

# 国内外数据治理模型对比分析\*

肖洁琼 奉国和

(华南师范大学经济与管理学院信息管理系, 广州 510006)

**摘要:** [目的/意义] 数据治理是数据科学时代关注的重要问题之一, 研究数据治理模型是实现数据价值最大化和数据治理最终目标的有效且必要途径。[方法/过程] 采用文献调查法分别对八种数据治理模型在内涵、要素、功能及范围四方面进行对比分析。[结果/结论] 归纳与总结出不同模型的优势与不足, 为数据治理工作者在选择与使用时提供参考。

**关键词:** 数据治理模型 数据治理内涵 数据治理要素 数据治理功能 数据治理范围

**分类号:** G203

**DOI:** 10.31193/SSAP.J.ISSN.2096-6695.2020.02.02

## 0 引言

当前世界正处于工业 4.0 时代, 各国经济发展战略白皮书都在强调利用大数据实现工业转型, 国外有美国的《大数据研发战略》<sup>[1]</sup> 和欧盟《朝向开放科学体系》<sup>[2]</sup>, 我国也颁布了《国务院关于促进大数据发展的行动纲要》<sup>[3]</sup>, 这些政策都旨在通过大数据、云计算和人工智能等技术, 促进数据共享和利用, 实现工业协作和生态创新<sup>[4,5]</sup>。

在此背景下, 从各种来源获得的以 ZETTABYTES 为单位的大型复杂数据集呈指数型增长, 这些数据在政府、医疗、教育、管理和物流等不同领域都有许多应用, 例如, 在美国被用于预防犯罪和反恐, 在韩国被用于支持智能政府运营和战略规划<sup>[6]</sup>, 但使用者在数据处理和管理方面面临巨大的问题, 实施大数据战略的组织和企业都亟需一套新的数据处理政策和模型。因此, 数据治理 (Data Governance, DG) 作为一个新兴话题成为研究人员和行业领导者关注焦点<sup>[7]</sup>。国内学者张宁等<sup>[8]</sup> 对数据治理概念、体系、内容和应用的相关研究进行述评, 指出“框架设计”“价值探讨”是目前研究重点, “框架体系”“成熟度模型”“框架模型设计优化”是未来重点研究领域。

\* 本文系广州市科技计划基础与应用基础研究专题项目“突发词探测理论、方法与应用研究”(项目编号: 202002030384) 成果之一。

[作者简介] 肖洁琼 (ORCID: 0000-0001-5783-2893), 硕士研究生, 研究方向为数据挖掘、信息计量学; 奉国和 (ORCID: 0000-0002-0774-1544, 通讯作者), 教授, 博士, 研究方向为信息计量学、数据挖掘, Email: ghfeng@163.com。

笔者以“数据治理”或“Data Governance”或“Information Governance”为关键词检索相关中外文文献,发现数据治理研究大致始于2004年,最先开始于企业实践,随后在学术界围绕模型构建和案例分析展开讨论,主要集中于计算机技术、企业管理以及图书馆管理等领域<sup>[9]</sup>。关于数据治理模型构建的研究,最早由大型信息咨询公司和标准化组织发起,随后延伸到学术界。业界对于构建数据治理模型的研究成果集中在各组织机构提出的治理框架和模型,大体可以分为三类:数据治理成熟度评估模型<sup>[10,11]</sup>、数据治理过程模型<sup>[12,13]</sup>和数据治理要素模型<sup>[14,15,16]</sup>。相对来说,学术界对数据治理模型的研究显得分散且独立。国内包冬梅<sup>[7]</sup>对国际影响力较大的DAMA框架和DGI框架的主要内容和要素架构进行了研究,并在此基础上提出针对我国高校图书馆数据治理的CALib数据治理框架;刘越男和闫慧等<sup>[17]</sup>就大数据情境下的政府数据治理展开研究并以贵州省抢险救灾实例进行论证;杨琳<sup>[18]</sup>在分析国内外数据治理理论、方法和需求的基础上,提出面向大数据的治理框架并列出了相应的应用场景。国外K.Wende<sup>[19]</sup>提出了描述和说明企业数据治理的弹性模型,认为模型应该由角色、决策域与主要活动以及职责三部分组成;S.Kim<sup>[20]</sup>提出将商业和IT技术结合起来构造数据治理模型。相比于对数据治理模型的研究,分析数据治理成功案例的研究则显得略少,夏义堃<sup>[21]</sup>对比分析部分发达国家政府数据治理政策和治理模型,提出我国政府数据治理框架和治理内容的独特观点;K.Weber<sup>[22]</sup>基于6家不同类型跨国公司有关数据治理的社区行动研究项目,提出了一个由数据质量角色、决策域和责任三部分组成的数据治理责任分担矩阵模型。此外,还有一些学者从数据治理的内涵、要素、模型等角度进行述评,指出治理体系和治理模型的设计是未来研究的重点方向,数据治理实践是终极目标<sup>[8,23,24]</sup>。

结合国内外文献调查可发现,首先,在数据治理领域,学界和业界的关联并不紧密。学界和业界对数据治理的定义并没有随着时间推移而集中发生改变或达成共识,业界各机构对数据治理的定义更多取决于业务需求和商业竞争而不是学术界研究成果;但是,学界的研究却是随着业界的发展产生波动,不少学术研究成果都是在业界出现新理论、新模型和新案例的基础上产生的。由此可见,在数据治理领域,业界对数据治理理论和模型的研究进度快于学术界,且研究成果被学术界认可。其次,构建大量的数据治理模型并不能促成大批成功的数据治理案例产生,究其原因,学界对于数据治理的研究与推崇并不能代表企业和组织愿意接受理论使用模型。组织机构想要实现数据价值最大化,就需要进行数据治理,构建具有高度可操作性的模型<sup>[20]</sup>;然而,现有研究成果并没有阐述组织机构如何进行数据治理模型的选择、实施与维护。因此,笔者从组织机构实际应用出发,选择国内外大型知名标准化组织和咨询公司相关数据治理框架模型作为研究对象,使用文献调研法,选择深受业界和学界认可的4个标准化组织和4个咨询公司提出的8种数据治理模型,并对它们的发布时间、发布机构和模型名称做了简要梳理(见表1)。笔者希望通过分析表1所列8种数据治理模型所体现的治理内涵、构成要素、特点、功能和使用范围,探索不同模型优势与不足,并结合我国开展数据治理的现实环境,着重对组织结构构造数据治理模型提出新的思考,以期为相关研究提供借鉴与参考。

表1 数据治理模型基本信息

机构类型	发布时间	发布机构	模型名称
标准化组织	2004年	国际数据治理研究所 (DGI)	DGI 数据治理框架
	2008年9月	国际数据管理协会 (DAMA)	DAMA 数据治理模型
	2015年5月	信息技术服务标准 (ITSS) (中国)	《数据治理白皮书》模型
	—	高等教育统计局 (HESA) (英国)	HESA 数据治理模型
咨询公司	2007年10月	国际商业机器公司 (IBM) 数据治理委员会 (美国)	IBM 数据治理模型
	2008年12月	高德纳咨询公司 (Gartner) (美国)	Gartner 6 阶段成熟度模型
	2015年3月	Mustimuhw 信息解决方案公司 (Mustimuhw Information Solutions)	MIS 数据治理螺旋模型
	2015年9月	艾比埃公司 (Information Builders) (美国)	IB 数据治理简易模型

## 1 数据治理内涵分析

不同数据治理模型对数据治理定义差异很大,本节拟从整体认识和概念构成两方面进行分析,整理各模型白皮书中数据治理定义(见表2),发现数据治理的必要条件及其对模型构建的影响,为后续治理模型研究提供参考。

表2 数据治理内涵

发布年份	模型名称	数据治理内涵
2004	DGI 数据治理框架 <sup>[13]</sup>	数据治理是根据约定模型处理信息和实施决策的一个系统,包括实施者、实施步骤、实施时间、实施情境以及实施途径与方法。
2007	IBM 数据治理模型 <sup>[10]</sup>	数据治理是一种质量控制规程,用于在管理、使用、改进和保护组织信息的过程中添加新的严谨性和纪律性;有效的数据治理可以通过促进跨组织协作和结构化策略制定来提高公司数据的质量、可用性和完整性;它平衡了派系孤岛与组织利益,直接影响了任何组织最关心的四个因素:增加收入、降低成本、降低风险和增加数据信心。
2008	DAMA 数据治理模型 <sup>[14]</sup>	数据治理是数据管理的核心功能,指对数据资产管理行使权力和控制的活动集合(包括计划、监督和执行);数据治理功能指导所有其他数据管理功能的执行。数据治理在更高的层次上执行数据管理。
	Gartner 6 阶段成熟度模型 <sup>[11]</sup>	数据治理是 IT 治理的一部分,同时也属于公司治理的范畴,企业应该建立一个数据治理的规范流程。
2015	MIS 数据治理螺旋模型 <sup>[15]</sup>	数据治理是信息相关流程的决策权和责任制,根据商定的模型执行,这些模型描述了谁可以采取什么行动来处理什么信息,何时,在什么情况下使用什么方法。
	IB 数据治理简易模型 <sup>[12]</sup>	数据治理是新兴学科的总称,包括数据质量、数据管理、业务流程管理和风险管理等许多不同实践;其目标是确保数据以可持续的方式服务于商业目的。数据治理也可定义为人员、流程和技术的正式协调,使组织能够将数据用作企业资产。
	《数据治理白皮书》模型 <sup>[24]</sup>	数据治理是对企业的数管理 and 利用进行评估、指导和监督,通过提供不断创新的数据服务,为企业创造价值的活动。
—	HESA 数据治理模型 <sup>[16]</sup>	数据治理平衡了派系孤岛与组织利益,实现了公平访问所有人的数据。

## 1.1 大体认识

表 1 看出, 整体上, 业界对数据治理内涵的认识受发展时间的影响较轻, 内涵大体上可以分为四类: 一是数据治理是数据管理的一部分; 二是数据治理包含数据管理; 三是数据治理与数据管理相关但没有从属关系; 四是数据治理与 IT 治理相关。

第一种观点“数据治理属于数据管理”是由国际数据管理协会 (DAMA) 提出的, DAMA 是全球首个数据管理专业组织, 该内涵被广泛应用。一方面, 以 DAMA 数据治理理论为基础的数据治理模型都将数据治理归于数据管理之中, 认为数据治理是数据管理框架的核心职能, 是在高层次上执行数据管理制度; 然而, 这一内涵被认为具有一定的局限性。另一方面, 它对于数据治理活动的内涵则得到了广泛认同, 学界认同数据治理是对数据资产管理行使规划、监控和执行活动。

第二种观点“数据治理包含数据管理”比较少见, 主要体现在 Information Builders 数据治理简易模型中。Information Builders 是美国一家软件与咨询公司, 致力于帮助组织构建信息系统, 以形成巨大的竞争优势<sup>[25]</sup>。其数据治理模型就是在数据管理和数据治理的理论基础上提出的, 该模型认为人员、流程和技术的正式协调能够使组织将数据用作企业资产。不过, 这种理论与模型在业界的应用率较低。

第三种观点为“数据治理与数据管理相关但没有从属关系”, 该观点目前接受程度最高, 应用程度最为广泛, IBM 公司数据治理白皮书、数据治理协会 (DGI) 数据治理框架、ITSS 的《数据治理白皮书》和 HESA 数据治理模型等, 都以该观点作为基础。该观点认为, 数据治理是针对信息相关过程的决策和规范。

第四种观点为“数据治理与 IT 治理相关”, 这种观点常见于组织和企业进行数据治理实践。IT 技术和数据处理是不可分离的, 因此, 这一观点也具有客观存在性, Gartner 公司的成熟度模型认为 IT 治理是数据治理的一部分, 国际标准化组织 / 国际电工委员会的第一联合技术委员会 / 国际 IT 治理与服务管理标准组织 (ISO/IEC JTC1/SC40) 也指出 IT 治理模型可以运用于数据治理领域<sup>[26]</sup>。

对数据治理整体认识和理解是构建治理模型的基础。使用第一种观点的机构, 它们的数据治理模型从属于数据管理模型, 在数据管理模型中, 数据治理模型只是其中的一部分, 地位与其他数据管理功能模型相比略高或者相同, 但为了配合其他数据管理功能、实现数据管理目标, 构建的数据治理模型大多数情况下只是一个治理框架或治理要点, 而没有具体的治理内容或生命周期。使用第二种观点的机构, 它们的数据治理模型包含了数据管理模型, 构建的数据治理模型在每个方面都能体现数据管理的影响, 这类模型实质是通过有效的数据资产管理、数据质量管理、数据业务管理等内容来进行数据治理。使用第三种观点的机构, 它们的数据治理模型体系没有数据管理等字眼, 且构建要素主要由治理步骤或治理要素组成。使用第四种观点的机构构建的数据治理模型常常与 IT 治理模型类似甚至相同。

## 1.2 概念构成要素

通过对比分析发现, 在数据治理概念构成上, 笔者认为 8 种模型就数据治理定义而言大致都包含数据治理目标、数据治理职能、数据治理核心和数据治理规范四个方面内容。



在数据治理目标上,可以分为四类:第一种为提高数据质量、实现数据可持续发展<sup>[10]</sup>,IBM数据治理模型认为有效的数据治理可以提高组织机构数据的质量、可用性和完整性。第二种为指导数据管理,《DAMA数据管理知识体系指南》阐述数据治理是数据管理的核心功能,由此可认为其目标是更好进行数据管理<sup>[14]</sup>。第三种为实现数据公平访问,HESA数据治理模型重点关注角色分配以最终实现公平访问所有人的数据<sup>[16]</sup>。第四种为实现数据价值最大化,IB数据治理模型和《数据治理白皮书》都表明数据治理目标是确保数据以可持续的方式服务于商业目的<sup>[12,24]</sup>。

在数据治理职能上,8种数据治理定义比较相似,即通过各种治理活动对数据进行评估、指导和监督。虽然不同概念内涵侧重的关键活动不一致,但如《DAMA数据管理知识体系指南》、DCI数据治理框架和HESA数据治理模型等着重于数据治理的指导职能。关于数据治理的评估职能,IBM公司数据治理白皮书和《数据治理白皮书》都有所涉及,但与《数据治理白皮书》相比,IBM数据治理对评估职能的关注度更高。至于监督职能,都有强调,但如何监督和谁监督等问题都未涉及。

在数据治理核心上,可以分为四类:一为控制数据质量,如IBM数据治理模型直接指出,数据治理是一种质量控制规程,数据治理的核心在于数据质量的控制<sup>[10]</sup>。二为创新数据服务,《数据治理白皮书》认为数据治理是一个提供不断创新的数据服务,为企业创造价值的活动,其治理核心在于创新数据服务<sup>[24]</sup>。三为平衡组织利益,HESA数据治理模型着重解决角色分配就是为了平衡派系孤岛与组织利益<sup>[16]</sup>。四为对数据资产的决策和权利控制,数据治理委员会认为数据治理的核心是对数据相关事宜的决策制定与权利控制<sup>[14]</sup>。

在数据治理规范上,8种数据治理定义都显示要遵循一定的规范和流程,增强治理过程的严谨性和纪律性。上述分析可知,数据治理模型的构造必须从认知入手,并且建构数据治理模型也是为了更好遵循数据治理规范、解决数据治理核心、体现数据治理职能、最终实现数据治理目标。因此,本文接下来的分析也将从这四个方面开展。

## 2 数据治理模型构成要素和治理特点分析

总结归纳8种数据治理模型,得到各模型构成要素和治理特点(见表3)。对比分析后发现这8种模型的构成要素可以从数量和内容两个角度进行划分,同时通过分析这两大类模型构成要素的内涵及特点,探讨数据治理模式是否体现治理内涵中必须具备的治理核心和治理规范。

表3 数据治理模型的构成要素

模型类型	模型名称	构成要素	模型特点
单要素模型	IBM数据治理模型 <sup>[10]</sup>	由11种成熟度模型构成,分别为:1)支撑域:数据架构、分类与元数据、审计与日志与报告;2)核心域:数据质量管理、信息生命周期管理、信息安全与隐私;3)促成因素:组织结构与认知度、政策、数据相关责任人;4)成果:数据风险与合规、价值创造。	每种成熟度模型都有5个成熟度等级;整个模型即是成熟度框架,又是一组评估数据治理成熟度的问题/工具。

续表

模型类型	模型名称	构成要素	模型特点
单要素模型	Gartner 6 阶段成熟度模型 <sup>[11]</sup>	1) 规范; 策略管理、过程建模与设计、学习管理; 2) 规划: 需求分析、影响分析、存储库与元数据管理; 3) 构建: 模式设计、数据架构、数据质量规范; 4) 运行: 数据质量监控与实施、数据访问审计与报告、信息生命周期管理。	每个阶段都需要采取特定措施将组织提升到下一个成熟度; 并强调管理信息作为数据资产应得到高层重视。
	MIS 数据治理螺旋模型 <sup>[15]</sup>	由 6 个治理核心要素组成: 1) 原则与愿景; 2) 治理结构; 3) 责任机制; 4) 治理政策; 5) 隐私与安全政策; 6) 法律。	6 个治理要素中的任何一个进步都会使治理效果倍增。
	IB 数据治理简易模型 <sup>[12]</sup>	治理过程分为 5 个阶段: 1) 决议; 2) 分配; 3) 解决/审查; 4) 批准; 5) 关闭。	类似于数据治理生命周期模型, 但每个阶段都可重复或回退。
	HESA 数据治理模型 <sup>[16]</sup>	由 4 部分组成, 分别是: 1) 治理参与人员; 2) 治理授权与报告; 3) 操作管理; 4) 合规性。	关注治理过程中的角色构建, 治理关键在于分配角色。
多要素模型	DGI 数据治理框架 <sup>[13]</sup>	1) 人员与组织结构: 数据利益相关人、数据治理委员会、数据管理员; 2) 规则与协同工作规范: 使命、关注域、数据规则与定义、决策权、职责分工、控制; 3) 过程: 数据治理过程。	以一种直观地方式展示了组件间的逻辑关系, 将治理流程融入模型之中, 形成了一个从方法到时间的自成一体的完整系。
	DAMA 数据治理模型 <sup>[14]</sup>	1) 职能框架: 战略、组织和角色、政策和标准、项目和服务、问题、估值; 2) 环境要素子框架: 组织与文化、活动、主要交付物、角色与职责、实践与方法、技术、目标与原则。	定义了一个二维矩阵, 竖轴为职能组成, 横轴为环境要素。
	《数据治理白皮书》模型 <sup>[24]</sup>	1) 原则: 战略一致、风险可控、运营合规、绩效提升; 2) 范围: 大数据生命周期、大数据安全与隐私、大数据架构、数据质量、大数据服务创新; 3) 实施与评估: 促成因素、实施过程、成熟度评估、审计。	3 个治理核心要素共同构成三维模型; 同时, 每个治理要素细分为若干治理要点。

注: 数据治理模型按模型类型和发布时间升序排列。

## 2.1 模型构成要素数量

从构成要素的数量角度看, 可以划分为单要素构成和多要素构成两大类。单要素数据治理模型是指治理模型的主框架是由同类型要素构成的, 其构成模型主框架的要素主要有成熟度模型、参与要素模型和过程模型三种。例如, IBM 数据治理成熟度模型是典型的成熟度模型, HESA 数据治理模型是典型的参与要素模型, 而 IB 数据治理简易模型则属于过程模型。多要素数据治理模型是指由两类及以上要素构成模型主框架、多角度同步进行数据治理的模型, 主要的构成要素有职能、环境、人员、过程和规范等。例如, DAMA 数据治理模型结合了数据治理职能和环境要素, 《数据治理白皮书》模型则是从原则、范围和实施三个维度出发, 构造了一个三维度模型。

## 2.2 模型构成要素内容

从构成要素的内容角度看, 可以划分为过程要素、评估要素、角色要素、原则要素、职能要素及范围要素六种。在模型构成要素中, 范围要素使用最为频繁, 它常作为模型的单一要素或主要要素出现; 其次是过程要素、评估要素和角色要素三个要素均可作为模型的单一要素出现, 也常常与其他要素结合出现; 使用较少的是原则要素和职能要素, 原则要素是指数据治理过程中所遵循的、首要的、基本的指导性法则; 职能要素是指数据治理职能。由于这两个要素一个处于数据治理的初始阶段, 一个可以作为数据治理的结果阶段, 因此, 这两种要素往往不会作为唯一要素

独出现在治理模型中。

范围要素模型是指将数据治理过程中的关键实施领域模块化、规范化的治理模型,或者说,治理模型侧重于解决治理什么的问题。当范围要素作为单一要素出现在模型中,则它主要用于解决组织和企业关于治理域的问题。MIS 数据治理螺旋模型是一个典型的范围要素模型,整个模型仅包含原则与愿景、治理结构、责任机制、治理政策、隐私与安全政策和法律六个治理核心要素,共同构成一个螺旋上升式的治理场景。当范围要素作为多要素中的主要素出现时,对其他要素具有指引作用,例如,DAMA 数据治理模型是属于范围—职能模型,该模型中,范围要素作为主导要素,必须先满足范围要素,才能实现其职能。在大多数情况下,当范围要素出现在多要素模型中,会被认为是模型的主要素,而其它要素为一般要素。

过程要素模型是指将数据治理过程模块化,使治理者在进行数据治理时能够按部就班、有条不紊,这类模型主要解决治理流程的问题。IB 数据治理简易模型就是一个过程要素模型,它采用6个可重复的渐进式措施(包括:决议、分配、解决/审查、批准和关闭)将组织的数据治理水平提升到下一个阶段。Gartner 6 阶段成熟度模型属于过程要素模型,根据数据生命周期将数据治理分为6个阶段,每个阶段都需要采取特定措施将组织提升到下一个成熟度。

评估要素模型是指评估数据治理成熟度的治理模型。当评估要素作为单一要素时,此类模型是一组评估数据治理成熟度的问题或工具。IBM 公司的数据治理成熟度模型就是一个典型的评估要素模型,它由数个关键域组成,每个关键域都有若干个与之相匹配的成熟度模型,每个成熟度模型都有成熟度等级,从而对数据治理的评估贯穿于整个模型中。当评估要素作为一般要素出现在模型中,通常作为辅助要素存在,例如,《数据治理白皮书》模型由多个要素组成,其中包含评估要素。

角色要素模型是指将参与数据治理的组织和人员在治理活动中扮演的角色及其权责关系的划分作为重点关注的治理模型,当角色要素作为模型单一要素时,权责分配处于组织机构数据治理核心。HESA 数据治理模型认为数据治理的主要因素有4个环节,即治理参与人员、治理授权与报告、操作管理和合规性。因此,该模型将这4个环节贯穿于整个模型,致力于让治理参与者都能渗透到其中的一、二个环节,其治理特点在于分配角色而非分配工作。当角色要素处于一般要素地位时,大多数情况下会呈现出与其他要素势均力敌的状态。

若要为模型构成的6个要素进行排序,范围要素排在第一位,它是主要素,其他若干要素都是辅助范围要素;其次是过程要素,它为数据治理活动提供了思路和具体步骤;排第三位的是原则要素,它是数据治理过程中的准则;处于第四位的则是评估要素,它能为处于进行中或完成的数据治理工作给予一定的指导意见;角色要素处于第五位,它可以对数据治理活动所需的人员进行合理的分工,使治理活动显得更科学合理;处于末位的是职能要素,它可以充当数据治理结果的模板,评估治理工作是否全面。

上述对比分析表明,从是否体现数据治理核心和治理规范的角度出发,多要素数据治理模型比单要素数据治理模型更能体现和实现对数据资产管理的决策、协调和分工以及治理规范,特别是拥有职能或过程要素以及规范或原则要素;单要素模型中,以参与要素为主框架的数据治理模型因拥有更明确的角色分工而突出数据治理核心,以治理过程为主框架的数据治理模型更为详细

地描述治理方法和步骤, 因而更能突出数据治理规范。因此, 笔者认为, 组织在选择数据治理模型时, 可以优先考虑多要素数据治理模型, 同时, 根据自身实际情况, 模型的主框架组成要素可以选择组织相对薄弱的方面, 例如, 对于数据治理的过程和规范不甚明确的组织可以选择带有过程要素和规范要素的治理模型。

### 3 数据治理模型功能及适用范围分析

各模型的功能和适用范围见表 4。分析数据治理模型的功能, 既能体现治理目标和治理职能, 还能发现不同模型的适用范围, 为组织从需求的角度出发选择数据治理模型给予相应的帮助。

表 4 数据治理模型的功能与适用范围

模型名称	功能	适用范围
DGI 数据治理框架 <sup>[13]</sup>	将数据利益相关者的多样性和清晰的思想结合在一起, 制定实用的、可操作的数据治理过程并进行管理。	组织文化与环境独特。
IBM 数据治理模型 <sup>[10]</sup>	作为一个具体和客观的基准, 评估组织当前数据治理状态和有效性, 定义新机会、改善治理。	组织不清楚目前在治理方面的位置及其他他们为实现目标所需采取的步骤。
DAMA 数据治理模型 <sup>[17]</sup>	制定数据管理计划; 进行数据管理控制。	将数据治理纳入数据管理范畴的组织。
Gartner 6 阶段成熟度模型 <sup>[11]</sup>	提供在企业范围内管理信息的流程与方法。	处于企业信息管理的早期阶段。
MIS 数据治理螺旋模型 <sup>[15]</sup>	提供具有指定工具和流程的综合框架, 建立并实现组织的直接和积极参与, 最终拥有和控制相关数据。	组织数据获取渠道广来源多。
IB 数据治理简易模型 <sup>[12]</sup>	实现业务价值最大化并建立数据治理可持续发展计划的实用方式。	组织业务大多围绕各类数据展开。
《数据治理白皮书》模型 <sup>[24]</sup>	明确了治理活动中每个过程域的建设目标和度量标准。	企业组织和开展大数据治理实践。
HESA 数据治理模型 <sup>[16]</sup>	设置通用的治理活动的范围和职责范围。	组织需要管理的数据量不大且数据治理处于起步阶段。

注: 数据治理模型按发布时间升序排列。

从数据治理模型功能出发, 上述 8 种模型的功能大致可以分为以下 4 类:

一是评估组织当前数据治理状态, 即通过在组织内构建相应的治理模型, 评估目前的治理位置及为实现目标所需采取的步骤, 这是该类数据治理模型区别于其他模型的最大特点。8 个模型中, IBM 数据治理模型和 Gartner 6 阶段成熟度模型都具有评估功能, 然而, 后者适用于在企业信息管理的早期阶段, 而前者的适用范围则更加广泛。

二是规定治理内容, 即通过模型各组成要素, 为组织的数据治理设计和部署提供一定的范围和原则。大多数数据治理模型都具有这项功能, 例如 DAMA 数据治理模型认为其 8 个职能要素就是组织数据治理过程中应该实施的内容, 《数据治理白皮书》模型则是通过范围维度和原则维度具体阐述数据治理 7 大关键域和 4 大原则, HESA 数据治理模型则通过角色分配设置通用治



理活动范围和职责范围。

三是建立治理过程，即以流程图的方式明确地表述数据治理的步骤，为组织设计数据治理流程提供有力的支持。通常，模型构成要素中含有过程要素的数据治理模型都具备该功能，DGI 数据治理框架以访问路径的形式展示了治理目标、治理角色、治理内容、治理方法和治理流程五个问题，形成了一个从方法到实施自成一体的完整系统；Information Builders 数据治理简易模型提出了7步可重复可循环的治理过程，为以数据为支撑业务的组织实现业务价值最大化建立了较强实用性的过程；Gartner 6 阶段成熟度模型则更适合企业，它提供了在企业范围内管理信息的流程与方法。

四是控制与监督治理实施情况。数据治理必须遵循一定的过程和规范，治理情况也应该及时向项目负责人和组织进行反馈，该功能必不可少。表3中8个治理模型都涉及此功能，但是有些模型将该功能直接表现在治理流程中，而更多的则是隐藏在其他要素中。例如，IB 数据治理简易模型中直接制定流程改进反馈机制这一步骤；而 DGI 数据治理框架则是在人员和组织结构进行数据治理过程中通过决策和职责分工等活动间接对数据治理进行控制和监督。

通过上述分析，笔者认为，在数据治理目标上，8种模型的具体治理目标和适用范围有所不同，但都能够体现数据治理最终目标即实现数据资产价值最大化；在数据治理职能上，每种模型都具有2到3种不同的功能，虽然具体治理内容和过程存在一定的差异，但都能对数据资产进行规划、控制和执行的，只是规划、控制和执行程度不一致。因此，若组织从治理目标和治理职能的角度出发选择数据治理模型，则它们之间不具有明显的差异。笔者建议组织从自身需求出发，根据数据治理模型的适用范围进行选择更为合理。例如，如组织不清楚目前在治理过程中的位置，可以选择具有评估功能的治理模型；如不确定已有的治理内容覆盖是否全面，则可以选择侧重于规定治理内容的模型。

## 4 结论与启示

数据治理模型是实现数据治理的一个重要组成部分，其作用非常明显。通过上述对比分析发现，即使组织机构在数据治理领域进行了深入的研究，构造了广泛的模型，其研究成果不仅没有得到学界公认，彼此之间也有矛盾之处。每个数据治理模型都有其他模型不具备的优势，也有不可避免的不足（见表5）。

表5 数据治理模型的优势与不足

模型名称	优势	不足
DGI 数据治理框架 <sup>[13]</sup>	执行流程易于理解便于操作。	环境适应性较弱，不利于组织变更。
DAMA 数据治理模型 <sup>[14]</sup>	职能分工明确，环境要素清晰，具有较好的模仿性。	对环境的适应性较弱，需要其他条件配合。
《数据治理白皮书》模型 <sup>[24]</sup>	明确治理活动中每一个过程域的建设和度量。	模型没有辅助的实施流程与步骤。
HESA 数据治理模型 <sup>[16]</sup>	对组织规模和数据的要求较低，执行难度小。	治理角色定义复杂；不关注治理步骤。

续表

模型名称	优势	不足
IBM 数据治理模型 <sup>[10]</sup>	拥有一个具体和客观的基准, 有利于评估组织当前数据治理状态和有效性。	没有提及评估流程与整体治理过程。
Gartner 6 阶段成熟度模型 <sup>[11]</sup>	制定每个阶段的行动方案, 并强调管理信息作为数据资产应得到高层重视。	在数据治理的中后期适用性弱。
MIS 数据治理螺旋模型 <sup>[15]</sup>	对数据格式要求低; 模型可变性强。	治理要素不明确, 需要配合其他模型使用。
IB 数据治理简易模型 <sup>[12]</sup>	重点关注数据; 适用于宏观的治理方案。	模型过于简单, 需要制定详细步骤。

注: 数据治理模型按发布机构和发布时间升序排列。

综合观察表 1 至表 5 可以发现, 由标准化组织构建的数据治理模型大多都是多要素模型, 这类模型通常包含着数据治理范围要素, 能够为需求者提供清晰的思路和方案, 对缺乏治理经验的应用者友好, 可模仿性强, 在现实中容易理解和实操, 但同时, 由于标准化组织构建的模型更注重普适性, 因此这类模型通常缺乏针对性, 并且对环境的适应性较差, 需要实施者按照模型内容依次部署。由咨询公司构建的数据治理模型大多是单要素模型, 这类模型大多是范围模型、过程模型或评估模型, 针对性较强, 但同时, 这类模型发布机构受其商业性和盈利性的影响, 治理模型可变性强而可模仿性弱, 在实践中往往需要专业人员的指导。因此, 组织对于数据治理模型的选择不能仅看模型自身的优缺点, 而应结合组织自身的管理目标、文化背景和运作模式等, 让数据治理模型能够为整个组织的共享而服务<sup>[23]</sup>。笔者整合现有成果, 将模型中必须存在的要素进行分析和提取, 为组织选择治理模型或者构造治理模型提供一定的帮助。

首先, 组织在选择数据治理模型时, 要明确数据治理的目标和原则。明确数据治理目标能够帮助组织更好确定适用的治理功能和治理范围, 进而目的性更强, 准确度也更高。数据治理原则是数据治理活动中所应遵循的标准或法则<sup>[23]</sup>, 组织在选择数据治理模型时, 也要关注模型中是否有原则要素的存在, 如模型中直接列出了原则, 则整个模型运转的边界将更加清晰。其次, 模型的意义在于帮助组织明晰数据治理的流程, 实现治理清晰度; 创建明确的任务, 确保努力的价值<sup>[23]</sup>。因此, 选择带有过程要素的多要素数据治理模型对组织实现数据资产管理决策、协调和分工更有帮助。第三, 通过上述分析可知, 模型构建要素的本质都是人、活动和规范。在选择和构建治理模型时, 组织还应该根据自身的需求和短板, 选择更具有针对性的模型。

现阶段, 我国还没有健全的法律制度来规范和保障数据治理行为, 建议各组织机构在选择数据治理模型时选择原则要素模型, 如果模型中不包含原则要素, 更应该将治理原则、方法和流程等内容作为官方行动准则进行公示, 同时要注意准则的专业性和完整性, 并经常修订更新。为了有效实施数据治理活动, 各组织机构应该采取自上而下全员参与的行动方案, 建立职责清晰、运转高效的数据治理组织体系, 形成组织高层宏观领导、各业务部门协同合作的工作模式。为了更好地进行数据治理工作, 维护数据治理成果, 组织内的各业务部门应该定期开展交流分享经验, 根据组织实际情况和业务人员的反馈, 及时修改和完善数据治理内容体系和行动规范, 若对某些问题有重大意见分歧, 及时聘请业界专家进行针对性咨询, 同时, 高度重视数据安全和风险防

范,这是维护治理成果的必要手段。

本文分析数据治理常见的8种模型发现,尽管有些模型在必备要素之外还拥有其它相似要素,表面上类似,但实质有差别,每个模型都有其使用范围和优缺点。因此,组织在选择数据治理模型时应该充分考虑自身的条件和需求,确定治理模型后,科学地实施模型并根据实际情况对模型进行调整甚至更改。总之,数据整理治理模型设计合理与否,必将影响到组织基于数据进行价值增值活动的各个环节。

### 【参考文献】

- [1] NITRD. The federal big data research and development strategic plan [EB/OL]. (2016-05-20) [2018-09-05]. <https://www.nitrd.gov/PUBS/bigdatardstrategicplan.pdf>.
- [2] COUNCIL OF THE EUROPEAN UNION. The transition towards an open science system—council conclusions [EB/OL]. (2016-03-15) [2018-09-05]. <http://data.consilium.europa.eu/doc/document/ST-7082-2016-INIT/en/pdf>.
- [3] 国务院. 关于促进大数据发展的行动纲要 [EB/OL]. (2015-09-15) [2018-09-05]. [http://www.gov.cn/jzhengce/content/2015-09/05/content\\_10137.htm](http://www.gov.cn/jzhengce/content/2015-09/05/content_10137.htm).
- [4] 顾立平. 数据治理——图书馆事业的发展机遇 [J]. 中国图书馆学报, 2016, 42(5): 40-56.
- [5] YEBENES J, ZORRILLA M. Towards a data governance framework for third generation platforms [J]. Procedia Computer Science, 2019, 151: 614-621.
- [6] AL-BADIA A, TARHINIA A, KHANA A I. Exploring big data governance frameworks [J]. Procedia Computer Science, 2018, 141: 271-277.
- [7] 包冬梅, 范颖捷, 李鸣. 高校图书馆数据治理及其框架 [J]. 图书情报工作, 2015, 59(18): 134-141.
- [8] 张宁, 袁勤俭. 数据治理研究述评 [J]. 情报杂志, 2017, 36(05): 129-134, 163.
- [9] WATSON H J, FULLER C, ARIYACHANDRA T. Data warehouse governance: Best practices at blue cross and blue shield of north carolina [J]. Decision Support Systems, 2005, 38(3): 435-450.
- [10] IBM CORPORATION. The IBM data governance council maturity model: Building a roadmap for effective data governance [R]. 2007.
- [11] GARTNER. Gartner introduces the EIM maturity model [R/OL]. (2008-12-05) [2018-10-29]. <https://wenku.baidu.com/view/cc9ec16af111f18583d05a80.html>.
- [12] INFORMATION BUILDERS. Seven steps to effective data governance [R/OL]. (2015-09-23) [2018-10-29]. [https://www.whitepapers.em360tech.com/wp-content/files\\_mf/white\\_paper/wp\\_iway\\_7steps.pdf](https://www.whitepapers.em360tech.com/wp-content/files_mf/white_paper/wp_iway_7steps.pdf).
- [13] THE DATA GOVERNANCE INSTITUTE. Data governance: The basic information [EB/OL]. [2018-10-29]. [http://www.datagovernance.com/adg\\_data\\_governance\\_basics/](http://www.datagovernance.com/adg_data_governance_basics/).
- [14] DAMA International. The DAMA guide to the data management body of knowledge [M]. New York: Technics Publications, 2009.
- [15] MUSTIMUHW INFORMATION SOLUTIONS INC. Data governance framework framework and associated tools [R/OL]. (2015-05-31) [2018-10-29]. <https://static1.squarespace.com/static/558c624de4b0574c94d62a61/t/558c75a5e4b0391692159c81/1435268517023/BCFNDGI-Data-Governance-Framework.pdf>.
- [16] HESA. Data governance [EB/OL]. [2018-10-29]. <https://www.hesa.ac.uk/support/tools/data-capability/signposting/governance>.
- [17] 刘越男, 闫慧, 杨建梁, 等. 大数据情境下政府治理研究进展与理论框架构建 [J]. 图书与情报, 2017(1): 87-93.

- [ 18 ] 杨琳, 高洪美, 宋俊典, 等. 大数据环境下的数据治理框架研究及应用 [J]. 计算机应用与软件, 2017, 34(4): 65-69.
- [ 19 ] WENDE K. A model for data governance –organizing accountabilities for data quality management [EB/OL]. [ 2018-10-27 ]. <http://aisel.aisnet.org/acis2007/80>.
- [ 20 ] KIM S. The analysis of data governance model for business and IT alignment [J]. Journal of the Korea Society of Computer and Information, 2013, 18(7): 69-78.
- [ 21 ] 夏义堃. 政府数据治理的国际经验与启示 [J]. 信息资源管理学报, 2018, 8(03): 64-72, 101.
- [ 22 ] WEBER K, OTTO B. One size does not fit all—A contingency approach to data governance [J]. Journal of Data and Information Quality(JDIQ), 2009, 1(1): 1-27.
- [ 23 ] 刘桂锋, 钱锦琳, 卢章平. 国外数据治理模型比较 [J]. 图书馆论坛, 2018, 38(11): 18-26.
- [ 24 ] 张绍华, 潘蓉, 宗宇伟. 大数据治理与服务 [M]. 上海: 上海科学技术出版社, 2016.
- [ 25 ] Information Builders [EB/OL]. [ 2017-10-20 ]. [http://www.informationbuilders.com/about\\_us](http://www.informationbuilders.com/about_us).
- [ 26 ] ISO. ISO38505-1: Information technology governance of it—Governance of data [R/OL]. [ 2018-07-11 ]. <http://www.doc88.com/p-7972827513292.html>.

## A Comparative Analysis of Data Governance Models at Home and Abroad

XIAO Jieqiong FENG Guohe

(The Department of Information Management, School of Economics & Management,  
South China Normal University, Guangzhou 510006, China)

---

**Abstract:** [ **Purpose/significance** ] Data governance is one of the most important research topics under the data science. It is a necessary and effective way for research on data governance models so as to maximize the value of data and the ultimate goal of data governance. [ **Method/process** ] Using literature research, this paper compares eight typical models of data governance from four aspects of content, element, function and scope. [ **Result/conclusion** ] This paper summarizes advantages and disadvantages of different models in order to provide references for the selection and using of data governance models.

**Keywords:** Data governance model; Data governance content; Data governance elements; Data governance functions; Data governance scope

---

( 本文责编: 周 霞 )