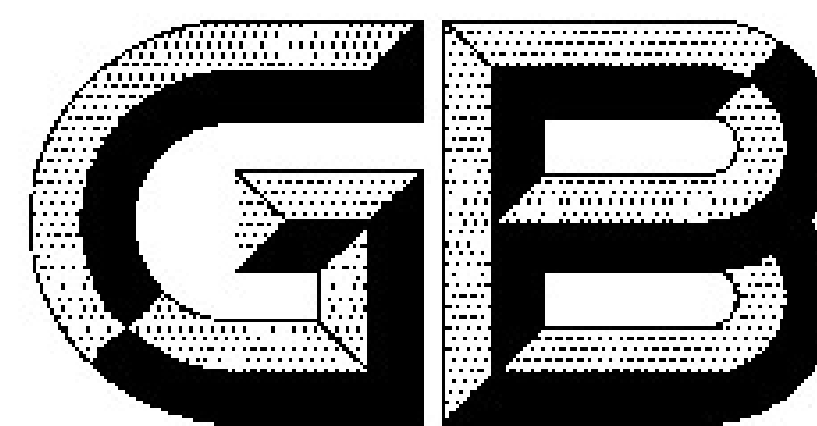




中华人民共和国国家标准

GB/T 42450—2023

信息技术 大数据 数据资源规划

Information technology—Big data—Planning of data resource

2023-03-17 发布

2023-10-01 实施

国家市场监督管理总局 发布
国家标准化管理委员会

目 次

前言 I

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 概述 2

5 数据资源规划流程 2

6 数据资源预配置相关活动 3

 6.1 数据资源现状盘点 3

 6.2 数据资源规划目标确定 3

 6.3 数据资源规划文件编制 3

7 数据资源规划相关数据活动 3

 7.1 数据处理相关活动 3

 7.2 数据管理相关活动 4

参考文献..... 6

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：沧州市大数据管理办公室、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、方正国际软件(北京)有限公司、美林数据技术股份有限公司、北京华宇信息技术有限公司、北京三维天地科技股份有限公司、深圳市华傲数据技术有限公司、中国航空综合技术研究所、中国人寿保险股份有限公司、山东润一智能科技有限公司、中汽数据(天津)有限公司、中电长城网际系统应用有限公司、中国长江三峡集团有限公司、北京北达软信息技术有限公司、成都四方伟业软件股份有限公司、源数科技(无锡)有限公司、无锡雪浪数制科技有限公司、河南科技大学。

本文件主要起草人：刘博胜、张群、李冰、杨瑛、戴炳荣、韩冰、张煜、李晓燕、曹珍、徐征、豆昱、彭晓刚、林志鹏、刘俊良、王兆君、黎俊茂、刘利达、楚思思、张凯、邵俊、杨芬、唐进、应志红、姚乐、姜向军、高洪美、旷志光、程宏斌、逢振、王峰、张志勇。

信息技术 大数据 数据资源规划

1 范围

本文件描述了组织用以支持实现其业务规划目标所需的数据资源规划流程,规定了数据资源预配置和数据资源规划相关活动。

本文件适用于指导组织建立满足其业务需要的数据资源规划相关文件及活动。

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 26816 信息资源核心元数据
- GB/T 35273 信息安全技术 个人信息安全规范
- GB/T 37722 信息技术 大数据存储与处理系统功能要求
- GB/T 38667 信息技术 大数据 数据分类指南

3 术语和定义

下列术语和定义适用于本文件。

3.1

数据资源 data resource

作为资源看待的用于支持实现组织业务目标的数据。

注：本文件的“业务目标”指的是组织围绕其业务和业务发展领域明确或隐含设定的目标。为达到此类业务目标,可能需要一系列资源提供保障,作为资源看待的数据是其中之一。

3.2

数据资源规划活动 data resource planning activity

以支持实现组织业务目标为目的,对当前和未来一段时期里组织内所需数据资源以及相关活动进行的系统化设计。

3.3

数据资源规划文件 data resource plan document

对数据资源开展规划活动产生的最终输出物。

注：数据资源规划文件是数据资源预配置的成果物,呈现为支持实现组织业务战略/业务目标所开展的明确数据现状梳理、数据资源规划目标及任务制定、相关路径及具体活动策划等。

3.4

数据资源目录 data resources catalog

与组织业务活动相适应的、便于检索的多层次分类、分级的数据资源编排形式。

3.5

数据管理 data management

在数据处理系统中,提供对数据的访问,执行或监控数据的存储,以及控制输入输出操作等功能。

[来源:GB / T 5271.1—2000,01.08.02]

3.6

数据资源预配置 data resource preconfiguration

针对数据资源规划确定的目标活动,梳理确定的数据资源当前状态和现状条件、数据资源未来组织方式和状态、数据资源配置路径和计划。

4 概述

数据资源与支持组织业务活动的其他各类资源类似,数据资源规划围绕组织业务规划中设定的目标以及数据本身特征来展开。数据资源规划活动所产生的最终产品是满足业务目标数据需求的数据资源预配置方案,即数据资源规划文件。数据资源规划目标是支持组织实现业务规划与目标,以确定的业务规划与目标为输入,系统性设计开展数据资源各要素的规划、描述,提出规范性要求。

5 数据资源规划流程

组织内数据资源规划流程见图 1,主要涉及两类活动。

a) 数据资源预配置相关活动:

- 1) 该类活动包含但不限于:数据资源现状盘点、数据资源规划目标确定和数据资源规划文件编制等;
- 2) 该类活动是开展数据资源规划所执行的活动,其目标是完成数据资源规划文件;
- 3) 该类活动是有顺序的。对于特定的数据资源规划任务而言,流程中的一些活动可能多次或反复执行,如从数据资源规划目标到数据资源配置方案可能会根据实际情况多次迭代重复。

b) 数据资源规划相关数据活动:

- 1) 数据处理的相关活动,如数据收集、数据存储、数据处理、数据交换、数据应用、数据删除和长期保存;
- 2) 数据管理的相关活动,如分级分类管理、数据模型管理、数据标准管理、元数据管理、主数据管理、数据目录管理、数据质量管理、数据安全保护和数据保障等;
- 3) 该类活动是各项独立的。活动顺序及是否执行所有活动,本文件只给出并描述一般需要执行的活动,是否确定活动的先后顺序,以及是否执行所列出的所有活动,由组织根据实际情况和要求考虑确定。

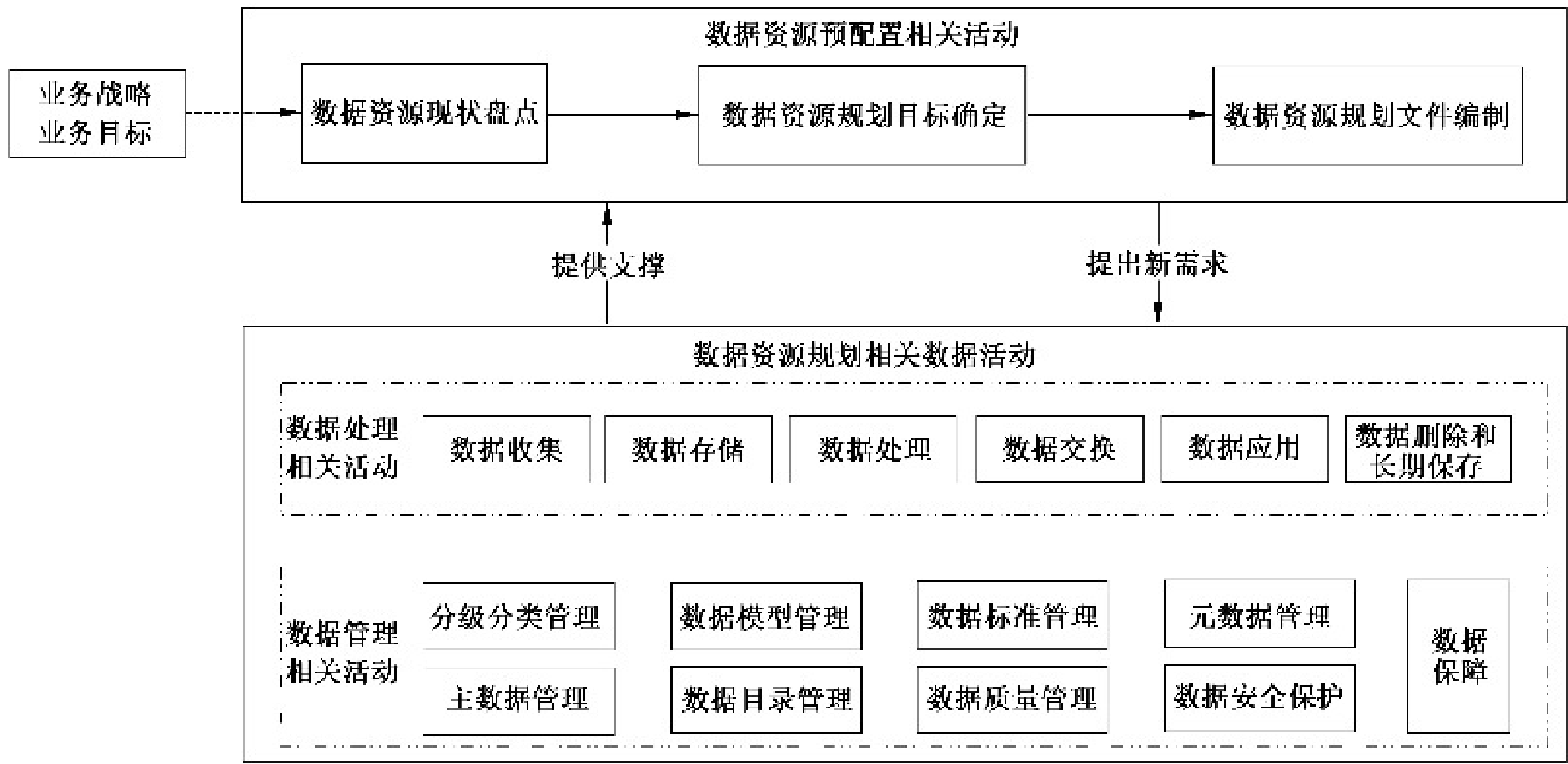


图 1 数据资源规划流程

6 数据资源预配置相关活动

6.1 数据资源现状盘点

基于业务目标，深入分析业务需求与问题，从业务视角对组织的数据资源盘点，明确业务域、数据流向等；从信息技术视角对组织的数据资源盘点，识别业务系统中的数据分布、数据存储。

6.2 数据资源规划目标确定

基于数据资源现状盘点结果，规划满足业务战略/业务目标的数据活动目标及任务，分解确定数据管理提升的各阶段以及子目标，包括数据管理能力成熟度提升相关计划、数据处理及数据管理相关子任务等。

6.3 数据资源规划文件编制

针对数据资源规划目标活动中确定的阶段目标及任务，结合各类数据资源的现状与需求，开展数据资源预配置方案，即数据资源规划文件的编制。数据资源规划文件中包括数据处理和数据管理相关内容以及对数据资源配置产生影响的其他相关因素，如数据架构、技术平台，业务数字化转型等相关内容。

数据资源规划文件的具体编写形式和内容，取决于组织对业务、技术以及其他管理等的要求。

7 数据资源规划相关数据活动

7.1 数据处理相关活动

7.1.1 数据收集

数据收集活动主要包含但不限于以下各项内容：

- a) 根据数据来源、数据量级、数据格式、更新方式与频率等因素进行数据收集；
- b) 在收集活动中针对所要收集数据的特点进行数据收集。

7.1.2 数据存储

数据存储活动主要包含但不限于以下各项内容：

- a) 根据数据结构特征、数据量级、时效性要求、读写性能需求、并发需求、访问频率和数据应用需求等因素,选择适当的存储方式；
- b) 建设数据资源库,对各类数据资源及数据流转过程中的数据资源进行分层分类存储；
- c) 数据存储活动应按照 GB/T 37722 规定的相关内容开展。

7.1.3 数据处理

数据处理活动主要包含但不限于以下各项内容：

- a) 根据数据量级、数据格式、数据时效性要求 and 安全性要求等因素进行数据加工、数据传输等；
- b) 针对所要处理的数据特点运用进行数据处理。

7.1.4 数据交换

数据交换活动主要包含但不限于以下各项内容：

- a) 根据组织内协同场景的业务需求进行交换；
- b) 根据交换的实时性需求、数据量级等因素进行交换。

7.1.5 数据应用

根据数据预处理结果,结合组织的实际应用需求,运用适当的数据处理技术为组织开发和(或)推行新应用、新模式提供所需数据的应用。

7.1.6 数据删除和长期保存

根据组织需求选用适宜的数据删除技术、存储媒体等进行数据删除、销毁和数据长期保存。

7.2 数据管理相关活动

7.2.1 分级分类管理

数据分级分类管理主要包含但不限于以下各项内容：

- a) 按照组织规定的分级分类原则、形式和方法,针对数据本身涉及的主题领域、数据来源以及预期的数据用途对数据资源进行区分、归类、分级；
- b) 确定具体的数据分类、数据分级方案,该方案中的数据分类可能是多维度的多种方案；
- c) 数据分级分类活动应按照 GB/T 38667 规定的相关内容开展。

7.2.2 数据模型管理

数据模型相关活动涉及模型建立和维护。此类活动主要包含但不限于以下各项内容：

- a) 根据组织拥有的数据本身特点和数据处理活动的需要建立相应的数据模型；
- b) 根据组织的业务需求、技术管理方面的变更以及相关数据当前特点的变化进行数据模型管理；
- c) 基于实际业务活动对实体增加关键属性,并体现模型间关系。

7.2.3 数据标准管理

数据标准相关活动主要包含但不限于以下各项内容。

- a) 标准体系规划,此项活动主要包含但不限于以下各项内容：

- 1) 建立覆盖组织数据资源规划过程中涉及数据的数据标准体系；
- 2) 划定标准分类及各类标准下的数据标准目录,并对各类标准进行统计汇总。
- b) 数据编码规划,此项活动主要包含但不限于以下各项内容:
 - 1) 建立组织级的数据编码体系；
 - 2) 对数据编码标准的执行情况进行动态监管,并根据需要更新编码标准。

7.2.4 元数据管理

元数据相关活动涉及元数据的建立和管理。此类活动主要包含但不限于以下各项内容：

- a) 建立组织内的元模型,用于描述本组织的元数据的统一的组织结构；
- b) 对元数据的更新变更进行动态监管；
- c) 对于元数据描述、元数据扩展等应按照 GB/T 26816 规定的相关内容开展。

7.2.5 主数据管理

主数据相关活动主要包含但不限于以下各项内容：

- a) 识别梳理主数据；
- b) 建立主数据标准,指导主数据的统一管理；
- c) 对主数据标准的执行情况进行动态监管以满足组织内各部门和(或)业务各环节使用；
- d) 规划主数据在业务系统间的集成关系、集成标准,指导主数据的集成。

7.2.6 数据目录管理

数据目录管理主要包含但不限于以下各项内容：

- a) 建立数据目录编制标准,明确数据目录的结构、层次和适用范围；
- b) 建立数据目录的维护、监测机制,使其所包含的数据内容持续满足业务对于数据资源的要求。

7.2.7 数据质量管理

数据质量活动主要包含但不限于以下各项内容：

- a) 建立并持续执行组织内数据质量校验规则；
- b) 建立组织内数据质量检验机制,及时发现并解决数据质量问题。

7.2.8 数据安全保护

数据安全保护活动主要包含但不限于以下各项内容：

- a) 建立数据安全管理体系,持续执行数据安全保障规章制度,保护、合规使用和处理敏感数据；
- b) 建立数据安全审查制度,定期开展数据种类、数量、面临风险、应对措施等的风险评估；
- c) 针对个人信息应建立相应的识别和保护机制,应按照 GB/T 35273 规定的相关内容开展。

7.2.9 数据保障

数据保障主要包含但不限于以下各项内容：

- a) 建立组织的数据保障机制,针对数据相关活动制定管理标准；
- b) 按照数据安全保障文件的规定执行数据安全保障活动；
- c) 提供支持执行数据保障的各类资源。

参 考 文 献

- [1] GB/T 5271.1—2000 信息技术 词汇 第1部分:基本术语
 - [2] GB/Z 18219—2008 信息技术 数据管理参考模型
 - [3] GB/T 18391.1—2009 信息技术 元数据注册系统(MDR) 第1部分:框架
 - [4] GB/T 36073—2018 数据管理能力成熟度评估模型
 - [5] GB/T 36344—2018 信息技术 数据质量评价指标
 - [6] GB/T 37973—2019 信息安全技术 大数据安全管理指南
-