

什么是 AI 智能体 (AI Agent)?

1. 核心定义：从“自动化”到“智能化”

在创客的世界里，我们习惯了编写**自动化脚本**。比如：“如果温度超过 30 度，就打开风扇”。这是一个固定的逻辑。

AI 智能体 (AI Agent) 则完全不同。它不仅能执行指令，还能**通过推理来决定做什么**。你可以给它一个模糊的目标，比如“让房间里的植物保持舒适”，它会自己去检查温度、查询该植物的习性、决定是否开窗或浇水，并最终去执行。

- **传统自动化**：“如果 A，就做 B”（死板的代码）。
- **AI Agent**：“这是目标，你自己看着办”（具备推理能力的逻辑）。

2. Agent 四大支柱

为了让 Agent 能够像人一样工作，它通常由四个部分组成：

- **感知 (Perception)**：获取信息的能力（传感器数据、用户输入的语音、摄像头画面等）。
- **大脑/推理 (Brain/Reasoning)**：核心的大语言模型（LLM）。
- **记忆 (Memory)**：记住之前的操作和偏好（历史对话记录和总结）。
- **行动 (Action)**：改变物理世界或数字世界（驱动电机、点亮 LED、发送邮件、调用 MCP 工具）。

AI Agent 的“朋友圈”：周边关键知识

在学习 AI 智能体硬件时，你会经常听到以下几个“黑话”概念：

1. 工具调用 (Tool Calling / Function Calling)

AI 本身只会“说”，不会“做”。工具调用是 Agent 能够操控硬件的关键。AI 会输出一段特定的代码格式，告诉硬件：“我现在需要调用 `turn_on_light()` 这个函数”。

2. 提示词工程 (Prompt Engineering)

这是你给 Agent 下达的“出厂设置”。对于创客硬件，提示词通常包含：

- **角色定义**：“你是一个智能管家”。
- **技能描述**：“你可以通过串口控制舵机”。
- **约束条件**：“每次动作前必须先检测传感器数据”。

3. 多模态 (Multimodal)

这意味着 Agent 不再只看文字。它能够通过摄像头“看”到你的手势，或者通过麦克风“听”到你的情绪。这对于开发交互式硬件项目（如桌面机器人）至关重要。

4. 边缘计算 vs 云端 AI

- **云端 AI**：大脑在远程服务器（如 OpenAI、豆包、通义千问等）运行，智商极高但需要联网，且有延迟。
- **边缘 AI (Edge AI)**：大脑就在你的硬件芯片上。虽然智商稍逊，但响应极快且不需要联网，非常适合 DIY 隐私项目。