

香港网络信息治理 之模式转型:从“二元规制”到“三元规制”*

沈新渊¹ 刘文轩^{2,3}

(1.山东政法学院法学院, 山东 济南 250014; 2.烟台大学法学院, 山东 烟台 264000;
3.南京大学犯罪预防与控制研究所, 江苏 南京 210093)

[摘要]现行香港网络信息治理“政府一个人”二元规制模式囿于事前预防机制缺位、平台监管权责失衡、执法权限配置疏漏三重积弊,导致特区网络空间乱象频出,治理质效不彰。究其原因,在历史因素上,“言论自由”思潮与“网络秩序”目标尚存价值抵牾;在规制思路上,并未足够重视平台对网络信息生成传播的中介影响;在规制方式上,并未充分发挥平台在网络信息治理中的关键作用。为此,香港网络空间治理亟需实现从二元规制模式向“政府—平台—个人”三元规制模式的转型和飞跃。具体而言,可以从以下四个方面协同推进:一是建立完善的网络治理法律规范体系,弥补现行法律框架在有害信息认定、举证及执法程序中的不足;二是夯实平台监管责任,包括明确平台内容管理义务、建立事前审查制度、增设事后处置通报机制,并加重平台监管失效的法律后果;三是强化行政执法力度,提高政府对网络信息治理的主动性和精准性,确保监管措施的可执行性;四是设立域外平台管辖保障机构,解决跨境执法难题,增强政府对境外社交平台的监管能力。

[关键词]香港特别行政区 网络信息治理 言论自由 二元规制 三元规制

[中图分类号] D911.01 **[文献标识码]** A **[文章编号]** 2096-983X(2025)03-0050-13

一、问题的提出

数字时代背景下,网络安全业已成为保障个人数据安全、维护社会和谐乃至国家安全的重要一环。自2019年香港“修例风波”以来,特别是全球新冠疫情爆发以来,香港社交网络平台不时被虚假新闻、政治谣言、煽动性信息、仇恨言论、恐

怖主义等有害内容侵蚀。这些有害信息^①不仅损害了香港市民的知情权益,甚至会动摇其对特区媒体甚至是特区政府的信任与信心。从2020年起,香港特区政府改变了长期惯行的不干预政策,加强了网络信息治理力度。从立法层面来看,陆续出台的《香港国安法》第9条、第43条第4项,《香港国安法第43条实施细则》“附表4

收稿日期:2024-09-18;修回日期:2024-11-21

*基金项目:2023年度教育部人文社会科学研究特别委托项目“香港规制社交平台危害国家安全信息研究”(JBF202304)

作者简介:沈新渊,讲师,法学博士,主要从事港澳基本法、数字法学研究;刘文轩,烟台大学副教授、优秀青年学者,南京大学犯罪预防与控制研究所特聘研究人员,法学博士,主要从事刑事诉讼法、数字法学研究。

①为了研究便利,本文主要以网络信息中的“政治性有害言论”“低价值的有害言论”作为研究对象。

关于移除危害国家安全的讯息及要求协助的细则”（下称“《附表4》”），《个人资料（私隐）条例》第9部《罪行及补偿》第64条刑事“起底”均涉及香港网络监管方面的内容。从执法层面来看，自2020年以来，香港警方通过网络在线巡查、多机构协作等方式，共查处利用Facebook、Instagram、YouTube等平台制作、散布有害内容的刑事案件多达21起^①。不可否认，如上举措对香港网络信息乱象起到了一定威慑效果。但从整体和长远来看，香港现行网络信息治理质效仍不甚显著。时至今日，“平台端”煽动性内容、仇恨言论、虚假信息仍层出不穷；绝大多数平台对香港警方发出的信息移除指令仍“不以为然”，主动履行有害信息制作者、传播者披露义务的平台更是少之又少。

实际上，香港网络信息治理成效不彰的根本原因在于其采用的“政府一个人”二元监管模式本身存在结构性缺陷。反观域外，德国、法国、英国、新加坡等国均陆续对网络社交平台展开了立法规制^②，在理顺“政府—平台—个人”三元关系的基础上，引入了平台规制模式，在预防有害信息生成、控制有害信息传播方面成效显著。本文在剖析特区网络信息治理模式现实缺憾与困境成因的基础上，依托规制理论与域外经验，提出香港网络治理三元规制模式的整体思路与具体方案，求教于学界同仁之余，以期助力香港网络空间治理质效整体性提升。

二、二元规制模式下香港网络信息治理之现实缺憾

网络信息治理框架一般涉及规制者、规制对象、规制命令（规制手段、规制目标）与规制结果四个要素。^③现行香港网络信息治理框架本质上可概括为“政府一个人”二元规制模式，

即香港特区政府凭借立法授权和自身的执法资源与能力，直接控制网络有害信息的传播。当线上危害引发线下“明显而即刻的风险”^④或对个体权益、社会利益及国家安全造成实质损害时，政府会动用公权力追究责任个体，并施以刑事处罚。这种源自传统农业和工业时代社会“个体化”的法律责任追究模式^⑤，虽然能够快速规范个人行为、短期内恢复社会秩序，但仍存在诸多缺憾。

（一）事前预防机制缺位

目前，香港涉及互联网信息具体监管措施的法律只有《附表4》和《个人资料（私隐）条例》第9部《罪行及补偿》第64条刑事“起底”。两部立法中均规定了网络服务提供者的信息禁制义务。以《附表4》为例，履行禁制义务的前提是警务处处长根据合理理由怀疑，筛选出相当可能构成或导致危害国家安全罪行发生的信息。经保安局局长批准后，授权指定人员要求平台服务商、主机服务商和网络服务商在指定期限内对相关电子信息采取禁制行动。这种规制机制属于互联网发展早期网络信息治理中的“通知—删除”模式。在此模式下，社交平台并不需要主动对平台内的信息内容作事前审查，只要在接到执法部门的信息移除通知后，及时对有害信息进行过滤、删除或屏蔽，即可进入“避风港”获得免责。质言之，社交平台仅具有被动配合执法的辅助地位，即使网络信息传播异化带来网络安全风险，也无须承担监管不力之责。只有当网络有害信息得到广泛传播，引起线上热议，或引发线下恐慌，甚至造成社会损害时，执法机关才会重视并采取相应规制措施。然而，在当今高度发达的互联网时代，信息的高速传播和广泛覆盖远超执法机关的响应能力。在应对突发事件或严重危机情况之时，这种事后监管模式不仅反应滞后，无法预防网络安全风险，还不能实时评

①数据来源：笔者以2020年以来港区危害国家安全的司法案件为样本，提取出因涉及社交平台信息内容获罪的案件数量。

②譬如，德国早在2017年出台《网络执行法》，法国于2018年颁布《反信息操纵法》，新加坡在2019年制定了《防止网络虚假信息和网络操纵法案》，英国于2023年发布《在线安全法案》。

③案件编号：Schenck v. United States[1919] 249U.S.47 (1919).

估网络环境,无法及时遏制有害内容的扩散。特区政府往往需要投入大量的法律资源和人力成本,并对有害内容采取严格的审查和制裁措施,这不仅加重了执法和治理的负担,还会加大“权力—权利”间的固有张力。不难看出,“政府一个人”二元规制模式“重事后监管、轻事前预防”的弊病尤为明显。

(二) 平台监管权责失衡

当前,港区颇为流行的社交网络平台,其运营商主要来自域外。以香港“修例风波”中助推力量较大的Facebook、Telegram、LIHKG等平台的服务行动轨迹察之,这些平台均未采取任何措施对有关香港服务区内煽动性信息、仇恨言论、虚假信息进行在先审查、删除或封禁。具体来说,LIHKG对言论自由持积极保障态度,欢迎不同政见用户发表立场言论,不会因政见及立场不同而无故封锁账号。Facebook虽建立了信息内容审查机制,但仅满足于全球内容审查的一般性需求,并未针对香港网络生态需求提供特定性审查。Telegram作为一种公益性、极具隐私性的社交软件,旨在为全球用户提供一个充分保障言论自由和私隐安全的社交空间。可见,香港的平台服务提供者漠视服务区内的网络生态环境,不会主动对网络有害信息进行筛查和处理,更不愿向香港特区执法部门通报违法活动、提供资料和技术协助。香港由于尚未制定专门维护网络安全的立法,也缺乏针对社交媒体平台的专门性规制法律,故而难以将平台的不作为行径视为违法。虽然《附表4》规定了社交平台须履行禁制义务和协助义务,但该规定之下的平台义务负担较轻,责任豁免幅度较大,不仅无须检测并消除网络服务平台中那些不能为网络用户所知的安全隐患,也无须事先采取必要的措施防止损害的发生或减轻损害造成的后果。可以肯定的是,在立法层面,网络平台责

任划定的严重失衡,加重了社交平台对社会责任的漠视与消极姿态,也同时间接助长了网络有害信息的滋生与传播。

(三) 执法权限配置疏漏

依据《附表4》之规定,警方可以要求相关平台或服务提供者移除或限制访问被认为是危害国家安全的信息,要求平台或服务提供者提供发布相关信息的用户的身份信息和其他相关资料,要求相关人士或机构协助调查、提供资料,以及要求网络服务提供者、平台服务商提供技术支援。然细究不难发现,警方执法权限不甚充足,强制力颇为薄弱,实践中平台服务商不配合、不作为甚至抵制的原因大抵如此。以“愿荣光归香港”禁制令一案^①为例,首先,依据《附表4》之规定,若服务商不遵从警方要求,警方并没有其他直接强制手段要求平台服务商履行禁制义务,除非通过起诉将其定罪。然而,这一过程相对漫长。其次,服务商更愿意接受法院的禁制令,但申请禁制令的前提是制定法中有明确规定,遗憾的是,《附表4》中并无禁制令的相关规定,这就意味着警方难以获批禁制令申请。实践中,警方申请法院禁制令须率先通过《香港国安法》第47条规定的行政机关证明文书程序,再经由法院审议并批出禁制令。这一来二往,警方获批禁制令还存在大量的时间成本。如果有害信息不能被界定为危害国家安全,警方还无法获得法院禁制令。假设警方最终以起诉方式迫令平台服务商履行有关义务,但由于《附表4》对平台服务商规定的惩罚措施为10万港元和监禁6个月,威慑性过低,^②平台也很有可能拒不执行。此外,平台服务商在香港并未设置联络代表或办事机构,行政执法和司法送达均存在管辖困境。可以发现,特区政府在网络信息治理上受到的牵制较多,执法颇为被动。虽然《香港基本法》规定了香港政治体制中行政、立法与司法既互相配合又互相制

^①案件编号: HCA 855/2023.

^②反观域外,德国《网络执行法》规定对社交网络平台最高5000万欧元的罚款措施;新加坡《防止网络虚假信息和网络操纵法案》规定了最高100万新加坡元的罚款和10年监禁的惩罚力度;英国《在线安全法案》设置了最高为社交媒体平台全球年收入10%或最高1800万英镑的累进罚款,以及对高级管理人员监禁的惩罚力度。

约的原则^[3]，但也强调“行政主导”的要求。^[4]从现有的行政执法权配给来看，特区政府在网络信息治理方面的“主导性”存疑，无法满足维护网络安全的要求，也与基本法精神相去甚远。

三、二元规制模式下香港网络信息治理困境之深层诱因

（一）历史因素掣肘：“言论自由”思潮与“网络秩序”目标尚存价值抵牾

信息内容属于言论自由的范畴，而网络监管会触及言论自由的核心价值。香港延续普通法的传统，重视保障言论自由。在“一国两制”框架下，《香港基本法》第27条、《香港人权法案条例》以及《公民权利和政治权利国际公约》保障香港居民享有言论自由权，香港社会极其重视对言论自由的保障。这种价值不仅具有法律规范的基础，更具有来自历史层面的沉淀。由于殖民时期，港英政府曾对香港的言论自由实施长达一个多世纪的高压审查^[5]，使香港社会饱受“人权”摧残。因此在回归后，香港社会格外珍视这一核心价值，并将其设定为一项受基本法保障的宪制性权利。目前，香港言论自由的审查主体为法院，只要不涉及“一国”这种原则性问题，法院对公民言论颇为包容。

除了受言论自由思潮的影响外，香港特区作为国际金融中心，还需要维护其作为自由港的地位，保障信息的安全和自由流动。在资本主义的政治经济体系中，信息自由是重要原则。根据《香港基本法》第151条规定，香港可以“中国香港”的名义继续保持和发展对外关系，签订和履行有关协议，维护其信息中心的地位。为确保市场充分竞争，香港的电讯市场全面开放，不设外资限制，信息得以自由流通，内容传递不受审查，且个人资料隐私得到适当保障。若政府强化市场干预，在技术上对网络信息进行封锁或限制，可能会加重网络服务提供者的责任，影响互联网产业创新，导致网络服务商疏远投资或撤出香港，进而损及香港国际金融中心地位。

（二）治理思路偏颇：并未足够重视平台对网络信息生成传播的中介影响

社交网络平台所产生的旁路效应（Bypass Effect）^[6]为香港用户带来了自由讨论公共或政治话题的便利。但缺少监管的社交平台容易沦为不法行为人集中宣泄仇恨、煽动组织暴力的网络“策源地”^[7]，足以对个人利益、网络安全、社会稳定乃至国家安全造成危害。具体来说，网络平台是一个社会关系薄弱的虚拟空间。进入网络平台的个体都以匿名方式传递信息、互动交流。在网络虚拟世界中，网民更容易表现出对他人、社会的冷漠、恶意或暴怒。公众的盲从性和情绪化会对网络信息产生碎片化、片面性认知，轻易被有害信息所“裹挟”。而一些群体更为了博取流量、吸引关注或激发仇恨，借助网络的特点借机炮制话题、煽惑人心、混淆视听，形成网络化攻击。凭借平台的瞬间集聚和传播效应，有害信息的风险被顷刻放大，届时平台内以假乱真、仇恨弥漫、舆论升级。加上网络信息问题是现实社会与网络社会发生“跨界性”连结所产生，并未与现实社会所割裂^[8]，经过网络发酵的社会问题，最终会从线上传导回线下，对公民权益甚至是社会安定造成损害。

实际上，在互联网发展初期，特区政府就已提出“硅港计划”，试图“把香港发展为国际创新科技中心”。但随着2000年科技产业股泡沫席卷全球，互联网版块进入萧条期，特区政府最终搁置科技转型计划，并削弱了对网络科技产业的投入力度。由于当时的香港正处于房地产、金融业红利期，特区政府也长期将大量优质资源投放至资本市场，从而导致其监管重心偏离网络。当全球的互联网发展浪潮席卷香港，香港的互联网发展需求刚好与美国互联网产业扩张无缝衔接。特区政府为积极“引进外资”，增强海外企业对香港营商环境的信心，巩固其国际金融中心及自由港地位，对互联网行业长期持不干预政策。正因特区政府在互联网发展初期对网络空间治理缺乏重视，后期更忽视了社交平台的场域作用^[9]，未能预判到社交平台强大的动员和组织力

量, 所以才会导致其网络空间秩序紊乱。

(三) 治理模式陈旧: 并未充分发挥平台在网络信息治理中的关键作用

“政府一个人”二元规制模式的底层逻辑在于: 在农业和工业时代, 信息传播主要依靠实体媒介。有害信息的散播者通常是具体的个人或实体, 责任主体明确。信息传播的范围相对有限, 多限于当地社区或特定社交圈。受侵害的主体之间的社会关系较强, 信息传播的影响范围较小, 便于追踪和处理。因此, 政府凭借其执法资源和执法能力能够解决传统社会问题。

但随着互联网的普及和信息传播方式的变革, 网络的去中心化特质使得信息的生成和传播主体更加分散, 许多内容以匿名方式发布, 增加了责任主体追踪难度。网络传播速度迅捷广泛, 信息可在瞬息内传播到全球, 传播范围远超传统媒介, 管辖过程将面临跨平台、跨国界的技术和法律问题, 需要更多专业的技术手段和国际合作, 这增加了政府管辖的复杂性。网络信息浩如烟海, 要实时监测海量级网络有害信息并在有害信息溢出效应之前及时管控, 仅凭政府有限的执法资源和信息技术能力, 难以实现。此外, 在网络聚合效应影响下, 有害信息的传播会导致严重的网络群体极化现象。伴随算法推荐技术的加持效应, 网络用户的观点内容更加类型化和同质化, 信息内容容易被推向极端而发生异化。受信息茧房的影响, 个体在群体非理性的网络空间内易被极端观点裹挟, 责任感降低, 情绪激化, 仇恨扩散, 导致从众参与非理性网络表达和集体行动。当个体淹没在群体之中, 个体责任就难以从群体中剥离开来, 导致离散个体与危害结果之间的因果关系不易界定, 执法机关只能选择性地对有恶劣表现的信息制作者或传播者定罪量刑, 无法快速、精准打击“始作俑者”, 也难以对群体追责, 降低整体性规制效果。这充分说明“政府一个人”二元规制模式积弊较多, 短板显著, 既不能起到事前预防网络安全风险的作用, 也无法实现事中全面监测、事后准确追踪违法责任主体的目标。

由此可知, 香港特区政府在网络信息治理过程中, 忽视了社交网络平台在网络有害信息生成和传播过程中的助推作用。其传统的事后监管模式又赋予了社交网络平台消极被动的角色, 对平台责任承担设定为高度豁免, 使本该承担着管理和控制内容传播的“网络守门人”^[10]置身事外。以此表明, 在Web3.0时代, 传统的“政府一个人”二元规制模式已经与数字时代的网络信息治理需求存在严重脱节现象, 这充分印证了二元规制模式的落后和失宜。为此, 特区政府亟需转变监管模式、革新监管方式。

四、香港网络信息治理模式之应然转向: 三元规制

社交网络平台在信息技术、用户数据、组织资源等方面具有天然优势。在技术赋权条件下, 对于网络信息内容的治理, 社交网络平台可以发布和执行平台规则自主调控平台生态环境, 真正采取有效措施管控有害信息的生成和传播, 起到治理效果。根据网络信息治理的普遍经验, “政府管平台、平台管用户”已成为一种默认共识和有效方法。^[11]对网络信息内容进行规制其实是“基于媒介”进行内容管理的一种方式。^[12]为此, 香港网络信息治理应及时引入平台规制, 从“政府一个人”关系下的政府直接监管信息模式转变为“政府—平台—个人”关系下的政府利用平台间接管控信息的模式。

(一) “政府—平台—个人”三元规制模式之理论溯源

在数字市场中, 网络平台扮演着关键角色, 既维护市场秩序, 又管理平台用户。这种核心作用使得越来越多的观点认为, 网络平台应当成为数字市场的守护者, 承担更多维护数字市场秩序的义务。将平台视为解决数字监管问题的中介主体, 经历了“平台是数字社会的看(守)门人”^[10]“平台是新的公用事业(公共设施)”^[13]“元规制”理论^[14], 以及“市民社会的联邦主义”^[6]。不管是何种学说, 都强调了当国家行

为能力受限、监管效用不足时，将监管活动的中心从国家主体转向非国家行为体，利用中间第三方主体的便利和能力，同时将非国家行为体的能力和责任相结合，行使有效的中间第三方监管，便可以弥补公权力缺陷、有效阻止违法行为、实现监管目的。

网络信息治理视域下，网络平台的“居间守门地位”并不是特定形成的，而是因其掌握的技术或作出的社会行为被立法授予的。例如，德国《网络执行法》为平台服务提供者设置通报义务、对违法内容的处置义务，并且提高处罚惩戒力度来强化网络平台服务提供者网络信息安全责任。英国《在线安全法案》将互联网服务提供商划分为I类服务和II类服务。I类服务提供商需要积极主动采取措施防止个人通过平台服务接触非法内容，管理和减轻非法内容用于犯罪或对个体造成的伤害，运用系统技术和流程服务尽量缩短非法内容暴露在网络空间的期限。当收到非法内容投诉，平台服务提供商应迅速删除该内容。欧盟《数字服务法案》要求平台对用户生成的内容进行审核和管理，移除非法内容，尤其是涉及仇恨言论、恐怖主义和儿童色情等内容。我国大陆地区《网络安全法》要求网络平台对其用户发布的信息进行审核，删除违法信息，防止其传播。《网络信息内容生态治理规定》明确规定平台应当建立健全信息审核机制，及时处理违法和不良信息。可见，现代数字化发达的国家或地区正是发现了平台在网络信息治理中的关键作用，从而赋予了平台合规审查义务，倒逼平台积极开展信息内容自主审查处置行动。

“政府—平台—个人”三元规制模式涵盖了政府对平台规制和平台自我规制的双重监督样态。香港传统的二元规制法律体系已为政府规制设置了多层次的监督机制，但在平台治理规范上呈现出规制盲区。因此，三元规制模式的研究重点强调要对平台自我规制进行规制。简言之，三元规制模式强调私主体在规制中的积极作用，旨在通过政府外部力量的引导和推动，促使被规制主体

制定内部规则，以弥补传统规制模式的不足。与二元规制模式相比，三元规制模式为被规制主体提供了更多的自主权，又允许政府参与规则制定过程。特别是在被规制主体对其面临的问题或对象的了解超过政府的情况下，三元规制能够更有效地找到经济合理的解决方案。

（二）“政府—平台—个人”三元规制模式之主要特征

经验证实，当网络用户存在严重违法行为，政府存在海量信息、网络技术的监管局限，如果加强监管力度则会带来更多的执法成本，既可能对无辜用户造成过度威慑，也可能无法有效阻止违法行为。利用网络平台这一中介的资源、信息或力量，能够防范不当行为的产生、有害信息的传播，可以形成有效监管方案，补足政府监管能力，达到监管目的。因而，赋予平台“数字守门人”义务实则是一种典型的替代执法。^①法院也承认替代执法模式已经成为解决互联网有害信息的有效办法。^②传统的执法手段不能有效解决互联网有害信息的传播和扩散，网络的全球性也导致同一种有害信息在不同地区是否构成犯罪的法律定性不同。将信息处理最终决定权回归到在线平台运营商手中，是一种“逆向思维”的监管模式。基于互联网信息内容治理的新模式，国家、互联网平台运营商和网络信息内容制作者形成了“政府—平台—个人”新型关系。其中，执法部门转介执法权属于有效的行政行为，如果互联网平台运营商不承担替代执法责任，政府也可以直接启用强制监管。

此外，并不是所有类型的网络平台都能被赋予“数字守门人”地位。一般而言，只有“权威守门人”^③级别及以上规模的网络平台才会被授予。例如，欧盟《数字服务法》就指定Facebook、Instagram、X(Twitter)、YouTube等拥有高流量和庞大市场份额的超级节点可成为“数字守门人”。这些“权威守门人”在拥有高流量、高信息流的同时也拥有雄厚的实力和先进的技术，能够根据其门控地位选择行使“守门”的程度。

①案件编号：HCJ 7846/19.

值得注意的是,平台规制具有监管优势的同时也伴生弱点。若无政府或其他外部调控的限制,而仅仅依赖平台内部调控,那么自我规制是很难成功的。^[17]因为企业经营的本质是追求利润。如果网络信息内容审核将显著增加平台企业的经营成本,或者平台企业本身漠视网络信息内容治理,亦或平台企业已成为具有垄断力量的头部平台,平台服务环境的优劣都不会影响其垄断地位,那么在没有任何外部机制约束的情况下,平台并不会自愿主动监管网络信息内容。对该问题的解决办法是强调政府在互联网监管方面的主导作用,对平台企业进行外部监督,为平台企业设定社会目标,赋予平台“公共守门人”地位,规范平台企业的经营行为。这种做法不仅要求平台对其疏于阻止违法行为的结果负责,还通过替代责任和高额罚款机制,促使平台在衡量违法成本和法律责任之间作出取舍,主动进行私人执法。这种反向倒逼机制可以推动平台在公共目的下更有效地履行其守门人职责,防止违法内容的传播,维护数字市场的健康和有序发展。

典型国家的实践也证明了引入平台规制需要增加政府干预,强化平台责任的重要性和必要性。例如,英国《在线安全法案》强调要加强社交媒体和其他在线平台的问责制,对受监管的内容须负注意义务和承担谨慎责任。要求社交媒体平台必须在通信管理局(Ofcom)发布的业务守则框架内行事,也要在Ofcom的指导下履行法案所规定的义务和职责,每年度向Ofcom提交透明度报告,同时在Ofcom的要求下派出技术人员向Ofcom具体阐明有关事项。Ofcom有权主动进入在线平台和网络服务提供者内部,调查其是否遵守法案规定的义务。这些网络服务提供者必须充分配合调查。Ofcom也可依法向网络服务提供者的主管人员发出通知,要求该人员在制定期限内到场接受质询。Ofcom还可以暂停不合规平

台在英国的运营,或撤销有关许可并限制其在英国市场的服务活动。因此,落实政府主导平台规制,可以规范市场中特定市场主体活动的行为,使市场主体能够按照政府规制的意图运作,达到政府通过规制实现社会正义的目的。

综上,“政府—平台—个人”三元规制模式的首要特征是引入平台规制。当政府监管出现执法资源和执法能力有限的情况,可以利用平台的技术和数据优势,弥补政府监管的无力之处。其次是赋予平台“守门人”义务,落实政府主导,强化平台责任。在整个网络信息治理过程中,由于平台自愿性规制激励效应不足,可以强化政府外部监督力度和督促效应,倒逼平台主动履行信息审查合规义务。

(三)“政府—平台—个人”三元规制模式之积极功能

1.平衡自由、秩序价值,推动网络产业创新

遍观数字法治发达区域,欧盟的《数字市场法》在制定理念上就展现了对自由主义秩序观的回溯。^[18]自由主义学说赋予公正的行为规则以重要性,且是法律规则维护自发秩序的基本条件。在数字经济法律治理中,虽然以美国和欧盟为代表的两大法域对数字市场治理形成了两种不同的自由主义面向,^①但自由主义的核心还是社会“不该被过度治理”。自由主义并非不治理或减少治理,而是倡导有节制的、审慎的、合理的治理,并时刻提醒以国家名义的理由之政治合理性要在合理范围和界限内实施治理。^[19]从比较研究的视角来看,德国、英国、新加坡、法国等典型国家均对数字市场和平台服务强化了政府监管,讲求数字经济秩序,较好地协调了与言论自由之间的矛盾。同样面临监管与自由话题的中国香港,认为对网络信息内容加强监管会进一步威胁香港的言论自由空间。虽然这种担心不是毫无道理,但这种认识具有将不同层级的问题混为一谈的嫌疑^②,

^①例如,“受芝加哥学派的新自由主义影响,美国整体的反垄断司法理念倾向于依赖市场的自我纠错机制,仅在极少数情形下进行司法干预。”

^②香港法院在陈宥羲案中认为,虚拟的互联网不具有“公共空间”性,不是“公共论坛”。政府并未限制公民表达言论,只是限制其表达方式;社交网络平台规制信息,是私法范畴,与公法中的言论自由限制体系属于不同领域。

同时也没有很好地理解监管和监督的意义。^[2]而在具体的制度设计和技术优化上，社交平台内部通常会设立用户投诉申诉友好型程序，当用户认为自己的言论被不当删除或限制时，可以通过申诉机制表达不满，申请复议。例如，Facebook内部的监督委员会，就是为应对日益增多的言论自由和内容管理上的争议所设立的独立组织。除了平台内部的权利救济机制，各国法律也规定了用户通过司法途径维护自身权益的程序。^①因此，香港的网络自由冲突也具备调和的空间。

在“政府—平台—个人”三元结构的互联网背景下，平台的自我规制机制是由企业或产业自发组织制定并在内部自愿实施，其内容、方式的选择以及内部监督激励机制的设计均由内部成员决定，而外部正式或非正式支撑体系构成其外在激励约束机制的一种特定规制形式。^[17]这种自我规制的特征在于较强的内生性和自组织性。其规制优势在于自由或自主度；规制主体拥有较高的专业技术水平、充分的“内部知识”和信息，规制针对性更强、有效性更高；可以依靠伦理标准、同行压力或自愿性的行为准则运作，更容易获得业界的自愿接受；在制定规则时不需要遵循严格正式的法定程序，具有便捷性和灵活性；规制的成本通常可以由被规制企业来承担，而不需要像二元规制那样由纳税人共同承担成本，更加公平合理。^[20]因此，与二元规制相比，平台自我规制负担较小，能够降低信息、监督和执行成本，更有利于提高企业效率和推动互联网产业创新。

2. 预防网络安全风险，控制有害信息传播

社交平台内部事前审查机制能够事先切断有害信息与网络社会的接达途径，预防有害信息的生成。用户在发布信息内容到网站上之前，平台会通过内容审核技术，对信息的编辑和制作过程进行监测和审查。若用户发布的是违反法律和平台规定的有害信息，平台会利用智能内容识别技术和过滤技术在信息上载的瞬间过滤掉

不符合法律和社区标准的内容，事先降低网络信息内容的危险程度，并在后台提示网络信息内容制作者，生成内容为违禁内容。若网络信息内容制作者无视平台规则，持续编辑生成有害信息，情况严重的，平台将在删除或屏蔽有害信息后，采取封禁技术直接封禁违规账号。正如2020年底至2021年初，Facebook和Twitter等社交平台封禁特朗普账号之例，展示了“断平台”（de-platform）的力量。^[21]因此，平台的事前审查机制能够有效预防网络安全风险。

社交平台内部事中监测机制能够控制有害信息的传播，防止负面影响扩大。平台会根据实时情况采取人工智能审查和人工审查相结合的在线审查模式监测网络信息内容。对具有敏感性、误导性或引起不适、不可控制的舆论信息，平台可以利用算法调控技术限制话题流量，控制信息传播的广度和深度。对于突发事件或紧迫危险，平台也可以利用过滤技术或封禁技术作出应急处置，令网络空间恢复秩序和保持平稳。

3. 引导平台积极履责，纾解执法被动困局

严厉的执法措施、严格的处罚力度会对平台形成强力威慑，倒逼平台积极开发更为先进的内容审查技术以避免承担审查不力的责任。例如，为防止触犯德国《网络执法法》而承担5000万欧元的高额处罚，Facebook投入了大量的时间和财力反复进行人工智能训练和机器学习用以信息内容审查。其人工智能技术可以在用户每天数十亿条的信息内容中反复检出并移除违规内容，且检测结果日益精准。在检测出违法内容时，Facebook会在收到用户投诉后的24小时内删除或阻止访问，如果是其他有害内容则会在7天内予以删除或者屏蔽。^[22]Facebook不敢再继续使用“否认、拖延和转移焦点”等惯技，在德国当地投入了其全球内容审查的六分之一人力专门执行德国的立法要求。2023年上半年，就有61149条内容被删除或阻止访问。^[23]可知，社交平台的

^①例如：德国《网络执法法》规定了社交媒体平台须设立机制，供投诉人及受影响用户对删除或不删除有问题内容的决定提出复检。平台营运商可向联邦司法办公室申请撤销对其的惩处，若未能成事，他们可向法院提出上诉。法国《反信息操纵法》规定就法官为停止传播某些网络内容所作出的命令，相关人士可向上诉法庭提出上诉。

内容审核和算法调控技术不仅审查速度惊人,而且报错率低,能够及时发现有害信息并迅速作出禁制措施,大幅度降低信息内容延时处置带来的社会成本。因此,采取严格执法措施,增加平台的违法成本,提高惩罚的威慑性,能够倒逼平台积极履责,及时禁制网络有害信息。

利用平台场域优势,要求用户遵守平台行为守则,能够提高公众的媒介素养和网络安全意识,帮助个人更好地理解 and 遵守网络规制。利用平台内部治理机制,要求平台审核和过滤用户生成的内容,能够有效打击虚假信息和仇恨言论,确保信息传播符合相关法律法规和社区标准。利用平台反馈机制,要求平台发布透明度报告,通报平台内容审核情况,可以随时评估平台“守门人义务”履行情况,以便执法部门及时调整执法力度,缓解执法被动状况。强化平台责任能够促使平台配合政府的法律法规,落实有关政策。必要时,可以增强平台披露数据信息的动力,以便政府更有效地进行监管和执法。

五、三元规制模式下香港网络信息治理之优化进路

诚如上文所述,香港亟需实施“以网管网”策略,强化平台责任、落实政府主导,如此才能有效治理网络信息、营造良好网络生态环境。但亟需解决的首要问题是监管限度问题。对于“低价值言论”的监管,因其属于“受到准确定义并严格限制的言论类型,其禁止和惩罚从来不被认为产生任何宪法问题。”^[24]但对于“高价值言论”的监管,特别是对香港网络空间的政治言论的监管,由于宪法保护的政治性言论和危害国家安全信息、煽动性言论、仇恨性言论之间的界限模糊,而严苛的平台合规审查义务会导致平台在实践中“层层加码”,形成“寒蝉效应”(Chilling Effect),进而危及用户言论自由。虽言论自由非绝对权利已成全球共识,但香港仍需考虑其国际化都市定位,避免过度监管,损害其开放性。由此,香港应坚守《香港基本法》言论自由基本原则,在

信息监管的严格程度上,可适度低于西方国家,如此既可保障其自身开放性地位,也可反击西方国家抹黑。在坚持适度监管的原则下,香港网络监管可从以下几个面向协同发力。

(一) 构建网络治理系统法律规范体系

在网络信息监管上,中国、美国、欧盟等典型国家或地区均出台了多项法律法规,形成了自成体系的规制模式。例如,美国出台了《爱国者法》《通讯规范法》等;欧盟发布了《处置网络虚假信息:欧洲进路》《欧盟反虚假信息行为守则》等;德国有专门性立法《社交网络中的法律执行改善法》;中国大陆通过了《网络安全法》一般性立法、《网络信息内容生态治理规定》《网络暴力信息治理规定》专门性规章等;英国有《在线安全法案》;法国制定了《反信息操纵法》等。这些国家或地区既规定了一般性的指导原则,又明确了社交平台信息内容规制的具体措施、程序以及人员配置等。

香港作为一个法治高度发达的普通法地区,特区政府的一切行为都奉行“法律主治”(Rule of Law),推崇“法律的至尊性”(Supremacy of Law)。^[25]在《香港国安法》颁行后,特区政府虽逐渐收紧网络信息监管力度,但却没有制定一部维护网络安全的一般性立法或一部针对社交平台的专门性法律作为执法依据。只有《附表4》直接涉及危害国家安全的网络信息具体监管措施。但《附表4》属于《香港国安法》下的后端规定,其主要任务是为了解决问题而非预防问题。其本身的属性就决定了它无法成为一个专门的针对社交平台进行规制的系统性法律。除了《附表4》以外,香港很多现行的法律是在互联网普及之前制定的,难以适应当前快速变化的网络环境和技术发展。因此,这些法律在应对新兴网络问题时显得颇为滞后和无力。典型国家的实践已经充分显示了网络安全立法的重要性和必要性。香港特区政府应借鉴典型国家的立法经验,尽早建立网络安全“一般性+专门性”法律规范体系。为此,建议制定《香港网络安全条例》,为网络监管搭建框架;制定《香港社交平台信息

安全条例》，明确社交平台信息规制的具体措施和程序，为政府执法、平台合规提供明确的权力（利）运行依据。

（二）夯实网络平台信息监管责任

1. 明确平台信息内容负责制度

《附表4》为社交平台设置了较高的责任豁免，其因缘可溯至美国《通讯规范法》第230条（下称“230条”），即网络平台尽到通知删除义务，便可对平台上用户发表的言论享有责任豁免。然互联网发展至今，230条已成为网络有害信息泛滥的温床，故而遭受存废争议。例如，美国最高法院对“冈萨雷斯诉谷歌”^①和“Twitter诉塔梅”^②两案迟迟不作裁判，其主要原因是考虑“互联网公司是否要为自己平台上用户发布的内容承担责任”这一核心问题。2020年美国总统特朗普的“预防审查制度”行政令以及美国司法部提出要对230条进行修改，都表明了平台责任高度豁免带来了严峻的网络信息安全问题，也阻隔了法律对互联网平台合理的规管路径，需要转变这一制度。此外，230条的责任豁免也存在例外情形。美国于2018年出台了《授权各州和受害者打击网络色情贩卖法案》和《停止授权性贩卖法案》，规定执法人员在得知有关平台淫秽色情内容泛滥，而该平台存在罔顾、怂恿或协助意图的情形下，州检察长可以代表该州居民直接提起民事诉讼或刑事指控，要求平台承担责任。这种信息内容负责制类似于“公益诉讼”，能够对社会的公共利益进行有效救济。

香港可以要求社交平台对受监管的内容负担注意义务和承担谨慎责任，探索建立社交平台对有害信息造成的社会危害的负责制。例如，在2019年“修例风波”中，反中乱港分子利用Facebook、Twitter、LIHKG、Telegram等社交网络平台制造“黑暴”事件，致使香港社会动荡不安，国家安全遭受严重威胁。香港检控部门有理由相信这些社交平台罔顾或协助反中乱港分子破坏中国的国家利益和香港的公共利益，可以在

前设法律基础上对涉事平台提起民事或刑事指控，要求其为香港的损失承担相应责任。

2. 建立平台事前审查制度

平台企业内部拥有较高的专业技术水平、充分的“内部知识”和信息，具有一定的组织行为规范，能够制定和实施组织内部的运作规则，并能有效约束被规制群体，使规制的针对性更强、有效性更高。而且，平台内部实施信息的在先审查治理机制，其成本通常由企业自身承担，能够减轻执法部门负担。因此，要求平台企业内部建立并执行正式的信息内容在先审查机制，采取责任倒逼方式，将“政府主动”转为“平台主动”，发挥平台自我规制方面的技术优势和数据优势，实时预警监测，提高用户网络安全意识，能够满足香港网络信息治理需求，提高网络监管成效。

具体而言，可以要求社交媒体平台对其提供的服务所载的内容进行非法内容风险评估，采取适当措施使非法内容风险评估保持最新状态。这种风险程度评估指标包括用户群体、每种非法内容单独评估、平台服务功能被用于实施或协助违法犯罪的可能性、不同种类的非法内容对个体的危害性、识别和评估存在较高风险的功能、社交服务的设计和运营等方面。可以要求平台制定明确的平台内容行为守则，采取有力措施防止个人通过平台服务接触非法内容。建立用户友好型投诉渠道，鼓励用户向平台投诉非法内容。当收到非法内容投诉，平台服务提供商应迅速删除该内容，做好有害信息的处理和数据存储。

3. 增设平台事后处置通报制度

为了让政府和公众了解平台在信息监管和内容处置方面的行动细节，避免信息处理过程中的不透明或不公正，防止平台滥用权力或不作为，要求平台定期发布透明度报告，通报平台在处理信息时所采取的具体措施、适用依据以及解释说明，有利于监督平台行为，确保其遵守

^①案件编号：Gonzalez v. Google, LLC, No. 18-16700 (9th Cir. 2021).

^②案件编号：Twitter, Inc. v. Taamneh, 598 U.S. 471 (2023).

公平的内容管理和处置程序。例如,德国《网络执行法》规定,社交网络平台应当将网络信息处理情况及时通报监督管理部门并定期向社会公开处理报告。凡是每年收到100起以上投诉的社交网络服务提供者都有义务每半年制作一次报告,并于制作完成后一个月内在联邦司法部公报以及平台网站首页上予以公布。报告义务应全面详尽,至少包括用户递交投诉的机制以及删除和屏蔽的判断标准、投诉数量、处理投诉的人员组织情况、最终被删除或者屏蔽的信息数量、让用户陈述的情况、转递其他专门机构的处理情况、处理周期以及通知投诉者和用户的情况。^[22]

借鉴德国《网络执行法》之规定,可以要求在香港提供服务的社交网络平台做好用户数据存储,制定内容处置报告,并定期向社会公布。同时,与政府保持友好联系和密切合作,通报信息审查、认定和处理情况,必要时共享个人信息数据。例如,可以规定社交平台必须每半年编写一次关于非法内容处理情况的报告,公开期限须在半年期结束后不少于一个月内,并置挂于香港特区政府新闻公报或官网。在官网上发布的报告应易于识别,可以直接访问或可永久查用。

4.加重平台规制失灵法律后果

法律责任和处罚力度的设定必须以事实为依据,与违法行为的性质、情节以及社会危害程度相当,遵循“大过重罚、小过轻罚、同过同罚”的过罚相当原则。^[26]大型社交网络平台拥有高流量、庞大用户量和广泛传播空间,对社会的影响极为深刻。若设置的处罚力度较轻,则对平台履行义务的威慑性和惩罚性不足,也会削弱执法的强制力。《附表4》规定的惩罚措施是最高为10万港元罚款及6个月监禁,与德国的《社交网络中的法律执行改善法》、新加坡的《防止网络虚假信息和网络操纵法案》以及英国的《在线安全法案》规定的惩罚力度相比明显过轻。在此,建议香港借鉴域外经验,提高社交网络平台违法惩罚力度,具体可根据平台的体量、市场流量,调整处罚额度,增加平台违法成本。

(三)强化网络治理行政执法力度

香港应以政府规制为主导,强化行政机关执法督促力度,监督平台的自我规制。对于香港本土的社交平台,如LIHKG,网络监管立法具有直接适用性;但对于Facebook、Telegram等域外社交平台,存在监管不能的可能性。若社交平台不能适应政府监管政策,或继续采取不配合、抵制的态度,政府应先与之协商,表明政府的监管动机和对基本权利的保障承诺,争取平台的协助。如果有关平台坚决表示反抗,特区政府应当严格执法,强制要求平台配合指令,而不需要再受制于行政机关证明文件和司法机关禁制令的双重程序。特区政府也可以约谈社交平台代表,在严重情况下,执法机关可以强制介入平台企业内部进行调查。在用户存在明显违法的情况下,可以强制要求社交平台提供个人信息数据。在极端情况或紧急情况下,执法机关可以强制要求社交平台封禁账户或屏蔽信息,甚至可以直接要求主机服务商切断数据存储或托管服务,网络服务商切断终端用户的网络传输服务。根据平台不配合执法的严重程度,追究平台违法责任,施以严厉处罚,或迫其主动退市,或对其进行封禁,限制其在香港地区提供服务,并借此机会推动新浪微博、微信、抖音等内地社交平台在香港地区的普及。

以Facebook和Telegram平台的监管思路为例:Facebook公司在德国提供服务时就基于商业利益作出了涉纳粹言论的特殊监管。虽然涉纳粹言论的管控与其“社群守则”精神相违背,但Facebook公司因不愿失去德国市场份额还是作出了一定妥协。香港可以借助Facebook内部成熟的信息内容审查机制,以香港市场份额和商业利益为筹码,要求其在香港提供服务时对某些类别的有害信息作出特殊监管。其中,网络平台所掌握的“地理封锁”技术就可以实现有害信息的特定监管。^[27]另外,对于较为特殊的Telegram公司,因其服务宗旨是为政治异见者提供发声平台,保障言论自由的极大化,不以营利为目标,不惧怕黑客破译和围攻,

协商招安不具备可能性，目前只能采取封禁。^①

(四) 设置域外平台管辖保障部门

由于Facebook、X (Twitter)、YouTube、LIHKG等社交平台的企业总部均不在香港地区，一旦执行有关网络平台信息安全审查法则，存在信息送达及联络沟通不便的可能。香港本土立法如何让域外网络服务提供商配合行政机关的执行命令？这将会是网络空间治理中跨境执法和司法实践的一大难题。虽然《附表4》规定了保护性管辖，但未指明具体的适用措施。香港的社交平台多为域外企业，且域外平台的市场份额占有量最大，又均未在香港设立办事处，虽被立法范围所覆盖，但域外平台企业的抵制、不配合行径已充分暴露现有立法执法效用的有限性。可以借鉴德国《网络执行法》规定的社交网络平台需在德国境内设置联系代表这种做法，要求香港社交平台在香港服务区内指定和授权联系代表和设置办事机构，并和政府保持联系，在行政处罚的情况下负责接收行政机关的处理决定和文书；在民事程序中负责接收法院的决定和文书；在刑事侦查机关和社交网络平台合作打击犯罪方面专门接受刑事侦查机关的询问。同时，应规定在确定的时间内作出答复，并详细叙述理由。

六、结语

现行香港网络信息治理“政府一个人”二元规制模式囿于事前预防机制缺位、平台监管责任失衡、执法权限配置疏漏三重缺憾，导致治理质效不甚显著。由此，本文建议香港网络空间治理注重挖掘网络平台监管自身优势，尽快实现从“政府一个人”二元规制模式向“政府—平台一个人”三元规制模式的转型、飞跃。囿于笔者研究能力及篇幅所限，本文呈现出“对策贡献有余、理论贡献不足”之问题，欢迎业内同仁批评、交流。

参考文献：

- [1]曹然. 从“规制”到“元规制”：社交平台用户生成内容治理的范式转型[J]. 苏州大学学报(哲学社会科学版), 2024, 45(3): 173-182.
- [2]劳东燕. 网络暴力刑法治理的基本立场[J]. 政法论坛, 2024, 42(3): 39-54.
- [3]王禹. 港澳政治体制中行政、立法与司法既互相配合又互相制约原则的探讨[J]. 政治与法律, 2018(8): 107-118.
- [4]曹旭东. 论香港特别行政区行政主导制[J]. 政治与法律, 2014(1): 11-21.
- [5]NG M. Political censorship in British Hong Kong: Freedom of expression and the law (1842-1997)[M]. Cambridge: Cambridge University Press, 2022: 1-196.
- [6]ABIRI G, GUIDI S. The platform federation[J]. Yale Journal of Law & Technology, 2023, 26(2): 240-323.
- [7]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.
- [8]刘少杰. 网络化的缺场空间与社会学研究方法的调整[J]. 中国社会科学评价, 2015(9): 57-64.
- [9][法]皮埃尔·布迪厄, 华康德. 实践与反思: 反思社会学导引[M]. 李猛, 李康, 译. 北京: 中央编译出版社, 1998: 133-134.
- [10]LEWIN K. Frontiers in groupdynamics: Concept, method and reality in social science; social equilibria and social change[J]. Human Relations, 1947, 1(2): 5-41.
- [11]李剑, 尹玉涵. 数字守门人理论: 溯源、误读与构建[J]. 社会科学研究, 2024(3): 11-24.
- [12]左亦鲁. 具体媒介具体管理——中国媒介内容管理模式初探[J]. 开放时代, 2023(2): 191-206.
- [13]高薇. 平台监管的新公用事业理论[J]. 法学研究, 2021, 43(3): 84-100.
- [14]PARKER C. The open corporation: Effective self-regulation and democracy[M]. Cambridge: Cambridge University Press, Chapter 8, 2002.
- [15]KRAAKMAN R H. Gatekeepers: The anatomy of a Third-Party enforcement strategy[J]. The Journal of Law, Economics & Organization, 1986, 2(1): 53-104.
- [16]BARZILAI-NAHON K. Gatekeepers, virtual communities and the gated: Multidimensional tensions in cyberspace[J]. International Journal of

^①例如, 巴西最高法院封禁X (Twitter), 这一事件表明, 当社交平台未能有效遏制有害信息时, 封禁可能成为维护社会秩序和保护公众权益的必要手段。

Communications Law & Policy, 2006, 11: 1-28.

[17]杨志强,何立胜.自我规制理论研究评介[J].外国经济与管理,2007,29(8):16-23.

[18]吴佩乘.数字平台反垄断的范式反思和规则调适——以欧盟《数字市场法》秩序自由主义面向为镜鉴[J].苏州大学学报(哲学社会科学版),2024,45(2):85-97.

[19]王少杰.福柯自由主义的逻辑谱系[J].社会科学家,2020,35(3):37-41.

[20]王茹.互联网经济规制的原则与多元规制体系的构建[J].行政管理改革,2018(1):42-47.

[21]左亦鲁.社交平台公共性及其规制——美国经验及其启示[J].清华法学,2022,16(4):95-111.

[22]查云飞.德国对网络平台的行政法规划制——迈向合规审查之路径[J].德国研究,2018,33(3):72-87.

[23]NetzDG Transparency Report (reporting period 1 January-30 June 2023)[EB/OL]. [2024-06-30].[https://scontent-hkg1-2.xx.fbcdn.net/v/t39.8562-](https://scontent-hkg1-2.xx.fbcdn.net/v/t39.8562-364139226_306330431844353_783773684531438147_n.pdf?_nc_cat=107&ccb=1-7&_nc_sid=b8d81d&_nc_ohc=GROVEleM1wYQ7kNvgFuiPrd&_nc_ht=scontent-hkg1-2.xx&gid=AtLCq7T3VxGW6HEQ9e33YLR&oh=00_AYDQ50L-6CDlv4rHLJFZ4V0GAuHyF3Ax10ov9O_QbQ0tyw&oe=668707C0)

364139226_306330431844353_783773684531438147_n.pdf?_nc_cat=107&ccb=1-7&_nc_sid=b8d81d&_nc_ohc=GROVEleM1wYQ7kNvgFuiPrd&_nc_ht=scontent-hkg1-2.xx&gid=AtLCq7T3VxGW6HEQ9e33YLR&oh=00_AYDQ50L-6CDlv4rHLJFZ4V0GAuHyF3Ax10ov9O_QbQ0tyw&oe=668707C0.

[24]张千帆.美国联邦宪法[M].北京:法律出版社,2011:384.

[25][英]戴雪.英宪精义[M].雷滨南,译.北京:中国法制出版社,2017:8.

[26]邹奕.行政处罚之惩罚性的界定[J].行政法学研究,2022(21):44-55.

[27]邹举,叶济舟.全球性抑或属地性:被遗忘权的空间范围争议与执行方案选择[J].新闻与传播研究,2020,27(8):75-93.

【责任编辑 苏聪文】

The Transformation of Hong Kong SAR's Online Content Moderation Model: From "Binary Regulation" to "Ternary Regulation"

SHEN Xinyuan & LIU Wenxuan

Abstract: The current "government-individual" binary regulation model for online contents moderation in Hong Kong SAR is constrained by three major issues: the absence of a preemptive prevention mechanism, an imbalance in the regulatory responsibilities of platforms, and gaps in the allocation of law implement authority. These situations have led to frequent chaos in the cyberspace of Hong Kong SAR. The underlying reasons can be traced to three aspects: First, historically, the discourse of "freedom of speech" in Hong Kong society remains in conflict with the governance goal of "cyberspace order". Second, in terms of regulatory thinking, the government has not sufficiently recognized the intermediary role of platforms in the generation and dissemination of online contents. Third, in terms of regulatory methods, the government has failed to fully leverage the key role of platforms in internet contents moderation. Therefore, it is imperative for Hong Kong SAR's cyberspace governance to transform and upgrade from the "government-individual" binary regulatory model to a "government-platform-individual" trinary regulatory model. Specifically, taking "moderate regulation" as the basic principle, efforts should be made in four directions: (1) Establishing a comprehensive legal framework for cyberspace governance; (2) Strengthening the online contents moderation responsibilities of social media platforms (including clarifying the platforms' responsibility for online contents, establishing ex-ante review systems for social media platforms, adding ex-post disposal and notification systems for social media platforms, and reinforcing the legal punishment of social media platforms' moderation failures); (3) Enhancing the administrative enforcement of cyberspace governance; (4) Setting up a specialized department for liaising with overseas social media platforms.

Keywords: Hong Kong SAR; online content moderation; freedom of speech; binary moderation model; trinary moderation model