

全球多国加强对未成年人使用社交媒体的管控——

# 青少年“社媒禁令”面临哪些挑战

自2025年12月10日澳大利亚率先实施全球首个针对16岁以下未成年人社交媒体禁令以来,法国、西班牙、英国等多国纷纷推动类似禁令。3个月过去,这场“社媒禁令”浪潮正面临不少质疑:一纸禁令真能“精准阻断”未成年人接触社媒吗?香港中文大学信息工程系副教授周亚金和中国社会科学院大学法学院副教授、互联网法治研究中心主任刘晓春,共同探讨“社媒禁令”面临的挑战与可能的优化方案。

▶近日,中国国家互联网信息办公室等八部门联合发布《可能影响未成年人身心健康的网络信息分类办法》,明确可能影响未成年人身心健康的网络信息类型及具体表现形式。专家认为,针对内容进行分级推荐或是在网络空间保护未成年人的更优方案。(商海春作 新华社发)



## 针对年龄的主要验证方案在效果上存在不确定性

禁令之下,TikTok通过面部分析技术估算用户年龄,并结合技术和人工审核的方法,检测可能未提供正确出生日期的青少年账号;Meta则结合行为数据与自拍双重机制验证……综合目前各国禁令和平台来看,年龄验证是主流方案,上传身份证、人脸分析、行为推断技术是判定用户年龄的主要技术路径。

这些技术方案的准确率如何?

澳大利亚实施禁令2个多月以来,许多自称未满16岁的用户依然活跃,在相关话题下发出抗议;也有一些青少年反映,自己仍可继续使用原有账号。澳大利亚广播公司的调查显示,约3/4的受访青少年表示他们打算继续使用社交媒体。

“对一个稍有技术常识的16岁少年来说,要突破禁令防线并不难。”周亚金认为,从技术原理上看,目前针对年龄的主流验证方案在实施效果上存在不确定性:

身份证件验证是目前最常见的方案,即用户上传证件,系统识别年龄。然而,一次性核验无法持续确认实际使用者身份,未成年人完全可以用家长的证件完成注册,后续使用中屏幕前坐的是谁,平台无从得知。

人脸年龄估算是另一种被寄予厚望的技术,平台通过摄像头捕捉用户面部特征,用算法推断年龄区间。但面部扫描技术对用户年龄的估算存在误差,需要大量样本训练且需持续迭代。“这意味着,刚满14岁的孩子可能被误判为‘已满16岁’而获得访问权限,而17岁的青年可能被误拦在外。”周亚金说,不同国家和地区、人种等数据差异也会导致模型效果不一,一个在欧美数据集上训练良好的模型,到了东南亚就可能准确率骤降。

行为推断分析技术则通过社交网络结构、互动模式、使用时段等推测用户年龄。周亚金举例说,如果一个账号的互动对象集中在13—15岁年龄段、使用时间集中在放学后、关注话题偏向校园生活,系统可以推断其可能为未成年人。“但用户行为易被刻意模仿或伪装,准确率有限,难以作为唯一判定依据。一个成年人完全可以伪装成青少年的行为模式,反过来,一个未成年人也可以模仿成年人的网络行为。”

刘晓春认为,各国目前推行社媒禁令的技术验证措施依赖实名认证或生物识别,有可能会直接导致未成年人生物信息被过度收集,进而引发隐私泄露隐患,这也是技术层面尚未解决的核心问题。“禁令的实施前提是精准的未成年人身份识别,如果未成年人可通过虚报年龄、借用账号等方式轻易规避,那么禁令既难以落地,还会削弱制度严肃性,并引发潜在风险。”

## “社媒禁令”忽视未成年人的社交和信息需求

数据显示,在澳大利亚社媒禁令生效前一周,澳大利亚VPN(虚拟专用网络)服务商搜索量飙升,部分VPN服务商的安装量短期内暴涨;一些先前鲜为人知的社交媒体APP下载量陡升……禁令一出,青少年开始各出奇招,有的宣称利用家中长辈照片绕过验证,有的利用VPN绕过地域和年龄验证,有的迅速转向了不在禁令名单内的平台。

过去数十年,网络社交推动了互联网向现实世界无限靠近,未成年人作为数字时代原住民,社交媒体已成为其获取信息、开展社交的重要载体。

“全面禁令无法消除未成年人的社交和信息需求,这是客观存在的心理发展规律。”刘晓春认为,封禁头部平台,未成年人转向即时通信群组、小众APP、加密群组等替代渠道是必然的,“一刀切”模

2月24日,英国信息监管局宣布,因美国社交化新闻和娱乐平台红迪网站非法使用儿童个人信息,对其处以1447万英镑的罚款。英国信息监管局调查发现,红迪网站可能使儿童接触到不恰当和有害的内容,且未能实施任何可靠的年龄核验机制。

社媒禁令之下,各国宣布了针对平台的治理措施。澳大利亚立法规定禁令执行的责任在于运营平台的网络社交媒体公司,未予配合的企业将面临最高约5000万澳元(约合2.32亿元人民币)罚款;西班牙政府称若非法或仇恨内容未从社交媒体平台上删除,将追究平台高管的法律责任……围绕网络空间的未成年人保护讨论由来已久,当法律要求平台采取“合理措施”时,这种“合理”究竟该如何定义?是只要平台部署了就算尽职免责,还是必须有能力对抗青少年技术绕过行为?

“平台可以声称‘已部署身份识别’,而实际上可能算法误差巨大,绕过成本极低,监管却无法追责。”刘晓春认为,如果没有配套详细的实施规则,压实平台责任,“合理措施”可能成为一句空话。

“对平台来说,这不是技术问题,而是商业问题。技术的钥匙在平台手里,但企业需要开门的动力。”周亚金分析,社交媒体的商业模式建立在用户规模之上,未成年人既是当下的重要用户群体,也是未来用户的蓄水池。让平台投入资源开发技术、部署系统、运营维护,只为了“赶走”自己的用户,这从根本上违背了商业理性。因此,如果没有对“合理”技术标准的统一,又缺乏细则监管、

式会推动风险向监管薄弱的“地下”社交空间转移。这类空间缺乏成熟的内容审核和安全防护,不良信息更易渗透,且事后追责、监管介入的难度远大于正规平台,反而让未成年人暴露在更高的安全风险中。

“青少年在禁令下规避监管的技术手段除了VPN,还有改机软件等黑灰产工具。”周亚金说,头部平台至少有系统审核、人工巡查、投诉响应等机制,而小众APP或其他技术绕过手段可能会使青少年进入一个风险更高的领域。

“‘一刀切’的禁令忽视了这部分青少年的发展权与参与权,甚至可能引发叛逆心理,适得其反。”刘晓春分析,一方面,网络是青少年连接外部世界、寻找同类、获取支持的重要途径,禁令会切断未成年人通过社交媒体获取知识、参与数字社会的渠道。另一方面,社交功能如今已嵌入各类互联网平台,视频APP有评论互动,游戏有好友系统,学习类工具也有社区板块。按平台类型“一刀切”禁止,不界定封禁范围,并无实际保护意义。

## “一刀切”封禁不符合商业逻辑,内容分级或是更优方案

执法追责,对平台而言执行禁令就是一笔“投入产出比为负”的生意。“需要为企业明确‘做到什么程度就算尽责’的相关细则,只有把‘社会责任’转化为可量化、可审计的合规成本,平台才能将其纳入商业决策模型。”

那么,监管应该在哪些方面制造动力?

专家一致认为,相比“一刀切”的年龄限制,监管层面更应该构建针对未成年人的“内容分级制度”。

周亚金指出,内容识别技术已经相对成熟,文本、图像、音频、视频的自动分类与审核在多数平台已有广泛应用基础。针对未成年人从“管控身份”转向“管控内容”更符合平台商业逻辑,是从“不可能”转向“可能”的关键一步。

“真正的保护不是把青少年关在门外,而是在门内划出安全区。”刘晓春认为,以青少年保护为核心的治理方案,核心是构建“分龄、分功能、分场景”的综合治理体系,而非单一维度施策。一是按不同年龄阶段实施差异化管控,匹配未成年人认知与使用需求,具体区分标准可以参考学龄阶段;二是对高风险社交场景严格限制,对低风险场景合理开放;三是探索灵活的身份验证方式,精准识别未成年人使用行为;四是兼顾保护与发展,保留未成年人合理的网络参与空间,发挥社交媒体的正向价值。

(据《人民日报海外版》)