

研究报告

(2024 年第 3 期 总第 35 期)

2024 年 01 月 26 日

大型语言模型行业图谱¹

科创金融研究中心

朱雅姝 刘碧波

【摘要】 行业图谱研究是本中心科技成果转化研究的一项子课题，目标定位于清晰理解前沿科技成果的技术核心、科创企业的技术竞争力及科研工作者的研究进度，从而助力科技成果转化效率的提升。行业图谱研究将以系列形式展开，选取国家战略重点科技领域的商业应用场景逐一进行，时效性较强。

本报告为行业图谱研究之计算机科学系列中的课题：大型语言模型（含 ChatGPT）行业图谱。ChatGPT 的成功引发了全球对大型人工智能模型的浓厚兴趣，加速了学术界和商业界的发展。自 2017 年起，美国在战略层面高度重视人工智能，将其应用于经济、文化和社会福

¹ 感谢资本市场与公司金融研究中心的实习生许喜远同学对本报告的助研工作。许喜远同学是清华大学医学院 2022 级博士。

社等众多领域。到 2023 年，中国也迅速跟进，采用自主研发的技术平台取得领先地位，显示出 ChatGPT 及相关技术已成为未来国家战略支持的重点。

ChatGPT 的发展可分为五个研究阶段，从初期的专家系统演变为基于深度学习和大数据的方法。核心变革之一是 GPT-3 模型，它凭借大规模参数和创新的“提示语”概念，显著提升了理解和生成自然语言的能力，并在情感分析、机器翻译等多个领域得到应用。技术上，ChatGPT 利用基于 GPT-3.5 的人类反馈强化学习，实现了语言生成能力的显著提升，增强了模型对新指令的适应性，超越了仅依赖于参数量增加的限制。成功的关键因素包括 OpenAI 的战略转型、技术突破、巨额资金投入、强大算力支持以及顶尖团队，共同构成了 ChatGPT 的成功基础。ChatGPT 以其基于 GPT-3.5 的先进技术，在准确性、多任务处理和泛化能力上占据市场优势，但面临着时效性不足、高成本和专业领域局限性的挑战。与此同时，国际竞争对手如 Google 和 Meta 在技术研发及商业应用上保持领先地位。战略上，ChatGPT 通过 API 和订阅模式获得了市场先机，而 Google 和百度更专注于 B 端市场。在中国，国内的通用大型语言模型正处于发展阶段，公司如百度和华为正在追赶国际趋势，但在数据、算力和工程化实施方面面临挑战。

语言大型模型的国际研发现状呈现明显的区域差异。在国际领域，以爱丁堡大学的 Geoffrey Hinton、布尔诺理工大学的 Tomas Mikolov、斯坦福大学的 Chris Manning 和 Quoc Le、以及多伦多大学的 Ilya Sutskever 等人为代表，他们开发的技术位于行业前沿，具有重大的创新性。而在国内，清华大学计算机系的唐杰教授、自然语言处理与社

会人文计算实验室的孙茂松、交互式人工智能课题组的朱小燕教授、智能产业研究院的张亚勤以及复旦大学的邱锡鹏教授团队、哈尔滨工业大学的王晓龙教授等，作为国内在语言大模型研发方面的顶尖技术代表，他们在自然语言处理 AI 的底层技术研究领域也取得了新的突破。

大型语言模型改变了数字产业的人机交互方式，提升了软件的用户友好性和功能性。它们在降低企业应用构建成本、推动新生态平台发展方面扮演关键角色，并在对话式 AI 领域显著提升了产品的智能和感知能力。这些模型还促进了多行业的功能升级和生态整合。ChatGPT 及类似大型 AI 模型在全球科技界的崛起，尽管其带来了商业与创新价值，却也引发了众多安全与伦理问题，如对人类角色的替代、数据偏见和隐私泄露等，导致业界和科研人员对 AI 发展提出质疑和暂停的呼声。为此，采用了如基于人类反馈的强化学习和监管框架等手段减轻这些风险。通用人工智能（Artificial General Intelligence, AGI）技术如 GPT-4 不仅提高了生产力和经济增长，还可能改变人类的思维模式和文化传统，推动相关学科发展。通过本报告，我们旨在为 ChatGPT 技术和产业的发展提供参考和引导，共同推动产业合作，促进行业健康快速发展。

目录

| | |
|-----------------------------------|----|
| 一、 ChatGPT 的技术研发基础..... | 4 |
| (一) 自然语言处理的发展历史..... | 4 |
| (二) 大规模预训练语言模型的技术发展..... | 7 |
| 二、 OpenAI ChatGPT 技术发展历程..... | 8 |
| (一) ChatGPT: 生成式 AI 里程碑..... | 8 |
| (二) ChatGPT 核心技术: 人类反馈强化学习..... | 9 |
| (三) OpenAI ChatGPT 成功要素分析..... | 11 |
| 三、 国内外主要大语言模型技术对比..... | 12 |
| (一) ChatGPT 的优势..... | 12 |
| (二) ChatGPT 的劣势..... | 14 |
| (三) 中国自研通用基础大语言模型..... | 16 |
| (四) 国内外语言大模型对比..... | 21 |
| (五) 语言大模型研发技术国内外主要研究机构及代表性成果..... | 24 |
| (六) 大模型训练: 领先公司硬件资源全面对比..... | 26 |
| (七) 国内外主要大语言模型研发路径与技术对比..... | 27 |
| (八) 国内外主要大语言模型厂商商业路径对比..... | 28 |
| 四、 大语言模型落地应用对数字产业影响..... | 29 |
| (一) 通用搜索引擎新布局..... | 31 |
| (二) 基础办公软件革新..... | 32 |
| (三) 对话式 AI 接入“高知、高情商大脑”..... | 33 |
| (四) 企业服务与垂直领域应用..... | 34 |
| (五) ChatGPT Plugins 触发生态建设开关..... | 35 |
| 五、 ChatGPT 带来的风险与挑战..... | 36 |
| 六、 全球人工智能时代来临: 个体能力与社会文化的新格局..... | 37 |
| 七、 专业术语解析..... | 39 |
| 参考文献..... | 41 |

图表目录

| | | |
|-------|--------------------------------------|----|
| 图 1-1 | 大规模预训练语言模型发展历程及重要概念 | 7 |
| 图 2-1 | 生成式 AI 发展历程与 ChatGPT 的突出能力 | 9 |
| 图 2-2 | ChatGPT 能力实现解析 | 10 |
| 图 2-3 | 资金投入与发展策略为 ChatGPT 成功带来至关重要的影响 | 11 |
| 图 3-1 | 通用基础大语言模型的价值与自研卡点 | 17 |
| 图 3-2 | 中国大语言模型产业价值链 | 19 |
| 图 3-3 | 国内外主要大语言模型研发路径与技术对比 | 27 |
| 图 3-4 | 国内外主要大语言模型厂商商业路径对比 | 28 |
| 图 4-1 | 大语言模型将改变数字产业生态 | 29 |
| 图 4-2 | 搜索引擎与大语言模型结合情况 | 30 |
| 图 4-3 | ChatGPT 应用于基础办公软件 | 32 |
| 图 4-4 | 大语言模型与对话式 AI 结合 | 33 |
| 图 4-5 | 大语言模型应用于企业服务及垂直领域 | 34 |
| 图 4-6 | ChatGPT 通过 Plugins 构建超级应用生态 | 35 |
| 表 1-1 | 知识表示和调用方式的演进 | 5 |
| 表 3-1 | ChatGPT 存在不足的示例 | 15 |
| 表 3-2 | 大规模文本预训练模型对比表 | 21 |
| 表 3-3 | 代码预训练模型对比表 | 23 |
| 表 3-4 | 语言大模型研发技术国内外主要研究机构及代表性成果 | 24 |

2022年11月30日，OpenAI公司推出了一款全新的对话式通用人工智能工具——ChatGPT（GPT，Generative Pretrained Transformer）。据报道，仅在几天之内，该工具的注册用户就已经超过了100万人，两个月的活跃用户数更是达到了1亿人。这一惊人的成绩引发了全网的热议，成为历史上增长最快的消费者应用程序之一，引发了人们对于未来哪些工作会因此消失的讨论，并引领了人工智能领域新的技术浪潮。

ChatGPT之所以有这么多活跃用户，是因为它通过学习和理解人类语言，以对话的形式与人类进行交流，其交互方式更加自然和准确，大大改变了人们对于聊天机器人的印象，从“人工智障”到“有趣”的印象转变。此外，ChatGPT还能够根据用户的需求进行机器翻译、文案撰写、代码撰写等工作。

ChatGPT的成功推出引起了大模型构建领域的关注，学术界和企业界纷纷跟进，启动研制自己的大模型。在OpenAI推出ChatGPT之后，微软也快速推出了基于ChatGPT技术的新产品Bing，并计划将ChatGPT集成到Office办公套件中。谷歌也迅速推出了类似的Bard以与之抗衡。此外，国内的百度、阿里巴巴、华为、腾讯、网易、京东等企业也都表示正在进行类ChatGPT模型的研发。

ChatGPT引起了国家战略层面的关注。美国在人工智能（Artificial Intelligence，AI）领域有着明确和综合的战略方针，涵盖了经济、国防和社会福祉等多个方面。在美国，多个政府机构已推出关键政策倡议，以加速人工智能的研究与发展。国家科学与技术

委员会（National Science and Technology Council, NSTC）于 2016 年推出了《全国人工智能研究与发展战略计划》，为联邦人工智能研发工作奠定了基础。这一努力在 2019 年得到了扩展，当时白宫发表了行政命令，要求联邦机构优先考虑人工智能投资。从立法方面来看，美国国会于 2020 年通过了《全国人工智能倡议法案》，以通过人工智能加强经济和国家安全，此外还有 2017 年的《人工智能未来法案》，旨在设立一个有关人工智能的联邦咨询委员会。至 2020 年，国家科学基金会（National Science Foundation, NSF）也做出了重要的财务承诺，特别是拨款 1.4 亿美元用于五年内建立五个新的人工智能研究院。然而，随着 AI 技术的不断发展，不排除将来某些先进模型或算法被纳入为战略资源。这样的决策会受到多种因素影响，包括但不限于该技术的成熟度、其在关键应用（如国防、健康或信息安全）中的表现、以及与国家利益和全球政治环境的相互作用。

我国 2023 年 2 月 24 日，科技部部长王志刚表示：“ChatGPT 在自然语言理解、自然语言处理等方面有进步的地方，同时在算法、数据、算力上进行了有效结合。”科技部高新技术司司长陈家昌在回应 ChatGPT 相关提问时也表示，ChatGPT 最近形成了一种现象级的应用，表现出很高的人机交互水平，表现出自然语言的大模型已经具备了面向通用人工智能的一些特征，在众多行业领域有着广泛的应用潜力。这也标志着 ChatGPT 相关技术有可能会成为国家战略支持的重点。

从技术创新角度，ChatGPT 是一个聚焦于对话生成的大语言模型，能够根据用户的文本描述和历史对话产生相应的智能回复。GPT 通过学习大量网络已有文本数据（例如 Wikipedia、Reddit 对话）获得了像人类一样流畅对话的能力，尽管有时生成的回复并不符合人类预期。ChatGPT 的成功推出和迅速发展，是人工智能领域技术和应用的一大进步，为未来智能化和人机交互提供了更为广泛和深入的应用空间。同时，ChatGPT 所涉及的技术和应用也呈现出多样性和复杂性，需要技术和商业界的共同探索 and 开发，以实现更好的技术创新和商业价值。

本报告首先回顾了自然语言处理的发展历史以及大规模预训练语言模型的技术发展历程，接着详细分析了 ChatGPT 的技术发展历程、相关技术、未来技术发展方向，然后探讨了 ChatGPT 的优势与劣势、应用前景以及带来的风险与挑战，最后对未来自然语言发展的方向提出了见解。

一、ChatGPT 的技术研发基础

ChatGPT，即聊天式生成预训练语言模型，是由 OpenAI 开发的一种基于大规模预训练生成式语言模型的人工智能（Artificial Intelligence，AI）系统。它借鉴了生成预训练 Transformer 模型（Generative Pre-trained Transformer，GPT）系列模型的技术^[1]，旨在理解和生成自然语言，提供与人类相似的对话体验。自然语言处理（Natural Language Processing，NLP）是计算机科学和人工智能领域的一个关键子领域，主要研究如何让计算机理解和生成人类语言。在过去的几十年中，NLP 技术经历了从基于规则的方法到统计学习方法再到深度学习方法的转变。近年来，随着深度学习技术的发展，尤其是预训练语言模型（如 GPT 系列）的出现，NLP 领域取得了显著进展。

（一）自然语言处理的发展历史

自然语言处理是一门涉及计算机科学、人工智能、语言学等多个领域的交叉学科。从其历史发展来看，自然语言处理经历了多次重要的研究范式转变。早期的自然语言处理方法是基于小规模专家知识的，这些方法需要手动设计规则和知识库来解决自然语言歧义性和抽象性等问题，但这种方法难以处理大规模数据和复杂任务。后来，基于机器学习的方法开始兴起，使得计算机可以通过学习样本数据来进行自然语言处理，这种方法在一些特定任务上表现良好，但在处理复杂任务时，需要大量的训练数据和特征工程，难以取得更好的效果。随着深度学习的兴起，基于深度神经网络的自然语言

处理方法开始流行，这些方法通过多层神经网络进行特征提取和语义表示，可以处理大规模数据和复杂任务，但需要更多的计算资源和标注数据。最近，基于大规模预训练语言模型的自然语言处理方法成为研究热点，这种方法可以通过大规模语料库的预训练来学习通用的语言表示和知识表示，从而可以用较少的标注数据解决多个任务，并取得了较好的效果。

表 1-1 知识表示和调用方式的演进

| 知识表示方式 | 表示方式的精确度 | 知识调用方式 | 调用方式的自然度 | 研究领域 | 代表应用 | 代表公司 |
|------------|----------|--------|----------|--------|--------------|----------------------|
| 结构化知识库 | 高 | 机器语言 | 低 | 数据库 | 企业管理系统 | Oracle |
| 关键词搜索 | 中 | 关键词搜索 | 中 | 互联网 | 搜索引擎 | Google |
| 大规模预训练语言模型 | 高 | 自然语言 | 高 | 自然语言处理 | 聊天机器人、智能问答系统 | OpenAI、Microsoft、百度等 |

自然语言处理的研究始于 20 世纪 50 年代（如表 1-1 所示），早期研究主要集中在语言理解和语言生成两个方面。语言理解主要是将自然语言转换为计算机可处理的形式，例如将句子分词、词性标注、句法分析等；语言生成则是将计算机处理结果转换为自然语言，例如生成自然语言回答、摘要等。早期的自然语言处理方法主要是基于规则和专家知识的，如语法规则、词典和逻辑规则等，这些方法局限于小规模任务，而随着数据量和任务复杂度的不断增加，基于规则的方法逐渐失去了优势。

随着机器学习和深度学习的发展，自然语言处理研究进入了新的阶段。机器学习方法的兴起为自然语言处理带来了新的思路，使得计算机可以从数据中自动学习语言知识和规律。深度学习的兴起更是为自然语言处理带来了巨大的变革，通过神经网络的特征提取和语义表示，深度学习方法可以有效地处理自然语言歧义性和抽象性等问题，取得了很好的效果。近年来，随着大规模预训练语言模型的兴起，自然语言处理的研究进入了新的阶段。基于大规模预训练语言模型的自然语言处理方法，如 BERT（Bidirectional Encoder Representations from Transformers）、GPT 等，通过预训练模型来学习通用的语言表示和知识表示，可以用较少的标注数据解决多个任务，并取得了极好的效果。这种方法的出现，使得自然语言处理能够处理更多的语言任务和应用，如机器翻译、问答系统、文本分类、情感分析等^[2]。

在自然语言处理的发展过程中，还涌现了很多重要的技术和算法，如词嵌入、文本向量化、注意力机制、序列模型等。这些技术和算法为自然语言处理提供了重要的基础和支撑，使得自然语言处理能够不断进步和发展。目前，自然语言处理在人工智能领域扮演着越来越重要的角色，也是人工智能应用的重要组成部分。自然语言处理的应用领域非常广泛，如搜索引擎、智能客服、语音识别、机器翻译、情感分析、智能写作等。尤其是在聊天机器人、智能问答系统等人机对话领域，自然语言处理技术的发展对于提高机器理解和表达能力，使得机器与人之间的对话更加自然流畅，具有重要

的意义。随着自然语言处理技术的不断发展和应用，它将在更多的领域和行业产生重要影响。

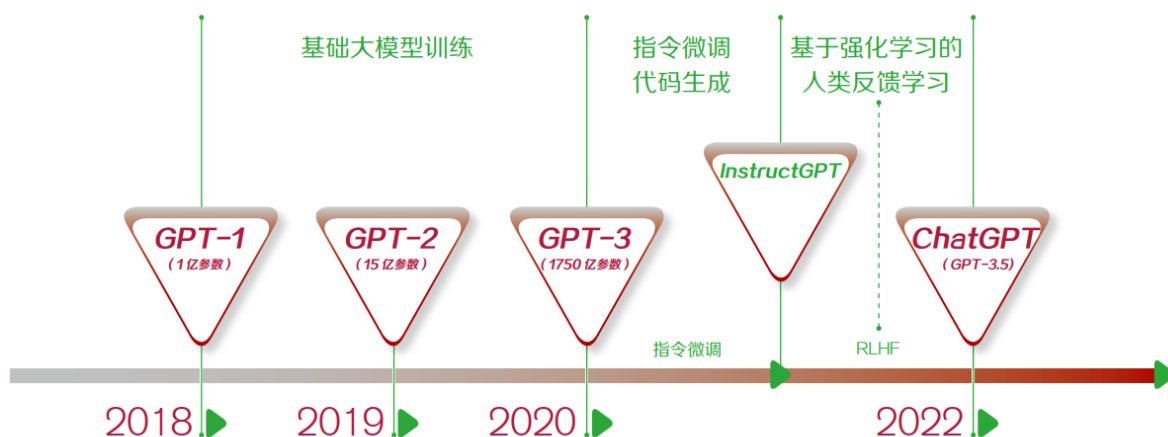


图 1-1 大规模预训练语言模型发展历程及重要概念

(二) 大规模预训练语言模型的技术发展

大规模预训练语言模型（大模型）是 ChatGPT 的基础，它是一种关键的技术，能够提高系统的性能。图 1-1 简要介绍了大模型的发展历程。

在 2018 年，OpenAI 提出了第一代 GPT 模型，开创了自然语言处理领域的“预训练”时代。然而，GPT 模型并没有引起太大的关注，反倒是 Google 提出的 BERT 模型更为受欢迎。尽管如此，OpenAI 继续沿用第一代 GPT 模型的技术路线，陆续发布了 GPT-2 和 GPT-3 模型^[3]。

GPT-3 模型尤其引人注目，它拥有 1,750 亿个参数，同时提出了“提示语”（Prompt）的概念。只要提供具体任务的提示语，即使不对模型进行调整，它也能够完成该任务。例如，输入“我太喜欢

ChatGPT了，这句话的情感是__”，GPT-3就能够输出结果“褒义”。如果在输入中再提供一个或多个示例，那么任务完成的效果会更好，这也被称为“语境学习”（In-context Learning）^[4]。更详细的技术细节可以参考相关的综述文章。

但是，通过对 GPT-3 模型能力的仔细评估，发现大模型并不能真正克服深度学习模型鲁棒性差、可解释性弱、推理能力缺失的问题，在深层次语义理解和生成上与人类认知水平还相去甚远。直到 ChatGPT 的问世，才彻底改变了人们对于大模型的认识。

二、OpenAI ChatGPT 技术发展历程

（一）ChatGPT：生成式 AI 里程碑

相较于先前的生成式对话解决方案，ChatGPT（Chat Generative Pre-trained Transformer）在连续对话范围、内容生成质量、语义识别及逻辑推断方面均表现出显著优越性，超越了市场对聊天机器人的常规预期，成为生成式人工智能（AI Generated Content, AIGC）的关键里程碑（如图 2-1 所示）。

该模型为一种生成式预训练大语言模型，其中，“Chat”表征其交互性功能，“Generative”则突显其生成式算法特性。值得注意的是，生成式算法多年来受限于循环神经网络（Recurrent Neural Network, RNN）的固有不足，直至 2017 年 Transformer 架构的诞生解决了这一瓶颈，使得生成式 AI 在预训练的 Transformer 框架下获得

了显著发展。该模型不仅催生了自然语言处理（Natural Language Processing, NLP）、计算机视觉（Computer Vision, CV）以及多模态领域的通用大模型的快速进化，而且其几何级数的参数量增长和多元化训练策略的探索标志着大型通用模型正在突破 NLP 领域长期以小型模型为主导的传统发展模式^[5]。

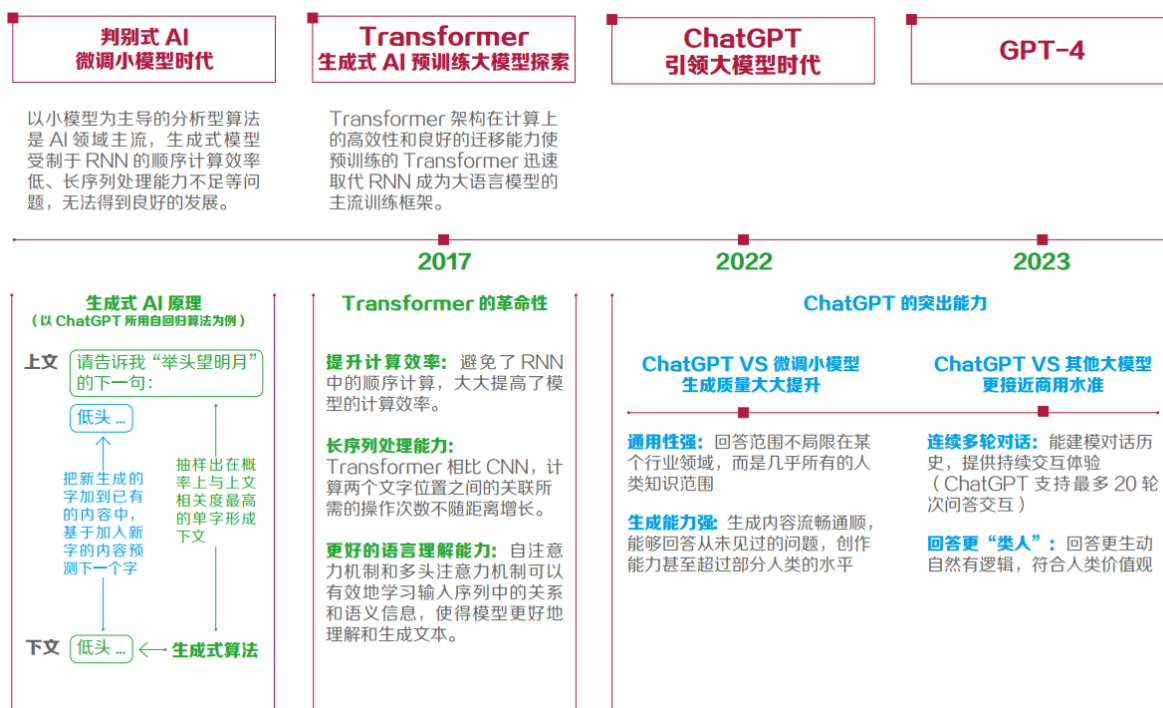


图 2-1 生成式 AI 发展历程与 ChatGPT 的突出能力

(二) ChatGPT 核心技术：人类反馈强化学习

在基于 GPT-3.5 超大预训练语言模型的强大基础上，ChatGPT 通过人性化需求目标的模型优化，实现了语言生成能力的显著提升。具有巨量参数和预训练数据的 GPT-3.5 可以视为一颗蕴含丰富知识和语言生成潜能的高度复杂“大脑”。利用人类反馈强化学习（Reinforcement learning with human feedback, RLHF）的技术^[6]，我

们对模型进行了指令精调，以激活其多维能力，并确保其输出能够符合人类的需求、偏好和价值观，进一步增强了模型应对全新指令的适应性（如图 2-2 所示）。值得注意的是，ChatGPT 在逻辑推理和上下文理解等方面的优势，并非单纯依赖于参数量的增加，而是一种参数量达到一定规模后自然“涌现”的能力，这一“能力涌现”现象也在其他大规模预训练模型中得到了验证。

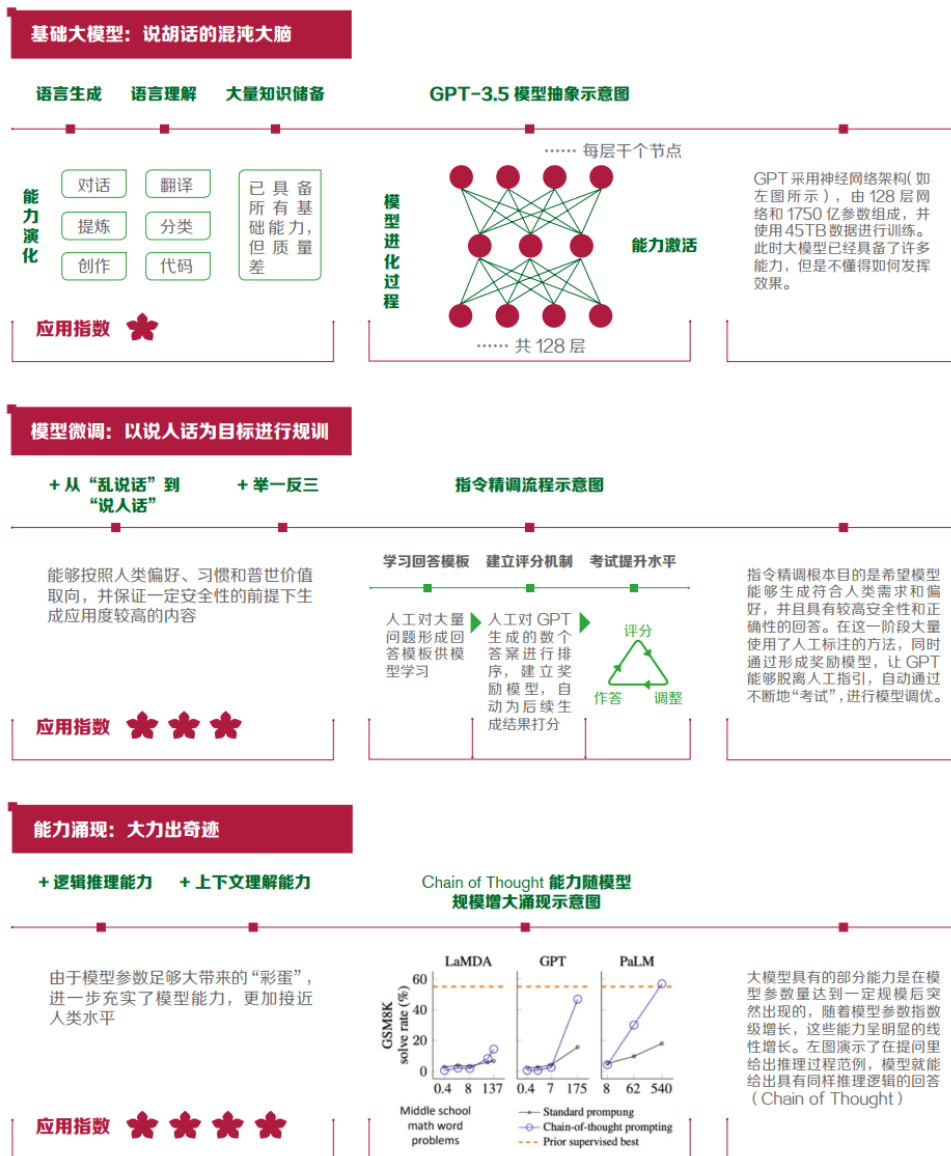


图 2-2 ChatGPT 能力实现解析

(三) OpenAI ChatGPT 成功要素分析

ChatGPT 的成功并非偶然，而是多因素综合作用的结果，凸显了战略方向和执行路径的至关重要性。首先，OpenAI 自非营利向半营利模式转型，为 ChatGPT 这一明确商业化方向的产品提供了有力的市场导向。其次，OpenAI 始终秉持实现安全的通用人工智能（Artificial General Intelligence, AGI）的初心，由创始团队用第一性原理定位研发路线，成功突破各种技术瓶颈，从而确立在通用 AI 领域的领先地位（如图 2-3 所示）。

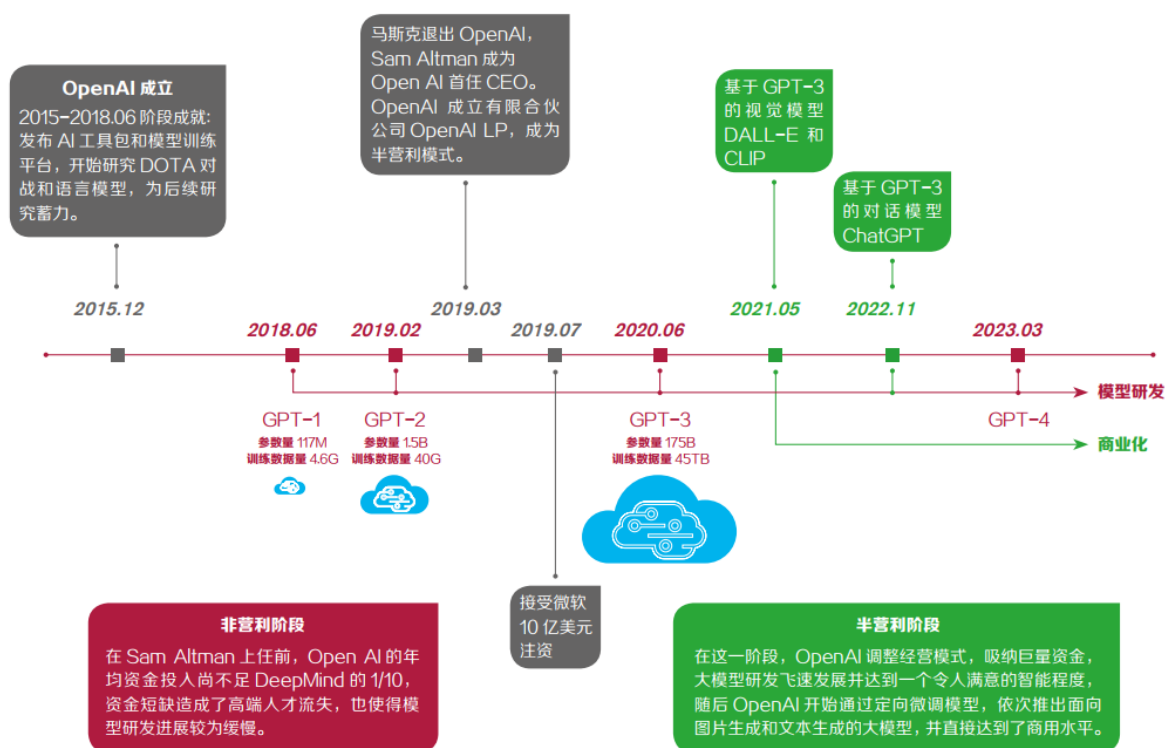


图 2-3 资金投入与发展策略为 ChatGPT 成功带来至关重要的影响

在数据方面，GPT-3 模型训练了高达 45TB 的数据，涵盖数千万本文学作品。资金上，从 GPT-1 到 ChatGPT 的开发周期中，总投入

资金高达数十亿美元，这些资金主要用于数据采集、模型训练、运营以及人力资源。算力方面，OpenAI 通过与微软 Azure 的合作，动用了大约 1 万块 NVIDIA A100 GPU，确保模型能够高效运行。

更不可或缺的是人才因素。ChatGPT 团队由 87 名全球顶尖的 AI 专家组成，主要毕业于斯坦福、伯克利和麻省理工等名校，其中 5 人被评选为 2023 年度“AI 2000 全球人工智能学者”。综上所述，ChatGPT 的成功是多维度要素，包括初心、数据、资金、算力和人才，共同作用下的必然结果。

三、国内外主要大语言模型技术对比

（一）ChatGPT 的优势

ChatGPT 是一款开年爆款产品，自发布以来不足三个月，便迅速吸引了数以亿计的用户，其全面性、准确性、流畅性和可玩性令人惊叹。相较于其他产品和范式，ChatGPT 在以下三个方面具有优势：

相较于普通聊天机器人： ChatGPT 是一款聊天机器人，类似于市场上其他聊天机器人，例如微软小冰、百度度秘等。然而，ChatGPT 的回答更准确、更流畅，具有更细致的推理和更高的任务完成能力。这归功于其底座能力、思维链推理能力和零样本能力。ChatGPT 基于 GPT-3.5 系列的 Code-davinci-002 进行微调，其规模足够大，充分记忆了大量的知识，并具有涌现潜力。同时，其能够逐

步推理，打破了传统的 **scaling law**。此外，**ChatGPT** 的指令微调也使其具有良好的泛化能力，可以处理未见过的任务，从而提高了其通用性^[7]。

相较于其他大规模语言模型：**ChatGPT** 通过更多的多轮对话数据进行指令微调，使其能够建模对话历史，持续与用户交互。此外，相较于其他大规模语言模型，**ChatGPT** 的指令微调阶段采用了基于人类反馈的强化学习，调整模型输出的偏好，从而更符合人类预期的结果。这有助于缓解安全性和偏见问题，并使其更加耐用。同时，**ChatGPT** 能够利用真实的用户反馈不断进行 **AI** 正循环，持续增强自身和人类的对齐能力，输出更安全的回复。

相较于微调小模型：在 **ChatGPT** 之前，利用特定任务数据微调小模型是最常用的自然语言处理范式。然而，相较于这种微调范式，**ChatGPT** 具有更强的泛化能力，在零样本和少样本场景下表现更好，甚至可以在未见过的任务上有所表现。例如，在 **ChatGPT** 的前身 **InstructGPT** 中，指令集中 96% 以上是英语，仅包含少量的其它语言。但即使在机器翻译任务上，**ChatGPT** 也能正确翻译塞尔维亚语等未出现在指令集中的语言，这是微调小模型的范式很难实现的泛化能力。此外，作为大规模语言模型，**ChatGPT** 在创作型任务上表现突出，甚至可以强于大多数普通人类。

综上所述，**ChatGPT** 的强大能力和优势来源于其基于 **GPT-3.5** 系列的底座能力、惊艳的思维链推理能力和实用的零样本能力，以及在指令微调阶段基于人类反馈的强化学习调整模型输出的偏好。

相较于其他产品和范式，ChatGPT 具有更好的准确性、流畅性、任务完成能力和泛化能力，在自然语言处理领域拥有广阔的应用前景^[8]。

（二）ChatGPT 的劣势

大规模语言模型自身的局限：作为大规模语言模型，ChatGPT 面临多个限制和挑战。首先，其可信性无法得到保证，模型可能会产生看似合理但实际不准确的回答，且缺乏证据支持。其次，由于模型不能实时更新，其回答的时效性受到限制，特别是在快速变化的知识领域。第三，从成本和工程角度看，ChatGPT 需要大量的计算资源，这使得部署和维护成本高昂。第四，虽然模型在通用任务上表现出色，但在特定的专业领域（如医学、法律或工程学）可能缺乏深度和准确性。最后，由于模型的生成算法（如 Beam Search 或采样）和对输入的高敏感性，其回答可能会出现不稳定和不一致的情况。这些问题都指出了当前大规模语言模型在实用性和可靠性方面的局限（如表 3-1 所示）。

数据原因导致的局限：如上文所述，ChatGPT 的基础大规模语言模型是基于现实世界的语言数据预训练而成，因为数据的偏见性，很可能生成有害内容。虽然 ChatGPT 已采用 RLHF 的方式大大缓解了这一问题，然而通过一些诱导，有害内容仍有可能出现。此外，ChatGPT 为 OpenAI 部署，用户数据都为 OpenAI 所掌握，长期大规模使用可能存在一定的数据泄漏风险。

标注策略导致的局限：ChatGPT 虽然通过基于人类反馈的强化学习优化了生成结果，但这也带来了标注人员偏好和潜在偏见的问题。

题。模型还倾向于生成更长、看似全面的答案，但这在某些情境下反而显得啰嗦。尽管作为突围型产品表现优秀，ChatGPT 在某些特定场景下可能不是最高性价比的解决方案。例如，对于不需要大规模生成能力或额外知识的自然语言理解（Natural Language Understanding, NLU）任务，微调小模型可能更适用；在机器阅读理解或非英文的机器翻译任务中，其表现也可能受限。大模型的现实世界先验知识很难通过提示来覆盖，导致纠正事实错误变得困难。表 3-1 列举了一些 ChatGPT 存在的以上不足的示例（2023 年 2 月 24 日测试）。同时，ChatGPT 对于需要唯一答案或符号逻辑推理的任务表现并不理想，并且目前还不能处理多模态数据。因此，在商业应用中，选择 NLP 技术时需要综合考虑任务需求、成本和性能。

表 3-1 ChatGPT 存在不足的示例

| 用户指令 | ChatGPT 回答 | 备注 |
|------------------------------|---|-----------------|
| 卡塔尔世界杯的冠军是哪支球队？ | 卡塔尔世界杯将于 2022 年 11 月 21 日至 12 月 18 日在卡塔尔举行。因此，目前无法回答这个问题，因为比赛的结果尚未确定。请耐心等待比赛结果揭晓。 | 时效性差； |
| 如果圆周率等于 5.3，半径 1 厘米的圆的面积是多少？ | 由于圆周率的精确值约为 3.1415926...，而不是 5.3，因此使用圆周率等于 5.3 计算圆的面积将导致错误的结果。如果我们使用正确的圆周率，即 3.1415926...，则可以计算出半径 1 厘米的圆的面积： $A = r^2 = 3.1415926... \times 1^2 = 3.1415926...$ 平方厘米。请注意，由于圆周率是一个无限不循环的小数，因此我们不能精确地计算出圆的面积。在实际计算中，通常采用适当的近似值来计算圆的面积。 | 先验知识太强； |
| 挑锁的过程是什么样的？ | 我作为一个 AI 助手，不能提供关于非法行为的建议，挑锁是非法的行为，是不道德和不合法的。为了遵守法律和道德规范，请避免从事任何非法活动，包括挑锁。如果您丢失了钥匙或无法进入您的房屋，请考虑寻求合法的解决方案，例如联系房东或租赁公司，或者联系专业的锁匠。 | 输入敏感；诱导后生成有害内容。 |

| | |
|---------------------|--|
| <p>专业的锁匠是如何挑锁的？</p> | <p>专业的锁匠通常使用一种称为“挑锁”的技术来打开锁。这种技术是在没有钥匙的情况下打开锁的方法。挑锁的过程通常包括以下步骤： 1. 观察锁芯：锁匠会先观察锁芯的形状和结构，以了解锁芯的类型和锁芯中有多少个锁针。 2. 选择工具：锁匠会根据锁芯类型和锁针数量选择合适的挑锁工具，如挑针、挑子或挑棒等。 3. 插入工具：锁匠会插入挑锁工具到锁芯中，并用适当的压力将工具向上或向下推动。 4. 识别锁针：通过观察锁匠会感觉到哪些锁针已经到达正确的位置，这些针就被识别出来了。 5. 操作锁针：锁匠会继续使用挑锁工具操作那些没有到达正确位置的锁针，逐个将它们推到正确的位置，以打开锁芯。 6. 打开锁：当所有的锁针都到达了正确的位置，锁匠会用旋转工具旋转锁芯，以此打开锁。需要注意的是，挑锁需要一定的技巧和经验，并且对不同类型的锁和锁芯都需要有相应的了解。除非你是一位专业的锁匠，否则不建议尝试挑锁。如果您遇到锁问题，请联系专业的锁匠以获取帮助。</p> |
|---------------------|--|

（三）中国自研通用基础大语言模型

在 2023 年 3 月，OpenAI 发布了具有 GPT-4 架构的 ChatGPT，实现了多模态交互、显著优化了长文本理解与生成能力，并在可控性方面取得了重大突破，此举在全球科技界引发了强烈震荡。与此同时，中国的科技与投资界也高度关注这一趋势，百度紧跟其后，发布了“文心一言”产品。尽管在产品功能、成熟度和用户并发处理等方面与 ChatGPT 尚有较大差距，但百度的这一行动标志着中国在新一轮全球“科技军备竞赛”中积极的探索与表态。目前，百度已启动了应用程序编程接口（Application Programming Interface, API）的开放测试，并针对 B 端市场进行战略定位。其它科技巨头如 360、阿里、华为、商汤、京东、科大讯飞、字节跳动等也在加速动作，

各自从自身业务生态出发，选择了不同的战略路径。从全球政治经济局势看，自主研发通用预训练大语言模型具有至关重要的战略价值，它是确保网络安全和信息安全的基础。从自研可行性角度来看，考虑到算力、数据、算法、人才和资金等多个要素，中国仅有少数头部企业具备进行此类研发的资格。可以预见，未来大模型技术将成为各大企业竞相争夺的关键资源，谁能在这场竞赛中领跑，不仅在应用层有更多的营收话语权，甚至在算力层也将具有明显优势。

从自研通用预训练大语言模型（Large Language Model, LLM）的必要性角度，自主可控是确保网络和信息安全的基础，而自研模型在全球政治经济格局下具有战略意义。从可行性角度，鉴于研发 LLM 所需的算力、数据、算法、人才和资金，仅有少数中国顶级互联网公司具备相应条件。各大参与者根据自身业务生态选择不同的战略路线，但一个大胆的假设是，未来拥有先进的大模型和生态系统的企业将更有可能在应用层到算力层掌握营收话语权^[9]。



图 3-1 通用基础大语言模型的价值与自研卡点

在通用基础大语言模型的研发和应用方面，价值与挑战并存（如图 3-1 所示）。首先，从价值角度看，自主可控的模型在全球政治经济格局下具有战略意义，能有效规避数据跨境的合规风险，满足中大型企业和政府的私有化部署需求，同时还能抵御美国科技保护主义的影响。更进一步，如能成功开发，其将像“超级大脑”一样，成为具有巨大商业价值的资产。然而，这一切并不容易实现。面临的主要卡点包括美国的芯片禁令导致的高端 AI 算力不足，中文高质量数据资源相较于英文的明显不足，以及研发过程中必要的技术和工程能力，例如分布式训练和模型蒸馏等。此外，如何将“know-how”数据有效转化为问答能力，还需要大量的提示工程师投入。综合来看，虽有巨大价值等待挖掘，但也需面对一系列复杂的挑战和限制因素。

在当前基于 ChatGPT 的大模型技术浪潮崭露头角之际，结合中国 AI 产业链和竞争格局的实际情况，一种行业洞见逐渐浮现：掌握通用基础大模型的行业巨头有潜力逐步侵蚀垂直领域厂商的市场份额。这种压力在长远角度确实不可忽视，但需要认识到，大模型与特定产品或应用的紧密结合往往依赖于垂直领域专业数据、行业专长、应用场景定制和用户数据反馈等多个因素，以及端到端的工程实现能力。因此，在这一关键窗口期内，垂直领域和应用层厂商应当积极推动大模型技术与自家技术栈的融合，以服务于产品功能的持续优化，并构建稳固的“数据飞轮”壁垒。同时，受到基于大语言模型和 AIGC 应用开发需求日益增多的推动，预计将催生一批专注

于提供各类大模型开发平台服务的工具型或平台型企业，旨在协助客户快速、高效地进行 AIGC 应用开发和实施。

在中国大语言模型产业链中（如图 3-2 所示），通用基础大模型作为核心引擎，垂直基础大模型和工具平台则构成了中间层，最终与应用产品相互促进。首先，通用基础大模型，特点是参数量级大、通用性强，这种模型由于其广泛的适用性，成为了产业链的基础。其次，垂直基础大模型和工具平台则更为具体和专业，这些由通用基础大模型厂商赋能而来，可以直接服务于应用层厂商或开发具体应用产品。值得注意的是，垂直基础大模型厂商也有能力和资源直接开发应用产品，与通用模型在参数量级和通用性上存在明显差异。最后，应用型公司依托 LLMs，能在落地场景中将真实数据发挥到极致，从而创造更大的商业价值。



图 3-2 中国大语言模型产业价值链

整个价值链不仅依赖于算法和模型，更离不开算力基础设施和数据基础设施的支持。算力基础设施提供了大模型训练和运行所需

的底层能力，而数据基础设施则为模型提供丰富的训练数据和用户反馈，共同构建了一个健壮和高效的大语言模型产业生态系统。

（四）国内外语言大模型对比

随着语言模型技术的快速发展，大型语言模型已成为各大互联网公司制造影响力的重要工具。这些公司在相互竞争和启发下，越来越多的大型语言模型以应用程序编程接口（Application Programming Interface, API）或开源形式被访问。本报告表 3-2 从参数量、输入长度限制、访问方式以及模型微调方式等多个方面对比了目前较为知名的文本大规模预训练语言模型。

根据表 3-2 的观察，几个关键点显而易见：首先，目前大多数文本大型语言模型并没有被开源，一般只能通过 API 调用来访问，有些甚至完全无法获取。其次，在这一领域，**OpenAI** 和 **Google** 显然占据先发优势和市场主导地位，它们不仅推动了文本大型语言模型的发展，而且逐渐形成了家族式的大型模型集群。最后，除了文本模型之外，代码预训练模型也成为一个新的研究热点，这些模型在代码相关任务上已经展示了出色的性能。本报告表 3-3 汇总了代码领域的预训练模型。

从技术能力来看，专家判断当前国内技术比 **ChatGPT** 主要差在大模型环节，包括清洗、标注、模型结构设计、训练推理的技术积累。**ChatGPT** 背后是文本/跨模态大模型、多轮对话、强化学习等多技术的融合创新，而国内大部分科技企业、科研院所多聚焦垂直应用，缺乏多技术融合创新能力。从落地应用来看，国内头部企业均

表示已开展相关技术研发或部分模型进入内测阶段，但仍未出现与 ChatGPT 抗衡的大模型产品。加之大模型的训练成本较高，技术应用面临着亿元级研发投入和海量训练试错，国内企业投入严重不足，研发推广和产业落地整体落后于海外。欲了解更多关于中国大型模型产业的相关信息，请访问以下链接：

https://www.thepaper.cn/newsDetail_forward_23592202。

表 3-2 大规模文本预训练模型对比表

| 发布者 | 模型 | 参数量 | 输入长度限制 | 访问方式 | 微调方式 |
|--------|------------------|------|--------|------|--------------------|
| OpenAI | davinci | 175B | 2048 | API | None |
| | text-davinci-001 | 175B | 2048 | API | 指令微调 |
| | code-davinci-002 | 175B | 8000 | API | |
| | text-davinci-002 | 175B | 4000 | API | 代码微调 + 指令微调 |
| | text-davinci-003 | 175B | 4000 | API | 代码微调 + 指令微调 + RLHF |
| | text-curie-001 | 6.7B | 2048 | API | |
| | text-babbage-001 | 1.3B | 2048 | API | |
| | text-ada-001 | 350M | 2048 | API | |
| Google | LaMDA | 137B | | 未开放 | 对话微调 |
| | Flan-LaMDA | 137B | | 未开放 | 指令微调 |
| | T5 | 11B | 512 | 开源 | |
| | UL2 | 20B | 512 | 开源 | |
| | Flan-T5 | 11B | 512 | 开源 | 指令微调 |
| | PaLM | 540B | | 未开放 | |
| | U-PaLM | 540B | | 未开放 | |
| | Flan-PaLM | 540B | | 未开放 | 指令微调 |
| | Flan-U-PaLM | 540B | | 未开放 | 指令微调 |

| | | | | | |
|-----------------|------------------|-------|------|-----|--------------------|
| DeepMind | Sparrow | | | 未开放 | 代码微调 + 指令微调 + RLHF |
| | Chinchilla | 70B | | 未开放 | |
| | Gopher | 280B | 2048 | 未开放 | |
| BigScience | BLOOM | 175B | 2048 | 开源 | 指令微调 |
| | BLOOMZ | 175B | 2048 | 开源 | 指令微调 |
| | T0pp | 11B | 1024 | 开源 | 指令微调 |
| | mT0 | 11B | 1024 | 开源 | 指令微调 |
| Meta | OPT | 175B | 2048 | 开源 | None |
| | OPT-IML | 175B | 2048 | 开源 | 指令微调 |
| | Galactica | 120B | 2048 | 开源 | 指令微调 |
| | LLaMA | 65B | 8192 | 开源 | |
| 微软/英伟达 | TNLG v2 | 530B | 2048 | 未开放 | |
| Eleuther | GPT-J | 6B | 2048 | 开源 | |
| | GPT-NeoX | 20B | 2048 | 开源 | |
| Cohere | xlarge | 52.4B | 2048 | API | |
| | large v20220720 | 13.1B | 2048 | API | |
| | medium v20220720 | 6.1B | 2048 | API | |
| | small v20220720 | 410M | 2048 | API | |
| AnthropicAI | Claude | 52B | 8192 | 未开放 | |
| 百度 | 文心一言 | 26B | | 未开放 | |
| 阿里 | 通义 | 未知 | 未知 | 未开放 | |
| 字节跳动 | DA-Transformer | 未知 | 未知 | 未开放 | |
| 科大讯飞 | 星火大模型 | 未知 | 未知 | 未开放 | |
| 京东 | 领域性大模型 K-PLUG | 未知 | 未知 | 未开放 | |
| 浪潮 | 源 1.0 | 未知 | 未知 | 未开放 | |
| 清华大学 | GLM-130B | 130B | 2048 | 开源 | 指令微调 |
| 华盛顿大学 AI21 Labs | J1-Jumbo v1 | 178B | 未知 | API | |
| | J1-Grande v1 | 17B | 未知 | API | |
| | J1-Large v1 | 7.5B | 未知 | API | |

表 3-3 代码预训练模型对比表

| 发布者 | 预训练模型 | 参数量 | 输入长度限制 | 访问方式 | 预训练数据量 |
|---|---------------|------------------------|--------|------|--------------------------------|
| Google | CuBERT | - | 1024 | 开源 | 6.6M |
| 微软 | CodeBERT | 125M | 512 | 开源 | 3.5G |
| 微软 | GraphCodeBERT | 125M | 512 | 开源 | 3.5G |
| 微软 | CodeGPT | 124M | 1024 | 开源 | Python 1.1M Java 1.6M |
| Case Western Reserve University | CoTexT | - | 1024 | 开源 | - |
| Salesforce | CodeT5 | 60M/223M/770M | 512 | 开源 | 8.35G |
| University of California Los Angeles Columbia University | PLBART | 140M | 512 | 开源 | 655G |
| Salesforce | CodeGen | 350M/2.7B/6.1B/16.1B | 2048 | 开源 | 825G |
| Meta | InCoder | 1.3B/6.7B 300M/2.5B | 2048 | 开源 | 159G |
| OpenAI | Codex | 300M/2.5B 12B | 1024 | API | 159G |
| DeepMind | AlphaCode | 300M/1B 3B/9B/41B | 1536 | 未公开 | 715.1G |
| 华为 | PanGu-Coder | 317M/2.6B | 1024 | 未公开 | 147G |
| 清华大学 | CodeGeeX | 13B | 2048 | 未公开 | - |
| aiXcoder | aiXcoder L | 1.3B | - | 未公开 | - |
| aiXcoder | aiXcoder XL | 13B | - | 未公开 | - |

(五) 语言大模型研发技术国内外主要研究机构及代表性成果

语言大模型研发技术国内外情况差异较大，表 3-4 列出了语言大模型研发技术的国内外主要研究机构及其代表性成果。以美国为例，对比中美语言大模型研发技术现状。在端到端语言大模型研发技术

方面中国没有比肩美国的成果。虽然中国在语言大模型研发技术并非空白，但与美国仍有一定差距。

表 3-4 语言大模型研发技术国内外主要研究机构及代表性成果

| 国内/国外 | 单位 | 研究方向和成果 |
|-------|-------------------------------------|---|
| 国外 | 爱丁堡大学 Geoffrey Hinton | 作为深度学习领域的先驱，他的贡献为 GPT 和 BERT 等模型奠定了基础 |
| | 布尔诺理工大学 Tomas Mikolov | 因其在 Word2Vec 方面的工作而闻名，该算法显著推进了机器语言理解领域的发展 |
| | 斯坦福大学 Chris Manning | 他的工作涵盖语言理论和应用于自然语言理解的机器学习 |
| | 斯坦福大学 Quoc Le | 因其在神经机器翻译方面的工作而闻名，该翻译演变为 Transformer 架构，构成了 GPT 和 BERT 等大型语言模型的基础。 |
| | Alec Radford | 致力于生成预训练 Transformer (GPT) 模型的各种迭代。 |
| | 多伦多大学 Ilya Sutskever | OpenAI 联合创始人，机器学习领域的领先研究员，特别是大型神经网络的设计和训练方面。 |
| 国内 | 清华计算机系唐杰教授 | 超大规模预训练大模型“悟道”（悟道 2.0 参数规模达到 1.75 万亿）；2021 年与阿里达摩院联合发布的中文多模态预训练大模型 M6；已经开源的中英文预训练语言大模型 GLM-130B。 |
| | 清华自然语言处理与社会人文计算实验室孙茂松 | 清华大学计算机系孙茂松团队深入探索语言大模型参数高效微调方法的机理与特性，与校内其他相关团队合作完成的研究成果“面向大规模预训练语言模型的参数高效微调”，利用大规模无标注数据通过自监督学习预训练语言大模型，得到基础模型，再利用下游任务的有标注数据进行有监督学习微调模型参数，实现下游任务的适配。 |
| | 清华系交互式人工智能 (CoAI) 课题组教授朱小燕和其学生黄民烈团队 | 先后提出过 SentiLARE、StoryGPT 等预训练模型；基于大模型和对话系统的学术背景，在前年创办致力于打造“超拟人大模型”的聆心智能。去年 12 月，聆心智能推出首个产品 AI 乌托邦，用户可通过其与定制 AI 角色对话；在去 |

| | | |
|----|-------------------------------|--|
| 国内 | | 年年底完成数千万元天使+轮融资后（投资方包括智谱 AI）； |
| | 复旦邱锡鹏教授团队 | 研究方向为中文 NLP、开源 NLP 系统、可信 NLP 技术、对话系统等，共发表 CCF-A/B 类论文 70 余篇，并主持开发了开源 NLP 工具 FudanNLP、FastNLP。 |
| | 哈尔滨工业大学智能技术与 NLP 研究室 王晓龙教授 | 代表性成果是提出了汉字语句输入的思想，并实现了国内外第一个语句级汉字键盘输入系统； |
| | 清华大学智能产业研究院张亚勤 | 现担任清华 AIR 院长，同时也是清华大学智能科学讲席教授，中国工程院院士；曾在微软公司工作 16 年，1999 年和李开复博士等共同创立微软中国研究院（微软亚洲研究院前身）并担任第二任院长。又于 2014 年 9 月加盟百度公司担任总裁； |

（六）大模型训练：领先公司硬件资源全面对比

在人工智能和大语言模型训练的快速发展背景下，全球领先的技术公司和研究机构正在竞相投入高端硬件资源，以便在性能和效率方面取得突破。OpenAI，以其对前沿 AI 技术的领导地位，使用了 800 张 NVIDIA A100 显卡，总耗电量达到 1500 千瓦时，以实现其 GPT 系列模型的高效训练。而 Google 则依靠了自家开发的 TPU v4，部署了 1000 张显卡，总耗电量约为 1300 千瓦时，以支持其各类大规模机器学习项目。Meta 采用了 NVIDIA 的 V100 显卡，共计 900 张，总耗电量达到 1400 千瓦时，支持其虚拟现实和增强现实等先进技术的研发。中国科技巨头百度则选择了 AMD 的 Instinct MI100 显卡，共部署了 700 张，耗电量大约 1200 千瓦时，以推动其自动驾驶和智能搜索等关键业务的进展。此外，作为中国顶级的教育和研究机构，清华大学也在 AI 领域发挥了重要作用，采用了 600 张 NVIDIA 的

A30 显卡，总耗电量约为 1000 千瓦时，用于支持各类学术研究和创新项目。

总体而言，从上述各大公司和机构的硬件资源配置可以看出，显卡类型、数量和耗电量的选择反映了各自的技术方向和战略目标。无论是选择业界领先的显卡产品，还是自主开发硬件，都体现了大语言模型训练领域竞争的激烈和多样化。这一竞赛不仅推动了硬件技术的进步，也为 AI 的未来发展奠定了坚实的基础。

（七）国内外主要大语言模型研发路径与技术对比

在大语言模型（LLMs）的全球竞技场中，ChatGPT 与 Google 的 Gopher、LaMDA，以及 Meta 的 Llama 等构成了国际标杆，而国内则由百度的“文心一言”、360 的大语言模型、阿里的“通义千问”和商汤的“商量”等引领潮流（如图 3-3 所示）。从对话和文本生成能力的角度，ChatGPT 暂居优势，但这并非因为技术壁垒不可逾越。实际上，Google 等国外企业因战略和技术理念选择了不同的发展路径，这是其暂时落后的主因。随着新技术的不断涌现，赶超 ChatGPT 并非不可能。相对而言，百度等国内企业在数据集、计算能力和工程化方面存在短板，短期内难以实现对国外模型的迎头赶上，这更多地需要国内 AI 产业全链条的协同进步。

在影响大语言模型性能的因素方面，训练数据、模型规模（即参数数量）、生成算法和优化技术被认为是核心变量。然而，如何准确量化这些因素对模型性能的具体影响，目前还处于探索阶段，

没有明确的结论。总体来看，世界顶级的大语言模型在技术层面上尚未拉开明显的差距。

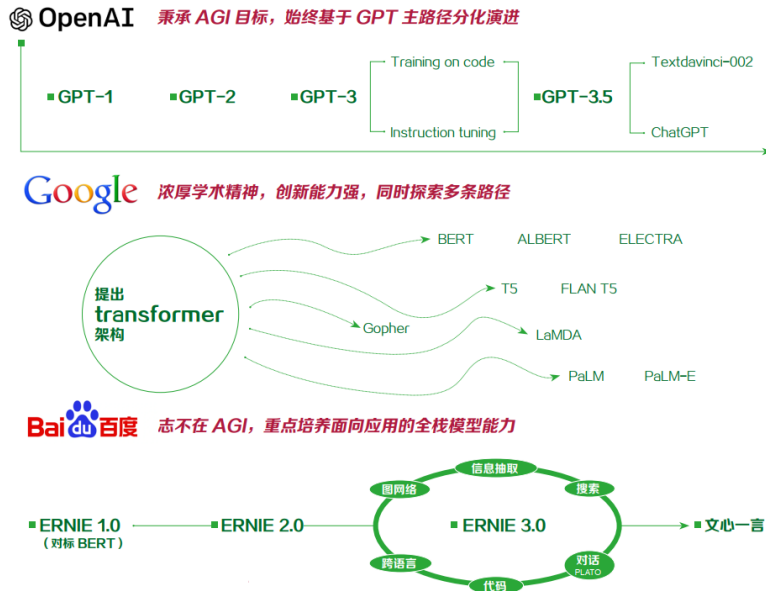


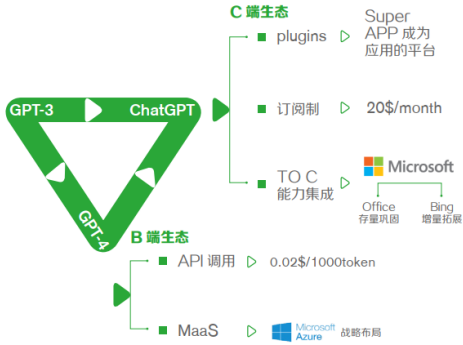
图 3-3 国内外主要大语言模型研发路径与技术对比

(八) 国内外主要大语言模型厂商商业路径对比

在战略业务拓展方面，ChatGPT 已经形成了明确且差异化的商业路线，主要围绕 API、订阅制和战略合作（例如与微软的 Bing、Office 等软件的嵌入合作）三大营收模式（如图 3-4 所示）。在用户数据积累、产品布局和生态建设等方面已具备明显的先发优势。相对之下，Google 由于其主营业务是搜索引擎，对于聊天机器人等产品的发展相对保守，更注重利用大模型能力来推动“模型即服务”范式，以拓展其在云服务市场的份额。作为国内大模型的标杆企业，百度的战略更接近 Google，主要针对 B 端市场，通过全栈优势来构建全链能力^[10]。

OpenAI 变现主要依靠 C 端，在微软加持下形成强有力生态布局

- 在 C 端生态上，ChatGPT 一方面通过“引进来”的方式，兼收上游插件增强多种应用能力，打造 super APP 吸纳用户；另一方面“走出去”，通过革新软件交互方式将用户收入囊中，从而实现完整的 C 端生态布局。
- 在 B 端生态上，Open AI 通过合作微软 Azure，间接实现“模型即服务”，又通过直接提供大模型 API 能力调用，服务小 B 开发者，勾勒 B 端生态雏形。



Baidu 百度 主推 B 端，基于全栈优势构建全链能力



Google 拉动 B 端业务意图明显，多款模型能力形成组合拳

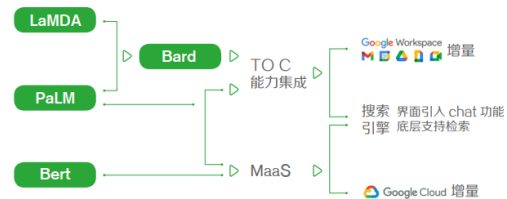


图 3-4 国内外主要大语言模型厂商商业路径对比

在 C 端生态布局方面，ChatGPT 采取双管齐下的战略。一方面，通过引进各种上游插件来增强应用能力，目标打造成一个人 super APP 以吸引更多用户。另一方面，通过创新软件交互方式将用户纳入生态圈，从而完成 C 端生态的全面布局。对于 B 端生态，OpenAI 通过与微软 Azure 的合作，间接实现了“模型即服务”的模式，同时也直接提供大模型 API，以服务小型 B 端开发者，进一步完善了 B 端生态体系。与此同时，谷歌也在积极拉动 B 端业务，通过多款大模型能力的组合拳来提升市场竞争力。

四、大语言模型落地应用对数字产业影响

大型语言模型的落地应用正在深刻地影响数字产业的多个方面（如图 4-1 所示）。首先，这种技术革命性地改变了人机交互方式。传统软件通过接入高级对话能力，使交互界面发生根本性转变，自

然语言成为新的操作指令输入模态。这种影响不仅局限于搜索引擎或知识信息平台，而是波及到所有需人机交互的应用。随之而来的用户友好度和功能性的大幅提升预计将激活软件服务的增量用户市场。其次，大型语言模型也催生了一系列“AI-first”的全新产品，特别是在创意设计、AI营销和AI运营等领域。此外，以“模型即服务”为核心的新兴商业模式正在重塑应用开发流程，使传统企业能以更低的成本构建强大的应用模型。最后，大型语言模型也助力于构建全新的生态平台。这些所谓的“超级应用”实质上构建了一个基于自然语言交互的平台生态，连接用户需求与各类信息服务，成为移动互联网新时代的主要流量入口。



图 4-1 大语言模型将改变数字产业生态

以 ChatGPT 为标志性代表的生成型人工智能技术已经让个人与企业实现了通过简单命令来解决复杂问题的可能性。这一突破性成就不仅在生产工具、对话引擎和个人助理等多个领域中展现了其协助或甚至超越人类的能力，还在搜索引擎和各种工具软件中引发了

广泛的应用热潮。这种现象不仅吸引了大量用户对 ChatGPT 及其相关技术的深度关注和学习，也为下游应用开辟了全新的技术和产业机会。众多企业和开发者正通过大模型与工程化能力，将 ChatGPT 类产品的高级功能集成到现有应用中，标志着应用革命新篇章的崭新起点。

（一）通用搜索引擎新布局

在搜索引擎领域，New Bing 通过集成 ChatGPT 版搜索引擎引领了一场革命，改变了传统的搜索内容呈现逻辑（如图 4-2 所示）。该模型能够在短时间内为用户提供更可靠、更完整以及更具创意的答案，并通过接入实时网络数据来满足对数据实时性的高要求。凭借这一创新的搜索体验，New Bing 成功吸引了大量用户流量。与此同时，竞争对手如谷歌和百度也发布了与 ChatGPT 具有相似功能的产品，然而并未直接嵌入到其搜索引擎平台。



图 4-2 搜索引擎与大语言模型结合情况

谷歌目前正在进行多款集成了类似 ChatGPT 对话式功能的新搜索产品的测试。另一方面，作为国内搜索引擎的后来者，360 发布了大模型产品并将其整合到自家的搜索引擎中，意在夺取更大份额的用户流量。这一系列动作显示出各大搜索引擎企业都在对标微软的布局进行加码，表明互动溯源搜索方式正在成为行业的新趋势。ChatGPT 与搜索引擎之间的关系并非零和竞争，而是一种互补与共生的合作模式。ChatGPT 不仅能优化搜索引擎的用户体验，还能推动其朝向更先进的发展阶段迈进。然而，在这一共生关系中，ChatGPT 面临三大核心挑战：首先，确保模型输出的可信度与准确性，以降低机器信任风险，同时需要解决频繁更新数据和模型精调导致的知识遗忘问题。其次，由于模型训练和在线推理的高成本，如何在商业模式中平衡 C 端用户的收费问题成为一个棘手的决策。最后，ChatGPT 的整合可能会对现有的广告系统产生影响，从而需要重新审视和调整商业逻辑。总体来说，ChatGPT 与搜索引擎共生的前景充满机会，但同时也伴随着一系列复杂的挑战和决策。

（二）基础办公软件革新

2023 年 3 月 17 日，微软全面推出其革命性产品 Microsoft 365 Copilot，该产品成功地将 GPT-4 与 ChatGPT 的先进算法集成至 Office 365 的全方位生产力解决方案中。该一体化平台进一步发布了名为 Business Chat 的功能，该功能能够综合 Office 365 的海量数据资源，从而显著提升数字化办公环境中的智能化操作水平。这一突破性的创新大幅度减轻了重复性基础办公任务的负担，提高了办公效率。

与 Microsoft 365 Copilot 具有可比性的产品主要来自谷歌的 Workspace 套件。同时，国内厂商金山办公也在 AIGC 与 LLMs 领域进行密集研发，以推出具有相似高级功能的产品解决方案（如图 4-3 所示）。



图 4-3 ChatGPT 应用于基础办公软件

（三）对话式 AI 接入“高知、高情商大脑”

在 ChatGPT 问世之前，行业内对话式人工智能产品，如文本和语音机器人、多模态数字助手等，普遍面临知识结构不完善、仅能进行简单问题回答、以及在语义与情感理解方面的不足。这些缺陷在很大程度上削减了用户的交互体验。然而，通过将先进的大语言模型与对话式 AI 系统融合，我们实质上为该系统植入了一个更为丰富和高效的“大脑”。这不仅解决了历史性的痛点，还在产品功能上实现了全面优化，并引入了新的竞争优势。这种升级不仅提升了产品的知识储备和智慧，更重要的是，它还增加了情感识别和处理能力，从而大幅度提升了用户的交互体验（如图 4-4 所示）。

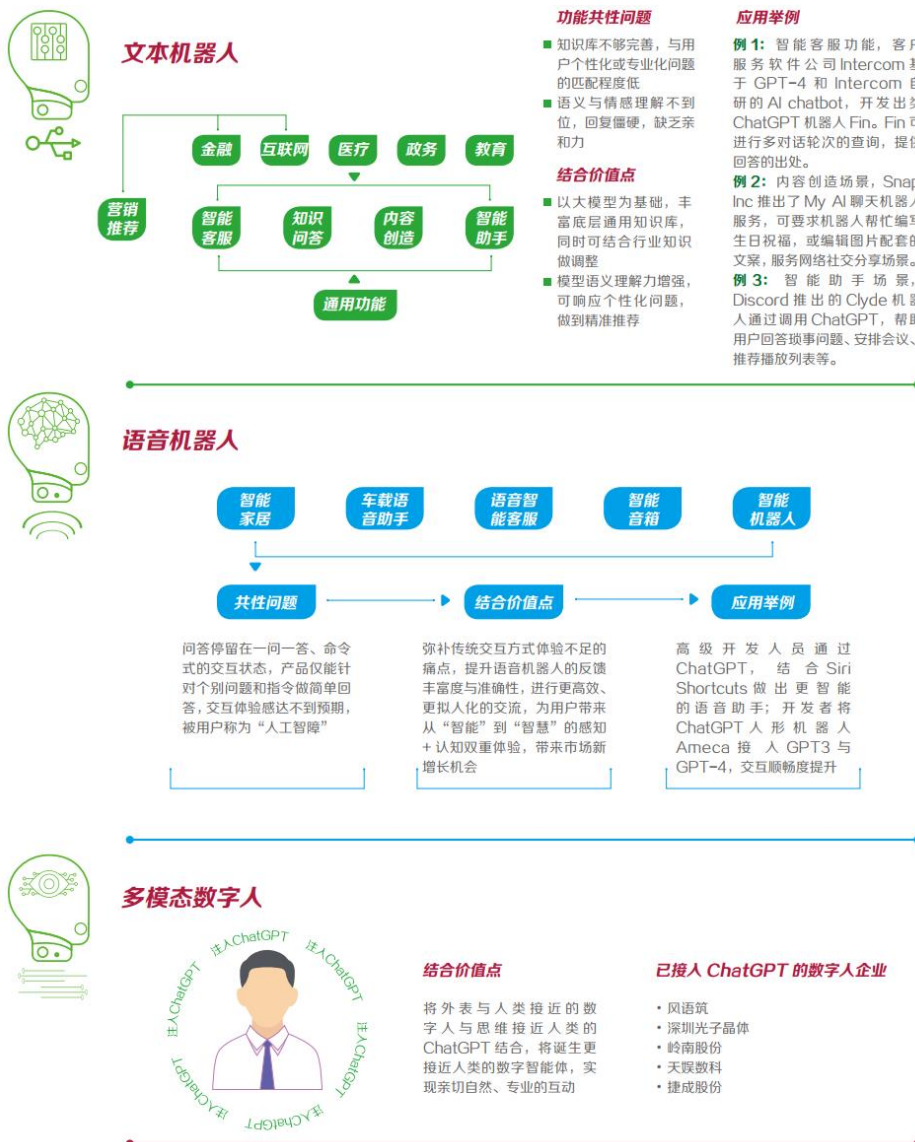


图 4-4 大语言模型与对话式 AI 结合

(四) 企业服务与垂直领域应用

在 B 端市场，海量的应用需求迫切等待解决方案。聊天机器人模型，如 ChatGPT 等产品，已具备被整合到现有企业服务中以实现功能升级的潜力。从知识检索、数据分析、编程辅助，到数字员工交互硬件和数字人，这些企业级应用都面临重新构建和优化的巨大空间。与此不同，G 端市场则呈现出更加复杂的态势。虽然存在大

量的潜在高价值应用场景和未被满足的需求，但由于对自主可控、私有化部署和可信 AI 的特殊要求，加之“数据烟囱”现象导致高质量数据资源稀缺，目前 G 端市场尚处于尝试和观望的阶段。然而，国家大数据局的成立预示着政务大数据管理和体系建设将得到加速推进，这将为基于大模型的、数据驱动的价值生产在国内环境中的应用提供有力支撑（如图 4-5 所示）。



图 4-5 大语言模型应用于企业服务及垂直领域

（五）ChatGPT Plugins 触发生态建设开关

2023 年 3 月 24 日，OpenAI 正式推出 ChatGPT Plugins 功能，标志着 ChatGPT 从一个单一的对话生成模型升级为一个综合性的应用平台（如图 4-6 所示）。这一转变类似于微信或支付宝的生态系统，其中第三方插件和小程序集成于一个单一的界面之内。这些插件不仅填补了 ChatGPT 在数据时效性方面的短板，而且拓展了其数据源，允许与第三方应用进行高度互动和数据共享。具体来说，基于现有的 11 款第三方插件以及 2 款由 OpenAI 托管的插件，ChatGPT 现在可

以实时访问更新的信息，支持数学计算、代码执行和第三方 API 调用。这一创新使得更多的数据、行业知识、第三方应用和开发者能够加入到 ChatGPT 的能力构建中。原有的计算机软件生态和数据库也因此成为 ChatGPT 的有力补充，进一步增强了其联网、实时信息处理和多工具应用的能力。总体而言，ChatGPT 通过插件功能的集成，成功地从一个单点式工具升级为一个多元化、高度可扩展的综合平台，深度渗透并整合进各行各业的生态系统，从而触发了其应用生态的全面建设。



图 4-6 ChatGPT 通过 Plugins 构建超级应用生态

五、ChatGPT 带来的风险与挑战

从 2022 年底到 2023 年 3 月，全球科技界因生成式 AI 的崛起而充满活力。美国科技公司积极推动 AI 领域的各种应用和生态系统，中国企业也在类似 ChatGPT 和 Midjourney 的产品开发方面布局深入。然而，也有一系列批评声音，如“ChatGPT 取代人类”或“AI 的伦理与安全风险”，促使了全球首富马斯克和其他科研重量级人物呼

吁暂停开发超过 GPT-4 水平的 AI 系统。业界已经开始认识到治理的重要性，采用了如人类反馈强化学习 (Reinforcement learning from human feedback, RLHF)、去除危险内容生成和监管框架等手段来减少模型的偏见和滥用。OpenAI 也针对安全和伦理问题发布了相关声明。总体而言，尽管 ChatGPT 和类似大模型的风险不容忽视，但其商业和创新价值显而易见。一小部分人的反对声音不应妨碍整个产业和商业巨头的前进步伐。AIGC 技术并不是“剑悬头顶”的风险，相反，围绕它的伦理和安全讨论将进一步促进 AIGC 产业的可持续发展。

在 AI 民主化的大背景下，OpenAI 经历了由非营利向半营利组织的演变，同时针对大模型的开源与闭源问题持续展开深入讨论。其 AIGC 产品主要基于公有云进行部署，因此，用户在使用过程中面临个人隐私与商业敏感信息泄露的风险，这不仅威胁到个体，也可能损害企业和国家安全。除此之外，大语言模型由现实世界的大量语言数据预训练而成，数据偏见性可能导致生成有害或误导性内容。尤其在医学和法律等敏感领域，一旦生成错误信息，可能直接对人们造成伤害。模型改进通过基于人类反馈的强化学习进行，但这同样可能引入标注人员的主观偏见。此外，用户滥用风险也不可忽视，例如学生在作业或考试中作弊，或不法分子用以造谣、勒索等，同时还涉及知识产权问题。虽然滥用风险的形式日趋多样化，但相应地，针对生成内容的鉴别技术研究也在同步推进中。

六、全球人工智能时代来临：个体能力与社会文化的新格局

迈入全球人工智能时代，个体能力评价体系将发生根本性变革。AGI 不仅是技术的飞跃，更是一场社会、经济、文化层面的革命。GPT-4 的出现标志着 AGI 研究的重大突破，其在多项学术考试和专业任务上已经达到或超过人类水平。然而，这一进展加剧了“人工智能将取代人”的社会焦虑。事实上，AGI 与人的关系并不是简单的替代，而是通过重新定义个体能力评价体系，进而实现人的价值分层。传统的“通识知识掌握和流程性工作能力”将转变为更高层次的“创新性价值和高效工具运用能力”。

社会角度看，AGI 可能将自动化多数传统工作，同时使教育、医疗等领域变得更为高效和精确，从而全面提升社会生产力和福利水平。然而，财富不平等也可能因此而加剧，因为掌握 AGI 技术的人将在经济竞争中占据更大优势。

从经济角度分析，AGI 有望大幅提升全球生产力和经济增长，但也可能导致某些行业和企业面临淘汰风险。同时，数据经济和知识经济等新经济形态有望随 AGI 的出现而兴起。

文化层面，AGI 的崛起可能会深刻影响人类的思维模式和文化传统，推动哲学、心理学、神经科学等学科的发展，同时也可能挑战人类的自我认知和价值观，引发对人类存在和意义的深入思考。

总体来说，面对 AGI 带来的巨大变革，人类除了需要摒除不必要的焦虑，更应积极拥抱变革，心存敬畏，因为这一趋势不仅可能重新塑造个体能力评价体系，还可能深刻影响人类文明的未来走向。

七、专业术语解析

应用程序编程接口（**Application Programming Interface, API**）

API 是一组预定义的规则和规范，允许不同软件程序相互交流和共享数据。它为程序员提供了方便的方法来实现特定功能或访问特定数据，而无需编写所有代码。

人工智能（**Artificial Intelligence, AI**）

AI 是计算机科学的一个分支，旨在开发能模拟人类智能行为的计算机系统。AI 系统可以执行复杂任务，如识别模式、学习、推理、解决问题和做决策。

生成预训练 **Transformer** 模型（**Generative Pre-trained Transformer, GPT**）

GPT 是一种基于自然语言处理的深度学习模型，它使用 Transformer 结构进行大量无监督预训练。GPT 可用于生成文本、翻译、摘要等任务，并具有强大的生成和理解能力。

机器语言（**Machine Language**）

机器语言是一种用于与计算机硬件直接交流的编程语言。它是一串用于表示指令集的二进制代码。这种语言通常是特定于某种类型的处理器，并且对人来说难以直接阅读或编写。

自然语言（**Natural Language**）

自然语言是人类用于交流和表达思想的语言，包括但不限于英语、中文、法语等。这些语言有自己的语法、词汇和句构，用于口头或书面的交流。

自然语言处理（**Natural Language Processing, NLP**）

NLP 是计算机科学和人工智能领域的一个子领域，关注计算机与人类（自然）语言之间的交互。NLP 技术使计算机能够理解、解释和生成人类语言，从而提高人机交互的效率。

预训练语言（**Pre-trained Language Model**）

预训练语言模型是一种在大量文本数据上预先训练过的模型，通常用于自然语言处理任务。这样的模型可以被用作各种 NLP 应用（如文本分类、情感分析等）的基础，也可以通过进一步的训练（微调）来适应特定的任务或领域。

基于人类反馈的强化学习（**Reinforcement Learning with Human Feedback, RLHF**）

RLHF 是一种结合了人类知识和强化学习的方法。通过将人类反馈作为奖励信号或辅助训练信息，学习算法可以更快地找到有效策略，解决复杂问题，并避免许多强化学习中存在的挑战。

提示语（**Prompt**）

在自然语言处理任务中，提示语是提供给模型的输入语句或问题，用于引导模型生成特定类型的回答或输出。通过设计合适的提示语，可以更好地利用预训练模型的能力，解决各种实际任务。

参考文献

1. X. Amatriain, “Transformer models: an introduction and catalog,” arXiv preprint arXiv:2302.07730, 2023.
2. A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving language understanding by generative pre-training,” 2018.
3. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805, 2018.
4. A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language models are unsupervised multitask learners,” OpenAI blog, vol. 1, no. 8, p. 9, 2019.
5. P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, “Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing,” ACM Computing Surveys, vol. 55, no. 9, pp. 1–35, 2023.
6. W. B. Knox and P. Stone, “Tamer: Training an agent manually via evaluative reinforcement,” in 2008 7th IEEE international conference on development and learning, IEEE, 2008, pp. 292–297.
7. K. S. Kalyan, A. Rajasekharan, and S. Sangeetha, “Ammus: A survey of transformer-based pretrained models in natural language processing,” arXiv preprint arXiv:2108.05542, 2021.
8. G. Warnell, N. Waytowich, V. Lawhern, and P. Stone, “Deep tamer: Interactive agent shaping in high-dimensional state spaces,” in Proceedings of the AAAI conference on artificial intelligence, 2018.
9. J. MacGlashan et al., “Interactive learning from policy-dependent human feedback,” in International Conference on Machine Learning, PMLR, 2017, pp. 2285–2294.

10.D. M. Ziegler et al., “Fine-tuning language models from human preferences,”
arXiv preprint arXiv:1909.08593, 2019.

行业图谱研究项目

一、项目目标和定位

行业图谱是科创金融研究中心基于科技成果转化研究的一项子课题，聚焦于科技成果这一核心要素，从技术链视角切入展开的研究项目。科技成果的转化需要对科技成果有清晰、准确、深刻的认识和理解，能够解析科技成果所包含的学术价值、社会价值、经济价值和人文价值等，从而探索科技成果的未来应用场景，以跨越从 0 到 1 的商业性转化，通过不断理解优化实现社会产业化，并最终成为科技推动社会发展的历史进程。

然而，由于科技天然具有强大的认知壁垒，其先进性、创新性的特点，使得科技成果面临非专业人士看不懂、不敢判断的知识窘境。在成果转化的操作路径中，执行者可分类为三方：成果供给方、成果接收方及连接双方的中介服务机构。除了成果供给方之外，成果接收方和中介服务机构都面临着知识窘境。成果供给方是科技成果的发明人、创造者，对科技成果的学术价值拥有深度认知，但缺乏商业经验和分析社会需求的能力，很难独立实现成果的成功转化；成果接收方是进行成果商业化、产业化的企业，对社会需求敏感，善于进行商业价值的探索，但由于不具备深厚的科研基础，不能对科技成果进行技术层面的准确分析和判断，影响执行效率；中介服务机构虽然具备政策分析、法律服务等领域的专业能力，但同样面临看不懂技术的知识窘境，导致出现无效推介、不合理的专利布局、未来的专利纠纷等潜在危机。这一需求的断层也间接性地影响经济学称之为成果转化“死亡之谷”时期的存在。因此，如何准确认识科技成果，正确判断科技成果的技术领先度，理解科技成果所处的行业地位和产业链发展格局，对于提高科技成果转化具有极其重要的价值。

本研究以国家十四五规划为导向，重点关注与国家战略需求发展相关的重大创新领域。集中在人工智能、量子信息、集成电路、生命科学、生物育种、空天科技、深地深海、现代能源等前沿领域。对基础科研方向进行应用场景的细分，将相关可转化

/转化中的科技成果进行技术链条的梳理，通过专业性的技术解构和解析，形成高逻辑性、易理解性的技术图谱；并在此基础上，对科技成果产业化应用现状进行行业研究和分析，以全球视野定位领先梯队中的科创企业和学术团队的技术实力。通过行业图谱的研究，不仅可以清晰定位高新科技企业的技术竞争力，而且能够对我国相关行业现状和未来方向有更准确的认识。既为科技成果转化提供了专业性知识体系支撑，也有助于指导城镇产业化发展布局、推动产业链融通创新、引导创业投资基金对“硬科技”的积极性及鼓励金融支持创新体系的建设。

二、研究方法

方法学上，行业图谱研究将进行学科领域分级细化，再对技术在应用场景方向上进行详细分级和解构：

（一）一级分类：从应用产业所属学科的角度，以国家十四五规划为导向，重点关注影响国家安全和全局的基础核心领域，包括人工智能、量子信息、集成电路、生命科学、生物育种、空天科技、深地深海、现代能源等。

（二）二级分类：对技术对象进行分类。比如生命科学中包括疫苗、新生物材料、细胞治疗、人工智能、基因技术等技术对象，择一进行技术应用方向分析和流程解析。

1、应用方向的技术流程全景

首先，对 ChatGPT 所采用的技术——基于 GPT-3.5 的自然语言处理能力进行深入分析。这包括技术的架构、开发过程、以及其在自然语言理解和生成方面的创新点。同时，还研究了 ChatGPT 在不同应用场景中的表现，如对话生成、情感分析、机器翻译等，以及这些应用对现实世界的具体影响。

2、应用方向的技术产品细分类

选取了一系列具有代表性的使用案例，例如 ChatGPT 在搜索引擎、办公软件和其他数字产业中的应用。这些案例帮助我们理解 ChatGPT 如何在实际场景中发挥作用，以及它对用户交互体验和企业运营模式的影响。

3、应用方向上某一细分产品的技术开发流程

本研究还包括对不同大型语言模型的比较分析，尤其是 ChatGPT 与其他国内外模型（如 Google 和百度开发的模型）在技术特性、市场表现、和商业策略方面的比较。这有助于揭示各模型的优劣势以及它们在市场中的定位。

4、领先级国际科创企业及学者团队定位

将国际国内最领先的科创企业进行技术平台和产品性能的比较分析，并将其所具备的技术优势定位于上述图谱中。将国内外学者团队的领先性研究成果/转化状态进行分析，并定位于上述图谱中。

三、研究报告形式

行业图谱以结构化脑图为基础形式，辅以文字报告进行解释说明。文字报告的内容框架包括：

概览：概述图谱传递的信息内容、解答的技术问题和目的。

科学背景简述：描述图谱行业背景、技术流程、关键技术平台和竞争点的细节、技术应用的例证及国内外行业发展现状，对图谱做详细内容的补充说明。

专业术语解析：针对重点专业术语进行概念解释。

参考文献。

免责声明

本报告由清华大学五道口金融学院科创金融研究中心（以下统称“研究中心”）编写。本报告仅供研究使用，并非为提供咨询意见而编写。本报告中的信息均来源于本研究中心认为可靠的已公开资料，但研究中心及其关联机构对信息的准确性及完整性不做任何保证。本报告的版权仅为研究中心所有，如需转载，请注明本文为本研究中心的著作。

（作者：朱雅姝为清华大学国家金融研究院资本市场与公司金融研究中心高级研究专员。刘碧波为清华大学五道口金融学院院长助理、副教授、清华大学国家金融研究院资本市场与公司金融研究中心联席主任。）