生成式AI幻象的制造逻辑及其超真实建构的文化后果

经羽伦　张殿元

【摘　　要】在人类文明史上出现的众多媒介都天然地具有欺骗性。正是媒介的欺骗性特征，使得媒介在中介化现实的过程中不断制造幻象。作为新型媒介的生成式AI，同样具有欺骗性，且在制造幻象方面已超越既往一切媒介，具体表现为“AI幻觉”“模拟影像”和“深度伪造”等形式。基于概率和模拟的自动化生产的软件系统，构成生成式AI复杂且庞大的欺骗机制，具有物质属性的交互界面，将此种欺骗隐匿其下。生成式AI已经建构起一个“超真实”的世界，其幻象正对应着鲍德里亚所预言的“分形文化”。最终的文化后果是，在生成式AI建构的幻象宇宙里，真假与否不再重要，它将我化引入一种“乐此不疲地制造幻象”的媒介环境与新文化形式之中。技术之恶的潜能亦可能被前所未有地激发。

【关 键 词】生成式AI；幻象；媒介的欺骗性；超真实；分形文化

【作者简介】经羽伦(1996–)，女，江苏扬州人，复旦大学新闻学院博士研究生；张殿元(1970–)，男，吉林长春人，复旦大学新闻学院教授，博士，博士生导师(上海 200433)。

【原文出处】《山东师范大学学报》：社会科学版(济南)，2024.5.113～126

【基金项目】本文为国家社科基金重点项目“新媒体文化对新时代中国创新型发展的理论支援和实践路径”(21AXW010)的阶段性成果。

2023年，OpenAI推出ChatGPT，让我们见证了文生文、文生图的快速发展过程。仅时隔一年，OpenAI发布文生视频大模型Sora。Sora能根据用户提示词(prompt)，快速生成连贯的60秒高质量视频，视频不仅可以呈现用户提示中所描述的内容，而且多角度镜头、分镜头的切换和运镜等也都符合逻辑。更为重要的是，Sora生成视频的流畅度、清晰度和逼真度与实景拍摄几乎毫无差异，如若不在视频中标注来源，用户根本无法看出视频是由Sora生成的模拟影像。在我们惊叹生成式AI的想象力和创造力之余，也会发现生成式AI似乎并不“完美”，ChatGPT和Sora都会制造出以错误再现、错误生成、AI幻觉等为代表的“生成式AI幻象”。[①]而具有涌现现实物理世界能力的Sora更是挑战了“何为真实”，生成式AI的欺骗性特征昭然若揭。那么，生成式AI的欺骗性和幻象之于真实而言意味着什么？对此，本文尝试在连续而非断裂的媒介史中获取启示，重思真实、幻象与媒介再现的关系，从媒介本体的欺骗性特征出发，考察不同媒介间欺骗性的共性和差异，并试图厘清媒介的欺骗性与错误再现、错误生成、虚假等幻象制造之间的关系，进而深度透视生成式AI幻象的本质、欺骗性的来源及其幻象制造可能引发的一系列潜在的文化后果。

一、真实、幻象与媒介再现

在进入媒介幻象、媒介的欺骗性与生成式AI的探讨之前，首先有必要厘清的一个关键问题是：何为真实？从哲学层面来看，“真实”表达的既是对“世界”的整体“真实”的诉说，又是对人的全面“真实”的

追求。[2]对此，现代性与后现代哲学家对“何为真实”的命题，呈现出巨大的学术分野。具体而言，现代性强调“理性”与“科学”，“逻各斯”(Logos)的通常解释是：理性、判断、概念、定义、根据、关系[3]，意味着“最真实的存在”和“最根本的真理”，构成人和万事万物都须服从的根本尺度与法则。在现代性的意义上，理性构成人的思想和行为的终极根据和源始出发点，使其进行自觉理解和反思。[4]基于理性的自然科学是现代性进程中的重要产物，如果一种科学原理在条件相同的情况下都能得到相同的结果，那么，这个科学原理就是真实的。[5]事实上，在现代性的视野中，由于科学和技术促使工业社会的发展与人类文明的进步，使其对唯一性、同质性、确定性表现出前所未有的坚持，因而现代性的观点认为存在绝对真实。同时，正是科学原理和技术再现了真实存在的东西，真实成为真理，进而真实变得可知、可确定。由此，“真实”总是与技术、媒介和符号的再现(representation)联系在一起。我们说一个东西是真实的，也就是说，这个东西能够被同样地再现出来。

然而，后现代哲学家的观点恰巧与之相反。后现代性是基于生产过剩的后工业社会背景而出现的，执着于“解构”，试图挑战、动摇和否定社会中一切传统、一致价值和普遍法则，这其中就包括现代性所强调的理性与科学，取而代之的则是多样性、异质性和区隔。同样，他们批判现代性的真实观，并以相对主义的视角审视“真实”。在他们看来，世界并不存在唯一的“真实”，“真实”是相对的、异质的和多元的，即客观真实(objective reality)是不容置疑的真理，不需要验证，也验证不来，他们并不执着于对唯一真实的追求，反而对此抱有警惕。同时，由于现代工业社会试图通过媒介和符号刺激消费，以解决生产过剩的问题，这使得后现代哲学家对真实与媒介再现之间的关系加以关注，并对由符号和媒介所建构的幻象展开强烈的批判。由此，西方学者将后现代性在文化上的思潮或者审美文化上的运动，概述为“再现的危机”(crisis of representation)。[6]

后现代哲学家纷纷指出，现实已被媒介再现或表征所建构的幻象取代。对此，法国思想家居伊·德波(Guy Debord)在《景观社会》中写道：“在现代生产条件主导的社会中，生活被表现为无数景观的堆积。一切鲜活的东西退居到表征当中。”[7]“景观”(spectacle)或后现代理论家所称的“拟像”(simulacra)，是一种由表象支撑的幻象，它遮蔽了社会存在本身。牛津词典对“拟像”一词的释义包括两个层面：一是事物的影像；二是具有欺骗性的代替物。[8]法国哲学家让·鲍德里亚(Jean Baudrillard)将“拟像”视为“编码规则要素及媒介技术操作的赝象”[9]，是“虚构的”或“模仿的”事物代替“真实”的过程。进入“仿真”阶段，“拟像”代表着后现代社会大量拷贝、极度真实而又没有本源、没有客体指涉、没有根基的虚拟存在或幻象，包括形象、符号、影像或景观等。正是“拟像”在替代“真实生活”和存在于真实世界中的客体与实在。

在厘清现代性和后现代性有关“何为真实”以及“真实与媒介再现”的学术分野后，可以更加清晰地洞察到：现代性的真实观是“真实是可知的、确定的和唯一的”，媒介可以再现真实；而后现代性的真实观是“真实存在但不可知”，媒介无法再现真实。具体来说，现代性的观点是技术、科学、媒介和符号可以再现真实。诚然，传统的大众传播媒介亦是现代性工业社会的产物，其之于真实的理解也是唯一的、确定的，主要体现在通过新闻客观性的方式和新闻专业主义的要求，追求唯一、确定的真相和事实。但后现代哲学家则尝试用“拟像”来阐释真实与媒介再现，并用“超真实”(hyperreality)的概念进一步揭示“符号真实”(symbolic reality)与“客观真实”的断裂与僭越。所谓“超真实”，就是一种没有原型和真实性的真实，由一种真实的模塑制造的真实[10]，它比真实还真实，真实不再是那个可以再现的东西，而永远都是已经再现的东西，即超真实的东西[11]。可见，对后

现代哲学家而言，符号真实与媒介再现所构造的幻象已然超越并替代真实本身。大众传播媒介并非在接近客观真实，它们只是在一味地炮制真实、制造幻象。

对“真实与媒介再现”的理解，更多的是基于后现代性的视野，且尤为关注“幻象与媒介再现”的关系，即由媒介(特别是生成式AI)所制造的虚假、错误生成、错误再现等幻象问题。当下，我们正身处数字媒介时代，人类生存的世界图式发生彻底性变革，相比于传统媒体时代的拟像、幻象，数字媒体恐有过之而无不及。后真相(post-truth)的来临足以说明，传统新闻媒体机构所坚守的唯一且确定的真相、真实正在被数字媒介和去中心化的大众节点所解构，真实被大众情绪所淹没，散落在虚拟空间的各个传播节点之中，每个个体感知并建构主观真实(subjective reality)，即“真实”成为个体内在的“真实”，是基于个体的情绪、感受的体悟而变动的，可感知的即为“真实”。因此，真实变得多元，随之而来的是幻象无处不在，而基于理性的唯一真实却难以寻觅，换言之，鲍德里亚断言的“超真实”在数字媒介更为显现。站在人工智能时代即将到来的临界点，作为技术图像的Sora横空出世，意味着媒介可以模拟出现实物理世界，生成式AI将制造出更多的媒介幻象，并以全新的幻象形式出现。对此，本文首先的观点是：媒介再现过程中所出现的幻象问题与媒介本体所固有的欺骗性特征有关。

事实上，媒介的诞生和演进总是无止境地缠绕着欺骗，尽管多数情况下人类不遗余力地识别并消除充斥于媒介中的虚假和幻象，以表征事实和真相，但不得不承认的是，在媒介建构现实世界的过程中，始终弥散着不同程度的、有意或无意的欺骗。诚然，媒介具有再现的功能，但“再现”代表了一个表述的世界而非现实的世界，是“真实”世界里事物的一种映象、类似物或复制品。[12]在以符号真实建构、再现世界的过程中，媒介与具体指涉物之间的真实是有所出入的，即使没有蓄意、恶意的欺瞒和误导，媒介技术也必然伴随各种形式的幻象，其中涉及虚假、错觉和幻觉等错误再现的形式。法国哲学家雅克·德里达(Jacques Derrida)认为，我们很难区分正确的再现与错误的再现，区分真理与谬误、现实与虚构，事实上，再现总是伴随着不可避免的虚构与错误。[13]

媒介之中出现的错误再现、错觉、蒙蔽和虚假等现象均可以归结为幻象，幻象对应着一种结果，映射在人类的感官上。所谓“幻象”(illusion)，指幻想、幻觉或梦境中产生的形象，更深层地说，“幻象”是表现欲望之实现的想象的场景，其与“现实”是一组相互对立的概念，而幻象也常常产生于媒介中介化现实的过程之中。在传统媒介研究中，学者往往将欺骗所导致的幻象结果归结为人为因素，即以错误再现为代表的媒介幻象是由人类有意或无意为之造成的。首先，人类通过感官和身体感受来认知和理解外部世界，但人类感官和身体存在局限性，这也致使人类常常陷于错觉和幻觉之中，因而“能够被欺骗”是人类体验世界和感受世界固有的生理特性，这也是幻象得以蒙混过关的前提。即人类拥有利用媒介技术操纵再现和制造幻象的能力。至此，“欺骗”在媒介史中主要作为一种不同寻常的情况被研究，往往强调人类对媒介的操纵能力，而非承认“欺骗”在现代媒介中的结构性作用。[14]这是人类中心主义视角导致的结果，它忽视了欺骗是人类和媒介技术交互关系的结构性组成部分，作为非人行动者的媒介，也同样具有欺骗性的特质以及制造、操纵幻象的能力。对此，我们选择回到制造幻象的原点，另辟蹊径地从媒介本体的“欺骗性”特征出发，寻求对媒介幻象制造过程的再阐释，“欺骗”一词暗示着某种形式的主体能动性，不仅是人类和技术开发者，还包括媒介和技术本身。

二、媒介的欺骗性与幻象制造

在人类文明史上出现的众多媒介也都天然地具

有欺骗性。在口语文字时代，语言中充满谎言，文字中始终存在虚假信息，包括绘画中的明暗光影都充满欺骗的意涵。进入印刷时代，古登堡印刷术带来机械复制，进而导致艺术作品“灵韵的丧失”，真品与复制品等同出现，媒介的欺骗性特征不言而喻。自工业革命兴起，随着各种电子媒介日益介入人类日常生活，媒介技术本身的欺骗性特征愈加明显。相对于传统的口语、书写和印刷系统，电子媒介是一个更擅长于呈现“真实”的符号系统，在再现和建构真实的过程中，电子媒介邀请人类的身体感官加入，从而使得“符号真实”在感官层面以更逼真的形式呈现。以留声机、电话和广播为代表的语音电子媒介，通过符号真实影响听众，使得听众将符号真实当作客观真实。其中，最具代表性的事件便是广播剧“火星人入侵地球”。[15]现场直播的媒介技术直接导致听众混淆虚构和现实，这说明广播不仅存在媒介欺骗性，而且其对人类行为的潜在影响力是巨大的。同时，照相机和摄像机等影像媒介技术，将“真实”以帧的形式存储、传输和呈现，为公众创造出口语、书写和印刷系统所不曾有的“真实感”。图像的出现似乎在摒弃人类的主观判断而忠实于世界的真实复刻，作为传播中介的影像媒介最初也被视为是客观中立的，这无疑强化技术主义“眼见为实”的论断，但当现实的本体落在影像层面，本原真实就会变得难以把握，透过影像洞察、获取的真实、事实和真相成为一种伪命题，最终，眼见未必为实、有图也未必有真相，影像媒介不断制造幻象，其欺骗性亦在所难免。

不同于语音和影像媒介以电视和电影为代表的影音媒介，同时实现了人类听觉和视觉感官的双重延伸，随之而来的是，人类的想象性被前所未有地激发，借由影音媒介技术，制造幻象的手段变得更加丰富，媒介的欺骗性开始以一种动态的、可见性的方式呈现。以电影为例，电影媒介拥有制造幻象的能力，弗里德里希·基特勒(Fredirch Kittler)在谈及电影媒介时认为，“在这些被欺骗的观众的眼中，在从未设想惊吓或谋杀的默片制作者的背后，电影从它的最初形态变为了一种幻觉媒介(illusionary medium)”。[16]在基特勒那里，善于制造和操控电影幻象的“大师”并不是卢米埃尔兄弟或梅里爱。诚然，他最关注的并不是导演们的个人艺术风格，或那些他们引以为傲的“经典”电影作品，而是那些构成“想象性”和“幻觉性”的物质基础，即媒介和技术本身，包括具体的电影拍摄技术或技巧，如叠化、多次曝光、倒放和停机再拍等。[17]一切电影媒介美学都源于后来被标准化的24帧的镜头，停机再拍与蒙太奇、慢动作与延时摄影只不过是把技术转译进观众的欲望，作为我们被欺骗的眼睛中的幻象，剪辑重新生产出运动的连续性与规则性。[18]幻象在所谓连续的时间过程中非连续地闪现。可见，在电影制作过程中，媒介和技术的欺骗性成为制造幻觉不可或缺的环节，它正对应着一种艺术创作的想象力，电影效果的好坏开始取决于媒介技术在多大程度上能够让人类心悦诚服地陷入欺骗和幻象之中。

进入互联网和数字技术时代，以计算机、智能手机和VR/AR为代表的数字媒介突破现实之维，使得人类社会日益脱实向虚。计算机媒介以数据、算法与计算机图像为数字基础设施，在柏拉图的洞穴之内的幻象世界与洞穴之外的现实世界之外，又重新创建起一个平行的虚拟世界。然而，在虚拟的外衣之下，伴随欺骗性技术手段的多样化发展，媒介的欺骗性日益介入人类日常生活，令幻象无处不在地充斥于人类的各种感官体验之中。首先，基于数字媒介技术的演进，影音媒介发生重大技术变革：运动影像从电子化走向数字化，这让影像由电子流变为二进制符码。[19]借助数字虚拟技术与数字媒介界面的交互系统，影像走出电影院和家庭客厅的物理边界，进入智能手机、平板、计算机和VR/AR等具有移动性和交互性的数字媒介之中，演化出一种新的影像形式：短视频。同时，数字媒介去中心化的生产模式，

使得短视频数量井喷,影像普遍化。然而,基于交互界面的指间实践,通过拼接、剪辑和修图等软件应用,借助移花接木的技术手段,短视频生产的影像更具迷惑性和欺骗性,其中制造的幻象更可谓千奇百怪、层出不穷。因此,数字媒介影像不仅完美继承电影和电视制造幻象的能力,而且还将媒介的欺骗性手段运用到极致,以致幻象开始达到"真假难分"的局面。

远不止是影像制造的幻象在数字媒介中更进一步,另一种更为隐蔽的、起到决定性作用的欺骗性技术系统被隐藏在交互界面之后,并决定着数字媒介影像和事实真相的前台再现,即数据与自动化分类、排序的算法机制。基于计算机的系统获取、处理、配置和再现的"数据",已经成为一种"事实来源",甚至在替代、侵蚀着事实本身。事实的时代被数字媒介和互联网终结:曾经由"事实"占据的位置正在被"数据"取代。[20]同时,"软件代理"的"自动化中介"实现大规模信息处理的计算资源配置,在相当大的程度上无需人工直接干预[21],即数据的收集、分类和排序几乎全权外包给机器。不置可否,数据不总是代表真实,算法及其权重也隐含价值偏见,在一系列的抽象机制和不可见的算法黑箱之中,虚假和偏见所引发的欺骗,被裹挟上技术中立的外衣而变得尤为隐秘,由此,计算机等数字媒介背后的数据和算法机制成为引发后真相的根源,建构出一种新的事实幻象。总体来看,通过"二进制符码-交互界面-影像幻象"的路径,作为元技术的数字媒介可以统一并超越既往媒介的"欺骗性"特征,并使得影像幻象无处不在、无时不在。在此基础上,数字媒介又出现新的欺骗性技术手段,基于数据和算法,数字媒介形成制造幻象的新路径:"数据/算法—交互界面—事实幻象",事实变得随机、弥散、不确定和无常,并具体演化成为一种名曰"后真相"新幻象。

穿越欺骗性媒介的发展,媒介史已经昭示不同媒介以欺骗为策略,从而实现特定的媒介传播效果。实际上,自媒介诞生之初,欺骗性就已经开始表现为某些媒介固有的特征,潜藏在技术系统和媒介运作机制的始末,但是由于媒介技术生成的许多形式的幻象,与真实并没有泾渭分明的轮廓,它们更多的是潜在、弥散、寻常且不易察觉的,因此在某些特定的情况下,它们并没有被视为欺骗。显而易见,区别于人类在媒介再现中蓄意、恶意制造的错误和幻象,广播、照相机、电影、电视、计算机和VR/AR等媒介的欺骗性,最初并不是负面的、恶意的,并且人类也总是积极、主动地陷入各种眼花缭乱的欺骗之中。对此,基于以上种种情况,本文尝试提炼出"媒介的欺骗性"(deceitfulness of media)特征,即"媒介总是被嵌入具有欺骗性的技术手段,欺骗性因而成为媒介本体的一种特质"。具体来说,正是由于媒介的欺骗性的特征存在,才使得媒介在中介化现实的过程中,永远在制造幻象,永远无法复刻现实,其中涉及媒介使用符号、二进制代码、数据和算法等表征系统,向受众或用户生产、制造并传达幻象;另一方面,媒介通过技术和设备创造出种种欺骗手段和机制,利用人类感官特质和心理特征,通过制造幻象邀请人类进入由不同媒介建构的感官体验之中,并为人类营造出自主性幻象[22],从而实现媒介快速融入人类的日常生活。

诚然,"欺骗"在中文语境中是一个充满负面联想的词汇,常常被认为是不道德的行为,但本文坚持使用"欺骗"一词,来解释媒介制造的诸如错误再现、错觉、虚假等幻象的根源,主要基于以下考量:"欺骗"本身存在两个不同的层面:一是形式层面的欺骗,欺骗性是物种存在于世的重要特征,对应着一种生存技能,比如人类天然的具有欺骗性的特质,会说谎、会隐瞒,不存在绝对诚实的人类;再如动物通过假死、拟态和声东击西等欺骗行为,迷惑捕猎者或猎物,以此争取更多的生存空间,这便是遵循"优胜劣汰、物竞天择"的自然法则,此种欺骗是物种生而有之的能力,不过多涉及道德批判,而媒介的欺骗性也

首先隶属于形式层面的欺骗，以便媒介能更广泛地为人所接受和使用，这是媒介之于人类的调试，也是媒介得以长久生存的法则。二是规范性层面，涉及欺诈和诈骗等恶意欺骗的情况，规范性层面的欺骗必须受到道德批判和法律约束。显然，媒介本体的欺骗性特征更偏向于形式层面的欺骗，但这并不代表媒介的欺骗性始终是无害的。当权力结构寄身于媒介之上，伴随人类的深度介入，媒介的欺骗性则会从“形式层面的欺骗”位移至“规范性层面的欺骗”，最终对社会产生深远影响。从此角度来看，使用“欺骗”一词，既可以准确呈现和阐释媒介本体的此种特质，又便于进一步探究由媒介欺骗性引发的媒介幻象的文化后果。

三、人工智能的欺骗性与生成式AI幻象

沿着媒介欺骗史轨迹，我们“媒介的欺骗性”特征这一视角来观照新型媒介“生成式AI”，便可以发现生成式AI同样具有欺骗性，它与其他媒介技术的欺骗做法是一种延续和继承关系，且在制造虚假、错觉和幻觉等幻象方面，生成式AI已超越既往一切媒介。追溯构思人工智能的起点“图灵测试”(Turing Test)，生成式AI的欺骗性开始变得有迹可循。早在1936年，英国数学家艾伦·图灵(Alan Turing)就在《计算机器与智能》一书中提出测试“计算机是否具有智能”的方法，即“图灵测试”[23]。“图灵测试”采用问与答的模式，如果计算机能够非常好地模仿人回答问题，以至于测试者在充分的交流中误认为它是人而不是机器，就可以称这台计算机能够思维、具备智能。[24]从图灵设置的测试规则来看，“欺骗”是人机交互的必然结果。在人机对话中，只有成功骗过人类审讯者，机器才可以赢得“游戏”。换言之，“图灵测试”本质正是一种“欺骗游戏”，它在一开始便要求人工智能具备欺骗人类的能力。因此，带着这样的一种前提和测试标准，“欺骗”成为生成式AI运作的核心，就像电路、软件和数据一样，是生成式AI运行的基本条件。[25]最终，“欺骗”被深深地根植于生成式AI的所有媒介实践之中。

在这场欺骗游戏中，“如何表现得更像一个人”成为生成式AI的终极目标。有趣的是，生成式AI不仅学习人类利用语言说话和回答问题的能力，也同样继承了人类的缺点，但这并不代表生成式AI不够智能，反而说明生成式AI离它的终极目标更近一步。具体来说，在面对不会答或者不知道的问题时，以ChatGPT为代表的生成式AI会“说谎”，在生成文本的过程中“一本正经地胡说八道”。对此，人工智能研究界将其称为“AI幻觉”(AI hallucination)，具体指大语言模型(large language model，LLM)生成看起来是合理的或与真实的推论一致的错误内容，如错误的参考来源、内容和陈述，可能与正确的信息交织在一起，并以一种有说服力和可信度的方式呈现，使人类在没有仔细检查和事实核查的情况下难以识别。[26]AI幻觉包含两类：一是内在幻觉(intrinsic hallucination)，指AI生成的输出与输入内容相矛盾，这种错误可以通过校验输入内容轻松识别；二是外在幻觉(extrinsic hallucination)，指生成内容的正确性无法从输入内容中验证，这类内容通常是由模型调用了输入内容之外的数据或文本而产生的，其中甚至带有某种倾向性的言论和态度，后者往往更为常见且难以识别。[27]在生成式AI的媒介实践之中，“AI幻觉”在极长时间内将是技术难以克服的，这标志着生成式AI已经将隐匿的欺骗性手段，彻底融入OpenAI文本生产、流通和消费等各个环节之中，在生成式AI应用和实践所能涵盖的范围之内，事实和幻觉的辨别不再由中介机构承担，而是被下放到每个与之对话的用户个体，真假分辨变得有门槛。最终，谎言、欺骗将无处不在，“AI幻觉”成为生成式AI制造的第一种幻象。

不仅是ChatGPT会出现欺骗性的“AI幻觉”，围绕生成式AI所展开的其他媒介实践也带有欺骗性，并不断地制造出不同表现形式的幻象。在文生文的基础上，生成式AI又迭代出文生图、文生视频的大模

型，其中以生成式AI应用Midjourney、Sora为代表，此类大模型能够深刻地理解文字语言，只是根据用户提供的文本和关键词，生成式AI就可以自动生成逼真的图片或视频，具有涌现和模拟真实物理世界的能力，文生视频甚至能够创造出“真实”且符合现实物理的动态模拟影像，被称为“世界模拟器”(world simulator)。事实上，电子影像和数字影像是基于现实世界的再现，生成式AI的自动化图片和视频只是模拟而非再现现实，“世界模拟器”自动生成的影像实际是“真实的虚假”，或者说是“真实的幻象”，因此，生成式AI所生成的影像本身就是一种“幻象”。当然，ChatGPT所出现的“AI幻觉”，在Sora中同样没有得到解决，生成式AI文生视频的大模型依旧会出现一些事实性的错误，它还不能准确模拟许多基本交互的物理现象，导致视频中出现穿帮镜头和画面。此外，由于映入眼帘的模拟影像过于逼真，足以达到以假乱真的水平，如若不在生成视频中标注AI来源，或者不是细究影像细节，便很难识别机器生成的影像，而这也会加剧另一种幻象制造，即“深度伪造”(deep fake)。所谓“深度伪造”，是指利用深度学习等强大的人工智能技术来操纵或生成更具欺骗性的视觉和音频内容。随着生成式AI的介入，深度伪造将更加难以识别。由此，“在生成式AI之中，出现的种种错误再现、错误生成、虚假再现的情形以及具有欺骗性质的现象”，可统称为“生成式AI幻象”(generative AI illusion)。

四、自动化生产、交互界面与生成式AI幻象

生成式AI的欺骗性和幻象缘何而来？事实上，媒介的欺骗性特征是由特定技术赋予的，因而媒介幻象根源于媒介技术，并且基于媒介实践和技术物质性、功能性的差异，最终，不同的欺骗性媒介也会制造出不同表现形式的幻象。具体来说，电子媒介时代，广播通过电磁波或导线等媒介物，向受众传播声音形态节目，幻象随之进入人类的听觉系统。自照相机开始，媒介的幻象制造进入人类视觉系统，照相机是利用光学成像原理形成影像并使用底片记录影像的媒介物，复刻现实成为照相机最重要的功能，但此种复刻是碎片化、片面化的呈现，不同的拍摄角度和拍摄方式也都可能造成虚假并生成幻象。在静态图像的基础上，电影和电视分别通过放映机、扬声器和屏幕等媒介物，结合拍摄视角、剪辑等特定的操作手段，在制造出连续动态影像的同时生成非连续性的幻象。数字媒介时代，计算机和智能手机等PC端被嵌入各式各样的代码和算法，形成隐秘的软件系统，基于不可见的算法自动化的装置，所有的欺骗机制均被隐匿于交互界面之后，幻象跃然界面之上。总体来看，电子媒介时代的媒介欺骗性与幻象制造，主要依赖媒介的物质性基础以及特定的操作手段和运行机制，而在此基础上，数字媒介则拥有更为复杂的软件系统，软硬件系统共同决定着其制造幻象的方式。

同样，进入人工智能时代，“AI幻觉”“模拟影像”和“深度伪造”等生成式AI幻象的出现也并非偶然，构成生成式AI欺骗性的技术涉及软硬件两套系统，即基于大数据、多模态大模型、自然语言处理和神经网络算法等实现自动化生产的软件系统，以及交互界面等硬件设备。聚焦生成式AI庞大且复杂的软件系统，即从技术逻辑和运行机制来看，不论是文生文、文生图，还是文生视频，生成式AI首先是建立在对人类元语言理解之上的，在人机交流中，如果元语言的理解出现障碍，便会导致交流的失败。[28]基于海量文本的原始数据，通过字符串(token)将各种形式的文本代码、数学和自然语言进行统一，其后，结合多模态大模型、优化神经网络算法和超强算力，利用深度学习方法对生成式AI展开训练，在此过程中，人类参与生成式AI大模型的监督训练，通过“人工数据标注”工作，针对字符串文本输出设定惩罚和奖励机制，进而调优预训练的语言模型，最终，以ChatGPT

为代表的自然语言处理大模型，具备模拟、理解人类意义及意向性能力，成功实现人机对话。而以Sora为首的大视觉模型(large vision model，LVM)，则通过补丁(patches，训练生成各种类型视频和图像模型的可扩展且有效的表示，对应ChatGPT的token)将视频和图片分解为最小的数据单元，再结合上述的大模型训练系统实现了类似的效果。

然而，海量数据集的来源存在事实性错误或结构性偏见，生成式AI并没有自主判断真实与虚假、公平与偏见的能力，这将影响输出文本和影像的准确性、真实性和客观性。首先是训练生成式AI原始数据集的真实性无法得到根本保证，即使数据集来自于相对权威的网站、机构或智库，事实性错误和结构性偏见依旧难以避免。其中，事实性错误能够通过事实核查的方式解决，但不能完全消除。相比之下，结构性错误则更为隐秘且难以感知，比如由于ChatGPT数据源中的英文资料占比最高，那么，大模型便会对英文建构的世界产生结构性偏向。因此，将事实性错误或结构性偏见直接投喂给大模型，必定会使其出现“AI幻觉”，甚至出现隐含歧视性、偏见性的输出内容。仅凭有限的人工数据标注工作，就想彻底消除海量数据的错误以及由此产生的AI幻觉，显然不切实际，甚至监督训练也会误导大模型，因为人类同样存在结构性偏见。简而言之，训练大模型的原始数据集为生成式AI产生“AI幻觉”等幻象埋下伏笔。

基于字符串和补丁的文字编码、影像编码和机器深度学习，生成式AI的自动化生产变得“智能”，但实际上，这只不过是一种基于概率计算组合并围绕文字、图像和视频展开的“接龙游戏”，以及基于人类语言和物理世界的“模拟游戏”。首先，就以ChatGPT为例的大语言模型而言，GPT(生成性预训练转换器，Generative Pretrained Transformer)是一种基于Transformer网络的自然语言处理模型，相比于其他深度学习的神经网络，通过多层自注意力机制(self-attention)和前向神经网络(feed-forward neural network)，GPT能够计算序列中各个词句间、各个位置间的关联，并对上下文信息进行捕捉。最终，通过深度学习海量文本数据，生成式AI能够根据一连串词语计算、预测下一个词、下一个句子和段落，并保证上下文衔接的相关或因果关系，即具备对人类说话方式的概率进行判断和预测的能力，以此实现生成式AI对人类语言能力的模拟。简单来说，生成式AI的内容生产逻辑就是用统计上最高概率的词来填补空白，它实际上并不可能真正理解词句和上下文之间的意义，更不用说其中的因果关系，因此，大语言模型只是一种基于模拟和概率接龙的“文字游戏”。生成式AI的软件系统更擅长于生成与真实事物、物理世界表面近似的答案，但它的回答与事实内容很容易发生断裂，从而使得所答非所问，其中的错误再现和幻象在所难免，且由于拥有庞大的计算参数和强大的算力，生成式AI知识在广度上超越任何一个人当用户没有相关领域足够的知识时，就会被错误的内容生成所欺骗，并对这些看似具有说服力的内容信以为真，但是对某个领域的专家而言，生成式AI所给出的回答常常充满荒谬。

其次，以Sora为代表的大视觉模型，其技术的底层逻辑依旧对应着一个概率模型。针对Sora的训练是通过Diffusion Transformer(扩散转换器)的架构来处理图像和视频的潜在代码的，且能在因果关系中处理时空互动。借鉴大语言模型中将文本转换为字符串的思路，Sora通过补丁实现多样化图像和视频内容的生成。因此，相比于大语言模型，大视觉模型只是在机器深度学习的过程中，将词句、上下文之间的关联性、因果性计算，拓宽成对物理规律和物理世界的计算，以学习视频画面的整体与局部、前后帧统一的像素级、物理与时空等维度的变换规律和表征间的相关或因果关系。同样，大视觉模型也能理解

人类语言，Sora整合GPT和DALL·E 3(文本提示生成图像)的重新标注技术，通过训练一个高度描述性的标题模型，在语言和视频、图像之间建立关联。最终，在人机对话框中，生成式AI对人类输入的提示词进行实时递归式的判断和估计，依据概率自动生成具备文本场景、色彩、时空、动作等信息的结构化图像或视频产品，最终实现对物理世界的深度模拟。但生成式AI对物理世界和外部世界的理解不是一个连续的完整体系，也没有真正推理不同概念与实体之间关系的能力[29]，因此，大视觉模型对现实世界的物理、因果关系和物体持久性的理解始终存在断裂，其输出内容同大语言模型一样经不起推敲。更深层次来看，基于概率和模拟的自动化生产的软件系统，只是使生成式AI变得更为“拟人”，“优先取悦人类反馈，而非基于事实逻辑实现内容生成”的原初设定，使人类极易陷入软件系统所设定的互动圈套。“事实让位于互动”预示着“内容”将变得极其廉价，而“真实”将会彻底成为一种稀缺资源。

聚焦生成式AI的硬件系统，即从物理组成和物质基础来看，人类与生成式AI的对话、交往借由交互界面实现。用技术术语来说，界面是计算机的硬件和软件内部或相互之间进行交互的地方，但它通常被描述为使计算机能够与用户交互的设备或程序。[30][31]事实上，从数字媒介开始，交互界面就极具迷幻性。作为一种制造连续性和可见性的欺骗性技术，交互界面将生成式AI的欺骗系统隐匿其后，给用户迷上一层幻象，这样用户就无法感觉到与生成式AI软件系统的“摩擦”，所有的欺骗机制变得透明。因此，建立在数字媒介交互计算系统基础上，延伸发展的生成式AI交互界面，并不是摒弃欺骗，而是将欺骗融入交互界面的设计和系统之中，欺骗与交互界面共生。在基于界面的人-机互动系统中，技术物既自我生成，同时又生成缔合环境。[32]表面上，用户通过交互界面实现生成式AI的操纵，实则也被生成式AI整合到它的软件系统之中，最终，生成式AI为用户制造出自主性的“幻象”。总体来看，基于概率和模拟的自动化生产的软件系统，构成生成式AI复杂且庞大的欺骗机制，具有物质性属性的交互界面将此种欺骗隐匿其下，至此，生成式AI的软硬件系统成为制造种种AI幻象的根源。

五、生成式AI幻象与超真实建构

鲍德里亚将“拟像世界”分为四个等级或序列：仿造(counterfeit)、生产(production)、仿真(simulation)和分形(fractal)。第一序列的拟像是从文艺复兴到工业革命“古典”时代，其显著特征是“仿造”，遵循自然价值规律，即“使用价值”。比如，彼时对艺术真品的复制只通过手工临摹完成，真品具有独一无二的客体身份，原作之外的仿造均是赝品，凭借人类经验能够细微感知客体与符号、现实与幻象的差异。在口语文字时代，媒介制造的幻象属于第一序列的拟像，其是对某种基本真实的反映。第二序列的拟像是工业时代，其显著特征是“生产”，遵循商品价值规律，即“交换价值”，机械复制的技术成为艺术生产的一种新方式或新媒介，机械复制使得艺术品的价值被低估，因为其可以被无限复制，最终导致灵韵丧失，原作与赝品之间的从属关系被扭转，客体和符号成为等同物。印刷和电子媒介制造的幻象对应着第二序列的拟像，其中的幻象开始掩盖和篡改某种基本真实。第三序列的拟像是当下受二进制代码系统支撑的时代，现代的符号生产阶段最显著的特征是“仿真”，遵循结构价值规律，即“符号价值”，“拟像”亘古有之，而“仿真”则是后现代社会的产物。所谓“仿真”，是指一种不以客观现实为基础但又极度真实的符号生产和行为过程。[33]“仿真阶段”意味着客体不再只是通过机械复制的方式被复制，借助二进制代码和算法，客体能够被计算机模拟生成，即二进制代码正在生产真品，客体开始被符号取代。当下，互联网媒介制造的幻象正对应着第三序列的拟

像，它正试图通过一系列的技术手段掩盖某种基本事实的缺席。

第四序列是鲍德里亚在《恶的透明性》一书中提出的，是对前三种拟像序列的补充，他认为拟像秩序已经进入“价值的碎形”的新阶段，其显著特征是“分形”。相对于其他的三个拟像序列，第四个序列的拟像界定并没有那么清晰，但从鲍德里亚对拟像分形阶段的论述来看，“拟像分形阶段”的特点大致可以概括为以下几个关键词：无限自我增殖、真实消失、价值形散和差异终结。具体来说，无限自我增殖是拟像走向分形的起点，亦是分形区别于仿真的最大特征。鲍德里亚将分形阶段的拟像比喻成细胞分裂或者病毒传播，是一种邻接性的(contiguous)、转喻式的(metonymic)、无性的自我增殖(proliferation)，“事物通过增殖、传染、饱和或透明化、衰竭或灭绝，通过模拟的泛行，在模拟中转为次级存在。而其结果，不是一种宿命般的消失，而是一种分形的散布”[34]。换言之，仿真阶段的拟像虽然告别机械复制可以生产真品，但是这种生产是有限度的，还未形成如分形一般的自我无限增殖模式。如果将仿真阶段的二进制代码和算法比喻为“黑箱”，那么，分形秩序便是这些“黑箱”正在不断地自我复制，是对仿真的不断增殖。

在分形阶段，符码按照分形系统的内在逻辑，永无止境地分裂、再生和扩散，客体完全被符码取代，客体是什么变得不再重要，它只是作为一种养料投喂给符码系统，以便符码系统能更完美地替代客体。随之而来的是真实的消失，再没有什么能得到真实的反映。其后，真实的消失直接导致价值的形散。各种网络中病毒式(virale)散布的逻辑，不再有所谓的价值法则，价值规律已经不存在了，因此也不是有关等值的逻辑，存在的只是某种价值的流行或转换，以及价值的随意性增值和扩散，“每一个价值或是价值的碎片，在模拟的天空[35]——中闪耀片刻，画出一条折线，几乎不与其他的线相交，随之遁入虚空。这就是分形化的模式，也是目前我们的文化所具有的形式”[36]。最后，分形文化意味着差异终结：所有的系统和所有的个体都在四处扩散，为了能够实现无限增殖、发展到最大的普遍化阶段，它们必须摆脱自身的本质与异质性，并被其他范畴吸收，才能完成“从特殊走向普遍”的统一。最终的结果是分形文化将我们带入超政治、超性别、超审美的时代，但在同一时刻，当一切都是政治的、性的、审美的，就意味着没有什么是政治的、性的和审美的。在这种存在于所有个体、事物、层次和系统的相互渗透中，一种主要的力量是媒介以及更为普遍的、一般性的传播。

虽然拟像的第四序列在推及整个社会的过程中可能会存在争议，但具体到人工智能媒介所制造的AI幻象之中，具有一定的解释力。与数字媒介相同，生成式AI“黑箱”以二进制代码和算法为基础，但是人工智能“黑箱”的复杂程度完全超越传统数字媒介，且生成式AI的“黑箱”已经进化出自动化生产模式，在万亿个参数的支撑下，通过多模态大模型、自然语言处理和神经网络深度学习等手段，生成式AI的“黑箱”正通过自动化生产实现无限生成，它可以根据人类的问题，源源不断地给出回答，这正对应分形阶段所说的“自我复制”和“无限增殖”。生成式AI试图通过吸纳海量大数据所描摹的现实世界，达到像现实世界一样永无声息地增生、扩散的状态，但其结果经常是混乱不堪、无法预料的，生成式AI幻象的出现正对应着这种结果。事实上，生成式AI的软硬件系统已经决定其无法准确回答问题，事实和真实在它那里并不重要，更重要的是能够反复实现人机对话与互动，只有这样生成式AI才能够保持“无限自我增殖”的状态，如果只是重复回答“我不知道”，或者只是直接读取数据库中的内容，那么，它与搜索引擎本质上并无差异。因此，分形阶段的第二种特征“真实消失”同样反映在生成式AI的自动化生产过程之中。

显然，生成式AI幻象已逐渐成为一种新的文化现象，其是当代“分形文化”的一种外在表现形式，意味着“价值形散”和“差异终结”。在生成式AI的分形文化中，不再存在真与假、美与丑、善与恶等二元参照对象，没有什么再能通过横纵坐标的参照系得以定位，价值通向各个方向和一切空隙，价值也不再需要任何参照，只需要顺从邻接性，以病毒式分裂和增殖的方式扩散开来。首先，面对的便是“真与假”的差异终结，或者说真与假的区分在生成式AI的意义场中变得不再重要且没有任何意义，生成式AI内容生成的最大魅力恰巧在于其“制造幻象”的能力，这正对应着机器的创造力和想象力，因而“制造幻象”是生成式AI媒介实践的重要组成部分。“生成内容而非真实”的程序设定已经注定“真实让位于幻象”的结果。在生成式AI幻象之中，“AI幻觉”意味着事实经不起推敲，Sora的出现代表着从“看见的世界”走向“模拟和制造的世界”，“模拟影像”中的画面都是模拟生成的，没有现实来源，“模拟影像”直接彻底消灭真实，而“AI深度伪造”更是直接制造虚假。至此，生成式AI制造的幻象本身就是“纯粹的拟像”，形成的是一种基于概率和模拟的“分形文化”。

鲍德里亚并未对分形阶段的“拟像”与“超真实”之间的关系进行论证，但我们可以先回到仿真文化的语境中，理解“超真实”概念本身，进而再探究生成式AI中“分形文化”所表现出的“超真实”。在谈及仿真文化的后果时，鲍德里亚认为，仿真产生了一种普遍“超真实”的幻象，即“真实与非真实间的区别模糊不清，非真实超过了真实，比真实还要真”，其乃是仿真的特性，是许多拟像共同组成的一种新的现实次序。[⑨]此时的真实已经不再是类似风景或海洋的自然之物，它还涉及人为生产或再生产的“真实”，而真实本身却在超真实中沉默。进入拟像的分形阶段，“超真实”从未消失，其亦是分形文化的一种结果和状态，并且较之仿真阶段，分形阶段的“超真实”将走得更远、更彻底，存在终结“真假差异”的可能性。诚然，大众媒介和数字媒介属于仿真文化的阶段，它们所创造的幻象同样会面临“超真实”的问题，比如大众传播媒介中的电影被视为“幻觉媒介”，尤其在科幻题材的电影中，超真实被演绎得淋漓尽致；而数字媒介的后真相也是一种“超真实”。

分形阶段的生成式AI幻象较之前者的根本性区别是：生成式AI将彻底夷平“真与假”之间的差异，即真假的区分变得没有意义。事实上，在大众传播和数字媒介所创造的媒介环境中，我们始终能够感受到来自“真实”与“虚假”之间的冲突和博弈，此消彼长。具体而言，大众媒介的真实和虚假均是由少数精英群体所操纵，但由于传统媒体新闻机构垄断着话语权，唯一、理性和确定的真实能够被公众所知晓，真实在表面上战胜着幻象，真实与虚假之间尚存在边界。而数字媒介的出现扭转了这一局面，在无数大众涌入传播网络后，真实变得极具不确定性，大众媒介时期所强调的“唯一真实”被众声喧哗的幻象淹没和吞噬，后真相凌驾于真相之上，即幻象比真实更普遍，真实与虚假共生。生成式AI则更进一步，在生成式AI建构的幻象宇宙里，真假与否的内容生成变得不再重要，更为重要的是“围绕生成式AI和幻象所展开的各种媒介实践及其构筑其上的全新的文化形式本身”，它可能将我们引入一种“乐此不疲地制造虚假、制造幻象”的媒介环境与新文化形式之中。而当生成式AI越来越多地介入人类日常生活，人类的境况是不断地被包裹在生成式AI所建构的幻象宇宙之中，分不清真假，也不在乎真假。“真假反转”和“真假倒置”的现象会比数字媒介时代的后真相更为普遍，最终的结果是“假作真时真亦假”。

生成式AI幻象之于大众媒介和数字媒介幻象的第二个区别是：生成式AI的内容生产主体是机器，生产机制是基于代码、数据和自然语言等技术物的无限自我增值。一方面，机器内容生成意味着制造幻

象的效率更高、成本更低，在具体的数量表现上，幻象的生成和制造将进入前所未有的阶段，人类将越来越多地被生成式AI幻象所包裹。另一方面，就生产机制而言，当基于代码和算法的计算机模拟系统，转变为算据、算力和算法更强的人工智能系统后，基于概率和模拟的自动化生产对真实世界的模拟与再生产变得更加事无巨细，“本真性”作为总体性媒介观的经验基础被彻底消解，不只是真实在被模拟，“超真实”本身也在无限模拟、增殖和扩散。最终，生成式AI的幻象打破数字媒介“真假共生”的幻象，表现出“以假乱真”和“无中生有”的“超真实”特征。

构筑在“真与假”差异的终结之上，在生成式AI幻象的分形文化中，更糟糕的后果是出现“善与恶”的差异终结。本文在选择使用“欺骗性”概括媒介的一般性特征时，就提及“媒介的欺骗性本身属于形式层面的欺骗，但其孕育着规范性欺骗的种子”。事实上，就形式层面的欺骗而言，由生成式AI产生的“AI幻觉”“模拟影像”和“深度伪造”等，正对应着它的“创造力”和“想象力”，人类在面对自己所不知道的问题时，会调动自己既有的知识体系来尝试回答问题，生成式AI亦是如此，它的魅力也恰恰源于这种“AI幻象”。对此，本文并不想从根源上否定生成式AI幻象，或者直接将其武断地归类为负面现象，这有悖于“媒介的欺骗性”内涵界定中反复强调的媒介欺骗性的中立属性。但是，这也并不意味着就可以肆无忌惮地制造生成式AI幻象，我们同样要看到它所带来的潜在风险。对此，需要警惕的是，当人类和权力结构介入生成式AI幻象的制造时，生成式AI的欺骗系统可能会助纣为虐，“生成式AI幻象”将从“形式层面的欺骗”走向“规范性层面的欺骗”，变得极具道德和法律争议，最为直接的例子就是“深度伪造”。

不同于传统数字媒介的“PS技术”，通过人类给定的素材，深度伪造在“自主性伪造”的基础上实现“有中生无”，并且存在滥用人类面部、声音、血型、指纹和DNA等个人生物识别信息的风险，动态的伪造影像使得蓄意、恶意的欺骗甚嚣尘上，这将对人类社会造成巨大威胁。至此，真实的消失、恶意的欺骗，非黑即白的“善恶之分”全在人与机器的一念之间，而这将会引发社会信任关系的全面崩塌。即使文字、图像和影像均是真实的，也可能会引发人类对其真实性的猜忌，因为什么看着都是真的，也就意味着什么都不是真的，曾经合一的“真实”与“信任”如今开始走向分离，何为真实并不重要，重要的是什么是我想要的，什么是我选择相信的。而这也同样警示我们，技术之恶的潜能已经被激发了，生成式AI的欺骗性很有可能被恶意利用而走向极端，我们在接受生成式AI幻象创造力和想象力的同时，必须揭露那些不恰当的、恶意的生成式AI应用，以抵抗公众或私人通过数据和算法对我们的深层控制。

结语

在人类文明史上出现的众多媒介都天然地具有欺骗性，这是媒介技术固有的特征。然而，在既往的有关“真实与媒介再现”的研究中，媒介再现中所出现的幻象和种种欺骗的结果往往被视作人为的，即人类利用技术制造幻象，欺骗只是一种人类行为，媒介与技术在此过程中保持中立状态，这是基于人类中心主义的论断，往往忽视媒介本身的能动性。当下，生成式AI无疑是具备主体能动性的非人行动者，而生成式AI幻象的出现也恰恰提示我们，应当突破人类中心主义的视角，重思真实、幻象与媒介再现之间的关系，从而深入探讨新媒介变迁所招致的新的文化后果。事实上，媒介再现过程中所出现的幻象并不只是单纯地由人类造成的，而是与媒介本体所存在的欺骗性特征的共同作用，才使得媒介在中介化现实的过程中制造幻象。作为新型媒介的生成式AI同样具有欺骗性，且在制造幻象方面，生成式AI已超越既往一切媒介，具体表现为“AI幻觉”“模拟影像”和“深度伪造”等形式。

从历时的层面来看，每一次技术的跃迁都会带来幻象制造的升级。相比于既往的电子媒介和数字媒介，生成式AI的软件系统使得欺骗机制和幻象制造变得更为复杂。未来，生成式AI也必定是走向全息化的。当更多的人类感官被邀请加入其中，生成式AI所制造的幻象就不仅仅是“眼见未必为实，耳听未必为真”，人类将会面临全方位、立体化的感官欺骗。随之而来的是，生成式AI建构起一个“超真实”的世界，生成式AI幻象形成一种新的文化现象，对应着鲍德里亚所预言的“分形文化”。最终的文化后果是，价值的形散，“真假与否”变得不再重要，“善与恶”差异终结。生成式AI幻象中将充满“以假乱真”和“真假倒置”的情形，一言以蔽之：假作真时真亦假。同时，技术之恶的潜能被前所未有地激发，生成式AI深度伪造存在，引发个体生物识别信息滥用的风险。当真假再难进行区分之时，当技术之恶和人类之恶相遇时，社会即将面临全面的信任危机。如何在生成式AI中破除恶意幻象、重识真实，如何钳制技术之恶、实现技术向善治理，也将成为人工智能时代的重大命题。

注释：

①目前，ChatGPT的用户页面已经明确标注：“ChatGPT也可能会犯错。请核查重要信息。”

②陆杰荣：《论哲学“真实”之规定》，《哲学研究》2003年第6期，第12-18页。

③[德]海德格尔：《存在与时间》(修订译本)，陈嘉映、王庆节译，北京：生活·读书·新知三联书店，2012年，第38页。

④贺来：《“关系理性”与真实的“共同体”》，《中国社会科学》2015年第6期，第22-44页。

⑤王晓升：《现代性、现代主义和后现代主义——概念的梳理》，《华中科技大学学报(社会科学版)》2017年第5期，第1-8页。

⑥M. Barnard, Graphic Design as Communication, New York: Routledge, 2005, p. 139.

⑦[法]居伊·德波：《景观社会》，王昭风译，南京：南京大学出版社，2006年，第3页。

⑧[英]Dorlling Kindersley：《牛津英语图解大词典》，北京：外语教学与研究出版社，1999年，第773页。

⑨[法]让·鲍德里亚：《消费社会》，刘成富、全志钢译，南京：南京大学出版社，2014年，第17页。

⑩马小茹：《“超真实”概念探析》，《哲学分析》2018年第5期，第115-134页。

⑪J. Baudrillard, Symbolic Exchange and Death London: Sage, 1993, p. 74.

⑫[英]大卫·麦克奈恩：《理解电视：电视节目类型的概念与变迁》，苗棣、赵长军、李黎丹译，北京：华夏出版社，2003年，第139页。

⑬罗钢、刘象愚：《文化研究读本》，北京：中国社会科学出版社，2000年，第20页。

⑭[意大利]西蒙尼·纳塔莱：《媒介的欺骗性：后图灵时代的人工智能和社会生活》，上海：复旦大学出版社，2023年，第11页。

⑮[美]希伦·洛厄里、梅尔文·德弗勒：《大众传播效果研究的里程碑》，刘海龙等译，北京：中国人民大学出版社，2009年，第29页。

⑯F. Kittler, Optical Media. A. Enns, Trans, Cambridge: Polity Press, 2010, p. 166.

⑰车致新：《“想象界”的物质基础——基特勒论电影媒介的幻觉性》，《电影艺术》2018年第4期，第94-100页。

⑱F. Kittler, Gramophone, Film, Typerwriter M. Wutz, Trans, Stanford, California: Stanford University Press, 1999, p. 119.

⑲傅懿瑾：《媒介幻象：媒介化视野下手机观影主体的纵向研究》，《中国电视》2022年第5期，第54-59页。

⑳胡泳：《论事实：美丽、谎言与ChatGPT》，《新闻大学》2023年第4期，第38-58页。

㉑R. Mansell, Imaging the Internet. Oxford: Oxford University Press, 2012, pp. 108-115.

㉒笔者注：自主性幻象是指，人类自以为可以自主地使用和操纵媒介，即人类之于媒介拥有绝对的统治权和控制权，但实际上，这种控制权只是人类的一种错觉，媒介通过欺骗性机制和手段，反向使人类沉溺于媒介使用之中。自主性幻象，亦是媒介欺骗性的一个重要面向，但不是本文的重点，故不重点展开讨论。

㉓A. M. Turing, "Computing machinery and intelligence", Mind, New Series, Vol. 59, No. 236, 1950, pp. 433-460.

㉔冯志伟、张灯柯、饶高琦:《从图灵测试到ChatGPT——人机对话的里程碑及启示》,《语言战略研究》2023年第2期,第20-24页。

㉕[意大利]西蒙尼·纳塔莱:《媒介的欺骗性:后图灵时代的人工智能和社会生活》,汪让译,上海:复旦大学出版社,2023年,第2页。

㉖S. Bubeck et al., Spark of Artificial General Intelligence: Early Experiments with GPT-4., http://arxiv.org/pdf/2303.12712, 2024-01-03.

㉗Y. Bang et al., A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity., http://arxiv.org/abs/2302.04023, 2024-01-03.

㉘刘海龙、连晓东:《新常人统治的来临:ChatGPT与传播研究》,《新闻记者》2023年第6期,第11-20页。

㉙方师师、唐巧盈:《聪明反被聪明误:ChatGPT错误内容生成的类型学分析》,《新闻与写作》2023年第4期,第31-42页。

㉚K. Jorgensen, Gameworld Interfaces. Cambridge, Massachusetts: The MIT Press, 2014, p. 4.

㉛B. Hookway, Interface. Cambridge, Massachusetts: The MIT Press, 2014, p. 14.

㉜胡翼青、谌知翼:《作为媒介性的生成性:作为世界的媒介》,《新闻记者》2022年第10期,第3-15页。

㉝赵元蔚、鞠惠冰:《鲍德里亚的拟像理论与后现代消费主体》,《社会科学战线》2014年第1期,第158-162页。

㉞[法]让·鲍德里亚:《恶的透明性:关于诸多极端现象的随笔》,王晴译,西安:西北大学出版社,2019年,第4页。

㉟笔者注:有趣的是,生成式AI"Sora"取自日语"天空",具有极强的隐喻。鲍德里亚在具体论述分形时提到"价值形散在模拟的天空之中",彼时自然没有Sora,这种历史性的预言与Sora命名的汇聚固然带有巧合的成分,但不可否认的是,在鲍德里亚写作的那个历史节点,他就已经看见Sora的诞生。

㊱[法]让·鲍德里亚:《恶的透明性:关于诸多极端现象的随笔》,王晴译,西安:西北大学出版社,2019年,第14页。

㊲[美]马克·波斯特:《第二媒介时代》,范静哗译,南京:南京大学出版社,2000年,第91页。

The Creating Logic of Generative AI Illusions and the Cultural Ramifications of Hyperreality Construction

Jing Yulun　　Zhang Dianyuan

Abstract: In the process of mediating reality, because of the deceitful features of media, media are always creating illusions, forever unable to replicate reality accurately. As a new type of media, generative AI similarly possesses this deceitful quality, and in terms of creating illusions, generative AI has surpassed all previous media forms. This is specifically manifested in forms such as "AI hallucinations", "simulated images" and "deep fakes". The software and hardware systems of generative AI become the root of creating various AI illusions. Generative AI has constructed a "hyperreality" world, with its illusions aligning with the "fractal culture" as predicted by Jean Baudrillard. The ultimate cultural ramification is that in the illusory universe constructed by generative AI, the distinction between true and false no longer matters. It ushers us into a media environment and a new cultural form where we are "enthusiastically engaged in creating illusions." The potential for the malice of technology could also be unprecedentedly stimulated.

Key words: generative AI; illusion;deceitfulness of media; hyperreality; fractal culture