

# MOOC环境下数据素养能力评价系统研究

刘爱琴, 王友林, 尚珊

(山西大学经济与管理学院, 太原 030006)

**摘要:** 开放性数据资源丰富民众的信息获取渠道, 但要求信息搜寻人员具备更加全面的数据素养能力, 从而能在内容不断丰富的MOOC平台上获取有价值的信息。本文通过构建数据素养能力评价系统, 采用因子分析法量化各项数据素养能力指标, 对MOOC数据素养教育的开展方式进行深入探究, 并从用户数据素养能力作用测评角度和课程内容评估、个性化推荐两个方面, 阐述数据素养评价系统对MOOC数据素养教育开展的重要意义。

**关键词:** 数据素养; 数据素养教育; MOOC; 个性化推荐

**中图分类号:** G255.73; G203

**DOI:** 10.3772/j.issn.1673-2286.2018.01.011

习近平<sup>[1]</sup>总书记在实施国家大数据战略中共中央政治局第二次集体学习时指出, 善于获取数据、分析数据、运用数据是领导干部做好工作的基本功。随着人民的生产、生活与数据结合得愈加紧密, 在互联网时代的新阶段加强对数据的利用和把握愈加成为一项必备能力。Prado等<sup>[2]</sup>将数据素养定义为, 作为信息素养的组成部分, 数据素养促使个体获取、解释、评估、管理、处理和合理利用数据, 该定义涵盖用户对数据信息的利用方式和途径。作为开展数据素养教育成效检验的有效手段, 数据素养能力评价系统可以衡量数据素养能力水平, 为开展数据素养教育的高校和科研机构指明现有数据素养教育体系中的短板, 并作为应对策略制定时的参考。

大规模开放在线课程 (Massive Open Online Courses, MOOC) 掀起了教育模式、教育体系的一场革命。《纽约时报》将2012年命名为“MOOC元年”<sup>[3]</sup>, 大量开放资源平台 (如Coursera、Udacity、EdX等)、高校 (如麻省理工学院、哈佛大学、耶鲁大学等)、科研机构、企业 (如微软、谷歌等) 都加入这场教育领域的变革。国内教育部门陆续启动“国家精品课程”“国家级精品资源共享课”等一系列有影响力的重大项目。一些门户网站如网易、新浪等都开设了特色鲜明的课程 (如网易公开课、新浪公开课)。清华大学的“学堂在线”和上海交通大学的“好大学在线”等高校MOOC平台不断将高校课程资源融入互

联网开放资源。大部分MOOC平台都收录了相关信息素养的课程。调查显示, 大部分MOOC平台均开设信息素养课程, 其中数据素养课程在被调查的14个平台开设的37门信息素养课程中占比超过50%<sup>[4]</sup>。MOOC平台上的数据素养课程与传统的F2F (Face to Face) 教学方式相比, 不受时间、空间的约束, 具有学习独立性和课程种类丰富的优势, 但缺乏与之相适应的评价系统框架和评价手段。数据素养研究领域仍处于起步和探索阶段, 研究成果较零散。本文结合孟祥保等<sup>[5]</sup>构建的数据素养研究整合框架, 参考国内外研究中衡量数据素养的相关指标, 对高校大学生数据素养调查结果进行量化分析, 以期MOOC数据素养教育提供一套有效的评估标准。

## 1 MOOC环境下数据素养教育研究

传统的数据素养教育方式以课堂授课为主, 不同的专业数据素养教育深入程度不同。数据素养是高校的公共基础课程, 通常由数学、计算机等相关专业的教师负责。教学模式是教师在一定周期内以固定时间、固定地点集中讲授, 对知识点在短时间内进行集中授课, 为防止为压缩授课时间而较少加入富媒体内容。MOOC环境下开展数据素养教育与传统教育模式相比有两大优势: 一是MOOC具有合作化特点, 从课上互动

交流到课后讨论, 都可加强学生对课程内容的理解, 显著提升学习效率; 二是MOOC不受时间、地点、媒介、教材和专职教师等的约束, 可以吸引大规模的学生。

传统教育模式对教材有很强的依赖性, 知识是有结构和层级的, 但对衔接性不强的小知识点, 没有必要开设一门课程进行系统性学习, MOOC可编辑碎片化、模块化的电子教材, 学生可灵活地选取学习内容。MOOC不是由一名专职教师完成全部授课工作, 而是由多名学科教师进行交叉授课。以数据素养课程为例, 其应用涉及数学、计算机等领域, 将不同学校甚至不同国家的教师集中到一起共同完成一门课程的授课工作, 使用户学到数据素养领域知识内容的精髓, 实时了解学科前沿动态。此外, MOOC通过与多种富媒体的有机结合, 突破传统以板书或PPT为主的形式, 使其平台课程可以和Word、PPT、PDF、Video、Photo等多种形式结合, 便于对课程生动形象的讲解, 提升教学质量。

但MOOC平台普遍存在对课程掌控不足, 对课程内容评价方式不明等问题<sup>[6]</sup>。通过评价系统的构建可以有效地解决上述问题。首先, MOOC因缺乏类似传统课

堂的教师监督, 授课教师难以对授课情况有较深切的感受。但随着互联网技术的发展, 实施远程操控摄像并利用人脸识别技术可以解决该问题; 再者通过评价体系的构建, 以结果为导向, 在用户进行相关课程的学习后进行在线考核, 以验证用户是否达到课程的教学目标。其次, MOOC的发展离不开课程证书体系的构建, 建立完善的评价体系对课程内容进行评估, 判定课程难易程度, 量化证书背后所代表的知识价值。

数据素养能力评价系统的构建不仅能提升数据素养课程的教学质量, 还有助于实现个性化课程定制。通过对课程的评价, 可以帮助用户在较短时间内获取适合自身学习能力与学习需求的数据素养课程。在数据素养能力评价系统与MOOC数据素养课程相互作用流程中, 用户在MOOC平台注册并根据自身学习需求选择课程, 通过数据素养能力评价系统对用户的数据素养能力进行测评, 判断是否达到相应的数据素养能力标准: 若未达到则需要进行该能力的课程学习, 学习后再再进行测评, 符合标准后才能授予其课程证书; 若学习后仍未达到标准, 则需进一步学习(见图1)。

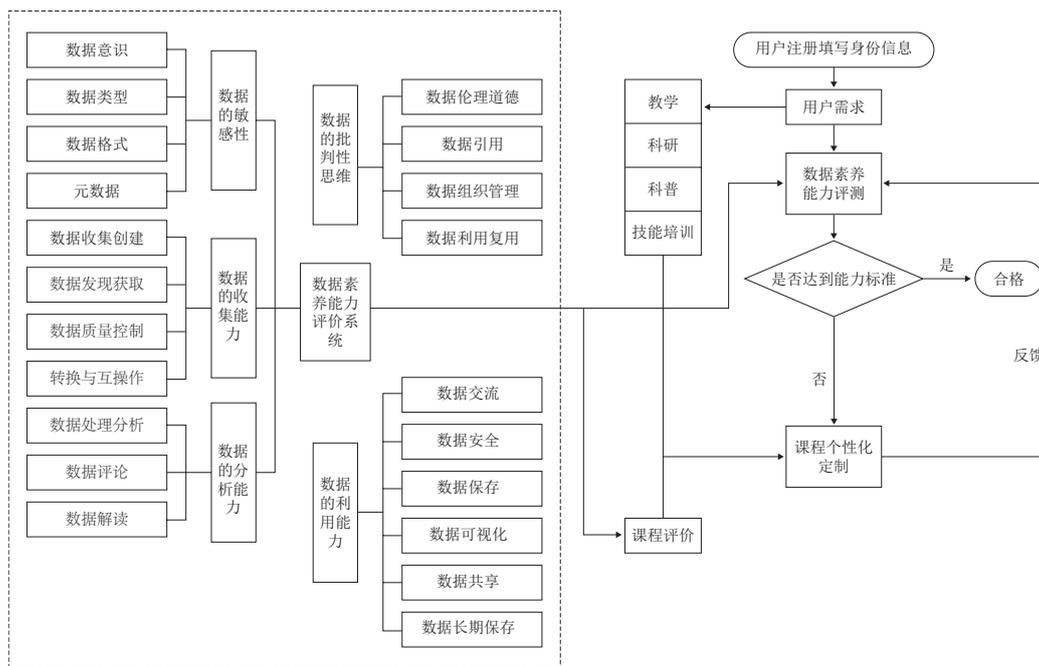


图1 数据素养能力评价系统与MOOC数据素养课程相互作用流程

数据素养能力评价系统用于用户能力评价和课程评价。数据素养能力评价系统对用户数据素养水平进行考量, 通过采取多种评价方式结合多项评价指标对用户数据素养能力进行客观评价, 并根据用户数据素养需求具体量化标准比对用户数据素养评价结果, 指

出用户数据素养能力的短板。在用户完成指定课程学习后, 数据素养评价系统将对用户数据素养能力进行再次测评, 判断用户是否达到要求的数据素养能力标准。数据素养评价系统对数据素养的课程设置起重要作用。数据素养评价系统要根据课程内容和要求, 对每门

课程涉及的数据素养能力指标进行评判,根据课程内容的难易程度,合理设置课程间的递进和衔接程度,做好课程学习前的准备。根据用户测评结果和用户需求,判断其是否需要学习相关准备课程,为用户定制个性化课程,使用户在完成MOOC平台上对应的数据素养课程的学习后,提升相应技能,达到教育目标。

## 2 数据素养能力评价系统指标体系的构成

数据素养能力评价系统指标由不同维度的数据素养知识和技能构成,本文根据胡卉等<sup>[7]</sup>对英美多所高校的数据素养评价模型指标和隆茜<sup>[8]</sup>和郝媛玲等<sup>[9]</sup>对国内高校本科生、研究生数据素养评价模型指标,提取到21种数据指标。具体见表1。

表1 数据素养能力指标因子分析标准权重

一级指标	二级指标	指标序号	标准权重
数据敏感性	数据意识	$X_1$	0.027 6
	数据类型	$X_2$	0.029 0
	数据格式	$X_3$	0.028 5
	元数据	$X_4$	0.028 3
数据收集能力	数据收集创建	$X_5$	0.052 7
	数据发现获取	$X_6$	0.050 4
	数据质量控制	$X_7$	0.050 5
	转换与互操作	$X_8$	0.050 1
数据分析能力	数据处理分析	$X_9$	0.045 1
	数据评估	$X_{10}$	0.046 3
	数据解读	$X_{11}$	0.052 1
数据利用能力	数据交流	$X_{12}$	0.048 3
	数据可视化	$X_{13}$	0.051 7
	数据安全	$X_{14}$	0.042 2
	数据共享	$X_{15}$	0.041 8
	数据保存	$X_{16}$	0.044 0
	数据长期保存	$X_{17}$	0.042 0
数据批判性思维	数据伦理道德	$X_{18}$	0.045 2
	数据引用	$X_{19}$	0.043 3
	数据管理组织	$X_{20}$	0.023 5
	数据利用复用	$X_{21}$	0.025 0
个人因素	学校	$X_{22}$	0.022 8
	城市	$X_{23}$	0.025 1
	专业学科	$X_{24}$	0.027 3
	年级	$X_{25}$	0.029 2
	参与科研情况	$X_{26}$	0.028 0

用户对数据的敏感性指用户面对数据时,在短时间内产生的相应反应,包括掌握数据的来源、类型、格

式甚至数据的意义等;数据收集能力是数据素养能力的重要组成部分,从对数据搜寻与获取到对数据格式与结构的转化,是数据素养能力链条的开端,其对数据分析结果有很强的束缚作用,数据收集能力会限制数据分析的广度和深度;数据分析能力指对数据的处理、分析,及对数据内容解释说明的能力;数据利用能力指用户可以对数据进行文本交流、共享及可视化等一系列操作;数据的批判性思维指在对数据进行利用时的辩证思维能力<sup>[10]</sup>。不同年龄阶段、学术阶段,以及不同的专业学科都会有与之相对应的数据素养能力标准,将个人因素加入数据素养能力评价指标体系十分必要。个人因素包含其所在院校所属的层次、城市级别、年级、专业和科研情况。

## 3 数据素养能力评价系统指标权重的确定

数据素养能力指标的权重采取LIKET五分法的问卷调查加以确定,问卷信度分析结果KOM值为0.787,表明存在显著的相关关系( $\alpha=0.000$ ),数据结果可以进行因子分析。在该数据基础上进行因子分析,经过因子旋转后提取到5个公共因子,其累计方差贡献率为0.902 87,表明该组数据有很强的代表性。将每组公共因子的因子载荷进行处理,获得26个指标的标准权重 $W_j$  ( $j=1, 2, 3, \dots, 25, 26$ ),结果如表1所示。

设 $S_j$  ( $j=1, 2, 3, \dots, 25, 26$ )为LIKET五分法下指标得分, $T$ 为数据素养能力得分( $T \in [1, 5]$ ),分值越靠近5,表明数据素养能力越强,计算方法如公式(1)。

$$T = \sum_{j=1}^{26} W_j \times S_j \quad (1)$$

根据因子分析得出的绝对权重结果,构建系统动力学流图,从中可见指标间影响作用关系和影响历程(见图2)。其中,针对数据素养能力构建的5个子系统间影响关系是间接的;个人因素与5个子系统间都有直接联系,可见个人因素在数据素养能力的体系中有极大的影响力。因而,为保证评价的客观性,评价系统要将个人因素纳入考虑。在课程设置上突出个人差异性,实现课程个性化定制,对用户数据素养能力提升有很大作用。

## 4 数据素养能力评价系统与MOOC数据素养课程个性化推荐

个性化推荐是一种根据用户的特征信息和需求动

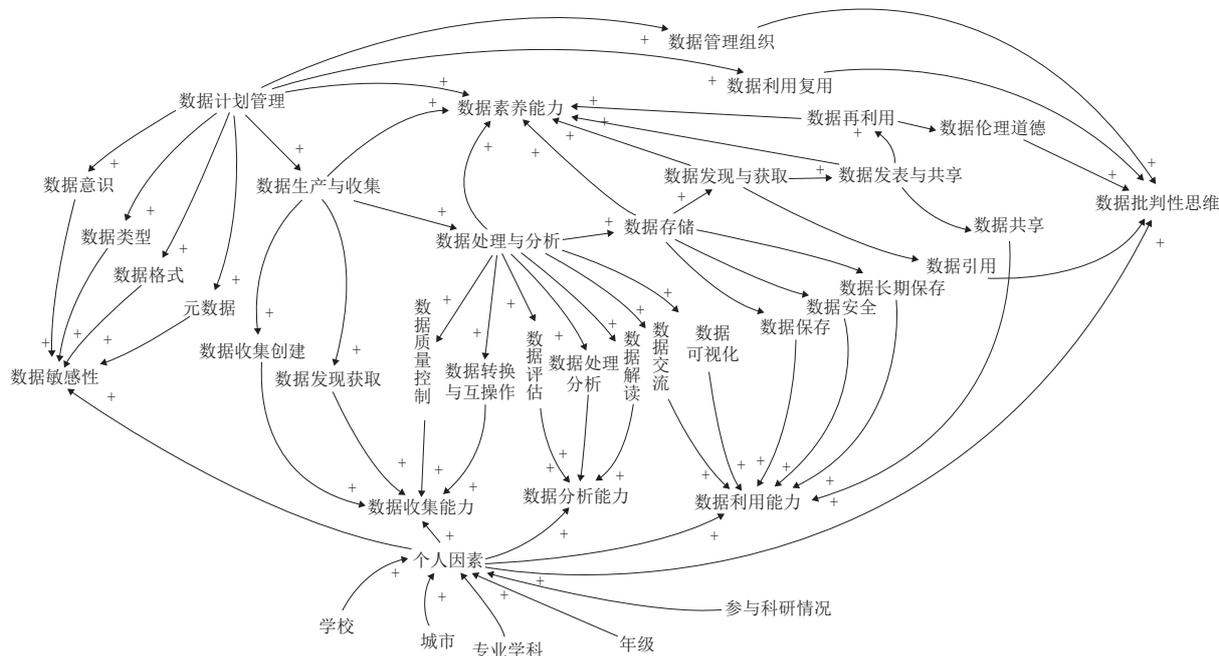


图2 数据素养能力指标系统动力学流图

态,为其推荐信息、资料或服务的方法<sup>[11]</sup>。个性化推荐可以分为记录用户行为模块、用户喜好分析模块和推荐算法模块,用户通过个性化推荐可一定程度地提升检索效率<sup>[12]</sup>。随着MOOC平台用户的增加,越来越多的信息输出者在各MOOC平台发布课程,在没有评价系统和个性化推荐的情况下,用户很难从众多课程中挑选所需课程。截至2017年12月22日,MOOC平台EdX涉及数据科学和统计的课程有181种;平台对课程水平进行划分,其中入门课程30种、进阶课程85种和高级课程66种,课程间缺乏过渡和衔接,少数系统化学习内容的课程在内部需要衔接相关课程的初步学习;涉及的语言有

英语、西班牙语、中文、法语和日语。

用户首次在MOOC平台注册时,平台便记录用户个人信息,并开始跟踪其浏览行为。在进行数据素养课程学习时,引导用户进行数据素养能力测评,将测评结果录入用户个人信息数据库,根据个性化推荐算法(基于MOOC课程内容的推荐算法、基于网络结构的推荐算法和混合推荐算法等<sup>[12]</sup>),完成用户个性化课程的推荐;用户进行学习则匹配完成,用户在课程学习结束后要再次进行测评,更新数据库的测评结果,若达到数据素养能力标准则结束教育课程,否则须继续进行学习,评价系统会反馈出能力指标供用户参考(见图3)。

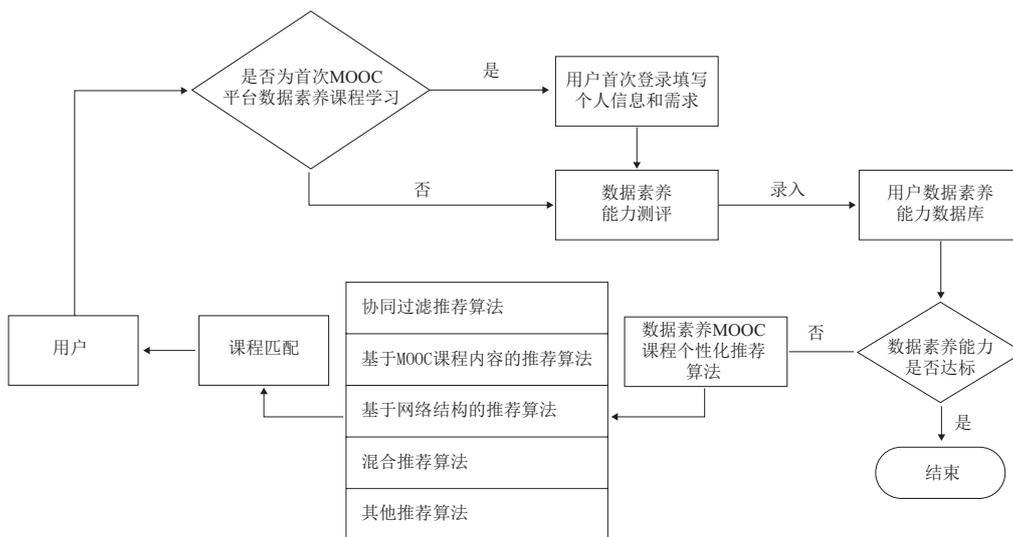


图3 数据素养能力评价系统与MOOC个性化推荐流程

## 5 结语

MOOC的课程教育方式打破传统教育格局,提供更多元和个性化的教育方式。庞大的数据资源,要求高校开展数据素养课程的学习和民众进行数据素养的普及工作。从民众的角度看,可以提高其在生活中辨别数据真伪的基本能力;从高校的角度看,对培养数据时代大众的数据分析能力有积极的影响。数据素养能力评价系统在MOOC数据素养教育中的辅助性作用,丰富MOOC数据素养课程的反馈系统和课程个性化推荐系统,使MOOC平台与用户间形成交互反馈。数据素养能力评价系统的影响与MOOC数据素养课程的作用机制是未来的研究方向。高校数据素养教育的开展还要从师资、MOOC课程资源等多方面进行把握,部分高校缺乏数据素养教育的师资力量或相关资源<sup>[13]</sup>,使高校无法正常开展MOOC数据素养教育,具体需结合MOOC课程高校实际情况进行深入研究。

## 参考文献

- [1] 习近平:实施国家大数据战略加快建设数字中国[EB/OL].(2017-12-09)  
[2017-12-10].[http://news.Xinhuanet.com/politics/leaders/2017-12/09/c\\_1122084706.htm](http://news.Xinhuanet.com/politics/leaders/2017-12/09/c_1122084706.htm).
- [2] PRADO J C, MARZAL M Á. Incorporating data literacy into information literacy programs:core competencies and contents[J].Libri,2013,63(2):123-134.
- [3] PAPPANO L.The Year of the MOOC[N].The New York Times,2012-11-02(1).
- [4] 黄如花,李白杨.MOOC背景下信息素养教育的变革[J].图书情报知识,2015(4):14-25.
- [5] 孟祥保,常娥,叶兰.数据素养研究:源起、现状与展望[J].中国图书馆学报,2016,42(2):109-126.
- [6] 唐菁,方东权,熊婵.由慕课引发的关于高校信息素养教育的思考[J].情报理论与实践,2015,38(10):42-45.
- [7] 胡卉,吴鸣,陈秀娟.英美高校图书馆数据素养教育研究[J].图书与情报,2016(1):62-69.
- [8] 隆茜.数据素养能力指标体系构建及高校师生数据素养能力现状调查与分析[J].图书馆,2015(12):51-56.
- [9] 郝媛玲,沈婷婷.高校文理科研究生数据素养比较[J].图书馆论坛,2016,36(8):126-132.
- [10] 张长海.基于批判性思维和创造力的我国大学生信息素养教育模式研究[J].中国图书馆学报,2016,42(4):102-116.
- [11] 胡国强,张旭虎.基于协同过滤算法的MOOC个性化推荐系统[J].信息技术,2017(6):99-103.
- [12] 刘建国,周涛,汪秉宏.个性化推荐系统的研究进展[J].自然科学进展,2009,19(1):1-15.
- [13] 张新红.MOOC环境下大学生信息素养教育模式研究[J].现代情报,2015,35(12):139-142.

## 作者简介

刘爱琴,女,1974年生,博士,讲师,硕士生导师,研究方向:信息资源管理。  
王友林,男,1993年生,硕士研究生,研究方向:信息资源管理,E-mail:18235115797@163.com。  
尚珊,女,1962年生,教授,研究方向:信息管理、虚拟咨询。

## Research on Data Literacy Evaluation System under MOOC Circumstance

LIU AiQin, WANG YouLin, SHANG Shan  
(School of Economics and Management, Shanxi University, Taiyuan 030006, China)

Abstract: Open data resources enrich people's access to information. It also requires information search personnel to have more comprehensive data literacy capabilities to obtain valuable information from the expanding MOOC platform. Through constructing the evaluation system of data literacy ability, this paper uses factor analysis to quantify the index of each data literacy ability and makes a deep exploration into the combination of data literacy education with MOOC platform. From the data literacy ability in user evaluation and course content evaluation, personalized recommendation of two perspectives, indicating that the data literacy evaluation system for carrying out MOOC data literacy education is of great significance.

Keywords: Data Literacy; Education of Data Literacy; MOOC; Personalized Recommendation

(收稿日期: 2017-12-23)