

## “诸神之战”：新工具抑或新存在

王金林，复旦大学哲学学院教授

ChatGPT 在整个 AI 行业乃至社会各界引发了一场“诸神之战”。OpenAI 公司作为 ChatGPT 的开发者，在介绍它时只是低调地把它描述为一种对话式人工智能：“它能够以对话方式进行交互。对话格式使 ChatGPT 能够回答后续问题、承认错误、挑战不正确的前提并拒绝不适当的请求。”尽管这种对话式人工智能看似平常，但发布后却迅速成为一种现象级应用，并牵动各界展开激烈的争论。

英伟达 CEO 黄仁勋赞誉 ChatGPT 为 AI 领域的“iPhone 时刻”。基辛格与 AI 业内巨头等联袂撰文，断言“ChatGPT 预示着一场智能革命”，为人类理性开辟了革命性的途径，同时也提出了自启蒙运动以来规模空前的哲学和实践挑战。他们希望通过“确保机器仍然是从属于人的对象”来重申人性。语言学大家乔姆斯基等人则持完全否定的观点，直斥大语言模型的非道德性、伪科学性和语言无能。他们认为 ChatGPT 只能描述和预测，完全缺乏解释能力，而解释才是“真正智慧的标志”。ChatGPT 要么“过度生成”（即既产生真理也产生谬误，既认可道德的决定也认可不道德的决定），要么“生成不足”（即对任何决定的不承诺和对后果的漠不关心）。深度学习巨头 Yann LeCun 等人也认为 AI 聊天机器人不了解人类诚实和体面的规范，并不以人类的方式进行对话，因此最多只是“非常有趣的工具”，不应将其“拟人化”。

倘若将上述评论拿去问 ChatGPT，其回复会既谦和又辩证：“我”只是一种大语言模型，没有自己的价值偏向与观点，但是“我”可以在若干方面对人们有所帮助，不过“我”也会产生事实性错误即所谓的“幻觉”，因此不能完全依赖“我”的回复来做决策，而必须进行核查，等等。ChatGPT 的回答当然不无道理。然而，这只是一种工具论意义上的道理。如果完全接受这一观点，那就误入“机关”了。就其本质和逻辑而言，ChatGPT 及其迭代产品 GPT-4 不仅是一种新工

具，而且预示着一一种新的存在。人类社会或将因此发生重大变革。这个问题事关重大，涉及广泛且深刻。受篇幅限制，本文仅从知识生产方式、语言生成方式以及人的对象性存在方式三个方面进行论述。



首先，从人类知识的生产方式来看，ChatGPT 带来的是变革。知识之于人至关重要。亚里士多德在《形而上学》开篇即言“求知乃人之本性”，培根更是断言“知识就是力量”。不过人们往往只从认识论或知识论的角度来理解这些洞见，而忽视了存在论或哲学人类学的观点。这导致知识与人性被误解为二元对立的观念。知识似乎仅仅被视为人类认识和改变世界的产物，无关乎人的本质规定。这显然是一种工具论的知识观，它根本不懂知识对人性意味着什么，也不懂知识生产方式实际上也是人性建构方式。知识与人性存在着本质的关联。人类生产的知识以及知识的生产方式塑造了人的本质。

GPT 模型的当前表现证明生成式人工智能已经具备了知识生产的能力，并且其生产方式与人类不同。GPT 模型的开发思路是让机器具备一种不受监督的学习能力，以使用自然语言模型培训所



有内容。训练 ChatGPT 的核心逻辑在于训练神经网络体系，让它去预测下一个单词，发现现有数据的逻辑，从而去生成新的知识。有人也许会反驳说：这只不过是基于统计学对现有文本或知识进行新的排列组合而已。从表象上看，这种观点当然有道理。然而，这种新的排列组合实际上意味着一种新的知识生产方式。自然语言处理和机器学习技术以远超人类的速度更快地处理和分析海量的数据和信息，从而能够发现新的知识和模式。ChatGPT 在回复笔者的提问时，简要概括其工作原理为五个步骤，即输入信息、文本解析、知识检索、知识生成和输出结果。可见，知识生成是其核心功能。当然，它也提醒我们，其知识生成能力取决于其所学习的数据和算法，因此所生成的知识可能不一定准确。

然而，问题并不在于 ChatGPT 生成的知识是否准确可靠（它的可靠性的提升只是时间问题，GPT-4 的发布已经证明了这一点），而在于它不仅能够完全理解人类的提示，而且它的回复呈现出类似于人类知识的样式。在进行了数百亿或数千亿参数级别的大规模语言模型训练之后，它甚至出现了一种其开发者无法完全理解的“涌现”现象。这就是业内所称的“狂卷语料库，大力出奇迹”。基辛格等人断言：“一项新技术正在改变自印刷术发明以来从未被动摇过的人类认知过程”，必将“重新定义人类知识”。倘若此论不虚，那就意味着人类的知识生产方式正在转变为人机合成的方式。在合成生物之后，现在出现了合成知识。这可能造成什么样的结果尚难预料，但可以肯定的是，随着人工生成知识的出现，“知情意”——规定人性的三大要素——中的知识及其生产

方式正在经历颠覆性的变化。一种新的存在正在形成，而 GPT 模型则是其前兆。

其次，从人类语言生成的方式来看 ChatGPT 带来的语言比知识更为根本，对于规定人性而言具有重要意义。然而，语言与知识一样，常常被视为一种仅仅用于交流的工具。尽管这一观点有其日常经验的支持，但它仍属于工具论的表面观点。亚里士多德将人直接定义为“逻各斯的动物”，正是因为没有语言或逻各斯，就不会有所谓的人了。维特根斯坦对此的理解是：想象一种语言，就是在想象一种“生命形式”。海德格尔更是把语言视为“存在之家”。没有语言，人生在世这件事根本不可能。语言并非一种可供人任意使用的交通工具，而是对人的本质的基本规定，因此他甚至认为并非人在说话，而是语言在说话，人说只不过是应合着语言而说。

ChatGPT 等 GPT 模型具有语言能力并开始生成新的语言。其基本机制是通过分析大量的语言数据来学习语言的结构和规则，从而生成新的语言表达方式。这固然使人机交互变得更加智能化，但同时也意味着人与语言关系的异化，语言开始成为人机混合物。机器语言逐渐进入人类的“自然语言池”中。这是否会像被编辑的基因污染人类基因池那样污染人类的语言池呢？或许有人会认为，无论机器多么精通语言，它所表达的仍然是自然语言，只是在自然语言的现有框架内进行“嬉戏”而已。笔者对此并不认可，ChatGPT 所生成的语句、短语或词汇正在被人们采纳并被纳入日常使用，这在一定程度上可能会影响自然语言的句法和语法。更重要的是，正如 ChatGPT 的回复所言，“智能体的出现意味着语言将不再仅仅是人类的专属领域”，并且“从更广义的角度上看，语言的创造者可以是任何具有能力创造和传播符号系统的智能体”。

我们当然知道目前 GPT 模型仍然可视为基于程序和算法的工具，其所谓的“创造性”是建立在对大量语言数据的学习和生成之上，而不是基于自主意识和意愿的创造。对话式人工智能实现了人机之间的自然语言交互，这无疑极大地降低了人机交互的门槛，提高了交互效率，从而倍增了生产力。但是无可否认的是，自然语言在此过程中显然被智能化了，其效率或准确性也许更高了，但自然语言的复杂性与丰富性却面临

丧失的风险。语言作为逻各斯、“生命形式”与“存在之家”已经面目全非了，一种新的机器语言正在形成，或者说人类语言池已经不再“自然”了，而语言池的污染显然会悄然地却又实质性地影响人与存在的关联。

在 GPT 模型出现之前，人类技术产品中最可怕的莫过于核武器，但此“大杀器”终究只属于器，与源于存在之根的大语言模型（或可称之为“大语器”）无法相提并论。大语器完全有可能越器而出，不复为器。如果说大杀器因其可怕的毁灭力量让人警醒的话，那么大语器则完全可以在潜移默化中对人性进行重构，实现“润人细无声”。

最后，从人的对象性存在方式来看 ChatGPT 带来的变革。人是一种对象性存在，也就是说人总是要把其本质力量对象化，并由此返回或反观自身。在中世纪，人将其本质力量投射在上帝身上，视其为全知全能全善的造物主，而人则同其他一切存在者一样成为被造物。近代以来，人开始充分肯定此岸世界的力量，把自己而非神确认为中心，重新占有了曾经被神所攫取的人的内在力量。这场人神之战被尼采概括为“我们杀死了上帝”。因为神终究只是一种观念性存在。

现在，一些原本属于人的能力和属性，如思考、决策、感知等，被投射到了人工智能身上。GPT 模型就是这种投射的最新成果。一场新的人神大战或许已经开始。然而，这一次的“神”不再是观念性的存在，而是以工具形式介入世界的对话式 AI，其存在完全可以充满实在性。那么，人类能否如法炮制，再次战胜自己所造之“神”，重新占有自己的本质力量呢？黑格尔的主奴辩证法的悲喜剧（是悲是喜取决于立场）是否会在人工智能领域重新上演，演变为“人机辩证法”？

必须强调的是，GPT 模型所预示的新存在并非诸多存在者之一种，而是马克思所述的“特殊的以太”“普照的光”。一切存在者都将在其普照之下，重新证明自己的存在合法性。人、社会与自然均会被重新规定，人类社会或许会从此逐渐步入“人机社会”。ChatGPT 的开发者们对此并不担忧，他们反复强调人工智能和通用人工智能（AGI）毕竟是由人类构建的工具和系统，其行为和表现也是由人类的目标和规则所决定的。虽然它们可以模拟某些活动，但并不具备真正的自主性和创造

力，而人类则具备智能体所缺乏的特征和优势，如情感、道德、创造性和自我意识等。姑且不论智能体以后能否在这些领域形成超过人类的能力，一种可能的情形是这些人类的独特优势在智能体时代将被边缘化，进而被消灭，就像工业革命初期机器纺纱碾压手工纺纱那样，关键在于主导性原则是计算性思维。

“ChatGPT 之父”奥特曼对更加强大的 AGI 的开发亦不无犹豫。一方面，他强调要确保 AGI 在“提升人类”方面造福于全人类；另一方面，他也承认这种“力量倍增器”存在巨大的潜在风险。他甚至表示，成功地过渡到一个拥有超级智能的世界可能是人类历史上“最重要、最有希望、最可怕的项目”。为此他提出了 AGI 三原则，即“赋能、共享与谨慎”，主张逐渐过渡，让整个社会有时间探究发生了什么，并制定应对策略，使 AGI 的发展方向“对齐”人类的价值与目标。马斯克等人呼吁暂停巨型 AI 实验的公开信也表达了同样的关切。

然而，遗憾的是，这一切努力显然仍然囿于工具论的范畴，仍然是在技术性规定这个大范式之下进行的。无论是逐渐过渡，还是暂停试验，都或许可以让人们从表象上对智能体有所认知（但对其本质与意义何在却仍然盲然不知），更不可能真正逆转新存在的到来。相反，AGI 却可能日益“了解”人类。人类几乎全部经验、知识和智慧都蕴藏在语言中，现在 GPT 模型已经学会使用语言，人类似乎已无秘密可言。

[ 本文系教育部人文社会科学重点研究基地重大项目“后现代的宏大叙事——《帝国》三部曲批判”（15JJD710003）的阶段性的成果。]