

DOI: 10.12120/bjutsxb20180665

# 人工智能发展的潜在风险及法律防控监管

马治国,徐济宽

(西安交通大学 法学院,西安 710049)

摘要: 人工智能技术的特殊性,使其对社会乃至人类文明可能造成不可逆转的严重危害;为了预防潜在风险的发生,中国亟需对人工智能的发展进行全面理性的评判,权衡其技术发展中附带和潜在的负面影响及背后蕴藏的巨大风险,并对人工智能的风险防控和监管的法律路径进行思考,构建起以事前预防、事中管控和事后监督为核心的监管体系;在抓住人工智能发展机遇的同时,能够有效防止不可逆的危害结果发生,促进人工智能的良性发展,进而以此造福人类社会。

关键词: 人工智能; 法律防控; 风险预防

中图分类号: D 920.4

文献标志码: A

文章编号: 1671-0398(2018)06-0065-07

近年来,人工智能技术迅速崛起并日趋成熟,我国积极响应人工智能的全球浪潮。2016年7月,国务院印发的《“十三五”国家社会发展科技创新规划》<sup>[1]</sup>(以下简称“‘十三五’规划”)中,人工智能被作为一项新一代信息技术列入规划。2017年3月,人工智能首度被列入政府工作报告;同年7月,国务院又下发了《新一代人工智能发展规划》<sup>[2]</sup>提出:“到2020年人工智能总体技术和应用与世界先进水平同步,到2025年人工智能基础理论实现重大突破,到2030年人工智能理论、技术与应用总体达到世界领先水平。”<sup>[2]</sup>可以预计,中国的人工智能产业即将迎来蓬勃发展的时代。然而,我们在期待这一领域蓬勃发展的同时,更应该保持理性且审慎的态度看待人工智能,务必准确认知其潜藏的危机和风险。著名科技先驱埃隆·马斯克(Elon Musk)认为在人工智能领域,政府监管的介入可能是一个非常明智选择。他认为,应该十分小心地看待人工智能,并猜测人类下一个生存威胁恐怕就是人工智能;而且对于人工智能应该在国际或者国家层面上有相应的监管措施,以防人类做出不可挽回的傻事来<sup>[3]</sup><sup>46</sup>。在现实中,人类社会刚刚兴起并迅猛发展的人工智能领域尚处在监管真空的境地,亟需一套科学的配套措施在人工智能的风险防控和过程监管方面进行规制。以人工智能为代表的新科技既为人类社会带来无尽便利和快捷,也延伸出诸多复杂的伦理、

法律和安全问题。2017年,围绕人工智能对社会发展影响和挑战的研究,成为我国哲学社会科学界探讨的一大热点<sup>[4]</sup>。

## 一、人工智能发展面临的机遇与挑战

目前,人工智能发展经过60多年的演进,突破瓶颈进入了新阶段。特别是在移动互联网、大数据、超级计算、传感网、脑科学等新理论新技术以及经济社会发展强烈需求的共同驱动下,人工智能飞速发展,呈现出深度学习、跨界融合、人机协同、群智开放、自主操控等新特征<sup>[2]</sup>。同时,人工智能的发展并不局限于某一特定产业,而是能够支撑所有产业变革的通用型技术。因此,人工智能具有广泛的社会溢出效应,在政治、经济、社会等各个领域都会带来深刻变革<sup>[5]</sup>。当下,人工智能已成为全球新一轮革命和产业变革的着力点,旧领域生发出新的产业形态,多领域催生了新兴的细分行业,由此创造出巨大的经济价值和社会财富。在谷歌、脸谱、IBM等世界知名企业的带领下,全球对人工智能的关注度不断提升。自2000年之后,全球人工智能企业的融资总额288亿美元。近五年来,人工智能企业融资总额达224亿美元<sup>[6]</sup>。到2017年,从细分领域来看,人工智能在融资上全面开花,包括芯片、技术平台等基础性技术,自然语言处理、语音识别、机器学习应用、计算机视觉与图像等人工智能应用,智能无人

收稿日期: 2018-04-11

基金项目: 西安交通大学人工智能与区块链技术法律新问题研究项目资助(ZGXJ201707)

作者简介: 马治国(1959—),陕西绥德人,西安交通大学法学院知识产权研究中心主任,教授,博士生导师

机、智能机器人、自动驾驶等重点行业应用都有频繁的融资和并购,研发和应用都取得了一定进展<sup>[7]</sup>。人工智能将从专业性较强的领域逐步拓展到社会生活的各个方面。

人工智能作为影响面广的颠覆性技术,可能带来并改变就业结构、冲击法律与社会伦理、侵犯个人隐私、挑战国际关系准则等问题,将对政府管理、经济安全和社会稳定乃至全球治理产生深远影响<sup>[2]</sup>。我们应当清醒地认识到,在抓住人工智能发展机遇的同时,还必须高度重视这一领域可能给社会带来的安全稳定等方面的巨大挑战。因而,我国亟需对人工智能发展加强前瞻预防与约束引导,从而最大限度降低风险,并确保其安全、可靠、可控的发展。这些问题如今在国际层面也逐渐受到重视。联合国教科文组织(UNESCO)与世界科学知识与技术伦理委员会(COMEST)在2016年联合发布的报告中,讨论了机器人的制造和使用促进了人工智能的发展,以及这些发展所带来的社会与伦理道德问题。人工智能机器人具有的类似人类的活动系统和学习能力使其成为了自动化、智能化的电子实体。这些实体不仅能够胜任复杂的决策过程,还能通过复杂的算法进行类似人类的具体活动。这些都潜藏着巨大的伦理和法律危机。

## 二、人工智能发展的潜在风险及防控

### (一) 人工智能自身特性与价值中蕴含的社会风险

自20世纪50年代人工智能诞生至今,这一领域的存续、发展与人类对人工智能“代替劳动”这一价值的期望密不可分。也就是说,人工智能的产生和发展最终仍归结于对人类劳动的解放性替代。根据马克思“劳动创造人”的原理,人之所以为人是因为劳动,是劳动创造了人;如果人远离劳动,也必然造成人的退化。惰性是人之常情,但是不能一味地任由惰性左右人的发展。人类经过数万年的演进,劳动机能是人类安身立命的本能之一。如果人类一味地靠技术和机器进行替代而不去提升自身固有的劳动创造能力,则很有可能对人类发展的正常演进造成重大误导。有些人工智能的支持者会认为,人工智能对人类劳动的替代只是部分取代,人们由此可以腾出更多时间做更有意义的事。但是,人们对人工智能的选择本身就是惰性驱使的结果,其他所谓更有意义的事情大多是排斥易疲劳的、相对较为轻松平和的劳动。这样,人们对技术和设备的依赖

程度必然会逐渐提高,从而在部分领域独立完成工作的能力势必减弱,易于陷入逐渐退化的怪圈。所以,人类对人工智能的选择本身就需要冷静理性的重新审视。许多学者预测,人工智能发展到一定程度,人类未来会在“万物皆互联、无处不计算”的环境下“精准生活”<sup>[8][129]</sup>,但是精确的生活背后必然是无穷尽的依附性。当人工智能渗透到人类的每个角落,人们对其依附程度逐步提高,依靠技术和产品过着所谓的“精确”生活。一方面,人类这样持续下去,自身的劳动机能必然会退化;另一方面,一旦人工智能自身出现问题或缺陷,必将对人类造成毁灭性的打击甚至有倾覆人类的可能。因此,我们在对人工智能极端推崇的同时,应以长远眼光理性权衡人工智能对人类影响的利弊;并且,对人工智能自身价值的认真考量和其自身特性的审慎对待,是人工智能风险防控和过程监管的前提性工作。

### (二) 人工智能风险防控的必要性

人工智能从产生至今已有60余年,学界仍然未见统一性或准确定义。这在很大程度上是因为人工智能在技术层面上有其自身的复杂性。具体而言,行为自主性是人工智能与人类先前科技成果最大的不同之处。例如,人工智能技术在现阶段已经可以做到在不需要人类控制或者监督的情况下,自动驾驶汽车或者起草一份投资组合协议<sup>[3][48]</sup>,甚至更为复杂的创造性行为。这些创造性行为又催生了人工智能的不可预测性与不可控性。人类决策系统与现代人工智能系统存在一些本质上的区别,正是这些区别使人类无法准确预知人工智能系统对于一个问题的解决措施。通常人类因受限于大脑认知能力,为了做出决策,无法在有时间限制的情况下分析所有的相关信息。所以,人类常常退而求其次,去选择一个自己满意的解决方法,而非一个最佳化的解决方法<sup>[3][49]</sup>。但人工智能因其技术特性,从而有能力在短时间内做出其自身认知的所谓最优化判断,但这些判断因为不受类似人类提前预置的概念、经验法则、传统智慧等因素的影响和限制,其最终决策很可能同人类现有认知程度的决策产生较大差异。所以,尽管人类设计了人工智能并制造了相关产品,其实很大程度上是无法预见人工智能可能做出的决策。2016年12月,世界经济论坛网站上的一篇报道称,根据对近900名专家的调查表明,当被问及哪些新兴技术需要进行更好的治理时,人工智能与机器人技术高居榜首,生物技术紧随其后<sup>[9][111]</sup>。

人类社会亟需理性认识人工智能正在发生的态

势和问题,更需要预估、评判和预防未来人工智能的发展中可能引发的潜在风险与危机。从长远来看,人们最担心的必然是未来人工智能对人类存续的颠覆性挑战。科学家霍金的“人类最后的文明”说和人工智能预测家库兹韦尔的全面超越人类智慧的“奇点”论<sup>①</sup>,都表现出对人工智能未来的“超人类性”“反社会性”以及“非可控性”的担忧<sup>[10]</sup>。所以,我们要避免这些灾难的出现,就必须要有有一套切实可行的风险评估和监管程序,同时还要构建人工智能复杂场景下突发事件的解决方案。

### (三) 人工智能潜在风险分析

#### 1. 伦理危机:拟人之祸

目前,人类对人工智能的相关探讨多与机器人技术相关联。机器人的智慧达到了一定程度就不可避免地对机器人进行“拟人化”的探讨,诸如对机器人的人权、劳动权等一些与自然人人身相关的权利进行了探讨论证。笔者认为,这个命题过于激进且缺乏论证基础,而且这些论证的本身就是对社会纲常伦理的巨大挑战。自然人与法人都是民事主体。自然人即生物学意义上的人,是基于出生而取得民事主体资格的人。如果仅凭人工智能的智能程度和产品的外形相似度就将人工智能与自然人相提并论,那么则是完全没有理论根据并且有违伦理的。2016年8月,联合国教科文组织与世界科学知识与技术伦理委员会联合发布了一份关于机器人伦理的报告。报告讨论了机器人在不同领域的广泛应用以及可能引起的伦理问题,希望能够提高公众意识,使其更多地关注和参与机器人伦理问题的研究与讨论。2016年9月,英国下议院科学技术委员会发布《机器人学与人工智能》报告,讨论了人工智能与机器人可能产生的伦理与法律问题,强调了适当的治理框架的必要性。2017年1月,英国政府针对《机器人学与人工智能》报告中的建议给予了回应,承认关于人工智能与机器人的社会、伦理与法律问题的重要性。2017年1月,霍夫曼(Reid Hoffman)、欧米迪亚网络以及奈特基金会等共出资2700万美元,成立人工智能伦理与治理基金,以支持美国及国际上关于人工智能伦理与治理项目及活动的跨学科研究<sup>[9][11]</sup>。

人工智能产品在运行过程中因其与自然人的本质差别,难免偶尔会做出违背人道主义的事情,妄

言、乱举等言行失范的恶性事件难免发生。与此同时,人们在使用智能产品的过程中,因先天的本质差异导致的优势心态会放任使用者在对机器人的使用过程中对其谩骂、虐待等不当行为,甚至当机器人变成发泄的工具、变态的玩具的时候<sup>[11][108]</sup>,人类逐渐异化的价值取向必然与当前社会伦理严重抵触。当人工智能在外形和智能程度都无限接近人类时,人类很难坚守住固有的伦理与道德体系和底线。

#### 2. 技术风险:滥用之患

随着社会的进步,科学技术的发展程度越来越超乎人类所能预知的范围。技术的重要性较以往的人类文明程度飞速提升,甚至在不久的将来或将成为控制人类社会的核心领域。当技术成为社会主导力量的时候,技术滥用的危机将成为人类面临的共同难题。比如,人工智能在各个领域的应用,一方面确实会带来前所未有的便利,减轻人类负担、提升幸福感;另一方面对社会各界的冲击同样是不可避免的。例如,智能机器人普及导致的技术性失业和经济动荡,智能创作平台或取代人脑创作导致的文化塌陷,并且人工智能技术发展一定程度且密切影响到人类工作生活后足以撼动现有的政治体制和社会制度。此种技术主导型发展一旦被滥用就极易失控,或将导致社会的变革甚至是对文明的颠覆和社会格局的重整。

人工智能技术本就其特性而言,有着不可避免的风险。这些风险中最为直接的现实挑战就是技术风险。一方面,智能产品的普及必然带来人与智能产品同在共存的生活常态,而技术自身的隐患同样会危及人类日常生活安全;另一方面,人工智能的技术风险也可能滋生侵权行为、信息泄露等一系列伴生问题。如果诸如此类的技术问题从源头上只能交给产品开发者来研究控制,不将所有不确定因素考虑进去,则人工智能技术便不能真正的推广普及。

#### 3. 战争灾难:军工之孽

从某种程度讲,军事与科学技术是相辅相成的。因战争中特殊而又高标准的技术要求,战争促进和带动技术的作用是不可小觑的,而技术的进步又会第一时间运用于军事。成熟之日的人工智能技术必将首先应用到军事领域。未来,一大批“战争机器”将会到达沙场,人与人正面冲突的几率将越来越少。这虽然一定程度上减少了士兵的伤亡,但并不能降

<sup>①</sup> 库兹韦尔认为,人工智能可以全面超越人类智能,而且超越的速度远超过人们的预期。他认为2045年是“奇点”到达的时期,这将是极具深刻性和分裂性的时间点,非生物智能在这一年将会10亿倍于今天所有人类的智慧。

低战争的残忍程度。高精密机器介入战场不仅会使国家或地区之间的摩擦变得更为频繁、交战门槛变得更低,战争或将更加频繁;而且人工智能自身的特性附带了战争属性之后,杀伤力和毁灭程度是难以想象的。因此其可能一个简单的指令就能在短时间内对平民造成大规模的毁灭性打击,极其残忍且危害后果惨痛。

阿姆斯特朗等人关于未来的人工智能军备竞赛专门预设了模型并进行了研究分析,得出较为直观的结果:如果控制能力本身的重要性远高于所采取的预警级别,那么整体结果会更安全<sup>[12]</sup>。笔者看来,如果人工智能的发展趋势依然不可逆转,政府或国际层面有责任出面把控人工智能的应用领域,避免肆意毁灭人类的人工智能武器被制造出来,避免更频繁的现代化战争,共同维护世界和平。

#### 4. 安全隐患: 异化之险

据牛津大学哲学家尼克·波斯特罗姆(Nick Bostrom)的调查,最常受访的前100位人工智能技术相关研究人员中超过半数的受访者认为,人类层级的机器智能对人类的影响将是“不太好的影响”或是“极其坏的影响(事关存亡的人类大灾难)”。“75%的专家认为走向超级智能是必然的趋势,且多数杰出的人工智能专家已经认识到人工智能具有威胁人类存续的可能性。”<sup>[13]</sup>国内外的诸多学者与专家认为,人工智能在未来或将超越人类智慧,甚至认为人类正在创造一个比自身更加强大的物种。不难想象,人工智能的智慧程度和情感能力一旦突破了质变,人类很难完全控制和预测其思维方式,人工智能的异化与畸变便具备了现实化的可能性。这种技术失控的局面在人类历史长河中早有先例,就如原子弹与克隆技术一样,科学家们创造出来的成果往往并不能由本群体来决定是否投入使用,因其一旦失控或异化将对人类造成毁灭性的颠覆并且往往是不可逆的。

当务之急,我们应当将这些风险扼杀在摇篮之中,运用制度思维与实践规范人工智能的开发与使用过程,尽可能避免潜在危机的现实化。“新一代人工智能发展规划”第五部分“保障措施”中提到,围绕推动我国人工智能健康快速发展的现实要求,是“妥善应对人工智能可能带来的挑战,形成适应人工智能发展的制度安排,构建开放包容的国际化环境,夯实人工智能发展的社会基础”<sup>[1]</sup>。面对智能革命,我们不能逃避、否定和阻止它。人工智能在造福人类的同时也加大了人类危害自身的可能性,

这是技术的负面性与风险的不确定性的联系。同时,传统社会治理体系无力解决工业社会过度发展而产生的社会问题,这即是法律的确信性与风险的不确定性的背离<sup>[8] 130</sup>。当今世界,人工智能的竞相发展是主旋律,但人工智能发展背后的风险防控同样是全世界共同的深刻命题。各国亟需做出针对风险的制度选择和法律安排,以法律防范风险治理危机,并且以制度监管过程控制风险。

### 三、人工智能风险防控与监管的法律思考

#### (一) 人工智能风险的防控监管现状及困境

我们对于人工智能技术的监管不能局限于事后,更重要的是预先防范、监管提前。目前,由于缺乏对人工智能潜在风险认识,国际上鲜有针对人工智能的专门的监管法律与机构,大多仍处于空白状态;同样,对于人工智能监管途径的法律研究也是匮乏的<sup>[3] 46</sup>。此外,我们对政府各机构各部门监管人工智能的技术能力考察相关研究也颇为少见,且很难有定论指向哪个机构能够更好地监管人工智能的开发、研究和使用的过程。

就目前社会监管基础而言,民法方面的产品许可制度和侵权责任等传统的监管措施都更注重事后监管,具体到人工智能的自身特性,并不能适用于管理有高度自主能力的智能产品所带来的风险。例如,人工智能产品完全根据其自主意识产生的自发性行为引发的侵权事件存在定性困境,这种情况下无法将责任单纯地归结为使用者、开发者或是销售者予以承担,而需要专门的定性和配套的归责原则。刑法方面,以智能机器人为代表的人工智能可以在自主意识和意志的支配下实施严重危害社会的犯罪行为。但是,无论是从刑法的目的分析,还是从刑罚种类分析,通过我国现行刑法对智能机器人直接加以惩罚是难以想象的<sup>[14]</sup>。然而,事前的监管是极其必要的但却困难重重,具体表现为人工智能的开发与研究可能具有隐蔽性、非连续性、分散性和易伪装性等。人工智能的自动化特征造成了风险预见和管理上的困难,尤其是在人工智能系统带来灾难性风险的情况下,事后监管措施极有可能失效。

#### (二) 法律在监管中的角色

尽管人工智能技术在发展过程中面临着各方面的潜在风险,但是发展趋势早已势不可挡。我们应当相信,法律制度能够在不阻碍技术创新与发展的前提下防控技术带来的负面问题,以及监管发展过

程的潜在危机。法律是规范人工智能技术研发与使用的有效选择,各领域潜在的风险和危机都可以依托法律的强制力和约束力来预防和化解。2017年1月,欧洲议会法律事务委员会已经认识到了监管人工智能技术的重要性,呼吁各成员国制定有关人工智能技术及其衍生产品的相关法案。我国依照“新一代人工智能规划”的精神,当前正在筹备和建立保障人工智能健康发展的法律法规和伦理道德框架;并开展与人工智能应用相关的民事与刑事责任确认、隐私和产权保护、信息安全利用等法律问题研究,同时正建立追溯和问责制度等<sup>[1]</sup>。除此之外,我国相关安全管理法规的研究制定也势在必行,从而为新技术的快速应用奠定法律基础。因此,我国相对应的司法系统也应做出适当调整。例如,法院也应解决可预见性与因果关系的问题,需要根据科技的发展和进步,不断调整其最接近因果关系的判断标准<sup>[3] 53</sup>。此外,法律除却将针对人工智能的研究与开发中的不透明和非连续性问题采取措施之外,法律也不能忽视对技术研发本身的限定和保护。从事科学研究是公民的自由与权利,是世界普遍认可的法律规则。但法律同样对科研自由有严格的界限,仅限定在那些有利于人类进步事业的范围。

### (三) 建立法律监管和评价体系

我国对待技术领域潜在风险的防控与监管,必须采用严格慎重的原则,并且法律监管措施应贯穿全过程。

#### 1. 事前预防

关于风险评估和预防,我国《新一代人工智能发展规划》已经体现出了对人工智能积极又审慎的态度,提出“人工智能未来规划中,加强人工智能对国家安全和保密领域影响的研究与评估是重要步骤,完善技术、相关人员、相关产品等配套的安全防护体系,构建人工智能安全监测预警机制必不可少。”<sup>[2]</sup>因此,我们应逐渐发展起对人工智能技术发展情况的预测、研判和跟踪研究,从而更准确地把握技术和产业发展趋势,从而将风险防患于未然。同时,我们还要增强风险意识,重视风险评估和防控,强化前瞻预防和约束引导;短期内的关注重点是对于就业的影响,长期重点则关注对社会伦理体系的影响,确保把人工智能发展规制在安全可控范围内。

#### 2. 事中管控

具体到研究内容,首要任务是建立评价体系,具体量化人工智能的智能等级,从而限制人工智能的应用范围以及研究领域,将一定强度等级的研究领

域严格控制在政府可监视范围之内。其次要构建动态的人工智能研发应用评估评价机制,围绕人工智能设计、产品和系统的复杂性、风险性、不确定性、可解释性、潜在经济影响等问题。同时,我国还应着力开发系统性的测试方法和指标体系,建设跨领域的人工智能测试平台,推动人工智能安全认证,评估人工智能产品和系统的关键性能;除此之外,应实施具体实行设计问责和应用监督并重的双层监管结构,实现对人工智能算法设计、产品开发和成果应用等的全流程监管。

#### 3. 事后监督

具体到产品应用,我国可以限制公司对不同级别智能产品的最大使用数量或占全体工作人员的比例,以降低技术性失业的风险和大规模下岗所带来的负面影响。同时,应设立监管部门进行专项监督,从保障从人类自身利益出发,规范市场行为,维护行业秩序,防止人工智能的肆意开发与过度使用<sup>[11] 108</sup>。此外,事后监管还应重视企业自身建设,大力促进人工智能行业和企业自律,同时加大对数据滥用、侵犯个人隐私等违背道德伦理等行为的惩戒力度。促进企业的良性发展和技术创新,鼓励企业加强人工智能网络安全技术研发,强化人工智能产品和系统网络安全防护。

#### (四) 构建与时俱进的科研伦理规范体系

广义的法律规范应当包括伦理规范的引导、补充和衔接,尤其对于科学研究而言。我国除了以法律层面为代表的强制监管,从长远来看,还需要在社会舆论层面进行正确的引导。对此,我国《新一代人工智能发展规划》提到,“未来工作的重点是要开展人工智能行为科学和伦理等问题研究,以此建立伦理道德多层次判断结构及人机协作的伦理框架。”<sup>[2]</sup>同时,还需制定人工智能产品研发设计人员以及日后的使用人员的道德规范和行为守则,从源头到下游进行约束和引导。由此看来,关于人工智能的“德治”与“法治”有着同样举足轻重的地位。目前,规范的伦理道德价值引导是极其可行并且必要的。短期来看,法律对于人工智能的研发、生产、使用等一系列过程的监管都尚属空白,亟需一套有效的约束机制来进行正确的价值引导,防止人工智能研发和使用的异化造成不必要的社会动荡甚至是灾难。在配套的法律法规还未成型生效之前,我们应倡导正确的价值引导,可以有效地避免违背伦理道德却未违反现行法律的恶性事件发生。长期来看,我国建立起关于人工智能的规范化伦理体系是人类文明进步的必经之路,是法律规

范的必要补充;人工智能发展到一定程度对传统纲常伦理和道德体系的冲击是不可避免的,是人类必须共同探讨的重大命题。

#### (五) 国际协同监管与合作

人工智能以其神速发展和已达到的智慧水平,给人意外之震撼。按照这条轨迹和速率,在未来的日子里,人工智能震惊人类的频率和烈度只会趋于密集和增高<sup>[15]</sup>。法律对人工智能现实问题的解决,关乎人与机器未来发展的走向,颠覆冷兵器时代下人们对机器的浅显认识,打破人机交互的传统相处方式<sup>[16]</sup>。目前,人工智能虽是各国提升竞争力的首选,但同时也是全球需要共同面对的巨大挑战。

人工智能的异化、极端化是危及全人类的重大危机和不可逆风险,这不仅仅是国家或地区层面的问题。目前,我国将积极参与人工智能异化及安全监管等人工智能重大国际共性问题研究,并积极参与全球治理;同时,在人工智能法律法规、国际规则等方面可以多方寻求国际合作,共同应对危机和问题。在人工智能今后的发展过程中,我国应当以负责任大国的姿态做出表率,审慎地做好人工智能风险把控和防范的

工作,并协同其他国家和地区共同努力维护人工智能技术的有序发展。同时,我们呼吁各国积极倡导“人类命运共同体”价值观,追求本国利益时兼顾他国合理关切,在谋求本国发展中促进各国共同发展,从而共同维护人工智能技术的有序发展,推动人类共同进步,避免不可逆的危机发生。

## 四、结论

人工智能的发展速度已超出了人们的预料,几乎每天出现在新闻头条,人工智能新技术、新进展向人类敲打着醒世警钟并不断展示着惊人成果。尽管科学技术本身是中立的,不易做道德评判。但是,我们应当对技术本身所伴随的附带性影响有清晰的判断、理性的认识以及审慎的对待。面对人工智能可能存在的事关社会倾覆甚至人类存亡的潜在不可逆风险,人类应当以更加谨慎的方法和态度对待,进行审慎的事前预防和严格的过程监管是极其必要的,应当充分运用法律对科技活动和后果的预判、评价、调整功能,积极应对高新技术发展中的未知因素和潜在风险,切勿当下漠视风险,未来自食其果。

#### 参考文献:

- [1] 中华人民共和国中央人民政府. 国务院关于印发“十三五”国家科技创新规划的通知[EB/OL]. (2016-07-28) [2018-07-29]. [http://www.gov.cn/zhengce/content/2016-07/28/content\\_5098072.htm](http://www.gov.cn/zhengce/content/2016-07/28/content_5098072.htm).
- [2] 中华人民共和国中央人民政府. 国务院关于印发新一代人工智能发展规划的通知[EB/OL]. (2017-07-20) [2018-07-29]. [http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm).
- [3] 马修·谢勒,曹建峰,李金磊. 监管人工智能系统: 风险、挑战、能力和策略[J]. 信息安全与通信保密, 2017(3): 45-71.
- [4] 光明日报理论部,学术月刊编辑部,中国人民大学书报资料中心. 2017年度中国十大学术热点[N]. 光明日报, 2018-01-27(11).
- [5] 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): 40-45.
- [6] 中国电子商务研究中心. 2017全球人工智能发展报告投融资篇[EB/OL]. (2017-08-11) [2018-01-29]. <http://b2b.toocle.com/detail-6410042.html>
- [7] 中国报告网. 2017年我国人工智能行业发展趋势及投资规模分析[EB/OL]. (2017-10-25) [2018-01-29]. <http://free.chinabaogao.com/it/201710/10252b3D2017.html>
- [8] 吴汉东. 人工智能时代的制度安排与法律规制[J]. 法律科学, 2017(5): 128-136.
- [9] 杜严勇. 论人工智能的自反性伦理治理[J]. 新疆师范大学学报(哲学社会科学版), 2018(39): 111-119.
- [10] 陈伟光. 关于人工智能治理问题的若干思考[J]. 人民论坛·学术前沿, 2017(20): 48-55.
- [11] 李博方. 人工智能产业发展的风险分析及对策研究[J]. 智能计算机与应用, 2017(3): 106-113.
- [12] ARMSTRONG S, BOSTROM N, SHULMAN C. Racing to the precipice: a model of artificial intelligence development[J]. *Ai & Society*, 2015, 31(2): 1-6.
- [13] 艾伦·达福,斯图尔特·罗素. 人工智能的真正风险[J]. 中国经济报告, 2017(2): 118-119.
- [14] 刘宪权. 人工智能时代的刑事风险与刑法应对[J]. 法商研究, 2018(1): 3-11.
- [15] 顾骏. 天问: 二元智能时代的一元未来[J]. 探索与争鸣, 2017(10): 56-65.
- [16] 马治国,田小楚. 论人工智能体刑法适用之可能性[J]. 华中科技大学学报(社会科学版), 2018(2): 108-115.

## Legal Precaution and Supervision about the Potential Risks in Artificial Intelligence Development

MA Zhiguo , XU Jikuan

( School of Law , Xi'an Jiaotong University , Xi'an 710049 , China)

**Abstract:** Now AI technology has become the new focus of international high-tech competition , major countries in the world regard the development of artificial intelligence as a major strategy to enhance national competitive ability and safeguard national security. With the breakthrough of AI technology , AI technology has mushroomed in the world in recent years ,and it have attracted an influx of capital and human resources of different industries. Besides that , the unique nature of the AI technology itself may cause in irreversible harm to the society and humankind. There is an urgent need to make a profound rational understanding and evaluation of the rapid development of artificial technologies and to weigh incidental or potential impacts. It is essential to analyze the huge risk behind its development on the basis of exploring the current situation of AI technology. Exploring the legal path of risk prevention and control of AI technology is also necessary , which means to construct an organic regulatory system that takes precaution , incident control and ex-post supervision as the core. As people seize the opportunity of AI technology development , it will help prevent irreversible harmful consequences effectively and then promote the healthy development of AI technology.

**Key words:** artificial intelligence; legal regulation; risk prevention

( 责任编辑 冯 蓉)

( 上接第 55 页)

## Value Goal and Historic Task of Socialism with Chinese Characteristics in New Era

LU Weiming , SUN Xihong , LÜ Fei

( School of Marxism , Xi'an Jiaotong University , Xi'an 710049 , China)

**Abstract:** In this article , we have made an explanation on that value and basic task of socialism , and holds that socialism not only takes the promotion of well-rounded development of the person as the highest goal , but also takes the pursuit of efficiency and equality as the specific value goal. The value goal and the nature of socialism require logically two tasks of historic importance of socialism. One is to release and develop further the productive forces which can meet the growing needs of our people. Another is to promote social equity and justice of the civil and international community. As socialism with Chinese characteristics has entered a new era , Xi Jinping gave new era connotation and asked new requirements to the goal value and historical task of scientific socialism , further developing the theory of scientific socialism.

**Key words:** Xi Jinping; socialism with Chinese characteristics; theory of scientific socialism; value goal; historic task

( 责任编辑 冯 蓉)