

# 宽严相济刑事政策下网络暴力犯罪的类型化规制

白丁羽

中央民族大学法学院, 北京, 100074;

**摘要:** 随着数字技术和信息通信发展, 网络暴力犯罪呈现规模化、系统化和技术化特征, 对公民人格权益及网络空间秩序构成严重威胁, 有必要依托宽严相济的刑事政策构建类型化规制路径。基于网络暴力的群体性、系统性与技术性的实质特征, 可将网络暴力行为划分为大众自发型、产业链型及技术驱动型三类, 其在参与主体、面向对象和行为类型上呈现差异化特征。大众自发型犯罪以网民偶发性的情绪化聚集为特征, 应当侧重教育矫治与非刑罚化处理; 产业链型犯罪依托职业化黑灰产运作, 需要强化全链条刑事打击; 技术驱动型犯罪涉及技术滥用, 应当穿透归责机制, 明确开发者、平台等技术主体责任。通过分层分类的刑事处遇策略, 期待平衡法益保护和人权保障, 实现网络暴力治理从粗放管控向精准治理的范式转型。

**关键词:** 网络暴力犯罪; 宽严相济刑事政策; 类型化规制

**DOI:** 10.69979/3041-0673.25.07.072

## 引言

随着数字技术的迅猛发展, 网络空间已经成为个人社交联络的第二空间, 其对个人生活、经济活动乃至社会治理的影响日益增强。然而科技是一把双刃剑, 与之伴随而来的是网络暴力犯罪问题的日渐抬头。网络暴力犯罪作为数字时代的次生问题, 对公民人格权益构成了前所未有的挑战, 并正以多样化、隐蔽化和跨地域化方式不断蔓延扩散。近年来, 从个体遭受侮辱谩骂、人身攻击和恶意诽谤到群体性人肉搜索与造谣传谣, 再到利用“AI换脸”“深度伪造”等技术手段进行谣言攻击和大规模传播, 网络暴力行为的手段不断升级, 覆盖面不断扩散, 危害程度呈现指数级增长趋势。例如在著名的“江歌案”中, 谭某不顾事实真相发表对江歌母亲江秋莲的侮辱性漫画文章, 并借助江歌案的影响力迅速走红全网, 江歌母亲的反对者也伺机加入对文章进行评论灌水 and 大规模转发, 江歌母亲的支持者也同步对谭某通过人肉搜索等多种方式网络暴力, 形成了系统性网络暴力事件。杭州女孩郑灵华因一张上传到网页的粉发照片而遭到网暴, 后续无良媒体未经本人同意发布不实报道进一步激化矛盾, 最终郑灵华选择自杀。传统的刑事治理模式在面对这一新型犯罪时, 往往显得力不从心, 一方面, 网络身份的匿名性和跨地域性导致行为主体难以被锁定, 技术手段的引入更导致犯罪成本大幅降低; 另一方面, 网络暴力的群体性参与特性使得责任界定模糊, 相当一部分人打着法不责众的旗号大行违法犯罪之事。

在此背景下, 以宽严相济的刑事政策嵌入网络暴力犯罪治理成为现实必要, 在面对网络暴力犯罪的多样化与复杂性转型的当下, 应当采用类型化思维, 尝试从网络暴力犯罪的发生逻辑入手, 构建分层分类的刑事处遇机制, 以实现惩治犯罪、修复社会关系和坚持刑法谦抑性的双重目标。本文试图以此为基础, 探讨如何在宽严并济的刑事政策指导下, 针对不同类型的网络暴力犯罪制定对应的规制路径, 期待为构建更加科学高效的网络空间治理体系提供借鉴。

## 1 网络暴力犯罪的概念

### 1.1 网络暴力的概念界定

对概念的理清有助于界定网络暴力的实质特征和行为模式。根据“两高一部”联合印发的《关于依法惩治网络暴力违法犯罪的指导意见》(下称《网络暴力意见》), 网络暴力是在信息网络中针对个人肆意发布谩骂侮辱、造谣诽谤、侵犯隐私等信息的行为, 其内核在于“贬损他人人格, 损害他人名誉”。宏观来看, 网络暴力是数字时代衍生的一种新型侵害形态, 其行为的违法性核心在于利用信息网络对特定个体或群体实施精神压制或人格贬损, 主要表现为具有人身侵害性的信息传播行为, 并对他人名誉权、隐私权等人格权益造成了实质性损害。从规范视角看, 尽管网络暴力并未直接施加物理强制, 但其通过持续性、规模性和恶意性的语言攻击和隐私曝光形成的精神压迫本质上与刑法规范下的“暴力”概念的法益侵害性并无二致, 无论是传统物

理暴力对人身权利产生的有形侵害，还是网络暴力对人格尊严的无形侵害，其实质内核都是剥夺或限制个体自主权，具有侵害人身法益的社会危险性。

以传统暴力为参照，网络暴力依托信息空间和技术手段，形成了区别于传统暴力犯罪的独有特征，其特殊性的表现体现在行为载体的虚拟性和危害后果的易扩散性上。一方面，传统暴力通常以物理空间的直接接触为界，网络暴力则可以依托数字平台的跨地域传播打破时空界限，使侵害行为和程度形成一对多、多对一甚至多对多的倍增效应。另一方面，网络暴力的法益侵害相比于传统暴力还具有社会性特征。其在直接损害个人人格权益的同时，也可能因为规模化的网络暴力事件导致网络生态秩序崩溃，削弱公众对数字空间的安全感。因此相较于传统暴力的个人权益侵害，网络暴力更可能触发社会秩序法益的保护需求。因此，网络暴力的复合法益属性要求刑事政策既要关注个体权利救济，又要兼顾网络空间的治理效果。

## 1.2 网络暴力的实质特征

结合前述网络暴力和传统暴力行为的异质性，从网络暴力治理的差异特征出发，网络暴力的实质特征可以总结为群体性、系统性和技术性。

网络暴力犯罪的群体性特征体现为参与暴力主体的广泛性和行为动因的多元性。在大量网络暴力案件中，事件的肇始可能只是部分网民发出欠考虑的不适言论，或少量网民的恶意引导，但随后大量网民可能出于道德义愤、娱乐心态、从众心理或利益驱动等因素而自发聚集，通过评论转发、二次加工等方式进一步助推了侵害信息的扩散，宏观上看形成了一种无组织意识群体。这种非组织化的临时性群体大多数时候并没有明确的指挥结构，只是由于网络空间的密集效应和大数据形成的茧房效应形成强大的舆论压力和情感光谱冲击，进而导致被害人在虚拟空间承受的非难压力映射到现实生活的精神折磨。群体性特征既是网络暴力犯罪的形式特征，又是使得责任主体难以精准定位的根本症结，部分参与者以法不责众为抗辩事由和合理化心理逃避法律追责和道义非难，进一步加剧了网络空间的治理难题。

系统性特征是指网络暴力中日益成熟的产业化运作模式。实践中，部分网络暴力事件已形成策划、引流、炒作、变现的完整产业链条，并以网络暴力的话题性为流量基础谋取非法利益。此类群体通过有偿删帖、刷量

控评、舆论操控等手段，雇佣集体网络水军编写煽动性文案、恶意利用算法推荐，人为制造话题热点，吸引不知情者参与到舆论骂战中，致使网络暴力从偶然冲突升级为盖然发生的规模化、组织化暴力生态，将网络暴力转化为牟利工具。例如刘学州案中，初始阶段围观群众对其寻母动机产生怀疑，对网传事实不加怀疑，进而发表对刘学州的攻击谩骂，但此时仍有部分网民对刘学州表示支持并建议其报警处理。但后续部分有影响力的用户为了博得流量，追逐热点甚至制造具有诱导性、情绪化的言论话题，导致大量网民聚集并参与到网暴中，使用大量侮辱性词汇对刘学州进行谩骂，同时大批网络水军加入并转发攻击性言论，使刘学州收到了数千条网暴私信，最终导致其在抑郁心理下选择自杀身亡。这一系统性的运作模式使无意识的网络群体演化为具有恶意倾向的暴力机器，使网络暴力的发生概率从偶然上升至盖然。

技术性网络暴力区别于传统暴力的实质特征。随着生成式人工智能、算法推荐等技术的应用化，网络暴力逐渐与技术手段深度融合，使侵害行为呈现出极强的隐蔽性、迷惑性和扩散性。例如，深度合成技术能够批量制造虚假图像、音频乃至视频，使虚假信息以极具迷惑性的方式迅速传播；生成式人工智能可以批量生产水军评论，且短评的文本生成与人类创作几乎没有差异，难以辨别真伪；AI换脸技术则可在短时间内伪造被害人的负面形象，使受害者面临几乎无法澄清的舆论压力。特别地，即便技术以中立性的方式被应用于网络空间，也难以豁免其在网络暴力中产生的助推作用，诸如基于用户偏好的大数据推送等看似无害的算法应用，在客观上仍然会强化单一叙事的传播，形成信息茧房，从而固化偏见、激化对立，引导群体极化，使传播者和被害者均深陷封闭的信息环境，进而进一步加剧偏见与敌意，放大网络暴力的杀伤力。因此，技术赋能不仅降低了网络暴力的实施门槛，还使网暴行为可以突破时空限制，形成远超传统暴力的侵害效果。

## 2 网络暴力犯罪的类型化梳理

从网络暴力的发生全貌来看，网络暴力的差异化特征会体现在主体、对象和行为上。网络暴力犯罪的类型化梳理应当结合前述网络暴力的实质特征，揭示不同形态网络暴力在主体、对象和行为三个层面的发生机理和社会危害面向。在互联网空间匿名性和即时性传送机制

的催化下,网络暴力已经从早期零散的个体谩骂或小规模骂战转变为具有准组织化或强技术性的复合型犯罪现象,根据网络暴力的群体性、系统性和技术性的侧重差异,可以将其分为大众自发型、产业链型和技术驱动型三类。

大众自发型网络暴力指由网民自发的非理性聚集形成的网络暴力群体事件,往往是由一个偶然事件触发蝴蝶效应,进而导致集体情绪共振。此类暴力行为通常始于某一个或部分个体对争议事件的片面解读或情绪宣泄,但宣泄内容通过社交平台的转发评论和推荐机制形成滚雪球。后续的参与主体既包含受误导的普通网民,也混杂着借机宣泄生活压力的非理性群体,有时也夹杂存在职业键盘侠刻意制造对立议题博取眼球。因此此类网络暴力的主体具有偶然性,且较为积极的参与者规模较大,对象通常是以具有一定社会影响力的公众人物、可能存在道德争议的新闻热点人物,但行为上通常较为温和,鲜有开盒、线下骚扰等恶性事件。例如指导性案例杭州快递案中,造谣诽谤信息起初仅是在小区、同城群聊中散布,使被害人遭到了小范围的排斥疏离,但后续相关虚假信息被扩散至数百个微信群聊和其他公号媒体等,进而吸引更多公众被误导转发,通过添油加醋后迅速极化,最终形成了阅读量达 5 亿,参与人次近 6 万人的舆论海啸,甚至连官方发布说明辟谣都难以弥补谣言发酵的后果,最终导致被害人患上重度抑郁。此类群体性侵害的司法认定难点在于,从微观视角上看单个网民的转发评论行为可能仅构成民事侵权,即便语言极化也难以论证到导致被害人自杀或陷入严重心理创伤的因果关系,但海量言论的叠加效应却直接造成了被害人的精神崩溃甚至自杀,因此如何处理个体行为和整体损害的处罚规制力度和范围成为难题。

产业链型网络暴力是利用无意识网络群体的从众性和平台推流和舆论传导的技术手段,通过程序化手段引爆舆论奇点的网暴方式,整体上展现出了犯罪产业化特征,并形成了从策划、引流、到造势、变现的完整利益链,与职业网络水军团队纠合系统性实施造谣引流、舆情操控等行为。主体上具有固定性和组织性,对象通常指向有敲诈空间的弱势群体或争议群体,行为的烈度远远高于自发型网暴。典型的运作模式是,首先由专业写手根据客户要求编造虚假信息,技术团队使用僵尸账号进行多平台传播,通过蹭热点话题的方式分流浏览量,数据分析师实时监测舆情热度并调整攻击策略,最终使

用敲诈勒索或流量分成等方式实现非法牟利。例如公众号“E 药学术车”发布《惊! 实名举报某公司贿赂江苏某医院主任+院长!》一文,使读者推测联想某药企,致使该企业名誉贬损,最终向其勒索巨额“撤稿费”和“公关服务合同”。由此可见,由系统性网络暴力带来的破坏力不仅及于个人和人身权益,也可能导致连带性的巨额经济损失。此外,部分灰产已经形成了反复敲诈的犯罪闭环,在制造网暴话题后要求被害人支付费用,负责有偿删帖的科技公司清除负面信息后又反向敲诈被害人,进而形成了造谣—删帖—利用删帖事件二次造势的闭合犯罪生态。

技术驱动型网络暴力则集中体现了网暴犯罪和科技发展的深度融合。特别地,在技术性网暴中,由于技术提供者和实际网暴者可能并不具有完全的意思联络,因此常常可能牵涉技术中立性争议。主体层面,存在技术人员直接或间接的参与,具有技术耦合的特点,对象具有较强的不特定性,兼具公众人物、弱势群体甚至一般网名。例如一些深度合成技术可以批量生成虚假音视频,这一技术既可以用于作品创作,生成具有艺术欣赏价值的流媒体作品,也可以用于违法犯罪行为中。或者通过 AI 换脸技术将公众人物或指定对象植入色情影片或存在道德争议的画面,由于技术嵌入的深度性导致其以假乱真的程度远远超过传统的诽谤手段;此外,平台的算法推荐机制也可能扮演不知情帮凶的角色。相当一部分平台基于强大的大数据分析,实现了基于用户画像进行的精准推送,导致侵权内容更可能被推送至具有相似阅读兴趣的用户,从而使具有争议性和话题性的网暴话题被指数级扩散。此外,更隐蔽的技术滥用存在于数据爬取领域,犯罪团伙利用网络爬虫、撞库、非法数据库等方式获取被害人的通讯录信息,进而实施“开盒”“人肉”等不法骚扰。技术型暴力的治理困境在于开发者常以技术无罪论为由规避责任,即便开发者确为帮助网络暴力开发软件,也可以通过不明知实际用途进行抗辩。

形式上,上述三类网络暴力具有异质性,但在危害性层面存在互相增强的强化效应。大众自发型为产业链的运作提供了必要的社会情绪基础,技术手段的嵌入又助推了暴力行为的破坏实效。例如德阳女医生案中,常某因不满公安机关的处理结果,将被害人个人信息曝光并借助网暴产业链扩大影响,后续由于媒体忽视公安机关的处理细节而直接作出主观判断,进而客观上引导力



网民群体对被害人发起侮辱谩骂，最终形成了规模化的网暴事件，最终导致被害人因无法承受精神压力自杀身亡。由此可见，网暴的三种类型并不孤立存在，而是具有形成情绪基础—定点爆破—技术强化犯罪增强回路的可能性，但从客观归责的角度对事件动因进行归纳，有助于刑事政策的精准介入。

### 3 宽严相济刑事政策与网络暴力犯罪的内在契合性

#### 3.1 宽严相济刑事政策的内涵与价值导向

宽严相济的刑事政策作为我国犯罪治理的核心指导原则，其内涵并非单纯的“宽”和“严”的二元对立或组合，而是要求面向差异化案件和罪名，通过科学的分层处遇实现刑罚适用的精准化、实效化。从刑事政策本质来看，“宽”的维度体现为对轻微犯罪及其背后可修复性法益的谦抑化处理，通常包括非犯罪化、非羁押化与非司法化三种方案。非犯罪化政策强调对于尚未达到严重社会危害程度的行为，应当避免刑事追究的过度介入。对于普通网民因为情绪宣泄而发布的偶发性的侮辱诽谤言论，如果没有导致严重后果并且及时消除影响，则不涉及严重的社会危害，因此可以通过私下处理、民事调解或者行政警告实现纠纷化解；非羁押化政策则侧重刑罚执行的人道主义转向，强调目的刑的适用。对网络暴力犯罪的初犯、偶犯或情节显著轻微危害不大的参与者可以适用社区矫正、赔礼道歉等替代性措施代替刑事处罚；非司法化则体现为刑事程序的前端分流，通过应用诉前和解等机制，降低司法资源消耗和社会关系二次恶化。而“严”的成分则关注对网络暴力犯罪生态的源头打击，针对存在较大恶意的组织化、职业化、技术化的网络暴力行为强化刑罚威慑，坚决打击惩处，形成特殊预防效应。对产业链型犯罪中的组织者、积极参与者和技术驱动型犯罪中具有主观故意的算法开发者，应当适用从严打击的刑事政策，不仅要从严量刑，还可以适用从业禁止、罚金刑等制裁手段和附随后果，斩断犯罪的再生产能力。

宽严相济刑事政策的价值导向在于实现犯罪治理中保护法益和保障人权的双向平衡。一方面，通过适用宽大的刑事政策避免网络暴力过度犯罪化对公民言论自由的挤压，平衡言论自由和社会秩序、人身权益的利益博弈。特别地，对于网络空间中涉及公共讨论和侵权行为的模糊地带，更应当警惕刑事司法的泛化倾向；另

一方面，通过适用严格刑事政策遏制中立技术异化和黑灰产业链对网络空间秩序的侵蚀，以防止局部犯罪盖然升级为系统性风险。例如，在大众自发型的网络暴力中，多数参与者只是受到群体情绪裹挟，主观恶性实际上较低，行为具有瞬时性和偶然性，如果一律适用刑罚可能引发罪责不相适应的原则性争议。而产业链型犯罪中，职业网络水军通过恶意炒作、舆论操控等手段牟取暴利，其行为构成了对网络公共空间的系统性破坏，只有适用刑罚处罚方能维护网络空间秩序和法治底线。

#### 3.2 宽严相济与网暴犯罪的内在契合性

宽严相济的刑事政策和网络暴力犯罪的治理需求具有多方面的契合性，这种契合性源于前述对网络暴力行为类型的类型化解构及其与传统暴力犯罪法益侵害的差异性。从行为性质看，大众自发型、产业链型与技术驱动型网络暴力分别对应着不同的危害层级和治理策略，因此刑事政策的适用需要通过类型化作为基础实现网暴治理的宽严相济。

对于大众自发型网络暴力，网暴参与者大多数仅仅是因为道德评判、信息误解或从众心理等非理性动机介入，其行为均表现出碎片化、去中心化的特征。从保护法益的角度，此类网络暴力侵害的法益通常集中于名誉权、隐私权等个体人格权，且多数情况下可通过删除信息、公开澄清、赔礼道歉等简单方式实现法益恢复。例如，网民在热点事件中跟风转发不实信息并评论，如果可以及时停止传播并主动消除影响，此时适用非犯罪化处理措施既符合比例原则的要求，也有助于引导社会自我修复，维持网络生态的自修复机能。此时，“宽”的政策导向能够避免由于刑事措施的刚性介入导致对正常信息互动的抑制，同时也可以通过教育矫治的方式促使参与者反思行为边界，实现“惩治一人、警示一片”的社会效果。

对于产业链型网络暴力，其危害性已超越个体权益范畴，演变为侵蚀网络生态的病原体。职业水军、恶意营销团队通过策划煽动、批量注册账号、虚假流量置换等手段形成产业化犯罪链条，其行为具有持续性和组织性。以刷量控评造势为例，黑产团队通过伪造用户评价干扰正常市场评论动向，最终不仅损害了商家商誉，更削弱了消费者对网络平台的信任基础。对此类犯罪，“严”的刑事政策应当贯穿侦查、起诉、审判全流程。在侦查阶段强化电子证据固定和资金流向追踪等措施，以摧毁

犯罪组织的经济基础；在量刑时加强对首要分子、积极参与者的惩处力度，且可以探索尝试从业禁止令在互联网领域的适用，防止犯罪主体改头换面后重操旧业。

对于技术驱动型网络暴力，其特殊性在于技术工具的双刃剑特点。深度合成、算法推荐等技术通常具有中立性，但是也可能被恶意利用，导致技术赋能犯罪。例如，不法分子利用AI换脸技术伪造他人不雅视频并借助算法精准推送，相较于言语诽谤和片面信息编造更有可能造成大面积的网络暴力攻击。在此类犯罪中，技术提供者的责任边界是刑事政策的关键和难点，如果开发者明知技术被用于违法活动仍提供支持，或未履行必要的内容审核义务，则应当承担共犯或至少过失犯罪责任。此时，“严”的政策导向还需要延伸至技术伦理层面，通过刑事责任倒逼技术研发者完善风险防控机制，以避免单纯依赖事后惩处制止技术越界。

刑事政策与网络暴力类型的深层契合性还体现在其作为社会事件时产生的社会治理面向上。“宽”与“严”的动态平衡实质上是法律效果与社会效果的统一，对偶发参与者适度从宽从而减缓社会对抗，节约司法资源；对组织化、技术化犯罪严厉惩治，维护网络空间的底线秩序。这一逻辑在《网络暴力意见》中也可以得到支撑，对普通网民以教育引导为主，对职业黑产以打击震慑为重，对技术滥用以源头治理为要。只有通过类型化响应机制，才能实现网络暴力治理从一刀切管理向精准治理的范式转型，最终构建兼具包容性与威慑力的网络法治生态。

## 4 宽严相济导向下的类型化规制路径

### 4.1 主体维度：责任分层与精准追责

网络暴力犯罪的治理需以主体特征为基础构建差异化的责任体系，通过精准识别不同参与者的社会危害性和主观恶性实现宽严差异化处理。对于职业化参与者、组织者和技术提供者等核心犯罪主体，由于其行为往往具有持续性、营利性和不可替代性，需要通过强化刑事责任形成高压震慑。例如，在产业链型网络暴力中，组织者通过招募水军、策划虚假话题操控舆论走向，其行为已突破言论自由边界，构成对网络公共秩序的蓄意破坏。对此类主体应从严适用刑罚，除判处自由刑外，可附加高额罚金、没收违法所得，彻底瓦解犯罪组织的再生能力。技术提供者的责任认定则需重点审查其主观明知程度，若开发者明知或应当知道深度合成工具、自动

化脚本等技术被用于捏造虚假信息，依然提供技术支持，则构成网络暴力的帮助犯，应比照主犯从轻减轻承担刑事责任。

而对于偶发性参与者或虽造成一定后果但主动修复法益的行为人，刑事政策应体现教育为主、惩罚为辅的宽缓面向。比如大众自发型网络暴力中，大量网民因信息不对称或情绪化表达卷入网暴活动，其行为虽然具有一定的违法性，但主观恶性较低且危害后果具有可逆性，因此可以不予追究或要求其承担民事、行政责任。例如，普通用户转发未经核实的谣言，在知晓真相后及时删除信息并公开澄清，则可以仅进行教育、训诫等方式予以规劝。对此类具有偶然性的主体可以构建阶梯式的责任体系，对于初次参与者以训诫、公益服务等非刑罚措施实现行为矫正，后续基于重复参与者则可以适用行政处罚，对屡教不改或客观上造成严重后果的主体启动刑事追责。同时，对于主动消除影响、赔偿损失的行为人，可以适用相对不起诉实现非罪化处理，从而实现既节约司法资源，又促进社会关系修复的客观实效。

### 4.2 对象维度：差异化保护原则

网络暴力犯罪的对象保护需要兼顾个体权益与公共利益，根据受害主体身份特征实施必要的规制策略。对于未成年人、残障人士等弱势群体，鉴于其身心承受能力较弱且缺乏有效维权手段，法律应倾斜配置保护资源。司法实践中可适当降低入罪门槛，例如将针对未成年人的网络侮辱、诽谤行为认定为“情节严重”时，相较于普通案件可以缩减转发量、浏览量等量化标准，同时加大对施暴者的精神损害赔偿支持力度。而对于公众人物、国家机关等特殊对象，则需在权益保护与言论自由之间寻求平衡。公众人物因主动置身于公共领域，其名誉权、隐私权保护范围应适当限缩。例如网民对官员廉洁性、明星职业道德的质疑性言论，只要未捏造事实或使用侮辱性言辞，即使存在部分失实亦不宜轻易入罪。国家机关作为公权力象征，对其网络形象的批评监督属于公民基本权利，司法实践中应严格区分恶意诋毁与正当监督的边界，从而避免将负面舆情简单等同于网络暴力。例如，如果网民基于客观事实对政策失误进行尖锐批评，即便用语稍有激烈也应当受宪法保护，只有以虚构事实、煽动仇恨的言论进行诋毁才需纳入刑事规制范畴。这种差异化保护既体现法律对弱势群体的特殊关怀，也维护了公共讨论空间的活力。

### 4.3 行为维度：类型化行为的分层规制逻辑

网络暴力犯罪的行为样态因其类型差异呈现不同的危害机理，需针对大众自发型、产业链型与技术驱动型的行为特征构建分层规制框架。对于大众自发型网络暴力，其行为多表现为网民个体或松散群体的情绪化表达，此类行为具有瞬时性和非理性，应重点关注及时干预和防止事件扩散。平台可以建立网络暴力的动态分级响应机制，对尚未形成规模传播的侵权言论，通过平台预警系统触发自动屏蔽或内容限流；对于已引发广泛传播的个案，则需要通过大数据溯源方法锁定关键的传播节点，将恶意加工、二次扩散的核心参与者交付公安机关并追究其法律责任，对单纯点击、浏览、评论的边缘化行为则以警示教育为主。

产业链型网络暴力的行为模式具有严密的组织分工与利益驱动特征，因此需要打破盈利链条，阻断非法收入链路为导向铲除犯罪生态。从策划编造虚假信息、批量注册账号引流，到操控舆论走向、实施敲诈勒索，此类行为的每个环节均存在明确的技术工具支持与利益分配机制。对此需采取全链条的打击策略，在刑事立法层面，将职业化网络水军的组织行为单独入罪，降低对下游具体侵权行为的证明责任依赖；在司法实践中，重点追查资金流向与技术服务合同，对提供虚假账号注册、流量刷量技术的第三方服务商以帮助信息网络犯罪活动罪定罪量刑。

技术驱动型网络暴力的行为特殊性在于技术工具对犯罪效能的指数级放大，所以应当从技术治理和责任穿透双重路径构建规制体系。深度合成、算法推荐等技术不仅改变了网络暴力的实施方式，更模糊了技术提供者与直接侵权人的责任边界。对于利用 AI 换脸、语音模拟等技术伪造他人不雅内容的行为，应当建立事前事后规制方案，事前要求技术开发者在产品发布前嵌入内容识别水印、使用日志存证等功能，从源头控制技术滥用风险；对未履行必要风险防控义务的开发者的，可参照自动驾驶基于客观归责理论追究刑事责任。而在算法推荐场景中，若平台利用个性化推送技术主动扩大侵权内容传播，则需突破传统中立帮助责任理论，将算法设计者的主观意图纳入考量。此外，对专门收集和生成僵尸账号生成软件的技术团队，即使其辩称技术中立，也可以依据软件功能的专用性、客户群体的特定性等证据认定主观明知，从而突破技术帮助行为的责任认定困境。

三类行为的规制逻辑最终可以指向宽严相济刑事政策的动态平衡，对大众自发型行为强调宽处理和为主，通过非刑罚手段实现社会关系修复；对产业链型行为坚持严厉打击和源头治理，摧毁犯罪组织的生存基础；对技术驱动型行为注重宽严相济和风险防范，在技术创新与风险防控间建立缓冲带。通过行为类型与规制策略的精准匹配，实现网络空间秩序维护与公民权利保障的有机统一。

## 5 结语

在数字技术重塑人类存在机制和交往逻辑的当下，网络暴力犯罪已从零散的个体冲突演变为威胁公共秩序的系统性风险。本文通过解构网络暴力的大众自发、产业链运作与技术驱动三种类型，揭示了其在主体、对象和行为危害三重维度的复合性和复杂性。本文基于宽严相济刑事政策的理论框架，提出类型化规制路径，对偶发性、低恶性的网民个体行为，通过非犯罪化、非刑罚化措施实现教育矫治；对组织化、牟利性的黑灰产业链条，强化全链条刑事打击与从业禁止等附随后果；对技术滥用引发的系统性风险，通过穿透式归责机制明确技术开发者与平台的主体责任。这种分层治理模式既体现了对轻微违法行为的刑法谦抑，又通过重点打击核心犯罪节点有效维护了网络秩序，实现了法益保护与人权保障的动态平衡。当前，网络暴力治理的核心矛盾已从单一的行为规制转向技术伦理与人性底线的深层博弈。唯有以法治之光照亮网络空间的暗面，才能让数字洪流中的每一粒尘埃，都不再成为压垮生命的最后一根稻草。

## 参考文献

- [1] 郭旨龙. 网络暴力刑法治理的解释原理[J]. 江汉论坛, 2023(5): 119-127.
- [2] 敬力嘉, 胡隽. 网络暴力法律规制的完善路径[J]. 中国人民公安大学学报(社会科学版), 2021, 37(5): 142-149.
- [3] 敬力嘉. 网络暴力公私合作治理模式的反思与调试[J]. 江汉论坛, 2023(5): 136-144.
- [4] 蹇昶, 杨宗科. 舆论失焦中网络暴力的法律规制[J]. 传媒, 2023(9): 94-96.
- [5] 李凯林. 谣言、信息茧房与群体极化：短视频平台网络暴力的治理[J]. 甘肃行政学院学报, 2023(3): 106-113, 123, 128.



- [6]李森. 网络暴力犯罪“法不责众”的规制难点与刑事应对[J]. 当代法学, 2024, 38(5): 112-123.
- [7]刘艳红. 理念、逻辑与路径: 网络暴力法治化治理研究[J]. 江淮论坛, 2022(6): 21-30, 2.
- [8]田宏杰. 网络暴力刑法治理中的“法不责众”困境及其化解[J]. 法学杂志, 2024, 45(1): 29-44, 2.
- [9]涂欣筠, 李政达. 环境犯罪学视阈下网络暴力行为研究[J]. 中国人民公安大学学报(社会科学版), 2023, 39(3): 16-29.
- [10]姚万勤. 客观归责理论与自动驾驶交通事故刑事责任归属的归属[J]. 大连理工大学学报(社会科学版), 2023, 44(6): 93-101.
- [11]张凌寒, 贾斯瑶. 人工智能生成内容标识制度的逻辑更新与制度优化[J]. 求是学刊, 2024, 51(1): 112-122.
- [12]Mingqi Z. Research on the Legal Regulation Dilemma and Governance Path of New Network Violence in the Era of Self-media[C]// SSHA 2024.
- [13]Yanjun L, Wei W. Research on Chinese Legal Regulation of Internet Violence in the Internet Era[J]. International Law Research, 2024, 13(1): 19-19.
- [14]李侯然. 网络暴力犯罪规制研究[D]. 南京财经大学, 2023.
- [15]李姿璇. 网络空间诽谤罪若干问题研究——以“杭州女子取快递遭诽谤”案为例[D]. 山东大学, 2021.
- [16]牛奔, 耿爽, 王红著. 数据科学导论[M]. 北京: 中国经济出版社, 2022.
- [17]杨文俊主编. 网络安全事件防范与应对[M]. 中国环境出版社, 2017.
- [18]被“抛弃”的刘学州的一生[EB/OL]. [2025-03-28]. <https://mp.weixin.qq.com/s/qWuiKXRaZPV358D1CpIN7g>.
- [19]“两高一部”联合印发《关于依法惩治网络暴力违法犯罪的指导意见》-中国法院网[EB/OL]. [2025-03-28]. <https://www.chinacourt.org/article/detail/2023/09/id/7550476.shtml>.
- [20]上海市第二中级人民法院(2020)沪02刑终672号刑事裁定书.
- [21]那个被网暴压垮的粉红色头发女孩, 走了\_新闻频道\_央视网(cctv.com)[EB/OL]. [2025-03-28]. <https://news.cctv.com/2023/02/22/ARTI6A4aGGV0h6oDNnUkT47V230222.shtml>.
- [22]自媒体发黑料, 不给钱不删稿? 用法治力量维护网络清朗\_中华人民共和国最高人民检察院[EB/OL]. [2025-03-29]. [https://www.spp.gov.cn/spp/zdgz/202503/t20250308\\_688311.shtml](https://www.spp.gov.cn/spp/zdgz/202503/t20250308_688311.shtml).
- 作者信息: 白丁羽, 女, 汉族, 2004-, 陕西榆林. 学校: 中央民族大学法学院 研究方向: 法学