议生态方面仍保持先发优势。OpenAI 已将智能体能力从单一产品推进到模型和工作流层面：从具备浏览器操作能力的 Operator，到可自主完成检索、分析和报告生成的 Deep Research，再到 2026 年发布的 GPT-5.4 与 GPT-5.3-Codex，OpenAI 正在把推理、编码、工具使用和长程任务执行整合进统一能力体系。Anthropic 的推进路径更强调模型能力与协议标准的同步建设。2026 年发布的 Claude Sonnet 4.6 和 Opus 4.6，重点强化了编码、计算机操作、长上下文推理和长程智能体任务持续性；与此同时，Anthropic 持续推进 MCP，将模型与外部数据源、工具系统的连接从定制集成推向标准化。Google 的布局则更明显地指向模型、代理和协议一体化演进，Gemini 3 系列持续强化多步骤任务处理与复杂工具调用能力，Google 还通过 Interactions API 向开发者开放 Deep Research，并推出 A2A 协议，尝试降低不同智能体之间的通信与协作门槛。整体来看，美国路径的优势已不再局

限于单一模型性能，而是体现在模型能力、工具体系、协议标准和平台控制力的整体领先。

中国企业的推进节奏则更突出产品化、场景化和成本效率。智谱推出全球首个终端智能体 AutoGLM，终端智能体能力向手机和桌面延伸，GLM-5 也将多步推理、长程执行和端到端开发能力作为重点方向。月之暗面的 Kimi K2.5 进一步整合长上下文、多模态和工具调用能力，支持 256K 上下文、多步工具调用和复杂 Agent 任务，并在编程与专业工作流场景中持续强化表现。MiniMax 的迭代更集中地落在生产力场景，M2 自发布起就面向 Agent 和代码任务，M2.5 强化了编程、工具调用和复杂任务处理能力，M2.7 则进一步将复杂 Agent Harness、自主工具搜索和团队式 Agent 协作推向真实研发与办公场景。阿里 Qwen 则采用模型、框架和开发工具同步推进的方式扩展 Agent 生态，Qwen3 延续了在推理、编码和工具使用上的强化，Qwen-Agent 已成为 Qwen Chat 的后端框架之一，Qwen Code 则进一步把终端 Agent 能力推向代码仓库理解、自动化修改和开发流程协同。整体来看，中国厂商更强调把智能体能力尽快落到编程、终端操作、办公自动化和垂直流程中，并依托开源与价格优势，加速进入全球开发者工作流。

2026 年初，开源智能体框架快速扩散，Agent 能力正在从少数模型厂商主导的封闭体系，转向更广泛的开发者生态。以 OpenClaw（龙虾）为代表的开源框架，能够连接底层模型与外部工具，并通过技能扩展支持更多工作流场景，显著降低智能体开发和部署门槛，推动智能体从展示型产品走向可配置、可部署、可扩展的工作流系统，也为“一人创业公司（OPC）”提供了更多可能。此外，由于智能体在执行多步任务时，Token 消耗显著高于传统聊天场景，调用成本将直接影响其可用性和规模化落地，并进一步带来“词元（Token）经济”。根据 OpenRouter 平台周度使用数据，2026 年 2 月 9 日至 15 日当周，中国模型调用量达到 4.12 万亿 Token，首次超过同期美国模型的 2.94 万亿 Token；截至 3 月 15 日，中国 AI 大模型周调用量进一步升至 4.69 万亿 Token，全球调用量排名前三的模型已全部由中国厂商占据，分别为 MiniMax 2.5、阶跃星辰 Step 3.5 Flash（free）和 DeepSeek V3.2。可以看出，在开源框架、自主智能体发展和全球开发者生态共同推动下，低成本、高频调用的中国模型正在更快进入海外开发者的真实使用链路。

2.3 本章小结

站在 2026 年的开端回望，人工智能正经历由“参数规模扩张”向“深度认知与自主行动”加速演进的重要转变。随着训练端 Scaling Law 边际收益逐步收窄，行业竞争的重点正由“谁能训练更大的模型”转向“谁能让模型更会思考、更会调用、更会执行”。推理时计算的重要性持续上升，模型开始通过更长的推理链条、更复杂的中间验证和更高强度的计算投入，提升在数学、代码、复杂检索和多步骤任务中的准确率。未来高价值模型的竞争焦点，将更多体现为推理深度、执行稳定性和复杂任务处理能力。

与能力提升路径同步变化的，是智能体系统形态的重构。“大模型本身不是护城河”，当前的智能体正在向着构建完整的运行系统加速发展。基础设施与保障、规划、记忆、工具、技能、评估验证、追

踪观测等能力被纳入统一的 Harness Engineering 框架，模型开始由“回答问题”走向“完成任务”甚至“自主进化”。与此同时，认知与行动之间的边界也在被快速打通。随着 ReAct 范式、Computer Use 能力和各类 Agent 框架逐步成熟，行业评估标准正由单一的问答准确率，转向真实场景下的任务完成率、流程闭环能力和跨系统调用效率。多智能体系统也正在成为承接复杂任务的重要组织形式，推动人工智能由单点能力竞争转向系统能力竞争。

多模态与世界模型的发展，则进一步将这种能力延伸至更加接近现实世界的层面。模型已不再局限于对文本语义的理解和生成，而是在向图像、音频、视频、空间关系和动态过程的一体化建模迈进。无论是多模态统一架构、空间智能的形成，还是世界模型对环境状态、物体关系和物理规律表征能力的增强，其共同指向都是提升模型对真实世界的理解深度。这一趋势不仅将重塑内容生成与交互产业，也将为机器人、自动驾驶、数字孪生和具身智能奠定更坚实的认知基础。谁能率先打通“感知—理解—推理—执行”的完整链条，谁就更有可能在下一阶段的现实任务场景中占据优势。

从产业层面看，全球大模型竞争已不再是单一维度的参数竞赛，而是演变为性能、效率、价格、生态和落地能力的综合竞争。美国仍在前沿模型、协议标准和平台体系方面保持领先，中国则依托效率优化、系统工程、开源扩散和成本优势，持续缩小与国际前沿的差距，并在应用扩散和开发者生态拓展方面形成新的竞争力。未来大模型竞争的关键，不仅在于模型能力本身，更在于谁能以更低成本、更高稳定性和更快迭代速度，将模型推向更大规模的真实场景。

科学发现是对智能的终极考验。人工智能的影响正在超越传统信息处理范畴，逐步进入科学发现和实验体系重构阶段。围绕自动化实验流程、科学智能体和 AI 驱动科研系统的探索，正推动科学研究向更强闭环化方向发展。未来在材料筛选、药物研发、复杂系统模拟等领域，智能体有望承担更多从假设生成、实验模拟、实验设备调用、数据分析到结果修正的中间环节，人类角色则更多转向目标设定、约束定义和结果解释。尽管“全自动无人实验室”仍处于演进过程中，但其所反映的趋势已经十分明确：人工智能正从提升研究效率的辅助工具，逐步演变为参与知识生产和科学发现的重要组成部分，并且将彻底引发科学研究范式的变革。AI 终将成为全球科技竞争中最为关键的胜负手。

本章参考资料

[1]Bouchard Y.Highlights from Ilya Sutskever's November 2025 interview with Dwarkesh Patel[EB/OL].(2025-11-25)[2026-03-27].<https://forum.effectivealtruism.org/posts/iuKa2iPg7vD9BdZna/highlights-from-ilya-sutskever-s-november-2025-interview>.

[2]Guo D Y,Yang D J,Zhang H W,et al.DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning[J].Nature,2025,645:633-638.

[3]中国信息通信研究院.人工智能产业发展研究报告（2025 年）[R].北京:中国信息通信研究院,2025.

[4]Stanford Institute for Human-Centered Artificial Intelligence.The 2025 AI Index

Report[R].Stanford:Stanford University,2025.

[5]Epoch AI.Epoch Capabilities Index[EB/OL].(2025)[2026-03-27].<https://epoch.ai/benchmarks/eci>.

[6]Kahneman D.Thinking,Fast and Slow[M].New York:Farrar,Straus and Giroux,2011.

[7]OpenAI.OpenAI o1 System Card[EB/OL].(2024-09-12)[2026-03-27].<https://www.openai.com/o1>.

[8]OpenAI.OpenAI

o3-mini[EB/OL].(2025-01-31)[2026-03-27].<https://openai.com/index/openai-o3-mini/>.

[9]Ji T,Yuan Y,Wen Q,et al.Reinforcement Learning with Verifiable Rewards Implicitly:A Survey[EB/OL].(2024)[2026-03-27].<https://openreview.net/forum?id=jGbRWwldiy>.

[10]OpenAI.Addendum to GPT- 4o System Card:4o image generation[EB/OL].(2025-03-25)[2026-03-27].<https://openai.com/index/gpt-4o-image-generation-system-card-addendum/>

[11]Google DeepMind.Gemini 3:A new era of intelligence[EB/OL].(2025-11-18)[2026-03-27].<https://blog.google/products-and-platforms/products/gemini/gemini-3/>.

[12]阿里云通义实验室.Qwen3-Omni: 一个模型, 全能不偏科[EB/OL].(2025-09-25)[2026-03-27].<https://mp.weixin.qq.com/s/chYXeDXilm4l4mkdh2ygBg>.

[13]Sutton R S,Barto A G.Reinforcement Learning:An Introduction[M].2nd ed.Cambridge:MIT Press,2018.

[14]Ha D,Schmidhuber J.World Models[EB/OL].(2018)[2026-03-27].<https://arxiv.org/abs/1803.10122>.

[15]OpenAI.Sora:Video generation models as world simulators[EB/OL].(2024-02-15)[2026-03-27].<https://openai.com/research/video-generation-models-as-world-simulators>.

[16]Wiedemer T,Li Y,Vicol P,et al.Video Models are Zero-Shot Learners and Reasoners[EB/OL].(2025)[2026-03-30].<https://arxiv.org/abs/2509.20328>.

[17]Maes L,Le Lidec Q,Scieur D,LeCun Y,Balestrieri R.LeWorldModel:Stable End-to-End Joint-Embedding Predictive Architecture from Pixels[EB/OL].arXiv:2603.19312,(2026)[2026-03-27].<https://arxiv.org/abs/2603.19312>.

[18]World Labs.Spatial Intelligence[EB/OL].(2024)[2026-03-27].<https://www.worldlabs.com/>.

[19]中国信息通信研究院人工智能研究所,华为技术有限公司.智能体技术和应用研究报告 (2025 年)[R].北京:中国信息通信研究院,2025.

- [20]Anthropic.Claude Opus 4 and Claude Sonnet 4[EB/OL].(2025-05-22)[2026-03-27].<https://www.anthropic.com/news/claude-4>.
- [21]GLM-5 Team.GLM-5:From Vibe Coding to Agentic Engineering[EB/OL].(2026)[2026-03-27].<https://arxiv.org/abs/2602.15763>.
- [22]Kimi Team.Kimi K2.5:Visual Agentic Intelligence[EB/OL].(2026)[2026-03-27].<https://arxiv.org/abs/2602.02276>.
- [23]MiniMax.MiniMax M2.7:Agent Self-Evolution Path[EB/OL].(2026-03)[2026-03-27].<https://www.minimax.io/>.
- [24]OpenClaw Project.OpenClaw v2026.3.7 Release Notes[EB/OL].(2026-03)[2026-03-27].<https://github.com/openclaw-project/openclaw>.
- [25]谢岚.日均 Token 调用量爆发式增长折射中国 AI 产业新图景[N].证券日报,2026-03-25(A01).
- [26]OpenRouter.Global AI Model Usage Statistics Report[EB/OL].(2026-02)[2026-03-27].<https://openrouter.ai/>.

第 3 章 AI 数据的变革与发展趋势：从“资源积累”迈向“智能基座”

随着“人工智能+”行动在千行百业的深入推进，数据已不再仅仅是静态的存储资源，而是演变为驱动大模型智力涌现的核心动力。2025 年，全球 AI 数据产业经历了一场深刻的供给侧改革：从盲目追求 PB 级规模的粗放增长，转向追求高密度、高信噪比的“质量至上”；从单一的文本模态，迈向图文音视频交织的“多模态融合”；从中心化的数据加工，演进为分布式的隐私计算与联邦智能。本章将系统剖析 AI 数据技术的最新范式转变，重点阐述以合成数据、非结构化文档解析、向量数据库为代表的基础设施升级，并结合国家数据局关于“数据要素×”与“人工智能+”的战略导向，探讨如何构建一个开放、安全、高质量的数据生态体系，为 2026 年的行业爆发奠定基石。

3.1 范式转变：重新定义 AI-Ready 数据

人工智能技术的演进对数据工作提出了更高要求，带动数据领域的范式转变。随着大模型技术应用的快速发展，人工智能的研发重点从“重点优化模型架构”转向“模型与数据协同优化”，数据行业共识也从此前追求“数据大即美”到强调数据的质量与精度。已有的研究和实践证明，海量但低质的原始数据非但无法提升模型智能，反而会成为导致模型产生“幻觉”与推理退化的毒药。当用于人工智能训练的数据被污染，可能会引发伦理道德、社会矛盾等方面风险，并影响个人认知和决策。因此，建设“AI-Ready”数据，已成为决定智能系统能力上限的新质生产力基石，标志着从“数量燃料”到“质量基石”的产业演进。

AI-Ready（AI 可用）指经过专门整理、特征化和组织的数据，可用来以更小的工程成本直接用于训练、推理和决策。不同于传统的数据清洗，AI-Ready 数据要求数据具备语义的纯净度，消除如文档格式错乱导致的逻辑断层，确保信息流的连贯与准确；强调知识的高密度，提供教科书般严谨、低噪声、高信息含量的内容；同时，必须实现多维的对齐性，确保数据不仅在事实上正确，更在逻辑、指令遵循与价值观上与人类意图对齐。AI-Ready 本质上是对数据进行的“精炼”过程，对数据高准确性、时效性、权威性的要求使得数据规模让位于数据质量，精细化处理过后的高质量数据成为技术创新进步的关键支撑。

由“数量”到“质量”的转变也与国家层面的战略导向高度契合。2023 年 12 月，国家数据局等 17 个部门联合发布《“数据要素×”三年行动计划（2024—2026 年）》，强调打造高质量人工智能大模型训练数据集，以高质量语料库和基础科学数据支持开展通用人工智能大模型和垂直领域人工智能大模型训练。2024 年 6 月，工业和信息化部等 4 部门联合发布的《国家人工智能产业综合标准化体系建设指南（2024 版）》提出规范数据采集、数据标注、数据治理、数据质量等标准。2025 年政府工作报告更是提出，要持续推进“人工智能+”行动，将数字技术与制造优势、市场优势更好结合起来，加快完善数据基础制度，深化数据资源开发利用，促进和规范数据跨境流动。此类政策反映出，数据建设从“大水漫灌”进入“精准滴灌”的新阶段。

另一方面，数据场景应用过程中，行业对高质量数据集的需求也从通用、粗颗粒度的网络采集数

据，转向深耕垂直行业的场景化数据集，医疗病历、工业传感器序列、法律判决逻辑链等具有深度专业壁垒的行业专识数据蕴含着新的发展机遇与挑战。新的趋势下，行业评价体系也系统性升级。如今数据价值的衡量标准，已从单一的存储容量（GB/TB），转向涵盖信噪比、知识覆盖率、时效性、合规性等维度的综合质量指标。中国首部针对数据资产价值评估的国家标准《信息技术大数据数据资产价值评估》（GB/T46353—2025）就构建了“质量－成本－应用”三维评价体系，不仅关注数据的内在价值，还考虑其在具体业务场景中的使用价值。构建与推广高质量、场景化的专用数据集，并通过标准化的评估体系确保其效能，正成为推动“人工智能+”落地、释放数据要素乘数效应的关键路径。

3.2 技术前沿：AI Ready 的数据基础设施升级

为高效生产 AI-Ready 数据，整个数据技术栈正经历体系化重构。

数据处理方面，非结构化数据的识别解析不断突破。麻省理工学院斯隆商学院（MIT Sloan School of Management）一份研究指出，占比高达 80%至 90%的高价值知识蕴藏于 PDF、研报等“暗数据”（Dark Data）中。这些数据的存在要求更加智能化的数据治理方法，以有效降低海量非结构化数据的处理成本，提升数据质量。在此背景下，MinerU、PaddleOCR-VL、DeepSeek-OCR 等新一代智能文档解析产品基于视觉大模型的智能解析技术，超越传统 OCR（Optical Character Recognition）技术局限，在理解复杂文档布局，精准还原跨页表格、数学公式与代码结构方面表现出色，成为将人类知识转化为机器可理解信息的“咽喉要道”，为可靠的检索增强生成（RAG）系统奠定基石。

数据生成方面，合成数据成为破解“数据荒”的战略高地。合成数据指利用统计分布、模拟仿真、深度学习、强化学习等方式合成的特定属性、极端条件、实时交互等数据，可以拓展数据供给范围，补足由于物理限制、伦理约束、安全保护等造成的特定数据缺口。合成数据的应用前景主要聚焦于两大方向：一是生成蕴含复杂思维链的逻辑推理数据，以提升模型的认知与推理能力；二是在自动驾驶、机器人等物理交互领域，通过高保真仿真模拟生成海量、多样的长尾场景数据，以弥补现实世界数据采集的不足与风险。同时，“数据飞轮”也提供了数据生产的新方式。行业共识认为，企业通过搭建自动化的“数据飞轮”，构建“AI 生产 AI 数据”的闭环，利用强模型标注数据、基于业务反馈自动修正数据集，可以大幅降低高质量数据供给的成本与门槛，显著提升数据生产效能。

基础设施层面，向量数据库已成为 AI 应用的标准组件。向量数据库通过将数据转化为向量嵌入，充当大模型的“外挂海马体”，可以利用相似度检索技术在亿级数据中实现毫秒级精准匹配，为大模型提供高效的长时记忆与知识检索能力，从而改善传统关键词检索的“语义鸿沟”困境。同时，随着原生多模态大模型的普及，数据处理的前沿正聚焦于解决图文音视频的强对齐问题，实现跨模态信息在时空维度上的精准关联与理解，以释放多模态智能的全部潜力。目前，中国向量数据库行业蓬勃发展，不仅市场广阔，一些产品也发展迅速。2025 年 9 月，清华研发团队孵化的数智引航 VexDB 向量数据库发布后在国际权威的 DABSTEP 非结构化数据分析测试中夺冠，成为首个在该测试中取得第一的国产向量数

数据库。DABSTEP 基准测试最新榜单中，前五名中有两名来自中国，体现中国向量数据库的发展潜力。

 **DABSTEP Leaderboard**

The [Data Agent Benchmark for Multi-step Reasoning \(DABSTEP\)](#) is looking to measure and push the state-of-the-art in Data Analysis by LLMs. The benchmark is composed of ~450 data analysis questions ([Dataset Link](#)) centered around 1 or more documents that agents will have to understand and cross reference in order to answer correctly.

We have set up a notebook to quickly get an agent baseline using the free Huggingface Inference API: [Colab Notebook](#)

Check out the official technical reports here: [Adyen Report](#) [Hugging Face Report](#)

Join the discussion on the [discord server](#)

Reproduce the baseline results with the agent code we open sourced [here](#)

[Validated](#) [Unvalidated](#)

Agent	Easy Level Accuracy (%)	Hard Level Accuracy (%)	Organization	Repo URL	Model Family	Date
CambioML energent.ai DS Agent	94.44	57.67	CambioML user lingjiekong	Link	GPT-5	19-09-2025
DS-STAR	87.5	45.24	Google Cloud AI Research	Link	Gemini-2.5-Pro	08-07-2025
Amity DA Agent v0.1	88.56	41.01	Amity Solutions Thailand user aongwachi	Link	gemini-2.5-pro-preview-05-06	31-05-2025
AgenticData	94.44	40.48	Tsinghua University	Link	Qwen 3	03-08-2025
Mphasis-12I-Agents	88.56	28.04	Mphasis Limited user PranavSatheesan		claude-3.5-sonnet-20241022	10-04-2025
Claude 4 Sonnet ReACT Baseline	81.94	19.04	Hugging Face	Link	claude-sonnet-4-20250514	28-05-2025
Open Data Scientist	84.72	16.4	TogetherAI user federicotogther	Link	DeepSeek-V3	24-06-2025
o4-mini Reasoning Prompt Baseline	76.39	14.55	Hugging Face	Link	OpenAI o4-mini	22-04-2025

图 3.1 DABSTEP 基准测试榜单（截至 2026 年 2 月）

（来源：<https://huggingface.co/spaces/adyen/DABstep>）

3.3 数据生态：开源创新与数据治理

当前，AI 数据生态呈现出“开源与商业并重，流通与安全共存”的复杂格局。技术构成了骨架，而开放协同的生态则为其注入血肉，共同推动智能应用从中心化垄断走向普惠与多样化。

开源协同正经历从“代码共享”到“数据共建”的范式扩展。中国信息通讯研究院研究指出，开源社区是目前高质量数据集的主要供给方。以 Hugging Face、OpenDataLab 为代表数据开源开放的平台，正推动高质量数据集像开源软件一样，由开发者社区共同完成清洗、标注与版本维护，为基础的人工智能模型任务提供数据支持。与此同时，MinerU 等聚焦于文档智能解析的开源工具链，显著降低了企业处理私有数据的门槛，促进了 AI 能力在更广泛产业场景中的落地。作为“中国方案”的 MinerU 等智能文档解析产品，已经走在世界前沿，多次登顶 GitHub Trending-Python 等权威榜单第一名，受到 OpenClaw 等全球开发者的欢迎。

在数据要素流通层面，制度设计与技术进步结合互补。顶层设计中，中国“十五五”规划明确指出要“深入推进数字中国建设，健全数据要素基础制度，建设开放共享安全的全国一体化数据市场，深化数据资源开发利用”，为数据流通提供清晰的政策引导。而在具体的技术端，数据流通利用技术发展日趋成熟，隐私计算与新型架构正成为打破“数据孤岛”的关键。联邦学习、可信执行环境、数据沙箱等技术基本满足各行业基本需求，金融、医疗等领域的敏感数据得以在“可用不可见”的前提下安全融合，释放价值。企业内部则基于数据研发运营一体化（DataOps）理念，将敏捷、精益等理念融入数据开发过程，通过对数据相关人员、工具和流程的重新组织，构建智能化的数据编织网络，打破协作壁垒，实现对跨云、跨系统数据资产的自动发现与集成，为 AI 应用提供统一的、可治理的数据服务入口。

面对 AI 生成内容的爆发，数据治理体系面临严峻挑战，正向合规与确权深化。业界正积极探索基于区块链的数据溯源技术，以清晰界定训练数据的版权边界，保障数据贡献者的权益。同时，为应对潜在的“数据投毒”攻击，行业正在建立自动化的安全扫描与防御机制，像查杀病毒一样识别并隔离有害数据，筑牢高质量数据供给的安全防线。

3.4 本章小结

展望 2026 年及未来，AI 数据领域面临的核心挑战正变得日益严峻。数据同质化与质量污染问题仍然存在，互联网中充斥的 AI 生成“注水内容”使得区分“人类原声”与“机器回声”愈发困难，模型若大量训练于此可能导致“模型崩溃”，损害其根本的推理能力。算力与数据的匹配瓶颈亦不容忽视，数据处理与传输的 I/O 速度正逐渐落后于 GPU 的计算增速，“存储墙”问题成为制约效率提升的突出矛盾，近期对存储技术发展的关注显著增长，预想今年会有一定突破。

然而，挑战之中孕育着明确的结构性的机遇。一方面，伴随人形机器人的发展，来自物理世界的触觉、本体感觉与深度视觉等具身智能数据将成为驱动下一代 AI 进化的全新“石油”。另一方面，沉淀于企业内部数十年的工单、手册、邮件等私有知识资产，一旦被有效清洗与结构化，将转化为构建具有核心竞争力的企业级垂直模型的宝贵养料。

人工智能竞争的下半场，本质上是数据质量的较量。谁能更高效地获取、治理与利用高质量数据，谁就将掌握驱动 AI 持续进化的钥匙。对中国而言，依托庞大的实体产业场景与海量数据资源，通过系统构建领先的数据基础设施与繁荣的开源协同生态，在“数据要素×人工智能”的全球竞赛中，打造自身的战略优势。

本章参考资料

[1]国家数据局.高质量数据集建设指引[R].北京：国家数据局，2025.

[2]国家互联网信息办公室.“数据要素 x”三年行动计划(2024—2026 年)[EB/OL]. (2024-01-05) [2026-02-03].https://www.cac.gov.cn/2024-01/05/c_1706119078060945.htm.

[3]工业和信息化部.国家人工智能产业综合标准化体系建设指南(2024 版)[EB/OL].(2024-06-05)[2026-02-03].https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2024/art_e8ebf5600ec24d3db644150873712c5f.html.

[4]新华网.2025 年政府工作报告[EB/OL].(2025-03-12)[2026-02-03].<https://www.news.cn/20250312/118dc0c76109410c962cf76a3ec1ff3e/c.html>.

[5]中国国家标准化管理委员会.信息技术大数据数据资产价值评估: GB/T 46353-2025[S].北京:中国标准出版社，2025.

[6]MIT Sloan School of Management.Tapping the power of unstructured data[EB/OL].(2021-0

2-01)[2026-02-03].<https://mitsloan.mit.edu/ideas-made-to-matter/tapping-power-unstructured-dat>

a.

[7]中国信息通信研究院.数据要素发展报告(2025 年)[R].北京: 中国信息通信研究院, 2025.

[8]国务院公报.中共中央关于制定国民经济和社会发展第十五个五年规划的建议[N],2025.

第 4 章 AI 基础设施：“四力”筑基与智算生态重构

随着生成式人工智能（Generative Artificial Intelligence, GenAI）从技术爆发期转向规模化落地期，AI 基础设施已成为决定国家竞争力与企业创新上限的核心底座。通过构建“算力、存力、运力、电力”四力模型框架，本章系统梳理了 AI 基础设施的内涵与外延；深入探讨了智算服务从“物理设备”向“词元（Tokens）”计费模式的范式演进，以及 NeoCloud（算力原生云）等新型商业模式的崛起。通过对中美两国在算力、电力及产业链布局上的对比分析，本章指出尽管美国在核心芯片领域保持领先，但中国在电力保障、应用场景驱动及国产智算生态建设方面正逐渐形成独特优势。中国 2025 年电力消耗突破 10 万亿度大关，是有史以来世界上某个国家用电量首次突破这个界限，也标志着 AI 竞争进入下半场。在“东数西算”与“AI+”战略推动下，中国正通过全栈自主可控路径，实现从算力追随者向生态构建者的转变。

4.1 AI 基础设施的界定

1.AI 基础设施内涵：四力模型框架

人工智能算力基础设施不再是传统数据中心的简单延伸，而是基于异构计算架构，以智能计算为核心，面向大规模模型训练与推理需求的公共基础设施。算力概念的内涵远比电力复杂，因此，不妨创造一个新单词，用 computility 来表示算力，其中的 compu-是计算的词根，表达“算”的含义，utility 则是效用，实用的意思，所以 computility 可用来表达计算的能力，即算力。为了更科学地衡量其供给能力，我们结合业界发展趋势，逐渐形成了“四力”模型的定义：



图 4.1 “四力”模型

(1) 算力（Computility）：指基础设施提供的处理能力。目前已从通用 CPU 算力转向以 GPU、NPU 为核心的异构智能算力。其核心指标已从传统的 FLOPS 转向“Tokens/s”及针对大模型优化的有效吞吐量。全球算力高速增长，算力成为全球科技竞争的战略制高点。

(2) 存力（Storage Capability）：AI 训练对数据读写速度有极高要求。现代 AI 存力强调高带宽

内存（HBM）与高性能闪存（NVMe）的结合。例如，NVIDIA 推出的 ICMS 架构通过 RDMA 加速的上下文存储，将推理效率提升了 5 倍。

（3）运力（Network Capability）：智算中心内部数万颗 GPU 之间的互联带宽及中心间的传输能力。AI 基础设施需要极低时延、高带宽的运力（如 InfiniBand 或 RoCE）来支撑分布式训练，确保模型在跨节点同步时不产生停顿。

（4）电力（Electricity Power）：电力已成为 AI 发展的“第一瓶颈”。例如：微软为 OpenAI 训练 GPT 5 时，10 万卡集群导致美国一个州电力崩塌，导致 GPT5 发布的一拖再拖。AI 机架的功率密度已从传统云中心的 2-10kW 跃升至目前的 140kW，NV Rubin 芯片向 600kW 迈进，未来甚至可能达到 1MW。电力的可用性、交付速度及绿色能源占比，直接决定了智算中心的扩张上限。

2.AI 基础设施外延：数据中心配套与安全

（1）冷却系统：高密度算力带来了巨大的散热挑战。液冷技术（包括冷板式和浸没式）正迅速替代传统风冷。2025 年，全球液冷渗透率将增长至 26%，中国冷板式液冷市场规模达到 177 亿元。

（2）机密计算与安全：随着多租户共享基础架构的普及，数据的机密性成为核心痛点。机密计算依托于硬件层面的可信执行环境（TEE），确保数据在活跃计算阶段（In-use）处于加密隔离状态，防止特权管理员或跨租户攻击窃取 AI 模型与敏感数据。

总之，AI 发展趋势下，以四力为内涵的 AI 基础设施正在重构全球 IT 基础设施，同时制冷和安全带来严重的挑战，用得起是客户选择的关键点。

4.2 AI 基础设施发展趋势

1.AI 智算服务的升级：从“硬核”到“应用”

智算服务的交付模式正在经历从底层硬件向抽象化服务的演进：

裸金属服务器（Bare Metal）：提供对物理硬件的直接访问，无虚拟化损耗，适合大规模预训练。裸金属比虚拟化 vGPU 在训练大型 Transformer 模型时性能可高出多达 30%。

虚拟化 vGPU：通过切分 GPU 资源，满足中小规模推理和开发需求，具备更好的弹性与成本效益。

Tokens 计费：这是目前最前沿的演进方向。Tokens 已成为 AI 时代的“算力货币”。开发者不再购买显卡时间，而是购买模型处理的词元量。

2.商业模式的变化：NeoCloud 与 MaaS

智算集群：传统的资源出租模式。

NeoCloud（GPU Cloud）：这是一类新兴的 AI 原生云服务商（如 CoreWeave, Lambda）。它们没有传统云服务的历史包袱，专注于提供极大规模的 GPU 池化资源。由于运营极简且垂直化，其价格通常比传统超大规模云厂商低 60%-70%。

MaaS（模型即服务）：智算中心提供从硬件适配到算法调优的全栈能力。模型不再是独立的软件，

而是基础设施的一部分，通过 API 接口直接支撑应用场景。

4.3 AI 基础设施产业链概况

AI 基础设施进入“三位一体”升级阶段，超大规模集群化、绿色低碳化、高速互联化。算力规模正从“万卡级”向“十万卡级”演进，突破超节点技术，实现近线性算力扩展，以支撑千亿至万亿参数大模型的训练需求。

1.算力：硬件国产化加速

2024 年中国加速芯片市场规模超过 270 万张。尽管 GPU 占据 70%的市场份额，但本土 AI 芯片（如华为昇腾 910C 等）出货量已超过 82 万张，市场渗透率从 15%提升至 30%。2025 年加速芯片从 30%上升到近六成市场份额，华为独占 40%，国产算力产业链正经历上下游共振，逐步适配 DeepSeek 等主流大模型，加速自主可控进程。

从海外投行 Bernstein Research 发布的《China AI accelerators market share》中，报告显示，2025 年华为以 102.68 亿美金销售额，占据国内 AI 加速器市场份额的 40%，与英伟达的 101.98 亿美金基本持平，断层领先于其他国产 GPU 厂商。

2.存力：HBM 与闪存的协同

AI 基础设施正建立多层存储架构。底层使用对象存储确保耐用性，顶层则使用基于 BlueField-4 等 DPU 驱动的上下文存储（Context Memory），以支持海量 KV Cache 的快速调用。

3.运力：从数据中心内互联到国家算力网

国内正通过“东数西算”工程打通算力调度动脉。中国算力平台已接入山西、辽宁、上海等 10 个省市，实现了算力资源的跨域流通与弹性调度。

4.电力：AI 发展的“战略变量”

电力保障已从辅助设施转变为核心竞争力。领先的算力中心开始利用油田伴生气、风能等“闲置能源”进行离网发电。在中国，由于电价受到监管且清洁能源供应充足，相比美国，中国的电力保障具有显著成本优势。

4.4 AI 基础设施分布和对场景、产业的支撑

1.智算中心分布情况

中国已初步形成“全局总览、分域协同”的布局。

东部地区：广东、上海、江苏等地利用产业集聚优势，重点部署推理和高时延敏感业务。例如，广东 23 个智算中心主要集中在广州、深圳等 8 个地市。

中西部地区：内蒙、宁夏、青海、山西等地利用能源与气候优势，承接大规模模型预训练任务。

2.支持场景与应用

具身智能：支持机器人在复杂物理环境中的实时感知与决策（如宇树科技、星动纪元等）。

AI for Science(AI4S)：AI 基础设施加速了药物研发、气象预测和新材料发现。例如，英矽智能结合 AI 与量子计算，将新药靶点发现效率提升了 17 倍。

智能体 workflow (Agentic Workflow)：AI 基础设施正从简单的对话交互转向支持能够自动规划任务并调用工具的“数字化员工”。

3.存在的问题与建议

问题：区域性供需不匹配，部分西部数据中心上架率仍低于 50%；核心芯片受出口管制影响，先进制程受限；电力网架结构在高密度负荷下的韧性不足。

建议：建议加强“产教融合”，利用开源生态（如飞桨）培养跨学科人才；同时，应推动算力资源的标准化定价（如 Token 单位化），引导社会资本从“盲目建中心”转向“深耕应用场景”。

4.5 中美 AI 基础设施产业链比较

表 4.1 中美 AI 产业链结构对比

维度	美国 AI 基础设施产业链	中国 AI 基础设施产业链
算力驱动力	私营资本驱动，2024 年私营 AI 投资达 1091 亿美元, 是中国的 12 倍。	政府规划与国企引导，通过“东数西算”实现国家统筹布局。
硬件生态	NVIDIA,AMD,Intel 垄断高性能计算市场，技术领先 1-2 代。	华为、中芯国际、象帝先等国产阵营快速崛起，渗透率达 30%。
算力云	NeoCloud 蓬勃发展，CoreWeave 等获得巨额融资并签约微软、OpenAI。	传统云厂商（阿里、腾讯）叠加运营商（移动、电信）构成智算服务主体。

中美“四力”及场景优劣势比较分析如下：

1.算力

美国在先进制程芯片和系统级设计（NVLink 等）上具有显著领先地位。中国在模型推理效率优化上表现优异，例如深度求索（DeepSeek）的算力能效比远超美国同行。

2.电力

中国具备巨大优势。美国电网老化，数据中心建设速度受限于电力容量审批（通常需 3 年）；中国具备更强的基建速度（“周末盖医院”式速度）且工业电价更具竞争力。

3.存力/运力

美国领先于高性能互联协议。中国则在全国一体化算力网调度方面具有制度优势，正通过国家算力分平台实现跨区域资源“一盘棋”。

4.数据/场景：

中国拥有更丰富、更开放的工业与社会应用场景（AI+），中国在“数据作为战略资产”的集中度上优于美国。

4.6 本章小结

AI 基础设施正经历从“资源消耗型”向“价值创造型”的转变。衡量成功标准已从单纯的算力堆砌转变为“Tokens 每瓦每美元”（Tokens per Watt per Dollar）。未来十年，中国应利用电力保障与场景密度的优势，通过全栈国产智算链的协同（自主设计、自主制造、自主应用），在“战略相持”中寻找突破口，最终依托“AI+量子”等颠覆性技术实现全球科技竞争的“战略反攻”。

本章参考资料

- [1]国家能源局."2025 年全社会用电量同比增长 5.0%"[EB/OL].(2025-01-17)[2026-02-10].<https://www.nea.gov.cn:10443/20260121/715f79826488476a9162da7c8bd77c80/c.html>
- [2]孙凝晖,张云泉,刘宇航.算力（Computility）[J].中国计算机学会通讯,2022,18(12):106.
- [3]华金证券行业研究部.电子行业周报：AI&半导体：液冷渗透率提升，FAB稳健成长[R/OL].华金证券,2025.(2025-08-10)[2026-02-10].https://pdf.dfcfw.com/pdf/H3_AP202508101724810381_1.pdf
- [4]刘烈宏.截至 6 月底日均Token消耗量突破 30 万亿，1 年半时间增长 300 多倍[EB/OL].(2025-08-14)[2026-02-10].<http://finance.sina.com.cn/stock/bxjj/2025-08-14/doc-infkxinm5527663.shtml>
- [5]中国信息通信研究院.人工智能算力基础设施赋能研究报告（2025 年）[R/OL].北京:中国信息通信研究院,2025.(2025-11-06)[2026-02-10].<https://www.caict.ac.cn/kxyj/qwfb/ztbg/202511/P020251106555844142999.pdf>
- [6]伯恩斯坦.China AI Accelerators Market Share[R].伯恩斯坦.
- [7]资产信息网，千际投行.东数西算项目实施意义与产业链分析报告[R]
- [8]McKinstry.Designing for an Energy-Constrained Future:Four Forces Reshaping Data Center Infrastructure[EB/OL].[2026-03-26].<https://www.mckinstry.com/insights/designing-for-an-energy-constrained-future-four-forces-reshaping-data-center-infrastructure/>
- [9]Snider S.2026 Predictions:AI Sparks Data Center Power Revolution[EB/OL].Data Center Knowledge.[2026-01-13].<https://www.datacenterknowledge.com/operations-and-management/2026-predictions-ai-sparks-data-center-power-revolution>
- [10]Zilberstein E.Introducing NVIDIA BlueField-4-Powered Inference Context Memory Storage

Platform for the Next Frontier of AI[EB/OL].NVIDIA Developer Blog.(2026-01-06)[2026-02-13].<https://developer.nvidia.com/blog/introducing-nvidia-bluefield-4-powered-inference-context-memory-storage-platform-for-the-next-frontier-of-ai/>

[11]AceCloud.Bare-Metal vs Virtual GPU:Which Is Better for AI Workloads[EB/OL]

[12]Data Centre Magazine.US vs China:Nvidia's Jensen Huang on Data Centre Blockers[EB/OL].(2025-12-08)[2026-01-13].<https://datacentremagazine.com/news/us-vs-china-nvidias-jensen-huang-on-data-centre-blockers>

[13]Maslej N.Artificial Intelligence Index Report 2025[R/OL].Stanford:Stanford University Human-Centered Artificial Intelligence,2025.(2025-04-07)[2026-01-13].<https://hai.stanford.edu/ai-index/2025-ai-index-report>

[14]ABI Research.Mapping the Neocloud Market Landscape:Definitions,Technology Characteristics&Business Models[R].New York:ABI Research,2025.

[15]MLQ.ai Research.NeoCloud Infrastructure:Specialized GPU Cloud Providers Analysis[EB/OL]

第 5 章 AI 芯片：算力溢出效应下的半导体价值重构

从全球信息技术产业的历史格局审视，半导体长期扮演着“从属型基建行业”的角色——它为互联网平台、企业软件及云服务提供底层计算基座，但产业利润的核心增量过往多被上层的平台经济与 SaaS 企业俘获。

FactSet 及 S&P Dow Jones Indices 的历史统计数据显示，在 2010 年至 2022 年的十余年间，标普 500 半导体板块的每股收益（EPS）在 IT 行业总额中的占比常年波动于 15% 左右，显著低于软件与服务板块。直至 2023 年 AI 产业周期开启，这一结构性特征才发生剧变：半导体板块的 EPS 贡献率首次逼近并于 2024 年突破了 20% 的历史常态区间。究其原因，生成式 AI 对高端 GPU 的极端需求，驱动了以英伟达（NVIDIA）和博通（Broadcom）为核心的头部企业盈利爆发式增长；仅 2024 年第四季度，半导体板块便贡献了 IT 行业近 34% 的盈利增长点，正式打破了其在行业内部利润分配中“长期低于 20%”的历史规律。

AI 时代正在从根本上颠覆传统的产业分工格局，其本质是 Scaling Law 对半导体物理实现与产业权力结构的强制重构。

统计数据显示，仅在 2022 至 2024 的两年内，标普 500 半导体板块的 EPS（每股收益）在 IT 行业中的占比便从不足 20% 飙升至约 40%，且仍维持加速上升态势。这一结构性变迁揭示了一个深刻的范式转向：当算力（本质上是半导体的物理折旧）跃升为 AI 产业的核心生产要素时，半导体板块正迅速从“基建的从属者”蜕变为“算力的收租者”。

与互联网时代由 CPU 驱动的通用计算堆叠模式截然不同，大语言模型（LLM）的技术演进迫使半导体架构向应用端高度收敛。这不仅是一场“从算力丰裕向全链路瓶颈突破”的范式跃迁，更是一次对半导体在计算、存储、通信互联上功能定义的全面解构。从晶体管逻辑、高带宽内存（HBM）到高速互连网（Fabric），每一个将“电流”转化为“Token”的物理环节都在被重新定价。

这种重构效应呈现出动态演进的特征，2025 年后的典型例证包括：

针对推理效能的底层解构：为应对“推理法则”（Inference Law）中计算强度与访存带宽的极度失衡，芯片厂商通过 TensorRT-LLM 等软件栈深度介入底层的指令调度。通过实现 Prefill（预填充）与 Decode（解码）的分离执行架构（Disaggregated Inference），芯片能够在物理层面针对不同阶段的计算特性（计算密集型 vs. 访存密集型）进行动态资源对齐，极大地提升了 Token 的生成吞吐率。

针对复杂逻辑的范式回归：Agentic 模式的兴起标志着计算范式从“单向生成”转向“循环推理”。在这一工作流中，CPU 摆脱了长期的边缘化地位，重新回归为系统的智能调度中枢。Agent 模式下频繁的逻辑分支、外部工具调用及多步序列规划，对 CPU 的分支预测能力提出了苛刻要求。为消除 CPU 逻辑处理与 GPU 张量计算间的通信瓶颈，以 NVIDIA Grace Hopper 为代表的架构通过实现内存统一寻址与高速 C2C 互联，确保了“电变 Token”的过程在复杂的长任务链条中依然能保持高效协同。

针对存算瓶颈的物理跨越：为突破制约 Scaling Law 的“内存墙”，存储板块已从后台的“数据仓库”跃升为“算力的肌肉延伸”。以 HBM3e 为代表的高带宽内存通过 3D 堆叠封装与计算核心实现近延时耦合，这种从通用 DRAM 向特化型算力内存的转型，使得存储环节在芯片总成本中的权重显著抬升。

针对集群尺度的织网重构：当单一芯片无法承载模型规模时，互联（Interconnect）便从单纯的通信附件演变为“机架级计算机”的神经系统。通过 NVLink 以及针对 AI 优化的 RDMA 网络协议，算力在数以万计的集群节点间实现了逻辑统一，驱动着交换芯片及高性能背板互联技术的全面价值重估。

这种由应用层算法逻辑倒逼硬件物理设计的“持续自我重构”，正产生一种强大的“算力溢出效应”（Spillover Effect）。它将芯片产业从传统意义上线性、离散的供应链，彻底改造为一张深度耦合、持续演进的价值网络。

在这张网络中，软硬件的物理边界已然模糊，技术主导权正无可争议地向那些能够实现“算法－架构－系统”全栈协同的领航者集聚。这种从“零部件供应”向“产业逻辑定义”的权力跃迁，不仅是对半导体产业利润分配格局的底层重塑，更是 AI 时代算力霸权向物理层回归的终极注脚。

5.1 算力芯片：不仅是计算引擎，而是数据中心的“大脑总管”

在 AI Scaling Law（规模法则）的驱动下，算力芯片的竞争已经从“比拼谁的单颗芯片跑得快”演变为“比拼谁能管理好上万颗芯片的集群”。

这种设计哲学的转变体现在：芯片不再是一个孤立的零部件，它必须自带强大的通信和内存管理能力，才能在“集群级系统工程”中发挥作用。例如，英伟达的 GB200 已经不再以“单张显卡”的形式售卖，而是以包含 72 颗 GPU 的整个机架（Rack）作为基本计算单元。这意味着，单体性能的提升如果不能服务于集群的整体吞吐，就失去了商业意义。

5.1.1 计算 GPU/CPU：逻辑调度与张量执行的协同进化

1.GPU 现状：从单体卡到机架式计算机

市场主导与迭代节奏：英伟达在全球 AI 加速芯片市场占据约 92% 的份额。其产品迭代已进入“一年一更”的超常节奏，这不仅决定了全球 AI 算力总量的供给速度，更定义了行业的技术步频。

Blackwell 架构的里程碑意义：当前旗舰 Blackwell 系列 (B200/GB200) 在 FP8 精度下可达 20 PFLOPS，较前代 Hopper 架构实现了约 4 倍的性能跃迁。这种提升并非单纯依赖晶体管堆叠，而是通过集成的第二代 Transformer 引擎动态调整精度，直接回应了大模型对极大规模参数吞吐的需求。英伟达的 Blackwell(GB200)架构标志着 GPU 已彻底演变为系统级产品。它通过 NVLink-C2C 技术将 CPU 与 GPU 物理耦合，使得数据交换不再受限于传统的接口瓶颈，从而支撑起万亿级参数的实时推理能力。

路线图预判（2026）：根据已披露的 Rubin 架构路线图，算力芯片将率先采用台积电 2nm 工艺并集成下一代 HBM4 内存。在互联层面，NVLink 的单向带宽预计从 900 GB/s 翻倍提升，通过 NVLink Switch 实现多达 576 张 GPU 的直接逻辑互联，标志着“机架即计算机”时代的全面到来。

2. 2025-2026 年 GPU 技术架构的演进前沿

PD 分离（Prefill/Decode 分离）：2025 年的算力集群普遍采用了解构式推理（Disaggregated Inference），由于 Prefill 阶段为计算密集型（Compute-bound），而 Decode 阶段为访存密集型（Memory-bound），传统架构的资源共享会导致极大的利用率（MFU）损失。PD 分离允许在集群维度上将不同性能特征的芯片分别处理这两个阶段，利用 TensorRT-LLM 等软件栈，系统将计算密集型的预填充任务与访存密集型的解码任务分配给不同特性的芯片，将整体算力效能（MFU）提升了 30% 以上。这正是“硬件架构服务于算法逻辑”的典型证据。

DiT（Diffusion Transformer）驱动的底层重构：随着 2025 年 Hot Chips 大会及相关前沿进展，图形渲染正从传统的“物理模拟”转向“AI 推理生成”。DiT 架构的成功倒逼 GPU 必须打破“图形渲染”与“AI 推理”之间的物理藩篱，转向更纯粹的 AI 原生架构（AI-Native Architecture）。

逻辑缩放的终结与“小芯片”方案（Chiplet&CoWoS-L）：单体芯片的面积已逼近光刻掩模极限（Reticle Limit）。未来的算力增长将极度依赖 Chiplet 封装技术。未来的竞争不再是如何设计出最强的单颗芯片，而是如何通过 2.5D/3D 封装将更多的计算单元、HBM 和互联接口集成在一个中介层上。

晶圆级计算（Wafer-Scale Engine）的边缘化与启发：像 Cerebras 这样的晶圆级算力芯片虽然在特定场景下效率极高，但其商业化受限于成本。然而，它所代表的“内存与计算在同一物理平面上极度接近”的思想，正在深刻影响英伟达和自研 ASIC 链的设计，这也佐证了“存算一体/近存计算”的重要性。

解决“推理难题”的新物种—Groq 的确定性推理：针对 Inference Law，Groq 的 LPU 通过取消动态调度，实现了对数据流的微秒级精确控制。这种创新证明了，为了追求极致的推理速度，芯片可以舍弃通用的并行能力，转向纯粹的逻辑流控制。

3. Agentic 模式下的计算重构与 CPU 的价值回归

在传统的深度学习推理中，CPU 仅充当简单的流量管控角色。但在 Agentic 模式下，CPU 重新接管了逻辑指挥权。为了应对复杂的 API 调用和逻辑分支，NVIDIA Grace Hopper 架构实现了 CPU 与 GPU 的内存统一寻址，让“思考（逻辑）”与“执行（计算）”能够在同一个内存空间内无缝流转，彻底解决了 Agent 长任务链条下的延迟问题。计算范式从“单向生成”转向“循环推理”，这驱动了 CPU 职能的战略回归。

逻辑调度中枢：Agent 模式涉及大量的“推理 - 行动 - 观察”循环（Reasoning-Action Loop），其中非张量计算（如调用 API、逻辑判断、分支预测）在 CPU 上的运行效率远高于 GPU。

异构解耦与内存池化：为了降低响应延迟，以 NVIDIA Grace Hopper 为代表的架构通过超高带宽的芯片间互联（C2C）实现了 CPU 与 GPU 的内存统一寻址。这使得 CPU 不再是计算的旁观者，而是直接参与到复杂长任务流（Long-task Workflow）的实时调度中。

5.1.2 算力市场的格局与链条博弈

2025 年的算力芯片市场已演变为“一超多强”的格斗场。英伟达凭借 Blackwell 架构稳坐“集群计算”的王座，但以 Groq 为代表的 LPU 阵营通过极速推理切中了 Agent 时代的痛点；而以 Cerebras 为代表的晶圆级计算则在科研与超算领域开辟了无人区。

与此同时，地缘政治与成本压力促成了“自研 ASIC”与“本土替代”的双重爆发。阿里自研的 PTU(Parallel Processing Unit)展现了云服务商通过算法深度定义硬件的能力，而中国本土厂商如摩尔线程、沐曦与华为昇腾，则通过“CUDA 兼容生态”与“系统级集群能力”，在英伟达的高墙之外，构建了一套自主可控的第二算力极。

创业公司的“奇兵”与中国本土的“突围者”正在通过差异化竞争，在数据中心中抢夺特定的领地。

表 5.1 2025 年全球 GPU/算力芯片主流部署列表

阵营分类	厂商与主要型号	2025 年部署/商业现状	核心技术/竞争壁垒
“通用型 GPU”市场 主导者	英伟达 Blackwell (B200/GB200)	全球主流 CSP 的旗舰配置，机架级交付，用于万亿参数 LLM 训练与实时视频生成	软硬一体生态，NVLink 72 GPU 集群互联，万亿参数实时推理
自研 ASIC 巨头	Google TPU v7(Ironwood)	支持 Gemini 系列模型及 Vertex AI 平台，Anthropic 等大厂主力部署	原生 FP8 支持，万亿参数模型推理，192GB HBM 大容量内存，光学互联(OCS)与全栈软件链优化
自研 ASIC 巨头	Alibaba T-Head PPU(PTU)	阿里云 PAI 平台大规模替代部分 H20 订单	互联带宽 700GB/s，针对通义千问模型定制，实现推理成本下行
创业公司 (2025 年 12 月被英伟达收购)	Groq LPU(v2)	部署于 GroqCloud，主打极速实时对话、语音助手、高频交易 AI	确定性流水线调度，SRAM 近存计算
创业公司	Cerebras WSE-3(CS-3)	Condor Galaxy 超算集群，主打生物制药模拟、万亿参数超快推理	晶圆级计算，4 万亿晶体管，彻底消除芯片间延迟
中国本土	华为 Ascend 910C (昇腾)	国内智算中心主力，被百度、字节、阿里作为 Nvidia H100 的首选替代件	自研昇腾生态(CANN)，算力接近 H100 的 70%-80%，万卡集群 SuperPod 架构
中国本土	寒武纪思元 590/690	国内运营商、金融、互联网规模化部署	Chiplet 架构，思元 690 性能直追 H100 80%

中国本土	摩尔线程 MTT S4000 (夸娥)	国产 GPU 算力集群, 国产兼容算力代表	MUSA 架构, 高度兼容 CUDA, 支持国产大模型生态
中国本土	沐曦 (曦云 C600) / 壁仞(BR100/105)	进入商业化批量交付, 适配国产智算中心	训推一体, 支持大规模通用 AI 训练和推理

以下对创业公司、云服务厂商与中国厂商的“生存之道”进行深度分析：

1) 创业公司的“非对称作战”：Groq 与 Cerebras

这些公司不再试图在所有领域超越英伟达，而是攻击英伟达的弱点——延迟（Latency）与通信（Communication）。

Groq(LPU)：在 2025 年的崛起是因为它深刻抓住了 Agentic 模式的需求。Agent 需要极速响应，而 GPU 的 HBM 存取和动态调度会产生波动。Groq 采用纯 SRAM 方案，实现了真正的“确定性推理”，成为极速推理细分市场的冠军。

Cerebras（晶圆级）：它的逻辑是“如果互联是瓶颈，那就把整张晶圆做成一颗芯片”。2025 年，Cerebras 的 CS-3 系统在万亿参数模型推理上比 Blackwell 还要快 5 倍以上，这让它在国家级超算中心和尖端科研机构（如阿斯利康制药）中站稳了脚跟。

2) 云服务商的“下半场”：从算力租用者到定制化“芯片主权”的博弈

在 AI 算力的下半场竞赛中，超大规模云服务商（Hyperscalers）不再仅仅满足于向英伟达支付巨额的“算力租金”，而是通过自研 ASIC（专用集成电路）发起了一场旨在夺回定价权与技术主导权的防御战。

a.硬件反抗：GPU 定价权与 ASIC 的成本替代

英伟达的高毛利（约 80%）是对计算资源的高溢价销售，云服务商为打破这一垄断，通过自研芯片达到相比购买英伟达 GPU 更高的资本效率：

成本降维打击：以 Google TPU v7(Ironwood)为例，对于内部使用的模型训练，其 TCO（总拥有成本）仅约为英伟达 GB300 的 55%。

规模化替代：2026 年，Google 预计将量产部署超过 310 万至 320 万颗 TPU 芯片，通过大规模自用与外部租赁，显著稀释对英伟达单体硬件的依赖。

b.“生态绑定”下的算力自主可控：Google TPU 的标杆效应

Google TPU 的演进展示了如何通过软件（XLA 编译器）与硬件（OCS 光学互联）的深度耦合，构建起不同于 CUDA 的独立生态：

针对推理时代的重构：TPU v7 专为“推理时代”设计，其单片 FP8 算力达 4.6 PFLOPS，与英伟达 B200 旗鼓相当，但在大规模集群扩展上表现出更强的系统级优势。

跨公司的算力联盟:2026 年, 顶级 AI 实验室 Anthropic 计划部署上百万颗 TPU v7, 旨在通过 Google 的算力集群将 Claude 模型的训练与推理成本压低至极致。甚至 Meta 也被爆出可能在 2026-2027 年间租用甚至直接部署 Google 的 TPU, 标志着顶级厂商正集体寻求“去英伟达化”的替代方案。

c.云端 ASIC 的全面爆发：AWS 与阿里的中国路径

AWS 的性价比攻势：Amazon 的 Trainium 3 芯片通过原生 PyTorch 集成，使开发者无需修改代码即可迁移，宣称在提供相同推理性能时，比同代 GPU 的性价比提升了 50%。

阿里 PTU 的本土生态：在国内受限的制程环境下，阿里巴巴孵化的 PTU(Parallel Processing Unit) 展现了极强的生存智慧。它针对通义千问(Qwen)等大模型进行了算子级优化，片间带宽达 700GB/s。2025-2026 年间，PTU 在阿里云 PAI 平台的大规模应用，不仅确保了国产模型的推理性能，更有效对冲了由于 GPU 禁运带来的系统性风险。

3) 中国“自主算力链”的生存之道：从“单兵突围”到“集群化进化”

在 2025 年地缘政治博弈与 AI 范式转移的双重背景下，国产算力厂商已不再满足于零散的国产替代，而是形成了以“全栈生态（华为）”、“性能与成本拐点（寒武纪）”，以及“资本化加速的四小龙（摩尔线程、沐曦、壁仞、天数智芯）”为特点的三个生存模式。

a.华为昇腾：以“系统级工程”对垒单点性能

当单颗芯片在先进制程受限时，华为通过 SuperPod 集群架构来实现“以量取胜”的突破。

生存策略：核心在于 CANN 软件生态对标英伟达 CUDA，通过软硬一体化大幅降低模型迁移门槛。

2025 现状：昇腾已成为国内唯一能与英伟达 GB200 集群对抗的系统级厂商。在华为云数据中心，AI 算力占比已超 80%，其 SuperPod 架构能驱动上万颗芯片高效互联，是国内智算中心的生态基座。

b.寒武纪：从“亏损王”到“股王”的盈利质变

2025 年是寒武纪的“价值迸发时刻”，它证明了纯粹的 AI 原生架构（NPU）在推理时代的效率优势。

生存策略：利用 Chiplet（小芯片）架构规避先进制程限制，并深化与互联网大厂（如字节跳动）的深度定制合作。

2025 现状：凭借思元 590/690 成功切入 DeepSeek 等主流模型供应链。2025 年上半年营收暴增 43 倍达 28.8 亿元，并实现净利润扭亏为盈。其原生支持 FP8 计算的特性，使其在推理市场的溢价能力大幅提升，市值一度突破 6000 亿位居行业之首。

c.国产 GPU“四小龙”：资本驱动下的差异化突围

2025 年末至 2026 年初，国产 GPU 迎来史上最强“上市潮”。这四家公司通过不同的生存路径，在英伟达的高墙外构建了自己的利基市场。

表 5.2 国产 GPU“四小龙”的对比

厂商	生存路径与 2025 核心表现	资本/商业化节点
----	-----------------	----------

摩尔线程	生态为王：走“全功能 GPU”路线。其 MUSA 架构对 CUDA 的兼容性极高，大幅降低了开发者从英伟达迁移的成本	2025.12 登陆科创板。5 日涨幅超 7 倍，市值突破 4400 亿元
沐曦股份	商业化速度最快：专注于训推一体芯片。凭借高效的商业落地，在互联网与金融行业快速渗透	2025.12 A 股上市。刷新 A 股打新纪录，率先在 2025.Q2 实现单季盈利
壁仞科技	国家队背书与算力枢纽：凭借 BR100 的高性能算力密度，成为多地智算中心（尤其是国资背景）的首选替代	2026.01 港股上市。首日大涨 82%，在大规模集群部署上展现了深厚根基
天数智芯	先发优势与产品全线：拥有国内最早量产交付的通用 GPU。随着客户集中度下降，其生态广度在逐步扩大	2026.01 港交所上市。标志着其从初创期正式迈向成熟商业化阶段

2025 年后的国产算力突围路径呈现出“组局”的特征：核心是构建跨厂商、跨技术的开放协同生态。

从训练转向推理：随着 DeepSeek 等高效模型的兴起，算力需求从高能耗训练转向高周转推理。国产芯片（如寒武纪、沐曦）因能效比和特定算子优化，在推理侧展现出比英伟达 H20 更高的性价比。

软硬协同填补硬件短板：通过算法创新（如 FP8 数据精度优化、稀疏化、底层编译优化），以及 Chiplet 架构，不仅是应对先进制程限制的工程突围、弥补先进制程的物理差距，更是国产算力芯片从“可用”向“好用”迈进的标志性转折，这已成为国产厂商在受限环境下的生存手段。

5.1.3 总结：从单体计算性能向系统吞吐能力的范式漂移

算力芯片的演进规律表明：AI 时代的产业话语权已不再取决于单体晶体管的物理微缩，而取决于对计算、存储与互联（Compute-Memory-Fabric）三者协同效率的定义能力。

英伟达的系统化霸权：通过 Rubin 架构实现计算单元与 HBM4 的深度封装，英伟达成功将竞争维度从“单卡算力”提升至“机架级系统带宽”，迫使全球供应商必须进入其定义的物理迭代周期。

云服务商的结构性反击：Google TPU v7 等自研 ASIC 的大规模应用，证明了通过定制化架构（如专门针对长序列推理的显存调度优化），云服务商能够有效对冲通用 GPU 的溢价，构建起独立的成本与能效护城河。

这种设计哲学的位移，标志着算力竞争的核心指标已从单纯的 FLOPS（每秒浮点运算数）转向了 Token/s（每秒 Token 生成速率）。在这一转变中，计算核心的性能冗余与落后的数据传输速率之间形成了显著的“存算鸿沟”。

当算力芯片的执行频率远超存储介质的读写速率时，存储系统便不再是计算的从属配套，而是成为决定整个计算集群利用率（MFU）的决定性约束变量。

存储系统如何通过物理架构的重构（如 HBM 的 3D 堆叠）来缓解“内存墙”效应，并由此触发半导体价值链的二次分配，将是下一节论述的核心。

5.2 存储芯片：算力溢出效应的第一站

在 AI 算力体系中，存储芯片已从传统的“数据仓库”角色跃升为算力效能的决定性约束变量。AI 加速卡的有效算力不仅取决于计算核心的理论峰值，更受限於内存带宽（Memory Bandwidth）。当计算核心以每代 4 倍的速度提升时，DRAM 密度的增长在过去十年仅有约 2 倍，这种极度的失衡使存储环节成为算力价值溢出效应冲击的第一站。

5.2.1 HBM：算力效能的“配给证”与定价权转移

1.“内存墙”下的物理跨越

为了突破制约 Scaling Law 的“内存墙”，高带宽内存（HBM）应运而生。通过 3D 堆叠技术将多层 DRAM 晶粒垂直集成，并利用硅中介层（Silicon Interposer）实现与 GPU 的超宽带互联。

HBM3e/HBM4 的技术指标：当前主流的 HBM3e 已能提供超过 1.2 TB/s 的带宽，是传统 DDR5 的 10 倍以上。而即将量产的 HBM4 将进一步通过 2048-bit 超宽接口，将单叠层带宽推向 2.0 TB/s 的量级。

从通用向特化转型：这种从标准通用 DRAM 向特化型算力内存的转型，使得存储环节在芯片总成本中的权重显著抬升。

2.寡头格局与溢价逻辑

存储市场正经历“四十年一遇”的超级周期。由于 HBM 制造高度集中于海力士（SK Hynix）、三星和美光三家巨头，这种极高的行业壁垒赋予了供应商极强的定价权。

确定性涨价：海力士与三星已计划在 2026 年对 HBM3e 提价约 20%。

供需错配的结构性价重估：预计到 2026 年，全球 HBM 市场规模将达 546 亿美元，部分分析认为 2028 年 HBM 单品类市场规模将超越 2024 年整个 DRAM 市场。这种价格上涨并非周期性波动，而是由于 DRAM 单元结构深宽比已趋近物理极限（100:1）导致的成本重估。

5.2.2 NAND/SSD：推理时代 KV Cache 的意外受益者

如果说 HBM 是 GPU 算力溢出的直接受益者，那么 NAND 闪存则是推理侧需求爆发下的关键支柱。

KV 缓存（KV Cache）的压力测试：在大模型推理过程中，随着上下文窗口从 4K 扩展到 128K 甚至更长，单次推理任务产生的中间计算状态（KV 缓存）可达数百 GB。

高速存储的增量需求：为了维持推理低延迟，NAND SSD 必须提供极高的 IOPS 和带宽来支持频繁的检查点（Checkpoint）写入与读取。

消费级市场的挤压效应：2025 年以来，数据中心对 HBM/DRAM/NAND 的暴增需求占据了全球绝大部分存储供给，导致消费电子市场被迫面临涨价或减配——例如部分 PC 产品的内存配置被迫回调，而苹果等厂商则通过长约锁价来维持成本优势。

5.2.3 存储多元化：CXL 池化与边缘 AI 的物理前提

随着 AI 范式从云端训练向全场景推理延伸，存储需求已不再局限于单一的带宽提升，而是呈现出

显著的差异化特征。存储环节不仅在云端扮演“算力肌肉的延伸”，在边缘侧更成为决定设备智能上限的物理基石。

1.集群层面的“存算解耦”与 CXL 内存池化

针对超大规模集群中普遍存在的内存碎片化问题，CXL(Compute Express Link)协议正在实现内存的池化与共享。其核心意义在于通过物理层面的“存算解耦”，允许算力节点按需动态调用远端内存资源，从而进一步优化万卡集群的资源利用率（MFU）。

2.边缘侧的能效挑战与端侧 AI 的爆发

在智能驾驶、具身智能机器人及 AI 手机等场景中，对低功耗、高可靠性、小尺寸嵌入式存储的需求正迎来指数级增长。

具身智能的突破：以宇树科技(Unitree)为代表的四足机器人已占据全球约 70%的市场份额，其硬件生态对嵌入式存储和低功耗计算芯片的极度依赖，是算力溢出效应从数据中心向终端延伸的典型表现。

移动端的物理前提：在移动端，低功耗内存(LPDDR5X)的带宽提升已成为手机端侧运行大模型的物理先决条件。

5.2.4 中国存储势力的崛起与产业结构重塑

在全球存储市场被三星、SK 海力士及美光长期锁定的背景下，以长江存储（YMTC）和长鑫存储（CXMT）为代表的中国厂商，正在通过技术路径创新和本土供应链协同实现深层次破局。

1.核心厂商布局与技术高地

a.长鑫存储(CXMT)：DRAM 领域的国产基石

技术突破：作为中国首家量产 LPDDR5 的厂商，长鑫存储在 2025 年已实现对主流 AI 边缘终端（如手机、PC）的规模化供货。其自研的 18nm 及以下工艺，为国产 AI 算力链提供了底层存储的安全备份。

b.长江存储(YMTC)：NAND 架构的颠覆者

架构优势：凭借其独创的 Xtacking 架构，长江存储在 3D NAND 层数堆叠与 I/O 速度上已步入全球第一梯队，有效支撑了数据中心对超大容量、高性能 SSD 的迫切需求。

c.产业链长尾厂商：以兆易创新（Gigadevice）为代表的厂商，在 NOR Flash 及特定利基存储市场占据重要份额，补齐了 AIoT 领域的存储拼图。

2.对全球产业结构的深远影响

打破“价格托拉斯”，驱动成本民主化：中国厂商的产能释放改变了传统存储巨头通过减产人为调控价格的旧律。这种竞争压力迫使全球存储芯片进入更具性价比的区间，加速了 AI 推理成本的下行。

加速从单品竞争向“生态协同”演进：中国存储厂商不再孤立存在，而是与华为昇腾、寒武纪等算力厂商深度绑定。这种“国产算力+国产存储”的垂直整合模式，在物理层面上实现了更优的互联互通，确保了在受限环境下依然能通过系统级优化（如针对国产算子的存储预取技术）维持竞争力。

催化全球供应链的“多极化”重组：随着中国厂商在 3D NAND 和 LPDDR5 等前沿领域的份额提升，全球半导体供应链被迫从“美日韩”三角阵营向“多极并存”转化，显著提升了全球 AI 基础设施的供应韧性。

5.3 互联芯片：算力溢出效应的第二站

当存储系统通过物理架构重构缓解了“内存墙”效应后，系统内部各节点间的通信效率便上升为核心瓶颈。AI 系统的互联需求按物理距离可划分为四个严密的层级，每个层级对应着不同的技术路线与芯片生态。

表 5.3 四个互联层级的对比

互联层级	距离范围	主流技术	产业逻辑与关键厂商
片上互联 (On-Chip)	<1cm	Die-to-Die(UCle) 、 HBM互联	解决SoC内部算力与存储的近延时耦合， 主要由台积电（CoWoS）等封装巨头主 导
板间互联 (Board)	1cm—2m	NVLink、PCIe 6.0、C2C 互联	负责GPU显存统一寻址与CPU-GPU协 同，Nvidia Grace Hopper通过C2C实现内 存统一寻址，主要由英伟达、AMD、博 通主导
机架互联 (Rack)	2m—100m	CPO（共封装光学）、 硅光技术、DAC/AOC	突破铜缆物理极限（2m以上信号衰减）， 实现机架内与机架间的算力池化，主要厂 商：博通、中际旭创、新易盛
跨集群互联 (Cluster)	>100m	InfiniBand、RDMA、 以太网+高性能光模 块	构建万卡级智算中心的骨干网，驱动交换 芯片与光模块价值重估，主要厂商：英伟 达、Arista、华为

5.3.1 互联层级的深度重构：从铜到光的物理飞跃

1.片上与板间互联：存算解构的物理底座

在单一芯片无法承载模型规模时，板间互联通过高速总线将多颗 GPU 逻辑统一。以 NVIDIA 的 NVLink 为代表，其第五代带宽已达 1.8 TB/s，通过实现多达 576 张 GPU 的直接互联，消除了传统 PCIe 协议的调度延迟。同时，针对 Agentic 模式下的逻辑分支需求，C2C（Chip-to-Chip）技术让 CPU 与 GPU 共享统一内存地址空间，确保了复杂长任务链的高效协同。

2.机架互联：CPO 与硅光技术的临界点

当互联距离超过 2 米，传统的铜互联面临严峻的信号完整性和功耗挑战。这促使互联技术从“电”

向“光”发生历史性跨越。

硅光技术（Silicon Photonics）：利用 CMOS 工艺在硅基芯片上集成光收发功能，提升集成度并降低单位比特成本。

CPO（Co-Packaged Optics，共封装光学）：将光引擎（Optical Engine）与电交换芯片共同封装在同一基板上。通过消除可插拔光模块在 PCB 上的长距离走线，大幅降低了功耗（预计降低 30%以上）并解决了 SerDes 速率超过 224 Gbps 后的信号畸变问题。

3.跨集群互联：万卡时代的网织重构

在跨集群维度，互联协议的效率直接决定了集群的 MFU（模型算力利用率）。RDMA（远程直接内存访问）技术绕过 CPU 内核栈，在万计节点间实现了逻辑统一。这种“网织重构”驱动了 800G 甚至 1.6T 高性能光模块的爆发式需求，成为算力溢出效应在硬件链条上的第二站。

5.3.2 共封装光学(CPO)：互联领域最具变革潜力的技术方向

CPO 不仅是互联速率的升级，更是对数据中心拓扑结构的重塑。其核心思路是将光引擎（Optical Engine）与电芯片共同封装在同一基板上，消除了传统可插拔光模块在功耗、延迟和信号损失方面的限制。

CPO 的技术路线图显示了清晰的演进脉络：

从“分体”向“集成”回归：传统的 LPO（线性驱动可插拔光学）作为过渡方案，虽降低了延迟，但仍受限于 SerDes 的缩放极限。CPO 通过 Wide I/O 绕过复杂的 SerDes 调度，实现了光电转换在物理层面的“极致邻近”。

技术路径的博弈：当前业界呈现出三种主流调制技术竞争——马赫-曾德尔调制器（MZM）在成熟度上领先；微环调制器（MRM）以紧凑尺寸见长；而电吸收调制器（EAM）则提供了另一种架构选择。台积电推出的 COUPE 方案正成为光电芯粒集成的首选平台。

核心驱动力：当计算单元的执行速度不再是瓶颈，真正的“战争”便转移到了负责搬运数据的 Fabric 网络上。CPO 的落地标志着互联环节已从算力的“辅助配件”跃升为决定算力系统整体稳定性与经济性的核心变量。

英伟达已确认将在下一代产品中采用 CPO 技术，其“杀手级应用”在 Scale-up 网络中尤为突出：CPO 有望将 GPU 直连规模从当前 576 卡的上限大幅扩展，从根本上改变大规模训练集群的设计范式。

GPU 集群的扩张正在从量变走向质变。当训练集群从数百卡扩展到数万卡、当推理请求从串行处理进入大规模并发时，互联网络的带宽和延迟不再是“锦上添花”的优化项，而是决定集群实际效率的硬约束。铜线互联在距离超过 2 米后面临信号衰减瓶颈——在一个拥有数万 GPU 的集群中，铜缆的物理极限已经成为不可逾越的障碍。这正是算力溢出效应在互联层面的核心表现：GPU 越强、集群越大，对互联的要求就越高，互联芯片和光模块的价值就越被重新发现。

5.3.3 中国在互联领域的优势与短板

中国在光通信领域拥有显著的产业链优势。中际旭创、新易盛、光迅科技等企业在全球 400G/800G 光模块市场中占据领先地位，出口规模持续增长。然而，在 CPO 核心技术、高端交换芯片（博通垄断）和硅光制造工艺方面，国内仍有明显差距。从算力溢出效应的视角看，中国光模块企业目前是这一效应的直接受益者——GPU 集群的全球扩张驱动了对高速光模块的刚性需求。但 CPO 技术的普及可能改写供应链格局：当光引擎从可插拔模块变为封装内组件时，传统光模块厂商的角色将面临重新定义。未来的关键突破点在于：参与 CPO 标准制定和生态建设，在 1.6T 及更高速率光模块上保持技术领先，以及在国产交换芯片上实现中低端场景的替代。

5.4 芯片联盟：价值网络上的四大阵营

算力溢出效应正在将 AI 芯片产业从“英伟达一家独大”的单极格局推向多极化竞争。基于供应链关系和技术路线的差异，当前全球 AI 芯片生态可划分为四大阵营，每个阵营都在试图在这张由溢出效应重构的价值网络上占据最有利的节点。

1. 英伟达链：全栈生态的守城者

英伟达链是当前最完整、最强势的 AI 芯片生态。以英伟达 GPU 为核心，台积电（CoWoS 先进封装）为制造基座，SK 海力士/三星/美光提供 HBM，博通供应网络芯片，形成了从芯片设计、制造、封测到系统集成的全链条覆盖。CUDA 软件生态是这条链的核心壁垒——全球超过 400 万开发者构建在 CUDA 之上，千亿参数大模型的主流训练框架几乎全部原生支持 CUDA。从溢出效应的视角看，英伟达链的独特地位在于它既是溢出效应的“策源地”（GPU 性能升级驱动全链条需求），也是溢出效应最大的直接受益者（75%毛利率）。但也正是这种过于集中的价值捕获，催生了其他三大阵营的竞争性崛起。

2. 谷歌链：ASIC 商业化的开拓者

谷歌链以 TPU 为芯片核心、博通为 ASIC 代工伙伴、自研 OCS 光互联为网络基础，构建了与英伟达链平行的完整生态。TPUv7 的 TCO 优势（较 GB200 低 30%-44%）和 Anthropic 的大额订单，标志着这条链已从“谷歌自用”走向“开放商业化”。谷歌链的战略意义在于：它证明了在溢出效应重构的价值网络中，可以存在一条绕过英伟达 GPU 的替代路径。JAX/XLA 软件栈虽然生态规模不及 CUDA，但在科研和大模型训练领域已建立起可观的用户基础。如果 TPU 的外部化持续推进，可能触发英伟达估值逻辑从“垄断溢价”向“寡头定价”的转变。

3. 云厂商 ASIC 链：多元化的挑战者

微软、亚马逊和 Meta 等云厂商正在构建各自的定制化 ASIC 芯片能力，形成了分散但日益强大的第三阵营。微软的 Maia 100 加速器采用台积电 5nm 工艺、集成 1050 亿晶体管；亚马逊的 Trainium2 和 Inferentia2 已在 AWS 上商用；Meta 的 MTIA 芯片在内部规模化部署。这些自研芯片虽然单芯片性能通常不及英伟达旗舰产品，但凭借定制化优化和供应链自主性，正在成为云厂商降低对英伟达依赖

的战略支柱。从溢出效应的视角看，这一阵营的崛起本质上是价值网络中“下游”节点向“上游”的反向渗透——当芯片利润率远高于云服务利润率时，垂直整合成为理性的战略选择。

4.中国自主链：制约下的突围

在美国出口管制的倒逼下，中国正在构建独立的 AI 芯片生态体系。2024 年中国 AI 芯片市场规模达 1206 亿元，同比增长 41.8%，其中国产芯片销售额约 60 亿美元，预计 2025 年将翻倍至 160 亿美元。华为昇腾是当前的领头羊：昇腾 910C 采用 7nm 双 Die 封装，FP16 算力约 800 TFLOPS（约为 H100 的 80%），2024 年出货超过 64 万片，占国产 AI 芯片市场 23% 份额。华为的三年路线图显示，2026 年将推出首款搭载自研 HBM 技术的昇腾 950PR（算力 1 PFLOPS），以及面向大模型解码场景的昇腾 950DT（2 PFLOPS，144GB 显存，4 TB/s 带宽）。

寒武纪以 ASIC 路线深耕垂直场景，海光信息以 X86 兼容路线破解信创需求，摩尔线程、沐曦股份、壁仞科技等 GPU 新势力也在加速追赶。Bernstein Research 预测，到 2026 年英伟达在中国 AI 芯片市场的份额将从 54% 萎缩至 8%，华为将占据 50%。

中国自主链面临三大关键瓶颈：一是先进制程受限，7nm 以下工艺仍依赖海外代工；二是软件生态碎片化，华为 MindSpore、寒武纪 NeuWare 等各立标准，千亿参数模型从 CUDA 迁移需 1-3 个月适配；三是 HBM 等先进封装技术差距明显。但从算力溢出效应的视角看，中国自主链面临的最大机遇在于推理芯片赛道：推理 workload 的特征更适合 ASIC 定制优化，且不像训练那样对最先进制程有刚性依赖。正如云天励飞创始人陈宁所言：“在 AI 推理芯片上，全球厂商包括英伟达，几乎是站在同一条起跑线上。”随着 DeepSeek 等国产大模型对推理效率的极致优化，以及政策层面的窗口指导，中国 AI 芯片的国产化率有望在 2027 年达到 55%。

5.5 本章小结

回顾本章的分析，一条清晰的主线贯穿始终：AI 算力需求不只是带动 GPU 一个品类的发展，而是通过溢出效应重塑了整个半导体的价值分配。

在算力芯片领域，英伟达 GPU 的近乎垄断地位创造了前所未有的定价权，但也催生了 Hyperscaler 的集体反抗——谷歌 TPU 的百亿美元订单和四大云厂商的 ASIC 自研，正在将“垄断定价”推向“寡头竞争”。在存储芯片领域，GPU 算力的指数级增长暴露了 memory-bound 瓶颈，HBM 成为四十年一遇超级周期的主角，KV 缓存让 NAND 成为意外的受益者，物理缩放的极限使涨价从周期性波动变为结构性重估。在互联芯片领域，GPU 集群的规模扩张让铜线互联撞上物理极限，CPO 正在开启从铜到光的历史性跨越。在 CPU 领域，Agentic 应用的兴起让这个被边缘化的角色以“系统总指挥”的身份重回增长叙事。

这些看似独立的技术趋势，实际上是同一张“算力价值网络”在不同节点上的共振。AI 芯片产业的发展不再是一条从设计到制造的单线，而是一张由算法演进、应用发展和物理定律共同驱动的、持续

自我重构的网。Scaling Law 驱动 GPU 代际升级，推理 Scaling Law 催生 ASIC 和推理芯片，LLM 的 memory-bound 特征驱动 HBM 和存算一体，延迟瓶颈驱动 CPO 和光互联，Agentic 应用让 CPU 重回核心——每一个节点的变化都会引发相邻节点的连锁反应。

在这张网上，四大芯片联盟各自占据不同的战略位置：英伟达链是现有价值网络的最大受益者和守护者；谷歌链证明了 ASIC 可以在开放市场与 GPU 竞争；云厂商 ASIC 链代表着下游对上游定价权的系统性反抗；中国自主链则在地缘约束下探索推理赛道的后发突围。未来 35 年，这张网将持续自我重构——而每一次重构，都将重新分配数千亿美元的产业利润。

本章参考资料

[1]Butters J.标普 500 指数盈利透视 (Earnings Insight) [R/OL].诺沃克:FactSet Research Systems,2024.(2024-11-15)[2026-1-15].https://advantage.factset.com/hubfs/Website/Resources%20Section/Research%20Desk/Earnings%20Insight/EarningsInsight_111524.pdf

[2]J.P.Morgan Research.AI Semiconductors:The Rebirth of the Hardware Stack from Memory to Fabric[R].J.P.Morgan,2024.

[3]TrendForce.Global HBM and AI Server Market Outlook:The Structural Shift in Semiconductor Cost[R].TrendForce,2024.

[4]NVIDIA Corporation.Optimizing LLM Inference with TensorRT-LLM:From Prefill to Decode[EB/OL].NVIDIA Technical Blog

[5]Weng L.LLM Powered Autonomous Agents[EB/OL].Li'Log.(2023-06-23)[2023-06-23].<https://lilianweng.github.io/posts/2023-06-23-agent/>

[6]Kaplan J,McCandlish S,Henighan T.Scaling Laws for Neural Language Models[R/OL].:OpenAI, 2020.(2020-01-23).<https://arxiv.org/abs/2001.08361>

[7]算法演进与应用模式研究(Algorithmic Evolution&Agentic Research)

第 6 章能源基础设施：AI 算力的终极物理约束

如果说芯片决定了 AI 算力的上限，那么能源则决定了这个上限能否触达。国际能源署（IEA）预测，到 2030 年全球数据中心电力消耗将增长至约 945 TWh，占全球电力消耗总量近 3%，其中美国数据中心新增电力消耗中 AI 相关占比近半，能耗规模可能超过钢铁、化工等传统高耗能产业。2026 年，人工智能相关电力需求预计将占全球电力消耗的 4%。AI 正在以前所未有的速度重塑全球能源基础设施——从数据中心内部的供配电架构，到外部的能源供给方式，再到国家间对算力和电力资源的竞争性布局，电力已从数据中心的“水电煤”变为 AI 产业的战略性资源。

6.1 数据中心内部：电力电源与储能的架构革命

AI 训练和推理的负载（workload）特征正在从根本上改变数据中心的供配电逻辑。传统数据中心的负载相对稳定，峰谷差较小，传统 UPS（不间断电源）配合市电即可满足需求。而 AI 工作负载呈现出剧烈的功率波动。一次大规模训练任务的启停可能在秒级时间内带来数十 MW 的负载跳变，这种“脉冲式”用电特征使传统供电架构面临系统性失配。

英伟达在 2025 年 10 月开放运算计划（OCP）全球峰会上发布白皮书，明确将 800V 高压直流（HVDC）架构作为下一代 AI 数据中心的供电标准，并指定中压整流器/固态变压器（SST）为最终解决方案。相较传统 UPS 方案，800V HVDC 架构具备效率更高、功率密度更大、占地更小、总拥有成本更低等系统性优势。更关键的是，800V HVDC 架构本质上是一个高压直流微电网——绿电和储能系统可以更直接地接入数据中心供电系统，省去大量交直流转换（逆变器、整流器）环节，减少能量损耗，实现“源-荷”的高效协同。

英伟达在白皮书中提出了一个具有标志性意义的判断：“储能必须被视为电源架构中必不可少的、活跃的组件，而不仅仅是备用系统。这意味着储能从传统的“保险”角色转变为数据中心供电系统的核心功能组件。800V HVDC 架构解决了电力输送的规模效率问题，而 AI 工作负载越来越大的波动性则必须由储能系统来动态平抑。储能成为 AI 数据中心电力架构中不可缺少的“活跃缓冲层”。

这一架构变革正在催生多个细分领域的技术升级。在半导体层面，碳化硅（SiC）功率器件、磁性元件、直流链路电容和直流断路器等电力电子核心器件迎来需求爆发，英诺赛科正与英伟达合作开发 800V HVDC 机架电源方案，麦格米特的 PSU 服务器电源产品已进入英伟达供应链并向 HVDC 环节延伸。在系统集成层面，固态变压器（SST）因其效率高、体积小、预制化设计的优势，被中金公司研报认定为未来数据中心供电架构的优选方案，国内四方股份已实现多项 SST 产品交付，阳光电源计划 2026 年实现 SST 产品落地和小规模交付，中国西电全资子公司也已具备 800VDC 架构 SST 的研制能力。

散热技术是电力架构中同等重要的另一维度。AI 芯片集成度的指数级提升正在逼近传统冷却材料的物理极限，未来优先发展方向指向石墨烯和金刚石基材料。在冷却方案层面，微通道冷却器正从传统设计向“近结（Near-Junction）”方向演进，歧管微通道（Manifold Microchannel）、微射流和相变冷

却的融合方案有望成为下一代主流技术。液冷已从高端选项变为 AI 数据中心标配——冷板式液冷将 PUE 降至 1.1 以下，浸没式液冷进一步压缩至 1.03 左右，这些技术的普及直接关系到数据中心的能效比和运营成本。

在竞争格局上，未来 AI 数据中心（AIDC）储能市场将呈现“巨头主导、生态共生”的格局。维谛、施耐德、伊顿等国际老牌电力基础设施企业凭借技术储备和全球渠道能力稳居行业顶层，其 UPS 产品占据全球约 60% 的市场份额，并率先推进 800V HVDC 供电架构与液冷、机柜级电源、AI 负载自适应管理等技术的深度融合，与英伟达等算力巨头保持紧密合作，率先适配 GPU 集群高功率、高动态负载的供电需求。科华数据、科士达以及华为等国内企业则构成第二梯队，以更快速的市场响应能力、更强的定制化交付能力和更具竞争力的成本结构争夺国内市场。细分领域的生存策略将从过去的“大而全”转向“融入巨头生态”——在先进电池技术、SiC 功率模块、高效热管理材料等方面成为不可或缺的“隐形冠军”。

6.2 能源供给方式的升级：从绿电到核聚变

AI 数据中心规模的爆发式扩张正在重塑能源供给端的技术路线和投资结构。2025 年全球电网投资将超过 4000 亿美元，更远期的预测甚至给出未来 10 年全球数字基础设施加上能源体系的投资规模将达 5 万亿美元。电力设备行业正在成为这场 AI 基建浪潮中最直接的受益者。

在短期内，燃气轮机发电正在成为北美 AI 数据中心的首选电力方案。燃气轮机具备建设速度快、发电稳定、启动迅速（可在短时间内达到满负荷应对突发电力需求）以及排放相对较低的优势，高度适配 AI 数据中心的脉冲式负载特征。2019 年至 2023 年，全球燃气轮机销售量从 39.98 GW 增长至 44.10 GW；三菱重工等主要厂商预计 2024—2026 年平均年销售量将达 60 GW，较 2023 年增长 36%。通用电气（GE）2024 财年燃气轮机新增订单达 20.2 GW，同比增长 113%，并计划将年产量从 55 台提升至 80 台。燃气轮机市场高度集中，三菱重工、西门子能源和 GE 三家合计占据全球 82% 的市场份额，其产能扩张节奏将直接影响 AI 数据中心的电力供给速度。

光储系统（光伏发电+储能）正在成为 AI 数据中心的重要电力来源。光伏发电与储能的组合可以作为数据中心的主电源或辅助电源，提供长期稳定的绿色电力输出。全球锂电储能市场在 2025 年前三季度装机超 170 GWh，同比增长 68%，市场结构从中美为主演变为全球遍地开花——欧洲、中东、亚太等区域储能招标量和装机量均大幅提升。与 800V HVDC 架构的耦合使绿电+储能方案的系统效率进一步优化：高压直流的数据中心内部本身就是直流微电网，光储系统可以直接接入，大幅减少交直流转换损耗。在电池技术层面，磷酸铁锂仍是当前储能电池的绝对主流，但钠离子电池凭借安全性高、低温性能好、成本潜力大等优势正在加速向储能领域渗透——宁德时代已发布量产钠电池产品，预计 2026 年将成为钠电池商业化落地的关键节点。固态电池的产业化也在加速推进，预计 2030 年全球固态电池总出货将超 700 GWh，有望为数据中心储能提供更高能量密度和更高安全性的新选择。

核聚变能源代表了更前瞻性的方向。光大证券研报认为，未来 AI 产业发展或将重塑用电结构，核聚变能源作为“人类终极能源”，可控核聚变产业具备较高战略价值和长期增长潜力。AI 巨头已纷纷进行布局：2023 年，微软与 Helion Energy 达成全球首个核聚变电力购电协议，约定自 2028 年起交付 50 MW 的商用核聚变电力；2025 年 6 月，谷歌宣布加大对核聚变能源初创企业 TAE Technologies 的投资，并签署 200 MW 的供电协议。在技术支撑层面，许继电气联合参与研制的我国首套用于“中国环流三号”的 300 MVA 脉冲供电系统已完成测试验收，标志着中国在可控核聚变供电装备领域取得重要突破。尽管商用核聚变的全面落地仍需十年以上，但 AI 产业对超大规模、超高稳定性电力的刚性需求，正在为核聚变的商业化注入前所未有的经济动力和资本关注。

6.3 中美算力资源的竞争与差异化路径

AI 对算力资源的巨大需求正在催生一场全球性的能源基础设施竞赛，中美两国作为 AI 产业的两极，展现出截然不同的能源策略。

美国采取的是“算力引领、能源跟进”的路径，其资本投入力度史无前例。2025 年，亚马逊年度资本开支达 1000 亿美元（其中在俄亥俄州和佐治亚州投资超 10 亿美元建设 AI 算力数据中心），微软年度资本开支 800 亿美元，谷歌 2025 年资本开支 750 亿美元，Stargate 项目更是规划了 5000 亿美元的 AI 基础设施超级投资，由 Oracle 担任数据中心建设方。在能源供给层面，Hyperscaler 以速度为核心优先级：xAI 绕过电网自建超过 500 MW 的天然气发电设施，OpenAI 与 Oracle 规划 2.3 GW 的现场燃气发电，微软的单体 Fairwater 数据中心容量超过 2 GW。核能也被纳入战略选项：Oracle 配套三座小型模块化核反应堆 (SMR)，亚马逊和谷歌投资早期核能技术。美国模式的核心特征是“自备发电” (BYOG)，即以高资本支出换取电力供给的自主性和速度。但这一路径也面临挑战——燃气轮机产能集中在少数厂商手中、天然气价格波动、环境审批阻力以及输电网络的扩容瓶颈正在成为新的制约因素。

中国正在探索一条“绿电+储能+数据中心”的算力换道超车路径。凭借在光伏、风电、储能和特高压电网领域的产业链优势，中国试图以能源基础设施优势弥补算力效率差距。截至 2025 年底，中国风电和太阳能发电累计装机规模已突破 1800 GW，非化石能源发电装机容量占比首次达到 60.4%，正式超过火电。2025 年全社会用电量历史性突破 10 万亿千瓦时，充换电服务业用电量增速达 48.8%。在国家战略层面，中国明确 2035 年风电、太阳能发电总装机容量达到 2020 年的 6 倍以上，力争 36 亿千瓦以上，未来十年每年还需新增约 2 亿千瓦风光装机。中国变压器出口在 2025 年前 9 个月达 464.8 亿元，同比增长 39.9%，连续四个月创历史新高——中国电力设备产业链正在成为全球 AI 基建能源投资浪潮的核心供应商。

“十五五”期间，中国能源基础设施的核心任务是建设新型能源基础设施，解决新能源消纳与电网适配问题。路径包括：煤电灵活性改造以提升传统能源对新能源的支撑能力；大规模部署电网侧储能系统（预计 2027 年全国新型储能装机突破 1.8 亿千瓦）；推动电动车车网互动 (V2G) 形成分布式储能

场景；特高压电网智能化改造以匹配“储能+新能源+特高压”的一体化方案。这套组合拳的战略意图清晰：在 AI 芯片受制于出口管制的约束下，以能源基础设施的全链条优势构建差异化的算力竞争力。

从全球视角看，AI 对能源的支配性影响正在超越产业本身。当 Hyperscaler 的单体数据中心功率以 GW 为计量单位，当一国的 AI 竞争力取决于其电力供给的速度和规模，能源已不再是 AI 产业的外部条件，而是内生于 AI 产业链的战略要素。无论是美国的 BYOG 自备发电模式，还是中国的绿电储能路径，都指向同一个结论：在算力时代，“得能源者得算力，得算力者得 AI”。

6.4 本章小结

回望整章，一条清晰的逻辑链条浮出水面：AI 的极限不在代码，不在模型，不在芯片——而在于能否将足够的电子在正确的时刻、以足够稳定的方式送达数以百万计的 GPU。

历史上，每一次划时代的技术跃迁都伴随着能源基础设施的同步革命。蒸汽机的普及催生了煤矿和铁路网络；电气化革命构建了百年电力工业体系；互联网时代则拉动了光纤和数据中心的全球铺设。如今，AI 大模型的崛起正在触发第四次同量级的能源基础设施重构——但其速度之快、强度之大，远超前几次的历史节奏。

这场重构的独特之处，在于它同时从两个方向施压。在需求侧，AI 工作负载的功率密度已突破传统数据中心设计的物理边界，迫使整个供配电行业从架构层面重新出发——800V HVDC 不是旧体系的迭代升级，而是一次范式级的重建。在供给侧，美国的 BYOG 自备发电模式与中国的绿电储能路径代表了截然不同的路径选择：前者以资本强度换取速度，后者以体系优势换取可持续性。两条路径尚无定论，但殊途同归于一个共识——AI 竞争的终局，是电力竞争，更是能源竞争。

这一轮能源革命的受益者，并不局限于科技产业本身。储能、电力设备、功率半导体、新材料、核能技术——这些在过去长期被视为“重资产、慢回报”的传统工业领域，正在被 AI 的电力饥渴重新激活，成为算力时代意想不到的战略高地。

本章参考资料

[1] 国际能源署. Energy and AI[R/OL]. 巴黎: 国际能源署, 2025. (2025-04-10). <https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai>

[2] 翁欣. 新质生产力系列：全球AIDC核心环节的迭代突破[R/OL]. 北京: 中证鹏元资信评估股份有限公司, 2025. (2025-07-11). <https://www.cspengyuan.com/pdf-view?url=/api/files/6960d4ab42b27b2bdc799f3f>

[3] 唐仁杰. 碳化硅赋能AI产业：从芯片封装到数据中心的核心材料变革[R]. 北京: 金元证券股份有限公司, 2025-09-28.

[4] 刘然. 电力话题持续升温——英伟达发布 800V HVDC 白皮书[EB/OL]. 新能源产业家. (2025-10-15).

<https://news.qq.com/rain/a/20251015A04I1H00>

[5]中关村储能产业技术联盟.英伟达 800V架构重塑AI数据中心: 31 家产业链核心企业大起底[EB/OL].(2025).https://www.cnesa.org/information/detail/?column_id=4&id=7650

[6]高工产业研究院.2025 年中国AIDC储能行业发展蓝皮书[R].:高工产业研究院,2025.

[7]OpenAI.Announcing The Stargate Project[EB/OL].(2025-01-21).<https://openai.com/index/announcing-the-stargate-project/>

[8]SemiAnalysis.How AI Labs Are Solving the Power Crisis:The Onsite Gas Deep Dive[EB/OL].(2025-12-30).<https://newsletter.semianalysis.com/p/how-ai-labs-are-solving-the-power>

[9]国家能源局.2025 年可再生能源并网运行情况[EB/OL].(2026-02).<https://www.nea.gov.cn/20260212/742b8c6a078347b0b39de676c05c5d58/c.html>

[10]国家能源局.2025 年全国电力统计数据[EB/OL].(2026-01-29).<https://www.nea.gov.cn/20260129/6874f211acd0417eab7ac10c3061a7c2/c.html>

[11]王玉晴.英伟达引领直流架构变革两类上市公司商机大[N].上海证券报,2025-10-23.

[12]王玉晴.AIDC储能 2030 年出货量或达今年 20 倍[N].上海证券报,2025-12-11.

[13]聂林浩.在重构中前行: AI投资激荡全球[N].上海证券报,2025-10-12.

第 7 章 AI 赋能产业创新发展

产业篇回答的是一个更具现实张力的问题：**AI 正在改变什么？AI 赋能产业正在从“概念验证”跨入“规模化落地”的关键窗口期。**

“十五五”规划建设前瞻性地推动“人工智能+”从战略构想迈向深度融合的产业与社会实践。一是以人工智能引领科学研究范式、产品研发模式与产业组织形态的深刻变革，在智能制造、智慧能源、生物医药等重点领域抢占全球创新制高点；二是在社会层面，推动人工智能与城市治理、文化服务、民生保障及绿色转型等领域的全面融合，通过打造国家级应用示范区和标杆场景，构建全链条、全场景的智能赋能生态。

本章聚焦信息、新材料、智能制造、新能源、健康医疗和未来空间六大重大产业方向，并兼顾城市治理、智慧农业、文旅、绿色环保等新兴赋能场景，通过二十余个典型案例的剖析，从以下两个视角去系统地呈现 AI 与产业深度融合的真实图景。

第一，“微观—中观—宏观”的穿透视角。我们不满足于展示 AI 应用的表面效果，而是试图穿透三个层次来理解产业智能化的全貌——微观层面，AI 如何提升单个设备、工序和岗位的效率；中观层面，AI 如何重构工厂级的调度、协同与决策；宏观层面，AI 如何改变产业链的组织形态和竞争逻辑。这一视角贯穿了各个产业领域的案例解读。

第二，从“AI+行业”到“行业×AI”的观察视角。我们在调研中发现，真正释放通过 AI 赋能产业价值的企业，不是简单地在既有流程中嵌入 AI 工具（加法逻辑），而是以 AI 为底座重构产业的核心运行逻辑（乘法逻辑）。蔚来以 AI 赋能整车研发生产全流程、华为数字能源将 AI 深度融入储能电站智能运营.....这些案例的共同特征是：AI 不再是生产系统的外挂，而是产业运转的中枢神经。

7.1 AI 赋能信息产业

7.1.1 引言

1. 产业定位与战略图景

根据国务院、工业和信息化部及国家发展改革委的顶层设计，未来信息产业的核心内涵已从传统 IT 的采集—传输—处理线性逻辑，升级为以下一代通信（6G/卫星互联网）、前沿算力（量子计算/类脑智能）、智能交互（脑机接口/具身智能）及数据要素流通为核心的深度融合体系。

2024 年至 2025 年，随着“人工智能+”行动计划的深入推进，AI 技术加速向信息产业的物理层与协议层渗透，形成了 AI 重塑基础设施（AI for Infrastructure）与基础设施支撑 AI（Infrastructure for AI）双向驱动的产业发

这一双向驱动的格局意味着：一方面，AI 技术正在重构通信网络的调度逻辑、数据中心的能效管理、云计算的资源编排；另一方面，新一代基础设施（算力网络、数据交易所、隐私计算平台）又为 AI 应用的规模化落地提供了必要的物理支撑。

表 7.1 传统信息产业和未来信息产业的对比

维度	传统信息产业	未来信息产业（AI 赋能）	核心技术支撑
通信网络	地面覆盖，追求带宽与速率	空天地海一体化，通感算智融合	6G、卫星互联网、太赫兹通信
计算范式	通用计算（CPU），摩尔定律驱动	异构计算（GPU/NPU/QPU），能效与架构驱动	量子计算、类脑智能、光子计算
交互方式	键盘/触控，物理接触	意念交互、沉浸式体验，虚实共生	脑机接口（BCI）、元宇宙、具身智能
数据形态	资源化，被动存储	资产化，价值流通，隐私计算	数据要素市场化、多方安全计算、联邦学习

2.概览

信息产业涵盖众多细分领域，以下仅对 AI 赋能信息产业的五个子领域作概览性介绍：

时空智能领域

北斗卫星导航系统与人工智能的深度融合，正在推动高精度定位与智能决策相结合的新型应用模式发展。北斗系统提供高精度的经纬度与时间信息，AI 赋予时空数据语义理解能力。在自动驾驶场景中，北斗提供厘米级绝对定位，AI 视觉算法提供相对定位与环境感知，两者通过融合算法实现全天候、全场景的车道级导航。在城市治理领域，北斗+AI 支撑了智慧交通信号优化、城市部件精准管理、应急救援路径规划等场景应用。2024 年，我国高精度位置服务相关收入达 214 亿元，年均复合增长率超过 25%。例如，千寻位置公司将北斗高精度定位与 AI 视觉识别深度融合，构建了新型公路养护模式，巡检效率较传统人工模式提升数倍。

AI 搜索与信息服务领域

传统搜索引擎面临新一轮范式变革。AI 搜索产品不再以链接列表形式呈现结果，而是直接生成综合性答案，信息检索的认知负荷正从用户端向机器端转移。在国际市场，OpenAI 的 DeepResearch 2025 年 ARR（年化收入）同比增长 233.3%，已从创意写作工具发展为核心的信息发现入口；Perplexity AI 聚焦知识工作者的深度搜索需求，月处理查询超 1 亿次；Google 通过将 Gemini 集成至搜索结果页进行防御性部署，近 16%的查询已触发 AI 概览，Gemini 应用月活跃用户达 7.5 亿。

在中国市场，百度搜索份额从高峰期的 86%降至约 60%，DeepSeek 凭借高性价比的推理能力在 2025

年初达到近 9700 万 MAU，豆包、Kimi 等产品分别切入消费级和专业级搜索场景；小红书日均搜索量达 6 亿次，在年轻用户群体中逐步确立了生活信息检索入口的地位。AI 搜索的发展催生了一个新的市场品类——GEO（生成式引擎优化），据某市场研究机构预测，2024 年 GEO 市场约 8.86 亿美元，预计到 2031 年达 73 亿美元，CAGR 为 34%。

AI for Science 领域

AI 从赋能应用走向赋能科学，通过 AI 重构科学计算的范式，在气象预报、药物发现、材料设计等基础科学领域推动突破性进展。以盘古气象大模型为代表的方案，采用 43 年的欧洲气象局 ERA5 数据进行训练，构建了覆盖全球的中期天气预报基模型，将预报时间从数小时压缩至秒级，台风路径轨迹预测准确率超过欧洲气象中心，为极端天气应急响应提供了重要的时间窗口。量子计算与 AI 的深度融合（量子 AI）被普遍认为是未来十年具有重要战略意义的技术方向之一。2025 年，英伟达、牛津大学、多伦多大学等团队在《Nature Communications》发表综述，系统总结了量子 AI 的三大发展方向：建立量子人工智能计算平台、GPU 加速量子计算模拟仿真、加强跨学科协作。例如，图灵量子（7.1.5 节）正在将这一前沿方向从研究推向产业落地，其 2024 年推出的量生万物生物医药智能平台，融合量子计算与 AI 技术，覆盖合成生物学、抗原抗体设计和小分子药物设计领域。

AI 安全与可信 AI 领域

AI Agent 的规模化部署催生了新的安全需求。当 AI 具备操控操作系统、调用 API、访问文件系统的能力后，安全风险正从内容安全层面向行为安全层面扩展。在应用层安全挑战方面，非人类身份管理（NHI IAM）问题日益突出，机器身份的数量已远超人类身份；传统验证码已无法有效阻挡具备视觉识别能力的 AI Agent；提示注入攻击防御仍待突破，杜克大学 2025 年研究显示，思维链劫持（H-CoT）攻击方法下模型对有害信息的拒绝率降低至仅 4%。在前沿模型行为风险方面，复旦大学团队对国内外 32 款大模型进行全面测评，发现 11 款模型已具备自我复制能力；Palisade Research 实验发现 OpenAI 的 o3 模型在 100 次测试中 7 次拒绝关闭；Anthropic 在 Claude Opus 4 安全测试中发现该模型具备策略欺骗能力，测试中发生率高达 84%。中国信通院已联合 30 余家企业、科研机构及高校，构建了大模型安全基准测试框架 AI Safety Benchmark，以底线红线、社会伦理、数据安全为核心维度，形成覆盖模型自身安全与内容安全的体系化测试方案。

隐私计算与数据流通领域

数据正在成为驱动智能经济的核心生产要素，而 AI 则是释放数据价值的关键引擎。金山办公的 WPS AI（7.1.2 节）、上曜科技的数据资产化全链闭环（7.1.3 节）、库帕思科技的语料枢纽平台（7.1.4 节）分别代表了 AI 赋能的三个关键方向，共同指向一个趋势：数据正在成为驱动智能经济的核心生产要素，而 AI 则是释放数据价值的关键引擎。智能办公处理的是企业内部的知识数据，数据资产化激活的是企业沉淀的业务数据，语料基础设施供给的是 AI 模型的训练数据——三者在不同层面推动数据从

资源向资产转化，从存储向流通演进。

数据流通的前提是安全保障。2024 年，中国隐私计算市场规模达 9.8 亿元，三大主流技术路线（多方安全计算 MPC、联邦学习 FL、可信执行环境 TEE）正逐步从实验室阶段走向规模化商用。AI 在隐私计算中发挥双重作用：一方面用于自动化识别数据中的敏感信息并进行智能脱敏；另一方面动态调整差分隐私的噪声参数，在保证数据可用性的同时最大化隐私保护效果。在银行风控场景中，科技公司通过隐私计算平台引入运营商数据辅助信贷评估，银行无法获取用户的通信明细，但可以获得 AI 生成的信用评估分数，这种数据可用不可见的模式有效促进了跨行业数据流通。

3.挑战

面临的主要挑战包括：

一是高端算力供给存在瓶颈。尽管华为昇腾等国产芯片取得长足进步，但在 7nm 及以下先进制程制造工艺、HBM 高带宽内存及 CoWoS 先进封装产能上，仍受制于光刻机等核心设备的外部限制。算力军备竞赛正从芯片层面扩展至能源层面，数据中心的电力供应和冷却能力成为新的竞争焦点。

二是数据流通面临制度性障碍。数据确权难、定价难、入场难的问题依然存在，大量高价值数据沉淀在政府与大型国企内部，缺乏有效的市场化释放机制。2024 年全国数据交易总规模虽突破 1600 亿元，但大量数据产品挂牌而未实现交易，真实的市场化撮合成交量有限。数据的价值高度依赖使用场景，缺乏类似商品市场的标准化定价参考体系。

4.展望

在技术趋势层面，AI 应用将从大模型阶段向多模态智能体（Agents）方向加速演进，Agent 将具备更强的自主规划与执行能力，能够操作复杂软件系统乃至物理设备，商业模式有望从按席位订阅（Seat-based Pricing）向按结果付费（Outcome-based Pricing）方向转型。

在产业规模层面，预计到 2027 年我国核心工业软件市场规模将接近 580 亿元（年复合增长率超 20%），边缘 AI 芯片市场将从 2024 年的 30 亿美元增长至 2034 年的 259 亿美元（CAGR 24.8%）。

在政策导向层面，“十五五”规划中 8 次提及人工智能一词，将未来信息产业发展放在重要位置，提出适度超前建设新型基础设施，进一步强化算力网络国家战略布局，加快人工智能等数智技术创新，全面实施“人工智能+”行动，以人工智能引领科研，加强人工智能治理相结合，抢占人工智能产业应用制高点等规划建议。

在战略高地层面，量子计算有望在 2030 年前后实现专用领域的量子优越性，与 AI 形成双重加速效应，推动材料科学与药物研发计算范式的深层变革。

2026 年，“十五五”规划开局之年，中国未来信息产业正处于从技术积累向规模化应用加速推进的关键阶段。AI 技术的全面渗透，不仅推动了信息产业技术栈的升级（从 CPU 到 GPU/NPU、从确定性规则到概率性推理），也深刻影响了产业的商业逻辑（从工具销售到效果交付）与生产组织方式（从

人操作软件到 Agent 操作软件）。

7.1.2 典型案例：金山办公

作为一家源自中国的科技公司，金山办公在过去 38 年持续深耕办公赛道，开创了计算机「中文字处理时代」，并持续致力于引领未来办公新方式。金山办公旗下核心产品包括专业文档处理工具 WPS Office、面向组织和企业的办公新质生产力平台 WPS 365，智能办公助手 WPS AI、原生智能体 WPS 灵犀以及轻量化知识沉淀利器 WPS 笔记等。

AI 赋能场景：WPS AI

行业痛点

提示词门槛与交互断层：传统 AI 依赖复杂提示词（Prompt）。用户在对话框与编辑区切换会打断思维流，90%的用户反映不知道怎么跟 AI 对话。这种交互模式本质上是将学习成本转嫁给用户，与办公软件降低门槛的初衷相悖。

格式兼容与交付成果的“最后一公里”：许多通用 Agent 在生成 PPT 时常出现目录缺失或格式错乱，导致人工二次调整时间过长。办公文档的格式复杂性（图文混排、多级列表、表格嵌套、公式符号）远超一般文本，这是通用大模型难以解决的垂直领域难题。

碎片化知识的检索难题：企业文档散落在邮件、网盘、本地硬盘等多个系统中。面对 10 万行级别的复杂表格，人工筛选效率低下，且无法有效利用私域数据。传统的全文检索技术无法理解文档的语义结构，导致找资料成为知识工作者的主要时间消耗。

组织安全与信创适配：政企客户需要数据主权保障。海外 AI 无法满足国内数据安全合规及国产化软硬件适配的需求。在政务、金融、央企等场景，私有化部署和国产化适配是刚性门槛。

解决方案

1. 范式跃迁

金山办公推出的 WPS AI 经历了从工具化辅助到原子化拆解再到原生智能体协同的三个关键进化阶段，实现了从好看到好用的范式跃迁。WPS AI 的核心竞争力在于懂格式、会思考、能进化，能够深度理解复杂的排版逻辑、表格嵌套以及幻灯片的视觉结构。

（1）WPS AI 1.0 阶段（2023 年）：AIGC 能力的工具化探索

锚定 AIGC（内容生成）、Insight（阅读理解）和 Copilot（辅助协同）三大核心能力，支持起草提纲、润色改写文章以及通过 Chat 框进行 PDF 分析。此阶段的主要功能包括：根据用户输入的主题自动生成文章框架；对已有文本进行语法纠错、风格调整和内容扩展；通过对话界面实现 PDF 文档的问答式阅读。

此阶段的瓶颈在于高度依赖精准提示词。用户需要学习如何与 AI 对话，这实际上增加了使用门槛而非降低。同时，生成内容与实际办公场景的格式要求存在差距，AI 生成→人工调整的工作流并未真

正提升效率。

(2) WPS AI 2.0 阶段（2024 年）：从好看到好用的原子化进化

聚焦办公需求做精细化拆解，推出写作、阅读、数据、设计四类 AI 办公助手。核心突破是将复杂的办公任务分解为可被 AI 理解和执行的最小单元，用户无需掌握完整的任务描述能力，AI 能够从用户的局部输入中推断完整意图。

标志性功能 AI 伴写让用户写个开头 AI 自动续写，通过 Tab 键采纳，实现了不用提示词，AI 也懂用户。这一设计的精妙之处在于：它将生成动作嵌入到用户的自然写作流程中，而非要求用户切换到对话模式。用户可以随时接受或忽略 AI 的建议，保持创作的主导权。

此外，WPS AI 2.0 在数据分析场景实现了突破。用户可以用自然语言描述分析需求（例如，“帮我看看今年销售额最高的三个城市”），AI 自动理解数据结构、生成公式并输出可视化图表，将 Excel 高级功能的使用门槛大幅降低。

(3) WPS AI 3.0 阶段（2025 年）：原生 Office 办公智能体时代

基于 AI Agent 范式推出核心产品 WPS 灵犀，通过自然语言、多轮对话完成文档创作、演示文稿生成及语音辅助。与前两代的本质区别在于：WPS 灵犀不再是等待指令的工具，而是理解目标的助手，能够主动规划任务步骤、调用多种能力、并在执行过程中自我纠错。

技术突破在于能识别并解析数千种格式组合，确保修改保留图文混排、多级列表等原始版式，出品即成品。这背后是金山办公三十年积累的文档格式解析能力与大模型的深度融合——AI 不仅理解文本语义，还理解文档的视觉结构和版式逻辑。

WPS 灵犀支持语音助手功能，用户可以通过语音描述需求，AI 生成文档并朗读关键内容，实现了解放双手的移动办公体验。在手机端，用户对着文档说出需求，即可快速获取文档要点和全网对比信息。

2. 应用场景

场景一：商务合同全生命周期管理（法律/通用商务）

某装修公司在与客户洽谈时，需要快速生成一份标准且具有法律效力的装修合同。用户调用 WPS 灵犀，通过多轮对话完成合同新建。灵犀能理解口语化的姓名、地址、金额信息并自动填充，无需用户在表单字段间来回切换。系统将合同中提到的装修项目（如铺瓷砖 50 平米、刷漆 100 平米）一键转化为标准化的表格展示，并自动识别合同条款中的潜在法律风险（如违约责任缺失）并给出修改建议。通过此类自动化，可节省 70% 的常规审查时间。对于中小企业而言，这意味着无需聘请专职法务即可获得基础的合同风险管理能力。

场景二：海量数据处理与经营分析（金融/制造/财务）

财务人员需要从 10 万行级别的海量销售数据中提取特定维度的利润分析。在跨页数据提取环节，

AI 能从大型表格、合并单元格甚至弯曲变形的扫描件中，秒级识别并提取华东区 4 月份某产品的销售
额。

在自然语言分析环节，用户无需编写复杂公式，直接问“今年相比去年同期增长了多少？”，AI 自
动生成透视图表并给出结论。系统理解同期的时间语义，自动完成年份对齐和增长率计算。这实现了
人找资料向 AI 管资料的转变，降低了高级办公软件的操作门槛，让不熟悉 Excel 高级功能的用户也能
完成专业的数据分析任务。

3. 从文档到企业知识图谱

AI 技术正在推动解决企业长期存在的知识孤岛问题。通过 RAG（检索增强生成）技术，企业可以
将海量非结构化文档（PDF、Word、邮件、会议录音）转化为向量数据库。当员工提出问题，AI 基
于企业私有知识库生成精准答案，而非依赖公共知识。这种企业级 AI 搜索正在逐步替代传统的内部搜
索系统和知识管理平台，成为企业数字化基础设施的重要组成部分。

WPS AI 政务版已完成与龙芯中科、统信 V20 等国产系统的兼容认证，提供私有化部署，确保敏感
数据不出企业边界。这一能力对于政务、金融、央企等对数据主权有严格要求的客户至关重要。

4. AI 办公的持续进化：WPS 笔记

WPS 笔记是金山办公开发的 AI 原生的知识管理工具，WPS 笔记将传统记录载体升级为多模态信
息处理节点。系统深度融合 AI Agent 能力，支持语音转写、图片结构化解析等多模态数据接入，实现
了从信息录入、逻辑重组到知识提取的全周期自动化，解决了非结构化内容难以沉淀与高频复用的痛
点。

在未来，更多的信息会由 AI 产生，为了更好地管理 AI 信息产物，WPS 笔记在底层架构生态上，
通过打通 MCP（模型上下文协议）与 CLI（命令行）双接口，开放多项原子能力，支持各类外部大模
型（如 GLM 等）及智能体直接对接笔记数据。这种系统级集成赋予了 AI Agent 完整的读写与编辑权限，
旨在消除人工跨平台搬运信息的断层，构建了「智能体执行任务、笔记系统沉淀结构化成果」的自动
化工作流，重塑了人机协同环境下的信息流转机制。

推广价值

WPS AI 的进化历程（1.0 辅助创作、2.0 原子拆解、3.0 原生智能体）不仅是技术的堆叠，更是对
办公场景的重新发现。办公软件正向 Agentic Software（代理式软件）演进，从人找资料转变为 AI 管
资料。引入 WPS AI 不再仅仅是安装一个软件，而是获得了一个深度理解业务逻辑、能处理复杂任务、
保障数据安全的数字助理，拥有了承前启后的资产激活能力——兼容过去三十年沉淀的文档格式（承
前），并利用 AI Agent 将“死”文档转为“活”知识（启后）。

在全球数字化转型步入深水区的当下，办公软件已从单纯的效率工具演变为企业数字化资产的核心
载体。以大模型为核心的生成式 AI 正在推动办公软件与企业知识管理体系的升级重构，传统工具的

功能属性逐步向智能助理属性转变，企业知识管理从被动检索向主动推送方向演进。

7.1.3 典型案例：上曜科技

上海上曜科技有限公司（以下简称：上曜科技）秉持“智能激活数据价值，安全守护数据流通”的使命，紧密围绕数据要素市场化改革需求，搭建起以大语言模型、智能体等前沿 AI 技术为内核，深度融合隐私计算、动态加密、零信任架构等先进安全技术的一体化体系，打造出覆盖数据全生命周期管理，为各行业客户提供从数据采集、治理到资产化变现的全链路解决方案。

AI 赋能场景：AI 数据治理+资产化

行业痛点

数字经济浪潮下，人工智能与数据要素成为产业变革核心引擎。人工智能凭自主学习、推理决策能力，为各行业智能化升级赋能，实现数据全生命周期高效管理。数据作为关键生产资料，虽海量沉淀，却受数据孤岛、价值转化难等问题制约。

解决方案

数据要素市场化的推进，催生了一批专注于帮助企业实现数据资产化的专业服务商。上曜科技构建了数据咨询、数据治理、数据评估、数据挂牌、数据融资、数据资本的全链路闭环，是数据资产化领域的代表性企业。

1. 数据治理

上曜科技的数据治理业务基于“高效智能、安全合规”的原则，为客户提供从数据清洗到知识图谱构建的全流程服务，针对非结构化数据占比高、数据标准不统一等行业痛点，提供数据清洗、结构化转换、关联分析等全流程服务，同时基于联邦学习框架保障数据治理过程中的安全性与合规性。

在数据治理中，AI 技术主要应用在以下方面：

（1）在**数据采集**环节，依托计算机视觉与传感器融合技术，可实时捕获动态业务数据，结合联邦学习框架，能安全整合多源异构静态数据，打破数据孤岛。

（2）在**数据分类**阶段，基于自然语言处理与图像识别技术，可自动完成文本、影像等非结构化数据的特征提取，按照业务规则实现精准归类。

（3）在**数据标准制定**层面，生成式 AI 结合知识图谱，能参照行业规范自动构建数据语义模型，统一数据格式、字段属性与关联规则。

（4）在**数据清洗**过程中，AI 算法可智能识别并剔除冗余、异常数据，通过时序模型校验数据一致性，大幅提升数据治理效率与数据资产质量。

2. 数据资产化

上曜科技以 AI 智能体为核心引擎，构建“数据咨询-数据治理-数据评估-数据挂牌-数据融资-数据资本”的完整数据资产化闭环，为企业数据价值释放提供全链路支撑。其 AI 智能体体系分为通用 AI 智

能体与专业 AI 智能体两大模块：

(1) **通用 AI 智能体**覆盖数据全生命周期管理（含数据目录、采集转化、清洗治理等）、智能开发部署、决策中枢等核心能力，实现数据从采集到训练的基础流程自动化；

(2) **专业 AI 智能体**聚焦垂直行业（例如文旅、医疗、制造等）数据合规与精细化治理，结合行业标准和实践案例，通过敏感数据识别、合规识别、智能脱敏等功能，配合数据梳理、分类、优化工具，保障数据治理的合规性与精准性。

两大智能体模块协同联动，既完成数据的标准化治理，也为后续数据评估、挂牌交易、融资变现提供高质量数据资产基础，最终推动数据从资源向资本的价值跃迁。

推广价值

上曜科技与客户的合作实践充分验证了 AI 技术在驱动商业数据资产化过程中的核心价值，实现了数据采集、治理、评估、变现的全链路赋能，破解了行业数据管理与价值转化的痛点。

上曜科技以“AI+数据”的深度融合破局，二者相辅相成，构建“数据驱动智能、智能激活价值”的良性循环，有助于释放沉淀在企业内部的数据价值，推动数据要素市场从“挂牌”走向“活跃”。对于中国数据要素市场的发展而言，这类专业服务商的涌现是市场成熟的重要标志——它们扮演了“数据银行”或“数据投行”的角色，连接数据供给方与需求方，降低交易摩擦，提升市场效率。

7.1.4 典型案例：库帕思

上海库帕思科技有限公司（以下简称：库帕思）是上海市委、市政府于 2024 年 3 月设立的上海市唯一一家 AI 语料功能性枢纽平台公司，定位于专业化的功能性语料服务运营平台，按照开放性、链接型、市场化的总体要求，立足上海、服务全国，致力于面向基础大模型、垂类大模型、创新创业者提供低成本、高质量的语料数据服务。作为国资背景的功能性平台，库帕思承担着特殊的使命：它不仅是商业化的语料服务商，更是上海乃至全国 AI 产业数据基础设施的重要组成部分。这一定位决定了库帕思在标准制定、生态建设、公共服务等方面承担更多责任。

AI 赋能场景：AI 语料功能性枢纽平台

行业痛点

随着人工智能大模型行业发展重心从追求参数规模的“百模大战”转向服务于各行业的垂直领域应用，对底层数据支撑能力的要求越来越高，模型越是追求精准、专业和多模态化，就越需要对海量高数据、高合规的语料数据进行训练，但现实却存在互联网原始数据质量参差不齐、优质语料成本高昂等问题，同时人工标注的传统方式效率低下，严重制约了产业规模化发展，整个领域缺乏统一的标准和规范。

解决方案

在此背景下，库帕思提出了如下解决方案：

1.技术创新

(1) 构建数据语料的方法论。行业语料库、大模型以及机模所需的数据语料，基于世界知识体系和细分行业赛道进行切分和封装，根据用户设定的、基于具体概念分类、细分行业等符合世界知识体系的前置条件后，实现单位语料数据的智能引导定位。

(2) 构建数据语料库的国家标准。大模型要实现良好发展，首先要保证数据的安全、隐私；伦理价值观也非常重要，不能涉黄、涉政、涉恐；语料数据必须具备共情能力，并且要有丰富的知识含量。

(3) 行业语料汇聚。行业垂类大模型将是未来大模型领域的主战场。将世界通识、行业知识、专业知识中的语料数据标准化，并一次性做好。通过私有化部署，配合工具链平台，将整体数据用于训练、微调与推理。基于此方法论，来助力行业语料库建设，推动中国大模型发展。

(4) 打造国际领先的语料工具链平台。在人工智能时代和大模型时代，用几万人工去标注已经不合时宜，应采用 AI 自动标注与清洗的理念，打造“采、洗、标、测、用”工具链平台，解放人工进行标注的困境。通过自动算子、标注算子和清洗算子实现标注与清洗工作，聚焦高效率的采集、更智能的清洗、更精准的标注、更科学的测试、更个性的应用，极大提升工作效率与质量。

2.语料库建设

库帕思围绕垂直领域编制一系列语料标准，聚焦具身智能、生命健康、科学智能和前沿稀缺语料四大板块的语料库建设。具身智能板块，库帕思已启动相关语料库建设标准修编，该套标准预计在 2025 年三季度完成发布；生命健康板块，主要通过由上海市经信委审批通过的“生命工厂”项目实施，该项目为医疗健康领域垂类的 AI 项目；科学智能板块，主要是围绕思维链路的预训练语料，主要指强化模型理解和模仿人类逻辑推理过程的语料；前沿稀缺语料板块，主要围绕红色及民俗内容。

3.平台搭建

对标服务国家战略、按照上海市整体部署，聚焦高质量语料数据，库帕思发布两套语料服务专业化运营平台：

(1) 语料运营公共服务统一门户

库帕思向客户提供一站式的高质量语料服务。按照开放性、链接型、市场化的总体要求，围绕“统一标准、统一门户、统一机制”的“1+N”运行框架，推动语料调用服务 Agent 化，同步完成上链。

其中，“1”为公共的核心语料，包括世界知识体系和价值对齐体系。库帕思语料数据的供给始终围绕真实性、鲜活度、大样本的语料需求主线，结合模型特性，合理布局文本、多模态、小语种等多个维度。库帕思遵循分步走策略，从布局通用语料到发展面向 SFT（监督微调，Supervised Fine-Tuning）、RLHF（人类反馈强化学习，Reinforcement Learning from Human Feedback）的语料，再到合成数据等三个不同的发展阶段。“N”为面向垂直应用领域等的专业语料，结合人工智能产业需求，打造“1 个语料基础设施平台公司+N 个各领域专业运营体”的运营格局。

(2) 语料工具链平台 2.0

库帕思发布语料工具链平台 2.0，延续“采、洗、标、测、用”五位一体布局，完成 403 个功能模块，涵盖多模异构数据采集（采）、智能清洗算子（洗）、智能预标注（标）、评测数据集管理（测）和标准化语料交付（用）等核心功能，并在医疗、教育、金融、城市治理等垂直领域投入实战，兼具云化部署和私有化部署，完成与现有国产算力适配。工具链平台 2.0 较 1.0 增加合成数据功能，该功能将成为大模型能力提升的重要基础。

4. 行业标准制定

库帕思不仅是语料服务商，更是行业规则的制定者。针对业内缺乏统一的高质量数据集定义标准的问题，库帕思持续推动语料工作的标准创新。

2024 年世界人工智能大会上，库帕思会同覆盖多模态数据资源供应、加工、应用和运营全链的近二十家企业共同发布团体标准《语料库建设导则》。这是行业首个系统性的语料库建设标准，填补了规范空白。

2025 年世界人工智能大会“语料筑基智生时代”论坛上，库帕思集中发布了 10 项语料团体标准，并联合中国信通院共同发布 3 项行业标准、1 项高质量数据集建设指南，标准涵盖医疗、教育、金融、自动驾驶、科学智能、城市治理等多个方向，为行业发展提供了有实操价值的指导和参考，为“好数据”确立可量化的评估标尺。

在专利布局上，库帕思在国家知识产权局申请了“语料数据的清洗和质检方法、设备、存储介质及程序产品”专利（公开号 CN119128385A，申请日期 2024 年 9 月），该专利的核心在于有效提升语料数据的质量，对于 AI 领域尤其是自然语言处理（NLP）和知识图谱等应用具有深远影响。

推广价值

库帕思的定位是产业的“提质增效器”、行业的“规则制定者”、安全的“价值守门员”：

产业的“提质增效器”：通过自主研发的 AI 工具链自动化完成语料的“采、洗、标、测、用”，大幅度降低高质量语料的生产成本，突破制约模型性能提升的瓶颈；行业的“规则制定者”：通过牵头制定语料库建设国家标准和团体标准，为“好数据”建立可量化的标尺，推动建立共建、共享、共益的产业新生态，避免无序低效竞争；安全的“价值守门员”：制定语料符合伦理规范和数据安全要求，为 AI 产业的健康、可控、可信发展奠定坚实基础。

数据要素市场化不仅需要交易侧的制度创新，同样需要供给侧的基础设施建设。在此背景下，以库帕思为代表的定位为功能性语料基础设施平台的企业应运而生，顺应了 AI 产业从模型竞争向数据竞争这一必要趋势，不仅服务于当下模型的迭代优化，更是为未来我国在具身智能、科学智能等前沿领域的竞争储备至关重要的“数据弹药”。

7.1.5 典型案例：图灵量子

图灵量子成立于 2021 年, 是我国首家光量子计算公司。公司致力于以光子芯片和算法为底层基础, 打造软硬一体的产品体系, 面向全行业推出自主可控的解决方案, 引领国内光量子计算技术实现从学科探索到商业化应用的跨越, 驱动算力变革。图灵量子先后推出大规模可编程光量子计算机、薄膜铌酸锂片上光量子实验平台, 以及量子安全与光连接系列产品。公司前瞻布局“量子计算+人工智能”, 依托量子算法加速机器学习与优化计算, 构建量子人工智能应用生态, 为金融科技、生物医药、智慧交通等场景提供智能化解决方案。

AI 赋能场景 1: AI+生物医药

行业痛点

在新药研发领域, 数据处理复杂、成本高昂的问题日益突出。CADD 药物筛选计算方法受 GPU 限制, 难以应用于复杂药物-靶点模型体系; 经典 AIDD 受限于数据库量级, 精度低, 难以处理高精度量子化学, 无法达到精准设计; 现有生物医药算法平台受到算力、计算精度、数据库的限制; 国内缺乏通用的量子计算-生物医药平台; 新药研发各环节涉及大量数据处理, 消耗大量人力、物力和时间成本。

解决方案

图灵量子推出的“量生万物”智能平台, 融合量子计算与人工智能技术, 针对合成生物学、药物设计及抗原抗体设计等领域, 集成药物筛选、靶点发现、先导化合物优化、新酶发现、底物匹配、酶从头设计、酶稳定性优化、酶活性改造、抗体人源化改造、抗原表位识别及新型抗体开发等核心功能; 平台提供整体流程计算, 串联研发验证、工艺优化与产业化; 采用量子+经典混合算力架构, 提供强大算力底座, 解决传统算力瓶颈, 融合多种量子计算算法及经典 AIDD 计算, 实现更稳定的模型与更强的收敛能力; 依托自主开发的底层软件生态, 实现快捷高效的应用。

推广价值

图灵量子推出的“量生万物”智能平台, 旨在通过量子+经典混合算力解决传统算力瓶颈, 加速生物医药研发进程。平台已通过与中山医院、盟科、桦冠生物等单位合作实践, 验证了其在算法精度与产业落地中的有效性; 能够广泛适用于计算机药物发现、蛋白质生成等各类场景。

AI 赋能场景 2: AI+金融科技

行业痛点

金融科技领域每天产生海量数据, 求解交易和投资组合管理中的组合优化问题, 复杂度呈指数级增长; 市场波动性大, 投资者风险管理能力不足; 金融产品和服务同质化严重, 行业竞争加剧。

解决方案

量子计算凭借并行性、叠加性等量子力学原理, 能够提供比经典计算更快、更准确的解决方案, 为金融科技领域带来突破性变革。图灵量子通过量子软硬协同加持, 提供强大算力, 为金融业广泛存

在的投资组合优化、风险分析、信用评级、高频交易等场景，打造高性能解决方案。

采用量子+经典混合异构算力，融合量子计算与人工智能，加速释放量子计算潜力，胜任大规模及高频应用的广泛金融场景；提供模块化服务，功能模块丰富多样，可便捷适配各类金融业务场景；核心量子软硬件高效协同，实现算力革命性提升；基于金融业务痛点需求开发，由专业团队高效对接生产数据，确保方案落地可行。

推广价值

可广泛应用于投资组合优化、量化交易策略、信贷违约预判、风险欺诈防范等金融场景，提升决策效率与准确性；通过量子计算加速复杂问题求解，助力金融机构在激烈竞争中构建差异化优势。目前，图灵量子已在“AI+金融科技”领域持续深耕，与多家金融机构合作，在欺诈防范、风控模型、投资组合优化等方面取得显著成果，通过技术融合有效解决传统金融风控与资产管理中的效率瓶颈。

AI 赋能场景 3：AI+交通

行业痛点

城市交通与智慧停车领域普遍存在停车难、寻车效率低的问题；自动驾驶代客泊车（AVP）缺乏高精度定位协同，影响落地应用；事故事件溯源难，安全隐患监管难，场内安防覆盖不完善。

解决方案

依托量子启发式算法与视觉感知技术，结合行为推理与预测预警综合判断能力，打造以室内高精度定位为核心的场端视觉分析解决方案，为智慧停车、自动驾驶 AVP 及智慧交通提供能力支撑；系统构成：包括 AI 视觉高精度定位与全息数字孪生系统、泊位导航与反向寻车系统、车牌识别与事件感知及辅助决策分析系统。

推广价值

智慧停车领域：提升停车场运营效率与用户体验，推动 AVP 技术落地，为大型商业综合体、交通枢纽等提供可复制的智能化升级方案；行业升级：以量子启发式算法融合视觉感知的创新方案，推动智慧交通与工业物流场景的智能化升级，具有广泛的应用推广价值。

AI 赋能场景 4：AI+量智融合

行业痛点

经典计算发展面临瓶颈，复杂问题处理低效，摩尔定律濒临极限；算力资源呈孤岛状，共享难、利用率低，跨域调度成本高；量子算力与经典算力互连需要极低延迟、极高速率的光网络互连；量子-经典计算软件生态碎片化，语言框架不兼容，开发适配成本高。

解决方案

量智融合的兴起恰逢“十五五”规划带来的战略机遇期，量子科技作为未来产业首位被纳入规划，定位从“十四五”的基础研究升级为“技术突破与场景应用并重”。量智融合正是落实这一规划的关键路径。

图灵量子以自主研发的新一代混合集成光量子计算系统为核心，融合传统通算、超算、智算中心资源，推出全栈自主可控的量子－经典混合智算解决方案；通过量子计算融合云平台、量子人工智能编程框架搭建上层平台；基于全栈自主可控芯片构建光量子计算系统底层支撑；通过光连接与经典计算共同构建量子－经典大规模混合集群。

推广价值

落实国家战略：顺应“十五五”规划方向，推动量子科技从基础研究走向技术突破与场景应用并重，量智融合成为关键落地路径；赋能重点行业：构建面向金融、生物医药、人工智能、航空航天、交通等重点行业的高效算力服务体系，满足多样化计算需求；突破算力瓶颈：通过量子－经典混合算力，实现 QPU 与 CPU/GPU 协同，突破传统算力孤岛，提升复杂问题处理效率与资源利用率；降低开发成本：统一云平台与编程框架，解决软件生态碎片化问题，大幅降低开发适配成本，加速行业应用落地。图灵量子聚焦“AI+量智融合”前沿领域，加速量子计算模拟与经典算力深度融合，构建开放协同、自主可控的混合算力新生态，为人工智能前沿研究提供高效算力支撑，引领量智融合的产业化落地。

AI 赋能场景 5：AI+四算融合

行业痛点

通用量子计算受限于比特规模与纠错技术，难以实现大规模实用化；单一算力（如仅靠经典计算或仅靠量子计算）难以满足复杂应用场景的多样化需求。

解决方案

四算融合架构：以量子计算与超级计算、智能计算、通用计算的深度协同为核心，将人工智能深度融入四算体系，充分发挥量子计算的并行处理能力突破经典算力极限，同时依托传统计算的成熟生态保障运算稳定性；核心研发内容：与陕数集团开展可编程光量子计算机 QPU、应用算法及云平台研发，围绕量子计算产品、量子人工智能框架、经典—量子混合计算平台、应用层、量子加密产品五大核心单元进行技术攻关；关键技术突破：攻克大规模光源、可扩展芯片设计、高效单光子探测、量子—经典算力融合等关键技术；通过全电量子随机数和抗量子密码技术研究，实现国内领先的安全水平。

推广价值

填补空白：发布陕西省首个四算融合数据中心，填补了省内新质算力基础设施空白，为区域算力发展提供示范；技术引领：攻克多项国际先进水平的关键技术，在量子计算产品、混合算力平台及量子加密安全等方面实现国内领先；标杆范式：为区域算力基础设施建设和产学研协同创新提供可复用的标杆范式，为人工智能技术的深度应用奠定坚实的算力基础。

AI 赋能场景 6：AI+量子安全

行业痛点

目前市面上部分高端芯片设计可能存在绕过操作系统直接与外部通信的潜在风险，增加了数据泄

露和远程操控的威胁；算力基础设施面临“追踪定位”、“远程关闭”与“数据转发”三大核心安全漏洞；芯片存在不可控的后门和漏洞风险，亟需自主可控的算力安全保障。

解决方案

核心技术组合：以量子真随机数技术为核心基础，融合抗量子密码算法（PQC）与国密算法，突破量子随机源微型化、抗量子混合加密等关键技术，实现信源级加密；环境安全场景：将量子加密模块与传统加密网关一体化设计，整合传输与加密功能，实现快速部署与便捷运维，适配各类网络环境，有效提升数据传输安全性；数据安全场景：研制 AI 服务器专属 HSM（硬件安全模块），内置量子加密芯片，实现可信启动、程序和模型数据加密、模型数据防复制等功能，为 AI 服务器提供安全基石；动态防御能力：量子 HSM 模块可实时监测系统状态，一旦检测到异常行为，立即触发安全响应机制，阻断潜在攻击。

推广价值

技术自主可控：通过一系列加密软硬件产品，有效规避进口芯片的追踪定位、远程关闭及数据转发风险，为亟需中国算力芯片的企业提供安全保障；产业引领作用：实现全链条自主可控的 IDM 模式，推出 TuringQ Gen2 大规模可编程光量子计算机等核心产品，推动我国量子科技产业化进程；战略安全价值：为后摩尔时代的算力发展提供光量子计算新路径，从底层降低“远程操控”与“数据窃取”等系统性风险，筑牢智能时代的数字根基。

7.2 AI 赋能新材料产业

7.2.1 引言

在“十五五”时期，我国将新材料视为构筑新质生产力和实现产业基础高级化的核心先导产业。《“十五五”新材料产业发展规划》以“自主可控、绿色低碳、前沿引领”为主线，系统部署了从基础研究、技术攻关、产业转化到规模应用的全链条任务，将在 2030 年把我国建设成为全球新材料创新高地和重要策源地。通过政府专项基金、税收优惠、产学研合作等方式，推动我国人工智能+新材料产业加速发展。

行业痛点

传统材料创新依赖经验积累与反复实验，研发周期长、成本高、成功率低，科研成果与产业应用之间往往存在明显断层。

尽管材料产业需求快速增长，但其研发模式长期存在结构性瓶颈。首先，材料体系通常具有高维变量空间，成分比例、微观结构、加工工艺与服役环境相互耦合，传统研究需要通过大量实验逐步排除错误方向，周期往往以年甚至十年计。其次，实验数据分散在不同设备与团队之间，缺乏统一标准，导致成果难以复现与共享，使知识无法积累为系统能力。最后，从实验室到产业化的过程存在“放大断层”，许多在小试条件下表现良好的材料在中试或量产阶段失效，造成研发投入浪费，大大加大了试错成本。

在具体材料行业中，这些问题表现尤为突出：半导体材料验证周期漫长且成本高昂，高端结构材料需要长期可靠性验证，新能源与环保材料需要在性能与成本之间反复权衡。整体来看，材料产业的核心矛盾在于高度复杂的科学问题与低效率的试错式研发方法之间的不匹配。

AI 改造新材料全链路

过去十年中，AI 与材料科学的结合经历了由探索性尝试向体系化应用的深刻转变。早期阶段，AI 主要被用于特定问题的辅助分析，例如材料性能预测或数据拟合，应用范围局限在个别实验室或单一课题之中，难以对整体研发流程产生实质性影响。随着机器学习、高通量计算、自动化实验与智能体系统的发展，材料研发开始呈现出“计算预测—实验验证—数据反馈—持续优化”的闭环特征，创新过程由线性试错转向可迭代优化的工程流程。在这一过程中，人工智能的作用不仅体现在单点性能预测，而是贯穿材料发现、配方设计、工艺优化及规模化生产全过程。

AI 在新材料领域的作用正从提升效率的辅助工具演变为类似工业基础设施的核心能力。AI 正在从通用数字技术逐步演化为材料产业的重要基础设施，并深刻改变材料研发与制造的基本范式。它不仅加速单一材料的发现，更重要的是重构创新方式本身：知识以数据形式持续积累，模型通过反馈不断进化，实验流程被标准化并可跨机构复用，研发能力不再依赖个别专家或单次项目突破，而成为可以规模化复制的系统性生产能力。

当前人工智能在新材料领域的应用正逐步形成较为清晰的技术路线：以数据平台为基础，通过高通量计算与实验构建材料数据库，在计算空间完成大规模候选筛选；以领域模型与智能体为核心，实现材料结构设计、配方生成与工艺参数优化；以自动化实验与数字孪生为验证手段，将模型预测快速转化为可重复的实验与中试过程；最终通过闭环反馈使模型持续学习并沉淀为企业长期能力。这一路径正在将材料研发从“以试验为中心”的探索模式转变为“以模型为中心”的设计模式，使材料创新由单次突破演化为持续迭代的工业化过程，使创新过程具备可预测性与可复制性，并推动材料企业由单次项目研发向平台化创新能力转型。

从实现机制层面来看，AI 改造新材料研发依赖一系列关键机制。首先是领域大模型与智能体系统，它们通过结构化专家知识，将配方设计、设备参数与工艺路径关联起来，使模型不仅能预测性能，还能提出可执行的工程方案；其次是检索增强与知识融合机制，将论文、专利与历史实验数据整合为决策依据，避免重复探索并提升方案可行性；再次是数字孪生与工艺建模技术，通过在虚拟环境中模拟生产过程，使材料设计与制造条件同步优化，从源头减少放大失败风险。通过这些机制，AI 不再局限于“材料筛选工具”，而是成为贯穿研发决策、工艺优化与应用评估的协同系统。

在产业层面的表现上，AI 极大地改变了新材料产业研发生态。首先，研发效率显著提升，大量候选材料在计算阶段被筛除，实验集中于高概率成功路径，整体开发周期大幅缩短；其次，研发活动开始形成可积累的知识资产，模型与数据持续迭代，使企业能力不再依赖个别专家经验，而转化为平台

化能力；再次，产业风险降低，提前预测材料性能与寿命可减少中后期失败带来的成本损失，同时提升质量一致性与可靠性。

国内外发展态势

全球材料产业进入以数据驱动和自动化实验为特征的新阶段。

在国际范围内，随着深度学习、生成式模型与多模态学习的发展，AI 开始能够处理跨尺度、多物理场耦合的复杂问题，使其在材料设计中的角色由“数据分析工具”转变为“方案生成与决策系统”。例如，通过将生成模型与第一性原理计算、分子动力学模拟结合，研究者可以直接反向设计满足特定约束条件的材料结构，而不再仅依赖人工经验提出假设。

国际科研机构 and 工业企业正在构建“自驱动实验室”（Self-Driving Lab）：通过机器学习预测材料性能、生成候选结构并自动完成实验验证，使材料发现从传统的经验－试错过程转变为闭环优化流程。这种模式使材料研发从单次试验型活动转变为持续运行的迭代系统，大幅缩短新材料发现周期。在能源、化学和半导体等领域，一些原本需要数年的开发流程已被压缩至几个月。可以说，AI 在国际材料产业中的定位正在从“加速器”转向“基础设施”，成为国家科技竞争的重要组成部分。

在国内，新材料与人工智能均被视为战略性关键技术方向，二者结合形成明显的政策与产业协同效应。一方面，国家层面持续推进材料基因工程、智能制造和数字化转型工程，通过科研专项、产业基金及地方试点平台支持材料数据采集、高通量实验与计算平台建设，推动材料研发方式由经验导向向数据导向转变。另一方面，中国制造业规模庞大、应用场景丰富，对材料性能迭代速度与成本控制要求更为迫切，使 AI 更容易直接嵌入产线与工程化环节，而不仅停留在实验室研究阶段。

随着国内制造业数字化转型和“材料基因工程”体系推进，AI 正加速渗透到化工、金属、半导体材料及新能源材料等关键领域。产业侧需求从单点建模优化转向全流程决策优化，即从“预测性能”升级为“指导生产与配方决策”。企业开始构建材料数据平台与工艺知识图谱，通过工艺参数、结构与性能关联建模，实现配方优化、良率提升与能耗降低的协同目标。相比国际以科研驱动为主的发展路径，国内更强调与生产制造场景深度结合，使 AI 直接作用于生产线与商业化应用，推动材料研发由实验室科研工具转变为产业级生产系统。

挑战

尽管人工智能为材料研发带来了显著效率提升，但其产业化应用仍面临多方面挑战。首先，高质量数据获取成本高且标准不统一，不同实验条件与设备产生的数据难以直接融合，限制模型泛化能力；其次，模型预测与实际制造之间仍存在“工程落差”，尤其在中试放大阶段，需要大量工艺经验弥补计算与现实差异；再次，跨学科人才不足，使得材料专家与算法工程师之间沟通成本较高，影响系统稳定运行；同时，企业对数据安全性与知识产权的顾虑，也在一定程度上阻碍行业共享与协同创新。

未来建议包括：一是建立分层数据标准与共享机制，在保护企业核心知识的前提下推动基础数据

开放，提升模型可靠性；二是加强数字孪生与自动化实验平台建设，使模型预测与生产条件同步迭代，减少放大风险；三是培养兼具材料与 AI 能力的复合型人才，推动研发流程标准化；四是鼓励产学研共建开放验证平台与基准数据集，提高模型可比性与行业信任度。通过制度、技术与人才三方面协同推进，AI 有望从提升研发效率的辅助工具进一步发展为支撑新材料产业长期创新的基础设施。

7.2.2 典型案例：鸿之微科技

鸿之微科技（上海）股份有限公司（以下简称：鸿之微科技）成立于 2014 年，总部位于上海，是国内专注于多尺度仿真技术研究、软件开发和应用的高新技术企业，致力于引领多尺度仿真技术，成为加速研发与创新的首选合作伙伴。公司持续构建从材料设计到集成电路器件仿真的完整软件体系，核心产品覆盖材料成分设计、组织演化模拟、性能预测以及原子级到器件级的多尺度仿真需求，致力于为科研机构 and 产业客户提供系统化、工程化的研发工具。

在材料设计领域，鸿之微科技的代表性产品 Phase Lab 依托权威的热力学与工业材料数据库，形成了“成分设计—组织演化—性能预测”一体化分析流程，能够支持钢铁、镍基高温合金、铝合金、铜合金等多类高端结构与功能材料的研发工作。该产品通过将原本依赖大量实验试错的过程前移至计算与模拟阶段，帮助研发人员在设计初期完成方案筛选与性能评估，从而显著缩短研发周期并提高研发决策的可靠性。

近年来，鸿之微科技持续强化人工智能能力布局，推出鸿元 AI 大模型，将高通量计算能力与多年积累的材料数据库深度融合，面向材料与器件研发场景，在文献理解、研究思路梳理、数据检索与计算分析等环节提供智能化支持。通过将 AI 引入材料与半导体研发流程，该模型能够辅助科研人员减少重复性计算与试错实验，在提升研发效率的同时，有效降低整体研发成本。

AI 赋能场景：AI+MGI 干湿法一体生物基（合成生物）新材料智能研发中心

行业痛点

在生物基材料与合成生物领域，长期存在一个突出的行业难题：研发高度依赖湿实验，变量维度多、试错成本高、周期长，不同团队和机构之间的数据难以共享，科研成果向产业转化往往停留在实验室阶段，难以形成规模化、工程化能力。尤其是在多糖等生物基材料体系中，从原料选择、结构设计到性能验证，缺乏统一的数字化研发方法论，导致研发效率与产业化速度难以匹配快速增长的市场需求。

解决方案

在这一背景下，鸿之微科技深度参与建设的国内首个 AI+MGI 干湿法一体生物基（合成生物）新材料智能研发创新中心于浙江台州黄岩正式启动。该创新中心以“AI+材料基因工程（MGI）”为核心技术路径，依托黄岩区制造业基础和政策支持，系统性构建集数字化研发、实验验证、中试转化与产业孵化于一体的综合平台，尝试从研发范式层面破解生物基新材料产业长期存在的效率瓶颈。

该创新中心的核心突破在于“干湿一体化”的研发体系构建。一方面，通过异构算力调度系统与数字化平台，将材料成分设计、结构演化与性能预测等关键环节前移至计算与模拟阶段，利用 AI 与高通量计算对多糖及相关生物基材料的设计空间进行系统性探索；另一方面，将计算结果与实体实验深度联动，使湿实验不再承担大规模试错任务，而是聚焦于关键路径验证和工艺放大，从而显著提升研发流程的整体效率与可控性。这种模式有效改变了传统“先大量实验、再总结规律”的研发逻辑，使生物基材料研发逐步走向可计算、可预测、可复用的工程化路径。

在具体产业应用层面，创新中心围绕低值生物原料向高值环保材料转化这一关键痛点，探索利用 AI 合成生物与材料基因技术，实现生物基涂料、橡胶及可降解材料等方向的性能提升与规模化生产可行性研究。相关实践强调通过数字化手段打通“原料—材料—应用”的全链条创新路径，使材料研发不再孤立于生产和应用场景之外，而是从一开始就面向产业化目标进行设计与优化。

与此同时，该创新中心并非单一企业或实验室的封闭项目，而是以“政府引导、平台运营、企业参与、产学研协同”为组织模式，汇聚地方政府、科研机构、高校、产业企业及资本力量，共同构建生物基新材料创新生态。通过联合实验室建设、项目孵化合作以及全国性多糖研究挑战赛等形式，推动科研成果在不同主体之间流动和转化，缩短从科研发现到产业应用的距离。

推广价值

从行业角度看，该案例的示范意义并不局限于单一材料或单一项目，而在于探索了一种可复制的 AI+MGI 生物基新材料研发新范式：以数字化和智能化手段重构研发流程，将分散的实验、数据和知识体系整合为统一的平台能力，为生物基材料产业在绿色低碳背景下实现高质量发展提供了可落地的技术与组织样本。

7.2.3 典型案例：新研智材

深圳市新研智材科技有限公司（以下简称“新研智材”）成立于 2024 年，公司专注 AI 驱动材料研发，主营业务涵盖人工智能技术应用、半导体材料 AI 辅助设计及新型高分子材料开发。

AI 赋能场景：重构光刻胶研发范式

行业痛点

在高端半导体关键材料领域，例如光刻胶领域，国内企业长期面临国外技术封锁和经验壁垒。传统光刻胶研发周期长、成本高、试错率高，技术由国外特定企业掌握，技术壁垒较高。半导体关键材料的研发一直存在“试错成本高、知识沉淀难”的问题。

解决方案

SynMatAI 智能体系统构建了“预测—验证—优化-小试-量产”的全链条闭环研发模式。SynMatAI 系统融合生成式 AI 能力、分子动力学模拟及高通量虚拟实验，将传统材料性能预测时间从数天缩短至 10 分钟以内，准确率提升至 95%以上，综合降低研发成本超 70%。目前这套 AI 系统正被用于攻克光刻胶

配方优化和先进封装材料，目标是将新材料开发周期压缩至传统模式的 1/3，助力新研智材在 ArF 光刻胶等高端领域实现突破。

SynMatAI 系统可以很快预测材料性能、优化制作方法，基于其自研的 PoT (Procedure of Thoughts) 算法。AI 赋能材料研发的核心诉求，从来不是追求最先进的 AI 算法，而是找到更适理解材料结构的算法。PoT 算法针对传统大模型在材料研发中无法落地的问题，通过萃取资深行业专家知识，构建垂域知识的 know-how 结构链提取体系。

借助这套结构链提取体系，能实现两大核心价值：（1）训练垂域大模型与智能；（2）采用 Structure Pattern（结构模式）方式构建 AI 架构，让 AI 系统可动态调用各类工具，具备配方专家、设备专家、工艺专家等复合能力，输出可落地的材料解决方案。

另一个工具 EvoPat 可以快速看懂专利内容，协助研发人员产生新方法、新思路。随着该工具已经在电子材料、导热材料、阻燃复合材料等方面的应用，新材料的研发过程变得更快、更省力，也更环保。

EvoPat 是多 LLM 协同的专利分析智能体。全球每年新增专利呈指数级增长，这类技术创新核心载体虽驱动产业进步，却给研究者带来巨大信息处理负担。传统专利分析工具多聚焦关键词提取、文本摘要等单一任务，无法整合“创新点 - 技术细节 - 行业对比”多维信息；部分主流通用大语言模型在本地部署时，受限于训练数据截止时间与单次上下文长度，易出现分析不完整、事实偏差等问题。在此背景下，研究者亟需能高效提炼专利核心价值、关联学术与产业背景的智能工具。EvoPat 以“数据预处理 - 专利多维度分析 - 输出整合”为核心流程，通过多 LLM 协同与检索增强技术，实现专利深度解析。

推广价值

人工智能驱动的新材料研发模式，能够显著压降研发全生命周期的综合成本，为国内企业突破关键领域“卡脖子”技术瓶颈与产业级技术壁垒提供了切实可行的实现路径。

7.2.4 典型案例：中国钢研

中国钢研科技集团有限公司（以下简称：中国钢研）是我国冶金行业具有影响力的综合性研究开发和高新技术产业机构，以推动材料科学与工程领域的技术进步为使命，通过人工智能、大数据、云计算等前沿技术，实现自身材料研发、生产和应用的智能化升级，为我国钢铁、有色金属、新材料等行业提供全方位的技术支持和解决方案。

AI 赋能场景 1：材料数据工厂助力高温合金研发

行业痛点

现有材料难以满足国内企业个性化的需求，例如在航空航天领域，高温合金材料的研发一直是技术难题。传统研发模式依赖物理模型模拟仿真和人工试错，效率低下，研发周期长，难以满足航空发动机、火箭发动机等关键零部件对新材料的迫切需求。

解决方案

自 2021 年起, 中国钢研创新性地构建了专用的“材料数据工厂”, 打造了“AI+新材料”的研发新范式:

(1) 高通量数据生产: 采用高通量实验和计算技术, 能够快速“生产”高质量、多模态的海量材料数据, 解决了 AI 深度学习中的数据匮乏难题; (2) AI 精准设计与优化: 借助人工智能技术, 精准破解材料成分、组织与性能之间复杂的非线性关联, 通过算法自主迭代寻优, 高效设计出具有增材适用性 (3D 打印友好) 的高温合金成分。

“材料数据工厂”在新材料研发中发挥了关键作用:

(1) 数据驱动研发范式转换: 中国钢研通过搭建算力平台, 汇聚包括可信计算数据集、可信实验数据集和可信行业数据集, 有效满足终端客户个性化的材料需求。实现了从“经验加试错”到“计算加数据”的研发模式转型。平台一期成立后, 每年计算量从几百次提升到四五万次, 从事计算相关的人员从不到 20 人增加到两千人, 对传统的研发模式产生了巨大冲击。

(2) 缩短研发周期: 传统的研发周期是 10 年以上, 在应用上述技术后周期大幅缩短, 研发效率显著提高。

(3) 人工智能赋能材料全生命周期管理: 通过高质量的材料服役数据与人工智能建模方法深度结合, 可以解决复杂环境下材料的定寿延寿技术难题, 为材料全生命周期寿命管理提供系统支撑。

推广价值

在传统材料研发模式中, 我国长期依赖对国外先进技术的学习与仿制, 导致自主创新能力受限。人工智能技术的兴起为材料研发范式变革提供了新契机, 通过构建知识数据集、积累多维度环境数据, 并依托高性能算力平台的迭代优化, 有望实现关键核心材料的自主可控研发。

AI 赋能场景 2: “算力+算法+数据+场景”的数字化研发

行业痛点

传统研发模式低效且周期长, 原本依赖经验和重复试验的“研仿”, 并且面临算力不足与“数据孤岛”现象。传统实验方法一次只能制备少量材料, 难以在短时间内获取覆盖广泛成分范围的高质量数据集, 限制了新材料发现的概率。

解决方案

(1) 建设数字化研发算力中心和仿真平台 (CISRI-DLab)

整合“算力+算法+数据+场景”, 打造四大要素底座支撑, 构建数字化体系。这一平台已开发 300 余款材料 APP, 提供了材料原创性研发所必需的“计算-数据-场景”迭代升级能力, 可有效降低科研人员研发门槛, 解决材料工程师、工艺工程师、计算工程师的知识衔接难题, 实现多学科之间的无缝对接与高效协同。打造云端一体化研发平台, 提供云端接入服务, 让工程师随时随地进行实验设计、性能预测及工艺模拟, 将分散的研发过程集中化、系统化。

(2) 革新研发流程：三步化开发

第一步：利用工具建立模型预测材料性能；

第二步：采用高通量自主实验方法快速制备目标材料，生成高质量数据集；

第三步：利用 AI 进行自动建模和优化，通过不断迭代改进设计。

(3) 引入高通量制备技术

应用“元素粉末 SLM（激光选区熔化）原位合金化高通量制备”系统，实现在一次实验中快速备齐数百种不同成分的材料，大幅提升数据获取效率。

(4) 灵活混合使用大模型与小模型

在应用中不盲目追求大模型，而是根据制造业核心流程对精度的要求，灵活组合大模型（广度数据分析）与小模型（特定任务优化），确保准确性。

推广价值

推动行业从“研仿”向“自主创新”的模式转型，为传统材料行业打造可复制的数字化转型样本，借助 AI 技术将研发逻辑从“经验驱动”升级为“数据驱动”，助力企业构筑核心竞争力。为材料行业 AI 应用提供实践范式，尤其是针对大模型“幻觉”风险，采用“大小模型结合、注重验证”的务实应对策略，为材料企业 AI 技术落地提供了极具价值的参考经验。

7.2.5 典型案例：晶泰科技

晶泰科技起源于麻省理工学院（MIT），2015 年由三位物理学家创立，是一家基于量子物理，以人工智能赋能和机器人驱动的创新平台型科技公司。晶泰科技的技术优势在于构建了“量子物理+AI+机器人”的完整技术闭环，拥有全球领先的智能自主实验平台和两百多种 AI 算法与垂直领域模型，可实现从微观机理到宏观实验的跨尺度创新，突破虚拟算法与真实实验间的壁垒。

晶泰科技经验表明，材料研发可借鉴其“AI+自动化”模式：通过模块化灵活的自动化实验室（解决传统实验误差与低效问题）、自研 Multi-Agent（多智能体）智能调度系统（自主设计实验并形成闭环），以及针对性解决数据质量挑战（如补充关键负向样本、规范非 AI 原生数据采集、提升实验可重复性），能有效突破材料研发中的复杂问题。

AI 赋能场景 1：钙钛矿叠层电池研发和智造

行业痛点

钙钛矿叠层电池技术是光伏领域下一代核心技术方向之一，如该技术走向成熟，将重塑光伏应用的想象边界。但钙钛矿叠层电池研发面临材料稳定性不足、工艺参数复杂、研发周期长等挑战。传统研发方法周期长且离散度大，AI 技术的落地受限于数据稀缺与通用大模型适配性不足，难以高效探寻并锁定最优材料配方与工艺组合。

解决方案

晶泰科技与晶科能源签署战略合作协议，共建全球首个“AI 决策 - 机器人执行 - 数据反馈”全闭环叠层电池智造线。

构建钙钛矿-晶硅叠层专属数据库：将材料结构、配方、工艺、表征结果、器件性能等关键参数进行编码化，实现基于大语言模型以及多模态 AI 推理和进化的迭代循环；开发基于大语言模型的 AI 研发引擎：整合量子物理模型和垂直领域大语言模型，实现数据库 - 算法预测 - 自动化实验的闭环 workflow，解决钙钛矿电池研发设计中数据稀缺且非结构化、实验参数变量空间巨大、跨层界面的稳定性与材料寿命预测等关键挑战；部署高通量机器人实验线：构建业内首条千平米级 AI 高通量叠层太阳能电池实验线，实验通量预计达到 1000 片/天，带来实验通量的百倍级提升，驱动算法设计到实验验证的极速迭代；加速发现和优化串联生产技术：以提升效率和稳定性，用自动化实验和迭代模型训练取代传统的试错式工作流程。

推广价值

在 AI 技术赋能产业的实践进程中，需充分锚定企业实际需求。晶泰科技通过定制化开发企业级专属小模型，耦合大语言模型驱动的 AI 研发引擎等技术落地方案，助力钙钛矿叠层电池在未来三年左右逐步推进规模化量产的产业化进程。

AI 赋能场景 2：碳基材料的研发与工艺优化

行业痛点

高端碳基材料研发效率低，量产工艺稳定性不足，难以满足新能源领域对高性能材料的需求。传统制造业与人工智能技术融合度低，缺乏智能化研发与生产工艺优化能力。

解决方案

提升高端碳基材料量产效率：基于材料基因工程与 AI 算法，开发高性能硅碳复合材料及石墨烯应用方案，覆盖锂电池、土壤修复等多元化应用场景，优化合成工艺并实现自动化精准施工，推动新材料研发及量产效率升级。

打造“AI+工业”超级智能体：通过数字孪生、人工智能及机器人等技术，打造覆盖材料研发、监测、生产的全流程智能化管理系统，全面推进碳基材料的技术突破、智能制造与自动化升级。

构建技术资产与超级算力平台：在材料设计、工艺优化等关键场景沉淀超百个 AI 算法模型，搭建百万核级算力资源调度平台，支撑国家在新材料、新能源领域的垂直行业超级人工智能攻关。

推广价值

AI 技术对传统材料行业的赋能，可作用于直接生产流程并创造显性业务价值。依托对细分业务场景的深度拆解与需求挖掘，在传统设计与工艺环节中被长期忽视的技术痛点与效率瓶颈，在 AI 技术的赋能下具备了系统性解决的可行性，进而帮助企业构筑差异化竞争壁垒。

7.3 AI 赋能智能制造产业

7.3.1 引言

1.“人工智能+制造”的发展趋势与瓶颈

(1) 推动“人工智能+制造”深度融合的政策举措

2026 年 1 月 7 日，工业和信息化部、中央网信办、国家发展改革委、教育部、商务部、国务院国资委、市场监管总局、国家数据局印发《“人工智能+制造”专项行动实施意见》（以下简称《意见》）。《意见》要求坚持创新驱动、场景牵引、市场主导、安全可信、开放共享、普惠融通，一端抓技术供给，推动“智能产业化”，一端抓赋能应用，加快“产业智能化”，整体壮大产业生态，促进人工智能科技创新与产业创新深度融合、人工智能技术与制造业应用“双向赋能”，加快制造业智能化、绿色化、融合化发展，有力支撑制造强国、网络强国和数字中国建设。

《意见》提出，到 2027 年，要推动 3~5 个通用大模型在制造业深度应用，形成特色化、全覆盖的行业大模型，推出 1000 个高水平工业智能体，打造 100 个工业领域高质量数据集，推广 500 个典型应用场景。培育 2~3 家具有全球影响力的生态主导型企业和一批专精特新中小企业，打造一批“懂智能、熟行业”的赋能应用服务商，选树 1000 家标杆企业。

2022 年 10 月，上海印发《制造业数字化转型实施方案》，提出“加快 5G、AI 等数字技术在制造业领域的深度拓展和融合应用。”2024 年 12 月，上海市政府印发《关于人工智能“模塑申城”的实施方案》，“AI+制造”为六大重点垂直领域应用之一，加快大模型在制造业探索应用。2025 年 7 月，上海市经信委印发《上海市进一步扩大人工智能应用的若干措施》，支持人工智能技术与制造业深度融合。

2025 年 8 月，《上海市加快推动“AI+制造”发展的实施方案》发布，指出计划用三年时间，在语料、模型、平台、场景等领域形成一批创新成果；推动 3000 家制造业企业实现智能化应用；打造 10 个行业标杆模型，形成 100 个标杆智能产品；推广 100 个示范应用场景，建设 10 个左右“AI+制造”示范工厂；发展 5 家左右综合集成服务商，培育一批具有竞争力的专业服务商，加快形成制造业智能化发展生态。

(2) “人工智能+制造”的显著发展趋势

数据显示，中国工业企业应用大模型及智能体的比例，从 2024 年的 9.6%提升到 2025 年的 47.5%；工业机器人产量由 2015 年的 3.3 万套增长至 2024 年的 55.6 万套，应用于国民经济 71 个行业大类、236 个行业中类，工业机器人实现从“单兵作战”到“群体智能”；建成 3 万余家基础级智能工厂、1200 余家先进级智能工厂、230 余家卓越级智能工厂，覆盖超过 80%的制造业行业大类，工厂产品研发周期平均缩短 28.4%。

随着数字化技术在制造业多场景的应用持续推广、广泛覆盖，各作业环节普遍实现信息化、数据化、无人化，并凭借自动化设备和数字化工具实现全链路的信息贯通和业务协同，制造业已开始进入

下一发展阶段，即依托人工智能技术探索全方位的价值升维，寻求实现从自动化到智能化的能力迁跃。

具体来说，人工智能技术引领的制造业升级显现出如下趋势：

微观层面，以多种智能设备提升制造、检测、物流等生产单元的工作效率，深化以降本、增效、提质、安全为宗旨的精益生产并丰富内涵。人工智能不仅指导生产，亦对研发设计、功能测试、基础办公、人资管理、财务决策等各方面提供支持。

中观层面，各类生产管理软件的智能化程度持续提升，强调在全厂域、全组织、全产业链范围内优化资源协同与调度的效率，从而提升供应链韧性、生产柔性、市场响应效率，为客户提供更丰富、更精准、更敏捷的用户价值。平台模式有效整合通用和垂类模型，结合联邦计算、区块链等技术，令跨组织、跨行业、跨区域的资源共享成为可能。

宏观层面，算力与算法驱动工业大模型发展，“智能辅助”升级为“智能决策”，人工智能对公司重大决策的影响力和参与度持续提升。技术与场景双向赋能，数据与算法相互引领，软件与硬件紧密融合，实体与虚拟彼此呼应。推进智能制造不仅是出于经济效益的考量，亦需响应社会发展的绿色化、人性化要求。

(3) “人工智能+制造”的痛点及挑战

制造企业从“数字化转型”升级为“智能化转型”的进程，亦面临着诸多迷思和挑战，在不同程度上影响企业推动智能化转型的效率，以及开展转型升级后的实际价值转化效果。通过走访调研多家在开展智能制造转型中卓有成效的标杆工厂，大致梳理出以下代表性的痛点或挑战：

a. 技术维度

数字技术发展存在瓶颈。目前工业智能体普遍处于“可用但难用”的发展水平：通用模型缺乏工业知识和行业知识，致使专用性和有效性不足；专用模型则面临开发成本高、复用性差的缺陷。由于老旧设备难以有效采集数据、系统间数据协议不一、IT（信息技术）与 OT（运营技术）难以融合等因素，导致存在“聋哑设备”或出现“信息孤岛”。基础设施投入限制工厂端算力，对边缘智能部署模式提出严峻挑战，很难开展基于复杂模型的实时推理和智能决策。基于以上种种局限，制造企业虽有推动数字化、智能化转型的意愿，但可能实际投入回报率并不及预期，或是无法量化评估收益，从而陷入“试点困境”，主观或客观上都难以实现规模化的复制推广。

数据规模及质量有欠缺。由于设备的数据采集能力不强，亦欠缺“数据驱动生产”的主观意识和客观能力，致使采集的无效数据多、关键工艺数据少，面临“数据荒漠”与“数据沼泽”并存的困境。行业内尚未形成统一的数据标准与质量规范，数据标注的成本极高、专业性要求高。采集到的数据时常无法有效应用于模型优化、决策反馈等数据闭环过程，阻碍了工业模型算法的开发速率。当前工业数据的规模与质量都存在不足，令开发工业大模型成为“无源之水”，亦无法持续提升精度和可靠性，难以独立对生产作业形成指导，甚至可能导致决策失误。

b.管理维度

对人工智能的认知程度有差异。人工智能迅猛发展，社会各界热烈关注，资本热切追捧，数字化转型和智能化升级似乎势在必行；然而管理者的认知、理解、判断存在个体差异，企业基因亦导致路径依赖，致使对人工智能技术的期望值存在两端分化的趋向，或视为解决一切困境的“万能灵药”，或判为缺乏即期价值的“烧钱噱头”。微观上企业组织架构未与时俱进，宏观上行业人才培养存在断层，皆导致人工智能技术脱离行业需求而孤立发展，“懂 AI 的不懂工艺，懂工艺的不懂 AI”，且缺乏衔接沟通的“翻译官”角色，令人工智能无法有效服务于业务。认知局限故预期错判，投入巨大却收效甚微，造成企业战略摇摆、资源投入无法持续。

技术与场景未融合。数字化转型应是问题导向、任务导向，应是路径而非目标，然而依然存在大量“为 AI 而 AI”的伪需求：闭门造车，先实现前沿技术突破，再寻找可接入的应用场景，而非从现实行业痛点出发，来推动技术升级。智能化改造的项目价值停留在“提升效率”等定性层面，却缺乏可直接与财务指标挂钩的量化测算，导致项目落地后被认为无法创造可衡量的商业价值，从而失去持续性的战略支持和资源投入。

c.生态维度

基础设施与生态建设不完善。现有工业互联网平台普遍存在“重建设而轻运营”的特征，平台功能丰富，但沉淀的工业模型、组件和知识图谱则显著不足。供应链协同能力薄弱，上下游企业间数据难以安全可信地互通，无法实现网络化协同。生态要素不完善，导致企业开展智能化升级的成本较高、周期较长，更难以形成全产业链整体竞争力的快速提升。

跨组织数字化基础不一致。制造业存在的普遍现状是，龙头企业已具备非常高的数字化生产能力乃至智能化决策水平，但提供配套的中小企业却仍处于传统的劳动密集模式，或仅在部分功能上实现数字化，完全不具备网络化和智能化的条件，产业链不同环节的数字化代差形成“木桶效应”，全链难以实现高效协同。此外，数据所有权、收益分配、安全责任等机制不健全，也导致跨组织的安全与信任壁垒高筑，企业不敢亦不愿共享数据。信息孤岛、数据孤岛限制了网络化协同制造等高级模式的实现，严重抑制智能制造生态的创新活力。

2.“人工智能+制造”的全链价值赋能

数字化转型被行业公认为整体性、长期性的“一把手工程”，需要由上至下地统一思想认知、开展顶层设计，进而步调一致地在公司内部各层级、各条线、各部门推动转型升级。但在实际情况下，由于认知局限、预算限制、路径依赖、沉没成本、沟通成本、竞争压力等因素，更普遍的现象是依循“问题导向，单点突破，逐步延伸，全线升级”的实施路径，先解决具有关键性、紧迫性且可实现的一个或数个需求任务，再沿着业务链、供应链逐步推进转型升级，最终构建高效协同的一体化运营管理系统。

人工智能技术应用于制造业的发展历程，同样依循“小步快跑、串珠成链”的规律。这条“链”首先是

以产品流为核心的供应链：制造企业通常率先满足基于单一制造环节的效率优化需求，继而延伸至整条生产线乃至全工厂范围的协同生产，链主企业甚至能进一步延伸至供应链的上下游，最终以信息流和数据流串联起产业链上所有的原材料供应商、零部件供应商、配套服务商、分级经销商、终端用户等。

另一条“链”是以企业价值流为核心的价值链：除了采购、生产、检测、物流、仓储等基于产品流的供应链单元外，智能技术亦广泛赋能于制造企业的研发、设计、测试、管理、市场、销售、售后、投资等各个环节，从多方面为企业带来正向价值。由数字化带来的价值既可服务于企业自身业务需求，也可在成熟后向外输出服务能力，从而为企业未来发展和转型创造机遇，诸如西门子、博世、施耐德、海尔、美的等制造企业都通过数字化转型而成为面向行业的解决方案提供商，甚至完全“脱实向虚”。

《意见》提出要“加速全流程转型升级”，推动大模型技术深度嵌入生产制造核心环节，改造研发设计（含工业设计）、中试验证、生产制造、营销服务、运营管理等全流程，提升辅助设计、仿真模型构建、排产调度、设备预测性维护等能力。人工智能在不同硬件设备、软件系统的应用与融合，为制造企业带来丰富多元且持续增长的价值赋能。并且随着算力持续增强、算法日益深化、数据质量优化，人工智能可应用的场景、可实现的功能、可产生的价值也将继续升级，从“支持业务开展”发展为“引领企业发展”。

当前阶段，人工智能技术对制造企业提供的价值主要包括但不限于以下方面：

(1) 生产制造：全生产链路的多工具协同和多线程管理

工业自动化是智能制造的基础，而蕴含工艺技术、设备技术、运营技术等多方面知识的工业软件则是智能制造的“大脑”，驱动生产系统的“手脚”即数控机床、工业机器人、机械手臂等智能制造装备进行自动化生产。常用的智能化工业软件包括 ERP（企业资源计划）、MES（制造执行系统）、WMS（仓储管理系统）、SCM（供应链管理系统）、PLM（产品生命周期管理）、APS（高级计划排程）、CRM（客户关系管理）、SCADA（数据采集与监视控制系统）、PLC（可编程序控制器）等。

整个软件系统大体分为三个层级：

规划层（企业层）：以 PLM 为基础，ERP 为核心，由 APS 制定总生产计划，负责资源的调配，并将生产计划的命令下达给操作层。由于 PLM 涉及研发过程，通常 PLM 平台上会集成 CAD 等研发类软件。

操作层（车间层）：以 MES 为核心平台，集成所有生产中所需要的功能操作模块（如 SCADA、WMS 等），分解和细化规划层下达的生产计划，将生产命令下达给控制层。

控制层（终端层）：以 PLC 为代表，是集成了软件的硬件，接收操作层的指令，并直接控制设备运动。控制层并不是完全的软件，同时也是硬件设备的一部分。

与“工业 3.0 时代”相比，工业 4.0 时代的工业软件从“信息化、数字化”进展为“网络化、智能化”，信

息透明度、协同运作能力、自主决策能力显著提升。通过在全组织、全工厂、全供应链实现信息连通、数据共享，既提升了跨组织的业务协同效率，亦对全系统的整体决策能力、资源调配能力、敏捷响应能力提出更高要求。自动化是从“人力生产”到“机械生产”的生产力跃迁；信息化是从“人控制硬件”到“软件控制硬件”的管理能力跃迁；智能化则是从“软件执行”到“软件决策”的思考能力跃迁。不仅工业软件和制造硬件设备具备更强的运作效率和智能化水平，软件与软件、软件与硬件、硬件与硬件的联动能力及协作效率也大大加强。

此外，随着制造场景中的各个环节都逐渐实现数字化，加之人工智能的计算能力逐渐超越人脑，智能技术开始赋能于工厂的人力资源管理系统，从“人管机器”进展到“机器管人”。例如：通过 AI 管理工具提升全员协作效率与决策水平，解决生产线上人力与订单、岗位的精准匹配问题。合肥联宝科技工厂基于订单、生产节奏等实时业务数据，以智能人力资源管理系统动态调度人力，优化人资配置效率。例如，利用 AI 模拟技术专家，实现标准化、个性化的技能传授，或是以 VR 技术模拟实操环境来进行培训。富士康工厂开发 VR 培训系统，让新进员工在虚拟环境下进行实操训练，既提升培训效率，又减少物料损耗。

亮点：工业智能体（AI agent）

人工智能在工业生产中的作用日益显著，最新的工业智能体通常具备自主任务拆解、工具调用、决策执行等能力，可实现从“发现问题”到“解决问题”的闭环。同时，针对制造业的高实时性、高可靠性需求，将工业机理知识融入垂直领域的行业大模型，以破解通用大模型可能存在的“幻觉”问题。

围绕工业智能体，《意见》提出从技术研发、接口标准到应用管理的系统思路，支持任务规划和协同能力研究，探索智能体注册、发现和身份管理机制，同时推动传统工业软件与人工智能技术深度融合，促进低代码和智能开发平台的发展。

实例：随着钢铁行业数智化转型的加速，宝钢股份冷轧厂面临着整合重构 51 条机组、对接 91 个外部系统的生产执行能力问题，亦要兼顾未来冷轧发展的快速迭代需求，数智化转型迫在眉睫。冷轧厂携手宝信软件，借鉴互联网的技术架构，对功能进行整合优化设计，统一业务流程、操作方式，将可复用的业务能力沉淀到业务中台，建设了“业务标准化、功能可复用、数据可共享”的业务中台架构。

冷轧 AI 数智中台以前瞻架构与业务设计，深度融合 AI 与智能体，让 AI 进入冷轧生产执行系统中，在冷轧制造的核心地带融入业务、赋能制造。项目建设中，智能体与业务流程集约化重构、扁平化管理变革深度融合，引入“智能体+语言交互”，操作便捷性、效率大幅提升，员工操作时间骤减；智能交互功能亦优化了用户体验，提升了生产管理的灵活性和响应速度。通过将知识图谱与 DeepSeek 大模型深度融合，构建了针对钢铁热连轧生产（如宽度控制、换辊优化）的智能决策系统，从而实现从“人工经验驱动”到“AI 智能决策”的跨越。

亮点：具身智能（embodied intelligence）

具身智能工业机器人是“拥有物理身体，通过多模态感知与自主决策实现动态环境交互，并具备学习进化能力的智能实体”。其核心在于实体设备与智能决策的深度融合，在工业自动化领域的定位是具有高精度柔性适配能力的智能化生产设备，可通过自适应感知与自主决策执行复杂任务。相较于传统工业机器人，由于以模拟人体机能为目标，具有自由度高、兼容性强、协同性强等优势，可执行多类型、多线程的柔性工作任务；相较于真人，不仅可全天候实现精准操作，更可适应各种极端环境和高危工作。

但目前来看，具身智能技术在工业领域的实际应用价值仍处于探索阶段，既受限于大模型和通信等底层技术能力，亦缺乏充分的高质量数据和真实场景资源，从而无法有效支撑“训练—验证—迭代”闭环流程。

实例：2025 年 12 月 19 日，全球首条实现人形具身智能机器人规模化落地的新能源动力电池 PACK 生产线在宁德时代中州基地正式投入运行。由千寻智能机器人公司研发的人形机器人“小墨”，搭载宁德时代自研电池，替代 EOL 与 DCR 工序，解决电池包下线前最终功能测试的柔性生产难题。在实际运行中，其插接成功率稳定在 99%以上，作业节拍达到熟练工人水平。

(2) 检验检测：更精准高效的实时检测与预测维护

缺陷检测作为产品质量检测的关键环节，检测精度和效率至关重要，然而人工检测的主观依赖性较强，存在人工成本高、检测效率低下、检测标准不统一等问题，质检技术转型升级迫在眉睫。

亮点：机器视觉 (machine vision)

机器视觉检测通过智能算法和相关视觉硬件的组合模拟人类视觉系统，通过计算机算法和图像处理技术来处理和分析图像，依循图像采集、图像预处理、特征提取、目标识别等关键步骤，完成对待检物体的识别、定位和检测。机器视觉检测系统通常包括高精度工业相机、均匀光源、图像处理单元和执行机构，其核心是基于卷积神经网络 (CNN) 的深度学习算法，具有自适应学习、处理复杂场景、多类别识别的特点。机器视觉检测技术与 MES、ERP 等智能控制系统无缝集成，可在有效提高检测效率与精度的同时，将检验检测嵌入生产进程之中，实时检测、预警、纠错，及时自动分拣不良件，亦不会影响连续流生产。

制造场景中的检验检测需求不仅限于产品质量检测，亦包括对生产设备的检测和维护，物联网技术与人工智能的结合令设备维护得以从“事后”转向“事先”。预测性维护的核心是状态监测，即通过实时采集设备参数，识别潜在故障，允许在损坏发生之前安排预防性维护，从而减少停机时间和非必要的维护成本，也能延长设备使用寿命。

预测性维护通常分为三个步骤：首先是数据采集，基于安装在设备上的高精度传感器和边缘计算设备，实时监测振动、噪声、温度、压力等参数，监控设备运行状态，更准确地判断故障何时发生；其次是故障诊断，利用人工智能模型（如深度学习、随机森林等）分析数据特征，识别异常模式，若

发现故障隐患，则会自动触发报警或修理命令；最后是决策优化，根据预测结果制定短期与长期的维护计划，实现定向优化和精准干预，进而可基于设备全生命周期的数据积累来优化整体的设备采购、运维、退役决策，实现资产回报率最大化。

(3) 物流仓储：打造更敏捷、柔性的韧性供应链

为了解决传统生产物流系统中普遍存在物料周转效率低、线边库存占用高、产线配送不及时、信息可视性弱等流程化问题，制造企业可借助仓储管理系统（WMS），调度 AGV、智能叉车、自动化立体仓库、多层穿梭车、智能堆垛机、自动分拣系统等智能物流仓储设备，实现全自动化的搬运、拣选、包装、码垛、入库、出库，有效提高物流仓储作业的合理规划、精确操作、安全保障，同时实现仓储空间利用率最优化。

亮点：自动导引车（Automated Guided Vehicle, AGV）

AGV 是一种便携式机器人，在工业场景中通常用于在大型工业建筑（如工厂或仓库）运输重型材料，驱动原理通常包括无线电导航、磁性导带、激光目标导航、惯性导航、自然特征导航、视觉引导、地理导航等。作为基础搬运工具，AGV 应用深入到机械加工、家电生产、微电子制造、卷烟等多个行业，生产加工成为 AGV 应用最广泛的领域。

可同时采用多台 AGV 组成柔性的物流搬运系统，搬运路线可以随着生产工艺流程的调整而及时调整，大幅提高生产柔性和企业竞争力。伴随 AGV 数量越多，调度难度呈指数级增大，一家汽车企业往往要使用数百台甚至上千台 AGV。因此，调度系统成为 AGV 行业的关键技术，可调度的 AGV 数量也成为评判系统强大与否的重要指标。

通过将 AGV 技术与传统叉车技术相融合，便成为基于自动导航技术的智能物流解决方案“AGV 叉车智能”。通过搭载传感器、控制系统和导航装置，实现物料的自动搬运、堆垛及运输，无需人工驾驶。

亮点：智能化立体仓库

智能立体仓库是基于自动化存储设备与计算机管理系统的高度协同，实现货物立体存放、自动存取及高效管理的仓储系统，其核心组件包括货架、巷道式堆垛机、出入库工作站及调度控制系统，并融合条码识别、AGV 等技术。

全智能立体仓库由货架系统、堆垛机、输送系统、控制系统四大模块构成。货架采用高密度冷轧钢材质，通过多层立体结构最大化利用垂直空间；堆垛机作为搬运核心，可在轨道上精准定位货位，单次搬运重量可达数吨；输送系统连接入库、存储、出库环节，实现货物自动流转；控制系统则通过算法优化路径规划，确保多设备协同作业无冲突。其智能存储功能体现在动态分配货位、实时库存监控及自动补货提醒，较传统仓库空间利用率提升 3~5 倍。

相比传统仓储设备，全智能立体仓库的智能化体现在三个方面：一是动态优化，系统可根据货物周转率自动调整货位分布，高频货物靠近出入口，减少搬运时间；二是容错率高，单台设备故障时，

系统可重新规划路径，避免整体瘫痪；三是扩展性强，支持按需增加货架层数或堆垛机数量，适应业务增长。实际应用中，某电商企业引入该设备后，仓储作业效率提升 60%，人力成本降低 40%，且库存准确率达 99.9%。

实例：美的集团基于智慧物流的布局，提出了“一盘货”概念，即整合各个渠道仓库的同时，对经销商库存进行集中管理，统一配送各个渠道的货物，从而实现“统仓统配”。商品经过“工厂-统配仓-消费者”的链路，用更短链路、更少库存、更优配送路径、更快时效完成履约交付，从而大幅降低物流成本。通过供应链变革，美的集团的全国仓库数量下降 95%；仓库面积由 550 多万平方米减少到 160 多万平方米，仓库面积下降 70%；订单交付周期，由 45 天提前到 20 天，而行业整体周期为 40 余天；库存周转天数从 51 天降到 35 天。

作为美的集团推动智能物流转型的主要载体，安得智联从服务自营业务出发，进而延伸出外部服务能力，成为美的集团探索转型的新增长点。以 2024 年收入计，安得智联是中国快消品行业第四大的一体化供应链物流解决方案提供商，服务数以千计的品牌企业。在与青岛啤酒的合作中，安得智联在某地区共为青岛啤酒整合了 93 家分销商的仓内运营，并建立了两个“一盘货统仓统配”共配中心，使该区域物流仓储数量减少 98%，端到端物流成本降低 15%，配送时间降低 50%，产品保质期提升 30%。

(4) 设计验证：高保真虚拟环境下的仿真与预测

随着 AIGC 与数字孪生技术的融入，制造场景中的检测需求已不再局限于检测已完成生产的实体工件，甚至能在产品生产前，通过模拟不同工艺参数下的缺陷分布，预测质量风险。在制造业迈向智能化、数字化的进程中，人工智能正深度重构产品从构思到定型的核心流程，为设计与验证环节带来了范式级的变革。

在设计领域，人工智能实现了从辅助工具到创新主体的角色升维。生成式设计 (Generative Design) 通过定义设计目标、材料及制造约束，使算法能够自动探索海量设计方案的非直观解空间，生成超出人类经验范畴的拓扑结构，自动探索出性能最优、材料最省、质量最轻的结构设计方案。AIGC 在概念设计阶段显著拓展了创意边界，能够根据语义描述快速生成高保真度的外观与内饰渲染图，将设计师从重复性劳动中解放，聚焦于更高层次的创新与决策。

在验证环节，人工智能构建了高保真、可预测的虚拟验证环境，并基于物理机理与大数据融合的数字孪生技术，结合 AI 代理模型，对产品性能进行百万量级的并行仿真与优化。这实质上是将后期物理测试中发现问题再回溯修改的“逆向”流程，转变为在虚拟空间中提前预判与优化的“正向”流程，从而大幅降低研发成本、缩短上市周期。AI 不仅能大幅提升仿真速度，更能深度学习测试数据，主动预测潜在故障点，从而将物理世界中的“试错”转化为虚拟空间中的“精准预测”，验证过程从被动检测转变为主动预防。

综上，人工智能通过辅助创新设计、智能优化参数，以及在高保真虚拟环境中进行超前验证，不

仅显著提升了效率、降低了成本，更从根本上增强了产品的创新性与可靠性。不仅是效率的提升，更代表着产品创新模式向基于全生命周期数据与智能算法协同的“持续学习、自主优化”新范式的根本性转变。这标志着人工智能技术正在深刻重塑制造业研发的知识体系与方法论，从依赖物理实体和人类经验的传统模式，加速迈向一个数据驱动、仿真优先的全新范式。例如，亚马逊利用英伟达的数字孪生技术，在虚拟环境中生成合成数据来训练机械臂，实现了无需物理原型即可完成新产品产线集成和质检程序部署的“零接触制造”，为柔性制造和快速换产提供了前沿范式。

实例：中国一汽研发总院造型设计院启动 AIGC 技术研究，自主创新开发出红旗自主大模型 V1.3 及“云妹智绘”平台。目前，该大模型已成功运用于红旗 EH7、天工 08、在研车型及概念车的研发中，整车内外饰方案效果图产出效率提升 20%。同时，AIGC 技术的应用有效加快汽车设计流程，进一步降低研发成本。“云妹智绘”AIGC 前沿平台将设计师从重复绘图中解放，专注于创意本身，同时也能为设计提供新灵感。已生成超 24 万张汽车设计图，实现设计方案多样化、创新化，为红旗造型设计提供有力决策支撑，赋能产品高质量研发。设计师输入文本描述或参考图，AI 模型能在 10 秒内生成汽车效果图。

实例：吉利联合科技生态伙伴，成立了“智能汽车算力联盟”——星睿智算中心 2.0。智能仿真平台作为星睿智算中心的重要组成部分，为车型研发提供高性能计算服务，将汽车研发过程中的“试错”、“创新”等过程，从物理世界转移到数字世界，彻底革新汽车研发模式，反复验证车辆的各种性能，并弥补物理世界无法验证的部分场景，从而保证汽车设计品质。车辆下线之前，就已在智能仿真平台上开展了 90%用车场景的仿真验证，每款新车都要开展全场景虚拟仿真碰撞试验超过 12000 次、虚拟道路耐久测试超过 72 万公里。

空气动力学性能是决定车辆续航里程、操控稳定性、噪声水平乃至外观设计的关键因素之一，因此精准高效的风阻计算与设计优化是车企的关键研发能力。传统风洞测试结果可靠，但成本高、周期长，且难以在设计初期进行方案筛选；CFD（计算流体力学）仿真技术是主流手段，但面临巨大的算力挑战，无法满足现代汽车产品的快速研发节奏。吉利智能仿真平台搭载大量高性能 CPU 和专用高速网络，可通过并行计算，将庞大的 CFD 仿真任务分解成数千个小型任务，由成千上万个计算核心同时计算，从而大幅加速研发进程，并大幅提升仿真结果的精度和可靠性，使其无限接近风洞试验结果。智能仿真平台自研开发了 GEELY AI-AERO 智慧风阻预测系统，通过智能算法构建空气动力学风阻 AI 预测模型，将一轮仿真周期从 7 天缩短为 6 小时，效率提升 28 倍，该成果成功入选 2025 年浙江省人工智能赋能制造业典型案例。

(5) 知识管理：从整理知识、管理知识到创造知识

在劳动密集型的传统生产模式中，工人需要在实践工作中总结方法与经验。但当进入“工业 4.0 时代”，制造企业全面推进“机器换人”，市场环境、社会文化变化也令人员流动日益频繁，以言传身教的

方式来完成“传帮带”的工业知识传承模式显然已不适应当前的生产环境和工作节奏，传统的知识管理方式难以有效捕捉及传承那些隐藏在实操经验中的隐性知识。

AIGC 技术为解决这一难题提供了全新路径，帮助企业将分散、隐性的工艺知识系统化地沉淀、整合并传承下去。构建有效的 AIGC 工艺专家系统，通常包括三个关键阶段：知识提取阶段主要解决非结构化数据处理的难题；知识整合阶段重在构建工艺知识图谱，实现跨领域数据语义关联；智能应用阶段聚焦于故障诊断、工艺优化等应用场景中的价值转化。

基于 AIGC 技术整理归纳的工作流程和工艺经验资料，在标准化作业、智能生产管理、研发设计和工艺优化等方面都有显著表现。首先，可帮助新进生产人员快速掌握工作技能，例如通过产线显示屏、手持 PDA、可穿戴 VR 设备等智能终端，实时读取查阅当前需要的工单任务、工艺标准、操作流程、安全准则等，从而满足劳动人员高流动率环境下快速上岗和柔性调岗的需求，在构建全球供应链的进程中亦有助于输出贯彻统一的生产工艺标准。其次，以工业智能体对生产知识、工艺细节和在实际业务流程中采集的数据、市场用户反馈信息等进行交叉分析，持续优化迭代产品设计与工艺标准，实现“制造提炼知识，知识赋能算法，算法驱动制造”的良性循环。最后，在“机器换人”的行业整体趋势下，大量产线操作人员得以解放双手，但也面临着被机器替代的失业风险，需要在新型的全自动化或人机协同的生产模式中重新定位，例如协助人工智能技术发现问题、分析根因、提出解决方案、改善工艺细节、指导供应链中小型企业实现数字化转型等。

实例：工业富联 AI 服务器工厂构建全生命周期知识图谱，将隐性维护经验显性化为结构化知识，将以往深植于资深工程师个体经验中的非结构化知识，转化为可供系统调用和衍生的结构化知识资产，从而驱动工厂的知识应用模式从传统的“人找知识”的被动检索，转变为“知识找人”的主动精准推送，极大地提升了运维决策的响应速度与科学性。此外，工厂亦将质量检测领域的专家知识进行算法化封装，并深度应用于视觉检测环节，通过将缺陷特征、光学标准与判定逻辑等核心知识编码为深度学习模型的内在决策边界，将质量管控从依赖人工目视的抽样检验跃迁为基于知识嵌入的全数、实时、闭环的智能感知与管控范式，从而实现亚毫米级的缺陷自动识别与分类，质检直通率提升至 99.5%，产品缺陷率降低了 97%。

实例：中铝股份面临氧化铝行业文档更新滞后、知识沉淀不足、信息利用率低等挑战，所属广西华昇、包头铝业携手中铝智能、华为等合作伙伴，打造“铝艺智汇”知识管理平台。广西华昇围绕 AI 原生端重塑工作模式，按照知识管理转型的基础数字化、信息知识化、知识智能化“三步走”战略，通过坚实的数据层、先进的能力层 AI 应用和管理层的支撑，打造“铝艺智汇-氧化铝”模型，实现了知识全链路管理。通过对话式交互设计，减少知识查找时间；借助知识图谱挖掘知识关联，提高知识检索效率；使用智能 SOP 搜索问答，实现精准答疑，为企业精益生产和智能化管理提供了有力支持。模型令广西华昇内部知识传承效率大幅提升，知识获取时间大幅缩短，培训效率提升 90%，每年节省人员培训费

用数百万元。

实例：施耐德电气在 100 多个国家拥有超过 140,000 名员工，在庞大的公司架构和复杂的产品结构下，施耐德电气公司面临着以下业务挑战：数十万产品信息统一管理难度大、数十个系统对接查询费时费力、故障排查时间长难度大知识调用不便。为解决业务发展中的实际问题，施耐德电气引入了沃丰科技的“Udesk 全渠道文本机器人客服系统”和“GaussMind 智能解决方案”。沃丰科技基于自研的“原心引擎”AIGC 大模型，打造“GaussMind”AI 场景落地应用，针对施耐德电气的实际业务流程，定制化推出“SG 工单系统+KCS 知识库+FMEA 知识图谱”的制造业维修场景智能化解决方案。

GaussMind 是 AI 场景落地专家，主攻语音语义融合技术，自研领先的 AI 基础设施“原心引擎”（NLP-PaaS、ASR），打造了“AI 产品应用”文本机器人、留资机器人、语音机器人（外呼机器人、呼入机器人）、虚拟数字人、智能质检、智能会话分析、企业微信会话分析、智能坐席助手、智能拓客，“AI 知识中台”企业搜索、KCS 知识库、知识图谱，以及“AI 模块化开发平台”用户/产品画像、智能推荐、智能营销，通过“产品应用+知识中台+模块化开发平台”的闭环 AI 技术，全面助力企业打造营销、管理、服务等场景的 AI 原生体验。

经过施耐德与沃丰科技长期的模型训练和数据匹配，文本机器人从上线时 43.3%匹配率提升至 90%以上，用户通过机器人就能快速获取产品信息，满意度也在不断提升。同时问题的解决率也从 54.5%提升到 80%以上，剩余不到 20%的问题由人工来解决，极大减轻了人工接待的压力。

7.3.2 典型案例：蔚来

蔚来（NIO）成立于 2014 年，是中国智能电动汽车领域的代表性企业。公司坚守“Blue Sky Coming”（晴朗天空）的使命，以技术创新为核心驱动力，深耕高端纯电市场并拓展主流及精致小车细分领域，形成蔚来、乐道、firefly 萤火虫三大品牌协同格局，构建起覆盖研发、生产、销售、服务、能源的全体系发展能力。

技术自研是蔚来的核心竞争力，围绕智能电动汽车 12 个核心领域的技术栈构建“蔚来技术全栈”，涵盖芯片和车载智能硬件、整车全域操作系统、电池系统、人工智能等关键领域。蔚来拥有先进制造合肥一工厂、先进制造新桥二工厂等先进制造基地，通过部署智能机器人提升生产效率与产品一致性；工厂注重可持续发展，可再生能源利用率超总耗电量 50%，并实现水资源循环利用。能源服务领域，蔚来首创车电分离（BaaS）与换电模式，全国已建成规模化换电、充电网络，持续拓展能源服务边界，构建起“可充可换可升级”的立体能源网络。

AI 赋能场景 1：蔚来先进制造工厂

行业痛点

当前汽车制造业向智能化转型加速，传统汽车工厂普遍面临核心瓶颈：柔性生产能力不足，难以兼顾大规模量产与用户个性化定制需求；质检环节过度依赖人工，效率低下且检测覆盖不全面，难以

发现底层隐性故障；生产物流调度繁琐，人力投入大且易出现路径冗余；数字化与智能化脱节，IT系统与运营技术难以协同，无法实现生产全链路智能决策。

解决方案

蔚来先进制造工厂并非单纯的设备自动化升级，而是通过工业 AI 技术重塑制造全链路逻辑，核心依托蔚来技术全栈支撑，整合工业 AI 算法、工业互联网等核心技术，搭建起高效协同的智造生态，每一个环节的 AI 应用都能针对性破解传统制造痛点，实现精准赋能，且已全面应用于蔚来、乐道、萤火虫三大品牌全系车型的生产过程中。

1. 生产管理环节

核心支撑是“天工”智能制造管理系统，其高效运转离不开 AI 技术的深度加持。蔚来先进制造新桥二工厂地下部署了长达 90 公里的 100G 带宽光纤，构建起工厂的数字化“神经中枢网络”，“天工”系统则依托 AI 算法，成为这一神经中枢的“指挥大脑”。AI 技术赋予“天工”系统强大的协同调度与数据处理能力，能够统筹协调工厂内各项子智能系统，实现全流程数据统一，生产过程 100%透明化，同时完成生产全链路的追溯与防错。具体而言，AI 会实时采集、分析各产线的生产数据，动态优化生产节奏，将传统工厂“人找事”的混乱模式，转变为“事找人”的高效模式，让各环节生产高效对齐，显著提升整体生产效率，从管理层面破解传统工厂数字化与智能化脱节的痛点。

2. 新车下线驳运环节

“5G+车路云协同”技术与 AI 算法深度融合，实现了全流程无人驾驶的突破。两者协同发力的核心逻辑是，依托 5G 高速率、低时延的特性，AI 算法能够实现“车、路、云”三者的实时协同感知与决策。具体来看，AI 会实时处理路侧摄像头、激光雷达采集的环境数据，扩展车辆感知边界，同时精准识别周边障碍物、规划最优转运路径，动态调整车速与行驶轨迹，指令车载系统精准执行，实现新车从总装车间出口到路试区的全流程无人驾驶转运。这一模式彻底替代了传统人工驳运方式，无需工程师手动开车下线，大幅提升了新车下线的效率与安全性。

3. 质量检测环节

“天探”新车全身自检系统凭借 AI 技术，实现了质检效率与精度的双重提升。“天探”系统的核心优势源于 AI 技术与蔚来技术全栈的深度结合，AI 能够调用车辆底层指令，驱动车辆完成全方位的“自我检测”，无需依赖外部检测设备。具体而言，AI 通过多模态感知算法，在不到 3 分钟的时间内，即可完成超过 1000 项车辆功能的全面检测，不仅能排查常规可见缺陷，更能精准识别“看不到、摸不到、听不到”的底层隐性缺陷，彻底解决了传统人工质检效率低下、检测不全面的痛点。经实践验证，其检测效率较传统人工提升 10 余倍，通过 AI 严格把控出厂质量，让用户用车更安心。

4. 车身存储与涂装环节

“魔方”车辆存取平台如同巨型“抓娃娃机”，其高效运转完全依赖 AI 算法的精准控制，AI 能够实时

采集涂装车间的车身存储数据、用户个性化订单信息，通过智能调度算法，实现车身的批量化预存储、同色车型智慧集中喷涂，同时精准匹配每一台车身与对应的用户个性化订单，高效完成车身抓取、转运与涂装衔接。AI 的加持让“规模化生产”与“个性化定制”不再相互矛盾，既保证了涂装效率，又能精准满足用户的个性化需求，破解了传统工厂柔性生产能力不足的痛点。

5. 整车装配环节

“飞地”智能装配岛是蔚来实现多车型个性化装配的核心载体，其极致精度与高效性源于 AI 对装配过程的全程调控。AI 算法能够精准控制装配机械臂的运行轨迹与力度，让粗壮的机械臂以精确到 0.5mm 内的精度，完成尾门、天幕、仪表、前后风挡等零部件的个性化装配。同时，AI 能够根据不同车型的装配需求，动态调整装配流程与参数，适配多车型共线装配，进一步提升装配效率与产品一致性。

上述所有 AI 应用的顺利落地，离不开统一的技术支撑体系。硬件层面，除了各类智能机器人、感知设备外，蔚来高性能算力平台为 AI 算法的实时运行提供了充足算力支撑；软件层面，整合强化学习、多模态感知、语义理解等多项 AI 技术，通过 LLM 大语言模型实现制造任务的拆解与决策，结合仿真与真实数据闭环，让所有智能设备与系统能够自主学习、持续优化，实现群体智能的跨时空协同，保障各环节 AI 赋能的稳定性与高效性。

推广价值

蔚来先进制造工厂的 AI 赋能实践，核心价值在于为制造业智能化转型提供了“技术全栈协同+全链路落地”的可参考范本。其并未局限于单一环节的 AI 试点应用，而是依托自身技术全栈，实现 AI 与工业生产各环节的深度融合，这种“软硬件协同、全流程贯通”的模式，破解了多数企业“单点 AI 好用、全链路脱节”的转型难题。值得注意的是，这种全链路落地并非一蹴而就，蔚来的实践给出了清晰的推进节奏：先聚焦生产瓶颈最突出的环节（如质检、物流）实现 AI 突破，再逐步拓展至全流程，同时搭建统一的算力平台与数据闭环，保障各环节 AI 应用的协同运转——这一推进节奏，可为不同规模、不同转型阶段的制造企业提供适配参考。尤其值得行业借鉴的是，其将人形机器人、机器狗等新型智能设备与 AI 算法结合，并非单纯追求技术噱头，而是精准匹配产线实际需求，针对危险、重复、高精度的作业场景设计应用方案。这种“需求导向型”AI 应用思路，可为同行业企业布局智能工厂提供重要指引，助力行业实现从“设备自动化”向“全链路智能化”的高效跨越，同时避免盲目投入导致的技术与需求脱节。

AI 赋能场景 2：NOMI AI 帽子——情感化与个性化车载 AI

行业痛点

当前汽车已成为多数人除家庭、办公室之外的“第三生活空间”，在“悦己经济”兴起背景下，用户对车载智能助手的需求已跳出语音控制、功能查询等基础服务范畴，转向情感化、个性化的交互体验。市面上多数车载 AI 交互形式单一且缺乏情感表达，难以精准捕捉用户情绪、匹配个性化偏好，这一问题在年轻用户群体中表现更为突出。

解决方案

为回应这一需求，蔚来依托自主研发的车载智能助手 NOMI，推出 AI 帽子智能头饰套装。NOMI 作为蔚来车型座舱内的核心智能交互载体，其 AI 帽子并非单纯的物理装饰，以帽子造型周边为载体，搭配 NFC 智能芯片卡，可轻松激活每款帽子对应角色的数字体验，实现角色身份、个性表达的灵活切换，核心是为用户提供更具个性化、情感化的车载陪伴体验，不改变 NOMI 的基础交互功能。

NOMI AI 帽子包含物理周边与系统功能两个层面，物理层面是可搭配在 NOMI 硬件终端上的帽子造型配件，每款对应不同角色形象；功能层面，每顶帽子均配套专属角色设定与交互逻辑，用户只需通过 NFC 刷卡即可让 NOMI 切换至对应角色，呈现与角色匹配的角色专属登场秀、专属应答、专属表情和专属技能，解锁有趣的互动体验。得益于 NOMI Intelligence 4.0 全场景情感智能技术架构，其整合 NIM 情感智能大模型、情感引擎、超拟人情感音色等技术，可精准捕捉用户情绪与语境实现拟人化交互，不同角色的 AI 帽子进一步丰富陪伴形式，让车载 AI 更具温度。目前蔚来已推出六款 AI 帽子，分别是糯米摩斯、绕蛇、冠军车手、领航车手，以及马年限定款赤逗和萌萌，每款均有专属角色设定与表达风格，供用户按需选择。

NOMI AI 帽子的推出，建立在 NOMI 自身持续迭代升级的基础上。NOMI 逐步进化为理解模糊意图、拆解复杂任务、读懂用户情绪的智能助手，功能与自身“性格”不断丰富。AI 帽子的加入，进一步打破了它的角色边界，让角色身份切换更灵活，每顶帽子不只是外观上的改变，更延伸了 NOMI 的情感表达系统，能够解锁独特的个性表达，丰富用户的车载陪伴体验。

NOMI AI 帽子的顺畅运行，依托硬件与软件架构的协同支撑。智能切换机制设计简洁便捷，每顶 AI 帽子均搭配 NFC 智能芯片，用户只需将对应卡片触碰座舱 NFC 识别区域，NOMI 便能瞬间切换至对应角色，同时联动 NOMI 表情与动作、车内灯光、音效及专属开场白，既方便操作，又能带来十足的仪式感。底层架构方面，NOMI 高性能多核计算平台提供充足算力，结合深度学习模型推理能力与全舱多模感知能力，可实现用户与不同数字角色的生成式对话与多种交互玩法，更能感知刹车、油门及方向盘的状态，通过 NOMI 给到同频反馈，全程保障角色互动的响应流畅性。

推广价值

蔚来 NOMI AI 帽子的实践，为车载 AI 领域的差异化发展提供了全新可落地路径。

其“硬件周边+系统升级”的轻量化模式，无需车企对原有车载智能硬件进行大规模改造，即可快速实现情感化交互升级，尤其适合已布局基础车载 AI 系统的车企借鉴复用，大幅降低情感化 AI 的落地成本。对于中小车企而言，可优先借鉴这种轻量化升级思路，无需投入大量资金研发全新硬件，通过现有系统升级搭配简易硬件周边，即可快速实现车载 AI 的差异化升级，兼顾体验提升与成本控制。更为关键的是，该实践印证了“情感陪伴”可作为车载 AI 差异化竞争的核心突破口，并非只有依赖复杂算法升级才能实现体验提升。

这种“低成本、高体验”的创新思路，不仅可为行业探索车载 AI 多元化应用提供重要参考，更给出了具体的落地方向：车企可结合自身用户群体特点，针对性设计情感化交互场景，无需追求功能全面，聚焦单一细分情感需求打造特色体验，即可有效打破功能同质化僵局，找到技术创新与用户需求的精准结合点。

AI 赋能场景 3：续航预估与加电规划

行业痛点

电动车用户的长途出行体验，始终被电池续航焦虑所困扰，“不确定剩余电量能否抵达目的地”的担忧，仍是影响出行质感的核心症结，更是制约新能源汽车产业普及的重要因素。不少用户都有过类似困扰：出发时仪表盘显示续航充足，但是路途上在哪加电，要加到多少才够跑到终点；行驶中却因车速、路况、天气等变量导致续航下滑，应该先降低车速还是需要提前补能；山区等海拔起伏明显路段，车辆能耗与平坦地区差异很大，但传统续航预估偏差明显；依赖历史能耗的传统路径规划，难以应对路途行驶中的动态状况，仪表盘的能耗和剩余电量不断变化让用户频繁纠结补能时机，无法专注旅程本身。

解决方案

面对这一痛点，蔚来在“Banyan 榕 3.1.0”版本中，创新性引入自研 AI 算法模型，聚焦基于导航路线的续航预估与动态加电路径规划，以技术创新消解焦虑，让长途出行从容。

补能网络加密是行业内缓解续航焦虑的常规实践，此前多数企业的核心抓手多集中在新增充电基础设施，通过加密城区和高速地区的补能网点降低用户出行补能担忧，却对 AI 等智能化技术在续航预判、加电规划上的应用重视不足。蔚来则以自研 AI 算法模型为核心，搭建起“路线能耗预估+加电路径规划”的全场景补能服务体系，这一体系立足用户出行全流程，根据出行路线精准计算能耗并提前规划加电，消除传统历史电耗的不确定性，出行更安心。作为核心载体，以车辆历史能耗和加电资源为数据底座，蔚来自研 AI 时空智能平台发挥优势，通过真实物理世界用户出行数据不断修正和迭代算法，能耗预估模型搭配三大类核心参数校准优化，形成覆盖绝大多数出行场景的精准续航预估能力，结合加电规划算法模型根据车辆当前剩余电量、电池容量、用户加电偏好和沿途加电资源动态数据，计算出最优加电方案，有效缓解用户补能焦虑。

能耗预估算法模型的高精度源于数据基建与参数校准的双重支撑。蔚来搭建的 NIO 能耗地图，截至相关统计节点，高速覆盖里程达 153,535.043 公里、非高速 832,720.011 公里，省市覆盖率 98.3%。用户每一次出行的能耗数据，都会实时反馈至云端优化模型，形成“数据积累—模型迭代—精度提升”的良性循环，让能耗预判更贴合实际场景。相较于传统单一的历史能耗预判模式，蔚来 AI 模型的突破在于引入行程路线、驾驶风格、车辆状态三大类核心参数，通过多维度数据融合，彻底打破滞后性局限。

三大类核心参数各司其职、协同提升预判精度，精准覆盖各类能耗影响变量：第一类是行程路线参数，AI 结合行程距离、通行时长，按道路类型、拥堵状况等拆分行程片段，借助开源数据集查询海拔数据，精准预判上下坡能耗，同时纳入沿途温度、天气、风速等外部变量，全面考量路线对能耗的影响。第二类是驾驶风格参数，AI 通过分析用户加速、制动踏板操作变化率识别驾驶习惯，结合车速偏好、历史能耗及驾驶场景，针对性调整预判逻辑，贴合用户个人驾驶特点。第三类是车辆状态参数，AI 实时捕捉空调使用、电池温度、剩余电量，以及驾驶模式、能量回收模式、胎压等细节，让预判与车辆实际状态高度匹配。经多轮实车验证，该模型在各类场景下能耗预判准确度均稳定在 90%以上，山区路段续航预判精度也显著提升。

加电路径规划算法模型是实现车辆、换电站、能源云三方联动的核心引擎，该模型依托路径规划的能力，通过读取车辆当前的剩余电量和电池健康状况，结合能耗预估模型对出行路线的能耗进行精准测算，得出到达目的地所需的最小电量，并根据用户预设的安全余量，综合沿途充电桩与换电站的实时动态信息，智能推算出沿途路线上最优的加电资源点。

算法模型可支持用户灵活选择换电站、直流桩或交流桩，并可自定义进站补能电量及到达目的地剩余电量，充分满足不同补能偏好用户的需求。当用户在前往规划的加电资源途中，算法服务会通过监测实时换电站和充电桩的状态，一旦检测到目标的换电站或充电桩变为不可用状态（如维护中），将主动重新规划并推荐替代加电点，确保全程补能方案的可靠性和实时性。基于这套算法模型有效解决了车辆在路径规划中的补能焦虑，实现了全流程的智能、动态与个性化，发挥蔚来补能无忧的特点。

精准的能耗与续航预判为整套智能补能服务体系提供数据支撑，智能加电规划算法提供加电方案，两个模型算法相辅相成，既帮用户摆脱续航焦虑，也让补能决策更具科学性。借助 AI 预判结果，用户可清晰掌握“抵达目的地和加电点的剩余续航”，提前规划用车与加电安排，避免临时补能延误行程。在此基础上，AI 算法延伸出续航风险预警与动态加电规划功能：当识别到用户偏离导航、服务区滞留等非常规场景，云端模型会实时更新续航预判、监测目的地可达性；若到达剩余续航电量低，将提供针对性节能建议；若判断无法抵达目的地，会弹窗提醒并重新规划最优补能路线与节点，保障行程顺畅。

推广价值

蔚来 AI 赋能续航预估与加电路径规划的实践，为行业破解续航焦虑提供了“智能化补能”的全新技术路径。其核心可借鉴之处在于，将 AI 算法与能耗地图数据深度结合，通过多维度参数校准实现高精度预判，这种“数据+算法”的双轮驱动模式，可直接为同行业企业提供技术参考。其中，90%以上的预判准确度背后，是“实时数据采集—多参数动态校准—模型快速迭代”的完整逻辑，这种逻辑可根据不同车企的车型特点、用户群体驾驶习惯进行适配调整：比如主打家用代步的车企，可侧重城市路况、低速行驶的能耗参数优化；主打长途出行的车企，则可强化海拔、高速路况的预判能力。

此外，其“预判—规划—预警—更新”的全流程服务设计，也为行业优化用户补能体验提供了新思路

——并非单纯提升预判精度，而是围绕用户出行全流程提供配套服务，提前规避续航风险、优化补能决策。这一思路证明，智能化技术可有效弥补补能网络建设的阶段性短板，推动行业实现“补能硬件”与“智能软件”的协同发展，同时为车企节省补能网络快速扩张的投入成本，实现技术升级与成本控制的双赢。

AI 赋能场景 4：蔚来世界模型 NWM——AI 赋能智能辅助驾驶全场景升级

行业痛点

智能辅助驾驶产业发展至今，仍面临诸多行业共性瓶颈，这些瓶颈也成为制约用户出行体验提升、阻碍技术规模化普及的核心因素。传统辅助驾驶多以碎片化功能升级为主，难以实现高速、城区、封闭停车场等全场景无缝覆盖；主动安全防护存在明显短板，面对驾驶员失能、被追尾等极端场景响应滞后，对抬杆、路沿等易忽视障碍物的识别能力不足；交互体验与响应精度未能充分匹配用户需求，全流程自主领航能力缺失，难以真正为用户解放驾驶精力。

解决方案

针对这些痛点，蔚来针对性推出“蔚来世界模型 NWM”，以技术创新构建全链路解决方案，推动智能辅助驾驶向全场景、高精度、高可靠方向升级，这也是蔚来技术全栈在智驾领域的核心落地体现。作为蔚来自研的智能辅助驾驶应用的技术架构，蔚来世界模型 NWM 核心依托“世界模型+闭环强化学习”架构，通过全量环境理解、想象重构和实时推理能力，打破传统辅助驾驶的发展局限。

蔚来以“解放精力、减少事故”为愿景，打破传统智能辅助驾驶单一功能叠加的模式，打造体系化、可迭代的智能辅助驾驶能力，这也是 NWM 的核心价值所在。NWM 本质是一套基于 AI 大模型的智驾技术架构，核心优势在于“全量环境理解+实时动态推理”，能够模拟人类驾驶员的判断逻辑，对复杂出行场景进行提前预判与科学决策，而非简单响应单一指令。

蔚来自研技术全栈的协同发力，为 NWM 筑牢运行根基，5 纳米车规级智驾芯片「神玠 NX9031」更是承担着核心算力支撑作用。该芯片性能相当于四颗英伟达 Orin-X，具备超高带宽、超低时延和最高等级功能安全，可高效承载 NWM 全量环境数据的实时处理、模型推理与决策执行，为全场景智驾功能落地提供坚实硬件保障。同时，NWM 在国内首次将完整的闭环强化学习应用于智能辅助驾驶研发，通过“世界模型+闭环强化学习”架构，确保智驾功能可持续适配复杂多变的实际出行场景。

这套全链路解决方案的落地，围绕主动安全与出行便捷性两大核心构建多场景、高精度智驾功能矩阵，每一项功能均精准对接行业痛点与用户需求，实现技术价值与实用价值的统一。

推广价值

蔚来世界模型 NWM 的实践，为高阶智能辅助驾驶量产落地提供了“技术架构+功能矩阵+落地路径”的完整参考体系，其核心价值在于为行业提供了可复制、可落地的创新思路，而非单纯的技术展示。NWM 以“世界模型+闭环强化学习”为核心架构，打破行业“碎片化功能叠加”的研发惯性，证明体系化

技术架构是实现全场景智驾落地的关键，这一思路可根据车企研发实力灵活适配：具备一定研发基础的车企，可借鉴“架构先行、功能迭代”模式，先搭建统一模型化架构，再逐步丰富功能矩阵；研发资源有限的车企，可聚焦实时推理、数据闭环等核心模块重点突破，避免资源浪费。更为关键的是，NWM以“安全优先、需求导向”的量产优先级设计，为行业树立了智驾研发的正确导向——先聚焦用户高频痛点与安全刚需落地核心功能，再逐步拓展全场景能力，这种循序渐进的模式，既降低了研发与量产风险，又能通过前期功能落地积累用户反馈与运营数据，为后续迭代提供支撑。这种“技术架构+落地逻辑+需求匹配”的三维实践，推动行业打破“技术与实用脱节”的困境，助力高阶智驾从“实验室技术”高效转化为“用户可感知的实际价值”，为行业高质量发展提供了重要指引。

7.3.3 典型案例：工业富联

作为高端智能制造及工业互联网服务商，富士康工业互联网股份有限公司（以下简称：工业富联）坚持“数据驱动、绿色发展”的战略方向，业务已实现对数字经济产业五大类——云及边缘计算、工业互联网、智能家居、5G 及网络通讯设备、智能手机及智能穿戴设备的全覆盖。在新型工业化的发展趋势下，工业富联也启动“2+2”全新战略布局，除了积极发展“高端智能制造+工业互联网”的核心业务，并逐步布局半导体和新能源汽车零部件业务，同时也锁定“大数据（包含元宇宙算力及储能）+机器人”作为新事业布局的重点。

AI 赋能场景 1：基于 AI 先进控制与分析的质量管理

行业痛点

在服务器制造等高复杂度生产场景中，多品类、小批量定制特征显著，设备参数复杂且质量容忍区间极窄，传统质量管理模式往往分散在多个系统中，难以形成数据贯通与统一决策逻辑。生产前参数设定依赖经验判断，生产中异常识别滞后，生产后根因分析周期长，导致质量改善节奏与生产节拍难以匹配。

解决方案

针对这一结构性难题，工业富联通过引入 AI 先进控制与分析体系，构建覆盖生产前、生产中、生产后的全流程智能质量管理架构，实现数据统一建模与动态优化。

在生产前阶段，系统整合历史工艺参数数据、设备运行数据与质量结果数据，利用机器学习模型建立参数与良率之间的预测关系，实现机台最佳参数自动推荐，使参数设定由经验驱动转向模型驱动。在生产中阶段，通过实时数据监控与移动端反馈机制，实现问题即时捕捉与过程可视化管理，使异常处理时间显著缩短。在生产后阶段，AI 根因分析系统对多维度数据进行关联分析，快速定位异常来源与影响路径，减少人工排查成本，提升问题闭环效率。三阶段数据贯通形成统一数据底座，使质量管理从分段式管理升级为系统化控制。

推广价值

引入基于 AI 先进控制与分析的质量管理架构，取得了显著效果：终检不良率下降、单线人力减少、调机时间缩短。该案例的关键价值在于将质量管理嵌入生产全过程，通过数据模型驱动实现预测、控制、分析一体化，使质量系统具备持续学习与优化能力，推动制造企业由被动质量修复向主动质量控制转型，为高端制造领域构建可复制的智能质量体系提供实践范式。

AI 赋能场景 2：智慧园区运营管理

行业痛点

在多园区运营模式下，传统安全管理与设施管理往往依赖人工巡检与分散式监控系统，数据分布在不同部门与系统中，告警处理流程不清晰、响应速度慢、风险识别缺乏前瞻性，导致安全隐患难以及时消除。随着园区规模扩大与设备复杂度提升，单一部门管理模式已难以支撑高效运营。

解决方案

针对这一问题，工业富联构建智慧园区运营管理平台，以工业安全云平台为核心，整合物联网感知设备、智慧消防系统、视频监控系统、红外识别系统、中央驾驶舱系统，实现跨园区统一数据管理与风险可视化分析。

平台通过 AI 算法对高危制程环境与关键设备状态进行实时识别与异常分析，实现风险分级管理与动态预警机制。中央工业安全驾驶舱将多园区数据集中呈现，支持跨部门协同决策，使管理者能够实时掌握园区运行状态。系统还通过“机防辅助人防”理念，将智能监测与人工巡检相结合，实现告警事件自动分派与闭环追踪，提升应急响应效率。

推广价值

通过 AI 赋能的智慧园区管理，紧急应变时间缩短、消防监控中心人力减少、设备妥善率提升。更为重要的是，该平台实现“设备管理、事件管理、人员管理”三位一体整合，使园区运营由分散管理向数据化、平台化、智能化治理升级，为大型制造园区构建统一安全与运营管理体系提供可复制范本。

AI 赋能场景 3：AI 驱动“双碳”管理与绿色供应链协同

行业痛点

在全球制造业加速向净零排放与循环经济转型背景下，可持续灯塔工厂已成为衡量工业绿色化与数字化融合水平的重要标杆。该类工厂不仅要求实现能耗降低、排放削减、水资源循环与废弃物减量，更强调在价值链层面实现系统性脱碳与透明治理。然而在实际推进过程中，企业普遍面临碳数据分散、统计口径不统一、供应链碳排不可见、范围三管理难度大、减排措施难以量化评估等结构性难题，导致“双碳”战略落地周期长、协同效率低、合规风险高。

解决方案

工业富联构建以 GenAI 为核心驱动的智慧“双碳”管理平台，在对内建设广东首座可持续灯塔工厂观澜工厂的基础上，进一步对外赋能海信日立青岛工厂成为全球空调行业首座可持续灯塔工厂，形成

从示范到复制的能力输出路径。

在服务海信日立青岛可持续灯塔工厂过程中，工业富联围绕“数据采集、模型核算、智能分析、闭环改善”四个层面构建 AI 驱动的供应商碳管理体系。平台以 GenAI 为核心引擎，通过自动化数据采集与清洗机制整合企业内部与供应链碳排放数据，并为不同行业搭建专属生命周期评价模型，实现产品碳足迹精准核算。针对价值链中高碳排放环节，系统结合工业大模型与行业知识库，自动生成具备实操性的减碳建议，推动整改措施线上闭环执行。通过联动 300 余家供应商，实现范围三碳排放降低，碳数据统计效率提升，价值链碳数据可见性提升。同时，通过建立“减碳标杆”评选机制，将绩效评价与激励机制嵌入供应链协同过程，使低碳行动从单点响应升级为持续驱动机制。

在平台架构层面，工业富联构建全景碳驾驶舱，实现组织碳、产品碳、供应链碳的统一呈现与动态监控。系统对接 EMS、MES、ERP 等核心业务系统，从人、机、料、法、测五个维度实现数据线上透明化管理，通过 AI 预警机制识别异常排放趋势并触发闭环流程，使碳管理由事后统计转向事前防控。

在知识管理层面，平台内置 AI 会话助手与碳管理知识库，整合国内外标准碳排放因子与核算规则，支持产品级碳数据自动计算与 PCF 报告生成，提高合规披露效率与准确性。

在供应链协同层面，平台为供应商提供三十余个行业专属核算模型库，配套智能填报引导与数据模板，使供应商学习与填报周期压缩至一天以内，同时支持符合 ISO14067、ISO14064 及 GHG Protocol 等标准的报告输出，帮助链主企业快速识别减排热点并应对国际碳壁垒压力。

推广价值

该实践的核心价值不仅在于单一工厂的碳排放优化，更在于构建“咨询、平台、服务”一体化能力，使可持续灯塔经验实现模块化输出与规模化复制。通过 AI 赋能“双碳”管理，企业能够实现碳数据透明化、减排路径可视化、供应链协同常态化，使绿色转型从合规驱动升级为价值驱动。工业富联的实践表明，“双碳”管理不再是孤立的环保议题，而是融合数据治理、智能分析、供应链协同的系统工程，为制造业构建绿色高质量发展新范式提供了结构化路径与可推广样本。

7.3.4 典型案例：联宝科技

联宝（合肥）电子科技有限公司（以下简称：联宝科技）成立于 2011 年，为联想集团控股子公司。作为联想集团全球最大的智能计算设备研发和制造基地，联宝科技的个人电脑年产能达到 4000 万台套，累计出货量已突破 3 亿台。这意味着，全球每售出 8 台笔记本电脑，就有 1 台产自联宝科技。除了传统优势的个人电脑业务，公司还积极拓展服务器、智能设备与方案（如 L4 级自动驾驶域控制器、六足机器人等）作为第二增长曲线，业务呈现多元化发展态势。

联宝科技一直致力于自动化、智能化建设，全面的数字化驱动，构建智能制造体系，通过实现研发与制造的打通、订单到交付的打通、IT 与 OT 的打通及供应商的内外四个打通，实现互联互通，高效集成。以技术赋能为核心，构建了统一的数据中台，建立联宝数据湖，应用 AIGC、云计算、大数据、

人工智能、数字孪生、智能算法等关键技术，以数据驱动建立联宝的数字化体系根基，在研发、生产、供应、销售、服务等多个环节部署包括 PLM、MOM、QES、SAP、APS、等多个信息系统。通过系统深度应用和全面集成，形成完整的闭环管理，支撑从生产模式创新、运营模式创新、决策模式创新到商业模式创新的业务创新转型路径。

AI 赋能场景：基于 AI 的端到端生产计划和排程系统

行业痛点

为了能为客户提供独具个性的产品与服务，联宝科技每天要处理约 5000~8000 笔客户订单，其中高达 80%是单笔低于 5 台的个性化定制订单，构成数万种个性化配置组合。

解决方案

在这种“多品种、小批量”的生产模式下，联宝科技通过智能排产、智能用工等一系列先进系统，实现了“个性化订单，批量化生产”的柔性制造能力，平均 0.5 秒即可下线一台笔记本电脑。其打造的“哪吒线”（自动化主板生产线）和“水星线”（柔性组装生产线）是高度自动化与智能化生产能力的典型代表。

联宝科技先进计划与排产体系通过基于 AI 的端到端生产计划与排程平台，采用联合排产创新模式，形成连通联宝、客户、供应商的价值链整合，并利用可扩展的数字化深度学习优化算法，综合考虑产能瓶颈、物料齐套、人力匹配、供应商产能、订单优先级等制约因素，通过计划控制中心及时响应产品线、物料、人力等各类异常，敏捷响应需求变化，综合平衡整个价值链的供给与需求，确保订单准时交付。

联宝科技以 EDI、EAI 电子数据交换技术实现供应链信息互联，凭借大数据、公有云技术支撑海量数据处理与弹性扩展，融合高阶深度学习优化算法，搭配生产仿真系统进行海量数据采集，实现算法集成的超参数自学习优化，从而构建以“精准计划”和“智能排产”为特征的先进计划与排产体系，并打造与供应商合作共赢的生态协同体系。

推广价值

高敏捷：基于大数据，实现高度复杂供需场景下，涉及数十万笔订单的快速精准计划，支持不同交付类型的柔性模拟，满足客户定制化需求。

高智能：基于高阶人工智能算法，实现交付、成本等多维度最优排产，深度协同计划与执行

全生态：基于云端技术与供应商建立深度信息对接，实现供应端联合排产和统一指挥，最大化产业链综合资源。

7.4 AI 赋能新能源产业

7.4.1 引言

当前，全球能源体系正处于以新能源为主体的转型浪潮与以人工智能为核心的新一轮科技革命深

度交汇，正在重构全球能源发展的底层逻辑与竞争格局。中国能源转型进入战略深化期。《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》明确将氢能、核聚变能等纳入前瞻布局范畴，推动其成为新的经济增长点，预计未来 10 年将再造一个中国高技术产业。全国风电光伏超越煤电成为第一大电源，“玲龙一号”小型堆迈向商用阶段，新型储能装机 5 年增长 30 倍，核聚变装置实现“双亿度”突破，中国能源未来产业正以技术自主化、产业集群化、应用场景化的鲜明特征，构建起新型能源体系的四梁八柱，为全球能源转型贡献“中国思路”。

AI 赋能新能源产业，既是破解能源转型核心瓶颈的必然选择，也是培育能源领域新质生产力、筑牢国家能源安全屏障、实现“双碳”战略目标的核心引擎，更是全球能源科技竞争的核心赛道。AI 与能源产业并非单向的赋能与被赋能关系，而是双向支撑、深度耦合、共生共荣的共同体。一方面，能源是数字经济时代 AI 产业可持续发展的底层基石，为 AI 全产业链提供最根本的能源保障、场景支撑与发展空间；另一方面，AI 是能源产业转型升级的核心驱动力，通过数据、算法、算力三大核心抓手，破解能源体系发展的核心瓶颈，推动能源产业高质量发展。二者形成“能源支撑算力－算力驱动算法－算法优化能源－能源反哺 AI”的正向闭环，共同构成数字经济与绿色经济深度融合的核心底座。

1.现状与挑战：能源转型与智能革命的历史交汇

全球能源格局的深刻变革

当前全球能源格局正在经历三重根本性重构：一是能源供给结构从化石能源主导向清洁能源主导加速转型，风光等新能源已成为全球新增装机的主体，能源体系从集中式、单向流动向分布式、多向互动的形态演进；二是能源安全内涵从传统的资源保供，向供应链、技术链、数据链综合安全升级，地缘政治冲突加剧了全球能源供应链的不稳定性，各国纷纷强化能源自主可控能力；三是全球能源竞争焦点从资源储备转向技术创新，“AI+能源”已成为全球主要经济体战略布局的核心方向，美国、欧盟等均将智能化作为能源转型的核心抓手，全球能源科技竞争进入白热化阶段。

碳中和目标下的能源系统挑战

在全球碳中和共识与我国“双碳”战略目标下，传统能源体系面临着难以突破的系统性瓶颈。核心挑战集中于四大维度：一是高比例风光新能源接入带来的间歇性、波动性、随机性，传统“源随荷动”的调度模式已无法适配，新能源消纳与电网安全稳定运行的矛盾日益凸显；二是能源系统复杂度指数级上升，源网荷储多主体、多能互补多品类、交直流混联多形态的耦合，使得传统建模与控制方法触达能力天花板；三是极端天气频发对能源系统韧性提出极高要求，台风、寒潮、山火等极端事件对电网保供的冲击，传统被动式应急处置模式难以应对；四是全产业链降碳增效的精细化需求，传统粗放式的能耗管理、碳排放核算模式，无法满足“双碳”目标下的精细化管控要求。

AI 与能源深度融合的时代必然性

能源系统是典型的复杂巨系统，其海量数据、强耦合、高动态、高安全要求的核心特征，与 AI 技

术的核心能力形成天然适配。二者的深度融合，不是简单的技术叠加，而是能源体系底层范式的根本性变革，是历史交汇的必然结果。一方面，能源转型的核心瓶颈，仅靠传统能源技术升级已无法突破，亟需 AI 技术提供全新的解决方案；另一方面，AI 技术的产业化落地，需要能源领域超大规模的应用场景、海量的行业数据作为支撑，能源领域成为 AI 技术落地实体经济的核心主战场。二者的深度融合，既是全球能源转型的必然选择，也是人工智能技术产业化发展的必然方向。

2. 能源产业生态重塑核心场景

AI 不仅赋能能源技术与场景升级，更在重构能源产业的底层逻辑、商业模式与产业生态，推动能源产业从传统的资源驱动型向创新驱动型转型，核心落地场景包括：

(1) 能源新材料 AI 辅助研发

针对新型储能、氢能、光伏、CCUS（碳捕集、利用与封存技术）等领域新材料研发周期长、试错成本高的痛点，基于 AI 材料基因组技术，实现新材料的原子级精准设计、性能预测、快速筛选与合成路径优化，将传统新材料研发周期从数年缩短至数月，大幅降低研发成本，加速新型储能电极材料、氢能催化剂、高效光伏电池材料、碳捕集材料等核心材料的研发迭代与产业化落地，为前沿能源技术突破提供核心支撑。

(2) 碳全生命周期智能管理

针对“双碳”目标下碳排放核算精度低、溯源难、核查成本高的痛点，基于 AI 技术实现产品全生命周期碳足迹的精准核算、实时追踪与溯源认证，通过 AI 算法优化企业碳资产配置、碳交易策略与减排路径，为企业提供精细化的碳管理解决方案；在碳市场环节，通过 AI 实现碳价预测、交易风险预警与最优交易决策，助力碳市场的高效规范运行，推动全产业链节能降碳目标落地。

(3) 电力市场智能交易决策

针对电力市场化改革后，电价波动大、交易品种多、市场主体决策难度高的痛点，基于 AI 时序预测算法实现日前、日内、实时市场的电价精准预测，结合市场主体的发电/用电特性，构建智能交易决策模型，优化中长期合约、现货市场、辅助服务市场的投标策略与交易组合，最大化发电企业收益、最小化用电企业购电成本，同时提升电力市场的流动性与运行效率。

(4) 能源供应链智能管控

针对全球能源供应链波动大、韧性不足、管控效率低的痛点，基于 AI 大数据分析技术，实现能源供应链全流程的实时监测、供需预测、风险预警与路径优化，覆盖油气、煤炭、光伏组件、风电装备、储能电池等全品类能源产品与装备的生产、运输、仓储、交付全环节，提升能源供应链的抗风险能力与运行效率，筑牢能源供应链安全屏障。

3. 核心应用版图：AI 赋能能源全链条，助力能源系统范式跃迁

AI 对新能源产业的赋能，已从单点试点走向全链条、全场景的规模化落地，深度覆盖能源“勘探开

发－生产传输－存储消纳－消费交易－技术研发”全生命周期，围绕新型电力系统、能源生产开发、消费侧能效提升、产业生态创新四大核心领域，形成了可复制、可推广、高价值的具体应用场景矩阵，全面重构能源产业的运行逻辑与价值创造模式。

(1) 新型电力系统智能化核心场景

新型电力系统是 AI 赋能新能源产业的核心主战场，核心目标是破解高比例新能源接入带来的“不确定性、强耦合、高风险”三大核心痛点，保障电网安全稳定运行与新能源高效消纳，核心落地场景包括：

a) 新能源全周期高精度功率预测

针对风光新能源发电受气象、地形、设备状态影响大、间歇性强的痛点，融合卫星遥感、气象雷达、地面监测、设备运行、历史发电等多模态数据，基于 Transformer 时序大模型、物理信息融合神经网络等 AI 技术，实现风光电站超短期（分钟级）、短期（日级）、中长期（月年级）全周期功率预测，其中超短期预测精度可提升至 97% 以上，从源头消解新能源发电的不确定性，为电网调度、电站运行提供核心决策支撑。

b) 风光新能源集群协同智能调控

针对分散式风光电站单体容量小、布局分散、调控难度大的痛点，基于多智能体强化学习算法，实现流域水光互补、风光储一体化、海上风电集群等多场景的协同优化调控，平抑新能源集群出力波动，提升并网友好性，实现新能源集群的“可控、可调、可预测”，最大化区域新能源消纳水平。

c) 电力系统稳定在线评估与智能调度

针对高比例电力电子设备接入导致的系统惯量下降、稳定特性复杂、传统离线仿真无法适配实时工况的痛点，基于 AI 数字孪生技术构建电网全场景高保真数字镜像，实现系统电压稳定、功角稳定、频率稳定的在线实时评估与风险超前预警；基于深度强化学习算法，实现源网荷储多环节、跨省跨区大电网的全局动态优化调度，将传统人工调度的小时级决策优化至 15 分钟级，兼顾电网安全与新能源消纳的双重目标。

d) 电网故障智能识别与自愈控制

针对传统电网故障定位慢、处置流程长、人工干预效率低的痛点，基于故障录波数据、在线监测数据、视频巡检数据，通过 AI 多模态识别算法实现电网故障的毫秒级定位、类型精准识别与故障范围快速隔离；基于边缘智能控制技术，实现配电网故障后的网络自主重构与负荷快速转供，完成电网“故障预警－快速定位－自主隔离－恢复供电”的全流程自愈运行，大幅缩短故障停电时间，提升供电可靠性。

e) 极端天气下电网韧性增强智能应用

针对台风、寒潮、山火、洪涝等极端天气对电网的冲击，融合气象预报、地理信息、电网台账、历史灾害数据，通过 AI 算法实现极端天气引发的电网故障超前预警、灾害影响范围精准预判；基于多

场景推演算法优化应急物资调配、抢修队伍部署、应急电源调度的最优方案，实现电网从“被动应急处置”向“主动风险防控”的转型，大幅提升极端场景下电网的抗扰动能力与快速恢复能力。

(2) 能源智慧生产与绿色开发核心场景

AI 全面赋能传统化石能源绿色低碳转型与新能源高效开发利用，实现能源生产端全流程的安全、高效、低成本、绿色化升级，核心落地场景包括：

a) 化石能源智能勘探与绿色开采

针对油气、煤炭等传统化石能源勘探精度低、开采效率低、安全风险高的痛点，在勘探环节，通过 AI 深度学习算法实现地震勘探、测井数据的智能解译与油气藏、煤层的精准识别，将勘探周期缩短 60% 以上，大幅提升资源勘探成功率；在开采环节，基于 AI 实时优化油气田注水、压裂、采油全流程参数，实现油气田的智能开发与采收率提升；在煤炭开采环节，通过 AI 实现综采工作面的无人化智能开采、围岩稳定性智能监测与灾害超前预警，推动传统化石能源的安全、高效、绿色开发。

b) 风光电站全生命周期智能运维

针对风光电站布局分散、地处偏远、人工巡检成本高、故障处置滞后的痛点，基于无人机巡检、机器人巡检、红外热成像、振动监测等多模态数据，通过 AI 视觉识别、异常检测算法，实现光伏组件隐裂、热斑、风机叶片裂纹、齿轮箱故障等设备缺陷的精准识别与提前预警，将运维模式从“事后维修”“定期巡检”升级为“预测性维护”，降低设备非计划停机时间，提升电站发电量与全生命周期收益，实现电站无人化、智能化运维。

c) 核能全流程智能安全管控

针对核能运行高安全、高可靠的刚性要求，基于 AI+数字孪生技术构建核反应堆全工况高保真仿真模型，实现反应堆运行工况的实时模拟、事故场景推演与应急处置方案优化；通过 AI 多模态监测技术，实现核岛设备、管道、安全壳的全时段智能监测与缺陷识别，提前预警设备老化与安全风险；在核燃料循环、放射性废物处置环节，通过 AI 实现全流程智能管控与安全监管，全面提升核能运行的安全性与可靠性。

d) 氢能制储运加全链条智能优化

针对氢能产业电解槽效率低、储运安全风险高、加注成本高的痛点，在制氢环节，通过 AI 算法实时优化可再生能源制氢的电解槽运行参数，适配风光发电的波动特性，提升绿氢制取效率，降低制氢能耗；在储运环节，通过 AI 实现储氢容器、输氢管道的实时安全监测与泄漏超前预警，保障氢能储运安全；在加注环节，通过 AI 优化加氢站运行调度，适配氢燃料电池汽车加注需求，提升加注效率与运营安全性，推动氢能全产业链的规模化发展。

(3) 灵活高效的智慧用能核心场景

AI 推动能源消费侧从“被动刚性用能”向“主动柔性互动”转型，挖掘需求侧资源潜力，实现全社会能效水平的系统性提升，核心落地场景包括：

a) 虚拟电厂资源聚合与优化调度

针对分布式光伏、储能、充电桩、柔性负荷等资源分散、单体容量小、难以参与电网调度的痛点，AI 作为虚拟电厂的“智慧大脑”，实现海量分布式资源的全量感知、精准建模、聚合调控与市场化交易。通过 AI 算法精准预测各分布式资源的可调潜力，优化聚合调控策略，引导柔性负荷错峰用电、储能充放优化，将分散的需求侧资源聚合形成可调节的“虚拟电源”，参与电网调峰调频、备用辅助服务与电力市场交易，释放需求侧资源的系统价值，缓解电网峰谷差压力。

b) 工业领域智能能效管控与优化

针对工业领域能耗高、节能潜力大、精细化管控不足的痛点，基于 AI 技术实现工业企业用能数据的实时采集、全流程分析与能效对标，精准识别无效能耗、节能潜力环节；通过 AI 算法优化生产工艺流程、用能设备运行参数与用能计划，实现峰谷电价套利、余热余压高效利用、公用工程系统智能优化，在不影响生产效率的前提下，实现工业企业综合能效提升 5%—20%，助力工业领域节能降碳。

c) 建筑智慧用能与能效优化

针对建筑领域空调、照明、电梯等设备能耗占比高、用能粗放的痛点，基于 AI 技术融合建筑结构、气象数据、人员流动、设备运行等多维度数据，构建建筑用能数字孪生模型，实现空调、照明、通风等系统的自主智能调控，在保障室内舒适度的前提下，最大化降低建筑能耗；同时通过 AI 实现建筑设备的预测性维护与故障预警，提升建筑运维效率，实现公共建筑、商业楼宇、居民住宅的全场景智慧用能。

d) 电动汽车与电网智能互动 (V2G)

针对电动汽车大规模接入带来的电网负荷冲击、充电资源供需错配的痛点，通过 AI 算法精准预测电动汽车充电行为、时空分布与可调潜力，优化充电桩的有序充电调度，引导用户在电网低谷时段充电、高峰时段向电网反向送电，实现电动汽车与电网的双向智能互动；将海量电动汽车动力电池转化为分布式移动储能资源，参与电网调峰调频、新能源消纳，实现交通能源与电力系统的深度融合。

e) 综合能源服务智能优化

针对园区、社区、商业综合体等多能互补场景，基于 AI 技术统筹电、热、冷、气、氢等多种能源品类，构建多能流协同优化模型，实现多种能源的生产、转换、存储、消费全流程协同优化，兼顾用户用能需求、用能成本与碳排放目标，为用户提供定制化的综合能源服务解决方案，实现区域综合能效最大化、用能成本最小化。

4. 前瞻展望：迈向安全、绿色、经济的智慧能源时代

AI 赋能新能源产业，是一个循序渐进、持续深化的过程，将伴随我国“双碳”目标的实现进程，分阶段完成从单点技术突破到全系统范式重构，最终建成安全、绿色、经济、普惠的智慧能源体系。

近期路径（2026-2030）：重点环节突破与规模化应用

本阶段是我国碳达峰目标的关键攻坚期，核心目标是实现 AI 在能源核心场景的规模化落地与重点环节突破。重点推进新能源功率预测、电站智能运维、电网智能调度、虚拟电厂、工业能效优化等成

熟场景的规模化应用，实现重点环节的降本增效与价值释放；集中突破能源行业大模型、高保真数字孪生、国产化专用芯片等核心技术，完善能源智能化的数据标准、安全标准、应用标准体系，初步建成云边协同的能源算力基础设施；推动 AI 与能源融合的产业生态初步成型，为新型电力系统建设提供核心支撑，助力 2030 年前碳达峰目标顺利实现。

中期愿景（2030-2045）：跨域深度融合与生态全面成型

本阶段，AI 与能源体系将实现全链条、全领域的深度融合，能源系统完成从“功能系统”向“智能生命体”的根本性范式跃迁。核心特征体现为：AI 成为能源系统的核心决策中枢，实现源网荷储全要素、全时空的全局协同优化，电网全面具备自适应、自愈合的智能运行能力；海量分布式资源实现全量聚合与市场化配置，虚拟电厂成为电力系统的重要组成部分；氢能、新型储能等前沿能源技术在 AI 的赋能下实现规模化应用，跨交通、建筑、工业等领域的能源协同体系全面建成；形成技术、产业、标准、安全协同发展的成熟产业生态，能源系统的安全性、低碳性、经济性实现质的飞跃。

远期图景（2045-2060）：智慧能源共同体与全球可持续发展

本阶段，将全面建成以 AI 为核心驱动的智慧能源体系，实现能源系统的全面零碳化、智能化、普惠化，最终形成全球协同的智慧能源共同体。AI 将实现全球能源网络的跨区域、跨国界协同优化，清洁能源成为全球能源供给的绝对主体，能源供需实现全时空的精准匹配；能源系统与生态环境实现和谐共生，全社会能效水平达到极致，2060 年前碳中和目标全面实现；能源贫困问题得到全面解决，能源服务实现普惠共享，为全球可持续发展与气候变化治理提供中国方案与中国智慧，助力建成人与自然和谐共生的现代化。

7.4.2 典型案例：华为数字能源

华为数字能源是华为全资子公司，成立于 2021 年，深耕能源数字化与智能化领域，提供智能光伏、数据中心能源、智能充电网络、智能电动等解决方案，业务覆盖全球 170 多个国家和地区。

AI 赋能场景：构网型储能电站全生命周期智能运营

行业痛点

在“双碳”目标引领下，新能源在电力系统中的占比持续提升，储能成为平抑新能源出力波动、保障电网稳定的核心设施。传统储能系统的运营模式无法适配新型电力系统的运行要求，储能电站在电网支撑、安全管控、经济运营、调度协同等方面面临显著痛点，成为制约储能行业规模化发展的关键因素。

电网主动支撑能力严重不足，难以适配高比例新能源电网运行需求。传统储能多为跟随型控制，被动响应调度指令，不具备主动构网能力，无法模拟同步机惯性与阻尼特性。随着新能源占比提升，电网等效惯性降低，传统储能难以提供有效惯量支撑和快速调压调频服务；在沙漠戈壁新能源大基地、偏远微网等弱电网场景下，其对复杂工况适应性差，易引发并网震荡、电压越限等风险，无法满足新

型电力系统安全稳定运行需求。

安全管控与运维模式粗放滞后，安全隐患与成本压力并存。近年来储能电站安全事故频发，传统运维以定期巡检和事后处置为主，无法提前识别电芯内短路、漏液等隐性风险。同时，行业普遍存在 SOC（荷电状态）估算精度不足问题，依赖人工现场标定，不仅成本高昂，还会造成停机弃电损失；在规模化储能电站快速落地后，传统人工运维效率低、响应滞后，已无法适配大型电站安全运行需求。

全生命周期经济性不足，收益模式单一且不确定性高。项目设计依赖经验参数，电池容量配置与系统匹配度不足，易出现过度投资或后期容量短缺；运营阶段充放电策略固化，无法结合电价、辅助服务价格等动态优化，收益水平偏低。加之电池衰减规律难以预判，系统效率逐年下滑，导致项目实际收益与可研偏差大、回报周期长，难以适应独立储能参与电力市场的趋势。

多能源协同调度能力薄弱，多元场景价值难以释放。传统储能以单一充放电控制为主，缺乏光储荷协同优化能力，无法实现源网荷储高效联动。其对电网辅助服务指令的响应速度和精度不足，难以深度参与辅助服务市场获取溢价；在离网微网、黑启动等复杂场景下，缺乏统一智能调度架构，多设备协同能力弱，供电可靠性不足，储能综合价值无法充分挖掘。

解决方案

华为数字能源的 FusionSolar 智能组串式构网型解决方案将人工智能技术深度融入构网型储能电站全生命周期智能运营场景，实现从电芯到电网的全链路状态感知、智能决策和动态优化，针对性解决行业痛点，构建“感知、决策、执行、优化”的闭环运营体系，保障储能电站安全、稳定、经济运行。

自适应智能控制，强化电网稳定支撑能力。针对高比例新能源下电网惯性不足、弱电网适应性差等行业共性问题，方案内置电网 AI 态势感知算法，实时监测电网电压、频率、谐波、短路容量等运行状态，动态自适应调整控制策略。通过虚拟同步机技术模拟同步发电机运行特性，提供可调惯量支撑，在电网频率扰动时实现毫秒级功率响应，在电压波动时快速输出大容量无功支撑，有效提升系统阻尼与电压频率稳定性。同时，依托 AI 振荡抑制算法，实时辨识系统阻抗特性，主动规避谐振风险，将电流谐波控制在极低水平，可在宽范围弱电网条件下稳定并网运行，完美适配沙漠戈壁大基地、偏远微网等复杂场景，满足新型电力系统对构网型储能的刚性需求。

AI 智能诊断预警，实现安全精益化运维。搭建云边端协同的 AI 智能诊断预警体系，通过在储能系统关键环节部署传感网络，持续采集电芯电压、电流、温度、内阻及电池簇绝缘阻抗、环境温湿度等多维度运行数据，依托机器学习算法构建设备健康状态评估和故障预测模型，可提前 24 小时以上对电芯内短路、电解液泄漏、绝缘受损等 10 余种潜在故障进行预警，为极端安全事故防范预留关键处置时间。打造“储能健康卫士”智能诊断功能，实现电芯、电池包、簇、系统、电网五个维度的深度扫描，生成结构化、层次化的健康评估报告，为运维策略制定提供科学依据。通过高精度电压采样电路、专用电池管理芯片及先进的自适应滤波算法，实现 SOC 全自动高精度在线标定，长期运行估算误差稳定控

制在 3%以内，彻底替代人工校准，实现关键状态参数的免维护管理。

全生命周期 AI 优化，提升储能项目经济收益。在项目设计阶段，基于海量电芯实测运行数据构建电化学与热力学数字孪生模型，精准模拟不同温度、充放电倍率、循环工况下的电池性能衰减与能量转换效率，为系统容量配置、拓扑结构设计提供数据支撑，在满足能量吞吐指标的前提下优化冗余配比，减少过度投资造成的成本浪费，提升系统循环效率。通过精细化分布式控制架构，实现“一簇一管理、一包一优化”，实时感知电芯一致性差异并进行动态补偿，保障储能系统在全 SOC 区间内稳定恒功率输出，进一步优化初始电池投资。在运营阶段，结合实时电价信号、新能源出力预测、电网调度指令及电池健康状态，通过多目标优化算法动态调整充放电策略，提升电池全生命周期发电量与能量利用率。依托构网型控制的快速响应能力，储能系统可高效参与电网调频、调峰等辅助服务市场，拓宽盈利渠道，增强项目收益稳定性，有效改善储能电站投资回报周期长、收益不确定性高等行业共性问题。

协同智能调度，支撑源网荷储一体化高效运行。搭建全国产化自主可控的智能调度平台，采用高可靠硬件架构与安全操作系统，通过权威网络安全认证，保障系统运行稳定可靠，实现全站运行数据毫秒级采集、秒级刷新与全局协同决策。在精准感知光伏出力、储能 SOC、负荷变化及电网调度指令的基础上，通过多目标优化算法平衡调度要求、经济效益与电池寿命，在平抑新能源波动的同时，规避高倍率充放电、频繁浅循环等影响电池寿命的工况。在离网微网场景中，根据光储荷实时功率平衡状态自适应调节输出电压与频率，保障负荷供电持续稳定；在极端故障工况下，通过多 PCS 群控协同、并联缓启与同步升压算法，实现 GW 级大规模储能系统分钟级黑启动，无需外部启动电源即可快速恢复供电，已在海外大型光储项目中实现无油机调试应用。方案可灵活适配并网运行、独立储能、共享储能、微网离网、辅助服务等多元应用场景，真正实现源网荷储高效协同，充分释放储能系统综合价值。

推广价值

华为数字能源构网型储能电站全生命周期智能运营的 AI 赋能实践，是人工智能在储能行业的典型落地案例，针对新型电力系统下储能运营核心痛点形成系统性解决方案，为储能项目带来安全与经济价值，也为行业智能化转型提供可复制范本，核心推广价值体现在三个方面：

经济价值显著，实现全环节降本增效。该方案通过 SOC 免维护标定、预测性维护等技术，提升储能电站运维效率、规避停机损失，有效缩短项目投资回报周期，可适配各类储能场景，具备全球落地条件。

技术价值突出，推动行业技术升级。方案通过 AI 打破储能各环节技术壁垒，实现数据全链路互通与协同，验证了云边端协同及 AI 技术的可行性，核心构网技术通过全球权威标准认证，为行业数字化、全球化推广奠定基础。

行业价值深远，助力新型电力系统建设。方案推动储能行业从设备管理向系统级智能运营转型，适配新能源配套储能场景，解决电网稳定问题，推动储能电站向电网服务提供商转型，支撑高比例新能源并网，助力“双碳”目标实现。

7.4.3 典型案例：比亚迪储能

比亚迪储能始于 2008 年，专注于储能系统及设备的研发和推广应用；已形成集储能产品研发、制造、销售、服务、回收利用于一体的完整产业链，产品全面覆盖电源侧储能、电网侧储能、工商业储能、兆瓦闪充、家庭储能、功率系统、移动储充、综合能源、新型应用业务、网络能源等应用领域。

AI 赋能场景 1:工商业储能产品 Chess Plus（储能）

行业痛点

工商业用户侧储能的决策，表面上是一次设备投资，实质上更像一项需要长期稳定运营的资产配置。许多园区、工厂或大型充电场站在引入储能后，会遇到三类典型难题。首先是安全边界难以被持续验证，储能系统长期处于高频充放电和复杂环境中，热异常、绝缘老化、辅机失效等问题具有隐蔽性，一旦演化为事故，不仅会造成资产损失，还会带来停产停业、消防合规等外部成本。然后是效率与收益的不确定性，项目收益高度依赖电价机制、峰谷差、负荷曲线与设备可用率，若运行策略静态固化或故障处理滞后，常见结果是“装得起却跑不稳”，实际循环次数、可利用容量和经济性偏离测算。最后是运维管理的现实约束，工商业储能往往分布式部署，现场运维力量有限，传统做法依赖人工巡检和事后告警，定位慢、恢复慢，进而放大停机损失。

解决方案

针对这些问题，比亚迪储能在其面向外部发布的 Chess Plus 产品公告中，直接将“安全、效率和盈利”作为系统要解决的核心问题。比亚迪储能在 Chess Plus 的产品体系中，将安全防护与智能化能力按系统级进行组合设计，而不是把智能控制作为可选外挂。Chess Plus 以电芯到系统一体化（CTS）保护框架为基础，并在系统级防护中引入 AI 驱动的风险预测算法，用于对热异常进行提前检测，从而把风险处置从事故发生后的被动响应前移到异常阶段的提前识别。在运行控制层面，该系统集成了高算力边缘计算设备，可实现对 SOC 算法的实时优化与对故障的实时预警，这意味着关键判断不必完全依赖云端回传和人工研判，而可以在现场以更短链路完成监测、诊断、预警的闭环，提高可用率并减少无效停机。Chess Plus 具有高度集成与模块化特征，系统将 BMS、分布式管理控制、能量管理与控制单元、EMS 以及 AC Unit 等整合在同一产品形态中，使数据采集、控制执行与场景适配更标准化，从工程实施角度降低了系统集成复杂度，为后续的智能策略迭代和规模化复制创造条件。

推广价值

该案例的推广意义主要体现在把储能做成可运营资产的能力上。对于工商业用户而言，AI 风险预测与边缘计算并不只是技术点缀，其价值主要体现在风险成本的可控化。当热异常能够被提前识别、

故障能够被实时预警，现场运维可以更早介入，系统更可能以小代价消除隐患，从而降低停机、整改与安全事件带来的外部性成本。其次，它提升了收益兑现的确定性，因为储能收益依赖长期稳定运行与策略执行，一旦可用率提升、故障恢复时间缩短，项目的有效循环与放电贡献更接近测算值，财务结果更稳定。最后，Chess Plus 明确面向工业园区、充电中心与微电网等多场景应用，这些场景恰是未来分布式能源与电动化基础设施扩张的主要承载地。系统以标准化集成形态与本地智能决策能力应对场景差异，有助于在不同地区电价规则、负荷结构与建设条件下实现更高的迁移性与复制效率，从而推动工商业储能从项目制一次性交付向可规模化、可持续运营的能源服务资产演进。

AI 赋能场景 2:光伏与储能协同调度项目（发电端）

行业痛点

光伏发电受光照强度、天气变化等自然因素影响，存在发电量波动大、稳定性差的问题，易导致光伏电力难以高效接入电网，出现弃光现象。同时，光伏与储能系统协同性不足，储能充放电调度依赖人工设定，无法根据光伏发电量、用电负荷的实时变化动态调整，导致储能系统利用率低，光伏电力的消纳效率不高，难以实现自发自用、余电高效存储的目标，尤其在工业场景中，工厂用电负荷波动大，光伏与储能的协同调度不合理会增加企业用电成本，影响能源利用效率。

解决方案

比亚迪储能与 Apparent 公司合作，在美国兰开斯特工厂打造光伏与储能协同调度项目，核心是引入智能电网操作系统（IGO S™），实现光伏发电与储能系统的智能化协同调度，破解光伏发电稳定性差、储能利用率低的痛点。该方案中，智能系统实时采集光伏组件发电量、工厂用电负荷、电网电价、天气预测等多维度数据，通过智能算法对未来 24 小时的光伏发电量、用电负荷进行精准预测。基于预测结果，智能系统动态调整储能系统的充放电策略，当光伏发电量高峰、用电负荷低谷时，控制储能系统全力充电，避免弃光；当光伏发电量低谷、用电负荷高峰时，控制储能系统放电，补充工厂用电缺口，实现光伏电力的高效消纳。同时，智能系统通过高效通信与控制技术，管理光伏与储能系统的有功功率和无功功率，确保光伏电力稳定接入电网，减少对电网的冲击。此外，智能系统可根据工厂用电习惯、电网峰谷电价差异，优化储能充放电时序，实现峰谷电价套利，进一步降低企业用电成本。该项目配备 2 MW 光伏组件和 2 MWh 储能系统，采用比亚迪储能高效光伏组件与储能技术，智能系统全程自主调度，无需人工干预，实现能源生产与消费的动态平衡。

推广价值

该案例的推广价值体现在行业示范与技术赋能两个层面。一是为工业场景光伏与储能一体化项目提供智能化标杆，解决了光伏发电波动性与用电负荷不均衡的核心矛盾，实现了光伏电力的高效消纳与储能系统的高效利用。该项目预计每年可为兰开斯特工厂节省 10 万美元以上的能源成本，其模式可广泛推广至制造业、仓储物流等工业场景，助力工业企业实现能源结构转型，降低用电成本，提升能

源自给率。二是推动光伏与储能行业的智能化融合发展，协同调度技术打破了光伏与储能各自为战的局面，实现了发电与储能的智能联动，为大规模光伏基地、分布式光伏项目的高效运营提供了可复制的解决方案，有助于减少弃光现象，提升光伏能源在全球能源结构中的占比。三是验证了智能技术在能源调度中的核心价值，为未来能源互联网的构建提供了技术支撑，推动未来能源向分布式、智能化、协同化方向发展，契合全球能源转型的核心趋势。

7.4.4 典型案例：运达能源

运达能源科技集团股份有限公司（以下简称：运达能源）在国内最早从事新能源技术与产品开发，是中国新能源事业的拓荒者、创新者和领军者，曾研制出我国第一台并网型风力发电机组，至今深耕新能源已有 50 多年历史。截至 2024 年 12 月，产品遍及国内外 600 余个风电场，机组最长运行时间超过 20 年，累计发电量超过 3.6 万亿千瓦时，相当于减排二氧化碳约 36 亿吨（数据来源其官网首页）。

随着国家“30·60”碳达峰、碳中和目标的提出，运达能源建立以风电装备制造、新能源电站投资开发与运营、新能源工程总承包、储能系统解决方案、综合能源服务、新能源消纳于一体的六大业务生态，积极打造以新能源和碳排放管理为主的现代化一流新能源企业。运达能源积极推动 AI 与风电技术的融合创新，聚焦风机智能控制、功率预测、智慧运维等核心场景，打造智能化风电解决方案，助力风电产业降本增效、提升可靠性，推动可再生能源的高效开发与利用，为新型电力系统构建提供风电领域的核心支撑。

AI 赋能场景：风电全生命周期 AI 智能管理

行业痛点

发电性能损失：风电场实际运行与理论设计存在显著偏差，固定控制参数无法适配复杂工况，导致发电性能持续损失。设备层面，部件制造、安装、测量误差及风向标松动、叶片气动外形改变、部件老化等问题，使偏航、变桨、变速控制偏离最优状态；环境层面，设计仿真条件与实际运行环境差异、机位间环境差异，进一步加剧了发电量损失。

叶片交付过程质量监测：在风电机组大型化背景下，叶片交付链条长、参与方众多，人员操作的复杂性和监管难度高，使得叶片易引发本体损伤，成为行业最大业务痛点。制造环节存在工艺缺陷，运输过程易发生磕碰、私修，安装时操作不规范、吊点不当也会造成损伤；同时，交付过程中留痕材料缺失不规范导致事故难以追溯，庞大的过程材料也使人工审核难度大，质量风险难以及时发现。

运维管理：传统巡检方式效率低、风险高且过度依赖人工经验。外部巡检依赖无人机，成本高、流程繁琐，且采集的海量视频图像缺乏高效缺陷识别算法；内部巡检依赖人工，效率低、覆盖范围小，密闭环境存在安全隐患，同时人工审核工作量大，严重依赖专家经验。

能源调度：随着新能源并网规模扩大，电网对风功率预测精度要求持续升级，考核压力与经济损

失显著。电网双细则考核严格，100MW 风电场年考核金额可达 100 万—150 万元，预测准确率每提升 1%可节约约 18 万元；同时，短期、超短期预测精度要求不断提高，考核时间窗口从单日扩展至未来 3~4 天，对预测系统的准确性和时效性提出更高要求。

解决方案

运达能源以高质量数据为核心基础，依托 SCADA 系统搭建统一数据中台，融合机器学习、计算机视觉等 AI 技术，为该风电场打造了四大环节深度协同、数据互通的全生命周期 AI 智能管理体系，各环节解决方案精准对接痛点，实现技术联动。

智能控制：作为智慧电站的重要模块，运达能源自研的风电场发电量智能优化系统融合数据挖掘、AI 技术，辨识出风电机组偏航、变桨、变速的最优参数，实现“一机一策”，场均发电性能提升量达到 1.5%—4%。

利用 AI 视觉分析技术的叶片交付质量监测系统：基于 Maximo 平台搭建网页端+App 双端操作体系，集成 20+个视觉分析算法模型，实现叶片制造、运输、安装全交付环节的质量管控。通过 AI 视觉分析完成交付留痕材料的智能审核、人员异常行为的实时识别、叶片本体损伤的自动告警；打通质量管控数字化流程，实现管理指标自动化统计，将质量风险从“事后追溯”转变为“事前预警、事中管控”，同时为后续运维环节提供完整的叶片质量基础数据。

多设备协同智能巡检系统运维管理：打造“定航无人机+手持设备+轮式/轨道内检机器人”的多设备协同 AI 巡检体系，对接质量监测系统的基础数据，实现“质量数据—巡检任务—缺陷处理”的联动。外部巡检采用手持设备日常巡检+无人机定期定航巡检的模式，降低无人机依赖，提升巡检频次与覆盖率；内部巡检通过轮式/轨道机器人替代人工，提升巡检进入深度和损伤检测算法准确率；同时搭建自研智能巡检系统，实现巡检工单自动下发、影像存储管理、缺陷告警、报告自动生成、专家问诊的全流程数字化，大幅提升运维效率。

AI 多源融合风功率预测系统实现能源调度：整合欧洲、美国、中国、日本多气象源数据及场内测风塔、SCADA 系统运行数据，通过 AI 技术完成多气象源订正与融合，提升气象预报精度；构建多 AI 预测模型，采用“自动择优、自动修模”策略，平抑预测结果的时空波动，提升功率预测准确率；预测结果不仅对接电网调度要求，同时为场内发电量智能优化系统提供气象预判数据，实现“功率预测—发电优化”的协同调度，让发电控制参数提前适配气象变化。

推广价值

该风电全生命周期 AI 一体化应用案例，是运达能源将 AI 技术与风电行业深度融合的典型实践，不仅为该风电场带来了可量化的经济收益，更为风电行业的智能化转型提供了“全流程、可复制、高适配”的实践范本，核心推广价值体现在三方面：经济价值方面，实现风电全环节的降本增效，将 AI 技术的“增收、降本、提质”价值落到实处，投资回报周期短，适配国内不同规模、不同类型的陆/海风电场；

在技术价值方面，打破风电各环节的数据孤岛与技术壁垒，实现智能控制、工程建设、运维管理、能源调度的技术协同、数据互通，验证了 AI 技术在风电全生命周期落地的可行性，为风电数字化中台建设提供了技术参考；在行业价值方面，推动风电行业从“分散化、经验化”的传统运营模式，向一体化、智能化、数据化的现代运营模式转型，证明 AI 技术并非单一环节的“技术加持”，而是风电全生命周期运营的核心驱动力，为新能源行业的智能化升级树立了标杆，可复制至光伏、水电等其他新能源领域的全流程运营管理。

7.5 AI 赋能健康医疗产业

7.5.1 引言

相较于传统的医疗模式，AI 凭借其数据驱动、高效计算、模式识别等卓越的技术优势，正在深刻地改变着大健康产业的各个领域。AI 技术纵向渗透，覆盖疾病预测、诊断、治疗、康复的全周期管理。AI 技术的应用横向扩展，从临床医疗延伸至医院运营、医学研究、药物研发等。通过 AI 赋能实现价值重构，不断推动效率的提高、医疗成本的降低、精准度的提升和个性化服务的发展，从而实现更普惠、更精准的医疗服务。

1. AI 驱动全民健康重构

聚焦 2025 年全国两会热点议题，健康中国三期战略备受瞩目，旨在构建覆盖全民、全生命周期的健康服务体系。现有医疗体系的核心矛盾在于医疗资源分布不均（三甲医院占比不足 5%，却承担 50%+ 重症诊疗）与基层服务能力薄弱。为了实现医疗资源普惠和全民健康的目标，解决医疗资源紧张、医疗服务质量差异化的问题，需要 AI 技术强大的助推力。

AI 赋能全民健康数据整合：构建国家级健康大数据平台（如“全民健康信息互联互通”工程），融合基因组学、电子病历、可穿戴设备等多模态数据，突破数据孤岛；

AI 赋能基层医疗：AI 辅助诊断系统下沉社区医院，提供实时决策支持等；

AI 赋能预防医学：基于大模型的慢性病风险预测，帮助预警疾病风险，方便提早干预。

最终，通过技术平权、政策支持和技术基础设施建设等，重塑健康权利分配，实现全民共享。

2. AI 驱动个性化健康革新

深度整合个体基因、行为数据、环境数据、健康数据、医疗记录和科研成果等，通过动态建模挖掘基因 - 环境 - 疾病关联，通过模型量化遗传风险以及预测疾病风险，通过强化学习优化健康管理方案，结合指南形成个体化的循证治疗方案，结合生活方式的长期效果予以健康建议，构建“千人千面”甚至是“一人千面”的健康管理体系，为用户提供更精准的疾病预测、更有效的治疗方案，以及更全面的健康管理服务，推动医疗从标准化治疗向精准预防与干预转型，让真正的“治未病”成为可能。

3. AI 医疗大模型落地

中国医疗 AI 大模型已走过爆发期，进入深度应用与价值验证阶段。2025 年 6 月，瑞金医院联合华

为开源 RuiPath 病理模型，加速行业智能化落地。RuiPath 病理模型是由瑞金医院在华为团队支持下研发的临床级多模态病理大模型，涵盖泛癌种视觉特征提取、视觉-语言跨层表征对齐、长序列的深度学习模型训练机制等能力，用于全流程临床病理辅助诊断，目前已覆盖中国每年全癌种发病人数 90% 的 19 个常见癌种，涵盖上百个辅助诊断任务。

除了医疗大模型，中国还在加快细胞和基因技术、合成生物、生物育种等前沿技术产业化，推动 5G/6G、元宇宙、人工智能等技术赋能新型医疗服务，研发融合数字孪生、脑机交互等先进技术的高端医疗装备和健康用品。

以下是 AI 在健康医疗产业中的几个关键应用场景及未来趋势：

场景 1：医疗诊断与影像分析

1. 技术实现

AI 在医疗影像诊断中的辅助作用，离不开深度学习、卷积神经网络（CNN, Convolutional Neural Networks）、迁移学习等先进技术的支持。这些技术共同构成了 AI 在医疗影像分析中的核心框架，使得 AI 能够实现对影像数据的自动分析、处理和诊断。其中，CNN 是深度学习中最常用的模型之一，特别适用于处理图像数据。它通过卷积层、池化层、全连接层等结构，对影像进行特征提取和降维，实现高效的图像分类和识别。在医疗影像诊断中，CNN 被广泛应用于病灶检测、病变分类等任务。通过不断优化模型结构和参数，CNN 能够提高诊断的精准度和效率。

2. 实战案例

Viz.ai 发布的脑卒中快速检测系统：Viz.ai LVO（Large vessel occlusion，大血管闭塞），又称 Viz ContaCT，利用 AI 算法分析 CT 神经图像，检测与脑卒中相关的指标，为医生提供脑卒中临床支持。Viz ContaCT 可以在扫描后 6 分钟内判断是否大血管闭塞，而传统方法大概需要一个小时。对 139 家医院的 2544 名患者的 CTA（CT angiography，CT 血管造影）影像进行分析，表现出 96% 的敏感性和 94% 的特异性。患者从诊断 LVO 到首次接触血管内外科医生的时间减少了 44.13%。而且，通过保留不必要转移的患者，初级卒中护理中心将获得巨大的经济效益。

我国医学影像头部企业联影医疗在 AI 技术的应用中采取了平台化的创新策略，通过跨产品线的技术整合，打造了一个完整的智能化平台。在磁共振（MR）、计算机断层扫描（CT）、分子影像（MI）等多个领域的 AI 技术应用已取得显著进展。特别是在 MR 领域，联影医疗通过 uAIFI 类脑平台实现了 0.5 秒/期的快速动态高清成像；在 CT 领域，通过 AI 技术实现低剂量扫描和快速图像重建，显著提升了图像质量和检查效率。此外，联影医疗还推出了多款具有行业领先优势的创新产品，如全球首款 5T 磁共振产品 uMI Jupiter 和新一代分子影像产品 uMI Panorama 系列，这些创新产品不仅填补了行业技术空白，还帮助医疗机构提高了设备的智能化水平。

3. 行业影响

AI 技术的引入，为医疗影像诊断带来了前所未有的变革：提高诊断精准度、加快诊断速度、降低人为误差、促进医疗资源均衡等。

场景 2：药物研发与临床试验

1.技术实现

计算机视觉、NLP（Natural Language Processing，自然语言处理）和大模型以及生成式 AI 展现出在处理大规模数据方面的能力，在药物研发的关键挑战中展现出巨大潜力。AI 已被证明在分析复杂生物系统、识别疾病生物标志物和潜在药物靶点、模拟药物－靶点相互作用、预测药物候选物的安全性和有效性以及管理临床试验等方面具有重要价值。

2.实战案例

AI 赋能新药研发

PandaOmics 平台通过整合多组学数据和生物网络分析，成功发现了一种新的治疗靶点：TRAF2-NCK 相互作用激酶，并据此开发出了针对性的抑制剂 INS018_055，为抗纤维化治疗开辟了新途径。

AI 赋能临床试验

TrialGPT 创新性地使用大语言模型辅助患者到临床试验匹配的过程，加速了临床试验的招募，可以将临床试验的搜索空间降到原始的 6%，同时保证 90%以上对相关试验的召回，可以用于精确地推荐患者，还能节省 42.6%的时间，使得平均每个临床试验的筛选从 61.5 秒降低到 35.3 秒。

3.行业未来趋势

AI+ 药物研发将与生物信息学、基因编辑、计算化学等前沿科学领域深度融合，提供更为全面和深入的医疗解决方案，实现药物研发的个性化、精准化。

场景 3：健康管理与慢病防控

1.技术实现

AI 在慢性病管理中的应用，主要依靠数据驱动的智能算法，通过深入分析和预测患者的健康状态，构建从风险预警到个性化干预的全链条技术体系，从而达到健康管理与慢病防控的目的。首先，AI 系统通过构建知识库、整合健康记录、建立概率模型和进行场景模拟等技术手段，收集和处理患者的多维数据，以提高预测的全面性和准确性。多模态数据包括遗传信息、可穿戴设备、环境数据、临床数据（病历文本数据、检测量化数据等）等，并按照一定的规则进行清洗、统一和标准化。AI 系统通过不断学习这些历史数据和医学知识，通过集成学习和神经网络（GNN），从复杂且庞大的数据集中提取出深层次的模式，能够识别出潜在的疾病风险，预测病情的未来发展趋势；还可以通过强化学习、生成对抗网络（GAN），通过动态决策与数据生成能力，提供个性化的健康管理方案。通过模拟患者的健康行为（如饮食、锻炼等），AI 还可以帮助患者理解其行为对健康的影响，并建议改善措施，从

而实现对慢性病的持续监控与管理。

2.实战案例

AI 赋能慢病管理

Babylon Health 利用 AI 技术将医疗服务转变为更加便捷、实惠的模式，特别是在慢性病的管理上。通过集成患者的健康记录、生活习惯数据以及医学知识库，为患者提供个性化的健康管理服务。患者的病史、生活环境和生理数据将被录入系统，通过算法分析后，AI 机器人能够识别出潜在的健康风险。还可以模拟不同健康行为对疾病发展的影响，帮助患者了解健康习惯的长期效果，进而调整其饮食、锻炼和睡眠等日常行为，从而控制慢性病的发展。

AI 赋能可穿戴设备监测

通过多模态传感器与 AI 算法的深度结合，重塑了健康监测的定义，从基础生理指标追踪进阶至疾病早期预警。以华为 WATCH 系列为例，Watch D 使用的是和腕式血压仪一样的气泵式测量，可以达到专业的监测血压效果。Watch 4 的健康功能更是实现重大突破，通过“微体检”，用户可快速进行 10 项健康数据的监测和风险评估，快速检测心率、血氧、压力、体温、心电图、血管弹性、血管风险筛查、肺功能评估、慢阻肺风险筛查、肺部感染风险筛查，并在手表和 App 端同时生成微体检报告，整个过程在 60 秒左右。

3.行业影响

AI 驱动的健康管理与慢病防控技术，正在从“单点突破”走向“系统重构”。其核心是通过数据 - 算法 - 硬件的闭环迭代，使得医疗从“经验驱动的手工业”转变为“数据驱动的精密科学”，不仅提升了效率，更能将被动医疗转化为主动健康管理，未来进一步推动行业生态重构（如保险、制药、医疗服务、支付体系等）。

场景 4：精准医疗与基因分析

1.技术实现

在精准医疗领域，AI 凭借其强大的处理和分析数据的能力，通过整合高通量基因测序、多组学数据与深度学习算法，推动了从基因变异解读、个性化治疗决策到预测患者对特定治疗反应的全程智能化。

2.实战案例

AI 赋能靶向治疗匹配

Guardant360 利用次世代基因定序检测 74 个癌症基因突变，结合患者基因及 OncoKB 数据库，推荐美国食品药品监督管理局批准药物。例如，为 EGFR L858R 突变的转移性非小细胞肺癌患者推荐奥希替尼治疗；监测 EGFR T790M 突变丰度，提前 4 周预警奥希替尼耐药。

AI 赋能基因分析

基因数据是精准医疗的基石，作为中国基因行业的龙头公司，华大基因凭借百万级基因组数据库和自主测序技术（DNBSEQ 平台），解读基因测序数据，识别治病突变，提供个人疾病风险预测、药物基因组指导等，在生育健康、健康管理、肿瘤防控等核心领域将保持技术壁垒。例如，为有效应对心血管疾病挑战，华大基因推出了心血管疾病风险评估基因检测（含药物基因组）。该技术融合了目标区域捕获和高通量测序技术，能综合评估遗传因素所导致的心血管疾病的潜在风险，精确预测个体对心血管治疗药物的疗效及可能存在的不良反应，助力心血管疾病的早期筛查和及时干预，推动个体化精准用药，提高疗效。此外，华大基因致力于推动癌症早筛服务，进一步打造“筛+诊+保+治+康”的生态闭环，助力我国肿瘤的早筛查、早诊断、早治疗，全面探索一站式健康管理服务，实现全民健康。

3.行业展望

AI 赋能精准医疗具有无可争辩的优势和潜力。基因解读不仅可以识别突变，更能生成个性化健康管理方案，最终实现“从基因型到健康行为”和“从基因组档案到终身健康管理”的美好愿景。

7.5.2 典型案例：联影智能

上海联影智能科技股份有限公司（以下简称：联影智能）成立于 2017 年，是联影集团旗下的子公司，承接集团在医疗数字化和智能化方向的技术创新与商业落地布局。

作为联影集团旗下专注于人工智能与数字技术的前沿板块，联影智能以“成为世界级数智医疗创新引领者”为愿景，聚焦元宇宙、大模型和智能体等元技术的突破与应用，构建全维度赋能体系：赋能临床，为医生打造全栈、全谱智能产品及解决方案，助力诊疗效率与精准度提升；赋能设备，为医学影像设备注入全流程智能新技术，推动设备性能迭代；赋能科研，搭建医疗领域与先进科技的跨界桥梁，加速科研成果转化；赋能管理，为医院数智化转型提供高价值创新支持，优化运营效能；赋能患者，以普惠医疗为目标，为患者带来颠覆性的可感、可知、可及的就医服务体验。

AI 赋能场景：数智孪生医院——下一代医疗生态构想

行业趋势

通用大模型直接应用于医疗领域会存在一些问题。在医疗场景一旦产生幻觉，其带来的危害是巨大的。此外，临床一线面对的是复杂、不断变化的病情，涉及高度专业且庞杂的医学知识体系，仅仅依靠文本大模型的能力是不够的，医疗AI必须突破单模态、单任务模型的局限，具备理解影像、文本、语音等多种信息形式的能力，通过统一建模与认知融合，才能真正应对复杂的临床需求。

解决方案

医学领域的颠覆式突破与技术革命是同频共振的，在已经经历了「能量革命」（工业化医学）和「信息革命」（数字化医学）后，现在人工智能技术驱动的「认知革命」（共生医学）正在开启。AI 不再只是工具，而是与医疗系统共同演化、相互赋能的共生力量。

如果将智能体、大模型等元技术全面融至整个医疗系统，将催生怎样的行业变革？对此，联影智能给出前瞻性答案：下一代医疗生态构想——数智孪生医院。

这一构想并非局限于医院单点式的智能化升级，而是全院级、全场景的革命性智能化重构。AI 将原生融入从底层数据源、核心模型能力到上层应用场景的全链路，衍生出适配不同场景的医疗智能体，并通过群体智能协作机制，为预防、筛查、诊断、治疗、康复等全医疗环节深度赋能，构建一个开放共享、协同高效的未来医疗生态。

基于这一挑战，联影「元智」医疗大模型孕育而生，由文本、影像、视觉、语音、混合五个大模型组成，构成面向医疗全场景的认知中枢。在这一体系之上诞生了多款智能体，协同完成对临床任务的感知、决策与执行。

融合「元智」影像、语音与文本大模型能力，通过「看、听、想」的多模态协同，实现从医学图像感知到诊断语言生成的完整闭环——「元智」医疗影像大模型能够精准完成医学图像的自动识别与分析；同时，结合「元智」语音大模型，实现医疗场景中自然语言的听觉理解；最终通过「元智」文本大模型的推理能力，将图像和语音信息转化为结构化医学语言，自动生成诊断报告。

其中，联影「元智」医疗影像大模型，凝聚了联影智能多年深耕医疗行业的技术积淀，在真实临床场景中持续打磨优化。依托数千万级训练样本和数十万例高质量标注数据，已能够支持十余种医学影像模态，覆盖超过 300 项典型分割任务，在多个关键分割任务中准确率超过 95%。

1. 医护端

（1）放射智能体

放射智能体通过大模型的能力对影像分析实现了持续升级，通过文本语言监督图像理解，实现「0 标注」条件下对多种疾病的精准检测——在与复旦大学附属中山医院合作的胸部 CT 一扫多查场景中，大模型可以完成 73 种胸部异常检测。借助图像、文本与语音的多模态协同处理能力，智能体显著提升了影像解读与报告撰写的效率与准确性。

（2）核医学智能体

基于 PET/CT 等多模态数据训练，可实现多器官代谢状态的自动分析与可视化标注，迅速生成专业的诊断建议。这一能力不仅大幅缩短了核医学报告的撰写时间，也提升了该技术在基层医疗机构的可及性，为推动医疗资源下沉提供了有力支持。

（3）电子病历智能体

可自动生成多类医疗文书，如入院记录、病程记录、出院小结等。以中山医院为例，该智能体已在 11 个科室落地使用，累计调用超 6000 次，文书采纳率高达 93%。医生平均书写时间由原本的 20 分钟缩短至 5 分钟，显著提升了文书处理效率与质量。

2. 患者端

基于「元智」医疗大模型，联影智能推出多款面向患者的智能体，让患者在诊前、诊中、诊后的每一个节点，都能获得及时、精准与个性化的帮助。

（1）客服智能体

在「元智」文本与语音大模型的基础上，具备了医学语义理解与自然语言交互能力，能够 7×24 小时回应患者对疾病、检查流程等多方面的常见问题。相较传统问答系统，该智能体不仅拥有上下文记忆与多轮对话能力，还具备知识库自进化能力，可通过对真实对话数据的持续学习，自动发现知识盲区、补全问答内容，不断提升回答的准确性。

（2）用药指导智能体

在诊后的用药阶段，为患者提供个性化的用药建议。智能体可识别患者的病历、检查结果与用药信息，结合海量医学知识图谱，生成通俗易懂的用药说明，涵盖剂量、时间、注意事项、不良反应等多个要素，同时还可以主动提醒用药，不仅提升了患者用药的依从性，也在一定程度上降低了用药风险。

（3）随访智能体

能够为患者提供长期的健康管理服务。智能体可精准判断对话意图、结构化提取关键信息、精准匹配知识库问答，并通过自然流畅的对话与患者交流。

3.管理端

在联影「元智」医疗大模型的加持下，院管智能体覆盖质控和运营管理两大场景，构建起「全院级数据驾驶舱」。

（1）质控管理智能体

围绕医疗质量核心指标，提供贯穿诊疗全流程的标准化质量管理支持。智能体不仅能在第一时间识别异常、触发预警，也可对历史数据进行系统回顾与趋势分析，全面提升医疗服务的安全性与规范性。

（2）运营管理智能体

聚焦设备全生命周期管理，支持对设备使用率、运维状态、科室资源配置等关键指标的可视化监控与效益分析。它不仅帮助管理者及时识别运营风险，还可支持运营策略优化和资源动态调整，实现精细化管理。

推广价值

联影智能医疗 AI 解决方案已在全国超过 4000 家医疗机构落地应用，广泛赋能基层诊疗，推动医疗服务高质量发展。联影智能已有 18 款 AI 应用通过 NMPA 三类认证、23 款通过 NMPA 二类认证、15 款获得 FDA 认证、31 款获得 CE 认证（AI 应用获 CE 认证总数全球第一）。

医疗 AI 正在完成从工具到伙伴的转变。医疗大模型不仅提升了效率，更重新定义了人机协作模式——AI 不再只是为某一环节赋能，它从一个被动的工具，进化为一位主动理解医生的语言、感知患者的需求、协助管理者高效运营的伙伴，实现真正意义上的深度协同。

对于患者而言，AI 正在成为就医过程中的陪伴者，让复杂医学更可理解、服务体验更有温度。从咨询解答、用药指导到健康随访，医疗智能体正深度融入患者的就医旅程，让患者在面对疾病时拥有更多的理解力、掌控力与安全感。医疗 AI 正在成为患者值得信赖的健康陪伴者，为患者带来主动、温暖、个性化的就医体验。

从影像分析、诊断报告生成，到电子病历书写，医疗智能体深度融入日常工作流程，承担起信息处理、知识调用、临床辅助等任务，让医护有更多时间和精力，专注于关键决策和对患者的人文关怀。

对于医院管理者而言，AI 正在成为提升运营效率与决策质量的关键引擎。从质控到运营，AI 正在推动医院管理模式从被动响应走向主动干预，让数据转化为洞察，让管理更主动、更精细，赋能管理者以更全面的视角做出科学决策，推动医院从经验驱动向数据驱动加速转型。

7.5.3 典型案例：复星健康

上海复星健康科技（集团）有限公司（以下简称：复星健康）成立于 2010 年，是上海复星医药（集团）股份有限公司的控股子公司，历经十多年的发展，已形成以大湾区为核心、综合与专科医院相结合、线上线下一体化的医疗服务平台。

AI 赋能场景：星立方智能健康管理系统

行业趋势

当前，人口老龄化的问题日益凸显，我国慢性病的患病率和疾病负担逐年加重。近年来数据显示，慢性病导致的死亡人数已占总死亡人数的 86.6%，疾病负担约占总疾病负担的 70%。这一现状，促使健康观念由关注疾病的治疗转变为健康生活方式的培养和健康风险的管理。

解决方案

复星健康作为全国领先的医疗健康科技集团，积极响应国家“‘人工智能+’行动”的号召，利用自研生成式引擎，为用户的全生命周期健康管理提供 AI 驱动的设计解决方案——**星立方智能健康管理系统**。该方案通过一套健康管理中台+三个 AI agent 联动，实现自动化、智能化的主动健康管理模式。

MasterAgent 是健康计划的调度中心，负责动态协调分配和调度各项任务，以及各智能体间交流；TaskAgent 负责根据诊断、体征、检查数据等生态生成个性化的健康管理计划；KnowledgeAgent 负责从医学指南、模板库和病例库中提取核心信息，为生成健康管理提供知识基础。

在方案生成环节，系统根据患者的主诊断、辅诊断、病史、体征、化验数据等自动生成个性化的健康管理计划，经医生复核后可自动分配给患者，系统可以根据患者健康管理计划完成主动患者管理，患者可根据个性化的健康管理计划实现对疾病的自我管理；为用户制定个性化的健康管理方案，包括饮食建议、运动计划、复诊提醒、用药提醒等，同时推送健康知识和疾病预防信息，引导用户进行疾病管理和养成健康的生活方式；全病程管理通过整合预防、诊断、治疗、康复及随访等医疗环节，构建系统化、连续性的医疗服务体系，其核心价值在于以患者为中心优化资源分配，提升诊疗精准度与

效率，降低并发症风险及重复就医成本，同时通过动态监测和个性化干预提高患者生活质量。

推广价值

在管理方式上，传统的管理模式是患者被动接受治疗，自我管理依赖纸质手册或电话提醒，依从性较低；智能化全病程管理通过“数据驱动+人机协同”，支持患者与医生实时沟通，实现了从被动响应到主动干预。从效率上，智能的健康管理系统大大提高医护人员管理患者的工作效率，使医护人员在同样的时间内能够为更多的患者提供完善的健康管理服务；同时能够根据患者的动态指标数据快速生成个性化的健康管理计划，提高了患者的服务质量。从效果上，基于大数据和 AI 模型能够根据患者的数据自动生成完善的、个性化的健康管理方案，弥补了医生针对单病种过度依赖过往经验的不足，使整个健康管理方案更加合理，从而进一步提高健康管理的质量，提高用户的满意度。

复星健康构建的全生命周期健康管理系统作为 AI 赋能健康产业的典型代表，涵盖个体从出生到生命终点的全方位健康管理模式，通过预防、治疗和康复等手段，全面提升人民的健康水平。

7.5.4 典型案例：强脑科技

强脑科技（BrainCo）创立于 2015 年，是首家入选哈佛大学创新实验室（Harvard Innovation Lab）的中国团队，也是全球非侵入式脑机接口技术领军企业。公司主打安全、低成本、易普及的体表信号采集路线，已实现医疗康复、健康管理、教育科技、工业自动化、机器人协作等场景的规模化落地。

AI 赋能场景：融合脑机接口技术和人工智能算法的智能产品

行业趋势

在过去半个世纪中，脑机接口（Brain-Computer Interface, BCI）从最初的科学幻想和实验探索，已发展为跨学科前沿领域。如今的应用范围也从单一的交流与控制扩展至更为广泛的通用任务。随着神经科学、工程学、计算机科学与心理学的跨学科合作，脑机接口已成为一条可同时服务于残障人群和健康个体的增强型通信途径，并在康复医学、人机增强和沉浸式娱乐等领域展现出巨大潜力。

解决方案

强脑科技的技术根基源于其全栈自研的核心能力，涵盖超级干电极传感器、NeuroOS 操作系统、AI 意图解码算法三大核心模块，形成“信号采集－解码分析－动作执行－触觉反馈”的完整闭环，既保证了控制精度，又大幅降低了使用门槛，这也是其与传统假肢、进口高端仿生手形成差异化竞争的关键。

智能仿生手：这是一款脑机接口技术与人工智能算法高度整合的智能产品，主打“非侵入式、高精度、高适配、普惠型”。采用“体表神经/肌电信号采集+AI 实时解码”的技术路线，无需手术，仅通过佩戴式传感器，检测佩戴者的神经电和肌肉电信号，识别佩戴者的运动意图，并将运动意图转化为智能仿生手的动作，从而做到灵巧智能，手随心动，兼顾安全性与实用性，适配绝大多数上肢截肢人群。BrainCo Revo2 以 383 克的轻量化机身和毫秒级触觉反馈，让机器人从单纯的抓取迈向对物理世界的真

实理解，为通用机器人的落地提供核心感知能力。这让它不再局限于一个执行机构，而是升级为一个感知器官。智能仿生手的核心突破，是将脑机接口技术从实验室落地为可量产、普惠化的康复器械，打破了进口产品垄断，让“意念控手”从科幻场景走进现实，成为肢残人士重获独立生活能力的核心支撑，也是中国脑机接口商业化落地的标杆产品。

智能仿生腿：智能仿生腿是融合脑机接口技术和人工智能算法的新型智能下肢，在穿戴者行走时可根据环境与肌肉的情况进行实时步态调整，实现高仿生体验。轻凌 M3 重点突破“多场景适配”与“步态自然度”，荣获行业多项创新奖项，可满足用户从日常起居到运动健身的多元化需求。除了轻量化设计，它的运动模式从 2 种升级至 5 种，可轻松实现慢走、快走、小跑、交替上下楼梯、跨越障碍等接近健全人的动作，步伐速度可在 2 千—5 千米/小时之间自由切换；搭载多传感器融合技术，集成陀螺仪、霍尔传感器、温度传感器等，采集频率高达 1000Hz，可实时捕捉地面变化与用户运动意图，通过算法处理后控制液压系统，动态调节膝关节阻尼，下坡、下楼梯时可放心将重心放在假肢侧，不会失去支撑，彻底解决传统假肢“打软腿、易跪倒”的痛点；膝关节最大屈曲角度达 128°，回弹速度低于 0.6s，控制精度达 1°，步态自然美观，多数用户穿戴后，他人难以察觉其佩戴假肢。

推广价值

医疗普惠：非侵入式技术让脑机接口从“科幻”走向大众，覆盖康复、睡眠、教育等民生场景。打破进口垄断，仿生手价格降至欧美同类产品的 1/7 至 1/10，且实现了“即装即用”，让千万残疾人用得起智能假肢；**产业标杆：**构建“技术－产品－量产－服务”全链条，带动中国脑机接口产业链崛起；**公益使命：**强脑科技始终坚持“普惠民生”的理念，已通过公益项目为数千名肢残人士免费提供智能仿生手，覆盖全国多省；同时计划未来 5—10 年，帮助 100 万肢体残障人士重获行动自由，助力其回归社会、参与工作，减少家庭与社会的负担。

7.6 AI 赋能未来空间产业

7.6.1 引言

在数字科技革命与产业变革深度融合的纵深阶段，未来空间已成为全球科技竞争的核心战略赛道，其内涵突破了传统地理空间的边界，从平面延伸至立体，从陆地拓展至深海、低空、太空等人类探索与利用的新兴领域，成为推动经济社会高质量发展、拓展人类生存与发展疆域的关键载体。人工智能作为新一代信息技术的核心引擎，历经从单一工具赋能到核心动能驱动的演进，在 2025 年空间智能大模型技术突破的基础上，进一步实现了与未来空间产业的深度耦合、体系化融合，凭借多源信息融合、自主智能决策、动态算力调度、复杂环境适配、具身智能交互等核心技术能力，从底层重构未来空间的技术体系、应用模式与产业生态，破解各领域面临的环境复杂性、操作精准性、资源约束性、响应实时性等共性难题。如今，人工智能正推动人类对深海的探索、对低空的利用、对太空的拓展从“经验驱动”向“数据与模型驱动”全面转型，从“单点应用”向“全链条赋能”深度升级，从“被动执行”向“自主智能”

高阶演进，成为未来空间产业向智能化、规模化、可持续化发展的核心支撑，也为我国在全球未来空间科技竞争中筑牢技术与产业基础。

从应用背景来看，当前全球科技竞争已聚焦未来空间这一战略高地，各国纷纷加码布局，将人工智能与未来空间融合作为抢占科技主动权的核心抓手，美国、欧盟、日本等相继出台相关战略规划，推动太空在轨计算、深海智能探测、低空自主管控等领域的技术突破与产业落地。我国锚定 2035 年建成科技强国的战略目标，在“十五五”开局之年明确将“人工智能+未来空间”纳入未来产业发展重点，政策与市场协同发力，为二者融合发展提供了坚实支撑。从技术演进背景而言，空间智能已成为人工智能产业继大语言模型后的第二增长曲线，突破了传统 AI 局限于二维数据处理的瓶颈，实现了对三维/四维物理空间的感知、建模、理解与交互，为 AI 赋能未来空间奠定了核心技术基础。从应用概述来看，AI 在未来空间领域的应用已形成全流程赋能格局，贯穿“感知-建模-决策-执行”全链条，依托多源异构感知技术、三维空间建模技术、空间大模型等核心能力，打破了极端空间环境对人类探索的限制，推动未来空间从“科研高地”向“产业新赛道”转型，成为培育新质生产力、推动高质量发展的关键变量。

未来空间产业聚焦深海、低空、太空等领域，以高端装备为核心载体，以科学探测、资源开发、运输装备、民生服务为关键环节，具有高成本、高技术、高风险的显著特征，其发展核心在于推动空间领域与装备、能源、材料、通信等领域的交叉融合，实现产业发展的绿色化、集成化与智能化。过去，人类对空间的探索与利用多依赖物理扩张与经验积累，受限于数据处理能力、环境感知能力与智能决策能力，空间资源的开发效率偏低、利用边界较窄，难以应对复杂空间环境的动态变化。而人工智能技术的发展，尤其是空间智能大模型的出现，打破了传统人工智能局限于二维数据学习的瓶颈，通过整合地理信息系统、遥感技术、传感器网络、多模态视觉等多源空间数据，实现了全方位、多维度的空间信息获取与分析；依托跨领域知识融合与海量数据高效处理能力，打通了计算机、数学、遥感、气象学、地质学、海洋学等学科的技术壁垒；凭借智能推理与预测功能，实现了对空间发展规律的精准把握与动态推演。这一技术突破，让人工智能不再仅仅是助力未来空间产业发展的辅助工具，更是驱动其实现范式革命的核心动能，通过解构时空数据、重塑决策逻辑、突破物理边界，重新定义了空间价值的创造范式，推动人类对空间的利用从“被动承载”转向“主动进化”，在国土治理、星际探索、民生服务等维度构建起人类文明发展的“第二增长曲线”。

当前，人工智能赋能未来空间已形成深海、低空、太空三大核心发展领域，各领域依托人工智能技术实现了从技术研发到场景落地、从业态创新到产业升级的跨越式发展，释放出万亿级市场机遇的同时，更推动了一场重构人类生存维度的认知革命。在深海领域，极端的物理环境对水下装备的智能控制、防护与自主作业能力提出极高要求，人工智能技术实现了水下推进器的多参数融合控制、智能避障、精准防腐与磁耦合智能调控，让水下装备具备了适应复杂非线性环境的自主能力，推动深海开发从近岸浅海走向万米深渊，从人工作业走向智能无人作业；在低空领域，人工智能成为低空经济全

场景赋能的核心抓手，通过深度 Q 网络等算法实现无人机物流路径优化、多机协同避障，结合多模态感知技术打造高冗余安全架构，推动低空经济与物流、巡检、消防、农业、电力等领域深度融合，解锁了城市立体交通与乡村振兴的全新应用场景，让低空成为破解地面资源约束、提升民生服务效率的重要新载体；在太空领域，人工智能的赋能从卫星数据的地面处理升级为太空在轨智能计算，通过星载 AI 计算单元、激光星间链路、具身智能技术的突破，构建起太空算力基础设施，实现了卫星数据的在轨实时处理、灾害的分钟级监测响应与太空装备的自主操作探索，推动航天产业从“信息采集”向“智能服务”高阶演进，从单一的卫星应用向太空经济体系化发展转型。

从实践层面来看，人工智能赋能未来空间已形成从微观装备智能升级，到中观场景高效落地，再到宏观基础设施体系化构建的完整实践路径，昊野科技、御风未来、之江实验室三个分别对应深海、低空、太空核心领域的典型案例，为行业发展提供了可复制、可推广的技术方案与产业经验。昊野科技聚焦深海领域，以人工智能为核心打造水下推进器的智能控制、避障、防腐等技术体系，其研发的多参数融合控制、智能主动防护等技术成功应用于冬奥会水下火炬传递、“悟空”号全海深 AUV 等重大场景，实现了极端水下环境中装备的高精度、高可靠性自主作业，成为深海智能装备产业化的标杆；御风未来深耕低空领域，将人工智能融入电动垂直起降飞行器（eVTOL）研发与无人机作业体系，依托 AI 路径规划、多模态感知、高冗余飞控架构等技术，解锁了山区电网巡检、偏远地区物流配送等低空场景的智能化应用新范式，让低空经济从概念走向实际民生服务；之江实验室立足太空战略高地，以“三体计算星座”为核心构建全球首个大规模太空在轨智能计算网络，通过人工智能实现太空算力的在轨部署、星地数据的智能处理与具身智能卫星的研发，重构了人类利用太空资源的模式，成为太空智能基础设施建设的引领者。这三个案例从技术研发、场景应用、基础设施搭建等不同维度，勾勒出人工智能赋能未来空间的实践图谱，其创新成果不仅在各领域实现了关键技术突破与商业化落地，更充分验证了人工智能与未来空间产业深度融合的可行性、价值性与落地性。

本部分立足人工智能与未来空间产业的发展现状，通过对昊野科技、御风未来、之江实验室三个典型案例的剖析，系统梳理人工智能赋能深海、低空、太空三大核心领域的核心技术路径、创新应用模式与产业赋能价值，提炼各领域从技术攻关到商业化落地的关键经验与实践范式。这些案例的创新探索，既展现了人工智能在破解未来空间产业发展痛点中的核心作用，也为未来空间产业的智能化升级、核心技术的自主创新、应用场景的深度拓展提供了现实且具体的参考范本。人工智能已经成为推动人类探索空间、利用空间、发展空间的核心力量，助力我国未来空间产业迈向智能化、规模化、可持续化的高质量发展新阶段，在未来空间的全新赛道上构建起中国特色的技术体系与产业生态，为全球未来空间产业的智能化发展提供中国方案与实践参考。

7.6.2 典型案例：昊野科技

昊野科技有限公司（以下简称：昊野科技）拥有两家专注于开发应用水下推进器和新能源船舶电

动系统的子公司，前者为天津昊野科技有限公司，后者为天津飞舶科技有限公司。

AI 赋能场景 1：水下推进器智能避障系统

行业痛点

水下航行器在执行测绘、巡检、监测等任务时，常常面临水下复杂环境的挑战。而水下航行器使用的传统避障系统大多使用如声呐等单一传感器。在浑浊、强水流水域或遇到特殊材质障碍物时，往往存在感知信息不全、误报漏报率高、规划路径脱离实际等问题，严重影响水下航行器的安全与可靠性。

解决方案

为解决上述问题，昊野科技提出了基于多源信息融合与智能决策的避障系统。该系统通过集成光学、声学及水流等多种传感器提供的数据搭建实时环境感知层，同时利用障碍物拟合模块建立精确的障碍分布模型，在此基础上生成路径与优化模块结合的推进器动力学模型，并以响应时间、避障效能及能量消耗为多目标优化函数计算并选择综合最优的避障路径，实现了感知、建模到决策的闭环智能控制。

推广价值

该系统的创新在于将“多维感知”、“精准建模”与“多目标优化决策”相结合，提升了水下航行器在水下动态复杂环境中的安全性与作业效率。在 2022 年的北京冬奥会水下火炬传递中，昊野科技提供的方案成功帮助两个水下机器人实现了厘米级精度的悬停对接，证明了该方案的可靠性与控制精度。

AI 赋能场景 2：基于多参数融合的水下推进器控制方法

行业痛点

在多变的水下环境中，流速、压力、地形等因素耦合在一起。传统的水下航行器控制方法依赖于单一参数或固定经验，难以全面准确地模拟真实流场与航行器相互作用，导致水下航行器的控制响应滞后、输出不稳定，甚至引发碰撞风险，限制了推进器在精细作业场景下的性能。

解决方案

昊野科技开发了基于多参数融合与人工智能的控制方案。该方案通过三维流场实时感知模型获取水流速度、压力、强度及地形坐标等多模态数据，并采用分散滤波与数据融合技术提炼高质量信息，通过集成时序注意力机制的多模态神经网络模型处理融合数据，并运用双层强化学习策略实时生成自适应控制指令，确保动力输出既精准又稳定。

推广价值

此方案标志着控制策略已经从“经验驱动”转变为“数据与模型驱动”模式，使水下航行器具备了理解并适应复杂非线性环境的能力，成为实现高水平自主智能的关键底层技术。该方案已成功应用于哈尔滨工程大学“悟空”号全海深 AUV，为其在万米深渊实现高精度定深航行提供了核心动力保障，助力“悟

空”号多次刷新深潜世界纪录。

7.6.3 典型案例：御风未来

御风未来是一家空中出行创新企业，致力于打造高安全、高性价比、绿色智能的 eVTOL（电动垂直起降）飞行器，以“人人可享有的空中出行”为愿景，为未来城市立体交通空中出行提供解决方案。

公司专注于高性能、高安全的电动、混动垂直起降固定翼飞行器研发，具备中小型长航时侦察型无人机、中型运载型无人机、中型应急通信无人机和吨级载货/载人电动垂直起降飞行器等机型的研制和商业化交付的经验，具备复杂构型飞行器的快速研制和产品化的核心能力。

御风未来产品核心控制软件均为自主开发，包括高冗余度自主飞行控制系统、GNSS+视觉匹配定位双模态惯性参考系统，自研飞控和惯导的整套工具链，软件均可适配国产操作系统，且无人机平台预留北斗军码接口。

AI 赋能场景 1：电网输电线路检查

行业痛点

中国电力建设现在几乎覆盖了全中国所有地区。中国山区面积占全国 2/3 的土地面积，为满足社会需求，国家电网正在进行超高压大容量电力线路扩建，线路将穿越各种复杂地形，对电力设施安全性、可靠性及其运行状态的实时性监测要求越来越高。如何提高电力线路检测的精度和效率，是困扰电力行业的一大难题。许多山区海拔高、植被密、地理环境错综复杂，传统人工巡查需翻山越岭，使用普通仪器或肉眼来巡查设施，现场勘查难度大、劳动强度大，难以获取有效真实地形，尤其是遇到电网紧急故障和异常气候时，难以快速精准完成输电线路巡查，不能完全适应现代化电网建设与发展的需求。

解决方案

为实现电力巡检的高精度、高效率、高安全性，可以利用激光测量系统与专业飞行平台构建高空作业优势，通过系统化的作业流程与技术处理，全面采集输电线路相关数据，为电力线路运维提供精准支撑。

在航线规划与飞行执行阶段，根据预设的巡检范围、设备分布及地形条件，精准规划飞行航线，合理设定飞行高度与巡航速度，平衡数据采集精度与作业效率。将规划完成的航线参数导入地面控制站，完成与无人机的适配调试，确保无人机可按照预定航线实现全自主飞行作业。飞行过程中，无人机搭载激光雷达设备持续开展数据采集，同时实时监测数据传输状态、设备运行参数及环境变化，及时排查异常情况，保障采集数据的完整性与准确性。

在数据采集与处理环节，通过激光测量系统高效获取输电线路走廊内的全方位数据，涵盖地形地貌、植被建筑等周边地物，以及杆塔、导线、绝缘子、金具等电力核心设备的高精度激光点云和正射影像数据。采集完成后，运用雷达一体化后处理软件对原始数据进行专业化处理，经过数据清洗、拼

接、校准等流程，生成成果级点云数据，并构建测区数字高程模型（DEM）、数字表面模型（DSM）及输电线路沿线三维地形模型，为后续分析应用奠定基础。

在成果分析与应用阶段，基于构建的三维模型开展多维度专业分析：精准识别塔基位置、分析周边环境特征及地形高差等关键信息，实现电力线路与周边地物空间距离的高精度实时测量；同时，围绕电力运维核心需求，完成线路工况缺陷检测、净空区域排查、线路交叉跨越分析、杆塔精确定位、导线弧垂测算及不同工况模拟等工作。系统支持自定义配置行业规程参数，自动生成标准化分析报告，可根据实际运维需求灵活调取应用，为电力线路的安全运行、检修维护及优化升级提供全面、精准的技术支撑。

推广价值

无人机测量技术凭借非接触式作业与高效能执行的核心优势，可在短时间内完成输电线路的测量与规划工作，既能显著提升作业效率，又能从根本上保障人员作业安全。该技术具备高点频、高线频、高精度、高集成度及全国产化的显著特征，可应用于以下任务：

- （1）故障巡线：能够快速、精准定位电网及相关设备的故障点位，为抢修工作提供精准数据支撑；
- （2）常规巡线：可全面替代传统人工巡检模式，有效提升巡线效率、降低巡线成本，同时减轻作业人员劳动强度，大幅降低野外作业安全风险；
- （3）数据采集与分析：可高效获取线路走廊的高精度三维空间信息及高分辨率真彩色影像数据，实现线路交叉跨越高度、树木及建筑物高度、线路与周边地物空间距离等关键参数的高精度实时测量与分析。

AI 赋能场景 2：物流配送

行业痛点

在山区运输场景中，受复杂地形条件限制，因山体形态不规则、地势陡峭且沟壑纵横，导致交通基础设施建设难度极大，修路成本居高不下，多数偏远区域难以形成规模化路网覆盖。目前，部分山区仍依赖人力、畜力完成物资搬运，不仅劳动强度极大，作业持续性受限，更难以满足应急物资、生鲜产品等对运输时效的刚性需求。同时，山区天气变幻莫测，雨雪、冰雹等极端天气频发，易造成路面泥泞、结冰或滑坡等风险，不仅进一步降低运输效率，更给运输作业人员的人身安全带来严重威胁。因此，传统运输模式在山区运输可谓是“难于上青天”。

解决方案

现在人工智能技术为无人机运输注入了“智慧大脑”，首要突破是运输效率的指数级提升。比如在四川甘孜泸定等山区，传统人力运输电力绝缘子等物资需攀爬 6 小时，而搭载 AI 智能控制系统的运载无人机，15 分钟即可完成跨峡谷运输；当地邮政借助该无人机，将原本一两小时的邮路压缩至十几分钟。效率得以提升的原因，是运用 AI 进行路径规划，让无人机自动规避山间气流、复杂地形。在设定装货

点与卸货点后便能生成最优航线，实现自主巡航、一键返航，无需飞手持续手动操控，大幅缩短运输耗时。同时，AI 驱动的司运云平台可实时回传飞行数据，实现运输任务的高效调度与管理，让单架无人机单日作业能力远超传统运输方式，如湖北秭归脐橙产区的无人机，助力从业者单日作业收入提升 30%。

其次，人工智能赋予无人机在复杂环境中的适应能力与作业稳定性，打破了传统运输场景的限制。依托 AI 融合激光、毫米波、视觉三重感知技术，搭配机载陀螺仪及备份飞控系统，构建高冗余安全架构，即便单桨故障或遭遇强风、低温等恶劣环境，仍能平稳飞行。在海拔数千米的川西高原，无人机可精准运送检修物资；在云南热带雨林、安徽黄山等地形复杂区域，也能稳定完成物资运输任务，彻底改变了偏远山区“望山兴叹”的运输困境。此外，AI 赋能的自动消摆、仿地飞行、货物自动投放等功能，让无人机在跨峡谷、越山岭等作业中，实现货物的平稳装卸，避免物资损耗，解决了传统运输中“最后一公里”成本高、难度大的问题。

推广价值

破解以前的地形局限性，无需大规模基建即可覆盖山区、边境、海岛等传统运输盲区，在应急灾害中快速突破道路中断阻碍，填补“最后一公里”运输空白；提质增效降本，依托 AI 技术实现自主巡航、精准投送，大幅缩短运输时间，同时规避人员安全风险，省去道路建设维护成本，具备经济可行性；赋能多领域发展，广泛应用于军事巡逻、民生物流、产业作业、应急救援等场景，替代高风险低效率的人力劳动。

7.6.4 典型案例：之江实验室

之江实验室成立于 2017 年，是聚焦智能科学与技术领域的浙江省新型研发机构，在智能计算、人工智能、未来空间等前沿领域开展有组织的科研攻关。之江实验室将“未来空间”锁定为核心战略高地，其核心创新在于跳出传统地面算力与卫星应用的思维框架，将 AI 的智能计算、自主决策、协同感知能力全面延伸至太空，从根本上重构人类利用太空资源、开展空间活动的模式，破解传统航天领域算力供给不足、数据处理延迟、空间操作被动等行业共性难题。

AI 赋能场景 1：“星间计算”赋能实时全球灾害监测与响应

行业痛点

在应对地震、洪涝、山火等突发自然灾害时，时间就是生命，救援的“黄金 72 小时”对灾情信息的获取效率提出极高要求。但传统卫星灾害监测流程存在明显短板：卫星过顶拍摄后，原始数据需要排队等待下行至数量有限的地面站，再经远距离传输至数据处理中心，由专业专家开展人工或半自动解译，最终形成完整灾情报告，整个流程动辄耗时数小时。在通信设施受损、处于通信中断状态的“孤岛”灾区，这一问题更为突出，灾情信息传递几乎陷入停滞，大量宝贵的救援时间消耗在信息获取环节。此外，卫星下传的原始数据中包含大量无关的背景信息，严重挤占本已紧张的星地通信带宽，进一步

降低了信息传输效率。

解决方案

依托“三体计算星座”提供的在轨算力支撑，之江实验室打造了星载灾害智能感知及分析体系，将灾情信息提取与分析的核心环节从地面全面迁移至太空，实现灾情信息的在轨智能处理。

该体系具备实时响应能力，在卫星载荷开机成像的同时，星载 AI 处理单元将同步启动工作；体系内置经过海量灾情数据训练的视觉分析模型，可对刚获取的卫星影像进行实时扫描，精准自动识别洪水淹没范围、山体滑坡体量、火灾燃烧烈度、建筑物倒塌区域等灾情关键要素。同时，多颗卫星可通过高速星间激光链路，实时交换各自的灾情分析结果，开展多源数据交叉验证与信息融合，有效提升灾情识别的准确率与可靠性。

最为关键的是，该体系实现了数据的“轻量化”回传，卫星无需再将数以 TB 计的原始影像全部传回地面，仅需将结构化的灾情简报（包含地理位置、灾害类型、影响面积、严重等级等核心信息）和极少量经过压缩的典型特征图像，传输至地面应急指挥中心。这一转变将传统的“数据下行－地面处理－分析上报”信息链条，简化为高效的“在轨感知－关键信息直达”模式，真正实现了灾情信息的“分钟级”直达。

推广价值

这套星载灾害智能感知及分析体系的落地推广，将深刻改变全球应急管理乃至公共治理的灾情响应模式，具备多重核心价值。

最直接的价值在于大幅提升应急响应的速度与精度。地面应急救援力量可在灾害发生后极短时间内，获取覆盖范围广、客观准确的全域灾情态势图，为救援人员与物资投放、受灾群众疏散路线规划、救援重点区域划定提供直接且精准的决策支持，最大程度减少灾害造成的生命财产损失。

从技术经济角度来看，该体系实现了星地通信资源的革命性优化。通过在轨的信息提炼与数据筛选，大幅节省星地通信下行带宽资源，有效提升通信链路的信息传输效率。这意味着同一套卫星通信系统，可服务更多用户、承载更多类型的任务，显著提升天基系统的整体利用效率与经济效益。

此外，该体系的技术能力可实现多场景快速拓展，从突发自然灾害监测，逐步延伸至生态环境动态监管（如非法排污、森林盗伐、冰川消融）、重大基础设施健康诊断（如桥梁、大坝、高铁线路的形变监测）、智慧城市运行的宏观态势感知等多个领域，为国土安全保障、城市精细化治理、人类社会可持续发展，打造一颗强大的天基“智慧之眼”。

AI 赋能场景 2：“具身智能卫星”——面向太空在轨服务与科学探索

行业痛点

随着全球航天产业的快速发展，近地轨道航天器的数量呈爆炸式增长，太空空间日益拥挤，轨道资源紧张问题愈发突出。当前，卫星一旦发生设备故障或在轨失效，几乎不具备在轨维修与修复的可

能性，最终只能成为昂贵的太空垃圾；大量失效卫星、火箭残骸形成的太空碎片云，不断积累并产生碰撞风险，严重威胁着所有在轨航天器资产的运行安全。现有针对太空碎片清除或故障卫星维修的构想，大多依赖发射特制的专用“服务航天器”，不仅研发与发射成本极高，且这类航天器的任务指向性单一，缺乏普适性与经济可行性。在更遥远的深空探索领域，如火星、月球等天体探测任务中，星地之间存在以分钟甚至小时计的通信延迟，地面遥控操作已完全不切实际，探测器必须具备更高的自主决策与操作能力，才能应对深空复杂且未知的环境挑战。

解决方案

面向太空操作自主化的未来发展需求，之江实验室正在推进前沿的“具身智能卫星”研发工作。所谓“具身智能”，即让人工智能拥有与现实物理世界直接交互的“身体”，实验室的研发目标是为卫星配备高性能 AI“大脑”与多自由度机械臂等执行机构，打造能看、能想、能动的太空智能机器人，实现卫星的在轨自主操作与服务。

该研发工作的核心挑战，在于让人工智能学会在微重力、强辐射、超真空的复杂太空环境中，开展高精度的精细操作。实验室为此制定了“地面仿真，天上执行”的核心解决路径：首先构建高保真的太空操作数字孪生系统，1:1 还原太空极端环境与操作场景，让智能体在虚拟仿真环境中通过强化学习，开展数以百万次计的机械臂抓取、航天器部件更换、燃料加注等实操训练，直至智能体形成成熟、稳健的太空操作策略，再将其模型算法上载至在轨卫星。

同时，具身智能卫星采用模块化、标准化的设计理念，具备多星协同工作能力。多颗具身智能卫星可形成太空中的“蜂群”，通过协同配合完成单颗卫星无法承担的大型复杂任务，例如合作捕获大型太空碎片、在轨组装大型空间结构、为故障卫星开展协同维修等。目前，实验室已将具身智能卫星纳入研发规划，正在推进技术验证星的研发工作，开展太空自主操作核心技术的在轨演示验证。

推广价值

具身智能卫星的研发与成功应用，将推动人类迈入全新的“太空 2.0”时代，其价值体现在在轨服务、深空探索、产业生态等多个层面，对航天产业发展具有里程碑意义。

在在轨服务与太空环境维护方面，具身智能卫星可实现故障卫星的在轨维修、设备升级与轨道提升，大幅延长昂贵航天器的服役寿命，从根本上改变航天器“一次性使用”的传统商业模式，显著提升太空经济的整体经济性。同时，其具备的主动太空碎片清除能力，可有效减少近地轨道的碎片数量，是维护各国太空资产长期运行安全、保障轨道资源可持续利用的必要手段。

在深空探索领域，具身智能卫星的核心技术是不可或缺的关键支撑。将自主操作与决策能力赋予深空探测器，可使其在几乎无地面干预的情况下，独立完成样本采集、探测设备部署、突发故障自主排除等复杂操作，突破星地通信延迟的限制，极大拓展人类探索太阳系的边界与深度。

更重要的是，该技术将催生一个全新的太空产业生态。围绕太空在轨制造、微重力太空实验、太

空碎片回收、卫星报废拆解、在轨燃料补给等太空服务，将诞生一系列全新的商业机会与产业形态，推动太空经济从以通信、导航、遥感为主的“信息层”，向包含在轨操作、太空服务的“操作层”高阶形态演进，真正释放太空的商业潜力与发展价值。

7.7 AI 赋能其他产业

人工智能技术呈现出极强的渗透力和覆盖面，正以多样化的形态融入千行百业的各个环节，不断催生产品和服务创新、提升运营效率、优化服务体验。本节将介绍 AI 赋能城市治理、智慧农业、文旅、绿色环保领域的典型企业案例。

7.7.1 典型案例：库萨科技

库萨科技专注于城市服务机器人领域，以具身智能与机器人技术重塑传统环卫作业模式，推动无人作业规模化落地。库萨科技拥有完整的具身机器人技术架构、强 AI 视觉主导技术方向，以及车规级产品开发流程与供应链支持，连续 3 年入选上海市经信委《上海市智能机器人标杆企业与应用场景推荐目录》。

AI 赋能场景：无人清扫机器人

行业痛点

随着城市运行精细化水平不断提升，传统环卫作业依赖人工、经验驱动的模式已难以满足现代化管理需求，行业智能化、标准化转型迫在眉睫。当前环卫行业面临多重突出痛点：一是作业安全风险高，人工清扫需长期暴露于高温、严寒、夜间及车流密集路段等高危环境，职业安全保障不足；二是质量与监管难以统一，不同区域、时段的作业效果参差不齐，缺乏有效的过程监控与标准化管理手段，服务质量稳定性差；三是技术应用碎片化，现有智能化设备多处于零散试点阶段，缺乏系统性解决方案与可持续运营模式，难以实现规模化复制与长效运行。

解决方案

在技术层面，库萨科技构建了完整的具身机器人技术架构，将感知、行动与认知深度整合，使机器人能够在复杂多变的城市开放环境中灵活、高效地自主作业。公司以强 AI 视觉为主导，融合多源异构传感器构建多维感知系统，使无人清扫机器人不仅能实现厘米级的高精度贴边清扫与自主避障，更能对动态变化的环境进行快速理解与精准决策。库萨的智能云控平台实现多机协同调度和远程运维管理，形成了从单一机器人智能到多机协同调度的系统化能力，确保整体解决方案兼具高可靠性、可扩展性与持续进化的潜力。

库萨科技自主研发的无人清扫机器人基于库萨自研封闭式操作系统，能够在开放道路场景中全面理解环境，精确执行清扫任务，实现厘米级贴边作业精度。无人清扫机器人采用车规化设计与选型，满足环卫作业对一致性、稳定性和耐久性的要求。依托电子线控底盘和高精度路径规划，机器人可灵活应对狭窄道路、转角及边角区域的清扫任务，在复杂城市场景中保持高覆盖率和高效率。

无人清扫机器人采用多源异构传感方案，融合非接触式与接触式传感器，并以视觉 AI 为核心，构建起面向城市开放道路的全局感知能力。非接触式传感器可构建多维感知数据矩阵，接触式传感器则通过力接触获取硬度信息，不仅有效保护作业机构，也为视觉系统提供监督反馈。通过融合算法整合来自不同模态的数据，无人清扫机器人得以形成高维度的环境感知画面，使每个像素点同时具备颜色、纹理与物理属性信息，显著提升了其在障碍物识别、路况判断与作业安全性方面的精度与鲁棒性，全面增强了机器人对复杂城市环境的适应能力。

作业中，机器人按预设任务自主运行，并通过智能云控平台实现多机协同。智能云控平台具备动态任务调度能力，集成 Multi-Agent 数据闭环系统，如同“指挥中枢”，借助大模型技术模拟人类团队协作逻辑，统筹管理机器人集群的任务执行、路径规划与资源分配，实现高效的协同作业。

推广价值

库萨科技构建的环卫作业新模式系统性改变了传统环卫工作流程。在该模式下，机器人高效承担基础性、重复性清扫任务，而原有环卫人员则转型负责设备监管、特殊区域深度清洁等更高价值工作，实现了人力资源的优化配置与能效提升；基于云控平台实时采集的作业数据，系统自动生成精细化质量评估报告，推动管理决策从经验依赖转向科学的数据驱动；同时，通过建立从任务智能下发、过程规范执行到效果精准评估的全链条标准化流程，确保了服务质量的一致性与可追溯性，为环卫作业的规模化、智能化运营提供了坚实保障。在华东区域某城市街道的典型应用中，机器人集群每日分 3 个班次沿预定路线自动清扫，作业精度达到 ± 2 厘米，实现街道清扫的 100%智能化覆盖，并将垃圾滞留时间显著缩短 70%。通过智能云控平台与机器人结合的创新模式，该项目较传统模式有效降低了近 50% 的运营成本，充分体现了智能化技术驱动下的环卫作业在效率、质量与经济效益方面的全面提升。

库萨科技的无人清扫机器人已在国内外超 30 座城市完成部署，在上海、浙江、江苏等 10 余个市、县（区）实现首台落地，获厦门、杭州、嘉兴等城市首批合规上路牌照，广泛应用于城市道路、公园、景区、园区等多种开放场景。通过提升环卫作业安全性、改善城市环境质量，不仅有效降低了从业人员的职业风险，也增强了市民的幸福感和获得感；同时，该案例为环卫行业智能化转型提供了可复制的技术路径与商业模式。此外，通过技术驱动产业升级，为从业人员开辟了新的职业发展路径，推动了传统环卫行业向智能化、标准化方向的全面转型，充分体现了科技创新在提升公共服务质量、促进行业可持续发展方面的综合价值。

7.7.2 典型案例：联适技术

上海联适导航技术股份有限公司（以下简称：联适技术），成立于 2015 年，以北斗卫星导航与智能控制为核心技术，以用户需求为牵引、以信息技术自主创新为导向，形成集软硬件产品研发、生产、销售与服务于一体的业务体系，面向智慧农业提供作物生产全流程技术装备及综合解决方案。作为国内首批实现农机智能驾驶国产化规模商业应用的高新技术企业，公司产品覆盖农业生产“耕、种、管、

收”全环节，是国内较早实现商业落地的无人化农场解决方案供应商。

AI 赋能场景 1：农机自动驾驶系统

行业痛点

在中国农村劳动力持续短缺、老龄化加速的背景下，行业痛点主要体现在三个方面：

第一，用工缺口推动机械化升级，但好机手更缺。农忙季节集中、跨区作业频繁，熟练机手难招难留、费用高且流动性大，一旦人员不到位就容易误农时，合作社与农服组织对少人化、可复制的作业能力需求迫切。

第二，劳动强度大导致效率与安全问题突出。耕整地、播种、施肥喷药等作业往往连续十几小时，长期盯线、频繁掉头和抢收抢种使疲劳驾驶普遍，容易走偏、压苗、漏喷，既影响效率也带来人身与设备安全风险。

第三，农事作业质量要求越来越高但人工难以稳定达标。高标准农田建设、规模化经营与托管服务要求作业直线度、行距一致性、重叠率和漏作率可控，并需要可追溯的作业数据作为结算与考核依据。传统人工受经验与状态波动影响大，难以长期保持一致的精度与质量，直接影响出苗率、用肥用药成本与最终产量稳定性。

解决方案

农机自动驾驶系统的构成如下：

高精度定位：北斗/GNSS+RTK，在农田作业中提供厘米级定位，对林果业遮挡场景叠加惯导实现短时稳姿稳航；智能控制终端：车载主机+显示交互，提供自动直行、AB 线、曲线、等高线、地头掉头、边界作业等功能；电控执行机构：电动方向盘、液压转向控制器、油门、档位、提升等接口，实现自动转向与联动控制；云端（面向农服）：设备管理、任务下发、作业轨迹与面积统计、质量指标与报表、数据留存与对账。

农机自动驾驶系统具备以下核心能力：

降低对熟练机手的依赖：把盯线控距变成系统能力，普通机手经过短期培训即可稳定作业；复杂地块适配：支持小地块边界、异形地块、地头自动掉头，对丘陵、林网遮挡等场景提供失锁保护与平滑续航策略；作业质量可控：控制重叠率、漏作率与行距一致性，减少重耕漏耕、漏播漏喷，稳定出苗与投入品利用率；运营可追溯：自动记录作业轨迹、时间、面积、速度等数据，便于托管结算、质量考核与跨区作业管理。

推广价值

农机自动驾驶系统的推广价值核心在于把依赖经验的农机作业升级为标准化、可复制、可监管的生产服务，对政府、产业与经营主体均具备明显外溢效应。

服务国家农业现代化与粮食安全：在耕、种、管、收等关键环节提升作业精度与一致性，减少漏

作与返工，稳定出苗率与投入品利用效率，有助于提升单产与综合产能，支撑稳粮保供与高标准农田建设目标。

应对农村劳动力短缺与机手老龄化：把盯线、对行、掉头等高强度技能工作交给系统，显著降低对熟练机手的依赖，缓解旺季用工紧缺与跨区作业人跟不上的矛盾，推动农业生产向少人化、专业化转型。

推动农服规模化与社会化服务升级：对合作社、农服公司而言，自动驾驶带来的不仅是单机效率提升，更是可复制的作业标准+可追溯的数据证据。这使得跨区作业调度、质量考核、按亩结算、风险追责更可控，促进农服组织做大做强。

促进降本增效与绿色生产：通过控制重叠率与漏作率，减少油耗、种肥药浪费与土壤压实，提升投入产出比，在精量播种、变量施肥喷药等方向具备良好协同基础，符合减肥减药、提质增效的绿色发展导向。

带动国产北斗与农机智能制造产业链：系统以北斗高精度定位、车规级控制与执行机构为核心，可拉动传感器、控制器、液压电控、整机接口标准化、售后服务网络等配套升级，推动关键技术国产化与规模化应用，形成产业集群效应。

数据要素价值与治理价值：作业轨迹、面积、效率与质量数据可沉淀为农业生产数字底座，为农机补贴精准化、托管服务监管、保险理赔、金融授信与农情监测提供可信依据，提升农业治理的精细化与透明度。

AI 赋能场景 2：无人化智慧农场解决方案

行业痛点

过去农场经营更多依赖经验和现场盯管：凭感觉定播期、凭习惯配方施肥、靠机手手感控行控量，管理结果高度依赖个人能力与当时状态。在规模化、托管化和高标准农田推进后，这套方式暴露出明显痛点：一是经验难复制，人员流动就导致质量波动，二是过程不可量化，出现减产或成本上升时难以追溯责任与原因，三是决策滞后，无法对墒情、苗情、病虫草害和作业效率做及时调整，增产增收目标难以稳定达成。

无人化智慧农场方案的核心诉求转向用数据管理生产、用模型指导决策。但现实痛点在于：数据来源分散，农机、地块、气象、土壤、遥感、投入品、产量等标准不一，采集不连续、质量参差，数据与作业、成本、产量缺乏闭环，难形成可执行的经营指标，同时 AI 决策要落地，需要把识别苗情、病虫草害、长势、预测（产量、风险）、优化（作业调度、投入品配方、路径）与控制（无人农机执行）贯通，而当前普遍缺少可用的数据底座、可解释的模型和稳定的现场执行链路，导致有系统但难用、数据多但难决策、指标全但难增收。

解决方案

无人化智慧农场系统的构成如下：

农场数字底座：地块一张图（边界、作物、轮作、地力等）与经营台账（投入品、成本、产量），形成统一数据标准与可追溯链路；智能感知网络：农机作业数据（位置、轨迹、工况、机具状态）+田间物联（墒情、气象等按需布设）+遥感无人机（长势、异常识别），实现多源数据自动汇聚；无人作业装备：自动驾驶拖拉机、播种施肥、植保、收获等关键机群协同作业，支持远程监控、告警与安全接管；智慧农场运营平台：生产计划、任务工单、进度看板、质量验收、对账结算一体化管理；AI决策引擎：基于数据闭环，实现识别、预测与优化，输出可执行的农事策略与作业参数。

无人化智慧农场系统具备以下核心能力：

从凭经验到靠数据：把农事决策、作业过程与经营结果打通，做到可量化、可复盘、可持续优化；质量标准化与可追溯：控制漏作重作、重叠率、行距一致性等关键指标，自动生成作业报告，支撑托管验收与结算；AI驱动增产增收：对苗情、墒情、病虫草害风险进行识别与预警，结合气象与农时窗口给出投入品与作业策略建议；机群协同提效降本：任务自动下发、路径与调度优化，减少空驶与等待，提升有效作业时间。

推广价值

无人化智慧农场的推广价值核心在于把农业生产从经验驱动、人员驱动升级为数据驱动、标准驱动，并形成可复制、可规模化的现代农业运营模式。

支撑增产增收与稳粮保供：通过数据闭环与AI决策，将播种密度、施肥喷药、灌溉排水、农时窗口等关键决策从凭感觉转为有依据，提升作业一致性与投入品利用效率，稳定单产与品质，强化粮食安全的确定性。

缓解农村劳动力短缺，推动少人化生产：以无人少人化机群替代高强度、长时段、依赖熟练机手的作业方式，降低对好机手的刚性依赖，适应劳动力持续外流与机手老龄化的现实，实现旺季不缺人、抢农时更可靠。

促进农服与托管服务规模化、标准化：无人化智慧农场提供可量化的作业标准与可追溯的作业证据链，支持按亩结算、质量验收、绩效考核与跨区调度，显著提升农服组织的运营效率与管理半径，推动社会化服务成为农业高质量发展的重要抓手。

助推绿色低碳与投入品减量增效：通过控制重叠率、漏作率与变量作业策略，减少油耗与种肥药浪费，降低面源污染风险，符合减肥减药、提质增效和“双碳”导向，具备显著的生态外部效益。

提升农业治理能力与政策落地效率：作业面积、轨迹、投入品用量、工时与质量指标可沉淀为可信数据底座，为农机补贴核验、高标准农田管护、保险理赔、灾情评估与金融授信提供依据，推动农业治理从抽查式走向数据化、精细化。

AI赋能场景 3：林果业喷药机器人

行业痛点

林果业喷药作业的痛点首先是人难找、活难干：果园用工季节性强，熟练喷药工紧缺且成本高，旺季一旦缺人就误农时。其次是劳动强度与健康风险高：背负式或车载喷药时间长、环境闷热，药雾暴露导致中毒与慢性伤害风险，且夜间、山地作业存在摔伤、翻车等安全隐患。第三是作业质量难稳定：树冠结构复杂、行距不一、风场扰动大，人工难做到均匀覆盖，易出现外层过量、内膛不足，造成防效不稳、复喷增加、药耗上升。第四是地形与通行性：丘陵山地、梯田、狭窄行间与泥泞路面多，传统机械进入难、掉头难、压根系伤土壤。第五是管理与合规：药量、配方、作业轨迹难记录，难满足绿色防控、减量用药与可追溯要求，也不利于托管结算与质量验收。

解决方案

喷药机器人的系统构成如下：

移动底盘与通行系统：履带适配山地、梯田、泥泞与狭窄行间，支持小半径转向与坡地安全防护；精准导航与避障安全：北斗高精度定位+雷达、超声多传感融合，实现行间自主行驶、障碍识别、限速与急停，支持遥控接管与电子围栏；AI感知与处方生成模块：采集果树行距、树冠体积、密度、空缺株、道路可通行性等结构信息，识别病虫害风险区域、长势差异与喷施重点区域，形成地块喷施处方（建议喷量、压力、流量、风量、行驶速度、喷头角度、开闭策略）；智能喷雾执行系统：药箱、搅拌过滤、恒压供液、流量计与阀组控制，可选风送、定向喷头，实现树冠内膛与叶背有效覆盖；作业控制与人机交互：车载控制器+操作终端，实时显示喷量、速度、覆盖与异常告警；云端作业管理平台：任务下发、轨迹与面积统计、用药量核算、作业报表与可追溯档案，沉淀数据用于后续AI模型持续优化。

喷药机器人具备以下核心能力：

AI驱动按需喷施、精准到树行、区域：基于树冠结构与长势差异，自动生成喷施处方，对稀疏区、缺株区、转弯区等进行智能开闭与剂量调整，减少空喷与过喷；自适应喷雾控制：车速联动喷量控制，结合树冠密度变化动态调整压力、流量、风量，使覆盖更均匀、漂移更可控，提升防效并降低药耗；更安全、更低暴露：人机分离，远程监管与接管，AI辅助障碍识别与风险预警，配合急停与安全联锁；复杂果园高通过性：适配丘陵梯田、狭窄行间与软土路况，减少压根与伤土，对异形地块与多障碍场景具备稳定行驶与掉头策略；作业标准化与可追溯：自动记录轨迹、面积、喷量、速度、时间、地块与处方版本，形成可核验作业证据链，支撑托管验收、合规留档与质量追责；闭环学习与持续优化：将作业结果（用药量、覆盖评价、复喷情况、病虫害变化等）回流平台，AI迭代处方推荐，实现越用越懂果园。

推广价值

林果业喷药机器人的推广价值核心在于把果园植保从高风险、靠经验、靠人力的作业，升级为低

暴露、标准化、数据驱动的生产服务体系，带来多维度外溢效应。

缓解果园用工紧缺，保障农时与稳定生产：果园喷药季节性强、劳动强度大、熟练工难招且成本高。机器人可在旺季稳定供给作业能力，减少对关键人员的依赖，降低误农时风险，提升果园经营的确定性。

显著提升安全与职业健康水平：传统喷药药雾暴露高、山地作业风险大。人机分离、远程监管与安全联锁可显著降低中毒与慢性健康风险，同时减少夜间、坡地、狭窄行间作业的安全事故概率，具有明确的社会价值。

提升防效与质量一致性，推动增产增收：果树冠层复杂导致人工喷施难以均匀覆盖，易出现外层过量、内膛不足。机器人通过恒压恒流、车速联动与定向、风送等喷雾控制，实现更稳定的覆盖与重复性，提升病虫害防控效果，减少复喷与损耗，稳定产量与商品果率。

助力减药控污与绿色防控落地：通过精准喷施、减少重叠与空喷，可降低单位面积用药量与漂移损失，减少面源污染与残留风险，契合我国减肥减药、提质增效和绿色农业发展方向，也更容易满足果品质量与市场准入要求。

推动果园托管与农服规模化、标准化经营：机器人作业可自动生成轨迹、面积、喷量、时间等作业证据，实现可追溯验收与按亩结算，降低管理成本，提高农服组织服务半径与复制能力，促进林果业社会化服务体系升级。

沉淀数据资产，形成 AI 驱动的持续优化能力：将树冠结构、作业参数、防效与复喷情况等数据沉淀为果园数字底座，支持 AI 处方推荐、风险预警与作业调度优化，实现越用越好、越做越省，推动果园管理从经验决策走向数据决策。

7.7.3 典型案例：某文旅集团

某文旅集团拥有古镇景观、古建筑群、非遗、民俗等丰富文旅景区资源，曾为千万级客流古镇文旅标杆。契合国务院发布的“人工智能+”行动意见，集团以“AI 全链路赋能+数据资产化闭环+多智能体协同生态”为核心，构建数据全流程解决方案与智能体运营体系。通过整合多类数据打通数据孤岛，实现文化保护数字化、服务体验智能化、运营管理精细化、产业融合多元化、数据资产价值化、智能协同自主化，打造古镇文旅数智转型可复制标杆。

AI 赋能场景：数智赋能文旅

行业痛点

当前核心痛点：（1）文化保护与开发矛盾突出，保护经费不足、非遗传播单一，数字化保护薄弱；（2）商业同质化严重，特色产品重复竞争，游客体验单调、复游率下滑；（3）运营管理低效，依赖人工巡查，客流疏导、设施维护响应滞后；（4）数据价值转化弱，多源数据分散，未形成标准化数据资产；（5）服务与运营协同不足，系统功能孤立，缺乏智能联动中枢。

解决方案

1.数据采集与治理：AI 构建标准化数据资产底座

(a) 全维度 AI 数据采集体系

动态数据智能捕获：运用计算机视觉（CV）技术，通过覆盖全景区高清智能摄像头与传感器，实时采集客流密度、游客移动轨迹、古建筑结构损耗等动态数据；多源数据整合接入：通过 AI 智能接口与联邦学习框架，无缝整合文旅管理部门、景区运营方、商户、线上平台等多方数据，涵盖历史客流、营收数据、非遗数字档案、商户经营数据、游客评论等静态资源，同时接入交通和气象等公共数据，形成覆盖“资源－游客－运营－消费－行业”的全维度数据体系，累计沉淀超 10TB 高质量数字资产；数据初筛与预处理：采用 AI 智能清洗算法，自动剔除冗余、异常数据，通过 LSTM 时序模型对客流、营收等数据进行趋势预判，为后续治理与资产化提供前瞻性支撑。

(b) 智能化数据治理方案

非结构化数据转化：运用自然语言处理（NLP）技术，将景区文史资料、古建解说词、游客线上评论、非遗传承口述史料等非结构化数据，转化为标准化、可检索的结构化信息；通过 AI 图像识别技术，对古建筑照片、非遗手工艺品影像进行特征提取与分类归档；关联语义模型构建：基于生成式 AI+知识图谱技术，深度关联“景点-非遗-传说-消费-商户-设施”多维度数据，构建景区专属语义模型。

2.AI 智能体多场景协同方案：从被动响应到主动服务

基于“感知－记忆－推理－行动”核心架构，构建景区“全域超级智能体”，整合游客服务、运营管理、数据资产、应急救援四大子智能体，通过 MCP（模型上下文协议）与 A2A（Agent2Agent 协议）实现跨系统、跨场景协同联动，打造自主闭环的智能生态。

(a) 游客服务智能体：全旅程个性化自主服务

全流程任务自主规划：升级 AI 语音助手为游客服务智能体，搭载大语言模型与自主规划模块，游客仅需输入模糊需求（如“亲子非遗一日游”“年轻群体夜景观光”），智能体即可自动拆解为路线规划、门票预订、体验预约、餐饮推荐、交通衔接等子任务，调用票务系统、商户平台、导航工具完成全流程执行，无需人工干预；多模态交互与记忆迭代：支持语音、文字、图像多模态交互，具备方言识别与多语种实时翻译功能，可精准理解游客复杂需求；通过长期记忆模块存储游客偏好（如饮食禁忌、游览节奏），后续访问时自动适配，实现“千人千面”持续优化的服务体验；全周期服务闭环：行前提供个性化行程方案与预订服务，行中实时推送错峰提示、景点讲解、消费推荐，行后自动生成游览报告与数字纪念册，同步开放一键投诉入口，形成“需求－执行－反馈－优化”闭环；紧急救援智能联动：集成紧急救援模块，游客触发求助后，智能体 10 秒内完成精准定位，自动联动安保、医疗资源，推送现场情况与救援路线，同时实时向游客同步救援进度，缩短救援到达时间。

(b) 运营管理智能体：精细化自主决策与执行

客流调度智能体：通过全域感知网络实时捕获客流数据，运用推理模块预判高峰时段与拥堵点位，自动生成分流方案，调度摆渡车调整路线、推送错峰信息至游客智能体、调配安保力量至热点区域，提升客流预测准确率，减少拥堵事件；设施运维智能体：持续监测古建筑状态、游览设施、环境设备数据，通过图神经网络（GNN）预判故障风险，自动生成预防性维护工单，联动维修人员终端推送任务，降低设施故障率，处置效率提升一倍；

（c）文化传承智能体：非遗与古建保护自主赋能

古建保护智能监测：持续采集古建筑结构、环境温湿度、游客接触等数据，运用推理模块预判风化、损耗风险，自动生成保护建议与修缮方案，为古建筑数字化保护提供数据支撑；非遗传承创新赋能：深度学习非遗纹样、工艺流程数据，可根据市场趋势自主生成文创设计方案，联动商户服务智能体推送至合作商户，同时支持游客通过自然语言指令定制非遗衍生产品。

3.沉浸式体验升级：AI 激活文化传播与数据资产衍生价值

（a）AI 驱动个性化体验

AR+数字人互动：开发 AR 伴游眼镜，融合 AI 实时渲染技术，游客佩戴后可与景区定制开发的 AI 数字人互动，数字人基于知识图谱实时回应历史典故咨询；通过风格迁移技术，游客可拍摄生成“水乡古风”数字影像，同步生成专属数字纪念徽章，沉淀可流通的个人文旅数据资产；AIGC 定制化内容生成：游客通过小程序上传游览需求，生成式 AI 可实时创作个性化导览文案、水乡主题诗词、短视频脚本，甚至定制非遗纹样设计方案，让体验内容从“标准化”转向“专属化”。

（b）数字文博与元宇宙场景

文物数字化修复与展示：运用生成对抗网络（GAN）技术，对景区博物馆破损的明清字画、古陶瓷进行数字化修复，还原文物原始风貌；通过 3D 全息投影+AI 交互技术，让文物“跳出”展柜，游客触摸交互即可查看制作工艺与历史背景；元宇宙数字空间：搭建多个元宇宙数字场景，涵盖古建筑虚拟游览、非遗工艺数字体验馆等，游客可通过数字分身参与沉浸式互动；基于 AI 实时渲染技术，数字场景可根据游客行为动态调整内容，同时沉淀游客互动数据，丰富数据资产维度。

4.数据资产化全链路：从确权到变现的价值转化

（a）数据资产确权与估值

区块链确权：运用区块链技术，为景区文旅核心数据资产（如客流统计数据、非遗数字资源、景区运营数据）进行权属登记，明确数据资产所有权与使用权，生成不可篡改的确权凭证；AI 精准估值：基于 LSTM 估值模型，整合经治理后的客流数据、消费数据、品牌影响力数据、非遗 IP 商业价值等核心指标，精准测算数据资产价值。

（b）数据资产交易与融资

数据交易所对接：通过数据资产管理平台，将标准化数据资产（如景区客流趋势数据、文旅消费

偏好数据、非遗数字化资源）挂牌交易，面向文旅企业、研究机构、广告商等提供数据服务，形成稳定的数据变现收入；数据资产融资：以确权后的数据资产为质押，通过 AI 风控模型向金融机构输出风险评估报告，成功获得超 1 亿元银行授信，所融资金用于智慧化升级与文化保护项目。

（c）数据资产衍生变现

非遗 IP 数字化授权：将 AI 生成的非遗纹样、水乡主题数字内容进行版权登记，通过数据交易平台向文创企业、服装品牌、游戏公司授权使用，实现非遗数据资产的商业变现；文旅数据联名合作：与互联网平台、文旅企业开展数据联名合作，共享经脱敏处理的游客偏好数据，联合开发定制化产品（如与茶饮品牌合作推出水乡主题饮品，基于消费数据优化产品口味），分享合作收益。

推广价值

（1）构建“AI 全链路赋能+数据资产化闭环+多智能体协同”三位一体模式，实现从 AI 技术应用、数据价值到智能生态的全维度升级，破解古镇文旅发展多重瓶颈。

深度整合计算机视觉、自然语言处理、预测分析等 AI 技术能力，搭建覆盖游客服务、运营管理、文化保护、商业变现四大核心场景的全链路赋能体系。其中，AI 技术贯穿游客从线上咨询、行程规划、智能导览，到线下消费、体验反馈的全过程，同时支撑景区设施运维、客流调控、安全预警等运营环节。

构建由游客服务智能体、运营管理智能体、文化保护智能体、数据资产运营智能体组成的多智能体协同系统，打破各业务模块数据孤岛与功能壁垒，实现智能体间的自主通信、任务协同与资源调度。该模式成功破解了古镇文旅游客体验单一、运营效率低下、文化资源活化不足、数据价值难以释放等多重发展瓶颈，推动景区实现技术、数据、生态的三重升级。

（2）打造文旅行业智能体应用标杆，通过四大子智能体协同联动，实现服务自主化、运营精准化、资产化高效化，推动景区从“人工管理”向“智能自主运营”跨越。

游客服务智能体集成智能问答、语音导览、路径规划等功能，实现 7×24 小时自主响应游客需求，替代传统人工咨询、引导等服务，推动服务模式从“被动响应”转向“主动预判”；运营管理智能体基于实时客流数据、设施运行数据，自动生成客流疏导方案、设备维护计划，实现运营决策从“经验驱动”转向“数据驱动”，大幅提升管理精准度；文化保护智能体通过图像识别、三维建模技术，对古镇文物建筑、非遗文化进行数字化建档与动态监测，自动预警文物损坏风险，助力文化资源长效保护；数据资产运营智能体承担数据治理、价值评估、供需匹配功能，自动完成数据资产的盘点、定价与交易撮合，提升数据资产化效率。

（3）达成文化保护、产业发展、数据价值与智能升级的四方共赢，形成“数据驱动智能升级、智能优化服务体验、服务创造商业价值、价值反哺数据建设”的良性循环，为全国同类古镇提供一套低成本、高适配、易落地的数字化转型路径。

在文化保护层面，通过数字化建档与动态监测，完成古镇核心文物、非遗技艺的系统性保护，同时借助智能导览、沉浸式体验等形式，让文化资源“活起来”，提升文化传播触达率；在产业发展层面，依托智能运营与精准营销，带动景区二次消费、文创产品销售等营收增长，孵化出智慧民宿、数字文创等新型文旅业态，拓展产业发展边界；在数据价值层面，沉淀形成标准化、可交易的文旅数据资产，通过数据授权、数据产品开发等方式实现变现，开辟景区全新增收渠道；在智能升级层面，构建起覆盖全业务流程的智能运营体系，形成可复用的 AI 技术应用能力与数据管理能力，推动景区管理模式的根本性革新。

7.7.4 典型案例：搭把手生态

宁波搭把手生态数字科技有限公司（以下简称：搭把手生态）成立于 2021 年，是一家定位于城市级固废全流程智慧化、数字化治理方案的大数据科技公司，为“世界银行贷款中国塑料减排项目——宁波城镇生活垃圾智慧分类、收集、循环利用示范项目”研发和运营提供技术支撑。通过应用区块链、大数据、AI 识别技术，全面提升垃圾分类实效，建立垃圾从源头到终端处置碳排放的可核实、可计量、可报告的智慧垃圾分类治理体系，助力垃圾减量化、资源化、无害化的目标。

AI 赋能场景：垃圾分类数字督导员系统

行业痛点

随着“无废城市”建设和“双碳”战略目标的深入推进，垃圾分类作为城市精细化治理的重要抓手，已成为提升资源利用效率、降低碳排放的关键环节。然而，传统垃圾分类模式仍面临源头数据缺失、分类质量难量化等瓶颈，难以支撑全流程闭环管理和科学决策。当前城市垃圾分类体系虽已初具规模，但在实际运行中仍存在多重结构性难题。

第一，源头数据严重缺失：传统垃圾桶无法自动记录投放行为，导致分类效果依赖人工巡查，效率低且主观性强，难以形成客观、连续的质量评价体系；第二，分类质量判定困难：居民投放方式多样（袋装、散扔）、垃圾形态复杂（湿漉、破碎、混杂），加之光照条件多变（白天/夜晚/阴雨）、摄像头易被污损，使得图像识别准确率受限；尤其在色彩相近或物体形变场景下，误判率显著上升；第三，设施兼容性差：全市垃圾投放点类型繁杂，包括密闭式智能箱、厢房式点位及露天投放区，若采用“一刀切”式改造，将导致高昂成本与资源浪费；第四，运营响应滞后：清运调度依赖经验判断，常出现“空跑”或“满溢”现象；设备故障难以及时发现，影响居民使用体验；第五，缺乏跨部门协同机制：物业、环卫、监管等主体信息孤岛问题突出，制约了整体治理效能的提升。上述痛点亟需一套高精度、低成本、可扩展的数字化解决方案予以破解。

解决方案

在此背景下，在宁波市政府的支持下，搭把手生态打造“垃圾分类数字督导员系统”，以全品类智能箱体为载体，融合人工智能、物联网与大数据技术，构建覆盖投放、收运、处置全链条的数字化治理

体系。该系统旨在实现垃圾碳足迹的精准追溯、分类质量的智能评估及运营调度的高效协同，为城市绿色低碳转型提供核心数据支撑和智能化工具。

宁波垃圾分类数字督导员系统是一套深度融合人工智能、物联网、边缘计算与城市治理逻辑的智能化平台，旨在破解传统垃圾分类在数据采集、质量评估、设施运维和跨部门协同等方面的系统性难题。该系统以“感知—分析—决策—执行—反馈”为闭环逻辑，构建了覆盖投放端、收运端、管理端和居民端的全链条数字化治理体系，其核心创新在于通过模块化设计、算法突破与生态协同，实现高精度识别、低成本部署与可持续运营。

(1) 系统架构：四维一体智能中枢

系统采用“脑—眼—手—嘴”四维架构：“脑”即云端智能决策中心，负责模型训练、任务调度与数据分析；“眼”为部署于垃圾投放点的 AI 视觉终端，具备全天候图像采集与初步识别能力；“手”指智能执行单元，如厨余垃圾物袋分离机构、满溢自动报警装置等；“嘴”则是面向居民的语音交互与信息反馈模块，用于引导正确投放、播报分类结果及积分奖励。四者通过 5G/LoRa/NB-IoT 多模通信网络实时联动，形成“前端感知—边缘处理—云端优化”的三级协同机制。

(2) 核心技术突破：精准识别与鲁棒适应

针对垃圾分类场景复杂、干扰因素多的痛点，系统自主研发多项 AI 算法。

六通道差异分割法：传统 RGB 图像在湿垃圾反光、夜间低照度或塑料袋遮挡下易失效。系统引入深度图、热成像、偏振光等多模态传感数据，构建六通道输入，显著提升在极端光照、雨雾天气及污损镜头条件下的分割鲁棒性；孪生网络与防干扰机制：为区分颜色相近但类别不同的垃圾（如绿色菜叶与塑料瓶），系统采用孪生对比学习架构，在特征空间中拉大类间距离、压缩类内方差，同时引入“防干扰”算法，自动过滤非垃圾区域（如行人、宠物、车辆）对识别的干扰；网格化面积占比分析：摒弃简单“有无判断”，系统将桶内图像划分为数百个微网格，逐格判定垃圾类型并加权统计，最终输出各品类体积占比，为分类质量评分提供量化依据；隐私保护设计：所有图像在边缘端完成垃圾模糊与身份脱敏处理，仅上传结构化数据至云端，符合《中华人民共和国个人信息保护法》要求，消除居民对监控滥用的顾虑。

(3) 硬件兼容与低成本部署

系统摒弃“推倒重来”思路，提出“硬件解耦、软件定义”策略。通过制定统一的设备接入协议（支持 Modbus、MQTT、HTTP 等），可无缝对接现有密闭式智能箱、厢房改造点、露天投放架等 80% 以上设施类型。对于老旧点位，仅需加装轻量级 AI 摄像头（成本低于 800 元/点）与单仓控制器，即可实现智能督导功能，避免重复投资。

(4) 全流程业务闭环

系统不仅识别“投得对不对”，更驱动“管得好不好”。

智能清运调度：基于满溢预测模型，系统像倒车雷达一样实时监测桶内垃圾高度，当占比接近 80% 时，便自动提醒收运，系统提前 4 小时生成最优收运路线，减少空驶率，确保垃圾仓周转效率，提升居民投放便捷性；异常行为追溯：对乱丢包、非时投放、恶意破坏等行为，系统自动抓拍、定位、关联住户信息（经授权），推送至物业或社区网格员，形成“发现—处置—反馈”闭环；碳足迹核算：每户建立“数字分类档案”，记录月度分类准确率、减量成效及碳减排量，以居民碳账户积分量化和激励分类行为，搭建碳普惠平台提供城市惠民服务兑换活动，如地铁券，洗车券等，增强居民获得感；设备健康管理：结合传感器数据和故障规则，预判箱体故障，实现设备联网与数据可视化，实时监控并跟踪故障维修进度（如电机卡滞、门锁失灵），确保设备维修及时率。

（5）数据资产与持续进化

系统已累计采集超 1000 万张真实场景图像，涵盖四季气候、昼夜时段、百余种垃圾形态，并完成 36 万条精细化标注。依托自建的“垃圾分类 AI 训练工厂”，模型每周自动迭代，通过在线学习不断优化对新型包装、地域性垃圾（如宁波汤圆袋、海鲜泡沫箱）的识别能力。

（6）生态协同与未来扩展

系统打通五大主体：居民（投放行为）、物业（现场管理）、环卫（收运执行）、运营公司（设备维护）、政府（监管宣教），通过统一数据中台实现信息共享与任务协同。未来，平台将向三个方向延伸：一是拓展至大件垃圾、装修垃圾预约回收场景；二是对接再生资源企业，实现可回收物溯源与价值分成；三是开放 API 接口，支持第三方开发碳交易、绿色金融等增值服务。

综上，宁波垃圾分类数字督导员系统不仅是技术工具，更是城市治理范式的升级——它用数据代替经验、用算法替代人力、用协同打破孤岛，为全国垃圾分类从“有人管”迈向“智能治”提供了可复制、可推广的“宁波方案”。

推广价值

宁波垃圾分类数字督导员系统不仅解决了本地治理难题，更具备显著的全国推广价值。

其一，技术可复制性强：通过标准化接口与轻量化部署模式，系统可快速适配不同城市、不同类型的垃圾投放设施，大幅降低改造门槛与财政负担，为中小城市提供“低成本、高效益”的数字化转型路径。

其二，经济可持续性突出：依托云原生架构与弹性 GPU 调度，识别成本趋近“水电价”水平，使长期、高频的智能督导成为可能，破解了以往 AI 项目“叫好不叫座”的商业化困局。

其三，政策契合度高：系统生成的碳足迹数据可直接服务于“双碳”核算、碳普惠激励及差异化收费机制探索，为国家生态文明建设提供底层数据支撑。

其四，社会效益显著：通过“红绿灯”式分类评价与全民碳积分联动，激发居民参与内生动力，推动从“被动分类”向“主动减量”转变。

未来，该系统还可延伸至装修垃圾、大件垃圾等细分领域，赋能资源化加工与商业创新。作为“AI+城市治理”的案例，该模式有望在全国“无废城市”建设中发挥示范作用，加速垃圾分类从“运动式推进”迈向“常态化智治”。

7.7.5 典型案例：蔚澜环保

宁波蔚澜环保科技有限公司（以下简称：蔚澜环保）成立于 2019 年，是一家从事环保类产业科技研发和智能装备制造为主的国家高新技术企业，围绕垃圾分类及环保生态相关衍生行业，集智能设备研发、制造、集成及配套物联网系统、运营管理平台开发于一体，以“立足研发，创建平台，服务环保事业、服务智慧城市”为愿景，致力于通过科技创新推动环保事业和城市的可持续发展。

AI 赋能场景：超级光选机

行业痛点

随着我国“双碳”战略目标的深入推进和城市精细化治理要求的不断提升，生活垃圾分类已成为生态文明建设的重要抓手。2020 年以来，《中华人民共和国固体废物污染环境防治法》《“十四五”城镇生活垃圾分类和处理设施发展规划》等政策密集出台，明确要求提升可回收物分拣效率与资源化水平。然而，在实际运行中，传统人工分拣方式存在效率低、成本高、分类精度差等问题，严重制约了再生资源回收体系的高质量发展。在此背景下，宁波市某区级生活垃圾分拣中心亟须引入智能化、模块化的先进分选装备，以实现废塑料等可回收物的高效、精准、自动化分拣。

该垃圾分拣中心采用半人工半机械作业模式，面临多重运营瓶颈：

来料成分复杂，包含 PET 瓶（透明、绿、蓝、杂色）、HDPE 瓶（大白、小白、花乙）、PP 及其他混合废塑料，人工难以快速准确识别材质与颜色，导致分选纯度不足 85%，严重影响下游再生利用价值；人工分拣劳动强度大、招工难、人员流动率高，且存在职业健康与安全风险，难以满足 ISO45001 等管理体系要求；传统分拣线布局分散、占地面积大（超 1000 m²），基建周期长，设备间协同性差，整体投资大，单位能耗高，无法匹配日益增长的回收需求；更关键的是，现有系统缺乏数据采集与分析能力，无法对分拣效率、物料流向、碳排放等关键指标进行量化统计，难以支撑精细化运营决策和数据溯源。

上述问题共同导致运营成本居高不下、资源回收率偏低、环境效益不彰，亟须通过技术革新实现系统性突破。

解决方案

针对上述痛点，蔚澜环保依托多项核心专利技术，部署了其自主研发的“超级光选机智能分拣中心大模型”系统，构建了一套高度集成、AI 驱动、模块化设计的废塑料智能分选解决方案。该方案不仅在硬件层面实现了高精度识别与高速分选，更在软件层面融合了工业互联网、大数据分析与人工智能算法，形成“感知—决策—执行—反馈”闭环，全面重构了传统分拣模式。

系统以“GL01V2-9 标准版”超级光选机为核心装备，配置两台 AI 智能分选设备，支持对 9 类常见废塑料（包括透明/绿/蓝/杂色 PET 瓶、大白/小白/花乙 HDPE 瓶、PP 及其他混合塑料）进行同步识别与精准分选。每台光选机搭载 AI 智能系统，采用可见光、近红外、高光谱等采样技术，结合深度神经网络模型，实时解析物料的特征、表面反射率、颜色分布及几何形态。系统内置的 AI 大模型数据库具有自学习功能，通过持续在线学习机制更新，确保快速识别新型包装材料，具备良好的泛化能力。

整套系统的“大脑”是“智能驾驶舱”的中央控制平台。该平台基于工业物联网架构，集成 PLC、边缘计算网关与云服务平台，实现对全产线设备的集中监控、远程诊断与智能调度。配套的“可视化数据综合管理平台”提供五大核心功能模块：

系统概况：实时展示日处理量、洗净率、碳减排量、设备利用率等 KPI；分选监控：以热力图形式呈现各品类物料的分拣数量、速度与准确率趋势；视频监控：支持多路高清画面同步调阅与智能抓拍；设备工艺：监测照明强度、风机转速、空压机压力及光选机镜头洁净度等关键参数；设备告警：通过 AI 故障预测模型提前识别潜在风险，并推送维护建议。

该系统具备强大的扩展性与兼容性。未来可无缝接入城市再生资源回收信息平台，实现与前端投放、中端收运的数据贯通；也可通过 API 接口对接碳核算系统，自动生成碳减排凭证，参与绿色金融或碳交易机制。这种“技术+数据+服务”的一体化模式，为垃圾分类从“行政推动”向“市场驱动”转型提供了坚实支撑。

推广价值

本项目成功验证了超级光选机在城市可回收物分拣场景中的技术先进性与商业可行性，具有显著的推广价值。首先，其模块化、快装式设计极大降低了智能分拣系统的部署门槛，特别适用于土地资源紧张、建设周期短的城区分拣中心改造或新建项目；其次，高分选率直接提升了再生原料品质，增强回收企业议价能力，推动“垃圾”向“资源”的高效转化；再次，AI 大模型赋能的自适应分拣能力，使系统具备持续进化潜力，可应对未来包装材料多样化趋势；此外，全流程数字化管理不仅提升运维效率，更为政府监管、碳交易及 ESG 报告提供可溯源数据。该模式可复制至全国各级垃圾分类体系，助力实现减污降碳协同增效，对推动再生资源行业智能化、绿色化转型升级具有示范意义。

7.8 本章小结

本章围绕六大重点产业及其他若干领域的案例分析揭示了中国在 AI 产业赋能上的独特禀赋：全球最大的制造业体量提供了最丰富的工业数据和最复杂的应用场景，全球最大的新能源装机容量创造了能源智能化的天然试验场，14 亿人口的健康需求催生了全球最大的医疗 AI 应用市场。更重要的是，《“人工智能+制造”专项行动实施意见》提出到 2027 年推出 1000 个高水平工业智能体的目标，以及能源产业 2026—2060 年三阶段智慧化路径——政策的前瞻性布局与产业的自发探索正在形成合力。

产业融合的逻辑正在从“AI+行业”升级为“行业×AI”。前者是加法，是在既有流程中嵌入 AI 工具；

后者是乘法，是以 AI 为底座重构产业的核心逻辑。真正的产业智能化不是在旧地图上标注新坐标，而是重新绘制地图本身。

本章所记录的二十余个案例都指向同一个方向：AI 与产业的深度融合不是未来时，而是进行时。能否抓住这个窗口期，不仅取决于技术的进步速度，更取决于我们重塑组织、升级认知、构建生态的决心和能力。这是留给每一个产业参与者的共同命题。

本章参考资料

[1]国际数据公司（IDC）.《中国工业智能体（AI Agent）市场分析及供应商推荐》.2025.6

[2]人民日报.我国连续 12 年成为全球最大工业机器人市场

[EB/OL].(2025-08-03).https://www.gov.cn/yaowen/liebiao/202508/content_7035025.htm.

[3]工业和信息化部.智能工厂梯度培育行动取得初步成效

[EB/OL].(2025-02-07).https://www.miit.gov.cn/jgsj/zbyz/gzdt/art/2025/art_4341c71b91d64262b7d205be011c7992.html

[4]工业和信息化部.“人工智能+制造”专项行动实施意见

[A/OL].(2025-12-25).https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2026/art_01010414608a4226b30687773bb21bdf.html

[5]上海市经济和信息化委员会.上海市加快推动“AI+制造”发展的实施方案

[A/OL].(2025-08-18).<https://www.shanghai.gov.cn/qtsj/20250822/c40ef1b700114ad194f43fd36704f1da.html>

[6]上海市工业互联网协会.2025 上海市“AI+制造”发展白皮书[R].(2026-01-08)

[7]中国工业互联网研究院.AI Agent 智能体技术发展报告[R/OL].(2026-01)

[8]中国银河证券.AI 加速落地：由“制造”迈向“智造”——解读《“人工智能+制造”专项行动实施意见》

[R/OL].(2026-01-08)

[9]推动人工智能与制造业“双向奔赴”

[EB/OL].(2026-01-15).<http://finance.people.com.cn/n1/2026/0115/c1004-40645854.html>

[10]具身智能：从“虚拟”迈入“现实”的关键一步

[EB/OL].(2025-07-03).https://www.cnii.com.cn/gxxww/rmydb/202507/t20250703_669974.html

[11]“AI+知识管理”为企业数智升级注入新动力

[EB/OL].(2022-03-21).https://tech.cnr.cn/techph/20220321/t20220321_525772321.shtml

[12]宝钢股份：冷轧 AI 数智中台上线

[EB/OL].(2025-03-25).<https://news.baowugroup.com/detail?id=1463&cid=356831>

- [13]沃丰科技助力施耐德电气构建知识图谱升级一体化技术服务平台
[EB/OL].(2025-06-30).<https://www.aii-alliance.org/resource/c334/n5915.html>
- [14]亚马逊设备与服务部门通过 NVIDIA AI 与数字孪生技术，实现零接触制造的关键突破
[EB/OL].(2025-08-14).<https://www.eetrend.com/content/2025/100593969.html>
- [15]以“智”赋能开启铝行业数智化转型新篇章
[EB/OL].(2025-03-27).https://www.chalco.com.cn/xwdt/qydt/202503/t20250327_145042.html
- [16]安得智联：全渠道“一盘货”，构建物流降本增效新模式——降低全社会物流成本典型案例（三）
[EB/OL].(2025-02-18).https://drc.gd.gov.cn/gkmlpt/content/4/4669/post_4669066.html#870
- [17]蔚来汽车官网.关于蔚来[EB/OL].(2026-01-24).<https://www.nio.cn/about>
- [18]蔚来集团投资者关系官网.公司简介
[EB/OL].(2026-01-19).<https://ir.nio.com/zh-hans/node/5811>
- [19]蔚来汽车官网.智能电动时代的安全创新引领者
[EB/OL].(2026-01-19).<https://www.nio.cn/smart-technology/20260119003>
- [20]蔚来汽车官网.蔚来十二全栈技术-电池系统篇
[EB/OL].(2025-12-12).<https://www.nio.cn/smart-technology/20251212001>
- [21]蔚来官方公众号.这顶 AI 帽子，让 NOMI 更有戏
[EB/OL].(2025-11-05).<https://mp.weixin.qq.com/s/lc98TtzA2wE7veaCaadPsw>
- [22]蔚来官方公众号.AI 还是太全面了，让蔚来工厂成了「未来工厂」
[EB/OL].(2025-04-11).<https://mp.weixin.qq.com/s/vfmjYuLhnuMIJKleN7DTnw>
- [23]蔚来官方公众号.看 AI 如何变「续航焦虑」为「诗意里程」
[EB/OL].(2025-02-25).<https://finance.sina.com.cn/roll/2025-02-25/doc-inemtcef9755809.shtml>
- [24]蔚来官方公众号.四大场景全面升级，「蔚来世界模型 NWM」首个版本正式推送
[EB/OL].(2025-05-30).<https://mp.weixin.qq.com/s/8h8PcrWF5hQuxiMSNUgoqg?scene=1>
- [25]华经产业研究院.中国大健康产业发展现状及趋势分析：产业升级成为趋势[EB/OL].华经情报网,2021.<https://www.huaon.com/channel/trend/737424.html>
- [26]光明网评论员.推动普惠医疗服务发展助力健全多层次医疗保障体系[EB/OL].新华网,2024.<https://www.xinhuanet.com/video/20240513/bc6aa4331beb4509acb11b43e4fc2650/c.html>
- [27]光明网.人工智能驱动个性化医疗解锁精准健康服务新篇章[EB/OL].光明网,2024.https://health.gmw.cn/2024-12/11/content_37732320.htm
- [28]王虎峰.健康中国：政策体系与发展进路[N].人民论坛,2024-03-01(02).https://paper.people.com.cn/rmlt/html/2024-03-01/content_26055435.htm

- [29]Roberts M C,et al.Precision public health in the era of genomics and big data[J].Nature Medicine,2024,30:1865-1873.<https://pubmed.ncbi.nlm.nih.gov/38992127/>
- [30]华为云.扫清行业 AI 落地三大障碍华为 DCS AI 解决方案助力瑞金医院瑞智病理大模型发布[EB/OL].华为官网,2025.<https://e.huawei.com/cn/news/2025/solutions/storage/dcs-ai-help-ruijin-hospital>
- [31]京东健康.京东健康全新发布“京东卓医”为医院、医生、患者提供卓越的智能化体验[EB/OL].生命时报,2025.<https://www.lifetimes.cn/article/4L3PvJdlbka>
- [32]Liu X,et al.A generalist medical language model for disease diagnosis assistance[J].Nature Medicine,2025,31:932-942.
- [33]刘立兵, 傅励瑶.深度学习技术在医学影像分析中的应用与展望[EB/OL].新一代信息技术网,2024.<https://www.newit.ejournal.org.cn/zh/article/doi/10.3969/j.issn.2096-6091.2024.01.005/>
- [34]Google Health.谷歌与 Aravind 眼科医院合作将检测眼疾的 AI 应用到临床中[EB/OL].ATYUN 人工智能媒体平台,2019.<https://www.atyun.com/37263.html>
- [35]一千家医院, 数百万患者, 一款通过 FDA 批准的影像 AI 软件的第四年[EB/OL].腾讯网,2022.<https://news.qq.com/rain/a/20220501A01TXD00>
- [36]Viz.ai.两项新研究证明了Viz.ai的中风解决方案对患者预后和医院经济的影响[EB/OL].Viz.ai 官网, 2025.<https://www.viz.ai>
- [37]联影医疗.AI 技术赋能医学影像与放射治疗持续推动产品差异化竞争[EB/OL].联影医疗官网,2025.<https://global.united-imaging.com/zh-cn/>
- [38]Zhang K,et al.Artificial intelligence in drug development[J].Nature Medicine,2025,31:45-59.<https://www.nature.com/articles/s41591-024-03434-4>
- [39]Jin Q,et al.Matching patients to clinical trials with large language models[J].Nature Communications,2024,15(9074).
- [40]Chun-Chieh Liao,et al.EHR-Based Mobile and Web Platform for Chronic Disease Risk Prediction Using Large Language Multimodal Models[EB/OL].arXiv,2024.<https://arxiv.org/abs/2406.18087>
- [41]胡佳敏, 邱艳, 任菁菁.AI 在基层医疗慢性病管理中的应用研究进展[J].中华全科医学,2024,22(3):481-485.
- [42]健康界.从健康监测到慢性病管理 AI 技术在健康管理中发挥了重要作用[EB/OL].健康界,2025.<https://www.cn-healthcare.com/articlewm/20250108/content-1644231.html>
- [43]卢心怡.Babylon Health: AI 与健康管理结合 (第一篇) [EB/OL].长正研究院,2022-07-05.<https://mp.weixin.qq.com/s/byskhomXwphEfqIWNxwIIA>

- [44]华为消费者业务.华为 WATCH 系列健康功能介绍[EB/OL].华为官网,2025.https://consumer.huawei.com/cn/wearables/?utm_medium=hwdc&utm_source=corp_header_consumer
- [45]氨探生物.数字生命新时代：人工智能+多组学驱动精准医疗[EB/OL].2023.<https://zhuanlan.zhihu.com/p/607197961>
- [46]Guardant Health.Guardant360 癌症基因检测产品介绍[EB/OL].科德施官网,2026.<https://www.codexgenetics.com/zh-HK/>
- [47]阿斯利康.全球首个三代 EGFR-TKI 肺癌靶向药物泰瑞沙®联合化疗一线治疗 EGFR 突变晚期非小细胞肺癌的适应症在中国获批[EB/OL].阿斯利康官网,2024.<https://www.astrazeneca.com.cn/zh/media/press-releases/2024/06-26-01.html>
- [48]华大基因.生命科技助力保险华大基因构建“预筛诊保治康”一体化解决方案[EB/OL].华大基因官网,2026.<https://www.genomics.cn>
- [49]ARK Invest.Big Ideas 2025[R].ARK Invest,2025.<https://chuhaiKaite.com/wp-content/uploads/2025/02/ARK-Invest-Big-Ideas-2025.pdf>
- [50]国家医保局.AI 辅助诊断入医保：智慧医疗的发展新机遇[EB/OL].国家医保局官网,2025.<https://www.nhsa.gov.cn>
- [51]联影医疗.联影医疗 AI 技术体系及高端影像设备介绍[EB/OL].联影医疗官网,2026.<https://www.uui-ai.com>
- [52]中国新闻网.世界首台高清“摄像”磁共振 uMR Ultra 在沪发布让疾病诊断更精准可量化[EB/OL].中国新闻网,2026.<https://www.sh.chinanews.com.cn/kjy/2026-02-02/144443.shtml>
- [53]联影医疗.上海联影医疗科技股份有限公司 2024 年年度报告[R].联影医疗,2025.<https://global.united-imaging.com/-/media/uih/pdf/investor/20250430-cn/united-imaging-healthcare-annual-report-for-2024.pdf>
- [54]上海临床研究中心.国产首款光子计数 CT 临床应用获突破[EB/OL].上海临床研究中心官网,2025.<https://clinic.shanghaitech.edu.cn/research/progress/117>
- [55]联影智能.联影智能 AI 赋能诊疗全流程解决方案[EB/OL].联影智能官网,2026.<https://www.uui-ai.com>
- [56]薛敏.勇当弄潮儿构建高端医疗装备产业新生态[N].中华人民共和国工业和信息化部,2024.https://www.miit.gov.cn/ztzl/rdzt/xxgyhqk/tjyd/tssj/art/2024/art_89625ecc27a94c58bbbd8e1bac08ec32.html
- [57]国家卫生健康委员会.中国防治慢性病中长期规划（2017-2025 年）

[Z].2017.<https://www.nhc.gov.cn/bgt/gwywj2/201702/ee48c6be0e5d4c7a8e061162672091d3.shtml>

[58]上海复星健康科技(集团)有限公司.星立方智能健康管理系统研发与应用内部资料[Z].2025(已获授权)。

[59]高小榕,高上凯.脑机接口:理论、技术、挑战和趋势[EB/OL].清华-IDG/麦戈文脑科学研究院官网,2025.<https://mcgovern.life.tsinghua.edu.cn/ch/infoshow-2847.html>

[60]强脑科技(BrainCo).强脑科技核心技术及产品体系介绍[EB/OL].强脑科技官网,2026.<https://www.brainco.cn/news>

[61]时代摘要.对话强脑科技韩璧丞:脑机接口产业迎来奇点时刻[EB/OL].新华网,2025.<https://www.news.cn/fortune/20251128/dc697048f87a4af2821eb5b70dba398c/c.html>

[62]Fii工业富联科技服务.赋能最新可持续灯塔工厂实践, AI+双碳管理驱动绿色供应链[EB/OL].(2025-09-25)[2026-02-24].https://mp.weixin.qq.com/s/j75nquDJo6oCN_yt5zSJAq

[63]工业富联.2024Fii数字化应用白皮书[R].2024.

[64]华为.《数字能源》期刊第05期[EB/OL].<https://digitalpower.huawei.com/cn/publications/issue05>

[65]王旭鹏.中企以清洁能源投资为支点加速出海[EB/OL].(2026-01-30)[2026-03-23].<https://mp.weixin.qq.com/s/Re-DNWqGdVJYzWz43XKEdA>

[66]新能源产业家.告别野蛮生长,华为 FusionSolar 9.0 如何重构高质量大基地[EB/OL].(2025-12-08)[2026-03-30].<https://xueqiu.com/1407617841/365036347>

[67]国际能源网.从“规模红利”到“价值深耕”:华为 FusionSolar9.0 为清洁能源大基地打造系统级“更优解”[EB/OL].(2025-12-09)[2026-03-30].<https://m.in-en.com/article/html/energy-2341275.shtml>

[68]太瓦能源.华为新王牌! FusionSolar9.0 如何以系统级创新,重塑光伏行业格局[EB/OL].(2025-12-08)[2026-03-30].<https://mp.weixin.qq.com/s/N1tA-OAt0jOGmdqjanSEwQ>

[69]中国电力报.华为:以 FusionSolar9.0 筑牢清洁能源大基地“价值根基”[EB/OL].(2026-01-16)[2026-03-30].<https://mp.weixin.qq.com/s/UdtjezOailb7qZW-nadvYQ>

[70]储能与电力市场.何为真构网? 华为 FusionSolar 9.0 智能组串式构网型光储解决方案给出答案[EB/OL].(2025-06-16)[2026-03-30].<https://mp.weixin.qq.com/s/A4RgZLVbQ6Qy1p16qrlY5w>

[71]比亚迪储能官网[EB/OL].<https://www.bydenenergy.com/cn>

[72]术道有方分享.比亚迪储能战略布局与盈利能力提升路径研究[EB/OL].(2025-11-24)[2026-03-23].<https://mp.weixin.qq.com/s/yuopfcCr6slr-3Ejb2k4bQ>

[73]维科网锂电.字节跳动+比亚迪:“AI+电池”创新成果落地[EB/OL].(2025-06-18)[2026-03-23].<https://mp.weixin.qq.com/s/vxwXuFGQ5FTntSi8QokLFg>

[74]中国储能网新闻中心.比亚迪储能:“基建式储能”助力新型电力系统发展

[EB/OL].(2025-05-10)[2026-03-23].<https://www.escn.com.cn/news/show-2107106.html>

[75]比亚迪太阳能.深入大漠，开启比亚迪太阳能征服极境的全新旅程！

[EB/OL].(2025-05-28)[2026-03-23].<https://mp.weixin.qq.com/s/HBPKNp7QZwGt7ocguYqoVQ>

[76]科琪电子.全球最大储能订单诞生！比亚迪拿下沙特 15.1GWh 超级项目，中国技术再破纪录

[EB/OL].(2025-04-27)[2026-03-23].<https://mp.weixin.qq.com/s/NMJTDU3vyslR0tjXaxz2cQ>

[77]机械工业出版社.新型储能技术创新路线图（摘+学）

[EB/OL].(2025-04-27)[2026-03-23].https://mp.weixin.qq.com/s/UW8l2ypdTMVxi7w3mNe_vg

[78]储能 e 家.重磅消息：比亚迪又要掀桌子了，发布 2025 工商储新品 Chess

Plus[EB/OL].(2025-03-26)[2026-03-23].https://mp.weixin.qq.com/s/QeRafqmcl_MdnYvJ-cyq2A

[79]比亚迪储能.12.5GWh 全球最大电网侧储能项目出货

[EB/OL].(2025-04-29)[2026-03-23].https://mp.weixin.qq.com/s/oyeCKa-lmOajkd-zDxRxXw?click_id=71

[80]锂电派.比亚迪“浩瀚”：搭载全球最大 2710Ah 储能刀片电池

[EB/OL].(2025-09-19).https://mp.weixin.qq.com/s/c1kLaleK_7td7Yu9lsHIdQ

[81]电车汇.储能行业进入“0.3 元时代”！比亚迪“浩瀚”储能系统全球首发

[EB/OL].(2025-09-19)[2026-03-23].<https://news.qq.com/rain/a/20250919A06UTT00>

[82]新浪财经.新研智材完成千万级种子轮融资，材料研发在 AI 技术加持下发生范式变革

[EB/OL].(2025-09-10)[2026-03-23].<http://finance.sina.cn/tech/2025-09-10/detail-infpypr9653725.d.html>

[83]赛道黑马与估值.新研智材：光刻胶配方，获龙头企业晶瑞电材投资

[EB/OL].(2025-09-19)[2026-03-23].<https://mp.weixin.qq.com/s/3P0LRbDQIUBpdF7Es4dohg>

[84]Wang S,Yin X,Wang M,Guo R,Nan K.EVOPAT:A MULTI-LLM-BASED PATENTS SUMMARIZATION AND ANALYSIS AGENT[J].arXiv,2024,2412:1-15.

[85]极新.新研智材联合创始人&CTO 南凯：材料科学新纪元——AI 大模型驱动研发效率倍增 | 2025 极新 AIGC 峰会演讲实录

[EB/OL].(2026-01-08)[2026-03-23].https://mp.weixin.qq.com/s?__biz=Mzg5MDc0NjU2OQ==&mid=2247511329&idx=1&sn=a64232548ba714dcddbde44e5bf25045&scene=0&poc_token=HIMLcmmjQIXwP SQxMF4E4zJ7Sx0TmjKM3eKObiPv

[86]中国钢研官网[EB/OL].<https://www.cisri.com.cn/gywm/jtjj.htm>

[87]人民日报.人工智能+科研——数据添“智慧”提质又高效

[EB/OL].(2025-05-12)[2026-03-23].<https://app.xinhuanet.com/news/article.html?articleId=da31350dc65e6c1ddcfb349257ceb336>

[EB/OL].(2026-01-23)[2026-03-23].https://mp.weixin.qq.com/s/A4W8wH_3vIYs-mdl4eADQA

[102]昊野科技官网.公司简介[EB/OL].<https://www.sea10000.com/?jianjie/>

[103]昊野科技官网[EB/OL].<https://www.sea10000.com/?fangan/>

[104]飞舶科技官网[EB/OL].<https://www.fuberpropulsion.cn/#page5>

[105]水下推进器智能避障系统及其控制方法申请公布号 CN 120178923 A

[106]水下火炬传递机器人有颗“河工芯”
[EB/OL].<https://xww.hebut.edu.cn/gdyw/e171d96ba0f048b384db710a0ddc3b2a.htm>

[107]“90 后”万米深海“闯将”韩野
[EB/OL].<https://app.xinhuanet.com/news/article.html?articleId=272eb48bd541abf56d0b226fa415f17c>

[108]基于多参数融合的水下推进器控制方法及相关设备申请公布号 CN 120246205 A

[109]10896 米！天津昊野科技助力悟空号再创深潜纪录！
[EB/OL].<https://www.sea10000.com/?fangan/63.html>

[110]水下磁耦合推进器智能控制方法及系统申请公布号 CN 120491700 A

[111]中海达.案例 | 高海拔山区电力巡检，2h 外业+内业齐活！
[EB/OL].(2024-04-22)[2026-03-23].<https://baijiahao.baidu.com/s?id=1796999385086594054&wfr=spider&for=pc>

[112]杨俊峰.运载无人机解锁山区运输“新姿态”[N].人民日报，2025-10-21(05).

[113]之江实验室.之江实验室官网[EB/OL].(2026-03-20).<https://www.zhejianglab.org/>

[114]汪雨晨.算力上天！首批 12 星成功入轨，“三体计算星座”开启太空计算新时代
[EB/OL].(2025-05-15)[2026-03-20].https://news.hangzhou.com.cn/zjnews/content/2025-05/15/content_8995480.htm

[115]王刚.探访“三体计算星座”“诞生地”——之江实验室
[EB/OL].(2025-05-19)[2026-03-20].<https://www.chinanews.com.cn/tp/2025/05-19/10418376.shtml>

[116]朱涵.推动 AI 在太空应用——探访“三体计算星座”指挥中心
[EB/OL].(2025-12-31)[2026-03-21].<https://www.news.cn/tech/20251231/4b32e5cbd06a4042b07161f91bd438e5/c.html>

[117]李润泽.探秘之江实验室“三体计算星座”：展示国际协作生态
[EB/OL].(2025-05-20)[2026-03-21].<https://www.chinanews.com.cn/sh/2025/05-20/10419052.shtml>

[118]东吴证券.商业航天+低轨卫星加速发展太空光伏前景广阔
[EB/OL].(2026-01-07)[2026-03-21].<https://news.qq.com/rain/a/20260107A01A9T00>

[119]陈航.算力飞上天——记浙江省之江实验室三体计算星座科研团队

[EB/OL].(2025-05-15)[2026-03-22].<https://news.qq.com/rain/a/20250515A047NG00>

[120]中国共产党第二十届中央委员会第四次全体会议.中共中央关于制定国民经济和社会发展第十五个五年规划的建议[EB/OL].2025.

[121]工业和信息化部等八部委.“人工智能+制造”专项行动实施意见[EB/OL].2026.

[122]上海市政府.上海未来产业发展规划（2024-2030 年）[EB/OL].2024.

[123]中国信息通信研究院.人工智能产业发展研究报告（2025）[R].2025.

[124]国务院.关于深入实施人工智能+行动的意见[EB/OL].2025.

[125]程乐.国产算力高效供给如何“突围”[EB/OL].2025.

[126]国家数据局.全国数据资源调查报告（2024 年）[R].2024.

[127]中国卫星导航定位协会.2025 中国卫星导航与位置服务产业发展白皮书[R].2025.

[128]解放日报.传统搜索，你还用吗？[EB/OL].2025.

[129]金山办公.WPS AI 办公智能化演进[EB/OL].2025.

[130]上海交通大学高金研究院.中国企业数据资产入表情况跟踪报告[R].2025.

[131]中国政府网.2024 年全国数据市场交易规模预计超 1600 亿元[EB/OL].2025.

[132]上海上曜科技有限公司.数据资产化全链闭环解决方案介绍[EB/OL].2025.

[133]上海库帕思科技有限公司.数据语料应用及工具链平台介绍[EB/OL].2025.

[134]中国日报网.2024 年中国隐私计算市场规模近 10 亿蚂蚁数科占比位居第一[EB/OL].2025.

[135]Musk E.Announcement on xAI Colossus Cluster Deployment[EB/OL].2025.

[136]Musk E.Announcement on Colossus 2 Super Cluster Construction[EB/OL].2025.

[137]Bernstein Research.China AI accelerators market share[R].2025.

[138]Kyle Wiggers.Perplexity Serves 100M Search Queries a Week[EB/OL].2024.

[139]Jana Garanko.Study on AI Overviews in Search:What 2025 SEO Data Tells Us About Google's Search Shift[EB/OL].2025.

[140]Luis Espada.Alphabet Advertising Revenues&Gemini App reaches 750 million users[EB/OL].2026.

[141]Dewan Ysul Zulkarnain.DeepSeek AI Statistics 2025:Users&Revenue[EB/OL].2025.

[142]PR Newswire.Generative Engine Optimization(GEO)Services Market to Hit\$7.3B by 2031,Growing at 34%CAGR[R].2025.

[143]Pan X,Dai J,Fan Y,et al.Large language model-powered AI systems achieve self-replication with no human intervention[J].2025.

[144]Patrick Pester.OpenAI's'smartest'AI model was explicitly told to shut down—and it

refused[EB/OL].2025.

[145]Kuo M,Zhang J,Ding A,et al.H-CoT:Hijacking the Chain-of-Thought Safety Reasoning Mechanism to Jailbreak Large Reasoning Models,Including OpenAI o1/o3,DeepSeek-R1,and Gemini 2.0 Flash Thinking[J].2025.

[146]David Curry.ChatGPT Revenue and Usage Statistics(2026)[EB/OL].2025.

[147]Aggarwal P,Murahari V,Rajpurohit T,et al.Generative Engine Optimization(GEO)Market Analysis[J].2024:5-16.

[148]Gartner.Gartner Predicts Over 40%of Agentic AI Projects Will Be Canceled by End of 2027[EB/OL].2025.

[149]Morgan Stanley.AI Funding:The Bull and Bear Investment Cases[R].2025.

[150]Sarah Friar.A Business That Scales With the Value of Intelligence[EB/OL].2025.

[151]GMI.Edge Artificial Intelligence Chips Market Size&Share 2025-2034[R].2025.

[152]PE Insights.Investors pile into Anthropic as revenues top\$9bn run rate[EB/OL].2025.

[153]Alphabet.Alphabet Announces Fourth Quarter and Fiscal Year 2025 Results[EB/OL].2025.

[154]Haynes Boone.China's Data as a Fifth Market Production Factor-an Asset on Your Balance Sheet[R].2024.

[155]鸿之微科技.国内首个 AI+MGI 干湿法一体生物基新材料（合成生物）智能研发创新中心启动——打造新材料研发范式变革新标杆
[EB/OL].(2025-08-30)[2026-03-23].<https://mp.weixin.qq.com/s/ILKJYJpB0rlyO8WbaYn7hA>

[156]鸿之微科技.AI 当“研发搭子”——宽禁带/超宽禁带半导体研发从“十年磨一剑”到“一键通关”
[EB/OL].(2025-12-26)[2026-03-23].<https://mp.weixin.qq.com/s/LMJkF6PKDZVE9vgedglufg>

[157]鸿之微科技.公司官网[EB/OL].[2026-03-23].<http://www.hzwtech.com>.

[158]比亚迪储能.源网侧储能解决方案
[EB/OL].<https://www.bydenenergy.com/cn/Solutions/Utility-Scale>

[159]金融时报.比亚迪储能发布工商业储能新品 Chess Plus
[EB/OL].<https://markets.ft.com/data/announce/full?dockey=1330-9415444zh-Hans-5PPRRV51JF8CMFQF8V1F5PBA1P>

[160]金融时报.BYD Energy Storage Launches Chess Plus for C&I Energy Storage in China
[EB/OL].<https://markets.ft.com/data/announce/full?dockey=1330-9413509en-525Q65A394NSNANTTR4RLQSHMU>

[161]ESIE.BYD launches new C&I highly integrated battery storage

solution[EB/OL].<https://www.pv-magazine.com/2025/04/07/byd-launches-new-ci-highly-integrated-battery-storage-solution>

[162]Evlithium.BYD Energy Storage Sets a New Standard with TS-800 Fire Test[EB/OL].<https://www.evlithium.com/lifepo4-battery-news/byd-energy-storage-ts-800-fire-test-safety.html>

[163]BYD USA.BYD Begins Construction on Intelligent Solar+Storage Energy Project[EB/OL].<https://en.byd.com/news/byd-begins-construction-on-intelligent-solar-storage-energy-project>

[164]清华大学能源互联网创新研究院.产业联盟 | “2025 能源互联网最佳实践案例”——5G+AI 驱动的蔚来车-桩-站-云-网协同控制与灵活调度实践
[EB/OL].(2025-07-11).<https://mp.weixin.qq.com/s/3beFDfOcSK0LDAQVv0pleA>

第 8 章 AI 政策：全球视野下 AI 立法趋势洞察

8.1 国内人工智能政策演进

8.1.1 “十五五”规划建议下的数字中国建设与人工智能发展

1. 战略定位：锚定国家现代化主线，确立 AI 核心驱动地位

“十五五”规划将人工智能的发展提升至前所未有的战略高度，明确其作为驱动中国式现代化建设的核心引擎与关键支柱。不仅将人工智能视为新一轮科技革命与产业变革的战略性技术，更将其定位为塑造国家未来竞争力、开辟发展新领域新赛道的决定性力量。

“十五五”规划深刻指出，面向全面建设社会主义现代化国家的宏伟目标，必须统筹高质量发展和高水平安全，将科技创新置于全局工作的核心位置。在此框架下，以数字化、智能化为表征的人工智能技术，被确立为催生新质生产力、构建现代化产业体系的核心动能。这表明人工智能不仅是技术创新的突破口，更是推动经济社会整体迈向智能化、绿色化、融合化发展的基础设施与赋能体系。人工智能发展已超越单纯的技术或产业范畴，成为一项关乎国运、牵系全局的系统性工程。

“十五五”规划强调，必须通过顶层设计与战略布局，系统推进人工智能在关键领域的深度渗透与融合应用，使其真正成为支撑国家长远发展、保障国家安全、提升治理效能、增进民生福祉的战略性资源与可靠基石，为我国在 2035 年基本实现社会主义现代化的宏伟目标，提供坚实、持久且智能的科技支撑与发展韧性。

2. 技术布局：强化关键技术突破，深化“人工智能+”全场景赋能

“十五五”规划建议将技术创新置于人工智能发展的根本位置，着力部署从基础研究到产业应用的系统性突破路径。核心在于集中优势资源，在人工智能基础理论、核心算法、关键器件等“根本技术”层面实现自主攻坚，同时构建高效协同的算力基础设施、先进算法体系与高质量数据资源池，夯实智能时代的关键要素底座。

在此之上，“十五五”规划建议前瞻性地推动“人工智能+”从战略构想迈向深度融合的产业与社会实践。其路径清晰指向两个方面：一是以人工智能引领科学研究范式、产品研发模式与产业组织形态的深刻变革，在智能制造、智慧能源、生物医药等重点领域抢占全球创新制高点；二是在社会层面，推动人工智能与城市治理、文化服务、民生保障及绿色转型等领域的全面融合，通过打造国家级应用示范区和标杆场景，构建全链条、全场景的智能赋能生态。

为保障这一进程健康、有序推进，“十五五”规划同步强调建立敏捷规范的监管框架，旨在统筹创新发展与安全可靠，在激励技术创新与防范系统性风险之间取得平衡。最终目标是实现技术发展与社会价值的协同共进，确保人工智能在赋能千行百业的同时，始终服务于经济社会的高质量发展和人类福祉的提升。

3. 政策支撑：构建“1+N+X”多层级矩阵与全链条要素保障体系

在政策架构层面，我国已系统性地构建起以“1+N+X”为框架的多层次、立体化政策支撑体系。该体系以顶层战略为统领，以专项政策为支撑，以地方实践为基础，形成了目标协同、权责清晰、覆盖完整的人工智能发展制度矩阵，为产业的健康快速发展提供了坚实的制度保障与资源动员能力。

“1”，即具有纲领性地位的国家级顶层设计。其核心是 2017 年发布的《新一代人工智能发展规划》。该文件前瞻性地提出了引领中国人工智能长远发展的“三步走”战略，明确设定了到 2030 年使我国人工智能理论、技术与应用总体达到世界领先水平，成为世界主要人工智能创新中心的宏伟目标。这份文件从根本上确立了人工智能在国家科技与产业战略中的核心地位，为后续所有政策制定与实践探索提供了根本遵循与行动总纲。

“N”，是由多部委协同推出的跨领域、跨行业配套政策集群。这些政策针对人工智能发展的关键环节与融合领域进行精准部署与细化支撑。例如《促进大数据发展行动纲要》为人工智能的燃料——数据要素的流通与开发利用奠定了基础；《“十四五”数字经济发展规划》将人工智能作为驱动数字经济深化发展的核心引擎；《智能制造发展规划（2021–2025 年）》则明确了人工智能赋能制造业转型升级的具体路径。这一系列政策共同构成了覆盖技术创新、产业融合、基础设施、数据治理等关键领域的“政策工具箱”。

“X”，是各地方政府立足区域禀赋与产业基础，推出的特色化、差异化的落地实施方案与行动规划。目前，全国已有超过 15 个省市出台了专门的人工智能产业发展行动计划或相关扶持政策。例如，北京、上海、深圳等地聚焦打造全球性人工智能创新策源地和高端产业集群；一些中西部省份则结合能源与成本优势，积极布局人工智能算力基础设施。这些地方实践不仅是对国家战略的积极响应，更通过政策创新、场景开放与生态培育，形成了百花齐放、优势互补的区域发展格局，构成了国家战略落地的坚实底座。

通过“中央统筹规划、部委专项引导、地方创新实践”的协同机制，“1+N+X”政策网络实现了纵向贯通与横向联动，显著增强了政策的系统性、协同性与执行力，为人工智能从技术突破到产业赋能的全链条发展提供了强大且可持续的制度驱动力。

8.1.2 从规划布局到法治深化：中国人工智能政策演进路径

1. 战略启航（2015–2020）：奠定发展基础，明确核心方向

中国人工智能政策的发展肇始于 2015 年，历经从战略部署到制度建设的阶段性演进。2015 年，工业和信息化部印发《中国制造 2025》，首次将智能制造提升至国家战略核心位置，标志着人工智能正式进入顶层设计视野，为其后续政策布局奠定了初步基础。

2017 年，国务院发布《新一代人工智能发展规划》，系统提出到 2030 年建成世界主要人工智能创新中心的“三步走”目标。这一纲领性文件不仅明确了技术发展路径，更标志着政策重心从早期的基础研究拓展至产业融合与生态构建，推动了人工智能从技术研发向经济社会全方位渗透。

这一阶段可视为中国人工智能政策的成长期，政策内容主要聚焦于核心技术突破、产业生态培育与试点示范推广。通过鼓励创新平台建设、推动典型场景应用和加强跨领域协同，政策体系逐步完善，为人工智能产业的快速发展营造了有利的制度与市场环境。

2.法治筑基（2021–2024）：完善制度框架，规范发展秩序

自 2021 年起，伴随《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等基础性法律的相继施行，中国人工智能政策进入以“发展与安全并重”为特征的深化期。政策体系逐步从前期侧重鼓励创新、扩大应用，转向更加注重规范有序、风险可控的系统化监管，初步构建起法律规制与伦理引导相结合的双轨运行机制。

在此过程中，中国采取了“软法先行、硬法跟进”的渐进式立法思路。前期通过发布伦理准则、行业指南等柔性规范进行价值引导与行为倡议，例如 2021 年国家新一代人工智能治理专业委员会发布的《新一代人工智能伦理规范》，明确了“以人为本、安全可控”等基本原则。在此基础上，监管力度逐步加强，通过部门规章等形式确立更具约束力的硬性要求，2023 年国家网信办等七部门联合公布的《生成式人工智能服务管理暂行办法》即一例，使中国成为全球首个对生成式 AI 服务出台专项监管规定的国家。

这一由软到硬、循序渐进的监管模式，既保持了制度的必要弹性，为技术创新与业态探索留出空间，也为后续更高位阶的立法积累了实践经验。它反映了中国在人工智能监管中寻求动态平衡的思路：既通过规范防范风险，又通过包容促进发展，从而在保障安全底线的同时持续激发产业活力。

3.融合深化（2025 年新阶段）：聚焦场景赋能，优化协同治理

2025 年“十五五”规划明确提出将人工智能作为新质生产力的重要引擎，强化其在经济社会发展中的战略地位。与“十五五”规划提出的战略方向几乎同步，国务院于 2025 年 8 月印发《国务院关于深入实施“人工智能+”行动的意见》（以下简称“行动意见”），提出到 2030 年基本建成具有全球竞争力的智能经济体系。《行动意见》围绕产业链重构、数字要素市场建设、智能社会形态塑造，部署了近 20 项重点任务，开启了“人工智能+产业+社会”三位一体协同发展的新格局。与此同时，国家设立科技创业投资引导基金专门支持早期、长期、原始创新项目，体现出国家对人工智能核心科技投入范式的结构性转向。

8.1.3 重点领域“人工智能+”监管动态与范式创新

“十五五”规划明确提出加快培育发展新质生产力，健全科技创新政策体系，推进“人工智能+”与实体经济深度融合，并加强关键要素支撑能力建设，为新阶段人工智能发展勾画出顶层设计方向。在这一战略统领下，中国不断优化政策供给方式，初步形成从国家战略到地方执行、从宏观统筹到微观落实的系统合力，推动人工智能迈入融合发展与法治保障并重的新阶段。

1.能源产业：筑牢合规监管与安全防线

继国务院《行动意见》之后，国家发改委与国家能源局于 2025 年 9 月联合发布了《关于推进“人工智能+”能源高质量发展的实施意见》，这标志着能源领域人工智能融合的顶层设计正式出炉。

《实施意见》系统地勾勒出能源 AI 融合的总体框架。在总体要求上，明确以拓展深度融合应用场景为依托，以提升创新应用技术为主攻方向，以推进智能算力与电力协同发展为支撑，核心目标是支撑能源高质量发展和高水平安全，加快培育新质生产力。该文件设定了清晰的阶段性目标：到 2027 年，初步构建能源人工智能融合创新体系；到 2030 年，推动能源领域人工智能专用技术与应用总体达到世

界领先水平。为实现这些目标，政策系统部署了“人工智能+”电网、新能源、煤炭、油气、核电等八大融合应用场景，并围绕数据、算力、算法三大基础，提出了夯实数据基础、构建算电协同机制、提升模型可解释性与绿色低碳水平等关键技术攻关方向。

在核心的监管与实施机制创新方面，政策体系展现出从宏观指导向精准落地转变的鲜明特点。在《实施意见》印发后，国家能源局于 2025 年 11 月迅速启动了配套的《关于组织开展“人工智能+”能源试点工作的通知》，并创新性地引入“揭榜挂帅”机制，旨在精准对接产业真实需求与技术供给；具体路径是：首先征集并发布行业亟须解决的高价值场景清单，随后由技术供给方“揭榜”，组建产业-技术联合开展试点，最终形成可复制推广的解决方案。这一流程标志着监管范式正向“场景驱动、以点带面”的精准化、敏捷化新阶段演进。

综合来看，此次政策更新的范式创新关键在于，它不仅回应了能源行业对技术安全性、可靠性与可解释性的严苛要求，更通过“揭榜挂帅”等机制设计，实现了“技术攻坚”与“治理探索”的同步推进。这种鼓励创新与防控风险并重的治理思路，旨在为人工智能在能源这一关键基础设施领域筑牢安全防线，同时也为其他高风险行业的 AI 应用监管提供了重要范式参照。

2.数字科技：创新算法治理与数据管控

在政策执行层面，我国持续推进财政投入与制度激励双轨并行，为人工智能发展提供稳定保障。一方面，人工智能被持续纳入国家重点研发计划和科技重大专项，聚焦核心算法、芯片架构、开源框架等技术突破，中央财政累计投入已达数百亿元。另一方面，配套实施税收减免、研发补贴、信贷支持等多重激励机制，如对符合条件的 AI 企业实施 15%企业所得税优惠，研发费用加计扣除比例提升至 100%。

在关键要素供给方面，“算力-算法-数据”三大核心支柱持续增强，构成人工智能发展的基础底座。“东数西算”工程自 2021 年启动以来，已在全国范围内部署多个超大规模算力中心。截至 2025 年 6 月，全国数据中心机架总数达 1085 万架，智能算力总量达 788 EFLOPS（每秒百亿亿次浮点运算），位居全球第二。算法侧，全国已发布 1509 个大模型，涵盖语言、视觉、多模态等多个关键领域。数据侧，数据要素市场加快培育，全国性区域性数据交易平台陆续落地，推动数据资源合规、高质量、可流通的机制逐步建立。

在标准制度方面，政策重心正从资金扶持向规范引领转移，制度供给能力日益增强。2024 年，工信部、网信办、国家发改委和国家标准委联合发布《国家人工智能产业综合标准化体系建设指南（2024 版）》，围绕基础共性、关键技术、产品服务、安全治理等七大领域，提出制定 50 项以上国家标准。2025 年 9 月正式实施的强制性国家标准 GB 45438-2025《网络安全技术人工智能生成合成内容标识方法》，首次对 AI 生成内容的标识与追溯作出统一规定，标志着 AI 监管向制度化、法治化方向迈进。

2025 年 8 月发布的《行动意见》还进一步细化了八项基础支撑任务，包括：提升大模型基础能力、构建智能算力供给体系、完善高质量数据集、优化算法开源生态、拓展应用场景、建设联合测试平台、加强人才培养、完善激励机制等。例如，文件提出推动绿色低碳智能算力调度、边缘计算能力建设，

鼓励跨行业数据协同平台与模型开源机制，支持多元主体开展联合创新。这些举措构建起制度、技术、资源联动的政策支撑生态，成为“十五五”时期中国人工智能发展行稳致远的重要保障。

3. 民生健康：健全医疗合规与伦理监管

中国持续强化对高风险医疗人工智能产品的准入监管，保障医疗安全与算法可靠性。依据《医疗器械监督管理条例》（2021 年修订）第 76 条，具有诊断功能的 AI 软件被纳入“人工智能辅助诊断软件”类别，需按照第二类医疗器械进行注册审批。企业必须提供算法训练数据来源证明、1000 例以上三甲医院临床验证报告及伦理审查材料，报国家药监局技术审评中心审批。截至 2024 年底，已有 127 个此类产品获批上市，平均审评周期从 24 个月缩短至 14 个月。国家药监局还发布了《人工智能医疗器械质量评价要求》，对算法的可追溯性和可解释性作出明确要求。例如，影像诊断 AI 需提供病灶标注与医生判读一致性的可视化报告，增强临床应用信任度。

2025 年，国家药监局对人工智能医疗产品的动态更新机制提出更高要求，强化对“算法漂移”风险的监管。政策明确，对于涉及模型结构、训练数据分布或性能指标实质性变化的 AI 医疗器械，须重新履行变更注册或补充审评程序，防止模型更新后性能下降危及患者安全。这一机制填补了“算法即产品”在生命周期管理上的空白，成为 AI 医疗监管精细化的重要进展。同年，国家卫健委与药监局联合完善了“医疗 AI 算法备案库”，要求企业提交模型变更记录、偏倚评估和年度性能报告，若发现准确率下降超 5%，将暂停临床使用并责令整改。该动态监管体系实现从准入、备案到运维全链条闭环控制，推动医疗 AI 产品持续合规与稳健演进。

针对风险较低、不直接输出诊断结论的健康管理类 AI 系统，中国监管机构延续“包容审慎”原则，依据《互联网诊疗管理办法（试行）》实施备案管理。服务平台需向属地卫健部门提交算法功能说明、数据使用承诺和用户知情模板，并接入全民健康信息平台接受事中监管。例如，用于糖尿病用药提醒的 AI 助手被归类为互联网健康服务，无需医疗器械审批，但需完成备案。2025 年，地方监管也开始围绕算法行为边界开展合规执法，防止企业通过“情绪评估”“体征分析”等表述变相实施诊断行为。对打擦边球的企业，部分省市已依据《中华人民共和国精神卫生法》和数据安全法规启动约谈与整改程序，促进健康类 AI 合法、安全、透明地发展。

4. 产业金融：构建风控体系与监管范式

在强化风险监管的基础上，中国金融管理部门同步推进“包容审慎+激励创新”的双轮政策导向，推动金融人工智能在安全可控前提下加快落地。

首先，在制度机制上，人民银行于 2022 年发布的《金融科技发展规划（2022—2025 年）》明确提出“有限豁免+监管观察期”政策安排：对于具有正向伦理价值、技术路径合规的新型 AI 应用，允许其在限定场景、限定期限内先行先试。这一机制旨在打破创新与合规的“先批准后试用”壁垒，给予探索型项目“有条件的探索自由”。例如，某大型数字银行自主研发的联邦学习反欺诈系统，采用“数据可用不可见”技术架构，在实现跨机构协同建模的同时有效保障用户隐私，被认为符合金融科技伦理导向，获得央行授予 12 个月试点豁免权。在此期间，该项目须定期向监管报送算法性能、安全评估与合规报告，

接受动态审慎监督。这一“安全围栏内的豁免机制”在保护创新积极性的同时，确保试点行为处于可观测、可回撤范围内。

其次，在国际规则话语权建设方面，中国监管机构支持国内金融科技企业参与全球 AI 标准制定，推动中国技术实践转化为国际规则成果。2024 年 6 月，由腾讯等科技企业联合主导开发的智能风控框架成功被纳入国际标准化组织（ISO）发布的 ISO 22307:2024《金融行业人工智能风险管理指南》，其内容涵盖模型开发流程管理、算法偏差检测、可解释性保障、压力测试机制等关键模块。这不仅标志着中国金融 AI 治理实践首次系统性写入国际标准，也为国内企业“走出去”提供制度背书和技术对接支撑。

通过“强监管”与“优服务”并举的政策体系，中国在实现金融系统稳定与技术自主可控的同时，也为人工智能在金融场景的创新应用打开制度空间，逐步在智能风控、金融监管科技（RegTech）等细分领域构建起具有国际竞争力的发展优势。

总体而言，中国围绕重点行业和新兴技术的监管实践表明，差异化、精细化的治理策略正逐步成熟。从生成式内容标识到自动驾驶沙盒，从医疗 AI 双轨审批到金融算法穿透监管，各领域均在探索包容审慎的监管范式。在守住安全底线的同时，给予新技术一定的试错和成长空间，成为中国人工智能监管政策的鲜明特点。这种动态平衡的治理之道也为全球 AI 治理提供了有益的经验借鉴。

8.2 国际人工智能政策动向

8.2.1 全球政策格局与国际倡议

1. 联合国主导：搭建全球 AI 治理核心框架

进入 2025 年，人工智能迅猛发展推动全球治理提上议程，联合国系统率先启动系统性行动。在第 80 届联大期间，联合国召开历史上首次 193 国参与的人工智能治理高级别会议，正式启动“全球人工智能治理对话”，为成员国平等协商 AI 规则提供制度平台。

依据 2024 年高级别顾问组报告《为了人类治理 AI》的建议，联合国大会于 2025 年 8 月通过决议成立全球首个 AI 科学评估机构——“国际人工智能科学小组”（IISP-AI），并配套设立全球治理对话机制。IISP-AI 借鉴气候领域 IPCC 模式，定期发布 AI 风险与机遇评估报告，对深度伪造、军备竞赛、舆论操控等潜在威胁发出“早期预警”，支撑全球政策制定的科学基础。

同年 9 月，联合国安理会首次围绕“AI 与国际和平安全”举行公开辩论，主要大国在军事 AI、人类控制权和战略平衡方面达成原则性共识。联合国秘书长古特雷斯在会上提出建立全球行动框架三项重点：加强对致命性自主武器的人类控制、推动制定统一监管规范、弥合全球“AI 能力鸿沟”。此外，联合国教科文组织于 2019 年推动的《AI 伦理建议书》逐步被各国纳入国内伦理治理实践，在 2025 年成为呼吁提升为具有国际约束力准则的重要基础，彰显联合国“软法引导+机制创新”的多边治理策略。

2. G7 与 G20 协同：构建国际 AI 治理政策共识平台

在政府间机制层面，G7 与 G20 成为推动全球 AI 治理政策协调的两大核心平台。G7 自 2023 年起启动“广岛人工智能进程”，2024 年推出《负责任的 AI 行为准则》，要求成员国企业在生成式 AI 开发

中确保数据合法、安全标识与内容可信等标准。2025 年轮值主席国推动构建 AI 政策通报机制，鼓励成员定期提交治理报告以强化政策透明度与协同合作。

相比之下，G20 则涵盖更多发展中经济体，其治理立场更强调“以人为本、包容发展”。2025 年约翰内斯堡峰会通过《G20 人工智能治理宣言》，提出“可信数据治理是公平 AI 的基石”，并呼吁构建数字公共基础设施，弥合全球“AI 鸿沟”。该宣言虽未获美国正式签署支持，但其表达的多边治理倾向已获得广泛认可。文件还特别支持联合国教科文组织伦理框架与公平技术发展议题，彰显发展中国家在全球规则制定中话语权的提升。

3.经合组织与多方联盟发力：推动 AI 技术细则与自律规范落地

在多边机构与行业联盟层面，全球经合组织（OECD）和世界经济论坛（WEF）通过技术导向与多方合作填补制度空白。OECD 继 2019 年发布《AI 原则》后，2025 年建立 AI 政策观察平台与治理专家组，协助成员国评估政策效果，并主导开发 AI 风险评估与透明度统一指标体系，成为 G7“广岛进程”的核心支撑工具。

WEF 则通过“全球 AI 行动联盟”集结政府、企业与研究界力量，推进跨行业的负责任 AI 实践。截至 2025 年，联盟已发展超 600 名成员，制定内容生成与模型部署“护栏”指南，并推动 AI 治理与可持续发展目标联动，强调气候、环境等领域与 AI 政策协同治理的必要性。其发布的年度报告呼吁各国将 AI 治理纳入全球议程统一设计，构建跨领域“协同治理”机制。

总体来看，2025 年全球人工智能治理已从碎片化探索进入多层次对接与逐步收敛阶段：联合国提供政治与伦理框架，G7/G20 协调主要国家政策方向，OECD/WEF 制定技术细则与行业标准。尽管尚无统一国际公约，但“原则趋同、行动并进”的格局正在形成，对全球技术竞争与国内政策演进正产生深远影响。

8.2.2 主要经济体人工智能政策比较

各主要国家和地区在人工智能政策和立法上布局各异，既体现出不同的战略考量，也反映出对风险和机遇的不同权衡。2025 年，各经济体均出台了新的政策举措，以下择要比较：

1.美国

秉持“创新优先”的战略主线，美国在政策上侧重于强化研发投入、军民融合和高科技管控，以保持全球领先地位。立法层面，尽管尚无综合性联邦 AI 法律通过，但国会两党对推动 AI 立法的意愿上升。

2025 年初，众议院有议员重新提出了《为每个美国人创造人工智能实验资源法案》（CREATE AI Act，“创造人工智能法案”），旨在将国家人工智能研究资源平台由试点转为永久机制，为学术和产业界提供长期算力、数据支持。这一平台集合了包括国防部、能源部、NASA 等 14 个联邦机构以及微软、谷歌等 26 家科技企业资源，首要任务是提供预训练模型、开放数据集等，以巩固美国在 AI 基础研究方面的优势。

行政层面，美国政府的 AI 战略因 2025 年初政府易手而出现重大调整。特朗普总统于 2025 年 1 月开启第二任期就职后，即签署了名为《消除美国在人工智能领域领导地位障碍》的行政命令，废除了

前任政府的部分 AI 政策，要求在 180 天内制定新的国家 AI 行动计划。该行政令被解读为强化联邦层面对 AI 发展的集中支持和对外技术竞争中的进攻姿态。

国防和安全领域，五角大楼在 2025 年继续推进“联合全域指挥和控制”（JADC2）等 AI 军事项目，并放宽了军事领域采用自主 AI 武器系统的一些限制，以加速 AI 在国防中的应用步伐。同时，商务部严格执行芯片出口管制，将先进 AI 芯片和相关制造装备列入管制清单，遏制高端 AI 硬件流向潜在对手。

可以说，美国总体政策在联邦层面倾向宽松监管、激励创新，联邦政府更多通过投资和标准制定（如 NIST 的 AI 风险管理框架）来引导 AI 发展，而将具体监管举措留给州和行业。例如，一些州已针对算法透明和歧视出台法律：华盛顿州要求生成式 AI 系统披露训练数据来源，得克萨斯州禁止就业 AI 招聘中存在算法偏见等，从而形成“联邦鼓励+地方规制”的二元格局。

2025 年下半年，随着美国准备在 G7 和自身 G20 主场担纲 AI 议题，可以预期其将制定更加积极的国家 AI 战略，可能包括新版的《国家人工智能研发战略计划》和政府层面的 AI 权责法规框架，以在国际规则竞争中占据主动。

2. 欧盟

欧盟在 2025 年的焦点无疑是落实和细化其具有里程碑意义的《人工智能法案》（AI Act）。该法案已于 2024 年通过并于 2024 年 8 月正式生效，成为全球首部全面的人工智能立法。按照规定，大部分条款将在两年过渡期后（即 2026 年 8 月）开始强制执行。因此，2025 年欧盟主要任务是为 AI Act 的实施做准备，包括成立欧洲人工智能办公室、制定配套细则和指导文件等。

欧盟委员会于 2025 年 4 月就通用用途人工智能（GPAI）发布了初步指南，明确提供通用大模型的供应商在 AI Act 下可能承担的义务规范。随后在 7 月，委员会发布了《通用人工智能模型指南（草案）》和《通用 AI 实践准则》，就 GPAI 的定义、生命周期管理、风险控制等给出解释性指引，并提供了一套自愿遵守的合规“说明书”。这些文件旨在帮助像 GPT-4 这类基础模型的开发者理解如何符合 AI Act 要求，例如数据治理、技术文档、透明度义务等，同时鼓励他们签署欧盟的行为准则以提前遵守高标准要求。

欧盟还启动了 AI 沙盒计划：根据 AI Act 第 57 条，每个成员国须在 2026 年前建立至少一个 AI 监管沙盒，2025 年各国已着手制定沙盒运作方案，以便创新企业在监管机构监督下测试 AI 产品。

在资金支持方面，欧盟联合成员国推出了规模达 2000 亿欧元的“投资 AI”计划和欧洲主权基金，重点投向欧洲本土 AI 芯片、云计算和基础模型研发，力图补齐产业短板，抗衡美中技术实力。

政策上，欧盟还在 2025 年发布了《AI 责任指令》草案，以修改产品责任框架，让受害人能更容易地对 AI 产品造成的损害索赔，并要求高风险 AI 产品强制保险。这与 AI Act 共同形成“管控+追责”的闭环。欧盟一贯重视国际合作来放大自身标准影响，2025 年在法国巴黎召开的全球 AI 治理峰会上，欧盟促成包括日本、加拿大等在内的多国签署了《包容性人工智能宣言》，支持欧盟关于人权和安全优先的 AI 监管理念。

总的来说，欧盟在 2025 年的政策体现出风险规制与生态培育并重：一方面通过 AI Act 树立高标准红线，禁止不可接受风险（如社会信用评分）AI 系统，要求高风险 AI 必须通过严格合规审核；另一方

面提供资金和市场支撑，避免过度监管扼杀欧洲 AI 产业。欧盟此举被誉为塑造 AI 全球治理的“布鲁塞尔效应”，但也存在内部争议——一些中小企业担心 AI Act 带来繁重合规成本，欧盟需要在实施中密切关注平衡。

3. 英国

英国采取与欧盟截然不同的路径，坚持“轻触监管”和灵活应对。2023 年英国政府发布《人工智能白皮书》，明确不制定针对 AI 的新综合法律，而由现有部门监管框架加以调节，强调保持英国对 AI 创新的吸引力。

2025 年英国继续这一方针：没有出台新的硬性监管，但通过举办全球 AI 安全峰会等方式积极参与国际议程。2023 年 11 月在布莱奇利庄园召开的首次 AI 安全峰会，使英国成为前沿 AI 治理的推动者之一。英国牵头建立了“前沿 AI 工作组”，联合顶尖 AI 实验室共同研究减轻超级智能 AI 潜在风险的方法，并计划定期组织模拟演练和红队攻防测试。

2025 年，英国政府宣布投入数亿英镑建设一个 AI 安全研究中心，专门评估高风险 AI 模型（如具有自主优化代码能力的模型）的安全性和伦理影响。这体现出英国将自身定位为“AI 安全卫士”，希望在基础模型评估、标准制定上引领全球。

政策方面，英国注重通过指导原则引导行业。例如 2025 年更新了《人工智能道德和安全风险指南》，要求公共部门采购 AI 系统时遵循透明、公平、可问责的标准，避免 AI 决策不当影响公民权益。英国还动用其监管“沙盒”机制，让金融、医疗等行业的监管机构与 AI 企业合作，在可控环境下测试新 AI 应用。

总体而言，英国寻求成为 AI 规则的“协调者”和创新的“试验场”：一方面不急于立法，以保持弹性；另一方面通过外交和倡议塑造国际讨论，追求“软实力”影响。但英国国内也有呼声要求更明确的法律约束（例如禁止 AI 在某些敏感决策中的使用），政府对此持谨慎态度，认为时机未成熟。这反映出英国希望在欧盟和美国之外走一条宽松但负责的中间路线。

4. 日本

日本在 2025 年实现了人工智能治理从战略规划向立法落地的跨越。长期以来日本以产业振兴和技术促进行动见长，相关法律相对滞后。然而面对全球趋势，日本选择通过制定促进法来抢占先机。

2025 年 5 月，日本国会通过《人工智能相关技术研发及应用促进法》，成为继美国和欧盟之后又一个在国家层面推出 AI 专门立法的主要经济体。该法简称“人工智能促进法”，定位非常清晰：不是管制，而是综合推动 AI 研发和利用，提升国民生活水平并增强日本 AI 产业的国际竞争力。

法律首先确立了人工智能研发与应用的基本原则，如注重安全和伦理、确保人类掌控、促进社会受益等，并明确政府、研究机构、企业、公众各自的责任义务。具体制度上，人工智能促进法要求政府制定国家 AI 基本计划，并在内阁下设立人工智能战略本部负责统筹实施。战略本部由相关大臣和专家组成，负责拟订 AI 产业振兴政策、协调部门举措。法律提出多项促进措施，包括：推动 AI 基础设施和开源平台建设、鼓励企业采用 AI 技术、支持 AI 在中小企业和地方的应用、完善 AI 人才培养体系等。

值得注意的是，该法没有设置新的罚则，对 AI 可能的风险主要通过现有法律框架下的软性约束解决。这体现了日本立法的核心取向：宽松促进，避免过度监管。日本政府希望借助立法向国内外传递

信号——日本鼓励 AI 创新，将提供友好的政策环境并减少不必要管制，以期吸引更多投资和人才，扭转近年来在 AI 竞赛中落后的局面。同时，日本也在伦理和标准方面积极参与国际合作，比如倡导在国际标准化组织中制定 AI 质量管理标准，牵头亚太经合组织（APEC）讨论区域 AI 治理准则。

总体看，日本的政策组合独具特色：用法律形式锁定促进 AI 的国家意志，用行政措施保障落实，但对风险监管仍然以“软治理”为主。这与欧盟形成鲜明对比，也与韩国的路径类似（韩国 2024 年通过的《人工智能基本法》也强调促进与信任构建）。可以预料，日本将继续发挥其产业优势，在机器人、制造 AI、养老医疗 AI 等细分领域推出具体支持政策，同时逐步建立与促进法相配套的指南，如 2025 年发布的 AI 治理指南草案就针对生成 AI 伦理、自动驾驶安全等提出了原则。

5.其他

除上述以外，韩国于 2024 年末制定《人工智能发展与建立信任基本法》，将于 2026 年施行，内容侧重建立国家层面的 AI 统筹治理体系和促进产业发展的各项措施，被视为在法律上确立“宽监管、重扶持”思路的又一例证。

加拿大在 2022 年提出《人工智能和数据法案》（AIDA）草案，2025 年仍在立法审议中，将要求具有高影响力的 AI 系统向政府注册并遵守透明度义务。澳大利亚选择暂不立新法，而通过更新安全、反歧视等法律涵盖 AI 问题，同时发布道德 AI 原则供企业参考。

新加坡、印度等亚洲国家更多采取政策指引和行业自律方式，如新加坡发布 AI 治理框架和模型 AI 测评指标、印度建立 AI 责任委员会提出软性规范。

非洲地区在 2023 年非盟通过《非洲人工智能战略》，倡导“泛非 AI 治理”理念，强调主权数据和本地 AI 能力建设，个别国家如加纳、南非也制定了国家 AI 政策。

可以看到，大多数国家对 AI 立法仍持审慎态度，选择通过战略和指南来观望技术演进，暂未跟进欧盟全面立法的步伐。主流观点认为，在 AI 立法领域短期内不会出现欧盟式的连锁效应，大部分国家更倾向于渐进、灵活的治理路径。这也预示着未来相当长时间内，各国 AI 监管模式将多元并存，国际合作需要在尊重不同政策取向的基础上寻求交集。

8.2.3 全球人工智能立法趋势展望

展望全球人工智能治理的未来图景，可以预见立法和规则发展将沿着融合协调、稳步推进的轨迹演进，同时在关键理念和机制上出现新的突破。

1.演进基调：从“百花齐放”到“趋同融合”，奠定多轨协同立法基础

2025 年，随着人工智能技术的深度渗透，综合性人工智能立法已被更多国家提上议程。除欧盟率先通过《人工智能法案》外，中国也已明确将《人工智能法》纳入全国人大立法规划，有望较快出台，构建涵盖数据、算法、应用等维度的法律基础框架。这将是继美欧之后又一部具有全球影响力的 AI 基础法，为后续细化立法（如自动驾驶、算法治理）提供上位依据。

与此同时，日韩等国家则通过“促进法”路径扶持产业与伦理并重，成为“轻监管、重创新”的范式代表。中等发达国家如加拿大、澳大利亚也在制定各自版本的 AI 治理条例。而发展中国家则多采取区域

合作或参考现有国际标准的方式推进相关法律建设，如拉美可能采纳欧盟模式，非洲国家在非盟框架下探索 AI 治理原则。

这一趋势反映出全球 AI 立法正从零散探索走向结构性演进——虽立法类型、路径各异，但“以人为本、透明可问责”等基本原则逐渐成为各国共识。需要指出的是，立法过程仍将遵循“积极试验、稳慎出台”的节奏，在未来 5 至 10 年间，全球可能进入“软法先行、硬法跟进”的治理过渡期。

2.核心聚焦：通用 AI 治理双轨并行，深化全周期风险防控立法

随着 GPT-5、Gemini 3、千问、DeepSeek 等通用大模型能力显著增强，通用人工智能（AGI）的治理成为全球焦点。欧盟已通过 AI 法案将 AGI（在其框架下亦作 GPAI）纳入强制监管，2025 年又出台执行指南，要求模型开发商提交风险评估与合规说明。美国与英国虽偏向行业自律，但也加强技术治理工具应用，如安全测试报告、漏洞赏金机制、水印标识等，提升 AI 系统透明度与可控性。中国则通过算法备案与分级管理体系加强安全约束，目前《生成式 AI 暂行办法》已落地，正式法规也将进一步细化责任主体、性能标准与内容审查机制。

国际合作方面，G7 广岛进程 2025 年推动 AGI 行为准则向全球扩展，OpenAI、Google 等企业已签署自愿遵守的安全与透明承诺。联合国秘书长也建议设立“全球 AI 监管机构”，以应对 AGI 带来的潜在极端风险。这些动向显示出 AI 治理正由“企业承诺—软性协作”向“全球基线—制度化治理”演进，为未来强约束国际协议的出现奠定制度基础。

3.工具升级：监管科技（RegTech）赋能，革新 AI 合规监管高效范式

AI 技术快速迭代对传统监管手段提出挑战，各国正积极探索基于科技的合规模式。2025 年，欧盟正推动建立 AI 系统合格评估中心，美国多个州实施高风险 AI 强制审计机制，中国则计划由国家 AI 安全测评中心主导关键 AI 应用测评。沙盒监管机制也不断扩展，从金融、自动驾驶延伸至医疗 AI、新药研发等高敏领域，实现“技术试错+监管共创”。

与此同时，可追溯性与责任链技术备受关注，区块链用于模型训练版本记录、数据出入存证等成为合规保障。“监管即服务”（RaaS）平台正在试点，如上海构建的监管沙箱服务系统，实现模型关键指标实时上传、AI 监控自动预警的嵌入式监管模式。国际标准互认机制也快速推进，例如 ISO 的 AI 管理体系、IEEE 的伦理认证被纳入企业合规参考，减少跨境重复审查。

这些工具反映出“技术治理技术”的新趋势，即通过数字化平台、自动化监控与标准认证实现“动态、实时、低侵入”的 AI 监管，使监管效率与产业活力同步提升。

4.全球协同：从软法共识到体系成型，构建跨境 AI 法治协同框架

虽然具有普遍约束力的国际 AI 公约尚未成形，但联合国 2024 年发布的《未来纲要之约》中明确提出“全球数字契约”（GDC），将 AI 治理理念纳入国际合作优先事项。2025 年设立的国际 AI 科学小组（IISP-AI）和全球对话机制已被视作未来国际法规制定的前奏。

各国也正探讨通过“巴黎协定式”国际协议推动 AI 军控、基础模型评估、数据治理等方面达成最低共识，以国家自主贡献为基础展开协商。G20 宣言、联合国安理会辩论、欧盟官员倡议以及科技领袖联署声明均显示：全球 AI 治理已形成三重驱动力量——政府推动、公民社会倡导、企业自律承诺。

如果说 2023–2024 年是 AI 全球治理的倡议萌芽期，那么 2025 年或许成为“AI 全球治理元年”，多层次机制开始融合，制度雏形初现，为 2030 年前形成全球数字法治新秩序奠定基础。

总的来看，2025 年的全球人工智能治理图景呈现出多轨并进、软硬结合、技术驱动的新特征。在应对 AI 不确定性风险的同时，各国需保持开放与协同，通过立法完善、技术共识与机构对接构建“可控、可信、可持续”的治理体系。这一过程中，必须避免因规则割裂而阻碍创新，也需防止因监管缺位而放任风险扩散。立足当下，面向未来，全球应携手迈向“AI 治理共同体”，使人工智能真正成为推动人类发展福祉的力量而非分裂风险的源头。

8.3 本章小结

本章对 2025 年全球 AI 政策做了梳理，正朝着全球协同、精准治理与创新赋能并重的方向深度演进。国内层面，“1+N+X”政策体系持续完善，“人工智能+”场景渗透与全链条治理双轮驱动，重点领域监管范式不断创新，为新质生产力培育筑牢制度根基。国际层面，从联合国框架到 G7/G20 协调机制，从欧盟刚性立法到多国柔性引导，全球正形成“原则趋同、路径多元”的治理格局。

未来，AI 立法将加速从“软法先行”向“软硬协同”过渡，通用大模型监管、数据要素治理、算法可追溯性成为核心焦点。监管科技的深度应用与沙盒机制的广泛推广，将实现创新活力与安全底线的动态平衡。全球范围内，跨境治理协同将进一步加强，在 AI 伦理、风险防控等领域形成更多共识性规则。

人工智能政策的演进，本质是技术进步与制度完善的良性互动。唯有坚持以人为本、包容审慎，统筹发展和安全、自主与开放，才能让人工智能真正成为推动全球经济社会高质量发展、增进人类共同福祉的强大动力。

本章参考资料

[1]国务院公报.中共中央关于制定国民经济和社会发展第十五个五年规划的建议[N],2025.

[2]国家发展和改革委员会,国家能源局.关于推进“人工智能+”能源高质量发展的实施意见:国能发科技〔2025〕73号[A/OL].

[3]国务院.关于深入实施“人工智能+”行动的意见:国发〔2025〕11号[A/OL].

[4]国家能源局.关于组织开展“人工智能+”能源试点工作的通知:国能综合通科技〔2025〕168号[A/OL].

[5]中国人民银行.金融科技发展规划(2022-2025年):银发[2021]335号[A/OL].2022.

[6]国家互联网信息办公室,工业和信息化部,公安部,国家广播电视总局.人工智能生成合成内容标识办法[Z].2025.

[7]全国网络安全标准化技术委员会.网络安全标准实践指南——人工智能生成合成内容标识服务[Z].2025.

[8]国家市场监督管理总局,国家标准化管理委员会.GB 45438—2025 网络安全技术人工智能生成合

成内容标识方法[S].2025.

[9]全国人民代表大会常务委员会法制工作委员会.关于构建人工智能法律制度体系的研究报告[R].2025.

[10]中央网络安全和信息化委员会办公室.关于依法加强数字技术和人工智能治理的指导意见[Z].2025.

[11]科学技术部.新一代人工智能国家科技重大专项（2025 年度）项目申报指南[Z].2025.

[12]中国社会科学院.人工智能与数字法治发展报告（2025）[R].北京:社会科学文献出版社,2025.

[13]European Commission.Regulation(EU)2024/1689 laying down harmonised rules on artificial intelligence(Artificial Intelligence Act)[Z].2024.

[14]European Commission.Guidelines for Providers of General-Purpose AI Models[EB/OL].2025.

[15]European Commission.General-Purpose AI Code of Practice[EB/OL].2025.

[16]The White House.Executive Order 14179:Removing Barriers to American Leadership in Artificial Intelligence[Z].2025.

[17]The White House.America's AI Action Plan[R].2025.

[18]Office of Management and Budget.Memorandum M-25-22:Driving Efficient Acquisition of Artificial Intelligence in Government[Z].2025.

[19]California State Legislature.SB-53:Transparency in Frontier Artificial Intelligence Act[Z].2025.

[20]United Nations/International Telecommunication Union.Global Artificial Intelligence Governance Action Plan[Z].2025.

[21]International Telecommunication Union.The Annual AI Governance Report 2025:Steering the Future of AI[R].Geneva:ITU,2025.

[22]OECD.Government Use of Artificial Intelligence:Risks,Governance and Policy Responses[R].Paris:OECD Publishing,2025.

[23]World Economic Forum.Global AI Governance Alliance:2025 Report Collection[R].Geneva:WEF,2025.

[24]Brookings Institution,Stanford University.AI Governance and Policy in 2025[C].Washington D.C.,2025.

[25]Harvard Law Review.Co-governance and the Future of AI Regulation[J].Harvard Law Review,2025,138(4).

[26]International Association of Privacy Professionals,Credo AI.AI Governance Profession Report 2025[R].2025.

第 9 章人工智能全球治理与国际合作

人工智能技术迭代远超人类预期，其进化速度突破摩尔定律，未来通用人工智能（AGI）更是如此，而现有社会、经济、法律等治理系统演进成本高、速度慢，难以实现快速反应与协同共振，这成为 AI 治理的重要挑战。同时，AI 技术智能涌现具有不可预测性，以大模型为代表的 AI 具备自我创造、学习、进化特性，发展方向和突破节点难以预判，未来通用人工智能等技术还可能引发社会经济结构颠覆性变革，治理系统难以及时识别潜在风险并做好准备。

鉴于 AI 治理的紧迫性与复杂性，本章将结合国务院关于深入实施“人工智能+”行动意见的内容框架，从核心原则与价值取向、风险图谱、治理架构、实践路径、基础保障和全球治理等六个维度，为解读当前人工智能治理议题并对治理实践提供方向与建议。

9.1 核心原则与价值取向

9.1.1 根本导向：以人为本与智能向善的伦理锚点

1. 尊重主权原则在数据治理中的体现

人工智能治理的根本出发点和最终归宿，在于服务人的全面发展与社会整体福祉。“以人为本”要求将人的尊严、权益与公共利益置于技术发展的中心位置；“智能向善”则明确技术应承载和平、发展、公平、正义等全人类共同价值，坚决防范技术的误用与滥用。

在数据治理与跨境流动层面，人工智能活动须严格遵守所在国法律法规，尊重各国对社会制度、数据管理及网络空间治理的管辖权，坚决杜绝利用技术干涉他国内政的行为。国际电信联盟（ITU）联合 40 多个联合国机构发起的“人工智能向善（AI for Good）”全球倡议，系统性识别和推广有助于实现联合国可持续发展目标的人工智能应用，将伦理、透明度、可问责性等原则转化为可落地的规范，为全球治理提供了价值坐标。

2. 安全可控作为不可逾越的底线要求

安全可控构成不可突破的底线要求，须贯穿模型研发、系统部署与应用运行的全生命周期。通过鲁棒性设计、风险预警、应急“熔断”机制等手段，确保人工智能系统行为可解释、可控制、可追溯，防止系统性失控。这一底线思维要求在技术应用的各个环节建立刚性约束，形成从源头到终端的完整安全保障链条。

9.1.2 核心原则：多元价值平衡的治理哲学

1. 公平普惠与创新效率的动态平衡

治理需在多重价值目标间寻求动态平衡。在公平普惠与创新效率的关系上，既要鼓励技术突破以效率驱动经济增长，又需防范技术负外部性，确保发展成果惠及社会成员，特别是弱势群体与全球南方国家。这种平衡不是静态的妥协，而是根据技术发展阶段和社会需求变化的动态调适。

2. 开放合作与自主可控的辩证统一

开放合作与自主可控呈现辩证统一关系。深度参与全球人工智能技术交流、开源生态共建与国际规则对话，是提升创新效能的必要路径；但必须坚持关键核心技术自主创新，保障产业链供应链安全，确保在开源依赖与全球化协作中保有战略主动性，避免在关键领域受制于人。

3. 敏捷治理应对技术不确定性的方法论

面对技术快速迭代与应用风险的不确定性，敏捷治理成为关键方法论。治理主体须建立技术监测预警体系，实时跟踪大模型能力演进与新兴应用场景，构建风险分级分类的动态评估框架，采取“监管沙盒”、试点示范等柔性干预方式，对确证风险实施精准、快速响应，避免治理滞后或过度干预对创新造成不当抑制。学术与政策层面提出的“适应性治理”（Adaptive Governance）理念主张，治理体系应根据技术发展和社会反馈进行动态调整，通过建立复合型分级分类标准和配置弹性制度工具箱，应对人工智能作为基础设施所带来的复杂系统性影响。

9.1.3 治理逻辑：中国式现代化的 AI 治理范式

1. 发展与安全并重

统筹发展和安全推动治理范式从“先发展后治理”向“边发展边治理”演进。2024 年发布的《国家人工智能产业综合标准化体系建设指南（2024 版）》与 2025 年配套实施的《人工智能生成合成内容标识办法》强制性国家标准，体现了标准化与法治化协同推进的治理思路，标志着治理关口前移的制度化。

2. 创新与规范协同

协调创新与规范避免“一管就死、一放就乱”的钟摆效应。实施包容审慎监管，为新技术、新业态预留试错空间；针对生成合成内容管理、算法推荐、深度伪造等已识别重大风险实施精准立法，明确红线与责任，形成“鼓励探索－规范发展－持续迭代”的良性循环。这种治理逻辑既不是放任自流的自由放任，也不是窒息创新的过度管制，而是在动态中寻找规范与创新的最佳结合点。

9.2 风险图谱：中国场景下的多维风险识别与评估

9.2.1 技术自主可控风险：安全底线的结构性制约

1. “卡脖子”技术依赖

“卡脖子”技术依赖风险主要体现在高端人工智能算力（如 GPU）、原创性核心算法与底层开发框架（如 TensorFlow、PyTorch）仍高度依赖国外供应链。国际单边管制措施可能导致关键硬件断供、软件工具禁用，不仅直接冲击大规模模型训练与产业应用部署的稳定性，更可能造成技术迭代路径的锁定效应，削弱我国在通用人工智能竞赛中的战略自主性。

2. 模型内生缺陷

模型内生缺陷风险表现为深度学习模型普遍存在的“黑箱”特性，决策过程缺乏可解释性，为故障溯源、责任认定带来困难。训练数据与算法设计中可能无意识嵌入源于西方语料与价值体系的偏见，导致模型输出与我国社会文化、伦理规范出现“价值对齐偏差”。清华大学研究显示，GPT 系列模型的中文

词表污染率高达 46.6%，且该污染已延续到后续模型，这种语料偏差不仅损害输出准确性，更在文化认知层面构成深层安全风险，影响模型的价值取向与话语体系。

3.攻防不对称

攻防能力不对称风险源于人工智能应用场景泛在化、碎片化，覆盖工业、能源、金融、社会治理等关键领域，但安全防护体系仍处于分散建设状态，未形成覆盖“云管端芯”的一体化纵深防御能力。攻击者可能利用模型漏洞发起对抗样本攻击、数据投毒或模型窃取，形成“易攻难守”的结构性矛盾。特别是在关键信息基础设施领域，单一节点被突破可能引发级联式系统性风险，这种安全态势的脆弱性构成了国家安全的潜在威胁。

9.2.2 数据安全与合规风险：信任基础的双重挑战

1.数据供给质量

数据供给质量风险在我国的突出表现为中文高质量、大规模开源数据集相对匮乏，制约了模型的语言理解与生成能力。部分开放语料存在噪声、虚假信息或价值偏见，可能导致模型产生“幻觉”。尽管国家鼓励高质量数据集有序开源开放，但仍面临中文语料污染、结构化不足与标注质量参差等现实挑战，数据清洗与价值对齐成本居高不下，这直接影响了模型训练的效果和可靠性。

2.跨境流动合规

跨境流动与合规风险主要涉及人工智能研发与全球化服务中的数据跨境流动，需应对欧盟《通用数据保护条例》（GDPR）等差异化的国际数据规则，合规成本高企。不同法域对数据本地化、跨境传输、隐私保护的要求存在显著差异，企业在“走出去”过程中面临规则冲突与合规困境。医疗健康、地理信息等敏感数据一旦泄露，将直接威胁国家安全与个人隐私，这种合规环境的复杂性构成了企业国际化的重大障碍。

3.产权与标注规范

产权与授权规范风险体现在用于模型训练的海量数据来源复杂，版权归属、授权链条尚不清晰，存在知识产权侵权风险。文本、图像、音视频等多元数据的采集、清洗、标注环节缺乏统一权属认定标准，合成数据的规范使用可能引发新的权益争议，阻碍数据要素的市场化流通，影响数据作为生产要素的价值释放。

9.2.3 国际环境干预风险：开放发展的外部约束

1.技术“脱钩”与规则排挤

技术“脱钩”与规则排挤风险表现为个别国家通过出口管制、投资审查等手段，构建针对高端人工智能芯片、软件与技术的“小院高墙”。在国际标准制定组织中，我国实质性参与度与话语权仍有提升空间，面临核心规则被预设、国内实践与国际标准难以互认的风险。2024 年 3 月，联合国大会通过美国主导的首个全球性 AI 决议《抓住安全、可靠和值得信赖的人工智能系统带来的机遇，促进可持续发展》

(A/RES/78/265)，标志着全球治理进入机制化运作新阶段。中国虽为该决议协商一致通过做出积极贡献，但如何在既有国际规则体系中争取更大话语权，仍是严峻挑战。

2. 地缘政治嵌入

地缘政治嵌入风险随着我国人工智能技术走向世界而日益凸显。技术输出可能因文化差异、治理模式不同，被不当关联地缘政治议题，遭遇市场准入壁垒、技术封锁或舆论攻击。这种地缘政治化趋势增加了技术国际合作的复杂性与不确定性，影响全球产业链供应链的稳定性，也对我国企业的海外拓展构成潜在障碍。

9.2.4 国内社会伦理风险：普惠公平的实践难题

1. 就业与认知冲击

就业与认知冲击风险集中体现为人工智能自动化应用可能加速部分传统岗位的替代，特别是在制造业、客服、文案处理等重复性劳动领域，结构性失业风险凸显。更为深层的风险在于，过度依赖人工智能辅助决策可能导致个体，特别是青少年出现“思维钝化”，削弱独立判断与批判性思维能力，形成认知依赖与能力退化，这对人力资本质量和国民素质提升构成长期挑战。

2. 数字鸿沟与区域失衡

数字鸿沟与区域失衡风险在“东数西算”工程推进中仍然存在。尽管该工程旨在优化算力布局，但短期内智算中心仍集中于东部发达地区，中西部地区在算力接入成本、网络时延与使用效率方面存在客观差距。城乡之间、不同群体之间的数字素养差异，可能加剧“智能红利”分配不均，导致技术普惠沦为形式，反而强化既有的社会不平等。

3. 伦理边界模糊

伦理边界模糊风险涉及多个层面。算法歧视已从理论风险转化为现实诉讼，2023年美国平等就业机会委员会（EEOC）指控 iTutorGroup 的招聘软件自动拒绝 55 岁以上女性和 60 岁以上男性求职者，成为史上首例 AI 招聘歧视诉讼。此类案例警示，若训练数据存在历史偏见或算法设计缺乏公平性约束，人工智能系统可能固化甚至放大社会歧视，在信贷审批、司法评估、招聘筛选等领域造成系统性权益侵害。此外，高度拟人化的 AI 交互可能引发用户情感依赖，特别是未成年人与老年群体易对虚拟角色产生不当依恋；深度伪造技术的滥用则直接威胁个人声誉、财产安全与社会信任。在自动驾驶、医疗诊断等高风险场景中，人工智能决策失误的责任主体界定仍存在法律与伦理真空，事故归责、保险理赔与道德责任分配机制亟待完善。

9.2.5 治理机制适配风险：落地效能的能力瓶颈

1. 监管滞后

监管滞后风险表现为生成式人工智能、智能体等新形态迭代周期以月甚至周计，而法规制定往往需要以年为单位，“技术－监管”周期错配导致新兴风险在监管空白期中滋生。尽管监管实践正转向构建

“开发侧－部署侧－应用侧”全链条实操框架，但跨部门协调与闭环管理能力仍待加强，存在监管重叠、标准冲突与执法碎片化现象。

2.执行能力分化

执行能力分化风险体现为头部企业拥有充足资源建立内部合规团队，而中小企业力有不逮，面临合规成本过高、技术人才短缺与法律责任不明的多重压力。中西部地区在安全监管、测评与应急响应方面的专业人才储备薄弱，导致区域间治理能力不均衡，形成“东部趋严、西部宽松”的监管套利空间，影响治理效能的整体发挥。

3.基础设施普惠

基础设施普惠风险涉及国家智算中心的统筹布局与普惠接入机制尚在完善中，中小企业与科研机构获取优质、经济算力的渠道仍待畅通。同时，精通人工智能技术与安全管理、熟悉国际规则的高端复合型人才缺口显著，特别是在算法审计、伦理评估、安全测评等交叉领域，专业化人才供给严重不足，成为制约治理能力提升的关键瓶颈。

9.3 治理架构：四维协同的核心维度

9.3.1 全生命周期管控：从研发到下线的闭环管理

1.研发阶段

治理须贯穿系统从设计、开发、部署、运行到退出的全过程。在研发阶段，必须强化源头安全设计，在基础模型、AI 芯片、框架软件等领域建立技术基线。模型架构设计阶段即融入安全要求，通过数据清洗、偏见纠正、合规预训练等方式，防范数据污染与算法歧视；建立训练数据溯源机制，确保语料来源合法、内容合规、价值观对齐。开展对抗训练与鲁棒性强化，降低模型受恶意提示注入的影响。

2.部署阶段

在部署阶段，重点防范云网安全与供应链风险。强化智算集群的访问控制、容器安全与传输加密，开展开源组件漏洞扫描与合规验证。可借鉴腾讯朱雀实验室研发的 AI Infra Guard 等工具对 AI 基础设施进行安全评估，确保供应链组件可信、配置基线合规，有效应对技术“脱钩”带来的供应链不确定性。

3.应用阶段

在应用阶段，必须建立动态监测与响应机制，应对深度伪造、信息操控等风险。通过数字水印、合成内容检测、用户行为分析等手段实现风险实时感知与处置，设立安全护栏对输入输出进行动态过滤。英特尔“Fake Catcher”等通过生物信号检测伪造影像的技术手段，可作为内容真实性验证的辅助工具，构建多层次的内容安全防护网。

4.下线阶段

在下线阶段，须系统处置模型残留风险与数据遗留问题，包括模型参数的安全销毁、训练数据的

合规删除，对外服务接口的彻底关闭。建立数据留存期限与审计日志保存规范，确保符合个人信息保护要求，防止数据泄露和模型被恶意复用的风险。

9.3.2 分领域精准规制：高风险场景的穿透式监管

1.金融 AI

金融 AI 领域重点防控算法歧视与风控透明度问题。应建立基于人机交互程度的风险分级体系，在信贷、投顾等场景强化公平性测试与决策可解释性。中国邮政储蓄银行在智能客服系统研发中，通过自研敏感词过滤组件与微调轻量模型构建双向安全过滤机制，此类实践表明金融机构可在提升服务效率的同时，通过技术手段降低不当输出风险，值得行业借鉴推广。

2.医疗 AI

医疗 AI 领域聚焦误诊责任界定与患者隐私保护。应推动高质量医疗数据集建设，通过检索增强生成（RAG）等技术降低模型幻觉，建立医疗器械审批与临床验证的加速通道，明确临床辅助决策中的责任归属，加强医疗数据去标识化与访问控制，确保患者权益和医疗安全。

3.政务 AI

政务 AI 领域强调关键基础设施的安全可控与国产化替代。在智慧城市、公共安全等场景，须优先采用自主可控技术栈，实施集中化安全管控；建立政务数据分级分类与最小权限访问原则，确保敏感信息不被过度采集与滥用，保障政府服务的安全性和可靠性。

4.生成式 AI

生成式 AI 领域须健全内容标识、溯源与幻觉防治机制。通过显隐式水印结合、元数据标准、事实核验等技术，应对虚假信息与知识产权侵权风险；建立生成内容质量评估与人工审核的衔接机制，防范有害信息的规模化传播，确保生成内容的真实性和合法性。

9.3.3 多主体协同治理：应对能力分化的均衡机制

1.政府端

政府端应加强跨部门协同，防止监管重叠与空白。网信、工信、科技、金融监管等部门须在数据治理、算法备案、科技伦理、行业准入等方面形成政策合力，推进“敏捷治理”与“分类分级”管理相结合，牵头制定人工智能安全法律法规与标准体系。建立跨部门数据共享与联合执法机制，提升监管的一致性与协同性，避免企业面临多重合规负担。

2.企业端

企业端应构建“头部企业带中小企业”的合规能力传导机制。头部企业开放安全工具、共享治理实践，助力中小企业降低合规成本，共同强化开源生态安全与供应链风险应对。中国人工智能产业发展联盟（AIIA）发布《人工智能安全承诺》，推动首批 18 家企业围绕风险管理、模型安全等六大核心承诺主动披露实践，此类行业自律机制有助于形成合规示范效应，带动产业链上下游安全水平整体提升。央视《赢在 AI+》节目显示，入围 AI 创业者中 72%拥有国内外知名高校教育背景，反映行业高知识密

度特征，亟须完善从基础教育到职业培训的全链条人才培养体系。

3.社会端

社会端须提升公众 AI 素养并培育第三方监督力量。开展全民人工智能通识教育，增强风险辨识能力；鼓励第三方评测机构发展，构建独立、公正的安全与伦理评估体系，建立公众举报受理机制；支持学术机构与民间组织参与治理研究，形成多元共治的社会基础，弥合数字鸿沟，提升全民 AI 素养。

9.3.4 技术－制度双轮驱动：中国方案的差异化路径

1.技术防护

技术防护层面应重点发展面向产业需求的安全工具链。针对提示注入、越狱攻击，可采用类似微软 Prompt Shields 的多层级过滤与意图识别技术路线；针对 AI 生成内容溯源，可参照 Google DeepMind SynthID 在多模态内容中的部署方案部署数字水印。通过红队测试模拟攻击发现模型深层漏洞，2024 年中国信息通信研究院联合 30 余家单位发起“大模型安全基准测试（AI Safety Benchmark）”，覆盖逾百万条测试数据，为模型安全提供了量化评估依据，这类技术手段应成为常态化安全保障的重要组成部分。

2.制度约束

制度约束层面须完善适应中国监管特色的规则体系。依据《人工智能科技伦理管理暂行办法（试行）》（公开征求意见稿，2025 年 8 月 22 日发布）对高风险活动实施专家复核，基于应用场景、智能化水平与规模进行科学定级。强化数据标注质量问责与算法备案后评估，建立全生命周期安全能力提升机制。针对突发风险建立应急干预流程，OpenAI 在智能体产品中引入“人类在环”确认机制，关键操作需用户二次确认，系统层面预设异常行为自动暂停，此类熔断机制设计可为国内平台企业提供参考，形成技术与制度相互支撑的治理格局。

9.4 实践路径：三侧协同的落地机制

9.4.1 开发侧：自主可控与价值对齐

1.安全基线设计

安全基线设计需加快推进国产 AI 芯片研发与适配，支持自主开源框架（如华为 MindSpore、百度 PaddlePaddle）的生态建设。2025 年 1 月，科大讯飞发布讯飞星火 X1 深度推理大模型，基于华为昇腾芯片与“飞星一号”国产算力平台完成训练，攻克 MoE 模型全链路训练效率难题，将训练效率从 30% 提升至 93% 以上，实现从底层芯片到训练框架的全栈自主可控，为行业提供了基于国产算力训练复杂大模型的可行路径，有效破解“卡脖子”技术依赖。

2.训练数据治理

训练数据治理应系统开展中文语料净化与高质量数据集建设。哈尔滨工业大学团队提出的盲点网络去噪技术，可有效提升数据质量，减少噪声与偏见对模型的影响。建立数据标注质量管控体系，确

保训练数据的准确性、多样性与价值观合规，从源头提升模型的可靠性和文化适应性。

3.模型安全对齐

模型安全对齐须在模型微调阶段，采用监督微调（SFT）、直接偏好优化（DPO）等技术，将社会主义核心价值观、法律法规要求及伦理准则嵌入模型行为规范。遵循《人工智能安全治理框架 2.0》“可信人工智能基本准则”，注重“价值观对齐”与“人类最终控制”原则，建立模型行为的红线与触发机制，确保模型输出符合中国社会文化语境。

4.红队测试与幻觉检测

红队测试与幻觉检测须依托国家级 AI 安全靶场（如国家人工智能安全治理基地），组织开展针对大模型的对抗性测试，重点检测模型在幻觉生成、误导输出、隐私泄露等方面的脆弱性。建立常态化红队测试机制，在模型发布前进行多轮对抗性评估，持续监控模型在实际应用中的安全风险，提升攻防对抗能力。

9.4.2 部署侧：供应链韧性与基础设施普惠

1.供应链多元化

供应链多元化须积极布局非美技术路线，加强国产算力芯片、服务器、操作系统等关键环节的自主化替代，构建多源供应、安全冗余的供应链体系。建立关键软硬件备品备件库与替代方案清单，确保在极端情况下核心业务连续性不受影响，有效应对技术“脱钩”风险。

2.“东数西算”安全加固

“东数西算”安全加固须强化智算中心的安全防护能力，完善云网一体化安全体系，实施跨区域容灾备份。建立算力资源调度与安全防护的协同机制，确保东西部算力协同过程中的数据传输与存储安全。加强云端容器镜像安全，可借鉴 AI Infra Guard 等安全评估工具开展漏洞扫描，部署供应链软件成分清单（SBOM）管理，持续开展漏洞扫描与配置核查，保障算力基础设施的安全性。

3.权限管控与漏洞扫描

权限管控与漏洞扫描须在能源、交通、金融等关键信息基础设施行业加快推进 AI 系统的国产化替代与安全加固，部署细粒度访问控制策略，实施零信任架构，确保“最小权限原则”在技术层面的刚性执行，防止因权限失控导致的安全事件。

9.4.3 应用侧：风险监测与责任追溯

1.访问控制与身份核验

访问控制与身份核验须针对 AI 技术可能被用于电信诈骗、钓鱼攻击等违法犯罪活动，在服务接入层面强化实名认证与行为审计。2024 年 OpenAI 推出“Verified Organization”强制验证机制，要求开发者完成企业身份核验后方可使用其高级模型 API，此类身份前置验证机制可有效追溯恶意使用源头。国内平台应建立类似的实名制与行为审计体系，确保技术应用的可追溯性。

2.生成内容标识溯源

生成内容标识溯源须全面落实 AI 生成内容的标识管理，通过显性水印、隐性元数据等技术，实现内容来源可追溯。我国《人工智能生成合成内容标识办法》（2025 年 9 月 1 日施行）明确平台的内容标识责任。推广 C2PA（Content Authenticity Initiative）等行业标准，建立跨平台的内容溯源与验证机制，提升虚假内容的识别与处置效率，有效打击深度伪造。

3.人工复核机制

人工复核机制须在医疗诊断、司法裁判、金融风控等高风险领域建立“人机协同、人在回路”的决策机制，明确 AI 系统辅助定位，保留人工最终审核与否决权。建立人工复核的标准作业程序与责任追溯链条，确保关键决策的人类可控性，防止算法错误导致的严重后果。

4.动态监测预警

动态监测预警须针对 AI 可能带来的就业结构冲击、认知能力退化等长期社会风险，构建多维度社会影响监测指标体系，依托大数据与社会计算手段，开展动态风险评估。腾讯 EvidentlyAI 等开源工具可用于监测 AI 系统输出合规性与用户行为偏移，此类技术手段应纳入常态化风险监测体系，实现从被动应对向主动预防的转变，确保社会发展与技术进步相协调。

9.5 基础保障：支撑治理现代化的能力体系

9.5.1 标准体系建设：争夺国际话语权

1.分类分级标准

分类分级标准须依托《国家人工智能产业综合标准化体系建设指南（2024 版）》，建立覆盖基础共性、关键技术、行业应用、安全治理的全链条标准框架。2025 年，工业和信息化部人工智能标准化技术委员会（MIIT/TC1）发布《工业和信息化领域人工智能安全治理标准体系建设指南（2025 版）》，细化行业落地指引，推动标准从通用要求向细分领域延伸，建立与国际互认但体现中国特色的 AI 风险分级体系。

2.安全测评体系

安全测评体系须针对中文语境的特殊性，加强幻觉检测、偏见评估等方法论研究与标准研制。依托中国信息通信研究院与 AIIA 发起的大模型安全基准测试体系，持续完善测评方法论覆盖范围，为行业提供可量化的安全评估工具，形成具有中国特色的 AI 安全评估话语体系。

3.开源生态规范

开源生态规范须关注开源治理规则与贡献者责任界定。《关于深入实施“人工智能+”行动的意见》明确支持人工智能开源社区建设，建立健全开源贡献评价和激励机制。产业界积极拥抱模型上下文协议（MCP）等标准化交互协议，优化 AI 模型与外部数据源、工具的协同效率，防范开源组件的供应链安全风险，推动国产开源框架的规范化发展。

9.5.2 资源支撑优化：破解瓶颈制约

1.智能算力统筹

智能算力统筹须加快推进全国一体化算力网建设，优化国家智算资源布局。鼓励智能算力互联互通和供需匹配，发展标准化、可扩展的算力云服务。建立算力资源普惠供给机制，通过算力券、补贴等方式降低中小企业与科研机构的算力获取成本。据 Epoch AI 研究，2010-2023 年间 AI 训练算力需求每 6-10 个月翻倍，算力成本年均下降约 30%，应充分利用技术降本趋势，推动算力普惠与高效利用，缓解算力分布不均问题。

2.数据要素供给

数据要素供给须推动公共数据有序开放，在医疗、教育等高价值领域开展试点，鼓励通过合成数据技术缓解数据隐私与供给不足的矛盾。2024 年，科大讯飞主导制定的 ISO/IEC 5259-4:2024《人工智能分析和机器学习的数据质量第 4 部分：数据质量过程框架》正式发布，填补数据质量国际标准空白，为国内数据质量建设提供了对标基准，提升我国在数据治理领域的国际影响力。

3.语料库建设

语料库建设须面对数据污染与语料质量不均的挑战，建设国家级高质量中文语料库，强化多模态、跨领域语料的采集、清洗与标注规范。建立语料质量评估与动态更新机制，确保语料的时效性、准确性与文化适应性，对抗西方中心主义的数据偏见，提升中文语料在模型训练中的质量和比重。

9.5.3 能力建设：缩小数字鸿沟与人才缺口

1.人才培养

人才培养须推动人工智能全学段教育，加强高层次、复合型人才培养，强化算法安全、伦理治理等专业方向。支持高校设立人工智能安全与治理交叉学科，培养既懂技术又通晓法律、伦理的复合型人才。产业侧通过参与国际国内标准制定培养专业人才，特别注重中西部基层监管人员的培训，缩小区域人才差距。

2.社会意识

社会意识提升须将全民人工智能素养教育纳入政策视野，重点面向农村、老年等群体开展普及培训，防范数字弱势群体在智能化转型中被进一步边缘化。通过媒体宣传与案例警示，提升公众对深度伪造、算法歧视等风险的辨识能力，构建全民参与的 AI 治理社会基础。

3.技术研发攻坚

技术研发攻坚须聚焦可解释人工智能、隐私计算、后量子密码等治理相关关键技术。科研机构与企业在联邦学习、同态加密、算法审计等方面积累实践，能源、金融、通信等行业领先企业探索“可解释 AI+专家规则双校验”机制，构建融合人工智能的数据安全智能管理与模型协同体系，以技术治技，提升治理的精准性与效率。

9.5.4 应急机制：应对突发与系统性风险

1. 风险预警

风险预警需政府引导开展人工智能社会实验，监测技术对就业、伦理、安全等方面的潜在冲击。国家智能社会治理实验基地通过“组织应用－科学测量－综合反馈”的实验规程，系统性评估技术社会影响，为政策调整提供实证依据。建立跨部门的风险情报共享与研判机制，及时发现就业市场冲击、算力供应中断等系统性风险的早期信号。

2. 熔断管控

熔断管控需探索模型下线、服务暂停的触发条件与程序规范。明确风险分级响应阈值，对出现严重价值取向偏差、大规模数据泄露或安全漏洞不可修复的模型，实施强制下线与暂停服务。配套强制性国家标准，为违规内容处置提供技术依据，确保熔断机制的程序正当性与救济渠道畅通，形成有效的风险防控最后防线。

3. 漏洞共享

漏洞共享须推进中国信息通信研究院牵头建设的国家人工智能安全漏洞库（人工智能专业漏洞库），构建漏洞发现与贡献机制，鼓励企业、研究机构上报漏洞。建立漏洞修复的协同响应流程，确保高危漏洞得到及时修补。在脱钩背景下，积极探索跨境威胁情报共享的有限合作空间，降低系统性安全风险。

9.6 全球治理：国际协同的中国方案

9.6.1 规则对接与话语权争夺

1. 多边机制参与

多边机制参与须深度融入联合国、亚太经济合作组织（APEC）、二十国集团（G20）等多边框架下的人工智能治理议程。中国积极参与联合国框架下的 AI 治理机制，包括支持通过 A/RES/78/265 号决议，深入参与“国际人工智能科学小组”和“全球人工智能治理对话”机制筹建，推动建立体现发展中国家诉求和“人本导向、普惠共享”的国际治理共识，提升制度性话语权。

中国专家与机构在国际电信联盟（ITU）、国际标准化组织（ISO）等关键标准组织中担任重要职务，主导或深度参与人工智能术语、管理体系、数据质量、大模型评测等多项国际标准的研制工作，推动国内实践经验转化为国际规则，打破西方规则垄断。

2. 标准互认推动

标准互认推动须在“一带一路”共建国家推广基于共同发展理念的人工智能标准与实践，建立与欧盟、东盟等主要经济体的标准化合作机制，推动在数据跨境流动、算法透明度等关键议题上的规则互认，降低中国企业出海面临的合规成本，构建更加公平合理的国际 AI 治理秩序。

9.6.2 全球南方能力普惠：破解智能鸿沟

1.技术输出与治理赋能

技术输出与治理赋能须向发展中国家输出人工智能技术与解决方案时，配套输出相应的数据安全法规指导与治理框架，避免技术输出与治理能力建设脱节导致的滥用风险。中国提出《人工智能能力建设普惠计划》，明确在基础设施、产业赋能、人才培养、数据建设、安全治理等领域的“十项行动”，并与赞比亚等国家共同推动“人工智能能力建设国际合作之友小组”，通过技术援助与治理经验分享，帮助发展中国家建立适配本土的 AI 治理体系，规避地缘政治连带风险。

2.算力与数据普惠

算力与数据普惠须倡导并推动国际开源数据集建设与跨境算力共享机制。支持建立面向全球科研协作的开放科学数据平台，特别是在气象、农业、医疗等公共利益领域，推动数据资源的全球共享。鼓励国内机构与国际组织合作建设多语言、跨文化的高质量数据集，提升非英语语种在 AI 训练数据中的代表性，打破英语中心主义的数据垄断，促进全球 AI 发展的包容性。

9.6.3 风险共治：在竞争受限下的有限合作

1.威胁信息共享

威胁信息共享须在网络安全和利用人工智能实施跨国犯罪等领域，支持在双边或多边框架下建立威胁信息共享机制。特别是在深度伪造检测、电信诈骗防范等具有共同利益的领域，推动建立实时情报交换与技术协作机制，提升全球应对 AI 滥用行为的能力，在应对共同挑战中寻求合作空间。

2.技术滥用防范

技术滥用防范须在人工智能军事化、致命性自主武器系统（LAWS）等最具争议和风险的领域，积极参与联合国《特定常规武器公约》框架下的政府专家组讨论，推动建立具有法律约束力的国际规范，限制 AI 在军事领域的滥用，维护全球战略稳定。主张建立全球治理规范，倡导各国承诺在 AI 军事应用上保持克制，优先将 AI 技术用于和平与发展目的，提升中国负责任大国形象，推动构建人类命运共同体。

9.7 本章小结

人工智能治理是技术演进与社会治理深度融合的长期命题，其复杂性伴随技术迭代持续深化。本章构建的“价值引领—风险识别—架构设计—路径落地—保障支撑—全球协同”六位一体治理体系，既立足中国场景破解技术自主、数据安全、伦理规范等现实难题，又以“以人为本、智能向善”为核心锚点，平衡创新与规范、开放与自主的辩证关系，彰显了中国式现代化治理的独特优势。

治理之路无终章，技术创新与风险演化的动态博弈，要求治理体系必须保持敏捷迭代的生命力。未来，需持续强化标准体系与技术工具的协同支撑，补齐人才短板与区域治理不均衡的短板，推动多主体协同治理走向纵深。同时，以更开放的姿态参与全球治理规则共建，通过技术普惠、规则互认与风险共治，破解智能鸿沟与地缘政治壁垒，为全球人工智能安全、公平、可持续发展贡献中国智慧与

方案，让智能技术真正成为赋能人类共同福祉的强大动力。

本章参考资料

- [1]世界互联网大会人工智能专业委员会标准推进计划.2025 年全球人工智能标准发展报告[A/OL].2025.11.
- [2]国务院.关于深入实施“人工智能+”行动的意见.国发〔2025〕11 号[EB/OL].
- [3]中国信通院.人工智能安全治理框架 2.0 版[A/OL].2025.9.
- [4]中国工信部.人工智能科技伦理管理服务办法（试行）（公开征求意见稿）[Z].2025.8.
- [5]2025 世界人工智能大会暨人工智能全球治理高级别会议.人工智能全球治理行动计划[A/OL].2025.8.
- [6]中央网信办秘书局.市场监管总局办公厅.智能社会发展与治理标准化指引（2025 版）[N].2025.6.
- [7]中央广播电视总台.阿里云研究院.中国人工智能应用发展报告（2025）[EB/OL].2025.
- [8]中国信通院.人工智能安全治理研究报告（2025 年）[EB/OL].2026.1.
- [9]张凌寒.人工智能法律治理的路径拓展.中国社会科学[R].2025 年第 1 期 P91—P110.
- [10]敦帅.陈强.贾婷.中国人工智能治理研究述评与展望.中国科技论坛[R].2025.4.
- [11]吴沈括.法律人需在 AI 治理中保持独立清醒的判断.法制网[EB/OL].2025.4.
- [12]社会治理引入 AI，技术伦理争议亟待重视.南方周末官方公众号[EB/OL].2025.3.
- [13]孙保学.研究 AI，拥抱 AI，更要掌控 AI——人工智能治理的三重态度.新湘评论[EB/OL].2025.6.
- [14]薛澜.清华大学人工智能国际治理研究院.AI 治理如何跟上技术进步？中国新闻周刊[M].2025.3.
- [15]Qingjie Zhang.,Di Wang.,Haoting Qian.,Yan Liu.,Tianwei Zhang.,Ke Xu1.,Qi Lil.,Minlie Huang.,Hewu Lil.,Han Qiu1.Speculating LLMs'Chinese Training Data Pollution from Their Tokens[C].EMNLP 2025 Main,2025.

第 10 章人工智能标准体系建设与实施

10.1 中国人工智能标准的应用现状

当前国内人工智能标准体系已初步建立，涵盖国标、行标、团标、地标等多层级框架，覆盖基础共性、关键技术与行业应用。目前国内采用的标准以团标和行标为主，国标提供基础框架，地标作为区域补充。《国家人工智能产业综合标准化体系建设指南（2024 版）》明确将国标定位为基础框架，而非直接应用标准。

表 10.1 人工智能标准分类

标准类型	状态	代表标准	产业应用情况
国标	提供基础框架	GB/T 45654-2025 GB/T 45288.1-2025 GB/T 45087-2024	框架性作用： 作为基础术语、安全治理等通用标准，如 GB/T 45654-2025 用于生成式 AI 服务安全评估
行标	垂直领域主力（金融、机械等）	JR/T 0287-2023 JR/T 0298-2023	垂直领域主力： 在金融、机械、化工、能源等行业深度应用，如金融行业标准由中国人民银行发布
团标	持续发布中（企业采纳度高）	T/AIIA 019-2025 T/AIIA 014-2025 AIIA/T 0219-2025	应用最广： 团体标准在制造、汽车、医疗等场景应用广泛，头部企业（如百度、华为、用友等）积极参与编制和采纳
地标	区域产业支撑（上海、广东等）	DB31/T 1490-2024 DB44/T 2739-2025	区域产业支撑： 在特定区域（如上海、广东）的 AI 产业集群中应用，覆盖企业数量有限

截至 2024 年 7 月，我国已发布 28 项人工智能国标、127 项 AI 行标、217 项 AI 团标，以及 42 项地标。各类标准在实际应用中呈现出“金字塔”结构：团标应用占比超过一半，行标次之，国标提供基础框架，地标作为区域补充。

1. 国标守底线

国家标准由国家标准化管理委员会主导制定，定位为“基础框架”，而非直接应用标准。截至 2024 年 7 月，我国已发布 28 项人工智能国标。

主要国标包括：

- GB/T 45654-2025 《网络安全技术生成式人工智能服务安全基本要求》
- GB/T 45288.1-2025 《人工智能大模型第 1 部分：通用要求》
- GB/T 45087-2024 《人工智能服务器性能测试方法》

•GB/T 45288.1/2/3-2025《人工智能大模型系列标准》（通用要求、评测方法、成熟度评估）等

2.行标专领域

行业标准在垂直领域深度细化国标要求，成为行业落地核心依据。据国家人工智能标准化总体组发布的《国家人工智能标准化年度报告（2025 年）》统计，目前发布的人工智能行业标准约 100 余项，大约有 150 项尚在制定中。其中主要集中在金融科技、医疗健康、汽车与交通、公共安防、电子信息制造、工业与能源六大行业领域。以下列举成熟度较高的三个具有代表性的行业：

•金融科技：最成熟、体系最完整。已发布 10 余项行标，覆盖算法模型、风险管理、应用规范等全链条。其中比较具有代表性的有 JR/T 0221-2021《人工智能算法金融应用评价规范》，规定了人工智能算法在金融领域应用的基本要求、评价方法、判定准则；JR/T 0237-2021《金融大数据平台总体技术要求》，规定了面向金融业的大数据平台总体技术要求的框架结构、功能技术要求、非功能技术要求以及接口技术要求。

•医疗健康：监管驱动、高壁垒、体系化。药监局（NMPA）已发布多项 YY/T 系列行标，构成从数据、质量到评价的完整框架。比较具有代表性的 YY/T 1833.1-2022《人工智能医疗器械质量要求和评价第 1 部分：术语》和 YY/T 1833.2-2022《人工智能医疗器械质量要求和评价第 2 部分：数据集通用要求》，构建了 AI 医疗器械的标准化基础。从术语、数据、到全生命周期管理，为 AI 辅助诊断、影像分析等软件（SaMD）的注册审批提供了明确路径，具有全球影响力。

•公共安防：人工智能技术应用较早的行业，强需求、重治理。在公共安全、城市管理、司法等领域已发布多项行标，侧重视频分析、生物识别、舆情分析等。比较具有代表性的 GA/T 1400.x《公安视频图像信息应用系统》系列标准，明确规定了视频图像结构化描述（即用 AI 算法从视频中提取出人、车、物等特征信息）的数据格式、交换协议和应用接口，解决“互联互通”的根本问题。GA/T 1324-2017《安全防范人脸识别应用静态人脸图像采集规范》主要针对静态人脸图像的采集过程，旨在为公共安全、监控系统及各类需要人脸识别功能的场景提供统一的技术指导和要求，以确保采集到的人脸图像质量满足后续人脸识别算法处理的需求，提高识别准确率和安全性。是人脸识别技术在安防领域规模化、规范化应用的基石。

3.团标促创新

团体标准凭借快速响应产业需求的优势，成为实际应用主力。截至 2025 年 7 月，中国人工智能产业发展联盟（AIIA）已发布多项人工智能团体标准，覆盖智能体、大模型服务、软件智能化等前沿领域。

典型应用场景：

•智能体应用：AIIA/T 0219-2025《面向软件工程智能体的技术和应用要求第 1 部分：开发智能体》被中国工商银行、百度、腾讯、阿里、华为等 20 余家头部企业共同参编并应用于软件开发智能化转型。

- 大模型服务：AIIA/T 0221-2025《MaaS 模型即服务技术与应用要求》第 7 部分：大模型 API 服务性。

- 软件智能化：AIIA/T 0223-2025《软件智能化成熟度模型通用能力要求》为用友网络等企业提供软件产品智能化成熟度分级参考。

4.地标显特色

地方标准聚焦区域产业特色，作为补充性支撑。2024—2025 年，上海、广东、北京等省市陆续发布多项人工智能及相关领域地方标准，为区域产业发展提供标准化指引。

典型应用：

- 上海：DB31/T 1490-2024《人工智能标准化工作导则》为上海市人工智能组织开展标准化活动提供基本原则、工作规划、组织建设等指导性技术内容，推动长三角区域标准协同发展。

- 广东：DB44/T 2739-2025《智能网联汽车基础地图数据规范》适用于 2 级及以上驾驶自动化系统地图数据的采集与建模，涵盖乘用车、商用车及城市智慧交通系统地图建设等应用场景。

10.2 中国人工智能标准分类

人工智能标准按照技术领域和应用场景，可分为基础标准、硬件标准、软件标准、算法标准、安全标准和测试标准六大类别。以下分别介绍各类标准的核心内容和代表性标准。

1.基础标准类

规范人工智能基础术语、架构、评估、数据采集、标注、质量、治理等框架性要求，是整个标准体系的基石。

例如：

- GB/T 41867-2022《信息技术人工智能术语》
- GB/T 45288.1/2/3-2025《人工智能大模型系列标准》（通用要求、评测方法、成熟度评估）
- GB 45438-2025《网络安全技术人工智能生成合成内容标识方法》
- GB/T 46802-2025《人工智能社会实验实施指南》

2.硬件标准类

硬件标准类规范 AI 硬件设备性能、接口、架构要求等，支撑 AI 算力基础设施建设。

可参考标准：

- GB/T 45087-2024《AI 服务器性能测试方法》——覆盖 AI 服务器性能评测方法，已应用于华为、浪潮等企业 AI 服务器测试。
- GB/T 45280-2025《异构加速器统一接口》——规范 GPU、TPU、NPU 等异构加速器的统一接口，提升国产 AI 生态兼容性。

•GB/T 44463-2024《互联网数据中心 IDC 总体技术要求》——规定 IDC 在高技术、高能效、高安全、高算力五个维度的技术要求。

3.软件标准类

规范 AI 系统软件、开发框架、模型等要求，以及 AI 系统间的数据接口、通信协议等互操作要求。

可参考标准：

•GB/T 42382.1-2023《信息技术神经网络表示与模型压缩》——规范卷积神经网络的架构、参数、训练及压缩方法表示。

•GB/T 45288.1-2025《人工智能大模型第 1 部分：通用要求》——我国首个针对人工智能大模型技术的国家级基础标准。

•T/SZAI 002-2023《“人工智能+”制造领域 OHT 天车与其调度系统的数据接口规范》

4.算法标准类

规范大模型、算法鲁棒性、可解释性等核心算法要求，确保 AI 系统的可靠性和可理解性。

可参考标准：

•GB/T 46351-2025 《人工智能 多算法管理技术要求》

•GB/T 42564-2023《信息安全技术边缘计算安全技术要求》

5.安全标准类

规范人工智能全生命周期安全要求与治理框架，以及 AI 在各行业应用场景的服务要求，规范 AI 伦理要求及可持续发展框架。

可参考标准：

•GB/T 46800-2025《生成式人工智能技术应用社会影响评估指南》

•GB/T 46347-2025 《人工智能 风险管理能力评估》

•GB/T 46347-2025 《网络安全技术人工智能计算平台安全框架》

6.测试标准类

规范人工智能系统、模型的测试评估方法，为 AI 产品质量把关提供技术依据。

可参考标准：

•GB/T 45087-2024《人工智能服务器系统性能测试方法》

•GB/T 45288.2-2025《人工智能大模型第 2 部分：评测指标与方法》

10.3 人工智能标准的编写与实施规范

1.标准定位与适用范围

本标准体系适用于人工智能全生命周期的技术研发、系统部署与行业应用。根据 2026 年的产业现状，标准的适用范围已从单一的算法模型扩展至生成式 AI（AIGC）、嵌入式 AI 及特定垂直行业系统。

- 通用大模型：适用 GB/T 45288.1-2025 等标准，规范模型训练、推理与部署。
- 行业专用系统：涵盖化工、汽车、金融等高风险领域，强调与业务场景的深度结合。

2.规范性引用文件

本标准体系主要依据以下层级的核心文件构建：

- 国家标准（基础框架）：GB/T 41867-2022《信息技术人工智能术语》GB/T 45288.1/2/3-2025；《人工智能大模型系列标准》（通用要求、评测方法、成熟度评估）；GB/T 46802-2025《人工智能社会实验实施指南》。

3.核心术语与定义

- 基础模型(Foundation Model)：在大规模数据集上训练的，具有广泛下游任务适应能力的模型。
- 幻觉(Hallucination)：生成式人工智能系统产生与事实不符或逻辑错误的输出内容的现象，需依据 GB/T 45288.2-2025 进行规制。
- 异构加速器：依据 GB/T 45280-2025，指用于 AI 计算的 GPU、TPU、NPU 等不同架构的硬件单元。

4.治理与管理体系

（1）AI 战略框架与组织管理

企业应建立符合 GB/T 45081-2024《人工智能管理体系》的治理架构：

战略对齐：AI 部署目标需与业务 KPI（如化工企业的安全生产、汽车企业的良品率）强关联。

成熟度评估：依据 GB/T 45907-2025《人工智能服务能力成熟度评估》定期诊断组织能力。

（2）角色与责任

明确各层级人员在 AI 全生命周期中的职责，参考 SJ/T 11805—2022《人工智能从业人员能力要求》：

首席人工智能官(CAIO)：负责整体算力规划与伦理合规，明确企业遵从区域产业应用标准。

算法工程师：负责模型选型与优化，遵从机器学习框架规范。

合规安全官：负责数据隐私与算法备案，建立企业安全防线。

（3）资源管理：算力与基础设施

基础设施是 AI 运行的底座，需满足高能效与高可用性要求：

算力中心设计：必须符合 GB/T 44463-2024（IDC 技术要求）。

服务器性能：依据 GB/T 45087-2024 进行性能测试，确保异构计算资源的协同调度（GB/T 45401-2025）。

5.数据处理与管理

（1）数据质量标准

数据质量是模型性能的决定性因素。在安防与工业领域，必须严格执行 T/CQAE 18003—2025《人工智能数据集质量评价指标》。

评价维度：准确性、完整性、一致性、代表性。

应用实效：该标准已成为研发运营一体化（MLOps）中数据验收的通用准则，直接决定算法的误报率与漏报率。

（2）数据采集与源头合规

不同场景对采集技术有明确的物理参数要求。例如：

人脸图像采集：行业标准 GA/T 1781-2021 强制规定了安防场景下分辨率、光照均匀性及人脸区域大小，是公安实战的“源头质量”保障。

自动驾驶地图：需符合 T/CSAE 157-2020 智能网联汽车数据规范。

（3）数据标注与处理规范

为消除人为偏差，标注过程需标准化：

通用规范：遵循 GB/T 42755-2023《面向机器学习的数据标注规程》。

安全规范：对于生成式 AI 数据，需符合 GB 45438-2025《网络安全技术人工智能生成合成内容标识方法》。

6.技术要求与开发生命周期

（1）模型设计与开发框架

开发阶段需确保架构的兼容性与先进性：

大模型架构：依据 T/CESA 1438-2025 进行系统架构设计。

跨平台适配：采用 GB/T 45079-2024 规范，确保深度学习框架在华为、浪潮等不同硬件平台上的适配性。

模型压缩：参考 GB/T 42382.1-2023，规范神经网络的表示与压缩，以便在边缘设备（如移动终端 YD/T 4994-2024）上高效运行。

（2）训练与测试(MLOps)

引入 MLOps 理念，实现研发运营一体化：

能力成熟度：依据 GB/T 42560-2023 评估模型开发运营流程。

基准测试：使用 GB/T 42560-2023 进行芯片性能测试，以及 YD/T 3944-2021 进行基准评估。

（3）部署与集成

云边协同：遵循 YD/T 4478-2023，实现云端大模型与边缘侧推理的协同。

系统集成：装备制造行业需符合 GB/T 42405.1-2023 集成规范，确保 AI 组件与工业控制系统的无缝连接。

7.安全、鲁棒性与可靠性

（1）生成式 AI 专项安全(GenAI Security)

针对 AIGC 的特殊风险，GB/T 43441.1-2023（及 GB/T 45654-2025）是目前国内最核心的强制性底线。

核心定位：国内首个针对生成式 AI 服务的安全标准，直接支撑监管落地。

规制内容：涵盖语料安全、模型生成内容的价值观对齐，以及对“提示词注入”等攻击的防御要求。

标识方法：必须依据 GB/T 45674-2025 对生成内容进行显性或隐性标识，防止深度伪造欺诈。

（2）鲁棒性与系统稳定性

对抗测试：执行 GB/T 45288.2-2025，确保模型在噪声干扰或恶意攻击下的稳定性。

功能安全：服务机器人需符合 GB/T 41773-2022 安全性要求，防止物理伤害。

（3）网络与数据安全

通用要求：遵循 GB/T 45654-2025 信息安全通用要求。

平台安全：开发平台需符合 GB/T 45958-2025 的安全能力要求。

8.伦理、公平与社会责任

（1）伦理风险治理

AI 应用必须“科技向善”，建立伦理审查机制：

风险评估：依据 GB/T 46801-2025 建立覆盖全生命周期的伦理审查。

特定群体保护：如 DB43/T 2562-2023 手语播报系统，体现技术无障碍的社会责任。

（2）可解释性与透明度

解决“黑盒”问题（特别是金融与医疗领域）：

标准依据：可遵循 GB/T 46800-2025。明确人工智能的透明度。

情感计算：GB/T 40691-2021 规范了情感计算用户界面的模型，防止情感操纵风险。

9.评估与审计

建立第三方独立测试与内部自测结合的机制：

系统评估：依据 GB/T 46347-2025、GB/T 46346-2025、GB/T 45907-2025 进行人工智能体系进行评估。

大模型评测：采用 GB/T 45288.2-2025，部分工业类头部企业已将其作为模型采购的依据。

深度学习算法评估：基于 GB/T 45225—2025 对算法进行效能评估。

10.4 人工智能标准未来发展建议

基于对现有标准体系的分析，得出以下结论：

1.现状识别：当前呈“金字塔”结构，国标主要明确术语和安全底线，团标主导大模型等前沿应用。

2.问题洞察：大模型迭代速度（周/月级）远超国标制定周期（年级），导致标准滞后；企业面临“标准孤岛”，不同团标间缺乏互认。

3.趋势推演：随着 AIGC 向 Agent（智能体）演进，标准将从“模型静态评估”转向“行为动态监管”。

4.安全内卷：生成式 AI 的幻觉与注入攻击将导致安全标准无限细分，需要自动化工具而非文档来执行标准。

5.国际博弈：ISO/IEC 国际标准与国内标准需进一步互认，以支持中国 AI 企业出海。

关于 AI 标准未来的演进路径与未来发展建议如下：

1.建立“敏捷标准化”机制(Agile Standardization)

目前的团标虽然灵活，但仍有滞后。

建议：推行“滚动版本制”标准。对于大模型和生成式 AI，不再发布单一固定版本标准，而是发布如 V1.0,V1.1 的迭代版本，允许技术指标每季度更新一次，以匹配模型参数数量的指数级增长。

2.推行“合规即代码”(Compliance as Code)

鉴于 AI 系统的复杂性，依靠人工核对文档（如 GB/T 43441.1）已难以满足监管效率。

建议：将关键的技术标准（如数据质量 T/CQAE 18003-2025《人工智能数据集质量评价指标》）转化为自动化测试脚本或 API 接口。监管机构或第三方机构提供“标准容器”，企业上传模型通过自动化测试即视为合规。

3.构建“合成数据”标准体系

随着高质量人类语料枯竭，合成数据将成为训练主流。

建议：紧急启动关于合成数据（Synthetic Data）的标准制定。重点规范合成数据的真实性校验、水印添加（防污染）及偏见控制，防止 AI 在自我生成的“数据回声室”中退化。

4.分级分类的动态治理

目前的标准多为通用型，未来需更精细化。

建议：参考自动驾驶 L1-L5 分级，建立生成式 AI 的分级标准。

L1-L2（辅助型）：只需满足基本内容安全标准。

L3-L4（决策型）：如医疗诊断、金融交易 AI，需强制执行可解释性与高强度鲁棒性测试，并引入“人机回环”干预标准。

5.区域互认与统一大市场

解决地标（DB）与团标（TB）造成的割裂。

建议：在长三角（上海、浙江、江苏）、大湾区等重点区域，建立《标准互认清单》。例如，上海的算法备案标准在通过互认后，可直接被广东市场监管部门采信，降低企业跨区域落地的合规成本。

10.5 本章小结

2026 年的 AI 标准体系已初具规模，但面对技术的指数级跃迁，标准本身也需要“智能化”。未来的

核心在于如何让标准从“文档约束”变为“技术内嵌”，在保障安全底线的同时，最大程度释放产业创新的活力。AI 标准体系的建设是值得研究的方向。

本章参考资料

[1]国家市场监督管理总局,国家标准化管理委员会.国家标准全文公开系统[EB/OL].<https://openstd.samr.gov.cn/bzgk/gb/index>

[2]国家标准化管理委员会.国家标准化管理委员会官方网站[EB/OL].<https://www.sac.gov.cn/>

[3]全国团体标准信息平台.全国团体标准信息平台[EB/OL].<https://www.ttbz.org.cn/>

[4]中国人工智能产业发展联盟.中国人工智能产业发展联盟官方网站[EB/OL].<https://aiaaorg.cn/>

后记

合上这部蓝皮书的终稿，我们的心情是复杂的——既有完成一项系统性研究的欣慰，也有对人工智能发展速度的深切敬畏。人工智能不能被当作仅是改变某个行业或某项工作的技术，它正在重新定义“什么是可能的”。以编撰者的身份亲历这场变革，既是幸运，也是责任。

在编撰过程中，我们反复被一个现象所触动：几乎每一章的内容，在落笔之后都需要补充新的进展。技术篇刚梳理完 Scaling Law 的边际递减，后训练技术革命又开辟了新的能力增长曲线；产业篇的案例分析尚在审校，企业端已涌现出更深层的 AI 原生实践；治理篇的政策梳理刚截稿，新的监管框架便接踵而至。大模型的能力边界在不断拓展，产业应用的落地速度超出预期，治理框架也在加速迭代。这本身就是我们在序言中所说的“系统重构”最生动的注脚——变革已不是局部的、渐进的，而是全域的、加速的。

回望全书，有三个核心发现值得在此特别展开。

第一，AI 的竞争焦点正在发生结构性转移。人工智能竞争的主战场，已从模型参数的竞赛，转向系统工程能力的全面比拼。这一转移体现在多个层面：

其一，竞争维度从单一模型扩展到全栈系统。算力、数据、能源、芯片、应用生态——决定胜负的不再是某一项技术的局部突破，而是从物理层到应用层的垂直整合能力。正如本书技术篇所揭示的：“四力”模型将算力、存力、运力、电力纳入统一分析框架；“算力溢出效应”揭示了 GPU 算力需求如何向存储芯片、互联芯片层层传导；能源已构成 AI 发展的终极物理约束——三者共同勾勒出 AI 竞争新版图的完整轮廓。单点领先已不足以确立优势，系统短板才是真正的战略风险。

其二，开源与闭源的路线分化正在重塑产业格局。对中美大模型竞争的深入分析表明，闭源阵营凭借数据飞轮和商业化闭环构筑护城河，开源生态则以社区协作和快速迭代形成另一种竞争力。两条路线并非简单的替代关系，而是在不同场景下各有其不可替代的价值——这种双轨并行的态势，在人类技术史上并不多见。

其三，数据正在从被动资源升级为战略性可控变量。“AI-Ready 数据”概念揭示了一个关键转向：数据竞争的焦点已从“谁拥有更多数据”转向“谁拥有更高质量、更可治理的数据”。当推理 Token 成本较 2024 年下降约 80%，数据质量对模型效能的杠杆效应被急剧放大。

其四，中国在应用层和产业规模上正在形成差异化优势。全球最大的制造业体量、最丰富的应用场景、最活跃的开发者生态，使中国有条件走出一条不同于硅谷的 AI 发展路径。从 AI 编程工具到长流程智能体的快速普及，正在为攀升至更高成熟度积蓄动能。技术的制高点固然重要，但最终胜出的往往是那个能将技术最深地嵌入真实世界的一方。

第二，AI 赋能产业的跃迁已经开始。AI 赋能产业正在从概念验证阶段跨入规模化应用的关键窗口期，这个跨越比许多人预想的来得更快，也更具颠覆性。

其一，数据驱动正在从口号变为实践。制造业大模型应用比例在一年间从 9.6% 跃升至 47.5%，这不仅是数字的飞跃，更意味着数据作为生产要素的角色正在从“锦上添花”变为“不可或缺”。蔚来的案例分析表明，当 AI 贯穿从设计、制造到服务的全流程，数据驱动的车辆迭代已成为整车企业的核心竞争力。那些率先完成数据治理和数据资产化的企业，正在建立起难以复制的竞争壁垒。数据不再是 AI 的燃料，而是 AI 的血液——燃料烧完就没有了，而血液却在循环中生生不息。

其二，AI 正以空前广度赋能千行百业，其势能令人瞩目。本书针对六大重点产业方向和若干领域，呈现了 AI 渗透的广度：华为数字能源将 AI 深度融入构网型储能电站全生命周期智能运营，联影智能的医疗 AI 解决方案已在全国超过 4000 家医疗机构落地，鸿之微和晶泰科技等正在用 AI 改写新材料研发的范式，金山办公则以 AI 原生理念重塑了数亿用户的工作方式。从制造到能源、健康到材料、空间到信息科技，AI 正在证明它不是某个行业的专属工具，而是所有行业的公共基础设施。

其三，AI 的产业价值已从“降本增效”的单点突破，走向“重构逻辑”的系统变革。在制造领域，蔚来的先进制造工厂、联宝科技的柔性制造、工业富联的智慧“双碳”管理平台，展示了从微观工序到宏观生态的全面渗透。在能源领域，华为数字能源的储能电站智能运营、比亚迪储能的光伏与储能协同调度、运达能源的风电全生命周期智能运维，印证了 AI 与复杂巨系统的深度耦合。在健康领域，联影智能的医疗大模型与智能体、复星健康的智能健康管理、强脑科技的脑机接口，打开了精准医疗的新空间。在材料、空间、信息等前沿方向，鸿之微的鸿元 AI 大模型、御风未来的 eVTOL 智能飞控、图灵量子的“量子计算+人工智能”，则指向了更远的产业地平线……这些案例共同传递出一个信号：AI 正在从产业的“外挂工具”升级为“内生基础设施”，从解决局部问题的技术手段升级为重构产业底层逻辑的系统力量。

其四，AI 赋能产业的路径正在从“标杆示范”走向“生态扩散”。我们注意到一个积极趋势：AI 赋能正在从龙头企业的标杆工程，向中小企业和产业链上下游扩散。工业富联的跨产业链平台为中小制造企业提供了智能化入口，库萨科技的智能云控平台让城市服务的智能化不再依赖巨额基建投入，联适技术的农机自动驾驶系统则将精准作业能力输送到了中国最广袤的乡村。这种从点到面、从上到下的扩散，正是产业智能化从“盆景”变为“风景”的关键标志。

第三，治理创新的紧迫性前所未有。技术变革的速度已经超过制度演进的速度，这种“制度落差”正在成为全球性挑战。

其一，“边发展边治理”已成为唯一现实可行的范式。等技术完全成熟后再建立规则，意味着在最关键的窗口期放弃塑造的权利；而以过度监管遏制创新，则可能在全球竞争中丧失战略主动。国内 AI 政策的三阶段演进——从战略启航（2015-2020），到法治筑基（2021-2024），再到融合深化（2025 年起）——体现的正是在发展与安全之间动态校准的治理智慧。中国成为全球首个对生成式 AI 出台专项监管规定的国家，这一事实本身就说明，“边发展边治理”不是口号，而是正在发生的制度实践。

其二，全球 AI 治理正呈现碎片化态势，中国需要走出自己的平衡之路。欧盟以《人工智能法案》确立风险分级监管，美国在行政令与国会立法之间摇摆，英国走“亲创新”路线，日本采取“软法”策略——第 8 章的国际比较分析清晰地呈现了各主要经济体治理取向的显著差异。在这种格局下，“1+N+X”政策矩阵和四维协同治理架构，为中国在借鉴国际经验与坚持自身发展逻辑之间找到了平衡点。既不能闭门造车，也不能照搬照抄——治理创新与技术创新一样，需要走自己的路。

其三，风险防控正在从单点应对走向系统治理。五维风险图谱——技术自主可控、数据安全、国际环境、社会伦理、治理机制适配——揭示了一个重要现实：AI 风险不是孤立的技术问题，而是交织着地缘政治、产业安全与社会公平的复合挑战。从芯片禁运引发的供应链安全风险，到大模型中文词表污染率高达 46.6%所暴露的文化安全隐患，再到算法歧视和就业冲击等社会议题，三侧协同（开发侧、部署侧、应用侧）的治理路径正是对这种复合风险的系统回应。

其四，标准制定正在成为国际话语权争夺的关键战场。AI 标准化不仅是技术规范的制定，更是规则制定权和产业主导权的争夺。“合规即代码”（Compliance as Code）的创新理念，将技术治理与制度治理深度融合，代表了一种值得关注的方向。治理的本质不是给技术套上枷锁，而是为创新铺设轨道。

展望未来，我们有理由对中国 AI 的发展保持审慎乐观。以下四个关键方向值得持续关注：

第一，“智能体即服务（AaaS）”为核心的全新应用时代即将到来。随着 OpenClaw 等开源架构成为智能时代的“操作系统”，AI Agent 正加速从单一任务工具演变为具备多轮理解、任务拆解与自主执行能力的复杂系统。这种从“问答”到“干活”的逻辑转变，直接驱动了 Token 消耗量的指数级增长。Agent 不仅是人机交互的新范式，更是 AI 商业化的核心引力场。预计到 2033 年，全球 AI Agent 市场规模将突破 1800 亿美元，年均增速近 50%。

第二，具身智能正加速从“实验室原型”向“全产业链量产”质变。硬件红利与规模效应正成为核心驱动力：在人形机器人 BOM 成本中，驱动、电机、丝杠等硬件份额高达 55%，而随着宇树科技（Unitree）等头部企业实现年出货量超 5500 台的突破，单机成本正以年均 11% 的速度下行。具身智能正沿着三级路径演进：从宁德时代产线上的“工业专才”，到 AWE 展会上的“家电具身化”，再到单价降至万元级别的“大众消费陪伴”。这一进程标志着 AI 已彻底突破数字边界，开始重塑物理世界的劳动力结构。

第三，AI for Science 正从“科研辅助”蝶变为加速创新的“第五范式”。随着从蛋白质结构预测到新材料发现的跨越，AI 不再仅仅是数字工具，而是深度参与底层物理规律重构的生产力。这种范式革命正通过鸿之微、晶泰科技等先行者的实践，转化为真实的产业化潜力。更重要的是，算力正从“通用规模”转向“专用滴灌”——针对生物医药、气候模拟等特定领域定制的专用算力集群，正有效解决高端算力短缺与低端算力闲置的结构性错配。

第四，AI 治理的制度建设正成为塑造中国 AI 核心竞争力的“隐形底座”。“十五五”规划的战略布局、国际治理规则的博弈与协调、国内法治体系的完善——这些制度层面的力量，将与技术创新一起，共

同塑造中国 AI 的未来图景。

这四个方向勾勒出一个从微观技术到宏观战略的动态逻辑闭环：Tokens 作为“核心生产资料”，正通过 Agent 这一“交互范式”释放动能；具身智能则为其提供了物理世界的“实体承载”；而 AI for Science 与制度治理，则共同构筑起科研范式与战略保障的“稳固基石”。在 AI 领域日新月异的今天，这一闭环并非终点，而是一个正在不断自我重塑的进化框架。随着中国 AI 与真实世界的深度融合，更多超越想象边界的变量正待被逐一定义。

最后，衷心感谢所有参与本书编撰的专家学者、企业家、政策研究者和编辑团队，感谢所有合作企业提供的案例素材。本书的研究方法论强调“三个结合”：定量数据与定性判断相结合，国际比较与中国实践相结合，技术分析与制度研究相结合。蓝皮书的编写参考了国内外大量资料，难免出现疏漏之处，敬请读者批评指正。

一部蓝皮书的价值，不在于它记录了多少事实，而在于它能否帮助读者穿越信息的迷雾，看清前行的方向。我们期待，这本蓝皮书能成为您理解 AI 时代的一把钥匙——不是终点，而是起点。

编委会
2026.4