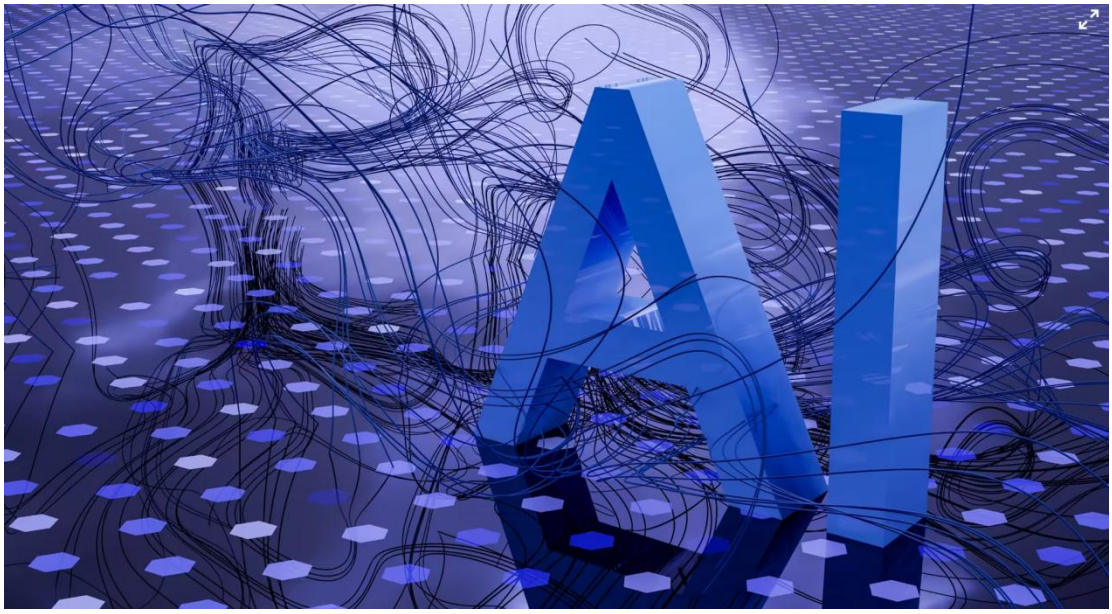
AI 领域的 Agent 是什么?

Agent（代理）像一个具有“手、脚”的智能体，它能够进行思考、决策，并且能执行具体的任务。是一种能够感知环境、进行决策和执行动作的智能实体。不同于传统的人工智能，Agents 具备通过独立思考、调用工具去逐步完成给定目标的能力。



一.Chat 和 Agent 的区别

Chat（聊天）：纯粹的 Chat，像是一个主要由“大脑和嘴”构成的智能体，专注于信息处理和语言交流。比如 ChatGPT 这样的系统，它能够理解用户的查询，给出有用和连贯的回答，但它本身不直接执行任务。

Agent（代理）：像一个具有“手、脚”的智能体，它能够进行思考、决策，并且能执行具体的任务。是一种能够感知环境、进行决策和执行动作的智能实体。不同于传统的人工智能，Agents 具备通

过独立思考、调用工具去逐步完成给定目标的能力。在有 LLM 作为其大脑之后，Agents 更是具备了对通用问题的自动化处理能力。

Agents 与大模型的区别在于，大模型与人类之间的交互是基于 Prompt 实现的，需要有输入才会产生输出。当输入的 Prompt 不清晰时，会明显影响大模型回答的效果，大多数需要多轮输入才能得出一个效果较好答案，甚至对于部分问题，大模型甚至无法处理，如问大模型今天天气怎么样。

而对于 Agents，仅需要给出一个目标，它就能根据目标进行独立、自主的思考，它会根据给定任务详细拆解出每一步的计划步骤，依靠来自外界的反馈和自主思考，自主创建 Prompt，来实现目标，比如问今天天气怎么样，它分解为多个步骤，通过确定你所在地点，然后调用天气查询 API 等步骤，为你获得你所需要的信息。

Agents 目前可分为自主智能体（Autonomous Agent）和生成智能体（Generative Agent）。

自主智能体：如 Auto-GPT，主要是为人类服务，自动执行任务并实现预期结果。

生成智能体：如斯坦福和谷歌的西部世界小镇，它们在同一环境中生活，拥有自己的记忆和目标，不仅与人类交往，还会与其他机器人互动。

我们可以简单粗暴的理解为，Chat 强调的是“说”，Agent 强调的是“做”。

自 ChatGPT 发布后，从 plugin 的推出，到 Function Calling 再到 Assistant API 的面世，OpenAI 这一系列动作就充分表明，有这么强大的 LLM 作为基本盘的情况下，人们就不可能仅仅满足于让它“嘚啵嘚”。

历史总是惊人的相似。从 2014 年亚马逊开创性推出 Amazon Echo 开始，智能音箱横空出世。一开始的智能音箱，也只是有个“嘴”，只能实现播放音乐、查询信息、设置提醒等功能。而且“脑子”还不太灵光。

但是随着阿里、百度、小米等科技巨头的纷纷加入，智能音箱在竞争中卷出了新高度。打通支付、和智能汽车、智能家居互通，智能音箱不断地突破和扩展功能边界，逐步坐到了智能家居生态的“大总管”位置上。

随着应用场景的持续拓展，智能音箱又延展到儿童教育、养老关怀等领域，深刻影响了人们的日常生活。

相信有一天，智能音箱会强大和多样化到一个程度，以至于“智能音箱”这个名字不再适合这个品类，那将是新一轮故事的开始。

同样惊人相似的，还有从单纯的 AI 智能语音助手、智能客服（只会说）到以 AI+RPA 为核心技术的的 AI 数字员工（会说又会做）的发展史。

这些，都是人工智能走向多元化和融合化的一个个缩影。

因此，随着技术水平的不断进步和场景化落地的不断挖掘，Chat 和 Agent 的界限必定会越来越模糊，生成式 AI 会融合 Chat 和 Agent

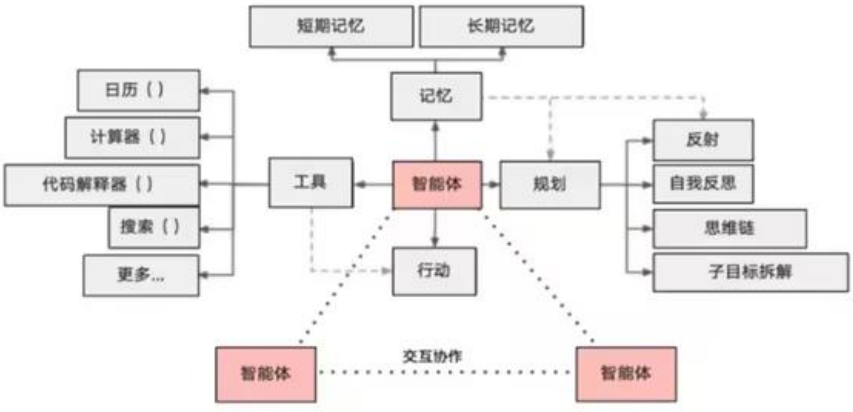
的特点，形成既能进行高质量、高人格化对话，又能高效执行复杂任务的 AI 自动化系统，为人们提供融合、互补、多样化的解决方案。

二.AI Agent 的崛起

AI Agent 的崛起不仅仅是技术上的突破，更是对软件开发理念的一次深刻变革。

在传统的软件开发中，程序员需要预先定义所有的逻辑和规则，然后进行代码实现。而 AI Agent 的出现，要求我们对软件进行充分地“放权”：它由一颗大脑（LLM）来进行自主支配运行，并在运行时自动学习、适应和调优。这种前所未有的开发范式的转变，让程序员不得不重新思考软件开发的本质，也重新思考软件开发的未来。

典型的 AI agent 分为 Memory（记忆）、Tools（外部工具） 、Planning（计划） 和 Action（行动）四个模块。



当前学习 AI Agent 基本上分作两条路径：

基于 OPenAI 技术路线，以及基于开源技术路线。建议每个技术人员，都选择一条路，亲自趟一趟。大模型爆发之后，AI Agent 的发展也可谓是一日千里，各种项目层出不穷。

三. 部分 Agents 盘点

AutoGPT

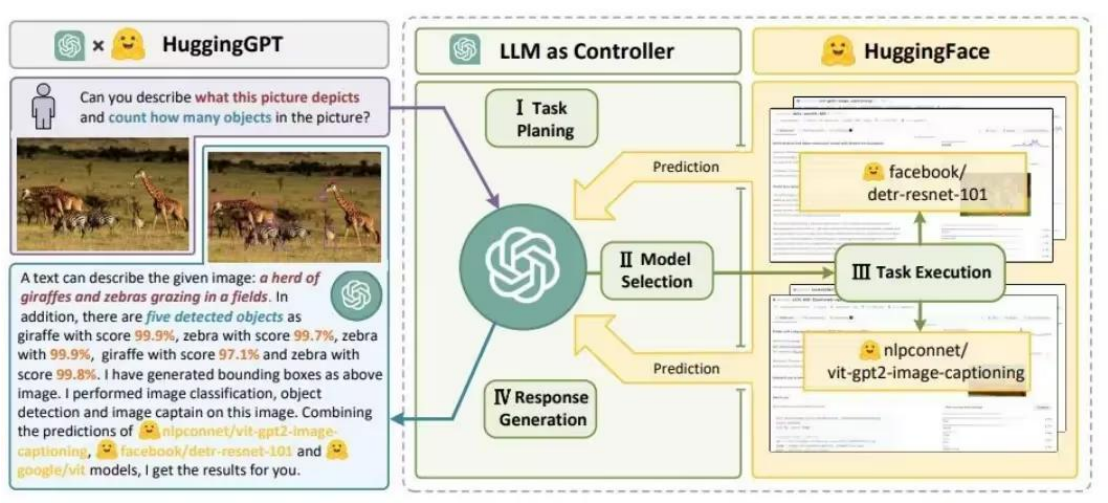
项目地址：https://github.com/Significant-Gravitas/AutoGPT

可以根据你设置的目标，将实现这个目标的任务进行拆解，再采用搜索、浏览网站、执行脚本等方式一条条去执行任务，帮你完成目标。

JARVIS

项目网址：https://github.com/microsoft/JARVIS

一个非常有意思的“模型选择”Agent。它将用户要求拆解成子任务，再到 Huggingface 上选择合适的专家小模型执行任务，最后对结果进行处理和返回给用户。



由于 JARVIS 可以调用其它模型工具，因此它可以执行多模态任务。

MetaGPT

项目网址：https://github.com/geekan/MetaGPT

MetaGPT 是另一个开源人工智能体框架，试图模仿传统软件公司的结构。与 ChatDev 类似，Agent 被分配产品经理、项目经理和工程师的角色，并且他们在用户定义的编码任务上进行协作。

工具、平台、社区的不断成熟，为个体开发者提供了一个全新的舞台。程序员与人工智能之间的距离从未如此之近。AI Agent 的崛起，让有想法、有技术的人能够以前所未有的方式释放自己的创造力，打造出各种有趣、实用的 AI 原生应用。