

《现代远程教育》网络首发论文

题目： 本体、认识与价值：智能教育的技术伦理风险隐忧与治理进路
作者： 韦妙，何舟洋
DOI： 10.13927/j.cnki.yuan.20211012.005
网络首发日期： 2021-10-13
引用格式： 韦妙，何舟洋. 本体、认识与价值：智能教育的技术伦理风险隐忧与治理进路[J/OL]. 现代远程教育. <https://doi.org/10.13927/j.cnki.yuan.20211012.005>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

本体、认识与价值：智能教育的技术伦理风险隐忧与治理进路

韦妙 何舟洋

(湖北工业大学, 湖北 武汉 430068)

【摘要】人工智能技术引领的智能教育革命方兴未艾, 各类智能教育新应用为教育发展增值赋能的同时也隐匿着巨大的伦理风险, 不断侵蚀着人类教育几千年来所坚守的道德基础。研究从本体论、认识论、价值论三重哲学视角对智能教育的技术伦理问题进行本质还原, 探寻智能教育伦理风险的原点逻辑和治理之道。从本体论出发, 智能技术的本质是以符号世界替代真实世界, 这会引发盲目的科学崇拜, 只有保有对智能技术的理性与慎辨才能指向智能教育的澄明进路; 从认识论出发, 智能技术颠覆了主客二分的“人-技术”关系, 奇点超越后混乱的人机关系将导致不可测的道德危机, 只有在承认智能技术“类人”身份后赋予其道德观念和道德能力才能形塑智能教育的道德进路; 从价值论出发, 智能技术描绘了教育公平的美好愿景, 但其深藏的技术偏向可能把教育带向更深层次的歧视和不公, 只有依赖制度力量构建问责机制对智能技术应用予以监管, 才能促进智能教育的福祉进路。

【关键词】智能教育; 技术伦理; 本体论; 认识论; 价值论

一、前言

作为新一轮科技革命、产业革命、乃至未来几十年人类社会经济发展的重要驱动力, 人工智能技术近年来在一向“因循守旧”的教育领域同样掀起了一场智能教育革命。各类智能技术加持的教育新产品、新模式、新业态不断见诸于各级教育发展规划, 有些甚至已落地应用并初具成效。知识生产、流动和利用方式的颠覆性变革让每一个人都仿佛触摸到了未来智能教育新生态中教学效能大幅提升、学习方式随心所欲、优质教育资源普惠大众的美好画面。然而, 在智能教育革命汹涌向前的同时也出现了越来越多的争论和非议, 如智能校园监控妨碍人权

【基金项目】教育部人文社科研究项目“农村小规模学校信息化发展模式研究”(编号: 15YJC880094)

【作者简介】韦妙, 博士, 硕士生导师, 湖北工业大学职业技术师范学院副教授; 何舟洋, 湖北工业大学职业技术师范学院硕士研究生。

[1]，智能头环侵犯隐私^[2]、智能机器人教师扭曲学生情感认知发展^[3]、智能教育应用导致更隐蔽的“智能鸿沟”有违教育公平^[4]，种种反对的声音都指向了智能教育发展背后所暗含的巨大伦理风险。如何在享受智能技术红利的同时避免智能技术破坏人类教育几千年来所坚守的伦理道德基础成为未来智能教育发展必须重视和首要解决的问题。

每一次新技术与教育的结合和碰撞在表面上呈现出来的都是教育功能的扩展和效率的飞跃，实质上却是教育内在的文化、秩序及伦理发生了不可逆的改变。前者“见效快”极易唤起人们对技术革新的热情而被大肆宣扬和追捧，后者不易感知往往被淹没在技术狂欢的喧嚣之中。人工智能对教育“功用性”的提升幅度远超以往任何一种技术，其必然对现有教育价值秩序和伦理道德带来前所未有的冲击。技术的“可用性”和“隐蔽性”往往相互交织，人工智能的“功用”越显著，其对教育伦理秩序的影响就越隐匿。正如哲学家鲍尔格曼（Albert Borgman）所指出的“技术进化的过程导致人们在获得功能的同时失去了与‘与境’的交互和参与，这必将导致物与‘与境’的剥离、手段和目的的割裂^[5]。”

因此，思考智能教育潜在的伦理风险必须摒弃技术功能和应用层面的干扰，回归事物缘起的哲学本质：从本体论层面追问“智能技术的本质是什么”，从认识论层面追问“智能技术与人的二元关系是否已改变”，从价值论层面追问“智能技术能否忠于教育的终极价值”，以此为基础才能真正揭示智能教育技术伦理风险形成的内在规律和可能的治理之道。

二、智能技术的三重哲学反观

在哲学的基本问题中，本体论问题具有根本性的意义，起着奠基的作用，认识论问题围绕本体论展开，构成了主客体二元关系的基本论述，而公平正义作为终极价值问题的理性追求正是本体论和认识论这两个维度的落脚点，也是哲学所指向的终极目标。其中，本体论是一元的，回答的是我们所生活的世界上最真实的事实是什么样的，旨在从事物和世界的本质上确立概念的终极存在；认识论是二元化的，回答的是人和真实事实之间的关系是怎样的，旨在确立概念自身所展开的主客体关系，为世界的解释模式；价值论是多元的，回答的是现实世界中价

值关系的运动变化规律及主体的价值取向，旨在确立公平正义为客观世界必然进程的内在目的和理性所追求的终极价值^[6]。只有从这三个哲学维度对智能技术进行反观，才能跳出固有的认知框架，从根源上厘清智能技术的伦理问题。

（一）本体论反观：智能技术是真实映现世界的吗？

作为哲学思考的经典理论形态，本体论诉求于客观理性和客观知识的理论思辨。技术本体论于技术哲学而言处于基础性地位，哲学的技术转向表明技术开始成为了哲学的核心问题，所谓技术转向是要从存在论的维度来阐释技术、理解技术^[7]。智能技术在我们的生活世界中的本质究竟是什么？其形而上（Metaphysics）的定义是什么？智能技术的“智能”和人的“智能”有区别吗？这些本体论追问会横亘在每位技术哲学家面前。

在现象学的视角中，智能技术不再只是单纯的、赤裸裸的工具，而是一种生活方式，是现代世界构成的主要环节。智能技术是根据科学和数学来形塑的，是通过实验、因果方程式和符号关系所维系的逻辑。例如，数字孪生技术从物理空间、云端服务和数字空间三个逻辑层次进行结构建模，从而构建出具备可定量特征的“客观”世界，数字化模型越精确，该世界就越“客观”。这也恰恰印证了胡塞尔（Husserl）所说的技术创造出一种“自主（Eigenständige）的绝对真理”^[8]，它用纯符号逻辑的操作取代“可见的”几何图形。智能技术在数学框架内定量化地对本质进行解释，又把真与善、科学与伦理的关系逐渐分离开来。此时，人们认为智能技术似乎映射着科学真理。

从本体论而言，“科学-智能技术”被熔接成一种新型的社会控制形式，其本身的运行扮演着“真理”，但究其根本，智能技术背后的数学科学作为观念面具（Ideenkleid）代表着一种既表现生活世界又粉饰（Vertritt）生活世界的符号^[9]。智能技术的符号运算就能代表绝对正确和绝对真理吗？智能技术能完美地映现真实世界吗？斯蒂恩·马克（Steen Marc）在《打开黑匣子却发现它是空的：社会建构主义和技术哲学》一文中提到，技术设计实践中内在的伦理品质往往是隐含的和不能检验的^[10]，因此我们不能按照终极因（Final Cause）来科学地设想智能技术，认为智能技术的数理逻辑完全摆脱了生活世界中的不确定性和特殊性。因此，为了使人在使用智能技术时不迷失自我，无论智能技术作为观察、测

量和计算的中心主体多么重要，它都不能作为伦理、审美或教育的完全行为者来发挥科学作用。

（二）认识论反观：人如何与智能技术交往？

智能技术是人创造的产物，从现象学的视角分析智能技术与人的关系是极具复杂性的，这种复杂性突出体现在智能技术兼具自然属性与人工属性。随着技术人（Technological Man）的出现，技术像人一样通过改变自己或环境的方式对某个刺激做出反应，具有自主行动力、自主决策力和自然反馈机制，即认为技术存在某种智能。智能技术对于人而言有什么关联和超越？从认识论反观智能技术就是要遵循自然规律和社会发展规划这两个方面，从“人-技术”二者的关系对智能技术进行判定。

任何一种技术的进展都不仅仅局限在技术上，随着仿生科学的不断进步和发展，在造出一个与人类极为相似的“身体”时，它还表现在文化观念的适应和调整^[11]。德福雷斯（Dreyfus）对海德格尔（Martin Heidegger）关于技术的论述进行了创造性解读，提出智能技术的“具身化”方向，即智能技术具有一个“身体”^[12]。后认知主义科学认为“意识”并不存在于身体和大脑，而存在于具体的交互活动中。因此，“具身化”的智能技术作为一种“类人”的智能，是对人类生存状态和生活世界的功能性模拟，人们并不能笃定地排除其不具有自我意识。由于“算法黑箱”^[13]的限制，我们不禁发问，技术体是否有“意识”且知道自己在做什么吗？究竟要如何定义智能技术的身份角色？智能技术的意识活动会超越人类吗？

从认识论而言，赛博格（Cyborg）的出现模糊了自然物与人造物之间的界限，打破了物理世界与非物理世界之间的壁垒^[14]。智能技术不完全是一种人工机器，这种无法归属技术身份的溢出现象，正是技术革命导致文化调适过程中产生纠结心理的真实写照。人与智能技术在交往的过程中，智能技术不再是主客二分的客体，而是处于人和机器“共生关系（Symbiosis）”下具身化、情境性的主体，对人与智能技术的关系向“相互依赖”“相互渗透”“相互嵌入”的演变臻于成熟^[15]。当智能技术逐渐进化为人类历史上前所未有的“自主性”技术维度，成为反思性人工智能（Artificial Reflexive Intelligence）时，其便脱离了他者

的支配，具备了自主反思力和自我修改完善程序的能力，达到或超越笛卡儿“我思”的标准，届时人类社会将不再是自文艺复兴运动后以人为中心的一元主体世界，而将演变成双元或多元的主体世界。当人类无法揣测智能程序的行为意图，使其脱离了人类的掌控时，将会带来无法想象的伦理灾难。因此，人类亟须在智能技术实现奇点超越之前给它限定责任范围和行为规范。

（三）价值论反观：智能技术能否忠于人类价值需求？

价值来自于事物系统复杂的目的性，从不同的视角来看，价值的内涵和导向不尽相同。在哲学视域下，价值论旨在确立追求以公平正义为内在目的的终极理性价值。人对智能技术越来越高的依赖程度必然成为未来智能社会发展的基本走向。智能技术的价值取向是怎样的？智能技术能否忠于人类终极的理性价值？只有这些价值论的追问得到确定回答，人类才能将智能技术导向公平正义的发展进程。

正如赫伯特·马尔库塞（Herbert Marcuse）所分析的社会现象——技术嵌入多层次规则后会建立起一个带有偏见的运行体系，即技术和统治是合理性和压迫性特有的融合^[16]。卢梭（Rousseau）的《论不平等》认为形式的偏向隐藏在技术体系的各个方面，也就是说在科学和技术的物质先验论中（In Materiellen Apriori）潜藏着一种阶级利益和历史状况所决定的技术设计^[17]。人们通常认为偏见是一种对公平的偏离，技术的偏见是指在不顾及个人感受的情况下，将相同的标准用于所有人。与普通技术相比，智能技术往往带有更微妙更隐蔽的偏向形式，智能技术的不可解释性导致即使其在设计时被蓄意以牺牲某一群体利益为代价而使另一群体获益，使用者也往往无从察觉。未来社会中，智能技术可能以一种极其隐蔽的方式偏袒特定社会群体的特权利益而侵害普通用户的合法权利。

从价值论而言，技术偏向通常在传统的技术认知中以“技术中性论”^[18]无伤大雅地体现出来，不被人们所重视。智能技术却是将无意识的偏向暗自隐藏在助长偏见的程序设计中，使用智能技术的一方或双方往往由于经济利益、不诚实的主张、不合理的恐慌、种族或性别偏见以及科学和公众人物的腐败而被不公平地对待。如果智能技术在不知不觉中变成了无耻行动者实现价值歧视的阴谋、借口或工具，那么它的运用越广泛就会导致越多的偏见和歧视，最终侵蚀掉整个人类

社会公平正义的价值基石。

三、智能教育的技术伦理风险隐忧

时至今日，人在使用智能技术的过程中，往往只能感知到自身施加给技术的影响，却很少体验到技术的反作用，这就带来了一种主体无视关联性（Connectedness）的技术幻觉。这种幻觉向人们保证，人们在运用智能技术对世界进行改造时不会引发后果。在智能教育领域，诱人的技术幻觉一开始就掩盖了人们对技术潜在伤害的关注，等人们最终发现并重视智能技术潜在的伦理风险时可能已悔之晚矣。

（一）符号世界的预设：智能教育的科学崇拜

智能技术的本质就是以数学和科学为基础，将参数和模型不断精确和细化到可以被理论严格描述。智能技术通过符号建构客观世界，把客观事实压缩放入符号体系中，让客观事实在符号世界中呈现意义，最终无遗漏地刻画出世界的物理模型。这意味着对物理世界最深刻的描述不再包含任何人为的不确定的参数，届时物质世界的终极模型将与符号世界完美同构。就此意义而言，智能技术将会把世界本身内化为一个庞大而精密的符号世界，对符号表达和数字计算的信任和依赖必然引起人们对智能教育的科学崇拜。

智能技术所带来的教育效能具有无可辩驳的合理性，在社会生活的所有层面都能得到认可和尊重^[19]。在智能教育环境下，一切都强调从社会需求的角度出发来考虑问题，智能技术的应用被概念化为用标准的技术范畴对所处理的教育问题进行社会解释，如“可操作的和不可操作的”“有效率的和无效率的”。“效能的颂歌”成为控制的辩解和托辞，谁敢质疑智能技术进步作出的教育贡献和建立的科学崇拜呢^[20]？卢德派成员和其他“浪漫主义者”的呼吁很容易因为现代科学技术的压倒性成果而被忽略，此时的智能教育在符号世界的构建中义无反顾地走向盲目的科学崇拜。比如数字孪生校园、全息智能教育空间及数智融合等智能教育应用，它们无不凸显智能技术在物理世界与数字世界之间的镜像化关联，将物理实体对象的数字模型通过仿真和数据分析进行实时感知和诊断，预测物理实体

对象的状态并通过性能优化发送指令，进而调控物理实体对象的行为^[21]。在数智融合技术驱动下，教育中的数字画像走向数字孪生体，学生、教师和教育中介等成为数字化、镜像化的符号^[22]，这种以效能为主导的智能教育理念暗中助长着人们的技术崇拜，不断将符号世界中的实验方程、算法逻辑以及因果关系简单嵌套到真实的教育物理场域中来。

从本体论视角出发，智能技术根据科学和数学来形塑和同构教育，使其在客观性方面变得越来越依赖技术。人们依靠智能技术追求在教学实践中获得技术预设的“效能”，从而建构起一个绝对的教育观念实在，即对智能教育“理想化”的技术理论。例如，解读和研究人类大脑内部学习机制的脑机接口技术（Brain-Computer Interface, BCI），它能够不依赖外围神经和肌肉组成的通信通道直接让大脑与外界进行交互，来分析学生的学习状态、注意力水平和学习风格等^[23]。这样的数据分析结果必定不可能是完全可靠的，人作为独创性的主体不应被剖解为符号数据去代入科学技术逻辑。盲目的技术崇拜与不理智的宗教狂热无异，我们需要一种更加理性的态度去看待智能技术背后的算法体系和所谓的教育“效能”。

（二）人机关系的重构：智能教育的奇点超越

与以往的技术不同，智能技术作为一种“类人”的智能，与人的智能在形式上具有同构性、在逻辑上具有顺承性、在行动效果上具有统一性。这预示着智能技术组织模式的发展能淡化主体性智能思维与意识的社会历史根基，其超强的模仿力能消解智能生物基础的唯物主义限定。人与智能技术之间的关系从人与非人的二元论关系，逐渐发展成人与智能技术相互生成的“共生”关系，不再存在绝对的鸿沟和界限。

“人-技术”关系的转变建构了一种人类社会必然走向智能技术奇点超越的观念，从技术逻辑的视角预设了未来教育的“人-机”双主体共存的情形。一个在教育领域著名的案例是给初等几何设计证明程序，马文·明斯基 (Marvin Minsky) 和国际商业机器公司 (International Business Machines Corporation) 设计了一个被称为“驴桥” (Pons Asinorum) 的程序^[24]，它得出了史无前例的证明方法，最令人惊讶的是这个“新想法”是由程序得出的，而不是设计者给出的。

这是否可看作是技术拥有了作为主体的创造性思维呢？此时证明方法的功劳到底应该归于设计者还是程序自身？技术无意识论认为，程序本身没有创造力，这种表面的创造性可能是通过不同符号或符号串的组合意外得出来的结果，或是设计者深藏的潜意识想法被程序算法呈现出来了。但是，如果人类在对智能教育的设计初始没有给技术赋予“创造性”指令的话，我们就不能排除智能技术通过“控制——反馈”循环机制的不间断改进和更新对符号串进行重组和计算，进而生成“创造力”的可能性。此时，智能教育产品的设计者和使用者的大脑与智能技术的“心智”必须要区分开来，程序的设计归于人类，但程序产生的新的教育理念和方案却不能完全归于设计者，设计者是“元作者”，智能技术则是“作者”^[25]。当智能技术完成奇点超越，拥有了“类人”的思维，那么教师在面对智能技术时就不能简单的将其判定为“机器”，想当然的忽视智能技术可能的“自主”行为。

人工智能奇点论宣称第一台超智能机将是人类最后一个发明^[26]。从认识论视角出发，拥有类人性的技术作为一个不依赖生物代谢的超生物体，是独立于社会历史经验的自治系统，奇点的超越在逻辑上是必然的，在现实上是可能的。在技术乐观主义者畅想的未来教育场景中，达成奇点超越的智能技术可以依据人类预设的教育属性和目的对自身的操作系统进行自主选择及更新，从而对学生进行自主教育。至于如何去教、怎样去教、该教什么等问题，这些不再是人类需要考虑的范畴，人类要做的仅仅是设定教育目的即可，因为智能技术强大到能够演算和运行出趋于完美的教学逻辑，甚至会超越人类教师的大脑智慧。但是，当智能技术拥有了不同于人的超生物特性和逻辑自治系统时，人类又怎敢轻易将教育权完全让位于智能技术呢？如果遇到“师德败坏”的机器人教师该怎么办？智能技术在未来教育中的身份认定和权责归属都将成为极其棘手的问题，更令人担忧的是人类目前甚至都无法判断智能技术的奇点超越会在何时何地发生。

（三）教育公平的“神话”：智能教育的价值歧视

智能教育产品及应用由于其共享性和自动化设计经常被认为能有效缩小各类教育差距，在教育领域助力人类实现对公平正义终极价值的追求。智能技术真的能在未来教育发展中创造出教育公平的“神话”吗？公平存在的地方也是不公

平现象时常出没的地方，公平与不公平是在价值秩序中共生的。公平的价值属性无疑使教育场域变得更加复杂，教育公平和不公平的界线很难确定^[27]。人们寄希望于技术促进教育公平的同时，却也可能陷入一种符号化的公平陷阱中。智能技术看似合理公正的数据、计算、量化等会干扰人们对教育公平的真实判断，所营造的教育公平愿景背后可能隐匿了更多的价值歧视，带来更深层次的教育不公平。

马克斯·韦伯（Max Weber）将技术的偏向性划分为实质的偏向和形式的偏向两种类型^[28]。在智能教育时代，实质的偏向是直接建立在智能技术中的不平等，它着眼于技术的内容和目的偏向。在智能时代，操纵人的不是技术，而是资本，资本的好恶即代表了技术实质的偏向。算法程序的背后是资本以增加用户粘性来牟取利益的商业逻辑，当人接收的信息内容再也逃不过算法有目的的推荐机制时，人就被带向了智能技术的实质偏向。自文艺复兴后，人文主义号召人类“倾听自己内心的声音”，而现在甚至未来，人们却被灌输“信服算法的痕迹”等理念。算法会向受教育者推送其感兴趣的内容，但同时也会放大最具分裂性和情感操控性的内容。在一次又一次的反复推送中，受教育者固有的观念被不断强化，其能听到理性且全面的声音将会越来越少，最终被困在了智能算法编织的“过滤气泡（Filter Bubble）”中^[29]。当受教育者接收的信息越来越单一，学习路径越来越窄，观点日趋偏激和狭隘时，所谓的开放性、自主化、个性化的智能教育反而可能导致受教育者整体的价值偏见和歧视。

智能技术的形式偏向是现代社会独一无二的存在。这种偏向着眼于智能技术本身在形式和手段上的偏向，其本质是指在时间、地点和由相对中性要素组成的系统等方面的引入方式上具有偏向的选择^[30]。智能教育的形式偏向是由技术运行体系表面的公平掩盖体系形式中暗含的不公平，在客观运转上与社会现存的价值歧视不存在直接关联，而是聪明地为自己披上了科技的外衣，即使稍显端倪也会因“算法黑箱”而无迹可寻或无据可查。智能教育的形式偏向往往暗藏在智能技术的筛选设定中，在不知不觉中就完成了对待不同人群的区别化对待。例如，智能技术对学生进行“大数据画像”建模、分类和身份标识的过程其实就是给学生烙下暗含技术偏向的刻板印象的过程^[31]。教师运用学生“大数据画像”辅助教学时自然就形成了教育中的偏向性认识。智能技术的形式偏向将导致智能教育中更

难以察觉的价值偏见和歧视。

从价值论视角出发，智能技术能够促成的教育公平只能停留在“技术公平”的理想层面，智能技术的实质偏向和形式偏向投射在现实层面则意味着更多隐性而深远的教育不公平。

四、智能教育的技术伦理治理进路

当智能技术被视为纯粹的工具时，人们便更倾向于以技术的视角去表达和看待所有的教育问题，越容易偏信智能教育中出现的问题仅仅是因为技术层面，认为只要技术持续进步就自然会有新的解决问题的办法。但是，技术红利永远和技术风险共生，技术进步也许不一定是人们原本期望的那种“进步”。正如雅克·埃吕尔(Jacques Ellul)提出：技术进步本身的性质是模糊的，不能评判技术进步本身是好或是坏，每一次技术进化在给生活世界带来福利的背后总伴随着新风险的生成^[32]。随着智能教育技术的迭代增强，其潜在的负面效应也会不断放大，如何在预见智能教育技术伦理风险的基础上寻求治理之道，才是避免人们在“技术幻觉”操控下对智能教育技术偏信和滥用的唯一出路。

（一）理性与慎辨：智能教育的澄明进路

在本体论层面，智能教育背后所构建的符号世界，使人们达到了在实际教学经验中达不到的东西——算法的精确性，利用智能技术实现教育的理想算法由此成为可能，使得人们自以为掌握了绝对的“教育恒等式”。智能技术让观念世界与物质世界相协调，使人们能够去预测和规划教育规律，拥有在教学实践中所期望的“先见之明”，通过将虚拟的符号世界与现实的物质世界挂钩建构起一个绝对的教育观念实在。当人们在教育过程中运用智能技术进行科学计算、在具体的教育实践中进行符号构建时，技术将会把教师、学生和教育中介等各种元素抽象概括为具有精确性和普遍有效性的质的量化，这便使智能技术在有目的的实际教育情境中对教育有了预期（Voraussehen）和谋划（Vorhaben）。在这种教育预期和谋划中，普遍的可定量性是掌握教育规律的前提，算法的精确性是产生教育预期的必要条件，此时智能教育就变成了一种精神和物质的纯工具手段。然而，

人作为独立鲜活的人并不能被完全抽象为数学符号，教育中也并不存在纯粹合理的科学秩序，智能技术所构筑的工具性的世界无法涵盖教育的全部内容。人只有保有这种理性，对智能技术慎思明辨，才能为智能教育发展开辟一条澄明之路。

首先，人们不能任由智能技术对教育过程中出现的人或物随意地进行表征，要尊重人的主体价值性和主观能动性，要深刻意识到智能技术意图通过数学上的简化模型来开展教育仅仅是对人脑错综复杂的化学、物理信号交互的低层次复现^[33]。单纯地将鲜活的教育现象符号化、抽象化并不能真正认识并解决教育问题，要按照教育本来的面目和规律去认识和改造它。其次，不能因为教育过程的复杂性和多变性就在技术算法中舍弃或遗漏某些教育因素，智能技术的算法程序既是“过滤器”，也是一把“无形的枷锁”，这易使教师囿于程序中已设定好的内容去认识了解学生，受到技术算法的遮蔽^[34]。要用不确定性思维来指导智能技术的教育应用，挖掘和关注教育中的各种不确定因素，从而破除技术的遮蔽^[35]。最后，要用现象学的思维突破自然主义思维、现成性思维，这两种思维把智能教育的科学合理性视为“理所当然”，从而不予置评，在不自觉中坚持了思维定式。教育家们要利用现象学中“悬搁”的方式，打破这种“理所当然”，进而对智能教育进行本质还原，找出技术预先被给予的意义结构和世界构成。

当然，人们不能全盘否定智能技术在教育应用中的算法优势，辩证看待并加以利用才是对智能技术伦理治理的理性态度。如果说人类的教育经验构成了一个地平面的话，那么智能教育的技术科学就是在这个土壤上长成的大树，这棵树只有从土壤中吸取营养，才能长出四面延伸的新枝杈。智能教育的技术科学这棵树之所以能够长这么大，正是因为它根植于人特殊的教育实践活动中，只有扎根越深，才能挖掘出智能教育领域数据算法里人的交互、人的教育思想价值、人的教育经验与实践，进而打破对智能教育盲目的技术崇拜，冲出智能教育对“效能颂歌”的狭隘追捧。

（二）身份与责任：智能教育的道德进路

在认识论层面，智能技术意味着技术算法具有自我保存、更迭和进化的理性需求和行为能力，也意味着技术算法具身化的完全体能够产生“自我意识”。海量的数据是其自主性的根基、精准的算法是其自主性的体现、理性的决策是其自

主性的核心，人与智能技术正在形成一种“人机共生”的新型关系^[36]。

当具有自主性的智能技术应用于教育领域时，首先要改变之前对智能技术“外在客体”的刻板认识，对智能技术的“类人”身份进行认同。这是教育者应秉持的对智能教育新型“人机关系”最基本的认识，在此基础上师生才能和人工智能体和谐共处、共生发展。其次，人类应该赋予智能教育技术道德判断力和道德行动力，在“人机共生”的智能教育环境下将其塑造为具备基本道德素养的“类人”，相反，缺乏道德的“类人”教育者就要被人类拒绝甚至停止“生命”。最后，要想满足人类对智能教育的道德期待，就需要将人类深邃的教育道德观嵌入到智能技术体中。在人工智能体辅助教师教育的过程中，既要追求道德判断力和道德行动力，还要保证其伦理敏感性、安全性和可靠性，使之更好地服务于教育实践。詹姆斯·摩尔(James Moor)将人工道德智能体划分为四类，依次是伦理效果智能体、隐含式伦理智能体、显现式伦理智能体和完备伦理智能体^[37]。这说明人类赋予人工智能体道德能力的过程是渐进性的，将人类的教育道德观嵌入到智能技术体的过程会有一个漫长的过程，我们最终所追求的完备伦理智能体是让技术通过遵循“道义逻辑”拥有和人类一样的道德意识，建立起设身处地为人类着想的道德教育观。

如何让智能技术最终走到完备的伦理智能体这一形态呢？阿明·格伦瓦尔德(Armin Grunwald)认为，科学技术进步必然会出现种种技术道德失范和责任归属不明的问题^[38]。现实生活中，人类会被赋予各种复杂的教育责任，出现任何教育问题都要担责。那么，既然承认了智能技术的“类人”地位，就应权责一致，智能技术也要与人一样“类责任化”，即承担和人一样的教育责任，这样才能有效促成智能教育的道德成长。温德尔·瓦拉赫(Wendell Wallach)在《道德机器》一书介绍了“自上而下”和“自下而上”的技术伦理原则，“自上而下”是指在设计教育伦理程序时，选取具体的伦理理论以保证后续算法和子系统的执行；“自下而上”的模式是指智能技术体在模拟的教育场景中所生成和创造的教育道德代码^[39]。在实际的智能教育场景中，“混合式”的道德进路更能结合上述两者的优点，即通过普遍的教育伦理规则与具体鲜活的教育道德情境双向发力，从而使人工道德智能体生成特定的教育价值观，既保证了基本教育伦理理论的吸收和真实案例的灵感需求，又能在教师、学生、技术开发设计者和学校管理者的

多方参与中，对智能教育的道德发展进路进行导引和形塑。

（三）公平与法治：智能教育的福祉进路

在价值论层面，智能技术赋能教育的最终意义都应指向实现“将一切知识教给一切人”的教育终极价值。然而，智能技术高效能的背后却藏匿着不同往日的技术偏向性，如若不加以行之有效的伦理规范予以限制和规训，必将导致智能教育发展背离教育公平的初衷。

技术算法可能会犯错，甚至带有偏见，但由于“算法黑箱”的复杂性及不透明性却很难被教师、学生、家长或教研设计者审查发现。因此，智能教育技术应用首先要保持算法透明性。迪亚克帕罗斯(Diakopoulos Nicholas)、科利斯卡(Koliska Michael)在《新媒体的算法透明性》中的算法透明性框架就提出要杜绝智能教育技术中的算法垄断^[40]，智能教育技术的设计和开发者要保持透明操作，建立和完善人工智能的透明性规范，秉持教育的公平性，让不同知识背景的主体都能理解并接受算法。其次，要进一步强化智能教育技术应用的算法可解释性。蒂姆·米勒(Miller Tim)等把可解释性称作人工智能算法输入的某些特性引起某个特定输出结果的原因^[41]。美国国防部下属的国防高级研究计划局(Defense Advanced Research Projects Agency)一直致力于可解释的人工智能(Explanatory Artificial Intelligence)^[42]，该研究正力图开发新的机器学习系统，该系统将能够把模型转换为可理解的、对用户有用的解释。因此，技术和相关行业管理机构要鼓励和统筹可解释人工智能的研究人员、行动科学研究人员与教育主体之间的合作，为可解释的智能教育模型设计提供助力，如此才能防范智能教育产品应用中隐含有不公平、歧视性的信息。最后，要想解决智能教育潜在的价值歧视问题不能只是技术部门的内部约束，更需要外部力量的强力监管^[43]。技术的透明度和解释性只是实现算法伦理和问责制的一种方法，智能技术的价值判断力归根结底取决于智能教育设计团队。利益驱动致使出现了部分技术精英或资本力量人为控制智能技术、侵害普通人权益的恶性事件^[44]，算法程序打着建立个性化推荐机制的幌子越权读取甚至出卖用户信息及隐私来牟取私利，需要警惕。鲁宾·宾斯(Reuben Binns)等提出规避算法偏差必须提倡问责与追责^[45]，关注教育数据、算法持有者和师生使用的主体责任。因此，在智能教育产品发布

之前，应由政府主导，各方公众代表参与到“公民陪审团”或“混合治理平台”以评估拟议智能技术应用的价值预期，给予教育工作者对智能教育产品的质询权和建议权，鼓励其参与和监察合作设计过程，主动干预算法，消除智能教育产品中潜在的歧视性或者偏见性因素。在智能教育产品发布后，如果出现了损害普通消费者教育权益的不良后果，政府应根据算法的透明性开发原则追溯相应的主体责任，做到追责时“有迹可循”^[46]。除此之外，政府还可以通过在智能教育领域提供专门法律和规章条例来约束智能教育企业、设计者及其他主体的行为规范，问责时“有法可依”，以法治的力量确保智能教育的公平福祉。

五、结语

近一个世纪前，阿道司·赫胥黎（Aldous Huxley）就在《美丽新世界》一书中幻想了未来社会利用智能技术给婴儿灌注思想的睡眠教育场景^[45]，今日读来仍发人深省。如果人类沉溺在智能技术构筑的“完美教育”中不可自拔，任由技术接管全部教与学的工作，那么思想、情感、道德等这些人类教育最为珍视的追求也必然会被技术控制，人性之自由将在机器的轰鸣声中灰飞烟灭。人类必须警惕智能教育发展背后巨大的伦理风险，预见性地创设技术伦理原则、操作规范以及约束机制。智能技术到底是赋能教育还是奴役教育，或许只在人的一念之间。

[参考文献]

- [1] 陈畅. 人脸识别进校园引发争议:对“人工智能+教育”应持审慎态度[EB/OL]. 光明网. (2019-10-10)[2021-09-01]. https://legal.gmw.cn/2019-10/10/content_33220501.htm.
- [2] 郝孟佳. 小学生戴“头环”防走神? 教育部门: 责令暂停使用[EB/OL]. 人民网. (2019-11-01)[2021-09-01]. <http://edu.people.com.cn/n1/2019/1101/c1053-31432543.html>.
- [3] 李华锡. 课堂人脸识别让孩子变“戏精”? “AI+教育”伦理难题待解[EB/OL]. 中国青年网. (2019-05-02)[2021-09-01]. http://edu.youth.cn/jyzx/jyxw/201905/t20190502_11942636.htm.
- [4] 邱晨辉. 人工智能会对教育造成多大的颠覆? [EB/OL]. 新华网. (2017-12-04)[2021-09-01]. http://www.xinhuanet.com/tech/2017-12/04/c_1122051525.htm.
- [5] 吴国盛. 技术哲学讲经典读本[M]. 上海: 上海交通大学出版社, 2008: 412.
- [6] 王治东. 人工智能研究路径的四重哲学维度[J]. 南京社会科学, 2019(09): 39-47.
- [7] 吴国盛. 技术哲学讲演录[M]. 北京: 中国人民大学出版社, 2009: 165.

- [8] Husserl E. Ideas: General Introduction to Pure Phenomenology [M]. London: Routledge; Reprint, 2010: 234-260.
- [9] Foucault M. The ethic of the care for the self as a practice of freedom: An interview with Michael Foucault on 20th January 1984 [J]. Journal of Materials Chemistry, 1987, 8(11): 2507-2514.
- [10] Steen M. Upon Opening the Black Box and Finding It Full: Exploring the Ethics in Design Practices [J]. Science Technology & Human Values, 2015, 40(3): 389-420.
- [11] 王峰. 人工智能: 技术、文化与叙事 [J]. 上海师范大学学报(哲学社会科学版), 2019, 48(04): 97-103.
- [12] Hubert L. Dreyfus. What computers still can't do: a critique of artificial reason [M]. Massachusetts: The MIT Press, 1992: 56.
- [13] [美] 布莱恩·阿瑟. 技术的本质: 技术是什么, 它是如何进化的 [M]. 曹东溟, 王健译. 浙江: 浙江人民出版社, 2014: 26-50.
- [14] 朱彦明. 后人类主义对教育的挑战与重塑 [J]. 南京社会科学, 2018(11): 129-136.
- [15] 余聪, 张亮. 人工智能时代意识形态批判的哲学反思——“人机关系”视野中的批判向度 [J]. 浙江学刊, 2021(02): 4-11.
- [16] Marcuse H. One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society [M]. Boston: Beacon Press, 1991: 3-10.
- [17] [法] 卢梭. 论人类不平等的起源和基础: A discourse on inequality [M]. 高煜, 译. 广西师范大学出版社, 1958: 157-160.
- [18] 吴致远. 有关技术中性论的三个问题 [J]. 自然辩证法通讯, 2013, 35(06): 116-121+128.
- [19] [34] [美] 尼尔·波兹曼. 技术垄断: 文化向技术投降 [M]. 何道宽, 译. 北京: 北京大学出版社, 2007: 153.
- [20] Chant C. The Mantra of Efficiency: From Waterwheel to Social Control [J]. Journal of design history, 2010, 23(2): 211-213.
- [21] 田锋, 段海波, 杨振亚等. 数字孪生体技术白皮书(2019) [R]. 北京: 数字孪生体实验室, 2019: 1-135.
- [22] 郑浩, 王娟, 王书瑶, 顾雯. 认知数字孪生体教育应用: 内涵、困境与对策 [J]. 现代远程教育, 2021(01): 13-23.
- [23] 任岩, 安涛, 领荣. 脑机接口技术教育应用: 现状、趋势与挑战 [J]. 现代远程教育, 2019(02): 71-78.
- [24] Minsky M. The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind [M]. New York: SIMON & SCHUSTER, 1988: 80.
- [25] 涂良川. 马克思历史唯物主义视阈中的人工智能奇点论 [J]. 东北师大学报(哲学社会科学版), 2020(01): 78-84.
- [26] [美] 雷·库兹韦尔, 奇点临近 [M]. 李庆诚, 等译. 北京: 机械工业出版社, 2011, 116-117.
- [27] 王鑫, 熊和平. 教育公平为何是个“假问题”: 价值哲学的视角 [J]. 教育发展研究, 2018, 38(24): 9-13.

- [28]王琴,罗甜田. 马克思与韦伯资本主义合理性批判的理论路径比较[J]. 四川轻化工大学学报(社会科学版), 2020, 35(03):55-72.
- [29]Eli Pariser. The Filter Bubble: What the Internet Is Hiding from You[M]. London: Penguin Press, 2011:50-55.
- [30]Feenberg A. Critical Theory of Technology: A Companion to the Philosophy of Technology[M]. New York: John Wiley & Sons, Ltd, 2009:90-106.
- [31]孟翀,王以宁. 教育领域中的人工智能: 概念辨析、应用隐忧与解决途径[J]. 现代远程教育, 2021(02):62-69.
- [32]Jacques Ellul. The Technological Order, in Philosophy and technology: Readings in the philosophical problems of technology, edited by Carl Mitcham and Robert Mackey[M]. New York: The Free Press, 1983:86.
- [33]靖东阁. 人工智能时代教育研究的计算主义及超越[J]. 电化教育研究, 2021, 42(02):18-24.
- [35]白蕴琦. “互联网+教育”推进的不确定性与应对研究[J]. 现代远程教育, 2020(04):19-26.
- [36]韦妙,何舟洋. 技术现象学视域下人工智能对教师角色的重塑[J]. 电化教育研究, 2020, 41(09):108-114.
- [37]Moor J H. The Nature, Importance, and Difficulty of Machine Ethics[J]. IEEE Intelligent Systems, 2006, (24):18-21.
- [38][德]阿明·格伦瓦尔德. 技术伦理学手册[M]. 吴宁译, 北京: 社会科学文献出版社, 2017: 156.
- [39][美]温德尔·瓦拉赫, 科林·艾伦. 道德机器: 如何让机器人明辨是非[M]. 王小红, 译. 北京: 北京大学出版社, 2017:61.
- [40]Diakopoulos N, Koliska M. Algorithmic Transparency in the News Media[J]. Digital Journalism, 2017(7):809-828.
- [41]T Miller, P Howe, L Sonenberg. Explainable AI: Beware of Inmates Running the Asylum Or: How I Learnt to Stop Worrying and Love the Social and Behavioural Sciences[J]. Artificial Intelligence, 2017(12):10-17.
- [42]Goodman, B, & Flaxman, S. European Union Regulations on Algorithmic Decision-Making and a “Right to Explanation”. AI Magazine, 2017, 38(3):50-57.
- [43]赵磊磊,张黎,代蕊华. 教育人工智能伦理: 基本向度与风险消解[J/OL]. 现代远程教育:1-15[2021-10-08]. <https://doi.org/10.13927/j.cnki.yuan.20210831.005>.
- [44]金民卿. 网络论争背后的资本控制与技术依赖[EB/OL]. [2021-09-01]. <http://theory.rmlt.com.cn/2015/0720/395375.shtml>.
- [45]Binns R, Van Kleek M, Veale M, et al. 'It's Reducing a Human Being to a Percentage'; Perceptions of Justice in Algorithmic Decisions[EB/OL]. [2021-09-01]. <http://www.cs.ox.ac.uk/publications/publication11851-abstract.html>.
- [46]Diakopoulos N. Accountability in algorithmic decision making[J]. Queue, 2016,

13(2):56-62.

[47]Huxley A. Brave New World[M].New York:HarperCollins, first Perennial Modern Classics edition,1932:1-5.

Secret Worry and Governance Approach of Intelligent Educational Technology Ethics: Ontology, Epistemology and Axiology

WEI Miao, HE Zhouyang

(Hubei University of Technology,Wuhan 430068)

【Abstract】 The revolution of intelligent education led by artificial intelligence technology is in the ascendant, all kinds of new applications of intelligent education not only add value and empower the development of education, but also hide huge ethical risks, which constantly erode the moral foundation of human education for thousands of years. From the perspective of ontology, epistemology and axiology, this paper studies the essence of technological ethics in intelligent education, and explores the origin logic and governance of ethical risks in intelligent education. Starting from the perspective of ontology, the essence of intelligent technology is to replace the real world with the symbolic world, which will lead to blind science worship, only by maintaining the rationality and careful discrimination of intelligent technology can we point to the clear path of intelligent education; starting from the perspective of epistemology, intelligent technology subverts the "human-technology" relationship between subject and object, and the chaotic human-technology relationship that occurs after the singularity is exceeded can lead to unpredictable moral crisis, only by recognizing the "human-analogy" identity of intelligent technology and endowing it with moral concepts and moral ability can we shape the moral approach of intelligent education; starting from the perspective of axiology, intelligent technology describes beautiful vision of educational equity, but its deep technology bias may lead education to a deeper level of discrimination and injustice, only by relying on institutional forces to build an accountability mechanism to supervise the application of intelligent technology can we promote the well-being of intelligent education.

【Keywords】 Intelligent Education; Technical Ethics; Ontology; Epistemology; Axiology