

人工智能是伦理道德建设的“助推器”

孙伟平

摘要：人工智能是异常复杂、深具革命性的高新科学技术，社会智能化是人类文明史上前所未有的社会伦理试验。人工智能的快速发展，智能时代的到来，对社会生产方式、生活方式乃至休闲娱乐方式造成了全面、系统的冲击，对包括伦理道德建设在内的精神文明建设也产生了巨大、深远的影响。从积极、乐观的建设视角看，以智能技术的发展和广泛应用为基础，伦理道德建设的物质基础更加夯实，道德手段和工具更加丰富多样，社会伦理道德建设和人自身的提升面临着难得的历史机遇。

关键词：人工智能；伦理道德建设；助推器

中国分类号：TP18 **文献标识码：**A **文章编号：**1673-5706(2019)04-0098-04

一、智能技术的发展、社会智能化的推进，正在不断夯实伦理道德建设的物质基础

智能技术的发展、社会智能化的推进与生产力的发展是相伴随的。社会的智能化程度已经成为一个国家、地区发展水平的标志。随着智能时代的到来，人工智能不断向社会生产方式、生活方式渗透，智能产业更是崛起为新的重要的经济增长点，并将极大地提升社会生产力水平。人工智能的广泛应用，促使产业结构不断调整，经济发展不断转型升级，劳动生产率空前提高，所提供的产品和服务日益丰富。在经济快速发展的背景下，社会积累的财富越来越多，人们的生活更加富裕，生活品质不断改善，社会治理水平也不断得以提升。经济、社会的快速发展，人们生活质量的提升，虽然不能与道德进步

直接划等号，但为伦理道德的提升、人与社会的自由全面发展奠定了更加坚实的物质基础。^①

根据唯物史观理论，生产力是社会发展中起决定性作用的因素，也是人与社会自由全面发展的决定性力量。物质文明建设与包括伦理道德在内的精神文明建设向来是相互联系、相互作用、相辅相成的。古人云：“仓廩实而知礼节，衣食足而知荣辱。”（《管子·牧民》）只有在一定的物质生产基础之上，例如，只有在解决了吃穿住行等基本需要之后，人们才可能产生更高层次的精神文化需求，才可能有时间和余力充实和发展自己，才可能逐步改变混沌、愚昧、迷信、落后的状况，在精神文化和伦理道德方面不断取得进步。否则，就如同马克思所说的：“就只会有贫穷、极端贫困的普遍化；而在极端贫困的

本文为教育部哲学社会科学重大课题攻关项目“人工智能的哲学思考研究”（18JZD013）阶段性成果。

^①孙伟平：《关于人工智能的价值反思》，《哲学研究》2017年第10期。

情况下，必须重新开始争取必需品的斗争，全部陈腐污浊的东西又要死灰复燃。”^①

此外，人工智能的应用不仅促进产业结构调整和经济转型升级，为新型伦理道德建设提供良好的前提和基础，而且可以直接应用于与伦理道德相关的领域，为伦理道德建设提供直接支持。例如，迈入智能时代，一个国家、地区往往更有条件增加财政投入，加大智能教育、培训的力度，创新思想政治教育方式，提高大众的文化素质、科技水平和伦理道德修养；能够提供更加丰富多样的精准化智能服务，促使社会运行更加安全高效，人们能够最大限度地享受高质量服务和便捷生活；有条件建立健全社会福利和社会保障体系，尽可能帮助和扶持贫困地区、弱势群体，促进社会公正与社会和谐；可以强化人工智能对自然灾害的实时监测，构建食品药品和公共安全智能化监测预警与控制体系，保障社会公众的生命财产安全；可以综合运用技术、经济和法律手段，对突破道德底线的行为（特别是智能犯罪行为）予以必要的惩处，维护基本的伦理、社会秩序等等。

总之，随着智能时代的到来，智能经济和智能文化快速发展，智能治理和智能服务不断改进，这前所未有的满足了人们不断增长的精神文化需求；特别是极大地拓展了人们的生存和活动空间，改善了人们的生活品质，可以帮助人们更好地发展自己，实现自身的潜能与价值，在儒家倡导的“成人”“成己”方面取得实质性进步。

二、智能技术的发展和應用可以帮助人类更加理性地进行道德评价和选择

与无机器辅助的人类相比较，人工智能正在表现出日益增强、超越人类的趋势。人工智能可以其“客观的立场”“冷静的态度”“丰富的知识”“敏捷的思维”“绝对的执行力”等帮助人们更好地把握时代发展趋势，细致地掌握事实情况，以此为基础更好地认识自我，不断提升道德评价、选择和决策水平。

人工智能的任务是理解自然智能的工作原理和

工作机制，开发具有类似人类智能的机器，为人类提供各种各样的产品和服务。迄今为止，人类智能是已知的最高级、最复杂的自然智能，因而人工智能的发展方向是模拟人类智能。当然，这种模拟不是刻板地模拟人脑的物质结构，也不是简单地模仿人脑的工作机理，而是一种启发式、创造性的模拟，是模拟支撑人类思维的信息转换和智能创造原理。

基于学习算法的人工智能系统可以系统地掌握一个人的生活经历和处世原则，帮助自己在面临一定道德情形时，准确地分析自己的道德需要，“发现”自己的道德动机，明确自己的道德权利、责任和义务。假设智能助手知道某个人坚持儒家的“己所不欲，勿施于人”之类的“道德金规则”，那么，在处理相应的人际关系时，它便会将这一伦理原则和当下的具体情形联系起来，提示“主人”应该怎么做。例如，一个人不赞成自己的孩子脾气暴躁，对自己不讲礼貌，如果某一天自己因为某种原因，也对年迈的父母过于性急，态度恶劣，这时，智能助手便可以适时地提醒：这样做违背了自己信奉的“己所不欲，勿施于人”的伦理原则，是一种不理智、不恰当、不文明的行为。

相对人类的某一方面、甚至相对一定的个人整体来说，一定的智能系统可能拥有的立场更加客观，视野更加开阔，知识更加丰富，了解的事实更加完整，后果预测更加可靠，对道德规范的适用更加到位，并更少受到各种社会关系、“人情世故”的干扰，犯各种人为的低级错误，据此可以理性地帮助人们进行道德评价、选择和决策。事实上，对于日常生活中的一些道德难题，包括债务纠纷的裁定、交通事故中责任的划分、法律案件的具体量刑等以及不同道德选择的后果预测，特定的智能系统正在显现出一定的优势。它们虽然尚嫌稚嫩，却已经成为人们进行道德评价、选择、决策的参谋和助手。我们完全可以预测，今后的智能系统拥有的数据将更多更好，将会更加“聪明”“敏捷”，能够更好地理解伦理原则和道德规范，更有效地帮助人们进行道德评价、选择和决策。

^①《马克思恩格斯选集（第1卷）》，北京：人民出版社，1995年，第86页。

三、智能技术的发展和广泛应用，有利于人们的道德自律与人自身的道德修炼

一般而言，道德依靠两种方式加以维系，一是自律，二是他律。人工智能为道德自律和道德他律提供了新的实践基础、实践方式。

道德自律以人们的自我认知和道德评价为基础，是人们自我约束、自我提升的一种价值追求。智能系统的视野更加开阔，拥有的知识更加系统，可以帮助人们深刻洞察所处的社会环境，准确把握面临的事实情况，从而依照习得的伦理原则和道德规范指导自己的行为，并不断在生活实践中修正、完善自己。

如果人们有心向善，希望在伦理道德方面自我修炼，自我提升，那么，可以借助智能秘书、学习助手、智能保姆、智能管家之类贴心的智能工具，及时提醒自己抑制不合理的欲望，督促自己切实实践履道德规范。例如，通过程序设定，提醒和督促自己“应该去看望父母了”、“应该信守诺言按时赴约”、“应该依据协议偿还债务”、“应该完成自己的本职工作”等。这可以帮助自己时时自警，严格要求自己，锻铸自己的道德人格，做一个有德之人。对于一些自己不满意、却难以自制的不良陋习，例如办事拖拖拉拉、网络游戏成瘾、网络购物上瘾等，还可以设计相应的智能程序，在超过预设的阈值后，帮助自己“强制执行”。例如，在每天或每周的网络游戏达到设定的时间阈值后，设定的程序可以黑屏甚至关机，从而强行中断游戏。

人类正在与越来越多的智能机器“生活”在一起。有了各种“知规知礼”、“铁面无私”的智能小伙伴，人们就将心中的“大法官”外在化了，时时警醒自己，修身养性，止恶扬善，令道德自律水平跨上一个新的台阶。

四、智能技术的发展提供了新型的道德他律方式，有助于重建新型的伦理秩序

道德他律主要依靠社会舆论以及相关机构的教育、管理加以维系。智能时代为道德他律提供了更加透明的社会环境，也提供了越来越多的道德工具。

因为整个社会日益数字化、网络化、智能化，数据采集、存储、处理、传输能力空前强大，人们的一举一动几乎都处在“聚光灯”之下，无所遁形。

这种空前“透明”的社会环境，这种方便追索的道德执法能力，可以形成强大的道德舆论压力，敦促人们主动抑制不良动机，自觉规范自己的言行。

在一定的道德情形中，如果某个人、某个组织的行为有违既定的伦理原则和道德规范，借助人工智能，社会公众和相关机构可以及时查询当事人的背景信息，准确还原事情的经过，从而以事实为依据，形成强大的社会舆论压力，迫使行为人弃恶扬善，维护正常的伦理秩序。

道德失守底线是当今世界、特别是社会转型时期的中国面临的突出问题。迈入智能时代，对于那些肆意践踏道德底线、甚至触犯法律的行为，例如故意伤害、见死不救或“救命索要救命钱”、被人施救却恩将仇报以及公务员索贿行贿等恶行，可以借助各种智能系统，包括各种智能监测系统、大数据分析系统，方便快捷地弄清事实真相，并对当事人采取系统性的、强有力的他律手段。例如，除了通过各种媒介曝光当事人，进行大范围、强有力的道德谴责外，还可以对其采取包括技术、经济、法律制裁等在内的惩处措施，切实让作恶者成为“反面教材”，既没有“面子”，又无利可图，从而将“善有善报，恶有恶报”落到实处。

可见，迈入新颖、独特的智能社会，面对现代社会“信仰缺失”、价值多元、“道德底线”屡屡失守的窘境，人工智能可能成为道德教育、管理的“好帮手”，成为维护正常的社会秩序的“道德利器”。或许，囿于社会的智能化水平，我们今天从中受益还不够多，但人工智能的发展一日千里，新颖别致的智能教育、管理方式正在涌现，更加合理、更加公正的伦理秩序有望形成。

五、人与智能机器之间的道德竞争，特别是道德教育、管理权之争，迫使人类不断加快“进化”的步伐，打造新型的人机关系和“人机文明”

伦理道德领域一直存在着道德进步论与道德退步论的争论。今天，即使从道德进步论的视角看，伦理道德的发展也难以跟上人工智能发展的速度。在人工智能的研发和应用中，出现了越来越多的既有道德从未涉及的问题，包括智能无人系统的道德主体地位问题、人的隐私权受到侵害的问题、数字鸿沟和社会分化导致的社会不公问题，等等，在一

一定程度上出现了“道德失范”和“伦理失序”。如何立足我们的文化传统、应对问题和挑战，重塑智能时代的道德教育、管理方式，是一个亟待解决的难题。

随着智能时代的来临，智能系统越来越多地渗入了人们的组织管理过程，教育、管理方式正在迎来新的变革。智能系统能够记忆（存储）大量的政策和法规文件，自动处理管理工作中的行政协调和控制任务，承担管理工作中的程序性任务，减少管理工作中的人为失误，从而节省管理成本，提高管理效率。不过，运用各种智能系统所进行的道德教育和管理，有时可能忽视被管理者的文化传统和心理特征，甚至忽视被管理者是一个具体的活生生的“人”，从而缺乏人类特有的同情心，缺少人们特别看重的“人情味”。

具有冲击力的是，随着人工智能的突破性发展，一个问题突显出来：在自然人与智能机器人之间，谁的道德表现更为优异、更令人信服？谁能占据道义制高点、掌握话语权？谁更有资格拥有道德裁判权，以及相应的教育、管理权力？虽然人类号称“万物之灵”，过去在伦理道德领域一直牢牢地掌控着话语权，但发展到智能时代，面对这些本来有确定答案的问题，人类已经变得不那么自信，不那么有把握了。

更值得警惕的是，随着今后可能出现的超级智能及其发展，它是否可能依仗自身的体力和脑力优势，以及远超人类的活动效率，强行抢夺道德评价、决策的“话语权”，以及道德教育、管理的“资格”，甚至自以为是地对创造它的人类进行道德训诫和道德管理，将人类强行纳入智能系统的道德范畴？若果真如此，那么人类将在一定程度上失去道德尊严和主动权。这也将是有史以来翻天覆地的“伦理大变局”。

在日益严峻的形势面前，面对智能系统、智能机器人可能带来的道德挑战，面对各种智能机器为人类所树立的更好的参照物，人类作为“万物之灵”，必须未雨绸缪，采取有效的行动，切实维护人类的道德

地位和主动权。具体地，我们必须立足智能时代，认真反思人类在历史上的道德表现以及人自身的缺陷和局限性，主动运用各种智能技术加强人自身的修炼和“监管”，以一种建设性的姿态投入与智能机器系统的道德竞争，从而加快人自身的进化历程，令人类在道德完善方面取得实质性进步，并据此打造更加先进、更加合理的人机关系和“人机文明”。^①

参考文献：

- [1] 孙伟平. 关于人工智能的价值反思 [J]. 哲学研究, 2017, (10).
- [2] 马克思恩格斯选集(第1卷) [M]. 北京: 人民出版社, 1995.
- [3] [美] 皮埃罗·斯加鲁菲. 人类 2.0: 在硅谷探索科技未来 [M]. 牛金霞, 闫景立, 译. 北京: 中信出版社, 2017.

作者：孙伟平，上海大学社会科学学部特聘教授、博导

责任编辑：周修琦

^① [美] 皮埃罗·斯加鲁菲：《人类 2.0：在硅谷探索科技未来》，牛金霞、闫景立译，北京：中信出版社，2017 年，第 375 页。