ICS 35.240

|  |
| --- |
| CCS L 70 |

DB21

辽宁省地方标准

DB21/TXXXX—XXXX

|  |
| --- |
|  |

工业数据流通 制造企业大数据价值管理规范

Industrial data circulation manufacturing enterprise big data value management standard

（征求意见稿）

|  |
| --- |
| 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

辽宁省市场监督管理局　发布

目 次

前言 III

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 缩略语 2

5 数据价值管理 2

6 数据价值发现要求 3

6.1 数据类型 3

6.2 数据质量评价 3

6.3 数据价值评估 4

7 数据价值交换要求 5

7.1 数据表达与索引 5

7.2 数据统一可靠存储 6

7.3 数据可信认证 7

8 数据价值服务要求 7

8.1 数据共享 7

8.2 数据分层管理 8

8.3 数据安全管控与隐私保护 8

9 数据资产化管理要求 9

9.1 元数据管理 9

9.2 数据标准管理 10

9.3 数据模型管理 10

9.4 数据治理 10

9.5 数据生存周期管理 11

10 制造大数据价值化应用要求 12

10.1 数据应用需求管理 12

10.2 数据应用项目实施 12

参考文献 13

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。

本文件由辽宁省工业和信息化厅提出并归口。

本文件起草单位：东北大学、四川长虹电器股份有限公司、四川长虹电子控股集团有限公司、四川启睿克科技有限公司、四川大学、上海交通大学、北京理工大学。

本文件主要起草人：杨晓春、毕可骏、唐潮、唐军、杜科、蒋玉明、高晓沨、张志威、钟全德。

本文件发布实施后，任何单位和个人如有问题和意见建议，均可以通过来电和来函等方式进行反馈，我们将及时答复并认真处理，根据实际情况依法进行评估及复审。

归口管理部门通信地址：沈阳市辽宁省沈阳市皇姑区北陵大街45-2号。

归口管理部门联系电话：024-86913384。

标准起草单位通信地址：辽宁省沈阳市和平区文化路3巷11号。

标准起草单位联系电话：024-83687392。

工业数据流通 制造企业大数据价值管理规范

1. 范围

本文件规定了制造企业大数据价值管理规范以及相应的管理要求，包括数据价值发现、数据价值交换、数据价值服务、数据价值化应用等领域及相关子模块。

本文件适用于各组织和机构从数据价值角度开展制造企业大数据管理及应用。

1. 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 35273—2020 信息安全技术 个人信息安全规范

GB/T 35295—2017 信息技术 大数据 术语

GB/T 36344—2018 信息技术 数据质量评价指标

GB/T 37973—2019 信息安全技术 大数据安全管理指南

1. 术语和定义

GB/T 35295界定的以及下列术语和定义适用于本文件。

1.

数据治理 data governance

对数据进行处置、格式化和规范化等相关操作。

1.

数据标准 data standard

数据的命名、定义、结构和取值的规则。

1.

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 36344-2018，2.3]

1. 缩略语

下列缩略语适用于本文件。

ADS：数据应用层（Application Date Service）

DDL：数据定义语言（Data Definition Language）

DAM：数据资产管理系统(Data Asset Management)

DWD：细节数据层（Data Warehouse Detail）

DWS：数据服务层（Data Warehouse Service）

ERP：企业资源计划（Enterprise Resource Planning）

ETL：抽取、转换和加载（Extract-Transform-Load）

FTP：文件传输协议（File Transfer Protocol）

IT：信息技术（Information Technology）

KPI：关键绩效指标（Key Performance Indicator）

MDM：主数据管理系统（Master Data Management）

ODS：操作数据存储（Operational Data Store）

OLAP：联机分析处理（On-Line Analytical Processing）

TMP：临时（Temporary）

SQL：结构化查询语言(Structured Query Language)

3NF：第三范式(Third Normal Form)

1. 大数据价值管理

大数据价值管理包括数据价值发现、数据价值交换、数据价值服务、数据资产化管理和数据价值化应用。大数据价值管理模块及子模块见表1。

表1 大数据价值管理模块及子模块

|  |  |
| --- | --- |
| 模块 | 子模块 |
| 数据价值发现 | 数据类型 |
| 数据质量评价 |
| 数据价值评估 |
| 数据价值交换 | 数据表达与索引 |
| 数据统一可靠存储 |
| 数据可信认证 |
| 数据价值服务 | 数据共享 |
| 数据分层管理 |
| 数据安全管控与隐私保护 |
| 数据资产化管理 | 元数据管理 |
| 数据标准管理 |
| 数据模型管理 |
| 数据治理 |
| 数据生存周期管理 |
| 数据价值化应用 | 数据应用需求管理 |
| 数据应用项目实施 |

1. 数据价值发现要求
	1. 数据类型

因制造大数据存在基础性、重要性、可靠性和应用范围等方面的差异，不同类型数据其生产、经营、发布和提供服务的方式不同，数据的来源渠道、采集方式、保密程度也不同。数据类型如下：

1. 用户数据：组织在开展业务服务过程中从个人用户或组织用户收集的数据，以及在业务服务过程中产生的归属于用户的数据。如个人用户信息（即个人信息）、组织用户信息（如组织基本信息、组织账号信息、组织信用信息等）；
2. 业务数据：组织在业务生产过程中收集和产生的非用户类数据。参考业务所属的行业数据分级分类，结合自身业务特点进行细分，如产品数据、合同协议等；
3. 经营管理数据：组织在机构经营管理过程中收集和产生的数据。如经营战略、财务数据、并购及融资信息等；
4. 系统运行和安全数据：网络和信息系统运维及网络安全数据。如网络和信息系统的配置数据、网络安全监测数据、备份数据、日志数据、安全漏洞信息等。
	1. 数据质量评价

数据质量评价的框架、过程及负责方如下：

1. 数据质量评价指标框架由数据规范性、完整性、准确性、一致性、时效性、可访问性6个评价指标构成。数据质量评价指标说明应符合GB/T 36344--2018中的规定。
2. 数据质量评价完整过程参见GB/T 36344--2018附录A。
3. 制造大数据数据质量管理组织负责对制造大数据生存周期各阶段开展数据质量评价。
	1. 数据价值评估

数据价值评估的主要维度、过程及负责方如下：

1. 制造大数据主要以数据所提供的服务实现业务需求、改善系统性能、提高竞争力的能力或程度来衡量数据的价值。数据价值评估主要维度包括：
2. 数据成本：制造大数据产生、采集、存储及数据处理所耗费的全部成本；
3. 数据质量：围绕6.1中的数据质量评价指标开展数据质量评价的结果；
4. 数据市场价值：数据在交易流通过程中实际的市场交易价格或产生支付意愿；
5. 数据需求：业务活动对数据的需求程度，包括数据使用率、数据业务价值、数据可溯性等维度；
6. 数据所有权：数据所有权所属主体对数据价值产生影响，由于数据应用范围和分层属性的差异，存在不同的价值标准。
7. 数据价值评估过程：数据价值评估完整过程包括建立数据价值评估规范、确定评价指标、实施评价、数据价值变现、数据价值反馈、评价指标闭环优化；
8. 制造大数据数据价值评估组织负责对不同主体不同类型的制造大数据开展数据价值评估。
9. 数据价值交换要求
	1. 数据表达与索引

数据表达及访问的整体过程由以下环节组成：

1. 数据建模：相关技术部门根据业务的变化，需要新增或修改数据模型来支撑业务的发展，新增或修改的数据模型需要做到表、字段命名的标准化，模型的版本管控和模型变更管理等，以方便后续在多次数据的ETL操作及数据分析过程中能有效使用数据；
2. 基于统一数据中台进行数据模型管理，主要包括：
3. 模型类目管理：支持有权限的用户创建、修改和删除数据模型目录；支持用户浏览数据模型目录；
4. 模型查询：支持用户通过精确搜索或模糊搜索查询数据模型，并展示数据模型的具体信息，包括概念模型、逻辑模型、物理模型及映射关系；
5. 概念模型管理：支持有权限的用户创建、修改和删除概念模型，包括业务实体、业务属性等；
6. 逻辑模型管理：支持有权限的用户创建、修改和删除逻辑模型；支持逻辑模型和概念模型的映射关系展示；
7. 物理模型管理：支持有权限的用户创建、修改和删除物理模型；支持物理模型、逻辑模型和概念模型的映射关系展示；
8. 数据模型标准管理：支持数据模型实体及属性的标准创建、修改、删除和查询，包括自定义、自动和手动创建；支持数据模型标准与数据标准的映射；
9. 数据模型版本管理：支特记录数据模型变更的记录；支持数据模型版本的查询，以查看数据模型的历史信息。
10. 定义数据生存周期模型：根据设计阶段的数据模型，结合业务、管理、财务、审计、合规、监管、法律等对于数据访问时效要求，定义各类数据的生存周期模型；
11. 定义数据策略：根据数据模型生存周期的定义，结合数据之间的关联关系，定义各业务数据的保存、存储、访问和迁移策略；
12. 在表结构设计完成后，索引设计是数据访问性能的关键要素。索引的建立应对每个索引的必要性经过仔细分析，具备充分的依据；
13. 设计索引应满足的条件：每个表都应要有主键，主键索引在建立的时候一定要明确的指定名称，不应让系统默认建立主键索引；数据量超过300的应有索引，索引应该建在选择性高的小字段上；索引表空间应与数据表空间分离。
	1. 数据统一可靠存储
		1. 数据库设计范式

数据库设计应使用关系数据库理论，达到较高的范式，避免数据冗余。如果在数据量上与性能上无特别要求，考虑到实现的方便性宜有适当的数据冗余，但最低要达到3NF。

* + 1. 字段存储

一个字段中不宜存储多个标志，如11001表示5个标志的一种取值。

* + 1. 命名规则

命名应使用英文单词，不应使用汉语拼音和拼音简写；应选择与对象本身意义相对或相近且最简单或最通用的英文单词，不应使用毫不相干的单词来命名，不应使用SQL语言及对应数据库的特有关键字。

* + 1. 用户与权限

数据库用户应根据需要分配能覆盖所需的最小权限（系统权限、角色权限及对象权限），不应授予非系统用户管理员等权限。

* + 1. 数据表空间

应根据业务场景和实际数据量来定义数据表空间使用固定大小方式或自增长方式。

* + 1. 海量数据分区

应对单表数据大于2000万或者单表占用空间大于3G的表进行表及索引分区设计。

* 1. 数据可信认证

数据可信认证应符合以下要求：

1. 基于统一数据中台构建数据及用户认证体系，实现数据安全统一管控；
2. 统一认证/用户管理，可支持以下方式：
3. 支持创建用户账户，设置密码和权限；
4. 支持定义用户组，设立密码标准管理用户、密码、用户组等；
5. 支持用户身份验证与鉴权。
6. 日志审计：支持监控用户身份认证和访问行为；支持对系统日志记录及查询功能；支持经常性分析；
7. 角色授权：支持权限配置、查询功能权限的分配情况；支持权限的禁止和互斥关系的配置；
8. 租户隔离：管理用户、密码、用户组和权限。
9. 数据价值服务要求
	1. 数据共享
		1. 基本内容

制造大数据共享服务通过统一数据中台进行，主要涵盖制造大数据资产登记及管理、目录编制和平台应用等工作。基于数据中台的数据共享需进行统一管理和维护，由相关部门提供技术支撑。

* + 1. 制造大数据资产登记及管理

制造大数据资产登记及管理，主要包括对各数据源采集、使用、产生、管理的数据资产清单进行登记及管理。

* + 1. 制造大数据目录编制

制造大数据目录编制主要包括通过统一数据中台对开放共享的大数据目录进行分级、编制、上传、更新及维护等工作。

* + 1. 制造大数据共享服务平台应用

制造大数据共享服务平台应用主要包括数据申请、获取、提供、使用、评价等工作。

* + 1. 制造大数据共享服务方式

制造大数据共享服务方式包括数据库方式、FTP方式、数据接口方式，每种方式需提供数据字典及格式说明等数据说明文档。

* + 1. 数据共享场景

数据共享场景主要包括同一组织上级部门与下属部门之间的数据共享，及各不同组织和机构或同一组织不同部门相互之间的数据共享。

* 1. 数据分层管理

不同组织或机构根据自己的业务需求可以分成不同的层次，来满足不同的业务需求。数据分层的类型如下：

1. ODS层：对操作系统数据不处理直接存放，为DWD层提供基础原始数据，可减少对业务系统的影响；
2. DWD层：存放历史明细宽表数据，主要对ODS数据层做一些数据清洗和规范化的操作，提供业务系统细节数据的长期沉淀，为未来分析类需求的扩展提供历史数据支撑；
3. DWS层：公共数据汇总层，用于提供后续的业务查询，OLAP分析，数据分发等；
4. ADS层：存放数据产品个性化的统计指标数据，该层主要是提供数据产品和数据分析使用的数据；
5. TMP层：ETL临时数据层。
	1. 数据安全管控与隐私保护

数据安全管控与隐私保护应符合以下规定：

1. 在使用数据过程中，数据安全应符合 GB/T 37973--2019的规定，个人隐私信息保护应符合GB/T 35273--2020的规定；
2. 数据分类分级，按照数据价值和重要程度，将数据划分为P0、P1、P2、P3四个等级。根据数据安全需要，在数据建模时进行等级标识，在数据应用时进行等级分类管控。具体定义如下：
3. P0：非常重要的主题域数据和非常重要的应用数据，具有不可恢复性，如交易、日志、集

团KPI数据等；

1. P1：重要的业务数据和应用数据，具有不可恢复性，如重要的业务产品数据；
2. P2：重要的业务和应用数据，具有可恢复性，如ETL产生的中间过程数据；
3. P3：不重要的业务和应用数据，具有可恢复性，如某些可通过历史明细数据生成的报表。
4. 数据访问授权，对生产系统的用户实施严格的访问授权管控，不应通过超级管理员权限在系统后台直接修改数据，要求各信息系统保留日志记录供数据安全审计，对非授权访问等不当行为将定期进行通报考核；
5. 对包括重要系统信息在内的数据，要求技术支持相关部门进行数据存储安全加固防护，定期进行备份，以保障核心敏感数据不易丢失、外泄。批量使用生产系统数据应通过申请和审批，开发测试环境使用生产系统的数据，则应进行数据转换变形处理；
6. 对于核心商业秘密、个人隐私等敏感数据的传输实施严格的审批流程，经数据管理相关部门审批同意后方可进行传输，并要求技术支持相关部门采取必要的数据加密、数据脱敏等措施。
7. 数据资产化管理要求
	1. 元数据管理
		1. 系统注册管理

所有现网系统，应由技术相关单位在数据资产管理平台进行元数据注册登记。

* + 1. 模型管控

数字化转型相关部门定义定时采集任务，将现网数据库的数据表结构采集并存入管理平台，对数据模型的变更进行稽核比对。技术相关部门对现网模型进行变更时，需提交模型变更申请，并将变更原因，变更内容（DDL脚本），相关需求等信息在模型管理平台中进行提交审批。

* + 1. 数据字典标准管控

数字化转型相关部门组织对业务系统的用语、单词、域进行梳理，形成常用的数据字典标准，并录入数据资产管理平台。当数据字典标准不完善时，技术相关部门可进行定义，并在平台内提交新增数据字典标准申请。

* + 1. 元数据分析

数字化转型相关部门定义采集任务，将现网数据库的技术元数据采集并存入管理平台，进行补充完善，并组织业务相关单位整理业务系统的管理元数据，业务系统将表之间的依赖关系梳理完善后，可在数据资产管理平台上查看各类血缘分析、影响分析。

* 1. 数据标准管理

数据标准应优先遵循国家、行业等相关标准，由数字化转型相关部门组织制定共性通用的数据标准，

各领域数据管理相关部门组织制定领域数据标准，技术支持相关部门和业务单元参与标准的制定，以实现数据语言的统一，保证跨系统数据的一致性和共享使用。

数据标准包括主数据、元数据、指标数据等标准规范，不同的数据标准规范分类分批发布后执行，并根据业务需求持续完善：

1. 主数据标准通过MDM的标准模板进行固化，除ERP系统外，其他信息系统不应自行维护主数据信息表，需要时由MDM进行主数据分发调用；
2. 元数据标准通过DAM进行管理控制，主要对数据库表、字段命名等进行规范检查，新建信息系统需通过DAM进行元数据注册、审核；
3. 指标数据由领域数据管理部门负责数据业务规则定义，对每一个指标数据需确定至少1名业务负责人负责指标定义及解释，并明确数据技术负责人，负责数据来源链路及处理过程的解释。指标数据的标准应通过数据仓库和大数据平台进行固化、维护。
	1. 数据模型管理

数据模型管理应符合以下要求：

1. 数字化转型相关部门统一规划和建设数据仓库和大数据平台，集中汇聚内外部的各类数据资源，按照业务模型，通过标准化用语、单词等数据要素来规划设计数据模型，并通过数据基础管理平台进行标准固化；
2. 在信息系统建设和运行维护过程中，技术支持相关部门通过数据模型管理工具，统一管控数据模型的新增和变更，各业务单元需为数据模型梳理提供业务支持；
3. 为维护数据架构的统一，各业务单元未经同意，不应擅自建设大数据、数据仓库、商务智能等数据基础平台；各业务单元及技术支持相关单位需遵从和维护数据模型，并根据需求变化及时对数据模型进行更新维护。
	1. 数据治理

数据治理应符合以下要求：

1. 关于组织智能战略和数字化转型，结合业务变革和IT项目实施，由数字化转型相关部门牵头进行数据治理，制定数据治理方案，组织业务单元和技术支持单位实施；
2. 各业务单元依托数据管理相关人员协同参与数据治理，明确相关数据的责任人等角色，以完成本单位数据治理工作，并形成数据治理的常态工作机制；各业务单元可提出领域内的数据治理需求，跨部门、跨系统的数据治理由数字化转型相关部门评估后组织实施；
3. 数字化转型相关部门定期组织开展数据治理相关交流和培训，并对数据治理过程和结果进行检查、总结，形成数据治理简报，定期发布；对共性问题和解决方案进行知识沉淀，组织整理和制定相应的数据知识库、管理规范等。
	1. 数据生存周期管理

数字化转型相关部门牵头组织制定数据生存周期管理规范，制定不同的管理策略，明确不同数据的保存时限，数据清理流程，定期执行数据的归档、迁移和销毁。

* + 1. 数据的保存时限

除外部监管和内部管理已明确了数据保存时限外，还需确保数据至少有一个完整年度的数据周期，同时根据分析需要加上原先的数据，应保存不低于相关文件规定的时限，特殊情况宜根据相关规定做出调整，使得业务能够分析对应时间段的数据。

* + 1. 数据的冻结

2年未使用的主数据可进行冻结标识，经公示无异议后，在MDM系统进行冻结处理；若需解冻，则通过MDM系统提出解冻申请。

* + 1. 数据的归档

每年组织一次数据清理，每两年执行一次数据归档，对待归档的数据，经清理确认无相关业务数据及公告后，执行离线归档或存储迁移。

* + 1. 归档数据查询

根据业务管理或监管需要，技术支持相关部门负责对归档数据的查询请求进行管理，确保数据在需要时可恢复查询。

* + 1. 数据销毁的执行

对于已归档、冻结或需要清除的数据，根据数据生存周期管理流程，经申请审批及公示后，由技术支持相关部门执行数据销毁操作，并做好删除记录。

1. 制造大数据价值化应用要求
	1. 数据应用需求管理

数据应用需求管理要求如下：

1. 数字化转型相关部门统一收集各类业务数据应用需求，基于业务驱动原则，组织各类数据主题分析项目建设及推广应用，为各级管理者的决策提供数据支持；
2. 业务单元应根据业务需要向数字化转型部提交数据应用需求申请，经评估后开展相关项目实施或提供数据应用服务；
3. 各业务单元宜借助大数据、人工智能等技术对海量制造大数据进行处理和价值化应用。
	1. 数据应用项目实施

数据应用项目实施应符合以下要求：

1. 数据应用类项目，包括数据基础平台建设、维护、优化等，由数字化转型相关部门统一组织实施；
2. 各业务单元投资且仅限于本单位使用的数据报表、指标分析、数据应用系统等，为保证数据模型和系统应用级数据模型的一致性，经数字化转型相关部门审核备案后，由业务单元组织实施；
3. 对涉及需要上报的指标，要求应使用统一的数据源和指标口径，不应自行进行数据加工处理。

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 35273-2020 信息安全技术 个人信息安全规范
3. GB/T 36344-2018 信息技术 数据质量评价指标
4. GB/T 37973-2019 信息安全技术 大数据安全管理指南
5. GB/T 37988-2019 信息安全技术 数据安全能力成熟度模型
6. GB/T 38667-2020 信息技术 大数据 数据分类指南
7. GB/T 39400-2020 工业数据质量 通用技术规范
8. GB/T 40685-2021 信息技术服务 数据资产 管理要求
9. GB/T 38667-2020 信息技术 大数据 数据分类指南

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_