



人工智能技术治理应抓住要害

□ 赵精武 周瑞珏

自 ChatGPT-4 上线之后, 人工智能技术创新浪潮再度掀起, 如果说之前还有观点质疑当下并非规制人工智能技术的最佳时机, 那么 ChatGPT-4 所展现出的高度智能化功能使得这些质疑观点逐渐消散。正如过去各类信息技术发展规律, 生成式人工智能技术在不断进行创新的同时, 也正在扩大社会风险的波及范围和损害程度。尤其是在网络信息内容生态领域, 人工智能技术能够根据用户提出的具体需求, 在短时间内生成海量文本信息、图片信息乃至视频信息, 一旦这些生成信息在网上传播, 网络空间将不得不面对新一轮的信息浪潮的冲击。一方面, 生成类信息具有典型的同质性特征, 加之平台以用户偏好推送相关信息, 这使得用户可以获取的信息内容、信息类型均有所衰减; 另一方面, 人工智能技术所生成的信息越发贴合人类的表达逻辑, 这使得一般用户无法轻易辨别信息内容的真伪。当然, 人工智能的技术风险远不止于此。例如, 在科研教学领域, 人工智能技术能够成为学生自主学习的辅导工具, 但过度依赖人工智能技术完成课程作业或文献撰写, 从长远来看只会削弱学生批判意识和独立学习能力。在网络安全领域, 人工智能技术已经能够为部分代码片段编写提供辅助指导和查漏补缺功能, 这无疑会降低恶意代码编写的技术门槛与时间成本, 进而导致网络安全态势的恶化。

事实上, 人工智能技术所引发的诸多社会风险问题早已不再是理论空想, 在当下的社会生活中, 已经存在滥用人工智能技术进行违法犯罪活动的事件。在今年4月侦破的一起利用人工智能发布网络虚假信息

息的犯罪活动, 犯罪嫌疑人为了规避平台的信息内容审核机制, 利用 ChatGPT 对新闻素材进行修改编辑, 上传至平台非法获利。无论是专家学者预估的社会风险, 还是现实生活已经发生的犯罪活动, 无不说明当下人工智能技术需要规范治理。国家互联网信息办公室也在4月发布了《生成式人工智能服务管理办法(征求意见稿)》, 用于指引和规范生成式人工智能技术的应用活动。

然而, 这远远不足以规制潜在的技术滥用风险, 传统的治理理念和治理工具在人工智能技术治理领域正在逐渐失效。其中的原因正是技术服务提供者无法完全控制用户的使用行为, 作为技术工具的人工智能应用, 其输出结果是否合理合法, 在很大程度上取决于用户所输入的问题和需求。更重要的是, 对于部分海外人工智能服务提供者而言, 国内立法难以在域外予以适用, 非法使用者利用境外技术服务从事违法活动, 监管者往往难以在事前阶段就予以预防。继续僵化沿用“危险制造者承担责任”的治理逻辑只会适得其反, 既会限制人工智能技术的创新应用, 也会浪费既有的社会治理资源。因此, 根据技术创新应用实践特征, 适时调整治理理念和治理方式越来越重要。

在治理理念层面, 首先需要明确的是, 人工智能技术治理应当秉要执本, 优先预防和规制核心的技术滥用风险。在人工智能技术引发的诸多社会风险中, 危害性最大的风险类型当属网络信息内容生态风险。AI换脸、AIGC生成仿真照片等信息服务内容正在重新塑造网络信息内容生态, 这些服务所生成的信息内容“半真半假”, 这使得用户更加难以在信息大爆炸时辨别信息真伪。同样的, 平台运营者乃至监管机构也需要花费更多的时间和精力去一一核对这些信息的可靠性与真实性。而人工智能技术所引发的数据安全、网络安全、算法歧视等风险类型本质上并没有超出现有的技术风险治理框架, 数据安全法、网络安全法和《互联网信息服务算法推荐管理规定》等法律法规对此均有提及, 并且监管机构和平台也能够采取技术手段实现高效的事前监管效果。进一步而言, 人工智能技术治理活动应当分阶段、分重点, 在不同的技术创新发展阶段, 技术风险的类型、特征具有不同特点, 倘若“眉毛胡子一把抓”, 显然无法从根本上规制技术风险, 只会让技术风险危害性逐步加深。因此, 在现阶段人工智能技术治理的首要任务应当是全面预防和治理滥用人工智能生成违法信息内容, 保障网络信息生态安全。

在治理方式层面, 监管机构、平台运营者、技术研发者以及用户应当共同参与到人工智能技术治理活动中, 以网络信息生态安全为基本导向, 细化多元主体共治的基本内容, 避免陷入单一主体责任承担的治理误区。对于技术研发者而言, 应当将技术研发与安全评估同步实施、同步进行, 在相关信息服务商业化应用之前, 进行算法安全风险评估、科技伦理审查, 重点对生成内容的合法性和可控性进行评估和调整。对于平台运营者而言, 应当严格落实网络实名制的基本要求, 在用用户注册和使用, 提醒用户合理使用人工智能信息技术服务。同时, 平台运营者应当定期更新优化信息内容审核机制, 一方面以关键词、图片自动化识别等方式屏蔽、删除明显违法的网络信息内容, 另一方面要求用户在上传人工智能生成信息时进行标注和说明, 用以提示其他用户该信息内容的来源和产生方式。对于监管而言, 应当适时调整监管模式, 除了社会公众举报投诉等传统方式之外, 还可以充分利用监管技术, 以人工智能监管技术辅助识别可能涉嫌违法违规的人工智能生成信息。此外, 还需要根据技术创新发展趋势, 周期性调整人工智能技术重点监管事项, 重视科技伦理审查机制的灵活性, 引导人工智能产业良性发展。对于用户而言, 应当合理使用人工智能信息服务, 树立法治观念, 充分认识到发布网络虚假信息可能承担的法律责任; 也需要保持警惕意识, 不轻信非正规来源的网络信息, 理性发布网络评论信息。

人工智能技术治理难以“毕其功于一役”, 需要在技术创新发展的进程中抓住技术风险的主要内容和作用方式, 适时调整和更新治理理念和治理方式, 方能消解人工智能技术滥用所导致的社会危害性。

弥合城乡“数字鸿沟”该怎么做

□ 张培培

在数字化进程中我国取得了举世瞩目的成绩。然而, 也应该看到, 数字化发展不平衡、特别是农村地区信息化进程明显落后于城市地区, 城乡“数字鸿沟”问题日益凸显。“数字鸿沟”是指不同人群、行业、区域等在信息基础设施的接入、数字技术的开发与应用、数字资源的获取使用等方面的不平等现象。城乡“数字鸿沟”本质上是城乡经济社会发展二元结构在数字时代的反映。城乡“数字鸿沟”不仅成为当前数字中国建设的痛点, 也成为未来我国实现高质量发展的短板。分析城乡“数字鸿沟”的表现, 探索弥合城乡“数字鸿沟”的有效路径, 成为新时代乡村振兴战略实施过程中亟待解决的重要现实问题。

(一)

不同数字化发展阶段, “数字鸿沟”的表现也不一样。接入鸿沟、使用鸿沟和能力鸿沟代表了“数字鸿沟”的不同类型。随着数字化进程的加快, 我国城乡“数字鸿沟”已从最初的接入鸿沟逐渐转向更深层、更新型的使用鸿沟和能力鸿沟, 特别是近年来随着大数据、区块链、人工智能等新一代信息技术的广泛应用, 城乡“数字鸿沟”有了新的表现。

第一, 农业领域产业数字化水平滞后于非农产业。产业数字化是指将现代数字技术应用于农业、工业和服务业等传统领域以实现其数字化转型。尽管有农村电商、数字农业等发展亮点, 但受制于“三农”领域基础数据资源体系建设不完备、涉农信息技术工具匮乏、数字技术的农业应用范围有限等原因, 农业领域的数字化程度仍然相对滞后。根据《中国数字经济发展研究报告(2023年)》显示, 2022年我国服务业、工业的数字经济渗透率分别为44.7%、24.0%, 而农业领域数字经济渗透率仅有10.5%, 作为第一产业的农业生产数字化程度远低于工业和服务业, 农业领域的数字经济发展落后已成为数字中国建设的突出短板。

第二, 乡村治理数字化以及公共服务数字化水平与城市存在差距。农村地区在线上与线下互联方面滞后于城镇, 乡村数字孤岛现象突出, 这些都妨碍了信息化与乡村治理的深度融合, 导致农村地区党务、村务的数字化水平、涉农事项在线办理能力、数字网格化治理水平偏低。另外, 乡村公共服务的数字化供给能力相对较弱, 特别是在数字教育、在线医疗、数字普惠金融等领域的城乡差距依旧巨大。这不仅阻碍了数字时代农村基层治理模式创新, 也影响了村民利用数字技术参政议政的热情, 农村居民无法像城市居民那样分享到治理数字化和公共服务数字化带来的新型数字红利。

第三, 农村居民在数字素养以及使用信息资源的机会和能力方面有待提高。一方面, 高新技术产业使得相关人才向城市聚集, 乡村数字人才短缺已成为制约乡村数字经济发展、农业数字化升级的明显瓶颈。另一方面, 农村劳动力素质、受教育水平总体较低, 加之近年来乡村人口结构性变化, 农村空心化现象不容忽视, 农村居民在深层次利用网络, 特别是使用信息资源和数字技术就业、创收方面远远落后于城镇居民。随着人工智能、区块链等新一代信息技术的发展, 技术使用门槛再一次提高, 缺乏数字技能以及相关培训的农村劳动者可能处于不利的位置, 进一步拉大农村居民在就业、教育、医疗以及整体经济机会方面的差距。

(二)

以数字乡村战略弥合城乡“数字鸿沟”, 积极探索数字乡村发展新模

式, 促进城乡协调发展, 不仅有助于广大农村居民公平地参与数字经济、共享数字普惠成果, 更是推进乡村全面振兴、实现共同富裕的重要途径。

第一, 强化农村信息基础设施建设, 推动城乡互联互通。着力提升农村“新基建”的发展速度, 加强光纤网络、移动宽带等基础服务, 布局农业物联网、智慧物流、人工智能等新型基础设施建设; 加快农村大数据与农业综合信息服务平台建设, 开发适应“三农”特点的信息终端和应用软件, 建立数据共享机制、拓展数据获取途径, 不断提升农村地区的数据获取和共享水平; 加快乡村交通、水利、气象、电力、冷链物流等基础设施的数字化转型, 完善农村电商配套基础设施; 鼓励多元主体参与农村及偏远地区的信息基础设施建设, 扩充资金来源; 积极推动城乡信息基础设施互联互通以及同步规划建设。

第二, 加快农业产业数字化发展, 实现数字农业升级。推进数字技术与农业农村经济深度融合, 将数字化融入包括农作物种植、营销等农业全产业链各环节、各领域; 推动云计算、大数据、区块链、物联网等现代信息技术在农业生产经营中的应用, 打造无人农场、数字农场、数字果园、数字渔场等场景; 推动涉农数字产品的研发生产, 因地制宜发展“互联网+”特色产业; 持续推进“数商兴农”, 以电商为抓手, 加快农业生产、加工、销售、物流等产业链数字化升级, 让手机成为新农具, 线上直播等电商模式成为新农活; 加快数字乡村创新创业基地建设, 拓宽乡村产品服务的销售渠道; 推进数字乡村发展投融资创新, 鼓励企业投入乡村振兴事业, 把社会资本和公共投资整合起来; 推进城乡要素双向自由流动, 以信息流带动技术流、资金流、人才流、物流流向农村地区集聚, 形成以城带乡、城乡融合的数字经济发展模式。

第三, 推进乡村治理数字化, 构建数字乡村社会。配合乡村基层治理的相关目标, 推动数字党务、村务、商务在各地乡村个性化实践; 建设乡村治理数字化平台, 同步接入国家和省、市级平台, 打破涉农职能部门之间的信息壁垒, 推动涉农政务数据的开放共享, 基本实现“数据多跑路、群众少跑腿”; 打造农村社区服务平台, 推进涉农服务事项在线办理, 完善“互联网+网格治理”服务管理模式, 优化政务服务的民生小程序; 完善国家智慧教育平台, 重视农村教育数字化建设, 完善农村中小学互联网教学机制, 将城市优质教学资源引入乡村; 提高农村医保信息数字化建设, 发展数字健康, 探索优质资源通过远程医疗、互联网医院、在线问诊等方式进一步下沉到乡村的方案路径。

第四, 提升农村居民的数字素养, 共享新型数字红利。构建和完善农村数字人才体系, 建立城乡人力资源要素双向流动机制, 鼓励数字专家和人才向乡村基层流动, 支持高校为乡村振兴提供智力服务, 通过远程培训、委托实习等方式提升乡村教师素质, 培养与现代农业深度融合的数字技术应用人才; 加强农村地区的数字教育投入, 开展农民数字技能培训, 提升乡村居民在日常生产生活中运用数字技术的能力; 抓住人工智能发展契机, 开发人工智能基础教育, 发挥人工智能在推动优质资源普惠中的覆盖作用; 开发针对边缘人群及地区的数字化工具, 更好地服务农村地区; 推动“互联网+”教育、医疗、法律、金融、文化等向农村延伸, 让农民更好地分享互联网发展成果。



守好网络平台舆论阵地

□ 冀翠萍

互联网新媒体快速发展, 不断塑造着日渐繁荣的网络舆论场。同时, 不可忽视的是网络谣言滋生以及网络暴力的发生, 破坏了舆论生态环境。党的二十大报告鲜明指出, 健全网络综合治理体系, 推动形成良好网络生态。网络空间不是“法外之地”, 是亿万民众共同的精神家园, 需要牢牢守住网络舆论场的“边界线”, 扎紧网络法治的“安全网”, 促进多元主体协同参与, 共同营造清朗网络空间。

平台成为网络舆论的重要承载地

在传统媒体时代, 媒体生产信息同时也把关信息, 信息遵循“先过滤后发布”的原则, 由新闻从业者规范信息生产、严把信息内容, 由主流媒体管理分发渠道、控制传播规模。进入互联网时代, 生产权和传播权向普通公众转移, 社会各主体都可以生产信息、制造内容, 从而带来网络信息泛滥、信息过载和信息爆炸。海量信息如何高效触达用户? 互联网平台运用其强大的数据采集、计算能力, 承担起网络信息的分发与传播功能, 依靠算法模型筛选出符合用户兴趣与偏好的信息, 进而精准推送给用户。网络信息遵循“先发布后过滤”的方式, “算法推荐”成为信息分发的主流模式, 能够有效干预信息流向。然而, 触达用户的网络信息并未结束传播历程, 用户阅读、转发、点赞和评论信息又形成了二次传播。专业内容生产者还将信息作为内容生产的原材料, 叠加用户的评论信息, 进行剪辑、视频化等二次创作, 推动网络信息发酵为舆论热点议题, 于是专业内容生产者也成为舆论议题的设置者和

引导者。由此, 互联网平台成为聚合信息内容、聚集内容生产者的中心节点, 成为公共事务讨论的主体空间, 成为网络舆论场的重要承载地。从某种程度上说, 互联网平台引导着舆论的呈现与流向, 应当承担起网络舆论治理的责任。

“流量至上”下的网络舆论变化

互联网时代, 尤其是在网络平台, 评价传播效果最有效、最直接的指标就是“流量”。无论是基于用户兴趣爱好的推荐, 还是基于事件热度的推荐, 都将激发用户的关注和参与, 形成网络“流量”。一方面, 用户的参与度、活跃度等流量数据, 使平台流量变现, 也成为资本市场评价平台的重要一环。另一方面, 用户关注度成为互联网时代流量密码, 各个媒体都在竞相争夺用户注意力, 而争议性、话题性内容往往能够迎合用户心理需求, 吸引用户关注, 互联网平台也乐于持续将热点事件和信息推送给用户。有自媒体生产者指出, 在内容创造中加入“热点”更容易获得关注和推荐, 用户关注与平台推荐相互配合, 推升事件的热度值和传播力。

在“流量至上”和经济利益的驱动下, 一些自媒体追求奇观化、景观化的内容生产, 如编造信息、恶意炒作等现象。可见, 互联网平台并未呈现客观、全面的信息, 反而产生了一些低效率、谬误性信息, 甚至一些谣言占据互联网平台的醒目位置, 严重侵害网络舆论生态。一些互联网平台、自媒体等在“流量”的追逐中, 丢弃了对公共价值的追求, 脱离了秩序化传播的轨道, 构建了偏向性的信

息环境, 产生虚假舆论共识。

构建良好网络平台舆论生态

增强互联网平台的“把关”效能。互联网平台要清醒地认识到舆论秩序的重要性, 明确作为国家、社会与公众“连接器”的角色定位, 坚持正确舆论导向, 把严、把实、把好人内容生产与传播关口, 打破单纯追逐“流量”的行为。对于符合社会主义核心价值观、客观呈现事件全貌的予以积极推荐。探索建立透明化的内容来源引用机制, 推动自媒体主动规范内容生产与传播。此外, 探索从信息内容治理转向对互联网账号及自媒体所属机构治理, 加强对互联网账号运营团队的资质审核, 备案团队的资金来源、公司资质等, 加强对自媒体的审查管理, 降低对同一机构账号的推荐权重, 推动网络产品形态机构主动规范账号管理和内容生产。

探索推动互联网信息分类分级管理。信息分类分级体现了科学化、精细化、针对性的治理思路, 可以有效规避传统治理中的“一刀切”问题, 是平衡互联网自由和秩序的重要保障。探索互联网信息的分类管理制度, 明确信息内容所属范畴。对于用户发布的信息内容, 依据信息类别建立适应性的审核与研判机制, 加强对内容相似比高的个体信息的审核, 加强对带有新闻性质信息的审核与监管, 加强对求助举报信息传播链条的监管。同时, 探索互联网信息的分级管理方式, 依据信息真实性级别, 建立针对性的审核与监管机制, 优化监管力量和资源配置, 提升审核监管效能。

健全互联网法律法规体系。我国网络法治路径逐渐清晰, 网络安全法、《网络信息内容生态治理规定》和《互联网信息服务算法推荐管理规定》等法律法规的制定, 初步构建起较为完整的互联网法治体系。但是, 在构建平台互联互通规则、界定违法行为的内涵与外延、明确违法行为的规范方式、法律运用实践等方面还存在不足, 迫切需要深入分析各类型违法行为的构成要素, 探索制定专门法的立法模式, 加快重点领域的网络立法进程, 推动互联网法律法规向细分领域、舆论生态全域延伸。在此基础上, 还要加大网络执法力度, 对公众反映强烈的、带有主观恶意的违规行为, 对严重影响网络舆论秩序的恶意行为, 从严从重处罚。发挥惩戒和警示作用, 推动网络执法“落地有声”, 彰显网络舆论场“边界线”的威力。

营造有序参与网络舆论的社会氛围。从治理行动上引导舆论参与, 一次治理行动, 也是一次网民舆论参与的再教育, 引导网民健康有序地参与舆论监督。提升网民媒介素养, 突破“信息茧房”, 拓宽信息接收视野, 了解倾听多元化声音。在突发事件或热点面前, 保持理性审慎态度, 多接触有明确引用来源的信息, 提升自身理性思辨能力。有效利用“违法和不良信息举报”“风纪委员会”“用户众议”“专家顾问团”等方式, 不断提高信息甄别的准确性。还要及时吸纳用户意见建议, 引导用户自觉维护和发展理性有序的舆论生态。

