

ChatGPT： 通用人工智能的典范

第 1 章





随着技术的发展，人工智能应用取得了巨大突破。2022 年 11 月，由 AI 驱动的自然语言处理工具 ChatGPT 横空出世，以高性能和巨大的应用潜力吸引了全球关注。越来越多的人开始思考，ChatGPT 究竟能够带来什么？

通用人工智能指的是具有广泛的智能性、能够处理多种任务的智能系统。作为通用人工智能的典范，ChatGPT 的出现不仅改变了人机交互方式，还对各行各业的发展产生了深远的影响。本章将对 ChatGPT 的概念、通用能力、GPT-4 大模型进行讲解，全面拆解 ChatGPT 背后的大模型。



1.1 追根溯源：ChatGPT 是什么

作为一款性能强大的 AI 大语言模型，ChatGPT 上市后引发了各界的关注。ChatGPT 火爆的背后，离不开人工智能研究公司 OpenAI 的不懈努力。同时，ChatGPT 强大性能的背后，也离不开 GPT 系列大模型的迭代。

1.1.1 ChatGPT：AI 驱动的自然语言处理工具

ChatGPT 是一款由 OpenAI 推出的、AI 驱动的自然语言处理工具。它通过深度学习技术进行训练，可以完成各种自然语言处理任务，如多轮智能问答、文本生成、语言翻译等。

ChatGPT 的核心逻辑是通过模型训练学习语言的规律，模型训练是基于网络上的书籍、新闻、文章、博客等海量文本数据实现的。基于此，模型能够掌握语法、海量词汇，理解上下文之间的关系。模型会通过无监督学习的方法学习各种语言知识，并通过海量文本数据推测语言中的规律。这使得 ChatGPT 成为一种强大的自然语言处理工具，能够理解、生成各种内容。

用户与 ChatGPT 对话时，ChatGPT 会接收用户输入的文本信息，并对信息进行处理。模型将基于这些文本信息的预测结果转化为文本形式，作为回答反馈给用户。在这个过程中，模型理解语言的能力、生成回答的能力都会不断迭代进化，并根据与用户的交流不断调整回答。

但是 ChatGPT 生成内容也存在一些局限性。模型基于训练数据进行学习，

如果数据中包含存在偏见、错误的信息，模型会基于此产生存在偏见或者错误的回答。为了解决这一问题，OpenAI 不断地使用更高质量的数据优化语言模型，设计更加完善的监管机制，以提高模型的可靠性。

综上所述，ChatGPT 是一款基于大规模文本训练的大语言模型，可以基于自然语言理解、深度学习等技术，理解和生成各种内容，完成多种自然语言处理任务。

1.1.2 从 GPT-1 到 GPT-4，ChatGPT 的前世今生

ChatGPT 的出现和发展离不开 GPT 系列大模型的支持。GPT 全称为 Generative Pre-Trained Transformer (预训练生成式转换器)，是一种生成式预训练 Transformer 模型。从初代版本的 GPT-1 到 GPT-4，GPT 系列大模型的性能和能力持续迭代，ChatGPT 在这一发展过程中出现并不断进化。

1.GPT-1

2018 年 6 月，OpenAI 发布 GPT 系列大模型的初代版本 GPT-1。GPT-1 在训练过程中依赖数据标注以及模型微调，语言泛化能力不足。GPT-1 并不是一种通用语言模型，更像一种处理特定语言任务的专用模型。

GPT-1 的模型训练分为两个阶段。首先，通过无监督学习进行预训练，生成语言模型。其次，根据问答、自然语言推理等特定任务的要求，对模型进行微调。GPT-1 在处理多种语言任务方面有不错的表现，在问答、自然语言推理等方面超越了之前的语言模型，可以根据提示或上下文生成流畅的回答。但是，GPT-1 在生成文本方面存在局限性，如容易生成重复性文本、无法对多轮对话进行推理等。

2.GPT-2

GPT-2 于 2019 年 2 月发布。相较于 GPT-1，GPT-2 是一个泛化能力更强的语言模型，具有一定的通用性。GPT-2 的通用性体现在可以应用到多种任务场景中，且不需要经过专门的训练。相较于 GPT-1，GPT-2 可以通过对大规模数据的预训练，具备解决多种语言任务的能力。

GPT-2 能够生成连贯且自然的文本，但在复杂的文本推理、上下文理解等方面存在缺陷，难以在较长的文本输出中实现上下文连贯。



3.GPT-3

GPT-3 于 2020 年 5 月发布。GPT-3 在训练过程中引入了 in-Context Learning (上下文学习)，即在训练模型时，在输入的文本中加入多个示例，引导模型输出相关内容。in-Context Learning 包括三种学习模式，分别是零样本学习、单样本学习和少样本学习。其中零样本学习指的是没有示例，只给出提示；单样本学习指的是只给出一个示例；少样本学习指的是给出多个示例。in-Context Learning 的优势在于，可以让模型从示例中进行学习，无须进行模型微调和数据标注，降低模型训练成本。

GPT-1 的参数为 1.17 亿个，GPT-2 的参数为 15 亿个，GPT-3 的参数量有了显著提升，高达 1750 亿个，是一个规模超大的语言模型。在机器翻译、智能问答等自然语言处理任务中，GPT-3 都有出色的表现。同时，在海量参数的支持下，GPT-3 能够完成更加复杂的任务，如生成新闻报道、生成代码等。

GPT-3 功能强大，但存在滥用的风险，一些不法分子可能会基于 GPT-3 生成虚假新闻、恶意软件等。

4.GPT-4

2023 年 3 月，OpenAI 发布了 GPT 系列大模型的新版本 GPT-4。相较于之前的版本，GPT-4 在各项能力上有了质的突破。除了文本生成能力、对话能力等大幅提升外，GPT-4 还迈出了从大语言模型向多模态模型转变的第一步。除了文本输入外，GPT-4 还支持图像输入，可以实现图像优化、图像转文字等。

在图像识别和理解方面，GPT-4 可以实现图像输入，理解图像内容并生成相关分析。例如，GPT-4 可以根据一张草图，生成一个完整的网站；可以根据食品照片，分析出其制作过程；可以根据植物照片，分析植物的病症等。

在内容生成方面，GPT-4 可以生成歌曲、绘画作品、剧本、营销文案等，内容更加专业。在内容创作过程中，GPT-4 能够模仿不同用户的创作风格，满足用户的个性化需求。

GPT-4 的推理能力也有了大幅提升，在各种专业考试中展现出了与人类相当的推理水平。例如，GPT-4 通过了美国律师资格考试，得分超过 90% 的考生。

此外，GPT-4 接受了大量恶意提示的训练，具有更强的内容辨别能力，在内容真实性、风险可控性方面有了一定的进步。

回顾 GPT 系列模型的发展历程，从 GPT-1 到 GPT-4，GPT 系列模型的性能实现了质的飞跃。ChatGPT 在 GPT 系列模型发展的过程中应运而生。初代 ChatGPT 搭载的是 GPT-3.5 模型，可以完成智能对话、文本内容生成、图片内容生成等多种任务，但推理能力和智能性有待提升。而在 GPT-4 模型出现之后，ChatGPT 在内容创作、图像理解、逻辑推理等方面的能力都实现了飞跃，这为 ChatGPT 的广泛应用奠定了坚实的基础。

1.2 通用能力：ChatGPT 四大功能

ChatGPT 的通用能力主要表现在四个方面，分别是基于海量数据的内容智能生成、区别于传统搜索方式的智能搜索、支持多种语言批量翻译的智能翻译和赋能智能机器人。

1.2.1 内容智能生成：基于海量数据生成多种内容

基于对大量文本数据的预训练，ChatGPT 能够更加智能地理解和生成语言，完成各种内容生成任务。同时，基于 GPT-4 强大的深度学习能力，ChatGPT 可以从海量数据中提取信息，优化模型，提供更加智能的解决方案。

ChatGPT 在内容智能生成方面的表现十分出色。它主要可以生成以下几种类型的内容，如图 1-1 所示。

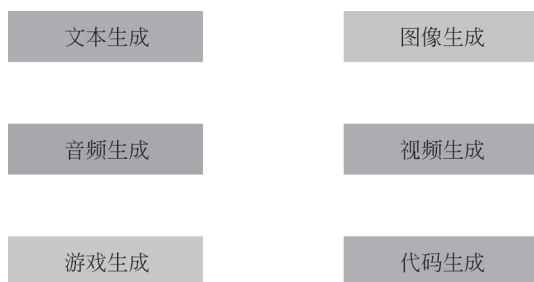


图 1-1 ChatGPT 可以生成的内容类型



1. 文本生成

在文本生成方面，ChatGPT 生成的内容十分多样，可以根据用户提问与用户对话、搜索海量教学资料生成教案、根据新闻采编内容生成新闻报道、根据小说内容完成小说续写等。

2. 图像生成

在图像生成方面，ChatGPT 可以根据用户指令生成相关图像。在此基础上，ChatGPT 还可以接收用户指令，对图像的细节进行调整。ChatGPT 也可以根据用户要求生成图文结合的海报、画报等，还可以生成复杂的设计图样。以建筑效果图为例，用户只需输入建筑物尺寸、材料、颜色等信息，ChatGPT 便可根据这些信息生成逼真的建筑效果图。

3. 音频生成

ChatGPT 可以生成流畅的语音内容：一方面，ChatGPT 可以实现自然、真实的语音合成，用户可以根据自身需求定制个性化的语音风格；另一方面，ChatGPT 可以实现多样的音频创作，如编曲、音乐创作、生成音效等。

4. 视频生成

在视频生成方面，ChatGPT 可以根据用户的要求生成富有创意的视频，助力用户创作。以制作短视频为例，用户需要准备好短视频脚本、音频素材、文字等，同时在脚本中说明各种素材的使用场景。ChatGPT 能够根据用户输入的脚本生成个性化、符合用户需求的短视频。

视频软件 Wondershare Filmora 在其网页版中上线了智能脚本小助手功能。该功能以 ChatGPT 为支撑，可以智能生成故事、演讲稿、两人对话等多种脚本，为用户创作短视频提供助力。

5. 游戏生成

在游戏生成方面，ChatGPT 能够提供多种助力。ChatGPT 可以提升游戏内 NPC（Non-Player Character，非玩家角色）的智能性，生成自然的对话内容，提升玩家的互动体验。ChatGPT 也可以实现游戏剧情的智能生成。ChatGPT 可以助力游戏人物背景、故事剧情等内容的创作，并生成游戏地图、关卡、道具等。总之，ChatGPT 可以大幅提升游戏开发的效率。

6. 代码生成

借助自然语言处理技术，ChatGPT 可以快速了解用户需求，生成满足用户需求的代码。用户可以将烦琐、重复的代码生成任务交由 ChatGPT 处理，提高代码生成效率。同时，ChatGPT 生成代码支持用户自定义模板，便于用户调整代码生成的规则和格式。

ChatGPT 在内容智能生成方面具有生成内容自然流畅、生成多元化内容高效等特点。同时，ChatGPT 还能够根据用户需求生成定制化内容，满足用户的个性化需求。

1.2.2 智能搜索：ChatGPT 颠覆传统搜索方式

AI 的发展使得搜索引擎发生重大变革。ChatGPT 将颠覆传统的搜索方式，实现智能搜索。

传统搜索引擎将用户输入的关键词与资料库中的关键词进行匹配，并按照一定的算法对结果进行排序。然而这种方法存在搜索引擎无法理解复杂关键词、无法解析长文本、缺乏交互能力等弊端。

而 ChatGPT 催生一种新的智能化搜索方式，实现搜索引擎与用户的自然交互。ChatGPT 在智能搜索方面具有以下优势。

(1) 精准搜索。与传统搜索引擎相比，基于强大的自然语言处理能力，ChatGPT 能深刻理解用户的搜索意图，提供更加精确的搜索结果。

(2) 个性化搜索。ChatGPT 能够结合用户的个人信息、浏览记录和搜索偏好，为用户提供个性化搜索结果，满足用户的需求，提高用户的体验。

(3) 对话式交互。用户提出问题后，ChatGPT 能够以对话的形式向用户提供答案，省去了用户反复点击链接的麻烦，使搜索更加便捷。

基于精准搜索、个性化搜索、对话式交互等优势，ChatGPT 能够极大地提升用户的搜索体验。用户能够以自然语言交互的方式，获得精准的搜索结果。

ChatGPT 将成为下一代搜索引擎的催化剂。在搜索引擎发展的过程中，更加便捷、便于交互的搜索引擎更能满足用户的需求，也更具发展潜力。而 ChatGPT 不断进化，实现了从“模糊搜索”到“精准搜索”的跨越，为未来搜



索引引擎的发展指明了方向。

1.2.3 智能翻译：支持多种语言批量翻译

ChatGPT 具有智能翻译功能，支持多种语言批量翻译，用户能够快速获得高质量翻译结果。ChatGPT 的智能翻译功能能够帮助用户在多种语言环境下开展工作或学习。

用户使用 ChatGPT 翻译文本时需要遵循以下几个步骤：首先，用户需要准备好需要翻译的文本，并确认翻译的语言，例如，中文翻译成英语或者日语；其次，将文本输入 ChatGPT 对话框中，并说明需求；最后，点击发送输入的内容，ChatGPT 将会自动输出翻译好的文本。

用户在使用 ChatGPT 翻译文本时，需要注意以下几点。

(1) 虽然 ChatGPT 性能强大，但仍无法与人工翻译相媲美，因此，用户需要对翻译的结果进行核查，自行修改不合理之处。

(2) 用户需要对文本进行整理，将具有关联性的文本整合在一起，这样更有利于 ChatGPT 准确翻译。

(3) 为了确保翻译结果的准确性，用户可以先让 ChatGPT 翻译少量内容，如果结果相对准确，再进行批量翻译。

ChatGPT 进行批量翻译具有许多优势，如图 1-2 所示。



图 1-2 ChatGPT 进行批量翻译的优势

(1) 翻译质量高。ChatGPT 拥有基于海量数据训练的大模型，能够保证翻译结果的准确性。

(2) 便于操作。ChatGPT 能够为用户提供简洁、明了的操作界面，方便用户快速操作，获得翻译结果。

(3) 提高效率, 降低成本。用户只需要在 ChatGPT 对话框中上传文本, 便可获得翻译结果。相较于耗时、耗力的人工翻译, ChatGPT 能够降低翻译成本, 提高效率, 适合个人用户或者企业使用。

总之, ChatGPT 的智能翻译功能适合批量处理文本翻译任务。随着 ChatGPT 的发展和智能翻译功能的升级, 其将更加高效、高质量地完成文本翻译任务。

1.2.4 赋能智能机器人: 提高服务质量, 提升智能性

机器人是一种通过编程、自动控制执行任务的机器。随着多模态感知、深度学习、定位导航等技术的发展, 机器人变得越来越智能。智能机器人被广泛应用在智能制造、酒店餐饮、医疗等场景中。日常生活中的教育机器人、伴读机器人等也属于智能机器人的范畴。

智能机器人具备以下四大要素, 如图 1-3 所示。

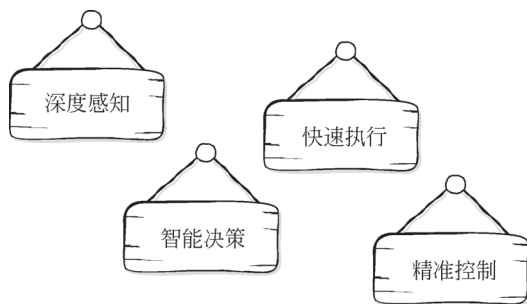


图 1-3 智能机器人具备的四大要素

其中, 深度感知指的是智能机器人能够接收各种传感信息, 获取周围环境、目标对象的数据; 智能决策指的是智能机器人能够对接收的各种数据进行分析; 快速执行指的是基于完善的结构和驱动系统, 智能机器人能够依据指令快速完成任务; 精准控制指的是智能机器人融合了定位导航、障碍规避等技术, 能够合理规划行动路径, 实现精准控制。

ChatGPT 的一大功能就是能够为机器人赋能, 提高机器人的智能性。ChatGPT 的技术优势主要体现在以下几个方面。



(1) 具有强大的自然语言理解能力。基于大规模数据训练，ChatGPT 能够更好地理解自然语言，不受语法、词汇等的限制，对自然语言的理解更加准确。

(2) 具有更强的自我学习能力。ChatGPT 能够将学习到的语言规则应用于新任务中，不断学习和自我完善，提高自身智能化水平。

(3) ChatGPT 能够根据用户需求针对某些领域（如医疗、金融等）进行定制化训练，为用户提供更加专业的服务。

ChatGPT 的以上优势对机器人的发展有深刻影响。一方面，ChatGPT 可以提高机器人的理解能力，机器人可以更高效、准确地理解自然语言和语义，智能化程度更高。同时，机器人可以更加准确地回答用户的问题，提高运行效率。另一方面，ChatGPT 可以降低开发和维护机器人的成本。在 ChatGPT 的赋能下，企业可以通过深度学习技术和大模型的应用节省开发和维护机器人的成本，减轻负担。

ChatGPT 可以从多个方面为机器人产业赋能。例如，在智能客服机器人领域，ChatGPT 可基于海量数据提高智能客服机器人自主处理任务的能力，如准确回答用户问题、为用户提供专业的解决方案等，提高智能客服机器人的服务水平；在智能工业机器人领域，ChatGPT 可以用于训练智能工业机器人，使其能够完成更加复杂的工业任务，提高生产质量和效率。

总之，ChatGPT 能够为机器人赋能，提高机器人的自然语言理解能力、自我学习能力、交互能力，提升机器人的服务质量和智能性。在此基础上，企业可以降低机器人的运营成本，用户可以获得更好的服务体验。



1.3 GPT-4 引领通用人工智能风口

在 GPT 系列大模型迭代的过程中，提高大模型的通用性始终是 OpenAI 的目标。相较于此前的 GPT 系列大模型，GPT-4 在通用性方面有了很大提升，拉开了通用人工智能的序幕。

1.3.1 通用人工智能成为 AI 发展的下一阶段

GPT-4 大模型的通用能力让人们窥见了通用人工智能的冰山一角，也展现了人工智能从专用人工智能到通用人工智能的发展趋势。通用人工智能将成为 AI 发展的下一阶段。

通用人工智能是什么？和只能处理单一任务的专用人工智能相比，通用人工智能可以处理更加广泛和复杂的任务。在处理任务时，通用人工智能可以跨越多个学科、领域、模态，大幅提升任务处理效率。基于强大的任务处理能力，通用人工智能更适用于多元化场景，发展潜力巨大。

专用人工智能和通用人工智能有何区别？专用人工智能的优势在于，在某一领域优势显著，能够高效处理专业任务。但其缺点在于，具有很强的局限性，无法解决新问题。通用人工智能不针对特定的领域，具有灵活分析问题和解决问题的通用基础能力，包括自然语言理解能力、逻辑推理能力等。相较于专用人工智能，通用人工智能具备底层通用能力，可以灵活地分析、解决新问题。

通用人工智能不局限于任何具体应用，生成内容只是其无数应用中的一种。在实际应用中，通用人工智能可以作为一种生产力工具或者基础设施，赋能各行各业，提高全社会生产效率。

以软件行业为例，通用人工智能将大幅提升软件价值。当前，To C（To Consumer，面向个人）软件运行效率低，学习门槛高。以 Office 套件为例，很多用户只使用其中的一些简单功能，而较少使用高级、复杂的功能。在接入通用人工智能后，用户可以直接向软件提出需求，软件可以根据用户的需求自动调用相关高级功能，更好地满足用户的需求。因此，对于 To C 软件而言，很多以前用户很少使用的功能都需要重新定价。

To B（To Business，面向企业）软件也存在一些问题，如定制化服务周期长、成本较高、软件难以复制等。而接入通用人工智能的 To B 软件，能够在一定程度上解决以上问题。

通用人工智能之路通向何处？随着通用人工智能不断迭代，其可能会在未来



拥有类似人类的能够完成各种任务的通用学习能力，通过自适应学习、跨模态学习，不断提升自身通用能力，在更多领域落地。

1.3.2 大模型：实现通用人工智能的最佳路径

通向通用人工智能的路径是多元化的，其中，大模型是实现通用人工智能的最佳路径。大模型具有强大的通用性、实用性，能够降低人工智能应用开发门槛，提高模型精度，提高内容生成效率。大模型是人工智能技术的一大突破，成为通用人工智能发展的起点。

以 GPT-4 为例，除了完成各种语言任务外，GPT-4 还可以完成编程、医学、心理学等多个领域的复杂任务。GPT-4 在一些领域的表现甚至可以达到人类的水平。GPT-4 还可以自动学习各种新内容，不断自我优化，提升自身智慧性。

以往的 AI 应用难以全面覆盖产业流程，而大模型能够覆盖产业流程的每个环节。以自动驾驶场景为例，在输入层，大模型可以感知环境，生成丰富的实景图片；在输出层，大模型可以重构 3D 环境、进行路径规划等。大模型能够实现自动驾驶感知决策一体化，做出科学的驾驶行为预判断，提升自动驾驶的可靠性。

大模型正在从单一模态数据输入向多模态数据输入迭代。文本、语音、图像等多模态的联合训练，可以实现不同模态之间的互补。这有助于提升模型的效果和泛化能力，为通用人工智能的发展奠定基础。

多模态大模型能够感知多种模态的数据，具备识别、处理数据的能力，能够根据数据分析结果做出决策或执行决策。多模态大模型具有通用性，但需持续迭代，以拥有更强大的通用能力。

未来，随着大模型的发展，用户与模型的互动将变得更加自然。大模型将不断提升自己的理解能力、推理能力，更好地理解用户的意图并解锁更多技能，输出高质量的内容。通用人工智能将在大模型的支撑下更具智慧。

1.3.3 OpenAI 公布通用人工智能规划

OpenAI 始终以“确保通用人工智能造福全人类”为使命，不仅自身致力于通用人工智能技术研发，还愿意帮助其他机构研发通用人工智能应用。2023 年 2 月，OpenAI CEO 萨姆·阿尔特曼公布了对通用人工智能的规划。

通用人工智能能够为用户赋能，激发用户充分展现聪明才智和创造力。但是，通用人工智能存在滥用的风险。因此，通用人工智能应用的开发者需要在发挥通用人工智能优势的同时防范其可能带来的风险。

OpenAI 认为，想要发展通用人工智能，需要做好以下三个方面的准备工作。

首先，通用人工智能的发展过程是循序渐进的。人们有时间了解通用人工智能的应用情况，体验通用人工智能的优势并发现其缺点，基于此可以适时调整经济组织形式，并对通用人工智能进行有效监管。循序渐进的方式能够实现通用人工智能与社会的共同发展，人们也可以在这个过程中明确自身需求。

OpenAI 希望更多的人能够体验通用人工智能并积极推动它进一步发展。基于此，OpenAI 开放了 API (Application Programming Interface, 应用程序编程接口)，实现了模型开源。这将加速人们对通用人工智能的研究，让更多人积极贡献新想法。

其次，在模型构建方面，OpenAI 对模型的部署变得越来越谨慎，除了探索模型的通用性之外，还在努力创建可控程度更高的模型。OpenAI 认为，社会需要就如何使用人工智能达成共识。OpenAI 将积极进行外部投入实验，为通用人工智能进行复杂决策做好准备。

未来，OpenAI 将开发新的对齐技术，并测试当前的技术何时失效。从短期来看，OpenAI 将通过 AI 帮助人们评估复杂的模型和监控复杂的系统；从长期来看，OpenAI 将基于 AI 提出新的想法，以推动对齐技术不断迭代。

最后，OpenAI 希望聚焦如何治理通用人工智能、如何分配通用人工智能产生的收益、如何共享通用人工智能带来的发展机遇三个方面进行全球对话。



除了以上三个方面外，OpenAI 通过各种规定明确了其运作结构。例如，OpenAI 为其他组织提供助力的目的是提高组织的安全性，而不是在通用人工智能应用开发方面与其竞争。此外，OpenAI 对股东获得回报的上限进行了界定，以规范股东的行为。