

ChatGPT 的运行模式、关键技术及未来图景

朱光辉¹ 王喜文²

(1. 北京理工大学人文与社会科学学院, 北京 100081; 2. 北京华夏工联网智能技术研究院, 北京 100085)

摘要: 美国人工智能实验室 OpenAI 开发的人工智能聊天机器人应用 ChatGPT 引发广泛热议, 被认为是继互联网、智能手机之后, 带给人类的第三次革命性产品。互联网开辟了“空间革命”, 智能手机的出现带来“时间革命”, ChatGPT 的横空出世有望形成“思维革命”, 通过替代人类进行创作、创意、解答、咨询、翻译和客服等改变人类思考和处理问题的方式方法, 由此重塑各行业生态乃至整个世界。

关键词: ChatGPT; 运行模式; 关键技术; 未来图景

中图分类号: TP18 **文献标识码:** A **文章编号:** 1005-9245 (2023) 04-0113-10

DOI: 10.14100/j.cnki.65-1039/g4.20230217.001

埃隆·马斯克称 ChatGPT 将颠覆世界; 微软公司以数百亿美元投资 ChatGPT, 并计划将其整合到微软的 Office 办公软件和 Bing 搜索引擎之中; 部分高校和学术机构开展了关于使用 ChatGPT 写论文是否合规的大讨论; 还有部分咨询公司担忧是否会被其替代。2023 年, ChatGPT 的应用热情被点燃, 应用场景不断快速拓展。

一、ChatGPT 的运行模式

2022 年 11 月上线的 ChatGPT 是由美国人工智能实验室 OpenAI 开发的人工智能聊天机器人应用, 上线不到一周用户突破 100 万, 两个月时间吸引活跃用户超亿人, 打破抖音 9 个月吸引用户过亿的记录, 成为历史上用户增长速度最快的应用程序。

回顾 ChatGPT 的发展历程可知, OpenAI 自 GPT 1.0 开始, 就将大型语言模型 (Large Language Model, LLM) 视为通往通用人工智能 (Artificial General Intelligence, AGI) 的必由之路。具体而言, 在 OpenAI 看来, 未来的 AGI 应拥有一个与任

务无关的超大型 LLM, 可以从海量的数据中学习各种知识, LLM 以生成一切的方式解决各类实际问题。除此之外, AGI 能够听懂人类的命令, 便于人类使用。

(一) 幕后: 大型语言模型

ChatGPT “无比强大”的能力主要得益于其依托的大型语言模型。尽管 ChatGPT 加入了人工标注数据, 但量级只有数万, 这一规模的数据量和训练与 GPT-3.5 模型使用的几千亿级别的数据量相比, 几乎可以忽略不计, 基本不会对提升 GPT-3.5 的基础能力发挥作用。因此, ChatGPT 的强大功能主要源自隐藏其背后的“巨无霸”模型——LLM。

对于 LLM 的发展理念, 可以将其理解为“构建一个与任务无关的超大型 LLM, 让它从海量数据中学习各种知识”。图 1 展示了大型语言模型 LLM。第一, 这一 LLM 模型的规模必然是巨大的, 有能力开发出该模型或改动该模型参数的机构较少。对于任务需求方而言, 无论是无数的中小机构还是个人, 即使有能力开源出大型语言模型, 也无力部署这一模型, 更遑论用微调 (Fine-tuning) 技术模式修改模型参数。因此, 追求不修正模型参

收稿日期: 2023-02-16

作者简介: 朱光辉, 北京理工大学人文与社会科学学院研究员; 王喜文, 北京华夏工联网智能技术研究院院长, 高级工程师。

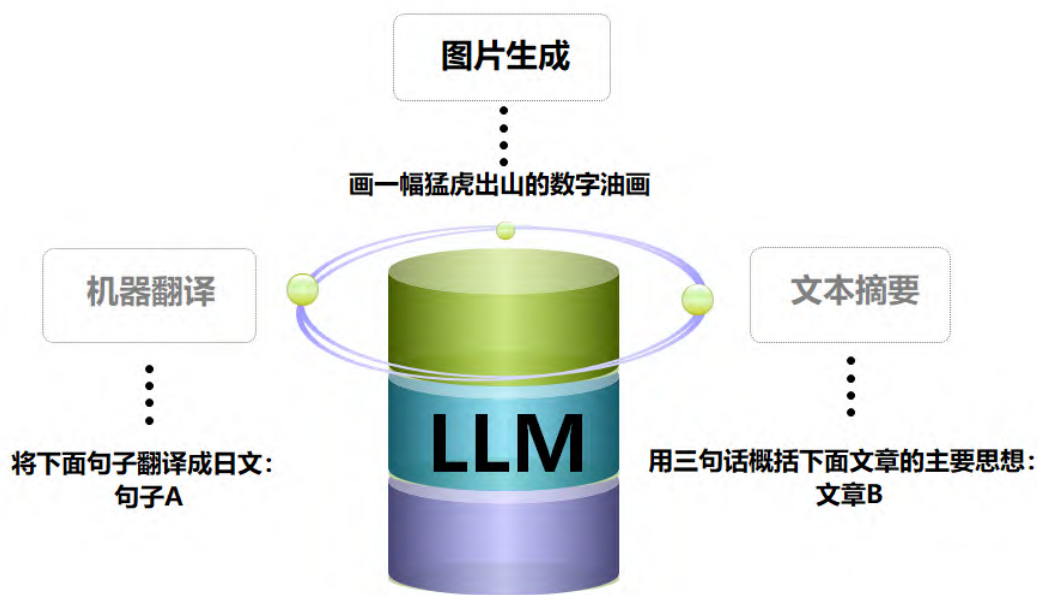


图1 大型语言模型图

数,即能让任务需求方顺利完成的方式,应采取提示词(Prompt)模式,而非微调模式。模型制作方将LLM作为公用基础设施服务,以基础设施即服务(Infrastructure as a Service, IaaS)的模式运行。与此同时,作为服务提供方,要考虑千变万化的用户需求,LLM模型开发方追求让LLM完成尽可能多类型的任务,成为大型语言模型追求走向通用人工智能的现实因素。

第二,LLM应具备强大的自主学习能力。假设人类向其灌输世界上所有能够获得的文本或图片等不同类型的数据,LLM应自动学习其中蕴含的知识点,学习过程无需人的介入就能灵活应用所学知识解决实际问题。数据是海量的,要吸收所有知识,就需要足够多的模型参数存储知识,因此,这一模型必然会“巨无霸”式的模型。

ChatGPT是否向GPT-3.5模型注入新知识?答案是肯定的,这些知识包含在揭秘ChatGPT时提到的“几万人工标注”的数据中,但注入的不是世界知识,而是人类偏好知识。所谓“人类偏好”,包含两方面含义。一是人类表达任务的习惯说法。例如,人们习惯性表达:“把下面句子翻译成日语”,以此表达机器翻译的需求,但LLM并非人类,它如何理解这句话的含义?人类要尽可能让LLM理解这句命令的含义,并正确执行。因此,ChatGPT通过人工标注数据的方式向GPT-3.5

注入这类知识,方便LLM理解人的命令,这是其“了解人类意图”的关键。二是对于什么是好的回答,什么是不好的回答,人类有自己的标准。例如,比较详细的回答是好的,带有歧视性内容的回答是不好的,诸如此类。但这是人类自身对回答质量好坏的偏好。人工标注通过打分模型(Reward Model)将这类信息反馈至LLM数据库。总体而言,ChatGPT将人类偏好知识注入GPT-3.5,以此获得能够听得懂人类语言、自身拥有判断标准的LLM。

就具体过程而言,首先,创建人类偏好数据。随机挑选部分问题,并由标注人员给出高质量回答,形成“人类表达—任务结果”的标注数据,反馈至模型,让其学习——这批数据数量仅有数万,并通过提示词模式进行,即模型参数不发生变化。其次,训练一个反馈模型。随机挑选部分问题,由原始模型输出答案,再由标注人员基于“人类偏好标准”(例如,相关性、信息丰富程度、答案有害、负面情感等)对原始模型的答案进行排序。最后,利用标注好的“人类偏好”数据,训练一个打分模型,这一打分模型会对原始模型的结果进行打分,告诉它什么答案分高,什么答案分低。

以此为基础,整个过程通过循环式地强化学习,将反馈模型和原始模型相链接,当原始模型输出的结果在打分模型中获得较低分值时,它将受到惩罚,同时,被要求重新学习。通过不断循

环，原始模型逐渐迭代升级，直至“脱胎换骨”，彻底掌握人类偏好，变成人类满意的模型，即 ChatGPT。

（二）台前：新型人机交互接口

目前，相关研究已经证明大型语言模型 LLM 对于知识具有强大的记忆能力。但现实世界中，一般不会将记忆能力的强弱作为判断人是否聪明的标准。是否具有强大的推理能力，通常是判断一个人是否聪明的重要标准。ChatGPT 要取得令人惊艳的效果，其背后强大的推理能力必不可少。推理能力的本质是综合运用较多相关知识，推导出新知识或新结论。当模型规模足够大时，LLM 本身具备相应的推理能力。

ChatGPT 的最大贡献在于较好地实现了大型语言模型 LLM 的接口层，让 LLM 适配人类习惯的命令表达方式，而非让人类适配 LLM，绞尽脑汁地想出要达成目标的命令。由此，能够增加 LLM 的易用性和用户体验。

这种交互方式的演变，是一种较为理想的新型人机交互模式。不需要专业的能力和高端的设备，只要开口表达人类诉求，人工智能就能够理解并帮助人类进行解答。在 2022 年 12 月的媒体通稿中，对 ChatGPT 的评价集中于“仿真性”，俨然通过图灵测试一般。这种仿真性，可以认为是 ChatGPT 的“智力”得到进一步提升，变得更加聪明。

二、ChatGPT 的关键技术

有研究发现，ChatGPT 在自然语言处理的系列任务方面，例如，文本理解与生成、对话问答、机器翻译和程序代码生成等方面都有较大进步。从技术层面讲，得益于近几年深度神经网络、大型语言模型研究的不断发展，即海量数据加之巨大的算力催生这样一个大型语言模型的落地应用。换言之，ChatGPT 背后的关键技术离不开大模型算法、大数据和大算力。

（一）算法

ChatGPT 由 GPT-3.5 模型提供支持，GPT（Generative Pre-trained Transformer，生成式预训练转换模型）是一种基于互联网可用数据训练的文本生成深度学习模型。在算法方面，该模型使用“利用人类反馈强化学习（Reinforcement Learning from Human Feedback, RLHF）”的训练方式，包括人类提问机器回答、机器提问人类回答，并不断迭代，让模型逐渐具有对生成答案的评判能力。RLHF 的训练过程可以分解为三个步骤（见图 2）。

1. 预训练语言模型

选取经典的预训练语言模型（LM）作为初始模型。在预训练模型出现之前，深度学习不够成功的原因主要在于两方面：一方面，匹配给某一具体任务的训练数据总量不够多。随着模型容量的增加，对训练数据的需求随之攀升，否则即使达到理



图2 RLHF的训练过程图

想深度,也无法取得预期任务效果,进而成为自然语言处理领域无法逾越的难题;另一方面,深度学习的特征抽取能力不够强。换言之,即使有再多的数据也无济于事,因为模型不能有效吸收数据中蕴含的知识。这两方面原因阻碍了深度学习在自然语言处理领域的突围。GPT 预训练模型的出现,无论是从学术研究角度审视,还是从场景应用角度观察,都代表自然语言处理领域的技术飞跃,并带来整个领域研究范式的转换。

2. 打分模型的训练

基于初始语言模型产出的数据训练打分模型(Reward Model, RM)的目标是评估模型的输出就人类而言是否表现不错。即输入[提示(Prompt),模型生成的文本],输出一个评估文本质量的标记数字。用于训练打分模型的提示词数据一般源自预先富集的数据集,ChatGPT 的 Prompt 数据主要是调用 GPT API 的用户。上述提示词会被放进初始语言模型(第一阶段的模型)中生成文本。可以将打分模型视为判别式的语言模型,从预训练语言模型出发,对 $x=[\text{Prompt}, \text{模型回答}]$, $y=\text{人类满意度}$ 构成的标注语料进行微调;也可以随机初始化,在语料基础上进行直接训练。

3. 基于强化学习进行语言模型优化

在初始的语言模型上生成文本,通过打分模型(RM)判断模型生成的文本是否优质(迎合人类偏好)的基础上,可以使用强化学习(RL)基于打分模型优化初始语言模型。

将初始语言模型的微调任务建模为强化学习

(RL)问题,需要定义策略(Policy)、动作空间(Action Space)和打分函数(Reward Function)等基本要素。策略指基于该语言模型,接收 Prompt 作为输入,再输出一系列文本(或文本的概率分布);动作空间是词表标记在所有输出位置的排列组合;观察空间是可能的输入标记序列,即 Prompt 为词表全部标记在所有输入位置的排列组合;打分函数是基于已设定的 RM 模型,配合部分策略层面的合约进行的打分计算。基于这一打分,可以根据策略优化算法更新模型参数。

通过上述过程,可以迭代式地更新打分模型(RM)和策略模型(Policy),让打分模型对模型输出质量的评估愈加精确,策略模型的输出不断与初始模型拉开差距,使输出文本愈益符合人类的需求和认知。

(二)算力

ChatGPT 能够成为新一代人工智能里程碑,离不开算力发展和数字经济时代形成的大数据共同支持的大型语言模型训练。在算力方面,ChatGPT 使用的 GPT-3.5 模型在微软云计算服务 Azure AI 的超算基础设施(由 V100GPU 组成的高带宽集群)上进行训练,总算力消耗约 3640 PF-days(即按每秒一千万亿次计算,运行 3640 天)。

由此带来两个问题,即巨大的算力需求与资金消耗。训练和运行模型均需要庞大的算力,有研究估测,训练 1750 亿参数大型语言模型的 GPT-3,需要有上万个 CPU/GPU 24 小时不间断地输入数据,所需能耗相当于开车往返地球和月球,且一次运算要花费 450 万美元(见图 3)。

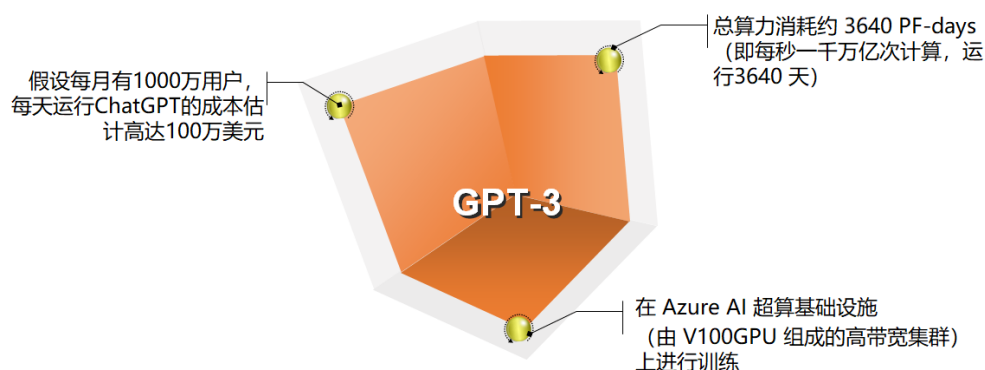


图3 巨额的算力图

此前，企业多通过自行发电的方式满足运营的电力需要，此举不仅耗资巨大，而且还需具备某些与企业业务关联不大的相关专业能力。电网基础设施的运行使供电成为一项公共事业，也使企业可以通过购买电量代替自行发电，就其实质而言，企业是将自行发电变为购买发电服务。集中发电可以使电力的使用更为高效，也意味着更多企业甚至个人可以根据自身需要购买电，不用为其他电量支付任何费用。电力供应的公共化提高了各部门的生产力，改善了社会生活质量，也为新兴产业的发展创造了机会。

信息和通信技术行业正经历与此类似的演进过程。几十年来，公共部门、私人部门、组织和个人等通过投资电脑软件和硬件，像购买商品一样购买信息和通信技术。过去 10 年，随着高速宽带基础设施的普及，信息和通信服务的供给方式快速更新，通过互联网可以将信息和通信技术作为一项服务进行购买。

现阶段，算力如同被广泛使用的电力一般，与此同时，算力也是一项具有潜在破坏性与变革性的创造。未来，各行各业的用户若要在不购买、安装和运行昂贵的电脑硬件的基础上使用服务，就可借助无处不在的有线或无线网络——从“云端”获取算力，这与使用其他公共基础设施服务没有区别（见图 4）。

（三）数据

有资料显示，ChatGPT 拥有多达 1750 亿个模型参数，并在 2023 年年初完成训练。模型训练的

背后离不开大数据的支持，OpenAI 主要使用的公共爬虫数据集拥有超万亿单词的人类语言数据。正是基于海量数据，ChatGPT 展示了三种强大的能力。一是语言生成能力。遵循提示词生成补全提示词的句子。这是目前人类与语言模型最普遍的交互方式。二是上下文学习（In-context Learning）能力。遵循给定任务的几个示例，为新的测试示例生成解决方案。值得注意的是，GPT-3 虽然是语言模型，但上下文学习才是 ChatGPT 的真正重点，而不是“语言建模”（Language Modeling）。三是世界知识能力。包括事实性知识（Factual Knowledge）和常识（Commonsense）。

上述三种能力均来自大规模预训练。在有 3000 亿单词的语料上预训练拥有 1750 亿参数的模型（60% 的训练语料来自 2016–2019 年的 Common Crawl 语料库 + 22% 来自 WebText 语料库 + 16% 来自书籍和报刊杂志 + 3% 来自维基百科）。其中，Common Crawl 是 2008 年至今在一个网站抓取的大型数据集，数据包含原始网页、元数据和文本提取，其文本来自不同语言、不同领域。重点研究实验室一般会优先选取纯英文过滤版（C4）作为数据集。其中，WebText 是一个大型数据集，其数据从社交媒体平台 Reddit 所有出站链接网络中爬取，每个链接至少有 3 个赞，代表流行内容的风向标，对输出优质链接和后续文本数据具有指导作用。

关于 ChatGPT 上下文学习的能力来源及为什么上下文学习可以泛化，现阶段尚未有明确剖析。

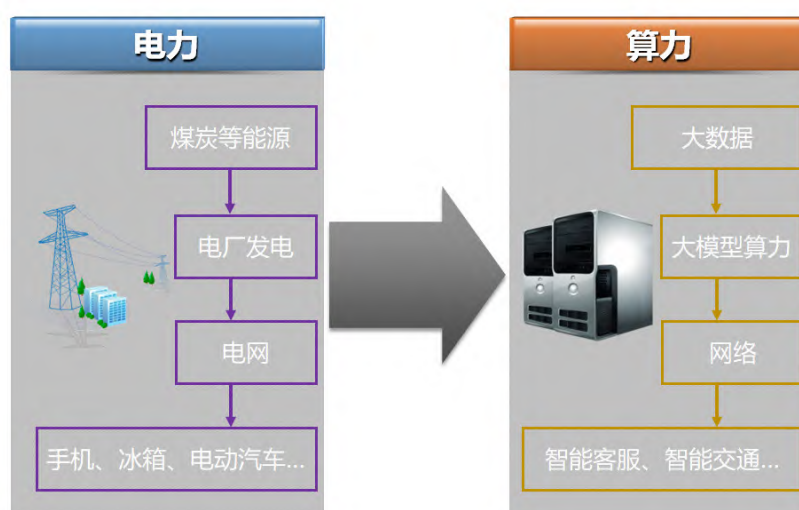


图4 从电力到算力图

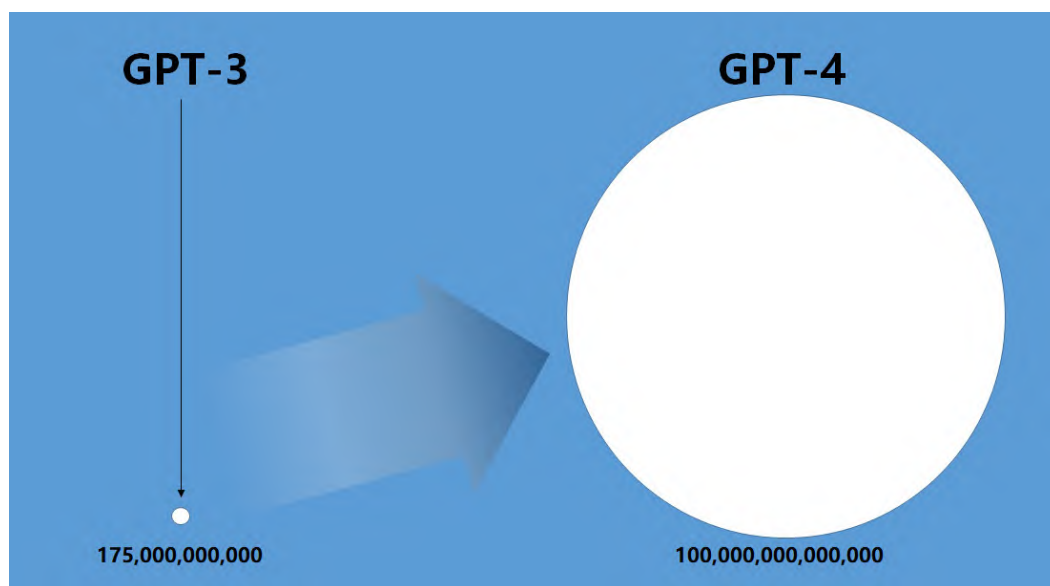


图5 大量的数据图

有人工智能领域专家推测,这种能力可能来自同一个任务的数据点在训练时按顺序排列在同一批处理中。未来,对语言模型预训练促进上下文学习的原理以及上下文学习行为与微调(Fine-tuning)的协同原理值得进一步研究。

现阶段的 ChatGPT 是在拥有 3000 亿单词的语料基础上预训练拥有 1750 亿参数的模型, GPT-4 将是一个拥有超 100 万亿级别参数的大模型(见图5)。根据学术界的既有研究可知,深度神经网络的学习能力和模型的参数规模呈正相关。人类大脑皮层有 140 多亿个神经细胞,每个神经细胞有 3 万余个突触,因此,大脑皮层的突触总数超 100 万亿个,神经细胞通过突触相互建立联系。一旦 GPT-4 实现 100 万亿参数规模,就可以堪比人类大脑,意味着它将达到与人类大脑神经触点规模的同等水平。如果上述假设成为现实,不仅意味着 GPT-4 系统可以改造人类的思想和创作能力,形成人工智能超越专业化能力和大众化趋势,而且意味着这一系统开始具备人类思维能力,并有可能在某一方面或其他方面替代人类。

三、ChatGPT 的未来图景

相较以往,人工智能进化的深度学习能力,对大部分人而言只是一个高深的概念。ChatGPT 通过生成式预训练转换模型 RLHF 基于人类反馈的强化学习这一方式,让所有人真正接触到“人工智能+深度学习”带来的变化以及对于人类的生活产生的

影响。因此, ChatGPT 可能会加速人工智能和深度学习理论在经济社会各领域的普及应用。

学术界普遍认为, ChatGPT 的未来应用场景充满无限可能。从社交媒体到广告创意,从游戏到影视娱乐,从编程到深度写稿,从平面设计到产品工业设计,从文字翻译到外事同声传译等,每个原本需要人类创作的行业都可能被 ChatGPT 颠覆性重塑。ChatGPT 被公认为是继互联网、智能手机之后,带给人类的第三次革命性产品。互联网开辟了“空间革命”,使人类可以实时与全世界链接,不必奔赴现场,通过互联网就可以沟通、教学、视频会议,使政治、社会和商业等领域发生连锁变化;智能手机的出现带来“时间革命”,通过可拓展安装的各种 APP 应用软件,可以实现最快交易、最速送达,为人类的生活、工作和消费带来巨大变化; ChatGPT 的横空出世,有望形成“思维革命”,替代人类进行创作、创意、解答、咨询、翻译、客服等,改变人类思考和处理问题的方式方法,并由此重塑各行业生态,甚至重塑整个世界。

现阶段的 ChatGPT 以高度拟人化的对话问答模式带来更好的交互体验,短期内将促进金融、媒体、医疗等诸多领域自然语言处理的应用。例如,在金融领域, ChatGPT 利用其大模型能够大幅提升语义搜索能力,面对复杂多变的投资理财咨询,能够准确找到满足用户需求的咨询结果;又如,招商银行信用卡已经基于 ChatGPT 撰写宣传稿件,写出“生命的舞台上,我们都是基因的载

表1 ChatGPT教学应用举例表

序号	能力	用途
1	回答问题	ChatGPT不仅可以作为教师的人工智能助教，而且可以作为学生的人工智能学伴。
2	论文写作帮助	ChatGPT可以帮助学生写论文，建议文本题目，概述结构，提供思路，并提供改进帮助和建议。
3	学习材料的生成	ChatGPT可以生成学习材料，例如，摘要、抽认卡和测验，帮助学生复习和学习课程材料。
4	个性化辅导	ChatGPT可以通过回答问题和提供个性化反馈，为学生提供个性化辅导。
5	会话练习	ChatGPT可用于语言学习，提供与虚拟语言导师的对话练习。
6	基于游戏的学习	ChatGPT可以被整合到教育游戏中，以提供有吸引力的互动学习体验。
7	支持学生的研究性学习	ChatGPT可以通过提供信息来源、搜索策略和其他资源建议，帮助学生进行研究项目。
8	创意写作协助	ChatGPT可以通过产生想法、改进建议和对书面工作的反馈，协助创意写作；也可以为学生提供书面工作的反馈，包括改进建议、语法修正和结构建议。
9	评估和考核测验的生成	ChatGPT可以生成评估，例如，小测验和考试，帮助教师评估学生的理解。
10	课程材料生成	ChatGPT可以生成完整的课程材料，例如，教学大纲、课程计划和阅读材料。
11	内容翻译	ChatGPT可以将教育内容从一种语言翻译成另一种语言，使学生更容易理解和获取信息，通过提供学习不同语言的信息和资源，协助进行多语言教育。
12	复习、练习和考试准备	ChatGPT可以帮助学生准备考试，提供练习题、考试策略和其他资源。
13	解决问题的帮助	ChatGPT可以帮助学生解决问题，提供分步指示、策略建议和提示。
14	职业指导	ChatGPT可以通过提供潜在职业、工作要求和相关信息帮助学生进行职业指导。
15	时间管理协助	ChatGPT可以帮助学生进行时间管理，提供提示和策略，以保持组织和任务的优先次序。
16	家庭作业协助	ChatGPT可以通过提供解释、例子和额外资源帮助学生完成作业，也可以通过提供解释、分步指示和其他资源帮助学生完成家庭作业。

体”、“如果说基因给我们的生命带来了基础，那亲情便是对生命的深刻赋予。它不由基因驱使，而是一种慷慨的选择”等富有诗意的文案。在投研方面，业内首份采用 ChatGPT 撰写的行业研究报告完成度较高，但距专业研究报告仍存在较大差距。财通证券团队介绍，“ChatGPT 在文字表意、标题撰写等方面均具有较高水平”。在媒体领域，大量的稿件均可以通过 ChatGPT 进行自动化生产，其独创性和创造力并不输专业人员。未来，文字工作者应积极探索新技术帮助其提高生产效率，让 ChatGPT 起草初稿，人类只需在其基础上进行修改完善。在医疗领域，ChatGPT 可以替代

专业人员为患者提供心理咨询、问诊和解答服药建议，等等。

以教育领域为例，从媒体报道中可以整理出 16 种 ChatGPT 教学应用用法（见表 1）。据媒体报道，ChatGPT 除在高校占有一席之地外，其适用范围已经下探至学龄前儿童和中小学阶段。据报道，有一位居住在国内某城市的 4 岁孩子的妈妈向记者明确表示，她每天都会登录 ChatGPT，和“它”聊天，并把“它”推荐给其他妈妈，有助于解答孩子的教育问题。部分中小学教师也在思考将 ChatGPT 融入自己的教学工作，一名刚入职的小学语文教师解释道，小学生由于年龄小、心智

尚未发育成熟,上课时不仅需要教师在知识学习上提供帮助,而且需要在心理层面进行全方位培养,她经常询问 ChatGPT 诸如“如何矫正小学生行为习惯”、“教师如何与内向的小学生沟通”等问题。

由此可见,ChatGPT 在各行各业均具有无限的未来应用场景,正在快速走进人类的工作和生活,会成为继互联网、智能手机之后,人人都离不开的工具。

四、ChatGPT 的影响分析

科技进步造福人类经济社会,让人类可以提质增效、实现高质量发展。同时,科技是一把“双刃剑”,使用不当,会对经济社会产生负面效应。

(一) 正面效应:大幅提质增效

科技创新可以提高经济结构的效率和灵活性,提高企业竞争力,节约能源、资源和人工成本,推动产业升级,促进供给侧结构变革,改善质量和效率,催生新产品、新技术和新模式,进而促进经济增长。ChatGPT 作为一种科技创新,通过分析大规模数据,找到诸多规律,生成新作品。其不仅限于分析已经存在的东西,在某些情况下,较人类更具创造力,且创造得更好。未来,ChatGPT 处理的领域包括所有知识工作和创造性工作,可能涉及数亿的人工劳动力,使相关领域人工劳动力的效率和创造力得到大幅提高,不仅较以往更快、更高效,而且更完美、更具创意。

第一,ChatGPT 能够带来更高的生产率。与传统的文本识别或语言理解系统相比,聊天机器人系统更容易使用,可以更有效地了解输入的问题,提高结果的准确性,改变传统的生产和服务方式,让企业更有效地利用有限的资源产生更多的产品和服务,在提升效率的同时,降低成本。第二,ChatGPT 新技术的应用可以提升企业的核心竞争力,带来增量利润,促进业务增长,不断改善经济社会结构,带动产业数字化转型和智能化升级。此外,技术创新可以满足更多的消费需求,促进投资市场的发展。例如,ChatGPT 可以帮助客服代言人进行个性化会话,有能力快速了解客户的需求、分析重点,及时回答客户的提问,能够提高客户的满意度和对公司的信任度。在此基础上,ChatGPT 新技术的应用可以使企业拥有更多的消费者,拉动营

收和利润的增加,促进投资市场的发展,最终实现经济的可持续发展。

(二) 负面效应:作弊与知识产权纠纷

ChatGPT 被滥用的问题愈发明显,最常见的是作弊问题。在美国,北密歇根大学一名学生使用 ChatGPT 生成的哲学课小论文“震惊”教授,得到全班最高分。有调查显示,89% 的美国大学生承认使用 ChatGPT 做家庭作业,53% 的学生用它写论文,48% 的学生使用 ChatGPT 完成测试。

据媒体报道,近期,有多所欧美高校对 ChatGPT 发出禁令。法国巴黎政治学院宣布禁止学生使用 ChatGPT 和其他人工智能产品完成报告,除非教师有特定课程需求,否则学生使用 ChatGPT 完成报告,最严重将面临退学处罚;美国纽约市公立学校只有在进行人工智能与科技相关教学时,才能由教师申请在课堂上使用 ChatGPT;澳大利亚、印度、英国等多所大学也限制学生使用 ChatGPT,尤其是在校园内以及考试期间。

ChatGPT 这类人工智能产品虽然能为学生提供快速简洁的答案,但无法帮助学生培养批判思维与解决问题的能力。与此同时,ChatGPT 创作的内容所有权归使用者所有,但如果产生的内容有侵犯他人知识产权的行为,那么,将沦为知识产权纠纷事件。例如,使用 ChatGPT 在未经授权的具有知识产权的图片或文字基础上创作的内容,可能出现知识产权纠纷问题。对使用 ChatGPT 生产内容而言,知识产权纠纷不可忽视,知识产权保护、合法性检查、协调和解决、技术防作弊以及持续的监管等措施,将是未来应对 ChatGPT 知识产权纠纷可能性时采取的重要措施。

五、政策建议

ChatGPT 是新一代人工智能的代表性应用,是未来的高科技风口。唯有把握此次技术红利,前瞻性布局,才能引导 ChatGPT 更好地服务经济社会发展,创造更大更多价值。

一是引导企业融入 ChatGPT 浪潮。ChatGPT 的重点在于创造,通过大量语料训练和强大的纠错能力,对素材进行消化、整理和再输出,与传统的 IT 行业云计算模式完全不同。以阿里云为例,其主要负责为客户提供云存储和数据库服务,程序运作的底层逻辑是先有问题再找答

案。ChatGPT 是从 0 到 1 主动创作内容,侧重点区分较为明显。因此,笔者建议各类企业通过组织企业负责人培训等工作,让相关负责人意识到 ChatGPT 时代的到来和重大意义,促使企业将人工智能主导的内容创作变为公司核心业务的重要组成部分。

二是推动 ChatGPT 技术融入元宇宙产业布局。元宇宙被认为是数字经济的下一个支撑点,受到地方政府高度重视,纷纷将其列为新兴产业,不断谋划元宇宙发展布局。据统计,浙江省、江西省、湖北省、河南省 4 省以及北京市、上海市等均从省级层面对元宇宙进行布局。杭州市、合肥市、成都市、武汉市将元宇宙写入 2022 年政府工作报告。ChatGPT 的出现为所有人提供了以自然语言对话方式进行文本生成的新方式和新工具,大幅降低了构建元宇宙的门槛,可以用与构建互联网完全不同的方式构建元宇宙。大量非专业人员可以通过语言描述自身需求,ChatGPT 据之自动生成设计图和代码,使效率得到极大提高,成本大幅下降,元宇宙内容数量将得到进一步丰富。笔者建议各地元宇宙产业发展的主管部门定期追踪研究 ChatGPT 技术进展和国内外先进应用场景案例,以更快的速度实现元宇宙发展布局。

三是促进 ChatGPT 技术融入数字人(Digital Human/Meta Human)功能。随着 AR/VR(增强现实/虚拟现实)、虚拟引擎、3D 建模等硬件技术的日臻成熟,数字人将迎来重大发展契机,要充分利用新科技提供更加完善的沉浸式体验,极大地提升数字人的商业价值,拓展其应用领域。例如,数字人可以承担向导角色,通过数字人与用户进行多模态互动,“面对面”实现低延迟实时交互交流,为游客提供路线规划、信息查询、导览讲解等智能服务,在社交和娱乐过程中持续为用户带来新鲜体验,有效提高景区吸引力,提升景区的品牌价值和商业价值。以北京市为例,2022 年 8 月,北京市发布《北京市促进数字人产业创新发展行动计划(2022-2025 年)》,明确提出到 2025 年,北京市数字人产业规模突破 500 亿元的发展目标。与此同时,全国诸多省区市纷纷将数字人纳入新科技发展目录,谋划进一步推动其普及落地。ChatGPT 将彻底改变数字人的全部构建技术。从数字人的面貌、穿着、形态设计、行动和姿态以及自然语言沟通,到数字人的任务执行与对环境的感知和交互等,ChatGPT 将使数字人

的制作更加便捷和简单,让数字人的功能更加丰富。笔者建议政府、企业、各大高校和科研院所研究落实 ChatGPT 技术在数字人中的应用,超额实现数字人产业发展目标。

四是重视 ChatGPT 隐患,将其纳入社会科学研究课题。由于 ChatGPT 训练数据源自互联网,企业或科研所在使用时需要对 ChatGPT 生成的作品进行把关,防范知识产权风险。尤其是高校和科研院所工作人员在使用 ChatGPT 撰写论文、发明专利等过程中,一方面,建议其在使用 ChatGPT 创作内容时,务必核实内容不违反他人知识产权,并对相关内容拥有授权,避免出现知识产权纠纷;另一方面,应该有第三方帮助其“验明正身”,通过技术手段建立一个有效的、具有针对性的持续监管体系,以确保 ChatGPT 生成内容的合法合规。例如,通过检测工具,明确文本是由人类撰写还是由 ChatGPT 生成。科技在进步,社会科学也要及时跟进,建议高校、科研院所等单位将 ChatGPT 列为重大课题,注重科技创新过程中的伦理问题,注重科技创新可能造成的不利社会影响,制定相关政策法规,促进自然科学与社会科学的交融共振。

参考文献:

- [1] ChatGPT: Optimizing Language Models for Dialogue [A/OL]. <https://openai.com/blog/chatgpt/#rf2>.
- [2] The Day the AGI Was Born [A/OL]. <https://space.swyx.io/p/everything-we-know-about-chatgpt>.
- [3] OpenAI invites everyone to test ChatGPT, a new AI-powered chatbot—with amusing results [A/OL]. <https://arstechnica.com/information-technology/2022/12/openai-invites-everyone-to-test-new-ai-powered-chatbot-with-amusing-results/>.
- [4] 中小学教师 ChatGPT 的 23 种用法 [A/OL]. <https://new.qq.com/rain/a/20230215A003EP00>.
- [5] 后 GPT 3.0 时代,主流大模型技术精要详解,走向 AGI 之路的大门已开 [A/OL]. <https://www.51cto.com/article/744516.html>.
- [6] 张俊林: 由 ChatGPT 反思大型语言模型(LLM)的技术精要 [A/OL]. <https://www.toutiao.com/article/7186818289258021434/?wid=1676446398108>.
- [7] ChatGPT 正“侵入”国内教育界,是敌是友? 咋对待? [A/OL]. <https://baijiahao.baidu.com/s?id=1757680347167478557&wfr=spider&for=pc>.

ChatGPT: Operation Mode, Key Technology and Future Prospects

ZHU Guang-hui¹ WANG Xi-wen²

(1.School of Humanity and Social Sciences, Beijing Institute of Technology, Beijing 100081 ;
2. Beijing Huaxia Institute of Industrial Internet Smart Technologies, Beijing 100085)

Abstract: The AI chatbot application ChatGPT developed by OpenAI, an American AI lab, has caused widespread discussion and is considered to be the third revolutionary product after the internet and smartphones. The internet has opened up “space revolution”, and the emergence of smart phones has brought about “time revolution”. The advent of ChatGPT is expected to form a “thinking revolution” that will change human thinking and problem-solving methods by replacing humans in producing composition and other creative activities, answering questions, and providing consultation, translation and customer services, thereby reshaping the ecology of various industries and even the world.

Key words: ChatGPT ; Operation Mode ; Key Technology ; Future Prospects

[责任编辑: 王文秋]

[责任校对: 曹晶晶]