

网络暴力下互联网社交平台安全保障义务的构建研究

李 傲

青岛科技大学法学院, 山东 青岛

收稿日期: 2026年3月15日; 录用日期: 2026年4月8日; 发布日期: 2026年4月17日

摘 要

数字化的大背景下, 网络暴力事件的频发对公民的人格权保护以及网络空间秩序的维护等构成了严峻挑战。由于网络暴力群体性、匿名性以及其本身的法益复合性等特征, 加之网络空间的主要管理者本身义务认定模糊、履行标准不一以及技术能力与责任不匹配等问题, 导致现行的法律规制不足, 治理机制也存在一定的缺陷。因此需要明确互联网社交平台在网络暴力下安全保障义务的履行, 通过不同阶段具体义务的设置和法律制度的健全, 实现平台主体责任的强化、遏制网络暴力以及保障公民人格权益, 构建更加健康有序的网络环境。

关键词

网络暴力, 社交平台, 互联网侵权, 安全保障义务

A Study on the Construction of Internet Social Platforms' Duty of Safety Guarantee against Cyberbullying

Ao Li

Law School of Qingdao University of Science and Technology, Qingdao Shandong

Received: March 15, 2026; accepted: April 8, 2026; published: April 17, 2026

Abstract

Against the backdrop of digitalization, the frequent occurrence of cyberbullying incidents poses severe challenges to the protection of citizens' personality rights and the maintenance of cyberspace order. Owing to the characteristics of cyberbullying such as its group-based nature, anonymity and

compound legal interests involved, coupled with issues including the ambiguous determination of obligations of the primary administrators of cyberspace, inconsistent performance standards, and the mismatch between technical capabilities and liabilities, the current legal regulation is insufficient and the governance mechanism has certain defects. Therefore, it is necessary to clarify the performance of the duty of safety guarantee by internet social platforms in the context of cyberbullying. Through the formulation of specific obligations at different stages and the improvement of legal systems, we can strengthen the primary liabilities of platforms, curb cyberbullying, safeguard citizens' personality rights and interests, and construct a healthier and more orderly internet environment.

Keywords

Cyberbullying, Social Platforms, Internet Infringement, Duty of Safety Guarantee

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

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

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



Open Access

1. 引言

互联网社交平台的全民普及使网络已成为公众生活的特殊公共场所，网络暴力也从偶发的个体行为演变为群体性和结构性的治理难题，AI 伪造以及跨平台联动等新型样态层出不穷。互联网社交平台在网络暴力的滋生中有着双重角色，它既是网络暴力过程中信息传播的主体，又是网络暴力风险防控的主体，实质性参与到网络暴力从生成到防治的全过程。

网络暴力作为一种互联网侵权形式需受到法律的严格规制。尽管我国先后颁布了《关于切实加强网络暴力治理的通知》以及《网络暴力信息治理规定》等规范性文件，但网络暴力事件仍层出不穷。《民法典》中对互联网侵权的规制集中于以第 1194 条到第 1197 条为核心的“通知 - 删除”规则与“知道规则”，这些规则主要坚持被动响应的治理逻辑，难以应对网络暴力的群体性与持续性，导致受害者陷入维权困境。此外，不同的社交平台技术能力、用户人数等差异显著，现有规制通常采取一刀切模式出现责任义务与平台规模不匹配等问题。

面对上述问题，需要为互联网社交平台构建安全保障义务以规制网络暴力问题。基于此，本文将互联网社交平台安全保障义务为研究对象，分析现有的网络暴力的特殊性以及立法和实践中对解决该问题存在的不足之处，为社交平台引入安全保障义务，论证该方法的理论可行性和具体内容，并配合配套措施，从而守护公民合法权益，推动网络暴力治理的规范化。

2. 网络暴力与社交平台安全保障义务引入的基础理论

网络暴力和安全保障义务这两个名词并不复杂，网络暴力可能引起侵权纠纷，安全保障义务则是《民法典》中的一类特殊侵权责任。但要通过互联网社交平台安全保障义务治理网络暴力，需要厘清网络暴力的法律定位以及社交平台在网络暴力中的角色定位。

2.1. 网络暴力的法律界定与侵权特征分析

网络暴力这一概念并非严格意义上的法律规范概念，其事实上是对网络空间中各类侵害他人合法权益的线上行为的统称。在 2022 年中央网信办发布的《关于切实加强网络暴力治理的通知》中指出网络暴力是针对个人集中发布的侮辱谩骂、造谣诽谤、侵犯隐私等违法信息及其他不友善信息而侵害他人合法

权益和扰乱正常网络秩序的行为。

与传统的人格权侵权相比,网络暴力主要依托互联网平台的传播功能,形成了具有特殊性的侵权征。

一是行为主体的匿名性与群体性。由于网络空间本身具有虚拟性,施暴者多以匿名或者化名的形式参与,再加上社交平台的用户庞大且互动性强,使得网络暴力信息极易引发群体跟风,造成群体性攻击的同时滋生法不责众的侥幸心理,导致个体责任意识弱化。这种匿名性与群体性相较于传统侵权而言不仅增加了侵权主体的认定难度,而且也增加了受害者的维权难度。

二是损害后果的扩散性和不可逆性。网络技术打破了传统侵权的空间限制,暴力信息可以通过单一社交平台甚至多平台联动扩散。此外,平台算法本身以流量为导向,通常会将高互动性、高争议性的暴力信息标记为热门内容并通过首页推送和话题上榜等方式持续扩散;再加上如今网络水军等灰色产业链,甚至可能形成裂变式传播。且网络暴力虽然是即时的名誉贬损和隐私泄露等,但会固化为负面的社会评价渗透至受害者的职业和生活;部分受害者因此受到严重的心理创伤,有时甚至具有致命性。这种扩散性和不可逆性已超出传统侵权的损害填补范畴。

三是权利救济的复杂性。网络暴力的匿名性和群体性增加了侵权主体认定的难度,与此同时网络暴力信息的传播特性使得受害者在举证上也存在困难。网络暴力维权程序往往复杂繁琐消耗时间长,甚至在此期间受害者还要承受网络暴力的后续影响,即时性危害与维权的滞后性形成鲜明对比,程序的繁琐性使得权利救济沦为事后补救,难以实现及时止损的核心目标^[1]。

2.2. 互联网社交平台的角色演变

互联网社交平台的角色定位随着技术发展和用户规模扩大而逐渐转变,并影响着其责任的内容与边界。

在社交平台发展早期,其核心功能是提供信息存储和传输的场所,例如早期的新浪博客仅作为内容发布的工具,对用户的博文内容并不进行主动地审查。这一时期的平台仅作为信息传输交流的渠道,法律也并不要求其承担主动审查义务。此背景下针对平台内用户被侵权的情况遵循谁侵权谁担责的原则,对平台则采取避风港原则¹进行规制,减轻了互联网平台发展前期的法律风险和经济负担^[2]。这时的平台角色可界定为一种技术中立的服务提供者。当前,随着用户规模的爆炸式增长,社交平台已成为公众交往和舆论形成的核心场所,其功能与物理性的公共场所已无本质差异。随着算法技术的升级,平台对于网络信息内容的传播范围有了实质影响,平台自身权力不断扩张,开始有了空间治理权,即平台通过用户协议、社区规则和算法控制等方式主导着平台内的信息秩序。这表明如今社交平台已向特殊公共空间的管理者这一角色转变。

随着平台角色的转变,其早已不是网络暴力被动的旁观者,其行为直接影响着网络暴力的生成与扩散。首先,互联网社交平台实际上是网络暴力的工具提供者,其所提供的信息发布和互动评论等功能是网络暴力侵权的必要工具,因此若平台未对自身设置必要的风险防控极有可能成为网络暴力侵权的“帮凶”;其次,社交平台作为网络暴力信息的信息管理者,其内容审核以及算法推荐决定着网络暴力的扩散范围,其限流和删除措施的实施影响着损害后果的扩大程度;最后,社交平台在网络暴力侵权下也可能成为潜在获益者,网络平台的运行本身就是一种商业盈利,而网络暴力内容通常具有高流量可以为平台带来广告收益和用户留存等商业利益,平台在客观上可能会在网络暴力侵权中间接获益。

¹参见《民法典》第1195条:网络用户利用网络服务实施侵权行为的,权利人有权通知网络服务提供者采取删除、屏蔽、断开链接等必要措施。通知应当包括构成侵权的初步证据及权利人的真实身份信息。网络服务提供者接到通知后,应当及时将该通知转送相关网络用户,并根据构成侵权的初步证据和服务类型采取必要措施;未及时采取必要措施的,对损害的扩大部分与该网络用户承担连带责任。权利人因错误通知造成网络用户或者网络服务提供者损害的,应当承担侵权责任。法律另有规定的,依照其规定。

3. 我国互联网社交平台网络暴力防治的现状与困境

网络暴力防治是我国数字治理的核心议题之一，虽然当前立法和司法中初步构建了治理框架，但仍存在被动应对和效果有限等问题。因此本章将从立法和司法两个方面展开，分析其网络暴力防治的现状及存在的问题，为后续平台安全保障义务的构建提供问题指引。

3.1. 立法现状梳理与评析

我国网络暴力防治的法律规范涵盖了民事、行政和刑事多个领域，但针对互联网平台责任的相关立法主要集中于民事法律和行政法规。

首先，《民法典》中直接规制平台的内容在第 1194 到 1197 条，主要规定了平台的“通知 - 删除”规则与“知道或应知”规则。平台“通知 - 删除”规则要求网络服务提供者在接到被侵权人要求采取删除、屏蔽等必要措施的通知时及时采取措施，否则应承担相应的责任，在此规定下平台承担的是一种事后责任。然而这一规则在网络暴力泛滥的今天其局限性逐渐凸显，即若平台仍恪守消极的事后责任对用户的保护就不够周全^[3]，此外，被网暴者向平台发出通知后若网暴未构成重大侵权或产生较大社会影响，平台往往敷衍了事拒绝履责^[4]；平台本身也逐渐滥用“通知 - 删除”规则，将其本应承担的义务变成逃避责任的常用方法。上述种种因素导致该规则呈现出显著的滞后性与局限性。为适配网络技术的发展，《民法典》中还规定了“知道或应知”规则，该规则将平台对于网络暴力侵权信息的知晓情况作为其责任认定的重要因素，对于知道或应当知道网络暴力侵权行为而未采取必要措施的平台需要与侵权人承担连带责任。这一规定一定程度上给网络平台增加了主动应对的义务，但存在一定的司法适用问题。此外，《民法典》在第 1032 到 1038 条中关于隐私权和个人信息保护的相关规定也为“人肉搜索”类的网络暴力提供了直接规制依据。

其次，在行政层面的立法以强化平台治理责任为导向，形成了专门的规范。2024 年出台的《网络暴力信息治理规定》是我国首部针对网络暴力的专门规章，在其中明确了平台的监测预警义务、快速处置义务和受害者保护义务。该《规定》要求平台建立健全网络暴力信息特征库和典型案例样本库，加强对网络暴力信息的识别监测²；对于明显的网络暴力信息采取快速的删除、屏蔽限流等措施，并要求平台建立一系列保护机制从而保障用户权益。

根据梳理可以看出，现有的立法对平台在网络暴力下所承担的责任规定存在一定的问题。首先，民事法律中对于平台责任认定的条款相较于个人侵权人而言数量较少。然而根据前文论述可知从个体网民着手进行网络暴力侵权的法律规制路径存在重重障碍，对作为一种社会现象存在的网络暴力的遏制效果相对有限^[5]。其次，现有的对平台进行规制的条文规则散见于各类法律之中，未能形成统一的逻辑体系，且各条文中对于平台义务的规定存在原则化的特点，不够具体。最后，各个规则之间也存在着适配和衔接问题。比如《网络暴力信息治理规定》中已经要求平台承担主动的监测义务而《民法典》中的民事责任仍以“通知 - 删除”规则的被动响应为核心，导致义务和责任脱节。

3.2. 司法实践考察与检视

我国涉及网络暴力的司法案件数量正逐年上升，司法实践围绕网络暴力形成以民事追责为主，刑事罪责为辅的格局，但主要围绕网络暴力的源头侵权人展开，对平台的责任追究数量相对较少。

司法实践中对平台的责任认定主要围绕上文提及的“通知 - 删除”规则与“知道或应知”规则展开。

²参见《网络暴力信息治理规定》第 12 条：网络信息服务提供者应当在国家网信部门和国务院有关部门指导下细化网络暴力信息分类标准规则，建立健全网络暴力信息特征库和典型案例样本库，采用人工智能、大数据等技术手段和人工审核相结合的方式加强对网络暴力信息的识别监测。

“通知-删除”规则是法院适用的主要条文；而“知道或应知”规则则适用不常使用，且即使适用该规则，法院在判断是否构成知道或应当知道时的认定标准也相对保守。以北京互联网法院审理的一起高中生被造黄谣的典型案件为例，法院最终基于对未成年人利益最大化的原则审慎适用了“知道或应知”规则以追究平台的责任。但实践中许多网络暴力侵权所造成的损害后果远重于上述案例法院却未能在裁判时适用该规则，导致为数不多的要求平台主动响应的规则形同虚设。

此外，当下网络暴力侵权案件的司法实践中责任方多为个体用户，然而网络暴力的匿名性等特点导致受害者在司法裁判中面临举证障碍。一是侵权主体难以锁定，施暴者往往通过匿名账号和境外服务器等手段发布信息，通常情况下受害者需要向平台索取后台数据才能锁定真实身份，而平台常以保护用户隐私为由拖延甚至拒绝提供。二是因果关系证明困难，由于网络暴力的群体性特征导致其损害后果是多主体行为共同作用的结果，受害者难以证明其中某一施暴者的行为与自身精神等损害的直接因果关系，法院本身也难以确定单个主体的责任份额。三是证据认证难，网络暴力信息极易被删除或修改，且海量的网络暴力信息保存成本远超普通受害者的承受能力。

综上，可看出我国网络暴力司法实践中对平台责任的适用不足，而对个体的责任追究又困难重重，使网络暴力受害者难以有效维护自身权益。

4. 安全保障义务的法理基础和引入正当性分析

安全保障义务是为公共场所的经营者和管理者设立的特殊责任，而随着数字技术的发展网络空间已成为公共活动的重要场所。在网络暴力侵权事件中，绝大多数视角都集中于最初的侵权人和受害者之间，往往忽略了互联网社交平台发挥的作用和承担的角色。因此下文将从传统安全保障义务的内涵切入，并基于此论证社交平台安全保障义务的引入。

4.1. 传统安全保障义务的内涵

安全保障义务的法理基础可追溯至德国法的“交往安全义务”，其核心逻辑是开启或维持危险者需要承担防止危险现实化的义务，我国吸收借鉴了相关规定形成了现在的安全保障义务。根据规定可以看出，安全保障义务的内涵本质上是义务主体承担的对特定危险、风险防范或免除的义务，这种危险或风险由安全保障义务人开启、维持或主导，因此需要其积极采取合理措施保障处于危险、风险可能性中的相关主体人身、财产等合法权益[6]。

传统的安全保障义务是民事侵权责任体系中针对物理空间的场所管理者的特殊责任，规定在我国《民法典》侵权责任编第1198条，主要包括义务主体、义务内容以及责任形态三方面。

条文首先规定了义务主体，即对物理场所有实际控制能力的经营者和管理者如商场的运营方，根据规定可知安全保障义务主体限定的核心依据是风险控制能力，即只有对场所具有支配权的主体，才可能通过设施维护和人员管理等方式管控风险[7]。其次对安全保障义务的内容进行了规定，可以分为两大类，一类是对场所内的设施 and 环境的维护与管理，确保其符合安全标准；另一类是针对场所内第三人侵权行为的合理干预和救助。最后针对不同的侵权场景规定了不同的违反安全保障义务的责任形态，对于因自身过错导致的损害要求义务主体承担直接责任，对于因第三人过错导致的损害未尽到合理义务的承担相应的补充责任。

4.2. 平台安全保障义务的引入

相对于传统的物理空间，网络空间与其在控制方式、风险来源以及损害形态等方面存在一定的异同。首先在控制方式上，二者均以保障参与者人身、财产及人格权益为目标，遵循“谁管控空间、谁承担义

务”的核心逻辑，但物理空间具备闭环性、地域性，管理者通过物理安保、门禁等实体手段实现直接管控；网络空间呈开放性、无界性，平台仅能依托算法、技术代码实现间接管控。其次在风险来源上，两者均是由于相关主体的不当行为与管理者防控缺失导致的，但物理空间风险传播范围小、易感知，而网络空间风险以虚拟性、扩散性为特征，可跨地域极速裂变传播，风险来源更多元，还易叠加次生危害。在损害形态上，物理空间损害主要为有形、即时的人身财产损伤，可精准量化、波及范围固定；网络空间损害多为无形的名誉贬损、精神损害，难以量化，且损害具有持续性，波及范围更广，举证难度也更大。据此可以看出，安全保障义务引入社交平台责任具有正当性和可行性，但需要结合平台自身特点进行规制。

首先对于安全保障义务本身而言，引入平台这一主体具有法理上的正当性。安全保障义务的内核是义务人在领域范围内主动采取危险防控措施规避由其引发的风险，法理基础是危险开启与危险控制理论[8]。该理论认为，开启或持续危险源的主体应对危险源具备最强的控制能力，应当承担防范危险、消除损害的义务。网络社交平台构建了网络交往空间，为用户信息传播、互动交流提供技术载体与服务支持，本质上是网络暴力危险源的开启者与控制者。平台掌握信息审核、传播管控、账号管理等核心技术，能够提前防范网络暴力信息发布、阻断暴力信息传播、处置施暴账号，是唯一具备全方位防控网络暴力风险的主体；相比普通用户，平台在技术、资金、权限上具有绝对优势，最有能力预见、识别、防控网络暴力风险，因此必须承担与之控制能力相匹配的安全保障义务，这与物理场所经营者基于场地控制承担安保义务的逻辑相一致。且根据风险和收益一致性原则，从危险活动中获得利益的主体应当承担相应的风险管控责任。社交平台的商业盈利与平台内用户的活跃和内容的流量直接相关，即使平台并非积极主动地利用网络暴力，但其客观上已从暴力信息的传播中间接获利，需要为其获利承担一定的责任。

其次，对于互联网社交平台而言，承担安全保障义务规制网络暴力侵权具有现实必要性。一是用户进入社交平台开展活动时，基于平台提供防范恶意攻击和防止隐私泄露的服务对平台存在一定的信赖，这种信赖利益的存在便要求平台应当承担相应的安全保障义务以实现用户的合理期待。二是网络暴力侵权的特点导致受害者权利救济面临多重困难，传统的互联网侵权责任的被动性和高门槛使其在规制网络暴力时存在局限性，与此同时平台的算法技术和资源控制却日益加强，这种不对等的局面需要通过平台的责任限制加以改善。最后，网络空间已成为社会公共生活的核心场所，平台承担安全保障义务已成为维护网络生态和保障公共利益的客观要求。

5. 网络暴力下互联网社交平台安全保障义务的具体构建

网络暴力的技术特征与治理困境，要求互联网社交平台安全保障义务的构建要突破传统“通知-删除”等规则的被动逻辑而转向主动防治。因此下文将以社交平台安全保障义务构建的基本原则为基础，构建平台安全保障义务的核心内容并明确义务违反后的责任承担方式，从而实现网络暴力治理需要。

5.1. 社交平台安全保障义务构建的原则指引

在网络暴力侵权的治理中，社交平台安全保障义务的构建并非简单的责任堆积，而需要依托合理的原则进行指导从而使义务设定具备法理正当性的同时又能满足数字空间的复杂治理需要。

首先为保障安全保障义务与平台规模和能力等要素向适配，需要将比例原则作为核心准则，即要求平台所承担的义务必须与其自身的规模、技术能力以及网络暴力信息的风险向对应。对不同规模的平台设置不同的义务，同时将义务的履行标准以平台实际技术能力为限，使之合乎比例避免一刀切的责任分配。其次为发挥安全保障义务维护权益的功能，在引入安全保障义务时要坚持权利平衡原则，即要注意在防控网络暴力与保障用户言论自由和个人信息权益之间寻求平衡，防止过度管控损害合法言论表达自

由，保障用户的基本权利。最后为保障平台在安全保障义务下的治理公信力，要坚持透明原则。平台掌握着规则设置和申诉通道设立等重要的网络暴力治理方式，需要使平台的治理行为能够被用户知晓，从而保障用户知情权和救济权的同时促使平台规范治理行为。

5.2. 社交平台安全保障义务的核心内容

在上述原则的指导下，围绕网络暴力侵权的全过程并结合现有的法律规定，对社交平台安全保障义务进行细致展开，并针对不同规模平台设定差异化的履行要求。

5.2.1. 事前预防义务

社交平台在网络暴力治理中存在技术优势和信息优势，因此应充分利用其数字技术在信息传播和内容管理方面的技术特质，在网络暴力风险防范过程发挥治理前端的效应[9]，因此需要平台在网络暴力的生成环节承担事前预防义务。

一是要坚持用户账号风险管控，账号是网络暴力信息的实施载体，需要平台通过精细化的管理防范恶意账号。通过严格落实《网络安全法》中的账号实名认证要求，并对境外账号的登录注册进行IP核验与身份备案；对于高风险账号利用技术进行精准识别和限制，同时建立并严格执行账号信用制度，对于低信用的账号进行流量限制和讨论禁止等措施。二是要实现信息内容的主动前置审核，平台因开启或维系社交网络空间而负有包括主动审查在内的安全保障义务，既要排除正在发生的侵权行为，也要审查和控制未来可能发生的侵权行为。平台需要突破对网络暴力信息审核的被动模式，通过对用户发布的内容进行前置审核，并建立分级审核机制，对明显暴力内容实行发布前强制拦截，而对模糊性内容实行人工复核，并结合平台历史信息、司法典型案例与网信部门规范指引需建立动态更新的网络暴力特征库。

5.2.2. 事中监控与处置义务

事中的监控与处置义务是社交平台安全保障义务的核心环节，影响着网络暴力侵权的走向。网络平台在获得具体知情前的侵权风险只是抽象上的，但若其知晓率用户的具体侵权行为和信息，这种危险便由抽象转化为具象，那么该危险便不再是意外，而转为发布容许的风险[10]，需要采取必要的措施，在网络暴力发生过程中通过主动检测和快速干预从而阻断网络暴力信息扩散。

首先，社交平台应当对明显违法的侵权信息进行主动处置。这是对现有的《民法典》中红旗规则的具象化体现和优化要求，即平台应当承担主动发现义务，加强对违法内容的管理。对于尚无定论的社会争议事件、涉及未成年人以及残障人士等敏感群体的相关内容以及与历史网暴事件相关联的内容等具有高风险的话题信息进行主动的审核和风险提示；对于一些直接辱骂、公开他人个人隐私和信息、捏造事实诽谤等具备高度辨识度的暴力信息无须等待受害人反映便采取措施进行删除处置。

其次，社交平台需要强化侵权内容投诉机制，建立高效公正的投诉举报和响应机制。海量且高速运转的信息仅凭平台自主主动识别不具备现实性和可行性，因此需要平台为用户提供必要的投诉举报机制。对不予处理的投诉需要提供书面理由，并明确可行的申诉渠道。此外，对于所有涉及网络暴力的投诉举报数据应当进行留存，以此作为平台是否履行安全保障义务的证据之一，从而降低受害方的举证难度同时倒逼平台及时合理处理投诉。

5.2.3. 事后配合和救济义务

事后配合和救济义务是社交平台安全保障义务的最后一步，针对网络暴力已经发生的情况，通过证据保全和配合救济等方式为司法追责提供有效支撑。

首先，证据保全是网络暴力侵权救济的基础前提，更是应对网络暴力匿名性导致的责任追究困境的关键。平台需要依据自身的技术优势建立网络暴力信息证据留存机制，配合司法机关调查取证从而协助

锁定侵权主体。其次，对网络暴力侵权人的惩戒是平台事后救济义务的必要手段，也是规制网络暴力中法不责众的重要保障。平台应根据网络暴力信息的情节轻重程度实施差异化惩戒，且考虑到网络暴力的跨平台传播特性，各社交平台应进行协同，共享违规账号信息，实现一处违规、多处受限，破解单一平台治理的局限性。此外平台也可通过算法推荐扩大事实澄清内容的传播范围，为受害者的正面信息提供流量支持，联动权威媒体开展正向引导，协助受害者恢复社会关系与人格尊严。

5.2.4. 义务履行的差异化要求

互联网高速发展的情况下各类社交平台层出不穷，不同的社交平台中用户规模、技术能力以及市场地位各不相同，平台安全保障义务的构建需要充分考虑技术可行性和现实操作性。引入风险分级与能力相称原则，对不同层级平台的安全保障义务进行差异化设置。我国市场监督管理总局于 2021 年发布了《互联网平台分类分级指南(征求意见稿)》，里面涵盖了可供参考的平台划分标准。基于此，可参照该指南的内容将平台划分为超大型平台、大型平台以及中小平台三个层级，针对各层级平台的能力设定具体的义务。

超大型平台作为网络空间的头部力量，典型平台如微信和抖音，其技术实力雄厚且资源充分，应当承担最严格的主动治理的安全保障义务。在事前预防中依托自身数据自主研发一定的网络暴力识别技术，将自身依据规定构建动态的网络暴力信息库和典型案例库并向行业开放，发挥领导导向作用；在事中义务中建立全面实时的内容审核检测机制；大型平台具备一定的技术能力但资源有限，因此应当重点聚焦于承担安全保障义务的履行举措，在事前预防中落实严格的实名认证制度，事中通过重点监测和风险响应对网络暴力信息加以处置；中小型平台用户规模有限且相对前两层级平台风险较小，其安全保障义务便应当以最低合规义务为核心，避免过度加重运营负担。

此外，对于超大型平台和大型平台，应当严格遵循“知道或应知”规则，且法院在判断是否违反安全保障义务时对此类平台应当降低“知道或应知”规则的认定标准；对于中小型平台，则应当要求其严格遵循“通知-删除”规则，并对“知道或应知”规则的认定标准进行适当降低。通过上述一系列举措实现网络暴力侵权的精准化平台治理。

5.3. 社交平台安全保障义务的责任承担

在明确社交平台安全保障义务的具体内容后，还需要匹配相应的责任以确保义务的履行。因此需结合立法规定明确社交平台安全保障义务的责任承担。

首先是责任的触发。根据现行的民事法律规范，对于互联网侵权责任的规制是上文提到的“通知-删除”规则和“知道或应知”规则。“通知-删除”规则规定相关的平台经通知后未采取必要措施的需要承担责任，这里的通知实际上是网络服务提供者即本文中的社交平台应尽的安全保障义务的触发因素^[11]，在此基础上如果相关社交平台在具体知情后未及时采取必要措施，那么其在客观上便构成了对安全保障义务的违反，需要承担侵权责任；“知道或应知”规则亦是同理，该条文将应知纳入归责的范围本身就表明立法者不希望社交平台对侵权信息始终处于被动知晓状态而应当承担某种审查义务^[12]，故可以将其作为责任触发的法律依据加以解释。

其次是归责原则，对于互联网社交平台安全保障义务的归责原则的选择需要结合自身特点进行分析。无过错责任原则作为最为严格的归责原则，通常在产品缺陷致人损害等侵权中适用。但对于互联网社交平台而言使用无过错责任原则进行归责过于严苛，这会在无形中为社交平台施加压力，随后转接到对用户个体的严苛监管，最终可能会对公民侵犯言论自由。传统的安全保障义务坚持过错责任原则，以义务主体存在过错为责任的承担前提，由受害方承担证明义务主体未尽到安全保障义务的举证责任。但是对

于网络暴力侵权而言平台和用户之间的信息和技术不对等加剧了受害者的举证困难。综上,为了维护用户的合法权益同时又能督促社交平台积极履行安全保障义务,采用过错推定原则较为适宜,由社交平台自身承担履行安全保障义务的举证责任,解决受害者举证困难的同时又不会过度限制平台。

最后是责任承担方式。结合《民法典》规定的安全保障义务条款和网络侵权条款,并根据平台自身的过错与损害后果进行针对性分析。一方面,对能够确定单一侵权人的网络暴力侵权,平台明知或应知网络用户利用其服务实施网络暴力,却未采取必要措施,或主动推荐、放大暴力内容以获取流量利益,应与施暴者对受害者的全部损失承担连带责任;对于平台已履行基础义务但存在轻微疏漏的情形,则应当承担补充责任,即先由侵权人承担赔偿责任,不足部分由平台在过错范围内补充赔偿,既保障受害者权益,又避免平台承担过重。另一方面,对于群体性网络暴力中难以确定具体侵权人的,应当要求平台对受害者承担全部赔偿责任,后续可向能够确定的实际侵权人追偿,从而倒逼平台严格履行安全保障义务的全方面内容,破解追责难困境。

6. 总结

网络暴力的蔓延与升级,既暴露了数字时代网络空间治理的结构性矛盾,也凸显了互联网社交平台安全保障义务构建的紧迫性与必要性。作为网络公共空间的经营者和管理者,社交平台不仅是信息传播的核心载体,更是网络暴力滋生、扩散的关键场域,其安全保障义务的合理构建与有效履行,是破解网络暴力治理困局、守护公民人格权益、净化网络生态的关键。而网络暴力的行为不法性、法益复合性与责任竞合性特征,决定了平台安全保障义务不能局限于传统“通知-删除”等规则的被动响应模式,而需构建“事前预防-事中处置-事后救济”全链条覆盖的义务体系。网络空间的治理是一项系统工程,平台安全保障义务的有效落地,未来还需依赖立法、司法、监管与行业自律的协同发力,从而促进解决网络暴力侵权这一具体问题的同时推动数字社会的整体治理。

参考文献

- [1] 张新宝. 顺风车网络平台的安全保障义务与侵权责任[J]. 网络与信息法学研究, 2018, 4(2): 3-17.
- [2] 冯延有. 角色定位与制度流变——论互联网平台治理权的形成与实现[J]. 师大-西部法治论坛, 2022, 7(00): 203-216.
- [3] 石佳友. 网络暴力治理中的平台责任[J]. 法律科学(西北政法大学学报), 2023, 41(6): 14-23.
- [4] 周昕怡. 网暴治理下平台责任的民法规制[J]. 南方论刊, 2025(12): 74-76.
- [5] 王华伟. 网络暴力治理: 平台责任与守门人角色[J]. 交大法学, 2024, 53(3): 51-64.
- [6] 苏成慧. 信息处理者安全保障义务的体系阐释[J]. 河北法学, 2026, 44(1): 120-138.
- [7] 杨立新. 侵权责任法[M]. 北京: 法律出版社, 2021.
- [8] 王磊. 安全保障义务的解释论展开[J]. 现代法学, 2024, 46(3): 99-112.
- [9] 赵鹏. 私人审查的界限——论网络交易平台对用户内容的行政责任[J]. 清华法学, 2016, 10(6): 115-132.
- [10] 李依怡. 网络服务提供者共同侵权责任的重构[J]. 数据法学, 2025, 8(1): 215-237.
- [11] 冯楚奇. 网络服务提供者违反“交往义务”的间接侵害著作权责任——基于《民法典》第 1194 条至第 1197 条的法教义学分析[J]. 法律适用, 2020(14): 71-84.
- [12] 汪倪杰. 论网络服务提供者安保义务的边界及其构造——以《民法典》网络侵权规则的解釋论为视角[J]. 法治研究, 2022(1): 102-117.