

【专题：融智能+活数据+万联网+创新设计】

负责任的人工智能与设计创新

王 韞^{1,2}, 徐迎庆²

(1.北京航空航天大学, 北京 100191; 2.清华大学, 北京 100084)

摘要: **目的** 快速发展的人工智能技术的应用将改变人们在未来的思考和生活方式, 重塑社会礼仪和道德伦理, 产生巨大而不可估量的影响。数据噪音所带来的潜在社会道德和设计伦理问题可能会对设计创新中的公平性和包容性造成干扰和威胁, 潜移默化地加剧社会矛盾, 加深社会的不稳定性。因此, 在全球化、普适计算和物联网时代, 设计需要理性地重新思考其在人文艺术和科学技术之间的桥梁作用, 并从新的角度来观察与探究人工智能应用下的“不确定性”问题。**方法** 面向应用的人工智能的本质是工具性的, 然而设计师群体对这种新工具的理解和思辨还有很大的改善空间。设计具备将价值观融入造物的能力和义务, 应充分地发挥其特质和作用, 借鉴负责任创新中的预期、反思、协商和响应四个责任维度, 将责任风险评估、管理和干预全方位地嵌入到人工智能应用设计创新的决策和流程中, 以一个包容开放的动态适应机制来应对未来的挑战。**结论** 设计师是社会创新的重要驱动力之一, 需要积极地投身于倡导创新价值、技术及设计伦理的前沿。负责任的人工智能和设计创新应协同制定更具有前瞻性的目标和任务, 引领聚光灯照亮人类前进的方向。

关键词: 人工智能; 数据偏见; 负责任创新; 负责任设计; 设计伦理; 设计决策; 社会责任

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2021)06-0001-06

DOI: 10.19554/j.cnki.1001-3563.2021.06.001

Responsible AI and Design Innovation

WANG Yun^{1,2}, XU Ying-qing²

(1.Beihang University, Beijing 100191, China; 2.Tsinghua University, Beijing 100084, China)

ABSTRACT: Rapid development of Artificial Intelligence (AI) application will change how we think and live in the future and reshape social protocols and moral ethics, resulting in an immense but immeasurable impact. The potential issues of social morality and design ethics brought by data noise may cause interference and threat to the fairness and inclusiveness in design innovation, and subtly aggravate social contradictions and deepen social instability. In the era of globalization, ubiquitous computing and IoT, designers need to rationally rethink their role as a bridge between humanities and technology, and observe and explore the “uncertainty” in AI applications from a new perspective. The nature of application-oriented AI is implemental, however, designers’ understanding and speculation of such a new tool still have large room for improvement. Having the ability and obligation to integrate values into its creation, design should give full play to its characteristics and impact, learn from the four dimensions (anticipatory, reflective, deliberative and responsive) of Responsible Innovation, for embedding assessment, management and intervention of responsibility risks into design decision and process of AI applications, and for establishing an inclusive, open and dynamic adaptive mechanism in order to meet the future challenges. As an important driving force of social innovation, designers should actively engage in the forefront of promoting innovation value and design ethics. Responsible AI and design innovation should work together to set more forward-looking goals and tasks, and lead the spotlight to illuminate the future of human progress.

收稿日期: 2020-12-09

作者简介: 王 韞 (1985—), 女, 北京人, 清华大学博士生, 北京航空航天大学讲师, 主要研究方向为感知觉体验设计和跨学科设计思维教育。

通信作者: 徐迎庆 (1959—), 男, 北京人, 博士, 清华大学教授, 主要研究方向为用户体验设计、触觉认知交互、文化遗产数字化以及自然用户界面设计。

KEY WORDS: artificial intelligence; data bias; responsible innovation; responsible design; design ethic; design decision; social responsibility

信息技术革命以来,现代信息科技改变了世界的社会经济结构,新的技术开发和产品应用层出不穷,物理世界与数字世界的界限逐渐模糊,学科间的交叉创新日益增多。设计师群体也从未像今天一样,触手可及数量如此众多的优秀技术专家、高效技术应用手段以及海量规模的大数据。随着计算能力和算法的提升,人工智能(Artificial Intelligence, AI)成为21世纪的尖端技术之一,试图探索人类智慧的本质,全面地革新信息和知识的处理、分析、理解、加工和应用等能力。AI除了在数据分析和应用上开辟了诸多学科交叉研究方向以及设计机会,也作为一种更加智能的计算机辅助设计技术进一步地融入设计主体。与此同时,人工智能所带来的数据伦理问题可能会对设计中的公平性和包容性造成干扰和威胁。面对这一挑战,设计师群体需要针对其所应秉持的观点以及采取的策略等问题展开充分的讨论和探索。

1 人工智能应用中的伦理挑战

人工智能无疑是信息科学中最具有未来色彩的分支之一,与其相关的机器学习和神经网络等技术的开发和应用逐渐成为学科交叉创新中的热门方向。这些技术的发展大大地提升了设计师对于数据的挖掘和分析能力,同时对相关工具的掌握程度和使用方式也逐渐对设计创新的效率和质量有了更为直接的影响。尤其是在学科交叉较为深入、人工智能应用广泛的交互设计研究领域,面对社会学、认知心理学、脑科学和大数据计算等新老领域密集的知识网络,设计从业人员、专家和学者大都会产生压迫感,并对设计活动本身的定位感到困惑。

事实上,人工智能为设计所带来的挑战远不止于新知识和新工具的冲击。它推动了科学技术的发展以及人类未来思维模式的变化,也对社会秩序、道德伦理和人类生活方式等进行着重塑,从而造成了规模巨大且难以估测的影响。计算机信息处理和知识加工能力的增强,使得人们参与社会互动的范围和频率大大增加,信息和知识的冗余给人带来的选择困境甚至会反过来支配人的需求^[1]。更危险的是,基于没有使用责任维度进行调节的数据而进行的决策和应用,会潜移默化地加剧社会矛盾。

尽管来自不同领域的大批专家和学者已经敏感地看到了技术应用中的伦理问题并开展了积极的讨论,但是设计思潮对于相关主题的探索仍然相对滞后。在全球化、普适计算和物联网时代,设计确实需要理性地重新思考其在人文艺术和科学技术之间的桥梁作用,并从新的角度来观察与探究人工智能应用

中的“不确定性”问题。

1.1 数据偏见

Batya Friedman 在价值敏感设计(Value Sensitive Design/ASD)理论中^[2],梳理了计算机系统三种偏见:源于社会组织、实践和态度的先存偏见;源于技术设计中工具缺陷的技术偏见;以及在用户实际使用时产生的突发偏见。在图形用户界面发明之前,人们主要通过键盘与计算机进行交互。图形用户界面的推广虽然改善了许多人对于计算机的使用体验,但是在早期的图形用户界面中,视觉受损的用户难入其门,这是由于技术缺陷所造成的偏见的一个案例。技术偏见的可见度高,能够通过系统的迭代进行修正,上述图形用户界面就增加了针对视觉受损用户的功能。然而,先存偏见很可能更难被意识到,并且会产生更深远的影响。这是因为它来源于社会中的一些概念性和结构性问题,比如人的身份、人与技术之间的关系以及个人或群体之间的关系等。

人工智能会忠实地反映出原始数据的特征,因而把社会中的不平等与偏见明确地展示在世人眼前,使得那些以往人们视而不见或者习以为常的社会问题原形毕露。与此同时,人工智能往往是在大数据的基础上产生结论,由于其自身并没有分辨数据真伪的能力,因此也就进一步地放大了那些由来已久的偏见,从而加剧了社会的不平等。基于数字化信息的AI决策系统可能会让少数族裔、女性或者有认知病史的人更难找到工作^[3],而残障人士、老年人和低收入人群等落后于数字化时代的无数据群体则处在聚光灯外的阴影中。数据偏见和算法的“黑箱性”反映了数据权利之间的失衡现象,这将不可避免地导致算法歧视^[1]。

1.2 负责任的人工智能

数据偏见固然具有破坏潜力,但面向应用的人工智能具有工具性的本质,为其自我矫正提供了可能的手段。计算机科学领域的学者和从业者最先行动起来,针对人工智能的负面影响进行了评估和干预。2018年,牛津大学、剑桥大学和Open AI公司等多家机构共同发布了《人工智能的恶意使用:预测、预防和缓解》^[4]。2019年电气和电子工程师协会(IEEE)最新版的《符合伦理的设计》^[5]中,列举了8项自动和智能系统的设计、开发和应用原则:人权、福祉、数据安全、效率、透明、问责、慎用和资质。2019年英国艾伦图灵研究所出版《人工智能伦理和安全原则手册》^[6],阐述了透明原则,包括流程透明度(建立基于流程的治理框架)和成果透明度(解释成果、澄清内容和负责任地实施)。中国对人工智能的伦理

高度重视,2002—2017 年间,与人工智能伦理相关的国家级课题研究共计 26 项^[1]。2017 年发布的《新一代人工智能发展战略规划》^[7]中强调,必须认真管理好 AI 的技术和社会属性,以确保其可靠性。2019 年,中国科技部发布了《新一代人工智能治理原则——发展负责任的人工智能》^[8]。

在世界范围内,公平 AI、可解释 AI 和以人为中心的计算(HCC)等概念和开发框架陆续被提出,建立以人为中心的数据伦理,强调算法的透明性,消除数据孤岛,提倡有规范的数据共享,防止数据滥用^[1]。

2 设计创新的社会责任

2.1 负责任创新

越来越多的学者开始关注科技所给予的未来和人类作为未来的实际承受者之间的责任鸿沟,对该问题的认识和辩论可以对责任融入社会创新起到积极的推动作用。“负责任创新”,又称“负责任的研究和创新(Responsible Research and Innovation, RRI)”,自 2003 年起就出现于学术与政策文本中^[9]。《欧洲联盟条约》为负责任创新框架的建立提供了一些规范性的锚点^[10],这些锚点及其相互关系为定义研究和创新应追求的“正确”影响提供了一些初步的合法性基础,基于《欧洲联盟条约》的规范性锚点见图 1。

欧洲委员会成员 Rene Von Schomberg 对负责任创新的定义是“一个透明、互动的过程,社会行动者和创新者对创新过程和成果的道德可接受性、可持续性和社会期望性进行互相响应,使得科学技术的进步更好地为人类社会服务”^[10]。为了将当前的创新实践与其所许诺的未来联系起来,负责任创新促使创新者思考正在做的工作可能导致的不同结果,以及个人的责任和义务;主张打破传统的责任分工,强调各方的参与,从监管者、创新者和用户等不同视角来探讨科技将会给世界带来怎样的未来、应当关心怎样的未来、这些未来的价值观是什么以及关于未来的探讨本身是否具有足够的公平性和透明度。负责任创新既是名词性的基本价值观,又是动词性的程序方法,其本

质特征离不开学科交叉及跨领域合作^[11]。

与负责任创新有关的理论框架中,比较有代表性的是英国学者 Richard Owen 提出的可持续的集体承诺的 4 个维度^[10]。(1) 预期(Anticipatory):描述和分析可能产生的预期性的,以及潜在的非预期性的影响。这需要在探索不同途径时自问“如果……,是否会……”,以及“是否还有其他的可能”。预期是对创新所带来的不确定的未来的一种暂时的先验感知和表达,并保持选择的开放性。(2) 反思(Reflective):尽可能地反映已知的和未知的潜在目的、动机和影响,以及考虑相关的风险、假设和困境等因素。反思比预期更贴近于在当下资源和条件下的相对确定的创新成果的直接影响。(3) 审慎(Deliberative):通过将不同的直接和间接利益相关者纳入到集体讨论和审议中,引入更广泛多元的视角,来对问题进行更加全面的界定。(4) 响应(Responsive):建立一个迭代的、包容的、开放的、具有动态能力的适应性机制,利用一种集体的自反过程来确定创新的方向,进而影响随后的创新步伐和轨迹。

2.2 负责任的设计

负责任创新是“通过对当前科学和创新的响应性管理来关心未来的一种集体承诺”^[10],设计作为创新的主要驱动力之一,需要积极地参与到定义创新价值取向和设计伦理的最前线。事实上,负责任创新的行动纲领与设计学思潮中激进主义设计、批判性设计、思辨设计、负责任设计、社会责任设计和参与式设计等诸多观点不谋而合。

1971 年 Victor Papanek 在其专著《为真实的世界设计》^[12]中,尖锐地审视了设计的责任困境和使命,并着重强调了跨学科研究的必要性。他指出设计师对于其工作的支配力比其设想的要大得多,必须意识到其所承担的社会和道德责任。受到 Papanek 等学者的启发,世界范围内的诸多学者呼吁设计应从伦理、情感和以人为本的价值观出发,响应社会责任问题,包括可持续发展、社会包容性、发展中世界问题、两性平等和环境保护等,以及事实上占人口绝大多数的被排斥、受歧视和低经济水平的弱势群体的需求。

Richard Buchanan 在 1992 年将 Horst Rittel 的“抗解问题(Wicked Problems)”理论作为探讨设计思维特质的一个切入点^[13]。他指出,设计主题的“不确定性”和“抗解性”将设计学和其他基于观察和理解既定主题的学科区分开来。设计师在实践中要面对“问题定义”这一个动态任务。设计的历史不仅仅是人造物的历史,还是设计师对设计主题的不断变化的观点的历史,设计的产物是为了表达这些观点而构思、规划和生产的具体对象。设计行为是对人们生活的一种精辟的和编排好的干预,而不是直接的抗议或示威,设计的公共影响潜力在于提高意识^[14]。现实世界中没有撤销按钮,设计师有在其创造的事物中融入价值观

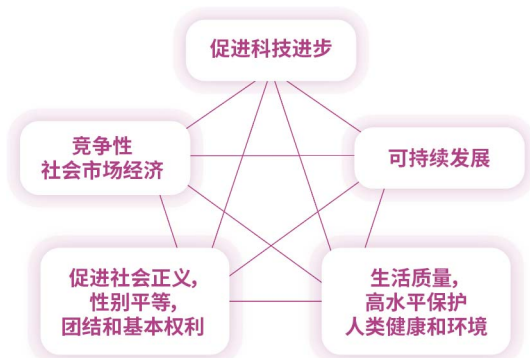


图 1 基于《欧洲联盟条约》的规范性锚点
Fig.1 Normative anchor points derived from
“The Treaty on the European Union”

的能力和义务,因此对于这种能力的谨慎反思是极其必要的。

产品、服务、交互、体验和流程等复杂系统作为当今设计创新的产物,已经成为社会环境的组成部分,而人工智能是一把双刃剑,在不断地加速这些系统的迭代速度的同时,也将放大每一个设计决策可能对公共未来产生的影响。可以预见的是,人工智能应用的潜在风险将会在上述的经典设计伦理领域中加深已有的困境,同时引发一系列新的问题。当技术以一种压倒一切的态势影响人们的生活时,设计师首先应秉承一种理性的应对态度,以适时适度地应用技术构建人与产品以及系统之间的和谐关系作为设计的基本原则,发挥技术变革的正面能量,减少其对于人类发展的副作用^[15]。

2.3 人机交互第三范式下的设计创新任务

人机交互的范式模型在人工智能时代的定义尤为重要,因为人工智能正是试图效仿人类的思考模式,为人机耦合提供新的场景和接口,促使普适计算的未来逐步实现。范式的存在可以有助于理清哪些问题是重要的、如何评价解决方案的效果并进一步指导下一步的工作。

人机交互的前两波浪潮分别是以优化人机匹配为重点的人因工程 and 将理论研究从计算机内部转向人类头脑中的认知革命。而为了处理诸如价值观和社会问题之类的前两个浪潮难以应对的现象,人机交互研究需要借鉴其他学科的方法,比如情境行动、民族志研究以及上文提到的价值敏感设计和参与式设计等^[16],这体现了将交互视为一种意义创造的形式第三种范式。

在大规模的数据运算下,人工智能应用中信息结构和处理流程的复杂程度大幅提升,传统的“以任务为中心”的可用性评估可能已不再适用。同时,自然人机交互的发展强调人机关系的和谐性和交互通道的多样性,将人类情感交换也作为信息流来建模,使情绪和感受成为可被计算的数据。人机交互的第三波浪潮正是一种情境视角的设计范式,关注在交互系统中建构意义、设计者/用户/接口在情境中应占据的位置以及明确设计中的价值观。在情境视角下,系统周围发生的事情比界面上发生的事情更值得关注,“不做”和“做”的决策一样重要。

三种范式并不互斥,不同范式的工作可以互相交叉,并对设计采取不同的立场。而设计师的工作方法则具有着一贯性:关注产品、服务、交互和流程的背后发生了什么,并从设计研究中提炼出一系列原则来指导具体的设计实践,所谓的“设计原则”即是交互范式中意义和价值观的体现。设计的社会责任潜力使其在“某种程度上自然适合第三种范式”^[16]。第三种范式重视并应对设计情况的复杂性,使得对情境的理解和建构成为设计创新的核心任务。

3 设计决策的责任维度框架

负责任创新与人工智能伦理都具备“抗解问题”的典型特征:当触及表面问题时必然会暴露出其他新的问题。而设计思维已被证明是解决“抗解问题”的有力途径之一。法国学者 Pavie 和 Carthy 试图将责任风险假设的定义、规范和测试嵌入到设计思维经典的五步创新流程中,以促进创新的参与者更好地了解人类和社会的总体需求,明确考虑组织的经济必要性,有助于将创造力的需要与影响的监测结合起来^[17]。

另外一个极具参考价值的项目是社会技术整合研究(STIR)^[10]。该研究是一个负责任创新实验项目,将人文社科学者纳入了三大洲超过25个科学和工程研究实验室,旨在将责任和价值观的规范和调节从创新的上游(政策制定)和下游(成果监管和市场选择),引入中游——“大型、分布式和动态决策过程的实现阶段”^[18]——也就是创新发生的过程中。STIR 为跨学科合作者开发了一套责任协议,使科学家和创新者可以实时地将责任嵌入决策。STIR 实验室定期使用该协议,责任对决策的具体影响体现在研究方向、实验设计、安全和环境的变化上,并被记录为“生产性研究中断”。

学科交叉研究使设计从业者和学者可以在系统开发的早期参与到规则的制定之中,将参与式和以人为本的设计理念和方法整合到创新的各个阶段,并将后果评估纳入设计决策。受 STIR 协议的启发,结合信息技术背景下主要的四个设计创新维度(4P):产品、流程、目的和人,也可以制定一个初步的设计决策责任协议,ICT 领域的设计决策责任协议见表1。从确认做什么、怎么做、为什么以及涉及到哪些利益相关者的角度,采用相应的责任指标进行全程调节。在此基础上,可以针对人工智能应用的设计进行更加谨慎的责任管理。

1) 产品。在市场调研、竞品分析和用户研究等设计流程前期的产品机会点的定位和定义决策上纳入反思维度。这意味着设计师需要突破自己先前知识和经验的局限,考虑对于用户需求的选择和描述是否具备客观性、包容性和公平性,尽可能地从多元的角度来评估设计的切入点是否有责任隐患。比如,如果交互界面的用户模型的大数据分析固化在某一年龄段和社会阶层,这可能会对其他用户群体的使用体验

表1 ICT领域的设计决策责任协议
Tab.1 Design decision responsibility protocol in the ICT field

设计创新	设计决策构件	责任维度
产品 (做什么?)	机会判断	反思性
流程 (怎么做?)	方案选择	响应性
目的 (为什么?)	考量因素	反思性+审慎性
相关者 (涉及谁?)	后果评估	预期性

造成持续的负面影响；在制定聊天机器人任务时，如果没有对潜在的有害内容进行充分地规范，机器人的过滤和应对策略则可能会出现偏离设计初衷的发展。其中，人工智能辅助设计的产品和系统的包容性需要被重点关注，包括如何在构架设计中尽可能地纠正先存偏见以及缩小不同用户群体对于相关技术应用的认知偏差。

2) 流程。响应性维度的重要特征在于情境化和互动性以及互动中对于可能产生的新知识和新观点保持一种灵敏性^[19]。在产品的研发过程中，响应式设计对不同的方案路线保持开放的态度，对可能产生的技术和突发偏见问题进行实时地测试、分析和干预，通过多个短期的反馈循环及时地对产品进行迭代。响应性的一大挑战是如何应对人工智能应用中大数据的规模经济问题，这给设计师的个人反应能力带来了压力。因此在这一环节，设计师更加需要与学科交叉团队协作，高效地使用信息化的设计工具，建立更具有适应性的可用性试验策略。特别是在使用人工智能进行辅助决策的系统中，比如人力资源管理、医疗诊断和自动驾驶等，设计师对信息处理模型和情境变化的敏锐洞察力需要发挥作用，以便参与到对决策可能产生的偏差的紧密追踪、解释和修正中。响应式设计是对于抗解问题中“不确定性”的一种积极回应，它采用务实的、试错的和缓进的原则不断地自我调整与完善。

3) 目的。在设计目的决策中往往涉及多个考量因素，反思性和审慎性维度将有助于将责任嵌入不同因素的权重评估中，尽可能地减少产品和系统的副作用。设计的目标策略很容易陷入“提供某种技术应用”和“高效地获得盈利”优先的商业考量中^[19]，进而造成包容性的缺失和潜在责任风险的提升。审慎性维度在这一决策环节的引入也意味着，参与式设计应该在设计创新的早期方向规划和阶段性成果评价中都发挥更大的作用，用于矫正人工智能可能激化的社会伦理问题和用户体验问题。设计师通过对各个考量因素进行谨慎地假设和推演，来更有针对性地指导参与式设计活动的目的、标准和方法。比如，基于用户消费行为的个性化智能推送因为缺少规范和制约，可能在某种程度上削弱了用户的自主探索能力和判断力，设计师可以通过与用户协作观察整个使用体验中实际发生的情况，来探讨是否可以在用户个人消费习惯的健康程度和商业销售目标之间形成一种制衡关系。

4) 相关者。人工智能时代的信息技术变革具有迅速而广泛的传播潜力，与传统的设计师和用户之间的简单二元关系相比，智能产品和系统的间接利益相关者范畴越来越大。前文已提到，预期性与反思性有很多共同点，但前者更强调可能的以及理想的未来。预期性设计既要考虑设计决策对相关利益相关者可能产生的计划内、外的影响，也要使这些群体参与到

对未来进行预先推测和前瞻性讨论中，包括何时采用或拒绝某种技术。对设计师来说，人工智能应用不应被简单概念化为一种可被设计、开发、交付以及最后被选择的技术产品，而是一种不同利益相关者在持续学习和协作过程中不断自我更新的一个“开放空间”^[19]。

4 反思与期望

创新实践是价值实践的同时，也是利益驱动的价值选择。如果忽略创新基于利润追逐的动力来源，过度的伦理中心主义和生态中心主义倾向可能会导致负责任创新被束之高阁^[11]。而设计是一个极佳的切入点 and 有效的实用方法。根据设计思维理论，衡量设计创新的三个重要维度是市场可取性 (Desirable)、商业价值可行性 (Viable) 和技术可行性 (Feasible) 考量。负责任设计能够在满足基本的创新动力和条件的基础之上，将创新的目标转向责任敏感的主题，快速地将责任评估嵌入整个设计流程，最终以智能产品、服务和系统的产出形式对于技术和设计伦理问题进行直接的干预。综上，负责任的人工智能应用的设计创新主题趋势包括以下方面。

1) 为更广泛的人群设计。目前，人工智能主要应用在消费市场领域，包括内容推荐、聊天机器人、图像识别、机器翻译、自动驾驶、运输物流、智能医疗和信息化教育等，其中一些技术路线已趋于成熟，开发成本也相应降低。设计师可以将目光转向更广泛的人群，在高度产业化的技术水平和配套资源的基础上，引入创新设计的责任评估机制，开发稳定和可持续的智能产品和系统，加速发展中国家良性的数字化基建发展。除此之外，农业养殖、工程建筑和生态管理等机械化改造仍不彻底的传统领域，也可以在负责任设计框架下，依托 AI 技术进行价值敏感的数字化转型。

2) 更深入的无障碍设计。人工智能可以协助交互设计突破图形用户界面传统 WIMP 接口模式的局限，通过基于隐喻和转喻等多种语言表达方式以及基于自身认知的多模态交互体验，进行信息空间与物理世界的认知迁移，并对人类的情境、意识和情绪感知进行计算^[20]。也就是说，可以在为障碍人士设计物理辅助用品和环境之外，从意识和感知的角度更深入地探索障碍人士对外部世界的编码解码方式。这有助于设计师从全新的角度来改善和提高障碍人群的生活质量和生命体验，更好地保障他们对世界、社会及文化生态的理解、赏析、创造的能力和权利。

3) 更包容的全球化文化生态设计。人工智能可以通过情感计算、人格计算、社会计算和文化计算等科技人文的方法，研究人类的体验模式与愿景，进而创造新的文化形态，拓展人类认知边界^[20]。这使得设计师可以更有效地将包容性设计融入全球化文化生态的叙事当中，比如社会责任主题的新媒体艺术创作

以及非物质文化遗产的数字化保护工程等。同时需要注意的是,人工智能经过 60 多年的发展,正在以其特有的方式高效、精准和稳定地实现人类的创作闭环^[21]。在艺术创作领域, AI 已经从素材处理工具逐步进化为人类的协同创作者。怎样在负责任设计的框架下与人工智能进行共创,也是设计师需要深思的问题之一。

5 结语

人工智能的发展为世界提出了诸多新挑战,也提供了全新的操作手段。但是,设计从业者和学者对手中这样的新工具的深度的理解和多元的思辨仍有很大的改善空间。面对科技发展所带来的不可预测的未来,设计师是这场社会责任唤醒运动中不可或缺的重要角色,应走到跨学科团队的建设前线,思考技术应用的伦理基础,参与决策评估和准则制定。为了靠近这个目标,设计师需要勇敢地踏出舒适区,探寻与科学技术和人文社科领域学者共享知识和共担责任的新方式,成为将聚光灯转向世界盲点的活跃引导者和先锋思潮的积极贡献者。负责任的设计必然离不开负责任的设计师,设计教育也应相应地注重培养设计师对于个人、职业、社会和技术伦理的理解。

在充满不确定性的技术应用浪潮中,设计师始终会面临设计主题的定位定义不清晰、与其他学科的分工分责不明确的问题。这些问题并没有标准答案,在每一次设计范式的变革中,它们随着设计师对于技术发展的不断反思和干预而慢慢化解。而更重要的问题是,设计师秉持怎样的价值观,以及怎样承担将价值观嵌入设计决策的义务。

参考文献:

- [1] 谢洪明, 陈亮, 杨英楠. 如何认识人工智能的伦理冲突? 研究回顾与展望[J]. 外国经济与管理, 2019, 41(10): 109-124.
XIE Hong-ming, CHEN Liang, YANG Ying-nan. How to Comprehend the Ethical Conflict of Artificial Intelligence? A Literature Review and Prospects[J]. Foreign Economics & Management, 2019, 41(10): 109-124.
- [2] FRIEDMAN B. Value-Sensitive Design[J]. Interactions, 1996, 3(6): 16-23.
- [3] BUOLAMWINI J, GEBRU T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification[J]. Proceedings of Machine Learning Research, 2018(81): 1-15.
- [4] BRUNDAGE M, AVIN S, CLARK J, et al. The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation[EB/OL]. (2018-02-01) [2021-01-20]. <https://maliciousaireport.com>.
- [5] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems[EB/OL]. (2019-03-01) [2021-01-18]. https://ethicsinaction.ieee.org/?utm_campaign=EAD1e&utm_medium=PR&utm_source=Web&utm_content=geias.
- [6] LESLIE D. Understanding Artificial Intelligence Ethics and Safety: A Guide for the Responsible Design and Implementation of AI Systems in the Public Sector[EB/OL]. (2019-06-11) [2021-01-05]. <https://doi.org/10.5281/zenodo.3240529>.
- [7] 中华人民共和国中央人民政府. 新一代人工智能发展规划[EB/OL]. (2017-07-20) [2021-01-05]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm. Central People's Government of the People's Republic of China, Development Planning for a New Generation of Artificial Intelligence[EB/OL]. (2017-07-20) [2021-01-05]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- [8] 中华人民共和国科学技术部. 发展负责任的人工智能: 新一代人工智能治理原则发布[EB/OL]. (2019-06-17) [2021-01-05]. http://www.most.gov.cn/kjbgz/201906/t20190617_147107.htm. Ministry of Science and Technology of the People's Republic of China. Developing Responsible AI: New Generation AI Governance Principles Released[EB/OL]. (2019-06-17) [2021-01-05]. http://www.most.gov.cn/kjbgz/201906/t20190617_147107.htm.
- [9] 晏萍, 张卫, 王前. “负责任创新”的理论与实践述评[J]. 科学技术哲学研究, 2014, 31(2): 84-90.
YAN Ping, ZHANG Wei, WANG Qian. Review on Theory and Practice of “Responsible Innovation”[J]. Studies in Philosophy of Science and Technology, 2014, 31(2): 84-90.
- [10] OWEN R, BESSANT J, HEINTZ M. Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society[M]. West Sussex: John Wiley & Sons, Ltd., 2013.
- [11] 刘战雄. 负责任创新研究综述: 背景、现状与趋势[J]. 科技进步与对策, 2015, 32(11): 155-160.
LIU Zhan-xiong. Review of Responsible Innovation: Background, Current Situation and Trend[J]. Science & Technology Progress and Policy, 2015, 32(11): 155-160.
- [12] 维克多·帕帕奈克. 为真实的世界设计[M]. 北京: 北京日报出版社, 2020.
PAPANEEK V. Design for the Real World[M]. Beijing: Beijing Daily Press, 2020.
- [13] BUCHANAN R. Wicked Problems in Design Thinking[J]. Design Issues, 1992, 8(2): 5-21.
- [14] MARKUSSEN T. The Disruptive Aesthetics of Design Activism: Enacting Design between Art and Politics[J]. Design Issues, 2013, 29(1): 38-50.

(下转第 36 页)