

doi:10.13582/j.cnki.1672-7835.2018.06.012

人工智能在行政治理中的作用及其法律控制^①

宋华琳,孟李冕

(南开大学 法学院,天津 300350)

摘要:目前对人工智能的定义尚缺乏共识,但可按发展程度分为弱人工智能、通用人工智能和超人工智能。在行政治理中,人工智能可以参与行政立法过程、政策形成和行政决定,推进行政治理智能化。但同时亦需认识到人工智能治理适用范围的局限。人工智能治理中不仅存在诸如歧视、安全和隐私等风险,甚至可能导致政府控制能力减弱。因此,需对人工智能治理加以法律控制,施加正当程序约束,以实现其良性运作。

关键词:人工智能;行政法;正当程序;说明理由

中图分类号:D912.1 **文献标志码:**A **文章编号:**1672-7835(2018)06-0082-09

一 人工智能概说

1945年,冯·诺依曼首次提出“奇点”,即机器达到人类智力水平的点^①。1950年,艾伦·图灵定义了被广泛解读为人工智能“成人仪式”的图灵测试^②。1956年夏,麦卡锡、闵斯基、西蒙、纽厄尔、香农、索洛莫诺夫等人召开了为期一个月的达特茅斯会议,讨论主题为用机器来模仿人类学习及其他方面的智能,这标志着人工智能的正式诞生^③。

(一)人工智能的概念与特征

何为人工智能,目前尚缺乏共识。有学者在研究中,将人工智能定义为,机器能够完成人类需要智能才能够完成的任务^④。此定义避开对

“智能”这一概念的解释。定义“人工智能”的困难不在于“人工”的概念,而在于“智能”的概念模糊^⑤。作为人工智能的领军人物,斯坦福大学约翰·麦卡锡教授认为,“任何关于智能的定义都依赖人的智能”,因为“目前为止,还不能总括出怎样的计算机程序可被称为智能”^⑥。而“智能”的定义又与人类本身就难以定义的特征相关联,这涉及模式识别、分类、学习、记忆、归纳、类比、泛化、联想、规划、优化、创新、演绎推理、问题求解、语言处理等^⑦。据此,或可将智能视为主体在纷繁复杂的环境中实现目标的能力^⑧。

根据发展程度的强弱,可将人工智能分为弱

① 收稿日期:2018-02-25

基金项目:国家社会科学基金项目(17BFX004);南开大学跨学科建设项目(63185030)

作者简介:宋华琳(1977-),男,河北黄骅人,博士,教授,博士生导师,主要从事宪法学与行政法学研究。

②参见(美)雷·库兹韦尔:《奇点临近》,李庆诚、董振华、田源译,机械工业出版社2011年版,第10页。

③在图灵测试中,一台计算机将试图让一个人认为它是人类。判定者需要在的一组人类选手中找出计算机。电脑和人类选手在物理上都和评委相分离,他们之间只通过文字来交流。如果有超过30%的测试者不能做出正确选择,那么这台计算机就通过了测试,并被认为具有人类智能。

④集智俱乐部:《漫谈人工智能》,人民邮电出版社2015年版,第102页。

⑤Matthew U. Scherer. “Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies”. *Harv. J. L. & Tech.*, 2016 (29): 362.

⑥Matthew U. Scherer. “Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies”. *Harv. J. L. & Tech.*, 2016 (29): 359.

⑦转引自 Matthew U. Scherer. “Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies”. *Harv. J. L. & Tech.*, 2016 (29): 359-360.

⑧参见集智俱乐部编著:《科学的极致:漫谈人工智能》,人民邮电出版社2015年版,第103页。

⑨参见集智俱乐部编著:《科学的极致:漫谈人工智能》,人民邮电出版社2015年版,第118页。

人工智能(ANI)^①、通用人工智能(AGI)^②和超人工智能(ASI)^③。通用人工智能被认为拥有“合理的自我理解和自控能力,能够在各种环境中解决复杂问题,并能学会解决其在设计时不知道的新问题”^④。目前还不存在通用人工智能和超人工智能。

(二) 需要厘清的要点

需厘清的要点是:第一,人工智能不单单是一项技术。如果一个系统拥有语音识别、图像识别、检索、自然语言处理、机器翻译、机器学习中的一个或几个能力,那就认为它拥有一定的人工智能^⑤。

第二,人工智能不能取代人,人工智能是将人的特征、技能加以片段化、分割化和分散化。或许可以将人工智能视为“特定目的的人”(special-purpose human beings)^⑥,但人工智能只能在某个特定领域、特定方面、特定治理环节上发挥作用,它至多可以“部分”取代人。一个完全脱离人类的人工智能程序并没有独立的意义。当务之急是开发人机协同增强智能,加强人机混合智能的基础理论研究,提高人机交互能力^⑦。

二 人工智能在行政治理中的作用

人工智能推动行政治理智能化,在行政规则制定、政策形成和行政治理过程中,人工智能都有可能发挥相应的作用。

(一) 人工智能在行政规则制定中的作用

在行政规则制定过程中,可借助人智能系统来检索信息,听取公众意见,改进规则的审查与备案工作。

1. 信息检索的智能化

在行政规则制定过程中,有必要进行法律规

范的检索、搜集、阅读、整理和分析。目前已有商业化的“北大法宝”系统,原国务院法制办公室的“中国政府法制信息网”^⑧上的“法律法规库”也有检索功能,但对此只能进行最为原始的手工检索。随着法律信息检索的智能化进程,基于自然语言处理和深度机器学习的语义检索和法律问答正在改造传统的法律检索服务。让机器去评估法律规范文本时,去逐步适应文本;通过E-discovery,进行自动化的关键检索,寻找确切的词汇或文本;通过自然语言处理,来分析规范的相关性,这有助于为行政规则制定中的政策学习,为规则的移植和借鉴,提供必要的信息和知识基础。

2. 听取意见的智能化

古语云:“兼听则明,偏听则暗。”听取意见有助于改进公众对行政规则制定的参与,能更为全面、均衡地搜集规范制定所需信息,使得在规则形成过程中更好考虑各方观点,进而提高规则质量,改进规则的可接受性^⑨。行政机关纵然在行政规则制定中,吸纳了公众提出的意见,一般也殊少反馈。对公众意见的反馈总体上也较为概括,很少阐述具体的理由依据和考虑的因素,更缺少与公众的沟通交流^⑩。

人工智能的引入,有助于推进意见听取和意见回应的智能化。2018年2月2日,上海市立法信息平台投入使用,此平台对规则制定过程全程留痕,实现过程控制,还可以通过“一键推送”电子流程改变草案征求意见的繁冗,提高及时性和反馈效率^⑪。公众未来有望借助人智能系统,提交对行政规则的意见。人工智能系统可以对公众意见进行收集、筛选和归并,运用自然语言处理

①弱人工智能(Artificial Narrow Intelligence,简称ANI)不真正拥有智能,也不会有自我意识,如AlphaGo围棋程序。

②通用人工智能(Artificial General Intelligence,简称AGI),也称为强人工智能。

③目前对超人工智能(Artificial Superintelligence,简称ASI)的理解是,不仅具备人类的某些能力,还拥有自我意识和情感等。

④Ben Goertzel, Cassio Pennachin. *Artificial General Intelligence*, Springer-Verlag New York, Inc. VI (2007).

⑤参见腾讯研究院,中国信通院互联网法律研究中心,腾讯AI Lab,腾讯开放平台:《人工智能》,中国人民大学出版社2017年版,第20页。

⑥Jack M. Balkin. “The Path of Robotics Law”. *Calif. L. Rev.*, 2015 (6): 59.

⑦参见《国务院关于印发新一代人工智能发展规划的通知》(国发〔2017〕35号,2017年7月8日)。

⑧<http://www.chinalaw.gov.cn/col/col4/index.html>.

⑨参见P. P. Craig. *Administrative Law*, 2nd edition, Sweet & Maxwell, 1989, p.185.

⑩参见曾娜:《电子化行政立法研究》,知识产权出版社2017年版,第89页。

⑪参见罗培新:《上海率先建成政府立法信息平台,打通运用法制大数据的“任督二脉”》,《上观新闻》,2018年2月7日。

程序来阅读和总结意见,甚至帮助对意见加以回应^①。

3. 审查阶段的智能化

在政府法制机构人力资源极其有限的情况下,无论是对行政法规和规章的事前审查,还是对行政规范性文件的备案审查,都给法制机构带来沉重的负担。2014年9月起,天津市人大常委会法制工作委员会开始使用北大法宝智能立法支持平台^②。通过立法支持平台,可以把待审查文件、“敏感词词库”、法律数据库三者相对照,标出需要引起注意的“敏感词汇”,同时,自动呈现与这些“敏感词汇”相关的法律规定,立法工作者再对这些信息进行比对考量^{③④}。这是在行政立法中应用人工智能系统的先声,未来通过人工智能系统的改进,人机交融智能的增强,将有助于推进规则制定审查工作的智能化,提高审查质量和效率。

(二) 人工智能在政策形成中的作用

传统的行政法学者也许永远无法告诉我们,什么才是一个好政策,什么才是美好的政策蓝图^⑤。通过开发适用于政府服务与决策的人工智能平台,研制面向开放环境的人工智能引擎,有助于改进公众对公共政策形成的参与,改进政策形成的内容,优化和改进政策的备选方案,有助于行政机关制定出更好的政策。

首先,人工智能系统有助于改进公众对公共政策形成的参与。正如英国学者哈洛和罗林斯所说:“行政官员和政客都不是从法律而是从政策中寻求他们的授权的。换句话说,他们是以政策为中心的。积极地看,行政官员视法律为一套挂衣钩,将政策挂在上面;消极地看,法律也许是在政策能够得以贯彻实施之前要跨越的一连串栏

架,法律在此意义上就是一种控制。”^⑥人工智能系统的建构,有助于畅通政府与公众的交互渠道,使公众更好反映对公共政策的诉求,政府也可以更好地分析公众对公共政策的诉求,使政策更具灵活性、一贯性,让政策更为平等地对待行政相对人。

其次,人工智能系统可以通过自动化模型的建构,来模拟相关的情境和政策备选方案,以及相应情境和备选方案下的应答,进而去评估不同备选方案可能带来的成本,给经济社会可能产生的影响,以及可能带来的预期效果,这有助于推进政策形成过程中的影响评估,改进政策形成的内容。人工智能系统可以通过建立模型,通过机器学习,通过对大型数据集的分析,输入不同的政策选择,来判定未来可能的不同结果,进而有助于在不同政策方案之间进行比较与选择。

再次,通过引入人工智能系统,来开展数据分析或行为预测,有助于研判复杂政策问题,评估风险,进行应急处置,开展战略决策,为动态制定和修订政策提供必要的参考。例如,可以利用“智慧信用城市政府信用管理平台”,进行智能信用分析;可以将模拟算法应用于招标投标、投资融资、日常监管、行业分析等应用场景;可根据政务服务大数据,自动形成涵盖市场主体存量、资金流入以及市场活跃度等信息的走势分析,帮助政府制定有效的产业调控政策^⑦。学者已经证明,机器学习算法可以用来预测财务报表欺诈、选举舞弊,甚至非法捕鱼行为^⑧。通过行为预测,政府可以制定更为有的放矢、量体裁衣的政策。

^①Cary Coglianese & David Lehr. "Regulating by Robot: Administrative Decision Making in the Machine-Learning Era". *Georgetown L. J.*, 2017(105): 1172.

^②高绍林:《人工智能如何辅助地方立法》,2018年1月14日, <http://lzzfyjy.cupl.edu.cn/info/1038/8016.htm>, (最后访问时间:2018年3月21日)。

^③高绍林:《人工智能如何辅助地方立法》,2018年1月14日, <http://lzzfyjy.cupl.edu.cn/info/1038/8016.htm>, (最后访问时间:2018年3月21日)。

^④参见罗培新:《上海率先建成政府立法信息平台,打通运用法制大数据的“任督二脉”》,《上观新闻》,2018年2月7日。

^⑤参见(美)约瑟夫·P. 托梅恩,西德尼·A. 夏皮罗:《分析政府规制》,苏苗罕译,载方流芳主编:《法大评论(第三卷)》,中国政法大学出版社2004年版。

^⑥(英)卡罗尔·哈洛、理查德·罗林斯:《法律与行政(上卷)》,杨伟东、李凌波、石红心、晏坤译,商务印书馆2004年版,第165页。

^⑦参见艾琳,王刚:《重塑面向公众的政务服务》,社会科学文献出版社2015年版,第245-298页。

^⑧转引自 Cary Coglianese, David Lehr. "Regulating by Robot: Administrative Decision Making in the Machine-Learning Era". *Georgetown L. J.*, 2017(105): 1161.

(三)人工智能在行政治理中的作用

行政主体在行政过程中,可进行行政检查,进行行政执法,发布行政命令。在这些环节中,人工智能系统均有可能发挥作用,来辅助确定行政执法资源的配置,在一定程度上做出自动化的行政决定。

1.在行政检查中的应用

行政执法过程中,存在执法资源稀缺,执法人员工作负荷重,执法资源与执法任务不匹配的情况。通过人工智能系统的引入,有助于分析已有的行政执法案例,已有的违法案例和情形,判断甄别出相对更高的风险。通过引入人工智能系统,建构智能化监测平台和智能化预警系统,对数据集进行处理,通过预测违法行为发生的概率及分布^①,有助于选择性地将稀缺行政执法资源配置在风险更高的地方,针对更有可能发生的违法行为、违法主体,加强行政检查频率,提高行政监管力度,确保“好钢用在刀刃上”。

在实践中,美国纽约建立了市长办公室的数据分析系统,其中包括与城市消防部门合作利用机器学习来决定应派遣建筑检查员的地点^②。在我国,上海市食品药品监督管理局于 2018 年初通过网络订餐平台的数据分析,甄别出第一批存在较高食品安全风险的“网红”餐厅^③,开展线下检查,发现部分此类餐厅存在管理状况差乃至无证经营的情况^④。人工智能的介入,有助于改进行政检查的针对性。

2.在行政执法中的应用

通过人工智能系统、行政执法监测平台、信息收集系统的建构,行政机关引入人工智能系统,引入视频监控、遥感、无人机、无人驾驶汽车、机器人等,有可能更好地查明事实,查明违反人工智能系

统所设定规则和指令的行为,进而确证违法情节、违法行为及其幅度。以道路交通违章的自动识别为例,可对监控设备采集的数据进行智能识别和分析,按照技术标准形成车辆违法数据,继而经交警部门审核后,对违法车辆进行处理,以违法处理单的形式将处罚通知送至违法者。最近,上海市食品药品监督管理局探索运用人工智能图像识别技术,自动识别食品监控拍摄的违法行为,加大了对食品安全违法行为的威慑力^⑤。

在未来,如能将梳理好的行政执法依据、行政裁量基准、行政执法流程、行政执法文书整合入人工智能系统中,则有可能在人工智能系统的辅助下,做出自动化的行政执法决定^⑥。所谓自动化,是指“自行运作,很少需要或不需人的控制”^⑦。行政机关甚至还可以借助人工智能系统,对行政相对人发出行政指导,告知行政相对人应如何遵守法律;告知行政相对人是否存在轻微违法现象,对其进行劝诫,责令其改正。这有助于削减行政执法的成本,向行政相对人发布精准化、情景化的指令^⑧。

3.在行政命令中的应用

行政命令是行政主体在行政管理过程中依法作出的、要求相对人为一定行为或不为一行为的单方意思表示^⑨。例如,交通警察在城市交通路口以形体动作指挥交通的行为,即为较为典型的动作命令^⑩。人工智能使得行政命令有了智能化的可能。2018 年,上海进行智能信号灯系统试点,该系统能通过多渠道自行感知和收集相关交通数据,再进行分析研判,给出信号控制和交通组织的最佳方案,其初步成果之一就是优化信号灯配时,缓解交通拥堵。该智能系统在一定程度上替代了以形体动作指挥交通的交警,甚至比交警的指挥更加科学合理。

①徐骏,苗运卫:《智能辅助:破解环境刑事司法证据因局之匙》,《常州大学学报(社会科学版)》2018 年第 2 期。

②转引自 Cary Coglianese & David Lehr. “Regulating by Robot: Administrative Decision Making in the Machine-Learning Era”. *Georgetown L. J.*, 2017 (105): 1161.

③“网红”餐厅是指开店速度极快、人气火爆、有大量外卖的餐厅。

④https://www.thepaper.cn/newsDetail_forward_2089280, 澎湃新闻,最后访问时间 2018 年 4 月 26 日。

⑤https://www.thepaper.cn/newsDetail_forward_2089280, 澎湃新闻,最后访问时间 2018 年 4 月 26 日。

⑥袁雪石:《建构“互联网+”行政执法的新生态》,《行政管理改革》2016 年第 3 期。

⑦Thomas A. Smith. “From Law to Automation”. *Criterion J. on Innovation*, 2016 (1): 535.

⑧Anthony Casey & Anthony Niblett. “Self-Driving Laws”. *University of Toronto Law Journal*, 2016 (66): 432-433.

⑨闫尔宝:《行政行为的性质界定与实务》,法律出版社 2010 年版,第 86 页。

⑩闫尔宝:《行政行为的性质界定与实务》,法律出版社 2010 年版,第 90 页。

(四) 人工智能治理适用范围的局限性

在目前的行政活动中,人工智能的引入相对更多集中于事实认定环节,相对更多适用于羁束行为而非裁量行为。在关注人工智能在行政治理中的作用时,亦应冷静审视其局限性。

首先,在法律要件中,需要对不确定法律概念进行认知判断。诸如公共利益、重要依据、可靠性、能力、必要性、特殊困难等不确定法律概念首先是一个主观认识问题,个案适用离不开权衡,有时甚至需要预测未来^①。只有在谨慎、全面考虑、评估和权衡各种观点的情况下,才可能做到这一点^②。不确定法律概念会影响法律要件中事实认定的准确性,人工智能对此难以个案权衡,有可能加剧事实认定的不准确。由于不确定法律概念亦存在于羁束行政之中,因此人工智能并非完全适用于羁束行政。

其次,在法律效果中,存在决定是否采取法定措施和采取何种法定措施的“裁量”,即“决定裁量”和“选择裁量”。行政机关有选择“是否”在行政决定过程引入人工智能技术的裁量权,还可以选择“如何”利用人工智能方式作出行政决定^③。当行政机关做出裁量决定时,需要综合权衡案件的不同因素,而人工智能系统在行政活动中的引入,可能使行政执法人员将复杂现实空间中遭遇到的案件案情,归约为输入计算机的若干指标或参数,然后利用人工智能系统像“自动售货机”式输出决定。自动化的行政决定可能会给行政相对人带来便利,但在特定情形下,这通过适用于简化的,更容易转化为代码的规则,而忽视了问题的复杂性^④。

无论是羁束行政行为,还是裁量行政行为,人工智能皆非完全适用。因此,应根据不同的具体

行政领域,来决定人工智能系统在行政过程中适用的阶段和范围。在羁束行政行为中,人工智能系统需关注不确定法律概念。在需要个案裁量的行政领域,人工智能系统只能发挥辅助性的作用,它应更多被适用于行政决定的准备过程,而非实质性的决定制作过程^⑤。

三 人工智能治理的风险

人工智能治理存在歧视、安全和隐私等风险,有可能减弱政府控制能力。理性考量将人工智能引入行政治理带来的可能法律风险,有助于更为理性审视人工智能,并给出相应的制度规范之道。

(一) 歧视

1. 算法歧视

人工智能算法本质上是“以数学方式或者计算机代码表达的意见”。

算法的公平性存在以下困境:第一,公平本身是一个不确定的概念。很难将公平量化、形式化,并转译为操作性的算法^⑥。第二,算法不是一种完全价值中立的活动,算法难免蕴含价值判断,有时算法会不当地将特定对象、特定项目、特定风险给予不当的权重^⑦。第三,由于设计者的偏见、信息不充分等因素,设计者会有意或无意在算法中植入不当目的。例如,由于人工智能算法的设计、目的、成功标准、数据使用等是设计者、开发者的主观选择,他们有可能将自己偏见嵌入算法系统^⑧。第四,算法是预测性的模型,算法过于依赖规则和数字,而无法对规则和数字之外的因素加以充分考量。第五,对于一个自主学习的人工智能治理系统而言,算法是深度学习的产物,通过学习而不断寻求自身的调适,算法结果因此也存在

①(德)哈特穆特·毛雷尔:《行政法学总论》,高家伟译,法律出版社2000年版,第133-134页。

②(德)哈特穆特·毛雷尔:《行政法学总论》,高家伟译,法律出版社2000年版,第133-134页。

③参见谢硕骏:《行政机关以电子方式做出行政处分:以作成程序之法律问题为中心》,《台大法学论丛》2018年第4期,第1793页。

④Danielle Keats Citron. "Technological Due Process". *Wash. U. L. Rev.*, 2008(85): 1297.

⑤参见(德)汉斯·沃尔夫,奥托·巴霍夫、罗尔夫·施托贝尔:《行政法》(第1卷),高家伟译,商务印书馆2002年版,第43页。

⑥参见腾讯研究院,中国信通院互联网法律研究中心,腾讯AI Lab,腾讯开放平台:《人工智能》,中国人民大学出版社2017年版,第242页。

⑦参见腾讯研究院,中国信通院互联网法律研究中心,腾讯AI Lab,腾讯开放平台:《人工智能》,中国人民大学出版社2017年版,第242页。

⑧(英)玛格丽特·博登:《AI:人工智能的本质与未来》,孙诗惠译,中国人民大学出版社2017年版,第244页。

不可预见的可能^①。结果的不可预见性会在一定程度上影响算法的公平性。

此外,算法存在三种形式的不透明性:第一,因商业秘密或国家秘密而产生的不透明性;第二,因公众不懂代码以及算法等专业技术产生的不透明性;第三,从机器学习算法的特征中产生的不透明性,这是熟知的“黑箱难题”,可能连设计者都不知道算法如何决策^②。算法不透明也会导致算法歧视。

2. 数据歧视

对于以机器学习为代表的人工智能治理系统而言,控制数据就控制了算法。掌握数据越多,算法越精准。当数据不全面、不准确时,也会出现歧视^③。

首先,数据可能缺少均衡性和代表性。数据的遗漏、偏差或过时可能会造成数据样本的代表性缺失,代表性缺失的数据样本训练出的算法极可能存在歧视。例如,谷歌公司的图片软件曾错将黑人的照片标记为“大猩猩”,其原因可归结为缺乏足够代表性的数据样本,所以软件才不能识别出黑人。

其次,数据可能存在偏见。如果数据存在偏见,由偏见数据训练的算法则可能存在歧视,这种偏见可能是数据自身具有的,也可能是行政主体对数据的不当处理所致。在公共安全智能化监测预警和风险预防过程中,都有可能存在数据偏见。

所以,当在特定情况下使用预测性算法时,应在一定程度上对基础性的数据来源、数据信息、数据利用方式加以公开,且能对不准确的数据加以

修正。

(二) 安全

人工智能治理的安全风险涉及网络安全(security)和人身安全(safety)。

1. 网络安全

根据《中华人民共和国网络安全法》第 76 条的规定,网络安全“是指通过采取必要措施,防范对网络的攻击、侵入、干扰、破坏和非法使用以及意外事故,使网络处于稳定可靠运行的状态,以及保障网络数据的完整性、保密性、可用性的能力”。应保障人工智能系统的安全性,算法不被黑客控制,系统和算法不能被被黑客攻击或改变^④。“友好人工智能”是系统安全所追求的目标,要求算法易于理解、值得信赖、稳健性强^⑤,且不易被黑客操纵^⑥。

2. 人身安全

人工智能治理中的人身安全,即确保人工智能技术不会伤害人类。人工智能的失控会给人身安全造成威胁。阿西莫夫的机器人三原则、德国自动驾驶的 20 条伦理准则^⑦和阿西洛马人工智能原则^⑧(Asilomar AI Principles)皆是出于保护人身安全方面的努力。

人工智能的安全问题与其控制问题密不可分。人工智能系统包含了很多不可控制的机制,如文件损害、输入设备受损等故障、安全系统漏洞、计算机系统相比于人类更快的反应时间、有瑕疵的编程等等^⑨。人工智能系统的自主性也带来了控制难题。在人工智能系统存在一定程度自主

^①Matthew U. Scherer. “Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies”. *Harvard Journal of Law & Technology*, 2016 (29): 367.

^②Jenna Burrell. “How the Machine ‘Thinks’ Understanding Opacity in Machine Learning Algorithms”. *Big Data & Society*, 2015 (3): 1.

^③*AI in the UK: ready, willing and able*, <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>, p.41.

^④参见(英)玛格丽特·博登:《AI:人工智能的本质和未来》,孙诗惠译,中国人民大学出版社 2017 年版,第 190 页。

^⑤稳健性是指控制系统在一定(结构、大小)的参数摄动下,维持其他某些性能的特征,这是在异常和危险情况下系统生存的关键。

^⑥参见(英)玛格丽特·博登:《AI:人工智能的本质和未来》,孙诗惠译,中国人民大学出版社 2017 年版,第 196 页。

^⑦德国自动驾驶的 20 条伦理准则第七条要求:在被证明尽管采取了各种可能的预防措施仍然不可避免的危险情况下,保护人的生命在各种受法律保护的权利中享有最高的优先性。因此,在技术上可行的范围内,系统必须被编程为在权益冲突时可以接受对动物和财产的伤害,如果这样可以防止人身伤害的话。

^⑧在加利福尼亚州阿西洛马 Beneficial AI 会议上,来自全球 2000 多人,包括 844 名人工智能和机器人领域的专家联合签署了 23 条 AI 发展原则,呼吁全世界的人工智能领域在发展 AI 的同时严格遵守这些原则,共同保障人类未来的利益和安全。该系列原则称为“阿西洛马人工智能原则”,名称来自本次会议的地点。这一系列原则目前共 23 项,分为三大类:科研问题、伦理和价值观、长期问题。

^⑨Matthew U. Scherer. “Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies”. *Harvard Journal of Law & Technology*, 2016 (29): 366.

性的情况下,人仍应保留对其的控制^①。否则,可能导致行政治理的混乱,乃至决策形成的异常,对公共安全产生危害。

(三) 隐私

在现代行政国家和信息社会背景下,行政治理过程中集聚了纷繁的信息,有的信息来自行政机关的日常管理工作,有的来自市场主体和社会主体的呈报。当行政治理中引入人工智能系统后,不仅获取了更多信息,还使得行政机关有可能对更多信息加以分析和利用,使信息成为“源头活水”。

同时,需要看到,人工智能与大数据治理可能会危及个人的隐私权。隐私权是指自然人享有的私人生活安宁与私人信息秘密依法受到保护,不被他人非法侵扰、知悉、收集、利用和公开的一种基本人格权利。《民法总则》第110条规定,自然人享有隐私权;《网络安全法》第45条规定,依法负有网络安全监督管理职责的部门及其工作人员,必须对在履行职责中知悉的个人信息、隐私和商业秘密严格保密,不得泄露、出售或者非法向他人提供。

因此,为了保护隐私,在人工智能治理过程中,需要明确政府数据开放的范围与标准,并探索数据匿名化等新型数据保护工具,以更好保护相对人隐私。

(四) 政府治理能力的弱化

人工智能在治理中的应用,会在某种程度上弱化政府的治理能力。因为人工智能系统可能由市场主体加以开发设计,政府机关在利用人工智能系统时,可能不理解人工智能系统中的算法与数据,乃至不知道有些决策是怎样发生的^②。某种意义上,将政治考量、政策决断和个案处理交给

了机器。进一步地,算法的不透明性,使得某些互联网平台企业、互联网商业公司不仅集聚了数据和财富乃至分享了公权力,但对这些商业主体缺乏明确的公共问责机制。为此,应要求人工智能的研究者、设计者和实际运营者向应用人工智能系统的行政主体,说明人工智能系统的设计和运营要点,说明其算法所考虑的要素。

此外,“每个机器人背后都是一个良好的人”^③,行政执法人员应在数据的选取,算法的设计、运行和修正中,特别要考虑个案裁量因素,在适用法律规范的环节发挥作用。行政执法人员还要在人工智能所不能为、所不适为的领域,依法开展行政治理。行政执法人员还要学习人工智能系统的相关知识,使得人工智能系统在行政治理中得到更好的使用。

四 人工智能治理的法律程序约束

公权力的运行离不开正当程序。正当程序蕴含着尊严、平等、秩序、公正和效率基本价值,这构成了法治国家的基石^④。人工智能在行政治理中的应用,亦应纳入法治的轨道。

(一) 透明度的要求

通过人工智能和大数据进行行政治理,有助于改进公共政策,改进公共服务的提供^⑤。在多数情况下,需要以政府购买公共服务的方式,向市场主体购买人工智能系统的设计和运营^⑥。因此,透明度的要求是对政府和企业的要求。在许多情况下,透明是治理的有益工具^⑦。

首先,应公开算法流程,公开验证算法时所生成的适当记录,向公众披露如何开发算法,开发算法时有哪些考量^⑧。通过开放记录,说明算法体现了怎样的政策判断,并可以评估算法的效用和

^①Matthew U. Scherer. "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies". *Harvard Journal of Law & Technology*, 2016 (29): 366.

^②R. Brauneis & EP Goodman. *Algorithmic Transparency for the Smart City*, Social Science Electronic Publishing, 2017, p.6.

^③Arthur R. Miller. "Copyright Protection for Computer Programs, Databases, and Computer-Generated Works: Is Anything New Since CON-TU?". *Harv. L. Rev.*, 1993 (106): 1045.

^④章剑生:《行政程序正当性之基本价值》,《法治现代化研究》2017年第5期。

^⑤Lilian Edwards. Privacy, "Security and Data Protection in Smart Cities: A Critical EU Law Perspective". *European Data Protection Law Review*, 2016, pp.28-58.

^⑥R Brauneis, EP Goodman. *Algorithmic Transparency for the Smart City*, Social Science Electronic Publishing, 2017, p.10.

^⑦Danielle Keats Citron & Frank Pasquale. "The Scored Society: Due Process for Automated Predictions", *Wash. L. Rev.*, 2014 (89): 8.

^⑧R. Brauneis, EP Goodman. *Algorithmic Transparency for the Smart City*, Social Science Electronic Publishing, 2017, p.6.

公平性^①。但不建议完全公开代码和数据,因为完全公开的不利后果在于:第一,有可能泄露个人的敏感隐私数据;第二,有可能使得行政相对人采取行动,来规制人工智能治理;第三,有可能损害人工智能系统设计主体的商业秘密和竞争优势;第四,有可能危及国家安全^②。

其次,为了使算法决策可理解,要求技术上的透明度,就包括事先透明度和事后透明度。事先透明度,即在使用人工智能之前,对决策程序加以解释^③。事后透明度,即事先不知决策程序,但可以通过对人工智能绩效的监测,来确证出决策程序^④。

再次,透明度要求对人工智能系统运行过程予以完整记录。例如,德国《自动驾驶法》要求,带有自动驾驶功能的汽车需安装一个类似飞机“黑匣子”的装置,用来记录具体驾驶情况,以保证在出现交通事故时,明确责任划分。同理可知,当在行政治理过程中运用人工智能系统时,需要留存必要运行记录,以便追溯管理。

最后,人工智能系统某种意义上体现了实验主义治理,通过自主系统进行不断学习,来改进算法。自主系统不是静态地执行固定操作,而是在动态地变化改进。因此需要定期公开相关数据和信息,保障公民了解人工智能系统的真实运作。

(二) 测试和评估程序

正当程序原则对人工智能的约束,也体现在引入测试和评估程序,以确保被写进代码中规则的准确性以及系统的安全性。

测试和评估的前提条件也是体现正当程序原则的关键,即不仅需要查看人工智能系统所使用的数据集,还需要查看算法中描述变量、相关性和结论的源代码与程序员的注释^⑤。

1. 测试

应通过测试平台对人工智能系统加以测试,来测试系统是否违背法律法规及伦理规范,是否准确写入相关规则。测试平台还可测试人工智能系统的偏差、偏见和任意性。测试应该包括预期性能检测和意外场景假设检测^⑥。

2. 评估

行政机关应该对人工智能产品和系统的关键性能进行评估,必要时可发布评估报告,来说明人工智能系统蕴涵的安全、歧视和隐私等风险。围绕此评估报告展开风险交流,有助于弥合专家与公众认知的分裂^⑦。

(三) 说明理由

说明理由制度对正当程序原则的践行具有重要的意义。说明理由是良好行政的基本原则之一,有助于实现程序性权利的工具性价值与非工具性价值,有助于保障行政相对人权利^⑧。

在特定情形下,有必要就通过人工智能系统做出的行政治理决定说明理由,特别是当人工智能行政治理决定有可能给行政相对人带来不利影响,涉及行政相对人实体权利、程序权利及救济权利,以及涉及“最低限度的公正”时,应就人工智能系统的运营过程、运营结果、做出行为的推理,给出理由说明^⑨。说明理由的欠缺,可能导致相应行政决定的无效或被撤销。

说明理由要求行政机关向相对人就人工智能系统的结果或行为加以解释。首先,应解释决定时考虑的主要因素有哪些,以及各因素的不同权重;第二,当某个因素改变时,是否会改变行政决定,哪些因素才是决定性的;第三,为什么相似的情况可能有不同的决定,为什么不同的情况可能有相似的决定。^⑩

①R. Brauneis, EP Goodman. *Algorithmic Transparency for the Smart City*, Social Science Electronic Publishing, 2017, p.2.

②Paul B. de Laat. "Big Data and Algorithmic Decision-Making". *Acm Sigcas Computers & Society*, 2017 (47): 39.

③AI in the UK: ready, willing and able, <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>, p.38.

④AI in the UK: ready, willing and able, <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>, p.38.

⑤Danielle Keats Citron & Frank Pasquale. "The Scored Society: Due Process for Automated Predictions". *Wash. L. Rev.*, 2014 (89): 25.

⑥Danielle Keats Citron & Frank Pasquale. "The Scored Society: Due Process for Automated Predictions". *Wash. L. Rev.*, 2014 (89): 25.

⑦沈岍主编:《风险规制与行政法新发展》,法律出版社2013年版,第193页。

⑧参见宋华琳:《英国行政决定说明理由研究》,《行政法学研究》2010年第2期。

⑨参见宋华琳:《英国行政决定说明理由研究》,《行政法学研究》2010年第2期。

⑩Doshi-Velez F, Kortz M, Budish R, et al. *Accountability of AI under the Law: The Role of Explanation*, Social Science Electronic Publishing, 2017.

在对借助人工智能系统做出的行政决定给出理由说明时,还需要做出两个层面的解释。首先是局部的解释,是指对某一特定决定解释,不需要对整个系统活动进行解释^①。其次是对因果关系的解释,说明存在哪些因素,为何这些因素导出这样的结果^②。但并不必对系统技术细节加以过多解释。

五 结语

本文是一项初步研究的尝试,试图去勾勒人工智能在行政治理中的作用及其法律控制的面貌。本文力图梳理在行政规则制定、政策形成和做出行政决定的过程中,人工智能可以发挥怎样

的作用,也力图指出人工智能治理所面临的歧视、安全和隐私等风险。在未来,应理性对待传统行政治理和人工智能治理的关系,践行正当程序,强化理由说明,以规范人工智能在行政治理中的作用。

如何对行政治理中人工智能系统的引入加以法律控制,其间依然存在诸多未竟的问题,例如,人工智能在行政治理中,还可以发挥哪些作用;人工智能对传统行政决策过程产生哪些冲击;人工智能治理与传统行政治理之间是怎样的关系;人工智能治理责任主体、归责措施、救济途径为何,等等。这些问题都有待我们继续观察人工智能治理实践的发展,继而做进一步的探究。

Artificial Intelligence's Function in Administrative Governance and Its Legal Control

SONG Hua-lin & MENG Li-mian

(School of Law, Nankai University, Tianjin 300350, China)

Abstract: There lacks of consensus on the definition of artificial intelligence. However, artificial intelligence can be classified as Artificial Narrow Intelligence, Artificial General Intelligence and Artificial Superintelligence by its level of development. In administrative governance, artificial intelligence may be involved in procedures for administrative rulemaking, policy-making, and administrative decision in order to promote intellectualization of governance. Meanwhile, governance by artificial intelligence has its limitation in application. Artificial intelligence governance may cause risks of discrimination, security and privacy, and even decrease the regulatory capacity of governments. Artificial intelligence governance need legal control and constriction by due process for good operation.

Key words: artificial intelligence; administrative law; due process; reason-giving

(责任校对 朱正余)

^①Doshi-Velez F, Kortz M, Budish R, et al. *Accountability of AI under the Law: The Role of Explanation*, Social Science Electronic Publishing, 2017.

^②Doshi-Velez F, Kortz M, Budish R, et al. *Accountability of AI under the Law: The Role of Explanation*, Social Science Electronic Publishing, 2017.