

ChatGPT：研究进展、模型创新及医学信息研究应用场景优化^{*}

陈凌云¹ 姚宽达¹ 王 茜¹ 方 安¹ 李 刚²

(¹ 中国医学科学院/北京协和医学院医学信息研究所 北京 100020 ² 中国农业银行 北京 100005)

〔摘要〕 目的/意义 梳理分析 ChatGPT 训练流程与理论模型，为医学研究提供参考借鉴。方法/过程 系统梳理 2018 年以来 GPT-1 发布至今相关模型和流程文献资料，分析 ChatGPT 核心流程、理论模型及其创新点。根据现有资料分析 ChatGPT 预训练监督、自动评估以及强化学习近端策略优化（proximal policy optimization, PPO）模型 3 个层次技术构成。结合医学研究需求，分析人工智能技术面向医学信息领域应用的优化方向。结果/结论 ChatGPT 技术应用的突破是流程、算法和模型有效组合和不断迭代累积的结果，其模型及研究方法可以应用于医学文献自动化阅读与知识提取、基因研究与疾病风险评估等方面。

〔关键词〕 ChatGPT；预训练监督模型；强化学习近端策略优化模型

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2023.07.003

ChatGPT: Research Progress, Model Innovation, and Optimization in Medical Information Research Applications

CHEN Lingyun¹, YAO Kuanda¹, WANG Qian¹, FANG An¹, LI Gang²

¹ Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China; ² Agricultural Bank of China, Beijing 100005, China

〔Abstract〕 **Purpose/Significance** To analyze and outline the training process and theoretical model of ChatGPT, and to provide references for medical research. **Method/Process** The literatures on relevant models and process of GPT-1 released since 2018 are systematically reviewed. The core process, theoretical model, and innovative aspects of ChatGPT are analyzed. The three-level technical components of ChatGPT are examined, which include pre-training supervision, automatic evaluation, and proximal policy optimization (PPO) for reinforcement learning. Combining with the needs of medical research, the optimization direction of artificial intelligence (AI) technology for medical information application is analyzed. **Result/Conclusion** The breakthroughs in the application of ChatGPT technology result from effective combinations of processes, algorithms, and models through continuous iteration and accumulation. The models and research methods of ChatGPT can be applied in automated reading and knowledge extraction of medical literature, gene and disease risk assessment and so on.

〔Keywords〕 ChatGPT; pre-training supervised model; reinforcement learning proximal policy optimization (PPO) model

〔修回日期〕 2023-06-09

〔作者简介〕 陈凌云，硕士，发表论文 4 篇；通信作者：方安，博士，研究馆员。

〔基金项目〕 国家重点研发计划项目子任务（项目编号：2022YFF0711902-2）；国家社会科学基金项目（项目编号：21CTQ016）；中国医学科学院医学与健康科技创新工程（项目编号：2021-I2M-1-056）。

1 引言

自 2022 年 11 月, ChatGPT 迅速成为现象级焦点, 引发市场广泛关注, 掀起新一轮人工智能浪潮。推出仅两个多月, ChatGPT 全球用户已突破 1 亿, 成为历史上用户增长速度最快的应用程序。

ChatGPT 是一种基于语言模型的人工智能系统, 由开放人工智能 (OpenAI) 公司开发, 是 GPT 系列模型的一部分, 旨在进行对话式交互。ChatGPT 经过大规模预训练, 可以理解和生成自然语言文本, 能够与用户对话, 并根据上下文生成连贯的回复。可被用于提供智能客服、虚拟助手、在线聊天等场景, 为用户提供实时的问题解答和交流服务。ChatGPT 在自然语言处理领域具有广泛的应用前景, 可以帮助人们更方便地与计算机交流和获取信息。

ChatGPT 引入基于人类反馈的强化学习模型,

并使用高质量大数据, 特别是其训练流程与理论模型在学术界引起极大关注^[1], 在需要算法与技术、组织与检索、人机智能交互^[2]与协同支持的医学领域产生了范式转换、方法创新的巨大影响^[3]。其预训练监督、自动评估以及强化学习近端策略优化 (proximal policy optimization, PPO) 模型 3 个层次技术创新, 在医学信息研究应用方面具有重要意义, 能够在医疗大数据临床文本处理和知识发现、医学文献知识提取、优化医学问诊、改善医患交流和沟通, 以及在医院管理和科研中提供决策支持和数据分析。

2 模型研究进展

自 2018 年 GPT 被 OpenAI^[4] 公司开发以来, 经历神经网络、深度学习, 以及 Transformer 模型等多种模型迭代累积, 逐步完善与成熟^[5], 见图 1。

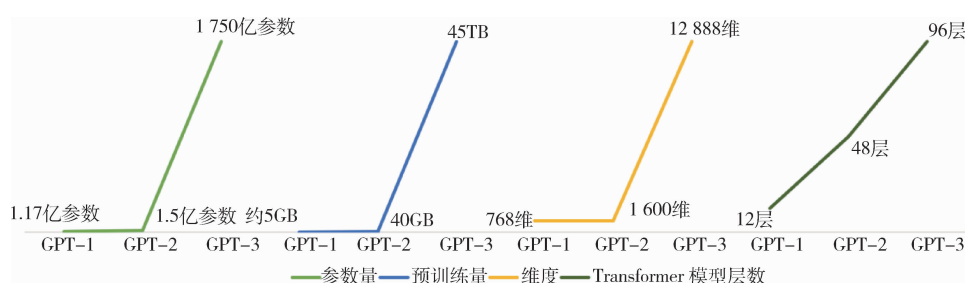


图 1 GPT 迭代对比

GPT-1 是 GPT 系列第 1 个版本, 使用 12 个 Transformer 编码器, 采用自回归语言模型, 创新性地将语境信息编码为隐藏向量, 在此基础上生成任务^[6]。GPT-2 是 GPT 系列的第 2 个版本, 于 2019 年发布, 具有更多参数和更多层网络结构, 使用 48 个 Transformer 编码器和 1.5 亿个参数, 进一步使用无监督预训练方法, 自动学习语言规律和特征, 获得更强的语言泛化生成能力^[7]。GPT-3 是 GPT 系列第 3 个版本, 于 2020 年发布, 使用 1 750 亿个参数, 采取零样本和多样本学习方法, 甚至在没有任何样本的情况下生成解决方案, 具有较强对话交互能力, 并在对话中实现逻辑推断, 获得生成连贯和富有逻辑性的语言能力^[8]。

2022 年 ChatGPT 继续基于 Transformer 模型架

构, 在微软 Azure AI 超级计算基础设施上训练, 吸收 GPT-1、GPT-2、GPT-3 模型经验, 参数由亿级迭代至千亿级, 使用人类反馈的强化学习模型, 应用强化学习 PPO 模型, 成功实现多问题提示、动态交互的大规模预训练语言能力^[9]。

GPT 模型系列从 GPT-1 到 ChatGPT, 其算法区别体现在: GPT-1 采用单向 Transformer 模型, 利用自回归方式预训练和生成。GPT-2 引入更大模型规模和参数数量, 采用更多层 Transformer 模型, 提升生成能力和语言表达连贯性。GPT-3 进一步增加模型的规模和参数数量, 引入零样本和小样本学习概念, 展现更强的通用性和泛化能力。ChatGPT 是基于 GPT-3.5 架构训练的版本, 具备更强的语言理解和生成能力, 但具体算法细节并未公开。应

用方面区别体现在：GPT - 1 主要应用于生成文本、回答问题和完成填空题等自然语言处理任务。GPT - 2 在生成任务方面取得显著进展，可用于生成新闻、故事、代码和对话等任务。GPT - 3 具备较强的通用性和泛化能力，可适应多个自然语言处理任务，包括翻译、问答、摘要生成、对话交互等。ChatGPT 作为语言模型，可被应用于回答问题、生成文本、提供建议等任务，但其基于已有数据进行训练。

近期，OpenAI 公司公布 GPT - 4 模型^[9]，继续优化大规模预训练模型，支持图文混合输入，实现更准确、更真实的机器翻译、自然语言输出等能力，表现出专业和学术方面与人类基本相同的能力。

ChatGPT 采用多种模式叠加的方法。其中，对比预训练是指预训练多个模型，并将其结果对比和融合。通过对比预训练，ChatGPT 能够获得更全面和多样化的语言表达能力。ChatGPT 还进行多种微调，包括人工加工数据微调、指令微调和代码数据

微调。这些微调过程可以提高 ChatGPT 在特定领域或任务中的性能表现，使其更适用于医学信息应用等特定场景。人工加工数据微调是指利用人工标注或编辑的数据微调，以增强 ChatGPT 在特定领域的理解和表达能力。指令微调是通过针对特定任务或指令的训练样本微调，使 ChatGPT 能够更好地理解和遵循给定的指令。代码数据微调是指使用编程代码相关的数据微调，使 ChatGPT 能够更好地理解和生成与编程相关的内容。

3 模型创新

ChatGPT 由 GPT - 3.5 模型提供支持，采用“人类反馈强化学习”训练方式，包括人类提问机器回答、机器提问人类回答，不断迭代和反馈，模型逐渐具有生成答案和相应的评判能力，见图 2^[10]。为下一阶段医学研究提供思路和方法。ChatGPT 创新主要体现在以下 3 方面。

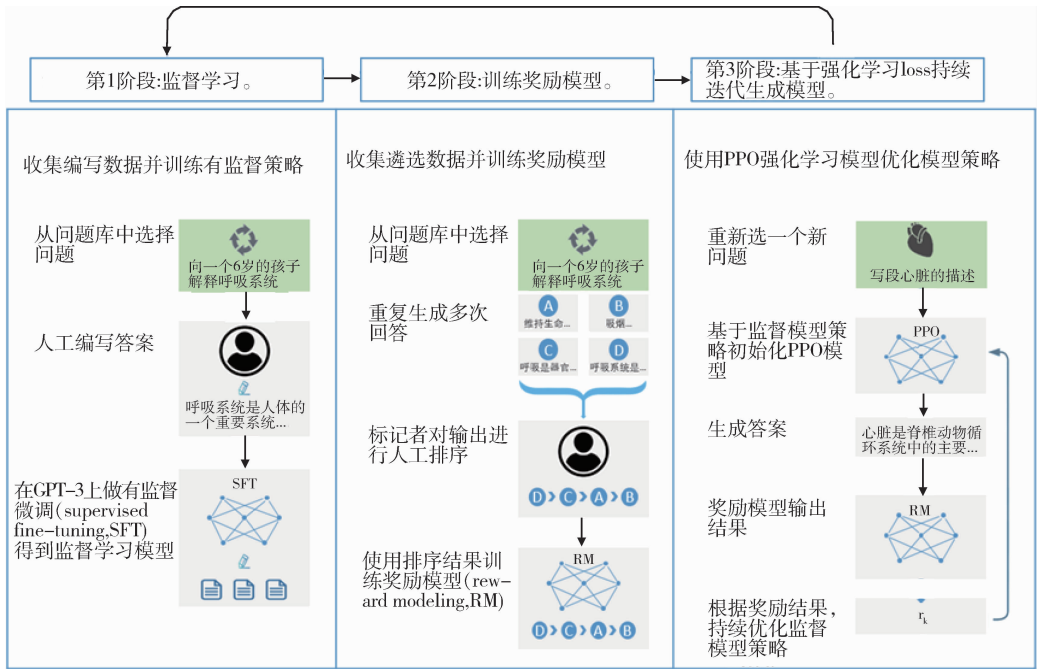


图2 ChatGPT 训练流程

3.1 微调 GPT 模型，更好适应任务

微调预训练模型是 ChatGPT 第 1 层级创新。ChatGPT 预训练模型初始层主要负责提取原始文本

中特征，输出层和附近层级负责生成语言任务，并在模型输出层添加具体分类器，将生成的响应与真实响应比较和计算打分，调整模型参数，适应特定对话目标。同时，ChatGPT 不断微调学习率和正则

化方法，初始阶段学习率较大，后续阶段逐渐减小，以帮助模型更好地适应对话任务；正则化则分别采用 dropout、L1、L2 等方法，避免模型过度拟合训练数据，提高泛化能力，降低模型复杂度^[11]。

3.2 通过人工标注，训练生成打分模型

生成自动奖励模型、采用自动评估机制是 ChatGPT 第 2 层级重要创新。首先，ChatGPT 模型对输出结果进行人工标注排序，舍弃常用的直接打分方法，提升模型的普适性和准确性。其次，奖励模型对排序结果两两组合，计算差值，遍历所有排序靠前与排序靠后结果的差值，对数处理计算整体值，生成奖励模型。奖励模型的损失函数：

$$\text{loss}(\theta) = -\frac{1}{\binom{K}{2}} E_{(x, y_w, y_l) \sim D} [\log(\sigma(r_\theta(x, y_w) - r_\theta(x, y_l)))] \quad (1)$$

其中 $r_\theta(x, y)$ 是奖励模型对提问信息 x 和输出信息 y 的标量输出， θ 为参数， y_w 是 y_w 和 y_l 输出信息对的排序靠前的项， D 是比较数据集^[12]。

最后，ChatGPT 采用梯度下降法最小化损失函数，学习优化策略，生成奖励模型。据此，采用此自动奖励函数打分，ChatGPT 有效改进预训练模型，即当原始模型输出的结果在打分模型中获得较低分值，或者输出结果偏离人类认知时，模型将受到惩罚，被要求重新学习。通过不断循环，初始模型逐渐被迭代升级，逐渐掌握人类偏好，成为人类满意的模型。

3.3 应用强化学习 PPO 模型，持续优化自主学习能力

应用强化学习 PPO 模型是 ChatGPT 第 3 层级关键创新。基于上述奖励模型，ChatGPT 获取预训练模型输出结果的分值后，返回强化学习 PPO 模型，进一步优化微调 GPT 模型参数。在此阶段，强化学习 PPO 模型为 ChatGPT 的发展起到极大推动作用。强化学习 PPO 模型目标函数：

$$\begin{aligned} \text{objective}(\varphi) = & E_{(x, y) \sim D_{\pi_{\varphi}^{RL}}} [r_\theta(x, y) - \\ & \beta \log(\frac{\pi_{\varphi}^{RL}(y|x)}{\pi^{SFT}(y|x)})] + \\ & \gamma E_{x \sim D_{\text{petrain}}} [\log(\pi_{\varphi}^{RL}(x))] \end{aligned} \quad (2)$$

其中 π_{φ}^{RL} 是学习后的策略， π^{SFT} 是监督训练模

型， D_{petrain} 是预训练模型概率分布。 β 是惩罚系数， γ 是预训练损失系数，以分别控制 KL 惩罚和预训练梯度的强度^[12]。

对于传统 PPO 强化学习模型，包含策略、动作空间和奖励函数等基本要素^[13]。策略指基于语言模型，接收 prompt 提问作为输入，输出一系列文本（或文本的概率分布）；动作空间是词表标记的排列组合；传统 PPO 模型目标通常包括模型值和“裁剪项”，以获得稳定输出结果：

$$J_{PPO2}^{\theta'}(\theta) \approx \sum_{(s_t, a_t)} \min \left(\frac{p_{\theta'}(a_t | s_t)}{p_{\theta^{\text{prime}}}(a_t | s_t)} A^{\theta^{\text{prime}}}(s_t, a_t), \right. \\ \left. \text{clip} \left(\frac{p_{\theta'}(a_t | s_t)}{p_{\theta^{\text{prime}}}(a_t | s_t)}, 1 - \varepsilon, 1 + \varepsilon \right) A^{\theta'}(s_t, a_t) \right) \quad (3)$$

相比较而言，ChatGPT 模型的目标函数是创新性地增加 KL 惩罚和预训练梯度两个部分，成为奖励函数、模型值与人工值差距，以及特定场景的泛化值 3 个部分组成的函数，可准确获得符合人类需求和认知的结果，最重要的是能够主动适应泛化场景，获得自主学习能力^[14]。

4 面向医学信息领域应用的优化方向

目前，医学研究主要基于规则、机器学习和深度学习方法进行研究，大都可以归为有监督学习模型。医学文献、病理研究、基因研究等在无监督学习研究方面，已经取得长足进步，取得较好研究成果^[15]。但结合 ChatGPT 训练流程和模型创新点，基因、文献、问诊平台建设等将能产生新的研究成果。ChatGPT 赋能医学信息应用可体现在以下两方面。一是数据加工与语料库构建。ChatGPT 可以用于处理和加工医学领域大规模数据，包括医学文献、临床数据、病历信息等，通过 ChatGPT 的语言理解和生成能力，可以实现对数据的自动化整理、分类和标注，构建更丰富、结构化的医学语料库。二是医学信息或情报智能分析。通过 ChatGPT 的语义理解和推理能力，可以理解和解析医学文本、病历描述等信息，并从中提取关键信息、推断相关关系和趋势，进而辅助医学专业人员进行信息检索、知识发现和决策支持，提供更准确和智能化的医学信息分析服务。ChatGPT 面向医学信息领域应用的

优化方向主要涉及以下 3 个方面。

4.1 进一步研究自动奖励模型，提升动态准确性

OpenAI 公司爬虫数据集超万亿，GPT - 3 模型拥有多达 1 750 亿个参数，ChatGPT 模型参数达数十亿级。但是模型参数越大，拟合效果和准确性未必越好。ChatGPT 创新性地应用自动奖励模型，进一步发挥自动校准作用，准确、真实实现语言生成能力。ChatGPT 自动奖励模型也不是一成不变的，而是不断动态生成的，其本身就是“GPT”模型，随着 ChatGPT 预训练模型的改进，打分奖励模型也在优化改进。

医学文献自动化阅读与知识提取方向的应用潜力巨大。通过应用 ChatGPT 等技术，可以实现文献筛选、知识抽取与整合，疾病诊断与治疗指南的自动化，探索新兴研究领域，促进学术交流与合作，提高医学研究效率和质量，为医学领域带来巨大的变革和进步。

在医学研究中，无监督学习模型无法最大成效地发挥作用，原因之一是没有自动评估机制和自动奖励模型，不能对初始模型动态校准，更无法对模型参数迭代升级。而 ChatGPT 模型中的自动奖励模型为此类研究提供了思路。

进一步看，在医疗大数据临床文本处理与知识发现研究中，数据准备、文本挖掘、评估反馈等研究都较深入，成果也较丰富，后续可应用 ChatGPT 自动奖励模型方法，研究知识发现领域相应的奖励模型，动态评估文本处理与知识发现研究，进一步提升准确性和自动化。

在生物医学科技文献知识对象挖掘中，数据和信息量较大，专业术语较复杂。前期在结合专业术语和语言特征语料库标注的基础上，应用 ChatGPT 自动奖励模型方法，训练相应的自动奖励模型，见图 3，既能提升对医学文献评估的准确性，又可以提升专业人员梳理文献资料的效率。

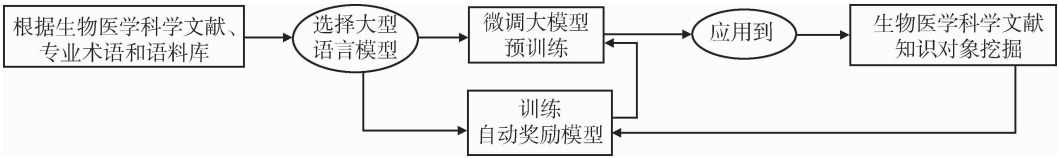


图 3 训练生物医学科技文献自动奖励模型流程

4.2 进一步研究强化学习，提升模型的学习能力

现阶段，人工智能的研究过程中存在 3 种不同研究思路。一种是基于人工神经网络的方法，试图从结构方面模拟人类智能。另一种是基于物理符号系统的方法，从功能方面模拟人类智能。第 3 种是基于感知动作系统的方法，试图从行为方面模拟人类智能^[16]。随着数据信息量不断增长，未来 10 年生成式人工智能的能力和 demand 可能会增长超过 1 000 倍，因此自动学习知识技能已成为人工智能研究的主要方向。ChatGPT 采用强化学习模型，从行为方面模拟人类智能的感知动作系统，为人工智能赋予自主学习能力。进一步分析，ChatGPT 获得自主学习能力和逻辑推理能力，得益于应用强化学习 PPO 模型，特别是马尔科夫链决策过程不断动态

调整前后语义策略，使前后策略更加一致，真实准确生成结果。

目前，国内医学信息量呈指数级增长，但仅能通过关键词和问题检索方式访问医学数据库。在庞大的数据海洋中，精确检索所需的数据信息仍是一项相当困难的任务，缺乏针对性。为此，医学研究可以借鉴 ChatGPT 强化学习 PPO 模型，标注人工检索过程、停留时间等，获得前后概率分布，并以检索时间和对象生成检索模型，为下一步人工检索提供支持。基于此，指数级的医学数据可以整合成强大而智能的知识库，见图 4。



图 4 PPO 模型处理海量医学数据的流程

ChatGPT 也可以成为辅助医生诊断和选择治疗方法的智力支持。目前,在医学研究中,慢性病等病因形成时间较长、较多、较复杂,前后病因的逻辑和相互影响对疾病治疗都非常重要。借鉴 ChatGPT 中 PPO 强化学习模型,应用马尔科夫链决策过程,动态学习前后病因影响大小、不同地域病因的影响关系,强化模型值与人工值差距研究,可以建立有成效的分析和决策系统,为慢性病等病因分析和治疗提供人工智能支持。

强化学习还可以为基因研究和疾病风险评估提供帮助。目前医学研究已经从成千上万患者样本中,生成数以万计的数据集,数据集通常包括突变 DNA 序列、DNA 甲基化、单个基因活性以及细胞通路中蛋白质相互作用信息。近年来的研究表明,蛋白质和基因的相互作用可以映射成一个数学网络,即相互作用网络图,因此也可以借鉴 ChatGPT 中 PPO 强化学习模型,应用马尔科夫链决策过程,在前期生物信息学和人工智能算法结合的基础上,强化泛化场景研究,突出基因序列突变之外的隐性基因研究,为基因研究和疾病风险评估提供全场景的研究方法。

4.3 加强上下文学习能力,提升问答对话系统能力和交互体验

以往医学问答系统存在数据缺失^[17]、简写、拼音代替、错别字^[18]等问题,影响准确性。借鉴 ChatGPT 提问等模式,开发基于 ChatGPT 模型的医学问答系统,系统以高度拟人化对话问答模式,带来较好交互体验,还可以收集整合问题集。此外,医生结合患者病历信息以及回答,给出患者问诊、解答等建议,进一步提升问答对话系统能力和交互体验,同时丰富“答案”集,为后续标注和自动评估机制打下基础。

在医院管理和卫生系统管理方面,ChatGPT 可应用于医院内部的自动化客服和咨询系统,为患者提供在线预约、挂号、查询检查结果等服务,减轻人工接待压力,提高工作效率;医院知识库管理和知识查询,帮助医护人员快速获取医学知识和临床指南,支持决策和诊断过程,提升医疗质量和安全

性;卫生系统公众健康教育,通过智能问答和健康咨询,为大众提供个性化健康建议和预防措施,促进公众健康管理和健康意识。这些应用可以提高服务效率、增强医学知识的传播和应用、促进公众健康意识,为医疗行业的改进和发展作出积极贡献。

借鉴 GPT-N 预训练大模型应用程序接口(application programming interface, API),开放接口研究和对话框提示研究,患者输入个人信息、健康状况信息以及病历信息,拓展信息采集时间宽度和灵活性,对话框提示能提升问答准确性,API 接口能提升医学数据自动化传输水平。同时,医生可自主灵活选择时间,根据患者问题,起草诊断计划,提高治疗效率。根据问题数据集和答案集,再结合脱敏临床知识数据,可进一步提高医学科研效率。

5 结语

ChatGPT 迅速成为现象级焦点,掀起新一轮人工智能浪潮,需要客观分析其流程、模型和创新价值,并探讨其在医学研究中的应用。自动奖励模型可以提高评估准确性和效率,强化学习 PPO 模型突显研究对象的多样性和相互影响,通过应用马尔科夫链决策过程、分析前后逻辑,缩小模型值与人工值的差距,增强特定场景的泛化能力。问答系统可以提升对话能力和交互体验,并拓宽时间范围和数据集。综合考虑 ChatGPT 的训练流程和模型创新点,可以将其应用于医学研究领域,从而开创出更多、更好、更有价值的研究。

参考文献

- 1 郑永和,丁雨楠,郑一,等. ChatGPT 类人工智能催生的多领域变革与挑战(笔谈)[J]. 天津师范大学学报(社会科学版), 2023 (3): 49-63.
- 2 桑基韬,于剑. 从 ChatGPT 看 AI 未来趋势和挑战[J]. 计算机研究与发展, 2023, 60 (6): 1-10.
- 3 VAN DIS E A M, BOLLEN J, ZUIDEMA W, et al. ChatGPT: five priorities for research [J]. Nature, 2023, 614 (7947): 224-226.
- 4 ChatGPT Pro. ChatGPT: The most advanced AI chatbot by OpenAI [EB/OL]. [2023-04-12]. <https://chatgpt.pro/>.

(下转第 29 页)

- in a health care context [J]. *Journal of the academy of marketing science*, 2013, 41 (3): 338–356.
- 24 KUIPERS S J, CRAMM J M, NIEBOER A P. The importance of patient – centered care and co – creation of care for satisfaction with care and physical and social well – being of patients with multi – morbidity in the primary care setting [J]. *BMC health services research*, 2019, 19 (1): 1–9.
- 25 KIM J. Customers’ value co – creation with healthcare service network partners: the moderating effect of consumer vulnerability [J]. *Journal of service theory and practice*, 2019, 29 (3): 309–328.
- 26 LIU W, FAN X, JI R, et al. Perceived community support, users’ interactions, and value co – creation in online health community: the moderating effect of social exclusion [J]. *International journal of environmental research and public health*, 2019, 17 (1): 204.
- 27 刘娜. 互联网环境下患者价值共创对患者满意度的影响 [D]. 天津: 天津大学, 2020.
- 28 OSEI – FRIMPONG K. Patient participatory behaviours in healthcare service delivery: self – determination theory (SDT) perspective [J]. *Journal of service theory and practice*, 2017, 27 (2): 453–474.
- 29 KIM J. The effect of patient participation through physician’s resources on experience and wellbeing [J]. *Sustainability*, 2018, 10 (6): 2102.
- 30 MCCOLL – KENNEDY J R, HOGAN S J, WITTELL L, et al. Cocreative customer practices: effects of health care customer value cocreation practices on well – being [J]. *Journal of business research*, 2017, 70 (1): 55–66.
- 31 DING B, LIU W, TSAI S – B, et al. Effect of patient participation on nurse and patient outcomes in inpatient health-care [J]. *International journal of environmental research and public health*, 2019, 16 (8): 1344.
- 32 BEIRÃO G, PATRÍCIO L, FISK R P. Value cocreation in service ecosystems: investigating health care at the micro, meso, and macro levels [J]. *Journal of service management*, 2017, 28 (2): 227–249.
- 33 YI Y, GONG T. Customer value co – creation behavior: scale development and validation [J]. *Journal of business research*, 2013, 66 (9): 1279–1284.
- 34 RANJAN K R, READ S. Value co – creation: concept and measurement [J]. *Journal of the academy of marketing science*, 2016, 44 (3): 290–315.
- 35 ČAIĆ M, ODEKERKEN – SCHRÖDER G, MAHR D. Service robots: value co – creation and co – destruction in elderly care networks [J]. *Journal of service management*, 2018, 29 (2): 178–205.

(上接第 23 页)

- 5 OpenAI. Introducing ChatGPT [EB/OL]. [2023 – 04 – 12]. <https://openai.com/blog/chatgpt>.
- 6 RADFORD A, NARASIMHAN K, SALIMANS T, et al. Improving language understanding by generative pre – training [EB/OL]. [2023 – 06 – 09]. <https://www.cs.ubc.ca/~amurham01/LING530/papers/radford2018improving.pdf>.
- 7 RADFORD A, WU J, CHILD R, et al. Language models are unsupervised multitask learners [J]. *OpenAI blog*, 2019, 1 (8): 9.
- 8 BROWN T, MANN B, RYDER N, et al. Language models are few – shot learners [J]. *Advances in neural information processing systems*, 2020 (33): 1877–1901.
- 9 OpenAI. OpenAI API [EB/OL]. [2023 – 04 – 12]. <https://platform.openai.com>.
- 10 OUYANG L, WU J, JIANG X, et al. Training language models to follow instructions with human feedback [EB/OL]. [2023 – 04 – 14]. <http://arxiv.org/abs/2203.02155>.
- 11 ZIEGLER D M, STIENNON N, WU J, et al. Fine – tuning language models from human preferences [EB/OL]. [2023 – 04 – 12]. <https://arxiv.org/pdf/1909.08593.pdf>.
- 12 SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms [EB/OL]. [2023 – 04 – 12]. <https://arxiv.org/pdf/1707.06347.pdf>.
- 13 KNOX W B, STONE P. Augmenting reinforcement learning with human feedback [C]. *Bellevue: ICML 2011 Workshop on New Developments in Imitation Learning*, 2011.
- 14 WANG F Y, MIAO Q, LI X, et al. What does ChatGPT say: the DAO from algorithmic intelligence to linguistic intelligence [J]. *IEEE/CAA journal of automatica sinica*, 2023, 10 (3): 575–579.
- 15 钱力, 刘熠, 张智雄, 等. ChatGPT 的技术基础分析 [J]. *数据分析与知识发现*, 2023, 7 (3): 6–15.
- 16 钟义信. 人工智能范式的革命与通用智能理论的创生 [J]. *智能系统学报*, 2021, 16 (4): 792–800.
- 17 李贺, 刘嘉宇, 李世钺, 等. 基于疾病知识图谱的自动问答系统优化研究 [J]. *数据分析与知识发现*, 2021, 5 (5): 115–126.
- 18 孙博, 赵犇, 张飞. 基于自然语言处理的医学智能问答系统设计与实现 [J]. *中国卫生信息管理杂志*, 2021, 18 (6): 738–742.