

数据管理能力成熟度模型

李冰¹, 宾军志²

1. 中国电子技术标准化研究院, 北京 100007

2. 御数坊(北京)科技咨询有限公司, 北京 100007

摘要

为促进大数据产业持续深入发展,提高政府、企事业单位大数据资产管理意识,借鉴国内外成熟度相关理论思想,结合数据生命周期管理各个阶段的特征,对数据管理能力进行了分析、总结,提炼出组织数据管理的八大过程域,并对每项能力进行了二级过程域和发展等级的划分以及相关功能介绍和评定标准的制定。

关键词

大数据;数据管理;能力成熟度

中图分类号:TP399

文献标识码:A

doi: 10.11959/j.issn.2096-0271.2017039

Data management capability maturity model

LI Bing¹, BIN Junzhi²

1. China Electronics Standardization Institute, Beijing 100007, China

2. Data Governance Workshop, Beijing 100007, China

Abstract

To promote the continuous development of big data and improve the government and enterprises' awareness of data management, the organizational data management of the eight process areas were extracted by analyzing and summarizing the data management capabilities, combining with data lifecycle management at all stages of the characteristics and the theory from domestic and abroad. Otherwise, each capability was divided into two process areas and development levels, introduced the functions and formulated the standardization of assessment.

Key words

big data, data management, capability maturity

1 引言

在信息化的时代,特别是如今与大数据相关的研究和应用层出不穷,数据已经成为各个单位最重要的资产,国务院也于2015年8月正式印发了《促进大数据发展行动纲要》,在纲要中明确指出了大数据已经成为推动经济转型发展的新动力,大数据持续激发商业模式创新,不断催生新业态,已成为政府、企事业单位促进业务创新增值、提升核心价值的重要驱动力。但是随着大数据行业的蓬勃发展,国内的相关部门正面临越来越多的挑战。

首先,由于大数据是相对较新的行业,目前大数据相关理论的发展相对滞后,特别是数据管理理论,目前国内各单位更多的是采用国际咨询公司的理论框架或者国际数据管理协会的数据管理知识体系作为引导,但是这些理论基本没有考虑国内数据行业发展的现状和特性,同时,普及程度也有待提高。这导致目前国内很多行业在数据管理方面的意识薄弱、管理方式各异、发展相对落后的局面。

其次,由于目前在数据管理能力成熟度模型的研究中普遍缺少一个统一、系统、适应现代信息环境的数据管理和质量保证体系的专业标准(如类似制造业的ISO 9000等),所以国内外的学者在借鉴软件能力成熟度模型(capability maturity model for software, CMM)的基础上,在不同研究领域尝试提出各种数据能力成熟度模型,用以研究、指导具体的数据生产过程的数据管理,在国际上比较有名的数据能力成熟度模型有美国IBM公司的数据治理能力成熟度模型等,该模型在充分借鉴CMM的基础上,针对数据管理的不同领域进行详细的定义,每个领域都按照CMM

的模式进行阶段划分。

随着数据应用的逐渐增多,国内对于能力成熟度模型的研究也在逐渐增多。目前,针对数据能力的评价依然没有一个完整、全面的模型,现有的模型有的本身就存在特定的倾向性,有的是针对数据管理的特定领域。由于信息化的快速发展,数据的重要性已经体现得越来越明显,特别是大数据、物联网时代,数据已经成为国家的战略资源。针对这样的战略资源,我国迫切需要建立一种通用的能力评价模型,帮助各个企业、单位更好地进行数据资源的评估和规划,进而使我国的信息化在国际上占据更有利的位置,在国际信息化的标准化领域有更大的话语权。

2 数据管理能力成熟度研究现状

2.1 国外研究现状

由于数据的重要性越来越高,数据管理的重要性也在逐渐提升,所以国际上一些组织在借鉴软件能力成熟度模型的基础上也提出各自的数据能力成熟度模型,用以规范、指导具体的数据生产过程的数据管理。目前在国际上关于数据能力成熟度评估模型方面的研究比较著名的有以下两个。

(1) 软件工程研究所(Software Engineering Institute, SEI)的数据能力成熟度模型

数据能力成熟度(data management maturity, DMM)模型是由卡耐基梅隆大学旗下机构研究所以能力成熟度模型整合的各项基础原则为基础开发的,并于2014年8月正式发布。软件能力成熟度集成模型(CMMI)是一项拥有20多年历史、经过实践检验的绩效改善以及软件和系统开发的

黄金卓越标准。DMM模型是一个能实现业务部门利益与IT相互匹配的强大加速器（如图1所示），可为公司组织提供一套最佳实践标准，制定让数据管理战略与单个商业目标相一致的路线图。从而确保能强化、良好地管理并更好地运用关键数据资产来实现商业目标。

DMM包含以下六大职能域^[1]：

- 数据管理战略；
- 数据质量管理；
- 数据操作；
- 数据平台和架构；
- 数据治理；
- 支撑流程。

由于CMMI在软件过程成熟度（SW-CMM）评估过程中取得了巨大的成功，DMM模型一经发布就引起了各方的关注，当前已经在国际上培训了一批评估师，包括中国、巴西、美国等，并且在房地美（美国联邦住宅贷款抵押公司）、微软等公司进行了模型验证。

（2）企业数据管理协会的数据管理能力成熟度模型

企业数据管理协会（EDM Council）是北美地区的一个主要面向金融保险行业数据管理的公益性组织，在数据内容标准制定、数据管理最佳实践等方面有丰富的经验，是业界的倡导者和领导者。组织内部的成员大部分都是数据管理行业和金融保险行业的企业。

数据管理能力评价模型（data management capability assessment model, DCAM）是由企业数据管理协会主导，组织金融行业企业参与编制和验证，基于众多实际案例的经验总结来进行编写的，并于2015年2月正式发布。DCAM首先定义了数据能力成熟度评估涉及的能力范围和评估的准则，然后从战略、组织、技术和操作的最佳实践等方面描述了如何成功

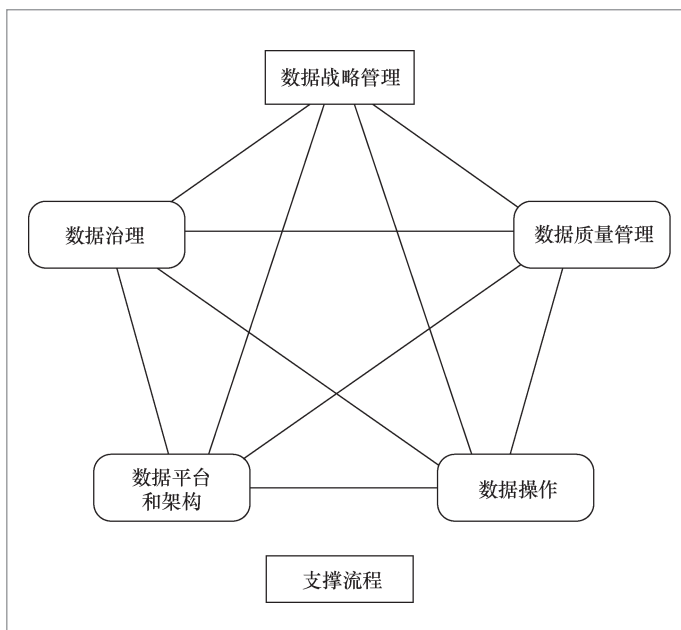


图1 DMM能力模型职能域的划分

地进行数据管理。最后，又结合数据的业务价值和数据操作的实际情况定义数据管理的原则。

如图2所示，在DCAM中，主要分为以下8个职能域^[2]：

- 数据管理策略；
- 数据管理业务案例；
- 数据管理程序；
- 数据治理；
- 数据架构；

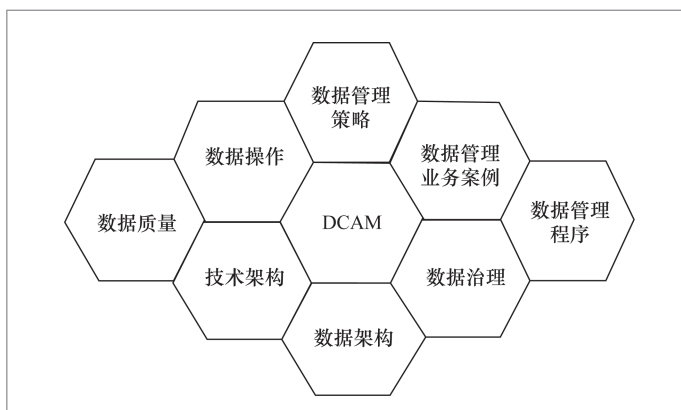


图2 DCAM中数据管理职能域的划分

- 技术架构;
- 数据质量;
- 数据操作。

由于金融是监管驱动的行业,各金融公司都会面临大量的监管需求,例如巴塞尔协议、各国自身的监管需求等,所以DCAM在金融业具有很大的影响力,在DCAM的推广过程中,EDM也在尝试把DCAM和监管需求进行映射,从而可以帮助金融企业更好地满足监管需求。

2.2 国内与国外差异

数据管理能力成熟度模型(data management capability maturity model, DCMM)与以上2个模型最大的差异在于它既吸收了行业公认的部分,又结合了国内数据发展的现实情况,增添了“数据标准”“数据安全”和“数据应用”3个独立的能力项。

- 数据标准:国外的数据管理相关工作中对于数据标准的强调非常少,DAMA数据管理知识体系指南(DAMA guide to the data management body of knowledge, DMBOK)、DMM或者DCAM等文件中都没有关于数据标准的内容,而在国内恰恰相反,在国内很多行业,特别是银行、政府等行业在开展数据治理的过程中,往往会首先制定各自的数据标准。2017年是我国的标准化大年,面对诸多的数据孤岛,数据开放、共享、融合是当前要务,强调数据标准就是强调夯实数据的基础。

- 数据安全:随着数据在单位之间的流动性越来越高,特别是《中华人民共和国网络安全法》的发布和执行,数据安全和隐私的保护也引起了大部分单位的重视,国家也在制定数据安全相关的标准,为此,DCMM也把数据安全作为数据能力

的一个重要维度,意图通过评估来提升各单位的数据安全能力状况。

- 数据应用:数据应用是数据资产价值体现的重要方式,也是数据管理的重要目标,国内很多单位也把数据管理和数据应用放在统一的团队中进行开展,同时也可以通过数据应用来保证数据管理工作的针对性,更利于体现数据管理工作的价值。

3 数据管理能力成熟度模型概述

3.1 数据管理能力成熟度模型含义

数据管理能力成熟度模型是通过一系列的方法、关键指标和问卷来评价某个对象的数据管理现状,从而帮助其查明问题、找到差距、指出方向,并且提供实施建议。

3.2 数据管理能力成熟度模型内容

数据管理能力成熟度模型定义了数据能力成熟度评价的八大能力域:数据战略、数据治理、数据架构、数据标准、数据质量、数据安全、数据应用、数据生命周期管理,如图3所示。这8个能力域又包括29个能力项,见表1。

3.3 数据管理能力成熟度等级划分

数据管理能力成熟度评价划分为5个等级,包括:初始级、受管理级、稳健级、量化管理级和优化级。在此基础上,确定每个层次的基本特征如下,数据能力成熟度等级定义如图4所示。

- 等级一:初始级。组织没有意识到数据的重要性,数据需求的管理主要是在项目级来体现,没有统一的数据管理流

程,存在大量的数据孤岛,经常由于数据的问题导致低下的客户服务质量、繁重的人工维护工作等。

- 等级二:受管理级。组织已经意识到数据是资产,根据管理策略的要求制定了管理流程,指定了相关人员进行初步的管理,并且识别了与数据管理、应用相关的干系人。

- 等级三:稳健级。数据已经被当作实现组织绩效目标的重要资产,在组织层面制定了系列的标准化流程以促进数据管理的规范化,数据的管理者可以快速地满足跨多个业务系统、准确、一致的数据要求,有详细的数据需求响应处理规范、流程。

- 等级四:量化管理级。数据被认为是获取竞争优势的重要资源,组织认识到数据在流程优化、工作效率提升等方面的作用,针对数据管理方面的流程进行全面的优化,针对数据管理的岗位进行关键绩效指标(key performance indicator, KPI)的考核,规范和加强数据相关的管理工作,并且根据过程的监控和分析对整体的数据管理制度和流程进行优化。

- 等级五:优化级。数据被认为是组织生存的基础,相关管理流程能够实时优化,能够在行业内进行最佳实践的经验分享。

4 数据管理能力成熟度模型特征

数据管理能力成熟度模型在制定过程中充分研究了国外理论和实践的发展,同时,充分考虑了国内各行业数据管理发展的现状,并引入了国内数据管理发展相对领先的金融行业的实践经验,保证了模型的创造性、全面性和可操作性。

创造性:数据能力成熟度评价模型是

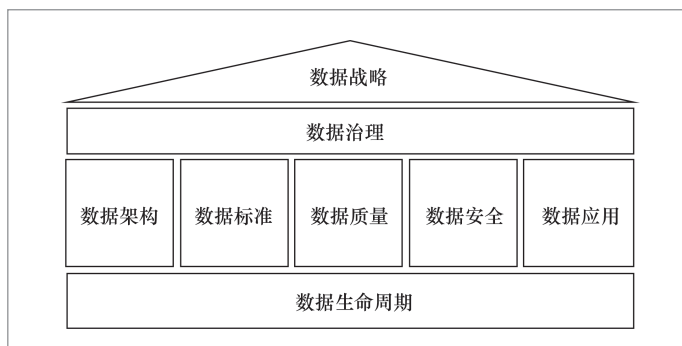


图3 数据管理能力成熟度模型

表1 DCMM的能力域和能力项

能力域	能力项
数据战略	数据战略规划 数据职能框架 数据战略实施 数据任务效益评价
数据治理	数据治理组织 数据制度建设 数据治理沟通
数据架构	数据模型 数据分布 数据集成与共享 元数据管理
数据应用	数据分析 数据开放共享 数据服务
数据安全	数据安全策略 数据安全治理 数据安全审计
数据质量	数据质量需求 数据质量检查 数据质量分析 数据质量提升
数据标准	业务术语 参考数据和主数据 数据元 指标数据
数据生存周期	数据需求 数据设计和开发 数据运维 数据退役

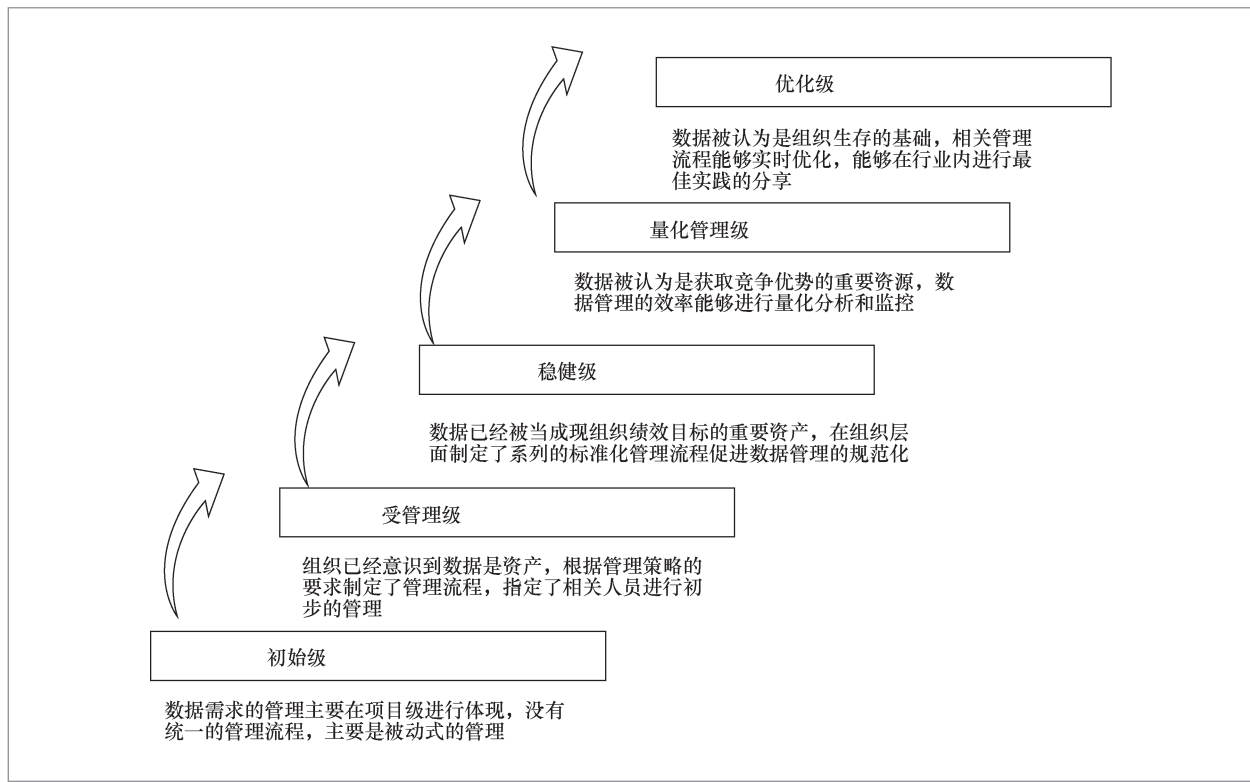


图4 成熟度等级

国内外数据行业发展的崭新事物，目前体系化的数据能力成熟度评价模型基本都处于起步阶段，该模型是国内第一份完整的数据能力成熟度评价标准，对规范国内大数据行业的发展具有重要意义。

全面性：在标准研制的过程中，对数据管理相关的理论进行了充分研究，包括DMBOK^①、DMM、DCAM^③、Gartner（高德纳咨询公司）报告^②等资料，并且充分考虑了国内数据管理行业的发展，包括国家大数据领域的政策以及标杆企业数据管理的整体发展历程。

可操作性：在标准研制的过程中，标准化研究院召集了国内数据管理行业产学研相关的单位，都具备丰富的理论和实践经验，同时，进行了多次标准的试点验证工作，结合试点验证工作的总结，有针对性地对标准进行了完善和修改，保证标准的可操作性。

5 数据管理能力成熟度模型建设目的

数据管理能力成熟度模型是数据管理和应用的基础，将在行业里起到很大的作用。

（1）准确评价各地大数据发展现状

通过对地方上各单位数据管理、应用情况进行评估，可以掌握各单位大数据管理和应用的现状，发现具备的优势和存在的问题，为更好地利用本地的数据资源和进行针对性的指导提供支持。

（2）培养大数据发展人才

大数据产业的发展是技术驱动式的，对人员的技能和素质有很高的要求，通过DCMM的评估，可以对各地方和单位的数据从业人员进行培训，提升数据管理和应用的技能，进而从整体上促进地方和单位数据行业的整体发展。

① www.dama.org

② <https://www.gartner.com/doc/829515/gartner-introduces-eim-maturity-model>

(3) 规范和指导大数据行业发展

大数据行业是相对较新的行业，理论和知识都处于发展阶段，特别是数据管理和应用的知识体系。通过DCMM的评估，可以规范和指导大数据行业的发展，提升从业人员数据资产意识，提升数据技能，推广和传播数据管理最佳实践，从而促进整体行业的发展。

数据管理能力成熟度模型也会对评估企业带来极大的促进发展作用。

(1) 发现存在的问题，指明发展方向

通过对企业DCMM的评估，可以发现企业数据管理过程中存在的问题，并且结合其他企业的最佳实践经验，给出针对性的建议。同时，也可以发现企业数据管理过程中的优点，并加以强化和宣传。

(2) 提升人员技能，建立数据能力提升体系

通过DCMM的评估和培训，可以加强企业内部技术人员、业务人员以及管理人员的数据资产意识，提升相关从业者的技能，理清数据管理、应用建设的思路和框架，规范和指导相关工作的开展。

(3) 持续提升数据能力

开展DCMM评估之后，可以免费获得后续能力提升服务。服务将会从行业专家、最佳实践、行业研讨会、行业报告等多个层面开展相关服务，持续推动行业和公司数据能力水平的提升。

6 应用案例

为了应对大数据环境下数据资产整合、数据标准化管理、数据质量提升等多方面的挑战，某通信企业于2015年9月成立了大数据中心，统一负责某地区大数据相关业务，包括大数据管理、应用产品开发和业务运营，为内部数据服务和外部数据

变现统一提供支撑，并于2016年4月开展了数据能力成熟度评估，评估结论如下。

评估组根据DCMM标准的定义，从DCMM的8个数据能力域展开详细的评估，并根据该企业数据能力相关的制度建设、过程监督和管理、组织人员的建设、工具的应用等多个方面进行评分，根据评分的结果汇总成如图5所示的数据能力成熟度等级分布情况。该企业在数据战略管理域、数据生命周期管理域以及元数据管理域获得了较高分数。这表明该企业制定了明确的数据战略，并且结合元数据、数据架构、数据标准等内容实现了数据资产全生命周期的管理，明确了数据管理组织和职责，对三域数据实现了统一的管理，提升了数据质量，奠定了数据应用和分析的基础。同时，在评估过程中也发现了存在的一些问题，在数据质量考核、数据安全标准、数据标准落地等方面需要进一步的加强和提升，需要在大数据中心平台建设的重点过程中重点进行关注和改善。

通过数据能力成熟度的评估，该企业更加准确地发现了自身存在的问题、与相关公司在数据管理和应用方面存在的差异

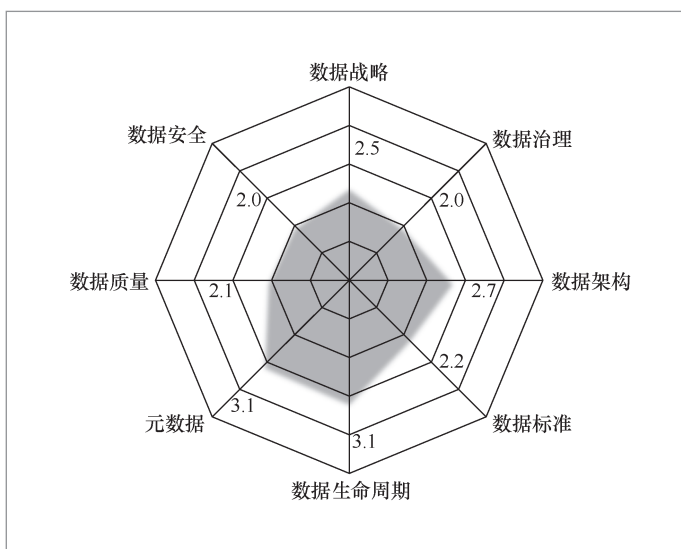


图5 某通信企业评估数据能力成熟度等级分布

以及自身存在的优势,明确了下一步改进的方向,为数据资产的价值变现和提升奠定了基础。

7 结束语

数据管理能力成熟度模型作为一种新兴的技术手段,将为大数据产业带来一种全新的推动力量。该模型的推广应用,将吸纳国内各行业各领域的专业人才,形成专家库,为整个产业的技术力量做人才积蓄;丰富行业数据库,并对行业发展趋势分析提供参考依据。在此过程中还会不断

发现和培育行业标杆,推广优秀案例,使得大众企业受益。

参考文献:

- [1] CMMI Institute. CMMI institute data management maturity model v1.0[R]. [S.l.:s.n.], 2014.
- [2] EDM Council. EDMC DCAM version 1.0[R]. [S.l.:s.n.], 2015.
- [3] AIKEN P, ALLEN M D, PARKER B, et al. Measuring data management practice maturity: a community's self-assessment[J]. Computer, 2007, 40 (4): 42-50.

作者简介



李冰 (1989-), 女, 中国电子技术标准化研究院工程师, 主要研究方向为项目管理、大数据标准化、数据开放共享、数据管理。



宾军志 (1976-), 男, 御数坊(北京)科技咨询有限公司联合创始人, 主要研究方向为数据治理、数据标准、数据质量、主数据等与数据相关的领域。

收稿日期: 2017-01-20