

数智化时代网络空间治理的风险防控与价值引领

李文菁, 赵宽容

四川城市职业学院, 四川 成都

收稿日期: 2025年12月7日; 录用日期: 2025年12月30日; 发布日期: 2026年1月9日

摘要

随着我国互联网普及率突破78%, 网络空间成为风险防控的新阵地, 其为社会带来发展机遇的同时, 也带来了主流价值被稀释、风险防控难度加大等全新挑战。本研究旨在探索数智化时代下的网络空间治理的理论和机制创新。通过构建“技术-制度-价值”三层融合的智能治理理论模型, 明晰数据智能、规则协同与目标共识间的系统化关联路径, 进而提出“底线安全-柔性引导”本土化双维治理机制, 力求在筑牢意识形态的基础上, 实现圈层化、情感化、精准化的价值传播。研究通过对国际数字治理经验的批判性借鉴, 设计包含可量化指标(如公共认可度PAR、引导效能指数GEI)的实践框架与评估体系。最终, 从顶层设计、人才培养、动态评估等方面提出推广措施, 为提升网络空间治理工作的精准性、前瞻性与有效性提供理论支撑与实践方案。

关键词

网络空间治理, 大数据驱动, 风险防控, 治理机制

Risk Prevention and Value Guidance in Cyberspace Governance in the Digital Intelligence Era

Wenjing Li, Kuanrong Zhao

Sichuan Urban Vocational College, Chengdu Sichuan

Received: December 7, 2025; accepted: December 30, 2025; published: January 9, 2026

Abstract

With my country's internet penetration rate exceeding 78%, cyberspace has become a new frontier

for risk prevention and control. While it brings development opportunities to society, it also presents new challenges such as the dilution of mainstream values and increased difficulty in risk prevention and control. This study aims to explore theoretical and mechanistic innovations in cyberspace governance in the digital intelligence era. By constructing a three-tiered integrated intelligent governance theoretical model of “technology-institution-value,” this research clarifies the systematic relationship between data intelligence, rule coordination, and goal consensus, and then proposes a localized two-dimensional governance mechanism of “bottom-line security-flexible guidance.” This mechanism seeks to achieve layered, emotional, and precise value dissemination while strengthening ideological foundations. Through critical analysis of international digital governance experiences, the study designs a practical framework and evaluation system including quantifiable indicators (such as Public Acceptance Rate (PAR) and Guidance Effectiveness Index (GEI)). Finally, it proposes implementation measures from the perspectives of top-level design, talent training, and dynamic evaluation, providing theoretical support and practical solutions for improving the accuracy, foresight, and effectiveness of cyberspace governance.

Keywords

Cyberspace Governance, Big Data-Driven, Risk Prevention and Control, Governance Mechanism

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言：智能时代网络空间治理工作的战略意义与概念框架

随着互联网技术的迅速发展,网络空间已成为国家治理和意识形态的重要场域。截至 2024 年 12 月,我国互联网用户规模达 11.08 亿,互联网普及率达 78.6%,移动互联网用户占比高达 99.7% [1]。在这一时代背景下,以《中华人民共和国网络安全法(2025 年修订版)》为基石的国家网络治理法规体系不断完善,为规范网络空间秩序、保障公民合法权益提供了坚实法律支撑和根本遵循依据[2]。网络空间不仅为风险防控工作提供了新的路径和方法,也带来了数据超载、价值冲突等多重挑战。传统的网络治理工作在新时代下面临人才储备不足、技术发展缓慢和组织形式单一等问题,亟须通过理论和机制创新实现风险防控与价值引领的双重目标[3]。大数据、人工智能等新一代信息技术为网络治理工作带来了前所未有的机遇,通过数据驱动的精准感知、风险预警和价值引导,网络治理工作可以从“模糊画像”转向“精准刻画”,从“事后处置”转向“事前预防”。在此背景下,本研究以“大数据驱动”为核心视角,探索网络空间治理工作机制的创新路径,具有重要的理论意义与现实紧迫性。

本研究聚焦网络空间治理、大数据驱动、风险防控和价值引领,系统分析当前网络治理工作的现状与理论困境,提出“技术-制度-价值”三层融合的理论模型,探讨其创新价值与推广路径,为新时代网络治理工作的数字化转型提供理论支撑与实践参考。

2. 现状与理论困境：网络治理工作的实际进展与研究短板

2.1. 实践现状：技术赋能与多元探索

近年来,随着大数据、人工智能等新一代信息技术的快速发展,网络治理工作在平台建设、技术应用与机制创新方面取得了显著进展,形成了“智能感知、风险预警、精准引导”三大方向的实践探索:

在平台建设载体创新方面, 各地方与机构依托技术优势, 构建了多层次、智能化的治理工作平台[4]。例如, 湖南红网搭建“三湘统战网”“湘商频道”等平台, 采用“专职 + 兼职”模式来广泛联系社会阶层人士, 其“时评之家”汇聚超过 2000 名评论员, 半数来自民营经济与新媒体等领域; 天津中医药大学开发 AI 治理平台, 运用自然语言处理技术实现舆情热点自动识别与快速响应; 陕西高校设计“一库一平台两端三中心”智慧治理系统, 实现治理。

数据的动态管理与成员全覆盖, 这些平台初步体现了“数智 + 治理”的转型趋势。

在风险防控层面, 各类 AI 内容检测工具和舆情监测系统已用于敏感信息识别、虚假账号追踪与预警阈值设定等领域[5]; 在价值塑造层面, 以用户画像分析、资源热力图技术助力圈层划分、内容精准推送与人力配置优化, 如武汉市委党校模型有效提升引导资源的投放效率。

在实践路径与案例积累方面, 各地涌现出一批具有代表性的创新案例。除前述平台外, 生成式 AI 如 ChatGPT 已被尝试用于政策解读与多语言宣传材料生成, 均体现出技术工具在提升治理工作效率方面的潜力。这些实践初步构建了网络治理工作的案例库, 为理论提炼提供了实证基础。

2.2. 理论困境：从“理论构建”到“技术应用”差距

尽管网络治理工作与实践层面取得了一定成效, 但从理论创新与范式转型的视角审视, 仍面临以下几大理论困境, 制约其向纵深发展:

1) 理论构建滞后：“工具理性”与“价值理性”失衡

现有研究多停留在将网络治理视为传统工作方式的“线上移植”或技术工具的简单叠加, 缺乏对“智能 + 治理”这一范式转型的系统性理论阐释[6]。研究往往侧重于“如何用技术”, 而疏于探讨技术如何重塑治理主客体关系、如何与制度规则协同、如何有效承载价值驱动使命(价值理性)。例如, 对于生成式 AI 在政策解读中的伦理边界、算法推荐对共识形成的影响机制等关键问题, 尚未形成融合政治学、传播学与人工智能科学的理论框架, 导致实践缺乏深层次理论指引。

2) 实践研究失衡：“顶层设计”与“基层落地”脱节

现有研究成果多聚焦于省市层级或技术条件较好的标杆案例(如天津、湖南的实践), 对县域、乡村等基层治理场景的关注严重不足。基层面临着网络覆盖率低、对象老龄化导致数字素养不足、方言文化多样性等特殊挑战, 而当前基于大数据、AI 的技术方案往往预设了较高的数字化基础, 适配性差, 导致“最后一公里”问题突出。理论与基层实践的脱节, 使得许多先进理念难以在广大基层落地生根。

3) 比较视角缺失：“本土特色”与“国际经验”隔阂

国内研究较少系统借鉴国外在网络政治参与、计算宣传治理等领域的研究成果与实践经验, 而国际研究又往往忽视中国治理工作的政治特色与制度语境。例如, 欧盟基于风险分级的 AI 监管模式强调平台问责与个人权利, 与我国强调治理效能与社会稳定优先的实践导向存在结构性差异; 南非的“预防 - 恢复”双轨治理框架依赖于活跃的公民社会, 而我国基层动员则依托党组织体系与行政资源。如何批判性吸收国际经验, 进行本土化改造, 构建既符合中国国情又具备国际视野的网络治理理论, 仍是当前研究的薄弱环节。

4) 风险应对薄弱：“技术防御”与“制度 - 技术协同”不足

面对当下社会生成式 AI 技术带来的深度伪造、算法偏见与认知操纵等新型风险, 现有研究提出的对策多集中于技术层面的防御(如虚假内容检测), 缺乏从制度规范、伦理准则与技术治理协同角度系统性解决方案的提出。例如, 针对俄罗斯 AI 水军网络所展示的系统性认知攻击, 仅靠单点技术难以有效应对该难题, 亟须建立涵盖数据安全、算法审计、跨境流动管理等制度 - 技术协同治理体系。理论研究的滞后使得网络治理工作面临不可控的技术安全与伦理风险。

3. 理论创新与难点突破：融合模型与双维治理机制的构建

3.1. 理论创新：构建“技术 - 制度 - 价值”三层融合的智能治理系统模型

本研究提出“技术 - 制度 - 价值”三层融合的智能治理模型(见图 1)，该模型将网络治理视为系统工程，强调技术赋能、制度适配与价值塑造的协同互构，突破了现有理论的局限性。以下对三层融合智能治理模型进行阐述：

- 1) 技术层：依托大数据、自然语言处理、生成式 AI 等技术，构建智能感知、风险预警与精准引导三大模块。通过动态画像库、情感分析和个性化内容推送，实现治理对象的精准识别与定位。
- 2) 制度层：通过建立数据共享机制、跨部门协同处理流程和安全伦理规范制度等，确保技术能够在法治轨道上建立数据共享机制、跨部门协同处置流程与安全伦理规范，保障技术发展始终遵循法治原则和规范。例如，制定《分布式 AI 治理应用数据安全共享协议》，明确各部门权责边界。
- 3) 价值层：根据公共认可度、文化认同强度、引导效能指数量化指标，评估价值层对社会认知结构的塑造成效，推动社会主义核心价值观在网络空间落地生根，达到空间网络治理的终极目标。

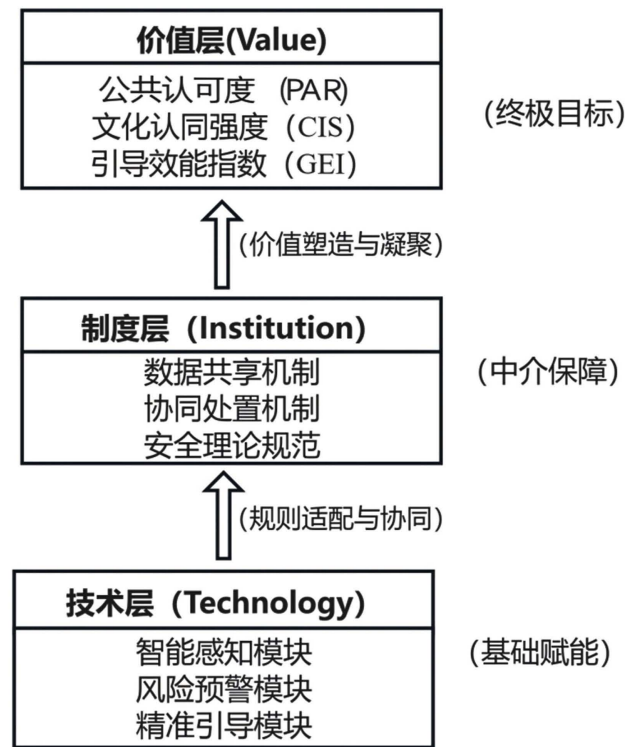


Figure 1. The three-layered integrated intelligent governance model of “technology-institution-value”
图 1. “技术 - 制度 - 价值”三层融合智能治理模型

该模型突破了“技术工具论”的局限，将治理工作从工具层面提升至范式层面，为实现“键对键”与“心连心”的深度融合提供了理论框架。

3.2. 机制创新：构建“底线安全 - 柔性引导”本土化双维治理机制

通过对欧盟、南非、俄罗斯治理模式的批判性分析，本研究提出“底线安全 - 柔性引导”双维治理框架(见图 2)：

1) 底线安全：设定算法透明度阈值，建立虚假信息分级响应机制，构建反制机器人识别模型，筑牢意识形态与技术安全防线。

2) 柔性引导：开展圈层化精准引导，赋能关键意见领袖，构建情感认同空间，以价值塑造增强文化认同与政治共识。

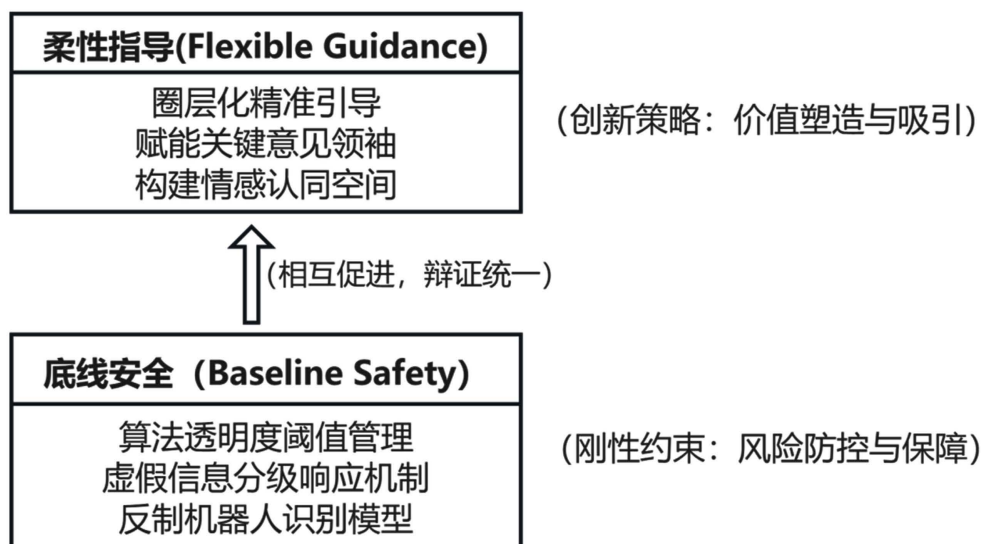


Figure 2. “Bottom-line Security-Flexible Guidance” two-dimensional governance framework

图 2. “底线安全 - 柔性引导” 双维治理框架

该框架既吸收国际规则的规范性，又立足中国政治实际情况，实现了刚性约束与柔性策略的辩证统一。

3.3. 难点突破：应对模型构建与落地过程中的核心挑战

在模型的构建与验证过程中，本研究突破了以下两大核心难点：

1) 理论构建的复杂性与量化难题

a) 难点体现：技术、制度、价值三层之间的互动关系具有高度动态性和情境依赖性，难以用传统线性模型进行刻画。特别是生成式 AI 的自主性特征与治理工作必须坚守的政治性、政策性之间存在内在张力。

b) 突破路径：对模型的核心维度进行反复迭代与操作化定义，将抽象关系转化为可观测的指标。针对 AI 的自主性问题，明确提出“人机协同、算法可控”原则，设定人机交互的伦理边界与审核流程，确保技术应用始终服务于政治目标。

2) 技术安全与伦理风险的不可控性

a) 难点体现：生成式 AI 可能被用于制造虚假信息或错误解读政策，大数据应用涉及大量敏感信息，存在隐私泄露、算法偏见乃至国家安全风险。

b) 突破路径：通过构建“底线安全 - 柔性引导”双维治理框架，在底线安全维度上，设定算法透明度阈值、虚假信息分级响应机制，并基于对抗性案例分析构建机器人的识别模型，筑牢技术安全防线。

在柔性引导维度上，强调通过圈层化传播、赋能关键意见领袖和构建情感认同空间等策略，来提升文化认同的内生动力，实现价值引导与情感共鸣的有机统一。这一框架实现了刚性约束与柔性策略的辩证统一。

证统一, 为应对不确定的技术风险提供了兼具原则性和灵活性的治理手段。

4. 价值与推广路径: 智能治理模型的内在贡献与实践转化

4.1. 实际价值: 推动智能治理工作范式转型的三大贡献

本研究构建的“技术 - 制度 - 价值”三层融合智能治理模型, 通过系统性的理论创新与机制构建, 为新时代网络治理工作的范式转型与效能提升提供了重要理论支撑, 其贡献主要体现在以下三个方面:

第一, 拓宽了网络空间治理工作的理论维度。该模型将大数据、人工智能等前沿技术深度融合在治理工作的理论机制中, 突破了将技术视为单一工具的狭窄视角, 构建了“智能赋能 - 制度适配 - 价值塑造”三位一体的理论框架。不仅丰富了数智时代网络空间治理的核心内涵, 还实现了将治理工作从“经验驱动”向“数据驱动”的范式跃迁, 为构建适应数字时代的治理理论体系奠定了坚实基础。

第二, 创新了网络空间治理工作的方法体系。模型提出技术层、制度层和价值层之间的动态关系和操作路径, 为破解了“重技术轻价值”、“重顶层轻基层”等实践难题提供了系统方法论。例如, 通过公共认可度和引导效能指数量化指标评估价值塑造成效, 通过数据共享机制和协同处置流程化制度协同, 实现了理论构建与实践方法的统一, 显著提升了治理工作的科学性、前瞻性和精准性。

第三, 强化了网络空间治理工作的复制潜能。该模型源于对多层次实践案例的深度解构与提炼, 具有鲜明的实践基因。所提出的国际经验本土化框架, 旨在解决现实工作中的实际痛点, 形成“理论指导实践 - 实践反馈理论”的良性闭环, 为各地区、各层级治理部门提供了可复制、可适配的实践指南, 有力推动了网络治理理论从学术建构向操作方案的转化。

4.2. 推广路径: 从“理论模型”到“实践转化”的策略

为促使智能治理模型在全国范围内落地生根, 实现从理论创新到普遍范式的转化, 可重点从以下三条路径协同推进:

1) 构建“理论转化与标准共建平台”

联合中央相关部门、国家级智库(如中国社会科学院、中央社会主义学院)及重点高校研究力量, 成立“智能时代网络空间治理工作创新实验室”或专项研究基地。系统梳理与完善模型的理论体系与操作指南, 牵头编制《智能治理系统建设参考架构》《网络治理数据安全与伦理规范》等标准性文件。

2) 建立“分层分类的试点推广机制”

遵循“因地制宜、循序渐进”的原则, 在全国选择不同发展水平的地区(如东部数字化先进地区、中部崛起地区、西部生态脆弱或多民族地区)及不同层级(省、市、县、乡)开展差异化试点。鼓励试点地区基于统一模型框架, 结合本地资源禀赋与核心需求进行应用创新, 例如在发达地区重点探索生成式 AI 在政策解读与国际传播中的应用, 在基层地区推广离线 APP 与“数字乡贤”模式。建立试点经验交流平台, 定期总结最佳实践, 形成可复制的“工具包”或“解决方案库”, 供其他地区对标学习。

3) 完善“政策保障与能力提升体系”

建议将智能治理能力建设纳入相关国家规划与政策文件体系, 明确发展目标与支持措施。推动中央和地方财政设立专项经费, 支持治理部门的基础设施升级、数据资源采购与复合型人才培养。联合知名高校与龙头企业, 建立“智能治理人才培训基地”, 重点提升治理干部的数据思维、技术理解能力与网络履职能力。同时, 将智能治理工作的成效纳入相关工作考核评价体系, 形成有效的激励约束机制, 确保模型的持续优化与长效运行。

5. 伦理风险及其应对策略

5.1. 隐私风险：数据过度采集与滥用

智能治理模型依赖大规模用户数据采集，可能导致个人隐私泄露。例如，用户画像分析若缺乏严格的数据脱敏机制，易引发身份信息、行为偏好等敏感数据的滥用。应对措施：建立“数据最小化”原则，明确数据采集范围与使用期限；推行匿名化处理与差分隐私技术，降低数据关联风险；完善用户授权机制，允许用户随时撤回数据使用许可。

5.2. 算法偏见：技术歧视与公平性缺失

算法模型可能因训练数据偏差或设计缺陷产生歧视性结果，如特定群体的内容推送偏差。应对措施：构建多元化训练数据集，纳入不同地域、阶层、文化背景的数据样本；建立算法审计制度，定期评估模型的公平性指标；引入第三方伦理审查，确保算法决策透明可解释。

5.3. 数据安全：系统漏洞与跨境流动风险

治理平台的集中化数据存储易成为攻击目标，跨境数据流动可能违反主权国家的数据管辖规则。应对措施：采用分布式存储与区块链技术，增强数据抗攻击能力；制定《数据跨境流动安全评估办法》，明确数据出境的安全标准；建立实时监测与应急响应机制，及时处置数据泄露事件。

5.4. 思想控制：技术赋能下的意识形态操纵

生成式 AI 与个性化推送技术可能被用于精准传播特定意识形态，形成“信息茧房”。应对措施：推行“内容多样性指数”，强制平台展示多元观点；设立“算法伦理委员会”，审查推荐算法的价值观导向；加强公众媒介素养教育，提升对技术操纵的识别能力。

5.5. 综合应对策略：制度 - 技术 - 教育协同

1) 制度层面：出台《智能治理伦理规范》，明确技术应用的红线；将伦理评估纳入项目审批流程，实行“一票否决制”。

2) 技术层面：研发“伦理增强型 AI”，内置公平性、透明性算法模块；建立“算法沙盒”机制，在可控环境中测试新技术的伦理影响。

3) 教育层面：在高校开设“技术伦理”课程，培养复合型治理人才；面向公众开展“智能时代伦理素养”普及活动，提升社会整体认知水平。

6. 结论与展望

本研究围绕“大数据驱动下的网络空间治理工作机制创新”这一核心议题，通过理论构建、实证分析与国际比较，系统探讨了在智能化时代背景下治理工作面临的机遇、挑战与转型路径。主要研究结论如下：

1) 本研究构建“技术 - 制度 - 价值”三层融合智能治理理论模型，突破了传统网络空间治理的“技术工具论”局限，将技术赋能、价值塑造和制度适配相统一，为推进数智化网络治理工作提供了系统性框架。

2) 本研究通过分析国际治理经验，提出符合我国政治特色与文化背景的“底线安全 - 柔性引导”双维治理框架，为实现网络治理工作的安全可控与价值塑造的有效平衡提供了理论参考和实践支撑。

总之，网络治理工作机制的创新是一项长期性和系统性工程。未来研究将继续坚持“以我为主、兼

容并蓄”的原则, 在技术赋能与价值塑造之间、制度规范与实践创新之间、中国特色与国际经验之间寻求动态平衡, 持续推动风险防控在网络空间深化拓展, 为实现中华民族伟大复兴凝聚更广泛、更深层次的共识与力量。

基金项目

眉山市 2025 年度统一战线研究项目“网络空间统一战线工作机制创新研究——基于大数据驱动的风险防控与价值引领”。

参考文献

- [1] 中国互联网络信息中心. 第 54 次《中国互联网络发展状况统计报告》[EB/OL]. 2024-08-29.
<https://www.cnnic.net.cn/n4/2024/0829/c88-11065.html>.
- [2] 《习近平总书记关于做好新时代党的统一战线工作的重要思想学习读本》(2025 年)系统阐释网络统战方法论[M]. 北京: 人民出版社, 2025.
- [3] 顾楚丹. 社会网络视角下社交平台社群的互动仪式链研究——以粉丝社群为例[J]. 中国青年研究, 2022(2): 37-43+29.
- [4] 李腾飞. 新质生产力与新时代统一战线的深度耦合机理与作用研究[J]. 山西社会主义学院学报, 2025(2): 47-52.
- [5] 王萌, 胡艳霞, 曹玥. 冲击与应对: 网络热点事件中网络人士统战工作研究[J]. 天津市社会主义学院学报, 2025(2): 48-51.
- [6] 钟伟军, 郭玉飞. 统合式技术适配: 数字化平台赋能基层权力监督的逻辑——基于 J 省基层监督数字化改革的实践分析[J]. 电子政务, 2025(6): 78-88.