



# 元宇宙的伦理风险与应对

田少波,田 野

(武汉理工大学马克思主义学院,武汉 430070)

**摘 要:** 2021 年元宇宙作为资本和科技公司合力推动的热门话题,引起了社会高度关注。关于元宇宙的本质、价值和伦理风险尚处在分析界定与阐明之中。从理论渊源来说,元宇宙是一个平行于物理世界的一个社会化赛博世界;从技术本质来看,元宇宙是整合多种技术产生的下一代互联网应用和社会形态;在价值论方面,元宇宙的五种“超越”构建了一种虚实联动、平行互动的“第三世界”新形态,其发展是科技和人类文明的必然选择;从伦理风险角度来说,元宇宙会遭遇科林格里奇困境、去中心化、精神迷失、“化身”、虚拟侵害的风险。为了应对元宇宙潜在的伦理风险,注重人文精神的注入和负责任创新是迫切的选择。无论元宇宙最终走向何处,价值与伦理的反思是必要的。

**关键词:** 元宇宙;伦理;伦理风险;技术风险;技术伦理

**中图分类号:** B82-057

**文献标志码:** A

**文章编号:** 1673-3851 (2023) 12-0687-09

## Ethics risk and response in the metaverse

TIAN Shaobo, TIAN Ye

(School of Marxism, Wuhan University of Technology, Wuhan 430070, China)

**Abstract:** In 2021, the metaverse, as one of the most popular topics promoted by capital and tech companies, attracted widespread attention in society. The nature, value and ethical risks of the metaverse are still under analysis, definition and elucidation. From the theoretical origin, the metaverse is a cyber world parallel to the physical world; from the essence of technology, the metaverse is an internet application and a new social form that combine multiple forms of technology; in axiology, the five "transcendences" of the metaverse build a new form of "the Third World" with virtual and real linkage and parallel interaction and the development of the metaverse is an inevitable choice for technology and human civilization; from the point of view of ethical risk, the metaverse will encounter a series of difficulties, such as Collingridge's dilemma, decentralization, mental disorientation, avatars and virtual infringement. Therefore, it is an urgent choice to propose a response path of humanistic spirit injection and responsible innovation in response to potential ethical risks. No matter what the ultimate development of the metaverse will be, reflection on values and ethics is necessary.

**Key words:** metaverse; ethics; ethics risk; technology risk; technical ethics

2021 年 Facebook 推出“Meta”元事业,一时掀起“元宇宙”风暴。元宇宙在科技公司和资本的合力推动下成为了热门话题。因此,人们称 2021 年为元宇宙元年<sup>[1]</sup>。关于元宇宙的含义以及它试图表达的

内在本质和图景,曹刚<sup>[2]</sup>认为元宇宙建立了一种更广泛的“在场”关系,一是开放性,二是具象性;同时元宇宙展开生活空间的引领性内容是主体自主性:一是人类进入元宇宙的主要方式是数字分身,二是

收稿日期:2023-07-11

基金项目:中央高校基本科研业务费专项资金(20231V013-012)

作者简介:田少波(1971—),男,湖北仙桃人,研究员,博士,主要从事科技哲学方面的研究。

人类超越了社会现实身份带来的局限和束缚,三是元宇宙提供了一个开放创造的平台。喻国明等<sup>[3]</sup>从联系的视角阐明,元宇宙实现了现实世界和虚拟世界的革命性连接,基于对现实世界的超越以构造出更高层次的现实世界。但是,上述观点缺乏对元宇宙的理论渊源和技术架构以及其技术演变的内在逻辑的分析。

在元宇宙潜在的伦理风险方面,谢程远等<sup>[4]</sup>认为元宇宙可能会引发人类在虚拟世界的迷失和摆脱人类控制的问题。王建红<sup>[5]</sup>指出,沉浸感强烈所造成的成瘾性,现实世界的匮乏和虚拟世界感官刺激满足的反差,形成的“补偿效应”可能引发精神迷失。当前,学界对元宇宙伦理风险方面的思考主要围绕在人的精神或心理方面,缺乏对去中心化、虚拟“化身”和侵害风险方面的讨论以及伦理风险应对路径的思考。因此从哲学的视角洞见元宇宙的未来,重新思考存在、自我、价值、伦理等问题是必要的。马尔库塞在《单向度的人》中提到:“当技术成为物质生产的普遍形式时,它就制约着整个文化,它设计出一种历史总体——一个世界。”<sup>[6]</sup>如此“一个世界”的图景在未来试图被元宇宙所占据,而通达未来的道路需要首先把握住当下,亦即人们期望中的元宇宙仍有许多哲学层面的问题有待思考与解决,否则所构造的不过是一个“乌托邦”。本文试图从哲学的视角探讨元宇宙技术本质层面的含义,为什么我们需要元宇宙,分析元宇宙可能潜在的伦理风险,针对可能存在的伦理风险进行反思并提出风险应对路径。

## 一、元宇宙的渊源与内涵

由于学科知识背景和哲学立场不同,对元宇宙的理解呈现多元结构。从技术发展视角来看,元宇宙是诸多技术汇聚和整合的产物,是技术发展到一定程度的呈现。技术发展内生的变革力会改变经济和社会的发展,对生产过程中人与人的关系产生质的影响,这是一个客观规律<sup>[7]</sup>。

### (一)元宇宙的理论渊源

元宇宙的最早构思来自莱布尼茨的“命运之宫”。莱布尼茨在《神义论》中不是用技术的手段,也不是用哲学的逻辑推理,而是用寓言(朱庇特、塞克斯图斯和忒奥多鲁斯的故事)的方式来描述一个元宇宙系统,他称之为“可能世界”<sup>[8]</sup>。在莱布尼茨的描述中,可能世界的整体安排非常像计算机程序,尤其像虚拟实在的系统程序。

元宇宙的早期文化源头来自1992年由美国科

幻作家斯蒂芬森(Stephenson)创作的小说《雪崩》,“Metaverse”一词正是来自这部著作<sup>[9]</sup>,实际就是一个平行于物理世界的社会化赛博世界。元宇宙一词英文为“Metaverse”,形而上学一词英文为“Metaphysics”,“形而上者谓之道,形而下者谓之器”,元宇宙想要表达的是相对于物理世界而建构的一个数字平行宇宙,在真实宇宙和虚拟宇宙之间来回穿越,实现虚实相交。

元宇宙的早期概念源头来自赛博空间(Cyberspace),是控制论和空间的组合,而后催生的虚拟现实(Virtual Reality)在1984年被提出,钱学森先生在书信中将虚拟现实技术翻译为更符合中国传统文化的“灵境”(Virtual Reality)<sup>[10]</sup>。赛博空间催生了“镜像世界”的软件技术概念,即把真实世界的场景都投影到软件程序中,通过显示器与镜像世界中的“我”进行互动。1991年耶鲁大学计算机科学教授Gelernter出版了著作《镜像世界》(Mirror Worlds),他认为互联网的终极形态就是“镜像世界”——物理世界的虚拟映射,就像是一个小镇倒映在平静的湖面上,但对不同的观者,它夹杂了每个人不同的生命体验,倒影中包含了人们真实生活中的社会、机构和家庭结构<sup>[11]</sup>。“镜像世界”从哲学上回答了主体性内涵分布式智能在沉浸式虚拟环境中的感觉性存在(是一种特殊的 reality),并在交互条件下形成了社会性。

### (二)元宇宙的技术架构与概念内涵

元宇宙的技术初始概念——赛博空间,历经了视觉耦合机械模拟器VCASS(1953—1984年)到VR技术(1984年),可见元宇宙技术不是凭空出现的,而是具有历史性和可追溯性。

从元宇宙的技术架构来说,元宇宙需要六种技术来支撑。一是区块链技术——基于智能合约,去中心化的结算平台和价值传递机制,保障价值归属与流转(对等网络架构);二是人工智能技术——支撑基于价值、决策和评价的一个学习过程;三是电子游戏技术——图像、声音、动画作为资源的游戏引擎和起支撑性的程序代码;四是交互技术——实现人机对话,形成沉浸式虚拟环境;五是网络和算力技术——空间定位算法、虚拟场景拟合、实时网络传输;六是物联网技术——通过各种传感器,实现万物互联,元宇宙需要万物互联技术,同时实现化身反哺物理世界,形成一种平行且相交的世界。元宇宙现象需要多种技术的支撑,单一的技术层面无法完美阐释元宇宙的本质。

从理论渊源和技术架构可以看出,元宇宙的概念在20世纪就已经被提出,并伴随着技术架构的发展而演变。总之,元宇宙是一个根植于物理世界、通过多种技术汇聚与整合出的平行于物理世界的镜像虚拟世界,现实世界是元宇宙的客观物质基础和物质载体的原型,它基于拓展现实技术和数字孪生技术实现时空拓展;基于AI和物联网技术实现虚拟人、自然人和机器人的人机融生性;基于区块链、Web3.0、数字藏品/NFT等实现经济增值性。元宇宙在社交系统、生产系统、经济系统上虚实共生,每个用户可进行世界编辑、内容生产和数字资产所有<sup>[12]</sup>。元宇宙就是ER拓展现实,把整个人类活动的方式用另一种方式表现出来。

## 二、元宇宙产生的必然性

技术不断发展促使社会变革以及经济形态转变,这样一种由内而外的转变呈现给人类的是第三世界——人工世界。波普尔<sup>[13]</sup>曾提出“三个世界”理论:世界1是物理世界,包括物质和能量,物理世界是“我信故我思”;世界2是主观知识世界,即意识状态和主观经验的世界,心理世界是“我思故我在”;世界3是客观知识世界,包括由各种载体记录并储存起来的文化、文明、科学技术等理论体系的人类精神产物,人工世界是“我在故我信”。波普尔阐明了物理世界作为世界1的优限性,世界2是随附其后的展开,而后形成世界3;三种世界是彼此相互作用的<sup>[14]</sup>。

### (一)元宇宙的超越性

元宇宙内在结构的超越性展开为五种超越:一是超越界面:从二维到三维(空间感)。人类早期的连接方式是二维形态的,人们通过书信、驿站等方式进行人与人之间的信息传递与连接。元宇宙本质上比人与人二维通信方式的意义更宏大,它打通了狭义上人和物的关系;二是超越静态:从读屏到互动(情境感)。人类对于外部世界客体的感知在视网膜上留下的图像都是二维的,元宇宙从技术上创造一种“幻觉”,让人们感觉到视网膜上的是一种三维图景,在这种三维虚拟场景中,产生一种情境感;三是超越单一:从静思到沉浸(场景感)。当下的初级形态的“元宇宙”,是通过VR设备达到一种体验感和情景感,但是这种感受属于被绑定式的,还不足以达到无缝穿越。而“元宇宙”是凸显“身体感知”特性的“沉浸式体验式文化”,是一个沉浸式符号化和虚拟化的空间,体验型虚拟影像诉诸体验者的感官,符号化

的虚拟场景仿佛是可触及的“超现实”场景<sup>[15]</sup>。四是超越身体:从具身到“化身”(穿越感)。我们现在可以做到3D虚拟投影,这种虚拟影像可以营造一种对现实肉身不在的人的体现,但还仅仅是一种被局限于某一场合的技术;五是超越现实:从理想到满足(满足感)。这种超越现实给人们提供了一个低成本试错的方式,它区别于物理世界的不可逆性,元宇宙世界允许后悔,同时又能保证现实物理世界的完整性。从元宇宙的五种超越可以看出,虚实不再是作为哲学范畴的对立存在,而是以引导知识体系为内容的虚实联动。

在以五种超越性生成的虚拟空间中,“化身”身处人类云的虚拟世界,能够超越时间、空间进行交互,主体通过数据传输把其全部信息附着到“化身”之上,“化身”可以是机器人或信息人,从而实现“化身”对肉身不在的人的体现,作为信息的“化身”继续保存着这些信息。区别于传统机器人的代理思维和代理意识,传统机器人在结构上和人不一样,而“化身”在效果和功能上和人一样。因此在元宇宙空间,意识的数据还原到场,通过算力产生的学习与模仿而形成的沉浸式仿真实验,实现对肉身不在者的再现,正如Virilio<sup>[16]</sup>所说:我们正在观看的并非“历史的终结”,而是“地理的终结”。

### (二)元宇宙是人性所致

正如杜威所说:“人生活在危险的世界之中,便不得不寻求安全。”<sup>[17]</sup>“缺乏”使得技术涌现,人类感受到了自身的不足,试图通过技术来弥补生理上的不足。从农业社会到工业革命,技术发展脉络呈现出五种社会形态:交通网(Grid1.0)、能源网(Grid2.0)、信息网(Grid3.0)、物联网(Grid4.0)、智能网(Grid5.0)。由此可见,人类增强的内在活力是科技进步的首要驱动力,人作为有缺陷的动物,通过技术抚慰心灵,增强自身,技术它不仅延伸人类的器官,还能够强化甚至代替人类的器官<sup>[18]</sup>。元宇宙从技术上打破了“人与技术主客体分明”和“一般意义的工具”的关系,从一般的意向性关系“人——技术——世界”,转变为具身关系“(我——技术)——世界”。“机器作为工具,不是我们各部分的简单的物质延伸。”<sup>[19]</sup>

通过历史可以看出,人类的发展历经了三个“轴心时代”。雅斯贝斯认为:“从古代文明中,或者从他们的活动范围中,在公元前800年至200年的轴心期,东西方两极分化的人类精神基础在西方、印度和中国这三个彼此独立的地区产生了。”<sup>[20]</sup>第一轴心

时代是人性中对“真理”的渴求开始觉醒,人文学科、哲学、文学开始突破,这一阶段发生在物理世界;第二轴心时代是文艺复兴时期,理性开始大觉醒,科学开始突破,这一阶段发生在心理世界;在第三轴心时代我们将面临第三世界,也就是人工世界,智性会进行一场大觉醒,技术方面也会突破,从而产生虚拟世界。科技的进步是人类独立思考的结果,科技发展的历史告诉挑战者们,科技根植于社会发展之中,渗透于社会生活之中。人作为难逃一死且变化无常的生物,通过人造物,世界的营造为人自身提供了坚实、稳固而安全的栖息地。正如阿伦特所说:“它又是人为的,而不是一个威严冷漠、不受触动的大自然。”<sup>[21]</sup>人性内部的驱动力是元宇宙涌现的自然所示,是全球化的必然结果。因此,利用科学技术推动人类文明进程的发展,是人类文明的必然选择<sup>[22]</sup>。

综上所述,元宇宙是一个根植于物理世界、通过多种技术汇聚与整合出的平行于物理世界的镜像虚拟世界,现实世界是元宇宙的原型,具有客观物质基础和物质载体,是能够破除人类生存束缚、实现人类价值最大化的场域,是能够沟通虚拟世界与现实世界的连接,实现以超越现实为抓手和以高纬度为内容的新型虚拟世界。元宇宙具有“在场性”、开放性(身份互确、信息共享、人人平等)、具象性(感官沉浸、实时参与)、自由性(数字身份、多元创造)和超越性(摆脱身体局限和物理束缚)。由此看来,元宇宙在本质上是一种数字化虚拟空间,其所展现的虚拟空间是人和物理世界之间的交互场域,它沟通连接了人的感觉与现实世界。

### 三、元宇宙的伦理风险

人类立足自然世界千万年,除了用认知官能和自然世界打交道,作为一种群体形成了社会,通过让渡一部分权利形成社会契约,而伦理共同体内部的形式与内容是法律法规和道德规范的展开与运用。当前元宇宙面临一个严肃的挑战,随着元宇宙技术的发展,人们需要思考其内部性质以及由内部性质和外部性质交互而涌现的伦理风险。元宇宙作为一种技术形态即展开为人工世界,这必然会带来社会的革命并由之生发出全新的环境,人类必然要面对一些全新的伦理问题。诚然,元宇宙的出现有其必然性,物理世界资源本质的有限性和虚拟世界数字化资源的无限性之间形成了某种张力,物理世界不允许虚拟世界占用太多资源,而虚拟世界试图找到能为物理世界创造价值的方式以表明其存在的正当

性。因而元宇宙要想回答正当性,必须进行伦理层面的考量。

#### (一)元宇宙的科林格里奇困境

英国技术哲学家 Collingridge(科林格里奇)曾提出一个有关技术控制的困境:“一项技术产生的社会后果不能在技术生命的早期——研发阶段被准确预见。然而当技术产生社会后果时,它往往已经成为整个经济和社会结构中难以抽离的一部分,以至于难以对它进行控制。”<sup>[23]</sup>

首先,通过回溯其内部结构发现,元宇宙在其设想中表达为对有限世界的超越。物质世界作为有限世界,迫使人们始终面临着有限性,例如资源的有限性、生命的有限性、能力的有限性等。元宇宙作为数字化的存在,本质上是数学以及数学所附带的无穷性,因此一个由数字化呈现的世界通达为一个无穷的世界,其实践的路径就是想象力的发挥。但是,元宇宙背后的技术支撑仍然是一种物质形态,这意味着元宇宙内部结构试图彰显的无限性无法摆脱物质自身的有限性。同时人们始终处于时间性的发问和框定之中,时间性在诸问题之中的优先性以及时间性使得我们在一定时间内拥有资源的有限性,使得人们仍然要面对由有限性而带来的问题。

其次,元宇宙的“化身”具有无限性,“化身”形成于表层经验结构中主体和客体的动态交互,而意识作为以自我为对象的作为绝对统一的自我意识(主体)具有唯一性,以及在认知神经科学中所表达的神经结构也是独一无二的。无论是主体自然感知框架与人工感知框架的切换,还是经验本身可能呈现为多种不同的形式,人的心灵总是在最深层次保持其自身的统一性。由此而言,“化身”仍然需要面临伦理问题,因为无限性具有无限种解决问题的可能性,无限的可能性意味着对负面后果的全然消解。但是,元宇宙是无法离开物理载体而独立存在的(这个结论具有自明性)。元宇宙是符号系统的一种特殊配置,在本体论意义上离不开物理载体。因此,元宇宙的属性在物理载体中可以找到它存在的充分条件,但是反之不成立。所以,虚拟无限性对问题的全然消解无法冲破物理世界的种种羁绊。

总之,元宇宙仍然会面临科林格里奇困境,其主要原因是早期风险难以预测。虽然研发人员会基于当前归纳的伦理风险来预测可能潜在的风险,而后进行规避,但是这种规避只是基于过往对风险的归纳,难以预测全新的技术会产生何种程度的社会危害。尤其是,元宇宙是对物理世界的超越,元宇宙不

是寻常的当代事件,而是一个存在论事件,意味着元宇宙蕴含着某种新问题的起点,构成了人类生活和思想的一个新本原<sup>[24]</sup>。这也就意味着我们很难全面分析其应用于社会后会造成何种风险。

## (二)元宇宙的去中心化风险

元宇宙是 Web3.0 即运行在区块链技术上的虚拟世界,通过区块链技术实现去中心化的机制,采取分布式的核算和存储方式,即不存在中心化的硬件或管理机构,进而形成了去中心化的特征。而区块链技术以互联网为硬件依托,以分布式数据库技术为软件支撑,同时实现了交易的去中心化和信息传播的去中心化,极大地提高了达成共识和价值传递的效率<sup>[25]</sup>。

首先,基于去中心化的视角,元宇宙把现实世界的中心化社区变为分散成无数个去中心化的社区,这种去中心化的底层逻辑已然颠覆 Web2.0 中的价值机制,进而传统社会中的效率机制也被颠覆。因为在现有社会结构中,事件的推进需要社区所有人达成共识,而共识的机制是不追求效率的,集权才是追求效率的。元宇宙的价值机制是通过消费信用来维护的,当用户去消费信用的时候,区块链会发挥效用,同时为虚拟世界购买商品的归属权提供支撑。因为现实世界的产品是有归属权证明的,在现实世界购买任何商品都是需要证明,以证明产品归属权。但是在虚拟环境里如何证明?答案是一套秩序和标准,区块链可以提供一种电子证明。虽然从表象结构来看,元宇宙是去中心化的,但是从另一个角度来说却是中心化的。随着基础用户的增多,数据逐步庞大,用户的数据对背后的技术承载者(公司)来说是透明化的,通过掌握用户的数据由此产生了精准的推送服务,因此在这个层面来说是中心化的。

其次,由中心化结构中衍生出的另一个间接风险是歧视风险。唐·伊德首次将“技术”纳入到具身性的概念范畴之内;真正的“具身关系”是技术嵌入人身体经验后的人、技术及世界的关系<sup>[26]</sup>。元宇宙作为一种技术物和媒介物,其背后也隐含着一套数字逻辑,通过与个体的“具身实践”再生产着某种虚拟社会结构和社会关系。去中心化的结构试图说明其是一个平等的交互空间,但是从技术架构的实质内涵来看,元宇宙本身可能隐蔽着不平等的权利关系。首先,设计者是基于健全人的屏幕互动和设备佩戴能力来建立元宇宙的,潜在的问题是当人们通过思维、知觉、视觉具身参与到虚拟构建的空间时,在先地需要肉身具有健全的感知框架和思维功能。

虚拟空间对感知框架和思维框架的占据是不可控的,而老龄化群体首先面临着肌体老龄化和感知功能老龄化,这就意味着老龄化群体在进入元宇宙时,实际上是不具备健全的感知能力和思维能力的。其次,进入元宇宙时是需要佩戴外部设备的,但是对于一些不具备视觉能力和听觉能力的人,技术发展强制地抛弃了这些少数群体,他们不断被边缘化,实际上内部隐含着一种不平等关系。由此而言,元宇宙自身在先地携带着一种“顺从”和“接受”的技术逻辑。

最后,元宇宙是去中心化的,那么该由谁来承担元宇宙的所有权和管理权?元宇宙世界产生的问题由谁来承担?这些问题尚不清晰。如果元宇宙的运行不依附某一权威,而是依靠使用者来承担,那么将会出现无人承担责任,因为使用不代表“拥有”。技术的决定性社会作用逐步演变为统治性的社会作用,技术的解放力量转而成了解放的桎梏。去中心化最大的问题就是责任归属权的流失。“技术既是人类自身的力量同时也是人类自我毁灭的力量。”<sup>[27]</sup>在技术发展过程中,责任伦理是重要一点,技术的发展指向是以人的生存与发展为最高目的的,但忽略了伦理道德的一面。因此,元宇宙的发展必须加入责任伦理的考量,提醒人们对后果和行为进行考量负责,责任与伦理环境连接起来,“谁对伦理环境的负面作用最大,就必须受到相应的谴责和惩罚”<sup>[28]</sup>。总的来说,去中心化只是元宇宙的内部空间形式,而其外部掌控形式仍然是中心化的。假设最终的元宇宙是由一家公司控制,那么它将主宰着全体人的精神,是一个超政府的存在,是一种专制的权利。

## (三)元宇宙的精神迷失风险

翟振明指出:“一旦我们进入虚拟实在的世界,虚拟实在技术将重新配置整个经验世界的框架,我们把技术当成一个独立物体——或‘工具’——的感觉就消失了。这样一个浸蕴状态,使得我们第一次能够在本体层次上直接重构我们自己的存在。”<sup>[29]</sup>由此观之,元宇宙区别于传统互联网的二维、局部和“虚假”。虚拟现实的创始人之一拉尼尔说:“VR 非常像物理世界,其中你在外表上感知的那种现实就是你通过感觉器官和物理世界所感知到的。通过实验表明:小白鼠在显示着障碍墙图像的虚拟环境中,会把虚拟障碍当作真实障碍物,像走真实迷宫一样来回寻找出路。”<sup>[30]</sup>这些作为显象的伦理风险诚然也会递进到与之结构具有相似性的元宇宙技术之中。随着技术不断发展成熟,这种“由虚到实的浸蕴

感”逐步占据主体,而当真与假失去分辨力,随之而来的是“真实性”意义的消散,肉身被动单线程的抵达感觉之中,从而丧失考虑感觉背后的“实在性”,拒斥实在性如同拒斥世界的真理。

首先,自然主义的意识阐释立场指明,大脑神经元组织的信息处理过程为人的一切意识活动。元宇宙正是通过脑机接口对人的生物层面上的干预来影响人的意识状况,达到高度的沉浸感。这意味着意识的数字化,而意识的数字化解构了意识的能动性。例如,用某种感受的物理事实就能够对某种感受进行重新定义,在新的定义中,数字化的“意识还原物”到场,把意识指向某物的主观能动性消解了,意识变为被动性的接受被还原为数字信号的内容。从某种程度上讲,这时的内容是他者生成的内容,它失去了主体在所处的世界当中对实在的客体的交互性。基于此,人的意识不再是私密且不可显现的,而变为可传输的电信号。当意义的实现是通过不同的电信号输入时,意味着建立在对“实在”刻画过程中的意义彻底消散了。因为现实世界的价值和意义是多元化的,是通过实践等方式实现的,其价值背后具有的实在性保证了其价值性。

其次,海德格尔把生命的特征论述为“被抛”:自我的惊愕感意味着对当下具体的感知,这个感知意指没有人请我们来;也没有人“准”我们来,我们每个人都是被扔到这个世界上来的<sup>[31]</sup>。亦即,我们突然发现无人邀请或批准我们到来,没有挑选家庭,在一定的时间、社会和历史,生命就是一种“被给予”。由此伴随着事件的偶然性和生命的有限性,构筑起了价值。我们被抛到这个世界上,自身就带有意义的缺陷,所以我们需要不断调节自身,追求意义。而元宇宙是一种选择和创造无限性的世界,选择和创造无限,意味着意义也是无限,但无限的意义本身就消解了意义的价值。一方面,元宇宙是一种主体赋值的意义世界,是一种通过意识来意构的世界,其形式类似于梦境中的乌托邦,于是价值不再是劳动的凝聚和经验世界实践的产出,而是欲望通过想象智能被实现。这个时候只是为了欲望本身而去做出选择,失去了现实世界之中选择和努力的意义,那么就会带来欲望的沉沦和精神的迷失,人们会逃避现实世界,逃避经验的生活,不再创造价值。另一方面,元宇宙是一种“虚拟空间”,能实现“化身”在虚拟空间中的“转场”,然而这种具身“转场”在一定程度上也加速了一种“多线程”的生存模式和交往模式。例如,元宇宙可能会提高个体的工作强度,因为虚拟空

间的“转场”性实质是一种多线程的便利性,会造成虚拟世界的社交倦怠和结构压力。

最后,元宇宙作为一个全新的虚拟世界,元宇宙在其中作为一种新的中介方式出现,连接主体和事物之间的关系,这种中介的连接,就意味着人被管理和操控。因为当人们选择进入元宇宙,享受平台的服务时,就意味着必须要让渡权利,元宇宙作为一种虚拟世界,把控着全平台用户。同时,生物体的存在与对象的存在互为因果关系,感知的活动是感知者利用感知系统去感知环境中对象的一个整体过程。感知得到的意识不是简单地发生在大脑的思考过程中,也不是发生在有机结构的生态水平上。感知者所感受的客体会受到感知者的影响,因此对客体的感知必须在感知者的行动中加以理解。结论是:感知觉的问题是人与环境的关系问题(环境就是生物体周边的事物,但是它不等同于物理世界)。用吉布森的生态心理学来解释便是,在元宇宙中,参与者将足够的注意力资源分配到虚拟环境中,只要有足够的感觉保真度,那么参与者可以“放弃怀疑”,并将虚拟环境视为一个实际的地方,可以产出一种存在感。这种难辨真假的沉浸感不断渗透到人们的生命之中,经验生活和身心全面迁移,人们必然会进入一种依赖,一种不自觉的被迫的依赖,在沉浸的虚拟中忘记自我。电脑、互联网、手机作为被广泛采用的现代工具,与“短期意识”完全和睦共处。持这种意识的人善于幻想,但不愿意思考。用俄罗斯教育家乌申斯基的话来说:“幻想容易,思考难。”<sup>[32]</sup>

#### (四)元宇宙的“化身”风险

在数字乌托邦主义者看来,数字生命是宇宙进化的最终选择<sup>[33]</sup>。那么,数字孪生中,人一旦被数字化,他的环境应该是什么?他被什么环境物竞天择?他是否还会像传统人类遵循自然法则一样去遵循一个元宇宙里原生的“自然法则”?技术的发展引发社会的变革,改变人类接受的多样性,我们是否会接受这样一个数字人?是主动还是被动地接受?还是随着技术在社会的渗透,形成的环境进而改变人们的认知结构?

首先,传统世界是主客二元的世界,人们依靠第一视角对世界认识、感受和经历。第三视角——“他者”,是一个可能威胁到我生存的朋友或敌人的关系,现在缺少第二视角的“化身”,这个“化身”很可能会变成我和他,尽管“化身”并不真实存在。但是在元宇宙空间里,每个人都被提供了你我他三个视角,是集于一身的视角。元宇宙为我们提供了“化身”、

替身,这个“化身”、替身不是一个被奴役的账号,而是这个账号具有自我独立的人格。这样一个万物在场的元宇宙空间,“化身”面对“化身”的同时还要面对背后的肉身主体,在一个以去中心化形式的系统的元宇宙世界,达摩克利斯之剑失去了对“化身”的规约力,被理性中的自我约束力所替代,这种依靠自我约束的形态是对人性的绝对考验。

其次,卢梭在社会契约论中指出,人类并不能产生新的力量,而只能是结合并运用已有的力量;所以人类便没有别的办法可以自存,除非是集合起来形成一种力量的总和才能够克服这种阻力,由一个唯一的动力把他们发动起来,并使它们共同协作<sup>[34]</sup>。人们之所以会生活在一起,是因为人们相互需要,而人们之所以相互需要是因为每个人在能力上是有限的、有差异的,无法完全靠自己某方面的优先能力来满足和维持自己的生活,因此人们需要分工合作来战胜自然。但是,元宇宙打破了人的差异性,使人趋于同构化,虽然其提供了多重“化身”与可选择主体的多样性,但是这种多样性带来的并不是差异性,而是同一性,是一种数字化的同一性。在物质世界中,身体的结构是由自然给予的,人们拥有不同的身高、肤色、性格、样貌,虽然元宇宙可以如此展现,但追问其本质,不过是一种同质化的数字代码。

最后,相较于自然世界的交互方式,元宇宙改变了人们的交互模式。元宇宙主要是基于拓展现实(XR)等技术,创造了一种“presence”的感觉。但是这种社会性主体在场呈现的是脸部不在场的状态。人类过去遵循自然法则、社会契约和伦理道德规范来生活,脸部是交互的窗口,因此具有伦理意味,当脸部不在场时,人们往往会忽视伦理。虽然虚拟空间为主观能动性提供了发挥空间更广的实践领域,以及在此领域中更为深入的自由性,但是这种自由并不是排除了外部权威和外部干涉的抽象自由,元宇宙依然要遵守伦理共同体的法律和伦理道德。

### (五)元宇宙的虚拟侵害风险

物理世界的规律从第一性原理推动之后就不变了。在物理世界,人们不会有本质上的差别,都不能违背物理定律和生物原理。但是元宇宙作为一个人工世界,它底层的代码由谁来写?即使采用去中心化的方式,但是仍然需要代码去建构,如果底层代码的实际操控者出现风险,那么整个元宇宙就会出现危机。设计者作为元宇宙底层代码的编写者,发现自己能够创造这样一个人工世界,站在一个上帝视角,他者可以定义这个世界的规则,定义虚拟数字人

的命运等。

首先,互联网在部分悲观主义者看来具有沦为“电子圆形监狱”(electronic panopticon)的风险,意指它可能导致“无处不在的监视”<sup>[35]</sup>。那么我们如果确保作为元宇宙里的主体性的我是否是被底层代码修改?如何确保这个我还是真实的我?如何确保产生的行为、感觉、思想的真实性,以及我怎么确保我没有处在这个被修改的困境之中?元宇宙试图通过沉浸感形成“在场”(presence),形成虚拟环境中主体“在场”的感觉(关于存在的意识状态和心理感觉)。如果主体全身心沉浸的程度越高,那么主体感觉存在的程度也越高。这个感觉性存在一定是建立在人机交互的过程中实现,利用交互产生心理感觉,进而产生感觉性存在。这似乎说明虚拟世界中造成的伤害、冒犯和侮辱具有和现实世界中一样的伤害性,在虚拟世界中,侵害是真实的。例如,在虚拟世界对其他“化身”进行侵害(强奸或侮辱性侵害)时,这种感觉性的存在使得主体会遭受如现实中一样的伤害,尽管其伤害是一种电信号和数字化的传输,但是精神的伤害是可以改变其心理状态,从而导致抑郁或自杀等行为发生。

其次,虚拟侵害下的隐私权怎么保障?虚拟空间的隐私问题历经智能革命的转变之后不再具有绝对性,因为机器人从简单的监视进展为可以对人采取拘押行动<sup>[36]</sup>。所以,智能技术隐性的问题是带来一种被称之为“电子圆形监狱”的风险<sup>[22]</sup>。因此,无论如何人们行动,都会在元宇宙里留下痕迹,而且,利用区块链的方式作为记录,这些痕迹将无法消除。但这并不意味着进入元宇宙就等于放弃隐私,相反,由于元宇宙本质是把信誉和信用度变成商品,所以隐私并不会逐步瓦解,元宇宙会变成一座新的“牢狱”。同时,元宇宙是一种可形式化的假设,即是诸事物的运作过程包括生命的本质特征都隐含着某种形式的机制的预设。这个形式的机制是什么呢?就是一种数字化的机制,通过基于生物或物理的算法,意指通过数学语言来表征事物及其过程的形式,生命、意识乃至人的心理状态本身也能被数字化,进而可以通过计算机的算法来表达。这种数字化的表达使得数据变得透明化,数据隐私的侵害始终存在。

最后,如果两种物质载体的属性存在着物理层面上的不可分辨性,那么在载体之上附随的元宇宙现象也是不可分辨的。只要物质载体的属性在结构或状态上不可分辨,即使不同的物质载体,元宇宙现象也是不可分辨的。这就是说同一个元宇宙现象可

以依附于不同的载体。这使得元宇宙具有可传递、可存储、可转换的特点,也意味着元宇宙始终存在着不稳定性。虽然自然世界也具备不稳定性,但是这种不稳定性是自然世界强加给我们的,同时我们无法窥探自然世界背后的本质结构。元宇宙的本质结构则是一个被人类建构和描述的系统,那么始终隐含着一个危险,就是技术的脆弱性,这种脆弱性是因为人类行为的不端。如果系统背后的某个主体控制了这一系统(关闭机器或毁坏机器),或是肆意编造和修改数据代码,都会造成巨大的灾难。

#### 四、元宇宙伦理风险的应对路径

科技的发展必然会面临诸如此类的伦理风险,但风险并不等于危险,要针对风险提出可行的伦理风险解蔽路径。皮特·克罗斯(Peter Kroes)和安东尼·梅耶斯(Anthonie Meijers)在讨论技术人工物时指出:“技术人工物具有双重属性,技术人工物既是具有物质性的实体,又是具有特定功能的客体,前者必须符合物理的、自然规律;后者与人的意向性相关,其功能可以被善使用,也可以被恶使用,人们会对其做出规范性判断。”<sup>[37]</sup>那么,虚拟世界的数字化建构,意味着伦理、道德和法也要数字化建构,因为设计者在设计元宇宙之初,是一种意向性置入。伊德<sup>[38]</sup>在分析技术意向性的“他者关系”时避免了海德格尔式的分析,引入了列维纳斯的他者概念。他者意指以自我为中心,我之外的诸事物于我而言都是作为他者而存在的,技术显然是外在与我的存在物,人对于世界的感知不是借助外在的技术,而是对于技术本身的直接感知。如果我们将人的伦理道德拓展到技术物,将技术纳入人的伦理分析范围,基于意向性这样一种主体与外在世界交遇的能力,设计者在元宇宙技术设计之初带着在先的人文理念与之交遇之后,元宇宙内部此时蕴涵着应然的伦理理念。因此,对设计者进行人文价值的灌输,也就意味着元宇宙的框架被置入了精神人文理念,它的人工性需要从人文价值高度上进行更具深度、更为开放的不间断诠释和理解。

此外,元宇宙技术作为一种全新的技术形态,基于传统框架“人—技术—世界”的审视已然不足以面对虚拟世界所构造的“人(技术)—世界”的框架。对元宇宙技术的人工性或工具性的诠释和理解,需要放置于精神人文价值的层面进行更为开放和更具深度的反思。我们要以精神人文价值全面审视元宇宙技术的“人工性”,发展与素质并重,全面提高素质素

养,普及数字生活理念,针对相关责任主体进行“意向性”置入。将人工性限定在人类的精神人文限度内,前提是未雨绸缪,因此规范、教育和监管都要走在技术前面,尽可能地早规避、早监管,以规避元宇宙隐蔽的伦理风险。唯有如此,元宇宙技术对人类的意义才不会仅仅像马克斯·韦伯所强调的作为“中立性质的手段”和“一般意义的工具”,而转变为内具深厚人文价值尺度的技术展现形式。从这个视角出发,元宇宙技术的“人工性”之称谓就存在某种程度的误导,需要以“负责任创新性”视角去发掘其中负载的精神人文价值选择和责任。这是规避元宇宙技术隐蔽的伦理风险的前提。例如,在“去中心化”风险、隐私风险、精神迷失风险等议题上,如果不能从“负责任创新”高度上重新梳理界定元宇宙技术的人工性内涵,就不可能对其产生的伦理风险作出预估,并在设计过程中有的放矢。在不久的将来,当社会逐步被元宇宙技术所渗透时,对人类面临的全新问题的解决寄希望于人文社会科学研究成果,而对全新问题的界定需要由人类共同体中的人类生活实践来引领。由此,有必要回顾一下著名的人类学家克劳德·列维—斯特劳斯的断言:“21世纪将会是人文与科学世纪的世纪,或者它将根本不存在。”<sup>[32]</sup>

#### 五、结 语

元宇宙在今天产生的社会波动,一方面是一些资本和科技公司一起炒作的一个概念游戏,另一方面是技术内在逻辑的演变与形式的展现。诚然,无论是 Mate,还是 Cyber,对元宇宙的概念还无法达成一致解读。但从整体而言,元宇宙是与物理世界绝缘,在物理世界之外构建的一个比特世界,与自然时空相平行,由多种技术汇聚而成。从价值角度而言,元宇宙的五种“超越”形成了一种全新形态的第三世界,通过虚拟映射,实现反哺物理世界,有利于科技的发展和知识的传递,用“虚”服务于“实”。从伦理角度而言,元宇宙带来了全新的伦理风险挑战,理清可能存在的伦理风险是首要任务。伦理本身并无确定答案,伦理天然的是主观决定。元宇宙技术可以葬送人类,也可以把人类送往星辰大海,产生何种结果取决于人类。也就是说,从伦理风险应对的视角出发人依旧是智能系统的尺度,智能系统不能离开人,同时需要理清元宇宙的人工性与精神性,以意向性为抓手实施人文精神的置入,以“负责任创新”为出发点尝试提出伦理风险的应对路径。总之,元宇宙的出现是技术发展的必然,作为一种全新的“宇

宙”,它给人们提供了一个思想实验室,一种社会性的存在。思考元宇宙潜在的伦理风险及其应对路径,是推进元宇宙发展的必要路径。

### 参考文献:

- [1] 赵国栋,易欢欢,徐远重. 元宇宙[M]. 北京:中译出版社,2021:5.
- [2] 曹刚. 元宇宙、元伦理与元道德[J]. 探索与争鸣,2022(4):83-85.
- [3] 喻国明,耿晓梦. 元宇宙:媒介化社会的未来生态图景[J]. 新疆师范大学学报(哲学社会科学版),2022,43(3):110-118.
- [4] 谢程远,马彩慧.“技术正义—权力正义”视域下元宇宙的价值意涵和实践向度[J]. 理论导刊,2022(5):69-75.
- [5] 王建红.“元宇宙”现象的历史唯物主义探析[J]. 江淮论坛,2022(2):152-157.
- [6] 马尔库塞. 单向度的人:发达工业社会意识形态研究[M]. 刘继,译. 上海:上海译文出版社,2006:142.
- [7] 曹新,向文华. 21世纪的高新技术发展与经济社会变革[J]. 南方经济,2002(3):23-24.
- [8] 莱布尼茨. 神义论[M]. 朱雁冰,译. 北京:生活·读书·新知三联书店,2007:413-416.
- [9] Stephenson N. Snow Crash[M]. New York: Penguin Random House,1992:14-15.
- [10] 张晖. 从钱学森对VR的译名看科技译名的“中国味”[J]. 中国科技翻译,2020,33(1):9-12.
- [11] Gelernter D. Mirror Worlds: Or the Day software Puts the Universe in a Shoebox... How It Will Happen and What It Will Mean[M]. New York: Oxford University press, 1993:21-28.
- [12] 沈阳. 元宇宙时代的语言变化[J]. 语言战略研究,2022,7(2):5-6.
- [13] 波普尔. 客观知识:一个进化论的研究[M]. 舒炜光,卓如飞,周柏乔,等,译. 上海:上海译文出版社,1987:164-165.
- [14] 中国科学院沈阳自动化研究所简明信息技术百科辞典编辑组. 简明信息技术百科辞典[M]. 北京:知识出版社,1992:15.
- [15] 付茜茜. “元宇宙”:赛博空间的技术趋势与文化症候[J]. 学习与实践,2022(4):132-140.
- [16] Virilio P. The Information Bomb[M]. Chris T, Trans. London and New York: Verso, 2000:9.
- [17] 杜威. 确定性的寻求[M]. 傅统先,译. 上海:上海人民出版社,2004:1.
- [18] 盖伦. 技术时代的人类心灵:工业社会的社会心理问题[M]. 何兆武,何冰,译. 上海:上海科技教育出版社,2008:3.
- [19] Mournier E. Be Not Afraid[M]. Cynthia R, Trans. London: Rockcliffe, 1951:195.
- [20] 雅斯贝斯. 历史的起源与目标[M]. 魏楚雄,俞新天,译. 北京:华夏出版社,1989:31.
- [21] 阿伦特. 人的境况[M]. 王寅丽,译. 上海:上海人民出版社,2009:106.
- [22] 刘大椿,成素梅,段伟文,等. 智能革命与人类深度智能化前景(笔谈)[J]. 山东科技大学学报(社会科学版),2019,21(1):1-11.
- [23] Collingridge D. The Social Control of Technology[M]. Milton Keynes: Open University Press, 1980:16-17.
- [24] 赵汀阳. 假如元宇宙成为一个存在论事件[J]. 江海学刊,2022(1):27-37.
- [25] 孙国茂. 区块链技术的本质特征及其金融领域应用研究[J]. 理论学刊,2017(2):58-67.
- [26] 吴国盛. 技术哲学经典读本[M]. 上海:上海交通大学出版社,2008:373.
- [27] 斯蒂格勒. 技术与时间[M]. 裴程,译. 南京:译林出版社,2000:100.
- [28] 晏辉. 伦理环境的原始发生及其归责问题[J]. 伦理学研究,2020(3):19-26.
- [29] 翟振明. 有无之间:虚拟实在的哲学探索[M]. 北京:北京大学出版社,2007:1.
- [30] Morie J. Ontological implications of being in immersive virtual environments [C]//Conference on Engineering Reality of Virtual Reality. San Jose: Proceedings of SPIE, 2008,6804:5.
- [31] 麦基. 思想家:与十五位杰出哲学家的对话[M]. 周德明,翁寒松,译. 北京:生活·读书·新知三联书店,2004:88-89.
- [32] 瓦切斯拉夫·谢苗诺维奇·斯乔平,章含舟,陈欢,等. 科学哲学如何塑造更好的“后人类”社会:访斯乔平教授[J]. 哲学分析,2018,9(6):160-175.
- [33] 泰格马克. 生命3.0:人工智能时代人类的进化与重生[M]. 汪捷舒,译. 杭州:浙江教育出版社,2018:40.
- [34] 卢梭. 社会契约论[M]. 何兆武,译. 北京:商务印书馆,2003:18.
- [35] Katz J, Rice R. Social Consequences of Internet Use: Access, Involvement, and Interaction[M]. Cambridge, MA: MIT Press, 2002:3-8.
- [36] 李保艳,刘永谋. 元宇宙的本质、面临的风险与应对之策[J]. 科学·经济·社会,2022,40(1):15-26.
- [37] 张灿. STS视域下的技术安全哲学研究[M]. 北京:中国社会科学出版社,2018:94.
- [38] 伊德. 技术与生活世界:从伊甸园到尘世[M]. 韩连庆,译. 北京:北京大学出版社,2012:34-37.

(责任编辑:陈丽琼)