

如何开始

Start with one clear use case

先选一个具体场景

明确谁在用、解决什么问题、需要连接哪些数据，再确定模型、算力和部署方式。

01 场景确认

明确目标、用户和业务流程

02 方案设计

确定应用、模型、数据与算力

03 部署上线

完成开发、接入、测试和交付

04 持续优化

根据使用数据迭代效果与成本

咨询部署



扫码沟通具体需求

电话：027-87805210

手机：18627001072

网址：xserver.cn

可提供现有系统、数据类型、使用人数及目标场景，我们据此给出部署建议。

先验证价值，再按需要扩大部署范围。

部署需要多少成本

Cost structure

成本由两部分组成

一次性部署成本 + 持续运行成本。规模越大、数据要求越高、私有化程度越深，投入越高。

一次性部署成本

需求梳理、应用开发、知识库建设、系统对接、测试与上线。

持续运行成本

模型 Token、GPU 算力、存储带宽、运维支持和后续扩展。

三种常见投入方式

轻量验证

调用现成模型与云端算力

投入较低，上线快

标准部署

应用开发 + 知识库 + 系统接入

兼顾效果与成本

私有部署

专属模型、服务器与独立环境

控制力更强，投入较高

成本可控的关键

先做小场景验证，再按实际使用量扩展模型、算力和用户规模。



AI 应用端部署

把 AI 真正部署到业务里

不是只提供模型，而是完成场景梳理、应用开发、系统接入和部署交付。



从一个场景开始，让 AI 产生结果

需求梳理 / 应用开发 / 部署上线 / 持续运维

获取部署方案与成本测算

电话：027-87805210 | 手机：18627001072

网址：xserver.cn

我们做什么

AI application deployment

把 AI 部署到应用端

连接模型、数据、算力和现有系统，形成可直接使用的 AI 应用。



01 场景梳理

明确问题、用户和使用流程

02 应用开发

智能问答、知识库、助手与自动化

03 系统接入

连接业务系统、文档、数据库和接口

04 部署交付

支持云端、本地及私有化部署

最终交付的不是概念

而是能够登录、使用、管理和持续优化的 AI 应用。

可以部署到哪些领域

Deployment scenarios

从办公到科研，从服务到研发

优先选择数据明确、流程重复、结果可衡量的场景。

政府

政策知识库、材料辅助、政务问答

高校

教学助手、科研检索、知识服务

研究所

文献分析、数据处理、算力平台

企业

办公助手、客服、营销与经营分析

软件开发

代码生成、测试、文档与研发辅助

信息技术服务

系统集成、智能体与项目交付

常见应用形态

知识库

智能问答

业务助手

流程自动化

数据分析

研发辅助

一个原则

不追求“大而全”，先把一个高频问题真正解决。

部署后有什么好处

Business outcomes

同样的工作，投入更少

以下为知识检索、材料整理、内容生成等重复型工作的参考对比。



传统方式与 AI 部署后对比

对比指标	传统方式	AI 部署后	变化
人员成本	100%	65%	-35%
时间成本	100%	40%	-60%
任务准确率	85%	95%	+10个百分点
结果一致性	70%	95%	+25个百分点
权限可控覆盖	60%	95%	+35个百分点

参考目标值；准确率和安全性会受数据质量、权限方案及人工复核机制影响。

一年可以节省多少？

约 13.6 万元 / 年

10 人 × 1.2 万元/月 × 12 月 × 30%重复工时 × 50%效率释放

= 21.6 万元等效人力价值 - 约 8 万元年度部署与运行成本

说明：以上为参考测算，不含因响应提速、服务增量带来的额外收益。实际节省金额以岗位成本、工作量、使用率和部署方案为准。