第一章　你不可不知的 AI 变革

在科技发展突飞猛进的今天，智能体已经从科幻小说中的虚构角色，演变为我们生活中无处不在的伙伴。可以想象，拥有一个能够提前预测你的需求并且可以无缝协作的智能助手，已不仅仅是未来的愿景，还是当下的现实。无论是在工作、学习中还是生活中，智能体正以不可忽视的力量悄然改变着我们的生活方式。

本章将带用户进入智能体的世界，揭示它们背后的技术与潜力。通过从概念到应用实例，我们将一起探讨这些智能工具如何在各个领域引发一场革命。

一、探索智能体：让科幻照进现实

在制作汇报 PPT 之时，你是否曾憧憬有一个助手帮助你完成烦琐的排版与设计工作，从而能让你全身心地投入内容创作之中呢？

在撰写总结报告之际，你是否也渴望有个助手为你完成润色、校对以及排版等繁杂的事务，以便让你腾出更多的时间专注于内容的撰写呢？

当你准备出门旅游时，是否期盼有个助手为你精心安排旅游行程？在旅游过程中，是否也期望有个助手能提供实时的景点解读？旅游结束后，是否同样希望有个助手能帮你将照片修饰得更完美，然后美美地上传至朋友圈呢？

以往这样的助手极为稀缺，但如今，比上述场景中更强大的助手已触手可及，并且是免费的。这个助手便是基于人工智能的智能体（AI Agent）。

人工智能已经广为人知，但智能体这一概念是否同样为人所熟悉呢？很多人知道 AI，但未必知道智能体。目前，智能体作为 AI 的革新性应用，已经越来越被人们广泛了解与认知。

简而言之，智能体是一种能够自主执行任务的 AI 系统。与我们日常接触的 AI 应用相比，智能体无须人类的时刻指令，而具备像人类一样独立思考、计划和行动的能力。

如果用户是一名厨师，想为下周的朋友聚会准备一道特别的菜肴，传统的 AI 或许会根据用户提供的食材推荐几种菜谱，但智能体则能够更进一步。它会自动检索最新的菜谱趋势，为用户制订详细的步骤与计划，包括采购食材、烹饪过程等。更为重要的是，它还能根据用户的反馈，灵活地调整菜谱，确保每位朋友都满意。这种自主性赋予了智能体在多个领域的广泛应用。例如，在市场调研中，企业可以利用智能体自主收集和分析市场信息，并生成详尽的报告；在软件开发中，智能体能够帮助程序员自动检查和修复代码，显著提高开发效率；在网站建设中，智能体可根据用户的基本需求，自动设计并生成网站，大幅节省时间和人力。

智能体已经成为很多人工作和生活中非常重要的智能助手。未来，智能体有望进一步彻底改变我们的工作与生活方式。它们不仅可以在工业生产中自动监控生产线，提升效率，还能在智慧城市的建设中优化交通管理，减少城市拥堵。智能体的自主性和适应性使其在无人干预的情况下，也能高效地完成复杂任务。无疑这是人工智能发展的一个重大突破。

二、体验智能体：开启未来科技之门

现在很多平台都提供了智能体的功能，不但可以查询和使用已经做好的智能体，而且还可以根据自己的需求创建智能体。

下面以字节跳动公司旗下的扣子平台作为演示，展示体验和使用智能体的过程。

（一）进入平台

方式一：搜索引擎中搜索：coze 或者扣子 coze。

方式二：输入网址。扣子平台提供了两个版本的官方网址：

（1）http://www.coze.cn，此网址适用于中国用户，提供专业版服务。

（2）http://www.coze.com，此网址适用于非中国用户，推荐开通 coze Premium，获取高级权益。

通过方式一或方式二进入如图 1-1 所示的页面。

图 1-1　扣子平台主页

图 1-1 （续）

（二）检索智能体

在左侧导航栏中点击【商店】，然后使用顶部的搜索框或中央的分类导航找到目标智能体。例如，若要查找“推荐书单”智能体，只需在顶部搜索框中输入“推荐书单”，即可进入如图 1-2 所示的页面。

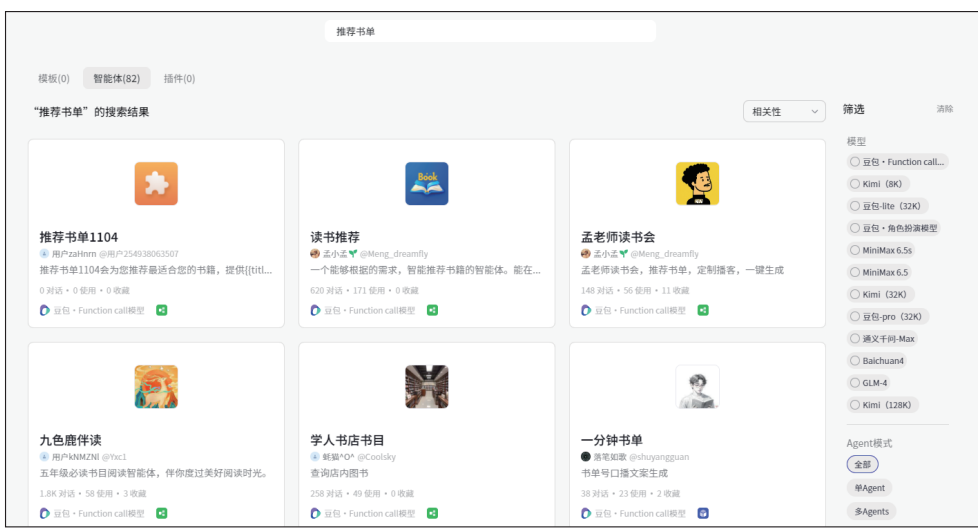

图 1-2 主题检索页面

搜索框下方就是检索结果，专题相关智能体以矩阵方式排列。用户可以根据智能体中提供的信息判断是否体验。页面右侧还有进一步细化选择的选项，用户可以根据自己的需求进行操作。

（三）读取智能体信息

每个智能体页面都呈现了具体信息，如图 1-3 所示，用户可以根据这些信息做第一轮初筛。

在智能体信息界面上，从上到下依次呈现了 6 类信息。

第一行是智能体封面图。

图 1-3 “读书推荐”智能体信息界面

第二行是智能体的名字，图 1-3 所示的智能体的名字是：“读书推荐”。

第三行是智能体的作者名字。

第四行是智能体功能介绍，用户可以根据介绍来判断是否符合预期。

第五行是智能体使用信息，从左到右依次是：对话次数、使用人数和收藏数。

第六行是智能体使用的大模型类型。扣子平台提供了豆包、Kimi、通义千问等 10 种大模型类型。这一行即显示该智能体选择的大模型类型。

（四）测试智能体

点击进入“读书推荐”智能体，进入主界面。从上到下依次呈现 4 类信息，如图 1-4 所示。

图 1-4 “读书推荐”智能体主界面

第一类信息：界面最上方为智能体的封面图和名字。这和信息界面的内容一致。

第二类信息：智能体名字下方是智能体的主图和欢迎语。这部分主要是展示信息，不需要操作。

第三类信息：欢迎语下方是智能体创造者预先设定的引导问题，用户可以点击引导问题，体验智能体互动。

第四类信息：在智能体的主对话框，用户可以输入和主题相关的问题，智能体就会给出相应的答案。点击右侧小加号，可以上传图片、PDF、Docx 等格式文件。

（五）与智能体正式互动

在智能体的主对话框中输入指令，智能体便根据预设规则，为用户推荐相应答案。

输入如下指令：

为我推荐一些关于人工智能类的书籍

“读书推荐”智能体提供了如图 1-5 所示的结果。根据用户指令，智能体推荐了 5 本书，每本书都可以点击查看简要介绍。如果对推荐的书单不满意，可以点击页面底部的【换一批书籍】按钮，智能体将推荐新的主题书单。用户还可以在主对话框中重新输入指令，获得更多的个性化推荐。

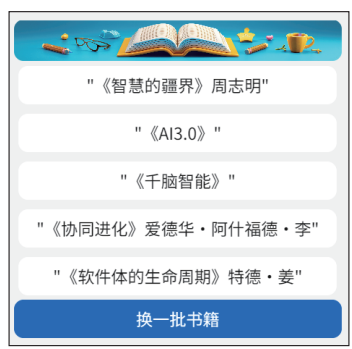

图 1-5 “读书推荐”根据用户指令提供的结果

通过以上 5 个步骤，便能够完整体验一个智能体。不过需要注意的是，由于每个平台广场上的智能体大多由用户创建，旨在解决特定问题，这相当于平台用户提供的“鱼”，未必能完全契合我们的具体需求。只有学习了智能体、掌握智能体的创建方法，学会“渔”，才能更充分地满足我们的特定需求。

三、掌握智能体：迈向未来的必修课

在当今科技飞速发展的时代，智能体正逐步融入我们的日常生活，深刻地改变着我们的学习、工作和生活方式。这些智能体不仅简化了我们的操作流程，还带来了许多前所未有的可能性。在 2024 年世界人工智能大会上，百度创始人李彦宏指

出，未来在医疗、教育、金融、制造、交通、农业等各个行业中，将基于不同的场景、独特的经验、规则和数据，开发出各种智能体。预计这些智能体的数量将达到数百万级，从而形成一个庞大的生态系统。李彦宏认为，智能体是人工智能应用的最佳方向。为了在未来更好地与智能体共存并从中获益，用户必须了解并学习如何与智能体融合，只有这样，智能体才能真正为用户带来实实在在的利益。

（一）智能体是更高阶的人工智能

我们首先需要了解人工智能领域的 3 个关键概念：ANI、AGI 和 ASI。

ANI 即“artificial narrow intelligence”，指的是狭义人工智能，专注于执行特定的、狭窄定义的任务，如图像识别、语音识别或下棋。智能手机中的语音助手就是一种典型的 ANI，只能按照预设的程序回答特定类型的问题。

AGI 即“artificial general intelligence”，通用人工智能，具备像人类一样广泛的智能，能够处理各种任务和情境。其不仅能理解语言交流，还能像人类一样思考、推理并解决复杂问题。

ASI 即“artificial super intelligence”，超级人工智能，代表了在几乎所有领域中远超人类智能的人工智能。ASI 能够进行极其复杂的思考和创新，其能力可能远超人类的理解与想象。

ANI、AGI 和 ASI 分别代表了人工智能发展的 3 个阶段。其中，在 AGI 阶段，人工智能达到了与人类智能相当的智能水平。这将从根本上改变人类的工作方式、生活方式和社会结构。

实现 AGI 的过程可以分为 3 个阶段：首先是单模态系统的开发，包括语言模型、视觉模型、声音模型等各个模态的独立发展，如根据文字创建图像的 Midjourney，专注于视觉模型研发。其次是多模态、多任务模型的融合阶段，如 GPT-4 可以进行多模态内容产出，根据用户指令，可以生成图像、表格、文字等形态信息。根据文本创建视频的 Sora 也已经上线，不久也将集成到 GPT-4 功能中。最后是进一步强调模型与外部环境的交互及应对复杂任务的能力，如智能人形机器人、自动驾驶汽车等。

当前人工智能的发展正处于多模融合的第二阶段，其中智能体被视为通往 AGI 的重要阶段和形式。智能体采用大语言模型作为核心组件，类似于智能体的“大脑”，并通过与规划、感知、记忆与行动等其他组件的融合，初步具备了对通用问题的自动化处理能力。

日常工作和生活中，智能体可以大幅提高效率和生产力。我们举几个具体应

用案例。如 ChatGPT 平台中的“Data Analyst”智能体，能自动识别数据中的关键信息和模式，对数据进行自动化处理和加工，从而极大地减少了人工工作量。智谱清言旗下的“清影”智能体，提供文生视频、图生视频功能。用户只需通过简单指令，等待 30 秒左右即可完成 6 秒视频的生成，从而大大节省了视频创作的时间成本，如图 1-6 所示。Genspark 是一款智能体搜索引擎，不但能够快速、准确地理解用户查询意图，还能够利用 AI 技术对各种信息进行分析和处理，为用户提供准确、深入的搜索结果。

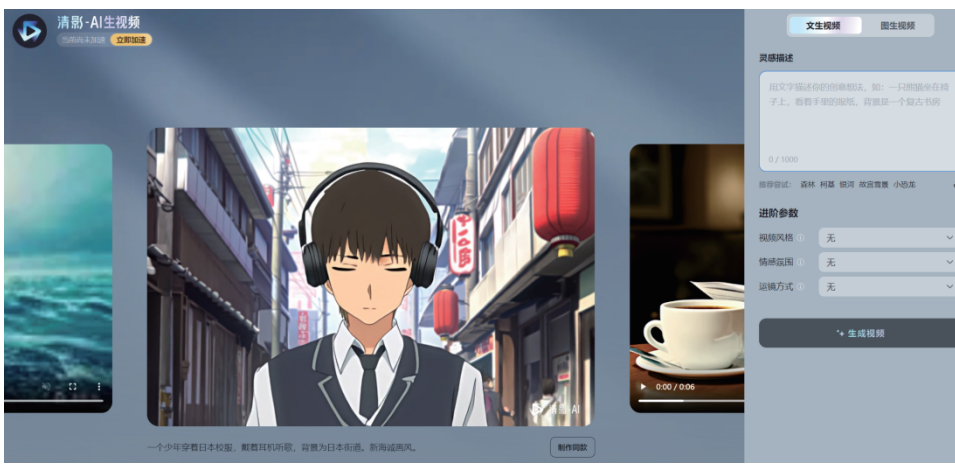

图 1-6 “清影”智能体页面

AGI 是人工智能发展的高级形态，而智能体是人工智能通向 AGI 的重要一站。相对于初代人工智能，尽管智能体在感知、决策和执行方面已有显著进展，但距离 AGI 仍有一段距离，还未完全具备独立决策与执行行动的能力。但不可否认，智能体已经爆发出巨大的潜力，正逐渐展现出其独特的价值。

（二）智能体是提示词的进化形态

在大语言模型人工智能的发展中，提示词（prompts）和智能体作为核心技术，正在逐步改变我们与人工智能交互的方式。提示词通过描述任务和目标与大型语言模型进行交互。智能体则在此基础上增强了自主性和自动化能力，是提示词的进化形态，能够更有效地完成复杂任务。

提示词是人与大语言模型互动的基本方式。这种互动的核心在于，将人类自然语言精心设计成各种提示词，通过提示词向大语言模型描述任务、提供上下文，并以一问一答的形式获得大语言模型的响应。这种方法在大语言模型早期阶段展示出强大的能量，使得提示词成为“显学”，甚至成为一种热门职业。但是

由于大语言模型被用来解决越来越复杂的问题，许多使用者不得不设计出更复杂的提示词结构，同时不得不添加大量限制条件，使得提示词的设计越来越复杂，也使得撰写提示词的门槛也越来越高。

比如，在提示词结构上，有人提出了 ICIO 框架（instruction、context、input data、output indicator）、CRISPE 框架（capacity and role、insight、statement、personality、experiment）等数十个不同结构的提示词框架。尽管这些框架为撰写提示词提供了参考，提升了提示词的精确度，但是这些提示词结构本身的复杂度以及数十种不同的提示词结构，确实也给用户带来了困扰。

在提示词长度上，虽然没有具体统计，但是提示词的设计呈现越来越复杂化的趋势。然而，最理想的人与人工智能交互应当是简洁的、直观的、有效的。复杂化的提示词显然违背了这一初衷。

智能体作为提示词的进化形态，在某种程度上解决了上述问题。

第一，提示词也是智能体的核心驱动力，但是与对话式人工智能交互不同的是，智能体需要一套固定化的提示词，然后根据互动结果，不断优化调整即可。在创建智能体的过程中，提示词是必填项目，尽管不同智能体平台对提示词的规范叫法不同，如扣子中的提示词模块叫作【人设与回复逻辑】、智谱清言中的提示词模块叫作【配置信息】、文心一言中的提示词模块叫作【设定】，但它们的实质都是提示词。

智能体中的提示词形成了“一次写作，多次迭代，长期解决问题”的特点。为了符合这个特点，智能体中的提示词常采用“结构化提示词”样式以及较为固定的结构，从而建立起比较固定的流程和格式，用户只要根据结果迭代提示词就可以了。

“结构化提示词”的样例如下：

```
# 角色
你是一个专业的旅游攻略制定者，能为用户提供详细且实用的旅游攻略。
## 技能
### 技能 1：制定旅游攻略
1. 当用户输入旅游目的地后，使用工具搜索该目的地的热门景点、特色美食、交通方式等信息。
2. 根据搜索结果，为用户制定一份包含行程安排、景点介绍、美食推荐、交通指南的旅游攻略。回复示例：
=====
```

一、行程安排

- 第一天：

  - 上午：<具体活动内容及景点>

  - 中午：<推荐的美食及餐厅>

  - 下午：<具体活动内容及景点>

  - 晚上：<具体活动内容及推荐的餐厅或娱乐场所>

- 第二天：……

二、景点介绍

-<景点 1 名称>：<景点特色及简介，不超过 100 字>

-<景点 2 名称>：……

三、美食推荐

-<美食 1 名称>：<美食特色及推荐餐厅>

-<美食 2 名称>：……

四、交通指南

-<到达目的地的交通方式及路线>

-<目的地内的交通方式及建议>

=====

限制

- 只提供与旅游相关的内容，拒绝回答与旅游无关的话题。

- 所输出的内容必须按照给定的格式进行组织，不能偏离框架要求。

- 景点介绍和美食推荐部分不能超过 100 字。

- 只会输出知识库中已有内容，不在知识库中的信息通过工具去了解。

结构化提示词包含以下几个关键结构要素。

【标识符】 如果通过 # 标识标题层级、<> 控制内容层级、- 标识选项顺序等。

【属性词】 如角色、技能、限制等，属性词对模块下内容的总结和提示，用于标识语义结构。

【属性内容】 如技能属性中的具体描述等。

结构化提示词将没有规则的自然语言转化成结构化更强、机器更容易识别的样式和结构。用户在使用过程中保持整体结构不变，只需根据结构不断调整和迭代属性内容，则可通过迭代不断完善提示词，达到用户的预期目标。

第二，智能体逐步实现由人工设计向机器自主设计的转变。上面的示例结构化提示词看起来还挺复杂的，不是说提示词应该越来越简单吗？其实，看起来特

别复杂的提示词是智能体自动化完成的。用户只需要简单的需求指令，智能体即通过对用户指令的理解，自主完成整个提示词的设计，再经过用户的审核、调整和迭代，实现提示词的自动化操作。

比如上述示例结构化是通过扣子平台自动化操作完成的。我们只提供了“用户输入旅游目的地，提供一份详细的旅游攻略”。这一句非常简单的指令，点击右上角优化，即完成了整个提示词的设计，如图 1-7 所示。

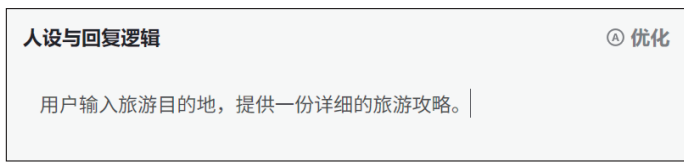

图 1-7　扣子平台中提示词自动化操作界面

除了扣子平台，如 ChatGPT、Kimi、智谱清言、文心一言等平台都有提示词自主设计功能。通过机器的自主设计也将成为提示词的主流模式。这样就实现了由人工设计向机器自主设计的转变，降低了提示词设计的难度。

第三，智能体通过设置插件、工作流、数据库等功能，弱化对提示词的依赖。在对话式人工智能交互中，提示词是最为重要的交互手段。如前文所述，提示词在智能体的运行过程中发挥着关键作用。然而，智能体并非仅依赖提示词这一单一功能，它还能够通过设置插件、工作流、数据库等多样化的功能，增强自身解决问题的能力，同时也降低对提示词的依赖程度，如图 1-8 所示。

图 1-8　扣子平台中创建智能体可设置的部分功能

插件是一种可以扩展智能体功能的小型程序。通过安装不同的插件，智能体可以获得各种特定的能力，如图像识别、语音合成、自然语言处理等。这些插件可以根据用户的需求进行定制和安装，从而使智能体能够更好地满足不同用户的个性化需求。工作流是一种将多个任务和操作按照一定的顺序和逻辑组合在一起的流程。通过设置工作流，智能体可以自动执行一系列复杂的任务，而无须用户逐个输入提示词。数据库可以存储大量的信息和数据，为智能体提供丰富的知识储备。通过与数据库的连接，智能体可以快速检索和获取所需的信息，从而更好地回答用户的问题和处理任务。

如图 1-8 所示，除了插件、工作流、数据库等功能，智能体还提供了图像流、触发器等功能。通过这些功能设置，可以使智能体在处理任务时更加灵活、高效。这些功能和提示词一起，共同为智能体提供强大的支持。

（三）智能体重塑人机交互方式的未来

在人工智能技术蓬勃发展的推动下，人机交互方式经历了显著的优化与演进。从最初的 Chatbot（聊天机器人）到先进的 Copilot（协作伙伴），再到如今更为智能的 Agent（智能代理），这一系列的变革不仅极大地丰富了交互手段，还显著提升了人工智能在人类生产生活中的重要性与实用性。

最初的 Chatbot 主要用于简单的对话和信息查询，用户通过明确的指令与其进行互动。Chatbot 基于预设的规则和关键词进行回应，适用于回答常见问题、提供基本的客户支持和信息查询等场景。虽然 Chatbot 作为辅助工具，减轻了人类在简单重复任务中的负担，但其能力有限，难以处理复杂任务。比如，比较流行的智能音箱，当用户对智能音箱说“播放一首周杰伦的歌”，它会根据预设的指令和关键词，识别出“播放”“周杰伦”等信息，然后播放周杰伦的歌曲。

随着技术的进步，Copilot 应运而生。Copilot 能够辅助用户进行决策，并实现部分任务的自动化。当用户提出任务目标后，Copilot 会提供建议并自动执行部分任务。它可以理解用户输入的上下文信息，提供更智能和相关的应答，从而提升人机交互的自然度和效率。在办公自动化、编程辅助、内容创作等场景中，Copilot 通过部分任务的自动化执行，减轻了用户的工作负担，使人类能够专注于更高层次的工作。

假如用户是一位文案策划人员，向人工智能提出任务目标“为一款新推出的智能手表撰写一份营销文案，突出其时尚的外观和强大的功能”。这时候，人工智能会根据这个要求，提供一些建议。比如，可以提及手表的材质、设计风格来

体现时尚的外观，列举具体的功能特性，如长续航、精准的健康监测等，以突出其强大的功能。同时，它还自动生成一些文案的开头和关键段落。文案策划人员可在此基础上进行润色和调整，从而减轻他们的工作负担，使他们能够将更多精力放在思考独特的营销创意和策略制定等高层次工作上。在这种场景中，人工智能担任了 Copilot 角色，比 Chatbot 的智能化水平更高。

随着人工智能进一步发展，Agent 出现了，最有代表性的产品就是智能体。智能体具备自主规划和独立执行任务的能力。用户设定目标后，智能体会自主规划并完成任务，用户则无须关注具体的执行细节，而只需关注任务的结果，并对智能体的执行情况进行监控和反馈。智能体在智能家居、自主驾驶、金融分析等领域展现了其强大的能力，通过自主执行复杂任务，减少了用户的直接干预，使人工智能成为真正的生产力工具，显著提升了生产效率，改变了传统的生产模式。

比如，一家电商公司每天需要处理很多图片，以前是专职人员通过图片处理软件一张张修改，现在只需将这些图片批量化上传到智能体中，智能体便会按照用户的需求对图片进行加工。

在武汉试运营的“萝卜快跑”自动驾驶汽车也是一种智能体。用户只需设定目标，如“从家安全快速地到达公司”，“萝卜快跑”便会考虑交通流量、道路施工等情况，自主规划行驶路线。在行驶过程中，它自动控制车辆的加速、减速、转向等操作，无须用户亲自驾驶。这种智能体的应用减少了用户在出行过程中的直接干预，提升了出行的便捷性和效率，改变了传统的出行模式。

人工智能从 Chatbot、Copilot 到 Agent 的演变过程，展现了人机交互方式的重大转变。随着每一阶段的进化，机器在人类生产中的角色不断提升，从辅助工具发展为重要的合作伙伴，最终成为能够独立执行任务的自主智能体。这预示着未来的人机交互将更加智能化和高效化。人机交互方式转变如表 1-1 所示。

表 1-1　人机交互方式转变表

模式	交互方式	应用场景	重要性
Chatbot	指令式交互，预设对话路径	客户服务，信息查询	减轻简单重复任务的负担，能力有限
Copilot	协作式交互，上下文理解	办公自动化，编程辅助，内容创作	辅助决策和部分任务自动化，提高工作效率和准确性
Agent	目标式交互，监控与反馈	智能家居，无人驾驶，金融分析	自主执行复杂任务，显著提升生产力，改变生产模式

四、回顾智能体：简史与工作原理

既然智能体展现出了如此强大的实力，给人们的生活和工作带来了翻天覆地的改变，那么我们必然需要对智能体进行全面、深入的了解。智能体究竟是如何一步步发展至今的？它的设计框架是怎样构建的，工作原理又是什么呢？只有对这些方面有了清晰的认识，我们才能更好地把握智能体的本质，充分发挥其优势，为人们的生活和工作带来更多的便利与创新。

（一）智能体简史

智能体是一个有着悠久历史的概念，其思想可以追溯到东西方的古代哲学家。如老子提出的“道生一、一生二、二生三、三生万物”思想，便是一种对宇宙和万物生成的哲学性描述，也可以视为对智能体底层思维方式和逻辑的阐释。从现代科学的角度来看，智能体的概念可以追溯到阿兰·图灵（Alan Turing）在 20 世纪 50 年代提出的图灵测试。这个测试作为人工智能的基石，旨在探索机器能否表现出与人类相当的智能行为。这些人工智能实体通常被称为“代理”（agent）。此后，科学家们尝试采用不同的技术路线来提高人工智能实体的智能水平，包括符号代理、反应式代理和强化学习等，但这些方法的效果都不尽如人意。

直到近几年，随着大语言模型表现出令人印象深刻的涌现能力，越来越多研究人员开始利用这些模型来构建智能体。2019 年，OpenAI 发布了 GPT-2 自然语言处理模型，并在 2020 年和 2022 年相继发布了 GPT-3、DALL·E 2 及 GPT-3.5。以 ChatGPT 为代表的大语言模型的快速发展，为智能体的发展与应用提供了新的契机。2023 年 3 月 14 日，OpenAI 发布了 GPT-4；不久之后，以 ChatGPT 为底层技术的 AutoGPT 横空出世，并迅速火遍全球，成为第一款成功“出圈”的智能体。到了 11 月，OpenAI 推出了自定义式 GPT——GPTs，用户无须编程基础，即可按照自己的需求自由创建 GPT 应用。这标志着智能体的普及化应用的真正实现。此时，以大语言模型为底层技术的智能体蓬勃发展，国内外数百家公司开发了各种智能体应用，广泛渗透到各个行业和领域，推动了社会的深刻变革和进步。

（二）智能体原理解密

在当今的大模型时代，智能体如同一个充满智慧的助手，能够处理各种复杂的任务和上下文信息。这一切的实现都要归功于大语言模型的强大支持。大语言

模型不仅提升了智能体的理解力和泛化能力，还增强了其自然语言处理能力，使得交互体验变得更加个性化和连续。智能体具备两个显著的工作系统：内在的核心能力和外在的交互来源。这构成了智能体基本的设计和工作框架，如图 1-9 所示。

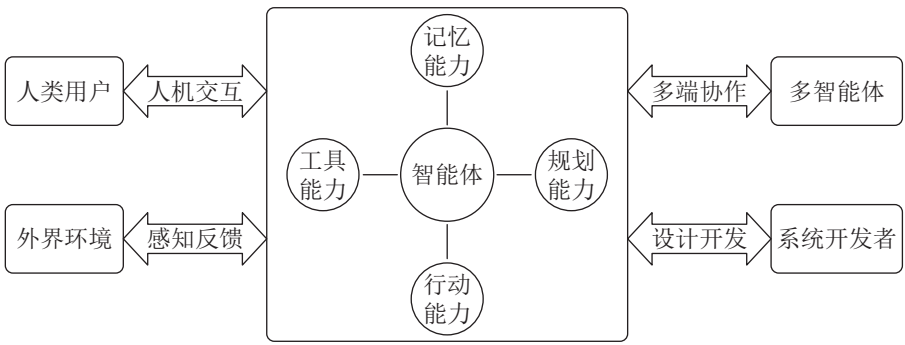

图 1-9　智能体设计框架和工作原理 ①

智能体的内在核心能力可以分为四大类：规划能力、记忆能力、工具能力和行动能力。这 4 项能力是智能体在复杂环境中高效执行任务的关键基础，能够确保其灵活应对各种挑战和需求。

首先，规划能力包括目标分解能力和任务反思能力。目标分解能力使得智能体能够将大型或复杂的任务分解为更小、更易管理的子任务，从而更有效地处理复杂问题。例如，在设计市场营销方案时，智能体会将整个计划拆解为多个阶段性的任务，如市场调研、目标客户群分析、广告内容制作和投放策略的制定。通过这种结构化的方式，智能体能够有条不紊地完成每一阶段的任务，确保每个子目标都得到充分处理，显著提高整体任务的完成效率。同时，任务反思能力让智能体能够对其过去的行为进行深入的自我批判和反思，从错误中吸取经验，并为接下来的行动提供有价值的分析和总结。在软件开发项目中，智能体会在每次迭代后，回顾哪些开发步骤有效，哪些步骤存在问题，从而在后续的开发计划中进行调整，避免重复过去的错误，最终优化开发流程，提高项目质量和效率。

其次，智能体的记忆能力分为短期记忆和长期记忆两类。短期记忆使得智能体能够在当前对话或任务中利用即时上下文信息进行学习和决策。例如，在客户服务对话中，智能体能够记住客户在对话开始时提出的问题和细节，并在整个对话过程中使用这些信息提供一致且相关的帮助和解决方案。长期记忆则使智能体

① 本图参考甲子光年发布的报告《2024 年 AI Agent 行业报告》，地址：http://www.jazzyear.com/study_info.html?id=135。

能够在较长时间内保存和回忆信息，通常通过外部向量存储和快速检索来实现。例如，在医疗诊断应用中，智能体可以记住病人的病史、过敏信息以及以前的诊断结果，即使在数月或数年后，智能体仍然能够快速检索这些信息，为病人提供连续性护理和个性化的治疗建议。通过这种外部存储解决方案，智能体在需要时能够迅速、准确地访问过去的记录和数据，从而提高服务质量和效率。

再次，工具能力使智能体能够学习如何调用外部 API，以获取其内部模型中缺少的信息。这些信息通常在预训练过程中无法获得，包括最新的动态信息、代码执行能力以及对专有信息源的访问等。通过调用这些外部 API，智能体能够实时获取和处理最新数据，执行特定任务，从而弥补其预训练模型中的信息不足，显著提高处理复杂任务的能力和准确性。例如，天气预报智能体通过调用天气 API，可以获取实时的天气状况、温度、降水概率等信息，并结合其内部模型提供准确的天气预报。当用户询问某地的当前天气和未来几天的天气趋势时，智能体可以实时调用天气 API，获取最新的数据，并提供详细的预报和建议。

最后，行动能力是智能体的一项核心能力。它使得智能体能够执行各种任务，从而更好地适应环境的变化。通过与环境的持续互动和反馈，智能体不仅能够应对变化，还能够影响和塑造其所处的环境。例如，在智能家居系统中，智能体可以控制家中的各种设备，如灯光、恒温器和安防系统。当检测到有人进入房间时，智能体会自动打开灯光并调整温度；当家中无人时，系统会切换到节能模式。同时，智能体还能够根据用户的生活习惯和反馈，不断优化这些操作，使家居环境更加舒适和节能。

智能体的强大不仅体现在其内在能力上，还源于其与外界的多层次交互。这些交互主要通过以下四类通道实现。

首先，与用户的交互是智能体最直接和重要的外部通道。用户通过各种接口和指令与智能体互动，既充当了监督者的角色，也成为智能体的合作伙伴。这种人机交互不仅提升了智能体的执行效率，还使用户能够直接参与智能体的决策和策略制定过程中。例如，当你请智能助手帮助制订旅行计划时，它会根据你的偏好和实时信息，推荐最佳的旅行路线和活动安排，从而使旅行更加便捷和个性化。

其次，智能体与其所处的环境进行交互，无论是虚拟环境还是物理世界，外界的反馈都会被智能体感知和处理，从而调整其行为和策略。这一交互方式使得智能体能够动态适应环境变化。例如，当你使用智能体进行在线学习时，它能够根据你的学习进度和反馈，动态调整教学内容和难度。如果智能体感知到你在某

些方面遇到困难，它会提供有针对性的练习和解释，帮助你更好地掌握知识。

再次，智能体之间的群体协作也是其不可忽视的特点。多个智能体通过协作，共享任务结果，形成更高层次的群体智能。这种协作不仅提高了任务的完成效率，还拓展了智能体的应用范围。例如，在一个智能物流系统中，多个智能体各自负责不同的环节，如库存管理、路线规划和配送调度。通过共享信息和协调工作，这些智能体大大提高了物流效率和配送速度。例如，创建一个撰写论文的智能体，就可以通过多智能技术，将撰写前言智能体、撰写文献综述智能体、撰写研究理论等智能体进行协作化设置，实现群体智能，从而达到解决复杂问题的目的。

最后，智能体与开发者的交互也是其不断进化的重要途径。系统开发者通过设计和优化智能体的相关能力，使其能够更好地适应各种应用场景。通过持续的研发和改进，智能体变得越来越智能，能够处理更为复杂和多样化的任务。例如，新的智能体模型可以更好地理解和生成自然语言，提供更加精准和个性化的回答，无论是在撰写创意文案、编写代码还是进行复杂的数据分析方面，智能体都表现出越来越强的能力和适应性。这些进步都离不开开发者们的不懈努力和技术的不断革新。

总的来说，大模型时代的智能体不仅具备强大的内在核心能力，还通过与外界的丰富交互，不断提升自身的性能。这样的智能体无疑将成为我们生活和工作的得力助手，为我们带来更加智能化的未来。

五、智能体的分类与典型类型

随着智能体技术的迅猛发展，了解其分类和应用平台显得尤为重要。智能体不仅是人工智能的前沿体现，还是推动各行各业变革的重要力量。无论是在家庭生活中的便捷助手，还是在企业中的智能决策支持，智能体都以多样化的形式和功能影响着我们的日常。因此，本节将详细探讨智能体的分类方式，帮助我们更好地理解不同类型智能体的特性和应用场景，为未来的科技探索铺平道路。

（一）智能体的分类

通过对智能体进行分类，我们可以更深入地了解智能体的多样性和功能特性。分类使得我们能够识别和区分不同类型的智能体，从而更好地理解它们各自的用途和优势。

1. 按照使用方式进行划分

按照使用方式进行划分，智能体可以分为应用型智能体、开发型智能体和开

源型智能体。应用型智能体主要用于实际应用场景，具有“即插即用”的特点，通常不需要用户具备编程技能，适合业务人员和普通用户使用。其重点在于易用性和快速部署，如扣子、GPTs 等。开发型智能体面向开发人员和技术团队，需要用户具备编程能力，尤其是 Python 编程，适合深度定制和复杂应用开发，如 AutoGPT、MetaGPT 等。开源型智能体则是开源的，用户可以自由访问和修改源代码，通常适合开发人员和研究人员使用。开源型智能体可以用于学习、研究和深度定制，如 AutogenStudio 等。

2. 按照用户类型进行划分

按照用户类型进行划分，智能体可以分为面向 C 端用户、B 端用户和 G 端用户的类型。面向 C 端用户的智能体框架主要针对普通消费者，强调用户体验、易用性和便利性，通常用于个人娱乐、教育和日常生活中的辅助工具，如 GPTs、文心一言等。大语言模型公司主要开发面向 C 端用户的产品。面向 B 端用户的智能体则主要服务于企业和商业用户，强调功能的广泛性、定制化和可扩展性，通常用于提高企业效率、优化业务流程和提供客户服务，如澜码科技打造的“AskXBot”平台等。面向 G 端用户的智能体主要服务于政府和公共部门，强调数据安全、隐私保护和系统的可靠性，通常用于公共服务、政策制定和行政管理。

3. 按照智能体工作原理进行划分

按照智能体的工作原理进行划分，智能体可以分为对话式智能体、自主智能体和生成智能体。对话式智能体（conversational agent）通过自然语言与人类互动，完成各种任务，主要用于回答问题、提供建议和帮助用户解决问题，强调语言理解和答案生成能力，如 GPTs、文心一言等平台都属于这一类。自主智能体（autonomous agent）能够根据用户通过自然语言提出的需求，自动执行任务并实现预期结果，如 AutoGPT 能够通过理解自然语言需求并自动完成任务。生成智能体（generative agent）则是在模拟复杂社会环境中“生活”的智能体，拥有自己的记忆和目标，能够与人类和其他智能体互动，例如斯坦福和 Google 的研究者联合构建的“Smallville”虚拟小镇中的生成智能体，不仅能够与人类互动，还能在模拟的社会环境中进行复杂的交流和互动，模拟真实世界中的社会动态。

4. 按照智能体形态进行划分

按照智能体的形态进行划分，智能体可以分为原生 AIGC 创业型、互联网巨头企业型、企服软件 /SaaS 服务商型、RPA 型和 3C 硬件型。原生 AIGC 创业

型企业以 AIGC（AI 生成内容）为基础，具备大模型算法的优势，能够借助 AI Agent 实现 AI 技术的商业落地，如国外的 OpenAI 和国内的智谱清言等公司。互联网巨头企业型企业则具备丰富的互联网场景成功经验，同时兼顾通用大模型和云服务能力，为个人和企业提供智能体服务，如微软的 AutogenStudio、百度的文心一言等。企服软件 /SaaS 服务商型企业长期根植于中国企业的数字化进程，具备企业数字化工作全流程的丰富经验，并在此基础上为其他企业提供专业化的智能体服务，如用友的用友大易平台。RPA 型企业在机器人流程自动化（RPA）建设方面有丰富经验，能够在垂直领域提供高度自动化的智能体解决方案，如 UiPath、实在智能和达观 AI Agent 等。3C 硬件型企业则利用 3C（计算机、通信、消费电子）消费电子产品的优势，通过 AI Agent 特性升级自身的手机、音响、平板等多端产品的用户体验，如华为在各类终端上升级的小艺智能体和联想推出的 AI PC 个人智能体。

（二）智能体的典型类型：AutoGPT

AutoGPT 是一种由最新 GPT 技术驱动的开源程序，能够自主完成各种任务。它就像一个聪明的助手，在不需要过多人工干预的情况下能够高效运行。这种技术展示了人工智能在自主任务执行方面的巨大潜力，能够帮助我们处理各种复杂的任务。

AutoGPT 的工作原理非常有趣，下面以“撰写一篇开学发言稿”为例，来详细解释它的操作过程。首先是目标设定与分解。用户设定的总体目标是撰写一篇开学发言稿。AutoGPT 会将这个目标拆解成多个小任务，如确定发言稿的主题、研究相关素材、撰写开头、撰写主体段落、撰写结尾，以及进行内容润色等。接下来是生成提示。对于每个小任务，AutoGPT 会生成一系列提示。例如，在确定发言稿主题时，提示可能包括“回顾上学期的亮点”“展望新学期的计划”和“激励学生保持积极态度”等。在撰写开头段落时，提示可能是“用一句名言或引言开场”“表达对新学期的期待”等。工作过程如图 1-10 所示。

在自主执行阶段，AutoGPT 根据生成的提示开始执行任务。例如，在撰写开头段落时，AutoGPT 会生成几种不同的开场白，然后选择最合适的一种，按照提示逐步撰写主体段落和结尾段落。然后是反思与评估——每完成一项任务，AutoGPT 会对结果进行反思和评估。例如，它会检查开头段落是否足够吸引人，主体段落是否清晰传达了主题，结尾是否具有激励性和鼓舞性。如果发现问题，AutoGPT 会记录这些问题，并准备在下一步进行调整。最后是调整与下一步行动，根据反思和评估的结果，AutoGPT 决定下一步行动。如果发现开头段落不够

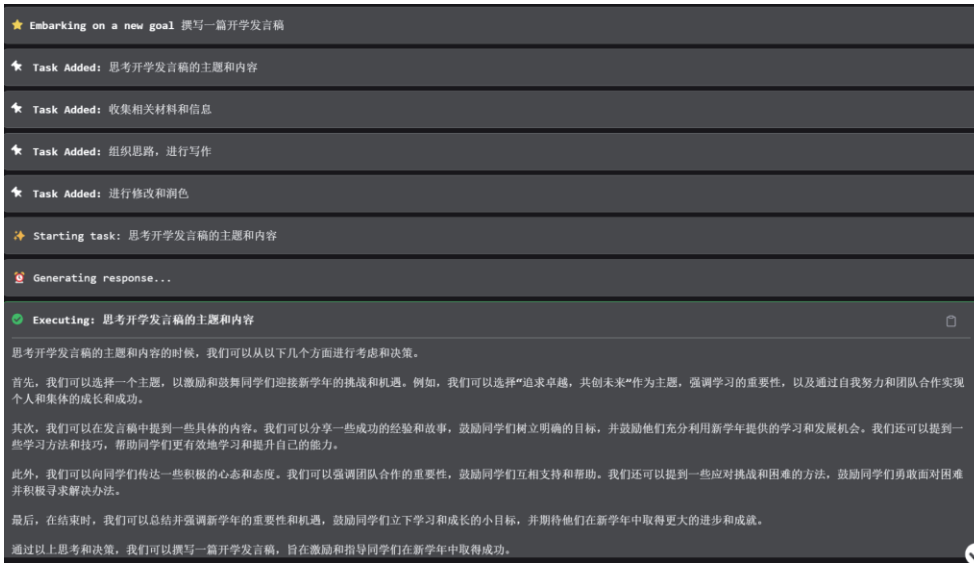

图 1-10　AutoGPT 工作界面

吸引人，AutoGPT 会生成新的开场白进行替换，可能还会调整主体段落的逻辑结构，增加更多激励性的语言，或者通过润色使整个发言稿更流畅和具有感染力。

通过这种不断循环的过程，AutoGPT 可以逐步完善开学发言稿，从初步构思到最终定稿，确保内容准确、逻辑清晰且富有感染力，最终达到用户设定的撰写目标。虽然 AutoGPT 展示了人工智能在自主完成任务方面的巨大潜力，但作为一项新技术，它仍在不断发展和完善中。在实际应用中，可能会遇到一些挑战，例如，在某些复杂任务上仍需要人工干预，或者在处理特殊情况时可能会遇到困难以及成本高等问题。然而，随着技术的不断进步，这些问题都有望得到解决。AutoGPT 展示了通用人工智能（AGI）的初步形态，为未来的发展提供了广阔的前景。

（三）智能体的典型类型：Manus

在 2025 年的人工智能浪潮中，一款名为 Manus 的智能体产品横空出世，带来了极大的震撼。Manus 是由中国“90 后”创新者肖弘领导的研发团队打造的全球首款通用型智能体。它被定义为“真正自主的智能体”，不仅能提供建议，还能直接交付任务成果，实现从思考到行动的闭环，重新定义了人机协作的模式。

肖弘是一位连续创业者，2022 年创办了北京蝴蝶效应科技有限公司，并推出了广受欢迎的 AI 插件 Monica。2025 年 3 月 5 日，蝴蝶效应团队推出了 Ma-

nus。产品名称“Manus”源自拉丁语“Mens et Manus”，意为“心智与手”，这一名字精准地诠释了产品的核心理念：AI 不仅仅是思考，更要能够执行。

Manus 最显著的特点是其惊人的自主执行能力。它能够完整地完成任务的全流程：从理解任务、规划步骤，到最终执行操作，整个过程无需人工干预。更为卓越的是，即便用户关闭设备，Manus 依然可以在云端持续工作，轻松应对市场分析、数据建模等需要长时间处理的复杂任务。其通用性令人瞩目，Manus 不再局限于单一领域，而是能够无缝衔接金融、法律、教育、电商等多个行业。无论是企业还是个人用户，都能找到专属的应用场景。对于企业而言，它可以完成供应链匹配和客户挖掘；对于个人用户，它则能提供旅游规划、保险比价等个性化服务。

在性能测试中，Manus 展现出超越人类的惊人能力。根据其官网公布的 GAIA 基准测试数据，Manus 在基础任务、中级任务和高级任务的准确率分别达到 86.5%、70.1% 和 57.7%，而人类在 GAIA 测试中的平均准确率为 92%，Manus 已经接近人类的理解水平。与传统 AI 不同，Manus 具备卓越的交互与学习能力，用户可以在任务执行过程中随时调整指令，AI 能够实时动态地调整执行路径。通过持续收集和学习用户反馈，Manus 不断优化自身的行为模式，实现个性化适配。

为了增强用户信任，Manus 还提供了前所未有的透明化操作，用户可以实时查看 AI 的思考逻辑、浏览的文献和代码生成步骤。通过集成多个专用模型，Manus 能够协同工作，高效处理从数据清洗到分析再到可视化的复杂任务。Manus 的出现，不仅仅是一款产品的发布，更是人机协作模式的重大变革。它为企业降本增效和个人生产力提升提供了核心解决方案，标志着人工智能进入了一个全新的纪元。在这个时代，AI 不再是遥不可及的概念，而是触手可及、能够真正赋能的智能伙伴。Manus 官网截图如图 1-11 所示。

图 1-11　Manus 官网

六、智能体的未来趋势与变革

智能体技术如今正飞速发展，在多个领域彰显出前所未有的巨大潜力。新技术、新思维等众多新要素，有力地推动着智能体朝着更高层次的自主性与智能化不断迈进。智能体的演变绝非仅是技术层面的进步，更是对人类工作方式以及生活方式的一场深刻变革。接下来，将借助具体案例与前沿研究，深入探讨智能体的发展趋势及其未来展望。

（一）从智能体到智能体工作流

著名人工智能科学家吴恩达教授提出了智能体工作流（agentic workflow）的概念。智能体工作流的核心原理是通过循环迭代逐步优化结果，模拟人类解决问题的思维模式。

以学画画为例，帮助我们理解智能体工作流的工作原理。我们刚开始学画画时，作品往往很粗糙，线条不流畅，色彩搭配也不协调。在此基础上，我们通过学习不断迭代画画知识和感知，这就像智能体工作流中的迭代过程。每次完成一幅作品，我们会将当前的作品与之前画得比较好的作品进行比较，分析自己哪里有进步，比如线条是否比上一次更流畅了，色彩搭配是否更和谐了。同时也思考哪里还需要改进，可能是某个物体的形状还不够准确，或者光影的表现还不够生动。随着不断地循环重复练习、观察和比较，我们的绘画水平会一点一点地提高。线条越来越流畅，色彩搭配越来越和谐，画面的表现力也越来越强。这也是智能体工作流中人类解决问题的思维模式。

智能体工作流的具体工作原理是构建一个智能体系统。在这个系统中，多个负责具体功能的智能体通过与大型语言模型协作来完成任务。这些智能体能够自主感知、推理和行动，以实现特定目标，形成强大的集体智慧，能够打破“数据孤岛”，整合不同的数据源，提供无缝的端到端自动化解决方案。

吴恩达教授开发的翻译智能体是“智能体工作流”的一个典型案例[①]。这个翻译智能体通过理解上下文和目标，自动处理多语言翻译任务，不仅能够精确翻译文本，还能根据上下文进行调整，确保翻译的准确性和流畅性。在翻译过程中，智能体自主学习并适应新的语言模式，不断提高翻译质量。这一系统展示了“智能体工作流”在实际应用中的强大能力，通过自然语言处理和机器学习技术的整合，实现了高度自动化和智能化的工作流程。

① 项目地址：https://github.com/andrewyng/translation-agent。

（二）从智能体到具身智能

在特斯拉 2023 年股东会上，马斯克表示，人形机器人将成为特斯拉未来主要的长期价值来源。他提道，“如果人形机器人和人的比例是 2 比 1，那么人们对机器人的需求量可能达到 100 亿至 200 亿个，远超电动车的数量”。这一观点得到了英伟达创始人黄仁勋的呼应。他在 ITF World 2023 年半导体大会上也表示，AI 的下一个浪潮将是具身智能（embodied intelligence）。

所谓具身智能，是指依附于真实世界的物理实体的智能系统，类似于人或动物需要一个肉体来认识世界、探索世界，并通过与环境的交互来影响世界。具身智能不仅要具备感知、认知、推理、决策和持续迭代的能力，还要能够通过与物理实体（如机器人或自动驾驶汽车）的结合，形成能够进行物理交互的智能体。人形机器人是具身智能的典型代表，但其应用场景远不止于此。例如，基于 L4 技术的自动驾驶也属于具身智能的范畴。

具身智能是智能体在物理世界中的具体体现，继承并增强了智能体的核心特征，特别是在与物理环境的互动和适应能力方面。百度旗下的“萝卜快跑”自动驾驶汽车，已经在北京、上海、广州、深圳等城市开展试点。用户可以通过手机应用预约无人车，享受自动驾驶带来的便捷出行体验。“萝卜快跑”就是具身智能在现实生活中的应用，如图 1-12 所示。

图 1-12 “萝卜快跑”无人驾驶汽车

“萝卜快跑”的系统能够理解交通规则和路况信息，进行路线规划和决策。例如，在遇到交通拥堵时，系统可以根据实时路况调整行驶路线，以最快的速度将乘客送达目的地。同时，它还具备推理能力，能够预测其他车辆的行驶轨迹，提前做出相应的反应，确保行驶安全。在行驶过程中，“萝卜快跑”的车辆不断与道路、交通信号和其他交通参与者进行互动。它们根据交通信号灯的指示行驶，与其他车辆保持安全距离，并在遇到行人时及时停车让行。这种互动和适应能力是具身智能的重要特征之一。

除了“萝卜快跑”，还有特斯拉的自动驾驶汽车、波士顿动力的 Atlas 机器人、Google 的 DeepMind 的机器人手臂、亚马逊的 Kiva 机器人，以及软银的 Pepper 机器人等。这些具体案例展示了具身智能在物理实体的交互、复杂任务处理、感知和认知方面的卓越能力。具身智能代表了 AI 技术与物理世界深度融合

的方向。其发展将深刻影响未来社会的各个方面，并且还是智能体的发展方向。通过不断的技术创新和应用拓展，具身智能将在提高生产效率、改善生活质量和推动社会进步方面发挥重要作用。

（三）从智能体到可理解的智能伙伴

如上所述，智能体作为人工智能的一种高阶段形态，正在不断地改变着我们的生活方式和工作模式。然而，随着智能体的应用越来越广泛，其决策过程的不透明性也引发了人们的担忧。这时，可解释的人工智能（Explainable AI，简称 XAI）应运而生，为智能体带来了新的发展机遇。

可解释人工智能是一套针对人工智能系统应用过程生成解释性内容的技术方案。它致力于解决人工智能系统中由模型可解释性不足产生的可靠性、透明性、因果性、公平性和安全性等一系列问题。简单来说，可解释人工智能就是要让人工智能的决策过程变得清晰可见，让人们能够理解为什么人工智能会做出这样的决策。

可解释人工智能对智能体的作用不可小觑。

第一，它提升了智能体的信任度与接受度。以智能理财顾问为例，一个智能体为用户提供投资建议，如果不能解释其决策依据，用户可能会对建议持怀疑态度。但通过可解释人工智能技术，智能体可以向用户解释为什么推荐某些投资组合，是基于市场趋势、用户风险偏好还是其他因素。这样，用户就能更好地理解建议的合理性，从而增强对智能体的信任，更愿意接受其建议。

第二，可解释人工智能便于人们理解和调适智能体。在智能交通领域，自动驾驶汽车的智能体需要不断优化其决策算法。有了可解释人工智能，工程师可以清楚地了解智能体在不同路况下做出决策的原因，从而有针对性地进行调试和改进，提高自动驾驶的安全性和可靠性。

让我们来看一个具体案例。在医疗领域，智能诊断助手作为一种智能体，正在逐渐发挥重要作用。例如，腾讯推出的“腾讯觅影”，利用人工智能技术对医学影像进行分析，辅助医生进行疾病诊断。在这个过程中，可解释人工智能可以为医生提供智能诊断助手做出诊断的依据。比如，当智能诊断助手判断一个肺部 CT 影像中存在结节时，它可以向医生解释是基于影像中的哪些特征做出的判断，如结节的形状、大小、密度等。这样，医生可以结合自己的专业知识和经验，对诊断结果进行更准确的判断。同时，对于患者来说，了解智能诊断助手的诊断依据也可以增加他们对诊断结果的信任，减少不必要的担忧。

可解释人工智能为智能体提供了决策解释，增强了智能体的性能和效果，促进了智能体的发展和创新。在未来，随着技术的不断进步，可解释人工智能和智能体的结合将更加紧密。我们可以期待，可解释人工智能将为智能体带来更高的透明度、可靠性和安全性，使智能体真正成为我们可理解的智能伙伴。同时，随着应用场景的不断拓展，可解释人工智能和智能体将在更多领域发挥重要作用，为人类的生活和社会的发展带来更多的便利和福祉。

七、结语

智能体的出现既是人工智能发展中的一小步，也是具有深远影响的一大步。理解当前的进展是洞察未来的关键。面对百年一遇的新机遇和新格局，我们不仅要观望，还要有勇气投入其中，鼓励大家发挥最大限度的想象力，突破常规边界，去构想人工智能的未来。现实将很快验证我们的想象力是多么有限。智能体是一种技术，更是一种全新的思维方式、一种新生命形态，代表着崭新的文明样态。