

---

# “人工智能+”政府决策:挑战与应对

刘宇轩<sup>1</sup> 张乾友<sup>21</sup>

(1. 浙江大学, 浙江 杭州 310012;

2. 南京大学, 江苏 南京 210023)

**【摘要】:** 人工智能对人类社会正在产生巨大的影响,并在不同的方面降低人类的决策成本,提高决策精度。政府决策作为社会治理的核心内容之一,是人工智能应用并造福人类的重要领域。近年来,各地政府不断加大对人工智能应用的资金投入,利用人工智能技术辅助政府决策在短期内已经取得较好的效果。但与此同时,人工智能应用于政府决策带来的诸多问题也同时引发了巨大的争议。人工智能技术应用于政府决策仍面临一些难题和潜在的伦理困境,这要求我们对人工智能做出批判性的思考,进而寻找政府如何更好地促进人工智能在政府决策中的应用路径。

**【关键词】:** 人工智能 政府决策 社会治理

**【中图分类号】:** D630.1 **【文献标识码】:** A **【文章编号】:** 1002-6924(2021)4-0014-08

经过数十年的发展,人工智能这项厚积薄发的技术随着运算能力的飞速提升在近年来越来越多地造福人类的日常生活,并使“人工智能可以帮助人类降低决策成本”的观念获得了日益广泛的接受。在“注意力稀缺”的今天,在浏览视频或是购物网站时,便捷的推荐系统可以帮助浏览者高效地锁定迎合自己偏好的内容,从而大大降低选择过多条件下的决策成本。而当人工智能可以在日常生活中大幅降低使用者的决策成本成为一种日常经验时,尝试将人工智能引入到政府决策中来也成了顺理成章之事。2017年7月20日,国务院正式印发了《新一代人工智能发展规划》,确立新一代人工智能发展三步走战略目标,将人工智能的发展上升至国家战略层面。《人工智能标准化白皮书》(2018版)中写道:“人工智能可帮助人类准确感知、预测、预警基础设施和社会安全运行的重大态势,及时把握群体认知及心理变化,主动作出决策反应,显著提高社会治理能力和水平,同时保障公共安全。”<sup>[1]</sup>不过,政府决策从来都不是一件纯技术层面的事。尽管在政府决策中引入人工智能带来了许多可预期的积极影响,却并不意味着人工智能就是改进政府决策的最佳答案。人工智能在面临社会治理的大考之前,人类更应该以人工智能创造者的身份对人工智能进行批判性的思考,从而对人工智能引入政府决策可能带来的机遇与风险有更清醒的认识。

## 一、人工智能的两种治理想象

人工智能之所以引起如此大的关注,甚至成为很多国家的国家战略,一个很大的原因在于它激发了人们的治理想象,让人们意识到了使现有的治理模式或至少是治理方式发生根本性变革的可能性。不过,在展开这种想象时,很多人头脑中其实存在一种误解:人工智能是真正具有智慧的。这样的误解在AlphaGo战胜韩国棋手李世石所引发的大范围讨论中就有体现。有人认为AlphaGo在围棋上取得的成就意味着其在其他领域也一定能大展拳脚,更有人戏称AlphaGo面对“神之一手”落败的一盘是人工智能主动做出的退让。这种误解将人工智能与通用人工智能的概念混为一谈。所谓通用人工智能更加接近于科幻作品中无所不

---

<sup>1</sup>**作者简介:** 刘宇轩,浙江大学公共管理学院博士研究生,主要研究方向:数据治理与公共管理方法论;;张乾友,南京大学政府管理学院副教授,博士生导师,主要研究方向:政治哲学与公共行政理论。

**基金项目:** 教育部人文社会科学重点研究基地重大项目“创新驱动与政府角色”(16JJD0009)

---

能,或者至少是多才多艺的超级人工智能。通用人工智能当“具有人类水平,甚至超人的智力;将他们的知识推广到不同的领域;反思自己;这创造了根本性的创新和见解。”<sup>[2]</sup>而现在为大众所体验到的人工智能则是基于算法训练出的数学模型,更加极端地说,现有的人工智能和机器学习都只是统计学发展的产物,而统计学模型显然与智能或智慧并不属于同样的范畴。人工智能一词似乎也并不是该领域内唯一出现类似歧义的概念,深度学习原先更多代指深层的神经网络,而如今,深度学习则被赋予了多元的含义,以至于可能需要新的词来代指原先的深度学习。人工智能与通用人工智能之别意味着:当谈及通用人工智能时,我们所做的论述建立在一种科幻式的畅想之上,就可行性或者可实现性而言,这样的论述更像是空中楼阁;而当谈及人工智能时,所涉及的则更多是技术和操作层面的内容。有观点认为,“人工智能的价值不在于模仿人类行为,而是其具备自主的学习和决策能力”<sup>[3]</sup>。而事实上,以机器学习为基础的人工智能恰恰是在模仿人类行为——人类不仅提供可供训练的数据,模型的特征往往也是人类选择的结果。这表明,在谈及人工智能时,人们往往并未意识到人工智能与通用人工智能的区别。之所以如此,一个重要原因在于人工智能并没有呈现出超越其他现有技术的应用前景。所以,只有似乎全知全能的通用人工智能才能激发人们的想象力。但另一方面,如果不能明确讨论的对象,那所有的治理想象可能都仅仅只是一种想象。

对于政府而言,一个现实的困难在于,无论官员还是其智囊,我们都很难奢望他们在处理工作事务之余再对人工智能进行技术层面的深入学习,即使他们有意愿这么做,可能也缺乏足够的时间与知识储备。这意味着对于人工智能与通用人工智能概念的混淆同样可能发生在政府官员对于人工智能应用前景的预期上。从表面看,人工智能为政府决策带来的帮助和利用技术提高政府行政效率似乎有相似之处,而技术改进在提高行政效率上的成功经验似乎为人工智能在政府决策中的应用注入了信心。但严格说来,两者并非完全相同。通过现有的技术,我们可以在物理学意义上复现人类的一些行为以及这些行为所产生的结果,工业革命就是这种复现并喷的代表。但以当前的人类科技,没有哪种技术可以精确地复现人类大脑的决策过程。这意味着,人工智能即便是辅助政府的决策,只要其或多或少承担了决策的任务,决策的过程就有可能因此发生变化,而这种变化发生的概率和利弊都是暂时无法给出准确答案的。其原因在于,现有的人工智能本质上是以数据分析为基础的,是一个数学的黑盒子,而人是一种生物,并非数学公式的组合,因此,对于人工智能的应用理应建立在更加成熟的考量之上。从另一个角度而言,对于同样的政策难题,在被告知使用人工智能进行决策时官员的心理预期与被告知使用数据分析时官员的心理预期可能是存在差异的,但是就当前的技术进展而言,两者又十分相似,甚至在一些情况下趋近于相同。如果这样的差异确实存在,这也从另一个方面警醒政府对于人工智能应用的乐观预期可能受到了对于人工智能的误解的影响,错误地把人工智能和通用人工智能混为一谈。

在人工智能减少政府决策成本的问题上,担心的声音认为,本应由人类经过思考做出的判断变成由算法计算得出,这背后可能涉及更深层次的伦理问题;也有更为畅想式的声音,谈论政府最终被人工智能取代的种种。现已有研究提出了“人工智能的决策伦理”<sup>[4]</sup>或者算法伦理的概念。而这样的概念在通用人工智能出现之后也许会成为有意义的议题,但是针对现有的人工智能,算法并不涉及伦理问题。当然,这并不是说对于通用人工智能的伦理担忧毫无意义,而是说,由于现有的人工智能在原理上并不支持人们所担忧的那些场景,在通用人工智能的底层技术没有为研究人员勾勒出一个化虚为实的框架之前,过多的伦理担忧可能束缚这样一种蕴含着为人类带来重要治理变革之可能性的技术的发展。

区分人工智能与通用人工智能的意义在于,通用人工智能才是真正的“智能化”,而现在的人工智能更多是一个“凸优化”的问题。“智能化”意味着通用人工智能可以自行定义它要解决的问题,自行设定对所定义问题之解决所需回应的价值标准,并在此基础上做出决策。“凸优化”则是在既定价值标准前提下寻找给定问题的最优解。这表明,对现有的人工智能不应抱有和对真正媲美人类智慧的通用人工智能一样的期待。同样,人们对通用人工智能所蕴含潜在威胁的回应也不应落在以机器学习为基础的人工智能身上。算法作为一种演绎推理能力的反映,并不能到达人类形象思维、另案思维的能力。<sup>[5]</sup>现在的人工智能大多是通过机器学习算法,对数据进行拟合,其本质是对于真实智能的尽可能的“形似”。然而无论多么极致地“形似”都不意味着“神似”。人工智能模型的决策方式和智能主体思考的方式尽管有着颇为相似的外表,其内核却不尽相同。因此,人类无法也不应该理所应当的期望现有的人工智能必然会发展成为想象中的通用人工智能。同时,通用人工智能也未必是人工智能发展到一定阶段的产物,也许两者并不共享同样的底层技术,这一问题只有在通用人工智能真正出现的那天才会有答案。因此,尽管两者的称谓十分接近,但是在学术探讨中,应当将两者严格区分。从社会科学的角度,两者也截然不同。人工智能是现实

---

问题，今天的人们有着运用这一技术来解决某些实际问题的旺盛需求，也要求社会科学家从实际出发来评估这一技术的发展和应用；通用人工智能目前则更多停留在想象阶段，它更需要的是技术层面的研发，以获得可实践性。而由于其实践前景尚不明晰，社会科学家尽可以围绕其展开想象，同时不必抱有过于强烈的道德预设，而应以更加宽容的心态来迎接其发展。

在对人工智能进行评估时，人类面对同样问题的表现是一个重要的参考维度。很多对于人工智能的伦理责难实际上是把通用人工智能可能解决的问题强加给了人工智能。在一些研究中，人们把诸如“归纳主义者火鸡”“电车难题”<sup>[6]</sup>等难题交给人工智能而得到人工智能可能存在伦理困境这一结论。这样的论证是不恰当的。实际上，面对这些困境，即便是最聪明的人类也会难以决策。在人工智能的具体应用中，尽管往往模型的表现能略好于人类的平均水平，甚至是优秀人类的水平，但是面对人类也无法给出明确回答的伦理难题，过多的苛责模型和算法是没有意义的。面对“电车难题”，无论是人类还是人工智能都只能二者取其一。不同的是，人类在抉择时很清楚这是一种伦理选择，而在人工智能那里，这种抉择本身并没有伦理含义。当然，如果人类确认了人工智能的选择结果，就赋予了这一选择以伦理含义，但这并不意味着人工智能本身做出了伦理选择。反过来说，寄希望于人工智能来帮助人类做出那些困难的伦理选择，这可能是对人工智能的一种错误期待。同时，即使通用人工智能出现后做出了从特定角度来看更优越的选择，也并不意味着人类应当接受通用人工智能所做出的选择，正如我们并不会接受某个公认“智者”的所有建议一样。人类的存在一直面临许多伦理挑战，而人工智能的任务并不是解决这些伦理挑战，而是做出一个“更像人类”的选择。在这个意义上，人工智能可以帮助人类反思自己的决策推理——即将思想实验变成技术实验，从而改进人类的决策质量。但寄希望于人工智能为人类提供一种更优越的决策推理则是不切实际的。同时，即使是使用人工智能改善政府决策的质量，也面临着巨大的挑战与难题。

## 二、“人工智能+”政府决策的现实挑战

### （一）前景评估难题

在以往政府行政改革的尝试中，不乏技术提高行政效率的成功案例。而大数据、人工智能等技术显然承载了更高的期待，人们希望它们可以“使政府管理的体制、结构、职能、流程和方式发生革命性的变化。”<sup>[7]</sup>在过去，电话的出现无疑大大降低了政府内部的沟通成本，带来了沟通效率的提升，但是显然不会有人将电话与“革命性的变化”这样的字眼联系起来。这意味着，相较于绝大多数技术推动政府变革的案例，人工智能获得了更为特殊的对待，但这不仅应体现在对其的推动上，也应表现为更为审慎的态度。

与人工智能不同的是，绝大多数技术虽然最终成为了政府的帮手，但是这些技术从未在政府决策的舞台上发光发热。反之，这些技术带来的裨益往往表现为对于社会治理中的一些物理性过程的重复——老百姓网上办事省去的是跑腿的功夫，而安全方面的加密技术则是为信息的安保省下了人力。但自始至终，政府的决策成本并未因此减少，甚至在特定的情况下，一项技术的使用从起步到成熟的过程，反而需要大量的决策成本。在过去，技术承担着减少政府职员体力劳动的任务，而人工智能则是朝向减少脑力劳动这一目标的尝试。

在预估人工智能在降低政府决策成本方面的可能前景时，首先应当对于人工智能可能的应用做出预判。不难发现，并不是政府使用的所有人工智能应用都会降低决策成本。举例而言，上海市政府开发的“上海刑事案件智能辅助办案系统”建立了“16.55万份电子卷宗、裁判文书和庭审笔录组成的数据库”，而该人工智能政法系统的应用能使“庭审时间缩短20%至30%，疑难复杂案件的审判时间更是缩短了50%”。<sup>[8]</sup>显然，人工智能在改变司法审判过程的直觉主义，推动法律符号系统的完善方面有独特的作用。<sup>[9]</sup>在人工智能政法系统的应用中，除了辅助决策，人工智能同样承担着将书面稿转化为电子稿的任务。而这种转录任务，并不属于决策的范畴，人工智能实际上扮演了一个打字员的角色。诸如此类的人工智能应用并不存在太多的伦理困境，因为这种应用本身也只是取代了一些体力劳动。同时，此类应用事实上也像电话一样并不会带来革命性的变化。但是在手写的卷宗、裁判文书和笔录转化为电子稿件之后，如何通过自然语言处理的方法得到一个关于量刑的精确化判断才是人工智能应用的争论焦点，也是人工智能独有的应用前景。

---

可见，一方面，人工智能到底能在政府决策中得到何种程度的应用，很大程度上取决于技术的进展，而技术能进展到何种程度，这本身是不确定的。另一方面，我们虽然要求政府工作人员应当是人工智能方面的专家，但他们中的大多数又并非如此，从而使他们在如何应用人工智能的判断上也存在不确定性。如果政府依靠自身进行判断，那么技术难题便难以解决，可能形成对决策智能化的错误预期。而如果政府选择寻求外部技术力量的帮助，这些外部技术力量对于政府业务的了解又有欠缺。如果要培养精通人工智能和工作业务两方面知识的复合型人才，一方面，成本会极其高昂，另一方面，应用前景的不确定性又会造成个体缺乏投身这一领域的动力。这是导致这一问题到目前未能很好解决的原因。

## （二）数据质量难题

关于现有的机器学习模型最终表现如何，机器学习从业者中流传着这样一句话：数据本身的质量和特征工程决定了模型的上限，而模型的选择和参数调整则决定了模型最后是否能逼近这个上限。如果数据质量低，那么模型表现就很难尽如人意，一个精度和准度不足的模型，显然不足以帮助甚至辅助人类做出决策。人工智能应用政法系统作为一个在一定程度上成功的案例，离不开卷宗、文书等资料的妥善保存，对于人工智能应用政法系统的开发者来说，优质的数据资源相对易于获取。但是如果我们的人工智能用于研判哪些官员更应该被优先提拔，便会发现很多工作信息并没有以数据的形式得以保存，因此获得一个优质的数据集便成为难题。因此开发一个提拔官员的人工智能就比人工智能应用政法系统要更为困难。随着诸如 Facebook 泄露用户信息等丑闻的频频爆出，各国公民都愈发意识到我们正处在一个信息保护尤为重要的时代，隐私权的问题也被频频提及。这也意味着想要大范围获取数据不仅仅是一个技术难题，同时也是一个伦理难题和政治难题。很多人都不愿“赤裸”地生活在监控之下，无论是人的监控还是算法的监控。这意味着人工智能和隐私保护潜在的争斗在未来可能会愈演愈烈，也意味着获取一个高质量的数据可能会面临更高的成本和更为严厉的道德挑战。

数据质量难题的有些方面表现得更为隐蔽。如，谷歌曾经尝试开发一款人工智能面试官为企业中的 HR 提供人才录用方面的帮助。谷歌有着高质量的数据集和好的初衷，试图通过人工智能技术模拟出一个人工智能 HR，从现有的考评资料中学习 HR 的工作模式。但事与愿违，训练出的人工智能面试官尽管在数学意义上有着出色的表现，但在实际模拟之后却表现出了性别歧视、种族歧视等特征。人工智能面试官也因此夭折。在这里，指责人工智能面试官是不公平的，因为人工智能并没有自己的价值观和意识形态，其所表现出来的种族歧视恰恰源于其学习的对象，也就是那些人类面试官。人工智能面试官的失败指明，即便数据在数学意义上表现出质量，但是数据本身在伦理意义上可能并不达标。这也是人工智能在社会问题的应用中面临的独特挑战。人可以将自身的意识形态通过合理的解释加以隐藏，例如，对于女性的歧视可以解释为职位对于力量的需求，但人工智能则会将性别歧视这一本源表现出来。当人工智能被引入政府的事务中，那么不可避免的是模型对于政府官员的学习，而在这些学习对象本身可能存在道德不完善的前提下，学习的结果就很可能重蹈人工智能面试官的覆辙。而这一结果的影响显然会比在企业之中严重得多。针对这样的情况，部分学者提出了在算法中嵌入伦理机制的方法，但问题在于，算法本身并不存在伦理问题，算法是工具性的，仅仅为达到某一目标提供路径，伦理困境并不发生在算法上，而只发生在人类自身。如果试图简单地消除人工智能所表现出来的意识形态，必然要去改变人工智能所学习的数据源，从而模型的精准度不再得到保证，我们本寄希望于人工智能来达成的目的也会受到扭曲。

## （三）极少数偏差难题

一个优秀的人工智能模型，最终输出结果的精度和准度通常能达到很高的水平。但是，再优秀的人工智能输出的结果也一定会有偏差，换言之，错误在所难免。如：购物 App 使用的推荐系统便会偶尔推荐用户并不感兴趣的东西，或是在用户购买了一台冰箱之后频繁地推荐各种品牌的冰箱，但一个家庭往往并不需要那么多冰箱。这是人工智能犯错的实例之一。但推荐系统的这类错误无伤大雅，对于淘宝、京东等公司来说，真正重要的是推荐系统正确推荐的那些“大头”。同样，各类人工智能医生在诊断时也不能做到尽善尽美，但是通过对精度和准度的调节取舍，至少可以确保人工智能医生基本不会将病患诊断为健康，对于医院来说，这已经可以在很大程度上节省检查设备的负荷。对于政府而言，能否找到类似的应对措施决定了人工智能究竟能在多大程度上帮助政府节省决策成本。如果人工智能的决策可能存在的错误导致政府不得不对其逐一审核，那么人工智能辅

---

助政府决策就失去了意义。人工智能辅助政府决策具有巨大的价值是因为可以将政府官员从大量重复性的决策过程中解放出来，从而保证其他更重要的政府决策的决策质量。如果对于人工智能的决策全部需要人工进行验证的话，就意味着这些重复性的决策过程其实并未被省去，人工智能便也就仅仅能对决策提供验证作用。因此人工智能辅助政府决策对于人工智能决策的精度有着极为苛刻的要求。

如果存在一个人工智能法官，它虽然不可避免地会犯错，但相较于人类法官则具有明显更低的错误率。那么，人工智能法官的错误似乎就是可以容许的，因为至少错判率已经大大得到降低。但其中的问题在于，人类的错判在确认之后可以要求当时的法官提供解释，并予以矫正和相应的惩处。但对于人工智能法官，个案的错误最终只能成为模型调整的众多因素之一，没有人能为此给出解释，追责于技术人员似乎也毫无道理。人工智能是一个强大的黑盒子，但是在使用之前，政府更应该思考人工智能是否能够“增加政府管理社会、经济、文化事务活动的公开性和透明度。”<sup>[10]</sup>通常来说，人工智能如果能够在某方面降低人类的决策成本，一个隐含的前提条件是在这一方面的判断上人工智能已经拥有了超过人类的判断力。但是紧接着的问题是，在意识到人工智能有着更为出色的判断力时，人类是否还能保持对于自身判断的自信。尤其是当人工智能辅助政府的决策时，政府职员能否对这一强大的黑盒子保持一定距离，避免过度依赖。这决定了人工智能的罕见犯错是否能够被及时地发现和纠正。

#### (四) 预期陷阱

一方面，即便人工智能在技术层面就此不再取得任何突破，凭借现有的经验已经足够在方方面面为人类提供大量的帮助。另一方面，人工智能能否在技术上取得突破，又关乎人类对人工智能的期望。这其中的危险之处在于，对于技术越是陌生的人，对技术的期望越是线性。换言之，他们的期望随着技术不断地进步越来越高。但是反观人类技术发展的历史，技术进步往往呈现一种螺旋式上升。事实上，技术发展困境一词略有不妥，因为技术发展的停滞是常态，而技术突破相对少见。这意味着，对于某一项特定技术的进步抱有线性的期待是危险的。从1949年Donald Hebb提出Hebb学习规则至今，机器学习经历了一段螺旋式发展的历程，机器学习的应用也从1952年的跳棋人工智能扩散至社会的各行各业。但是在20世纪60—70年代，人工智能遭遇了前所未有的寒冬，技术人员一度认为其不具备足够的可行性。尽管今天人工智能爆炸性的话题度使得该行业进入了前所未有的繁荣时代，但这并不意味着第二次寒冬不可能到来。计算机视觉领域的专家Filip Piekniewski认为人工智能的寒冬必将到来，并给出了三条理由：深度学习技术没有突破进展，深度学习无法扩大规模，自动驾驶濒临崩溃。<sup>[11]</sup>Gary Marcus也对人工智能的热潮提出了质疑，他认为识别和理解有着不同的含义，“智能机器不需要复制人类的技术，但是更好地理解人类的常识推理可能是一个很好的起点。”<sup>[12]</sup>对于人工智能的担忧并不仅仅停留在纯粹技术的层面，“国防、国土安全和刑事执法部门使用的机器学习已经开始引发警报，并引发了严肃的伦理考虑。”<sup>[13]</sup>

更严重的期望陷阱在于，人们期待人工智能尽可能多地解决那些较为复杂的问题，却忽视了莫拉维克悖论的存在。莫拉维克悖论是指，对机器人来说，简单的动作比复杂的推理更难。Steven Pinker也认为：经过数十年人工智能的研究，人们学到的主要经验是，“困难的问题是简单的，简单的问题是困难的。”<sup>[14]</sup>早期的人工智能研究者乐观地认为思考机器很快就会出现，因为他们通过计算机巧妙地解决了复杂的代数与几何难题，创造出了像人一样下棋的人工智能。但随后他们发现，恰恰是幼儿也能做到的简单的识别与感知，对计算机来说要比推理难得多。人类在直觉上有着巨大的禀赋，有着对于周围事物强大的感知力，同时有着出色的行动能力。感知与行动并不需要复杂的计算，但对于计算机来说却难以做到。人类无法以自身的常识去判断，对于人工智能而言，怎样的突破是困难的。因此人类对于人工智能的期待可能和早期研究者的自信一样，看起来理所当然，实际上却存在巨大的壁垒。同样，人工智能也无法解决那些人类也束手无策的问题。Alpha Zero可以下围棋，但如果存在一种人类都没有发明的棋类玩法，Alpha Zero是无法学会这种玩法的。先会有会下围棋的人，然后才会有会下围棋的人工智能。

人工智能技术发展的停滞期本身对于社会整体而言并不会构成巨大的危害，但如果对于人工智能的期待呈泡沫式增长，那么泡沫破灭的巨大落差是值得警惕的。当政府越来越多地将决策任务委以人工智能时，如果技术的应用达到了底层技术所能达到的极限——事实上人工智能的底层技术已经多年未有突破，整个政府的改革可能会因此受限。因为在这样的情况下，政府已经把很多原本复杂的改革问题简单地转化为了对于算法进步的依赖。以人工智能降低政府决策成本的改革尚处在早期阶段，在

---

这一阶段更为谨慎地对待人工智能的利与弊能够帮助政府在诸多问题出现的时候更加游刃有余。换言之，政府更应该对人工智能，或者是其他任何一项技术，例如区块链等，抱以更低的期待和更大的耐心。因为对于政府而言，惊喜远比失望更加安全。在提倡推动治理能力现代化、智能化的今天，对于智能技术前景的评估显得尤为重要。政府的数字化、智能化改革必然是由一系列子集组成的长期的进程，需要前瞻性指引确保改革的平稳推进。如果智能化技术研究的进程在未来出现停滞，政府的数字化改革则可能会因此陷入预期之外的困境中。

### 三、不确定前景下“人工智能+”政府决策如何作为

#### (一) 理性看待人工智能前景

面对人工智能浪潮，政府需要明确的一点是：人工智能并不是政府的救命稻草，而是政府决策锦上添花的帮手。在人工智能这一概念出现以前，政府已经存在了数千年，在数千年的历史中，政府不需要人工智能的辅助也完成了大量的治理任务，这其中的决定性因素恰恰是人类自身。在政府演变的过程中，有大量的改革设想曾被提出。人工智能在政府决策中的应用，也像这些改革一样，也许最终会成为政府决策中稳定的一环，也有可能由于其他原因最终不能为政府所用。但毋庸置疑的是，人工智能已经呈现出的巨大应用前景对于政府来说是一笔宝贵的财富，即便从保守角度看，人工智能强大的数据挖掘能力也能在很大程度上弥补人类有限的信息处理能力。借助人工智能，政府可以更加准确掌握社会情况，从而更好地完成其职能。但目前并没有任何一种论证能够证明人工智能是通向一个更好政府的必由之路。人工智能更可能是众多路径中的一种，是推动政府前行的无数双手中强而有力的一只。在反思人工智能带来的机遇与挑战时，政府应当忘记人工智能这个称谓，如果这项技术不叫人工智能，而是仅仅叫统计学习或是机器学习，其对政府的吸引力也许会发生一定的改变。

#### (二) 促进技术理性与政治理性的统一

当代各国政府面对的一个共同挑战在于，公众对于政府的期望越来越高，要求政府总是能在所有事务上高效地做出正确的决策。与此同时，传统的集体决策机制则日益僵化，尤其公众内部的决策分歧往往导致政府在面对棘手问题时难以抉择。正是在这两方面压力的驱动下，当代公共决策日益明显地走上了技术替代政治的发展方向。如成本—收益分析这一决策技术就因其能够更直观地帮助人们理解决策的后果而显著地改变了各国政府的监管决策，在提高了监管决策技术合理性的同时，也在一定程度上因其排斥了公众参与而削弱了监管决策的政治合理性。在这种发展的支持者看来，这并不必然是一件坏事，因为公众经常受到错误认知的影响，而成本—收益分析等决策技术则可以帮助公众矫正其错误认知，由此，虽然最终的决策并不符合公众最初的认知，却能够符合公众的反思理性。而在这种发展的反对者看来，究竟哪些监管事项可以纳入成本—收益分析从而进入决策过程，这是一个不透明的问题，而在这一根本性问题上排斥公众参与，就等于剥夺了公众的决策参与权。从实践来看，任何决策技术的应用都是有成本的，尤其当决策所需的计算需要大量专家参与时。在某种意义上，正是由于计算的成本非常高昂，监管机构才无法对每一种监管提议进行准确计算，结果导致公众只能获取特定监管提议的事实信息，而在有限信息基础上做出的判断也就很难是一种符合技术理性的判断，更不用说符合政治理性了。在这个问题上，人工智能可以被期望替代专家，从而大幅降低决策计算的成本。由此，公众的每一种合理或不合理的监管提议都可以借助人工智能以较低的成本计算出准确的决策信息，而在这些信息的基础上，公众就能更加理性地反思自己的决策认知，从而让作为非技术专家的公众的决策判断也更符合技术理性，进而，决策过程就可能同时体现技术理性与政治理性。当然，这种发展并不会自然发生，它既需要政府探索适用于人工智能的更多决策领域，也需要政府更乐于倾听公众的合理或不合理的意见，并借助人工智能来帮助公众提升自身的决策理性，而不是以技术为理由排斥公众的决策参与。

#### (三) 注重技术人才的分散化储备

作为一种被期望可在很大程度上改进政府决策的技术，人工智能的发展不可避免地会和政府治理产生强互动。任何一个政府都不可能放任这一技术自发发展，而必须寻找更好地规范和利用这一技术的办法。这就意味着，政府应当提高自身的技术能

力，否则，如果政府无力审核人工智能在政府决策中的应用内容，就不可能对这种应用的结果负责。如此，出自提高决策质量目的的技术应用反而可能造成政府履责的不力。要提高政府的技术能力，可能的途径之一是政府增设庞大的技术部门，吸纳最顶尖的人工智能专家为政府决策服务。在 20 世纪的特定时期，许多国家都采纳过这种做法，在特定领域，也取得了很好的成效。但问题在于，人工智能的发展具有极大的不确定性，它到底是在政府领域还是在市场领域更能产生价值，是尚未有定论的问题。因此，一方面，如果将优秀人才都留在政府内部，既不有利于技术本身的发展，也是对公共资源的浪费。另一方面，政府也不应放任市场对人工智能专家的争夺，因为这会导致政府在要应用市场中的人工智能创新成果时处于与大企业间的谈判劣势，甚至使大企业有能力不当地介入到政府决策中来。对政府来说，如果既要鼓励市场驱动的技术创新，又要保证技术应用的问责性，就需要促进技术人才的分散化储备。一方面，政府自身应当设立专门的人工智能技术部门，这个部门不一定在规模上很庞大，但应与产业界有经常性的项目合作，以掌握最新的技术进展。另一方面，政府应限制市场中人才的过于集中，避免优秀人才过度集中在少数组织。在这方面，当前，大学等科研机构与大企业有着天然的利益关系，通过二者间的项目合作推动技术进步可能是最有效的途径选择。但要避免技术力量过于集中，政府就应通过政策手段促进多元化的网络发展，借助“思想市场”来规范人工智能技术的发展，从而保证当技术的进步提供了可操作的政府决策应用前景时，政府可以有多种选择来实践这一前景。

#### (四) 设立针对性更强的人才培养渠道

无论人工智能是一个多么技术化的难题，当其被应用于辅助政府决策时，其面临的困难都不仅仅是技术层面的。政府面临的决策难题是复杂的，其中蕴含着大量的伦理议题，如果仅仅把人工智能作为一个技术问题和政府决策割裂开来，它可能会造成更糟糕的伦理问题。人工智能 HR 就是一个典型的例子。为了尽量避免人工智能造成这类问题，技术人员应当对政府议题有足够地了解，对于政府在不同决策中可能面临的道德困境有充分地认识。在过去，当人工智能应用于某一具体领域的时候，跨学科的合作是常见的方式。但是由于政府决策的严肃性，政府应当有针对性地搭建人才培养渠道，培养人工智能与政府事务都精通的复合型人才。复合型人才可以有效地避免技术专家与政务专家之间可能的沟通鸿沟，更能在开发之前对于可能产生的结果做出预测，规避风险。由于高校内公共事务与人工智能分属于不同的学部，并且共享的基础知识较少，短期而言，政府应该加强对外包公司技术人员的伦理培训，而从长期来看，高校内的跨学科人才培养才是输送高质量人才的长久之道。

总之，在本文提到的种种问题中，技术性问题更多是次要矛盾，而政府本身相关的问题才是主要矛盾所在。技术是客观的，而政府对于技术，尤其是人工智能这样变革性的技术应该有着更加充足的准备。一味地强调技术自身的不足并不能带来实际的改变，但是因地制宜地使用技术可以以技术为推力构建更高效、甚至更加公正的政府。桑斯坦在《网络共和国》中对以协同过滤为基础的推荐系统表达了担忧。他认为，协同过滤机制导致了“越来越多的人只听到他们自己的回音，这样的情形比分裂来的更糟糕”<sup>[15]10</sup>。因此，桑斯坦提出了公共论坛原则，希望通过“赋予演讲者接触场所和人的权利”<sup>[16]18</sup>来解决问题。协同过滤恰恰在推荐系统中得到广泛使用，因此，桑斯坦所说的也是一个人工智能应用的问题。但是，显然该问题的症结在于民众的政治冷漠，协同过滤仅仅是社会的显微镜，将矛盾放大而已。可以预见的是，在未来社会中，诸如此类的争议不会是个例，各种各样的社会问题可能会通过人工智能被放大。因而，政府只有更加谨慎、理性地对待人工智能与社会、政府的关系，才能让人工智能造福于人类。

#### 参考文献:

[1] 人工智能标准化白皮书(2018 版)[EB/OL]. (2018-01-24) [2019-05-15]. <http://www.cesi.ac.cn/201801/3545.html>.

[2] Stan Franklin. A Foundational Architecture for Artificial General Intelligence[C]//Conference on Advances in Artificial General Intelligence: Concepts. Amsterdam: IOS Press, 2007: 37.

[3] 贾开, 蒋余浩. 人工智能治理的三个基本问题: 技术逻辑、风险挑战与公共政策选择[J]. 中国行政管理, 2017(10): 42.

- 
- [4]柳亦博. 人工智能阴影下:政府大数据治理中的伦理困境[J]. 行政论坛, 2018(3):102.
- [5]钱铁云. 人工智能是否可以超越人类智能?——计算机和人脑、算法和思维的关系[J]. 科学技术与辩证法, 2004(5):44-47.
- [6]Philippa Foot. The Problem of Abortion and the Doctrine of Double Effect[J]. Oxford Review, 1967(2):152-161.
- [7]陈振明. 政府治理变革的技术基础——大数据与智能化时代的政府改革述评[J]. 行政论坛, 2015(6):4.
- [8]人工智能:让政府更高效[EB/OL]. (2018-07-30)[2019-03-12]. <http://edu.people.com.cn/n1/2018/0730/c1053-30177984.html>.
- [9]葛翔. 司法实践中人工智能运用的现实与前瞻——以上海法院行政案件智能辅助办案系统为参照[J]. 华东政法大学学报, 2018(5):67-76.
- [10]蒋录全, 吴瑞明, 王浣尘. 电子政务中的政府信息公开[J]. 情报杂志, 2004(4):52.
- [11]Filip Pieniewski, The AI Winter Is Well on Its Way[EB/OL]. (2018-06-04)[2019-02-03]. <https://venturebeat.com/2018/06/04/the-ai-winter-is-well-on-its-way/>.
- [12]Ernst Davis, Gary Marcus. Commonsense Reasoning and Commonsense Knowledge in Artificial Intelligence[J]. Communications of the ACM, 2015(9):92.
- [13]Cary Coglianese, David Lehr. Regulating by Robot: Administrative Decision Making in the Machine-Learning Era[J]. Social Science Electronic Publishing, 2017(5):1161.
- [14]Pinkers. The Language Instinct: How the mind Creates language[M]. Penguin Uk, 2003:192.
- [15](美)凯斯·桑斯坦. 网络共和国:网络社会中的民主问题[M]. 黄维明, 译. 上海:上海人民出版社, 2003.