

# AI Agent

## Agent 结构:

1. 终端: Web、App、Terminal、Plugin
2. Server: 存储、检索、token 扫描、调优
3. 底层: 通用模型

从 Claude Skill、Claude Actionbook、google Hybrid 模式调用  
接口 ,Agent 目前主要竞争方向:

- 1、更少的 token 消耗，更快速的内容生成成为主要目标。
- 2、通过工具包开源调用接口以开发社区建立产品生态，进而占领市场。

## 目前针对第二层 Server 层有两种实现思路:

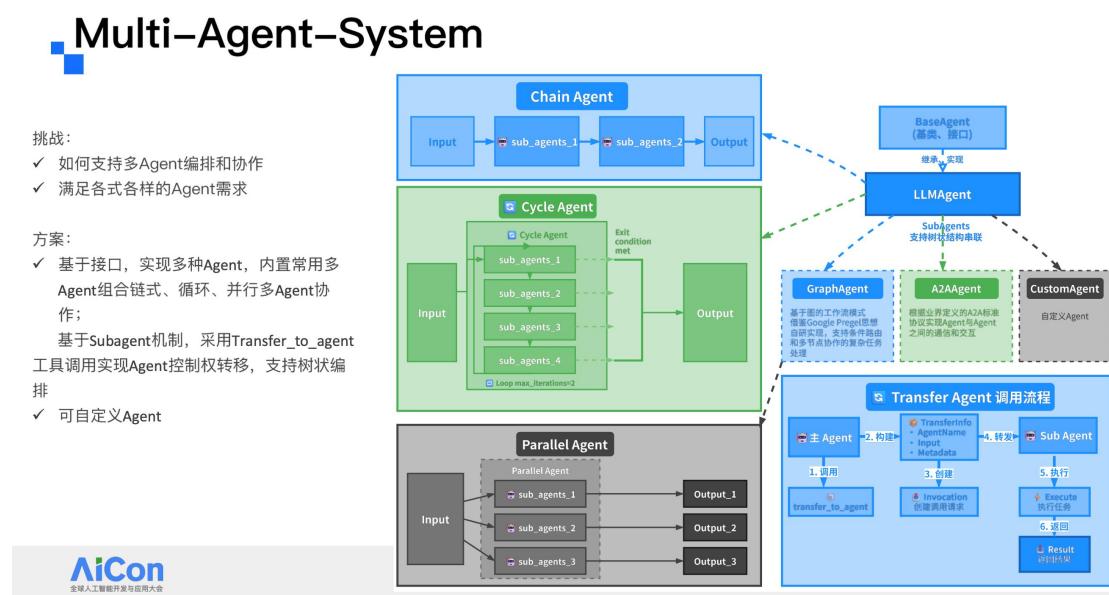
1. 做成通用的 AI Coding 平台: 阿里 RDS Supabase、Google Chrome Gemini nano
2. 定制化开发: 小红书营销 AI 平台、基于 Claude SDK 开发应用  
第一种方案更适用 C 端产品，对于细分领域的精读不高，对广度要求较高

第二种方案更适用 B 端产品，对于私有数据、定制物料的要求高，

对细分领域要求高，对业务流程结合的定制化要求高。

企业场景下的 AI 应用面临很多不确定性	
C 端	B 端
常规知识	训练大模型阶段直接被训练进入了模型内
新知识	互联网几十年的发展，有无数专业媒体、自媒体网站
搜索工具	Google、百度等搜索引擎经过了几十年发展，非常强大易用
权限	互联网基本上是公开的
	大量权限制约了 AI 对企业内知识的访问

Agent 架构设计：



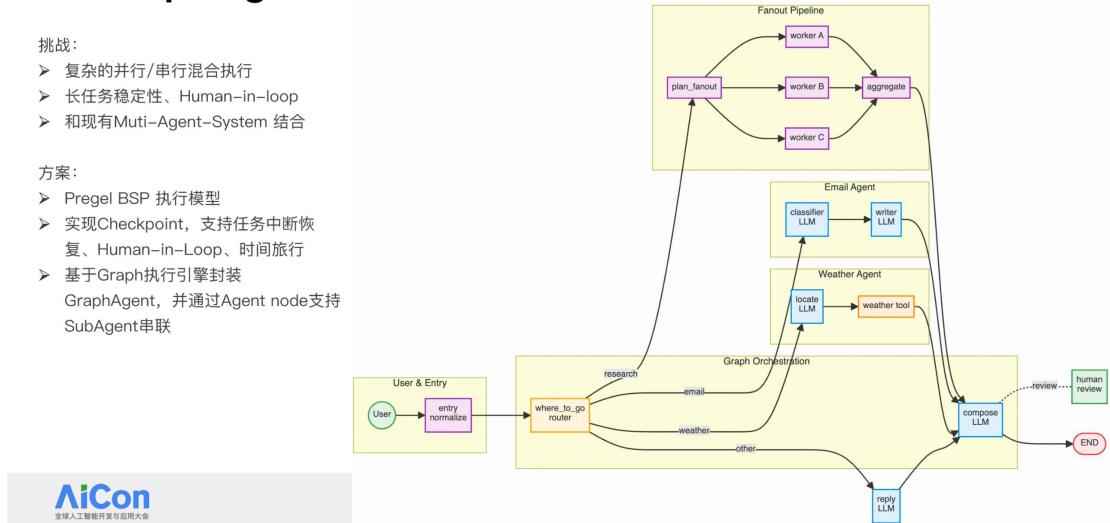
## GraphAgent

挑战:

- 复杂的并行/串行混合执行
- 长任务稳定性、Human-in-loop
- 和现有Multi-Agent-System 结合

方案:

- Pregel BSP 执行模型
- 实现Checkpoint, 支持任务中断恢复、Human-in-Loop、时间旅行
- 基于Graph执行引擎封装  
GraphAgent, 并通过Agent node支持  
SubAgent串联



AiCon  
全球人工智能开发与应用大会

## Planner

挑战:

- 如何让不同模型的Agent都具备自主性?

方案:

- 内置思考的built-in planner，  
开启模型思考
- 支持React planner，React 通过  
显式指令引导模型按标签格式输出，  
流式累积完整消息后再正则  
匹配标签

支持自定义Planner



AiCon  
全球人工智能开发与应用大会

InfoQ 极客传媒

用 AI 的思维, 以 AI 为中心构建技术和产品:

范式转变:

我想怎么做: 以我为中心, 建立我的 design system。

我想 AI 怎么做: 以 AI 为中心, 建立 AI 的 design system。

【关联思考-知行合一--20260102】

前端的需求:

1. 物料整合

## 2. d2c

以 AI 为中心的思维模式的实现路径:

1. prompt

2. fine-tune

3. skill

prompt 更加考验使用者的逻辑能力、表达能力。

逻辑能力: 清晰化目标, 只能自己想要什么。

表达能力: 把目标讲清楚, 让 AI 完全能理解。

例如:

case1: 生图应用:

### 生图应用：图生特效

基于文生特效模型的衍生能力

- 1. 黑白涂鸦作为输入
- 2. 彩色涂鸦作为输入
- 3. 黑白稿作为输入
- 4. prompt+材质参考图作为输入

文生图和图生图的区别

- 文生图的门槛更低, 更适合普通用户体验。
- 图生图对于颜色、形状、结构可以更精准的控制。
- 专业画师更擅长用绘图表达创意和概念。
- 对于专业画师, 图生图比文生图更受欢迎。

线稿生特效

淡金色预期的拖尾特效。金线流光, 粒子特效, 条形, 蔚蓝效果, 梦幻, 散发出神秘的光芒, 风格是写实。方向是左向右。

黑白稿生特效

一个青金色的特效, 风格是中国风, 顶部是一个铃铛, 云雾缭绕, 底部水花四溅。

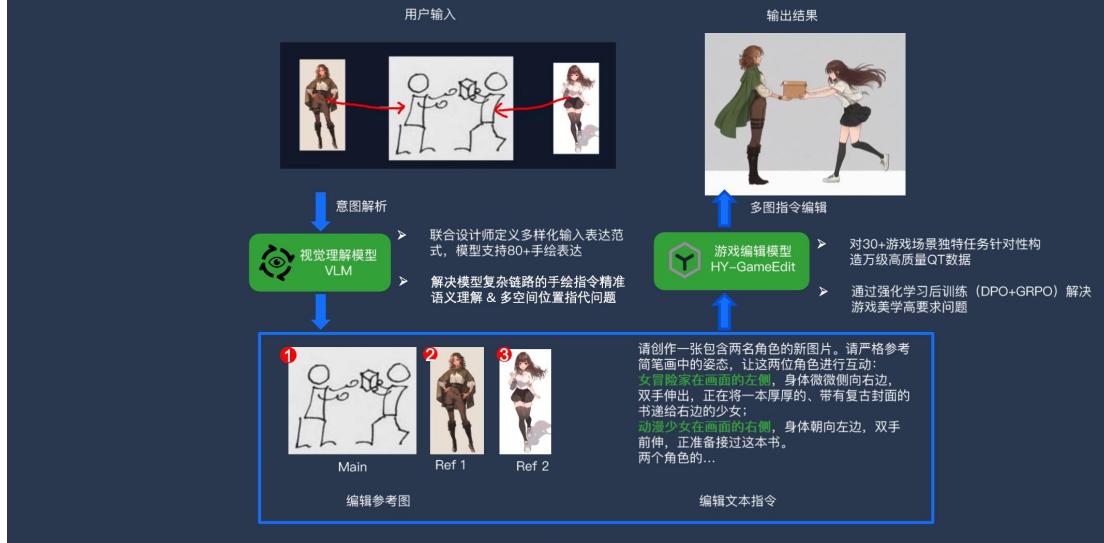
彩色线稿生特效

一个冲刺特效, 头部是金色的箭头带着浅蓝色的烟尾, 特效周身带着星星点点。

参考材质生特效

绿色的拖尾特效, 形状锥形, 写实风格, 灰色背景。

# 新一代游戏图像编辑功能——绘意画板



case2: prompt 模板生成后台页面

skill:

prompt 工程，将物料拆解，告诉 AI 怎么使用物料，生成什么内容。

前提：使用者自己清晰目标；使用者自己清晰如何使用物料。

类似：workflow

fine-tune