

# 公共治理视域下自动化应用的法律规制

唐林垚\*

## 目次

|                      |                        |
|----------------------|------------------------|
| 一、自动化应用介入公共治理领域的法律问题 | 三、公共治理领域自动化应用法律法规的全新面向 |
| (一) 新公共治理语境下的自动化应用   | (一) 宏观技术共治：以程序透明取代代码透明 |
| (二) 自动化应用的角色与功能变迁    | (二) 中观价值位阶：场景理论助推技术性规制 |
| (三) 自动化应用规制的法律聚焦     | (三) 微观权利保障：避免自然人决策名存实亡 |
| 二、自动化应用治理导向法律规则的构建机理 | 四、结语                   |
| (一) 自动化应用的三重维度区分     |                        |
| (二) 公共治理理论的立场辨析      |                        |
| (三) 传统法律应对的局限与失灵     |                        |

**摘要** 技术飞跃与治理提升之间的关联通常被假定,也早已被吸纳进“数字中国”“智慧社会”等政策实践,但法律的习惯性滞后带来了监管套利的空间。公共治理领域的自动化应用经历了从基于数据库编码的计算机自动化到基于机器学习的算法自动化再到基于神经网络的超级自动化的逐步跃迁,从早先自下而上专业人士的“辅助/参考”和个别部门的“部署/应用”嬗变为自上而下的社会化“嵌入/集成”,公权力、私权力及私权利的关系发生了结构性转换。相应地,国际通行、日益趋同的个人隐私保护(事前同意)、算法可解释性(事中监测)和完整履责链条(事后追责)的监管范式也应依照宏观技术共治、中观价值位阶和微观权利保障三位一体的规制路径进行体系化重构。

**关键词** 公共治理 人工智能 自动化应用 算法规制

党的十九届五中全会提出,“十四五”时期经济社会发展要以推动高质量发展为主题,这是根

\* 中国社会科学院法学研究所助理研究员、法学博士。本文系国家自然科学基金青年项目“自动化应用提升现代化治理的法律保障研究”(项目编号:20CFX006)的阶段性研究成果。

据我国发展阶段、发展环境、发展条件变化做出的科学判断。<sup>〔1〕</sup> 自动化应用能够提高经济社会的智能化水平,有效增强公共服务和城市管理能力,是提升现代化治理不可或缺的技术加持。在“黑天鹅”和“灰犀牛”无处不在的风险社会,依靠大数据、云计算、机器学习以及各类算法工具营造的行政“社会共律”,是独立个体普遍“自律不足”的必要补充;健康码等自动化应用介入公共治理具有了充分的合法性根基。《法治中国建设规划(2020—2025年)》要求在2022年年底建成全国一体化的电子政务服务平台,推动人民群众参与社会治理,打造共建共治共享的社会治理格局;电子化审批、算法化决策、智能化评估和服务等,已在多个领域先行试点。

自动化应用提升现代化治理,需要强大的经济和法律制度作为后盾。“由于技术变革带来的‘外部性’不断变化,行为主体难以就风险程度、治理事项的优先排序、治理目标与手段的选择等问题形成共识,这使得安全治理在起点的原则规范上就缺乏坚实基础。”<sup>〔2〕</sup>囿于对不断涌现的新形势、新业态、新模式的认知不足,当前个人信息保护法和数据安全法的立法和阐释进路,正不可避免地步入过于关注隐私保护的欧盟《通用数据保护条例》、强化信息披露的法国《行政法典》以及确保责任主体链条可追踪识别的美国《算法问责法案》等“问题驱动型立法”的路径依赖,<sup>〔3〕</sup>“头痛医头、脚痛医脚”的立法范式,难以为自动化技术的跳跃式发展和管理部门的跨界式应用提供保障。研究表明,完全从信息保护入手建立规制体系的尝试,已经被证明面临加剧社会割裂的风险与徒增合规成本等流弊。<sup>〔4〕</sup> 如何突破公私二元界限、挣脱传统法律规制的制度惯性,采取整体化应对而非单纯数据保护的思路,以期自动化应用介入现代化治理保驾护航,无疑是迈向全面建设社会主义现代化国家新征程之时亟待思考的重要问题。

## 一、自动化应用介入公共治理领域的法律问题

公共治理领域自动化应用的是非利弊之争,由来已久、反复无常、难分伯仲,始见于1978年法国《数据处理、文件和个人自由法案》颁布前夕、兴盛于1995年欧盟《数据保护指令》适用之后、延续至2018年欧盟《通用数据保护条例》的监管实践,与新公共治理理论的由盛转衰、新公共服务理论的迅速迭代和公共价值管理理论的异军突起相伴相随。

### (一) 新公共治理语境下的自动化应用

早在1971年,莫顿(Michael Morton)在《反思智能辅助系统:基于计算机的自动化决策》一书开头就发出了“切勿盲目迷信机器自动化决策”的警告,<sup>〔5〕</sup>并直接促成了法国《数据处理、文件和个人自由法案》对司法行政自动化应用的一概禁止。然而,计算机自动化的发展和应用契合了早期新公共治理理论为解决监管机构资源匮乏而被迫演化出的“按需监管”路径,用清晰有序的风险

〔1〕 王昌林:《深入学习贯彻党的十九届五中全会精神 以推动高质量发展为主题》,载《人民日报》2020年11月17日,第9版。

〔2〕 刘杨钺:《技术变革与网络空间安全治理:拥抱“不确定的时代”》,载《社会科学》2020年第9期,第41页。

〔3〕 参见苏宇、高文英:《个人信息的身分识别标准:源流、实践与反思》,载《交大法学》2019年第4期,第54—71页。

〔4〕 参见韩新远:《个人行为轨迹信息的法律属性与分类保护研究》,载《交大法学》2021年第3期,第69—79页。

〔5〕 Morton M. S. S., *Management Decision Systems: Computer-Based Support for Decision Making*, Division of Research, Graduate School of Business Administration, Harvard University, 1971, p.1.

评估框架取代了过去凭空想象的资源分配顺序,实乃既定禀赋下公共治理能力提升之“自然选择”,在法国之外的其他国家发展迅猛。彼时学界前瞻性地指出,自动化应用“人为打压理性数据分析的地位,为获得似是而非的简化结果毫无原则、不计后果地对某些因素进行加权并赋予相对重要性”;〔6〕对美国大型城市公共信息系统的早期实证研究也表明,鉴于某些公共任务本身所蕴含的政治属性,不能一味地让计算机系统在“涉及公共权利的事项上为所欲为”。〔7〕

不过,这些理性的呼声几乎完全淹没在新公共治理理论学者对“解构”“非领地化”和“变样”的狂热之中,技术层面的高歌猛进与传统古德诺行政二分论和韦伯科层制论的节节败退形成鲜明对比,很大程度上满足了激进派“掀起一场官僚制度革命”的主观想象。即便是与新公共治理理论学者所倡导的“层层契约制”格格不入的新公共服务理论学者,也大多坚信技术进步将促进社会福利的雨露均沾——算力提升让过去难以协同一致的计算机系统间可以合乎逻辑地彼此联动,极大地消除了管理部门之间的隔阂,使得“系统遗留问题得以自然消解”。〔8〕在技术裹挟治理的大趋势下,公域软法的必要性,日益凸显。

## (二) 自动化应用的角色与功能变迁

将集体协商和网络治理贯穿于公共服务的获取和传递之中、轻“个体偏好的简单叠加”重“利益相关者的对话与参与”,新锐公共价值管理理论学者不约而同地主张应跟随技术变革的脚步不断修正民主与效率的关系,但他们对自动化应用的评价却呈现出两极化趋势。莱顿·安德鲁斯(Leighton Andrews)认为,技术进步促进了公共服务有效和规范地传递,是对缺乏专业决策、响应迟延和服务标准化匮乏的官僚主义的矫正,能够有效促使“技术官僚的角色从管理者向公共利益共同创造者和维护者转变”。〔9〕对此,迈克尔·利普斯基(Michael Lipsky)针锋相对地指出,“公共服务的核心,在于自然人可以而机器人却难以实现的‘推理和决策’”,借此唱衰公共行政领域的自动化应用。〔10〕

利普斯基的上述悲观言论至今已经整整过去十一年,在公共治理领域,无论是行政部门、人事部门还是公共企业乃至司法部门的自动化应用不仅有增无减,还大有百花齐放之势。2016年,美国威斯康星州最高法院在“州政府诉卢米思案”(State v. Loomis)中正式认可法官借助自动化系统进行量刑裁判的合理性与正当性,意味着人工智能对司法的影响从辅助层面正式进入实质层面。长期引领工会斗争之风的威斯康星州的这一举动,旋即激发了学界对公共治理领域自动化滥用的担忧。连司法量刑裁判都可以通过自动化系统完成,人工智能技术的应用是否还存在边界?自动化帝国主义侵入传统的司法审判领域引发了几乎是规律性的部门法学者的反思和诘问,例如,人工智能能否真正模拟法律议论?〔11〕机器学习何以辅助事实认定?〔12〕在管理部门“数字化”的大背景下,自动

〔6〕 Christie G. C., *An Essay on Discretion*, 5 *Duke Law Journal* 747 (1986).

〔7〕 Northrop A., Kraemer K. L., Dunkle D. & King J. L., *Payoffs from Computerization: Lessons over Time*, 50<sup>th</sup> year *Public Administration Review* 505 - 514 (1990).

〔8〕 Dunleavy P., Margetts H., Tinkler J. & Bastow S., *Digital Era Governance: IT Corporations, the State, and e-Government*, Oxford University Press, 2006, p.1 - 15.

〔9〕 Andrews L., *Public Administration, Public Leadership and the Construction of Public Value in the Age of the Algorithm and “Big Data”*, 97 (2) *Public Administration* 296 - 310 (2019).

〔10〕 Lipsky M., *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*, Russell Sage Foundation, 2010, p.5 - 6.

〔11〕 参见季卫东:《人工智能时代的法律议论》,载《法学研究》2019年第6期,第32页。

〔12〕 参见栗峥:《人工智能与事实认定》,载《法学研究》2020年第1期,第117页。

化应用做出价值判断的合法性基础何在?<sup>[13]</sup> 如何针对平台用工、电子交易等公共治理样态引入新的类型化规则?<sup>[14]</sup> 在智能社会中消除异化、促进人的自由全面发展所需要的法律保障是什么?<sup>[15]</sup> 上述法学论题在回应和补充公共治理学科相关研究的同时,也带来了方法论上的启示:我们究竟应当如何定性公共治理领域的自动化应用?管理效率的提升能在多大程度上正当化自动化应用对公众权利的侵蚀?

### (三) 自动化应用规制的法律聚焦

推进国家治理体系和治理能力现代化,是我国全面深化改革的总目标。个人信息一直是一种非稀缺的公共物品,<sup>[16]</sup>从效用增益的角度来看,公共治理领域的自动化应用符合从传统社会治理体系向科技社会治理体系转型的大趋势,因为自动化系统能以更高效率和更低成本实现“温故而知新”。在北京,城市管理综合行政执法局通过对海量投诉举报电话的机器分析,得出每年夏季晚间某固定时段占道经营违法行为上升的结论,结合指挥系统的位置信息,集中调配执法力量,实现科学执法;在上海,政务服务“一网通办”、城市运行“一网统管”,将“菜篮子、米袋子、水电煤气”等基本民生保障类的“获得感”数据作为治理数字化的重要抓手;在江苏,公安部门的犯罪预测系统已覆盖超过百种常见的违法犯罪行为,犯罪治理活动从事后打击转为事前预防;在河北,导诉机器人、手机立案平台、多元解纷结果预测系统为群众提供了像“新零售”“网络店”一样方便的纠纷解决服务渠道。

需要特别强调的是,各类自动化系统远非尽善尽美,也并非在所有领域都比自然人做得更好,甚至在多个应用场景中问题频现。<sup>[17]</sup> 例如,在缺乏科学论证的基础上,纽约州曾率先利用预警保护系统来鉴别具有“严重家暴倾向”的父母,并对“高危儿童”进行隔离保护。该系统不仅没有大幅降低虐童事件的发生概率,还造成了上万正常父母同其子女的被追分离,上线不到半年就引发众怒。美国宾夕法尼亚州用于提供替代性制裁矫正犯罪管理画像的坎帕斯风险评估系统(CAMPAS),被指控存在严重种族歧视的倾向:在该系统做出的犯罪高风险判定中,黑人数量竟然占到白人数量的两倍以上;在系统看来,即便是有多次前科和犯罪指控、经济状况欠佳的白人,也比接受过良好教育、无任何不良嗜好的黑人,犯罪风险更低。2019年4月,逾四万居民对密歇根州政府提起了集团诉讼,起因是政府用于检测骗补行为和欺诈行为的米达思预测系统(MIDAS)的出错率高达93%,致使数万申请人被无端惩戒。<sup>[18]</sup>

从世界范围来看,有关自动化应用的立法建议早已渐次成型,但多墨守自动化应用为自然人辅助之成规,致使“组织性无序”和“程序性滥用”屡禁不止,“这是一种极为寻常的经验,但也是全球化所指”。<sup>[19]</sup> 究其根源,碎片化的规则探寻容易缺乏对自动化应用特征共性的整体把握,造成个别小体制对抽象的大体制的掣肘和抵消。“公共事业”已经成为自动化应用的核心领域的美国

[13] See Veale M. & Brass I., *Administration by Algorithm? Public Management Meets Public Sector Machine Learning*, in Yeung K. ed., *Algorithmic Regulation*, Oxford University Press, 2019, p.121 - 149.

[14] 参见王天玉:《互联网平台用工的“类雇员”解释路径及其规范体系》,载《环球法律评论》2020年第3期,第85页。

[15] 参见孙伟平:《人工智能与人的“新异化”》,载《中国社会科学》2020年第12期,第119页。

[16] 参见刘迎霜:《大数据时代个人信息保护再思考——以大数据产业发展之公共福利为视角》,载《社会科学》2019年第3期,第100页。

[17] 参见袁康:《可信算法的法律规制》,载《东方法学》2021年第3期,第5页。

[18] 参见唐林垚:《算法应用的公共妨害及其治理路径》,载《北方法学》2020年第3期,第51页。

[19] [英] 卢恰诺·弗洛里迪:《信息伦理学》,薛平译,上海译文出版社2018年版,第427页。

和加拿大,仍“屡败屡战”,积极探寻框架式治理和协同治理共存并行的全周期评估框架。<sup>[20]</sup>即使试错成本高昂、技术性矫正手段相对匮乏,公共治理也早已不可回头地染上了数字时代的自动化底色,“人类正在从持续了千百年的物理实体社会跨入新兴的虚拟数字社会”,<sup>[21]</sup>涵盖法学理论、司法实践和制度规范的法律转型升级早已是“箭在弦上”。

## 二、自动化应用治理导向法律规则的构建机理

长期以来,面对公共治理的数字化、公式化和模型化,新公共治理理论学者和新公共服务理论学者倾向于笼统地将技术进步视为官僚主义从韦伯式的公共利益行政廉洁和公正模式向新公共治理范式转变的驱动力,而公共价值管理理论学者却不由自主地沉浸在对技术变革远期风险的悲观情绪之中,此间的价值观割裂,须在日新月异不同类型的自动化技术区别对待的基础上予以修复。

### (一) 自动化应用的三重维度区分

以人工智能颠覆传统公共治理手段为视角,自动化应用可以由远及近、由浅入深地划分为三个维度。第一维度是基于数据库编码的计算机自动化(1980年代开始),可以近乎完美地替代基层公务员实现“唯手熟尔”的重复性基础工作和流程性工作,例如,将电话或通信分类到正确的联系人、审阅数份公共文件后提炼出要点和关键词、根据以往的公共文书来修改生成更新的交易文书等。<sup>[22]</sup>重复、简单的工作日益被高效、廉价的机器人逐步取代,大大提升了劳动生产率。在第一维度,自动化决策对自然人的取代在于计算机具有更高数据处理能力的事实。

第二维度是基于机器学习的算法自动化(2010年代开始),其实质是作为编程理论典范的透过“规则本原的逻辑”,在范围不确定的环境中进行逻辑推导或模式识别,替代自然人实现“温故而知新”的基础性预测工作与规范性鉴别工作,但常因系统失灵多次给社会带来极高的“矫正成本”。在欧洲地区不止一次出现这样的情况:例如,法国银行间通用的客户资信考察系统,曾一度拒绝为单亲家庭成员提供住房贷款;英国铁路部门的智能审核系统把关不严,超额滥发多张特许经营执照。近年来,亚洲各国也相继出现了“人工智能热”,技术外部性效应日益凸显,例如,超过1/4的韩国公共公司在人员招聘中使用智能面试系统。该系统可以检测面试者恐惧、喜悦和回避的面部表情,并结合其措辞、眼神活动和声调变化等细节,对面试者进行抗压能力高低以及是否具备职场灵活性的判断。重压之下,各类面试练习班层出不穷,求职者支付天价补习费用,只为学习如何在机器面前“用眼睛微笑”。在第二维度,自动化应用导致的公权力扩张在于管理部门对机器学习技术能够做出准确而中立预测的不当信赖。

第三维度是多个机器学习模型相互交织、基于神经网络的超级自动化(2020年代开始),现已彻底突破利普斯基所认为的自动化决策系统的“智能上限”,<sup>[23]</sup>其获得识辨特定时期的法律和社会运行宏观规律、以超乎常人的洞察力来提供制度解决方案或进行价值判断的能力的奇点临近。在第三维度,自动化应用以管理部门为载体,表面上看起来是被公权力所“收编”,实则将“权威从

[20] 参见张欣:《算法影响评估制度的构建机理与中国方案》,载《法商研究》2021年第2期,第102页。

[21] 吴霁虹:《众创时代——互联网+、物联网时代企业创新完整解决方案》,中信出版社2015年版,第1页。

[22] 参见左卫民:《关于法律人工智能在中国运用前景的若干思考》,载《清华法学》2018年第2期,第104页。

[23] Lipsky, *supra* note [10], at 5-6.

个人转向由算法构成的网络”。<sup>[24]</sup> 人类将不再是自主的个体,而是一种生化机制集合体的组成部分,逐渐由自动化决策系统实时评价、定义、监测和指挥。

### (二) 公共治理理论的立场辨析

以上述三个维度作为基本线索,交叉比对相关研究后不难发现,新公共治理理论学者出于对第一维度机器自动化成功的认可、公共价值管理理论学者出于对第二维度算法自动化失灵的担忧,对第三维度超级自动化崛起的判断大异其趣——要么过于乐观、要么过于悲观。实际上,公共治理领域第二维度或第三维度的自动化应用,其前景定然介乎于二者之间;正如法学家们多倡导对技术变革的“理性冷漠”,跳出人工智能的应然与实然之争,转而就已经出现的问题和即将出现的问题寻求应对之策,提出了代码正义观和算法正义观等新型法权理念。<sup>[25]</sup> 在智能社会语境下,需要密切关注的是公共治理领域自动化应用已经产生的第二维度困境和必然面对的第三维度风险,借此反思人工智能促进公共治理体系和治理能力现代化所需的必要法律保障,进而为公共治理领域自动化应用的制度设计明确方向。

### (三) 传统法律应对的局限与失灵

对于自动化决策崛起所带来的法律挑战,世界各国主要采取了三种传统方式加以应对:以个人数据权为核心展开的隐私保护立法、以因果辨析和流程透明为导向的可解释性合规要求,以及覆盖数据处理全过程的完整履责链条。这种看似完备的“事前一事中一事后的”规制范式,在实践中面临可行性和可欲性难题:一方面,法律规则之于技术嬗变的更高站位,源于立法者制定规则时运用的思维和智识合理性,即立法者是否可以突破当前自然科学以及宗教伦理所框定的篱藩,超前布局应对技术必然带来的权力异化与权利式微;但就目前而言,立法者在自动化应用第一维度和第二维度已经完成的知识积累,还尚不足以应对第三维度自动化应用可能带来的法律挑战。另一方面,“问题驱动型”的立法应对模式并不自然催生令人向往的制度生态,即便具备技术可行性的法律规则,也未必长期符合相关主体的合理利益诉求;大众的狂热和盲从常因理性的回归得以修正,而追逐自身利益最大化导致的立场转变,或让原本就举步维艰的立法回应不堪一击。在这个意义上,既有“事前一事中一事后的”规制范式的脆弱性显而易见——“可欲性”沦为了为达到“可行性”而进行的说理手段,但“可行性”却难以带来“可欲性”的法律效果。具体而言,无论是以个人数据权为核心展开的隐私保护立法,以代码公开和流程透明为导向的可解释性合规要求,还是覆盖数据处理全过程的完整履责链条,都不具备严格意义上的“规则圆满性”,在实践中屡被架空。

其一,“告知—选择框架”广泛被各国立法者前置为数据处理全周期的“第一闸口”,使得隐私保护争议在司法审判实践中常以合同纠纷论处。在引人瞩目的“刷脸第一案”中,杭州市人民法院判处动物园不当采集公民生物特征信息,但胜诉的原告郭兵仅就合同利益损失和交通费补偿获得了1038元的赔偿。这样的判决结果显然无助于遏制被告动物园继续“店大欺客”,强推其他“技术入园”的自动化应用尝试。更重要的是,长篇累牍用户协议下的“告知”实际上是毫无诚意的“告知”,授权俘获机制下的“选择”也未必就是真正意义上的“选择”。<sup>[26]</sup> 现行法判定数据主体是否因充分“被告知”而做出真实意思表示“选择”的审查标准,只是技术方是否履行了最低信息披露要

[24] Harari Y. N., *Homo Deus: A Brief History of Tomorrow*, Random House, 2016, p.16.

[25] 参见马长山:《智能互联网时代的法律变革》,载《法学研究》2018年第4期,第20页。

[26] 姚佳:《知情同意原则抑或信赖授权原则——兼论数字时代的信用重建》,载《暨南学报(哲学社会科学版)》2020年第2期,第48—49页。

求,这实际上是以自动化应用技术主体的行为违法性评价取代了数据主体的法益保护评价,难以对数据主体的隐私侵害提供充分救济。

其二,强行代码公开将必然导致知识产权保护和投机式对抗方方面面的阻力,可解释性合规要求取代算法公开成为“退而求其次”的法律选择。立法者看到的是一个规定可解释性合规要求就足以让自动化应用可解释的世界,但正是这种不切实际的理想愿景导致了技术发展与法律滞后之间的巨大鸿沟。随着自动化应用从第一维度和第二维度逐渐向第三维度转换,机器学习模型和卷积神经网络的复杂程度呈指数级递变,就如同再强势的父母也无法完全预设子女的人生选择那般,一旦进入全自动更新换代的状态,自动化应用的造物主也可能退居次席,成为仅仅掌握底层代码的“局外人”,甚至对自动化应用的脱缰束手无策。<sup>〔27〕</sup>即便是在自动化应用复杂度相对有限的第一维度和第二维度,考虑到数据主体通常不具备相关理论知识背景,要就自动化应用的运作机理与决策方式做出准确且外行人也能理解的解释,往往徒劳无功。更有可能的情况是,信息和技术的双重优势地位必然带来监管套利的空间,解释和说理的方式最终将完全取决于经济和政治目标,致使数据主体在“半哄骗半威吓”下掉入任人宰割的陷阱。

其三,当自动化应用失灵损害当事人利益时,自动化应用的使用方未必是唯一应当承担责任的主体,技术开发方、数据处理方乃至看似中立的平台也是潜在的责任方。<sup>〔28〕</sup>但是,在实践中,每一层法律关系之间,可能存在着环环相扣的免责条款,出于对当事人意思自治的尊重,此类条款不宜随意被强行法排除。此外,理想化的完全履责链条高度依赖侵权责任,其本质是基于合同相对性的信义义务约定,通常以具有直接近因的法律关系为限。况且,一旦自动化应用进入第三维度,自动化应用所依托的非自然人载体将取得独立法律地位,<sup>〔29〕</sup>此时,传统的信义义务关系又可能陷入自指化悖论——依照数据主体要求提供专业化服务的自动化应用,究竟应当被视为对其载体享有支配权的运营方的代理人,还是为数据主体实现特定目的的代理人?在双方代理绝对禁止的情况下,基于信息、知识和能力的不对等而产生的信义义务极有可能出现“指向劈叉”的情况。总而言之,功能主义的立法范式,长期落后于技术突飞猛进带来的破坏式创新,夹杂着概念层面的模糊和失序,最终导致了个人权利的虚化。

### 三、公共治理领域自动化应用法律法规的全新面向

以自动化应用的维度区分为基础、以公共治理理论的立场辨析为坐标、以传统法律应对的失灵为警醒,公共治理领域自动化应用法律规制的完善方向,可以从宏观、中观和微观三个层面展开。

#### (一) 宏观技术共治：以程序透明取代代码透明

由于缺乏数据积累,管理部门经常需要同掌握大数据的私营企业或者平台合作,这种权力“外包”颠覆了传统的权力专属原则和正当程序原则。<sup>〔30〕</sup>具有准公权力的企业和平台横亘在政府与市场之间,公权力、私权力(准公权力)和私权利的多方博弈不仅没有给普罗大众带来实质性利好,反而极有可能造成高度技术化的管理部门和持“尚方宝剑”的平台寡头对身处信息弱势地位的普

〔27〕 参见李晟：《略论人工智能语境下的法律转型》，载《法学评论》2018年第1期，第98页。

〔28〕 参见胡凌：《平台视角中的人工智能法律责任》，载《交大法学》2019年第3期，第5页。

〔29〕 参见彭诚信、陈吉栋：《论人工智能体法律人格的考量要素》，载《当代法学》2019年第2期，第52页。

〔30〕 参见张凌寒：《风险防范下算法的监管路径研究》，载《交大法学》2018年第4期，第49—62页。

通民众的“轮番算计”。<sup>[31]</sup> 外溢出来的公权力难以监管,却迅速呈现出私有化、商品化和官僚化的异化趋势;生产数据和积累行为剩余,或将成为个体受众在“监视资本主义”下的唯一价值。自动化应用构建于海量个人数据的处理之上,而个人数据处理活动包含了数据主体、数据处理主体以及社会公众三重利益。从数据安全的角度来看,自动化应用规制的微妙性体现在,当数据处理主体以压倒性的技术和信息优势侵蚀数据主体的合法利益时,政府不宜直接介入数据主体和数据处理主体之间的利益分配,而只能以社会公众利益为导向对数据主体和数据处理主体之间的利益失衡进行力所能及的纠偏。依照计算风险规避效用的汉德公式,对管理部门“公共使命”和平台企业“社会责任”的笼统强调,远不如反其道而行之直接从代码入手、从源头解决问题具有更高法益。如何以权利保护和责任落实为核心、以“技术倒逼技术”为手段,让国家公权力与社会私权力所形成的在一定程度上“去中心化”的共谋治理机制更加贴近人民群众对技术变革的理想期盼,是未来立法工作的重中之重。

无论是哪个维度的自动化应用,算法程序的运行皆以代码为基础,因此,通过代码对自动化应用实施逆向管理在理论上可行。外观上,代码是具备技术中立性特征的系统语言,但这并不意味着代码的组合就不能进行价值负载。实际上,基于代码的自动化应用,或多或少包含了所有者、利用方、程序员以及数据集本身的私心与偏见。“法官对于法律用语不可附加任何意义,毋宁须以受法律及立法者拘束的方式,来发现法律的语义内容”,<sup>[32]</sup>但是,自动化应用的创造者和改造者,是清一色未经过任何系统法律训练的程序员,他们将鲜活的语义转化为毫无情感的代码,进一步完成机器模型的建立,使得自动化应用背后的计算机语言有时不能精确地转译规则的实际含义,甚至偶尔还会因为程序员的个人好恶、理解偏差使之与原本预期相去甚远。以美国公共福利系统的失策为例,美国科罗拉多州公务员曾经将“无家可归”错误解读为“行乞为生”,使得本应获得政府救济的流浪汉们被自动化审核系统拒之门外。<sup>[33]</sup> 在代码化过程中因语言理解和事实认定产生的巨大差异,以过失和疏忽为构成要件的一般侵权责任以及以信义义务为核心的事后审查机制都于事无补,只能通过更精确的代码语言来减少转译过程中的偏差和耗散。在宏观层面,与其让算法“透明”,不如让程序员的代码写作过程“透明化”;可以考虑以技术手段构建专属于程序员群体的阐释和说理场域,通过交流和相互探讨的方式将模糊的概念精确化,使得自动化应用底层代码中的逻辑连锁推论,可以更好地适用于意义涵摄,并通过完善中间步骤的方式更为准确地定义终局概念。

## (二) 中观价值位阶: 场景理论助推技术性规制

自动化应用对公共治理的介入,需要对各类信息进行数据化改造,在克服表达障碍后,主要依据概率计算而非演绎推理的基本逻辑来行使公共职权。在机器学习和神经网络的加持下,第二维度的算法自动化已经可以熟练进行法律规则适用的三段论推理乃至辩证推理,但当下第三维度的超级自动化即便通过模型累积递归、交叉话语博弈也难以做出公正的价值判断。理论上,价值判断的命题群可被结构性改造以实现整体客观性,但不同的应用场景和语境必然对应不同次序、不同权重的价值元规则。在技术共性规则之外,法律应就不同公共治理领域(例如卫生、医疗、教育、司法、社会保障等)所面对的技术社会化应用场景,分门别类地填补具体的、颗粒

[31] 参见黄尹旭、杨东:《超越传统市场力量: 超级平台何以垄断? ——社交平台的垄断源泉》,载《社会科学》2021年第9期,第100页。

[32] [德]卡尔·拉伦茨:《法学方法论》,陈爱娥译,商务印书馆2003年版,第23页。

[33] Pasquale F., *The Black Box Society*, Harvard University Press, 2015, p.165.



度更细的法律规则以实现公共治理的提质增效，<sup>[34]</sup>场景理论介入自动化应用规制，具有了合理性基础。

场景理论提倡以情境为导向考察自动化应用的风险水平，相应制定与风险程度相一致的规制策略和规则，亦即“贴合现实的法律应对”，其理论根源是传播学家梅罗维茨·约书亚(Meyrowitz Joshua)的情境理论(Contingency Theory)。<sup>[35]</sup> 海伦·尼森鲍姆(Helen Nissenbaum)等人将情境理论延伸至信息科学领域，发展出了基于“场景一贯性”的社群主义隐私产权理论。<sup>[36]</sup> 其提倡结合参与者、数据类型和流动样态三大关键要素，判断自动化应用的风险是否超过数据主体的合理预期。美国《消费者隐私权利保护法》与欧盟《通用数据保护条例》是将场景理论引入立法的典型代表，前者对“场景自治”进行了专章规定，认为个人隐私受侵害程度同因违反信息传播规则导致的场景破坏程度息息相关；后者要求自动化应用的利用方在事前考察数据处理活动的“性质、范围、语境与目的”，并主动向监管者提交潜在风险的应对措施。近年来，场景理论还延伸到了人脸识别技术的法律规制当中，为网格和网络作为治理之理想类型的二分提供了理论基础。<sup>[37]</sup>

不过，上述立法实践，均未跳脱出传统的自动化应用的规制思维，不仅未能真正实现场景理论介入自动化应用规制本应带来的监管效率提升，甚至起到了同预期完全相反的效果。《消费者隐私权利保护法》所规定的“场景自治”主要作用于传统的“告知—选择”框架，在实践中反而导致了自动化应用“目的拘束原则”的扩大解释，<sup>[38]</sup>扩大了某些自动化应用和个人数据处理活动的合法性基础，使得自动化应用的流弊变本加厉。《通用数据保护条例》的相关规定，本来旨在明确自动化应用利用方的风险管理责任，以区分情境适当增强或降低外部执法威慑的手段，保障自动化应用的风险“可防可控”；但是，过于教条的法律规定带来了“执行难”的问题，欧盟各成员国依照《通用数据保护条例》制定本国信息保护法时，只能着眼于依照情境适当强化或削弱自动化应用利用方的信息披露义务，并同步赋予数据主体是否继续授权的权利，这就又回到了毫无新意的“告知—选择”框架上。

回到自动化应用的本质，无论是第一维度的自然人决策辅助，第二维度的人机交互决策，还是第三维度的高等自动化，自动化应用必然会因为运用主体的不同、针对对象的不同以及所涉及问题的不同具有大相径庭的性质，就此而论，场景理论介入自动化应用规制的根本目的，应当是根据不同场景确定不同的规制方式。由于自动化应用的风险与日俱增且难以被当事人察觉，不同的风险场景“依其位阶衡量轻重”的规则设计需以系统化构筑技术风险管理体系为目标，将主要风险管理责任配置给自动化应用的使用方、次要风险管理责任配置给相关管理部门，实现齐心共治、多措并举的合力局面。公共价值管理理论学者虽然对自动化应用的评价褒贬不一，但大多都认可辩证地看待自动化应用的社会效益。从场景理论的视角来看，不同的场景，一定对应着不同的自动化应用的“社会最佳活动水平”，场景化法律规则，正是以宽严相济的动态规范方式，促使自动化应用无限趋近符合现实情况的“社会最佳活动水平”，借此来实现公共效益的“帕累托最优”。在这个意

[34] 参见唐林垚：《公共卫生领域算法治理的实现途径及法律保障》，载《法学评论》2021年第3期，第95—107页。

[35] See Meyrowitz J., *No Sense of Place: The Impact of Electronic Media on Social Behavior*, Oxford University Press, 1986, p.1-12.

[36] See Nissenbaum H., *Privacy as Contextual Integrity*, 79 *Washington Law Review* 119, 119 (2004).

[37] 参见胡凌：《刷脸：身份制度、个人信息与法律规制》，载《法学家》2021年第2期，第41页。

[38] 参见郭春镇：《对“数据治理”的治理——从“文明码”治理现象谈起》，载《法律科学》2021年第1期，第58页。

义上,法律是技术领域国家宏观调控的重要手段,动态的风险控制机制是确保自动化应用“技术向善”不可或缺的法律保障。

### (三) 微观权利保障:避免自然人决策名存实亡

目前,公共治理领域的自动化应用处于从第二维度向第三维度的过渡阶段,具体表现在:自动化系统占有、处理承载着社会利益的大数据,从单纯的技术工具逐步升级为通过嵌入社会权力结构发挥作用的“准公权力”。随之而来的格局很可能是,自动化系统的应用范围越来越广,专属于人类的决策空间和社会沟通被渐次压缩:公共治理和服务中的各个环节越来越多地被算法事先预设,通过大数据反映出来的趋势和评价会限制和迷惑自然人的推理和判断,也必然会削弱管理部门的自由裁量权。本来希望通过技术来确保公共治理的中立和客观,结果却很可能完全抹除自然人独有的通时合变的人情味,让冷漠、机械、独裁的自动化决策大行其道,<sup>[39]</sup>同新公共治理理论所倡导的“公民权理念观”和公共价值管理理论所强调的“协商机制”都背道而驰。反过来,自动化系统的专断又将逐渐影响大众群体的行为规范,在人机相互适应的过程中不断强化“算法绝对主义”。职是之故,法律首要任务,从社会整体利益的角度防范公共治理领域中的自动化霸权,亦即,从宏观层面防止自然人决策名存实亡。

早在自动化应用的第一维度时期,人们就已经意识到,只有在决定自身利益的关键事务上享有积极参与的权利,才能在知识、经验和道德上维持一以贯之的卓越。<sup>[40]</sup>出于对机器超越人类的无形恐惧,欧盟立法者在1995年的欧盟《数据保护指令》第15条率先规定了“脱离自动化决策权”的雏形:“数据主体享有免受基于用户侧写的完全自动化决策的权利。”在此后的20年中,虽然自动化应用已从第一维度向第二维度跃迁,但“脱离自动化决策权”在各成员国法院的审判实践中几乎从未发挥任何作用,<sup>[41]</sup>被戏称为欧盟《数据保护指令》中的“二等权利”。即便如此,2018年的欧盟《通用数据保护条例》第22条概括继承了《数据保护指令》第15条,还进一步拓展了“脱离自动化决策权”的行权对象和适用范围,赋予相关人完全脱离自动化应用对其做出的具有法律后果的决策的权利,足见“脱离自动化决策权”在自动化应用从第二维度向第三维度跃迁时代背景下的特殊意义。

为了使“脱离自动化决策权”脱离“二等权利”之称谓,欧盟各成员国通常将《通用数据保护条例》第22条作为赋权性条款解读,但由于普遍缺乏对第三维度自动化应用的准确理解,原本就过于抽象的“脱离自动化决策权”在实践中广泛存在与“更正权”竞合之表象、与“获解释权”补强之曲解以及与“被遗忘权”混同之误判。界定“脱离自动化决策权”的内涵应当回归自动化应用的维度划分,“在基于数据库编码的计算机自动化场景为脱离数据采集,在基于机器学习的算法自动化场景为脱离用户侧写,在基于神经网络的超级自动化场景为脱离平台交互”。<sup>[42]</sup>在“更正权”“获解释权”和“被遗忘权”之外,公共治理学者还提出了“陈情权”“离线权”和“注意力保护权”等概念,这些形形色色的具体权利,既是“脱离自动化决策权”的细化表达,也是在自动化应用大行其道的背景下,以更高程度的人类干预取代较低程度的人工介入的理性呼吁。总之,自动化应用的维度跃

[39] 参见许中缘:《论智能机器人的工具性人格》,载《法学评论》2018年第5期,第153页。

[40] 参见唐林垚:《〈个人信息保护法〉语境下“免受算法支配权”的实现路径与内涵辨析》,载《湖北社会科学》2021年第3期,第135页。

[41] 虽然“脱离自动化决策之诉”客观存在,但是双方的争论焦点是决策是否完全由自动化应用做出,而非“脱离自动化决策权”的行权方式与行权效果。See Cour de Cassation, Chambre Criminelle, *Audience publique No de pourvoi*, Publié au bulletin, 1998, p.97.

[42] 唐林垚:《“脱离算法自动化决策权”的虚幻承诺》,载《东方法学》2020年第6期,第18页。

迁，亟需微观层面的新型个体权利予以应对；坚持问题导向和目标导向、围绕“促进人的全面发展”构建权利簇，正是“以良法促进发展、保障善治”的题中之意。

#### 四、结 语

公共治理领域的自动化应用，完全不同于私人部门的自动化应用。传统侵权责任规定和信义义务创设，足以应对大多数私人领域的自动化妨害，基本能在合同相对方的信义义务框架内解决问题。但是，公共治理领域的自动化失控，却极有可能给整个社会带来类似环境污染一样的“公共妨害”，其本质是技术外部性扩散所引发的社会成本向广大弱势群体的不合理转嫁，若不加以控制，定然会随着公共数据体系的污染日趋积重难返；以当代公共治理理论已确立的规范价值作为对标，跨学科对公共治理领域自动化应用的动态法律思考迫在眉睫。

“法者，治之端也。”法律的终极原因是社会福利，技术亦然。自动化应用介入现代化治理的持续与有序，有待整合社会各方力量，形成跨领域的、立体的协同合作体系来推动技术和法律的共同演进。<sup>〔43〕</sup> 在这个协同体系中，社会公众利益拥有压倒性考量，而权力制约和权利保护应共存并行，平衡自动化应用载体的内部治理系统和外部监察制度，既是实现以“智”提“质”治理的重要基石，也是用新发展理念引领发展全局的内在要求。如此这般，方能真正促进全面依法治国的制度优势与势在必行的技术治理之间的交流与相互增益。

---

**Abstract** The link between technological leaps and governance improvements is commonly acknowledged and assumed by the general public, and it has long been captured in the current *digital China policy* and *intelligent society practice*, yet the perennial lack of applicable rules has created abundant room for regulatory arbitrage. In a worldwide sense, automated applications in public administration have gradually shifted from the data-encoding-based computational automation to machine-learning-based algorithmic automation and then to neural-network-based super automation, and previous bottom-up occasional *auxiliary/reference* and partial *deployment/application* had given way to top-down social *embedded/integration*, leading to a structural transformation in between the public authority, private power and personal rights. Accordingly, the internationally accepted regulatory paradigm of personal privacy protection (*ex ant consent*), algorithmic interpretability (*interim monitoring*) and full accountability chain (*ex post accountability*) should be systematically reconstructed in a trinity manner of macro-technical co-governance, meso value spectrum and micro right protection.

**Keywords** Public Governance, Artificial Intelligence, Automated Decision, Algorithmic Regulation

---

（责任编辑：曹博）

---

〔43〕 参见彭中礼：《智慧法治：国家治理能力现代化的时代宣言》，载《法学论坛》2020年第3期，第29页。