

基于 DCMM 的高校数据治理能力评估应用与研究*

张国宝

(河海大学, 江苏 南京 210098)

摘要: 针对在高校实施数据治理的评估评价相对方法不足和较少开展的问题,通过文献梳理研究了数据治理评估在相关应用领域研究进展,列举了 5 种主流数据治理模型及其作为数据治理水平评估模型的适用性。选取 DCMM 数据管理能力成熟度模型,面向高校进行数据治理评估指标体系的问卷设计,采用 AHP 层次化方法对问卷数据进行结果分析,展示出高校数据治理能力的建设状况。通过分析说明基于治理模型的高校数据治理评估方法的可行性,并且给出数据治理提升的措施和建议。

关键词: 数据治理;治理评价;治理模型;数据管理能力成熟度评估模型;层次化分析法

中图分类号: TP393

文献标识码: A

DOI: 10.19358/j.issn.2097-1788.2022.04.004

引用格式: 张国宝. 基于 DCMM 的高校数据治理能力评估应用与研究[J].网络安全与数据治理, 2022, 41(4): 26-30.

Application and research on DCMM based evaluation of university data governance capability

Zhang Guobao

(Hohai University, Nanjing 210098, China)

Abstract: Aiming at the problems that the evaluation of implementing data governance in colleges and universities is relatively insufficient and less developed, this paper studies the research progress of data governance evaluation in related application fields through literature review, and lists five mainstream data governance models and illustrates the applicability as a data governance evaluation model. The DCMM data management capability maturity model is selected, the questionnaire design of the data governance evaluation index system is carried out for colleges and universities, and the AHP hierarchical method is used to analyze the results of the questionnaire data, showing the construction status of the data governance capacity in colleges and universities. The analysis shows that the evaluation method of university data governance based on governance model is feasible, and the conclusion and strategy suggestions for improving data governance are given.

Key words: data governance; governance evaluation; governance model; DCMM; AHP

0 引言

2021 年 12 月中央网络安全和信息化委员会印发《“十四五”国家信息化规划》,提出“建立高效利用的数据要素资源体系”“加强数据治理,提升数据资源开发利用水平”。这些政策都对高校的数据治理和管理的建设提出了明确要求。

数据管理和治理能力评估是数字校园建设水平的重要衡量方面。数据管理能力成熟度评估模型^[1]

(Data Management Capability Maturity Assessment Model, DCMM)是我国首个数据管理领域国家标准,将组织内部数据治理管理能力划分为八个能力域。

数据治理模型研究为数据治理实践活动提供理论依据,高校数字校园建设在数据治理理论、框架与模型等方面的研究与实践主要有:许晓东等人^[2]从战略高度提出对高等教育的数据治理,应且必将成为高等教育治理的一个重要组成部分;刘桂锋等人^[3]对国内外数据治理的要素、模型、框架等研究归纳,分析数据治理模型的类型、特点、问题、挑战

* 基金项目: 江苏省教育科学研究所“智慧校园专项课题”(2019-R-75632)

与机遇;余鹏等人^[4]提出设计教育大数据平台与数据治理框架;董晓辉等人^[5]设计了一个高校教育大数据治理的参考框架,刻画高校教育大数据治理的主要内容,并从活动视角提出高校教育数据治理体系构成要素^[6];如孙嘉睿等人^[7]对数据治理的体系架构、数据治理保障梳理分析基础上分析具体业务领域的实践。

除高校数据治理之外,还有其他行业基于 DCMM 或其他模型的治理评估研究,万方等人^[8]在警务领域依据 DCMM 进行了数据治理评估的研究探索;胡成等人^[9]在电力行业基于“AHP+熵权”耦合构建供电企业数据治理成效评价模型。然而高校开展数据治理如何进行有效治理水平的量化评估,推动针对贯彻 DCMM 模型的数据治理,目前可获知的研究还较少,构成本文研究的问题背景。

1 数据治理评估的理论依据

在已有的数据治理模型中,研究关注较多的是五种数据治理模型,即 IBM 数据治理委员会提出的四层要素(核心要素、促成要素、支撑要素、成效)治理能力成熟度模型、国际数据管理协会(DAMA)提出的 DAMA-DMBOK2 数据治理体系模型、国际数据

治理研究所(DGI)提出的十大关键要素的 ORF(组织、规则、流程)流程化模型、企业数据管理协会(EDM Council)发布的 DCAM 模型以及中国电子技术标准化研究院(CESI)牵头制定的 DCMM(数据管理能力成熟度评估)模型(GB/T 36073-2018),如表 1 所示。

根据表 1 所示,五种主流数据治理模型都有涉及对数据治理能力和水平的评估,其中具有量化评估评估参考性的模型中,以 DCMM 较为符合高校的特点。DCMM 从制度设计、组织和人员、技术与服务、标准与规范的维度对数据管理能力进行解构,分解为具体的过程能力项,各个能力项对应划分不同等级。DCMM 能力成熟度模型对高校的数据治理水平评估更具参考性。文献^[12]在数据治理的成熟度模型比较方面进行了模型内容的差异比较研究。

2 高校数据治理评估的问卷设计

为了进行高校的数据治理水平的调查采样,开展高校数据治理评估问卷设计。通过调查采样把高校的数据管理和数据治理的实际现状与数据治理评估模型相结合,以实现对高校的数据治理的建设

表 1 数据治理模型的评估内容比较

数据治理模型	主要内容	相关参考文献资料	治理能力的评估	治理评估的参考性
IBM: 四层要素模型	核心准则包括数据质量、生命周期管理、安全和隐私;支撑准则包括数据架构、元数据和日志审计;促成要素包括组织、管理和政策;成效为风险管理和业务价值;提出 14 个步骤组成的数据治理统一流程理论	IBM 数据治理成熟度评估模型	步骤包括成熟度评估、治理分析、度量结果等内容,有明确的治理水平量化评估模型 CMM, CMM 描述了 5 个等级: 初始化、已管理、已定义、量化管理、持续改进	可整体地量化评估
DAMA: DAMA-DMBOK2 数据治理体系	定义了 11 个主要的管理职能,并通过 7 个环境元素对每个职能进行描述;构建以数据为中心的数据治理组织	DAMA 数据治理框架 ^[10-11]	监控和指导政策合规性、数据使用和管理活动,没有明确的治理水平量化评估	局部定性评估
DGI: ORF 治理模型	DGI 数据治理框架的设计采用 5W1H 法则;数据治理框架的 10 个通用组件构成组织、规则、流程 3 个层次;流程划分为数据治理生命周期 DGLC 方法的 7 个阶段	DGI 数据治理框架	数据治理流程中只包含数据治理的监控、评估和报告	局部定性评估
EDM: DCAM2.0 成熟度模型	DCAM2 将数据管理能力划分为 7 个职能域,分属协作、执行、资金支持 3 个层面;将企业数据管理成熟度划分为 6 个层级	EDM 数据管理能力成熟度评价模型 DCAM 2.0	将企业数据管理成熟度划分 6 个层次: 未启动、概念性、发展性、已定义、已达成、增强型	可整体地量化评估
DCMM: 数据管理能力成熟度评估模型	该标准将组织对象的数据管理划分为八大能力域;对每项能力域进行了二级能力项和成熟度等级划分	标准号: GB/T 36073-2018 ^[11]	对数据管理能力划分为 5 个成熟度等级: 初始级、受管理级、稳健级、量化管理级、优化级	可整体地量化评估

水平予以量化评估。

DCMM (GB/T 36073-2018) 是国家标准化管理委员会于 2018 年发布的现行国家标准,它给出了数据管理能力成熟度评估模型以及相应的成熟度等级,定义了数据战略、数据治理、数据架构、数据应用、数据安全、数据质量、数据标准和数据生存周期 8 个能力域。它适用于组织和机构对数据管理能力成熟度进行评估。通过图 1 所示的成熟度能力指标模型可更直观地看到,8 个不同的能力域可大致划分为三类指标项,即从目标和职责、评估检查、实施应用三个方面进行衡量。衡量数据管理成熟度水平通过指标项的建设实施情况对照评估,建设情况覆盖指标项越多总体的成熟度水平越高,反之建设情况覆盖能力指标项越少则总体的成熟度水平越低,总体可划分为 1 至 5 个等级。

问卷基于 DCMM 并参考全国 DCMM 符合性公共服务平台 (<http://www.dcm.org.cn/>) 的数据治理宣贯要求,设计面向高校数据治理水平的评估指标体系问卷,共 8 个方面 28 项能力指标。面向教育部直属高校和江苏省内高校进行问卷调研,并采用 AHP 方法进行调研数据分析。

3 数据治理评估调研与结果分析

3.1 面向高校的数据治理水平调研

依据 DCMM 设计数据治理水平调研问卷。按照 DCMM 分为 8 个一级指标、28 个二级能力指标,对每个能力指标进行指标权重重要

性 (Weight) 以及该指标的本校建设情况自评评估 (Evaluation) 两个维度的调研。课题组选择调研样本为教育部直属高校和省内高校共计 116 所,通过问卷星平台进行问卷发布和调研,回收有效问卷并统计分析,结果如表 2、表 3 所示。

表 2 调研问卷结果——按学校统计

学校	Σ IndexWeight 指标权重	Σ IndexEvaluation 指标自评评估
U1	60	97
U2	137	56
U3	139	84
U4	139	83
U5	134	90
...	116	81
U39	99	98

表 3 调研问卷结果——按指标统计

Index - Weight	Σ IndexWeight 指标权重 (1~49)	Σ IndexEvaluation 指标评估 (1~39)	同一化一级指标维度
IW1	171	130	数据战略 A1=485
IW2	164	125	数据治理 A2=488.25
IW3	150	112	数据架构 A3=492
IW4	173	122	数据应用 A4=497
IW5	159	117	数据安全 A5=487
...			数据质量 A6=465
			数据标准 A7=483
IW28	157	102	数据生存周期 A8=459.75

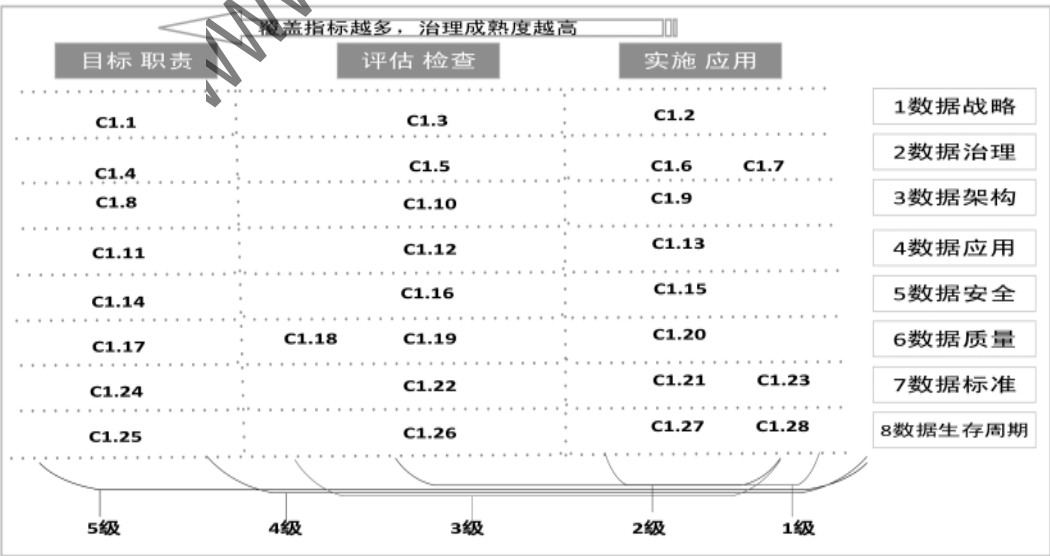


图 1 数据管理成熟度能力指标模型

为简易比较分析,按照一级指标的维度初始化判断矩阵 M ,如下。

$M=$

A11	A12	A13	A14	A15	A16	A17	A18
A21	A22	.	.	.	.	.	.
A31	.	A33	.	.	.	.	.
A41	.	.	A44	.	.	.	.
A51	.	.	.	A55	.	.	.
A61	.	.	.	.	A66	.	.
A71	.	.	.	.	.	A77	.
A81	.	.	.	.	.	.	A88

$=$

1	0.99	0.99	0.98	1	1.04	1	1.05
1.01	1	0.99	0.98	1	1.05	1.01	1.06
1.01	1.01	1	0.99	1.01	1.06	1.02	1.07
1.02	1.02	1.01	1	1.02	1.07	1.03	1.08
1	1	0.99	0.98	1	1.05	1	1.06
0.96	0.95	0.94	0.93	0.95	1	0.96	1.01
1	0.99	0.98	0.97	1	1.04	1	1.05
0.95	0.94	0.93	0.93	0.94	0.99	0.95	1

3.2 数据治理评估结果计算

把每个调研学校看作不同的评估方案,汇总每个学校评估方案的数据作如下分析:

利用和积法计算判断矩阵 M 的最大特征向量为特征向量近似解:

$W=$

$(0.126, 0.127, 0.128, 0.129, 0.126, 0.120, 0.125, 0.119)^T$

(1)

矩阵的最大特征根为:

$$\lambda_{\max} = \sum_{i=1}^n \frac{(BW)_i}{nW_i} = \frac{1.006}{8 \times 0.126} + \frac{1.011}{8 \times 0.127} + \frac{1.02}{8 \times 0.128} + \frac{1.032}{8 \times 0.129} + \frac{1.009}{8 \times 0.126} + \frac{0.962}{8 \times 0.120} + \frac{1.003}{8 \times 0.125} + \frac{0.953}{8 \times 0.119} = 7.995$$

(3)

判断矩阵的一致性指标和随机一致性比率:

$$\begin{cases} C.I. = \frac{\lambda_{\max} - n}{n - 1} = -0.0007 \\ C.R. = \frac{C.I.}{R.I.} = 0 < 0.10 \end{cases}$$

(4)

通过分析 $U1 \sim U39$ 的结果得分验证可知, AHP 方

法分析结果和初始的治理水平评估得分(自评结果)存在差异,结果比较如式(5)所示。由于 AHP 的分析结果结合了其他学校建设情况的因素,学校 $U1$ 较 $U39$ 的数据管理和治理水平更高,结果更具相对合理性。

Evaluation: $U39 > U1 > U5 > U3 > U4 > U2$

AHP: $U1 > U39 > U5 > U3 > U4 > U2$ (5)

另外从调查收集的指标数据分析,在 28 项指标中数据战略制定、数据管理职责明确、数据集成共享架构建设、制定数据服务的标准规范并实施、统一数据分析平台建设分别获得调查高校的更高权重,得分分别为 170、173、170、170 和 167。从高校的自评的数据分析,数据战略制定、数据战略目标实施、数据集成共享架构建设、统一数据分析平台建设、制定数据服务的标准规范并实施的建设得分最高,分别为 0.65、0.625、0.67、0.675、0.625。分析可知,被调查高校认为数据管理职责明确是数据治理最重要的方面,而在实际建设情况中未能很好地实现职责明确,数据集成平台及数据分析实施普遍更好。

4 提升数据治理的建议策略

(1) 从调查评估数据看,高校的数据管理治理职责需进一步明确,如责任到人的数据管理机制、跨部门数据治理的协同机制等。从数据治理体系、治理实施工程的视角看,数据治理工作需要系统化和整体性的构建和实施,不仅包括治理的核心要素(质量管理、安全与隐私管理、生命周期管理)、支撑要素(数据架构、元数据与分类、数据日志和审计),还要包括外围促进要素(组织架构、政策、数据权属)等方面。文献[13][14]对高校的数据治理的策略也在具体分析基础上给出了数据治理的参考经验。结合具体治理实践,本文认为应坚持“流通为用,权责匹配,突出成效”的思路,加强以下 3 方面建设,不失为提高数据治理水平的策略选择。

4.1 构建数据治理的协同模式

基于数据实现开放和规范化的共享方式,建立形成数据源头部门、数据管理部门、数据使用部门三方协同参与的数据治理的机制,是实现数据“流通为用”“权责匹配”的重要保证。明确全生命周期下的数据权属与责任匹配,实现数据提供者负责数据的质量、数据管理者负责数据的共享策略与安全、数据使用者明确数据的使用需求与反馈,三方协同使得数据的治理更加高效、数据的责任更为明

确、数据的质量更有保证,从而实现更大程度数据挖掘和发挥数据效用。

4.2 建立数据要素的治理规范

数据列为与土地、劳动力、资本、技术同样重要的第五生产要素。2019年4月国务院发布的《关于构建更加完善的要素市场化配置体制机制的意见》,提出加快培育数据要素市场,探索建立统一规范的数据管理制度。需要建立围绕数据资产的管理规范制度体系,加强其权属、流通、交易、销毁等全生命周期的规范管理。从技术、制度、法理等不同层面落实数据安全责任、防范数据安全风险,形成权属明确、安全规范的数据制度和管理规范并有效施行,保障数据治理成效的最大发挥。

4.3 深入推进数据治理场景化

场景化治理即围绕数据应用的场景实现精细化数据治理。精细化过程是不断地对数据提出“质疑”和需求确认的过程,最终能够使得数据技术、管理与业务场景需求完全融合、无缝衔接,避免形式化的治理、概念上的治理,最终达到通过数据治理“精准”满足业务需求。

从高校的实践来看,更多呈现数据价值的业务场景需求按用户对象可划分为三类:一是直接面向师生个人的数据应用需求,如个人数据查询、个人画像等;二是面向各个业务部门维度的数据使用需求,如部门数据统计、部门数据共享等;三是面向学校全域的数据统计与决策分析需求,如数据统计报表、校情概览、绩效评价等。这些应用场景需要结合实际数据治理的实施深入实践,通过数据应用的成效实现治理的目的。

5 结论

本文基于DCMM进行高校数据治理水平的量化评估,具有较为明确的标准性、指导性和可操作性。对高校的数据治理进行了实践探索与理论补充,为推进高校实施DCMM的标准,提升高校数据治理水平提供了一定借鉴与参考。不足之处在于:一是数据治理是一项长期的系统工程,因校因地差异较大,用一个具体标准衡量不同高校的水平高低,有“削足适履”之嫌;二是问卷采样的样本数量

还较少,指标权重不能完全反映高校整体的数据治理的现状水平,更科学合理地设定评价指标的权重是下一步需要研究的内容之一。

参考文献

- [1] 中国国家标准化管理委员会.GB/T 36073-2018,数据管理能力成熟度评估模型[S].2018.
- [2] 许晓东,王锦华,卞良,等.高等教育的数据治理研究[J].高等工程教育研究,2015(5):25-30.
- [3] 刘桂锋,钱锦琳,卢章平.国内外数据治理研究进展:内涵、要素、模型与框架[J].图书情报工作,2017,61(21):137-144.
- [4] 余鹏,李艳.大数据视域下高校数据治理方案研究[J].现代教育技术,2018,28(6):60-66.
- [5] 董晓辉,郑小斌,彭义平.高校教育大数据治理的框架设计与实施[J].中国电化教育,2019(8):63-71.
- [6] 董晓辉.活动理论视角下高校教育数据治理体系构成要素研究[J].中国电化教育,2021(3):79-87.
- [7] 孙嘉睿.国内数据治理研究进展:体系、保障与实践[J].图书馆学研究,2018(16):2-8.
- [8] 万方,周西平.基于DCMM的我国警务数据管理能力成熟度评估[J].云南警官学院学报,2021(1):90-96.
- [9] 胡成.基于“AHP+熵权”耦合TOPSIS法的供电企业数据治理成效评价体系构建及其应用研究[J].科技创新导报,2017,14(27):172-173.
- [10] 冉冉,刘颖,胡楠,等.大数据环境下的数据治理框架研究及应用[J].电子世界,2018(24):129-130.
- [11] 雷斌,陆保国.面向大数据的治理框架研究[J].电子质量,2018(6):1-3,7.
- [12] 叶兰.数据管理能力成熟度模型比较研究与启示[J].图书情报工作,2020,64(13):51-57.
- [13] 刘银喜,赵淼,赵子昕.政府数据治理能力影响因素分析[J].电子政务,2019(10):81-88.
- [14] 赵亚伟.大数据环境下高校数据治理策略探索[J].电脑知识与技术,2019,15(1):22-23.

(收稿日期:2022-10-01)

作者简介:

张国宝(1980-),男,博士生,工程师,主要研究方向:校园信息化、数据分析、工作流。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.cn