

中宣部首批“哲学社会科学期刊重点专栏”：元宇宙与数字经济研究（四）

从 ChatGPT 到 Sora： 生成逻辑、哲学本质及世界图景

黄欣荣^{1, 2}

(1. 浙江科技大学 马克思主义学院, 浙江 杭州 230031; 2. 江西财经大学 马克思主义学院, 江西 南昌 330013)

摘要: ChatGPT 和 Sora 的出现是人工智能发展史上的两件轰动性事件, 共同展现了基于大模型的生成式、通用性人工智能的强大功能。ChatGPT 和 Sora 在生成逻辑、哲学本质及世界图景上既有共性和继承, 又有突破和创新。从生成逻辑看, 二者从演绎推理式转向学习生成式, 均基于大模型的学习生成, 是生成式人工智能的杰出代表, ChatGPT 主要生成静态文本和图像, Sora 能够生成复杂的动态视频。从哲学本质看, 二者均涉及学习、理解、意识和思维等哲学本质问题, ChatGPT 主要是因果、语义、语境的理解和生成, Sora 能够理解时间和空间等复杂关系并生成动态时空视频。从世界图景看, 二者都可能带来创新创造的广阔未来, ChatGPT 主要生成静态的语义世界, Sora 则带来更加复杂的元宇宙世界, 甚至能够重演创世造物的动态过程。

关键词: Sora; ChatGPT; 生成式人工智能; 生成逻辑; 哲学本质; 世界图景

中图分类号: G206; B15 **文献标识码:** A **文章编号:** 1005-9245 (2024) 06-0072-09

Sora 是美国 OpenAI 继 ChatGPT 后发布的再次震惊世界的生成式人工智能产品, 其能够根据文本描述生成高清、连贯、逼真的一镜到底视频, 彰显了其对现实世界的深刻理解和模拟能力^①。生成式人工智能何以在短时间内连续震撼世界? ChatGPT 已足够强大, 吸引广泛关注, 为何 Sora 的出场仍如此惊艳? 二者存在何种继承与关联? Sora 实现了哪些突破和创新? 从 ChatGPT 到 Sora, 在人工智能发展史上意味着什么? 笔者试图基于科技哲学的视角和方法, 从生成逻辑、哲学本质和世界图景三个维度分析 ChatGPT 与 Sora 间的关联及突破。

一、生成式人工智能：从 ChatGPT 到 Sora 的轰动效应

自 1956 年达特茅斯会议正式提出人工智能以来, 其经历了三次高潮和两次低谷。人工智能在 21 世纪的前 10 年默默无闻, 自 2016 年 AlphaGo 战胜世界围棋冠军李世石后, 人工智能开始走出低谷, 进入激动人心的时代。在 AlphaGo 余温未消的 2022 年 11 月 30 日, OpenAI 推出 ChatGPT, 人们震惊于其强大的功能, 注册用户数在 2 天内即达百万^②。当人们热议 ChatGPT 的强大功能并展开深入研究时, OpenAI 于 2024 年 2 月 16 日强

收稿日期: 2024-03-19

基金项目: 本文系国家社科基金重大项目“负责任的人工智能及其实践的哲学研究”(21&ZD063)的阶段性成果。

作者简介: 黄欣荣, 浙江科技大学马克思主义学院特聘教授, 江西财经大学马克思主义学院教授、博士生导师。

① 令小雄、王鼎民、唐铭悦:《ChatGPT到Sora:Sora文生视频大模型对影视创作的机遇、风险及矫治》,《新疆师范大学学报(哲学社会科学版)》, <https://doi.org/10.14100/j.cnki.65-1039/g4.20240318.001>。

② 朱光辉、王喜文:《ChatGPT的运行模式、关键技术及未来图景》,《新疆师范大学学报(哲学社会科学版)》, 2023年第4期。

势推出 Sora^①，业界甚至发出“再见，人类”的哀叹^②。随着计算机、人工智能等新技术的发展日新月异，人们惊叹“各领风骚数十年”。从元宇宙的再次爆火到 ChatGPT，再到 Sora，仅相隔 1 年时间，发展速度之快令人应接不暇。人们还来不及消化元宇宙，ChatGPT 就踏浪而来；人们还未弄清 ChatGPT，Sora 又强势来袭。现在只能用“各领风骚仅一年”对这种后浪速推前浪的“智能爆炸”现象进行刻画。

（一）ChatGPT 是聊天式文本生成器

OpenAI 发布的聊天机器人模型 ChatGPT，意为预训练生成式聊天转换器^③。首先，ChatGPT 是人机沟通的转换接口。人类的自然语言只有持相同语言的人才能理解，智能机器只能用二进制的比特语言传递信号，因此，不同语种之间、人机之间只有通过“翻译”转换才能进行交流，ChatGPT 就是这样一种通用转换器。其次，ChatGPT 主要基于大规模预训练语言模型。ChatGPT 能够翻译转换、理解并重组和创新人类的自然语言，主要基于海量语料库的训练。ChatGPT 能够从人类自然语言中找到最基本的语言构成要素和语法规则，基于涌现生成，按照人类的意图进行生成创新。最后，ChatGPT 与人类沟通交流的方式是聊天。聊天是人类最便捷、最简单的交流方式，ChatGPT 即通过聊天的方式与人类沟通交流。ChatGPT 让人类无需掌握二进制机器码就能与机器沟通交流。以往的人类需要适应机器并学会机器语言，现阶段，ChatGPT 的出现使机器能够适应人类，即 ChatGPT 可以用人类的文字或语音与人类沟通交流。

（二）Sora 是图像世界的视频模拟器

Sora 是 OpenAI 推出的生成式人工智能产品。ChatGPT 的名称出自英文缩写，Sora 是隐喻式命名。有观点认为，Sora 来自日语そら，意为空^④，不仅隐喻空间、时空，而且包括生长空间，意为发展机遇、充满希望。如果说 ChatGPT 像一粒种子，能够涌现生长，那么，Sora 将为其提供空间和表现机会。

首先，OpenAI 将 Sora 作为视频生成器推出。ChatGPT 解决了文字生成问题，可根据需要生成带有意义的文字或图片；Sora 能够根据人类的需要或意图生成富有想象力和创造力的、带有时间坐标的动态视频，解决了图像生成问题。因此，从技术层面看，Sora 是视频生成器。其次，可将 Sora 理解为世界模拟器。Sora 不仅能够生成图像和视频，其终极目的是模拟世界图景。元宇宙提出孪生世界和创构世界的理想，但因缺少技术支持尚未付诸实践。Sora 能够模拟现实世界，生成真假难辨的世界图景，既解决了世界孪生问题，又解决了世界创构问题。最后，Sora 可以理解和创建时空。ChatGPT 解决了自然语言理解、意义涌现和因果关系问题，但真实世界万物的存在必然通过更加形象、沉浸的时间和空间进行呈现。Sora 在 ChatGPT 的基础上增添了时空要素，不仅为万物的涌现生成提供了条件和机遇，而且为之奠定了坚实基础。

（三）Sora 是对 ChatGPT 的继承和创新

ChatGPT 能够轰动世界，主要基于两大创新：一是生成，二是通用。20 世纪的人工智能研究让人工智能具备推理演绎能力，21 世纪的 ChatGPT 让人工智能具备经验归纳能力，在此基础上，人工智能具备归纳和演绎的双螺旋创新能力，并形成“经验—学习—知识—推理—创构”的人工智能学习和创新的闭环系统，开辟了人工智能自我组织、自我学习、自我创新的自组织发展之路。由此可见，ChatGPT 是人工智能从推理演绎式走向归纳生成式的转折点。Sora 引起轰动的原因在于其具备文生图、文生视频的能力，使归纳生成从自然语言推广到图片和视频领域。

以往的人工智能虽然功能强大，但只具备某种特定功能，只能完成特定任务。以 ChatGPT 为代表的生成式人工智能已经具备通用能力，只要把任务交给生成式人工智能，其他一切都可以由其自动完成，标志着人工智能已从专用走向通用。例如，ChatGPT 可以完成编程、文秘、论文写作、文学创作等任务。Sora 在 ChatGPT 的基础上实现新飞跃，

① 《OpenAI 发布首个视频生成模型：输文字出视频，1 分钟流畅高清》，https://www.thepaper.cn/newsDetail_forward_2637061。

② Raphaël Millière. Are video generation models world simulators? <https://artificialcognition.net/posts/video-generation-world-simulators/#concluding-thoughts>.

③ 成生辉：《ChatGPT：智能对话开创新时代》，北京：中信出版社，2023 年版，第 3 页。

④ Video generation models as world simulators, <https://openai.com/research/video-generation-models-as-world-simulators>.

从生成视频走向创造世界。因此，未来的生成式人工智能可能至大无外、至小无内，带给人类无穷的想象空间。

二、生成逻辑：从 ChatGPT 到 Sora 的延展

20 世纪的人工智能因缺少经验数据的采集、存储和处理等能力，虽然发展过程曲折，但解决了机器的推理演绎能力问题，使智能机器能够像人一样进行逻辑演绎和推理。但是，这一时期人工智能演绎的大前提及各种理论仍需人类输入“喂养”，带有明显的人工性。进入 21 世纪，随着数据技术的发展，智能感知、数据挖掘、数据算法、云存储等技术迅速兴起。大数据时代的到来，使数据采集、存储、处理和传输等问题迎刃而解，人工智能已具备经验归纳能力。ChatGPT 与 Sora 在生成机理上遵循共同的逻辑，因此，OpenAI 在 ChatGPT 成功之后迅速推出 Sora，并取得重大突破。

（一）ChatGPT 和 Sora 在生成逻辑上的共通性

ChatGPT 和 Sora 在技术路径上遵循的生成逻辑相似，即大语言模型、涌现生成、通用智能三大技术逻辑。换言之，在技术逻辑上，ChatGPT 和 Sora 都是在大模型的基础上进行机器的深度学习，由此找到涌现生成所需的积木模块、生成规则、迭代算法和优化路径。

第一，大语言模型。人类之所以拥有智慧，得益于向经验学习并不断积累经验。这种经验包括自身的直接经验和他人的间接经验。人类自诞生至今，积累了巨大的经验宝库。生成式人工智能一改以往依靠的推理演绎路径，转向经验学习，踏上归纳创新之路。人类的经验知识已经通过语言的形式储存在各类文本中，ChatGPT 和 Sora 通过汇聚人类的各种经验，形成人类语言、图像、影音等大模型，使之成为 ChatGPT 和 Sora 学习的基础，即 ChatGPT 和 Sora 找到了生成要素、生成模型、生成规律的基础。因此，基于大数据技术的兴起和大数据时代的到来，人类的经验被数字化并形成大模型，为 ChatGPT 和 Sora 等生成式人工智能奠定了经验基础。

第二，涌现生成。ChatGPT 擅长语言文字生成，Sora 擅长图片视频生成，但二者在技术本

质上均基于大模型的涌现生成过程。ChatGPT 和 Sora 如同人类一般学习、思维和创新，从大模型中找到用于推理演绎所需的经验要素和理论前提。从技术层面看，无论 ChatGPT 还是 Sora，都在应用机器深度学习的基础上实现基于大模型的学习。ChatGPT 和 Sora 通过大模型的学习，获取人类学习和创新的奥秘，进而找到涌现生成所需的三大要素：积木、模型和主体，即作为生成元的积木模块、作为特征构造和生长规律的模型以及具有环境适应性的主体。ChatGPT 通过预训练（Pre-trained）学习，基于 Transformer 完成模型功能，以此实现生成（Generative）功能。Sora 与 ChatGPT 遵循相同的生成逻辑，通过预训练数据和 Transformer 找到视频积木并生成视频。其中，Sora 的视频生成更为复杂，需要用数据包 Patches 代替 ChatGPT 的令牌 Token，除 Transformer 外还增加了扩散模型^①。

第三，通用智能。以往的人工智能皆为专用智能，只具备某种特定功能，例如，AlphaGo 只能下围棋。ChatGPT 和 Sora 均具备通用能力，可以完成不同任务。以往的人工智能虽具有一定的学习能力，但不具备生成功能，主要靠人类“喂养”，按规则行事。而 ChatGPT 与 Sora 主要从方法论、生成逻辑上实现通用功能，只要学会归纳、生成和涌现等能力，就能像人一样根据给定任务寻找要素、规则，并生成答案。换言之，以往的人工智能只学到人类的“术”，是人授予机器以“鱼”；而 ChatGPT 和 Sora 学会了人类归纳、生成之“道”，是人授予机器以“渔”，即只要人类提出任务，ChatGPT 与 Sora 就能完成。

（二）Sora 在技术逻辑上对 ChatGPT 的创新

虽然 ChatGPT 和 Sora 都属于生成式通用人工智能，但后者在前者的基础上实现了突破。笔者从模型结构和训练方法两方面探讨二者的差异与进步。

1. ChatGPT 的模型结构和训练方法

ChatGPT 是基于 GPT 的对话生成模型，具有复杂的模型结构和精细的训练方法。ChatGPT 的模型结构和训练方法可能在细节上作出定制化调整，以适对话生成任务需求，并不断优化模型的性能和生成质量。从模型结构看，ChatGPT 采用 Transformer 架构，包括多层 Transformer 编码器

^① Video generation models as world simulators, <https://openai.com/research/video-generation-models-as-world-simulators>.

和解码器。Transformer 模型的输入是对话历史文本序列，经过编码器处理后，得到语境表示。解码器根据语境表示和自身状态生成下一个词的概率分布，进而作出响应。ChatGPT 的 Transformer 层通常包括自注意力机制、前馈神经网络等子层，用于捕捉全局依赖性和提取特征^①。从训练方法看，ChatGPT 的训练数据通常是大规模的对话文本语料库，包括对话历史和对应回复。在训练过程中，模型通过最大化生成下一个词的概率学习语言模型。ChatGPT 可能采用基于梯度下降的优化算法，例如，利用 Adam 优化器更新模型参数。为提高模型的对话生成能力，ChatGPT 可能进行多轮对话的微调，以逐步优化模型在对话任务上的表现^②。从细节优化看，ChatGPT 可能采用特定的对话生成任务技巧，例如，采用 Token 表示对话开始和结束，引导模型生成合理的对话。模型可能进行 Beam Search 或 Nucleus Sampling 等解码策略，生成更加流畅和多样化的对话响应^③。ChatGPT 还可能通过结合强化学习等方法引入奖励信号指导模型生成更加符合语境和逻辑的对话。

2.Sora 对 ChatGPT 的突破创新

与 ChatGPT 相比，Sora 在多模态支持、知识库、语境理解和用户控制等方面进行创新，使其能够提供更全面、更准确和更具个性化的对话服务^④。从模态支持看，Sora 是多模态对话模型，可以同时处理文本和图像输入，意味着用户可以通过文本或图像与 Sora 进行交互，Sora 可以理解和生成与图片相关的对话内容。这种多模态支持使 Sora 能够更好地理解 and 回应用户需求。从知识库看，Sora 对知识库进行改进，拥有更广泛和更准确的知识。Sora 不仅能够从大量的结构化和非结构化数据源中获取信息，而且可以在对话中使用相关知识提供更准确、更详细的答案，使 Sora 能够回答更复杂、更具体的问题。从语境理解看，Sora 对语境理解进行改进，能够更好地理解对话中的语境信息。Sora 可以更准确地回应用户的追问或进行多轮对话，并保持对话的连贯性。Sora 能够记忆之前的对话内容，并利用相关信息生成更具针对性的回复。从用户控制看，Sora 允许用户控制对话的方向和风格。

用户可通过指令明确期望或偏好，并引导对话发展。这种用户控制使 Sora 能够更好地满足用户需求，提供个性化的对话体验。

综上所述，Sora 继承了 ChatGPT 的生成逻辑，又在其基础上作出改进和优化，能够处理文本、图像和视频等更多模态以及更加复杂的因果、时空问题。文本和视频是人类经验和学习的基础，视频较文本更加形象生动，Sora 将 ChatGPT 的文本生成拓展为视频生成，在三维动态视频生成领域大有可为。ChatGPT 的文本生成和 Sora 的视频生成相互补充，共同构成更加完整的人工智能生成能力。

三、哲学本质：从 ChatGPT 到 Sora 的深化

ChatGPT 与 Sora 揭示了语言理解、意识和思维、学习本质等诸多维度的哲学本质，ChatGPT 主要是因果、语义、语境的理解和生成，Sora 能够理解时间和空间的复杂关系并生成动态时空。

（一）ChatGPT 与 Sora 在哲学本质上的共性

随着数据采集、存储和处理能力的不断提高，人工智能的神经网络派和统计派逐渐兴起，特别是英国人工智能专家杰夫里·辛顿等学者的深度学习不仅可以模仿人类如何从经验中归纳、提炼和学习，而且能够从海量的经验数据中发现新知识。2016 年，AlphaGo 对海量数据进行深度学习，初步证明了从经验数据中学习的重要性。2022 年爆火的 ChatGPT 不仅改变了人们对人工智能的看法，而且颠覆了当时认为能够自我学习、自我提升的人工智能距离人类十分遥远的认知^⑤。Sora 进一步强化了机器学习、语言理解、机器思维等具有深刻哲学意义的问题，因此，有必要探讨 ChatGPT 与 Sora 在哲学本质上的共性。

ChatGPT 与 Sora 作为基于大语言模型的对话或视频生成系统，在人类思维和语言理解方面具有哲学意义，即引发对语言、智能、意识和人机交互等领域的深入思考和讨论，推动人类探索自身认知和技术发展。首先，在语言理解的模拟方面，ChatGPT 与 Sora 通过对海量文本数据的学习和模拟，能够生成类似人类语言的文本，表明人类

①②③ 黄佳：《GPT图解 大模型是怎样构建的》，北京：人民邮电出版社，2023年版，第135、201、225页。

④ Video generation models as world simulators, <https://openai.com/research/video-generation-models-as-world-simulators>.

⑤ 成生辉：《ChatGPT：智能对话开创新时代》，北京：中信出版社，2023年版，第4页。

语言理解的一部分可以被模拟和复制，进而引发对语言本质和理解过程的深入思考。其次，在智能与意识的关系方面，ChatGPT与Sora展现了在不具备意识和情感的情况下也能产生看似智能对话的能力，引发关于智能与意识之间关系的讨论，促使人们重新审视智能和本质的本质。再次，在语言塑造思维方面，语言是思维的载体，ChatGPT与Sora的文本生成能力表明了语言对思维的影响和塑造作用，进一步引发对语言与思维的关系以及语言如何塑造人类认知等问题的探讨。最后，在人机交互的未来，ChatGPT与Sora代表人机交互技术的最新进展，通过自然语言交互实现信息传递和问题解决，对智能助手、虚拟现实等领域的发展具有重要意义。

通过揭示ChatGPT与Sora的机器学习、语言理解和机器思维背后的哲学本质可知，上述问题均与归纳问题密切相关。ChatGPT与Sora的成功再次证明经验是新知识的源泉，归纳是发现新知识的科学方法，一切新知识都是从经验归纳中提炼生成的；否定经验、排斥归纳是错误的，只有坚持归纳和演绎双螺旋式发展，才能真正揭示人类智能的奥秘。

第一，ChatGPT与Sora开启了人工智能归纳—演绎的双螺旋式发展历程。以往的人工智能能够基于现有知识实现演绎和推理，但无法如人类一般发现新知识。ChatGPT成为人工智能发展的转折点，此后的智能机器学会了人类的归纳提炼能力，能够从经验大数据中实现自我学习和自我归纳，提炼出演绎推理所需的一般性知识前提。即ChatGPT与Sora不仅可以依据一般性知识进行演绎和推理，而且能够实现归纳提炼，由此找到作为一般性知识的推理前提。ChatGPT与Sora如人类一般开始坚持归纳和演绎的双螺旋式发展，不再依赖人类为其提供作为演绎前提的一般性知识。智能机器可以离开人类的“喂养”自我发展，人工智能的人工性渐弱，智能性渐强。

第二，ChatGPT与Sora使人工智能初步实现归纳跃迁，从智能走向智慧。以往的人工智能经常被人们诟病为会计算没常识，有智能没智慧。要么依靠海量数据加暴力计算，从中发现有限的算法；要么从现有知识中进行演绎性推理，缺少人的灵活性。智能是通过理性从有限数据中找到有限算法或

有限规律，需要海量数据的学习才能归纳提升，属于完全归纳。智慧除理性的归纳或演绎外，还可依靠联想、直觉、猜想等非理性思维实现从有限到无限的跃迁，属于非完全归纳。ChatGPT与Sora不同于以往的人工智能，二者在一定程度上学会了人的非理性思维能力，能够从有限数据中归纳跃迁全称命题，进而实现非完全归纳。ChatGPT与Sora向前迈进的一小步意味着人工智能向人工智慧迈出一大步。

第三，ChatGPT与Sora开启了智能机器快速自我迭代的时代。人类虽然能够从经验中归纳、学习，并从归纳得出的一般性知识出发进行演绎和应用，但人类的学习能力、记忆能力和计算能力有限，因此，人类知识积累的速度缓慢，经过数万年的经验积累和数千年的文化积淀才达到现阶段的知识规模。ChatGPT与Sora拥有海量的经验数据库，具备超强的学习能力和推理能力，一方面，通过大模型进行快速归纳，提炼出模型，并探索得出规律；另一方面，利用这一模型和规律进行推理演绎应用。ChatGPT与Sora在归纳学习和推理应用中快速迭代，实现自我更新，实现归纳与演绎的迭代闭环，其学习能力超越了任何个人、甚至人类群体。ChatGPT与Sora摆脱了以往人工智能依赖人类学习提升能力的局限，取得了自我学习、自我迭代的突破。

（二）Sora在哲学本质上对ChatGPT的深化

虽然ChatGPT与Sora在机器学习、语言理解和思维等方面具有相同的哲学本质，但深入分析后可知，ChatGPT与Sora在语言理解、创新创造和表现形式等方面存在差异。

ChatGPT作为基于人工智能的对话模型，在人类思维和语言理解方面具有重要的哲学意义，揭示了人类思维和语言理解的哲学本质。ChatGPT引发了人类对语言与思维关系、意义与语义、智能与自我意识关系以及人类思维和认知能力的限制与可能性的深入思考^①。通过对ChatGPT的研究与分析，可以更好地理解和探索人类思维与认知的本质，推动哲学和认知科学的发展。

第一，在语言与思维关系方面，虽然ChatGPT可以生成流畅的自然语言，但不具备真正的思维能力，致使人们重新审视语言与思维间的关系。ChatGPT在某种程度上可以模拟人类的语言表达

^① 黄欣荣、刘亮：《从技术与哲学看ChatGPT的意义》，《新疆师范大学学报（哲学社会科学版）》，2023年第6期。

能力，但不具备内在的主观意识或意图，由此引发哲学家和认知科学家对语言和思维间的本质联系的探讨。第二，在意义与语义方面，尽管 ChatGPT 可以生成看似合理的回答，但其并未真正理解回答的含义。这使人们开始思考意义和语义的本质。关于 ChatGPT 的回答是仅基于统计模型的概率预测，还是存在更深层次的语义理解这一问题，引发关于意义如何产生并理解哲学和认知科学的讨论。第三，在智能与自我意识方面，虽然 ChatGPT 在某种程度上可以模拟人类的智能表现，但其不具备真正的自我意识，由此引发关于智能和自我意识之间关系的讨论。ChatGPT 如何产生智能，是否存在智能的内在本质？上述问题与哲学和认知科学中有关智能及自我意识的问题相关。第四，在人类思维和认知能力的限制与可能性方面，ChatGPT 的存在引发人们对人类思维和认知能力的限制及可能性的思考。ChatGPT 在某些方面表现出令人惊讶的能力，但也存在明显的局限，致使人们开始思考 ChatGPT 的存在是否对人类思维的理解构成挑战，其能否揭示人类自身认知的局限性和可能性。

ChatGPT 主要模仿人类的语言理解和因果推理等能力，涉及语言本质、理解本质、思维和自我意识等哲学问题。但 ChatGPT 尚未涉及时间、空间，即时空问题。康德将哲学范畴归为四大类，即量、质、关系和模态^①，ChatGPT 已经涉及前三类范畴，尚未涉及模态问题。因此，还需具备能够解决时空问题和模态问题的人工智能突破 ChatGPT 的局限。从哲学层面看，Sora 突破了 ChatGPT 的局限，在 ChatGPT 的语义、因果问题基础上实现了对时空观和模态范畴的突破。

在 ChatGPT 的基础上，Sora 激发了对多模态思维、知识与创造相结合、语境感知与对话、用户控制与个性化体验以及迭代改进与自我学习等方面的思考和创造。基于 Sora 的能力和特点，人类可以拓展智慧和创造的边界，实现更加丰富、创新的智慧和创造成果^②。

第一，多模态思维。Sora 作为多模态对话模型，能够同时处理文本和图像输入。这种多模态思维能力启发人们以更全面和多样化的方式思考和创造。人类可以借鉴 Sora 的多模态思维，融合不同

形式的信息和感知，进而产生更富创意性和创新性的思维与创作。第二，知识与创造的结合。Sora 拥有广阔的知识库，可从海量数据源中获取信息，启发人们将知识与创造相结合，产生更富深度和广度的创造性成果。第三，语境感知与对话。Sora 在对话中能够理解和记忆语境信息，可根据之前的对话内容生成回复，启发人们将语境感知和对话作为一种思维和创造工具。人类可以借鉴 Sora 的语境感知能力，将对话和交流视为创造性过程，从中汲取灵感、获得反馈，不断迭代和改进自己的思考与创作。第四，用户控制与个性化体验。Sora 允许用户更好地控制对话的方向和风格，以满足个性化需求，启发人们将用户控制和个性化体验应用于智慧和创造领域。人类可以借鉴 Sora 的用户控制能力，将用户的偏好与需求纳入智慧和创造过程，提供更贴合用户个性化智慧和创造的体验。第五，迭代改进与自我学习。Sora 采用迭代改进和自我学习的方法，不断优化自身性能和表现，启发人们将迭代改进和自我学习应用于智慧和创造过程。

由此可见，ChatGPT 通过大数据、大模型的学习，开始具备因果、语义、语境理解和生成基础，Sora 能够进一步理解时间和空间的复杂关系并生成动态时空。要使人工智能在真正意义上具备人的智能，还需具备时空思维和想象能力，在理解真实世界时空关系的同时，根据真实世界的时空关系创造数字时空。Sora 在 ChatGPT 正确处理因果关系的基础上突破了时空关系，不仅能够生成真实世界的孪生世界，而且能够创造出真假难辨的虚拟世界。Sora 生成的视频图像之所以令人感到惊讶，主要原因在于其已真正具备理解事物因果、时间、空间等复杂关系的能力，能够反映更多模态，生成较文本更加复杂的三维时空。

四、世界图景：从 ChatGPT 到 Sora 的完善

ChatGPT 和 Sora 作为代表性的智能交互工具，具有强大的语言理解和生成能力，可以为用户提供个性化、智能化服务。ChatGPT 是基于大语言模型的对话生成工具，Sora 是智能虚拟助手，二者的结合可以为用户提供更加智能化和个性化

① [德]康德：《纯粹理性批判》，邓晓芒译，北京：商务印书馆，2017年版，第56页。

② Raphaël Millière. Are video generation models world simulators? <https://artificialcognition.net/posts/video-generation-world-simulators/#concluding-thoughts>.

的交互体验。根据 ChatGPT 和 Sora 在智能交互领域的优势和特点，可以构建开放、创新和多维度的世界图景，为智能交互的发展和应用开辟新的可能。

（一）ChatGPT 和 Sora 是构建世界图景的新工具

ChatGPT 和 Sora 建构世界图景包括五个步骤。

第一，知识库和信息获取。ChatGPT 和 Sora 通过知识库和信息获取能力，可以帮助构建世界图景，为用户提供全面准确的信息。第二，语义理解和推理。ChatGPT 和 Sora 具有一定的语义理解和推理能力，尽管其推理能力相对有限，但可以理解和解释用户的问题，并尝试从已有的知识中进行推理和回答。通过语义理解和推理，ChatGPT 和 Sora 可帮助用户构建对世界的理解及图景。第三，语境理解和对话连贯性。ChatGPT 和 Sora 在对话过程中能够理解和记忆语境信息，进而生成连贯且具有针对性的回复。这种语境理解和保持对话连贯性的能力使其能够建立起与用户的连贯对话，逐步构建更完整和准确的世界图景。第四，个性化体验和用户控制。ChatGPT 和 Sora 在对话中允许用户通过明确的指令或偏好引导对话的发展，以满足用户的个性化需求。这种用户控制和个性化体验能力使 ChatGPT 和 Sora 能够根据用户的需求和偏好构建更符合用户期望和兴趣的世界图景。第五，迭代改进和自我学习。ChatGPT 和 Sora 采用迭代改进和自我学习的方法，不断优化自身的性能和表现。ChatGPT 和 Sora 通过与用户的交互和反馈，能够不断改进自己的回答和行为，提高对话的质量和准确性。这种迭代改进和自我学习能力有助于构建更加准确和完善的世界图景。

ChatGPT 和 Sora 是描述真实世界、创生世界图景的新工具，但二者的生成能力存在一定差别，ChatGPT 主要是文本生成器，只能生成文本，形成静态的文本世界和语义世界；Sora 作为世界模拟器，能够生成真假难辨的动态三维时空视频，是世界图景的最佳生成器。

（二）ChatGPT、语义世界与因果性

在 ChatGPT 的世界图景中，语义世界和因果性共同构建了 ChatGPT 与用户交流的基础。ChatGPT 的语义世界由词汇、短语、句子等语言元素构成，通过对语言元素的理解和分析，ChatGPT

建立起对话的语义框架^①。在这一语义世界中，ChatGPT 通过对话与用户交流，不断学习和理解语义信息，进而更好地理解用户需求并提供个性化服务。ChatGPT 的语义世界是动态空间，伴随与用户的交流不断演化和完善，可以为用户提供更加智能和贴近需求的服务。因果性是构建语义世界的基础。在 ChatGPT 的世界图景中，因果性指 ChatGPT 在处理信息时遵循的因果关系和逻辑规律。通过对因果关系的认知，ChatGPT 能够更好地理解问题的本质并提供解决方法，进而提供更加准确和有效的回答^②。因果性认知使 ChatGPT 能够进行推理和逻辑推断，帮助用户解决问题和获取信息。

ChatGPT 通过不断学习以及与用户的交流，逐渐完善了语义世界和对因果性的认知。在与用户的对话中，ChatGPT 不仅可以获取新的语义信息和知识，而且能够通过因果性的推理和逻辑推断，为用户提供更加准确和有效的答案。ChatGPT 的学习与交流是动态过程，通过持续的实践和反馈不断提升自身的智能水平和服务质量。在 ChatGPT 的世界图景中，语义世界和因果性的交织能够为用户服务提供更加智能化、个性化的体验。通过对话与用户交流，ChatGPT 能够更好地理解用户需求并提供定制化服务，帮助用户解决问题和获取信息^③。ChatGPT 的世界图景不仅拓展了人工智能技术在自然语言处理领域的应用范围，而且为用户提供了更加便捷和高效的服务体验。

由此可见，ChatGPT 的世界图景是由语义世界和因果性相互交织组成的世界，ChatGPT 通过不断学习和交流，与用户共同探索和理解世界，为用户提供更加智能化、个性化的服务。语义世界和因果性的交织为 ChatGPT 的发展和应用提供了新思路 and 可能性，将进一步推动人工智能技术在自然语言处理领域的发展和应用。ChatGPT 的世界图景不仅是技术模型，而且是智能化交流和服务的新范式，为人们的生活和工作带来更多便利和可能。

（三）Sora、时空观与元宇宙

Sora 作为高度智能化的虚拟助手，其世界图景充满想象力和科技创新。作为智能虚拟助手，Sora 的世界图景融合了世界模拟器、时空观和元宇宙等概念，能够为用户提供全新的智能交互体验。

①②③ [美]斯蒂芬·沃尔弗拉姆：《这就是ChatGPT》，WOLFRAM传媒汉化小组译，北京：人民邮电出版社，2023年版，第3、96、82页。

在 AI 官网关于 Sora 的说明中，明确表达了 Sora 是一款世界模拟器，意味着 Sora 是为模拟世界而生^①。世界模拟器指 Sora 构建的虚拟现实环境，通过模拟和仿真技术，为用户呈现与现实世界相似、甚至较现实世界更加丰富多彩的虚拟空间。在这一虚拟环境中，用户可以进行交互、学习、娱乐等活动，与 Sora 展开智能对话，并获取个性化服务。世界模拟器的核心在于智能化和个性化。Sora 通过对用户需求及行为进行学习和分析，不断优化世界模拟器的内容与功能，使其更加贴近用户的需求和期望^②。在世界模拟器中，用户可以享受 Sora 提供的定制化服务，获得更加智能和便捷的体验。

Sora 要实现对世界的模拟，除要实现语言理解、把握因果关系外，还需理解时空关系，用视频图像表达物体在三维空间的关系及演变。因此，时空观是 Sora 世界图景中的重要概念。在 Sora 的时空观中，时间和空间不再是静态概念，可以被扩展和重新定义。Sora 通过对时间和空间的重新构想，为用户提供跨维度的智能交互体验。在 Sora 的时空观中，用户可以在不同的时间和空间维度中与 Sora 进行对话和互动。无论过去、现在还是未来，无论地球、外太空还是虚拟世界，用户都可以通过 Sora 的时空观进行跨越式智能交流。这种跨维度的交互体验不仅拓展了用户的认知和想象，而且为智能助手的发展和應用提供了新的可能。

2021 年，元宇宙再度出现在大众视野并引发轰动。有观点认为，元宇宙是虚假概念，不可能实现。囿于技术限制，元宇宙预示的数字技术未来愿景仍在探索之中。Sora 的推出为元宇宙提供助力，元宇宙是 Sora 世界图景中最具前瞻性的概念之一。在元宇宙中，用户可同 Sora 进行更加深入、丰富和多样化的交流与互动，体验前所未有的虚拟现实世界。元宇宙的核心在于开放性和创新性。在元宇宙中，用户可以自由探索、创造和交流，与其他用户共同构建充满活力和创意的虚拟社区，特别是在现实世界中难以实现的理想，通过 Sora 建构元宇宙，并在元宇宙中创世造物，让思想自由驰骋^③。Sora 作为元宇宙的智能助手，有望引领智能交互的未来蓝图。Sora 是实现元宇宙的重要工具，元

宇宙为 Sora 描述了未来世界图景。

综上所述，Sora 的世界图景是融合世界模拟器、时空观和元宇宙等概念的智能虚拟空间，为用户提供了全新的智能交互体验。通过对上述概念的深入探讨，可以更好地理解 Sora 的世界图景蕴含的深层次意义和未来发展方向。伴随人工智能技术的不断进步和创新，Sora 的世界图景将不断完善拓展，能够为用户提供更加智能化、个性化的服务，推动智能交互的发展和應用。Sora 的世界图景不仅是一种技术模型，而且是一种智能化交流和服務的新范式，将为人们的生活和工作带来更多便利和可能性。如果说 ChatGPT 是用文本创构一维的语义世界，那么，Sora 将使其变成更加形象生动、更容易理解和沉浸的三维动态的虚拟世界。未来，基于 ChatGPT 的智能创作文字脚本，Sora 可以将其变成形象生动的视频图像，甚至可能创生元宇宙虚拟世界。

五、结 语

作为生成式人工智能的代表，ChatGPT 和 Sora 的相继出场让世界为之轰动。通过上述分析可知，ChatGPT 和 Sora 是人工智能的关联产品，二者在生成逻辑、哲学本质和世界图景三个维度存在紧密关系。ChatGPT 和 Sora 是基于大模型的生成式人工智能，依靠海量的人类经验数据进行机器的自我学习，从中找到生成元、生成规则和生成模型，进而为智能机器提供演绎和推理所需的前置条件。ChatGPT 和 Sora 能够理解人类的自然语言、因果关系，进而理解并完成人类交付的各项任务。但是，Sora 除继承 ChatGPT 的语言理解能力和因果关系处理能力外，还在真实世界所需的时空关系认知和表述上获得重大突破，生成人类更为熟悉、更易理解的三维时空视频图像。因此，Sora 通过处理多模态信息理解和表征因果性与时空观，生成真假难辨的虚拟世界。Sora 的出现说明人工智能已具备因果观念和时空观念，在一定程度上具有实践理性和综合判断力。综上所述，ChatGPT 和 Sora 开创了人工智能的生成式路径，预示着人工智能正对人类智能靠近，无论生成逻辑、哲学本质还是世

①② Video generation models as world simulators, <https://openai.com/research/video-generation-models-as-world-simulators>.

③ 黄欣荣：《元宇宙、意托邦与数字创世造物》，《上海师范大学学报（哲学社会科学版）》，2022年第5期。

界图景都向前推进了一大步。ChatGPT 和 Sora 的闪亮登场正式开启了走向通用人工智能之路，美国未来学家雷蒙德·库兹韦尔预言的“奇点”可能很快就会来临。

From ChatGPT to Sora: Generative Logic, Philosophical Essence, and World Prospects

HUANG Xin-rong^{1, 2}

(1.School of Marxism, Zhejiang University of Science and Technology, Hangzhou Zhejiang 230031 ;

2.School of Marxism, Jiangxi University of Finance and Economics, Nanchang Jiangxi 330013)

Abstract: ChatGPT and Sora are two consecutive sensational events in the history of artificial intelligence development, jointly demonstrating the power of generative and universal artificial intelligence based on large models. In terms of generative logic, philosophical essence, and world landscape, ChatGPT and Sora have both commonalities and inheritances, as well as breakthroughs and innovations. From the perspective of generative logic, they are all examples of artificial intelligence transitioning from deductive reasoning to learning-generating. They are all based on large models and are outstanding representatives of generative artificial intelligence. However, ChatGPT mainly generates static text and images, while Sora can generate more complex dynamic videos. From a philosophical perspective, both involve issues of philosophical essence such as learning, understanding, consciousness, and thinking. However, ChatGPT mainly focuses on the understanding and generation of causality, semantics, and context, while Sora can further understand complex relationships such as time and space and generate dynamic spatio-temporal videos. In terms of the world prospects, both of them may bring a vast future of innovation and creation, but ChatGPT is mainly a static semantic world, while Sora brings a more complex metaverse world and can play the role of the creator to replay the dynamic process of creation.

Key words: Sora ; ChatGPT ; Generative AI ; Generative Logic ; Philosophy Essence ; World Prospects

[责任编辑: 王文秋]

[责任校对: 曹晶晶]