

教育评估信息采集与处理

吴新林

(上海市教育评估院教育评估研究所, 上海 200031)

摘要: 基于构建的教育评估信息 DIKW 价值体系, 文章着重分析了采集与处理阶段从数据到信息的价值提升过程中需关注的多个方面, 包括数据采集的原始性、充分性、精准性, 信息处理的导向化、标准化等, 并就当前工作中存在的一些问题及应对策略和技术试作探讨与展望。

关键词: 教育评估; 数据采集; 信息处理

中图分类号: G40-058.1 **文献标识码:** A **文章编号:** 2095-3380(2020)02-0050-06

Information Collection and Processing in Education Evaluation

Wu Xinlin

(Education Evaluation Research Department, Shanghai Education Evaluation Institute, Shanghai 200031)

Abstract: Based on DIKW pyramid system of education evaluation, this paper focuses on the analysis of the important aspects in the process of data collection and information processing to improve data value - originality, sufficiency and accuracy of data collection, as well as the guidance and standardization of information processing. The paper also discusses existing problems and the corresponding coping strategies.

Keywords: Education evaluation; Data collection; Information processing

当前我国推进教育现代化的步伐不断加快, 教育评估作为实施教育质量保障的重要手段和构建现代教育治理体系的重要组成部分, 受到社会各界的高度关注。同时, 随着教育实证研究文化的兴起, “以事实为依据、用数据说话” 日益深入人心, 教育评估领域也越来越注重通过数据来提升评价的科学性和决策的精准性。由此, 评估信息的采集与处理对教育评估工作的重要影响愈发显现。

实践证明, 真实、准确、完整的评估信息是做出客观、公正、合理的价值判断的前提, 是开展评估工作的基础所在; 而明确、充分的数据采集方式和科学、规范的信息处理手段, 是取得真实可靠的评估信息的有效保障。评估信息的采集与处理工作的

成效已成为影响教育评估的信度和效度、决定评价结果可靠与否的决定性因素之一。^[1,2]

一、评估信息的价值体系

信息是用于表征物质客体的成分、结构、状态、特性、行为、功能和演化趋势等。评估信息则是在评估活动中, 用于有意识地认识评估对象, 反映评估对象的客观情况。

狭义上的评估信息, 主要指通过对评估对象进行分析、研究、了解、认识后, 选择数据、文字、图像、形态等恰当的表现形式, 按照一定的组合结构, 来表现、显示、代示、启示评估对象的某一属性。^[3]

而在实际工作中, 评估信息的范畴往往更为泛

化,不仅包括评估对象的具体客观属性,还扩展到经过一定加工处理、已被赋予逻辑判断和价值导向的各类过程性、结果性信息。教育评估信息的概念或可延伸至整个教育活动中所产生的、满足教育评估需求且可采集的、一切可用于价值判断的数据集合。

为了更好地讨论信息采集与处理工作,有必要理清数据和信息的概念,参照信息科学中 DIKW 体系(关于数据、信息、知识及智慧的体系),广义上的评估信息根据其价值体系可以分为四个层次。

数据层次:接近狭义上评估信息的概念,仅指对评估对象特定属性的客观表征,包括对评估对象的数量、位置或关系的抽象表示。它可以被保存、传递和处理。

信息层次:按照评估目的、方案、方法和评估指标的要求,通过对相关数据的萃取、分析、处理,所形成的对评估对象某一方面情况的评价或可供直接评价的依据。它可以是基于某种算法逻辑直接计算得出的数值,也可以是专家通过经验做出的主观判断。但无论何种形式,评估信息都是基于客观数据,又经过一定价值判断所形成的。

知识层次:通过对评估信息的归纳、演绎、比较,对评估对象的相对整体情况做出的具有较强的判断性和结论性的意见。

智慧层次:针对现实问题,基于已有的评估结论、判断意见,通过分析、对比得出相关解决方案的能力。

例如,学生数、教师数是一种数据,生师比是一种信息。普通高等学校的生师比的合格标准,则被认为是一种知识。而上升到智慧层面,某阶段某类学校生师比处在一个较高的数值,成为制约人才培养质量提升的问题之一,应扩大专任教师数的规模。

通常,高价值层次的信息具备更强的抽象性、概括性,从更为宏观、整体的角度反映评价对象,相应地更易进行评价对象间的对比对照。而低价值层次的信息则更为具体、客观,较易快捷、准确、直接地反映出评估对象的某个方面的运行状况。

评估信息的采集与处理,处于数据价值由数据层向信息层提升的阶段,主要为将表征评估对象某一属性的原始客观的数据,转化为符合评估指标导向而被赋予逻辑的信息,为评估提供素材和依据。信息处理可以分为采集、处理和解释等多个环节,采集与处理则是最核心部分,下文将作进一步探讨。

二、数据采集环节的关注重点

数据采集环节,主要基于评估内容(包括评估目的、评估方法、组织方式等),从涉及评估对象的属性表征数据中萃取出相关的数据。

采集环节的着眼点是数据,关键要提高所采集数据的质量。衡量数据质量的维度有很多,包括时效性、一致性、准确性、完整性、可靠性、有效性等等。^[4]高质量数据的核心是真实、精确、全面。结合工作实际,教育评估的数据采集环节可从原始性、充分性、精准性等角度切入,予以重点关注。

1. 数据采集的原始性

数据是表征评价对象的客观属性,客观真实地反映评价对象的实际情况。在数据采集阶段,保持被采集数据未经处理的原始形态,^[5]对于后续评估工作选择适当的处理方式、掌握真实可靠的信息,都至关重要。关注采集数据的原始性有主客观两方面的考虑。

(1) 选取源头数据

客观层面,在数据来源选择时追寻信源,选择数据产生最初的源头。缩减数据传输的链条,既便于对数据的追溯、核查,也尽可能地避免了数据在传输过程中丢失、被篡改等失真问题。

例如,一些评奖结果发布后,都会被许多媒体转载,尤其在新媒体快速发展的当下,传播力最广、最易被获取的信息源往往不是最初的发布者。而在转载过程中,由于媒体自身站位角度、内容篇幅、排版效果甚至疏忽大意等原因,有意无意之间数据有被删减、散佚的可能,很难完全保证“拷贝不走

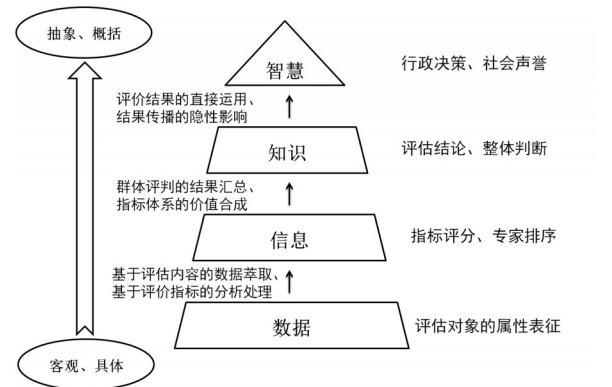


图1 教育评估信息 DIKW 价值体系

样”。此时,选取权威且接近源头的信源,是对数据完整性的最大保障。

(2) 不过早介入处理

主观层面,尽可能不过早地对数据进行处理,保留数据的原始面目。与“信息”的概念相对照,数据本身并不涉及价值判断,而任何处理或多或少都隐含一些价值评判的意味。数据采集环节处于评估信息处理工作的早期,不过早地介入价值评判,有助于避免采集主体由于先入为主、利益驱动等方面的原因对所采集数据的真实性施加影响。

目前评估实践中,数据的采集主体常常是被评对象自身。作为利益直接相关者,其主观态度会对数据采集的准确性造成很大影响。而要求其提供界定明确的原始数据,能在一定程度上避免弄虚作假的现象。

当评估组织者是采集主体时,其好比一个裁判员,在比赛开始前已经预判了比赛的结果,必然会对比赛过程中的判罚施加干扰。在采集环节致力于追求数据的客观、准确,与对数据的清洗、整理工作区分开来,后者留待信息处理环节再作考虑,避免先入为主、未评先判。这也为后续评估中个别价值判断的调整留有余地。

2. 数据采集的充分性

数据采集的充分性包含量大、面广、时长等多方面的角度。规模上,大容量、全方位的采集,提高数据的完整性;范围上,科学的抽样方法、合理的样本分布,确保数据的代表性;时间上,长时段的跟踪、监测,检验数据的时效性、连续性。

(1) 运用新技术手段辅助全方位采集

数据的完整性关心的是是否存在数据缺失情况,及其缺失程度是否在可接受的范围之内。数据缺失包括整条记录的缺失或部分字段的丢失。

随着大数据、云计算、深度学习等新概念、新技术的不断涌现并在各个应用领域取得重大进展,越来越多的学者认为,在新一轮信息技术革命的支撑和驱动下,教育评估将从基于小样本、片段化数据的推测,升级变革到基于全方位、全程化数据的证据性决策。^[6,7]理论上,大容量、全方位的数据采集方式,或将从根本上解决数据完整性的问题。

(2) 采用科学的抽样方法

现实中教育领域涉及的数据项,数据量大、覆盖面广,且非结构化数据占比非常高,显然暂时无法做到对全集数据的采集,抑或即使采集了也不可能对海量数据进行有效、及时的处理,实际往往通过对样本数据的观测来反映总体情况。

一般容易认为,抽样时尽可能多地收集数据,提高样本量,就能更接近真实的总体水平。然而统计学研究表明,大型的非随机样本创建,不一定比样本量较小的随机抽样好。

在抽样工作中,当做不到真正意义上的普查数据、全数据时,科学合理的样本分布要比大样本容量更值得关注。比如,在进行问卷调查时,调查样本抽取的随机性会受时间、空间、调查方式甚至调查者偏好习惯的影响。有时,单纯为了提高问卷的回收量,使得较易采集区域的数据被大量、重复采集,反而会使抽样结果有所偏颇,削弱其代表性。

(3) 坚持长时段的跟踪、监测

数据的时效性描述的是数据的更新程度,它将对分析处理环节产生重大影响。如果数据从建立到分析周期的时间间隔过长,就可能导致由此得出的判断、结论失去现实指导意义。某种程度上,评估就是用过往的数据指导当下、预测未来,这既是评估工作的价值所在,也是一个无奈之处。但随着人工智能、深度学习等技术的发展和在教育评估领域的应用,相信将大大缩短数据的延迟时长(即数据产生到可以被观测的时间),使我们所采集的数据不断趋近当下的状况。

由此,数据采集的时机就显得很重要。现阶段评估实践中,数据采集一般是在评估项目立项后再启动,这会直接对数据的真实性产生影响,增加数据核查的负担。同时,采集的数据也往往只能以某一个时间点或较短周期内的数据为主,无法对数据的长期变化情况进行观测,不利于数据校验和趋势分析。

大数据时代,顺应评价最重要的意图“不是为了证明,而是为了改进”的理念,教育评估的预测功能将发挥更大作用,得到更多重视。对于数据采集来说,从对节点数据的即时采集,转向对长线数据的实时跟踪、监测方向拓展,是大势所趋。

3. 数据采集的精准性

评估数据的本质是对评估对象的一种抽象表示。精准性反映了所采集的数据准确表示评估对象的能力。

在理想的状态下,采集的数据越全面,越能够准确地反映评估对象的状况。但事实上,评估工作实践由于受到采集方式、渠道的制约或出于采集成本(时间成本、劳务成本)的考虑,当评估对象分散、评估内容宽泛时,数据采集很难做到面面俱到,往往是全面与局部的折中妥协。此时,更应在提升数据采集的精准性上下功夫。数据采集的精准性包含准度和精度两层含义。

(1)提高数据采集的准度

准度,指数据采集的无差错率。准确是数据的生命,一组准度无法达到要求的数据,对评估工作的伤害至深。所谓“偏见比无知离真理更远”,不准确的数据,不如没有数据。对于评估组织者,主观上切忌为了在形式上迎合指标要求的全面性,采集无法确保准度的数据。

在具体实践中,可以通过数据间的逻辑校验、平行数据比对、显著性差异分析等技术方法,以及同行间数据公示等举措,加强对所采集数据真伪的甄别。

(2)提升数据采集的精度

精度,即数据采集的细化程度。细化程度越高,数据粒度就越小。数据粒度本是数据库设计时需要考虑的问题,它会对数据库中存储的数据量的大小及其查询、处理的速度产生极大的影响。但随着信息技术的飞速发展和普及,相关存储、处理的成本已经不再是主要问题。

在数据采集环节,应尽可能地提升采集数据的细度。关切的重点是数据细分的方法、界定,以及由此带来的分类数据间的独立性、相关性等问题。一般而言,细粒度数据利于明确界定,提高数据的单纯性,也便于数据核查、校验工作的开展,相应地有助于提高数据的准度。

同时,数据精度的提升也为后续分析处理工作打下了更好的基础。例如,将学生数分解为专科生数、本科生数、硕士研究生数、博士研究生数等,则可在不同指标测算时,赋以相应的折合系数,形成更为合理的折合学生数。

三、信息处理环节的关注重点

基于 DIKW 体系,采集环节的着力点是数据,处理环节则将从数据转向信息。如果说数据采集环节的主要任务是去伪存真,解决数据的准确性问题;到了处理环节,目标则是去粗取精,解决数据的有效性问题的。信息处理是一个选择合适的处理方法、手段、工具等,将客观数据转化为有效信息的价值提升过程。这个环节需要关注的重点,可从导向化、标准化等角度予以切入。

1. 导向化处理

信息是经过加工处理的数据,必然带有一定价值取向和判断。评估数据纷繁芜杂,因此,评估信息是否有效,关键取决于其是否符合教育工作的整体导向,是否利于、易于反映评估指标体系的价值判断。

(1)核心导向,紧扣立德树人根本任务

“培养什么样的人、如何培养人以及为谁培养人”是任何时代教育需要回答的首要问题。我国的国家性质决定了教育必须把培养社会主义建设者和接班人作为教育工作的根本任务。教育评估是教育发展的“指挥棒”,落实立德树人根本任务,也是教育评估工作最核心的导向目标。

在信息处理环节,需要更加重视对能有效反映立德树人落实情况的客观数据的挖掘、梳理,弱化或删除与核心导向不一致、对核心价值观反应不敏感的数据,抑或建立“负面清单”,对负面数据作逆向处理。

(2)多元导向,破除“五唯”,扭转片面量化

2018年9月,习近平总书记在全国教育大会上提出了“扭转不科学的教育评价导向,坚决克服唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾”的要求。破除“五唯”现象的关键是建立科学、全面的评价体系,而落实到评估信息,关键就是确立多元化导向的采集与处理理念。

“五唯”现象的形成,有其操作简单、结果直观的原因,也有一味追求“量化评价”的时代背景,但随着我国教育从普及数量向提升质量方向的转变,单一僵化的数字已无法反映教育日新月异的发展状况。破除“五唯”要求教育评估应该从片面极端的“数字化”回归多元导向的“数据化”。

数据化指通过记录、分析、重组数据,实现对业

务的指导,其核心内涵是对大数据的深刻认识和有效利用。所有对事物定性或定量的描述,都可以是数据。对于评估信息处理来说,要进一步拓展思路,加强对定性数据、非结构化数据的收集、研究,探索对此类数据更为科学、有效的处理方式。

(3) 价值导向,遵循评价指标体系要求

教育评估的本质是一种价值判断,评价指标体系就是为一定的价值判断确立相应的评判标准和尺度。同样一组数据,采用不同处理方式,可能会得出完全不同的结论。因此,信息处理的方法、手段的选取,必须严格遵循评价指标体系的要求,为更好地展现其所蕴含的价值导向服务。不同的价值导向,有着不同的信息处理的方式,大致有以下几种考量。

水平导向与绩效导向:水平导向只考虑产出的数据,不考虑相关的投入情况;绩效导向,则一般采用投入产出比的模型,通过计算单位投入下的产出水平来比较其优劣。

规模导向与效益导向:规模导向一般关注数据的总量;效益导向,则更关注均量的情况。比如,学校的科研经费投入,与专任教师队伍的规模直接相关,测算师均科研经费,更利于比较不同规模学校的实际情况。

历史导向与发展导向:历史导向统计当前所有的存量数据;发展导向更注重一定周期内数据的增量变化。比如学科领域的很多成果,都是长期积累的结果,与所处的基础环境密切相关,但也需要时刻关注发展趋势,以及与同类标杆学科的比较情况。

2. 标准化处理

数据是对所描述对象的一种解释,同时又需要被解释,这种解释方式,即定义,必须是权威、标准、通用的,才有意义。只有建立明确定义的数据界定标准,才有可能掌握可靠、真实的数据。

数据的一致性反映数据质量的一个重要维度,主要考察数据是否遵循了统一的规范、保持了统一的格式,数据是否符合逻辑。数据的标准化是提高数据一致性的重要保障,包含数据本身的标准化和数据关系的标准化两方面的内容。

(1) 建立明确的数据格式规范

数据本身的标准化指的是数据本身具备规范的格式,每一项数据存在它的格式界定,包括数据

项的涵义、时限、量纲等都要清晰明确,尽量减少歧义,避免产生不同的理解。

教育中对某项数据的定义,往往会随时代的变迁、价值取向的调整而有所变化。行政管理部门和理论学术界应当携起手来,建立更为统一、明确的各级各类教育元数据标准。不同来源、不同主体、不同时段采集的数据也有可能存在界定的差异,这些必须在信息处理前予以厘清,进行一致性检验,建立统一标准进行数据清洗,保证数据的可比性。

(2) 进行数据间的相关性分析

数据关系的标准化指的是,多项数据之间存在着固定的逻辑关系,数据间相对稳定的逻辑关系是对其进行分析处理的重要基础。

教育评估中,为了更全面地反映实际发展状况,要同时兼顾数量和质量,同一评价指标中往往涉及多个数据项,各项数据的波幅、量纲、甚至类型均不尽相同。在将多项数据合成时,需要先对数据相关性进行分析,明确数据间的依赖关系,选用合适的效用函数对其进行无量纲化等处理。

四、评估信息采集与处理中的技术展望

社会发展对教育评估的科学性、专业性提出了更高的要求。作为评估工作的基石,评估信息的采集和处理必须与时俱进,实现评估信息的“智能采集”“智能处理”进而走向“智慧评估”。

1. 综合性基础状态数据平台与过程化监测系统

当前数据的采集工作的主体多是评价对象自身,其弊端显而易见。究其原因,主要是可供教育评估领域使用的权威性、综合性基础信息平台及其数据资源不足,数据质量跟不上当前评估工作科学化、精细化的要求。

教育领域应当统筹各方资源,加快权威性、综合性的基础信息平台 and 状态数据库,以及过程化监测系统的建设。采集主体从评估对象向第三方转变,确保数据采集环节的客观、公正。采集时段从实施过程中采集,向独立于项目实施过程转变。数据内容从节点数据向状态数据转变,从结果数据向过程数据转变,数据选择范围从有限性向宽泛性转变,从迫于时间、劳力成本只能有什么数据用什么数据,到

评什么内容选什么数据。这将为评估工作真正走向智能化、精确化、个性化,奠定坚实的数据基石。

2. 物联网感知技术

物联网是把物体与互联网连接起来进行信息交换和通讯的一种网络,正在成为信息采集的重要渠道。

在教育评估工作中,存在着大量的分布范围广、观测成本高、管理难度大的数据,如教学场地、场景的日常状态,设施设备、实验器材的使用情况,等等,这就使得物联网感知技术大有用武之地。通过传感器感知采集点的环境参数,使用电子标签对采集点的数据进行标识,可有效采集场景、设备的状态数据。^[8]运用物联网感知技术将有助于切实推动对更广泛的评估数据的智能化识别、采集、跟踪、监测和评价。

3. 自然语言处理技术

教育领域由于其行为、场景的复杂程度,其非结构化数据占据着重要的地位,但在评估工作中由于受到数据价值密度低、存储体量大、处理成本高等因素的制约,非结构化数据一直不得不被策略性地忽略。

自然语言处理是非结构化数据处理的重要技术之一,主要研究如何让机器对自然语言信息进行处理,包括信息成分的发现、提取、存储、加工、传输等,旨在将人们对语言结构规律的认识用精确的、形式化的、可计算的方式加以表示。^[9]

自然语言处理技术在教育评估领域有着广泛的应用前景。比如,建立各级各类教育专家评审语言的规律模型,对于采用协商论证等方式的专家评审工作,辅助专家口头意见的速记、专家组意见的快速汇总整理、评审意见报告模板创建等。引入自然语言处理技术进行语义学分析,能使统计结果更为有效。

参考文献

- [1] 丁毅强,王战军.构建评估信息的可靠性架构[J].学位与研究生教育,2005(10):42-46.
Ding Yiqiang, Wang Zhanjun. Construction on Reliability Framework of Evaluation Information [J]. Academic Degrees & Graduate Education, 2005(10):42-46.
- [2] 冯晖.教育评价中的敏感性问题研究[J].华东师范大学学报(教育科学版),2007,25(2):37-41.
Feng Hui. Research on Sensitivity Problems in Education Evaluation [J]. Journal of East China Normal University (Educational Sciences), 2007, 25(2):37-41.
- [3] 廖文武,刘文,李冰如.略论专业学位研究生教育评估信息的采集与处理问题[J].当代教育论坛,2014(3):77-83.
Liao Wenwu, Liu Wen, Li Bingru. A Brief Discussion of Professional Degree Education Assessment [J]. Forum on Contemporary Education, 2014(3):77-83.
- [4] 丁小欧,王宏志,张笑影,等.数据质量多种性质的关联关系研究[J].软件学报,2016,27(7):1626-1644.
Ding Xiaou, Wang Hongzhi, Zhang Xiaoying, et al. Association Relationships Study of Multi-Dimensional Data Quality [J]. Journal of Software, 2016, 27(7):1626-1644.
- [5] 沈忠华.新技术视域下的教育大数据与教育评估新探——兼论区块链技术对在线教育评估的影响[J].远程教育杂志,2017(3):31-39.
Shen Zhonghua. A Probe into the Education Big Data and Educational Evaluation from the Perspective of New Technology: Also on the Influence of Block Chain Technology on the Evaluation of Online Education [J]. Journal of Distance Education, 2017(3):31-39.
- [6] 郑燕林,柳海民.大数据在美国教育评价中的应用路径分析[J].中国电化教育,2015(7):25-31.
Zheng Yanlin, Liu Haimin. Path Analysis of the Application of Big Data for Education Evaluation in the United States [J]. China Educational Technology, 2015(7):25-31.
- [7] 冯晖.教育评估现代化的内涵特征与推进策略[J].上海教育评估研究,2019(3):1-4.
Feng Hui. Connotation Characteristics and Promoting Strategy of Modernization of Education Evaluation [J]. Shanghai Journal of Educational Evaluation, 2019(3):1-4.
- [8] 邢蓓蓓,杨现民,李勤生.教育大数据的来源与采集技术[J].现代教育技术,2016(8):14-21.
Xing Beibei, Yang Xianmin, Li Qinseng. Source and Acquisition Technology of Big Data in Education [J]. Modern Educational Technology, 2016(8):14-21.
- [9] 非结构化数据的处理[EB/OL]. (2018-07-01) [2020-02-10]. <https://wenku.baidu.com/view/063f347ef342336c1eb91a37f111f18583d00c75.html>.
Processing Unstructured Data [EB/OL]. (2018-07-01) [2020-02-10]. <https://wenku.baidu.com/view/063f347ef342336c1eb91a37f111f18583d00c75.html>.