ICS 03.060
A11

JR

中华人民共和国金融行业标准

JR/T 0176.1—2019

证券期货业数据模型
第1部分：抽象模型设计方法

Securities and futures industry data model—
Part 1：Abstract model design method

2019-11-18 发布　　2019-11-18 实施

中国证券监督管理委员会　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　总体设计架构 3
5　整体设计方法 4
6　公共部分设计方法 4
7　交易部分设计方法 4
7.1　概述 5
7.2　行为及过程梳理 5
7.3　数据项梳理 8
7.4　数据模型设计 8
7.5　模型验证 10
8　监管部分设计方法 10
8.1　概述 10
8.2　方法描述 10
8.3　步骤简述 10
8.4　模型验证 11
9　信息披露部分设计方法 11
9.1　概述 11
9.2　模板（Template） 11
9.3　映射（Reflect） 12
9.4　提炼（Refine） 12
9.4.1　通用维度的设计 13
9.4.2　原子数据梳理 13
9.4.3　复合数据梳理 14
9.4.4　可复用数据表梳理 14
9.5　重组（ReCombine） 15
9.6　回归（Regress） 15
9.7　步骤简述 15
10　元语定义 15
10.1　各层次数据元语定义 15
10.1.1　数据元语定义 15

10.1.2 表元语定义.....17
10.2 代码元语定义.....19
11 产出物说明.....20
附录A （规范性附录） 数据类型.....21
附录B （规范性附录） 引用类别.....22
参考文献.....23

前　言

JR/T 0176《证券期货业数据模型》分为8个部分：

--第1部分：抽象模型设计方法；

--第2部分：逻辑模型公共部分 行业资讯模型；

--第3部分：证券公司逻辑模型；

--第4部分：基金公司逻辑模型；

--第5部分：期货公司逻辑模型；

--第6部分：证券交易所逻辑模型；

--第7部分：期货交易所逻辑模型；

--第8部分：监管机构逻辑模型。

本部分为JR/T 0176的第1部分。

本部分依据GB/T 1.1-2009给出的规则起草。

本部分由全国金融标准化技术委员会证券分技术委员会（SAC/TC180/SC4）提出。

本部分由全国金融标准化技术委员会（SAC/TC180）归口。

本部分起草单位：中国证券监督管理委员会信息中心、中国证券监督管理委员会证券基金机构监管部、中证信息技术服务有限责任公司、中国期货市场监控中心有限责任公司、申万宏源证券有限公司、中证机构间报价系统股份有限公司、中软国际科技服务有限公司、深圳市致远速联信息技术有限公司、上海吉贝克信息技术有限公司、上海立信维一软件有限公司。

本部分主要起草人：张野、刘铁斌、周云晖、谢晨、罗黎明、孙宏伟、黄璐、汪萌、张春艳、王辉、曹雷、陈楠、富子祺、乔蔚、黄文璐、朱旭、刘佳、李光涛、刘国勇、杨诚、李婷婷、李海、杨洪峰。

引　言

证券期货业数据化程度相对较高，机构多、类型广、交易方式多样，机构内及机构间数据交换频繁、业务发展迅速，为提高数据交换效率、规范行业机构数据应用系统建设、提升行业数据标准化水平，证券期货行业组织开展了行业数据模型建设工作，旨在清晰描述整个市场的数据流向、数据名称、数据定义、结构类型、代码取值和关联关系等，为行业机构内部系统建设和机构间数据交换提供指导。本部分是证券期货业数据模型系列标准的第一部分：抽象模型设计方法，基于此方法能够形成一套符合监管规范的模型框架，以及一套依据监管规则提炼形成的行业数据字典。相关成果是行业数据模型的重要组成，是行业标准的数据审核依据，是行业逻辑模型的映射基础，对于规范行业数据语言、推进行业数据治理、辅助行业监管科技建设等都具有十分重要的意义。

证券期货业数据模型
第1部分：抽象模型设计方法

1 范围

本部分规定了证券期货业抽象模型设计方法，包括总体设计架构、整体设计方法、公共部分设计方法、交易部分设计方法、监管部分设计方法、信息披露部分设计方法、元语定义以及产出物说明。

本部分适用于证券期货行业抽象模型建设工作。

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。凡是注有日期的引用文件，仅所注日期对应的版本适用于本文件。凡是未注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

《企业会计准则》财政部会计司

3 术语和定义

下列术语和定义适用于本文件。

3.1 业务

3.1.1

抽象模型 abstract model

以证券期货行业相关法律法规、部门规章、业务规则、指导性文件等为依据，采用自顶向下的梳理方法，对资本市场各类业务活动进行遍历，对关键业务流程和数据要素进行识别而形成的一系列具有通用性、稳定性和扩展性的数据集合。

3.1.2

主体 identity

参与资本市场业务的主要对象。

注：包括机构、个人及产品。

3.1.3

品种 variety

参与资本市场业务的各类金融工具及服务等。

3.1.4

财务 finance

各类机构参与资本市场业务过程中涉及的资金活动及资金关系。

3.1.5

交易 trading

各类主体参与资本市场业务所涉及的转让行为。

3.1.6

监管 regulation

主管部门依照法律法规，统一监督管理全国证券期货市场、维护市场秩序、保障其合法运行的行为。

3.1.7

信息披露 information disclosure

披露主体将其自身的财务变化、经营状况等信息和资料向监管部门和交易所报告，并向社会公开或公告的行为。

3.1.8

行为 behavior

各类市场主体在交易、监管、信息披露过程中的一系列动作。

3.1.9

过程 process

行为发展所经过的环节。

3.1.10

监管对象 regulation object

主管部门依法负责监管的对象。

3.1.11

监管主题 regulation theme

对监管对象的某一方面或某一领域的监管职责。

3.1.12

监管方式 regulation method

对监管对象的某个监管主题所采用的监管手段或措施。

注：包含现场检查、行政许可等。

3.1.13

监管业务 regulation business

对监管对象的某个监管主题进行监管的具体工作内容。

注：包括法规制定、行政许可、稽查处罚、日常监管、宏观监管和内部管理。

3.1.14

监管过程 regulation process

对监管对象的某个监管主题采用某种监管方式进行监管的过程。

3.2 数据

3.2.1

元语 meta semantic

数据的描述信息。

3.2.2

实体关系图 entity-relationship diagram

以图形方式描述数据模型实体关系的一种方法。

注：实体关系图简称ER图。

3.2.3

数据字典 data dictionary

对标准中涉及的各类数据进行定义或描述的数据集合。

3.2.4

原子数据 atomic data

独立的、不可再分的最小数据单元。

3.2.5

复合数据　compound data

由原子数据通过某种关系组合或加工而成的数据。

3.2.6

可复用数据表　reusable table

由原子数据或复合数据组成的，在多个业务环节中被共同使用的数据集。

3.2.7

语义独立的数据表　semantic specified table

由原子数据或复合数据组成的描述某一特定业务逻辑的数据集。

3.2.8

驼峰命名法　camel-case

混合使用大小写字母为变量和函数命名的方法。

4　总体设计架构

抽象模型遵循“1+3+N”的总体设计架构如图1所示，其中：

——“1”是指抽象模型的行业公共部分，包括主体、品种、财务三个主题及其他公共部分。行业公共部分可以被“交易”、“监管”、“披露”中至少两条业务主线复用，且自身保持语法、语义一致。

——“3”是指以“交易”、“监管”、“披露”三条业务主线分别梳理形成的模型，三组模型彼此耦合度不高，组间个性化部分允许出现同名不同义的冲突情况，不要求在语法语义上保持一致。每组模型再分为该业务范畴内的公共部分及覆盖其下应用系统的专用数据模型。其中公共部分可以在该业务范畴内被共用，应在组内保持语义、语法一致。

——“N”是指在每个业务主线梳理形成的模型中，特定应用系统的专用数据模型，该部分的数据和表为某一应用特有的，只在该应用中被使用，与业务主线中其他应用的模型不复用。每个“N”中涉及的数据、表在语法、语义上不要求一致。

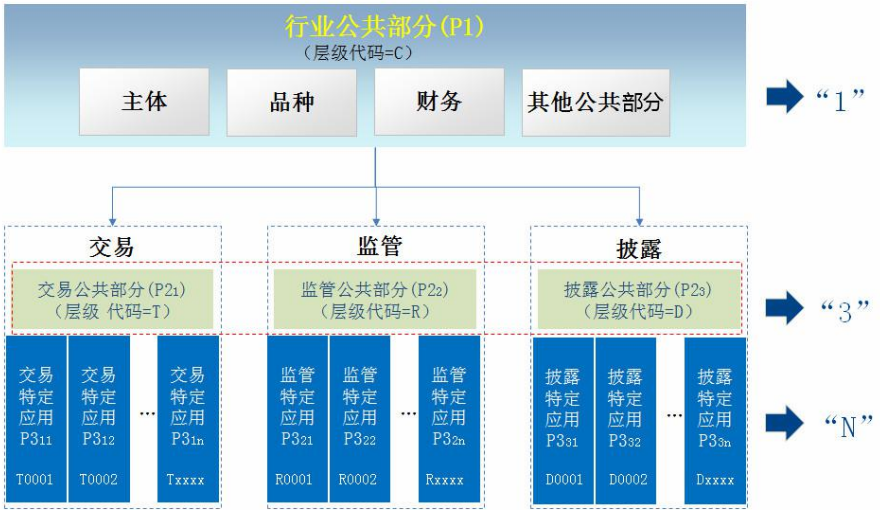


图 1　“1+3+N”设计架构

“1+3+N”的设计模式可以表达为：

$$SDOM = P1 + \sum_{i=1}^{3}\left(P2_i + \sum_{j=1}^{n_i} P3_{ij}\right)$$

式中：

SDOM ——抽象模型；

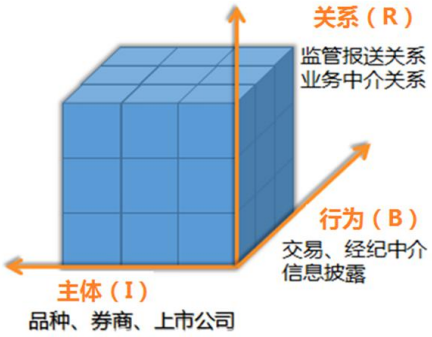
P1 ——∈{主体，品种，财务，其他公共部分}；

$P2_i$ ——{交易公共部分，$i = 1$；监管公共部分，$i = 2$；披露公共部分，$i = 3$}

P3ij ——j代表每个业务主线下，特定应用系统的专用数据模型。

5 整体设计方法

抽象模型整体采用“主体-行为-关系”（Identity-Behavior-Relevance，简称IBR）设计方法进行设计，即以主体（Identity）为核心，外延其各类市场行为（Behavior），并总结其中的相关关系（Relevance），抽取、划分形成一系列针对特定业务场景和应用的数据表，最终形成数据模型总体框架。具体内容如图2所示。



		交易	监管	信息披露
I	主体	品种	监管对象	披露主题 披露主体
B	行为	行为 过程	监管主题 监管方法	披露
R	关系	业务流 程关系	监管 报送关系	披露树的关系

图 2 IBR 方法论

6 公共部分设计方法

抽象模型的公共部分主要包括品种、主体、财务三个主题，具体设计方法如下：

a) 品种主题。参照GB/T 35964-2018，结合国内资本市场实际情况调整、补充品种分类体系，经过补充实体、属性、代码等进一步细化，形成品种主题模型。

b) 主体主题。结合资本市场实际，依据最新发布的法规文件，梳理形成的主体分类体系，并补充实体、属性、代码等，形成主体主题模型。

c) 财务主题。依据《企业会计准则》，结合资本市场在信息披露、交易和监管领域的实际应用梳理形成财务报表体系，并补充实体、属性、代码等，形成财务主题模型。

7 交易部分设计方法

7.1 概述

抽象模型交易部分采用的是“证券品种-过程-行为”（Securities-Process-Behavior，简称SPB）设计方法，首先根据公共部分的品种主题筛选待梳理的品种，其次根据各品种梳理业务主干流程，对流程图中各个交易环节进行归纳和抽取，整理出行为列表；将各个行为按照交易前、中、后的各个过程进行切分，形成各品种的行为过程二维图。针对行为过程二维图中行为，梳理业务流程的环节和各环节的产出，形成业务流程的泳道图；细化业务流程泳道图中的产出物涉及的数据项，明确数据项的定义及属性信息。将数据项按主题归并提炼形成数据对象，数据对象及其之间的关系以实体关系图的形式表达，形成理论数据模型。最终通过现实系统的数据模型对理论模型进行回归验证及完善。具体梳理步骤如图3所示。

图 3 SPB 梳理方法

7.2 行为及过程梳理

针对每个品种所涉及到的业务主干流程进行梳理，以融资融券为例，具体行为及过程梳理步骤如下：

a) 按照投资者融资买入、投资者融券卖出、证券公司提供服务三条主干流程进行分析，形成三大主干流程图，梳理结果示例如图4所示。

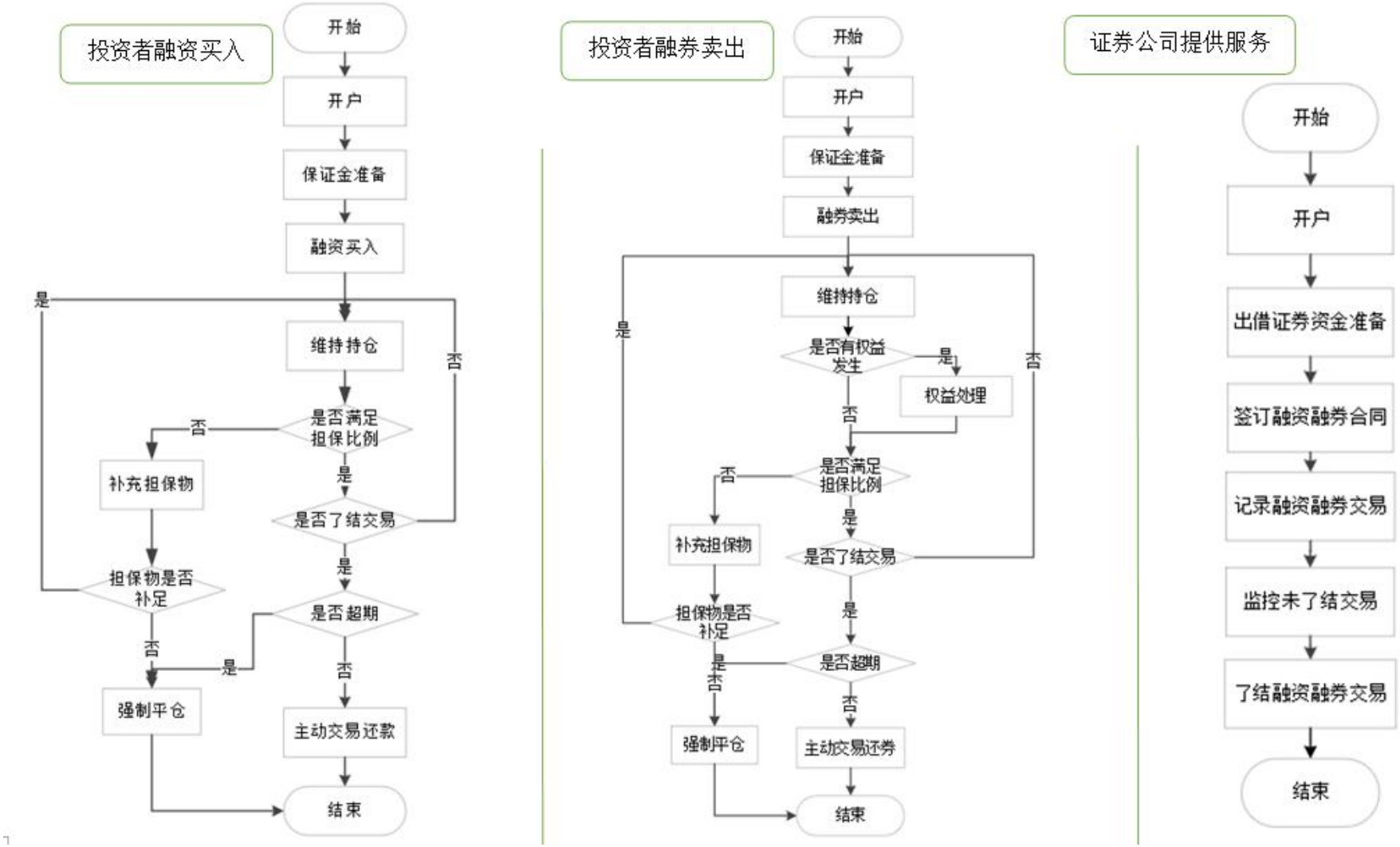

投资者融资买入
开始
开户
保证金准备
融资买入
维持持仓
是否满足担保比例
否
是
补充担保物
是否了结交易
是
担保物是否补足
是否超期
否
是
否
强制平仓
主动交易还款
结束
是
否
投资者融券卖出
开始
开户
保证金准备
融券卖出
维持持仓
是否有权益发生
是
权益处理
否
是否满足担保比例
否
是
补充担保物
是否了结交易
是
担保物是否补足
是否超期
是
否
否
强制平仓
主动交易还券
结束
是
否
证券公司提供服务
开始
开户
出借证券资金准备
签订融资融券合同
记录融资融券交易
监控未了结交易
了结融资融券交易
结束

图 4　主干流程样例

b) 对主干流程图中的各个交易环节进行归纳和抽取，整理出各品种涉及的行为列表；在归纳和抽取过程中需保证每个行为的独立性，行为与行为之间不应存在包含关系。梳理结果示例如图5所示。

交易全过程 \ 行为		开销户	资金准备	信用交易					信用非交易						强制平仓	维持担保比例	权益处理
				融资买入	融券卖出	买券还券	卖券还款	信用账户普通交易	直接还款	直接还券	担保证券划入	担保证券划出	券源划转	余券划转			
投资者融资买入	开户																
	保证金准备																
	融资买入																
	维持持仓																
	补充担保物																
	强制平仓																
	主动交易还款																
投资者融券卖出	开户																
	保证金准备																
	融券卖出																
	维持持仓																
	权益处理																
	补充担保物																
	强制平仓																
	主动交易还券																
证券公司提供服务	开户																
	出借证券资金准备																
	签订融资融券合同																
	记录融资融券交易																
	监控未了结交易																
	了结融资融券交易																

图 5　行为列表样例

c) 对每个行为按照交易前、交易中和交易后的时序进行切分，识别出各行为涉及的基本操作，最终形成该品种的行为过程二维矩阵图。梳理结果示例如图6所示。

融资融券			开销户	资金准备	信用交易					信用非交易						强制平仓	维持担保比例	权益处理
					融资买入	融券卖出	买券还券	卖券还款	担保证券普通交易	直接还款	直接还券	担保证券划入	担保证券划出	券源划转	余券划转			
交易前	信用账户准备		√															
	保证金准备			√														
交易中	委托				√	√	√	√	√									
	成交				√	√	√	√	√									
交易后	登记	申请								√	√	√	√	√	√	√		√
		过户			√	√	√	√	√		√	√	√	√	√	√		√
	结算	清算	√		√	√	√	√	√	√	√	√	√	√	√	√	√	√
		交收	√		√	√	√	√	√	√	√	√	√	√	√	√	√	√

行为

过程

图 6　行为过程二维图样例

d) 在行为过程二维矩阵图形成后，参照相关业务流程及规则定义梳理出具体的细化流程图，便于后续进行数据项梳理及提炼。梳理结果示例如图7所示。

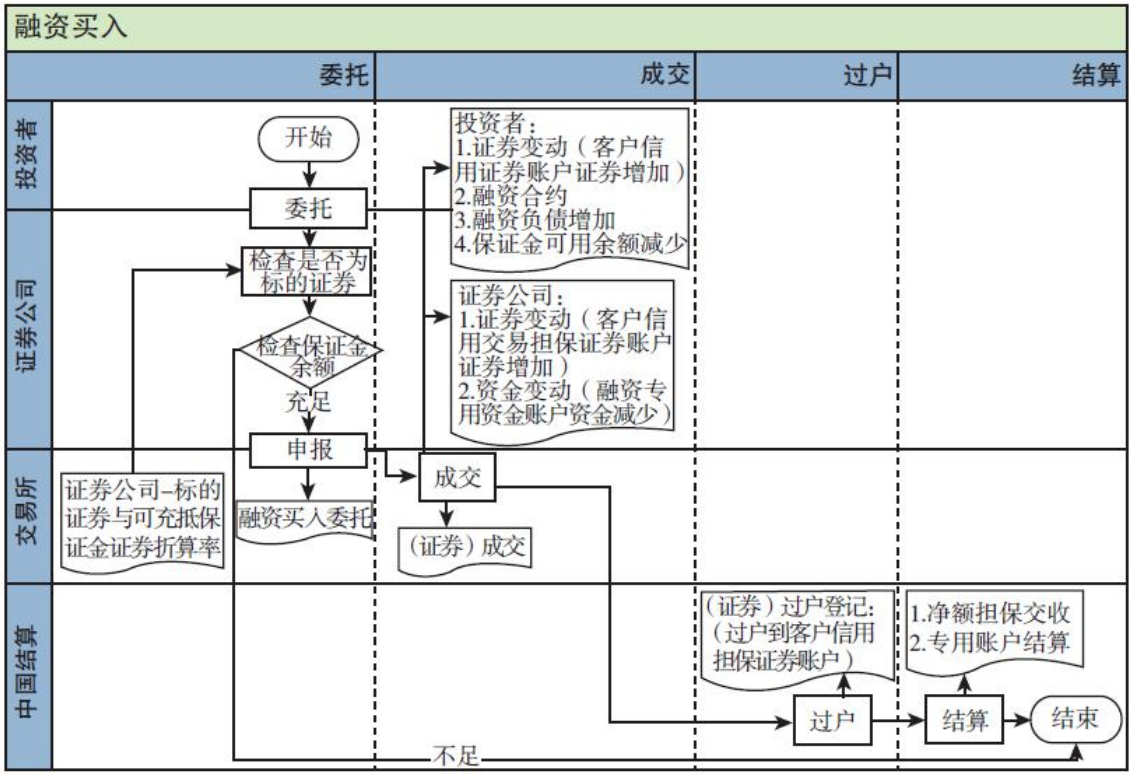


图 7　业务流程图样例

7.3　数据项梳理

参照业务流程图，对每个业务环节梳理和提炼数据项，并根据数据项的业务特征进行分类。梳理结果示例如图8所示。

品种	行为	过程	业务环节	业务数据类	数据项中文名称	数据项英文名称	数据层次	复合类型	详细定义	取值与公式	是否引用代码表	数据类型	数据长度	参考资料	备注
融资融券	融资买入	成交	成交	(融资融券)成交	证券成交编号		原子数据								
				(融资融券)成交	成交日期		原子数据								
				(融资融券)成交	成交时间		原子数据								
				(融资融券)成交	证券交易代码		原子数据								
				(融资融券)成交	买方证券账户编码		原子数据								
				(融资融券)成交	买方证券委托编号		原子数据								
				(融资融券)成交	证券交易单元标识		原子数据								
				(融资融券)成交	卖方证券账户编码		原子数据								
				(融资融券)成交	卖方证券委托编号		原子数据								
				(融资融券)成交	成交价格		原子数据								
				(融资融券)成交	成交数量		原子数据								
				(融资融券)成交	成交金额		原子数据								
融资融券	融资买入	成交	成交	融资买入合约	融资开仓单据号码		原子数据								
				融资买入合约	发生日期		原子数据								
				融资买入合约	发生时间		原子数据								
				融资买入合约	证券公司编码		原子数据								

图 8　数据项提取样例

7.4　数据模型设计

数据模型设计步骤如下：

a)　针对全部业务流程中梳理出的数据项分类，按照合并同类项的原则进行聚合，形成独立的数据实体；

b)　根据业务的关联度，构建实体与实体之间的关系；

c)　以实体关系图的形式对模型进行描述和表达。

模型设计结果示例如图9所示。

说明：

——● 标识关系，当一个外键从父实体迁移到子实体的主键区域，在两个实体之间就形成了一个标识关系，标识关系用连接两个实体间的实线来表示。

◇----● 非标识关系，当一个外键从父实体迁移到子实体的非主键区域，在两个实体之间就形成了一个非标识关系，非标识关系用连接两个实体间的虚线来表示。

图 9　融资融券实体关系图示例

7.5 模型验证

以现有的业务系统作为回归验证的对象，检查模型对业务系统全部数据范围的涵盖情况。主要考察模型符合实际业务的情况，包括对现有业务的覆盖情况和对未来业务的兼容性，通过实际系统对理论模型进行验证。

8 监管部分设计方法

8.1 概述

监管部分主要采用“监管主题/监管方式”（Theme/Method，简称T/M）方法，以相关法律法规为基础，遵循有法可依的原则，构建监管模型。

8.2 方法描述

依照法律法规，对监管对象（如会计师事务所）、监管主题（如获取证券期货业资格）、监管方式（如行政许可）进行三维分析，并在此基础上梳理出对应监管流程及相关信息，从而得出监管模型。

监管数据梳理空间=T×M，其中T=I×B×V

其中，T（Theme），监管主题

M（Method），监管方式

I（Identity），监管主体

B（Business），监管业务

V（Variety），品种

T/M方法的原理如图10所示。

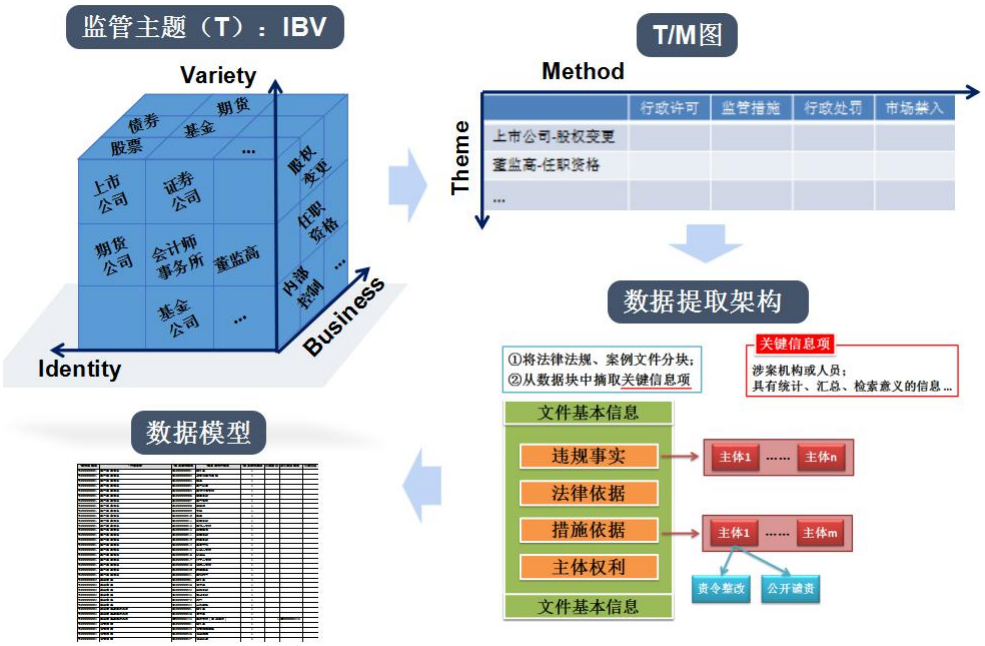


图 10 T/M 方法

8.3 步骤简述

T/M梳理方法用四个步骤梳理出监管模型：

a) 参考行业抽象模型主体部分梳理监管主体（I），针对每个主体梳理相应监管业务（B）及涉及品种（V），I×B×V 交汇的点即为监管主题（T）；

b) 按事前、事中、事后梳理监管方式（M）；

c) T 与 M 正交展开遍历，针对 T×M 的每个项，根据法律法规及详细案例梳理监管流程，并依据数据提取架构找出每个流程相关数据集；

d) 根据模型设计方法，按每个监管方式，提取公共的流程和公共的数据表、数据项。

8.4 模型验证

考察监管部分模型符合实际业务的情况，包括对现有业务的覆盖情况和对未来业务的兼容性。

以数个有代表性的业务为例，对模型进行检验，检查该业务在模型中的体现情况，业务中涉及的数据项、数据对象的整理提炼情况。

9 信息披露部分设计方法

9.1 概述

信息披露部分主要采用T4R（Templete Reflect Refine Recombine Regress，简称T4R）方法。该方法是采用基于信息披露模板的模型梳理方法，包括模板（Templete）、映射（Reflect）、提炼（Refine）、重组（Recombine）、回归（Regress）等五个重要环节。其中模板是信息披露的基础，映射基于模板制定出语义独立的数据表，进一步提炼出原子、复合数据及可复用数据表，重组后进行逻辑推演和测试验证，经多次迭代后形成最终的披露信息模型。各环节的关系如图11所示。

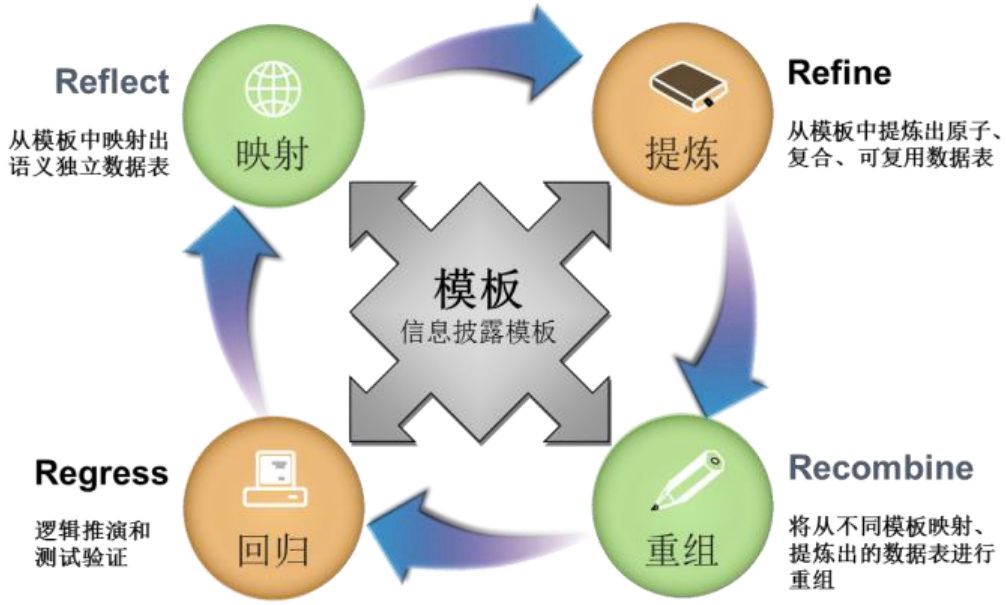


图 11 披露模型各环节关系

9.2 模板（Template）

模板是信息披露内容的具体体现，构建模板，首先以统一的主体和品种分类为依据，结合主体和品种运行周期，分别整理归纳和信息披露有关的法律法规，将梳理出的信息披露业务规则和要求进行规范后形成的文档即为信息披露模板，如图12所示。

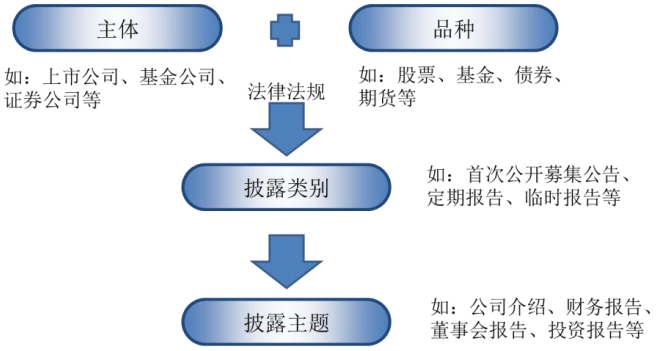


图 12　信息披露模板

模板用于全程指导后续的信息披露流程的各个环节应用。通常情况下，一种信息披露业务对应一个信息披露模板，模板由标题、段落、文字和表格构成，既对业务科目做出了定义，也对业务科目产生的场景进行了描述。

模板中的业务科目和表格可被复用，业务含义相同的科目和完全一致的表格，按照先提出先定义的原则，只定义一次，后续定义的模板只对此科目进行引用，而不进行重复定义。模板是信息披露模型梳理和创建的起点和依据，通过后续的建模方法，完成信息披露模板和标准到信息模型的转换。

9.3　映射（Reflect）

在选定信息披露模板后，对模板的结构进行梳理，定义出语义独立数据表，以及数据表间的多层次构成关系。如：一般企业上市公司年报包括董事会报告、监事会报告、财务报告等，其中财务报告包括审计报告、财务报表、财务报表附注等。

语义独立数据表可面向应用进行建模指导，它可由其他语义独立数据表、可复用数据表、复合数据、原子数据共同组成。以一般企业合并财务报表为例，根据业务梳理出的具体组成，如图13所示。

语义独立数据表编码	语义独立数据表中文名称	成员数据编码	成员数据中文名称	成员数据层次	数据重定义说明
D4FR101100	财务报表（一般企业合并）	D4FR101110	资产负债表（一般企业合并）	4	沿用
D4FR101110	资产负债表（一般企业合并）	C3FR100010	资产负债表	3	沿用
D4FR101110	资产负债表（一般企业合并）	C3FR100050	报告信息表	3	沿用
D4FR101110	资产负债表（一般企业合并）	D1FR000123	报告数据日期	1	沿用
D4FR101110	资产负债表（一般企业合并）	D1FR000084	资产负债表一般企业合并标识	1	沿用
D4FR101110	资产负债表（一般企业合并）	D2FR000001	货币资金	2	沿用
D4FR101110	资产负债表（一般企业合并）	D2FR000002	结算备付金	2	沿用
D4FR101110	资产负债表（一般企业合并）	D2FR000003	拆出资金	2	沿用
D4FR101110	资产负债表（一般企业合并）	D2FR000005	应收票据	2	沿用
D4FR101110	资产负债表（一般企业合并）	D2FR000006	应收账款	2	沿用
D4FR101110	资产负债表（一般企业合并）	D2FR000007	预付款项	2	沿用

图 13　一般企业合并财务报表组成样例

9.4　提炼（Refine）

9.4.1 通用维度的设计

提炼原子数据的工作，是在每个模板内根据披露信息的业务含义，提炼出不需要再分的最小数据单元。

在信息披露的规范要求中，对于数据的报表类型、数据状态、分级类别等背景信息，采用通用维度的形式标注，同一份报告中不同背景信息的数据通过通用维度加以标注和区分。

在模型定义中对于多背景信息的原子数据、复合数据采用“原子或复合数据[维度]”的形式标注。例如：基金信息披露中“下属分级基金的基金简称”的通用维度标注形式，如图14所示。

基金名称[12]	（0009）		
基金简称	（0011）		
基金主代码[13]	（0012）		
基金代码[14]	（0014）/（0015）		
基金动作方式[15]	（0017）		
基金合同生效日	（0018）		
基金管理人	（0186）		
基金托管人	（0213）		
报告期末基金份额总额	（1702）		
基金合同存续期	（0023）		
基金份额上市的证券交易所（若有[6]）	（0120）		
上市日期（若有）	（0121）		
下属分级基金的基金简称[17]	（0011）	（0011）	（0011）
下属分级基金的交易代码	（0012）/（0014）/（0015）	（0012）/（0014）/（0015）	（0012）/（0014）/（0015）
报告期末下属分级基金的份额总额	（1702）	（1702）	（1702）

→ D5FD000110［D5GD030000］

图 14 “下属分级基金的基金简称”通用维度设计图

9.4.2 原子数据梳理

根据以下原则对原子数据的粒度适当性进行判断：

——财务类数据以财务报表及财务报表附注中的独立财务科目以及相关的独立概念为原子数据；

——非财务数据选择具有统计与分析价值的概念作为原子数据，当一个概念理论上可细分但其细分概念不具有统计与分析价值的时候，则不再细分；

——当信息披露相关法律法规明确列举一项需披露的内容时，该项内容至少应被定义为一个原子数据，如有进一步细分的必要时可定义为复合数据。

注：上述细分原则中列举的财务类数据，由于大多数可由计算得到，因此可能被定义为复合数据。

原子数据样例如图15所示。

* 编码	* 中文名称	* 英文名称	详细定义	* 数据层次	复合类型	取值与公式	枚举值编码	* 数据类型	格式定义	* 层级代码	引用类别	被引用项编码	引用说明
D5GD010000	报表类型	Report Type	报表类型	1			DIM7100	E		C			
D5GD020000	数据状态	Data Status	数据状态	1			DIM7200	E		C			
D5GD030000	分级类别	Classification Dype	分组类别	1			DIM7300	E		C			
D5GD040000	所有者权益维度	Equity Member	所有者权益维度	1			DIM7400	E		C			
D5GD040100	归属于母公司所有者权益维度	Equity Attribuable To Owners Of Parent Member	归属于母公司所有者权益维度	1			DIM7410	E		C			

图 15 原子数据样例

9.4.3 复合数据梳理

复合数据的梳理工作方法是在原子数据梳理的基础上，把可以通过计算、构成以及其他因素加工而成的数据识别出来，确定为复合数据。

财务数据以复合数据为主。

复合数据样例如图16所示。

基础属性								技术属性		
* 编码	* 中文名称	英文名称	数据层次	* 复合类型	* 详细定义	取值与公式	引用代码编号	* 数据类型	数据长度	格式定义
D2FR000001	货币资金	Bank Balances And Cash	2	计算类复合	企业会计准则 33 号《合并财务报表》			N	(38,4)	
D2FR000002	结算备付金	Settlement Reserve	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000003	拆出资金	Due From Banks And Other Financial Institutions	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000004	交易性金融资产	Financial Assets Held For Trading	2	计算类复合	企业会计准则 33 号《合并财务报表》			N	(38,4)	
D2FR000005	应收票据	Notes Receivable	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000006	应收账款	Accounts Receivable	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000007	预付款项	Advances To Suppliers	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000008	应收保费	Premium Receivable	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000009	应收分保账款	Amounts Due From Reinsurers	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000010	应收分保合同准备金	Reinsurance Contract Reserve Receivable	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000011	应收利息	Interest Receivable	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000012	应收股利	Dividends Receivable	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000013	其他应收款	Other Receivables	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000014	买入返售金融资产	Financial Assets Purchased Under Agreements To Resell	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000015	存货	Inventories	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	
D2FR000016	一年内到期的非流动资产	Current Portion Of Non Current Assets	2	计算类复合	企业会计准则 30 号《财务报表列报》			N	(38,4)	

图 16 复合数据样例

9.4.4 可复用数据表梳理

可复用数据表梳理工作是在相同主体或相同类别的模板范围内，对每个模板内容进行分析，对相关的语义上独立的数据表进行抽象与归纳，提炼出逻辑上可复用的数据表。

在数据提炼阶段，暂不考虑跨披露主体和模板的复用问题，关注重点在于单一模板内的可复用性。在信息披露实践中会出现同一内容在多处表现的情况，在单一模板内的原子数据和复合数据不能出现重复定义的情况。

可复用数据表样例如图17所示。

* 编码	* 中文名称	* 英文名称	* 详细定义	* 数据层次	* 层级代码	引用类别	被引用项编码	引用说明
T5FR100010	资产负债表	Balance Sheet	资产负债表	3	C			
T5FR100020	利润表	Income Statement	利润表	3	C			
T5FR100030	现金流量表	Cash Flow Statements	现金流量表	3	C			
T5FR100040	所有者权益表	Statement Of Changes In Equity	所有者权益表	3	C			
T5FR100050	报告信息表	Report Information	报告信息相关信息表	3	C			

图 17 可复用数据表样例

9.5 重组（ReCombine）

从模板映射出语义独立数据表，提炼出原子数据、复合数据、可复用数据表后，不同披露主题、产品的披露模板中仍然存在一定的共性内容，应对比其共性内容进一步识别和提炼可复用的数据表，识别业务含义相同的原子数据和复合数据，替换模型中的相应元数据定义，使得同一业务含义的数据只定义一次。

从模板映射和提炼出来的可复用数据表、语义独立的数据表均属于根据模板梳理的“展示”层面的模型，这些表和原子数据、复合数据的对应关系也属于“展示”层面，实际存储层面应根据数据之间的关系对这些数据进行重新组织，确定原子数据、复合数据的表存储结构，并确定表间关系以及主键。

9.6 回归（Regress）

对于已经分析得出的信息披露模型，应对其进行回归验证。信息披露模型回归验证方法包括逻辑推演和测试验证两种方式：

a) 逻辑推演：深入分析信息披露业务的原始需求，罗列信息披露业务的场景与流程，推演信息披露模型对现有业务需求的适用性，以及对未来可能业务需求的兼容性；
b) 测试验证:通过搭建原型测试系统,测试验证模型对典型信息披露业务的场景与流程的适应性，覆盖典型信息披露业务的全部数据范围以及数据之间的关系，能够方便、高效地承载信息披露业务的数据与应用。

通过回归验证的内容，纳入最终的信息披露模型。未能通过回归验证的内容，应重新按照梳理方法进行梳理，直至通过回归验证。

9.7 步骤简述

披露部分模型设计步骤如下：

a) 针对全部模板中梳理出的数据项分类,按照合并同类项的原则进行聚合,形成独立的数据实体；
b) 根据业务的关联度，构建实体与实体之间的关系；
c) 模型以图形加表格的形式作为主要的描述方法。

10 元语定义

10.1 各层次数据元语定义

10.1.1 数据元语定义

数据包括原子数据和复合数据，数据元语包括基础属性及技术属性和复合逻辑属性两大类，其中原子数据具有基础属性及技术属性，复合数据具有复合逻辑属性，数据元语及元语描述如表1所示。

表1 数据基础属性及技术属性

元语类别	元语	元语描述
基础属性	编码[a]	XXXXXXXXXX: 第一位：字符型，长度1位，代表数据和表分类（D数据，T表）； 第二位：字符型，长度1位，用于梳理过程中分配给不同的小组以便协同工作，“0”分配给公共主题，“1”和“2”分配给交易组主题，“3”和“4”分配给监管组，“5”和“6”分配给披露组，其他暂不使用； 第三到十位：字符型，长度8位，由字母和数字组成，表示顺序编码。
	中文名称[a]	X[100] 字符型，长度不超过100。

表 1　数据基础属性及技术属性（续）

元语类别	元语	元语描述
基础属性	英文名称 [a]	X[300] 字符型，长度不超过 300； 使用驼峰命名法，即：每个单词的首字母大写，只允许出现大小写字母、数字和“_”（英文下划线）填写此项。
	详细定义	X[1000] 字符型，长度不超过 1000； 用于对该原子数据进行详细的描述。 对于约定俗成的常识概念，可以不填写。
	数据层次 [a]	X[1] 字符型，长度 1 位，含义如下： —1 表示原子数据； —2 表示复合数据。
	复合类型	X[1] 字符型，长度 1 位，含义如下： —1 结构性复合； —2 计算类复合； —3 离散型复合。 （对于复合数据此项为必填项）。
	取值与公式	X[1000] 字符型，长度不超过 1000； 用于表示数据的统计口径、取值范围、枚举类型等； （对于复合数据此项为必填项）。
	枚举值编码	X[7] 字符型，应符合表 7 代码表编码元语描述。（当数据项为代码时，此项为必填项）。
技术属性	数据类型 [a]	X[1] 字符型，长度 1 位。 如：C 字符型、B 布尔型、N 数值型、D 日期型、T 日期时间型、E 枚举型。 数据类型定义应符合附录 A。
	格式定义	针对某种特定数据类型的格式定义和约束。 如：D 日期型[YYYYMMDD]。
	层级代码 [a]	X[5] 字符型，长度 5 位，含义如下： 用于表示模型中的层次。 —C 代表行业公共部分；—T 代表交易公共部分； —R 代表监管公共部分；—D 代表披露公共部分； T0001-T9999 代表某一交易类特定的应用编码； R0001-R9999 代表某一监管类特定的应用编码； D0001-D9999 代表某一披露类特定的应用编码。 应用编码和应用名称由交易、监管、披露分别对各自部分进行维护。
	引用类别	X[1] 字符型，长度 1 位，含义如下： 用于描述该数据引用其他数据的类别。 —1 具象引用； —2 应用引用； —3 同义引用。 引用类别定义应符合附录 B。
	被引用项编码	XXXXXXXXXX: 字符型，应符合表 1 中对编码的元语描述； 当引用类别非空时，此项必填。

表 1　数据基础属性及技术属性（续）

元语类别	元语	元语描述
技术属性	引用说明	X[1000] 字符型，长度不超过 1000； 用于表示数据间引用关系的说明。
	参考数据级别	X[10] 字符型，长度不超过 10，含义如下： 用于表示数据的定级，即数据重要程度的标识，一般使用等级描述标识进行描述； —1 4 级（极高）：数据主要用于行业内大型或特大型机构中的重要业务使用，一般针对特定人员公开，且仅为应知悉的对象访问或使用； —2 3 级（高）：数据用于重要业务使用，一般针对特定人员公开，且仅为应知悉的对象访问或使用； —3 2 级（中）：数据用于一般业务使用，一般针对受限对象公开，一般指内部管理且不宜广泛公开的数据； —4 1 级（低）：数据一般可被公开或可被公众获知、使用。
[a]为必填项		

复合逻辑属性用于说明复合数据项与原子数据项之间的一对多的逻辑关系，如表2所示。

表 2　数据复合逻辑属性

元语类别	元语	元语描述
复合逻辑属性	复合数据编码[a]	按照“基础属性-编码”。
	复合关系序号[a]	X[1] 数字型，长度 1 位，编号为 1-9，用于表示某一复合属性下的每一种复合关系，每个复合数据的关系序号不延续。
	中文名称	按照“基础属性-中文名称”。
	成员数据编码[a]	按照“基础属性-编码”； 用于建立复合数据与成员（原子或复合）数据之间的对应关系。
	成员数据中文名称	按照“基础属性-中文名称”。
[a]为必填项		

10.1.2　表元语定义

表包括可复用数据表和语义独立数据表，表元语包括基础属性、组成属性、表内关系属性和表间关系属性，表元语及元语描述如表3所示。

表 3　表基础属性

元语类别	元语	元语描述
基础属性	编码[a]	XXXXXXXXXX: 第一位：字符型，长度 1 位，代表数据和表分类（D 数据，T 表）； 第二位：字符型，长度 1 位，用于梳理过程中分配给不同的小组以便协同工作，“0”分配给公共主题，“1”和“2”分配给交易组主题，“3”和“4”分配给监管组，“5”和“6”分配给披露组，其他暂不使用； 第三到十位：字符型，长度 8 位，由字母和数字组成，表示顺序编码。
	中文名称[a]	X[100] 字符型，长度不超过 100。
	英文名称[a]	X[100] 字符型，长度不超过 100。

表 3　表基础属性（续）

元语类别	元语	元语描述
基础属性	详细定义 [a]	X[1000] 字符型，长度不超过 1000。 用于对该表进行详细的描述。
	数据层次 [a]	X[1] 字符型，长度 1 位，含义如下： —3 表示可复用数据表； —4 表示语义独立数据表。
	层级代码 [a]	X[5] 字符型，长度 5 位，含义如下： 用于表示模型中的层次： —C 代表行业公共部分； —T 代表交易公共部分； —R 代表监管公共部分； —D 代表披露公共部分； T0001-T9999 代表某一交易类特定的应用编码； R0001-R9999 代表某一监管类特定的应用编码； D0001-D9999 代表某一披露类特定的应用编码。 应用编码和应用名称由交易、监管、披露分别对各自部分进行维护。
	引用类别	X[1] 字符型，长度 1 位，含义如下： 用于描述该数据表引用其他数据表的类别。 —4 继承 引用类别定义应符合附录 B。
	被引用项编码	XXXXXXXXXX: 字符型，应符合表 3 中对编码的元语描述； 当引用类别非空时，此项必填。
	引用说明	X[1000] 字符型，长度不超过 1000； 用于表示数据间引用关系的说明。
[a]为必填项		

组成属性表用于描述表的组成成员，以及成员与其他被引用项之间的关系表，如表4所示。数据表可由原子数据、复合数据或者被引用的表组合而成。

表 4　表组成属性

元语类别	元语	元语描述
组成属性	数据表编码 [a]	按照“基础属性-编码”。
	中文名称 [a]	按照“基础属性-中文名称”。
	成员数据编码 [a]	按照“基础属性-编码”。 用于建立可复用数据表与成员（原子、复合或可复用）数据之间的对应关系。
	成员数据中文名 [a]	按照“基础属性-中文名称”。
	成员数据层次 [a]	X[1] 字符型，长度 1 位，含义如下： —1 原子数据； —2 复合数据； —3 可复用数据表； —4 语义独立数据表。

表 4　表组成属性（续）

元语类别	元语	元语描述
组成属性	引用类别	X[1] 字符型，长度 1 位，含义如下： 用于表示该数据表的组成成员与被引用项之间的关系。 —1 具象引用； —2 应用引用； —3 同义引用。 引用类别定义应符合附录 B。
	被引用项编码	XXXXXXXXXX: 字符型，应符合表 1 中对编码的元语描述；当引用类别非空时，此项必填。
	引用说明	X[1000] 字符型，长度不超过 1000； 用于描述引用关系。
[a]为必填项		

表内关系属性用于描述各表内的各数据之间的关系，如表 5 所示。

表 5　表内关系属性

元语类别	元语	元语描述
表内关系属性	数据表编码[a]	按照“基础属性-编码”。
	中文名称[a]	按照“基础属性-中文名称”。
	表内关系序号[a]	X[2]为顺序号，由 00-99； 每个数据表的关系序号不延续。
	表内关系说明[a]	X[1000] 字符型，长度不超过 1000； 用于对表内逻辑关系进行详细的描述。
[a]为必填项		

表间关系属性用于描述表与表之间的关系，如表 6 所示。

表 6　表间关系属性

元语类别	元语	元语描述
表间关系属性	数据表编码[a]	按照“基础属性-编码”。
	中文名称[a]	按照“基础属性-中文名称”。
	表间关系序号[a]	X[2]为顺序号，由 00-99； 每个数据表的关系序号不延续。
	表间关系说明[a]	X[1000] 字符型，长度不超过 1000。 用于对表间逻辑关系进行详细的描述。
[a]为必填项		

10.2　代码元语定义

代码用于描述某些在取值上具有枚举特征的原子、复合的取值范围及内容。代码分为公共代码和交易、监管及信息披露专用代码，代码表元语包括基础属性及代码取值属性两部分，代码表基础属性如表7所示。

表 7　代码表基础属性

元语类别	元语	元语描述
基础属性	代码表编码 [a]	XXXXXXX: 第一位到第三位：字符型，长度 3 位，固定为“DIM”用于标识代码类数据表； 第四位到第七位：数值型，长度 4 位，顺序编号，由 0001-9999；公共代码可用区间为 0001-0999，交易类可用区间为 1000-3999，监管类可用区间为 4000-6999，信息披露类可用区间为 7000-9999。
	中文名称 [a]	X[100] 字符型，长度不超过 100。
	上级代码表编码 [a]	
	代码表说明 [a]	X[1000] 字符型，长度不超过 1000。 用于该代码表的描述。
	参照标准	代码定义参考的相关标准或文件。
	备注	X[1000] 字符型，长度不超过 1000。 用于说明代码及其取值的适用范围及其他关系说明。
[a] 为必填项		

代码取值属性如表 8 所示。

表 8　代码取值属性

元语类别	元语	元语描述
取值属性	代码表编码 [a]	XXXXXXX: 第一位到第三位：字符型，长度 3 位，固定为“DIM”用于标识代码类数据表； 第四位到第七位：数值型，长度 4 位，顺序编号，由 0001-9999；公共代码可用区间为 0001-0999，交易类可用区间为 1000-3999，监管类可用区间为 4000-6999，信息披露类可用区间为 7000-9999。
	代码表中文名称 [a]	X[100] 字符型，长度不超过 100。
	代码取值编码 [a]	X[10] 长度 10 位，用于描述代码表中各取值； 采用分级编码，每两位表示一个层级，由 01-99。
	代码取值描述 [a]	X[100] 字符型，长度不超过 100； 用于该代码的取值的含义。
	备注	X[1000] 字符型，长度不超过 1000。 用于说明代码取值的适用范围及其他关系说明。
[a] 为必填项		

11　产出物说明

根据抽象模型设计方法论，以元语定义为基础，按照公共、交易、监管、披露四个部分梳理形成了数据模型产出成果，包括实体、属性、代码等。这些模型产出成果通过专门的数据模型管理平台进行存储及管理，并提供了浏览、查询、修改、删除、评审等功能。证券期货业数据模型管理平台的访问地址为http://sdom.csisc.cn。

附 录 A
（规范性附录）
数据类型

A.1 字符型：C

字符型（Character）数据是不具备计算能力的文字数据类型，用字母C表示。
字符型数据包括中文字符、英文字符、数字字符和其他ASCⅡ字符及其他不同语种的字符。

A.2 布尔型：B

布尔型（Boolean）表示条件的有效性，用字母B表示。
它的值只有两个：false（假）和true（真）。false的序号为0，true的序号是1。

A.3 数值型：N

数值型（Numeric）是表示数量、可以进行数值运算的数据类型，用字母N表示。
数值型数据由数字、小数点、正负号和表示乘幂的字母E组成。

A.4 日期型：D

日期型（Date）是表示日期的数据类型，到年/月/日，用字母D表示。

A.5 时间型：T

时间型（Time）是表示时间的数据类型，到时/分/秒，用字母T表示。

A.6 枚举型：E

枚举型（Enum）是指将变量的值一一列出来，变量的值只限于列举出来的值的范围内，用字母E表示。

附 录 B
（规范性附录）
引用类别

B.1 具象引用

引用项与被引用项名称不同、语义的核心内容相同、语法相同，被引用项的取值范围在引用项的取值范围进行了缩小。

语法涉及计算公式、数据类型、格式定义，语义涉及取值范围、详细定义。

B.2 应用引用

引用项与被引用项名称不同、语义上限定了特别的应用场景、语法相同，引用项与被引用项计算方式及基本公式相同。

B.3 同义引用

引用项与被引用项名称不同、语义完全一致、语法相同。

B.4 继承引用

用于描述两个数据表之间的关系，二者语义不同；引用项必须包含被引用项的所有元语定义及成员，并增加了被引用项所没有的元语定义及成员。

参 考 文 献

[1] GB/T 35964—2018 证券及相关金融工具 金融工具分类（CFI编码）
[2] JR/T 0124—2014 金融机构编码规范