

企业AI资产管理解决方案

多模型聚合· 权限管控· 额度分配· 消耗统计· 智能分析

2026/04

深圳领驭科技有限公司

Shenzhen LEINWIN Technology Co., Ltd

LEINWIN
领驭科技

领驭科技

人工智能与智能制造领域先锋企业
专注于为企业提供前沿的数字化解决方案

90%

管理者工业软件
行业经验超15年

70%

技术研发
人员占比

“专精特新”企业

CS1/CMMI3级认证

科创型中小企业

ISO9001/27001认证

50+ 软件著作权/发明专利

300+ AI项目案例

深圳市瀚鹏 人工智能科技有限公司

聚焦工业AI的深入应用与行业落地

在人工智能方向，领驭科技自主研发了瀚鹏AI应用解决方案，融合多模态大模型、知识库、智能体等核心能力，广泛应用于快销品、高科技电子、智能装备等行业

苏州领驭 信息技术有限公司

专注于半导体行业，为头部制造企业提供本地化支持与交付能力

在智能制造方向，公司以工业软件为核心，构建涵盖数字化工厂、系统集成、工业数据分析与AI算法的综合解决方案，已服务面板半导体、3C电子、汽车装备等多个领域

新加坡&香港 分公司

响应客户全球化布局

具备东南亚区域项目实施与交付能力，积极推动中国企业出海，用数字化能力助力全球制造。

发展历程

LEINWIN

公司使命：以人工智能与数字智造加速全球企业生产力升级

核心价值观：创新 高效 专业 分享

领驭1.0 IT信息化集成服务商

- 专注于MOM和IT信息化集成，奠定行业基础
- 西门子工业软件合作伙伴，深化数字化工厂解决方案

领驭2.0 2020年-2024年 数字化转型服务商

- 扎根智能制造，开启工业系统“自主研发元年”
- 深耕行业，打造行业解决方案，沉淀Know-How经验
- 连续多年获得西门子匠心服务奖
- 微软(中国)最佳出海云服务解决方案合作伙伴

领驭3.0 2025年 企业智能体全栈服务商

- 瀚鹏工业AI平台荣获深受企业欢迎的制造业信息化优秀产品
- 加大研发投入，华东研发子公司正式启航
- 成立海外业务中心，助力中资企业出海
- 全面布局AI，携手企业共创AI+场景

荣誉资质

领驭已获得国家高新技术企业、专精特新企业
知识产权管理、信息安全管理、质量管理等国家体系认证
发明授权5项，申请中15项、软件著作权40+项
并通过了CMMI3级认证。

微软相关资质

HKCSP 1T、21V NCEI、21V 2T CSP

MAICPP: Infrastructure、Data & AI、Modern Work、Digital & App
Innovation

ASP: AI Platform on Microsoft Azure、Infra and Database Migration
to Microsoft Azure



企业AI资产管理挑战

AI模型快速爆发，企业面临全新的资产管理难题

背景

LEINWIN



安全风险

多模型接入缺乏统一管控，数据泄露风险高，权限边界模糊



成本失控

各部门分散采购AI服务，Token消耗无监控，预算超支频发



模型孤岛

不同供应商模型各自为政，无法统一调度与横向对比



使用黑盒

缺乏使用追踪和审计日志，AI调用行为不可追溯



运维复杂

多模型API管理复杂，版本迭代、故障切换难以自动化



合规压力

数据隐私法规趋严，模型使用需满足合规审计与追溯要求

方案定位与愿景

一站式AI资产治理中枢，让每一笔AI投入都可衡量、可管控、可优化

定位

LEINWIN

MISSION

构建企业级AI资产 统一治理中枢

整合企业所有AI模型资源，提供从接入、管控、计量到分析的全生命周期管理能力，实现AI资产的可见、可控、可优化。



全可见

多维度数据看板
实时掌握AI资产全貌



全可控

精细化权限与额度
AI使用安全合规



全可优

智能分析与推荐
持续优化AI投入产出

产品架构全景图

架构

LEINWIN

四层架构设计，从接入到决策的完整能力链路



多模型聚合网关

核心能力

LEINWIN

统一接入层，屏蔽多厂商差异，实现模型资源的无缝聚合与调度



统一API接入

一套API标准对接所有模型，消除多厂商接口差异



智能路由引擎

基于任务类型、成本、延迟等多维度自动选择最优模型



负载均衡

多实例自动分发请求，保障高并发场景下的稳定性



故障自动切换

模型服务异常时秒级自动切换至备用实例，业务零感知

模型能力对比矩阵

模型	推理	创意	代码	多模态
GPT-5.4	★★★★★	★★★★☆	★★★★★	★★★★★
Claude 4	★★★★★	★★★★★	★★★★☆	★★★★☆
GLM-4	★★★★☆	★★★★☆	★★★★☆	★★★★★
Qwen	★★★★☆	★★★★☆	★★★★★	★★★★☆
DeepSeek	★★★★★	★★★★☆	★★★★★	★★★★☆

AI模型权限管控体系

基于RBAC模型的细粒度权限管理，确保AI资源安全有序使用

权限管控

LEINWIN

超级管理员

全局配置 / 全部模型 / 无限额

部门管理员

部门模型 / 成员管理 / 额度分配

普通用户

授权模型 / 个人额度 / 日志可查

审计员

只读权限 / 全量日志 / 合规报告



模型级别授权

可精确到单个模型版本的访问权限控制



部门/团队隔离

按组织架构划分模型资源池，互不干扰



审批 workflow

高权限模型申请需经审批，防止滥用



异常行为告警

实时监控异常调用模式，自动触发告警

额度分配与管控

额度管控

LEINWIN

多维度的Token额度管理体系，精细化控制AI使用成本

5,000万

企业总Token池

12

分配部门数

78%

平均使用率

¥32.5万

本月节省成本



部门级配额

按部门业务需求分配月度Token额度，支持动态调整和超额预警



用户级配额

针对个人设置每日/每周调用上限，防止资源垄断与滥用



模型级配额

对高成本模型设置独立配额，引导用户优先使用性价比模型



智能预警

使用率达80%自动通知管理员，达95%触发熔断保护

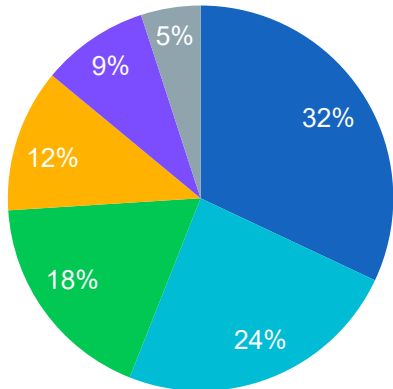
智能体使用统计

全方位追踪企业内部AI智能体的使用情况与效能表现

使用统计

LEINWIN

智能体使用分布



■ 代码助手 ■ 文档写作 ■ 数据分析
■ 客服机器人 ■ 翻译工具 ■ 其他

热门智能体排行

排名	智能体	调用量	满意度
1	代码助手 Pro	12,450	96%
2	智能写作官	9,830	94%
3	数据分析师	7,220	93%
4	客服精灵	5,610	91%
5	多语翻译官	4,350	95%

满意度趋势



Token消耗统计与管控

实时监控Token消耗趋势，智能分析成本结构，驱动成本优化决策

LEINWIN

瀚鹏AI资产管理平台

销售AI资源账户

当前余额
\$1900.04

历史消耗
\$115.95

销售智能体使用统计

请求次数
1842

统计次数
727

模型Token消耗

统计额度
\$281.25

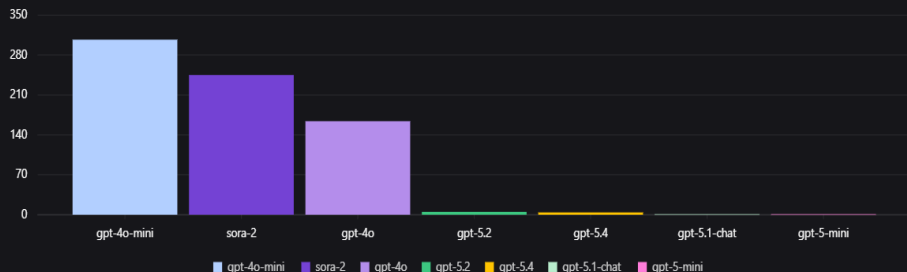
统计Tokens
6,624,585

模型数据分析

消耗分布 / 消耗趋势 / 调用次数分布 / 调用次数排行

模型调用次数排行

总计: 727

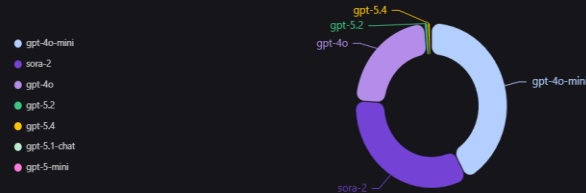


模型数据分析

消耗分布 / 消耗趋势 / 调用次数分布 / 调用次数排行

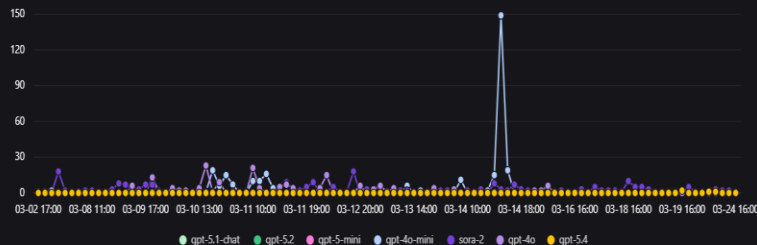
模型调用次数占比

总计: 727



模型消耗趋势

总计: 727



部门AI使用日志与分析

全链路审计追踪，多维度交叉分析，满足合规与运营需求

LEINWIN



全量调用日志

记录每一次AI调用的完整上下文：用户、时间、模型、输入摘要、Token数



合规审计报告

一键生成部门/个人级别的AI使用报告，满足内部审计与外部合规要求



使用趋势分析

按时间维度分析各部门AI使用趋势，识别增长热点与异常波动



部门对比看板

横向对比各部门的AI使用效率、成本与价值产出，辅助资源优化

瀚鹏AI资产管理平台

销售AI资源账户

当前余额
\$1900.04

历史消耗
\$115.95

销售智能体使用统计

请求次数
1842

统计次数
727

模型token消耗

统计额度
\$281.25

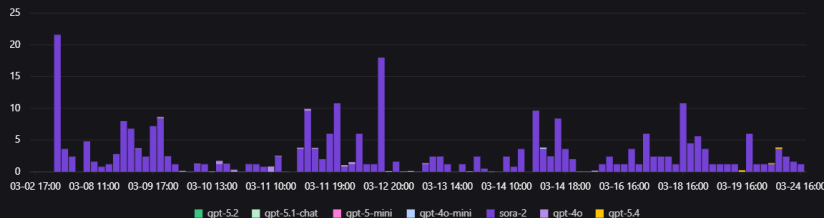
统计tokens
6,624,585

模型数据分析

消耗分布 / 消耗趋势 / 调用次数分布 / 调用次数排行

模型消耗分布

总计: \$281.25



数据安全与合规保障

LEINWIN

端到端的安全防护体系，满足企业级数据隐私与合规要求



数据脱敏传输

敏感信息在发送至AI模型前自动脱敏，支持正则与语义识别双引擎



私有化部署

支持完全私有化部署，所有数据不离开企业网络边界



审计追溯

全量操作日志留存180天+，支持按需导出与取证分析



加密存储

AES-256加密存储所有模型配置、用户数据与审计日志



合规报告

内置等保、GDPR等合规报告模板，一键生成审计材料



灾备容灾

主从架构+异地灾备，RPO<1分钟，RTO<5分钟

典型应用场景

LEINWIN

覆盖企业核心业务场景，AI资产管理平台释放组织生产力



研发团队

效率提升40%

统一管理代码生成、代码审查、技术文档等AI工具的使用权限与消耗



产品运营

内容产出3倍

为内容创作、用户分析、A/B测试等场景提供智能体与模型资源支持



客户服务

响应速度60%

管理客服智能体的规模与质量，监控响应效果与用户满意度



数据分析

分析效率50%

支持数据清洗、报表生成、趋势预测等AI辅助分析工具的统一调度

核心技术优势

LEINWIN

企业级架构设计，高可用、高性能、高安全

01 高性能网关

单节点支撑10万+QPS，P99延迟
<150ms，异步流式输出

02 高可用架构

微服务+K8s容器化部署，自动扩缩容
, 99.99%可用性

03 低代码集成

丰富的SDK与API，5分钟完成接入，
零侵入业务系统

04 智能降本

基于使用模式自动推荐最优模型组合
, 降低30%+Token成本

05 开放生态

支持私有模型接入，插件化架构便于
功能扩展

06 实时分析

亚秒级数据处理管道，实时计算消耗
与成本指标

实施路径规划

分阶段实施，快速见效，持续优化

LEINWIN



Thank you!

开启企业AI治理新时代

让AI资产可见、可控、可优化



400-8889-356

www.leinwin.com



深圳领驭科技有限公司

SHENZHEN LEINWIN TECHNOLOGY CO.,LTD

广东省深圳市南山区粤海街道国人通信大厦 9F

9F, Guoren communication building, Yuehai street,
Nanshan District, Shenzhen City, Guangdong Province