介绍与评论

网络过滤技术的正当性批判*

——对美国网络法学界一个理论论争的观察

时　飞

内容提要:尽管立基于社会公共秩序的考虑的网络过滤技术有其合理的一面,但在互联网上设置网络过滤技术装置,即便是声称对信息自由流动损害最小、最大程度尊重网民自由选择、兼顾社会安定需求和言论自由的内容选择平台,但仍然无法避免恣意判断,并进而对公民言论自由权造成损害。其是否具有正当性,仍是一个富有争议的话题。尽管语境不同,但网络内容选择平台在美国引起的争议,应当成为我国在进行有关网络言论管制时所应当借鉴的参照系。

关键词:网络过滤技术　内容选择平台　事先限制　言论自由

时飞,华中科技大学法学院副教授。

关于自由言论的后果,耶鲁法学院的哈里·威灵顿教授指出:"乍看之下,直觉或许会暗示说,与他在其他领域所拥有的自由相比较,一个人所拥有的言论自由应当要更多一些。下面这句格言可能包含着一些真理:'棍棒和石头会令我骨头折断,但语言将永远也无法给我造成伤害。'而事实上,语言经常带来伤害。语言会侵犯一个人的名誉并对其造成伤害,煽起偏见或者是激情,可以'星星之火,可以燎原'之势传遍全世界。政府所管制的绝大多数其他行为,它们所具有的伤害性潜能与言论所具有的伤害性潜能相比较,无疑小巫见大巫。"[1]威灵顿在讲述这段话的时候,言论主要借助报刊杂志、电视广播等媒介进行传播。即便这些媒介中传播的言论具有伤害可能性,但由于它们事先都经由编辑审查,其危险性也就相应降低。在互联网取代传统媒体成为主导性的传播媒介后,言论对秩序威胁的风险也就相应增加。网络空间的信息编制,主要是由用户推动的,[2]为降低网络空间的危险言论给人们的日常生活、社会秩序或国家安全带来的冲击,仅靠内容上的分门别类并不足以解决

* 本文先后获得北京大学法学院互联网法律中心2008-2009学年度博士论文资助计划和易继明教授主持的"国家网络空间安全及信息内容安全监管体制研究"([242]2009A61)课题资助,特致谢忱。

〔1〕 参见 Harry H. Wellington, On Freedom of Expression, *Yale Law Journal*, vol. 88, 1979, pp. 1106-1107。

〔2〕 参见 Jennifer L. Brenner, True Threats: A More Appropriate Standard for Analyzing First Amendment Protection and Free Speech When Violence Is Perpetrated Over the Internet, *North Dakota Law Review*, vol. 78, 2002, p. 762。

言论内容的危险潜能。网络空间所带来的危险言论,其后果可能会是灾难性的。以网络威胁为例,"这种威胁不仅仅给受害人带来伤害,还会对网络空间的言论自由造成伤害:如果人们担心遭到报复的话,这就会使得他们就算进到了这个论坛里,也不敢贸然发表他们的看法。"[3]为防止不受宪法保护的威胁为祸网络空间,必要的技术支持有利于事先降低网络言论的危险程度,促进言论自由的健康发展。因此,利用过滤软件有效地消除网络空间的危险性言论对我们生活世界、社会秩序或者是民主政治的破坏性影响,也许就拥有了正当性的理由。但技术运用过程中无法逃避的命运是技术反过来吞噬孕育技术的自由。如果将所有的危险言论得到根除的希望全部寄托在网络过滤上,我们就是在制造一个"自由易逝,压制常存"的全息监控世界。

本文不欲全面分析网络过滤技术与言论自由的关系,只是试图梳理美国网络法学界对由万维网联盟推出的网络内容选择平台进行的争论,洞悉那种据说从技术层面对言论自由威胁最小但对社会公共秩序有着重大保护功能的网络过滤技术所隐含的对言论自由的威胁,说明除非理顺这种技术背后的法律机制,否则,网络过滤技术会通过自我强化的方式普遍化地压制言论自由。

一 为什么过滤?

过滤不是网络时代的新发明,媒介技术的每一次进步,都伴随着审查技术和过滤技术的发展。编辑是印刷媒介中最好的过滤装置;在广播、电视等电子媒介中,节目编制就是过滤装置,但人力仍有所不及,因此即使是美国,也通过专门立法规定必须在有线电视接口安装反暴力芯片,这种装置从三个层面强化对言论内容的审查:之于孩子,是一种阻塞性质的过滤装置,防止孩子直接和暴力或低级趣味的节目接触;之于父母,则是一种选择性的过滤装置,其目的旨在帮助父母确信哪些节目是孩子可以看的;之于媒介信息的制造者,它是一种组织性的过滤装置,是帮助制造者确信那些信息是可以不受反暴力芯片阻塞的。[4] 在网络空间,由于网络传播的技术特质,要阻止危险言论的传播,只能依赖自动化的审查和过滤技术。一种常见的过滤技术就是使用过滤软件,根据事先拟定的"黑名单",将那些列入清单的被认为非常不适宜的网站给过滤掉。另一种技术手段就是根据关键词来过滤相关内容。第三种技术就是通过分析网站内容的标签页来过滤相关内容,这种类型的过滤一般是由互联网内容选择平台实现的,在互联网内容选择平台中,过滤软件会有效地阻止那些根据这个平台筛选出来的内容网站上的信息,防止它进入不应当进入的个人生活世界里。

由于"因特网既不是自由的工具也不是一边倒的武器",[5]强制安装过滤技术就会带来重要的宪法难题。"黑名单"或以关键词为基准的过滤技术对言论自由的破坏作用远远胜过

---

〔3〕 参见 Prana A. Topper, The Threatening Internet: Planned Parenthood v. ACLA & A Context-Based Approach to Internet Threats, *Columbia Human Rights Law Review*, vol. 33, 2001, p. 228。

〔4〕 参见 J. M. Balkin, Media Filters, the V - Chip, and the Foundations of Broadcast Regulation, *Duke Law Journal*, vol. 45, 1996, p. 1143。

〔5〕 [美]曼纽尔·卡斯特:《网络星河:对互联网、商业和社会的反思》,郑波、武炜译,社会科学文献出版社 2007 年版,第 178 页。

它带来的秩序安定,它们是一种不折不扣的代表制度暴政的网络代码暴政。[6] 它们不应当成为保障网络自由言论不向危险言论或者是暴力威胁方向发展的工具。而通过贴标签来防止危险言论的互联网内容选择平台,也难逃偏见或歧视的诘难:"贴标签就是一种偏见态度以及类似态度的尝试,它是审查的工具。"[7] 当然,过滤技术未必就会钳制公民自由。防止不受欢迎的言论对人们生活世界的骚扰,在一定程度上还是能为这样一款软件提供正当性证明的。毕竟,在一个可能偏激化或群体极化的网络环境中,对一些不必要的信息的过滤,可能是面向民主参与的信息机制所必需的。

想要审查人们想些什么、表达些什么,这很困难,但要防止人们接收什么,相对比较容易。[8] 因此,开发一款便于使用的网络内容识别和选择的平台,让它帮助互联网用户选择性地控制在线内容就成为过滤软件的原初目的。但问题不在于它做了些什么,而是它是如何运作的。因为,一种建立在对网络言论内容歧视性的识别基础上的过滤技术,尽管言之凿凿地要保护受众的利益,但在一个由言说者、受众所形成的言论自由链中,过滤技术的目的能否实现,则远非清楚明了。

二 过滤技术规制言论的程序

万维网联盟推出的互联网内容选择平台,并不是过滤软件,它是一种确保分级系统和过滤软件能够兼容运行的代码协议。内容选择平台是要详细说明过滤软件是怎样与分级系统互相作用的,它本身并不会就具体言论的内容来设置分级标准或者是过滤有关内容。因此,人们往往会认为内容选择平台就是一种内容中立的网络言论管理程序,编制这套代码的初衷就是为对言论的管理乃是基于一种内容中立的用户的自我管理,它容许用户可以根据自己的偏爱、喜恶程度来设置自己拒绝的东西。

与直接给言论的内容分级并在此基础上过滤言论不同的是,内容选择平台的设计原理就是为了让网络内容服务商和独立的作为第三方存在的分级组织生产给予内容选择平台而形成的标签,而软件制造商则主要制造那些可以阅读这些标签并根据对这种标签的分析对有关言论进行过滤的软件。[9] 贴标签和过滤的生产环节彼此独立,定制者可以分别挑选软件及其标签资源。[10] 这样一来,就会形成两个独立的市场,彼此之间就识别内容的软件和如何给言论贴标签展开竞争,言论的集权化管制的危险就因此而实质性降低。[11]

内容选择平台要完成内容筛选和过滤,只需要贴标签然后过滤即可。[12] 其具体操作可以分为四个步骤:(1)由内容服务商提供内容定制,并且这些内容应当经过内容分级系统的

---

〔6〕 参见 Lawrence B. Solum Minn Chung, The Layers Principle: Internet Architecture and the Law, *Notre Dame Law Review*, vol. 79, 2004, pp. 908 - 909。

〔7〕 参见 American Library Association, Statement on Labeling: An Interpretation of the Library Bill of Rights (最后一次修改为 1999 年 12 月 11 日) http:// www. ala . org/ alaorg/ oif/ labeling. html。

〔8〕 参见 Paul Resnick, *Filtering Information on the Internet*, Science Annual, Mar. 1997, p. 108。

〔9〕 参见 Lawrence Lessig, *Tyranny in the Infrastructure. The CDA Was Bad—but PICS May Be Worse*, 5. 07 Wired (July 1997), http://www. wired . com/ wired/ 5. 07/ cyber rights . html。

〔10〕 参见 Keith Weidner, Mandatory Self-rating or Zoning: Which Way to Empower Parents? *Georgia Journal of Law & Public Policy*, vol. 2, 2004, p. 681。

〔11〕 参见 Lawrence Lessig, *Code* 2. 0, Basic Books, 2006, p. 256。

〔12〕 同上注。

不同形式的分级服务器的认可;[13](2)定制者可以根据他们的特定偏好选择分级服务器或服务系统;(3)定制者配置他们的固化内容选择平台的过滤软件,按照他们选择的内容分级体系创制一套"挑选规则";(4)软件根据网站上的内容等级和挑选规则之间的异同开始过滤。[14]

首先,通过内容分级系统来给特定的网络言论的内容进行分级。分级系统是一连串类型化的、序列化的软件,按照该软件的编码系统,网络言论可以在其中得到分门别类的级别定制;分级就是按照这种类型化、序列化的软件,用特定的语言、专门词汇来描绘特定的互联网上的言论的具体内容。内容选择平台的工作原理是,分级系统根据多元化的言论类型区分出不同的内容,并给予这些不同内容以相应级别的价值分,它并不是简单地给这些网上言论贴上是否合适孩子阅读的标签这样简单,它会给不同言论添加不同标签:暴力、威胁、性、赤裸、成人话题等不同级别的内容,得到的价值评分是不一样的。这种软件的最新发展是,即使是传统意义上的威胁,也同样可以给予不同的情景作进一步的区分:色情威胁还是暴力威胁;人身威胁还是政治威胁;有价值的财物损毁之威胁还是不特定物品的威胁。在伤害程度方面,是见血即可,还是身体组织的进一步伤害,或者是以动物的鲜血作为恐吓的手段,或者是外在于这些威胁手段的其他手段。这是带有价值评判的分级系统,它绝非简单地贴标签、设置黑名单或者是根据关键词实施过滤。

由于这种类型化、序列化的内容评级体系带有价值判断的因素,个人网站上的具体言论之分级体系的设定,可以由自己独立完成,也可以由独立的第三方代为完成。如果是个人独立完成,内容服务商应当提供一份问卷,该问卷应当按照内容服务商平日使用的标准来设计,根据个人在这份问卷上返回的答案,相应的标签就会嵌入网站的标题里。如果是独立的第三方来完成的,一个独立的分级服务器就应当按照当前的分级系统或者是根据特定的优先续或价值标准而开发的分级系统所设定的标准,给相应的网站上的标签赋予一定的价值权数。父母就可以使用这种已经反映了他们的价值偏好和特定价值取向的、且由专门组织定制了的标签。由于这种软件的开发商的目的是根据不同人的价值偏好、价值导向来设定内容识别系统的,它希望一个网站可以贴上多种多样的标签,只要每个人的价值偏好不同,同一内容的网站在不同人的分级系统中价值侧重点也就会有所不同。它试图防止单一化的价值偏好在确定每个人的个性化视窗浏览的时候,会约束、压制其他人基于另一种价值偏好和价值导向而需要的另一种不同风格的网络言论。

其次,每个人都可以根据其价值偏好,在众多的分级系统或分级服务器中挑选一个适合自己的分级系统或服务器。按照莱斯格的说法:"如果你想要遵循基督徒右翼的分级标准,你可以挑选这种分极系统;如果你想要遵循无神论左翼的分级标准,你就选它好了。一旦选定分级系统,我们就会把我们希望交由过滤软件加以过滤的内容挑选出来。"[15]

再次,定制者可以在这种以内容选择平台为操作基础的过滤软件上配置一套"挑选规则",他可以凭借这种分级系统中使用的每个类型的言论的最大化价值准则,把那些在分级系统中不曾出现的网站排除即可。除了个人能够自我设定之外,独立的分级组织也可以根

---

〔13〕 参见 Fernando A. Bohorquez, Jr., Note, The Price of PICS: The Privatization of Internet Censorship, *New York Law School Law Review*, vol. 43, 1999, p. 530。

〔14〕 参见 Keith Weidner, Mandatory Self-rating or Zoning, p. 681。

〔15〕 参见 Lawrence Lessig, Code 2.0, Basic Books, 2006, p. 256。

据一般价值体系和规范意涵,提供一些不需要亲自设定标准即可直接采用的"挑选规则",当然,前提是人们愿意相信这种组织的独立、公正和中立。

最后一个步骤就是开始过滤网络言论的有关内容。这是一个纯技术阶段,在这个阶段,如果某个网络用户试图登录某个网站,已经定制了的过滤软件就会根据已经设定了分级标准将这个网站的内容和定制者事前设定了的内容识别规则加以比较。如果这个网站的内容符合挑选规则中的排除规则设定的基本参数,用户就有权登录,但如果该网站的内容不符的话,则该用户将被禁止登录盖网站,且被告知这个网站已经被封锁了。

按照内容选择平台的设计,有三项因素极其重要:第一,根据价值偏好和价值导向来设定分级系统,并在此基础上给有关网络言论的内容贴标签;第二,定制者挑选一款适当的分级系统软件,并设定一组反映个人价值偏好和价值导向的挑选规则;第三,一旦有人在定制系统锁定的网络中登录网站,软件就会自动运行,并识别有关网站的内容,从而做出是否准予登录的决定。尽管研发这款软件的目的是为了方便定制者(尤其是父母保护孩子心理健康之需要)基于一定的价值偏好和价值定向而创制一个自己感到惬意的网络环境,但是,在技术发展的过程中,由于价值多元化和规范意涵的多层次性的不易把握,个体化的价值标准往往被统一化的软件编制者或者是独立的第三方分级组织的价值准则取而代之。尽管最后的操作权可能依旧被订制者持有,但他已经是一个用户,而不是这个程序的主导者和参与者。

三 个人化的内容选择向强制性过滤发展的可能性

从技术架构的结构功能来看,网络空间的基础架构包括如下几个方面:(1)网络技术协议(例如,TCP/IP),(2)网络技术协议的标准以及标准的具体适用(例如,网络浏览器或者是数字认证标准),以及(3)不轻易被改变的受保护的治理结构和社会使用模式。[16] (1)和(2)固然是整个网络空间的支配性技术架构,但只有(3)的存在才使得(1)和(2)的决定性作用得以发挥。这些基础架构并非固定不变,它们会根据法律管制、决定基础架构的政治经济结构的变化而变化。[17] 因此,尽管本意是"自我规范、由市场需求驱动的技术装置",但这种技术装置并非完全中立。尽管人们误信说,由于存在着个体化的控制功能,所以内容选择平台已经否定了政府管制网络言论的需要;由于技术扩展本身伴随着的管制权力的弥散化,已经使得在传统媒介中需要国家加以管制的言论,只需根据技术锚定即可解决。但在过滤技术的发展过程中,面临着激烈的竞争,而网络空间本身是一个持续分层的结构,缺少了硬件装置,代表民主参与的自由言论就可能无法在网络空间里发展。[18] 这些代表个人自由意志选择的软件,会受控于大型的网络服务商,是网络服务商而不是自由软件决定人们该如何

---

〔16〕 参见 Lawrence Lessig & Paul Resnick, Zoning Speech on the Internet: A Legal and Technical Model, *Michigan Law Review*, vol. 98, 1999, p. 397。

〔17〕 参见 Laurence H. Tribe, The Constitution in Cyberspace: Law and Liberty Beyond the Electronic Frontier, *The Humanist*, September-October 1991。

〔18〕 因此,尽管一直在强调网络协议层是怎样代替传统的媒介机制来实现民主参与、民主审议所必须的理性对话的,但是,弗努姆金教授还是着意提醒我们注意到"互联网上的软件并不是生存在真空状态中。它需要硬件支持其运转,需要(网络基础设施提供的)链接性能来支持其浏览功能"。参见 A. Michael Froomkin, Habermas@ Discourse. Net: Toward A Critical Theory of Cyberspace, *Harvard Law Review*, vol. 116, 2003, p. 858。

选择规避危险言论的渗透。

网络内容选择平台作为一种过滤技术，它只是过滤技术的开端。技术永远都做不到中立，它也不是单纯的对问题的直观反映。莱斯格指出："如果你建造了一个，那么，一大批的同类事物就会出现"；换言之，如果有了第一个技术协议准许贴标签和过滤软件形成互动格局的话，接下来，一个蔚为壮观的技术体系就会自动形成。[19] 没有一款软件能把所有人的价值偏好进行准确的预测并把它们都压缩到格式化的软件编码中；也没有一款软件对所有充满变数的价值偏好的限度能做出精准的计算，并对其扩张空间预留足够的编码变化。即使是愿意从事个人化的软件生产的软件商，也支付不起这种多元化的、随时充满变数的价值偏好的过滤系统所需要的核实成本。更何况，软件生产商还必须面对有多少人愿意购买这种耗时费日的软件来约束那不着边际的网络世界里的浩如烟海的信息流。

这就意味着，要创制具有变数可能的过滤系统，必然要求软件生产上投入可观的资源：(1)必须有一个组织来统一创制一个体系化的软件系统；(2)这个组织必须根据这个系统的编码定制个人网站内容的级别，不管它采取说服内容提供商对其所属网站进行自我分级，还是通过巨额投入时间、精力和金钱来设定这样一个分级系统，将所有网站实时运行的数据进行解码、分级；(3)系统必须对所有人公开。这就意味着如果指望社会大众会去购买这些软件的话，则无异于缘木求鱼，因此，没有任何一个组织担负得起这笔巨额花销，从而也就意味着真正对个人化的信息隔离有用的过滤技术实在有限。与此同时，要形成大规模的个性化的过滤软件，已经失去了有效的经济刺激，不会再有哪个逐利的软件生产商会去考虑这种分级系统，因为适于每个人的价值偏好和价值导向的具体化的软件，想要对症下药地被那些愿意利用它的人购买，这无异于大海捞针。想要建造一个多元化的基于个人价值偏好和价值导向为主的、以内容识别的标签化和随后的个体化过滤为核心的分级系统，首先就面临着经济效益上的重重障碍。

其次，个人也无法准确地识别自己的价值偏好，要将这些识别出来的价值系统按照一定的价值序列进行排序，也非轻而易举的事情。因此，确定何谓正当与不正当、适当与不适当、优先与靠后的价值设定基准，往往最后就会落到作为独立的第三方而存在的分级组织的手中。问题恰恰在于，独立的第三方不是一个独立的仲裁人，它也不是网络言论内容的定制者的法定代理人或协议代理人。如果我们还相信过滤软件生产的市场机制神话的话，我们就会发现，独立的第三方本身就是市场活动的一个主体，它识别价值偏好、给内容贴标签的活动就是市场参与机制的一部分。除非有明确的法律安排，内容服务商的自我贴标签行为对它的活动并不会带来任何收益的时候，它本身是缺乏任何动机去给自己的网站输出的言论贴标签的。[20] 这就意味着，并不存在一个完全以网络冲浪者的价值偏好为前提的第三方，第三方作为市场交易的一方，其设定的价值排序乃是以自己运行网络为前提的。

第三，指望网络冲浪者愿意进行自我分级，从网络空间的运行结构上来说，完全说不通。内容选择平台的运作程序中的价值前导，实际上很容易蜕变为第三方的自由裁量或恣意妄为，这样一来，即使初衷是好的，但在实际运行过程中，给予内容识别以一定的价值偏好，就和黑名单、关键词过滤一样，既简单粗暴又不敷足用。要想让这套软件运行良好，其前提就是绝大多

---

〔19〕 参见 Lawrence Lessig, What Things Regulate Speech: CDA 2.0 VS. Filtering, *Jurimetrics Journal*, vol. 38, 1998, p. 659。

〔20〕 参见 Timothy Wu, Application-Centered Internet Analysis, *Virginia Law Review*, vol. 85, 1999, p. 1185。

数网站都已经被分级了。英国一个网络过滤机制功能研究小组发现,基于个人价值偏好而形成的网络过滤软件如果其定制功能要得到全面发挥的话,只有等到没有受到分级审查的网站的比例足够低,而且它们究竟允许登陆还是不允许登陆已经在公共网站的醒目位置贴上显眼标签之后才可能,但这样一来,这种过滤软件的远处功能及其价值平衡机制,就会遭到彻底的破坏,过滤软件的功能发挥取决于它对这款软件的原初目的的破坏程度。[21]

第四,由于自我分级的运行机制本身的局限,除非政府以强制性的自我分级制度积极主动介入网络言论的分级制度中,否则,要想让一种个性化的过滤软件普遍化生产就几乎不可能。而政府的强制分级系统的建构,必须依托一整套规范化的法律装置作为后盾。而一如我们所知,政府不可能把个性化的分级设计机制作为维系其技术支撑法律体系的核心,因此,内容识别就不可避免地带上政府本身的价值选择标准。而政府如果想让一种过滤软件介入到人们的生活世界,是因为它认为这种言论本身已经是足以令人倍感憎恶的;或者是说这种言论已经偏离了政府所认为的安全港的基本规则太远。

因此,第五,如果导入强制性的法律安排,软件开发商的软件开发就已经不再单纯属于商业行为,而变为带有政府管制色彩的管制手段。商业机制一旦被权力机制所捕获,基于客户的价值偏好而提供的自由选择的系数就会降低。不仅如此,网络上基于内容识别的贴标签和过滤就可能变成这样两种非此即彼的选择:或者是整个软件开发产业都会聚集在一种统一标准下,依从这种标准给网络上的言论进行自我分级;或者是内容服务商按照所有浮现出来的分级标准来对网络言论进行分级,从而人人自危,导致整个网络空间的言论可能被多个版本的过滤软件在不同架构中加以磨灭。果真如此,人们就可能会在确认所有的过滤软件都不可取的情景下,不得不使用这些他们根本信不过的软件来过滤一些有人认为他们必须过滤的言论。

四 过滤背后的事先限制风险

恰如莱斯格指出的:"工程师编写代码;代码定义网络架构;网络架构定义一个确定的社会空间里可能的事物的限度。但是,没有哪一个民主过程可以定义社会空间,如果我们把代码市场也看作是一种民主过程的话。"[22]只要政府发现网络用户开始增长,它就开始强调通过技术控制网络用户的言论活动而获得实现的政府利益。[23] 因此,即便是从技术上来说,个别化的内容选择和分区制度可以有效地减轻危险言论在网络空间的传播,但它们只是一枚硬币的一面,而这枚硬币的另一面,则隐藏着事先限制的风险。

这种依靠个人化的选择机制确立起来的内容选择平台或者是依托于网络中集体选择机制建构起来的网络分区技术,都无法避免陷入事先限制的嫌疑之中。由于它们的技术运作原理是自动区分哪些内容是危险的且不应当受到宪法保护,哪些内容则是应当受到宪法全面保护的,它们也难逃基于内容的歧视性对待的事先限制的嫌疑。

---

[21] BBC ET. AL., *Striking a Balance: The Control of Children's Media Consumption* (2002), at http://www.icra.org/_en/press(2008 年 10 月 29 日最后一次访问).

[22] 参见 Lawrence Lessig, The Zones of Cyberspace, 48 *Stan. L. Rev.* 1403, 1410 (1996)。

[23] 参见 Lawrence Lessig & Paul Resnick, *Zoning Speech on the Internet: A Legal and Technical Model*, 98 Mich. L. Rev., 395, 396 (1999)。

事先限制意味着:第一,在言论还没有得到沟通或传播之前,就禁止该言论的形成与传播;[24]第二,在司法机关还没有充分决定该言论是否被有关言论自由的基本原理保护之前,就作出禁止其进行传播的决定。[25] 这种事先限制防护秩序毫无功效,但贻害无穷,举其要害,大致有如下几个方面:第一,事先限制封死言路;第二,比刑事控罪容易实施,因此有滥用之虞;第三,缺乏联邦宪法对刑事控罪的程序性保障,如控辩制、陪审团制和独立审判;第四,侵犯了公众的知情权;第五,不适当地把政府的权力扩展到个人领域。[26] 即使是自称对言论自由危害最小的内容选择平台,也无法避免陷入事先限制的情景。不宁唯是,这种伴生于网络空间的过滤技术,对言论自由所造成的伤害远比传统媒介所造成的伤害不仅大了许多,其影响所及,极难消除。

(一)事先锁闭言路

从其本质上来说,网络空间具有这样一种价值:"很多人都力争创造一种'公共领域',使不同的人不再只跟相同思想的人讨论",[27]因此,即使是那些在一些人眼里看起来不受欢迎的言论,未必就能代表所有的人都不欢迎它。技术化的管制手段,往往注意的是类型化言论管制的普遍化、一般化和规范化。即使是对言论自由伤害最小的内容选择平台,也需要通过事先确定有关言论的内容是否属于应当隔离的类型,这就不可避免地要给有关言论贴上标签,进而采取阻塞的方式来将其隔离于特定听众的视听范围之外。这就意味着个人的选择判断已经被技术的自动化选择和判断所取代,这是一种传统的事先审查制度的复辟,只是数字技术取代了人工的审阅而已。当然,它可能还是一种个人化的自我信息隔离机制,但是,这种技术本身还可以控制信息的上流,[28]换言之,它所导致的结果不单纯是个人对信息的自我隔离,因为它并不禁止(实际上也不可能!)对上游信息的过滤,借助于中立化的控制格局,它取得的效果要远远超过那些平面媒介中的事先审查手段所获得的效果。

但自动把关人的出现,会导致言论播散渠道的自我堵塞:即使是自我选择的过滤技术,也无法将所有的价值判断分解成不同情境下的过滤参考,这样一来,机械化的自我扩大标签识别范围可能就是一种自然而然地产生的附带后果。另一方面,如果是第三方的价值评判,则很少会考虑到具体言论的基本作用,而是采行一种规模经济的做法,扩大标签内容的范围、延长标签识别体系的运作时间。因此,基于不同社会规范、不同的政治社会意义而制作的言论,就可能因此而永久地消失在读者的世界中:因为软件编程取代了人的大脑识别体系,因为软件编程的自动化取代了意义建构体系的多元化运作空间。

(二)真空装置下言论自由的自我窒息

政府可以强行就某些言论内容对特殊群体的心灵伤害程度规定必须在网络端口安置自动过滤软件,它的基本功能就是对这些令人讨厌、反感的言论进行分类评级,它可以对你所要访问的网址所代表的具体言论进行测评,一旦对该网站上的有关言论的关键词进入它的屏蔽程序,则你就不能访问该网址。过滤软件是在神不知鬼不觉地过程中自动执行这样的

---

[24] Southeastern Promotions, Ltd. v. Conrad, 420 U.S. 546, 559 (1975).

[25] Pittsburgh Press Co. v. Pittsburgh Commission on Human Relations, 413 U.S. 376, 390 (1973).

[26] 参见 Vincent Blasi, Toward A Theory of Prior Restraint: The Central Linkage, *Minnesota Law Review*, vol.66, 1981, pp. 24-47。

[27] [美]卡斯·孙斯坦:《设计民主:论宪法的作用》,金朝武等译,法律出版社2006年版,第50页。

[28] 参见 Lawrence Lessig, What Things Regulate Speech: CDA 2.0 VS. Filtering, p.661。

任务的。它的目的不是阻止有关言论的制作,而是防止该言论的扩散。对制作该禁止言论的人来说,它的功能就是有效阻止他的言论的外溢效应。如果这项过滤命令是由一个过滤技术公司在网络空间的信息传递链条的上游制作的话,那么,我们可能永远都没有机会知道那个被锁定的网站上的具体内容,因为它将会在网页快照服务中也会被连根拔起。

在你无法浏览网络上的言论时,你不可能知道哪些内容被屏蔽掉,那些内容在上游某处就被一个网络内容选择平台的过滤器给滤掉了,但你是不可能知道所发生的事情的。这就带来一个问题:我们自以为知悉全部信息,实际上我们只是生活在过滤装置的下游;我们以为我们了解世间的真相,但我们所了解的只是那些愿意将这些消息发布给我们的人留存下来的。网络空间的中央化控制模式如果配置上过滤系统,,这对日益依赖网络空间来认知社会、获得必需信息并根据这些信息制作各种言论的网民们来说,就不啻一场彻头彻尾的灾难。而在现实世界里,过滤并不能将我们全部隔离于真相之外。恰如莱斯格指出的:"在现实空间里……过滤是不完全的:无论我多么忽视无家可归的现象,我在去银行的路上仍不可避免地会看到无家可归的人;无论我多么忽视不平等的现象,我去机场的路上所看到的其他社区的生活现状,还是提醒我美国是一个不平等的国家。所有这些我不愿意正视的问题还是会浮现在我面前。在现实生活中,它们要求我予以关注,不管我是否适用了过滤装置来遮蔽我的视线。"[29]

即便是基于个人对信息的精确筛选的要求而出台的信息过滤技术或网络分区技术,也同样潜藏着压制信息自由流动的风险:个人对信息的量身定制,可能也会给政府或者是基于商务延伸的网络服务供应商在和政府合谋的过程中对网络空间的言论进行事先的过滤,从而导致政府厌恶的、为服务商所心领神会的那些危险言论自动消失在网络空间。信息的窄化最终导致的结果就是整个网络言论按照政府设计的模式走向沉寂化[30]或者是碎裂化。[31] 尽管凯斯·桑斯坦说,政府未必就是言论自由的敌人,[32]但他本人从来就没有明确地说过政府就一定是言论自由的朋友。尤其是那些对由网络草根发起的自由言论的舆论造势心有余悸的国家,借口为个人享有一片祥和的网络舆论环境而巧制手法来钳制网络言论,更是政府所愿意乐见的。如果技术上的支撑可以得到这样的个人的默许的话,政府就会肆无忌惮地扩充这种令个人言论自由的网络言论过滤技术的适用空间。

(三)言论自由保护的私法化障碍

如果每一个人的网络都能自成一体,不需要来自服务商提供的基础架构的支持就可以自行接入全球网络,也许我们的担心是多余的。但网络服务商提供的各种服务是网络空间得以建构的根本,而网络服务商的基本利益又与个体化的网络用户完全不同。因此,这种过滤技术就带来另一个问题,那就是把公民所享有的基本权利交由独立的第三方来加以审查,从而导致言论自由受到损害的时候,获得的救济力度就会大大减少。过滤技术既然可以由自己来审查,由自己按照个人喜好来加以定制,那么,网络服务提供商也就找到了管制的逻辑:网络服务提供商也有着自己的经营利益、自己因为触碰具体危险观点而遭致的利益受损,因此,在自己的网络服务器的代码层配置内容选择平台,它们会按照自己利益最大化的

---

〔29〕 [美]劳伦斯·莱斯格:《代码》,李旭等译,中信出版社 2004 年版,第 220 页。

〔30〕 参见 Cass Sunstein, *Republic. com* 2.0, Princeton University Press, 2007, p. 44。

〔31〕 参见 Marshal Van Alstyne & Erik Brynjolfsson, *Electronic Communities: Global Village of Cyberbalkan*? MIT Sloan School Work Paper, 1997。

〔32〕 [美]凯斯·桑斯坦:《偏颇的宪法》,宋华琳、毕竞悦译,北京大学出版社 2005 年版,第 299 页。

方向去筛选那些它们认为适于提供给社会大众分享的信息。[33]

由网络服务商来加以定义的过滤系统对言论自由的直接损害主要体现在以下几个方面:第一,由于它的利益并不是以帮助公民实现其言论自由来实现的,因此,它对网络空间里的危险言论的定义本身就有着极大的随意性,其在具体的网络环境中的技术标准的运用常常前后不一致;第二,技术编码是一种自动运行的体系,它与言论自由所必需的社会规范支撑体系有着极大的差别,现实世界里的社会规范可以支撑的言论,可能仅仅因为系统无法识别就有可能被禁止;第三,如果网络服务商的利益和公民的言论直接发生抵触,服务商更可能基于价值偏见和主观臆断来区分言论内容,并且,它会让愿意接收这些信息的人无从了解它的具体价值偏好,因为一个匿名的网络世界往往是众口难调,每个人的选择空间的多元化和选择标准的特定化,导致信息散佚的现象层出不穷,如果人们不明就里,就很难知悉信息受阻原来只是因为网络服务商的私心所致。[34]

但问题还不仅于此,网络服务商的过滤行为往往是一种单纯的民事行为。由于这里适用的是私法的准则而非公法的准则,要证明它和政府行为扯上关系,证明责任就不可避免地落到了受害人一方。但一方面,宪法理论关于公民言论自由权利的讨论,主要的设定对象是政府的滥用权力行为。宪法并不支配私人行为。[35] 其次,由于过滤技术的复杂操作系统并非一般的网络用户所能洞悉,要想真正查找出过滤的任意诱因,在既有的法律规则体系内,普罗大众尚不具备这种能力,也缺乏相应的工具支撑。

即便是网络服务商没有包藏祸心,由于网络的快速发展,要想在所有的端口上安置内容中立的过滤器防止危险信息进入人们的视野,本身就是一件明知不可为而为之的事情。如何设置有效的过滤标准的问题就出现了:既然个体化的审查无法做到,为了最大努力地完成审查的责任,基于内容的歧视性过滤就不可避免,而且,这种歧视性的过滤还带有黑名单或关键词审查的味道。这样一来,最令我们感到恐怖的事情就会发生:我们的言论自由之所以受到限制,只是因为网络服务商提供的自动软件的审查、过滤的缘故;而我们的言论自由之所以会受到持续的损害只是因为这是一种私人行为,是一种受到私法规则保护的自我保护行为所致。早在上个世纪70年代卡尔文就预见了私法规则体系中的公法问题,[36]但他始料未及的是,这种紧张会出现在他一生珍视的言论自由上,而且使用的手段较他所处时代有过之而无不及,完全超出他的想象之外。因此,“如果政府想控制言论,那就应当让大众知道这种控制的存在。”[37]

(四)为政府滥用权力提供口实

过滤未必就全是坏事一件:在众口难调的情况下,网络过滤技术“使人们能够设计出自己高度个性化的通讯包,把有麻烦的问题以及自己不喜欢的声音过滤掉”,[38]这属于个人正当权利的范畴,倒也值得肯定。网络言论内容控制的自动化只是一种技术装置,如果是我

---

〔33〕 参见 Fernando A. Bohorquez Jr., The Price of PICS: The Privatization of Internet Censorship, 43 *N. Y. L. Sch. L. Rev.* 523, 537(1999)。

〔34〕 [美]劳伦斯·莱斯格:《代码》,李旭等译,中信出版社2004年版,第218页。

〔35〕 详见凯斯·桑斯坦:《偏颇的宪法》,第187页。

〔36〕 参见 Walter J. Blum and Harry Kalven, Jr., Public Law Perspectives on a Private Law Problem: Auto Compensation Plans, *The University of Chicago Law Review*, vol. 31, 1964, p. 641。

〔37〕 参见劳伦斯·莱斯格:《代码》,第221页。

〔38〕 [美]凯斯·孙斯坦:《设计民主:论宪法的作用》,金朝武等译,法律出版社2006年版,第42-43页。

们应要求安置在自己的网络端口的话,我们还是可以将这种装置拆除出去的,毕竟,这种自我强化的内容控制技术只是相应审查私人化的要求而已,个人毕竟还有很大的选择空间,相应地,个人就算是愿意选择过一种掩耳盗铃的生活,这还是他的自由选择。

但政府有可能成为内容选择平台的买家,[39]因此,这种以约束公民言论自由为目的的个体化装置,就会成为政府限制言论自由的总体性装置:如果政府不是要求每个人必须在自己的网络端口设置这样的过滤装置,而是通过间接管制的手段,要求处于网络上游的网络服务提供商和网络内容提供商也安装这种过滤装置的话,我们就只能生活在一个信息净化空间里了,该净化装置可能是为了防止我们直接暴露在危险言论的侵袭下、直接成为危险言论的受害者。但是,过滤技术更多的是在考虑听众听不到有关危险信息,试图将听众和危险信息直接隔离开来,从而对这种危险信息形成真空隔离状态。另外,这种做法本身也是对社会大众的不负责任,没有理由认为国家有能力为社会大众提供最全面的保护,因此,如果社会大众因为信息过滤的缘故而不能了解哪怕是一丝的危险信号,从而导致他们在危险发生的时候没有任何防护能力的话,这无论从任何层面上都可以说是国家的一种失职。从技术和政治结合的层面上来看,允许政府立法设置一种只是局限于过滤某种特定类型的言论,这种做法潜藏着一个雷区,那就是政府并不满足于只是强化一种限制最小的技术,它会更倾向于将过滤软件所过滤的内容作最大程度的扩张:"如果政府在过滤 X 类型的言论上享有正当的利益,但不具有过滤 Y 类型和 Z 类型的正当利益,那么,这就存在两种架构,其中的一种可以过滤 X、Y、Z 型言论,而另一种则将只能过滤 X 类型的言论,于是,国会就会依据宪法权力来强化第二种类型的过滤技术,而第一种则不会受到重视。"[40]

立基于此推论,政府会通过要求网络服务商安装类似的过滤装置来确保网络上不至于出现政府所憎恶的危险言论,而网络服务商则会因为自身的商务机制考虑,乐于接受这样的要求。两者之间的短路式合谋,会使得政府的权力不仅大大扩张,而且政府的管制更有了口实可言:它不仅仅要确保公民的言论自由的实现,而且还要确保网络服务商的合法财产权益不因危险言论的扩散而遭受损失。危险言论是针对公权力的,私人财产不应当为公权力买单,也不得为公民的宪法权利之行使买单。

(五)危及民主参与

公众观点持续流动和相互影响对民主参与意义重大。网络空间里的信息传递,尽管首先是以一种一对多的格局呈现的,但它一经生成,就不再受到初始传播者的控制。这些言论有可能在其他网站上,以多种形式来加以阅读和评论,在不同的读者群或者是在言说者和听众之间形成多程式的信息反馈。[41] 因此,民主参与如果要实现最起码的合法性,建立在"知情的、非强制性的一致同意"[42]或"合理地可以接受的"[43]基础上的决策机制就显得极为重要。

民主参与中的审议,包括如下三个层面的内容:"(1)参与审议的过程是由平等性和对称性的规范支配的;所有人都拥有展开言语行动、质疑和辩论的机会;(2)所有人都有权质

---

〔39〕 Lawrence Lessig, Code 2.0, Basic Books, 2006, p.257.

〔40〕 参见 Lawrence Lessig, What Things Regulate Speech: CDA 2.0 VS. Filtering, p. 665。

〔41〕 参见 A. Michael Froomkin, Habermas@ Discourse. Net: Toward A Critical Theory of Cyberspace, *Harvard. Law Review*, vol. 116, 2003, p. 860。

〔42〕 [美]詹姆斯·博曼:《公共协商:多元主义、复杂性与民主》,黄相怀译,中央编译出版社 2006 年版,第 23 页。

〔43〕 [德]于尔根·哈贝马斯:《后民族结构》,曹卫东译,上海人民出版社 2002 年版,第 260 页。

疑预先确定好的对话主题;(3)所有人都有权发起对对话程序规则以及被采用的或实施的方式的反思性论辩。"[44]要保证审议质量的提高,从而实质性地提高民主参与的真实程度和决策机制的完善,就不仅仅需要和相同意见进行交流,相反,只有不同意见的充分交流才会允许真正的辩论和思想的交换。这种充分的交流才会使得参与者有机会"从更大的菜单中作出选择",从而导致一种社会意识而非个人私利。[45]

过滤技术使喜欢或厌恶某种言论的简单心理逻辑获得了坚实的技术支持。它固然能为个人自由空间带来好处,但从民主参与的理性化程度来说,则是弊害多多。异质性的意见的充分交流和沟通,是促进民主参与的基本前提。在一个媒介沟通机制不足的世界,人们放弃理性参与可能是一种自我保存的需要,但是,在技术已经提供了充分保护的网络空间,如果意欲通过个人对网络外部世界的多元意见的隔离而实现个人的自治,只会一方面强化人的过分的孤独隔离感,另一方面也会导致人们如果一有机会将意见表达的时候,则会采取一种歇斯底里的爆发方式,而不是采取理性的磋商机制。

如果国家借助这种过滤技术将即时性的公民发言进行内容分级和过滤,网络空间里充斥的言论可能就是一些谎言。国家如果借口网络空间存在着危险言论,从而不仅仅对制作言论的人进行内容上隔离,并且采取过滤技术防止公民接触这样的信息,民主参与和审议的质量就会降低:如果指望民众在政府思想的呵护中长成独当一面的自主公民的话,则无异于与虎谋皮;不经一事的公民,想要直接面对问题的本质,谈之何易?

五 结 语

"完全不理睬当下,这从来不是一个可行的社会秩序原则。"[46]制度化的技术方案的设置,可以在一定程度上解决危险言论在传播过程中所造成的危害,但它必须以对宪法上记载的权利的有效运行给予切实保障为皈依。毕竟"技术始终只是工具和武器,正是由于它服务于任何一个人,它才不是中立的。从技术之内在性质得不出任何一种人性的和精神性的决断。……技术本身在文化上是盲目的。因此,从纯然的技术之外别无一切得不出任何一个以往精神生活的中心领域得出的结论:不论是文化进步的要领,或精神领袖的抑或某一政治体制的类型。"[47]技术特性取决于人们对技术背后的规范体系的理解和运用,技术在扩大我们的自由选择范围的时候,可能也在扼杀我们的自由选择。[48] 这既源于技术不可控的特点,也源于权力机制对技术的塑造。

但网络言论也有不受发言者的意志控制的特性。人们的行动世界固然外在于网络虚拟世界,但他们的行动指南和观念构图却有可能深受网络世界里的言论模式的影响。在这个

---

〔44〕 参见 S. Benhabib, ed., *Democracy and Difference: Contesting the Boundaries of the Political*, Princeton University Press, 1996, p. 70。

〔45〕 参见 Michael MacKuen, Speaking of Politics: Individual Conversational Choice, Public Opinion, and the Prospects for Deliberative Democracy, in John A. Ferejohn & James H. Kuklinski eds., *Information and Democratic Process*, University of Illinois Press, 1990, p. 60。

〔46〕 [美]理查德·波斯纳:《超越法律》,苏力译,中国政法大学出版社 2001 年版,第 582 页。

〔47〕 [德]卡尔·施米特:《论断与概念:在与魏玛、日内瓦、凡尔赛的斗争中(1923 - 1939)》,朱雁冰译,上海人民出版社 2006 年版,第 125 - 126 页。

〔48〕 "完美的控制技术并不必然蕴含着完美的实现正义的技术"。Lawrence Lessig, Zones of Cyberspace, *Stanford Law Review*, vol. 48, 1996, p. 1411.

意义上,网络并非完全独立于现实世界,并非无拘无束的自由放任世界。“网络空间的东西并不完全是非竞争性的。网络容量就是一个制约,带宽也不是无限制的,但这些小小的缺陷并不能证明从广泛自由走向完全控制的合理性。网络空间存在着合作及资源短缺的问题,但解决这些问题的必要办法并不是采用控制体系或更好的独占手段。网络空间的繁荣在很大程度上得益于公共资源这一事实使我们不禁要问:我们是否应当向这一空间注入更多的自由,而不是来自现实空间的约束?”[49]

过滤技术并没有向网络空间注入适量的自由或者是给予人们更多的自由保证。即使是完全私人化的过滤装置,也不免将其所不需要的信息隔离掉的危险。但个人化的过滤是一种规模不经济,它缺乏必需的支撑架构,即使个人愿意掩耳盗铃,但这种掩耳盗铃未必就是出于他的自愿选择。政府也许可以适度地推行这样一种软件系统,但又怎能保证不至于出现南辕北辙的后果呢?

如果我们相信言论自由并非一个完全理论化的协议,[50]而是一种随时随地都在尝试自我改进、自我完善的理论草案的话,自由言论在制造者和听众之间所形成的信息循环就是其题中应有之义;如果我们能够认识到网络空间的言论自由乃是一种双向的流动:没有单纯的听众,尽管并非每个人都是发言者,但是,他们的确可以及时回应发言者,以发言者的姿态,我们就会发现,如果将一些言论隔离出我们的世界,我们将会被装到一个永远只有自己声音的“回音室”。信息过滤的对象是听众,这样的做法无疑是在宣布:国家比个人拥有更为整全的判断体系来判断什么信息对他是最安全的。因此,当我们在推行“绿坝”等据说有利于净化网络环境的过滤装置的时候,当我们对其技术架构隐晦不言的时候,有否想过其可能引发的问题,比起其所要解决的问题,则是多出了不知几许?

[**Abstract**] Based on consideration of public order of society, the internet filtering technique has its rational dimension. However, the internet filtering device set on the internet, though claiming that it has minimum negative effect on the free flow of information, respects free choices of internet users to the greatest extent, and gives consideration to both needs of social stability and platform of internet content selection (PICS) of freedom of speech, still can not avoid arbitrary judgment and infringement upon citizens' right to free speech. Whether the internet filtering technique has legitimacy remains a controversial topic. Although in different contexts, yet the controversies over PICS can be used for reference when China takes steps to control internet speech.

(责任编辑:支振锋)

〔49〕 [美]劳伦斯·莱斯格:《思想的未来》,李旭译,中信出版社 2004 年版,第 121 - 122 页。

〔50〕 详见孙斯坦:《设计民主:论宪法的作用》,第 2 章。