

# 媒介即模型：“人—ChatGPT”共生自主系统的智能涌现

金庚星

(上海理工大学 网络与新媒体系, 上海 200093)

[摘要]以大型语言模型(LLMs)为代表的生成式人工智能的出现,在物质技术、内容模态、组织制度与媒介性四个面向上对当前的媒介理论构成了挑战。在海德格尔的“媒介即语言”、曼诺维奇的“媒介即软件”的基础上,大型语言模型之后的媒介理论变革可以进一步通过“媒介即模型”来阐明。在人工智能的语境中,“媒介即模型”意味着一切软件都会被接入大模型进行改造。这一观点为媒介理论注入了面向机器学习时代的新解释力。在“媒介即模型”的理论视野下,所有人类行动者都参与到了一个共同信息库的编码之中,并和其他过去的、同时的、或将来的交流行动者共同塑造一个共享的社会信息系统。作为现代社会中最具影响力的知识生产制度之一,媒介系统与社会系统之间的关系范式也将从内容模态对现实世界的知识表征,转向语言模型对现实世界的计算生成。从更一般的社会科学知识论层面来看,“人—模型”之间相互叠加增强的交互方式预示了“人—模型”共生自主系统的崛起,并进一步涌现出群体智能,这是理解生成式人工智能时代社会认识论转型的关键。

[关键词]生成式人工智能;ChatGPT;语言模型;共生自主系统;社会认识论;媒介理论

DOI:10.3969/j.issn.1002-1698.2023.04.007

## 一、问题提出及意义:大型语言模型之后的媒介理论

关于数字革命的叙事,通常会在信息时代(第三次工业革命)和智能时代(第四次工业革命)之间作出分期,前者见证了“摩尔定律”支配下的算力增强,直至行星尺度的云计算出现,<sup>[1]</sup>以及人类社会通过联网沉淀起海量的在线大数据;而后者则强调计算机在尽可能少的人类干预

和控制的情况下,通过自主学习和自适应来完成更加复杂的任务。尤其是当计算机能够像人类一样以自然语言而非代码的方式进行人机沟通并完成任务时。有研究者宣称 GPT-4 已经成为第二个通过图灵测试的聊天机器人,<sup>[2]</sup>也正是在这一点上,无怪乎基于超大规模预训练语言模型的自然语言交互界面 ChatGPT 一经发布,就被赋予人工智能的里程碑式意义,甚至 GPT-4 更是被视为“通用人工智能的早期版本”。<sup>[3]</sup>

作者简介:金庚星,传播学博士,上海理工大学网络与新媒体系讲师,研究方向:通信基础设施与现代中国、媒介理论、数据诠释。

在有关人工智能的技术图景中,计算机主要被训练用来完成两类任务:决策和生成,这分别对应于机器学习的两种模型。<sup>[4]</sup>其中判别模型(discriminative model)主要通过监督式学习来将现有的带标签的数据点进行分类,是在给定观测值的前提下关于目标值的条件概率分布模型。另一种生成模型(generative model)则假设观察变量和目标变量之间遵循联合概率分布,主要通过无监督学习来发掘并抽象出无标签的输入数据的底层模式,然后使用这些知识来生成新的内容。存在两种广泛使用的生成模型:一种是生成性对抗网络(GANs)和变分自编码器(VAEs)这样的深度学习模型,可以从图像和文本输入数据创建出可视化和多模态人工物;另一种是诸如GPT(Generative Pre-trained Transformer)这类基于转换器的生成性预训练语言模型,可以利用海量在线数据来生成新的单模态文本。

新一代生成模型可以用于包括文本、代码、图像、语音、视频等在内的所有单模态或多模态的内容生成,在这个意义上,生成模型已经真正地成为了人工智能语境中各种内容模态的生成媒介,即元媒介(metamedium);<sup>[5]</sup>在此基础上,生成模型甚至还可以用于图像、视频、文本间的跨模态转换,这进一步打破了内容模态之间的边界,成为“终结所有媒体内容(模态)的媒介”。<sup>[6]</sup>这是继海德格尔的“语言作为元媒介”,列维·曼诺维奇(Lev Manovich)的“软件作为元媒介”之后对媒介理论的又一次重大挑战。<sup>[7]</sup>

为了回应这种挑战,本文在生成式人工智能语境下提出“媒介即模型”的理论假设,并结合“媒介转向”<sup>[8]</sup>的各个维度来回答三个相互关联的研究问题:1.当前以ChatGPT为代表的生成式人工智能语言模型的出现将在物质技术、内容模态、组织制度等方面给理解媒介带来何种变化?在何种层面上将更新现有媒介理论的知识?2.媒介是人类交流的场域,媒介领域的这些变化,即大型语言模型成为人工智能时代的基础媒介后,将如何重新界定人类交流的意义?3.既然

媒介在技术和组织面向上都可被视为现代社会中最具影响力的知识生产的基础设施和机构之一,那么模型介入的社会交流,在更一般的社会科学知识论层面引发了知识生产范式及社会认识论的何种变革?

## 二、从“媒介即语言”到“媒介即模型”： 媒介生成主义的理论演变

媒介理论存在两种截然相反的立场:“媒介边缘主义(media marginalism)”和“媒介生成主义(media generativism)”。<sup>[9]</sup>前者假设媒介是中性的透明工具,不会影响其传输的内容,因此相较于其传递的信息或内容居于次要地位。该立场假定尽管客观现实可以藉由媒介获取,但仍然存在一种超越媒介之外的客观现实并不由媒介构成或改变。二战以来的主流大众传播学正是基于这样的认识论假设。然而,自麦克卢汉的“媒介即讯息”向“媒介边缘主义”的核心假设发起挑战以来,这一立场已经遭到了各方的挑战,当前的媒介理论倾向于假设媒介构成并在某种程度上影响了交流过程,这使得媒介自身的物质性被重新放置在聚光灯下。

相反,“媒介生成主义”以媒介作为一种先验条件,主张媒介并非表征现实的中性的透明工具,而是被视为无处不在且先在于我们对世界的体验,它们生成而非仅仅是传递给定的客观现实。与“媒介生成主义”的假设相伴随的则是近年来人文社会科学领域的“媒介转向”。

而当前以大型语言模型为代表的生成式人工智能所引发的媒介变革,意味着媒介的指涉对象已然发生了变化。相应地“媒介生成主义”也需要面向生成式人工智能时代释放出新的解释力,以至于已经有媒介理论家们开始探索“机器学习之后的媒介理论”,甚至提出“媒介理论的终结”。本文在媒介生成主义的脉络下首先回顾了作为元媒介的语言,继而讨论软件如何在元媒介的层面改变了技术媒介的配置,并在二者的基础上进一步阐明模型的生成性(见下页表1)。

表1 “媒介即模型”的理论脉络

(元)媒介理论	人/媒介交互界面	物质性	交流实践	(社会)认识论
媒介即语言	身体	语法、语义、语用	语言的边界即世界的边界	语言转向
媒介即软件	命令行/GUI/浏览器	仿真与发明	对着系统言说	计算转向
媒介即模型	感知/自然语言交互	无监督学习、有监督微调	人机共生自主系统	群体智能转向

(一)媒介即语言

在任何关于媒介的讨论中,语言都是一个重要而持久的话题。语言被认为是人类交流的首要媒介——元媒介,<sup>[10]</sup>海德格尔挑战了我们对媒介的惯常假设,他把语言的传统观念——“我们说话”翻转过来变成了——语言“说话”,来揭开直接经验中隐藏的东西。媒介也以超越单纯表征的方式表达这些东西。

沃尔特·本雅明早就观察到了语言的非工具性,<sup>[11]</sup>认为并非意义藉由语言为工具而传达,而是意义因语言才得以实现。语言是最原型的媒介,甚至是意义得以成立的场所,语言的物质面向,其实正是语言本身。今天将算法描述为逻辑和控制的总和,<sup>[12]</sup>正是将其形式逻辑框架对应于语言的语法层面,而构建算法作为完成任务的一系列精确控制步骤则对应于语用层面的结果。

语言总是呈现自己作为所有过去和未来媒介的生成性元媒介。维特根斯坦有句著名的表述:语言的边界即世界的边界,语言构成了我们对世界的认知和表达。因为它允许说话者使用有限的一组基本元素和规则创建无限个独特而有意义的组合的表达。例如,可以用“主语—动词—宾语”的顺序规则造一个句子,比如“人训练模型”。然而,也可以使用相同的元素来创建一个不同的句子,例如“模型被人训练”。这种生成新句子的能力使语言能够适应不同的语境。

(二)媒介即软件

在语言之外,元媒介对于技术媒介的指涉首先来自于工业媒介的软件化。<sup>[13]</sup>在软件化之前,特定物理、机械和电子媒介的技术属于其“硬件”部分,由两部分组成:用于存储信息的物理材料,以及在材料上刻写、记录、编辑修改并且能呈现和查看信息的器具。材料和器具一起决定了某

个特定的媒介可以做什么。例如,书写(媒介)技术就是由纸张这样的物理材料,以及钢笔或打字机这样的书写器具的特性决定的。

在软件中,仿真物理、机械和电子媒介意味着仿真其工具和界面,而非“材料”。通过软件的仿真,原本特定的媒体创建和交互技术纷纷从其硬件中独立出来,摇身一变成为软件(即独立的算法)。而媒介的各种物理材料在仿真的过程中似乎被淘汰了,软件只处理唯一一种物理材料,即数字化数据。软件化之后,物理材料成为数据结构;物理、机械和电子工具被转换成操作这些数据结构的软件工具(算法)。

如果说特定的媒介效应以前是物理材料属性和工具属性之间相互作用的结果,那么现在则是不同算法修改单一数据结构的结果。软件定义的“媒介”是一种特定的数据结构和用于创建、编辑和查看存储在该数据结构中的内容的算法的配对组合,即(软件)媒介 = 数据结构 + 算法。

(三)媒介即模型

生成式人工智能语言模型汇集了先前两种元媒介,变成了一种更基础的元媒介:一方面将自然语言建模为拥有计算能力的语言模型;另一方面所有的软件都将被接入大模型进行升级改造。<sup>[14]</sup>实际上,软件自身的运作方式也已发生了变化,一个流行的说法是,软件1.0正在吞噬世界,<sup>[15]</sup>而软件2.0(人工智能模型)正在吞噬软件。<sup>[16]</sup>所谓“软件1.0”的运行方式全部来自程序员在软件代码中预先植入的明确规则指令。而在“软件2.0”中,<sup>[17]</sup>其运行规则不再需要程序员的预先植入,而是以人工神经网络的框架被开发出来,从而拥有了从大量交互结果(数据)中学习乃至识别出隐藏的规则和模式的能力,且运行规则在递归的学习过程中动态地修订。

“软件 2.0”意义上的“语言模型”实际上即利用计算机去重新建模人类的自然语言，<sup>[18]</sup>把自然语言转化为机器能够计算的一种机制。人类自然语言一方面有规则，另一方面又随时可以打破规则，并在不同的语境下呈现出不同的语义。大型语言模型实际上并不能真正理解语言，而是通过把整个句子的联合分布概率拆解成其中每个词的条件概率的连乘，大幅降低语言模型的建模难度。这样语言模型的计算生成机制就转换为给定上文预测下文的机器学习任务。

在实际运作中，大型语言模型通过人类的“提示(prompt)”来生成任务，执行任务的方式是将这些任务转化为下一个单词预测的形式。例如，一项名为“把数字经济翻译成英文”的任务，在语言模型中则被转化为“数字经济的英文翻译是……”这样的下一个单词预测的任务。这种生成范式代表了语言模型从专门执行单项任务的领域模型，向可执行多项任务，甚至未明确包含在其训练数据中的新任务的通用模型的转变。

在此意义上，大型语言模型目前还无法说出“属于自己的”语言，它只能说出底层的数据关联。<sup>[19]</sup>

### 三、模型的物质结构与涌现：“提示—响应” 作为一种媒介生成模式

大型语言模型是在“预训练—微调”范式下产生的技术人工物，其依托神经网络与微调机制的物质结构设计通常涉及三个步骤：<sup>[20]</sup>基于神经网络“转换器”框架预训练(pre-training)一个语言模型、收集问答数据并训练一个奖励模型、用基于人类反馈机制的强化学习(reinforcement learning from human feedback)微调(fine-tuning)语言模型以迭代式改进。而模型的生成性运作则是在“提示—响应”交流对话轮中实现的。

#### (一)“预训练—微调”范式下的物质结构

与传统意义上呼应了自柏拉图和笛卡尔以来假设存在先在观念的理性主义，<sup>[21]</sup>将计算理解为“受规则控制的形式符号—转换序列”的符

号人工智能不同的是，<sup>[22]</sup>基于神经网络的转换器框架的人工智能语言模型沿袭了洛克所发展的经验主义假设，相信机器可以通过模仿人脑中神经元之间的连接和信息传递来实现认知和行为的模拟，不需要预先指定具体的规则和知识，而是通过在一个大型无标签文本语料库中进行预训练来自动学习语言的基本模式和结构，并自动调整神经元之间的权重和连接，从而实现输入数据的分类和预测。

预训练过程将自上而下的知识框架与自下而上的数据结合起来。首先，大型语言模型以一套基于语言学理论和知识的预定义规则和结构进行初始化。这些规则和结构为模型提供了关于语言的基本语法和句法的先验知识。然后，模型通过在一个如维基百科或新闻文章这样的大型文本语料库上进行训练，以学习自然语言中的单词和短语之间的统计模式和关系，从而将现实世界的知识和背景纳入其对语言的理解。

一旦预训练过程完成，语言模型就可以使用基于人类反馈机制的强化学习在具体的如文本挖掘、情感分析或语言翻译等下游任务上，持续地进行微调，以适应新的任务场景和数据集。并通过自我学习和改进来提高对话文本的质量，以确保文本的语法正确，且在语义层面有意义。在微调过程中，模型首先与人类进行交互，并生成针对用户输入的回复。在每轮交互结束后，模型将回复展示给人类专家，以便专家对回复质量进行评分和反馈。强化学习算法可以帮助模型收集人类专家的评分和反馈作为强化学习的奖励信号(reward model)，并利用历史对话数据和用户反馈，持续微调模型的对话策略、参数和权重，以更好地适应不同的对话场景和用户需求。

然而，当下游任务的目标与预训练目标差距过大时，用基于人类反馈机制的强化学习来微调并不总是有效。于是在“提示—响应”会话轮中，一种更灵活的使用人工编写的提示词来指导模型进行微调的方法出现了。面向预训练模型的提示能力某种程度上可被视为雷蒙德·威廉斯



所提及的识字能力在人工智能语境中的新变种。

## (二)“提示—响应”交流会话轮中的涌现

总体而言,ChatGPT 语言模型无论在框架还是方法上都与先前的模型没有本质的区别,<sup>[23]</sup>不同之处仅仅在于把预训练数据量和模型参数量作了根本的提升。实际上,ChatGPT 使用的数据涵盖了多种语言和各真实领域。它们可能是维基百科等知识图谱数据、文艺影视对话字幕、社交媒体和在线问答平台的对话记录,以及新闻报道、科技论文、博客等文本数据。而第三方插件功能进一步让 ChatGPT 拥有了“眼睛和耳朵”,可以实时访问全网数据。<sup>[24]</sup>此外,前后四代 GPT 模型的参数量已经呈现出了爆炸式增长,分别是:1.17 亿、15 亿、1752 亿和 100 万亿。正是因为 ChatGPT 拥有了相当大规模的参数量且在足够大的数据集上进行了训练,最终致使其在“提示—响应”会话轮中自发地涌现(emergence)出一些小模型所不具备的能力,<sup>[25]</sup>比如在和用户的对话中呈现出的强大自学能力,以及在一定的提示词引导下展示出的较为复杂的推理能力。据说当大型语言模型在其参数量达到 200 亿左右,计算量大概达到 10 的 22 次方之后,便会完成从量变到质变的飞跃,涌现出这些新能力,<sup>[26]</sup>它们构成了 ChatGPT 的独特可供性(affordances)。

在这些涌现的能力中,最核心的是一种情境学习(in context learning)能力,<sup>[27]</sup>表现为在每一轮人机对话中,ChatGPT 都能将当前对话的历史和上文作为输入,使用转换器模型中的自注意力机制来聚焦对话中最重要的部分,从而理解人类对话的语义和意图,这使模型能够更好地适应不同对话场景和语境,并且生成更准确、自然的回复。比如,“小样本提示”(few-shot prompting)策略仅仅通过给模型提供一些少量的示例样本(shots),以及与任务的上下文信息和领域知识有关的一些关键词、短语、句子或段落来辅助模型进行学习,即无需引入额外的参数,就能引导模型参考输入示例返回关于特定任务的答案。

此外,由于经过海量数据预训练的大型语言

模型已经习得了多种推理方式,所以对于某个复杂的推理问题,用户只需将其分解为一步一步的思维链(chain-of-thought prompting),并输入给模型,就能引导模型做一些相对复杂的推理任务。但是,并非大型语言模型中所有的任务都会发生涌现,并且涌现出来的“小样本提示”能力也是不可预测的,因为这些能力原本就没有包含在预训练任务中。<sup>[28]</sup>所以如何更好地理解“提示—响应”的运作机制对于人机交流至关重要。

## 四、通向“人—模型”共生自主系统： 重塑人类交流

正是因为 ChatGPT 基于多层神经网络与微调机制的物质结构设计使其在一定程度上能够自主地(autonomous)生成响应,<sup>[29]</sup>只是这种自主性受到其在预训练过程中所学习到的模式和规律的限制,也就意味着它不能脱离其所学习的模式和规律产生新的想法或知识。或者说,生成形式是自主的,生成内容是受限的。

在某些情况下,模型还会自主与网络空间的其他模型互动、可从网络空间检索数据、可独立于人类主体进行自己的活动。这意味着拥有学习和适应能力,以及与环境、自身、人类互动能力的模型呈现出了自主系统(autonomous systems)的特征。这种自主性(autonomy)在 AutoGPT 上进一步展现出来,后者是可在不受人干扰的情况下完全自主运行的“智能体”(AI agent)。<sup>[30]</sup>

### (一)“人—模型”共生自主系统

当然,所有这些自主行动最终都会影响人类主体的行为。当这一自主系统与人交互时,不管是基于人类反馈的强化学习机制来看,还是从“提示—响应”的涌现生成模式来看,一种“人—模型”的共生自主系统(symbiotic autonomous systems)已经出现。这不同于人与搜索引擎系统的交互,在这种交互中,机器仅仅是执行人类指令的被动工具,而“人—模型”的共生自主系统是一种相互协作、彼此增强的模式。<sup>[31]</sup>

1. 在预训练过程中,模型高度依赖由人类的

言语行为产生的大型在线数据集来训练,然后再向人类输出响应,对人类的认知产生反馈作用。而不管人类认知在多大程度上受到了模型的影响,只要他们再次说话,那么新产生的自然语言数据将不断地被反馈给模型,成为模型用以训练的新语料,从而进一步影响模型对人类语言的理解和生成。但是,一旦这些预训练的语料库包含偏差或不准确的信息,那么这种共生循环还会导致进一步的认知偏差。

2. 在基于人类反馈机制的强化学习进行微调时,机器通过与人类专家的互动来收集人类反馈,用以指导机器学习算法,不断优化自身的决策和行动,更好地适应人类需求。

3. 在“提示—响应”交流会话轮中,模型只回复不提问,人具有发起交流的主动性。模型通过学习人类输入的“小样本提示”或“思维链”来自我微调,经过多轮人机会话,就可以更好地理解人类意图,生产更加准确的响应。其中决定人机交流会话可持续性的关键因素之一,是用于发起和引导对话的提示词的质量。具体而明确地表明用户的话题和意图的提示有助于语言模型生成符合用户预期的响应。如果提示词过于含糊不清,或充满歧义和误解,模型可能会生成与用户意图不符合的响应,导致“交流的无奈”。

所以这种共生关系不是偏利共生,<sup>[32]</sup>而是一种互利共生,模型和人都会从“提示—响应”会话轮中受益。语言模型意味着能动性 with 智能的解绑。最好的行动者既不是语言模型,也不是人,而是一个熟练地、有见地地使用语言模型的人。<sup>[33]</sup> 机器和人的相互增强重叠进行。通过人机交互的连贯共生,模型表现出自主的集体智能,<sup>[34]</sup> 以至于人工智能和人类的自然智能之间的二分法将不复存在,在这个意义上,模型缓解了智能机器<sup>[35]</sup>与心智之间的潜在对立。<sup>[36]</sup>

## (二)模型“微调”的人类交流

在审视“人—模型”共生自主系统给人类交流带来的变化时,必须考虑到模型的物质性如何在“元交流 (metacommunication)”层面影响人类

交流。“元交流”存在于通过任何媒介开展的所有形式的人类交流与传播活动中,它不仅完成了对于内容的编码 (codification),而且实现了对于社会关系的维系。<sup>[37]</sup> 这可以通过对“人—搜索引擎”与“人—ChatGPT”交流的比较来阐明。

在“人—搜索引擎”的交流关系中,<sup>[38]</sup> 搜索引擎算法一方面为可访问的大量数据点赋予了编码结构,这样用户就能以信息 (或不以信息) 的形式访问这些数据点,并且假设访问的信息具有一定的相关性。编码是一种关于检索结果的元数据,已经在很大程度上塑造了我们按照特定的交流行为模式,比如搜索、归档、排序和过滤逻辑,来进行思考、内化和行动。<sup>[39]</sup> 另一方面,搜索引擎还使交流关系的建立得以可能。用户不一定是明确地与那些具名的个体或机构,而是与分布式的信息资源建立了一种交流关系,而后者可能会引发更多后续步骤的交流和互动。更重要的是,用户在向搜索引擎系统提供输入时,实际上重新配置了这一交流系统,无论程度多么微小。

与“人—搜索引擎”的交流方式相比,“人—ChatGPT”的交流出现了新的变化。在“人—搜索引擎”的交流过程中,“关键词”的输入指令仅仅被用来调用算法规则以匹配搜索请求与网页内容,不会改变搜索系统本身的存在状况;但在“人—ChatGPT”的交流中,“提示词”的输入在语义层面无时无刻不在引导 ChatGPT 语言模型本身进行“微调”,模型与人类行动者交叠增强,在这种交流关系中出现的是共生能动性 (symbiotic agency)。<sup>[40]</sup> 在这个意义上,提示成为一种 ChatGPT 介入的人类交流的元语言——既微调语言模型本身的配置,又在一定程度上增强 (同时也是限定) 模型介入的人类交流能力。在这样的“元交流”中,人类交流与人机交流之间的边界不再那么泾渭分明,所有人类行动者都参与到了一个共同信息库的编码之中,并和其他过去的、同时的,或将来的交流行动者共同塑造一个共享的社会信息系统。这个递归式交流系统也模糊了传播与传承的区别,进一步赋予媒介学 (mediology)

在生成式人工智能时代的基础性地位。<sup>[41]</sup>

### 五、结语：群体智能——大型语言模型驱动的新型社会认识论

假如我们接受媒介学的相关主张，将媒介（物）视为知识生产、传播和传承的基础物质材料，那么媒介（物）指涉对象在生成式人工智能时代的变化，不但会重新界定人类交流的意义，还将在一定程度上构成对认识论的冲击，人文社会科学的“媒介转向”也藉此更具范式更替的意义。

“媒介即语言”可被视为前数字时代的一种认识论基础，并在20世纪后半叶哲学和社会科学领域的“语言转向”（linguistic turn）中大放异彩。<sup>[42]</sup>这一转向不再接受语言是一种表征现实的透明工具，而是相信语言是主客体之间的中介，相应地，语言不是对现实的直接反映，而是一种对现实的阐释和构建，是社会交往、权力斗争和文化认同的场域。这一转向将知识视为特定历史、文化和语境下的相对不确定的“叙事”。

在曼诺维奇关于“媒介即软件”的论述中，软件之后的媒介的永久可扩展性对“计算机化社会”中的知识技术产生了重要影响。<sup>[43]</sup>（软件）媒介的逻辑是把一切都变成数据，并使用算法来分析这些数据，这已经改变了认识事物的意义，并导向一种新的软件认识论。<sup>[44]</sup>认识论提出的问题包括什么是知识，如何获得知识，以及我们对于一个主题的认知可以达到何种程度。计算机以“元媒介”之姿引入了获取知识的新方式，并在此过程中重新定义何谓知识。例如，总是有可能发明新的算法或者扩展现有的算法来分析已经存在的旧数据，从中发掘以前的算法未能发掘的新知识。当今整个艺术人文领域的“数字人文”和社会科学领域的“计算转向”正是立足于计算媒介这一物质基础之上。

不管是“媒介即语言”，还是“媒介即软件”，所沿袭的认识论传统更多地都是关注个体如何通过感官、记忆和推理等方式获取知识，大型语言模型的“提示—响应”的运作方式意味着人类

行动者都参与到了一个共同信息库的编码之中，知识的生产有赖于他人的参与，不管参与者本人是否承认这一点。因此，当大型语言模型介入社会交流后，人机共生的交互与生成模式挑战了传统认识论的个体预设，引发了认识论向社会认识论的转变。社会认识论将认识论的范围从主体所经历的认知过程扩展到了其所处的直接、远程、当前和过去的社会环境中他人的社会活动。<sup>[45]</sup>这种社会认识论建立在分层有序的深度神经网络架构的物质基础之上，并跨越语言、国家、地域、性别等边界，通过无时无刻不在进行的“提示—响应”会话轮，进行递归式的知识的集体协作生产、分享和修订，体现为一种群体智能（swarm intelligence），而模型成为协调群体智能的基础媒介。当现实只能通过涉及模型转换的社会认识论来体验时，模型的媒介性将成为人类存在的基本状况，这也再次回到了“媒介转向”的理论假设中。在此意义上，（模型）媒介真正将生成式人工智能时代的人类联结成了命运共同体。

#### 注释：

[1] Bratton, B. H., *The Stack: On Software and Sovereignty*, Cambridge, MA: MIT Press, 2015, pp. 109 - 146.

[2] 2022年12月，BuzzFeed的数据科学家马克斯·伍尔夫（Max Woolf）在twitter上表示，人工智能聊天机器人ChatGPT成为第二个通过图灵测试的聊天机器人。

[3] Bubeck, S., et al., "Sparks of Artificial General Intelligence: Early Experiments with GPT-4", <https://arxiv.org/abs/2303.12712>.

[4] 周志华：《机器学习》，北京：清华大学出版社，2016年，第2-12页。

[5][37][38][丹麦]克劳斯·布鲁恩·廷森：《如何与数据打交道——元数据、元媒介和元传播》，刘君等译，《中国传媒报告》2014年第4期。

[6] 英文原文为“medium to end all media”，参见 Winthrop - Young, Geoffrey., *Kittler and the Media*, Cambridge: Polity, 2011, p. 131.

[7] Hepp, A., "Artificial Companions, Social Bots and Work Bots: Communicative Robots as Research Objects of Media and Communication Studies", *Media, Culture & Society*, 2020, 42 (7 - 8), pp. 1410 - 1426.

[8] “media (1) turn”参见 Weber, Stefan, "Medial Turn; die Medialisierung der Welt", *Medien Journal: Zeitschrift für Kommunikationskultur*, 1999, 23(1), pp. 3 - 8; 另见黄旦主编：《媒介转向》，



《中国传播学评论(第八辑)》,北京:中国传媒大学出版社,2019年,第1-5页。

[9] Krämer, Sybille. , *Medium, Messenger, Transmission: An Approach to Media Philosophy*, Amsterdam University Press, 2015, pp. 27-38.

[10][美]戴维·J·贡克尔、[英]保罗·A·泰勒:《海德格尔论媒介》,吴江译,北京:中国传媒大学出版社,2019年,第29-37页。

[11][日]吉见俊哉:《媒介文化论——给媒介学习者的15讲》,苏硕斌译,台北:群学出版社,2009年,第7页。

[12] Fuller, M. , *Software Studies: A Lexicon*, Mit Press, 2008, pp. 15-20.

[13][44] Manovich, L. , *Software Takes Command: Extending the Language of New Media*, London: Bloomsbury, 2013, pp. 207, 337.

[14]《阿里巴巴张勇:“所有软件都值得接入大模型升级改造”》,多知网,http://www.duozhi.com/industry/insight/2023041115076.shtml。

[15] 曼诺维奇所指的“媒介即软件”对应的是软件1.0。

[16] Andrej Karpathy , *Software 2.0*, medium, https://karpathy.medium.com/software-2-0-a64152b37c35.

[17] Dilhara, M. , Ketkar, A. , & Dig, D. , “Understanding Software-2.0: A Study of Machine Learning Library Usage And Evolution”, *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 2021, 30(4), pp. 1-42.

[18][26]《复旦MOSS大模型拟4月中旬开源,邱锡鹏详解如何构建》,澎湃新闻,https://m.thepaper.cn/wifiKey-detail.jsp?contid=22626797&from=wifiKey#。

[19] 赵汀阳:《GPT推进哲学问题了吗?》,《探索与争鸣》2023年第3期。

[20] Lambert, Nathan, et al. , “Illustrating Reinforcement Learning from Human Feedback (RLHF)”, huggingface. co, Retrieved 20230304.

[21] Crockett, L. J. , “‘All AI’s Are Psychopaths’? Reckoning and Judgment in the Quest for Genuine AI”, 2nd Annual European Conference on the Impact of Artificial Intelligence and Robotics, Instituto Universitario Lisboa, Portugal, Oct. 21-23, 2020, published in Proceedings.

[22] Denning, P. J. , “Opening Statement: What is Computation?”, *The Computer Journal*, 2012, 55(7), pp. 805-810; Fortnow, L. , “The Enduring Legacy of the Turing Machine”, *The Computer Journal*, 2012, 55(7), pp. 830-831.

[23]《吴军:ChatGPT七问七答》,Redian 新闻,https://redian.news/wxnews/344259。

[24]《ChatGPT插件重磅上线:不只“联网”,还将打造强大AI生态》,界面新闻,https://m.jiemian.com/article/9133930.html。

[25] Wei, J. , et al. , “Emergent Abilities of Large Language Models”, https://arxiv.org/abs/2206.07682.

[27] 张俊林:《大语言模型的涌现能力:现象与解释》,知乎,https://zhuanlan.zhihu.com/p/621438653。

[28] 张江等:《AI何以涌现:复杂适应系统视角的ChatGPT和大语言模型》,搜狐网,https://m.sohu.com/a/652716004\_121124373/? pvid=000115\_3w\_a。

[29][美]兰登·温纳:《自主性技术——作为政治思想主题的失控技术》,杨海燕译,北京:北京大学出版社,2014年,第25-36页。

[30] Kevin Jiang, “What’s Auto-GPT? New, autonomous ‘AI agents’ can act on their own, rewrite their own code”, https://www.thestar.com/business/technology/2023/04/14/whats-auto-gpt-new-autonomous-ai-agents-can-act-on-their-own-rewrite-their-own-code.html? utm\_source=share-bar&utm\_medium=user&utm\_campaign=user-share.

[31] S Shyam Sundar, “Rise of Machine Agency: A Framework for Studying the Psychology of Human-AI Interaction (HAI)”, *Journal of Computer-Mediated Communication*, 2020, 25(1), pp. 74-88.

[32] 偏利共生(commensalism)意味着人机互动仅仅对人有益,对孳生体没有影响。而互利共生(mutualism)意味着人机互动中孳生体和人都从中获得增强。

[33] Floridi, L. , “AI as Agency Without Intelligence: On ChatGPT, Large Language Models, And Other Generative Models”, *Philosophy and Technology*, 2023, 36(1), p. 15.

[34] Wang, et al. , “On the Philosophical, Cognitive and Mathematical Foundations of Symbiotic Autonomous Systems”, *Philosophical Transactions of the Royal Society A*, 2021, 379(2207), 20200362.

[35] 既表示物理机器,如robot,也表示软件程序,如Chatbot。

[36] Saracco, R. , Grise, K. and Martinez, T. , “The Winding Path Towards Symbiotic Autonomous Systems”, *Philosophical Transactions of the Royal Society A*, 2021, 379(2207), 20200361.

[39] Andersen, J. , “Archiving, Ordering, and Searching: Search Engines, Algorithms, Databases, and Deep Mediatization”, *Media, Culture & Society*, 2018, 40(8), pp. 1135-1150.

[40] Papacharissi, Z. , *A Networked Self And Human Augmentics, Artificial Intelligence, Sentience*, Routledge, 2018, pp. 97-106.

[41][法]雷吉斯·德布雷:《媒介学引论》,刘文玲、陈卫星译,北京:中国传媒大学出版社,2014年,第1-32页。

[42] 陈新:《真之效用:哲学与史学中的“语言转向”》,《社会科学战线》2018年第7期。

[43][法]让-弗朗索瓦·利奥塔尔:《后现代状态——关于知识的报告》,车槿山译,南京:南京大学出版社,2011年,第1-17页。

[45] 关于菲利普·基切耳(Philip Kitcher)的观点,转引自 Godler, Y. , Reich, Z. & Miller, B. , “Social Epistemology As A New Paradigm For Journalism and Media Studies”, *New Media & Society*, 2020, 22(2), pp. 213-229.

[责任编辑:刘毅]