

# 智能广告的技术伦理风险及其治理新探

杨先顺 李 婷

**摘要** 随着人工智能技术在广告领域的广泛应用,智能广告中技术应用引发伦理风险的问题逐渐成为焦点。从技术伦理学视角来看,目前智能广告中主要存在两类伦理风险:一是基于多端数据挖掘产生的消费者隐私侵犯、信息价值侵害和创意知识产权侵犯,二是基于深度学习算法产生的广告运作壁垒、消费者阶层分化和消费异化。智能广告中技术伦理风险的治理应先从观念化价值、经验化价值以及技术化价值三个方面建立治理原则,再从智能广告的设计、测试、应用、复盘四个环节制定治理策略,即伦理嵌入、伦理评估、伦理调适和伦理确认,以期规避和约束智能广告中的技术伦理风险。

**关键词** 智能广告;技术伦理学;伦理风险;伦理价值;伦理治理

**中图分类号** G206;F713 **文献标识码** A **文章编号** 1672-7320(2023)03-0100-11

**基金项目** 国家社会科学基金项目(19BXW100)

在人工智能掀起技术创新和产业革新浪潮的今天,人工智能技术已经成为国际竞争的新焦点,中国也进一步将人工智能的发展上升至国家战略层面。在广告领域,早在2012年人工智能技术已经在程序化购买、实时竞价等互联网广告中开始使用,随后逐渐渗透到整个广告产业发展的方方面面。智能广告能够实现消费者洞察、内容生产、精准投放及效果反馈的智能化,进而促进广告生产和传播效率的最大化,并深刻地改写了广告运作的逻辑。然而人工智能技术在优化广告运作效率的同时也带来了诸如数据泄露、隐私侵犯、算法黑箱、抄袭侵权等伦理风险,如何规避智能广告中的技术伦理风险,解决伦理问题,减少甚至规避人工智能技术对广告行业带来的负面影响,促进广告业健康、有序发展,是当今广告学界和业界必须高度关注并重点解决的问题。

## 一、研究综述

目前,学界对于智能广告伦理的研究成果较少。从国外研究情况看,截至2022年11月Web of Science数据库有关智能广告伦理的论文仅7篇(经主题检索,剔除1篇研究医学伦理的论文,标题检索为0篇),经甄别仅2篇论文直接论述了智能广告伦理,其他5篇只是提到而没有充分论述,国外其他数据库的相关文献不超过此数。在直接论述智能广告伦理的2篇文献中,Rodgers和Nguyen探讨了6种与道德哲学结合的、为线上消费者购买产品的算法购买决策路径<sup>[1]</sup>。Vlacic等人讨论了人工智能在市场营销中的数据保护和伦理问题<sup>[2]</sup>。从国内研究情况看,根据在中国知网(CNKI)检索的数据,截至2022年11月国内与智能广告伦理直接相关的论文仅有5篇(主题检索,剔除不直接相关论文),无相关研究专著。如果将范围拓展到相近领域,会发现数字广告伦理、计算广告伦理的相关研究论文分别为5篇和3篇(主题检索),无相关研究专著。这些论文主要探讨了智能广告、数字广告或计算广告等存在的伦理问题,如蔡立媛等人认为智能广告中的技术应用不仅强化了旧有的广告伦理失范问题,而且引发了新的广告伦理危机——广告泛滥、广告欺骗、“时空剥夺”“视听暴力”以及广告资源无效配置等“时空侵犯”问题<sup>[3]</sup>。姜

智彬等人指出当前智能广告的数据困境,如数据孤岛、数据安全堪忧、隐私数据泄露问题,并提出“数据脱敏”、偏差矫正、法治监管、促进共享开放等解决措施<sup>[4]</sup>。李名亮借助弗洛里德的RPT模型,从信息伦理的角度对智能广告的伦理风险问题进行了探讨<sup>[5]</sup>。针对智能广告的诸多伦理失范现象,有学者强调需要加强智能广告的监管工作,比如通过对智能广告进行数字化处理实现内容分类,然后建立标签分明的数据库系统,以便快速准确地做出监管行为<sup>[6]</sup>,等等。

值得注意的是,国内外有关智能广告伦理的研究均忽视了一个重要的视角,即技术伦理学的视角。德国技术哲学和技术伦理学家格伦瓦尔德认为:“技术伦理学的任务,就是依据理性论辩的原则,建立起技术评价和技术决策的一套规范基础,目的是借此为经过伦理思考和能够负责任的决策提供帮助。”<sup>[7]</sup>(P7)目前虽然已有学者开始从技术伦理学视角观照人工智能伦理,但未能将这一视角延伸到智能广告的伦理领域。我们认为该视角更利于揭示智能广告伦理问题产生的内在机制,从而提出更有针对性和专业性的治理策略。

本文在前期深度访谈和参与式观察的基础上,从技术伦理学的视角探讨智能广告中的技术伦理风险及其治理。首先,通过对参与智能广告运作的企业人员进行了深度访谈,以获得第一手的行业应用资料,了解行业状况。按照“理论饱和原则”,一共对13名被访对象(年龄分布在25-40岁之间)进行了线上或线下的一对一访谈,其中女性7人、男性6人。被访者基本上都是深度参与到智能广告运作中的工作人员,其中非技术人员10人、技术人员2人,还有1人为人工智能公司的算法研究员(限于篇幅具体过程和不在本文赘述)。其次,选取深圳市腾讯计算机系统有限公司和深圳市柒彩科技有限公司作为参与式观察的对象,本文第二作者主动参与到智能广告的具体运作中去,亲身体验和发现智能广告中技术应用的伦理风险,从实践中获得宝贵的资料,从而深化对伦理治理的思考和研究。

## 二、智能广告中的技术伦理风险

伦理风险是指在人与自身、他人、自然、社会的伦理关系方面所产生的不确定的负面伦理效应,诸如伦理关系失调、社会秩序失衡、自然失控、人们行为失范、心理失衡等伦理现象<sup>[8]</sup>。将风险置于伦理的范畴内考量,其研究范围则在“人—社会—自然”这一复杂的体系中展开,这些不确定的负面伦理效应一旦产生,将会严重威胁到这一复杂体系的稳定。技术伦理学视域下的伦理风险主要是指技术在应用过程中给“人—社会—自然”这个复杂体系带来的伦理关系方面不确定的消极影响。

### (一) 基于多端数据挖掘的伦理风险

“数据”首先是一个科学的概念,是指“能够被数字化传递或处理的数字形式信息”<sup>[9]</sup>(P9),是对于已经发生的行为或者事件的客观记录与呈现。数据与信息之间存在着密切的关联,客观的数据中包含的是人或物的信息。因此“数据代表着人或物所具备的特征”<sup>[9]</sup>(P10)。从社会视角来看,这些特征都是基于人而产生的。随着人工智能技术的广泛应用,消费者的各种网络行为和社会活动都能够转换成量化的数据,成为广告主实时追踪和监测的对象。

#### 1. 基于数据监视的消费者隐私侵犯

消费者隐私侵犯在互联网广告中业已存在,但在智能广告中人工智能技术更容易从各种网络行为中获取大量的原始数据,使得消费隐私信息触手可得,因而更容易导致消费者隐私权受到不同程度地侵犯。消费者隐私数据,一般指消费者姓名、年龄、证件号码、联系方式等与个人身份密切相关的敏感信息。目前智能广告中使用的消费者数据是将消费者在网络上留下的搜索、浏览、点赞、分享等行为痕迹进行实时监控,然后以数据的形式存放于服务器中。智能广告中的消费者智能洞察以及广告的智能投放都是建立在对消费者个人数据在线实时监控的基础之上,比如微信通过对所有用户的点赞、分享和社交数据进行搜集,然后按照一定的标准对所有用户进行标签分类,从而将消费者的各类数据同相应的广告主需求相连接,实现广告主在微信朋友圈的精准广告投放。

在所有智能广告技术伦理问题中,隐私侵犯问题是最普遍也是危害最大的伦理风险。这主要是因随着人工智能技术的日益强大,消费者个人信息数据获取的方式越来越丰富,获取的成本也越来越低,挖掘潜在价值的可能也越来越大,但是这也为数据开发和收集过程中的控制和责任划分增加了难度。尤其是广告主们在追求更大经济利益的过程中,往往会侵犯消费者的个人隐私。为了遏制对消费者的隐私侵犯,我国也加快了有关法律的立法进程,如《中华人民共和国个人信息保护法》的颁布实施等。然而,因为互联网上的数据是海量的,加之执法人员的相对不足以及有的消费者隐私保护意识和能力不够,这使得对隐私侵犯的监督管理面临新的挑战。

### 2. 基于数据开发的信息价值侵害

这是智能广告中出现的新的伦理风险。随着人工智能技术在广告中的应用和深入,消费者数据的获取和开发方式已经不仅仅是直接从网络上进行记录和收集,还可以通过算法来进行预测性分析。智能广告洞察是基于全平台信息共享所建立起的消费者数据库。广告主为了精准预测消费者的需求,会充分挖掘消费者数据的现有信息价值,甚至想方设法对其潜在价值进行预测,一旦消费者个人数据被无节制地开发或利用,消费者的信息价值将会遭到不同程度的侵害。信息价值开发很大程度上都依赖于网络信息数据的开放性和共享性,然而正是这些属性给消费者信息价值带来了诸多安全隐患。人工智能技术在追踪和记录消费者数据的过程中,消费者的个人信息价值在生产、储存和应用的任一环节都有可能被一些别有用心的企业或个人进行利用,如果发生任何开发不当的行为,后果将不堪设想。当消费者数据成为智能广告运作过程中的重要战略资源时,其具有的利益和诱惑性也会越来越大,因此消费者信息价值遭到侵害的可能性也就更大。同时,随着人工智能技术的发展和成熟,信息价值侵害的方式也变得更加隐蔽和多样,这也就意味着保护消费者信息价值安全的难度更大。

### 3. 基于数据重组的创意知识产权侵犯

创意知识产权侵犯在传统广告中就已出现,但在智能广告中又呈现出新的特点。智能广告的创意生产依赖强大的语音识别、自然语言处理等人工智能技术获取庞大的广告创意基础元素数据库,具体包括颜色、字体、场景、图片、视频等内容。参照程序化目标变量,根据不同场景,经过机器自动识别对各个元素数据进行排列组合随机生成新的广告创意,整个过程不仅节省了时间和人力,而且由于花费时间短,广告主可以根据自己的需要来实时修改广告创意内容,极大地提高了广告创意的效率。然而在机器智能化排列组合中,是对已搜集的数据进行重组,即对已有的全网公开的基础创意元素进行直接复制和修改,事先难以征得原有元素创作者的同意,因此,智能广告创意生产中的数据重组往往会对这些广告创意原始创作者们的知识产权造成侵犯。

知识产权具有双重属性,一是相对权利,二是绝对权利。前者表现在知识产权只是在特定的时空范围内有效,超出限定的范围后知识产权将失去其约束作用;后者则表现在知识产权是绝对的排他和垄断的,在没有取得产权所有者同意或授权之前,任何人都不得擅自使用。也就是说,知识产权是一种既能自由共享又能合理使用的无形资产,这就导致知识产权在实践过程中始终存在自由共享和限制使用的矛盾,而在智能广告中的技术应用过程中,这一矛盾变得更为突出。特别是在智能广告的内容生产环节,一些企业或个人为了追求经济利益,往往会在获取基础广告创意元素时,强调知识产权的相对权利属性,加上人工智能技术的发展和成熟使得广告创意作品的获取、复制和传播变得更加方便和迅速,会导致越来越多的知识产权侵权行为,不利于知识产权的保护。尤其是随着智能广告中数据挖掘技术和算法的日益强大、数据重组的日益频繁,广告作品复制、传播将会变得更加便捷和多样,这使得侵犯知识产权行为变得更加难以控制,也增加了对图片、视频等广告作品版权保护的难度。

### (二) 基于深度学习算法的伦理风险

基于大数据算法的智能广告系统为广告内容生产和个性化推荐提供各种服务信息和选择建议,能够辅助广告主找到最佳的广告投放方式,帮助消费者获得准确适合的广告信息。然而,在算法驱动智能

广告发展的同时,由算法引发的伦理风险也日益明显。

### 1. 基于“算法黑箱”的广告运作壁垒

基于“算法黑箱”的广告运作壁垒是智能广告特有的伦理风险。算法决定了人工智能技术运算的结果,是人工智能技术应用和发展的基础。“算法就是一系列指令,告诉计算机该做什么。”<sup>[10]</sup>(P3)算法的本质就是为了获得目标结果而按照设定好的程序运行的一系列指令,一般难以转化为通俗易懂的语言,这也导致智能广告中技术背后的决策逻辑难以让人理解、预测和评估,特别是与云数据、人脑互联等前沿技术结合后,技术系统所产生的日益增长的风险数量使得技术变得更加不可预测<sup>[11]</sup>,也就是普通消费者不能判断人工智能信息的来源真伪和运作过程,称之为“黑箱”。“黑箱”作为控制论中的一种隐喻的概念,主要指不能够被大众知晓或直接从外部观察到内部运行状态的系统<sup>[12]</sup>。深度学习算法是直接从事物外部或原始特征入手,能够进行自动学习并获得一定的认知。而人工智能系统在输入数据和输出认知结果之间,往往都存在着人们无法洞悉的“隐层”,这就是“算法黑箱”<sup>[13]</sup>。

“算法黑箱”是由于算法的自主性所引发的伦理风险。传统意义上的程序员是以手动方式编写代码,通过这种方式来定义简易算法的决策规则和权重分配。而在智能广告系统中,广告主和消费者并不知道人工智能算法的工作原理,不知道智能化广告创意是怎么生产出来的,这就带来基于“算法黑箱”的广告运作壁垒的伦理风险。智能广告系统在人工智能新算法的优势作用下取得了快速发展,然而面对迅速更新发展的新算法,广告行业的相关规范还未及时更新,以至于广告从业者在对“算法黑箱”所带来的伦理风险时不能迅速提出解决方案。同时智能广告的运作因严重依赖代码与数据样本的机器学习,难以保证稳定性,多番重复运行的样本数据也极易误导机器,机器可能因假设误导而做出错误的道德决策,一旦技术“自主”失控,风险便随之而来。

### 2. 基于算法歧视的消费者阶层分化

随着消费社会的到来,消费者阶层分化现象日趋显著,但只有在智能广告中才出现了基于算法歧视的消费者阶层分化。智能广告运作过程中的数据如果预先存在偏见,则往往会导致算法歧视。算法歧视是指程序的设计者将自身的偏见有意或无意地嵌入到算法的程序设计中,使本应保持客观中立的算法带有了某种偏向,从而在具体应用中表现出某种歧视。也就是说,算法歧视在设计算法之前就已经存在。算法歧视一般具有隐蔽性和无意识性的特点,因此在智能广告运作的一开始很难被察觉,往往可能会造成广告信息传播的不公正、不对称,从而使部分群体的利益受到损害。分类是算法决策的重要前提,但是分类的标准是基于消费者实时的行为数据设定的,是临时变动的,所以每次算法分类标准的更新都会导致其预测结果的变动,进而导致相关群体的利益也会受到影响<sup>[14]</sup>。在智能广告的算法运作中,即使是身份、职业相同的不同消费者也可能遭受不平等的对待。例如,电商平台在营销过程中的“大数据杀熟”与价格歧视,商家基于消费大数据对用户进行精准的差别定价,以隐蔽的方式对价格不敏感人群实施高定价,或者对平台的活跃用户进行正常定价,而对新晋用户采用低价或优惠策略,这是商家利用信息的不对称最大限度地攫取消费者剩余价值的一种手段,但是无形之中已经对消费者的身份阶层进行了重新划分。在智能消费者洞察过程中,对消费者进行标签分类就是使用算法本身所带来的歧视,也就是说,歧视往往伴随着算法的无意识属性就能够对消费者身份产生影响,无形中对消费者进行了不同程度的阶层划分,并且这种阶层划分只是基于企业或个人的经济利益,缺乏对人文价值的考量,极有可能导致对消费主体的刻板印象和阶层区隔。

### 3. 基于算法预测的消费异化

消费异化现象虽然自人类进入消费社会即已出现,但基于算法预测的消费异化则是智能广告的“原罪”。前文提到,消费者数据一方面可以通过对消费者在网络上留下的行为痕迹进行记录和收集获得,另一方面还可以通过算法的预测性分析获取,从而对消费者需求进行精准预测,达到为消费者“创造需求”的目的。但是在实际操作过程中,很多个人和企业为了追求更大的经济利益,通过不断为消费者推

送和消费者需求强相关的信息,来引导消费者产生持续不断的购买行为,这不仅稀释了原本广告信息的内容价值,而且会误导消费者冲动、盲目消费,导致消费者的购买行为不再基于自身的真实需求,而是基于算法预测为消费者精心设计的“虚假欲望”。

“异化”这一概念源远流长。马克思在《1844年经济学哲学手稿》中深入研究了“劳动异化”,提出了“劳动所生产的对象,即劳动的产品,作为一种异己的存在物,作为不依赖于生产者的力量,同劳动相对立”的深刻洞见<sup>[15]</sup>(P52)。马尔库塞、弗洛姆、鲍德里亚等人则对“消费异化”进行了反思和批判,鲍德里亚认为“消费是用某种编码及某种与此编码相适应的竞争性合作的无意识纪律来驯化人们”<sup>[16]</sup>(P78)。中国学者陈力丹指出,异化“是这样一种现象:人们自己创造的某种东西,反过来却被它所控制。”<sup>[17]</sup>本文作者之一曾探讨了网络广告的技术异化问题<sup>[18]</sup>。

在智能广告领域,算法预测技术本身就是为了挖掘和刺激消费者潜在的消费欲望。在广告推送过程中,算法预测为抢占消费者注意力,会想方设法对消费者个人信息进行深度开发,甚至会多次修改和重组,从而为消费者量身打造个性化的需求环境,这些需求既包含了消费者的真实需求,也包含了大量被表象覆盖的虚假需求,一旦消费者的生活被这些虚假需求吸引和控制,并在算法优化的过程中不断进行周期循环和需求扩张,不仅会引发消费者对算法的隐性依赖,还会使其不断陷入盲目消费的陷阱,自觉成为广告的附庸和消费的奴隶,从而引发严峻的消费异化问题。此外,在算法预测的推荐机制下,消费者看到的广告信息都是与其所在场景高度匹配的个性化内容,以至于在面对这些广告信息时缺乏足够的抵抗力,无形中也剥夺了其在购买行为中的选择权利,削弱了人们在消费时的自主性。

### 三、智能广告中技术伦理风险治理的伦理价值与伦理原则

智能广告中技术伦理风险治理需要宏观的伦理价值和伦理原则作为指导。因此,我们首先从技术伦理学的视角确立智能广告中技术应用的伦理价值,在传统技术伦理学基本原则的基础上对智能广告中技术应用的伦理原则进行建构,为智能广告中技术伦理风险的治理提供依据。然后,在此基础上从智能广告中技术应用的各个环节分别对其伦理治理进行探讨和研究,形成合理、规范的伦理治理策略,从而有效解决或规避智能广告中的技术伦理风险,促进智能广告的进一步发展和完善。

#### (一) 智能广告中伦理风险治理的伦理价值

要厘清智能广告中技术伦理风险的治理策略,就必须要从技术伦理学的角度,对智能广告中技术应用的伦理价值进行反思,进而为其伦理评估和伦理调适提供可参考的原则和标准。建构智能广告中技术伦理风险治理伦理原则的首要前提是确立人工智能技术在广告应用过程中应当遵循的伦理价值。价值就是“人们为了获得美好生活而为之持续奋斗的信念”<sup>[19]</sup>(P72)。弗雷德曼(B. Friedman)等人曾提出“人类福祉、财产权、隐私、公正、普适性、信任、自主性、知情同意、可解释性、同一性、平衡和环境可持续性”这12项技术伦理价值<sup>[20]</sup>(P55-59)。在技术伦理学领域,技术应用的伦理价值一般可以分为观念化价值、经验化价值和技术化价值三类<sup>[21]</sup>。

观念化价值指的是在理论上或观念上应该重视的概念价值,智能广告中技术应用的观念化价值应和人类一般伦理价值保持一致,也就是人类本身的核心诉求,在智能广告领域表现为相关利益人的诉求,其中和技术直接接触的设计者、开发者、使用者等直接利益相关者往往将尊重、责任、共享等价值作为诉求,而对于受技术间接影响的消费者和其他间接利益相关者则将隐私保护、公正、安全等价值作为诉求。经验化价值指的是从具体理解语境和操作经验中对观念化价值进行实践确认,如各种技术规范、技术制度等具有的约束性、预防性和包容性等,智能广告中技术应用的经验化价值应该以利益相关者的诉求为前提,对涉及人工智能技术的具体行为进行价值设定,也就是指导技术应用的规范制度,其本质是促进智能广告更好地为人类服务。技术化价值指的是在智能广告中与人工智能技术相关的价值,包含可解释性、可控制性、可持续性、可预防性等,这一价值可归结为实现人工智能技术的“负责任创

新”。“负责任创新”是近年来技术伦理学的热点概念,最早出现于2002年美国学者加斯顿(D.Guston)在哥本哈根螺旋会议上宣读但尚未正式出版的文章中,后被德国学者赫尔斯特罗姆(Hellstrom)于2003年引用并公开发表后,受到国内外学者广泛关注<sup>[22]</sup>,英国学者理查德·欧文(Richard Owen)等人提出:“负责任创新是一个关于未来的集体追求,它通过对当前的科学和创新进行响应性的管理来实现”,而“响应性,即利用集体自省的过程来设定创新的方向,并影响创新之后的路径和速度”<sup>[23]</sup>。负责任创新旨在寻求技术创新与社会责任的价值平衡,其本质是可持续发展和健康有序发展理念的延伸。负责任创新既是一种抽象的价值理念,又是一种具体的伦理实践。

### (二) 智能广告中技术伦理风险治理的伦理原则

目前学术界已有一些学者在探讨人工智能的伦理原则,并提出了一些富有建设性的观点,如杨庆峰认为:“人工智能伦理的原则可以划分为两个来源:强调如自由、公正、安全的人本主义原则以及强调如可持续、负责任的、可信的、可靠的、可治理的技术主义原则。”<sup>[24]</sup>孙伟平等人主张:“以‘人本原则’为‘基点’和‘统领’,以公正原则和责任原则为主干,这就构成了人工智能发展的整体性的伦理原则体系。”<sup>[25]</sup>本文在学界已有成果的基础上,通过以伦理价值为引领、以伦理治理需求为问题导向的双向推导法建构智能广告中技术伦理风险治理的伦理原则。

伦理原则指的是为了实现伦理价值而制定的具体规范标准,因此是伦理价值在实践过程中的具体呈现。伦理价值是伦理原则的来源,同时要从实践过程中的伦理风险来追溯伦理治理的需求,从侧面来支撑伦理原则。智能广告主要是由人工智能技术驱动发展的,因此需要在明晰智能广告中技术应用的观念化伦理价值、经验化伦理价值以及技术化伦理价值的基础上,结合技术伦理学的一般原则和人工智能技术的特殊性,以及实践过程中产生的伦理风险转化而来的伦理治理需求,建构其伦理原则,具体包括:尊重、安全、责任、公正、可约束性、可预防性、可协同性、可解释性、可控制性、可持续性(见图1)。

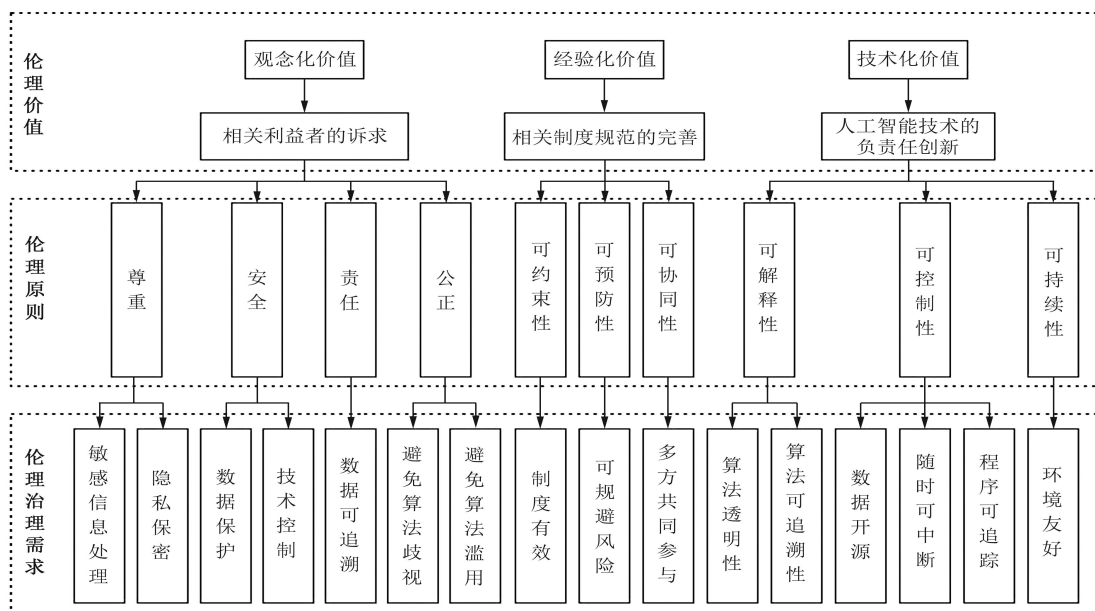


图1 智能广告中技术伦理风险治理的伦理原则<sup>[21]</sup>

从观念化伦理价值来看,实现智能广告“利益相关者的福祉”需建立尊重、安全、责任以及公正的伦理原则。在人工智能技术的应用过程中,消费者作为智能广告的目标受众,其隐私侵犯和信息安全遭到破坏,因此需要遵循隐私和安全的伦理原则对消费者的敏感信息采取保护性措施。比如,在广告智能洞察过程中,在建立消费者数据库时不可避免地要对消费者个人数据进行开发和收集,技术开发者和设计

者在法律允许的范围内合理利用消费者数据并保护消费者隐私,将有利于建立消费者与广告从业者之间的尊重和信任,促进广告行业的健康发展。作为人工智能技术的开发者,需要遵循公正的伦理原则,避免算法歧视和算法滥用所带来的伦理风险;同时,落实责任划分需要在据可追溯的基础上通过记录智能广告运作过程中的数据变化,来明确事故或过失的责任主体。更重要的是,尊重与公正的伦理原则可以有效避免由于算法特殊性带来的消费异化和消费者阶层分化等伦理风险。

从经验化伦理价值来看,相关制度规范是对智能广告中技术应用实践性价值的确认和保障,由此提炼出可约束性、可预防性和可协同性等原则。相关制度规范既要智能广告运作中的人工智能技术具有约束作用,以此规避可能引发的伦理风险,同时也要尽可能对这些潜在的伦理风险进行预测性防范,紧跟技术发展的速度甚至超越技术发展进行前瞻性的预防。值得注意的是,相关制度规范的制定不应局限于广告行业,而应该进行学界与业界、参与者和接受者等多元主体的有效整合,即遵循可协同性的伦理原则,从而尽可能从多维度对智能广告中的技术伦理风险进行规避和约束。

从技术化伦理价值来看,实现人工智能技术的“负责任创新”需建立可解释性、可控制性和可持续性伦理原则。其中,可解释性原则是治理算法黑箱带来的伦理风险的需要,即保持算法的公开透明及数据的可追溯,让消费者能够对智能广告中人工智能技术的具体运作有清晰的认知;可控制性原则是基于数据开源、随时可中断、程序可追踪的伦理治理需求,保证人类对技术的绝对控制,防止技术失控而引发一系列伦理风险;可持续性原则是基于环境友好的需求,在维护生态文明的前提下推进技术文明的发展。

#### 四、智能广告中技术伦理风险的治理策略

人工智能技术是智能广告发展的直接驱动力,其发展的不确定性也是导致诸多伦理风险的直接因素。这对以往的伦理范式形成新的挑战,一定程度上也阻碍了智能广告发展的进程,对此必须进行有效的伦理治理。“伦理对技术规约的过程性,主要体现在它的阶段性特征上,即伦理对技术的规约依据不同的技术发展阶段具有不同的规约手段和规约机制。”<sup>[26]</sup>因此,智能广告中技术伦理风险的治理需要针对人工智能技术应用的不同环节提出相应的治理策略。通过对技术人员的深度访谈了解到,人工智能技术在应用过程中的流程主要包括设计、测试、应用和复盘四个环节,据此本文在王钰、程海东研究<sup>[27]</sup>的基础上,结合智能广告每个环节的特性提出伦理嵌入、伦理评估、伦理调适和伦理确认四大伦理治理策略(见图2),并进行详细阐述,以期能够对智能广告中的技术伦理问题起到规范和约束作用。

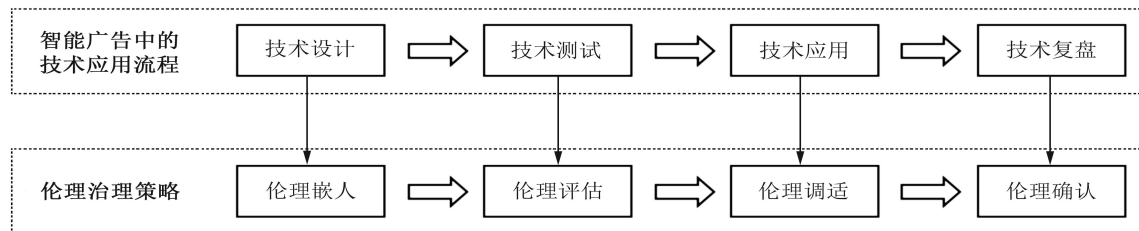


图2 智能广告中技术伦理风险的治理策略

##### (一) 智能广告技术设计阶段的伦理嵌入

在技术设计阶段进行伦理嵌入,需要技术的开发者和设计者根据一定的伦理价值和伦理原则对技术进行价值观念和道德观念的输入,使技术在具体运行过程中能够自动开展“道德判断”工作。也就是说,将具体的价值原则嵌入到人工智能技术的结构之中,使技术拥有一定的价值倾向。这其实就是费雷德曼提出的“价值敏感设计”,它“是一种以价值理论为基础的技术设计方法,它强调在整个设计过程中以价值为原则,并全面地解释人的价值在技术中的展现”<sup>[28]</sup>。智能广告中技术设计阶段的伦理嵌入包括

前文所述的观念化价值、经验化价值以及技术化价值。好的技术设计不仅能够为相关利益者的诉求服务,还能够促进相关制度规范的完善,并引导人工智能技术向好的方向发展,实现技术的升级和优化。在完成对技术目标价值的分析之后,需要和伦理学家一起跨领域探讨如何在技术设计的架构中贯彻和执行伦理原则,进而引起广告智能化的积极道德变化。

但是技术设计阶段的伦理嵌入是基于以往经验的预测和判断,往往在实际运行过程中并不一定会按照预期的设定充分发挥作用,无法保证伦理嵌入对伦理风险治理的效果,所以在具体伦理嵌入的操作过程中,技术开发者往往需要根据实际应用的场景和语境,不断对原始的设计嵌入方案进行优化完善。这就要求在伦理测试环节对伦理风险进行更加精准的评估。

## (二) 智能广告技术测试阶段的伦理评估

任何技术在正式投放应用之前都必须进行测试,智能广告中的技术应用也不例外。对智能广告技术的测试不仅仅是测试其功能效果,更是将伦理风险纳入评估的指标体系之中;而且技术的测试评估是一个持续性的过程,往往从技术开始测试之前评估工作就已经展开,随后在技术测试的过程中以及测试完成后都需要对技术进行实时监测和评估。

1. 技术测试前的伦理风险预测。一方面,伦理风险预测是建立在大数据基础之上的机器智能预测,人工智能技术能够通过算法学习和判断各种伦理风险,在形成自己的认知和理解后精准预测出每一种伦理风险产生的条件。这个过程就如同消费者智能洞察一样,人工智能技术通过特定的数据挖掘方式和分析逻辑将消费者的信息定制化地呈现出来,形成广告主对于消费者信息全面的感知与理解,从而实现为消费者个人量身定制的个性化广告内容生产和推荐投放。在技术测试前的风险预测环节,人工智能技术需要将伦理风险的相关数据同消费者数据一样进行挖掘和深度分析,从而建立基于各种伦理风险信息数据库,为科学的伦理风险预测提供基础。另一方面,技术测试前的伦理风险预测工作不仅需要技术开发者的参与,还需要专业的伦理学专家的指导和帮助。在技术测试前需要联合技术伦理委员会等专业人士进行人工判断和预测,减少机器判断的误差,以确保伦理风险预测的准确性。

2. 技术测试中的伦理风险识别。在技术测试过程中开展伦理风险识别是为了发现和判断可能妨碍技术应用的目标伦理风险,并对识别的伦理风险进行收集记录、分析研判和审核审查,为后续的工作提供合适的应对方案。智能广告技术测试中的伦理风险识别具体包括以下几个方面:首先,对技术测试中出现的潜在伦理风险进行全面识别和判断。其次,详细追溯和分析不同类型伦理风险背后所蕴含的伦理困境或价值冲突,发现不同伦理风险之间的共识性伦理原则,从而准确判断其形成的原因。比如,智能广告运作过程中由于人工智能技术的应用引出的消费者隐私保护和知识产权保护,设计者要坚持公正原则、责任原则等共识性的伦理原则。最后,通过利益相关者的诉求和反应对这些伦理价值和原则进行重要性排序和评估,不断对现有的人工智能技术应用的规范性共识框架进行更新,促进人工智能技术伦理原则框架的优化和升级,为之后的技术伦理风险识别提供极具操作性的参考。

3. 技术测试后的伦理风险处理。对技术测试过程中伦理风险进行伦理评估的目的并不一定是为了判断新技术的好坏与优劣,而是为了探讨如何使技术更加负责地为人类和社会造福<sup>[29]</sup>。也就是说,智能广告中人工智能技术的伦理评估并不是简单的判断和识别技术伦理问题,而是在识别风险的基础上能够对其进行有效的伦理治理。因此,在准确识别技术测试中的伦理风险之后,需要对其进行处理:一方面,可以根据前文建立的智能广告技术伦理原则的框架,结合人工智能技术应用的具体道德语境,通过解释和权衡来实现伦理原则对具体伦理风险的分析与澄清;另一方面,还可以根据人工智能技术应用的具体道德语境的特殊性对伦理原则框架进行检验、修正与调整。此外,还可对技术伦理风险进行分类后进行针对性地处理,比如对于能够通过算法避免的伦理风险可以直接删除,而对于比较复杂的、难以避免的潜在伦理风险进行实时追踪,密切关注其在技术应用过程中的演变,以便及时做出应对措施。同时,技术测试阶段的伦理评估是一个连续性过程,所以在出现争议性问题时无需一味追求共识,而是应



在整个人工智能技术测试和应用的环节进行持续地追踪和关注,防止因认识不足和信息缺失而造成片面性判断,从而对智能广告中的技术伦理风险形成全面性认识,不断提高伦理风险评估的准确性和有效性,最终为技术应用阶段的伦理调适创造更加有利的工作条件。

### (三) 智能广告技术应用阶段的伦理调适

人工智能技术经过伦理评估与广告运作系统进行融合,获得特定的“职业角色”,在智能广告运作过程中不断完成伦理价值和伦理原则的嵌入。就智能广告技术应用的实际情况而言,要真正实现有效嵌入,就必须充分考虑智能广告运作过程中各种社会因素的影响,比如政府监管、行业约束等。也就是说,在智能广告技术应用过程中需要根据广告行业发展的社会要求不断调整相应的伦理规范,从而更好地和整个社会的伦理价值相适应。王钰、程海东曾提出伦理调适的三种路径:制度调适、舆论调适和教育调适<sup>[27]</sup>,限于篇幅此处仅探讨智能广告技术应用阶段的制度调适和舆论调适。

1. 技术伦理体系的制度调适。制度调适是最具有约束力和效力的伦理调适方式。面对当前发展迅速的新技术,传统的政策制度已无法对人工智能技术的具体应用进行有效规制和约束,同时对于其在应用过程中产生的伦理风险也无法做出准确预测和有效应对,因此迫切需要更新和完善相关政策法规。2021年,《中华人民共和国个人信息保护法》《中华人民共和国数据安全法》的颁布实施在社会上产生广泛而积极的影响。但是由于部分企业法律意识淡薄,在经济利益驱动下心存侥幸、不严格遵守法律,甚至触犯道德底线。这是制度调适面临的难点问题。当前制度调适需从以下三方面发力:首先,持续推进智能广告的法制化和标准化,需进一步通过立法明确人工智能技术发展的规范,并针对广告行业出现的新问题修订《广告法》,增加有关智能广告的法律规定;还需制定保护消费者隐私及信息安全等具体技术标准,在政府主导下由标准化研究机构组织专家研究制定。其次,提升对智能广告监管的数据化和智能化水平,用大数据、人工智能技术建立监管平台,做到全方位、全过程的监管。最后,增强行业规范的约束力,各地广告协会或其他相关协会除了制定相关伦理规范之外,还需把企业守法合规表现、遵守伦理道德情况与企业的等级评定、奖项评选乃至会员资格直接挂钩,以避免让行业规范成为一纸空文。

2. 技术价值考量的舆论调适。在人工智能技术伦理嵌入的过程中,良好的舆论场能够对技术主体起到正确引导作用,帮助他们在技术开发和应用过程中做出符合技术伦理原则和价值的决策,从而在一定程度上对人工智能技术伦理风险进行规避和制约。因此,舆论调适作为人工智能技术伦理调适的辅助支撑,在具体实践中往往也发挥着不可忽视的重要作用。在智能广告技术应用舆论调适的过程中,舆论监督是覆盖范围广、影响企业声誉的调适方式。广告传播也是一种社会信息传播的方式,信息传播的内容和形式都需要接受社会舆论的监督,尤其是公众监督和媒体监督。在智能广告技术应用的伦理调适环节,公众监督主要表现为广大消费者的监督,也就是消费者在面对智能广告中的技术伦理问题时形成一定的群体共识,这种群体共识是社会公众的多数意见之集合,也是社会伦理评价的集中体现。而媒体监督往往通过借助传播工具来实现监督的目的。要充分发挥舆论监督的调适作用,就要从以下三个方面入手:第一,政府统筹并引导媒体监督相关企业单位在信息传播的过程中要做到公开透明,杜绝数据造假,消除“算法黑箱”,促进公众对大数据和算法的认知和了解,同时对正确的伦理价值进行有效的宣传和推广,有助于消费者形成并加强对智能广告中技术价值考量及相关伦理规范的自觉监督意识,并为智能广告技术应用中的考核及责任追究提供支撑依据。第二,拓宽舆论监督和反馈的渠道,充分发挥新媒体平台的优势,建立消费者投诉和治理反馈的有效机制,即“一键投诉、有错必纠”机制,该平台需与前文所述的监管平台合而为一,才能发挥有效治理作用。第三,发布权威性新闻报道和研究报告,对企业的违法失德行为进行曝光或分析,给相关企业造成舆论威慑力,例如中央电视台每年的“3·15晚会”对企业侵犯消费者权益的行为进行深度报道,这种做法可借鉴到智能广告伦理风险的治理中。研究机构可定期发布智能广告企业声誉报告(或法律伦理舆情报告),还可在企业社会责任报告中增加企业伦理的指标和内容。总之,最终要形成具有良好伦理导向的道德舆论场,有效防治技术伦理风险。

#### (四) 智能广告技术复盘阶段的伦理确认

在智能广告运作过程中,人工智能技术的每一次应用都需要进行全面深刻的复盘,通过对技术设计、技术测试及技术应用等不同阶段中的技术伦理风险进行回顾总结,避免同样的风险和问题再次发生,不断完善和优化人工智能技术伦理价值框架,形成统一的技术伦理认知思维。智能广告中技术复盘阶段的伦理确认不仅是对技术不同发展阶段伦理问题的总结,更是对每一阶段工作内容的检验和确认,因此需要形成完整的伦理确认路径,才能够保证结论的科学性和全面性。

一方面,从智能广告运作的不同环节来看,完整的伦理确认路径是从伦理道德的角度审视智能广告运作的流程,因此这种伦理审视要贯穿智能广告运作的全过程:首先,在消费者智能洞察的过程中,是否符合社会的整体价值观念,对消费者的数据是否履行了道德上的保护义务,在对数据的应用过程中又是否能够在满足个体需要的基础上实现价值的创造。其次,在广告内容智能化生产过程中,是否使用的所有广告创意素材都获得了原作者的同意和授权。再次,在智能化广告推送过程中,智能技术下的算法是否存在歧视问题,在传播途径、传播符号、传播检测等方面是否能够建立伦理监督机制,形成问题导向化的预警解决方案。最后,在广告效果智能监测环节,当出现技术伦理问题时,消费者的反馈渠道是否完善,技术开发者是否能及时进行调整等,这些都需要进一步进行检验和确认。

另一方面,从智能广告中技术应用的不同阶段来看,技术应用复盘阶段的伦理确认路径主要包括对技术设计阶段伦理嵌入的合理性和规范性的再验证、对技术测试阶段伦理评估的再核对,以及对技术应用阶段伦理调适的再考察。通过对技术应用不同阶段的伦理嵌入、伦理评估及伦理调适的完整复盘和检验,形成智能广告技术应用过程中不同利益群体的不同伦理价值的伦理共识框架,并以此为基础为下一次的技术应用提供伦理导向和价值考量。如此循环往复,不断优化和完善智能广告中技术伦理风险的治理原则和治理策略。

综上所述,在智能广告迅速发展的今天,我们必须充分认识到其中已经存在或潜在的技术伦理风险,从观念化价值、经验化价值以及技术化价值层面建构智能广告中技术伦理风险治理的伦理原则,从智能广告的设计、测试、应用、复盘四个环节进行伦理治理,包括伦理嵌入、伦理评估、伦理调适和伦理确认,从而实现智能广告的健康有序发展,促进人工智能技术更好地为人类服务。

#### 参考文献

- [1] W. Rodgers, T. Nguyen. Advertising Benefits from Ethical Artificial intelligence Algorithmic Purchase Decision Pathways. *Journal of Business Ethics*. 2022, 178, (2).
- [2] B. Vlacic, L. Corbo, CES Susana, M. Dabi. The Evolving Role of Artificial Intelligence in Marketing: A Review and Research Agenda. *Journal of Business Research*, 2021, (1).
- [3] 蔡立媛,周慧. 人工智能广告的“时空侵犯”伦理危机. 青年记者, 2019, (15).
- [4] 姜智彬,马欣. 领域、困境与对策:人工智能重构下的广告运作. 新闻与传播评论, 2019, (3).
- [5] 李名亮. 智能广告信息伦理风险与核心议题研究. 新闻与传播评论, 2020, (1).
- [6] 邵敏,赵韵文,林雪冬. 试论智能广告的形式、特点及监管. 湖南大众传媒职业技术学院学报, 2017, (5).
- [7] 阿明·格伦瓦尔德. 技术伦理学手册. 吴宁译. 北京: 社会科学文献出版社, 2017.
- [8] 陈爱华. 高技术的伦理风险及其应对. 伦理学研究, 2006, (4).
- [9] 戴维·赫佐格. 数据素养: 数据使用者指南. 沈浩等译. 北京: 中国人民大学出版社, 2018.
- [10] 佩德罗·多明戈斯. 终极算法: 机器学习和人工智能如何重塑世界. 黄芳萍译. 北京: 中信出版集团, 2017.
- [11] J. L. Garcia, H. M. Jerónimo. Fukushima: ATsunami of Technological Order//*Jacques Ellul and the Technological Society in the 21st Century*. Springer Netherlands, 2013.
- [12] 张淑玲. 破解黑箱: 智媒时代的算法权力规制与透明实现机制. 中国出版, 2018, (7).
- [13] 许可. 人工智能的算法黑箱与数据正义. 社会科学报, 2018-03-29

- [14] 刘培,池忠军.算法歧视的伦理反思.自然辩证法通讯,2019,(10).
- [15] 马克思.1844年经济学哲学手稿.北京:人民出版社,2014.
- [16] 鲍德里亚.消费社会.刘成富、全志刚译.南京:南京大学出版社,2001.
- [17] 陈力丹.准确使用“异化”概念.当代传播,2003,(5).
- [18] 杨先顺.技术异化中的人性残缺——对当前网络广告的追问与反思.现代传播,2005(2).
- [19] I. Van de Poel, L. Royakkers. *Ethics, Technology and Engineering*. Oxford: Wiley Blackwell, 2011.
- [20] B. Friedman, P. H. Kahn, A. Borning, A. Hultgren. Value Sensitive Design and Information Systems//N. Doorn, D. Schuurbers, I. van de Poel, etc. *Early Engagement and New Technologies: Opening up the Laboratory*, vol16, Springer series: *Philosophy of Engineering and Technology*. Dordrecht: Springer, 2013.
- [21] 于雪,段伟文.人工智能的伦理建构.理论探索,2019,(6).
- [22] 费艳颖,刘彩薇.负责任创新视角下我国人工智能产业政策的解构与重构.情报杂志,2021,(7).
- [23] 理查德·欧文等.负责任创新的理论框架.李福嘉译//王前,菲利普·布瑞.负责任创新的理论与实践.北京:科学出版社,2019.
- [24] 杨庆峰.从人工智能难题反思AI伦理原则.哲学分析,2020,(2).
- [25] 孙伟平.论人工智能发展的伦理原则.哲学分析,2022,(1).
- [26] 王健.技术伦理规约的过程性.东北大学学报(社会科学版),2003,(4).
- [27] 王钰发,程海东.人工智能技术伦理治理内在路径解析.自然辩证法通讯,2019,(8).
- [28] 张贵红.价值敏感设计与大数据伦理.伦理学研究,2019,(6).
- [29] P. Brey. Philosophy of Technology After the Empirical Turn. *Techne: Research in Philosophy & Technology*, 2010, 14(1).

## A New Exploration on the Risk Management Of Technology Ethics in Intelligent Advertising

Yang Xianshun (Jinan University)

Li Ting (Shenzhen Qicai Innovation E-commerce Company Ltd)

**Abstract** With the wide application of artificial intelligence in the field of advertising, the problem of ethical risk caused by the technology in intelligent advertising has gradually become the focus. From the perspective of technology ethics, there are mainly two types of ethical risks in intelligent advertising: one is based on multi-end data mining, which will lead to consumer privacy infringement, information value infringement and creative intellectual property infringement; the other is based on deep learning algorithms which will result in barriers of advertising operation, consumer class differentiation and consumption alienation. With regard to the governance of ethical risks in intelligent advertising, principles should be first established from the three aspects of conceptual value, empirical value and technical value, and then subsequent strategies should be developed in terms of the four links of design, testing, application, and review in intelligent advertising, namely ethical embedding, ethical assessment, ethical adjustment, and ethical confirmation, so as to avoid and restrain the technology ethical risks in intelligent advertising.

**Key words** intelligent advertising; technology ethics; ethical risk; ethical value; ethical governance

■ 收稿日期 2022-05-17

■ 作者简介 杨先顺,哲学博士,暨南大学新闻与传播学院教授、博士生导师,暨南大学传播与国家治理研究院院长;广东广州 510006;

李婷,深圳市七彩创新电子商务有限公司, TikTok 海外直播项目合伙人;广东深圳 518100。

■ 责任编辑 杨敏