



新闻爱好者

Journalism Lover

ISSN 1003-1286,CN 41-1025/G2

## 《新闻爱好者》网络首发论文

题目：作为新基础设施的 AGI：以 GPT-4o 等新一代生成式人工智能为例  
作者：郭全中，苏刘润薇  
DOI：10.16017/j.cnki.xwahz.20240527.001  
网络首发日期：2024-05-27  
引用格式：郭全中，苏刘润薇. 作为新基础设施的 AGI：以 GPT-4o 等新一代生成式人工智能为例[J/OL]. 新闻爱好者.  
<https://doi.org/10.16017/j.cnki.xwahz.20240527.001>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

## 作为新基础设施的 AGI: 以 GPT-4o 等新一代生成式人工智能为例

郭全中 苏刘润薇

**【摘要】** GPT-4o 与谷歌 Project Astra 的相继发布, 展现出大模型交互能力的本质性提升和实用性的突破, 预示着以生成式人工智能为代表的 AGI 将如同互联网和人工智能一样, 成为支撑社会运行的新基础设施, 并对传媒业的技术应用、业务实践以及产业融合带来全景式的影响。

**【关键词】** GPT-4o; AGI; 基础设施; 生成式人工智能; 传媒业

### 一、引言

继 Sora 文生视频模型惊艳世界之后, 2024 年 5 月 14 日, OpenAI 揭开了最新生成模型 GPT-4o 的“神秘面纱”, 其强大的实时语音和视频交互能力再次引发震撼。次日, 谷歌在 I/O 2024 开发者大会上也发布了一系列引人瞩目的产品, 包括最新的 AI 智能体 Project Astra。作为新一代生成式人工智能, GPT-4o 和 Project Astra 在技术原理、交互功能和场景落地方面均实现了新的跃升, 代表着通用人工智能 (AGI) 发展的最新突破。随着 AI 技术不断朝着 AGI 演进, 其广泛的认知能力、自我进化能力以及在多样化任务中的应用潜力, 将逐步成为与互联网、人工智能类似的社会新基础设施。对于传媒业而言, 技术始终是驱动行业发展的动力, AGI 作为新基础设施将对传媒业的技术应用、业务实践以及产业融合带来全景式的影响。

### 二、GPT-4o 与谷歌 Project Astra 的新跃升

#### (一) 从“拼接”到“原生”的多模态模型转向

GPT-4o 和 Project Astra 两款产品分别代表了 OpenAI 和谷歌在生成式人工智能领域的最新进展。GPT-4o 以其“Omni” (意为“全能”) 之名, 彰显了卓越的多模态处理能力, 能够接受文本、音频、图像和视频的组合输入, 并生成相应的组合输出, 在实时语音对话、情感表达及视觉交互等方面取得了显著进步。在性能上, GPT-4o 是 GPT-4 的第三次主要迭代, 相较于 GPT-4 Turbo, 其在文本推理、多语言音频处理和视觉理解方面表现出色。与 OpenAI 的 GPT-4o 类似, 谷歌的 Project Astra 也是一款多模态生成式人工智能产品, 作为一个“通用 AI 智能体”, 其基于 Gemini 模型同样展现出多模态理解和实时对话的强大性能。用户可以借助手机摄像头、智慧眼镜等设备捕捉现实世界中的事物, 并通过 Project Astra 获得对所见内容的自然语言描述, 从而实现一种全新的交互体验。

在此之前, 尽管 DALL·E、Midjourney、TTS、Runway、Pika、Sora 等生成式人工智能模型和产品已经能够模拟人类实现跨模态的信息转换, 但都是以单一功能为导向, 如从文本到图像、从语音到文本、从图像到视频、从文本到视频的转换等, 用户在不同任务之间切换时的体验十分碎片化。并且, 这些不同模态之间的信息传递主要依赖于大语言模型 (LLM) 生成的离散文本, 在级联过程中不可避免地会引入噪声并出现传播错误, 同时由于缺乏整体训练, 在理解复杂且隐含的用户指令、进行多模态生成方面能力有限。<sup>[1]</sup>例如, 原先 ChatGPT 在与用户进行语音模式的对话交谈时, 需要首先通过语音识别模型 Whisper 将用户输入的语音转换成文字, 再由 GPT 大模型处理文字, 最后通过 TTS (文字转语音) 模型输出语音。这种由三个模型“拼接”的处理方式不仅速度慢、效率低, 而且还会永久性地过滤掉语调、口音、语速、说话人数、音色、情绪等微妙的信息, 同时输出的语音缺乏抑扬顿挫和情感表

现力，显得机械感强且平淡无味。而 GPT-4o 和 Project Astra 则通过跨文本、视觉和语音的训练，构建了一个“端到端 (end-to-end)”的原生多模态大模型，实现了模型的“三合一”。这意味着所有输入和输出的信息均由同一个神经网络直接处理，无需经过多个独立模型的级联过程，不仅极大地降低了响应的时延，还显著提高了信息处理的准确性和连贯性，并增强了信息输出的表现力。例如，GPT-4o 能够进行实时的语音对话，其输出的语音既可以包含喜怒哀乐等情感元素，甚至还能模拟笑声、唱歌和插话等丰富的情感表现。

## （二）大模型交互能力的本质性提升

保罗·莱文森提出了媒介发展的“人性化趋势”概念，认为媒介必然沿着人类传播要求的方向进化，且媒介具有“延伸性”，其使用功能将越来越符合人类感官愉悦的要求。<sup>[2]</sup>以 GPT-4o、Project Astra 为代表的新一代生成式人工智能则进一步验证了媒介进化论。正如 360 集团创始人周鸿祎的形象阐述，GPT 赋予了 AI 对知识的理解能力，让其拥有了“大脑”；GPT-4V 赋予了 AI 视觉能力，让其拥有了“眼睛”；GPT-4o 赋予了 AI 看懂、听懂和表达情感的能力，让其拥有了“眼睛”、“耳朵”和“嘴巴”。由此可见，生成式人工智能的“人性化”演进趋势将人机交互的体验推进到了前所未有的高度。

第一，实时视觉识别扩展了人类的视觉边界。谷歌发布的视频显示，Project Astra 能够准确识别摄像头捕捉到的物体、代码、行为和场景等，并作出相应的解读，与人类进行实时沟通。同样，在 OpenAI 发布的视频演示中，GPT-4o 具备处理图像分类、物体检测和场景理解等复杂视觉任务的强大能力，用户可以与其“视频通话”，通过摄像头直接解决各种问题，如解答数学题、逗弄宠物、唱生日歌和辅助视障人士等。

第二，类人对话提升了人机交互体验的质量。一是响应速度提升使对话更加流畅。根据 OpenAI 官方数据显示，GPT-4o 针对音频输入的平均反应时长为 320 毫秒，最快可达 232 毫秒，相较于 GPT-3.5 (2.8 秒) 和 GPT-4 (5.4 秒) 的平均延迟，反应时间大幅缩短，与人类日常交流的时序更加契合。二是与 Siri 等传统语音交互系统相比，GPT-4o 不仅无需唤醒词启动，还能记住用户的名字和偏好，从而无需反复输入指令；同时更在音色、音调和表达习惯（如语气词）上几乎与真人无异，甚至能够感知并回应情绪（如开启嘲讽模式），在情感理解和表达上取得了显著突破。三是 GPT-4o 在哲学思辨、科学阐释、工作和日常生活等领域展现出广泛而深入的理解能力。其能够根据场景氛围和用户情绪灵活调整语调和措辞，塑造了一种接近真实人际交流的互动体验。例如，在辅助“石头剪刀布”桌面游戏时，GPT-4o 能够胜任主持人的角色并调动游戏氛围；在即兴创作歌曲时，其不仅能够编写歌词，还能即时生成旋律和谐的曲调，甚至自动进行和声以及实现模型之间的对唱等。

GPT-4o 和 Project Astra 在交互能力上实现了本质性的提升，包括对多维感官的识别、对复杂任务的理解、更自然的对话流畅度和感知情感的能力，标志着生成式人工智能向更加智能化和人性化方向发展的重要一步。

## （三）大模型实用性突破与大规模应用落地

场景落地是大模型发展的关键。然而，无论是 ChatGPT 还是国内的文心一言等大模型，尽管在技术层面上取得了一定进展，但目前主要局限于生成类场景的应用，尚未实现预期中广泛的场景价值。以 ChatGPT 为例，自 OpenAI 在 2022 年 11 月 30 日发布以来，其用户数量在短短五天内突破百万，两个月内迅速飙升至亿级规模。可是，随着时间的推移，这一强劲的增长势头并未持续。根据数据分析公司 Similarweb 的消息，在 2023 年 5 月 ChatGPT 全

球总访问量达到 18 亿次的峰值之后，其流量增长便开始放缓，<sup>[3]</sup>增长势头也再未恢复至先前的惊人速度。这一现象的背后，很大程度上是由于 ChatGPT 在场景落地方面的不足。而 Sora 作为文生视频领域的技术前沿惊艳于世，目前仍处于内测阶段，暂未对外开放使用，这也表明大模型从研发到实际应用的转化过程中面临较大困难。

GPT-4o 与 Project Astra 有效解决了大模型的实用性问题。第一，在使用操作上，此前，用户主要通过键盘输入提示词与 AI 互动，这种操作相对复杂，对于很多普通用户来说存在较高的使用门槛。OpenAI 创始人、CEO 山姆·奥特曼也曾表示，提示工程(prompt engineering)只是生成式人工智能发展的过渡阶段，未来将不再需要这一过程。GPT-4o 和 Project Astra 则实现了从文字输入到口语互动的交互方式转变，类似于智能手机的出现彻底颠覆了基于键盘的操作方式，变为手指触摸操作，AI 技术也迎来了新的“iPhone 时刻”。如今，用户可以通过自然语言直接与 AI 进行口头对话并获得即时反馈，人机交互变得更加直观和自然。由此也预示着生成式人工智能的接受度将大幅提高，更多用户能够真正接触和使用 AI 技术，从而帮助其解决生活难题、提供娱乐和情感陪伴等。第二，在使用成本上，例如 GPT-4o 进一步强化了“免费+收费”的商业模式，践行了“OpenAI 创造 AI，人们利用它创造新奇事物并惠及所有人”的愿景，目的在于降低用户的使用成本，以服务于更广大的社会公众。此外，通过开放 API 接口和降低计费价格，OpenAI 鼓励并支持开发者基于 GPT-4o 构建多样化的 AI 应用生态，从而在大模型的基础上，真正地实现向各个细分商业场景的输出，在教育、广告、搜索、娱乐等各种领域更好地落地。

### 三、AGI 何以成为基础设施？

实现 AGI 不仅需要 AI 具备深厚的知识储备和高度的推理能力，更关键的是要能够与人进行交互。GPT-4o 与 Project Astra 的问世代表着类人 AI Agent（智能体）的雏形显现，展现出了真正能够为人所用的潜力，进一步拉近了人类与 AGI 的距离。那么，AGI 是什么？AGI 如何成为社会的基础设施？AGI 将带来哪些影响？在探讨这些问题之前，首先需要明确基础设施的定义与标准，立足于理论之维与实践之基探寻依据，并从互联网和人工智能的演进中寻找规律，进而理解 AGI 作为基础设施的合理性与必然趋势。

#### （一）基础设施的定义与标准

对于基础设施的理解和研究通常始于技术功能，但其内涵远不止于技术层面。<sup>[4]</sup>从字面上看，“基础设施”由“基础”和“设施”构成，“基础”指的是事物发展所必需的、基本的、不可或缺的支撑要素；“设施”则是指能够满足社会生活需求的操作系统。在英文中，“infrastructure”对应基础设施的概念，由前缀“infra-”（意为“在下面”）和“structure”（意为“结构”）组成，强调其作为“底层技术架构”的特性。综合来看，本文将基础设施定义为支撑社会运行所需的底层技术架构和操作系统。

第一，在理论层面，基础设施需要具备以下特征：一是物质性。基础设施是由一系列中心、节点、线路和终端构成的网络<sup>[5]</sup>，是可以使货物、思想、人员等实现空间交换的物质形式<sup>[6]</sup>。无论是电网、供水、燃气、交通等作为物质工程的硬基础设施，还是由信息基础设施、融合基础设施、创新基础设施构成的“新基建”（新型基础设施建设）等软基础设施，都是人类赖以生活和行动的物质条件，为推动经济社会发展起到了至关重要的基础作用。二是关系性。基础设施不仅可以是物质层面的基础设施建设，还可以是隐喻的深层结构。在 Susan Leigh Star 的研究框架中，技术、基础设施与组织变革三者之间存在着互动关系，拓展了理



解基础设施的关系性视角,<sup>[7]</sup>即基础设施并非是孤立存在的,而是嵌入到了社会结构、配置和技术之中<sup>[8]</sup>,既源自社会关系,也形塑了社会关系<sup>[9]</sup>。因此,在探讨某一事物何以成为基础设施时,还需考量其在落地过程中的物质性特征与其所处的社会环境及不同主体之间的复杂关系,<sup>[10]</sup>只有其与社会实践活动紧密相连,才能真正发挥基础设施的作用。三是历时性。一方面,基础设施的建设与发展是一个逐步成熟和完善,进而基础设施化的过程,无论是规模的扩大、功能的增强,还是技术的革新,都需要历经时间的沉淀。另一方面,基础设施的发展路径还沿袭了既有的惯性<sup>[11]</sup>,但这种惯性并非简单重复过去,而是受到历史经验、社会惯例和传统文化等多种因素的综合影响,形成了相对稳定的基础设施发展模式,为其持续发展提供结构性保障。四是公共性。随着基础设施的发展,其在经济和社会进程中扮演着越来越重要的角色,提供广泛而普遍的公共服务<sup>[12]</sup>,并在整合公众方面发挥着关键作用<sup>[13]</sup>。同时,基础设施也在一定程度上构建了自下而上的权力生态,承担了部分政府的管理职能。

第二,在实践层面,基础设施的建设还深受三大要素影响:一是用户数量。根据梅特卡夫定律,网络的价值与用户数量的平方成正比,基础设施作为承载社会各种资源的网络,其高效运行依赖于用户的广泛参与,用户规模越大,其发展的根基就越牢固,从而带来更显著的规模效应和范围效应。二是用户的使用频度。如前所述,基础设施的“设施”代表满足社会生活需求的操作系统,许多经济社会活动都需要依托于基础设施来实现,因此,高频率的用户使用是评估基础设施有效性的重要指标,也是衡量其功能和服务质量的重要标准。三是对用户的影响程度。用户的行为模式、生产和生活方式等都可能因基础设施的进步而发生变化。例如,公共交通系统的优化和共享出行系统的兴起,改变了人们的出行习惯;随着信息化和网络化的发展,远程办公、在线会议等新型工作方式逐渐成为趋势。由此可见,基础设施对用户行为、观念、习惯等的深刻影响反映了其在社会中的重要性和实际价值。

## (二) 互联网、人工智能的基础设施化演进

### 1. 互联网平台化:作为“连接一切”的基础设施

互联网作为当代社会发展的重要驱动力,其本质在于“连接”<sup>[14]</sup>。这种连接不仅仅是信息层面的交流互通,更涵盖了资源、服务乃至整个社会的全面联通。随着互联网技术的不断发展,平台的功能和作用也在逐渐扩展和深化,影响力日益凸显,平台化的互联网逐步成为“连接万物”、支撑社会发展的重要基础设施。因此,互联网的基础设施化主要体现在平台的展现形式上。

从学理角度而言,平台被视为基础设施的隐喻<sup>[15]</sup>,在现代社会中,互联网平台发挥着越来越基础性的作用。第一,互联网平台的建立基于一系列互联网相关的技术标准,如网络基站、服务器、数据中心、互联网协议,以及互联网接入与输出的终端设备等<sup>[16]</sup>,构成了互联网平台的物质性基础。第二,互联网因其连接性、开放性和去中心化的特征,削弱了传统权力结构中的中心节点,为不同阶层和背景的个体提供了展示和互动的平台,赋予了个体成为信息传播主体的可能性。随着社交媒体平台的出现和发展,互联网日益成为连接和维系社会关系网络的基础纽带,不仅引发了社会关系维度的深刻变革,也使互联网逐渐嵌入更广泛的社会关系网络中,成为社会互动的重要媒介。第三,互联网的发展经历了从PC互联网到移动互联网,再到智能互联网的演进过程,逐步形成了相对稳定的平台发展模式。第四,互联网平台正在自我构建一种庞大的生态系统,逐步融合现有的基础设施,成为兼具信息传递、社会交往、公共传播、政治沟通、产业经营等的公共空间<sup>[17]</sup>。一些超级平台甚至凭借强大的

资本、经济与数据优势，获得了改造乃至创造基础设施的权力，开始承担起部分政府职能。

[18]

而在实践过程中，互联网之所以能够成为基础设施，首先在于其拥有大体量的用户规模。根据 Statista 的数据，截至 2023 年，全球互联网用户数量约为 54 亿，覆盖了全球三分之二的人口。其中，自 2008 年以来，我国一直是全球互联网用户数量最多的国家，这为我国互联网的高速发展奠定了基础。CNNIC 的数据显示，截至 2023 年 12 月，我国网民规模达到 10.92 亿人，互联网普及率达 77.5%，其中使用手机上网的网民比例高达 99.9%，互联网在我国社会具有较高的渗透度。其次，互联网的用户使用频度高。Statista 的数据显示，截至 2023 年第四季度，全球互联网用户每日平均上网时间为 6 小时 35 分钟，其中日均使用社交媒体的时间为 143 分钟。根据 CNNIC 发布的数据，在中国，网民的人均每周上网时长从 2012 年的 19.9 小时增加到 2023 年的 26.1 小时，可以看出互联网在用户日常生活中占据越来越多的时间。此外，在互联网的影响下，人们的生产、生活和工作逐渐从线下转移至线上，生存方式从数字化生存向数据化生存转变<sup>[19]</sup>，短视频平台的兴起更是激发了视频化生存的转向，从而重塑了人们的社交和消费的观念和行为等。

## 2. 平台 AI 化：人工智能技术平移的基础性作用

随着信息技术的快速发展，人工智能技术逐渐取代传统互联网技术成为平台运作的基础性技术。换言之，互联网平台的展现形式依然保持不变，但从互联网向人工智能的技术平移使得当前各大互联网平台都朝着 AI 化的方向发展。

在学理角度上，第一，人工智能并非只是后台运行的代码和抽象的技术，相反，其与物质世界的协同交互与影响是非常深入的。一方面，AI 的运行需要依赖充足的电力供应、低延迟的网络、高性能的硬件、可靠的存储解决方案以及大规模的数据中心等基础性物质资源。另一方面，通过物联网设备（如传感器、摄像头）收集物理世界中的数据，AI 算法能够进行数据分析和处理，进而在复杂的场景中提供精准、高效的智能化服务，如手机地图平台运用 AI 技术可以分析实时交通数据，优化交通流量，减少拥堵和事故发生的可能性。在这个过程中，数字劳动的物质网络也被同时构建了起来<sup>[20]</sup>，在平台经济的驱动下出现了外卖员、网约车司机等新型职业群体。第二，人工智能推动了人类社会关系的重构和结构的重组，实现了从机器辅助人类到人机协同、人机共生的转变，同时虚拟与现实的界限也随之进一步消弭，虚实交互不断深化。第三，人工智能技术由于受到认知水平、技术水平和资金水平等多方面不足的制约，经历了半个多世纪的持续深潜，才逐渐成为社会的底层技术。从翻译软件、AI 围棋模型 AlphaGo，到智能推荐算法、电子设备智能助手，再到生成式人工智能的出现，AI 的应用领域日益广泛。第四，人工智能技术在互联网平台的广泛应用下，正在赋能教育、医疗、交通、农业等社会各领域的数字化、智能化转型，催生出了新理念、新用户、新模式、新产业和新应用。在 2024 年《政府工作报告》中，我国将“人工智能+”行动上升为国家战略，未来将形成“人工智能+各行各业+各应用场景”的深度融合。

在实践角度上，无论是在线学习、远程办公，还是智能家居、智慧医疗等领域，人工智能正在对人类生活进行全方位渗透。自从 ChatGPT 问世以来，特别是新一代生成式人工智能如 GPT-4o 和谷歌 Project Astra 的推出，AI 直观易用的界面设计、高效准确的信息处理能力、自然流畅的对话体验，进一步提升了用户接纳度。

### （三）“生产即分发”：AGI 也将成为新基础设施

自 1956 年达特茅斯会议上首次提出“人工智能”概念以来，为了实现 AGI 的终极目标，人工智能技术经历了专用人工智能（ANI）的经验积累，以及 AI 大模型的能力涌现，奠定了技术、资金、人才等基础条件。所谓 AGI 就是一个能够执行人类所有工作，甚至超越人类能力范畴的智能系统，它不再局限于特定的任务或领域，而是拥有广泛的认知能力、自我进化能力，并在多样化的任务和复杂环境中展现出人类级别或更高的智能水平。<sup>[21]</sup>也就是说，AGI 的最终意义是创造出一个终极智能世界<sup>[22]</sup>，将成为未来的社会基础设施。

从学理角度出发，互联网和人工智能技术推动了互联网平台迅速崛起，出现了“基础设施的平台化”和“平台的基础设施化”<sup>[23]</sup>，平台成为基础设施的主要展现形式。但是，相对于未来的 AGI，则会在现有平台基础上衍生出“生产即分发”的新形式。具体而言，AGI 基础设施将不再局限于当前平台的展现形式，而是进一步把现有的基础设施化、媒介化的平台“再中介化”。这一过程彻底改变了传统的生产和分发流程，实现了生产之后直接分发的模式，无需依赖传统平台作为“中介”。例如，基于个体化模型和个性化 AI 进行的人机交互，生成的内容可以直接面向用户进行分发，而不再需要经过任何中介平台的处理。AGI 通过“去中介化”，及其在全面泛化的过程中具备的环境配置性、嵌入基底性、动态生成性、关系公共性、透明生态性<sup>[24]</sup>的特征，在各类复杂环境中自适应地实现更高层次的资源配置效率和用户体验优化。正如凯文·凯利所言：“当技术变得隐形的时候才是最强大的。”未来，AGI 会成为“看不见、摸不着、离不开”的基础设施，深度渗透并融入社会的每一个细微脉络之中，无缝连接并塑造社会运作的每一个层面，真正实现全面的智能化。

在具体实践方面，有学者通过实证研究发现用户会下意识地将人际交往的原则和期望迁移至人智交互情境中，因此，在 AI 的研发与设计过程中，能力和情感的双重兼顾尤为必要。<sup>[25]</sup>目前，GPT-4o 与谷歌 Project Astra 已经突破了语音、听觉和视觉感官能力的界限，展现出强大的情感理解能力，甚至能够提供情绪价值，带来了丰富的交互体验。未来，AGI 将可能模拟人类“眼耳鼻舌身意”六识的感知能力<sup>[26]</sup>，来获得更多的感知数据，从而带来更加直接、自然、真实的沉浸式体验，以更加细腻、人性化的方式融入人类社会的方方面面。虽然，目前的 AI 技术与真正的 AGI 尚存在距离，但根据 ChatGPT、Sora、GPT-4o、Project Astra 等引发的社会反响，可以预见，随着 AI 能力的不断提升和 AI 技术的日益普及，AGI 的用户群体将持续扩大，使用频度亦将大幅度提升。

### 四、AGI 作为新基础设施对传媒业的全景式重构

GPT 基于背后的预训练大语言模型，能够对知识进行理解，表现为二维的信息处理能力；Sora 作为“世界模拟器”，具备视觉感知能力，能够对物理世界进行更加立体和直观的认知和模拟，实现了对现实的三维理解。但是，GPT 和 Sora 更多是在技术和底层架构上产生了影响，并没有真正实际落地，或者说在实际操作层面仍然面临一定难度。而 GPT-4o 和谷歌 Project Astra 在技术、性能上的跃升为大模型增加了一个全新的维度——能够实际应用和实现场景落地，预示着 AGI 正在从理论走向现实，有望真正成为新的基础设施。对于传媒业而言，具备更为有效落地场景的 AGI 新基础设施将引发行业的全景式变革。

#### （一）AGI 将成为传媒业的技术主导方向

传媒业的发展史本质上是一部技术进化史，技术一直是传媒业发展的核心驱动力。古登堡发明印刷术促成了报纸的诞生，拉开了大众传播时代的序幕；广播电视技术的出现推动了



电子媒介的繁荣；进入互联网时代，数字化浪潮带来了翻天覆地的变化，传统媒介渠道的传播力被削弱，用户连接的失效导致了以“二次销售”为核心的商业模式的坍塌，以互联网技术为主导的数字化转型成为媒体发展的目标和方向。近年来，5G、物联网、大数据、算法、人工智能等前沿技术的飞速发展，催生了短视频、直播等新型媒介形态，新闻内容的视频化转向、直播带货的商业模式等为传媒业带来了新的收入增长点。

随着人工智能技术的进步，传媒业迎来了以生成式人工智能为代表的新一轮技术革命。从 ChatGPT 人机对话应用程序横空出世到 Sora 文生视频大模型惊艳四方，再到 GPT-4o 和谷歌 Project Astra 的能力跃升，AGI 的曙光日益显现。尽管目前还未真正实现 AGI，但自然语言处理（NLP）、机器学习等基础性技术，以及具备泛化任务处理能力的类 AGI 已经在传媒业得到广泛应用，如撰写新闻稿件、生成采访提纲、翻译文件、分析数据、优化广告创意等，为传媒业实现了降本增效。此外，一些媒体还积极推出了传媒垂类大模型，如人民网与百度发布的“人民网-百度·文心”、中央广播电视总台与上海 AI 实验室联合发布的“央视听媒体大模型”、传播大脑科技（浙江）股份有限公司研发的“传播大模型”等。未来，AGI 将成为传媒业的技术主导方向，推动传媒业在内容生产、传播模式、营销模式等方面的全面智能化。

## （二）传媒业务实现智能化、精细化与场景化再造

第一，在 AGI 的加持下，新闻生产主体泛化，形成“UGC+PGC+AIGC+AI Agent”协同生产的格局。其中，AIGC 已在体育、财经报道中有所应用，通过 AI 自主内容创作，实现全天候、实时内容产出，提升了内容生产效率。AI Agent 作为一种高级应用程序，能够理解人类指令，并以类人的方式作出回应、互动和决策。原先的围棋机器人 AlphaGo、语音助手 Siri 等可被视为 AI Agent 的初级形态，而 GPT-4o 和谷歌 Project Astra 则具备了更高级的类人 AI Agent 特性，能够在与用户进行自然、流畅、即时的互动交流，将成为全新的内容生产主体。

第二，在 AGI 的驱动下，新闻传播的互动方式和场景也展现出显著症候。一是 GPT-4o 和 Project Astra 这种类 AGI 技术在人机交互方面的显著进步，已经调整并重塑了传媒业的新闻分发模式，即在人与机器的一问一答过程中，形成了“生产即分发”的传播形式。具体来说，用户通过即时语音对话和实时参与主动提出需求，AGI 在时延、表达方式上能够进行类人响应，针对性地满足其需求，从而颠覆了传统新闻机构“点对点”的新闻分发模式，以及依赖算法推荐而导致的内容分发同质化现象，实现了真正意义上的“点对点”传播，满足了个性化、精细化、“千人千面”的互动体验<sup>[27]</sup>。二是 AGI 进一步深化了元宇宙新闻场景的“脱虚向实”进程。一方面，作为元宇宙最早落地的应用场景之一，以 GPT-4O 为代表的新一代生成式人工智能赋能虚拟数字人的类人化属性和智能化水平，能够克服当前一些数字人依靠幕后工作人员操控、动作和表情不自然、语音交互不流畅等不足，给予用户更优质的体验。另一方面，通过手机摄像头、智慧眼镜或数字替身，用户可以参与沉浸式的新闻互动，新闻传播从单一的信息传递转变为多感官的综合体验，在一定程度上增强了新闻的现场感和真实性。

## （三）产业融合与传媒功能深度延伸

在 AGI 的技术语境中，传媒业作为相对传统的产业，需要通过与 AI 技术的深度融合，注入新的活力与动力。AGI 作为基础设施将延续“互联网+跨界”的模式，推进传媒业在产



业层面的深度融合。第一，互联网企业通常是新技术的先行者，无论是西方发达国家还是我国的互联网企业，都在积极探索 AI 大模型及其相关业务，致力于实现产业智能升级，推动 AGI 的全面快速发展，为传媒业的跨界融合提供了技术、人才和资金支持。第二，面对技术对传媒业的冲击，尤其是 AGI 应用下“生产即分发”的新模式，将进一步动摇媒体的用户基础。因此，传媒业需要把握住 AGI 的机遇，将功能的“触角”延伸至社会更多领域，如文化创意产业、教育、医疗、智慧城市、数字治理等，拓展自身的影响力。

## 五、结语

塞巴斯蒂安·特伦曾言：“人工智能更像是一门人文学科，其本质在于尝试理解人类的智能与认知。”以 GPT-4o、谷歌 Project Astra 为代表的新一代生成式人工智能赋予了 AI 以人类的感官，不仅在多模态处理和人机交互能力上实现了显著突破，更在实际应用场景中展现了广泛的实用性，代表着 AGI 的最新发展方向。在 AI 技术的演进趋势下，AGI 将逐渐成为无处不在的现象和实体，成为继互联网和人工智能之后的新基础设施。然而，任何技术的发展都伴随着挑战，在实现 AGI 的过程中，诸如数据安全、隐私保护、伦理道德等问题亦不容忽视，在未来的演进过程中还需进一步跟踪研究。

[项目基金]: 本文为北京市社会科学基金规划重点项目“首都互联网平台企业社会责任与协同治理体系研究”的阶段性成果，批准号(22XCA002)

## 参考文献:

- [1] Wu S, Fei H, Qu L, et al. Next-gpt: Any-to-any multimodal llm[DB/OL].(2023-09-13)[2024-05-18]. <https://arxiv.org/abs/2309.05519>.
- [2] 保罗·莱文森.人类历程回放:媒介进化论[M].郭建中,译.重庆:西南师范大学出版社,2016.
- [3] 财联社.OpenAI 放开限制! 用户无需注册即可使用 ChatGPT[EB/OL].(2024-04-02)[2024-05-18]. <https://finance.eastmoney.com/a/202404023031708894.html>.
- [4] 柯小俊.“技术网络”与“意义点位”:作为纪念碑的基础设施研究[J].新闻大学,2023,(7):86-100+120.
- [5] [9][11]张磊,贾文斌.互联网基础设施研究:元概念、路径与理论框架[J].中国社会科学院研究生院学报,2021,(5):22-32+145.
- [6] Larkin B. The politics and poetics of infrastructure[J]. Annual review of anthropology, 2013, 42: 327-343.
- [7] 王家东.基础设施、边界对象与科学技术研究的人类学路径[J].自然辩证法研究,2023,39(6):112-117.
- [8] Star S L. The ethnography of infrastructure[J]. American behavioral scientist, 1999, 43(3): 377-391.
- [10] 段世昌.媒介基础设施视角下的技术编码研究——教育网 PT 站的“分享”规则及用户实践[J].新闻界,2020,(10):81-94.
- [12] 张超义.能源网、交通网、信息网:作为媒介的网络型基础设施与三网融合[J].全球传媒学刊,2023,10(5):56-70.
- [13] Collier S J, Mizes J C, Von Schnitzler A. Preface: Public infrastructures/infrastructural publics[J/OL].(2016-07)[2024-05-18].<https://limn.it/articles/preface-public-infrastructures-infrastructural-publics/>.

- [14] 彭兰.“连接”的演进——互联网进化的基本逻辑[J].国际新闻界,2013,35(12):6-19.
- [15] 刘战伟.凸显什么?遮蔽什么?——作为隐喻的“平台”:连接、中介与基础设施[J].新闻记者,2022,(6):54-66.
- [16] 束开荣.互联网基础设施:技术实践与话语建构的双重向度——以媒介物质性为视角的个案研究[J].新闻记者,2021,(2):39-50.
- [17] 殷琦,国秋华.从静态机构到动态功能:互联网平台的内涵演变与治理转向[J].现代传播(中国传媒大学学报),2023,45(4):135-142.
- [18] 刘战伟,刘洁.“平台/platform”:一个概念史的溯源性研究[J].新闻与写作,2023(8):70-82.
- [19] 彭兰.“数据化生存”:被量化、外化的人与人生[J].苏州大学学报(哲学社会科学版),2022,43(2):154-163.
- [20] 孙萍.如何理解算法的物质属性——基于平台经济和数字劳动的物质性研究[J].科学与社会,2019,9(3):50-66.
- [21] 郭全中,张金熠.作为视频世界模拟器的 Sora:通向 AGI 的重要里程碑[J].新闻爱好者,2024(4):9-14.
- [22] 郭全中,袁柏林.从 GPT 看 AGI 的本质突破:传媒业新挑战与未来[J].新闻爱好者,2023(4):30-35.
- [23] Plantin J C, Lagoze C, Edwards P N, et al. Infrastructure studies meet platform studies in the age of Google and Facebook[J]. New media & society, 2018, 20(1): 293-310.
- [24] 陈昌凤,袁雨晴.智能新闻业:生成式人工智能成为基础设施[J].内蒙古社会科学,2024,45(1):40-48.
- [25] 张妍,赵宇翔,吴大伟,等.人智交互情境中用户对生成式人工智能的心智感知及反应研究[J/OL]. 情 报 理 论 与 实 践 ,1-12[2024-05-18].  
<http://kns.cnki.net/kcms/detail/11.1762.G3.20240510.1836.002.html>.
- [26] 张诗瑶,沈阳.Sora:传媒生态的镜像进化与认知变革[J/OL].编辑之友,1-8[2024-05-18].  
<http://kns.cnki.net/kcms/detail/14.1066.G2.20240419.1104.002.html>.
- [27] 郭全中,苏刘润薇.大模型驱动下的传媒再造[J].出版广角,2024(3):4-12.

#### 作者简介:

郭全中:中央民族大学新闻与传播学院教授,互联网平台企业发展与治理研究中心主任,江苏紫金传媒智库高级研究员

苏刘润薇:中央民族大学新闻与传播学院硕士研究生