

基于ECD的学生信息素养评价研究框架*

朱莎¹, 吴砥¹, 杨浩^{1,2}, 孙泽军¹, 余丽芹¹, 杨洒¹

(1.华中师范大学 国家数字化学习工程技术研究中心, 湖北 武汉 430079; 2.美国纽约州立大学奥斯威格分校 教育学院, 美国 纽约 13126)

摘要: 学生信息素养评价是全面提升学生信息素养水平的重要前提, 而获取学生信息素养真实表现相关的过程性数据是实现精准评价的重要依据。针对传统的学生信息素养评价范式中存在的评价内容有限、评价结果效率较低, 且缺乏对过程性数据的采集与分析等问题, 该研究基于“证据中心设计”(Evidence-Centered Design, ECD)的核心思想, 采用理论研究与实践应用相结合的方法, 设计多类型的复杂情境任务, 基于xAPI采集与存储复杂、细粒度的过程性数据, 利用机器学习算法挖掘学生信息素养评价指标与过程性数据的关联机制, 识别并提取学生信息素养行为特征变量, 构建贝叶斯网络并训练样本数据, 实现数据驱动的学生信息素养精准评价, 从而为培育与提升学生信息素养提供决策参考。

关键词: 证据中心设计; 学生信息素养评价; 情境任务; 机器学习; 贝叶斯网络

中图分类号: G434

文献标识码: A

信息素养是融合意识、知识、能力、态度与价值观的综合素养, 是信息社会公民的基本素养, 是人们应对信息化环境下的生活、学习所必备的核心素养之一。当前, 大数据、人工智能等新技术的发展正不断重塑社会生活和学习的新形态, 提升学生信息素养水平对实现我国从信息社会向智能社会的转型升级、促进创新人才培养以及贯彻落实我国教育信息化相关政策方针至关重要。义务教育阶段学生信息素养的监测、评价与培养, 是新时期我国义务教育阶段的重要任务。2018年, 教育部发布《教育信息化2.0行动计划》, 明确提出“制定学生信息素养评价指标体系, 开展规模化测评, 实施有针对性的培养和培训”^[1]。2020年, 教育部发布了《2020年教育信息化和网络安全工作要点》, 提出“大力实施信息素养全面提升行动”的工作任务, 即“培养提升学生的信息素养, 发布中小学生信息素养评价指标体系, 在‘智慧教育示范区’创建区域开展评测工作”^[2]。2020年初疫情防控期间大规模在线教学活动对学生信息素养提出了较高的要求, 这进一步凸显了学生信息素养培育提升工作的重要性和紧迫性^[3]。

学生信息素养评价可以量化反映学生信息素养培养的成效与不足, 是学生信息素养水平提升的关键环节, 是制定学生信息素养全面提升计划的基础和前提。目前, 学生信息素养评价主要沿用传统的教育测量与评价范式, 评价方法以标准化测试及自陈式量表为主, 其评价结果效率低、解释力度不强, 难以准确把握学生信息素养水平^[4]。尽管有少数研究利用情境任务测评系统获取学生信息素养表现数据, 但侧重考察学生的信息获取、处理、应用等知识与低阶能力, 对智能时代需要的信息创造、问题解决、计算思维等高阶能力的评价较少涉及, 更难获取与信息素养相关的态度、价值观方面的数据。总体而言, 现有的学生信息素养评价基于“依据规则给潜在特质赋值”的经典测量观, 采用的是传统的“试题—答案”的测验设计范式, 缺乏对复杂、细粒度的过程性数据的采集与分析, 无法实现精准评价, 难以为学生信息素养培育和提升提供决策支持。

随着大数据、人工智能技术在教育领域的日益发展与融合, 数据驱动的学生综合素质评价成为当前教育测量与评价领域的主要趋势。“证据中心

* 本文系中央高校基本科研业务费专项资金“基于证据中心设计理论的学生信息素养评价研究”(项目编号: CCNU20XJ014)、国家重点研发计划“基于大数据分析挖掘技术及其智慧教育示范应用”项目课题“基于大数据的教学效果评价技术”(课题编号: 2018YFB1004504)阶段性研究成果。



设计”(Evidence-Centered Design, ECD)的理论强调“基于证据进行推理”,力图构建复杂的任务情境,获取多类型的过程性数据,致力于解决复杂测验设计的理论问题,适用于高度抽象、结构复杂的高阶思维能力的评价,是数据驱动的新型教育评价的一种范式。本研究拟基于ECD,探索多类型的评价任务如何诱发学生在情境任务中的真实反应,如何挖掘信息素养评价指标与关键过程性数据的关联机制,旨在构建过程性数据驱动的学生信息素养自动评价机制,从而精准、客观、全面地评价学生信息素养水平,为学生信息素养水平提升提供决策支持。

一、学生信息素养评价的研究现状

(一)信息素养的内涵与组成要素

1974年,美国信息产业协会主席保罗·泽考斯基(Paul Zurkowski)最早提出信息素养的概念,自此之后,信息素养的概念在世界范围内得到了广泛传播和使用,其内涵得到了不断的延伸和发展。信息素养的内涵演变大致可划分为:20世纪70年代以前的萌芽阶段,强调手工检索文献的技能^[5];20世纪70年代至80年代末的发展阶段,强调运用信息工具及主要信息资源解决问题的技术和能力^[6];20世纪90年代以后的成熟阶段,人文属性得到彰显,批判、评价信息的能力被充分重视^[7]。随着信息社会的快速发展,信息素养被赋予了新的时代内涵,信息意识以及信息社会的伦理、道德被充分重视。

在20世纪90年代中期,信息素养的概念被引入国内。王吉庆认为信息素养是一种在信息社会中获取信息、利用信息、开发信息方面的修养和能力^[8];陈维维等认为信息素养是对信息活动的态度以及对信息的获取、分析、加工、评价、创新、传播等方面的能力^[9];钟志贤等认为信息素养是一种为实现特定目的而合理合法地利用各种信息工具的能力^[10];吴砥等提出信息素养是信息社会的公民适应信息化环境下的学习、工作和生活所需的基本素养,同时也是中小学生必备的核心素养之一^[11]。近年来,人工智能技术推动人类社会从信息时代向智能时代发展,信息素养被赋予了新的时代内涵。熊璋认为编程能力、计算思维及对智能化社会的深度认知,是人工智能时代学生信息素养内涵的重要内容^[12];陈凯泉等提出智能化社会要求学生的信息素养内涵进行转型与再造,其中计算思维、编程能力和人机协同能力是核心^[13];于晓雅提出,智能时代的信息素养应强调利用信息技术解决问题的思维和行为以及应对智能社会的伦理道德问题的

能力等^[14]。

关于信息素养组成要素,国外比较典型的有:(1)二要素说,包括认识信息环境、参与和丰富信息两个要素^[15]。(2)三要素说,包括信息行为、信息获取、信息道德三个要素^[16]。(3)四要素说,包含寻找和获取利用信息、信息意识、信息社会责任和信息伦理四个要素^[17];或者包括信息检索、信息获取和交流、信息传播、信息重组四个要素^[18]。(4)七要素说,包括信息技术经验、获取信息源、信息处理、信息控制、信息知识建设、信息知识延伸、信息智慧体验七个要素^[19]。(5)九要素说,包括认识信息、确认信息完整、确定潜在信息来源、信息搜索策略、获取信息来源、评价信息、整理信息、整合信息、运用信息九个要素^[20]。国内学者对信息素养组成要素的看法主要包含:(1)二要素说,包含信息意识和信息能力^{[21][22]}。(2)三要素说,由信息知识、信息意识和信息能力构成^{[23][24]};或者由信息意识、信息能力和信息伦理道德构成^{[25][26]}。(3)四要素说,包含信息知识、信息意识、信息能力和信息道德^[27-29]。

(二)学生信息素养评价方法研究

标准化测试、基于自陈式量表的评价是应用较早的学生信息素养评价方法。典型的标准化测试工具包括美国肯特州立大学设计的大学生信息素养测评工具(SAILS)和中小学生学习素养测评工具(TRAILS)。还有针对特定对象,如医学生的信息素养标准化测评工具。标准化测试尽管是了解信息素养的最佳评价方法^[30],却不能有效评价学生在真实场景中的信息素养能力^[31]。基于自陈式量表的评价主要根据被试的主观报告来确定其信息素养水平,例如PISA的ICT素养测评中采用自陈式量表评价学生ICT的熟练程度,以及其他研究者自主研制信息素养量表。

鉴于基于自陈式量表评价的结果具有主观性,其测试结构与其实际水平之间的相关性较低等问题。近年来研究者开始关注学生在真实场景的信息素养表现,并采用电子档案袋、基于计算机的情境任务等对学生信息素养进行过程性评价。有研究者设计信息素养能力档案袋评价学生的信息、媒体和技术技能或信息素养水平^[32],还有研究者综合采用书目分析与档案袋评价学生的信息素养^[33]。利用档案袋收集学生的信息作品与测评反思,反映了学生信息素养水平的提升过程^[34],能够考查学生对信息知识或技能的掌握程度,还能对学生的信息思维、信息意识等非智力因素进行评价^[35]。然而,档案袋评价需要教师投入大量时间和精力,且难以制定客

观、公正的评价标准,限制了档案袋评价在学生信息素养评价中的应用^[36]。

基于计算机的情境任务测评系统采集并分析学生完成任务时的过程性数据,从而评价学生的知识、技能与能力。根据任务呈现的形式,情境任务一般包含以下三类:(1)基于信息应答的任务,即简单建构响应任务以及拖拽响应任务等,适合评价信息素养相关的知识和简单技能等^{[37][38]},其测评结果比较客观、准确,有利于评分和比较。(2)模拟真实情境的任务,即模拟真实情境设计的技能操作任务与作品创作任务,适合评价学生的软件应用技能以及检索、选择、评价、转换、创造、交流与共享信息的能力^[39],可提供关于任务过程与结果的丰富观测数据。(3)基于游戏的任务,即在游戏化的活动中设置的任务,其具有丰富性、可玩性、仿真性等特点,适合评价学生的信息素养知识^[40],信息检索等操作技能^[41],以及问题解决能力、逻辑推理能力等高阶思维能力^{[42][43]}。

(三)目前学生信息素养评价的局限性

作为21世纪核心素养的重要组成部分,学生信息素养成为了当前倍受关注的热点问题,其评价也受到国内外学者的广泛重视。现有研究在学生信息素养评价领域取得了较丰硕的成果,为本研究奠定了坚实的研究基础。但是也具有一定的局限性。具体表现为:(1)现有研究局限于信息素养相关的知识和低阶技能的评价,对信息思维、信息行为、信息创造、问题解决等高阶思维能力的评价较少涉及,缺乏一套面向智能时代的学生信息素养评价标准;(2)现有研究主要采用标准化测验、自陈式量表等评价工具,忽视了对学生信息素养评价指标与关键过程性数据的关联机制挖掘;(3)现有的研究已将ECD广泛应用于核心素养、计算思维等高阶思维能力的评价,但是在学生信息素养评价领域的应用极其少见,基于ECD的学生信息素养自动评价技术尚处于探索阶段;(4)现有研究局限于对学生信息素养水平的监测、评价,忽视了对学生存在问题的诊断和动态反馈,难以指导学生信息素养的培育和提升。

二、ECD及其在教育测评中的应用

ECD是一套强调“基于证据进行推理”的系统化评价方法,它力图借助信息技术获取证据,应用数理统计模型对个体的高阶能力进行基于证据的推理^[44]。ECD的概念评价框架包含三个核心模型,分别是学生模型、证据模型和任务模型。学生模型回答“测量什么”的问题,包含需要测量的知识、技能和能力的相关变量及其相互关系,研究者通常采

用文献研究法来确定所测能力的评价指标和对应的可观察行为。证据模型回答“如何测量”的问题,包括证据规则和测量模型。证据规则是指评分标准,描述的是如何从个体完成情境任务的行为过程中抽取观测变量,形成与拟测量的学生能力水平相对应的证据;测量模型是将观测变量的数值和学生模型变量以特定结构和方式联系起来的概率模型^[45]。任务模型回答“用什么测量”的问题,探讨怎样构建复杂的任务情境,刺激学生在评价指标上的行为反应,从而获取证据模型所需的证据。概念模型框架不是一个简单的线性结构^[46],而是一个不断完善、迭代的环形结构。总的来说,基于ECD模型的测验逻辑主要包括三个核心:一是建构情境任务诱发个体行为表现,二是依据表现提取反映个体能力水平的可靠证据,三是依据证据对个体能力水平进行推理^[47]。

证据是ECD的核心,是建立基于学生模型、任务模型和证据模型推理链条的关键所在。证据抽取是基于ECD评价的核心,包含定义证据规则,依据证据规则完成评分,为观测变量赋值。当前研究采用的证据抽取方法主要包含两种:一是依据证据规则编写算法程序^[48],主要针对简单的证据规则;二是利用机器学习、深度学习等高级统计模型从复杂数据中挖掘、推论与需要测量的能力相关的模式与规则,从中识别与提取可靠证据^[49],主要针对复杂的证据规则。基于证据的推论提供学生模型变量和观测变量之间的关联信息,通常采用多维项目反应理论(IRT)模型、诊断分类模型、贝叶斯网络和双因子模型等高级统计测量模型。

ECD已被广泛应用于国际教育测评项目中。PISA、ATC21S、NAEP等许多国际大规模测评项目基于ECD模型开发了在线测评系统,用于评价学生的合作解决问题能力;美国教育考试中心基于ECD开发了计算机自适应(CAT)版本的研究生入学考试(GRE)和iSkills评价项目。此外,国内外学者也将ECD广泛用于核心素养、21世纪技能、合作问题解决能力、计算思维、数据素养、逻辑推理能力等高阶思维能力的评价。

三、基于ECD的学生信息素养评价的研究内容

基于ECD的学生信息素养评价的研究内容主要包含五个方面,即学生信息素养评价标准研究、学生信息素养过程性数据采集与存储研究、学生信息素养观测变量的抽取与处理研究、基于贝叶斯网络的学生信息素养自动评价关键技术研究、学生信息素养自动评价技术的应用效果研究。研究内容框架

如图1所示。

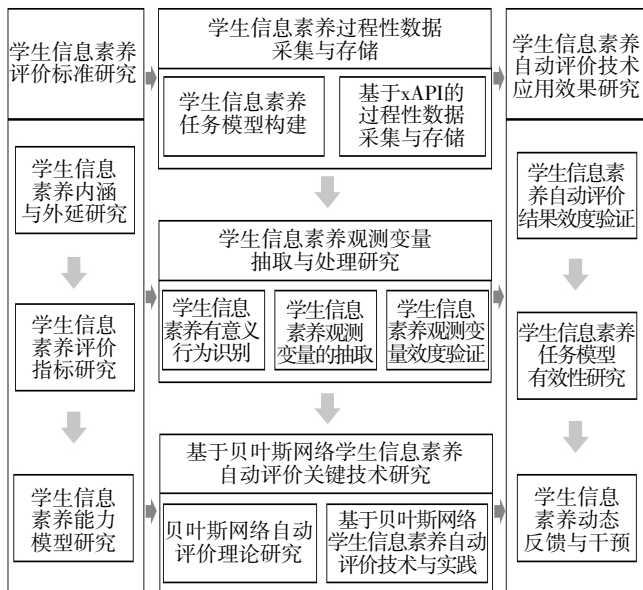


图1 基于ECD的学生信息素养自动评价研究内容框架

(一)学生信息素养评价标准研究

1.学生信息素养评价指标研究

为了确定需要测量的学生信息素养相关的知识、技能和能力的相关变量，需要首先研究学生信息素养评价指标。结合国内外学生信息素养评价指标及国内学生信息素养培育现状与发展需求，研究团队研制了《中小学生信息素养评价指标》，并于2017年6月在CELTS获得正式立项。截至2019年，在全国累计征求了70名不同领域专家的意见，通过多轮修订，标准的框架、内容已得到专家普遍认同。《中小学生信息素养评价指标》包括信息意识与态度、信息知识与技能、信息思维与行为和信息社会责任四个一级指标，每个一级指标分别包含2—3个二级指标。信息意识与态度主要考查学生对信息的感知意识、应用意识和安全意识，以及学生在学习和生活中利用信息与技术的态度；信息知识与技能主要考查中小学生对信息科学知识和信息应用技能的理解和掌握程度；信息思维与行为主要考查学生利用信息技术来辅助自身思维的综合意识与能力倾向，如计算思维、非线性思维、创造性思维、智能化思维等，以及学生使用信息技术的行为与习惯；信息社会责任主要考查学生对信息道德与伦理、信息法律与法规的认识与践行水平。

2.学生信息素养能力模型研究

在学生信息素养评价标准研究基础之上，构建学生信息素养能力模型，即从教育测量的视角，描述需要测量的内容及其相互关系。学生信息素养能

力模型研究包含以下三个关键环节：一是在学生信息素养评价标准的基础上，深入分析学生信息素养相关的知识、技能、情感等领域的评价指标及其相互关系；二是厘清知识、技能和情感领域的学生信息素养评价指标与其能力表现的实质性联系；三是确定每个评价指标对应的不同水平等级的行为表现特征，界定可观察的行为及其操作性定义。表1为初步确定的学生信息素养各评价指标对应的可观察行为。

表1 学生信息素养评价指标及其对应的可观察行为

一级指标	二级指标	可观察行为
信息意识与态度	信息感知意识	能够对信息进行识别、分类 能够利用网络来寻找、筛选和判断信息 能够确定信息来源的正确性和可靠性
	信息应用意识	能够利用信息技术的相关知识与方法解决问题 能够利用思维导图工具开展头脑风暴活动 能够使用信息技术工具辅助学习
	信息安全意识	能够保护自身和他人隐私 能够分辨健康与有害信息
信息知识与技能	信息科学知识	了解各类操作系统、文字处理软件、图形图像处理软件、视音频处理软件等的操作方法 了解信息技术的发展历程、基本现状及未来趋势 掌握信息检索与评测的基本知识与技术、信息分类与存储的方法 掌握动画、视频、课件、网页等数字化学习资源的设计方法
	信息应用技能	能够利用各类搜索引擎和网络平台查找所需信息 能够对信息分类并用表格的形式呈现信息 能够通过各种途径和方法鉴别和分析信息 能够基于特定内容或围绕特定主题，创造有价值的信息资源 能够创造新产品，如主题幻灯片、数据图表等
	信息思维	能够定义及识别信息中的隐含假设，对信息进行演绎推理 能够在信息活动中采用计算机处理问题的方式界定问题、抽象建模、组织数据 能够通过整合资源并运用合理算法构建问题解决方案 能够总结计算机解决问题的过程方法并迁移到相关问题解决
信息思维与行为	信息行为	能够积极使用广泛的交流工具(电子邮件、即时通讯、社交网络)进行在线交流 能够使用协作工具创建和管理内容(如项目管理系统，共享文档等) 能够使用先进的通讯工具与人交流(如视频会议、数据共享、应用共享)
	信息伦理道德	尊重知识，认同信息劳动的价值 遵守网络文明公约，净化网络语言，文明礼貌地学习交流 约束自己的信息道德行为和监督他人信息行为 管理及保护个人资料，尊重他人信息安全
信息社会责任	信息法律法规	遵守有关信息使用的法律和法规 清楚平等访问、存取信息的权力，尊重他人知识产权

(二)学生信息素养过程性数据采集与存储

1.学生信息素养任务模型构建

学生信息素养任务模型旨在通过设置不同类型的任务,刺激学生在信息素养评价指标上的行为反应,因此,构建学生信息素养任务模型是采集学生信息素养过程性数据的基础和前提。根据信息素养评价指标的不同类型,即知识相关指标、技能相关指标和情感相关指标,分别设计三种类型的任务,如表2所示。第一类是信息应答任务,适合评估学生对信息素养相关知识的掌握情况,其测评结果较为客观、准确,有利于评分和比较。如学生在多个主题中选择并拖动相应的选项来完成任务;第二类是模拟真实情景任务,适合评估学生的信息检索、信息处理、信息应用等低层次技能,其为任务测评过程与结果提供复杂观测数据,如利用文字处理软件或图形处理软件制作电子海报、利用搜索引擎搜索信息等;第三类是基于模拟游戏的任务,适合评估学生的高水平技能和情感相关的指标,包括信息沟通、信息评价和信息伦理,其具有丰富性、可玩性、仿真性等特点,如学生与游戏伙伴的合作。

表2 任务样例

任务类型	测评内容	任务特点	任务设计
信息应答任务	信息科学知识、信息法律法规等	测评结果客观、准确	学生单击并拖动相应的选项以完成多项选择项
真实情景任务	信息意识、信息思维、信息行为、信息应用技能、信息感知意识、信息应用意识、信息安全意识等	测评情景真实、融入生活	学生设计并制作一个介绍演讲的学术海报,并与他人分享
模拟游戏任务	信息思维、信息行为、信息应用技能、信息伦理道德等	测评过程丰富、有趣	学生需找出游戏角色在特定场景中的对错行为

2.基于xAPI的过程性数据采集与存储

在任务模型构建的基础上,构建基于xAPI的过程性数据模型,进而对学生测评任务中产生的过程性数据进行采集与存储。如图2所示,该数据模型以Task(任务)为中心,Task的执行者是参与测评的学生(Actor),Actor元素包含学生的身份文档(Learner Profile),身份文档中记录了学生的个人基本信息。Actor元素与Task元素之间补充了任务发生的情境信息(Context)。Task元素由学生的动作(Verb,表示学生的一系列任务操作行为)、对象(Object,行为的执行对象)、工具(Tool,学生在测试中使用工具的情况)、时间戳(Timestamp,动作发生的时间)构成。

结合情境任务测评系统获取数据的可能性,拟采集的学生信息素养过程性数据如表3所示。

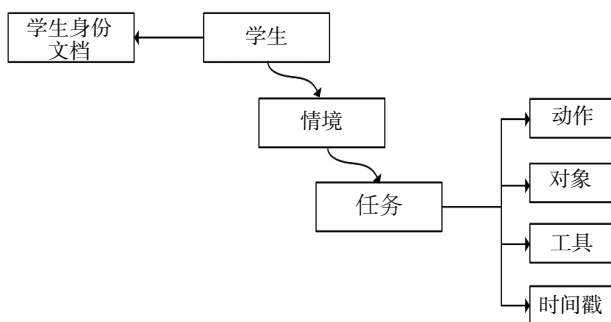


图2 xAPI过程性数据模型

表3 情境任务测评拟采集的过程性数据示例

属性	说明	举例
学生(Actor)	学生的姓名/学号	如:王琳/20200320
动作(Verb)	学生在完成情境任务过程中出现的各种行为动作	如:浏览、单击、保存、选择、键入、查询、标记等
对象(Object)	情境任务中的各个交互对象,即角色、系统、文档、电子表格、演示文稿、图片、视频、音乐等	如:中小学生信息素养测评系统、文档、电子表格、演示文稿、图片、视频、音乐等
情境(Context)	提供任务的上下文情境	如:扮演一名社团统计员,完成期末考试成绩统计任务
时间戳(Timestamp)	记录行为动作发生的时间	如:0:00:00—0:00:22; 0:00:23—0:00:47

(三)学生信息素养观测变量的抽取与处理

在情境任务测评系统中,学生与测评系统、情境任务交互的多类型、细粒度、全过程数据都会被自动采集、记录下来,对这些复杂的数据进行梳理、整合和汇聚,挖掘学生信息素养评价指标与过程性数据的关联机制,抽取学生信息素养行为特征变量,是精准评价学生信息素养水平的前提。学生信息素养观测变量的抽取与处理,包括图3所示的三个关键步骤:一是识别有意义行为;二是抽取观测变量;三是观测变量的效度验证。

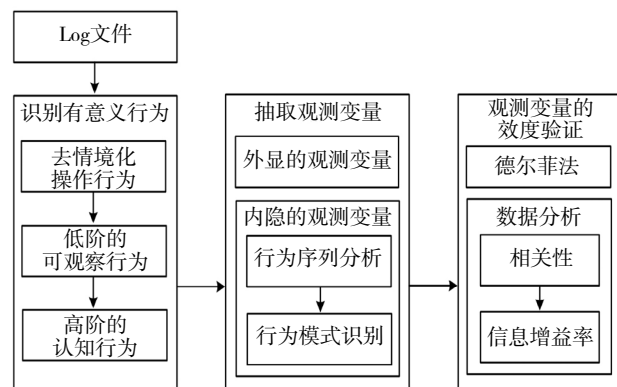


图3 学生信息素养观测变量的抽取与处理流程

1.学生信息素养有意义行为识别

在情境任务测评的过程中,学生的信息素养集



中体现于一些关键的有意义行为。然而,由于测评环境和被试学生行为习惯的复杂性,有意义行为的表现极其复杂。因此,如何在情境任务测评系统记录的过程性数据中识别这些有意义行为是本研究面临的一个关键问题。学生在各类任务中有意义行为的识别可视为一种序列分类问题,即判断给定的行为序列是否属于某一个类别。我们将Log文件中记录的输入、点击、拖拽等去情境化的操作行为数据提炼为信息查询、任务完成、信息发布等低阶的可观察行为;将学生在测试中表现出的犹豫、猜测等状态行为提炼为对信息的感知、判断、反思等高阶的认知行为。高阶的认知行为很难通过编写统一的算法来对数据进行处理,需要依靠人工识别部分行为,并对相关数据进行标注;再以这部分数据作为训练数据,运用机器学习算法建立模型,在模型达到满意的精度以后进行其余数据的自动标注。

BILSTM+CRF模型因其具有对上下文的理解能力而被广泛应用于序列数据的标注。如首先基于xAPI的过程性数据模型将Log数据转换为包含Actor、Verb、Object、Context、Timestamp的向量;其次定义情境中的有意义操作集合,再选取集合内的操作对每一个包含此情境的向量进行人工标注;再次,使用训练数据对模型进行训练,将向量输入BiLSTM模型,并将BiLSTM的输出作为CRF层的输入,识别标签之间的关系,进一步提高预测的准确性;最后,输出向量对应每个操作的概率,选择概率最大的作为最终标签。

2. 学生信息素养观测变量的抽取

通常基于任务结果的观测变量能直接反应被测学生知识、技能的掌握情况,基于任务过程的观测变量可以解释学生完成任务中的思维过程与情感状态。我们可从行为序列抽取基于任务结果的观测变量,如任务完成与否、完成进度、消耗时间等;也可直接通过统计分析、聚类分析、时间序列分析和模式识别等复杂的算法处理获取基于任务过程的观测变量,如低阶的认知行为指标(动作总数、交流的数量等)、高阶的认知行为指标(对信息的感知、判断、反思等)。此外,我们还可通过频繁模式挖掘、因子分析等方法对行为序列模式进行处理,发现学生行为的自然分类并总结其特征。

在基于情境任务的学生信息素养测评中,既包含了有预设正确或最优步骤的结构化操作过程,也包括没有预设正确或最优步骤的非结构化操作过程。结构化过程体现了被试学生信息知识与技能的熟练程度,而在非结构化过程则更能体现被试学生更深层次的思维与习惯。对于数据分析,结构化与

非结构化的过程可抽象为结构化与非结构化的行为序列。对于结构化序列,可利用SPADE算法,采用广度优先搜索(BFS)和深度优先搜索(DFS)策略来产生频繁序列;对于非结构化的行为序列,可先采用用户行为模式检测(UBPD)算法,将每个测试者Log数据按任务进行切分;再按一定的窗口大小和滑动步长转换为子序列集,统计每人每种子序列出现次数,形成用户矩阵;然后,综合方差解释、基序数选择合适的因子数量,并进行因子分析;最后总结每一类共同的基序及这一类测试者的行为模式,解释其意义。

3. 学生信息素养观测变量的效度验证

由于信息素养的复杂性和机器学习算法的不确定性,某些观测变量与信息素养的相关性、可解释性可能较差。同时在以上步骤中,通常会将尽可能多的信息纳入分析,导致观测变量在不同维度分布可能不均衡,甚至出现冗余。基于以上原因,需要对观测变量的有效性进行检验。验证观测变量的有效性可采用主客观相结合的分析方法:一是通过德尔菲法,即专家判断观测变量是否与可观察行为关联恰当,可观察行为的观测变量是否充足;二是通过数据分析,分析基于任务过程的观测变量的值与基于任务结果的观测变量的值的相关性。若基于任务过程的观测变量与基于任务结果的观测变量的相关性低或与其他观测变量存在高度共线性,即认定为质量不佳的观测变量,可考虑删除;接下来计算每个观测变量的信息增益率,结合上一步的分析结果,剔除信息增益率较低的观测变量,从而获得最终的观测变量。由此确定的观测变量能够有效表征学生信息素养,增强其可解释性,同时也能避免观测变量的冗余,这对提高学生信息素养评价结果的准确性至关重要。

(四) 基于贝叶斯网络的学生信息素养自动评价关键技术研究

为了建立学生信息素养评价指标与观测变量之间的关联规则,完成基于证据的学生信息素养推理过程,本研究利用贝叶斯网络实现学生信息素养自动评价。贝叶斯网络具有较强的建模功能以及较完善的推理机制,能够建立证据累积自动评分过程,最终通过有效融合先验知识和当前观察值来完成各种查询、预测与评估^[50]。贝叶斯网络在评估研究和实践中得到了广泛的应用,特别是在ECD理论框架下,研究者大多采用基于贝叶斯网络的自动评分技术。基于贝叶斯网络的自动评分技术主要包含以下几个关键阶段:首先构建能力模型和证据模型的贝叶斯网络片段;其次,确定贝叶斯网络先验概率,并计算条件概率表,构建

完整的贝叶斯网络并训练样本数据;最后,输入新的样本数据更新条件概率表,选择最大概率值作为该条数据的自动评价依据。

在基于贝叶斯网络的学生信息素养自动评价中,首先根据学生信息素养评价标准,分别建立信息意识与态度、信息知识与技能、信息思维与行为、信息社会责任的贝叶斯网络结构片段,明确每个贝叶斯网络结构片段中的可观察行为及其观测变量,并构建对应的能力模型贝叶斯网络片段。根据能力模型贝叶斯网络片段,明确观测变量的证据抽取规则,构建对应的证据模型贝叶斯网络片段。接下来,将贝叶斯网络模型片段进行整合,形成完整的贝叶斯网络结构图。在这个过程中,专家需要结合情境任务的难度、区分度,对每个情境任务下各项操作的难度系数进行评价,确定各个观测变量的参数。然后根据专家确定的变量参数及先验条件概率,确定父节点与子节点的关系以及每个变量的概率估计值,进一步细化和更新更多的观测变量值。最后,可将实证研究中获取的观测变量数据有序输入已训练完毕的贝叶斯网络中。贝叶斯网络将根据数据更新各节点的条件概率表,形成该数据的贝叶斯网络拓扑图,并选择最大概率值作为该条数据的评价依据,从而实现学生信息素养的自动评价。

(五)学生信息素养自动评价技术的应用效果研究

为了验证基于ECD的学生信息素养评价的科学性和有效性,需要开展过程性数据驱动的学生信息素养自动评价实证研究,验证学生信息素养评价结果的效度,分析学生信息素养任务模型的有效性,基于评价结果建立动态反馈与干预策略。

1.学生信息素养自动评价结果效度验证

首先,遴选试点中小学校,采用标准化测试、自陈式量表和情境任务测评的方法同时开展学生信息素养评价实践。基于测评结果,以标准化测试分数和量表分数作为效标分数,分析情境任务测评结果的效标关联效度;依据测评结果,从不同信息素养水平等级学生中分别遴选出5—10名学生进行持续追踪调查,采用课堂观察法分析学生在真实的学习场景中应用信息技术的能力,采用半结构化访谈法让教师评价不同类别学生信息素养的日常表现,分析不同类别学生的测试结果与其实信息素养表现之间的关系,验证情境任务测评结果的区分效度。

2.学生信息素养任务模型有效性研究

学生在完成情境任务时会生成大量的证据数据,这些证据数据可用于验证学生信息素养任务模型的有效性。通过多面Rasch分析得到项目难度值

与学生能力值,将项目难度值与学生能力值比较,考察项目难度是否与大部分学生能力水平匹配。此外,任务难度的分隔指数(Separation)与分隔信度(Reliability)代表了任务之间的难易程度的差异,可据此调整项目难易程度或进行筛选,从而使整个测试具有适中的难度与较好的区分度。

3.学生信息素养动态反馈与干预

为了促进学生信息素养的培育和提升,必须及时反馈学生信息素养评价结果,并给学生提供个性化的信息素养提升建议。基于学生信息素养情境任务测评的结果,利用数据可视化工具将学生信息素养评估的全过程以直观明了的图形、图像信息呈现出来,包括学生信息素养水平的整体评估结果,学生在完成每个情境任务时的可观察行为表现,以及学生在各个观测变量上的量化得分,并向学生提供全面的个人信息素养评估报告;基于学生信息素养评估结果,科学诊断学生信息素养整体水平,深入学生在完成情境任务时表现出的信息素养不足的问题,研究有针对性的、切实可行的学生信息素养培育与提升策略。

四、总结与展望

学生信息素养是一种复杂的高阶能力素养,对其进行科学、精准的评价是极具挑战性的任务。幸运的是,人工智能时代的教育测评理论、数理统计方法与技术的发展为实现学生信息素养精准评价提供了契机。本研究试图构建基于ECD的学生信息素养的新型评价范式,通过设计任务、构建情境、提取证据、建立证据与中小学生信息素养评价指标的推理链条,实现面向复杂任务情境的学生信息素养自动精准评价,从而为培育与提升学生信息素养提供参考。尽管本研究提出了一个理论可行的研究框架,但是本研究在实践过程中可能面临以下几个主要的困难和挑战:一是在初期建立机器学习模型时,可能存在样本量有限、样本不均衡的问题,且初期的数据标注可能大部分需要人工完成;二是应用BILSTM+CRF模型对序列数据标注时,由于数据处理过程较多,不同步骤间的误差可能累积,能否在有限的时间内达到满意的精度将是一大挑战;三是对行为序列挖掘结果进行解释时,如何有效地将行为序列分析结果与有意义行为及其对应的信息素养评价指标建立直接或间接的相关关系,也是一大挑战。下一阶段,研究团队将重点围绕基于过程性数据的学生信息素养特征变量抽取与处理方法,以及学生信息素养评价指标与观测变量的关联规则挖掘展开研究,设计基于复杂情境任务的学生信息素



养测评系统,获取大量实测数据,初步建立过程性数据驱动的学生信息素养精准评价机制。

参考文献:

- [1] 教育部.教育部关于印发《教育信息化2.0行动计划》的通知[EB/OL].http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425_334188.html,2020-06-01.
- [2] 教育部.《2020年教育信息化和网络安全工作要点》[EB/OL].https://www.edu.cn/xxh/focus/rd_xin_wen/202003/t20200303_1714814.shtml,2020-06-01.
- [3] 吴砥,余丽芹等.大规模长周期在线教学对师生信息素养的挑战与提升策略[J].电化教育研究,2020,(5):12-17+26.
- [4][31] Samson,S.Information literacy learning outcomes and student success [J].The Journal of Academic Librarianship,2010,36(3),202-210.
- [5] 柯平.信息素养与信息检索概论[M].天津:南开大学出版社,2005.
- [6] Zurkowski P G.The Information Service Environment Relationships and Priorities [R].Washington,DC:National Commission on Libraries and Information Science.Related Paper No.5,1974.
- [7][20] Doyle C.Outcome Measures for Information Literacy Within the National Education Goals of 1990:Final Report of the National Forum on Information Literacy [R].Washington,DC:US Department of Education.Related ERIC document NO.ED 351033,1992.
- [8][23] 王吉庆.信息素养论[M].上海:上海教育出版社,1999.
- [9] 陈维维,李艺.信息素养的内涵、层次及培养[J].电化教育研究,2002,(11):7-9.
- [10] 钟志贤,汪维富.Web2.0学习文化与信息素养2.0[J].远程教育杂志,2010,28(4):34-40.
- [11] 吴砥,许林等.信息时代的中小学生信息素养评价研究[J].中国电化教育,2018,(8):54-59.
- [12] 熊璋.加强青少年信息素养教育的重要意义[J].国家治理,2016,(3):41-45.
- [13] 陈凯泉,何瑶等.人工智能视域下的信息素养内涵转型及AI教育目标定位:兼论基础教育阶段AI课程与教学实施路径[J].远程教育杂志,2018,(1):61-71.
- [14] 于晓雅.人工智能视域下教师信息素养内涵解析及提升策略研究[J].中国教育学报,2019,(8):70-75.
- [15] Lloyd A.Working (In)formation:Conceptualizing Information Literacy in the Workplace [C].Rockhampton,Queensland:Central Queensland University Press,2003.218-224.
- [16] Johnston B,Webber S.Information Literacy in Higher Education:a review and case study [J].Studies in Higher Education,2003,28(3):335-352.
- [17] Lupton M,Glanville C,et al.Information Literacy Toolkit[DB/OL].http://www.griffith.edu.au/centre/gihe/griffith_graduate,2004-12-11.
- [18] Bawden D.Information and digital literacies:a review of concepts [J].Journal of Documentation,2001,57(2):218-259.
- [19] BRUCE,C.The Seven Faces of Information Literacy [M].Adelaide:Auslib Press,1997.
- [21] 任燕丽.信息与信息素质[J].北京成人教育,1999,(6):43-44.
- [22] 霸桂芳.大学生信息素养的内涵与培养策略研究[J].教育探索,2007,(2):15-16.
- [24] 张亚莉.信息素养内涵的建构[J].图书馆论坛,2005,(5):53-55+117.
- [25] 吕庆阳,刘孝文.国内外信息素质概念的界定[J].河北科技图苑,2008,(2):40-42.
- [26] 钟志贤.面向终身学习:信息素养的内涵、演进与标准[J].中国远程教育,2013,(8):21-29+95.
- [27] 仇诚诚.大学生信息素养及评价标准研究[J].图书情报论坛,2011,(Z1):38-41.
- [28] 夏秋萍,谌新华.高校图书馆提升大学生信息素养的策略研究——大数据时代背景下[J].学理论,2017,(12):181-182.
- [29] 石映辉,彭常玲等.中小学生信息素养评价指标体系研究[J].中国电化教育,2018,(8):78-82+98.
- [30] 卢敏.大学生信息素养评价方法浅析[J].现代情报,2014,34(12):141-144.
- [32] Newell Terrance S.Learning in Simulations:Examining the Effectiveness of Information Literacy Instruction Using Middle School Students' Portfolio Products [J].Evidence Based Library and Information Practice,2010,5(3):20-38.
- [33] Scharf D,Elliott N,et al.Direct Assessment of Information Literacy Using Writing Portfolios [J].The Journal of Academic Librarianship,2007,33(4),462-477.
- [34] 杨阳.“互联网+”时代下新型学生档案袋评价的探索[J].华中师范大学研究生学报,2018,25(3):109-113.
- [35] 魏铁军,吕蕊.教育档案式评价研究[J].教育教学论坛,2011,(9):163-164.
- [36] 江彬,邱立中.科学认识档案袋评价[J].上海教育科研,2003,(11):37-39.
- [37] Katz I R.Testing information literacy in digital environments:ETS's iSkills assessment [J].Information technology and Libraries,2007,26(3):3-12.
- [38] Gobert J D,Kim Y J,et al.Using educational data mining to assess students' skills at designing and conducting experiments within a complex systems microworld [J].Thinking Skills and Creativity,2015,(18):81-90.
- [39] Guo Y R,Goh D H L,et al.Tertiary students' acceptance of a game to teach information literacy [J].Aslib Journal of Information Management,2017,69(1),46-63.
- [40] Siddiq F,Hatlevik O E,et al.Taking a future perspective by learning from the past-A systematic review of assessment instruments that aim to measure primary and secondary school students' ICT literacy [J].Educational Research Review,2016,19:58-84.
- [41] 孙鑫,黎坚等.利用游戏log-file预测学生推理能力和数学成绩——机器学习的应用[J].心理学报,2018,50(7):761-770.
- [42] [英]维克托·迈尔-舍恩伯格,肯尼斯·库克耶,盛杨燕,周涛译.大数据时代:生活、工作与思维[M].杭州:浙江人民出版社,2013.
- [43] 李菲茗,张浩等.计算机交互式评价的开发研究——以科学探究为例[J].中国远程教育,2020,(5):29-37+77.
- [44] Mislevy R J,Almond R G & Lukas J F.A Brief Introduction to Evidence-centered Design [R].Princeton,NJ:ETS Research & Development Division.RR-03-16,2003.
- [45] 张娜,李峰.合作问题解决能力评价的本土化研究——基于以证据为中心的设计模型[J].中国考试,2019,(8):59-65.
- [46] Shute V J,Masduki I & Donmez O.Conceptual framework for modeling, assessing and supporting competencies within game environments [J].Technology,Instruction,Cognition & Learning,2010,8(2):137-161.
- [47] 袁建林,刘红云.核心素养测量:理论依据与实践指向[J].教育研

究,2017,38(7):21-28+36.

- [48] Yuan J, Liu H, et al. Assessment of Collaborative Problem Solving Based on Process Stream Data: A New Paradigm for Extracting Indicators and Modeling Dyad Data [J]. *Frontiers in psychology*, 2019, 10: 1-14.
- [49] Robert J & Mislevy, et al., Design and Discovery in Educational Assessment: Evidence-Centered Design, Psychometrics, and Educational Data Mining [J]. *JEDM-Journal of Educational Data Mining*, 2012, (4): 11-48.
- [50] Pearl, J. Probabilistic reasoning in intelligent systems: networks of plausible inference [M]. San Mateo: Morgan Kaufmann, 1988.

作者简介:

朱莎: 讲师, 博士, 研究方向为信息素养、教育信息化绩效评估(zhusha@mail.ccnu.edu.cn)。

吴砥: 教授, 博士生导师, 研究方向为师生信息素养评价、区校教育信息化评价(wudi@mail.ccnu.edu.cn)。

杨浩: 教授, 博士生导师, 研究方向为数字公民素养、数字化学习理论与实践(harrison.yang@oswego.edu)。

Research Framework of Information Literacy Evaluation of Students Based on Evidence-Centered Design

Zhu Sha¹, Wu Di¹, Yang Hao^{1,2}, Sun Zejun¹, Yu Liqin¹, Yang Sa¹

(1. National Engineering Research Center for E-learning, Central China Normal University, Wuhan 430079, Hubei;
2. School of Education, State University of New York at Oswego, New York, USA 13126)

Abstract: Students' information literacy assessment is the prerequisite for improving students' information literacy comprehensively, while obtaining relevant process data related to the real performance of students' information literacy is an important basis for accurate assessment. Previous study found that the traditional evaluation paradigm of students' information literacy lacked of analysis of process data, which lead to problems such as, limited evaluation content, and low validity of evaluation results, etc. Thus, this study intends to adopt the idea of Evidence-Center Design (ECD), combining theoretical and practical methods, design multiple types of complex situational tasks, and collect and store complex and fine-grained process data based on the xAPI. Then, machine learning algorithm is used to mine the correlation mechanism between the evaluation index of students' information literacy and process data, and identify and extract the behavioral characteristic variables of students' information literacy, then build Bayesian network and train sample data to achieve data-driven evaluation of students' information literacy, so as to provide decision-making reference for cultivating and improving students' information literacy.

Keywords: evidence-centered design; students' information literacy assessment; situational task; machine learning; bayesian network

收稿日期: 2020年5月5日

责任编辑: 邢西深