

DB 21

大 连 市 地 方 标 准

DB 21/T XXXX. 2—XXXX

大数据标准体系框架

Big data standard system framework

（征求意见稿）

（本草案完成时间：2021.05.17）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX – XX – XX 发布

XXXX – XX – XX 实施

大连市市场监督管理局 发 布

目次

前 言..... II

1 范围..... 3

2 规范性引用文件..... 3

3 术语和定义..... 3

4 总体要求..... 3

5 标准体系结构图..... 4

 5.1 概述..... 4

 5.2 分体系结构图..... 6

附录 大数据标准明细表..... 13

前 言

本文件按照GB/T1.1—2020《标准化工作导则第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由软件和信息技术服务标准化技术委员会（辽宁）提出。

本文件由大连市工业和信息化局归口。

本文件起草单位：大连理工大学计算机科学与技术学院、东北财经大学、大连海洋大学、大连大学、大连高新区大数据产业协会、大连市人工智能产业协会、东软集团（大连）有限公司、大连数据湖信息技术有限公司、大连景泰科技有限公司、大连开一信息科技有限公司（待定）。

本文件主要起草人：王宇新、祝志杰、于红、于凤岐、刘长虹、钟将、刘炳军、李俊、尹宏、周东生、曲鹤（待定）。

大数据标准体系框架

1 范围

本标准给出大连市的大数据标准体系框架。

本标准适用于以大数据作为技术要素的标准研制工作。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T13016 标准体系表构建原则和要求

GB/T35295—2017 信息技术大数据 术语

3 术语和定义

GB/T35295-2017 界定的以及下列术语和定义适用于本文件。

下列术语和定义适用于本文件。

3.1

标准 standard

为在一定的范围内获得最佳秩序，经协商一致制定并由公认机构批准，共同使用的和重复使用的一种规范性文件。

3.2

标准体系 standard system

一定范围内的标准按其内在联系形成的科学的有机整体。[GB/T13016—2018，定义 2.4]

3.3

大数据 big data

具有体量巨大、来源多样、生成极快、且多变等特征并且难以用传统数据体系机构有效处理的包含大量数据集的数据。

4 总体要求

要求如下：

DB 11 FORMTEXT 11 21/T1 11 FORMTEXT 11 XXXX.21 —11 FORMTEXT 11 XXXX1

- 标准体系编制原则应符合 GB/T13016 要求；
- 标准体系应系统全面、层次恰当、划分明确、相互协调。
- 标准体系结构图应由一个总体系结构图和多个分体系结构图构成体系结构图内方框间用实线连接,表示方框间的层次关系、序列关系。

5 标准体系结构图

5.1 概述

大数据标准体系框架由九个类别的标准组成,分别为:基础、数据、技术、平台/工具、管理、安全和隐私、行业应用、行业人才、行业企业,总体系结构图见图 1。

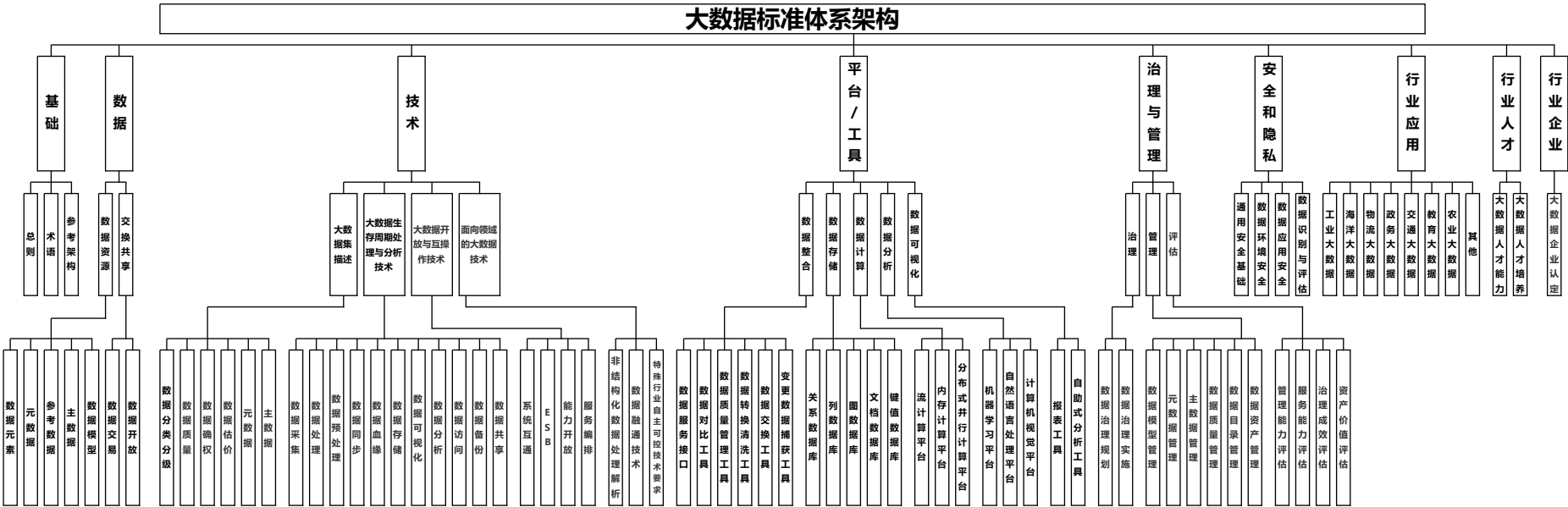


图1 总体系结构图

5.2 分体系结构图

5.2.1 基础

基础类标准：为整个标准体系提供包括总则、术语、参考架构等通用的基础性标准。为大数据其他部分的标准制定提供基础遵循，支撑各行业、各领域对大数据达成统一理解。

基础类标准分为（分体系结构图见图2）：

——**总则类标准：**针对大数据地方标准化指南等进行规定的标准，包括大数据技术参考模型、标准化顶层设计和总体要求、标准体系中列出的数据、技术、平台/工具、管理、安全和隐私、行业应用、大数据人才、大数据企业等各类标准之间的关系与要求，适用于大连市大数据领域建设规划、实施和管理；

——**术语类标准：**针对大数据及其应用领域术语和定义、密切相关的通用术语和定义等内容进行规定的标准；

——**参考架构类标准：**针对大数据技术参考模型、参考框架和应用指南、参考框架用例和需求及基于参考架构的接口框架等内容进行规定的标准。

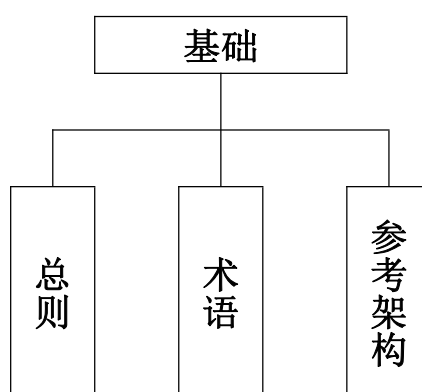


图2 基础类标准分体系结构

5.2.2 数据

数据类标准：为大数据标准体系提供数据层面的准则，数据标准主要包括数据资源标准和交换共享标准两个部分。

数据类标准分为（分体系结构图见图3）：

——**数据资源类标准：**数据资源标准主要从元数据、主数据、数据模型和数据口径四个方面进行规范。

——**交换共享类标准：**主要针对数据交易、数据开放和数据共享功能进行标准化。

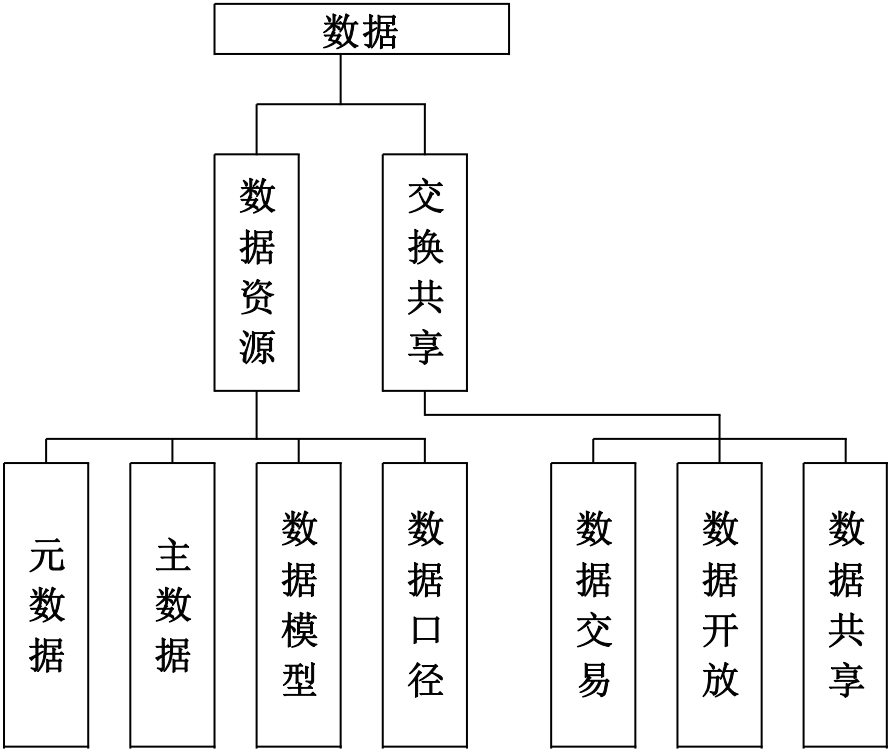


图3 数据类标准分体系结构图

5.2.3 技术

技术类标准：针对大数据通用技术进行规范。包括大数据集描述、大数据生存周期处理与分析技术、大数据开放与互操作技术、面向领域的大数据技术四类。

技术类标准分为（分体系结构图见图4）：

- 大数据集描述类标准：**主要针对多样化、差异化、异构异质的不同类型的数据建立标准的度量方法，以衡量数据质量，包括：数据分类分级、数据质量、数据确权、数据估价、元数据、主数据等相关技术标准。
- 大数据生存周期处理与分析技术类标准：**主要针对大数据产生到其使用终止这一过程的关键技术进行标准制定，包括数据采集、数据预处理、数据存储、数据分析、数据可视化、数据访问等标准。
- 大数据开放与互操作技术类标准：**主要针对不同功能层次功能系统之间的互联与互操作机制、不同技术架构系统之间的互操作机制、同质系统之间的互操作机制等相关标准以及通用数据开放共享技术框架等标准，包括：系统互通、ESB、能力开放、服务编排等。
- 面向领域的大数据技术类标准：**主要针对工业、海洋、电子政务等领域或行业的共性且专用的大数据技术标准。

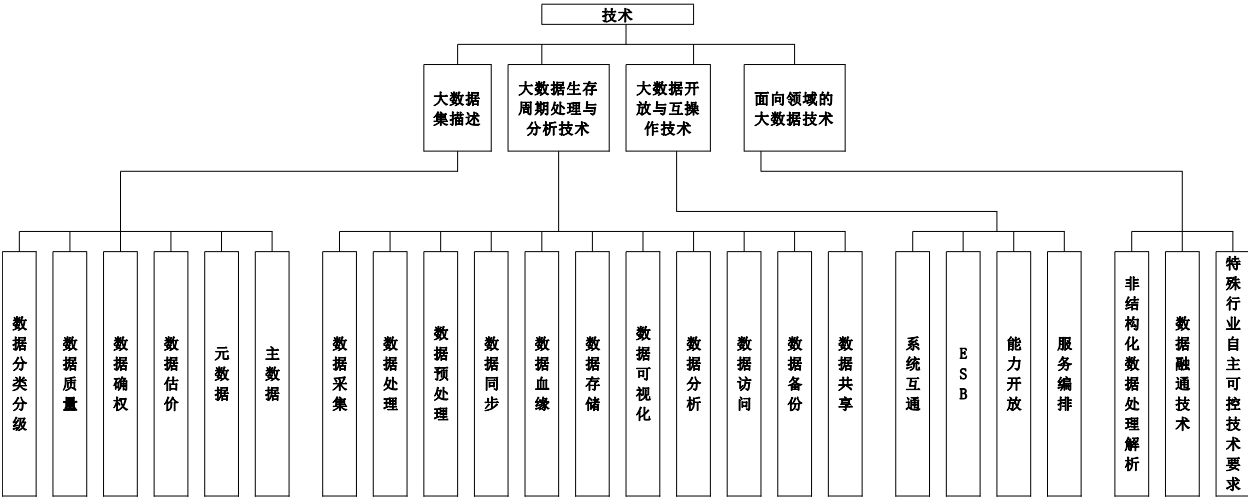


图 4 技术类标准分体系结构图

5.2.4 平台/工具

平台/工具类标准主要针对大数据相关平台及工具产品进行规范。

平台或工具类标准分为（分体系结构图见图 5）：

- 数据整合类标准：**针对数据服务接口、数据比对工具、数据质量管理工具、数据转换和清洗工具、数据交换工具和变更数据捕获工具的功能和性能进行规范。
- 数据存储类标准：**针对关系数据库、列数据库、图数据库、文档数据库和键值数据库的功能和性能进行规范。
- 数据计算类标准：**针对流计算平台、内存计算平台和分布式并行计算平台的功能和性能进行规范。
- 数据分析类标准：**针对机器学习平台、自然语言处理平台和计算机视觉等平台的功能和性能进行规范。
- 数据可视化类标准：**针对报表工具和自助式分析工具等数据可视化平台或工具的功能和性能进行规范。

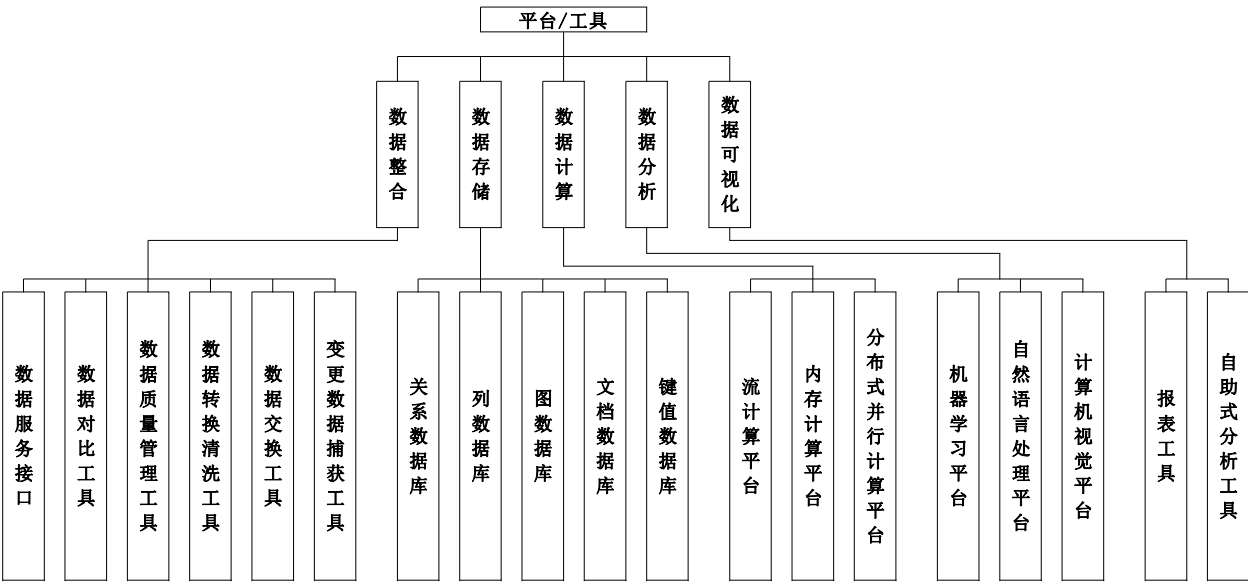


图 5 平台/工具类标准分体系结构图

5.2.5 治理与管理

治理与管理类标准：贯穿于数据生存周期的各个阶段，是大数据实现高效采集、分析、应用、服务的重要支撑。该类标准主要包括治理标准、管理标准和评估标准三部分。

治理与管理类标准分为（分体系结构图见图 6）：

- 治理类标准**：主要对数据治理的规划和具体实施方法进行标准研制。
- 管理类标准**：主要面向数据管理模型、元数据管理、主数据管理、数据质量管理、数据目录管理以及数据资产管理等理论方法和管理工具进行规范。
- 评估类标准**：在治理标准和管理标准的基础之上，总结形成针对数据管理能力、数据服务能力、数据治理成效、数据资产价值的评估方法。

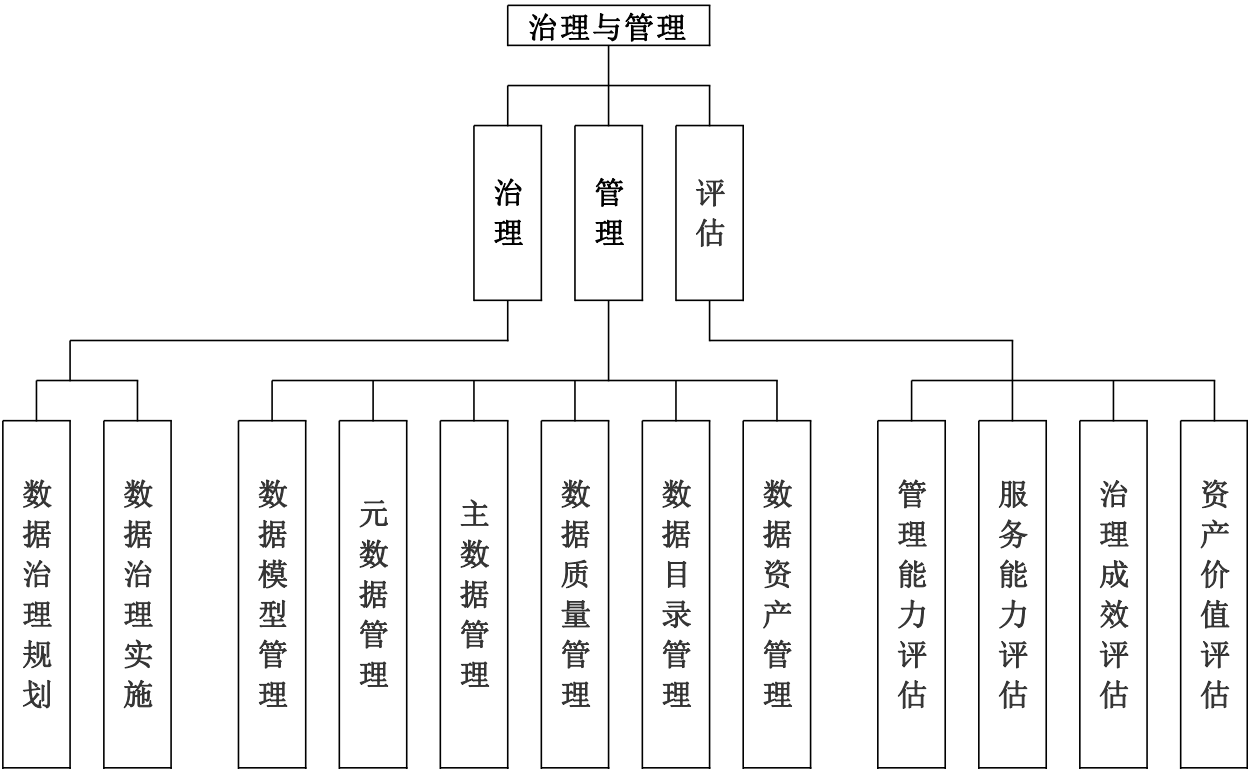


图6 治理与管理类标准分体系结构图

5.2.6 安全和隐私

安全和隐私类标准主要针对通用安全基础、数据环境安全、数据应用安全和数据识别和评估等方面进行规定。

安全和隐私类标准分为（分体系结构图见图7）：

- 通用安全基础类标准：针对数据和隐私安全方法提出的基础类标准，主要用于指导数据安全普遍性要求标准制定；包括数据识别、存储、传输、交换、运算、销毁等数据生命周期安全管理；
- 数据环境安全类标准：针对数据环境、平台提出的标准，主要用于指导数据存储物理、硬件、平台要求标准制定；
- 数据应用安全类标准：针对个人和组织数据等提出的数据应用安全类标准；用于指导个人信息和组织数据保护要求标准制定；
- 数据识别与评估类标准：针对个人和组织数据安全评价、评估类标准，用于指导数据安全的评价、评估要求标准制定。

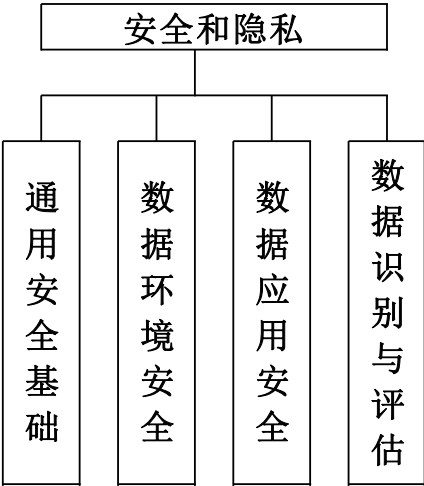


图 7 安全和隐私类标准分体系结构图

5.2.7 行业应用

行业应用类标准：主要从大数据为各行业提供的服务角度出发，根据本地各行业领域特性产生的专用数据标准，包括工业大数据、政务大数据、海洋大数据、交通大数据等领域。

行业应用类标准分体系结构图见图 8。

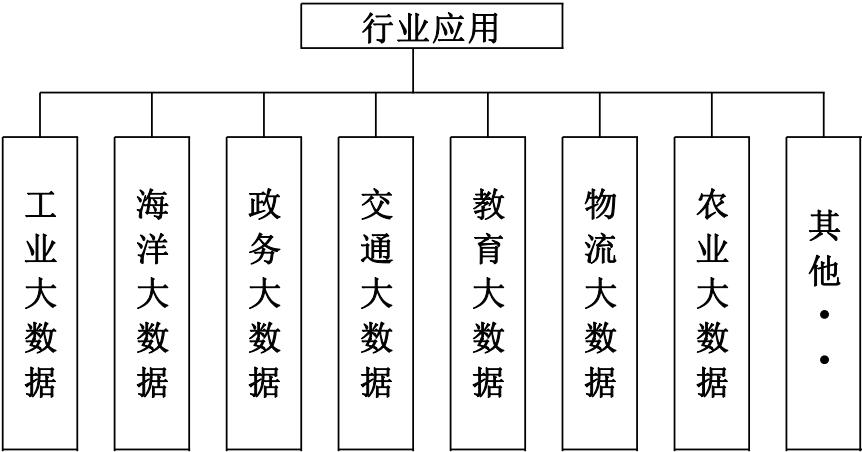


图 8 行业应用类标准分体系结构图

5.2.8 行业人才

行业人才类标准：主要是针对大数据人才的能力、培养等内容进行规定。

行业人才类标准分为（分体系结构图见图 9）：

- 大数据人才能力类标准**：针对大数据人员职业定义、技能等级要求、数据素养等进行标准化规定。
- 大数据人才培养类标准**：针对大数据人才知识体系架构、大数据人才培养、大数据人才认证等进

行规定的标准。

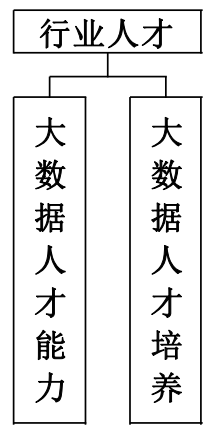


图 9 行业人才类标准分体系结构图

5.2.9 行业企业

行业企业类标准主要是针对大数据企业认定进行规范。针对大数据企业的认定原则、大数据企业能力要求、大数据企业分类等内容进行规定的标准。

行业企业类标准分为（分体系结构图见图 10）：

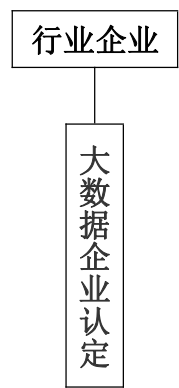


图 10 行业企业类标准分体系结构图

附录 大数据标准明细表

序号	一级分类	二级分类	国家标准编号	标准名称	状态
1	基础	总则	——	信息技术大数据标准化指南	拟研制
2		术语	GB/T35295-2017	信息技术大数据术语	发布
3		参考架构	GB/T35589-2017	信息技术大数据技术参考模型	发布
4			GB/T38672-2020	信息技术大数据接口基本要求	发布
5			ISO/IECTR20547-1	信息技术大数据参考架构第 1 部分架构与应用指南	发布
6			ISO/IECTR20547-2	信息技术大数据参考架构第 2 部分用例和需求	发布
7			ISO/IECTR20547-3	信息技术大数据参考架构第 3 部分参考架构	发布
8			ISO/IECTR20547-4	信息技术大数据参考架构第 4 部分安全与隐私	发布
9			ISO/IECTR20547-5	信息技术大数据参考架构第 5 部分标准路线图	发布
10			——	信息技术大数据基于参考架构下的接口框架	在研
11	数据	数据资源	GB/T18142-2017	信息技术数据元素值表示格式记法	发布
12			——	信息技术数据元素值表示——格式记法	在研
13			——	信息技术元数据属性	在研
14			——	信息技术大数据开放共享第 4 部分：共享程度评价	拟研制
15			——	信息技术大数据开放数据集基本要求	拟研制
16			——	信息技术大数据开放数据集标识管理	拟研制
17			GB/T18391.1-2009	信息技术元数据注册系统(MDR)第 1 部分：框架	发布
18			GB/T18391.2-2009	信息技术元数据注册系统(MDR)第 2 部分：分类	发布
19			GB/T18391.3-2009	信息技术元数据注册系统(MDR)第 3 部分：注册系统元模型与基本属性	发布
20			GB/T18391.4-2009	信息技术元数据注册系统(MDR)第 4 部分：数据定义的形成	发布
21			GB/T18391.5-2009	信息技术元数据注册系统(MDR)第 5 部分：命名和标识原则	发布
22			GB/T18391.6-2009	信息技术元数据注册系统(MDR)第 6 部分：注册	发布
23			GB/Z21025-2007	XML 使用指南	发布
24			GB/T23824.1-2009	信息技术实现元数据注册系统内容一致性的规	发布

序号	一级分类	二级分类	国家标准编号	标准名称	状态
				程第 1 部分：数据元	
25			GB/T23824.3-2009	信息技术实现元数据注册系统内容一致性的规程第 3 部分：值域	发布
26			GB/T32392.1-2015	信息技术互操作性元模型框架(MFI)第 1 部分：参考模型	发布
27			GB/T32392.2-2015	信息技术互操作性元模型框架(MFI)第 2 部分：核心模型	发布
28			GB/T32392.3-2015	信息技术互操作性元模型框架(MFI)第 3 部分：本体注册元模型	发布
29			GB/T32392.4-2015	信息技术互操作性元模型框架(MFI)第 4 部分：模型映射元模型	发布
30			GB/T32392.5-2018	信息技术互操作性元模型框架(MFI)第 5 部分：过程模型注册元模型	发布
31			GB/T32392.7-2018	信息技术互操作性元模型框架第 7 部分：服务型注册元模型	发布
32			GB/T32392.8-2018	信息技术互操作性元模型框架第 8 部分：角色与目标模型注册元模型	发布
33			GB/T32392.9-2018	信息技术互操作性元模型框架第 9 部分：按需模型选择	发布
34			GB/T30881-2014	信息技术元数据注册系统(MDR)模块	发布
35		交换共享	——	数据开放共享第 1 部分：总则	在研
36			——	数据开放共享第 2 部分：政府数据开放共享基本技术要求	在研
37			——	数据开放共享第 3 部分：开放程度评价	在研
38			——	信息技术大数据开放共享第 4 部分：政府资源目录体系	拟研制
39			——	信息技术数据交易平台交易数据描述	在研
40			——	信息技术数据交易通用概念描述	拟研制
41			——	信息技术数据交易交易流程描述	拟研制
42			——	信息技术数据交易数据管理规范	拟研制

序号	一级分类	二级分类	国家标准编号	标准名称	状态
43			——	信息技术数据交易技术规范	拟研制
44			——	信息技术数据交易风险评估	拟研制
45			——	信息技术数据交易交易质量评估	拟研制
46			——	信息技术数据交易数据价值评估指引	拟研制
47			GB/T36343-2018	信息技术数据交易服务平台交易数据描述	发布
48			GB/T37728-2019	信息技术数据交易服务平台通用功能要求	发布
49	技术	大数据集描述	GB/T32909-2016	非结构化数据表示规范	发布
50			——	信息技术数据质量检测	拟研制
51			——	数据能力成熟度评价模型	在研
52			——	大数据解决方案基本评估规范	拟研制
53			——	数据溯源描述模型	在研
54			ISO8000-110	主数据的语法和定义	发布
55			ISO8000-110	主数据源	发布
56			ISO8000-110	主数据准确性	发布
57			ISO8000-110	主数据完整性	发布
58			——	信息技术大数据分类指南	在研
59			GB/T34952-2017	多媒体数据语义描述要求	发布
60			GB/T38667-2020	信息技术大数据数据分类指南	发布
61			GB/T36344-2018	信息技术数据质量评价指标	发布
62			GB/T35294-2017	信息技术科学数据引用	发布
63			GB/T34945-2017	信息技术数据溯源描述模型	发布
64		大数据生存周期处理与分析技术	GB/T32908-2016	非结构化数据访问接口规范	发布
65			GB/T36345-2018	信息技术通用数据导入接口规范	发布
66			20190841-T-469	信息技术大数据面向分析的数据检索与存储技术要求	在研
67			——	信息技术通用数据导入接口测试规范	拟研制
68		大数据开放与互操作技术	——	信息技术大数据分析总体技术要求	拟研制
69			——	信息技术大数据互操作技术指南	拟研制

序号	一级分类	二级分类	国家标准编号	标准名称	状态
70	平台/ 工具	数据整合	GB/T34949-2017	实时数据库 C 语言接口规范	发布
71			GB/T32633-2016	分布式关系数据库服务接口规范	发布
72			——	实时数据库通用接口规范	在研
73		数据存储	GB/T37722-2019	信息技术大数据存储与处理系统功能要求	发布
74			GB/T38676-2020	信息技术大数据存储与处理系统功能测试要求	发布
75			GB/T38673-2020	信息技术大数据大数据系统基本要求	发布
76			GB/T28821-1012	关系数据管理系统技术要求	发布
77			GB/T32630-2016	非结构化数据管理系统技术要求	发布
78			——	信息技术大数据图数据库系统技术要求	拟研制
79			GB/T12991-2008	信息技术数据库语言 SQL 第 1 部分：框架	发布
80			——	信息技术大数据存储与处理平台技术要求	在研
81		数据计算	GB/T38675-2020	信息技术大数据计算系统通用要求	发布
82			——	非结构化数据查询语言	拟研制
83			——	信息技术大数据批流融合计算系统技术要求	拟研制
84		数据分析	GB/T37721-2019	信息技术大数据分析系统功能要求	发布
85			GB/T38643-2020	信息技术大数据分析系统功能测试要求	发布
86			——	信息技术大数据分析系统功能测试规范	在研
87			——	信息技术大数据分析系统基本功能要求	在研
88		数据可视化	——	大数据可视化工具通用要求	拟研制
89	治理与 管理	治理	20194186-T-469	信息技术大数据数据资源规划	在研
90			——	信息技术大数据数据治理体系建设指南	拟研制
91			——	城市数据治理能力成熟度模型	拟研制
92			——	信息技术大数据数据治理工具基本要求	拟研制
93		管理	GB/T36073-2018	数据管理能力成熟度评估模型	发布
94			——	信息技术大数据资产管理指南	拟研制
95			——	信息技术大数据系统运维与管理功能要求	拟研制
96			——	信息技术服务大数据运维服务元数据	拟研制
97			——	云计算服务运行监管指标及接口要求	拟研制

序号	一级分类	二级分类	国家标准编号	标准名称	状态
98			——	信息技术大数据主数据管理规范	拟研制
99		评估	20190840-T-469	数据管理能力成熟度评估方法	在研
100			——	信息技术大数据数据服务能力评估第 1 部分：评估模型	拟研制
101			——	信息技术大数据数据服务能力评估第 2 部分：评估过程	拟研制
102			——	信息技术大数据数据治理绩效评价	拟研制
103			——	信息技术大数据数据资产评估	拟研制
104	安全和隐私	数据应用安全	——	信息安全技术健康医疗信息安全指南	在研
105			——	信息安全技术个人信息安全工程指南	在研
106			GB/T37964-2019	信息安全技术个人信息去标识化指南	发布
107			GB/T34978-2017	信息安全技术移动智能终端个人信息保护技术要求	发布
108			GB/T35273-2020	信息安全技术个人信息安全规范	发布
109			——	信息安全技术个人信息保护管理要求	在研
110			——	信息安全技术个人信息保护指南	拟研制
111			——	个人信息告知同意指南	拟研制
112			——	政务信息资源安全分级指南	拟研制
113			——	信息安全技术金融信息保护规范	在研
114			——	重要数据业务运营安全规范	在研
115			——	政务信息共享数据安全规范	在研
116		数据识别与评估	GB/T35274-2017	信息安全技术大数据服务安全能力要求	发布
117			GB/T37932-2019	信息安全技术数据交易服务安全要求	发布
118			——	信息安全技术云计算服务持续监管框架及技术规范	在研
119			——	大数据业务安全风险控制实施指南	拟研制
120			20180840-T-469	信息安全技术个人信息安全影响评估指南	在研
121			GB/T37988-2019	信息安全技术数据安全能力成熟度模型	发布
122			20180840-T-469	信息安全技术个人信息安全影响评估指南	在研
123			GB/T20009-2005	信息安全技术数据库管理系统安全评估准则	发布

序号	一级分类	二级分类	国家标准编号	标准名称	状态
124			——	信息安全技术个人隐私保护评估方法	拟研制
125			——	具备数据留存功能的网络安全产品个人信息保护测评技术要求研究	在研
126			——	信息安全技术云计算服务安全能力评估方法	在研
127			——	信息安全技术数据出境安全评估指南	在研
128			——	大数据服务安全可控评价指标	拟研制
129		数据应用安全	——	政务信息资源共享安全标准体系研究	拟研制
130			——	大数据安全参考框架	拟研制
131			——	能源企业大数据应用安全防护指南	拟研制
132			——	云服务数据安全指南	拟研制
133			——	数据安全分类分级实施方法	拟研制
134			——	网络安全态势感知和数据规范	拟研制
135			——	信息技术大数据安全参考架构	拟研制
136			——	信息技术大数据安全分级指南	拟研制
137			——	信息技术大数据脱敏指南	拟研制
138			GB/T22080-2008	信息技术安全技术信息安全管理体系要求	发布
139			GB/T22081-2008	信息技术安全技术信息安全管理体系实用规则	发布
140			GB/T31496-2015, ID T	信息技术安全技术信息安全管理体系实施指南	发布
141			GB/T37973-2019	信息安全技术大数据安全管理指南	发布
142		数据环境安全	——	信息安全技术网站安全云防护平台技术要求	在研
143			——	大数据基础软件安全技术要求	拟研制
144			GB/T20273-2006	信息安全技术数据库管理系统安全技术要求	发布
145			GB/Z28828-2012	信息安全技术公共及商用服务信息系统个人信息保护指南	发布
146			——	信息安全技术大数据中的隐私保护框架	拟研制
147			——	基于大数据的数字化转型能力要求	拟研制
148			——	区块链安全标准体系研究	拟研制
149			——	基于区块链的电子货币安全技术标准研究	拟研制
150			——	信息技术大数据系统软件安全技术要求	拟研制

序号	一级分类	二级分类	国家标准编号	标准名称	状态
151			——	信息技术大数据跨集群安全技术框架	拟研制
152			——	信息技术大数据平台安全技术要求	在研
153			——	工业互联网平台安全要求及评估规范	拟研制
154			——	混合云安全技术要求	拟研制
155			——	政务信息共享安全标准体系研究	拟研制
156			——	人工智能安全标准研究	拟研制
157			——	政务云网络安全服务接口指南	拟研制
158			——	工业互联网平台安全标准体系研究	拟研制
159	行业应用	工业大数据	GB/T38666-2020	信息技术大数据工业应用参考架构	发布
160			GB/T38555-2020	信息技术大数据工业产品核心元数据	发布
161			20180988-T-469	信息技术工业大数据术语	在研
162			20182054-T-339	智能制造工业数据空间模型	在研
163			20182040-T-339	智能制造多模态数据融合系统技术要求	在研
164			20182053-T-339	智能制造工业大数据平台通用要求	在研
165			20182052-T-339	智能制造工业大数据时间序列数据采集和存储框架	在研
166			——	信息技术工业大数据数据分类	拟研制
167			——	信息技术工业大数据工业订单元数据规范	拟研制
168			——	工业大数据数据质量评价指南	拟研制
169			——	工业互联网标准体系框架	拟研制
170			——	工业大数据装备数据采集规范	拟研制
171			——	信息技术大数据产品要素基本求	在研
172			——	智能制造对象标识要求	在研
173			——	智能制造标识解析体系要求	在研
174			——	数字化车间通用技术要求	在研
175			——	数字化车间机床制造信息模型	在研
176			——	信息技术工业云服务模型	在研
177			——	信息技术工业云服务能力总体要求	在研
178			——	机器人大数据采集规范	拟研制
179			——	智能制造制造对象标识解析体系应用指南	在研

序号	一级分类	二级分类	国家标准编号	标准名称	状态
180		政务大数据	GB/T38664. 1-2020	信息技术大数据政务数据开放共享第 1 部分：总则	发布
181			GB/T38664. 2-2020	信息技术大数据政务数据开放共享第 2 部分：基本要求	发布
182			GB/T38664. 3-2020	信息技术大数据政务数据开放共享第 3 部分：开放程度评价	发布
183			20190842-T-469	信息技术大数据政务数据开放共享第 4 部分：共享评价	在研
184		海洋大数据	——	信息技术生态环境大数据数据分类指南	拟研制
185			——	信息技术生态环境大数据数据体系	拟研制
186		物流大数据	DB/2102T 0022-2021	海运中转数据交换 中转申请报文	发布
187			DB/2102T 0024-2021	集装箱海运动态查询-数据交换标准	发布
188			DB/2102T 0025. 1-2021	检验检疫口岸电子数据交换标准 第 1 部分：分卸	发布
189			DB/2102T 0025. 2-2021	检验检疫口岸电子数据交换标准 第 2 部分：移箱	发布
190			DB/2102T 0025. 3-2021	检验检疫口岸电子数据交换标准 第 3 部分：放行	发布
191			DB2102/T 0023. 1-2021	集装箱码头堆场自动化改造 第 1 部分：接口规范	发布
192			DB2102/T 0023. 2-2021	集装箱码头堆场自动化改造 第 2 部分：业务流程	发布
193			DB2102/T 0021-2021	电子放箱业务 放箱单报文	发布
194			DB2102/T 0013-2020	散杂货汽车衡海关监管信息传输格式	发布
195			DB2102/T 0012-2020	码头无人值守汽车衡消息传输格式	发布
196		教育大数据	——	信息技术服务大数据教育行业督导平台技术规范	拟研制