

人工智能教育大脑模型构建及实践路径

郭胜男¹ 吴永和¹ 张治²

(1. 华东师范大学 教育学部教育信息技术学系, 上海 200062;
2. 上海市宝山区教育局, 上海 201999)

[摘要] 人工智能教育大脑是整合新一代智能技术的类脑复杂智能教育系统, 可为解决教育领域数据治理难题提供新方向。智能时代的教育数据治理在教育数据质量、安全风险和算法伦理等方面面临诸多现实难题, 极大限制了教育大数据价值的发挥。受互联网大脑和城市大脑的启发, 本研究通过解构教育数据治理的内涵、困境与成因, 从教育大脑结构模型、轻量级教育大脑信息模型和教育大脑的核心模块三个层面构建人工智能教育大脑模型, 以厘清教育大脑内部的结构、信息流转和技术关键。本研究还从建立数据管理组织架构、创新安全风险防范体系及构建算法伦理监管机制三方面提出教育大脑赋能教育数据治理的实践路径。人工智能教育大脑通过系统性整合教育数据、智能算法和算力资源, 强化教育数据治理体系和智能技术的耦合性, 保障教育数据质量与安全, 充分发挥大数据在教育治理中的“数字红利”, 为不断推进教育数字化转型、实现教育治理现代化提供新思路。

[关键词] 人工智能; 教育大脑; 教育数据治理; 教育治理; 数字化转型

[中图分类号] G434

[文献标识码] A

[文章编号] 1007-2179(2022)04-0022-09

一、问题提出

人工智能技术发挥作用的三要素为数据、算法和算力。正如人工智能专家周志华(2020)所言, “今天的人工智能热潮是由于机器学习, 尤其是深度学习技术在大数据、大算力的支持下发挥出了巨大威力”。这些均需建立在数据质量、数据安全及算法合乎伦理的基础上, 因此, 如何进行有效的教育数据治理成为教育数字化转型的必经之路。人工智能教育大脑系统整合了教育数据、智能算

法和算力资源, 以保证最大限度地挖掘大数据的潜在价值, 为教育数据精准治理提供新方向。

“智能大脑”最早由雷·库兹韦尔(Ray Kurzweil)提出, 指借助海量数据和算法, 模拟人脑思考和判断(何怀宏, 2018)。之后在城市治理领域, 有研究者提出城市大脑概念, 指在互联网大脑架构的基础上, 以云神经网络和云反射弧为建设重点, 为解决城市治理难题提供全新工具(胡坚波, 2021)。在教育领域, 顾小清等(2021)以“人工智能教育大脑”为隐喻, 将智能技术与教育融合, 致力于解决

[收稿日期] 2022-05-24 **[修回日期]** 2022-07-04 **[DOI 编码]** 10.13966/j.cnki.kfjyyj.2022.04.003

[基金项目] 国家社会科学基金 2021 年度重大项目“面向未成年人的人工智能技术规范研究”(21&ZD328); 2020 年上海市科学技术委员会科研计划项目“教育大数据治理体系与面向大规模智慧学习的教育大脑研究”(20511101500)。

[作者简介] 郭胜男, 博士研究生, 华东师范大学教育学部教育信息技术学系, 研究方向: 智能教育与学习分析研究, (Shining_guo123@163.com); 吴永和(通讯作者), 博士, 研究员, 博士生导师, 华东师范大学教育学部教育信息技术系, 研究方向: 智能驱动的教育研究与教育数字化转型(yhwu@deit.ecnu.edu.cn); 张治, 博士, 上海市宝山区教育局局长, 研究方向: 教育智能化、教育大数据(87476088@qq.com)。

[引用信息] 郭胜男, 吴永和, 张治(2022). 人工智能教育大脑模型构建及实践路径[J]. 开放教育研究, 28(4): 22-30.

高等教育数据治理和教学创新问题。张治等(2022)将人工智能本质和人脑特质结合,构建教育大脑生态架构,强调基于人工智能的教育大脑是实现教育数字化转型的关键基础设施。由此可见,人工智能等新一代信息技术支撑的“教育大脑”具备规范化整合数据资源和算法模型、提高数据质量和保障隐私安全的强大能力。

数据作为新的生产要素,在渗入教育治理结构时,必然会引发一系列新的问题,如数据质量参差不齐、隐私泄露、算法失衡。现阶段教育领域尚缺乏有效的教育数据治理方案。为了提高教育数据质量,保障算法安全,打破数据孤岛之间的壁垒,充分挖掘教育大数据的潜在价值,教育治理亟需更宽广的视野和更强大、灵敏的治理工具。本研究遵循“问题解构→方法模型→实践路径”的逻辑思路,在解构教育数据治理的内涵、困境及成因的基础上,构建人工智能教育大脑模型,提出教育大脑赋能教育数据治理的实践路径。

二、概念阐释、困境与成因

(一)概念阐释

数据治理(data governance)是数字化技术和治理理论融合形成的新概念,学界尚未给出统一的定义。国际数据管理协会(DAMA International, 2009)认为,数据治理指在管理和使用层面对数据进行规划、监督和控制。国际著名数据治理专家桑尼尔·索瑞斯(2014)认为数据治理是制定与大数据相关的数据优化、隐私保护与数据变现等政策。国内学者梁芷铭(2015)认为数据治理是为了应对大数据带来的种种不安、困难与威胁,运用技术工具对大数据进行管理、整合、分析并挖掘其价值的行为。有学者将数据治理等同于数据管理,认为二者均指采取必要的步骤、度量 and 策略管理数据的整个生命周期,涉及数据识别、收集、准备、组织、分类、处理、分析、存储、发布、管理和重用等(Gajbe et al., 2021)。事实上,数据治理蕴含明确相关角色、工作职责和 workflows 等含义,是过程性概念,强调做出哪些决策以及由谁做出决策(Abraham et al., 2019)。数据管理是目标性概念,强调决策是数据执行的一部分。从本质上讲,数据治理是对数据管理进程行使权利和控制,指导组织内和跨组织的数

据和算法整个生命周期的正常运行,强调数据资产的价值和数据安全管控(Carretero et al., 2016)。

随着教育信息化的发展,教育领域积聚了海量数据,呈现“大数据”形态,“教育数据治理”概念应运而生。教育数据治理作为崭新的研究领域,其概念一直是学界讨论的议题。教育数据治理涵盖教育的数据治理和数据驱动的教育治理两层内涵,数据治理为教育治理提供技术和程序,教育治理为数据治理提供目标(刘金松,2018)。两者交叉融合、密不可分,共同服务于教育发展需求,提升教育治理水平和治理能力。

(二)现实困境及成因

研究发现,现有教育治理体系与大数据算法系统的耦合度不高(Janssen et al., 2020),教育数据治理在数据质量、安全风险、算法伦理等方面面临难题,未能充分发挥大数据的价值,严重阻碍了教育数字化转型的现代化进程。

1.教育数据质量不高,限制了数据价值释放

教育大数据在科学决策、支持多样化教学场景、创新教学服务和提升教学效率等方面蕴含巨大的教育价值,正如徐宗本院士(2017)所言,“大数据可以带来超凡应用价值”。在教育领域,只有切实将数据应用于解决教育发展问题,才能真正彰显数据的价值。但实际状况与理论设想“相距甚远”,“虽然大数据在教育领域的热度日趋增加,真正利用大数据解决教育问题的成功案例却很少”(祝智庭,2017)。究其原因,数据质量是影响数据价值发挥的重要因素。

质量是教育数据的灵魂,是大数据发挥价值的基础和保障,涉及数据规范性、准确性、完备性和一致性等。在智慧教育时代,人工智能、云计算、虚拟现实和物联网等智能技术扩展了教学管理和实践的时空边界,政府部门、社会组织及各类院校等产生的海量数据在数据规模、结构类型、数据收集和存储工具、时效价值等方面存在很大差异,教育数据的规范性、准确性、完备性、一致性等数据质量问题叠出,导致难以充分发挥数据的价值。这可归因为以下几点:其一,各教育主体间的“烟囱效应”。人工智能技术赋能多元化数据生产主体和治理主体,但各主体间因业务需求不同往往“各自为政”,未能形成统一的数据规范标准。在这种

“烟囱式”治理模式下,数据冗余、数据片面、死数据等问题频出,导致教育决策科学性不高。其二,学习分析技术与数据之间的“弱耦合性”。区别于一般大数据,教育大数据是在教育场景中产生的数据,结构多样(结构化、半结构化、非结构化),模态丰富(行为、生理、心理),格式繁杂(视频、音频、文本、平台日志、生物传感)(吴永和等,2021),现有学习分析技术难以捕获数据之间的关联性和一致性。此外,高效的学习分析工具推广不易,如果没有利益相关者的批准和监管,高效的学习分析技术不会被教育系统采纳。可见,学习分析技术与数据之间的“弱耦合”极大地限制了数据价值释放。其三,从技术角度看,机器学习算法偏好“好数据”。教育领域可视为数据驱动下的智能算法的运行场域(赵磊磊等,2021)。流行的机器学习算法,如K-最近邻(KNN)、朴素贝叶斯(NB)、支持向量机(SVM)和深度学习等对特征空间的变化非常敏感,高质量的训练数据集才能形成高性能的特征向量,进而将机器学习算法训练成“好”的模型(Singh et al., 2022)。然而,教育实践中通常存在降低数据质量的因素,如噪声、缺失值、数据不一致或异常值等,严重影响机器学习算法的学习性能,造成模型训练数据的欠拟合或过拟合,进而导致决策失误。

2. 开放共享程度扩大,加剧数据安全风险

2017年,国务院印发《国家教育事业发展规划“十三五”规划》指出,要加快推进教育大数据建设与开放共享,形成教育数据资源开放共享机制(教育部,2017)。人工智能时代,教育系统内部自组织、自生产形成的“熵减”现象和自我加强的“马太效应”均蕴含着教育数据开放共享的迫切诉求。开放共享对数据资源的快速流转与价值挖掘至关重要(陈雅云等,2021)。扩大教育数据的开放共享空间有利于促进教育数据的整合、流转,打破各教育子系统之间的数据壁垒,提升数据的应用价值,形成相互支持的产业链。

移动互联网、区块链、边缘计算、人工智能等技术深度发展,各教育系统、平台和信息门户等的加强,为教育数据的开放和共享提供了强有力的技术支撑,开放数据获取、传输、存储、处理和应用的力度不断增强。在此过程中,数据泄露和数据破

坏等数据安全风险问题不可忽视。一方面,教师、学生等数据生产者的隐私泄露风险加大。从宏观上看,数据的开放共享具有高度复杂性和系统性,涉及各利益相关者的隐私。随着数据开放共享力度的提升,学校、教师、学生等数据提供者的隐私空间被严重压缩(田贤鹏,2020)。从技术视角看,学习分析系统的设计未能完全保证师生的隐私安全。随着数据资源更加多模态,学习分析工具更加移动,原本无法追踪、统计和检索的踪迹变得有迹可循,隐私侵犯现象变得普遍。另一方面,数据在共享过程中被破坏的风险增大。其一,开放共享为不法之徒窃取、篡改、滥用数据提供了可能,尤其是一些教育机构为了利益窃取开放共享的数据以换取商业报酬。其二,数据在传输和交换过程中会丢失。教育系统各组织间数字媒介和数据维度等不一致,使得数据在交换过程中易丢失或损坏,造成数据缺失或不完整。总而言之,数据安全治理是教育数据开放和共享过程中不可或缺的重要环节,也是教育数据治理亟需解决的核心问题。

3. 人工智能技术带来算法失衡风险

人工智能时代是一个让数据“开口说话”的时代,数据和算法是人工智能技术发挥作用的重要基石。“如果说大数据是信息时代的石油,那么算法就是从中提炼衍生产品的炼油厂”(肖凤翔等,2020)。大数据为算法提供底层支撑,算法从大数据中提炼知识和信息,可见,算法优劣严重影响数据治理的效能。然而,随着智慧学习、智慧课堂等人工智能新应用的出现,算法在教育实践中不可避免地遇到诸多困境,如算法失灵、算法偏好、算法滥用,给教育实践带来了不可预估的风险,严重限制了教育数据治理效能的发挥。

首先,机器学习算法在为数据治理提供强大算力的同时,其应用存在算法失灵的风险,具体表现为算法运行失序,背离教育德性伦理或者不利于教育正向发展。教育领域算法大多移植于其他领域,这些算法的数理逻辑并不完全适用于教育系统的技术架构和教育逻辑,无法完全满足教育的实际需求,有的会背离以“人的发展”为核心的目标。其次,算法偏好风险。“再好的算法也不能完美反映客观实在,其中必然存在诸多偏差与偏见”(谭维智,2019)。教育领域的算法应用是将教育现象抽

象成数据进行表征, 其中的程序设计和计算逻辑是人为设计的, 不可避免存在偏见, 从而导致算法结果出现偏差。机器学习算法以教育数据为支撑进行模拟和诊断, 诊断的准确性往往与数据质量相关, 偏好“好数据”, 造成“茧房效应”。最后, 算法的误用和滥用风险。从教育实践者的视角看, 算法分析过程是个“黑匣子”, “知其然而不知其所以然”, 这无疑增加了算法使用的盲目性, 无法针对性地服务于教育教学。算法的滥用还模糊了公共领域和个人隐私边界, 可能侵犯师生的正当权益。

三、模型构建

数据质量、安全风险和算法伦理是教育数据治理亟需解决的难题。人工智能教育大脑本质上是类脑复杂教育系统, 面对系统性和功能性如此复杂的教育大脑, 本研究从结构模型、轻量级信息模型和核心模块三方面剖析人工智能教育大脑的整体结构、信息流转调度及核心功能, 为教育数据治

理实践提供支架。

(一) 人工智能教育大脑结构模型

任何时期的教育变革都离不开物质技术基础的支持, 新一代信息技术驱动的教育新基建是支持教育大脑运转的“基石”和“底座”(见图1)。教育大脑的底层支撑技术以人工智能为核心, 通过与物联网、5G、云计算、边缘计算、大数据、数字孪生等新技术的叠加, 为教育提供了一套全新的基础设施——教育新基建。从功能性结构来看, 教育新基建主要包括信息基础设施、融合基础设施和创新基础设施(祝智庭等, 2022)。其中, 信息基础设施可细化为网络通讯基础设施、新技术基础设施(如区块链、数字孪生)和计算基础设施。融合基础设施指传统教育基础设施在人工智能等技术加持下形成的集时空、虚实环境于一体的融合型基础设施。创新基础设施指支持科技创新的基础设施, 可不断为教育数字化转型提供新型工具。

正如城市神经网络和云反射弧是城市大脑

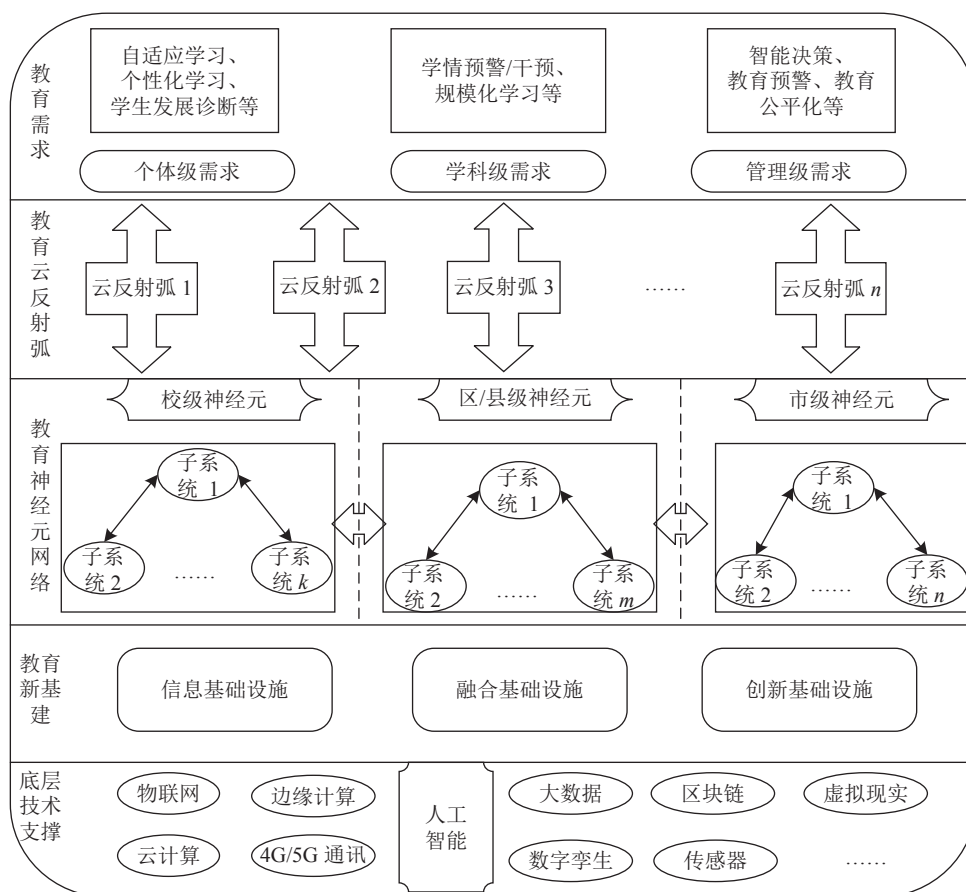


图1 教育大脑结构模型

建设的重点内容,教育神经网络和教育云反射弧是教育大脑结构的核心内容。神经元是教育神经网络的基本单元,每个神经元代表具备一定数据收集、存储、分析和决策功能的教育子系统。它从层级上可分为校级神经元、区/县级神经元和市级神经元,每个神经元按照一定的功能标准和安全标准相互联通、资源共享,形成教育神经网络,为教育云反射弧提供统一的运转基础。每条通过技术联通的教育神经元之间的链条通路便是一条教育云反射弧,对应不同教育需求,包括个体级需求、学科级需求和管理级需求。

(二)轻量级教育大脑信息模型

教育大脑本质上是一个类脑复杂智能教育系统,将大量的教育系统、子系统和平台等神经元链接起来,涉及众多的复杂技术和因素。为了明晰教育大脑内部的数据流转和信息安全调度,本研究分别针对学生成长管理、面向特定学科的自适应学习管理和线上线下融合智慧教育模式的需求,构建轻量级教育大脑信息模型(见图2),以厘清数据和

算法有效治理的交互机制。

教育大脑作为互联网大脑的进化产物,其构建初衷是解决教育领域的重难点问题,响应教育需求。轻量级教育大脑信息模型包括教育神经网络、数据湖、数据仓库、智能处理器和智能服务等模块,遵循“问题驱动→数据输入→数据共享→数据分析→数据服务→需求响应”的逻辑,通过模拟生物的反射弧结构,深度挖掘教育数据的潜藏价值,对教育问题或需求进行及时且智能化的反馈。教育神经网络由无数个相互联结的教育神经元构成,每个教育神经元信息由教育内容、数据来源和数据集成规则等构成。教育内容可借助普适设备(如平板电脑等)、多模态传感器、平台日志、问卷和访谈等方式收集视频、语音、文本、生物传感等多模态数据(吴永和等,2021)。每个教育问题对应的云反射弧均需要汇聚多个教育神经元的数,具体操作流程为:首先,将教育神经元中产生的海量异构数据像水一样汇聚在数据湖中,数据湖以原始格式存储数据,无需对数据进行结构化处理,保留数

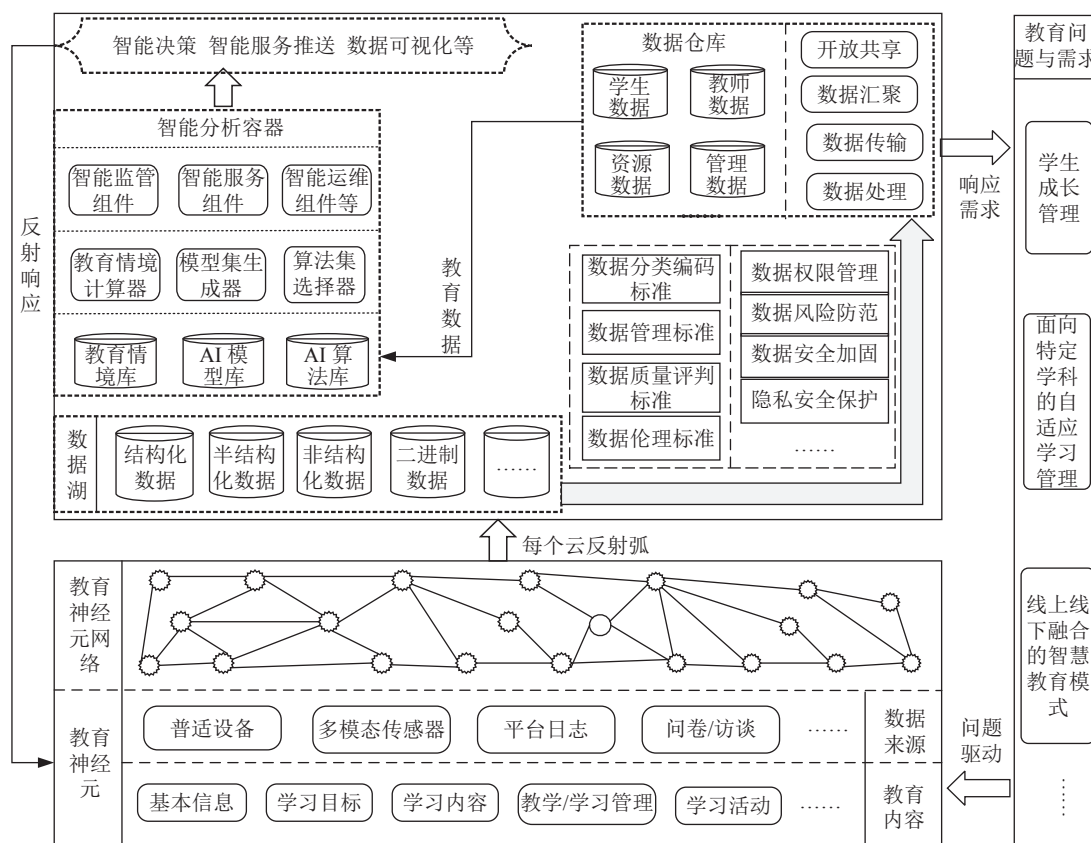


图2 轻量级教育大脑信息模型

据之间自然形成的“血缘关系”,包括结构化数据、半/非结构化数据和二进制数据(如视/音频);其次,教育神经元之间的数据共享和融通严格遵循数据质量、数据伦理和数据安全等规范和标准,以保障数据在开放共享过程中的质量和安全;再次,数据共享和融通将数据按一定的规范和标准格式化处理后存储在数据仓库中,作为智能分析容器的数据原材料;智能分析容器是将数据原材料转为智能服务的“工厂”,由自底向上的三部分组成,分别是底层算法/模型库、中层算法/模型匹配器和顶层服务组件;最后,提供智能决策、智能服务推送和数据可视化等智能化服务,将反射响应结果反馈给神经元以完成云反射,同时响应对应的教育问题,实现教育的精准化治理。

(三)教育大脑核心模块

人工智能的发展是不断模拟、扩展人类智能的过程,反观教育领域的现实境况,人工智能技术“有计算而无算计”“有智能而无智慧”“有感知而无认知”。人工智能教育大脑要实现“眼观六路、耳听八方、融合思考”的教育发展愿景(顾小清等,2021),离不开人类群体智慧和机器群体智能的深度协作,更离不开智能算法和高质量数据的

底层支撑。具体而言,教育大脑主要包括智能分析和数据融通两大模块(见图3),它们可类比为教育大脑的“左右脑”,左脑追求理解与认知,右脑追求快速大量记忆和自动处理机能,左右脑相互协调,提升教育大脑的“智商”和“数商”。

1. 数据融通模块

教育大脑的数据融通模块是以数据中台技术为架构的数据汇聚中心,致力于实现教育数据的共享融通、规范统一和风险防范,主要功能包括:其一,基于区块链、云计算和物联网等技术构建数据湖,在逻辑上联通各神经元以不同方式收集的多模态异构数据,既保证数据的“进化”和“血缘关系”可追踪,又保障数据“物种”多样性;其二,数据共享和流通需严格按照数据权限管理和风险防范制度,明确权责边界,加强隐私保护,保障数据的安全流通;其三,严格遵循数据分类编码标准、数据管理标准、数据质量评估标准和数据伦理标准,对数据生命周期全流程进行规范和监控,采用统一的标准化规则清洗、规范化、标签化数据,将满足规则的数据存入数据仓库中,为后续智能分析持续提供“燃料”。显然,数据融通模块通过清洗、筛选、关联多模态异构数据,打通数据流通渠道,提升数

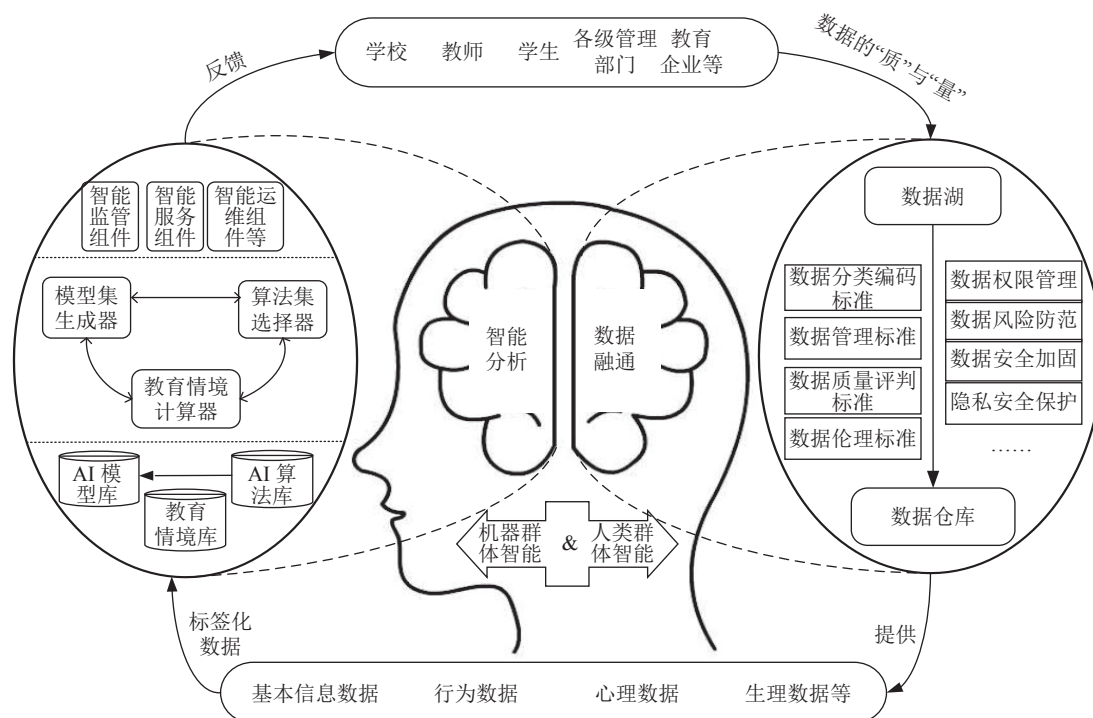


图3 教育大脑的核心模块

据质量,为教育数据的精细化治理奠定了基础。

2. 智能分析模块

智能分析模块是人工智能教育大脑的数据“加工厂”,它通过算法训练、情境匹配、决策服务等,充分发挥机器群体智能以挖掘教育数据的潜在价值。首先,智能分析模块储存了合理、科学、丰富的资源库,包括算法库、模型库和教育情境库。其中,人工智能模型是人工智能算法(移植的开源算法或教育领域的创新算法)根据教育情境产生的教育数据训练而得的。其次,智能时代的教育情境是多元变化的,只有在特定情境中选择適切的人工智能算法/模型,才能提供精准的决策服务。因此,智能分析模块提供了算法集选择器、模型集选择器和教育情境计算器,保证教育大脑能够自主提供适当的计算和决策服务,提升教育大脑的“智商”。再次,人工智能算法的监管功能,通过设置算法的智能监管组件监督算法使用过程中的伦理问题,防止出现算法滥用、算法误用和算法失灵等风险。最后,人工智能算法的创新功能,用于不断迭代解决教育问题的人工智能算法,更新算法库和模型库,确保教育大脑的高效运转。

四、实践路径

教育数据治理是一个复杂且反复迭代的系统工程,目的是实现数据应用价值和风险管控。本研究提出的人工智能教育大脑以新一代智能技术为基础,充分整合数据、算法和算力,打通教育各部门的业务壁垒,为数据资源共享与规约整合、数据

安全和隐私保障、算法“善治”、决策智能化等提供治理支架。因此,基于新一代智能技术,将智能治理的理念融入教育治理,在对数据和算法进行治理的同时,将“冰冷”的数据和“呆板”的算法转化为数据质量和风险防范的“利器”,是智能时代摆脱教育数据治理困境的实践逻辑(见图4)。

(一)加强顶层设计,建立数据管理组织架构

人工智能教育大脑为教育数据生命周期全流程提供了一系列质量规范,如数据分类编码、数据标准、数据质量评估等,但我们仍需要开展顶层设计和战略规划,建立权责明晰的数据管理组织架构,配合教育大脑形成完备的管理体系。其一,建立科学合理的数据管理组织架构。进入智慧教育时代,教育数据管理模式已由“逐层上传”的科层级模式逐步转向“网格管理”的扁平化模式,各级教育部门需建立权责分明的数据管理中心,落实各级部门的权力和责任。这包括:一是确保数据生命周期的每一流程都“有人负责,有人监督”,一旦发现数据质量等问题能够及时溯源;二是搭建各层级业务间的“沟通桥梁”,相互验证数据的真实性和有效性,保障数据的质量和安全流转。其二,完善数据生命周期全流程规范化制度。数据质量参差不齐归根究底是缺乏规范化的数据标准和流程。这就有必要从政策和制度着手,加强数据全生命周期的制度建设:1)以数据标准、质量管控、数据共享权限等政策为依据,进行整体性制度设计,制定通用的数据规范格式;2)全流程遵循“数据收集→数据存储→数据处理→数据传输→数据交换→数

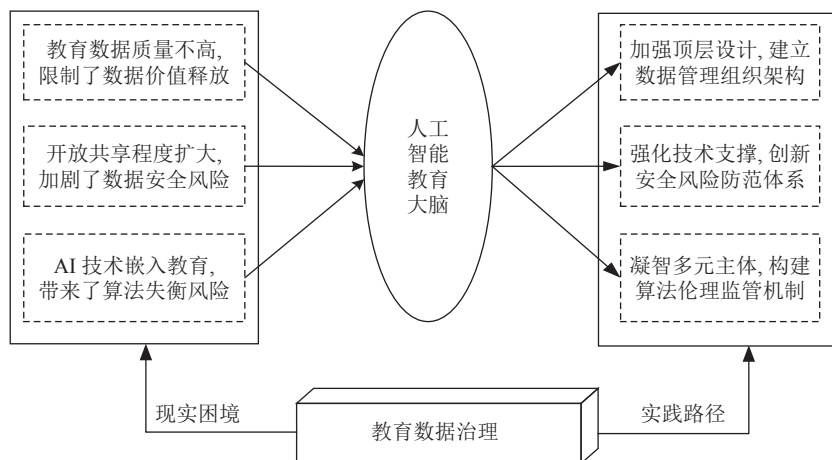


图4 教育大脑赋能数据治理的实践路径

据销毁”的数据标准制度,各层级数据管理中心采用统一的规范进行数据流通,保障数据的规范性、准确性、完备性和一致性,提升数据的应用价值。

(二)强化技术支撑,创新安全风险防范体系

人工智能时代,教育数据安全问题日益突出,如何利用技术手段防范安全风险成为教育数据治理的当务之急。人工智能教育大脑通过教育神经元网络响应教育需求的过程中,为规避数据治理的安全风险需设置防范措施,如共享数据权限管理、数据风险防范、数据安全加固和隐私安全保障等,这些措施以智能技术为依托,遵循“风险识别→风险预警→风险管控”的系统流程,将风险防范的“先手”与风险消解的“后招”有效结合。第一,建立数据安全风险智能识别机制。这需要明确数据安全风险类别,通过数据挖掘和特征提取等技术对风险源进行特征提取,并建立数据安全风险模型准确识别数据安全风险。第二,建立数据安全风险实时预警机制,即基于数据挖掘、物联网、大数据和机器学习等技术,快速整合多源的多模态风险数据,构建数据安全风险预警平台,完善“快速感知、实时监测、事先预警”的一体化流程,实现对数据安全风险的实时预警。第三,建立数据安全风险智能管控机制。风险管控指针对数据泄露和数据破坏等风险预先采取的挽救措施,如梳理数据湖中的数据来源,进行数据血缘分析,对重要性和敏感程度不同的数据启动不同级别的数字加密等。

(三)凝智多元主体,构建算法伦理监管机制

人工智能技术嵌入教育系统带来的不仅是技术问题,更多的是源于人类自身的人性和伦理问题。事实上,人们对技术应用于教育的监管能力远不及技术的革新速度(谭维智,2019)。因此,为了确保人工智能教育大脑的良性运转,我们需要凝聚多元教育主体的智慧,对算法/模型进行伦理监管,以保证其符合“以人为本”的教育价值观。这包括:1)算法设计需符合教育逻辑。算法设计者需以“人的发展”为设计理念,根据“人的体验”不断改善算法设计的影响因素。2)算法选择需进行评估。算法对大多数教师、学生或教育管理者等使用者来说是“黑匣子”,这就需要对算法进行质量评估和风险评估。一方面,算法使用者应根据具体教育情境进行算法迁移、模型训练和算法更新,评估模

型计算是否丢弃了数据携带的重要信息以及从其他领域迁移的算法是否符合教育情境需要;另一方面,算法使用者应具备自主评估和管控算法风险的能力,厘清算法或模型的局限性及可能带来的风险,预设风险应对方案。3)算法应用需先经测试。智能产品进入学校,需经过情境测试和教育效果评估,接受家长、教师、管理者和专家的审查。可见,人工智能教育大脑为凝聚教育管理者、教师、学生、家长等的集体智慧提供了接口,从“算法设计、算法选择、算法应用”三个层面构建算法伦理监管机制,为教育数据治理提供算法伦理保障。

[参考文献]

- [1] Abraham, R., Schneider, J., & Brocke, J.(2019). Data governance: A conceptual framework, structured review, and research agenda[J]. *International Journal of Information Management*, 49(Dec.): 424-438.
- [2] Carretero, AG., Gualo, F., & Caballero, I.(2016). MAMD 2.0: Environment for data quality processes implantation based on ISO 8000-6X and ISO/IEC 33000[J]. *Computer Standards & Interfaces*, 54(PT.3): 139-151.
- [3] 陈雅云,郭胜男,马晓玲,吴永和(2021). 数智融合时代学习分析技术的演进、贡献与展望:2021学习分析与知识国际会议(LAK)评述[J]. *远程教育杂志*, 39(4): 3-15.
- [4] DAMA International (2009). *The DAMA guide to the data management body of knowledge*[M]. New Jersey: Technics Publications: 37-62.
- [5] Gajbe, S. B., Tiwari, A., & Singh, R. K. (2021). Evaluation and analysis of data management plan tools: A parametric approach[J]. *Information Processing & Management*, 58(3):102480.
- [6] 顾小清,李世瑾(2021). 人工智能教育大脑:以数据驱动教育治理与教学创新的技术框架[J]. *中国电化教育*, (1): 80-88.
- [7] 何怀宏(2018). 奇点临近:福音还是噩耗:人工智能可能带来的最大挑战[J]. *探索与争鸣*, (11): 50-59+117.
- [8] 胡坚波(2021). 关于城市大脑未来形态的思考[J]. *人民论坛·学术前沿*, (9): 50-57.
- [9] Janssen, M., Brous, P., Estevez, E., Barbosa, LS., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence[J]. *Government Information Quarterly*, 37(3).
- [10] 梁芷铭(2015). 大数据治理:国家治理能力现代化的应有之义[J]. *吉首大学学报(社会科学版)*, 36(2): 34-41.
- [11] 刘金松(2018). 数据治理:高等教育治理工具转型研究[J]. *中国电化教育*, (12): 39-45.
- [12] Singh, D., & Singh, B. (2022). Feature wise normalization: An effective way of normalizing data[J]. *Pattern Recognition*, 122: 108307.
- [13] 桑尼尔·索瑞斯(2014). 大数据治理[M]. 匡斌译. 北京:清华大学出版社: 286.
- [14] 谭维智(2019). 人工智能教育应用的算法风险[J]. *开放教育研究*, 25(6): 20-30.

- [15] 田贤鹏(2020). 隐私保护与开放共享: 人工智能时代的教育数据治理变革 [J]. 电化教育研究, 41 (5): 33-38.
- [16] 吴永和, 郭胜男, 朱丽娟, 马晓玲(2021). 多模态学习融合分析(MLFA)研究: 学理阐述、模型样态与应用路径 [J]. 远程教育杂志, 39 (3): 32-41.
- [17] 肖凤翔, 张双志(2020). 算法教育治理: 技术逻辑, 风险挑战与公共政策 [J]. 中国电化教育, (1): 76-84.
- [18] 徐宗本(2017). 大数据应用价值与挑战并存 [J]. 中国战略新兴产业, (25): 96.
- [19] 赵磊磊, 张黎, 代蕊华(2021). 智能时代教育数据风险治理: 实然困境与实践路径 [J]. 湖南师范大学教育科学学报, 20 (6): 94-102.
- [20] 周志华(2020). “数据、算法、算力”人工智能三要素, 在未来要加上“知识” [R]. 2020 全球人工智能和机器人峰会.
- [21] 中华人民共和国教育部(2017). 国务院关于印发国家教育事业发展“十三五”规划的通知 [EB/OL]. [2017-01-10]. http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/201701/t20170119_295319.html.
- [22] 张治, 徐冰冰(2022). 人工智能教育大脑的生态架构和应用场景 [J]. 开放教育研究, 28 (2): 64-72.
- [23] 祝智庭, 孙妍妍, 彭红超(2017). 解读教育大数据的文化意蕴 [J]. 电化教育研究, 38 (1): 28-36.
- [24] 祝智庭, 郑浩, 谢丽君, 吴慧娜, 吴永和(2022). 新基建赋能教育数字转型的需求分析与行动建议 [J]. 开放教育研究, 28 (2): 22-33.

(编辑: 魏志慧)

A New Approach of Educational Data Governance in the Intelligent Age: the Construction and Practice of Artificial Intelligence Educational Brain Model

GUO Shengnan¹, WU Yonghe¹ & ZHANG Zhi²

(1. Department of Education Information Technology, Faculty of Education, East China Normal University, Shanghai 200062, China; 2. Education Bureau of Baoshan District, Shanghai 201999, China)

Abstract: Artificial intelligence educational brain can be regarded as a complex, intelligent education system like the human brain by integrating a new generation of intelligent technology, which can provide a new approach to solving the data governance problem in education. In the intelligent age, educational data governance is facing many issues, such as quality, security risks, and algorithm ethics, which seriously limit the value of educational big data. Therefore, inspired by the Internet brain and urban brain, this paper deconstructs the connotation, difficulties and causes of educational data governance, and constructs the artificial intelligence educational brain model from three aspects: The educational brain structure model, lightweight educational brain information model, and the core module of the educational brain, so as to clarify the internal structure, information flow and technical key of the educational brain. Based on this, this paper puts forward the practical path of the educational brain empowering educational data governance from three aspects: establishing data management organization structure, innovating security risk prevention system, and constructing algorithm ethics supervision mechanism. AI educational brain strengthens the coupling between educational data governance system and intelligent technology by systematically integrating educational data, intelligent algorithms, and computing resources, ensures the quality and safety of educational data, and gives full play to the "digital dividend" of big data in educational governance, so as to provide the new solving ideas for continuously promoting the digital transformation of education and realizing the modernization of educational governance.

Key words: artificial intelligence; educational brain; educational data governance; educational governance; digital transformation of education