

DOI: 10.53104/fxsy.2026.02.04.005

网络暴力治理中的平台责任边界

陈宇航¹, 李悦¹

1. 浙江工商大学, 浙江 杭州 310018

摘要: 网络空间已经成为公众交流、信息传播和社会互动的重要场域。开放性、即时性和裂变式传播特点在提升信息流动效率的同时, 也使网络暴力问题更加突出。网络暴力不同于传统侵权行为, 其形成通常涉及内容发布者、传播参与者、网络平台以及技术机制等多个因素。网络平台作为信息传播的重要组织者, 掌握内容分发、用户管理和技术治理能力, 在网络暴力治理过程中承担着重要责任。当前, 平台责任已经从单纯的信息服务义务逐渐转向网络空间治理义务, 但平台应承担何种范围的责任仍存在争议。平台责任过轻, 难以及时阻止网络暴力扩散; 责任过重, 又可能导致平台采取过度限制措施, 影响正常网络表达。如何在保护用户人格权益、维护网络秩序与保障表达自由之间形成合理平衡, 成为网络暴力治理中的重要法律问题。本文围绕网络暴力治理中的平台责任边界展开分析, 在梳理网络暴力传播特点和平台治理基础的基础上, 探讨平台在内容审核、用户管理、个人信息保护等方面承担的具体责任, 并从平台控制能力、风险发现可能性、损害后果程度以及责任类型区分等角度分析平台责任边界。文章认为, 平台责任应建立在合理注意义务基础上, 根据平台实际治理能力和风险防控能力确定责任范围, 避免将平台作为网络暴力后果的替代承担者。在制度完善方面, 应进一步优化平台治理规则, 完善责任分层机制, 推动网络暴力治理由事后处置向全过程风险防控转变。

关键字: 网络暴力; 网络平台; 平台责任; 责任边界; 网络治理

The Boundary of Platform Responsibility in the Governance of Cyber Violence

CHEN Yu-hang¹, LI Yue¹

1. Zhejiang Gongshang University, Hangzhou 310018, P. R. China

Correspondence to: LI Yue

Abstract: The internet has become an important space for public communication, information dissemination, and social interaction. While the openness, immediacy, and viral nature of online communication have improved the efficiency of information exchange, they have also intensified the problem of cyber violence. Unlike traditional forms of infringement, cyber violence is usually formed through the interaction of multiple factors, including content creators, participants in information dissemination, online platforms, and technological mechanisms. As important organizers of online information circulation, platforms possess capabilities in content distribution, user management, and technological governance, and thus undertake significant responsibilities in the governance of cyber violence. Currently, platform responsibility has gradually shifted from simple information service obligations toward broader obligations of maintaining order in cyberspace, yet the appropriate scope of such responsibility remains controversial. Insufficient platform responsibility may make it

引用格式: 陈宇航, 李悦. 网络暴力治理中的平台责任边界[J]. 法学视野, 2026, 2(4): 34-44.

difficult to prevent the spread of cyber violence, while excessive responsibility may lead platforms to adopt overly restrictive measures that affect normal online expression. Achieving a reasonable balance between protecting users' personality rights, maintaining online order, and safeguarding freedom of expression has become an important legal issue in cyber violence governance. This paper analyzes the boundaries of platform responsibility in the governance of cyber violence. Based on an examination of the characteristics of cyber violence dissemination and the foundations of platform governance, it discusses the specific responsibilities of platforms in content moderation, user management, and personal information protection. It further explores the boundary of platform responsibility from the perspectives of platform control capacity, possibility of risk detection, severity of damage consequences, and differentiation of legal liabilities. The paper argues that platform responsibility should be based on reasonable duty of care, with the scope of responsibility determined according to platforms' actual governance capabilities and risk prevention abilities, so as to avoid treating platforms as substitutes for the direct perpetrators of cyber violence. In terms of institutional improvement, it is necessary to optimize platform governance rules, establish a differentiated responsibility system, and promote the transformation of cyber violence governance from post-event response to comprehensive risk prevention.

Key words: cyber violence; online platforms; platform responsibility; responsibility boundaries; cyber governance

引言

互联网技术的发展改变了社会信息传播方式,也改变了公众参与社会交流的路径。社交媒体、短视频平台以及网络社区降低了信息发布门槛,使个人能够快速参与公共讨论。网络空间开放性和互动性的增强,也使部分社会矛盾、情绪冲突和意见分歧更容易集中呈现,并进一步演化为针对特定个人的网络攻击行为。

网络暴力通常表现为侮辱、诽谤、恶意评论、隐私曝光以及群体性攻击等形式。与传统侵权行为相比,网络暴力具有传播速度快、影响范围广、行为主体分散等特点。受害者面对的不再是单一侵权主体,而可能是大量匿名用户形成的持续性攻击。这种传播模式增加了责任认定和权利救济的难度,也对传统法律治理方式提出新的要求。

在网络暴力事件的发展过程中,网络平台发挥着特殊作用。平台不仅提供信息发布和交流空间,也通过内容审核、算法推荐、账号管理和社区规则影响信息传播过程。当攻击性内容经过平台机制持续扩散时,平台是否需要承担相应责任,成为当前网络治理领域的重要问题。

马永强指出,网络暴力的生成受到平台结

构、现实社会冲突网络化以及网络身份关联等因素影响,仅依靠技术治理难以完全解决问题,需要通过法律治理和多元协同机制进行规制^[1]。这一观点说明,网络暴力并非单纯由用户行为造成,其形成过程与平台运行机制存在密切联系。

近年来,关于网络暴力治理的讨论逐渐从行为人责任追究延伸至平台责任界定。平台拥有较强的信息管理能力和技术控制能力,在风险发现、内容处理和用户管理方面具有其他主体无法替代的优势。但平台责任强化并不意味着平台需要承担所有网络行为产生的后果。网络空间中的表达具有复杂性,部分争议性言论、负面评价和观点冲突并不当然构成网络暴力。

如果平台承担过高治理责任,可能导致平台为了规避风险而限制正常表达;如果平台责任不足,又可能造成违法信息持续传播,使受害者难以及时获得保护。因此,明确网络暴力治理中的平台责任边界,需要在平台治理能力、用户行为责任以及权利保护之间建立合理联系。

本文以网络暴力治理中的平台责任为研究对象,围绕平台责任产生依据、具体内容、现实困境和边界判断展开分析,探讨平台如何在法

律框架内履行治理义务，并在保护网络秩序与维护表达自由之间实现平衡。

1 网络暴力治理中平台责任的现实基础

网络暴力并不是单纯的信息违法传播行为，而是在技术环境、用户互动和平台机制共同作用下形成的复杂现象。网络传播方式改变了传统信息流动模式，使个体发布的信息能够在短时间内获得大量关注。某一具有攻击性的言论经过评论、转发和推荐后，可能迅速形成群体性讨论，并扩大对受害者的影响。

网络平台在这一过程中逐渐成为信息传播的重要组织者。早期互联网平台主要承担信息存储和技术服务功能，但随着社交平台、内容平台不断发展，平台已经具备规则制定、信息管理和用户约束能力。平台决定用户如何参与互动，也影响信息如何被传播和呈现。

网络暴力治理中的平台责任首先来源于平台对网络空间具有一定控制能力。平台掌握用户账号体系、内容审核工具以及数据分析能力，能够通过技术手段和管理措施降低违法信息扩散风险。郭芙蓉、计玲芳指出，数智技术环境下网络话语暴力呈现出对象范围扩大、施暴者身份难以确定、情绪传播增强等特点，技术发展在提高传播效率的同时，也增加了治理复杂性^[2]。从网络治理实践来看，平台已经不再是完全中立的信息传递者。平台通过推荐机制影响内容曝光程度，通过社区规则影响用户行为，通过账号管理维护网络秩序。这些管理能力使平台获得了较大的网络空间影响力，也要求其承担与能力相匹配的治理责任。平台责任的承担并不意味着平台需要替代用户承担全部法律后果。网络暴力的直接实施者仍然是发布攻击内容、传播违法信息的用户，平台责任主要体现为对自身控制范围内风险的预防和处置义务。如果平台已经具备发现风险和采取措施的能力，却长期放任相关内容传播，则可能需要承担相应责任。从法律关系来看，平台责任实际上是一种风险防控责任。平台并非网络暴力行为的直接实施主体，但由于其处于信息传播链条的重要位置，对违法行为具有一定预防和控

制能力。因此，平台责任的合理确定，需要以平台能力和实际影响为基础，而不能简单依据损害结果进行判断。

2 网络暴力治理中平台责任的理论基础

网络平台责任的形成基础，在于平台权力与治理义务之间的对应关系。平台拥有较强的信息组织能力，也承担维护网络生态秩序的责任。随着平台功能不断扩展，单纯将平台视为技术服务提供者已经难以解释其在网络治理中的现实作用。

平台责任首先体现为技术控制能力与风险承担之间的联系。网络暴力虽然由用户行为直接产生，但传播范围和影响程度往往受到平台机制影响。平台能够通过内容审核、传播限制和账号管理降低风险，这种能力构成平台承担治理责任的重要基础。从平台运行方式来看，平台具有类似网络空间“守门人”的作用。平台通过制定规则、管理用户和调整信息传播路径，对网络环境产生实际影响。赵凌认为，网络暴力治理中的平台责任需要结合平台自身能力和治理优势进行完善，使平台从被动处理转向主动承担治理义务^[3]。平台责任还需要受到法律边界限制。平台并不能对所有网络表达进行全面审查，也不能因为网络暴力发生在平台内部，就承担全部损害责任。网络信息数量庞大，表达形式复杂，要求平台提前识别所有潜在风险并不符合现实情况。合理的平台责任应建立在注意义务基础上，即判断平台是否具备发现风险的条件，是否采取了合理措施，以及未采取措施是否扩大了损害后果。只有当平台具有治理能力却未履行必要义务时，才具有责任追究的正当性。从责任类型来看，平台责任具有层次性。一般网络侵权可能涉及民事责任；违反网络安全管理义务可能涉及行政责任；严重违法定义义务并造成重大危害后果的情形，才可能进入刑事责任评价范围。不同责任类型对应不同违法程度，需要根据具体情况进行区分。

明确平台责任的理论基础，有助于避免两个极端。一种情况是认为平台完全没有责任，将所有问题归结于用户个人；另一种情况是认

为平台应对所有网络风险负责。合理的平台责任体系，应当在平台治理能力与用户行为责任之间建立平衡。

3 网络暴力治理中平台责任的具体内容

网络暴力治理中的平台责任并不是抽象的治理要求，而是体现在内容管理、用户管理以及个人信息保护等具体环节。平台承担相关责任的核心目的，是利用自身技术能力和管理优势降低网络暴力风险，而不是替代行为实施者承担全部法律责任。

3.1 平台的信息审核与风险识别责任

内容审核是平台治理网络暴力的重要环节。网络暴力通常具有明显的扩散过程，攻击性内容在最初出现时可能影响范围有限，但经过评论、转发和推荐后，容易形成持续传播。如果平台能够及时发现具有明显违法风险的信息，并采取限制传播、删除内容等措施，可以降低损害进一步扩大的可能。平台审核主要依靠人工审核与技术审核相结合的方式。人工审核能够结合具体语境判断内容性质，对于隐晦表达、情绪攻击以及复杂争议具有较强识别能力；技术审核则能够处理海量信息，提高风险发现效率。但网络语言具有较强变化性，一些网络暴力行为并不会直接表现为明显侮辱性词汇，而是通过暗示、隐喻、群体标签等方式实施攻击，单纯依靠关键词识别难以及时发现。

网络暴力治理中的审核责任，应当理解为平台合理范围内的风险识别义务，而不是绝对的信息过滤义务。平台无法提前预测所有用户行为，也无法完全消除网络空间中的冲突表达。判断平台是否履行审核责任，需要结合平台规模、技术能力、信息传播程度以及是否收到有效投诉等因素。对于已经形成明显风险的信息，平台应当采取更积极的处理措施。例如，针对集中辱骂、恶意造谣、持续攻击特定个人等行为，平台应及时降低传播范围，并根据用户违规情况采取警告、限制功能或者账号处置措施。

3.2 平台的用户管理与传播控制责任

网络暴力通常具有群体化特点，单个用户

的攻击行为经过大量用户参与后，可能形成持续性的网络围攻。因此，平台治理不能仅关注单条内容，还需要关注用户行为模式以及信息传播过程。用户管理责任主要体现为账号管理、异常行为识别和违规处置。平台可以通过完善账号规则、限制恶意账号活动、识别异常互动行为等方式降低网络暴力发生概率。例如，短时间内大量账号集中发布相似攻击内容，或者某一用户持续针对特定对象进行骚扰，都可能属于需要平台关注的风险行为。

网络暴力中的组织化传播使传统责任追究方式面临困难。李航宇、刘素霞指出，网络暴力具有群体参与和精神损害特点，容易出现责任主体分散、行为难以对应的问题，需要通过更加完善的治理机制提高规制效果^[4]。平台作为信息传播的重要节点，可以通过技术手段识别异常传播路径，在一定程度上弥补传统治理方式的不足。但平台用户管理责任也需要保持合理范围。网络空间中的争论、批评和意见表达并不当然属于网络暴力。平台不能仅因为内容存在负面情绪或者观点冲突，就采取限制措施。只有当行为具有明显攻击性、违法性或者造成较大损害风险时，平台治理措施才具有正当性。

3.3 平台的个人信息保护责任

个人信息泄露是网络暴力中较为严重的一种表现形式。部分网络暴力行为并不限于语言攻击，而是通过公开个人身份信息、工作信息、家庭信息等方式扩大伤害范围，使受害者面临网络空间和现实生活的双重压力。网络“开盒”行为就是典型表现。行为人通过非法方式获取并传播个人信息，使网络攻击从言语层面延伸至现实风险。孙道萃指出，网络“开盒”行为本质上涉及非法公开公民个人信息，部分情形可能伴随网络暴力，并需要根据具体行为承担相应法律责任^[5]。在此类情形中，平台承担的信息保护责任主要包括两个方面。平台需要加强自身数据安全保护，防止用户信息因系统漏洞或者管理不足发生泄露。平台需要在发现非法公开个人信息行为后及时采取措施，避免相关信息进一步传播。个人信息保护责任说明，平台治理网络暴力不能局限于删除攻击性言论。网络暴力背后的信息泄露、身份暴露和现实风险，同样属于平台需要关注的治理范围。

3.4 网络暴力事件中平台责任的实践考察

网络暴力中的平台责任并非停留在理论层面,近年来发生的一些网络事件表明,平台在信息传播、用户管理以及风险处置过程中发挥着重要作用。通过具体案例分析,可以进一步观察平台责任的实际表现,并明确平台治理义务的边界。某网络事件中,个体因社会热点话题受到大量网友关注,相关讨论逐渐从事件本身转向对当事人的个人评价。随着话题热度提升,大量用户通过评论、转发等方式发表攻击性言论,部分内容涉及侮辱、恶意猜测甚至个人信息传播。事件发展过程中,平台成为信息扩散的重要渠道。一些攻击性内容在短时间内获得较高曝光,并形成持续性讨论,对当事人的社会评价和正常生活造成影响。

从平台责任角度来看,该类事件反映出平台内容审核和传播控制义务的重要性。网络暴力并非完全由某一条违法信息造成,而是在大量互动行为共同作用下形成。当平台发现相关内容已经出现明显攻击倾向,或者相关话题已经形成集中性攻击时,应当通过限制传播、处理违规账号、优化推荐机制等方式降低风险。该类事件也说明,平台责任不能简单理解为删除违法内容。网络暴力具有明显的传播链条特征,单纯处理单条攻击性言论,可能无法阻止后续扩散。例如,部分用户可能通过修改表达方式、重新发布内容继续实施攻击,如果平台缺少针对异常传播行为的管理机制,治理效果仍然有限。

近年来,“开盒”行为成为网络暴力治理中的典型问题。部分行为人通过非法获取并公开他人身份信息,引发大量用户围绕受害者展开攻击。与一般网络争议相比,“开盒”行为突破了网络言论范围,将个人隐私暴露和现实风险联系起来。孙道萃指出,网络“开盒”行为的核心在于非法公开公民个人信息,部分情况下会与侮辱、诽谤等网络暴力行为结合,需要结合具体行为进行法律评价^[6]。在“开盒”类网络暴力中,平台责任主要体现为个人信息保护和风险处置义务。平台需要防止用户信息被非法传播,并在发现相关内容后及时采取措施。如果平台已经收到举报或者能够明显识别相关信息具有侵害风险,却未及时处理,则可

能扩大损害结果。

从上述案例可以看出,平台责任并不是要求平台对所有网络争议进行主动干预,而是要求平台在具有治理能力和风险识别条件时,采取合理措施降低损害。网络暴力治理中的平台责任,应当围绕平台能够控制的传播环节展开。对于用户实施的直接侵权行为,责任主体仍然是行为人;对于平台未履行合理管理义务导致风险扩大,则应根据具体情况承担相应责任。

4 网络暴力治理中平台责任承担的现实困境

网络平台已经成为网络暴力治理的重要主体,但平台责任的落实仍面临规则、技术和实践层面的多重障碍。平台需要在快速处理违法信息与保障正常网络表达之间进行平衡,也需要在用户规模扩大、传播机制复杂化的情况下承担合理治理义务。当前平台责任困境并非来源于平台缺少治理意愿,而是受到责任标准不明确、治理能力有限以及责任协调不足等因素影响。

4.1 网络暴力认定标准不明确,平台审核面临判断困难

网络暴力与正常网络表达之间并不存在绝对清晰的界限,这是平台治理面临的首要问题。在实际网络环境中,部分内容可能同时具有批评、讽刺、情绪表达等多重属性。例如,用户针对社会事件发表激烈评论,可能属于正常意见表达;但如果表达对象转向特定个人,并伴随侮辱、造谣或者持续攻击,则可能演变为网络暴力。二者在语言形式上可能较为接近,平台需要结合具体语境、传播范围和行为目的进行判断。

现有平台治理规则通常依靠社区规范、关键词识别以及用户举报进行处理,但这些方式难以完全解决复杂情境下的判断问题。关键词审核容易忽略语境差异,例如某些具有攻击性的表达可能通过谐音、隐喻或者网络用语进行包装;单纯依据用户举报处理,也可能受到群体情绪影响,使平台难以准确判断内容性质。

网络暴力认定标准的不明确,使平台容易

面临两种治理风险。一种是治理不足，对明显攻击性内容处理不及时，导致受害者权益持续受到侵害；另一种是治理过度，将正常争议和负面评价纳入限制范围，影响用户正常表达。因此，平台审核需要从单纯判断文字内容，转向综合考察行为背景、传播方式和实际影响。网络暴力治理不能只关注某一句话是否具有攻击性，也需要判断该行为是否已经形成持续性侵害。

4.2 算法推荐机制可能扩大网络暴力传播影响

当前网络平台已经不再只是信息储存空间，而是通过算法推荐影响信息传播范围。算法机制能够根据用户兴趣、互动程度和传播热度调整内容曝光，这种机制提高了信息传播效率，但也可能放大具有强烈情绪色彩的内容。网络暴力事件中，攻击性内容往往具有较高互动性。争议、愤怒和冲突容易吸引用户参与，相关内容可能因为较高点击量和讨论度获得更多推荐。如果平台算法过度关注流量指标，可能使具有攻击倾向的信息获得更大传播空间。这使平台责任问题从传统的“是否删除违法内容”扩展到“是否影响内容传播”。如果平台只是被动处理用户发布的信息，责任范围相对有限；但如果平台通过推荐机制持续扩大某类攻击性内容的影响，则需要进一步考察平台是否履行了传播风险控制义务。

当然，算法推荐并不意味着平台需要对所有推荐结果承担责任。算法运行具有复杂性，平台无法完全预测每一条内容可能造成的社会影响。问题的关键在于，平台是否建立了合理的算法治理机制，例如是否对异常传播趋势进行识别，是否降低明显具有攻击风险内容的扩散范围。郭芙蓉、计玲芳指出，数智技术环境下网络话语暴力呈现出传播速度提升、对象泛化以及责任认定困难等特点，技术发展在推动信息流动的同时，也增加了治理难度^[2]。这说明，平台技术能力既是治理优势，也是责任来源。

4.3 平台用户管理能力有限，难以及时控制群体性攻击

网络暴力通常不是单一用户行为，而是多个用户共同参与形成的群体性攻击。在部分事件中，最初发布攻击内容的用户可能只是少数，

但经过转发、评论和话题聚集后，会形成大量用户参与的攻击环境。这种特点增加了平台治理难度。平台可以删除某条违法内容，但难以完全阻止类似内容通过不同账号重复出现。部分网络暴力行为还具有组织化特点，用户可能通过群组、私信或者其他隐蔽方式协调攻击行为，使平台难以及时发现。

李航宇、刘素霞指出，网络暴力具有群体参与特征，传统责任追究容易面临“法不责众”的困境，需要进一步完善治理方式^[4]。对于平台而言，群体性攻击意味着治理对象已经从单个内容转向用户关系和传播网络。

目前部分平台主要依靠账号处罚处理违规行为，但账号处罚存在一定局限。例如，被限制的账号可能重新注册，攻击行为可能转移至其他平台。如果缺少跨环节风险识别机制，平台治理效果容易受到影响。因此，平台用户管理不能仅依靠事后处罚，还需要关注异常行为模式。例如短时间大量账号围绕同一对象发布攻击内容、相似文本集中传播等情况，都可能成为平台判断风险的重要依据。

4.4 平台责任强化与表达自由保护之间存在冲突

网络暴力治理最终需要面对一个价值平衡问题，即如何在保护人格权益和保障表达自由之间寻找合理范围。网络空间的重要功能在于提供公共讨论渠道。用户对社会事件发表意见、对公共行为进行评价、表达个人情绪，都属于网络交流的重要组成部分。如果平台责任设置过高，平台可能为了降低法律风险而采取保守治理策略，对大量具有争议但并不违法的内容进行限制。这种情况可能导致“寒蝉效应”。用户担心表达受到处罚，减少参与公共讨论，网络空间的开放性受到影响。但表达自由也并非没有边界。针对特定个人进行持续侮辱、造谣、曝光隐私信息等行为，已经超出正常表达范围。平台治理的目标不是限制所有负面表达，而是识别并处理已经侵害他人合法权益的行为。平台责任边界的确定，需要建立在区分正常表达与网络暴力的基础上。只有明确治理对象，才能避免平台责任扩大化，也能够保障网络空间保持必要的讨论活力。

5 网络暴力治理中平台责任边界的判断标准

网络暴力治理中的平台责任认定,核心并不在于判断平台是否能够完全阻止网络暴力发生,而在于明确平台在何种情况下应当承担治理义务,以及违反该义务后应承担何种程度的责任。网络空间中的信息传播具有高度复杂性,平台虽然掌握技术资源和管理能力,但并不直接参与大多数网络暴力行为。如果简单依据损害结果要求平台承担责任,容易造成责任归属偏移;如果完全否认平台责任,又难以回应平台在网络传播过程中的实际影响。因此,平台责任边界需要建立在平台能力、风险程度和行表现为表现相结合的判断体系之上。

5.1 以平台实际控制能力作为责任基础

平台责任首先来源于其对网络信息传播过程具有一定控制能力。传统网络服务提供者主要承担信息存储和技术支持功能,平台与用户行为之间的关系相对间接。但在当前网络环境中,平台已经通过算法推荐、内容排序、流量分发以及账号管理等方式参与信息传播过程。平台不仅影响信息是否能够被传播,也影响信息能够被多少用户看到。这种控制能力决定了平台不能完全以“技术中立”为理由排除自身责任。网络暴力事件中,某些攻击性内容之所以能够快速扩大影响,与平台的信息传播机制存在联系。例如,具有强烈情绪倾向的内容更容易获得互动和曝光,如果平台推荐机制持续扩大相关内容传播范围,平台实际上已经参与了风险扩散过程。但平台控制能力并不意味着平台需要承担无限责任。平台能够影响信息传播,并不代表其能够预见和控制所有用户行为。互联网用户数量庞大,信息内容不断变化,即使平台采取审核和管理措施,也难以完全避免网络暴力发生。因此,平台责任判断应关注平台是否具有现实控制能力,而不是要求平台达到绝对治理效果。

不同类型平台的控制能力存在差异,其责任程度也应有所区别。大型社交平台和内容平台拥有更强的数据处理能力、用户管理能力和技术资源,能够建立更加完善的风险识别机制,

其承担的治理注意义务应相对较高。对于规模较小、功能较单一的平台,则需要结合其实际运营能力判断责任范围。

5.2 以风险发现可能性判断平台注意义务

平台责任的另一个重要判断因素,是平台是否具有发现网络暴力风险的可能性。平台责任并不是基于网络暴力结果产生后的自动承担,而是建立在平台能够预见风险并采取措施的基础上。

在网络暴力形成过程中,不同阶段的平台注意义务并不完全相同。网络暴力初期,相关内容可能仅由少量用户发布,影响范围有限,平台发现风险的难度较高。随着相关内容大量传播、用户集中举报或者出现明显攻击特征,平台发现风险的可能性逐渐提高,此时平台应承担更积极的处置义务。例如,对于公开发布个人身份信息、组织用户集中攻击某一对象、持续发布侮辱诽谤内容等行为,平台通常能够通过用户举报、关键词识别、异常传播监测等方式发现风险。如果平台已经掌握相关情况,仍未采取删除、限制传播或者账号处理措施,则可能说明平台未履行合理治理义务。

网络暴力治理中的“知道”标准需要避免过于严格或者过于宽泛。如果要求平台必须主动发现所有违法内容,实际上会形成无限审核责任;如果仅承认平台收到明确通知后的责任,又可能导致平台长期忽视明显风险。因此,需要结合信息传播程度、平台技术能力、用户投诉情况以及违法内容明显程度进行综合判断。平台注意义务的核心,在于要求平台对自身能够控制的风险保持合理关注。平台并不需要成为网络空间中的全面监管者,但需要在其能力范围内发挥治理作用。

5.3 以传播机制影响程度确定平台责任范围

网络平台与传统信息传播主体的重要区别,在于平台不仅提供传播渠道,还通过技术机制影响信息流动。算法推荐、热点排序和流量分配等机制,可能改变网络暴力内容的传播范围。因此,在判断平台责任时,需要考虑平台传播机制对损害扩大的影响。如果网络暴力内容仅由个别用户发布,尚未形成明显扩散,平台责任主要集中于内容管理义务。如果平台通

过推荐机制持续提高相关内容曝光度,使攻击行为从个体表达演变为群体性攻击,平台在传播过程中的作用明显增强,其责任判断也应有所不同。当然,算法推荐并不意味着平台对所有推荐结果承担直接责任。算法运行具有复杂性,平台无法完全预测每一条内容可能产生的社会影响。判断平台责任时,需要关注算法机制是否存在明显风险,例如是否长期鼓励极端化内容传播,是否缺乏针对异常传播行为的调整措施。网络暴力治理中的平台责任,应当关注平台对传播过程的影响程度。平台越主动参与信息筛选和传播排序,对风险控制承担的要求也越高;平台仅提供基础技术服务时,其责任范围则应相应受到限制。

5.4 以损害后果严重程度确定责任强度

网络暴力造成的损害具有不同层次,不同损害后果对应的平台责任也应有所区别。部分网络冲突可能仅表现为短时间内的观点争论,而严重网络暴力可能导致个人社会评价降低、心理压力增加、现实生活受到干扰,甚至涉及个人信息泄露和人身安全风险。对于一般性网络争议,平台可以通过提醒、限制互动、降低传播范围等方式进行治理。对于恶意造谣、持续骚扰、公开隐私信息等严重行为,平台则需要采取更强措施。

网络“开盒”行为体现了损害程度对平台责任判断的重要影响。孙道萃指出,网络“开盒”本质上涉及非法公开公民个人信息,部分行为可能伴随网络暴力,并需要根据具体行为性质承担相应法律责任^[5]。这说明,涉及个人信息安全的网络暴力行为,其危害已经超出一般网络争议范围,平台在信息保护和风险控制方面应承担更高要求。损害后果并不是判断平台责任的唯一标准,但能够影响责任强度。当损害结果越严重,平台采取积极治理措施的必要性越高。

5.5 区分民事、行政与刑事责任边界

平台责任边界最终需要落实到不同法律责任类型的区分。网络暴力治理中的平台责任并不意味着所有治理不足都需要通过刑事手段解决。不同责任类型具有不同功能,应根据平台行为性质和违法程度进行适用。

民事责任主要解决平台未履行必要管理义务导致用户权益受损的问题。例如,平台未及时处理侵害人格权益的信息,导致损害进一步扩大,可能需要承担相应赔偿责任。

行政责任主要针对平台违反网络安全管理义务的行为。平台作为网络服务提供者,需要遵守相关监管要求,对违法信息传播承担管理责任。行政监管能够通过处罚、整改等方式推动平台完善治理机制。

刑事责任则需要更加严格的条件。高昱东认为,平台刑事归责需要围绕主观明知、义务违反以及损害后果等因素进行判断,不能仅依据网络暴力发生在平台环境中就直接追究平台刑事责任^[7]。平台只有在明确知悉风险、违反法定义务并造成严重后果的情况下,才具有刑事责任评价基础。

区分不同责任类型,有助于实现平台责任的精准化。责任过轻,难以发挥法律约束作用;责任过重,则可能导致平台采取过度治理措施。合理的平台责任体系,应当根据平台行为、治理能力和损害后果形成层次化责任结构。

6 网络暴力治理中平台责任完善路径

网络暴力治理中的平台责任完善,需要建立在责任边界明确的基础上。平台治理并不是简单增加平台义务,而是在现有法律框架内,根据平台控制能力和网络传播特点,进一步完善治理机制。相比单纯依靠事后处罚,更有效的治理方式应当关注风险预防、过程控制和权利救济,使平台责任真正落实到网络暴力形成和传播的关键环节。

6.1 完善平台风险预防与全过程治理机制

当前网络暴力治理较多集中于违法内容出现后的删除和处罚,但网络暴力具有传播速度快、影响范围扩大的特点,仅依靠事后处理往往难以及时消除影响。平台需要将治理环节前移,通过风险识别、内容审核和传播控制降低网络暴力发生概率。平台可以结合用户行为特征、信息传播路径以及异常互动情况,对可能形成网络暴力的事件进行提前识别。例如,短

时间内大量相似账号集中攻击特定对象、相关话题出现明显情绪化扩散等情况,都可能属于风险信号。平台通过技术工具发现风险后,应结合人工审核进行判断,避免单纯依赖算法造成误判。全过程治理并不意味着平台需要监控所有用户行为,而是在平台能够合理识别风险的范围内履行注意义务。平台责任的重点在于提高治理能力,使其在网络暴力形成初期能够及时介入。

6.2 完善平台责任规则与处置程序

平台责任边界不清,很大程度上源于治理规则缺乏具体化。实践中,平台需要面对大量复杂情况,例如普通网络争议与网络暴力如何区分,轻微违规行为与严重侵害行为如何处理,这些都需要更加明确的规则支持。平台应根据网络暴力行为的性质、传播范围以及损害程度建立分级处理机制。对于一般性争议,可以采取提醒、限制互动等柔性措施;对于恶意造谣、持续骚扰、公开个人信息等严重行为,应采取删除内容、限制账号功能甚至封禁账号等措施。

同时,平台处置程序需要提高透明度。用户有权了解内容被处理的基本原因,也应拥有申诉和复核渠道。透明的治理程序不仅能够保护用户权益,也能够避免平台利用管理权进行过度限制。

6.3 强化个人信息保护与受害者救济机制

网络暴力治理不能仅关注公开攻击性言论,还需要关注个人信息泄露带来的持续性伤害。“开盒”等行为说明,网络暴力已经从语言攻击延伸至人格权益和现实安全领域。平台需要进一步强化个人信息保护责任,防止用户身份信息被非法传播。

平台在发现涉及个人信息泄露的内容后,应及时采取限制传播措施,并根据情况协助相关部门进行调查。同时,应完善针对网络暴力受害者的救济机制,为用户提供更加便捷的投诉渠道和处理反馈。受害者救济机制的完善,有助于改变网络暴力治理中“用户发现问题但难以及时解决”的情况。平台不仅需要管理内容,也需要承担保护用户合法权益的责任。

6.4 推动平台责任与多元治理机制衔接

网络暴力具有复杂的社会成因,仅依靠平台单独治理难以达到理想效果。平台掌握技术和数据优势,但政府监管、司法救济以及社会监督同样不可缺少。在治理体系中,平台应承担与自身能力相匹配的管理责任,政府需要通过法律规范和监管机制明确平台义务,司法机关则需要通过案件处理进一步明确责任标准。多主体协同能够避免平台责任过度集中,也能够提高网络暴力治理的整体效果。网络暴力治理最终需要形成平台主动治理、政府有效监管、用户理性参与的责任体系。平台责任的完善不是简单提高处罚力度,而是通过制度设计使平台治理能力与法律责任相互匹配。

7 平台刑事责任介入的边界限制

网络暴力造成的损害具有较强社会影响,部分严重情形可能涉及刑事责任追究。在治理实践中,强化平台责任已经成为提升网络空间治理效果的重要方向,但平台刑事责任的适用需要保持必要限度。刑法作为最严格的法律制裁方式,应当针对具有较高违法性和社会危害性的行为,而不能成为解决所有平台治理不足问题的工具。

平台承担刑事责任的理论基础,主要来源于其网络安全管理义务。当平台具有管理能力,已经明确知悉违法行为存在,却拒绝履行必要治理义务,并造成严重后果时,其不作为可能具有刑法规制必要性。高显东指出,平台刑事归责需要围绕主观明知、义务违反以及损害后果等因素展开,通过判断平台是否具有防控能力以及是否实际违反法定义务,确定刑事责任范围^[7]。

当前实践中,网络暴力治理涉及的平台刑事责任主要与拒不履行信息网络安全管理义务等相关制度联系。该类责任强调平台对于网络安全风险的管理义务,要求平台在法律规定范围内采取必要措施。对于平台明知存在违法信息传播风险,却未按照要求采取处置措施的情况,可以通过刑事责任进行约束。

但平台刑事责任不能无限扩张。网络暴力形成过程通常涉及多个主体,用户发布内容、其他用户参与传播以及平台传播机制共同影响

最终结果。如果仅因为损害发生在平台环境中,就要求平台承担刑事责任,会造成责任归属偏移,也可能弱化直接实施者的责任。

平台刑事责任认定需要坚持责任与行为相对应原则。平台需要对自身违反管理义务的行为负责,而不是对所有用户违法行为承担替代责任。例如,平台未及时处理明显违法内容,与平台已经采取合理审核措施但仍发生隐蔽性攻击行为,二者在法律评价上应当有所区别。

刑事责任介入还需要考虑网络治理与表达自由之间的关系。网络平台承担治理责任,并不意味着平台具有无限制的信息控制权。过度依赖刑罚推动平台治理,可能导致平台为了避免风险而采取过度审查措施,使正常网络表达受到影响。

原新利、陈燊龙在分析侮辱诽谤型网络暴力犯罪时指出,网络暴力具有区别于传统侵权行为的传播特点,刑法规制需要关注行为性质、危害后果以及责任主体之间的差异^[8]。这一思路同样适用于平台责任判断。平台刑事责任的设置,应当关注平台是否违反明确义务,而不能仅以网络暴力造成严重影响作为追责依据。

从治理效果来看,平台刑事责任应当发挥补充和保障作用。民事责任和行政监管能够处理大部分平台治理不足问题,刑事责任主要针对严重违法法定义务、造成重大危害后果的情形。通过明确不同责任层级,可以避免刑事治理过度介入,也能够促使平台更加积极履行网络治理义务。

8 网络暴力治理中平台责任体系的优化方向

网络暴力治理中的平台责任,本质上是网络空间治理结构变化带来的责任重新分配问题。平台掌握信息传播的重要资源,也承担维护网络生态秩序的现实义务。未来平台责任体系的完善,需要在强化治理能力与保障用户权益之间保持平衡。

参考文献:

[1] 马永强. 论网络暴力的生成机制与刑法规制[J]. 武汉大学学报(哲学社会科学版), 2026, 79(5): 164-

网络暴力治理应当从单纯依赖事后处理转向全过程风险控制。传统治理模式往往关注违法内容出现后的删除和处罚,但网络暴力具有快速扩散特点,单纯事后处置可能难以及时恢复受害者权益。平台需要将风险识别、传播控制和用户救济纳入日常治理体系,通过更加主动的方式降低网络暴力发生概率。

平台责任体系还需要进一步明确责任边界。平台承担治理义务的基础在于其技术控制能力和信息管理能力,而不是网络空间中的所有风险都由平台承担。合理的平台责任制度,应当区分用户行为责任与平台管理责任,使直接实施网络暴力的用户承担相应法律后果,同时要求平台对自身能够控制的风险履行必要义务。

实名制和用户身份治理也是完善网络暴力治理的重要方向。余人、裴晗、刘曦柯指出,网络实名制在强化责任追踪、改善网络治理方面具有积极作用,但同时也伴随着隐私保护和表达空间变化等问题,需要在安全与权利之间寻找平衡^[9]。这一问题说明,网络治理不能简单追求风险控制,而需要兼顾用户合法权益。

未来平台治理还需要进一步加强多主体协同。网络暴力涉及技术、法律、社会心理等多个层面,仅依靠平台自身难以完全解决。政府需要通过法律规范明确平台义务,司法机关需要通过案件处理形成责任认定标准,平台需要利用技术和管理优势落实治理措施,用户也需要提高网络行为责任意识。

网络暴力治理中的平台责任边界,不应理解为平台责任范围的简单扩大,而应理解为平台权力与责任之间的重新平衡。平台拥有越强的信息组织能力,就越需要承担相应治理义务;平台责任越明确,用户权利保护和网络表达自由越需要得到制度保障。通过合理划分责任范围,可以推动平台从被动合规走向主动治理,使网络空间治理更加稳定和有序。

176.

[2] 郭芙蓉, 计玲芳. 数智时代网络话语暴力发展新态势与矫治研究[J]. 山东工商学院学报, 2026, 40(4): 125-132.

[3] 赵凌. 网络暴力治理中的平台责任完善[D]. 导师: 孙庆春. 山东师范大学, 2026.

[4] 李航宇, 刘素霞. “法不责众”困境下网络暴力刑事规制进路[J]. 哈尔滨学院学报, 2026, 47(8): 62-66.

[5] 孙道萃. 网络“开盒”行为的罪责要义和定性止争[J]. 数字法治, 2026, (4): 128-138.

[6] 李佳音. 网络暴力协同治理中平台义务研究[D]. 导师: 李继扬, 申青山. 郑州航空工业管理学院, 2026.

[7] 高显东. 网络暴力治理中平台刑事归责研究[D]. 导师: 孙杰. 山东政法学院, 2026.

[8] 原新利, 陈燊龙. 侮辱诽谤型网络暴力犯罪的刑法规制路径[J]. 山西省政法管理干部学院学报, 2026, 39(3): 63-67.

[9] 余人, 裴晗, 刘曦柯. 从后台隐匿到前台可见: 中国网络实名制的制度演进与综合治理[J]. 决策与信息, 2026, (8): 48-59.

版权声明

© 2026 作者版权所有。本文依据“知识共用署名 4.0 国际授权合约”(CC BY 4.0) 以开放获取方式发布。该许可允许使用者在任何媒介中自由使用、复制、传播与改编文章(含商业用途), 惟须明确署名原作者及出处, 并注明所作修改(如有)。完整协议详见: <https://creativecommons.org/licenses/by/4.0/deed.zh-hans>

出版声明

所有出版物中的陈述、观点及资料仅代表作者及供稿者个人立场, 与 Brilliance Publishing Limited 及/或编辑人员无关。Brilliance Publishing Limited 及/或编辑人员对因内容所提及的任何理念、方法、说明或产品所导致的人身或财产损害概不负责。