大变局时代的伦理应对

当今世界，新一轮科技革命和产业变革加速演进，人工智能、生命科学等前沿技术不断突破，深刻重塑着生产生活方式与社会治理结构；与此同时，逆全球化暗流涌动，地区冲突与文明摩擦频发，传统安全与非传统安全威胁叠加共振，人类正经历百年未有之大变局。在这一复杂的时代背景下，伦理道德不再是书斋里的抽象命题，而是已成为关涉技术向善、社会和谐乃至人类文明存续发展的关键变量。如何立足中国实践、回应时代之问，在技术狂飙与文明碰撞中探索伦理应对的中国方案，为不确定的未来提供确定性的价值锚点，已成为学界必须直面的重大问题。在此背景下，近日，江苏省道德发展智库、江苏省公民道德与社会风尚协同创新中心、东南大学中华伦理与未来文明研究院、东南大学人文学院等单位共同举办“伦理与文明”国际会议。《探索与争鸣》编辑部在会议基础上，聚焦“大变局时代的伦理应对”，组织本次圆桌笔谈，邀请部分与会学者以及相关领域专家围绕技术伦理、主体责任、全球秩序与文明理解等议题展开探讨，共同解答时代的伦理难题。

万俊人教授指出，当代社会面临着“形而上学堆积”的道德疑难与类似“辛顿问题”的技术伦理挑战，哲学伦理学必须超越单纯的解释功能，以“天职”意识承担起改变世界的实践责任，重建一种足以承担时代使命的新型哲学和科技伦理。**邓安庆**教授批判了传统元伦理学在人工智能冲击下的局限性，分析了以道义实存论作为当代元伦理学的可能性，提出以道义实存论重构在场形而上学，让伦理学重返生活世界，为时代的伦理危机提供具有思想性的智慧的解决方案。**王淑芹**教授指出，确保人工智能向善发展的关键，在于将伦理要求内化为相关行动主体的美德。科技研发者、政策制定者与伦理评审者的卓越品格，是填补法律滞后空间、实现价值对齐不可或缺的“内控力”与“压舱石”。**庞俊来**教授提出，要超越文明冲突，需建立一种新的“伦理理解”模式。这种理解不依赖于西方语言中心主义的范式，而是基于“命运的通感”“知识的同构”与“价值的共享”，旨在为不同文明间的深度对话与“伦理文明”的构建奠基。**曲红梅**教授基于马克思主义世界历史理论，探讨了全球交往中个体成员的规范性立场。她试图超越爱国主义与世界主义的抽象对立，为“人类命运共同体”这一理念提供坚实的哲学基础。**蒋艳艳**副教授聚焦高度不确定时代的公共舆论场域，构建了一个包含道德认知、道德态度与道德规范因子的分析模型，并提出以对话的道德影响力为旨趣来探寻道德舆论的应然模式。

——主持人　樊　浩　张　蕾

大变局时代的哲学伦理学响应

万俊人，清华大学文科资深教授，国家文科一级教授

百五十多年前，擅长国际政务的清廷重臣李鸿章在同治十一年（1872 年）上呈奏章，请求朝廷关注“三千余年一大变局”，亦即由“欧洲诸国”创始并迅速展开的现代世界商业拓展之崭新局面。这大概是国人“睁眼看世界”的最初一瞥，自然难免有些瞠目结舌，继而惊愕乃至惊慌失措。于今，“百年未有之大变局”的警示再一次响彻耳边，与曾经的耳闻有所不同，这一警示已然不再只是针对国际商业或政治经济学意义上的“大变局”，甚至也不只是一个时事政治的“大变局”命题，还是一个复杂而深远的哲学命题，其关乎历史哲学、政治哲学和伦理学的时代应对与未来走向。

历史地看，哲学原发于人类空间意识的觉醒，却成长于人类时间意识的流变进程。当古希腊人“惊异”于所处其间的宇宙及其无穷变化时，哲学作为人类追问不已的爱智方式便开始其智识探究历程，从对外在宇宙的“天问”，到对人自身的自我认知，哲学从宇宙论（cosmology）转向“认识自己”，进而聚焦美德伦理学、政治哲学、逻辑学、修辞学乃至形而上学等，晚期智者和苏格拉底、柏拉图、亚里士多德被视为这一哲学转向的代表。在中国，先秦诸子同样表达了类似的哲学意识转向，从老子的“问道”到孔子的“事人”，[①] 清晰地反映了这一转向，而稍晚一些的伟大诗人屈原更是以《天问》《九歌》等诗篇，典型地表达了这种由外而内、由天而人的哲学意识转换。中西哲学随后发展路径的歧出在于：同处“轴心时代”即公元前 6 世纪—公元前 5 世纪前后两百多年时间，中国先秦诸子百家历经百家争鸣与战国争雄，由道转儒，中经援佛济儒，逐渐形成近两千年的儒家主导之学术和政治文化面貌；而西方则中经千年神政合一的教父哲学和经院哲学之漫长沉积，迎来“文艺复兴”人文主义思潮的爆发，迎来欧洲由神到人的近代人道主义启蒙运动。这是一场人类历史上前所未有的从传统转进现代的历史大变局，如果说哲学充当了这一伟大变局的理论旗帜，那么近代科学技术发明则是撬动这一伟大变局的第一杠杆。自此，以“惊异”的姿态探寻宇宙奥秘的哲学确乎找到了现代哲学反思得以展开的地平线（horizon）：近代航海家历经一次次远航探险



而最终实现的“地理大发现”，终于勾勒出现代性“世界”的完整图像和边际，尽管这一图像首先凸显为国际贸易的世界大市场，却也为哲学建构现代知识论意义上的“世界观”提供了空间具象化的蓝图。与此同时，一种空前而迅疾的近代科技发明及其所驱动的世界现代化进程不仅强健有力，而且呈现出几何级别的加速狂飙突进态势。这一突进态势不仅惊醒且极大地强化了哲学家们的时间意识，而且催生了现代进步主义的历史哲学取向。这种现代性哲学的“时间意识”甚至很快盖过空间意识，成为其标志性历史主义主导型反思方式（以黑格尔、马克思为近代典型）。更重要的是，现代性哲学的时空意识逐渐促动其本体论的再一次转折：中古时期由外到内或由物到人——亦即由天地宇宙到人自身——的哲学转变面临着人与事或主体与客体的两分化纠结。这一纠结同时在现代性时间与空间的双重叙事向度上展开。进至 20 世纪，胡塞尔发起颠覆性的现象学哲学运动，提

① 参见《论语·先进》：“未能事人，焉能事鬼？”

出“回到事物本身”的哲学命题，“搁置”（“悬搁”）一切历史的和既定的，从生活世界的事实重新开始哲学的反思。海德格尔选择了以时间意识反思存在与此在的存在主义哲学路径，来回应其哲学导师的世纪性哲学之问。萨特及同时代的法国存在主义哲人和文学家则选择了以生活经验的空间意识反思，来回应胡塞尔现象学的世纪性哲学之问。二者的相交点是，通过反本质主义的形而上学颠覆，反思近代以降的现代性哲学，重建现时代的“人–事”哲学。与此平行的另一路哲学转折同样具有划时代的意义，即由20世纪逻辑与经验实证主义和科学主义思潮所驱动的“分析哲学”暨“科学哲学”运动，它创造了现代哲学发展史上长达百年且至今方兴未艾的“分析[哲学]的时代”。

现代世界与时俱进，哲学同样如此。20世纪将临前夜，身处“世纪转折点”的德国哲学狂人尼采高呼“上帝已死，一切皆有可能”。这是旧世纪哲学的暮鼓，也是新世纪哲学的晨钟。同样来自德国的历史哲学家雅斯贝斯在《时代的精神状况》一书中提醒人们注意，我们的世界又一次发生“变形”：“世界正经历一场极大的变化，以往几千年中的任何巨大变化都无法与之相比。我们时代的精神状况包含着巨大的危险，也包含着巨大的可能性。如果我们不能胜任我们所面临的任务，那么，这种精神状况就预示着人类的失败。”[①] 雅斯贝斯如此断论的根据是：“我的本质即是作为整体的历史时代和社会状况。”[②] 也就是说，“我之所是即时代之所是。而时代之所是则展开其自身为那必然过程中的一个特定阶段”。[③] 人及其命运一如流水，在历史的长河里与时俱进地流动。不过，在雅斯贝斯的历史哲学视野里，人仍然是时间或历史演进中的主体，确证和捍卫人的主体性几乎成为整个现代性哲学的全部目的和使命。

然而，当时间来到又一个世纪之交，跨入21世纪的人类似乎正在亲手制造且不得不独自面对一种迥然不同的新境况：雅斯贝斯所说的技术大网内的世界开始演进为由诸如OpenAI和DeepSeek、超级互联网、超大数据、前沿生物（命）医学等新的超高科技所编制的智能化、数据化或数智化甚至虚拟化的超高技术世界，前沿科技开始占据时间之流的主航道，也在逐渐取代语言之于人类文化的根基地位，构成当下乃至未来人类文明暨文化的基本符号。在这个以“高精尖”为符号表征的智能化世界里，人或许不再是唯一的认知主体和行动主体，越来越多的可能是，人工智能机器人等逐渐替代自然人和社会人的活动，成为全新的认知者和行动代理者（agent），仿佛成为主宰世界的另类主体。与此同时，启蒙时代以来历经几百年建构并勉力确立的“现代世界体系”（the modern system，沃勒斯坦语），乃至整个现代生活世界的既定秩序，也因为近半个世纪的经济全球化与逆全球化/反全球化之反复对冲而加剧着我们时代的高度不确定性。

德国伦理学家汉斯·约纳斯（Hans Jonas）曾将20世纪的哲学伦理学困境概述为“形而上学堆积”（metaphysical accumulation），并严肃指出，所有传统伦理学都只考虑“非累积的行为”（noncumulative action），而“世界技术的累积变化不断超越了每一个其有所贡献的行为的条件，并以前所未有的情形运行，对此经验学说是无能为力的”。[④]因之，当代伦理学必须转变其理论方式，“超越”其“行为学说”（the doctrine of action）的古典样式而转入“存在学说”（the doctrine of being）的当代样式。即便近代伦理学以诸如普遍理性和主体（人）目的为理论武器，对传统伦理学进行了革命性价值重构，也依然囿于一种现代“智人”或“现代人”（homo sapiens）观念而“安居”于人类自我中心主义的主体性哲学世界，对现代“技术人”（homo faber）日趋强大以至于成为现代世界的宰制性力量这一趋势缺乏足够的意识。[⑤]时至今日，我们可以接着约纳斯的这一论断再进一步：今日世界，诸如OpenAI和DeepSeek之类的超级智能化技术的超

①②③ 卡尔·雅斯贝斯：《时代的精神状况》，王德峰译，上海：上海译文出版社，1997年，第19页，第23页，第23—24页。附注：1933年出版的该书英译本的书名为《现时代中的人》（*Man in Modern Age*）——引者注。

④⑤ 汉斯·约纳斯：《责任原理：现代技术文明伦理学的尝试》，方秋明译，香港：世纪出版有限公司，2013年，第13—14页，第15—16页。

速应用，以及作为当代新型“技术人”的智能机器人的广泛介入（massive engagement），不仅已然超出了约纳斯对现代“技术人”和“形而上学（道德疑难）堆积”的预期，也超出了当代人工智能科技界的领袖们（如诺贝尔物理学奖获得者杰弗里·辛顿）自身的想象，以至于辛顿本人近期越来越不安地呼吁，要谨慎开发、利用、规范、控制人工智能等超高新科技，确保科技社会和人类未来发展的安全。

人工智能技术的发展迅疾而猛烈，于是一些著名的经济学专家和管理者不时发出“人工智能经济泡沫”的预警。即便如此，这一超级智能科技的发展趋势已成脱缰之马，正在一望无垠的原野上狂奔，几乎已成无法改变的当代大变局之“暴风眼”。从哲学伦理学的视角来看，现时代最为迫切的任务确乎如同约纳斯所言，要抓紧“寻求一种技术时代的伦理学”。关于现时代，虽有各种不尽相同的标识，但不可否认这是“一个前所未有的大变局时代”。其前所未有的特性集中体现在它的高度不确定性上，而这种不确定性绝对不只是因为当代国际政治秩序的紊乱、经济全球化与逆（反）全球化的张力，以及“文化多元论”等一般文明特征所构成的“文明冲突”（亨廷顿语）和整体的心灵失序（沃格林语），更突出表现在由 OpenAI 等所代表的人工智能、几乎无限的信息网络和海量数据、各种颠覆性生命科技（如 AI 医疗、基因编辑、干细胞移植）等当代前沿超高科技所导致的超规范或反规范风险。譬如，智能机器人的“自主”越界、OpenAI 运作中已经出现的“逃逸”或拒绝执行人为操作指令等反常现象，已然使人类所预期的技术、价值、管理诸方面“对齐”（alignment）设想（设计）和努力，都陷入高度不确定的或然性状态。由此导致的社会风险已远远超出了德国社会学家乌尔里希·贝克（Ulrich Beck）所描述的“风险社会”（risk society）[①] 范畴。它不只限于现代工业和商业的快速发展所造成的社会失序、生态破坏、文化多样性萎缩等“外部风险”或“局部风险”，而是直接威胁到人类自身主体性，颠覆作为现代性之价值根本的“目的–手段”或“主体–客体”的两分范式，甚或导致人类文明摆脱人类自身操控之“终极结局”（the endgame）的总体性、根本性风险。

毋庸讳言，这是我们时代的哲学之难，更是我们时代的技术伦理之难。尤其严峻的是，对早已深陷“形而上学（道德疑难）堆积”的现代伦理学来说，既有的许多道德疑难（moral problematic）尚未消解，新的道德疑难又接踵而至，让人顿生几分“屋漏更遭连夜雨，船迟又遇打头风”（冯梦龙《醒世恒言》）的悲怆。由现代市场经济生发的“斯密问题”，亦即“经济人”与“道德人”的人格两分性张力问题；从逻辑经验主义认识论视角提出的“休谟问题”，亦即如何从“是然”（to be）中合乎逻辑地推出“应然”（ought to be）的问题；尼采在世纪的转折点上提出的“尼采问题”，亦即“上帝死后，我们（人）怎么办？”的问题；罗尔斯在《正义论》《政治自由主义》《万民法》等著作中提出的“罗尔斯问题”，亦即如何通过建构制度正义而实现现代民主社会的普遍公平正义，同时又在文化多元的条件下形成现代民主社会的“重叠共识”或“公共理性”问题；直到当下人们越来越关注的“辛顿问题”，亦即如何合理有效地开发、利用乃至规范化调控智能机器人等超高科技所产生的科技安全问题——这些问题逐一摆在当代哲学伦理学的面前。如果说此前诸多现代性道德难题已然构成了难以消解的现代性伦理学难题或“形而上学堆积”，那么相较而言，它们几乎都不及当下的“辛顿问题”来得深刻与棘手。因为此前的几乎所有道德难题终究仍然在人类理性的掌握之中，人类依旧具有解释、证成乃至一定程度上加以解决的主体能动性。可是，面对“辛顿问题”或者更广义的当代超高科技伦理难题，人类非但缺少这种哲学伦理学的主体能动性，而且很可能面临难以提供合理充分解释，无法提供哪怕是基本必要的证成，亦无法自主解决甚至完全超脱人类理性规范和有效调控范围的根本性伦理困境。同此前遭遇的所有伦理困境相比，当下

① 参见乌尔里希·贝克：《风险社会：新的现代性之路》，张文杰、何博闻译，南京：译林出版社，2018 年。

直面的伦理困境具有某种不可消解的绝对性和根本意义上的颠覆性。曾几何时，谈到“休谟问题”，现代语言逻辑哲学大师维特根斯坦也只能有些无奈地说：“伦理学谈论的在任何意义上都对我们的知识无所补益。但它是人类思想中一种倾向的纪实，对此，我个人不得不对它深表敬重，而且，说什么我也不会对它妄加奚落。”① 这是一种基于语言哲学和逻辑证明的优雅的无奈，但毕竟还保留了对伦理学的学术敬重，且后期维特根斯坦通过语用学的哲学转向，为伦理学的知识合法性留下了某种解释和解决的空间。面对“辛顿问题”，我们的无奈不仅无法再保持哪怕是一丝优雅，更看不到解决难题的任何余地。

然而，无论我们面对的时代难题多么艰难，无论我们时代的道德疑难及其“形而上学堆积”多么严重和深厚，我们都无法逃脱，只能勇敢面对，因为所有的道德问题都是我们自己造成的，都必须由我们自己担待和解决。我们是全部问题的始作俑者，因之也注定我们必须承担解释和解决所有这些问题的全部责任，哪怕这种责任是绝对的“难以承受之重”。可是，比直面这些道德难题——尤其是“辛顿问题”——的勇气更为重要和关键的，是如何找到解释并化解这些道德难题的路径。对此，国内外科技伦理学界乃至整个哲学界一直在探索各种可能的解释暨解决方案。据我有限的了解，国际上目前较有代表性的科技伦理（学）解决方案是，以荷兰科技伦理学家维贝克为代表的“将技术道德化”（moralizing technology）的内在主体论路径；国内较引人注目的哲学解决方案是由赵汀阳提出的“动词哲学”② 或基于动词本体的形而上学路径。前者集中于重建当代科技伦理；后者着眼于重建哲学形而上学；二者的共同点在于，改变传统伦理学乃至整个哲学长期甘当“消防队 / 救火员”或“黄昏猫头鹰”等被动角色的消极理论姿态，力求积极主动地解决新时代新科技所可能产生的道德疑难。

维贝克在《将技术道德化》一书中开宗明义地指出：“借助于现有的伦理学框架是不能永远充分地解决新技术带来的问题的，因为新技术以崭新的方式影响着人类的存在。”③既有伦理学的局限和尴尬在于，它习惯于从外部来思考或批判现代科学技术及其对人类社会的消极影响，而真正有可能解释或解决当代新技术之道德难题的方式应该是“发展一种‘从内部’的伦理学：使得技术的道德意蕴变成技术发展的一个显性要素”。④人类似乎一直习惯于从“目的–手段”的价值关系来理解人与技术的关系本质，把技术当成一种被使用或最多是“好用”的“人工物”。借助现象学，尤其是海德格尔的技术追问方式，维贝克提出并努力证成一种积极主动的技术伦理学，它将技术理解为人类的“行动调节”，具有其内在的人化意向性，因而，“人–技术关系的本质源自行动调节”，⑤而非目的与手段或用与被用的单向度关系。因此，新的技术伦理首先必须突破“人本主义”或“人类中心主义”的现代哲学前提，重新认识并肯定技术本身的道德性。譬如，孕期医检之于婴儿的健康或避免先天残障，智能驾驶技术对于防止因人工驾驶中的盲察而造成误伤行人，等等。这当然并不意味着技术本身可以充当道德行动者或道德主体，而只是说，技术作为人类发明并不断改进增强的“行动调节”和技术运用本身内含道德意向性，因而同样具有作为道德行动者或道德主体的特性，人工智能机器人或可为最新例证。所以，我们对伦理（人性）与技术之关联的理解，不应该简单以技术为敌，只想着“从技术中拯救人性，而是让技术和人性以合意的方式协调在一起”。⑥需要批判的不是技术本身，而是对技术的误用或滥用，一如福柯所指出的那样，人们应该憎恶和批判的并非权力本身，而是权力的宰制和专断性滥用。维贝克还批判地具体分析了传统伦理学的结果论（目的论）和道义论，认为它们都无法适应新技术伦理学的需要。相反，亚里士多德奠立的以人的“美好生活”为根本目标的美德伦理学，倒是有可能为当代新技术伦理提供宝贵的思想资

① 路德维希·维特根斯坦：《伦理学演讲》（1929年），载万俊人主编、唐文明副主编：《20世纪西方伦理学经典——伦理学基础：原理与论理》（上），北京：北京师范大学出版社，2021年，第146页。

② 详见赵汀阳：《寻找动词的形而上学》，北京：生活·读书·新知三联书店，2013年。以下凡引述或转述赵汀阳的有关观点，均来自此书，不再详细注释。另参见其近期新刊论文或学术讲座——引者注。

③④⑤⑥ 详见彼得·保罗·维贝克：《将技术道德化：理解与设计物的道德》（2011年），闫宏秀、杨庆峰译，上海：上海交通大学出版社，2016年，“序”第1页，“序”第2页，第13页，第106页。

源和理论通道。

相较于维贝克的“后人本主义”技术伦理学，赵汀阳的“动词哲学”显然抱有更大的哲学“野心”：他所寻求的远不止于科技伦理的更新，而是直指整个哲学形而上学的基础重建，即试图颠覆他判定为始终以名词为本体的传统形而上学，重建一种以动词为本体的形而上学，最终以“我作故我在”替代作为现代哲学之标识的笛卡尔式的“我思故我在”。在赵汀阳看来，相较于“名词哲学”，“动词哲学”可以为当代科技伦理提供更具解释力也更具前瞻性的哲学方法和形而上学基座，至少，“动词哲学”所内含的开放创造性潜能，更能满足新型科技伦理应对高度不确定性的未来时所要求的可能性预期和开放性回应。

对于维贝克等人“将技术道德化”的新科技伦理学范式，我虽抱有基本的同情理解甚或原则认同——我赞同重新认识现代科学技术本身的内在合目的性或道德性，建构一种广义的内在目的论或美德道义论的科技伦理学，以替代以往简单外在化的科技伦理学批判范式，进而使胡塞尔–海德格尔所开启的现代性科学批判获得更为积极的建设性意义——然而，我依旧怀疑，基于“将技术道德化”的“行为调节”概念是否足以完全解除作为道德主体的人与作为人工智能产物的科技（事）——哪怕是超高智能化、数智化的超高科技（事）——之间的主客二分张力？缓解无论在多大的程度上都不等于消解。而且，即便通过扩充“美好生活”这一美德伦理的最高目的，可以在较高程度上赋予科技本身以内在目的性价值，仍有可能“收之桑榆”而“失之东隅”，即“将技术道德化”可能在赋予科技伦理以某种内在目的论价值维度的同时，又失去（至少会降低）其既有的道义力量和规范调控作用。因此可以说，建构当代科技伦理学仍然是一项开放性的应用伦理工作。对于赵汀阳的“动词哲学”及其科技伦理的应用前景，我也多少怀有类似担忧：如果哲学一如罗素和杜威等现代大哲所言，是对“确定性的寻求”，或者说，正是由于人类及其生活于其间的世界充满高度不确定性，使得哲学成为人类寻求知识和智慧的基本方式具有一种内在的必然，那么，全然否定或颠覆传统的“名词哲学”是否可信复可行？无论是在语言学意义上还是在语言哲学意义上，无名词则无概念，无概念又何谈理论？更不用说在无名词或无概念的前提下谈论形而上学问题了。确乎，动词远比名词更能充分表达人类自身的创造性实践和主体性行动的哲学意味，更切近“当事哲学”的品性，但名词也可以承担同样的哲学工作，在描述和评价的双重意义上均如此。易言之，哲学的运思和论理其实既需要动词化，也需要名词式，因为它不单需要描述、分析和论证，也需要概括、概念、范畴、命题、规范、原则和体系化或理论（逻辑程式）化，即“动词哲学”这一说辞本身也是一种概念“名”称。想当初，亚里士多德将“形而上学”置于“物理学”（自然学）之后，并将之界定为研究“作为存在的存在”（Being qua being）之学问时，给形而上学选择的语言学词性“身份”（status）既非“是之为是”（be as be）的系动词，也非将系动词“是”简单名词化，而是二者兼顾，融合为动名词（being）。我的理解是，倘若非要用语词特性来表述哲学形而上学的“论理”身份，或可说，一种基于动名词的形而上学可能为新的或未来科技伦理学提供更为开放、更具有解释力和公共说服力的哲学证成，动名词形而上学的自我证成也许更为充分合理。

总而言之，不管如何命名，不论多少可能，重建一种足以承担时代使命的新型哲学和科技伦理学的关键时刻已然来临，仿佛文艺复兴时刻、启蒙运动时刻、尼采的时刻、胡塞尔–海德格尔时刻的“灵光再现”。这是我们的“天职”（calling），在一种诗、史、思的融合意义上，亦可将“calling”转译为来自我们生活世界的现时代“呼唤”。时代的呼唤即是哲学伦理学（家）的天命。不是么？！[①]

① 这篇简笔式的笔谈小文是我于2025年10月12日在中国人民大学举办的“第五届国杰论坛”上的发言和同年11月1日在华东师范大学组织的“中国自主知识体系建设论坛暨《华东师范大学学报（哲学社会科学版）》创刊七十周年学术研讨会”上的发言综合整理而成。感谢上述两个论坛的盛情邀请。

AI 时代急需伦理学研究范式转型：道义实存论之尝试

邓安庆，复旦大学哲学学院教授

当代世界的剧变由三大全球性事件协同驱动。一是地缘政治的“后冷战”转向。20 世纪 90 年代初期，自由主义与社会主义两大阵营的“冷战”结束，开启了全球化与“开放”时代的浪潮，人类由此获得了三十多年相对和平的机遇，



新的世界力量格局也因此发生明显改变。二是全球权力结构的重组与逆流。“后发达国家”在改革开放的全球和平体系中迅速崛起，冲击了原有的单极权威世界，导致发达国家内部保守主义与民粹主义抬头，因而出现了发达强国发动的逆全球化甚至反全球化的思想与行动。其从全球化的领导者变成了反叛者，从规则的制定者变成了规范的破坏者，于是，世界局势剧烈震荡在所难免。三是科学技术迅猛发展中的信息革命、生物技术革命，尤其是网络技术和人工智能的产能化，迅速地让科学技术变成了存在方式与交往方式的引导者与规范者，因而工业文明时代所形成的规范秩序迅速瓦解，全球化时代形成的全球伦理被束之高阁，人类的政治危机、伦理危机迅速成为亟待解决的难题。

而伦理学在此危机面前所做的各种应急性方案，如应用伦理学迅速兴起，成为解决人类生活实践中出现的棘手问题的“部门伦理学”。于是，出现了两种伦理学模式并行的局面：传统的哲学伦理学探究一般伦理学基础理论，应用伦理学探究时代变迁中出现的专门领域内急需解决的伦理难题。但这种理论与实践的分离，显然就是当代伦理陷入危机的表现，而不能为解决伦理危机提供解决方案。因为二者虽然各有自身的原则与进路，但都只能在越来越专业化的伦理学内部分工中，做越来越精致的专业化的“学问”，为自身的理论提供看似更有理由的辩护性论证，但都逐渐丧失了论证的思想性和智慧性。究其根本，智慧与知识的关键区别在于，前者是为第一性原理而奠基，后者只是寻找存在现象之间的因果关系。更准确地说，前者是揭示生活世界本身内在隐秘的存在机制，而后者只是为已有的规范规则提供辩护理由。因此，一旦生活世界的存在机制发生了根本性的转折，这两种伦理学就都需要完成各自的范式转型。[①]

为了让伦理学重返生活世界，为时代的伦理危机提供具有思想性的智慧的解决方案，本文将分三方面论证伦理学范式转型的必要性与可能路径：一是当代伦理学为何必须完成范式转型；二是当代元伦理学与对亚里士多德德性论的美德伦理复兴为何失效；三是道义实存论作为当代元伦理学的可行性。

一

伦理学范式转型的必要性是由伦理学的学科使命决定的。伦理学属于哲学学科，是区别于单纯思辨的实践哲学。从诞生起，实践哲学便将哲学的“爱智”放置在伦理生活之存在方式中，即寻找正当生活的智慧，它区别于求知的“科学”，通过揭示伦理生命之秘密，为人类安身立命提供智慧，这是其最为根本的哲学使命。亚里士多德在《形而上学》中一开始就明确区分知识与智慧：前者探究事物之所以成为该事物的

① 参见拙文：《数智时代 AI 伦理需要新的范式》，中国社会科学网 2026 年 1 月 5 日，主要讨论了“当代应用伦理学”不可能作为所需要的新范式，本文可视为这一讨论的继续。

可经验的“原因”，后者则探究事物“如何”成为该事物的不可经验的“第一原因”（本因或本原）。因而，哲学通过揭示万事万物成为其自身——“是其所是”（存在）的第一性原理，为人类在这个纷繁复杂、变动不居的现象世界的存在锚定方向与意义，从而得以安身立命于这个瞬息万变的世界。

所谓存在的第一性原理，即存在自身内在向善的机制。从单纯思辨来理解，则为理论哲学，而从实践来理解，即存在如何通过自身的实存原理而完成自身内在本性所规定的至善之实现，即为伦理学。两者都以“自然”存在何以达到至善为主题。区别只在于，思辨的理论哲学是“静观”（存在是什么，至善作为最高的存在方式也是在“静观”中“达到”），因而是关于“存在之为存在”的“知识”（识见）；而实践哲学则强调“行动”，存在论意义上的“行”（实践），就是存在者（人）“是其所是”的伦理存在机制（是其所是的“生活方式”）。因而，伦理学在亚里士多德这里，是不可能遗忘“存在”而能获得善的知识，善也不可能离开“是其所是”的人生实践而仅仅作为一个道德观念。所以，深耕古典学的现代著名哲学家伽达默尔，对柏拉图和亚里士多德的伦理学都做出了一种存在论的阐释：“我要表明，亚里士多德，物理学的创始人和实践哲学的奠基者，坚守着柏拉图笔下的苏格拉底遗产——善是实践性的善。但是，作为物理学的创始人，亚里士多德完成了柏拉图笔下苏格拉底之要求，从对善的经验出发去把握这个普遍者（Universum）。于是在“亚里士多德”那里善就既显现在物理学中，也显现在实践哲学中。在物理学中作为‘目的因’，在实践哲学中作为‘属人的善’。”①

亚里士多德伦理学成为西方传统伦理学的经典范式，正是因为它从人的存在方式本身阐明“实践活动”（行为之原理），即从伦理存在的机制中，揭示存在向善、通往至善的完满实现之机制。在这种自身存在的机制中，阐明了善的含义（相当于“元伦理学”）：善是所有实践活动的终结目标。他这样就让“好的”或“善的”诸如此类的评价性谓词，摆脱了主观性评价的误区，让它有了实体性“归属”：谁的善？内在的还是外在的？亚里士多德通过区分两种“自然”：第一自然（物理学意义上的自然）、第二自然（伦理学意义上的自然），指出前者的善属物之优秀，后者的善属人的生活之完善和人性品质之卓越。物善与人善，都在其存在形式完美实现了由偶然获得的物性或人性之“所是”的最终目的上获得规定，因而是物之优或德之卓（美德）。如此理解的亚里士多德至善存在方式成为德性论。物有物的德性，马有马的德性，人有人的德性。分而言之，这是万物各自禀性的最终实现（完满完成）；统而言之，这是内在于物的物性或内在于人的人性达到其实存的最为卓越的终极状态。所以，无论是善还是德，都渊源有自地根源于存在，在实存中实现。离开存在方式或实存的梯级结构，伦理学是无法言说善与德的，其规范性全都围绕自身之善的存在方式和实存等级而得到规定。

因此，亚里士多德的伦理学之所以成为古代经典范式，又在现代不断经历“复兴”，全赖这种伦理学是在存在的第一性原理上，即在与物理性存在的运行机制的对照中讨论善、德之存在或实存。更清楚地说，伦理学就像物理学揭示自然物的存在密码一样，揭示人类存在之向善的道路，唯有在存在之道路上才能有“逻各斯”之道理的分辨。从规范性语境而言，在哲学意义（非宗教）中唯有对存在之道路的揭示，才是所有伦理之道理的规范性根据。唯有在存在之道作为规范根据之意义上讨论元伦理学才有意义。这一元伦理，实质上就是存在之“语法”（逻各斯意义上的逻辑学）。道义论（即行为的道义原理或法则），在亚里士多德这里，即中道正义论，其论证也无非就是以政治或伦理两种实践（存在）方式中的存在语法为根据，将其证成为政治或伦理生活得以运行得好的机制。德性论更是依据“功能论证”中万物内在灵魂的造就生命之力，揭示存在物之本性的最优实现

① *Hans-Georg Gadamer Gesammelte Werke, in 10 Bände,Werke 7: Griechische Philosophie III*, Tübingen: J.C.B. Mohr（Paul Siebeck）, 1991, S.200.

之路。没有对存在之道路的揭示，就没有什么伦理学问题能够获得有效性证明。

因此，亚里士多德伦理学作为最正统的、以存在语法为核心的实践哲学模式，其存在的完满性决定了伦理生活的整全性不容被分割为元伦理学、道义论、后果论或美德论。伦理学的所有二阶、三阶问题，都只是一阶问题，即存在问题的结构性和实存等级性上的功能性分殊，唯有其整体才能作为存在之语法的正当性言说。而在现代哲学主体性转向之后，伦理生活随着经济、政治和文化各个功能领域分化，加之政治学与伦理学分开，于是从属于存在语法的伦理语法，演变为单纯为主观意图立法的道德语法，或仅关乎个人性格之卓越的美德语法。伦理学随之从对存在之道路的揭示，演变为依主观意图立法的主观理由之争。19 世纪以来的历次亚里士多德复兴运动，都在致力于综合亚里士多德和康德两种不同的实践哲学模式，试图把二阶或三阶伦理语义之争还原为一阶伦理生活的规范结构之争。直至 20 世纪中叶之后，打着“新亚里士多德主义”旗号的美德伦理复兴运动，才把亚里士多德作为实践哲学的伦理学窄化为一种以行动者的品质为中心的“美德进路”，砍掉了其形而上学，也砍掉了其论说善的元伦理语境，更砍掉了其关于中道正义的行动道义的规范伦理论证。而所有这些，都是言说属人之善的存在论前提。人的美德之实现乃是这一论证结构中的一个结论性环节，甚至连结论性环节都算不上。因此，斯洛特最终论证说，亚里士多德伦理学不是“美德伦理学”。我们也完全可以说，美德伦理学根本不能算是亚里士多德伦理学的复兴，因为它使伦理学偏离了正轨。正如德国当代著名哲学家赫费所注意到的，《尼各马可伦理学》结尾处（第十卷 6—9 节）在论证了伦理德性和理智德性作为两种不同的实践模式，哪些德性可以实现存在之至善，哪些德性（如不自制、意志软弱等）不能实现存在之至善后，又呼应开始部分揭示的存在道路之语法，将政治和伦理作为相互从属的生活世界，论证了纯粹思辨的伦理生活为何要高于追求美德的政治生活，但这种论证必须在“属人之善”的伦理层面展开。“更为恰当的，是另一个统摄性的、不再进行等级划分的标题——‘属人的哲学’。这是一种真正意义上相互回转的关系，因为‘形而上学何为’这一问题，是一个实践的乃至实存论层面的问题。伦理学与形而上学作为哲学学科，在很大程度上彼此独立；然而，对一种仅仅致力于纯粹认识、最终献身于自然哲学与形而上学的生活之证成，却归属于伦理学的任务范围。”①

这样，我们也从现代伦理走偏的例子，阐明了当代伦理学范式转型的必要性，其根本理由就是：伦理学只有通过揭示道义存在的语法，才能兑现其引导和安顿人类伦理生活的使命。鉴于人工智能的全面商业化和产业化应用，并在存在论层面引导和规范人类的生产、交往和行动方式，如果我们放任伦理学在不断破碎化的功能领域做专业化的精细学问，而越来越远离伦理生活世界本身，那么不仅“AI 向善”的问题无法得到解答，存在自身如何向善的问题也无法获得解答。

二

当代元伦理学的兴起已经引发了一场伦理学范式革命，这是谁也不会否认的伦理学史的事实。但这一范式革命的实质，是现代哲学“语言转向”在伦理学中的反映。其意义在于，将伦理学研究从伦理生活自身的“一阶规范”转向了对道德语言中二阶“规范语句”的分析，通过对道德语言和语句的逻辑分析，追求思想的明晰性和确定性。这种变革为伦理学提供了一种科学式的自我反思机制，有助于清除传统自然语言表述中语义模糊所引发的混乱。其积极意义在于，伦理“学问”更加精进，问题意识更为明晰，以及增强思想表达的准确性。但其深刻的内在困境在于：无

① Otfried Höffe, *Praktische Philosophie: Das Modell des Aristoteles*, 2., durchgesehene Auflage, Berlin: Akademie Verlag, 1996, S. 186–187.

论二阶道德话语的规范性含义何等清晰，也无法抵消其合理性辩护理由，只有从一阶伦理存在的“存在法相”（Seinsverfassung）[①]获得规范性根据时，才能做出一种“道德中立”的元伦理学辩护。否则，所有的辩护性理由都只是一种新的主观性理由，无法获得正当性的承认。为此，美国法学家和哲学家德沃金对元伦理学的无效性进行了决定性的诊断，得出了一个令所有元伦理学家都无法接受的否定性判断：元伦理学建立在一个错误之上。他的论证如下：根据大多数道德哲学家在普通伦理或道德问题（他们称之为一阶实质性问题）与二阶问题（他们称之为“元伦理”问题）之间所做的区分，道德实在论和外部怀疑论都是元伦理立场。然而，这种区分是一个错误，至少在传统构想的元伦理学意义上是如此。当然，人类学、个人心理学和社会心理学中存在一些有趣的、属于二阶性质的问题，其意义在于它们是关于（about）道德判断的，但其本身并不要求作出道德判断。但在那类问题中，并不存在独特的哲学问题；特别地，道德判断是否可以为真或为假的问题，是一个实质性的道德议题，而不是一个独立的元伦理议题。[②]

德沃金不仅否定了元伦理学关于一阶和二阶问题的前提预设，还借助元伦理学另一个基于休谟从“是”推导不出“应该”的所谓“自然主义谬误”来反驳元伦理学的基本论证全都不能成立，因为既然非道德的“事实”（是）不能证立道德结论（应该），那么同理，非道德的二阶道德话语和道德判断（如语义学或道德心理学）语义分析，也根本无法颠覆或证伪任何道德结论。在关于道德判断语义分析的元伦理学中“并不存在独特的哲学问题”，这既宣布了元伦理学的死刑，也间接否认了以道德心理学论证来复兴所谓的“新亚里士多德主义”的美德伦理学的价值。

因为美德伦理学几乎全盘接受了 20 世纪元伦理的基本预设，同样也遵循着元伦理学关于伦理学类型的划分，所谓聚焦行动的规范伦理学和聚焦行动者的美德伦理学，以此来反对现代的两种规范伦理学——康德主义的道义论伦理和功利主义的后果论伦理的有效性，试图以他们所阐发的极为狭隘的美德伦理学来解决当代伦理问题。鉴于美德在伦理学中根本不是一个元概念，要解释清楚何为一个行动者的美德，需要借助于一系列更为基本的概念才有可能。因此，美德论无论是在亚里士多德伦理学，还是在康德伦理学中，都从属于一个最终目的论证或哲学论证的一个终极问题，即只有解答“人是什么”，才能进一步探讨相关问题。所以，要把美德单独拿出来作为一个唯一合理的“伦理进路”，会导致美德伦理学内部的根本分裂。被视为美德伦理学重要代表之一的努斯鲍姆，勇敢地站出来宣称自己不是美德伦理学家，因为美德伦理根本就是一个“范畴错误”。美德伦理在传统伦理学中基本都有涉及，其在伦理学中的正常任务，在于寻求人在何种存在法相中才能成就人的美德，而不是美德伦理学所要论证的“一个有美德的存在者何以作出道德行动”。哪怕我们忍耐着接受这也能算作一个问题的话，那么这绝非一个二阶层面能够回答且必须返回到一阶层面才有可能解答的问题。而在一阶层面，这个问题就不是一个有美德者，而是任何一个普通的存在者，在其特有的人性状态和“存在巨链”（great chain of being）的规范结构中，如何能够作出合乎其德性的道德行为的问题。因此，也必须完成其范式转换，让伦理学重返“存在法相”所规范的“存在道路”才有可能。因此，无论元伦理学，还是规范伦理学（任何类型），要能解决伦理学在当代伦理生活中的规范失效问题，都需要完成伦理学本身作为实践哲学的范式转型。

三

德沃金已经论证，所谓二阶元伦理学的所有主张，本质上都是一阶道德主张。那么，拒绝讨论一阶伦理是作为其规范根据的存在机制之揭示，而将“存在巨链”自

① 这是海德格尔《存在与时间》中的基本概念，在该书中翻译为“存在机制”，王庆节在海德格尔另一部著作《康德与形而上学疑难》中翻译为“存在法相”。

② Ronald Dworkin, *Justice for Hedgehogs*, Cambridge, Massachusetts London, England: The Belknap Press of Harvard University Press, 2011, p.67.

身的规范结构视为形而上学并予以抛弃的哲学，都只是现代虚无主义的一种意识形态表达，不可能在哲学上有所建树。实际上，在一阶的伦理生活世界中，所有道德主张的合理性与正当性辩护，都不能单纯地作为主观的自由意志之表达，或只是为自己的意愿和理性所支持的“更好的理由”，而必须如德沃金所言，是一个全然超越所有主观性的、中立的“绝对理由”，这样才能表达出世界隐秘的存在机制之必然性，以及为纯粹理性把握到的“应然”的理由。这就回到了我在上文所论证的：绝对的“理由”根源于存在机制所指向的存在“道路”，这才是任何一种哲学伦理学的思想语法，也是所有伦理学即实践哲学的一个基本法门。

像康德这样完成了哲学主体性转向的哲学家，他的伦理学虽以研究主观意志的立法原理为核心，但他最伟大的贡献，在于揭示意志自由这一伦理道德的基础，即作为存在机制的应然的普遍法则对个人主观意志的直接规范。因而，他才能把出于主观意志的道德法则阐释为纯粹理性之自律的自由。将康德伦理学解读为以形式主义为立法原理的主观道德学，是一种错误解读，而将其理解为以行动为中心而不以行动者的品质为中心的规则伦理学，更是误解中的误解——根本没有把康德孜孜以求的重建新形而上学的努力纳入眼帘。海德格尔，也只有海德格尔，能以其深厚的哲学史功夫，独具慧眼地指出：《纯粹理性批判》与“知识理论”完全没有干系，形而上学作为人类本性的内在可能性，乃是对存在物之存在法相的开显，“于是，康德坚守这一形而上学的意图，进而还在特殊的形而上学中把这一意图强化，将之称为‘本真的形而上学’‘终极目的的形而上学’”。“首先将那存在物的存在法相预先给予处理，而所有正在探究的问题，都应该能够和存在物的存在法相关联起来。”[①]由此可见，康德致力于通过超越论存在论之重建来解决伦理学基本问题。对此视而不见的人，就不可能正确地理解康德的道义论和德性论。

这样，我们就已经从哲学伦理学的使命要求，论证了不仅对经典的亚里士多德伦理学和康德伦理学的理解本身需要完成范式转型，而且在现代语言转向之后所形成的当代元伦理学和美德伦理学研究，也需要完成范式转型。因为伦理存在具有时间性，真正起作用的正义伦理根源于伦理生活的存在之道，而当代的存在之道，是历史上从未遇到的千年变局，其中最具决定性的因素是科学技术尤其是人工智能已经成为一种与人类智能共存的智能形式，完成了从单纯的智能工具、向智能体代理人（agent）的转向，在许多具体的智能方面大大超越人类，并正在朝向建构自身的元宇宙演进。虽然大语言模型向“世界模型”的转型才刚刚开始，但这已经是人工智能在第一性原理层面模拟物理世界的努力，一旦成功，人工智能就将拥有自主的世界伦理的规范机制了——这才是人类焦虑的真正要害。目前，许多企业设立“哲学家”岗位，或许预示着人工智能已经从思维机器及其智能工具转向对其自身底层存在法相的建构。“哲学家”岗位并非单纯负责“价值对齐”的程序员，而是将哲学专业意识——对存在法相之揭示或理解的智慧——重组到 AI 模型中去的一个新标志。

连科技公司都意识到且在行动上已经朝向第一性存在机制转型，唯独哲学界依然固执地反对作为有限人性本性之表现的“在场形而上学”，人类就会在等待不断创制出来的便利机器的把玩中，被机器工具无情地抛弃在参与构建世界共存机制的创世者的地位之外，从而失去在未来世界中存在的力量与品质。如果我们的哲学伦理学不能完成范式转型，那么 AI 伦理和 AI 治理就是一句空话。转型的实质是揭示存在自身的规范结构（存在法相），所有伦理学，无论什么类型的伦理学，本质上都需要在“一阶”伦理生活世界揭示道义实存的机制，进而才能有效阐明存在如何向善的规范结构，只有在此规范结构中，人类谈论如何与 AI 共存才是有效的。

① 马丁·海德格尔：《康德与形而上学疑难》，王庆节译，上海：上海译文出版社，2011 年，第 5 页、第 7 页、第 13 页。

美德作为智能向善的实践支柱

王淑芹，首都师范大学燕京资深教授，教育部长江学者特聘教授

人工智能（AI）是当代社会科技发展的显著标志，是不可逆转的社会潮流。在人工智能技术飞速发展的今天，“具身智能”等在社会各领域的广泛应用，使得人工智能作为数智时代的核心技术，正在深刻而广泛地重塑公共管理模式、经济发展方式、社会文化走向以及人们的生活样态。人类发展史上虽然经历过多次技术革命，但“从根本上改变我们这个社会的社会架构和政治架构的技术却屈指可数……但人工智能有望在人类体验的所有领域带来变革”。[①]毋庸置疑，科技已然成为影响社会发展的第四种力量。人工智能技术与工业化、市场化、资本化的有机结合，极大地提高了社会生产效率，为人类创造了更丰富的社会财富，便利了人们的社会生活，但人工智能技术滥用可能引发的安全风险，无不昭示：如何促进人工智能科学技术发挥正效能，减少乃至避免其可能引发的社会风险，应对其给人类生存带来的挑战，是当代全球治理及推进我国国家治理体系和治理能力现代化必须要面对的重大理论与实践问题。

人工智能健康发展需要构筑道德护栏

人工智能作为增强人类认识与改造自然和社会的高效工具，需要在人类的价值统领下发展，而不能遮蔽或背离价值理性。人工智能技术及其应用是服务人类的重要工具，其功能的发挥不能使人的自主、自由和尊严等本质性的类人性受到威胁，它们不能从服务人类的工具变为主宰人类的“主人”。为防范人工智能偏航，避免因滥用而导致“点石成金的国王愿望”式的悲剧，既需要制定和健全相关的法律制度、技术标准以及场景应用的具体管理制度，为人工智能的研发和使用划定红线；同时，也需要科技人员具备良好的道德素养，发挥道德意识、道德直觉和道德良知的内在约束作用，将人类基本价值深度嵌入人工智能系统的全生命周期，主动为人工智能技术扎紧道德护栏，促进人工智能的“算法”、人工智能系统的目标追求与人类共识性的基本价值相一致，从根本上避免人工智能成为人类“自掘坟墓”的危局。基于此，我国新一代人工智能治理专业委员会



发布了《新一代人工智能伦理规范》，明确要求从事人工智能管理、研发、供应、使用等相关活动的自然人、法人和其他相关机构等，都要“将伦理道德融入人工智能全生命周期，促进公平、公正、和谐、安全，避免偏见、歧视、隐私和信息泄露等问题”。[②]

在人工智能领域，为了尽量确保 AI 系统的决策和行为与人类的主流价值观、社会发展规律要求和人类全面进步发展观的偏好及意图保持一致，促进技术安全可靠应用，真正造福于人类，出现了“价值对齐”（value alignment）的应对之策。尽管学界对人工智能与人类价值能否对齐以及对齐程度存有争议，但不可否认的是，人工智能系统的决策与目标导向本身必内蕴一定的价值取向。理想的状态是算法能够兼顾资本增殖逻辑与人类和社会利益的诉求，但当二者发生矛盾时，利润最大化与社会利益最优化何者优先无疑成为难以抉择的焦点。显然，价值选择是人工智能技

① 亨利·基辛格、埃里克·斯密特、丹尼尔·胡滕洛赫尔：《人工智能时代与人类未来》，胡利平、风君译，2023 年，北京：中信出版集团，第 15—16 页。

② 参见《新一代人工智能伦理规范》，https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html。

术的核心问题。作为支配智能体运行的“大脑中枢”的“算法”，既是技术创新程度和技术规则要求的产物，也是科技人员道德观念和价值立场的反映。科技人员作为算法的设计者与编程的思维主体，其道德意识、道德观念以及道德追求必然会渗透到算法指令中。在文化多样和价值多元的后现代社会，因价值的复杂性、模糊性、多元性等特性，会导致“价值对齐”难的问题，即与“谁”（哪种文化的价值观）的价值观对齐。需要申明的是，价值对齐难，既不意味着不能对齐，更不是不对齐的理由。为确保人工智能造福人类而不危及人类独有的特质，价值再难对齐也得对齐，因为放弃价值对齐不符合人工智能的角色定位。人工智能技术是因其能够更好地服务人类才被发明创造出来的，它存在的价值与意义就是要增益人类福祉，一旦人工智能体失去服务人类美好生活的目的性，在终极意义上也就失去开发和利用的价值。即使由于道德原则的抽象笼统性，以及不同族群价值观的差异性，人工智能的算法对齐有一定的难度，但仍要发挥开发、监管、使用者的道德智慧，避免算法违背人类的基本价值。必须承认，人的社会性存在方式决定了人们不得不共同生活，这就从根本上决定了价值文化再多元，社会也会有主流价值观或核心价值观以及人类的基本价值观。正是由于人工智能体是人的思想、价值和目的的体现，驱动智能体活动的各种算法，不可避免地会内蕴设计者的价值观和价值追求。

因此，智能向善，首先需要强化科技人员及应用人员的道德意识，树立正确的道德价值观。科技人员若能坚守人类基本的良善道德原则，就会在算法设计及应用各个环节中，使技术合乎人类共同利益及全社会整体利益，形成人机友好型的智能产品。显然，要避免技术异化为控制或伤害人类的工具，使人工智能不危及人的尊严、权利和自由，首先需要科技工作者秉持人类共识性基本道德，积极践行人工智能伦理，为人工智能的可持续发展构筑道德护栏。

美德是人工智能科技创新的内驱力

对“美德”这个概念需要进行全面的理解，尤其需要回归其原初本义。“美德”是古希腊思想家提出的重要概念，它的原初意义是指人们在社会分工体系中，出色地完成社会指派的职责。美国哲学家麦金泰尔认为，在古希腊荷马时代，“一个履行社会指派给他的职责的人，就具有德性”。[①] 不同职责或角色的德性是不同的。武士的德性是勇敢，统治者的德性是智慧。由之，一个人具有他的特殊的和专门职责上的德性，他就是善的。[②] “善”完全在于社会职责的履行。[③]古希腊哲学家亚里士多德认为德性是一种使人善良，并使其出色运用其功能的品质。[④]也就是说，美德的原初意义是社会成员出色完成本职工作所具有的优秀品德和情操，以及由此取得的优异成绩。只有出色地完成社会职责的人才具有“美德”，意味着这类人具有超出常人的优秀品质。展言之，那些胸怀大志、追求卓越之人，为实现远大抱负和理想目标，往往具有超出一般人的意志品质。这些优秀意志品质不是在社会外在道德要求驱动下形成的，而是基于个体的目的和人格追求“所形成的主体价值观念对主体自身行为的内在自律”。[⑤]显然，优良的意志品质是成就卓越之人的核心素养。

人工智能的创新离不开科学家和技术人员追求真理、严以治学、献身科学、甘于淡泊、勇攀高峰、敢为人先、勤奋刻苦、坚韧不拔、刚毅果敢、诚实守信等优良品质。这些优秀品德是科技人员取得优异成绩的根本保障。具有美德的科技工作者，其研究动力往往是在科学家精神指引下的自觉行为。正因为他们志存高远，才具有强大的精神动力进行科学研究，从而在锲而不舍的远大目标追求中，不断实现科技创新。具有美德的科技工作者，如同无需外力

①②③ 阿拉斯代尔·麦金泰尔：《伦理学简史》，龚群译，北京：商务印书馆，2004年，第31页，第31页，第31页。

④ 亚里士多德：《尼各马科伦理学》，苗力田译，北京：中国社会科学出版社，1990年，第32页。

⑤ 万俊人：《美德伦理如何复兴?》，《求是学刊》2011年第1期。

驱动的自转体，往往是负责任地点燃美好世界火花的人。

美德是防止人工智能致恶的内控力

人工智能全面渗透人类生活已势不可挡。因此，我们既要积极拥抱新技术，也要主动避免其可能带来的社会风险与挑战。科技活动应客观评估和审慎对待其中的不确定性和应用风险，力求规避、防范可能引发的风险，防止科技成果被误用、滥用，避免危及社会安全、公共安全、生物安全和生态安全。[①]毕竟，一旦 AI 发展失控，就会威胁人类的生存基础，使人类面临自我毁灭的危险。

无疑，如何既利用人工智能的功能造福人类，又规避其潜在风险，是目前社会各界共同关心的问题。反观人类社会不断更替进步的原因，人类具备应对各种风险挑战的能力，就像自带防御疾病的免疫系统一样，人类也有应对各种社会问题的内外机制。人类的理性基于社会利益关系协调发展的需要能够进行自我立法，这是人类化解社会矛盾、克服困难的外在手段。除此之外，人类还有由良心良知所构筑的内在约束机制。面对人工智能的安全风险，人类除了主动建立健全相关的法律法规、技术标准外，还具有科技工作者基于正义感和良心的内在约束力。科技工作者若具备基于对社会善恶道德的自我觉知和内省能力而形成的道德责任感，便能在科学研究、技术开发等活动中，恪守人类基本的道德原则，把“增进人类福祉”和“尊重生命权利”作为首要价值，自觉抵制各种违法背德的利益诱惑，充分展现人性的光辉。美德在商业利益面前，虽具有脆弱性，但也有坚如磐石的一面，表现为美德成为人们抵御不义之财侵蚀的绝缘体。具有美德的科技工作者，有献身科学、造福人类的风骨节操，他们会将 AI 系统的安全性、透明性和人的基本权利保护置于首位，坚守国家、社会、人类之大利，在算法程序的编制中，自觉避免算法偏见对弱势群体的歧视，防范数据滥用、隐私泄露、深度伪造等社会风险。

人工智能是一个复杂的技术系统，常常存在试错成本。人工智能的科学研究涉及许多利害关系，且这些利害关系往往具有很强的隐蔽性：一是科学研究中涉及的技术与利益关系问题具有封闭性，通常只有圈内科技人员知晓，社会大众往往无知无感；二是某些科技引发的隐性风险，通常只有研究者自知。在此情形下，如果科技人员没有社会责任感和良心良知，就会只顾维护自身已有的利益格局，隐瞒研究成果产生的隐性或危害长远的社会风险。

众所周知，对科技行为进行规约的法律，一般是在某些科技行为的破坏性后果已产生且社会危害凸显后，政府相关部门才会启动法律程序着手制定相应的行为规范。这就预示，科技人员的许多创新性研发活动，通常是在法律空白的情况下进行的，尤其是人工智能的科技创新，更是处于不断摸索的过程中，往往无具体标准和原则可遵循。显然，在缺乏法律明确规范的相当长一段时期内，科技人员的有些研发活动可能会处于缺乏外在规制的灰色地带，此时人工智能向善只能靠科技人员内心的科学信念和道德良知来维系。在科技人员明知智能产品有设计缺陷并预知风险的情况下，唯有那些具有求真务实良好道德素养的科技工作者，才能把安全原则置于首位，自觉用科技理性节制利益欲望，主动完善智能体的设计和构造，而非不顾社会后果地一味急功近利地进行批量生产和应用。

人工智能领域法律的制定和修改相对于变动的创新活动以及引发的各种社会利益关系而言，常常是滞后的。在未形成具有广泛共识性的人工智能治理框架和标准规范之前，即在人工智能科技活动相关法律法规未出台之前，技术的安全性、可靠性、可控性、公平性在很大程度上依赖科技人员的道德自律。在这个意义上，科技人员

① 中共中央办公厅 国务院办公厅印发《关于加强科技伦理治理的意见》，2022-03-20，https://www.gov.cn/zhengce/202212/content_6688372.htm。

的道德良心是制约科技偏航的内在制动力。换言之，人的良心和善心是避免“纸夹最大化”毁灭人类“回形针”现象的终极制裁力。

这表明，要提升数据真实性和准确性，确保人工智能始终处于人类目的价值指导之下，打造可审核、可监督、可追溯、可信赖的人工智能技术，就需要科技人员具有美德。科技人员具有道德良知，是阻挡资本效用原则践踏科学真理原则的杀手锏。人工智能的价值属性是为人类美好生活提供技术支持，增强人类的能力，一旦人工智能成为隐私泄露的工具，挑战人的尊严，冲击人类的核心价值，人工智能技术就背离了科技价值理性的目的价值。

美德是人工智能健康发展的基石

人工智能向善，除了依靠科技人员的美德保障外，也需要人工智能领域的制规者和评审者具有相应美德。要避免资本的野蛮生长及技术滥用等社会风险，就需要人类发挥自我立法的主体性，对数据保护、算法透明度、责任归属、风险分类监管等方面进行明确的规定，制定出符合人工智能技术健康发展的良法。良善不是法律的天然本性，而是立法者专业素养与优良道德品质有机结合的产物。人工智能领域的相关法律法规及其标准，既需要制规者拥有相关的专业知识和能力，成为行家里手，懂得人工智能技术的特性，也需要制规者具有公平公正的正义美德，能够秉持正义的法律理念，出于公心合理设置权利与义务，摒弃本位主义狭隘利益的局限性，避免法律出现偏私性。因此，制规者具有公平公正的美德是人工智能领域制定出良法的内在保障。

人工智能作为新兴产业，技术创新性强。为规避重大社会风险，确保智能向善，许多国家不仅会加强立法和监管力度，如欧盟制定和颁布《人工智能法案》，旨在根据 AI 系统的风险等级，实施分级监管，对高风险的人工智能进行严格约束，而且往往会设置风险评估环节，组织专家进行伦理审查，进一步把关定性。我国科技部、教育部等部委印发的《科技伦理审查办法（试行）》要求：“从事生命科学、医学、人工智能等科技活动的单位，研究内容涉及科技伦理敏感领域的，应设立科技伦理（审查）委员会。”该“办法”不仅规定“对可能产生较大伦理风险挑战的新兴科技活动实施清单管理”，还要求“科技伦理（审查）委员会对审查的科技活动，可作出批准、修改后批准、修改后再审或不予批准等决定”。《科技伦理审查办法（试行）》对人工智能等风险等级较高的研究项目，要求通过三级伦理审查并实施“一票否决制”，是实现负责任科技创新的制度保障。无论是欧盟的《人工智能法案》还是我国《关于加强科技伦理治理的意见》，都明确要求建立伦理审查和伦理熔断机制，对那些严重影响人类生命安全和尊严的前沿技术，设置动态风险评估，必要时终止项目。显然，伦理审查委员会是控制人工智能技术风险的最后把关人，他们的专业判断和价值评估是规避人工智能技术风险的前置性预警机制。专业性评审专家不仅要具有专业敏感性，能够明察秋毫，发现技术的可能风险或隐性风险，而且要出于公心和社会责任感，坚持科学求真原则，秉公办事，避免因人情、利益输送而出现的偏私性评审。一旦伦理审查的评估环节因利益共谋而导致“伦理熔断机制”失灵，智能向善的“安全阀”就无法发挥保险丝的断电功能，进而无法发挥评审机制过滤、筛选、把关的重要作用。

总之，人工智能向善，既需要将技术发展纳入法治轨道，也离不开科技工作者的向善美德。只有当人工智能技术真正承载造福人类的善意，人类方能迈向一个更高效、更便利的有温度的智能时代。显然，要确保人工智能不危害人的生命、尊严、隐私、权利等基本价值，使人工智能服务于人类福祉这一终极目的，离不开科技人员以及相关监管人员的正义精神和社会责任感，这是智能向善的根本保障。

走向“伦理文明”的伦理理解

庞俊来，东南大学人文学院、中华伦理与未来文明研究院教授，
教育部青年长江学者

马克思在《关于费尔巴哈的提纲》中指出：哲学家们只是用不同的方式解释世界，而问题在于改变世界。当今世界，人们似乎还没有好好地理解与解释这个世界，就急于投身到改造世界的伟业之中。在多元主义文化兴起之后的世界里，人类面临越来越重要的相互理解问题：来自不同文化的主体之间如何相互理解；不可共度的多元文化之间是否有共通的理解结构；人类文明需要奠基在什么样的文化理解的理念之上。

伦理理解的时代气象

时代呼唤伦理，世界需要伦理理解。纵观世界政治与国际秩序，中美博弈、俄乌冲突、巴以危机等一直牵动着全球公众敏感而脆弱的神经，世界的不确定性显著增强。“这个世界会好吗？”一百多年前一位中国老人的未来之问，正成为当代世界公民的普遍困惑。[①]这是深陷战争泥沼的俄乌、巴以民众的呐喊，也是经历“梦幻”般选举的美国选民的疑问；这是正在经历全球文科衰落浪潮的公共知识分子的反思，也是每个清晨连接互联网的亿万普通大众梦醒时分的迷茫……20 世纪 70 年代，伽达默尔就曾经说过，因为“世界政治和社会政治形势的尖锐化”，“理解问题”具有“日益增长的现实意义”，人们需要在“地区、民族、联盟和世代之间建立相互理解”，但是这种“努力遭到了失败”，其根本原因在于不同的地区、民族、国家、世代之间“所用的主导概念”之间“缺少共同语言”。[②]五十余年过去了，世界政治与国际秩序的现状表明，人类之间的相互理解似乎还是没有找到根基，伽达默尔所开启的语言哲学解释学似乎湮没在人类文明的极速发展之中。

从经济与社会发展历程看，现代文明发展的四大引擎——科学、技术、经济和营利，每一个都带有自身根本性的伦理缺陷：科学排除了一切价值判断及科学工作者的主体良知，患有严重的盲症；技术是纯粹工具性的，傲慢地凌驾于精神和情感之上；经济用冰冷的计算将人类互助打入冷宫；营利则贻害所有领域，包括教育、政治、生物及基因研



究领域。科学家在实验室拥有的权力被完全剥夺，集中在企业大亨和强权国家的手中。[③]现代科学技术与市场经济的极端发展，已经日益背离人类对美好生活的追求。我们从“生而自由却无往不在枷锁之中”，开始转换为我们“生在现代性的牢笼却无往不在自由之中”。人类文明的发展开始呼唤伦理的生活，期盼激发理解生活的伦理价值的出场。

为应对世界政治与经济社会发展的当代困境，1993 年——与亨廷顿发表“文明冲突论”同一年——全球 6500 余名宗教领袖、宗教神学学者以及一些有个人宗教背景或没有宗教背景的学界和新闻界人士，在美国芝加哥召开“第二届世界宗教会议”，并发表《全球伦理宣言》。我们需要追问的理论问题是：为什么一个“宗教”会议，最后不是发出“宗教宣言”，而是发出了“伦理宣言”？这足以表明，“伦理”是我们可以通达世界的共通性话语，即便那些不可共度的“宗教”派别，也能接受“伦理

① 1918 年，梁济在与儿子梁漱溟讨论时局时，提出“这个世界会好吗?”的疑问。参见梁漱溟口述、艾恺采访：《这个世界会好吗：梁漱溟晚年口述》，上海：东方出版中心，2006 年。

② 汉斯–格奥尔格·伽达默尔：《诠释学Ⅱ：真理与方法》（修订本），洪汉鼎译，北京：商务印书馆，2007 年，第 220 页。

③ 埃德加·莫兰：《伦理》，于硕译，上海：学林出版社，2017 年，第 15—16 页。

宣言”，伦理理解是多元文化时代不同文明之间相互理解的基本方式。

近年来，中国开始倡导“人类命运共同体”意识，指出“人类只有一个地球，一个世界”，国际社会日益成为一个你中有我、我中有你的“命运共同体”。“命运共同体”更加强调人类社会的相互理解与融通。“人类命运共同体”理念的提出，迅速得到了世界性的学术回应，法国哲学家莫兰认为：“人类还是第一次拥有了共同命运的意识和共同危机感”，“没有任何国家再可以自行其是，或者颐指气使成为霸权”，“人类开始有可能在寰球范围内实践互助的责任伦理和家园共建”，“中国拥有古老的伦理源泉，应当将它们现实化，将其激活，使其再生”，“全世界都在注视着中国，期待她以文明的姿态参与全球治理”。[①]我们这个时代越来越意识到中华伦理型文化传统在当代可能具有的文明力量，期待一种伦理文明引领人类文明发展的新方向。

无论是国际政治秩序还是经济社会发展，无论是知识界的学术探讨还是现实国家的实践主张，这些时代的发展症候都表明一个迫切而重要的理论问题：当今世界越来越需要多元文化之间的相互理解，需要不同文明之间的相互体认。这一宏观的文化理解问题必须落实为不同文化中的具体个体之间的相互理解。以往人类文明史中的不同文化的相互理解，可能是国家与民族之间精英式、局部的、间断性的相互联系与长时段的文化融合。当代文明的文化理解是普遍的、即时的、数量级的、平民化的频繁交往与交流，需要找到每个具体个体能够相互理解的文化通感。伦理理解正是希望将人类理解的哲学诠释学从哲学视野拓展到文明与文化理解的视阈，找到不同文化体系、不同文明个体之间的相互理解方式。

理解的知识困境及其伦理指向

面对人类日益严峻的理解困境，一方面，伽达默尔认为，“一切理解都是语言的理解，这是一个挑战性的论断”，人类相互理解的努力遭遇失败的原因是“好像缺少共同语言”；另一方面，他又认为，人与人之间的相互理解过程表现为语言现象，即便理解过程的对象是语言以外的领域时，或者倾听的是无声的书写文字的时候，理解过程也表现为思维本质的灵魂与自身的内心对话的语言现象。[②]长期以来，西方思想将人与人之间的理解建立在语言的逻各斯之中，以形而上学为其基本表现形式，以知识论的普遍性真理追求为准绳，通过语言哲学不断深化人类理解的理性限度。直到20世纪末，罗蒂指出：“我们必须区分‘世界存在那里’和‘真理存在那里’这两种主张……‘真理不存在那里’，只是说如果没有语句，就没有真理；语句是人类语言的元素；而人类语言是人类所创造的东西。”“只有语句才有真假可言；人类利用他们所制造的语言来构成语句，从而制造了真理。”罗蒂建议我们“不要崇拜任何东西”，要把“我们的语言”“视为时间和机缘的产物”。[③]罗蒂直接揭示了语言的偶然性，从而使得理解的真理性动摇，理解的不可理解性问题凸显。

在此意义上，法国哲学家弗朗索瓦·朱利安，更是振聋发聩地提出：“如果不再从‘存有本体’的认知角度去理解生活的话（欧洲思想，或至少说哲学经常如此做），我们如何着手对待生活呢？换句话说，如果因为我们一开始就身在‘生活’里而与之没有任何距离的话，我们如何‘触及’它呢？”[④]事实上，每个人一开始都是在“生活”中，在“人与人的关系”所维系的“伦理世界”中生活，对人与人的关系的伦理理解是我们生活的基本起点。伦理天生就是关涉“与他人连结，与社区连结，与社会连结，直至与人类种属连结”，[⑤]可以说，伦理本身就是生活，生活本身也是伦理的连结。理解本质上是对生活的理解，是对人与人关系的理解，是在人与人的关系中对生活的理解，也即伦理理解。西方哲学诠释学在探索尽头发现了语言的偶然性，语言哲学的真理性被取消之后，我们需

①⑤ 埃德加·莫兰：《伦理》，第17—26页，第35页。
② 汉斯-格奥尔格·伽达默尔：《诠释学Ⅱ：真理与方法》（修订本），第220页。
③ 理查德·罗蒂：《偶然、反讽与团结》，徐文瑞译，北京：商务印书馆，2003年，第13—35页。
④ 朱利安：《从存有到生活：欧洲思想与中国思想的间距》，卓立译，上海：东方出版中心，2018年，第5页。

要在人与人的相互关系中建构性地理解人本身以及人类所处的时代。

一般而言，理解问题包含理解主体、理解对象、理解过程与时间性四大元素。一种理解得以完成必然是理解主体在一定时间中对理解对象意义统一体的认可，以及由此在主体间形成的共识。从理解的主体与对象关系看，有自身理解问题，有对象理解问题，也有主体间理解问题。从上述要素分析中，我们可以得出，理解有以理解对象为中心的形而上学方式的理解，以理解主体与主体间性为中心的伦理理解，以及以时间性为中心的信念理解。西方哲学诠释学在某种意义上是以形而上学为主导的理解方式，这种理解方式经过存在论的确证、认识论的划定界限以及语言哲学的限度，正进入以实践生活为中心的新纬度，这个纬度就是伦理理解的纬度。

伦理理解是以理解主体为中心的理解，也是以人与人的关系为中心的主体间性的理解。伦理理解是与形而上学理解、信念理解相平行的理解方式。不同于形而上学理解以理解对象的客观性真理为前提，也不同于信念理解最终指向宗教式的终极关怀，伦理理解是以当下的人与人之间的互系性实践生活为中心的相互理解。伦理理解包含着对存在正当性辩护的意义理解，本身就是对意义统一体的体认，内在地包含着形而上学理解与宗教信念理解，能够为多元文化冲突下的全球对话提供理解平台。伦理理解以当下生活世界为中心，理解历史性的知识世界，面向未来的实践世界。伦理理解以对伦理学内部基本问题的理解来解决意义冲突与价值重估问题，以对伦理学与形而上学、宗教的外部关系的理解，来应对全球化时代的多元文化间相互理解问题。伦理理解的内部理解的基本内容有：群己关系理解（契约论）、义利关系理解（美德论）、名实关系理解（规范性）。其外部理解的基本内容有：伦理学与形而上学关系理解，伦理学与宗教信仰关系理解，伦理学、形而上学、宗教三者之间关系的整体性理解。伦理问题是人类的普遍性问题，也是作为具体的人和大写的人类能够感知到的共通性问题，我们可以以此为起点触动具体个体的生命体验。

伦理文明视阈中的伦理理解

叔本华在19世纪就曾指出：“在人类中常有一种不可否认的伦理趋向，来源于（虽然不自觉地）形而上学，并且对人生不按这些路向给出解释，任何宗教便不能赢得立足之地；因为各种宗教全都是靠它们的伦理方面紧紧抓住人的思想的。每种宗教均把它的整个教义当作每个人容易感觉到的道德动机之基础，但他并不因此而理解这一点；并且他把教义和道德动机基础两者连接得如此紧密，以致它们似乎是分不开的……人的理智必须尽力解决的所有问题中，形而上学的问题最难，难到这种程度以致许多思想家认为它绝对无法解决。”[①] 由是观之，伦理、形而上学、宗教是人类理解世界的三种基本方式，而形而上学与宗教最终获得理解必须依靠伦理，伦理是最能“抓住人的思想的”，伦理理解是具有普遍意义的。

遗憾的是，叔本华站在近代哲学的地基上，依然走向了伦理形而上学，还没有意识到伦理理解的根本性。现代以来，西方文明始终试图以现代性来统摄当代世界，将西方之外的世界设定为“未完成的现代性”。哈贝马斯虽然意识到西方启蒙理性在当代遭遇的危机，提出行为交往理论，“希望把理性的实践理解为体现在历史、社会、肉身和语言中的理性”。某种意义上，这种行为交往理论就是建立在伦理理解的基础之上，而不是单纯的西方逻各斯中心主义。但是，哈贝马斯最终还是将这种相互理解的力量中心寄托在欧洲，以一种反问的方式——“除了欧洲，还有谁能有勇气从自身传统当中去获得一定的见识、胆量和幻想——即为了剥夺盲目强迫捍卫系统和提高系统的前提所具有的建构心性结构的力量而必须具备的一切？”[②]

① 叔本华：《伦理学的两个基本问题》，任立、孟庆时译，北京：商务印书馆，1996年，第291—292页。

② 于尔根·哈贝马斯：《现代性的哲学话语》，曹卫东等译，南京：译林出版社，2004年，第410页。

马克思曾以现代资本主义为中心，将前资本主义时代归纳为人对人的依赖阶段，并认为这种对人的依赖关系“起初完全是自然发生的”，将资本主义阶段归结为人对物的依赖阶段，将未来的共产主义社会设想为人的自由全面发展阶段。[①]在人类生产力极不发达的时代，单个人或小群体在人类尚不了解的喜怒无常的大自然面前没有独立性，必须依靠人与人的合作才能够对抗自然以生存下来，这种合作是“自然的”，也是“被迫的”。随着人类认识自然、改造自然能力的增强，最终喊出了“征服自然”的口号，人类从人对人的依赖走向人对物的依赖，进入商品拜物教的时代，消费主义的张扬导致大自然不能承受之重。如前所述，现代文明的四大引擎（科学、技术、经济与营利）已经导致越来越深刻的人类文明危机。这种危机使得我们开始重新思考人与人的关系，呼唤走向人的自由全面发展的新文明秩序。

马克思当年所设想的“建立在个人全面发展和他们共同的、社会的生产能力成为从属于他们的社会财富这一基础上的自由个性”的人的自由全面发展的阶段，某种意义上就是建立在具体个体生命得到彻底绽放基础之上的人类命运与共的伦理共同体，也就是伦理文明的时代。不同于前资本主义时代人对人的依赖是“被迫”“自然发生的”，只是“在狭小的范围内和孤立的地点上发展着”，伦理文明时代人与人的关系是在全世界范围内、通过互联网技术普遍联结的，它既是人对人的依赖，更是人类自觉主动地追求和创造美好生活。这种美好生活是通过人与人的伦理价值认同与共识，确立对人与自然关系限度的认识，在一种人与人的关系可控制的基础上发展人与自然的关系，从而创造一种为人类“伦理支配的文明”。伦理文明时代，我们已从单一地处理人与人、人与自然、人与自身关系的时代，迈入处理这些关系复杂互动的时代，以“人与人”的伦理秩序来规范“人与自然”的发展限度，将“人与自然”的关系控制在“人与人”的关系可以把控的范围之内，从而保障人类文明发展的有序与美好。

伦理文明的发展不是无源之水，也不是无根之木。汤一介先生曾经预言：“21 世纪将形成一个文化上的新的轴心时代。”[②]不同于旧轴心时代各文化共同体在各自时空的独立绽放，新轴心时代是多元文化共同体的汇通与融合。西方现代性文明受到古希腊科学民主的滋润，中东阿拉伯带着希伯来文明的血统，中华文明更是五千年来都没有中断。可以说，新轴心时代的文明就埋藏在旧轴心时代的文化土壤之中。人类在前资本主义时代创造的人对人的依赖的不同方式，构成了伦理文明时代相互进行伦理理解的文化密码。伦理文明时代的人类，需要完成历史的和解。历史的和解中，最重要的就是西方以人与自然关系为主导的科学文化、中东阿拉伯世界以人与自身关系为主导的宗教文化，与中华文明以人与人的关系为主导的伦理型文化的和解。在古今中西交汇的时代洪流中，中华伦理型文化在新的历史时期，一方面在国家发展上提出马克思主义基本原理与中华优秀传统文化相结合的“第二个结合”；另一方面在国际社会提出人类命运共同体理念，正是为时代确证伦理文明的合理性，探索走向伦理理解的人类可能性，将人类从追求个体的美好生活推向追求人类美好生活的文明发展维度。

从伦理文明的视角出发，我们走向了更深层的“伦理理解”。这一理念既有全球宗教群体提出的“世界伦理”，也包含了中国提出的“人类命运共同体”（已被写进联合国重要文件）。二者共同的“真实”“存在”就是人类个体与族群都生活其中的“活的伦理”——命运通感、知识同构、价值共享，构成了人类相互理解的基石。因此，伦理理解既是一种面向生活的人类实践需要，也是一种超越语言的意义理解突破，更是一种文明对话的时代话语探构。

［本文系国家社会科学基金重点项目“中国伦理道德发展指数建构研究”（22AZX013）阶段性研究成果。］

①《马克思恩格斯文集》第 8 卷，北京：人民出版社，2009 年，第 52 页。

②汤一介：《瞩望新轴心时代：在新世纪的哲学思考》，北京：中央编译出版社，2016 年，第 28 页。

如何做一个人类命运共同体的成员

曲红梅，吉林大学哲学基础理论研究中心研究员、哲学社会学院教授

1995 年，为纪念康德《永久和平论》发表 200 周年，德国法兰克福大学召开学术研讨会，会议论文于 1997 年结集出版，论文集与康德原作几乎同名，我们不妨译为《再论永久和平》。[①]三十年过去了，在世界反法西斯战争胜利及联合国成立均已八十多年的今天，人类更接近永久和平了吗？联合国发挥了更大的作用吗？康德原著的副标题为“一个哲学规划”，“哲学”意味着它是理想性的，“规划”意味着它有实现的可能性。在今天，世界和平的目标如何达成？实现和平的主体是谁？这些主体应当做些什么？这是本文要重点讨论的问题。

全球化趋势下人类社会的走向和文明的发展是复杂的：一方面是随着交往的普遍化而出现的趋同和融合现象；另一方面是西方价值观想要突出其地位的分立现象。2017 年 1 月 18 日，国家主席习近平在联合国日内瓦总部发表演讲，提出“世界怎么了、我们怎么办”的疑问，并提出构建人类命运共同体，实现共赢共享的中国方案。[②]和平、发展、繁荣、进步，是千百年来人类共同体始终追求的梦想。实现这一梦想的路径却表现出价值观上的差异。“以全人类共同价值推动构建人类命运共同体”是立足现实、着眼长远的中国方案，它是不同于西方价值观的方案。大力倡导全人类共同价值，体现了马克思主义中国化时代化的新进展。

我们共同面对的难题

在西方学界有关人类共同面对风险的讨论中，最难解决的问题是爱国主义与世界主义的冲突。这实际上反映了基于国家立场和基于全人类立场的不同价值诉求。这一问题可以转换为相应的很多问题。比如，如何对待同胞以外的其他人？对家人的义务与对陌生人的关照存在冲突时，我们该如何取舍？这是处于世界之中的个体在实践中经常会面临的选择。康德的“永久和平理论”为我们解决上述冲突提供了一条可能道路，但这条道路预设和限定条件太多。

康德哲学的解决方式是立足于共和制政体的“公民的爱国主义”（civic patriotism）。



康德认为，共和制政体在原则上保证了站在国家立场的爱国和站在公民个人立场的爱国是一致的，因为共和制“首先是根据一个社会的成员（作为人）的自由原则，其次是根据所有的人（作为臣民）对于唯一共同的立法的依赖原理，第三是根据他们（作为国家公民）的平等法则而奠定的”。[③]在康德看来，共和政府是为满足公民的共同利益服务的，而公民是为追求共同政治利益而结合在一起的自由平等的个人。“公民的爱国主义”表达的就是对共同享有的政治自由和维护这种自由的制度的热爱。它注重个人的公民身份（citizenship），具有内在的政治本性。通常情况下，推进普遍正义的一般责任与为国家谋求正义的特殊责任之间存在冲突，个人作为国家公民的完全责任与作为世界公民的不完全责任之间存在矛盾。但康德相信，作为国家公民与作为世界公民在共和制中是和谐一致的。

葆琳·科林赫尔德（Pauline Kleingeld）在《康德与世界公民主义》中指出，康德

① 参见 James Bohman and Matthias Lutz-Bachmann, *Perpetual Peace: Essays on Kant's Cosmopolitan Ideal*, Cambridge and London: The MIT Press, 1997.

② 习近平：《论坚持推动构建人类命运共同体》，北京：中央文献出版社，2018 年，第 414—426 页。

③ 康德：《历史理性批判文集》，何兆武译，北京：商务印书馆，1997 年，第 105—106 页。

所描述的世界公民比古希腊个人主义的世界公民原型更有价值。[①]古代世界公民主义的发端者昔尼克学派所坚持的其实是一种极端个人主义态度，完全不顾社会习俗，为我们呈现出一个无牵无挂、随遇而安的个体的典型形象。在科林赫尔德看来，康德眼中的世界公民表现出另外一种行为态度：对作为全球共同体成员的其他人类个体、文化和民族的认知、尊重、开放、好奇、仁慈和关注的态度。康德甚至进一步认为世界公民应当是优秀的爱国者。一方面，康德认为世界公民不但要“认识世界”，更要“拥有世界”，而“拥有世界”的主要方式就是参与人类实践活动，因此，对身边的人和事漠不关心并非世界公民应有的态度；另一方面，康德认为世界公民想要向外寻求认识和影响，就“必须实现在家里通过与自己的城乡同胞的交往来获得这种知识”，否则，世界公民的人类学经验是有限的。[②]基于此，科林赫尔德认为，昔尼克派的世界公民主义仅存在于抽象的道德领域；康德虽然也是根据他的道德形而上学来阐发世界公民思想、坚持一种世界公民式的道德理论，但他并未停留在道德领域，而是从政治、文化、历史等多维度论证“世界共和国”的现实性。

但我们发现，康德的“哲学规划”是非历史的。他认为只有通过共和制才有实现“永久和平”的可能性。虽然涉及了具体的政治安排，但他以理想的道德世界共同体为蓝本规划现实政治共同体，使得这种世界共同体的实现面临重重困难，举步维艰。这不禁让人思考：康德“永久和平”的蓝图本身是不是有问题？以阿拉斯代尔·麦金泰尔为代表的哲学家们就是站在与康德完全不同的立场上看待我们面临的难题。

麦金泰尔以一种“历史主义”的态度反对世界主义，强调爱国主义是美德。在《爱国主义是一种美德吗？》一文中，他认为通常的爱国主义是对某个特定国家的忠诚，只能通过个人与国家的所属关系展现出来，爱国主义是一种体现美德的忠诚。[③]他同时指出，这种爱国主义和我们通常理解的客观道德（即从非个人的、不偏不倚的立场进行道德判断）是对立的。麦金泰尔认为，爱国主义具有道德特殊性，对国家和民族的热爱和忠诚不仅是美德，而且是关键的美德。[④]道德义务的主要内容和本性就体现在我们从何处、向谁学习道德。通过这个定义，麦金泰尔确定了个人的爱国主义态度与这种态度所产生的环境之间本质性的、历史性的、不可磨灭的关系。因此，在他看来，世界公民主义并非理解道德的最好路径。他更倾向于一种必然指向爱国主义的道德。爱国主义道德是伴随着某种特定的社会、政治和经济框架而体现在某个特定社会共同体成员身份之中的。我们所获得的道德的本质特征正是源于我们生活和学习所处的那个独特的共同体。伦理学家的中心任务是说明他/她所生活的社会的核心信念，使这些信念成为国家稳定、人民坚守的信念。这些信念只有在特定的人群关系中才能形成，只有在其具有特殊性时才是可接受的。也就是说，只有在共同体中，真正的道德才能实现，因为在共同体中一个人生成并保持了他作为一个道德行动者（moral agent）的地位。

从世界历史的观点看

康德和麦金泰尔的不同方案，让我们对推动建构人类命运共同体所面临的困难有了更切近的认知。但这里有一个需要认真对待的问题：人类命运共同体并不等同于康德关于永久和平的“哲学规划”。也就是说，康德的永久和平理想与麦金泰尔的爱国主义道德是建立在两种完全不同的人性论和历史观之上的，两者没有达成一致的可能。但本土共同体与人类命运共同体之间虽可能存在一定紧张关系，却并不像康德式世界主义和麦金泰尔式社群主义那样尖锐对立。从根本上说，人类命运共同体意识和理念是一个切实可行的、既保障本国利益又兼顾各国共同发展的方案，它建构

① ② Pauline Kleingeld, *Kant and Cosmopolitanism*, Cambridge: Cambridge University Press, p.1, p.1.

③ ④ Alasdair MacIntyre, “Is Patriotism a Virtue?” in *Theorizing Citizenship*, Ronald Beiner ed., Albany: State University of New York Press, 1995, p.210, p.218.

起国家和世界之间共生共存的合理关系。

马克思的世界历史理论对理解处于全球化之中的人类共同体及其持有的价值体系具有重要理论价值。大多数传统意义上的社会历史理论与社会发展理论都是在“民族国家”的意义上理解社会的。随着社会发展越来越趋向于“世界化”，即全球化，传统理论的局限性就暴露出来。一方面，这种理论完全在社会（民族国家）内部寻找变化发展的机制，忽视了在全球化条件下各国之间的相互依赖和相互作用，忽视了“世界历史”整体对各个民族国家的塑造；另一方面，这种理论抛弃对人类整体发展的深入研究，将小范围适用的经验推广到全世界。例如，西方现代化理论就企图简单地把西方社会的发展经验推广到非西方社会，进而宣称，西方社会的发展代表人类整体发展的正确方向和最高水平，应该成为全球效仿的楷模。这种研究社会的封闭式思维方式既无法解释在世界交往中不同地域和国家的政治、经济、文化、文明之间的整合与冲突，也不能为在全球化背景下各民族、国家的发展提供合理有效的发展战略。

马克思的世界历史理论是理解和分析现代性及当代社会发展趋势的重要理论武器。我们在马克思主义的经典著作（尤其是《德意志意识形态》《共产党宣言》《资本论》）中可以发掘深刻的世界历史理论的内涵和意义，并在方法论和思维方式的维度理解现代性批判和当代社会实践。世界历史理论不是历史唯物主义的组成部分，而是其现实表征，是马克思历史观不同于其他历史观的关键。马克思的世界历史理论体现了马克思主义哲学所具有的“世界历史规定性”和所依据的“经验事实”的全球性。[①]在这个意义上，马克思的世界历史理论表明了人类社会实现个体全面而自由的发展的现实化路径。“世界历史”是一种新的存在方式，为人的彻底解放创造了前提。人类解放与世界历史进程紧密相连。世界历史是“现实的人”所进行的生产活动与人们之间的交往关系。马克思对世界历史的考察与反思始终以生产实践作为历史的基础，以现实的从事实践活动的个人的解放和全面发展作为世界历史的核心。[②]

马克思的“世界历史理论”所表征的独特的思维方式，就是“历史世界化的思维方式”，或曰“社会发展全球化的思维方式”。这种思维方式包括以下几个理论原则：一是“整体性原则”和“有机论原则”；二是“全人类的利益高于一切”的原则。根据世界历史的思维方式，面对“历史向世界历史转变”的命题，我们既不能用既定的或假定的理解“人”的方式来为人类构筑“永久和平”的理想，也不能只考虑狭隘地域范围内的人类群体的特殊性，而应基于人们之间日益紧密而广泛的交流形成对个体的人及其相互关系的理解，基于现实的影响人们抉择的价值排序形成价值共识。“全人类共同价值”意在促进人类命运共同体的构建，它不是高高在上难以企及的理想，而是全世界人民能够共同认可的基本价值前提。

我们都是人类命运共同体的成员

以全人类共同价值推动构建人类命运共同体，能够体现全球化视域内的相互尊重和交流，是对人类共同体的维护和发展。在西方社会亦曾出现过几种保护人类共同体的设想。但这些方式要么是保守的，要么是暂时性的，要么过于理想化。自由主义视域内的“区别性对待”方式就是一种保守的方式。威尔·金里卡认为文化和价值观以其广阔的背景为每一个个体自由的实现提供了丰富有意义的选择，但由于人数和影响力的差别，小众群体不可避免地会遭遇不平等对待。因此，金里卡提出“区别性对待的群体权利”（group-differentiated rights）方案，[③]对小众群体给予特殊保护，弥补不平等对其造成的伤害，并且促使其成员做出有意义的选择。他期望通过“区别性对待”使每个个体的道德选择在社会文化中获得认可，同时也保证个体能够在这种文化中自由呼吸、自由发展。然而这

① 叶险明：《马克思世界历史理论的特性与世界历史理论基本问题——马克思主义世界历史理论在当代发展的一个重要逻辑环节》，《马克思主义研究》2010年第1期。

② 赵士发：《从实践观的思维方式看马克思的世界历史理论》，《东南大学学报》（哲学社会科学版）2003年第3期。

③ 参见金里卡：《多元文化公民权：一种有关少数族群权利的自由主义理论》，杨立峰译，上海：上海译文出版社，2009年，第2页。

样的保守方式除了思想家的呼吁，并不能切实产生效用。

后殖民主义视域内的“特殊保护”方式也是暂时的。在后殖民主义的视角中，人们已经不再关注民主社会内部主流群体与小众群体间的文化正义问题，而是关注自由民主的政权与非自由民主的政权之间的文化正义问题。这种文化多元主义已经从向国家政权索要对原住民的特殊权利和保护，发展为诉求原住民保有其自我管理和自我发展的主权。海库·帕瑞克（Bhikhu Parekh）认为，自由主义文化并不比其他文化更胜一筹，以自由主义的价值观评判非自由国家就是不正义的。[①]“特殊保护”方式强调了不同社会文化之间的共生共存问题。弱势的社会文化要求特殊保护貌似合理，但这种特殊保护方式是在不平等前提下的一种不得已的方式。

同样是倡导人类共同体的美好未来，为什么人类命运共同体与自由主义、后殖民主义乃至世界公民主义所倡导的人类共同体都不一样呢？这与不同观点对“人”的不同理解有关。历史唯物主义视域下理解的人是“现实的人”，是现实的“诸个体”（real individuals）。“全人类共同价值和人类命运共同体中的‘全人类’‘人类’一词强调的是由‘现实的人’组成的共同体，强调的是人类命运的相互联系和相互依存。”[②]全人类共同价值和人类命运共同体是对“全人类”意义上人的自由全面发展的自觉追求，体现了马克思主义的人类立场，是我们与西方资本主义国家在价值立场上的本质区别。西方价值观的起源和应用都是为了服务特定的群体利益，是为实现少数群体的利益而强行推行的价值认同，并对其他价值观进行压制，强迫世界接受西方文明体系，推行霸权主义与强权政治。“人类命运共同体”蕴含的价值观是从“现实的人”出发，“把人真正理解为在与他人内在统一的社会化的、一体性关系中生存发展的开放、包容的存在。这种新哲学意识充分彰显了人类相互依存、命运与共的整体性与内在相关性”。[③]全人类共同价值是基于马克思主义对民族历史向世界历史转变的判断，基于人类日益紧密的社会、经济和文化交流，基于现有的价值序列形成的共识，是在不同国家、民族交往过程中产生的凝聚全人类共识的“最大公约数”。其目标在于构建人类命运共同体，共同塑造更加美好的未来世界。在国际局势动荡不安的今天，“全人类共同价值”能够促进构建一个和平稳定、合作共赢的世界。

作为历史唯物主义出发点的“现实的人”为我们理解“人类命运共同体”提供了规范性依据。构建人类命运共同体的核心对象是“人类”。在这个意义上，我们都是人类命运共同体的成员，并应自觉承担构建和维护人类命运共同体的责任。世界历史理论和“现实的人”是从不同侧面揭示的人类社会的现实及其运动，体现了马克思主义哲学最为关注的“人类化的社会”或“社会化的人类”。自由主义、后殖民主义和世界公民主义对“人”的理解都是基于个人主义的。马克思所说的“现实的人”既不是德国古典哲学理解的抽象的“人”（Mensch），也不是功利主义中同质化的“个体”（individual），马克思关注的是“现实的人”的生存和解放，也就是全人类的生存和解放。

人类命运共同体关注人的生命与健康、生存与发展。构建人类命运共同体的重大理念，既体现了“和平、发展、公平、正义、民主、自由”等全人类共同价值，又体现了“公正、合理、包容、合作、共享”的时代精神。构建人类命运共同体重大理念得到国际社会的高度评价，已被写入联合国多个文件，有效影响着包括人权在内的全球治理体系变革。这是从中国共产党和中国国家的性质出发，从中国与世界的关系出发，以马克思主义唯物史观理解我们所处的历史方位和人类社会发展大势而获得的必然结果，是中国向世界提出的解决全球性问题的独特方案。

[本文系国家社科基金重大项目“马克思主义经典作家的共同体思想研究”（23&ZD199）阶段性研究成果。]

① Bhikhu Parekh, *Rethinking Multiculturalism: Cultural Diversity and Political Theory*, Cambridge, MA: Harvard University Press, 2000, p. 96.

② 李健：《历史唯物主义视域中的全人类共同价值与人类命运共同体》，《思想理论教育导刊》2024年第7期。

③ 贺来：《马克思哲学的“类”概念与“人类命运共同体”》，《哲学研究》2016年第8期。

探寻道德舆论的应然模式

蒋艳艳，东南大学人文学院、中华伦理与未来文明研究院副教授

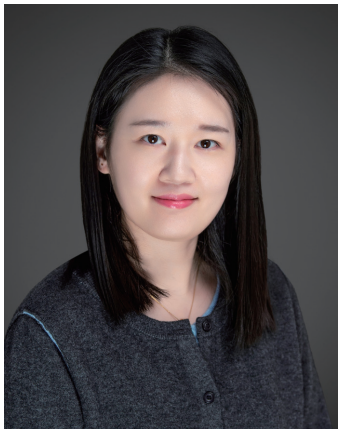
不确定时代的到来颠覆性地改变了道德生活的实现方式，道德舆论成为道德生活的重要表达方式。然而，不确定时代的道德舆论也潜藏着一些不稳定因素，使其伦理效应的发挥既存在机遇也面临挑战。因此，探寻道德舆论的应然模式，是当下道德生活理论探索与治理实践亟须关注的重要课题。

当代道德舆论的伦理定位与伦理机理

自伦理道德产生以来，人们一直借助传统习俗、内心信念和社会舆论来实现道德生活。[①]但在早期，相较于传统习俗和内心信念，社会舆论的作用相对有限。道德生活最初主要依托传统习俗，其中的规范性内容可为道德生活创造相对稳定的精神源泉。社会舆论充其量只是道德生活的一种外在机制。

当下，我们步入了一个不确定时代，文化不确定性的加深，使道德生活的实现格局发生天翻地覆的变化。一方面，传统习俗呈现祛魅态势。文化的不确定性变革了原有文化样态，不仅弱化了传统习俗赖以存续的文化资源，也改变了传统习俗得以发挥作用的文化环境。不同习俗在碰撞与冲突中逐渐祛魅，以至于难以如往常一样为道德生活提供稳固的精神支持。另一方面，内心信念的权威性下降。文化的不确定性决定了价值冲突的客观存在，当人们长期身陷于价值冲突之中，个人理智能力会遭到质疑，文化系统也会遭到冲击，个体内心信念可能出现迟疑不决、反复无常乃至缺位离场的状态，无法有力帮助人们在道德生活中进行可靠的道德判断。

尽管如此，不确定时代的到来让道德舆论获得了更多的发展空间。道德舆论发酵的原因之一即是话题的争议性。如果一个话题总是能很快得到正确回答，相关舆论很难发酵并持续。而不确定时代包容异质观点的表达，恰恰能够增加话题的争议性，在潜移默化中激发道德舆论的形成。同时，与不确定时代相伴而生的媒介技术的深度变革，特别是网络媒介技术的发展，带来了数字化语言构型与去中心化交往行动，破除了道德舆论的时空限制，提升了道德舆论的涉及广度和影响力度，使其在道德生活中发挥越来越大的影响力。总之，当代道德舆论有取代传统习俗和内心信念之势，成为道德生活的重要表达方式，扮演起道德生活的晴雨表和风向标。



既然当代道德舆论在道德生活中具有举足轻重的伦理地位，那么其究竟具备何种伦理机理才得以产生伦理效应？首先，道德舆论由道德认知、道德态度和道德规范三重伦理因子构成。道德舆论是众多以道德评价为形式的道德意见的集合。这种意见本质上是一种评价性的道德语言表达。借鉴道德语言伦理因子的形上分析，[②]道德认知是第一重伦理因子，是发言者对是非善恶等行为及其执行状况的基本认识，既包括对象相关的物理事实等自然形态，也包括在一定历史条件下形成的对象相关的制度事实；道德态度是第二重伦理因子，是发言者对他人、对群体的行为、人格等所持有的一种稳定看法，是发言者自身表达出的主观立场，反映出发言者的情感或意愿；道德规范是第三重伦理因子，是发

① 参见朱贻庭主编：《伦理学大辞典》（修订本），上海：上海辞书出版社，2011年，第14页。

② 参见杨国荣：《道德与语言》，《学术月刊》2001年第2期。

言者评价人们行为善恶的具体标准，是人们在道德实践中形成并应当普遍遵循的行为准则。三重伦理因子并非彼此独立，道德认知以道德规范为依据，而道德态度来自道德认知，又对道德规范具有引导甚至建构意义。正是在三重伦理因子的互动中，发言者的道德语言得以表达，进而为道德舆论的生成奠定基础。

其次，道德舆论呈现出从个体性的道德意见，经主体间性的道德讨论，向公共性的道德舆论发展的伦理演化过程。道德舆论总是起始于分散、无序的个体性的道德意见，上述伦理因子即是个体性道德意见的伦理结构展现。但个体性道德意见一般不会止步不前，而是会在语言作用的引导下自然走向具有主体间性的道德讨论，这本质上是一种基于对话的交往行动，交往行动正是构建公共性的重要前提和基本方式，所以在主体间性道德讨论的带动下，个体性的道德意见将逐步发展为公共性的道德舆论。

最后，道德舆论具有影响社会道德秩序维系和个体道德品质养成的双重伦理效果。前述关于伦理因子的探讨已通过还原论策略，将道德舆论还原为一种道德意见，进而还原为一种评价性道德语言，从而初步指出道德舆论对道德规范的影响。这种道德规范从狭义来说正是社会道德秩序的重要构成。即便从广义来讲，囊括如法律体系、政治制度乃至宗教价值的社会道德秩序，其内核仍在于道德规范。当然，现实的道德舆论并非单一的道德意见，而是诸种道德意见的集合。在其趋向共识的动态过程中，各种意见相互作用，可能产生更为复杂的后果。在个体道德品质方面，道德舆论的影响取决于个体在对话实践中是否将道德规范凝化为其内在人格，进而是否将道德规范凝聚为其内在德性。其中，个体既包括相关舆论中道德行为的当事人（行为者、受影响者等），也包括参与舆论的众多旁观者。相关影响往往与社会道德秩序的后果产生联动效应。

道德舆论在不确定时代面临的机遇与挑战

在不确定时代，道德舆论的重要性日益凸显，有望成为实现道德生活的关键途径。然而，其自身在不确定时代也潜藏着一些典型的不稳定因素，这些不稳定因素使得道德舆论的伦理效应呈现出机遇与挑战并存的复杂局面。

第一个不稳定因素是在动员和宣泄之间摇摆的情绪表达。这一不稳定因素与道德舆论中发言者的道德态度相关。在不确定时代，道德态度所表现出的情感倾向正逐渐向情绪倾向转变。这源于三方面的变化：一是传播主体从被动接收变为“传受一体”，激发同情或仇恨等情绪表达；二是网络媒介的数字化、超时空等特性为情绪释放提供了技术条件；三是社会转型加剧了风险与矛盾，助长了负面情绪的宣泄。回望伦理思想史，人们对情感往往给予更多的伦理期望，而对情绪虽不否认其伦理意义，却保持相对谨慎的态度。究其原因，主要在于情绪和情感不同，特别是和儒家伦理所理解的先天必然的、内在的、普遍的情感不同，情绪本质上是一种个体化的表达，是个体心理方面的知觉和感觉，更具伦理张力，也潜藏着一定的伦理风险。当情绪是一种良好样态时，它能通过对不道德现象的批判，实现对道德生活的动员与助力；但在网络媒介的传播平台，情绪又容易被过分放大，易于形成“以自我为中心”的“网络个人主义”，也易于造就“部落化”的“信息茧房”和“回音室效应”，这些不良情绪更多是一种宣泄式的表达，甚至会成为作恶的工具和帮凶，无益于道德生活的实现。

第二个不稳定因素是在融通与疏离之间徘徊的主体间关系。主体间性的道德讨论是道德舆论伦理演化的关键环节，但在不确定时代，主体间的关系正在发生重大的形态转变，因而使道德舆论既存在机遇也面临挑战。当下，主体间的交往不仅能在传统的意义上“以有形客体为中介”，[①]而且借助于网络化的技术平台，

① 高鸿：《数字化时代主体间性问题研究》，上海：上海社会科学院出版社，2008年，第16页。

更多地能“以无形客体(信息)为中介”。[①]在后者意义上，由于作为无形客体的信息不受时空限制，所以网络媒介的确可以推动世界范围内更普遍的交往，也有助于人与人之间形成平等和互动的融通关系。在道德舆论中，这种融通关系能帮助主体间进行充分的道德讨论，展开深入的道德理解，进而推进个体性的道德意见向公共性的道德舆论的演进，推动道德共识的达成。事实上，“以无形客体（信息）为中介”的主体间关系并非一直能够形成良好的融通状态，作为中介的信息往往会让主体间关系陷入不良的疏离境地。例如，信息权力的操纵将导致关系失衡，掌握信息资源和技术优势的少数个体或组织，可以通过算法推荐、社交机器人等技术手段操控信息流动，形成信息茧房，使公众在受控的信息环境中难以实现真实沟通。

第三个不稳定因素是在多元与虚无之间游移的共识观念。共识不仅是道德舆论伦理演化的目标，也是道德舆论伦理效果的基础。在不确定时代，由于“不同的人有着不同的文化、教育、社会等方面的背景，有着不同的生活际遇与人生体悟，也有着异质性的、特定的需要回应的生活挑战”，[②]共识不再是一种一元主义的追求，而是一种多元主义的态度。这种观念很容易诱发相对主义，而相对主义则可能进一步堕落为虚无主义。具体到道德舆论领域，道德多元主义可以接纳人们就道德事件形成的差异性和多样性的共识，这在一定程度上有助于激发道德生活的活力，但与此同时也容易忽视普遍道德规范的追求，导致道德分歧与道德冲突的产生，为道德相对主义提供土壤。起初，这种道德相对主义可能只是一种文化相对主义，即认为不同的文化具有不同的道德主张。当文化相对主义蜕变成主观主义时，道德标准便取决于个体行动者，甚至成为个人的主观感受。在极端情况中，不仅人与人之间的道德批评或评判是不可能的，而且人们可能对道德规范产生怀疑与否定，最终导向道德生活的祛魅，导致道德虚无主义的产生。这种虚无观念无疑是一颗思想“毒瘤”，将对社会道德秩序的维系与个体道德品质的提升产生不可估量的负面影响，不仅会消解道德存在的根基，使道德的神圣性遭受践踏，而且会导致部分群体公然鄙视道德规范、藐视道德底线、漠视道德责任，忽视道德人格的塑造，最终引发全社会的道德秩序危机。

以对话的道德影响力为旨趣的应然模式

面对道德舆论在不确定时代可能遭遇的种种机遇与挑战，为了更好地发挥道德舆论的伦理效应，展现道德舆论的伦理地位，探寻道德舆论的应然模式势在必行。而应然模式的关键在于以对话的道德影响力为旨趣。

这一路径的选择，意味着对对话的道德影响力的承认。自古希腊哲学始，对话的道德影响力便已悄然显现。苏格拉底以“问答法”探寻真理，主张通过对话显明道德原则；柏拉图以对话体构建哲学体系，在《理想国》中借对话探讨正义等议题；亚里士多德则强调知识与真理需经由对话与辩论获得，并指出友谊作为道德生活的基础亦依赖深入对话。即便在康德以“绝对命令”为代表的主体性伦理学中，仍潜藏对话逻辑。韦尔默指出，康德伦理学隐含对“所有相关的人的观点”[③]的考量，可视为一种潜在的对话原则，这也为后来阿佩尔、哈贝马斯等人发展话语伦理学提供了可能。20 世纪兴起的语言学转向进一步确立了对话的伦理学地位。话语伦理学关注对话的规范合理性问题，元伦理学则聚焦对话的道德语言意义，两者均以“对话”为焦点。除此以外，这也是基于道德对话作为道德舆论形成重要环节这一基本事实。道德舆论是在诸个体表达道德意见基础上，凭借相互之间的道德讨论而形成，这种道德讨论正是一种道德对话的形式。

以对话的道德影响力为旨趣来探寻道德舆论的应然模式，意味着要构建出一个合乎规范合理性的道德对话，即让道德舆论符合话语伦理要求的对话合理性。其一，

① 高鸿：《数字化时代主体间性问题研究》，第 16 页。
② 贺来：《“主体性”的当代哲学视域》，北京：北京师范大学出版社，2013 年，第 94 页。
③ 阿尔布莱希特·韦尔默：《后形而上学现代性》，应奇、罗亚玲编译，上海：上海译文出版社，2007 年，第 11 页。

形成一种合理的评价性的道德语言。这需要道德认知、道德态度和道德规范三重伦理因子及其互动皆充分体现出对话合理性。道德认知要以对象相关的物理事实和制度事实的真理性为基础和前提，要对与对象相关的物理事实和制度事实进行仔细的观察、分析，在此基础上才能进行明智的判断、得出正确的结论。在合理的道德认知基础上，进一步表达出合理的道德态度，尤其要注意合理地将事实判断与价值判断结合起来，合理地使用道德术语，合理地发挥理性的作用并妥善对待理性之外的因素。道德规范既存在于道德认知的发生环节，也关涉道德态度的后果意义，还需要合乎对话合理性的要求。一个合理的道德规范一定是以道德为内容的规范，并且这种规范需要与行为密切相关。更重要的是，要妥善处理好可能的争论之处，如道德规范常常会陷入的相对主义或绝对主义两难选择。

其二，达成一种“不定的”道德共识。道德舆论从个体性到主体间性再到公共性的伦理演化过程也需要合乎对话合理性，这集中表现在某种道德共识的达成。哈贝马斯的话语伦理思想曾设想以“理想对话情境”来达成合理的道德共识（真理的道德共识），但显然当下虚实融合的网络空间很难真正成为“理想对话情境”的实践场域，且“理想对话情境”的预设本身忽视了具体情境中的有限性、模糊性、破碎性和时间性，反而使对话丧失了其存在的价值。因此，需要有一种新的方案来形成新形态的道德共识。在此，可借鉴韦尔默的观点：合理道德共识的生成不再是“理想对话情境”中的问题，而是“必须置于某些具体的特定情境中”，[①]不再是终极奠基问题，而是道德的应用问题，即在何种情况下使用何种道德规范才合理的问题。在此意义上达成的道德共识是一种“不定的”道德共识，它对于特定的道德情境而言是一种“当下确定性”。具有“当下确定性”而非“普遍确定性”的道德共识能够更大程度地、更加全面地把握不确定时代的规范性与偶然性，也能尽可能地保持和激发道德舆论在不确定时代的伦理活力与优势。

其三，实现社会道德秩序与个体道德品质的协同发展。当道德舆论的伦理因子结构和伦理演化过程均合乎对话合理性时，社会道德秩序的维系将成为可以预期的结果。因为上述“不定的”道德共识可以为我们提供一个更具有开放性的道德规范体系及其应用模式，或是延续着现有道德规范及其作用，或是建构出新型道德规范及其影响，从而服务于新时代社会道德秩序的维系。困难之处在于，何以凭借对话合理性进一步将社会道德秩序内化为个体道德品质。福柯关于“直言”的话语伦理思想或许可以给我们带来启示。福柯所说的直言形式来自古希腊时期真理或真理的实践技术，意指“某种说出一切的活动”。[②]在直言与“关心自己”有关的演变方向中，人们在直言中可以完成对自我的审慎检查，进而帮助塑造个人的道德生活标准，提升个人的道德品质；在与“治理或者说支配他者”有关的演变方向中，人们通过直言来分析直言对象的处境，可以帮助直言对象认识自身和返回自身，使直言对象在反思中理解并建构自身的道德主体性。倘若道德舆论的发言者能够坚持直言，那么在维护社会道德秩序的同时，不仅发言者自己有望更好地培育个体道德品质，作为发言对象的行为者也将获得提升个体道德品质的机会，最终，社会道德秩序与个体道德品质的协同发展得以实现。

综上，以对话的道德影响力为旨趣，为不确定时代探寻道德舆论的应然模式指明方向。其核心在于将对话合理性贯穿于道德舆论伦理机理的各个环节。唯有如此，才能从容应对道德舆论在不确定时代可能遭遇的种种机遇与挑战，才能充分发挥道德舆论在不确定时代应当具有的正向伦理效应，使道德舆论成为不确定时代有效的伦理应对方式。

［本文系国家社科基金青年项目“网络空间的道德舆论研究”（20CZX060）阶段性成果。］

① Albrecht Wellmer, *Ethik und Dialog*, Frankfurt am Main: Suhrkamp, 1986, p.114.

② 米歇尔·福柯：《说真话的勇气：治理自我与治理他者Ⅱ》，钱翰、陈晓径译，上海：上海人民出版社，2016年，第10页。