

中华人民共和国国家标准

GB/T 44109—2024

信息技术 大数据 数据治理实施指南

Information technology—Big data—Guidelines of data governance implementation

2024-05-28 发布

2024-12-01 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言 III

引言 IV

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 数据治理实施过程 2

5 规划 3

 5.1 现状调研及分析 3

 5.2 确立目标 3

 5.3 明确治理内容 3

 5.4 确定实施路线 3

6 执行 4

 6.1 概述 4

 6.2 建立治理组织 4

 6.3 制定制度规范 4

 6.4 开展治理活动 4

 6.4.1 数据架构设计 4

 6.4.2 元数据管理 5

 6.4.3 数据标准管理 5

 6.4.4 数据质量管理 6

 6.4.5 主数据管理 6

 6.4.6 数据应用 7

 6.4.7 数据安全治理 7

 6.4.8 数据生存周期管理 8

7 评价 8

 7.1 明确评价目标和范围 8

 7.2 设计指标体系 8

 7.3 评价治理绩效 8

8 改进 9

 8.1 进行差异分析 9

 8.2 制定改进方案 9

 8.3 执行改进活动 9

附录 A（资料性） 数据治理组织架构 11

A.1 数据治理组织建立模式 11

A.2 数据治理组织层级划分 11

A.3 数据治理组织职责参考 11

附录 B（资料性） 数据治理制度及管理规范 13

B.1 概述 13

B.2 数据治理组织管理制度 13

B.3 数据架构相关制度及管理规范 13

B.4 元数据相关制度及管理规范 13

B.5 数据标准相关制度及管理规范 13

B.6 数据质量相关制度及管理规范 13

B.7 主数据相关制度及管理规范 13

B.8 数据应用相关制度及管理规范 14

B.9 数据安全相关制度及管理规范 14

B.10 数据生存周期相关制度及管理规范 14

B.11 数据治理考核评价管理制度 14

附录 C（资料性） 数据治理评价指标体系 15

参考文献 16

前 言

本文件按照 GB/T 1.1—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。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本文件起草单位：中国电子技术标准化研究院、上海计算机软件技术开发中心、北京亿信华辰软件有限责任公司、美林数据技术股份有限公司、方正国际软件(北京)有限公司、北京三维天地科技股份有限公司、北京华宇信息技术有限公司、普元信息技术股份有限公司、贵州图智信息技术有限公司、阿里云计算有限公司、浪潮软件科技有限公司、成都四方伟业软件股份有限公司、中国石油化工集团有限公司、上海软中智链数字科技有限公司、上海逸迅信息科技有限公司、国能信息技术有限公司、广东电网有限责任公司、安徽省铁路投资有限责任公司、宁夏交投科技发展有限公司、四川发展数字金沙科技有限公司、山东润一智能科技有限公司、北京易华录信息技术股份有限公司、广州赛宝认证中心服务有限公司、中电科大数据研究院有限公司、上海观安信息技术股份有限公司、西北工业大学、成都九洲电子信息系统股份有限公司、深圳市飞速创新技术股份有限公司、三亚海兰寰宇海洋信息科技有限公司、南京维数软件股份有限公司、成都数之联科技股份有限公司、广州芳禾数据有限公司、暨南大学、北京全路通信信号研究设计院集团有限公司、首都医科大学附属北京地坛医院、杭州数梦工场科技有限公司、云上贵州大数据产业发展有限公司、浙江省数字经济发展中心、山大地纬软件股份有限公司、中国联合网络通信有限公司广东省分公司、中国融通集团信息技术有限公司、以萨技术股份有限公司、航天网安技术(深圳)有限公司、南方电网储能股份有限公司、福建福清核电有限公司、数字重庆大数据应用发展有限公司、中国南方电网有限责任公司。

本文件主要起草人：张群、戴炳荣、李冰、高洪美、訾津津、李晓燕、张煜、王兆君、蒋楠、史玉良、应志红、刘庆会、谢江、朱松、刘陈宇、王金超、李保东、刘影、王兰、闭珊珊、王红梅、杨琳、宋俊典、余永升、张延生、李书超、康丽丽、柳富、刘俊良、刘宇峰、陈彬、杨秋勇、李军、姜向军、袁雷、刘彦华、陈瑛琪、曹亦伟、郝建明、刘利达、旷志光、吕将男、陶平、王凌、寿阳、金荣华、李冰、逢辉、黄海峰、余楷、孙丽娟、肖让、张艳宁、王鹏、刘婷、刘川杰、向伟、石桂华、胡芑、傅彦、周俊临、童瑶、张昊天、翁健、夏志华、林鹏程、郭宁、黄明峰、刘军、杨媛媛、田燕翔、黄勇、吕梁、李勇路、鄂梅、严宇平、黄德辉、李智宇、李瑞、肖政、姚巍、杨旭、张黎明、王宇静、王成亮、王伟、刘祯、杨伟伟、郭林元。

引言

目前,数据治理标准主要有:

- GB/T 34960.5—2018 信息技术服务 治理 第5部分:数据治理规范;
- GB/T 36073—2018 数据管理能力成熟度评估模型。

本文件作为数据治理标准之一,在研制过程中数据治理实施过程主要借鉴了 GB/T 34960.5—2018 中的数据治理过程,包括规划、执行、评价和改进四个过程,并进一步明确了各个过程的实施活动和内容。其中执行过程中开展的治理活动,主要借鉴了 GB/T 36073—2018 中的能力域,进一步明确了各个能力域的实施过程和内容。

数据治理实施过程对应关系见表 1。

表 1 数据治理实施过程对应关系表

GB/T 34960.5—2018	GB/T 36073—2018	本文件	
统筹和规划	数据战略	现状调研及分析	规划
		确立目标	
		明确治理内容	
		确定实施路线	
构建和运行	数据治理	建立治理组织	执行
		制定制度规范	
	数据架构	开展治理活动:数据架构设计	
		开展治理活动:元数据管理	
	数据标准	开展治理活动:数据标准管理	
	数据质量	开展治理活动:数据质量管理	
	—	开展治理活动:主数据管理	
	数据应用	开展治理活动:数据应用	
	数据安全	开展治理活动:数据安全治理	
	数据生存周期	开展治理活动:数据生存周期管理	
监控和评价	—	明确评价目标和范围	评价
	—	设计指标体系	
	—	评价治理绩效	
改进和优化	—	进行差异分析	改进
	—	制定改进方案	
	—	执行改进活动	

信息技术 大数据 数据治理实施指南

1 范围

本文件提供了大数据环境下开展数据治理实施的过程指南,包括规划、执行、评价和改进四个过程的相关活动及内容。

本文件适用于指导组织开展数据治理实施工作。

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/T 35295—2017 信息技术 大数据 术语

3 术语和定义

GB/T 35295—2017 界定的以及下列术语和定义适用于本文件。

3.1

数据治理 data governance

对数据资源管理行使权力和控制的活动集合(计划、监督和执行)。

3.2

数据管理 data management

数据资源获取、控制、价值提升等活动的集合。

[来源:GB/T 34960.5—2018,3.2]

3.3

元数据 metadata

关于数据或数据元素的数据(可能包括其数据描述),以及关于数据拥有权、存取路径、访问权和数据易变性的数据。

[来源:GB/T 34960.5—2018,3.6]

3.4

元模型 metamodel

规定一个或多个其他数据模型的数据模型。

[来源:GB/T 36073—2018,3.9]

3.5

数据生存周期 data lifecycle

数据获取、存储、整合、分析、应用、呈现、归档和销毁等各种生存形态演变的过程。

[来源:GB/T 34960.5—2018,3.7]

3.6

数据质量 data quality

在指定条件下使用时,数据的特性满足明确的和隐含的要求的程度。

[来源:GB/T 36073—2018,3.10]

3.7

数据标准 data standard

数据的命名、定义、结构和取值的规则。

[来源:GB/T 36073—2018,3.7]

3.8

主数据 master data

组织中需要跨系统、跨部门进行共享的核心业务实体数据。

[来源:GB/T 36073—2018,3.12]

3.9

参考数据 reference data

对其他数据进行分类和规范的数据。

[来源:GB/T 36073—2018,3.13]

4 数据治理实施过程

数据治理实施过程框架包括以下规划、执行、评价和改进四个过程，见图 1。

- a) 规划过程主要是确定数据治理目标、内容和路线，主要活动包括现状调研及分析、确立目标、明确治理内容、确定实施路线。
- b) 执行过程是在规划过程基础上开展数据治理工作，主要活动包括建立数据治理组织、制定制度规范、开展数据治理活动。其中治理活动包括数据架构设计、元数据管理、数据标准管理、数据质量管理、主数据管理、数据应用、数据安全、数据生存周期管理等。
- c) 评价过程是对数据治理执行情况进行评价，主要活动包括明确评价目标和范围、设计指标体系、评价治理绩效。
- d) 改进过程是对数据治理执行情况进行分析和改进，主要活动包括进行差异分析、制定改进方案、执行改进活动。

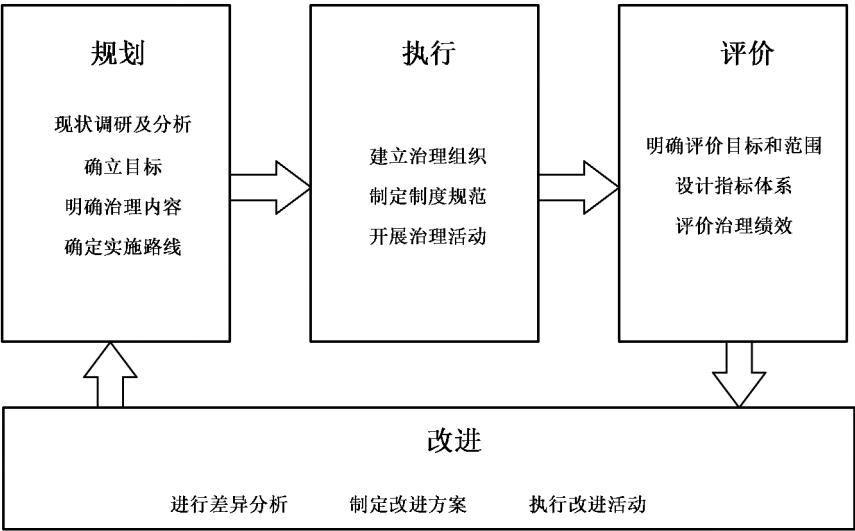


图 1 数据治理实施过程

5 规划

5.1 现状调研及分析

围绕业务战略,组织充分调研和分析当前的数据管理现状及其存在问题,为规划确定数据治理目标、数据治理内容和数据治理实施路线提供依据和指导。

a) 现状调研的形式及内容包括:

- 1) 通过向技术人员、业务人员发放调查问卷、现场访谈等方式调研组织的数据治理现状,从技术、业务、管理视角理解数据治理的问题、痛点和核心需求;
- 2) 通过中高层领导访谈方式调研管理层的数据治理需求和规划,深刻理解组织的业务战略和数据战略,掌握管理层的数据治理思路、目标、方向、期望和要求;
- 3) 通过信息系统盘点、相关资料查阅等方式调研组织的数据资源及管理现状,包括数据量、采集方式、数据分布、数据标准、数据质量、数据应用及其业务场景等,以及数据管理运营模式、组织架构、流程规范、业务现状、信息化现状、安全现状等。

b) 组织结合数据战略,根据调研情况对数据问题及治理需求进行分析,内容包括但不限于:

- 1) 分析数据战略对数据治理提出的需求,结合业务特点分析数据治理的相关法律、规定、方针和政策要求;
- 2) 分析数据治理组织及运营、制度规范制定与执行、数据治理沟通机制等方面存在的问题和治理需求;
- 3) 分析组织在数据架构设计、元数据管理、数据标准管理、数据质量管理、主数据管理、数据安全治理、数据应用、数据生存周期管理等方面存在的治理问题;
- 4) 参考国家技术标准、行业最佳实践、技术发展趋势等,分析数据治理存在的不足和改进方向;
- 5) 形成数据治理调研分析报告,包括但不限于调研方法、调研对象、调研内容、现状分析、存在问题、结论建议等。

5.2 确立目标

组织根据现状调研和分析,发现其在数据资源及管理中的问题,并与其他组织进行对标和差异分析,围绕数据治理目标细化自身的数据治理实施目标。数据治理实施目标包括但不限于实现满足外部监管要求、保障数据安全合规、提升内部工作效率、提升数据管理能力等级、增强数据共享和应用能力等。

5.3 明确治理内容

组织根据数据治理实施目标明确数据治理的内容,包括对组织、制度、管理能力等方面的治理。

5.4 确定实施路线

组织通过制定符合数据治理目标的实施路线,保障数据治理工作的有序、有效进行。制定数据治理实施路线时需考虑以下内容:

- a) 结合业务需求和组织架构实际情况,确定数据治理组织架构的建立路线,包括组织建立模式和分层模式;
- b) 参考 GB/T 36073—2018 中数据制度建设的成熟度等级要求和最佳实践,确定数据治理相关制度、流程和规范的制定路线;
- c) 根据现有的资源、环境和能力,围绕数据架构设计、元数据管理、数据标准管理、数据质量管理、

主数据管理、数据应用、数据安全管理和数据生存周期管理等,确定相应的实施活动及目标,制定分步实施策略,明确每个实施任务的优先级;

- d) 根据已有技术路线、技术发展趋势和经费保障情况,选择合适的技术路线;
- e) 形成数据治理实施路线图,包括但不限于划分阶段、每个阶段的目标、工作任务、时间节点、里程碑标志物、相关参与部门和角色等。

6 执行

6.1 概述

数据治理执行过程包括建立治理组织、制定制度规范、开展治理活动。建立治理组织旨在确定组织模式、组织架构以及人员角色分工,定义各方职责;制定制度规范旨在制定数据治理实施依据的制度、规范和标准文件;开展治理活动旨在执行数据架构设计、元数据管理、数据标准管理、数据质量管理、主数据管理、数据应用、数据安全管理和数据生存周期管理等活动。

6.2 建立治理组织

组织参考 GB/T 36073—2018 中数据治理组织的成熟度等级要求建立有效的数据治理组织,保证数据治理目标达成和数据治理活动的有效实施。建立数据治理组织时需考虑以下内容。

- a) 确定组织模式:数据治理组织模式可采用集中式、分布式和混合式等方式,参考附录 A 中 A.1 中各种模式的特点进行选择。
- b) 划分组织层级:明确组织模式后,划分组织层级。数据治理组织层级一般可分为决策层、管理层、执行层,参考 A.2 中各层级的特点进行确定。
- c) 定义组织职责:为数据治理组织的各个层级定义相应的职责,参考 A.3 进行细化和扩展。
- d) 提升人员能力:定期提供数据治理相关制度规范和理论方法的培训,以及实施经验分享,不断提高人员专业能力。

定期复盘和审视数据治理组织运营的情况,以保证数据治理组织的有效运行。需综合考虑当前组织管理模式是否与战略目标一致、是否权责分明、协同是否高效、外部法律法规是否对组织管理存在风险和影响、人员能力匹配程度等因素,根据战略目标和内外部环境的要求对数据治理组织进行相应调整。

6.3 制定制度规范

组织参考 GB/T 36073—2018 中数据制度建设的成熟度等级要求和最佳实践,制定相适应的管理规范和制度体系,确保数据治理各项活动的有效执行。制度规范范围覆盖数据治理组织管理制度、数据治理考核评价制度以及数据架构设计、元数据管理、数据标准管理、数据质量管理、主数据管理、数据应用、数据安全管理和数据生存周期管理等活动相应的制度规范。制度规范的思路内容可参考附录 B。

数据治理组织定期复盘和审视相关制度规范的执行情况,发现存在的问题和不足,有针对性地及时调整和更新完善。

6.4 开展治理活动

6.4.1 数据架构设计

组织参考 GB/T 36073—2018 中数据架构的成熟度等级要求对其进行治理。数据架构设计活动包括业务域及流程梳理、数据实体识别、数据分布确定、数据集成流向梳理、数据模型设计及应用,具体如下。

- a) 业务域及流程梳理:收集组织架构、业务架构、管理手册或内控手册等资料,梳理组织职能域,结合实际情况划分业务域,梳理相应的业务系统、业务流程划分等。
- b) 数据实体识别:基于业务流程中各业务活动的输入输出数据,识别数据实体;基于业务系统相关资料梳理物理表清单及字段清单。
- c) 数据分布确定:确定数据实体在业务域间的分布地图、在业务流程间的分布地图,确定物理表在业务系统间的分布地图。
- d) 数据集成流向梳理:梳理数据在业务域之间、业务流程之间及业务系统之间的流转关系,确定权威数据源;同时定期维护和更新组织中的数据集成流向。
- e) 数据模型设计及应用:收集和理解数据需求,制定模型规范和设计数据模型,制定数据统一编码,建立数据模型映射关系并开展应用;周期性开展模型符合性检查,并对模型变更管理。数据模型包括主题域模型、概念模型、逻辑模型、物理模型。

6.4.2 元数据管理

组织参考 GB/T 36073—2018 中元数据管理的成熟度等级要求对其进行治理。元数据管理活动包括元数据识别、元模型构建、元数据创建整合、元数据变更管理和元数据应用,具体如下。

- a) 元数据识别:识别并确定元数据及管理范围,包括技术元数据、业务元数据、管理元数据等。
- b) 元模型构建:对要管理的元数据进行结构化、模型化,抽象出稳定的元模型结构,理清元模型结构及元模型间关系。
- c) 元数据创建整合:建立审核流程以验证元数据创建、维护的结果;根据元数据的特点进行多种方式采集和存储,并对不同来源、不同类型的元数据进行整合。
- d) 元数据变更管理:建立元数据变更流程,以提升元数据的准确性和时效性,保障元数据的质量。元数据变更流程一般包括元数据变更申请、元数据变更审核、元数据维护和元数据发布四个阶段。
- e) 元数据应用:依托采集整合的元数据,通过去重、冲突处理、关系还原等优化操作,开展链路分析、差异分析、热度分析、数据地图展示、数据资产发布等应用建设,以发挥元数据价值,具体如下:
 - 1) 链路分析,构建以数据流向为主线的上下游链路,在此链路上进行血缘追溯、变更影响评估、全链关系分析;
 - 2) 差异分析,比对两个同类型的元数据,分析相似元数据间的细微差异;
 - 3) 热度分析,分析元数据间的关联情况,根据关系数值,评估元数据热度与重要性;
 - 4) 数据地图展示,将元数据整合成一致的脉络图,按实体关系、系统关系、数据库关系等不同维度展示;
 - 5) 数据资产发布,围绕元数据进行数据资产的编目、使用、更新等全面管理。

6.4.3 数据标准管理

组织参考 GB/T 36073—2018 中数据标准的成熟度等级要求对其进行治理。数据标准管理活动包括数据标准分类设计、数据标准制定、数据标准发布、数据标准应用和数据标准维护。

- a) 数据标准分类设计:根据外部监管要求、行业通用标准、专家实践经验和业务实际情况确定数据标准的分类和范围。数据标准一般包括业务术语标准、参考数据和主数据标准、数据元标准和指标数据标准。根据业务主题域进一步细分数据标准,原则上尽可能覆盖组织的主要业务活动和生产系统中产生的所有业务数据。数据标准从业务域、业务活动、对象实体、实体关系等方面层层递进、逐步展开。
- b) 数据标准制定:通过资料收集、调研访谈、分析评估等方法梳理业务指标、数据项、代码等数据

要素。依据数据标准模板,定义各类数据的业务属性、技术属性、管理属性,如中英文名称、业务定义、业务规则、统计维度、制定依据、数据类型、数据长度、数据精度、标准制定人员、标准归属部门、标准使用部门等。

- c) 数据标准发布:数据标准制定后面向各业务部门征求意见,并由评审小组进行审查,判断数据标准是否符合业务需求、是否具有业务前瞻性、是否有歧义、是否具有权威性、是否具有可落地性等。审查通过后方可发布和使用。
- d) 数据标准应用:针对发布的数据标准面向相关业务人员和技术人员组织培训和宣讲,促进数据标准在各部门间达成一致和理解。针对新建系统,直接采用已发布的相关数据标准;针对已建系统,通过分析数据标准与数据现状的差异,根据差异结果,从业务流程、系统多个方面分析数据标准落标后的影响,制定落标具体执行方案,如落标形式、操作流程、数据清洗规则等。
- e) 数据标准维护:基于内外部需求变化,对数据标准及时更新和管理。一般包括数据标准的变更管理和版本管理。变更管理包括数据标准的新增、修改和废止等;版本管理包括版本号管理、标准变更记录等。

6.4.4 数据质量管理

组织参考 GB/T 36073—2018 中数据质量的成熟度等级要求对其进行治理。数据质量管理活动包括数据质量需求确立、数据质量检查、数据质量分析和数据质量提升。

- a) 数据质量需求确立:根据业务操作需求、业务管理需求、数据管理目标和行业监管要求,明确组织数据质量管理目标和范围,梳理各类数据质量需求和责任部门,确定数据质量评价维度,设计各类数据的质量评价指标、校验规则与方法,并持续更新。
- b) 数据质量检查:根据数据质量需求制定数据质量检查计划,并根据数据质量规则对数据质量情况进行校验和剖析,发现数据质量问题,生成数据质量检查报告,并反馈给数据管理人员。
- c) 数据质量分析:针对发现的数据质量问题及相关信息,确定数据质量分析方法和要求,判断数据质量问题产生的原因及影响的范围,确定数据质量问题的优先级,生成数据质量分析报告。
- d) 数据质量提升:根据数据质量分析结果,结合数据质量目标,制定并实施数据质量提升方案,跟踪数据质量改进过程,优化数据质量相关制度、流程、标准等,确保数据质量改进成果得到有效保持。

6.4.5 主数据管理

组织参考 GB/T 36073—2018 中主数据标准的成熟度等级要求对其进行治理。主数据管理活动包括主数据识别、主数据标准制定、主数据采集、主数据清洗、主数据集成、主数据应用管理,具体如下。

- a) 主数据识别:界定主数据需求与范围,依据主数据的定义及特征,对主数据进行识别,包括识别组织级主数据类型及属性、识别不同主数据之间的关系、识别主数据与业务系统之间的关系、识别主数据与业务管理之间的关系等。
- b) 主数据标准制定:需在充分考虑当前信息系统的现状及未来信息系统发展规划的基础上,根据业务需求的紧急程度分期建设主数据标准。
- c) 主数据采集:根据制定的主数据抽取策略和采集方案,实现主数据的抽取、处理及任务调度设置。在实时数据采集或批量数据采集时满足各业务部门、各信息系统对主数据的及时性需求。
- d) 主数据清洗:按照主数据标准对现有原始数据进行清洗,保证主数据的唯一性、准确性、完整性、一致性和有效性,然后通过系统校验、查重及人工比对、筛查、核实等多种手段对主数据代码的质量进行检查,建立高质量的主数据代码库。
- e) 主数据集成:通过服务总线、数据接口、数据库等方式与各个目标信息系统进行主数据集成,通过相应的工具平台或共享数据库实现主数据在多个系统之间的调用、分发,满足业务系统数据

共享需求。

- f) 主数据应用管理:包括明确管理要求、实施有效的管理、强化保障服务,以及转化和切换存量系统主数据代码等内容。

6.4.6 数据应用

组织参考 GB/T 36073—2018 中数据应用的成熟度等级要求对其进行治理。数据应用活动类型包括数据分析、数据开放共享、数据服务等。

- a) 数据分析实施活动包括数据准备、工具准备、提供分析、流程编排等,具体内容如下:
 - 1) 数据准备:对采集的原始数据进行预处理,保证数据能被后续分析直接使用;
 - 2) 工具准备:提供建立数据模型和应用模型的算法库或者工具库;
 - 3) 提供分析:通过离线数据分析、流数据分析、交互式联机分析、预测分析、描述分析等数据分析方法或者中间件,将数据准备模块输出的数据以及数据建模过程中产生的中间数据转变成知识或者决策;
 - 4) 流程编排:对数据分析的各环节进行流程编排操作。
- b) 数据开放共享实施活动包括需求分析、设计资源目录、数据资源准备以及开放共享数据等,具体如下:
 - 1) 需求分析:根据组织内各部门、外部部门及社会其他组织的需求,明确数据开放共享需求和范围;
 - 2) 设计资源目录:根据数据开放共享需求,设计和编制资源目录。资源目录的内容包含数据来源、数据格式、业务层级、业务类别、共享方式、共享条件、开放方式、开放条件等;
 - 3) 数据资源准备:将原始数据通过抽取、清洗、转换、比对、加载等处理,变为有价值的、可进行开放共享的数据资源;
 - 4) 开放共享数据:建设数据服务封装能力,通过采用数据文件、API 接口、数据报表、数据订阅等多种方式为数据使用方提供数据。
- c) 数据服务实施活动包括需求分析、服务规划、过程管控和能力提升等,具体如下:
 - 1) 需求分析:分析组织的数据战略和发展方向及内外部的数据需求、现有资源和环境,明确数据服务的内涵、范围、类型、用户和方式;
 - 2) 服务规划:制定数据服务目录、服务级别、访问权限和实施方式,明确数据服务的内容和能力;
 - 3) 过程管控:建立数据服务管控流程,监督数据服务的安全性、合规性,并对实施过程进行审核和控制;建立数据服务支持流程,通过标准化、自动化等方式支撑数据服务的交付,满足服务需求;建立数据服务决策机制,支持用户通过多种方式输出结果,辅助业务决策;建立数据服务管理机制,对数据服务的过程、质量和安全等进行管理、持续改进和优化;
 - 4) 能力提升:建立数据服务的分析机制,从用户分布、服务响应、实施效果、运维等角度进行分析,定期对数据服务的能力和价值进行评估、改进和优化,促进服务创新。

6.4.7 数据安全治理

组织参考 GB/T 36073—2018 中数据安全的成熟度等级要求对其进行治理。数据安全治理活动包括数据安全风险评估、数据安全策略制定、数据安全运营、数据安全审计。

- a) 数据安全风险评估:包括数据处理活动识别、数据安全风险源识别、数据安全风险分析等,具体内容如下:
 - 1) 数据处理活动识别:包括业务系统和 IT 资产识别、业务流识别、数据处理相关场景分析、数据分类分级等;

- 2) 数据安全风险源识别:参考 GB/T 37973—2019 开展数据资产识别、数据安全脆弱性识别、数据安全威胁识别、现有数据安全措施识别;
- 3) 数据安全风险分析:基于数据安全风险源识别对数据安全风险进行分析和赋值。
- b) 数据安全策略制定:依据国家、行业等监管需求,结合自身的数据安全业务需要,制定适合的数据安全管理目标、原则和策略。
- c) 数据安全运营:围绕数据应用过程中识别的风险,开展数据合规及数据安全风险与事件监测、数据安全事件分析、数据安全事件响应和应急处置等,并进行数据安全平台运营、数据安全基线检查等。
- d) 数据安全审计:包括过程审计、规范审计、合规审计、供应商审计、审计报告发布、数据安全建议等。

6.4.8 数据生存周期管理

组织参考 GB/T 36073—2018 中数据生存周期的成熟度等级要求对其进行治理。数据生存周期管理活动包括数据采集、数据存储、数据加工、数据利用、数据退役,具体如下。

- a) 数据采集:明确数据来源、数据体量、数据变化频率、数据结构等信息,对多源异构数据制定数据采集方案,包括数据采集技术手段、数据采集频率等;采集过程涉及数据预处理等;遵循元数据、数据建模等管理规范。
- b) 数据存储:对不同类型的数据进行存储,包括结构化与非结构化数据、实时与批量数据等,明确数据存储策略、安全策略等。
- c) 数据加工:根据不同的使用需求及数据管理规范,进行数据建模、数据整合、数据清洗、数据安全防范等,以满足数据应用需要。
- d) 数据利用:依据相关制度规范保障数据传输、提供、公开、分析、呈现等过程中的安全性、稳定性、时效性和准确性,防止数据泄露风险,保证数据应用效果。
- e) 数据退役:对没有使用价值的数据,依据法律法规、业务、技术相关需求制定数据归档、迁移、销毁规则,执行数据退役相关工作。同时制定数据恢复及检查机制,在数据需要时可恢复。

7 评价

7.1 明确评价目标和范围

在数据治理评价工作启动前,组织根据其愿景、战略目标、业务发展需要、法律法规和行业规范等制定整个评价的目标和范围。

7.2 设计指标体系

组织依据数据治理评价目标和范围,结合实际情况,参照规划和执行过程的相关活动,定义适用于自身的评价要素。基于定义的评价要素,组织从数据治理的执行情况和取得成效等角度,需综合考虑定量和定性两个维度的评价指标,设计数据治理评价指标体系。数据治理的评价指标体系可参考附录 C。组织可根据实际情况,选择和设计适用于自身的数据治理评价指标体系。

7.3 评价治理绩效

组织基于设计的数据治理评价指标体系对数据治理工作的绩效进行整体评价。数据治理绩效评价需考虑以下内容:

- a) 定义数据治理绩效评价等级;
- b) 选择合适的评价方法,如专家打分法、层次分析法或混合法等;

- c) 对每个评价指标和评价要素分配合适的权重；
- d) 基于评价方法计算组织的数据治理绩效评价得分；
- e) 形成数据治理评价报告,为差异分析提供依据。

8 改进

8.1 进行差异分析

组织采用差异分析方式对制定的数据治理目标与实际执行取得的治理结果进行比较,分析两者之间是否存在差距。若存在差距,则组织根据 7.3 的评估结果逐项开展差异原因分析,找到影响数据治理效果的原因,为数据治理改进方案提供依据。组织开展差异分析时,需考虑以下内容:

- a) 差异分析要点贯穿数据治理执行过程的各项活动,确保数据治理执行过程得到全面改进;
- b) 差异分析在数据治理评价之后开展,且差异分析要点覆盖 7.2 评价要素的全部内容;
- c) 在数据治理执行的各阶段对阶段性的执行情况和取得成效进行记录,把记录作为阶段性开展数据治理活动的基线信息,并与该阶段的数据治理期望目标进行差距分析;
- d) 形成数据治理差异分析报告,包括但不限于差异分析要点、所属治理执行环节、预期目标、实际执行记录、执行成效、差异问题、差异原因、问题干系人、改进优先级、记录时间等。

8.2 制定改进方案

组织依据数据治理差异分析报告,制定合理的数据治理改进方案,指导数据治理改进工作有序进行。组织制定数据治理改进方案时,需考虑以下内容。

- a) 数据战略规划改进要点需考虑法律法规、组织愿景和使命、组织业务发展规划等的变更调整,具体情形包括但不限于:
 - 1) 分析与数据战略规划相关的法律法规的变更范围,调整数据战略使之与法律法规相符合;
 - 2) 根据组织愿景和使命的变更,制定数据战略目标调整和改进规划;
 - 3) 分析变更的业务内容,梳理并调整与之对应的数据战略规划范畴,保证业务规划与数据规划的协调性。
- b) 数据制度规范的改进要点需考虑数据策略、数据标准、技术规范和管理流程等的重新设定或优化,具体情形包括但不限于:
 - 1) 评估发现当前的数据策略与数据战略规划不相匹配,无法有效支撑数据战略实现;
 - 2) 数据标准、技术规范无法满足数据治理、数据安全等方面的要求;
 - 3) 数据治理执行过程中对数据标准、技术规范产生了新的规范性需求;
 - 4) 管理流程的效率无法满足组织业务发展需求。
- c) 数据治理执行过程的改进要点需考虑组织架构、人员、技术应用、沟通和推广等方面因素,从规划层明确组织可以对执行过程提供的组织保障、资源保障、技术保障、有效沟通和治理成效推广方案等。具体情形在实际执行过程中具有多样性、不确定性,无法穷举,本标准仅给出考虑要点供参考。
- d) 数据治理改进方案包括但不限于改进目标、改进范围、改进内容、改进策略或优先级、改进责任人、改进周期及计划安排等。

8.3 执行改进活动

组织依据数据治理改进方案执行数据治理相关改进活动,且不断持续优化和改进。组织执行数据治理改进活动时,需考虑以下内容。

- a) 确认数据治理改进内容,包括需求、技术、工具、知识、流程、人员职能、团队建设等方面,具体

如下：

- 1) 完善数据需求、元数据管理、数据标准管理、数据质量管理、数据安全治理、数据应用等数据治理需求；
 - 2) 提升数据管理、数据校验、数据整合应用、数据质量监控、数据分析等方面的数据治理技术；
 - 3) 根据数据治理活动需求，分析当前市场上现有的数据治理工具，选择符合组织数据治理需求的工具，并应用到数据治理活动的相关环节，以提升数据治理能力和效率；
 - 4) 总结数据治理执行过程中的经验，针对数据治理知识、数据治理实例、数据治理专家人才形成相应的数据治理知识库、数据治理案例库及数据治理专家库等，并持续丰富数据治理知识，为持续提升数据治理能力提供知识积累；
 - 5) 持续规范数据需求管理、元数据管理、数据采集管理、数据标准管理、数据质量管理等的流程，并进一步完善业务流程；
 - 6) 获得管理层对数据治理执行工作的认同、支持，依据数据治理组织和人员的职能履行情况，调整业务部门和技术部门的数据管理责任，完善数据管理组织职责和相关规章制度等；
 - 7) 基于数据治理执行过程中存在的人员能力不足、业务理解不到位、人员治理技术欠缺等情形，通过完善人才培养体系，培养团队人员的专项能力，建设业务、技术、管理等各方面均有相关人才支撑的复合型数据治理团队。
- b) 根据数据治理改进内容，制定改进执行计划，明确改进的任务、时间及实施人员等，按照项目管理制度下达并完成各项改进任务，达到设定的目标。
 - c) 根据数据治理改进任务的执行情况，验证分析各阶段的问题定位是否准确、改进任务是否取得成效等，从而验证数据治理改进方案的有效性，并总结数据治理改进执行过程中的问题、方法、经验等，持续完善数据治理改进改进方案。
 - d) 对数据战略目标、数据治理相关制度及规范、数据治理策略、数据治理方法、数据治理流程、数据治理人才建设、数据治理执行环节的各项保障措施等不断矫正，逐步构建完善的数据治理体系，形成数据治理闭环，保证数据治理工作的持续性、有效性。
 - e) 形成数据治理改进报告，包括但不限于执行环节、改进要点、存在问题、改进优先级、改进依据、改进方式、改进成效、改进负责人、记录时间等。

附 录 A
(资料性)
数据治理组织架构

A.1 数据治理组织建立模式

- A.1.1 集中式治理组织。数据的决策和执行统一管理,参与人员负责各类数据管理工作,并向所在主管领导汇报。集中式的优点是具备专职的部门管理,驱动力强,决策容易,能保证数据治理工作的持续性、一致性。其缺点是对组织架构影响最大,需要投入较多的人员。
- A.1.2 分布式治理组织。分散的数据管理部门,每个部门独立负责本范围的决策和实施权力。数据管理组织结构与现有的业务结构完全相同。分布式的优点是对当前组织架构无较大影响,起点较低,在业务领域中明确数据管理职责及人员即可。其缺点是数据治理力度较弱,缺乏组织级视角。
- A.1.3 混合式治理组织。决策工作由集中的数据管理部门负责,具体的执行工作由下属的各部门完成。混合式的优点是有专职的部门管理,可以达成组织层面的整合,以较小的投入保证数据治理工作开展。其缺点是数据管理部门影响力有限,如何监督及制定绩效直接影响数据治理的成效。

A.2 数据治理组织层级划分

- A.2.1 决策层是数据治理工作的决策机构,一般为数据治理委员会。其目的是确立数据治理的目标、资源,保证涉及部门的负责人涵盖其中,从而快速搭建数据治理体系,在组织内形成共识并朝着一致的目标推进。决策层一般由组织的中高层业务领导、首席数据官、首席信息官参与。
- A.2.2 管理层是数据治理工作的管理机构,一般为数据治理管理部门或办公室。其目的是负责数据治理相关的日常管理工作,向数据治理决策层负责。管理层可以由一个专门设置的独立部门担任,也可由相关业务部门和信息技术部门共同承担。
- A.2.3 执行层是数据治理日常工作的执行机构,其目的是推动执行数据治理要求,参与、配合数据治理日常管理工作,向数据治理管理层负责。执行层由业务部门和信息技术部门共同组成。业务部门在数据治理管理层的指导和协调下,负责推动各自业务条线的数据治理相关内容,并定期进行治理成效评估;信息技术部门与业务部门合作负责发现数据质量问题,确定数据问题的原因,并将问题归类汇总后发送相关执行部门进行数据治理。

A.3 数据治理组织职责参考

- A.3.1 决策层的主要职责包括但不限于:
- 负责数据战略制定、重大事项的协调和决策;
 - 从战略角度进行统筹和规划,确定数据治理目标、范围、原则;
 - 明确数据治理的组织、角色和职责;
 - 负责数据治理人才建设及能力提升;
 - 负责审查数据治理工作规划、阶段性目标;
 - 审批数据管理的相关制度、标准及流程;
 - 负责确定数据治理的工具、技术和平台;
 - 负责确定数据治理的评估指标、方法;
 - 听取并评价数据治理工作情况汇报;
 - 接受风控、审计部门监督,保证数据治理工作合规。
- A.3.2 管理层的主要职责包括但不限于:

- 负责数据治理的牵头,组织、指导、推广和协调数据治理工作;
- 综合数据治理管控办法、数据治理考核机制等有关规章制度的牵头制定、修改等;
- 负责数据的分析整理并出具数据指标报告;
- 负责数据的监测预测工作;
- 建立数据冲突的处理流程和数据变更控制流程;
- 负责对基础数据质量的检测、发布、考核和清理完善工作;
- 负责管理、评估各部门和信息系统的数​​据治理工作情况。

A.3.3 执行层的主要职责包括但不限于:

- 贯彻执行数据治理相关制度、流程;
- 为开发人员定义数据规格及标准;
- 定义引证/参考数据;
- 有效的追踪数据质量问题;
- 实施正确的数据质量规则;
- 不间断地监控数据质量水平和问题;
- 保证数据的质量、审计、安全的合规性。

附 录 B
(资料性)
数据治理制度及管理规范

B.1 概述

数据治理实施遵循相关的制度和管理规范,以确保数据治理组织建立、数据架构设计、元数据管理、数据标准管理、数据质量管理、主数据管理、数据应用、数据安全治理、数据生存周期管理等相关活动得到有效执行。数据治理相关的制度及管理规范涵盖数据治理组织管理制度、各治理活动相关制度和管理规范以及数据治理成效评价与考核相关制度等。

B.2 数据治理组织管理制度

数据治理组织管理制度定义数据治理组织架构,明确组织职责、角色分工及岗位要求,指导、执行和监督数据治理工作。

B.3 数据架构相关制度及管理规范

B.3.1 数据分布设计规范。根据实际情况,确定数据在系统、组织和流程等方面的分布关系,定义数据类型,明确权威数据源,明确数据发布及权威数据源确定的流程及管理要求等。

B.3.2 数据集成共享规范。明确数据集成共享的原则、方式和方法等。

B.3.3 数据模型设计规范。对主题域模型、概念模型、逻辑模型、物理模型的设计规范进行说明,包括命名规范、模型的更新/停用流程、模型的归口管理组织及责任人等。

B.4 元数据相关制度及管理规范

B.4.1 元数据定义规范。对元数据含义、元数据分类定义,明确各类元数据(技术元数据、业务元数据、操作元数据)定义规范,为元数据维护及分析应用提供依据。

B.4.2 元数据管理规范。明确元数据管理目标及范畴,明确元数据的责任组织及归口人,定义元模型管理规范,定义各类元数据的采集、存储、变更、分析等管理流程规范,保障元数据的应用。

B.5 数据标准相关制度及管理规范

B.5.1 数据标准规范。定义业务术语、主数据、参考数据、指标数据等各类数据的标准、分类,以及业务含义、业务规则、取值范围等属性。

B.5.2 数据标准管理办法。明确数据标准的责任组织及归口人,明确数据标准的发布、审批、执行、变更等流程、工具规范,为数据标准维护及应用提供依据。

B.6 数据质量相关制度及管理规范

B.6.1 数据质量规则管理规范。定义数据质量的一致性、完整性、准确性、及时性、规范性等相关规则。

B.6.2 数据质量管理规范。定义数据质量评价的指标体系,明确质量评价机制以及质量问题分析、跟踪和改进等流程规范。

B.7 主数据相关制度及管理规范

B.7.1 主数据标准规范。明确每类主数据的分类规则、编码规则、属性定义、值列表等,为主数据维护及应用提供依据。

B.7.2 主数据管理办法。明确主数据管理的组织与职责,明确主数据的识别、创建、变更、停用、清洗、分发、应用、考核等流程、工具规范,为主数据维护及应用提供依据。

B.7.3 主数据应用管理制度。明确主数据的应用范围、应用规则、管理要求和考核标准等,有效实现主数据应用管理。

B.8 数据应用相关制度及管理规范

B.8.1 数据开放共享管理规范。明确数据开放共享的类型、条件、范围、规则、流程、工具等,指导并规范数据开放共享过程。

B.8.2 数据服务管理办法。定义数据服务目录、类型、标准规范,明确数据服务的发布、上架、第三方调用与审批流程以及状态标识等。

B.8.3 数据分析管理规范。根据实际业务需求,定义数据仓库各层的命名规范、基础数据规范、事实表规范、维表规范,明确数据分析的流程、方法及工具。

B.9 数据安全相关制度及管理规范

B.9.1 数据安全管理制度。明确数据所有者权利、利益相关者的责任与义务,定义数据分类管理、分级保护、授权使用相关原则。

B.9.2 数据安全规范。按照数据的分级规范和相应的安全保护标准,明确权限管理、数据加密等安全技术规范、操作流程、操作规范,设立安全风险评估机制和应急响应机制,并明确安全体系的动态维护机制。

B.10 数据生存周期相关制度及管理规范

数据生存周期管理办法明确数据的采集、存储、加工、使用、归档和销毁等的管理方法、流程及工具,保障数据在生存周期中被一致、有效地管理。

B.11 数据治理考核评价管理制度

数据治理评价与考核管理办法对组织数据治理过程中的数据架构设计、元数据管理、数据标准管理、数据质量管理、主数据管理、数据应用、数据安全、数据生存周期管理等活动定义评价指标体系、考核机制等,持续提升组织的数据治理能力。

附 录 C
(资料性)
数据治理评价指标体系

组织在设计数据治理评价指标体系时宜参考表 C.1。

表 C.1 数据治理评价指标体系

一级评价要素	二级评价要素	评价指标
战略规划	战略目标	数据治理的战略级统筹规划,战略目标清晰等
	路线规划	数据治理实施策略,技术路线选择,数据治理实施路线图等
治理组织	组织架构	组织架构情况,分工职责的界定,运行情况等
	沟通机制	沟通机制的设置情况,执行情况等
制度规范	制度流程	数据治理相关制度流程的制定情况,覆盖范围,执行情况等
	标准规范	数据治理相关标准规范的制定情况,执行情况等
治理活动	数据架构设计	数据分布情况,数据模型建设情况,数据架构平台建设情况等
	元数据管理	元模型构建情况,元数据采集情况,元数据分析情况,元数据变更维护情况,元数据应用情况,元数据管理工具建设和使用情况等
	数据标准管理	数据标准制定情况,数据标准覆盖情况,数据标准应用情况,数据标准维护情况,数据标准管理工具建设和使用情况等
	数据质量管理	建立数据质量评价指标,数据质量检核和分析情况,数据质量预防和改进措施,数据质量管理工具建设和使用情况等
	主数据管理	主数据识别,主数据标准制定,主数据采集,主数据清洗,主数据集成、主数据应用管理情况,主数据管理工具建设和使用情况等
	数据应用	数据分析管控能力,数据开放共享目录,数据开放共享机制,数据开放共享情况,数据服务管控能力,数据开放共享平台建设情况等
	数据安全 管理	数据风险识别和分析,数据安全应对策略,数据安全管控措施,数据安全审计,数据安全技术工具建设和使用情况等
	数据生存周期管理	数据的采集、传输、存储、使用、开放共享、销毁等生存周期活动的管理策略与实施方法,相关技术工具建设和使用情况等

参 考 文 献

- [1] GB/T 34960.5—2018 信息技术服务 治理 第5部分:数据治理规范
 - [2] GB/T 35274—2017 信息安全技术 大数据服务安全能力要求
 - [3] GB/T 36073—2018 数据管理能力成熟度评估模型
 - [4] GB/T 37973—2019 信息安全技术 大数据安全管理指南
 - [5] GB/T 38667—2020 信息技术 大数据 数据分类指南
-