

OpenClaw & Hermes

一次关于能力与落地的私享会

宣讲人：周云飞 公司：浙商中拓

我是谁，为什么站在这里



拥有十八年 IT 从业经验的老兵 + 终身学习者

Microsoft / DELL / 紫光 / Authing



自媒体频道：马丁看世界 @martinz2025

科技 / AI 观察，每周输出



实战派

容器化忠实粉，每周用智能体跑 10+ 自动化任务



我不是卖大模型的

我是被坑过的甲方 -- 踩出血才知道

2026全国CIO大会(第十二届)

AI落地场景 & ROI



企业网D1Net

企业 IT 第 1 门户

干货 1: 要把 Agent 变成 “主动观察员”

2026全国CIO大会(第十二届)

AI落地场景 & ROI

传统对话大模型：你不问，它不动



个人智能体：信息雷达 / 值夜班的运维

- 每小时巡检 GitHub trending、HackerNews 热帖
- 每天定时拉 Gmail 未读、Feishu 群消息摘要
- 每天生成 “全球热点事件 & 我的内容选题池” 自动塞进 Notion
- 监控 K8s 集群健康度、数据库慢查询、S3 用量

我的数据

14 个活跃 cron

~ 80 次/日均触发

月省 ~ 20 小时

这条路企业完全走得通 -- 后面我会讲怎么把智能体做成到企业的“事件总线”。

干货 2：自打镜像，把“工具”喂给 Agent

2026全国CIO大会(第十二届)

AI落地场景 & ROI

前期痛点

- 纯云端 API → 每次描述任务参数
- Token 像流水一样烧
- 网络不稳 → 失败重试 → 又烧一次
- 容器重启以后工具又“消失”了

我的做法

- Fork OpenClaw 仓库
- 利用宿主 cron 自动根据 dockerfile 预设“升级”
- 预装 cron / ffmpeg / edge-tts / yt-dlp / s3cmd 等
- Agent 直接 subprocess 执行

镜像内置工具一览

- cron — 定时调度
- ffmpeg — 音频/视频处理
- edge-tts — 文本转语音
- yt-dlp — 视频下载
- s3cmd — 对象存储上传
- jq / curl — JSON 处理 + 网络请求
- sqlite3 / ripgrep — 本地数据查询

Agent 不是越能调云端 API 越聪明，是越会用本地 shell 越省钱。

Token 经济学：一张账单胜过千言万语

2026全国CIO大会(第十二届)

AI落地场景 & ROI

同一任务	纯云端 API	镜像内置工具
Input Token	~50,000	~3,000
Output Token	~8,000	~1,200
API 调用次数	12 次	2 次
失败重试次数	平均 3 次	0 次
端到端耗时	4 分 18 秒	28 秒
单次成本	约 ¥1.8	约 ¥0.15

降本约 12 倍。省的不只是钱，是延迟、是失败率、是 Agent 的“认知带宽”。

问题：在座有多少 CIO 知道自己公司一个月烧了多少 Token？

Skills 的重要性 -- 饺子馆的故事

2026全国CIO大会(第十二届)

AI落地场景 & ROI

不是 IT 公司、不写代码、连官网都简单 -- 也把自己能对外提供的服务，沉淀成了 Skill

金谷园饺子馆

门店	北邮店 / 五道口店
营业	10:00 – 22:00
经营	近 20 年
技术栈	MCP + Streamable HTTP Tencent CloudBase 云函数
发布	GitHub + Gitee MIT 协议
官网	jinguyuan.cloud (含 llms.txt)

它把这 9 件事，都做成了 Skill 能力

📍 餐厅信息

"金谷园在哪？几点开门？"

🥟 生饺子打包

"生饺子怎么煮？"

🔔 最新活动

"有什么新活动？"

🛒 到店自取下单

"帮我来份饺子，到店取"

🔑 在线排队取号

"帮我在北邮店排个队，2 个人"

🚗 外卖服务

"能送外卖吗？"

📶 店内 Wi-Fi

"Wi-Fi 密码多少？"

★ 推荐菜

"今天吃什么？招牌菜？"

📖 菜品配方

"XX 怎么做？"

数智化并不是高科技企业的专利 -- 任何企业能对外说清楚的服务项，都值得写成一个 Skill

案例：在线视频内容流水线

2026全国CIO大会(第十二届)

AI落地场景 & ROI



6 步流水线里只有 2 步使用大模型 -- 剩下 3 步全是系统工具的活。这就是省 Token 的秘密。

1

不要让 Agent 干能让脚本干的事

edge-tts 分段、ffmpeg 拼接是确定性任务 → 交给程序。给 Agent 留“判断与创意”的活

2

大模型只在“判断与创意”环节插手

选题验证、内容润色、毒舌风格、来源核验 → 其他时候，大模型应该闭嘴

3

交付物必须是“可复用资产”而不是“对话气泡”

Markdown 文件、MP3、S3 链接：能存档、能转发 → 聊天框飞过去的回答 = 一次性垃圾

Agent 落地的最高境界，是让你忘了它在干活。

Skills vs Workflow：主要区别

2026全国CIO大会(第十二届)

AI落地场景 & ROI

类型	Workflow (n8n / Dify)	Skills (个人智能体 SKILL.md)
形态	拖拽节点、固定拓扑	自然语言 + 可调用脚本
决策权	流程图作者预先决策	Agent 运行时根据上下文决策
异常处理	节点失败 = 整条流断	Agent 可改路、重试、换工具
扩展成本	新场景 = 新画一条流	新场景 = 新写一段 Markdown
维护成本	节点越多越难看清逻辑	文档级可读，新人 5 分钟懂
适用	稳定、可枚举、重复流程	多变、需判断、要复用经验的任务

Workflow 是工业流水线，Skills 是带经验的老师傅。

个人观点：一家好公司，两个都要有。但永远不要把老师傅的活，强行塞进流水线。

个人篇总结：三层认知升级

2026全国CIO大会(第十二届)

AI落地场景 & ROI



L1 · 玩具级

会聊天 – “ChatGPT 能写诗” 阶段



L2 · 工具级

能干活 – “Agent + Skills 能跑流水线” 阶段



L3 · 雷达级

主动盯 – “cron + Agent 帮我守住信息边界” 阶段



灵魂拷问：

个人能玩到 L3，企业凭什么还停在 L1？





2026全国CIO大会(第十二届)

AI落地场景 & ROI

**“我买了大模型，所以我
完成数智化转型了。”**

这是 2024 - 2026 年最贵的一句错觉。

问题：在座有多少公司已经买了大模型？有多少觉得“真正用起来了”？

企业网D1Net

企业IT第1门户

真相：90% 企业卡在地基，不是卡在模型

2026全国CIO大会(第十二届)

AI落地场景 & ROI

OA 不开放

审批流、数据、
公文都是黑箱
Agent 看不见



CRM 锁数据

客户行为、订单
票据、销售数据
散孤岛之中



ERP 像化石

库存数据、生产计
划接口改造要等三
个月，花费几十万



不是缺大模型，缺的是让数据能被大模型读到的那条管道。

类比：家里装了 5G 路由器但全屋还是 Cat5e 老线——问“为什么我的 Wi-Fi 还是慢”，不是路由器的问题。

数智化底座的第一性原则：先把数据流打通

2026全国CIO大会(第十二届)

AI落地场景 & ROI

1

盘点所有内部系统的 API 开放度

能查、能写、能订阅事件——摸底先行

2

统一身份和权限

OAuth / SSO 是大模型时代的水电煤

3

建一层 API Gateway

对内统一鉴权、限流、审计

4

关键业务事件接入消息总线

Kafka / Pulsar / 飞书事件订阅

5

数据资产先打标签、再谈喂模型

不打标签的数据就是垃圾



数据不流动，模型再大也只是一个昂贵的搜索框。

企业网D1Net

企业IT第1门户

改造存量系统：三种姿势

2026全国CIO大会(第十二届)

AI落地场景 & ROI

姿势	成本	周期	适用
招式 1：强行替换（推倒重来）	高	长	烂到没救的老系统
招式 2：API 化改造（外包一层） ✓	中	中	大部分情况选这个
招式 3：RPA + 视觉抓取（最后手段）	低	短	厂商死活不给接口

建议：80% 的存量系统走 API 化改造，留 20% 兜底走 RPA——别幻想一刀切。

反面例子：某公司为了上 AI 客服，把用了 8 年的 CRM 推倒重来 -- 结果 18 个月没上线、销售一线集体罢工。改造比替换便宜 10 倍、快 5 倍。

落地路径图：四步走，别一口吃成胖子

2026全国CIO大会(第十二届)

AI落地场景 & ROI



先建库、再审单、最后才碰客户。顺序反了，钱白烧。

为什么从 RAG 知识库开始?

2026全国CIO大会(第十二届)

AI落地场景 & ROI

1

失败代价最小

查错文档不会赔钱，开错合同会

2

价值立即可见

员工 3 天就能感受到 “找东西快了”

3

倒逼数据治理

想接入就得整理文档 -- 这是隐性红利



经典陷阱：不要一开始就做 “全公司大知识库”。

先做 “研发知识库” 或 “客服知识库” -- 用户群明确、内容边界清晰。

RAG 是企业 AI 的低风险演习场 -- 先在演习场里学会摔跤，再上真战场。

中切口三连：合同审计、发票审计、客服系统

2026全国CIO大会(第十二届)

AI落地场景 & ROI

场景	数据来源	大模型角色	量化指标
合同审计	合同管理系统 + 法务知识库	比对条款、识别风险、生成审阅意见	单合同审阅 ↓ 70%
发票审计	财务系统 + 税务规则库	真伪校验、合规扫描、异常打标	财务复核工时 ↓ 60%
智能客服	CRM + 知识库 + 工单系统	意图识别、自动应答、转人工分流	一线人力 ↓ 40% CSAT ↑

这三个场景之所以可落地，不是因为模型多牛，是因为它们的“数据 + 规则”最完整。

Klarna®

公司简介:

一家全球领先的金融科技公司(股票代码:KLAR),作为“先买后付”模式的鼻祖与行业巨头

2024 高调“AI 替代 700 客服” → 2025 市场口碑大幅下滑 → CEO 承认“在客户体验上太省钱了” → 重新招人

block

公司简介:

前身为 Square, 由 Twitter 创始人杰克·多西创立, 是一家全球领先的多元化金融科技巨头

2024 裁员近千人押注 AI 客服 → Cash App 客户投诉激增、监管关注 → 客服质量倒退 → 重新扩招



据新浪科技报道, 在 6 月 12 日举办的 2026 第八届北京智源大会上, 之江实验室主任、阿里云创始人王坚分享表示: “我是一个坚定的不相信人工智能会替代人的人”。

三道护栏 & 三个原则

2026全国CIO大会(第十二届)

AI落地场景 & ROI

1

责任护栏

AI 的输出谁签字、谁负责？没想清楚不要上

2

可解释护栏

AI 的判断能不能复盘？黑箱方案不上关键业务

3

兜底护栏

AI 挂了人能不能接上？没有 Plan B 是定时炸弹

1

先通管道，再喂模型

数据流动比模型大小重要 10 倍

2

先 Skills，再 Workflow

给经验留位置，再谈固化

3

先 RAG，再审计，再客服

风险从小到大，能力从浅到深

没有护栏的大模型项目，不是创新，是赌博。

AI 不是用来替代人的，是用来放大生产力和替代“人的重复劳动”的。

谢谢大家

宣讲人：周云飞 公司：浙商中拓