

· 专辑: 人工智能与情报学 ·

人工智能应用的风险评估与应对策略 *

蒋 洁

(1. 南京信息工程大学中意网络侵权研究所 江苏南京 210044)

(2. 南京信息工程大学法政学院 江苏南京 210044)

(3. 中国人工智能产业发展联盟 北京 100191)

摘 要: 大数据生态中多元应用社会生活的人工智能技术,在勾勒健康恒存的愿景框架、推进开源聚力的知识传递、驱动协作共享的产业变革以及实现精准递归的社会治理的同时,也逐渐暴露出算法决策内嵌安全隐患、责任分散置险族群永续、精英趋向引发异化强权以及智能迭代侵害人格权益等诸多风险。亟待确立人工智能应用有益族群永续的立法目标,遵循技术进步、经济增长和公众权益交互规律,平衡国家安全、社会稳定、企业发展与个人成长的矛盾冲突;构筑张弛有度的政府监管机制、颁行透明可信的数据核馈规范、完善有区别的责任制度、实施适应全新需求的层级税制并加强人工智能领域的宣传教育,加快提升中国的软实力和巧实力,积极建设创新型科技强国。

关键词: 人工智能;多元应用;情报交流;风险表征;政策法律

中图分类号: TP18; G203

文献标识码: A

DOI: 10.11968/tsyqb.1003-6938.2017122

Risk Assessments and Countermeasures of AI Applications

Abstract In the worldwide surge of digitization, artificial intelligence is applied in all aspects of social life. On the one hand, they have contributed to improve the vision of good health, push forward the process of open source knowledge transmission, promote the collaborative sharing industrial transformation, and realize the precise and recursive management of society. On the other hand, there are a number of serious risks involved, such as hidden dangers in algorithm decision, the survival crisis of mankind caused by the diffusion of responsibility, alienated power raised by elite tendency and personal rights violated by iterative intelligence. The article proposes to establish the legislative purpose of obeying the interactive law among technical development, economic growth and social stability, equating the collision of national security, social stability, business development and personal growth, and helping ethnic sustainability through intelligent technology research and development. It's especially urgent to make regulations of the balanced government supervision, the transparent and credible data auditing and feedback system, the distinguishing accountability mechanism, the tax mechanism satisfying new needs. Above countermeasures would help to rapidly enhance China's soft power and smart power, and build an innovative technological power.

Key words artificial intelligence; multiple applications; information exchange; risk characterization; policy and law

信息纪元中迅速提升的大数据处理技术、开源互补的广域算法、持续作用的香农定理和摩尔定律以及人类族群探索未知的执意和趋利避害的本性合力引领历经两轮大幅起落的人工智能技术作为数字生态关键要素再获新生,不仅构筑共生、竞争、合作、依存的协同进化生态圈,亦在情报研判、决策支持与

劳动替代中暴露出诸多风险。亟需采取一系列行之有效的应对策略,积极推动我国人工智能战略的革命性发展。

1 人工智能多元应用的情状画像

万物感知的爆发式增长推动人工智能技术迅速

* 本文系国家社科基金项目“云数据隐私侵权问题研究”(项目编号:13CFX083)、司法部国家法治与法学理论研究项目“大数据背景下的电子取证问题研究”(项目编号:15SFB3018)、国家社会科学基金特别委托项目“大数据治国战略研究”(项目编号:15&ZH012)、中国计算机协会——腾讯联合科研基金项目“人工智能的隐私侵权风险与法律应对研究”(项目编号:CCF2017-01)阶段性研究成果与江苏省“333 人才工程”研究成果之一。

收稿日期:2017-12-15;责任编辑:魏志鹏

摆脱“计算机展示的一系列同质智能行为”^[1]的负面标签。智联图景中能效率次递增的海量数据集成训练样本嵌入神经网络,大幅拓展人类思维界域并增强群体行动量能,助力整个社会迎来跨越式发展的起飞点。多元异质的崭新智能应用全方位颠覆国家安全情报构成与交流过程,积极推进科学、个性、精准的嵌入式业态革命,逐步构筑良性循环的智能可再生能源系统,产生巨大经济效益与社会影响力。相关应用的情状画像内蕴勾勒健康恒存的愿景框架、推进开源聚力的知识传递、驱动协作共享的产业革命以及实现精准递归的社会治理等点睛之处。

1.1 勾勒健康恒存的愿景框架

寒武纪大爆发的实力出现营建的全新生物对抗环境倒逼防卫竞争下多样生物协同进化。信息互通场景中遍布全球的物理传感载体实时监控巨量节点、互链数据资源池自动整理并优化组合海量信息、多层微化算法积极洞察复杂社会机能。诸多原力聚合跃迁长期停滞的智能机器识别能力与覆盖空间,驱动人类族群加速勾勒增强自我健康的宏大愿景。

人工智能多元应用积极抽取并持续更新生命体征情报图谱,充分发挥循证医学在理想状态下高效挖掘与传递疾病规律的核心优势,逐渐开辟超越资深专家水平的全维度个性化诊疗局面。疾病检测领域,沃森智医系统通过集中检测海量图片数据归纳糖尿病性视网膜病变和黄斑水肿等致盲因素的初发迹象^[2];疾病治疗领域,初步智能化的达芬奇手术系统已经辅助完成300多万例微创手术^[3]。事实上,从“23与我”公司有重点的基因芯片筛查服务到核林公司尝试通过神经假体引领的脑细胞通信协助认知缺陷群体,几何乘数拓展的人工智能应用正在迅速拉升整个社会的泛在健康指数,拉开璀璨的永生序幕,“把生命稍纵即逝的光华写进我们的遗传密码,从此成为永恒”^[4]。从体内芯片情报介质、入脑起搏器、植入式温度计等起步的旨在结合生物潜能和机器优势

塑造具有无穷适应力的半机械人的探索活动正在开拓可遗传栖变下的恒存通路。

1.2 推进开源聚力的知识传递

随着制造、发售和维护实时信息捕析设施的成本骤降、计算能力大幅提升和云端存储巨量扩容,持续汇总的密集数据为打破纵横割裂的信息孤岛奠定了坚实基础。大数据环境能够规避非注意盲视^①的不良影响、运用人工神经网络架构演算识别和精准整合隐嵌用户终端的海量知识原点而形成睿智洞见的多样智能应用,积极打造人机融合的个性化知识传承网络,开拓重建巴别塔的绿色路径^②。如明略数据在全国落地了大量智能化的公安情报知识图谱,准确高效地研判警务情报,切实维护公共安全。又如,如火如荼的人工智能图书馆建设实现了个性化推荐推送,积极强化馆藏稀缺资源的延伸性,全面整合碎片信息并推进读者互动,逐渐形成强大的知识流转格局。

1.3 驱动协作共享的产业变革

从蒸汽机、计算机到人工智能的技术跃迁递次趋零商品和服务的边际成本,“催生出新的业态和商业模式,引发产业结构的深刻变革”^[5],日渐构筑数据驱动个性定制、人机协同生产服务和跨界共享产品价值的智能经济形态,加快深化供给侧结构性改革。众多企业基于用户活跃度、权威度与兴趣方向等动态画像进行集优销售、信用预测与分类补偿等等。如人工智能技术广泛应用于发行、印刷、数字加工与阅读等领域,通过机器协助写作校对和智能增强交互体验等重塑新闻出版行业;又如,综合运用精准对象检测与追踪技术、生物特征自主感知技术以及意图识别和交易技术等积极开发具有高识别力、高触达力、高洞察力和高服务力的新零售实践样本,如Take Go、淘咖啡等无人值守智能门店。此外,自动驾驶技术尝试嵌入共享交通工具等现象进一步激励了“使用而不购买”的新兴数字消费方式,节约巨量资源,促进环境改善。

①非注意盲视(inattention blindness)主要是由于注意力分散或注意任务繁重而在处理相对复杂的任务时,不能察觉到背景中出现的醒目意外刺激的现象。

②《圣经·旧约·创世记》第11章记载,当时语言相通的全球人类齐心协力地意图建造通往天堂的高塔。上帝担忧人类若是真能修成宏伟的通天塔,或许将统一强大、无所不能,就改变并区别了人类的语言,阻止人类相互之间有效沟通,计划因此半途而废。

1.4 实现精准递归的社会治理

《新一代人工智能发展规划》(国发〔2017〕35号)中指出,“人工智能技术可准确感知、预测、预警基础设施和社会安全运作的重大态势。”整合大数据、云计算、深度学习与虚拟现实的集智平台充分发挥社会治理最优解的隧道效应,以结合环境复合因子的智析节点为“瓦斯报警鸟”^①,审慎聚化、分类、压缩和处理万物互联中众多传感器生成数据,精准权衡按需数据搜索、标记和使用的方差偏离和计算负荷,批量保真并递归提升数质精度,近看助力提升国家安全防卫、完善灾难响应方式、解决人口老龄化与教育平权化等难题,远期助力人类族群摆脱资源环境束缚。如融合图像识别、语音识别、自然语言处理等创新技术的自动情报生成系统有助于突破传统情报收集和流转的真实性、透明性与全面性障碍,俨然是颠覆性的公共安全保障力量。

2 人工智能多元应用的风险表征

随着经济奇点下爆发式扩张的人工智能应用持续展现出精准、高效、节能等优势,各级公权机构、海量社会组织、众多企事业单位、乃至广大居民开始依赖复合影响变量的权重计算进行评估和决策,引发了侵害生命健康权、受教育权、就业权、平等权、婚姻权以及参与政治生活和社会生活的权利等灾难性负面效应。

2.1 算法决策内嵌安全隐患

基于深度学习引擎在特定场景进行智能推理的全新范式改变了人类族群自由抉择的传统模式,看似科学客观的算法决策内嵌着形形色色的故障因子。事实上,非正常性黑盒输入导致智能应用的处理与输出模块处于不可知的自主状态,过度依赖机器决策存在巨大的安全隐患。

(1)公权机构、企事业单位、其他社会组织乃至

众多居民在近因效应作用下盲目认为依据短时尺度的数据体验就能做出准确的趋势判断,往往积极删除缺乏新颖性和时效性的信息资源,导致运用人工智能技术分析环境变迁、国家安全、公共事务、市场趋势以及特定个体状况等情报信息时采获的近期数据过于庞杂,与之对应的远期数据寥寥无几。缺乏完整的持续数据集聚致使宏大规模的过滤数据质量低下,相应的算法决策(尤其是少见且难以预测的极端事件中情境理解与准确抉择)的真实性与可靠性大打折扣。

(2)参与人工智能运作中学习样本拣选与标注的从业人员本身就可能存在种族歧视、性别偏见、宗教倾向或多元价值观,甚至通过不经意输入错误或恶意微调细微因素等欺骗人工神经网络。由于人工智能技术的一次突破意味着无穷复制的替代物,渗透在算法框架中的固有偏见与歧视必会随着规模化发展扩散到诸多应用领域。

(3)目前投入应用的弱人工智能尚处于技术发展初级阶段,几乎不可能进行逆向操作的算法运作不断暴露出容易导致安全隐患的技术缺陷。如具有锁定潜在不良行为强大能力的恶意购物智能筛查决策程序存在因错误跟踪辨识用户行为而限制或拒绝交易,进而影响营销关系的情况;又如,众多黑客不仅可以自主学习识别算法框架的偶发错误,甚至可以利用已有的受信代码预测智能系统未来更新可能出现的零日漏洞^②,进行最大程度的安全渗透。

2.2 责任分散置险族群永续

图灵测试迄今的众多人工智能技术水平量度方案相继揭示智能应用难以突破莫拉维克悖论^③而达至获得心灵并获取思想的超人工智能,阅面(Read-Face)、蓝眼(Blue Eyes)等嵌入互动智能硬件的情感认知引擎本质上是以简单方式认知世界的单色符号表征,但诸多合并人类智慧和机器智能的人机协作

①金丝雀对瓦斯气味非常敏感,即便极其微量的瓦斯,也会停止发声;而当瓦斯含量达到一定限度(人类尚不能察觉的量值)就会毒发身亡。17世纪采矿设备简陋的情况下,英国工人在每次下井时会带上一只金丝雀作为“瓦斯检测指标”。

②零日漏洞(zero-day)是当前科技高速发展背景下随着黑客群体对安全漏洞的知识提升与技术共享,一经发现后立即被恶意利用的安全瑕疵,往往具有更大的突发性和更强的破坏力。

③莫拉维克悖论认为,人工智能实现逻辑推理等人类高级智慧只需要较少计算资源,而实现感知和运动等低级智慧却需要巨大计算力。

模式已经能够独立感知基本情感颗粒和识别陌生生物体,甚至通过蜂巢思维将习得成果汇聚中央知识库并实现全域共享。技术迭代下持续趋向自动化的人机协作中主动伪造数据与被动黑客劫持等现象愈加隐蔽,零乱不堪且疏漏重重的相应责任机制进一步刺激参与者不择手段地汲取跨物种、跨代际的物化生命价值与量化环境对价,严重威胁人类族群健康永续。如在高效的人工智能技术应用与多方位的规范巢白激励控制权让渡的新形势下,外科医生使用机械臂手术和驾驶员使用辅助驾驶系统等大幅削弱参与人员对于关键环境要素的识别和反应能力,甚至产生智能系统可以自动弥补全部遗失错漏的错觉,进而引发严重的伤亡事故。

2.3 精英趋向引发异化强权

大数据联姻深度学习背景下人工智能应用中高度聚合的数据资源与高新技术正在派生掌握绝对资源优势的精英特权阶层,必然加剧人际异化、经济危机和社会动荡,拦阻整个社会迈向应然王国的步伐。随着智能机器人和算法框架愈加集中于少数精英群体,巨额收益亦飞速向其靠拢。持续异变的分配模式甚至逐步解体现有社会结构,导致技术精英日渐占据社会强势地位。然而,绝大多数技术专家与杰出的应用人物类似大航海时代开疆拓土的哥伦布和达伽马,相对缺乏宏观的政经洞见与审慎态度,不仅助力技术发展、知识传递、经济增长与环境改善,亦可能疯狂地推进整个世界的黑盒化。

同时,人工智能多元应用实现创意人才向上流动,日渐优化精英阶层的生存环境。如昂贵的高精尖智能运载工具不仅节约时间并延长出行里程,更突破了传统交通运输根深蒂固的4D风险^①,意味着更高的安全系数,在一定程度上赋予精英阶层在任何其他生命面前的优先通道,严重侵害弱势群体的基本生存权和发展权;又如,通过基因编辑技术微调关键致病因子的昂贵智能医疗铺开应用意味着精英阶层健康生命的大幅延长,使得收入成为寿命的最佳预测指标,反向加剧贫富鸿沟,甚至可能影响整个族群有序繁衍。

人工智能广域渗透导致稀缺资源苛刻分配、社会上升阶梯大面积断裂,逐渐剥夺低收入、低水平且前景有限的弱势群体的生存空间。如掌握覆盖整个市场智能监控系统的权威精英拥有上帝视角,通过设置准入门槛和预警系统抑制竞争,或者采取“击毙式收购”方式消灭潜在对手,变相剥夺其他群体的发展机会,形成后现代的马太效应生态圈;又如,现代社会中劳动者享有不因种族、性别、宗教、外形、性取向和政治见解等遭受不同待遇的平等就业权。有意义的工作是个人自我认同感的重要组成部分,是人生意义的构成要素。伏尔泰曾言“工作撵跑了三大魔鬼:无聊、堕落和贫穷”,失业率较高的国家和地区中滥药、自杀和精神患病概率均较高。人工智能多元应用的高速发展导致整个社会就业结构变迁,秘书、快递员、操盘手、保险定损员、图书管理员、一线制造工人等具有常规性和重复性特征的工种逐渐被机器替代(如多家高校图书馆已经用智能管理员取代传统人工),数以亿计丧失经济价值的失业群体难以自力打通重入职场的渠道。不同于圈地运动之后大工业生产需要海量雇佣工人(彼时普遍用工荒导致整个社会加大职业技能培训力度和辐射面),目前人力资源相对冗余、具有创新性或需要复杂沟通的岗位人员几近饱和,强势精英阶层缺乏主动投资全民职业能力培养的动力,大幅压缩了弱势群体的生存空间与上升通道。

2.4 智能迭代侵害人格权益

人机物三元融合发展进程中,数据成为驱动经济增长的关键要素,数字化资讯积累成本与周期耗费是能否最大限度释放智能红利的关键。如立足基因分析搭建生命周期监控与疾病预防机制的智能诊疗系统中内蕴海量受体遗传信息的数据资源池助力持有人在漫长而封闭的研发周期中逐渐打造行业技术门槛,具有巨大的使用价值和交换价值。“23与我”公司能够与辉瑞、基因泰克等业界巨头以及众多科研院所达成基因组测序合作协议的核心筹码就是其在长期亏损检测中通过附随使用协议攒下的80万消费者的基因数据。

“大数据往往混合了来自多个数据源的多维度

^①4D风险具体指醉酒(Drunk)、吸毒(Drugged)、瞌睡(Drowsy)以及走神(Distracted)。

信息。”^[6]人工智能广域应用带来的巨额数据红利进一步刺激众多公司持续挖掘数据隐私,诸如脸书悄然扫描用户隐私信息和谷歌“深思”非匿名获取和识别海量急性肾损伤患者数据等恶性事件层出不穷。尤其是在数字追踪线上惯例影响之下,优步、易到和滴滴等企业的城市规划数据中心长期监控用户行踪数据,大量网贷平台公然追踪并评级用户日常行为,无人车、无人机、无人商店等均以 7/24 模式疯狂收集用户身份和行为习惯,植入体内的芯片情报介质更是不间断地开展监控和存储。缺乏高精尖技术知识的普通民众难以深刻意识数据隐私无序暴露和肆意使用可能造成的恶劣后果,往往为了节约时间和经济成本、获得更多便利或更高质量而忽视隐私侵害。

此外,“搜索一旦得以进化,就有可能成为最终连接人性和机器的‘铰链’”。^[7]去中心化巨网笼罩下人工智能技术应用于全网海量数据资源搜索领域导致记忆外部化。智搜系统对于散落网际空间的个人数据永不遗忘的记忆传承减弱了人们对于自身隐私的控制权。基于海量零散数据处理的智能识别软件持续迭代,进一步助力深层的关键性隐私挖掘,迫使整个族群被长期禁锢于透明的数字化圆形监狱。如形形色色的面部识别算法不仅用于情报收集研判(如从“天眼”系统中分辨恐怖分子)、判断和引导对象政治选择(如 2016 年美国大选和英国脱欧投票期间通过面部识别分类发送资讯而试图影响投票者),甚至以高达 91% 的准确率识别性取向^[8],催生外表决定论的粗暴法则,严重侵害当事人的知情权、自主权和人格尊严,在特定的时间或环境下(如乌干达、尼日利亚等全民歧同且酷法反同的国家应用同性恋识别软件)可能产生灾难性后果。

事实上,人格尊严是个人价值的源泉,是自由和独立的基石。企业通过智能系统持续监视员工(如彭博监控所有员工整个职业生涯的全部办公室活动)并收集一切移除人类直觉和情感的客观数据,精准评估其未来潜力、货币价值和人格特征,继而引发偏见、歧视和其他不可预期的侵害对方尊严的恶性事件,导致员工丧失劳动热情和主人翁精神,甚至引发严重的心理疾病。此外,包括类人性伴侣和老幼陪伴

机器人等在内的替代人类接触的人工智能应用导致弱势群体被物化忽略,侵害其基本的人格权益。

3 人工智能应用风险的应对策略

万物互联时代的科技乌托邦愿景大力推动现实与虚拟交融,极客思维映射人工智能全方位渗透社会生活各个领域,在“促进国家拥有经济、金融、意识形态、政治和其他资源……促进国家繁荣”^[9]的同时,充满变化性、危险性与不确定性。从实质侵害配偶权益的性爱机器人到有违伦理的葬礼诵经机器僧人^[10];从上线不足一日就变身不良少女的微软小冰到编写具有自我保护意识的计算机程序,“我们现在已经进入通向未知世界的兔子洞了”^[11]。面对人工智能多元应用的已知风险与未知威胁,亟待围绕《网络安全法》《新一代人工智能发展规划》的指导思想、战略目标、重点任务和保障措施,构筑具有中国特色的公权主导、多方参与、域内外合作规制的智能应用政策法律调整框架。运用复合技术标准检测、认证和评估人工智能安全指数,确立妥善处理国家安全、公共安全、企业利益和个人权益的协同联动和相关责任划分的政策法律体系,推动各级监管部门的能力建设、参与企业的自我审查与个人权益的全面保障,形成健康有序的智能生态培育环境,大幅提升整个社会的文明水准,加快发展中国的软实力和巧实力,积极建立创新型科技强国。

3.1 确立和谐永续的立法目标

数据资源与处理技术爆炸式增长中创新红利、市场红利和人口红利交叠起效,为我国人工智能多元应用在全球范围内弯道超车创造了有利条件。亟待确立人工智能研发和应用有益族群和谐永续的立法目标,遵循技术进步、经济增长和公众权益交互规律,平衡国家安全、社会稳定、企业发展与个人成长的矛盾冲突,秉持安全性、可用性和可溯源性原则制定指引人工智能规划和部署的法律规范,明确智能应用的法律主体及相关权责,实行设计问责和应用监督并重的行政监管机制与法定认证和预警标准,制定人工智能研发行为守则并布局人工智能人才发展战略。如海量数据和人工智能多维融通在一定程度上导致技术引领的信息垄断和数字鸿沟。“随着各平台不断

专业化自己的服务,未来社会的不平等问题将更严重”。^[12]亟待进一步完善公平准入和有序竞争规则,避免滥用垄断地位,持续增进公众福祉。

3.2 构筑张弛有度的政府监管机制

激进的技术变革往往促使体现集体理性的政府公器适用全新的监管模式。迅速迭代的人工智能既是社会发展的核心动能,亦是人类文明的根本风险。亟待迅速构筑张弛有度的政府监管机制,积极权衡技术进步、经济增长与公众权益的交互关系。以最优方式划分部门权限,坚持公众参与原则,确保监管连贯、合理、公平,兼顾人工智能固有的军民两用性,既保密和保护突破性军事创新,又扩展研发和应用民用颠覆性技术;既制定防止人工智能造假的技术标准,又补白数据侵权的保障性规范。坚持将大数据作为国家基本战略资源与将人工智能多元应用的关键部分作为公共基础设施管理的基本方针。运用合理而透明的可靠系数标准从网络层、数据层到算法层等全面量化人工智能系统的安全程度,逐步建立关键基础设施保护体系、推进弹性保障的非歧视用工制度、灵活现代的福利体系以及审慎高效的数据开放规则,防御并降低偶发事件带来的灾难性风险,逐步提升我国的全球位次。

3.3 颁布透明可信的数据核证规范

人工智能产业顺遂成长的关键在于获得数据权人、其他使用者乃至普通民众的数字信任。随着中国科技行业逐渐走出加拉帕戈斯群岛式封闭状态,大量内嵌丰富功能的智能应用的国际竞争力愈加强大。亟待颁布透明可信的自动化数据审核和反馈规范,积极完善冗余自检机制、构筑轻量密钥算法标准、健全高性能模块分离流程和安全防御指标。不仅立法督促科技企业全面公开智能技术部署和渐进测试过程,助力其依据用户体验和机器错误积极调整设置;还要求智能企业预先锁定并嵌入高度模块化的安全设计,遵循最小阻力路径识别和控制黑客注入易受攻击的疏漏集群,避免利用情感或意图错觉操控弱势用户^[13]。

3.4 完善有区别的责任制度

完全人工智能(Full AI)和应用型人工智能(Applied AI)图景下能够观察环境并做出行动以达到目

标的智能主体绝大多数行为不受人类直接指令约束,而是基于所获信息的分析判断。创造者难以预料和事先控制其不同情境中的应激反应与决策选择。单纯以产品具有缺陷追究生产者的产品缺陷责任或以未尽到合理注意义务追究使用者的过错责任等,既不符合追求公平和正义的法治精神,亦不利于科学划分参与各方的权利义务,还会因缺乏可操作性而影响法律公信力并打破智能应用有序发展的稳定格局。亟待出台系统化的调整人工智能多元应用的特别规范,赋予智能主体法律人格。完善智能应用的设计、制造与使用中有区别的归责体系,对于商用智能系统适用严格责任制度,给予公益相关的智能主体(如救援机器人、智能国安情报员和智能图书管理员等)特殊宽待。细化智能运作的标准流程和数据隐私保障制度、智能系统登记制度、危险责任制度、强制保险制度和赔偿基金制度等。

3.5 实施适应全新需求的层级税制

人工智能的广域应用在很多领域替代了传统人力劳动,间接导致国家大量税收流失。有必要依据新纪元发展特征,考虑实现社会公允与均衡进步的宗旨,积极改革现有税收制度,实施有条件的层级税制。如对于有利环境改善、疾病治疗和知识普及的人工智能技术应用免税,对于其他工业和服务业智能系统(如无人商店的自助结账系统)的算法设计、机械制造和使用企业征税;又如,增收资产税和奢侈品消费税以避免财富和资源过度集中于技术精英群体,避免智能企业的应税业务转移到其他国家和地区。进而使用上述部分税收收入搭建弱势群体进阶知识储备的绿色通路,实现社会公平。

3.6 加强人工智能领域的宣传教育

人工智能推动社会发展的核心价值使得相关宣传教育至关重要。全面提升整个社会获取、理解和整合数字信息的素养和参与研发应用的能力是维持人工智能可持续发展的重要举措。有必要立足国家人工智能发展战略,通过相关宣传教育,积极培养大批守法的“人工智能+”的复合型人才,尤其重视提升女性、少数民族同胞和欠发达地区居民的相关法律法规认知水平,防止智能纪元的性别歧视、种族歧视和收入偏见等等。此外,积极加强高等院校培养智能人

才中的党性教育、爱国主义教育和集体主义教育,重点强化私营科技企业的党组织和工会系统,亦有利于降低科技企业集结逐利与技术专家群体泛自由主义过度泛滥引发的诸多危险。

4 结语

进化的互联网与迭代的信息科技推动人工智能多元应用广域开启整个族群改造世界的第二次契机。智能应用已经从最初的军事、科技和工业应用渗

透到普通民众的日常生活之中。在充分发挥加快知识传递、驱动产业变革、完善社会治理以及推动族群永续等积极效用的同时,亦暴露出大量阻碍社会和谐发展的严重风险。亟待建立智能系统设计者、制造者、使用者以及其他利益相关者合作共赢的权责体系,充分发挥各级各地公权部门合法规制与激励发展的双重作用,积极深化智能应用领域的国际管控合作,确保运作流程公开透明与安全可靠,统一协调公权力、私权力、多样生态和社会秩序等。

参考文献:

- [1] Elish M C, Tim H. An AI Pattern Language: Accounting for Human Factors & Human Frames[R]. Data & Society Research Institute, 2017: 8.
- [2] Michelle K. American Diabetes Association and IBM Watson Health Join Forces to Reimagine How Diabetes are Prevented and Managed [R]. American Diabetes Association Report, 2016: 31.
- [3] Da Vinci Surgery: Minimally Invasive Surgery[EB/OL]. [2017-12-08]. <http://www.davincisurgery.com/>.
- [4] 王立铭. 上帝的手术刀: 基因编辑简史[M]. 杭州: 浙江人民出版社, 2017: 1.
- [5] 周志敏, 纪爱华. 人工智能: 改变未来的颠覆性技术[M]. 北京: 中国工信出版集团, 北京: 人民邮电出版社, 2017: 27.
- [6] 李开复, 王咏刚. 人工智能[M]. 北京: 文化发展出版社, 2017: 3.
- [7] (美) 斯特凡·韦茨. 任颂华, 译. 搜索: 开启智能时代的新引擎[M]. 北京: 中信出版社, 2017: 101.
- [8] Yilun W, Michal K. Deep neural networks can detect sexual orientation from faces [J]. Personality and Social Psychology. Preprint [EB/OL]. [2017-12-08]. <https://psyarxiv.com/hv28a/>. 同性恋识别算法使用深度神经网络提取特征, 在给定同一测试者五张面部图像的前提下, 分类器区别男性同性恋的准确率高达 91%、女性 83%, 远优于人类判断表现(男性 61%、女性 54%)。
- [9] (美) 乔舒亚·库帕·雷默. 罗康琳, 译. 第七感: 权力、财富与这个世界的生存法则[M]. 北京: 中信出版社, 2017.
- [10] Robot Priest's Funeral Role[N]. Hindustan Times, 2017-08-23(4).
- [11] (英) 卡鲁姆·蔡斯. 张尧亮, 译. 人工智能革命: 超级智能时代的人类命运[M]. 北京: 机械工业出版社, 2017: 173.
- [12] (印) 阿鲁·萨丹拉彻. 周恂, 译. 分享经济的爆发[M]. 北京: 文汇出版社, 2017: 8.
- [13] (英) 玛格丽特·博登. 孙诗惠, 译. AI: 人工智能的本质与未来[M]. 北京: 中国人民大学出版社, 2017: 1955.

作者简介: 蒋洁(1979-), 女, 南京信息工程大学中意网络侵权研究所、南京信息工程大学法政学院副教授, 中国人工智能产业发展联盟政策法规组副组长, 意大利信息法学博士后, 腾讯学术顾问。