

## 中金-AI 十年展望（五）：从 ChatGPT 到通用智能，新长征上的新变化

原创 赵丽萍 于钟海等 中金点睛 2023-02-03 07:51 发表于北京



2012 年深度学习元年以来，AI 进入学术和商业界发展新阶段，2017 年大模型路线以“通用智能”思路降低算法边际成本，逐渐成为学界与产业共识。**2022 年下半年，AIGC 及 ChatGPT 关注度大幅提升背后，我们认为其本质是弱人工智能到强人工智能的阶跃，海外及国内商业落地处在初期，但我们认为新的产业趋势值得关注。OpenAI 的技术进展及投资方向、海外 AI 独角兽业务进展、国内龙头 AI 公司的跨模态布局是重要风向标。**

**本文作者：赵丽萍，于钟海，魏鹤霏**

### 摘要

**技术层面，ChatGPT 和 AIGC 的持续升温依托于大模型的技术红利。**从 2012 年的深度学习元年以来，各界产生了海量的数字化需求，大模型技术路线是降低边际成本的核心，2017 年大模型（Transformer）路线逐渐成为学术界与国内外巨头的发展共识。ChatGPT 由 GPT-3.5 大模型加入基于人类反馈的强化学习训练而成，带来弱人工智能向通用智能的阶跃。

**商业层面，以 OpenAI 投资方向为风向标，落地仍在探索期，跨模态打开场景空间。**对于 AIGC 领域，AI 作画等跨模态应用是未来的发展趋势，全球范围内

AIGC 独角兽已初步具备 B 端为主的变现能力，海外落地节奏显著快于国内。ChatGPT 推出后，海外已有早期合作案例，例如为 BuzzFeed 提供个性化测试、为 Amazon 解决客户和工程师技术难题等。此外，以 OpenAI 前沿投资方向为锚，我们认为文本到图片/视频的跨模态生成、垂直领域 AI 写作、智能笔记、AI 语言学习平台或为潜在落地方向。

**展望未来，行业格局有望走向底层集中、垂类多点开花，MaaS 是商业模式演进的潜在方向。**我们认为，ChatGPT 为代表的模式背后，成本、算力、场景、数据等多维度需求铸就高门槛，大模型路线下，未来行业格局趋向集中，泛化通用大模型能力的厂商有望呈现多强格局；关键垂类仍可能有场景、数据优势，精细调优后的垂类模型仍有差异化竞争潜力。展望未来的商业模式，我们认为以海外 Hugging Face 为代表的 Model-as-a-Service 是潜在方向，具备大模型基础的龙头有望成为 AI 开源基础设施提供商、社区生态建设者。

## 风险

技术进展不及预期，行业竞争加剧，商业化落地节奏不及预期。

## 正文

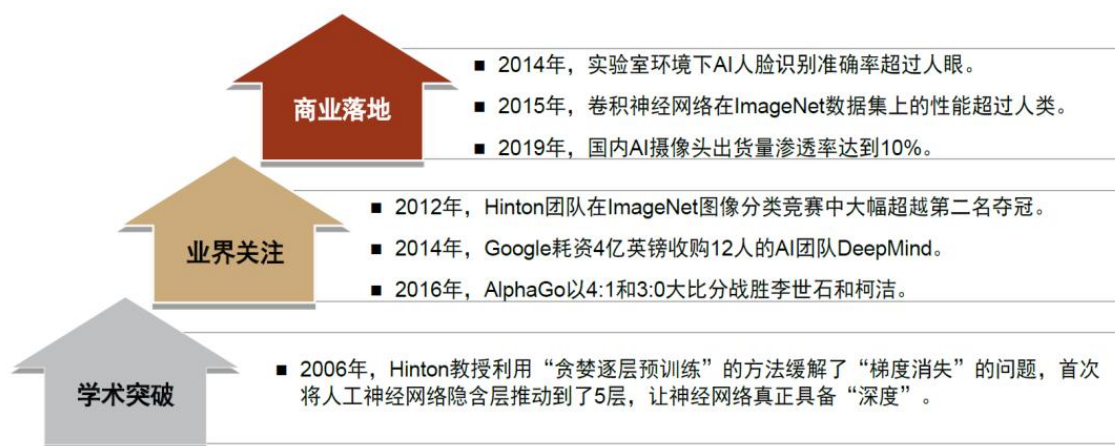
**技术篇：ChatGPT 和 AIGC 持续升温背后，本质上是大型模型的技术红利**

**ChatGPT 和 AIGC 持续升温的本质是背后的 AI 大模型进入新范式。**ChatGPT 是基于 GPT-3.5 的 InstructGPT，基于人类反馈的强化学习 (RLHF) 训练的语言类大模型。本文梳理了 ChatGPT 背后的技术演进与应用落地节奏，并对未来 AI 行业的格局加以展望。

**从 2012 年的深度学习元年开始，上一轮红利已经持续 10 年**

**2012 年是深度学习元年，以 Hinton 团队在 ImageNet 大赛中引入深度学习算法为重要节点。** Hinton 带领的团队在 ImageNet 挑战赛中，通过构建深度神经网络 AlexNet 成功将图片识别错误率降低了 10.8pcts，以大幅超越第二名的成绩夺冠，证明了深度学习的发展潜力，开始被业界和资本关注。

**图表 1：AI 从引起业界广泛关注到目前不足 10 年**



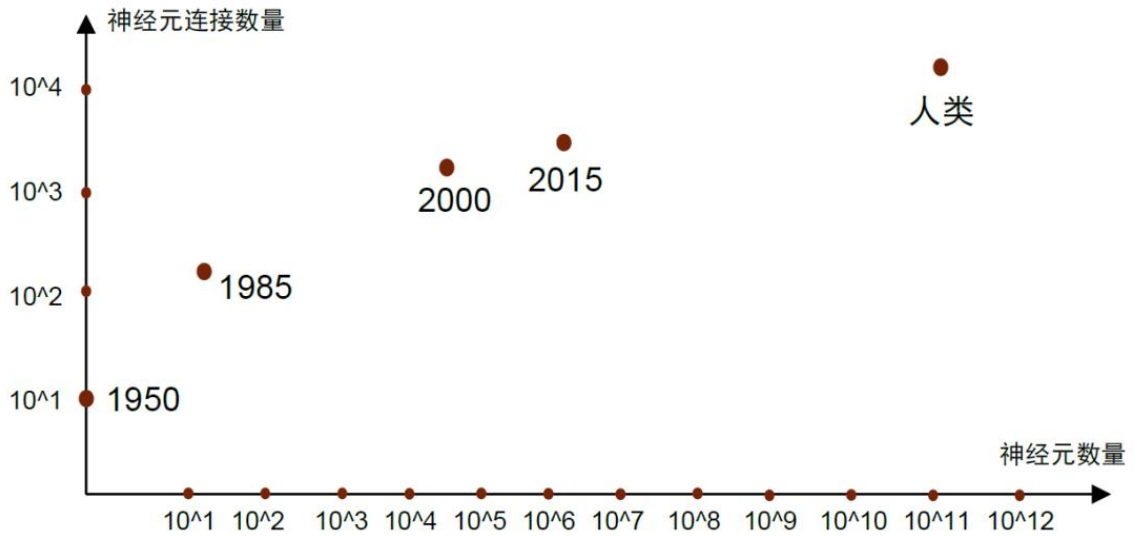
资料来源：量子位，Omdia，中金公司研究部

**早年业界一度认为小模型、小算力是方向。**此前的共识是通过精妙的算法和更高的模型精度，减少对算力的需求，但行业落地实践证明其短板明显：

▶ **精妙模型路线导致重复研发，无法让 AI 赋能千行百业。**在 AI 公司进入大量新场景时，对每个场景都重复研发大大增加了边际成本。

▶ **精妙模型路线无法解决长尾场景，完善 AI 商业化价值闭环。**以份额占人工智能行业近一半份额的智慧城市为例，其中的长尾场景如防火防灾、电梯事故、垃圾乱扔由于研发难度高，大部分 AI 公司难以低成本解决，严重影响到客户对整套解决方案买单的意愿。

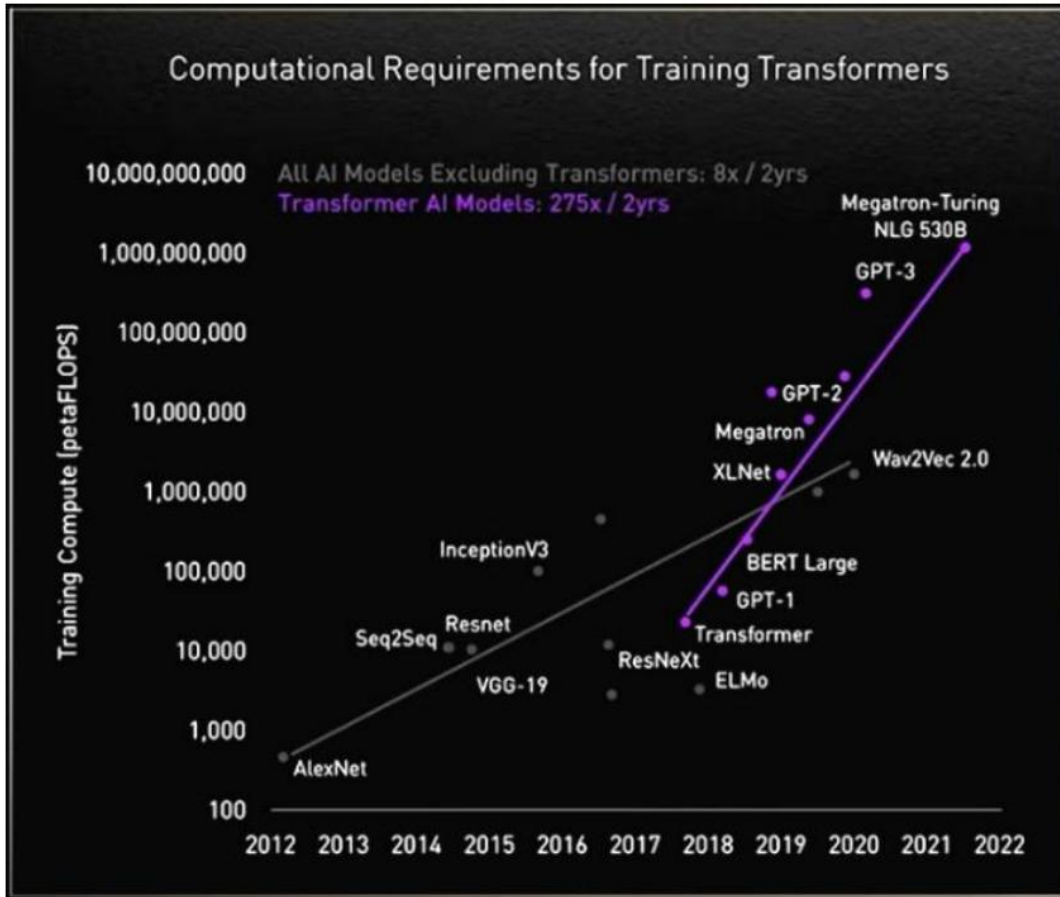
**图表 2：目前 AI 模型的神经元总量与人类差距较大**



资料来源：阿里云《中国企业 2020：人工智能应用实践与趋势》（2019 年），中金公司研究部

**技术演进视角，大模型是实现通用 AI 的重要方向。**虽然距离完全达到人类智能水平的 AI，还有很长一段路要走。但近几年在长尾场景等问题导致了对更通用的人工智能的刚需，在国内外巨头纷纷投入大量资源攻克通用人工智能难题的推动下，通用的语言模型、视觉模型甚至多模态模型也开始逐渐取得突破。2020 年 OpenAI 推出 1750 亿参数的 GPT-3 大模型是行业重要里程碑，为后续大模型迭代、加入人为干预，以及 AIGC、ChatGPT 的推出打下技术基础。

**图表 3：大规模预训练模型对于算力需求呈指数增长**



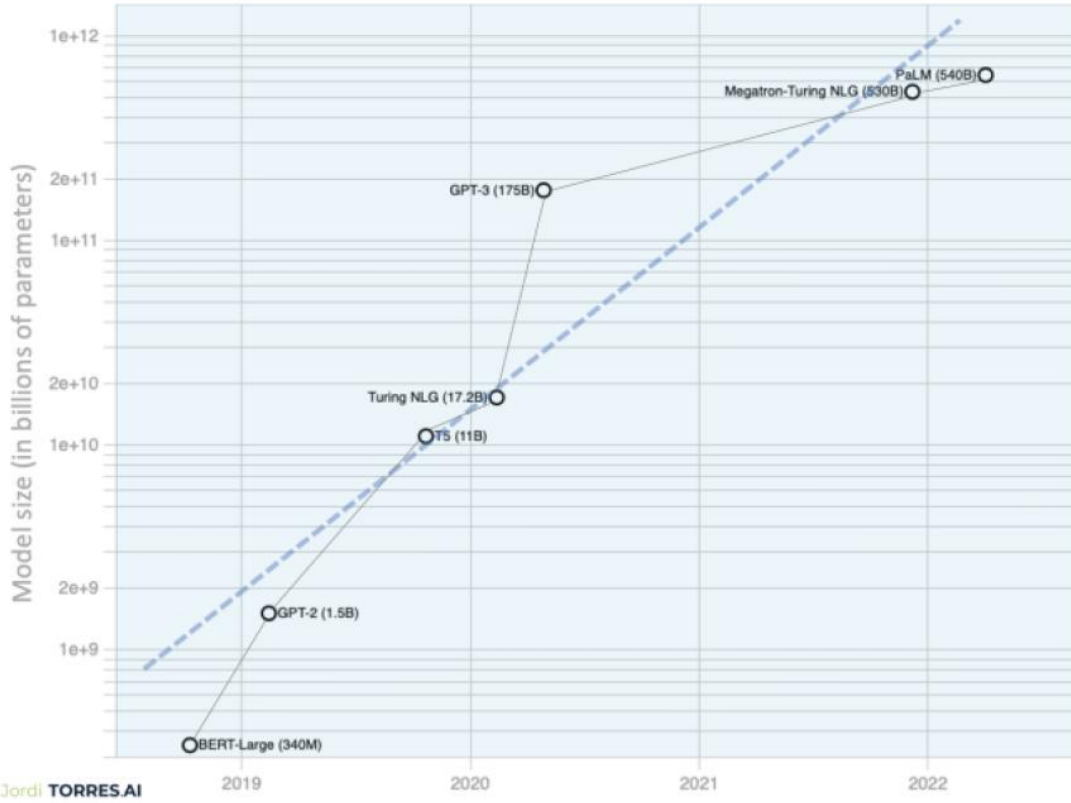
资料来源：英伟达官网，中金公司研究部

### 大模型将 AI 从感知提升到理解的维度

过去的深度学习在 CV（计算机视觉）领域大放光彩，但本质上还是更多的解决感知问题。CNN（卷积神经网络）已经能够较好的解决常规的人脸识别、物体识别问题，但在 NLP（自然语言处理）领域，对于物体、概念的抽象理解、语义的识别和理解，乃至推理和逻辑仍然是巨大的挑战。Transformer 算法（大模型路线）在 NLP 领域成果卓越，催化 NLP 的能力从简单的感知，向深层次的理

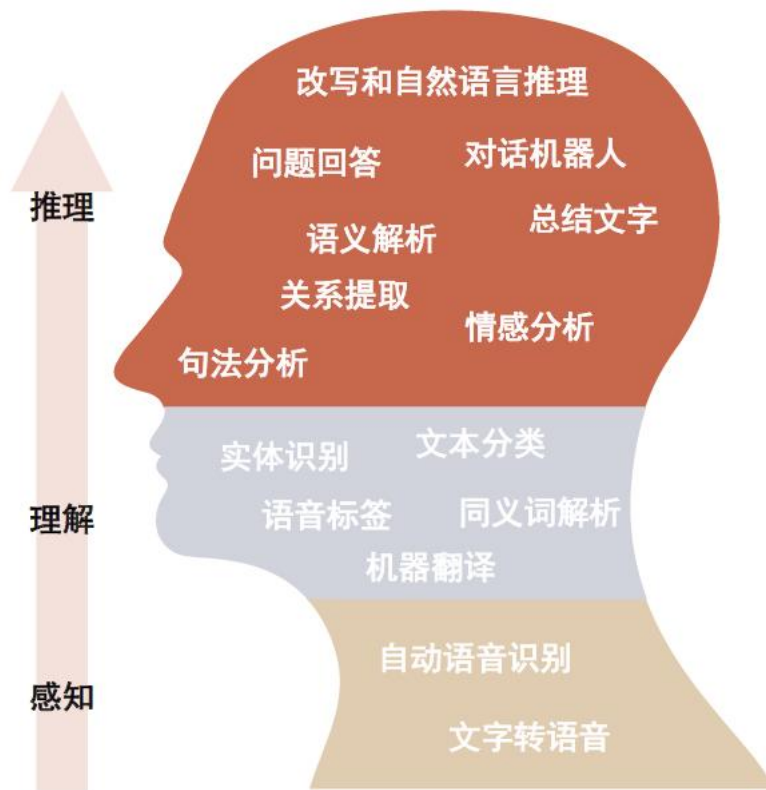
解乃至推理不断发展。

图表 4：模型的参数大小在不断提升，性能也在与时俱进



资料来源: Jordi Torres.AI, 中金公司研究部

图表 5: AI 的能力也在从感知向理解、推理、生成攀爬



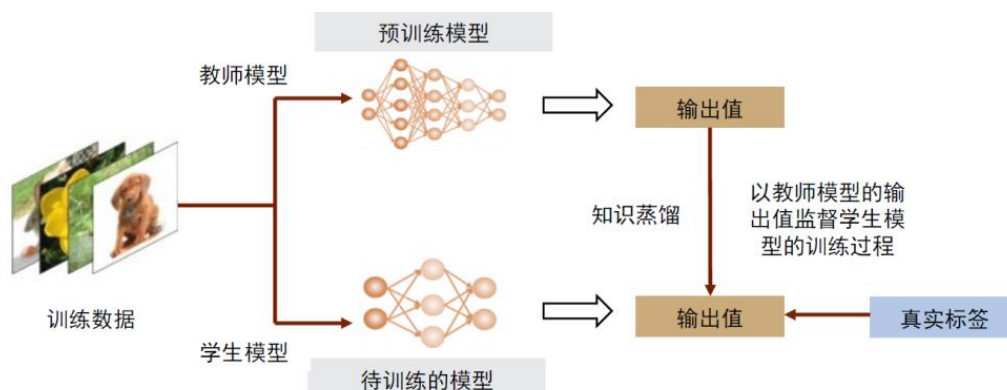
资料来源：中金公司研究部

## 如何简单的理解大模型（大规模预训练模型）

**基于具备一定通用性的大模型，通过少量的增量训练蒸馏出小模型，是解决长尾问题的关键技术架构。**从模型训练到部署，需要通过剪枝、量化、蒸馏等模型压缩技术实现更高的经济性及快速推理。以蒸馏为例，可以将结构复杂、参数规模庞大的大模型，压缩成结构简单、易于部署的小模型，相比于直接生产的小模型，大模型蒸馏出的小模型具有更强的泛化能力，边际成本大幅降低。

**大模型+小模型的方式能有效降低 AI 落地边际成本。**由于避免了“手工作坊”式的 AI 生产方式，不需要每做一个项目就派出大量专家花数月驻场收集数据、调试模型、训练模型，并且对长尾场景的解决为客户创造了更多价值。用足够多的数据和足够大的算法去训练一个足够大的通用模型，再通过量化、剪枝、知识蒸馏等模型压缩方法把大模型变小，高效的进行模型生产，并且由于算法足够多，能够覆盖各种长尾场景，大大降低了复制成本。因此我们认为，大模型是未来 AI 行业的必然趋势。

图表 6：蒸馏技术是类似于老师-学生传递知识的过程



资料来源：ICCV2019，华为云，中金公司研究部

**2020 年超大模型 GPT-3 发布，衍生出 DALL·E 与 CLIP 模型。**随着 GPT-3 在多个自然语言处理基准上展现超强性能，该模型延伸至图像领域，衍生出匹配文字和图像的 CLIP 模型，以及由文字生成图像的 DALL·E 模型。2022 年 4 月，OpenAI 推出 DALL·E 2，基于 CLIP 和 Diffusion 模型，分辨率提升了 4 倍，准确率更高，业务更广：除了生成图像，还能二次创作。

**图表 7：GPT-3 衍生出 DALL·E 与 CLIP 模型**

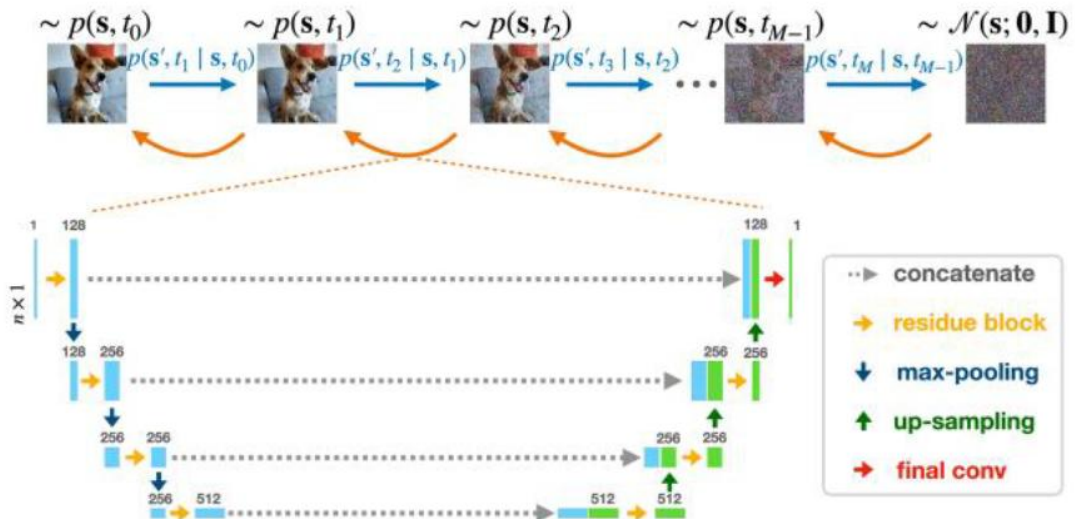


资料来源：OpenAI 官网，中金公司研究部

**Stable Diffusion 由 Stability AI 公司提出并于 2022 年 7 月开源，是 AIGC 热度提升的重要驱动力。**Stable Diffusion 的重要贡献在于使用空间降维解决内存和模型推理时长痛点，采用 Latent Diffusion Model (LDM) 方式，在模型复杂度和图像质量之间达到平衡。Stable Diffusion 不仅使用户仅在消费级显卡上就能够快速生成高分辨率、高清晰度图像，而且建立开源生态，大大降低用户的使用门槛。

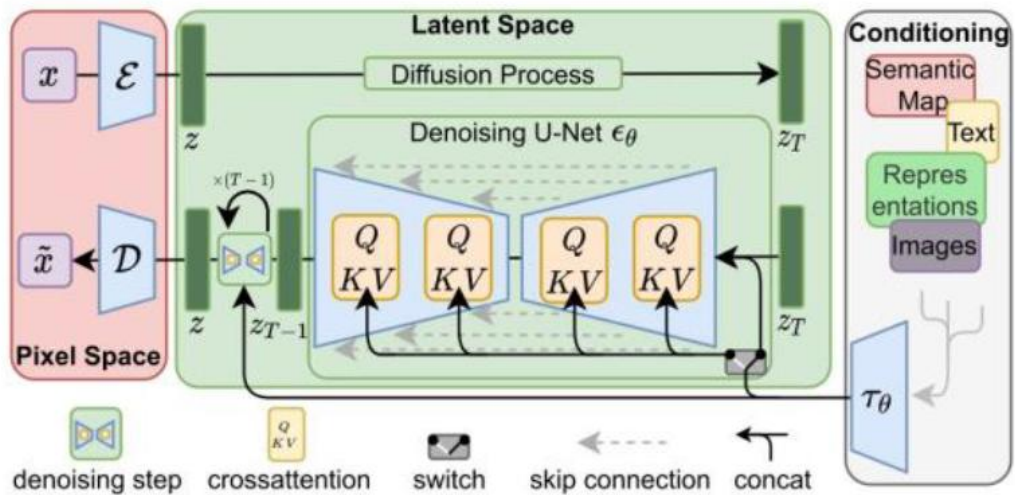
**图表 8：Diffusion 模型原理示意图**





资料来源：Stability.AI，中金公司研究部

图表 9：Stable Diffusion 模型原理示意图



资料来源：CVPR 2022 《High-Resolution Image Synthesis with Latent Diffusion Models》，中金公司研究部

**ChatGPT：基于 GPT-3.5 版本的 InstructGPT，开放接口，风靡全球**

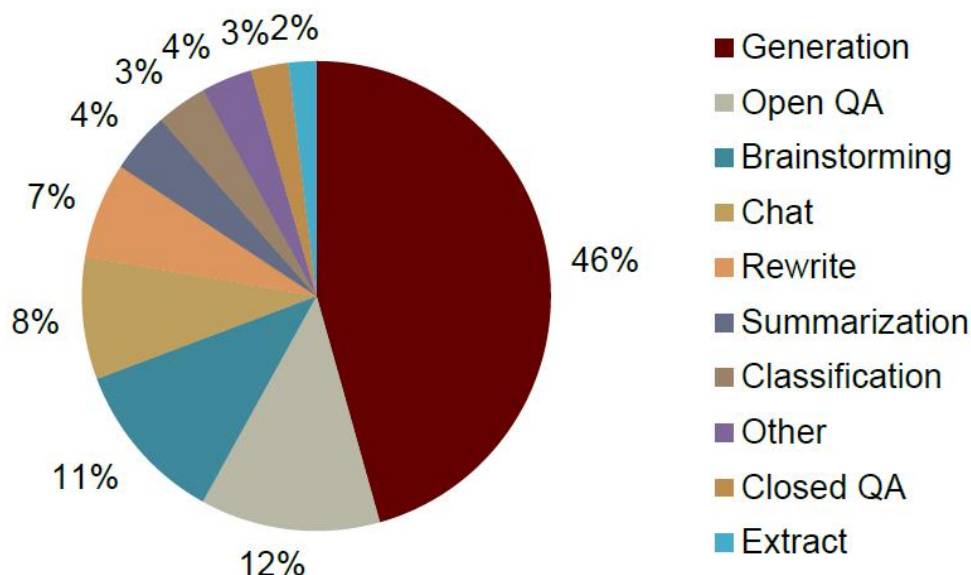
**2022 年 11 月 30 日，OpenAI 发布了语言大模型 ChatGPT，通过大规模预训练(GPT-3.5)和自然语言生成技术实现多轮对话问答。**该模型以对话的形式与用户进行自然交互，核心方法是基于人类反馈的强化学习(RLHF)，能够实现“回答后续问题、承认错误、质疑不正确的前提和拒绝不适当的请求”的能力。

**图表 10: ChatGPT 拥有智能问答、多轮对话的能力**



资料来源: ChatGPT, 中金公司研究部

**图表 11: InstructGPT 公开 API 后用户的使用功能和频率**



资料来源: 《Training language models to follow instructions with human feedback》 (OpenAI,2022) , 中金公司研究部

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：AI十年展望（五）：从ChatGPT到通用智能，新长征上的新变化.pdf

请登录 <https://shgis.com/post/1721.html> 下载完整文档。

手机端请扫码查看：

