

中华人民共和国金融行业标准

JR/T 0107.1—2014/ISO 17369-1:2013

统计数据和元数据交换 (SDMX)
第1部分: 框架

Statistical data and metadata exchange (SDMX)—

Part 1: Framework

(ISO 17369-1:2013, Framework for SDMX technical standards, IDT)

2014-08-28 发布

2014-08-28 实施

中国人民银行 发布

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 SDMX 标准框架结构	2
5 流程和业务范围	3
5.1 流程模式	3
5.2 SDMX 和自动处理	4
5.3 统计数据和元数据	4
5.4 统计交换的 SDMX 概述	5
5.5 SDMX 注册服务	6
5.6 Web 服务	7
6 SDMX 信息模型	7
7 SDMX-EDI	8
8 SDMX-ML	8
9 一致性	9
10 对于 SDMX 内容导向指南的依赖性	9
10.1 跨域概念	9
10.2 元数据常用术语	10
10.3 统计专题领域	10
参考文献	11

前 言

JR/T 0107《统计数据和元数据交换（SDMX）》分为七个部分：

- 第 1 部分：框架；
- 第 2 部分：信息模型 UML 概念设计；
- 第 3 部分：SDMX-ML 模式和文档；
- 第 4 部分：SDMX-EDI 语法和文档；
- 第 5 部分：注册表规范 逻辑功能和逻辑接口；
- 第 6 部分：SDMX 技术说明事项；
- 第 7 部分：Web 服务用法指南。

本部分为 JR/T 0107 的第 1 部分。

本部分依据 GB/T 1.1-2009 规则起草。

本部分等同采用 ISO 17369-1:2013《统计数据和元数据交换（SDMX） 第 1 部分：框架》。

本部分由中国人民银行提出。

本部分由全国金融标准化技术委员会（SAC/TC 180）归口。

本部分主要起草单位：中国人民银行调查统计司、中国金融电子化公司。

本部分主要起草人：盛松成、徐诺金、姚力、巴运红、任全忠、潘润红、李曙光、韩建国、贾树辉、李兴锋、曹小艳、吴隼、廖燕平、王媛、司燕翔、刘蔚、张艳、吴永强、邓琳莹、李静。

引 言

统计数据和元数据交换（SDMX）标准由 SDMX 国际组织发起并提出。SDMX 国际组织是由国际清算银行（BIS）、经济合作与发展组织（OECD）、欧盟统计局（Eurostat）、欧洲中央银行（ECB）、国际货币基金组织（IMF）、联合国（UN）和世界银行（WB）七个国际组织联合建立，其制定发布的《统计数据和元数据交换》标准规定了统计人员在采集、处理和交换统计数据时所使用的统计概念和方法，规范了对外披露统计信息时统计数据的机构范围、地理区域、存量性质、时间属性、频度以及文件格式等内容。

SDMX 标准提供了统计数据及元数据交换和共享的标准化格式，可以达到更好地扩展和高效率使用的目的。目前 SDMX 标准主要应用领域为部分国家中央银行和统计部门。本部分的作用在于规范我国金融统计标准体系的内部处理和对外发布，促进金融统计的互联互通、信息共享和业务协同，提高信息共享的效率，满足金融综合统计的需要。

统计数据和元数据交换（SDMX） 第1部分：框架

1 范围

本部分规定了统计数据和元数据交换的框架结构、流程和业务范围、SDMX信息模型、SDMX-EDI、SDMX-ML、一致性、对于SDMX标准的依赖性等内容。

本部分适用于金融统计中数据和元数据的交换和共享。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是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/T18391 信息技术 元数据注册系统（MDR）

3 术语和定义

下列术语和定义适用于本文件。

3.1

数据集 data set

由数据组成的离散集合（包含特定时间周期内的典型观测结果），即一段固定时间周期内共享一种结构的相似数据的集合。

3.2

数据结构定义(DSD) data structure definition

每个数据集都具有一系列结构元数据，这些说明性内容包括概念如何与数据“立方体”的观测结果、维度以及属性有关的信息，以及数据和相关识别及说明（结构）元数据表示形式的相关信息。

注：在SDMX V2.1以前的版本中，数据结构定义称为关键字族key family。

3.3

代码表 code list

枚举维度、属性和SDMX其他结构部分的表示形式中使用的一系列值，可用其他结构元数据（表示编码如何组合分层次）补充。

3.4

类别方案 category scheme

包括用于数据和元数据组织的任一种有用分类，由不同等级的类别组成。

3.5

概念方案 concept scheme

一系列概念的可维护列表，这些概念用于数据结构定义和元数据结构定义。

注：概念方案可有很多种。概念的核心表述能够被详细说明。例如，核心代码表或其他表示形式，如“日期”。该核心表示形式在使用概念的数据结构定义或元数据结构定义中会被取代。实际上，希望保持原先版本数据结构定义方案规范的组织，将继续使用数据结构定义中的表示形式。

3.6

元数据集 metadata set

统计交换的正式SDMX概述中几乎所有对象的信息集合。

注：元数据集可描述数据或结构定义的维护方，也可描述数据发布的计划安排，可给出一段时间内单一类型数据流，也可以给出数据质量等。在SDMX中，参考元数据的生成方可获取其关心的任何概念或强制报告，并提供包含该信息的参考元数据集。

3.7

元数据结构定义 metadata structure definition

标识了公布的参考元数据内容的元数据集。

注：参考元数据集也有一组描述自身如何组织的结构元数据。这些概念之间的关系（特别作为层次结构），表示的结构，发布的表示格式（自由文本、编码值等），关联的SDMX对象类型。

3.8

数据供应方 data provider

生成数据或参考元数据的组织。

3.9

约束条件 constraints

给出了数据来源或元数据来源的子集，并且也可提供数据预计发布的信息。

注：约束条件与数据供应方、供应协议、数据流、元数据流、数据结构定义和元数据结构定义有关。

3.10

结构集 structure set

提供了一个机制，该机制可在结构元数据集中，形成一个对于特定、相关数据和元数据集关系的完整描述。

注：结构集可用于维度和属性、概念、代码表、类别方案的相互映射。

3.11

层级代码表 hierarchical code list

支持代码层级的规范。

注：代码引自代码表。一个或多个层级中，层级代码表规范的是代码组织，但并不能定义代码本身。

3.12

SOAP Simple Object Access Protocol

简单对象访问协议。

3.13

WSDL Web Services Description Language

网络服务描述语言。

3.14

WADL Web Application Description Language

网络应用程序描述语言。

3.15

URL Uniform Resource Locator

统一资源定位符。

4 SDMX 标准框架结构

统计数据和元数据交换（SDMX）规定了能够促进统计数据和元数据交换的标准，这些标准使用现代信息技术，重点关注于聚合数据。

SDMX技术规范包括以下部分：

SDMX框架：介绍了SDMX及其范围。为了满足一致性要求，这部分内容将有所修正。

SDMX信息模型 UML概念设计：其他部分中给出的具体执行的语法是基于该信息模型的，目的是为技术人员按中性语义格式理解整体技术标准。该部分包含以附录形式给出的UML指南。该部分是资料性的。

SDMX-EDI：定义了SDMX结构的数据和元数据交换的UN/EDIFACT格式，包含描述SDMX报文中UN/EDIFACT语法用法的规范性章条。该部分包含规范性章条。

SDMX-ML：定义了SDMX结构的数据和元数据交换的XML格式，包含描述SDMX报文中XML语法用法的规范性章条，并提供一系列规范性的XML模式和非规范性的XML文件示例。

SDMX注册表规范：给出了可用的数据和参考元数据的集中注册信息，并给出了结构元数据和供应信息的存储库；定义了SDMX注册提供的基本服务，包括数据和元数据注册、数据和元数据查询及注册服务更新的订阅和通知。该部分包含规范性章条。

SDMX技术注解：用于帮助希望使用SDMX规范的有关人员。该注解提供各种报文和语法间表述性差异的描述，版本变更，维护机构，以及SDMX注册。该部分是资料性的。

Web服务使用指南：用于帮助希望使用Web服务技术实现SDMX标准的有关人员。该指南重点给出了Web服务技术的相关内容（包括但不强制SDMX一致性注册），该技术能够在创建Web服务时不用考虑开发环境或平台的差异性。该部分包含部分规范性章条。

5 流程和业务范围

5.1 流程模式

SDMX设计了统计数据和元数据交换的三种基本流程模式。这三种模式说明如下：

- 双边交换：交换流程的所有内容均被双方所认可，包括数据和元数据的交换机制、格式、频率或计划安排以及交换通信使用的模式。该模式可能是最常用的流程模式。
- 网关交换：网关交换是一系列有组织的双边交换，在网关交换中数据和元数据收集机构或个人遵循将收集到的信息以单一的、已知的格式并按照单一的、已知的流程进行交换。该模式通过共享机构/个人可以减少多重双边交换（在数据和元数据收集过程中）的管理负担。该模式也是统计领域常用的类型，在该模式中，各机构均同意以该方式获得其共同责任范围内的高效性。
- 共享数据交换：已知且标准的开放式、可随意获得的数据格式和流程模式。因此，任何组织或个人可使用任一当事方的数据及元数据（假定当事方已被允许获得相关数据）。该模式不要求双边协议，仅要求数据和元数据供应方及使用方遵守该标准。

本部分规定了用于促进基于任一流程模式交换的SDMX标准，并指明了SDMX在任意情况下的优越性。协议双方可能共同使用一种标准格式（例如SDMX-EDI或SDMX-ML），也可能对于网关流程中的数据发送方使用彼此认可的一种标准格式进行数据交换，或同其他同意使用该标准格式的数据供应方进行数据交换；也可协议使用完整SDMX标准集合以支持交换中常用的数据共享流程，该流程基于SDMX一致性注册服务或其他架构。

SDMX标准通过集中注册服务支撑数据共享流程交换。注册服务为群体中现有数据和元数据提供了可视性，并通过为自动处理过程提供一系列触发程序以支持数据和元数据的访问和使用。数据或元数据本身在集中注册时并不存储（这些服务仅在已知位置提供了关于数据以及附加元数据的元数据有用集合），

用户/应用软件都能方便地定位并获取已注册的数据和/或元数据。显然,该标准可用于所有数据、元数据及注册服务,从而允许在数据共享群体间内高度的自动化。

不同的流程模式并不互斥,使用SDMX格式表述数据和元数据的单一系统可支持所有的三种交换场景。不同标准可适用于不同流程(例如,很多注册服务接口仅适用于数据共享场景),但所有流程的共享信息模型都是相同的。

除关注于收集和报告外,数据发布也很重要。不论统计数据 and 元数据在生成和开发过程中当事双方如何交换,最终都是提供给某些类型的终端用户。通常情况下,该过程通过机构特定的应用软件实现。然而,统计数据和元数据越来越频繁地以各种格式被发布到网站上。Web上数据及其伴随元数据的发布是SDMX标准的主要内容。统计数据和元数据标准允许对数据发布的更新,一旦获得数据,处理标准格式将更轻松可行,且数据和元数据彼此关联,数据理解和进一步处理变得更为容易。

在讨论统计数据时,数据发布的很多内容会影响到数据质量,这些内容包括数据发现、使用的便捷性以及及时性。SDMX标准为数据发布的所有内容均提供了支持。标准数据格式促进了使用便捷性,并提供了与相关元数据的链接。注册服务概念意味着更便捷地发现数据和元数据。通过在整个数据生命周期提高效率,数据及时性得到了改善,并使元数据可用性和使用便捷性得到促进。

应注意,虽然SDMX主要关注于统计数据和元数据的交换和发布。然而,此处规定的标准模型和格式在数据内部处理范围中也可能有多种应用,这些数据处理与机构和用户间的交换无关。显然,用于交换和发布目的的数据和元数据的标准格式也能促进机构和用户的内部处理,只是该内容并非SDMX关注的重点。

5.2 SDMX 和自动处理

统计数据和元数据交换有很多不同的自动流程,其中某些流程比其他流程应用得更广泛。此外,某些信息技术在今天的信息系统中几乎无处不在。SDMX 的目标是为那些自动流程和技术提供最为常用的标准。

简言之,这些标准可描述如下:

- 数据和元数据的批量交换:当事双方之间的整个或部分数据库的传送,包括增量更新。
- 互联网上数据和元数据的供应:互联网技术(包括其在私有或半私有TCP/IP网络中的使用)是相当普遍的,该技术包括作为数据和元数据自动供应主要机制的XML和Web服务,以及更传统的静态HTML和数据库驱动的发布。
- 通用流程:许多应用和流程是针对某些特定的数据和元数据集而设计,而其他类型的自动服务和流程设计可用于处理任何类型的统计数据和元数据。一旦门户网站和数据传输专线可在互联网上使用,则此项内容更为明显。
- 数据的表示和转换:统计数据和元数据若想真正对客户发挥作用,必须支持将其转换为特定应用的处理格式、其他标准格式以及表示格式的自动流程。虽然并未对自动处理的各个方面提出,但如果必须支持对手间的信息交换,则自动处理提出了一系列必须支持的要求。

5.3 统计数据和元数据

为避免对SDMX格式包含哪些“数据”和“元数据”产生混淆,特别提供了范围声明。统计“数据”通常是数值型观测结果的集合,这些观测结果与时间相关。统计数据与一系列元数据值和特定概念的表示形式有关,这些内容起到了数据的标识符和说明符的作用。这些元数据值和概念可被理解为一个多维坐标系的已命名维度,通常将该坐标系称为数据“立方体”。

SDMX标识了用于建模、表示以及理解该多维“立方体”结构的标准技术,从而允许各种来源数据的自动处理。该方法广泛适用于各种数据类型,并试图提供最简便、最易于理解的技术来支持大量数据和元数据的交换。

术语“元数据”实际上是非常广泛的概念。“结构”元数据和“参考”元数据是有区别的，前者用于统计数据 and 元数据的说明和识别，后者具有更宽泛的意义，它更为普遍地描述和定性了统计数据集和流程，且通常不与数据的特定观测结果或序列相关，但是与数据的整个集合或者甚至是提供这些数据的机构有关。

SDMX信息模型不仅支持数据的结构化，也支持“参考”元数据的结构化。当这些参考元数据结构独立于数据及其结构元数据存在时，他们通常是相关联的。SDMX信息模型为任一部分数据和结构元数据附加了参考元数据，也提供了参考元数据及其结构说明的报告和交换模型。SDMX标准的这种作用用于支持数据质量倡议的很多方面，允许其以最广泛的方式用于元数据的交换，在该数据质量倡议中与质量相关的元数据是最主要的部分。

元数据不仅与数据相关，也与数据流的提供和管理流程相关。SDMX信息模型提供了一系列与“数据供应”有关的元数据，这些元数据对于需要理解数据供应方输出的内容和形式的人员非常有用。每个数据供应方能以标准形式描述其产生的数据和元数据集的内容及其依赖关系，并提供关于其数据和元数据提供的计划安排和机制信息。该要求允许某些确认和控制函数的自动化，以及数据报告的支持管理。

SDMX也指出了分类模式在组织和管理数据及元数据的交换以及发布的过程中的重要性。该方案能表述SDMX中分类模式和域类别的信息及其与数据和元数据集的关系，并对模型中的其他对象进行分类。

SDMX标准提供了支持符合任一类型统计数据交换的常用模型、语义选择和基于XML的数据格式选择，SDMX-ML内容给出了基于每种实现方式的特殊要求进行的优化格式。

图1表示的是信息模型里的对象，后文还有更加详细的说明。

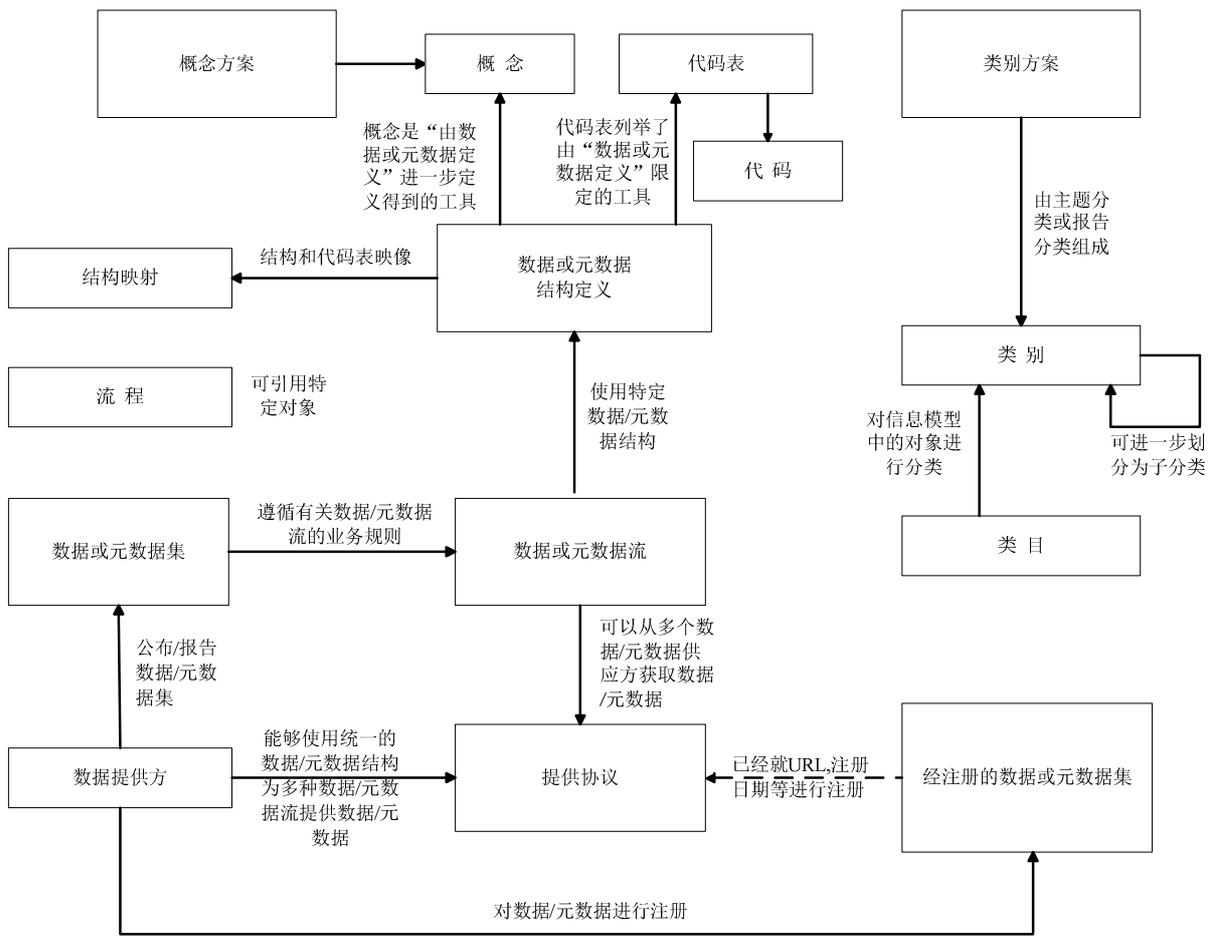


图1 SDMX 信息模型中主要工具示意图

5.4 统计交换的 SDMX 概述

ISO/TS 17369 SDMX的1.0版本覆盖与这些数据集结构有关的所有统计数据集和元数据。该范围在支持不同统计交换模型（双边交换、网关交换以及数据共享）方面非常有用。SDMX信息模型同时提供了更为广泛的一系列正式对象，这些对象描述了统计交换中的相关方、流程和系统，从而支持了更为完整的统计交换视图，因此，可完全支持开放的数据共享模型，并且其他交换模型可更为完全地自动化。

应注意理解正式对象的集合不仅是技术领域的，也是其在真实的统计数据和元数据交换中表示的内容。

SDMX给出了数据集（在特定时限内，根据特定结构报告的适用于数据结构定义的具体统计数据）以及描述统计数据集结构的元数据。统计交换中与此有关的一个概念是“数据流”（在持续更新的基础上为数据概念的报告或发布提供支持）。“数据流”可被理解为不绑定时间的数据集。数据集被机构拥有和维护（维护机构以同样的方式拥有数据流）。

版本2.0和2.1，类似于版本1.0，允许发布统计数据（和相关的结构元数据），也提供标准，参考元数据体系。参考元数据并不是数据集完整的一部分，但它独立于统计数据。SDMX提供了参考“元数据集”、“元数据结构定义”、“元数据流”。这些项是非常相似的数据集、数据结构定义和数据流，但它们关注的是参考元数据，而不是统计的观察值。同样的方式，数据提供者可以公布统计数据，也可以公布参考元数据。元数据结构定义的维护，与数据提供机构对数据结构定义及数据集结构定义相类似的方法维护。

数据和参考元数据的结构定义使特定统计概念与其表示形式产生联系，不论是文本还是代码等。这些概念从特定机构负责维护的“概念方案”中获得。概念方案组合了一系列概念，给出其定义和名称并当某些概念较其他概念非常特殊时，给出表述中的语义关系。某一单一概念方案可能同时用于数据结构（数据结构定义）以及参考元数据结构。

任一统计交换以及很多发布活动的本质是“服务等级协议”的概念，即使该概念未被定性或尚未明确。SDMX将该理念吸收进术语“供应协议”。数据供应方可向很多不同数据流提供数据。数据流可合并来自多个数据供应方的数据。供应协议是告诉你哪些数据供应方给哪些数据流提供了什么数据。该协议同样也适用于元数据流。

供应协议考虑到了各种可用的信息：统计数据或元数据报告或发布的计划安排，理论上可能的数据集内（如数据结构定义或参考元数据结构定义中描述的）关于数据或元数据报告的特定主题，以及统计数据和元数据覆盖的时限。这些信息集合在SDMX信息模型中的术语为“约束条件”。

下文定义一些概念：

- 组织方案：定义组织和组织结构的规划。其中，维护机构、数据供应方、数据使用方和组织单位使用特定的组织方案。
- 数据流定义：标识了数据结构定义并可能与分类中一个或多个专题域有关（该内容有助于根据分类模式查找数据）。
- 元数据流定义：与数据流定义非常类似，但其描述、分类并限制的是元数据集。

5.4.1 数据构建的注意事项

“立方体”是复杂的多维结构，可延其任一轴线（或“维度”）对其进行观察。同时，立方体数据的完整结构可在SDMX中描述，其目的是为传送而对数据进行格式化。多数SDMX数据格式为时间序列，即一系列围绕时间维度组织的观测结果，因此，每个观测结果按时间依次出现。

然而，某些类型的统计数据，并非按此方式进行交换，这样的统计数据称作“横截面”数据，其数据组成围绕立方体的其他非时间维度。SDMX为数据集中由单一维度组成的数据提供统一的格式。在这一环境中，时序是统一格式的一种特例。

数据统计“立方体”中经常出现的另一类型结构是层次分类，该分类用于描述沿其任一维度（或轴线）的点。此外，只要支持“稀疏立方体”的优化处理，则该标准也支持立方体不同维度间的函数依赖表示形式。该内容为“约束条件”的一部分，该约束条件考虑到了立方体区域的构架系统，或者考虑到了数据结构定义描述的所有数据结构中的有效数据结构集合的供应。

5.4.2 参考元数据构成的注意事项

元数据结构基于能组成语义层和表示层的概念，这些层次是构建XML报告格式的基础。SDMX-ML中包含用于此目的的3种报文类型：结构报文（提供元数据结构定义）、通用元数据报文（为任一元数据结构定义提供单一格式）以及特定结构元数据报文（提供元数据结构定义特定格式）。通常，该机制用于支持参考元数据报告和发布。

元数据结构定义从概念方案中抽取任一概念，并描述其如何形成作为元数据属性的报告或发布结构——以平面列表或层次方式。为这些元数据属性赋值以数据类型（编码、文本等）和出现次数。该元数据的“目标”是流程、信息、组成方式、交换等的分类，这些都是元数据的描述对象。由于SDMX信息模型给出了统计交换和发布的统一形式，该模型可用作统计活动中不同执行方和来源的类型分类法。因此，参考元数据集和元数据流的“目标”（对象）可作为参考该模型的某些标准分类的对应内容进行描述。

就数据结构而言，元数据集通用格式提供了一种已知的文档结构，同时，特定结构格式来源于元数据结构定义并能够进行高度方案确认。

5.5 SDMX 注册服务

为使得统计交换的SDMX模型中的大量数据和元数据可见，基于一系列注册服务的架构可能会非常有用。“注册表”（如在Web服务术语中给出）是一种存储了用于维护和查询的元数据并能被具有足够访问权限的网络中的其他应用程序所使用的应用文件（应注意的是，访问控制的机制在SDMX标准范围之外）。其可理解为分布式数据库或元数据库的标志，该数据库或元数据库由因特网上或类似网络中统计领域内的数据供应方的数据集和参考元数据集组成。

注意，SDMX注册服务不关心数据或参考元数据的存储。假定数据和参考元数据取决于其数据供应方的网站。SDMX注册服务考虑到了其为数据和参考元数据提供的可见性以及用于访问数据和参考元数据的信息。因此，已注册数据集在注册系统中将有其URL，但并非数据本身。需要访问数据的应用程序可查询注册系统，或可通过注册数据源的URL在类别方案和数据流中查找，就能够直接从数据供应方检索数据（借助SDMX-ML查询报文或其他机制）。

SDMX不要求注册系统的具体技术实现方式，只规定了注册系统可支持的标准接口。因此，用户可以用自己选择的任何方式进行SDMX一致性注册，只要支持该接口，如此处规定。这些接口以XML文档表示并形成新的SDMX-ML语言。

这里讨论的注册服务可以简略的归纳为：元数据结构维护、数据和元数据源注册、查询、订阅/通知。

- 元数据结构维护：该注册表服务允许有维护机构使用权限的用户提交和修改结构元数据。从这一层面上来讲，注册表类似于一个结构元数据存储库。但是，在SDMX结构中，允许提交结构对象的一个“根”（比如一个代码表），不论是一个文件还是结构元数据源（如另一个注册表），这个根都可以从元数据检索引用出实际的位置。
- 数据和元数据源注册：该注册服务允许具有维护访问权限的用户去通知注册系统已有数据集和参考元数据集的位置（为检索目的）。注册系统储存了关于该对象的元数据，并将该元数据同结构元数据联系，这些结构元数据为处理及查找其位置的应用程序提供充分的结构信息。注册系统中的对象按照一个或多个类别方案进行组织和分类。

- 查询：注册服务具有对库中包含的元数据进行查询的页面，因此，应用程序和用户能够发现数据集和参考元数据集、结构元数据是否存在，以及与这些对象有关的供应方/代理机构，以及描述数据和元数据如何获得以及如何进行分类的供应方协议。
- 订阅/通知：可能需要在存储库中“订阅”特定对象，因此每当存储库内容更新时，将向所有订户发送通知。

5.6 Web 服务

该服务允许计算机应用程序直接在互联网进行数据交换，特别是允许以比过去更加灵活的方式进行模块或分布式计算。然而，为使得Web服务发挥其功能，需要很多标准，这些标准用于申请和提供数据、用于表示用来打包交换数据的封装数据及用于Web服务的互相描述，可以使Web更容易地集成到使用其他的网络服务作为数据资源的应用程序中。

SDMX为这些使用标准提供了指南，这种方式提升了SDMX Web服务的互用性。允许通用客户应用程序的开发，该程序能够同实施这些指南的任何SDMX Web服务进行有意义的交流。

更确切地说，SDMX Web服务指南提供了：

- 基于简单对象访问协议（SOAP）的Web服务标准接口（WSDL）：2.0版本的Web服务指南包括了一系列Web服务功能，但在实施中，这些功能不再适用于当前SDMX基础上的Web服务。此外，WSDL的操作及其有效载荷已经成为规范（WADL）。
- RESTful Web服务标准接口（WADL）：RESTful的应用程序接口注重简洁，其目的不在于复制SDMX-ML查询报文的全部语义，而是能够简单地执行一个标准查询的有限集合。同时，不同于SDMX规范的其他部分，restful的应用程序接口仅注重数据检索（通过HTTP获得）。

一个常见的错误代码规范表：在Web服务使用中，为解释遇到的各种问题，应当建立一个错误代码。在SDMX标准2.1版本之前，没有统一的SDMX Web服务使用错误代码，目前的2.1版本填补了这一空白。

6 SDMX 信息模型

SDMX提供了统计数据的建模方式并定义了用于此目的的元数据结构集。由于SDMX规定了两种语法格式来表示数据和结构元数据，则该模型被用作保证不同格式间无损传输的一种机制。所有格式都是通用信息模型的语法表示形式。

SDMX承认统计数据是结构化的；在SDMX中，该结构被称为“数据结构定义”。根据其相似程度，“数据集”由一个或多个下层“组”构成。每个组依次包含一个或多个数据“序列”。每个序列或部分具有一个“关键字族”——每个概念簇的值，也称为“维度”——该维度标识每个序列或部分以及一个或多个“观测”，该观测结果通常组合了观测的时间和观测结果。此外，元数据可作为说明性“属性”附加到该结构的任一等级。代码表（枚举值）及数据和元数据表示的其他形式也被建模。

通常用于处理统计数据的“立方体”结构与SDMX信息模型中的“数据结构定义”理念间存在某些相似性。应注意，根据SDMX信息模型构建的数据为交换而被优化，可能与不具备处理来自复杂统计系统的数据“立方体”的能力的合作机构有关。SDMX时序可被理解为立方体的“一部分”，该“部分”由其关键字识别。关键字包含除时间外，数据结构定义预见的所有维度值。当然可对来自SDMX结构化数据的数据立方体进行重建和描述，并根据推荐的标准交换这些数据库。在V2.0中，更多的通过层级代码表、约束条件、数据结构定义的关系来描述结构化的立方体，都成为了可能。

SDMX标准也提供了参考元数据的视图，包括：对哪些附加元数据参考了统计交换流程（数据供应方、结构、供应协议、数据流、元数据流等）的SDMX范围内有意义“对象”的机制；描述一系列有意义概念的机制；将其组成提示结构并显示其值如何表示的机制；该内容基于参考元数据的简单层次概念，该概

念对于很多元数据系统和分类/类别方案非常常用。SDMX为描述参考元数据结构并根据该结构报告参考元数据提供了模型（以及XML格式）。

SDMX标准也引入了与统计交换的流程内容有关的元数据的支持内容。可对逐步处理过程进行建模，并对提供数据和参考元数据的人员以及他们如何提供的信息进行表述，同时也可给出服务水平协议（与供应协议相同类型）的技术内容。

SDMX信息模型非常正式地描述了上述列出的所有内容，以给出统计交换流程的标准视图。

SDMX信息模型采用UML进行表示也同样以正式文体进行描述。尽管信息模型是非规范性的，但它是理解和使用规范性格式说明非常有价值的工具。

7 SDMX-EDI

SDMX-EDI格式是SDMX标准支持的一种格式：

- 统计定义：SDMX-EDI采用UN/EDIFACT格式表示SDMX信息模型覆盖的结构元数据；
- 统计数据：对当事双方大量时序数据的批量交换进行了优化动作，该动作考虑到了大型的整体或部分数据集的极致压缩表示形式。非时序数据也可支持，例如横截面数据，只要其作为重新包装后的时序表示，但这种格式不直接支持横截面数据；
- 数据集表：数据集及其结构元数据列表。

SDMX信息模型提供了用于SDMX-EDI的EDIFACT语法中给出的构造，以及SDMX-ML的XML语法中给出的构造。由于这两种语法实现方式反映了相同的逻辑构造，SDMX-EDI数据和元数据报文可被转换为相应的SDMX-ML格式，反之亦然。因此，这些标准提供了用于处理和交换统计数据和元数据的基于UN/EDIFACT和XML系统间的互操作性。

8 SDMX-ML

虽然SDMX-ML格式最初设计用于支持批量交换，但SDMX-ML支持更广泛的需求。XML格式用于很多不同类型的自动处理，因此，必须支持更加多样的处理场景。这就是为什么SDMX-ML格式有很多报文类型。每种报文类型用于支持一系列特定的处理需求。

- 结构定义：所有SDMX-ML报文类型共享一种元数据的常用XML表示形式来理解和处理数据集或元数据集，关于类别方案和组成方式的附加元数据也包含在内。同样，数据和元数据供应的结构内容（数据流和元数据流）也使用该格式进行描述。
- 通用数据：同结构定义报文的内容一样，以SDMX-ML表示的所有统计数据均能根据此数据格式进行标记。该数据设计为需要用单一格式对其进行处理的任一场景的数据供应。在进行处理前，这一应用程序可能需要数据集结构的独立访问权限。以这种格式标记的数据并非压缩型，但其可容易地获得数据集的所有方面。该格式不但提供数据集及其使用通用XML解析器的结构定义间的严格确认，并考虑到了部分数据集（增量更新）和完整数据集的传输，也支持时序和横截面用例。
- 特定结构数据：该格式专门针对数据集的数据结构定义（即特定DSD），且通过结构定义报文和技术规范格式定义的元数据构造间的映射关系生成。该数据支持XML格式（特别是当数据集中50%都以通用数据表示）的大型数据集的交换，通过通用XML解析器严格验证了同DSD的一致性，并考虑到了部分数据集（增量更新）和完整数据集的传输，也支持时序和横截面用例。

很多XML工具和技术对XML模式实现的功能充满期待，其中一个是XML模式中描述的XML构造和XML实例中标记的数据间直接的联系。强大的数据类型化也被认为是常规的，支持标记数据的完整验证。该报文类型支持验证和其他预期的XML模式功能。

- 通用元数据：所有以SDMX-ML格式表示的参考元数据均可根据此模式进行标记。该元数据仅进行少量确认并略微冗长，但其支持处理参考元数据的通用软件工具和服务的生成。
- 特定结构元数据：对于每个元数据结构定义，XML模式是用于创建这样的结构，以进行报告元数据集的验证。该结构不像通用元数据格式一样冗长，因为XML标记直接与报告的概念有关，适宜于专为处理某一特定类型元数据报告的应用程序。类似于使用XML数据的特定结构数据格式。
- 查询：数据和元数据通常在数据库中发布，这些数据库可在Web上获得。因此，必须具有标准查询文件，来允许对数据库进行查询并返回SDMX-ML数据、参考元数据或结构报文。查询报文是SDMX信息模型用于Web服务和数据库应用的实现方式，考虑到了将发送给使用该技术的数据供应方的标准申请。
- 注册：SDMX-ML界面支持所有可能的SDMX注册服务交互。该文件几乎全部基于报文的同步交换——由“响应”报文对“申请”报文进行回答。有两种基本类型的申请：“提交”，向注册服务书写元数据；“查询”，用于发现元数据。注册交互为所有类型的供应元数据提供格式，同时为订阅/通知、结构元数据和数据及元数据注册提供格式。非同步的（注册）通知报文是一个特例。

由于所有SDMX-ML格式采用相同信息模型且所有数据和元数据报文均来源于结构定义报文，该报文描述了数据集或元数据集，则每个类似格式间可能存在标准映射关系。这些映射关系能通过通用转换工具实现，对于所有SDMX-ML用户均非常有用，且并非专门针对特定数据集的数据结构定义或元数据集的结构定义。SDMX-ML部分包是特定数据格式、元数据格式和结构定义格式间映射关系的集合，结构定义格式是所有数据格式的来源。

9 一致性

本部分包含一个规范表述，阐述应用程序应有哪些功能以便同SDMX规范保持一致。这里强调了应用程序必须能够实现的功能，以及实施者根据SDMX一致性所作出的一致性表述内容。

10 对于SDMX内容导向指南的依赖性

为了同统计数据交换内容与语义联系更加密切的其他SDMX指南联合使用，设计了该技术标准。SDMX信息模型对任一统计概念同等地行之有效，但是为了鼓励互用性，其还必须对特定概念和术语的使用进行标准化和统一化。为实现该目的，SDMX将制定跨域概念、术语和结构的指南并对其进行维护。该成果主要有三部分内容。

10.1 跨域概念

SDMX跨域概念是关于统计领域交叉使用概念的内容指南。该列表预计将增长并随着SDMX应用领域数量的不断增加，该列表将进行修订。SDMX跨域概念的使用提供了促进SDMX标准概念间的互用性的框架。统计概念的协调性不仅包括概念的定义及其名称，也包括其标准代码表的表示形式以及其在数据结构定义和元数据结构定义中扮演的角色（适用时）。

SDMX跨域元数据概念指南的目的有2个：提供可用于结构统计数据和元数据的核心概念集，以促进系统（如上所述，“结构元数据”）间的互用性；通过对不同类型的元数据（如上定义，“参考元数据”）使用统一的概念名称和定义，以促进元数据更广泛的交换。

10.2 元数据常用术语

元数据常用术语是提供术语定义的SDMX指南,这些术语用于比较和映射在数据结构定义中以及统计元数据管理的其他方面发现的术语。特别是,该指南给出了GB/T 18391中大量统计术语的一致性定义,这些术语可被直接使用或对照可被映射的其他术语学系统。该术语集合包含SDMX技术标准中使用的术语,也给出了SDMX跨域元数据概念建立所基于的术语定义。

10.3 统计专题领域

统计专题领域是统计信息范围的列表,其目的是组织广泛的统计交换和分类。为促进互用数据和元数据交换,统计专题领域可作为对照不同当事双方可互相映射的类别方案的标准方案。统计专题领域的另一目的是允许组织有对应的“域组”,每个域组可定义其域中的标准数据结构定义、概念等。这些组在国际领域内已经存在。SDMX将使用统计专题领域列表来促进这些组制定各类实质标准,并提升这些标准在统计领域内和统计领域间的SDMX一致性技术系统的互用性。在SDMX中,这些方案的内容组织被支持为类别方案。SDMX统计专题领域由SDMX倡议提出并维护,并且应接受调整。

参 考 文 献

- [1] GB/T 19256.3 基于XML的电子商务 第3部分：消息服务规范
 - [2] GB/T 19256.4 基于XML的电子商务 第4部分：注册系统信息模型规范
 - [3] GB/T 19256.5 基于XML的电子商务 第5部分：注册服务规范
-