

STONE TEK G5208 AIO Qianfan一体机

大模型产业变革带来的挑战

企业端挑战



数据重要
无法上云



知识散乱
效率低下



部署困难
成本高昂



应用构建
集成复杂

大模型创新需打造一套私有化软硬一体 + 企业知识赋能 + 低代码应用开发+ 行业生态协同的全链路解决方案

供应端挑战



环境繁杂
交付迟缓



模型杂乱
管理失序



生态薄弱
落地艰难



运维高耗
盈利承压

G5208 Qianfan

G5208 Qianfan产品特性

百度千帆大模型Lite平台

大模型管理与调用

Agent 开发

知识库构建

更多应用...

G5208 Qianfan



G5208 PCIE5 Qianfan



百度千帆 × STONE TEK



开源模型应有尽有



轻量化开发便捷高效



私有化部署开箱即用

海量开源模型预置，高效部署

Agent应用开发与知识库构建，轻量易用

百度自研千帆Lite平台，私有化部署，安全合规

百度千帆大模型Lite平台

产品亮点

「轻」

- 一键部署，从上电到上线以“小时”计算
- 部署即拥有可用的AI大模型平台，降低探索成本

「省」

- 单机形态、免集群，显著降低环境搭建与运维复杂度
- 用有限预算覆盖主流AI模型推理、应用开发需求

「稳」

- 企业级管理能力，精细化监控。
- 私有化部署，保障数据安全和行业合规性；内置百度自研内容安全围栏，符合TC260-003标准

「开箱即用」

- 内置模型、知识库与Agent构建能力

低成本 · 快交付 · 安全可控

大模型平台
性价比之选

模型常新 · 应用共生 · 生态放大

产品架构

基础管控

监控看板

应用层：开箱即用

应用开发

Agent

Agent开发 指令优化

知识库

知识库检索 知识库增强 解析切片

体验测试

知识库关联 prompt交互测试

调用统计

模型层：兼备大小模型

模型推理

模型管理

模型广场 模型管理

模型服务

生成式AI 判别式AI 重排序 向量表征 第三方

预置模型

生成式AI模型（大模型）

ERNIE ChatGLM系列 Baichuan系列 Llama系列 ...

判别式AI模型（小模型）

分类 检测 自定义

算力层：适配算力

算力服务器

产品优势

模型多样性

使用便利性

安全性

企业级软件

丰富的生态应用

成熟案例

帮助企业快速迭代

支持一键部署

安全算子保障客户使用安全

支持报表统计等应用

后续支持多种不同生态应用集成

成熟产品，具备各行业落地经验

公司总部：深圳市南山区滨海大道3378号盈峰中心16F

公司总机：0755-26903345
技术支持：400-0151-918

http://www.stoneholdings.cn

