

《超越看门人：社交媒体平台网络暴力治理比较研究》

研究过程部分信息

胡岑岑 黄雅兰

本研究分析的社交媒体平台基本情况^①

平台名称	上线年份	月独立设备数（万）
微博	2009.8	61941
哔哩哔哩	2009.6	17255
小红书	2013.6	19990
抖音	2016.9	91266
知乎	2011.1	6955
豆瓣	2005.3	939
快手	2011.3（原 GIF 快手）	47109
今日头条	2012.8	33973
百度贴吧	2003.12	-
Soul	2015.9	1790
陌陌	2011.8	9573
虎扑	2004	389
Lofter	2011.8	596
斗鱼	2014.1	2559
YY	2008（原 YY 语音）	1670

^① 月独立设备数数据来自艾瑞咨询的“移动 app 指数”，<https://index.iresearch.com.cn/new/#/app>，数据时间为 2023 年 9 月。
作者简介：胡岑岑，北京体育大学新闻与传播学院讲师，主要从事新媒体与青年文化研究、视听传播研究；黄雅兰，暨南大学新闻与传播学院副教授，主要从事女性主义媒介研究、数字新闻学研究。

本研究的编码类目

1. 网络暴力的定义。主要考察平台相关政策文件中如何界定“网络暴力”，分为三个变量：（1）是否有明确定义（是/否）；（2）是否结合案例对网络暴力进行具体说明（是/否）；（3）如何表述网络暴力与侵权行为之间的关系？（a. 并列关系，指网络暴力与造谣诽谤、恶意骚扰恐吓等侵权行为并列出现在文件中；b. 包含关系，指将网络暴力作为概括性、总体性条目，下辖造谣诽谤、谩骂侮辱等多种侵权行为；c. 两者都有；d. 无明确表述）。

2. 网络暴力的类型。主要考虑平台相关政策文件具体提及的网络暴力类型及范围，共分为9类：（1）谣言诽谤/恶意揣测；谩骂侮辱/贬低歧视；（3）道德绑架；（4）骚扰拱火；（5）公开隐私/信息泄露；（6）煽动滋事；（7）威胁恐吓；（8）内容宣扬暴力；（9）以上都没有。

3. 网络暴力的治理措施。主要考虑平台相关政策文件具体提及的网络暴力治理规范及措施，共分为16类：（1）内容过滤；（2）风险提示；（3）宣传倡导；（4）一键防护；（5）一键取证/举报；（6）评论转发管理；（7）私信设置与防护；（8）黑名单；（9）关键词屏蔽；（10）违规内容清理；（11）账号封禁；（12）治理公示；（13）向有关部门报告；（14）关怀安慰受害者；（15）用户协管员/社区志愿者；（16）账号价值评估。其中1-3类属于事前监测预防措施，4-9类属于事中干预处置措施；10-14类属于事后救助措施；15-16类属于常态治理措施。

各平台文件提及率最高和最低的网络暴力行为类型及其比例

	最高	提及率	最低	提及率
微博	骚扰拱火	46.15%	道德绑架	7.69%
小红书	公开隐私	50%	道德绑架	0.00%
哔哩哔哩	谣言诽谤/恶意揣测	68.18%	道德绑架	9.09%
抖音	谩骂侮辱/贬低歧视	64.29%	道德绑架	21.43%
知乎	谩骂侮辱/贬低歧视	61.11%	道德绑架	0.00%
	骚扰拱火			
豆瓣	公开隐私	58.82%	道德绑架	0.00%
快手	谣言诽谤/恶意揣测	80%	道德绑架	0.00%
今日头条	谣言诽谤/恶意揣测	57.14%	道德绑架	0.00%
	谩骂侮辱/贬低歧视			
百度贴吧	谩骂侮辱/贬低歧视	75%	道德绑架	0.00%
soul	谩骂侮辱/贬低歧视	69.23%	道德绑架	7.69%
陌陌	谩骂侮辱/贬低歧视	100%	道德绑架	0.00%
	公开隐私			
虎扑	谣言诽谤/恶意揣测	71.43%	道德绑架	14.29%
	谩骂侮辱/贬低歧视		煽动滋事	
Lofter	谩骂侮辱/贬低歧视	71.43%	道德绑架	0.00%
斗鱼	谩骂侮辱/贬低歧视	69.23%	道德绑架	0.00%
YY	谣言诽谤/恶意揣测	68%	道德绑架	4%

各平台文件提及的网络暴力治理措施统计结果

		微 博	小 红 书	哔 哩 哔 哩	抖 音	知 乎	豆 瓣	快 手	今 日 头 条	百 度 贴 吧	So ul	陌 陌	虎 扑	Lo ft er	斗 鱼	YY
事前 监测 预防	内容过滤												×		×	×
	风险提示									×		×	×	×	×	
	宣传倡导												×			
事中 干预 处置	一键防护											×	×	×		×
	一键取证 举报															
	评论转发 管理		×									×	×	×	×	×
	私信设置 与防护							×		×		×	×	×	×	×
	黑名单										×	×		×		
	关键词 屏蔽		×					×		×		×		×		
	内容清理															
事后 惩治 救助	账户封禁															
	治理公示											×			×	
	关怀安慰	×				×	×	×	×	×	×		×	×	×	×
	向有关部 门报告															
常态 治理 机制	用户协 管员							×				×			×	
	账号价值 评估				×		×	×		×				×		

（标注“×”表示相关平台的文件中未提及该治理措施。）

各平台事前预防监测功能测试结果

平台名称	发出端		接收端	
	提示违规	提示发送失败	拦截成功	不友好评论提示
微博			2/5	
小红书			4/5	
哔哩哔哩			2/5	
抖音	1/5		1/5	1/5
知乎			5/5	
豆瓣				
快手	5/5		4/5	
今日头条		2/5	4/5	
贴吧	5/5		2/5	
Soul	1/5		4/5	
陌陌	1/5		4/5	
虎扑	2/5		3/5	
Lofter			4/5	
斗鱼		5/5	5/5	
YY	4/5	1/5	5/5	

（平台测验中，我们尝试在每个平台发送 5 条违规评论，表格中的数字表示被警告/被提示/被拦截的评论数量占违规评论总数的比例。）

各平台“一键防护”功能的便捷度和内容测试结果

平台名称	设置点击次数	保护时限（天）	陌生人不可评论/发弹幕	陌生人不可私信	陌生人不可分享/转发/收藏我的内容	陌生人不可@我/我不会出现在转发信息中	陌生人不可搜索我/加我好友
微博	4	7	√	√		√	√
小红书	4	7	√	√	√	√	
哔哩哔哩	5	7	√	√		√	
抖音	3	7	√	√			√
知乎	4	7	√	√			
豆瓣	6	14	√	√	√		√
快手	6	无	√	√			
今日头条	4	无	√	√	√		
百度贴吧	4	无		√			
Soul	4	无			√		√
陌陌	无						
虎扑	4	无	√	√			
Lofter	无						
斗鱼	5	无	√	√			
YY	无						