



中国互联网协会
Internet Society of China

智能体应用发展报告（2025）

中国互联网协会

2025年12月

主编单位

中国互联网协会

专家组

邵广禄 高同庆 黄澄清 曾剑秋 孙明俊

主编

武锁宁

副主编

裴 玮

执行编辑

周姝妍 李金穗 冯 杰 李景亚 郭晚平

参编人员（按姓氏拼音排序）

巴成福 毕 然 曾西平 曾勇波 常润博

陈光炎 陈珂羽 陈 逸 邓 凯 杜冀中

杜永刚 黄 莺 胡军锋 姜 昊 贾晓博

靳 伟 李 波 李胜菊 刘杰靖 李玉杰

吕尚佳 马 超 南 瑞 石 峥 孙宇航

王 琳 王睿哲 王晓思 吴 娜 许立东

袁 博 袁永强 张红光 张思源 赵 峰

赵冀昌 朱 迪 朱 捷

参编单位（按拼音首字母排序）

阿里云计算有限公司

二六三网络通信股份有限公司

国家能源集团物资有限公司

京东科技信息技术有限公司

北京市商汤科技开发有限公司

北京百度网讯科技有限公司

北京泰尔英福科技有限公司

北京智谱华章科技有限公司

联想集团

联通在线信息科技有限公司

联通数据智能有限公司

荣耀终端股份有限公司

深圳市和讯华谷信息技术有限公司

支付宝（杭州）数字服务技术有限公司

中关村智用人工智能研究院

中国铁塔股份有限公司

中电信人工智能科技有限公司

中移九天人工智能（北京）科技有限公司

中国移动信息技术有限公司

中国邮政储蓄银行股份有限公司

引言

近年来，全球人工智能的竞争焦点正经历一场深刻的结构性转变。其核心已不再仅限于基础大模型的“参数竞赛”，而是迅速转向以构建智能体（AI Agent）生态系统为目标的方向拓展。智能体，作为连接数字智能与应用场景的“神经系统”，已成为驱动大模型技术创新、商业化落地以及塑造未来产业格局的战略制高点。

大语言模型（LLM）虽然在语言理解、逻辑推理和知识储备方面展现了惊人的能力，但其本质上是一个知识层面的系统。它们虽拥有丰富的知识储备，仍缺乏与外部环境交互的能力，极大地限制了在产业应用中的融合深度。智能体的出现解决这一核心痛点。它构建了一条连接知识与应用的高效通道，通过对末端感知、记忆、分析、规划等能力的加持，提升了人工智能的执行力。智能体将 AI 的角色从被动的“信息响应者”升级为主动的“任务执行者”，使其能够真正与外部世界进行交互，并自主完成复杂任务。

面对这一趋势，国内外科技巨头对智能体的战略价值已经形成高度共识。OpenAI 认为智能体是人工智能的下一个重大突破口；微软总裁萨提亚·纳德拉在 Build 2025 大会上宣告，微软正式进入 AI 智能体时代，致力于让 AI 智能体像互联网一样自由链接、自主运作。谷歌则在开发者大会上宣布全面拥抱智能体，其 Gemini 模型的研发目标是旨在推动 AI 从“信息工具”进化为“通用智能体”。与此同时，国内电信运营商、科技互联网公司、云厂商以及终端厂商也都分别投入

到了智能体技术研发、生态构建和场景落地等领域。例如，阿里云百炼平台构建了多模态智能体生态，提供“模型调用+Agent 平台”的双重服务；中国电信发布“星辰智能体应用开发平台”，提供一站式智能体解决方案。联想在 2025 年 5 月正式发布面向消费者的“天禧个人超级智能体”，以及面向开发者的“擎天超级智能体工厂”。AI 领军企业将智能体作为战略布局的核心，全面加速其技术研发与应用。

同时，全球超级智能体市场的增长速度惊人。市场咨询公司 Markets and Markets 的预测数据显示，全球 AI 智能体市场将从 2024 年的 51 亿美元增长到 2030 年的 471 亿美元，年复合增长率达 44.8%。根据赛迪顾问近期发布的研究报告显示，2024 年中国智能体市场规模达 47.5 亿元，同比增长 64.4%；预计 2025 年将达 78.4 亿元，增速持续超过 60%，连续两年实现了翻倍的增长，到 2026 年将接近 150 亿元。多项市场数据、行业预测和技术发展趋势共同指向一个结论：**2025 年是全球智能体应用爆发元年**，标志着这项技术从实验室概念正式迈入大规模商业落地的新阶段。

我国高度重视人工智能的应用落地，将智能体的发展视为实现国家“人工智能+”战略的核心抓手。2025 年 8 月，国务院发布了《关于深入实施“人工智能+”行动的意见》，为我国人工智能产业的发展提出了明确的目标。《意见》指出，到 2027 年，要率先实现人工智能与六大重点领域的广泛深度融合，其中新一代智能终端、智能体等应用的普及率需超过 70%。展望 2030 年，这一普及率目标将提升至 90%以上，

届时智能经济将成为中国经济发展的关键增长极。这一系列目标的实现，其核心抓手正是智能体的大规模落地与应用。由此可见，实施“人工智能+”行动的关键是发展蓬勃的“智能体经济”。

中国互联网协会（以下简称“协会”）持续关注 and 关心智能体的应用推广与生态建设。早在今年1月举办的2025中国互联网产业年会上，协会前理事长邬贺铨院士就指出，互联网将进入“Internet of Agents”时代。因此，协会秘书处将智能体生态构建及推进智能体应用作为今年的工作重点。今年以来，协会致力于推动科技创新与产业应用的深度融合。先后组织了多场智能体相关研讨会，并调研走访数十家人工智能企业。今年6月，启动了“AI领航杯”“人工智能+”应用与技能大赛，收到参赛项目近2000份，其中人工智能及智能体相关项目近千份。在今年7月召开的中国互联网大会上，协会联合业界有代表性的11家单位正式启动了“智能体创新推进计划”，目前计划合作伙伴单位近百家。今年10月，启动“智能体创新应用案例征集”活动，短短半个多月，智能体案例申报数量超两百份。这些积累与沉淀，为协会分析研究智能体应用的一系列课题提供了现实基础。

调研发现：我国的智能体应用尚处于起步阶段，市场碎片化特征明显，生态建设尚未成熟；智能体的创新不是一次“单点技术突破”，而是“复杂的系统工程”；尽管数十年的数字化建设取得了一定成果，但“数据孤岛”和“生态壁垒”仍然存在；各行各业的数据采集与治理水平参差不齐，高质

量的领域数据严重缺失；加之，智能体的架构设计复杂、算力网络部署成本高昂等问题也制约了智能体的规模应用。目前，智能体市场需求呈现突出的“碎片化”特征，难以快速实现规模化应用推广。同时，商业路径仍在探索中，安全法规也须未雨绸缪。企业在没有清晰的“商业模式”和“安全治理”框架之前，难以大规模投入资源推动智能体应用落地。这一切要求要求相关行业产业的协同共建，提前研究解决标准化、商业化和安全治理等架构性课题。

智能体是人工智能打通产业落地“关键一公里”的必然选择，更是“人工智能+”战略落地的“先遣队和主力军”。只有通过技术、产业、政策等多层次的协同攻坚，才能充分释放智能体的变革性力量。为此，协会联合“智能体创新推进计划”合作伙伴单位共同研究撰写了《智能体应用发展报告（2025）》。本报告旨在系统性地剖析智能体从技术创新走向产业应用所面临的核心挑战，并尝试为产业提供跨越阻碍的战略思考及路径，推动我国在“人工智能+”的新浪潮中行稳致远，共同迎接智能体经济时代的到来。

目 录

引 言	4
第一章：智能体是 AI 落地的“关键一公里”	10
（一）智能体概念的提出、演进与定义	10
（二）L3 级智能体是现阶段的创新主体	12
（三）模型和数据是智能体构建的关键要素	19
（四）“三位一体”的智能体基本架构	22
第二章：智能体带来数字化效能全面跃升	26
（一）从数字化到智能化，破解信息化困局	26
（二）交互革命：类人感知重塑一切人机界面	30
（三）协同革命：内外贯通实现端到端智能化	32
（四）效率革命：自主学习力重构软硬件	35
（五）服务革命：“人找服务”到“服务为人”	35
（六）智能体在智能经济中的地位显著	40
第三章：智能体创新是复杂的系统工程	43
（一）智能体产业覆盖广泛、服务模式多样	43
（二）内部协同：从“单兵作战”到“军团协同”	48
（三）平台互通：打破“数据孤岛”与“生态壁垒”	50
（四）数实融合：跨越数字产业与实体经济	51
（五）算力支撑：从集中到分布的基础设施需求	53
（六）生态构建：全产业链的深度合作必不可少	54

第四章：推进智能体应用落地的市场研究	57
（一）异中求同，探索批量化路径和工具	58
（二）点面结合，解决算力网络部署难题	60
（三）由内到外，从主权充分的场景做起	61
（四）由虚向实，逐步从数字向物理渗透	62
（五）因势利导，充分释放体制政策优势	64
第五章：构建智能体生态的对策建议	68
（一）积极推进开源开放的生态环境建设	68
（二）构建“智能体互联网”的协议与标准	70
（三）探索多样化的商业合作及服务模式	73
（四）推动形成多元化智能体能力评估体系	75
（五）构建全面的标识管理和安全治理体系	76
附录：2025 智能体创新应用案例	4

第一章：智能体是 AI 落地的“关键一公里”

现阶段，人工智能产业发展正处在从理论走向实践的关键节点，智能体被认为是打通 AI 落地“关键一公里”的核心引擎。一直以来，智能体的概念与边界正随着技术迭代持续拓展。近年来以大语言模型为代表的人工智能技术，赋予了智能体自主感知、规划执行的重要能力，打破了传统 AI 的应用局限，重新定义了“新一代智能体”。当前，L3 级智能体凭借可自主决策、闭环解决复杂任务的特性，成为行业创新的核心主体，引领着技术落地的主流方向。模型与数据作为智能体构建的两大基本要素，奠定了其智能能力的根基，而“模型、工具与编排”的三位一体架构，则进一步扩展了其场景适配的潜力，让技术赋能更加广泛。

分析机构和行业专家纷纷提出，2025 年已成为智能体应用爆发的元年，政策支持、技术成熟度与市场需求共同推动智能体在千行百业加速渗透。

（一）智能体概念的提出、演进与定义

1.1 溯源：智能体发展的来龙去脉

智能体作为能够自主设定并实现目标的智能代理，其概念的演进之路与人工智能技术的发展紧密相连，实现了从自动化工具到自主性实体的升级演进。其源头可追溯至 20 世纪 50 年代，控制论、博弈论、信息论以及图灵测试的提出，为智能体的研究奠定了理论基础。这些早期理论探讨了系统如何通过反馈机制来调节自身行为以达成目标，这正是智能体自主性的核心思想。

1986年，人工智能领域先驱马文·明斯基(Marvin Minsky)在其著作《心智社会》(The Society of Mind)中，首次系统性地引入了“智能体”概念，将其描述为构成心智的简单处理单元。进入20世纪90年代，随着机器学习和自然语言处理技术的突破，智能体的研究领域不断扩展，其定义也随之演变。1995年，为消除智能体概念在研究界引发的混淆，英国著名科学家伍德里奇(Michael Wooldridge)在论文《智能代理：理论与实践》(Intelligent agents: theory and practice)中对智能体展开深入探究并提出定义：“在特定环境中感知并自主行动以实现目标的计算系统被称为智能体”，该表述至今仍被广泛引用。

2023年以来，大语言模型的出现改变了传统智能体的模块化架构，使得人们对智能体的设想不再停留于理论概念，而是具象为可落地的智能助手、自动化流程代理等实际应用场景的实体，对其研究重点扩展至多模态感知交互和自主学习能力。尤其是以OpenAI o1、DeepSeek-R1为代表的推理大语言模型迅速发展，为智能体提供了“思考大脑”，使其能够处理和推理海量的非结构化信息，从而催生了今天所讨论的、具备通用问题解决能力的新一代智能体范式。

1.2 定义：新一代智能体

为厘清边界，本报告对智能体给出以下限定内涵：**借助大语言模型为核心认知引擎，能够自主感知环境，拥有记忆存储能力，可以进行多步推理和规划，并调用外部工具来执行任务，以达成用户设定的高层级目标的计算实体。**该计算

实体可以由软件或硬件构成，核心能力包括感知、记忆、规划和执行等。

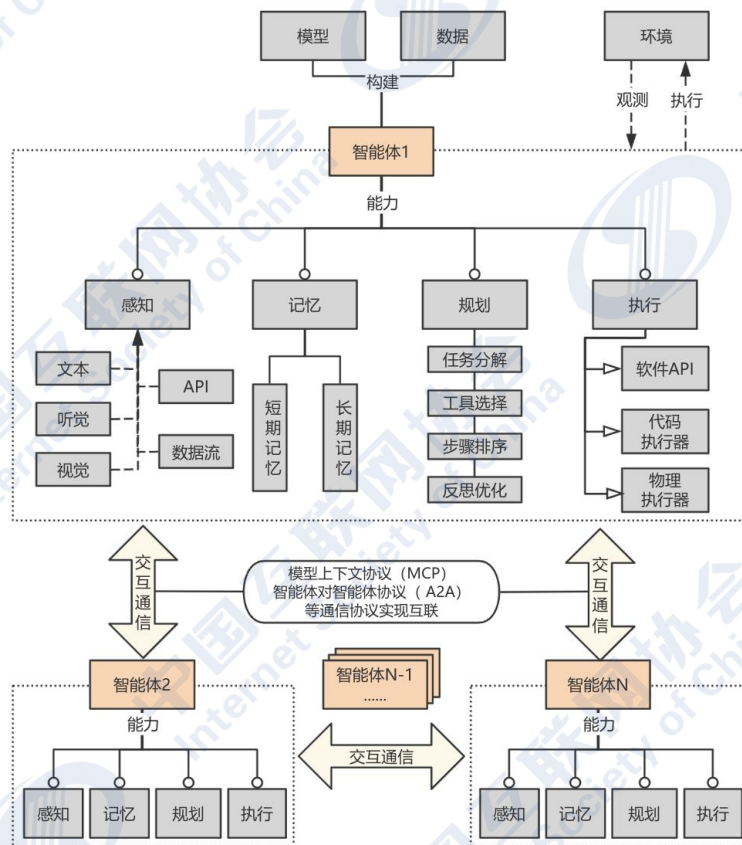
智能体与其他 AI 应用（如聊天机器人、推荐系统）的核心区别在于**自主性和开放性**。传统 AI 应用大多为反应式，它们被动地响应用户的明确指令或查询。例如，聊天机器人只有在被提问时才会回答。而智能体则是主动式和目标导向型系统，用户只需设定一个高层级目标，例如“帮我规划一次为期五天的北京旅行并预订机票酒店”，智能体便能自主地将目标分解为一系列子任务，并主动规划、执行至目标完成。同时，传统 AI 应用通常在封闭的系统内运行，其能力受限于预设的规则和数据。智能体则是一个开放的系统，其核心能力之一就是与外部世界进行动态交互，通过调用 API、访问数据库、执行代码等方式来获取信息和完成任务，从而超越了模型本身的知识 and 能力边界。总体而言，智能体完成了从“对话式 AI”到“行动式 AI”的跃迁，它不再只是信息传递者，而是真正的任务完成者。

（二）L3 级智能体是现阶段的创新主体

2.1 智能体的功能模块及构成

一个功能完备的智能体，依赖其背后复杂而精密的“技术栈”协同工作，共同构建起从感知到执行的完整闭环。感知、记忆、规划和执行是智能体的四大核心能力。**感知**可以分为环境感知和意图理解两项能力模块，是智能体的“五官”；**记忆**包括信息存储、更新以及共享，是智能体积累的“过往经验”；**规划**主要是指任务推理及分析决策能力，帮助智能

体确立从目标到路径的“蓝图”；执行包括工具调用和互联网协同，是智能体完成任务的“手和脚”。



注：中国互联网协会根据公开资料整理

在实际应用中，面临行业领域复杂的任务场景，往往需要通过多个智能体和工具、行业应用系统以及外部环境配合才能完成任务。**多智能体系统 (Multi-Agent Systems, MAS)** 应用而生，它模仿人类社会的分工协作模式，让多个具有不同专长的智能体组成一个“团队”，共同完成复杂的系统性任务。系统首先对复杂任务进行智能分解，然后将子任务分配给最适合的智能体。例如，一个“市场分析报告生成”任务，可以分解为：由“数据搜集智能体”负责从数据库中抓取信息，

由“数据分析智能体”负责处理数据并生成图表，最后由“报告撰写智能体”将所有结果整合成文。

未来，随着任务复杂度的提升，多智能体协作将成为主流应用范式。为实现各方的高效协同，业界正致力于通信协议的标准化工作，旨在打破技术孤岛，使其能够无缝地交换信息、分配任务和对齐目标，形成高效的“智能体网络”。

2.1.1 感知：连接模型与外部环境的“五官”

感知能力是智能体的“五官”，负责从环境中收集信息，用于任务的处理和执行。基于多模态大模型技术，智能体能够同时理解文本、图像、音频等多种信息输入，从而更全面地感知环境。同时，智能体广泛采用检索增强生成（Retrieval-Augmented Generation, RAG）技术，通过实时查询外部知识库，如企业内部文档、实时数据库，来克服大模型自身知识陈旧及容易产生“幻觉”的缺陷，确保其决策的准确性和时效性。强大的感知能力是智能体理解用户意图、适应动态环境的前提，智能体是否能自主准确地理解用户的需求和深层目的，并将用户指令翻译成一个结构化、明确的任务目标，是后续决策与行动的基础。

2.1.2 记忆：智能体累积过往经验的能力

记忆是指智能体对系统内部的信息进行存储、更新、检索、共享和管理。记忆是实现长期连贯性、个性化服务和持续学习的根本保障。它使得智能体能够超越单一对话轮次的限制，利用历史经验和知识进行推理和决策，进而理解用户意图，在处理长时间序列任务时带来更好的用户体验。

智能体的记忆体系通常分为短期和长期两种模式。**短期记忆**主要由基础大模型的上下文窗口负责。短期记忆会存储用户在当前对话中最近的输入、前几轮的对话历史以及正在进行的任务状态；**长期记忆**是指存储在外部的知识系统，包括向量数据库、文件存储系统和知识图谱等。长期记忆可以存储海量的、经过向量化处理的领域知识、历史记录、用户画像和行为偏好等。通过检索增强生成（RAG）机制可以快速、精确地调取相关信息，实现“外挂知识”的增强。

更高级别的智能体在记忆管理中会融合“**反思优化**”（Reflexion）技术，持续提升智能体的学习迭代能力。这个过程是将任务执行成功或失败的经验教训记录在长期记忆里，让智能体通过学习错误来优化自己的行动计划，就像给智能体建了一个“错题本”，让它面对新任务时能够做出更合理的决策，实现智能体的自我迭代和能力演进。

2.1.3 规划：从目标到路径的“蓝图”

规划能力是区分智能体与简单自动化工具的关键所在。它的具体过程是指智能体将一个高层级的目标，自主分解为一系列具体的、可执行的步骤或子任务，根据分解出的子任务，评估当前环境及可用工具，通过推理、反思等技术，输出可执行的最优行动策略。目前广泛采用的规划框架包括思维链（Chain of Thought, CoT）、回顾链（Chain of Hindsight, CoH）、ReAct（Reasoning and Acting）、Reflexion 等技术。ReAct和Reflexion提升智能体的自主规划能力。ReAct强调推理和反思的重要性，让智能体能够在多变的环境中，透过理

解任务、推理规划、执行行动和反思调整的循环过程，从而做出有效决策。**Reflexion**技术则让智能体通过动态记忆管理和自我反思机制，利用学习错误来优化自己的行动计划。

2.1.4 执行：调用万物、外部互联的“手脚”

执行能力是智能体的“手脚”，负责将规划好的“蓝图”付诸实施，对外部世界产生实际影响。这是智能体任务实现的闭环。执行主要通过调用工具来完成，这些工具包括：软件 API，例如调用企业内部的 ERP、CRM 系统，或外部的天气查询、地图导航等公共 API；代码程序，例如编写并运行代码（如 Python 脚本）来处理数据、分析问题或控制其他程序；硬件设备，比如具身智能场景下，向机器人或无人机等物理设备发送控制指令。通过执行能力，智能体真正打通了从数字智能到应用场景的“最后一公里”，使其能够完成订票、发送邮件、管理服务器、控制生产线等真实世界的任务。

在场景应用中，不仅需要考虑如何选择匹配的模型、写出合适的提示词，更需要有效地集成智能体的感知、记忆、规划、执行四大核心能力，精细化的构建智能体以满足特定场景的需求，目前“人工智能+”应用的挑战已从“算法问题”转变为“系统工程问题”。

2.2 智能体的等级划分

随着智能体技术从理论走向应用，为加强推动智能体行业的有序发展，对其能力进行系统性的分级和评估变得至关重要。中国互联网协会参考已发布的国家标准《汽车驾驶自动化分级》、自动驾驶领域的 SAE 分级标准，提出按照智能

体的自主性水平划分为 L0 至 L5 六个等级，为行业各方对智能体能力分级形成共识提供参考。

2.2.1 智能体自主性水平划分

从高度依赖人工运营管理的 L0 阶段到实现全智能化运营管理的 L5 阶段，智能体的核心特征、关键技术及典型应用存在差异，为理解智能体的能力成熟度提供了参考框架。

等级	类型	核心特征	关键技术	典型应用案例
L0	非智能体 (Non-Agent)	完全依赖预定义规则和人工操作，不具备任何 AI 能力。	传统编程、固定逻辑	传统的恒温控制器、批处理脚本。
L1	半智能体 (Semi-Agent)	在预设规则下完成简单、单一的任务，缺乏学习和适应能力。	规则引擎、基础的语音/文本识别	早期的客服问答机器人、基于关键词触发的语音助手。
L2	准智能体 (Quasi-Agent)	具备模仿学习或强化学习能力，能从有限数据中学习并做出基本推理。	机器学习、路径规划算法	天气查询系统、地图导航工具、简单的流程自动化 (RPA)。
L3	基础智能体 (Basic Agent)	基于大语言模型 (LLM)，能够理解复杂指令，进行记忆和多步推理，自主规划并调用工具完成任务。	大语言模型 (LLM)、任务规划、工具调用 (API)	智能办公助手 (如自动整理邮件、安排会议)、自动化代码生成工具、企业内部的“数字员工”。
L4	增强智能体 (Enhanced Agent)	具备较强的自主学习和泛化能力，能理解用户的长期意图，具备一定的模拟和共情能力，并能从任务中持续学习优化。	强化学习、持续学习、情感计算	个性化私人健康顾问、能够主动提供建议的财务规划助手、高级客户关系管理系统。
L5	完全智能体 (Fully)	完全自主，能够独立进行长期战	多智能体系统 (MAS)、世界	智慧城市管理系统、全自动

	Autonomous Agent)	略规划、执行和反馈，并能与其他智能体进行复杂协作，共同完成宏大任务，展现出“群体智能”。	模型、自主学习与演化	化的科学研究平台、自适应的全球供应链网络。
--	--------------------------	--	------------	-----------------------

2.3 L3 级智能体是目前产业创新的主战场

当前，全球产业界创新的焦点和主战场集中在 **L3 级智能体**。这一级别的智能体，是人工智能时代的智能体，而非计算机时代的智能体，它们以大语言模型为核心驱动，实现了从被动响应到主动规划和执行的质变。L3 级智能体能够完整地执行“从用户提问（输入层）到意图识别（理解层）、知识检索（知识层）、工具调用（执行层）再到结果生成（输出层）”的全流程，在部分应用场景中实现了过程的可视化与可干预。在企业应用中，L3 级智能体以“数字员工”的身份扮演不同层级的角色，既是一线用户的工具助手，也能成为管理者的辅助决策专家。

L3 级智能体的普及，标志着人工智能应用进入了一个高级阶段，AI 不再仅仅是辅助工具，而是能够独立完成工作的协作伙伴，这正是当前智能体经济形态的核心。

展望未来，智能体的发展将朝着 L4 和 L5 级别迈进，其核心演进方向是**多智能体系统**，即从“个体智能”走向“群体智能”。多智能体系统由多个独立的、可能具备不同专长的智能体构成，它们在一个共享的环境中交互与协作，共同完成单个智能体无法胜任的复杂任务。其核心优势在于扩展灵活性和专业化分工。通过增加或减少智能体数量，系统可以灵活地适应任务规模和环境变化；同时，每个智能体可专注于特

定领域或任务，通过协作发挥各自专长，实现“1+1>2”的系统效能。

目前，多智能体系统已经在智能交通、智能电网、供应链管理等领域初步应用并展现出巨大潜力。在未来的智慧城市管理中，交通管理智能体、公共安全监控智能体、能源调度智能体将协同工作，共同优化整个城市的运行效率和应急响应能力。多智能体协同无疑是实现 L5 级完全自主智能体的必由之路，代表了智能体应用的终极形态。

（三）模型和数据是智能体构建的关键要素

3.1 模型：认知与推理的“大脑”

大模型是新一代智能体的认知引擎，扮演着“大脑”的角色，基础模型的能力决定了智能体能力的上限。它为智能体提供了三大基础能力：首先，大模型使智能体能够理解人类用自然语言表达的复杂、模糊甚至隐含的意图；其次，经过海量数据预训练的大模型，沉淀了大量的常识性通用知识和专业领域知识，为智能体的决策提供了丰富的知识信息；最后也是最重要的，大模型强大的逻辑推理和思维链能力，使智能体能处理多步骤复杂任务，进行因果推断，并生成行动计划，提升了智能体的自主性。

但是，随着场景应用的加深，业界正涌现出专门为智能体任务而优化的模型。它们不再仅仅追求模型的通用能力，而是聚焦于增强复杂推理能力与工具调用能力，致力于提升智能体在特定复杂 workflows 中的表现。并且，支撑智能体的底层模型并非单一形态，而是根据应用场景对算力、成本和专

业深度的不同要求，形成了一个涵盖全面的选择组合。例如，通用大模型用于处理需要广泛知识的任务；行业级模型用于构建行业垂直智能体；轻量小模型用于执行快速响应任务。智能体开发者可以根据具体需求，灵活搭配使用这些模型，以在性能、成本和延迟之间找到最佳平衡点。

3.2 数据：行业应用的“知识燃料”

智能体的运行依赖多维度的数据支撑，包括内部业务数据（结构化/非结构化）、外部公开数据、感知采集数据（如视频/音频/传感器）以及动态实时数据等，为智能体执行任务提供所需的上下文、事实依据和不可或缺的领域知识。数据缺失以及质量缺陷会直接影响智能体输出的准确性与专业性。因此，数据层是确保智能体能够从通用智能走向垂直行业的关键支撑，是智能体进入千行百业的“燃料”，为智能体构建起了强大的知识图谱和行动基础。

现阶段，智能体落地难点就在于如何把企业内部多来源、多形态、多格式的各种数据，转化成能够支持 AI 模型后续训练和智能体应用的数据资产。虽然大模型在各类标准化竞赛中能够取得不错的成绩，但对于企业应用中的术语、黑话、业务逻辑等结构化困难并不了解，这也使得智能体在场景落地时可能不达预期。所以场景和应用是“标”，数据资产是“本”。数据层的广度、深度和质量，直接决定了智能体在复杂环境中的决策质量和专业深度。

目前，应用侧正积极做好数据层面的基础建设，完善多源数据治理体系，强化数据采集、清洗、标注等全流程能力，

释放高质量数据要素价值。例如，企业围绕销售、业务、技术、综合等多个线条构建企业级高质量数据集，并通过 RAG 技术构建企业级知识底座。智能体能力演进的关键驱动力是“数据飞轮”，接入实时业务数据，产生推理和决策数据，获得用户即时反馈，形成驱动领域模型和智能体迭代强化的闭环，才能使智能体的工作能力不断向上攀登。

3.3 灵活开放的智能体架构设计

在企业级应用中，需要构建一个清晰、分层的技术架构来支撑智能体系统的稳定运行和高效迭代。业界通常采用三层架构模式来构建智能体平台，实现功能解耦和资源优化，最终形成了以“用户为中心，模型为基础，数据为核心”的一站式服务平台。

上层为用户端口/服务接口，提供用户交互的界面以及轻量化 API 调用接口，确保服务类型的多样性和灵活性，是智能体能力的对外输出窗口。中间层为模型训练与推理服务，承载多模态模型、混合模型框架、知识图谱与智能体集群，是智能体实现认知、决策和执行的核心引擎。底层为数据采集与治理层，负责多源异构数据的统一接入与清洗，核心目标是保障数据的质量、一致性和可访问性。

通过分层架构的设计，使得智能体开发者可以针对性地升级和迭代任意层级的技术，而无需全面重构整个系统，从而大大提高了智能体平台的敏捷性和可维护性。目前，松耦合架构成为了主流的平台设计理念，通过对“AI 领航杯”参赛项目的分析，其中 80.95% 的项目支持轻量化 API 调用，

71.03%的项目支持插件化扩展，多以开放性、高适配性的技术方案来构建灵活的“积木式”生态。

(四) “三位一体”的智能体基本架构

随着行业实践的深入，企业的关注点已从“如何选择更好的模型”转向“如何设计规范化、标准化的智能体架构”。这标志着智能体开发已从算法探索阶段进入系统工程阶段。一个设计不合理的架构，可能导致后期高昂的重构成本。当前，业界普遍认可的智能体搭建是一个包含模型层、工具层和编排层的“三位一体”基本架构。

4.1 模型层：智能体的“大脑”

企业在选择模型时，需要根据成本、性能、响应速度和数据隐私等多重因素进行权衡。模型的选择与组合方式直接决定了智能体的认知上限和运行成本。对于性能要求高、且不涉及高度敏感数据的场景可以直接调用强大的通用大模型。部分企业场景对于隐私安全要求高，则适合选择私有化微调方案。通过在企业内部的私有数据上对开源或商业模型进行微调，以提升模型在特定领域的专业性和适配性，同时保障数据安全。“大模型+小模型/专用模型”的混合模式是目前常用的一种方式。采用一个强大的通用大模型作为任务分解和推理的“总指挥”，调用多个成本更低、响应更快的专用小模型来执行具体任务（如文本分类、情感分析），以实现成本与性能的平衡。但是，管理不同模型间数据格式转换、处理专用模型的调用失败、以及设计高效任务切换逻辑会带来大量的工程开发工作。

无论采用哪种模型组合方式，由于大模型或定制模型的虚幻现象导致智能体的决策失误是不可避免的。目前业界最先进的解决方案是基于本体+智能体的解决方案，实现企业的运营数字孪生，确保智能体的运营有着确定性的逻辑。

4.2 工具层：智能体搭建的“脚手架”

工具层包含一系列智能体可以调用的能力，如内部系统 API、外部服务接口、代码执行环境、数据库等。它是智能体与数字世界和物理世界交互的桥梁。其中，检索增强生成（RAG）技术是工具层最为关键的组成部分，它通过从外部知识库中检索相关信息，为大模型提供决策依据，有效缓解了模型的“幻觉”问题，并使其能够访问最新的私有知识。但是，构建一个真正有效的 RAG 系统，其挑战远超简单的向量检索。应对企业级应用的复杂需求，工具层的搭建也充满挑战。企业内部文档格式多样，包括报告、表格、图示甚至视频。如何设计最优的分块，策略以在切割文档的同时保留其完整的语义结构，是一大难题。一个真正有效的 RAG 系统必须能够统一索引并检索文本、表格、图表和图像中的信息，这需要复杂的多模态模型和检索技术，技术门槛也极高。

4.3 编排层：智能体的“神经中枢”

编排层负责协调模型层和工具层的工作，管理从任务理解、规划、执行到结果反馈的整个流程。它是智能体的“总调度师”，是智能体的“神经中枢”，决定了智能体工作的逻辑和效率。编排层通过动态协作框架，实现任务分解、动

态规划、角色扮演和协同推理等高级功能。当前，市场上的编排框架可以分为两大阵营：灵活的开源框架和集成化的商业平台。目前有多种主流的开源智能体框架可供选择，它们在设计理念和适用场景上各有侧重。

框架	核心焦点	优势	最佳适用场景
LangChain	构建模块化的 LLM 应用和智能体	生态系统庞大，社区活跃，集成度高	通用型 AI 应用、RAG 流程、聊天机器人
LangGraph	基于图的工作流编排	支持多智能体、重试、人机协同	需要分支逻辑的复杂企业工作流
CrewAI	基于角色的智能体协作	角色分工明确，共享上下文	需要多个具有不同职责的智能体团队协作的任务
AutoGen	异步的多智能体对话	支持实时并发对话，事件驱动	需要多个 LLM“角色”进行实时交互的场景
Semantic Kernel	面向企业的“技能”编排与集成	多语言支持（C#, Python），与微软生态深度融合	企业级应用，尤其是在 .NET 生态中需要将 AI 与非 AI 服务结合的场景

图表：主流的开源智能体框架

企业级应用中的核心诉求是追求确定性和可靠性。每一次操作都必须有据可查，以确保安全合规与与审计追溯要求。然而，基于大模型的智能体本质上是概率性系统，其行为具有内在的不可预测性。这使得编排层的设计挑战从简单的“ workflow 管理”提升到了“为概率系统强加确定性行为”的高度。编排层因此成为一个复杂的工程化“约束系统”，其首要任务就是驯服 AI 的不确定性。在企业环境中，“能用”是远远不够的，必须能够证明系统“为何以及如何”工作，尤其是在出现问题时。如何在保留智能体自主性的前提下，设计并实施这些强制性的控制系统，是编排层设计的核心难

题之一。

“三位一体”的架构为智能体提供了从思考到行动的完整闭环。然而，将这一概念架构转化为投入生产面临着更深层的工程问题。真正的困难并不是选择性能最强的模型或最先进的工具，而在于如何将这三个层面无缝集成，并对其进行系统化管理。这需要解决模型间的协同、工具的可靠调用、以及编排逻辑的鲁棒性等一系列棘手问题，其复杂性远超模型选型本身。

第二章：智能体带来数字化效能全面跃升

智能体不仅代表着人工智能技术的演进，更为重要的是，它正引领一场经济社会数字化效能的全面跃升。数字化浪潮奔涌十年成果颇丰，但企业仍面临“数据孤岛、系统割裂”的深层困局。智能体的崛起，正为这一困局找到突破口。

自 2025 年起，在政策推动与技术迭代的双重驱动下，企业正加速从“信息化升级”转向“智能化重构”。这场由智能驱动的效率革命，从交互方式、协同模式、执行效率和服务形态四个维度为传统信息化系统所带来的变革性升级，初步实现穿透系统孤岛、打破组织壁垒，构建起人机共生、自我演进的新数字化形态。从破解困局到重塑范式，智能体正在重新定义数字化的价值内核。智能体的核心价值在于全面提升企业数字化转型的效能，推动组织从被动响应的工具使用者，进化为具备自主学习与持续优化能力、人机深度协同的“自进化组织”。

（一）从数字化到智能化，破解信息化困局

自 2015 年国家首次提出数字化转型以来，这一战略成为国家经济发展和现代化建设的核心战略之一。经过十年发展，数字中国建设取得很大成绩。然而，随着新一轮科技革命和产业变革的深入，数字化进程正迈向数智化、智能化全面跃升的新阶段，传统数字化解决方案的局限性也日益凸显。

1.1 数字化转型面临的瓶颈与痛点

尽管中国企业的数字化转型已从早期的单点技术应用，逐步走向全域生态重构，但区域发展不平衡、中小企业转型滞后、数据治理体系不完善等宏观问题持续制约着行业整体的效能提升。即便是在已完成数字化基础设施建设的大型企业中，数字化解决方案也长期面临着三大核心痛点。

首先是严重的**数据孤岛**。企业的核心业务数据，如客户关系、供应链、财务和人力资源等，往往分散在为特定功能而独立设计的系统中，例如客户关系管理（CRM）、企业资源计划（ERP）、供应链管理（SCM）等，系统在建设时缺乏统一规划，形成了信息壁垒。此类“孤岛型作业方式”导致了严重的数据碎片化和信息不互通，使得管理层难以形成一个全面、统一的业务视图。例如，某些制造企业因生产系统与供应链系统的数据不互通，导致对市场需求的误判和生产计划的滞后，最终造成库存积压超过 20%，束缚了企业的利润提升。数据孤岛的存在，使得跨领域的数据分析和洞察变得异常困难，数据作为核心资产的价值很难被最大化。

其次是普遍存在的**流程分割**。在数据孤岛的基础上，业务流程被迫在不同的系统和部门之间进行人工“接力”。例如一个典型的“从订单到收款”流程，可能需要销售人员在 CRM 系统中创建订单，随后手动通知财务人员在 ERP 系统中开具发票，再由物流部门在 SCM 系统中安排发货。每一个环节的转换都须依赖或等待人工干预，不仅效率低下、容

易出错，更构成了企业内部巨大的、难以量化的隐性运营成本。这些断裂的流程消耗着大量的人力资本，成为组织敏捷性的主要障碍。

最后是**决策高度依赖人工经验**。传统的组织架构与文化惯性制约了数据驱动决策的真正落地。尽管许多企业引入了商业智能（BI）等先进工具，能够将历史数据可视化，生成各类报表和仪表盘，但这仅仅是将数据“呈现”给人。最终的复杂决策，尤其是面对动态市场变化时的战略决策，仍然高度依赖管理层个人的经验、直觉和判断力。数据在这一模式下扮演的是“支撑”角色，而非“驱动”角色。决策的质量和效率受限于人的认知边界和处理信息的能力，难以应对日益复杂的商业环境。

总体而言，过往的信息化和数字化方案在一定程度上解决了企业流程的“标准化”和“线上化”问题，是数字化转型不可或缺的基础。然而，其固定、封闭和被动的本质，使其在面对复杂多变的需求时显得力不从心。系统被动地响应用户的指令，依赖预定义的规则和固化的逻辑，无法应对动态变化，最终导致决策效率低下和组织僵化

1.2 数字化向智能化迈进

与传统信息化方案的被动响应模式截然不同，智能体解决方案凭借其三大核心特性——自主性、交互性和持续学习能力，正成为破解数字化困局的关键，改变数据与业务流程

之间的关系，引领企业运营模式“数据支撑决策”到“数据驱动行动”的根本性转变。在传统模式下，系统将数据呈现给人，由人进行分析、决策，并最终手动执行。而智能体则能够直接“阅读和理解”数据，基于预设的目标自主进行推理、规划，并调用所需工具来完成任务。通过构建由多模态数据融合、知识图谱和多智能体协同组成的智能应用集群，智能体能够主动打通数据孤岛，连接并操作那些原本相互割裂的业务系统，将过去需要人工处理的复杂决策流程实现端到端的自动化。

然而，实现这一愿景并非易事。这意味着企业必须拥有良好的数字化转型基础。目前明确的实施路径是：企业需要先“修路”，再“跑车”。即首先构建规范化、标准化的数字化基础设施，包括清晰的业务流程、开放的 API 接口以及统一的数据治理体系。只有在基础牢固的“数字高速公路”上，智能体这辆高性能的“汽车”才能安全、高效地行驶，真正融入业务、参与决策、释放执行力，成为推动组织进化的新引擎。

1.3 生态的繁荣推动智能化进程

AI 应用开发工具和应用生态的繁荣，大幅降低了智能体行业应用的门槛，并加速了产业智能化落地进程。一方面，众多开放的开发平台，汇聚了丰富的开发框架、工具组件、算法资源、海量的数据集等，开发者可便捷调用模型、微调

模型、快速验证想法、构建 AI 应用。另一方面，开源模型因其具备可定制性强以及社区支持丰富等优势，正吸引越来越多的企业选择将其作为技术解决方案的重要组成部分。得益于国家战略支持与产业生态形成的强大合力，为企业从“数字化”迈向“智能化”铺平了道路，智能体的在千行百业的应用正在快速验证落地，预示着一个由智能体驱动的全新商业时代的到来。智能体正通过四层重构颠覆企业经营模式：交互层通过自然语言取代复杂界面；协同层通过目标导向替代固定流程；执行层从软件到硬件双重提效；服务层从“人找服务”到“服务为人”。有机构预测，到 2026 年，60%的企业将部署智能体，其核心价值不仅是效率提升，更是通过突破数字化的局限，释放组织创新潜能。

(二) 交互革命：类人感知重塑一切人机界面

智能体正在引领人工智能入口的交互革命。传统的图形用户界面（GUI）和命令式交互，要求用户学习并适应机器的逻辑。现在，智能体类人化、多模态的感知能力能够以更自然、更主动的方式服务于人。从“人适应系统”到“系统适配人”，实现了“焕新一切交互”。

这种交互变革为消费科技领域（toC）和企业应用领域（toB）都带了改变。首先是消费科技领域，交互模式正从早期的 Siri 等功能有限的语音助手，演进为能够跨应用执行复杂任务的下一代智能体设备。例如，在大模型赋能下，集成

了智能体的手机能跨应用自主操作完成订票、点餐等复杂任务。未来的 AI 眼镜、智能耳机将成为无所不知的百科全书，能够听、看、感知并与用户共情，类似科幻电影中“贾维斯”式的智能伙伴体验正逐渐成为现实。在各类服务中，用户无需逐一操作设备或跳转平台，只需通过一个简单的指令，智能体便能理解其意图，并主动编排和执行一系列跨设备跨平台的操作，将服务丝滑地融入用户的日常生活。

另外，在企业应用领域，交互的变革带来了全新的工作模式。例如，车间的经理想要了解机器操作情况，只需要举起手机对准设备提问：“这台机器的维护历史是什么？有没有未完成的工单？当前生产批次推荐的压力设定值是多少？”智能体利用计算机视觉识别设备，通过自然语言处理理解问题，然后分别调用维护日志系统、ERP 系统和工艺手册数据库的 API，最终将来自多个后台系统的信息整合成一个简洁、直接的答案，呈现在用户面前。无论员工的技术背景如何，都能通过简单的对话与强大的后台系统进行交互，从而获取数据、洞察信息、驱动业务。

国内领先的协同办公平台钉钉，已经将智能体助理深度集成到产品中，成为企业员工的“超级助理”。员工不再需要登录复杂的 BI 系统，拖拽维度和指标来制作报表。他们可以直接在钉钉的对话框中向 AI 助理提问，例如：“帮我分析一下华东大区上个季度的销售额和利润率，并与去年同期

进行对比。” AI 助理会自主理解问题，调用后台数据库 API，进行数据分析，并以图表或文字摘要的形式直接在对话中呈现结果。这种交互方式的变革，极大地提升了企业内部知识和数据的流动性。当每一个员工都能轻松地与企业的数据和流程进行交互时，才能真正释放企业在数字化基础设施上巨额投资的全部潜力。

(三) 协同革命：内外贯通实现端到端智能化

交互革命解决的是“人与系统”之间沟通的鸿沟，而协同革命解决的则是“系统与系统”、“数据与行动”之间的壁垒。智能体通过连接不同系统的 API，实现了数据与信息的自由流动，将原本需要人工执行的跨系统复杂流程自动化，从而实现真正意义上的端到端智能化。这场革命的核心在于智能体的两大能力：“灵活调用工具”和“深度释放数据价值”。

3.1 灵活调用工具

“工具调用”（Tool Calling）是赋予大语言模型与外部世界互动的核心技术。它使得智能体不再局限于其内部的静态知识库，而是能够像人类一样，在需要时主动使用外部工具，例如调用一个 API、查询一个数据库或执行一段代码来获取信息或完成任务。在企业环境中，每一个拥有 API 接口的软件系统，无论是内部的 ERP、CRM，还是外部的供应商平台、天气服务，都变成了智能体可以调用的模块化能力。

“工具调用”是对传统 API 经济的一次智能化升级。过去，API 的调用逻辑需要由程序员预先在软件中编码。而现在，智能体可以根据用户的动态需求和上下文环境，自主决定调用哪些工具，并以何种顺序调用，从而实现高度灵活和智能化的流程编排。目前，国内外出现了丰富的开发框架及工具组件，能够提供标准化的接口和预置的连接器，使得开发者可以轻松地将任何 API 或内部函数封装成智能体可以调用的工具，从而将整个企业的 IT 能力向智能体开放。

在企业应用中，以物流调度系统为例，智能体通过工具调用不仅能让数百台机器人同时作业，还可以管理车间的大门、道闸、红绿灯等物联网设备，真正实现多种设备和车辆的互联互通与协同化工作。以企业采购和供应链管理为例，这是一个典型的长链条、跨部门、内外协同的复杂流程。传统的数字化系统可以提供各个环节的数据，但流程的推进和决策仍需大量人工干预。现阶段的供应链智能体能够实时解析销售数据及库存水平，自动调用供应商 API 或访问比价网站进行询价、查询比对候选供应商的历史履约记录、认证资质等为供应商确认下单，最终可以实现从需求产生到采购下单的协同超级自动化，极大地提升了决策质量和运营效率。

3.2 释放数据价值

智能体的另一核心价值在于其“激活一切数据，释放数据价值”的能力。在智能体出现之前，数据的主要服务对象

是人，人通过分析数据报表来做出决策。而在智能体时代，数据可以直接服务于机器，智能体不仅能呈现“是什么”，还能解释“为什么”，并预测“会怎样”，最终将洞察直接转化为行动。这种转变的逻辑原理在于，智能体将企业的各个系统和数据源连接成一个有机的“企业神经系统”。数据孤岛和流程碎片化，就像生物体中大脑与四肢之间的神经连接不畅，智能体平台则扮演了中枢神经的角色。它通过 API 接口感知来自各个业务“器官”（ERP、CRM、SCM、物联网传感器等）的数据信号，在其“大脑”（大模型与业务逻辑）中进行处理、分析和决策，然后通过工具调用向各个“器官”发出精确的行动指令。智能体通过推理能力激活数据进行分析，帮助企业更好地理解自身的底层数据资产，提升企业的智能化水平及协同效率。

智能体改变了数据的使用范式，企业采购是体现数据价值释放的典型场景。传统的采购流程高度依赖采购员的个人经验和谈判技巧，而采购智能体或采购供应链一体化平台则可以通过全流程的数据驱动，实现多环节协同优化。国家能源集团物资公司自主研发了“基于人工智能模型的能源采购供应链数智一体化平台”，平台依托“多智能体+专有数据+专家知识”的架构逻辑，实现了能源采购各关键业务流程的智能化重塑。从需求预测、供应商评估到合规管理与辅助决策，平台不仅提升了采购业务的自动化、标准化与透明化水

平，还显著降低了研发投入与运维成本。据测算，平台在上线后通过压缩评审会议、减少调研成本、优化培训与差旅资源配置，节省了大量人力物力资源，同时提升了服务响应效率与用户体验，助力企业建立起以 AI 为核心驱动的新型业务模式。智能体将原本沉睡在各个系统中的供应商数据、价格数据、履约数据等激活，并直接转化为降低采购成本、提升供应链韧性的商业价值。这正是从“数据支撑决策”到“数据驱动行动”的核心所在。

（四）效率革命：自主学习力重构软硬件

智能体拥有强大的自主学习能力，在其与环境持续交互的过程中，能够自动获取知识、改进决策策略、优化行为模式，无需人工干预。这项能力为软硬件的效率提升开辟了动态演进的新路径，核心变革在于将传统“固化配置”转变为“自适应调优”。在数字世界，它正从根本上改写软件的生产 and 运行方式，开启“重构所有软件”的进程；在物理空间，它深度赋能硬件设备，尝试“驱动一切硬件”，将智能制造推向新的高度。

4.1 重构软件开发运营

在软件开发领域，智能体正从两个层面带来颠覆性变革。一方面，它催生了全新的 AI 原生应用；另一方面，它正成为开发者的强大“编程助手”乃至“自主程序员”，深刻重构了软件的开发、测试与运维（DevOps）全生命周期。

编程智能体（Coding Agent）能够根据开发者用自然语言描述的需求，自动生成代码、编写注释、调试错误、优化性能，甚至部署上线。其中，智能体的自主学习能力是提升软件开发效率的关键驱动力，它通过引入分层记忆管理系统（如情节记忆、工作记忆和语义记忆），使智能体能够从历史交互中总结项目经验和个人偏好，实现自我学习和进化。例如，阿里发布的 Qoder 平台具备长短期记忆系统，可记录用户“完成任务后自动生成单元测试”等习惯，后续执行时无需重复指令，直接优化 workflow。这种人机协同的编程新范式，极大地解放了程序员的生产力。研究与实践表明，编程智能体可以将程序员的编码效率提升 30% 至 50%，显著缩短研发周期，降低开发成本。智能体的自主学习能力进一步体现在其对复杂任务的适应性上。以通义灵码为代表的 Coding Agent 通过反思迭代（ReAct 模式）和工具调用能力，能够根据代码执行结果动态调整策略，而非依赖固定 workflow。

编程是智能体驱动外部软硬件工具的桥梁。Anthropic 公司的 Claude 模型把代码编程作为核心竞争力，其能力与智能体的自主学习特性相契合。国外 AI 编程平台如 Cursor 持续迭代并接入更强大的模型。在国内，阿里巴巴的“通义灵码”已广泛应用于吉利汽车、中华财险等企业，其核心优势在于 Agentic Coding LLM 的演进——它不仅基于静态代码训练，更能从软件开发动态过程（如 Bug 修复、测试用例生成）中

学习，从而处理仓库级复杂任务。百度“文心快码”、华为云 CodeArts 等平台也集成类似能力，共同构建了融合自主学习的 AI 辅助编程生态。未来，随着多智能体协同协议的发展，具备自主学习能力的智能体将进一步打通需求分析、设计、开发、测试全链路，形成更高效的自动化软件开发体系。

4.2 驱动硬件智能升级

工业智能体与物理空间的深度融合，尤其在自主学习能力的赋能下，正在为智能制造带来一场深刻的效率革命，推动生产线从自动化向自主化跃升，加速“黑灯工厂”愿景的实现。

工业智能体区别于传统自动化系统的核心在于，它是一个专为工业生产设计、具备自主感知、认知、决策和学习能力的软硬一体系统。它不再仅仅依赖预设指令，而是能够通过融合多模态感知、大模型任务规划等技术，在复杂环境中动态适应并优化任务执行策略。尤为关键的是其数据闭环持续学习机制。中国移动紫金研究院研发的“灵御”巡检机器人，将大模型驱动的智能体融入云边端巡检体系，实现了自然语言交互、知识增强决策、灵活任务调度与闭环优化执行，推动了巡检机器人向更高层次的智能化与人机协同演进。

智能体的自主学习能力有助于生产任务自主化提升。传统工业机器人依赖预设程序，在稳定环境中执行固定任务。而融合了自主学习能力的智能体，例如，一个部署在生产线

上的智能体，可以实时分析设备传感器数据，预测潜在的机械故障，并自动安排维护计划，从而大幅减少非计划停机时间。另外，智能体可以赋能生产流程的系统性优化。智能体能够基于实时数据和历史经验进行全局优化自主学习能力，是打通工厂全链路自动化的关键。未来的智能工厂将是以模型为知识和决策中枢，能够自主组织资源、应对变化的有机系统。

(五) 服务革命：“人找服务”到“服务为人”

智能体的出现颠覆了原有的人机协作模式：用户不再需要主动提出需求、寻找所需的服务，而是由智能体主动预测需求，并直接向用户提供个性化、情境化的解决方案。这标志着 AI 应用从一个被动的“辅助工具”演变为一个主动的“协作伙伴”，服务将变得无处不在且高度智能。

5.1 从“辅助工具”向“协作伙伴”转变

大语言模型发展初期，AI 作为辅助工具生成内容、回答问题，其余工作任务仍由人类完成。在智能体的推动下，AI 扮演独立工作的协作伙伴，不仅能够响应人类的请求，还能根据环境变化独立决策、主动行动，向人类直接提供任务结果，在数字系统建设中凸显“以人为本”的理念。过去的模式是“人找流程”：人类主动登录多个系统，在复杂界面里找入口、导数据、提申请，主动迁就机器和流程，费时又费力；智能体模式翻转为“流程找人”：智能体主动理解目标，

主动调度后台各系统和服务，完成所有步骤，仅将唯一需要人来决策或确认的节点精准推送给人。

服务范围也从提供单点赋能，转变为提供一站式服务，服务链条更加完整，定位也从“过程工具”升级为“结果交付”。人们可以将重复性、繁琐的事务交给智能助手，或使用智能体优化流程、协助任务处理，从而节省精力，专注于创新性和前瞻性工作。通过分析中国互联网协会举办的“AI领航杯”大赛项目，我们发现，人工智能落地应用时，“提供一站式服务，形成统一入口”的项目接近 77%；而“AI 自主执行”已逐渐成为主流的人机协作模式，占比 49.40%。权威分析机构 Gartner 预测，到 2029 年，80%的常见客户服务问题将由智能体自主解决，无需人工干预。

5.2 “数字劳动力”市场正在兴起

在企业内部，智能体正以“数字员工”的身份被整合到各个业务流程中，承担起日益复杂的工作，一个庞大的“数字劳动力”市场正在兴起。这些数字员工 7x24 小时不知疲倦地工作，以极高的效率和准确性处理着从财务、人力资源到客户服务的各类任务，深刻重塑了企业的内部服务模式。有机构预测，到 2028 年，预计 33%的企业级软件应用将内嵌智能体能力。

邮储银行基于 AI Agent+数字员工系统创新性探索人机交互服务新模式。在线下（云柜数字员工）、线上（数字客

服、数字人直播等)场景中打造“数字员工”体系,为客户量身定做每一次有温度的交互,通过生动拟人化的形象、自然的交流方式,与客服、营销、运营、宣传等场景结合,打通人与机器之间的情感,激发业务创新发展活力、推动金融服务从“功能覆盖”向“体验升维”转型,实现服务自动化、智能化与人性化,打造企业级的数字员工范式应用。融合“数字员工”作为前台交互载体与“AI Agent”提供中台智能决策引擎,实现“体验无感化、服务自动化、风控智能化”三大目标。

交互的自然化、协同的无缝化、执行的高效化以及服务的主动化——可以清晰地发现,智能体为企业带来的绝非简单的降本增效,而是一场深刻的组织范式革命。它正在推动企业从一个由人类主导、层级分明、流程固化的传统组织,向一个数据驱动、敏捷适应、人机深度协同的“自进化组织”演进。在这个新的组织形态中,智能体构成了组织的“数字神经系统”,负责感知内外部环境变化、执行海量标准化任务。而人类员工的价值则发生了根本性的转移。当重复性的流程执行工作被大规模自动化后,人类的经济价值不再体现在“动手”的能力上,而是更多地体现在那些机器尚无法企及的高阶认知能力上。

(六) 智能体在智能经济中的地位显著

在中国，智能体、多模态模型与模型部署共同构成人工智能产业的三大核心增长引擎，标志着智能体广阔的市场前景获得了行业的认可。有观点认为，未来十年，业界对智能体的投入规模，将是今天对大模型投入的百倍甚至千倍。如果说 AI 服务市场规模可达万亿美元，那么由智能体构成的“AI 数字劳动力”市场规模则有望超过十万亿美元，是移动互联网市场规模的数十倍。

智能体不再是未来的愿景，而是正在发生的产业革命。2025 年，正是这场革命全面拉开序幕的引爆点，抓住了智能体，就抓住了“人工智能+”时代的核心机遇。智能体作为连接基础模型能力与千行百业应用的“路”和“桥”，是推动“人工智能+”战略、构建智能经济的核心支点之一。国家与地方层面密集出台的支持政策，为智能体的发展注入了强劲动能，也展示了智能体在智能经济推动中的影响力和战略地位。

北京市经济和信息化局在 2025 年 4 月发布的《支持信息软件企业加强人工智能应用服务能力行动方案》中，明确提出“支持通用智能体发展”，鼓励开发能够跨领域、多任务、自规划的智能体系统。上海市经济和信息化委员会同期启动的新一代 AI 创新任务揭榜挂帅工作，也将“复杂开放环境下异构智能体的协同与群智涌现”问题列为重点攻关方向。

与此同时，行业共识也在加速形成。中国信通院联合工商银行、百度、腾讯等二十余家行业领军企业共同发布的《面向软件工程智能体的技术和应用要求》，首次从技术能力和

服务能力两大维度对智能体开发框架进行了规范，为企业提供了宝贵的能力建设与技术选型指南。

市场预测数据进一步印证了智能体的战略重要性。德勤咨询预测，到 2025 年，使用生成式 AI 的企业中将有 25% 部署 AI 智能体，这一比例到 2027 年将攀升至 50%。Gartner 的预测更为大胆，认为到 2028 年，将有 33% 的企业软件应用程序内置智能体 AI，并且 15% 的日常工作决策将由智能体自主完成。在部分垂直领域，行业已从“百模大战”迅速迈入“千体之争”的新阶段。今年，在中国互联网协会举办的“AI 领航杯”“人工智能+”应用与技能大赛中，涌现了大量智能体或准智能体的应用案例。数据显示，48% 的项目以开发智能体作为“人工智能+”应用的主要载体，搭建智能体已经成为企业应用 AI 的重要方式之一。这标志着智能体作为 AI 应用的核心载体，其战略地位已无可替代。

第三章：智能体创新是复杂的系统工程

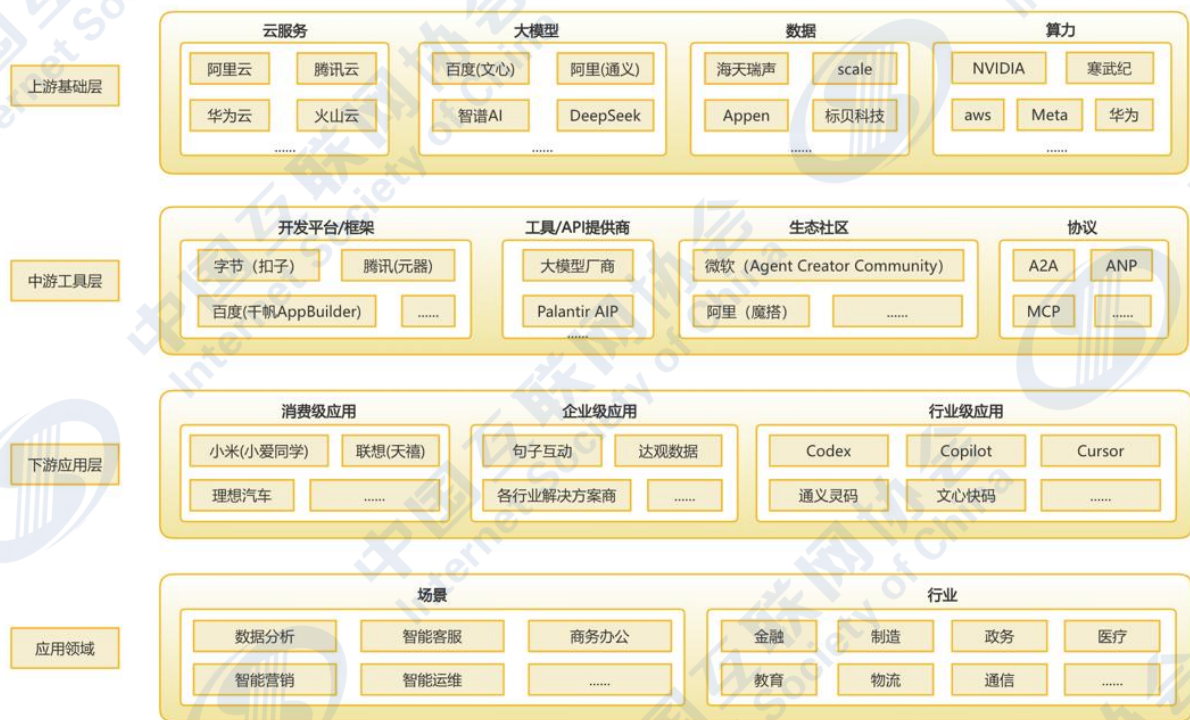
智能体的功能是革命性，为企业带来了效能的提升，其市场前景美好、潜力巨大。但是就像“樱桃好吃树难栽”，要实现这些目标需要系统创新。构建智能体并投入生产应用并不是简单地接入模型或者调用 API 就足够了。智能体创新是跨流程、跨企业、跨产业的系统性工程。从上游算法研发到下游场景落地，产业链条上各类企业需要紧密配合；组织内部必须打通 workflow，实现研发、业务与运营的一体化；不同平台间需建立通用协议与信任机制，破解长期存在的“数据孤岛”与“生态壁垒”，让数据流动起来；智能体技术需要跨越数字世界与物理世界的鸿沟，无缝嵌入制造、医疗、金融等实体业务中；算力与网络设施需要从集中式数据中心向分布式边缘节点演进，为智能体提供实时响应的基础保障。这些遍布在智能体创新应用之路上的复杂问题，需要以系统性思维统筹谋划，只要构建全产业链深度协作的生态体系，才能将技术可能性转化为可持续的创新动能。

（一）智能体产业覆盖广泛、服务模式多样

1.1 智能体产业链可划分为三层

智能体产业是一个由上游基础资源、中游开发工具与下游场景应用构成的生态系统。中国互联网协会根据公开资料整理了智能体生态图谱，梳理了人工智能的应用价值从创造、传递到交付全过程的参与者，展示了从技术研发到商业落地的完整链条，供业界参考。

智能体产业链可划分为三层：上游基础层提供核心的算力、数据和模型；中游工具层提供开发平台与框架，赋能生态发展；下游应用层则聚焦于将智能体技术与具体场景结合，实现最终的商业价值。



备注：篇幅有限，仅以部分企业为例展现生态特征，不代表全貌

1.1.1 上游基础层：构建智能体的“基石”

上游基础层是整个智能体产业的“基石”，为所有中下游创新提供了不可或缺的基础条件。这一层由算力基础设施、数据资源和基础大模型三大支柱构成，其发展水平直接决定了整个生态的能力上限。

国内基础大模型市场以百度“文心”、阿里巴巴“通义”、商汤科技“日日新”以及智谱 AI、DeepSeek 等为代表的模型厂商，构成了模型供给的核心力量。当前，竞争的焦点已从单纯追求参数规模的“百模大战”，转向针对智能体特定需求

的能力优化，例如增强模型的复杂推理与工具调用能力，以更好地支撑下游应用。通用模型虽知识广博，但在执行精确、复杂的行业任务时仍显不足。因此，专门为智能体任务优化的推理模型，如 DeepSeek-R1 被行业广泛调用，标志着模型层正从“通用化”向“专用化”和“场景化”演进。

其次，算力与云基础设施是驱动智能体的“动力源”。大模型的训练和推理需要海量的高性能计算资源，这使得以阿里云、百度智能云、腾讯云为代表的云服务巨头在上游扮演了关键角色。这些厂商不仅提供底层的 GPU 算力，更通过其完善的云平台服务，将算力、模型与开发工具深度整合，形成了一种强大的生态绑定效应：下游的智能体开发者在选择某一家的基础模型时，往往也深度绑定了其云服务生态。这种“模型即平台，平台即生态”的策略，使得少数头部厂商不仅掌握了核心技术，更掌握了市场的话语权与控制力。

1.1.2 中游工具层：生态繁荣的“连接器”

中游工具层在产业链中扮演着“连接器”的角色，它将上游原始的技术能力，转化为开发者可以便捷使用的工具、平台和框架，极大地降低了智能体应用的开发门槛。例如，智能体开发与编排平台，通过集成化的开发环境，将复杂的提示词工程、检索增强生成（RAG）、记忆管理、工具调用和多智能体 workflow 编排等功能模块化、可视化，让开发者能够像搭建“乐高积木”一样，快速构建和部署智能体。以字节跳动“扣子（coze）”、百度“千帆 AppBuilder”、腾讯“元器”、用友“BIP 智能体构建平台”、为代表的平台，正是这一趋势

的典型代表。它们不仅提供了丰富的预置工具和模板，还支持通过轻量化的 API 和插件化扩展，与企业现有系统进行灵活对接。对“AI 领航杯”参赛项目的分析显示，超过 80% 的项目支持 API 调用，超过 71% 支持插件化扩展，这充分说明了市场对开放、灵活架构的偏好。

从更深远的战略层面看，中游平台之争本质上是未来智能体经济“操作系统”的占位竞争。一个成功的平台，如同 PC 时代的 Windows 或移动互联网时代的 Android，通过提供标准化的接口、丰富的开发工具和繁荣的应用生态，吸引海量开发者。一旦形成规模效应，平台将成为连接上游模型和下游应用的核心枢纽，掌握其生态定义权。因此，中游不仅是技术工具的集合，更是决定未来产业格局的关键战略高地。

1.1.3 下游应用层：AI 商业价值的“交付者”

下游应用层是智能体技术与真实场景需求相结合、创造并兑现商业价值的最终环节。当前，下游应用呈现出面向消费者（toC）和面向企业（toB）两大方向。面向企业端的智能体将聚焦逻辑严谨性和工作效率提升，主要承担重复性任务处理，其核心价值在于工具实用性和流程优化；面向消费者端的智能体则侧重交互创新和服务体验，具备更强的思考能力和记忆功能。

消费级应用主要以个人助手和智能硬件为载体。软件层面，各类智能助理正变得更加主动和全能；硬件层面，终端厂商正将智能体深度嵌入智能手机、AI PC、新能源汽车等设备中，代表企业包括小米、联想和理想汽车等。企业级应

用则展现出更为广阔和多样化的前景，其核心形态是“数字员工”或“行业专家助手”，以“智能体即服务”（AaaS）或集成于 SaaS 软件的形式交付。这些智能体被应用于智能客服、数据分析、营销自动化、智能运维等多种场景，旨在提升企业运营效率和决策水平。

1.2 国内业务的服务模式多样

1.2.1 以初创企业为代表的轻量化开发

由技术敏锐、行动迅速的中小型科技创新企业主导。它们通常不自研基础大模型，而是通过调用上游厂商的 API，聚焦于解决特定业务场景中的高频痛点，以 SaaS 订阅的模式提供轻量化、标准化的智能体解决方案。其核心竞争力在于对特定业务流程的深刻理解和快速产品化的能力。

1.2.2 以终端厂商为代表的端侧智能体

传统硬件终端厂商正积极拥抱智能体，将其作为产品差异化和重塑用户体验的核心战略。该模式通过在智能手机、PC、汽车等设备上部署端侧 AI，实现更低延迟、更具个性化且隐私保护更强的智能交互。

1.2.3 以大型科技公司为代表的平台化赋能

大型企业软件和云计算公司凭借其深厚的客户基础、技术积累和生态系统，正在构建企业级的智能体开发与服务平台。它们提供的不是单一的智能体应用，而是一整套“授人以渔”的工具，赋能千行百业的客户构建深度契合自身业务流程的定制化智能体。这种模式的核心是平台化、生态化和深度集成。

1.2.4 以实体行业为代表的系统性解决方案

智能体应用的最高阶形态不再局限于单个智能体的任务执行，而是通过构建多智能体系统，对庞大、动态的复杂系统进行整体性优化与管理。该模式常见于工业制造、智慧城市、通信网络等领域，通常由掌握关键基础设施和拥有大规模系统集成能力的国家级企业或行业巨头主导。

(二) 内部协同：从“单兵作战”到“军团协同”

2.1 多智能体系统正成为企业级智能体的主要形态

随着应用场景的复杂化，单一智能体已难以满足部分企业级的综合业务需求。现实世界中的复杂任务，例如智能物流、智能城市管理和智能工厂等，需要多个具备不同专业能力的智能体协同工作。相比于 AaaS（智能体即服务），MAaaS（多智能体即服务）提供了更加复杂和灵活的智能体应用，允许多个智能体协同工作完成复杂任务。例如，在智能物流领域，多个智能体可以协同工作，优化仓储管理、运输路线和库存控制等环节。MAaaS 能够提高任务执行的效率和精确度，特别适用于需要多智能体协作的行业。通过多智能体的协作，工厂能够实现智能化生产，减少人工干预，提高生产效率。

数据显示，“AI 领航杯”大赛中 43.58% 的智能体项目采用了多智能体系统的解决方案，这表明多智能体协同正成为智能体应用的主要形态，特别是在复杂场景下，多智能体协同工作能够更好地处理分布式任务和协作问题。这种从

“单兵作战”向“军团协同”的模式转变，将挑战从构建单个智能体升级为管理一个高效的多智能体系统（MAS）。

2.2 MAS 存在复杂的协同难题

构建和管理一个高效的多智能体系统，意味着必须解决其内在的四大协同难题。随着智能体数量的增加，它们之间的潜在交互关系呈指数级增长。首先，如何设计高效的通信协议和协作机制，确保信息能够准确、及时地在智能体之间流转，成为第一难题。其次，不同智能体可能拥有各自的局部目标，而这些目标之间可能存在冲突。例如，在智能电网中，一个负责最大化发电效率的智能体，可能与一个负责保障电网稳定性的智能体产生目标冲突。系统必须具备有效的冲突消解和协商机制。另外，在动态变化的环境中，如何将复杂的总任务动态地分解并分配给最合适的智能体，以及如何在某个智能体失败后进行任务的重新调度，是确保系统鲁棒性的关键。最后，系统也需要考虑如何客观评估每个智能体在集体任务中的贡献，并设计合理的激励机制以促进协作而非竞争，是实现系统整体目标最优化的核心。

为了应对多智能体系统协同带来的挑战，业界正在探索创新的架构模式，例如交通银行构建“1+1+N”的多 Agent 智能体系统，其中一个核心智能体以知识图谱和 RAG 决策引擎为核心，负责战略决策；一个流程编排调度智能体作为枢纽，负责任务分解和调度；N 个场景化智能体集群作为触手，负责具体任务的执行。各行各业也涌现出不同的 MAS 构建方案，尝试解决多智能体协同问题，推动智能体从“单点智能”

向“群体智能”的跃迁。

(三) 平台互通：打破“数据孤岛”与“生态壁垒”

智能体的核心在于其调用工具和访问数据的能力，但这一能力的前提是能够与外部世界进行无缝互通。然而，在现实的企业环境中，智能体面临着由“数据孤岛”和“应用壁垒”构成的巨大屏障。

3.1 企业内部异构系统兼容性问题

企业级智能体的构建和应用，并非对数字化转型方案的简单升级，而是系统性重构。尽管部分大型企业凭借资金与技术优势，率先完成数字化基建，但在接入大模型能力以及构建智能体的过程中依旧困难重重。例如，企业内部往往存在大量不同时期、由不同供应商开发的异构系统（如 ERP、CRM、MES），这些系统通常使用不同的数据格式、API 协议和认证机制。让智能体能够理解并调用这些多样化的系统，是一项艰巨的集成工程。有统计显示，高达 86% 的组织需要进行基础设施升级才能有效支持 AI 智能体。要让智能体从一个封闭的“大脑”变成一个能干的“行动者”，就必须解决跨系统互通这一基础性难题。

3.2 与外部工具存在通信或生态壁垒

智能体与外部工具之间的通信缺乏统一标准，导致每个集成项目都需要进行定制化开发，成本高昂且难以扩展。为了构建一个灵活的“积木式”生态，行业亟需标准化的通信协议。目前，业界正在探索如模型上下文协议（Model Context Protocol, MCP）和智能体对智能体协议（Agent-to-Agent, A2A）

等方案，以解决兼容性问题。MCP 侧重于为智能体提供对工具、API 和数据的结构化访问，而 A2A 则更侧重于实现智能体之间的点对点协调、任务共享和协商。

另外，智能体在快速发展并尝试融入各行各业时，还面临着显著的生态壁垒。各大平台例如电商、社交等应用构建起了“围墙花园”，数据接口不开放，智能体难以获取跨平台的完整数据以执行复杂任务。同时，各平台间存在商业利益与信任壁垒，企业将用户数据和行为数据视为核心资产，不愿共享，用户对让智能体直接操作账户进行交易存在安全和隐私担忧。这些壁垒不仅限制了智能体能力的发挥，也影响了整个产业协同发展的效率。

(四) 数实融合：跨越数字产业与实体经济

智能体与实体经济的深度融合，是“人工智能+”战略落地的核心，也是智能体创造最大价值的主战场。交通、能源、工业制造等实体企业，拥有海量的行业数据和复杂的物理世界交互场景，是智能体技术应用的“沃土”。然而，数实融合的过程也面临着层层阻碍。

智能体的核心能力是“自主执行任务”。如果企业数字化转型不彻底，所依赖的业务流程本身是割裂的、非标的，数据是孤立的、不一致的，那么智能体非但不能提升效率，反而会成为新的“技术债”。智能体的规模化会成为企业数字化转型的障碍，因此，实体企业数字化转型与智能体应用须协同规划、协同建设。

4.1 数据质量不高制约智能体的广泛应用

实体行业可用的高质量、有效标注数据稀缺。现有数据大多来源多样、格式不一，可能涵盖生产日志、传感器读数、视频监控、甚至纸质文档。大量设备数据因传感器缺失、接口陈旧或仍依赖纸质记录而未被数字化。这些质量参差不齐、未被标注的数据难以直接用于构建高精度的智能体构建。同时，通用模型缺乏融合了行业知识（Know-how）和业务原理的“颗粒度更细”的专业数据集，使得智能体难以理解并可靠执行特定行业的复杂场景任务。

数据基础的薄弱，直接导致智能体在实体经济场景中“难用”——技术“通而不专”与业务“两张皮”的情况难以避免。许多先进的通用大模型在进入具体工业领域时，由于缺乏高质量的行业数据有效供给，难以深入理解特定的工艺逻辑和“老师傅”的隐性经验，出现技术与业务脱节的现象。但是，在工业生产中，一个微小的误判可能导致重大安全事故或质量缺陷。由于高质量行业数据稀缺，智能体在复杂、非标工况下的决策准确性和可靠性面临挑战，导致企业“不敢用”。

4.2 收益不确定性限制智能体的规模化落地

对于大型企业而言，生产场景复杂且极具差异，即便同一条产线，生产不同型号产品时所需的数据和参数也完全不同。这种碎片化导致为特定场景开发的智能体解决方案难以简单复制到其他场景，每一次适配都意味着巨大的数据清洗、标注和模型再训练成本，使得智能体的规模化落地变得困难。

对于中小企业而言，面临数据治理和高性能算力投入成

本高、投资回报周期不确定的困境，缺乏进行大规模数据基础建设的意愿和能力。尽管行业提出了“轻量化+模块化”的 AI 数据服务与智能体解决方案来降低中小企业使用门槛，也在探索基于云服务或公共平台的模型托管、数据治理服务等商业模式，可以帮助中小企业以更低成本和更灵活的方式获取所需的数据能力和 AI 服务。但是，许多企业对于路径和收益的不清晰，加之数据安全、技术适配性等顾虑，仍然“不敢用、不愿用、不会用”，导致智能体在与实体经济融合过程中困难重重。

(五) 算力支撑：从集中到分布的基础设施需求

5.1 算力从集中式向分布式协同演进

智能体的大规模部署对底层算力基础设施提出了新的要求。大模型的训练阶段，需要依赖于由数千甚至数万个 GPU 组成的超大规模智算中心，这是一种典型的集中式计算模式。然而，当智能体进入应用服务阶段，其运行场景将变得高度分散和无处不在，数以亿计的智能体将运行在各种环境中，这要求构建一个云、边、端协同的分布式算力网络。这种“训练在云端，推理无处不在”的模式，要求我们科学布局算力、模型、数据等基础设施的协同建设。需要以发展智算云服务为抓手，构建一个云、边、端协同的分布式算力网络，提高智能基础设施的共享服务水平，以支持未来智能体互联网的互联互通。

云端负责模型训练和知识更新，而边缘和终端则成为智能体感知世界、服务用户的“神经末梢”。一部分智能体将

部署在公有云或私有云数据中心，处理高并发的公共服务或企业内部的敏感业务数据。在“AI领航杯”的参赛项目中，云端部署占比高达45.2%，基于企业对数据安全定制化的要求，私有化部署占32.8%。另一部分智能体将部署在靠近数据源的边缘服务器上，如工厂车间、变电站、交通路口等，以满足低延迟和高可靠性的要求。还有大量智能体将直接运行在终端设备上，如智能手机、自动驾驶汽车、智能家居设备等，以保证极致的低延迟、离线运行能力和用户数据隐私。

5.2 网络设施和运维管理成为难点

在分布式智能体时代，系统的整体性能不再仅仅由单个计算节点的处理速度决定，而是越来越受限于连接这些节点的网络的物理特性——数据传输的稳定性、时延控制能力等核心指标成为关键制约因素。在分布式系统中，数据和指令需要在地理上分散的节点间长途跋涉，任何网络拥塞或连接缓慢都会造成严重的时延。同时，部署、监控和保护一个大规模、异构、地理分散的设备集群，其运维管理的复杂性也带来了巨大的成本投入和工程工作量。大量的调研表明，企业对大模型、智能体应用到生产实际的痛点是数据出域导致的风险，因此算力及网络的合理布局是关键。智能体应用的未来图景是数以亿计的智能体运行在从云数据中心到工厂车间、自动驾驶汽车、乃至个人智能手机的每一个角落，在现实世界中必须直面物理定律的严苛约束和运维管理的巨大复杂性。

(六) 生态构建：全产业链的深度合作必不可少

释放智能体的全部潜力，绝非单一技术或单一企业能够完成，它高度依赖于一个健康、开放、协同的全产业链生态体系。从芯片制造商、云服务商，到模型开发商、应用开发商，再到最终的用户，每一个环节都至关重要。生态体系的建设是保障智能体创新和应用落地的根本。然而，当前智能体生态的构建正面临两大系统性障碍：一方面，商业竞争驱动各大平台构建封闭的“围墙花园”；另一方面，跨企业协作缺乏法律保护导致信任基础薄弱。这使得实现智能体跨区域协作、形成真正“价值网络”困难重重。

6.1 大型平台“围墙花园”难以打破

智能体生态的发展由于企业利益竞争关系正陷入困境。从整个生态的集体利益出发，行业亟需开放、统一的标准，以促进不同智能体间的互联互通，释放网络效应。然而，从企业的战略视角看，构建封闭的生态系统更有利于形成竞争优势。大型互联网科技平台（例如淘宝、百度、抖音）的核心竞争力，源于其拥有海量的、高质量的专用户数据。这些数据是训练和优化智能体的“燃料”，是它们最宝贵的资产，通过控制专有数据和锁定用户来建立坚固的商业“护城河”。这种个体与集体利益的矛盾冲突，是阻碍开放生态形成的最根本原因。

6.2 协议标准化存在系统性障碍

即使企业有开放的意愿，技术标准化的滞后也使得大规模协作难以实现。当前的智能体生态，正面临着发现机制缺失和通信协议碎片化的双重困境。在智能体能够相互通信之

前，它们必须先能“找到”彼此。然而，目前全球范围内尚不存在一个像互联网 DNS 域名系统那样，能够为数十亿智能体提供注册、发现和地址解析的公共基础设施。同时，业界正在积极探索适用于智能体网络的协议标准。其中，模型上下文协议（MCP）和智能体对智能体协议（A2A）是两个备受关注的通用协议，阶段性的解决了部分通信问题。但是，各大科技巨头纷纷推出自家的智能体通信协议，导致了严重的市场碎片化。开发者被迫在不同生态系统之间做出选择，或者投入大量精力开发脆弱的、定制化的“适配器”来连接不同的系统，这极大地阻碍了跨平台、跨厂商的智能体协作。

为了加速智能体在千行百业的应用，需要共同构建一个开放、协作、共赢的智能体生态系统。如何化解生态合作困境，推动开放标准和开源生态的建立，促进产业链资源的优化整合，是当务之急。数据分析显示，“AI 领航杯”的参赛项目中，80.95%支持轻量化 API 调用，71.03%支持插件化扩展，开放和协作已成为行业初步共识。但是，建立繁荣的技术生态，需要整个产业链的深度合作与共同努力，从底层技术标准的制定，到上层应用场景的协同，缺一不可。

第四章：推进智能体应用落地的市场研究

当智能体技术走出实验室，如何在复杂多变的应用场景中实现规模化落地，成为决定产业变革成败的关键命题。今年以来，中国互联网协会组织了一系列的研讨和调研，尝试为智能体应用推进谋篇布局，总结了智能体应用落地的五大战略思路：异中求同，在碎片化需求中提炼共性，探索可复制的批量化实施路径与工具链；点面结合，单点建设关键枢纽节点，攻克算力网络的部署难题，构建分布式智能基础设施；由内到外，优先在数据主权充分、可控性强的业务场景深耕细作，积累核心能力后向外拓展；由虚向实，从界面成熟的数字应用逐步向场景复杂的物理世界渗透，实现虚实融合；因势利导，充分发挥体制与政策优势，为创新营造有利环境。通过五步推进思路的研究总结，旨在为行业提供一套可操作、可演进、可持续的智能体应用战略框架。

当前，我国智能体的应用尚处于起步阶段，规模化推广面临着多重挑战：市场碎片化特征明显、数字化技术的兼容性差、终端成熟度不一、以及各领域推进体制强弱不等。其中，行业需求的高度碎片化与规模化推广所需的标准化之间的矛盾尤为显著。千行百业的应用需求呈现出个性化、碎片化的特征，以现有的大模型能力，开发一个万能的智能体难以实现。智能体需要“术业有专攻”，通常基于特定的行业领域，定制能够完成单一的任务的智能体，这类模式更易落地应用，且能更好的融入业务流程。但是不同企业在业务流

程、数据架构、组织文化和决策机制上的巨大差异，导致智能体难以像传统软件产品一样实现“即插即用”。若为每个场景定制开发，不仅研发成本高昂、周期漫长，更难形成可复制的商业模式，严重制约智能体经济的规模化发展。

（一）异中求同，探索批量化路径和工具

首先，智能体规模化落地的突破口是推进“异中求同”的策略。从纷繁复杂的个性化需求中，识别出跨行业、跨企业普遍存在的高频、通用、流程标准化的共性任务，优先构建可复用、可组合、可扩展的智能体“能力基座”，实现规模化验证与快速部署，再以此为基础向高价值个性化场景延伸形成“共性打底、个性增值”的渐进式演进路径。

在互联网经济时代，由于存储和分销成本的急剧下降，大量非热门、小众商品（即“长尾”）的累积市场份额可以超过少数畅销商品（即“头部”）的市场份额。这一理论颠覆了传统的“爆款”思维，展示了小众市场的巨大商业价值。将长尾模式映射到智能体应用市场，为应对个性化需求，应摒弃“一刀切”的智能体产品设计思路，与其倾尽资源打造一款覆盖所有场景的“爆款”超级智能体，不如构建一个开放、灵活、可扩展的模块化智能体开发平台，将智能体的核心能力解耦为标准化的功能组件，形成类似“乐高积木”的能力单元库。

智能体的底层交互模式、数据调用逻辑和任务执行流程等能力存在着高度的共性。例如，无论是电商领域的导购智

能体，还是旅游领域的行程规划智能体，都可能需要调用用户认证、支付、地理位置查询、日历同步等基础功能。将跨场景的共性能力进行抽象、沉淀，开发成标准化的、可复用的软件模块或服务接口，形成“能力积木”。开发者在构建新的智能体时，无需再从零开始“造轮子”，而是可以像“叠乐高”一样，通过拖拽、配置和简单的逻辑编排，快速调用、组合这些“能力积木”，从而以一种工业化、低成本的方式，高效地搭建出满足特定长尾市场需求的智能体应用。这种模式解决了完全定制化开发的成本困境，实现了规模化生产与个性化需求的统一。

目前模型厂商纷纷推出基于自研大模型的开发平台，例如阿里巴巴百炼、腾讯元器、字节跳动扣子等。同时，开源平台也发展迅猛，例如谷歌 Vertex AI、LangChain、Dify、FastGPT 等，这些备受行业关注的开发平台正是“能力积木”理念的商业化体现。平台通过提供一系列预构建的工具、模板和无代码/低代码开发环境，使企业，特别是缺乏专业 AI 技术团队的中小企业，能够像订阅云服务一样，快速部署用于特定业务场景的智能体。例如，通过智能体开发平台，企业可以轻松构建用于自动化客户服务、营销传播的智能体，而无需从零开始开发底层技术。从开发单一应用转向构建可组合服务生态的思路，代表了智能体时代“异中求同”开发范式的变革。繁荣的智能体生态系统，其核心竞争力将不再是拥有一个强大的超级智能体，而是拥有一个丰富便捷的“能力积木”库。

（二）点面结合，解决算力网络部署难题

随着智能体的规模化应用，算力网络等基础设施的部署将是巨大的考验。数以亿计的智能体若投入使用，将产生前所未有的、持续性的、分布式的算力需求，传统的中心化云计算架构在延迟、成本和数据主权方面将难以为继。要支撑智能体时代的到来，必须解决算力及网络的合理布局问题，采用“点面结合”的方式构建新一代智能算力网络。

智能体的算力需求具有三大特征。首先是高频推理，不同于以“训练”为中心的传统AI，智能体应用以“推理”为中心。它们需要7x24小时在线，持续感知环境、进行决策，这将产生海量的、常态化的推理负载；其次是低延迟，在“由虚向实”的应用中，如工业控制、自动驾驶、远程医疗等，智能体必须在毫秒级别做出反应，任何网络延迟都可能导致严重后果；最后是算力迁移，尤其在政务、金融、医疗等“主权充分”的场景中，海量的原始数据因其敏感性、保密性或体量巨大而无法轻易离开本地。为避免数据跨域传输的风险与成本，通过“算力迁移”替代“数据流动”，这要求算力必须向数据侧移动，而非数据向算力中心汇聚。

为应对这些挑战，一个“点面结合”的算力网络部署策略势在必行。“点”是指建设国家枢纽节点，打造高算力战略中心，例如国家级的、集约化的超大规模智能算力中心。这与我们建设“算力网络国家枢纽节点”的国家战略相一致。这些“点”如同国家算力电网的“大型发电厂”，集中部署最先进的计算集群，专注于承担最繁重的算力任务。其核心功能包括

大规模模型训练、复杂批量推理、算力储备与调度等。以贵州为例，其作为全国首个国家大数据综合试验区，已建成算力网络国家枢纽节点，综合算力位居全国第一方阵，这为“点”的建设提供了坚实的基础。“面”是指构建泛在边缘网络，形成低延迟、高可靠的算力矩阵。它如同算力电网的“配电网”和“变电站”，由遍布在基站、园区、工厂、乃至终端设备侧的边缘计算节点构成。其核心功能是满足低延迟需求、解决数据出域问题等。

“点面结合”的核心在于智能协同。它不是“点”和“面”的简单叠加，而是通过一个强大的算力网络操作系统，将中心化的“点”与分布式的“面”智能地连接起来，形成一个统一、高效的“算力一体化网络”。这个系统能够根据智能体任务的特定需求——如延迟、成本、带宽、数据位置等——进行智能路由和动态调度。一个高延迟、高通量的训练任务会被自动分配到远端的“点”（国家枢纽），而一个低延迟、高实时的工业控制任务则会立刻在本地的“面”（边缘节点）上执行。这种“点面结合”、合理布局的算力网络，是支撑未来亿万智能体落地的唯一可行路径。

（三）由内到外，从主权充分的场景做起

智能体的在大型企业的应用落地涉及复杂的流程整合与数据贯通，既包括机构内部的系统打通，也涉及与外部合作伙伴的互联互通，是一项复杂的工程。面对这种复杂性，最务实且风险最低的推广对策是“由内到外”，即首先从企业内部能够创造显著价值、且组织拥有充分数据和流程“主权”

的场景切入。拥有更高的自主权、更低的管理协调成本和更可控的数据环境，这为智能体应用提供了一个理想的“沙盒”环境。允许企业在风险可控的前提下，通过试点项目快速验证技术的有效性和商业回报，并积累宝贵的开发、部署与运维经验。

“内部试点”的策略在业界已得到广泛推进。通过对“AI领航杯”参赛项目的统计分析显示，接近四分之三的项目选择先进行内部试点，待标杆案例打造完成后，再进入行业复制或多场景拓展阶段。在大型央国企中，这一趋势更为明显，超过三分之一（35.7%）的智能体方案首先服务于内部需求，待模式成熟后再向外部市场输出。选择合适的内部试点场景至关重要。理想的切入点应具备“高频、高需、高收益”的特征，即业务流程相对高频、自动化需求迫切，且智能化改造能带来显著的效率提升或成本节约。企业内部流程链条相对较短、数据和系统相对集中、且能带来显著效益的场景被认为是高价值场景，例如内部 IT 运维系统、设备故障预警、企业经营分析、人力资源管理等领域，目前已涌现出大量具有显著量化成果的成功案例。

这些内部试点项目所带来的价值远不止于直接的成本节约或效率提升，本身也在推动企业进行数据治理、流程优化和跨部门协同，从而在根本上提升了整个组织的数字化成熟度。因此，“由内到外”不仅是一种低风险的推广策略，更是企业迈向全面智能化所必需的、打牢基础的关键一步。

（四）由虚向实，逐步从数字向物理渗透

在人工智能的应用体系中，如果说大模型是“大脑”，那么智能体就是连接“大脑”与现实世界、执行“最后一公里”任务的“神经系统”。一个功能完备的智能体不仅需要聪明的“大脑”和敏捷的“神经”，还需要形态各异的“手脚”——即终端设备。然而，不同形态的终端（如纯软件、虚拟数字人、物理机器人）其技术成熟度和应用条件存在显著差异。因此，智能体的推广应用应遵循一条“由虚到实”演进路径，即从当前技术最成熟的数字虚拟世界起步，随着机器人和具身智能技术的不断突破，再逐步渗透到复杂的物理世界。

第一阶段：聚焦虚拟世界，从数字服务切入。当前，大模型最成熟的能力仍然是基于语言的“问答交互能力”，与之适配度最高的终端是电脑、电视、手机、平板等显示设备。智能体应用的初期阶段，主要从**数字客服**领域切入，能够最大化发挥大模型的语言服务能力。通过部署能够 7x24 小时在线、理解复杂自然语言查询、并能跨系统执行操作（如查询订单、修改账户信息、办理业务）的“超级客服”智能体，企业可以极大地提升客户满意度，同时显著降低人力运营成本。同时，结合虚拟数字人技术，可以推出形象生动、交互自然的**数字人智能体**，将强大的会话 AI 能力与生动的虚拟数字人形象相结合。这些形象可亲、交互自然的智能体，正在被广泛应用于政务服务大厅、银行网点、企业展厅、线上直播等场景，提供引导、咨询、业务办理乃至情感陪伴等服务，极大地优化了用户体验。

第二阶段：迈向物理世界，拥抱具身智能。随着计算机视觉、传感器融合、运动控制以及强化学习等技术的飞速发展，人工智能正从数字领域大步迈入物理世界。特别是以人形机器人为代表的具身智能技术，正迎来其商业化的前夜。行业预测显示，具身智能市场有望在 2025 年左右进入快速膨胀期，这将成为智能体“由虚到实”演进的关键拐点。AI 正从数字领域步入物理世界，重塑工业、物流、家庭服务等多个领域。企业应密切关注这些技术进展，在条件成熟时，逐步在仓储物流、生产制造、巡检安防等领域，试点和推广能够执行物理任务的具身智能体，最终实现数字智能与实体经济的深度融合。

“由虚到实”的演进路径不仅是技术成熟度的自然体现，更是一条战略性的数据收集与模型训练的有效路径。在第一阶段，大规模部署的虚拟客服和数字人智能体，与数以亿计的用户进行交互，产生了海量的、关于人类语言、意图和解决问题流程的真实世界数据。这些宝贵的数据，恰恰是训练第二阶段物理智能体如何理解人类指令、如何与人类协作的“教科书”。一个成功部署了虚拟客服体系的企业，无形中已经为其未来的工厂机器人或仓库机器人积累了最核心的训练资产。这使得两个看似独立的阶段之间形成了一条强大的战略纽带，率先在虚拟世界取得成功的组织，将在迈向物理世界的竞争中获得显著的先发优势。

（五）因势利导，充分释放体制与政策优势

智能体技术的全面推广，不仅依赖技术与市场的驱动力，

更需要发挥特定领域的体制机制、产业政策和社会需求的引导作用。通过“因势利导”，选择那些基础好、需求强、具有体制优势的重点领域进行集中突破，打造一批具有影响力的“灯塔项目”。这些标杆案例的成功，不仅能解决关键领域的迫切问题，更能形成强大的示范效应，为智能体技术在千行百业的广泛渗透和深度应用提供可复制的模式与宝贵的经验，从而带动全局发展。

政务服务：“一网通办”驱动下的政务智能体创新

我国在数字政府建设，特别是“一网通办”改革方面，形成了强大的顶层设计和推进体制，这为政务数据的互联互通和流程标准化奠定了坚实基础。这种全国范围内统一规划、协同推进的体制优势，为政务智能体的规模化应用提供了得天独厚的“土壤”。各地政府把握我国统一大市场建设和“一网通办”深入推进的利好条件，大力发展政务智能体。目标是让政务服务从“能办”走向“好办、智办”，实现服务的智能化、精准化和个性化，不断满足企业和民众的新需求。在政策咨询层面，可基于用户画像与需求标签，提供千人千面的政策解读、精准匹配与智能推送；在业务办理层面，可引导用户完成在线申报，并自动进行材料预审；在城市治理层面，可深度融入智能交通调度、公共安全预警、环境质量监测、应急事件处置等场景，助力提升城市治理的精细化、智能化水平。

医疗健康：迈向个性化、精准化的“AI私人医生”

医疗健康关系国计民生,是国家“人工智能+”行动的重中之重。我国拥有庞大的人口基数和丰富的医疗数据资源,为训练精准的医疗大模型和智能体提供了基础。智能体的应用将推动医疗健康服务从“信息化”向更高效、精准、人性化的“智慧化”阶段跃升。医院都可以通过建设高质量的医疗健康数据中心和专属大模型,开发运营覆盖诊疗、管理、医保等全流程的智能体,提升运营效率。同时,医生和护士都将配备 AI 助理,提供智能辅助诊疗、医学影像辅助诊断、日常护理协助等服务,从而将医护人员从重复性工作中解放出来。例如浙江大学医学院附属邵逸夫医院的智慧医疗解决方案,基于大语言模型的 AI 导诊分诊系统将首诊准确率提升至 97%,有效解决患者“挂错号”难题;智能预问诊 5 秒生成结构化病历,较传统模式效率提升 40 倍,年均可释放医护生产力超 10 万小时,使其聚焦高价值诊疗行为。“AI 私人医生”,提供覆盖全生命周期的个性化健康管理服务,整合用户的健康数据、可穿戴设备信息,提供疾病预防、慢病管理、康复指导等服务,成为人人可及的高水平健康助手。例如郑州大学医学人工智能团队研发的心血管健康多模态平台,实现可穿戴、康养、运动等五大类硬件设备接入并构建包含 7 大类、78 项健康指标的个人健康云档案。通过数据链路打通,助力心梗患者赢取 90 分钟黄金救援时间,目前已成功处置中高危患者超过 1.5 万人。

工业制造：加速“两化融合”的智能制造新范式

工业是实体经济的主战场，也是智能体应用价值最直接的领域。我国作为制造业大国，拥有全球最完整的工业体系和丰富的应用场景。国家大力推进的“两化融合”战略，为工业智能体的应用提供了强大的政策支持。产业应充分发挥“工信体制优势”和制造业大国的产业基础，积极推进智能体在“两化融合”领域的应用优势，打造智能制造新范式。

工业智能体正赋能产品从设计、生产到运维的全生命周期，其应用价值直接体现在降本增效和质量提升上。一系列面向工业制造的大模型和智能体正在涌现，推动 AI 在产品设计、中试、生产、服务、运营等全环节落地应用。另外，工业智能体的应用能带来了实实在在的经济效益，避免不必要的损失。联通数据智能公司联合中国船舶集团以元景工业万悟智能体开发平台为核心底座，打造多个智能体。例如，结构翻模设计智能体，通过元景多模态、视觉大模型对 2D 设计图纸进行参数和设计要素识别，输出 3D 建模所需的数据要素，通过匹配相应建模语言，通过专业 CAD 软件重新构建全船精确的 3D 数字模型，精确还原船体结构等所有要素的空间位置与几何关系，节省 40%设计人力，效率提升 30%。

基于“因势利导”的推进对策，通过在公共属性强、监管明确的领域率先垂范，不仅解决了关键的社会问题，也为整个智能体产业的技术标准、数据规范和商业模式探索了道路，从而为后续在更广泛的商业市场中的应用铺平了道路，形成了一个正向循环的“飞轮效应”。

第五章：构建智能体生态的对策建议

展望未来，构建开放繁荣的生态体系已成为决定智能体长远发展的核心命题。这不仅关乎技术突破，更是涉及商业创新、标准建设与安全治理的全方位体系构建。近期，中国互联网协会组织多次研讨，综合各方观点，提出了五大对策建议：一是推进开放合作，打破壁垒激活多方共创；二是探索商业模式，创新价值分配激励持续投入；三是引领标准制定，以通信协议与接口规范夯实互联互通底座；四是推动能力评估，建立多元化标尺引导健康竞争；五是建设安全治理，依托标识体系守住可信运行底线。五大对策环环相扣，涵盖机制创新、技术规范、市场培育与风险防控全链条，旨在为产业各方提供一套系统性、可操作的生态构建行动框架。未来，只有形成开放协同、标准统一、安全可信的繁荣生态，才能释放智能体的网络倍增效应。

（一）积极推进开源开放的生态环境建设

繁荣的智能体生态是开源创新与开放合作两大核心要素深度融合、协同发力的结果。智能体技术的未来走向，不仅取决于算法模型的持续精进、算力基础设施的迭代升级，更依赖于一个繁荣有序、安全可信、协同共生的生态系统的成功构建。智能体生态走向成熟，需要政府、高校、科研机构、企业及开源社区等产业各方秉持开放包容、合作共赢的态度，共同开启智能体产业发展的新篇章。

首先，激励开源与创新，激活生态发展的内生动力。开源是推动智能体技术快速迭代、降低产业准入门槛、加速技术普及落地的关键力量，更是生态创新活力的重要源泉。当前，全球智能体领域的技术竞争，已逐步延伸至开源生态的布局与争夺，一套成熟的开源体系，能够快速聚集全球开发者资源，形成“技术共享—协同创新—成果复用”的良性循环。为此，需从政策引导、资源支持、机制保障三个层面发力：在政策引导上，政府可出台专项扶持政策，将开源贡献纳入企业技术创新评价体系，鼓励企业、科研机构将自主研发的优秀模型、工具组件、数据集及智能体开发框架向社会开源；在资源支持上，设立人工智能开源专项基金，为开源社区提供资金支持，搭建开源技术测试平台，降低开发者的创新成本；在机制保障上，建立健全开源贡献评价与激励机制，通过揭榜挂帅、技术认证、商业合作优先等方式，激发开发者的开源热情。同时，通过举办智能体创新竞赛、场景应用挑战赛等活动，赋能中小企业与初创团队，鼓励其聚焦细分领域的“小而美”场景（如中小商户智能运营助手、社区医疗辅助诊断 Agent 等）开发创新应用，形成“头部企业引领开源、中小企业深耕场景”的百花齐放的产业创新格局。

另外，深化开放合作，拓展生态发展的全球格局。智能体技术的通用性与渗透性，决定了其生态构建必然是一项全球性事业，任何国家或地区都难以仅凭自身力量完成全链条

布局。在全球化背景下，唯有打破地域壁垒、行业壁垒与技术壁垒，深化国际交流与合作，才能推动智能体生态实现高质量发展。一方面，需积极参与并主动引领智能体相关国际标准的制定，在通信协议接口、安全治理规则、能力评估基准等关键领域，发出中国声音、贡献中国方案，推动形成兼顾安全性与包容性的全球统一标准体系；另一方面，加强与全球领先研究机构、科技企业的深度合作，围绕大模型“幻觉”治理、多智能体协同、跨语言跨文化交互等共性技术难题开展联合攻关，共享研究成果与技术经验。更重要的是，构建面向全球开放的开源技术体系与社区生态，通过优化开源社区的国际化运营机制、提供多语言技术支持等方式，吸引全球开发者参与中国智能体生态的建设，最终形成“技术共享无边界、价值共创促发展、责任共担保安全”的全球智能体生态系统，让智能体技术跨越行业、连接全球，为人类社会的进步与发展贡献力量。

开源开放是智能体生态从“单点突破”走向“全局繁荣”的必由之路，更是智能体技术从“实验室”走向“产业界”的核心保障。持续推动智能体生态突破技术瓶颈、完善产业体系、拓展应用边界，真正实现“技术赋能千行百业、生态助力全球发展”的目标，为经济社会高质量发展注入源源不断的智能动力。

（二）构建“智能体互联网”的协议与标准

构建可互操作、可扩展的智能体经济，其核心前提是实现智能体的互联互通。随着智能体应用向纵深发展，仅仅追求单个智能体能力的极致突破，已不足以应对日益复杂的现实世界任务。未来智能体应用的真正潜力在于如何将无数个不同领域的智能体互联互通，形成一个能够协作共生、涌现出群体智能的协同网络。然而，当前主流网络协议已无法满足智能体互联互通的核心需求。传统的互联网协议（如 HTTP）是为用户的“浏览-点击”交互模式和“用户驱动、请求-响应”的无状态设计而构建的。对于需要大规模、高频、自主协作的智能体而言，这些协议存在诸多天然障碍，例如协作成本高昂，语义标准缺失等。构建一个互联互通的“智能体互联网”，必须优先建立统一的网络基础设施和标准化的接口协议。这要求我们突破当前以用户为中心的网络协议局限，合力打造一套专为智能体间及智能体与工具间的通信标准。

构建智能体互联网的“通用语言”不仅是技术升级，更是关乎未来智能经济“通用语言”定义权与生态系统控制权的战略竞争。现阶段，有两方面的关键标准需要重点关注。一是统一的上下文数据结构标准，确保变量值、执行历史、临时文件引用等状态信息传输无丢失、无冲突；二是完善的版本管理标准，适配智能体与工具的快速迭代需求。一套高效的智能体通信协议，可以显著提升任务可追溯性与可迁移性，支持用户随时接管任务，或转移给更专业的智能体，进一步

提升智能体应用的灵活性。

国家战略层面，我国已将智能体及其协同技术的标准化提上议程。最新发布的《国家人工智能产业综合标准化体系建设指南（2024版）》，明确将“智能体”和“群体智能”纳入关键技术标准体系，并对多智能体分布式一致性等关键技术规范提出了要求。这为“智能体互联网”的实现提供了坚实的政策保障和标准基础。在产业实践中，全球统一的“智能体互联网”仍在酝酿，但国内领先企业已开始积极探索。例如，阿里巴巴百炼平台为企业级用户提供了多智能体编排能力，允许开发者用代码灵活定义智能体之间的“工作流程”（如顺序、并行、路由）。腾讯云平台的“Multi-Agent 模式”则将更多主动权交给大模型，使其自主规划并支持“自由切换”和“ workflow 编排”。

目前，引领通信协议与接口标准化，需依靠头部企业、开源社区、行业协会三大力量协同推进，形成“创新-验证-落地”的完整闭环。头部企业凭借市场影响力与技术积累，推广内部成熟的 Agent 框架、工具描述格式，快速形成“事实上的行业标准”，降低行业落地门槛。开源社区以敏捷开发与协作优势，推动通信协议、工具描述规范的快速创新与验证，例如围绕 Agent-to-Agent 通信的行为语言（ACL）探索，为标准升级提供技术灵感。行业协会与联盟可以汇聚企业与社区的创新成果，转化为正式、可信赖的行业标准与国

家标准，同时融入安全合规要求，确保标准的权威性与适用性。呼吁三大主体深度协同，共同引领国内智能体通信标准体系建设，为整个智能体经济构建统一、高效、安全的“通用语言”，抢占全球智能经济生态竞争的战略制高点。

（三）探索多样化的商业合作及服务模式

活跃的智能体生态，需依托灵活多样的商业模式满足不同用户、场景需求，实现资源高效分配与生态可持续运营。当前人工智能市场已涌现三类核心商业化策略——以计算资源消耗为核心的按量付费模式、提供稳定服务与固定预算的订阅模式、与业务需求深度绑定的项目制交付模式。三者并非相互排斥，而是构成多层次、互补的商业矩阵，为供需双方提供丰富选择，共同推动智能体商业模式的创新与成熟。

按量付费模式以计算资源消耗（如按 Token 付费）为核心，成本与实际使用直接挂钩，透明度高，用户可按需付费，无需承担闲置资源成本，尤其适配初创企业、小型项目。其核心优势是灵活度高、入门门槛低；但存在成本波动性大、预算难评估的问题，且复杂推理、多步骤规划类任务易导致成本快速累积。适用场景包括智能体研发与测试、即时问答系统、高频低复杂度功能调用、流量突发性强的应用。

订阅模式是指用户支付固定周期性费用，换取稳定的智能体服务，部分方案还包含专用工具套件。该模式预算可预测性强，既能帮助服务商维持稳定现金流，也能让用户获得

持续、可靠的服务支持，适配有长期稳定需求的用户。其核心优势是预算可控、服务稳定；但灵活度较低，难以满足用户突发的大规模资源需求。适用场景包括企业内部 Agent 平台、SaaS 化智能体服务、有稳定工作负载的生产环境、预算固定的大型开发团队。

项目制交付模式以定制化解决方案为核心，将智能体技术直接转化为商业成果，深度匹配传统企业对定制化服务、行业集成的需求。其核心优势是贴合业务场景、价值转化直接；但存在前期投入高、交付周期长、评估与验收标准复杂的问题，对服务商的行业理解能力、项目管理能力要求极高，需通过严格合同约定风险。适用场景包括工业控制系统集成、金融风控定制化模型、大型企业流程再造与数字化转型项目、需深度定制及现场部署的实体经济应用。

智能体商业模式的探索，本质是技术价值与市场需求的精准匹配过程。未来，单一模式的局限性将进一步凸显，“混合模式”与“场景化创新模式”有望成为行业主流。例如为企业客户提供“订阅基础服务+按需付费增值功能+定制化深度优化”的组合方案，既保障服务稳定性，又满足灵活拓展需求。同时，随着技术成熟与生态完善，商业模式还需向“价值共创”升级，服务商需深度融入用户业务全流程，通过数据反馈迭代服务、共享价值增长，最终形成供需共赢、可持续发展的智能体商业生态，全面赋能千行百业的数字化转型。

(四) 推动形成多元化智能体能力评估体系

随着智能体应用加速落地，产业需求正从“能用”向“好用”跨越，核心关键在于建立一套客观、全面、多维度的能力评估体系。当前基础大语言模型的评估基准已相对成熟，但这类静态、以知识为中心的测试，远无法衡量智能体作为“行动者”在动态环境中的真实表现。因此，构建面向智能体特性的全新评估框架，是建立市场信任、引导技术优化、加速产业落地的核心基石。

当前基础大语言模型的评估体系已形成业界共识，核心聚焦模型“大脑”能力，即固有知识储备与推理能力。行业广泛采用 MMLU（大规模多任务语言理解）、GSM8K（数学推理）、HumanEval（代码生成）等基准测试，本质为静态、单轮问答，类似对模型进行“学术考试”；效率指标则主要关注 Token 生成速度（Token/s）与训练/推理算力消耗，难以覆盖智能体的动态行动特性。

近年来，以 AgentBench、APT Bench 为代表的新型基准测试，正引领智能体评估范式的根本性转型。人工智能行业对“智能”的定义已从“知道什么”转向“能做什么”，业界逐渐认识到，智能体能力是复杂的涌现特性，更侧重规划、工具使用、环境交互、动态决策等面向行动的技能，而非静态知识测试所能概括。

评估范式	代表性基准	评估焦点	任务性质	核心衡量指标
------	-------	------	------	--------

静态知识评估	MMLU, GSM8K, HumanEval	知识储备、静态推理、孤立技能	单轮问答、静态问题求解	准确率
动态能力评估	AgentBench, SWE-Bench, APTbench	规划、行动、环境交互、动态决策	多轮、交互式、目标导向型任务	任务成功率、执行效率、鲁棒性

为准确评估智能体的真实效能，行业亟需从静态知识问答，全面转向动态化、场景化的模拟演练。新一代评估框架必须能够还原应用场景的复杂任务与潜在风险，为开发者优化技术、企业选型决策、监管机构制定规则，提供可靠、准确的衡量标准。

（五）构建全面的标识管理和安全治理体系

随着智能体深度融入关键经济社会活动，构建坚实可信的安全基座成为其规模化落地的核心前提。这需要建立统一全面的标识管理和安全可信体系，为智能体身份验证、权限管控与行为审计提供机制保障，确保其行为可追溯、可问责，夯实社会信任根基。

统一标识管理是智能体安全治理的首要环节，需为每个智能体赋予唯一数字标识，实施标准化能力清单与权限定义，将自主性约束在预设范围以降低失控风险。2025年实施的《人工智能生成合成内容标识办法》已推动形成覆盖全链条的标识治理体系，通过显式标识与隐式标识相结合的方式，实现生成内容可识别、可溯源。智能体行为日志应接入统一

审计平台，完整记录决策链与执行结果，对高风险操作引入数字签名技术确保不可篡改，为事后追溯与问责提供技术支撑。当前头部企业已形成覆盖多模态的标识技术体系，发布相关标准实践指南，有效降低了行业实施门槛。

数据与网络安全是智能体稳定运行的关键保障。数据层面需建立智能体自优化机制，适配企业数据持续积累的需求；网络层面则要求企业重新审视安全体系布局，应对规模应用带来的新挑战。针对大模型“幻觉”导致的决策失误，业界采用“本体+智能体”架构构建数字孪生系统，注入确定性逻辑保障。《终端智能体安全 2025》白皮书提出了数据去毒、可信隐私沙箱等技术路径，为数据安全防护提供了具体方案。同时，需加强标识安全防护技术研发，防范标识信息被篡改或绕过风险，提升系统整体安全性与稳健性。

强化高风险场景管控，完善人机协同机制。例如工业、医疗等领域的智能体直接关联物理设备与人身安全，其可靠性、准确性与鲁棒性至关重要。对此，应在关键操作节点设置“人类批准”环节，构建人机协同机制，实现人类经验与智能体效能的优势互补。《AI 智能体安全治理》白皮书已归纳出智能体在任务执行等环节的安全风险，并提出 33 项针对性防治措施。在终端智能体快速渗透的背景下，还需聚焦单智能体安全与多智能体可信互连，通过安全护栏、跨设备可信连接等技术，强化高风险场景的可控性与可靠性。

智能体发展始终面临自主性、安全性与可解释性的“三难困境”，治理关键在于依据场景风险动态调整平衡点：医疗诊断类智能体优先保障安全与可解释性，娱乐类智能体可适当放宽自主性限制。这需要建立覆盖全生命周期的安全治理体系：输入阶段实施过滤约束，拦截非预期信息；运行阶段通过权限限制与安全沙箱隔离防护；工具调用阶段实行API权限校验与频率限制；合规层面嵌入法律法规与行业标准，完善日志审计功能。

安全治理是智能体走向成熟普惠、规模化应用的必经之路。当前相关法律法规和行业标准正逐步完善，智能体设计需前置考虑合法合规要求，开发过程中确保系统日志的完整性、可追溯性。通过系统的标识管理、动态的人机协同与完善的监管合规体系，将有效化解智能体安全风险，为其健康有序发展保驾护航，推动人工智能技术更好地服务经济社会发展。随着智能体逐步融入关键经济社会活动，构建坚实可信的安全基座成为其规模化部署的前提。这要求建立统一、全面的标识管理和安全可信体系，为智能体的身份验证、权限管控与行为审计提供机制保障，确保其行为可追溯、可问责，从而奠定社会信任的基础。

附录：2025 智能体创新应用案例

1.10 千伏线路状态智慧评估体系

申报单位：国网河北省电力有限公司石家庄市栾城区供电分公司

案例介绍：本项目充分利用公司现有计量、测量、采集装备，通过收集接入包括来自用采系统、D5000 系统、同源系统、一体化系统等多项数据，建立智能体专家模型，并将专家模型与国网光明大模型结合，综合利用大模型的推理生成能力以及专用模型的特定分析能力，实现大小模型协同工作，最终构建出“10 千伏线路量测问题排查”智能体，实现对 10 千伏线路量测问题的精准排查。

2.5GC “智擎”网络智能化运维体系创新与应用

申报单位：中国移动通信集团陕西有限公司

案例介绍：为应对 5G 核心网云化架构复杂、跨域协同低效等挑战，陕西移动自主研发 5GC “智擎”网络智能化运维平台。该平台构建“1-2-1-4-5”方法论，集成 GNN、大模型与 Agent 技术，突破多任务双流图嵌入等难点，并采用 GraphRAG 补全知识图谱，通过解耦式智能体框架实现资源分析、性能优化、故障诊断等五类场景的智能化运维。目前平台功能模块已 100% 嵌入陕西移动生产流程，实现网络运维效率提升 30%，全年零重大故障。经济效益显著，节约人工耗时 15 人/年，降本 386.1 万元；同时有效减少因网络故障导致的直接间接损失。该创新成果已获多项奖项与知识产权，为 5G 运维从“被动响应”向“主动预防”转型树立了行业标杆。

3.AI 安全垂域大模型赋能网络安全运营

申报单位：杭州安恒信息技术股份有限公司

案例介绍：苏州大学附属第一医院通过安全垂域大模型实现对安全告警事件的自动化分析、研判、处置 100%覆盖，日均研判近 80 万数据，准确率 96.8%，人力投入持续减少。利用安全垂域大模型进行多维度风险监测，逻辑研判网络环境风险，深度剖析从扫描探查阶段到资产破坏阶段资产失陷过程。并进行自动化人机交互，AI 解读海量安全日志、威胁事件跟踪预测、各类事件意图识别与解读、安全知识问答、风险推理演算。最终构建体系化安全生态，模块化对接既有云网数用端安全体系，形成联防联控自适应生态。建立智能化安全运营体系，借助于大数据模型、数字化引擎，结合感知、推理、决策、执行，将安全运营推向自动驾驶 APS 新高度。

4.AI 大模型驱动的城市治理智能问答实践——智慧问数

申报单位：苏州市伏泰信息科技股份有限公司

案例介绍：伏泰科技“智慧问数”智能体案例，旨在破解城市治理领域业务人员数据分析门槛高、响应慢、数据与知识脱节的痛点。该案例通过构建以 AI 大模型为核心的智能体，集成 SQL 查询、RAG 知识检索、多智能体协同决策等能力，实现了从“用工具”到“用对话”的数据使用范式转型。系统具备多模态交互、会话式问答、动态数据洞察、智能编排等核心功能，在徐州、武汉、蓝德等多个项目中落地，

具体应用于多模态业务查询（如语音询问车辆位置、工单进度）、收运路线优化、企业画像构建与风险预警、视频内容智能研判及知识库问答等场景，有效支撑了环卫、市政等业务的日常监管与战略决策，最终实现了数据分析的平民化、实时化和智能化。

5.AI 赋能 康养共生，轻松健康服务创新实践

申报单位：北京轻松健康网络科技有限公司

案例介绍：轻松健康集团以 Aicare 技术栈为基础，将人工智能、大数据、云计算、知识图谱、自然语言处理、计算机视觉等多项技术进行高强度集成，形成覆盖保险与健康服务全链条的智能化平台。这不仅是单一功能的集合，而是一个深度融合、高度协同的数字化服务生态，在统一架构下支撑智能核保、理赔、健康评估、慢病干预、精准营销等多类应用场景。轻松健康集团构建了一个覆盖保险与健康服务全流程的智能化平台，实现了从健康风险评估、智能核保、智能理赔，到慢病管理、心理健康干预、精准营销的全链条服务升级。通过高集成度、多模态的技术体系，不仅实现了保险服务的自动化、智能化与个性化，还打通了保险与健康管理的壁垒，重构了保障服务的价值链条。

6.AI 建筑师智能体赋能行业“智惠”升级

申报单位：中国电信股份有限公司大客户事业部

案例介绍：AI 建筑师智能体针对建筑行业规划可研低效、设计合规查询不便、施工监管粗放、运维响应滞后等痛点研发，其采用基于 Agent 的 ReAct 理念与 LangChain 架构，构建覆盖规划、设计、施工、运维全生命周期的智能体系，实现统一入口调度与任务编排。研发中突破多模态数据整合、行业专属算法优化等难点，规划阶段实现可研报告小时级生成，设计阶段提升绘图效率 80%，施工阶段可秒级识别 650 种违规行为，运维阶段构建工单闭环服务。该智能体应用于建筑全流程，推动行业从“经验驱动”向“数据驱动”转型，适配设计单位、施工企业等多场景，为住建领域数字化提供一体化支撑，具备显著技术创新性与行业示范影响力。

7.AI 驱动的科技数智化转型实践

申报单位：中国邮政储蓄银行股份有限公司

案例介绍：邮储银行 2023 年 5 月启动大模型能力孵化，在 2023 年底科技板块形成“两部两中心”布局后，紧握金融科技抓手，紧随以数字技术和人工智能为核心的时代浪潮，精细施策，建立了三维立体式的金融科技管理体系，借助大模型、机器学习、D2C 等新质生产力技术，积极探索银行科技领域应用场景，基于已有 30 余个系统基础上，打造了“AI 驱动的科技数智化中台”，以“用户”为中心，统一流程、统一数据、统一视图为抓手，构建专业领域十四大能力中心，打通管理链、研发链、运维链全生命周期流程，依托“邮智”大模型，

构建七大类场景应用智能体，推动 AI 驱动转型变革。该项目目前已面向总、分行推广，已成为支撑邮储银行科技金融服务的重要抓手。

8.AI 助手小牛

申报单位：途牛（南京）信息技术有限公司

案例介绍：途牛于 2025 年推出并开源“AI 助手小牛”智能出行管家，实现旅游行业 LLM 应用的重大突破。该智能体通过 13 个专业化智能体协同工作，覆盖机票、酒店、跟团游等全品类旅游服务。项目实现三大业务突破：实现旅游行业全品类覆盖；完成从查询到预订的全流程对话式服务，用户通过纯交互即可完成需求咨询、产品推荐到最终预订；实现多资源动态打包，智能协同完成机票、酒店等资源的实时匹配与组合。基于多智能体平台集成架构，支持自研与第三方智能体统一管理。“AI 助手小牛”已通过生成式人工智能服务备案，以一站式智能服务重塑传统旅游预订模式。

9.AI 助推江苏文旅行业智能化

申报单位：江苏省数字文化和智慧旅游发展中心

案例介绍：江苏省数字文化与智慧旅游发展中心作为江苏省文旅厅信息化建设单位，与蚂蚁集团、中电鸿信、南京欣网联合打造“四位一体”智能生态体系，包括：行业监管领域的“AI 监控”智能体、文旅数据分析领域的“智能问数”智能体、政务办公的“智问文旅”智能体、公共服务领域的“苏心游”智能体。该智能体体系面向政府、

企业、公众三端，提供天（无人机）、地（视频监控）的 AI 立体监测能力，提供智能查询、数据分析的 AI 分析能力，提供智能文件检索、业务场景适配的 AI 文档撰写能力，提供游前（云游景区、智能推荐）、游中（AI 伴游、智能讲解）、游后（“随手拍”投诉、智能反馈）的 AI 全链路支持能力。

10. 百度智能云数字员工：赋能千行百业服务营销提质增效

申报单位：北京百度网讯科技有限公司

案例介绍：面对企业营销、服务环节中存在的响应机械、数据割裂与效率瓶颈等难题，百度智能云数字员工以“结果交付”为导向，致力于将智能体从工具升级为可交付业务成果的新型生产力。研发过程中，数字员工深度融合大模型、数字人技术与行业知识，实现了高拟真形象生成与自然流畅的无感延时交互（语音识别准确率>98%，响应时延<1秒）；同时数字员工以“岗位化”形态切入企业高频场景，以“结果交付”为导向，内置上百个行业 SOP，支持开箱即用与极速培训，显著降低了企业应用门槛。目前，数字员工已广泛应用于金融、航空、教育、政企等行业，覆盖智能服务、培训讲解、内容运营、营销宣传等高频场景，为企业提质增效与用户体验升级提供了可复制的智能化路径。

11. “拔山”多模态智能笔录辅助系统

申报单位：中国刑事警察学院

案例介绍：为应对传统笔录制作中效率低、智能化不足等挑战，中国刑事警察学院研发了“拔山”多模态智能笔录辅助系统。该系统以多模态大模型和智能体为核心，集成知识图谱问答、深度语义质检、多模态内容解析及案件要素抽取四大核心模型，融合语音识别、语义质检等关键技术，构建了集语音交互、逻辑校验与线索结构化于一体的智能辅助平台。系统实现了低延迟实时语音交互与跨模态信息理解推理两大关键功能，有力支撑起笔录内容自动校验、案件要素智能抽取等上层应用。这一完整的技术闭环已赋能于刑事案件初查、行政案件办理等场景，满足多警种的复杂业务需求。

12.CarrotAI 大模型在多元场景中的落地应用

申报单位：江苏汇智智能数字科技有限公司

案例介绍：江苏汇智智能数字科技有限公司开发的 CarrotAI 大模型，针对教育、政务及智能制造领域的智能化需求，进行了适应性训练与优化。该案例重点解决了行业专业知识融入、交互意图精准理解等共性技术问题，形成了可快速部署的行业解决方案。目前已在部分高校的智能教学辅助、政务热线智能问答及工业流程咨询等场景中投入应用，运行稳定，有效提升了相关业务的自动化处理效率。

13.超级工程工地智能体生态创新实践

申报单位：中国联合网络通信有限公司软件研究院

案例介绍: 该案例面向雅鲁藏布江下游水电超级工程极端复杂的施工管理挑战, 由中国联通软研院构建了以数字孪生为中枢、多智能体协同的“智慧工地智能体集群”。案例通过部署“感知智能体”、“轨迹智能体”、“运维智能体”与“交互智能体”, 实现了对工地安全、进度、人员、设备的全域感知、智能分析与主动干预。案例创新性地 将多模态 AI 技术与工程管理深度融合, 攻克了传统管理模式“看不见、管不全、响应慢”的行业痛点, 实现了从被动处置到主动预警、从经验决策到数据驱动的管理模式变革, 为国家级重大工程的数字化建设提供了可复制、可推广的智能化范式。

14. “大模型+IPA” 双核驱动的数智化运维创新与实践

申报单位: 中移信息技术有限公司

案例介绍: 面对传统运维模式下人工操作效率低、流程管理不闭环、系统规模庞大复杂等挑战, 提出了“大模型+IPA 双核驱动的数智化运维”方案, 聚焦事件管理、变更管理、应急演练、系统巡检、告警处置、容量评估等运维场景, 通过构建岗位智能体与工具智能体协同的架构, 实现了运维任务的自动执行与智能决策, 创新性地 将大模型的认知能力与 IPA 的执行能力深度融合, 形成了“感知-决策-执行-闭环”的运维新范式。该方案已应用于中国移动计费账务、一级结算中心、集中结算等 12 个业务系统, 实现了运维流程的标准化、自动化与智能化。在提升运维效率、降低人力成本、保障系统稳定性方面成效显著, 为行业数智化转型提供了可复用的实践路径。

15.打造多模态“客服智能体”构建通信全场景服务新体验

申报单位：中国联合网络通信集团有限公司

案例介绍：“客服智能体”以元景大模型为核心底座，深度融合联通海量知识，通过 RAG、MCP、IVVR 等关键技术，构建“大小模型结合+多智能体协同”的架构体系，具备五大核心服务能力：支持多问题同时解答；实现上下文感知的多轮畅聊；在 10010 热线中提供视听一体服务；具备智能联想与追问推荐功能；根据用户状态提供差异化定制服务。自 2025 年 7 月发布以来，该智能体已全面赋能中国联通 APP、10010 热线等核心触点，为全国 4 亿用户提供精准高效、体验优质的智能服务，显著提升了客服效率与用户满意度。

16.得物商品智慧查验鉴别智能体

申报单位：上海得物信息集团有限公司

案例介绍：得物平台为确保消费者购买到全新正品，创新应用自研多模态大模型，构建了全球最大的查验鉴别体系，包括查验基地、样品库及智能设备。通过商品智慧查验鉴别智能体，得物攻克了自动化效率与模型通用性难题，率先实现 AI 查验鉴别的大规模产业应用。累计鉴别超 8 亿件商品，多品类准确率高达 99.999%，流程速度从小时级提升至秒级，应用规模与精度全球领先。2024 年成功拦截 160 万件假货，2025 年获世界人工智能大会最高奖 SAIL 奖。该智能体可拓展至众多产业及政府监管场景，以科技赋能产业高质量发展。

17.电信大模型应用开发平台-星辰智能体平台

申报单位：中国电信人工智能科技（北京）有限公司

案例介绍：星辰智能体平台作为一站式智能体引擎，致力于破解大模型技术门槛高、通用能力与业务需求脱节的落地难题。平台通过四大核心能力实现智能体全生命周期管理：提供零代码敏捷开发环境，通过可视化画布与场景模板，降低开发门槛 90%以上；深度业务适配，集成企业知识库与工具广场，实现从对话到业务执行的升级；支持闭环调优与敏捷部署，通过数据回流与运营看板精准优化，一键发布应用；具备全自动工作模式，超级智能体可自主理解指令、规划流程，7x24 小时执行跨系统任务。平台通过技术普惠化、能力业务化与进化自动化，精准匹配业务需求，驱动智能体持续进化，已成为赋能各行业数字化转型的关键基座。

18.多模型与智能体聚合及服务引擎(MoMA)

申报单位：中移九天人工智能科技（北京）有限公司（九天人工智能研究院）

案例介绍：企业用户面临模型与智能体选择组合难题，以及模型落地阶段成本与效益平衡的挑战。为此，研发了多模型与智能体聚合及服务引擎 MoMA，汇聚并管理多种模型与智能体，面向复杂场景动态规划，通过标准化接口调度模型与智能体，提供高准确、高安全、高可靠且高效的大模型推理服务。目前，基于 MoMA 的推理服务消耗

Token 数达 447 亿，API 调用次数 7000 万+。移动集团内部推理订单请求累计 2344 单，真实支撑移动内部真实业务场景使用。

19.多智能体融合平台驱动工业智能排程创新

申报单位：中科斯欧（合肥）科技股份有限公司

案例介绍：本案例旨在解决制造业在排程环节面临的数据异构、系统割裂及决策效率低等核心问题。基于自主研发的多智能体协同平台，该案例构建工业数据融合智能体，通过“协议自适应转换+语义动态对齐”双引擎实现 IT/OT/CT 全域数据融合，兼容 50 余种工业协议。平台采用五步标准化实施路径，结合业务模型，通过多智能体协同编排实现从数据接入到智能排程的全流程闭环。在 200+订单并发场景下，平台具备多约束条件自动排程、冲突预警与产能实时计算能力，使排程效率提升 70%，数据利用率从 31%跃升至 89%，有力支撑了制造业的数字化转型。

20.多智能体协同安全运营平台（Agentic AI SOC）

申报单位：中通服创发科技有限责任公司

案例介绍：中通服创发科技有限责任公司案例：天翼云安全运营中心项目，构建多 AI 智能体实时协作体系，通过华为昇腾 AI 技术认证，支撑国家云 SOC 能力提升。该中枢纳管 500+资源池，对接 IPS、WAF、蜜罐等 2500+安全设备，日均接入日志超 5 亿，实现 7*24 小时的网、数、信、反诈等全专业安全运营线上化，配套实战剧本 200+，工单

闭环率 100%。纵向拉通中国电信集团重保平台，累计处置公司内部威胁事件超 1 万起；在国家级攻防演练中连续三年 0 攻破、0 失分，国家公有云专项演练排名第二，有效提升威胁检测精度、缩短事件响应延迟、降低团队运营成本，成为国家云安全运营的核心技术中枢。

21.伏泰基于 AI 大模型的无人清扫智能应用

申报单位：苏州市伏泰信息科技股份有限公司

案例介绍：伏泰基于 AI 大模型的无人清扫智能应用，针对传统环卫“人力密集、感知决策闭环断裂、多机协同低效”痛点，基于 AI 大模型，研发无人清扫智能体。通过攻克多模态感知同步、大模型端侧适配等难题，通过 AGI 泛化能力提升机器人在复杂城市场景中的感知、认知与行动闭环能力，构建“环境理解-任务生成-作业执行-质量反馈”闭环，具备动态路径规划、多机协同调度、作业质量识别与数据闭环优化等功能。目前已在全国 11 省 30 余区域落地，覆盖城市道路、景区、校园等场景，实现扫净率 96%以上，运营成本降低 50%以上、作业效率提升 20%~60%的显著成效，推动环卫从“机械式作业”向“场景化智能协同”升级，为城市治理提供可复制的智慧环卫方案。

22.GPTBots 客服 AI 智能体

申报单位：深圳市和讯华谷信息技术有限公司

案例介绍: GPTBots 平台为一家国际知识产权服务与咨询公司解决了海量跨时区在线咨询导致的非工作时间客户服务响应不及时、潜在销售机会易流失的痛点。通过与 Zoho SalesIQ 深度集成, 实现 AI 智能客服 7x24 小时自动响应咨询, 复杂对话或销售机会自动转交人工跟进, 提升潜客留存和转化效率。创新点在于“AI 即服务”, 通过“无代码/低代码”界面降低 AI 应用门槛, 提供场景化解决方案并构建生态, 采用灵活付费模式, 降低企业风险与成本。核心技术优势在于 AI 智能体对 CRM 系统的“代理员工级”操作能力。GPTBots 团队深入调研客户需求, 在研发中攻克 AI 与 CRM 系统深度交互的技术难题, 优化知识检索与生成模型, 实现多模态融合和可视化 workflow 设计, 建立全生命周期模型管理体系。

23.国内首款面向企业的 L3 级 AI 服务智能体

申报单位: 联想(北京)有限公司

案例介绍: 当前, 智能体产业正以年均复合增长率超 40% 的速度蓬勃发展, 成为推动数字化转型的关键力量。但受限于算法瓶颈与技术能力, 智能体多停留在 L1-L2 级别, 难以真正介入到工作场景中。联想全新推出的联想百应智能体 2.0 版本能够深度融入企业高频场景, 突破预设 workflow 限制, 以多模态感知理解复杂需求、动态生成决策链条、闭环执行解决问题, 具备 AI 感知规划、AI 上下文增强、AI 多工具调用、AI 自生成、AI 自迭代五大 AI 核心能力, 真正让 AI 服务落地业务实处。联想百应智能体整体呈现出主动性、创造力和执行力三大

L3 级产品表现，将智能体从“响应式助手”升级为“协作级伙伴”，为中小企业智能化转型带来跨越式提升。

24. “AI+回声”智能体

申报单位：中国移动通信集团安徽有限公司、安徽省通信行业暨互联网协会

案例介绍：安徽移动 10086 热线每月产生 300 万通来话，其中 10% 会生成投诉工单，对于 90% 未立单的通话，存在大量的问题盲区。传统的人工抽查方式，合规检查覆盖率低，主观依赖性较大，溯源分析成本较高。解决方案：依托“聚智”平台，打造“AI+回声”智能体，解决投诉的源头治理问题。基于“全量”10086 来话，重塑“听音、回声、溯源、根因修复”四大流程。从用户诉求识别、用户态度感知等维度，及时识别风险，输出服务推诿、违规收费等 200+ 标签，赋能大音等业务平台，开展满意度修复、优化营销设计、辅助业务决策，实现了投诉治理“抓早、抓全、抓小”。

25. 敬业钢铁磁粉探伤：浪潮云帆无损检测智能体

申报单位：浪潮企业云科技（山东）有限公司

案例介绍：敬业钢铁有限公司针对钢化工件磁粉探伤作业中存在的缺陷漏检、误判及人工依赖度高等问题，研发了磁粉探伤全过程智能检测与评估系统。技术层面，采用高灵敏度荧光磁粉与复合磁化技术，构建基于磁痕目标检测与缺陷分类的多模态大模型，实现亚毫米级缺

陷的精准识别与在线自学习迭代。运维层面，通过自动化识别装备与智能工艺编排，建立覆盖磁化到复验全工序的智能检测体系，完成从“发现缺陷”到“闭环处置”的快速响应。该系统应用于钩缓装置零部件的磁粉探伤，有效降低了质量损失率与返工率，提升了检测环节的决策效率、准确率与可视化水平，显著改善了质量把控能力。

26. 经营管理智能体集群

申报单位：数族科技（南京）股份有限公司

案例介绍：该项目聚焦生产经营、财务融资、政策匹配等核心环节的智能化需求，通过融合多源数据与行业知识，构建了企业健康度分析、财务风险预警、政策精准匹配等八大核心模型的智能分析体系。研发中攻克了多源数据实时融合、边缘云端协同推理等关键技术，创新性地知识图谱与深度学习结合，实现经营管理场景的深度建模。系统支持从数据采集到决策执行的端到端服务，具备高实时性、强泛化能力和业务可解释性。该系统已在多家制造企业完成部署，有效提升了企业经营韧性、运营效率与风险应对能力，为区域产业治理与产业链协同提供了数据智能支撑。

27. 晋商银行智能保函审核

申报单位：晋商银行股份有限公司

案例介绍：2023年以来，晋商银行大力发展分离式保函业务，随着业务规模的不断扩大，粗放的业务审批流程与精细化的管理要求矛盾

逐渐显现。为解决传统保函审核流程中业务审批效率受限以及自然语言处理技术存在局限性等问题，晋商银行积极探索并创新解决方案。通过引入自动化风险评估、大模型增强语义理解能力以及负面案例检索与规则引擎结合等技术手段，晋商银行成功实现了保函审核流程的智能化升级。实践结果显示，该方案不仅显著提升了业务场景应用成效，还带来了明显的社会效益，并为金融行业落实“五篇大文章”战略提供了有力支撑。

28.九天智能体平台

申报单位：中移九天人工智能科技（北京）有限公司（九天人工智能研究院）

案例介绍：九天智能体平台基于智能体自主任务规划与执行引擎、知识增强等前沿技术构建，致力于帮助个人与企业用户以零代码或低代码方式，低成本、高效率地开发定制化 AI 应用智能体，推动大模型技术在千行百业的广泛落地。平台提供多样化的应用构建支持，包括插件集成、知识增强、 workflow 编排等核心能力，并可通过 API/SDK 灵活对接业务系统，全面降低智能体构建门槛。目前，平台已累计服务用户超 47 万，为 38 家以上省专公司提供稳定服务，累计调用量突破 7000 万次，知识库规模达 6300 余个，段落集超 2480 万，插件与工具总量分别达到 633 个和 608 个，展现出卓越的平台服务能力与生态成熟度。

29.九霄智基大模型开发管理平台

申报单位：浪潮智慧城市科技有限公司

案例介绍：九霄智基大模型开发管理平台是全栈式 AI 服务平台，致力于解决智能体开发面临的技术门槛高、研发效率低等核心问题，为数字政务、社会治理等场景提供从 AI 资源管理到智能体开发的一体化服务。平台由三大核心模块构成：训推中心提供模型调优、评估及算力数据管理；算法中心支撑 AI 算法全生命周期管理；智能体开发中心集成知识库与工具开发功能，支持通过拖拽式操作快速构建智能体应用。该平台显著降低了大模型技术的应用门槛与实施成本，有效提升了开发效率，助力政府及企业构建坚实的数智能力底座。

30.基于 AI 的因岗因人安全应急知识学习系统

申报单位：中安华邦（北京）安全生产技术研究院股份有限公司

案例介绍：本案例针对我国企业数量多、员工规模大、岗位结构复杂、安全素质不一的现状，以 AI 技术为核心，构建了“华文安典”大模型，建立岗位学习知识图谱，并运用认知诊断技术实时监测员工知识掌握状态与学习过程，精准评估学习成效。案例配备丰富的内容知识库，全面覆盖各类企业需求，为企业安全培训提供资源保障。该案例构建了融合知识、问题与能力关联的岗位知识图谱，实现对员工能力精准量化评价，开发矩阵式岗位安全知识库与智能推荐系统，实现培训个性化和智能化。目前，该案例主要应用于矿山、化工、制造等生

产型企业，支撑员工安全素质提升与安全培训管理数字化，已获国际先进科技成果评价。

31.基于操作智能体构建云网自智运维体系

申报单位：中国移动通信有限公司

案例介绍：中国移动云网运维智能体针对云化核心网在保障电信级高可靠性时面临的运维挑战，构建了“训·推·用”三位一体的云网协同自智运维体系。该体系结合操作智能体与高稳运维体系，在全球最大规模 NFV 算网中实现任务自受理、能力自编排与服务自监控。其核心是基于 ReAct 闭环机制的操作智能体系统，突破了端到端故障自愈的技术瓶颈，实现秒级隐患发现与分钟级定位抢通。目前已规模化部署 31 省，纳管近 50 万台设备，每年监测数据近 300 亿条，云端直控超 7000 万次，自动化处置故障 18.4 万次，将平均故障处置时长从 113 分钟压降至 15 分钟内，年节约运维成本超 2000 万元，显著提升了运维效率与系统可靠性。

32.基于多模协同的漏洞利用攻击识别与验证

申报单位：恒安嘉新（北京）科技股份有限公司

案例介绍：针对网络流量中存在的漏洞利用攻击识别及验证程序 PoC 生成需求，解决目前“漏洞判定 - CVE-ID 推理 - PoC 生成”全流程自动化不足、算力成本高、结果难溯源等问题。该案例通过采用多模型融合（BERT、Qwen3-32B、SecGPT）与检索增强生成（RAG）技

术，突破低算力部署与高准确率平衡难点，创新研发了 Agentic RAG 一体化链式框架智能体，覆盖三大流程处理环节。该成果应用范围包括网络安全领域的流量漏洞利用检测、CVE-ID 匹配及 PoC 自动化生成，具备结构化信息可审计、低成本快速部署特性，内置 6000+CVE-ID 相关数据支撑推理，已用于威胁流量监测与攻击溯源场景，提升漏洞利用攻击识别的效率。

33.基于多智能体协同的融媒内容智创平台

申报单位：宁波广播电视集团

案例介绍：宁波广播电视集团基于多智能体协同的融媒内容智创平台，充分利用自然语言大模型和视音频 AI 等技术，研发出大量适用于媒体行业的 AIGC 创作工具，通过科学的平台管理，形成一个涵盖文字、图片、音乐、视频等多种形式的，能够辅助或自动生成高质量、多样化的媒体内容生成智能体。通过自研与能力接入，本平台实现了智能对话、AI 绘画、AI 音乐、AI 视频以及传统媒体能力等多种工具集的对外服务，研发并推行了一套平台级 AI 能力的统一 API 接口规范，实现 AI 能力的即插即用与服务化供给。目前，已有超过 200 家企事业单位入驻平台，生成式人工智能生成次数超过 250,000 次，在省内多家媒体机构多个栏目中进行日常内容生产应用。

34.基于工道智能体+AR 技术的设备智能维护

申报单位：苏州协同创新智能制造科技有限公司

案例介绍：面对电子信息制造业在点检、巡检环节依赖人工导致的效率低、漏检率高等问题，本方案创新性地将多模态大模型、AR、数字孪生与智能体技术深度融合。方案以 AR 眼镜为硬件入口，通过视觉识别实现图纸与实物的自动叠加比对，构建“扫描即识别、所见即所得”的交互体验。核心由多智能体协同工作，自动完成从异常发现、报告生成到任务下发的全流程闭环，将传统质检效率从 15 分钟大幅提升至 30 秒。该方案通过“云-边-端”架构将技术深度融合入 workflow，构建出具备持续进化能力的专家级辅助系统。作为核心基础设施，已助力美的获得“世界首个多场景覆盖智能体工厂”认证，并在多家产业链企业成功推广应用。

35. 基于混合双模型与智能体的企业服务全链路智能客服系统

申报单位：北京中科汇联科技股份有限公司

案例介绍：慧政大模型作为中科汇联慧政超脑的核心，是基于合规国产基座模型、利用高质量政务数据微调得到的行业大模型。它通过自然语言交互理解用户需求，结合政务知识搜索技术，为用户提供精准的政策解读与问答服务。相比通用模型，慧政大模型显著提升了政务文本理解与生成能力，将政务问题解答率从 85% 提升至 95%，回复可读性提升 50%，知识梳理效率提高 2 倍。结合 DeepSeek-R1 模型后，增强了推理与多模态交互能力，实现“深度思考”。通过检索增强生成（RAG）技术，进一步保障了知识检索的精准性与回答的权威性。有效推动了服务向精准化、个性化、高效化迈进。

36.基于“京办”的政务 AI 智能体技术及其应用

申报单位：京东科技信息技术有限公司

案例介绍：该智能体面向政府工作人员及公众用户，旨在通过自然语言交互提供高效、准确的政务智能服务，提升服务效率与质量。该智能体采用“大模型+领域知识库”技术路线，核心功能包括：智能问答，基于多源政务数据构建的知识库，为工作人员和市民提供精准的政策解读与业务咨询；智能问数，依托高质量政务数据库，支持自然语言的数据查询、分析与可视化展示，助力数据驱动决策；智能写作，结合素材召回与大模型生成能力，提供续写、摘要、大纲等辅助功能，提升公文写作效率与质量。

37.基于九霄知言大模型的交通管理智能体

申报单位：浪潮智慧城市科技有限公司

案例介绍：基于九霄知言大模型的交通管理智能体，采用“1+1+N”技术架构，面向交通管理行业构建，旨在解决民意数据分散、处置效率低、决策支撑不足等痛点。该智能体具备四大核心能力：实现7×24小时交通民意智能问答服务；融合大模型与BALLTree聚类技术，精准识别热点问题及高发区域；自动构建民意画像并生成多粒度分析报告；借助RAG检索技术为执法人员提供精准法条查询与执法建议。通过以上能力，该智能体显著提升了交通民意处理的效率与规范性，有效助力交管部门科学决策，增强群众满意度。

38.基于内生智能的大小模型协同自智网络技术创新与应用

申报单位：中国移动通信集团安徽有限公司

案例介绍：随着网络规模发展，网络结构演进，无线网运维工作的要求不断提高，但还面临成本费用需进一步压降的矛盾，要在优化效率及运营效益方面进一步提升，亟需依托智能化手段实现网络运维降本提质增效。安徽移动项目团队开展了基于内生智能驱动的多智能体协同无线运维系统研发和应用探索，通过“内生智能+大模型+数字孪生”技术的结合使用，构建了具有更细粒度的问题挖掘、更精准的分析定位、更安全的方案执行特征的无线网运维智能体，提升网络质量和运行效能。此外，创新了多智能体协商、数字孪生仿真等技术，提升方案准确性和执行的安全性。

39.基于运维行业大模型的智能运维管理平台

申报单位：联通数字科技有限公司

案例介绍：该案例通过构建“动态感知-智能分析-自主执行”的闭环体系，利用自研采集器和健康度评估模型，打造了新一代 AI 驱动的智能运维管理平台。平台采用 SM4 加密与自动化引擎，实现对多品牌关键基础设施的标准化管理。同时融合大语言模型、RAG 及 Multi-Agent 技术，基于多年运维知识构建 AIOps 智能体框架，打造覆盖故障预警、网络感知、巡检管理等场景的同舟智能运维助手。目前，该项目已成功应用于 10 余家央国企及政府单位的关键基础设施

运维，显著提升了运维效率与管理水平，有力推动了智能化运维与运营全要素生产力的升级，为国家数字化建设提供了安全可靠的运维支撑。

40.基于智能体技术的云平台智能运维系统

申报单位：厦门渊亭信息科技有限公司

案例介绍：本项目由渊亭科技针对同方有云重点运维业务的痛点进行行业大模型建设。具体的，处理大量的云平台运维数据（日志、监控信息、应用信息等），利用开源大模型训练运维基座模型，并结合业务需求，孵化面向不同场景的专业运维大模型，能够进行自动发现故障和自动异常检测，探索了基于运维大模型的智能运营。同时，设计了数据回流机制，能够在运维过程中从海量运维数据中不断进行处理加工和提炼，反哺专业运维大模型，形成数据飞轮。

41.卡易无极之枢影像智能体解决方案

申报单位：浙江卡易智慧医疗科技有限公司

案例介绍：项目旨在解决传统医疗影像流程效率低、资源分配不均等痛点，以“无极之枢影像智能体”为核心，面向影像检查全流程，构建综合影像智慧应用平台。智能体通过整合多源数据、采用对象存储与分布式数据库，打破数据壁垒，提升数据融合效率，并在保障数据安全与隐私的前提下，解决单点技术局限、泛化性不足、环节依赖人工经验、主观性强、特殊场景适配难等问题。平台运用知识增强推理

技术实现跨模态协同分析，结合双深度网络等决策模型，为医疗机构和区域提供影像智能预约、智能技术质控、智能辅助诊断、智能报告书写、语音图像调阅处理、智能报告解读、智能随访召回等一体化智能服务，优化资源配置，减轻医护人员工作负担，推动影像医疗服务提质增效。

42.科技项目智能治理

申报单位：国家能源集团物资有限公司

案例介绍：科技项目智能治理案例针对能源行业科技项目管理中存在的“评审标准不统一、过程管控薄弱、知识资产沉淀低效”等难题，开展了覆盖科技项目管理“研立执结”全生命周期的智能治理体系建设。该案例深度融合自主研发的多模态认知智能与动态知识演化技术，构建了包含研发智库、立项智审、执行智管、结项智收四大核心模块的一体化管理平台。通过系统整合国家标准、行业规范与企业最佳实践，建立了包含 100 余项量化指标的评审体系和 200 余个可复用模板库，实现了从技术可行性评估、立项智能审查、执行动态管控到成果转化评估的全流程数字化治理。该案例突破了传统依赖人工经验的管理模式局限，显著提升了科技项目管理的科学性、效率与透明度。

43.联通公众 APP “智办”引擎，全场景 AI 交互新范式

申报单位：中国联合网络有限公司广州软件研究院

案例介绍：传统 APP 系统菜单繁多，业务办理效率低，各类应用功能菜单种类繁多、菜单层级深用户难以找到，增加了操作的复杂性和时间成本。此外，在办理复杂业务过程中，步骤繁多、业务产品组合多样，存在诸多规则限制和互斥依赖关系、且依赖手工录入信息，营业员办理业务时极易错选、漏选，降低业务办理效率，也降低了整体运营效率。以一线实际使用感知作为唯一标准，通过人工智能全面提质增效打造端侧智能化架构体系，助力中国联通业务效能提升。基于统一技术架构，面向一线用户打造“公众通通助手”，为客户提供咨询、办理、营销等联通智慧服务，为一线提供“看得到、查得准、办得快”多元场景支撑。

44. 联通 U 爱——面向一老一小一特群体的新一代智能终端

申报单位：联通在线信息科技有限公司

案例介绍：联通在线公司响应国务院人工智能+行动意见，聚焦新一代智能终端及智能体服务，本年度打造一款无屏化、一键式语音交互的创新型产品-联通 U 爱。聚焦“一老一小一残”三类群体在安全、健康、陪伴的核心痛点需求，覆盖用户的可信通信、安全守护、信息娱乐、文化教育与生活健康五大核心场景，成为用户身边情感属性与功能属性兼备的随身伴侣。该产品围绕终端、语音交互、固件系统、模型算法、生态能力等方面推进产品研发，终端侧采用紫光国产化 AI 芯片和有源 NFC 芯片，基于 RTOS 轻量化系统进行固件开发，在语音方面，打造了 36 个高意图场景，11 个高频智能体服务，使用了

4 个参数量模型进行场景化微调，以提高适用性和回复效率，有效识别率达 93%。

45.Lumo AI 合规智能体

申报单位：上海淇毓信息科技有限公司

案例介绍: Lumo AI 合规智能体是面向金融及强监管行业的智能平台，通过 AI 与数据治理体系的深度融合，构建新一代“AI+安全”合规方案。该平台技术路径清晰，构建了覆盖法规知识图谱、制度模板库、风险评估模型与合规算法引擎的智能中枢，在数据安全、内容检测与模型优化方面形成独特优势。其应用成效显著，能在 48 小时内完成新规制度化落地，政策适配准确率高达 99.2%，日均可处理 20 万级安全请求，效率远超人工。该智能体不仅帮助金融机构显著降低合规成本、提升风控能力，还推动了行业合规能力的均衡化与标准化，助力监管部门实现实时、动态的穿透式监管。

46.MaaS 为基，AaaS 为翼，释放智算无限潜能

申报单位：中国移动通信有限公司

案例介绍: 本案例围绕大模型与智能体技术变革，创新构建了融合异构算力底座、MaaS 与 AaaS 的三位一体智能体算力服务平台，有效驱动中国移动网络云数智能化转型。系统核心包括：高带宽、低时延的异构智算资源池，奠定坚实算力基座；覆盖大模型训推全流程的 MaaS 服务，实现资源的快速弹性供给；支持低代码开发的 AaaS 能力，推

动智能体应用规模化落地。该平台实现了千亿级大模型并行训练性能提升 36.5%，推理性能提升 59.4%，集群长稳训练达 14.3 天，达到国内领先水平。每年节约智算成本超 3350 万元，并通过智能体应用节省人工运维成本 2000 余万元，成效显著。

47.金融教育基地智能体“消百科”

申报单位：蚂蚁科技集团股份有限公司

案例介绍：蚂蚁集团秉承金融普惠使命，积极推动“三个一”普惠教育行动，即构建金融必修课程体系、打造金融教育开放平台、建立金融素养评估与研究体系。依托支付宝平台，面向消费者与行业机构推出金融知识学习服务平台“金融教育基地”，并通过 AI 技术赋能，由传统教学模式向“主动式学用结合”模式迈进。基于大模型与数字人技术，打造了金融教育智能体“消百科”，不仅能够精准解析用户需求，还能实现金融知识解答、金融课程以及实用工具精准推荐，实现“从学到用”的高效转化。同时，蚂蚁集团将这一技术能力开放赋能行业，支持行业机构实现金融教育模式的数智化升级，共同构建开放共享的金融教育生态，为金融普惠理念的深入落地贡献力量。

48.面向通信网络监控感知分析的运维智能体

申报单位：中国移动通信集团天津有限公司

案例介绍：本案例旨在解决通信网络运维中数据治理难、监控存在盲区、知识应用断层及人机协同效率低等核心问题。案例创新性地将九

天大模型与数字孪生、多模态交互技术深度融合，构建了集“网管数字孪生监控系统、虚拟数字人、网络运维知识大模型”于一体的智能运维平台。平台通过构建智能数据治理体系，实现了TB级运维数据的标准化处理与全生命周期管理；利用RAG增强生成等技术，攻克了专业知识的动态更新与精准应用难题；并通过多模态虚拟数字人提供了自然高效的统一交互门户。该成果有效实现了故障的精准预测与智能诊断，在试点应用中显著提升了运维响应效率与决策准确性，为通信网络的智能化升级与数字化转型提供了具备高可复制性的实践路径。

49.面向维护作业和设备管理的视觉大模型智能体应用

申报单位：中国铁塔股份有限公司

案例介绍：目前，中国铁塔智能维护涉全国55个地市，覆盖17万个站址的使用及维护系统。传统运维模式面临着如下挑战：1.效率瓶颈：为实现全国2000余万设备设施监控，依赖人工巡检导致运维成本攀升且时效性不足，传统方式难以满足实时响应需求。2.安全隐患加剧：基站设备暴露于复杂环境中，存在非法入侵、设备被盗等风险，人工监控难以实现全天候精准防范。为了进一步实现精细化的智能维护，本项目通过视觉大模型智能体来实现对机房内人员行为、及资产情况做进一步的识别，能够快速判断和识别人员行为，机房场景下的人员行为识别准确率提升，准确率达到85%以上。

50.某省医保局跨模态医保 AI 智能客服

申报单位：二六三网络通信股份有限公司

案例介绍：某省医保局跨模态医保 AI 智能客服案例，针对群众医保服务便捷化需求，尤其老年群体数字渠道使用困难、非工作时间人工服务缺失的问题研发。研发中融合 AI、数字人、5G/4G、OCR 等技术，构建含“视频经办”的智能客服系统，突破业务流程智能化、智能调度、医保大模型构建等难点，创新引入电子签名、智能外呼功能。应用于医保经办服务领域，覆盖异地就医备案、参保登记等高频业务。其以多模态交互优化体验，政策误答率 $\leq 0.05\%$ ，7×24 小时服务提升覆盖与效率，助力医保服务数字化转型，实现了基层经办窗口的智能化服务升级。

51.全流程医疗健康管理大模型系

申报单位：中国移动通信集团天津有限公司

案例介绍：天津移动案例全流程医疗健康管理大模型系统面向国家战略与行业转型需要，针对智慧诊疗中数据整合困难与智能生成能力不足等问题，采用“微调+知识增强”技术路径，结合中国移动九天与 Deepseek 大模型，实现了体检报告自动化生成、权威医疗指南查询与病历图像理解分析等核心功能。项目创新构建健康管理知识脑图系统，融合思维链与医学知识图谱技术，有效抑制大模型推理幻觉，并基于患者画像智能体生成个性化健康报告。平台采用模块化架构，支持灵

活部署与快速对接各类体检中心及医疗机构，具备良好可复制性与推广能力，为健康管理领域提供了高效、低成本的数字化解决方案。

52.全栈故障定位智能引擎

申报单位：中国移动通信集团甘肃有限公司

案例介绍：全栈故障定位智能体基于中国移动聚智平台，旨在提升甘肃移动支撑系统的故障定位能力。该智能体采用“采集-处理-分析-展示”四层架构，通过各类 Agent 及 API 自动采集 IaaS、PaaS、SaaS 全栈多源异构数据。系统对非结构化日志进行实时清洗与标准化处理，转换为统一格式以降低 Token 消耗。基于 Flink 流式处理框架进行数据聚合与异常检测，并利用九天大模型结合运维知识图谱进行智能推理，生成根因分析与处置措施。诊断结果通过 API 实时推送至监控大屏与移动办公平台，实现信息共享。该方案无需逐个开发故障智能体，具备全栈通用性。应用后显著缩短了平均修复时间，提升了“1-5-10”运维能力，推动了 AIOps 落地，有效降低了运营成本。

53.融信通智能消息中台解决方案

申报单位：联动优势科技有限公司

案例介绍：当前数字金融的发展进入了深水区 and 高质量发展阶段，金融机构需要摆脱对于外部流量和数字化能力的依赖，迫切需要建立自主可控、成本合理、数字化服务入口和生态营销环境。联动优势融信通智能消息中台解决方案依托 5G 消息、AI、区块链、云计算和物联

网等数字化技术，充分融入信用卡、理财、基金购买、生活缴费等金融服务场景，从而帮助金融机构客户建立全新的数字化经营体系，实现金融机构服务的轻量化、智能化、全媒体化、全量化，帮助金融机构实现智能营销，实现降本增效。

54.入口级 AI 智能体赋能行业创新应用

申报单位：中国移动通信集团江苏有限公司

案例介绍：江苏移动针对大模型领域 AI 入口繁多、应用碎片化及生态协同不足等痛点，打造苏晓伴 AI 智能体平台。平台以“轻量化入口、资源整合、能力融合、生态贯通”为核心，汇聚主流大模型与 120 余项 AI 应用，构建统一智能入口。基于多模态融合架构及 MCP 协议，实现智能体能力标准化集成与协同调度，支持 SaaS、API 及私有化部署，可灵活嵌入既有系统。平台支持“千人千面”的 AI 服务按需匹配，根据企业数智化成熟度提供适配路径，显著降低应用门槛。目前已覆盖 4000 余家集团客户，深度赋能政务、教育等领域，成功打造“常相伴”、“扬小关”等 50 余个标杆案例，为行业智能化转型提供了可复制的实践路径。

55.涉诈网络资源分析智能体

申报单位：恒安嘉新（北京）科技股份有限公司

案例介绍：涉诈 APP 与网址呈“变种快、伪装深、关联密”特征：新型涉诈 APP 篡改代码、换图标避监测，近 60%涉诈网址仿正规域

名且短期注册，给传统反诈研判带来挑战。传统方式中，涉诈 APP 需提取包名、签名等信息比对，难追踪通联地址变化；涉诈网址靠核查备案、比对黑名单，难辨高度伪装网址且易因网址失效延误研判，还缺乏关联分析能力，导致研判碎片化。该方案依托 DeepSeek、千问等大模型，采用“通用大模型赋能+垂直小模型专精”双层架构，实现涉诈 APP/网址从“线索收集-深度分析-精准研判-决策输出”全流程智能化，同时沉淀涉诈 APP/网址情报特征库，有效破解传统研判难题。

56.世运会 AI 赛事管理平台

申报单位：成都明途科技有限公司

案例介绍：明途科技为成都 2025 年世运会开发的 AI 赛事管理平台，基于“任务全覆盖、过程全留痕、结果全运用、AI 全赋能”的设计理念，依托 Workbrain 大模型，攻克了多系统整合、多角色协同等难点，贯穿赛事筹备、执行到复盘的各个环节，实现任务智能分配、实时填报、自动预警等功能，并与政务、OA 等系统无缝对接，打破了信息壁垒，提升了运营效率。该案例验证了 AI 在大型赛事中的应用价值，推动体育产业与人工智能融合，助力成都智慧城市建设。目前平台应用活跃，未来将继续深化 AI 技术在赛事服务中的创新。

57.水滴保智能电话&在线客服-保小慧

申报单位：北京水滴科技集团有限公司

案例介绍：“保小慧”聚焦保险服务场景中客户咨询量大、响应时效要求高的痛点，基于自然语言处理与深度学习技术，构建覆盖电话与在线客服的智能服务平台。研发过程中突破多轮对话理解与业务知识融合难题，实现问题流转、产品解读等复杂场景的精准响应。系统已全面应用于售前、售中及售后服务环节，显著提升服务效率与用户满意度，月均交互量超百万次，响应准确率持续优化，推动保险客户服务向智能化、精细化发展。

58.数据融通、模型融汇、场景融智，打造招采一体化智能体

申报单位：中国移动通信集团陕西有限公司

案例介绍：为响应国家招投标采购领域数智化转型战略，陕西移动针对传统招采流程中文档处理效率低、智能辅助弱、风控滞后及供应商寻源难等痛点，打造了多模态“招采一体化智能体”。该智能体基于人工智能与大数据技术，构建多源智能中枢与动态知识图谱，具备六大关键能力：采购方案智能推荐、文件自动生成、评审流程指引、风险实时预警、文档合规稽核及智能问答。通过需求洞见、策略推荐与风险防控，实现了智慧采购、合规与监管。该方案有效解决了数据孤岛、效率低下等难题，显著提升了招采流程的智能化水平，为行业数智化转型提供了可复用的技术范式，具备广泛的应用推广价值。

59.数智驱动普惠金融服务全链路革新

申报单位：成都优卡数信信息科技有限公司

案例介绍: 针对普惠金融机构面临的线下拓客难、运营效率低等痛点, 以及用户产品选择难、借款流程复杂等需求, 优卡科技投入超 4000 万元研发, 构建“1+N”产品布局, 整合智慧营销、AI 外呼、CRM 等六大 SaaS 产品, 形成覆盖“获客-转化-风控-运营-管理”的全链路数字化解决方案。该方案已服务 1200 多家金融机构, 覆盖 372 座城市, 累计服务用户超 1.2 亿人次。智能风控模型有效助力机构实现精准获客与降本增效, 2021 年 SaaS 收入达 7000 万元。案例入选成都大数据产业发展蓝皮书, 获评四川省新经济重点平台等多项荣誉, 持续推动普惠金融服务“最后一公里”的落地实践。

60.问法湖南--5G 智慧精准普法平台

申报单位: 湖南红网新媒体集团有限公司

案例介绍: “问法湖南”是由湖南省委网信办、湖南省司法厅指导, 红网打造的以“问法”促“普法”的精准法律服务平台。聚焦媒体、法律与人工智能交叉领域, 构建集普法宣传、法律咨询、法律服务为一体的智能化平台, 通过法律咨询、类案检索、合同审查、法律文书生成等多种智能体, 着重解决法律人才资源供需失衡、服务响应存在局限、普法工作精准度低等问题, 实现法治需求与普法供给精准对接, 让群众足不出户即可享受全业务、全时空的精准普法服务。平台自 2023 年 8 月上线以来, 累计入驻律师 1151 人、入驻律所 375 家, 法律咨询服务的 7.4 万余次。获 2025 年王选新闻科学技术奖(进步奖)、

2024 年中央网信办“全国网信系统十大网络普法创新案例”及中记协“媒体+优秀案例”荣誉。

61.小智 AI-实时 AI 交互系统

申报单位：深圳新智未来科技有限公司

案例介绍：该系统旨在解决智能软件开发门槛高、技术封闭及二次开发难等行业痛点。项目集成语音识别、语音合成、自然语言理解、情绪感知及声纹识别、记忆体功能等多模态交互模块，通过模型剪枝与量化优化，实现大模型在低算力环境下的高效运行。研发过程中，团队突破了实时多模态语音交互和情绪分析的技术难题，将回复时间压缩至 500 毫秒以内，并构建了全栈式开源软件生态，使开发者能够快速二次开发并适配多场景应用。截至目前，小智 AI 已在教育陪伴、智慧养老、智能家居等领域广泛落地，累计部署设备超 50 万台，日均处理对话量达 300 万次。项目在 GitHub 社区累计收获 20.4K stars、4.1K forks，并荣登 GitHub 全球 C++ 开发者 Trending 榜榜首。

62. “兴海”——船海学科教育大模型

申报单位：哈尔滨工程大学

案例介绍：船海学科面临师资队伍从单学科背景向多学科交叉转变，培养体系从序贯式灌输型向递进式自主型培养转变，实习实践从船厂见习向数字化智能化实操转变；原有吸收欧美船型的增量创新已进入瓶颈，船海人才原始创新能力亟待提升。作为 A+ 一流学科，与全国

最大船海试验中心（中船 702 所）、民船设计国际领先的上海船舶研究设计院、“爱达·魔都号”建造方上海外高桥造船有限公司等共建学科大模型。大模型首创四层框架，提供“五特四通”功能，面向学生：提供泛在学习、实践训练、职业规划；面向教师：提升教学效果、教学效率、教学效益；面向学科，促进资源整合、体系重构、交叉创新；面向社会：提供行业支撑、全球标杆、教育公平。大模型已上线国家平台。

63. “一案多查”刑事办案智能辅助系统

申报单位：北京市房山区人民检察院

案例介绍：该案例开发背景源于 2023 年最高检“高质效办案”要求，聚焦检察机关在“一案多查”中需同步审查刑事、民事、行政、公益诉讼等多重法律关系线索的实践需求。研发以“通用大模型+小包公法律垂直大模型”为技术基座，通过改进模型架构突破 token 长度限制，实现对长文本证据的有效处理。系统构建了不同罪名的知识图谱，接入实时法规数据库和类案检索系统，形成专业法律 AI 应用。创新性地将大模型能力应用于“四大检察”融合履职场景，支持复杂案件中不同性质法律问题的同步分析与线索关联。目前该系统已在检察办案环节投入应用，为检察官处理复杂案件提供技术辅助。

64. 以决策智能为核心的无线运维技术

申报单位：中国移动通信集团设计院有限公司

案例介绍：中国移动无线网络运维智能体案例，聚焦通信行业网络运维转型需求，针对当前运维人力依赖强、成本高、故障响应慢等痛点，构建以大模型为大脑，智能体协同决策的网络智能运维体系，打造以决策智能为核心的无线运维解决方案，构建“感知-分析-决策-执行-高阶演进”全生命周期体系。技术上首创基站态势智能感知与故障智能分析等核心能力，有效解决故障预测与根因定位难题。该智能体已应用于安徽、江苏、福建等 10 个省份，实现无线故障自闭环率超过 20%，隐患识别效率提升 22%，故障工单处理时长压降 15%，显著提升运维效率，有力推动通信产业数智化转型，具备广泛的行业推广价值。

65.一款面向微生物检验场景的智能体案例

申报单位：迪安诊断技术集团股份有限公司

案例介绍：该案例开发了一套基于深度学习的革兰氏染色智能分析系统，通过卷积神经网络实现细菌/真菌/细胞的自动分类与定位。系统整合图像预处理、特征提取和智能判读模块，能够快速准确地完成革兰氏染色图像的质量评估，以及对细菌/真菌/细胞进行识别，同时通过智能算法，对全玻片图像进行分析，给出具体感兴趣区域，实现人工智能采图，智能报告总结等功能。该系统可辅助检验人员提高工作效率，减少人为误差，并为临床提供更可靠的细菌/真菌/细胞鉴定结果。未来可进一步拓展至耐药性预测等应用场景，为抗感染治疗决策提供智能化支持。

66. 依托大模型的智能巡检及根因排查

申报单位：北京中电汇通科技有限公司

案例介绍：2025 年国家推进“人工智能+”，国网发布专项方案，8 家网省电力试点 AI 场景。电力运维存在“事多人少成本高”的痛点：某信通公司半年检修两千余次，人均日检修 7 小时，人工巡检 1 小时/次、故障排查 2 小时/次，且工单审核易漏判。故研发以“LLM+行业数据+专家经验”构建智能体实现以下功能：工单风险评估自动核验工单合规性，智能巡检整合多平台数据生成报告，故障分析快速定位根因，薄弱环节分析挖掘隐患；借 MCP 封装工具、RAG 优化知识库突破数据集成与幻觉问题。已在 3 省电力落地，巡检效率提升 80%，故障处理 ≤ 10 分钟，工单故障率降 90%，每月根据新增工单、巡检数据微调模型 1-2 次，已适配 12 类新型设备操作场景。

67. 一站式大模型应用构建平台

申报单位：马上消费金融股份有限公司

案例介绍：为应对金融行业智能化转型中 AI 技术功能单一、适应性不足的挑战，马上消费金融股份有限公司研发了“一站式大模型应用构建平台”。该平台通过异构算力调度、低精度量化等技术提升算力利用与推理性能；借助智能体代理机制实现任务分解与动态优化，增强多模态交互与自适应学习能力；同时采用隐私计算与合规框架，保

障数据安全及模型可解释性。平台融合了异构调度、自主决策与行业应用模板等创新，有效兼顾了性能与合规要求，已在客服、风控等场景验证成效，有力推动了金融 AI 的规模化应用与技术门槛降低。

68. 邮储银行全流程测试智能体生态群

申报单位：中国邮政储蓄银行

案例介绍：为应对金融业数字化转型与监管趋严的挑战，邮储银行构建了深度集成于自研平台的智能测试引擎。该引擎通过大语言模型、智能体平台及软件工程技术，构建全流程、体系化的智能测试应用能力，推动测试模式向人机协同升级。首先将 AI 能力融入测试全流程，形成由 40 余个智能体组成的生态，覆盖需求分析、用例生成、智能执行及缺陷分析等环节，实现测试工作全面智能化。其次基于积累的测试资产构建领域知识库，推动应用场景从通用向知识化深化。最后通过 AG-UI、MCP 等协议搭建多智能体协同框架，以“邮小测”超级助手实现技术应用标准化，助力测试人员专注业务问题，深化敏捷研发转型。

69. 邮储银行业务需求智能体

申报单位：中国邮政储蓄银行

案例介绍：在“数智化”转型不断深化的背景下，邮储银行面临企业级系统数量增多、规模扩大、业务复杂度持续提升的挑战，需求管理环节存在对个人经验依赖过重、学习成本高、质量不均、协同效率低

等问题。为应对上述挑战，邮储银行启动科技数智化 AI 能力提升专项，综合运用大语言模型、智能体平台与软件工程等前沿技术，自主研发智能需求引擎；深度集成于统一需求流程入口——业务需求中心，将 AI 能力无缝嵌入需求编制、检查、审核及资产管理等关键环节。同时，打造“邮小需”多任务智能体助手，有效支撑复杂任务协同处理；并通过构建业务领域知识库，贯通企业级数据资产，实现场景调优与智能检索，全面赋能业务需求管理的全链路数智化转型。

70.邮件网关嵌入式钓鱼检测智能体

申报单位：北京兴云数科技术有限公司

案例介绍：针对 AI 生成钓鱼邮件攻击激增、传统检测误报率高、员工防范意识不够导致中招率高等痛点，该智能体通过大模型与邮件网关深度集成，构建“语义解析 + 动态沙箱”双引擎。实际应用中单日研判 2000+ 封高危邮件（较人工提升 20 倍），Q3 演练 100% 检出 69 封钓鱼邮件（人工仅 45 封），9 月取样检出 14 封钓鱼邮件（人工仅 2 封），有效破解加密附件隐蔽、AI 伪装等攻击难题，重构企业邮件安全运维模式。

71.语义调度多智能体自迭代平台内容创新

申报单位：中移（杭州）信息技术有限公司

案例介绍：在数字化时代，数字内容服务面临海量运营素材制作需求，传统制作方式存在制作周期长、人力成本高、内容质量参差不齐等问

题。中移杭研案例通过构建 AetherMind 语义协同智能体平台，汇聚多个 AI 智能体，并实现基于任务的智能体 AI workflow 任务规划，构建可编排、高自动化、智能调度的内容生成-编辑一体化智能体工作平台，并创新数字内容高时效生产，高自动转换的智能运营模式。该模式创新性地将数字内容生产重构为可编排智能体协作任务，将 AI 技术深度融入内容制作全链路，实现了从创意到成品的高效转化目前服务已覆盖全国 31 省，8000 万家庭用户，月 API 调用量达 1500 万次。平台累计生成高光片段 31.7 万个、AI 海报超 10 万张，内容生成效率较人工提升 30 倍，降本增效显著。

72.政务多模态智能体：案件全流程+办公图文处理创新解决方案

申报单位：中网华信科技股份有限公司

案例介绍：传统案件处理模式因案件量激增面临效率瓶颈，人工处理证据审查、法规检索等环节耗时且易出错，亟需技术手段优化全流程。本案例以 AI 为核心推进建设：需求分析聚焦案件分析、证据鉴定等高频场景，技术选型采用生成式 AI，系统集成覆盖立案、证据鉴别的全流程模块。技术难点为深度伪造证据识别与检索精准性，创新思路为构建人机协同机制。目前系统已应用于智慧监督、智慧政法及政务办公领域，创新性体现在首创政务垂直领域大模型与一体化办案平台。

73.支付宝政务 AI 助手

申报单位：支付宝（杭州）数字服务技术有限公司

案例介绍：支付宝政务 AI 助手——晓政，率先提出“一句话办事”的下一代政务智能化愿景，推动政务服务从“人找服务”向“主动办成”跃迁。该方案构建了五大核心能力：面向高可信政务场景的白盒化智能中枢、动态演进的政务知识大脑、深度理解业务的智能帮办引擎、跨部门协同的服务引擎，以及开放可复制的智能政务生态体系。用户只需表达诉求，系统即可自动调取材料、编排流程并完成办理。晓政重新定义了“高效、可信、有温度”的政务服务标准，目前已与政府客户建设了十多个省级一网通办平台和 80 多个省市垂直部门智能体，覆盖全国超 16000 项政务服务事项，累计服务用户超 7 亿人次。

74.智能体革新 城市管理智能服务新范式

申报单位：中国电信股份有限公司大客户事业部

案例介绍：城市管理智能服务智能体依据国家智慧城市与新型基础设施建设政策，针对传统城市管理中体检效率低、社区服务碎片化、视觉治理精度不足等痛点，整合三大核心能力：城市体检报告生成大模型自动生成差异化评估报告，为决策提供依据；智慧社区智能体精准落地民生服务并收集居民诉求；城市视觉治理大模型基于 PB 级视觉数据监测服务与决策执行效果。通过智能任务调度实现三者协同运作与数据互通反哺，形成治理闭环。该智能体有效提升了城市管理的精细化与智能化水平，契合国家数字化转型战略，具备显著的行业示范价值。

75.智能物流的“超级连接器”——满帮 AI 调度重塑公路物流实时互动体验

申报单位：江苏满运软件科技有限公司

案例介绍：满帮集团开发 AI 调度系统，以应对货主传统人工调度模式的瓶颈，如找车难、效率低、成本高及服务能力有限。基于数月场景探索和满帮十多年物流数据沉淀，系统从文字沟通演进至语音消息、RTC 通话和 AI 电话，突破行业术语、司机多样性、信息非标及实时性高等难点。为货主与司机间搭建了 AI 连接器，赋能调度员聚焦复杂任务，实现 100%接通、并发处理、主动触达沉默运力，显著降本增效。应用范围覆盖运满满和货车帮 APP，日均发货量超 1.6 万票，日活司机数 1.3 万，体现高创新性、广泛影响力和活跃度。

76.智网星链：面向现实场景的具身智能体平台

申报单位：上海氮川科技有限公司

案例介绍：本项目面向驾校、文旅、医疗等传统行业企业，通过智网星链平台提供全链路解决方案，助力其数字化转型与智能化升级，有效提升运营效率与竞争力。对于 ISV 客户，其可基于该智能体平台进行二次开发，以满足下游用户的个性化需求。项目的核心创新在于独创的“基建定制+SaaS 授权”双循环模式，通过 MCP、A2A、RVP 三大关键技术，实现了工具生态、协作网络与虚实边界的融合贯通。该模式不仅深度契合国家“人工智能+”行动方针，更为智能体应用

的未来奠定了坚实基础。作为市场上首创的城市智能体虚实映射平台，其独特创新性在行业内独树一帜，引领着智能体应用的发展方向。

77. “智优”自优化智能体赋能 5G-A 高阶自智网络

申报单位：中国移动通信集团河南有限公司

案例介绍：中国移动以 2025 年达到 L4 级高阶自智网络为目标，加速网络运维数智化转型。随着 5G 网络场景越发复杂，运维挑战愈发明显。现有固定基站配置参数组合难以匹配细粒度的网络场景变化，部分时间段内 5G 网络速率体验明显劣化。基于 AI 强化学习、持续学习、安全探索、KPI 预测等先进 AI 关键技术，构建自适应、自学习“智优”无线自优化智能体，实现互操作、调度/功控、RF 等配置参数随时匹配网络环境，确保 5G 上行、下行速率最优。“智优”创新实践效果明显，应用区域日均低速率小区占比减少 25%，下行速率平均增益 9.7%。案例共荣获发明专利 4 项，计算机软件著作权 2 项，国际奖项 2 项，省通信学会奖 2 项。

78. 中国铁塔塔娃智能体平台案例

申报单位：中国铁塔股份有限公司

案例介绍：中国铁塔智能体平台是一个开源的大语言模型应用开发平台，融合 BaaS 与 LLMOps 理念，专注加速生成式 AI 应用从原型到生产的全流程。平台核心特点包括：提供低代码/无代码开发环境，通过可视化工具和预置模板大幅降低技术门槛；支持数百种专有及开

源模型，提供云服务与企业级私有化部署方案，保障数据安全；采用模块化架构，内置 RAG 引擎、Agent 框架和工作流编排等功能，支持复杂任务自动化。凭借 API 优先设计与完善的可观测性工具，平台便于系统集成与全生命周期运维管理。其开源生态、灵活扩展及企业级支持，使其成为生成式 AI 应用开发的高效“加速器”，广泛应用于智能客服、内容生成等多种场景。

79.中国移动九天大模型赋能家宽运维智能化升级

申报单位：中移九天人工智能科技（北京）有限公司（九天人工智能研究院）

案例介绍：中移九天公司（九天研究院）案例 AI 家宽装维智能应用，聚焦家庭宽带领域，采用九天网络自然语言大模型，研发“多智能体协同决策+动态流程调度+上下文感知交互”的人机交互系统，攻克了多模态意图精准识别、流程中断恢复、跨智能体通信协同等关键技术。自主研发上下文感知交互机制，通过追踪用户历史对话与场景参数，结合流程控制的智能跳转逻辑，攻克了复杂场景下的意图偏差矫正与环境感知编排难题。提供多模态语言交互、业务知识问答、现网数据核查、处理方案推荐、数据自服务五大场景支撑能力，变革综调支撑流程。上线后，问题自动处理率高达 50%，装维支撑平均时长降低了 68%，有效提高了工作效率和服务量，大幅减少了用户的等待时间。

80.自动化测试桩生成智能体

申报单位：中国移动通信集团甘肃有限公司

案例介绍：自动化测试桩生成智能体基于中国移动聚智平台与九天大模型构建，旨在解决系统集成测试痛点，提升测试案例生成效率与应用范围。该智能体针对传统静态测试桩开发成本高、维护难、周期长等缺陷，融合软负载数据采集、API 接口规范注入、大模型智能生成与 CI/CD 自动化技术，构建了从数据感知到能力部署的完整闭环。整个过程无需人工编码与手动发布，将测试桩上线周期从传统的数天甚至数周缩短至小时级，实现了“测试即代码”。通过构建“数据驱动、AI 生成、自动验证、持续部署、自我进化”的正向循环，有效提升了软件测试效率与质量，推动 DevOps 敏捷开发应用落地，显著降低了运营成本。