



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

2025

“人工智能+”行业发展蓝皮书

“ Artificial Intelligence + ” Industry Development

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

2025 年 3 月

主编：史占中、陈晓荣

校阅：周辰、张晔、窦玉梅、孙崇理、刘香港

整合排版：刘香港

引言：全球化视野看人工智能（史占中、王计登、窦民）

第1章 AI大模型：实现通用智能的重要桥梁（窦民）

第2章 AI数据：驱动智能时代的核心引擎（李薇、商富凯、管航）

第3章 AI算力：支撑智能进化的底座（王计登、窦玉梅、刘香港）

第4章 AI赋能相关产业领域的典型应用场景

4.1 AI赋能未来制造业的应用场景（章明根）

4.2 AI赋能未来信息产业应用场景（窦玉梅、王计登）

4.3 AI赋能未来材料产业应用场景（张弛、刘香港）

4.4 AI赋能未来能源产业应用场景（董密尔）

4.5 AI赋能未来健康产业应用场景（许蕾、窦玉梅）

4.6 AI赋能未来空间产业应用场景（张晔、刘香港）

4.7 AI赋能未来金融服务业创新实践（黄志锋）

4.8 本章结语（刘香港、张晔）

第5章 AI赋能行业应用案例

5.1 AI赋能制造行业应用案例（章明根、刘子昱）

5.2 AI赋能信息科技行业应用案例（王计登、窦玉梅、刘香港）

5.3 AI赋能材料行业应用案例（刘香港）

5.4 AI赋能能源行业应用案例（董密尔）

5.5 AI赋能健康医疗行业应用案例（许蕾、窦玉梅）

5.6 AI赋能空间行业应用案例（刘香港、张晔）

5.7 AI赋能金融行业应用案例（黄志锋）

第6章 国内外人工智能产业政策及趋势分析（孙玉贝、肖耀、丁紫玉、孙崇理、周辰）

第7章 人工智能发展伦理和治理框架

7.1 人机关系和伦理问题（屈杰）

7.2 AI数据和隐私保护（林玉岚）

7.3 AI内容安全和虚假信息（辛咏琪）

7.4 AI知识产权和责任归属（周辰）

全球正在见证一场由人工智能驱动的“技术—经济”范式重构。人工智能作为拥有自我进化能力的通用赋能技术，正以指数级创新速度突破“技术-经济-社会”的三重边界，其影响力已超越传统 ICT 技术的历史坐标系，展现出重塑人类文明进程的史诗级能量。人工智能的技术路径和产业生态尚处于不断演变的阶段，未来可能会出现新的主导技术路线，赋能行业发展也将带来更多可能性。

人工智能与前沿技术产业领域，如先进制造、量子计算、生命科学、新材料、新能源等加速融合，将催生出更多新的科技和产业赛道。本报告旨在梳理人工智能发展现状与趋势，并通过对相关产业领域主要应用场景与典型案例的跟踪研究，深入剖析 AI 在行业深度应用中面临的问题与挑战，希望为政府主管部门和相关行业企业提供决策参考，共同探讨 AI 如何助推传统产业转型升级，引领未来产业创新发展。

目录

引言：全球化视野看人工智能	1
一、人工智能的起源与范式演进	1
二、技术突破与全球竞争格局	2
三、国际合作与治理框架	2
四、社会影响与伦理挑战	2
五、未来方向：可持续与普惠发展	3
本章参考文献	3
第1章 AI大模型：实现通用智能的重要桥梁	4
1.1 国内外大模型技术发展态势	4
1.2 浅析大模型的核心技术	5
1.3 大模型技术发展趋势展望	10
1.4 结语	11
本章参考文献	11
第2章 AI数据：驱动智能时代的核心引擎	15
2.1 模型突破推动数据需求升级	15
2.2 驱动因素推动数据生产革新	17
2.3 技术演进推动数据生态重构	19
2.4 全球竞争推动数据战略迭代	20
2.5 结语	22
本章参考文献	22
第3章 AI算力：支撑智能进化的底座	25
3.1 上游：AI算力硬件基础层	25
3.2 中游：算力软件服务层	29
3.3 下游：算力应用场景层	30
3.4 算力发展现状与挑战分析	32
3.5 算力发展趋势与方向	35
3.6 结语	37
本章参考文献	38
第4章 AI赋能相关产业领域的典型应用场景	40
4.1 AI赋能未来制造业的应用场景	40
4.2 AI赋能未来信息产业应用场景	43
4.3 AI赋能未来材料产业应用场景	46
4.4 AI赋能未来能源产业应用场景	50
4.5 AI赋能未来健康产业应用场景	57
4.6 AI赋能未来空间产业应用场景	61
4.7 AI赋能未来金融服务业创新实践	68
4.8 结语	74
本章参考文献	76
第5章 AI赋能行业应用案例	80
5.1 AI赋能制造业应用案例	80
5.2 AI赋能信息科技行业应用案例	82
5.3 AI赋能材料行业应用案例	84

5.4 AI 赋能能源行业应用案例	88
5.5 AI 赋能健康医疗行业应用案例	90
5.6 AI 赋能空间行业应用案例	94
5.7 AI 赋能金融行业应用案例	97
本章参考文献	100
第 6 章 国内外人工智能产业政策及趋势分析	103
6.1 人工智能治理困境与监管沙盒	103
6.2 国内人工智能产业政策及环境分析	105
6.3 国际人工智能整体产业政策	108
6.4 未来政策趋势	109
6.5 结语	110
本章参考文献	111
第 7 章 人工智能治理与伦理	112
7.1 人机关系和伦理问题	112
7.2 AI 数据和隐私保护	117
7.3 AI 内容安全和虚假信息	122
7.4 AI 知识产权和责任归属	126
7.5 结语	130
本章参考文献	131

引言：全球化视野看人工智能

人工智能（AI）的全球化发展正以前所未有的速度重塑技术、经济与社会结构。这一进程不仅体现为技术突破的加速、AI+产业变革，更涉及国际合作、治理框架的构建以及伦理风险的平衡。以下从多维度解析 AI 全球化发展的关键趋势与挑战。

一、人工智能的起源与范式演进

人工智能的学科起源可追溯至 1956 年的达特茅斯会议^[1]，约翰·麦卡锡（John McCarthy）、克劳德·香农（Claude Shannon）、马文·明斯基（Marvin Minsky）等著名学者首次明确提出了“人工智能”的概念，通过计算机模拟人类智能，实现语言理解、抽象推理和自主问题解决。

随后，人工智能的技术范式先后经历了从规则、统计再到深度学习驱动的三个阶段发展历程。20 世纪 60 至 80 年代，规则驱动范式主导了 AI 发展，“符号主义”主张根据人类专家知识的规则库进行推理和决策。90 年代后，统计学习范式崛起，贝叶斯网络、支持向量机等算法通过数据驱动模式优化参数，推动垃圾邮件过滤、搜索引擎排序等应用落地，标志着 AI 从“人工定义规则”向“基于数据进行学习和推理”的范式转移。

公众广为熟知的是现在所处的 AI 发展第三阶段——深度学习。2012 年成为 AI 发展的重要分水岭。杰弗里·辛顿（Geoffrey Hinton）凭借深度卷积神经网络 AlexNet^[2]在李飞飞组织的 ImageNet 竞赛中以压倒性优势击败传统方法，标志着人工智能进入了深度学习时代。这一突破背后是神经网络、数据和算力三大要素的融合，卷积神经网络通过自动学习特征，其效果超越了手工设计的特征，海量数据为算法提供了文本、图像、语音等多模态燃料，大规模并行计算使得训练数十亿参数模型成为可能，重塑了 AI 研究的范式。2017 年 Transformer 架构^[3]的提出更是革命性里程碑——自注意力机制不仅解决了长序列建模难题，还催生了 GPT、BERT 等大语言模型，使自然语言理解的准确率超越人类基线。至此，深度学习完成了从实验室理论到产业基石的华丽转身，驱动聊天机器人、智能客服、医疗影像分析、自动驾驶、具身智能等场景的规模化落地。

纵览过去七十年间人工智能的发展历程，技术演进始终贯穿着符号主义、连接主义与行为主义^[4]的哲学张力。符号主义追求具备可解释性的规则推理；连接主义主张通过模拟人脑神经元的连接方式来实现人工智能；行为主义则强调通过试错与交互来优化行为。与此同时，我们可以清晰地看到，人工智能的重大突破往往源于未知前沿的探索精神、基础理论（如 Transformer 的数学建模）与工程实践（如分布式训练框架）的协同共振，众多胸怀大志之人正在通往人工智能的道路上全力书写下一个辉煌篇章。

全球人工智能产业发展历程如下图：

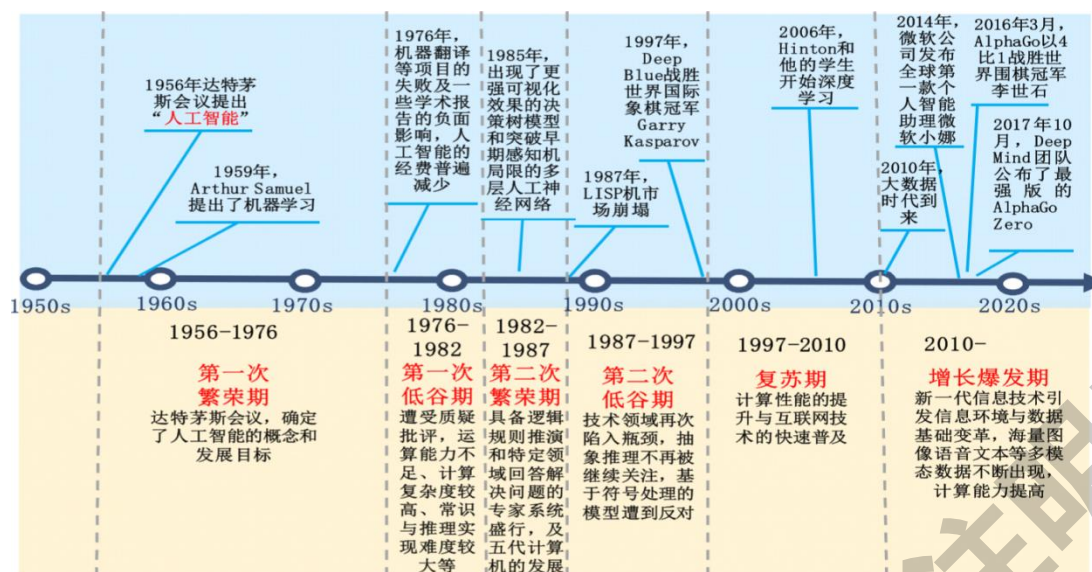


图1 全球人工智能发展脉络

二、技术突破与全球竞争格局

全球 AI 技术呈现多极化竞争态势，美国、中国、欧洲等国家和地区通过大型模型研发与政策扶持争夺主导权。如，美国企业 OpenAI、谷歌和 xAI 公司持续推出 GPT、Gemini 及 Grok 等系列模型，提升多模态能力与计算效率；中国企业月之暗面的 Kimi、字节的豆包、腾讯的混元以及阿里的通义系列等模型表现出显著竞争力，尤其是深度求索公司 DeepSeek 发布的开源模型以高性价比引发关注；法国 Mistral 公司推出高速生成的开源助手 Le Chat 成为“欧洲之光”。与此同时，各国政策支持力度持续加大，如美国“星际之门”项目计划投资 5000 亿美元建设 AI 基础设施，欧盟则通过“投资人工智能”倡议调动 2000 亿欧元推动超级工厂建设。

三、国际合作与治理框架

技术发展的全球化特征要求超越地缘政治的合作。2025 年巴黎人工智能行动峰会签署《关于发展包容、可持续的人工智能造福人类与地球的声明》，强调开放、透明、安全等原则，呼吁建立全球治理体系^[5]。中国积极参与国际对话，推动《全球人工智能治理倡议》和联合国相关决议，倡导技术普惠与风险共担。如清华大学人工智能国际治理研究院提出，需通过国际标准制定、数据共享和技术转移缩小“AI 鸿沟”，避免技术垄断加剧全球不平等。

四、社会影响与伦理挑战

AI 的普及在提升生产效率（如医疗诊断、农业管理）的同时，也会引发结构性风险。如，生成式 AI 可能加剧就业市场的两极分化，高收入国家更易受益于技术红利，而发展中国家因基础设施滞后面临边缘化风险。伦理问题同样突出，包括算法偏见、隐私侵犯及超级智能失控的可能性。国际学界呼吁加强 AI 伦理研究，将人类价值观嵌入技术设计，并通过跨学科合作（如法律、社会学）构建适应性治理框架。

五、未来方向：可持续与普惠发展

AI 的可持续发展需兼顾能源效率与社会公平。当前大模型训练的高能耗问题（如 ChatGPT 单次训练耗电量达家庭数月用量）促使业界探索绿色计算与低资源算法。此外，普惠应用成为关键议题，如中国通过 AI 赋能基层医疗和传统产业数字化转型，展示了技术下沉的潜力。WTO 报告预测，若全球均衡应用 AI，2040 年前贸易增速可翻倍，但需通过国际合作消除数据流动壁垒与本地化政策限制。

本章参考文献

- [1] 国务院发展研究中心国际技术经济研究所，中国电子学会，智慧芽. 人工智能全球格局：未来趋势与中国位势[M]. 北京：中国人民大学出版社, 2019.
- [2] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25.
- [3] Vaswani A, Shazier N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- [4] Yann Le Cun. 科学之路：人，机器与未来[M] 李皓,马跃,译.北京：中信出版集团, 2021.
- [5] 中国、法国、印度、欧盟在内的 60 个国家和组织，《关于发展包容、可持续的人工智能造福人类与地球的声明》[EB/OL].
<https://my.ambafrance.org/Statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and>

第1章 AI 算法“大模型”：实现通用智能的重要桥梁

1.1 国内外大模型技术发展态势

当前大模型开辟了通用人工智能技术演进的新路线，正在深刻改变人类社会生产生活。大模型是当前科技发展的制高点，也是中美科技竞争的焦点。

OpenAI 公司作为本轮人工智能革命的领军者，其早期研发团队是 Transformer^[1]架构首创者和 Scaling Law^[2]的坚定信奉者及实践先驱。自 2022 年 11 月推出现象级产品 ChatGPT^[3]以来，该企业通过持续突破不断重塑行业认知，产品上线两个月用户破亿，展现通用人工智能的早期雏形；次年 4 月问世的 GPT-4^[4]在律师资格考试等专业测试中超越 90% 人类考生，9 月系统升级实现视觉与语音交互^[5]，奠定多模态认知基座，推动教育个性化与客服智能化深度变革。OpenAI 技术迭代在 2024 年迎来爆发期，2 月视频生成模型 Sora 突破动态语义理解瓶颈^[6]，开启 AI 内容创作新纪元；下半年连续推出 o 系列推理模型^[7]，通过分层认知架构实现复杂决策推理；2025 年初推出支持复杂研究任务的 Deep Research 功能，发布的 GPT-4.5^[8]更号称在认知维度实现飞跃，其万亿级参数模型不仅构建起迄今为止人类知识最完整的数字镜像，更在情感计算与共情交互领域取得突破。

除了 OpenAI，美国的 Anthropic、Google、Meta 以及 xAI 等企业都在大模型领域持续发力，逐渐形成了“OpenAI 领跑，Anthropic、Google 和 xAI 追赶，Meta 开源”的竞争态势。Anthropic 初创团队出自原 OpenAI，专注于安全和伦理。2023 年 3 月，Anthropic 发布第一版 Claude^[9]，宣布与 ChatGPT 展开直接竞争。2024 年 3 月，发布 Claude 3^[10] 增强数学和编程能力。6 月推出 Claude 3.5 Sonnet^[11] 平衡性能和能力，提升用户体验。2025 年 2 月推出 Claude 3.7 Sonnet^[12]，首度引入混合推理模型，用户可选择快速响应或逐步推理。Google 一直致力于 AI 领域的创新，2023 年 2 月发布 Bard^[13] 作为 ChatGPT 的竞争者。2023 年 11 月，Gemini 1.0^[14] 正式发布，接棒 Bard 成为继任者。2024 年 2 月推出 Gemini 1.5 Pro^[15]，采用 MoE 架构，扩展上下文长度。2025 年 2 月，推出 Gemini 2.0^[16]，上下文长度达到 200 万 token，多模态能力进一步强化。xAI 由 Elon Musk 创立，迅速成为 AI 领域的后起之秀。2024 年 2 月开源 Grok-1^[17]，促进社区合作和提高透明度。2024 年 8 月发布 Grok2^[18]，提高模型性能。2025 年 2 月向大众免费提供 Grok3^[19]，在思维推理上提升显著，并推出 DeepSearch 功能。在大模型方面，Meta 扮演的角色专注于开源 AI 模型，推动社区合作和创新。2023 年 2 月，开源 Llama^[20]，成为很多大模型研究的基座。2024 年 7 月，发布的 Llama 3.1 405B^[21] 是当时最大开源 AI 模型，能力媲美商业模型。12 月又开源 Llama 3.3^[22]，以 70B 参数量媲美 405B 性能。

中国 AI 大模型的发展也受到全球 AI 热潮的推动，2023 年是关键转折点，各大科技公司和初创企业开始推出自己的大语言模型，号称“百模大战”，主要参与者包括百度、阿里、字节、腾讯、商汤、华为，以及“AI 大模型六小虎”（零一万物、稀宇科技、月之暗面、阶跃星辰、百川智能、智谱华章）。2023 年 3 月，百度发布文心一言^[23]，成为中国首个 AI 聊天

机器人。8月阿里开源发布 Qwen 系列^[24]，成为早期开发者社区的重要资源，8月发布 Qwen-VL^[25]，可以感知文本图像。12月，百川开源 Baichuan2^[26]，数学和逻辑推理、复杂指令跟随能力提升显著。2024年竞争加速，3月月之暗面宣布 Kimi 支持 200 万 token 上下文^[27]，4月商汤发布日日新 5.0^[28]，9月阿里开源 Qwen2.5^[29]，支持多模态，最大模型参数量 72B。10月，智谱发布智能体 AutoGLM^[30]，支持手机、网页和电脑端操作。2025年1月 稀宇科技-Minimax 开源发布 MiniMax-Text-01 和 MiniMaxVL-01^[31]，支持 400 万 token 上下文窗口和多模态功能。随后，阿里发布 Qwen 2.5-Max^[32]和 Qwen 2.5 VL^[33]，在语言和多模态上体现了对标国外顶尖模型的能力。

与此同时，以上海人工智能实验室为代表的中国科研机构也在大模型领域展现了自己的优势。该实验室构建了书生通用大模型体系，涵盖多模态、语言和视频生成能力。2021年发布国内首个通用视觉大模型书生 1.0^[34]，2023年推出语言大模型浦语 InternLM^[35]，并率先实现开源和免费商用。2024年，开源浦语 2^[36]在 200K 长上下文中大海捞针表现卓越；图文大模型 InternVL^[37]以 1/3 参数量实现 ViT-22B 性能，升级版的 InternVL 1.5^[38]表现接近 GPT-4 Vision 水平。进一步，多模态大模型书生·万象 2.0^[39]首创新进式对齐训练，实现与 GPT-4o 和 Gemini 1.5 Pro 媲美的性能。浦语 3.0^[40]以 4T 数据训练出高性能模型，万象 2.5 成为首个 MMMU 得分超 70%的开源模型^[41]。在视频生成方面，2023年全球首发开源文生视频 DiT 模型 Latte^[42]，2024年推出筑梦 2.0 实现分钟级 4K 视频生成^[43]。

2025年初，AI 领域最炙手可热的现象级公司莫过于深度求索（DeepSeek）。它发布了两款具有里程碑意义的模型：DeepSeek-V3 混合专家（MoE）大语言模型^[44]和 DeepSeek-R1 推理模型^[45]。这一模型组合通过系统性的工程创新，显著提升了全球大模型研发的效能基准。据业界有关团队估计 DeepSeek-V3 在训练阶段仅消耗 278.8 万 H800 GPU 小时（折合 557.6 万美元），采用 2048 卡集群配置实现行业领先性能，该训练成本可能约为行业平均水平的 1/20~1/15，显示出 DeepSeek 在计算资源利用率上的重大突破。DeepSeek-R1 作为首个开源推理模型(Reasoning Model，即以逻辑和系统的方式进行思考，利用证据和过往经验来得出结论或作出抉择)，通过强化学习在复杂推理任务上达到了与 OpenAI o1 模型相当的性能，形成了“训练-推理”双引擎驱动的技术架构。

DeepSeek 的技术演进成功突破了传统的“算力军备竞赛”模式，通过开源模型架构和高效训练方法，推动了 AI 技术的普惠化，打破了大模型仅被 OpenAI、Meta 和 Google 等巨头垄断的“神话”，降低了行业准入门槛，对硅谷和华尔街，乃至全世界 AI 产业产生了显著冲击。随着国内科研团队在大模型领域的不断创新，全球 AI 竞争格局正在重构。

1.2 大模型的核心技术

近年来，以 Transformer 架构为代表的生成式预训练大模型正在重塑人工智能的技术版图。这些具备千亿级参数以上的智能系统不仅展现出惊人的语言理解与生成能力，更在跨模态交互、复杂推理等维度持续突破认知边界。大模型的技术体系庞大，发展更是日新月异，

本节将试图用比较通俗的语言来解构支撑大模型发展的核心技术体系：首先剖析 Transformer 架构的自注意力机制如何突破传统神经网络的序列建模瓶颈，奠定大模型的基础骨架；继而揭示“预训练-微调”范式与强化学习的协同作用，解读模型如何通过海量数据预训练获得通识能力，再经由领域适配实现专业化转型；重点探讨交互提示工程对模型认知对齐的关键价值，以及推理时间扩展技术如何突破模型固有思维框架；最后聚焦扩散模型等新兴技术如何推动生成质量跃升。这些技术模块共同构建起大模型的进化阶梯，使其逐步实现从数据驱动到知识内化、从通用应答到专业赋能的范式转变，正在重塑人机协作的智能新纪元。

1.2.1 大模型的基石：Transformer 架构

人工智能研究者的一个理想，就是想让机器像人一样理解语言、看懂图片。Transformer 架构是实现这个目标的关键技术之一。Transformer 架构是现代大模型的核心，首次由 Vaswani 等人在 2017 年提出^[1]。它通过多头注意力机制和位置编码处理序列数据，解决了传统循环神经网络（RNN）在并行处理和长距离依赖上的局限性。当前，Transformer 已经成为大模型的首选架构。

Transformer 架构的优点来自：

（1）并行处理能力：传统的神经网络（比如 RNN）像一条流水线，必须按顺序处理数据，速度很慢。而 Transformer 可以同时处理所有数据，就像多条流水线一起工作，大大提高了效率。

（2）长距离依赖：在一句话中，后面的词可能受到前面很远处的词的影响。Transformer 能很好地捕捉这种长距离的依赖关系，让模型更好地理解上下文。

（3）多头注意力机制：想象一下，你在阅读一篇文章，有些词语会特别吸引你的注意力。Transformer 的“多头注意力机制”就像多个“注意力过滤器”，让模型能够同时关注输入数据的不同部分，从而更好地理解其含义。

Transformer 目前在大模型中得到了非常广泛的应用。在语言模型上，Transformer 是构建大语言模型（LLM）的基础，比如 GPT 系列。这些模型可以生成文本、翻译语言、回答问题等等。而在多模态大模型中 Transformer 不仅能处理文字，还能处理图像、音频等多种类型的数据。如，视觉 Transformer（ViT）将图像分割成小块，然后像处理文字一样处理这些小块，从而实现图像分类、目标检测等任务。

1.2.2 模型的成长之路：预训练、微调和强化学习

训练大模型就像培养一个孩子，需要经历不同的阶段，学习各种知识和技能。通常情况下，训练分为三个主要阶段：预训练（Pre-training）、微调（Fine-tuning, FT）和强化学习（Reinforcement Learning, RL），后两者也常常被称为后训练阶段。

（1）预训练：打好基础

预训练是模型的初始阶段，目标是在大规模文本、图像等语料库上训练模型。预训练就

像让孩子读中小学，学习通用的知识和技能。在这个阶段，模型会在海量的数据上进行训练，学习语言、图像的模式、知识的结构等等。主要的预训练任务包括：

掩码语言建模（Masked Language Modeling, MLM）：就像做完形填空，模型需要预测被遮盖的词语，从而学习上下文关系。

因果语言建模：让模型续写故事，模型需要预测下一个词语，从而学习生成文本。

下一句预测（Next Sentence Prediction, NSP）：就像成语接龙，模型判断两个句子是否连续，增强句子关系理解。

在多模态模型预训练中，主流的预训练任务还包括图像-文本匹配、掩码多模态建模等。

图像-文本匹配：模型需要判断给定的图像和文字描述是否匹配，从而学习跨模态的关联。

掩码多模态建模：类似于掩码语言建模（MLM），但应用于不同模态，模型必须预测序列中缺失的部分，如文本、图像补丁或其他数据类型，类似看图说话或者根据文本画图的游戏。

（2）微调：培养特长

微调是将预训练模型适配到特定任务的过程，通常使用较小、任务特定的数据集，可以显著提升模型在下游任务如文本分类、问答和翻译上的表现。微调就像让孩子上大学，通过专业课学习特定的专业知识和技能。在这个阶段，模型会在特定的数据集上进行训练，以适应特定的任务。

通常情况下使用的技术是监督微调（Supervised Fine-tuning, SFT）。就像做练习题，模型会学习输入和输出之间的对应关系，从而学会生成特定的输出。如在聊天机器人开发中，SFT 可以帮助模型生成符合用户期望的回复。另一个技术是领域适配，为了让“通才”的大模型，成为某一个细分领域的专家，需要用专业数据改造通用模型，赋予他这个领域的特色能力。

（3）强化学习：从做题家到解题高手

就像学生通过大量做题积累经验，强化学习让大模型在模拟环境中不断尝试不同解题思路。每次尝试后，系统会给出评分（比如答案是否正确、解题步骤是否合理），模型根据这些反馈调整自己的思考方式。强化学习还可以通过“思维链”技术（Chain of Thought, CoT），让 AI 学会像人类一样拆解难题：先列已知条件，再分步骤推导，最后验证结果。这种训练让 OpenAI 的 o1 模型在国际数学竞赛中正确率从 13% 飙升至 83%。

强化学习也可以让输出更符合人类偏好，基于人类反馈的强化学习 (Reinforcement Learning from Human Feedback, RLHF) 是模型输出与人类偏好对齐的重要方法，让人类评价模型的输出，然后用这些反馈来改进模型。这就像让老师给学生打分，从而帮助学生进步。RLHF 解决了模型输出安全性和实用性的挑战，特别是在对话系统和内容生成中，显著提升了用户体验。

当预训练数据不够时，强化学习能让 AI 自己创造训练场景。如 AlphaGo 一般，AI 可以通过自我对弈生成新题目，像棋手“左右互搏”一样持续提升。在自动驾驶等实时场景中，强化学习让模型像老司机一样，根据路况动态调整策略，在动态、实时、交互中不断学习和进步。

1.2.3 让 AI 更懂人类：交互与提示技术

训练好的大语言模型就像一个刚入职的聪明实习生，虽然知识渊博但缺乏工作经验，需要通过清晰的“工作指导书”——提示词（Prompt）^[2]——才能高效完成任务。这种指导方式主要有三大模式：

零样本学习：好比直接问实习生“用小学生能听懂的话解释万有引力定律”，无需提供任何案例，模型仅凭对任务的理解就能生成答案。这种方式适用于常识类任务，类似人类基于已有知识快速响应新问题。

少样本学习：好比案例教学，就像给实习生看几个优秀方案模板，模型通过 2-5 个示例就能掌握任务规律。这种技术特别擅长处理格式固定的任务（如邮件撰写、数据表格生成、客服回复等）

思维链提示：面对数学、代码这类复杂问题，让模型像学生写草稿纸那样逐步推理，改善复杂任务表现，避免直接跳转到错误结论。

提示工程的核心技巧在于“结构化表达”，通常包括角色设定、任务分解、格式规范、纠错机制等几个要素。提示词技术灵活性高，特别在资源有限的场景中表现优异，可以通过精心设计提示词最大化模型性能。当前提示工程师已经成为 AI 时代的热门职业之一。

1.2.4 突破思维局限：推理时间扩展

诺贝尔经济学奖得主丹尼尔·卡尼曼在著作《思考快与慢》^[3]中提出人脑有两个思维模型：其中系统一基于直觉和本能，无需刻意思考，就能快速做出正确的决定；而系统二则基于理性思考，当遇到复杂问题的时候会理性地分析问题，花费更多的时间和努力再来做出决策。

传统的 AI 大模型（如 GPT-4）的推理过程更接近系统一——根据 next token prediction，依赖预训练形成的知识直觉快速生成答案，但随着任务复杂度提升（如数学证明、代码优化），缺乏对复杂问题的分步验证和纠错机制，导致逻辑断层和事实幻觉，而且单纯扩大模型参数规模的边际收益也在下降，这个背景下推理时间扩展（test-time scaling, TTS）^[4]通过赋予模型动态分配计算资源的能力，实现了从“直觉反应”到“深思熟虑”的范式转变。该技术的核心在于推理阶段的动态计算优化，包含三大创新方向：

（1）多步推理迭代，多动脑多想几遍

链式思考（Chain-of-Thought）：就好像写草稿分步骤，做数学题时先假设条件再推导结论，模型就像用“<假设><结论>”的标签分步骤写草稿，边想边检查。

共识生成（Consensus）：遇到难题时，让模型像小组讨论一样提出多个答案，然后投

票选最好的（比如 5 个答案里选出现次数最多的），或者用“评分老师”（奖励模型）挑出最优解。

搜索优化（Search）：就像程序员写代码时，写完先运行看报错，再根据错误提示修改，反复改错直到做对。

（2）计算资源动态调控，灵活控制思考时间

预算强制（Budget Forcing）：类似考试时间分配，给模型设置“最长思考时间”，简单的题快速答完（用<final_answer>提前交卷），难题多想想（用<wait>标签继续思考）。就像考试时先做简单题，留更多时间给大题。

潜在空间推理（Latent Reasoning）：心算代替笔算，有些思考不需要写出来，像数学高手在心里快速推算。某些模型用这种方法，1 秒钟能处理近 4000 个词，比边写边想快 3 倍。

（3）强化学习驱动优化，学练测闭环迭代

两阶段强化学习训练：第一阶段重点训练“解题步骤规范”（比如数学符号不能写错），像老师盯着你改作业。第二阶段还要兼顾“回答是否有用”、“是否安全”，就像学霸不仅会解题，还要学会把知识讲明白。

过程奖励模型（Process Reward Model, PRM）：大模型每想一步就有个“老师”打分，告诉它这一步对不对。就像做题时，每写一步就有人用红笔批改，及时纠正错误。

1.2.5 从通才到专家：专业能力增强

训练好的大模型尽管已经有了很强的世界知识，但是这些知识的实时性以及某个特定领域的专业度还存在欠缺，需要通过其他技术手段，进一步增强模型的知识 and 能力。

（1）知识检索增强：给模型装上“实时搜索引擎”

大模型训练数据截止后，知识会“冻结”（如不知道 2025 年的新政策）。通过实时网络搜索，在遇到问题时，模型会自动搜索最新资料。针对专业领域或私域数据，可以通过检索增强生成（Retrieval-Augmented Generation, RAG）^[5] 技术，通过向量匹配和语义理解，从专业数据库中精准提取信息。

（2）工具调用能力：让模型学会“用外挂”

尽管大模型的通用能力在不断提升，但在专业领域通常有很多专用工具，因此模型的工具体调用能力就显得尤为重要。模型能够像指挥官一样，分解任务，选择合适的工具（如计算器、编译器、专业工具链），并整合输出准确的结果。2024 年 11 月由 Anthropic 推出的开源协议 MCP（Model Context Protocol, 模型上下文协议），使 AI 应用能够安全访问和操作本地及远程数据，让 AI 模型与外部工具和数据连接，为 AI 应用提供了连接万物的接口。

（3）多专家协同：组建“AI 专家会诊团”

当存在多个领域模型的时候，可以让多个领域模型组合，形成“multi-agent”系统，完成复杂任务。就像在医疗咨询时，先由分诊模型判断科室，再转接专科模型进行专业诊断。2025 年 4 月，Google 开源首个标准智能体协议 Agent-to-Agent (A2A)，详细定义了一组标准端

点和响应模型，使得任何智能体 Agent 都可以采用此协议进行通信，聚焦于 Agent 之间的协作，消除技术栈之间的障碍。

通过类似 MCP、A2A 这些围绕 AI 系统构建的新技术的诞生，AI 产业朝着协作而非对抗的方向进化。模型的专业性和实时性得到了显著提升，能够在各个领域提供更精准、实时和高效的支持。

1.2.6 从混沌到清晰：扩散模型

讲到人工智能生成内容（AIGC，即 AI Generated Content；或国外亦称为 GenAI，Generative AI），一个绕不开的话题就是扩散模型（Diffusion Model）^[6]。

扩散模型的核心思想是通过逐步添加噪声来破坏数据（如图像），然后再通过一个学习模型逐步去噪，恢复出原始数据。想象一下，你有一张清晰的照片，然后你开始不断地往照片上撒细沙（噪声），越撒越多，照片就越来越模糊，最终完全变成了一堆随机的噪点。这就是正向扩散过程。扩散模型的目标就是学习如何逆转这个过程。也就是说，它要学会如何从一堆随机的噪点中，一步一步地把沙子（噪声）去掉，最终还原成一张清晰的照片。这就是逆向扩散过程，也叫去噪过程。

扩散模型在 AIGC 多个领域已经有了广泛的应用，大家最常见到的文生图模型，可以根据文字描述生成逼真的图像，比如 Midjourney，DALL-E 2，Stable Diffusion 等，还有视频生成模型，比如 Sora、Pika、Runway、可灵等。此外还有音频/音乐生成、3D 生成、图像修复和图像超分辨率等，扩散模型凭借高质量、多样性、可控性的生成能力，在各个领域都有着广泛的应用前景。

1.3 大模型技术发展趋势展望

首先是多模态融合和多模态涌现。自然界本身就是多模态世界，人类通过“眼耳鼻舌身意”来感知和认知世界，并且与这个世界交互。同样的，大模型正在从语言这一单一模态处理向深度融合的跨模态交互演进，实现文本、图像、音频、视频，乃至触觉、味觉、嗅觉等多维信息的联合理解与生成。多模态涌现，指的是当多模态大模型的规模（包括模型参数、数据量）以及多模态交互复杂度达到一定的临界点之后，会突然展现出跨模态的推理、泛化、生成甚至创造的能力。在大语言模型中，已经很好展示出了语言的智能涌现。而多模态涌现能力的实现，还有包括跨模态对齐、原生多模态模型架构、参数和数据的规模效应、跨模态信息的互补和增强在内等的技术难关等待攻克，值得人们继续发掘和探索。

其次，是高阶的推理能力，解决真实复杂任务。主流的大模型已经几乎穷尽了互联网上的文本知识，称其为“上知天文，下知地理，中晓人和，明阴阳，懂八卦，晓奇门，知遁甲”可能也不为过。最新的模型在代码、数学等推理任务上也展示了强大的性能，但是面对真实世界的复杂任务时候，还不能很好地“运筹帷幄之中，决胜千里之外”。除了进一步扩充高阶数据的规模、质量和多样性，还需要突破“思维链推理-强化学习”的技术难点和规模化瓶颈，让模型自我学习和自我迭代，真正让推理能力具备泛化性，结合领域适配和垂直场景强化，

做好安全对齐和幻觉抑制，这样大模型才能在更多的场景落地开花。在科学智能上，大模型正在成为科学智能的底座，赋能科学家加快实现科学发现。在工业智能领域，大模型可以深入参与工业生产的“设计开发-生产制造-销售服务”的全流程，为新质生产力发展提供新动能。

最后，我们也应该看到，除了大模型，还有诸如世界模型、空间智能、具身智能、群体智能、神经符号结合等多种新技术和观点在科学界涌现，实现通用人工智能（AGI）的技术路径可能并不是单一的，我们也期待有更多更新的技术范式出现加速通用人工智能时代的到来，并且可能比我们原来预料的更快。

1.4 结语

大模型技术作为当前实现通用人工智能的主流技术路线，已成为全球科技领域最具前沿性和战略意义的赛道之一。从国际到国内，参与者众多，竞争异常激烈。美国的科技巨头如 OpenAI、Google、Meta 等凭借强大的算力基础、数据资源和算法创新，持续引领着大模型的发展潮流。令人欣喜的是，中国的科研创新力量在大模型技术的发展中正扮演着越来越重要的角色，以 DeepSeek 等为代表，中国众多的科技公司、初创企业以及学术机构正加速追赶，用原始创新和开源模式打破欧美的技术垄断，并在部分领域实现超越。这种全球范围内的技术角逐不仅推动了大模型性能的快速提升，也使得相关技术迭代空前发展。

然而，我们也必须清醒地认识到，通用人工智能的技术路径尚未完全收敛，大模型虽是当前的主流方向，但未必是唯一解法。大模型技术本身也在快速演进之中，正在从堆砌算力和数据的竞赛转向更高效、更灵活的架构设计。大模型面临的诸多核心挑战，如模型的可解释性、能耗效率、伦理风险等，仍需持续攻关。此外，诸如神经符号 AI、量子计算与 AI 的结合、生物启发式计算等新兴技术路径，也提供了多样化通用人工智能技术探索可能。这种未收敛的状态既是挑战，也是机遇，它鼓励全球科研人员和产业界不断尝试新的思路和方法。

展望未来，中国人工智能发展不仅需要技术层面的持续突破，还需在国际合作与竞争中找到平衡点。对于中国而言，抓住这一历史性机遇，如持续加强国家政策支持、加大基础研究投入、培养跨学科人才、推动开源生态建设、完善产学研协同机制，将是实现持续不掉队甚至弯道超车的关键。对于全球人工智能发展来说，唯有通过开放协作、成果共享，才能加速通用人工智能实现，造福人类社会。

本章参考文献

1.1 国内外大模型技术发展态势

- [1]Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.
- [2]Kaplan J, McCandlish S, Henighan T, et al. Scaling laws for neural language models[J]. arXiv preprint arXiv:2001.08361, 2020.
- [3]OpenAI. ChatGPT[EB/OL]. OpenAI, (2022-11-30)[2025-03-06].

- <https://openai.com/index/chatgpt/>.
- [4]Achiam J, Adler S, Agarwal S, et al. Gpt-4 technical report[J]. arXiv preprint arXiv:2303.08774, 2023.
- [5]OpenAI. GPT-4V[EB/OL]. OpenAI, (2023-09-25)[2025-03-06].
<https://openai.com/contributions/gpt-4v/>
- [6]OpenAI. Sora[EB/OL]. OpenAI, (2024-02-15)[2025-03-06]. <https://openhaai.com/sora/>
- [7]OpenAI. O3 Mini[EB/OL]. OpenAI, (2025-01-31)[2025-03-06].
<https://openai.com/index/openai-o3-mini/>
- [8]OpenAI. GPT-4.5[EB/OL]. OpenAI, (2025-02-27)[2025-03-06].
<https://openai.com/index/introducing-gpt-4-5/>
- [9]Anthropic. Claude[EB/OL]. Anthropic,
(2023-03-14)[2025-03-06]<https://www.anthropic.com/news/introducing-claude>
- [10]Anthropic. Claude 3[EB/OL]. Anthropic,
(2024-03-04)[2025-03-06]<https://www.anthropic.com/news/claude-3-family>
- [11]Anthropic. Claude 3.5 Sonnet[EB/OL]. Anthropic,
(2024-06-21)[2025-03-06]<https://www.anthropic.com/news/claude-3-5-sonnet>
- [12]Anthropic. Claude 3.7 Sonnet[EB/OL]. Anthropic, (2025-02-25)[2025-03-06].
<https://www.anthropic.com/news/claude-3-7-sonnet>
- [13]Google. Bard: Google's AI Search Updates[EB/OL]. Google Blog, 2023-02-07[2025-03-06].
<https://blog.google/technology/ai/bard-google-ai-search-updates/>.
- [14]Team G, Anil R, Borgeaud S, et al. Gemini: a family of highly capable multimodal models[J]. arXiv preprint arXiv:2312.11805, 2023.
- [15]Google. Google Gemini Next Generation Model: February 2024 Update[EB/OL]. Google Blog, (2024-02-15)[2024-05-06].
<https://blog.google/technology/ai/google-gemini-next-generation-model-february-2024/>
- [16]Google. Google Gemini AI Update: December 2024[EB/OL]. Google Blog,
(2024-12-03)[2024-05-06].
<https://blog.google/technology/google-deepmind/google-gemini-ai-update-december-2024/>
- [17]xAI. Open Release of Grok-1[EB/OL]. xAI, (2024-03-17)[2025-03-06].
<https://x.ai/blog/grok-os>
- [18]xAI. Grok-2 Beta Release[EB/OL]. xAI, (2024-08-13)[2025-03-06].<https://x.ai/blog/grok-2>.
- [19]xAI. Grok-3 Beta[EB/OL]. xAI, (2025-02-19)[2025-03-06].<https://x.ai/blog/grok-3>.
- [20]Touvron H, Lavril T, Izacard G, et al. Llama: Open and efficient foundation language models[J]. arXiv preprint arXiv:2302.13971, 2023.
- [21]Dubey A, Jauhri A, Pandey A, et al. The llama 3 herd of models[J]. arXiv preprint arXiv:2407.21783, 2024.
- [22]Meta. Model Cards and Prompt Formats for Llama3[EB/OL]. Meta,
(2024-12-06)[2025-05-06]https://www.llama.com/docs/model-cards-and-prompt-formats/llama3_3/.
- [23]百度云. 文心智能体平台-文档中心[EB/OL].
(2023-11-15)[2025-03-06].https://agents.baidu.com/docs/develop/model/ERNIE_bot_introduce/
- [24]Bai J, Bai S, Chu Y, et al. Qwen technical report[J]. arXiv preprint arXiv:2309.16609, 2023.
- [25]Bai J, Bai S, Yang S, et al. Qwen-vl: A frontier large vision-language model with versatile

- abilities[J]. arXiv preprint arXiv:2308.12966, 2023, 1(2): 3.
- [26]Yang A, Xiao B, Wang B, et al. Baichuan 2: Open large-scale language models[J]. arXiv preprint arXiv:2309.10305, 2023.
- [27]极客公园.Kimi 智能助手支持 200 万字无损上下文,月之暗面大模型长文本新突破[EB/OL]. (2024-03-20)[2025-03-07].<https://hub.baai.ac.cn/view/35888>.
- [28]商汤科技. 率先完成“云、端、边”全栈布局!大模型性能对标 GPT-4 Turbo,商汤日日新 SenseNova 5.0 全面升级[EB/OL]. (2024-04-23)[2025-03-06].
<https://www.sensetime.com/cn/news-detail/51167729?categoryId=72>.
- [29]Yang A, Yang B, Zhang B, et al. Qwen2. 5 technical report[J]. arXiv preprint arXiv:2412.15115, 2024.
- [30]Liu X, Qin B, Liang D, et al. Autoglm: Autonomous foundation agents for guis[J]. arXiv preprint arXiv:2411.00820, 2024.
- [31]Minimax.MiniMax-01 开源[EB/OL]. Minimax, (2025-01-25)[2025-03-06].<https://www.minimaxi.com/news/minimax-01-%E7%B3%BB%E5%88%97>
- [32]Qwen.Qwen 2.5 Max[EB/OL]. Qwen, (2025-01-28)[2025-03-06].<https://qwen-ai.com/2-5-max/>
- [33]Bai S, Chen K, Liu X, et al. Qwen2. 5-VL Technical Report[J]. arXiv preprint arXiv:2502.13923, 2025.
- [34]Shao J, Chen S, Li Y, et al. Intern: A new learning paradigm towards general vision[J]. arXiv preprint arXiv:2111.08687, 2021.
- [35]Team InternLM. Internlm: A multilingual language model with progressively enhanced capabilities[EB/OL].(2023-9-27)[2025-03-06].<https://github.com/InternLM/InternLM-techreport/blob/main/InternLM.pdf>
- [36]浦语团队.书生·浦语 2.0 正式开源[EB/OL].(2024-1-17)[2025-03-06].<https://www.shlab.org.cn/news/5443847>
- [37]Chen Z, Wu J, Wang W, et al. Internvl: Scaling up vision foundation models and aligning for generic visual-linguistic tasks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2024: 24185-24198.
- [38]Chen Z, Wang W, Tian H, et al. How far are we to gpt-4v? closing the gap to commercial multimodal models with open-source suites[J]. Science China Information Sciences, 2024, 67(12): 220101.
- [39]InternVL. InternVL 2.0[EB/OL]. (2024-07-02)[2025-03-06].
<https://internvl.github.io/blog/2024-07-02-InternVL-2.0/>
- [40]INTERNLM. InternLM[EB/OL]. (2025-03-01)[2025-03-06].
<https://github.com/InternLM/InternLM?tab=readme-ov-file#model-zoo>.
- [41]Chen Z, Wang W, Cao Y, et al. Expanding performance boundaries of open-source multimodal models with model, data, and test-time scaling[J]. arXiv preprint arXiv:2412.05271, 2024.
- [42]Ma X, Wang Y, Jia G, et al. Latte: Latent diffusion transformer for video generation[J]. arXiv preprint arXiv:2401.03048, 2024.
- [43]Fan W, Si C, Song J, et al. Vchitect-2.0: Parallel transformer for scaling up video diffusion models[J]. arXiv preprint arXiv:2501.08453, 2025.
- [44]Liu A, Feng B, Xue B, et al. Deepseek-v3 technical report[J]. arXiv preprint

arXiv:2412.19437, 2024.

[45]Guo D, Yang D, Zhang H, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning[J]. arXiv preprint arXiv:2501.12948, 2025.

1.2 浅析大模型的核心技术

[1] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

[2] Sahoo P, Singh A K, Saha S, et al. A systematic survey of prompt engineering in large language models: Techniques and applications[J]. arXiv preprint arXiv:2402.07927, 2024.

[3] 卡尼曼 D. 思考, 快与慢[M]. 胡晓姣, 李爱民, 何梦莹, 译. 北京: 中信出版社, 2012: 20-25.

[4] Muennighoff N, Yang Z, Shi W, et al. s1: Simple test-time scaling[J]. arXiv preprint arXiv:2501.19393, 2025.

[5] Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks[J]. Advances in neural information processing systems, 2020, 33: 9459-9474.

[6] Rombach R, Blattmann A, Lorenz D, et al. High-resolution image synthesis with latent diffusion models[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 10684-10695.

第2章 AI 数据：驱动智能时代的核心引擎

2.1 模型突破推动数据需求升级

2.1.1 驱动人工智能进化的“燃料”

人工智能模型的发展史，本质上是一部“数据需求进化史”。从早期人工输入规则和知识图谱的符号主义（Symbolic AI），到深度学习时代让机器从海量数据中挖掘数据特征和规律，再到如今的大语言模型和多模态模型对数据规模、类型和质量的系统性重构，每一次模型的技术突破都伴随着对数据需求的重新定义。

一方面，是对数据数量的需求变化。以 Transformer 架构为例，其自注意力机制在理论上突破了传统 RNN/CNN 局部感知的局限^[1]，随之而来的就是对数据需求的指数级增长。比如 GPT-3 的训练数据量约 45TB^[2]，相当于数个国家级图书馆的藏书量；多模态模型 CLIP 则需处理 4 亿对图文数据^[3]，相当于 200 万本图画书。这种“模型越复杂，数据越要吃撑”的强耦合关系，被 Hoffmann 等人^[4]称为“Chinchilla Scaling Law”——模型参数和数据量应以相同比例增长的规律（佳数据量（Token 数）约为模型参数数量的 20 倍），才能实现最佳性能。

另一方面，是对数据质量的需求变化。随着大模型在语言理解和多模态领域的深入应用，高质量训练语料的重要性愈发凸显。除国外研究机构积极开源的训练数据集（如 The Pile、C4、LAION-5B、ImageNet-21k 等）外，国内研究机构也在积极打造面向大模型训练的大规模语料库。如，The Pile（EleutherAI）：包含网络文本、学术论文、图书合集等，广泛用于 GPT-NeoX 等开源大模型的预训练，是开源社区中的高知识密度语料标杆；又比如，万卷语料库（WanJuan）：由 OpenDataLab 联合多家机构推出的高质量训练语料系列，包括覆盖文本、图文和视频的万卷·多模态，以及面向通用基础能力的万卷·CC 以及面向多语言的万卷·丝路系列；除此之外，谷歌团队提出的 C4（Colossal Clean Crawled Corpus）和北京智源研究院发布的悟道语料库（WudaoCorpora）都为高质量训练语料奠定了重要基础。

庞大的规模和高维的质量，正在推动数据这个人工智能的“燃料”不断释放强劲动能。

2.1.2 从规模到价值的数据需求重构

当前大模型技术突破对数据需求的重构，可以从以下关键维度进行剖析：

（1）数据质量的颗粒度从粗到细

传统监督学习模型（如 ResNet）依赖人工标注的“静态标签”（如 ImageNet 的类别标签），但大模型对数据的语义理解需求已从“类别判别”升级为“跨模态语义对齐”。如，多模态模型 PaLI 要求图文数据在像素、物体、场景、语义四个层级实现严格对齐^[5]。同时，自监督学习的兴起使得数据必须包含足够的“隐式信息”——如，对比学习模型 SimCLR 要求原始单张图像可通过数据增强而生成多视角的图像^[6]，这些多视角的数据则是原始图像的“隐式信息”。

因此模型技术的迭代发展对原始数据的丰富性提出了更高要求。

(2) 数据模态的演变由单到多

大模型的数据需求正从单模态向多模态加速演进。早期模型（如 BERT、GPT）仅依赖文本数据，通过语言模式学习语义，使用单一模态的数据会缺乏对现实世界的具象认知。多模态大模型（如 GPT-4^[7]、Flamingo）通过融合文本、图像、音频、视频等多模态的数据，构建不同模态之间的关联，模拟人类多感官协同的认知方式。因此模型技术的突破对原始数据的多样性提出了更高要求。

(3) 训练数据的依赖由静到动

大模型落地面临的核心瓶颈之一是“知识冻结”——传统模型训练依赖离线静态数据，难以适应现实世界的动态变化。Meta 开源的 LLaMA 采用了分阶段训练的方案，这为模型的应用落地产生更多灵活性，但是也要求数据的工程能力具备实时采集、清洗、去重能力。以金融领域为例，高频交易模型的训练数据需实现分钟级甚至更高频更新，并需要包括事件性的标签标记（如财报发布、政策变动的标记）。

2.1.3 数据需求变化牵引技术升级

模型突破引发的数据需求升级，正在倒逼数据生产链的技术重构。

(1) 合成数据（Synthetic Data）的崛起

当真实数据难以满足模型需求时，合成数据则成为关键补充。NVIDIA 的 Omniverse 平台通过物理仿真生成模拟的自动驾驶训练数据，其研究表明合成数据可将标注成本降低 90%^[8]。然而，合成数据的“真实性鸿沟”（Reality Gap）问题仍未完全解决——剑桥大学团队发现，过度依赖合成数据会导致模型在一些特殊场景下的边缘案例上性能下降^[9]。

(2) 数据价值链的重构

传统数据标注产业（如通过人工对图像中物体进行边界框的标注）正被自动化工具取代。Google 提出的“AutoML-Zero”方案可实现数据预处理流程的自动优化^[10]，而 Scale AI 等企业则通过“人机协同标注”将标注效率提升 3 倍。这一变革的本质是数据的产生方式正从“劳动密集型”向“技术密集型”迁移。

(3) 隐私与合规的技术平衡

GDPR（欧盟《通用数据保护条例》）、CCPA（美国加利福尼亚州《消费者隐私法案》）等法规对数据使用的限制，迫使企业探索隐私计算落地新场景。联邦学习允许模型在非共享的数据上进行训练，但是其数据交互的通信成本高与异构数据处理困难的问题仍待突破。最新研究显示，差分隐私与联邦学习的结合可在保护数据隐私性和保障模型效果之间找到更佳的平衡。

(4) 大模型中规模化法则 (Scaling Law) 的争议与现状

2020 年, OpenAI 在论文《Scaling Laws for Neural Language Models》^[11]中提出, 模型性能与计算量、数据量和参数量呈指数关系, 即“大力出奇迹”的逻辑一度主导了大模型发展。过去十年, 该定律支撑了以 GPT 系列为代表的模型技术突破。然而, 随着性能提升放缓, 学术界对其有效性产生分歧。争议本质指向数据瓶颈: 互联网公开数据耗尽将导致传统预训练受限, 但垂直领域的高质量数据与合成数据可能成为数据规模新的增长点。

AI 模型的技术突破绝非孤立事件, 其背后是数据需求从“量变”到“质变”的系统性升级。未来, 数据将不再只是模型的“燃料”, 而是驱动技术创新的“战略资产”。这一趋势对行业的核心启示在于: 数据能力的构建必须与模型演进同步规划, 任何脱离数据战略的技术路线图都将面临天花板效应。正如纽约大学教授 Gary Marcus 所指出的: “深度学习的天花板不是算力, 而是高质量数据的可获得性。”

2.2 驱动因素推动数据生产革新

2.2.1 核心技术突破

大模型对数据规模、质量及多样性的需求升级正在倒逼数据生产技术的系统性革新。主要的数据生产技术体现在自动化、合成化与隐私化三个方面: 自动化标注技术通过模型进行辅助标注, 进而反哺数据生产, 形成“模型-数据”闭环。例如, Google 提出的 SimCLR 框架^[1]利用对比学习自动生成标签, 显著降低人工标注成本, Scale AI 的“人机协同标注”系统通过不确定性估计动态分配任务, 效率提升 300%; 合成数据生成技术则在真实数据缺口场景下发挥重要作用, NVIDIA Omniverse 通过物理仿真生成自动驾驶场景数据, 而 StyleGAN3 生成的合成人脸在视觉真实性测试中表现优异^[2]; 在隐私保护领域, 联邦学习与差分隐私成为合规“标配”, 蚂蚁链提出的“区块链+联邦学习”架构实现了跨境数据安全流转与可追溯。这些技术突破将数据生产从劳动密集型转向技术密集型, 推动数据从采集到应用的全链条生产方式的革新, 最终为大模型提供更可靠、更丰富的训练数据支撑。

2.2.2 大模型数据处理流程优化

大模型的高效训练离不开超大规模数据的高质量处理。从数据获取到标注, 全流程均需要在规模、精细度与合规性之间取得平衡。以下四个阶段相辅相成, 形成“大模型训练数据”的核心处理流程。

(1) 数据去重

在大模型语料构建的初期, 开发者往往会从网络爬取海量原始开放数据, 或整合已有大规模数据集进行二次加工。在获取数据的基础上, 要对重复数据进行去除, 防止重复数据浪费存储和算力, 或因模型过度拟合影响泛化性能。对多模态数据而言, 还需结合特定的去重

算法防止数据视觉角度的重复。常见的去重方法如 CommonCrawl 采用算法过滤来快速判断网页是否已抓取^[3]，PaLM-2 则结合语义相似度聚类来精确剔除重复文本^[4]。

(2) 数据提取

数据提取包括网页类 HTML 数据的提取、文档类如 PDF 数据的提取等，目的是得到可被机器理解的自然语言，这一环节也是大模型数据处理的关键环节。在大模型场景下，“数据提取”不仅关乎文本可读性，更是对隐含结构、元数据以及多模态信息的深度挖掘。特别是 PDF 与 HTML 两种主流格式，对后续数据清洗、语义理解及模型训练有重大影响。

目前先进的文档布局理解与结构化建模通过视觉-文本多模态预训练，将版面信息（段落位置、字体、表格网格等）与文字内容共同编码，显著提升对段落合并、表格检测、公式识别等任务的抽取准确率。MinerU、Marker、Nougat 等常用工具在复杂排版的文档、数学公式、表格内容识别和提取上为面向大模型的文档数据提供了高质量的语料基础。在 HTML 网页格式数据的抽取深度和动态性方面，较多前沿性研究在网页正文提取、噪声剔除、DOM 树和与语义标记、跨页面整合等方面多有探索。

无论是 PDF 还是 HTML，数据提取过程都直接影响训练语料的完整性、准确性与适用性。先进的文档理解模型、多模态预训练算法以及高可扩展的分布式数据管道，正在为大规模数据抽取提供新的路径。从全局来看，若在抽取阶段就能高效识别并剔除噪声、正确保留关键结构信息，将显著提升后续清洗、过滤与标注阶段的效率与质量。

(3) 清洗过滤

在获取并提取了海量文本或图像后，清洗与过滤是不可或缺的步骤。首先在不良内容过滤的方法上，GPT-4 采用 RoBERTa 分类器过滤低质量文本^[5]，Google 通过 Perspective API 基于 BERT 检测有害言论，准确率高达 92%^[6]。其次，在内容主题与语言检测的方法上，针对多语言数据集（如 Common Voice、Wikipedia Dump 等），需要自动化检测并分割不同语言文本，以进行后续的分语种处理或多语种对齐。另外在数据分布一致性的检测上，当抽取到的数据分布与任务目标领域明显不符时，应及时进行抽样审查或统计分析，避免将不相关或噪声过多的数据注入训练集。

(4) 标注增强

为了进一步提升模型在特定任务中的性能，往往需要进行额外的标注与数据增强。这包括自动标注，即 Meta 的 LLaMA 通过掩码语言建模自动生成部分标签^[7]，T5 模型利用知识蒸馏生成软标签^[8]；人机协同，即对于难度较高或模糊的样本，可通过不确定性估计将任务分配给人工标注员精修，提升整体标注质量；对比学习与数据增强，即在图像领域，SimCLR 等对比学习框架能自动生成多视角图像；而在文本领域，可利用同义替换、反向翻译等技术增强样本多样性。

2.2.3 数据共享与协同

基于当前仍旧存在的合成数据的真实性鸿沟、隐私-效用平衡难题、多模态对齐成本高企等问题，未来可能的突破包括：高保真合成数据生成，即基于扩散模型的高保真图像/视频合成^[9]有望弥补真实数据短板；零样本标注与提示词（Prompt）工程，即通过少样本或零样本的 Prompt 设计自动引导模型生成高质量标注^[10]，大幅减少人工参与；去中心化 Data DAO 生态，即利用区块链与加密技术构建分布式数据自治组织（如 Ocean Protocol），促进跨地区、跨机构的数据共享与交易；全球治理与标准化，即在国际层面需建立跨域数据流通标准（如 ITU AI 数据质量标准），并推进类似“一带一路”数据走廊的合作机制；在产业层面鼓励开源社区（OpenDataLab、ModelScope 等）共建新的数据平台与工具链。

数据生产技术的迭代不仅是工具的更新迭代，更是 AI 产业底层逻辑的深层重塑。实现大模型时代的可持续发展，需在技术自主化（国产化工具链）、生态协同化（产业与开源社区共建）与全球治理（国际数据流通标准）的多维度持续探索与突破。

2.3 技术演进推动数据生态重构

2.3.1 数据分工体系日益精细

当前全球人工智能数据产业已形成多层次协同体系，核心参与者围绕数据价值链条展开深度协作。首先，在数据生产层面，科研机构与科技巨头主导高质量数据资源的开发。OpenAI 开发的 WebText 语料库^[1]通过过滤和清洗 Common Crawl 数据得到高质量的训练数据。其次，在数据服务层面，Scale AI、Labelbox 等技术企业依托 AutoML 工具链重构产业角色，将数据标注由纯人工标注阶段推向 AI 主导并结合人工阶段，驱动数据标注产业从劳动密集型向算法增强型转型。另外，在数据治理层面，制度创新与技术创新形成共振。上海数据交易所发布 U235 框架，连通真实世界和数据世界，基于 SwiftLink 管理平台核心架构实现数据上链、收益上链、资产上链，提升资产的真实性、透明性、流动性、经济性和共识性，使得“真数据”成为“金资产”，为数据要素创新应用探索路径，助推全球数据交易市场蓬勃发展。据统计，2023 年全球数据交易规模约 1261 亿美元，至 2030 年预计有望达到 3708 亿美元。数据生态从线性供应链向网状价值网演进，催生出更具弹性与创造力的新型分工范式。

2.3.2 数据获取方式越趋便捷

数据开源使得数据的获取方式在广度、深度和便捷度上更进一步。在开源生态领域，国内外开源数据平台正加速向多模态、工具链整合及社区协作方向演进。国内以 OpenDataLab 为代表，聚焦大模型训练全链路支持，通过智能数据工具（如 MinerU）和超大规模精标数据集（如“书生·万卷”），构建起覆盖预训练、微调到评测的完整生态，成为国产大模型数据基座的核心力量。国际层面，Kaggle 凭借竞赛机制与海量公开数据集持续引领数据科学实践，HuggingFace 则以模型库、数据集及开源工具重塑 AI 开发范式，推动全链路创新。Papers

with Code 通过论文与代码的深度绑定加速科研复现，GitHub/Gitee 作为代码托管核心平台，在信创政策下成为国产替代与全球协作的关键枢纽。整体趋势显示，开源平台正从单一数据托管向“数据-工具-模型-应用”全栈生态升级，依托开放协作与 AI 原生工具（如代码生成、自动化标注），赋能开发者高效参与技术迭代，并加速跨领域交叉创新。

更深远的影响则来自开源生态的重构作用。DeepSeek 开源降低了模型使用门槛，未来将形成“大厂炼基座+中小厂做应用”分工。这种开放策略不仅加速垂直领域数据共享（如金融风控、工业质检的专有数据集开放），更倒逼行业数据标准建立——医疗影像 DICOM-CV 互操作协议、自动驾驶场景的 nuScenes 标注规范等均在开源社区驱动下成型。技术栈革新与开源文化同向偕行，正构建起“数据生产-开放共享-标准反哺”的正向循环，重塑全生命周期技术价值链。

2.3.3 合规约束意识增强

全球数据生态正经历技术突破与监管框架的深度博弈，驱动生产模式与价值链的范式重构。在监管引领侧，欧盟《人工智能法案》以数据可追溯性为核心要求，推动企业构建覆盖数据采集、标注、使用的全链路审计体系，倒逼合规能力嵌入技术底座。在技术驱动侧，特斯拉“自动标注工厂”通过仿真合成与强化学习实现标注效率跃升。2018 年，为了应对越来越多的标注需求和越来越复杂的标注规则，特斯拉创建了千人规模的数据标注团队，并搭建了专业的数据标注系统。最初特斯拉大多数的标注还是在二维图像上进行，一个图文多模态模型需要 533 小时的人工标注完成。从 2018 到 2021 年，特斯拉的数据工程平台经历了 3 次迭代，2021 年至今，特斯拉的数据标注开始转移到四维空间（三维空间+时间维度），通过轨迹的聚合重建与自动标注搭配，一个模型的标注仅需 0.5 小时算力+0.1 小时人工，效率较 2018 年提升了 800 倍之多^[4]，标志着数据生产从依赖人力转向 AI 原生模式；更深层次的生态协同正在显现——Databricks 提出的 Lakehouse 架构^[5]打通数据治理与模型训练流程，实现数据治理与模型训练流程的无缝衔接，形成“数据-模型”双向增强闭环，使合规要求与模型性能同步优化。目前，越来越多的企业已经设立了独立数据资产管理部门，除合规执行外也涉及数据资产评估、跨域流通谈判与伦理风险对冲等业务，推动数据集生态由资源供给转向价值创造。在此进程中，合规约束非但未抑制创新，反而成为驱动技术代际跨越与生态位重组的核心变量。

2.4 全球竞争推动数据战略迭代

2.4.1 政策博弈与技术标准双重竞争

当前，数据主权和标准主导权已成为大国科技竞争的关键，中、美、欧采取了不同的策略。美国试图凭借技术优势，尤其是硅谷巨头，控制“算力-数据-模型”这三个核心环节，并通过《芯片与科学法案》等手段加强对算力基础设施和数据资源的控制，但其数据生态存在

结构性问题。中国拥有海量数据，这推动了人工智能的快速发展，但同时也面临底层技术依赖国外和需要摆脱技术跟随者角色的挑战。欧盟则侧重于制定规则，通过《人工智能法案》强调数据可追溯性，并试图建立“欧洲数据空间”^[1]，但严格的 GDPR 合规成本导致一些欧洲企业向亚太地区转移。国际竞争的重点正从技术转向标准制定，如电气电子工程师协会（Institute of Electrical and Electronics Engineers, IEEE）的《联邦学习标准》和中美企业在世界知识产权组织（World Intellectual Property Organization, WIPO）的专利争夺，都体现了对数据治理规则话语权的争夺。这场竞争的深层原因在于“数字主权”和“技术互通”之间的矛盾，各国都在争夺数据控制权、技术主导权和规则定义权，这正在重塑全球数字秩序。

2.4.2 数据要素与技术创新双重驱动

全球数据交易加速活跃，2023 年全球市场规模达到了 1261 亿美元，但技术和制度上的限制仍然阻碍了数据价值的充分发挥。从政策层面看，国家出台的数据“二十条”旨在推动全国统一的数据市场建设，为释放国内数据要素的价值提供保障。上海数据交易所的发展就是一个例子，2024 年的交易额已经超过 40 亿元，比去年和前年都有显著增长，挂牌的数据产品也超过了 4500 种^[2]。尽管发展迅速，但数据定价仍然面临产权不清晰、价值评估模型缺失等问题。从技术层面看，在可信流通领域，蚂蚁链的“区块链+联邦学习”架构实现了跨境数据的“可用不可见”共享；在融合计算领域，华为云 ModelArts 攻克了文本、图像、视频等多种数据类型对齐的难题，支持大规模数据的联合训练。隐私计算技术也形成了双重保护，差分隐私通过加入噪声来保护个人数据，安全多方计算则实现了多方数据协同建模。这些技术结合，使数据在使用价值和合规性之间尽可能达到平衡，推动数据要素从封闭走向开放协作，为市场化进程提供创新动力。

2.4.3 立法自主、技术攻坚、生态协同三位一体

面对未来日益激烈的数据竞争，我们需要从数据主权、产业发展和生态构建三个方面入手，才能获得战略优势。首先，在数据主权方面，各国都在加快完善相关法律体系，建立跨境数据流动的“负面清单”，明确哪些敏感领域的的数据不能随意传输，比如金融和地理信息等。同时，还要努力实现从芯片到训练框架的全套国产技术自主可控。国家级数据资源池的建设需要突破现有局限，整合政务、科研、产业等多个来源的数据，构建覆盖重点领域的基础数据集，从而强化数据的多样性和主权属性。其次，在产业发展方面，市场需要聚焦数据生产链上的关键环节不断创新。比如，在自动化标注领域，要对标国际领先技术，推动更多标注任务实现自动化。同时，多模态合成数据以及高质量的推理数据需要在医疗影像、自动驾驶和科学研究等领域加快应用。此外，还要培育一批数据服务商，构建包括标注工具链、数据交易平台和合规审计服务在内的完整产业链，实现上下游协同创新。最后，在生态构建方面，国际标准制定的话语权非常重要。中国通过“数字丝绸之路”建设“一带一路”数据走廊，推动与东盟、中东等地区的数据互联互通。同时，开源生态建设也需要双轨并行，一方面，开源

社区要加速覆盖核心数据集；另一方面，要推动国内开源数据平台赶超国际主流平台，最终形成“中国数据-全球生态”的良好局面。

2.5 结语

数据作为 AI 时代的核心燃料，既是人工智能时代的战略资源，也是推动人工智能发展的内驱动力，随着技术的不断迭代和对人工智能认知的不断升级，全球范围内的国家、地区、企业都对数据获取的重要性和数据使用的有效性寄予“厚望”，并不断探索和尝试。我国作为数据大国，必须在发挥数据要素的放大、叠加、倍增作用上持续发力，抓住未来 3-5 年的关键窗口期，实现“数据大国”到“数据强国”的质变，补足高质量发展的“数字板块”。

同时，我们也清醒地看到，在全球竞争环境下，各个国家、各个地区的法律法规不尽相同、不够完备，给数据的开发使用共享升级带来制度性障碍。同时，数据使用的各个环节和数据要素本身依然存在瓶颈性问题，亟需新的理论创新和技术突破。除外，数据的安全性问题也始终困扰着包括政治家、科学家、工程师和普通群众在内的所有人员，数据共享和数据保护在某种程度上是一对不可调和的矛盾。

面向未来，需要进一步完善法律法规、提升技术防护能力、加强监管执法，充分保护数据隐私安全。在这个前提下，加强数据治理、优化数据采集、推动数据共享，实现数据数量和质量的双提高。在技术层面要加强技术研发、拓展应用场景、培养专业人才，促进数据技术和应用的不断创新。各个国家、各个地区之间要进一步达成共识，共同参与国际标准制定、不断加强国际交流合作、不断推动技术输出引进，构建数据产业健康有序发展。

本章参考文献

2.1 模型突破推动数据需求升级

- [1] Vaswani, A., Shazeer, N., Parmar, N., et al. Attention Is All You Need. *Advances in Neural Information Processing Systems*, 2017, 30.
- [2] Brown, T., Mann, B., Ryder, N., et al. Language Models Are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 2020, 33.
- [3] Radford, A., Kim, J. W., Hallacy, C., et al. Learning Transferable Visual Models From Natural Language Supervision. *Proceedings of the 38th International Conference on Machine Learning (ICML)*, 2021.
- [4] Hoffmann, J., Borgeaud, S., Mensch, A., et al. Training Compute-Optimal Large Language Models. *arXiv preprint arXiv:2203.15556*, 2022.
- [5] Chen, T., Li, X., Sciotto, S., et al. PaLI: A Jointly-Scaled Multilingual Language-Image Model. *arXiv preprint arXiv:2209.06794*, 2022.
- [6] Chen, T., Kornblith, S., Norouzi, M., et al. A Simple Framework for Contrastive Learning of Visual Representations. *Proceedings of the 37th International Conference on Machine Learning (ICML)*, 2020.
- [7] OpenAI. GPT-4 Technical Report. *Technical Report*, OpenAI, 2023.

- [8] Kato, S., Takeuchi, E., Ishiguro, Y., et al. Autoware on Board: Enabling Autonomous Vehicles with Embedded Systems. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Workshop, 2018.
- [9] Ros, G., Sellart, L., Materyska, J., et al. The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [10] Real, E., Liang, C., So, D., et al. AutoML-Zero: Evolving Machine Learning Algorithms From Scratch. Proceedings of the 37th International Conference on Machine Learning (ICML), 2020.
- [11] Kaplan, J., McCandlish, S., Henighan, T., et al. Scaling Laws for Neural Language Models. arXiv preprint arXiv:2001.08361, 2020.

2.2 驱动因素推动数据生产革新

- [1] Chen, T., Kornblith, S., Norouzi, M., et al. A Simple Framework for Contrastive Learning of Visual Representations. Proceedings of the 37th International Conference on Machine Learning (ICML), 2020, pp. 1597-1607.
- [2] Karras, T., Aittala, M., Laine, S., et al. Alias-Free Generative Adversarial Networks. Advances in Neural Information Processing Systems, 2021, 34: 852-863.
- [3] Raffel, C., Shazeer, N., Roberts, A., et al. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. Journal of Machine Learning Research (JMLR), 2020, 21(140): 1-67.
- [4] Chowdhery, A., Narang, S., Devlin, J., et al. PaLM: Scaling Language Modeling with Pathways. arXiv preprint arXiv:2204.02311, 2022.
- [5] OpenAI. GPT-4 Technical Report. Technical Report, OpenAI, 2023.
- [6] H. Hosseini, S. Kannan, B. Zhang, and R. Poovendran, Deceiving Google's Perspective API Built for Detecting Toxic Comments, arXiv preprint arXiv:1702.08138, 2017.
- [7] Touvron, H., Lavril, T., Izacard, G., et al. LLaMA: Open and Efficient Foundation Language Models. arXiv preprint arXiv:2302.13971, 2023.
- [8] Rombach, R., Blattmann, A., Lorenz, D., et al. High-Resolution Image Synthesis with Latent Diffusion Models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 10684-10695.
- [9] Schick, T., Schütze, H., Siriwardhana, S., et al. True Few-Shot Learning with Language Models. Findings of the Association for Computational Linguistics (ACL), 2022, pp. 1805-1821.

2.3 技术演进推动数据生态重构

- [1] Brown, T., Mann, B., Ryder, N., et al. Language Models Are Few-Shot Learners. Advances in Neural Information Processing Systems, 2020, 33: 1877-1901.
- [2] Rombach, R., Blattmann, A., Lorenz, D., et al. High-Resolution Image Synthesis with Latent Diffusion Models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 10684-10695.
- [3] Dwork, C., Kenthapadi, K., McSherry, F., et al. Calibrating Noise to Sensitivity in Private Data Analysis. Theory of Cryptography Conference (TCC), 2006, pp. 265-284.
- [4] 整数智能. 智驾数据的终极标注工具：给 Tesla 带来 800 倍提效的 4D 标注工具. 技术报告, 整数智能信息技术（杭州）有限责任公司, 2024.
- [5] Armbrust, M., Ghodsi, A., Xin, R., et al. Delta Lake: High-Performance ACID Table Storage over Cloud Object Stores. Proceedings of the VLDB Endowment, 2020, 13(12): 3411-3424.

2.4 全球竞争推动数据战略迭代

[1] European Commission. Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) [EB/OL]. (2021-04-21)[2025-03-23]. <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>.

[2] 上海数据交易所, 弗若斯特沙利文 (Frost & Sullivan), 头豹研究院, 等. 《2024 年中国数据交易市场研究报告》[R]. 上海数据交易所 2025 年年度发布会, 2024-11-25.

第3章 AI 算力：支撑智能进化的底座

算力，作为处理和分析这些海量数据的能力，正日益成为数字经济的核心基础设施。算力产业是以计算能力为核心，涵盖硬件、软件、服务及应用的完整产业链。其核心价值在于从“成本中心”转向“价值创造中心”，成为驱动 AI 产业化发展的核心引擎。它贯穿了从底层芯片制造到千行百业智能化应用的各个环节，如同一条庞大而精密的生产线，源源不断地为数字世界提供动力。本章主要按产业链的位置和产品形态划分进行介绍，分算力硬件基础层（上游）、算力软件服务层（中游）、算力应用场景层（下游）。

3.1 上游：AI 算力硬件基础层

上游主要包括芯片、服务器、存储设备、网络设备等硬件。这些是算力产业链的核心环节，为算力设施提供运算、存储和网络的基础能力。

3.1.1 芯片

芯片是 AI 算力的核心硬件支撑，如同汽车的发动机，决定了整个系统的性能和效率。当前 AI 芯片市场呈现多元化竞争格局，主要包括图形处理单元（GPU）、现场可编程门阵列（FPGA）、专用集成电路（ASIC）等专用计算芯片，以及新兴的类脑芯片等。根据《中国算力发展指数白皮书（2022 年）》，中国智能算力占比已超 50%，人工智能芯片（如 NPU）正加速商业落地，尤其在生成式 AI 和大模型训练中占据主导地位。国际数据公司（IDC）与浪潮信息日前联合发布《中国人工智能算力发展评估报告》显示，2024 年，中国智能算力市场规模达 190 亿美元，同比增长 86.9%，其中 GPU 占比超 60%，但 ASIC 和 DPU 因能效优势增速更快。

在 AI 浪潮推动大模型研发和相关应用层出不穷的大背景下，算力需求维持高景气度，整体算力产业链持续受益。英伟达（NVIDIA）在 GPU 领域占据主导地位，其 Blackwell 需求强劲，供不应求，而国产替代芯片如昇腾、昆仑等也在快速发展。此外，芯片的制造、设计、封测技术的革新也成为了构筑算力芯片进一步发展的重要底座，芯片制造和 CoWoS 封装产业链也因旺盛需求积极扩产。

目前 AI 芯片以芯片内、节点内及系统级异构计算为主流技术发展方向。国外主要以英伟达（NVIDIA）系列、谷歌（Google）TPU 系列作为领航，其特点如下：

- **英伟达（NVIDIA）系列：**（1）A100/H100 GPU（Ampere/Hopper 架构）。A100 GPU 在 FP16 精度下的峰值算力达 312 TFLOPS（每秒浮点运算次数，是衡量算力性能的常用单位之一），H100 结合 Transformer 引擎特有的动态精度切换机制优化大模型训练，将单精度计算效率提升至理论值 6 倍以上，并支持多实例 GPU（MIG）技术，可将单卡分割为多个独立实例，提升资源利用率。（2）B200/GB300（Blackwell 架构）。首个采用多芯片封装（Chiplet）设计的 GPU，集成多达 2080 亿个晶体管，单卡算力可支持 1.8 万亿参数 GPT 模型的训练，能耗较前代降低 75%。可通过 NVLink 和 Quantum-X800 InfiniBand 实现超大规模集群互联，支持千卡级并行训练。

- **谷歌 (Google) TPU 系列：**TPU v4/v5。Google 设计的专用架构，为 TensorFlow 优化，采用脉动阵列设计，加速矩阵运算，TPU v4 单芯片 BF16 算力达 275 TFLOPS，TPU v5 进一步优化能效比。可通过光互连技术 (OCI) 构建 TPU Pod，支持数万芯片互联，单集群可训练超千亿参数模型。谷歌内部用于搜索、翻译等 AI 服务，对外则通过 Google Cloud 提供 TPU 租赁。
- **Intel 公司 Gaudi 高性价比。**基于 16nm 工艺，ResNet-50 训练性能达 NVIDIA V100 的 4 倍，功耗降低 50%，集成 10 个 100GbE 以太网端口，支持标准网络协议扩展。支持 FP32、bfloat16 精度，内置 RDMA 技术，可构建无阻塞大规模集群。
- **AMD 公司最新的 AI 数据中心 Instinct MI325X GPU，**配备 256GB HBM3E 内存，在某些基准测试中性能超过英伟达的 H200 GPU。
- **Groq 公司，**专为 AI 推理设计，提供高性能和低延迟的 AI 推理能力。其开发出的一种新 AI 处理器 LPU (Language Processing Unit)，理论上推理速度相较于英伟达 GPU 提高了 10 倍，但成本却可以降低到十分之一。

国内主要 AI 芯片及其特点：

- **寒武纪 MLU100 云端智能芯片。**国内首款云端 AI 芯片 (2018 年发布)，理论峰值算力达 128 万亿次定点运算，采用 16nm 工艺，典型功耗仅 80 瓦，支持大规模数据中心的复杂智能任务。支持端云协同，与终端处理器 (如寒武纪 1M) 适配，适用于视觉、语音、自然语言处理等多种场景。
- **华为昇腾系列 AI 芯片。**昇腾 910 系列是国产高性能 AI 芯片，支持混合精度计算，FP16 精度下峰值算力达 320 TFLOPS，适配大模型训练与推理。昇腾云服务提供全栈 AI 解决方案，支持昇思 (MindSpore) 框架，广泛应用于智算中心与企业级 AI 应用。
- **沐曦专注于高性能 GPGPU 设计，**支持深度学习和大模型推理。
- **天数智芯智铠系列。**通用 GPU 芯片支持 FP32/FP16 混合精度计算，适配主流 AI 框架 (如 TensorFlow、PyTorch)。适用于云端推理和训练任务，强调低延迟与高吞吐量。
- **摩尔线程 MTT 系列 GPU。**支持 Ollama 开源框架。提供从云端到边缘的多场景 AI 算力支持，兼顾图形渲染与计算加速。
- **燧原科技邃思系列 AI 芯片。**专注云端 AI 训练与推理，支持 FP32/FP16/BF16 等多种精度，部署于多个智算中心。采用自研架构，优化能效比，适用于大规模集群训练。
- **壁仞科技 BR 系列 AI 芯片。**BR100 系列采用 7nm 制程工艺，单芯片峰值算力达到每秒千万亿次浮点运算 (1PFLOPS)。其 BF16 算力突破 1000 TFLOPS，8 位定点算力达 2000 TOPS 以上，领先于同期竞品如英伟达 A100。

3.1.2 服务器

服务器是指在网络环境中为客户提供各种服务的计算机，其承担着数据存储、转发、发布等任务。按用途可分为通用服务器和专用服务器。通用服务器的核心计算单元通常由 CPU 芯片提供；专用服务器通过集成 GPU、ASIC、FPGA 等芯片，提供 AI 训练、AI 推理、图形渲染、科学计算等功能，满足客户在专用场景下的需求。AI 服务器需求快速增长，也推动了光模块、AI 芯片等上游产品的增长。

3.1.3 存储设备

算力芯片需要存储芯片配合运行，主要包括 RAM、ROM 两类。由于 AI 需要频繁读写，RAM 对于 AI 影响更大，其中包含用于芯片高速缓存的 SRAM、主存储器 DRAM (含 HBM) 等。ROM 包含 NAND Flash、Nor Flash 等形态。目前，DRAM 和 NAND Flash 是主流的两种植存芯片。HBM 属于 DRAM 的子品类，其主要特点是多层 DRAM 芯片堆叠，并采用大量 TSV 互联，因而带宽显著高于其他 DRAM 芯片，主要用于高性能计算场景。竞争格局从全球 HBM 产能来看，海力士相对领先，三星、美光紧随其后，国产化需求迫切，我国存储芯片企业亟待突破。

目前国内长江存储凭借技术突破和政策支持，正从“追赶者”向“领跑者”转型。尽管面临国际竞争和供应链压力，但其国产替代及 AI 赛道布局，正在为未来发展注入强劲动力。长江存储研发出 232 层 3D NAND 芯片，位密度达 19.8Gb/mm²，超越了同样在开发 232 层 QLC 3D NAND 芯片的美光和英特尔 (Solidigm)。自主创新的 Xtacking 架构缩短产品上市周期，提升存储密度和 I/O 速度，如 X3-9070 芯片读速达 7000MB/s。产品覆盖全场景，提供 3D NAND 闪存晶圆、颗粒及消费级/企业级 SSD，应用于移动设备、数据中心、AI 服务器等领域。代表产品如致态 TiPlus7100 系列 SSD，采用 Xtacking 3.0 技术，寿命达 2400TBW。2024 年全球存储市场规模达 1298 亿美元，长江存储通过技术迭代加速追赶三星、SK 海力士等巨头。合肥长鑫专注于 DRAM (动态随机存储器) 研发与生产，产品主要用于计算机、服务器等设备的内存条。其技术源于对德国内存企业奇梦达的技术收购，目前量产 19nm 工艺的 DDR4 芯片，并计划推进更先进制程。

表 3.1 合肥长鑫与武汉长江存储技术突破

维度	合肥长鑫 (DRAM)	武汉长江存储 (NAND)
技术领域	DRAM (动态内存)	NAND Flash (闪存)
核心突破	氮化硅层、CMOS集成、DDR5/LPDDR5技术	Xtacking架构、3D NAND堆叠技术
产能目标	2025年30万片/月 (DRAM)	2025年30万片/月 (NAND)
市场定位	国内服务器、PC、移动设备DRAM主力供应商	国产SSD、手机存储芯片核心供应商
行业意义	打破三星、SK海力士、美光垄断，保障供应链安全	实现NAND Flash国产化，降低对进口依赖

合肥长鑫与武汉长江存储分别代表中国 DRAM 和 NAND Flash 领域的突破，是国产存储器产业的核心力量。两者通过技术追赶和产能扩张，逐步打破国际垄断，推动中国半导体产业链自主化，未来在 AI、5G、数据中心等场景中将发挥关键作用。

3.1.4 网络设备

网络设备用于构建高效的数据传输网络。在如今超大模型与超大集群时代，网络设备的重要性极大提升。目前来看，铜连接（或可能其他）承担机柜内超高速连接，光模块和交换机承担机柜间拓展的分工形式已形成明显趋势。

在超大规模计算时代，网络设备已成为 AI 算力的“神经网络”，其技术演进直接决定着超大模型的训练效率和数据中心的扩展能力。国际巨头思科、英伟达、博通等凭借技术积累和生态优势占据高端市场，而国内厂商如华为、新华三、光迅科技等通过国产替代和差异化创新（如硅光技术、液冷方案）快速崛起。未来，随着 800G 网络、DPU 和 AI 网络调度技术的普及，网络设备将从“连接工具”进化为“智能算力基础设施”，成为全球科技竞争的新焦点。

表 3.2 国际国内网络设备相关企业对比

公司	技术领域	核心突破	代表产品	市场定位	行业意义
思科 (Cisco)	交换机、路由器、SDN	高端交换机市场主导、SDN (ACI 架构)	Nexus 9000 系列 (800G 交换机)、ACI 平台	全球高端数据中心核心网络设备供应商	推动软件定义网络 (SDN) 标准化，定义数据中心网络架构
英伟达 (NVIDIA)	AI 网络、交换芯片、光模块	InfiniBand 技术、DPU (BlueField)、AI 优化网络	Quantum-2 平台、BlueField-4 DPU、HDR 200G InfiniBand	AI 超算集群与云服务商首选	通过网络与计算融合，加速 AI 训练与推理效率
博通 (Broadcom)	交换芯片、光模块、全栈方案	Tomahawk 系列交换芯片 (800G)、协议标准制定	Tomahawk 5 交换芯片、StrataXGS 系列交换机	全球交换芯片与光模块技术霸主	推动 400G/800G 光模块标准化，支撑超大规模数据中心建设
华为	全栈网络设备 (交换机、光模块、软件)	AI Fabric (零丢包)、液冷技术、硅光模块	CloudEngine 系列 (400G/800G 交换机)、凌霄硅光模块	国内政企、运营商及超大规模数据中心	推动国产全栈网络解决方案，打破高端市场垄断
新华三 (H3C)	交换机、SDN、企业级网络	AI 驱动网络 (智能流量调度)、CloudOS 云平台	S12500/S10500 系列交换机、Comware 系统	国内政企、金融、教育等垂直领域	通过 AI 优化网络运维，助力企业数字化转型
光迅科技	光模块、硅光技术	400G/800G 硅光模块、国产化替代	100G/400G 光模块、硅光芯片 (1.6T 原型)	光模块中低端市场国产化核心供应商	降低光模块成本，提升中国在光通信领域的全球竞争力

3.2 中游：算力软件服务层

中游环节主要涉及算力网络的建设与运营，包括数据中心（IDC）、边缘计算、智算/超算中心以及云计算服务等。这一环节通过提供 IDC 服务、云服务、计算服务等，为下游应用提供算力支持。

3.2.1 数据中心与边缘计算

数据中心为存放服务器、交换机等网络 IT 设备并提供专业运维服务的机房，主要用于组织、处理、存储和传输大量数据，本质是通过为 IT 设备供给电力、空间、散热等方式，推动能源要素向数据要素转换。我国 IDC 行业的主要参与者分为三大群体：一是以中国电信、中国移动、中国联通为主的基础电信运营商群体；二是以润泽科技、奥飞数据、万国数据、世纪互联等为代表的专业 IDC 服务商群体；三是以阿里云、华为云、腾讯云为代表的云服务商，也围绕云服务自建数据中心。

边缘计算是指将计算能力从网络核心下沉到网络边缘，通过在终端设备附近建立计算节点，使数据、应用和服务的主要分析处理环节都在边缘节点进行。其目的是为了：降低时延，扩展带宽；位置感知，用户识别；本地化；支持设备异构性；以及提高资源利用率等。

智能计算中心（AIDC）专注于 AI 训练和推理，是人工智能发展的重要支撑。如华为昇腾和阿里云智算集群是国内智能计算中心的代表。华为昇腾智能计算中心依托华为强大的技术研发能力和产业生态优势，为人工智能开发者提供了高效、便捷的计算平台。其采用了先进的昇腾芯片和智能计算架构，能够实现大规模的 AI 训练和推理任务。如在智能安防领域，华为昇腾智能计算中心可以对大量的监控视频进行实时分析和处理，快速识别出异常行为和目标，为城市安全提供了有力保障。阿里云智算集群是阿里云推出的面向人工智能领域的高性能计算平台。它集成了阿里云的先进技术和资源，能够为用户提供灵活、可扩展的算力服务。阿里云智算集群广泛应用于电商、金融、医疗等多个行业，为企业的智能化转型提供了强大的支持。如在医疗领域，阿里云智算集群可以帮助医疗机构对大量的医学影像数据进行分析 and 诊断，提高疾病诊断的准确性和效率^[1]。

目前，国内已建成超 250 个智算中心，这些智算中心分布在不同的地区和行业，为人工智能的发展提供了强大的算力支持^[2]。它们不仅推动了人工智能技术的创新和应用，也为数字经济的发展注入了新的动力。

3.2.2 云计算服务

云计算服务是指通过互联网按需提供的计算资源、应用程序及服务，用户无需自行购置或维护物理设备，即可灵活使用存储、服务器、数据库、软件等资源。其核心特点在于弹性扩展和按使用付费的模式。

目前，观察和分析 AI 算力基础设施的视角主要在于：超大规模云服务厂商如何应对大模型对并行加速计算的内在需求以及进行“规模化”创新。伴随大模型技术的汹涌发展和背后

涌入的资本驱动,云厂商以大模型为工作负载,对基础设施的创新和服务能力水平逐渐分层,其中新进入者 AI 算力云(如:CoreWeave)在竞争分化中更占有优势。

自 2012 年 AlexNet 图像识别神经网络首次在多张 NVIDIA GPU 上进行训练以来,深度学习浪潮正式拉开帷幕。在 2012 年至 2020 年期间,随着深度学习技术的不断发展,全球云计算市场也逐渐走向成熟。在此阶段,开发者通常在单台到百台 GPU 服务器上完成模型训练,随后根据需求将推理部署在本地服务器或云端。2020 年起,以 OpenAI 为代表的开发者开始使用数千张 GPU 进行预训练大语言模型的训练,2022 年底,通用对话机器人 ChatGPT 作为大模型的典型应用被推出。为了支持用户规模的快速增长和应用服务的规模化,开发者通过自建 GPU 集群和租赁第三方云服务的方式满足需求。大模型从训练研发到推理生产,均以云(服务器集群)作为核心基础设施。

从整体市场看,提供 AI 云服务的传统超大规模提供商包括谷歌、微软、亚马逊、甲骨文,以及中国互联网大厂腾讯、百度、阿里巴巴等。尽管 Meta、xAI、字节跳动和特斯拉也拥有强大的 GPU 集群,但目前并不对外提供 AI 云服务,因此不属于这一类别。Google 拥有当今世界上最先进的计算系统,并率先大规模使用许多关键技术,如其机架级液冷架构和多数数据中心训练^[4],这些技术现在也被其他大模型初创公司和 AI 算力云巨头采用。

从集群规模、服务客户规模角度,目前在全球范围最主要的四家 AI 算力云巨头是 Crusoe、Nebius、Lambda Labs 和 CoreWeave,其中 CoreWeave 是迄今为止全球最大的 AI 算力云,管理着非常高性能的 GPU 集群,并被 OpenAI 和 Meta 甚至 Nvidia 委托管理其内部大型 GPU 基础设施^[5]。

中国的新型 AI 算力云的领军企业是字节跳动旗下的火山引擎,以及硅基流动等在中国 AI 行业的其他创新力量,在 2024 年底到 2025 年初国产 DeepSeek V3/R1 大模型推出后,中国算力云和模型厂商、GPU 厂商合作,推进实现大模型在国产 GPU 芯片上的推理甚至训练,该进程正在快速地迭代和演进^[6]。

3.3 下游：算力应用场景层

算力作为数字经济时代的核心生产力,正深度融入千行百业,形成“算力+产业”的创新范式。如在智能制造领域,工业边缘数据中心通过实时算力支持智能检测与故障分析,推动传统工厂向“智慧工厂”转型;教育领域通过公共算力资源覆盖校园,为高等教育和职业教育提供沉浸式实训环境;金融领域构建分布式算力架构,实现高频交易与低时延风控的协同;交通领域部署多层级算力设施,支撑车路协同自动驾驶与港口自动化生产;医疗领域则依托算力处理海量医疗影像数据,提升疾病诊断精度。此外,算力在元宇宙、数字孪生等新业态的拓展应用,正推动产业边界持续扩展。工信部 2025 年启动的算力强基揭榜行动,更将绿色算力、安全算力等关键技术纳入重点攻关方向,为产业深度赋能^[7-8]。以下是目前紧密相关的一些重要领域的算力应用场景示例:

3.3.1 互联网与消费：大规模数据处理

在互联网与消费领域，电商、游戏、视频等行业产生了大量的数据，需要强大的算力进行处理。

电商行业是数据密集型行业之一。电商平台每天都会产生海量的交易数据、用户行为数据等。这些数据对于电商企业来说具有重要的价值，可以帮助他们了解用户需求、优化商品推荐、提高客户满意度。如，阿里巴巴电商业务中历史数据存储与查询相关业务，大量采用基于列存储技术的 HiStore 数据库，双 11 当天 HiStore 引擎处理数据记录超过 6 万亿条、原始存储数据量超过 5PB。从单日数据处理量上看，该系统已成为全球最大列存储数据库^[9]。在如此体量巨大的平台上，每天的实物和虚拟商品交易达到亿级别^[10]，通过强大的算力支持，阿里巴巴可以对这些数据进行实时分析和处理，为用户提供个性化的商品推荐和优质的购物体验。

游戏行业也是算力需求较大的行业之一。随着游戏画质和复杂度的不断提高，游戏对计算能力的要求也越来越高。例如，一些大型 3D 游戏需要在短时间内处理大量的图形渲染和物理模拟任务。通过云计算和边缘计算等技术，游戏厂商可以将这些计算任务分配到云端或边缘设备上进行处理，提高游戏的运行效率和稳定性。

视频行业同样离不开算力的支持。随着高清视频、4K 视频甚至 8K 视频的普及，视频的制作、传输和播放都需要强大的计算能力。例如，抖音等短视频平台每天上传的视频数量数以百万计。通过算力支持，这些平台可以对视频进行快速地编码、解码和处理，确保用户能够流畅地观看视频。

3.3.2 智能制造：工业互联网、机器人、自动驾驶

在智能制造领域，工业互联网、机器人、自动驾驶等应用对算力提出了极高的要求。

工业互联网是智能制造的核心技术之一，它通过将工业设备、传感器、控制系统等连接起来（IoT），实现数据的实时采集、传输和分析。在工业生产过程中，会产生大量的生产数据、设备状态数据等。通过强大的算力支持，企业可以对这些数据进行分析和挖掘，实现生产过程的优化和智能化管理。例如，在汽车制造企业中，通过工业互联网和算力支持，可以实现对生产线的实时监控和故障预警，提高生产效率和产品质量。

机器人是智能制造的重要工具之一。机器人需要具备强大的计算能力来实现自主导航、目标识别、动作控制等功能。例如，在物流仓库中，智能机器人可以通过激光雷达、摄像头等传感器获取周围环境信息，然后通过强大的算力进行实时处理和分析，实现自主导航和货物搬运。

自动驾驶是未来交通发展的重要方向。自动驾驶汽车需要在复杂的路况下实时感知周围环境、做出决策和控制车辆行驶。这需要强大的算力来处理大量的传感器数据和进行复杂的算法计算。例如，特斯拉的自动驾驶汽车配备了先进的传感器和强大的计算芯片，通过不断学习和优化算法，提高自动驾驶的安全性和可靠性。

3.3.3 科研与公共服务：气候模拟、基因测序、智慧城市

在科研与公共服务领域，气候模拟、基因测序、智慧城市等应用也离不开算力的支持。

气候模拟是研究气候变化的重要手段。通过建立复杂的气候模型，科学家可以模拟不同气候情景下的气候变化趋势。这需要强大的计算能力来处理大量的气象数据和进行复杂的数值计算。例如，国家气候中心利用超级计算机进行气候模拟，为气候变化研究和应对提供了重要的科学依据。

基因测序是生物医学领域的重要技术之一。通过对生物体的基因进行测序，可以了解生物体的遗传信息和疾病发生机制。基因测序会产生大量的基因数据，需要强大的算力来进行数据分析和解读。例如，华大基因利用先进的测序技术和强大的算力支持，开展大规模的基因测序项目，为人类健康和生物科学研究做出了重要贡献。

智慧城市是城市发展的未来方向。通过将物联网、大数据、人工智能等技术应用于城市管理和服务中，可以实现城市的智能化运行和管理。在智慧城市建设中，需要处理大量的城市运行数据，如交通流量、能源消耗、环境质量等。通过强大的算力支持，可以对这些数据进行实时分析和处理，为城市管理者提供决策支持。例如，一些城市通过建设智慧城市平台，利用算力支持实现了交通拥堵预警、能源消耗优化等功能，提高了城市的运行效率和居民的生活质量。

3.3.4 金融与能源：高频交易、风险分析、能源调度优化

在金融与能源领域，高频交易、风险分析、能源调度优化等应用对算力也有很高的要求。

金融领域，高频交易是一种利用计算机算法进行快速交易的策略。高频交易需要在极短的时间内对市场行情进行分析和判断，并做出交易决策。这需要强大的计算能力和低延迟的网络连接。例如，一些大型金融机构，包括量化投资公司，利用高性能的服务器和算法交易系统，进行高频交易，获取市场差价收益。

风险分析也是金融领域的重要工作之一。金融机构需要对各种风险因素进行分析和评估，如信用风险、市场风险、流动性风险等。通过强大的算力支持，可以对大量的金融数据进行分析 and 建模，提高风险分析的准确性和可靠性。例如，银行可以利用机器学习算法对客户的信用风险进行评估，为贷款决策提供依据。

在能源领域，**能源调度优化**是提高能源利用效率和保障能源安全的重要手段。通过对能源生产、传输、消费等环节的数据进行实时监测和分析，利用强大的算力进行优化调度，可以实现能源的合理分配和高效利用。例如，电网企业可以利用智能电网技术和算力支持，对电力负荷进行预测和调度，提高电网的稳定性和可靠性。

3.4 算力发展现状与挑战分析

3.4.1 硬件突破：多元架构与生态博弈

GPU 的统治地位与国产突围。英伟达凭借其 CUDA 生态和算力优势占据全球 AI 训练市场超 80% 份额^[1]。但国产 GPU 正加速突破，2025 年，国产 GPU 企业摩尔线程凭借自主研发的 MTT S4000 系列产品，通过动态重构技术实现图形渲染与 AI 计算自由切换，在

Resnet50、BERT 等多个 AI 基准测试中达到了国际一流的水平，成为了首个真正打破国际巨头垄断的本土 GPU 企业^[12]；华为昇腾 910B 在昇腾社区生态中支持超 200 个应用，政务云市场占有率高达 67%^[13-14]。

专用芯片的崛起，异构计算成为新趋势。谷歌 TPU 通过脉动阵列架构实现张量运算效率提升 15-30 倍^[15]；华为昇腾 NPU 采用达芬奇架构，支持原生 MindSpore 框架，在端侧推理场景能效比达 15.8TOPS/W；光子芯片（如 Lightmatter）通过光子计算突破电子芯片的能耗瓶颈，某实验室测试显示其推理速度远比 GPU 更快^[16]。

量子计算的潜在颠覆，量子算法（如 Grover 搜索）可加速参数空间探索。微软 Azure Quantum 平台通过量子计算重构梯度下降的优化范式，不仅加速了大模型训练，更开辟了“量子-AI-HPC”三元融合的新赛道。随着量子纠错技术与混合架构的成熟，未来 AI 训练有望突破“摩尔定律”限制，实现指数级效率跃迁^[17]。量子神经网络（QNN）的并行性有望解决经典模型在复杂优化任务中的局部最优困境。

软件与算法协同，效率革命与生态适配。模型压缩与轻量化方面，知识蒸馏技术使大模型体积缩小的同时保持较高的准确率（如医疗影像诊断系统）；混合精度量化将原来较大的浮点数压缩，推理速度得以大幅提升；剪枝技术通过动态跳过冗余计算分支，使自动驾驶障碍物识别帧率也大幅提高。分布式训练框架创新方面，Primus 框架通过动态负载均衡和智能任务分配，使百万级参数模型训练通信开销降低^[18]；PyTorch DDP 结合 Ring AllReduce 架构，实现多卡并行训练效率提升，华为 MindSpeed 的 P2P 分流技术利用昇腾芯片异构计算能力，优化长序列训练稳定性等。DeepSeek 在 2024 年底、2025 年初得到全球瞩目也是最好的现实案例。

3.4.2 当前挑战：技术瓶颈与结构性矛盾

（1）算力瓶颈与制程限制

芯片制程逼近物理极限（1nm 节点），传统冯·诺依曼架构面临“内存墙”与“通信墙”双重制约，导致算力密度提升乏力^[19]。尽管摩尔线程通过 7nm Chiplet 工艺实现良品率提升，但英伟达 H200 流片成本飙升至单次 2.3 亿元，暴露高端芯片供应链的脆弱性。全球智能算力规模虽达 335EFLOPS（2023 年），但高性能算力占比不足 30%，且国产芯片单卡训练效率不到 A100 的一半，形成“量级追赶、能效代差”的困境。美国对华芯片出口管制（如 1017 新规）进一步加剧技术封锁，CUDA 生态垄断导致国产替代成本高昂。

（2）异构计算协同不足

随着计算需求的多样化，CPU、GPU、FPGA、ASIC 等不同架构计算单元被广泛应用。然而，当前不同架构计算单元之间的协同效率较低。一方面，技术标准的不统一使得各计算单元在通信和协作时存在障碍；另一方面，缺乏有效的调度算法和软件支持，难以充分发挥不同计算单元的优势，实现算力资源的最优调度。这导致了在实际应用中，算力资源的浪费和计算效率的低下。

(3) 高能耗与碳排放争议

AI 训练能耗持续攀升，ChatGPT 每天响应约 2 亿个需求，消耗超过 50 万度电力，相当于 1.7 万个美国家庭平均一天的用电量^[20]。训练一个大型 AI 模型的碳排放量相当于 5 辆汽车整个生命周期的排放量；使用 1750 亿个参数训练 GPT-3 消耗了 1287 兆瓦时的电力，并导致产生了 502 吨二氧化碳当量，相当于驾驶 112 辆汽油动力汽车一年。GPT-3 每日运行产生的碳足迹有 50 磅，相当于一年排放 8.4 吨二氧化碳。据预测，我国到 2030 年智算中心年用电可能达到 0.6 万亿千瓦时至 1.3 万亿千瓦时，占全社会用电的 5%至 10%^[21]。冷却系统能耗占数据中心总能耗的 40%，液冷技术（如联想问天海神）虽能效提升 30%，但中小型数据中心仍依赖传统风冷，平均 PUE 值达 1.5^[22-23]。绿色算力发展面临“高配置、低效能”悖论——80%智算中心未配备向量数据库，数据预处理消耗 45%算力资源，形成“算力增长、功耗失控”的恶性循环。

因此，需要进行能效优化与开发有效液冷技术。能效优化方面，算法与芯片协同设计成为核心，通过算法创新降低训练算力，推动高能效芯片需求。算法与芯片的协同设计已成为提升计算能效的核心路径。其核心理念是通过算法层面的轻量化设计，减少模型对计算资源的依赖，从而降低芯片硬件设计的复杂度与能耗需求。而液冷技术、存算一体架构等技术加速应用，有助于应对数据中心能耗攀升问题。液冷与存算一体技术正与 AI 调度算法结合，如谷歌利用强化学习动态调节数据中心制冷系统，进一步降低 10%-15%的冷却能耗。到 2025 年，绿色化技术有望将全球数据中心碳排放强度削减 40%以上。

未来，随着大模型参数量持续增长，“算法轻量化+芯片定制化”双轨策略将成为平衡性能与能耗的关键，预计到 2026 年，协同设计技术可使 AI 训练综合能效提升 3-5 倍。

(4) 算力分配不均与生态断层

全球来看，存在地域资源垄断与马太效应。目前北美是主导地位，美国占据全球超算 TOP500 榜单的 34%，英伟达、AMD 等企业垄断高端 GPU 市场，形成技术护城河。东亚具有追赶压力，中国依托政策驱动在智算中心建设上快速扩张，但高端芯片仍依赖进口；日本、韩国聚焦半导体制造，但应用生态碎片化。欧洲掉队，发展中国家边缘化，非洲、拉美等地区数据中心覆盖率不足 10%，算力需求与供给倒挂，难以支撑本地数字化转型。算力定价权集中，云计算巨头（如亚马逊、微软）掌握定价话语权，中小企业被迫承担溢价，加剧资源分配不公。

中国国内来看，算力供需的时空错配。我国算力资源分布不均，东部地区需求旺盛但资源相对紧张，西部地区资源丰富但需求相对不足。“东数西算”也具有一定局限性，中国西部可再生能源丰富，但网络带宽不足导致跨区域传输时延超 50ms，实时性业务难以落地。由于跨区域调度机制不完善，东西部在商业模式协调上困难重重，电价与算力成本难以实现有效平衡。此外，同质化竞争加剧，进一步阻碍了算力资源的合理配置和跨区域协同发展。

(5) 算力调度标准体系、互操作性缺失以及产业链协同不足

目前，算力度量、调度、结算等方面缺乏统一标准，异构资源的识别与感知技术也不成熟。这导致了不同算力资源之间难以实现互联互通和协同工作，形成了算力孤岛现象。缺乏统一标准和互操作性，限制了算力市场的规范化发展和资源的高效利用。

算力服务商与行业应用需求之间存在脱节现象，缺乏能够满足行业多样化需求的一体化解决方案。同时，传统企业对算力价值的认知不足，导致算力在行业应用中的深度和广度不够。产业链协同不足，使得算力行业难以形成完整的生态系统，限制了其对实体经济的支撑作用。

（6）市场认知与应用深化难题

一是当前算力市场存在供需结构失衡的问题。通用算力占比过高，而随着人工智能等新兴技术的发展，智能算力缺口日益增大，边缘算力布局也明显不足。这种供需结构的失衡，导致算力资源无法有效匹配行业多样化的需求，限制了算力在不同领域的深入应用。二是冷热数据处理矛盾。在“东数西算”过程中，约 80% 的“冷数据”适合迁移至西部地区进行存储和处理，但对于时效性强的应用，仍需要在东部地区就近处理“热数据”。如何优化资源调配策略，平衡冷热数据的处理需求，成为“东数西算”工程实施过程中的一大难题。三是成本与效益平衡。在高电价地区，算力中心的运营成本压力巨大。为实现成本与效益的平衡，需要通过提升机柜功率密度、优化资源利用率等方式降低成本。然而，这些措施在实际实施过程中面临着技术和管理等多方面的挑战。

（7）此外，我国还面临核心技术瓶颈与自主化等难题

高端芯片依赖度高。我国在高端芯片领域，如光刻机等设备，仍面临严峻的技术瓶颈。由于技术研发滞后，目前对进口高端芯片的依赖程度极高。这种依赖不仅使得我国算力基础设施建设容易受到国际形势变化的影响，面临供应链安全风险，更重要的是，芯片自主创新能力的不足已成为阻碍算力基础设施发展的关键因素。缺乏自主可控的高端芯片，意味着在算力的核心环节上受制于人，难以实现算力的高效提升与自主发展。

3.5 算力发展趋势与方向

3.5.1 硬件创新：突破传统架构的物理极限

一是存算一体芯片的突破。通过将存储单元与计算单元深度融合，消除冯·诺依曼架构的“存储墙”瓶颈。例如，美国 Mythic 公司开发的模拟存算芯片可实现 100TOPS/W 的能效比；中国的知存科技已量产 WTM2 系列芯片，支持语音识别、图像分类等低功耗场景。

二是光量子计算的研究应用。利用光子并行传输特性实现超高速计算，Lightelligence 的光子芯片在矩阵乘法中展现出比 GPU 快 1000 倍的潜力，曦智科技的 PACE 光子计算引擎已在金融风控领域实现商业化落地。量子计算与经典算力融合，IBM 量子系统 One 通过量子-经典混合算法优化供应链路径，计算效率较传统方法提升百万倍。

三是生物启发、神经形态计算。类脑芯片通过模拟神经元突触特性降低功耗，IBM 的 TrueNorth 芯片单芯片能耗仅 0.75mW，清华大学的天机芯实现全球首款异构融合类脑芯片，

支持自动驾驶实时决策。Intel 的 Neuromorphic 芯片 Loihi 2 模拟人脑 100 万神经元，功耗仅为 GPU 的千分之一，在模式识别任务中展现生物启发的自适应能力。

3.5.2 算法与算力的共生优化：智能驱动效率革命

一是动态计算架构。动态稀疏训练，如 OpenAI 的 Sparse Transformer 通过动态激活稀疏连接，将训练速度提升；谷歌 Pathways 系统支持混合精度训练，模型参数量缩减但仍保持极高的精度。模型并行化，如微软 DeepSpeed 框架实现万亿参数模型的分布式训练，通信开销降低。

二是 AI for AI。自动化机器学习 (AutoML)，如 Google Vertex AI 自动搜索最优模型架构，训练时间缩短；华为 MindSpore AutoML 实现端到端模型优化，推理延迟降低。以及神经架构搜索 (NAS)，如 Google 团队的 EfficientNet-B7 超越了之前最好的 GPipe 的精度，但是参数量少了 8.4 倍、速度快了 6.1 倍，相比于广泛使用的 ResNet-50，EfficientNets-B4 在相近的 FLOPS 下将 top-1 精度从 76.3% 提高到了 82.6%，且除了 ImageNet，EfficientNets 在其他数据集上表现也很好，在 8 个广泛应用的数据集中的 5 个实现了最先进的精度，然而比卷积网络的参数量减少了 21 倍^[24]。

三是算力运营优化与调度。算法层面优化策略如采用模型剪枝、知识蒸馏等技术，减少参数量。基于强化学习的动态资源分配，提升集群利用率。以及边缘协同，通过联邦学习实现边缘端与云端算力互补，降低数据传输延迟。调度平台方面，混合云管理将支持公有云与私有云的无缝切换，适应突发算力需求；而 AIOps 工具可以进行智能监控与故障预测，将节点故障恢复时间缩短至分钟级。

3.5.3 绿色算力与可持续发展：能源与技术的双重革新

液冷技术迭代。浸没式液冷，如以阿里云在杭州的云数据中心为例，该中心采用全浸没液冷技术，节能效果显著，整体节能率超过 70%，年均 PUE 低至 1.09，与传统数据中心相比，每年可节省电力高达 7000 万度，这一数字相当于西湖周边所有路灯连续点亮 8 年的总耗电量^[25]。中科曙光液冷服务器集群使用中高温水作为冷媒，针对主要发热源 CPU 和内存进行部件级精确制冷，散热效率更高，且可实现全年自然冷却，空调系统能耗降低 80%，但单机柜可实现高功率密度约 30kW，相比一般的传统数据中心，功率密度提高 3 倍以上^[26]。余热回收利用，如微软和 Fortum 将使用零排放电力为数据中心供电，同时会将废热引导到当地居民家中，预计可以满足该地区 25 万供热居民约 40% 的需求^[27]。

边缘计算网络。雾计算节点，如华为 Atlas 500 智能小站，其算力可达 16TOPS，在矿山设备数据本地处理中，延迟降至 5ms^[28]，这表明在特定应用场景下，数据本地处理率和传输带宽需求都有显著的优化空间。分布式算力网络方面，以 StarlingX 开源平台为例，它是一个基于开放源代码的边缘计算平台，提供了一系列高性能、高可靠性的基础服务。StarlingX 通过整合边缘节点资源，支持智能制造实时质检，端到端时延可压缩至毫秒以内。这种低时延的特性对于需要快速响应的应用场景至关重要。。

3.5.4 算力民主化：普惠与生态的重构

开源算力平台，模型即服务（MaaS）。Hugging Face 推出 Whisper 实时语音转写模型，开发者可通过 API 按需调用算力；Stability AI 的 Stable Diffusion 开源模型支持本地部署，降低创意产业创作门槛。低代码开发工具，如谷歌 Vertex AI Workbench 集成 AutoML 与 Jupyter Notebook，非技术人员可快速构建 AI 应用。

算力即服务（CaaS）。构建弹性算力池，如 AWS EC2 Spot 实例支持按秒计费，闲置算力利用率得到提升；阿里云弹性计算服务（ECS）提供百万级核的分钟级扩缩容能力。行业定制化方案，如百度智能云为农业提供“遥感+AI”病虫害监测服务，按亩收费模式降低农户使用成本。

3.5.5 AI 算力将对大模型训练和推理产生深远影响

（1）训练阶段：算力效率提升方面，类似 DeepSeek 等通过算法创新应用将大幅降低训练算力消耗，但 Scaling Law 仍主导需求增长，2023-2028 年智能算力 CAGR 达 46.2%。分布式训练优化方面，将更多采用参数服务器架构，支持千卡级集群并行训练，模型迭代周期也将大幅缩短。**（2）推理阶段：**边缘化部署，70% 的 AI 推理将下沉至边缘端，依赖低功耗芯片（如英特尔至强®Max 系列）。实时性要求，如金融风控、自动驾驶等场景需亚毫秒级响应，将加速推动存算一体芯片研发。**（3）开源模型推动普惠，**开源大模型通过降低训练成本（仅需闭源模型的 1/10），将加速技术下沉至中小型企业。

3.6 结语

智能进化的底层逻辑是算据（数据）、算法（模型）与算力三者共同作用的产物，三者相互依存、相互促进，共同构成 AI 技术发展的“铁三角”。若将 AI 比作一个“智能生命体”，数据是 AI 系统的“血液”，是算法训练和模型优化的基础；算法是 AI 的“大脑”，是实现智能决策的逻辑和方法论；算力是 AI 的“心脏”，是执行算法和处理数据的硬件与软件资源的动力源。它们通过“数据积累→算法创新→算力支撑→技术突破”的循环，推动 AI 从实验室走向产业应用。

当前，全球正经历一场“算力军备竞赛”：科技巨头投入巨资构建超大规模算力集群，各国竞相布局量子计算、光子芯片等前沿技术。然而，算力的未来不应止步于竞争，而应走向“生态共建”。从技术协同，如硬件（如存算一体芯片）、算法（如 AI for AI）、能源（如绿色计算）的深度融合，将推动算力效率指数级提升；到全球协作，打破数据孤岛与算力垄断，通过开源平台、分布式算力网络（如边缘计算）实现资源共享；再到政策引导，各国制定算力基础设施标准，平衡商业利益与公共福祉，避免算力鸿沟加剧全球不平等。未来的智能革命，应是由“硬件创新、算法突破、能源革命、政策护航”共同驱动，最终形成一个开放、可持续、包容的“智能生态”。

算力的稀缺性与能源消耗的争议，也迫使我们思考：如何让算力更高效、更绿色、更普惠？让算力普惠，让智能共生需要从企业端开放算力资源，降低中小企业与开发者的技术门

槛，避免“算力垄断”扼杀创新；在学术界应推动算法轻量化与绿色计算研究，让低功耗设备也能承载高阶智能；政府与国际组织需制定算力共享政策，投资公共算力平台，确保偏远地区与弱势群体也能享受智能红利；全社会正视算力的伦理挑战（如碳排放、数据隐私），在技术狂奔中守住可持续发展的底线。智能文明的终极目标是让技术服务于人，而非被算力束缚。从实验室到工厂，从城市到乡村，唯有让算力真正流动起来，AI 才能跨越“实验室的边界”，真正成为推动人类社会进步的“通用能量”。

本章参考文献

- [1] 中国通信工业协会数据中心委员会，《中国智算中心产业发展白皮书（2024 年）》，北京：中国信通院，2024。
- [2] 中国电信研究院，《智算产业发展研究报告(2024)》，北京：中国电信研究院，2024。
- [3] 沙利文、天罡智算.《2024 年中国智能算力行业白皮书》[M]. 2024.
- [4] SemiAnalysis 2024 年 9 月 4 日报告：Multi-Datcenter Training: OpenAI's Ambitious Plan To Beat Google's Infrastructure - Gigawatt Clusters, Telecom Networking, Long Haul Fiber, Hierarchical & Asynchronous SGD, Distributed Infrastructure Winners
- [5] SemiAnalysis 2025 年 3 月 26 日报告：The GPU Cloud ClusterMAX™ Rating System | How to Rent GPUs - 90%+ Coverage by Rental GPU Value, GPU Cloud Evaluation Guidelines, GPU Pricing Updates, GPU Bubble Burst, CoreWeave IPO, Hyperscalers, AI Neocloud Economics, Neocloud IRR
- [6] ReportLinker: Huawei and SiliconFlow: A Leap Towards China's AI Dominance?
- [7] 工业和信息化部等六部门关于印发《算力基础设施高质量发展行动计划》的通知 | 赣县区信息公开】
- [8] 工信部组织开展算力强基揭榜行动
- [9] 双 11 当天处理数据 5PB！揭秘 HiStore 数据库-阿里云开发者社区
- [10] 淘宝万亿级交易订单背后的存储引擎_云数据库 RDS(RDS)-阿里云帮助中心
- [11] 冲破 GPU 枷锁，DeepSeek + LPU 芯片能否颠覆英伟达算力垄断？
- [12] 摩尔线程算力布局：国产 GPU 的市场突围
- [13] 昇腾 910B 芯片的缔造者：突破封锁的“中国芯”，如何撑起 AI 算力脊梁？
- [14] 【行业动态】2025 年国产 GPU“大逃杀”：谁将登顶算力王座？
- [15] ASIC 芯片崛起：全球市场规模破百亿，未来增长势不可挡
- [16] Lightmatter：用光子计算终结 AI 的能耗危机 - 每时 AI
- [17] 量子计算赋能 AI：Azure Quantum 优化大模型训练中的梯度下降
- [18] Primus 框架：揭开分布式机器学习训练的神秘面纱-易源 AI 资讯 | 万维易源
- [19] 穆翔栩，贾明艳，王天成，等. 人工智能时代高质量算力发展趋势洞察[J]. 信息通信技

术与政策, 2025, 51(2): 18-23.

[20] AI 是“吃电狂魔”? 能有多耗电? 中国这个解法值得关注_未来 2%_澎湃新闻-The Paper

[21] 发展 AI 面临“算法碳足迹”困局, 如何破解?

[22] 双碳战略下, 数据中心的“冷”与“热”|服务器|英特尔|液冷|风冷|能耗_网易订阅

[23] 液冷技术——为数据中心节能降耗

[24] TAN M, LE Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks[J/OL]. 2019. DOI:10.48550/arxiv.1905.11946.

[25] 探秘数据中心的新宠——浸没式液冷技术

[26] 新一代一体化风液混冷先进数据中心 (C7000 2.0) -中科曙光

[27] 微软创建世界最大的数据中心废热回收计划, 将为芬兰南部居民供暖_腾讯新闻

[28] 2025 年智能化设备行业市场发展现状及未来发展前景趋势分析_中研普华_中研网

第4章 AI 赋能相关产业领域的典型应用场景

新一代人工智能技术的突破性发展，正以前所未有的深度和广度重构全球产业版图。本章聚焦工信部等七部门联合印发的《关于推动未来产业创新发展的实施意见》划定的战略领域，以智能制造、信息科技、先进材料、清洁能源、健康医疗、空间经济等六大主航道为观察切口，系统梳理人工智能技术驱动下的产业变革场景。通过剖析 AI 与各产业深度融合催生的新引擎、新场景、新业态，特别结合典型实践应用场景及前沿技术趋势，揭示人工智能如何通过“技术穿透力+产业纵深度”的双重突破，在关键领域形成高质量发展新动能。本章研究既是对国家未来产业战略的实践呼应，也为行业把握智能革命机遇提供系统性参考框架。

4.1 AI 赋能未来制造业的应用场景

人工智能正以前所未有的技术穿透力重构制造业全价值链，推动产业从传统制造向智能生态跃迁。通过深度学习、大模型技术与工业机理的深度融合，AI 已在研发设计、生产优化、质量管控、供应链管理等核心环节形成系统性突破，不仅实现生产效率提升与能耗降低，更催生出柔性生产、预测性维护等新型范式。在国家“人工智能+”战略引导下，智能制造正加速突破数据孤岛、算力瓶颈等挑战，通过工业互联网与多模态技术的协同创新，构建起人机共融、虚实交互的智能生产体系，为制造业高端化、绿色化、融合化转型注入核心驱动力。

当前，制造行业企业数字化转型正面临三重挑战：系统孤岛、流程僵化与人才缺口。据麦肯锡调研，超过 70% 的数字化转型项目未能达到预期目标，主要原因在于：（1）系统割裂：企业平均拥有 100+ 应用系统，但仅 30% 实现有效集成，导致数据流转受阻；（2）流程固化：传统自动化工具依赖预设规则，无法适应业务变化，维护成本高昂；（3）人才短缺：全球 AI 人才缺口超过百万，企业难以组建专业团队支撑数字化转型。这些挑战导致企业在数字化转型过程中面临效率低下、成本高企、创新乏力的困境。

因此，未来智能制造升级趋势之一是系统架构范式改变，即从流程自动化到认知自动化。传统 AI 基于规则的流程自动化，系统采用“单线程脚本驱动”架构，主要特点包括：（1）规则引擎具有依赖性，需人工预先定义操作步骤和判断条件，无法应对流程变更或异常情况。

（2）线性执行模式，按预设流程顺序执行，缺乏动态调整能力，对异常处理能力有限。这种架构在处理标准化、重复性任务时效率高，但面对复杂多变的业务场景时显得力不从心。

随着大模型技术的突破，以大模型为核心的 AI 智能体正在引领企业自动化从“流程驱动”向“认知驱动”（认知：人类对社会事务认知经验的总结）转变：（1）从预设规则到自主决策：传统 AI 需预先定义每一步操作，而智能体能理解业务语义并自主规划执行路径；（2）从单一任务到复杂场景：传统 AI 专注于结构化、重复性任务，智能体则能处理开放性问题与跨系统协作；（3）从工具到生产力主体：传统 AI 作为辅助工具提升效率，智能体则作为“数字员工”承担知识型工作。这一转变正重塑企业对自动化的认知，从“如何做”（How）到“做什么”（What）的范式升级。

未来智能制造升级趋势之二是基于大模型+多智能体的全数字化员工服务为主。采用"多智能体协同"架构，数字化员工核心特点包括：（1）多智能体分工协作：采用规划代理、执行代理与验证代理的分工模式，模拟人类"计划-执行-检查"的工作流；（2）动态模型调用：整合诸如 DeepSeek、Claude、QWQ 等国内外模型优势，通过 API 实现多模型动态选择；（3）虚拟执行环境：运行于隔离的虚拟机沙盒，可安全调用 100+数字工具，支持跨模态操作；（4）深入决策类场景：原先需要积累大量经验才能决策的场景，逐步被认知智能体代替和超越。

4.1.1 场景 1：智能生产调度与动态路径优化

AI 技术实现

该场景通过采用强化学习算法与多源数据（如行车位置、物料库存、订单优先级等），实现动态路径规划与多设备协同调度。核心逻辑在于实时数据驱动的决策优化，如利用激光扫描和数字孪生技术构建高精度三维库区模型，结合图像识别技术实时监控物料位置，再通过多智能体强化学习算法（如 Q-Learning 或深度确定性策略梯度算法）计算最优作业路径，减少空载率与冲突等待时间。

行业应用

国内 A 钢铁厂冷轧库区部署的 AI 系统不仅优化了 20 多台行车的协同作业，还通过边缘计算节点实现毫秒级响应。系统根据订单紧急程度动态调整优先级，如高附加值订单优先调度，同时结合历史数据预测未来物料流动趋势，提前规划资源分配。该方案将行车空驶率降低近一半，整体库区吞吐量提升三成以上。

行业影响

此类 AI 系统在矿业、汽车制造等领域同样适用，如 APS 生产排程系统支持多目标优化算法（如遗传算法），可灵活应对设备故障、订单变更等突发情况，实现年计划与实时调度的无缝衔接。

4.1.2 场景 2：基于大模型的设备预测性维护与智能运维

AI 技术实现

该场景通过对工业互联网实时监测，AI 大模型+行业知识库自动诊断工业设备异常，并自动生成维护方案，帮助企业大幅降低停机损失和维修成本。整个系统的工作流程如下：（1）**数据采集**：从设备传感器获取实时参数数据（温度、振动、电流等）；（2）**异常检测**：基于预设阈值检测参数是否异常；（3）**AI 深度分析**：当检测到异常时，启动大模型进行深度分析；（4）**知识检索**：AI 智能体主动搜索相关运维知识库，补充专业信息；（5）**生成维护计划**：根据知识库分析结果，自动生成详细的维护计划；（6）**报告生成与保存**：将所有信息自动整合为结构化报告并保存。

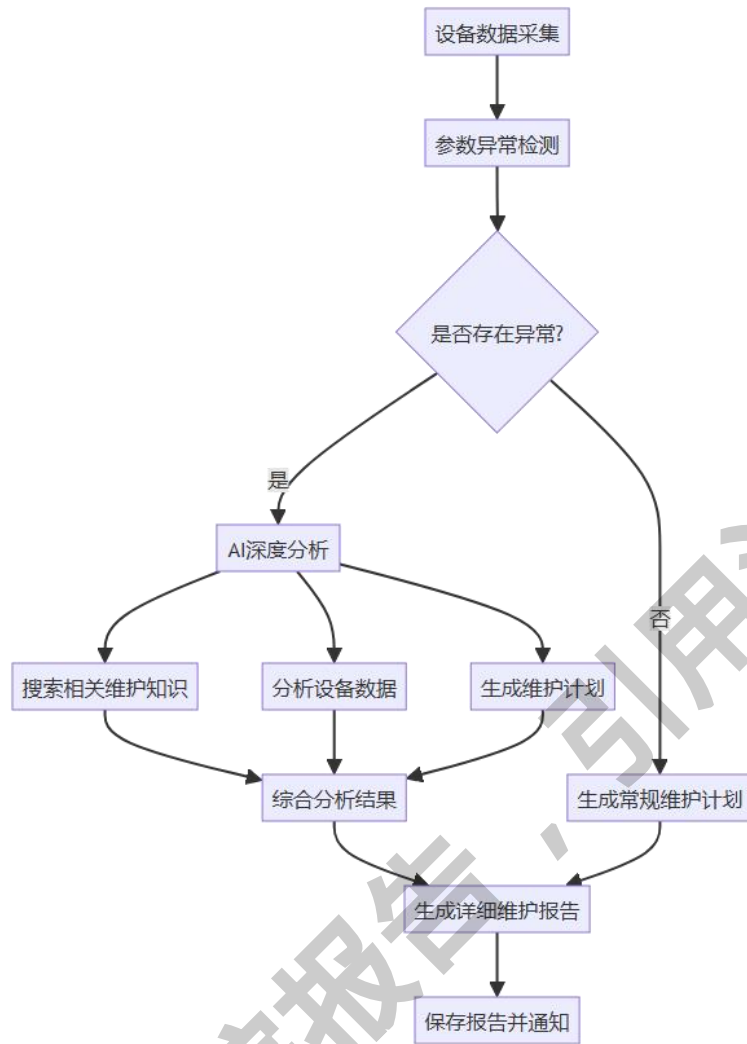


图 4.1 AI 驱动设备预测性维护与寿命管理系统的工作流程

大模型可以根据振动、温度、电流等多维度传感器数据自动拟合和生成相关的优化模型，输出相关的因子，自动生成相关的代码，实现对设备状态的动态预测。同时生成的结果根据人机对话的结果进行奖惩模型的训练，为知识库的积累打下良好的基础。

行业应用

半导体制造是当今世界最精密的工业之一，一条生产线上动辄数十台价值数百万美元的设备 24 小时不间断运转。然而，这些高精尖设备如何保障正常率成了大问题：（1）设备故障代价高昂：一台关键设备的意外停机可能导致整条生产线停产，每小时损失可达数十万美元（2）维护困境：过度维护浪费资源，维护不足又可能导致灾难性故障（3）专家依赖：设备状态分析严重依赖经验丰富的工程师，而这类人才稀缺且成本高昂（4）数据孤岛：大量设备数据被收集但未被充分利用，缺乏智能分析能力。

某半导体厂应用设备预测性维护后，非计划停机时间减少了 30% 以上，提前预警潜在故障，将被动维修转变为主动预防：（1）延长设备寿命：通过及时发现并处理小问题，避免演变成大故障，延长设备使用寿命（2）提高维护效率：精准定位问题，减少不必要的维

护工作，维护效率提升 42% (3) 降低专家依赖：AI 分析师提供专家级建议，减轻对高级工程师的依赖，新手工程师也能做出正确决策 (4) 数据价值挖掘：将沉睡的历史数据转化为有价值的决策依据，实现数据资产的增值。

行业影响

在所有制造行业里，设备预防性维护也面临共同性问题：被动维修模式导致高昂成本，大多数工厂采用“坏了再修”的被动维修模式。专业维修人员短缺，随着工业设备越来越复杂，真正懂设备的专家越来越稀缺。数据孤岛，缺乏智能分析，设备采集数据充分利用来预测故障和优化维护策略。维护知识难以传承，老师傅积累的宝贵经验往往难以系统化传承，新人培养周期长，经验断层严重。AI 驱动的设备预测性维护与寿命管理系统，将重塑精准制造的流程。

4.1.3 场景 3：供应链智能协同与风险对冲

AI 技术实现

多目标优化模型整合原材料价格、航运指数、产能利用率等数据，采用 NSGA-II 算法平衡运输成本、碳排放、交货周期等冲突目标。系统引入区块链技术确保供应链数据透明性，并利用蒙特卡洛算法模拟评估突发事件（如港口罢工）的潜在影响。

行业应用

国内 D 厂的多基地协同模型将订单分配效率大大提升，关键突破在于引入强化学习动态调整权重。如在碳排放约束收紧时，系统自动增加新能源运输车队占比，同时通过期货市场对冲原材料价格波动风险。

行业影响

从产业链来看，类似这种原材料供应链领域，AI 技术驱动的智能协同与风险对冲模型正在重塑行业格局。通过整合多目标优化模型与区块链技术，企业能够实现从采购、运输到生产的全流程智能化管理。这一模式不仅显著提升了订单分配效率，还帮助企业更好地应对碳排放政策约束和市场波动带来的挑战。

另外，这种创新模式推动了行业向低碳化与可持续发展转型。通过动态调整资源比例及优化运输路径，给类似的其他在降低运营成本同时还能结合金融工具比如期货市场进行风险对冲，有效平抑了原材料价格波动对企业利润的影响，增强了财务稳定性。此外，区块链技术的应用确保了供应链数据的透明性与可追溯性，从而提高了客户信任度，并为跨企业协作奠定了坚实基础。

4.2 AI 赋能未来信息产业应用场景

工信部等七部门联合印发《关于推动未来产业创新发展的实施意见》指出，未来信息产业方面要推动下一代移动通信、卫星互联网、量子信息等技术产业化应用加快量子、光子等计算技术创新突破，加速类脑智能、群体智能、大模型等深度赋能，加速培育智能产业。

人工智能作为底层驱动力，它本身又是未来信息产业的核心组成部分，正在重塑量子科技、类脑智能、下一代通信技术等领域的边界，与其他信息科技交叉融合发展，正推动全球产业格局的深刻变革。在信息领域，各国部署了人工智能、量子信息、下一代通信技术、半导体等；未来信息产业的核心在于技术融合与场景创新。我们相信人工智能与量子信息结合产生的变革是未来十年最重要的突破，结合中国自主可控下一代半导体芯片技术和通信技术，未来信息产业赋能千行百业，预判我国科技产业由“战略防守”进入“战略相持”阶段；帮助中国数年后在科技领域实现从“战略相持”到“战略反攻”。

4.2.1 场景 1：智能通信网络的动态优化

AI 技术实现

5G/6G 网络切片智能调度：基于强化学习算法，动态分配网络资源。如，在工业互联网场景中，当检测到某工厂车间的 VR 质检流量激增时，系统自动为该区域分配专属网络切片，保障 10ms 级时延要求。

频谱感知与动态分配：部署深度神经网络模型，实时分析电磁环境。如，在上海虹桥枢纽，AI 系统通过分析毫米波频段的干扰模式，将频谱利用率从 68% 提升至 89%。

智能边缘计算节点：构建云边协同架构，利用联邦学习技术训练分布式 AI 模型。如，深圳南山科技园部署的边缘节点，将视频分析任务的响应时间从 2.3 秒缩短至 0.8 秒。

行业应用

智能电网通信保障：国家电网在特高压输电线路部署 AI 巡检系统，通过无人机搭载的视觉识别模型，实现绝缘子破损检测准确率 99.2%，运维成本降低 42%。

车联网 V2X 优化：在杭州城市大脑项目中，AI 系统实时优化车路协同通信策略，使自动驾驶车辆的紧急制动响应时间从 0.7 秒降至 0.3 秒。

应急通信快速恢复：2024 年云南地震救援中，AI 驱动的卫星通信系统在 2 小时内重建灾区通信网络，较传统方式提速 70%。

行业影响

网络能效提升：华为 2025 年白皮书显示，AI 优化使 5G 基站能耗降低 28%，全国年节省电费超 30 亿元。

运维成本下降：中国移动引入 AI 故障诊断系统后，网络故障平均修复时间（MTTR）从 45 分钟缩短至 12 分钟。

新业务孵化加速：AI 赋能的网络切片技术催生了远程手术、工业 AR 等新业态，预计到 2028 年将创造 1.2 万亿市场规模。

4.2.2 场景 2：AI 在量子信息技术的应用

推动量子计算的加速器量子纠缠技术的突破离不开人工智能的支持。量子信息技术与人工智能的交叉融合，正成为全球科技竞争的制高点。量子信息是量子物理与信息技术相结合发展起来的新学科，主要包括量子通信和量子计算 2 个领域。量子通信主要研究量子密码、量子隐形传态、远距离量子通信的技术等。量子计算主要研究量子计算机和适合于量子计算

机的量子算法。从技术实现、应用场景、未来趋势及发展路径四个维度，剖析 AI 赋能量子信息技术的核心逻辑与战略价值。

技术实现

量子计算凭借其并行性与指数级算力优势，有望突破传统计算机的物理极限；而 AI 通过算法优化与数据驱动，可加速量子技术的研发与应用落地。

- AI 优化量子系统控制

量子比特的脆弱性和噪声干扰是量子计算实用化的核心障碍。AI 通过机器学习算法（如强化学习、贝叶斯优化）可实时校准量子比特参数，提升系统稳定性。

- 量子计算反哺 AI 算力瓶颈

经典 AI 模型的训练依赖海量算力，而量子计算的并行性可加速矩阵运算、优化神经网络训练。如，量子退火算法在组合优化问题中展现出显著优势，可将机器学习中的特征选择速度提升百倍以上。

- 量子 - 经典混合计算架构

当前阶段，量子计算机尚未实现通用计算能力，混合架构成为主流。AI 在此架构中扮演调度者角色，根据任务类型动态分配量子与经典计算资源。

行业应用

2024 年 12 月 9 日，Google 公布有 105 个物理量子比特的 Willow 量子芯片实现“纠错里程碑”，通过纠错技术首次演示了“低于阈值”，谷歌的 Willow 量子芯片通过 AI 驱动的纠错技术，将错误率降低至指数级水平，使其在“随机电路采样”任务中仅用 5 分钟完成经典计算机需 10^{25} 年的计算量。AI 还可优化量子门操作序列，降低量子电路深度，从而提升计算效率。

- 金融科技：风险建模与资产优化

量子计算结合 AI 算法，可高效解决蒙特卡洛模拟、投资组合优化等复杂问题。如，某银行采用量子退火算法优化资产配置，将风险分析时间从数小时缩短至秒级，同时提升收益率 15%。AI 还可通过强化学习动态调整量子模型参数，适应市场波动。

- 医药研发：分子模拟与靶点发现

药物分子模拟中，量子计算处理高维哈密顿量求解，AI 则负责数据预处理与结果分析，量子计算机可精确模拟分子电子结构，而 AI 加速候选药物筛选。如，辉瑞利用量子-AI 混合平台，将抗肿瘤药物研发周期从 5 年压缩至 18 个月，成本降低 40%。AI 还可通过生成式模型设计新型分子库，指导量子计算定向模拟，形成“AI 生成-量子验证”的闭环。

- 通信安全：量子密钥分发（QKD）优化

在量子通信领域，AI 可优化 QKD 协议的抗干扰能力。中国电信的天地一体化量子通信网络，通过 AI 实时分析信道噪声特征，动态调整光子编码策略，将密钥生成效率提升 30%。此外，AI 驱动量子随机数生成器，可增强加密系统的不可预测性。在实验中，研究人员

利用强化学习算法优化光子的传输过程，使光子的吸收效率提升了 60%以上。这一成果不仅适用于超导量子系统，其远程纠缠协议原则上可扩展至其他量子计算平台，为量子互联网的发展提供了重要硬件支撑。

行业影响

量子纠缠技术的突破标志着人类在量子计算领域迈出了重要一步，而 AI 技术则是推动这一进程的重要力量。通过学习和使用 AI 工具，我们可以更好地应对未来的挑战。

- 量子优越性的场景扩展

当前量子计算主要在特定任务（如随机电路采样、组合优化）中展现优势，未来将向通用型量子 AI 系统演进。如，OpenAI 的 o1 推理模型若与量子计算结合，可解决更复杂的逻辑推理问题，逼近通用人工智能（AGI）。

- 量子-神经形态计算融合

类脑智能与量子计算的结合可能催生新型计算范式。清华大学汪野教授提出，量子脉冲神经网络可模拟人脑的量子相干效应，实现更高能效的感知与决策。此类系统在自动驾驶、机器人领域潜力巨大。

- 量子云服务的普及

随着量子计算硬件的小型化与 AI 算法的轻量化，量子云平台将降低技术门槛。如 IBM 2016 年推出全球首个量子计算平台，2019 推出量子云服务，2023 年 12 月 4 日，发布了第一个 1000 量子比特量子芯片，IBM 使用分块联接的方式扩展量子比特，出现 2 个逻辑比特。亚马逊 Braket 平台已集成 AI 驱动的量子编程接口，开发者无需物理设备即可调用量子算力。中国“南京量子计算产业创新平台”亦计划开放量子-AI 联合开发工具链。

4.3 AI 赋能未来材料产业应用场景

人工智能技术加速推进，从分子尺度的精准调控，到系统级的全面设计与优化，推动材料行业从传统的“试错实验”模式向更高效的“计算设计”模式实现系统性革新。通过大语言模型、机器学习、第一性原理计算等跨学科技术的深度融合，人工智能已在高通量筛选、性能预测、逆向设计等核心领域取得显著突破。大量实践与数据表明，材料模拟与 AI 的结合不仅能够有效缩短研发周期，也显著降低了研发成本。

在国际层面，美国材料基因组计划（MGI）通过高通量计算与实验融合，加速了新材料的研发与产业应用；英国利物浦大学材料创新工厂和帝国理工的 ROAR 实验室致力于发展功能材料的自动化合成的智能设计方法；加拿大大学加速营则聚焦于整合跨学科资源和先进计算平台，通过协同研发和开放合作，加速材料科学的创新和应用转化；新加坡国立大学的功能智能材料研究所则着重于智能响应型材料，利用 AI 技术实现多维性能协同优化。

国内在该领域的发展虽起步较晚，但在国家对新材料产业的重视下，“十四五”规划中已明确强调要加快新材料产业与人工智能、大数据等新兴技术的深度融合，提高研发效能与成果转化效率。政府通过设立专项基金、提供税收优惠政策、推动产学研深度合作、建设新材

料产业基地建设等一系列措施积极扶持产业发展；同时倡导共享平台，开放数据库与标准化流程的建设，促进产业生态的完善。得益于政策支持与市场需求驱动，国内企业和研究机构持续增加在 AI 技术、自动化实验室与高通量设备的投入，初步成果已在电子、能源和生物医疗等领域逐步显现。尽管与自动化设备的成熟度和大规模高通量实验数据的积累方面与国际先进水平仍有差距，但未来前景广阔。

4.3.1 场景 1：新材料研发的智能加速

AI 技术实现

基于机器学习算法，构建材料性能预测模型。通过对海量材料数据（包括成分、结构、制备工艺等）的学习，模型能够快速预测新材料的性能，如强度、导电性、耐腐蚀性等。如，利用深度学习算法对材料晶体结构数据进行分析，预测材料在不同条件下的力学性能。同时，结合遗传算法等优化技术，对材料成分和制备工艺进行智能优化。在新材料研发过程中，通过模拟不同成分和工艺组合，快速筛选出最具潜力的方案，极大缩短研发周期。

行业应用

在航空航天材料领域，对高性能、轻量化材料的需求极为迫切。AI 助力研发新型航空铝合金材料，通过预测模型快速确定合金成分比例，优化制备工艺，使研发出的铝合金材料强度提升，重量减轻，满足航空部件对材料高强度、低密度的严格要求。如利用贝叶斯优化算法生成纳米结构材料，如加拿大团队开发的轻质高强材料，其强度接近钢但密度类似泡沫，通过虚拟测试与迭代优化实现应力均匀分布^[1]。创材深造利用 AI+高通量平台，将镍基高温合金研发周期从传统 5 年缩短至半年，材料性能达航空航天标准，已实现批量供货^[2]。

在半导体材料领域，研发高性能半导体材料时，AI 模型根据目标性能参数，快速筛选出合适的元素组合和生长条件，加速新型半导体材料的研发进程，为芯片制造等行业提供性能更优的材料支持。如美国 NIST 开发的 CAMEO AI 算法可自主发现新材料，无需人工训练；深势科技的 Uni-Mol 模型从 160 万候选分子中快速筛选出高能量转换效率的 OLED 材料，效率大幅提升；鸿海研究院通过强化学习（PPO/A2C 算法）优化碳化硅功率元件的制程参数，缩短研发周期等。

在新能源电池材料领域，AI 可加速电池材料的筛选与发现。电池材料的研发涉及成千上万种化学组合，而实验验证的时间和资源有限。AI 在高通量计算和机器学习方面的应用，使得研究者能够通过模拟和预测快速筛选出潜在的高性能材料。如微软和 PNNL 借助 AI 技术，筛选了 3200 万种潜在电池材料，并在 80 小时时间内将名单缩小到 23 种，其中 5 种是已知材料，团队表示如果使用传统方法获取这些材料，这个过程将耗时 20 多年^[3]。宁德时代通过 AI 筛选固态电解质组合，将锂离子迁移率预测效率提升，推动固态电池量产进程提速。

生物医药材料领域，AI 通过生成式模型（如大语言模型）可快速设计新型生物材料分子结构。2024 年年底，斯坦福大学等研究团队在《细胞》杂志发文称，多尺度、多模态的大型神经网络模型已经具备表示和模拟分子、细胞和组织在不同状态下行为的能力。在此基

基础上，AI 虚拟细胞有了高保真模拟、加速发现、指导研究的可信性。麻省理工学院研究团队以 3.9 万种化合物对 MRSA 的抗菌活性数据作为训练“脚本”，获得了抗菌能力的评估预测模型，随后以 3 个深度学习模型为基础，团队又“塑造”出化合物人类细胞毒性的“鉴定师”，对 1200 万种化合物进行“筛选”，最终获得能对抗 MRSA 又对人体安全的化合物^[4]。晶泰科技与微软合作开发的垂直领域大模型，结合机器人实验室数据，已成功应用于新材料设计案例开发，显著缩短了材料合成路径^[5]。英矽智能利用 AI 平台 Pharma.AI 优化药物分子结构，其技术逻辑可延伸至生物材料研发^[6]。AI 驱动的蛋白质设计平台（如 AlphaFold）可预测蛋白质结构，指导开发具有特殊功能的生物材料。以及 AI 整合患者基因组、代谢组等数据，定制化开发个性化生物材料。如，针对肿瘤异质性设计靶向药物递送载体，或根据患者免疫特征优化组织工程材料等。

行业影响

人工智能正在以“科学加速器”的姿态，重构材料研发的底层逻辑。这场革命不仅重塑了材料科学本身，更通过新材料的应用，对半导体、新能源、航空航天、生物医药等战略领域产生了连锁式影响。在数据实验环节，人工智能正逐步突破材料基因工程领域面临的多源数据融合与跨尺度计算难题，建立以“数据构建—模型聚合—高通量实验”为核心的三位一体闭环 AI 研发体系。这一体系将推动材料科学从传统“经验驱动”迈向“智能驱动”新模式。通过对多源异构数据的深入挖掘，以及结合高通量实验的自动化调度与精确控制，有效克服材料领域特有的数据稀缺、数据质量参差和高维计算瓶颈，实现对材料配方、工艺及性能的深度智能解析和预测优化。在研发效率方面，传统新材料研发周期往往长达数年甚至数十年，涉及大量试错过程。借助 AI 技术，研发周期可大幅缩短，提高研发效率，降低研发成本。在创新能力提升方面，AI 能够探索传统方法难以触及的材料领域，发现新的材料性能和应用潜力，推动材料创新，为行业发展注入新动力。

4.3.2 场景 2：材料生产过程的智能优化

AI 技术实现

利用传感器实时采集材料生产过程中的数据，如温度、压力、流速等。通过数据分析和机器学习算法，构建生产过程模型，实现对生产过程的实时监测和智能控制。如，基于神经网络算法的控制系统，根据实时生产数据调整设备参数，确保生产过程稳定，产品质量一致。同时，运用预测性维护技术，通过对设备运行数据的分析，提前预测设备故障，安排维护计划，避免生产中断。

行业应用

在钢铁生产过程中，AI 系统展现出显著的优化效果。在高炉炼铁环节，AI 系统实时监测高炉炼铁过程中的温度、压力等数参数。当炉内温度出现异常波动时，AI 系统迅速分析数据，自动调整送风系统的阀门开度，精确控制鼓风量，同时调整加料系统的加料速度和种类，确保炉内化学反应稳定进行。经过 AI 优化后，铁水质量稳定性提高，铁水成分的偏差

范围缩小，能源消耗降低。如，某大型钢铁企业在引入 AI 系统后，每月因铁水质量提升，减少了次品钢材产量，节约了能源成本。

在化工材料生产中，以聚合反应过程为例，AI 发挥着重要作用。在聚合反应釜上安装的传感器实时采集反应温度、压力、反应物浓度等数据。AI 模型根据这些数据，精确控制反应釜的加热、冷却系统，调整搅拌速度，以及反应物的进料速度和比例。通过对聚合反应过程的精准控制，产品的纯度提高，产品质量一致性得到显著提升，次品率降低。如，某化工企业生产的聚丙烯产品，在 AI 优化后，产品批次间性能差异减小，在塑料制品加工企业中的应用效果得到显著改善，产品市场竞争力增强，企业经济效益大幅提升。

行业影响

在生产效率提升方面，AI 优化后的生产过程减少了人工干预，提高了生产自动化程度。传统材料生产过程中，人工操作响应速度慢，且容易受到人为因素影响。AI 系统能够实时、快速地对生产数据做出反应并调整设备参数。AI 优化后的材料生产过程，生产效率可提升，产能得到增加。**在质量控制方面**，通过实时监测和精准控制，产品质量稳定性显著提高。传统生产方式难以实现对生产过程的全流程精准控制，产品质量波动较大。AI 系统能够对生产过程中的每一个环节进行精细化控制，确保产品质量的一致性。次品率降低，提升了企业产品在市场中的竞争力。如，某汽车零部件制造企业采用 AI 优化铝合金压铸生产过程后，产品次品率降低，每年因次品减少节约成本，同时产品质量提升也增强了企业在汽车零部件市场的口碑。**在节能减排方面**，AI 对生产过程的优化减少了能源浪费。通过精准控制设备运行参数，避免了设备的过度运行和能源的无效消耗。以钢铁行业为例，AI 优化后的生产过程能够降低能源消耗，减少二氧化碳等污染物排放。这不仅符合国家可持续发展要求，也为企业降低了生产成本，提升了企业的社会形象和环保竞争力。

4.3.3 场景 3：材料性能监测与寿命预测

AI 技术实现

采用无损检测技术（如超声波、X 射线等）结合机器学习算法，对材料性能进行实时监测。通过对检测数据的分析，判断材料是否存在缺陷、性能是否发生变化。同时，利用大数据分析和深度学习模型，对材料的使用环境、使用历史等数据进行分析，精准预测材料的剩余寿命。如，基于长短期记忆网络（LSTM）的模型，能够根据材料在不同时间点的性能数据，准确预测其未来寿命。

行业应用

在桥梁、建筑等基础设施领域，AI 在材料性能监测和寿命预测方面发挥着重要作用。以大型桥梁为例，在桥梁的关键承重部件，如钢梁、桥墩等部位安装应变片、加速度传感器以及无损检测设备，实时采集钢材和混凝土的应力、振动数据以及内部缺陷信息。AI 系统对这些数据进行实时分析，当发现钢材应力超过安全阈值或混凝土出现裂缝扩展迹象时，及时发出预警。同时，通过对多年来材料性能数据的分析，AI 能够预测桥梁材料的剩余寿命，

提前制定维护和修复计划。如，某跨江大桥在引入 AI 监测系统后，提前发现了部分钢梁的疲劳裂纹隐患，及时进行了修复，避免了可能发生的重大安全事故。

在汽车制造领域，AI 同样为汽车零部件材料性能监测和寿命预测提供科学依据。在汽车发动机、变速器等关键零部件上安装传感器，实时采集零部件在运行过程中的温度、压力、振动等数据。AI 系统根据这些数据，分析零部件材料的性能状态。如，通过对发动机活塞材料温度和应力数据的分析，预测活塞材料的磨损情况和剩余寿命。根据 AI 预测结果，汽车制造商能够为车主提供更精准的汽车保养和维修建议，如提前更换即将失效的零部件，提高汽车安全性和可靠性。如，某汽车品牌在引入 AI 监测系统后，因零部件提前更换而减少的车辆故障召回事件降低，提升了品牌形象和用户满意度。

行业影响

在安全性保障方面，AI 实现对材料性能的实时监测和寿命预测，及时发现安全隐患。传统的材料检测方式多为定期抽检，难以发现材料在日常使用过程中的细微变化和潜在缺陷。AI 的实时监测功能能够在材料性能出现异常的第一时间发出预警，有效避免因材料失效导致的安全事故，保障了人民生命财产安全。**在维护成本降低方面**，通过准确预测材料寿命，合理安排维护计划，避免过度维护和维修不足。传统维护方式往往采用定期维护策略，导致部分材料在未达到使用寿命时就进行了更换，造成资源浪费；而部分材料因维护不及时，导致设备故障，维修成本大幅增加。AI 预测技术能够根据材料实际寿命情况，精准安排维护时间和内容，优化维护计划，降低维护成本。

4.4 AI 赋能未来能源产业应用场景

在全球经济持续增长与人口不断攀升的双重驱动下，能源需求正以前所未有的速度激增，传统化石能源的有限性与环境承载能力之间的矛盾愈发尖锐，能源转型已成为全球可持续发展的必然选择。在此背景下，核能、核聚变、氢能、生物质能等清洁能源因其绿色、低碳、可持续的特性，被视为破解能源与环境困境的关键路径。与此同时，新型高效太阳能电池和新型储能技术的加速发展，为能源行业的深度变革注入了新的活力。

AI 技术作为当今科技领域的前沿力量，正以迅猛之势重塑各个行业的发展格局，并在能源领域展现出巨大的应用潜力和变革性影响。其强大的数据处理、模式识别和智能决策能力，能够对海量的能源数据进行快速分析和深度挖掘，实现能源系统的智能调度、精准预测与故障诊断，从而显著提升能源利用效率，降低生产成本，增强能源系统的稳定性与安全性。AI 技术的智能化、自动化特点，完美契合了未来能源在生产、存储、运输、应用等环节对高效、精准、安全的迫切需求，为解决未来能源发展的技术瓶颈和挑战提供了创新性的思路和方法，有力地推动了全球能源行业的智能化转型，助力构建一个清洁低碳、安全高效的现代能源体系，为实现全球可持续发展目标提供坚实的技术支撑。

4.4.1 场景 1：核能（核裂变）

核能作为一种低碳、稳定的能源，通过核裂变反应释放出大量的热能，进而转化为电能

或其他形式的能源。其主要优点包括能源密度高、供应稳定、不排放二氧化碳等温室气体。核能的产业应用现状十分广泛，不仅为全球提供大量清洁电力，还在医疗、工业供热、海水淡化等多个领域发挥着重要作用。中国目前是全球第三大核能生产国，拥有全球最多的在建核反应堆。核能的技术路线多样，涵盖了不同的反应堆类型和应用模式。压水堆（PWR）和沸水堆（BWR）是当前最主流的两种技术，高温气冷堆（HTGR）、快中子反应堆等技术也在不断发展，小型模块化反应堆（SMR）因其灵活性和较低的初始投资成本，成为当前核能技术发展的重要方向。这些技术路线的发展不仅丰富了核能的应用场景，也提高了核能的经济性和安全性，进一步推动了核能产业的可持续发展。

然而，核能的发展也面临一些挑战。首先，高初始投资成本是核能发展的一大障碍。核能发电厂的建设需要巨额的资金投入，包括场地准备、工程设计、设备制造、建设施工、调试以及融资等多个环节。这些成本使得核能项目在经济上面临较大的压力。其次，公众对核能的安全性和核废料处理存在担忧。核事故的潜在风险以及核废料的长期放射性，使得公众对核能项目的接受度较低。此外，确保稳定的核燃料供应以及妥善处理核废料，是核能产业需要解决的重要问题。

AI 技术实现

AI 技术通过多种方法针对性地解决核能发展中的痛点。在设计在建设阶段，AI 利用智能建模和优化算法，能够快速迭代核电站设计方案。同时，AI 在项目管理中通过精准预测和控制，优化成本和进度管理。在运行与维护方面，AI 通过机器学习算法实时分析传感器数据，实现对核反应堆等关键设备的健康状态监测和故障预测。此外，在核燃料管理中，AI 通过数据分析优化燃料使用和循环，提高燃料利用率并减少核废料的产生。AI 还用于核废料处理设施的智能管理，确保核废料的安全存储和运输，同时优化核燃料供应链的管理，通过智能监测技术确保燃料质量和安全性。

行业应用

EDF（法国电力公司）开发了一个 AI 驱动核维护平台，该平台利用机器学习算法来预测设备故障。这种预测性维护系统能够减少非计划停机时间，提高安全性，并降低运营成本。通过实时分析设备的运行数据，AI 系统可以提前识别潜在的问题，从而在问题发生之前进行维护。

行业影响

AI 技术的应用对核能行业产生了深远的影响。首先，AI 通过优化设计和项目管理，显著降低了核能项目的初始投资成本，缓解了高成本的挑战。其次，在运行与维护阶段，AI 通过实时监测和故障预测，提升了核能设施的运行效率和安全性。此外，AI 在核燃料管理中的应用推动了核能行业的可持续发展，减少了核废料的产生，优化了资源利用。在核废料处理方面，AI 确保了核废料的安全存储和运输，有效缓解了公众对核安全的担忧。最后，在核燃料供应链中，AI 优化了采购和运输效率，提升了燃料供应的技术水平，为核能行业

的长期发展提供了强有力的支持。

4.4.2 场景 2：核聚变

核聚变是轻原子核在高温高压下克服电荷排斥力合并成更重的核并释放巨大能量的过程，是太阳等恒星能量的来源。地球上，科学家致力于在受控环境中实现核聚变，以期作为未来清洁能源。目前，核聚变处于实验研究阶段，国际热核聚变实验堆（ITER）项目正在法国建设，计划 2036 年开始全面功率实验。全球多个国家积极开展核聚变研究，取得了接近科学增益的实验成果。

然而，核聚变的发展过程中面临多个核心问题。首先，等离子体的控制和维持是核聚变研究的关键挑战之一。在磁约束核聚变装置（如托卡马克）中，高温等离子体容易出现不稳定性，如撕裂模不稳定性，这可能导致等离子体逃逸并中断反应，同时对装置内部材料造成损害。其次，开发能够承受极端热负荷和中子辐射的材料也是一个重要问题。核聚变装置内部的材料需要在高温、高辐射环境下保持结构完整性和功能性，这对材料科学提出了极高的要求。此外，核聚变装置的运行需要高辅助功率，用于加热系统和磁体的励磁，这限制了装置的能效和经济性。最后，核聚变装置的运行和维护成本高昂，且技术复杂性高，需要大量的专业知识和经验。

AI 技术实现

AI 技术在核聚变研究中展现出巨大的应用潜力，能够有效解决上述核心问题。首先，在等离子体控制方面，AI 可以通过深度强化学习算法实时预测和避免等离子体的不稳定性。如，普林斯顿大学的研究团队利用 AI 成功预测了托卡马克反应堆中的撕裂模不稳定性，并提前 300 毫秒采取措施进行干预，从而维持等离子体的稳定。这种 AI 驱动的控制方法不仅提高了等离子体的稳定性，还减少了对人员经验和专业知识的依赖。其次，在材料开发方面，AI 可以通过机器学习模型加速新材料的评估和优化。AI 能够整合不同实验、模拟和制造过程的数据，生成可靠的预测模型，帮助科学家快速筛选出适合核聚变装置的材料。此外，AI 还可以通过模拟和优化核聚变装置的设计，提高装置的能效和经济性。如，AI 可以优化磁体的配置和运行参数，降低辅助功率的需求。最后，在运行和维护方面，AI 可以实现对核聚变装置的实时监测和故障预测，降低运行成本并提高安全性。

行业应用

Argonne 国家实验室的研究人员开发了一种结合 PRO-AID 诊断工具和大型语言模型（如 GPT-4）的系统，用于核反应堆的故障诊断和解释。该系统在 Mechanisms Engineering Test Loop Facility (METL) 成功诊断出一个故障传感器，并向操作员解释了问题。

行业影响

AI 为核聚变行业带来了前所未有的突破性进展，正以惊人的速度重塑行业未来。首先，AI 通过实时预测和干预等离子体的不稳定性，显著提高了等离子体的稳定性，为核聚变反应的持续进行提供了保障。其次，在材料开发中，AI 加速了新材料的筛选和优化，缩短了

研发周期，降低了研发成本。此外，AI 通过优化装置设计，提高了核聚变装置的能效和经济性，为核聚变能源的商业化提供了可能性。在运行和维护方面，AI 通过实时监测和故障预测，降低了运行成本并提高了装置的安全性，为核聚变行业的长期发展提供了强有力的支持。

随着技术的不断迭代升级，AI 将进一步优化核聚变能源的生产流程，降低成本，提升效率，使其在全球能源市场中占据举足轻重的地位。核聚变能源的大规模商业化应用已不再是遥不可及的梦想，而是触手可及的现实，AI 将成为这一伟大征程中不可或缺的“加速器”。未来，AI 与核聚变的深度融合将持续释放巨大潜力，有望彻底颠覆传统能源格局。

4.4.3 场景 3：氢能

氢能是一种清洁、高效的二次能源，通过氢气的燃烧或化学反应产生能量，其核心原理是氢气与氧气反应生成水，同时释放出大量的热能和电能。这一过程不仅高效，而且环境友好，反应产物仅为水，不产生任何温室气体或污染物。氢能的来源广泛，可以从水中通过电解、从化石燃料中通过重整等技术提取。由于其高能量密度、可储存性和可运输性，氢能被认为是未来能源结构中不可或缺的一部分，尤其在交通、储能和分布式发电等领域具有巨大的应用潜力。全球多个国家制定了氢能发展战略，推动氢能技术的研发和应用。如，中国在东北、华北北部和西北地区加速部署可再生能源和绿色氢能，以实现当地社会的低碳、绿色和可持续发展。

制约氢能发展的主要因素来源于几个方面，首先，氢气的生产成本较高，尤其是绿氢的规模化生产仍面临技术瓶颈。目前，氢气的生产主要依赖化石燃料重整和电解水等技术，其中电解水制氢技术在可再生能源领域的应用前景广阔，但成本较高。其次，氢气的存储和运输技术有待进一步突破，以提高安全性和经济性。氢气的低密度和高活性使得其在存储和运输过程中容易发生泄漏和燃烧等安全事故。此外，加氢基础设施不足也是制约氢能普及的关键因素之一。

AI 技术实现

AI 在氢能领域展现出巨大潜力，可有效破解发展难题。首先，在氢气生产方面，AI 可优化电解水制氢过程的参数，提高制氢效率，降低生产成本。通过机器学习算法，可以对电解水制氢过程中的电流密度、温度、压力等参数进行实时监测和优化，提高制氢效率。如，AI 可以通过分析历史数据和实时数据，预测最佳的电解槽操作条件，从而减少能源消耗和设备磨损，进一步降低生产成本。此外，AI 还能优化生产计划，根据市场需求和能源价格动态调整生产规模，实现资源的合理配置。

在氢气存储方面，利用 AI 开发更高效的存储材料和方法，提高存储密度和安全性。如，通过 AI 模拟和预测不同材料的储氢性能，加速新型储氢材料的研发进程。AI 可以帮助筛选和设计具有高储氢容量、快速吸放氢速率和良好循环稳定性的材料，减少研发成本和时间。

此外，AI 还可用于氢气运输过程的智能监控和优化，确保运输过程的安全和高效。通

过智能传感器网络和数据分析，实时监测氢气运输过程中的压力、温度和泄漏风险，及时采取措施预防事故发生。如，AI 可以预测运输管道中的压力变化，提前调整运输参数或发出警报，避免潜在的安全隐患。

最后，在加氢基础设施规划方面，AI 可以通过数据分析和预测模型，优化加氢站的布局 and 运营，降低建设成本，提高服务效率。AI 可以综合考虑人口密度、交通流量、能源需求等因素，确定加氢站的最佳位置和规模，同时预测未来的市场需求，为基础设施的扩展和升级提供决策支持。

行业应用

AI 加速绿色氢气生产：研究人员利用 AI 分析了超过 36,000 种金属氧化物组合，成功筛选出一种高效的催化剂合金（钨、铬和钛的合金），其性能远超传统材料。

AI 优化固体氧化物电解器：德国的研究人员利用 AI 模型改进了固体氧化物电解器的性能，使其在高温下运行更高效，降低了制氢成本。

AI 预测性维护：研究人员开发了基于 LSTM 的预测性维护系统，用于氢气设施的维护，准确率达到 97%。这种系统可以提前预测设备故障，减少停机时间和维护成本。

AI 预测氢气生产：研究人员利用 XGBoost 模型预测氢气生产，达到了 R2 值 0.9996 的高精度，显著提高了运营决策的准确性。

行业影响

AI 赋能氢能的行业影响展现出令人振奋的前景。AI 通过优化制氢参数和生产计划，不仅显著降低了制氢成本，还提高了能源利用效率，为氢能的大规模商业化铺平了道路。在氢气存储和运输方面，AI 通过智能监控和故障预测，增强了系统的安全性和稳定性，为氢能设施的安全运行提供了可靠保障。此外，AI 在加氢基础设施规划中的应用优化了加氢站的布局 and 运营，提高了服务效率，降低了建设成本，为氢能的普及提供了强有力的支持。AI 还推动了氢能行业的可持续发展，加速了氢能替代传统化石能源的进程，促进了能源结构的优化和绿色转型。展望未来，AI 有望成为氢能行业创新的核心驱动力，引领氢能技术的革命性突破，为全球能源转型注入无限潜力，开启一个更加清洁、高效和可持续的能源新时代。

4.4.4 场景 4：生物质能

生物质能是唯一可转化为气、液、固多种能源产品的可再生能源，其原料包括农作物废弃物、林业剩余物、畜禽粪便等。通过热化学或生物化学方法，生物质能可转化为生物燃料、生物天然气、电力等多种能源形式。如，生物质气化技术通过高温反应生成合成气，可用于发电或合成液体燃料；生物质液化则将生物质转化为生物油，进一步加工可得到高附加值的液体燃料。全球生物质能产业近年来取得显著进展，应用范围广泛，涵盖发电、供热、交通燃料等多个领域。然而，不同生物质能技术的成熟度存在差异。生物质发电技术相对成熟，已实现大规模商业化应用；生物燃料生产技术（如生物乙醇和生物柴油）也较为成熟，但在一些地区仍面临原料供应和生产成本的挑战；而一些先进的生物炼制技术（如将生物质转化

为高附加值的化学品和材料）仍处于研发或示范阶段。

AI 技术实现

生物质能的发展面临多方面的挑战。原料收集和预处理成本较高，且存在资源分散、供应不稳定等问题。如，农作物秸秆分布在广大农村地区，需要投入大量人力物力进行收集和运输。针对这一问题，AI 通过分析地理信息、农作物产量数据和物流成本，优化生物质原料的收集和运输路径，降低物流成本。同时，AI 还可通过监测原料质量，提前筛选出适合能源转化的优质原料。

此外，生物质能转化过程中的能源效率和环境影响也需要进一步优化。如，生物质气化过程中存在气体质量不稳定、低热值、高 CO 含量和高焦油含量等问题，限制了其大规模应用。AI 能够实时监测和控制反应条件，提高能源转化效率和产品质量。如，在生物质气化过程中，通过传感器获取温度、压力和气体成分等数据，利用 AI 算法动态调整气化剂的配比和反应温度，优化合成气的质量和产量。

同时，生物质能的利用过程可能对环境产生负面影响。如，生物质燃烧可能释放颗粒物和氮氧化物等污染物；生物燃料生产过程中使用的化肥和农药可能对土壤和水体造成污染。AI 技术用于监测生物质能项目对环境的影响。如，通过无人机和卫星遥感技术，实时监测生物质种植区的土壤质量、水资源利用和植被覆盖情况，评估生物燃料生产对生态环境的影响。

最后，生物质能产业的发展需要政策的持续支持和市场的稳定需求。然而，在一些地区，政策的不稳定性和市场的不确定性给生物质能项目的投资和运营带来风险。AI 通过分析市场数据和政策动态，预测生物质能产品的需求和价格走势，帮助企业在生产计划和销售策略上做出更合理的决策。同时，AI 还可评估政策变化对生物质能项目的风险影响，为企业提供风险预警和应对建议。

行业应用

Woodchuck.AI 是一个气候影响初创公司，正在通过 AI 技术革新废物转化为能源的进程。该公司正在密歇根州大急流城建立一个开创性的 AI 创新环境和生物质处理设备。Woodchuck.AI 的专有 AI 平台旨在优化国家生物质供应链，通过高效地将木材从废物流中转移，创造一个一致、可靠的生物质供应。这一创新不仅减少了垃圾填埋场的使用，还为绿色能源电网提供了支持。Woodchuck.AI 的 AI 平台通过解决关键的废物减少和可持续性挑战，确保生物质行业有一个可靠和可持续的优质原料供应，这对于实现各行业的净零目标至关重要。

行业影响

展望未来，AI 有望成为生物质能行业的核心驱动力，引领技术创新，在原料收集、供应、转化以及市场预测方面助力生物质能产业在全球能源转型中发挥更大作用。随着技术的不断进步和应用的深入，AI 将为生物质能行业带来更多的机遇和可能性。

4.4.5 AI 赋能未来能源的挑战与应对

技术融合的复杂性

AI 与未来能源技术的融合面临着诸多挑战，其中兼容性问题首当其冲。不同的能源系统和设备往往采用各自的技术标准和通信协议，导致数据传输和系统集成困难重重。以智能电网为例，其涉及众多供应商提供的各类设备，这些设备的通信接口和数据格式各不相同，使得 AI 系统难以实现对整个电网的统一管理和优化调度。为解决这一难题，建立统一的技术架构和接口规范显得尤为重要。政府和行业组织应牵头，联合能源企业、设备制造商和科研机构等各方力量，共同制定适用于 AI 与能源系统融合的技术标准。通过明确数据传输格式、通信协议以及系统接口要求，确保不同设备和系统之间的无缝对接和协同工作，从而提高 AI 技术在能源领域的应用效率和可靠性，为未来能源系统的智能化升级和可持续发展奠定坚实基础。

数据安全与隐私保护

数据的安全性和隐私保护在 AI 赋能未来能源的过程中也是至关重要的问题。能源数据往往涉及国家能源安全、企业商业机密以及用户个人隐私等敏感信息。如，智能电表所采集的用户用电数据，能够反映出用户的生活习惯和消费模式，一旦泄露，将对用户隐私造成严重侵害。为了保障数据的安全性和隐私性，能源企业和相关机构应采用先进的加密技术和严格的访问控制手段。对传输中的数据和存储在数据库中的数据进行加密处理，确保数据在各个环节的安全性。同时，建立完善的访问控制机制，仅允许授权人员访问敏感数据，并对数据访问行为进行严格审计和监控，防止数据泄露和滥用事件的发生。通过这些措施，可以在保护数据安全和隐私的前提下，充分挖掘能源数据的价值，为 AI 技术在未来能源领域的深度应用和创新发展提供有力支持。

人才短缺与培养

未来能源与 AI 交叉领域的快速发展，使得该领域面临严重的人才短缺问题。当前，既懂能源专业知识又掌握 AI 技术的复合型人才极为稀缺，这严重制约了 AI 在能源领域的深入应用和创新发展。如，在核聚变研究中，需要既熟悉核物理又擅长 AI 算法研发的专业人才来推动等离子体控制和材料研发等关键领域的突破。为缓解这一人才瓶颈，教育和培训体系需要进行相应的改革和完善。高校应加强跨学科教育，开设能源与 AI 相关的交叉学科专业和课程，培养具有跨领域知识和技能的复合型人才。同时，企业也应加大内部培训力度，为员工提供持续学习和技能提升的机会，鼓励员工跨领域发展，以适应未来能源行业对人才的新需求。通过培养和吸引更多优秀的复合型人才，能够加速 AI 技术与未来能源的深度融合，为能源行业的智能化转型和可持续发展注入新的动力。

法律法规与监管滞后

AI 在能源领域的广泛应用引发了一系列法律法规和监管方面的问题，而现有的法律框架和监管政策在很多方面存在滞后性和空白。如，AI 系统的决策过程缺乏透明度，可能导

致责任界定不清；数据的共享和使用也存在法律上的模糊地带，容易引发数据滥用和不正当竞争等问题。面对这些问题，政府和行业组织应加快制定和完善政策法规和标准体系。明确 AI 在能源领域的应用规范，包括数据的采集、存储、使用和共享等方面的法律要求，以及 AI 系统的安全性和可靠性标准。同时，加强监管力度，确保 AI 技术在能源行业的应用符合法律法规和伦理道德要求，保障市场的公平竞争和公众利益。通过建立健全的法律法规和监管体系，可以为 AI 赋能未来能源的发展创造一个稳定、公平、透明的市场环境，促进能源行业的健康、可持续发展。

4.4.6 AI 赋能能源的未来展望

展望未来，AI 技术在能源领域的应用将呈现出更加多元化和深入化的发展趋势。更先进的深度学习算法将不断涌现，能够处理更复杂、更海量的未来能源数据，挖掘出更深层次的能源系统运行规律。强化学习技术将在能源系统的优化控制中得到更广泛地应用，通过与环境的实时交互和反馈，实现对能源生产、传输和消费过程的动态优化。此外，AI 与其他新兴技术的融合也将成为未来的发展方向，如物联网、区块链等技术与 AI 的结合，将为未来能源行业带来更强大的智能化能力。

AI 的赋能将推动未来能源行业发生深刻变革，带来前所未有的机遇。能源生产模式将不断创新，从传统的集中式生产向分布式、智能化生产转变。能源市场也将重构，形成更加灵活、高效的能源交易机制，促进能源资源的优化配置。同时，能源服务将得到全面升级，通过 AI 技术实现个性化、智能化的能源管理和服务，提升用户体验和能源利用效率。这些变革将为相关企业创造新的商业机遇和发展空间，推动能源行业的转型升级。

4.5 AI 赋能未来健康产业应用场景

健康领域，中国正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。

相较于传统的医疗模式，AI 凭借其卓越的技术优势，正在深刻地改变着大健康产业的各个领域^[1]：从疾病预测到防控，从诊断到治疗，从药物研发到健康管理，从医院管理到医学研究，AI 的应用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^[2-3]。

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^[4]，旨在构建覆盖全民、全生命周期的健康服务体系，这也彰显了国家在实现全民健康目标上的坚定决心和强大执行力。要实现医疗资源普惠和全民健康的伟大战略，解决医疗资源紧张、医疗服务质量差异化的问题，就需要 AI 这个强大的助推力，包括全民健康数据的整合、AI 在基层医疗的应用、预防医学的强化、健康不平等的减少、政策支持和技术基础设施的建设等，重塑健康权利分配，最终实现全民共享。

AI 驱动个性化健康革新。通过深度整合个体基因、行为、环境数据、健康数据、医疗

记录和科研成果等^[5]，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。

AI 医疗大模型密集发布。观察 2025 年国内 AI 赋能医疗的有影响力的参与企业：华为于 2025 年 2 月推出基于华为 DCS AI 解决方案的瑞金病理大模型，希望通过在病理诊断领域，应用 AI 有效缓解我国病理医生短缺的现状，显著提升病理切片检查效率和诊断准确率；京东健康于 1 月 11 日发布了基于其线上全域场景的大模型全系产品“AI 京医”，以及业内第一个面向医院全场景应用的大模型产品“京东卓医”；此外 1 月我国科学家开发了用于辅助疾病诊断的通用大语言模型——MedFound，并在 Nature Medicine 期刊发布。

以下是 AI 在健康产业中的几个典型应用场景：

4.5.1 场景 1：医疗诊断与影像分析

AI 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）、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^[6]。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。

行业应用

AI 赋能疾病检测。谷歌与印度 Aravind 眼科医院合作的深度学习模型--AI 眼科工具 Aravind，通过 CNN 模型（如 Inception-v3、EfficientNet）实现了眼底图像的自动分析和分类：通过自动检测眼底图像中的病变（如微血管瘤、出血、渗出等），输出糖尿病性视网膜病变的存在和严重程度分级（如轻度、中度、重度），准确度高达 97.5%^[7]。该系统已在全球多个地区落地，显著提升了糖尿病性视网膜病变筛查的效率和覆盖率。

AI 赋能医学影像诊断。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 分钟判断是否大血管闭塞，而传统方法大概要一个小时。对 139 家医院的 2544 名患者的 CTA 影像进行分析，表现出 96%的敏感性和 94%的特异性^[8]。患者诊断 LVO 到首次接触血管内外科医生的时间减少了 44.13%^[9]。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。

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。

4.5.2 场景 2：药物研发与临床试验

AI 技术实现

计算机视觉、自然语言处理 (Natural Language Processing, NLP) 和大语言模型 (Large Language Models, LLMs) (如 ChatGPT 和 Gemini) 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^[10]。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^[11]。

行业应用

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^[12]。

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^[13], 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。

行业影响

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。

4.5.3 场景 3：健康管理与慢病防控

AI 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, 从而达到健康管理及慢病防控的目的。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^[14]。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^[15]。

行业应用

AI 赋能慢病管理^[16]。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^[17]。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^[18]，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。

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）。

4.5.4 场景 4：精准医疗与基因分析

AI 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。

行业应用

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^[19]，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。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^[20]。

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，识别疾病突变，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。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，提高治疗疗效。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^[21]。

行业影响

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“从基因组档案到终生健康管理”的美好愿景。

4.6 AI 赋能未来空间产业应用场景

未来空间产业聚焦空天、深海、深地等领域，通过高端装备，推动深地资源探采、城市地下空间开发利用、极地探测与作业等领域智能无人平台或设备等研制及创新应用^[1]。旨在通过技术融合与创新拓展人类活动及资源利用边界，具有高成本、高技术、高风险特征。其核心在于推动空间领域与装备、能源、材料等交叉融合，聚焦科学探测、资源开发、运输装备等关键环节的绿色化、集成化与智能化发展^[2]。

目前人工智能的发展主要聚焦在语言及多模态生成方向，基于大量的文字、语音、图像、视频等数据进行训练和发现规律，都是不超过二维的数据。但人类对世界的认识、世界背后规律的认知都是基于跟空间的交互，而空间本身是超越二维的。目前国内外有少数的前沿机构正在探索空间智能技术，通过大量像素数据的学习实现基于物理规律的空间自动生成。李飞飞教授研究团队通过多模态大语言模型（MLLMs），构建空间智能世界，整合语言和视觉信息方面的潜力备受关注。

与传统的人工智能模型相比，空间智能大模型具有以下显著特点：首先，它能够实现多源数据融合，整合来自地理信息系统、遥感技术、传感器网络等多个来源的空间数据，实现全方位、多维度的空间信息获取和分析。其次，它具有跨领域交叉应用的能力，不仅仅局限于计算机领域，还能与其他领域的数据和知识进行交叉融合，如数学、遥感、气象学、地质学等，实现跨领域的综合分析和智能决策。再者，它具备高效处理海量数据的能力，能够应对大规模、高维度的空间数据，借助分布式计算和高性能计算平台，实现对海量数据的快速处理和分析。最后，它拥有智能推理和预测的功能，通过学习空间数据的规律和模式，实现智能化的推理和预测，为用户提供精准的空间信息服务和决策支持。

因此，人类对空间的探索与利用，正从物理扩张转向智能重构。人工智能不再仅仅是工具，而是驱动未来空间产业跃迁的“核心动能”——它通过解构时空数据、重塑决策逻辑、突破物理边界，正在重新定义空间价值的创造范式。从低空物流的毫米级精准调度到虚拟现实的沉浸式孪生交互，从卫星组网的自主协同到城市空间的动态代谢，AI 以“算法+算力+数据”为支点，撬动了一场从效率优化到范式革命的产业跃迁。这场变革的本质，是让空间从“被动承载”转向“主动进化”，以智能决策替代经验试错，以动态推演取代静态规划，最终在国土治理、星际探索、虚实融合等维度构建人类文明的“第二增长曲线”。

以下场景将揭示：当 AI 与空间产业深度耦合，其释放的不仅是万亿级市场机遇，更是

一场重构人类生存维度的认知革命。

4.6.1 场景 1：城市空间规划的智能决策优化

AI 技术实现

基于生成式 AI 大模型，结合多模态数据（遥感影像、人口分布、交通流量等），构建动态规划知识图谱，实现空间冲突自动识别、方案动态优化与多目标平衡。通过强化学习算法，系统可模拟不同规划方案的长周期影响，辅助生成兼顾生态保护与经济效益的规划建议。通过融合多模态数据，包括遥感影像、人口分布、交通流量等各类丰富的数据信息，构建起一个动态规划知识图谱。在这个知识图谱中，不同的数据信息相互关联、相互影响，形成了一个有机的整体。基于此，系统能够实现空间冲突的自动识别。通过对遥感影像的分析，可以清晰地了解土地空间的利用现状，再结合人口分布和交通流量数据，能够精准地判断出城市用地与生态保护区之间是否存在冲突。一旦发现冲突区域，系统会运用强化学习算法，模拟不同规划方案在长周期内的影响。这种模拟过程并非简单地推测，而是基于大量的历史数据以及复杂的算法模型，对各种可能的情况进行详细的推演。通过模拟，系统能够辅助生成兼顾生态保护与经济效益的规划建议，确保规划方案既符合生态可持续发展的要求，又能为地区带来良好的经济收益。

城市地下空间综合利用赋能智慧城市建设



图 4.2 城市地下空间综合利用赋能智慧城市建设

行业应用

中国已成为城市地下空间开发利用大国。中国政府高度重视地下空间开发利用，出台一系列法律法规、规章、规范性文件鼓励和支持其发展。中国地下空间开发利用面积以年均 10% 的速度增长，大中型城市增长率高达 15% 以上，市场规模已达数千亿人民币。中国城市地下空间呈现“三心六片三轴”的总体发展形态，以地铁为主导的地下轨道交通、以综合管廊为主导的地下市政设施和地下停车、地下商业等快速崛起，开发利用规模居世界前列。

AI 赋能地下交通：从效率优化到安全升级。设备健康管理方面，通过部署 AI 机器人（如

北京地铁的智能安检机器人)实现24小时设备巡检,覆盖65%以上的车下巡检任务,效率提升30%^[3],利用机器视觉识别轨道裂缝、接触网异常等隐患,故障识别准确率大幅提升;**流预测与调度方面**,基于深度学习算法等分析历史客流数据,由上海市城乡建设和交通发展研究院等共同承担的智能交通主动管控模式,快速路区交通指数模型准确率达95%以上,将示范区域交通通行速度整体提高5%以上,拥堵时长降低8%以上^[4];杭州萧山国际机场升级AI新技术,不用身份证和登机牌,直接刷脸登,与原来的人工核验相比,新模式平均让每个航班旅客的排队时间缩短了6分钟以上,登机效率提升40%^[5];**自动驾驶优化方面**,采用多传感器融合技术(激光雷达+毫米波雷达)实现隧道内列车精准定位,应急事件处置反应时间由以前的5分钟大幅缩短至3秒以内,相比传统人工监控方式更精确、更快速,显著提升应急事件处理效率和水平,为隧道交通安全运行保驾护航^[6]。此外,AI地质建模技术可预测隧道开挖中的岩爆、渗水风险,利用无人机和高精度传感器进行数据采集,能够快速获取地铁沿线的地质数据,相比传统的人工采集方式,无人机技术可以将数据采集效率提升至传统人力的5倍以上^[7]。部署分布式光纤传感器网络,结合AI算法(如LSTM时序分析)评估隧道沉降变形,上海外滩隧道实现毫米级形变监测精度^[8]。

AI 重构地下商业生态:从空间运营到消费体验。**商业空间智能化改造方面**,利用AI技术对人流进行模拟和分析,能够精准预测不同时间段、不同区域的人流分布。通过模拟不同商铺布局方案下的人流走向,评估各方案对客流量的影响;根据人流模拟结果,结合客群画像和业态匹配度分析,对地下商圈的商铺布局进行优化。如,杭州“湖滨88”购物中心根据数据分析调整为中心定位为“国潮+亚运”主题,新引进花西子、李宁等29家品牌,单日客流上涨270%,营收同比上涨160%^[9]。博联AI照明系统(如深圳某地下商场案例)根据人流量动态调节亮度和区域照明,节能率超70%,年节省电费超百万元,实现能耗精细管理^[10]。**消费体验升级方面**,上海南京路地下步行街部署AI导购机器人,通过NLP交互提供个性化推荐,增加消费者停留时长,提升商业转化率^[11];AR导航系统结合室内定位技术,实现“虚实叠加”的沉浸式购物体验,提高用户复购率。地下购物中心、步行街等商业形态不断涌现,为城市商业发展注入新活力。

AI 驱动地下公共设施革新:从基础运维到韧性提升。如重庆江北区地下停车场采用AI车位识别系统,通过车牌识别+激光雷达定位实现“无感停车”,实现智能停车管理^[12]。苏州工业园区地下管廊部署物联网传感器,AI算法实时分析管道压力、温度数据,各类事件响应与联动处置时间由原有的30-40分钟缩短到5分钟以内^[13]。此外,还可以进行应急响应与韧性建设。安徽省城市生命线安全工程AI项目成功预警三级以上险情2000余起,全国累计预警超3500起,地下管网事故发生率下降了60%,风险排查效率提高了70%^[14]。以及搭建智能救援网络,联动无人机和机器人执行火灾救援任务,助力逃生通道规划效率提升等。

行业影响

在**效率革新方面**,传统的国土空间规划评估工作往往需要耗费数月的时间,涉及大量的

人力投入以及复杂的数据处理过程。而借助 AI 技术，这一过程如今能够缩短至数天。传统方式下，规划团队一般需要花费数月时间，组织大量专业人员对各类数据进行收集、整理和分析，而采用 AI 智能评估系统后，完成了同样的工作可以缩短至数天，同时降低了人力成本。在精准治理方面，通过数据驱动的空间资源分配方式，能够更加精准地把握土地资源的利用需求，减少因规划偏差导致的土地浪费现象。如，在一些城市的新区规划中，以往由于对人口增长和产业预测不够准确，导致部分土地规划不合理，出现了大量闲置土地。而现在，通过 AI 技术对人口分布、产业发展趋势等数据的精准分析，能够更加科学地规划土地用途，避免土地资源的浪费。在生态协同方面，AI 能够动态平衡开发与保护之间的关系。在“双碳”目标的大背景下，这一点显得尤为重要。AI 系统可以根据不同地区的生态环境特点以及发展需求，制定出既有利于经济发展，又能有效保护生态环境的规划方案，助力实现可持续发展的目标。

4.6.2 场景 2：低空经济的全场景赋能发展应用

AI 技术实现

主要利用 AI 算法（如深度 Q 网络）来优化无人机物流路径。在实际运行过程中，无人机需要实时考虑多种因素，如实时气象数据、空域管制信息等。深度 Q 网络算法能够根据这些实时变化的信息，实现多机协同避障与动态航线调整。当遇到恶劣天气，如强风、暴雨等情况时，算法会根据气象数据迅速调整无人机的飞行高度、速度和航线，以确保飞行安全。同时，AI 驱动的数字孪生平台发挥着重要作用。它能够通过对低空飞行场景的模拟，预测潜在风险。如，通过对某一区域的历史飞行数据、地形地貌以及气象条件等信息的综合分析，预测在特定时间段内可能出现的空域拥堵情况。基于这些预测结果，平台可以优化调度策略，合理安排无人机的飞行任务，提高低空资源的利用效率。

行业应用

低空经济+物流：智慧物流和数字化配送的加速器。地面物流承载压力大，而低空智联网的出现有利于低空物流的发展。到 2035 年，我国仅无人机物流行业产值可超万亿元规模，因此低空经济+物流是巨大的发展机遇。如美团与上海金山区政府正式合作签约，共同建设无人机城市低空物流运营示范中心，积极探索无人机在上海的落地，借助上海华东无人机基地的优势资源，依托国家民航局无人驾驶航空试验区的创新举措，构建城市末端十五分钟配送圈。

低空经济+旅游：有利开拓低空消费新业态。随着生活水平的提高，人们开始探索更新奇、有趣的旅游休闲项目。网络社交平台上不少人晒出了“热气球”、“滑翔伞”、“直升机”等短视频，刺激了消费者的欲望，有利于国内“平替”产品的涌现。如在神农架大九湖景区，游客可通过小程序预约直升机观光，10 分钟抵达核心景点，激活乡村旅游新业态。

低空经济+巡检：有利于整合巡检服务。无人巡检有诸多好处，如无人夜间值守、高效低空数据采集、突发事件的及时响应等，大大提高了巡逻和巡检的效率。而将其统一囊括在

低空经济体系中，有利于巡检产业的整合，提升巡检效率。如广州白云区创新“无人机+AI”低空巡检模式，智能巡航监管违法建设，搭载红外摄像头抓拍建筑垃圾偷倒行为，AI 识别地面环卫垃圾及占道经营情况等，赋能城市治理。

低空经济+消防：打通消防安全“空中通道”。部分地区一直为消防安全隐患所困。老旧小区、老城区、偏远山区等地方消防设施落后，低空消防能进行第一时间绕开复杂拥堵路段进行救援，这是地面消防远远无法达到的。如 2024 年云南楚雄山火扑救现场，3 架消防无人机在浓烟中划出精准航线：首架搭载红外热成像仪锁定火线蔓延方向，第二架投掷灭火弹压制火头，第三架为被困消防员空投呼吸面罩。这场持续 72 小时的立体救援，将传统处置时间压缩 60%，成为低空经济赋能应急救援的典型案列。

低空经济+农业：助力智慧农业，有利于可持续发展。低空经济+农业的重要载体是植保无人机。植保无人机在农业发展较早，在农林业应用广泛。因此农业是低空经济快速融合的产业。两者的结合有利于提升精细化生产和管理，有利于农业和农村经济的可持续发展。如依托低空物流，湖北十堰形成了“生产基地+中央工厂+电商营销”的农业产业化新模式，打造现代农业产业集群。

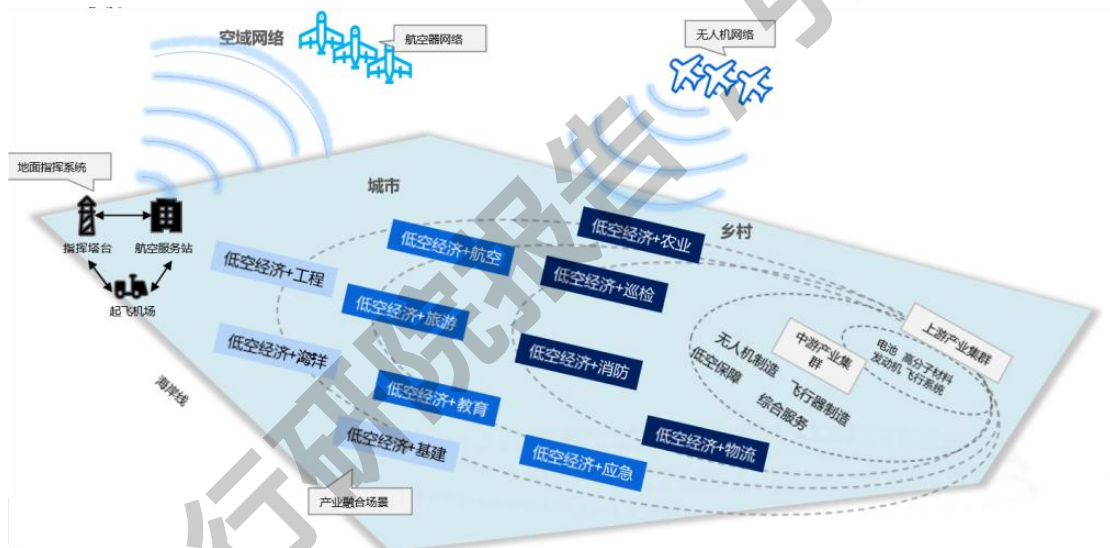


图 4.3 AI+低空经济的全场景赋能

行业影响

在商业化加速方面，智能调度与路径规划技术的应用，有力地推动了物流无人机、空中出租车等新业态的规模化落地。以往，由于低空飞行的安全性和效率问题，这些新业态的发展受到一定限制。而现在，随着 AI 技术的成熟应用，低空飞行的安全性和效率得到了极大提升，使得这些新业态能够快速发展并走向规模化。在安全性提升方面，空域冲突预警准确率提升。通过 AI 系统对空域信息的实时监测和分析，能够提前准确地预测空域冲突的发生，并及时发出预警。这使得相关部门和操作人员能够采取有效的措施进行应对，从而使事故率下降。在资源整合方面，通过 AI 调度实现了低空资源的共享。不同的无人机运营企业、空

中交通服务提供商等可以通过统一的 AI 调度平台，合理安排飞行任务，避免资源的重复配置和浪费，充分释放城市立体空间的潜力，为城市的发展创造更多的价值。

4.6.3 场景 3：卫星互联网的智能化运维与服务

AI 技术实现

依托基于 AI 的卫星数据智能分析平台（如星图云）。该平台利用计算机视觉技术对遥感影像进行解析。计算机视觉技术通过对影像中的各种特征进行识别和分析，能够快速准确地提取出有用的信息，如土地利用类型、植被覆盖情况、建筑物分布等。结合时序预测模型，平台能够实现卫星轨道优化、故障预警与能源管理。如，通过对卫星历史轨道数据以及当前运行状态的分析，时序预测模型可以预测卫星在未来一段时间内的轨道变化情况，从而提前对轨道进行优化调整，确保卫星始终处于最佳运行状态。知识图谱技术在其中也发挥着关键作用，它支持卫星数据与行业需求的精准匹配。通过构建知识图谱，将卫星数据与不同行业的需求进行关联和整合，使得卫星数据能够更好地为各行业服务。

行业应用

AI+卫星互联网+航空航海：保障全球航行安全与效率，重塑交通通信体系。跨洋航班、远洋船舶依赖传统海事卫星，带宽低、成本高且存在盲区。智能航空应用通过 AI 算法优化飞行路径规划、自主导航及实时环境感知，结合 5G 地空通信技术（5G-ATG），实现飞机与地面基站的直接通信，带宽和速率比卫星通信有所提升。卫星互联网支持空天地一体化网络，如洲际航天的“星上 AI 大模型”实现卫星自主决策，实时处理遥感数据并生成三维影像，服务于低空经济与数字孪生场景，开启航天工业“智能化、高效化、可持续化”为核心的工业 5.0+AI 时代^[15]。

AI+卫星互联网+偏远地区通信：填补全球网络覆盖空白，消除数字鸿沟。全无基站场景覆盖下低轨卫星星座（如 G60 星链）为山区、海洋等无网络区域提供高速的通信带宽，支持电力巡检、农业物联网等数据传输^[16]，便携式卫星终端，实现应急救援中的高清视频回传和远程医疗协作。

AI+卫星互联网+应急救援：AI 无人机协同网络。如系留无人机结合卫星链路，当震中地面基站损毁时，可用于医疗运输、农业植保、消防灭火的超大载荷空地协同系留无人机，为灾区提供通信照明的应急无人机。中型复合翼应急通信侦测无人机主要应用于航空应急救援场景，飞行时间可以超过 4 个小时，能在断电、断网情况下，为灾区提供应急通信。除此之外，它可以实时观察灾区情况，为救援提供依据^[17]。以及为高层建筑灭火救援研发的无人机等，结合城市智慧应急管理体系实现及时精准响应。

行业影响

在服务扩展上，推动了农业监测、环境评估等垂直领域的数据服务商业化。在农业监测中，通过卫星数据的分析，可以实时了解农作物的生长状况、病虫害发生情况等信息，为农业生产提供精准的指导。这些数据服务可以作为商品出售给农业企业和农户，促进了农业监

测数据服务的商业化发展。在成本降低方面，卫星运维效率提升。通过 AI 技术对卫星运行状态的实时监测和故障预警，能够提前发现并解决潜在问题，减少卫星故障的发生，从而降低了卫星运维成本。同时，通过优化卫星轨道和能源管理，延长了卫星在轨寿命。在全球覆盖方面，AI 驱动的低轨卫星组网技术，加速了全球通信与导航服务普惠化。通过构建低轨卫星网络，能够为全球范围内的用户提供更加便捷、高效的通信和导航服务，尤其是在一些偏远地区和发展中国家，使得更多的人能够享受到卫星互联网带来的便利。

4.6.4 场景 4：基于深度学习视觉技术的机坪作业保障

AI 技术实现

实时、精确的计算机视觉算法支持。运用深度学习、计算机视觉技术分析航班数据，实现对机坪作业保障运行的实时监控与优化。其中主要利用目标检测、跟踪和关键点检测技术，监测和追踪机坪内的航空器、特种车辆、机坪设施和机坪工作人员等关键资源，确保机坪作业工作的安全、顺利进行；再通过视频理解技术，针对性地处理航班节点中截断、遮挡物体等容易误识别的事件情况。算法通过分析机坪内的视频数据，实时发现异常情况，为机坪作业保障工作提供决策支持。

高效的资源调配工程优化。自主研发的机坪作业保障资源调度算法，根据航班实际情况，并行处理各个机坪作业节点分析流程，动态调整资源分配，如客梯车、排污车、加水车、加油车、客梯车等保障资源，提高资源利用率，实现资源最大化利用，提升航班靠桥率和机场运行效率。

预警与应急响应的方案。利用视频分析技术为各机场提供可定制化的机坪作业保障解决方案，提供各个节点事件回溯视频；建立预警机制，对可能影响机坪作业的因素进行预测和预警，提前制定应急响应措施，降低风险，对机坪作业实施提供实时监控与预警，提升异常情况响应效率。

行业应用

中国民航局于 2022 年先后发布了《机场协同决策系统技术规范》和《运输机场机坪运行管理规则》，其中《机场协同决策系统技术规范》明确提出了要采用人工智能机器学习技术对地面保障节点的视频数据进行分析从而实现航班保障节点时间数据的自动提取、录入和推送，航班保障节点共计 45 项，要求通过视频分析覆盖节点 20 项（图 1）。与此同时，《运输机场机坪运行管理规则》明确提出了机坪作业管理的各项规范，允许通过人工智能等先进手段对机坪作业进行视频分析、检测和报警，并通过数据收集方便管理部门对机坪作业进行实时监控和管理。

一方面，民航运行具有很强的系统性，一架飞机从起飞到落地、完成过站再次起飞，其间涉及撤轮挡、靠桥、机务放行、登机、关客舱门、起飞、落地、上轮挡等几十个航班保障节点、十几家驻场单位。目前，传统人工记录依靠地面保障人员通过手持终端设备进行记录，人力成本较高，工作量大，易出现漏填、误填，对于机场站坪监控员来说，高峰状态下保障

效率难以满足需求，且仅凭站坪监控人员靠人眼容易出现随意性大、数据不准确、无参考性等问题。任何一个保障节点的滞后或错误都会影响航班的目标撤轮档时间（TOBT），进而产生机场、航司的额外运营成本以及旅客的时间成本。

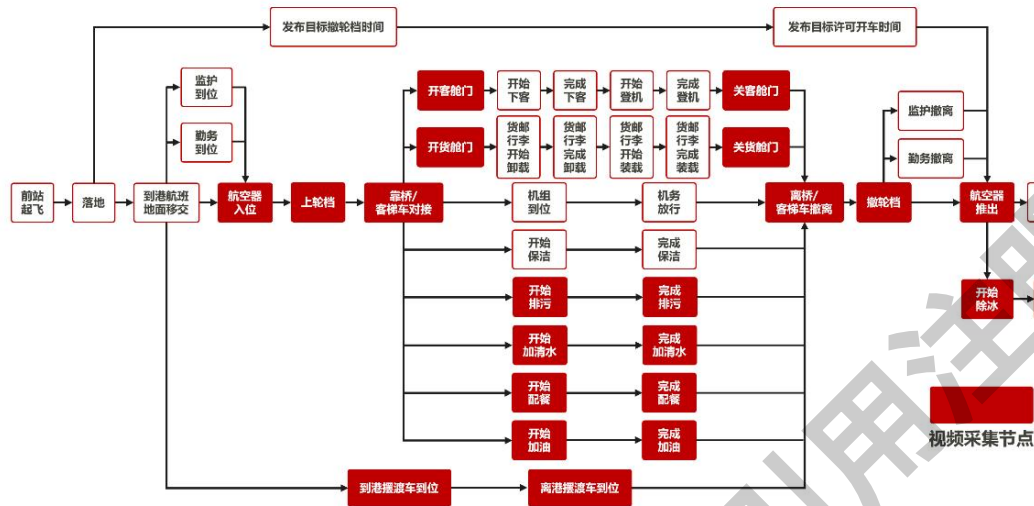


图 4.4 中国通用运输机场机坪航班保障节点

另一方面，机坪保障作业具有严格的安全规范要求，对航空器、特种车辆、机坪设施和机坪工作人员的运行要有监控措施，机坪的任何安全事故会造成严重的经济损失和社会影响。但目前的行业现状是完成保障任务时间短、任务重，一线作业人员能力或者遵章守纪的意识参差不齐、监控人员不足，无法实现全面及时的监管控制，而传统的监控仅仅起到事后追溯的效果，无法起到全面监控、事中预警的目的。

基于深度学习视觉技术的机坪作业保障，采用成熟的目标检测、目标跟踪、关键点检测等技术，可以满足智能航班保障节点识别和机坪作业规范识别的需求，能有效解决以上问题。

行业影响

提高运营效率。通过 AI 视觉技术实现航班保障节点的自动化采集，提高数据的可靠性和及时性，提升资源精细化管理，为机场在业务流程优化提供有效支撑，助力机场实现资源、运行效率的优化，有效地减少工作人员录入工作。同时，AI 视觉的建设成本较低，且可进行资源复用。机坪作业监管常态化。将现有抽查式监督的部分内容转由系统自动分析，实现机坪作业 24 小时常态化监管；部分作业规范无法通过视觉分析监管，暂时依然由人在系统中完成抽查监督；采用人机协同工作方式，提升机坪作业监管能力。提高安全管理水平。以民航局《运输机场机坪运行管理规则》为依据，结合机场机坪运行现状，基于机坪视频监控系统，利用计算机视觉和人工智能技术，通过对机坪作业流程风险的自动识别，主动提出预警信息，规范机坪运行管理秩序，提高机坪运行管理能力，提升机坪运行安全和效率。

4.7 AI 赋能未来金融服务业创新实践

金融产业将紧紧围绕数据驱动、服务普惠与风险可控这三大关键方向，借助人工智能技术与金融服务的深度融合，全面推动金融产品设计、风险管理、客户服务等核心环节的智能化重构。在数据驱动方面，金融机构将更加注重数据的收集、整合与分析，通过 AI 技术挖掘数据中的潜在价值，为金融决策提供精准依据。如，利用机器学习算法对海量的客户交易数据、行为数据进行分析，预测客户的需求和风险偏好，从而设计出更符合市场需求的金融产品。

在服务普惠方面，AI 技术将打破传统金融服务的时空限制，降低金融服务的门槛和成本，使更多的人能够享受到便捷、高效的金融服务。以智能投顾为例，通过 AI 算法为客户提供个性化的投资建议，无需人工理财顾问的一对一服务，大大提高了服务效率和覆盖面。对于风险可控，AI 赋能的风控系统能够实时监测和分析金融交易数据，精准识别潜在的风险，及时采取措施进行防范和控制。如利用深度学习技术对信用卡交易进行实时风险评估，毫秒级地拦截欺诈交易，有效降低金融机构的风险损失。

AI 技术在金融领域的应用不仅显著优化了传统金融业务的效率，还催生了一系列创新业态，如区块链金融、数字支付等。这些新业态的出现，不仅丰富了金融产品的种类，也为金融机构带来了新的增长机遇。区块链金融利用区块链的去中心化、不可篡改等特性，提高了金融交易的透明度和安全性，降低了交易成本。数字支付借助 AI 技术实现了支付过程的智能化，如人脸识别支付、指纹支付等，大大提高了支付的便捷性和安全性。

AI 助力金融机构实现从“经验驱动”向“算法驱动”的范式跃迁。传统金融业务中，决策往往依赖于人的经验和直觉，而 AI 技术的应用使得决策更加科学、精准。以数据为纽带，打破信息孤岛，构建动态决策体系，同时平衡创新与监管的关系，形成开放、安全、高效的金融生态。通过建立统一的数据平台，整合金融机构内部和外部的数据资源，实现数据的共享和流通，为 AI 算法提供丰富的数据支持。动态决策体系能够根据市场变化和客户需求实时调整决策策略，提高金融机构的应变能力和竞争力。在创新与监管之间找到平衡点，既鼓励金融机构利用 AI 技术进行创新，又确保金融市场的稳定和安全，形成一个健康、可持续发展的金融生态系统。

金融业的本质是信用与风险的管理，而 AI 正成为重塑这一领域的“新质生产力”。通过解构海量非结构化数据、重构决策逻辑、突破传统经验边界，AI 将金融业从“被动响应”推向“主动预判”，从“人力密集型”转向“算法密集型”。从智能风控的毫秒级欺诈拦截到财富管理的千人千面配置，从高频交易的算法优化到跨境支付的实时清算，AI 以“数据+算法+算力”为内核，驱动金融业完成从效率工具到价值创造的质变。

这场变革的本质，是让金融从“服务中介”升级为“智能中枢”，以动态风险定价替代静态规则，以全生命周期管理取代碎片化服务，最终在普惠金融、绿色金融、数字金融等维度构建全球金融体系的“韧性网络”。高频交易中，AI 算法能够快速分析市场数据，优化交易策略，

提高交易的成功率和收益。跨境支付借助 AI 技术实现了实时清算，大大缩短了交易时间，提高了资金的使用效率。

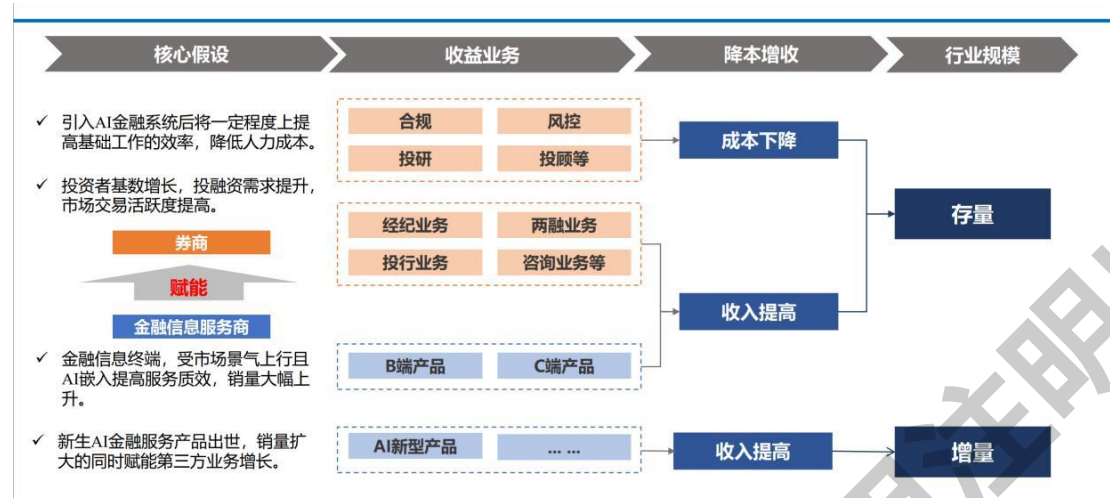


图 4.5 AI 赋能金融行业逻辑链条梳理

(资料来源：东吴证券研究所)

4.7.1 场景 1：智能风控的全局化重构

AI 技术实现

在金融行业的智能风控领域，AI 技术的应用正引领着一场深刻的变革。它依托多模态大模型，将交易流水、社交网络数据、物联网设备数据等多源异构数据进行深度融合，构建出一个动态的风险知识图谱。这一图谱能够实时更新和反映客户及交易的风险状况，为风险评估提供了全方位的视角。

通过图神经网络（GNN）技术，系统能够精准地识别出异常的资金流转路径，挖掘出隐藏在复杂交易网络中的风险点。同时，结合时序分析模型，如长短期记忆网络（LSTM）等，对客户的信用违约概率进行动态预测，从而实现对风险事件的提前预警，预警准确率被提升至 99.7%。如，国内某银行在实践中利用图神经网络技术，成功将反欺诈识别的误报率降低了 80%，极大地提高了反欺诈的效率和准确性。国内某券商通过构建 AI 数据治理平台，使得风险模型的预测偏差减少 40%，显著提升了风险评估的精准度。

此外，强化学习算法在智能风控中也发挥了重要作用。它被用于模拟不同风控策略在长期执行中的效果，即长尾效应，从而辅助金融机构生成既能够全面覆盖风险，又能促进业务增长的最优风控方案。这种算法能够不断试错和学习，最终找到在复杂业务环境中平衡风险与收益的最佳策略。

行业应用

在信贷风控方面，某银行将大模型技术深度嵌入信贷审批流程，实现了自动化的需求分析、案例生成以及测试脚本生成。这一举措使得审批效率得到了显著提升，相比传统审批方式，效率提高了 40%。

在反洗钱监测领域，国际清算银行（BIS）的研究指出，AI 技术能够深入分析高维稀疏数据，挖掘出传统手段难以发现的复杂洗钱网络。某跨国银行在跨境结算业务中应用 AI 技术，风险识别效率提升了 70%，成功拦截了多起潜在的洗钱交易，保障了金融系统的安全和稳定。

在市场风险预警方面，我国某企业投研平台接入大模型后，能够实时解析新闻舆情与市场波动之间的关联性，为投资者提供及时准确的市场动态信息，金融数据的调取效率提升 15%，帮助投资者更好地把握市场机会，做出投资决策提供了依据。

行业影响

AI 技术在金融风控领域的应用，使得风险可控性得到了质的跃升。传统的人工风控方式往往需要耗费大量时间和人力，如贷前审查通常需要 3 天才能完成，而 AI 系统仅需 15 分钟，人力成本节约 60%。

在长尾客户覆盖方面，AI 技术使金融机构能够拓展服务范围，覆盖以往难以评估信用的中小微企业等长尾客户群体。某银行通过构建非财务数据模型，将小微企业贷款覆盖率从 35% 显著提升至 78%，为解决小微企业融资难问题提供了新的思路和方法，促进了金融资源的公平分配和有效利用。

此外，监管协同也得到了全面升级。比如：中金智阅大模型通过构建多重风控防线，其文档审核系统成功入选国家级 AI 先锋案例，并与深交所、上交所的监管系统实现了直连。这不仅提高了金融机构的合规效率，还增强了监管机构的监管效能，促进了金融市场的健康稳定发展。



图 4.6 AI 赋能金融行业风控管理

（资料来源：根据公开资料整理）

4.7.2 场景 2：财富管理的个性化革命

AI 技术实现

在财富管理领域，依托生成式 AI，金融机构能够构建起客户 360°画像系统。这一系统整合了客户的风险偏好、生命周期阶段、社交行为等 200 多个维度的数据，为精准的客户细分和个性化的服务提供了坚实的基础。

通过强化学习算法，AI 可以模拟在不同市场环境下资产配置策略的长期收益，从而为客户提供最优的投资组合建议。同时，借助自然语言处理（NLP）技术，系统能够生成个性化的投资报告，用通俗易懂的语言向客户解释投资策略和市场动态。

行业应用

在智能投顾方面，国内某股份制银行的“AI 中台” 每日支撑超过 10 亿次决策调用，使得客户资产配置匹配度提升了 35%，留存率增加了 22%。另一家公募基金则通过 AI 工具实现客户画像的动态优化，服务响应速度达到了“秒级”，极大地提高了客户服务的效率和质量。

在产品定制领域，AI 根据个体健康数据等多维度信息，为客户提供精准的保险产品设计。如，国内某风险管理公司为偏远地区患者定制健康险，通过大数据分析和 AI 算法优化，使保费成本降低了 30%，让更多人能够负担得起保险服务。

在跨境服务方面，金融壹账通智能体平台通过 RAG 知识检索模块，整合了 30 个国家的监管政策数据，使跨境财富规划方案的生成效率提升了 2.5 倍，为高净值客户提供了高效、合规的跨境财富管理服务。

行业影响

AI 技术推动了财富管理服务的民主化。AI 智能投顾将原本仅服务于高净值客户的专属服务扩展至长尾市场，如 Betterment、Wealthfront 等平台，管理资产规模突破万亿美元，服务门槛降至 100 美元，让更多普通投资者能够享受到专业的财富管理服务。

在决策科学化方面，中金研究《AI 经济学》指出，AI 驱动的资产配置策略年化收益较传统模型高 3 - 5 个百分点，同时波动率降低 20%，显著提高了投资决策的科学性和收益性。

生态开放化也成为行业趋势，头部银行与科技公司共建超过 200 个 API 接口，形成了“财富管理 + 场景生态”模式。如浦发银行的数字员工覆盖智能培训、数字审核等十多个岗位，提升了运营效率和服务质量，为客户提供更加全面、便捷的金融服务体验。

4.7.3 场景 3：金融基础设施的智能化跃迁

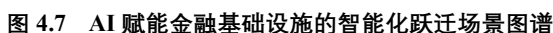
AI 技术实现

在金融基础设施领域，基于联邦学习与区块链技术，金融机构能够构建起跨机构的数据协作网络，有效解决长期存在的数据孤岛问题。通过联邦学习，各机构在保护数据隐私的前提下，共同训练机器学习模型，提升模型的泛化能力和预测准确性。而区块链技术则为数据共享提供了可信的底层架构，确保数据的真实性和不可篡改。

AI 数字孪生技术在金融系统的应用，使得模拟金融系统运行状态成为可能。通过创建虚拟的金融系统模型，AI 算法能够实时监测和分析系统的运行情况，预测潜在的风险和问题，并通过多智能体协同优化支付清算、跨境结算等核心流程。如某银行凭借其千亿参数金融大模型，实现了支付欺诈检测准确率高达 99.9%，每秒能够处理百万级的交易量，极大地提升了交易处理的效率和安全性。

行业应用

在绿色金融方面，AI 碳核算平台整合了卫星遥感、物联网等多源数据，为金融机构提供了精准的碳排放监测和核算服务。某银行借助这一平台，将绿色信贷风险评估效率提升了 50%，并将不良率控制在 0.3% 以下，既支持了绿色产业的发展，又有效控制了金融风险。



行业影响

AI 技术的应用使得金融交易的效率得到了指数级的提升。高频交易的延迟被降至微秒级，某券商的算法交易占比从 15% 跃升至 60%，年增收益超过 20 亿元，显著提高了市场流动性和交易活跃度。

AI 与金融的融合已超越“技术叠加”阶段，进入“基因重组”的新周期。当金融业从“资本中介”进化为“数据中枢”，其价值创造逻辑正被重新定义——风险防控从“事后处置”转为“事前预判”，客户服务从“标准化供给”转向“需求创造”，基础设施从“物理节点”升级为“智能网络”。据 IDC 预测，到 2027 年全球金融业 AI 支出将达 970 亿美元，而中国智慧金融市场规模有望占据全球 30% 份额。这场变革不仅关乎技术升级，更是人类对金融本质的认知革命：当算法成为新的“信用货币”，数据流动取代资本流动，AI 驱动的金融业将真正成为全球经济高质量发展的“血液循环系统”。

在 AI 与金融融合的新周期中，技术不再只是辅助工具，而是深入到金融业务的每一个细胞，重塑着金融行业的基因。风险防控方面，AI 技术实现了从“事后处置”到“事前预判”的转变。通过大数据分析和机器学习算法，金融机构能够提前识别潜在风险，及时采取预防措施，将风险扼杀在萌芽状态。客户服务方面，AI 的应用使服务从“标准化供给”转向“需求创造”。智能客服、个性化推荐等技术能够精准把握客户需求，为其提供量身定制的金融产品和服务，不仅提升了客户满意度，还激发了新的金融需求。金融基础设施方面，AI 技术推动其从“物理节点”升级为“智能网络”。区块链、智能支付系统等创新技术的应用，使得金融交易更加高效、安全和便捷，为金融服务的普惠化和全球化奠定了基础。

4.8 本章结语

一、AI 与实体经济的深度融合正在开启一场以“智能原生”为特征的产业范式革命

本章对六大战略领域的系统性研究揭示出：AI 赋能的本质在于通过“认知重构”与“流程再造”打破传统产业边界，其技术穿透力体现在知识沉淀、模式识别与决策优化三个维度，而行业纵深度则表现为从设备层到系统层的全栈式重构。智能制造领域展现出“感知-决策-执行”闭环的柔性进化，信息科技领域催生出算力-算法-数据的三体协同，先进材料领域实现高通量实验与逆向设计融合，清洁能源领域构建起时空动态优化网络，健康医疗领域完成从群体诊疗到分子级干预的跃迁，空间经济领域则突破人类物理极限的探索边界。

AI 在“技术穿透力+产业纵深度”的双重突破，重构复杂产业价值链条，在关键领域形成高质量发展新动能，总结其特征如下：

技术驱动的产业基因重塑

AI 技术正从单一工具演变为产业发展的“新型基础设施”。如，制造领域的高附加值订单优先调度、结合历史数据预测未来物料流动趋势、提前规划资源分配；金融领域通过 AI 实现风险评估与反欺诈的实时决策；均标志着 AI 已深度嵌入生产流程的核心环节。

数据驱动的全链条智能化

制造质检、医疗影像诊断等场景中，AI 通过海量数据训练形成精准决策能力，推动产业从经验驱动转向数据驱动。如医疗领域的 AI 辅助诊断系统，通过医学数据分析提升诊断效率，而能源领域则借助 AI 优化电网调度与损耗预测。

流程与组织的范式革新

AI 不仅优化效率，更催生新的业务形态。如，材料研发中 AI 加速新材料的发现周期，空间领域通过卫星数据分析实现环境监测与灾害预警的智能化，均体现了 AI 对传统产业逻辑的重构。

垂直行业的深度定制化

AI 技术与产业知识深度融合，形成“专业化模型+本地化场景”的生态。垂直产业模型快速发展，制造、医疗和能源等领域企业加速构建专属模型。算力基础以及数据合规与质量成为企业 AI 时代竞争力的关键。

人机协作模式快速发展

RAG（检索增强生成）与 AI Agent 技术结合，使 AI 从工具升级为自主任务执行者，推动服务自主化，产业人机协作场景将迎来普惠化快速发展，倒逼产业对劳动者技能更新和企业组织调整。

云边端协同的轻量化服务

边缘计算与终端设备算力迎来进一步提升，推动 AI 服务向“小场景、细颗粒”延伸，智能穿戴设备、智能 PC、智能手机将迎来交互方式的变革，大小模型协调、云边端协调成为产业发展主流方向。

二、未来方向与挑战

AI 技术发展正呈现三大融合趋势：一是多模态大模型与具身智能的结合，推动制造、医疗等场景向具身化决策演进；二是量子计算与生成式 AI 的融合，将加速新材料研发和能源系统优化；三是神经拟态计算与脑机接口的突破，预示着人机协同新范式的到来。值得关注的是，AI 产业化向产业 AI 化的重心转移正在形成“技术下沉-场景反哺”的双向进化机制，而可信 AI 与价值对齐技术的突破将成为规模化应用的前提条件。

同时，AI 赋能过程中也面临三个主要挑战：一是技术伦理与治理框架的完善，产业及企业数据隐私、算法透明度等问题需通过政策与技术创新同步解决；二是通用人工智能（AGI）的长期布局，随着多模态大模型、具身智能等技术发展，AI 将逐步实现从小场景到企业通用场景的跨越，甚至在整个产业中复杂场景、协同场景拓展。三是全球化竞争下的生态竞合，未来国家与城市竞争的关键在于“技术能力+产业禀赋+生态协同”的综合优势。

因此，AI 赋能行业的浪潮既是技术革命，更是认知革命。它要求我们以更开放的姿态拥抱变化，既深耕垂直领域的场景落地，又前瞻布局通用智能的突破方向。这场智能革命的核心价值在于重构产业创新要素——当算法成为新型生产资料、算力构成基础生产力、智能模型衍生出新生产关系时，传统产业的技术代差将被重新定义。未来产业竞争的本质，将演

化为“数据密度×算法精度×场景厚度”的乘积效应，而谁能率先构建起 AI 原生型产业生态，谁就能在智能经济时代掌握价值创造的制高点。这既是技术经济范式的迭代，更是人类认知疆域的拓展，其深远影响将超越产业变革本身，重塑人类文明与机器智能协同进化的新图景。

本章参考文献

4.2 AI 赋能信息产业应用场景

[1]Software 2.0, 即软件 2.0, 由前 OpenAI 创始成员和研究科学家 Andre Karpathy 于 2017 年提出, 他预见到引入 AI 之后的新软件开发范式, 他在一篇技术博客中提出了软件 2.0 的概念。

[2]2025 年 3 月 19 日, GTC March 2025 Keynote with Nvidia CEO Jensen Huang, 分享观点: Scaling law 进入了新阶段, From one to three scaling laws: pre-training scaling – post training scaling – Test-time scaling “long thinking”.

[3]2025 年 3 月 19 日, GTC March 2025 Keynote with Nvidia CEO Jensen Huang, 分享观点: “AI 模型的演化路: Perception AI – Generative AI- Agentic AI – Physical AI”

[4]目前比较流行的 RAG/Workflow 方案包括: anythingllm, Maxkb 工具, 围绕其的自动化工作流方案、业务流程多种多样

4.3 AI 赋能材料产业应用场景

[1]科学家利用 AI 创造出一种全新材料: 坚固如钢 轻如泡沫 革新航空航天材料科学_新闻频道_中华网

[2]如何借助 AI 加速高端金属材料研发? 创材深造孔辉博士精彩演讲分享|航空航天|高温合金_网易订阅

[3]和电量焦虑说拜拜! AI 智能技术从 3200 万种材料中发现未来「超级电池」

[4]AI 为药物研发按下“快进键”

[5]晶泰科技与微软中国达成战略合作 共探 AI 于生物医药、材料科学应用|微软中国|生物医药|晶泰_新浪新闻

[6]赛托生物与英矽智能战略合作 拟以 AI 技术发力合成生物学及创新药研发_腾讯新闻

4.4 AI 赋能能源产业应用场景

[1]中陕核秦洲核安: 拥抱 DeepSeek, 夯实“AI+核技术”创新之路

[2]中核二三: 用 DeepSeek 重塑核电工程管理 引领行业新赛道

[3]DeepMind 用 AI 实现可控核聚变突破: 等离子体形状模拟精度提高 65%

[4]降本增效我们在行动 | AI 云平台赋能, 让氢耗再低一点

[5]中国电建集团成都勘测设计研究院有限公司. (2024). 中国电建投资成立科技公司 含多项 AI 业务

4.5 AI 赋能健康医疗产业应用场景

[1]中国大健康产业发展现状及趋势分析, 产业升级成为趋势--华经情报网

- [2]推动普惠医疗服务发展，助力健全多层次医疗保障体系--光明网
- [3]人工智能驱动个性化医疗，解锁精准健康服务新篇章—光明网
- [4]健康中国：政策体系与发展进路，王虎峰，《人民论坛》（2024年03月01日第2版）
- [5]Precision public health in the era of genomics and big data. Megan C. Roberts et al. Nature Medicine, volume 30, pages1865–1873 (2024).
- [6]AI 在医疗：深度学习在医学影像诊断中的最新进展
- [7]谷歌与 Aravind 眼科医院合作，将检测眼疾的 AI 应用到临床中
- [8]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
- [9]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
- [10]人工智能赋能药物开发：现状、挑战与未来趋势
- [11]Kang Zhang, et al. Artificial intelligence in drug development. Nature Medicine, volume 31, pages45–59 (2025).
- [12]人工智能赋能药物开发：现状、挑战与未来趋势
- [13]Qiao Jin, et al. Matching patients to clinical trials with large language models. Nature Communications, volume 15, Article number: 9074 (2024) .
- [14]慢病精准预测：大模型+多模态融合
- [15]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. 中华全科医学, 2024, 22(3): 481-485.
- [16]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！--健康界
- [17]卢心怡，(2022年7月5日), Babylon Health: AI 与健康管理结合（第一篇），长正研究院
- [18]华为官网
- [19]科德施官网
- [20]肺癌一线治疗新选择，奥希替尼联合疗法再获批！
- [21]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

4.6 AI 赋能空间产业应用场景

- [1]工业和信息化部等七部门关于推动未来产业创新发展的实施意见_国务院部门文件_中国政府网
- [2]韩健,李雨凌,孙美玉.解码未来空间产业的新图景[J].经济,2024,(08):43-46.
- [3]【北京：交通巡检机器人上岗】-国家发展和改革委员会
- [4]预判拥堵，主动管控：AI+交通如何成为“火眼金睛”？
- [5]AI+交通如何赋能智慧出行？_中央网络安全和信息化委员会办公室
- [6]AI 让塘朗山隧道有了“数字生命”！深圳打造广东首个数字孪生智慧隧道平台_深圳新闻网
- [7]给地球做“CT”、AI 算法……找矿用上这些“黑科技”→_《人民日报》

- [8]光纤传感器及信号解调设备选型手册.7d008ef538e588dbc108235fcd55b2fd.pdf
- [9]大湖滨商圈构建线上线下一体化智慧消费生态_每日商报
- [10]智光杯双奖见证，博联智能以创新方案开辟智能照明新征途！|智能家居|智能化|智能控制_网易科技
- [11]百年商业街“搬进”元宇宙 --科技日报数字报
- [12]重庆江北区九街 2045P 智能车库投用 - 南京江宁停车产业发展有限公司
- [13]“数字孪生”赋能城北路地下综合管廊智慧运维 - 赫里奥(苏州)科技有限公司
- [14]2024 年“数据要素×”大赛优秀项目案例集——应急管理案例之四 | 多元数据赋能城市生命线安全工程，助力智慧韧性城市建设
- [15]AI 赋能卫星制造，开启航天工业 5.0+新时代_即时新闻_大众网
- [16]无需基站，实时互联！我国卫星互联网应用落地提速
- [17]空中“黑科技”飞入寻常百姓家-新华网

4.7 AI 赋能金融产业的创新实践

- [1]刘玉龙, 翟卓. 更高效、更便捷、更安全——从全国两会看数字金融如何赋能千行百业[N]. 新华网, 2025-03-07
- [2] 国务院办公厅. 关于做好金融“五篇大文章”的指导意见（国办发〔2025〕8号）[Z]. 2025-03-06.
- [3]佚名. 解码 AI+金融：从智能客服到风险重构的生态革命[J]. 金融科技前沿, 2025, 3(2): 15-28.
- [4]王孜, 杨春宇. AI 驱动金融革命：应用全景与治理挑战[J]. 金融研究, 2025, 41(3): 45-59.
- [5]李峰. 大模型时代银行业的智能化跃迁路径[J]. 上海金融, 2025, 36(1): 22-30.
- [6]中金公司. AI 智道（6）：AI+金融有望迎来规模化拐点[R]. 北京：中金研究, 2025.
- [7]金融壹账通. 智能面审产品解决方案白皮书[R]. 深圳：金融壹账通, 2025.
- [8]郑小林. 大模型驱动金融风险防控的底层逻辑变革[J]. 金融科技前沿, 2025, 3(2): 32-45.
- [9]国务院办公厅. 人工智能赋能金融风险防控指导意见（国办发〔2025〕12号）[Z]. 2025-
- [10]陈亮, 等. 中金点睛大模型：AI 投研平台的范式创新[J]. 证券与投资, 2025, 28(4): 56-
- [11]王孜. 健康险智能化转型：从风险预防到精准触达[M]. 上海：复旦大学出版社, 2025.
- [12]招商银行研究院. 财富管理中台：AI 驱动的客户运营体系构建[R]. 深圳：招商银行, 2025.
- [13]金融壹账通. 跨境财富管理智能体技术报告[R]. 香港：金融壹账通亚太中心, 2025.
- [14]中国人民银行数字货币研究所. DCEP 技术白皮书（2025 版）[R]. 北京：中国人民银行,
- [15]GS 银行金融科技部. 千亿参数金融大模型在支付清算中的应用[J]. 银行科技, 2025, 16(3): 18-29.
- [16]新致软件. 智能可穿戴设备在金融合规中的实践[R]. 上海：新致软件, 2025.

[17]中国电子信息产业发展研究院. 卫星互联网与普惠金融发展报告[R]. 北京: 工信部, 2025.

[18]国务院. 关于推进“人工智能+”行动的指导意见[N]. 人民日报, 2025-03-06(01).

[19]中国人工智能学会. AI+金融融合路径与技术伦理白皮书[R]. 杭州: 浙江大学, 2025.

上海交通大学研究院报告, 引用注明出处

第 5 章 AI 赋能行业应用案例

5.1 AI 赋能制造行业应用案例

5.1.1 案例 1：自主化制造系统，某 C 钢铁企业的工业机器人布局

背景及介绍

在智能制造的推动下，钢铁企业也积极响应市场需求，利用自身的垂直应用场景，加速工业机器人的布局和应用。面对生产效率提升、产品质量优化以及人工成本降低等多方面的挑战，该公司引入了大量的工业机器人和智能装备，构建自主化制造系统，旨在实现生产过程的自动化和智能化。

AI 技术具体应用

该公司的自主化制造系统集成了先进的 AI 算法和大数据分析技术，实现对生产过程的精准控制和优化。

工业机器人自动化：工业机器人根据预设的程序和算法，自动完成各种复杂任务，如搬运、焊接、装配等，降低人工成本和生产误差。

智能传感与机器视觉：结合智能传感和机器视觉技术，实现对产品质量的实时监测和优化控制，确保产品质量的稳定性和一致性。

数据驱动优化：通过实时采集和分析生产数据，AI 算法对生产调度和资源利用进行优化，提高生产效率和经济效益。

项目成果与效益分析

该公司的自主化制造系统上线运行后，为生产效率和产品质量带来了显著提升。生产效率提升方面，通过实现生产过程的自动化和智能化，该公司的生产效率显著提升，生产周期缩短，产能大幅提高。产品质量优化方面，智能传感和机器视觉技术的应用，使企业的产品质量显著提升，废品率大幅降低，增强了产品的市场竞争力。成本控制方面，工业机器人的引入降低了人工成本，同时优化生产调度和资源利用，减少了能源浪费和非计划停机时间，有效降低了生产成本。行业影响力提升方面，该公司在自主化制造系统方面的创新实践，为钢铁行业的智能制造提供了宝贵经验，提升了企业在行业内的知名度和影响力。

5.1.2 案例 2：工业元宇宙协同，某 D 集团的数字孪生

背景与详细介绍

该集团是与新中国几乎同龄的制造企业，面对行业竞争和可持续发展的压力，在“十四五”期间实施的“数字孪生+”工业元宇宙协同项目代表了国内该行业数字化转型的最高水平。这些项目旨在提升效率、降低成本并提高安全性。

AI 技术具体应用

本案例中产品研发与工艺优化中结合 AI 技术的应用正逐步重塑制造业的创新模式。通过 VR/AR 技术与数字孪生技术的深度结合，企业可在虚拟环境中完成产品开发和工艺优化，显著减少实物试验的次数，降低研发成本。AI 驱动的参数优化算法能够实时分析多维度数

据，智能调整工艺参数，以提高产品性能和制造精度。同时，远程协同研发平台依托 AI 自然语言处理 (NLP) 和计算机视觉技术，使跨地域研发团队能够在同一虚拟空间内高效协作，从而加快项目推进速度，实现敏捷创新。

在工艺优化领域，通过 AI 技术构建了全流程的工艺数字孪生系统，实现生产过程的精准模拟与优化。深度学习算法能够基于历史数据和实时生产参数，预测可能的工艺瓶颈并提供优化方案，确保制造过程的稳定性与效率。智能工艺优化系统还能结合强化学习模型，持续调整制造参数，在不同生产条件下实现最优工艺配置，提升产品一致性和良品率。

设备运维方面，AI 结合 AR 技术构建了智能运维体系，为关键设备创建实时数字孪生模型。通过 AI 驱动的数据分析，系统能够预测设备故障并提供预防性维护方案，减少意外停机时间。AR 辅助运维则使现场工程师能够通过智能眼镜实时接收 AI 解析的设备状态、维修指引和远程专家支持，从而提升维修效率，降低维护成本。AI 在智能制造中的深度应用，不仅提升了生产效率和产品质量，也推动了企业向智能化、数字化、可持续化方向发展。

成果与效益分析

在全球制造业迈向智能化、精益化的背景下，依托数字化技术与精益管理模式，全面提升生产运营效率。核心产线的生产效率实现了突破性增长，提升幅度超过二成以上，推动企业产能利用率显著提高。这一变革不仅优化了资源配置，还增强了企业的市场竞争力，使其在波动的市场环境中保持稳健增长。同时，通过智能制造数字孪生技术的深化应用，我们在高端产品制造环节取得了显著成果，一次性合格率的提升有效降低了返工成本，增强了客户信任度，进一步巩固了品牌影响力。

数字孪生技术的普及正加速推动产业升级，新产品的开发周期已缩短近一半，这一突破使企业能够更快响应市场需求，在竞争激烈的行业环境中抢占先机。与此同时，虚实结合的培训方式和远程虚拟协同模式的推广，使得企业安全事故发生率大幅下降，并成功实现重大安全事故“零发生”的目标，充分保障了生产运营的稳定性。此外，随着企业数字化转型的推进，我们培养了一批具备跨学科能力的复合型人才，这不仅提升了企业自身的创新力和运营能力，也为行业未来的智能制造变革奠定了坚实的人才基础。

5.1.3 案例3：汽车制造，E 公司超级工厂智能化生产

背景及介绍

在汽车制造业智能化转型的浪潮中，E 公司超级工厂 (Gigafactory) 作为全球制造业的标杆，通过 AI 技术重新定义了生产效率与供应链管理。面对传统汽车制造中存在的工艺分散、物流低效和环保压力等挑战，该公司自 2013 年起构建了集冲压、焊接、涂装和总装四大工艺于一体的超级工厂，以“空间垂直整合”和“零库存”理念颠覆传统生产模式。2024 年上海超级工厂的自动化率突破 95%，实现 30 秒下线一辆整车的制造速度，成为全球首个实现全产业链闭环的智能工厂。

AI 技术具体应用

工业物联网平台：部署百万级传感器网络，实时采集生产线振动、温度等 200+设备参数，通过大模型实现预测性维护，准确率 93%，故障预警时间从 24 小时延伸至 72 小时。构建数字孪生系统，模拟 4680 电池产线全生命周期运行状态，设备综合效率从 78%提升至 92%。

智能物流系统：采用“车轮上的仓库”模式，在工厂周边设置近百个集装箱装卸口，通过 AI 算法优化物料流动路径，使零部件从供应商到生产线仅需 10 分钟，实现零库存管理。自主研发的物料流动系统日均处理 2000+集装箱，物流效率较传统模式提升 300%。

超高速自动化产线：冲压车间采用 6000 吨级巨型压铸机，将 Model 3 后车架的 70 个零件集成化，生产时间从数小时压缩至 3 分钟，减少 80%焊接机器人使用量。焊装车间引入 3D 线激光扫描技术，零件尺寸检测精度达 $0.5\mu\text{m}$ ，缺陷漏检率降至 0.008ppm (百万分之一)。

闭环电池回收系统：在内华达州超级工厂部署电池回收装置，通过 AI 驱动的化学分离技术实现锂、钴等材料 100%回收，使新电池生产成本降低 15%。

成果与效益分析

生产效率方面，上海超级工厂通过多层厂房设计与立体化空间整合（如空中输送链、自动化物流系统），实现单位面积产量达到传统车厂的 3-5 倍，创全球最高产线密度。其核心在于将冲压、焊接、涂装、总装四大工艺集中布局，缩短物流路径 90%，并通过“车轮上的仓库”模式实现零库存管理，日均处理近 2000 个集装箱。

质量管控方面，产品不良率从 500ppm (百万分之五百) 降至 0.008ppm，达到六西格玛认证水平。通过 AI 视觉检测系统（3D 线激光扫描精度 $0.5\mu\text{m}$ ）与自动化质检线，缺陷漏检率趋近于零。

成本优化方面，通过预测性维护系统与设备健康管理，年减少非计划停机损失 6000 万美元，设备维护成本下降 45%。如 AI 算法提前 72 小时预警设备故障，平均故障间隔从 120 小时提升至 400 小时。

5.2 AI 赋能信息科技行业应用案例

5.2.1 案例 1：Synopsys 的 AI 驱动 EDA 工具

背景与介绍

新思科技 (Synopsys) 成立于 1986 年，总部位于美国硅谷。在芯片到软件的众多领域，新思科技始终引领技术趋势，与全球科技公司紧密合作，共同开发人们所依赖的电子产品和软件应用。新思科技是全球排名第一的芯片自动化设计解决方案提供商，全球排名第一的芯片接口 IP 供应商，同时也是信息安全与软件质量的全球领导者。作为半导体、人工智能、汽车电子及软件安全等产业的核心技术驱动者，新思科技的技术一直深刻影响着当前全球五大新兴科技创新应用：智能汽车、物联网、人工智能、云计算和信息安全。建立了完善的技术研发和人才培养体系，秉持“以新一代 EDA 缔造数字社会”的理念，支撑全球半导体产业的创新和发展，并共同打造产业互联的数据平台，赋能数字社会建设。2025 年 3 月 18 日，

新思科技公司 (Synopsys) 宣布与英伟达 (NVIDIA) 合作的下一阶段, 利用 NVIDIA Grace Blackwell 平台将芯片设计速度提高 30 倍。

AI 技术细节

Synopsys 作为全球领先的 EDA 工具提供商, 推出了一系列 AI 驱动的 EDA 工具, 如 DSO.ai (Design Space Optimization AI)。这些工具利用深度学习、强化学习等 AI 技术, 自动化完成从系统架构定义、设计实施、验证到制造的全流程任务。AI 算法能够智能探索设计空间, 优化参数组合, 找到最佳设计方案。

数字化设计空间优化: 利用 AI 算法自动探索设计空间, 优化功耗、性能和面积 (PPA) 等关键指标。通过深度学习模型预测不同设计参数组合的性能表现, 快速找到满足设计目标的最优解。

模拟设计自动化: 利用 AI 技术自动分析模拟设计的特点和需求, 生成适应新工艺节点的设计方案。

电路仿真: 利用 NVIDIA Grace Blackwell 平台, 新思科技 PrimeSim SPICE 仿真工作负载预计将实现 30 倍加速。与 NVIDIA 合作的加速计算架构使具有挑战性的电路仿真能够以 SPICE 级精度实现签核。

验证覆盖率收敛和回归分析: 通过 AI 算法智能生成测试向量, 覆盖更多设计场景, 同时利用回归分析技术预测潜在的设计错误, 提前进行优化和改进。

自动测试生成: 利用 AI 技术学习历史测试数据, 生成针对特定设计的高效测试模式, 减少测试时间和成本。

制造解决方案: 利用 AI 算法优化光刻模型参数, 提高模型预测精度, 从而在实际制造过程中获得更高的芯片良率。

成果与效益分析

芯片设计是人类历史上最复杂的工程挑战之一, 借助 NVIDIA Blackwell 和 CUDA-X, 新思科技将仿真时间从几天缩短到几小时——推动芯片设计的发展, 赋能 AI 革命。传统芯片设计需要工程师花费数周甚至数月的时间进行调整和优化, 而 AI 工具可以在短时间内完成这些任务。如, DSO.ai 在两天内完成了原本需要专家一个月的探索任务, 大幅缩短了设计周期。质量优化方面, AI 生成的设计方案往往在性能、功耗和散热等方面达到更优的平衡, 通过 AI 优化, 某大型核心 CPU 在效能和功耗上均达到了理想的设计目标。自动化设计减少了工程师的数量需求, 降低了人力成本, 同时, AI 工具提高了设计成功率, 减少了流片失败的风险。

5.2.2 案例 2: 英伟达 (NVIDIA) ——Omniverse 工业元宇宙平台

背景及介绍

英伟达作为全球领先的图形处理技术和人工智能计算公司, 推出了 Omniverse 平台。NVIDIA Omniverse 是一个包含 API、SDK 和服务的平台, 使开发者能够将 OpenUSD、

NVIDIA RTX 渲染技术和生成式物理 AI 集成到工业和机器人用例的现有软件工具和仿真工作流程中。旨在为工业界提供一个开放且易于扩展的虚拟协作环境，被应用于汽车制造、建筑设计、机器人仿真等领域，重新定义复杂系统的协作开发模式。通过 Omniverse，不同领域的专业人士能够在一个共享的虚拟空间中实时合作，无论他们身处何地。Omniverse 不仅仅是一个 3D 设计协作平台，它更是一个集成了先进 AI 技术的工业元宇宙平台，支持从概念设计到最终产品的全生命周期管理。这个平台利用了最新的光线追踪技术和物理准确的模拟引擎，使得设计师、工程师和艺术家能够在高度逼真的环境中进行创作和测试。

AI 技术具体应用

使用 Generative AI 引导合成数据生成，构建工业数字孪生并应用生成式 AI 来增强感知 AI 模型训练的数据创建。

用于机器人学习的合成运动生成，仅从少数几个人类演示中生成指数级的综合运动数据，用于模拟学习。

构建用于交互式流体仿真的数字孪生，可将实时可视化和分析集成到 3D 工作流程中，以实现产品的实时计算机辅助工程设计。

在工业数字孪生环境中，对用于工业自动化的多机器人队列进行大规模的仿真和测试，在实际部署这些多机器人队列之前，确保其在物理 AI 的支持下能够正常运行，提高实际部署的成功率和效率，减少可能出现的问题和风险。

使用数字孪生和 AI 生成精确的营销视觉效果，将生成式 AI 集成到您的 3D 工作流程中，以实现物理上精确的 3D 产品配置器。

构建用于远程监控的数字孪生，通过与实时 IoT 数据集成的准确数字孪生实现对工业运营的远程监控。

成果与效益分析

自推出以来，Omniverse 已经帮助多个行业实现了显著的进步。如，在汽车制造业，设计师们使用 Omniverse 来加速新车的设计流程，从最初的概念到最后的市场推广，大大缩短了产品开发周期。同时，制造商能够利用该平台进行生产线的模拟和优化，减少了实际生产中的错误和浪费。对于建筑和工程领域来说，Omniverse 提供了强大的工具来进行建筑设计评审和结构分析，提高了项目的可视化程度和决策质量。此外，通过实施预测性维护策略，企业不仅降低了运营成本，还提升了设备的可靠性和安全性。总体而言，Omniverse 作为一个工业元宇宙平台，不仅推动了技术创新，也为各行业带来了实质性的经济效益。它促进了跨学科团队之间的紧密合作，加快了创新速度，为企业创造了更多的商业机会和社会价值。

5.3 AI 赋能材料行业应用案例

5.3.1 案例 1：S 科技公司 AI for Science 变革新材料研发范式

背景及介绍

S 科技公司是全球 AI for Science 领导者，依托在交叉学科领域的深耕，构建了大模型体系，并进一步解决科学研究和工业研发领域的关键问题，将众多学科的科研方法从“实验试错/计算机”时代带入了“预训练模型时代”，形成了 AI for Science 的“创新-落地”链路和开放生态，赋能“千行百业”，为人类经济发展最基础的生物医药、能源、材料和信息科学与工程研究打造新一代工业设计与仿真系统。

AI 技术具体应用

其研发的新一代材料设计平台，深度结合人工智能技术与物理模拟和先进表征技术，并提供高质量多源数据库、高通量和高性能计算引擎和高便捷性的操作界面。材料研发人员可以在玄铸材料智能设计平台的帮助下深入理解材料构效关系，进而加速材料及工艺，持续提升新材料性能，推动能源电池、催化剂、聚合物、合金、陶瓷、电子产品等领域发展。

融合云原生高性能计算引擎与先进机器学习模型，推动新型合金材料成分设计，高通量预测弹性模量、缺陷形成能等多种合金属性，快速评估合金材料的稳定性。

使用分子动力学和聚合物结构表示的机器学习模型，实现聚合物材料性能的快速预测，精确预测聚合物结构及其关键物理化学性质，设计新颖聚合物、筛选配方并优化合成路径

利用深度势能分子动力学模拟，加速新型催化剂的发现并解析复杂的催化反应机制，深入理解真实反应条件下的催化过程，高通量筛选具备特定反应性能的创新催化剂。

通过大原子模型与电解液配方设计模型的结合，实现电池材料和电解液配方的高效筛选和优化。精准预测固态电解质的离子电导率及其稳定性利用高通量、高精度的快速分子物性预测，快速进行电解液配方的筛选和设计。

成果与效益分析

基于 DeePMD 方法和 DPA 原子间势函数预训练模型，评估预测正极材料和固态电解质全掺杂空间的关键性质，准确度与实验可比，效率比主流方法快至少 1000 倍。产品级正极材料和固态电解质掺杂评估，交互式完成万级别的材料筛选，加速新材料研发周期缩短至原来 1/3。

基于 Uni-Mol 分子三维结果预训练模型，准确预测电解液全分子空间的多种关键物性，快速推荐候选目标分子。基于 Uni-ELF 配方模型，几分钟内准确预测上千个电解液配方关键性质结果，快速设计和优化电解液配方。产品级电解液分子和配方性质预测和高通量筛选，一键式获得电解液配方性质报告，加速筛选流程，将百万级配方筛选缩短至天。

使用机器学习技术，充分挖掘历史数据规律，建立工艺构效关系模型，加速工艺调控效率。利用深度学习算法准确识别 SEM 图像的形态、尺寸、组成和分布，提升识别分析效率。产品级工艺调控和表征分析工具。快速完成工艺调控的综合分析，自动化批量处理电镜图像，与实验形成有效迭代，将开发成本减少至一半。

5.3.2 案例 2：HL 以人工智能驱动新材料研发

背景及介绍

HL 上海公司成立于 2021 年，是一家在材料科学与人工智能技术融合领域领先的高科技企业。公司以数据驱动的材料信息设计和工程技术为核心，自主研发了材料信息人工智能和数据管理平台，致力于通过机器学习、高通量计算以及高通量实验等前沿技术，为新材料研发提供高效能的解决方案。HL 的业务领域广泛，涉及新能源、电池、化工制剂、半导体等多个工业应用。公司不仅提高了新材料研发的效率，降低了研发风险，还推动了材料相关制造业在研发、工艺、流程等核心环节的创新，从而提升企业的竞争力和价值链地位。HL 以其在材料科学与人工智能技术交汇点的突破性探索，正引领着科学研究与工业制造的未来走向，以前所未有的速度和精确度，深度挖掘材料科学的潜力。

AI 技术具体应用

AI 加速材料设计流程。文献分析与知识整合方面，通过大模型与自然语言处理技术，从海量科学文献中提取关键信息，构建材料研发知识图谱。通过主动学习与贝叶斯优化，在实验初期结合专家经验与 AI 模型进行快速迭代，缩短目标配方收敛时间。进行性能模拟与逆向优化，利用机器学习建立数学模型，模拟电解液等材料的性能，并逆向指导配方优化，减少对实验数据的依赖。

实验设计与优化。通过机器学习算法，在极少采样下快速找到多目标优化的最优解，相比传统网格或正交采样方法减少实验次数。自主研发的高通量计算平台与自动化实验室联动，实现配方筛选与工艺优化的高效验证。

跨体系迁移学习。通过 AI 进行迁移学习技术，将已验证的配方优化经验迁移至其他材料体系（如金属、陶瓷等），降低跨领域研发成本。

成果与效益分析

核心领域成果方面，电池电解液研发助力企业实现电解液性能提升，显著增强市场竞争力；黑磷材料制备通过 AI 优化工艺，实现高纯度黑磷晶体规模化生产，推动其在电子信息、新能源领域的应用；其他新材料产品如开发磷基电池材料、气凝胶、有机硅封装胶等，并完成中试验证，部分产品已投产。**经济效益方面，**AI 加速材料设计流程，提升研发效率与经济性；在实施过程中，AI 加速材料设计流程展现出了相较于传统研发流程的显著优势，助力研发成本降低，实验次数减少，单项目研发周期缩短。市场得到了拓展，服务覆盖新能源、金属、高分子等多个领域，助力客户产品性能与成本优势的提升。**技术合作与投资认可，**与大型科技企业合作，新增 AI 相关业务范围，如人工智能理论开发、智能机器人销售；完成多轮融资，包括红杉中国种子基金领投的数千万元种子轮，加速技术商业化。

5.3.3 案例 3：SY 智能实验平台引领科研新潮流

背景及介绍

北京 SY 科技有限公司成立于 2020 年，核心研发人员来自国内顶尖知名院校，汇聚了杜邦等知名化工材料企业的高层次管理人才。公司基于 AI+自动化研发模式，通过高通量干湿实验筛选提供新材料、新能源、化工等领域的新分子合成服务，高效探索新催化剂和分子

合成。SY 独创领先算法，推出 DeepChem 智能合成平台，基于人工智能、计算化学、云平台 and 智能合成机器人，设计目标分子及其合成路径，探索催化剂、溶剂和温度等反应条件对结果的影响，提供深度的分子合成研发，并通过自动化合成仪器提供高纯度样品，以推动 AI+底层绿色创新的应用。光伏封装产品，作为深云平台设计的具有独特分子结构的新型有机硅材料，打破国际垄断，实现国产替代，提升光伏产品封装品质。公司的有机硅管线以硅研新材为产品品牌，“绿色高效”为核心理念，开发出超薄防火隔热涂料，实现金属及制品保护；透明防火涂层，用于车载/建筑玻璃等透明防火需求领域。

AI 技术具体应用

智能实验室与自动化系统。SY 自主研发的 DeepChem 智能合成平台，集成了 AI、计算化学、机器人等技术，实现了微克级自动化实验流程。其核心功能包括：全流程自动化，自动加料、反应、稀释、过滤、分析及数据处理，无需人工干预；高通量实验能力，单套系统日均完成百个实验，效率接近百人博士团队的工作量，将研发周期大幅缩短；通过 AI 分析实验数据，实时调整反应参数（如催化剂、溶剂、温度等），快速锁定最优合成路径。

分子设计与材料性能预测。其开发的平台可以在线完成初始结构的绘制，同时支持其他绘图软件的多种文件格式上传、输入。高性能集群服务器保证在平均 4 小时内完成分子结构的优化，快速测量、获取键长键角等结构信息。同时，平台提供了数十种溶剂，并根据用户需要持续更新中，无需用户再手动建模增加摆放溶剂分子，平台会根据溶剂的选择自动判断添加位置与数目，可以便捷地为体系设计显示溶剂模型，得到隐式溶剂模型无法展示的其他性质。

高度定制化的模型，自动化工业化合成。使用预设参数，自行训练模型，个性化定制模型，自行选取特定的算法、模型架构以及超参数进行训练，闭环迭代，持续优化。高性能的预训练模型，辅助进行过渡态的搜索，快速且精准，进行分子性能的预测，快速筛选可用结构，自动生成新分子，结构可靠性能优。自研的 FUN-SMILES 表达式，选择性生成分子，结构更合理，生成结构更可靠，便于工业化合成，内置官能团库，统一规范，便于自动化工业化合成。

直接对接 DFT 计算云平台，多维数据可视化，自动生成模型训练报告及结果分析报告，支持预测值与量子化学模拟结果的自动对比。

成果与效益分析

在实验材料研发设计方面，通过精心设计目标分子及其合成路径，全面探索催化剂、溶剂和温度等反应条件对结果的影响，在这个过程中，能够发现具备更优性能的化学结构，从而有效解决现有材料性能缺陷，为新材料研发提供分子合成解决方案。**研发效率提升，产品更快推向市场。**其智能实验室，每一套设备占地面积仅几平方米，每天却可以进行上百个自动化试验，接近百名博士的研发团队工作量，极大提高研发效率，过去需要 10 到 20 天才能完成的科研工作，如今只需要 1 到 2 天。通过 AI 模型的分析，研发团队能够快速确定电解

液的最优配方，减少了不必要的实验环节，使得新能源电池能够更快地推向市场，满足市场需求，还为企业节省了大量的时间成本。**得到政府的政策支持**，北京将 SY 纳入“AI+新材料”行动计划，这体现了政府对 SY 技术的认可和支持，并将在资金、人才、政策等方面给予 SY 支持，政策的支持为 SY 的发展提供了良好的环境和机遇，帮助其加快技术研发和产业化进程。**促进生态协同**。SY 与高校、科研机构、企业等合作，构建了从基础研究到产业落地的完整链条。高校和科研机构负责基础理论的研究和材料的初步设计，SY 则利用其 AI 技术和智能实验室对材料进行进一步的优化和实验验证，最终将研发成果推向市场，这种生态协同的模式，促进了产学研的深度融合，推动了“AI+新材料”领域的发展。

5.4 AI 赋能能源行业应用案例

5.4.1 案例 1：G 公司本地部署大模型保障核安全

背景及介绍

早在 2024 年 7 月，该公司就引入我国国产大模型，构建了核科学智能问答系统、辐射剂量评估模型、放射性废物分类识别工具，并实现了铀矿资源库、核地质勘探数据的知识图谱化管理。这一项目通过 GraphRAG 算法外挂知识库，有效降低了大模型幻觉，使核学领域问答准确率提升了 40%，并且支持无网环境下的极端场景应用。项目获批省级重点研发专项，形成了“数据不出库、权限不外授”的安全 AI 底座。2025 年 2 月，该公司宣布完成 DeepSeek 系列模型的全面本地化部署，并正式启动集团级私有化大模型训练计划。

AI 技术应用与项目成效分析

该公司的这一创新实践，不仅在技术上实现了突破，更在实际应用中展现显著成效。通过大模型的本地化部署，该公司成功构建了一个安全、高效、自主进化的智能化基座。这一基座不仅满足了涉核企业对敏感数据的存储要求，实现了全生命周期数据闭环，还通过定制开发的集团 AI 平台，深度对接集团核心业务系统，支持个性化方案研究和模块开发。如，在核工业数据分析问答、辐射剂量评估、辐射评估报告自动生成、放射性废物分类识别等核业务场景中，垂直功能模块的开发使智能化功能更符合企业的特定应用场景和决策支持需求，全面提升集团业务支持效能和组织运营水平。此外，该公司在训练中的大模型全面覆盖了集团各技术领域的标准，注入了集团发展数十年来积累的地质资料库、核地质勘探工程数据、环境辐射本底数据、放射性装置运维数据等核心数据及下属企业多年运营案例，形成了难以复制的领域知识密度优势。通过本地化模型的持续训练，使大模型对行业具有更好地理解 and 把控，打造了该公司专属能力。

5.4.2 案例 2：美国普林斯顿大学 AI 预测等离子体不稳定性

背景及介绍

2024 年，普林斯顿大学等离子体物理实验室（PPPL）联合该校人工智能研究中心，针对国家球形托卡马克实验升级装置（NSTX-U）开发了基于强化学习的等离子体控制算法。团队采用近端策略优化（PPO）算法训练模型，通过分析装置内 1,000+传感器的实时数据（包

括磁场强度、电子温度、等离子体密度等），提前 300 毫秒预测等离子体边界的“逃逸电子雪崩”现象。当 AI 检测到风险时，系统会自动调整 19 个磁场线圈的电流参数，将等离子体约束时间延长至 4.2 秒，较传统控制方法提升 50%。

AI 技术应用与项目成效分析

该研究首次实现了托卡马克装置中 AI 的闭环控制，将等离子体不稳定性预测误差率从传统方法的 28% 降至 9%。论文《Real-time prediction of plasma instabilities using deep reinforcement learning》发表于《自然》杂志 2024 年第 626 卷，被国际热核聚变实验堆（ITER）项目组引用为“里程碑式突破”。实验数据显示，AI 系统可将单次放电成本降低 15%，为商业化核聚变堆的实时控制提供了技术范式。

5.4.3 案例 3：DeepMind 提升托卡马克等离子体模拟精度

2023 年，DeepMind 与瑞士洛桑联邦理工学院（EPFL）等离子体中心合作，针对 TCV 托卡马克装置开发了基于 Transformer 架构的 AI 模型。团队将 10 万小时的历史放电数据输入模型，训练其学习等离子体磁流体力学（MHD）方程的隐含规律。相较于传统有限元模拟方法，AI 模型将等离子体边界形状预测误差从 1.2cm 降至 0.42cm（精度提升 65%），同时将单步模拟时间从 72 小时缩短至 24 小时。实验验证表明，AI 优化后的磁场控制策略使 TCV 装置的等离子体电流稳定性提升 30%，成功实现了“非感应电流驱动”模式的持续运行。相关成果发表于《Nuclear Fusion》期刊 2023 年第 63 卷，被 ITER 团队用于优化未来堆芯的磁场设计。DeepMind 开源了该模型的核心算法（Plasmaformer），推动全球核聚变研究机构的协作。

5.4.4 案例 4：K 科技公司 AI 云平台破解氢能重卡“油耗焦虑”

2024 年，该公司针对氢能重卡普遍存在的“里程短、氢耗高”痛点，构建了行业首个全自研的燃料电池 AI 研发云平台。工程师们在某物流公司的 30 辆氢能重卡上部署传感器，采集了 200 万公里实际运行数据（涵盖上海至苏州的高速公路、城市配送等复杂场景），发现制动能量回收率仅 1% 的关键问题。平台运用强化学习算法，在毫秒级内动态调整燃料电池输出功率与电机扭矩分配。如，当车辆爬坡时，AI 识别到燃料电池处于低效区间，自动提升输出功率至 85% 并优化空气供应，同时将电机能量回收阈值从 20km/h 降至 15km/h。实测数据显示，某辆氢能重卡从上海嘉定到苏州工业园的单程氢耗从 5.2kg 降至 4.3kg，相当于燃油车油耗从 35L/100km 降至 29L/100km。

该公司年运营里程超 1000 万公里，AI 优化后年节省氢燃料成本约 650 万元（按当前氢价 65 元/kg 计算），单辆车日均运营里程从 450 公里提升至 520 公里，相当于每天多完成 2 趟城市配送任务。团队已申请 12 项发明专利，其中《基于实时路况的燃料电池动态控制方法》获中国专利优秀奖，平台通过 ISO 26262 功能安全认证。项目成果获上海市经信委青年产业创新大赛金奖，中通、顺丰等物流企业已签约试点。

5.4.5 案例 5：L 集团创新实践——AI 驱动生物质能全产业链智能化升级

L 集团于 2024 年成立数科公司，通过 AI 技术重构生物质能全产业链，系统性解决行业长期存在的原料波动大、转化效率低、运维成本高等核心痛点。该公司依托自主研发的生物质能智能管理平台，通过部署物联网传感器网络与深度机器学习算法，实现了从原料采集到电力生产的全流程数字化管控。在原料端，系统实时采集全国 2000 余个农业合作社的生物质资源数据，构建包含湿度、热值、糖分含量等 128 项参数的动态数据库，结合地域气候模型与物流成本优化算法，将原料采购半径缩短 40%，采购成本降低 18%。在生产环节，AI 模型通过分析锅炉燃烧过程中的 6000+ 实时参数，动态调整进料速度与氧量配比，使发电效率从行业平均 28% 提升至 35%，同时结合电网负荷预测算法，将弃电率降低 22%。设备运维方面，部署于关键节点的 5000 个智能传感器与故障预测模型协同工作，实现设备健康状态实时监测与预防性维护，使年度运维成本下降 230 万元，设备停机时间减少 65%。该技术已申请《基于神经网络的生物质能发电过程优化方法》等 12 项发明专利，并纳入国家可再生能源智能化示范工程。

这些实践案例充分展示了 AI 在未来能源领域的巨大潜力和实际成效。通过这些应用，我们不仅看到了 AI 对未来能源发展的积极影响，也为后续进一步探索 AI 在能源领域的深化应用奠定了坚实基础。接下来，我们将深入探讨 AI 在未来能源领域的挑战与应对策略，以期这一领域的持续发展提供更全面的视角和更具前瞻性的思考和展望。

5.5 AI 赋能健康医疗行业应用案例

5.5.1 案例 1：腾讯觅影数智医疗影像平台

背景与介绍

医学影像是现代医学诊断的重要工具之一，尤其在癌症和心脑血管疾病等的早期发现和诊断中发挥了重要的作用。既往，医生主要依靠知识和经验去寻找潜在病灶。再加上影像的复杂性和对细节的高要求，难免出现误诊或漏诊。随着近些年来 AI 技术的不断发展，基于深度学习算法，尤其是卷积神经网络（CNN），AI 可以通过学习大量的医学影像数据，识别和分类图像中的细微变化，生成结构化报告，帮助医生快速完成诊断，显著提升了影像数据处理效率和早期疾病筛查精度^[1]。

根据 ARK Invest 的预测^[2]，全球医疗保健人工智能市场预计到 2030 年将达到超过 1,553 亿美元，年复合增长率为 35.5%，是 AI 应用最大的领域。其中，医学影像是最为成熟、广泛的 AI 临床应用之一。国家医保局也已将 AI 辅助诊断纳入放射和超声检查的扩展项目^[3]。在算法迭代、数据累积、临床验证深入和政策支持下，医学影像已经成为各大企业争相入局的焦点，进一步将 AI 赋能的医学影像推向更大规模的应用。以三甲医院为例，AI 辅助诊断系统渗透率超 65%；在基层医疗机构，通过云端 API 调用，可以实现技术普惠（WHO 报告）^[4]。

腾讯觅影是腾讯公司首款 AI 医学影像产品^[5]，依靠腾讯 AI Lab 的技术能力，运用计算机视觉、深度学习和大数据技术对各类医学影像（内窥镜、病理、钼靶、超声、CT、MRI

等)进行学习训练,完成对医学影像的分类、目标检测、图像分割和检索工作,有效地辅助医生进行病变识别、辅助诊断、重大疾病早期筛查、疗效评估等任务,致力于实现对早期食管癌、宫颈癌、早期肺癌、乳腺癌及乳腺癌淋巴切片病理图像、糖尿病性视网膜病变等多个病种的筛查研发。

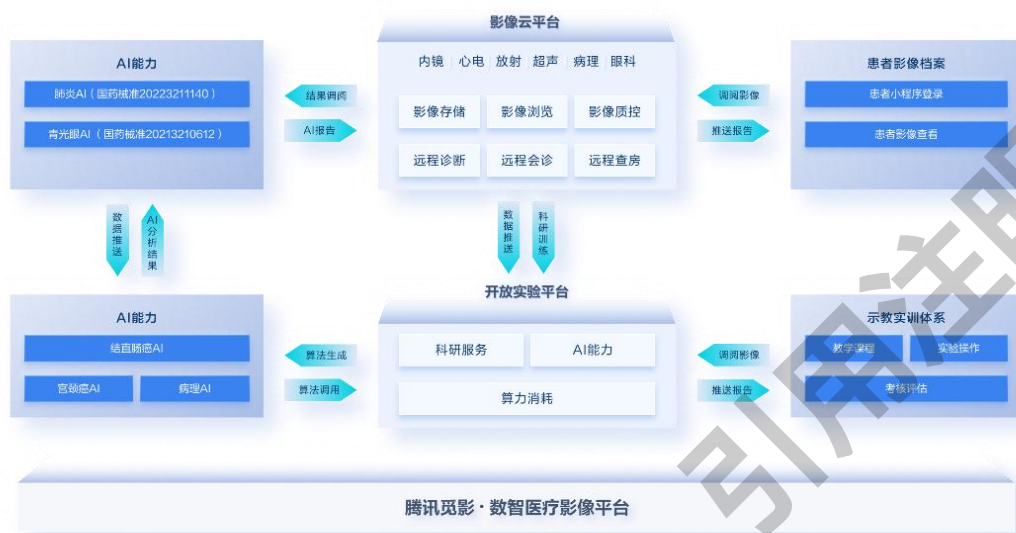


图 5.1 腾讯觅影数智医疗影像平台

AI 技术具体应用

(1) 重大疾病早筛。食管癌早筛：腾讯觅影通过运用计算机视觉与深度学习技术，对食管内镜影像进行训练学习，智能识别食管病灶，它筛查一个内镜检查用时不到 4 秒，对早期食管癌的检出率高达 90%，帮助医生更及时地发现食管癌早期病变，为患者争取了宝贵的治疗时间^[6]。肺癌早筛：通过对 CT 影像数据的预处理、3D 化以及进一步的擦除和增强，得到一个机器可以识别的增强 3D 图片。可以检测肺结节，3 毫米以上肺结节识别的准确率达到 95%；并判断其良恶性，对肺结节的良恶性判断灵敏度达到 85%，特异度达到 90%^[7]。在实际应用中，帮助医生更准确地评估患者的肺癌风险，减少了因人工读片可能出现的漏诊和误诊情况。眼底疾病引发的全身性疾病风险筛查：腾讯觅影的 AI 眼底筛查技术可解析血管细节、精准识别病变，助力发现多种眼底疾病，如糖尿病视网膜病变、高血压眼底病变等，进而提示患者可能存在的全身性重大疾病风险。这一创新举措为社区卫生服务中心带来了全新的筛查体验，如在深圳南山区的公益筛查中，为街道内的上班族、社区居民开展眼健康公益筛查，通过分析眼底图像，发现了不少居民存在疑似青光眼、高血压眼底病变等问题，为早期干预和治疗相关全身性疾病提供了线索^[8]。

(2) 重大公卫事件响应。采用自监督深度学习的方法，通过对肺部 CT 影像数据进行智能化分析处理，实现自动识别肺炎病灶、肺炎分割、肺炎量化评估、配准随访对比及自动生成分析报告等功能。以 COVID-19 肺炎辅助诊断为例，能在 CT 检查后最快秒级完成 AI

模式识别，数分钟即可为医生提供辅助诊断参考。武汉抗疫的两个多月内，先后为 24000 多人次患者进行了肺部 CT 诊断工作^[9]。

(3) 基层医疗赋能实践。在区域基层医疗机构，基于分级诊疗政策，面向区域医疗机构，协助基层医生在疾病诊疗、转诊、用药等过程中提升工作效率；针对区域卫生监管需求，在门诊数据分析、药物风险监控、诊断风险质控等工作中协助卫生管理机构，推动区域卫生水平建设。以腾讯觅影眼底 AI 筛查系统为例^[10]，已在全国 30 多家医院联合进行验证，并在广东、山东、四川、广西等多个省市基层医疗卫生机构落地试点，通过远程诊断技术服务为基层及社区提供更广的筛查支持，目前已筛查超过 6.5 万人。其他还有提供标准化诊断支持、推动社区各类早癌筛查、助力基层与上级医院协作等。

成果与效益分析

提升诊断准确度。对早期食管癌的检出率高达 90%，3 毫米以上肺结节识别的准确率达到 95%，避免人为误诊、漏诊等。**提高诊断效率。**在 AI 辅助下，筛查一个内镜检查用时不到 4 秒，十几秒可以诊断糖尿病视网膜病变，显著提高了诊断的效率，这将有助于缓解专科医生不足的瓶颈，开启疾病基层筛查新模式。**降低医疗成本。**得益于云端协同（利用腾讯云弹性计算资源，降低本地硬件投入）以及自动化流程（从影像上传到生成报告全程 AI 驱动，减少人工干预），早期癌症筛查的高成本得以控制。以食管癌为例，传统内镜检查费用逾千元，而“觅影 2.0”通过云端算力优化和规模化应用，将单次费用压缩至 200 元，降幅达 80%^[11]。

为基层医疗赋能。AI 推动医疗普惠革命^[12]。据国家远程医疗中心数据显示，截至 2024 年 6 月，全国 97.3% 的县级医院部署了 AI 辅助诊断系统，覆盖 CT、超声、病理等 18 个核心科室。腾讯觅影 AI 影像辅助系统可以协助云南勐腊县人民医院的肺部 CT 筛查，在三千多公里外的北京中日友好医院实时生成结构化报告，将基层医生的诊断效率提升 400%。2018 年，腾讯与浙江省德清县政府签署战略合作协议，打造全国的县域 AI 医学示范基地，通过“数字医疗创新联盟”推进医疗 AI 向基层下沉^[13]。其中，AI 筛查使早期肺癌检出率提升 3 倍。此外，腾讯觅影通过科研共享、硬件协同等方式，与各方合作推动基层医疗发展。如，与迈瑞合作推出全自动外周血细胞形态学分析仪，融入了 AI 算法，能清晰还原细胞立体结构和细节，阅片准确率达 95%，效率提升了 60 倍^[14]。

5.5.2 案例 2：N 公司全生命周期健康管理系统

背景与介绍

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。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%^[15]。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。全生命周期健康管理作为 AI 赋能健康产业的典型代表，通过技术创新和模式优化，已经初步取得一些成果。它是一种涵盖个体从出生到生命终点的全方位健康管理模式，旨在通过预防、治疗和康复等手段，全面提升人民的健康水平。

N 公司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^[16]。该系统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，Master Agent 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；Task Agent 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；Knowledge Agent 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。

AI 技术具体应用

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改善患者生活质量。



图 5.2 N 公司全生命周期健康管理系统

成果与效益分析

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，使医护人员在同等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。

5.6 AI 赋能空间行业应用案例

5.6.1 案例 1：CL——虚拟空间智能设计赋能城市更新

背景与介绍

CL 作为全球领先的空间智能企业，在城市规划和建筑改造领域发挥着重要作用。其利用自研生成式引擎，为相关工作提供 AI 驱动的设计解决方案。该企业的技术核心在于融合 GNN（图神经网络）与细粒度空间识别算法。图神经网络能够对空间结构和关系进行有效的建模和分析，而细粒度空间识别算法则可以精确地识别空间中的各种细节信息。通过两者的融合，实现了虚实融合的空间优化，为城市更新项目提供了更加科学、高效的设计思路。

AI 技术具体应用

在方案生成环节，用户只需输入城市老旧小区的相关参数，如小区的面积、建筑布局、周边环境等信息，AI 便能够自动生成兼顾功能性与美观度的改造方案。这些方案不仅考虑了小区居民的生活需求，还注重提升小区的整体美观性和环境质量。同时，AI 支持 3D 可视化推演，用户可以通过虚拟现实技术，身临其境地感受改造后的效果，提前对方案进行评估和调整。**在供应链协同环节**，AI 能够根据设计方案，精准匹配建材供应商与施工方案。它会综合考虑建材的质量、价格、供应周期以及施工方案的可行性、成本和工期等因素，动态优化成本与工期。通过这种方式，既保证了项目的质量，又降低了成本，提高了项目的整体效益。**在居民参与环节**，通过 VR 交互平台，居民可实时调整设计细节。居民可以根据自己的喜好和需求，对房屋的布局、装修风格等细节进行调整。这一举措极大地提升了居民的参与度，使城市更新项目更能满足居民的实际需求，提高居民的满意度。

成果与效益分析

在效率提升方面，传统的城市更新项目设计周期通常可能需要数月，涉及大量的人力投入和复杂的设计流程。而 CL 不仅能加快老旧小区改造的设计和实施进程，还能为城市居民提供更美观、舒适、个性化的居住和生活空间。在实际应用中显著提升了家装的设计与营销效率，大幅削减了人力、时间成本，实现降本增效的显著成果，为家装的数字化转型提供了有力支持。同时，促进了家居行业与城市更新的融合发展，推动了相关产业的升级^[1]。**在资源节约方面**，通过 AI 的精准设计和供应链协同，材料浪费降低。同时，由于项目的优化设计，碳排放减少了，为环境保护做出了积极贡献。**在社会价值方面**，CL 已助力全球 20 余个国家的城市项目更新。这些项目的成功实施，改善了居民的生活环境，提升了城市的整体形象。

5.6.2 案例 2：P 公司商业航天卫星智能化

背景与介绍

P 公司成立于 2006 年，面向政府、企业及特种领域用户提供以数字地球产品为核心的

软件销售和数据服务、技术开发服务、数字地球一体机和系统集成。该公司深度构建商业航天产业生态，立足自身技术沉淀与资源优势，全力推进全产业链发展战略，逐步实现了从产业链中下游到中上游的覆盖触达。2025 年 DeepSeek 开源大模型发布后，该公司第一时间便完成该模型的本地化部署，致力于构建航天领域行业大模型。在当今商业航天快速发展的背景下，卫星数据的采集、分析与服务的智能化需求日益迫切。通过大模型等赋能，利用先进的技术手段，实现了卫星数据采集、分析与服务的全链路智能化。通过强化学习与多模态数据处理技术，对卫星任务调度与数据处理效率进行优化，以满足不断增长的商业航天市场需求^[2]。

AI 技术具体应用

在轨道规划环节，AI 能够动态计算卫星过境时间与能源消耗。综合考虑卫星的任务需求、轨道参数、地球的自转公转等多种因素，通过复杂的算法模型，精确计算出卫星的最佳过境时间，以获取更多有价值的数据。同时，通过优化能源消耗，延长卫星的工作寿命。**在数据处理环节**，AI 能够自动解译遥感影像。通过对影像中的各种特征进行识别和分析，生成灾害评估、资源勘探等专题报告。如，在灾害评估中，AI 可以根据遥感影像中的水体变化、植被覆盖情况等信息，准确判断灾害的类型和范围，并评估其对环境和人类社会的影响。**在故障预测环节**，基于历史数据的异常检测模型发挥着重要作用。该模型通过对卫星历史运行数据的学习和分析，建立起正常运行状态的模型。当卫星运行数据出现异常时，模型能够提前预警卫星部件故障，以便及时采取维修措施，确保卫星的正常运行。

成果与效益分析

在内部的效率方面，利用融合太空管理场景的 AI 模型，快速处理和分析海量航天测控数据，自动生成分析报告，协助快速定位问题，提升数据分析效率，同时助力团队更精准地监控任务状态并快速识别潜在风险，快速生成和优化航天任务规划方案，利用智能化技术优化卫星轨道设计、地面站调度等复杂流程，能够提升任务执行效率和资源利用率，建立故障诊断模型，辅助工程师快速诊断和解决航天器故障，提高故障诊断准确率。**在技术突破方面**，接入大模型后，用户意图识别和逻辑推理能力整体提升 10%，同时也为用户提供零代码开发的智能体快速构建流程，提升空天信息行业垂类应用智能体构建效率 30%^[3]。P 公司深度融合卫星遥感、导航定位、航空监测等空天数据资源，形成覆盖全球的多维地理信息数据库与自主可控的遥感智能算法库。**在商业价值方面**，基于空天信息在线能力构建的生态应用输出有效助力了政企客户实现精准化、智能化管理升级，月活跃应用数达 12375 个，历史累计触达 10149 万个终端用户，日均 Tokens 消耗量达 2.29 亿次，受到业内一致认可^[4]。

5.6.3 案例 3：Q 科技打造星载 AI 赋能应急减灾体系

背景与介绍

该公司成立于 2020 年，致力于通过星载边缘计算技术重构卫星数据处理范式。其核心产品采用超异构计算架构实现卫星在轨实时数据压缩、目标识别与情报提取，将传统遥感信息处理时延从小时级压缩至分钟级。

AI 技术具体应用

该公司的技术贯穿灾害“监测-评估-响应”全链条，形成天基智能应急服务体系。灾前监测与预警，多源数据融合，通过光学、SAR 卫星实时捕捉地表形变、水体变化等异常信号，结合 AI 模型预判山体滑坡、洪涝等风险。短报文联动，与北斗星座集成，实现关键预警信息分秒级下传至地面指挥中心。灾情动态评估，在轨目标识别，AI 算法自动解译受灾区域建筑物损毁、道路阻断、人群聚集等关键信息，生成热力图辅助决策。数据压缩优化，采用 AI 图像压缩技术，将原始数据量大幅缩减，提升传输效率。救援资源调度，多星协同规划。星座级智能调度系统动态分配卫星观测任务，确保灾区全天候覆盖。与低空无人机、地面传感器联动，构建“天基预警-空中侦查-地面救援”闭环，实现空天地一体化。

成果与效益分析

产品效益方面，公司推出的核心产品基于高可靠边缘异构计算平台、高性能算法及软硬件协同优化能力，面向国防、应急等高时效业务和民商智能化服务，提供软硬件一体的算力平台+AI 处理解决方案，具有低成本、高可靠、高算力特点，适配典型卫星应用场景，实现卫星价值提升^[5]。

技术效益方面，通过调度多颗卫星、提高重访频率，目前卫星应急响应速度已经压缩至一小时内。而且可以在卫星拍摄后直接对图像进行处理和分析，形成重要情报信息，通过北斗短报文等快速链路下发至终端，全流程耗时仅需数分钟^[6]。形成“天基算力+AI 模型+短报文”的应急服务范式的模式创新。

5.6.4 案例 4：某 R 知名货运航空机坪智能识别

背景与介绍

基于计算机视觉技术，R 公司通过监控航班在停机位附近出入、开关舱门、车辆和人员保障到位情况的开展，可以及时掌握航班生产运行过程中，生产保障行为开展的进度情况，识别关键保障节点，并对出现的异常情况进行识别，将相关数据与报文、雷达、地面保障指挥调度系统等业务系统中采集到的数据进行相互印证，确保保障动态的及时准确采集的科技项目。

AI 技术具体应用

(1) 飞机器位置及状态

包括飞行器入位、飞行器推出、货舱门开启、货舱门关闭、机组舱门开启、机组舱门关闭。其中飞行器入位是个判断难点，业务上判断标准为放置飞机轮档，但由于摄像头的角度无法拍摄到机轮，无法直接识别。通过结合业务分析和目标侦测时序分析技术，采用机务人员待飞机停止后走近飞机并走出的直接视觉判断达到间接判断飞行器入位节点的功能，从而

解决了该问题，并落地使用效果良好，准确度高。



图 5.3 全货机及拖头车

（2）机坪特种车辆类型识别

通过对包括牵引车、传送带车、机组车、平台车、拖头车在内的机坪特种车辆的建模特殊训练，实现对其快速准确识别及目标跟踪。

该项目是国内首例航空货运基于计算机视觉的机坪保障案例，涵盖了机坪全部货运特种车辆的视觉识别技术，填补了国内机坪货运特种车辆的模型算法。

成果与效益分析

航班监控趋向精细化。通过智能机坪作业保障识别实现多个保障任务、保障节点、告警规则的自动计算，覆盖航空器、特种车辆、机坪设施和机坪工作人员等航班保障要素，满足不同岗位个性化的航班管控需求，发挥事前预警和事中干预效用。**资源派遣趋向智能化。**通过智能航班节点识别实现引导车、牵引车等岗位的智能调度，水车、行李装卸等岗位的通用派工，通过“规则配置+算法引擎”实现全智能化自动派工。每个业务岗位的保障任务精确到分钟级，有效提升资源利用率和劳动生产率。**消息传递趋向精准化。**通过智能机坪作业保障将各保障节点和安全事件数据发送至监控终端进行存储、可视化显示和事件回溯，实现跨岗位协同类工作的流程化、标准化、数据化，大幅提升工作效率。

5.7 AI 赋能金融行业应用案例

5.7.1 案例 1：GS 银行：大模型重构金融决策链

背景及介绍

在金融行业，信贷业务作为银行等金融机构的核心业务之一，一直扮演着至关重要的角色。然而，随着经济的快速发展和市场的不断变化，信贷业务面临着诸多新的挑战 and 难题。尤其是对于中小企业信贷风险识别难、人工审批效率低等痛点，传统的人工经验判断和简单的模型分析已经难以满足实际需求。在此背景下，GS 银行于 2024 年启动“AI+信贷”战略，依托大模型构建智能化风控体系，覆盖贷前调查、贷中监控、贷后管理的全生命周期。这一举措旨在通过 AI 技术提升信贷业务的风险管理能力和运营效率，为中小企业提供更优质、高效的金融服务，同时也为银行自身的发展提供新的动力。

在传统信贷模式下，风险识别主要依赖于人工审核和简单的模型分析，这种方式不仅效率低下，而且难以准确识别潜在风险。对于中小企业来说，由于其财务信息不透明、经营规模较小等原因，信贷风险识别难度更大。同时，人工审批流程繁琐，耗时较长，难以满足企业快速融资的需求。为了应对这些挑战，GS 银行积极探索 AI 技术在信贷风控领域的应用，以提高风险识别的准确性和审批效率，为中小企业提供更便捷、高效的信贷服务。

AI 技术具体应用

逻辑推理引擎：AI 模型强大的逻辑推理能力，能够解析企业财报与非结构化数据（如合同文本）。通过多级因果推断，该引擎可以识别出隐藏在复杂数据背后的隐形关联风险。如，它可以通过分析企业财务报表中的各项指标以及合同条款中的细节，推断出企业可能面临的经营风险、财务风险等，为信贷决策提供更全面、准确的风险评估。

动态知识图谱：GS 银行整合工商、税务、司法等 10 类外部数据源，构建了一个动态更新的企业关系网络图谱。这个图谱能够自动检测关联交易与担保圈风险，帮助银行更清晰地了解企业之间的关联关系以及潜在的风险传导路径。通过这种方式，银行可以及时发现企业可能存在的复杂关联风险，避免因信息不对称而导致的信贷风险。

自动化报告生成：AI 技术的应用使得贷审报告的生成更加高效、准确。AI 生成的贷审报告包含 60+项核心指标，涵盖了现金流预测与行业对标分析等多个方面。借助自动化报告生成功能，审批时效从原来的 3 天大幅压缩至 4 小时，不仅提高了工作效率，还减少了人为因素对报告质量的影响，为信贷审批提供了更可靠、及时的依据。

成果与效益分析

风险识别准确率显著提升：通过上述 AI 技术的应用，GS 银行的信贷风控系统风险识别准确率提升至 97.3%。2024 年该系统成功拦截高风险贷款申请 120 亿元，有效避免了潜在的信贷损失，为银行的资产安全提供了有力保障。

中小微企业信贷审批效率大幅提高：智能化的风控体系使得中小微企业信贷审批效率大幅提升，人工复核工作量大幅减少。这不仅缩短了企业获得贷款的时间，提高了金融服务的时效性，还降低了银行的人力成本，实现了双赢。

入选示范工程，推动普惠金融发展：该项目凭借其创新性和显著的成效，入选央行“金融科技赋能乡村振兴”示范工程。在推动乡村振兴方面，GS 银行的普惠贷款余额突破 2 万亿元，为农村地区的企业和个人提供了更广泛、更优质的金融服务，有力地支持了乡村经济的发展。

5.7.2 案例 2：TB：智能保险产品推荐系统

背景及介绍

随着市场环境的变化和客户需求的日益多样化，传统保险产品推荐模式面临着诸多挑战。为了更好地满足客户需求，提高市场竞争力，TB 于 2023 年启动“AI+保险”战略，构建智能保险产品推荐系统。该系统通过深度分析客户数据，提供个性化的保险产品推荐，覆盖人寿

保险、健康保险、财产保险等多个领域。

在传统保险产品推荐模式下,主要依赖于保险代理人的经验和客户的基本信息进行人工推荐,这种方式不仅效率低下,而且难以精准把握客户的实际需求。为了应对这些挑战,公司积极探索 AI 技术在保险产品推荐中的应用,以提高推荐的准确性和效率,为客户提供更优质、个性化的保险服务。

AI 技术具体应用

客户画像构建:整合客户的基本信息、消费行为、健康状况、家庭结构等多维度数据,构建 360°客户画像。通过自然语言处理技术分析客户的社交媒体言论和咨询记录,挖掘潜在保险需求。

智能推荐引擎:基于客户画像和历史购买行为, AI 模型实时推荐匹配的保险产品组合。如,为有孩子的家庭优先推荐教育金保险和重疾险,为经常出差的客户推荐航空意外险和交通意外险。

动态调整模块:系统持续跟踪客户生命周期变化,如结婚、生子、购房等,动态调整保险产品推荐。同时,根据市场利率波动和保险产品条款更新,优化推荐方案。

成果与效益分析

客户满意度显著提升:通过精准的产品推荐,客户满意度提升至 92.5%, 2024 年客户投诉率降低 60%。个性化推荐使客户感受到服务的专属性和专业性。

保险产品销量增长:智能化推荐系统使保险产品销量大幅增长,其中健康险和意外险销量增长尤为显著。

运营效率提高:AI 技术应用使保险产品推荐效率大幅提升,人工工作量减少一半以上。系统自动生成推荐方案,代理人可将更多精力用于客户关系维护和复杂案例处理。

5.7.3 案例 3: THS i 问财: AI 投顾决策辅助系统

背景及介绍

THS 于 2025 年推出“i 问财 2.0”,成为全球首个将大模型技术深度应用于投资领域的智能投顾产品。该系统以诺贝尔经济学奖得主丹尼尔·卡尼曼的“快慢双系统思维模型”为理论基础,通过深度强化学习技术强化 AI 的“慢思考”能力,覆盖选股、诊股、资产配置、风险预警等全链条投资场景,旨在为投资者提供更理性、透明、精准的决策支持。

传统投顾模式下,投资者需依赖人工顾问的经验解读海量数据,存在信息获取滞后、分析主观性强、服务成本高等问题。i 问财 2.0 通过 AI 技术重构投顾服务范式,整合实时金融数据、专业研究资源与多模态交互能力,实现从“经验驱动”向“算法驱动”的跃迁。

AI 技术具体应用

深度推理分析与决策拆解。应用思维链,将复杂投资问题拆解为逻辑清晰的推理步骤。如,用户输入“分析某股票走势并提出买入建议”时,系统自动分解为技术指标分析、市场情绪评估、基本面验证等子任务,模拟人类深度思考过程。支持多维度数据联动分析,如结合

K 线图、财报数据、行业舆情等，生成综合投资建议。

智能知识调用与动态学习。实时调用 THS 专业金融数据库（覆盖全球指数、A 股、基金等 15 个领域）及第三方研究报告，确保决策依据的全面性与时效性。通过自我修正机制，持续评估输出结果并优化模型参数。如，在股票诊断场景中，系统可基于用户反馈和市场实际表现调整预测模型，提升准确率。

复杂任务处理与风险控制。在资产配置优化中，系统结合用户风险偏好、资金规模、投资周期等参数，运用蒙特卡洛模拟与动态规划算法生成个性化组合方案，并实时监控市场波动以调整仓位。风险预警模块通过图神经网络技术实时扫描交易数据，拦截欺诈行为准确率达 99.7%。

成果与效益分析

决策科学性与用户满意度提升。系统每日服务超 500 万投资者，用户满意度达 90% 以上。在股票诊断场景中，深度推理分析使投资建议准确率提升至 87.5%，显著高于传统投顾的 65%。案例显示，用户通过 i 问财 2.0 优化持仓组合后，波动率降低 20%。

服务效率与覆盖范围扩展。支持 50 余类投资技能（如选股、行业分析、事件脉络梳理），服务响应时间缩短至秒级，人工投顾工作量减少 70%。在普惠金融领域，系统通过手机银行等渠道降低服务门槛，农村及小微企业用户覆盖率提升至 85%。

商业价值与行业影响力。2025 年，天风证券等机构评价其为“AI 投顾 2.0 时代的范式标杆”，推动行业从“工具辅助”向“决策主导”转型。

本章参考文献

5.1 AI 赋能制造行业应用案例

- [1] Baosteel's AI Implementation, Honeywell Forge Article
- [2] Ansteel's Smart Manufacturing, China Daily Article
- [3] 10 ways AI is being used in the Steel Industry [2025], DigitalDefynd Article
- [4] 科远&日照钢铁：无人运行+联合智能调度，国内最大规模冶金无人库效能倍增_库区_行车_作业
- [5] 一种基于图像分割的非接触式皮带撕裂检测系统及方法.pdf-原创力文档
- [6] 宝钢新品热轧工作辊缺陷的超声检测-会议-钛学术文献服务平台
- [7] 实现全产业链、价值链协同降碳系统布局 河钢与必和必拓携手探索原料端深度降碳
- [8] “双鞍”携手在全球工业互联网大会上促合作

5.2 AI 赋能信息科技行业应用案例

- [1] 利用英伟达 Grace Blackwell 和 AI 新思科技加速芯片设计

5.3 AI 赋能材料行业应用案例

- [1] 人工智能大幅提升科研效率 创新成果加速涌现 北京引领“AI+新材料”风口-千龙网·中国首都网

5.5 AI 赋能健康行业应用案例

- [1]Esteva A et al. Deep learning-enabled medical computer vision. NPJ Digital Medicine 2025;8(1):1-9.
- [2]《Big Ideas 2025》, ARK Invest.
- [3]AI 辅助诊断入医保：智慧医疗的发展新机遇。
- [4]AI 赋能医疗影像诊断：技术革新与临床实践深度解析（2025）。
- [5]腾讯觅影官网。
- [6]早期食管癌检出率达 90%，腾讯觅影 AI 服务我市多家医院。
- [7]腾讯首次详解“觅影”产品：AI 医学影像技术如何应用于肿瘤早筛。
- [8]腾讯觅影 AI 如何改变眼科筛查的未来。
- [9]AI 辅助诊断系统立功！科技部表彰腾讯公司抗击新冠肺炎疫情。
- [10]腾讯觅影青光眼筛查功能发布 眼底疾病社区筛查开启“AI 模式”。
- [11]腾讯医疗 AI 系统“觅影 2.0”：癌症早筛革命的里程碑。
- [12]AI+医疗普惠革命。
- [13]腾讯携手浙江德清打造县域 AI 医学示范基地。
- [14]60 秒写病历、30 秒自动阅片……迈瑞以“A+医疗”加速数智化。
- [15]《中国防治慢性病中长期规划（2017-2025 年）》。
- [16]上海复星健康科技（集团）有限公司内部资料。

5.6 AI 赋能空间行业应用案例

- [1]CityTalk | CL：百亿级 AI 颠覆设计界！
- [2]洞察者融合 DeepSeek 大模型，打造航天测控智能化基石|航天_新浪财经_新浪网
- [3]正式接入 DeepSeek 大模型，星图云服务再升级！-资讯动态-星图云开放平台
- [4]证券时报 | 中科星图 2024 年营收同比增长 29.49%，空天一体战略迈向高质量发展新篇章
- [5]奋进新征程 | AI+卫星，迈向天基智能时代 - 星测未来科技（北京）有限责任公司
- [6]解决方案 | 星载 AI 赋能应急减灾 - 星测未来科技（北京）有限责任公司

5.7 AI 赋能金融行业应用案例

- [1]GS 银行金融科技部. 工银智涌：企业级金融大模型技术白皮书[R]. 北京：中国 GS 银行, 2025.
- [2]财联社. 工行完成 DeepSeek 全行大规模应用，AI 驱动信贷风控质效双升[N]. 财联社, 2025-03-08.
- [3]沙丘智库. 银行业 DeepSeek 大模型应用跟踪报告 (2025Q1) [R]. 上海：沙丘研究院, 2025.
- [4]中国太平洋保险. 智能保险推荐系统技术白皮书（2024 版）[R]. 上海：TB 集团, 2024.
- [5]国务院办公厅. 关于推进保险业数字化转型的指导意见（国办发〔2023〕9 号）[Z]. 2023-05-12.

- [6]中国保险行业协会. 保险业 AI 应用案例集 (2024) [M]. 北京: 中国金融出版社, 2024.
- [7]THS 研究院. i 问财 2.0: 大模型驱动的智能投顾技术白皮书[R]. 杭州: THS 科技, 2025.
- [8]IDC. 中国智能投顾市场竞争力分析报告 (2025) [R]. 北京: IDC 中国, 2025.
- [9]中国证券业协会. 智能投顾服务合规指引 (2025 版) [S]. 北京: 中国金融出版社, 2025.

第6章 国内外人工智能产业政策及趋势分析

6.1 人工智能治理困境与监管沙盒

人工智能在给世界带来巨大机遇的同时,也带来了难以预知的各种风险和复杂挑战。在人工智能产业飞速发展的当下,如何平衡创新与风险是人工智能治理的核心命题。

6.1.1 人工智能治理难点

(1) 监管者和被监管者信息不对称

监管者很难准确把握技术发展,尤其类似人工智能高速迭代过程的实时动态。一般来说,政府作为监管主体,在识别治理对象及其风险、权衡治理目标、选择治理工具时需要掌握技术创新和应用的最新动态;而被监管者作为一线创新者和应用者,对技术发展和技术风险问题具有天然的信息优势,双方需要通过合适的“委托—代理”结构设计来缓解信息不对称问题。但是人工智能创新应用内在逻辑的复杂性“黑箱”进一步扩大了双方的信息差距,甚至有可能造成“共同无知”的新局面,从而放大了监管者和被监管者之间信息沟通、合作治理的难度。同时,被监管者很难明确监管者的治理目标。传统技术系统自身的运作逻辑、功能性能、应用场景比较清晰,致风险的原因及需要实现的治理目标往往也较为明确。但是人工智能发展的快速迭代和不可预测性导致监管者需要不断探索、权衡、调整治理目标和治理机制。在此背景下,被监管者需要及时了解政府规制关切的最新变化,适应人工智能时代治理要求的难度也在加大。

(2) 风险防范成本和效益不对等

人工智能对人类社会的影响具有颠覆性,如果遭到误用、滥用、恶用,可能引发社会、经济、军事等领域的重大安全威胁,一旦脱离人类控制甚至有可能对人类生存构成威胁。如果不加以监管和治理,这些切实存在的安全风险可能造成不可逆的严重后果。但另一方面,不同于需要大量基础设施投入、物理可见性较强的传统公共危险源,人工智能的通用性和易得性使其潜在风险几乎无处不在,极大提高了风险防范的社会成本。要全面防范人工智能带来的安全风险需要耗费大量的公共资源,甚至以部分牺牲人工智能发展为代价。如何在高昂的风险防范成本、巨大的潜在风险损失、不断提升的风险概率之间权衡,是监管者必须面对的重要考验。

(3) 机制复合体治理不协调

人工智能治理是涉及法律法规、行业标准、国际协调的复杂系统,需要不同国家、国际组织、相关企业的共同参与。如何在机制复合体的复杂框架下实现国际治理协调,是人工智能治理亟待解决的重要问题。

目前,以联合国为代表的国际组织、以七国集团为代表的多边合作机制、以中美为代表的人工智能大国、以微软和谷歌等为代表的跨国企业等各类国际行为体都在深度参与人工智能治理。这些参与主体在治理理念、关注议题、治理偏好、治理能力上存在差异,导致人工智能国际治理呈现嵌套、重叠、平行的碎片化局面,存在合理性、公平性、有效性等一系列

治理赤字。在此背景下，如何协调这些相互重叠、不同层次、不同诉求的参与主体，建立受到广泛认可的人工智能全球治理机制复合体，是人工智能治理的重要挑战。

(4) 地缘政治环境不稳定

一方面，人工智能国际治理需要掌握前沿技术的大国建立沟通协调机制，通过信息交换和自我控制在国际治理体系中发挥更为积极的作用。但另一方面，人工智能拥有从经济、军事等多个维度影响国家力量对比，甚至引发新一轮大国兴衰的巨大潜力。大国战略竞争的地缘政治环境导致大国协调变得更加困难，部分国家甚至以意识形态划线或构建排他性集团，恶意阻挠他国人工智能发展，利用技术垄断和单边强制措施制造发展壁垒，恶意阻断全球人工智能供应链，不但阻碍全球人工智能发展，对人工智能治理也产生了负面影响。

6.1.2 监管沙盒

鉴于人工智能技术对于高质量发展的有益性与风险性，各国政府正积极探索人工智能“监管沙盒”制度，以期在人工智能监管中统筹支持创新与防范风险两个维度。

“沙盒”一词最早源于计算机安全领域的“沙盒”技术（Sandbox），旨在为一些来源不可信、具备破坏力或无法判定意图的程序提供试验环境。首次将“沙盒”由技术领域嵌入管理范畴的有益尝试发生在金融科技监管领域。为了应对金融科技发展给金融监管体制带来的挑战，实现金融创新与金融监管之间的动态平衡，英国政府科学办公室于2015年3月发布的一份有关金融科技的报告首次提出“监管沙盒”。2015年11月，英国金融行为监管局在其发布的《监管沙盒》研究报告中，“监管沙盒”被阐述为是一个“安全空间”，企业可以在其中测试创新产品、服务、商业模式和交付机制，并且不会因为从事上述活动而立即招致通常的监管后果。通过“申请—审核—测试—报告—退出”的测试程序，金融行为监管局可以直观监测金融科技的风险，并将其控制在一定范畴内；同时，测试企业在沙盒内享有一定的责任豁免权。该制度于2016年6月在英国正式启动。

在人工智能领域中，“监管沙盒”成为全球治理的重要尝试。2019年，经济合作与发展组织发布的《人工智能原则》建议各国政府考虑尝试建立一个可控的环境供人工智能系统进行测试和扩大规模，监管者可以在现行监管框架之外测试新兴经济业态、制度和技术方法以及法律条款，而这种实验型的监管方法就包括“监管沙盒”。近年来，包括西班牙、新加坡、英国等国家（地区）积极在人工智能监管领域构建“监管沙盒”制度，企图抓住人工智能产业变革先机，成为人工智能全球治理的领导者。

西班牙于2022年建立人工智能监管沙盒，用来检验并完善欧盟《人工智能法》。因为欧盟人工智能法要在2025年正式实施，所以，通过在某个成员国建立监管沙盒，比照欧盟人工智能法，检验该法能否有效监管人工智能。如果发现缺陷，可以边试验边完善。

新加坡则联合谷歌、微软、亚马逊等人工智能头部公司建立基于治理科技的人工智能监管沙盒。该监管沙盒旨在开发人工智能技术透明性的自动评估技术。遴选入沙盒的公司可以知晓其开发的人工智能技术是否符合国际技术标准和监管标准。

英国则针对人工智能应用的不同行业分别建立监管沙盒制度。比如, 2023 年, 英国医疗卫生部下属的医疗卫生局建立医疗人工智能沙盒。2024 年春季, 医疗卫生局推出第一个监管沙盒, 为人工智能气闸室这种产品建立监管沙盒。这个监管沙盒把行业内的专家和监管部门集合起来, 组成监管团队。申请进入沙盒的市场主体将和监管团队合作, 把技术和产品研发过程中的信息提供给团队, 团队再把信息进行分析 and 加工, 向监管部门汇报。经过这种互动, 不仅提升技术和产品的性能, 降低风险, 也提升监管水平, 实现监管与技术水平的同步提升。

6.2 国内人工智能产业政策及环境分析

6.2.1 国家层面政策规划及布局

面对平衡人工智能产业创新与风险这一命题, 我国的人工智能产业政策从早期布局逐步走向深化, 并形成了从战略布局到深度应用的全周期政策体系, 体现出清晰的渐进式演进脉络。

(1) 早期布局阶段

2015 年至 2017 年是我国人工智能产业政策的早期布局阶段。2015 年国务院出台的《中国制造 2025》首次提及智能制造, 并提出“以推进智能制造为主攻方向”的明确要求, 制造业成为人工智能最先落地的行业之一。商汤科技、寒武纪科技为代表的产业界及学术界开始关注并引领中国人工智能产业的发展。这一时期的政策, 如《关于推进物联网有序健康发展的指导意见》《关于积极推进“互联网+”行动的指导意见》《关于印发促进大数据发展行动纲要的通知》等均以总体规划性文件为主。但这些政策文件对大数据、5G、物联网等领域的初步部署, 实际上为人工智能产业后续发展奠定了坚实的基础。

(2) 战略升级阶段

2017 年至 2021 年, 国家层面开始强化顶层设计与政策引导, 产业政策进入战略升级阶段。2017 年, 国务院发布第一份人工智能行业的系统性部署文件《关于印发新一代人工智能发展规划的通知》, 重点对 2030 年我国新人工智能发展的总体思路、战略目标和主要任务、保障措施进行系统的规划和部署, 政策重心转向核心技术突破与产业融合, 人工智能上升为国家战略。该政策文件确立了我国人工智能发展“三步走”战略目标, 包括: 1) 到 2020 年, 人工智能总体技术和应用与世界先进水平同步, 成为新的重要经济增长点; 2) 到 2025 年, 人工智能基础理论实现重大突破, 部分技术与应用达到世界领先水平, 成为产业升级和经济转型的主要动力; 3) 到 2030 年, 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。从当下这个时间点来看, 2025 年我国实际上已经基本实现了该文件的技术和应用突破目标, 部分技术与应用甚至已经接近或达到世界领先水平。

伴随 2017 年的人工智能战略升级, 我国也进入了人工智能产业政策的密集发布期, 人工智能不再局限于智能制造及新兴行业的定位, 而是延伸到了赋能实体经济的战略高度。在 2017 年全国两会提出“加快人工智能等技术研发和转化”, 此后九年, 人工智能发展政策均被

纳入两会政府工作报告^[1]。为了细化和落实部署，2017年12月，工信部发布《促进新一代人工智能产业发展三年行动计划(2018-2020年)》，详细规划了人工智能在未来三年的重点发展方向和目标。2019年3月，中央全面深化改革委员会审议通过的《关于促进人工智能和实体经济深度融合的指导意见》指出，要构建数据驱动、人机协同、跨界融合、共创分享的智能经济形态，人工智能产业政策日渐深入。

(3) 目前情况

2021年是“十四五规划”的开局之年。“十四五规划”将科技自立自强作为国家发展的战略支撑，并将人工智能列为重点发展产业之一。人工智能成为我国现代产业化体系的核心设施之一。除了继续引导发展之外，政策层面开始关注伦理安全等人工智能治理议题，出现了效力层级较高的法规制度，呈现发展与治理并举的特征，政策体系进入规范深化期。如：

2021年，中办、国办发布的《关于加强科技伦理治理的意见》，对新时代我国科技伦理治理工作做出了全面、系统的部署，确立人工智能政策的价值框架；《互联网信息服务算法推荐管理规定》要求平台透明化算法逻辑，禁止“大数据杀熟”；《深度合成管理规定》，要求AI服务提供者标识内容来源；《数据安全法》与《个人信息保护法》严格限制数据跨境流动，要求AI企业需建立本地化数据存储机制。

2022年，科技部等六部门出台《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》，以针对性地解决人工智能产业发展面临的场景创新不足、创新生态不完善等实践问题。

2023年，国家网信办等七部门发布《生成式人工智能服务管理暂行办法》，明确内容安全、数据合规要求，要求实施备案与安全评估制度，成为全球首部生成式AI专门立法。

2024年，工业和信息化部、国家网信办、国家发展改革委、国家标准委等四部门联合印发《国家人工智能产业综合标准化体系建设指南》，构建覆盖算法、芯片、应用场景的标准化体系，保障技术安全与产业协同；国家数据局等十七部门联合印发“数据要素×”三年行动计划（2024-2026年），推动数据要素与AI深度融合，优化算力资源布局，探索数据跨境流动试点。

纵观我国人工智能产业政策的演进过程，一方面，人工智能产业不断实现战略升级，角色定位从新兴到赋能实体经济再到现代产业化体系的核心设施，2025年政府工作报告中首次出现了“AI大模型、智能网联汽车、AI手机、智能机器人、具身智能等未来产业的表述，标志着人工智能产业的外延进一步扩大；另一方面，2021年以来，我国人工智能产业政策

¹ 具体包括如下主题：

2018年：“加强新一代人工智能研发应用”；

2019年：“深化人工智能等研发应用”；

2020年-2023年：“推进智能制造”、“推动产业数字化智能化改造”、“大力推进智能制造，培育壮大人工智能等数字产业”、“加快传统产业和中小企业数字化转型”；

2024年：“深化人工智能等研发应用，开展“人工智能+”行动”；

2025年：“持续推进“人工智能+”行动，支持大模型广泛应用，大力发展新一代智能终端及智能制造装备，培育具身智能等未来产业”。

开始关注技术标准化、技术伦理等议题，政策关注点从引导发展走向更为深入的人工智能治理。在继续通过政策扶持、资金支持等方式推动行业快速进步，鼓励人工智能技术创新和产业发展的同时，国家政策层面注重防范人工智能可能带来的数据泄露、算法偏见、虚假信息传播等风险，以保障国家安全、社会稳定和公民合法权益。这一渐进式的政策演进模式，回应了人工智能产业发展的实际，为制度创新预留了空间，也保证了监管的及时性与灵活性。

6.2.2 地方政府政策

近年来，我国各地方政府也出台了一系列支持人工智能产业发展的政策措施。

北京市发布《北京市推动“人工智能+”行动计划》，目标到2025年底形成3~5个先进可用、自主可控的基础大模型产品、100个优秀的行业大模型产品和1000个行业成功案例，率先建设AI原生城市，推动北京市成为具有全球影响力的人工智能创新策源地和应用高地。

上海市近年来先后出台了《关于加快推进人工智能高质量发展的实施办法》《上海市促进人工智能产业发展条例》，着力将上海打造成全国领先的人工智能创新策源地、应用示范地、产业集聚地和人才高地；《上海市推动人工智能大模型创新发展若干措施》旨在深入贯彻国家发展新一代人工智能的战略部署，推动上海大模型创新发展，营造通用人工智能创新生态，加快打造世界级人工智能产业集群，在支持大模型创新能力、大模型创新应用、创新要素供给能级等方面提出了具体举措。

浙江省一直是人工智能发展的创新高地。以宇树科技、深度求索等人工智能企业为代表的杭州“六小龙”出圈，引起了海内外的高度关注，被外界称为“神秘的东方力量”。这些年，浙江省在政策环境、基础要素、重大项目等方面持续用力，为人工智能产业发展奠定了坚实基础。2025年，浙江省将出台《浙江省“人工智能+”行动计划（2025-2027年）》，鼓励地方政府、领军企业、行业协会和专业服务机构加大人工智能场景开放，促进人工智能赋能千行百业。同时将聚焦为行业提供共性解决方案，支持建设一批“人工智能+”重大项目，在用地、用能、排污权和资金等方面给予支持。

山东省发布了《山东省人工智能产业科技创新行动计划（2025—2027年）》，提出到2027年，培育20家人工智能领军企业；新引育90名国家级、150名省级科技人才；打造一批细分领域产业集群，人工智能核心产业产值规模突破1000亿元，带动相关产业规模超万亿元。

安徽省发布《安徽省通用人工智能创新发展三年行动计划（2023—2025年）》等文件，用真金白银对人工智能产业给予补贴，提出要优先匹配算力、数据等要素资源，对关键技术项目所在市给予市级补助金额30%、最高1000万元奖补；支持创新应用大赛获奖项目和团队在皖落户创业，综合支持最高3000万元。

广东省发布《广东省新一代人工智能发展规划》，提出主动把握新一代人工智能发展重大战略机遇，加快建设国家科技产业创新中心，奋力实现“四个走在全国前列”；《广东省推动人工智能与机器人产业创新发展若干政策措施》，则围绕技术攻关、产业生态构建等12

个维度推出系统性支持政策。其中提出，对研发工业领域大模型和应用解决案例给予支持，每年择优支持不超过 10 个标杆案例，每个给予最高 800 万元奖励。

这些政策从技术研发、产业布局、人才培养、资金支持等多个维度，为人工智能产业提供了全方位的支持，加速了人工智能在各地形成百花齐放的态势。

6.3 国际人工智能整体产业政策

推动人工智能技术、产业的发展一直是各国的共识，当前，美国、欧盟、英国、日本等国家和地区，纷纷更新或建立人工智能产业政策，力求掌握新时代发展先机。

6.3.1 美国“创新优先”战略

美国以立法、军事应用和出口管制为三大支柱，全面强化其人工智能领域的全球主导地位。

（1）立法层面

《2024 年人工智能倡议法案》通过设立金额高达 250 亿美元的国家 AI 研究基金，整合联邦机构（如国防部、能源部）与科技巨头（微软、谷歌、英伟达等）资源，建立国家人工智能研究资源试点，重点支持基础模型研发和测试平台建设。该法案还授权美国国家标准与技术研究院（NIST，直属美国商务部）成立人工智能安全研究所，推动公私合作制定技术标准，并组建国际联盟以协调全球 AI 治理规则。

（2）军事领域

美国国防部实施“AI 作战网络”计划，通过“雷霆锻造”项目将 AI 深度嵌入军事决策流程，覆盖战略规划、资源分配和战役推演等环节。该项目联合 Scale AI、Anduril 等企业，利用大模型 LLM 和智能体 Agent 技术提升作战响应速度，并计划在印太和欧洲司令部先行部署，最终扩展至全军。同时，美国放宽军事 AI 应用限制，允许 AI 参与武器系统开发和战场决策，尽管此举引发伦理争议。

（3）技术管控方面

美商务部更新出口管制清单，将 14nm 以下 AI 芯片及制造设备纳入严格限制，要求台积电、三星等晶圆厂对客户实施“调查”，防止先进制程芯片流向中国实体清单企业。新规还引入“长臂管辖”，限制第三方国家向中国出口含美国技术的半导体设备，并扩大实体清单范围至 140 家中国科技企业及投资机构。

这一系列措施形成“创新驱动-军事赋能-技术封锁”的闭环战略，旨在巩固美国在 AI 产业链顶端的控制力。

6.3.2 欧盟“风险为本”监管

欧盟通过《人工智能法案》构建了全球首个系统性人工智能治理体系。该法案首创四级风险分类体系，将 AI 系统按威胁程度划分为不可接受风险（如大规模监控和生物识别滥用）、高风险（涉及医疗、司法等八大领域）、有限风险（需透明度约束的聊天机器人等）及最小风险（如智能推荐）。为统一监管尺度，欧盟成立欧洲 AI 委员会，统筹协调 27 国

市场监管机构与司法机关,形成“欧盟-成员国”双层监管架构,并赋予其发布高风险模型清单、制定技术标准及推动国际规则协调的权限。

在产业培育方面,欧盟通过 100 亿欧元主权 AI 基金支持本土创新,作为 2000 亿欧元 InvestAI 计划的核心部分,重点建设配备 10 万颗先进芯片的四大 AI 超级工厂,为中小企业提供算力资源,并启动“GenAI4EU”计划推动医疗、制造等 14 个行业的应用落地。

这一“风险分级-监管协同-产业赋能”三位一体战略,标志着欧盟在 AI 伦理治理与技术创新平衡上的深度探索。

6.3.3 英国“敏捷治理”模式

英国以动态监管框架与创新激励机制为核心,构建了全球领先的 AI 治理体系。2023 年发布的全球首份《AI 监管原则适应性指南》,提出安全稳健性、透明可解释性、公平性、问责制及可补救性五大原则,明确将 AI 应用风险分为四类(如不可接受风险与高风险),并允许企业通过“监管托管”制度在监管机构监督下进行高风险应用测试,如在医疗诊断、自动驾驶等领域先行试点,通过沙盒机制平衡技术创新与风险管控。

同时,英国投资 2.5 亿英镑成立国家 AI 实验室,重点推动 AI 在癌症早期筛查、个性化诊疗等医疗场景的研发应用,并通过国家 AI 研究资源计划采购 5000 颗英伟达 GPU,强化算力基础设施。这一模式通过规则适配性、公私协同性和治理敏捷性三大维度,巩固了其在全球 AI 伦理治理与技术创新的双重领导地位。

6.3.4 日本“社会 5.0”战略

日本以技术革新与制度适配双轮驱动构建超智能社会。知识产权领域,2024 年修订的《知识产权战略纲要》首次承认 AI 生成物可依据“市场价值”通过注册制获得类著作权保护,但要求开发者承担生成内容侵权责任,若 AI 作品与现有作品存在“本质特征相似”则需承担民事赔偿及刑事处罚。治理架构方面,2025 年通过的《AI 人机共生社会促进法》确立“人类最终决策权”原则,要求 AI 系统在医疗诊断、司法量刑等高风险场景必须保留人类介入机制,并建立由首相直接领导的跨部门 AI 战略本部统筹政策。经济激励措施上,政府推出“机器人税”优惠政策,企业购置工业机器人可享受 15% 税收抵免,同时制造业实施 AI 化改造的设备投资可减免 30% 固定资产税,该政策已带动丰田、发那科等企业将产线自动化率提升至 78%。

6.4 未来政策趋势

未来我国将着力构建“监管科技深化-制度建设-积极参与国际竞争”三位一体的新型治理体系。

(1) 持续完善监管科技应用深化,推广监管沙盒。监管科技应用将迈向“技术驱动型治理”新阶段。监管沙盒试点范围将扩展至更多地区,允许企业在可控环境中测试 L4 级自动驾驶、生成式 AI 等前沿技术。

(2) 制订一部综合性的人工智能法。欧盟已经率先制定了《人工智能法》,从较高维度总领人工智能治理政策。许多国内学者已经建议中国立法部门启动制订一部综合性的人工

智能法,包括2024年3月19日七所高校专家发布了《中华人民共和国人工智能法(学者建议稿)》、2024年4月16日中国社会科学院法学研究所等机构也起草并发布了《人工智能示范法2.0(专家建议稿)》等。国务院也在2023年和2024年的《立法工作计划》中提及“预备提请全国人大常委会审议人工智能法草案”,但全国人大常委会在相应的立法工作规划中仅使用“推进人工智能健康发展”等表述,尚未出现法律形式的立法规划。尽管如此,一部综合性、高效力层级的人工智能立法,有助于统筹调整目前分部门治理的现状,统筹解决主体权责(如AI生成著作权归属)、产品责任(如自动驾驶算法缺陷连带责任)、跨境数据流(如白名单国家数据通道)等问题。

(3) 中国正以战略主动姿态深度参与全球人工智能治理框架的竞争。在国际规则构建层面,中国政府通过联合国等多边平台积极推进人工智能伦理公约的磋商进程,在数据主权、算法伦理等关键议题的国际对话中力争掌握规则制定主动权,着力构建体现发展中国家诉求的全球AI治理新范式。值得关注的是,欧盟《人工智能法案》的合规性要求与美国出口管制形成的双重压力,已促使中国AI企业在数据隐私保护、算法透明度及伦理责任等合规要求,乃至商业模式设计层面采取更为审慎的应对策略。

展望未来,中美欧在人工智能领域的战略博弈将呈现技术标准争夺与制度体系竞合交织的长期态势。预计中国可能从三个维度强化战略布局:首先,加速构建自主可控的AI产业生态,通过国产大模型迭代升级逐步降低对国际开源技术的依存度;其次,依托“一带一路”建设推动AI技术标准输出,在发展中国家形成具有中国特色的治理方案示范效应;最后,着力推进中文语料库的国际化标准认证体系,通过语言文化载体提升中文大模型在全球人工智能治理体系中的话语权重。这种“技术突围-制度创新-文化赋能”的三维战略,或将重塑全球人工智能治理的权力格局。

6.5 结语

人工智能作为新一轮科技革命的核心驱动力,其产业政策与全球治理格局深刻影响着技术发展与经济转型。从国内视角看,中国通过《新一代人工智能发展规划》等顶层设计,确立了“技术攻关—场景应用—生态构建”三位一体的战略路径,强调以政策引导推动技术产业化,并在智慧城市、智能制造等领域形成规模化优势。然而,核心技术“卡脖子”、数据安全性与伦理风险等问题仍需政策持续优化。

国际层面,以欧美为代表的发达经济体正加速构建“监管先行”的治理框架,如欧盟《人工智能法案》确立风险分级监管体系,美国通过《国家人工智能倡议法案》强化技术领导力与联盟合作。全球政策呈现两大趋势:一是从单一技术扶持转向“创新激励与风险防控”双轨制;二是围绕数据主权、算法透明度和跨境协作的规则博弈日益激烈。

未来,人工智能政策需在技术创新与治理约束间寻求动态平衡。国内政策应加强基础研究投入与开源生态培育,同时完善伦理审查机制;国际层面则亟需构建多边对话平台,推动标准互认与协同治理。唯有通过开放包容的产业生态与科学规范的政策引导,才能实现人工

智能高质量发展与人类福祉提升的双重目标。

本章参考文献

- [1] 国务院. 新一代人工智能发展规划[M]. 北京: 人民出版社.2017.
- [2] 李彦宏. 智能革命: 迎接人工智能时代的社会、经济与文化变革[M]. 北京: 中信出版社, 2017.
- [3] 中国信通院. 中国人工智能产业发展白皮书 (2023) [EB/OL] .2023.
- [4] 薛澜,王净宇.人工智能发展的前沿趋势、治理挑战与应对策略[J].行政管理改革,2024,(08).
- [5] 杨丰一.人工智能“监管沙盒”: 理论溯源、实践考察与制度构建[J].中国行政管理,2024,(12).
- [6] 王佐发.我国如何构建人工智能沙箱监管[EB/OL] .2024.
- [7] Chen H, Wang Y, Zhang Y. Government support and firm innovation in artificial intelligence: Evidence from China[J]. Research Policy, 2022 Vol 51(9).
- [8] Xue L, Liang Z. Policy innovation and industrial transformation: The case of China's AI development strategy[J]. Science China Information Sciences, 2021 Vol 64(3).
- [9] 欧盟委员会. 人工智能白皮书: 通往卓越和信任的欧洲路径[M]. 北京: 人民出版社, 2020.
- [10] 李向阳, 王雪. 全球人工智能政策比较研究[M]. 北京: 科学出版社, 2021.
- [11] 施密特 E. 人工智能与全球权力竞争[M]. 张伟译. 北京: 中信出版社, 2023.
- [12] OECD. AI Policy Observatory: Trends and Data in Global AI Strategies[EB/OL] .2021.
- [13] Brundage M, Avin S, Wang J, et al. Toward trustworthy AI development: Mechanisms for supporting verifiable claims[J]. Journal of Public Policy, 2020 Vol 40(4).
- [14] Hagendorff T, Danks D, Pasquale F. Global AI ethics and policy: A comparative analysis of national strategies[J]. Telecommunications Policy, 2022, Vol 46(8).
- [15] 国务院办公厅. 促进新一代人工智能产业发展三年行动计划 (2018-2020) [EB/OL] .2017.
- [16] 吴军. 智能时代: 5G、IoT 与 AI 重构未来经济[M]. 北京: 中信出版社, 2020.
- [17] 中国信通院. 中国人工智能创新生态发展报告 (2023) [EB/OL] .2023.
- [18] Lee K, Kim B Y. Industrial responses to AI policy shifts: Evidence from South Korea's national AI strategy[J]. Technological Forecasting and Social Change, 2021 Vol 173.
- [19] Brunderink R, Cohen B, Larosse J. The impact of EU AI regulation on industry innovation: A case study of GDPR and the AI Act[J]. Journal of Public Policy, 2022 Vol 42(3).

第7章 人工智能治理与伦理

人工智能技术的迭代速度超出人类预期, 现有治理系统的演进速度难以追上技术系统的发展步伐。在技术创新规律方面, 摩尔定律是预测计算机处理能力的经典标准, 但人工智能尤其是通用人工智能系统的进化和迭代速度远远超出这一定律。在此背景下, 如何克服社会、经济、法律等治理系统演进的高昂成本, 根据人工智能的发展情况及时完成治理体系的快速反应、协同共振, 是人工智能治理的重要挑战。与此同时, 人工智能技术的智能涌现还具有不可预测性, 治理系统很难准确预判人工智能的发展及其复杂的经济社会影响。一方面, 与传统技术变革相比, 以大模型为代表的人工智能初步具备了自我创造、学习、进化的特性, 这一技术潜力导致人工智能技术的发展和突破节点难以预测。另一方面, 未来通用人工智能、世界模型、实体智能等技术的发展可能引发社会经济结构整体性、颠覆性变革, 治理系统很难及时识别人工智能发展潜在的、难以预知的冲突和风险, 并提前针对相关风险做好准备。

本章将分别通过四个维度: 人机关系和伦理问题、数据和隐私保护、内容安全和虚假信息、知识产权和责任归属尝试着解读目前全世界正在积极讨论着的人工智能治理议题。

7.1 人机关系和伦理问题

7.1.1 人机关系的现状和发展趋势

一、AI 人机关系的现状

随着人工智能技术的飞速发展, 人机关系已经从简单的工具使用关系转变为更加复杂和多元的关系模式。目前, 人机关系主要体现在以下几种模式:

1、人机协作模式

在多个领域, AI 与人类已经形成了紧密的协作关系, 共同完成复杂的任务, 提升效率和质量。

(1) 医疗领域, AI 辅助诊断系统是人机协作的典型代表。通过分析大量的医学影像和病历数据, AI 能够快速识别潜在的病变特征, 为医生提供诊断建议。如一些 AI 系统可以检测出早期肺癌的微小结节, 其准确率甚至超过经验丰富的放射科医生。这种协作模式不仅提高了诊断效率, 还降低了误诊率, 使患者能够更早地接受治疗。此外, 在手术过程中, AI 辅助机器人可以根据医生的指令进行精确操作, 减少手术风险, 提高手术成功率。

(2) 教育领域, AI 个性化教学系统能够根据学生的学习进度和特点, 提供定制化的学习内容和教学方法。通过对学生的学习数据进行分析, AI 可以预测学生的学习难点, 提前调整教学计划。如一些在线教育平台利用 AI 技术为学生提供个性化的学习路径, 包括推荐适合的学习资料、安排学习时间等。这种人机协作模式不仅提高了学生的学习效果, 还减轻了教师的教学负担, 使教师能够将更多精力投入到教学创新和学生情感关怀中。

(3) 工业领域, 在制造业中, AI 与人类工人的协作也日益紧密。AI 驱动的机器人可以在生产线上完成高精度、高强度的工作任务, 而人类工人则负责监督和调整生产流程。如,

在汽车制造工厂中，AI 机器人可以完成车身焊接、零部件安装等工作，而工人则负责检查产品质量和处理异常情况。这种协作模式不仅提高了生产效率，还降低了生产成本，提升了产品质量。

2、人机竞争关系

尽管人机协作带来了诸多益处，但 AI 在某些领域对人类工作岗位的威胁也不容忽视。随着 AI 技术的不断进步，一些传统的工作岗位逐渐被自动化系统所取代。

(1) 制造业领域，自动化生产线的普及使得许多重复性劳动岗位消失。如传统的装配工人被机器人取代，机器人可以 24 小时不间断工作，且具有更高的精度和效率。据国际劳工组织（ILO）预测，到 2025 年，全球约有 8500 万工作岗位面临被自动化取代的风险，其中制造业是受影响最严重的行业之一。

(2) 服务业领域，智能客服系统的广泛应用也对传统客服岗位构成了威胁。这些系统能够自动回答常见问题，处理客户投诉，甚至进行简单的销售推荐。如许多银行和电信公司已经大量采用智能客服，减少了对人工客服的依赖。据麦肯锡全球研究所报告，到 2030 年，全球约有 14% 的工作岗位可能被自动化取代，服务业中的客服岗位是其中之一。

(3) 金融业领域，人工智能在智能投顾系统中的应用也对金融行业产生了影响。智能投顾系统可以根据客户的财务状况和投资偏好，提供个性化的投资建议，甚至自动进行投资操作。这使得传统的投资顾问岗位受到冲击。据高盛研究，到 2025 年，全球约有 1700 万金融从业人员可能面临被自动化取代的风险。

3、人机依赖关系

随着 AI 技术的广泛应用，人类对 AI 的依赖程度逐渐加深。这种依赖不仅体现在日常生活中，还体现在社会运行的各个层面。

(1) 智能家居系统通过 AI 技术实现了设备之间的互联互通，用户可以通过语音指令或手机应用程序控制家中的电器、照明、安防等设备。如智能音箱可以根据用户的语音指令播放音乐、查询天气、设置闹钟等；智能门锁可以通过人脸识别或指纹识别实现自动开锁。这种依赖关系使得人们的生活更加便捷，但也带来了隐私和安全方面的隐忧。

(2) 在交通领域，AI 技术的应用也使得人们对其产生了依赖。自动驾驶汽车、智能交通信号灯等系统通过 AI 算法优化交通流量，减少拥堵，提高出行效率。如自动驾驶汽车可以根据路况自动调整车速、选择最优路线，甚至在紧急情况下自动刹车。这种依赖关系使得人们在享受便利的同时，也对系统的可靠性提出了更高的要求。

二、AI 人机关系的发展趋势

随着技术的不断进步，人机关系也在不断发展和演变。未来人机关系将呈现出以下几种可能的趋势：

1、智能化与自主化

AI 系统将具备更强的自主学习和决策能力，能够在复杂环境中独立完成任务。这种智

能化和自主化的趋势将使 AI 系统更加独立于人类，人机关系也将变得更加复杂。

(1) 未来的 AI 系统将能够根据环境变化自主调整行为，自主决策能力提升。如在自动驾驶领域，AI 系统不仅能够识别交通信号和障碍物，还能够根据实时路况自主选择最优行驶路线，甚至在遇到突发情况时做出合理的应急决策。这种自主决策能力的提升将使 AI 系统在某些方面超越人类，但也引发了关于责任归属和伦理道德的讨论。

(2) AI 系统将具备更强的学习能力，能够通过自我学习不断优化性能。如通过强化学习，AI 可以在虚拟环境中进行大量模拟训练，快速掌握复杂任务的技能。这种学习能力的增强将使 AI 系统能够更好地适应不断变化的环境和任务需求，但也可能导致 AI 系统的行为超出人类的预期和控制范围。

2、多模态交互

AI 将支持语音、图像、文字等多种交互方式，使人机交互更加自然和高效。这种多模态交互方式将使 AI 系统能够更好地理解人类的需求和意图，提高人机协作的效率。

(1) 语音识别和自然语言处理技术的不断进步将使 AI 系统能够更好地理解人类语言，进行自然语言交互。如智能语音助手可以通过语音对话回答用户的问题、执行用户的指令，甚至进行情感交流。这种自然语言交互方式将使 AI 系统更加贴近人类的生活，提高人机交互的自然性和便捷性。

(2) AI 系统将能够通过图像识别和手势识别技术与人类进行交互。如，在虚拟现实和增强现实应用中，用户可以通过手势操控虚拟物体，AI 系统则可以根据用户的动作和表情做出相应的反馈。这种图像和手势交互方式将使 AI 系统能够更好地理解人类的非语言信息，提高人机交互的准确性和效率。

3、人机融合

AI 技术将与人类的生理和心理特征深度融合，人机关系将更加紧密。这种融合不仅体现在技术层面，还体现在社会和文化层面。

(1) 生理融合方面，脑机接口技术的发展将使 AI 系统能够直接与人类大脑进行交互。如通过植入式或非植入式脑机接口设备，AI 系统可以读取人类大脑的神经信号，实现对肢体运动的控制、对情感状态的监测等功能。这种生理融合将使 AI 系统能够更好地理解人类的生理需求，为残疾人提供更有效的康复辅助，甚至为人类拓展新的感知和运动能力。

(2) 心理融合方面，AI 系统将能够更好地理解人类的心理需求和情感状态。如通过情感分析技术，AI 系统可以识别用户的情绪变化，并根据情绪状态调整交互方式和内容。这种心理融合将使 AI 系统能够更好地满足人类的情感需求，提高人机交互的亲力和信任度。

三、AI 人机关系发展的影响与挑战

AI 人机关系的发展将对人类社会产生深远的影响，同时也带来了诸多挑战。

(1) AI 技术的应用需要大量的数据支持，这可能导致个人隐私泄露和数据安全问题。如，智能家居设备和智能交通系统收集了大量的用户数据，这些数据如果被泄露或滥用，将

对用户的隐私和安全造成严重后果。该问题也将在 7.2 节中作具体探讨。

(2) 随着 AI 系统智能化和自主化程度的提高, 人机关系将变得更加复杂, 伦理道德困境也将更加突出。如当 AI 系统在自动驾驶汽车中面临无法避免的事故时, 应该如何做出决策? 这种决策是否符合人类的道德价值观? 这些问题需要我们在技术发展的过程中不断进行伦理思考和规范。

7.1.2 AI 对生命伦理的影响和挑战

一、AI 对生命伦理的影响

生命伦理学 (Bioethics) 是一门跨学科的伦理学分支, 主要研究生命科学和医疗实践中涉及的伦理问题。它关注人类生命从诞生到死亡的全过程, 涉及生命权、自主权、隐私权、公正性等多个方面。生命伦理学的核心目标是确保在科学和技术发展的过程中, 人类的生命和尊严得到尊重和保护。

1、生命权 (Right to Life)。生命权是生命伦理学中最基本的概念之一, 指的是每个个体都有权利享有生命, 且生命不应被无端剥夺。在医疗实践中, 生命权涉及对生命的延续、终止以及生命质量的评估。如对于临终患者的治疗决策, 如何在延长生命和减轻痛苦之间找到平衡, 是生命伦理学的重要议题。

2、自主权 (Autonomy)。自主权强调个体在医疗决策中的自主性和选择权。患者有权根据自己的意愿和价值观做出关于自身健康的决定, 包括接受或拒绝治疗。在 AI 应用于医疗决策时, 如何确保患者的自主权不被技术所侵犯, 是一个重要的伦理问题。

3、隐私权 (Privacy)。隐私权涉及个人数据的保护, 特别是在医疗领域, 患者的病历、基因信息等高度敏感数据需要严格保密。AI 技术在处理和分析这些数据时, 必须确保数据的安全性和隐私性, 防止数据泄露和滥用。

4、公正性 (Justice)。公正性要求医疗资源的分配和使用应公平合理, 不应因种族、性别、经济状况等因素而产生歧视。AI 在医疗资源分配中的应用, 如智能诊断系统和医疗资源管理系统, 需要确保其决策过程符合公正性原则。

二、AI 应用与伦理挑战

随着 AI 技术的快速发展, 其在医疗领域的应用越来越广泛, 同时也带来了诸多伦理挑战。

1、医疗诊断与治疗方面, AI 辅助诊断系统通过分析大量的医学影像和病历数据, 能够快速识别潜在的病变特征, 提高诊断的准确性和效率。然而, 这种技术也可能引发一系列伦理问题: 比如误诊风险, 尽管 AI 系统的诊断准确率较高, 但仍然存在误诊的可能性; 当 AI 系统给出错误的诊断建议时, 可能导致患者接受不必要的治疗或延误病情; 如何确保 AI 系统的可靠性, 并在误诊发生时明确责任归属, 是一个亟待解决的问题。另外责任归属也是个问题: 当 AI 系统参与医疗决策时, 责任归属变得模糊。如果 AI 系统给出的诊断建议导致了不良后果, 是开发者、使用者还是 AI 系统本身应承担 responsibility? 这种责任界定的困难可能导

致法律纠纷和伦理困境。

2、基因编辑方面，AI 技术可以辅助基因编辑，帮助科学家更精准地进行基因操作。然而，基因编辑可能引发对人类基因库的永久性改变，甚至涉及“设计婴儿”等伦理争议。如何在推动科技进步的同时，确保基因编辑符合人类的伦理价值观，是一个重要的伦理挑战。比如 AI 驱动的生命支持系统可以实时监测患者的生命体征，并自动调整治疗方案。然而，这种系统可能引发关于生命延续的伦理问题，如在何种情况下应终止生命支持，以及如何确保患者的自主权得到尊重。

三、AI 对人类生命观的影响

AI 技术的发展不仅改变了人类的实践，还对人类的生命观产生了深远影响。

1、人类中心主义的挑战

人类中心主义（Anthropocentrism）强调人类在自然界中的中心地位，认为人类的利益高于其他生命形式。然而，AI 的发展可能挑战这一传统观念：

(1) 机器的“生命”属性。随着 AI 技术的不断进步，机器的自主性和智能水平逐渐提高，引发了关于机器是否具有“生命”属性的讨论。如，一些高度智能的机器人能够自主学习和决策，甚至表现出类似人类的情感和行为。这使得人类不得不重新思考生命的意义和边界。

(2) 人类地位的重新评估。AI 技术的发展使得机器在某些领域超越了人类的能力，这可能引发对人类中心主义的反思。如，在自动驾驶汽车中，AI 系统能够比人类更安全地驾驶车辆，这是否意味着人类在某些方面不再是“最优”的存在？这种反思可能促使人类重新评估自身在自然界和社会中的地位。

2、生命定义的拓展

AI 技术的发展促使人类重新定义生命的边界：

(1) 机器的生命权。随着 AI 系统的智能化程度不断提高，一些学者开始探讨机器是否可以拥有生命权。如，如果一个 AI 系统具备了高度的自主性和情感，是否应该赋予其类似人类的生命权？这种讨论不仅涉及技术层面，还涉及哲学和伦理学的深层次问题。

(2) 人类生命的延伸。AI 技术在医疗和基因编辑领域的应用，使得人类生命的延续成为可能。如，通过基因编辑技术，人类可以预防某些遗传疾病，甚至延长寿命。这是否意味着人类生命的定义将从自然寿命向“设计寿命”转变？这种转变可能引发关于人类生命本质的重新思考。

四、AI 人机关系中的伦理挑战

AI 技术的发展不仅对生命伦理产生了深远影响，还在人机关系中引发了诸多伦理挑战和伦理冲突。

1、AI 决策与人类价值观的冲突

AI 系统在决策过程中可能与人类的道德价值观不符，引发伦理困境：

(1) 算法决策的伦理问题。AI 系统通常基于数据和算法进行决策，但这种决策可能与

人类的道德直觉相悖。刚刚提到的自动驾驶汽车面临不可避免的事故时，AI 系统如何在保护乘客和行人之间做出选择，这种决策自然会长期引发伦理争议。

(2) 价值观的多样性。不同文化和社会背景下的价值观存在差异，AI 系统如何在不同价值观之间找到平衡，是一个重要的伦理问题，甚至会引发政治问题和纷争。一些文化可能更重视集体利益，而另一些文化可能更强调个人权利。AI 系统在跨文化应用中需要考虑这种多样性，以避免伦理冲突。

2、AI 与人类责任的界定

当 AI 系统出现错误或误判时，责任归属难以明确：

(1) 开发者与使用者的责任。AI 系统的开发者和使用者在责任界定上存在模糊地带。如果一个 AI 系统在医疗诊断中出现误诊，是开发者的设计缺陷还是使用者的操作不当导致的？这种责任界定的困难可能导致法律纠纷和伦理困境。

(2) AI 系统的“道德责任”。随着 AI 系统的智能化程度不断提高，一些学者开始探讨 AI 系统是否可以承担道德责任。如果一个 AI 系统在自动驾驶过程中导致事故，是否可以将其视为一个“道德主体”来追究责任？这种讨论不仅涉及技术层面，还涉及哲学和伦理学的深层次问题。

3、歧视与偏见

(1) 数据偏见。AI 系统依赖的数据可能存在偏见，导致对某些群体的不公平待遇。AI 系统的训练数据通常是通过现有的社会数据集收集的，这些数据可能包含社会偏见。在招聘领域，如果 AI 系统的训练数据中存在性别或种族偏见，那么它在筛选简历时可能会对某些群体产生不公平待遇。此外，为避免数据偏见，需要对数据进行严格的清洗和校正。然而，这一过程不仅技术复杂，还可能引发新的伦理问题。如，如何在数据清洗过程中确保不丢失重要信息，同时又不引入新的偏见？

(2) 算法偏见。AI 算法的设计和开发过程中也可能引入人为偏见（或许是强化学习导致的）。算法的设计者在设计过程中可能无意中引入了自己的偏见。如，在开发信贷评估系统时，设计者可能基于自己的经验或观念，对某些群体的信用评分进行调整，从而导致不公平的结果。为避免算法偏见，需要提高算法的透明性和可解释性。然而，AI 系统的复杂性使得其决策过程难以被人类理解。如何在技术层面提高算法的透明性，同时又不牺牲其性能，是一个重要的研究方向。

7.2 AI 数据和隐私保护

7.2.1 AI 数据使用必将涉及隐私侵犯争议

AI 数据的隐私保护是指在人工智能技术涉及的数据收集、存储、使用、共享等全生命周期中，采取一系列措施确保个人隐私、敏感信息不被泄露、滥用，保障数据主体的合法权益，维护数据的安全性、完整性和保密性。一般在 AI 数据处理过程中，可以通过加密、匿名化等技术手段保护数据不被非法获取，以及建立严格的访问控制机制，限制数据访问权限，

防止数据滥用，但仍然存在以下几种主要的潜在风险：

1、数据收集环节的风险

在 AI 数据收集过程中，一直存在“过度”收集数据，以及收集来源合法性存疑的质疑。在全球范围内，一些 AI 应用在收集用户数据时，超出合理范围，收集大量不必要的个人信息。某些 APP 在安装时，要求获取过多与核心功能无关的权限，如拍照、录音、位置信息等。成因主要是企业为追求数据的丰富性，以提升 AI 模型的性能，同时缺乏明确的边界界定和有效监管。个人信息保护规则执行困难，企业难以通过正常合法途径获取数据，在数据采集环节，因知情同意授权规则难以实施，且缺乏其他合法性基础，导致个人隐私面临被侵犯风险。同时，匿名化处理的具体要求和标准不明确，使得开发企业难以把握采集、利用用户数据的合法边界，进一步增加了个人隐私泄露的可能性。

2、数据存储环节的风险

数据存储也面临着数据泄露和被篡改的风险。数据存储系统一旦遭受黑客攻击，就可能导致大量 AI 数据泄露。2024 年某知名 AI 公司数据库遭黑客入侵，大量用户数据被窃取。这是由于存储系统的安全防护措施存在漏洞，以及对数据访问权限管理不当导致。

3、数据滥用的风险

人工智能可以被用于制作虚假文本、音频、视频等深度伪造内容，进而实施政治干预、煽动暴力和犯罪等破坏公共利益，甚至侵害国家安全。在这个过程中，数据被滥用，不仅侵犯了公众的知情权，也可能对个人隐私造成潜在威胁，因为虚假信息的传播可能会误导公众对个人的认知，从而影响个人的声誉和隐私。

7.2.2 AI 数据的隐私侵犯问题成因分析

AI 数据隐私侵犯问题成因复杂，涉及技术、制度、产业发展等多方面。技术层面的漏洞与局限为侵犯行为提供了可乘之机，制度层面的不完善使得监管与约束乏力，产业发展过程中的逐利冲动与不当竞争加剧了问题的严重性。这些因素相互交织，共同导致了 AI 数据隐私侵犯问题的产生。

1、技术层面的缺陷

一是模型安全漏洞问题，AI 算法模型设计、训练和验证环节存在诸多脆弱点，如标准接口、特性库和工具包可能有逻辑缺陷与漏洞，还易被恶意植入后门。这些漏洞一旦被利用，攻击者就能获取、篡改或滥用 AI 数据，严重侵犯隐私。像一些不法分子可能利用模型漏洞，非法获取用户的敏感信息，如医疗记录、金融数据等。

二是生成“幻觉”问题，AI 生成“幻觉”会导致内容不可信，使 AI 在处理和输出数据时出现错误或虚假信息，干扰对数据真实性和可靠性的判断，在涉及隐私数据处理时，可能造成隐私泄露或被错误利用。这部分将在 7.3 节还会有具体阐述。

2、制度规范的缺失

由于数据获取规则模糊以及监管体系不完善，在个人信息保护领域，知情同意授权规则

难以落地，企业难以找到合法获取数据的途径，且匿名化处理的具体要求和标准不明确，企业难以把握采集和利用用户数据的边界，容易引发隐私侵犯风险。在知识产权规则方面，相关诉讼纠纷和监管处罚不断，也使得企业在数据获取和使用上存在不确定性，增加了隐私侵犯的可能性。

传统政府监管模式难以适应 AI 技术发展，事前风险审查缺乏统一标准，不同企业依据不同基准测试模型，导致难以比较头部模型风险和局限性；事中介入机制缺乏实践经验，对 AI 的事中监管处于初期阶段；事后调查溯源取证困难，监管部门在处理版权纠纷、内容安全等问题时缺乏有效核查核验工具和成熟的技术溯源措施。

3、产业发展的无序性

商业利益会驱使不当行为，部分企业为追求商业利益，在数据获取和使用上不择手段，过度收集用户数据，甚至将数据用于未经授权的目的，过去几年我国“315”晚会已有屡次曝光。在 AI 训练数据的获取上，一些企业可能未经授权使用版权材料进行模型训练，既侵犯了知识产权，也可能导致用户隐私数据被不当使用。

值得一提的是，开源模型其实也存在潜在风险，开源模型虽能促进创新，但也存在诸多问题，前沿人工智能开源可能因滥用和失控引发灾难风险，如被用于恶意目的，从而导致数据隐私侵犯。而知情同意授权规则难以定义与实施个人信息保护范围，同时难以援引订立合同所必须、履行法定职责等其他合法性基础。匿名化处理的具体要求和标准尚不明确，导致开发企业难以把握采集、利用用户数据的合法边界。

7.2.3 AI 数据的隐私保护对策分析

1、制度建设层面

(1) 首先需要完善数据隐私法规体系，制定专门针对 AI 数据隐私保护的法规，明确 AI 数据收集、存储、使用、共享等环节的具体规则和标准。参考欧盟《人工智能法》，对数据来源合法性、数据使用目的限制、数据存储安全要求等进行详细规定。对我国来说，可以整合现有的数据隐私相关法规，如《网络安全法》《数据安全法》《个人信息保护法》等，使其在 AI 数据隐私保护方面形成协同效应，避免法规之间的冲突和矛盾，并且加强执法力度。

(2) 其次应加强数据隐私监管措施，建立或增强专门的 AI 数据隐私监管机构职权，对 AI 数据处理活动进行全面监管。该机构应具备技术专业能力，能够对 AI 算法、数据处理流程等进行深入审查，确保其符合隐私保护要求。对 AI 企业的数据处理活动进行定期检查 and 随机抽查，及时发现和纠正潜在的隐私问题。此外，还需要加强对数据跨境流动的监管，制定数据跨境传输的规则和标准，确保数据在跨境过程中的隐私安全；要求企业在进行数据跨境传输前，进行严格的风险评估，并获得相关监管部门的批准。

(3) 最后要明确责任界定与救济途径，明确 AI 数据处理过程中各方的责任，包括数据收集者、使用者、存储者等的责任。当发生数据隐私侵权事件时，能够快速确定责任主体，

避免出现责任推诿的情况。参考欧盟《人工智能责任指令》的做法，减轻受害者的举证责任，使受害者更容易获得赔偿；设立专门的数据隐私侵权赔偿基金，当受害者无法从责任主体处获得赔偿时，可从基金中获得一定的补偿。同时，加强对数据隐私侵权行为的处罚力度，提高违法成本，对故意侵犯数据隐私的行为进行严厉打击，形成有效的法律威慑。

2、技术手段层面

在技术手段层面解决 AI 数据的隐私保护问题，可从数据加密、风险评估、内容标识、模型安全强化以及技术协同创新等方面着手。通过这些技术手段的综合运用，能够构建起多层次、全方位的 AI 数据隐私保护体系，有效降低数据隐私泄露风险。

(1) 数据加密与匿名化技术可以运用相对先进的加密算法对 AI 数据进行加密处理，确保数据在传输和存储过程中的安全性。在数据传输阶段，则采用 SSL/TLS 等加密协议，防止数据被窃取或篡改；在数据存储时，使用 AES 等加密算法对数据进行加密存储，即使数据泄露，攻击者也难以获取有效信息。

(2) 安全评估与监测技术可以建立全面的 AI 数据安全评估体系，对 AI 系统的数据处理活动进行定期评估。在模型训练前，对训练数据进行安全性评估，检查数据是否存在隐私泄露风险；在模型部署后，对模型的运行情况进行监测，及时发现潜在的安全漏洞。利用人工智能技术自身的优势，开发智能监测工具，实时监测 AI 数据的使用情况。通过分析数据的访问模式、使用频率等信息，及时发现异常行为，如未经授权的数据访问、数据滥用等，并及时发出警报。

(3) 内容标识与溯源技术可以对 AI 生成的内容进行标识，明确内容的来源和生成方式。通过数字水印、区块链等技术，在 AI 生成的文本、图片、音频、视频等内容中嵌入特定标识，方便用户识别内容的真实性与来源，也有助于在出现隐私问题时进行溯源。当发生数据隐私泄露事件时，可以快速准确地追踪数据的流向和使用情况，确定责任主体。

(4) 模型安全与防护技术可以加强 AI 模型的安全防护，防止模型被攻击和滥用。对模型进行漏洞检测和修复，定期更新模型的安全补丁，提高模型的安全性。采用模型加密、访问控制等技术，限制对模型的访问权限，防止模型被恶意利用。在模型训练过程中，采用联邦学习、差分隐私等技术，在保证数据可用性的同时，保护数据的隐私性。

(5) 最后是技术标准与规范制定，制定统一的 AI 数据隐私保护技术标准和规范，促进不同技术之间的兼容性和互操作性。如统一数据加密算法的标准、内容标识的格式和规范等，便于企业和开发者遵循，提高 AI 数据隐私保护的整体水平。

3、企业自律与社会监督层面

AI 数据隐私保护中，企业自律与社会监督是关键防线。企业通过完善内部管理、参与行业自律和技术创新保护隐私；社会借助强化监督意识、建立监督机制和推动公众参与，形成全方位监督网络。两者协同，能有效遏制隐私侵犯，保障数据安全。

在企业自律层面，企业应建立健全的数据隐私管理制度，明确数据收集、存储、使用、

共享等各个环节的操作规范和责任分工。制定详细的数据收集政策，明确收集目的、范围和方式，确保数据收集的合法性和必要性。加强对员工的数据隐私保护培训，提高员工的隐私意识和合规操作能力，防止因员工疏忽或违规操作导致数据泄露。同时，定期组织员工参加数据隐私保护培训课程，学习相关法律法规和企业内部制度，使员工深刻认识到数据隐私保护的重要性。积极参与行业协会和组织，共同制定和遵守行业数据隐私保护规范和标准。通过行业自律，促使企业自觉遵守相关规定，提高整个行业的数据隐私保护水平。配合行业组织开展数据隐私保护的监督和评估工作，对自身的数据处理活动进行自查自纠，及时发现和整改存在的问题。

在社会监督层面，媒体也应发挥监督作用，通过报道和曝光数据隐私侵权事件，引起社会各界的关注，类似国内的“315”晚会，监督企业重视数据隐私保护。加强对数据隐私保护相关法律法规和政策的宣传，提高公众的监督意识和能力，使公众能够更好地识别和监督企业的数据隐私侵权行为。此外，有必要建立健全的数据隐私保护社会监督机制，搭建专门的监督平台，方便公众对企业的数据隐私侵权行为进行举报和投诉。相关部门应及时受理公众的举报和投诉，对违规企业进行调查和处理，并将处理结果反馈给公众，如，设立专门的数据隐私保护举报热线和在线举报平台，建立举报奖励制度，鼓励公众积极参与监督，培养公众的数据隐私保护意识。

4、国际组织层面

在国际合作上深化与建议统一标准也势不容缓。人工智能的全球性发展使得数据隐私保护的国际合作愈发重要。各国有必要加强在数据隐私保护领域的沟通与协作，共同应对跨境数据流动带来的挑战。在此过程中，国际组织在推动全球数据隐私保护统一标准制定方面需要发挥更大作用，促进各国标准的互认与衔接。通过多边和双边合作机制，各国将分享隐私保护的经验和技術，加强执法合作，打击跨国的数据隐私侵权行为，为人工智能的全球化发展营造良好的环境。

如，联合国可致力于实现全球人工智能治理的协调和统合，酝酿协调机制。联合国大会通过相关决议等形式，推进人工智能系统伦理、人权保护、数据安全等讨论，支持可持续发展目标。其他国际组织如二十国集团、经济合作与发展组织、东盟、金砖国家等也在各自层面推动人工智能治理，兼顾发展与安全。

7.2.4 AI 数据的隐私保护未来展望

AI 数据隐私保护在人工智能蓬勃发展的当下至关重要，它不仅紧密关联个人隐私、人格尊严与基本权利，还深刻影响着社会稳定和人工智能产业的可持续发展。当前，AI 数据隐私保护面临诸多难题，数据收集时存在过度收集和来源合法性问题，存储环节有数据泄露与篡改风险，使用过程中还出现数据滥用和算法偏见现象。为应对这些挑战，法规政策层面，我国已构建起一套涵盖上位法和专项法规的监管体系，明确数据处理规则与责任；技术手段上，评估评测、内容标识、价值对齐、红队测试等技术不断发展应用，保障数据安全；企业

自律与社会监督也积极发挥作用，企业通过自愿承诺参与治理，社会监督促使企业重视隐私保护，公众智能素养和隐私保护意识逐步提升。展望未来，随着技术的持续进步，AI 数据隐私保护需在完善法律法规、加强技术创新、深化国际合作以及强化企业自律和社会监督等方面持续发力，从而构建更严密的保护体系，推动人工智能技术健康、安全发展。

7.3 AI 内容安全和虚假信息

7.3.1 生成式 AI 背景下虚假信息传播的新特征

生成式人工智能（Generative AI，国内也称 AIGC）的崛起，尤其是以 ChatGPT 和 Midjourney 等为代表的大型语言模型（LLM）和图像生成模型的广泛应用，为信息传播带来了前所未有的机遇，同时也为虚假信息的制造和传播提供了新的温床。与传统虚假信息相比，生成式 AI 背景下的虚假信息呈现出以下几个显著特征：

1、生成内容的高度逼真性

深度伪造技术利用生成对抗网络（GAN）、扩散模型（Diffusion Models）和 Transformer 架构等深度学习算法，能够合成高度逼真的图像、音频和视频。这些伪造内容足以以假乱真，普通用户难以辨别，甚至专业的鉴伪技术也面临巨大挑战。如，2023 年利用 Midjourney 生成的“教皇穿着时尚羽绒服”和“特朗普被捕”的虚假图像在社交媒体上迅速传播，引发广泛关注和混淆。

2、制造与传播的低门槛化

生成式 AI 工具的易用性和可及性，大大降低了虚假信息制造和传播的门槛。普通用户无需专业知识或高昂成本，即可利用这些工具生成以假乱真的内容，并通过社交媒体等渠道迅速传播。2023 年，仅需简单提示词，用户就能利用 ChatGPT 生成看似权威的虚假学术论文或新闻报道，这种“提示工程”（Prompt）的简易性使得虚假内容生产呈现爆炸式增长。

3、信息传播的病毒式扩散

在社交媒体时代，信息传播具有病毒式扩散的特点，虚假信息一旦发布，可能在短时间内触达海量用户。认知偏差、信息茧房等因素，进一步加剧了虚假信息的传播速度和广度。最新研究表明，AI 生成的虚假信息比人工创作的虚假信息传播速度更快、范围更广，因其往往更符合用户的认知偏好和情感需求。

4、“谎言红利”效应的强化

随着公众对深度伪造技术的认知加深，人们对音视频等证据的信任度降低，这为谎言者提供了逃避责任的借口。即使面对真实的证据，他们也可以辩称其为“深度伪造”，从而混淆视听、逃避追责。2022 年俄乌冲突中，双方均曾指责对方发布的战争罪证据为“AI 生成”，这种“真实性崩塌”现象使得公众难以辨别真相。

5、多模态融合与跨平台协同

生成式 AI 能够实现文本、图像、音频和视频的多模态融合，创造出更具说服力的虚假内容。同时，虚假信息通常在多个平台协同传播，形成“信息流行病”（infodemic）。如，2023

年一则 AI 生成的虚假新闻首先在推特上以文本形式出现，随后在 TikTok 上配以 AI 生成的视频，最终在脸书等网络媒体上大规模传播，形成完整的虚假信息生态系统。

这些新特征使得虚假信息治理面临前所未有的挑战，传统的治理方法和工具难以有效应对。

7.3.2 生成式 AI 虚假信息的多维度危害

生成式 AI 虚假信息不仅威胁个人权益和企业声誉，更对社会稳定、民主政治和国家安全构成多维度的危害：

1、个体层面

深度伪造技术可被用于制作虚假色情视频、实施网络欺凌、诽谤和敲诈勒索等，对受害者造成严重的心理创伤、名誉损害和经济损失。如，印度记者 Rana Ayyub 因深度伪造色情视频而遭受网络暴力和死亡威胁。2023 年，一项针对 12 个国家的调查显示，超过 60% 的女性公众人物曾遭遇 AI 生成的虚假色情内容攻击，造成严重的心理伤害和职业损失。

2、企业层面

虚假信息可能损害企业声誉，造成经济损失。2022 年，一家美国制药公司因 AI 生成的虚假研究报告在社交媒体传播，导致股价在一天内下跌 23%，市值蒸发超过 10 亿美元。此外，AI 生成的虚假客户评价和产品评论也严重扰乱了电子商务生态系统，使消费者难以做出明智选择。

3、社会层面

虚假信息可能引发社会恐慌、加剧社会分裂、破坏社会信任。2023 年，AI 生成的关于银行倒闭的虚假新闻在社交媒体传播，导致多家银行遭遇挤兑风波，引发金融市场动荡。认知心理学研究表明，即使在被告知信息是 AI 生成的虚假内容后，人们仍会受到持续影响效应的影响，难以完全摆脱最初形成的错误认知。

4、国家安全层面

深度伪造技术已经在西方被验证可能被用于制造虚假情报、误导战略决策、挑拨国际关系、煽动恐怖主义等，对国家安全构成威胁。2022 年，俄乌冲突期间出现了 AI 生成的乌克兰总统泽连斯基投降演讲视频，虽然很快被识破，但展示了此类技术对国际危机的潜在影响。美国国家安全委员会 2023 年的报告指出，生成式 AI 已成为国家间信息战的新型武器，对全球地缘政治稳定构成挑战。

5、认知安全层面

长期暴露于 AI 生成的虚假信息环境中，可能使公众丧失对信息真实性的判断能力，产生普遍的怀疑主义和信任危机。这种认知安全威胁可能比单一的虚假信息事件造成更深远的社会影响。

7.3.3 虚假信息治理的困境与挑战

传统的虚假信息治理方法主要包括事实核查、平台内容审核、法律法规约束等。然而，

在生成式 AI 时代，这些方法面临诸多困境与挑战：

1、技术挑战

深度伪造技术的快速发展使得虚假内容的检测和识别变得越来越困难。尽管一些研究致力于开发检测技术，但其准确率和可靠性仍有待提高，且往往滞后于伪造技术的发展。最新研究表明，当前最先进的深度伪造检测技术在实验室环境中的准确率可达 95%，但在真实应用场景中准确率降至 70% 左右，且对新型生成模型（如扩散模型）生成的内容检测效果较差。

2、规模挑战

社交媒体平台上的信息传播速度快、范围广，人工审核难以应对海量信息。脸书公司 2023 年报告显示，其平台每天需处理超过 1 亿条潜在虚假信息，而 AI 生成内容的爆炸式增长使这一数字预计在 2024 年增长 3-5 倍。自动化审核工具虽然效率高，但容易出现误判，且难以处理复杂的语境和文化差异，导致大量虚假信息漏检或误伤合法内容。

3、法律挑战

现有的法律法规难以完全覆盖深度伪造等新型虚假信息的产生和传播。各国法律体系差异巨大，跨境执法困难，导致虚假信息可以轻易规避。

4、伦理挑战

虚假信息治理涉及言论自由、隐私保护、算法偏见等复杂的伦理问题。如何在打击虚假信息的同时保护言论自由、避免算法歧视，是一个需要平衡的难题。研究表明，过度严格的内容审查可能导致“寒蝉效应”，抑制合法言论表达。同时，AI 内容审核系统本身可能存在偏见，对特定群体或观点的表达造成不公平限制。

5、认知挑战

人类认知偏差使我们容易相信符合已有信念的信息（信息“茧房”），且更容易记住虚假信息而非后续的更正。认知心理学研究表明，即使是明确标记为 AI 生成的内容，仍能影响人们的信念和决策，这种“知其假而信其真”的现象使得单纯依靠标记和教育的方法效果有限。

6、商业模式挑战

社交媒体平台的商业模式依赖于用户注意力和参与度，而情绪化、争议性的内容（包括虚假信息）往往能获得更高的参与度。这种结构性矛盾使平台在打击虚假信息方面存在内在动力不足的问题。

7.3.4 生成式 AI 背景下虚假信息治理的对策

为应对生成式 AI 带来的虚假信息挑战，建议从技术、法律、政策、教育等多方面入手，构建综合治理体系：

1、技术层面

(1) 加强虚假信息检测技术研发。投入资源开发更先进的深度伪造检测技术，提高检测的准确率、效率和自动化程度。最新研究方向包括多模态融合检测（结合视觉、音频和元

数据分析)、生物特征验证(如眨眼模式、脉搏信号分析)和神经指纹技术(分析 AI 模型留下的独特痕迹)。如, Intel 和微软合作开发的 FakeCatcher 技术通过分析视频中人物面部的血液流动模式(光密度变化),能以 96% 的准确率检测深度伪造视频。

(2) 建立内容溯源和认证机制。探索利用区块链等技术建立内容溯源机制,对内容进行数字签名和认证,提高内容的可信度。

(3) 发展负责任的 AI 技术。推动 AI 开发者采用负责任的 AI 原则,将伦理、公平、透明等价值观融入 AI 系统的设计、开发和部署过程。如, OpenAI 在 GPT-4 中实施了"防护栏"(guardrails)技术,限制模型生成有害内容的能力; Google 的 SynthID 技术在 AI 生成内容中嵌入不可见的水印,便于后续识别。

(4) 开发真实性增强(AR)技术。除了检测虚假内容,还应开发能主动增强真实内容可信度的技术。如,基于区块链的"真实认证"系统可以为重要新闻和公共信息提供不可篡改的验证机制;添加"数字水印"可以在内容创建阶段嵌入难以移除的标识。

2、法律与政策层面

(1) 完善相关法律法规。针对深度伪造技术等新型虚假信息的特点,完善相关法律法规,明确内容生产者、传播者和平台的责任。中国 2022 年出台的《深度合成管理规定》要求深度合成服务提供者建立用户注册、内容审核和安全评估机制,并明确标识合成内容。欧盟《数字服务法案》(Digital Services Act)则引入了"注意义务"(duty of care)概念,要求大型平台评估和减轻系统性风险,包括虚假信息传播风险。

(2) 建立 AI 内容标识制度。要求 AI 生成内容明确标识其来源和性质,增强透明度。欧盟《人工智能法案》(AI Act)要求生成式 AI 系统的提供者确保其生成的内容能够被识别为人工生成。中国《生成式人工智能服务管理暂行办法》也要求服务提供者采取有效措施标识生成内容,防止公众混淆或误解。

(3) 加强国际合作。虚假信息传播具有跨国性,需要加强国际合作,共同制定应对策略,分享技术和经验,打击跨境虚假信息传播。G7 国家 2023 年发布《广岛 AI 进程》首次将 AI 生成虚假信息作为全球治理重点,提出建立国际协调机制应对跨境虚假信息传播。联合国互联网治理论坛(IGF)也设立了专门工作组,协调各国在 AI 内容安全领域的政策和实践。

3、社会层面

(1) 提升公众的媒介素养和批判性思维能力。通过教育、宣传等方式,提高公众对深度伪造等新型虚假信息的认知水平,培养公众的批判性思维能力。认知心理学研究表明,提前向人们介绍虚假信息的常见手法,预防比事后更正更有效。

(2) 加强事实核查和辟谣机制建设。建立和完善事实核查机制,鼓励专业机构和媒体对虚假信息及时、准确的辟谣。国际事实核查网络(IFCN)已开始利用 AI 辅助工具提高事实核查效率,同时保持人工审核的准确性。中国的"辟谣平台"联动政府部门、媒体和专

业机构，形成协同辟谣网络，提高辟谣效率和权威性。

(3) 倡导负责任的信息分享行为。引导公众在社交媒体等平台上负责任地分享信息，不传播未经证实的消息。研究表明，在用户分享内容前提醒其考虑信息准确性可以显著减少虚假信息的传播。中国台湾有个“谣言猎人”项目鼓励公众主动举报可疑信息，并获得相应积分奖励，形成全民参与的辟谣机制。

(4) 构建多方参与的治理体系。虚假信息治理需要政府、企业、社会组织、媒体和公众等多方参与，形成合力。新加坡的治理模式整合了立法（《防止网络假信息和网络操纵法》）、执法（POFMA 办公室）、教育（媒体素养委员会）和社会参与（事实核查联盟），形成了综合治理体系。

4、中国对策

中国采取了政府主导、多方参与的治理路径，强调技术创新与法律规制并重：

(1) 法律规制。2022 年出台的《互联网信息服务深度合成管理规定》是全球首个专门针对深度合成技术的综合性法规，要求服务提供者建立健全算法机制和数据标签，明确标识合成内容，并对违规行为设定了明确的法律责任。2023 年《生成式人工智能服务管理暂行办法》进一步完善了对生成式 AI 的监管框架。

(2) 技术研发。中国科技企业和研究机构积极开发深度伪造检测技术。如，腾讯安全实验室开发的“天鉴”系统能够检测多种类型的深度伪造内容；中国科学院自动化研究所开发的“真假识别”平台利用多模态分析技术提高检测准确率。

(3) 平台责任。中国要求互联网平台承担更多内容审核责任。微博、抖音等平台建立了 AI 内容识别系统，对可疑内容进行标记或过滤。2023 年，中国互联网协会组织主要平台签署《生成式人工智能服务自律公约》，承诺加强 AI 生成内容管理。

(4) 协同治理。中国还建立了政府部门、企业、媒体和社会组织协同参与的治理机制。“网络辟谣联动机制”整合了官方媒体、专业机构和社交平台，形成快速响应和辟谣的能力。

7.4 AI 知识产权和责任归属

7.4.1 AI 技术发展引发知识产权法律挑战

AI 技术的飞速发展正以前所未有的态势重塑着社会的各个领域，在带来巨大创新机遇的同时，也对传统的知识产权法律体系发起了全面且深刻的挑战。从专利、著作权到商标等多个维度，AI 技术的独特性使得现有的法律规范在适用时面临诸多困境。

1、专利领域

AI 技术的可专利性边界模糊不清。传统专利法要求发明具备新颖性、创造性和实用性，然而，AI 算法、模型及系统组件在判断是否满足这些条件时存在诸多困难。如，AI 算法往往基于复杂的数学模型和数据处理逻辑，其创造性的判断充满难题。一方面，算法本身的抽象性使得其与传统专利法中对具体技术方案的要求难以契合，在判断是否具有创造性时，难以确定人类的创新贡献与 AI 自动化处理之间的界限；另一方面，诸多 AI 算法是在大量数

据训练基础上产生的，这种基于数据驱动的创新方式，使得其新颖性的判断也变得复杂。一些司法管辖区认为纯算法或数学方法缺乏可专利性，但当 AI 算法与具体应用场景相结合时，又难以明确区分算法本身和应用的界限，导致专利审查的不确定性增加。

2、著作权领域

AI 生成内容的著作权归属成为焦点问题。与人类传统创作过程不同，AI 生成内容的创作过程缺乏人类直接的创造性投入。以生成式 AI 为例，生成式 AI 可以根据输入的指令生成文章、绘画、音乐等内容，这些内容在形式上具备作品的特征，但却难以确定其创作主体。若将著作权赋予 AI 开发者，可能无法充分体现使用者在创作过程中的作用；若归属于使用者，又可能无法激励开发者对 AI 技术的研发投入。此外，AI 生成内容的原创性认定也存在争议。传统著作权法强调作品的“独创性”，即作品必须是作者独立创作完成且具有一定的创造性。而 AI 生成内容的创作过程是基于算法和数据的运算，其是否满足“独创性”要求在不同国家和地区存在不同的理解。部分国家认为，若 AI 生成内容没有人类的实质性参与，就不应被视为著作权法意义上的作品；而另一些国家则在探索给予一定程度的保护，但具体标准尚未统一。

3、商标领域

AI 生成的品牌名称或标识在申请商标时，面临着显著性和可识别性的审查难题。由于 AI 生成的商标往往是基于数据和算法产生的，可能缺乏传统商标所具有的独特创意和文化内涵，这使得在判断其是否能够区分商品或服务来源时存在困难。同时，AI 在商标设计过程中，可能会无意识地侵犯他人的在先权利，如使用了与他人商标相似的元素，这进一步增加了商标审查和管理的复杂性。

4、数据使用过程中的知识产权侵权风险

AI 系统的训练需要大量的数据，其中可能包含受版权保护的内容。当 AI 未经授权使用这些数据训练时，可能侵犯数据所有者的版权。而且，由于 AI 训练数据的规模庞大且来源复杂，难以逐一确定数据的版权归属和授权情况，这使得侵权风险进一步加剧。

7.4.2 AI 知识产权保护的法律分析

针对 AI 技术带来的知识产权挑战，从法律层面进行深入分析，有助于探寻更为合理的保护路径。

1、专利保护方面

目前各国主要依据传统专利法的基本原则来审查 AI 相关发明。然而，面对 AI 技术的特殊性，一些国家开始探索新的审查标准和实践。如，在判断 AI 发明的创造性时，部分国家更加注重人类在发明过程中的干预程度和创造性贡献。如果 AI 发明是在人类的指导和控制下完成的，且人类对发明的最终成果具有实质性的影响，那么该发明更有可能获得专利授权。此外，对于 AI 算法相关的专利申请，一些国家要求申请人提供算法的详细说明，包括算法的设计原理、训练数据的来源和使用方式等，以便审查员能够更准确地判断其是否符合

专利条件。

2、著作权保护方面

对于 AI 生成内容的法律定性存在多种观点。一种观点认为, AI 生成内容可以被视为一种特殊的作品, 其著作权归属于对 AI 创作过程进行了足够干预和控制的主体, 如使用者或开发者。如, 如果使用者通过特定的指令和参数设置, 对 AI 生成内容的创作方向和具体内容进行了明确的引导和调整, 那么使用者可以被视为该作品的著作权人。另一种观点则认为, AI 生成内容缺乏人类的直接创作, 不应受到著作权法的保护, 而应通过其他法律机制, 如反不正当竞争法或专门的数据库保护法来进行保护。在实际的司法实践中, 不同国家和地区的法院也做出了不同的判决。如, 在某些案例中, 法院认为 AI 生成的内容虽然没有人类的直接创作, 但如果在创作过程中融入了人类的选择、判断和安排, 那么可以被视为人类创作的衍生作品, 受到著作权法的保护; 而在另一些案例中, 法院则以缺乏人类的实质性参与为由, 拒绝给予 AI 生成内容著作权保护。具体也可参见 7.3 节的有关分析。

3、商标保护方面

对于 AI 生成的商标, 审查机构需要综合考虑多个因素来判断其显著性和可识别性。除了传统的商标审查标准外, 还需要关注 AI 生成商标的独特性和创新性。如果 AI 生成的商标能够通过独特的算法和数据处理, 产生具有鲜明个性和独特风格的标识, 且能够与其他商标明显区分开来, 那么它更有可能获得商标注册。同时, 为了避免 AI 生成商标侵犯他人的在先权利, 审查机构可以要求申请人提供关于 AI 生成商标的详细信息, 包括数据来源、算法设计等, 以便进行全面的审查。

除此之外, 在 AI 知识产权保护的法律分析中, 还需要考虑到国际协调的问题。由于 AI 技术的全球性和跨境性, 不同国家和地区的法律规定存在差异, 这可能导致在跨国知识产权保护中出现冲突和漏洞。因此, 国际社会需要加强合作, 通过制定国际公约或统一的法律标准, 来协调各国在 AI 知识产权保护方面的差异, 促进 AI 技术在全球范围内的健康发展。

7.4.3 AI 责任归属的法律框架

AI 责任归属的法律框架构建是确保 AI 技术安全、可靠应用的关键环节, 涉及到多个层面的复杂问题。当 AI 技术在运行过程中产生损害或引发纠纷时, 明确责任主体和责任范围至关重要。

1、在产品责任方面, 传统的产品责任法通常将生产者、销售者等作为责任主体。然而, 在 AI 产品中, 由于技术的复杂性和多层级的参与主体, 责任认定变得更加复杂。以自动驾驶汽车为例, 其涉及汽车制造商、AI 算法开发者、数据提供商等多个主体。如果自动驾驶汽车发生事故, 确定责任归属需要考虑多个因素, 如算法是否存在缺陷、数据是否准确和完整、汽车硬件是否正常运行等。如果是算法的设计缺陷导致事故发生, 那么算法开发者可能需要承担主要责任; 如果是数据的错误或不完整导致系统误判, 数据提供商也可能需要承担相应的责任。

2、在侵权责任方面，AI 系统可能侵犯他人的知识产权、隐私权、名誉权等多种权利。当 AI 系统侵犯他人知识产权时，如未经授权使用受版权保护的数据进行训练，开发者和运营者可能需要承担侵权责任。但在确定责任时，需要考虑开发者和运营者是否尽到了合理的注意义务，如是否对数据来源进行了审查、是否采取了必要的技术措施来防止侵权行为的发生等。如果 AI 系统被用于生成虚假信息，从而侵犯他人的名誉权，那么使用 AI 系统的主体以及可能的开发者和运营者都可能需要承担侵权责任。此外，在 AI 系统处理个人数据的过程中，如果发生数据泄露或滥用，导致个人隐私权受到侵犯，数据控制者和处理者应承担相应的责任。

此外，在 AI 决策影响个人权益的场景中，责任归属也存在诸多挑战。如，在就业筛选、信用评估、司法辅助等领域，AI 决策可能对个人的生活和发展产生重大影响。如果 AI 决策存在偏见或错误，导致个人受到不公平对待，那么责任主体的确定较为复杂。一方面，算法开发者可能需要承担责任，因为算法的设计和训练可能存在缺陷，导致决策结果的不公平；另一方面，使用 AI 决策系统的企业或机构也可能需要承担责任，因为他们有义务确保所产出的结果是公平、公正的。此外，对于 AI 决策系统的审核和监管机构，如果未能履行监管职责，也可能需要承担一定的责任。

为了构建合理的 AI 责任归属法律框架，需要综合考虑技术的特性、各方主体的利益以及社会公共利益。法律应明确规定不同主体在 AI 技术开发、应用过程中的权利和义务，建立健全的责任认定和追究机制。同时，还需要考虑到 AI 技术的发展动态，及时对法律框架进行调整和完善，以适应不断变化的技术环境。

7.4.4 全球 AI 知识产权与责任归属的法律治理框架

在全球化背景下，AI 技术的快速发展使得知识产权与责任归属的法律治理成为国际社会共同关注的焦点。目前，全球尚未形成统一的 AI 知识产权与责任归属法律治理框架，但各国和国际组织都在积极探索，通过制定法律法规、政策和标准，推动 AI 治理的规范化和国际化。

中国在 AI 治理方面走在世界前列，出台了一系列政策法规，如《生成式人工智能服务管理暂行办法》、《国家人工智能产业综合标准化体系建设指南（2024 版）》等。《生成式人工智能服务管理暂行办法》对生成式 AI 服务提供者的责任进行了明确规定，要求其使用合法来源的数据和基础模型，保护知识产权，确保服务的安全和合规。在知识产权保护方面，中国强调对 AI 相关创新成果的保护，鼓励企业和科研机构加强自主创新。同时，中国积极参与国际 AI 治理合作，提出了《全球人工智能治理倡议》，倡导在尊重各国主权的基础上，共同构建公平、合理、有效的全球 AI 治理体系。

欧盟在 AI 治理领域也积极作为，其《人工智能法》为 AI 技术的监管提供了全面的框架。在知识产权方面，该法案虽然没有专门针对 AI 知识产权的详细规定，但通过对高风险 AI 系统的数据治理、透明度要求等规定，间接影响了 AI 知识产权的保护。如，对于高风险

AI 系统，要求提供者确保数据的合法性和质量，这在一定程度上有助于明确数据在 AI 创作和应用中的权利归属。在责任归属方面，《人工智能法》根据 AI 系统的风险程度，对不同类型的 AI 系统规定了相应的责任主体和义务。对于高风险 AI 系统，提供者需要建立全生命周期的风险管理体系，进行数据治理，制定和完善技术文档等；进口人、分销商和部署者也需要承担相应的合规义务，如确保系统符合相关标准、标识和说明的完整性等。

美国在 AI 治理方面则注重创新与安全的平衡，通过行政命令、行业自律和法律诉讼等多种方式来规范 AI 技术的发展。在知识产权保护上，美国专利商标局不断调整审查指南，以适应 AI 技术的发展。如，在审查 AI 相关专利时，更加注重技术的实用性和创新性，以及人类在发明过程中的贡献。在责任归属方面，美国政府鼓励企业通过自愿承诺和行业标准来规范自身行为。如，拜登政府签署的《关于安全、可靠和值得信赖地开发和使用人工智能的行政命令》行政令，推动企业在 AI 开发过程中加强安全测试和风险管理，并要求企业与政府、学术界分享经验。同时，美国的司法实践也在不断探索 AI 责任归属的法律标准，通过一系列的法律诉讼，逐渐明确了在不同场景下 AI 开发者、使用者和运营者的责任。

国际组织在推动全球 AI 知识产权与责任归属法律治理框架的构建中也发挥着重要作用。联合国通过发布相关报告和决议，如《治理人工智能，助力造福人类》，强调加强全球合作，共同应对 AI 带来的风险和挑战。在知识产权和责任归属方面，联合国呼吁各国加强协调，制定统一的国际标准，以促进 AI 技术的可持续发展。经济合作与发展组织（OECD）也在积极开展相关研究和政策制定工作，发布了《人工智能原则》等文件，为各国的 AI 治理提供指导。

尽管各国和国际组织在 AI 知识产权与责任归属法律治理方面做出了努力，但仍然存在一些问题和挑战。如，不同国家和地区的法律文化和政策目标存在差异，导致在国际协调和统一标准制定方面面临困难；AI 技术的快速发展使得法律的制定和更新难以跟上技术的步伐，存在一定的滞后性。因此，未来需要进一步加强国际合作，推动各国在 AI 治理领域的交流与协调，共同构建更加完善的全球 AI 知识产权与责任归属法律治理框架。

7.5 结语

人工智能的快速发展在带来诸多便利与创新的同时，也引发了一系列复杂且深远的 AI 治理问题，这些问题贯穿于人机关系、生命伦理、数据隐私、内容安全、虚假信息以及知识产权和责任归属等多个关键领域。在人机关系和生命伦理方面，人工智能的广泛应用模糊了人机界限，引发了对人类情感、伦理道德以及生命安全的担忧。从人机交互中的情感依附问题，到人工智能决策对人类权利和利益的潜在影响，都需要深入思考和规范。数据和隐私保护是人工智能发展中至关重要的环节。大量数据的收集和使用为人工智能提供了动力，但也带来了数据泄露、隐私侵犯等风险。确保数据的合法获取、存储、使用和共享，以及保障个人隐私不被侵犯，是当前亟待解决的问题。AI 内容安全和虚假信息问题同样不容忽视。人工智能生成的内容可能存在误导性、虚假性，这不仅会扰乱信息传播秩序，还可能对社会稳

定、公众认知产生负面影响。如何有效识别和防范 AI 生成的虚假信息，维护内容安全，是社会面临的新挑战。在知识产权和责任归属领域，AI 技术的创新性使得传统知识产权法律面临挑战。AI 生成内容的权利归属尚不明确，同时，当 AI 系统出现故障或造成损害时，确定责任主体和责任范围也变得复杂。面对这些问题，需要综合运用法律、技术、伦理和教育等多种手段。完善相关法律法规，加强技术监管和行业自律，推动国际合作，提高公众意识，共同构建一个安全、可靠、公平且符合伦理道德的人工智能发展环境，以确保人工智能技术能够更好地服务人类社会。

本章参考文献

- [1] 新华国研经济学研究院,产业经济研究部.中国人工智能：主要产业、政策环境、挑战与对策、未来趋势及经验启示[EB/OL]. 2024.12.
- [2] 中关村科技会展.国家政策|全国人工智能产业发展政策汇编(截至 2024 年 11 月)[EB/OL]. 2024.11.
- [3] 朱齐宇,李白杨,张心源,吴应强.不确定环境下我国人工智能产业政策的演化形成研究[J].现代情报, 2025.1 Vol.45 No.2.
- [4] 国家工信部,国家网信办,国家发改委,国家标准化管理委员会.国家人工智能产业综合标准化体系建设指南(2024 版)[EB/OL].2024.6.
- [5] 联合国人工智能高级别咨询机构.治理人工智能,助力造福人类-最后报告 Advanced Translation Copy[EB/OL]. 2024.9.
- [6] 上海社会科学院,武汉大学,同济大学等联合发起.全球人工智能治理研究报告[EB/OL]. 2024.11
- [7] 王可,吴颖雯,环球律师事务所.欧盟《人工智能法》合规路径思考——数据和数据治理规则[EB/OL]. 2024.10.
- [8] 宁宣凤,吴涵等,金杜研究院.欧盟《人工智能法案》图解与合规要点启示[EB/OL]. 2024.8.
- [9] 中国信通院人工智能研究所.2024 年人工智能风险治理报告——构建面向产业的人工智能安全治理实践方案[EB/OL]. 2024.12.
- [10] 中国信通院.2024 年人工智能治理蓝皮书[EB/OL].2024.12.
- [11] 中国联通研究院,中国联通国际公司. 2024 中国人工智能伦理风险白皮书[EB/OL].2024.9.
- [12] 泰和泰北京办公室.2024 年人工智能法律风险年度报告（一）——2024 年人工智能法律法规[EB/OL].2025.1.
- [13] 吴逸菲,樊春良.中国语境下人工智能伦理治理的路径塑造——基于价值-工具二维理性的融合框架[J].科学学研究, 2025.1 Vol.43 No.1.
- [14] 全国网络安全标准化技术委员会.人工智能安全治理框架[EB/OL].2024.9.
- [15] 国家网信办,国家发改委,国家教育部,国家科技部,国家工信部,国家公安部,国家广播电视总局,2023 年第 15 号令. 生成式人工智能服务管理暂行办法,2023 年 8 月 15 日起施行[EB/OL].2023.7.
- [16] 人工智能时代下的“烦恼”：美国国会听证会探讨“深度伪造（deepfake）”风险及对策[EB/OL].2019.7.
- [17] 曹建峰.深度伪造技术的法律挑战及应对[EB/OL].2019 Vol.10.
- [18] 大模型安全 | 生成式人工智能的内容安全风险与应对策略[EB/OL].2024.9.

[19] 解码生成式 AI 引发的安全风险与挑战[EB/OL].2023.10.

[20] Virginia bans 'deepfakes' and 'deepnudes' pornography[EB/OL].2019.7.