

大数据治理标准体系研究

代红, 张群, 尹卓

中国电子技术标准化研究院, 北京 100007

摘要

大数据逐步从概念导入期转入深化务实应用的新阶段, 大数据治理成为大数据产业生态系统的新热点, 其发展亟需标准体系建设基础以及标准化的支撑。梳理了当前我国大数据治理标准化面临的问题, 明确了与大数据治理相关的概念和定义, 提出了大数据治理标准体系框架, 并给出了下一步标准化工作建议, 从而帮助业界构建涵盖大数据治理的新大数据标准体系, 为我国新阶段的大数据技术、产业发展提供标准化支撑。

关键词

大数据治理; 数据管理; 数据治理; 标准

中图分类号: TP30

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2019023

Study on big data governance standard system

DAI Hong, ZHANG Qun, YIN Zhuo

China Electronics Standardization Institute, Beijing 100007, China

Abstract

As big data has gradually stepped from the concept introduction period to the new stage of deepening pragmatic application, big data governance has become a new hot spot in the big data industry ecosystem. Its development requires the foundation of standard system construction and the support of standardization. The problems faced by China's big data governance standardization were sorted out, the concepts and definitions related to big data governance were clarified, the big data governance standard system framework was proposed, and suggestions for the future standardization work were given, which can help the industry to build a new big data standard system which covering big data governance, and provide standardized support for China's new stage of the development of big data technology and industrial.

Key words

big data governance, data management, data governance, standard

1 引言

当前,我国经济发展步入新常态,创新成为驱动经济社会发展的新引擎。随着信息技术与经济社会的深度融合,数据对社会各行业、各领域创新的支撑作用日益凸显。

《中华人民共和国国民经济和社会发展的第十三个五年(2016—2020年)规划纲要》(以下简称“十三五”规划)中提出“实施国家大数据战略”;2015年国务院专门出台了《促进大数据发展行动纲要》(以下简称《行动纲要》),指出:大数据成为推动经济转型发展的新动力,成为重塑国家竞争优势的新机遇,成为提升政府治理能力的新途径^[1];中国共产党第十九次全国代表大会报告明确指出:“加快建设制造强国,加快发展先进制造业,推动互联网、大数据、人工智能和实体经济深度融合”^[2]。

在我国全面实施国家大数据战略、构建数字经济、建设“数字中国”的大背景下,伴随着各行业的数据量急剧增长,大数据逐步从概念导入期转入了深化务实应用的新阶段。大数据技术与产业的融合发展既是热点也是焦点,既是难点也是重点,既是挑战也是机遇。如何推动大数据在各行业的实际应用,真正发挥大数据对各行业的支撑作用,涉及技术能力、基础设施、建模方式、软件理论、标准等各个方面。

学术界和产业界均意识到,一方面大数据给现有信息技术体系带来了系列挑战,另一方面有利于大数据产业健康有序发展的良好环境还需要营造。为此,大数据治理受到了人们的极大关注,成为大数据产业生态系统的新热点。大数据治理将会深入影响我国工业、政务、城市、金融、能源等领域的发展和进步。在此过程中,如何更好地发挥标准化工作的全面支撑作用,构建大数据治

理标准体系,充分释放数据中蕴含的巨大价值,是一个需要持续关注和解决的问题。

本文围绕“十三五”规划和《行动纲要》的要求,总结了大数据治理标准化面临的问题,结合我国目前大数据治理标准化的现状,针对性地提出了未来在数据治理标准化方面的几点工作建议。

2 大数据治理标准化面临的问题

2.1 缺乏对大数据治理的整体认识

在“互联网+”、大数据与人工智能的科技浪潮之中,随着政府、企事业单位对数据重视程度和对数据价值预期的提升,大数据治理相关问题的严重性、解决问题的迫切性得到了业界一定程度的关注和讨论。我国数据治理实践始于2003年,在2008年之后进入快速发展阶段,至今已覆盖政府、通信、银行、证券、保险、能源、制造、互联网等主要领域。然而,对于大数据治理的相关理念、方法和工具,国内仍然缺乏广泛、深入、一致的理解和行之有效的方案,实践道路曲折艰辛,治理成效并不显著。

2.2 大数据治理标准体系规划尚不完善

大数据治理的核心目标是将数据作为政府及企业的核心资产进行应用和管理。合理的数据治理能够建立规范的数据应用标准,消除数据的不一致性,提高组织内部的数据质量,推动数据的广泛共享,充分发挥大数据对政府及企业的业务、管理以及战略决策的重要支撑作用。目前我国政府、企业的数据治理能力普遍不足,需要建立大数据治理标准体系,通过标准化的手段为政府、企业的数据治理提供指导和规范,

不断扩大标准化在数据治理领域的广泛应用,促进组织完善数据治理机制,提升数据治理能力,加强组织间的数据交换共享,提升数据价值。

2.3 重点标准在行业应用中的支撑力度不够

大数据治理的相关标准需要全社会的广泛参与,同时标准化的推进程度和应用深度对区域中行业产业发展的健康性和持续性有积极的促进作用。标准的制定归根到底要落实到产业应用中。当前,对大数据治理领域重点标准的研制工作正在推进,多个在编大数据治理标准相继发布。然而,随着各行业大数据治理体系的建设,这些标准尚未形成一套有机整体,无法全面支撑大数据在各领域中的应用。围绕重点标准的推广宣贯、试点示范工作有待加强,需不断强化标准对市场培育、服务能力提升和行业管理的支撑作用。

3 大数据治理概述

大数据治理体系建设是一项系统工程,不是简单地针对一个单独的个体、组织或者某一个领域,而是一个全生态体系的运作,大到大数据技术平台的搭建、组织的变革、政策的制定、流程的重组,小到元数据的管理、主数据的整合、各种类型大数据的个性化治理和大数据的行业应用,无不需艰苦细致的工作。梅宏院士在全国信息技术标准化技术委员会(以下简称全国信标委)大数据标准工作组第五次会议中强调了国家大数据治理体系建设及其标准化的重要性。完善大数据治理体系建设的价值在于更好地解决大数据目前面临的问题,如数据资产定位、数

据管理机制、数据开放共享、数据安全保障等,释放海量数据中蕴含的巨大价值,更好地推进大数据和实体经济的深度融合。大数据治理的核心概念目前各方面说法不一,梅宏院士强调的大数据治理体系建设涉及国家、行业、企业层面的工作,需要国家层面的法律法规、标准规范、应用实践和技术进行支撑。

大数据治理还涉及数据治理和数据管理能力,国家标准《信息技术 大数据 术语》(GB/T 35295-2017)^[3]中对数据治理和数据管理能力的定义如下。

- 数据治理:对数据进行处置、格式化和规范化的过程。数据治理是数据和数据系统管理的基本要素;数据治理涉及对数据全生命周期的管理,无论数据处于静态、动态、未完成状态还是交易状态。

- 数据管理能力:指组织和机构对数据进行管理和应用的能力。

数据治理涉及数据治理组织、数据制度建设、数据治理沟通等方面的内容。数据治理组织包括组织架构、岗位设置、团队建设、数据责任等内容,是各项数据职能工作开展的基础;对数据管理和数据应用进行规划和控制,并指导各项数据职能的执行,以确保组织能有效落实数据战略目标。数据制度建设的主要目的是保障数据管理和数据应用各项功能的规范化运行,建立对应的制度体系;数据制度体系通常分层次设计,遵循严格的发布流程,并定期检查和更新;数据制度建设是数据管理和数据应用各项工作有序开展的基础,是数据治理沟通和实施的依据。数据治理沟通旨在确保组织内全部利益相关者都能及时了解相关政策、标准、流程、角色、职责、计划的最新情况,开展与数据管理和应用相关的培训,掌握与数据管理相关的知识和技能;数据治理沟通旨在建立与提升跨部门及部门内部的数据管理能力,提升数据资产意识,构

建数据文化。

4 大数据治理标准体系框架

4.1 在大数据标准体系框架中的位置

标准体系是指在一定范围内的标准按其内在联系形成的科学有机整体。前期,全国信标委大数据标准工作组结合国内外大数据标准化情况、国内大数据技术发展现状,组织研究并发布了大数据标准体系框架^[4],如图1所示。

大数据治理是国家实施大数据战略的重要保证,是发挥大数据作用、做大做强大数据产业的重要因素。为了更好地发挥大数据在各行业中的应用价值,当前各领域大数据治理都开始进行初步实践,大数据标准体系框架应该随着时间的推移和条件的改变不断发展和更新。因此,围

绕国家大数据治理体系的建设,根据大数据自身标准化的特点以及未来大数据发展的趋势,张群^[5]提出了更加完善的大数据标准体系框架(如图2所示),从而使得各项标准在相关内容方面衔接一致、互相补充、互相依存。完善的大数据标准体系框架能够便于了解一个系统内标准的全貌,从而指导大数据标准化工作的进行,提高大数据标准化工作的科学性、全面性、系统性和预见性。

与原有的大数据标准体系框架相比,完善的大数据标准体系框架由8个类别的标准组成,分别为:基础标准、数据标准、技术标准、平台/工具标准、管理标准、安全标准、大数据治理标准和行业应用标准。其中,大数据治理标准作为一项支撑体系,在原有的大数据标准体系基础上可构成一个3层的支撑结构:最底层的是基础标准、数据标准,为大数据治理标准提供底层的术语、参考架构、元数据等标准基

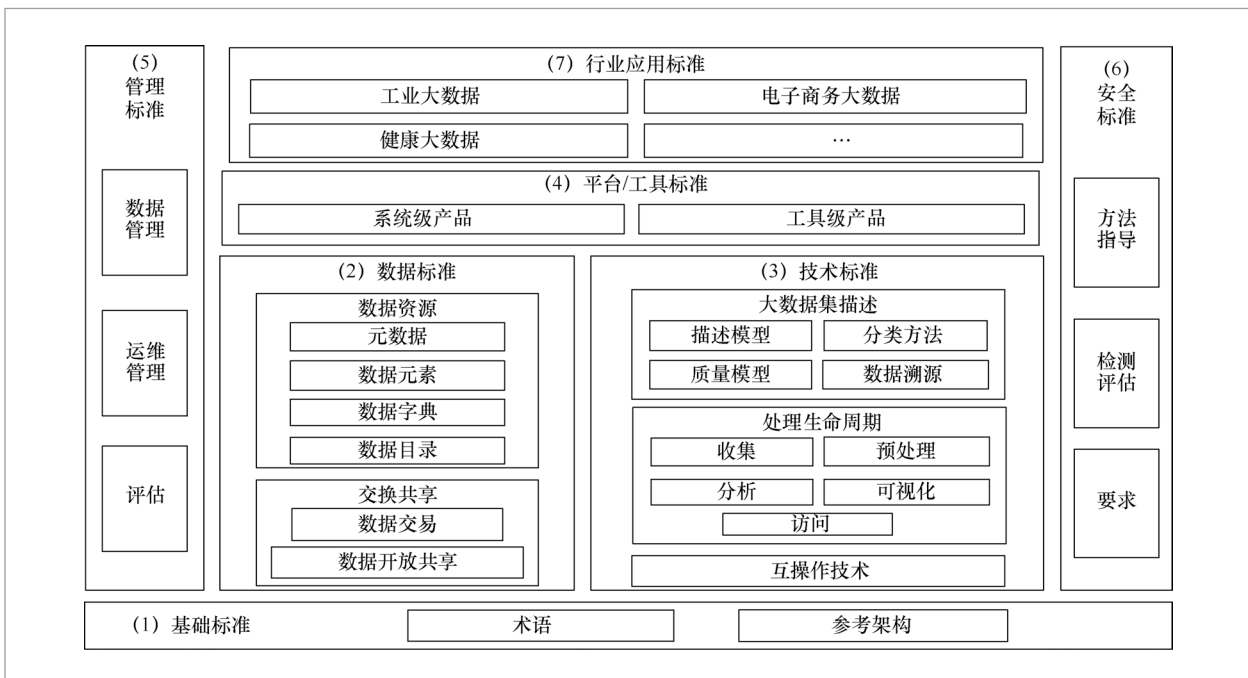


图1 大数据标准体系框架

础；第二层是技术标准、平台/工具标准，从技术和产品的角度为大数据治理提供数据处理分析技术及平台支撑，包括数据全生命周期处理、数据开放与互操作等；第三层是管理标准、安全标准，为大数据治理标准提供管理机制和安全机制。大数据治理标准主要包括咨询/规划标准和解决方案标准两部分，咨询/规划标准主要对大数据治理的路径规划、评估咨询等活动进行规范；解决方案标准则对与大数据治理相关的解决方案及产品进行规范。

4.2 大数据治理重点标准推进

全国信标委大数据标准工作组已经开展了大数据治理标准的研制工作，目前发布了国家标准9项，在研23项，申请立项3项，见表1。

针对大数据治理，全国信标委大数据标准工作组重点围绕《数据治理能力成熟度评估模型》(GB/T 36073-2018)国家标准开展了标准应用推广工作。《数据管理能力成熟度评估模型》规定了组织进行数据管理、评价的能力成熟度模型，包含数据战略、数据治理、数据架构、数据应用、数据安全、数据质量管理、数据标准、数据生命周期管理8个关键过程域，描述了每个过程域的建设目标和度量标准，可以作为组织进行数据管理工作的参考模型。《数据管理能力成熟度评估模型》适用于数据拥有方，使其对数据进行规划、管理和能力提升的规范和指导；同时也适用于数据应用方案提供方，使其进行数据解决方案的建设以及人员的培养等。目前该标准已经发布，并在贵州、广东、四川、上海、深圳等地区以及电力、通信、金融等行业进行了推广应用，有力提升了地方和行业的数字管理水平。

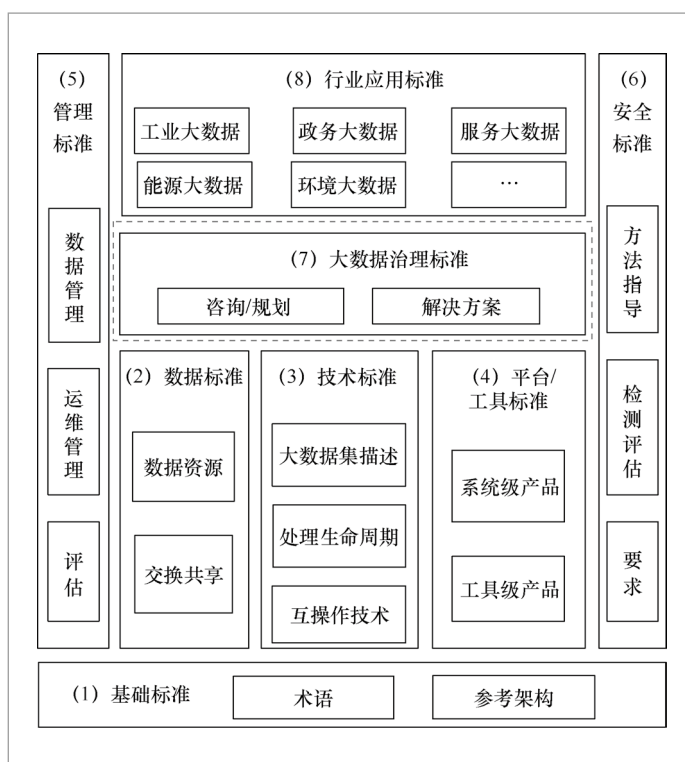


图2 完善的大数据标准体系框架

5 标准化建议

大数据治理标准化工作要坚持问题导向原则，为了完善大数据治理标准体系建设，应重点做好以下工作。

(1) 多元共治，支撑大数据治理标准体系建设

建设“国家层面、行业层面、组织层面”多元共治的大数据治理体系，以标准化工作为切入点，以全国信标委大数据标准工作组、中国电子工业标准化技术协会数据管理应用推进分会为依托，联合各地方主管部门、行业应用单位、科研机构以及相关企业等各方资源，共同解决大数据治理技术与产业融合难题，对接国家大数据标准化工作，为完善大数据治理能力提供整体解决方案。针对目前关键的数据质量、数据管

表1 大数据国家标准

序号	国标号/计划号	标准名称	状态
1	GB/T 35295-2017	信息技术 大数据 术语	发布
2	GB/T 35589-2017	信息技术 大数据 技术参考模型	发布
3	GB/T 34952-2017	多媒体数据语义描述要求	发布
4	GB/T 35294-2017	信息技术 科学数据引用	发布
5	GB/T 34945-2017	信息技术 数据溯源描述模型	发布
6	GB/T 36073-2018	数据管理能力成熟度评估模型	发布
7	GB/T 36343-2018	信息技术 数据交易服务平台 交易数据描述	发布
8	20141201-T-469	信息技术 数据交易服务平台 通用功能要求	报批稿
9	GB/T 36344-2018	信息技术 数据质量评价指标	发布
10	GB/T 36345-2018	信息技术 通用数据导入接口规范	发布
11	20160597-T-469	信息技术 大数据分析系统基本功能要求	报批稿
12	20160598-T-469	信息技术 大数据存储与处理平台技术要求	报批稿
13	20171083-T-469	信息技术 大数据 基于参考架构下的接口框架	草案
14	20171082-T-469	信息技术 大数据 分类指南	草案
15	20171084-T-469	信息技术 大数据 系统通用规范	报批稿
16	20171081-T-469	信息技术 大数据 存储与处理系统功能测试规范	草案
17	20171065-T-469	信息技术 大数据 分析系统功能测试规范	草案
18	20171066-T-469	信息技术 大数据 面向应用的基础计算平台基本性能要求	草案
19	20173818-T-469	信息技术 大数据 系统运维和管理功能要求	草案
20	20171067-T-469	信息技术 大数据 开放共享 第1部分: 总则	草案
21	20171068-T-469	信息技术 大数据 开放共享 第2部分: 政府数据开放共享基本要求	草案
22	20171069-T-469	信息技术 大数据 开放共享 第3部分: 开放程度评价	草案
23	20173819-T-469	信息技术 大数据 工业应用参考架构	草案
24	20173820-T-469	信息技术 大数据 产品要素基本要求	草案
25	20180988-T-469	信息技术 工业大数据 术语	草案
26	20182054-T-339	智能制造 工业数据空间模型	草案
27	20182040-T-339	智能制造 多模态数据融合系统技术要求	草案
28	20182053-T-339	智能制造 工业大数据平台通用要求	草案
29	20182052-T-339	智能制造 工业大数据时间序列数据采集和存储管理框架	草案
30	20190840-T-469	数据管理能力成熟度评估方法	草案
31	20190842-T-469	信息技术 政务信息资源共享评价指标	草案
32	20190841-T-469	信息技术 大数据 面向分析的数据存储与检索技术要求	草案
33	—	信息技术 服务大数据 运维服务元数据	申请立项
34	—	信息技术 大数据数据资源规划	申请立项
35	—	工业企业主数据管理指南	申请立项

理机制、共享开放、安全保障等问题,研究更完善的大数据治理体系,实现全覆盖。

(2) 夯实基础,打通数据管理机制

数据管理机制是大数据治理标准体系建设的重中之重,无论是政治环境导向还是实际应用需求,数据管理都将为

大数据发挥应用价值奠定良好的基础。目前,《数据管理能力成熟度评估模型》国家标准的推广应用工作已经进入了实质性推进阶段,应毫不松懈地继续完善标准推广与持续服务机制,做好标准的应用推广工作。通过国家标准的研制和应用实施,推动提升各组织的数据管理能力水平,为大数据产业的推进夯实基础,为数据在各行各业的流转保驾护航,推动大数据应用在国内各行各业的繁荣和发展。

(3) 标准先行,促进数据流通开放共享

数据开放共享是大数据治理工作中的重要内容,涉及各个行业。一是建立各领域统一的数据开放标准,界定数据开放边界。数据开放是大数据应用体系、服务体系、产业创新体系的基础。开放标准将指导各行各业完善数据开放的策略和计划,制定数据开放管理的流程和制度,构建数据开放服务平台和技术体系,实现数据的清晰、便捷和高效应用。二是建立各领域数据的共享标准,指导建设面向各行各业不同组织机构、不同业务场景、不同复杂度的数据共享系统,支撑各类业务流程,打破跨行业之间的数据壁垒,达成提高全行业整体运营效能的目标。同时,在数据标准和技术要求的指导下,有效地管理共享数据资源,加快数据资产的价值化进程。

(4) 紧跟需求,带动实体经济发展

与国际数据治理专业领域在理论、知识、方法和技术工具的原生性发展特点不同,我国信息技术发展和信息化进程起步较晚,大数据治理的发展主要体现在各行业信息化进程中的数据治理实践不断全面和深化。为了更好地满足实际需求,未来将重点在制造、城市、能源、金融等行业开展大数据治理研究,以标准化为抓手,打

造良好的产业生态环境,为产业转型升级贡献力量。

6 结束语

本文总结了我国大数据治理标准化面临的主要问题,给出了大数据治理概述,包括数据治理的定义。结合大数据标准体系现状,完善了大数据标准体系框架,给出了大数据治理标准体系在大数据标准体系中的位置以及重点相关标准的推进情况。最后,结合我国大数据治理的技术、产业发展现状,给出了针对大数据治理的下一步标准化建议,为后续的标准化工作提出了方向。

参考文献:

- [1] 中华人民共和国国务院. 促进大数据发展行动纲要[R]. 北京: 中华人民共和国国务院, 2015.
The State Council of the People's Republic of China. Action platform for greater data development[R]. Beijing: The State Council of the People's Republic of China, 2015.
- [2] 中国共产党中央委员会. 中华人民共和国国民经济和社会发展第十三个五年规划纲要[N]. 北京: 人民日报, 2016-03-17.
Central committee of the Communist Party of China. The 13th five-year plan for economic and social development of the People's Republic of China[N]. Beijing: People's Daily, 2016-03-17.
- [3] 《信息技术 大数据 术语》: GB/T 35295-2017[S]. 北京: 全国信息技术标准化技术委员会, 2017.
“Information technology-big data-terminology”: GB/T 35295-2017 [S]. Beijing: China National Information Technology Standardization Network, 2017.
- [4] 张群, 吴东亚, 赵菁华. 大数据标准体系[J]. 大数据, 2017, 3(4): 11-19.

ZHANG Q, WU D Y, ZHAO J H. Big data standards system[J]. Big Data Research, 2017, 3(4): 11-19.

- [5] 张群. 大数据标准化现状及标准研制[J]. 信息技术与标准化, 2015(7): 23-26.

ZHANG Q. Big data standardization current situation and standards development[J]. Information Technology & Standardization, 2015(7): 23-26.

作者简介



代红 (1966-), 女, 中国电子技术标准化研究院信息技术研究中心主任, 全国信息技术标准化技术委员会委员和国家民委民族语文工作专家咨询委员会委员, 主要从事信息技术标准化工作, 负责起草研制了150余项信息技术国家标准和行业标准, 组织承担多项省部级科研项目, 获得国家科技进步奖二等奖及多项国家和部委科技成果奖, 在国家级刊物上发表过多篇科技论文。



张群 (1988-), 女, 博士, 中国电子技术标准化研究院大数据与人工智能研究室高级工程师、主任, 主要从事信息技术相关标准的制定、修订及研究工作, ISO/IEC JTC1/SC 42/WG 2大数据工作组专家, 重点研究方向为大数据和人工智能领域相关技术及标准。



尹卓 (1990-), 男, 中国电子技术标准院研究院标准化工程师, 主要从事信息技术相关标准化工作, ISO/IEC JTC 1/SC 42/WG 2大数据工作组专家, 重点研究大数据领域技术及相关标准。

收稿日期: 2019-01-04