

2025

中国Data&AI数据 基础设施白皮书

AI创生时代，数智融合新底座

出品机构：甲子光年智库

发布时间：2025.10

在AI创生时代，Data&AI数据基础设施是新世界的底座

- 当前，全球正经历一场由地缘政治重塑和人工智能技术革命双重驱动的深刻变革。这两大趋势交织并进，正以前所未有的速度重构世界经济与社会发展格局。面对全球化向区域化、供应链向本土化转型的外部环境，以及人工智能从一项前沿技术向核心生产力跃迁的内部需求，中国正处于一个关键的历史转折点。麦肯锡研究显示，生成式AI有望为全球经济贡献约7万亿美元的价值，而中国有望贡献其中约2万亿美元，接近全球总量的三分之一。IDC的预测也印证了这一趋势，预计到2028年，全球AI领域IT总投资将增至8159亿美元，而中国的AI总投资规模将突破1000亿美元，五年复合增长率高达35.2%。这些数据表明，AI技术革命不仅是一场技术革新，更是一场将催生全新生产模式和经济增长点的大变局。
- 在这场变革中，各类组织机构，无论是企业还是政府，都面临着巨大的挑战和机遇。埃森哲的研究表明，中国企业正经历一场“挤压式转型”，转型窗口期更短，挑战更多。传统的、孤立的数据系统（如数据仓库或割裂的数据平台）已难以支撑AI应用对实时性、多模态数据处理和高弹性算力的需求。企业迫切需要打破数据孤岛，解决高质量数据供给不足、模型与业务场景割裂以及数据安全合规等痛点。这些挑战正倒逼数据基础设施实现一次范式跃迁，从单一的数据存储与分析工具，向支撑全业务流程的智能化底座演进。
- 数据厂商要想抓住机遇并在新一轮竞争中脱颖而出，关键在于构建一种全新的数据基础设施。这种基础设施应具备以下核心能力：
 - 必须将组织内部的数据生产、加工治理与业务终端AI模型的应用视为一个动态持续的生产过程，以满足业务的动态连续性需求。
 - 需要摒弃当前许多企业采用的单场景Agent应用建设模式。正如传统数据库时代的“数据烟囱”，这种模式将导致数据和能力再次割裂，无法形成全局协同效应。Gartner预测，到2028年，至少15%的日常工作决策将由Agentic AI自主做出，这要求基础设施必须能够支持大规模、跨场景的智能体部署。
 - 一体化的数据基础设施需要具备集中式建设和集中式管理的能力基盘，从而确保数据资产的统一治理、安全合规与高效利用。同时，它也必须面向业务侧即价值端，形成分散式的赋能能力，让各业务单元能够灵活调用和创新。
- **在技术层面，新型数据基础设施的核心路径是“湖仓一体”与“AI原生”的深度融合。它将数据湖的开放性和灵活性与数据仓库的结构化管理、ACID事务特性相结合，形成一个统一的数据底座。同时，它将AI模型的训练、部署与数据治理紧密耦合，实现从数据到智能的无缝流转。这种架构不仅是技术上的整合，更是理念上的革命，它为Data&AI深度融合体提供了坚实基石。支撑数据价值逐级跃升，从“点”（功能提效）到“线”（组织创新），到“面”（行业协同），再到“体”（产业带动），最终形成支撑未来智能社会的国家战略，成为智能革命的核心引擎。**
- 本白皮书旨在系统梳理中国Data&AI数据基础设施的发展现状、技术路径与实践经验，为大型组织提供前瞻性参考与可行性指南。人类开启信息化革命后的几十年里，全球竞争格局一直都由美国引领。然而，在进入AI新时代后，中国持续追赶的成果逐渐显露，尤其在通信、互联网、人工智能等领域出现了引领全球的创新技术产品和服务，而Data&AI数据基础设施正处在中国这些优势领域的交叉地带，必将成为中国在AI竞争中持续保持优势、赢下智能革命的必争之地。我们相信，当数据与AI在新底座的支撑下真正融合一体，新世界的无限可能便将由此开启。

数据应用与人工智能从相互独立到逐渐融合

- 在数据作为生产要素被广泛认知后，数据应用和人工智能加速融合。高质量数据集的积累加快了大模型的成熟，大模型又反过来让各组织机构更好地发挥数据要素的生产力，开始从深度融合走向一体化。

数据应用与人工智能的融合，推动架构与能力实现跨越式发展

- 第一阶段的感知智能主要依赖分类的非结构化数据，第二阶段生成式AI推动多类混合数据的治理体系发展完善，到了第三阶段的Agentic AI则对多模态数据融合与动态调度提出更高要求。数据应用与人工智能技术正从“AI+数据”的松散组合迈向“数据即AI、AI即平台”的深度融合阶段。
- 新阶段需要一个AI原生的、一体化的技术底座平台，为Agentic AI的场景化应用提供持续的工程化能力和稳定的技术底座支撑，类似建筑的地基工程，为组织长期提供多模态数据融合处理、动态异构资源调度，构建“数据-模型-Agent-业务”闭环系统。这个平台我们称为Data&AI一体化基础设施，它正飞快成为新一代智能系统的核心，围绕“数据流+Agent流”帮助组织构建闭环、自主的智能生态。

数据基础设施成为AI时代的“核心生产工具”

- 数据基础设施逐步从“支撑决策”升级为“协同驱动智能”，成为组织机构实现数据价值流转与智能闭环的“核心生产工具”。
- 其中，Data&AI一体化数据基础设施是最终能够长期持续支撑人工智能规模化落地的一体化基础软件平台，不仅能提供基础底座支撑，更重要的是通过平台持续、实时、动态工程化能力实现打通数据存储、治理、计算与AI模型开发的全链路，构建出了“Data for AI”和“AI for Data”的双向赋能体系。

Data&AI数据基础设施的价值链条呈现“点-线-面-体”递进

- Data&AI数据基础设施通过湖仓一体架构统一纳管全域数据资产，并以AI原生设计支撑模型的高效训练与实时推理，实现数据到智能的闭环转化，驱动业务持续创新，进而支撑数据价值逐级跃升，从“点”（功能提效）到“线”（组织创新），到“面”（行业协同），再到“体”（产业带动），为数字经济高质量发展提供软件技术基础设施，是智能革命的核心引擎。

Data&AI数据基础设施的能力体现在开发、架构、调度、AI原生和安全运营

- Data&AI数据基础设施让数据平台和AI智能工具融合成为统一系统性底座，其关键能力主要体现于融合开发、架构、资源调度、AI原生和安全运营五个方面。五个能力方向侧重点不同，各厂商在其专业方向上各有侧重。