

# 法社会学视角下的算法规避及其规制

邱遥堃

**摘 要** 算法权力虽然强大,但仍可能受到规避,与其治理对象之间持续存在动态博弈关系。规避算法的基本策略,包括避免成为治理对象、调整满足治理要求与混淆迷惑治理主体。这一现象的成因在于技术的刚性有限,且不匹配处理边际性治理问题所需要的利益衡量,无法将所有“标准”转化为“规则”,解决不了多元规范秩序背后的根本性社会矛盾。然而,大多数人并非算法规避者,算法权力总体有效;少数的算法规避也有利于限制算法权力,平衡治理与自由,促进社会稳定与治理的综合效率。只需坚持包容审慎的规制原则,在通过技术加强规制简单的算法规避之同时,对复杂的算法规避保持容忍,同时发挥社会规范的规制作用。

**关键词** 算法规避 法律规避 互联网治理 多元规范 法律与科技

作者邱遥堃,法学博士,首都经济贸易大学法学院讲师。

DOI:10.16094/j.cnki.1005-0221.2023.03.002

## 一、问题的提出

由于大数据杀熟<sup>①</sup>、新就业保障<sup>②</sup>、机器人法官<sup>③</sup>等现实问题层出不穷,当前对算法的法学研究主要着眼于权力限制的一面:或者要求算法的运行过程可监督、可控制,以算法透明与算法解释反对算法黑箱与算法独裁;<sup>④</sup>或者要求算法的运行结果合法且公正,杜绝算法失灵、算法歧视与算法损害;<sup>⑤</sup>甚至未雨绸缪地进一步思考(强)人工智能的法律地位,主张或有必要将其作为法律主体

\* 本文系司法部2021年度法治建设与法学理论研究青年课题“互联网内容治理中的平台责任研究”(21SFB3004)的研究成果。

① 参见史洪举“以司法裁判向大数据杀熟说不”,载《人民法院报》2021年7月17日;谢尧雯“网络平台差别化定价的规制路径选择——以数字信任维系为核心”,《行政法学研究》2021年第5期,第20-21页。

② 参见孙萍“‘算法逻辑’下的数字劳动:一项对平台经济下外卖送餐员的研究”,《思想战线》2019年第6期,第52-56页;谢增毅“平台用工劳动权益保护的立法进阶”,《中外法学》2022年第1期,第108、119页。

③ 参见 Melissa Hamilton, “The Biased Algorithm: Evidence of Disparate Impact on Hispanics”, *American Criminal Law Review*, Vol. 56, No. 4 (2019), pp. 1554-1558; 刘艳红“人工智能技术在智慧法院建设中实践运用与前景展望”,《比较法研究》2022年第1期,第1-3页。

④ 参见张凌寒“商业自动化决策的算法解释权研究”,《法律科学》2018年第3期,第72-73页; Robert Brauneis & Ellen P. Goodman, “Algorithmic Transparency for the Smart City”, *Yale Journal of Law and Technology*, Vol. 20, No. 1 (2018), pp. 163-175; Cary Coglianese & David Lehr, “Transparency and Algorithmic Governance”, *Administrative Law Review*, Vol. 71, No. 1 (2019), pp. 29-55; 丁晓东“基于信任的自动化决策:算法解释权的原理反思与制度重构”,《中国法学》2022年第1期,第115-117页。

⑤ 参见 Aziz Z. Huq, “Racial Equity in Algorithmic Criminal Justice”, *Duke Law Journal*, Vol. 68, No. 6 (2019), pp. 1111-1133; 郑智航、徐昭曦“大数据时代算法歧视的法律规制与司法审查——以美国法律实践为例”,《比较法研究》2019年第4期,第113-118页; 曹博“算法歧视的类型界分与规制范式重构”,《现代法学》2021年第4期,第123-126页。

而非单纯客体进行规制,由此衍生出人工智能作品的著作权保护、自动驾驶汽车的责任分配等问题。<sup>⑥</sup> 这些研究最终在政策上凝结为《个人信息保护法》第24条<sup>⑦</sup>、《关于落实网络餐饮平台责任切实维护外卖送餐员权益的指导意见》第2条<sup>⑧</sup>和《互联网信息服务算法推荐管理规定》《关于加强互联网信息服务算法综合治理的指导意见》等法律法规。

上述限权研究的潜在假定皆为:算法权力可能过于强大,使人无力抵抗,对个人权利与改革创新具有负面影响。然而经验观察显示,在特定条件下,算法权力也可能相对弱势,被治理对象发现规律后加以规避,致使算法权力无法实现既定目标。该规避可能借助同样强大的反向技术手段,但大多数情况下往往没有什么技术含量,仅仅是“上有政策,下有对策”在算法时代的新变体。于是相比之下,人工智能治理中的“智能”二字饱含反讽意味。

因此,研究算法权力具有局限性,不能仅仅关注权力失控的可能性,还必须注意其效用的边际;不能只进行静态描述,将受算法权力治理者视为有待拯救的受害者,还应注意其能动性,动态地分析权力关系双方你来我往的博弈过程。事实上,这正是一种法社会学的研究视角:关注“书本上的法”(law in books)与“行动中的法”(law in action)之间的差距,在多元规范的语境中审视各方以法律这一特殊规范为背景的互动,甚至不急于提出对现状的改进方案,因为知晓对当前平衡的打破会带来难以预期的后果,而且通常认为的有害行为往往具有意想不到的社会功能。<sup>⑨</sup> 该视角有助于全面认识相关法律现象的深层原因与实际影响,进而使据此提出的解决方案更审慎、有效,避免公共权力的行使不当或过分扩张,反而新增规制风险,使权力与权利关系落入更糟糕的平衡,新秩序的好处未体现,而破坏旧秩序的弊病却已先发生了。<sup>⑩</sup> 因此,本文将从法社会学的视角切入,通过剖析算法规避的主要策略、内在成因、实际后果与功能,揭示算法权力的运行边界,进而审慎地提出对算法规避的规制方案,借此反思算法规制乃至法律对技术的一般规制。

值得预先说明的是,算法权力的主体包括公、私两方面,算法所支持的规范亦包括法律与非法律的社会规范,故而存在“算法的法律”与“法律的法律”之区分。<sup>⑪</sup> 但对本文所论述的主题而言,上述区分意义有限,不论被规避算法所支持的主体及其规范性质如何,其类型、成因与后果皆类似,何况许多社会规范是法律治理的承包者,二者目标一致。<sup>⑫</sup> 此外,由于互联网平台作为运用

⑥ 参见袁曾“人工智能有限法律人格审视”,《东方法学》2017年第5期,第52-53页;司晓、曹建峰“论人工智能的民事责任:以自动驾驶汽车和智能机器人为切入点”,《法律科学》2017年第5期,第172-173页;梁志文“论人工智能创造物的法律保护”,《法律科学》2017年第5期,第161-163页;郭少飞“‘电子人’法律主体论”,《东方法学》2018年第3期,第41-49页。

⑦ 我国自2021年11月1日起施行的《个人信息保护法》在第24条规定“个人信息处理者利用个人信息进行自动化决策,应当保证决策的透明度和结果公平、公正,不得对个人在交易价格等交易条件上实行不合理的差别待遇。通过自动化决策方式向个人进行信息推送、商业营销,应当同时提供不针对其个人特征的选项,或者向个人提供便捷的拒绝方式。通过自动化决策方式作出对个人权益有重大影响的决定,个人有权要求个人信息处理者予以说明,并有权拒绝个人信息处理者仅通过自动化决策的方式作出决定。”

⑧ 市场监管总局等七部门于2021年7月联合印发的《关于落实网络餐饮平台责任 切实维护外卖送餐员权益的指导意见》(国市监网监发〔2021〕38号)在第2条规定:“……优化算法规则,不得将‘最严算法’作为考核要求,要通过‘算法取中’等方式,合理确定订单数量、在线率等考核要素,适当放宽配送时限。”

⑨ 参见Lawrence M. Friedman, *Impact: How Law Affects Behavior*, Cambridge: Harvard University Press, 2016, pp. 1-3; 刘思达:“美国‘法律与社会运动’的兴起与批判——兼议中国社科法学的未来走向”,《交大法学》2016年第1期,第20-26页。

⑩ 此处化用自费孝通先生的经典表述“法治秩序的好处未得,而破坏礼治秩序的弊病却已先发生了。”参见费孝通《乡土中国 生育制度》,北京大学出版社1998年版,第58页。

⑪ 参见郑戈“算法的法律与法律的法律”,《中国法律评论》2018年第2期,第67-69页。

⑫ 参见Seth F. Kreimer, “Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link”, *University of Pennsylvania Law Review*, Vol. 155, No. 1 (2016), pp. 16-27; 孔祥稳“网络平台信息内容规制结构的公法反思”,《环球法律评论》2020年第2期,第135-140页。

算法最重要的主体，具有私人治理公共化的特点，模糊了公与私之间的边界，更无必要进行界分。<sup>⑬</sup>故而，本文的论述将杂糅公、私主体的算法应用，仅在少数情况下进行必要的区分。也正因为如此，本文所谈算法权力的“治理”，主要意指通过算法进行治理或算法权力作为治理手段，区别于对算法权力进行限制的治理，但本文亦会论及对算法权力与算法规避进行限制的问题，主要以“规制”为关键词，故而在此特别对用语进行说明，以免读者产生混淆。

## 二、算法规避的主要策略

算法规避 (evasion of algorithm) 是指本应受算法权力治理者通过实施特定行为，最终免受算法权力的治理。<sup>⑭</sup>例如，网站根据搜索引擎排序算法的原理，嵌套符合条件的内含隐藏内容的页面，从而达到即使与搜索内容并不相关，亦可被检索到并排名靠前的目的。又如，用户通过总结社交网络内容审查算法的规律，以谐音、洗稿、添加空白内容等方式，发布本应被屏蔽的具有政治敏感性、抄袭盗版或低俗色情内容。再如，“网络水军”按照与评分平台之反刷分算法的博弈经验，通过账号培育、手工刷分等方式，继续堆砌虚假好评或差评。该行为的核心，是受算法权力治理者在违反算法要求的同时避开算法权力的后果。要实现这一目的，在经验上主要有以下三种基本策略。

### (一) 避免成为治理对象

算法规避的首要手段是避免成为算法权力的治理对象，使自己落在算法适用范围之外，就像在传统线下场景中的违法犯罪行为人为避开监控摄像、警察巡逻及逃脱公权力的可能执法那样。事实上，互联网兴起之初的“不可规制性”问题，便是这一策略的极端典型例证。早期的互联网架构保障了网上行为人的充分匿名，其身份、位置与行为皆不可获知，故无法对其进行规制，更加不用说采取算法手段。<sup>⑮</sup>然而，随着网络架构从开放走向封闭，“在互联网上没有人知道你是一条狗”的故事已成为历史，网络空间已具备充分的可规制性，并且已深度向线下延伸，模糊了传统的线上与线下之别。<sup>⑯</sup>当下，除非选择极少数完全不涉及互联网的行为乃至生活方式（例如使用现金、传统的书信），否则，要想彻底避免成为算法权力的治理对象，极为困难。<sup>⑰</sup>

但也并非完全不再有可能，平台选择即为避免成为算法权力治理对象的新型有效手段之一。由于不同平台的发展阶段有别，故而在网络空间趋于封闭的整体环境下，仍然存在新兴平台“非法兴起”的可能性，尤其是当破坏性创新与跨界竞争发生时。<sup>⑱</sup>即使不同平台的法治水平相当，平台间

<sup>⑬</sup> 参见 Jack M. Balkin, “Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation”, *U. C. Davis Law Review*, Vol. 51, No. 3 (2018), pp. 1173-1182; 刘权 “网络平台的公共性及其实现——以电商平台的法律规制为视角”, 《法学研究》2020 年第 2 期, 第 43-47 页。

<sup>⑭</sup> 参见 Jane Bambauer & Tal Zarsky, “The Algorithm Game”, *Notre Dame Law Review*, Vol. 94, No. 1 (2018), p. 10。该文对算法博弈 (algorithm game) 的讨论，启发了本文对算法规避之主要策略的论述，但本文结合互联网特别是中国互联网发展的现实，更深入地探讨了各类算法规避的内涵、机理与关系，并通过“算法规避”这一关键词，接续了有关法律规避、法律影响的既有法理学研究，进一步揭示了其成因与后果，对该文的研究进行了理论推进。

<sup>⑮</sup> 参见 [美] 劳伦斯·莱斯格 《代码 2.0: 网络空间中的法律》，李旭、沈伟伟译，清华大学出版社 2009 年版，第 35-42 页。

<sup>⑯</sup> 参见胡凌 “‘分享经济’的法律规制”，《文化纵横》2015 年第 4 期，第 113-114 页；Ignas Kalpokas, *Algorithmic Governance: Politics and Law in the Post-Human Era*, Switzerland: Palgrave Pivot, 2019, pp. 39-40。

<sup>⑰</sup> See Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*, New York & London: W. W. Norton & Company, 2015, p. 146。

<sup>⑱</sup> See Thomas Eisenmann, Geoffrey Parker, and Marshall Van Alstyne, “Platform Envelopment”, *Strategic Management Journal*, Vol. 32, No. 12 (2011), p. 1270。

的数据与信息壁垒,仍然能阻断跨平台的算法权力的治理行为,帮助跨平台的违法违规行。<sup>①⑨</sup>因此,平台间的边界犹如线上与线下的边界,将统一的网络空间再度分割,使统一的算法适用不再成立。算法规避者或可利用新兴平台的制度不健全,继续发布低俗抄袭内容、贩卖假冒伪劣商品;或可利用不同平台间的信息不对称与数据壁垒,从事跨平台抄袭、“山寨”行为,一如上述那种选择抛弃互联网的行为与生活方式。

类似地,不同国家法制环境的差异,同样可以奠定算法规避者空间选择的基础。某些国家认定的低俗色情内容,在另一些国家可能被认为是性自由与性解放的象征;<sup>②①</sup>某些国家认定的抄袭内容、“山寨”产品,在另一些国家可能被认为是合理使用与借鉴。<sup>②②</sup>彼此以他国规则为不合理,对行为人而言就产生了狡兔三窟的可能性。事实上,国际私法上的法律规避,便是利用不同辖区法律之差异来实施不为当地法律所认可之行为,其典型例子如同性伴侣选择同性婚姻合法化的地区结婚。<sup>②③</sup>由于各个国家网络主权建设的不断推进,区域化互联网或国域网正在形成当中。<sup>②④</sup>这意味着线上与线下的融合不仅仅是线上治理手段与治理逻辑对线下区域的整合,反之亦然,意欲规避算法者得到了又一算法权力的裂缝。

综上,避免成为治理对象的本质是寻求算法权力的“法外之地”,或是其他平台,或是其他国家,从而规避当前平台与当前国家的算法权力。然而,随着国内法制与国际法制的日趋统一,这一规避手段正如彻底断绝互联网的行为与生活方式那样越发难以持续。因此,如何在既有算法权力的治理体系下继续规避算法,就需要或者调整行为、或者迷惑主体。

## (二) 调整满足治理要求

算法规避的第二类手段是调整行为以满足治理要求,使自己虽然在算法适用范围之内,但仍可以隐秘实施算法所欲禁止的行为,就像传统经济法上的虚假纳税申报及其他虚假信息披露行为,或者传统民法上的变相借贷、变相融资、与实质内容不符的名义合同关系等“以合法形式掩盖非法目的”的行为。<sup>②⑤</sup>互联网的兴起,不仅为此类行为调整提供了辅助工具,例如自动化的洗稿软件及其他文本转换软件、媒介转换软件,<sup>②⑥</sup>而且更使得此类技术及其影响的规避行为广泛普及,大数据支配下的用户生产丰富多样且更新速度极快,<sup>②⑦</sup>大大提高了算法权力的治理难度与滞后性,正所谓“道高一尺,魔高一丈”。更具体而言,调整行为以满足治理要求,可被进一步细分为如下两种子类型。

其一是直接改变算法输入数据。用户在网上活动需要注册账号,提供基本的个人信息,例如性别、年龄、位置等,为算法基于前述信息的治理提供基础;如果谎报此类数据,那么就能实现对禁止未成年人浏览特定内容、向特定性别或地域用户进行推荐等算法之规避。诚然,网络实名制及其

<sup>①⑨</sup> 参见杨东“论反垄断法的重构:应对数字经济的挑战”,《中国法学》2020年第3期,第209-210页。

<sup>②①</sup> 参见桑本谦“网络色情、技术中立与国家竞争力——快播案背后的政治经济学”,《法学》2017年第1期,第81-82页。

<sup>②②</sup> 参见冯象、李一达“知识产权的终结——‘中国模式’之外的挑战”,《文化纵横》2012年第3期,第53-56页。

<sup>②③</sup> 参见徐崇利“法律规避制度可否缺位于中国冲突法?——从与强制性规则适用制度之关系的角度分析”,《清华法学》2011年第6期,第123页;许光耀“略论国际私法上的法律规避制度”,《法学评论》2012年第6期,第79页。

<sup>②④</sup> 参见刘晗“域名系统、网络主权与互联网治理:历史反思及其当代启示”,《中外法学》2016年第2期,第522-525页;张新宝、许可“网络空间主权的治理模式及其制度构建”,《中国社会科学》2016年第8期,第141-145页。

<sup>②⑤</sup> 参见朱广新“论‘以合法形式掩盖非法目的’的法律行为”,《比较法研究》2016年第4期,第175-176页;冉克平:“‘恶意串通’与‘合法形式掩盖非法目的’在民法典总则中的构造——兼评《民法总则》之规定”,《现代法学》2017年第4期,第74-76页。

<sup>②⑥</sup> 参见周勇“智能洗稿法律规制研究”,《当代传播》2019年第4期,第77页。

<sup>②⑦</sup> See Yochai Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom*, New Haven and London: Yale University Press, 2006, pp. 59-63.

罚则的存在,<sup>②7</sup> 以及通过身份证号、实名手机号与邮箱或大数据分析甄别用户申报信息之真伪,已在最大限度上减少了虚假数据申报量,但依旧存在前述制度与技术未能覆盖的网络空间,例如仅仅在浏览前要求确认年龄或者提供退出信息收集选项的网站。而且,保护个人信息的法律与社会规范,要求限制前述治理者的信息收集与分析。<sup>②8</sup> 因此,虚假信息披露仍有可能,可以通过直接改变输入算法的数据,从而对其进行规避。

其二是通过改变行为来间接改变算法输入数据。此类算法规避最为常见,例如通过谐音、拼音、首字母、拆字、联想词等方式规避算法对敏感词汇的筛查;通过词语替换、语序变更、意译等方式抄袭文字,或者通过翻转、旋转、裁剪、噪声、变速和帧率变换等手段改变音视频,又或者通过将不同类型的素材或不同来源的作品进行拼接等方式,规避算法对版权侵权的审查;通过对其他作品尤其是小众作品的随意打分来培育账号,通过手动刷分、以量取胜来控制评论,规避算法对虚假评分的审查。以上复杂的规避,皆属于广义的“以合法形式掩盖非法目的”之行为,以算法的技术水平为前提,又或者正因为算法对简单违法违规行之禁绝而产生,与之持续博弈、互为因果。<sup>②9</sup> 因此,提高技术水平,只会导致算法规避的继续升级,甚至由于触及个人自由之边界而难以为继,对算法规避行为的限制有限。

上述调整行为以满足治理要求的本质,是“以合法形式掩盖非法目的”,或直接改变输入数据,或通过改变行为来间接改变输入数据。然而,并非所有数据与行为都易于改变,又或者如若改变则将使规避者的本来目的部分地或完全落空,因此需要其他手段在不直接改变核心数据与行为的同时依然规避算法权力,此即混淆迷惑治理主体这一手段。

### (三) 混淆迷惑治理主体

算法规避的最终手段是混淆迷惑治理主体,使自己虽然直接从事了算法所禁止的行为,但仍可以蒙混过关,落在治理主体注意范围之外。在传播学上,这一策略被称为信息洪流(information flooding),意为通过大量无关、低质或虚假的信息的堆砌,使受影响者丧失对关键问题的注意力。<sup>③0</sup> 互联网是实施信息洪流策略的完美工具。相较于过去的电报、电话、广播、电视等媒介,互联网传播信息的数量更大、形式更多、速度更快、范围更广,而且传播方式是双向、多向而非单向、单一的,一对一、一对多与多对多传播混合,大大降低了信息传播的成本,增进了不同主体间的信息流通与共享,产生了所谓信息爆炸。尽管算法权力有利于应对海量内容的治理负担,但在具体个案中,仍然可能被信息洪流欺骗或蒙蔽。

例如,意欲传播淫秽色情内容者,若按法律或平台规则要求进行打码、删截或者改为软色情创作,则将违背其传播本意,因此调整空间有限。但他仍可以将少量的色情内容与大量无害或空白的内容组合发布,又或者对色情内容进行明暗、色调处理,规避算法对色情内容的审查。又如,意欲

<sup>②7</sup> 我国自 2017 年 6 月 1 日起施行的《网络安全法》的第 24 条规定“网络运营者为用户办理网络接入、域名注册服务,办理固定电话、移动电话等入网手续,或者为用户提供信息发布、即时通讯等服务,在与用户签订协议或者确认提供服务时,应当要求用户提供真实身份信息。用户不提供真实身份信息的,网络运营者不得为其提供相关服务。”该法第 61 条规定“网络运营者违反本法第二十四条第一款规定,未要求用户提供真实身份信息,或者对不提供真实身份信息的用户提供相关服务的,由有关主管部门责令改正;拒不改正或者情节严重的,处五万元以上五十万元以下罚款,并可以由有关主管部门责令暂停相关业务、停业整顿、关闭网站、吊销相关业务许可证或者吊销营业执照,对直接负责的主管人员和其他直接责任人员处一万元以上十万元以下罚款。”

<sup>②8</sup> 我国《个人信息保护法》第 16 条规定“个人信息处理者不得以个人不同意处理其个人信息或者撤回同意为由,拒绝提供产品或者服务;处理个人信息属于提供产品或者服务所必需的除外。”

<sup>②9</sup> 这类似于法律规避中法律作为规避的阴影而存在,参见 Robert H. Mnookin & Lewis Kornhauser, “Bargaining in the Shadow of the Law: The Case of Divorce”, *Yale Law Journal*, Vol. 88, No. 5 (1979), p. 950.

<sup>③0</sup> See Finn Brunton & Helen Nissenbaum, *Obfuscation: A User's Guide for Privacy and Protest*, Cambridge: The MIT Press, 2015, pp. 38-39.

传播盗版抄袭内容者，除上述内容调整策略外，还可采用将盗版抄袭内容与无关的内容或自己原创的内容组合发布，降低重复率，规避算法对版权侵权的审查。前述行为虽然不涉及复杂的行为调整，对人工审查而言事实上相对简单低级，但由于算法权力的刚性与人工审查对海量内容的应对乏力，故而依然存在广泛的应用空间，并且同样存在博弈升级现象，体现出“人工智能”治理最为不“智能”的一面。

综上，混淆迷惑治理主体的本质是信息洪流，通过将违法信息与合法信息加以组合来规避算法审查，相对简单低级，体现出人工智能权力缺乏弹性。行文至此，算法规避的主要策略已总结完毕，更进一步的问题便是：为什么算法规避者可以避免成为治理对象、调整行为以满足治理要求、混淆迷惑治理主体？下一节将从技术、规范与社会这三方面入手进行解析。

### 三、算法规避的内在成因

探究算法规避的内在成因，通常可以从主客观两方面入手：主观方面，规避者的守规意识薄弱、不积极遵守算法要求；客观方面，算法之治存在漏洞，故而被规避者利用。但此种老生常谈流于表面，未揭示所谓“算法漏洞”究竟是什么，以及所谓“守规意识薄弱”到底原因何在。更重要的是，如果有必要假定被治理者为“坏人”，<sup>③</sup>那么就不应过分关注甚至改造其主观意识，而应当在客观的制度建设与技术提高方面更多着力，借此改变其行为激励，使之与算法要求激励相容。因此，本部分将从技术、规范与社会三方面入手，解析算法规避的形成原因。

#### （一）技术刚性的相对性与绝对性

算法规避在技术上的表面原因是，正如代码与架构的刚性仅是从相对意义而非绝对意义来讲，算法权力的严格性亦仅是相对于一般人类智能而言，故而更聪明的少数人可以调整行为来满足治理要求，或者混淆迷惑治理主体。例如，防火墙或许能够阻碍许多人在网上漫游，但少数人仍能凭借虚拟专用网络等技术手段予以跨越；<sup>②</sup>缓冲带也许能够妨碍许多人在现实中超速行车，但少数人仍能凭借过人的车技不予理会。<sup>③</sup>算法权力亦然。内容审查、版权执法、犯罪预测等算法对大多数人与在大多数情况下有效，并不意味着其无往不利；对别有用心且又技高一筹的“坏人”而言，它们不过是提高了其规避或者违反算法的成本，迫使其付出更多的努力、采取其他方式来达成目的，例如更为复杂的内容生产（谐音、洗稿、拼图等），有针对性地加以规划的行动轨迹、个人形象设计，等等。

诚然，任何技术及其执行的规范都不可能做到完全实施，能够对大多数人的违法违规行为进行阻却即已实现其治理目标，并有利于避免被治理者的反弹，<sup>④</sup>但完美执法的无法实现也更深刻地揭示出：算法之治及更一般性技术治理的刚性，无法从事处理边际性治理问题时所需的利益衡量。<sup>⑤</sup>例如，将原文表达重新组装，固然能达到不被著作权法直接禁止的地步，但根本上还是抄袭了原文

<sup>①</sup> See Oliver Wendell Holmes, Jr., “Path of the Law”, *Harvard Law Review*, Vol. 10, No. 8 (1896–1897), p. 459.

<sup>②</sup> See Margaret E. Roberts, *Censored Distraction and Diversion Inside China’s Great Firewall*, Princeton: Princeton University Press, 2018, pp. 2–4.

<sup>③</sup> 参见注①，第369页。

<sup>④</sup> See Thomas M. Keck, “Beyond Backlash: Assessing the Impact of Judicial Decisions on LGBT Rights”, *Law & Society Review*, Vol. 43, No. 1 (2009), pp. 152–155.

<sup>⑤</sup> 参见梁上上“利益衡量的界碑”，《政法论坛》2006年第5期，第67页。

思想的洗稿行为，这究竟属于著作权的合理使用，还是对著作权的侵犯？用户为培育账号而发表的评论，究竟是言论自由的正当行使，还是违反平台规则、有意规避算法的不正当行为？这些问题对人类治理者而言都相当棘手，因为其触及治理权力与个人自由的边界，难以划定非黑即白的行为疆域，而技术的刚性所追求的正是这样的划界，故而产生了极大的不兼容。这正是智能化的算法反而容易被智能化水平不高的手段所简单规避的原因所在：刚性的边界划定太容易留下可以施展操作的灰色地带，而裁量的存在反而有威慑的效力遗留。

技术刚性与利益衡量的不兼容，也阻断了通过技术的进步来克服这一局限的可能。通过机器学习使利益衡量类型化、使其中的规则明晰化，是算法权力有益的改进方向，也是其对权力滥用之抑制的承诺所在。<sup>⑤</sup>但这并不足以取消边际性治理问题的存在及其对利益衡量的需求：解决旧的问题，明确现有规范，多阻却一部分人的规避与违法，还会产生新的边际，需要新的衡量，总有人会想方设法规避新的规范。前述洗稿与养号问题的产生，正是因为算法权力对较为初级的规避与违法行为进行了有效治理，而对这些问题的有效回应，也必不会是算法权力的发展终点。这正是算法之治及其规避的上述博弈所揭示的动态过程，更完美的执法引发更完美的规避。

## （二）规则规范与标准规范的鸿沟

从法学理论对“规则”与“标准”之区分的角度观之，上述不兼容的实质，是试图通过算法将标准完全转化为规则的不可能。规则具有事前明确的一项或多项具体事实及其未能实现时的规范后果；标准在事前则仅有诸多与立法目标相关、需要综合考虑的事实因素，其规范后果之确定更为复杂。由此可见，规则在事前的清晰明确，虽然充分体现了法治的应有之义，但它不仅有利于行为主体守法用法，而且也有利于规避与违法；对满足形式要件而违背实质目的、以合法形式掩盖非法目的之行为，受限于文义的规则是无能为力的，除非继续通过“解释”将规则在事实上又转化为标准。标准的模糊性与事后确定性，虽然不能提供精确的行为指引，但能够通过自由裁量在具体语境中平衡不同的利益，抑制上述规避行为，避免规则过宽或过窄的问题，从而成为规则治理、法治的必要补充。<sup>⑥</sup>

算法权力的追求是更为清晰明确的规则，使违法行为无所遁形。这也就将规则的上述利弊都指数级放大，为执法与治理者施加了更多限制，为规避与违法者提供了更多指导，使双方权力关系发生转化，大大增强被执法与被治理一方的权力。因此，在规则力所不逮的适用范围之外，裁量性的标准仍有其存在的必要与作用。就此而论，机器学习颇有机器裁量的色彩，并无明确的算法权力规则，而是综合多种因素决策，似乎不易被规避，但机器学习除黑箱不透明的问题外，其实仍是可以发现规律并对其有意加以规避的，除非其决策完全任意，但这将取消机器学习的正当性，使之完全无法监管。事实上，所有基于标准的决策都是可以总结规律的，并且都倾向于使之转化为规则，司法经验之归纳便是如此，算法不过是促进并加速了这一进程。这也就使规则与标准之间的鸿沟问题更为凸显。

也正因为如此，为增强算法问责性而采取的算法透明与算法解释之解决方案，具有将进一步增

<sup>⑤</sup> 所谓通过降低立法信息成本而使法律成为个性化规则的原理，便是如此。参见 Ariel Porat & Lior Jacob Strahilevitz, “Personalizing Default Rules and Disclosure with Big Data”, *Michigan Law Review*, Vol. 112, No. 8 (2014), pp. 1422-1433。

<sup>⑥</sup> 关于规则与标准的区分，参见 Louis Kaplow, “Rules versus Standards: An Economic Analysis”, *Duke Law Journal*, Vol. 42, No. 3 (1992), p. 557; Ward Farnsworth, *The Legal Analyst: A Toolkit for Thinking about the Law*, Chicago and London: The University of Chicago Press, 2007, pp. 163-171。事实上，德沃金（Ronald Dworkin）所谓“原则”（正义或公平的要求，具有开放性与相对性）与“规则”（基于特定事实，以全有或全无的方式适用）的区分，亦与此类似，参见 Ronald Dworkin, *Taking Rights Seriously*, Cambridge, MA: Harvard University Press, 1977, chapter 3。

强算法规避者之博弈能力的不利影响。当前并不透明也无解释的算法权力，已被意欲规避者总结出规律，如果治理算法完全透明并被解释到人人都能理解的水平，那么算法规避的难度将大大降低。算法规避的现实存在与算法透明的负外部性，或许提示了算法黑箱并非一个亟待解决的重要问题，毕竟不论运行过程透明与否，算法权力的治理后果都是现实存在的，而该后果可以成为治理或规避的起点。然而，这并非等于主张算法问责毫无必要、倡导“法不可知，则威不可测”，仅是认为认清算法透明的局限性与副作用，合理定位其在算法问责中的辅助作用，乃是更为可取的算法治理之道。<sup>③⑧</sup> 故而，应当通过有限度的算法透明来辅助算法监管，确保权力运行的公开透明，但避免对算法规避者提供过多的信息，防止其以人工学习或机器学习方式创造更复杂高级的规避行为。

随着技术的发展，个性化规则开始出现，标准与规则开始融合，但这一趋势仅仅缓和了效率与公平的矛盾，甚至反倒加剧了个案公平与普遍公平的冲突。<sup>③⑨</sup> 这是因为，以技术手段确定的个案区分是算法权力精细化的极致，可以使每个人在具体场景下如何行为更具指引，但更难横向或纵向对比不同场景的不同处理是否合理、一致，导致治理权力缺乏体系性、统一性。故而，从标准到规则的鸿沟仍然存在，仅仅是转变为个性化规则与普遍化规则的鸿沟，而技术的精细化刚性区分，使这一鸿沟更难通过技术之外的其他方式予以弥合。这不仅是算法问题的治理所面临的挑战，更是未来社会所有问题之治理所共同面临的挑战。

### （三）多元规范社会秩序根深蒂固

将技术及其所执行的规范进一步还原至其所源自的社会语境，可以发现算法规避的根本原因在于多元规范社会秩序的根深蒂固。总有其他规范对立于算法权力所欲实施的规范，使其影响无法充分发挥。这对于避免成为治理对象的规避手段而言尤其如此。由于社会分群与分层的现实存在，多元规范秩序不可避免，不同平台、不同国家的规范分殊仅仅是个中一例。更由于互联网对强关系的增强与弱关系的增加、表达渠道的空前丰富，甚至由于信息茧房的不断加剧，社会分群、分层乃至分裂、极化愈发严重，统一的规范共识愈发稀缺。<sup>④⑩</sup> 因此，即使试图以较为刚性的算法技术实施包括正式法律在内的特定规范，也会遭遇对立面规范的强烈冲击，呈现出明显的规范竞争格局。

也正因为如此，算法规避不必然具有非正当性，仅仅是规避者对不同规范与不同秩序的选择差异而已，即使被规避的算法所执行的规范是法律时亦然。算法规避现象也不必然是一个有待解决的问题，或者说对此的解决方案不必然是一切向书本上的正式法律或规范看齐，承认规范多元、允许规范竞争亦无可。规避算法者亦不必然是坏人或法盲，他们不过是有自己的利益与价值偏好，甚至精通算法及其背后的规范。应当承认，算法所支持的规范在算法规避中仍然起作用，是算法规避的背景、起点与约束条件；没有前述规范的存在，算法规避不会采取特定的形式，正如“在法律的阴影下讨价还价”。所以，算法规避不仅揭示了算法预期目的与实际效果之差距的存在，而且还体现了多元规范与多元秩序对行为主体的多重影响。

更进一步来说，在多元规范与多元秩序的视角下，算法只是影响法律与规范秩序的一个因素，其本身并没有改变社会的独立力量，此为其最关键的局限性。受算法支持的法律与规范固然效率更

<sup>③⑧</sup> 参见 Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, and Harlan Yu, “Accountable Algorithms”, *University of Pennsylvania Law Review*, Vol. 165, No. 3 (2017), pp. 657-660; 沈伟伟 “算法透明原则的迷思——算法规制理论的批判”, 《环球法律评论》2019年第6期, 第34-35页。

<sup>③⑨</sup> See Anthony J. Casey & Anthony Niblett, “The Death of Rules and Standards”, *Indiana Law Journal*, Vol. 92, No. 4 (2017), pp. 1407-1423.

<sup>④⑩</sup> 参见陈昌凤、张心蔚 “信息个人化、信息偏向与技术性纠偏——新技术时代我们如何获取信息”, 《新闻与写作》2017年第8期, 第42-44页。



高、刚性更强，但强化治理也无法解决根本性的社会问题，无法将对立面的社会规范完全压倒。只要社会矛盾仍然存在，规避乃至违法的行为就不会消失，而只会转变形式、转移阵地，甚至向更极端、更剧烈的形式转化。创新与革新从来都不会停止，也不会被完美执法所阻却，而只会改变其出现的形态，要求社会付出或高或低的不同代价。<sup>①</sup>

不问准度、不计后果的算法权力滥用，乃至算法独裁，因此变得不可能。试图将规避所体现的灰色地带全部归为违法违规，纳入算法权力的作用范围，从而取消对边际性治理问题的利益衡量，这是诱惑极大且也危害极大的治理路径。这样做不仅可能危害到个人的权利与自由，而且也将反作用于算法权力本身，因为过分侵犯其他社会规范的算法将丧失其正当性，需要比日常治理更为巨大的资源投入，而治理效用的产出，则会因为被治理者持续的抗争行为而备受削弱。例如，为了打击用户账号培育而对评论发言施加过分的限制，其最终结果是用户出走其他平台；又如，为了打击洗稿而对著作权合理使用施加过分的限制，其最终结果或许不是大众创造性的完全枯竭，而是整个著作权保护制度被弃之不用。算法独裁固然令人害怕，但普遍性的失范状态同样令人恐惧。

综上，算法规避之所以存在，是因为技术的刚性有限，且不匹配处理边际性治理问题所需的利益衡量，无法将所有的“标准”转化为“规则”，解决不了多元规范秩序背后的根本性社会矛盾。这一规避既无法通过技术进步予以彻底解决，更可能因为算法透明等监管措施的出台而愈演愈烈。因此，鉴于算法规避将长期存在，在提出某种针对上述成因的技术、规范或社会解决方案之前，就必须先分析算法规避的后果如何及应当如何看待这一现象。

## 四、算法规避的后果与功能

通常而言，在提出问题、分析问题之后，便应当进入解决问题的阶段，探索应对算法规避的有效办法，但法社会学的视角并非如此积极实用，毋宁说更为“无情”。或许并非所有问题都存在或应当存在解决方案，算法规避作为与算法权力针锋相对的一种博弈，简单的解题思路可能收效甚微，甚至还会引起规避的继续升级。

更重要的是，规避的后果或许并不如想象的那样严重，又或者说规避本身可能也存在法社会学意义上的一种“功能”，具有现实的合理性，因此不必然是一个问题。算法被规避了又怎样？何必在新的平衡尚未可知时破坏现有的平衡？故而，本节将尝试更深层地理解算法规避的后果或功能，为后续的另一节中审慎地提出一些具有包容性的应对策略奠定基础。

### （一）算法规避的实际负面后果

一般来说，首先，算法规避意味着算法权力所欲实现的目标并未实现或并未全部实现，降低了算法实施的效率，例如内容治理算法本欲减少各种违法违规的内容，营造清朗的网络空间，但规避者仍然发布并传播了此类内容，即导致治理的目标落空，可能不得不回到传统上效率更低的人工审查上来。其次，算法规避导致一部分人能够获得比其他人更多的收益，或者能够更多避免算法带来的不利影响，引发算法实施的不公平，甚至强化社会既存的不平等结构。例如，有能力以选择平台

<sup>①</sup> 关于对技术带来完美执法、过分限制个人自由的担忧，参见 Lisa A. Shay, et al, “Confronting Automated Law Enforcement”, in Ryan Calo, A. Michael Froomkin, and Ian Kerr, eds., *Robot Law*, Cheltenham, UK Northampton, MA, USA: Edward Elgar Publishing Ltd., 2016, pp. 235-273.

或国家来规避算法者，往往具有更高的知识水平、更强的经济能力，而规避算法以后，其获得的信息与资源相比于未规避者往往又有所提高，导致分化更加剧烈。最后，在算法规避的过程当中，算法权力与规避者之间的持续博弈，可能造成社会资源的系统性浪费，导致社会福利的总体性下降。例如，“网络水军”与评分算法之间的持续博弈，最终并未显著提高评分算法的合理性，反而可能因为在反刷分领域的巨大投入，而使得整体的“成本—收益”分析之结果为负。<sup>⑫</sup>

但上述分析只能说明算法规避具有降低效率、引发不公、导致社会资源浪费的潜在可能，无法体现负面后果的实际影响范围与严重程度。既然即使在算法等技术权力的赋能下也没有任何法律或规范能够得到完全执行，“书本上的法”与“行动中的法”之间的差距始终存在，不同法律或规范之间的实效差别，仅仅在于违反的比例与后果的严重程度，那么，究竟有多少人规避算法、有多少人服从算法，就成为算法权力是否有效、算法规避是不是一个需要解决的问题之前提。对此，可以借助威慑曲线这一理论工具予以说明。

所谓威慑曲线，是以执法概率或惩罚后果为横坐标，以威慑效果为纵坐标，得到正向相关的一条曲线。其系曲线而非直线的这一本质，提示着因果关系的非线性，而且作为凸曲线而非凹曲线，表明边际效果是递减而非递增的。换句话说，提高执法概率或加重惩罚后果能够带来的威慑效果越来越少，在达到一定数值后甚至趋近于零。<sup>⑬</sup> 该特点在经验上的意义在于：许多人是受威慑影响的，其固有的偏好无法被威慑所改变，要么是极端的守法者（比如无论如何都不会想象自己杀人越货的普通人），要么是极端的违法者（比如某些无论如何都不会被惩罚阻却的恐怖分子、反社会分子）。其背后的原因在于社会规范、个人内心准则对法律的支持或否定。因此，对威慑问题的讨论，需要集中关注可被威慑之人，通过自变量的调节，边际地改变其行为，否则就会导致执法资源的浪费。

由于算法规避所针对的是借算法权力实施的法律或规范，故而同样可通过上述威慑曲线进行分析。又因为不同算法规避问题对应的威慑曲线亦有所不同，所以此处以平台内容治理算法为例进行说明，然后对所得出的结论予以拓展。当前互联网平台的内容治理已相对完善，相关的平台规则已较为健全，并在算法的协助下得到了有效执行，成功治理了许多网络内容问题，维持了良好的网络内容秩序，能够满足法律要求与用户需求，甚至存在平台算法权力过于强大、可能被滥用的风险。因此，平台内容治理算法相关的威慑曲线已变得相对平缓。

这一威慑曲线的平缓，在经验上体现为许多人甚至大部分人发布的网络内容都平凡无害，并不经常考虑如何最大限度地畅所欲言，因为社会规范或个人内心准则的约束，导致其不敢甚至从未想过发表过于越界的内容。即使某些低俗内容为其所乐见，他们在平台规则乃至法律的威慑下，也会更多地选择调整自身行为以满足治理要求的策略，而非明目张胆地通过选择平台或采取迷惑混淆的手段来直接实施违法违规的行为，而这在较为宽松的标准审视下仍使治理目的得以实现，并且平衡了网络治理与个人自由。只有少数别有用心且又技高一筹的“坏人”才会整日思考如何发表越界的内容，而在算法与法律的治理水平不断提高时，那些仍无论如何都会绞尽脑汁规避内容治理算法者，又是其中的少数，位于威慑曲线的末端，属于极端的无法被威慑者。如果将执法资源集中于这

<sup>⑫</sup> *Supra* note ⑭, pp. 25–33.

<sup>⑬</sup> 参见注⑨，Lawrence M. Friedman 书，pp. 105–112。但劳伦斯·弗里德曼（Lawrence Friedman）的这一理论仍然存在进一步探究与深化的空间，例如以经验观之，威慑曲线的严格单调递减性或许仍有疑问，执法概率从1%提升到2%，很可能比从89%提升到90%的边际威慑效果要大，但从99%到100%这个关键一跃，反而又有可能比从91%到92%的边际威慑效果要大；又如，执法概率和惩罚后果或许也不能混为一谈，二者与威慑效果很可能分别构成不同形态的曲线。但就本文的论述目标而言，引用威慑理论主要是为了论证在互联网治理情况并未发生重大问题的当下，算法规避或违反者属于相对少数，故而前述发展暂不需要在本文中详究。

一末端，那么在“成本－收益”分析视角下可能显得并不合乎经济理性。故而，过分夸大算法规避的负面后果，是不切实际、没有必要的。

统计数据亦能佐证算法规避在平台内容治理领域的相对少数地位。以 Facebook 发布的透明度报告为例。自 2018 年以来，违反 Facebook 关于成人裸露与性行为规定的内容在平台总内容中的占比，最高时仅有 0.14%，亦即 1 万条内容中仅有约 14 条违规，最低时则仅有 0.02%。<sup>④</sup> 诚然，以 Facebook 的内容总量计，再低的比例也会被放大为巨大的绝对数量，而且色情内容仅仅只是违规内容当中的一种，此外还存在许多其他类型的违规内容。但同样不能忽视平台的用户总量与用户受信息茧房所限而能够实际接触到的内容，于是，单个用户遭遇违规内容的概率就很低了。故而，该数据也进一步支持了上述理论分析与经验观察。

因此，只要仍然存在相对稳定的社会秩序，那么算法规避对效率、公平与社会资源配置的负面后果就不会像设想的那样严重，大多数人仍受算法权力治理，并由于算法的持续升级而越发受到控制。更高级的算法规避产生的另一面，是大量较低级的算法规避与违法违规已被阻却或惩罚，而这正是算法对规范实施效率的巨大提升。更进一步来说，通过信息茧房、信息洪流等认知操纵手段，大多数人不仅在“成本－收益”计算的基础上选择不去违反或规避算法，更可能从一开始就没有违反或规避算法的意识或心理选项，而是默认算法规定的合理性，由此实现执法与守法的统一。<sup>⑤</sup> 因此，对算法权力的反思与批判，并不等于支持虚无主义的法律观与技术观。过分放大该问题，不仅没有必要，而且更可能由于高估了人的主观能动性而降低了对算法权力的警惕性，阻碍将防范权力失控设置为更紧迫的算法治理议题。

## （二）算法规避的隐性正面功能

更重要的是，算法规避事实上同样发挥着限制算法权力、对抗算法滥用的功能。当前的算法规制方案各有局限性。也就是说，算法过程的透明，既可能因为技术过程的复杂与使用主体的阻碍而缺乏可行性，也可能因为与私人利益、公共利益的冲突而缺乏可欲性；<sup>⑥</sup> 而算法结果的公平公正、算法所用数据的合规清洁，不足以打消人们对人工智能分析与决策的全部怀疑。因此，适当的改进方案，除了将前述进路予以有机结合外，利用算法规避这一被治理者在多元规范秩序视角下的第二方控制，或许亦有助益。<sup>⑦</sup> 例如，如果用户认为内容治理算法过于严格，或者觉得自己受到大数据“杀熟”，缺乏透明性或公正性，除了选择向更高层级的权力主体申诉外，还可以通过前述规避策略身体力行地主动抵制，发布算法所不允许的内容，阻碍算法目的实现，对权利进行自我救济。相比于申诉进路的结果不确定与过程缓慢，直接规避显然更有效率，也更能引起算法权力主体的警觉，推动算法权力的不断完善。

更一般而言，算法规避的限权作用，提示着算法权力的实施仍然存在缝隙，尚未过分限制个人自由，而这可以缓和被治理者的反弹，有利于维护社会稳定。此为算法规避的更深层次的隐性功能。应当看到，即使是既有效率又合规范且满足过程透明、结果公正的规范执行，也是对个人自由

<sup>④</sup> 参见《Facebook 透明度报告》，载 Facebook 官网 <https://transparency.fb.com/data/community-standards-enforcement/adult-nudity-and-sexual-activity/facebook/#prevalence>，2022 年 3 月 22 日访问。

<sup>⑤</sup> 参见王凌峰、葛岩、秦裕林“多学科视角下的守法行为研究——兼论自动守法中的高效认知界面优化”，《浙江社会科学》2015 年第 8 期，第 37－41 页；应飞虎、李宣“行为经济学促进守法的机理与路径”，《深圳大学学报》（人文社会科学版）2021 年第 6 期，第 93－98 页。

<sup>⑥</sup> 参见解正山“算法决策规制——以算法‘解释权’为中心”，《现代法学》2020 年第 1 期，第 189－192 页；注④，丁晓东文，第 106－107 页。

<sup>⑦</sup> 参见〔美〕罗伯特·C. 埃里克森《无需法律的秩序——邻人如何解决纠纷》，苏力译，中国政法大学出版社 2003 年版，第 149 页。

的阻碍，可能引发个体对抗。算法权力可以大大提高规范执行的效率，也因而必然对个人自由产生更大的约束，使人们无法畅所欲言、来去自如，阻碍可能的破坏性创新与改革。然而，算法规避的存在舒缓了这一紧张关系。规范执行得不那么有效率（亦即规范直接效果的降低），在某些情况下反而更有利于间接社会效果的最大化：新权力的“非法兴起”重获可能，治理与自由的边界仍然处于健康的流变与对话之中，而非被迫进行非此即彼的激烈斗争。例如，网络用户通过使用谐音、拼音、首字母、拆字、联想词等规避内容审查，表达、呈现了许多法律法规未必真正意欲限制、但又因各类原因而未给出明确豁免的内容，使原本过宽的内容限制条款得以细化、限缩，将粗疏的法律规定落实为可操作的、决定具体场景下言说与否的网络表达规则。<sup>④⑧</sup>

如果算法权力过分强大，那么甚至即使算法权力可以通过认知操纵促使执法与守法相统一，也仍然会存在少数人醒悟并反弹的可能性，并且反弹将更为剧烈。这是因为，上文所述多元规范秩序的根深蒂固，并非技术所能彻底改变。试想一下，若将内容治理算法发展到极致，没有任何违法违规或规避能够逍遥法外，甚至无从发声，则网络空间是否将成为一潭死水？非也。少数人仍将寻找突破架构控制的办法，另起炉灶，开辟新的言论空间，甚至将言论表达从线上转移至线下，依靠更为原始的方式发声引众，传播思想。例如，由于版权制度过于严格，过分保护版权人利益而阻碍后来作者的创新，曾有人主张建立与版权（copyright）制度针锋相对的版权开放（copyleft）制度。<sup>④⑨</sup>又如，当下热议的元宇宙，事实上源于 Facebook 试图摆脱操作系统层面的掣肘，直接触达并管理用户，这也说明了建立一套全新的控制架构并非不可能。<sup>⑤⑩</sup>而且，互联网对多元规范社会秩序的增强，将提供更多替代性的社会空间与社会力量，赋权并非仅仅针对博弈中的任何一方，且更不利于试图控制网络空间的公共主体，威慑曲线的趋于平缓将更为迅速。

由此可见，包容一定比例的算法规避，有利于限制算法权力或调节算法权力的实施效率，适当限定其适用范围，平衡治理与自由，促进社会稳定与治理的综合效率，在相对柔性地促使大多数人不选择成本较高的突破架构阻碍选项时，仍为少数能力超群者留下规避的空间，从而避免完美执法后果的出现，为个人自由的充分舒展、法律与社会规范的良性互动，以及可能的改革与创新，奠定必要基础。

## 五、算法规避的分类规制

承认算法权力的总体有效性乃至算法规避的隐性正面功能，并不等于主张对算法规避无所作为、放任自流，否则权力关系此消彼长，负面后果就会变得严重。由于某些算法规避仅仅利用治理者在信息与资源方面的局限性，并未触及技术与规则的边界，不涉及难办的多元规范竞争问题，对效率、公平与社会福利的损害较突出，对算法权力限制的支持相对次要，故而应当通过技术能力的增强与规则执行的完善进行治理。另一些算法规避虽然确实触及边际性利益衡量与多元规范竞争，其负面后果与正面功能都相对不确定，甚至收益可能大于成本，但仍然应当通过非法律与非技术的柔性方式进行治理，并保持与对立面的沟通和相互谅解。这一区分可被概括为简单案件与困难案件

<sup>④⑧</sup> 参见戴昕 “重新发现社会规范：中国网络法的经济社会学视角”，《学术月刊》2019年第2期，第117页。

<sup>④⑨</sup> 参见常安、燕辉 “从 Copyleft 看开放获取的社会理想及其对图书馆发展转型的启示——以 Aaron Swartz 及其《Guerilla Open Access Manifesto》为引子”，《图书馆》2014年第3期，第15-16页。

<sup>⑤⑩</sup> 参见胡泳、刘纯懿 “元宇宙转向：重思数字时代平台的价值、危机与未来”，《新闻与写作》2022年第3期，第47-49页。

(或称难办案件、复杂案件)的经典二分,二者的区别在于是否存在明确的规则可供适用。<sup>⑤①</sup>此外,针对困难案件的社会规范治理进路,不仅有利于避免“困难案件制造坏的法律”,而且也符合对新兴互联网问题坚持包容审慎治理的原则。<sup>⑤②</sup>

### (一) 规则执行的技术增强

算法规避策略中的相对简单者,包括避免成为治理对象中的平台选择、调整满足治理要求中的虚假申报、简单的行为调整、混淆迷惑治理主体。四者皆存在较为明确的规则可供适用,所需者仅仅是提高规则的执行水平,而技术对此能够发挥最大的作用,并且因此较少存在过于刚性或不够透明的问题,辅助技术的人工成本并不高昂。

就平台选择而言,“非法兴起”尽管是小型新兴平台与大型垄断平台竞争的重要手段,但互联网早已不是法外之地,任何平台皆不能逾越法律的底线而为违法违规行为提供庇护。例如,对于淫秽色情内容、假冒伪劣商品,平台应当通过包括算法在内的各种手段予以阻却与惩罚。而利用平台间的数据与信息壁垒从事明确的跨平台违法违规行为,同样应当受到禁止,对此,应当通过超越平台的国家层面执法或行业层面自律进行规制,例如部委联合组织的“清朗”“净网”系列行动、中国互联网协会等网络社会组织发起的行业评议与奖惩等。

就虚假申报而言,隐私与个人信息保护,不能作为以虚假信息披露来规避法律强制性规定的理由。一方面,法律虽然限制了个人信息处理者收集、处理、储存、传播个人信息的相关行为,但对此并非完全予以禁止,甚至还要求其为服务或执法所需而收集最小必要的个人信息。另一方面,前台匿名、后台实名的架构设置,能够平衡隐私保护与网络治理,而大数据分析亦可减少乃至杜绝通过虚假申报来隐瞒必要身份信息的行为。因此,法律与技术双管齐下,要求每个人都应当为其在网络空间中的所做所言负责。

就简单的行为调整而言,由于其本质的违法违规易于识别,故而应当在区块链取证所奠定的信息基础与网络架构控制所奠定的执行基础上,不断提高大数据分析、人工智能识别的准确度,高效认定算法规避导致的违法违规行为,同时辅以必要的人工审查作为补充,不能因复杂的行为调整存在就放弃对简单者进行规制。并且,由于此类规避并未触及规则边界,故而黑箱技术的治理结果更易受到人类理性的承认,算法透明与算法信任方面的问题较轻微,能够在最大限度发挥技术治理准确高效之优势的同时,减轻其局限性带来的负面影响。

就混淆迷惑治理主体而言,由于规避者从事的亦为直接的违法违规行为,仅仅利用信息洪流进行掩饰,故而同样不存在正当化理由。可以遵循前述的相似规制方法,通过人工审查辅助与机器学习技术,提高数据分析与智能识别的精度,避免人工智能被不智能的简单处理所蒙蔽。事实上,随着内容治理的不断发展,通过与空白内容相整合或明暗、色调变化等处理进行迷惑混淆的技巧,已越来越为算法所熟悉,因此迷惑混淆将变得越发困难,进而同样转入复杂的行为调整。而对于此类规避,技术与规则的作用有限,需要更多发挥社会规范的作用。

### (二) 社会规范的治理应用

算法规避策略中的相对困难、难办、复杂者,主要是调整满足治理要求中的复杂行为调整。再次强调,此处所谓“复杂”,并不是从单纯技术意义上(例如跨媒介洗稿的各种变换与拼接手段)来讲的,而是即使所用技术简单或并未使用任何技术,但触及边际性利益衡量与多元规范竞争,法

<sup>⑤①</sup> 参见注⑤⑦, Ronald Dworkin 书, pp. 228 - 230; 苏力“法条主义、民意与难办案件”,《中外法学》2009 年第 1 期,第 93 页。

<sup>⑤②</sup> 参见刘权“数字经济视域下包容审慎监管的法治逻辑”,《法学研究》2022 年第 4 期,第 45 - 50 页。

律性质认定极为困难,例如规避敏感词审查的各种转化方式、软色情、意译式洗稿等人类智能尚难裁量、人工智能更难判断的问题。因此,需要保持规制宽容,并通过社会规范进行治理,从而更有利于相关规制维护政治秩序、保持公序良俗、激励创作创新等目的之实现。

具体来说,首先,平台应当继续利用用户举报、评分、自我组织与管理等方式,充分发挥平台直接治理体制之外的个人与组织的自我治理作用。这不但能够将公众的信息、知识及其支持的社会规范有机融入算法之治,有利于边际性治理问题的利益衡量,<sup>⑤③</sup>而且还能够强化用户对算法的认同,促进其自愿守规,甚至最终将算法要求内化为无意识行为,有利于提高算法治理的正当性,降低算法治理的成本。<sup>⑤④</sup>这一规范执行与规范遵守的高度统一,以及法律、平台规则、社会规范的相互协调,正是所谓“网络生态治理”<sup>⑤⑤</sup>的题中之意。

其次,行业协会应当继续完善行业自律规范的制定与执行。一方面,应当通过行业实践经验的总结,细化法律未能明确规定的各类违规行为为判断规则,为算法权力主体的治理提供指引,减少算法规避者之自由调整行为的空间。另一方面,应当通过行业评议与奖惩,对违规行为及相关的直接与间接责任者进行力度适中的激励或抑制,从而减少确实损害社会福利的算法规避行为。<sup>⑤⑥</sup>但是,不论是自律规范的制定还是执行,都应当始终保持开放的治理态度,及时适应互联网环境的变化,在实践中不断进行调整。

最后,政府应当继续推广并完善社会信用制度,使之从经济领域向更为广泛的社会治理领域延伸。就此而言,可以网络实名制为基础,为每个网络用户建立信用档案,对受到其他用户举报、经治理者审核确认的违法违规者,根据其情节与后果的严重程度进行扣分,并根据低分所处区间,施以相应程度的惩罚。同时,对某些影响较为广泛、具有公共人物性质的用户,可以考虑设置公开的用户评论区,展示纯粹的信用分数无法充分体现的用户意见,为其他用户的评价提供充足的信息参考,从而形成良性的市场反馈循环。<sup>⑤⑦</sup>

但应当注意,由于某些平台选择、虚假申报、简单的行为调整与混淆迷惑治理主体的策略所规避的算法执行的规范,同样是有争议或有问题的法律、平台规则或社会规范,但孰是孰非无法在短时间内根据单方判断决定,因此亦留有多元规范竞争、保持规制宽容的余地。例如,为规避平台歧视消费者或盘剥劳动者的算法而选择其他平台,又或者为规避对老人、女性或特定地域用户的歧视算法而谎报数据,以及其他为规避不合理的规范及其相关算法而混淆迷惑或调整行为。此时需要治理的,往往并非算法规避,而是被规避的算法本身;算法规避作为个人自由之行使,此时具有高度正当性,发挥着上述直接限制算法权力、对抗算法滥用、促使算法向善的隐性功能。

与之类似,避免成为治理对象中的国家选择,同样涉及不同法律规范之争,因此难以判断孰是

<sup>⑤③</sup> 例如,微信曾提出“合议机制”来解决洗稿问题“对于判定可能有‘洗稿’争议的内容,经投诉方确认发起公开投诉、被投诉方确认回应后,平台将随机邀请一定数量的‘合议小组’成员参与评定,在规定时间内回收结果大于等于10份,且大于等于70%认为被投诉方‘洗稿’,则按照合议的多数意见认定为存在‘洗稿’。反之,若达不到70%的规则比例,则不存在‘洗稿’。”参见《守护原创!微信公众平台“洗稿”投诉合议规则》,载“微信法务”公众号 <https://mp.weixin.qq.com/s/841maq18WGHRjv15U690Q>, 2023年3月3日访问。

<sup>⑤④</sup> 参见李娜“守法社会的建设:内涵、机理与路径探讨”,《法学家》2018年第5期,第17页。

<sup>⑤⑤</sup> 国家网信办2019年12月出台、自2020年3月1日起施行的《网络信息内容生态治理规定》,在第2条第2款规定“本规定所称网络信息内容生态治理,是指政府、企业、社会、网民等主体,以培育和践行社会主义核心价值观为根本,以网络信息内容为主要治理对象,以建立健全网络综合治理体系、营造清朗的网络空间、建设良好的网络生态为目标,开展的弘扬正能量、处置违法和不良信息等相关活动。”

<sup>⑤⑥</sup> 参见陈华“行业自律的北京样本——北京网络媒体协会行为研究”,《新闻与写作》2010年第11期,第45-46页。

<sup>⑤⑦</sup> 事实上,这正是微博与淘宝的信用制度之结合,而实践也证明了其效用。参见邱遥堃“网络洗稿规制的反思与重述”,《出版发行研究》2021年第8期,第78页。

孰非。算法规避在此种情况下具有缓和被治理者反弹、维护社会稳定的深层隐性功能，例如低俗色情治理中的性自由之争，或者洗稿抄袭治理中的信息自由之争。但由于不同法律选择背后皆有各自国家强制力作为支撑，故而在治理结果上更接近对简单低级算法规避的强化规制，只是由于争议较大而始终存在规避与反规避的相互博弈。

## 结 语

本文对算法规避的研究，始终坚持法社会学的视角，并非先验地主张算法规避有害、应当受到规制，而是经验地分析算法规避的三种策略、多重原因、实际后果与隐性功能，并据此提出简要的规制方案。事实上，这一法社会学的经验进路，适用于所有法学研究与现实规制，但在新兴领域，特别是互联网治理领域，特别值得强调。新兴互联网治理问题往往以特殊性为理由或借口，要求另行规制或不规制，从而掩盖其新瓶装旧酒的本质。<sup>58</sup> 但正如算法权力不过是算法对多元规范的技术赋能、算法规避只是用户权力对算法权力的反抗一样，从经验视角切入，互联网新问题的本来面貌即可还原，甚至进入法学教义旧瓶装新酒的既定套路。因此，在这个意义上讲，经验进路也成为法教义学与社科法学相结合的最大公约数，使传统法学与其他学科在具体的观察与分析中得以沟通与协调，进而共同提出全面的解决方案。而这也是本文通过算法规避这个细小的切口，所欲揭示的更为一般的研究与规制之发展道路。

### 【主要参考文献】

1. [美] 劳伦斯·莱斯格 《代码 2.0：网络空间中的法律》，李旭、沈伟伟译，清华大学出版社 2009 年版。
2. 胡凌 “‘分享经济’的法律规制”，《文化纵横》2015 年第 4 期。
3. 戴昕 “重新发现社会规范：中国网络法的经济社会学视角”，《学术月刊》2019 年第 2 期。
4. 沈伟伟 “算法透明原则的迷思——算法规制理论的批判”，《环球法律评论》2019 年第 6 期。
5. 邱遥堃 “网络洗稿规制的反思与重述”，《出版发行研究》2021 年第 8 期。
6. 丁晓东 “基于信任的自动化决策：算法解释权的原理反思与制度重构”，《中国法学》2022 年第 1 期。
7. 刘权 “数字经济视域下包容审慎监管的法治逻辑”，《法学研究》2022 年第 4 期。
8. Robert H. Mnookin & Lewis Kornhauser, “Bargaining in the Shadow of the Law: The Case of Divorce”, *Yale Law Journal*, Vol. 88, No. 5 (1979).
9. Lawrence M. Friedman, *Impact: How Law Affects Behavior*, Cambridge: Harvard University Press, 2016.
10. Jane Bambauer & Tal Zarsky, “The Algorithm Game”, *Notre Dame Law Review*, Vol. 94, No. 1 (2018).

(责任编辑：尤陈俊)

<sup>58</sup> See Eric Goldman, “The Third Wave of Internet Exceptionalism”, in Berin Szoka & Adam Marcus, eds., *The Next Digital Decade: Essays on the Future of the Internet*, Washington, D. C.: TechFreedom, 2010, pp. 165-167.

“similar cases” is the direct reappearance of the similarity of cases on the basis of “value unit”. The mechanism for realizing the “similar judgement” is tending to convergence of the judicial opinions in similar cases on the basis of comprehensive consideration of judicial authority and judicial rationality. The corresponding system arrangements are the institutionalization of judicial consensus ,focusing on the construction of litigation procedures for “like cases should be treated alike” ,the systematic improvements of the precedents application system ,and the informatization energization of “like cases should be treated alike”.

**Key Words** Like Cases Should Be Treated Alike; Application of Precedents; Guidance Case System; Guiding Case; Operation of the Court as A Whole

Li Zhenxian ,Ph. D. in Law ,Lecturer of Law School ,Southwestern University of Finance and Economics.

### ***The Evasion of Algorithm and Its Regulation: From the Perspective of Legal Sociology***

QIU Yaokun • 60 •

Although the algorithm power is strong ,it may still be evaded ,and there is a continuous dynamic game relationship between it and its objects of governance. The basic strategies of evasion of algorithm include avoiding becoming the objects of governance ,adjusting to meet the requirements of governance and confusing the subjects of governance. The reason for this phenomenon is that the rigidity of technology is limited ,and it does not match the balance of interests required for the treatment of marginal governance issues. It is unable to convert all “standards” into “rules” and solve the fundamental social contradictions behind the pluralistic normative order. However ,most people are not evaders of algorithm ,and the algorithm power is generally effective. A few evasions of algorithm also help limit the algorithm power ,balancing governance and freedom ,and promoting social stability and the comprehensive efficiency of governance. We only need to adhere to the principle of inclusive and prudent regulation ,strengthen the regulation of simple evasions of algorithm ,maintain tolerance for complex evasions of algorithm ,and explore the regulative role of social norms.

**Key Words** Algorithm; Evasion of Algorithm; Legal Evasion; Internet Governance; Pluralistic Norms; Law and Technology

Qiu Yaokun ,Ph. D. in Law ,Lecturer of Law School of Capital University of Economics and Business.

### ***On the “In Accordance with Law” in the Chinese Civil Code***

LI Chang • 75 •

There are 162 “Yifa (in Accordance with law) ” in the Chinese Civil Code. This mode of legislation has a long history in China. However ,there has been a lack of academic analysis of this term. The function of this term is to indicate other norms in positive law ,which includes all levels of norms. But in fact this term has no real normative function. When words appear in the language game of the system of norms ,the elements of the words have been influenced by other norms ,which replace the role of this term. And this term has no indicative function beyond the word itself. The customary use of this term originates from the General Principles of Civil Law and was intended to fulfil the task of propaganda and education about the rule of law at that time. However ,the current misuse of this term has created many practical problems. Legislators