

Doi:10.11835/j.issn.1008-5831.zs.2022.06.001

欢迎按以下格式引用:潘军,姚科敏. AI 智能层级与仿人实现的价值调控与治理研究[J]. 重庆大学学报(社会科学版), 2022(4):251-261. Doi:10.11835/j.issn.1008-5831.zs.2022.06.001.



Citation Format: PAN Jun, YAO Kemin. Artificial intelligence level and value regulation and governance of human-imitation[J]. Journal of Chongqing University (Social Science Edition), 2022(4):251-261. Doi:10.11835/j.issn.1008-5831.zs.2022.06.001.

AI 智能层级与仿人实现的价值调控与治理研究

潘军^{1,2},姚科敏³

(1. 贵州商学院 马克思主义学院, 贵州 贵阳 550081; 2. 清华大学 人工智能大数据中心, 天津 300467; 3. 贵州科学院 科技与经济战略研究院, 贵州 贵阳 550081)

摘要:人工智能(AI)是人们在致力于如何仿人和仿人哪些功能的过程中不断深度创新、探索的动态性和先进性器物。其智能层级水平基本上以仿人的维度、深度和效度来核定或计量;它的仿人价值调控及其社会治理是以技术价值、社会价值、伦理价值等研究成果的“人文+科技”方式逐渐实现的。正如麦克尤恩指出,人类有历史以来,无论处于什么样的文化背景,人们始终有一个共同的梦,那就是非生殖性地创造出“人造版的我们”,使得我们也像上帝造人一样做一回上帝,即造出多功能的人造人。从以人为本的发展观来看, AI 的演进发展史是人类以人这个“种类”为本,不断创新模仿人自身的体力、智力、“体力+智力”,以在更高水平和更深层次上智造“人”,进而实现进一步解放人自身的实践探索历史。AI 在整个演进过程中,能够同时在一定维度和深度上反映 AI 智能层级形成与演进的过程及运行范式便是“仿人”;同时, AI 仿人原则和仿人价值实现调控的技术价值、社会价值、伦理价值等工程知识在此过程中得到了逐渐的积累和完善。一方面, AI 由机械力学技术向着信息化智能化技术发展,使得单一解放体力劳动向着全面服务人类自由解放发展,一定意义上让 AI 获得了理论积淀、技术积累和初具规模的工程范式;另一方面, AI 较为普遍的深度融合应用引起了人们对自身未来社会活动的隐忧,掀起了以其仿人的智力强弱层级来反映智能层级的价值调控与治理研究热潮,如兴起了人工智能技术价值、社会价值、伦理价值等研究。正因如此,梳理 AI 演进发展历程及其价值调控与治理规律,特别是厘清各历程中仿人原则的价值标准、价值调控措施及价值调控质效等治理要素,并从每一种特定智能层级出发,用仿人原则的价值调控与治理知识保障其正面效应得以可持续发展,负面效应得到有效消解和遮蔽,有益助推 AI 在向上向善的价值调控与治理下实现全面深度的智能融合应用,实现 AI 更广泛和可持续的深度智能融合应用,实现 AI 在未来社会中只是人的社会活动的得力助手且是向上向善的线上线下 24 小时

基金项目:国家社会科学基金项目“AI 时代中国公民道德选择能力定性定量研究”(21BKS165)

作者简介:潘军,博士后,贵州商学院马克思主义学院,清华大学人工智能大数据中心,Email:1161141538@qq.com;姚科敏,博士,贵州科学院科技与经济战略研究院。

助力人类社会活动的道德选择之劳动者。

关键词:人工智能;智能层级;仿人原则;价值调控;社会价值;伦理价值;社会治理

中图分类号:C91;C93;G641

文献标志码:A

文章编号:1008-5831(2022)04-0251-11

一、问题提出

人工智能(AI)是人们在致力于如何仿人和仿人哪些功能的过程中不断深度创新、探索的动态性和先进性器物。它的智能层级水平基本上以仿人的维度、深度和效度来核定或计量;它的仿人价值调控及其治理是以技术价值、社会价值、伦理价值等研究成果的“人文+科技”方式逐渐实现的。正如麦克尤恩指出,人类有历史以来,无论处于什么样的文化背景,人们始终有一个共同的梦,那就是非生殖性地创造出“人造版的我们”,使得我们也像上帝造人一样做一回上帝,即造出多功能的人造人;或者正如“图灵测试是这样的,将一个人安置在一台计算机终端前,让他通过书面问答与几个未知的对象交互。如果在一段合理的时间内,提问者无法判断自己正在与计算机还是人类交流,那么,这台机器就可以被认为是‘智能的’”^[1]。由此可见,AI以人自身的灵活性和辩证性智力、智慧等作为机器智能性的参考和评判标准,即AI智能层级或仿人智慧的实现程度代表着AI的智能水平和自动化能力。

党的十八大以来,党中央高度重视我国人工智能在各个领域的应用与发展。习近平总书记曾强调,“要加强人工智能同保障和改善民生的结合,从保障和改善民生、为人民创造美好生活的需要出发,推动人工智能在人们日常工作、学习、生活中的深度运用,创造更加智能的工作方式和生活方式”^[2]。实践中,我国AI智能水平取得了诸多新进展,尤其是在一些特殊领域的智能研发中,AI能够模仿人的手、脚、眼、耳、鼻、舌、身等部分器官功能,以代替人的某些社会活动。例如:贵州已建成了大数据高地——数谷,建成了中国电信、中国联通、中国移动、苹果、华为等重量级企业的大数据中心,为AI的“智能”发展储存了数据资源;阿里云已建成城市大脑,利用实时智能视频分析技术实现了实时的交通管理、路况分析、道路维修监管,提高了城市的智慧运行效率;清华大学的脑接口智能技术已能搜集与分析课堂上学生的脑神经信息,呈现学生在课堂上的专注程度;华为公司已形成一系列智能产品,特别是与相关部门共同开发的StorySign可以支持十种手语,能让听力障碍儿童获得阅读能力,并能与普通人和谐互动;在抗击疫情中,AI呈现了在高密度人群中快速、准确识别体温异常者的智能能力,实现了既富有成效地保护医护人员安全,又及时有效地为感染者实施远程诊疗的医务智能能力。此外,在芯片研发方面,中国芯片几乎涵盖了所有门类^[3],除已经成功研制14纳米芯片外,7纳米芯片及更微观乃至精灵级芯片也正在研制中。AI在国外也取得了一系列重大进展,尤其是在交通、金融、医疗^[4]、教育、自身体态发展^[5-6]等方面均获得了重大突破,形成了诸多AI研究成果,并运用到政治、经济、社会和军事安全等众多领域^[7]。

可以认为,AI仿人并融合应用技术在全球范围内取得了实质性、引领性的突破发展,但AI的可持续性、稳定性、可信赖性和向上向善性发展还面临许多阻力、短板和偏见束缚。为了破除这些束缚,习近平总书记强调,“要整合多学科力量,加强人工智能相关法律、伦理、社会问题研究,建立健全保障人工智能健康发展的法律法规、制度体系、伦理道德”^[8]。这就要求我们,既要赋予人工智能技术价值,更要赋予人工智能社会价值和伦理价值。为此,迫切需要梳理AI仿人实现与AI智能层

级的演进过程及其价值调控与治理规律,从技术价值、社会价值、道德伦理价值等多方面助力 AI 发展。

二、AI 仿人实现与 AI 智能层级价值调控、治理及存在问题

(一) 前人工智能阶段

中国古代有诸多“仿人器物”,如历史上的“木人”“木偶”“木偶人”等。其中最早的记载是战国时期的《列子·汤问》,它讲述了西周国王周穆王向西巡视过程中巧遇关于“前人工智能”的典故。虽然,这个故事难以考证“前人工智能”存在的真实性,但是也足以证明其器物化的仿人工的木人,或者说类似机器人的仿人机已经在战国时期有了雏形,或者说在西周时期就有了人造“类人器物”的认知萌芽。然而,欧洲人普遍认为“前人工智能”最早出现在欧洲,其中“仿人机”被称为“似人自动机”,亦称为“类人自动机”,1920年在《RUR》剧本中最早被称为“做工的人”。可以认为,前人工智能阶段是客观存在的,它是后来 AI 发展的认知基础和实践基础。

综观中西方前人工智能阶段的发展情况,所有器物化的仿人工叙事,都有鲜明且共同的价值调控与治理特征和发展规律:一是前人工智能阶段的仿人器物,是利用物理属性制造简单器械来实现简单“仿人”,即仿人工程属于典型的朴素仿“人工”的低“智能”层级,是纯粹的以物理属性为工程知识的、实践化的仿人器官功能的价值调控型应用;二是前人工智能阶段的价值调控体系,主要依靠人为规定的禁忌、习惯法、纪律、行规、法律及道德伦理来实现。如中国皮影有演出的行规,有表演的职业道德,有表演内容价值的社会伦理引领和表演技术的价值引领,无疑它们反映、教育和映射的都是人本身。前人工智能阶段的后期出现了系统化、程序化的器械人工或机械人工,但依然是仿人的“人工”,只是因为这个“人工”是以诸多技术集合在一定工程架构中具有了一些“智能”,且按照一定“机制”程序化运行。但是,这个时期的仿人智能层级是低水平层级,仅是一些简单性、程序性、复制性甚至是单一性地依靠系统机械循环运行。但是,这个时期已经有了程序化的系统软件思维,这是区别于器物化仿“人工”时期的关键要素,旨在解决 AI 仿人的维度问题,没有太高深的仿人深度和仿人智慧。在前人工智能阶段,不仅仿人的体力行为,也仿人的部分脑力功能,只是存在很大的局限性,不是真正意义的建构人脑认知和思维,而是将一些知识性东西数字化后放入程序化系统或记录某一场景形成数据库等方便查找、搜集、显示。一方面它在仿人实践中逐渐实现为了人和服务人的价值旨归;另一方面它将人类娴熟的劳动技能通过较为复杂的技术实现转换,把一些体力和智力劳动让渡给人类生产生活的工具,使其机械地跟人学、跟人做,把人从重复、机械劳动中解放出来。从道德伦理学来看,其价值调控与社会治理主要表现为工具运行不能伤害人为最根本的价值原则,即让人工智能工具固定在某种劳动范围或运动轨道中运行。

(二) 弱人工智能阶段

这个阶段区别于前人工智能阶段的核心是仿人的维度、深度、价值调控治理体系和价值调控工程知识更加复杂,特别是运用大数据驱动仿人思想、情感和行为实现了更多维度、更加深化和更加细化,这时“仿人机”有了仿人的一些局部智慧,甚至这些局部智能一定程度超越了人的智慧。如人脸识别技术,机器人的“眼”比人的“眼”还“精”,它通过一张人脸照片的识别,而在海量图像数据中快速精准地查到与其有“血缘”关系的人脸,且识别差错率远远地低于人眼的识别差错率,识别速度也远远高于人眼的识别速度。

弱人工智能阶段何以取得在前人工智能阶段上的稳定性发展呢?首先,这是人类对人工智能认知升华的结果,是随着指数级数据和每两年计算机算法能力翻一番的摩尔定律出现的产物,是人的社会实践能力的积累和质的提升的必然结果。一方面,AI 计算能力和存储能力有了较大程度提高,其算法显著提速;另一方面,在大数据技术的助力下,AI 拥有了语音能力、视觉能力等先进深度学习功能,如:单词错误率低于 3.1%^[9],亚马逊 Alexa、谷歌助手 Google Assistant、苹果 Siri 和微软 Cortana 的语音识别系统均有较高的社会认可度,人工智能的视觉能力已经实现了 3.47% 的识别错误率^[10]。

事实上,仿人价值调控工程的知识基础是大数据技术支撑下的 AI 学习能力,亦称机器学习或称深度学习技术,即“深度学习通过建立类似于人脑的分层模型结构,对输入数据逐级提取从底层到高层的特征,从而能很好地建立从底层信号到高层语义的映射”^[11]。目前,机器学习在围棋、教育、金融信贷、健康医疗、气象学、城乡规划、交通、党建、太空探索等很多领域都取得实质性进展。但是,AI 智能层级仍然只是局部的仿人或仿人的局部功能,弱 AI 相当于人的“体力和智力”的智能化,需要把人的认知事物标准体系输入到机械终端,使其获得认知模型后才能形成仿人的机器智能。也就是说,弱 AI 仿人实现的价值控制体系主要源于一定体量的数据对机器的训练。当数据是真数据时,AI 能正常仿人劳动;当数据是恶数据时,AI 就不能正常仿人劳动。如此,数据的安全可靠、真实及其伦理价值便成为弱 AI 阶段价值调控与治理体系的主要内容和核心资源。

(三) 强人工智能阶段

强 AI 概念并非在计算机科学领域被最早提出,而是由哲学家约翰·塞尔在《心灵、大脑与程序》一文中第一次正式提出。哲学立场的强 AI 是“基于心智的计算模型,以通用数字计算机为载体的 AI 程序可以象人类一样认知和思考,达到或者超过人类智能水平”^[12]。这个阶段区别于以往任何阶段的关键问题和要素,是强人工智能具有了系统性仿人功能价值的调控价值、工程方法及工程知识,不再是弱 AI 阶段模仿人的局部功能,而是全方位仿人的一般性功能,具有了人的大多数和一般性处理事务的智力和能力,如认知、思考和道德选择行为。事实上,在人类世界已有了一个“类人类的机器世界”,其仿人原则、仿人价值调控治理体系和工程知识也不局限于一定结构化的“人型”,还包括了人性及彰显的万物联动时代的到来。即世界万物按一定程序、秩序和理性价值调控治理体系或规则进行“认知”“思考”并作出道德选择行为,全新地诠释着人们对 AI 仿人的“人工”“智能”和“人工智能”的爆发性和向上向善性深度智能融合的理解。这时的“仿人机”成了人类社会活动的得力助手,时刻解放并帮助人类解决线上线下的大多数社会活动事务。这个阶段的 AI 智能层级属于强人层级,其“体力+智力+”仿人实现的价值控制是因万物的互联和万物的自动辨识数据、收集数据、加工数据、数据驱动,从而实现 AI 具有一定的理性调控价值,能分辨并自动契合人的理性需求和自然宇宙运行的一般性规律。

(四) 超人工智能阶段

AI 发展到一定程度或一定阶段,受仿人智能水平实现全面超越的引领和制约,AI 仿人行为、思想、情感深度将更真更巧,AI 甚至要超越人的智慧,即出现 AI 能够创构新的 AI,亦称超人工智能。鉴于此,一些学者从哲学、伦理学等学科视角担忧 AI 发展可能出现的社会风险,纷纷阐述其主张并贡献治理和引领发展良策。其中,影响较为深远的主张是英国牛津大学人类未来研究院尼克·波斯特洛姆提出的超级智能理念,他认为超级智能将在几乎所有仿人维度上,远远地超过人类的智

慧,且超 AI 会给世界带来必然的社会风险:智慧生命灭亡并永久失去未来发展潜能^[13]。我们若能提前预防与干预,特别是把握好智能水平发展规律,以仿人原则实现有价值引领的社会调控,超 AI 将极大地释放人类全面自由发展的潜力,这样的超人工智能工具只能按一定社会价值调控与治理体系或规则有秩序性地为人类提供更加精细化的服务,其仿人原则的价值调控工程方法和工程知识升级为追求如何满足人的有限理性需求和向上向善需要。也许那时仿人的人工、智能和人工智能都超过了世界上最强体力的人和最聪慧的人,其体力劳动和脑力劳动都需要具有向上向善素质和智力水平,而运用体力与脑力服务于人类有限理性需求和向上向善需要,也就必然成为超 AI 的主要智能发展维度,这也是人类社会的发展过程中必然的需求变革使然。

三、AI 仿人的本然价值及其实现条件

由 AI 智能水平及价值调控与治理演进历程可见:第一, AI 仿人原则是推动智能水平演进的不竭动力;第二,仿人原则的内容逐渐从宏观走向中观,再到微观,价值调控与社会治理也更加复杂;第三, AI 的仿人现实性从单一的体力仿人向体力、脑力并重仿人方向发展,价值调控与治理推崇科技理性和善性;第四, AI 仿人实现手段和工程方法及工程知识也从器物的利用到器物现状的改变与塑造,再到用弹力、水力、风力、电力等根据能量守恒定律和物理属性来支撑研发应用的转变,以及依靠数据爆发功能来推进和彰显 AI“智能”层级甚至可能超越人脑,价值调控与社会治理更需要解放、助力和保护人的全面自由发展,即 AI 仿人迫切需要非理性、非逻辑体系与灵性规则融合发展。综述之, AI 一定程度上正一步步引领着人的全面自由解放与发展,当然,在智能水平发展过程及其仿人原则下的价值调控与治理演进过程中,由于价值调控与社会治理具有现实诉求性,难免出现一系列社会问题,如果不加以梳理和消解, AI 发展的奇点乃至人的贪欲等综合因素就会演化成疾,逆向酿成 AI 智能水平和仿人原则及价值调控与社会治理的乱象,从而导致人类社会的不和谐,乃至造成世界的动乱,最终导致人类自身面临毁灭性的生存危机。因此,遵循 AI 智能水平及价值调控与治理演进规律,我们势必从技术价值、社会价值、道德伦理价值等视角来研究并构建“人文+科技”的 AI 仿人价值调控与治理体系,力争最大可能和稳定性地推进 AI 在助力人的全面自由发展方面向上向善发展。

(一) 技术价值调控与治理是 AI 坚持仿人原则、提升智能水平的价值基础

特定社会的生产力是 AI 仿人的根本性前提和条件,如在仿人体力的前 AI 时期和弱 AI 时期,受特定历史时期的社会生产力本身的局限, AI 的发展必然是反映当时社会生产力现状的产物。即人类满足自身需要,只能从社会现实性上结合特定历史时期的社会生产力来实现价值调控与治理和生产生活实践应用。所以,没有特定社会背景下的物质条件作为基础来谈 AI 实践应用之做法、想法只能是空中楼阁,更何况人工智能必然要依靠社会生产力来推动其仿人体力和智慧的实践。正因如此,仿人体力时期的 AI 只能结合特定历史时期的社会生产力发展水平来研究,包括物质生产力和精神生产力的历史性发展水平。如前文的器物化仿“人工”时期和系统化机械人工时期,两个时期的“仿人机”重在模仿人的体力,目的是让人的手脚更长、更多、更持久代替人的一些体力劳动。总的来讲,“仿人机”主要是用弹力、水力、风力和电力等物理属性来实现工程建造。在农业上,让人类的重活、苦活、脏活和重复劳动得以被机器替代,或者得以减少人的劳动力消耗;在军事上,“仿人机”延伸士兵的手脚和武器,让士兵少流汗、少流血、少牺牲;在企业生产上,让恶劣环境的工人获得

较多解放、自由和保护,且进一步降低生产成本、提高生产效率、减少安全事故。当然,在仿人发展过程中也出现过负面效用,特别是侵略性的战争兴起,往往就是因为仿人体力的“仿人机”暂时性领先了弱国的国防实力。因此,这两个时期同样需要 AI 仿人的技术、社会和道德伦理等价值调控治理体系来实现社会调控。尽管人类文化中科学技术、社会制度文化和道德伦理起到了一定作用,尽管非 AI 和前 AI 阶段已经过去,但是弱 AI 发展仍然迫切需要结合仿人层级和仿人工程方法及工程知识来协同构建多元的价值调控与社会治理体系。

(二) 社会价值调控与治理是 AI 以仿人原则保障服务全人类的价值依靠

一切围绕人和为了人的大数据及融合应用器物只有在以全人类为中心展开助力活动时,才能全面地实现为了人这个旨归。特别是在仿人体力与脑力并重的弱 AI 时期和可能到来的强 AI 时期,有着复杂系统的弱 AI 阶段和强 AI 阶段,是“仿人机”已然开始利用信息大数据驱动的新融合发展阶段,不是简单地仿人体力,而是进入了对人的体力与脑力同步进行模仿,旨在解决机器具有“人工”和“智能”一体化的创构阶段和优化阶段。目前,以“电能+机械能+大数据驱动机器学习”模式来实现仿人价值工程。机械能是以电能、光能等转化为机械体能来代替人的体力;大数据驱动机器学习则是以电能、光能等带动传感器,搜集人们的行为数据和一些数据化了的知识定理,并根据已有的大数据进行学习策略的探索和潜在结构的发现,获得一定的运行模型进行预测和分析^[14]。当然,这种人工与智能不是简单的知识和技术叠加,而是利用大数据分析仿人维度,从而驱动机械能实现设计目标或仿人目标,进而彰显“智能”。由此可见,机器学习的必备条件是大数据和大运算,前者是学习的内容,是关于整个人类的思想、情感和行为数据,后者是学习的方法,是处理关于人的行为大数据的方式和模型。没有大数据,AI 没有可学习的资源和动力来源,也就不会有仿人的维度,就会脱离人;没有大运算,或者说没有算法,也就无法将仿人维度落实在具体机器行为上,更无法处理关于人的大数据和搜集关于人的大数据的爆发性价值,从而彰显人工智能的“超越性”智能。显然,这两个阶段的 AI 仿人价值调控与治理体系需要结合大数据与一定算法等 AI 仿人的价值调控工程知识来构建,实质是这两个阶段的 AI 仿人价值调控与治理体系需要结合全社会开放和共享数据来助力实现 AI 仿人的全面性和灵活性。因此,要求全社会要有开放和共享数据的制度保障,要有开放和共享数据的文化、习俗及价值观等生态保障核心要素,这便是 AI 仿人的本然价值——社会价值。

(三) 道德伦理价值调控与治理是人与 AI 向上向善发展不可或缺的价值条件

人类的理性需求和向上向善发展需要也是 AI 仿人的根本遵循和必备要件,所以在追求满足人的理性需求和向上向善发展需求的强 AI 和超 AI 时期,数据驱动下的 AI 功能性智能比较稳定,仿人维度全面,超越性较强,具有仿人的向上向善性及向上向善性具有鲜明的鲁棒性,即求真、求善、求美具有精准性、综合性、和谐性和全面性。当然 AI 的仿人和助力人类活动范畴也不再局限于一般性仿人普通事务的“仿人机”,甚至具有了前瞻性、预期性和德性的能“满足人类理性需求和向上向善需要”的 AI“服务机”。因为这个时期,仿人的体力和脑力都超越了普通人,其存在的关键价值是利用已有的仿人“体力+智力”的高度智能融合发展,为人类的理性需求和向上向善需要服务,即及时、客观地发现人的理性社会需求和向上向善需要,并提供精细化的供给服务。人类的理性需求和向上向善需要,主要指人的有限需求,以及没有贪欲和没有损人利己行为的美好生活需求。总之,这个时期的 AI 工作重心和道德伦理价值调控与治理目标是追求满足人的理性需求和向上向善需

要,助力实现人民对美好生活的向往。这就类似于人的餐饮,按照营养学和生理学的要求,人们要按计量按比例配送输入,而不是每日三餐大鱼大肉。当然,这种服务依赖于工程方式和工程知识,可能是“电能/光能/新能源”+“机械能”+“数据化的机器学习”,也可能是“机械能”+“新技术”/“新算法”/“新材料”并依托这种新技术或新算法或新材料来驱动机械能。因为超 AI 受仿人维度、深度的影响需要进一步精细化和微观化,其智能程度必然倒逼运算速度大大提高和更加稳定,这就要求其运算方法上也许要创新另一种新的方法。正如“量子计算机存储量子比特,一个量子比特可表示量子态 $|0\rangle$ 和 $|1\rangle$ 的叠加,一次运算就可同时处理两个状态的信息,以此类推亦然”^[15]。而当下的非量子电子计算机存储、计算等能力都还不高,如每次只能处理一个比特的状态数据等还需要进一步优化和技术突破。当然,也有可能是另外的新技术、新材料来实现超 AI 仿人维度的有价值且是安全、和谐的落实和落地。

四、AI 智能层级与仿人现实间的价值调控与社会治理策略

器物化的仿人工时期和系统化的机械人工时期都已成为人类社会的历史,而当前 AI 发展需要注意的是,AI 始终是人类社会的工具类具象事物,其性质不会发生漂移,不会形成变异,但是要把握好未来 AI 由弱到强、再到超强智能水平的依次演进发展准备。特别是要结合 AI 演进发展的仿人维度、深度的客观现实诉求,切实实施以一定社会生产方式加以价值调控与治理,才能保证 AI 与人类发展相互作用,达到与人类合伦理原则和可持续发展需求的预期,确保并彰显人类研发与利用 AI 的初衷和旨归。2019 年 7 月,中央全面深化改革委员会第九次会议审议通过了《国家科技伦理委员会组建方案》^[16],旨在从国家层面重视并引领和干预 AI 的价值调控、治理体系之演进发展。当然,国际上诸多国家和地区已经纷纷出台助力 AI 发展的国家战略和价值调控与治理措施,如美国、欧盟、英国、日本、西班牙等 20 余个国家和地区先后发布 AI 发展战略、发展规划和发展计划^[17]。但是,综合来看,AI 智能层级与仿人现实之间的价值调控策略主要表现在技术价值调控、社会价值调控和道德伦理价值控制三个方面,只是三个层面的价值调控与治理还需要形成一些创新性共识,以求在更大范围助力 AI 可持续发展。

(一) 社会价值的创新性保障共识及策略:倡导安全、可靠、开放、共享的数据观

如在弱 AI 阶段的价值调控与治理中,由于 AI 仅仅是仿人的一些局部功能,其仿人原则在其现实性上依然是通过“电能+机械能+人的局部行为大数据供给机器学习”来实现,容易出现问题的领域和地方是数据搜集、利用环节,只要做好数据搜集和利用环节的价值安全调控工作,就能保障 AI 仿人原则的价值调控和价值实现。为此,需要加强以下三方面工作。

首先,建立数据权益归属机制和保障机制。数据的产生者往往是个体,数据的搜集者、利用者往往是大型实力企业和政府机构,显然,数据的产生、搜集、利用三方的典型性非契合、非均衡事实,容易导致数据权属争议,或者说是所有权、收益权、处置权等法律制度和机制的严重缺位和缺失,可能引发数据垄断、数据权利和数据暴力等社会调控之失控现象。为此,需要“结合实际,以法律立法、政府引领、行规自律,以及利用大数据本身技术突破来支持大数据时代个体理性化进行协同治理”^[18],才能科学建立数据权益归属机制和保障机制。

其次,建立筛查数据隐私调控机制。由于各个智能等级的 AI 均需要大数据驱动,当然不能简单地“拿”个体的隐私数据用于驱动。为此,有必要用语法隐私技术和语义隐私技术来建立数据筛

查系统,阻止隐私数据外泄和被不自觉或自觉的不理性的应用场景所运用。前者,是为保护隐私信息的需要,利用现代科学技术直接改变和处置数据集中的某些属性的值,使得数据在搜集、储存、分析、交易、利用中将属于隐私部分的数据不被选中而被应用,或者需要人机对话后由有特殊管理职能和权限的人员授权后才能应用;后者主要通过严格定义的数学隐私保护模型且已经形成了的隐私保护理论框架的黄金法则来实现隐私数据的保护^[19]。

第三,建立国家级数据监管中心。“由于 AI 是一类需要大数据作为能源的智能器械,还是一个自带闭环式的数据训练模型器械,所以建立 AI 数据国家级监控中心,目的就是対 AI 内外的大数据数量、类型、价值、安全等进行最高级别的且是少数人掌握的管制,确保人工智能的正确价值观有‘源头活水’且不受外界‘污染’”^[20]。同时,又能由权威管理人员在应急中及时实现价值调控与治理,包括重大问题的溯源倒查,即保障高质量的价值调控与治理权在权威机构和人员手中。

(二) 技术价值的创新性保障共识及策略:融合现代科技强化人机同伦的技术调控

如在强 AI 阶段的价值调控与治理中,由于强 AI 具有了人的一般性功能,仿人原则在一定维度上反映的是人的一般性事务的能力,包括体力劳动与脑力劳动两个版块,自然在其社会实现上要注重人与机器关系的建设和调控。一方面,建立“人机同伦系统”并强制嵌入 CPU 系统。“虽然人工智能再一次迎来了发展的‘黄金期’,但我们不可否认这个世界上永远不会存在‘完美’的算法。机器和人一样也会犯错误。甚至在很长一段时间内。基于当前的人工智能发展水平,机器犯错的比例可能比一般人还要高。在这种假设前提下,人工智能的应用将带给人类的第一个伦理难题很可能就是:机器如果犯错,该如何承担责任与风险?”^[21]为此,有必要将人的道德伦理和社会核心价值观数据化为 AI 的有机性道德伦理及核心价值观,让人的道德伦理和社会核心价值观成为 AI 的学习资源,让 AI 与人一样需要遵循一定的道德伦理和社会核心价值观。同时,让这种遵循受“人机同伦系统”强制嵌入 CPU 系统的限制,AI 仿人价值调控体系功效是长期性、稳定性和秩序性的,“即将这样的系统像硬件一样融合入中央处理器,不能选择应用,不能删除,始终伴随智能终端应用生命期。即使这一系统出现问题不能应用,那也得使整个智能终端不能运行”^[20]。

另一方面,赋予 AI 人一样的“户籍身份”并从严落实备案制度和机制以强化管理。尽管强 AI 将为人的社会活动提供大量帮助,成为人类社会活动得力的道德工具和道德助手,但若不加像人一样拥有“户籍身份”的监管,它在线上线下的活动就难以溯源和接受必要的倒查管控,一旦发生贪欲、侵权等非理性、非道德现象甚至是犯罪行为,就会出现法治对象空白、缝隙、含糊空间,而无从或不利于精准追责问责。为此,为了防止非理性、非道德现象甚至是犯罪行为出现,有必要在全社会建立一种全覆盖的赋予 AI 人一样“户籍身份”的备案管理机制及技术,即以具有较强保密性功能的“区块链”技术来分辨和区别 AI 个体,从而实现 AI 具有人一样的“户籍身份”。当 AI 出现坏的、恶的或是犯法行为等问题时,相关管理部门就能在第一时间,甚至是通过算法在更早预期内知道是谁或将是准行使的 AI 工具,在什么情况下出现了或可能出现什么问题和责任是谁等事务,并能通过国家级数据中心直接实现一定的行为干预控制或提前预警防控。

(三) 道德伦理价值的创新性保障共识及策略:制订 AI 服务清单

如在超 AI 阶段的价值调控与治理中,受超 AI 必然超越世界上“最智慧的人”的限制,仿人维度必然从以仿人的体力、脑力为基础,向追求如何满足人的理性需求和向上向善需要服务转变,即“仿人机”向“服务机”过渡。这种过渡的目的在于指 AI 向着更高级的智能水平发展和向更高级的仿人维

度、深度发展,即实现对人的理性需求和向上向善需要的识别、供给服务。核心的价值调控与治理对策就是要进一步梳理 AI 的服务清单,只运行符合关于人的且是理性需求和向上向善需要的价值目录范围内的程序。这就需要在超 AI 时代来临之前,结合具有中国特色的道德伦理思想,西方的理性价值伦理,特别是运用马克思主义伦理学体系、社会主义核心价值观,通过研究,形成人类命运共同体的理性需求和向上向善需要共识目录和共识清单,并综合运用生理学、营养学、心理学、社会学及社会管理学等诸多学科理论对人的理性需求和向上向善需要进行分类,形成规范的共识目录和共识清单,让 AI 执行这一共识目录和共识清单,实现 AI 像人类一样具有道德伦理观念、核心价值观、辩证性和灵活性“智慧”,具有真、善、美情怀和价值追求,具有只为人类共同的理性需求和向上向善需要服务的意识、行为和情感价值趋向,充分发挥其机械的“能动性”,智能屏蔽为任何个人与组织去干坏事、恶事、丑陋事或犯法事的仿人维度、仿人深度的训练和智能。即让 AI 在道德伦理及一定的核心价值观的价值控制下成为增强和补充人的认知和参与社会活动的得力助手和道德助手,包括成为文化创构在内的理想工具。这就需要顺应超 AI 的仿人原则,对 AI 的道德伦理等价值调控与治理系统进行升级改造,让 AI 的道德伦理等价值调控与社会治理系统也有“超人”的智慧。

五、研究与总结

人工智能是由人在特定社会活动诉求中,结合特定社会生产力而诞生的为了人且是服务人的先进性理性价值工具,从其演进发展过程和未来发展的可能性来看,可分为前人工智能、弱人工智能、强人工智能和超人工智能四个阶段。而任何阶段的 AI,除受所处时代生产力引领和影响外,还受本身发展的仿人维度、深度引领和制约,从而呈现出不同智能层级水平、价值调控与社会治理手段和知识。当然,每一层级的智能水平都有其一定的技术价值、社会价值和道德伦理价值等重要价值调控与治理体系,以保障 AI 实现为了人和服务人的工具目标价值。技术价值调控与治理体系方面,要融合现代科技强化人机同伦的技术调控应用;社会价值调控与治理体系方面,要倡导安全、可靠、开放、共享的数据观;道德伦理价值调控与治理体系方面,则需要制订 AI 服务清单。综而述之,当这三个方面的价值调控与治理体系协同实践并实现其价值属性,AI 的发展与应用就能最大可能和最为稳定地推进其成为人的全面自由发展的向上向善的人类工具,成为人的得力助手和道德选择助手。

参考文献:

- [1] 约翰·马尔科夫. 与机器人共舞:人工智能时代的大未来[M]. 郭雪,译. 杭州:浙江人民出版社,2015:14.
- [2] 习近平. 推动人工智能深度运用[EB/OL]. (2020-07-16) [2022-01-09]. https://politics.gmw.cn/2020-07/16/content_33999204.htm.
- [3] 中国芯片设计云技术白皮书 2.0[EB/OL]. (2022-04-21) [2022-04-26]. <https://new.qq.com/rain/a/20220421A012I200>.
- [4] 科技前沿[J]. 信息技术与信息化,2013(4):5-7.
- [5] MISKIN M Z, CORTESE A J, DORSEY K, et al. Electronically integrated, mass-manufactured, microscopic robots[J]. Nature, 2020,584(7822): 557-561.
- [6] BOUMAN A, NADAN P, ANDERSON M, et al. Design and autonomous stabilization of a ballistically-launched multirotor [C]//2020 IEEE International Conference on Robotics and Automation. May 31-August 31, 2020, Paris, France. IEEE,

2020:8511-8517.

- [7] 王海龙, 张昕婧, 曹茜茜, 等. 美国《人工智能国家安全委员会中期报告》的重点与启示[J]. 机器人产业, 2020(3): 23-25.
- [8] 习近平. 加强领导做好规划明确任务夯实基础 推动我国新一代人工智能健康发展[N]. 人民日报, 2018-11-01(1).
- [9] AMODEI D, ANUBHAI R, BATTENBERG E, et al. Deep speech 2: End-to-end speech recognition in English and Mandarin [EB/OL]. (2015-12-08). <https://arxiv.org/abs/1512.02595>.
- [10] DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. June 20-25, 2009, Miami, FL, USA. IEEE, 2009: 248-255.
- [11] 余凯, 贾磊, 陈雨强, 等. 深度学习的昨天、今天和明天[J]. 计算机研究与发展, 2013(9): 1799-1804.
- [12] 陈自富. 强人工智能和超级智能: 技术合理性及其批判[J]. 科学与管理, 2016(5): 25-33.
- [13] 尼克·波斯特洛姆. 超级智能: 路线图、危险性与应对策略[M]. 张体伟, 张玉青, 译. 北京: 中信出版社, 2015: 143.
- [14] KOHAVI R, PROVOST F. Special issue on applications of machine learning and the knowledge discovery process[J]. Glossary of Terms Journal of Machine Learning, 2017, 30(1998): 271-274.
- [15] 黄一鸣, 雷航, 李晓瑜. 量子机器学习算法综述[J]. 计算机学报, 2018(1): 145-163.
- [16] 习近平: 紧密结合“不忘初心、牢记使命”主题教育 推动改革补短板强弱项激活力抓落实——李克强王沪宁韩正出席[N]. 人民日报, 2019-07-25(1).
- [17] 中国信息通信研究院政策与经济研究所. 全球人工智能战略与政策观察(2020) [EB/OL]. (2020-12-30) [2021-05-13]. https://www.sohu.com/a/441355353_781358.
- [18] 潘建红, 潘军. 大数据时代个体数据理性化悖论与消解[J]. 甘肃社会科学, 2018(2): 45-50.
- [19] DWORK C. Differential privacy: A survey of results[C]. Proceedings of International Conference on Theory and Applications of Models of Computation. Springer, 2008: 1-19.
- [20] 潘军. AI 理性价值智能的隐忧与消解[J]. 自然辩证法通讯, 2018(4): 20-25.
- [21] 崔鹏. 人工智能演进中的伦理悖论与社会危机[J]. 大数据时代, 2017(2): 27-33.

Artificial intelligence level and value regulation and governance of human-imitation

PAN Jun^{1,2}, YAO Kemin³

- (1. School of Marxism, Guizhou University of Commerce, Guiyang 550081, P. R. China;
2. AI Big Data Center, Tsinghua University, Tianjin 300467, P. R. China;
3. Center for Science, Technology and Economic Strategy Studies, Guizhou Academy of Sciences, Guiyang 550081, P. R. China)

Abstract: Artificial intelligence is such a dynamic and advanced tool as people are committed to continuous innovation and in-depth innovation exploration on the development path of how to imitate humans and which functions to imitate humans. Its level of intelligence is basically verified or measured by the dimension, depth, and validity of imitation; its imitation value regulation and social governance is gradually realized in the “humanities + technology” way of technological value, social value, moral and ethical value and other research results. As McEwan pointed out, since human history, no matter what cultural background they are in, people have always had a common dream, that is, to create an “artificial version of ourselves” non-reproductively, so that we can act like a God and create human beings as God did, that is, to create

multifunctional artificial men. It can be seen from the perspective of people-oriented development that, the evolution and development history of AI is based on the “type” of human beings, constantly innovating and imitating people’s physical strength, intelligence, and “physical strength + intelligence” to achieve higher and deeper levels, and “human beings” are made intelligently to achieve the practical exploration of history that further liberates human beings. AI in the entire evolution process can reflect the formation and evolution of AI intelligence levels in a certain dimension and depth at the same time, and the operating paradigm is “human imitation”. At the same time, engineering knowledge such as the technical value, social value, and moral and ethical value of AI imitation principles and imitation values regulation has been gradually accumulated and improved in this process. On the one hand, the development of AI from mechanical mechanics technology to informatization and intelligent technology has enabled the development of the single liberation of physical labor towards the full service of human freedom and liberation. In a sense, AI has obtained theoretical accumulation, technical accumulation and a large-scale engineering paradigm. On the other hand, the more common and in-depth applications of AI have caused people to worry about their own future social activities, which set off a wave of value control research that reflects the intelligence level with its imitation of human intelligence, such as the rise of research on technological value, social value, moral and ethical value of artificial intelligence. For this reason, we should sort out the evolution and development of AI and its value regulation laws, especially clarify the value standards, value regulation measures, and value regulation quality and effectiveness of the human imitation principle in each course. Starting from each specific intelligence level, the value control knowledge of the imitation principle should be used to guarantee the sustainable development of its positive effects, and the elimination and shield of the negative effects. It is helpful to promote the comprehensive and in-depth intelligent integration application of AI under the upward and good value regulation, to achieve broader and sustainable deep intelligent integration application of AI, and to realize that AI is only a helpful assistant in human social activities in the future society which is a morally chosen laborer who promotes people’s social activities 24 hours a day, online and offline.

Key words: artificial intelligence; intelligence level; humanoid principle; value control; social value; ethical value; social governance

(责任编辑 彭建国)