



DeepSeek-R1

企业级独立部署托管



1 概要

英智 DeepSeek-R1 的企业级独立部署，为用户提供独享的高性能计算资源，确保数据安全性与业务稳定性。通过独立部署，企业可完全掌控服务器的安全配置与访问控制，有效规避共享托管环境中的多租户安全隐患。同时，独立部署架构支持高并发请求与大吞吐量数据处理，无任何性能限制，满足企业大规模 AI 推理与训练需求。



2 应用场景

依托英智强大的智算算力，客户可以直接租用高性能算力，通过高速网络（专线/深港专线）快速对接企业内应用，开箱即用。企业现有应用可通过独享 API 接入 DeepSeek，实现智能化升级；让已有应用能具备通过独享API接入DeepSeek的能力。开发、测试及产品设计团队可基于 DeepSeek 协同构建 AI 驱动的工作流，提升办公效率；同时，企业还可基于 DeepSeek API 开发独立的 AIGC 应用，如AI知识库、AI售后、AI智能客服、AI培训、AI文案生成和智能设计等，全面赋能业务场景。



3 核心亮点

联网/联网搜索：支持实时联网/联网搜索，帮助企业快速获取互联网信息，增强数据获取与分析能力。

RAG 知识库：集成 RAG 技术，支持企业将海量知识库与生成式 AI 结合，实现精准信息检索与内容生成，提升知识管理效率。

全闪存储挂载：支持高性能全闪存储挂载，提供超高速的数据读写能力，适用于大规模数据处理与高吞吐量场景，显著提升存储与计算效率。

专线/深港专线：提供国内点对点专线与深港跨境专线接入，国内专线优化本地数据交互，深港专线支持跨境业务，满足企业全球化需求。

开源 Agent 应用：兼容多种开源 Agent 应用框架（如 OpenWeb UI、Maxkb、英智玲珑等），为企业提供灵活的应用集成选项，降低开发与部署成本。

4 服务支持

高性能配件支持：提供多种高端硬件配置选项，包括 H100、4090、H20 和 910B，支持高并发计算与大规模 AI 推理任务，确保卓越的性能与稳定性。

高速网络接入：通过国内点对点专线与深港跨境专线，实现低延迟、高带宽的数据传输，优化本地与跨境业务的数据交互效率。

免运维托管服务：英智提供全托管运维服务，涵盖硬件监控、故障排查与性能优化，企业无需投入额外运维资源。

开箱即用部署：支持快速部署与无缝对接企业现有系统，提供预配置环境与自动化安装工具，大幅缩短上线时间。

SLA 服务保障：严格遵守服务级别协议（SLA），保障业务连续性与数据安全。

专家技术支持：英智专家团队提供 7x24 全天候技术支持，包括 API 集成指导、性能调优与故障应急响应，确保系统高效稳定运行。

5

企业级独立部署托管报价

模型	机器型号	显存	内存	单机支持人数	单位	价格
DS 671B	H100	640GB	2TB	500-1000人	台	7.5W/月
DS 671B	H20	768GB	2TB	300-500人	台	3.2W/月
DS 671B	910B	1TB	2TB	300-500人	台	4.4W/月
DS 70B	H100	640GB	2TB	5000-12000人	台	7.5W/月
DS 70B	4090	192GB	512GB	300-600人	台	8500/月
DS 32B	4090	192GB	512GB	500-1000人	台	8500/月

6 机房实拍



联系方式

联系电话：400-9090-913

QQ：3799595584

邮箱：BD@ai2nv.com

地址：深圳市南山区听海大道4035号
前海恒昌大厦东座1702



扫描上方二维码添加微信