

三端协同整体解决方案



PART 01

三端协同



市场发展现状

政策与合规

国家大力推动算力基础设施均衡布局
数据安全与隐私保护法规强制要求“数据本地化处理”

用户对体验的极致追求

低延迟响应：自动驾驶/工业质检/AR/VR 等场景要求毫秒级响应，纯云端处理因网络传输无法满足。

行业走向

截至到2024年，将在全球孕育出超过5亿个新应用，AIGC应用正在慢慢覆盖各行各业

技术发展现状

技术落地困难点

人少

既懂算法又懂应用开发的AI开发人员稀缺

技术要求高

模型训练、压缩、多端部署技术复杂，调试周期长；

多方沟通整合时间
阔度大整理

业务、算法、运维、厂商多方拉
通耗时耗力

若无统一平台支撑，单靠手工“拼凑式”开发，不仅效率低，还难以保证系统稳定性和可维护性

AI应用落地的最后一公里仍然面临诸多挑战，全流程链路打造是企业实现快速AI落地的关键。

三端协同方案

RoycomONE智算一体机解决是针对**模型推理应用**，**模型微调处理**，**大规模数据训练**场景一体化解决方案；围绕**边-局-中心”三端协同**架构无缝打通全域算力与应用场景深度契合，实现应用开箱即用，资源按需扩展、为用户提供AI大模型落地应用三位一体联动工作的可视化一体式的智算管理功能，为AI大模型行业化私有化应用落地提供强大的算力支持！

边端一体机-文昌星



智用一体机

局端一体机-天枢星



智推一体机



智训一体机

数据中心-紫微星



数据中心

RoycomONE家族

文昌星

天枢星

紫微星

智用终端
智算一体机
智算一体机
数据中心



开箱即可应用
一键生成能力



一键生成大模型
一键部署大模型
一键调用大模型



资源按需扩展
训推任务管理
监控运维管理



远程数字化管理
模块化系统设计
一体化交付能力

通用类

知识库问答 PPT生成 文件分析 旅游规划

华为 Ascend

寒武纪 Cambricon

沐曦 META

海光 HYGON

燧原 Enflame

天数 天数智芯

DeepSeek

Qwen3

QWQ

Llama

ChatGLM

...

内置训练推理环境

内置开发环境

GPU资源切分

单机单卡，多机多卡

模型微调

优先层级

算力切分

集群管理

资源调度

权限管控

资源管理

异构管理



配电系统
网络布线
空调制冷

消防系统
机柜组合
应急通风

监控中心

三端协同架构图

文昌星

通用类：知识库问答，文件分析，旅游规划

行业类：合同审核，金融看板助手，中医问诊，公文写作，教学辅助agent。。。

安全防护中心：数据加密，APTGateway，服务保护

天枢星
智算一体机

轻量微调推理

微调框架：LLaMA-Factory, DeepSeek, PEET等

推理引擎中心：VLLM, Ollama, SGlang。。。

数据处理：数据标注，数据清洗，算法优化

基础
行业
大模型

应用开发

模型接入

Prompt提示词工程

工具插件定义

零代码/低代码开发

可视化工作流引用

RAG/检索增强生成

MCP/API

紫微星
数据中心

硬件设施层

异构算力池

智控中枢系统

存储矩阵池

可信安全域

超融网络域

智控散热中心

可靠
高效

任务训练层

KVM

docker

Kubernetes

深度学习

大规模训练

集群资源中枢

AI/HPC任务

高并发吞吐

三端协同

打造 “大模型再训练与数据存储
(中心端) - 重型推理与轻型再训练算力支撑 (局域端) - 场景应用
(边缘端) ” 的全链路闭环

方案理念

定制化

面向行业及场景，从基础算力到应用落地的全能定制！个性化配置适配不同行业、不同规模用户的独特需求，实现快速部署与持续迭代。

开箱即用

算力+模型+调优+Agent+应用
一站式AI部署应用落地！图形化界面与向导式操作，关键功能一键直达，配置简单灵活，降低使用门槛，提升用户体验。

运维无忧

一体化交付，实现全栈服务的保驾护航；解决了 AI 技术落地中的硬件适配难、部署成本高、行业适配弱痛点，让AI更简单

AI 更简单，就用
RoycomONE

PART 02

产品介绍



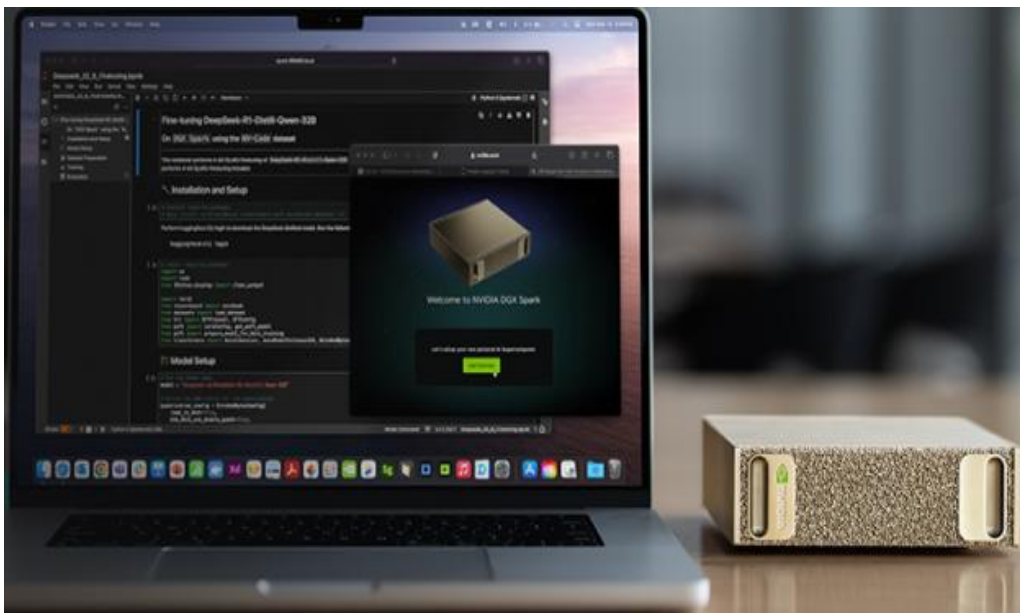


文昌星-
智用终端

文昌星智用终端

文昌星智用终端是针对企业个人用户打造**桌面一体化**智能终端解决方案；通过调用天速星模型与应用接口能力，实现终端软件产品的一键应用，以**定制方式**为企业提供AI场景的线路，根据不同的行业场景针对集成丰富智能化服务应用引擎，实现用户快速开机即用，推动 AI 与更多垂直领域的深度融合！

文昌星-智用终端拟形态



算力盒子

1 petaFLOP AI性能
显存内存共用机制

深度学习
集成工具框架

应用于需要多模态模型辅助大语言模型的场景
如：低空经济，天文，海洋



无算力盒子

软件APP应用

基础存储构建

应用于仅需要大语言模型交互的场景
如：政务公文，教育，金融

智用终端应用场景

已有场景

法意盒子

政务公文撰写

公文文件纠错

公文文件对比

合同文件审核

政务素材

法律知识库

天文盒子

天文模型知识对话

脉冲分析

光谱研究科学

教育盒子

小模型科研

AI数字人教育

海洋盒子

海洋模型问答

海洋环境检测

其他盒子

医疗问诊盒子

金融数据分析

规划场景

法律教学平台

风险点提示

合同自动化执行

.....

星系推测

行星轨迹分析

星图照片分析

行星图片采集

.....

AI辅助教学助手

个人学习成长盒子

改卷识别盒子

.....

水下数据采集盒子

洋流轨迹分析

智慧渔业

.....

挖掘中



天枢星-
智推一体机

智推一体机解决方案

RoycomONE智推一体机是针对AI大模型推理场景一体化解决方案；为企业是一套全面、易用且高性能的大模型推理基础设施，通过异构算力调度，动态扩缩容及多引擎优化技术专注于提升大模型推理效率与资源利用率，解决在复杂且异构的计算资源环境下高效、稳定地运行和管理各类大型 AI 模型所面临的挑战，并为上层应用提供统一、便捷的接口，加速生成式 AI 应用的落地。

通用应用场景

知识库问答

知识库问答:上传本地数据，构建本地库进行问答



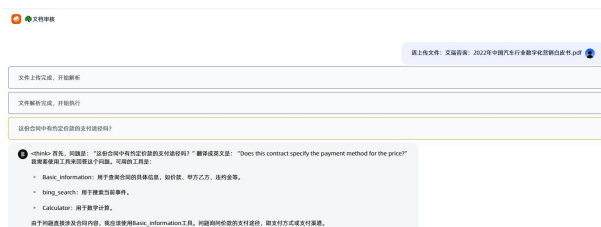
文稿撰写

文稿撰写: 基于输入内容要求进行撰写，如论文文章，演讲稿等



文档审核

文档审核: 提前设定要求，用户上传文档进行分析



大模型问答

大模型问答: 调用模型本身能力进行回答问题



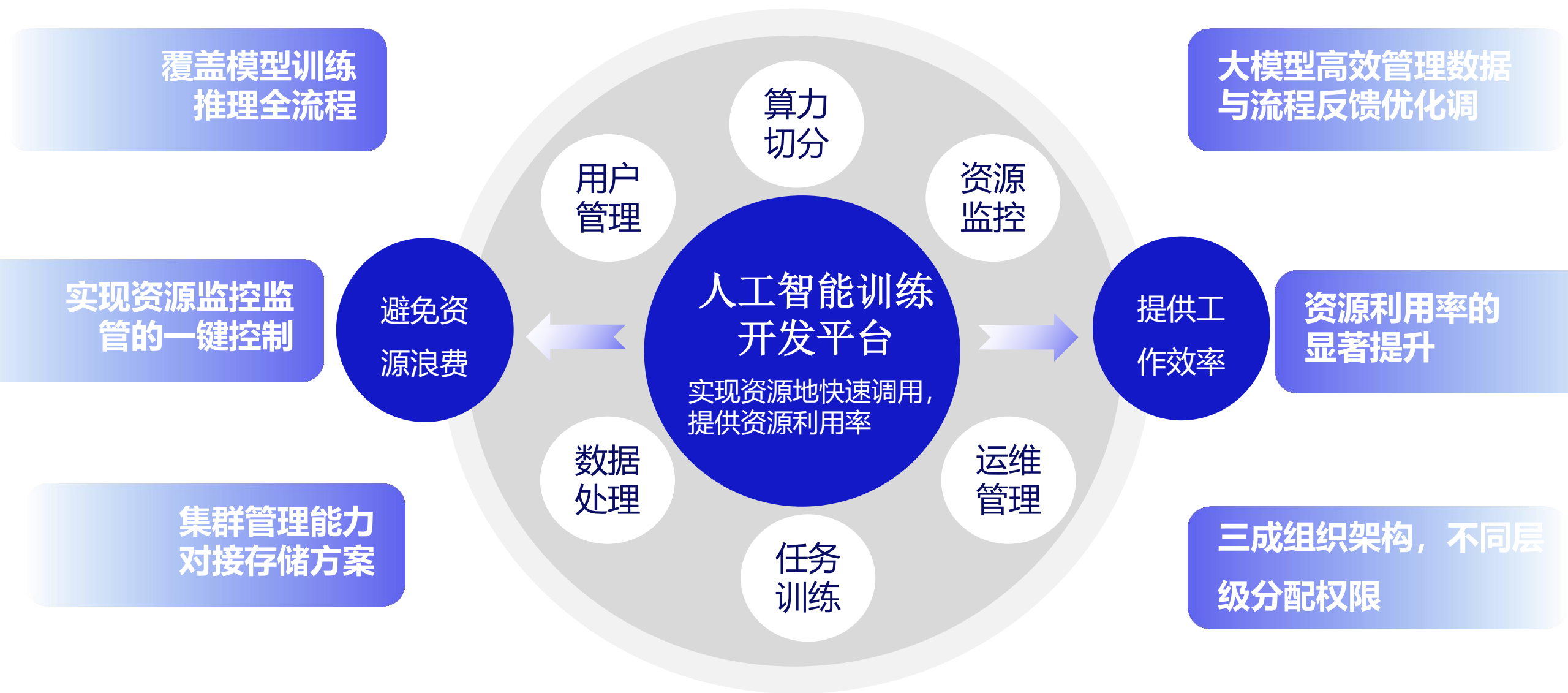


天枢星-
智训一体机

智训一体机解决方案

RoycomONE智算一体机是针对AI大模型训练推理场景一体化解决方案，资源按需扩展、算力切分，开机即插即用，能够为用户提供AI大模型训推任务管理、系统资源管理、监控运维管理等可视化的一体式的智算管理功能。内置多个模型，多种框架，让用户能够高效开展AI大模型训练开发调优、推理部署应用、数据存储管理等任务，为AI大模型行业化私有化应用落地提供强大的算力支持！

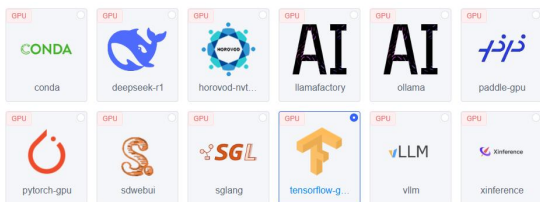
智训一体机-产品能力



通用应用场景

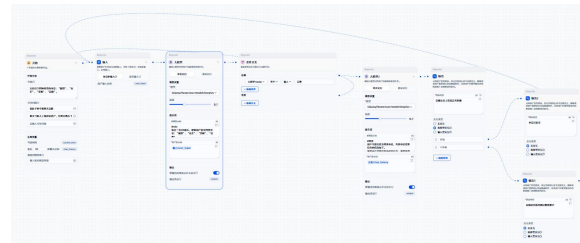
模型微调

通过内置的训练框架，可基于页面方式实现



应用开发

可通过工具流应用搭建智能体应用，结合AI应用实现场景落地



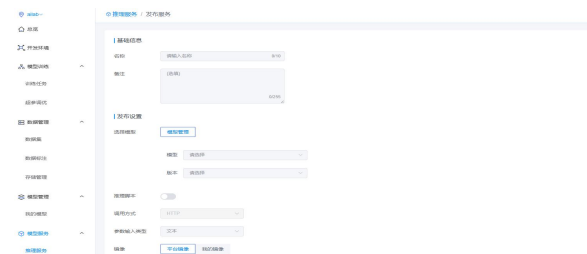
集群管理

集群管理算力资源，异构统一管理



API接口调用

模型部署服务的快速接口调用





紫微星-
数据中心

紫微星数据中心解决方案

RoycomONE数据中心方案，是集中存放计算、存储、网络设备并提供高可用、高安全、高可靠IT服务的核心基础设施；提供**大规模模型再训练与海量数据存储**，精准覆盖“边缘端感知交互推理 - 局域端算力支撑 - 云端大算力及存储调度”全场景需求，为客户提供快速响应的算力支持~

产品特点

更高效的电源系统设计：

在线模式UPS效率高达96%，可降低7.9%

模块化的精确配电设计：

模块化配电设计，确保IT按需部署

高效换热：

- 近端送风，全EC变频设计，按需制冷
- 高回风设计，较传统空调节能14.6%



模块化按需分步构建

先比传统的预先构建方式，按需分布构建可以节省成本并保持效率

TCO 通过分步部署节省的费用

总拥有成本节省 33%



38%

基建费用



19%

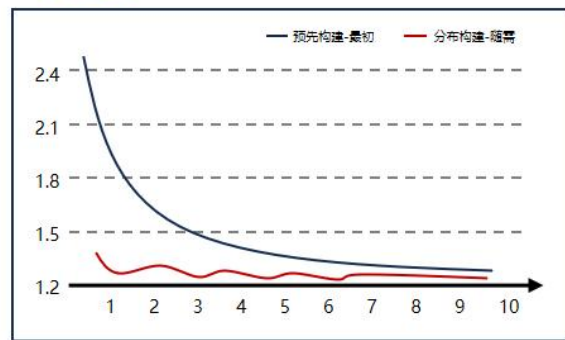
运营支出



44%

非能源运营支出

PUE



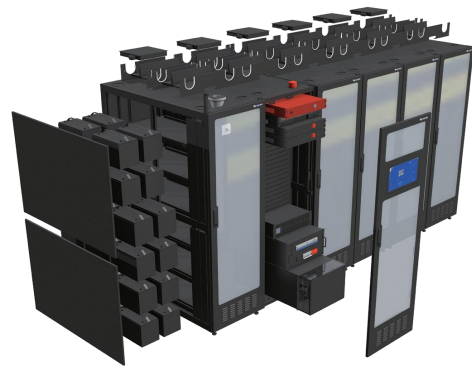


单柜智能微模块

- 独立冷热通道，高回风设计，不影响周边设备
 - 前门真空玻璃+精密空调采用高精度控制
 - 10寸触摸智能化监控屏幕
-

- 采用湿膜加湿方式，相比与电极加湿和远红外加湿节能约90%
 - 应急风扇和自动弹开门，达到设定高温阈值自动开启，辅助散热
 - 支持本地、远程Web界面访问
-

单排智能微模块



双排智能微模块

- 智能控制自动调节冷量输出，降低系统能耗
- 各模块与机柜间采用“热插拔”技术,所有模块可以在线热插拔
- 高清门楣显示屏搭配21寸大屏，制冷链路信息全局监控可视

数据中心-监控能力

