

中国网络诚信发展报告（2026）

中国网络社会组织联合会

目 录

第一章 政策理论：网络诚信的理论体系	- 1 -
一、网络诚信的内涵演变	- 1 -
(一) 基础诚信	- 2 -
(二) 协同诚信	- 2 -
(三) 全域诚信	- 4 -
二、2025 年网络诚信建设的基本特征	- 5 -
(一) 多元协同与责任明晰相统一：多方参与、共建共享	- 5 -
(二) 虚拟治理与现实赋能相统一：线上线下、深度融合	- 7 -
(三) 法治约束与文化浸润相统一：刚性保障、柔性引领	- 8 -
(四) 动态治理与前瞻防控相统一：与时俱进、精准适配	- 10 -
三、网络诚信建设的治理理念	- 11 -
(一) 坚持党的领导，把握网络诚信建设的正确方向	- 12 -
(二) 坚持以人为本，顺应为民惠民利民的迫切需求	- 12 -
(三) 坚持良法善治，筑牢网络诚信建设的法治根基	- 12 -
(四) 坚持系统观念，构建线上线下协同的治理体系	- 13 -
(五) 坚持创新驱动，提升网络诚信建设的治理效能	- 14 -
第二章 主体实践：多元协同的实践经验	- 15 -
一、政府监管实践：精准与协同治理推进	- 15 -
(一) 治理模式创新：精准治理重点领域失信问题	- 15 -

(二) 跨域治理实践：多方协同联动强化监管合力	19 -
二、平台主体履责：治理能力与机制升级	22 -
(一) 强化主体责任，完善内部诚信规范机制	22 -
(二) 创新治理模式，提升失信行为处置效能	24 -
三、社会协同参与：多元力量赋能与联动	26 -
(一) 行业组织引领：行业自律规范的制定与推广实践	26 -
(二) 公众监督赋能：公众监督的渠道建设与实践成效	29 -
四、人工智能信任：技术风控与诚信共建	31 -
(一) 强化技术合规风控，筑牢可信运行底层支撑	32 -
(二) 深化社会信任培育，构建全域协同诚信生态	34 -
第三章 重点专题：关键领域的探索研究	36 -
一、网络谣言治理：构建风清气正的网络生态	36 -
(一) 网络谣言治理成效	36 -
(二) 网络谣言治理面临的突出问题	40 -
(三) 网络谣言治理的优化建议	44 -
二、人工智能+医疗服务：规范数字化服务的诚信底线	47 -
(一) 人工智能+互联网医疗网络诚信的政策环境	47 -
(二) 人工智能+互联网医疗的发展现状分析	52 -
(三) 人工智能+互联网医疗网络诚信的实践经验	57 -
(四) 人工智能+互联网医疗网络诚信建设的政策建议	62 -
三、互联网金融：守护金融消费的诚信生态	64 -
(一) 互联网金融诚信监管发展成效	64 -
(二) 互联网金融诚信现存突出问题	67 -

(三) 互联网金融诚信治理建议	- 74 -
四、数字文旅：夯实文旅融合的诚信基础	- 82 -
(一) 数字文旅发展成效	- 82 -
(二) 数字文旅突出问题	- 86 -
(三) 数字文旅治理建议	- 91 -
五、网络公益：筑牢公益慈善的网络诚信防线	- 96 -
(一) 网络公益发展成效	- 99 -
(二) 网络公益突出问题	- 104 -
(三) 网络公益发展治理建议	- 112 -
第四章 发展建议：提质增效的优化路径	- 115 -
一、风险挑战	- 115 -
(一) 新技术新应用衍生风险与防控难点	- 115 -
(二) 跨域网络诚信治理的协同联动难题	- 117 -
(三) 长效治理机制的短板与适配性不足	- 119 -
二、发展建议	- 121 -
(一) 规范新技术新应用管理，化解技术带来的风险	- 122 -
(二) 健全跨域联动机制，提升多维度协同治理效能	- 123 -
(三) 凝聚多方合力，促进网上网下诚信有机融合	- 125 -

第一章 政策理论：网络诚信的理论体系

伴随着我国互联网从无到有、从小到大、从大到强的发展历程，网络诚信因网而生、与网同行、因网而兴。30多年来，网络诚信建设历经初步萌芽、规范发展、综合治理三个阶段，逐步形成政治引领、以民为本、良法善治、多元共建、技术赋能的基本原则，从互联网到移动互联网，从“互联网+”到“人工智能+”，网络诚信建设的内涵外延、价值理念、规范准则等持续拓展丰富，不仅包括网络空间思想引领、文化培育、道德建设、行为规范、生态治理等方面，还包括网上网下诚信建设有机融合、互相促进，逐步发展成为多维度、立体化、多主体、全方位的综合理论体系。

一、网络诚信的内涵演变

网络诚信的内涵演变与互联网发展历程紧密相连，伴随着数字技术的迭代、网络应用场景的拓展和社会治理需求的升级，网络诚信的核心内涵从单一维度的“线上守信”逐步延伸为多维度、多层次的理论范畴，是围绕“网络空间信任构建”展开的价值规范、行为准则与实践路径的理论凝练，本质上是将传统诚信理念与网络空间的虚拟性、开放性、跨域性、即时性特征相结合，形成适配信息时代的信任体系理论，总体上可划分为三个发展阶段，呈现出从“技术约束”到“制度规范”再到“全面治理”“高质

量发展”的递进逻辑，与我国网络空间治理同频共振、同向发力。

（一）基础诚信

技术维度的基础诚信是我国互联网起步发展、网络诚信初步萌芽的重要基石。1994年4月，我国全功能接入国际互联网。在互联网萌发发展初期，具有上网条件的网民数量有限，以门户网站为代表的第一代互联网应用涌现，互联网应用以PC端信息传播、基础社交为主，主要集中于信息浏览、即时通信等基础服务，客观上决定网络空间主体相对单一，“网络诚信”尚未形成明确的概念界定，核心内涵聚焦于技术层面的守信履约，主要体现为网络主体在基础网络活动中遵守技术规则、履行基本承诺，重点是不恶意攻击网络系统、不传播明显虚假信息、不盗用他人网络账号等。这一阶段，网络诚信的外延较为狭窄，主要局限于网络技术安全和基础信息真实性，相关治理工作以防范网络违法失信行为、规范基础网络操作为主，理论探索多围绕技术监管与行为约束展开，与现实社会诚信体系的关联性较弱。

（二）协同诚信

制度维度的协同诚信是在中央网信办统筹协调下，实现从制度体系初步构建到多元协同治理的纵深推进。党的十八大以来，习近平总书记从信息化发展大势和国内国际大局出发，高度重视互联网、大力发展互联网、积极运用

互联网、有效治理互联网，2013年底中央网络安全和信息化领导小组成立（2018年改为中央网络安全和信息化委员会），为制度维度协同诚信的构建提供了组织保障。2014年《社会信用体系建设规划纲要（2014—2020年）》（国发〔2014〕21号）出台，首次将网络诚信纳入社会信用体系整体布局，^①中央网信办协同相关部门推动政策落地。2016年以来，《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》相继出台，进一步完善网络诚信法治支撑体系，形成多维度、层级化的制度框架。

伴随着网络购物、网络支付等应用深度普及，网络诈骗、虚假宣传、恶意刷单等失信问题增多，“清朗”“剑网”等专项行动深入推进，聚焦重点领域失信问题强化监管执法，推动网络诚信治理从“制度约束”向“综合治理”延伸，实现制度协同、主体协同、治理协同的有机统一。

这一时期，网络诚信的内涵集中彰显制度维度协同诚信的核心特质。前期聚焦制度规范，明确网络诚信是网络空间各类主体尊崇道德、遵守法律、履行契约的行为状态，实现线上线下诚信建设联动，形成“技术+制度”双重约束框架；后期融入社会主义核心价值观，构建“思想引领、道德约束、法律规范、技术保障、行业自律”五位一体综合体系，外延拓展至数据诚信、算法诚信等新兴领域，覆

^① 国务院. 社会信用体系建设规划纲要（2014—2020年）[EB/OL]. 2014年6月14日[2026年3月10日]. https://www.gov.cn/gongbao/content/2014/content_2711418.htm.

盖网络生产、传播、交易、消费全链条，形成个人、企业、政务、社会组织诚信协同推进格局，成为推动数字经济高质量发展、营造清朗网络空间的重要保障。

（三）全域诚信

价值引领的全域诚信是我国网络诚信建设对标中国式现代化要求的赋能升华。党的二十大明确提出“健全网络综合治理体系，推动形成良好网络生态”；^①党的二十届三中全会对健全诚信建设长效机制作出系统部署。全国网信系统深入学习贯彻党的二十届历次全会精神，贯彻党中央对健全社会信用体系、强化网络空间诚信治理作出的系统部署，推动网络诚信建设全面融入新质生产力培育、数字经济高质量发展全过程。随着人工智能等技术加速演进成熟，信息化迈向数字化、网络化、智能化全面跃升的新阶段，网络诚信建设向全域覆盖、高质量发展方向迈进，网络诚信的内涵进一步丰富完善，突出“价值引领、全域覆盖、智慧治理”核心导向，不仅涵盖传统网络领域诚信建设，更聚焦人工智能生成内容诚信、数据要素诚信等领域，以构建“守信互信、共建共享”的网络生态为目标。同时，网络诚信的外延向全域延伸，贯穿网络空间全主体、全场景、全链条，治理模式从“被动整治”向“主动预防、智慧管控”转变，党委领导、政府管理、企业履责、社会监

^① 新华社. 习近平：高举中国特色社会主义伟大旗帜 为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告[EB/OL]. 2022年10月25日[2026年3月10日]. https://www.gov.cn/xinwen/2022-10/25/content_5721685.htm.

督、网民自律的多元共治格局更加完善，网络诚信建设的规范化、制度化、智能化水平显著提升，为依法管网治网提供强有力支撑。

二、2025 年网络诚信建设的基本特征

2025 年，我国网络诚信建设围绕“主体、载体、规则、保障”四大维度展开，社会多元主体共同参与、线上线下诚信协同推进、法治道德文化多措并举，在政策理论与工作实践层面，呈现出协同性、动态性、综合性等鲜明特点。

（一）多元协同与责任明晰相统一：多方参与、共建共享

多元协同与责任明晰相统一，涵盖政府部门、行业组织协会、互联网平台企业和广大网民等多元主体，明确各主体核心职责。政府部门承担监管、引导、政策制定等相关职责。2025 年，全国网信系统开展“清朗”系列专项行动，包括整治“自媒体”发布不实信息、整治短视频恶意营销、整治 AI 技术滥用乱象、整治涉企网络“黑嘴”^①等与网络诚信建设密切相关的重点领域。其中，中央网信办举报中心持续畅通网络举报渠道，强化涉仿冒假冒网站举报受理处置工作，根据网民提供的线索，协调有关部门逐一查证、及时处置违法违规仿冒假冒网站平台 1418 个，较去年增加 1.7 倍；^②2025 年，司法机关持续高压打击网络失

^① 中国网信网. 中央网信办发布 2025 年“清朗”系列专项行动整治重点[EB/OL]. 2025 年 2 月 21 日[2026 年 3 月 11 日]. https://www.cac.gov.cn/2025-02/21/c_1741837533079135.htm.

^② 中国网信网. 2025 年网信大事回眸之网络生态治理篇——共建清朗网络空间[EB/OL]. 2025 年 12 月 31 日[2026 年 3 月 11 日]. https://www.cac.gov.cn/2025-12/31/c_1768821889918450.htm.

信行为，成效显著。全国法院新纳入失信名单 233.98 万人次，266.96 万人次完成信用修复，连续 7 个季度修复人次超新纳入人次，全年追究拒不执行判决、裁定罪 4461 人，1865.18 万被执行人迫于惩戒主动履行义务。^①同时，严惩网络平台虚假宣传、账户隐匿财产等行为，联动推进协同治理，实现失信惩戒精准化与治理科学化。行业组织协会持续发挥桥梁纽带作用，积极推动行业自律建设。中国网络社会组织联合会牵头制定/修订团体标准十余项，覆盖网络直播、人工智能、互联网账号命名等重点领域；中国互联网协会发布《推动互联网平台互通互操作自律公约》，推动平台互通互操作，规范行业诚信经营。^②互联网平台企业积极履行主体责任，在主管单位指导下，微博全年清理谣言 31 万余条、涉黄低俗信息 2498 万条，处置谣言账号 3600 余个；反诈关闭涉诈账号 635 万余个、阻断高风险对话 544 万次；人工智能内容双轨标注月均 100 余万条；^③抖音直播首创“主播健康分”机制覆盖主播量级同比提升 60%，驱动生态持续向好；人工智能大模型将违规处置效率提升 31%，全年无限期回收 37 万个违规直播账号；通过线上线下协同，抖音直播年内协助警方捣毁诈骗、色情、赌博等黑产团伙 39 个；^④此外，北京市海淀区市场监管局、海淀

^① 最高人民法院. 2025 最高法深化改革：润物细无声[EB/OL]. 2026 年 3 月 7 日[2026 年 3 月 11 日]. <https://www.court.gov.cn/zixun/xiangqing/492071.html>.

^② 中国互联网协会. 中国互联网协会发布《推动互联网平台互通互操作自律公约》[EB/OL]. 2025 年 10 月 24 日[2026 年 3 月 11 日]. <https://www.isc.org.cn/article/26712104839278592.html>.

^③ 微博. 2025 微博社区信任报告[EB/OL]. 2026 年 1 月 31 日[2026 年 3 月 11 日]. <https://weibo.com/ttarticle/p/show?id=2309405260880345956798>.

^④ 中国新闻网. 抖音发布 2025 直播治理白皮书，无限期回收直播权限账号 37 万个[EB/OL]. 2026 年 1 月

区数据局与美团、腾讯、蚂蚁、叮当 4 家试点平台签署《信用监管数据开放试点合作备忘录》，标志着全国首个向平台企业开放信用监管数据的试点工作全面推进。^①广大网民是网络诚信建设的主力军，积极参与网络乱象投诉举报。

（二）虚拟治理与现实赋能相统一：线上线下、深度融合

虚拟治理与现实赋能相统一，既立足于网络空间虚拟性特点，强化线上行为规范、失信惩戒，也联动现实社会信用体系，推动形成“线上治诚信、线下促文明”良好局面，凸显网络诚信建设的现实导向价值。网络诚信建设的载体是网络空间，网络空间行为主体是“现实人”，网络空间是现实社会的延伸，网络失信行为会影响现实社会秩序、损害社会大众利益。2025 年，全国公安机关共侦破电信网络诈骗案件 25.8 万起，会同相关部门拦截诈骗电话 36 亿次、短信 33 亿条，紧急止付涉诈资金 2170.7 万元，各地公安机关累计见面劝阻 674.7 万人次，有力遏制犯罪高发势头。^②同时，我国持续推动网络失信信息与社会信用体系对接，2025 年，市场监管总局将 1.49 万户严重失信企业列入严重失信主体名单，^③对失信主体实施限制出行、限制贷款、限制招投标等联合惩戒措施，累计 1710 万人次迫于

20 日[2026 年 3 月 11 日]. <https://xinwen.bjd.com.cn/content/s696f8a3be4b0687a289008fc.html>.

① 中国质量报. 全国首个向平台企业开放信用监管数据 试点工作全面推进[EB/OL]. 2025 年 3 月 19 日[2026 年 3 月 11 日]. https://www.samr.gov.cn/xw/df/art/2025/art_d069746f95a543cebcb2385420e2741c.html.

② 新华网. 公安部: 2025 年全国共侦破电信网络诈骗案件 25.8 万起[EB/OL]. 2026 年 1 月 9 日[2026 年 3 月 11 日]. <https://www.xinhuanet.com/law/20260109/e5ac1ba872f244d9b01f53f9935c9f45/c.html>.

③ 央广网. 市场监管总局: 2025 年将 1.49 万户严重失信企业列入“黑名单”[EB/OL]. 2026 年 2 月 5 日[2026 年 3 月 11 日]. https://news.cnr.cn/native/gd/kx/20260205/t20260205_527516881.shtml.

信用惩戒和限制消费压力，自觉主动履行生效法律文书确定的义务，或者与申请执行人协商达成和解协议，^①彰显“线上失信、线下受限”的治理力度。绍兴嵊州市聚焦网约房治理难题，构建线上线下一体联动的信用治理模式，研发“嵊里安居”线上管服平台，生成网约房专属“安居码”归集信用数据、实现“一码溯源”，线下设立镇街共治服务点、成立行业促进会，通过线上信用评级、线下差异化监管，推动网约房行业诚信发展，既体现网络虚拟治理的便捷性，实现对现实行业的诚信赋能。^②

（三）法治约束与文化浸润相统一：刚性保障、柔性引领

法治约束与文化浸润相统一，既以法律法规为刚性底线，严厉打击网络失信行为，筑牢网络诚信法治屏障，又以诚信文化为柔性引领，通过宣传教育、典型示范，让诚信理念深入人心，实现“惩失信、扬诚信”，二者相辅相成、缺一不可。刚性规则聚焦法律法规落地与精准监管，是网络诚信建设的底线保障。2025年10月28日，十四届全国人大常委会第十八次会议表决通过《全国人民代表大会常务委员会关于修改〈中华人民共和国网络安全法〉的决定》，自2026年1月1日起施行。《中华人民共和国网络安全法》第七条明确提出“国家倡导诚实守信、健康文

^① 云浮市中级人民法院. 最高人民法院发布严格区分失信、失能被执行人强化信用修复典型案例[EB/OL]. 2025年7月28日[2026年3月11日]. <https://www.yfzy.gov.cn/courtweb/web/articleInfo/view?id=115014009>.

^② 嵊州市委改革办. 绍兴嵊州：构建线上线下一体联动信用治理模式 助力共享经济新业态健康发展[EB/OL]. 2025年12月30日[2026年3月11日]. <https://mp.weixin.qq.com/s/Ryn8GdabSqmzQOCeELp5xw>.

明的网络行为,推动传播社会主义核心价值观”。截至 2025 年底,我国已逐步构建起以《中华人民共和国网络安全法》《中华人民共和国电子商务法》《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》等为核心的网络法治体系,实现网络交易、数据治理、网络传播等重点领域全覆盖。柔性引导侧重为道德规范、行业自律和文化引领,通过弘扬诚信文化、开展宣传教育、树立诚信典型,引导网络空间多元主体自觉践行诚信理念,与刚性约束形成互补,推动诚信理念深入人心。近年来,各地各部门依托网络文明大会、网络安全宣传周、诚信兴商宣传月等载体,创新传播形式、贴合群众需求,开展线上线下网络诚信宣传教育活动,覆盖网民数亿人次。比如,重庆巫溪县委网信办指导本土团队推出方言普法短视频,围绕电信诈骗、网络安全、个人信息保护等基层治理重点,以方言、本土角色和生活场景为表达载体,采用“情景再现+法律解读”叙事结构,推出《说唱网安》等系列普法短视频作品,以反转、夸张等喜剧手法,将法律知识转化为群众可感知的生活情境,实现从“送到眼前”到“走进心里”的传播转变。系列作品累计播放量破千万,抖音相关话题浏览量超 8000 万次。^①福建莆田推出诚信情景短剧及特色普法作品,其中,莆仙戏《诚为媒》以诚信为主题传递法治理念。郑州、贵阳等地打造沉浸式网络普法街区与线上诚

^① 巫溪发布. 喜报! 巫溪案例入选 2025 年度重庆市网络法治宣传十佳创新案例[EB/OL]. 2026 年 2 月 12 日 [2026 年 3 月 11 日]. <https://mp.weixin.qq.com/s/YWgnyywWE45QagiHuaGEfQ>.

信宣传平台，^①将法治知识与诚信理念融入日常场景，让网民在潜移默化中提升诚信素养。

（四）动态治理与前瞻防控相统一：与时俱进、精准适配

动态治理与前瞻防控相统一，既紧跟数字技术发展和网络应用场景拓展，动态调整治理内容和优化治理方式，破解新型网络诚信问题，预判新兴领域诚信风险，完善防控措施，做到“防患于未然”。数字技术的快速迭代（如大数据、区块链、人工智能的应用）和网络应用场景的持续拓展（如直播电商、跨境电商、元宇宙等），使得网络诚信面临的问题不断更新，网络诚信建设呈现出动态发展的鲜明特征。截至 2025 年 12 月，数字经济核心产业增加值占 GDP 比重提升至 10.5%，生成式人工智能用户规模达 6.02 亿人，^②应用场景向生活、生产领域深度渗透，网络失信治理从早期的虚假信息、网络诈骗向数据滥用、算法歧视、跨境失信等领域延伸拓展，要求网络诚信建设的内容、方式精准适配技术发展和场景变化，防范人工智能技术应用带来的失信风险。比如，北京市市场监管部门查处首例利用人工智能技术进行虚假广告宣传案件，某公司通过人工智能技术伪造央视知名主持人形象，在网络视频账号上宣传普通食品“深海多烯鱼油”具有医疗功效，违反《中

^① 央广网. 中原大地“智绘”网络普法新图景[EB/OL]. 2025 年 7 月 15 日[2026 年 3 月 11 日].
<https://www.chinacourt.cn/article/detail/2025/07/id/8903751.shtml>.

^② 新华网. 数字经济规模稳步增长[EB/OL]. 2026 年 2 月 16 日[2026 年 3 月 11 日].
<https://www.xinhuanet.com/tech/20260226/d670989d4fcb41ce9e6e252f13b17a5a/c.html>.

华人民共和国广告法》相关规定被依法处罚，监管部门同步联动平台优化人工智能内容审核机制，动态应对新型失信行为。^①2025年4月，央视《财经调查》曝光跨境保健品领域网络失信典型案件，安徽全康药业、江苏蕲臻堂等企业通过境外空壳公司伪造“海外品牌”，将国内低成本保健品包装成“美国原装进口”进行虚假宣传，一款成本不足10元的维生素售价129元，销量达3.9万件，^②相关监管部门依法查处一批涉事主体，下架清理大量违规产品，有力整治虚假宣传、跨境造假等违法失信问题，彰显动态治理实效。同时，持续完善新兴领域诚信监管规则，依托全国信用信息共享平台，归集1.8亿经营主体的807亿条信用信息，^③为新兴领域违法失信风险预判提供数据支撑，推动网络诚信建设与数字经济发展同频共振，为数字经济高质量发展筑牢诚信根基。

三、网络诚信建设的治理理念

网络诚信建设的治理理念指在网络诚信建设全过程中，指导研究、推进实践的价值导向，是坚持以人民为中心的发展思想、推进国家治理体系和治理能力现代化在网络领域的具体体现，涵盖政治引领、价值导向、治理路径、实践方法等多个层面。

^① 中国市场监管报. 北京查处首例滥用 AI 技术发布虚假广告案[EB/OL]. 2025 年 10 月 17 日[2026 年 3 月 11 日]. https://www.samr.gov.cn/xw/df/art/2025/art_021c832229aa4e07a8cf55cd3df8fdfb.html.

^② 央视《财经调查》. 假进口的保健品[EB/OL]. 2025 年 4 月 27 日[2026 年 3 月 11 日]. <https://tv.cctv.cn/2025/04/27/VIDENVv8HyraoNnsYdrOK8EQ250427.shtml>.

^③ 人民日报. 信用信息共享水平大幅提升,全国平台建成 归集经营主体信用信息超 807 亿条[EB/OL]. 2025 年 4 月 3 日[2026 年 3 月 11 日]. https://www.gov.cn/zhengce/202504/content_7017040.htm.

（一）坚持党的领导，把握网络诚信建设的正确方向

坚持党的领导是网络诚信建设的根本保证，也是我国网络诚信建设的独特优势。习近平总书记指出，“网络空间是亿万民众共同的精神家园，网络空间天朗气清、生态良好，符合人民利益”，明确要求加强党对网络工作的全面领导，必须坚持以习近平新时代中国特色社会主义思想为指导，贯彻落实习近平总书记关于网络强国的重要思想和关于精神文明建设的重要论述，把党的领导贯穿网络诚信建设全过程、各方面，确保网络诚信建设始终沿着正确政治方向推进，凝聚全党全国人民团结奋斗的共同思想基础，为网络诚信建设提供坚强的政治保障和组织保障。

（二）坚持以人为本，顺应为民惠民利民的迫切需求

网络诚信建设的根本目的是保障人民群众的合法权益，满足人民群众对清朗网络空间的迫切需求。随着互联网深度融入日常生活，网民对网络信息真实性、网络交易安全性、个人信息保护要求不断提高，网络诈骗、虚假宣传、隐私泄露等网络失信行为严重损害群众利益、影响群众获得感。因此，网络诚信建设必须坚持以人民为中心的发展思想，聚焦群众反映强烈的网络失信问题，加大整治力度，完善保障措施，不断提升广大人民群众在网络空间的获得感、幸福感、安全感。

（三）坚持良法善治，筑牢网络诚信建设的法治根基

良法善治是网络诚信建设的重要保障。《中华人民共

和国网络安全法》《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》等法律法规，为网络诚信建设提供了坚实的法治基础。坚持良法善治理念，就是要将诚信从道德倡导上升为法定原则，明确各类主体的诚信权责，细化网络失信行为的认定标准、惩戒措施，构建“立法引领、执法严格、司法保障、守法自觉”的法治体系；强化执法力度，严厉打击各类网络失信违法行为，做到有法可依、有法必依、执法必严、违法必究；注重法治和德治相结合，以法治刚性约束规范行为，以道德引领凝聚共识，实现法治与德治相辅相成、相得益彰，筑牢网络诚信建设的法治根基。

（四）坚持系统观念，构建线上线下协同的治理体系

网络诚信建设是一项系统工程，涉及主体多元、领域广泛、问题复杂，必须坚持系统观念，统筹兼顾、综合施策。习近平总书记强调，“要坚持依法治网、依法办网、依法上网，让互联网在法治轨道上健康运行。”统筹网上网下治理，推动网络诚信与现实社会诚信体系深度融合，实现信用信息共享、奖惩措施联动；统筹事前预防、事中监管、事后惩戒，构建“预防为先、监管有力、惩戒有效”的全链条治理机制。同时，统筹技术、制度、道德、文化等多种治理手段，形成“技术保障、制度规范、道德引领、文化浸润”的协同治理格局，实现网络诚信建设的系统性、整体性、协同性推进。

（五）坚持创新驱动，提升网络诚信建设的治理效能

创新驱动是提升网络诚信治理效能的关键抓手，也是适应数字经济发展、应对网络失信新问题的必然要求。坚持创新驱动理念，就是要积极运用大数据、区块链、人工智能等新技术，打造信用信息共享平台、联合辟谣平台、炒信失信信息共享平台等，提升网络诚信监管的精准度和效率；创新监管方式，推行信用分级分类监管，对守信主体予以激励、对失信主体予以约束，实现“精准治理、科学治理”；创新宣传教育方式，打造多样化、接地气的宣传载体，提升网民诚信意识和网络素养；创新制度机制，完善信用修复、信用评价等制度，推动网络诚信建设常态化、长效化。

第二章 主体实践：多元协同的实践经验

2025年，我国网络诚信建设呈现政府监管、平台履责、社会监督、网民参与的工作格局，多元主体协同发力、同向发力，逐渐形成全链条、全方位、全覆盖的治理格局。政府部门持续强化制度供给与执法力度，平台企业积极履行主体责任，持续完善内部诚信规范与机制，行业协会组织广泛凝聚多元合力，助力营造依法办网、诚信用网的良好氛围。

一、政府监管实践：精准与协同治理推进

2025年，全国各级网信、公安、市场监管等部门坚持系统观念、法治思维与底线思维，以“精准监管、协同共治、信用赋能、技术支撑”为主线，持续创新监管模式、完善制度供给、健全协同机制、强化执法力度，推动网络诚信监管从被动处置向主动预防、从分散管控向全域协同、从运动式整治向常态化治理转型，逐渐构建起权责清晰、运行高效、覆盖全面、保障有力的协同监管体系，为维护清朗网络空间、保障数字经济健康发展奠定制度保障。

（一）治理模式创新：精准治理重点领域失信问题

2025年，聚焦网络谣言、AI造假、虚假宣传、网络水军、直播带货失信、刷单炒信、恶意抹黑企业等重点领域，监管部门采取多种方式，针对账号遮蔽、主体隔离、

技术伪装、链条拆分等问题，推动形成“查主体、查内容、查资金、查链条、查责任”的治理方式，有效破解“看得见管不住、管得住罚不到、罚得到难根治”的治理难题，推动网络诚信建设向精准化、纵深化、长效化发展。

1.多部门联动开展专项行动

聚焦自媒体造谣传谣、AI生成内容滥用、网络水军操纵舆情、直播营销虚假宣传等重点领域，形成“平台责任+内容源头+账号主体”三位一体的溯源，推动网站平台落实主体责任，依法处置各类违法账号。2025年，全国网信系统依法严厉打击网络信息内容违法行为，持续加大网络安全、数据安全、个人信息保护领域违法行为查处力度，依法约谈网站平台5811家，警告1646次，对521家网站平台实施罚款处罚，责令暂停功能或更新688次，下架App2133款，下线小程序192款，会同电信主管部门注销网站、App备案，以及关闭网站、App9637家。^①

国家市场监督管理总局聚焦网络交易市场诚信建设，从建章立制、专项执法、数字赋能、信用服务四方面构建系统治理格局。先后出台网络交易平台规则、直播电商等监管规范和外卖平台服务管理、网络交易平台合规评价相关标准，推动监管向事前引导、事中防控转型。开展直播电商专项整治，公布30件典型案例，严查虚假宣传、大数据杀熟、刷单炒信等行为。完善国家网络交易监管平台，

^① 中国网信网. 2025年全国网信系统严格规范公正文明执法 护航网信事业高质量发展[EB/OL]. 2026年2月14日[2026年3月17日]. https://www.cac.gov.cn/2026-02/14/c_1772800146205770.htm.

统筹推进直播电商“一库一码一系统”建设，实现对高风险直播活动的精准识别。推广合同示范文本，创建监管与服务示范区。^①同时，持续深化网络经营信用风险分类管理，全国超过6400万户市场主体被纳入管理范畴，高风险主体问题发现率达到85%以上，^②为推动线上线下协同建设提供了强有力支撑。各地积极相应，江西落实“亮照、亮证、亮标”，上海建立“三位一体”分类监管体系，对食品等重点领域实施全环节穿透监管，有效维护网络市场诚信秩序。

公安部开展“砺剑”“净网”系列专项行动，组织全国公安机关持续开展打击整治网络谣言工作，及时发现查处借热点舆情事件进行造谣传谣线索，重拳打击编造传播涉企业、涉民生、涉警情等领域网络谣言违法犯罪活动。截止2025年7月，依法清理涉企网络谣言信息26万余条次，关停违法违规账号1.1万余个，会同网信等部门，指导本地平台网站拦截清理涉企侵权信息10万余条。^③

国家发展改革委以社会信用体系建设为支撑，推进网络主体信用信息归集、共享应用、联合奖惩，将网络造谣、虚假宣传、恶意投诉、拒不履行平台责任等失信行为全面

^① 中国市场监管报. 织密直播电商“监管网”筑牢放心消费“防火墙”——2025年市场监管部门加强直播带货乱象整治工作综述[EB/OL]. 2026年1月7日[2026年3月20日]. <http://mt.cmrn.com.cn/wap/zggsb/20260314/142032.html>.

^② 央视新闻. 将1.49万户失信企业列入“黑名单” 2025年市场监管“成绩单”出炉[EB/OL]. 2026年2月5日[2026年3月20日]. https://content-static.cctvnews.cctv.com/snow-book/index.html?item_id=17256398526174187820.

^③ 公安部. 公安机关开展打击整治网络谣言工作 全力深挖线索斩断谣言传播链条[EB/OL]. 2025年7月10日[2026年3月20日]. <https://www.mps.gov.cn/n2254098/n4904352/c10136891/content.html>.

纳入信用记录，推动形成“一处失信、处处受限”的信用约束格局。

2.多措并举应对新技术失信风险

针对 AI 换脸、语音合成、虚假图文、虚实嫁接等新技术新应用带来的风险，监管部门运用自然语言处理、计算机视觉、大数据分析、多模态检测等技术，搭建智能化监测识别系统，穿透内容表层、追溯信息源头，实现对高仿真虚假信息的快速发现、精准识别与高效处置。2025 年，中央网信办部署“清朗·整治 AI 技术滥用生成虚假信息”专项工作，对未按规定标识、伪造溯源信息、恶意编造传播的 AI 内容实施穿透处置，有效遏制技术型失信乱象扩散。^①

针对网络使用主体，打破“虚拟账号—现实主体”隔离墙，对网络账号、直播间、自媒体运营者、MCN 机构、关联企业等实施身份核验，严查虚假身份、冒用资质、关联违规、批量操控等行为。市场监管部门建立直播电商平台常态化监管会商协调机制，加强对直播电商平台的行政指导，强化直播电商领域产品质量抽检，加大直播电商领域监管执法力度，强化直播电商合规管理标准引领，持续推动直播电商行业提升标准化、规范化水平。^②

3.双管齐下开展信用合规建设

^① 中国新闻网. 中央网信办部署开展“清朗·整治 AI 技术滥用”专项行动[EB/OL]. 2025 年 4 月 30 日[2026 年 3 月 23 日]. https://www.cac.gov.cn/2025-04/30/c_1747719097461951.htm.

^② 光明日报. 市场监管总局：加强直播电商监管力度[EB/OL]. 2025 年 4 月 21 日[2026 年 3 月 23 日]. https://www.samr.gov.cn/xw/mtjj/art/2025/art_32327fe9216947e58754513ed21aace5.html.

2025年，市场监管部门推动区域信用一体化协同发展，在共建“信用京津冀”“信用长三角”“信用大湾区”的基础上，推动共建“信用成渝”“信用东北”，实现区域信息共享、监管规则互认并开展联合监管。全国6400余万户企业完成信用风险自动分类，19个省（区、市）在食品、药品等12个重点领域推行“通用+专业”分类模式。全年对63.39万户经营主体实行非现场检查，开展触发式检查22.15万户，对13.31万户经营主体实施信用容错管理，提升柔性监管效能。同时，监管部门对诚信守法主体降低检查频次、提供便利服务，对失信主体提高抽查比例、强化重点监管、限制经营活动，对严重失信主体实施公示曝光、联合惩戒、市场禁入。全年对1801.16万户经营主体开展警示提醒，指导420.11万户企业加强信用合规管理。通过“年报服务月”“信用修复服务年”等活动，组建1.8万余支年报服务队，协调解决诉求14万多个，举办宣传6700多场次、教育培训5500多场次，凝聚“知信、守信、用信、增信”的社会共识。^①

（二）跨域治理实践：多方协同联动强化监管合力

网络失信行为具有传播跨域化、违法隐蔽化等特征，需要加快构建起“跨地区、跨部门、跨层级、跨平台”四位一体的网络诚信协同监管格局，实现“信息共享、标准互认、执法联动、结果互通、联合奖惩”，大幅提升网络

^① 人民网. 信用提升三年行动“提质增效”年成效显著[EB/OL]. 2026年3月16日[2026年3月23日]. <http://finance.people.com.cn/n1/2026/0317/c1004-40683279.html>.

诚信治理整体性、协同性与有效性，形成全国一盘棋、全域一张网的协同治理体系。

1.区域协同治理成效显著

国家市场监督管理总局推进全国网络交易监管一体化建设，实现跨区域消费维权、案件移送、执法协查、信用修复协同办理，推动网络诚信监管规则全国统一、信息全国共享、执法全国联动。哈尔滨、天津、杭州、石家庄、昆明、南宁、海口七市建立全国首个跨区域信用修复协作机制，实现信用修复“七城通办”，^①为全国跨区域信用协同提供示范样本。

长三角地区沪苏浙皖四省市共同建立长三角网络销售消费品安全与召回监管常态化协作机制，包括联合开展网络销售消费品伤害信息监测、联合建设网络销售消费品召回专家库、建立完善网络销售消费品召回工作交流会商机制、联合建立网络销售消费品缺陷技术分析研判机制、长三角网络销售消费品缺陷信息应急响应联动机制等。通过统一标准、对接共享，着力破解跨区域监管难题，为全国网络销售监管提供“长三角方案”。^②京津冀三地市场监管部门联合印发《京津冀市场监管部门关于实施深化信用提升助力经营主体高质量发展三年行动方案的意见》，在推动区域协同“信用+”服务、提升经营主体信用合规能力

^① 哈尔滨广播电视台. 哈市信用修复“七城通办”案例入选“全国优秀实践”[EB/OL]. 2025年12月15日[2026年3月23日]. <https://www.hrbtv.net/folder137/2025-12-15/1109891.html>.

^② 新华网. 构建长三角网络消费安全新生态[EB/OL]. 2025年4月29日[2026年3月23日]. <http://www.zj.xinhuanet.com/20250429/36cd7a5dfdef4e57a90863572e7f754f/c.html>.

和市场竞争能力、构建京津冀信用监管一体化机制等方面持续发力。^①中部六省省会城市郑州、长沙、武汉、合肥、太原、南昌签订跨区域合作协议，聚焦“打破地域限制、凝聚监管合力”，明确从共筑认知基础、共建协作机制、共推数据共享、共促标准互认四大维度构建跨区域联动体系，全面提升网络市场监管效能。^②

2.跨部门一体化监管形成合力

网信、公安、市监、工信等多部门建立联合研判、联合监测、联合办案、联合曝光一体化执法机制，凝聚监管合力实现全链条精准打击。网信与公安部门深化协同作战，依托“清朗”“净网”“护网”系列专项行动，搭建谣言溯源、账号处置、案件侦办快速联动通道，网信部门牵头开展内容清理、平台约谈、违规账号关停等治理工作，公安机关同步跟进刑事侦查、证据固定、涉案人员抓捕，形成治理闭环，有效处置AI造假、网络暴力、舆情敲诈等各类跨部门违法案件。市场监管、网信、公安三方聚力共治直播电商失信乱象，市场监管部门牵头严查虚假宣传、价格欺诈、假冒伪劣、刷单炒信等违法行为，网信部门同步处置违规直播间及营销账号，公安机关重拳打击直播诈骗等犯罪行为，2025年查办多起跨域重大案件，罚没金额

^① 人民网. 从“数据上网”向“数据驱动决策” 京津冀市场监管迈向数智化[EB/OL]. 2025年4月23日[2026年3月23日]. <http://politics.people.com.cn/n1/2025/0423/c458474-40466036.html>.

^② 中国市场监管报. 中部六省省会城市共建网络交易协同监管机制[EB/OL]. 2025年9月18日[2026年3月23日]. https://www.samr.gov.cn/xwxc/dfdt/art/2025/art_8be6a6e113c64f86ab7008e821cb7ee9.html.

超千万元，发布四批共 30 件典型案例，^①形成强力执法震慑。

二、平台主体履责：治理能力与机制升级

2025 年，随着《人工智能生成合成内容标识办法》（国信办通字〔2025〕2 号）及配套国家标准正式实施，“清朗”系列专项行动纵深推进，网站平台聚焦网络诚信建设痛点难点，创新探索“制度+技术+共治”融合路径，在社区规则体系完善、治理机制创新、协同共治推进等方面推出一系列突破性举措，促进网站平台履行主体责任从被动响应向主动建构、从分散治理向系统施策深度转型。

（一）强化主体责任，完善内部诚信规范机制

在主管部门监督指导下，网站平台以合规建设为基础，聚焦围绕各自业务特色、行业领域等特点，持续完善内部诚信规范机制，加强社区规则、用户协议建设，推动诚信建设融入平台运营各环节。

电子商务领域，聚焦商家诚信培育与交易秩序规范，创新惠商与治理并行机制，实现商家权益保护与消费者信任建设的有机统一。淘宝天猫集中出台 12 项惠商诚信举措，重点围绕诚信体系深化，创新退货退款处置机制、开通商家快捷申诉通道，升级识别模型治理“羊毛党”，建立账号诚信

^① 中国市场监管报. 织密直播电商“监管网”筑牢放心消费“防火墙”[EB/OL]. 2026 年 3 月 14 日[2026 年 3 月 23 日]. <http://pc.cmrrn.com.cn/shtml/zggsb/20260314/142032.html>.

管控体系，将异常售后、虚假退货等行为纳入信用管理，对职业“羊毛党”依法追责，^①持续优化商家营商环境。

民生服务领域，聚焦网民群体集中、使用频次高等领域，创新服务承诺与协同管控机制。食品安全方面，2025年11月，在市场监管总局指导下，京东、美团、拼多多等8家头部平台联合签署《网络食品交易第三方平台食品安全管理自律公约》，创新建立“统一资质审核标准、统一风险监测规则、统一失信黑名单、统一联合惩戒措施”机制，填补国内跨平台网络食品诚信自律统一标准的空白。^②针对互联网医疗、在线旅游等领域，创新建立“服务承诺刚性兑现机制”，将平台公示的服务标准、价格承诺等纳入合规管控，对失信行为实行先行赔付与违约金制度。飞猪新增45个细分违规场景判罚标准，针对租车无车、司机迟到等典型问题制定专项考核细则，明确接送机迟到补偿、服务缺位赔付等刚性措施，以制度约束倒逼商家诚信履约。^③京东买药推出“不瘦包赔”诚信服务，明确约定用药效果与赔付标准，最高赔付600元并返还邮费，以刚性承诺提升用户信任，^④推动医疗服务诚信升级。各平台精准对接民生需求，以场景化诚信举措筑牢群众信任根基，让诚信成为可感知、可兑现的民生福祉。

^① 周靖杰.聚焦平台经济高质量发展,2025 淘宝天猫集中公布 12 项惠商举措[EB/OL].2025 年 1 月 20 日[2026 年 3 月 17 日].<https://www.news.cn/tech/20250120/df83cf747351430ba5f09df78202d550/c.html>.

^② 国家市场监督管理总局.市场监管总局指导 8 家平台企业发起食品安全管理自律公约[EB/OL].2025 年 11 月 7 日[2026 年 3 月 17 日].https://www.samr.gov.cn:8080/xw/zj/art/2025/art_7cd83fd88a0b41c49717eb565f5b6b05.html.

^③ 扬子晚报搜狐号.飞猪更新多条商品治理规则严管“拒单”、“到无”等典型客诉问题[EB/OL].2025 年 2 月 26 日[2025 年 3 月 18 日].https://www.sohu.com/a/863956422_121448078.

^④ 王钰淇.京东买药推出“不瘦包赔”服务助力超重量人群科学减重[EB/OL].2025 年 4 月 7 日[2026 年 3 月 18 日].<https://www.xinhuanet.com/health/20250407/6f8edb570e214b788ead792511e602d2/c.html>.

政企协同领域，主动对接监管体系，创新信用信息互通机制，政企协同从单向监管转向双向赋能，为构建全网统一的诚信治理格局奠定基础。各平台落实《关于健全社会信用体系的意见》要求，对接全国信用信息共享平台，实现违法失信主体信息与平台内部信用管理体系实时互通、动态更新，依托平台归集的 807 亿条经营主体信用信息，^①强化失信行为精准管控。严格执行《互联网平台企业涉税信息报送规定》（国务院令 第 810 号），8000 余家境内外平台完成涉税信息报送对接，有效遏制刷单炒信等虚假交易行为，^②政企协同机制的创新，实现监管与平台治理的同频共振，提升网络诚信监管效能。

（二）创新治理模式，提升失信行为处置效能

网站平台以技术赋能为核心，创新内容审核、信用评价、协同共治等机制，推动网络诚信治理从“人工辅助”向“智能精准”转变、从“单一治理”向“多元协同”转变，不断提升治理效能与响应速度。

面对网络失信行为智能化、团伙化新趋势，头部平台上线内容治理专属大模型，构建“事前预警一事中拦截一事溯源”治理体系，实现从被动处置向主动防范的跨越。快手打通内容治理大模型调控链路，通过 AI 算法全量回扫高风险商品，将风险漏放率降至 0.1%，虚假宣传举报率下降 35%，

^① 中国青年报官方网易号.全国信用信息共享平台已归集经营主体信用信息超 807 亿条[EB/OL].2025 年 4 月 2 日[2026 年 3 月 18 日].<https://c.m.163.com/news/a/JS4T94QF0514CDBK.html>.

^② 周小涵.国家税务总局：新规实施后，“刷单”虚增业绩等无序竞争减少[EB/OL].2025 年 3 月 2 日[2026 年 3 月 18 日].<https://www.qlwb.com.cn/detail/27300033.html>.

创新剧本炒作识别处置机制，实时拦截直播间诱导下单行为，实时识别直播间演绎家庭矛盾诱导下单的行为，自动触发弹窗预警、直播遮罩乃至断播处罚，2025年，炒作卖货举报率同比降低61%。^①微博完善内容审核与投诉处置机制，新增3个投诉分类、完成5项产品升级，将5分钟、30分钟投诉受理完结率分别提升至86%、95%，^②以高效响应强化诚信导向。技术赋能下的治理机制创新，使平台对失信行为的识别更精准、处置更及时、震慑更有力。

生成式人工智能的快速发展，对网络诚信建设提出新挑战。网络平台严格落实《人工智能生成合成内容标识办法》（国信办通字〔2025〕2号），出台细化规则，创新“显式+隐式”双轨标识体系，采取多种举措保障用户知情权，规范AI内容传播秩序，在保障技术创新与维护内容真实性之间寻求平衡。微博多渠道推进AI内容标识，月均标识量超100万条。^③抖音平台上线内容标识功能和AI内容元数据标识读写功能，协助创作者为AI内容添加提示标识，可识别并写入AI内容的元数据信息，为内容溯源提供技术支撑。腾讯、B站、DeepSeek等平台明确禁止篡改、隐匿标识行为，防范AI虚假内容传播。^④

^① 中国消费者报.快手电商发布2025治理体验报告[EB/OL].2026年3月13日[2026年3月18日].<https://finance.sina.com.cn/jjxw/2026-03-13/doc-inhqvamt7965607.shtml>.

^② 微博官网.2025微博社区信任报告[EB/OL].2025年1月30日[2026年3月19日].<https://www.weibo.com/ttarticle/p/show?id=2309405260880345956798>.

^③ 同上

^④ 范佳来.腾讯、抖音、快手、B站、DeepSeek等平台官宣：上线AI标识功能[EB/OL].2025年9月1日[2026年3月17日].https://www.thepaper.cn/newsDetail_forward_31523045.

网站平台修订细化社区公约与诚信准则，公开审核算法逻辑、违规处置流程与申诉渠道，创新公开透明的诚信治理机制，提升合规预期，提供清晰诚信指引。抖音健康修订《社区医疗健康公约》，拓宽诚信管控范围，明确虚假科普等失信行为处置标准并公示。^①微博以“社区信任建设”为抓手，开展规则共创、大众评议等活动，推动诚信治理公开化、民主化。

各网站平台立足主体责任，在不同领域形成各具特色的网络诚信建设路径，积极推动规则透明、过程公开、结果公正，促进平台合规诚信运行从外部约束内化为行动自觉。AI内容标识、大众评议、健康分等多种举措，有效解决部分新技术新应用带来的显性失信问题，为应对更深层次的治理挑战积累宝贵经验、奠定技术基础。

三、社会协同参与：多元力量赋能与联动

行业组织以自律规范凝聚行业共识，公众以监督力量激活社会共治，二者形成互补联动格局，为网络诚信建设注入持续动力，推动构建共建共治共享的网络诚信生态。

（一）行业组织引领：行业自律规范的制定与推广实践

制定实施团体标准与公约倡议。中国网络社会组织联合会在中央网信办有关业务局指导下，按照《优质直播间认定规范》团体标准，2025年7月，发布第四批优质直播间（主

^① 千龙网百家号. 抖音健康发布新版《社区公约》虚假营销将被重点治理[EB/OL]. 2025年3月18日[2026年3月19日]. <https://baijiahao.baidu.com/s?id=1826912161537124259&wfr=spider&for=pc>.

播) 417 个,^①工程已累计认定优质直播间(主播)共计 1164 个,进一步推动网络直播行业健康发展,提升网络主播职业素养,促进直播行业高质量发展。中国互联网协会发布《安全可信 AI 云电脑能力要求》团体标准,^②帮助企业明确安全可信 AI 云电脑的整体框架、关键能力指标,从安全可信、产品质量双维度规范产品能力,并给出能力评价的成熟度模型。

广泛开展诚信宣传教育。2025 年 4 月,北京大学经济学院、北京大学中国信用研究中心及第一亚洲商人金银业有限公司联合主办“2025 中国信用 4.16 高峰论坛”,^③紧扣“数字时代社会信用体系建设重点和方向”主题,围绕社会信用立法、信用数据确权、区块链存证技术应用、数字信用评级标准等议题展开深度研讨。2025 年 9 月,全国“诚信兴商宣传月”活动在湖北省咸宁市正式启动,中国商业联合会、中国消费者协会等 6 家全国重点行业商协会联合发布诚信倡议,^④呼吁广大企业共同营造诚实守信的市场环境。迁安市以“网络安全宣传周”为契机,围绕网络诚信建设开展系列活动,将诚信理念宣介融入校园、社区与乡村,提升社会大众网络诚信意识。^⑤

① 中国网络社会组织联合会. 优质主播培育工程第四批优质直播间(主播)名单发布[EB/OL]. [2025 年 3 月 18 日]. <https://www.cfis.cn/20250730/310f4ceb97054c2aa0786cd8b63870aa/c.html>.

② 中国互联网协会. 标准 | 中国互联网协会关于批准发布《安全可信 AI 云电脑能力要求》等 8 项团体标准的公告[EB/OL]. [2025 年 3 月 18 日]. <https://www.isc.org.cn/article/27085831921070080.html>.

③ 新华网. 2025 中国信用 4.16 高峰论坛理论研讨会举办[EB/OL]. [2026 年 3 月 19 日]. <https://www3.xinhuanet.com/city/20250415/b995384303a140929aa64a5f61ce1cd6/c.html>.

④ 信用中国. 2025 年全国“诚信兴商宣传月”在湖北咸宁启动[EB/OL]. [2026 年 3 月 19 日]. https://www.creditchina.gov.cn/xydt/hangyexiehuishanghui/202509/t20250923_469559.html.

⑤ 唐山市行政审批局. 唐山迁安市开展“增强网络安全意识,营造绿色网络环境”主题诚信宣传活动[EB/OL]. [2026 年 3 月 23 日]. <https://xy.tangshan.gov.cn/cms/content/queryOfPreview.action?contentId=8371>.

探索社会信用数字化建设。丹江口市创新构建“信用+志愿服务”模式，以“小水滴”志愿服务联合会为核心，通过完善机制、数字赋能、星级激励等举措，推动志愿服务规范化、常态化发展，为信用建设与公益服务深度融合提供了“丹江口方案”。制定《环保志愿服务激励嘉许办法(试行)》，依据服务时长和质量积分，建立“线上记录+季度审核”信用管理机制，开发“丹江口小水滴”微信小程序与环保志愿服务数字可视化系统，实现志愿者注册、签到、积分查询“一站式”操作，平台数据与省、市各级互通，高星级志愿者可享星级志愿者权益，覆盖生活、教育、医疗、就业等领域。例如，五星级志愿者可获媒体专题宣传、子女公办学校优先安排、二级医院免费体检与专家“一对一”服务，将“信用值”转化为“实在福利”。^①天津市企业诚信协会信用交通分会，联合多方探索社会化信用评价机制，开发出租汽车信用等级 AI 自动评价模型，与天津市信用“海河分”平台实现互认。司机通过二维码即可享受充电、金融、法律、生活等全方位服务，实现“一码畅享”。拓展“N个”精准服务场景，围绕司机工作生活持续扩展“信用+”应用，整合充电资源，为信用良好的新能源车辆提供最低可至5折的服务费优惠。^②

^① 湖北日报. 丹江口市：以“信用+志愿服务”绘就水源地文明新图景[EB/OL]. [2026年3月19日]. https://news.hubeidaily.net/pc/c_4534644.html.

^② 信用中国. 天津依托信用全链条服务货车、出租汽车、网约车司机[EB/OL]. [2026年3月19日]. https://www.creditchina.gov.cn/zoujinxinyong/chengxinwenhua/chengshichengxinwenhua/202602/t20260202_473997.html.

（二）公众监督赋能：公众监督的渠道建设与实践成效

公众监督赋能与权益保障。中央网信办违法和不良信息举报中心公布数据显示，2025年，中央网信办举报中心指导全国各级网信举报工作部门、主要网站平台畅通举报渠道，加大违法和不良信息受理处置力度，推动构建良好网络生态，全年受理处置网民举报线索2.23亿件，其中，中央网信办举报中心受理举报501.7万件；各地网信举报工作部门受理举报917.7万件；全国主要网站平台受理举报2.09亿件。^①比如，北京市2025年12月共受理涉企网络侵权举报信息12599件，有效举报信息8472件，有效率67.2%；北京属地重点网站受理涉企网络侵权举报信息9.9万条，其中有效举报信息5.7万条，有效率57.6%。^②2025年9月，成都市互联网违法和不良信息举报平台及属地重点网站平台受理网民举报色情、赌博、侵权、谣言等违法和不良信息345838件，同比增加269%。其中，成都市互联网违法和不良信息举报平台受理举报650件，同比增加69%；属地重点网站平台受理举报345188件，同比增加270%。^③

网站平台积极引导网民参与信用建设。知乎推出“知友探真”栏目，上线年度“知友辟谣榜”和“知友探真榜”，

^① 中央网信办举报中心. 2025年全国网络举报受理情况[EB/OL]. [2026年3月19日]. https://www.12377.cn/tzgg/2026/b2faf3b5_web.html.

^② 北京市违法和不良信息举报平台. 2025年12月北京市互联网涉企侵权举报受理情况[EB/OL]. [2026年3月19日]. https://www.bjjubao.org.cn/2026-01/05/content_47941.html.

^③ 成都网络举报. 四川省成都市互联网信息内容监管情况[EB/OL]. [2026年3月19日]. https://www.12377.cn/wxxx/2025/75b410d3_web.html.

邀请积极的内容创作者参与网络辟谣等科普创作。“维生素C能预防感冒”“越早打生长激素越好”“窗户上贴麦字能抵抗台风”等话题被证实是伪科学，登上年度“辟谣榜”；“单位体积的人产生的热量是太阳的数倍”“歌曲 Scary Monsters And Nice Sprites 频率能驱蚊”“流感特效药不要用矿泉水和牛奶送服”等看似匪夷所思的“脑洞”话题登上年度“探真榜”。^①抖音集团推出“AI 抖音求真”功能，帮助平台用户精准识别并澄清视频内容中可能存在的误导性信息与不实谣言，为广大用户营造更为安全、可靠的信息消费环境。在用户日常使用抖音浏览内容的过程中，若系统检测到某一内容存在引发公众误解的潜在风险，或涉及未经证实的信息，将自动为该内容附加“AI 抖音求真”功能入口。用户点击该入口后，即可跳转至专属“求真卡”页面，便捷获取与该事件相关的完整、准确信息，助力用户全面、客观地了解事件真相，打破信息壁垒。^②

推动全民数字素养水平提升。中央网信办、教育部、工业和信息化部、人力资源社会保障部联合印发《2025年提升全民数字素养与技能工作要点》，提出到2025年底，我国全民数字素养与技能发展水平再上新台阶，数字素养与技能培育体系基本建成，数字资源供给能力显著增强，数字人才队伍进一步壮大，劳动者数字工作能力明显提高，群体间数

^① 新浪科技. 知乎发布 2025 年度“辟谣榜”和“探真榜”[EB/OL]. [2026年3月23日]. <https://finance.sina.com.cn/tech/2025-12-31/doc-inheshqt5319420.shtml>.

^② 骆倩雯. 2025 年度“自媒体”治理与发展创新案例发布[EB/OL]. [2026年3月23日]. <https://xinwen.bjd.com.cn/content/s698eb064e4b0687a289135f3.html>.

字技能鸿沟不断缩小，数字发展环境更加普惠包容，数字生活智慧便捷，网络空间安全有序，数字法治道德伦理水平持续提升。^①2025年6月10日上午，2025年全民数字素养与技能提升月在2025年中国网络文明大会主论坛上启动。提升月期间，围绕数字生活、数字工作、数字学习、数字创新等四大场景，组织举办全民数字素养与技能提升论坛、全民数字素养与技能培训基地品牌行动、政务领域人工智能部署应用行动、师生人工智能素养提升行动、高素质复合型人才培训计划、重点群体数字惠民便民公益活动、数字教育培训资源开放共享行动等系列主题活动，进一步推动数字技术普及应用，促进数字成果普惠共享，推进人工智能全学段教育和全社会通识教育。^②

四、人工智能信任：技术风控与诚信共建

2025年，人工智能（AI）技术从“规模化应用”向“可信发展”加速转型，生成式人工智能、深度合成、边缘智能等技术持续迭代，深度渗透政务服务、医疗健康、金融服务、教育科研、工业制造、社交传播等全场景，成为驱动数字经济高质量发展的核心引擎，也是网络诚信建设的关键领域。立足2025年产业发展与网络诚信实践，我国人工智能信任体系建设快速推进，跳出单一技术可靠的局限，贯穿治理规

^① 中国网信网. 中央网信办等四部门印发《2025年提升全民数字素养与技能工作要点》[EB/OL]. [2026年3月23日]. https://www.cac.gov.cn/2025-04/27/c_1747459876424202.htm.

^② 中国网信网. 2025年全民数字素养与技能提升月启动[EB/OL]. [2026年3月23日]. https://www.cac.gov.cn/2025-06/10/c_1751286501322657.htm.

范、社会认知全链条，形成体系化建设格局，整体呈现全新发展态势。

（一）强化技术合规风控，筑牢可信运行底层支撑

技术安全与规范运行是人工智能诚信建设的基础保障。2025年，我国全面升级AI全链条技术防控与管理制度体系，完成从被动处置向主动预警、从单点管控向全域覆盖、从技术防护向伦理约束的转型，夯实智能应用诚信根基。

健全全流程安全防控体系。行业构建“源头防控、过程监测、应急处置、溯源追责”闭环管理模式，针对深度伪造、AI诈骗、算法滥用等新型风险，搭建智能化预警监测体系。^①依托多模态识别技术，企业研发系统可快速甄别AI换脸、声音克隆等伪造内容；^②金融、支付平台依托大数据算法模型，实现AI诈骗事前预警、事中拦截、事后溯源；针对具备舆论属性与公共利益属性的算法，落实24小时动态监测，及时排查算法歧视与运行失控隐患。同时，防控范围持续延伸，覆盖政务、金融、医疗等国内重点细分场景，同步推进跨境AI应用协同治理，建立信息共享与联合处置机制，实现全域风险联防联控。

完善内容全周期规范管理。从源头严控训练数据质量，督促企业规范数据采集、审核与授权流程，剔除虚假侵权内容，建立可追溯的数据管理台账。落地AI生成内容审核机

^① 网易. 阿里AI应用旗舰夸克发布新产品“深度搜索”：让AI拥有“高搜商”[EB/OL]. 2025年5月8日[2026年3月23日]. <https://m.163.com/dy/article/JV3GHSH20517R7PC.html?clickfrom=subscribe>.

^② 腾讯云官网. eKYC-AI人脸防护盾上线公告[EB/OL]. 2025年3月19日[2026年3月23日]. <https://www.tencentcloud.com/zh/document/product/1061/68553?%21editLang=zh&lang=zh>.

制，推行“机器检测+人工复核”模式，严守内容上线关口。全面落实 AI 生成内容强制标识制度，细化标识规范与溯源要求，保障公众知情权；依托区块链存证技术搭建内容监测溯源平台，实时排查清理网络虚假合成信息。强化刚性追责与行业自律，完善法律追责链条，推动行业自律公约落地实施，建立信用评价联动机制，压实企业内容管理主体责任。

推进算法与系统长效治理。优化算法备案管理体系，建立算法公平性检测指标，排查各类歧视与偏见问题，保障算法决策公正合规。常态化开展 AI 模型性能监测，搭建实时运维监测平台，动态优化模型参数，适配多场景应用需求，稳定提升决策准确率与运行可靠性。清晰划分研发企业、运营平台、使用主体的责任边界，落实分级分类责任管理制度，探索建立 AI 责任保险机制，形成权责明晰、保障完善的运行责任体系。

融入伦理价值约束规范。将科技伦理嵌入 AI 研发、应用、监管全流程，明确技术发展的法律红线与道德边界，推动“技术向善”落地，例如，蚂蚁集团研发的可验证推理技术，^①在提升金融风控 AI 安全性的同时，嵌入伦理审查模块，避免算法歧视、数据滥用等伦理风险，同时确保算法公平性。企业在风控系统、芯片研发等核心环节增设伦理审查模块，兼顾隐私保护、公平公正与安全合规；建立全生命周期伦理

^① 央广网. 做好金融“五篇大文章”的 AI 答卷：蚂蚁数科以助力金融机构打造自主可控的 AI 大脑[EB/OL]. 2026 年 3 月 5 日[2026 年 3 月 23 日]. https://tech.cnr.cn/techph/20260306/t20260306_527544721.shtml.

评估机制，对高风险 AI 应用严格准入审核，通过宣传引导培育行业伦理自觉，实现技术安全与伦理合规深度融合。

（二）深化社会信任培育，构建全域协同诚信生态

认知普及与多元共治是人工智能信任长效发展的关键。2025 年，我国聚焦不同群体需求创新普及模式，校正社会认知偏差，凝聚多方治理合力，打造适配智能时代的网络诚信共生格局。

分层推进全民素养精准普及。摒弃同质化普及模式，针对不同群体实施差异化赋能。面向青少年开展 AI 诚信与伦理校园教育，培育合规用机、坚守诚信的基本意识；面向老年人聚焦 AI 诈骗防范，通过社区宣讲、一对一指导、^①简易工具教学等方式，提升风险辨别与自我保护能力；面向农村居民结合农业智能应用、政务便民服务开展场景化科普，缩小城乡认知差距；面向企业从业者开展合规与诚信专项培训，强化行业经营自律意识。

场景化校正社会认知偏差。围绕高频应用场景开展常态化科普解读，破除认知误区。在 AI 医疗领域，通过权威专题讲解技术适用边界与诊断局限，引导公众理性看待智能辅助诊疗；^②在政务服务领域，依托智能体验中心展示 AI 服务功能与业务边界，避免公众过度依赖；在内容传播领域，普及 AI 生成内容辨别技巧与标识规范，帮助公众区分真实信

^① 人民邮电报 . 智慧守护银龄族 筑牢反诈“安全线” [EB/OL]. 2025 年 11 月 14 日[2026 年 3 月 23 日]. https://wxb.xzdw.gov.cn/wlaq/aqdt/202511/t20251114_623379.html.

^② 法治日报. AI 医疗现状调查：“生病了问 AI”，出错了怎么办？ [EB/OL]. 2025 年 4 月 1 日[2026 年 3 月 23 日]. <https://news.cctv.cn/2025/04/01/ARTIya6pL3vs8EAUf6dFBhY1250401.shtml>.

息与合成内容，平衡消解“AI 万能”与“全盘不信任”两类极端认知。

多元化拓宽信任传播渠道。整合网络传播资源，构建短视频、图文专栏、专家直播一体化科普矩阵。平台上线 AI 防骗、诚信规范、风险识别系列轻量化内容，以情景演绎、案例解析形式降低理解门槛；邀请行业专家开展线上直播答疑，实时回应公众关切问题，增强科普互动性与传播覆盖面，全方位夯实社会信任认知基础。

第三章 重点专题：关键领域的探索研究

一、网络谣言治理：构建风清气正的网络生态

2025年，新技术迭代加速推动网络谣言形态向隐蔽化、技术化、产业化演变，人工智能生成合成虚假信息、剧情摆拍编造事件等新型谣言成为治理重点。

（一）网络谣言治理成效

2025年，有关部门通过政策精准部署、专项行动攻坚、司法重拳打击、技术创新赋能、多元主体协同等多重举措，构建起“事前预防—事中管控—事后追责”治理体系，有效遏制网络谣言传播乱象，取得阶段性成效。

1. 政策体系逐渐完善，专项行动成效显著

2025年7月，中央网信办在全国开展为期2个月的“清朗·整治‘自媒体’发布不实信息”专项行动，覆盖主流社交、短视频、资讯等全平台，精准锚定自媒体这一谣言主要传播载体，划定恶意蹭炒误导公众、歪曲事实捏造信息、不标注来源以假乱真、专业领域发布不实信息四大整治重点，将AI生成合成虚假信息、剧情摆拍、刷榜操纵舆情等新型谣言纳入重点整治范围，对谣言治理靶向发力，对重点领域谣言精准打击，自媒体恶意蹭炒热点、编造虚假信息等乱象大幅减少，AI生成虚假信息、模糊标注信息来源等谣言得到有效管控，灾情、疫情、民生政策等重点

领域谣言传播频次显著下降。

公安机关依托“净网 2025”，重点查处编造传播警情灾情谣言、借热点造谣带节奏、幕后操纵网络水军和 MCN 机构批量造谣三类违法犯罪行为，形成“以打开路、以打促治”的强大震慑。据统计，内蒙古自治区全年查处网络谣言案事件 1400 余起，关停违法违规账号 71 个，其中刑事案件 8 起，行政处罚 161 人；^①烟台公安历时四个月打掉炒作新能源汽车负面信息的 AI 造谣团伙，抓获 12 人、查扣资金百万元、关停账号 8000 余个。^②各地公安机关对情节轻微者批评教育、对违法者依法处罚、对涉嫌犯罪者立案侦查，让造谣传谣的违法成本显著提升，有效遏制谣言产业化、规模化传播趋势。

2025 年，网络谣言治理呈现三大特点：一是聚焦技术型谣言治理，针对性破解 AI 滥用带来的造谣难题；二是强化源头规制，明确要求平台健全信息来源标注、专业资质认证机制，从发布流程遏制不实信息传播；三是压实平台主体责任，通过细化违规处置标准、畅通投诉举报渠道，推动平台从“被动处置”向“主动防控”转型。

2. 司法实践创新突破，法律追责体系不断完善

积极回应技术发展带来的治理难题，探索出技术滥用法律追责新路径，让网络谣言治理有法可依、有规可循。

^① 央广网. 内蒙古：今年已查处 1400 余起网络谣言案事件[EB/OL]. 2025 年 9 月 19 日[2026 年 3 月 17 日]. https://nm.cnr.cn/yaowen/20250919/t20250919_527367779.shtml.

^② 央广网. 烟台公安破获新型网络水军案 借技术炒作车企负面牟利[EB/OL]. 2025 年 12 月 18 日[2026 年 3 月 17 日]. https://news.cnr.cn/native/gd/20251218/t20251218_527464527.shtml.

北京互联网法院在相关司法解释中，首次将“故意使用生成式人工智能工具捏造事实”明确列为量刑时的从重处罚情节，为惩治 AI 造谣行为提供清晰的法律依据，填补技术型谣言治理的法律空白。

检察机关发挥公益诉讼职能，针对网络谣言侵害社会公共利益的情形开展公益诉讼，督促相关主体履行监管责任。上海检察机关在办理上海奢侈品回收店摆拍“老人捡包勒索”造谣案时，开创性地对谣言引发的“社会修复成本”进行评估，将公共资源消耗、社会信任损耗等纳入追责范围，探索出让造谣者为其行为外部社会成本“买单”的新模式。^①司法实践的创新，让网络谣言治理的法律边界更加清晰，筑牢网络诚信建设的法治防线。

3. 技术赋能提质增效，智能治理能力大幅提升

2025 年，“以 AI 对抗 AI 谣言”成为治理创新的核心方向，智能化技术在谣言识别、溯源、辟谣等环节广泛应用，大幅提升治理效能。12 月，新华网联合中国科学技术大学、支付宝共同研发的“AI 探真”智能体正式上线，作为我国辟谣领域首个中央主流媒体 AI 智能体，深度整合权威辟谣资源，聚焦食品、健康、医疗等高频谣言领域，通过智能化对话场景为用户提供科普问答服务，实现权威辟谣信息的智能化、场景化传播。^②

^① 澎湃新闻. 上海网络辟谣惩处十大案例发布, AI 已成造谣传谣主要手段[EB/OL]. 2025 年 11 月 16 日[2026 年 3 月 17 日]. <https://baijiahao.baidu.com/s?id=1849859365693179636&wfr=spider&for=pc>.

^② 新华网. 新华网“AI 探真”智能体正式上线[EB/OL]. 2025 年 12 月 3 日[2026 年 3 月 17 日]. <https://www.xinhuanet.com/government/20251203/8f02a62252ec47e498203b03863208fe/c.html>.

网络平台持续优化智能识别系统,通过自然语言处理、图像识别、大数据分析等技术构建 AI 谣言识别模型,对 AI 生成内容、摆拍视频、虚假标注信息等进行精准识别;同时,完善算法推荐机制,明确“应标未标的信息不得进入算法推荐内容池”,从传播渠道遏制不实信息扩散。2025 年,多家主流平台运用该机制成功识别并拦截多起 AI 合成虚假灾情视频、摆拍卖惨短视频、伪造政策解读信息。网信部门督促网站平台深入排查整治,依法依规处置账号 13421 个,清理违法违规信息 54.3 万余条,^①有效阻断技术型谣言的规模化传播路径。

4.多元协同格局形成,社会共治合力持续凝聚

强化“社会共治”理念,推动形成“政府监管、平台履责、行业自律和公众参与”的多元协同治理格局,凝聚全社会共同参与的强大合力。在政府与平台协同方面,网信、公安等部门与平台建立常态化沟通机制,及时通报治理重点,督导平台落实主体责任;在行业自律方面,相关行业发起“诚信上网”倡议,引导自媒体账号规范运营,拒绝发布不实信息,推动行业形成诚信传播的自律氛围。公众参与层面,中国互联网联合辟谣平台依托全国网络辟谣联动机制,整合 111 家单位资源,实现谣言信息的联动发现、查证与辟谣,同时畅通用户举报渠道,鼓励公众成为网络谣言的监督者。各地权威媒体创新辟谣形式,

^① 中国网信网. 网信部门从严整治传播无 AI 标识的虚假不实信息问题[EB/OL]. 2026 年 2 月 12 日[2026 年 3 月 17 日]. https://www.cac.gov.cn/2026-02/12/c_1772636033171974.htm.

如信网推出《编辑部的故事》科普情景剧、《你听说的 VS 我知道的》辟谣短片，以生活化场景、强对比叙事拆解谣言套路，辅以 AI 科普辟谣、图说科普、互动 H5 等多元形式，让严肃的辟谣内容告别说教感，实现权威信息的广泛触达。^①通过多元主体协同发力，“不信谣、不传谣”的社会氛围更加浓厚，公众对网络谣言的辨别意识和能力持续提升。

（二）网络谣言治理面临的突出问题

随着生成式 AI 技术的深度应用和网络传播生态的持续变化，网络谣言的生成、传播模式不断迭代升级，治理工作仍面临诸多深层次、结构性挑战，新型谣言的治理难点凸显，网络谣言治理的长期性、复杂性、艰巨性依然突出。

1. AI 技术滥用，加大谣言识别溯源与处置难度

生成式 AI 技术的普及让造谣正式进入“工业化时代”，造谣门槛大幅降低，过去需专业技术团队完成的文案编撰、影像合成等工作，如今借助免费 AI 工具输入简单指令即可快速实现，AI 全流程造假成为最具欺骗性的谣言形式。山西网民王某利用 AI 生成地震灾情描述并合成“大同地震现场”虚假视频，细节高度逼真甚至能暂时绕过部分专业机构事实核查系统，在多平台获得了数百万次播放与转

^① 信网. 2025 年度网络热点谣言态势感知 警惕 AI 滥用下的造谣乱象[EB/OL]. 2025 年 12 月 30 日[2026 年 3 月 17 日]. <https://baijiahao.baidu.com/s?id=1852886331159756711&wfr=spider&for=pc>.

发；^①AI生成的“小男孩被埋”图片与西藏地震灾情关联传播，^②高度仿真的融合化内容，专业核查人员需投入大量时间和技术资源核实，普通网民仅凭直观感受更难甄别。同时，AI生成技术持续升级，虚假信息呈现“个性化定制”“精准推送”特征，造谣者可针对老年群体编造虚假惠民政策、针对灾难事故合成虚假现场画面、针对社会热点拼接迷惑性内容，让谣言的针对性和迷惑性大幅提升。此外，AI生成内容的无痕迹性让谣言溯源难度增大，部分造谣者利用技术手段隐藏发布痕迹，执法部门难以精准锁定违法主体，给司法打击和后续处置带来诸多现实挑战。

2. 谣言传播机制变化，防控窗口期压缩

算法推荐、跨界嫁接、产业化分工等多重作用下，谣言传播速度更快、范围更广、辨识难度更高，对传统的“72小时黄金处置期”产生一定冲击，事实核查和官方回应时间大幅减少。借助同城推送算法，云南刘某峰编造“入城需缴费”谣言在2小时内覆盖当地80%以上短视频活跃用户；^③包含“政府”“灾难”“腐败”“冲突”等关键词的内容完播率、互动率和转发率平均比普通内容高出37%，算法“无意识助推”让谣言极易形成指数级传播曲线，辟谣信息短时间内难以形成同等流量传播，加剧“谣言跑赢

^① 公安部. 公安部公布打击整治网络谣言10起典型案例[EB/OL]. 2025年10月21日[2026年3月17日]. <https://www.mps.gov.cn/n2254098/n4904352/c10267710/content.html>.

^② 新华网. 向“AI生成谣言”说不，让网络空间更清朗[EB/OL]. 2025年1月14日[2026年3月17日]. <https://www.xinhuanet.com/digital/20250114/3c9a1d10e3f845a7a1c81399b8cc9307/c.html>.

^③ 公安部. 公安部公布打击整治网络谣言10起典型案例[EB/OL]. 2025年10月21日[2026年3月17日]. <https://www.mps.gov.cn/n2254098/n4904352/c10267710/content.html>.

真相”的治理困境。“跨界嫁接”成为谣言主流叙事范式，造谣者将真实元素与虚构情节刻意混合，江苏曹某某将反腐败纪律要求与方便面消费记录强行关联，利用 AI 生成格式规整的虚假“纪委通报”，^①这种“半真半假”的叙事策略大幅增加公众和平台信息核实成本。谣言与网暴的黑灰产业形成技术化、链条化分工，犯罪团伙分设信息采集、数据清洗、AI 内容生产、流量运营等专属环节，造谣传谣速度更快、范围更广、危害性更大。与此同时，权威辟谣信息的传播效能不足，难以即刻对冲谣言影响，部分辟谣内容触达率较低，部分辟谣内容缺乏生动性和趣味性，对年轻群体的吸引力不足。同时，辟谣信息发布时机往往滞后于谣言传播，部分谣言已形成广泛传播并引发公众认知偏差，辟谣信息成效不足；针对 AI 生成的虚假视频、语音等新型谣言，普通公众难以理解专业技术核查过程，对辟谣信息的接受度和信任度有待提升。

3.重点群体防范薄弱，跨域治理存在盲区

我国网民规模已突破 11 亿，在网络空间形成不同类型的网络群体，对网络谣言的辨识能力参差不齐。不同网站平台间治理协同机制尚未完善，存在数据共享不畅、处置步调不一、责任界定模糊等问题，部分造谣者利用平台间的监管差异，在一个平台发布谣言后迅速在其他平台扩散，待某一平台完成处置时，谣言已在全网形成广泛传播；不

^① 新华网. 公安部网安局公布 10 起打击整治网络违法犯罪典型案例[EB/OL]. 2025 年 9 月 18 日[2026 年 3 月 17 日]. <https://www.xinhuanet.com/legal/20250918/f3260ac764174be89ce1ed94049dad3c/c.html>.

同平台 AI 识别模型、监测系统标准不一，对同一谣言识别和处置效率存在明显差异，难以形成跨平台的协同防控合力。各地各部门信息有效互通存在一定难度，部分跨区域传播的谣言因属地管理、部门职责划分等问题，出现处置不及时等情况，对跨区域造谣团伙的打击难以形成有效联动。“一老一小”、特殊人群、边疆边远地区等重点群体的谣言防范能力薄弱，成为谣言传播的“重灾区”。老年群体对网络信息的甄别能力较弱，对 AI 生成的虚假图文、视频缺乏辨别意识，极易被虚假养生知识、虚假惠民政策等谣言误导；偏远地区网络基础设施相对薄弱，权威辟谣信息传播渠道有限，当地群众获取真实信息的途径单一，媒介素养和诚信意识亟待提升。部分青少年群体因追求新奇、缺乏理性判断能力，在性别对立、婚恋矛盾等情绪化话题中，易被谣言引发的情绪裹挟，参与非理性传播和讨论，加剧谣言扩散。

4. 过度追求流量，经济利益与社会责任难以平衡

在流量至上不良导向影响下，部分网站平台、自媒体账号、网民等成为谣言失控扩散的重要推手。个别网民为增加账号流量，随意发布、转发未经验证的信息，内蒙古赤峰“免费捡白菜”事件中，网民黄某某捡到三棵白菜后发布视频并披露位置，引发网友大量转发，成为谣言传播的起点。部分自媒体账号受流量焦虑驱动，刻意编造、炒作虚假信息，将日常婚恋、家庭矛盾等极端化、戏剧化处

理，“山东女子索要5元早餐费遭拒轻生”“新郎因彩礼跳江”等谣言，通过塑造标签化角色、放大冲突点博取关注，挤占网络空间理性探讨空间。部分基层媒体未经核实便通过官方账号发布不实信息，如“免费捡白菜”事件中当地媒体发布的相关视频，直接推动谣言全网传播；部分专业领域账号冒用医生、律师等身份发布虚假内容，杜撰案例引流带货，利用专业身份误导公众。传播主体责任失守，叠加AI造谣、虚假摆拍等行为泛滥，让“眼见为实”的社会信任体系受到冲击。比如，某教育咨询公司利用AI编造教育领域虚假文章引发师生和家长焦虑，AI生成的虚假灾情画面扰乱灾后救援秩序等等，不仅严重误导社会舆论，更让公众对网络信息真实性产生普遍质疑，在很大程度上影响公众对官方权威信息和官方机构的信任度。青岛视频账号“苏星颜”摆拍送外卖卖惨涨粉并诱导卖课等行为，^①让公众对网上求助信息产生警惕，当真正的弱势群体需要求助、真实的社会矛盾需要关注时，公众可能因惧怕被欺骗而选择观望，削弱互联网传递正能量、解决实际问题的诚信根基。

（三）网络谣言治理的优化建议

网络谣言治理是一项系统工程，事关网络空间清朗建设，事关社会公共秩序和公众合法权益，事关国家治理体系和治理能力现代化。面对技术迭代和传播生态变化带来

^① 环球网. 刻意摆拍卖惨等虚假视频 网民“苏星颜”造谣牟利被处罚[EB/OL]. 2025年9月11日[2026年3月17日]. <https://society.huanqiu.com/article/4OHfj7KwBrb>.

的挑战，需要坚持问题导向、系统思维，从技术创新、制度完善、协同共治、素养提升等多维度发力，持续完善全链条治理体系，推动网络谣言治理向精准化、高效化、长效化迈进。

1.强化技术创新，构建 AI 谣言全流程监管体系

针对 AI 技术滥用导致谣言识别难、溯源难、处置难问题，以“技术对抗技术”为核心，构建覆盖生产、传播、溯源、处置的全流程技术防控体系。优化 AI 谣言识别能力，整合自然语言处理、计算机视觉、大数据分析技术，提升对 AI 换脸、语音合成、虚假图文等高仿真内容检测与预警水平。建立 AI 生成内容强制标识与可追溯机制，推动数字水印、内容溯源等技术标准化、普及化，为锁定造谣主体、开展司法打击提供技术支撑。探索建立跨平台、跨区域谣言监测预警平台，压缩谣言传播窗口期。优化官网辟谣内容生产供给模式，推动权威信息智能生成、精准触达。

2.优化传播机制，提升辟谣效能与算法治理水平

规范平台算法推荐机制，降低争议性、情绪化、对立化内容的推荐权重，提高权威信息与辟谣信息的流量支持，推动平台建立“算法向善”机制，从传播渠道遏制谣言扩散。健全谣言快速处置机制，对虚实嫁接、恶意拼接、黑灰产造谣等行为实施全链条打击，提高造谣传谣成本。创新辟谣表达方式与传播渠道，推动辟谣内容通俗化、场景

化、短视频化，提升公众接受度与传播力。建立权威信息快速发布与同步辟谣机制，确保信息发布跑在谣言前面，有效对冲谣言影响，破解“谣言跑得快、真相跟不上”的困境。

3.健全协同治理，强化重点群体网络保护力度

强化重点群体精准防护，面向老年人、青少年、偏远地区群众开展差异化媒介素养教育与辟谣科普，拓宽权威信息触达渠道，提升薄弱环节防范能力。健全激励与监督机制，畅通举报渠道，引导公众主动参与监督，形成全社会共同抵御谣言的合力。探索构建全网联动、全域覆盖、全民参与的协同治理格局，推动平台、地区、部门间数据互通、处置联动、标准统一，消除治理断点与盲区。完善跨平台协同处置联盟，实现谣言同步监测、同步下架、同步辟谣，提升全域治理效率。

4.压实主体责任，防范线上线下谣言传导风险

以压实责任、规范行为、修复信任为主线，严格规范网民、自媒体、媒体、专业账号等各类传播主体行为，强化内容审核与真实性把关，加大编造、摆拍、炒作、转发不实信息等处置力度，健全网络诚信约束机制，将信息真实性、社会公信力纳入账号运营核心评价指标，将造谣传谣行为纳入诚信档案，实施失信联合约束，强化“失信受惩、守信获益”导向。强化权威信息供给与正面宣传引导，及时回应社会关切，依法打击损害社会信任的造谣行为，

弘扬真实理性、诚信向善的网络风尚，修复社会信任根基，营造风清气正的网络生态。

二、人工智能+医疗服务：规范数字化服务的诚信底线

互联网医疗作为数字技术与医疗服务深度融合的重要形态，已成为优化医疗资源配置、提升医疗服务可及性的重要途径。通过在线问诊、远程诊疗和健康管理等服务模式，互联网医疗有效便利了群众就医，拓展了医疗服务供给，更好满足人民群众多层次、多样化的健康需求。近年来，人工智能、大数据等新一代信息技术加速应用，推动互联网医疗服务模式持续创新，数字分身、智能健康助理等新业态不断涌现。与此同时，互联网医疗兼具医疗服务的专业属性与互联网平台的开放属性，诚信建设成为行业健康发展的重要基础。在人工智能不断与互联网医疗服务场景融合的背景下，需要统筹推进技术诚信、用户诚信和平台诚信建设，构建规范、有序、可信的发展环境，促进互联网医疗规范发展和创新应用协同推进，更好地服务人民健康与数字健康治理。

（一）人工智能+互联网医疗网络诚信的政策环境

近年来，随着人工智能技术在医疗健康领域的深化应用，我国互联网医疗网络诚信政策体系也在不断完善，逐步形成了强调发展与安全并重、创新与规范协同推进的政策体系框架，为人工智能在互联网医疗领域的应用指明方向。

1.人工智能与算法治理相关政策要求

随着生成式人工智能和算法技术的快速发展，国家层面逐步建立起覆盖算法推荐、深度合成和生成式人工智能的综合治理体系，对人工智能在互联网医疗领域的应用发挥约束效力。

在算法治理方面，2021年，国家互联网信息办公室等部门联合发布《互联网信息服务算法推荐管理规定》（国家互联网信息办公室、中华人民共和国工业和信息化部、中华人民共和国公安部、国家市场监督管理总局令 第9号），要求建立算法备案、风险评估和用户权益保护机制，防止算法歧视和不当引导，“算法推荐服务提供者应当定期审核、评估、验证算法机制机理、模型、数据和应用结果等，不得设置诱导用户沉迷、过度消费等违反法律法规或者违背伦理道德的算法模型”。

在生成式人工智能治理方面，国家互联网信息办公室于2023年发布《生成式人工智能服务管理暂行办法》（国家互联网信息办公室、中华人民共和国国家发展和改革委员会、中华人民共和国教育部、中华人民共和国科学技术部、中华人民共和国工业和信息化部、中华人民共和国公安部、国家广播电视总局令 第15号），明确生成式人工智能服务提供者应当履行内容安全责任，防止生成虚假、误导性信息。《关于促进和规范“人工智能+医疗卫生”应用发展的实施意见》（国卫办规划发〔2025〕30号）提出，要推动医疗卫生领域

大模型规范备案，优化人工智能应用审核程序；创新监管方式和预警机制，加强对人工智能研发、审评、准入、应用等各环节监管，开展应用监测评估。

在数据与安全层面，《生成式人工智能服务管理暂行办法》要求生成式人工智能服务提供者应当依法承担网络信息内容生产者责任，履行网络信息安全义务。不得收集非必要个人信息，不得非法留存能够识别使用者身份的输入信息和使用记录，不得非法向他人提供使用者的输入信息和使用记录。应遵循《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等法律、行政法规的其他有关规定和有关主管部门的相关监管要求。

《关于促进和规范“人工智能+医疗卫生”应用发展的实施意见》要求建立临床数据授权运营管理制度，制订数据安全管理和个人信息保护负面清单，建立健全智能应用数据安全防护体系，促进数据规范流通共享。

在伦理方面，2021年《新一代人工智能伦理规范》提出将伦理道德融入人工智能全生命周期，促进公平、公正、和谐、安全，避免偏见、歧视、隐私和信息泄露等问题。《生成式人工智能服务管理暂行办法》要求在算法设计、训练数据选择、模型生成和优化、提供服务等过程中，应采取有效措施防止产生民族、信仰、国别、地域、性别、年龄、职业、健康等歧视，并明确要求防范技术滥用和伦理风险，避免生成内容“危害他人身心健康”。

2.互联网医疗网络诚信相关政策要求

随着人工智能技术的快速发展，互联网医疗的网络诚信管理并非单一政策目标，而是通过医疗行为规范、信息内容治理、信用监管和数据安全制度等多重机制予以体现。

在医疗信息真实性方面，《互联网诊疗监管细则（试行）》（国卫办医发〔2022〕2号）指出“人工智能软件等不得冒用、替代医师本人提供诊疗服务”，“处方应由接诊医师本人开具，严禁使用人工智能等自动生成处方”，这些规定明确了互联网医疗网络诚信的内容底线。明确了互联网医疗“以医疗机构和医务人员为责任主体，以医疗质量安全为核心，以技术为辅助工具”的监管逻辑。在现行政策框架下，人工智能在互联网医疗中的定位只能是辅助诊疗、信息支持或管理工具，而不能脱离医疗责任主体独立运行。

在信息内容处理方面，《生成式人工智能服务管理暂行办法》要求生成式人工智能服务提供者，应当按照《互联网信息服务深度合成管理规定》“对图片、视频等生成内容进行标识”，“发现违法内容的，应当及时采取停止生成、停止传输、消除等处置措施”，“防范未成年人用户过度依赖或者沉迷生成式人工智能服务”。2025年，国家卫生健康委先后发布《关于规范“自媒体”医疗科普行为的通知》《医务人员互联网健康科普负面行为清单》（国卫办医政函〔2025〕419号），提出“不得以健康科普形

式违法违规发布各类广告、导流导诊，或通过直播带货等形式推销和销售医药产品、养生课程、保健食品等牟利”，“不得发布未经科学验证、虚假错误内容，不得断章取义曲解专业指南、行业标准等，误导公众对科学知识的理解”，“不得滥用人工智能技术，发布未经核准真实性、科学性，或未添加显著人工智能生成合成标识的健康科普内容”等十条禁令，明确了互联网健康科普信息的红线。

在信用监管层面，国务院办公厅印发《关于加快推进社会信用体系建设 构建以信用为基础的新型监管机制的指导意见》（国办发〔2019〕35号），提出将信用监管嵌入行业治理全过程。2020年3月，国家互联网信息办公室发布的《网络信息内容生态治理规定》（国家互联网信息办公室令 第5号），提出网络信息内容服务平台应当建立用户账号信用管理制度，根据用户账号的信用情况提供相应服务。2025年修订后的《互联网信息服务管理办法》中明确了互联网信息服务的处罚机制。《网络暴力信息治理规定》（国家互联网信息办公室、中华人民共和国公安部、中华人民共和国文化和旅游部、国家广播电视总局令 第17号），明确指出应将涉网络暴力信息违法违规情形记入用户信用记录，依法依规降低账号信用等级或者列入黑名单，并据以限制账号功能或者停止提供相关服务。

表 1 人工智能+互联网医疗主要政策文件

类别	政策文件名称	文号	发文单位
人工智能+医疗	《卫生健康行业人工智能应用场景参考指引》	国卫办规划函〔2024〕420号	国家卫生健康委办公厅等
	《国务院关于深入实施“人工智能+”行动的意见》	国发〔2025〕11号	国务院
	《关于促进和规范“人工智能+医疗卫生”应用发展的实施意见》	国卫办规划发〔2025〕30号	国家卫生健康委办公厅等
人工智能监管	《互联网信息服务算法推荐管理规定》	国家互联网信息办公室 中华人民共和国工业和信息化部 中华人民共和国公安部 国家市场监督管理总局令 第9号	国家互联网信息办公室等
	《生成式人工智能服务管理暂行办法》	国家互联网信息办公室令 第15号	国家互联网信息办公室等
	《互联网信息服务深度合成管理规定》	国家互联网信息办公室 中华人民共和国工业和信息化部 中华人民共和国公安部令 第12号	国家互联网信息办公室等
网络诚信监管	《医务人员互联网健康科普负面行为清单》	国卫办医政函〔2025〕419号	国家卫生健康委等
	《关于加快推进社会信用体系建设 构建以信用为基础的新型监管机制的指导意见》	国办发〔2019〕35号	国务院办公厅
	《网络信息内容生态治理规定》	国家互联网信息办公室令 第5号	国家互联网信息办公室
	《互联网信息服务管理办法》（2024年第二次修订）	中华人民共和国国务院令 第292号	中华人民共和国国务院
	《网络暴力信息治理规定》	国家网信办、公安部、文化和旅游部、广电总局令 第17号	国家互联网信息办公室等

（二）人工智能+互联网医疗的发展现状分析

1. 主要应用场景的发展现状

随着人工智能技术的发展，人工智能+互联网医疗的融合创新正迈入深度落地的新阶段，技术赋能已从医疗服务的单点环节渗透至预防、筛查、诊断、治疗、康复和管理的全链条体系，成为重构健康服务模式、优化医疗资源配置的重要驱动力。目前来看，人工智能+互联网医疗已逐步形成包

含智能辅助诊断、临床决策支持、慢病管理、健康科普咨询等在内的核心应用场景。

（1）慢病管理

在慢病管理场景中，基于人工智能赋能，结合可穿戴设备、物联网等信息技术，可构建全周期慢病管理闭环，从而推动大众从“被动就医”向“主动健康管理”转变，如联动智能穿戴设备实时监测血压、血糖等指标，自动生成个体健康画像，提供饮食、运动、用药的个性化指导等，助力实现慢病早筛、早诊、早治、早管理，有效提高患者依从性。

（2）健康科普咨询与数字分身

健康科普与数字分身应用场景的核心在于实现医学知识的规模化传播与个性化服务供给。具体而言，一是面向公众提供通俗化、连续性的健康科普服务，涵盖日常健康管理指导、小病初步识别及健康信息甄别等功能；二是通过对专家知识与沟通逻辑的建模与模拟，形成可交互的数字化服务形态（如数字人），提升优质医疗知识的可及性与服务效率。

（3）智能预问诊和流程管理

在人工智能技术支持下，智能预问诊正逐步成为互联网诊疗与线上购药的重要前置环节。其核心功能在于通过结构化问询与症状信息采集，实现对患者病情的初步识别与分诊预判，从而为后续医生诊疗、处方开具及药事服务提供决策支持。基于此，智能预问诊可有效提升分诊导诊效率，规范诊疗流程，但其本身目前并不能替代后续医疗服务环节。

此外，医疗智能体（AI Agent）是融合医学知识库与医疗数据，能够在特定医疗场景中自主完成多步骤任务（如信息采集—分析判断—行动执行）的智能代理系统。医疗智能体在健康管理、全流程诊疗管理和专科深度应用等复杂场景中均有积极的应用前景。

（4）临床决策辅助支持

人工智能辅助临床决策类应用通过整合患者病史信息、检验检查结果及循证医学证据，为医生提供结构化的诊疗参考与风险提示，从而缓解临床信息过载问题，提升诊疗决策的科学性与效率。在基层或资源相对匮乏地区，可为医生提供知识补充与决策支持，弥补经验不足带来的能力差距。

（5）继续教育与技能提升

人工智能驱动线上医学继续教育平台已覆盖多学科领域，可根据医生的专业方向、职称等提供定制化课程，推送精准学习内容。如在接入人工智能大模型后，可根据考生答题记录、学习进度和考点掌握情况，进行个性化指导和易错知识点总结归纳，强化复习效果，高效分配备考时间，帮助基层、青年医生获取前沿医学知识，缩小医疗水平差距。

2.人工智能对互联网医疗运行模式的影响

（1）服务方式转变

人工智能+互联网医疗有助于推动医疗服务从“以治病为中心”向“以健康为中心”转变，强调全生命周期健康管理，促进优质医疗资源下沉，使更多患者有机会享受

优质服务。服务模式呈现多样化发展趋势，不仅更好满足患者个体化、差异化的健康需求，也逐步形成面向医生的临床决策支持、科研辅助等专业化人工智能应用。服务边界上从单一的在线诊疗逐步拓展至健康管理、保险对接、健康产品与服务供给等多环节。

（2）医患关系转变

随着人工智能在医疗健康行业的发展布局，传统的医患关系也在逐步发生改变。人工智能作为医患沟通的桥梁，医生角色将从“单一决策者”变为“人工智能辅助下的决策者”；而人工智能赋能患者将有效提升患者的健康素养水平，使得患者对于临床决策的参与度有所提升，医患共同决策、共同干预管理的图景将有望实现。

（3）运行模式转变

在人工智能赋能下，互联网医疗平台的运行逻辑正由以流量获取为主导，逐步转向以技术能力、数据资源、供应链整合、支付方协同及多场景应用为核心的综合竞争模式。人工智能大模型及高质量数据将成为重要的核心竞争力。平台有望深度连接医院、医生、药企、保险等多方，构建医疗健康生态；同时，顺应技术的快速发展与政策的监管要求，履行平台责任、确保医疗质量安全。

3.人工智能+互联网医疗中的网络诚信风险

与其他新型互联网服务形态相同，人工智能+互联网医疗同样面临技术、用户与平台不同层面的网络诚信风险。

（1）技术诚信风险

一是数据与模型可靠性风险。人工智能医疗高度依赖数据，但现实中数据质量不均、采集标准不统一，可能影响模型训练效果；同时，大模型特别是在复杂病症和罕见病场景中仍存在误判和“幻觉”问题，可能生成缺乏医学依据的诊疗建议甚至虚假证据，带来新的医疗质量与安全隐患。

二是可解释性与责任界定风险。部分人工智能模型呈“黑箱”特征，辅助诊断与建议缺乏清晰推理链条，影响医患信任；同时，研发方、平台与医生之间责任边界尚不十分明确，难以精准追溯责任主体。

三是数据安全与隐私风险。患者敏感信息在采集、传输和存储过程中存在泄露隐患，若缺乏完善保护机制，将影响医疗安全与公众信任。

（2）用户诚信风险

一是用户认知存在偏差。部分用户过度信任人工智能，忽视专业诊断，延误治疗。二是信息失真风险。用户为获取特定服务，故意隐瞒病史、症状等关键信息，影响人工智能分析结果。三是用户面临被误导与被导流风险。部分用户由于人工智能健康素养缺失易被人工智能生成的虚假健康信息、夸大宣传误导。

（3）平台诚信风险

一是平台治理与算法规范风险。互联网医疗平台在商业

化运营过程中，可能出现算法推荐偏向付费药品或服务的情况，在一般性咨询场景中推送关联度有限的产品，影响信息呈现的客观性。同时，人工智能应用相关技术规范与伦理规则仍不够完善，对证据解读、诊疗建议生成及使用边界缺乏统一要求，可能影响服务的规范性与一致性。

二是内容审核与责任落实风险。部分平台内容审核机制仍需加强，虚假科普、夸大宣传或技术合成形象用于推广等现象时有发生，相关责任界定和追溯机制尚不够清晰，增加管理难度。

三是用户使用与能力引导风险。人工智能在互联网医疗中的应用不断扩大，但平台在医生培训、使用规范提示及公众风险告知方面仍较薄弱，部分用户对人工智能生成内容的适用范围和局限性认识不足，容易产生不当依赖。

四是运行保障与应急处置风险。人工智能应用过程中可能出现系统异常或突发情况，部分平台在预警机制、应急流程和处置规范等方面准备不足，影响问题发现和响应的及时性。

（三）人工智能+互联网医疗网络诚信的实践经验

1. 京东健康互联网医院人工智能质控提升线上服务安全

针对互联网医疗中诊疗行为不规范、质量难以监控以及责任边界不清晰等网络诚信问题，京东健康互联网医院依托人工智能技术，构建覆盖诊前、诊中与诊后的全流程质量控

制与风险管理体系，强化医疗服务的规范性与可追溯性。

（1）建立三级质控体系

京东健康互联网医院基于人工智能技术，构建诊疗行为三级质控体系，防范医疗风险。^①一级人工智能质控，利用人工智能技术实现服务过程的全面质控，对医务人员应诊时间、诊疗过程用词、问诊全面性、诊断准确性、治疗方案的一致性、药品安全性、病历规范性等进行实时监测。二级人工质控，组建专业的质控团队，采取人工复核方式对人工智能识别出的高风险订单、医师反馈的问题订单、患者投诉差评订单进行重点监测，对医师的医患交流记录、开具的处方、电子病历等进行人工复核，弥补人工智能审核的局限性。同时建立医生对质控结果异议的申诉机制，确保质控管理结果的准确性。三级专家点评，邀请资深专家制定并修改医疗质量管理相关制度，修订互联网诊疗疾病的标准化诊疗路径，建立案例库，及时通报质控动态情况。通过以上措施，持续提升质控水平，不断优化服务流程，推动医疗质量的螺旋式上升。

（2）完善线上不良事件上报与处置机制

京东健康互联网医院建立线上不良事件上报与处置机制，对涉及专业、服务及合规等方面的问题，按照严重程度分为四级管理并给予相应处置。合规类不良事件，如医生出于盈

^① 互联网医院医疗质量管理要素专家共识(2025版)[J].中国医院,2025,29(09):1-7.

利目的，引导患者至互联网医院以外医疗机构就医；对线下医生诊疗结论和建议做负向/否定评论的情况（线下诊断明显有差错的除外）。专业类不良事件，如医务人员专业问题对用户身体、心理未造成伤害的情况。服务类不良事件，如在线医生与患者无端争执或过度承诺等服务态度不良问题，或出现了平台系统问题或就诊流程问题等未能及时处理，导致用户就诊体验不佳的情况。

京东健康互联网医院对于从不同渠道上报的不良事件，设立专门的不良事件管理部门，根据事件类型组织相关部门处理。涉及医疗质量的不良事件，会酌情组织临床专家进行专业评估，必要时启动医疗质量管理委员会专家参与评定。管理部门根据不良事件严重程度及专家反馈研判是否需要干预处理，必要时同步反馈法务、宣传等部门协同处理。通过定期收集、分析、总结、改进，以提升互联网医院医疗服务质量，避免不良事件的再次发生。

2. 百度健康咨询人工智能风险管理体系应对幻觉问题

针对人工智能医疗应用中内容生成不准确、算法“幻觉”以及信息误导风险等网络诚信问题，百度健康通过构建多层审核机制与风控体系，强化人工智能生成内容的真实性与可靠性。

（1）建立三层审核模式应对大模型幻觉问题

百度健康 2025 年开展人工智能医学科普工作，联合 10

万名医生共建超 1 亿条 AIGC（人工智能生成内容）权威科普。为解决大模型在医学知识问答中存在的“幻觉问题”，百度健康设计了一套严谨的 AIGC 内容生产审核流程，即“3+3”业务模型：第一层由人工智能审核师进行 7×24 小时智能审核与巡检，对比多模型差异实现不间断监控；第二层由百度健康医学审核中心专业人员进行高时效性人工审核；第三层由近万名同科室执业医师进行专家审阅，杜绝跨科室审阅确保权威性。完整流程为 AIGC 生成内容后，经智能与人工双重审核，再提交对应科室医生终审，形成“AIGC+真人审阅”模式，使用户满意度较传统医疗搜索内容提升超 15%。2025 年推出的百度健康人工智能管家进一步强化机制，融入 36 万名专业医生标注数据训练，并在生成健康建议后自动触发真人医生实时复核与专科二次核验，通过“人工智能生成+医生校验”双重保障解决准确性痛点，目前该三层模式已覆盖百度医疗搜索 80%的结果。

（2）完善全流程风险防控体系

此外，百度健康设计了统一的风控体系，在产品上线之前设定红线机制，对涉及隐私和敏感信息的内容进行严格限制和处理，平台在制度和流程上设计告知机制，用户在使用软件时会通过勾选授权条款的方式确认数据使用规则。对于人工智能不适合直接介入的高风险问题，设定了明确的咨询边界并及时转人工处理。

3.丁香园面向医生的“可信临床外脑”应用实践

针对人工智能在辅助医生决策中的“黑箱化”、证据不透明及信任基础不足等网络诚信问题，丁香园依托长期积累的专业数据与循证医学体系，构建“可信临床外脑”，并明确提出人工智能辅助临床决策不是“显得聪明”，而是“足够诚信”，在人工智能+医疗领域更要秉承“负责任创新”原则，将安全性、有效性和合规性置于优先位置。

(1) 坚持循证优先的决策支持机制

临床医生的诊疗活动通常具有高强度与高碎片化特点。在门诊或查房等时间受限的情境下，临床实践的关键在于能否在复杂病例处理中确保决策的质量。因此，具备可靠性与可追溯性的决策支持尤为重要。丁香园构建的“临床决策人工智能”应用，强调以循证医学为基础。系统在提供建议的同时，同步呈现证据来源及出处信息，如具体指南或药品说明书版本等；当在权威数据库中无法检索到充分证据时，则明确提示“暂无相关循证支持”，以避免因模型推断而生成缺乏依据的内容，从而提升决策支持的透明性与可信度。

(2) 构建高质量数据治理体系

针对网络医疗信息来源复杂、质量参差不齐的问题，丁香园“临床决策人工智能”建立严格的数据准入机制，采用循证医学“白名单”策略，仅纳入权威学术组织发布的临床指南及监管机构批准的药品说明书等高可信度文献。同时，

建立动态更新机制，持续跟踪医学指南与临床证据的迭代进展，保障知识库内容的时效性与准确性，从源头提升人工智能输出内容的可靠性。

（3）嵌入风险预警与干预机制

在医疗安全风险较高的应用场景中，系统强化风险识别与主动干预能力。例如在药物使用环节，当存在配伍禁忌或潜在不良反应风险时，系统将优先触发风险提示，而非简单响应用户提问。同时，针对儿童、孕妇及肝肾功能异常等特殊人群，系统提供更为详细的用药机制说明及风险提示信息，通过前置预警降低潜在医疗风险。

（4）明确人机协同的责任边界

在功能边界上，丁香园“临床决策人工智能”明确提示“仅供临床医生参考”。系统输出内容以结构化报告形式呈现，涵盖机制解释、证据来源与风险提示等信息，并支持一键溯源至原始文献，便于医生进行独立判断与复核。通过强化决策透明性与可追溯性，有效缓解算法不确定性带来的信任问题，促进人机协同下的规范应用。

（四）人工智能+互联网医疗网络诚信建设的政策建议

人工智能+互联网医疗网络诚信建设应坚持“统筹发展和安全”的理念，既要为技术创新留足空间，也要为风险防范筑牢底线，推动形成包容审慎、规范有序的发展环境。

1.宏观制度层面：建立分级分类的监管体系

对高风险应用场景建立负面清单，实行最严格的伦理审查和技术评估，实施全周期动态管控，明确禁止用于互联网诊疗；对中低风险场景，设立创新试验区，探索“沙盒监管”与柔性治理模式，建立容错纠错机制，在安全可控前提下鼓励合规应用。针对互联网医疗特点，制定适配线上服务特性的风险分级标准，建立实时预警与动态监测机制，建立跨部门协同监管规则，压实平台主体责任。严禁使用未备案的生成式人工智能服务，确保应用过程公平、可解释、可追溯，保障患者知情权与数据安全。

2.平台责任层面：强化透明度与流程规范

平台应主动履行算法透明义务，在人工智能交互界面动态提示安全风险，并清晰展示决策依据，为用户提供可溯源的循证信息。针对互联网医疗应用场景，构建“智能初审+人工复核+专家终审”多层次内容审核机制，对人工智能生成内容及人脸识别等应用依法采用数字水印技术进行标识，严禁商业导向的算法推荐偏离医疗本质。健全医疗质量内控体系，设立医疗质量管理委员会，建立不良事件上报与分级处理制度，将人工智能辅助诊疗纳入平台质量考核，确保服务安全可靠。

3.用户教育层面：提升人工智能素养与树立主体责任意识

政府、行业组织及医疗机构应协同推进公众人工智能素养教育，普及人工智能的辅助定位、技术边界与潜在风险，

引导用户理性认知人工智能健康工具的适用场景与局限性，避免将其等同于专业医疗诊断。应强化“每个人是自己健康第一责任人”理念，强化用户在个人健康管理中的主体责任，在授权使用人工智能健康助手、托管健康数据时主动了解信息收集与使用规则，增强对虚假信息、人工智能换脸诈骗等风险的警惕性，树立不盲从、不轻信、主动核验的自我保护意识，切实提升在数字医疗环境中的防范能力。

4. 社会协同层面：推动多元共治网络

政府应完善法规标准，研究制定“人工智能医疗应用管理办法”，建立多部委联合监管机制。鼓励平台签署诚信自律公约，定期发布透明度报告，共享失信名单。支持行业组织制定伦理指南与服务评价标准，开展医疗人工智能产品可信评估。强化媒体监督、公众举报与第三方评估，畅通社会参与渠道，加强国际协作，推动我国标准与国际接轨，共同筑牢人工智能+互联网医疗网络诚信基石。

三、互联网金融：守护金融消费的诚信生态

（一）互联网金融诚信监管发展成效

1. 互联网金融诚信监管提质。2025年3月21日，中共中央办公厅、国务院办公厅印发《关于健全社会信用体系的意见》，首次从国家战略层面对互联网金融领域信用建设作出专门部署。^①该文件明确将平台经济信用建设纳入社会信

^① 《中共中央办公厅 国务院办公厅关于健全社会信用体系的意见》，https://www.gov.cn/gongbao/2025/iss ue_11986/202504/content_7019258.html，2025年3月21日[2026年3月22日]。

用体系核心框架，要求加强公共信用信息与平台企业经营信息共享，建立平台间失信联合约束制度，互联网金融诚信监管进入了系统治理新阶段。2月28日实施的《互联网金融个人网络消费信贷贷后催收风控指引》（GB/T 45251—2025）作为我国首个贷后催收业务国家标准，将显著降低不良贷款率，有效防范系统性金融风险。4月发布的《关于规范供应链金融业务引导供应链信息服务机构更好服务中小企业融资有关事宜的通知》（银发〔2025〕77号）要求严防对核心企业多头授信、过度授信，利用供应链“数据信用”杜绝信息中介转变为信用中介，并侵害公民个人信息的现象。7月施行的《金融产品适当性管理办法》（金融监管总局令2025年第7号）针对互联网渠道销售特点，将适当性管理嵌入线上业务流程，严禁通过操纵业绩展示误导消费者，并对65周岁以上群体销售高风险产品采取更加审慎措施；与此同时，监管科技能力实现从“人防”向“技防”的跨越式转变，监管部门构建起基于大数据和人工智能的风险监测系统，通过机器学习模型识别隐性欺诈模式，实现对跨平台信用风险的实时监测预警，^①17个省级金融基础设施平台已接入全国黑名单共享系统，失信信息归集效率大幅提升。^②

2.市场主体信用意识增强。随着失信惩戒与信用修复机制的协同发力，市场主体的诚信意识正从外部约束向内生自

^① 谢可凡. 金融科技背景下商业银行信用风险评估应用框架初探[J]. 金融, 2025, 15(4): 716-725.
<https://doi.org/10.12677/fin.2025.154076>

^② 《2025年金融监管体系升级与失信惩戒机制深化：信用约束重塑市场秩序新生态》，
<https://www.chinabgao.com/info/1273440.html>, 2025年6月10日[2026年3月24日].

觉转化。2025 年全面落地的“黑名单+市场禁入”机制，将恶意逃废债经营主体依法依规纳入严重失信主体名单，在其申请政府资金、税收优惠、股票债券发行等方面予以严格限制。^①在跨平台信用联动方面，根据《中共中央办公厅 国务院办公厅关于健全社会信用体系的意见》要求，头部平台企业开始建立平台内信用管理制度和平台间失信联合约束制度，10 月实施的网贷新规进一步要求构建基于大数据的信用评估体系，对严重失信借款人实施跨平台联合惩戒，形成“一处失信、处处受限”的约束格局。值得关注的是，监管体系创新建立的弹性信用修复框架设置了 1 年观察期和 3 年退出周期，2024 年以来已有超过 85% 的失信主体在完成合规改造后实现业务重启，这种“惩戒+教育+修复”的闭环管理模式有效引导从业机构从被动合规向主动守信转变。^②在供应链金融领域，核心企业履约意识同样显著增强，政策要求严格遵守《保障中小企业款项支付条例》（国务院令 第 728 号公布，国务院令 第 802 号修订）及时支付款项，推动从主体信用向基于真实交易数据信用转型。^③

3. 金融消费诚信维权渠道日趋畅通。金融消费诚信维权渠道在 2025 年呈现出立体化、多元化、高效化的畅通态势，大消保工作格局的正式确立成为关键节点。3 月份国家金融

^① 《中共中央办公厅 国务院办公厅关于健全社会信用体系的意见》，https://www.mee.gov.cn/zcwj/zyygwj/202504/t20250403_1106416.shtml，2025 年 4 月 3 日[2026 年 3 月 24 日]。

^② 《2025 年金融监管体系升级与失信惩戒机制深化：信用约束重塑市场秩序新生态》，<https://www.chinabgao.com/info/1273440.html>，2025 年 6 月 10 日[2026 年 3 月 24 日]。

^③ 厦门国信信用大数据创新研究院，《一起回顾 2025 年 100 个“信用时光”》，<https://www.gx-credit.com/institute/1367>，2025 年 12 月 31 日[2026 年 3 月 24 日]。

监管总局会同央行、证监会召开的金融消费者和投资者保护监管联席会议标志着三方协作机制建立，消保工作从单一维度的事后补偿转变为穿透组织架构、重塑业务全链条的系统工作，有效解决了互联网金融消费维权中跨市场、跨业态、跨区域的监管协调难题。^①在纠纷化解层面，4月份中国人民银行、金融监管总局、中国证监会联合发布的《关于推进金融纠纷调解工作高质量发展的意见》（金发〔2025〕14号）明确加强对金融领域代理维权乱象的联合整治，各地金融纠纷调解组织与互联网金融平台建立的“线上调解+司法确认”对接机制大幅提升了纠纷化解效率。伴随金融消费者权益保护服务平台的不断完善，监管部门建立的常态化教育机制针对互联网金融特点重点普及理性借贷、防范非法代理维权等知识，主要平台投诉处理时效平均缩短40%以上，满意度提升至85%以上，形成了投诉、整改、提升的良性循环，消费者的诚信维权素养与平台的合规服务意识在互动中同步提升。^②

（二）互联网金融诚信现存突出问题

1.跨平台信用联动不足。当前，互联网金融领域跨平台信用联动机制仍面临结构性障碍，数据孤岛现象依然在一定程度上存在。2025年政府工作报告明确提出“出台健全社会信用体系的政策”，并将信用建设纳入全国统一大市场建设

^① 《金融消费者权益保护报告（2025）》，<https://news.qq.com/rain/a/20250316A035ME00>，2025年3月16日[2026年3月24日]

^② 北京前沿金融监管科技研究院. 2025 消费金融消费者权益保护创新与实践报告[R].

框架，但实践中公共信用信息与市场信用信息的融合仍存在一定的壁垒。

具体而言，首先从数据供给层面看，信用监管数据作为公共产品向平台企业开放的程度有限，政府难以精准掌握平台内商家的微观行为，存在监管真空地带，而平台企业则缺乏权威的政府信用数据，容易形成劣币驱逐良币的风险。以北京市海淀区试点为例，目前开放的数据主要集中在企业经营异常名录和严重违法失信名单两类信息上，纳税、知识产权、司法判决等关键信用维度数据尚未开放，平台难以形成商户的全景式信用画像，单一的数据维度限制了信用评价的准确性及应用场景的拓展。

其次从制度层面看，跨平台信用联动的激励机制设计尚不完善。信用监管数据开放通过数据赋能提升了平台治理效率，但缺乏平台参与合作监管的长期激励机制，平台企业主动共享信用数据的意愿不足。2025年3月，国务院办公厅发布的《关于做好金融“五篇大文章”的指导意见》（国办发〔2025〕8号）提出加大公共信用信息对金融领域的开放共享力度，探索发展相关行业特色企业征信机构，但公共信用信息与金融征信信息、市场信用信息之间的共享仍存在较大困难，信息孤岛现象依然严重。目前，公共信用平台收集的负面信用信息占到了大多数，正面信息严重不足，严重影响在信用监管中的应用。这种数据分布的失衡导致跨平台信用评估难以全面反映市场主体的真实信用状况，制约了信用联

动机机制的有效运行。

最后技术层面的障碍同样不容忽视。多平台间由于商业竞争、技术壁垒和隐私保护政策，导致数据难以共享，形成了各自独立的信息孤岛，这种现象严重阻碍了信用评估的全面性和准确性。虽然联邦学习等技术手段为跨平台数据共享提供了可能，但在实际应用中，数据标准化程度低、接口不统一、安全认证机制缺失等问题仍制约着技术效能的发挥。截至 2025 年底，市场监管部门累计归集经营主体各类信用信息 154.6 亿条，实现跨部门共享数据 122.7 亿条，但共享数据的时效性、完整性和准确性仍需提升，跨平台信用联动的实际效果与政策预期之间仍存在差距。^①

2.隐性失信行为难以管控。隐性失信行为的隐蔽性和复杂性对现行监管体系构成严峻挑战。这类行为往往披着合法合规的外衣，通过技术手段和制度漏洞规避监管，传统的事前审批和事后处罚模式难以有效应对。在供应链金融领域，虚构贸易背景套取资金现象频发，以承兴系案件为例，核心企业通过伪造应收账款凭证骗取资金逾 300 亿元，暴露了信用传导链条的脆弱性。2025 年 2 月，六部门联合发布的《关于规范供应链金融业务 引导供应链信息服务机构更好服务中小企业融资有关事宜的通知》（银发〔2025〕77 号）明确指出，应收账款电子凭证业务潜藏的风险包括相关业务开展情况处于监管盲区、核心企业存在信用过度扩张风险、供

^① 《国新办就市场监管服务经济高质量发展情况举行新闻发布会》，
http://www.china.com.cn/zhibo/content_118314860.htm，2026 年 2 月 5 日[2026 年 3 月 24 日]。

应链信息服务系统缺乏必要的设立标准和管理规范等。

技术依赖性风险的相生相伴加剧了隐性失信行为的识别难度。数字化平台在提升效率的同时，也成为了风险的集聚点和放大器，高度自动化的流程意味着一旦系统出现漏洞、算法存在偏见或数据被污染，其后果将被迅速、大规模地复制。2025年3月 OpenAI 推出的第三代金融预测模型在部署仅两周后突然失灵，多家投行依据其低风险评级重仓东南亚新兴科技股，结果遭遇当地政策突变引发的股灾，模型核心问题在于风险因子折叠缺陷，将政治动荡、汇率波动等非线性关联变量压缩为单一概率值，无法识别复杂关联。这种算法黑箱问题在互联网金融领域同样存在，金融机构对技术的依赖越深，技术本身的安全、稳定与可靠就越成为整个体系的薄弱环节。

多头借贷与以贷养贷行为的隐蔽化趋势值得警惕。年轻用户平均消费信贷负债达 1.8 万元，30% 负债超月收入 5 倍，2025 年消费金融行业年轻用户不良贷款率达 2.8%，较 2020 年上升 0.6 个百分点，以贷养贷用户不良率高达 8.5%。部分平台将借贷入口隐藏在领优惠券、免息分期、打车立减等界面背后，用户在不知情的情况下被贷款，这种行为涉嫌侵犯消费者的知情权和自主选择权。部分平台通过拆分费用、多头收费的方式规避监管，以增信服务名义收取担保费、服务费、咨询费，导致实际年化利率远超 36%，而监管红线明确

规定客户的综合融资成本不能高于 24%。^①

供应链金融中的信用风险传导机制进一步增加了隐性失信行为的管控难度。传统供应链金融高度依赖核心企业的信用背书，一旦核心企业因经营不善、财务造假或宏观调控等因素出现信用爆雷，其信用风险将迅速沿供应链向上下游大量中小微企业传导，引发连锁违约，形成系统性雪崩效应。信用风险扩散呈现系统性特征，核心企业财务波动会通过供应链票据迅速传导，如某车企商票逾期导致上游百家供应商融资成本骤升，核心企业信用风险扩散可能引发涟漪效应。这种链式传导与风险共振现象使得隐性失信行为的影响范围远超单一主体，传统以机构为单位的风险监测模式难以捕捉跨机构、跨市场的风险传染路径。

数据真实性核验依然是风控的薄弱环节。虚构贸易背景、伪造电子单据等欺诈行为，在数字化外衣的包装下更具迷惑性，对金融机构的穿透式核验能力提出了更高要求。尽管 2025 年监管政策要求加强贸易背景真实性管理，但面对海量交易数据和复杂的关联交易网络，金融机构现有的技术手段和人力资源仍显不足。特别是在跨境电商、平台经济等新业态领域，交易链条长、参与主体多、数据标准不一，隐性失信行为的识别和追踪面临更大挑战。

3.部分主体诚信履约意识薄弱。部分市场主体的诚信履约意识薄弱是互联网金融领域长期存在的突出问题。2025 年

^① 湖南日报，《互联网“借贷入口”，无孔不入，无处不在》，2025 年 3 月 22 日[2026 年 3 月 24 日]，<https://m.voc.com.cn/xhn/news/202603/31797740.html>。

政府工作报告明确提出加力推进清理拖欠企业账款工作，强化源头治理和失信惩戒，落实解决拖欠企业账款问题长效机制，这从侧面反映出拖欠账款问题的普遍性和严重性。在供应链金融领域，核心企业利用其谈判中的优势地位，在交货、价格、账期等贸易条件上强迫中小企业采取有利于自己的行为，导致中小企业资金紧张，一旦中小企业获得银行融资，供应链资金紧张状况得以缓解，核心企业可能进一步挤占中小企业资金，如果由此积累的债务负担超出中小企业的承债极限，核心企业就会由控制风险的变量转变为供应链系统性风险的震源。

借款人履约意识不足的现象在网络贷款领域表现突出。2025年网贷逾期现象频繁发生，根本原因包括经济压力加剧、财务规划不当、借贷门槛过低、高利率问题等。不少借款人在借贷过程中对利息、还款周期缺乏清晰认知，导致无力还款，部分借款人因侥幸心理逃避还款，造成征信不良。尽管2025年10月实施的网贷新规要求建立网贷黑名单制度，纳入失信被执行人名单，实施跨平台联合惩戒，同时限制高消费行为，但恶意逃废债行为仍时有发生。蚂蚁消金2025年初转让的5.81亿元不良贷款被全部归类为损失级，借款人平均年龄为38岁，部分贷款的逾期时长已经超过1000天，根据行业内的调查显示，超60%的消费贷违约者因收入减少无法还款，25%的借款人则因企业经营困难而受到影响。

中小微企业信用资质的内生性脆弱加剧了履约风险。中

小微企业普遍存在财务制度不健全、轻资产运营、缺乏合格抵押物、信用记录不完整、抗风险能力弱等问题，其信用资质的内生性脆弱，导致其在传统信贷体系中难以获得融资，即便在供应链金融场景下，其对外部环境波动的抵抗力也较弱，应收账款周转天数延长、订单被取消、上游原材料价格上涨等问题都可能迅速引发其资金链断裂。2025年供应链金融新规特别禁止核心企业强制接受非现金支付变相延长付款期限，侧面印证该问题的普遍性。核心企业若迟延支付供应企业款项，不仅需要支付逾期利息，还应当将逾期尚未支付中小供应商款项的合同数量、金额等信息纳入企业年度报告，依法通过国家企业信用信息公示系统向社会公示，拖欠中小供应企业款项的企业将可能受到县级以上人民政府及其有关部门的监督检查、函询约谈、督办通报、投诉处理等，依法依规被认定为失信企业的，其有关失信情况还将可能被记入其信用记录。

部分平台企业的合规意识不足也是诚信履约问题的重要表现。2025年1月发布的《小额贷款公司监督管理暂行办法》（金规〔2024〕26号）明确禁止小额贷款公司捆绑销售或附加不合理条件、将贷款列为支付结算的默认选项、诱导过度负债和多头借贷、以违法或不正当手段催收等，但实践中仍有平台通过欺诈或引人误解的方式进行营销宣传，片面宣传低门槛、低利率、高额度等，诱导借款人过度负债。个别平台将贷款列为默认支付选项，违反借款人意愿搭售商品、

服务或附加其他不合理条件，这些行为不仅损害了消费者权益，也破坏了市场的公平竞争秩序。2026年春节前夕，国家金融监督管理总局联合市场监管总局、中国人民银行对六家出行平台企业进行约谈，指出出行平台与金融机构合作借贷业务中存在的误导性营销、信息披露不充分、消费者权益保护缺失等三大核心问题。

履约意识薄弱问题的根源在于失信成本与收益的不对称。尽管2025年监管体系建立了黑名单共享机制，已有17个省级金融基础设施平台接入该系统，并设置了1年观察期和3年退出周期的弹性信用修复框架，但部分主体仍抱有侥幸心理，认为违约成本低于履约成本。部分企业将拖欠账款、逃废债务作为缓解资金压力的手段，缺乏对供应链整体利益和行业信用生态的考量。这种短期行为不仅损害了交易对手的利益，也通过供应链的链式传导机制放大了系统性风险，最终可能反噬自身。因此，构建有效的诚信履约激励约束体系，提高违约成本，强化信用约束，仍是互联网金融监管的重要课题。

（三）互联网金融诚信治理建议

1.搭建跨平台信用共享机制。破解跨平台信用联动不足的困境，需要从制度设计、技术架构和激励机制三个维度协同发力，构建覆盖政府、平台、金融机构的立体化信用共享体系。在制度层面，应加快落实2025年3月中共中央办公厅、国务院办公厅印发的《关于健全社会信用体系的意见》

中关于平台经济信用建设的部署，明确公共信用信息向平台企业开放的标准、范围和程序，建立跨部门、跨层级、跨地域的信用信息归集共享机制。具体而言，可借鉴北京市海淀区试点的经验，在确保数据安全的前提下，逐步扩大纳税、知识产权、司法判决等关键信用维度数据的开放范围，丰富平台企业信用评价的数据基础，同时建立数据开放的动态调整机制，根据平台企业的合规状况和信用管理水平，实施差异化的数据开放策略。

在数据融合机制建设方面，应着力打通公共信用信息、金融征信信息与市场信用信息之间的壁垒。2025年2月国家公共信用和地理空间信息中心与10家全国性银行机构合作共建的“国家信用信息归集加工联合实验室”提供了有益探索，应在此基础上进一步拓展参与主体，吸纳头部互联网平台、征信机构、数据服务商等多元主体参与，打造信用领域可信数据空间。针对多头借贷与隐性负债识别难题，可建立个人信用信息的统一归集平台，整合各互联网金融平台的借贷记录、还款情况、逾期信息等关键数据，形成完整的个人信用画像，为金融机构信贷决策提供全面参考。同时，应建立数据共享的负面清单制度，明确禁止共享的数据类型和情形，在促进数据流动与保护个人隐私、商业秘密之间寻求平衡。

技术架构的优化是跨平台信用共享的重要支撑。应加快制定信用数据的技术标准和交换规范，统一数据格式、接口

协议和安全认证机制，降低平台间数据对接的技术成本。可探索运用隐私计算、联邦学习等技术手段，在数据不出域的前提下实现信用信息的可用不可见，既保护数据安全，又释放数据价值。同时，建立信用信息共享的实时更新机制，确保各平台获取的信用信息具有时效性和准确性，避免因信息滞后导致的信用评估偏差。2025年市场监管部门累计归集的154.6亿条经营主体信用数据为跨平台共享提供了丰富的数据资源，关键在于建立高效的数据清洗、加工和分发机制，提升数据质量。

激励机制的设计直接影响平台企业参与信用共享的积极性。应建立信用共享的正向激励和负向约束相结合的制度安排，对主动共享高质量信用数据、积极参与联合惩戒的平台企业，在监管评级、业务准入、政策扶持等方面给予倾斜；对数据共享不积极、故意隐瞒关键信用信息的平台，采取监管约谈、限制业务扩张等措施。同时，可探索建立信用数据贡献度评价体系，根据平台企业提供数据的数量、质量、更新频率等指标，给予相应的数据使用权益，形成数据贡献与数据获取相匹配的良性循环。此外，应明确平台企业在信用共享中的权责边界，建立数据使用的行为规范和监督机制，防止平台利用准监管权打压竞争对手或形成数据垄断。

在供应链金融等特定领域，可建立核心企业信用信息的强制披露制度。2025年六部委发布的供应链金融新规已要求核心企业及时支付中小企业款项并公示相关信息，可在此基

础上进一步要求核心企业定期披露应付账款规模、账期结构、逾期情况等关键指标，通过信息公开约束其行为。同时，建立供应链上下游企业信用信息的联动共享机制，将核心企业的信用状况与其上下游企业的融资可得性、融资成本挂钩，形成信用风险的链式管理，防止核心企业信用风险沿供应链传导扩散。

2.强化隐性失信排查。针对隐性失信行为的隐蔽性和复杂性，应构建以技术赋能、穿透监管和协同联动为核心的排查治理体系，提升对隐性风险的识别、预警和处置能力。在技术赋能方面，应充分运用大数据、人工智能等监管科技手段，建立覆盖交易背景、资金流向、关联关系的多维度风险监测模型。可借鉴2025年3月OpenAI金融预测模型失灵事件的教训，在算法设计中避免简单压缩复杂变量，注重识别政治、经济、社会等非线性关联因素，提高模型对隐性风险的捕捉能力。具体而言，应建立贸易背景真实性的智能核验系统，整合工商登记、税务发票、物流信息、海关数据等多源信息，通过交叉验证识别虚构交易、伪造单据等欺诈行为。同时，开发资金流向的实时追踪工具，监测贷款资金的实际用途，防止资金被挪用或违规流入限制性领域。

穿透式监管是应对隐性失信行为的关键方法。应打破机构监管的传统思维，建立以功能和行为为核心的穿透式监管框架，透过复杂的交易结构和产品设计，识别业务实质和风险本质。在供应链金融领域，应强化对核心企业信用扩张的

约束，建立核心企业债务融资总额的监测预警机制，防止其通过多头开票、重复融资等方式过度加杠杆。针对平台经济中隐蔽的借贷入口和捆绑销售行为，应要求平台以显著方式披露借贷产品的关键信息，包括资金方、利率、期限、还款方式等，禁止将借贷设置为默认选项或隐藏于其他服务流程中。同时，建立算法审计制度，对平台使用的信用评估算法、推荐算法进行定期审查，评估其是否存在歧视性、误导性或诱导过度负债的倾向，确保算法应用的公平性和透明度。

协同联动机制的建设有助于形成隐性失信排查的合力。应加强金融监管部门与市场监管、公安、司法、网信等部门的协作，建立跨部门的信息共享和联合执法机制，提升对非法代理维权、职业反催收、恶意逃废债等黑灰产业链条的打击效能。2025年4月中国人民银行、金融监管总局、中国证监会联合发布的《关于推进金融纠纷调解工作高质量发展的意见》（金发〔2025〕14号），为部门协同提供了制度基础，应在此基础上进一步细化操作流程，明确各部门在风险监测、线索移送、案件查处、司法衔接等环节的职责分工。同时，建立区域间的监管协调机制，针对跨地区、跨平台的隐性失信行为，打破地域分割和监管壁垒，实现风险的协同处置。

行业自律与社会监督是隐性失信排查的重要补充。应支持行业协会制定隐性失信行为的识别标准和自律公约，建立会员单位的信用信息共享平台，对存在隐性失信行为的会员实施行业惩戒。鼓励第三方评级机构、会计师事务所、律师

事务所等专业机构参与隐性失信行为的排查，发挥其在财务审计、法律合规、风险评估等方面的专业优势。同时，完善举报奖励制度，激励知情者举报虚构贸易背景、伪造信用信息、非法代理维权等隐性失信行为，拓宽风险线索的来源渠道。2025年“清朗浦江·2025”网络金融信息治理行动的经验表明，社会监督和舆论监督在揭露隐性失信行为方面具有独特作用，应建立常态化的网络金融信息监测机制，及时发现和处置误导性宣传、虚假信息等风险隐患。

在制度层面，应完善隐性失信行为的法律界定和责任追究机制。针对现有法律法规对隐性欺诈、隐蔽性违规等行为界定模糊的问题，可通过司法解释、监管指引等方式明确认定标准，提高法律适用的确定性。同时，加大对隐性失信行为的处罚力度，提高违法成本，形成有效的威慑。对于技术性、隐蔽性较强的失信行为，应建立过错推定和责任倒置规则，减轻受害方的举证负担，强化行为主体的合规责任。

3.健全诚信履约激励约束体系。构建诚信履约的激励约束体系，需要从信用评价、激励引导、约束惩戒和修复救济四个环节入手，形成全周期、多层次的信用治理闭环。在信用评价环节，应建立科学合理的信用评分模型，将履约记录、还款意愿、合规表现等纳入评价指标，区分短期违约与长期失信、客观困难与恶意逃废债等不同情形，提高信用评价的精准性和区分度。2025年10月实施的网贷新规要求建立网贷黑名单制度，纳入失信被执行人名单，这一制度设计应与

信用评分体系有机衔接，对不同程度的失信行为实施差异化的信用惩戒。同时，应建立信用评价的动态调整机制，根据市场主体的信用行为变化及时更新信用等级，既体现失信惩戒的严肃性，又给予信用修复的机会。

激励引导机制的建设应注重正向激励与负向约束相结合。对于信用状况良好、履约记录优秀的市场主体，应在融资可得性、融资成本、业务准入等方面给予实质性优惠。具体而言，可建立守信主体的白名单制度，对纳入白名单的借款人提供利率优惠、额度提升、审批简化等便利，对守信的平台企业在监管评级、业务创新、政策支持等方面给予倾斜。在供应链金融领域，可将核心企业的履约记录与其上下游企业的融资支持挂钩，对及时支付中小企业款项的核心企业，鼓励金融机构为其供应链上下游企业提供更加便利的融资服务，形成守信受益的示范效应。同时，应加强诚信文化的宣传教育，通过典型案例宣传、信用知识普及等方式，提升市场主体的诚信意识和履约自觉，营造守信光荣、失信可耻的社会氛围。

约束惩戒机制应着力提高失信成本，确保惩戒措施与失信行为的性质、情节、影响相匹配。对于恶意逃废债、欺诈借贷等严重失信行为，应依法实施联合惩戒，在政府采购、招投标、行政审批、资质认定等方面予以限制，并纳入社会信用体系进行公示。2025年公安部与金融监管总局联合开展的金融“黑灰产”集群打击专项行动，对“职业反催收代理”

“非法代理退保理赔”等行为形成有力震慑，应将此类专项行动的经验做法制度化、常态化，保持对严重失信行为的高压态势。同时，应完善失信惩戒的程序保障，明确惩戒的适用条件、实施程序和救济途径，防止惩戒权的滥用，保护市场主体的合法权益。

信用修复机制是激励约束体系的重要组成部分，应体现惩前毖后、治病救人的原则。2025年监管体系建立的弹性信用修复框架，设置1年观察期和3年退出周期，为失信主体提供了改过自新的机会。应进一步完善信用修复的具体标准和操作流程，明确修复的条件、方式和时限，对主动纠正失信行为、积极消除不良影响的主体，及时开展信用修复，恢复其正常的信用等级。同时，建立信用修复的跟踪评估机制，对修复后的主体进行持续监测，防止其再次失信。对于因客观原因导致的短期违约，如自然灾害、突发疾病等，应建立容错机制，给予一定的宽限期和缓冲期，避免将其简单归为失信行为。

在供应链金融等特定领域，应针对核心企业与中小企业的不同特点，设计差异化的激励约束措施。对于核心企业，应强化其履约责任，建立应付账款的信息披露和监测预警制度，对拖延支付、强制非现金支付等行为实施约束，情节严重的纳入失信记录。对于中小企业，应注重培育其信用意识，通过信用培训、融资辅导等方式，帮助其建立规范的财务制度和信用管理制度，提升其在供应链金融中的信用可得性。

同时，建立供应链信用风险的共担机制，鼓励核心企业、金融机构、担保机构等共同参与，分散单一主体的信用风险，增强供应链整体的抗风险能力。

此外，应加强对平台企业履行中介责任的激励约束。平台作为连接借款人与出借人、核心企业与上下游企业的中介，其合规经营和诚信履约对整个生态具有重要影响。应建立平台企业的信用评价和分级分类监管制度，对合规意识强、风控能力好的平台给予业务创新和拓展的支持，对合规意识薄弱、风险事件频发的平台实施严格的监管措施，直至退出市场。通过强化平台责任，倒逼其加强对平台内参与主体的信用管理，形成平台自治与政府监管相结合的治理格局。

四、数字文旅：夯实文旅融合的诚信基础

数字文旅是数字技术与文化旅游产业深度融合的产物，是文旅产业高质量发展的重要方向。随着数字技术快速普及，文旅线上线下融合持续深化、文旅智慧服务覆盖面不断扩大，为游客提供了更加便捷、更加个性的服务与文旅体验。与此同时，数字文旅发展中虚假宣传频发、消费者维权困难等问题不断出现，破坏了文旅市场诚信生态，阻碍了文旅产业高质量发展。

（一）数字文旅发展成效

我国数字文旅产业在政策支持、技术创新和区域协同的推动下，正逐步形成以数字化技术为核心驱动力的高质量发

展格局，^①在国家层面，以《“十四五”旅游业发展规划》《“数字中国”建设整体布局规划》为核心纲领，明确提出推动数字技术与文旅产业深度融合、加快智慧文旅基础设施建设、强化文旅市场诚信监管、构建数字文旅信用体系等核心任务，为全国数字文旅产业发展划定战略方向、明确实施目标。省级层面，各地结合自身文旅资源禀赋与产业发展实际，出台针对性实施细则与配套政策，如河南省公布《河南省“十四五”文化旅游融合发展规划》，明确提出了实施数字化转型战略，将“数字先行”作为重要方针，布局元宇宙新赛道，建设智慧旅游体系，推动国家顶层设计在区域层面精准落地。地市层面作为政策执行的关键环节，立足本地文旅发展特色，将国家及省级政策导向转化为具体实践举措，因地制宜推动数字文旅项目落地，形成各具特色的区域发展模式。个体层面，景区、文旅企业、科技企业等市场主体主动响应政策导向，加大数字技术研发与应用投入，创新文旅服务形态与运营模式，成为数字文旅产业发展的重要载体。网络诚信理论及其制度体系是数字文旅的理论基础和制度机制基础，脱离了网络诚信理论指导及其制度体系的支撑，数字文旅难以持续和健康发展。

1. 文旅业线上线下融合不断深化

(1) 数字技术赋能传统文旅场景

数字技术为传统文旅场景注入新活力，推动产业转型升级升

^① 贺小荣、陈文灏. 中国数字经济与文旅产业融合的协同关系及其影响因素 [J]. 热带地理, 2024 年 (第 7 期).

级。国家大剧院打造“第二现场”超高清数字演艺新场景，依托超高清直播技术，采用 AI 自动追踪、多机位调度及 SRT 传输协议，实现远程一体化管控与分级信号推送，联动全国多个省份 20 余家剧院、百余家影院同步直播，创新舞台艺术数字演艺形式与文化传播模式。^①作为河南省文旅数字化创新标杆，“只有河南·戏剧幻城”在 2023 年入选国家首批“5G+智慧旅游”应用试点后，进一步深化智慧旅游建设。该项目通过“5G+智慧旅游”应用及声光电画、VR、AR 等科技手段打造沉浸式观剧体验，幻城剧场升降台、328 米夯土墙巨幅动态投影，用数字技术生动展现千年历史画卷，园区整合停车购票功能，实现“一部手机游幻城”。目前，该园区已实现全区域 5G 信号、无线 Wi-Fi、视频监控全覆盖，开城以来累计接待观剧人次达 5000 万，成为数字技术与文旅融合的标杆项目。^②

（2）线上线下一体化服务模式创新

各地积极探索线上线下一体化服务模式，借助数字技术整合文旅资源，为游客提供一站式便捷服务。敦煌莫高窟数字展示中心依托数字采集、高清投影、虚拟漫游等技术，对传统石窟文旅场景进行数字化赋能，构建“数字莫高”线上线下融合体系，实现洞窟文物的数字化永久保存与沉浸式展示。游客可通过 VR 设备“走进”未开放洞窟，近距离观察

^① 北京市文化和旅游局. 2025 年北京市文化和旅游科技创新优秀案例[EB/OL]. 2025 年 9 月[2026 年 3 月 13 日]. <https://whlyj.beijing.gov.cn/zwgk/tzgg/202509/P020250909599358474217.pdf>.

^② 河南省文化和旅游厅. 以“数”赋能 向“新”聚力——河南省文化和旅游数字化创新示范案例[EB/OL]. 2025 年 6 月 24 日[2026 年 3 月 13 日]. <https://hct.henan.gov.cn/2025/06-24/3173259.html>.

壁画细节，既有效保护了文物本体，又显著提升了游览体验，每年通过数字展示中心接待游客超百万人次，成功缓解了文物保护与旅游开发之间的矛盾，成为服务模式创新的重要实践。^①福建省文旅数智融合应用平台汇聚游前、游中、游后各类多源异构数据，整合 POI 空间数据、地理信息数据、文旅资源普查等基础数据，结合 OTA 价格数据、舆情数据、天气预报等准实时数据，以及景区闸机、交通路况、视频监控与无人机巡检等实时数据，建成跨界融合、字段统一的公共大数据平台，有效解决了省、市、县三级政府用户、旅行社和游客的数据采集、共享、预测难题。^②

2.智慧服务覆盖面拓宽

(1) 智能导览与个性化推荐服务

智能导览与个性化推荐已成为数字文旅的核心服务，为游客提供更便捷、多元的个性化体验，同时通过技术手段规范服务内容，明确服务标准，保障游客的合法权益。

贵州省文化和旅游厅推出“AI 游贵州”小程序，整合全省 9 个市州文旅资源，依托大模型算法，结合景点关联、交通衔接及游客体力等因素，为游客精准制定最优旅游路线，以技术赋能实现服务信息的精准推送，减少因信息不对称导致的虚假宣传、服务缩水等失信行为，践行《中华人民共和国电子商务法》中关于平台服务规范、信息披露义务的相关

^① 敦煌研究院. “数字敦煌”加载自动化和智能化[EB/OL]. 2025 年 10 月 10 日[2026 年 3 月 13 日]. <https://www.dha.ac.cn/info/1021/7085.htm>.

^② 人民网. 数智新场景引领文旅高质量发展[EB/OL]. 2025 年 4 月 22 日[2026 年 3 月 13 日]. <http://fj.people.com.cn/n2/2025/0422/c181466-41204701.html>.

规定。浙江省“云上西湖”小程序则为游客提供全程式、一站式、个性化旅游公共信息服务，覆盖西湖景区“吃、住、行、游、购、娱、研”全领域，包含场馆预约、一码通、智慧导览、多语言讲解、旅游舒适度、游玩指南等 20 多项功能，^①成为文旅个性化服务的典范。

（2）智慧监管与服务效能提升

智慧监管是数字文旅发展的重要保障，各地通过数字技术持续提升文旅市场监管水平与服务质量。三坊七巷官方旅游平台小程序创新上线全国首个人文景区智能体“三坊七巷 AI 文旅地图”，实现数字大模型技术与传统文旅资源的创造性融合，整合视听体验、出行规划、实时导航及 AI 互动伴游服务，覆盖景区旅游热点、特色展馆、演出、美食、购物及住宿等资源，进一步丰富了市民游客们的个性化与互动体验。^②

（二）数字文旅突出问题

数字文旅产业在快速发展过程中，网络诚信建设仍面临诸多突出问题，不仅破坏市场公平竞争秩序、侵害消费者与合规经营主体的合法权益，而且大幅提高了文旅消费者与文旅经营者之间的信任成本和交易成本，制约了数字文旅和实体文旅深度融合发展。

1. 平台虚假宣传与捆绑销售乱象丛生

^① 光明网. AI+文旅开启更智慧的旅程[EB/OL]. 2026 年 3 月 11 日[2026 年 3 月 13 日]. https://e.gmw.cn/2026-03/11/content_38640335.htm.

^② 福建省文物局. 福州：数字技术解锁古厝新玩法 千年坊巷“智慧旅游”丰富游客个性化体验[EB/OL]. 2025 年 4 月 16 日[2026 年 3 月 13 日]. http://wwj.wlt.fujian.gov.cn/xwzx/wbyw/202504/t20250416_6901127.htm.

线上文旅平台及入驻经营主体的虚假宣传行为，严重违背诚实信用原则，侵害消费者的知情权与公平交易权，已成为破坏数字文旅诚信生态的主要问题之一。

典型案例显示，消费者王某在机票代销平台上购买机票时，平台显示票面价格 280 元，机建+燃油费 70 元，另外可享受 40 元优惠，最终实际支付 310 元。但王某收到平台提供的客票信息后，在航空公司官方软件上查询，发现机票实际票面价格为 230 元，机建+燃油费 70 元，总计 300 元。多出来的 10 元是平台在销售机票过程中搭售的外卖服务包。北京互联网法院认为，平台在未设置醒目、清楚的提醒以及可以拒绝的选项的情况下搭售外卖服务包，属于故意隐瞒商品真实信息，构成价格欺诈。

另一案例中，消费者樵某在平台上预订酒店住宿，预订页面中酒店名称后被标注了五颗星图形（★★★★★），樵某认为该图形表示五星级酒店，后经查证该酒店并非五星级酒店。法院支持了樵某的诉请，认为平台经营者在描述商品及服务特征时应恪守诚信原则，采用容易让消费者混淆的内容且未以显著方式作出特殊标注和说明，足以对消费者造成误导，属于欺诈行为。^①

除明确的虚假宣传外，部分文旅经营主体线上宣传的产品与服务标准，与消费者实际体验严重不符，形成“货不对板”的失信乱象，严重影响游客体验，破坏行业诚信氛围。

^① 中国新闻网. 网上预定机票酒店套路多 电商平台如何担责[EB/OL]. 2025年2月14日[2026年3月14日]. <https://www.chinanews.com.cn/cj/2025/02-12/10367104.shtml>.

据人民网报道，消费者朱某在飞猪旅行网入驻旅行社“海南三亚环球国际旅行社”的直播间，下单三亚亚龙湾某豪华酒店套房，支付单价 2880 元/晚，直播间宣传该酒店“步行 3 分钟直达海边”“配套设施完善”。但朱某抵达酒店后发现，实际情况与宣传严重不符：酒店并非步行可达海边，需穿过小树林到达未开发的“临海”区域，距离最近的正规海滩需打车 4 公里，且酒店周边多为施工工地，配套设施匮乏，与“豪华酒店”的宣传定位差距较大。^①

2. 不实“避雷帖”滥用引发新型诚信争议

随着数字文旅的发展，网络评价成为消费者选择文旅产品的重要参考，但不实“避雷帖”“曝光帖”的滥用，逐渐成为扰乱市场公平竞争秩序、引发诚信争议的新型问题，违背《中华人民共和国民法典》诚实信用原则与《中华人民共和国反不正当竞争法》的核心要求。

2025 年 12 月，丽江市古城区文化和旅游局公开发函，指出小红书平台在履行信息监测及管理主体责任、审核义务方面存在不到位情况，导致大量消费者在企业无过错的情况下，为实现超合理诉求，利用平台发布不实“避雷帖”“曝光帖”，以舆论施压的方式威胁企业退钱退款，严重侵害了文旅经营主体的合法权益。据丽江市古城区旅拍行业协会调查摸排，截至 2025 年年底，丽江市当年因不实“避雷帖”导致相关企业退订金额达 100 余万元，间接经济损失超 500

^① 人民网. 直播间里订酒店货不对板，商家要担责吗？[EB/OL]. 2025 年 3 月 18 日[2026 年 3 月 14 日]. <http://gs.people.com.cn/n2/2025/0318/c183342-41167173.html?spm=5176.28103460.0.0.184175517ZAMbs>.

万元，凸显此类乱象对数字文旅诚信生态的严重破坏。^①

这一现象也折射出数字文旅评价系统的两难困境：平台既需依法保护消费者表达真实体验的言论自由权利，不能过度审核压制消费者的正当诉求与言论表达，又需切实履行主体责任，防止评价权被滥用于不正当竞争、恶意索赔等行为，否则将导致“劣币驱逐良币”的恶性循环，破坏数字文旅市场的诚信生态，最终损害消费者与合规经营主体的双重利益。

3.格式条款不公与维权追溯困难

(1) 格式条款不公平，提示说明义务缺失

北京互联网法院审理的相关案件显示，在线文旅消费中，经营主体多采用预先拟定的格式合同，部分格式条款存在不符合法律规定、不公平、不合理的内容，成为引发售后纠纷的主要原因之一，违反《中华人民共和国民法典》《中华人民共和国消费者权益保护法》中关于格式条款的相关规定。实践中，突出问题表现为经营主体对格式合同中免除或者减轻自身责任、加重消费者责任、限制或者排除消费者主要权利等相关条款，未尽到法定的提示说明义务，导致消费者在不知情的情况下接受不公平条款，合法权益受到侵害。

典型案例显示，方某在某票务平台上同时购买两张同场次演唱会门票，后因行程有变，向平台申请退票，其中一张门票退票成功，另一张却被平台告知无法退票。平台给出的

^① 南方网. 避雷帖引发文旅局喊话：丽江旅拍，套路多还是维权难？[EB/OL]. 2026年1月14日[2026年3月14日]. https://news.southcn.com/node_64549305f1/a9bcf52166.shtml?spm=5176.28103460.0.0.184175517ZAMbs.

理由是，购票页面的票务须知记载：“在销售阶段同一购票人、同一购票账户仅享有一次绿色通道权益，在产生一次退票后，如该购票人/购票账户再次购买同场次演出票后，将不再享有退票权益”。^①北京互联网法院审理后认为，依据通常理解，该条款的含义应为“退票成功后再次购买同场次门票，不再享有退票权益”，而涉案的另一张门票并非在方某退票成功后再次购买，因此平台不能依据前述格式条款免除自身的退票义务。该条款因存在歧义且平台未对其进行明确提示说明，不能作为免除平台责任的依据，这一判决也彰显了格式条款需符合公平原则、经营主体需履行提示说明义务的法律要求。

（2）恶意毁约频发与消费者维权追溯困难

在节假日等文旅消费高峰期，部分经营主体违背诚实信用原则，存在“恶意毁约、哄抬价格”的乱象，严重侵害消费者的缔约权与公平交易权。具体表现为，消费者下单并支付相关费用后，商家以“房东自住”“装修”等虚假理由拒绝确认订单，随后将产品涨价另行销售，从中牟取不正当利益。例如，消费者杨某在某平台预订三亚某酒店景观房2间4晚，共计支付464元，全款支付后收到平台发送的预订成功短信，但当日平台客服即电话通知杨某，称商户反馈房间满房无法安排入住，单方面取消了订单。杨某随后发现，该平台上仍然可以预订同一时间、同一酒店、同样房型的房间，

^① 人民网. 格式条款有争议 解释应有利消费者[EB/OL]. 2025年7月17日[2026年3月14日]. <http://society.people.com.cn/n1/2025/0717/c1008-40523607.html?spm=5176.28103460.0.0.184175517ZAMbs>.

但价格已涨至每间每晚 378 元，该商家的行为已构成违约，且涉嫌价格欺诈。^①

此类恶意毁约行为之所以屡禁不止，核心在于消费者维权追溯困难，难以形成有效约束。一方面，在线文旅消费涉及平台、经营主体等多个责任主体，一旦发生纠纷，各主体之间易出现相互推诿的情况，责任认定难度较大；另一方面，旅游消费具有较强的流动性特征，消费者与经营主体多为异地交易，证据收集、维权成本较高，加之部分平台数据管理不规范，相关交易记录、沟通记录等证据保存不完善，导致消费者往往面临“追责无门”的困境，这种维权困境也间接助长了经营主体的失信行为，加剧了数字文旅市场的诚信失衡。

（三）数字文旅治理建议

针对前述数字文旅领域虚假宣传频发、不实“避雷帖”扰乱市场、格式条款不公及消费者维权追溯困难等突出问题，立足产业实际，整合各地实践创新经验，从信用评价体系建设、消费纠纷快速处置、协同治理机制完善三个核心维度，提出针对性治理建议，推动数字文旅与网络诚信理论体系深度融合，规范市场秩序、保障消费者权益，促进产业高质量发展。

1. 建立健全数字文旅信用评价体系

信用评价体系是数字文旅网络诚信建设的核心支撑，也

^① 中国法院网. 文旅消费火出圈 权益保护不缺位[EB/OL]. 2025 年 3 月 16 日[2026 年 3 月 14 日]. <https://www.chinacourt.cn/article/detail/2025/03/id/8746020.shtml>.

是规范市场主体行为、防范失信风险的关键。构建科学完善的信用评价体系，需立足多元主体协同治理，覆盖市场经营主体、从业人员、消费者，实现信用信息全面采集、科学评价与高效应用，推动信用监管与行业治理深度融合。

构建多元主体信用评价指标体系需坚持全面、科学、可操作原则，明确各主体信用评价的核心指标、评分标准与权重分配，实现标准化、规范化评价。重庆市沙坪坝区围绕“诚信游沙磁”目标，聚焦景区、酒店、旅行社等7类文旅市场主体，出台《沙坪坝区文化旅游体育行业信用评价标准》，明确信用评价指标、评分标准与分级依据，实现信用评价标准化、规范化。该区完成区级信用平台与“渝快办”“惠游重庆”等政务服务、文旅服务系统对接，编制5个智能监管预案，整合多部门信用数据8000余条，实现对文旅市场失信风险的精准预警、动态监管；依托“信用沙坪坝”平台公示诚信企业38家、失信企业6家，构建跨部门联合约束机制，让守信企业获得更多发展机遇，让失信企业寸步难行，切实强化信用监管的震慑力。^①海南省琼海市在2026年重点探索深化特色应用场景，如“信用+民宿”“信用+共享农庄”，守信主体可以凭个人“金椰分”在海易办平台和“金椰分”小程序享受多样化的优惠便利。^②

除政府层面的体系构建外，企业层面的诚信实践也为信

^① 重庆市文化和旅游发展委员会. 创新信用驱动, 全面服务提质 ——沙坪坝区“诚信游沙磁”建设试点工作取得阶段性成效[EB/OL]. 2025年11月3日[2026年3月14日].

http://whlyw.cq.gov.cn/zwx_221/bmdt/qxfc/202511/t20251103_15134646.html.

^② 周长强. 以信用体系建设为引领擦亮“诚信琼海”城市名片[J]. 中国信用, 2026年(第3期):第33页.

用评价体系提供了重要支撑。云浮翔顺龙山酒店探索建立信用承诺践行机制，将诚信建设融入企业日常运营，定期对员工开展廉政文化、信用承诺、服务标准等相关培训，明确员工诚信服务要求；制定服务质量承诺、价格诚信承诺，公开承诺7天内凭消费发票或凭证无理由退货，保障消费者合法权益；在酒店内部公示监督投诉电话，主动接受消费者监督，同时将信用承诺的践行情况与员工绩效考核挂钩，形成“人人讲诚信、事事守诚信”的企业氛围，先后获得“文明诚信示范餐厅”“连续十一年广东省守合同重信用企业”等荣誉称号，成为企业诚信经营的典范。^①

2.构建高效便捷的消费纠纷快速处置机制

消费纠纷快速处置机制能破解维权追溯困难、化解市场矛盾，是网络诚信理论在文旅领域落地的重要体现。立足数字文旅消费特点，构建快速响应、便捷高效、多元协同的处置机制，简化流程、降低成本、提升效率，保障消费者权益。

(1) 构建消费纠纷快速响应机制

构建快速响应机制的核心是依托数字技术优化投诉处理流程，线上受理、线下处置，缩短处理周期。浙江省桐庐县在全国首创线下纠纷线上“一键和解”模式，为破解数字文旅消费维权难题提供了可复制、可推广的方案。该模式通过打通支付数据，建立消费者与商家的直接信息沟通渠道，

^① 广东省文化和旅游厅. 关于公布2025年文化和旅游市场信用应用创新案例名单的公告[EB/OL]. 2025年12月22日[2026年3月14日]. https://whly.gd.gov.cn/open_newgggl/content/post_4832694.html?spm=5176.28103460.0.0.184175517ZAMbs.

消费者在文旅消费过程中产生纠纷后，只需打开相关支付账单并点击“对此订单有疑问”，即可找到“消费纠纷一键和解”入口，系统自动导入订单时间、消费金额、商家信息等核心内容，消费者无需手动填写，选择退款、退货等具体诉求并上传相关证据后，即可实现与商家“面对面”在线沟通协商，大幅简化维权流程、降低维权成本。^①

（2）建立健全消费纠纷先行赔付机制

先行赔付机制能够破解维权“举证难、理赔慢”问题，需依托数字平台搭建体系，明确赔付范围、标准与流程，实现快速赔付、及时止损。海南省强化制度创新，依托数字技术打造“海南放心游”平台，建立“赔付原则、受理流程、办理时限统一”的先行赔付机制，整合全省文旅商家诚信保证金形成资金池，明确适用场景、标准与流程，实现3分钟受理、30分钟处置、48小时办结，提升游客满意度与信任度，形成“平台担保、先行赔付、后续追偿”的闭环。^②

（3）协同治理机制的完善路径

数字文旅网络诚信建设并非单一主体的责任，针对虚假宣传频发、维权追溯困难等突出问题，需参照网络信息内容治理，构建“政府—平台—社会”多元共治生态，^③强化各主体协同发力，推动治理模式从“事后惩戒”向“全过程治

^① 中国市场监管报. 桐庐试点“一键和解”机制数字赋能旅游消费提质升级[EB/OL]. 2025年9月24日[2026年3月14日]. http://www.cmrnn.com.cn/content/2025-09/24/content_278180.html.

^② 海南省人民政府. “海南放心游”春节升级不打烊[EB/OL]. 2026年2月15日[2026年3月15日]. <https://www.hainan.gov.cn/hainan/5309/202602/73e41249909d41a1a51fe9cdc4837987.shtml?spm=5176.28103460.0.0.184175517ZAMbs>.

^③ 王竹、郑兴华. 论网络信息内容的信用协同治理路径[J]. 山东社会科学, 2025年(第10期).

理”转变。重庆市渝中区创新推出“信用+商圈”赋能商贸文旅高质量发展，具有一定的典型意义。^①

监管部门应强化牵头引领作用，完善信息共享和执法协作机制。针对数字文旅中的虚假宣传、非法招揽、欺诈消费者、强迫消费等现象，监管部门加大联合执法力度。文化和旅游部在2025年12月审议通过了《旅游投诉处理办法》（文化和旅游部令第14号）并于2026年3月15日施行，对依法高效处置旅游投诉作出了规定。针对数字文旅中的问题完善了管辖制度，不仅完善了原有的地域管辖、级别管辖，新增了指定管辖。重庆沙坪坝区建立健全由区发改委牵头，区大数据局、区文化旅游委、区商务委、区市场监管局及各镇街共同参与的文旅信用建设联席会议制度，通过召开专题会议，协调推进预付费信用监管平台建设等工作；在磁器口古镇、歌乐山红色景区等重点文旅区域开展专项整治10次，检查经营主体210户，查处虚假宣传、违规收费等违规行为12起，推动涉旅投诉量同比下降25%；同时建立“24小时旅游投诉专线+先行赔付”快速处理机制，累计处理投诉800余件，办结率达100%，游客满意度达92%，形成“监管有力度、维权有保障”的良好局面。^②在接受、查处旅游投诉时，投诉受理部门可以针对每个案件建立诚信档案，为信用

^① 重庆市渝中区发展和改革委员会. 重庆渝中区：深化政务诚信建设，以“信用+商圈”赋能商贸文旅高质量发展[J]. 中国信用, 2026年(第3期):第37页.

^② 重庆市文化和旅游发展委员会. 创新信用驱动，全面服务提质——沙坪坝区“诚信游沙磁”建设试点工作取得阶段性成效[EB/OL]. 2025年11月3日[2026年3月14日]. http://whlyw.cq.gov.cn/zwx_221/bmdt/qxfc/202511/t20251103_15134646.html.

监管和信用市场累积数据资料。

线上文旅平台应切实履行主体责任，完善内部管理机制，规范平台经营行为。北京互联网法院结合相关案件审理经验，建议平台明确区分自营业务和非自营业务，以显著方式标记，避免消费者无法准确判断业务经营者的主体身份，减少信息不对称；严格履行商家信息核验、登记义务，对入驻商家的资质进行严格审核，对商家的违法违规行及时采取下架产品、限制流量、终止合作等有效处置措施，从源头遏制虚假宣传等失信行为；针对格式合同中涉及退款、售后等易引发纠纷的条款，若存在多种解释，应严格遵循《中华人民共和国民法典》规定，作出不利于提供格式条款一方的解释，保障消费者合法权益。

社会监督应发挥补充作用，构建多元共治格局，营造“人人讲诚信、人人守诚信”的行业氛围。重庆沙坪坝区创新监督模式，招募 30 名“诚信文旅体验官”，邀请其深入景区、酒店、旅行社等经营场所，参与企业信用监督，及时发现并反馈失信行为；计划在 A 级旅游景区布设信用查询终端，推动商户信用等级挂牌公示，提升信用信息透明度，方便消费者参考选择，引导经营主体诚信经营、游客诚信消费。

五、网络公益：筑牢公益慈善的网络诚信防线

网络公益是以互联网为媒介开展的公益活动，依托网络技术突破时空限制、整合线上线下资源、降低公众参与门槛，广泛覆盖物质援助、志愿服务、文化传承等诸多领域，旨在

构建开放共享的社会化公益生态。^①2020年印发的《法治社会建设实施纲要（2020—2025年）》提出深入实施中国好网民工程和网络公益工程，这意味着网络公益慈善被正式纳入国家政策框架之中。^②此后，随着“互联网+”公益慈善的深度融合，网络公益的形式与场景日益多元，逐渐形成“线上传播+线下执行”的协同运作体系。^③

“十四五”期间，我国网络公益实现新发展、新突破，重点体现在法律保障的持续夯实，尤其是慈善法的修改与落地。这部法律于2023年12月29日经十四届全国人大常委会第七次会议表决通过，并在2024年9月5日——我国第九个“中华慈善日”正式施行。这次修法工作并非小修小补，它完善了慈善的功能定位与监管机制，充实了慈善信托新制度，并针对近年来慈善实践中的痛点问题作出了系统性回应，不仅新增了“应急慈善”专章，更填补了网络个人求助的法治空白，补齐了重大突发事件中慈善力量如何有序参与的短板，还进一步厘清了公开募捐的规则边界，并首次明确了互联网公开募捐服务平台的法律责任。配套政策体系同步跟进，《慈善组织认定办法》（民政部令第73号）《慈善组织公开募捐管理办法》（民政部令第74号）等一系列规章的制订修订，让新法的原则性规定在实操执行的过程中得以被贯

① 百度百科. 网络公益[EB/OL]. 2022年3月20日[2026年3月19日].
<https://baike.baidu.com/item/%E7%BD%91%E7%BB%9C%E5%85%AC%E7%9B%8A/9231900>.

② 同上

③ 中国政府网. 中共中央印发《法治社会建设实施纲要（2020-2025年）》[EB/OL]. 2020年12月7日 [2026年3月18日]. https://www.gov.cn/zhengce/2020-12/07/content_5567791.htm.

彻落实。^①

在此基础上，“十四五”期间我国慈善事业实现有序增长，慈善力量深度融入国家大局。慈善组织从 9400 家增至超 1.6 万家，慈善信托增长近 4 倍，年度捐赠总额首破 2000 亿元，全国慈善组织年均支出约 1500 亿元，精准投向民生关切领域。^②互联网慈善同样交出亮眼答卷，民政部指定的 29 家互联网公开募捐服务平台和 3 家个人求助网络服务平台规范运行，累计超过 600 亿人次通过这些平台奉献爱心，捐赠总额逾 1600 亿元。^③

与此同时，主题宣传活动的传播力也在持续放大。2025 年第九个“中华慈善日”活动期间，全国共有 1800 万人次参与，654 家媒体发布 5800 余篇报道，互联网相关信息达 15.8 万条，覆盖 2.4 亿人次，成为历史上参与部门最多、传播效果最好的一次主题宣传活动。^④

地方实践同样涌现出具有全国乃至国际影响力的扎实经验。自 2018 年起，河南省慈善联合总会便将数字慈善作为重大战略强力推进，构建“五个慈善”工作格局，创新“会长分片指导+团队全员募捐”工作法，建立覆盖省、市、县、乡、村的“五级联动”机制。截至 2025 年 11 月底，该项目

^① 民政部.深学细悟做实修改后的慈善法[EB/OL]. 2024 年 5 月 7 日 [2026 年 3 月 18 日].
<https://www.mca.gov.cn/n152/n166/c1662004999979999335/content.html>.

^② 中国新闻网.慈善“成绩单”亮眼 第十二届中国慈善年会发布“十四五”成就与“十五五”新趋势[EB/OL]. 2026 年 1 月 18 日[2026 年 3 月 18 日].
<https://baijiahao.baidu.com/s?id=1854657338388118477&wfr=spider&for=pc>

^③ 民政部.在规范发展中持续释放温暖而坚实的力量——2025 年慈善事业发展综述[EB/OL]. 2025 年 12 月 29 日 [2026 年 3 月 18 日]. <https://www.mca.gov.cn/n152/n166/c1662004999980008629/content.html>.

^④ 同上

联动全省各级慈善组织累计募款 42 亿余元，带动 6172.8 万人次参与，实施项目 4000 多个，惠及群众 330 余万，连续数年在腾讯公益平台全国性网络募捐活动中筹款额位居第一。河南的公益实践，真正走出了一条政社协同、资源整合、规范发展的数字慈善新路。《“十四五”时期我国慈善事业发展成就和“十五五”时期发展趋势》报告则进一步揭示了“十五五”时期数智技术在慈善各环节的全面渗透。^①

（一）网络公益发展成效

1. 技术向善：AI 助力公益慈善创新发展

在公益慈善领域，人工智能正从辅助工具转变为解决社会问题的重要力量。2025 年 8 月 21 日，国务院印发《关于深入实施“人工智能+”行动的意见》（国发〔2025〕11 号），将“‘人工智能+’民生福祉”纳入顶层设计，明确鼓励运用人工智能赋能养老、托育、助残等社会服务领域，推动公益慈善向智能化方向演进。^②

2025 年 12 月 12 日，由字节跳动公益主办，腾讯公益、阿里巴巴公益等联合支持，聚焦技术与慈善深度融合的“数字慈善与 AI 赋能交流会”发布了“AI 慈善工具库”，该工具库整合了覆盖七大工作场景的 29 个智能工具，为公益慈善机构的日常运转提供科技助力。^③

^① 中国新闻网. 慈善“成绩单”亮眼 第十二届中国慈善年会发布“十四五”成就与“十五五”新趋势[EB/OL]. 2026 年 1 月 18 日[2026 年 3-月 18

日]. <https://baijiahao.baidu.com/s?id=1854657338388118477&wfr=spider&for=pc>

^② 中国政府网. 国务院关于深入实施“人工智能+”行动的意见[EB/OL]. 2025 年 8 月 26 日 [2026 年 3 月 18 日]. https://www.gov.cn/gongbao/2025/issue_12266/202509/content_7039598.html.

^③ 2025 年中国公益慈善项目交流展示会. 央广网: 2025 中国公益慈善项目交流展示会在广州举办[EB/OL]. 2025 年 12 月 16 日 [2026 年 3 月 18 日]. <https://www.charityalliance.org.cn/ds/2512dc1b1e.html>.

除工作效率的提高外, AI 技术更是在特殊群体的帮扶中逐渐展现出了它的不可替代性。唇腭裂(兔唇)是我国口腔颌面部最常见的先天性畸形, 内地每年有 2.5 万名新生唇腭裂患儿, 保守估计全国唇腭裂患者总数维持在 240 万人左右。唇腭裂的治疗周期漫长, 医疗成本高昂的同时言语治疗师还长期短缺, 很多偏远地区的家庭难以很好地接受术后的康复训练。^①字节跳动与四川大学华西口腔医院联合研发出了一款名为“光合愈言”的言语康复小程序, 它通过智能测评发音问题、精准推送纠音内容, 让患儿能够在家长陪伴下居家完成免费训练。^②截至 2025 年 12 月份, 它在多家唇腭裂机构落地, 已服务用户 2596 人, 实实在在弥补了康复资源的缺口。^③

自 2016 年起, 腾讯等企业便积极推动互联网、AI 等技术应用于公益场景。^④一方面互联网企业积极举办各类活动吸引开发者涌入“AI+公益”赛道, 如腾讯 Light·创造营、阿里巴巴联合魔搭社区主办的“大有可为”AI 开源公益挑战赛等。另一方面, 企业鼓励内部员工以专业能力入局, 成为公益项目的设计者、共创者。如阿里巴巴的“人人三小时”平台和字节跳动的“跳跳糖公益”, 前者十年累计公益时长

^① 文汇报.中国兔唇患儿约 240 万 治疗周期长达十余年[EB/OL]. 2026 年 1 月 21 日 [2026 年 3 月 18 日]. <https://www.wenweipo.com/a/202601/21/AP696fe581e4b04d7d56cf7767.html>.

^② 华西口腔医院.我院唇腭裂外科与字节跳动公益联合推出 AI 言语诊疗系统[EB/OL]. 2026 年 1-月 8 日 [2026 年 3 月 18 日]. <https://www.hxkq.org/Html/News/Articles/17562.html>.

^③ 北青网. AI 帮助唇腭裂儿童矫正语音[EB/OL]. 2025 年 12 月 29 日 [2026 年 3 月 18 日]. <https://baijiahao.baidu.com/s?id=1852803106300524098&wfr=spider&for=pc>.

^④ 百度百科.技术公益[EB/OL]. [2026 年 3 月 18 日].<https://baike.baidu.com/item/%E6%8A%80%E6%9C%AF%E5%85%AC%E7%9B%8A/62825967>

突破 500 万小时，催生了蚂蚁森林、“团圆”系统、轮椅导航、医疗 AI 癌症早筛等一系列深度融入业务的公益项目；^①后者则通过伙伴计划，支持员工开发出 AI 情绪认知小游戏、视障人士化妆辅助 APP “妆境”等产品。^②

2.打破壁垒：跨平台协同催生优质内容公益

2025 年，中国互联网公益迎来了一场更为深刻的转向。头部平台不约而同地弱化了对捐赠金额的追逐，转而把重心放在技术工具开发和资源联动上，试图以此回归平台对于公益的服务属性。正是在这个过程中，互联网公益平台之间的壁垒，开始被实质性打破。

在第九个“中华慈善日”期间，字节跳动公益联合腾讯公益、微博公益共同推出“慈善文化传播优质内容支持计划”，聚焦抖音、视频号和微博三家平台的优质公益内容，实施流量扶持与多层级的激励机制，让好内容被更多人看见。数据显示，截至 2025 年 8 月底，字节跳动平台已上线 3907 个公益项目，累计筹集善款 20 亿元，参与捐赠总人次超 17.3 亿，参与公益的创作者达 8.2 万名，相关公益内容累计播放量超 850 亿次，真正实现了优质公益内容的广泛传播，让更多人投身到公益事业之中。^③

作为“慈善文化传播优质内容支持计划”的一部分，字

^① 杭州网.从杭州出发的一场公益接力!十年志愿服务累计突破 500 万小时[EB/OL].2026 年 3 月 17 日 [2026 年 3 月 18 日].<https://baijiahao.baidu.com/s?id=1859891751667288621&wfr=spider&for=pc>.

^② 中国慈善家.AI 向善渐成共识，新的故事才刚刚开始[EB/OL]. 2026 年 1 月 5 日 [2026 年 3 月 18 日].
<https://baijiahao.baidu.com/s?id=1853437421809341739&wfr=spider&for=pc>.

^③ 民政部.字节跳动公益联合多方举办新媒体慈善文化传播交流会[EB/OL]. 2025 年 9 月 9 日 [2026 年 3 月 18 日]. <https://www.mca.gov.cn/n152/n166/c1662004999980006688/content.html>.

字节跳动公益搭建了一个双向筛选匹配的报名平台，让好项目和优质创作者得以互相选择，进一步推动公益内容的创作落地。2025年“DOU爱公益日”期间，平台成功撮合了近百位创作者为公益项目发声，这成为“平台托举公益——公益反哺创作者”的典范，帮助创作者在获得物质激励之余发挥更多的社会价值。^①

3.模式创新：交易捐与场景公益扩容升级

公益与商业场景的深度融合，让人们在日常消费中“无意识”实施善举。19年前，从淘宝第一笔公益宝贝“交易捐”落地，至今已累计达成数十亿元捐赠，捐赠订单量超600亿笔。2025年，交易捐的规模与形式持续扩容。在规模上，今年阿里系的闲鱼、闪购、1688等电商平台也相继加入公益宝贝；字节跳动“爱心好物”已有10.5万个商家参与，为超150个公益项目筹款达1亿元（截至2025年10月）；美团则带动178.6万家公益商家参与交易捐赠，累计建成乡村儿童操场4386座，惠及59.4万名乡村儿童。在形式上，随着直播带货、即时零售等新业态的兴起，交易捐也呈现出新特征，主要体现在两方面：一是与直播相结合，二是商家与公益建立更深层次的连接。海量中小商家通过平台完成公益启蒙，在与公益项目深度联动的过程中持续贡献微观层面的实践经验，为公益行业注入了丰富多元的价值增量。^②

^① 中国慈善家. 中华慈善日走过10年, 互联网平台间公益合作成为新常态[EB/OL]. 2025年9月29日 [2026年3月18日]. <https://www.zgcsj.com/gd/2025-09-29/1946.shtml>.

^② 中国慈善家. AI向善渐成共识, 新的故事才刚刚开始[EB/OL]. 2026年1月5日 [2026年3月18日]. <https://baijiahao.baidu.com/s?id=1853437421809341739&wfr=spider&for=pc>.

支付环节的创新也带来了新的活力。支付宝推出“碰一下公益开放计划”支付功能，用户在线下便利店、电梯广告等场景“碰一下”即可完成捐赠，简化线下捐赠流程的同时，将“无意识”的消费转化为“有意识”的公益互动，让公益活动更便捷、更个性化，更加触手可及。^①

场景公益的触角也在向基层社区延伸。延边州推进的“幸福家园”村社互助工程，依托“公益宝”平台实现八县（市）、1147个村（社）全覆盖，发布精准帮扶项目242个，募集善款210.31万元。这种“熟人社会+数字公益”的模式，让慈善真正扎根基层，探索出社区场景下的公益新路径。^②

4.透明化提升：区块链与反馈机制增强公众信任与热情

为使公益过程透明可信，民政部积极推动全国慈善信息公开平台与区块链技术有机对接，依托其不可篡改、全程可追溯的特性，破解善款去向不明等难题。在区块链技术赋能下，捐赠人可以实时追踪善款流向，了解善款用途，增强对公益组织的信任由此更加稳固。^③

当技术层面的透明逐渐成为公益标配，整个行业开始探寻更深层的连接——让善意被看见，更被感知。字节跳动公益发起的“好事成真”活动，便是一个“数字可见”到“情感可触”的反馈机制探索。该活动上线的“年度好事报告”，

^① 民政部.支付宝公益平台推出“碰一下公益开放计划”[EB/OL].2025年9月11日[2026年3月18日].
<https://www.mca.gov.cn/n152/n166/c1662004999980006738/content.html>.

^② 吉林省慈善总会.延边州全力推进“幸福家园”村社互助工程[EB/OL].2026年1月27日[2026年3月18日].
<https://www.jlcs.org.cn/nv.html?nid=1b848643-688b-4008-9193-72f93a60b786>.

^③ 钱晓田,沙勇.以“善治”推动数字公益行稳致远[EB/OL].2026年1月28日[2026年3月18日].
<https://sps.njupt.edu.cn/2026/0130/c10220a296942/page.htm>.

吸引了超 8200 名创作者投稿，60 万用户通过报告获取受助人反馈，回溯自己的那份善意究竟去了哪里。^①通过活动，抽象的数字被还原成了一个个具体而鲜活的人间故事，有了温度。可以说，区块链技术搭建起信任的骨架，反馈机制为公益活动注入情感的血肉，二者一起让公益的价值变得可见、可感、可信。

（二）网络公益突出问题

技术是一把双刃剑，在为网络公益带来便捷的同时，也给社会治理带来了一系列新挑战，如虚假募捐屡禁不止，伪公益营销花样迭出，数字鸿沟问题凸显，平台责任边界模糊不清等。下文将系统梳理当前阶段数字公益发展中存在的突出问题，为构建更加透明、高效、包容的数字公益生态提出针对性治理建议。

1. 虚假募捐、善款挪用、违规配捐等欺诈屡禁不绝

（1）虚假募捐

2025 年 12 月，陕西西安公安机关查处一起网红虚构事实骗捐案。拥有 10 万粉丝的网红王某莎，在多个互联网平台谎称孩子误食洁厕灵生命垂危，以急需 20 万元医疗费为由，在筹款平台发起募捐。经查，王某莎存在虚构事实、夸大病情的行为，致使 149 名爱心人士捐款。事发后，筹款平台立即暂停其筹款并将善款原路退还，公安机关依法对王某

^① 北青网. 字节跳动公益上线好事报告：60 万用户获取成果反馈[EB/OL]. 2026 年 1 月 9 日 [2026 年 3 月 18 日]. <https://baijiahao.baidu.com/s?id=1853821566438406738&wfr=spider&for=pc>.

莎作出行政处罚。^①除虚构紧急事件骗捐外，还有一类依托长期打造的慈善人设实施的骗捐行为，其隐蔽性、欺骗性与危害性更甚。2025年6月，浙江绍兴公安机关通报称，曾救助上百名孕妇及弃婴而被叫作“和尚爸爸”的道禄（原名吴某），因涉嫌诈骗被依法采取刑事强制措施。警方查明，2018年以来，道禄伙同他人长期在江浙多地活动，以“资助孕妇、助养儿童”为名骗取社会捐赠，且将大部分善款用于个人高消费。^②

机构层面的虚假募捐常以冒用正规慈善组织名义、仿冒机构名称、伪造官方资质的组织化形式实施，欺骗性更强、波及范围更广。2024年10月，山东省东平县民政局依法取缔非法组织“中国慈善联合会东平服务站”，该组织擅自冒用中国慈善联合会名义，以地方代表机构自居，伪造党和政府奖章及荣誉证书，并通过APP诱导群众参与所谓的捐赠投资，社会危害极大。无独有偶，福建漳州的“长泰区慈善会”故意使用与合法慈善组织高度近似的名称，通过设置“功德箱”诱导群众捐款；湖南溆浦的“中华志愿者协会溆浦服务站”则以“乡村振兴惠民招生”为幌子，虚构助学项目，以“读5年（大专）、读7年（本科），家里只交5000元，毕业分配工作”为噱头欺骗群众。^③

^① 公安部网安局.网红虚构病情骗捐?陕西西安网警依法查处一起捏造虚假信息骗取善款案[EB/OL]. 2025年12月8日 [2026年3月18日].

^② 央广网.网传“救助百名弃婴的道禄和尚被警方带走”，绍兴警方通报[EB/OL]. 2025年6月26日 [2026年3月18日]. <https://baijiahao.baidu.com/s?id=1835968431184968488&wfr=spider&for=pc>.

^③ 民政部.民政部公布8起公益慈善领域非法社会组织典型案例[EB/OL]. 2025年5月27日 [2026年3月18日]. <https://www.mca.gov.cn/n152/n164/c1662004999980005308/content.html>.

（2）善款挪用

在专项治理与监管收紧背景下，挪用善款的违规行为显著减少，据审计署披露，福建晋江慈善总会将 10 亿元善款“借”给地方融资平台长达五年未还，事件一出便引发舆论哗然。《中国慈善家》进一步深度调查发现，晋江市慈善总会自 2020 年起持续向晋江城市建设投资开发集团提供巨额资金，从最初的“借款”到后来的“委托”，金额始终超过 10 亿元，相当于其 22 年总筹款额的五分之一。^①这笔善款从慈善账户流向地方城市建设，不仅让公益初心沦为融资债务的筹码，更直接突破了法治红线、损害了社会公共利益。

（3）违规配捐

平台配捐机制被不法分子利用，形成盘剥患者的黑色产业链。2026 年 2 月，媒体曝光了浙江省梦守护公益基金会涉嫌“配捐”诈骗的事件。重大疾病患儿家属王琳（化名）为给孩子筹措医疗费，经他人介绍接触到该基金会所谓的“配捐”渠道。该渠道声称，家长需先自筹 4.4 万元交给中间人，随后个人账户会收到基金会拨付的 4.5 万元，以此获得 1000 元的“配捐”。在此次“配捐”过程中，王琳虽先收到了 4.5 万元善款，却需立即将 4.4 万元转账给中间人龚某，此外还得额外花费 200 余元购买药品发票以完成“配捐”流程，最终实得善款不足 800 元。4.5 万元善款中，中间人抽走了 4.4 万元，原本应助力困境家庭的“配捐”，被彻底异化为对弱

^① 中国慈善家杂志.直面|晋江市慈善总会 10 亿元巨额善款，怎么“借”给了地方融资平台？[EB/OL]. 2025 年 7 月 21 日 [2026 年 3 月 18 日].

势群体的二次盘剥与伤害。^①

2. 伪公益营销损害公众利益

利用公众对公益的信任开展商业营销，已成为直播电商时代的典型顽疾，其中虚假助农与坑老型伪公益是最为突出的两类表现形式。

2026年2月，成都市蒲江县市场监管局查处主播林某虚假助农案，开出《直播电商监督管理办法》施行后四川省首张相关罚单。林某在直播中虚构“蒲江县某村助农推荐官”“林主任”身份，伪造加盖公章的工作牌，编造“果子滞销烂在树上”“救当地村民的命”等话术，刻意标榜直播“不以营利为目的”。经查实，林某无固定工作单位，其“主任”身份与在蒲江驻村工作的经历均属捏造，直播带货的耙耙柑也并非滞销产品。^②此类虚假助农乱象并非个例，在四川大凉山地区已形成屡禁不止的黑色产业链。全国人大代表、四川省广元市苍溪县白驿镇岫云村党支部书记李君在2026年全国两会上指出，“假助农”直播已成为亟待整治的行业乱象。部分网红借助网络水军造势，通过虚构贫困、编造悲情故事收割流量，既欺骗消费者，也严重破坏了优质农产品的品牌声誉。^③

除虚假助农外，针对老年群体的坑老伪公益同样猖獗。

^① 红网.中间人何以沦为配捐的“吸血鬼”? [EB/OL]. 2026年2月14日 [2026年3月18日].<https://baijiahao.baidu.com/s?id=1856185204769947589&wfr=spider&for=pc>.

^② 国家市场监督管理总局.四川开出《直播电商监督管理办法》实施以来首张罚单[EB/OL]. 2026年2月3日 [2026年3月18日] https://www.samr.gov.cn/xw/df/art/2026/art_1b3e8b89fea14feaa77e221982472de6.html.

^③ 中国食品安全报.全国人大代表,四川省广元市苍溪县白驿镇岫云村党支部书记、村委会主任李君: 严打伪助农 守护真振兴 建议整治网红大V 借助农收割乱象[EB/OL].2026年3月8日[2026年3月18日].<https://www.cfsn.cn/news/detail/22/339197.html>.

湖北查处的系列案件显示，部分商家以公益关怀为幌子实施欺诈。例如，有土特产店以免费鸡蛋为诱饵，虚假宣传“东林沅熊胆粉”可治疗十余种疾病，经查该产品既无国药准字号，也不属于保健食品，最终被处以5万元罚款；另有养生保健馆宣称“智任堂亚麻酸型食用植物油凝胶糖果”可治疗高血脂、高血压、糖尿病、心脏病、癌症等多种重疾，却无法提供任何检验报告予以佐证，被处以3万元罚款。^①上述行为无一不是打着健康关怀的旗号，行虚假宣传、牟利欺诈之实。

3. 数字鸿沟问题依旧存在

数字化工具大幅提升了公益运作效率，但技术红利并未实现全民普惠，公益领域的数字鸿沟问题依然严峻，具体体现为五大核心断层：

一是技术接触断层。2025年夏天，贵州榕江遭遇“超30年一遇”的特大洪灾。腾讯基金会联合中国红十字基金会发起救灾物资发放项目，向重灾区发放“数字关爱卡”，村民持卡可通过微信小程序“刷脸”核验身份，自主选择物资。然而，这一数字化创新在落地时遭遇困境。志愿者们握着智能手机，穿梭在吊脚楼间，一遍又一遍地教第一次接触数字化的老人操作小程序。^②这个案例说明目前仍有相当数量的弱势群体未被数字世界所兼顾。

^① 光明网.“坑老”典型案例公布！免费领鸡蛋成标配……[EB/OL]. 2026年1月26日 [2026年3月18日]. <https://baijiahao.baidu.com/s?id=1855357154797955123&wfr=spider&for=pc>.

^② 央广网.一场跨越“最后一公里”的善意变革[EB/OL]. 2025年12月11日 [2026年3月18日]. <https://baijiahao.baidu.com/s?id=1851190515525766993&wfr=spider&for=pc>.

二是设计适配缺失。老年群体普遍面临“不敢用、不会用”的窘境，残障群体则受限于技术精细化程度不足的问题，其根源在于公益相关技术产品未能真正以用户需求为核心开展设计。

三是资源分配的马太效应。头部公益平台与大型机构依托技术优势占据流量高地，而县域慈善组织、乡村公益项目因缺乏技术支撑与运营能力，难以获得同等曝光机会，所以数字技术反而可能加剧公益资源分配的不平等。

四是人文温度存在消解的风险。公益项目若过度依赖算法推送与效率导向，极易陷入“苦难叙事+情感煽动”的模板化模式，从而忽视受助者的人格尊严与主体地位。

五是多元共治实践不足。弥合公益数字鸿沟，需要构建政府、企业、社会组织、志愿者与受益群体协同参与的生态体系，而当前各方联动的协同机制仍不完善。

当然，社会各界跨越公益数字鸿沟的努力从未止步。2024年，vivo与中国残疾人福利基金会联合打造的“声声有息”公益计划，成功探索出“科技+公益+生态”的创新模式，推动技术完成从“功能实现”向“体验优化”的切实升级；2025年，双方进一步深化合作，开启“五年三千万科技助残专项行动”，为无障碍融合教育发展持续赋能。^①尚网网络与免费午餐基金共同发起的“梦想钥匙”公益项目同样成果

^① 中国残疾人福利基金会.声声有息 有爱无碍——中国残疾人福利基金会携手 vivo 推动科技惠普，照亮无障碍未来[EB/OL]. 2025年10月17日 [2026年3月18日].<https://www.cdpcf.org.cn/xwzx/ywdt1/df813f4bd6824a0284a40067d2da6a73.htm>.

斐然。自 2014 年启动至今的十年间，该项目已覆盖青海、四川、广西等 18 个省份，累计为 349 所乡村中小学接入网络，惠及超 5 万名学生。2025 年，项目聚焦青海玉树地区，为当地 40 所学前教育学校、4000 余名儿童连通互联网，让数字教育在儿童启蒙阶段落地生根。^①

4. 平台责任边界不清与监管规范缺失

数字公益作为新兴领域，面临网络募捐规范不完善、个人求助监管薄弱、善款流向难以追溯等现实挑战。^②尽管新修改的《中华人民共和国慈善法》已将个人求助网络服务纳入监管范畴，但在实际落地过程中，受平台公益属性与商业本质冲突的影响，不同类型平台为追求流量与经济收益，仍遗留着各式各样的弊病，亟待进行针对性厘清与规范。

网络众筹平台的核心问题是审核把关不严与责任转嫁。据媒体报道，该类平台普遍存在“轻审核、重流量”的倾向，用户仅凭一张病历即可发起筹款，随后便套用悲情模板以博取网友同情。部分平台甚至主动为求助者提供筹款文案模板，变相降低审核门槛。更值得警惕的是，多数平台通过用户协议中的“免责声明”，将求助信息真实性的核验责任完全转嫁给用户，大幅弱化自身法定审核义务，造成“平台引流、用户担责”的不合理现象。^③

^① 新闻晨报. WiFi 万能钥匙探索公益新模式: 赋能乡村教师 激活全民公益[EB/OL]. 2025 年 5 月 19 日 [2026 年 3 月 18 日]. <https://baijiahao.baidu.com/s?id=1832528420137568787&wfr=spider&for=pc>.

^② 钱晓田, 沙勇. 以“善治”推动数字公益行稳致远[EB/OL]. 2026 年 1 月 28 日 [2026 年 3 月 18 日]. <https://sps.njupt.edu.cn/2026/0130/c10220a296942/page.htm>.

^③ 每日经济新闻. 水滴筹、轻松筹员工又在医院起冲突? 双方联合声明: 动口了, 但没动手[EB/OL]. 2020 年 5 月 21 日 [2026 年 3 月 18 日]. <https://baijiahao.baidu.com/s?id=1667257982301909397&wfr=spider&for=pc>.

直播平台则面临内容真实性核验与资金监管的双重缺位。2025年9月被媒体集中曝光的“云投喂”流浪动物造假乱象，^①便是典型的借公益名义实施的直播欺诈与伪公益行为。多家电商与直播平台上，大量个人店铺以“救助流浪猫狗”为名开设“云投喂”直播间，盗用正规救助基地图片与视频，虚构动物困境与救助场景，诱导爱心网友购买虚拟投喂商品、进行打赏。然而，这些商家大多无公开募捐资质，资金去向完全不透明，许多店铺在敛财后迅速下架商品、改名甚至直接关店，消费者的善款并未真正用于动物救助。平台未能有效审核商家资质、核验直播内容真实性，也未对善款（打赏）流向建立监管与溯源机制，客观上纵容了此类伪公益骗局的发生。

内容社交平台的突出问题集中在账号审核宽松、违规交易监管不力两方面。2025年，中央社会工作部等部门开展志愿服务领域突出问题专项整治行动，通报的典型案列涉及多家头部互联网平台，相关账号违规发布志愿服务时长虚假信息，暴露出平台信息审核不严谨，日常监管处置不及时、不到位等问题。此外，部分机构还通过注册“志愿阜南”“海南志愿者”等带有官方色彩的账号误导公众，混淆视听。这进一步此类平台在账号注册环节，存在名称合规性审核标准模糊、资质把关不严的监管短板。^②

^① 人民日报. “云投喂”流浪猫狗？莫被假慈善“云骗”了[EB/OL]. 2025年9月27日[2026年3月18日]. <https://www.peopleapp.com/column/30050386027-500007115531>.

^② 新华网. 中央社会工作部公布志愿服务领域7个突出问题典型案例[EB/OL]. 2025年12月22日 [2026年3月18日]. <https://www.xinhuanet.com/20251222/9ad3d4786ddc4f8db34c619a69b3a390/c.html>.

（三）网络公益发展治理建议

1. 建立信用评级与善款溯源机制

构建全链条信用评级体系。建议由民政部门牵头，建立覆盖捐赠者、筹款平台、慈善组织、受益方的全链条信用评级体系。对于个人求助者，引入大数据交叉验证，将虚构、谎报事实者纳入社会信用黑名单；对于平台和机构，将信息审核通过率、投诉率、善款执行率等指标量化评级并向社会公开。

积极推广区块链技术应用，争取将区块链技术变为行业标配。所有公开募捐的项目，都需要将关键信息上链存储，包括善款流向、物资签收凭证等。此外，还应推动全国慈善信息公开平台与各商业平台区块链节点的数据互通，努力让每一笔善款的流转都干净、透明、可见、可信。

2. 强化监管，形成多元共治的公益慈善体系

厘清平台的责任边界。立法与监管部门应进一步细化《中华人民共和国慈善法》及配套规章中关于平台责任的条款，明确平台对求助信息、直播内容、商户资质具有实质性、持续性的审核义务，而非简单的形式审查，禁止平台通过用户协议中的免责条款规避自身法定责任。对于审核不严导致诈骗频发的平台，有关部门应追究其相应的行政乃至民事责任。

构建多元协同的监管网络。打破政府单一监管的局限，构建“政府监管+行业自律+平台内治+社会监督”的多元共

治体系。民政、网信、公安、市场监管等部门应建立常态化信息共享与联合执法机制。同时，鼓励和支持第三方评估机构、行业观察组织发展，畅通媒体与公众监督举报渠道，让虚假募捐和伪公益行为无处遁形。

完善执法裁量与专项整治。落实《民政部登记认定的慈善组织行政处罚裁量权基准》（民发〔2025〕61号）等规章，确保行政处罚过罚相当。对骗捐、挪用善款等行为，依据新修改的慈善法从严处罚。针对虚假助农、坑老伪公益等顽疾，常态化开展专项整治行动。

3. 严厉打击伪公益行为

强化内容审核与技术监测。平台应利用 AI 智能监测系统，对直播、短视频内容进行全天候巡查，重点识别“卖惨带货”“虚构人设”等剧本套路。对此，可以借鉴成都市的成功经验，运用直播电商 AI 智能监测系统对海量直播画面、语音进行实时分析，及时发现违规线索。^①

落实黑名单与联合惩戒制度。建立跨平台的“伪公益黑名单”共享机制。对于查实存在虚假宣传、骗捐诈捐的主播、机构和个人，不仅要在其所在平台封禁，更要实现全行业禁入。同时，将相关违法信息记入市场主体信用记录，实施多部门的联合惩戒，切实强化警示震慑效果，严守公益慈善诚信底线。

提升公众媒介素养与辨识能力。政府部门和平台应联合

^① 封面新闻. 违法宣传、虚假广告跑不脱！成都上新 AI“慧眼”，为 7 万直播间戴上监管“紧箍咒”[EB/OL]. 2026-02-08 [2026-03-18]. <https://baijiahao.baidu.com/s?id=1856541782688591897&wfr=spider&for=pc>.

开展公众媒介素养教育，引导消费者理性看待“悲情营销”。建议短视频平台加大公益广告流量扶持，围绕社会主义核心价值观进行公益广告选题和制作，让优质正能量内容占据主流。新技术高速发展背景下的中国网络公益，正处于机遇与挑战并存的十字路口。新修改的慈善法为行业发展夯实法治根基，AI等新技术正在重塑公益的底层逻辑，平台的自我变革与跨域协同正在拓宽行业边界。然而，欺诈与鸿沟也时刻警示我们，技术向善的道路并非坦途，以善治促善行，用更完善的法治、更清晰的责任、更包容的设计，才能让网络公益真正成为汇聚爱心、温暖人心、促进共同富裕的强大载体。

第四章 发展建议：提质增效的优化路径

深化网络诚信建设，是加强网络空间道德建设的重要任务，是顺应信息时代发展潮流的必然要求。立足“十五五”发展，必须坚持问题导向与目标导向相结合，聚焦技术风险防控、跨域协同攻坚、长效机制完善等方面，统筹制度规范、技术赋能、联动执法、主体责任、文化培育等关键举措，持续提升网络诚信建设水平，推动实现网上网下诚信建设有机融合、共建共享，营造依法办网、诚信用网的良好氛围。

一、风险挑战

2025年，我国网络诚信建设在制度完善、技术应用和协同共治方面取得显著进展。随着生成式人工智能等数字技术的加速迭代，失信行为跨界演化、治理协同壁垒等问题交织叠加，新业态新模式层出不穷，亟需破解新技术衍生诚信风险、跨域治理碎片化、长效监管乏力等突出难题，推动网络诚信治理从被动应对向主动防控、从分散治理向协同共治、从短期整治向长效管护转变。

（一）新技术新应用衍生风险与防控难点

生成式AI等数字技术广泛普及应用，网络失信行为呈现出高技术化、隐蔽化、跨界化、产业化等特征，对深化网络诚信建设提出一系列新挑战。

生成式AI冲击知识产权规则，引发内容生态源头性诚

信失范。生成式 AI 打破传统内容生产与确权逻辑，在创作源头即衍生出大量侵权与造假行为，严重破坏内容领域的诚信根基。AI “洗稿”、抄袭行为泛滥，技术化“伪原创”手段使得侵权内容更具隐蔽性，严重挤压原创者生存空间，形成“抄袭获利、原创吃亏”不良导向。一些不法分子利用 AI 伪造专利、著作权、学术论文、科研数据等用于各类资质申报与商业活动，直接冲击科研诚信与商业信用体系。当前，针对 AI 生成内容的知识产权归属、侵权责任界定仍存在法律空白，开发者、平台、使用者责任边界模糊，导致权利人维权面临举证难、追责难困境，源头性失范侵蚀原创生态，所生成大量低质、虚假内容，滋生助长更深层次的网络失信行为。

AI 与网络黑灰产深度融合，致使失信行为扩散蔓延。AI 技术深度嵌入网络黑灰产各个环节，失信行为从单一环节演变为“技术供给—内容生成—引流变现—规避监管”恶性闭环。在互联网金融领域，基于深度伪造（Deepfake）的 AI “换脸”与拟声技术，使冒充熟人、领导的诈骗手段极具迷惑性，大幅降低犯罪门槛；^①在电子商务领域，AI 被滥用于批量生成虚假商品测评、用户好评与使用效果图，构建虚假信用评价体系，^②严重误导消费者决策；在舆论生态领域，AI “水军”自动化内容生成、矩阵式分发与全天候运营，恶

^① 以假乱真的 AI 诈骗，如何防范？（民生一线）.张天培、苏滨

[EB/OL].2025-04-08[2026-03-09].<http://society.people.com.cn/n1/2025/0408/c1008-40455106.html>.

^② 张雪泓.不法商家利用 AI 制作虚假广告北京互联网法院提醒消费者警惕新型网络消费虚假宣传 [EB/OL].2025-03-27[2026-03-09].https://legal.gmw.cn/2025-03/27/content_37930436.htm.

意诋毁竞争对手，^①利用动态调整生成内容、更换 IP 地址、虚拟主体身份等手段规避监管，导致传统监管手段面临发现难、溯源难、取证难、处置难等挑战，黑灰产所生成的虚假内容与诈骗信息，传播广度与社会危害程度，往往高度依赖算法推荐机制。

算法滥用使诚信失范行为更趋隐蔽。黑灰产聚焦于虚假内容“生产端”，算法机制则构成虚假内容“分发端”与“放大端”，进一步加剧网络失信治理的复杂性。部分平台利用算法实施价格欺诈、差异化定价、大数据杀熟等行为，通过动态调整商品价格、对不同用户差异化展示优惠力度等方式侵害消费者合法权益，消费者难以识别价格真实性，维权举证难度极大；部分内容平台为追求流量收益，通过算法推荐机制助推虚假内容、网络暴力信息、低俗不良信息扩散，形成“不良内容—流量变现—算法助推”恶性循环；部分企业利用算法黑箱规避监管，隐瞒算法推荐逻辑与数据使用规则，算法透明度不足、算法诚信核验标准缺失、算法责任界定不清等问题较为突出，难以遏制算法滥用本身引发的诚信失范，使得依赖算法分发的各类虚假内容与欺诈信息得以持续蔓延。

（二）跨域网络诚信治理的协同联动难题

当前，网络失信行为呈现出跨平台流转、跨地域扩散、跨领域渗透、跨境传导等显著特征，成为制约网上网下协同

^① 王潇. “水军”正在借助 AI 进化：谁在操控你的认知？
[EB/OL].2025-03-15[2026-03-09].<https://export.shobserver.com/baijiahao/html/869494.html>.

建设、跨平台跨领域协同治理的难题。

网络失信行为跨平台“换马甲重生”问题仍然频发。部分失信主体采用矩阵式账号布局，在一平台被处置后，通过变更主体与平台信息，进入其他平台继续违规活动，“一处失信、处处受限”的协同治理效果不足。网站平台间账号管理、处置标准与失信名单机制互不联通，常态化信息共享与联合处置机制尚未完全建立，部分中小平台为争夺市场，对失信主体入驻审核宽松，导致失信行为跨平台转移与扩散。

《关于健全社会信用体系的意见》提出“引导平台企业建立平台内信用管理制度和平台间失信联合约束制度”“加强对网络主播、自媒体、MCN 机构等信用监管”等要求，采取专项行动处置大量违规账号与机构，仍然很难破解平台间“信息孤岛”与标准不一、协同机制失灵等问题。

线上线下治理衔接不畅。^①网络诚信治理涉及网信、公安、市场监管、金融监管、广电、工业和信息化等多个职能部门，覆盖平台运营、内容生产、商业交易、金融服务等多个环节。在工作实践中，仍存在职责边界不清、监管权限交叉重叠、信息共享渠道不畅、执法协同效率不高等问题。电子商务领域虚假宣传、价格欺诈等失信行为，往往涉及平台属地、商家经营地、广告发布地、消费者所在地等多个地域，跨地域执法协作成本高、流程长、响应慢；互联网金融领域非法荐股、虚假借贷、跨境金融诈骗等失信行为，涉及网信、

^① 林瑜胜.推进网上网下协同治理
[EB/OL].2026-02-04[2026-03-10].https://www.studytimes.cn/jjsh/202602/t20260203_86090.html.

公安、金融监管等多个部门，常态化联合执法机制不健全，线索移送、证据固定、案件查办的衔接流程不畅。同时，大量线上违法失信线索，因线下主体核查难、落地处置难，难以实现“线上发现、线下查处”全闭环治理。

跨境治理防控难度持续加大。2025年，网络失信行为的跨境特征愈发明显，部分网络黑灰产将技术研发、服务器部署、运营主体设置在境外，通过跨境网络链路针对境内网民实施失信违法活动。境外操控AI水军集群批量散布虚假信息、煽动舆论对立，严重扰乱网络舆论生态；跨境网络诈骗团伙利用AI深度伪造技术实施精准诈骗，资金流转通过跨境渠道规避监管，溯源打击难度极大；境外虚拟货币交易炒作平台通过跨境引流、线下兑换等方式规避监管，诱导境内网民参与高风险交易，引发金融诚信风险；境外AI服务提供商未经合规审批，向境内用户提供生成式AI服务，成为虚假信息生成传播的重要源头。不同国家和地区的法律体系、监管标准、执法流程存在显著差异，跨境证据调取、执法协作、失信主体联合惩戒渠道不畅、流程繁琐、周期漫长，^①网络诚信风险跨境传导的防控压力持续加大。

（三）长效治理机制的短板与适配性不足

当前，网络诚信建设常态化、制度化、规范化的长效治理机制作用发挥仍有不足，亟须从信用体系、责任体系、共

^① 梁玮.区块链技术赋能与跨境新技术犯罪风险防控的司法平衡 [EB/OL].2025-10-11[2026-03-10].https://mp.weixin.qq.com/s?__biz=MzU2Nzk1NjA2Ng==&mid=2247525522&idx=3&sn=1d7f25ed9d0086490472255586bf43c6&chksm=fddfc08e671ac8153bccb1ff59685909cae63ba19f66f63a00b24f1ed8fa703fdff32fdfa0#rd.

治体系等多维度提升适配能力。

“守信激励、失信惩戒”机制支撑不足，信用评价标准不统一，各领域、各平台信用评价指标、评分规则互不兼容，跨领域、跨平台信用互认存在制度性与技术性障碍。信用信息共享壁垒突出，政务信用数据与平台企业掌握的交易行为、违规记录等市场信用数据互联互通不足，全国信用信息共享平台的网络领域信用数据归集覆盖面仍有明显缺口。分级分类监管精准度不足，差异化监管措施多停留在程序性层面，与网络经济特点适配的激励约束机制不完善，信用修复、联合惩戒的标准规范不统一，线上线下信用体系融合不够深入，对平台经济、直播电商等新业态的常态化、高效协同办案机制仍需加强。信用制度的碎片化与信息壁垒，在一定程度上削弱了对平台企业合规经营的外部约束力。

平台合规管理的内生动力与制度约束仍需强化，内生治理动力与能力存在结构性失衡。部分平台存在“重流量、轻合规”“重收益、轻责任”的价值偏差，对平台内虚假宣传、网络谣言、网络暴力等失信行为管控宽松，甚至通过算法推荐助推失信内容扩散。平台间治理能力不均衡问题突出，头部平台技术治理能力相对完善，中小平台在 AI 虚假内容识别、违规行为溯源等方面技术投入不足、管控能力薄弱。对 MCN 机构、入驻商家、账号主体的全生命周期管理机制不健全，资质审核不严、日常监管缺位、违规处置不力等问题普遍存在，平台合规管理的内生动力与制度约束仍需持续强

化。平台作为网络空间的第一道关口，其治理动力的提升与能力的完善，直接关系到前端信用制度的激励约束能否有效落地，也影响后端社会监督的压力传导。

社会共治格局尚未完全成型，全民网络诚信素养亟待持续提升。网络诚信建设多元共治体系有待完善，行业协会自律作用发挥不够充分，现有行业自律公约多为倡导性条款，强制性约束相对不足，对网络诚信新兴领域的自律规范覆盖有待拓展。社会监督渠道不够畅通便捷，举报受理、核查反馈、激励保障的全链条机制不够健全，社会监督的实效性有待进一步提升。全民网络诚信素养教育体系不够完善，宣传教育多为阶段性活动，常态化、体系化培育机制尚未完全建立，针对 AI 深度伪造等新型失信行为的辨别能力教育覆盖有待加强，部分网民网络诚信意识仍需增强，全社会共建共治共享的治理合力有待进一步释放。

总的来看，网络诚信治理的难点已从早期的“技术监管能力不足”逐步转向“制度协同效能不彰”。面对生成式 AI 等新技术带来的诚信风险，单一环节的治理努力往往难以完全覆盖问题的复杂性，需从更深层次破解制度性障碍。

二、发展建议

为进一步提升网络诚信治理体系和治理能力现代化水平，有效应对生成式人工智能、数据安全、新业态新模式带来的诚信风险，破解网络失信行为跨平台、跨地域、跨境传播等突出问题，需坚持系统思维、底线思维与创新思维相统

一，从技术监管、协同治理、长效建设三个维度综合施策。通过健全制度规范、强化技术赋能、完善联动机制、压实主体责任、厚植诚信文化，加快构建全方位、多层次、立体化的网络诚信建设格局，维护清朗网络空间，为数字经济健康发展提供坚实保障。

（一）规范新技术新应用管理，化解技术带来的风险

面对技术变革带来的新挑战，必须坚持技术治网与技术向善相统一，强化技术监管能力建设，建立适应新技术发展特点的诚信治理机制，将风险挑战控制在初步状态。

构建生成式人工智能全生命周期诚信治理体系。在现有《生成式人工智能服务管理暂行办法》《互联网信息服务深度合成管理规定》等生成式人工智能领域相关法律法规框架下，推动细化关于诚信建设内容条例，明确人工智能研发者、使用者、服务提供者的诚信责任和义务，建立涉及诚信算法备案、伦理审查、安全评估等制度。推动 AI 监管科技在人工智能治理中的应用，积极研发针对深度伪造、虚假信息生成的自动化检测工具，提升对技术性失信行为的发现和取证能力。推动建立算法透明度监管机制，要求对涉及公共利益和个人权益的关键算法进行逻辑公开或提供解释说明，打破“算法黑箱”。例如，全面深化算法备案制度，进一步推动所有日活用户超过一定规模的推荐算法必须备案并公示其基本原理、目的意图及运行机制，接受公众监督。

完善数据安全与隐私保护的信用监管机制。贯彻落实

《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》，推动数据合规情况纳入企业信用评价体系。探索建立数据流通交易信用档案，对数据交易主体进行信用分级分类管理，严惩数据违规采集、买卖、泄露等失信行为。推广隐私计算、区块链等技术应用，构建数据“可用不可见”的可信流通环境，从技术源头保障数据安全。例如，在金融领域推动“隐私计算+信用评估”试点，通过多方安全计算技术，在不交换原始数据的前提下，实现银行、税务等部门间的数据融合，精准刻画企业信用画像，既解决“数据孤岛”难题，又筑牢了隐私保护防线。

探索适应新业态新模式的包容审慎监管模式。针对元宇宙、Web3.0等新兴领域，坚持“鼓励创新、包容审慎”原则，划定网络诚信底线和红线，预留发展空间。推广“沙盒监管”机制，在可控范围内鼓励企业开展新技术新应用试点，对轻微失信行为以指导整改为主，避免“一刀切”式执法。建立新业态信用风险预警机制，加强对新业态发展动态的跟踪监测，及时发现苗头性、倾向性问题，快速响应、精准施策。

（二）健全跨域联动机制，提升多维度协同治理效能

针对网络失信行为跨平台、跨地域、跨境流动的特征，必须打破条块分割壁垒，构建协同治理体系，形成“一处失信、处处受限”的治理合力。

构建全国一体化的网络信用信息共享平台。指导互联网平台企业建立统一的信用数据交换标准与接口，推动账号主

体信息、违规失信记录、处置结果等数据在平台间的互联互通。充分发挥行业协会桥梁纽带作用，制定跨平台协同治理公约，明确平台间信息共享、线索移送、联合处置的权利义务。进一步发挥平台经济领域信用建设合作机制作用，组织头部平台开展联合打击行动，共享黑名单库与恶意特征库，形成打击网络失信行为的行业合力。加强对中小平台的监管引导，针对部分中小平台成为失信“避风港”的问题，加大监管执法力度，督促其严格落实主体责任，提升合规管理水平，通过技术赋能、培训指导等方式，提升中小平台治理能力，消除治理短板。

完善跨区域跨部门联合执法协调机制。建议持续完善相关部际联席会议，进一步厘清网信、公安、市场监管、金融监管等部门职责边界，建立权责清晰、衔接紧密的监管链条。健全线索双向移送、证据互认、联合办案制度，针对互联网金融、网络直播、电子商务等重点领域开展常态化联合执法行动，消除监管真空。优化跨地域治理流程，探索建立“云端立案、异地协查、结果互认”的跨地域执法模式，利用数字化手段降低跨地域执法成本，提高响应速度。针对跨地域网络诈骗、网络传销等顽疾，建立重点案件挂牌督办与区域联动处置机制，强化线上线下治理衔接，推动实现“线上发现、线下查处”，对线上发现的违法线索，及时通过线下核查落地处置；对线下查处的失信主体，同步实施线上经营活动限制。

持续深化跨境网络诚信国际合作。积极参与网络空间国际治理，在国际互联网治理框架下，主动发声，倡导构建多边、民主、透明的网络诚信国际规则体系。强化跨境风险防控，加强对境外操控的 AI 水军、跨境网络诈骗、虚拟货币炒作等风险的监测预警，建立跨境失信风险“防火墙”。联合金融机构、支付平台加强对跨境资金流向的监控，阻断网络黑灰产跨境资金链。推动跨境平台协同共治，督促跨境电子商务平台、社交平台切实履行主体责任，建立跨境消费纠纷快速解决机制，保障国内外消费者合法权益。加强与国际知名互联网企业的沟通协作，共同打击跨境虚假信息、网络暴力等失信行为，营造清朗的跨境网络空间。

（三）凝聚多方合力，促进网上网下诚信有机融合

坚持标本兼治、重在治本，从制度体系、信用机制、社会文化等方面入手，夯实网络诚信建设的根基，实现治理效能的常态化、长效化。

完善网络信用体系建设。推动制定网络信用信息采集、分类、评价、共享标准，消除各领域、各平台间的制度性与技术性障碍。探索网络信用评价与社会信用体系有效融合，实现信用信息跨领域互认。强化信用信息共享与应用，探索政务信用信息与平台市场信用信息共享渠道，扩大全国信用信息共享平台网络领域数据归集覆盖面。规范信用修复机制，明确网络失信行为信用修复的条件、程序与标准，建立便捷、高效的信用修复渠道，如开设“信用修复课题”，通过参加

诚信教育培训、签署信用承诺书等方式，鼓励失信主体主动纠错、重塑信用，营造包容审慎的监管环境。

压实平台企业主体责任。持续强化平台合规管理体系建设，督促大中小平台企业建立健全内部诚信管理制度，推动平台将信用要求嵌入产品设计、算法推荐、内容审核全流程，从源头防范失信风险。加强对 MCN 机构、入驻商家等主体的管理，指导平台完善入驻资质审核与日常巡查机制，对 MCN 机构实行信用分级管理，严控违规账号“复活”等问题。推动建立平台治理考核评价机制，对平台网络诚信治理成效进行评估。

构建多元共治社会生态。积极发挥行业组织引领作用，支持行业协会制定推广网络诚信自律公约，填补新兴领域规范空白。鼓励行业协会开展诚信教育培训、信用评价、纠纷调解等工作，提升行业自律水平。加强网络诚信文化建设，将网络诚信教育纳入培训内容，常态化开展网络诚信宣传进社区、进企业、进校园活动。针对 AI 深度伪造、网络诈骗等新型风险，普及科技知识与防骗技能，提升全民网络诚信素养与辨别能力，在全社会形成“崇尚诚信、践行诚信”的良好风尚。