

对人工智能的目的反思*

邓达奇

(深圳市社会科学院政法研究所, 广东 深圳 518028)

[摘要] 人工智能是一柄双刃剑, 反思其技术问题、引导其积极效能是哲学研究的基本任务。无论是人工智能的现状还是未来, 人类真正需要并能够控制的对象, 与其说是智能机器, 不如说是人类自身, 因此人工智能技术的真正问题在于人对待世界的问题。突破传统人工智能“技术漏洞—制度补强”的修补式思辨路径, 以人工智能作为反观人类自身得失、促成美好生活实现的窗口, 基于“人之主宰地位”及“人之本质定义”的视角重申敬畏之心、德性挺立之于人类生命及世界秩序的重要价值, 可得到人工智能于美好生活的新的哲学向度。就此而言, 人工智能作为一种人类发明的技术逻辑, 终究是需要回归社会实践的, 把握人工智能的安全性, 以确保负责任的技术创新, 是人工智能技术运用于生活的基本底线。而在美好生活的向度层面, 首先就必须正视和接纳人工智能为人类之秩序与安宁带来的系列福利, 其次要反思人类社会既有制度以及人性取向既有缺陷。在此基础上, 反省现代社会人的价值追求、增进其对于他者的同情、对于世界的包容, 既需要积极利用智能技术来提升生产效率与生活品质, 也必须完善与之相应的人文及伦理与法律制度建设, 从而最终实现美好生活所要求的, 效率与公平、技术与思想、物质与精神的双向并进。

[关键词] 人工智能 人类 反思 德性 美好生活

[中图分类号] B82-057; TP18 **[文献标识码]** A **[文章编号]** 2096-983X(2024)01-0139-09

既有的人工智能应被归为技术问题来看待, 因而无论是人工智能的设计过程还是其应用结果所引发的问题, 诸如大规模失业、专制的可能性增大、社会关系异化、智能武器促成战争频发、群体之间的文化歧视、权力和资源分化等^[1], 根源基本都可以归约到人及由人所构成的社会本身。换句话说, 人工智能并非新问题的制造者, 而是旧问题的放大镜或显微镜, 即它只是使得早已存在于人类社会的诸般问题加以凸显而已。更重要的是, 若是未来的人工智

能真正超越了单纯的技术定位, 即在情感、欲望、价值等层面实现了自主和自由, 人类也未必就能如预期那般, 保有对它的控制权, 反倒极可能使自身陷入无比危险的境地。^[2] 总而言之, 无论是人工智能的现状还是未来, 人类真正需要并能够控制的对象, 与其说是智能机器, 不如说是人类自身。

然而, 目前学界关注和研究重心主要侧重于, 对人工智能所导致的或可能导致的社会问题来设置规范。典型成果包括: 2016及2017年的年

收稿日期: 2023-05-31; 修回日期: 2023-11-07

*基金项目: 2023年深圳市社会科学院专项科研课题“深圳哲学社会科学领域意识形态阵地建设的发展路径研究”(2023AB006)

作者简介: 邓达奇, 法学博士, 西南政法大学法学博士后, 研究员, 主要从事法理学、法社会学研究。

底,在计算机、消费性电子产品等领域具有权威地位的电气和电子工程师协会(简称IEEE,全称Institute of Electrical and Electronics Engineers),先后发布的两版“人工智能设计伦理原则”^[3],以及2017年1月美国加州举办的主题为“福祉——人工智能(Beneficial AI)”的研讨会所发布的“阿西洛马人工智能原则”^[4]。同时,国内学者近两三年围绕人工智能的主体与客体之争,其实也主要是针对人工智能而发的研究成果。这也就意味着,即便是针对人工智能本身的探究,始终也难离人这一绝对的主体。

相较之下,借人工智能来反向审查人类自身,则不只关注人工智能所引发的负面效应,并将之作为人类自身问题的投射结果来对待,而且更进一步以整体的技术进化为窗口,反观人类自身的定位与完善之道,为美好生活的实现规划切实可行的实施方向。本文即以此为研究进路,首先分别解读人工智能对“人之主宰地位”及“人之本质定义”所带来的挑战与启迪,重申敬畏之心、德性挺立之于人类生命及世界秩序的重要价值。在此基础上,进一步总论人工智能之于人类美好生活向度一体两面的直接效能:既提供了相应的技术与物质支持,也扩大了社会的既存危机。为此,人类在积极发展和利用人工智能的同时,也必须完善与此相应的人文及制度建设,以期最终落实美好生活所要求的,效率与公平、技术与思想、物质与精神的双向并进。

一、人工智能所引发对“人之主宰”的反思

在人类改造世界的的能力相对落后的时代,往往会将个体或群体的命运交由外在神祇或超越性天道,以寻求自我及世界的解释与安定。因而早期人类文明,不约而同地以图腾崇拜、神化传说等神秘要素来建构认知及文化的体系。依照马克思的解读,宗教世界图式的本质是现实世界的“颠倒”,即人在面对现实的无奈和无力

时,寻求安顿和秩序的结果。正是因此,自近代工业革命开始,随着人类科技的发展、知识的累积,生存性欲求的满足变得更加便利,面对外部世界的无知感和无力感也趋于减少,宗教及宗教性也随之从自然探究、物质生产等领域不断退却。伴随“上帝”在现实世界的“退场”,人类开始逐步扮演自我欲望和外在于自然的主宰者角色,并借由技术和物质的持续扩张强化这一身份认同。

然而,人工智能却以其不同于以往技术物的能力及发展趋向,在某种程度上对人之主宰地位造成了挑战。尤其是随着近五年来深度学习神经网络算法的开发、物理及网络大数据的广泛运用,使得其内在运作逻辑,不再局限于固定化的专家系统或者数据模型,而扩充为一整套为人所不可知的算法过程;其呈现结果,不再表现为机械的模仿人的对话和行动,而延伸到远超于人脑的计算和创造能力。这在某种程度也预示着,人工智能在起点上虽是由人所制造,但最终展现的成果,却已然超出了人所能控制和理解的范围。换句话说,人类大脑在数据的储存、记忆及联想能力上的相对有限,使人工智能超出人之控制其实不仅是可能,而是正在日渐成为事实。当然,这并不是说人工智能已经发展到科幻小说及电影中所描写的,能够直接反向戕害人之身体的地步,而是说人工智能与其说是人所制造,不如说是“上帝”赐予人类新的礼物。因为,机器本身及其学习的数据或素材虽然都是人所提供的,但人自身却无法对之进行精细、完整的把握和呈现,而人工智能在既有的部分运用领域,却能如全知、全能的上帝一般,将之一一刻画和归类。更重要的是,这些刻画的过程和结果,最终只能由其制造者来反向猜测。所谓人工智能的“黑箱”,其本质也正在于此。如此一来,近代以后树立起来的人之创造主体、主宰地位等形象,到底还能体现在哪里,是否真的只有神秘、超越的上帝,才是万物真正的创造者和主宰者?近代以来所崇信的人之绝对主体及无限可能,是否只是人类

自身的臆想?

对于上述问题,可以从以下两个角度来回应:一者,将人工智能视作是人的创造物,将其全部成就最终都归于人,并借此进一步强化人在万物面前的存在感与优越性;二者,借由人工智能来重审人在世界关系网络中的定位,借之重审人类中心主义的生存态度,回归到对“未知世界”的敬畏上。前一种态度是坚持传统思维,通常会将人工智能目前的发展困境视为人类制度变迁之缺陷,因而需要加强人工智能的程序正义、标准建设,以消除算法黑箱带来的风险,这是“人工智能—人”双向的关系传统层面。相较之下,后者转换了视角,在本体论及方法论上无疑更为具有前瞻性与反思性。因为,即便就人工智能在特定领域的发展现状来看,人类智能已然无法匹敌。若是机器的深度学习,可以进一步扩张到德性、情感把握等领域,那么人类的优越性更是无从谈起。更重要的是,人工智能虽是由人所制造,但它却不同于任何以往的人造工具或机器,其在智能上取得的既有成就,以及未来可能的发展态势,使之不仅有助于增进人的理解力与行动力,强化人主宰万物的地位;而且同样也有机会作为独立的他者而存在,甚至反噬人类秩序。这意味着,科技的发展固然彰显了人的能力,无疑却也让人更清楚地知道世界之大,并非只有人而已;世界之复杂,更不是人所能随意支配的。

鉴于此,一方面,人需要重拾农业时代对于他者及世界的敬畏之心,承认并直面世界乃至人自身的生存图景,并非全然可控的对象。因此,如何遵循“知其不可为而不为”的自我约束,便应该成为人类自我定位的重要原则,切不可一心妄想做自我和世界的主宰。对此,庄子所谓“吾生也有涯,而知也无涯。以有涯随无涯,殆已;已而为知者,殆而已矣”^{[5](P104)},或许正可作为人类理解自我的核心态度。当然,这并不意味着人要放弃对于自我和外部世界的认知,只是要为理性套上必须的枷锁,以免其过于狂妄,占据了“头顶的星空和心中的道德律”本该有

的地盘。

另一方面,人需要培养克服欲望及虚荣的平淡心,扬弃近代工业革命以来将自身视作唯一中心,可以凌驾并操控他物的错觉。人作为万物的一分子,毋庸置疑拥有自身的独特乃至优势之处,但这并不意味着人可以遗忘自然界本身的功绩,更不代表人可以依照一己意图来改造外在世界。人在整个世界或许是居于食物链的顶端,但食物链的完整及良性循环,其实少不了任何一环的存在与价值。换句话说,万物跟人一样都有其适于自身的运作逻辑,如何在有效利用它们来为人类谋求福利的同时,更应当抱着一种“自然命运共同体”的态度来形成彼此的共生与互助。

再一方面,人需要培养独立人格、彰显大爱的道德心,以真正从内在发出对于包括自身在内的人类社会的拥抱与宽容。世界的不完全可知,并不止于外在自然现象,也体现在人类社会之国家、族群之间的文化差异上。人类中心主义在既往历史中的极端表现,尤其表现在人类内部的国家侵略、文化倾轧层面。人工智能所展现的超智力,无疑足以打破一切种族所可能持有的自我优越感,消解特定群体可以凌驾于其他群体的合法性。这也就意味着,人类世界内部应然的相处模式,应是在保持自身族群文化独特性与凝聚力的同时,对他者文化秉持足够的包容与同情,尽力消除国家、社会之间或内部人为制造的文明与野蛮、中心与边缘相对立的生存态势,努力以“人类命运共同体”的原则来建构各美其美、美美与共的大同世界。

需要强调的是,上述因人工智能对人之主宰地位的冲击,所提倡的对于“未知之地”的敬畏,并非是要人回归到巫术式信仰、盲目性崇拜的时代,而是要明了人不过是万物之一,有其特殊性,但未必有其超越性;人有智慧,但绝不意味着其他物种便缺乏智能;人可以寻找自身的特殊性,但万物也有其独属于自身的逻辑;人类社会族群可以不断强化内部的认同感,但也必须对他者的文化保持足够的同情。总而论

之,唯有依循“万物并育而不相害,道并行而不相悖”^{[6](P37)}的原则来安顿世间,才能成就天地生生之大德,也才是对于人之主体地位最为有益的解答。

二、人工智能所导致对“人之本质”的哲理审视

“人的本质是什么”的问题,背后其实是一个诸如“我是谁”“我从哪里来”“我将去向何方”“我与世间万物是什么关系”“我该跟世间万物如何相处”等的问题集。对于这些问题的回答,既可以是生活化、表面化的,也可以是哲理性、抽象性的。对于“本质”的追问,无疑属于后者,但却也必然借由前者来通达。思考方式及其所得结果,却可能会因为地域文化及时代发展的差异而出现不同的答案。这也就意味着,对于“人之本质”的诠释,其实是存在多重可能的。在此,仅以时代变迁为线索来予以简要展示。整体而言,人之本质,也即其区别于其他物种的核心特性,主要包括意识自觉与道德自觉两大面向。然而,纵观人类历史的发展趋势,人之本质的界定则存在着从“尊德性”到“道问学”的转向。

在雅斯贝尔斯所称之为的“轴心时代”,东西方文明都不约而同地迈入了人文大爆发的历史,其典型特征则是人类灵魂及其所蕴含的道德属性被高度推崇,并以之作为人与动物,乃至人类内部之间的等级划分的核心标准。例如,苏格拉底借由神谕提出“认识你自己”,正是要人重审自我的灵魂,将以往一切习以为常、自以为是的知识,放置到无限的追问和反省中,以抛却诸多似是而非的理解或行动,通达“美德即知识”的人生新标准。柏拉图的“理想国”,则将个体之善与城邦之正义对应起来:个体应以理性为主宰,引导意气,节制欲望,以实现灵魂的和谐与美德;城邦则应以代表智慧的哲学王作为统治者,引导勇敢的护卫者,管理节制的劳动者,以实现城邦的秩序与正义。亚里士多德

则一方面以幸福作为人类特有的终极目标,并以灵魂中的品性之优与智力之优,也即德性与理性,作为通达幸福的路径^[7];另一方面又提出“人天生是政治动物”,要人借由“总体的守法正义”与“部分的具体正义”,实现对他人的善,共同建构和维系城邦的善。

中国则出现了以儒、道、墨、法为主体的先秦诸子学,他们从“六经”之中各取所需,以求针对时弊提出各自的有效策略,最终让统一的“道术”裂变为春秋战国文化界的百花齐放。儒家之孔子提出的“为仁由己”、“克己复礼”、“孝悌也者,其为仁之本”,建构起一整套以“仁”与“礼”来定义人的价值体系。其后,孟子及宋明之际的陆王,沿着心性自主、德性自足的一面不断延伸,最终发展出一套以人之良知、良能来诠释世界的道德哲学。与之相对,荀子则通过对“化性起伪”、“隆礼重法”的强调,试图借助礼制的张扬,推动人心及政治的善治。道家之老庄,主张“道通为一”、“自然无为”、“致虚守静”等理念,以期削减现实之造作、化解人际的冲突,实现心灵的自足与逍遥。墨家则希望借助“非攻”、“非乐”等,反向降低人世的冲突与浪费现象;通过“交相利,兼相爱”、“尚贤”等,正面塑造现实之友好与良善。

相较之下,近代的文艺复兴与启蒙运动,则开启了人类观念中的欲望与理性;工业革命则直接让人体验到自身能力的强大。在此期间,虽说天文学的发展,让地球失去了宇宙中心的位置,牛顿、达尔文以来的自然科学,又让人类失去了生物意义上的特殊性。但整体而言,人从神所塑造的世界与形象中解放出来,为自身的独立自由、科学理性而振奋,无疑更应该被视为近代人类世界的主流。中国自鸦片战争之后,被迫打开国门来“开眼看世界”,自然也在自觉与不自觉中被此一主流所改造,以致“科学”的旗号一度遮蔽“启蒙”的价值。

概而言之,近代以来的人类重新审视和定位自己,主要表现为三方面的工作:一方面,人从神的掌控中解脱出来,丧失了天之骄子的地

位,以致自身的特殊和本质之处不得不面临重构;另一方面,文艺复兴、启蒙运动及宗教改革等思想观念上的革命,促使人逐步认可身体的欲望与内在的理性,并努力对之加以彰显及扩大;再一方面,工业革命的开启、自然科学的发展,使得人之欲望在改造自然的过程中不断被满足,人之理性则在征服外界的效果中得以确证。由此,近现代人类借助科技和市场所形成的强大实践能力,不仅再次确立了自身相对于万物的特殊性,而且使自身几近于取代上帝,成为世界的新主宰。人之为人的本质界定,也随之既包含着对于尘世物质欲望的追逐与满足,也统摄着理论理性与德性自主。近代以后所塑造的道德哲学体系,也呈现出与传统极大的不同:一者,人类自身被视作独立的价值设置与审定的源泉;二者,对于德性的诠释及规定,开始明显倾向于社会制度层面:以社会契约作根基、以社会正义为目标。此二者之成立的前提,又都在于认为人类才真正具有理性思维能力:人类理性不仅可以为自身和万物立法,而且促成个体借由契约构建相应的社会规范及历史文化共同体。然而,借用康德的概念可知,近代所诉诸的理性虽然依旧包含实践理性和理论理性,也即一般意义上的道德实践与科学认知两部分。但是,科学知识往往会披着一幅无所不知或不能知的外衣,褫夺本该隶属于道德实践的话语权。换句话说,是知识理性(尤其是自然科学知识),而非道德理性,才真正铸就了近现代世界的生产与生活方式,并塑造了与之相应的观念上层建筑。

马克思从政治经济学的视域出发,将人之本质诉诸于社会性及有目的的实践两大维度。其中,他对于前者的经典表述是,“人的本质不是单个人所固有的抽象物,在其现实性上,它是一切社会关系的总和”^{[8](P501)};对于后者的典型论证则为,“最蹩脚的建筑师从一开始就比最灵巧的蜜蜂高明的地方,是他在用蜂蜡建筑蜂房以前,已经在自己的头脑中把它建成了”^{[9](P208)}。从根本上看,马克思所发现的政治运作与经济

发展的内在逻辑,依旧得益于近代知识理性的极大增进。因为,无论是社会化大生产的形成,还是实践及实践检验能力的实现,都只能是近代科学大爆发所带来的特定结果。

总而言之,从传统到近现代,人类对于自身本质的界定,经历了从德性到知识、从个体到社会、从修养到生产的转向。此一转向的动力,主要来源于近代自然科学及工业生产的突飞猛进,使得人得以借由知识理性与劳动生产,在满足自我欲望的同时,实现对外在世界的控制。

然而,近些年人工智能的飞速进化,却猛烈冲击着近代以来的人之本质的定义模式^[10]。这一点,从人工智能在棋类领域的三次革命性进化,都造成人类对自身知识理性超越地位的质疑,便可证明。1962年,被誉为“计算机学习之父”的阿瑟·塞缪尔(Arthur Lee Samuel, 1901-1990)在IBM公司开发的西洋跳棋程序,战胜了美国当时著名棋手,既让人感觉惊奇,也不免引发民众对于人类命运的忧虑。随着信息的传播和技术的深化,人工智能会下西洋跳棋的事实,变成习以为常,不再显得智能。1997年,IBM开发的深蓝程序战胜了当时国际象棋大师加里·卡斯帕罗夫,人工智能进化的速度及其展现的能力,再一次震惊世界,人类智慧的优越感一时间被冲得七零八落,智能机器威胁论又一次喧嚣尘上。然而,随着国际象棋及中国象棋成为计算机的日常应用之一,人们再一次习以为常。在民众看来,无论是跳棋还是象棋,在步骤上都是可以穷尽的,以致无法完全代表人类智能。然而,2016年谷歌开发的AlphaGo,先后击败一众世界顶级围棋棋手,并且在对弈过程中表现得游刃有余,使得人类棋类领域最后保留的一块圣土也被人工智能所攻陷,人类智能的危机也因此再一次被引爆。2022年11月30日,美国OpenAI发布了ChatGPT聊天机器人程序,其具有理解和学习人类的语言的能力,实现根据聊天的上下文进行互动,真正像人类一样交流。再次引发对人类秩序规则建构的讨论,如防止学术欺诈和剽窃、防止数据泄露、引发

隐私权争议等等。

当代的人工智能以不可想象的速度发展着思维的能力和限度,在信息存储、运算、传输等多个方面,都已经远远超出了人类思维的极限。而且,目前部分经过深度学习训练之后的人工智能,已经可以根据劳动目的的需要来制造和使用工具。按照既有趋势的发展,只要人类提供的大数据足够丰富,人工智能很可能在工具的制造和使用能力、劳动的效率及效果等方面超越人类。就人工智能既有发展格局来看,未来的10—15年,将有40%—50%的任务可以被人工智能化。^[11]若是期间技术研究得到进一步突破,此比例自然也将得到进一步提升。待到那时,人之理性思维,以及劳动实践,都将不能再成为自身的专有标志。

综上,如果仅从理性思维能力,乃至有目的地改造自然等维度来定义人之所以为人。那么,在某种程度上,人工智能或许将比人类自身更为符合、也展现得更为淋漓尽致。为此,重新正视并培育人之本质所内含的德性维度,自然也就变得理所当然。毕竟无论就人工智能目前发展程度,还是从科幻题材对于人工智能未来形态的想象来看以及人工智能的本质危害而言,人之德性的培养都始终是有意义的。

首先,就人工智能的现有技术而言,不仅无法完整、真实地再现人之情绪与道德能力,更是缺乏自我的欲望表达、意志自由及价值追寻等,而且在未来很长一段时间内也会依旧如此。^[12]在此背景下,人工智能所造成的系列伦理问题,从根本上讲自然也就是人自身的问题,即人工智能的设计与运用,乃至道德主体的问题,其实不过是人类社会既有问题的投射或延伸而已。换句话说,现有人工智能其实就像金钱一样,本身确实是中立的,问题的产生及后果,主要都源于设计者或操作者本身,即是人类社会既存的不文明、不公正、自由的不完全等,不可避免地造成了人工智能领域的相应问题。如果个体能发动其良知、良能,挺立一己之德性;社会能践行公正契约、保障分配正义、维系自由原

则等,那人工智能所造成的问题,自然也将在很大程度上随之消解。或者至少可以,更大程度回归到技术层面,而非伦理层面来进行分析与改进。

其次,假如科幻题材所想象或预言的“强人工智能”时代终将到来,也即智能机器有朝一日能实现意识和情感的觉醒。那么,突显人之本质的德性一面,并以此来推进个体之人与全体人类的道德进化,同样会是有百利而无一害的。人类社会奉行的道德原则与正义制度,概括而言不过“天下为公”“己所不欲,勿施于人”“礼尚往来”等等。反之,若是不懂得或者不遵循此道的人或物,自然被排除在同道之外,无需也无法实现有效交往、建立亲密情感。而且,就全体自然界所展现的一般规律来看,基本也遵循相仿的正义逻辑。因此,若是人之德性能够得到足够的培育和扩充,逐渐实现推己及人、推人及物的普遍关爱,那不仅是人与人之间,而且人与物之间,也能实现和谐共存。更重要的是,由德性挺立而通达的“兼爱”,本身就是依照平等主体的原则来进行对话和交往的。这也就意味着,压根不存在人工智能因意识觉醒而与人争权夺利的问题,因为其权与利早已作为事实而存在。

最后,人工智能作为一种人类发明的技术逻辑,终究是需要回归社会实践的。但如果完全按照技术逻辑的胜任规则,重构社会秩序,经由文艺复兴、启蒙运动、宗教改革以及工业革命所建立的人本主义价值观及社会规则体系将会遭受颠覆性的打击。即“如果有能力以最小成本的最简单方式去获得最大利益,人就会理性地选择简单粗暴的方式去解决问题,而不会选择复杂的高成本的方法……绝对强者的人工智能系统不需要苦苦地通过复杂而高成本的制度、伦理和法律去解决社会矛盾,而将会‘理性地’选择简单粗暴的解决方式。”^[13]这即是人工智能的本质危害。对于这一点,需要的正是人之德性的输入。成熟的文明及道德伦理体系的复杂性,往往体现在社会公平的搭建、财富的合

理分配以及社会阶层的有效平衡等方面,因而出现了亚当·斯密的《国富论》、罗尔斯的《正义论》、哈贝马斯的《交往行为理论》等社会价值理论。虽然如何进行人之德性的输入,还有待持续研究,但是“计算机系统和其他人工制品具有意向性——设计师的意向行为赋予它们的意向性。毫无疑问,我们的社会有能力将机器人和其他人工智能定义为道德行为者和承受者。”^[14]就此而言,人工智能产品的开发者与决策者在治理和监管机构方面保护人类(特别是儿童)权利和安全的重要性,以确保负责任的技术创新。例如,在K12人工智能教育产品的设计过程中,需要注重AI课程的先行实验策略及风险评估机制,“注重国家战略引领、智能素养导向和实证结果取向”^[15],以保障中小学学生公平、公正受教育权的实现。

三、人工智能之于人类美好生活向度的思辨

对于美好生活的追求,不仅是人类社会建构的目标,同时也是其发展的动力。在苏格拉底看来,“德性是值得为之而生活的最高的价值,通过道德的完善,人们将会得到最大的好处。正是由于这种联系,‘德性’同时也是‘好处’,而且是‘完好’。正如苏格拉底所说:‘再没有什么比每天谈论人的完善更为幸福的了……没有经过如此审视的生活是不值得过的生活。’”^[16]

与之相对,孔孟之道虽然极力推崇“仁”、“义”的价值与作用,但却也明确指出为政之道,并不仅仅是单纯的道德教化便能了事的,而必须实现物质与精神的双重发展,甚至需要先保障物质基础,再推行道德教化。《子路》篇记载孔子回答弟子关于治理的先后次序的对话:“子适卫,冉有仆。子曰:‘庶矣哉!’冉有曰:‘既庶矣。又何加焉?’曰:‘富之。’曰:‘既富矣,又何加焉?’曰:‘教之。’”^{[17](P143)}由此可见,道德教化固然重要,但却绝不能将之当做空中楼阁来搭建,而必须率先实现百姓的物质

生活层面的必要安顿。相比较之下,孟子在《梁惠王上》中更是直言,普罗大众的德性教育不同于儒生士大夫,必须先使之拥有足够的物质条件,才能使之追求品质的高洁,否则便只能是将其推于巧言令色、作奸犯科的境地,是对之变相的加害:“无恒产而有恒心者,惟士为能。若民,则无恒产,因无恒心。苟无恒心,放辟,邪侈,无不为已。及陷于罪,然后从而刑之,是罔民也。”^{[17](P211)}

事实上,如果仅仅是有德性,却缺乏一定物质基础的生活,必然不可能是完全意义上的“美好生活”。党的二十大报告中指出,“治国理政有常,利民为本。为民造福是立党为公、执政为民的本质要求。必须坚持在发展中保障和改善民生,鼓励共同奋斗创造美好生活,不断实现人民对美好生活的向往”。此中界定的“美好生活”,正是指向物质与精神的充分、平衡发展,是兼顾人之身体与灵魂、社会之效率与正义的综合性表达。有鉴于此,前文所述人工智能所带来的两大反思:即解构人之主宰地位与重构人之德性本质,也并非要人一味舍弃科学、理性的一面,而恰恰是要人在既有技术的基础上,更为适宜地安顿自身的行为。质言之,人工智能带来的冲击也好,反思也罢,绝非是要人抛开技术不用,也不是要完全对既有生活样态加以革命性颠覆,而是要人学会如何重新审视自身,学会与人工智能,乃至与整个外在于我的世界,实现友好共存。

为此,首先就必须正视和接纳人工智能为人类之秩序与安宁带来的系列福利。人工智能之于美好生活的首要价值,无疑首先是技术的、物质的:人工智能的信息集成、归类功能,无疑可以极大程度提升人类视域的广度与深度,推进决策前提的全面性及决策结果的科学性,需要进行相应的制度改良,如在人工智能司法方面,采取“从软硬法结合视角建构司法系统的运行标准与制度规则、从全过程视角强化主体之间的协同治理及通过指导性案例和司法解释赋权法官的司法解释空间,提高法律解释

技术”^[18]等。智能化的生活,使得人的身体和时间获得进一步解放,以致有了更多时间和更大空间去进行自我喜爱的活动;智能化的生产,则让物质供应有了更为高效的保障,使得资源贫乏等问题得到一定程度的缓和;智能化的医疗,不仅让原本残缺的身体有了重新健全的可能,而且可以提前预防潜在的疾病,延伸人之寿命……凡此种种,无不预示着,人工智能及其进化,可以让人类生活变得更加美好。

其次,就人工智能对公私划分、就业、资源分配等既有秩序造成的冲击来讲,与其将之归罪于技术本身,不如借此来反思人类社会既有制度以及人性取向既有缺陷。就目前人工智能的发展状况而言,在它出现及应用之后,人类社会出现的新问题几乎寥寥无几,更多都是原本就存在的。换句话说,中心与边缘的分化,国家与社会集体与个人的冲突,效率与安宁的张力等一系列伦理困境的发生,只不过借由人工智能这个放大镜而更加凸显而已。这就意味着,若想真诚地、根本地解决前述诸般伦理冲击与问题,最基础也最核心的环节,乃在于审查和完善既有的人性与制度。如此,才能为美好生活的早一步实现,真正提供相应的精神与制度支持,并更为根本地解决包括人工智能在内的现代科技带来的诸般困扰。

一方面,唯有反省现代社会人的价值追求,尽力避免原子化的生存孤岛及虚无化的生存态度,才能尽力突破电子时代“信息茧”的普遍束缚,推动人的自由全面发展。另一方面,唯有普遍意义上个体道德的挺立,即增进其对于他者的同情、对于世界的包容,才能随之让人工智能的设计程序及呈现结果,也同样规避种族歧视等不和谐现象。再一方面,唯有推进政府运作的法治化,塑造资源分配的正义性,以及强化社会慈善事业的发展,才能在守住公、私之间红线的同时,也防止社会阶层的过度分化。最后一方面,唯有个体之人乃至全体之人类,消除莫名的骄傲感,实现同情心、同理心的推己及人、推人及物,真诚信仰万千世界都有其不可

知,更不可侵犯的运作逻辑,才能将各种层面的平等对话加以落实。

参考文献:

- [1]赵汀阳. 人工智能“革命”的“近忧”和“远虑”——一种伦理学和存在论的分析[J]. 哲学动态, 2018(4): 5-12.
- [2]赵汀阳. 人工智能会是一个要命的问题吗?[J]. 开放时代, 2018(6): 49-54.
- [3]BRYSON J, WINFIELD A. Standardizing ethical design for artificial intelligence and autonomous systems[J]. Computer, 2017, 50(5): 116-119.
- [4]阿西洛马人工智能原则——马斯克、戴米斯·哈萨比斯等确认的23个原则,将使AI更安全和道德[J]. 智能机器人, 2017(1): 20-21.
- [5]陈鼓应. 庄子今注今译[M]. 北京: 中华书局, 2017.
- [6]朱熹. 四书章句集注·中庸章句[M]. 北京: 中华书局, 2010.
- [7]唐热风. 亚里士多德伦理学中的德性与实践智慧[J]. 哲学研究, 2005(5): 70-79.
- [8]马克思恩格斯文集: 第一卷[M]. 北京: 人民出版社, 2011.
- [9]马克思恩格斯文集: 第五卷[M]. 北京: 人民出版社, 2011.
- [10]孙伟平. 人工智能对“人”的挑战[N]. 光明日报, 2018-01-29(15).
- [11]李开复: 人工智能可能带来哪些伦理问题? [EB/OL]. (2018-01-27) [2022-10-05]. <https://finance.qq.com/a/20180127/003303.htm>.
- [12]甘绍平. 机器人怎么可能拥有权利[J]. 伦理学研究, 2017(3): 126-128.
- [13]赵汀阳. 人工智能提出了什么哲学问题?[J]. 文化纵横, 2020(1): 43-57.
- [14]JOHNSON G D. Computer systems: Moral entities but not moral agents [J]. Ethics and Information Technology, 2006(8): 195-205.
- [15]楼又嘉, 李一航, 王永固, 等. 人工智能基础教育课程全球图谱: 开发框架与实施方略——UNESCO《K-12人工智能课程图谱》报告解析[J]. 远程教育杂志, 2022(3): 3-15.
- [16]赵敦华. 孔子的“仁”和苏格拉底的“德性”[J]. 北京大学学报(哲学社会科学版), 2003(4): 88-95.
- [17]朱熹. 四书章句集注[M]. 北京: 中华书局, 2010.
- [18]周媛, 张晓君. 人工智能司法的可解释性困境及其纾解[J]. 财经法学, 2023(2): 3-20.

【责任编辑 刘红娟】

Reflections on the Purpose of Artificial Intelligence

DENG Daqi

Abstract: Artificial intelligence is a general double-edged sword. The basic task of philosophical research is to reflect on its technical problems and guide its positive effectiveness. Breaking through the patchy speculative path of traditional artificial intelligence “technical loopholes-institutional reinforcement”, taking artificial intelligence as a window to reflect the gains and losses of human beings and promote the realization of a better life, and reaffirming the important value of awe and virtue in human life, we can obtain a new philosophical dimension of AI in a better life. It is necessary to actively use intelligent technology to improve the quality of production and life, but also to improve the corresponding humanities and system construction to ultimately achieve a better life.

Keywords: artificial intelligence; human; reflection; virtue; better life

(上接第76页)

Divided Justice: System Evolution Path of Enterprise ESG

CHEN Hongjie

Abstract: From a systematic perspective, the concept of ESG reveals the evolutionary characteristics of both functional differentiation and structural coupling between social systems. Functional differentiation means that different social functional subsystems operate independently based on different values of communication. For example, the market system adheres to the value of “profit is justice”; the ecosystem follows the value communication of “green is justice”; the ethical system, on the other hand, is filled with the imagination of justice based on human scale. Correspondingly, structural coupling and merging do not negate the autonomous operation of social subsystems with some ultimate value logic, but selectively translate the external environmental information disturbance into the default communication code of the system. The essence of enterprise ESG is to continuously improve the upper limit of China’s ESG capabilities by selecting a differentiation mechanism for market leaders. At the same time, as the “immune system” of the entire society, the law can ensure the lower limit of the legitimacy of China’s ESG practices through corporate compliance mechanisms.

Keywords: corporate governance; functional differentiation; structural coupling; information disclosure; reputation mechanism