

人工智能时代精准治理的隐忧与风险

李利文^{1,2}

(1. 广州大学公共管理学院, 广东广州 510006; 2. 广州大学大都市治理研究中心, 广东广州 510006)

摘 要:人工智能驱动精准治理给社会带来巨大福利的同时也潜藏着巨大风险。基于数据纹理和超级档案而展开的单粒度治理是这种风险的重要起源, 因为单粒度精准治理为人类提供了一个较为舒适的生活环境和制度环境, 个体在便利、舒适和安逸的环境中受个体逃避、懒惰和自私的本性影响而陷入单向度世界。单向度世界实际上开启了人类认知的半封闭机制, 在半封闭机制的选择和过滤下不符合个体喜好的事物被排斥掉, 而肯定自我的事物则被接纳, 进而造成个体批判能力的弱化和丧失。更重要的是, 这种风险在外部输入受到约束和内部认知变得固化的双重作用下, 使人工智能驱动的精准治理陷入内卷化困境, 造成个体陷入一种不断追求精致化却又难以突破和超越自我的平庸状态。

关键词:人工智能; 精准治理; 单粒度; 单向度; 内卷化

中图分类号: D630

文献标志码: A

文章编号: 1671-4970(2020)01-0082-09

精准治理是人工智能时代公共管理学关注的核心议题, 而精准治理风险则是精准治理研究的重要切入点。日本学者松尾丰曾指出: “真正恐怖的是, 最终有能力颠覆人类社会的人工智能, 既没有像人一样的外形, 也不会做出像人一样的行为, 而单纯只是预测精度非常高的东西。”^[1] 也就是说, 人工智能的核心优势在于“精准”, 而其潜在风险也来源于“精准”。人工智能驱动的精准治理在精准识别、精准匹配、精准决策、精准监督、精准服务和精准反馈等方面已经表现出不俗的潜力和价值, 而且在扶贫、社保、就业、教育、交通、治安、体育、反腐、药监等多

领域已展开应用, 但这些研究将注意力主要放在精准治理的精细化、个性化、清晰化等价值层面, 却忽视了其进一步发展可能带来的问题和风险。

随着大数据和人工智能的发展, 诸如淘宝等网络购物平台利用用户使用数据精准推送商品广告、今日头条等新闻网页根据用户的浏览记录和偏好推送个性化新闻以赚取流量、某旅行网站利用消费者使用数据进行“杀熟”以及非法组织利用网民浏览记录和个人信息等实施精准诈骗等行为的出现, 都已经开始显示人工智能驱动的精准治理可能隐藏的各种问题和风险。这说明人工智能驱动的精准治理

收稿日期: 2019-10-09

基金项目: 广州市哲学社会科学规划“羊城青年学人”项目(2019GZQN46)

作者简介: 李利文(1988—), 男, 湖南长沙人, 讲师, 从事城市基层治理与公共服务供给研究。

虽然很强大,但“在强大力量的背后,一般都潜伏着巨大的风险,风险可能随时出现,但我们又无法预测它何时出现。”^{[2][149]}传统上而言风险主要起源于模糊性,而人工智能时代的风险反而源自被视为“善治”的清晰性,那么人工智能驱动的精准治理到底隐藏着怎样的风险?又将把我们带向何方?这是本文所致力回答的问题。

一、文献回顾与分析框架

1. 人工智能驱动的精准治理研究现状

精准治理主要是指问题与对策、需求与供给、目标与手段之间的高效率、高精度和高吻合的匹配。本文中精准治理的对象比较宽泛,主要是指涉及公共性的商业行为、社会行为和政府行为。一旦这些行为可能有损公共性,政府就有干预的权力和义务。虽然学界对人工智能或大数据在精准治理上的价值认识基本达成共识(由于大数据与人工智能在精准治理议题上的研究彼此关联、难以分割,为在最大程度上吸纳相关文献,因此本文将两方面的相关文献都纳入综述),但专门对此进行研究的文献还较少,部分零散性的研究成果主要集中在以下两个方面。

一是价值分析层面。大数据为公共部门相关问题的解决提供了无限可能^[3],尤其是在提高绩效和效率方面作用显著^[4],而这些在很大程度上是建立在大数据的精准治理功能之上的。从公共服务视角而言,大数据或人工智能有助于公共产品和公共服务的精准化供给^[5]。一方面,大数据能够对闲置的、分散的、零碎的公共服务资源进行配置和重组,实现高效、个性和精细的公共服务供给^[6];另一方面,大数据有助于树立深入公众行为的需求识别理念,畅通双向的需求表达与识别制度,进而实现对公众真实需求的精准识别^[7]。从社会治理视角而言,大数据不仅强调以社会问题和社会需求为靶向,感知社会态势、畅通沟通渠道、辅助科学决策,从而实施精确预警、识别和管理^[8],而且强调主体利益的多样性,使“人民”的政治修辞具体化为差异性的利益诉求,进而提升社会治理的精度^[9]。还有研究提出人工智能驱动的精准治理必须坚持创新、适度、平衡和多元四大原则^[10],更有研究甚至认为技术驱动的精准治理是中国场景下政府治理范式的进化^[11]。总之,大数据或人工智能的精准治理功能已经作为一种重要的应用价值得到学术界和实务界的普遍认可。

二是应用分析层面。人工智能和大数据在具体领域的应用一开始就受到重视,如大数据驱动由政府决策被认为是精准决策模式,其“精准”贯穿于决策问题发现、决策咨询反馈、决策模拟实施、决策结果评估等过程和环节^[12];大数据驱动的精准扶贫在很大程度上减小了信息不对称,提升了精准扶贫的绩效^[13];大数据驱动由政府购买公共服务对提高公共服务的供给效率和效益,满足个性化和差异化需求具有重要意义^[14];基于机器学习的群租房治理,能够从公开的租房信息中有效区分群租房和非群租房^[15];基于犯罪大数据的解读和研判,能够提出犯罪的精准防控范畴,推动犯罪治理模式进化升级^[16];大数据驱动的图书馆服务,能够为读者群体提供千人千面的精准化服务,从而满足读者差异化、个性化和定制化的服务需求^[17];人工智能驱动的精准治理还能够提升应对自然灾害的回应和处置能力^[18]。这些都是大数据或人工智能精准治理功能积极层面的运用,同时我们也应注意到其还存在被消极利用的可能,如大数据运用于诈骗活动,使这种新型诈骗活动具有隔空、隐形、面广和精准等特点,将会对大众生活产生极大消极影响^[19]。总之,现有研究主要聚焦于大数据和人工智能在积极层面的应用而较少关注其在消极层面的非法应用。

由此可知,大数据与人工智能技术的发展是实现精准治理的重要前提和保障,与传统的治理模式相比,大数据或人工智能驱动的精准治理在很大程度上解决了传统治理中的信息搜集和信息分析问题,能够以较少的成本并轻而易举地进行大范围的个性化和智能化治理。因而,本文中后续的大数据或人工智能驱动的精准治理主要是从这一层面强调其与传统治理模式的不同而展开分析的。

2. 分析框架:单粒度治理—单向度悖论—内卷化困境

从以上综述可知,现有研究主要聚焦于精准治理的价值挖掘,而忽视了精准治理进一步发展可能隐藏的问题和风险。人工智能驱动的精准治理带来了个性化、精细化和智能化,但这是建立在清晰化基础之上的,也就是说清晰化是人工智能驱动精准治理的逻辑起点。社会科学在预测方面的表现向来令人失望,但是在大数据和人工智能的帮助下,“精准预测”变得不再那么困难^[20]。精准预测是建立在对事物的清晰化认知的基础之上的,而清晰化认知的

首要工作是建立超级档案。超级档案是人工智能将复杂、多元、持续的个体行为信息和数据进行清洗、联结和整合的超大数据集,这个超大数据集实际上是由众多的颗粒信息所组成的,超大数据集中的不同颗粒具有不同的数据纹理,正是根据这一纹理特征可以对多元复杂的个体实施个性化、精细化和智能化的单粒度治理。正如涂子沛的定义,单粒度治理就是根据一个人数据的纹理来治理^{[21]39},即基于数据纹理和超级档案而实行的个性化、精细化和智能化的分析和管理。这种单粒度治理是清晰化的重要体现,也是人工智能进一步发展隐藏风险的起源。

单粒度治理利用技术刚性为人类带来诸多便利、舒适和福利,而这种便利、舒适和福利主要建立在“精准”的基础之上。诸如根据个体特征提供个性化的服务、依据问题特征给予针对性治理以及基于不同记录推送差异化的信息等。这种具有广度、速度、效度、精度和温度的单粒度治理看似具有便利性和舒适性,但其本质上是根据人的特征和喜好来推送信息和提供服务的,在长时段的舒适区域和美好生活中人们可能会不断满足个人喜好而压制困难、恐惧和挫败,这将使人陷入具有过滤性、选择性和排斥性的半封闭状态,进而造成人格、思维和反思能力的缺陷,甚至诱发种种问题。这正如赫伯特·马尔库塞所说的单向度的思想和行为模式,即“在这一模式中,凡是内容超越了既定的话语和行为领域的观念、愿望和目标,不是受到排斥就是退化到这一领域。”^[21]由此可见,为促成社会福利最大化的精准治理反过来可能制约社会福利的增加和扩大,人工智能驱动精准治理就可能陷入这种单向度悖论,促成相关风险的形成。

一旦陷入单向度泥淖,人们接收到的信息将具有片面性和倾向性,尤其是基于人工智能算法推送的各种信息和做出的各种行为,将在最大程度上拟

合个体特征和满足个体喜好,这实际上限制了人们接受整体性、综合性信息的机会,在人的自主性之外增加了一层束缚。如果在这种情况下个体不能够敏锐地意识到这层束缚并迅速走出舒适区和突破自我,其认知将在既有的经验和知识上不断地细化深入和循环往复,即使经过过滤的新的信息不断输入,但个体的认知水平仍然不会有实质提升,这样就会促成刻板效应和认知固化,进而在这种外部受限和内部固化的状态下陷入平庸。此时,人工智能驱动精准治理所隐藏的风险将被进一步放大,陷入内卷化困境,即“一个系统在外扩张受到约束的条件下内部变得更加精细化和复杂化的状态。”^[22]基于以上“风险起源-风险形成-风险放大”的脉络,本文构建了一个“单粒度治理-单向度悖论-内卷化困境”的分析框架(图1)。单粒度治理是人工智能时代精准治理风险的重要起源,单向度悖论则是其风险形成的重要体现,而内卷化困境则是其风险放大的重要后果。接下来将基于该脉络和框架对人工智能驱动精准治理所隐藏的问题和风险进行细致地阐述和剖析。

二、人工智能时代精准治理风险的起源

一般而言,“风险的本质并不在于它正在发生,而在于它可能会发生”^[23],也即风险主要产生于不确定性。传统的国家治理是基于经验的模糊性治理,具有高度的不确定性,往往伴随着极大的风险,而现代的国家治理是基于技术的清晰化治理,不确定性急剧下降,伴随的风险也随之大幅减少。但人工智能时代的精准治理风险却恰恰与传统理论相悖,其主要起源于高精度的单粒度治理而非不确定性的模糊性治理。

1. 数据纹理分析

“在数据空间里,一个人或一个物体就是一个

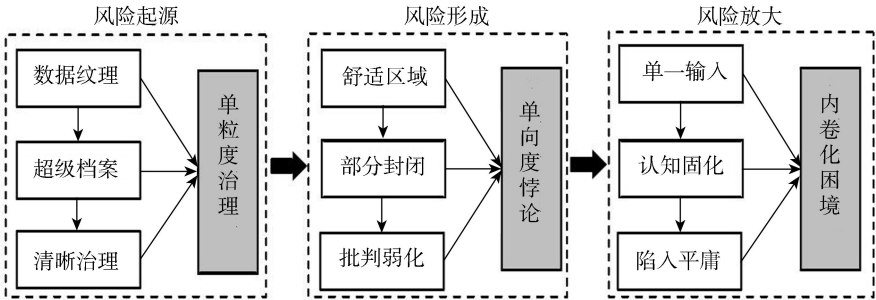


图1 分析框架

独特的“数体”,它可以被无数数据定义、支撑、背书,就像声波和指纹,每一组、每一条数据又都有自己的特征,这就是数据纹理。”^{[2][13]}人工智能驱动的精准治理之所以能够实现个性化和精准化的目标,就是因为人工智能对“数体”的数据纹理进行了精准剖解。一般而言,在对数据进行整合和建档之前,首先要对“数体”的纹理进行细致地分析,这样才能充分掌握“数体”的属性并为后续挖掘其潜在价值打牢基础。每一个人或每一个事物就是一个“数体”或颗粒,世界由无数“数体”或颗粒所构成,而每一个“数体”或颗粒所携带的信息又各不相同,基于上述特征所产生的多样性和差异性正是精准治理的基础。传统的数据分析不考虑构成大数据的单个“数体”的个性和特征,只对整体数据的属性和特征进行分析,并基于整体数据的“平均”特征对症下药,这种一刀切的治理方式缺乏针对性考虑。现代人工智能跳出了“大锅式”“平均式”和“漫灌式”状态,使基于“数体”特征的治理形式变得更加精细化、个性化和精准化。

正如像素是反应图片清晰度的重要指标一样,数据纹理是反应清晰度的核心指标。具体而言,数据纹理具有以下典型特性:①可读性。数据纹理并不是一种神秘的飘忽不定的事物,而是一种可认识、可识别和可读取的信息集合。②高维性。数据纹理并不是一个简单的一维或二维数据,它是一个包含诸多特征的高维数据集合。理论上而言,数据纹理的维度具有无限性,因为认识和测量“数体”的视角和维度理论上是无穷尽的,但也应该注意测量维度存在包含或交叉的状态。③规律性。数据纹理并不是一堆毫无关系、支离破碎的数据集合,而是一种具有一定次序和一定规律的数据集合,但这种数据规律需要通过学习和研究才能被发现。同时,对数据纹理规律性的探索是提升清晰度的重要方式。④独特性。世界上没有两片完全一样的树叶,每个“数体”的数据纹理也是一样,每个“数体”的数据纹理是独一无二的,这主要取决于数据纹理的高维性,高维的无限可能产生了多样性和差异性,而精准治理的实现也正是基于数据纹理的差异性和独特性。

2. 超级档案建立

超级档案是实现精准治理的重要前提和基础,没有精确的数据就无法提取治理所依赖的信息和知识,没有准确的信息和知识就无法实现精准高效的

治理。“数据是信息的载体,信息是有背景的数据,而知识是经过人类的归纳和整理,最终呈现规律的信息”^[24],超级档案就是“数体”在不同时刻不同领域活动所记录的庞大数据集合,也是人类提取信息和知识的资源库。超级档案和传统档案本质上都是一种数据和信息的集合,但超级档案的体量和规模是传统档案无法比拟的,两者最大的区别在于数据规模所引发的聚变和质变。理论上而言,生活在社会中的个体就是一个数据集合,个体的身份、性别、身高、婚否等基本信息,个体的单位、朋友、圈子、交往等关系信息,以及个体的兴趣、喜好、厌恶、倾向等情感信息,等等,这些集拢在一起的“数纹”就是个体颗粒的档案。如果将众多个体颗粒在不同时段的数据进行搜集和汇总则将形成超级档案。如果没有现代大数据和人工智能等信息技术的支撑,很难想象这种超级档案的存在。

超级档案是一种数据集合但并非简单的数据集合,其具有以下特征:①超大型体量。超级档案不是对某一领域的大数据或某类群体的大数据进行搜集和整理,其强调以个体粒度为单位的全方位、全领域和全时段的数据搜集和整合,因此其体量极其庞大。②结构化分类。超级档案中的数据类型具有多样性,既包括具有结构化特征的数据,也包括诸如图片、录音和视频等非结构化的数据,但超级档案中的大数据都经过了专业机构和公司的清洗和处理,具有较强的结构化特征。③连续性更新。超级档案并不是一个封闭的大型数据库,其会根据个体粒度内外部环境的变化而不断更新升级,以保证数据的信度和效度。④有意性关联。超级档案中的数据并不是孤立存在的,而是一种相互关联的复合型数据。只有“通过结合不同的数据进行分析,人工智能才能真正发挥出它的威力,找到人类难以发现的模式”^[25],因此人工智能的算法会刻意将超级档案中的数据进行关联,让数据在复合和叠加中产生最大的价值和威力。

3. 单粒度治理:从模糊性治理到清晰化治理

基于数据纹理特征而构建起来的超级档案犹如一张巨大、动态、清晰的图片,是个体颗粒行为和活动的一种立体投影。单粒度治理就是基于个体颗粒的数据纹理而进行的治理,其具有个性化、高精度和针对性的典型特征。人工智能驱动的精准治理本质上就是一种单粒度治理,其区别于传统的模糊性治

理模式,使得治理对象更加清晰、治理形式更加精准以及治理绩效更加显著。传统的经验治理和现代的技术治理两者在治理逻辑上存在明显差别,传统经验治理是从自上而下的整体性角度基于治理对象的模糊、含混、不确切的信息而作出的一种经验性判断,而现代技术治理是从自下而上的个体性角度基于治理对象的清晰、量化、高精度的数据纹理特征而作出的一种差异化处理。同样是应对不确定性,前者是基于模糊化的信息而采取的策略,后者是基于清晰化的数据而作出的判断。显然,现代大数据和人工智能技术使得信息获取更加容易、信息储存更加便捷、信息加工更加灵活、信息计算更加强大,这些都推动着现代社会从模糊性治理走向清晰化治理。

单粒度治理不仅提升了国家对社会识别能力和控制能力,而且改善了公共服务供给效率和社会资源分配效度,更为重要的是,其为处于复杂环境中的人们提供了一个较为舒适的生活环境和制度结构。单粒度治理的核心机制在于精准,即通过对个体颗粒的数据纹理进行分析和储存,进而转化为信息和知识,再基于这些信息和知识采取针对性对策。它改变了传统上信息不对称的状态,打通了治理主体与治理客体之间的信息阻隔。然而令人意外的是,正是这种被认为给社会带来诸多福利的单粒度精准治理却隐藏着巨大社会风险。从个体特征角度而言,基于个体独特性而展开的单粒度精准治理反而为个体独特性所俘获;从信息传输角度而言,为打破传统社会空间交流障碍和阻隔的单粒度精准治理反过来却制约着社会空间中个体的沟通与交流;从社会福利角度而言,为促成社会福利最大化的单粒度精准治理反过来可能制约社会福利的增加和扩大。由此可见,以单粒度治理为核心机制的精准治理既是效率、个性、精准的代名词,同时也是社会风险的重要起源。

三、人工智能时代精准治理风险的形成

人工智能驱动的单粒度精准治理犹如一柄达摩克利斯之剑,其既具有促进社会福利大幅提升的超强能力,同时也具有摧毁人类社会的巨大风险。那么这种风险到底是什么?它又是如何形成的?

1. 舒适区域形成

人工智能驱动的精准治理激发了人类懒惰、逃避和自私的本性,因为其通过个性化、精细化和精准

化的方式满足了人类的各种需求,使人们处于便利、安逸、舒适、悠闲的生活状态,人们缺乏动力和激情去改变这种便利、舒适和安逸的现状,进而选择了被动地接受各种便利、舒适和安逸。在人工智能时代,个体所产生的数据越多,其数据纹理就越清晰,而越清晰的数据纹理则可以促进更精准的治理。如在网络购物中,“一个人越多地使用这种服务,他就能获得更好的推荐,用户的行为训练了 AI 程序,帮助它们更好地了解用户喜欢什么。”^[26]另外,加入的个体越多,产生的数据也越多,从数据中提炼出的数据纹理和数据维度也越多,而基于大量数据纹理和数据维度的计算准确度也会大幅提升。“使用的人多起来以后,数据会越积越多,学习了大量数据的 AI 也会变得越来越聪明,我们用起来也会越来越便利。”^[27]由此可见,人工智能驱动的精准治理在追求治理绩效的过程中将人们带进了具有个性偏好和个性关怀的便利、舒适和安逸区域。

人工智能的优势是基于既有的数据纹理进行目标性治理,其缺乏对伦理价值、人道主义和同情怜悯等高度复杂思维的理解。“人与机器之间有一个无法逾越的鸿沟,即使我们可以在机器上模拟某种脑电波,可那根本不能构成像人一样的思维,人的思维是发生在心灵之中的。”^[28]因而人工智能驱动的精准治理实际上是一种功利导向的目标管理,支撑其背后的逻辑是人的懒惰、逃避、自私、不愿负责和追求快乐等本性。“那么会不会出现目标被设定为确保人类的快乐,结果机器智能就把人放进营养液,并且不间断地刺激人的微笑神经这种事情?答案是肯定的!”^[29]由此可见,与其说人工智能驱动的精准治理给人们带来的是舒适和便利,不如说是一种刻意的取悦。这样个体将处于一片祥和的氛围之中,似乎世界上的一切都按照个体的需要而得到妥善的安排。个体变得不思进取,沉浸在自我欣赏、自我认同和自我陶醉之中。在这种情况下,实际上个体的真实需要被个体的虚假需要所压制,个体的真实需求是追求社会进步和社会福利,而不仅仅只是技术进步带来的迎合个体本性的舒适和便利。因此,人工智能驱动的精准治理最容易造成虚假需求压制真实需求的社会风险。

2. 部分封闭开启

不合理的成分经常深深地隐藏于合理性之中,人工智能驱动的精准治理风险也是如此。人工智能

本身只是一种价值中立的信息工具,但它的这种工具性已经被肯定的性质所浸染,这种肯定的性质就是符合个体本性的价值体系。人工智能之所以受到人们欢迎,正是因为它所筹划的是在实践上顺应个体各种目的的东西。因此,人工智能驱动的精准确治理实际上内嵌着一个过滤机制或选择机制。这个过滤机制或选择机制正是根据个体需要将不符合个体本性的东西逐渐排除在外,剩下的都是符合个体喜好、个体欲望和个体倾向的东西。如果这种过滤机制或选择机制一旦建立,则意味着个体认知的半封闭机制已经开启。这种半封闭机制会自动屏蔽掉个体不喜欢的、不符合预期的、不能满足需求的信息,接受那些肯定自我的、强化自我的、美化自我的东西,进而使个体陷入一种单向度世界。更重要的是,一旦半封闭机制开启就很难停下来,因为在过滤机制和选择机制的强化作用下个体很难适应和接受具有多向度的真实环境。

这种半封闭机制并不只是一种单一的主体行为或客体行为,而是一种主客体互动行为。一方面,精准确治理的主体会根据客体的需求以及自身的目的来筛选和过滤信息,并把这种信息精准地推送给治理客体;另一方面,治理客体也会在治理主体所提供的信息中获取快乐和满足,并进一步向治理主体反馈其喜好倾向。正是在主客体的刺激-反馈循环中不断筛选和过滤信息,使个体接收到的基本上都是对其喜好持肯定性质倾向的信息。更重要的是,治理客体也有可能在治理主体不断推送的信息和服务中逐渐形成对该治理模式和治理价值的认同。这种半封闭机制中的主客体互动愈强烈,则单向度思维将会变得愈加纯粹,因为通过多轮次和高频率的筛选和过滤,治理客体的喜好和价值倾向基本能够被固定在限定范围,而这一限定范围实际上就是客体的舒适圈。一旦个体陷入舒适圈,其思维、认知和行动都将受这一圈内的单向度价值牵引,直至个体觉醒或者外部介入才能打破这层舒适圈,否则很难跳出该圈。

3. 单向度悖论:从崇尚理性到理性弱化

批判能力是理性思考的核心机制,如果社会缺乏批判能力,那么社会将失去理性。半封闭机制开启在某种程度上意味着批判能力的弱化,而批判能力的弱化则是社会进步停滞的开始。人工智能驱动的精准确治理给个体提供舒适服务和环境的同时,意

外地开启了个体信息的半封闭机制,使个体倾向于接受能带来快乐、喜悦和舒适的肯定性信息和知识,而正是这种肯定性信息和知识使个体陷入单向度世界。如在网络平台浏览新闻时,网络平台会根据个体的浏览记录、停留时长、评价内容等信息为其推送符合其兴趣和喜好的新闻,短时间内这种推送方式能够满足个体的兴趣欲望,但如果长时间陷入这种类似性新闻中,实际上就限制了个体接触信息的范围,使其在个人兴趣爱好中不断重复和徘徊,进而陷入单向度世界,甚至弱化和失去批判能力。在这一过程中,人工智能驱动的精准确推送服务只是为了更好地满足个性化的信息需求,然而正是这种看似合理的行为中却隐藏着极其可怕的不合理。

由此可见,人工智能驱动的精准确治理容易使社会陷入单向度悖论,即个体在追求目标的过程中反而为目标所俘获。实际上,“我们的目的不是用机器替代人,而是利用人工智能通过分析大规模数据最终找到那些难以发掘的规律那样的无与伦比的能力,来增强人类的能力”^[30],而人工智能驱动的精准确治理却恰恰存在于人类追求这种能力的过程中弱化这种能力的风险。本质上而言,人工智能只是一种价值中立的技术工具,这种技术工具受嵌入其中的价值所牵引,但意料之外的是这种价值在技术工具的作用下竟然走向自我弱化。这正如帕特所言:“人工智能真正危险的是让不具备常识的系统做出需要依靠常识判断的决定。”^[31]人工智能并不知道嵌入其中的价值要利用它做什么,它只知道按照既定的规则进行运作,正是这种价值中立的工具性在没有理解价值内涵的情况下将事务引向了危机。总之,人工智能驱动的精准确治理很可能将个体引向单向度泥沼,进而诱发批判能力弱化、理性思维变得不理性以及社会整体停滞不前等风险。

四、人工智能时代精准确治理风险的放大

以单粒度为核心的精准确治理机制推动现代社会治理从模糊性治理走向清晰化治理,然而正是这种蕴含巨大潜力的单粒度治理却隐藏着巨大社会风险。单粒度治理意外地引发单向度悖论,导致个体在追求目标过程中反而为目标所俘获,进而诱发人工智能驱动精准确治理风险的形成。那么这种风险将通过什么途径得以放大?又将把我们带向何方?

1. 单一输入束缚

人工智能驱动的精准治理风险主要是在外部条件受到约束而内部又无法自身突破的双重困境中得以放大的。人工智能驱动的精准治理为个体提供舒适、安逸和快乐的生活环境和制度环境,在这种环境中个体倾向于被动接受人工智能推送的迎合个体需求的各种信息,这样导致个体接收的信息实际上经受了筛选和过滤,最终进入个体视野的信息是符合个人喜好的狭小范围内的信息,最终造成个体陷入单向度世界。单向度世界中的个体认知能力受到限制、批判能力变得弱化,最终将因此而引发各种社会风险。人工智能驱动的精准治理风险继续放大的原因之一是外部输入的信息逐渐单向度化,即人工智能倾向于根据个体喜好特征源源不断地推送信息。这实际上已经开启了信息输入的半封闭机制,人工智能将根据个体喜好将其憎恶的东西排斥和过滤掉了。由此可见,如果长时间处于这种模式下,个体将陷入一种“无发展的增长”状态,批判能力也将逐渐弱化和丧失。

更重要的是,这种半封闭机制将在以下几个因素下进一步强化。一是人工智能的工具性。“1986年,约翰·希尔勒(John Searle)进行了一项名为‘中文房间’的思维实验,证明了能够操控符号的计算机即使模拟得再真实,也根本无法理解它所处的这个世界。”^[32]也即人工智能不能理解目标,只能为特定的目标服务,正是这种工具性强化了半封闭机制。二是人类的趋利避害特征。以市场营销为例,市场营销自动化的想法主要起源于集客营销和一对一营销两种理念^[33],这两种营销模式实际上都是根据个体的趋利避害的特性而建立的。趋利避害往往使个体追求对自己有利的事物而排斥对自己不利的事物,这样就进一步强化半封闭机制。三是半封闭机制的隐蔽性。个体往往从现实生活出发,倾向于计较微观上的利益得失而忽视宏观上的问题和危机。半封闭机制建立在个体寻求幸福的理性机制之上,因而个体也很容易忽视嵌入这种理性机制之中的非理性因素,也正是这种容易引发个体忽视的隐蔽性强化了半封闭机制。

2. 内在认知固化

人工智能驱动的精准治理风险放大的另一个重要原因是治理客体本身在认知上产生了固化效应,即个体对这种半封闭机制产生了认同和依赖。半封

闭机制带给个体的是与个体价值倾向和情感需求相一致的东西,个体在这种共同的价值认同和行为互动中逐渐融入这种价值。“它使我们更加关注那些可以强化我们现有观点的数据和想法,忽视其他反对它们的数据和观点”^[34],这样就让个体陷入“注意力盲区”。个体在半封闭机制中将注意力主要集中在自身感兴趣的和不断被接受的事物方面,从而忽视了那些自身不感兴趣的东西,久而久之,这种状态就会使个体陷入自我陶醉、自我欣赏和自我认同的狭隘范围,甚至产生认知上的刻板效应。更重要的是,如果个体在半封闭机制影响下完全信任被推送的信息,一旦这种被推送的信息是恶意的或虚假的,那么其将对个体形成巨大威胁。如利用大数据的“杀熟”和基于用户使用痕迹的精准诈骗就是典型的例证。

基于半封闭机制的固化效应强化了人工智能驱动的精准治理风险,这种自我强化主要体现在巨大启动成本、学习效应、协调效应、适应性期望4个方面^[35]。首先,巨大的启动成本。个体已经习惯于精准治理所带来的福利和舒适,如果要打破这种半封闭机制就必须打破人工智能驱动的精准治理所带来的“舒适圈”,这对于个体而言无疑是一种强大的成本。其次,学习效应。个体已经适应于人工智能驱动的精准治理所带来的便利和舒适,并已经习得了如何更好地适应这种便利和舒适。这种既有的学习成本也会强化半封闭机制。再次,协调效应。个体在接受人工智能推送的相关信息后,对这些信息产生了认同,并且在反复的信息接受中不断使单向度的东西变得更加精致化,个体与模式之间已经经过充分协调,这也强化了封闭机制。最后,适应性期望。个体在单向度事物和信息中使自己的观点和喜好等不断受到肯定和印证,这也将使个体沉浸在虚假乐观中,进而进一步强化封闭机制。

3. 内卷化困境:从追求卓越到陷入平庸

外部输入受到约束和内部认知变得固化的双重作用使人工智能驱动的精准治理陷入内卷化困境,即个体在半封闭机制中不断使自身变得更加精致化,但又难以突破自我和改变现状的状态。具体而言,这种内卷化困境主要表现为3种形式:①无效率。半封闭机制开启后,虽然仍会有源源不断的信息输入,但这种信息是经过筛选和过滤的单向度信息,其只能导致信息的累加,而不能从这些输入的信

息中进行有效地创新。这就导致虽然个体在忙忙碌碌,也从这种忙忙碌碌中收获了舒适和快乐,但这种忙碌往往是低效的和无意义的。②被锁入。在半封闭机制作用下,个体接收到的都是能够满足自身需求的信息,这就使得个体在满足中失去改变现状的动力,进而陷入一种自我满足的锁入状态。③强惯性。在半封闭机制作用下个体沿着自身的兴趣和喜好而不断发展和卷入,在这种模式的强大惯性支配下个体很难摆脱特定路径,进而在特定路径中陷入忙碌的无效率。

由此可见,追求卓越的人工智能沦为了人工愚蠢,曾经的雄心壮志却只能化为庸庸碌碌。“‘人工愚蠢’(Artificial Stupidity)可以定义为计算机学家们为开发各种程序所做的尝试,这些程序会导致某种类型的问题,这些问题通常都与人类思维有关。”^[36]人工智能驱动的精准治理之所以陷入内卷化困境,原因看似是技术治理本身发展所涌现出的问题,实际上这一问题仍然深深地植根于人类思维本身。人类的思维本质上是为改善社会福利所服务的,一旦人类福利得到大幅度改善,这种福利所带来的舒适就会反过来遏制这种福利的发展,只有当这种遏制所造成的危机状态足以威胁人类安全的时候,这种舒适才会被打破,第二轮发展才会得以继续。因此,社会每发展一段时间总会经历或长或短的停滞、无效和锁入等危机,这也说明社会发展是一个螺旋式上升的过程,而不是一种一蹴而就的交易。总之,人工智能驱动的精准治理不可避免地会将个体从卓越带向平庸,我们要做的就是不要让这种平庸占据太久的时间。

五、结 语

人工智能驱动的精准治理已经给人类带来巨大的收获和喜悦,但在这种收获和喜悦背后却隐藏着巨大的危机和风险。基于具有可读性、高维性、规律性和独特性的数据纹理以及由众多具有数据纹理特征的个体颗粒所构成的超级档案是单粒度治理的重要基础,而人工智能驱动的精准治理实际上就是一种单粒度治理,这种治理模式能够改善国家对社会控制能力和公共服务供给的绩效,为人类提供一个较为舒适的生活环境和制度环境,但这种单粒度精准治理同时也隐藏着巨大的社会风险,是人工智能时代精准治理的风险起源。这种风险就是单向度

悖论,即个体在追求目标的过程中反而为目标所俘获。半封闭机制是产生这种风险的重要路径,其通过过滤机制或选择机制将不符合个体本性的东西逐渐排除在外,接受那些肯定自我、强化自我和美化自我的东西。这种机制将导致个体批判能力弱化和丧失,进而诱发各种社会问题。更重要的是这种风险在外部输入受到约束和内部认知变得固化的双重作用下,使人工智能驱动的精准治理陷入内卷化困境,造成个体陷入一种不断追求精致化却又难以突破和超越自我的平庸状态。

令人意外的是,人工智能驱动精准治理的初衷是提高人类社会的福利,最后却抑制了福利的增长;其崇尚理性批判,最后却弱化了理性批判;其追求卓越品质,最后却陷入了平庸。由此可见,任何合理的事物中都孕育着不合理,而正是这种合理的正当性让我们忽视了其中潜藏的不合理。“技术创新必定是复杂的,我们必须接受它是一把‘双刃剑’这一事实,我们必须在探索 and 了解宇宙的同时,承受道德模糊甚至是恶意使用技术的可能性。”^[37]事物发展都具有两面性,我们必须平衡好这种两面性。同时,治理技术的发展不能离开治理内容,要平衡好“道”(治理理念)与“术”(工具意义)的关系,处理好权力与责任、权利与义务以及公共性与民生福祉之间的关系^[38]。人类社会的进程是一个发展与停滞相互交织的螺旋式上升过程,在取得进步的同时必然隐藏风险,在巨大风险中也必然蕴含新机。人工智能驱动的精准治理不可否认地会带来成功和舒适,但也不可避免地会带来危机和庸碌。因此,破解人工智能驱动的精准治理风险必须采取辩证的思维,既要接受其带来的成功和福利,也要承认其带来的停滞和损害,在此基础上预防和警惕风险,让损害降低到最低程度。

参考文献:

- [1] 松尾丰,盐野诚.大智能时代:智能科技如何改变人类的经济、社会与生活[M].陆贝旒,译.北京:机械工业出版社,2015:8.
- [2] 涂子沛.数文明:大数据如何重塑人类文明、商业形态和个人世界[M].北京:中信出版社,2018.
- [3] KEVIN C D, BENOY J. Big data in the public sector: lessons for practitioners and scholars[J]. Administration & Society, 2017, 49(7):1043-1064.

- [4] NICKY R, TOMMASO A, KRISTOF D W. Big data and the measurement of public organizations' performance and efficiency: the state-of-the-art [J]. Public Policy and Administration, 2017, 32(4): 263-281.
- [5] BERND W W, WILHELM M M. An integrated artificial intelligence framework for public management [J]. Public Management Review, 2019, 21(7): 1076-1100.
- [6] 李利文, 易承志. 公共服务供给的共享制创新: 类型、风险及其规避 [J]. 人文杂志, 2019(5): 109-118.
- [7] 王玉龙, 王佃利. 需求识别、数据治理与精准供给——基本公共服务供给侧改革之道 [J]. 学术论坛, 2018(2): 147-154.
- [8] 陈潭. 大数据驱动社会治理的创新转向 [J]. 行政论坛, 2016(6): 1-5.
- [9] 张海波. 大数据驱动社会治理 [J]. 经济社会体制比较, 2017(3): 64-73.
- [10] 庞金友. AI 治理: 人工智能时代的秩序困境与治理原则 [J]. 人民论坛·学术前沿, 2018(10): 6-17.
- [11] 李大宇, 章昌平, 许鹿. 精准治理: 中国场景下的政府治理范式转换 [J]. 公共管理学报, 2017(1): 1-13.
- [12] 段忠贤, 沈昊天, 吴艳秋. 大数据驱动型政府决策: 要素、特征与模式 [J]. 电子政务, 2018(2): 45-52.
- [13] 汪磊, 许鹿, 汪霞. 大数据驱动下精准扶贫运行机制的耦合性分析及其机制创新——基于贵州、甘肃的案例 [J]. 公共管理学报, 2017(7): 135-143.
- [14] 宁靓, 赵立波. 政府购买公共服务精准化的大数据应用模式研究 [J]. 山东大学学报(哲学社会科学版), 2018(3): 150-158.
- [15] 蒋林秀, 李泉. 基于机器学习的群租房精准治理研究 [J]. 公共管理与政策评论, 2019(2): 85-96.
- [16] 单勇. 基于犯罪大数据的社会治安精准防控 [J]. 中国特色社会主义研究, 2016(6): 54-61.
- [17] 曹树金, 刘慧云, 王连喜. 大数据驱动的图书馆精准服务研究 [J]. 大学图书馆学报, 2019(4): 54-60.
- [18] DAVID M H, EVAN R K, JUSTIN L, et al. Toward precision governance: infusing data into public management of environmental hazards [J]. Public Management Review, 2018, 20(5): 746-765.
- [19] 黄欣荣. 大数据时代的精准诈骗及其治理 [J]. 新疆师范大学学报(哲学社会科学版), 2017(4): 86-92.
- [20] 维克托·迈尔-舍恩伯格, 肯尼斯·库克耶. 大数据时代: 生活、工作与思维的大变革 [M]. 盛杨燕, 周涛, 译. 杭州: 浙江人民出版社, 2013: 16.
- [21] 赫伯特·马尔库塞. 单向度的人: 发达工业社会意识形态研究 [M]. 刘继, 译. 上海: 上海译文出版社, 2014: 12.
- [22] 刘世定, 邱泽奇. “内卷化”概念辨析 [J]. 社会学研究, 2004(5): 96-110.
- [23] 芭芭拉·亚当, 乌尔里希·贝克, 约斯特·房·龙. 风险社会及其超越: 社会理论的关键议题 [M]. 赵延东, 马缨, 译. 北京: 北京出版社, 2005: 3.
- [24] 涂子沛. 数据之巅: 大数据革命, 历史、现实与未来 [M]. 北京: 中信出版社, 2019: 298.
- [25] 日本日经大数据. AI 商业时代: 人工智能如何改变商业、就业和社会 [M]. 杨玲, 译. 北京: 机械工业出版社, 2017: 153.
- [26] 凯瑟琳·赫利克. 人工智能 [M]. 邹慧君, 梁庆华, 译. 上海: 上海科学技术出版社, 2017: 16.
- [27] 藤野贵教. 2020 年人工智能时代: 我们幸福的工作方式 [M]. 崔海明, 译. 北京: 机械工业出版社, 2018: 152.
- [28] 孟令朋. 论维特根斯坦关于人工智能的基本观点 [J]. 山东社会科学, 2012(5): 153-156.
- [29] 李智勇. 终极复制: 人工智能将如何推动社会巨变 [M]. 北京: 机械工业出版社, 2016: 86.
- [30] 沈向阳, 施博德. 计算未来: 人工智能及其社会角色 [M]. 北京: 北京大学出版社, 2018: 97.
- [31] 赫克托·莱韦斯克. 人工智能的进化 [M]. 王佩, 译. 北京: 中信出版社, 2018: 165.
- [32] 理查德·温. 极简人工智能: 你一定爱读的 AI 通识书 [M]. 有道人工翻译组, 译. 北京: 电子工业出版社, 2018: 81.
- [33] 野村直之. 人工智能改变未来: 工作方式、产业和社会的变革 [M]. 付天祺, 译. 北京: 东方出版社, 2018: 191.
- [34] 卡鲁姆·蔡斯. 人工智能革命: 超级智能时代的人类命运 [M]. 张尧然, 译. 北京: 机械工业出版社, 2017: 110.
- [35] 道格拉斯·C. 诺斯. 制度、制度变迁与经济绩效 [M]. 杭行, 译. 上海: 上海人民出版社, 2008: 129.
- [36] 雷·库兹韦尔. 机器之心: 当计算机超越人类, 机器拥有了心灵 [M]. 胡晓姣, 张温卓玛, 吴纯洁, 译. 北京: 中信出版社, 2016: 83.
- [37] 阿米尔·侯赛因. 终极智能: 感知机器与人工智能的未来 [M]. 赛迪研究院专家组, 译. 北京: 中信出版社, 2018: 53.
- [38] 周庆智. 道术之辨: 大数据治理的原则和边界——以基层社会秩序变革为中心 [J]. 学海, 2019(5): 107-113.

(责任编辑: 吴玲)