

ICS 35.240.01

A11

备案号:

JR

中华人民共和国金融行业标准

JR/T 0107.3—2014/ISO 17369-3:2013

统计数据 and 元数据交换 (SDMX) 第 3 部分: SDMX-ML 模式和文档

Statistical data and metadata exchange (SDMX)—
Part 3: SDMX-ML schema and documentation

(ISO 17369-3:2013, SDMX-ML: schema and documentation, 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 背景	1
2.1 概述	1
2.2 XML 的设计理念	1
2.3 促进 SDMX-ML 标准的使用	1
3 一致性	1
4 设计内容概述	2
4.1 要求范围	2
4.2 设计目的	3
4.3 SDMX-ML 包装方法：命名空间模块	3
5 相关文件	4
参考文献	6

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3部分。

本部分依据GB/T 1.1-2009规则起草。

本部分等同采用 ISO 17369-3:2013《统计数据和元数据交换（SDMX） 第3部分：SDMX-ML 模式和文档》。

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、刘彬、邓琳莹。

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、存流量性质、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。本标准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。

统计数据与元数据交换 (SDMX)

第3部分: SDMX-ML 模式和文档

1 范围

本部分规定了SDMX模式和文档,描述了背景、设计内容概述、非特定结构定义的通用模式、特定数据结构定义模式与元数据结构定义模式的XML规范性条款,并提供一系列XML方案和XML文件示例。

本部分适用于金融统计中数据和元数据的交换和共享。

2 背景

2.1 概述

本部分通过强调一组标准注册服务接口中的数据共享功能为所有“双边交换”、“网关交换”和“数据共享”模型提供更高层次的支持。SDMX系列标准通过集中研究与质量、方法论和其他相关问题的参考数据元以支持元数据交换和报告新的变化形式,提高了数据集和数据结构中数据元的提供能力,并为数据立方体提供更多支持。以往经验显示,现有结构型元数据和数据集结构的一些细微增添将有助于SDMX支持更多不同类型的统计信息。

2.2 XML的设计理念

在信息层,所有的文档类型共享“envelope”(SDMXMessage.xsd),是一组公共的低级别的组成部分(SDMXCommon[subset].xsd),因此所有信息的头文件信息和信息结构总是一致的。内容如下:

- 用于描述所有类型的结构元数据的图表和子模式,包括数据集(数据结构定义),元数据集(元数据结构定义),相关联的元数据和数据结构集;
- 登记交换数据所涉及的结构对象(SDMXStructure[subset].xsd);
- 数据共享、交换的属性数据模块(SDMXDataGeneric[subset].xsd);
- 特定属性的数据信息模块用于更新、版本修订和双向交换(SDMXDataStructured[subset].xsd);
- 用于调用网络服务的查询通用模块和子模块(SDMXQuery[subset].xsd);
- 注册界面的通用模块和子模块(SDMXRegistry[subset].xsd);
- 参考元数据集的通用模块(SDMXMetadataGeneric.xsd);
- 特有元数据结构定义模块(SDMXMetadataStructured.xsd)。

2.3 促进SDMX-ML标准的使用

除了上述不同的模式,为了从结构描述中创建特定关键字族模式,本部分开发了标准映射和相应转换工具,把XML数据实例从一个XML数据描述模式转化为另一个XML数据,并且将它们转化为相应的SDMX-ML报文。

免费工具在SDMX早期开发时期发挥了重要的推动作用,推动使用者在处理过程中更加便捷地使用数据。在满足SDMX使用者的不同需求,实际操作中所有格式都将共享通用信息模型。

3 一致性

本部分的第5章中提到的SDMX_2_1_SECTION_03A_PART_IV_DATA是规范性文档，为建立一致的SDMX-ML XML例子与W3C XML模块提供了规则。

4 设计内容概述

4.1 要求范围

通过了解以下XML设计要求，可以帮助理解不同文件类型之间的关系：

- 当数据库潜在规模发生交换时，必须使用合理的压缩格式获取大量数据；
- 发布数据库数量变化通知时无须重复发送完整的数据库。此外，对数据交换需要进行确认时不需要发送完整的交换数据集，只需传送完整的信息以确保准确升级和修订处理；
- 结构型信息及数据需要进行传输；
- 使用正确的转换方式转化GESMES/TS EDIFACT语法；
- 允许使用多种等效语言表达自然语言信息；
- 发送请求信息、数据和结构至信息源以支持网络服务和类似技术性服务；
- 使用者（和注册服务）可以不了解具体的关键字族，但可以处理关键字族内的数据，甚至（例如，对照性服务）可以根据单个XML实例中的多个关键字族进行数据结构化处理；
- 简单化XML以满足不熟悉统计领域的网络管理员和开发人员的使用需求；
- 规范化XML标准工具（如：网络开发环境、解析器、引导编辑性工具等）内的XML；
- 为满足不同用户的需要，结构化数据不仅可以作为时间序列数据使用，还可以作为横截面数据使用。数据可以依照单个关键字族进行数据结构化，并且转换为一种可结构优化的标准格式；
- 在不需要对所有的结构进行了解的情况下，数据结构和特定元数据结构格式应该能进行通用的处理过程。也就是说，特定结构模块应该来验证此类信息，但并不需要对其进行处理。
- XML格式可提高通用语法、概念和代码表的可重用性，同时作为识别某特定资源（代码表、数据结构定义、数据集等）维护的代理机构；
- XML格式在标准接口基础上连同标准注册服务共同支持应用程序的相互交流，同时还具备网站服务器和通过网址、类似协议服务运行的职能；
- XML格式必须支持没有被结构化的参考元数据报告，但它的构成连接了其他部分的统计集合、报告、处理、交换和分布的元数据主信息流。XML格式还必须支持和可配置的质量计划、方法元数据、管理元数据与类似元数据报告；
- XML格式通过结构间的映射概念和代码表描述元数据集和数据集群体之间的关系，并在影射的基础上通过非单一性、相关结构性定义的方式查询数据和元数据；
- 允许在数据结构内使用与时间观测值无关的时间关联概念；
- 用于非模糊性数据序列/观测值时允许使用数据结构定义中简单、未编码的递增性标识符，所使用的标识符在观测时间周期内并非为简单的一一对应关系；
- 允许未编码标识符和描述符与已建立了外部实体或标识符的数据结构定义相关联，用于消除其他序列/观测值（例如：当数据集描述一组代理机构或一组账户时可能产生同样的关键值）之间的歧义；
- 允许非数字型观测值进行编码；
- 允许“立方体”基础性系统（如OLAP）与一般性系统相互操作，避免丢失高级系统中的丰富元数据。

以上设计要求所涉及的内容具有广泛性，并且观察后可以发现某些要求之间存在一定矛盾。如无法满足单一功能（如数据交换、参考元数据交换、查询等）而设计出某个XML文件类型等。无论采用何种设计都需要明确与信息模型之间的关系。

4.2 设计目的

XML模块的基本设计从信息模型开始，而信息模型的建立也是根据其结构元数据所在的抽象类进行的。这就意味着如果更多的XML高级工具能够了解这种结构的本质，那么它们能使用这些设计轻松地将XML实例转换为基于模型的对象。

另外一个基本的方法是尽量使结构保持连续，这样的话用来运行XML的工具软件就可以多次利用。除了识别和版本等一些基础的信息模型特征会使用这种方法外，参考和询问等其他一些领域也开始使用这种方法了。

在标准化年代，典型的用户一般都使用CompatData Message或者GenericData Message.UtilityData很少被使用，而CrossSectionalData由于其定义的模糊性，通常会导致格式的不一致。因此，这个版本的设计方法就是为了能够尽量多的整合不同结构的信息，提高信息资源的利用率。总的来看，可以归纳为两种基本的数据模式：一般模式和特定数据结构定义模式，而且这些数据模式都已经被充分的结构化。CompatData Message和CrossSectionalData也都被整合为同一种数据模式。这种数据模式非常灵活且具有一定的弹性，适合不同角度的观察。为了使CompatData更加兼容，本版本中还创设了时间序列变量，并且保证这些时间序列变量可以像普通数据模块一样加工，无需引入其他补充变量。

同样的设计理念也运用于元数据参考信息。可以生成不同的数据和元数据结构的基础的特定结构模式已经构建成功，这样它们可以以普通的方式来处理特殊的结构。

本版本的另外一个改动是移除了大部分的网络查询功能，不同于适用于任何一些数据查询或者结构元数据查询的功能，本版本的目的是建立一个专属于查询最合适的代码表的数据库。

4.3 SDMX-ML 包装方法：命名空间模块

在提议的XML模式设计中，XML命名空间作为模块使用的构想为包装方案的基础，所有指定用户或申请者只需要为达到使用目的而熟悉整个文件库内的某个子集。这种方法不但适合前文3.2的内容，也经常用于其它领域的主要XML标准。

命名空间的另一个优点是，可以要求维护代理机构跟踪维护中的潜在可重用结构和数据，因此SDMX可以拥有自己的特定命名空间模块，并且允许其他维护代理机构根据所维护的特定关键字族或元数据结构拥有自己的命名空间。

根据以上的设计思路，设计出了一系列的命名空间软件包。为了方便使用，一些大的计划又分别被划分为了归属于同一命名空间的一组子模块，每个模块或者子模块都是W3C XML语言元素中单独的一个实体。当不同的模块或者亚模块之间有交互作用时，可以使用XML的引入机制来描述其他模块或子模块的结构。除此之外，命名空间模块本身还反映了不同的关系类型。举例来说，所有的数据结构都在基于http://www.sdmx.org/resources/sdmxml/schemas/v2_1/data的一个命名空间模块中。这种命名空间之间的关系同时也反映了不同结构之间的关系。

下面的这些例子具体描述了XML模式，并且都可以被其后所附的文件运行支持。这些可以划分为类属计划和数据结构特异性计划，前者可以提供完整的一系列计划定义，后者提供了核心结构和特定数据结构同核心结构之间的映射方法。内容如下：

- SDMX命名空间模块由共同的通用报文结构组成，包含相同的报文头信息，需要与其他所有SDMX-ML命名空间模块一起使用（SDMXMessage.xsd）；
- SDMX命名空间模块包含结构元数据中的描述内容，如数据结构定义、概念和代码表（SDMXStructure.xsd）；

- SDMX命名空间模块包含了例如关键字族、概念、代码表等一些结构元数据 (SDMXStructure [sub-module].xsd);
- SDMX命名空间模块包括了所有通用SDMX信息类型,这是所有其他SDMX-ML命名空间模型所需要的;
- SDMX命名空间模块为格式化数据 (SDMXGenericData.xsd) 描述通用 (非特定关键字族) 格式;
- SDMX命名空间模块描述了数据的通用格式 (SDMXDataGeneric.xsd);
- SDMX命名空间模块构造数据 (包括特定结构的数据);
- SDMX命名空间模块可以描述查询信息的结构 (SDMXQuery[sub-module].xsd),为网络支持和用户提供服务;
- 一系列的命名空间模块是由特定数据结构模块构建并维持,并非由SDMX维持;
- SDMX命名空间模块为参考元数据的描述提供了通用格式,这并未考虑到元数据结构的定义 (SDMXMetadataGeneric.xsd);
- SDMX命名空间模块为所有的特定结构元数据模块提供了通用的框架结构,这些都用于描述参考元数据 (SDMXMetadataStructured.xsd);
- 双边使用数据时,SDMX命名空间模块向所有特定关键字族模式提供通用框架,以用于数据库的交换、更新和修正 (SDMXCompactData.xsd);
- SDMX不负责维护模块时,SDMX中的命名空间模块将由特定关键字族“压缩”模式的创建者负责建立和维护;
- 处理和发布产品用途时,SDMX命名空间模块向所有特定关键字族模式提供通用框架,以便网站管理人员和开发人员使用标准的XML工具 (SDMXUtilityData.xsd);
- SDMX不负责维护模块时,SDMX中的命名空间模块将由特定关键字族“实用程序”模式的创建者负责建立和维护;
- 双边使用和处理横截面数据时,SDMX命名空间模块为横截面数据处理,向所有特定关键字族模式提供通用框架 (SDMXCrossSectionalData.xsd);
- SDMX不负责维护模块时,SDMX中的命名空间模块由特定关键字族“横截面”模式的创建者负责建立和维护;
- SDMX命名空间模块向参考元数据报告提供通用格式,但无需考虑元数据结构定义 (SDMXRefMetadata.xsd);
- SDMX命名空间模块为参考元数据报告,向所有特定元数据结构模式提供通用框架 (SDMXMetadataReport.xsd);
- SDMX不负责维护模块时,SDMX中的命名空间模块由特定元数据结构“元数据报告”模式的创建者负责建立和维护;
- SDMX命名空间模块为一系列的注册服务的交互作用提供了一个标准化的接口 (SDMXRegistry[sub-module].xsd)。

上述部分详细描述了本部分所建议的XML格式,需要结合所提供的文件档案共同检测。所建议的模式可分为通用模式和特定关键字族模式。通用模式可提供完整的模式定义组;特定关键字族模式可结合模式代码提供核心结构。此外,所建议模式还可提供由特定关键字族或元数据结构定义映射核心结构的方法指南。

5 相关文件

所有的相关文件大致可以包括以下的七大类:

- a) SDMX_2_1_SECTION_03A_PART_I_MESSAGE:包含了信息命名空间模块的描述;

- b) SDMX_2_1_SECTION_03A_PART_II_COMMON:包含了常用命名空间模块的详细信息;
- c) SDMX_2_1_SECTION_03A_PART_III_STRUCTURE:包含了命名空间结构模块的详细信息;
- d) SDMX_2_1_SECTION_03A_PART_IV_DATA:详细介绍了数据和元数据命名空间,以及构建特定结构模块的规则;
- e) SDMX_2_1_SECTION_03A_PART_V_QUERY:包括了查询命名空间模块的详细信息;
- f) SDMX_2_1_SECTION_03A_PART_VI_REGISTRY:包括了注册命名空间模块的详细信息;
- g) SDMX_2_1_SECTION_03A_PART_VII_SAMPLES:为现行简单的文件提供了简要的解释。
(相关文件的获取请参考www.sdmx.org或联系全国金标委jinbiaowei@pbc.gov.cn)

参考文献

- [1] W3C XML 模块定义语言, 版本 1.0. <http://www.w3c.org/XML/Schema#dev>, World Wide Web Consortium. 2004-10-28
- [2] W3C拓展的标记语言, 版本 1.0 第三版
<http://www.w3c.org/TR/2004/REC-xml-20040204/>, World Wide Web Consortium. 2004-02-04
- [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 18391.2-2009 信息技术 元数据注册系统 (MDR) 第2部分 分类
- [5] GB/T 18391.3-2009 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
- [6] GB/T 18391.4-2009 信息技术 元数据注册系统 (MDR) 第4部分 数据定义的形成
- [7] GB/T 18391.5-2009 信息技术 元数据注册系统 (MDR) 第5部分 命名和标识原则
- [8] GB/T 18391.6-2009 信息技术 元数据注册系统 (MDR) 第6部分 注册
- [9] GB/T 7408-2005 数据元和交换格式 信息交换 日期和时间表示法
-