

面向未成年人的人工智能技术规范研究： 学理阐释与探研进路

——面向未成年人的人工智能技术规范研究(一)

吴永和, 郭胜男, 许秋璇, 郑浩, 颜欢

(华东师范大学 教育学部教育信息技术学系, 上海 200062)

[摘要] 人工智能时代,如何协调好 AI 与未成年人之间的关系成为亟须解决的时代命题,而“面向未成年人制定什么样的 AI 技术规范”是回答好这一命题的关键。基于此,文章基于“面向未成年人的人工智能技术规范研究”国家社科重大项目,从学理上阐释面向未成年人的人工智能技术规范研究的核心内容与探研进路。文章由五部分组成:(1)面向未成年人的 AI 时代诉求;(2)面向未成年人的 AI 治理态势分析;(3)AI 技术规范概念阐释与定位;(4)研究的核心内容;(5)整体的探研进路。希望能够帮助读者理解面向未成年人的人工智能技术规范研究的重点和方向,期待能够为后续相关研究提供理论指导和参考,共同致力于解决 AI 与未成年人协调相融共进问题,助力构建人类命运共同体。

[关键词] 人工智能;技术规范;未成年人;核心素养;应用监管;安全伦理

[中图分类号] G434 **[文献标志码]** A

[作者简介] 吴永和(1969—),男,安徽巢湖人。博士,研究员,主要从事智能驱动的教育研究与教育数字化转型。E-mail:yhwu@deit.ecnu.edu.cn。郭胜男为通讯作者,E-mail:Shining_guo123@163.com。

一、引言

AI 技术作为引领第四次科技革命的重要技术之一,为社会建设和经济发展带来了重大而深远的影响。2019 年,我国科技部发布的《新一代人工智能治理原则》倡议发展负责任的 AI,协调好 AI 发展与治理的关系,确保 AI 安全可控可靠^[1]。2022 年 11 月,OpenAI 公司推出的以 ChatGPT 为代表的新一代人工智能工具,对社会和教育产生了广泛的影响。智能技术在教育领域中的应用政府高度重视,积极推动 AI 和教育深度融合,充分发挥其优势,促进教育创新变革,加快培养 AI 高端人才。而现阶段的未成年人是伴随着 AI 成长起来的一代人,正处于成长为具备创新能力和合作精神的 AI 高端人才的关键时期,承担着复兴中华民族的大任,因此,如何协调好 AI 与未成年

人的关系,保障未成年人健康成长是 AI 时代教育的重要命题。

面向未成年人的人工智能技术规范是 AI 时代确保未成年人健康成长的关键,是稳步推进 AI 技术创新发展和负责任的保障,也是我国落实未来 AI 发展战略的关键一环。为了深入推进面向未成年人的人工智能技术规范研究(后文简称:AI 技术规范)。笔者认为,首先要立足于 AI 技术发展的整体蓝图,深入了解未成年人的成长特征和 AI 技术发展的脉络,着眼于核心素养、应用监管和安全伦理,分别从人才发展、AI 向善和人机共融视角进行深度探析,最终落脚于 AI 技术标准规范制定,通过实践检验标准规范的可实施性。本文基于国家社科重大课题,从整体设计的视角阐释面向未成年人的人工智能技术规范研究的总体布局与探研进路,以期能够为后续相关研究提供理论

指导和参考。

二、面向未成年人的 AI 时代诉求

AI 时代,未成年人尚未成熟的理解能力、认知能力、辨别能力与心智发展等决定了其极易受到智能化环境的影响。因此,未成年人这一弱势群体对 AI 治理有了新的诉求。如图 1 所示,在 AI 技术发展蓝图的理论指导下,对面向未成年人的 AI 核心素养、应用监管和安全伦理进行研究,制定面向未成年人的人工智能标准规范,推进教育的数字化转型,促进 AI 技术与未成年人的深度融合与协调发展。在素养层面,一方面,培养未成年人正确对待 AI 的态度、知识等,防止受到不良信息侵害;另一方面,需要充分调动未成年人主动学习的积极性,通过个性化培养提升他们的智能创新能力,进而实现未成年人的全面发展。在监管层面,需要合理设计和有效约束 AI 产品,从根源上保障未成年人身心安全,实现智能应用的安全可控。在伦理层面,亟须技术伦理、社会伦理和隐私安全等方面的规范,保障未成年人的合法权益,推动 AI 与未成年人之间关系协调发展。因此,文章将从教育数字化转型、核心素养、应用监管和伦理安全以及规范需求五个方面具体阐述 AI 技术发展的时代诉求。

人工智能	未成年人
规范	AI技术规范体系协调人机融合、发展 通用体系
伦理	技术伦理 社会伦理 伦理隐私 人机协调
监管	合理设计 有效约束 信息保障 安全可控
素养	主动学习 个性发展 智能创新 全面发展
蓝图	AI与人才培养的深度融合、协调共进 数字化转型

图 1 面向未成年人的 AI 时代诉求

(一) AI 赋能教育的数字化转型

AI 技术是教育数字化转型的重要措施,借助智能化的方式重新整合各知识领域,集成多种新技术广泛联结的智能世界,推动教育的数字化转型。AI 为教育的数字化转型注入活力,通过资源设备的不断革新,以颠覆性的方式改变着学校的教学方式,决定了我国青少年的未来发展。国家教育“十四五”规划明确提出“教育高质量发展”的核心任务^[2],对学校建设、学生培养和教师发展等方面提出了更高的要求。AI 赋能学校建设人机协同和育人为本的未来学校,提升学校管理和决策水平;AI 赋能学生,促进培养全面、高素质、个性化创新人才;AI 赋能教师,释放教师的时间和精力;AI 赋能教学评价,实现精准画像和自适应智能教研。可见,AI 能够全方位赋能教育教学,赋智

家校互联,为推动教育数字化转型提供强有力支撑。

(二) AI 对核心素养提出更高要求

AI 技术对信息的智能化筛选,在传播和接收过程中能够帮助未成年人避免大量不良信息的侵袭。心智发展还未成熟的未成年人在信息良莠不齐的 AI 环境中,难以筛选辨别出正确、有效的信息,这些错误的不良信息影响到未成年人认知能力发展,对未成年人的核心素养发展提出更高要求。特别地,暴力信息对未成年人的影响深远持久,对犯罪行为具有长期潜移默化的“诱发效果”。例如,提高未成年人对暴力行为的赞成度,使得未成年人更容易产生暴力情绪^[3],对社会事件变得冷漠、缺乏社会责任感,甚至造成终身阴影^[4]。未成年人只有具备更全面的智能素养,才能在数智融合时代保障自身安全与权益,进而成长为高端人才。

(三) AI 给应用监管带来新挑战

智能时代,对 AI 产品的设计开发和应用进行监管,能够从源头上保障未成年人的心智发展。我国教育市场环境较为复杂,需要保证 AI 产品本身的健康安全,以保障未成年人基本权益。然而,未成年人缺乏足够的认知、辨别能力,极易受到不法企业、违法应用的诱导,泄露个人信息,给与不法分子可乘之机,在危害个人安全的同时还可能会给国家安全埋下重大隐患。从心理特征看,未成年人的好奇心强、自制能力差、渴望被认可,这无疑导致未成年人容易沉迷 AI 构建的虚拟世界。可见,AI 技术特性与未成年人的脆弱性均给 AI 应用监管带来新的挑战。

(四) AI 影响未成年人的伦理安全

AI 的可编辑性可能造成未成年人的伦理认知错误,对未成年人的伦理安全产生威胁。不仅因为 AI 本身算法、数据等方面的技术伦理风险,还因为未成年人的心理认知尚不成熟,过度信任与依赖 AI 产品,人、AI 与社会之间关系复杂性可能导致未成年人将人与 AI 的关系与行为规范相混淆,以与人相处的行为规范与 AI 相处,或是以与 AI 相处的行为规范与人相处,造成社会关系的紊乱、伦理认知的错误。因此,人类与 AI 的和谐共处是社会各领域都应当深入思考与研究的问题,是 AI 时代发展的必然诉求。

(五) 面向未成年人的 AI 技术规范需求

信息爆炸的 AI 时代,AI 技术凭借其跨时空、实时性和交互性的特点,拓展了未成年群体学习的空间和时间^[5],满足他们传统课程学习外的需求。事实上,AI 技术带给未成年人的影响具有两面性:一方面,在 AI 技术的赋能作用下,未成年人的求知欲和创造力得以激发,助力其个性化发展^[6];另一方面,AI 技术的

设计和应用给未成年人带来隐私、公平性、安全和伦理等方面的威胁。如果AI系统在设计过程中未能考虑潜在风险,潜在风险可能转变为现实威胁,给未成年人造成身心、情感和发展层面的伤害。简言之,作为“数字原生代”的当代未成年人亟须符合其权益和需求的技术规范,通过系统化培养未成年人的智能素养,加强AI应用监管力度,保障未成年人的伦理安全,进而实现AI与未成年人的深度融合与协调发展。

三、面向未成年人的AI治理趋势分析

近年来,联合国教科文组织发布了一系列报告文件,旨在促进AI与教育的融合并降低其潜在风险^[7-9]。在全球教育数字化转型的背景下,我国政府高度重视AI与教育的融合发展,出台了一系列青少年AI教育政策^[10]。在我国加速迈向智能化社会的进程中,迫切需要构建严密的AI治理体系,维护未成年人的健康成长^[11]。笔者认为当前围绕面向未成年人的AI治理主要体现了以下趋势:

(一)“技术伦理+社会伦理”,伦理规范持续建设

随着AI迅速渗透到未成年人学习、生活的各个场景,诸多学者从教育视角分析了面向未成年人的AI伦理困境与挑战,并提出相应的伦理原则^[12-13]。综合来看,当前面向未成年人的AI技术面临设计、算法、数据安全等方面的技术伦理风险,并可能会带来学生主体性缺失、师生角色失衡、教育艺术性消弭、数字鸿沟拉大等社会伦理挑战^[14]。近年来,我国政府在构建智能时代技术伦理秩序方面做出了巨大努力,在优化政、企、学界协同关系的基础上,组织成立了新一代人工智能治理专业委员会,并就AI治理原则、AI伦理规范以及全球AI体系框架的构建贡献了中国方案。但当前针对AI伦理规范的探讨多是从宏观通用的视角出发,尚未考虑未成年人心智不成熟、辨别能力低、社会经验少等特点,针对未成年人等弱势群体的伦理规范亟须完善。

(二)“政府立法+市场自治”,应用监管逐渐收紧

在树立AI伦理原则与准则的同时,还需要一套强有力的执行机制将AI治理落实到实践。在监管立法层面,以中国、美国、欧盟为代表的AI技术发展先进国家与地区都逐步收紧了对AI的应用监管。例如,我国近年来积极推动《个人信息保护法》《数据安全法》的有效实施,逐步建立对AI技术滥用与欺诈的处罚细则;美国强调监管的科学性和灵活性,致力于确保和增强美国在该领域的领先地位;继《通用数据保护条例》(GDPR)颁布之后,欧盟出台了全球首份综合

性人工智能法案,旨在构建严密的AI监管体系^[15]。在行业自治方面,伴随各国AI立法与监管的持续推进,国内外主流科技公司也逐渐加强内部AI伦理审查与规范。例如,微软、谷歌、商汤科技等公司纷纷设立了AI治理机构,负责企业的AI规则制定、案例研究、落地监督等AI伦理实践^[16]。尽管目前各国政府和企业推进AI技术应用监管方面作出了初步尝试,但专门针对面向未成年人的AI应用监管还较为缺乏。

(三)“通用规范+细分标准”,标准制定更为全面

为推动落实AI治理要求,包括行业协会、标准化组织、产业联盟等机构在内的行业组织较早地开展了AI治理相关研究,积极制定AI技术及产品标准。ISO/IEC JTC1在可信赖AI标准建立方面最为全面;IEC主要进行了初期探索;ITU-T聚焦在电信、医疗、多媒体应用等领域融合;IEEE以隐私、可解释为可信赖AI标准重点;NIST布局针对性较强的可信赖AI标准;欧盟以伦理为核心,形成法律法规为目的的AI国际标准化整体处于初期发展、快速部署阶段。AI可信赖、伦理方面备受关注,各标准化组织及机构工作侧重点有所差异但原则理念等整体趋同。现阶段,我国专家全面参与国际标准化组织活动并取得了一定成果。下一阶段,将以提升国内外标准一致性水平、加快我国标准和国际标准相互转化、推动我国先进标准化成果国际化为目标,持续深度参与全球AI标准化活动。

本研究在对当前AI治理现状梳理分析的基础上,发现了如下问题:第一,尚未勾勒出面向未成年人的AI技术发展蓝图。技术规范是立足当下面向未来的,因此,面向未成年人的AI治理必须首先把握好AI的发展趋势,准确勾勒出AI技术发展蓝图,指导建立相应的技术规范框架。第二,尚未建立完善的面向未成年人的AI素养体系。针对未成年人AI素养的培育,有利于培养更多智能社会要求的创新人才。然而现有的AI技术规范体系中往往忽略了提升未成年人的AI素养的重要性。第三,尚未建立完善的面向未成年人的AI应用监管体系。尽管当前各国AI监管机构都高度重视AI相关产品的应用监管,出台了大量的面向特定种类AI应用程序的监管规定,但从目前各国的政策与战略规划来看,尚未有一条明确的落实路径。第四,尚未建立完善的面向未成年人的AI安全伦理规范。从AI发展样态来看,AI技术的影响非常复杂,仅从某单一视角进行伦理审视无疑都将导致研究结果的局限性。第五,面向未成年人的AI技术标准较为缺乏。现有AI技术规范标准大多面向普通大众,专门针对未成年人的AI技术规范标准还未出台,对

于面向未成年人的 AI 技术标准研制的机理、影响因素和可能路径等方面尚不明确。

四、AI 技术规范的概念阐释及定位

(一) AI 技术规范的概念阐释

1. AI 技术规范的内涵

AI 技术规范是本研究的核心术语,学界迄今鲜有针对此术语的明确定义,因此,明晰这一术语的概念内涵是开展研究的基础。规范一词在《辞海》中释义为“标准:法式”,所谓技术规范则是规定技术产品、过程和服务应满足技术要求的文件。从法律层面来讲,技术规范也指有关人类在利用各类技术时应当遵守的行为准则^[7]。AI 时代,AI 技术具有高度的自主性和类人性,技术规范的内涵不断更新,既包括使用技术时应当遵守的行为规范和道德准则,也包含对 AI 的技术要求。故此,笔者认为面向未成年人的人工智能技术规范涵盖以下两种内涵(如图 2 所示):其一,未成年人学习或使用 AI 技术、产品或服务时应该遵守的道德准则和行为规范;其二,为保障 AI 产品质量,确保未成年人的健康成长,AI 产品在研发、制造和部署时应该遵守的操作准则和技术标准,其旨趣在于促进 AI 技术深度应用于未成年人,保障 AI 技术与未成年人的协调发展。

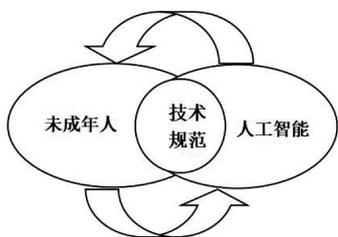


图 2 未成年人、AI 与技术规范的关系

2. AI 技术规范的外延

AI 技术规范内蕴对“人一机”“人一机—人”之间关系的规范,因此,笔者认为可以从核心素养、应用监管和安全伦理三个方面理解其外延。

首先,核心素养是从人才发展视角进行理解,指的是从未成年人学习/使用 AI 技术、产品或服务时应该具备的能力和素养。这要求未成年人需要具备正确使用 AI 技术的核心素养,丰富 AI 相关知识,增强底线思维和风险意识,避免滥用、误用、错用的可能。其次,应用监管是从 AI 向善视角进行理解,为保障 AI 产品质量和安全性,对人工智能产品、过程和服务的技术要求和规范,要求对从事 AI 活动相关的教育机构、企业和人员等进行监管,保障 AI 技术研发、供应和管理满足未成年人的成长特性,降低出现隐私泄

露、数据安全、算法歧视等风险的可能,以确保 AI 向善发展。最后,安全伦理是从人机共融视角进行理解,是指将道德伦理融入 AI 的全生命周期,对从事 AI 管理、研发、应用的企业、机构、学校和未成年人等进行道德规范,积极引导负责任的 AI 研发和应用,增强全社会的伦理意识,确保未成年人与 AI 的协调发展。

(二) AI 技术规范的概念定位

在推进 AI 技术规范的相关研究中,除了需要明晰其概念内涵,笔者认为,还需要从技术规范与社会规范、法律规范和技术标准的区别中透析技术规范的向度和限度,见表 1。首先,技术规范与社会规范是一般“规范”研究的一体两面,前者规范的是“人一机”和“人一机—人”之间的关系,后者只调整“人—人”之间的关系;其次,法律是以理性为特征的社会规范,随着 AI 和大数据规则的加入,法律规范应适当调整,需融合考虑技术规范和社会规范。从规范力度看,技术法规是将技术规范立法,两者可相互兼容,前者具有极强的约束性,一经制定便不容许违反,而后者是兼具“刚性”和“软性”的弹性规范,能更好地平衡 AI 技术与未成年人安全之间的关系,促进两者之间的协调发展。最后,技术规范是技术标准的一种形式,一旦通过标准制定程序,得到公认机构批准,技术规范便可以成为技术标准。

表 1 技术规范与社会规范、法律规范、技术标准的区别

术语	规范对象	规范力度	规范形式
技术规范	“人一机” “人一机—人”	“弹性”约束	制定技术操作要求、准则等
技术标准	“人一机” “人一机—人”	得到公认机构批准, 需严格参照执行	制定为标准
法律规范	“人—人” “人一机” “人一机—人”	通过了立法,具有 极强的约束性	制定为法律
社会规范	“人—人”	政府、机构等内部 制定,需遵照执行	人应该遵守的 道德准则和行 为规范等

五、研究的核心内容

面向未成年人的人工智能技术规范研究以未成年人和 AI 技术为研究对象,核心内容是对该课题的宏观把控。本文遵从“总—分—总”的逻辑,从理论奠基、人才发展、AI 向善、人机共融和规范实践等视角针对 AI 技术发展蓝图、核心素养、应用监管、伦理安全和标准规范进行基础性研究,致力于回答如何从技术规范视角促进未成年人和 AI 技术协调发展问题,如图 3 所示。

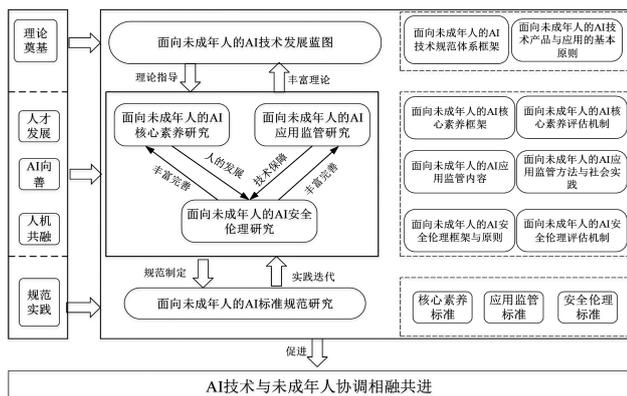


图3 研究的核心内容

(一)理论奠基,面向未成年人的AI技术发展蓝图

2019年,习近平总书记向国际人工智能与教育大会致贺信中提到:“人工智能是引领新一轮科技革命和产业变革的重要驱动力,正深刻改变着人们的生产、生活、学习方式”^[18]。现阶段的未成年人多是伴随着AI成长起来的“Z世代”,从出生起便与信息时代无缝对接,深受数字媒体、智能产品等的影响。因此,有必要从未成年人成长特征和AI应用特征着手,根植于智能时代的发展特性,绘制面向未成年人的AI技术发展蓝图,为AI技术规范研究奠定理论基础。具体地,可从以下几个方面展开:第一,深入文献和应用现状调研,探索AI技术发展脉络,并预判面向未成年人的AI发展趋势;第二,凝练面向未成年人的AI产品应用的基本原则,能够为其研发、设计、部署和应用提供理论指导;第三,通过分析面向未成年人的AI技术规范的构成要素,构建AI技术规范体系框架,为面向未成年人的AI核心素养、应用监管、安全伦理和标准规范研制提供理论指导。

(二)人才发展,面向未成年人的AI核心素养研究

AI技术减少了重复劳动,让人的主动性和创新性得到最大程度的发挥,这要求未来智能时代的人才能够运用智能技术,通过人机协同高效率、创造性地解决实际问题。从认知心理学的视角看,未成年人尚未形成对AI的理解,很容易被同化和顺应,造成未成年人失去批判性思维和创新性思维。因此,如何促使未成年人更好地适应智能化社会发展,对其施行AI核心素养的培育具有重要意义。笔者认为,面向未成年人的AI核心素养研究可从以下几点展开:理论框架层面,通过界定AI核心素养的概念内涵和构成,构建面向未成年人的AI核心素养框架;评估机制层面,从评估方法、评估机制和评估工具三个维度构建面向未成年人的AI核心素养评估机制,生成AI核心素养评估技术规范;实践层面,通过实证提出未成

年人AI核心素养培养与提升建议,以促进AI与人才培养的深度融合,最终培养和提升未成年人的AI核心素养。

(三)AI向善,面向未成年人的AI应用监管研究

《人工智能应用准则》中提到“人的监督”^[19],强调AI不应该超过人的自主性,人类应该监督AI技术的发展及所做决策,并在必要时进行干预。笔者认为,要做好AI应用监管就必须回应“是何”“为何”“如何”的问题。监管内容方面,第一,需要通过管控保证提供给未成年人合乎规范设计的AI产品,保证未成年人的健康和信息安全在知识获取和资源学习的过程中不会受到威胁;第二,采用数据驱动的方式监管未成年人使用AI产品的情况,通过监控应用数据的采集、分析过程以及决策结果,以了解未成年人使用AI产品过程的认知、情感情况,确定AI技术是否会对未成年人的日常行为和心理健康产生影响,以及应该采用什么样的干预措施;第三,基于数据分析的决策结果,监管AI产品的设计和开发过程,确保AI产品是未成年人所需的产品。监管方法与社会实践方面,首先,要构建面向未成年人的AI应用监管的信息模型,完善监管方案,统一监管流程;其次,开发具有一定技术规范的AI应用监管工具,有效进行监管数据的收集、规约、存储、分析和可视化,以数据不断优化智能化监管手段;最后,在明确监管内容和方法的基础上,结合社会实验方法和系统仿真法,在实践中验证监管内容的合理性和监管方法的有效性。

(四)人机共融,面向未成年人的AI安全伦理研究

随着AI技术在社会各领域的渗透,一方面提升了人们的生活、学习的服务质量,另一方面对人们的隐私安全、人身安全和享受公平的权益等伦理安全造成前所未有的冲击。未成年人正处于心智发展的不成熟阶段,具有较强的好奇心和求知欲,但又缺乏自我保护意识和能力,因此,针对这一类群体的AI伦理规范需从以下几点展开。首先,通过梳理适用于未成年人的AI安全技术和伦理原则,建立AI安全伦理整体框架,探析影响未成年人对AI认知和态度的因素;其次,从评估理念、评估方法和评估工具三个角度出发,面向未成年人构建一个AI安全伦理的整体评估机制;最后,基于AI伦理框架和评估机制,进行教育相关的推广与实践,例如,创新教育培训方案和教学模式、进行AI伦理安全课程设计及向政府建言献策。

(五)规范实践,面向未成年人的AI标准规范研制

风险社会视域下,发展与风险是事物发展的一体两面。诚然,AI技术创新发展与安全风险亦是相伴而

生的,故此,需要合规、合伦理的 AI 技术标准规范才能协调好未成年人与 AI 技术发展的关系。鉴于当前鲜有专门针对未成年人的 AI 技术规范,笔者认为可以以下几点进行探寻:一是通过探究影响 AI 技术标准规范研制的影响因素,从多种可行性路径中选出最优路径,为后期标准制定奠定基础;二是经过特定的标准制定程序,将 AI 核心素养规范、应用监管规范和安全伦理规范制定为标准,以构建针对未成年人的 AI 标准体系;三是在前面两者的基础上,提出面向未成年人的 AI 技术规范政策建议,以期增强该标准规范体系的可执行力度。

六、整体的探研进路

明晰面向未成年人的人工智能技术规范研究的整体思路,有助于全面系统地回答 AI 技术如何与未成年人和谐共生、共同发展这一核心问题。研究遵循“和谐友好、包容共享、尊重隐私、安全可控、敏捷治理”的治理原则,从系统观、发展观、治理观、融合观和实践观等不同视角,探索面向未成年人的人工智能技术规范。依此,研究借鉴太阳系结构模型,构建逻辑思路图,如图 4 所示。

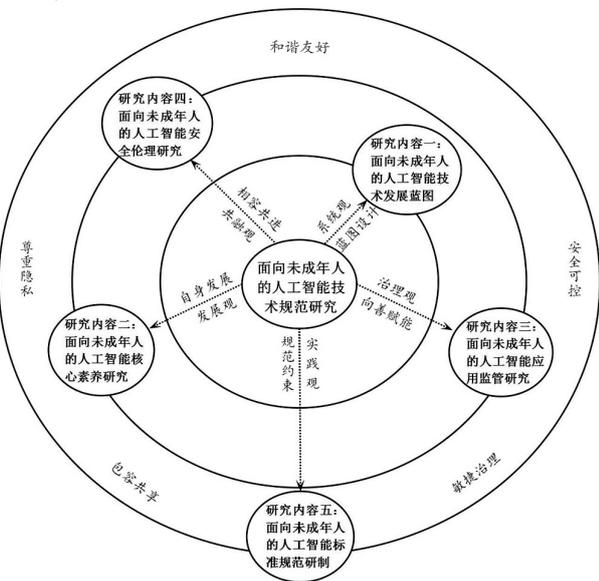


图 4 面向未成年人的人工智能技术规范研究的逻辑思路图

在图 4 中,研究分别以不同视角映射出五个核心内容:(1)基于系统观视角,勾勒面向未成年人的人工智能技术发展蓝图;(2)基于发展观视角,建立面向未成年人的人工智能素养体系;(3)基于治理观视角,建立面向未成年人的人工智能应用监管体系;(4)基于共融观视角,建立面向未成年人的人工智能安全伦理规范;(5)基于实践观视角,制定面向未成年人的人工智能标准规范。研究以复杂系统、社会实验、数据驱

动、场景应用、交叉融合、基于用例的分析(Use Case)等方法论为指导,依循“趋势预判,理论建构”“聚焦重点,分类推进”“标准研制,建言献策”路径,提出面向未成年人的人工智能技术规范研究新思路。各个核心内容的研究视角、方法论及路径,见表 2。

表 2 各个核心内容的研究视角、方法论及路径

研究内容	研究视角	研究方法论	研究路径
内容一	系统观	复杂系统	理论探索—实践生成—理论优化—迭代生成
内容二	发展观	社会实验、数据驱动、场景应用	理论框架构建—评估机制构建—评估实证
内容三	治理观	社会实验、数据驱动、场景应用	理论构建—方法探索—社会实践
内容四	共融观	社会实验、数据驱动、场景应用	理论框架构建—评估机制构建—实证和培养
内容五	实践观	交叉融合、社会实验、基于用例的分析	路径选择—标准研制—政策建议

第一,基于复杂系统理论,顶层架构面向未成年人的人工智能技术规范体系。复杂系统理论是解决以非线性为核心的复杂系统问题的理论,重在揭示系统的内部交互和外在演化规律^[20]。要解决复杂系统问题,需采用多尺度分析方法进行复杂系统建模,厘清其中错综复杂的依存关系和行为的多元因果联系^[21]。促进未成年人成长与 AI 技术治理协调相容共进的过程,在本质上属于复杂系统的协同演化过程,具有显著的非线性、涌现性和开放性特征^[22]。本研究将从复杂系统建模的社会计算角度出发,通过构建未来智能生态系统中的多主体系统,以仿真 AI 技术的发展样态,剖析系统所涌现出的 AI 发展蓝图。同时,探究不同外界变量及其组合模式对未成年人成长发展的影响,揭示 AI 技术与未成年人健康成长的协同共生机制,顶层架构面向未成年人的人工智能技术规范体系。

第二,统筹运用多元视角,分类推进面向未成年人的人工智能生态系统建设。面向未成年人的人工智能生态系统是由未成年人、智能支撑技术、社会环境与复杂应用场景等多要素构成的一个复杂系统。人机协同是面向未成年人的人工智能生态系统建设的核心和突破口,要求始终以未成年人的素养发展为本,对人工智能应用的风险边界进行有效管控,才能实现人机相融共进^[23]。故此,研究基于发展观、治理观、共融观的视角,以社会实验、数据驱动、场景应用的研究方法论为指导,共同遵循“理论构建—方法探索—实

践验证”的研究路径,分类推进面向未成年人的人工智能核心素养、应用监管、安全伦理研究内容,以重构AI社会生态,实现共建人类命运共同体的现实诉求。

第三,跨学科交叉融合,研制面向未成年人的人工智能标准规范和政策建议。跨学科交叉融合是不同学科在思维、方法、技术和语言等方面的全面融合,是超越学科边界以解决真实问题的学科整合研究^[24]。面向未成年人的人工智能技术规范研究既关涉AI这样的前沿科技,也涉及教育系统甚至道德伦理等人文领域,是一个跨科技、社会和人文的交叉领域研究。本研究融合系统科学、社会学、教育学、信息学、心理学等学科视野,将多门学科的研究方法进行系统性整合,生成适用于本研究的跨学科研究方法,以指导面向未成年人的AI核心素养、应用监管和安全伦理等方面的标准研制工作,积极为政府部门的政策制定和决策执行谏言献策。

第四,融合教育实践方法,以教育社会实践和教育工程实践推动标准规范研制与示范应用。教育社会实践是以育人目标为导向,识别因素差异对学习行为、人格塑造的改善路径^[25]。教育工程实践是指遵循“系统需求分析、标准范围确定、分类指导、系列标准采用、标准获取和解读、标准的集成和系统开发、系统完善示范应用”流程和步骤开展标准研制和应用工作的过程。面向未成年人的人工智能标准研制需融合以上两种实践方法,一方面,采用教育社会实验的设计思路,针对上海市宝山区、黄浦区、浦东新区,浙江温州市、安徽合肥市、北京海淀区、江苏南京市、新疆喀

什市、广西柳州市等建立实验区,以及多所学校建立试点校,实验区和试点校按照整个项目要求和自身发展需求设立子课题开展研究,如上海市求真中学计划开展智慧课堂(三个助手)、智能备考助手、智慧体育等方面研究;另一方面,充分利用标准委员会资源,依照标准研制流程和步骤规范化开展标准研制工作,并落实标准的应用与迭代。

七、结 语

数字化时代,AI和教育的深度融合促使教育发生全方位的创新与变革,加快培养AI高端人才是推动智能社会发展的必然之举。未成年人正处于AI高端人才培养的关键时期,如何协调好人工智能与未成年人之间的关系成为亟须解决的时代命题。由此,本文以促进人机协调相容共进为起点,基于理论奠基、人才发展、AI向善、人机共融和规范实践的视角,剖析面向未成年人的人工智能技术规范研究的核心内容和探研进路,不仅为后续各个核心内容的研究提供指导,同时引领实验区和试点校深入开展教育实践。然而,在人类命运共同体的视域下,面向未成年人的人工智能技术规范研究必将衍生出一系列理论和技术,如何协调好理论研究和教育实践的关系,是后续需要深入研究的重要议题。同时,以ChatGPT为代表的新一代人工智能工具将对教育产生广泛的影响,影响到我们的教育观、教学观、学习观、知识观、研究观、治理观,教师、学生、家长、管理者、研究者和决策者将要如何科学地应对,这也是本项目后面需要深入研究的内容。

[参考文献]

- [1] 中华人民共和国科学技术部.发展负责任的人工智能:新一代人工智能治理原则发布[EB/OL].(2019-06-17)[2022-05-29]. http://www.most.gov.cn/kjbgz/201906/t20190617_147107.html.
- [2] 中华人民共和国中央人民政府.中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议[EB/OL].(2020-11-03)[2022-06-09]. http://www.gov.cn/zhengce/2020-11/03/content_5556991.htm.
- [3] 贺建平,赵晓燕,黄肖肖.网络暴力游戏与青少年暴力行为的相关性[J].新闻界,2009(1):42-45.
- [4] 张振锋.网络不良信息对未成年人犯罪的影响[J].预防青少年犯罪研究,2017(1):31-37.
- [5] 刘亭.智能时代我国意识形态风险治理研究[D].马鞍山:安徽工业大学,2021.
- [6] 易凌云.互联网教育与教育变革[D].武汉:华中师范大学,2017.
- [7] UNESCO. Artificial intelligence in education: challenges and opportunities for sustainable development [EB/OL]. (2019-04-07)[2022-06-09]. <https://unesdoc.unesco.org/ark:/48223/pf0000366994>.
- [8] 中华人民共和国教育部.联合国教科文组织正式发布国际人工智能与教育大会成果文件《北京共识——人工智能与教育》[EB/OL].(2019-08-28)[2022-06-09]. http://www.moe.gov.cn/jyb_xwfb/gzdt/s5987/201908/t20190828_396185.html.
- [9] UNESCO. K-12 AI curricula: a mapping of government-endorsed AI curricula[EB/OL].(2022-02-16)[2022-06-04].<https://unesdoc.unesco.org/ark:/48223/pf0000380602>.
- [10] 杨宗凯,王俊,吴砥,等.发展智能教育学 推动教育可持续发展[J].电化教育研究,2022,43(12):5-10,17.

- [11] 赵磊磊,吴小凡,赵可云.责任伦理:教育人工智能风险治理的时代诉求[J].电化教育研究,2022,43(6):32-38.
- [12] 杜静,黄荣怀,李政璇,周伟,田阳.智能教育时代下人工智能伦理的内涵与建构原则[J].电化教育研究,2019,40(7):21-29.
- [13] 张立国,刘晓琳,常家硕.人工智能教育伦理问题及其规约[J].电化教育研究,2021,42(8):5-11.
- [14] 沈苑,汪琼.人工智能教育应用的偏见风险分析与治理[J].电化教育研究,2021,42(8):12-18.
- [15] 金玲.全球首部人工智能立法:创新和规范之间的艰难平衡[J].人民论坛,2022(4):44-47.
- [16] 宋婧.人工智能伦理之争[N].中国电子报,2022-04-01(03).
- [17] 李晓明.论人工智能刑法规制中的技术规范[J].东方法学,2021(2):54-68.
- [18] 中华人民共和国中央人民政府.习近平向国际人工智能与教育大会致贺信[EB/OL].(2019-05-16) [2022-05-29]. http://www.gov.cn/xinwen/2019-05/16/content_5392134.htm.
- [19] 旷视.人工智能应用准则[EB/OL].(2019-07-18)[2022-05-31]. <https://baijiahao.baidu.com/s?id=1639361752623817286&wfr=spider&for=pc>.
- [20] LOUIS C L, MANION K M. Research methods in education [M]. 6th ed. London; Routledge, 2007:28-30.
- [21] YANEER B Y. From big data to important information[J]. Complexity, 2016, 21(S2):73-98.
- [22] NORGAARD R B. Coevolutionary development potential[J]. Land economics, 1984, 60(2):160-173.
- [23] 祝智庭,韩中美,黄昌勤.教育人工智能(eAI):人本人工智能的新范式[J].电化教育研究,2021,42(1):5-15.
- [24] 岳伟.论教育科学研究的跨学科交叉融合[J].中国教育科学(中英文),2022,5(1):32-41.
- [25] 童莉莉,张晨,黄荣怀,雷凤宇,杨赞波.教育社会实验:人工智能融入教育的研究新探索[J].中国电化教育,2022(3):62-68.

Research on Technical Specifications of AI for Minors: Theoretical Analysis and Research Approach—Research on Technical Specifications of AI for Minors (1)

WU Yonghe, GUO Shengnan, XU Qiuxuan, ZHENG Hao, YAN Huan

(Department of Education Information Technology, East China Normal University, Shanghai 200062)

[Abstract] In the era of artificial intelligence, how to coordinate the relationship between artificial intelligence and minors has become a crucial issue, and the formulation of technical specifications for artificial intelligence (AI) designed for minors is the key solution to this issue. Based on the national social science major project of "Research on AI Technical Specifications for Minors", this paper theoretically explains the core content and research approach for developing the technical specifications of AI for minors. This paper comprises five parts: (1) the appeal of the age of AI for minors; (2) the analysis on the trend of AI governance for minors; (3) explanation and positioning of AI technical specifications; (4) the core content of the research; (5) the overall research approaches. The aim of this study is to help readers understand the focus and direction of the research on technical specifications of AI for minors, and provide theoretical guidance and reference for the future related research, ultimately contributing to resolving the coordination issues between AI and minors, and building a community with a shared future for mankind.

[Keywords] Artificial Intelligence; Technical Specifications; Minors; Core Competence; Application Supervision; Security Ethics