

论数字伦理体系的建构

江必新, 刘倬全

(湖南大学法学院, 湖南长沙, 410082)

摘要: 构建数字伦理体系, 有利于回应数字技术和人工智能“正反”二重性, 克服法律规范和传统治理方式的局限性, 防范数字网络安全和伦理技术风险, 保证数字社会良性运转, 应对数字技术对传统伦理体系造成的严峻挑战。在建构数字伦理体系时, 应处理好规范和发展等五大关系, 保障各相关利益主体的实质参与, 建立贯穿全链条的各类制度体系, 强化所有相关要素的标准化设计, 进行数字技术的全周期伦理考量, 遵循数字伦理框架性原则。要以元规则为出发点, 以求善、求真、求美为指引, 构建抽象化与具体化相结合, 适应不同时期、不同主题、不同环境、不同领域的数字伦理规则。

关键词: 数字社会; 数字伦理; 数字技术风险; 数字秩序; 数字价值

中图分类号:D90-053

文献标识码:A

文章编号:1672-3104(2024)01-0038-12

数字伦理是伴随着社会转型而兴起的新话题, 它既是一个理论问题, 更是一个实践问题。作为一个全新的命题, 数字伦理涉及许多全新的领域, 例如大数据杀熟、自动决策、数字透明、算法偏见等。无论是在个人层面还是在社会层面, 对于数字伦理的讨论比以往都更为现实而激烈。我国在当下的数字化转型过程中并未形成与之相匹配的数字伦理规则, 如何建立沟通现实世界与数字世界的伦理规则, 通过伦理法律在数字世界和现实世界之间架起沟通对话的桥梁, 融传统伦理于数字伦理之中, 是当下需要予以回应的重要问题。

一、数字伦理体系建构的意义

建构数字伦理体系的意义, 体现在以下几个方面:

(一) 回应数字技术和人工智能的“正反”二重性

数字伦理体系的建构, 是回应解决数字技术

和人工智能“正反”二重性冲突的现实选择。数字技术和人工智能的发展, 不仅可以帮助个体提高生产力和生产效率、优化协作和沟通、即时获取所需的信息, 还可以在商业发展上帮助企业简化流程, 节省各类资源, 提供更优发展路径。

随着技术的快速发展与应用, 与数字技术和人工智能的使用相关的潜在风险也与日俱增。

首先, 数字社会面临信息外泄的风险。随着人类文明对数字技术的依赖逐步加深, 未经授权的个体获取隐私数据的风险也在增加。近年来, 数据泄露事件的规模、成本以及频率都在快速增长。根据 Ponemon 发布的《2022 年数据泄露成本报告》, 2022 年全球信息安全事件的总量及平均经济损失均达历史最高, 每起信息安全侵害事件的经济损失为 435 万美元。由于恶意软件、黑客技术以及其他恶意活动的出现, 存储在数字设备上的任何数据都有可能受到攻击。

其次, 数字社会存在身份盗窃的可能性。身份盗窃被定义为使用他人的个人身份信息牟利或犯罪^[1]。当有人未经许可使用他人的个人信息

收稿日期: 2023-07-30; 修回日期: 2023-12-13

基金项目: 国家社科基金重大项目“建设中国特色社会主义法治体系研究”(232DA072)

作者简介: 江必新, 男, 湖北枝江人, 法学博士, 湖南大学法学院教授、博士生导师, 主要研究方向: 行政法、智慧法治; 刘倬全, 男, 湖南长沙人, 湖南大学法学院博士研究生, 主要研究方向: 智慧法治, 联系邮箱: 415499233@qq.com

(如身份证号码或银行账户信息)参与社会活动时, 就会发生身份盗窃。在数字时代, 犯罪分子比以往任何时候都更容易利用个人信息实施诈骗和其他犯罪。这不仅是对个人隐私的大规模侵犯, 更对组织和企业的运营发展构成了严重威胁。

再次, 数字社会的网络诈骗风险急剧攀升。在万物互联、大数据驱动一切的背景下, 越来越多的社会活动转移到线上。互联网技术与应用持续升级, 犯罪分子更容易将目标对准对数字社会不甚熟悉的“数字移民”和“数字难民”。网络诈骗与传统诈骗有一定相似之处, 但是其作案手法更加隐蔽, 诈骗的方式、手段也不断升级。当下的电信网络诈骗已形成了分工精细、体量庞大的黑色产业链, 打击难度愈来愈高, 成为令公众深恶痛绝的社会“毒瘤”和影响社会秩序的重要问题。

最后, 数字社会还伴生着数字成瘾的社会风险。随着社交媒体和其他数字活动的兴起, 新的休闲娱乐方式和信息获取渠道已为大众所接受。借助大数据与新型算法, 人们生活的资讯、社交、消费、休闲娱乐等各个渠道无一不被互联网所占领。数字化成瘾似乎已成为一股无法抗拒的裹挟力量。当人们过度依赖数字技术时, 数字成瘾弊端也随之显现, 例如出现认知障碍、人际关系淡薄、个人专注度下降甚至人身安全受到威胁等情况。如何发挥数字技术的积极作用, 抑制数字技术的负面影响, 妥善解决数字成瘾的问题, 将是未来数字社会面临的重要挑战。

(二) 消弭数字鸿沟和满足数字网络安全的需要

数字伦理体系的建构, 有助于抑制数字网络的无边界性、跨国性、专业性、非透明性。数字网络彻底改变了人们沟通、协作和开展业务的方式。数字网络具有无边界性、跨国性特征, 它模糊了国界限制, 促进了全球个人、组织和机构之间的合作与协调, 使其能够共享知识和资源。此外, 它的无边界、跨国性质还使商业组织能够接触到不同国家和地区的客户、合作伙伴和员工, 并快速、轻松地扩大规模并进入其他国家的市场。这有助于企业扩大影响力, 并从世界各地获

取宝贵的数字资源和专业知识。尽管数字智能技术的广泛应用全面提升了整个社会的生活质量, 但不同群体对数字红利的分享却存在显著差异^[2]。特别是对数字经济并不发达的国家来说, 意味着数字鸿沟的进一步加剧。与工业时代不同, 数字时代的数据资源成了超越其他传统自然资源的最宝贵资源。数据的开发和利用, 取决于人们运用数字技术的水平。世界银行最近对非洲、亚洲和拉丁美洲的二十多个中低收入国家的一项调查显示, 这些地区近70%的人不了解互联网是什么或如何使用互联网。这种数字技能差距使技术水平低、互联网普及程度差的中小型国家处于极其不利的地位。数据资源差异的扩大预示着综合国力之间的显著分化, 这进一步推动了不同国家和地区间数字不平等的现状滑向更深层次的失衡。这种差异加剧了社会不平等, 限制了全球经济和社会的可持续发展。解决这一问题不仅是技术挑战, 更是伦理挑战。为了消除数字鸿沟, 保障数字智能技术的可持续和负责任的发展, 需要构建数字伦理体系, 促进数字技术的普惠和包容, 实现数字红利的公平和合理分享。

数字网络具有不透明性, 还可能诱使技术型企业在商业活动中对客户和合作伙伴实施黑箱行为。数字网络是基于专业技术搭建的新型生态网络, 是新一轮工业革命与产业迭代融合的产物。与传统网络相比, 其不透明性更为彰显。譬如, 在数字金融领域, 数字网络的不透明特性将可能对金融市场投资者的信心和参与度造成冲击性影响。尽管数字网络在运作过程中存在诸多缺陷, 但数字网络仍是当下社会运转所必备的重要工具。因此, 筑牢数字网络安全防线, 发挥数字伦理的引导作用是数字社会良性发展的前提条件。

(三) 克服法律规范和传统治理方式的局限性

数字伦理体系的建构有助于克服法律规范和传统治理方式的局限性。这种局限性主要体现在: 无论是在物理世界还是数字世界, 法律规范对于善治都至关重要, 它为决策提供了制度框架, 有助于确保公平正义以可见的方式实现, 从而保护公民的合法权益。然而, 由于数字技术自

身的特点，法律规范在数字领域并不能发挥其在物理世界同等的作用。首先，数字技术变化更新的频率较高。智能生物识别、自动驾驶等人工智能新技术日新月异，不断改变甚至颠覆了人们的生产生活方式。由于法律法规必须确保一定的稳定性，若在制定修改相关法律时没有为新技术的发展留下一定的制度空间，就会导致法律相对滞后，并引发人们对信息安全和隐私保护的持续担忧。其次，数字网络的无边界性可能使执法活动变得困难。领土是传统国际法上主权国家行使管辖权的基础。国家主权原则要求在他国领土上行使管辖权时须获得该国同意。但事实上，数字技术使领土界限变得模糊，各国在网络空间和数据领域的管辖冲突在所难免^[3]。例如，数据隐私规范通常基于物理边界的假设而创设，但在数字世界中，数据可以在多个司法管辖区收集和共享，这使得传统属地管辖失效。因此，传统法律规范并不能当然适用于数字社会，而去地域化的管辖方式或将成为未来的趋势。最后，由于数字世界中的参与者往往因匿名而难以被追踪识别，在违法行为出现后，司法机关难以快速精准地定位到具体行为人，这可能使违法者逃避承担法律责任。

传统的行政管理模式也难以应对数字技术带来的伦理挑战。几个世纪以来，传统管制性治理理论和模式在社会治理方面发挥了巨大作用，但这些模式并不能当然适用于数字世界。首先，传统治理模式往往依赖政府在社会治理中发挥决定性作用。行政机关与个人、企业之间采用传统单向的意见交流方式，而数字权力的出现重新划分了公共权力运作的边界和模式。数字权力是权利关系的一种，以治理为主要目标。正如以色列著名学者尤瓦尔·赫拉利教授所言：“未来的时代将会进入算法主导的时代，权威将从个人转向由算法构成的网络。”^[4]其次，传统治理模式下的政府可能无法单独防范网络安全或数据隐私等新型科技风险。与传统治理模式不同的是，大型数字企业在数字社会中处于主导地位。例如，以阿里巴巴为代表的平台企业不仅提供交易平台服务，而且也是买卖双方在线争议解决的裁判者。具言之，阿里巴巴不仅可以被动地根据双方

提供的证据和事实确定违约方和相关权利义务，还可以根据大数据的预警主动发起相关调查，以维护平台的公正性。相较而言，行政机构的信息渠道和数字技术储备均显不足，无法高效独立地解决数字社会中产生的类似问题。最后，在数字世界中，数据可以在多个司法管辖区收集和共享。这也使得依赖物理边界的传统治理模式在数字社会难以发挥作用。

(四) 保证数字社会的良性运转

道德是一套指导人类行为和决策的原则和价值观体系，对人类社会的发展至关重要。它便利人类区分善恶是非，并指引人们以对社会负责的方式实施自身的行为。数字道德对数字个体的行为指引、数字商业盈利的可持续性、数字技术的良性开发利用以及数字社会成员取得互信，均具有十分重要的意义。进入风险社会阶段后，当法律等社会规则因社会的高度复杂性和不确定性而功能式微时，道德的规范作用就显得更加重要了^[5]。

首先，数字道德可以为数字个体提供数字伦理层面的行为框架。通过遵循这一行为框架，数字个体可以满足其在自身和社会层面的双重价值需求，包括但不限于自尊、自我满足、社会认同以及社会影响力等。当数字个体处于扭曲的社会关系当中，数字道德的存在及其价值功能的实现就显得尤为重要。数字伦理的规范化依赖于个体道德修养的增强。这是因为道德具有个体性质，侧重于个体的认知与决策。在整个社会结构当中，道德水平具有显著的分布差异。相对地，伦理具有社会性质，着重于集体的共识与准则。伦理标准实际上是一种非差异化的“社会契约”。在一个社会中，个体道德水准越高，伦理规则的文明程度越高。而如果道德水准差异化程度越大，形成高水平的伦理规则就越难。因此，持续提高个体道德修养应被视为构建数字技术伦理秩序的重要前提与基石^[6]。

其次，数字道德对数字商业行为的规范化、盈利的可持续性，具有十分重要的意义。因为它有助于确保决策符合所有利益相关者的最佳利益。合乎道德的企业更有能力与客户建立信任，吸引和留住顶尖人才，并在市场上创造良好的声

誉。由于数字经济本身具有虚拟性、高渗透性、边际成本递减和外部性等特点,商业组织在掌握先进数字化工具和自主知识产权的同时,更需以高尚的商业道德准则作为发展基础。

再次,数字道德对于数字技术开发利用的良性发展更是不可或缺。责任伦理学主张,伦理对技术的规约应是前置性的,汉斯·乔纳斯指出:

“新的类型和方面需要一种相应的预见和责任伦理学,它像必然遇到的突然事件那样惊奇。这种新的责任命令要求一种新的谦逊。它不像以前是由于我们能力的弱小而导致的谦逊,而是由于我们能力的过分强大。这种强大表示我们的活动能力超越了我们的预见能力、评价和判断力。”^[7]在开发新产品、新服务以及进行学术研究时,我们必须考虑伦理因素。这有助于确保科学和技术进步得到负责任的利用,进而保障人类与数字技术和谐共生。

最后,数字道德是数字社会成员之间取得互信的前提和基础。道德本身作为一个社会性概念,反映了人类基本的社会依存关系。共同体是人生存和发展的基本方式,本质上是一种伦理性存在,体现着伦理确定性、规范性和秩序性,既是一种组织性和约定性集合,又是一种精神与情感上的“默认一致”和“相亲相依”^[8]。数字社会更是如此,从数字生产者、开发者到使用者,每一个社会成员都不可避免地成为数字伦理共同体的一员。人类要实现整体的永续发展,必须遵从在普遍共识基础上采取共同行动这一集体性伦理准则^[9]。所有的数字成员都要遵循一定的数字伦理。它构成了数字社会成员之间信任和尊重的基础,帮助培育具有人性和社会道德责任的社会成员,使人们能够有效地开展协作活动,带动数字社会大步向前。

(五) 防范网络安全和技术伦理风险

数字伦理体系的建构有助于防范网络安全和技术伦理风险。网络安全是数字经济安全的重要内容,旨在保护计算机网络免受未经授权的访问、恶意活动和其他类型的威胁。虽然网络安全本质上主要是技术性的,但同样也取决于人类道德。在数字社会,人性的弱点可能会成为网络安全系统中最薄弱的环节。与严格执行代码的数字

设备不同,数字研发和运维者会出现过失甚至故意破坏网络安全的情况。例如,用户可能因遭受网络诈骗而错误传输信息,或故意与未经授权的第三方共享机密信息。此外,网络安全还取决于负责网络安全研发运营人员的道德伦理信念。网络安全管理员必须有意识地识别潜在风险,并采取适当措施保护系统。例如,定期更新安全协议、实施访问控制策略和监控用户活动等。因此,网络和技术伦理风险的预防,需要从网络使用者和技术研发运营人员入手,加强网络道德教育和科技伦理教育,使他们遵从道德和伦理规范,以道德伦理指导技术的运用和研发,最大限度地减少人工智能的伦理风险^[10]。此外,网络安全也只有在数字技术的运用符合数字伦理原则的情况下,才能得到有效保障。例如,它应该用来保护用户的隐私,防止数据被滥用;应以尊重个人权利、避免歧视和维护法治的方式使用,有效防止网络安全成为削弱个人权利的借口和工具等。

数字行为主体若缺失完整健康的数字伦理观,极易引发各种网络失范行为。网络的特征是数字技术发展的产物。互联网具有虚拟性、全球性、无边界性、共享性、隐蔽性等特征,为数字世界的发展提供了强大动力,但也引发了一系列道德失范行为。在这种环境下,部分数字主体因未形成完整健康的数字伦理观,缺乏道德自律,导致道德失范现象激增。例如,通过电信网络诈骗和网络赌博破坏正常的网络市场经济秩序,通过制造和传播网络谣言造成社会秩序不稳定,甚至通过网络教唆暴力恐怖活动、宣传民族分裂思想等。这些行为对国家安全和社会秩序都产生了负面影响。

(六) 应对数字技术对传统伦理体系造成的严峻挑战

数字伦理体系的构建是有效纾解数字技术对传统伦理体系造成严峻挑战的应然之举。随着数字技术深度介入各项社会活动,许多传统的社会概念不得不因此延展甚至颠覆其原有内涵。例如“隐私”转变为“数字隐私”,“好友”扩展到“微信好友”,“社群”扩展到“网络社群”,“裁判”扩展到“在线纠纷调解”等。数字技术开始作为新的生产要素,成为人类社会网络化、

信息化和智能化的根本驱动力。这为人类与社会的发展提供了前所未有的机遇，也重塑了人们现实的生产和生活方式。

然而，工业时代所形成的伦理准则无法规制新时代下各类数字化活动。数字技术正在挑战人的主体性地位和社会原有的技术伦理价值体系。一方面，数字技术侵蚀了个人的主体性。在大量的数字场景当中，人类将选择权让渡给了算法，算法的个性化推荐、营销、推送等功能能够很好地迎合个体需求^[11]。人们被动接受算法所构建的虚拟世界，这大大损害了个体的自主性和理性思考能力。此外，算法本身也蕴含着一定的价值取向，这使得算法可以潜在地操纵个体的意识并指导其行动。“空间对个人具有强大的管理和统治能力。”^[12]数字技术的发展，使我们生活在一个全景式的如同边沁所设计的环形监狱一般的数字空间。在这样的数字围城下，现实的人逐渐成为“数字世界的奴隶”。就此而言，数字技术可能会扭曲社会公平正义的具体实现。具言之，数字技术通过智能终端可能限制或妨碍人的全面发展，导致人在某种程度上被异化为从事数字化劳动、产出剩余价值的实体。以数字平台的工作者为例，他们似乎在从事自主的数字劳动，但实则数字资本高度支配了其劳动的全过程。尽管他们在表面上被视为数字经济的塑造者，实际上，他们被困于追求剩余价值的数字化符号的风险环境中。另一方面，广大数字用户为数字平台贡献了大量自由劳动时间，加剧了技术异化的程度。数字用户在网络空间进行聊天、购物、视频、游戏时，表面看来是用户自发实施的满足自身娱乐、消费、交流等需求的活动。然而，数字用户通过实施这一系列自发的行为，可以生产出数据产品。数字企业可以将其收集处理后形成数据商品，并卖给有需求的其他广告企业，使其对不同用户实施广告的精准投放。

二、数字伦理体系建构的基本准则

建构数字伦理体系，应遵循如下基本准则：参与、建立贯穿全链条的各类制度体系，对所有

相关要素强化标准化设计，对数字技术的全周期展开伦理考量以及遵循数字伦理框架性原则。

(一) 应处理好五大类关系

其一，要处理好规范和发展的关系。规范是数字技术发展的前提，良性发展必须以规范为基础。数字社会是风险社会，如果没有一定的规范，数字社会的发展就成为无源之水、无本之木。但规范也应该具有一定的灵活性，并定期进行审查和更新，以适应技术和社会的变化。

其二，要处理好问题导向和系统思维的关系。一方面，要强化问题导向，针对数字社会中出现的问题积极回应，补齐短板弱项。另一方面，要强化系统思维，统筹考虑数字技术所带来的伦理问题，确保数字伦理框架协调互促、形成合力。例如，在个人数据保护方面，不仅要运用数字伦理对数据泄露问题予以规制，也要考虑到技术发展和商业需要的平衡，以避免伦理规范过度限制创新与发展。

其三，要做到统筹运作与分类分级分阶段相结合。在构建数字伦理的基本原则和价值观时，要充分考虑到数字社会的多样性、复杂性、动态性和不确定性，完善数字伦理的整体框架。针对不同行业、不同场景、不同阶段因地制宜，根据行业和场景的特定需求细化具体规定，充分发挥数字伦理的指导作用。

其四，要处理好宣传引领与机制约束的关系。在面对数字技术这一新鲜事物时，宣传引领工作在统一思想、形成共识方面至关重要。开展好数字伦理的宣传教育，提升数字伦理的传播力、引导力、影响力，才能凝聚起强大的精神力量，使数字伦理真正发挥作用。然而，仅仅依靠宣传引领仍是不够的，还要将数字伦理有机融入具有国家强制力的约束机制之中，才能真正实现数字伦理所欲达成的社会目标。

其五，要处理好维护国家主权安全和发展利益以及构建人类数字命运共同体的关系。数字经济已成为全球新一轮科技革命与产业转型的核心动力。我国当前正处于数字化转型与智能化进阶的关键时期，数字安全的根本性影响日渐凸显，而数字技术的不平衡性在一定程度上有助于发达国家利用先进数字技术形成隐形的数字霸

权。这就要求我国必须积极参与数字技术的变革创新, 以捍卫国家的数字安全。人类命运共同体理念是习近平总书记站在全人类道义立场上, 回应全球百年大变局的时代诉求而提出的中国方案。人类命运共同体的基本目的在于保证人类系统的良性运转, 体现了对维系人类历史发展伦理义务的担当^[13]。数字命运共同体理念拓展了人类命运共同体的路径, 不仅有助于巩固世界的持久和平, 而且有助于实现数字技术成果的普惠性和共享性, 让数字主体都能享受到数字技术发展带来的红利。平衡好国家安全和人类命运共同体之间的关系, 要求我们在扩大国际合作的同时, 对数字技术设置负面清单, 对有可能影响到国家安全领域的数字技术予以严格保护。

(二) 保障各相关利益主体的实质参与

数字伦理体系的制定者在构建数字伦理体系的过程中, 应保障各相关利益主体实质参与进来, 表达自己的利益诉求, 并在相互协调中形成符合各方利益需求的数字伦理体系规则。这些利益主体包括但不限于设计者、生产者、应用者、消费者、管理者和监管者等。鉴于相关利益群体众多, 仅依靠政府部门难以完成构建数字伦理体系这一时代任务。“发挥群团组织、社会组织作用, 发挥行业协会商会自律功能, 实现政府治理和社会调节、居民自治良性互动, 夯实基层社会治理基础。”^[14]这是党中央对国家治理能力体系建设作出的重要安排。在这一背景下, 政府角色将发生进一步转变。数字伦理是未来数字社会的指导性纲领, 与相关利益主体之间具有密不可分的关联性, 应当鼓励不同利益主体参与数字伦理的制定。在各方主体追求最大化利益的博弈过程中, 治理主体可以将公共利益与私人利益两相权衡, 以提升伦理框架的科学性和可接受性, 形成各利益相关方可接纳的数字伦理。多元主体参与形成的《人工智能北京共识》, 就是当下社会共治的有益探索。在数字伦理初步形成之后, 行业自律和公众宣传是保证数字伦理发挥作用的关键步骤。马克思曾经指出: “科学绝不是一种自私自利的享乐。有幸能够致力于科学的研究的人, 首先应该拿自己的学识为人类服务。”^[15]这就要求数字工作者将提高人类的福祉作为一切工作

的出发点和落脚点, 加强行业自律, 对数字科技主动承担伦理责任。例如, 可以建立行业伦理投诉和伦理纠纷解决机制, 充分发挥数字伦理的作用, 推动数字伦理的进一步实施。数字伦理的最终形成还离不开社会公众的支持和认同。国家可通过各种渠道增强人民群众对数字技术、数字社会等概念的认知, 在加强数字伦理建设时, 同步更新公众的伦理观念, 营造良好的社会舆论氛围。

(三) 建立贯穿全链条的各类制度体系

数字伦理体系的制定者在构建数字伦理体系时, 需建立贯穿全链条的各类制度体系。所谓贯彻全链条的各类制度体系, 是指要建立标准体系、伦理审查体系、监测体系、问责体系、保障体系等。

在这些体系当中, 最基础的制度是标准体系。标准体系是衡量所有体系的基准, 涵盖了在研究与应用人工智能过程中应当遵守的基础伦理准则。在扬弃传统伦理框架的前提下, 我们需要尝试形成一套较为系统的数字伦理规则。数字伦理作为数字时代的行为准则, 发挥着方向型的基础作用。这些准则包括有利人类生存与发展原则, 和平与公共安全原则, 信息内容客观、准确、全面原则等。

审查体系用于逐案评估人工智能系统的道德性。它依据标准体系建立, 以确定新开发的人工智能系统是否符合伦理原则。人工智能系统的伦理审查体系应当分为技术伦理审查、责任伦理审查和利益伦理审查等几部分。鉴于人工智能系统的技术特征, 技术伦理审查将在综合评估体系中扮演主要角色。技术伦理审查是跨领域、系统性的审查活动, 旨在审查与技术有关的伦理难题。技术伦理审查的目标在于明确对数字技术及其应用后果的认知, 并在道德维度上为与数字技术相关的活动提供有效评价和指引。

监测体系用于持续跟踪人工智能的使用情况, 并确保其符合道德规范。它监测人工智能的开发和使用, 以确保人工智能系统按照标准体系的要求开展工作。具言之, 为防范人工智能潜在的伦理风险, 开发方和监管方需形成常态化的有效监测体系, 降低人工智能运行中发生伦理风险

的可能性，从而保护个体人身安全以及各类资产免受损失。

问责体系用于追究人工智能开发人员和运维人员实施违反伦理规则的行为后的责任。这一体系确保当损害后果出现时，可以精准地追溯到具体责任人并明确责任划分。这不仅要明确和细化人工智能开发主体履行义务的要求，还要完善相关的证据追溯流程，确定相应的免责条件和范围。在阿里巴巴宣布所有产品将接入AI大语言模型“通义千问”的同一天，国家互联网信息办公室发布《生成式人工智能服务管理办法(征求意见稿)》，并公开征求意见。这是生成式AI产业的首份监管文件，在我国监管体系发展中占据重要地位。

人工智能保障体系可为潜在的数字技术受害者提供基本保障。换言之，当人工智能利用不当造成他人人身和财产损害时，保障体系的存在可使受害者的损失得到一定程度的填补。人工智能保障体系不仅要保证公平，也要兼顾效率。国家要合理规划数字保障的范围，动态调整数字保障的统筹层次，建立系统完整的数字保障体系，以满足人民群众对数字社会日益增长的美好生活需要。

(四) 强化相关要素的标准化设计

数字伦理体系的制定者在构建数字伦理体系时，需对所有相关要素强化标准化设计。这些相关要素包括但不限于设计、算法和数据等。这意味着与人工智能、大数据相关的所有要素都应加以规范和设计。“标准”具有“清晰、简化、稳定”等特征，与泰勒的科学管理理论中所隐含的工具理性价值观相契合。人工智能系统标准化可以明确系统设计和应用中应遵循的伦理要求，引导人工智能系统的开发和应用符合伦理规范，确保人工智能系统产生可靠、准确的结果以及不会被用于恶意目的。人工智能系统的设计、算法和数据都应在一定标准体系下进行，以确保其安全有效。

其一，人工智能系统的设计应当标准化，以确保其一致性。标准化不仅有助于开发团队遵循统一的设计原则和实施细则，还能确保不同的系统或模块之间能更好地开展相互协作。这种一致

性有助于构建易于使用、高度可靠和安全的人工智能系统。

其二，人工智能系统中使用的算法产出过程应该标准化，以确保它们安全可靠。随着算法的应用场景越来越广泛，人类将大量的决策权力交给了算法。算法模型和推理框架的标准化，有助于实现算法产出过程的安全可靠，确保人工智能系统不会出现偏见或不公平的结果，作出符合道德原则的决策。

其三，人工智能系统使用的数据应该标准化，以确保结果的安全可靠。这是数字社会建设中的一项基础和前沿性的工作。在技术层面，要着手建立数据标准体系。数据的采集和处理同样要基于一定标准，确保数据来源真实可信。这有助于防范人工智能系统基于偏颇或不准确的数据作出决策。在人员和组织层面，要求从业人员严格按照标准化数据操作指南进行数据收集与处理，建立数据标准管理长效机制。

总之，人工智能系统的设计、算法和数据应进行标准化设计，以确保其安全有效，产生可靠和准确的结果。

(五) 对数字技术的全周期展开伦理考量

数字伦理体系制定者应对数字技术的全周期展开伦理考量。数字技术的全周期包括研发、测试、使用、监管等环节。人工智能的伦理考量是广泛而复杂的，应当将数字伦理贯穿在技术研发、技术测试、技术应用等全流程。

研发阶段主要由技术开发者主导，在这一阶段需要强调技术开发团队成员的伦理意识和责任感。数字伦理培训可以提高其对伦理问题的敏感性和深度理解，引导他们在技术研发过程中主动考虑伦理因素。数字企业还可以设置专门的数字伦理工程师岗位，明确其主要职责是将伦理价值目标落实到技术的具体架构中。

在技术测试阶段，应由独立的第三方组织(例如独立的伦理审查委员会)主导评估和验证人工智能系统的伦理合规性。委员会应由业界人士、政府官员和学界专家组成，通过制定技术规范和测试方法，来评估和验证人工智能系统的伦理合规性。例如，委员会可以进行公正性测试、隐私风险评估、透明度分析等，以确保人工智能系统

在各个方面符合伦理规范。委员会评估结果的通过与否, 成为其市场准入的先决条件。

在技术应用阶段, 政府、技术专家、公众的参与都必不可少, 他们都可以对数字技术的应用过程展开伦理监督。开发主体应当主动与用户共享有关人工智能产品的伦理信息, 并设置快速便捷的反馈响应渠道, 以不断完善伦理规则并防止用户受到显性或潜在的伤害。鉴于部分伦理审查落后于数字技术审查, 很多数字技术应用到终端市场后, 才显现出一系列的数字伦理问题。针对这一情况, 一方面, 政府可以通过行政指导和行政处罚等传统行政手段进行市场监管; 另一方面, 技术专家和公众可以通过公开的舆论监督对有悖于社会伦理的数字技术进行遏制。

总之, 只有对数字技术开展全周期的伦理考量, 才能使数字伦理不断优化完善。

(六) 遵循数字伦理框架性原则

除将伦理标准有机地融入行业标准和技术指南之外, 建构专门的数字伦理框架是系统解决数字伦理问题的重要举措。2020年7月, 五部门印发《国家新一代人工智能标准体系建设指南》, 涵盖了人工智能领域的安全与隐私保护、伦理部分的标准建设, 标志着我国在中央政府层面提出了新一代人工智能安全伦理标准体系建设的思路和内容。除此之外, 北京人工智能背景共识、中国新一代人工智能治理原则、欧盟提出的技术道德框架、谷歌的人工智能原则等各个层级的道德框架等, 也是在行业标准层面有益探索的代表性文件。数字伦理框架作为数字时代的行为准则, 发挥着方向性的基础作用。

在数字伦理框架中, 应当包含以下几个方面的原则: 一是有利人类生存与发展原则。数字伦理必须防止人工智能系统在数字社会出现异化、低俗和反人类。这一原则强调以增进人类福祉为中心, 符合马克思主义和新时代中国特色社会主义对数字伦理建设的应然要求, 它应当成为数字伦理的总体原则。二是和平与公共安全原则。在和平与公共安全原则的指导下, 各国在数字时代应当选择和平安全的发展路径。三是信息内容客观、准确、全面原则。信息传播的基础是真实性。合乎伦理性是指引信息传播的道德准则。四是诚

实信用原则。作为民法典的“帝王条款”, 诚实信用原则在数字空间比物理社会更具有独特意义。五是权利尊重与保护原则。权利尊重与保护原则作为一种道德框架, 具有普遍适用性。这意味着要将保障人的尊严放在数字发展的核心位置。数字发展应当使所有人在其权利得到尊重和保护的情况下参与社会活动。六是公平竞争与反垄断原则。算法垄断行为类型多样, 除通过法律规制和技术规制以外, 伦理规制是限制数字垄断的重要举措。七是透明及可审查追责原则。这一原则要求对借助数字技术、人工智能作出的重大决策, 尤其是存在潜在风险的决策, 其决策过程应当可追溯和被追责。八是公众参与和知情原则。公众是数字经济发展的参与者、监督者和受益者, 全社会应当广泛地参与数字技术的应用过程中, 建立有序互动的数字经济。九是社会最大利益和包容性原则。这一原则要求在数字技术发展的过程中, 不仅从利益最大化的角度出发增加社会整体利益, 而且要兼顾数字边缘群体和弱势群体。十是政治引领原则。政治引领有利于筑牢思想共识。国家应将政治道德教育融入日常的主题宣传、道德宣传、典型宣传、政策宣传等活动之中, 展现数字发展新成就、数字伦理新动态、数字从业者的新面貌, 从而为构建高质量数字伦理体系提供思想和政治保障。

三、数字伦理体系建构的具体路径

建构数字伦理的具体路径, 可以归纳为三个方面。

(一) 建构数字伦理体系中的“元规则”

数字伦理体系的元规则可以概括为两个方面: 一方面是“己所不欲, 勿施于人”。这一原则是指自己不愿意的, 不要施加给别人。“己所不欲, 勿施于人”有着深厚的历史渊源, 是我国一种普世的道德和伦理准则。随着时代的快速发展, 我们面临的道德和伦理问题也在不断变化。在这样的环境中, 一个固定和过于详尽的道德准则很可能迅速过时。然而, “己所不欲, 勿施于人”这一原则具有足够的灵活性, 能够契合数

字时代道德准则复杂多变的情境。另一方面是“推己及人，反求诸己”。这一原则是指你希望别人如何对待你，你就应当怎么对待别人。它与“己所不欲，勿施于人”有着相似内涵，但属于更为积极的道德准则。它不仅强调不对他人做你自己不愿接受的事，还进一步提倡在对待他人时应设身处地为他人着想，实施你自己希望接受的行为。这两方面构成的道德原则应当成为指导建设数字伦理体系的基本原则，或称为数字伦理体系建设的元规则。元规则在数字伦理形成以及应用的过程中起着基础性的指引作用。具言之，元规则可以发挥如下作用：

其一，元规则能为数字时代的技术人员和决策者提供基础性的方向指引。在数字世界中，许多数字系统和算法是高度复杂的，且常常缺乏透明度。这意味着，即使是接受过专业高等教育的专家，也可能难以完全理解一个系统的全部工作原理及其潜在影响。在这种情况下，“己所不欲，勿施于人”这一原则作为一个普遍的道德准则，能够为技术人员和决策者提供明确的指导方针。即便是在复杂和不透明的环境中，该原则也可以指导人们作出更道德、更负责任的决策。例如，将一个算法用于信贷评估，有可能因此影响某人能否贷款买房或购车，那么，算法的设计者需要考虑：我是否愿意自己或我的家人被这样一个我自己也不完全理解的系统评价？如果答案是否定的，意味着该算法系统可能需要修改，以确保其更为公平、透明和具备可解释性。

其二，元规则可以为机互动、机器互动的合理性判断提供理论依据。随着人工智能和机器学习的快速发展，数字伦理的复杂性也随之增加。这意味着我们不仅需要考虑人与人之间的互动，还需要考虑人与机器甚至机器与机器之间的互动。在此情形下，“己所不欲，勿施于人”也可以延伸适用于规范人与机器之间的交互活动。例如，当一辆自动驾驶汽车在决策时更倾向于保护车内乘客而不是行人，那么，设计者需要思考：如果我是行人，我是否会觉得这是公平的？此外，当一种人工智能系统开始与其他人工智能系统互动时，这一原则也非常有用。例如，在算法交易中，如果一个交易算法被设计成能够利用市

场敞口或者其他算法的弱点套利，那么，开发者需要思考：如果其他人使用相同的策略针对我的算法，我是否会觉得这是合适的？总体而言，在快速发展和多变的数字世界中，这一原则为如何作出道德和负责任的决策，提供了一种强有力且灵活的指导框架。

其三，在元规则的指导下，制定具体规则可以有效防止网络暴力行为的发生，保护人们的合法权益，维护社会的正常秩序。网络暴力是社会暴力在网络上的延伸，具有极强的社会危害性。随着数字网络的发展，网络暴力行为出现的频次愈来愈高。网络暴力的实施者以网络平台为载体，通过传播不实信息、诋毁的方式攻击他人，甚至引发社会秩序混乱及导致国家公共安全受损。网络暴力根源于伦理价值观，系因群体的价值判断而产生的群体极化行为^[16]。在元规则的引导下，数字网络参与者可以形成自我批判和反思的习惯，有助于铲除滋生网络暴力行为的文化土壤，塑造风清气正的网络环境。除此之外，它还能有效地防止他人的隐私被非法泄露。互联网技术的进步导致了个人隐私信息的数据化和分散化，信息的拥有者和处理者能够通过重复使用或未公开的交易将这些信息转化为经济收益。当个人信息在未经本人同意的情况下被盗用时，不仅会严重损害个人的尊严，而且使个人面临自主性丧失的风险。在元规则的指导下，企业应当充分考虑到使用者地位的不对等性，尽可能地将各类信息告知使用者，并通过技术手段保障用户的隐私信息。这些举措能够让用户在享受服务时感受到平等、透明和安全，有效地防止信息使用者和保存者实施不道德的行为，进而保护个人隐私，从源头上防止网络暴力现象的出现，维护社会的正常秩序。

其四，在元规则的指引下，大型企业滥用算法权力的冲动被遏制，倾向于作出合乎伦理道德的商业行为，保障消费者的合法权益。数据是对网络使用者网络行走的记录，当其成为数据时，也必然伴随一定的伦理风险^[17]。在数字经济背景下，平台企业因其固有的规模效应、网络效应及外部性等属性，倾向于利用数据上的优势进行跨领域竞争以实现规模扩张，从而导致事实性的市

场垄断发生。而垄断市场中算法权力的异化, 又会导致平台企业实施欺诈、歧视、排斥、杀熟等不合乎商业伦理的行为。例如, 在电子商务领域, 平台企业可以对消费者的浏览和消费记录进行收集分析, 以消费者可接受的最高价格推送商品信息, 以达到利润最大化的目的。这类行为不仅会侵犯消费者的知情权和公平交易权, 也会损害企业的声誉, 影响企业的长远发展。在元规则的指导下, 应当尽可能地降低消费者维权的成本, 畅通反馈渠道, 充分发挥消费者的个体监督作用。这可以促使企业认真对待并重新定位与客户之间的商业关系, 有效地防止不合乎商业伦理的行为发生, 进而维护消费者的合法权益, 保护企业的声誉。

总之, 确立元规则在数字伦理中的作用十分重要。它可以有效地保护公民和相关组织的合法权利和利益, 保障社会的正常运行, 促使数字社会中的不同主体承担相应的社会责任, 维护社会秩序。元规则为数字主体作出道德和负责任的决策提供了清晰易理解的指导。这使得它成为一个在当前复杂多变的数字环境中具有高度可操作性和应用价值的伦理准则。我们应该促进元规则在不同领域充分细化和发展, 引导数字主体实行合乎各方利益的行为, 营造美美与共的数字社会环境。

(二) 建构数字伦理体系中求善、求真、求美的共同价值目标取向

求善、求真、求美是数字伦理体系建构中共同的价值目标取向。第二次世界大战后, 作为起草方之一的中国, 与世界其他国家一同推动了世界人权宣言的颁布, 其中的很多原则已经渗透到全球的法律和政治文化中, 它们包括尊严、尊重、自由、安全、民主、隐私、身份、公平和平等。这表明, 即使各个国家拥有不同的文化背景, 仍存在共识性价值取向。在如何应对数据驱动社会所发生的变化的讨论中, 这可以作为集体反思的基础^[18]。人类对数字社会愿景的共识, 应当是求善、求真、求美。人工智能应该是真实的、可靠的、有利于人类发展的。人工智能系统需要真实, 才能提供一致和准确的结果, 并且应该有利于人类发展, 以便在现实世界中充分发挥作用。

其一, 这一取向有利于打破算法黑箱, 保证人工智能系统的真实可信。提高算法的透明度是实现数字伦理价值的重要途径, 也是解决算法黑箱问题的主要手段。一方面, 数据真实是人工智能系统真实透明的前提要素。“垃圾进, 垃圾出”作为一种基本运转方式, 说明人工智能系统只是处理它所提供的内容。因此, 算法收集信息的方式方法需要符合一定的形式和实质要求。另一方面, 在保证数据真实的情况下, 数据处理流程需要保证透明度。由于公开的算法信息有可能超出公众的专业认知水平, 导致公众难以理解算法的运行原理。因此, 处理数据的算法应当选择可解释性较强的模型, 以增强公众的理解程度, 以及当伦理问题出现时具备可追溯性。

其二, 这一取向有助于保障人工智能系统的可靠性及责任承担的可实现性。一方面, 人工智能应当只根据接收到的数据等作出决策。接收到相同数据应当给出相同判断, 接收到不同数据应当作出合理的差异性结果。即使面对有噪声或不完整的数据, 人工智能系统也应该提供大体一致的结果。另一方面, 人工智能在发展的过程中, 应当在追求效率和速度的同时兼顾责任与担当。鉴于算法在人类生产生活中扮演着愈来愈重要的角色, 为了防止算法霸权的产生, 算法开发者必须进行负责任的创新, 以协调人类和人工智能之间的紧张关系。

其三, 这一取向有助于促使人工智能系统满足人自由而全面发展的终极追求, 并在现实世界中反映出来。人民群众是历史的创造者, 社会的进步离不开人本身的进步。人类除了具有基础性的生存需求之外, 还有着发展上的需求。资本主义在追求发展的过程中, 过度地重视技术的作用, 忽视了人类的主体地位, 使得技术走向了异化的发展之路^[19]。因此, 人类在依靠人工智能获得更多自由时间的同时, 也要警惕人类自身价值存在被侵蚀的危险。马克思认为: 自由时间就是指可以自由支配的时间, 是人的闲暇时间, 是人可用来从事较高级活动的时间, 是人能够直接用于发展自身各种本质力量的时间, 是使个人得到充分发展的时间^[20]。人工智能在减少人们必要劳动时间的同时, 应当为人们提供更好的发展空间。

间,帮助人实现本身的价值。例如,帮助人类获得知识和技能,创造就业,链接人类情感,促进人类形成自我价值观等。

总之,人工智能系统不仅应该是真实和可靠的,还应该以人为本,有益于人类发展。在数字时代,这一传统的伦理原则不仅能够保持其基本的价值观和道德指导作用,还能适应新的技术和社会环境。因此,这一古老的伦理原则在数字时代可以焕发出新的活力。

(三)建构适应具体化、类型化、多样化需求的数字伦理规则

国家应根据不同时期、不同主题、不同环境、不同领域制定具体的数字伦理规则。数字技术的更新日新月异,当下的数字伦理极有可能无法约束未来的人工智能。因此,数字伦理的发展永远在路上,不可能毕其功于一役。新一轮数字技术发展的热潮,对传统伦理体系的冲击已经开始显现,部分冲击可以在现有的伦理体系架构内得以消解,但某些冲击已经对当下的伦理体系提出了颠覆性的挑战。例如,数据的真实性与可靠性问题、数字身份的隐私与污名化、信息茧房与自主性的丧失等。随着数字技术的不断发展,人们利用数字伦理调整社会关系的范围也在不断扩大。与复杂的数字关系、前沿技术和社会现实相比,数字伦理体系的建构可能是滞后的。要缩小两者之间的差距,不但要求我们在数字伦理的制定过程中有适当的预判,而且要在数字实践中与时俱进地完善数字伦理规则,填补已发现的漏洞。数字伦理的制定者应及时普及最新的数字伦理,确保人工智能系统的安全可靠,更好地服务于数字社会。

政府规制数字技术的具体行政手段和方法也应做到因地制宜。人工智能为推进国家治理体系和治理能力现代化提供了有力的技术支撑,数字中国建设可借助数字技术的优势加以实现。当前,数字经济成为我国经济增长的重要引擎,数字政务体系不断优化,数字文化产业让更多人感受到文化的魅力。我国数字基础设施建设取得了显著成果,但数字中国建设面临着诸多挑战和问题。我国地域辽阔,不同地区对于数字技术应用的水平不尽相同。对地方政府而言,解决不同地

域、不同产业中出现的具体问题,才是当下面临的主要挑战,例如如何将数字工具与本地现有产业有机结合起来,开展数字产业培育和集聚。数字中国建设不仅需要长远的战略部署和综合政策工具的支持,而且需要给予地方行政部门一定的自主发展权限。谁的战略思路正确、配套政策运用得好,谁就能在数字化建设的过程中脱颖而出。这不仅有利于最大限度地发挥不同类型数字产业的潜力,而且也为私人创新保留了更大空间。不同区域在互相竞争和学习的过程当中,能够不断完善数字治理制度,并在制度竞争的过程中逐步形成全面科学的数字治理方法和治理体系,从而为数字经济社会的高质量发展提供强劲动能。

四、结语

时移世异,数字技术的高速发展使其成为当代社会不可分割的一部分,而数字伦理正是伴随着社会转型而兴起的全新命题。通过建构完善的数字伦理体系,可以有效地消解数字技术带来的社会风险,保障数字技术的健康发展和实现数字社会的良性治理。一方面,社会不同利益主体要对数字伦理规则建构的基本路线、基本价值追求、基本形式等形成共识;另一方面,构建数字伦理体系是一项长期而艰巨的任务,要在实践中不断探索与完善,从而为数字中国的高质量发展保驾护航。

参考文献:

- [1] SALAM A F, DAI H, WANG L. Online users' identity theft and coping strategies, attribution and sense of urgency: A non-linear quadratic effect assessment [J]. *Information Systems Frontiers*, 2021, 24(6): 1929–1948.
- [2] 李志祥. 数字智能技术的正义挑战与伦理应对[J]. 上海师范大学学报(哲学社会科学版), 2023, 52(6): 94–102.
- [3] 吴玄. 云计算下数据跨境执法: 美国云法与中国方案 [J]. 地方立法研究, 2022, 7(3): 96–110.
- [4] 尤瓦尔·赫拉利. 未来简史[M]. 林俊宏, 译. 北京: 中信出版社, 2017.

- [5] 张康之. 在风险社会中看人的道德及其善的追求[J]. 河南社会科学, 2022, 30(6): 1–11.
- [6] 李正风, 王硕. 数字素养、数据权利与数字伦理[J]. 科普研究, 2022, 17(6): 8–14, 108.
- [7] JONAS H. The imperative of responsibility[M]. Chicago: The University of Chicago Press, 1984.
- [8] 田旭明. 数字社会的主要伦理风险及其应对[J]. 中州学刊, 2022(2): 87–93.
- [9] 张博. 数字治理的责任伦理向度: 情境转换、实践困境与破解之道[J]. 理论探讨, 2023(4): 76–82.
- [10] 冯建军. 网络公民教育: 智能时代道德教育的新要求[J]. 伦理学研究, 2022(3): 1–9.
- [11] 刘培, 池忠军. 算法的伦理问题及其解决进路[J]. 东北大学学报(社会科学版), 2019, 21(2): 118–125.
- [12] 米歇尔·福柯. 规训与惩罚[M]. 刘兆成, 杨远瀛, 译. 北京: 三联书店, 2007.
- [13] 罗理章. 构建数字命运共同体的伦理蕴含[J]. 人民论坛, 2022(4): 58–61.
- [14] 中共中央关于坚持和完善中国特色社会主义制度、推进国家治理体系和治理能力现代化若干重大问题的决定[EB/OL]. (2019-11-05) [2023-11-23]. http://www.xinhuanet.com/politics/2019-11/05/c_1125195786.htm.
- [15] 马克思恩格斯选集: 第2卷[M]. 北京: 人民出版社, 2012.
- [16] 王静. 数字公民伦理: 网络暴力治理的新路径[J]. 华东政法大学学报, 2022, 25(4): 28–40.
- [17] 颜世健. 数据伦理视角下的数据隐私与数据管理[J]. 新闻爱好者, 2019(8): 36–38.
- [18] ALLEN A L. Debating ethics and digital life[J]. Eur. Data Prot. L. Rev., 2019, 5(1): 7–12.
- [19] 宋建丽. 技术异化与自由追求之失[J]. 内蒙古大学学报(人文社会科学版), 2005(1): 94–100.
- [20] 中共中央编译局. 马克思恩格斯全集: 第46卷: 上[M]. 北京: 人民出版社, 1980.

On the construction of a digital ethical system

JIANG Bixin, LIU Zhuoquan

(School of Law, Hunan University, Changsha 410082, China)

Abstract: Constructing a digital ethics system is beneficial to responding to digital technology and dual nature of AI, overcoming the limitation of legal norms and conventional governance modes, guarding against digital network security and ethical technology risk, guaranteeing the healthy operation of digital society, and dealing with the severe challenges that digital technology poses on conventional ethics system. In setting up digital ethics system, we should effectively handle five types of relationships, ensure substantial participation of all relevant stakeholders, establish systems throughout the entire chain, strengthen the standardized design of all relevant elements, conduct full lifecycle ethical considerations of digital technology, and follow the framework principles of digital ethics. We should start from meta-rules, and should be guided by the pursuit of goodness, truth, and beauty so as to construct digital ethics rules that combine abstraction and concretization, and that are adaptable to different periods, themes, environments, and fields.

Key Words: digital society; digital ethics; digital technology risks; digital order; digital values

[编辑: 苏慧]