

数据点模型的原理及应用

杜美杰 刘凯 赵慧周

作为可扩展商业报告语言, XBRL 的应用范围不仅包括对外财务报告, 更应包括整个商业信息供应链上各种对内对外商业报告。如何将纷繁复杂的商业报告中隐晦的业务数据翻译为精确的、逻辑的、明晰的信息并进一步形成 XBRL 分类标准, 已成为大数据环境下商业信息沟通和使用的关键。鉴于此, 本文将介绍一种新的数据建模方法——DPM, 以期对 XBRL 分类标准的开发提供新的思路。

一、DPM 产生的背景

2004 年 6 月, 巴塞尔银行监管委员会发布了《统一资本计量和资本标准的国际协议: 修订框架》(International Convergence of Capital Measurement and Capital Standards: A Revised Framework) 的最终版本, 即巴塞尔新资本协议(也称 Basel II)。2005 年, 按照该协议的监管思路, 欧盟通过了资本要求指令(Capital Requirement Directive, CRD), 该指令明确规定将新资本协议的监管内容转换为欧盟法律。随后, 欧洲银行监管委员会(Committee of European Banking Supervisors, CEBS) 于 2005 年 12 月和 2006 年 1 月先后发布了关于通用财务数据的财务报告(Financial Reporting, FINREP) 框架指南和关于通用偿付比率的共同报告(Common Reporting,

COREP) 框架指南。尽管欧盟以法律的形式进行了限定, 但由于缺乏有效的数据管理模式, 在转化为欧洲各国法令并实施的过程中, 出现了各国报告模板差异较大、CRD I-III 数据之间缺乏可比性、缺乏统一的数据交换格式等问题。2009 年, 为应对全球金融危机, 在 G20 峰会的推动下, 国际社会制定了一揽子金融改革计划, 其中包括 Basel III。在采纳 Basel III 协议的过程中, 欧盟制订了统一的 CRD IV、单一规则手册和报告模板, 以加强欧盟境内银行业的监管协调工作, 促进跨境监管的正常运转和整个欧洲“以同一种语言沟通”的目标实现。作为单一规则手册的重要组成部分, 2011 年, 欧洲银行管理局(European Banking Authority,

EBA, 2011 年 1 月由 CEBS 更名) 启动了监管报表实施技术标准(Implementing Technical Standards, ITS) 项目, 极大地推动了从人工合规报送到电子格式报送的转变。2012 年, EBA 正式开展了有关 ITS 的数据点模型(Data Point Model, DPM) 建模, 2013 年发布了基于 DPM 的 COREP 和 FINREP 的 XBRL 2.0.0 分类标准, 并要求所有欧盟成员国分别自 2014 年 1 月 1 日和 2014 年 9 月 30 日开始执行。

二、DPM 的基本结构

统计(监管)方通常需要收集大量的信息进行分析, 其信息需求有三种常见的描述方式, 如图 1 所示。

为提高数据的管理和使用效率, 除

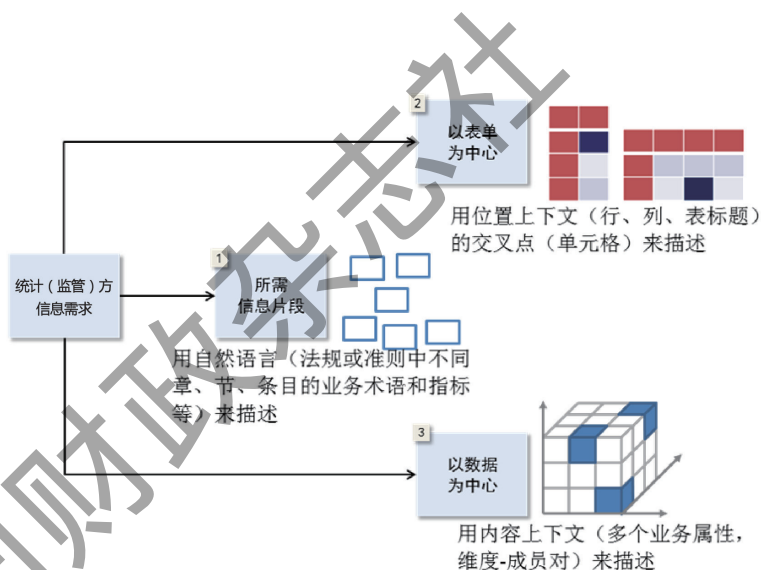


图1 统计(监管)方信息需求的三种常见描述方式

表1 以表单为中心和以数据为中心的数据结构界定模式比较

	以表单为中心	以数据为中心
数据描述方式	复杂的模棱两可的概念以数据的视觉化为导向	一致、清晰、明确的基本概念包括基础项和所有必要识别的附加维度
数据识别方式	所设计模板的行和列	多维的[维度][成员]对的集合
是否需要特定的位置上下文	如果缺乏位置上下文,数据就失去含义	无需位置上下文(模板、行、列),数据自身具有含义

了以法规或准则的形式进行说明以外,通常以表单或数据为中心来对数据的内容和结构进行描述。以表单为中心,是指表单中的每个单元格内容都由独一无二的行名和列名代表,离开了表单,这些单元格没有意义。以数据为中心,是指表单中每个概念都由自身相关业务属性决定,表单中的数据自身具有独立意义,不依附表单存在。二者之间的区别,如表1所示。

DPM是以数据为中心的建模方式,既表达了数据的内容,又表达了数据的结构。在DPM中,报表中每个数据点都由基础项(primary item)和维度(dimension)定义,关系如图2所示。该模型可视作n维空间,维度为轴,成员为轴坐标;每个数据点有一个精确的空间位置,由特定坐标(独一无二的[维度][成员]对的组合)定义。基础项是指每个数据点都需要界定的属性,包括:数据类型(data type),如数字型、日期型、文本型和货币型等;时间类型(period type),如时点、时期等;金额类型(amount type),如账面价值、名义价值、公允价值等。维度是指数据点基本属性以外其他需要界定的属性的类别,如活动发生地(location of the activity)、货币类型(currency)等。成员(member)是指维度的取值,每个成员都必须有准确的定义,且同一维度中的成员不能重复。域(domain)是指由本质含义相同的成员组成的集合。层级结构(hierarchy)是指域中成员之间的级别关系和运算关系,每个成员在域中都是通过父子节点关系组织起来的。由此可见,DPM中的每个

数据点都由属性独立界定,反映了数据的本质特点,有助于实现数据结构的明确性、一致性和连贯性。

三、DPM的建模和映射

数据建模通常是指对现实世界各类数据的抽象组织,即确定数据库需管辖的范围、数据的组织形式等并转化成现实数据库的过程。DPM是以数据为中心的建模方式,通过创建多维语义模型来反映统计(监管)方的信息需求。DPM通常有两种建模思路:一种是理想情况下,先开发DPM,即先以分解的方式识别不同的相关领域,然后将法规和配套指引(如代表目前报告要求的表单)作为这些分解数据的所有潜在逻辑组合的一个子集,来满足业务专家履行监管任务和政策制定的信息需求。另一种是先定义表单(包括详细指引),然后对它们进行分析,列出所有的业务属性并加以分解。第二种方法更常见,因为它反映了信息需求的演进过程。当然,在基于现有表单开发

DPM的过程中,可能需要调整一些模板的布局,以通过一致的模型实现多种灵活的表单视图。

以基于表单的DPM为例,其原型建模和向XBRL映射一般需要如下几个步骤:

(一) 分析原表单

1. 明确表单填报目的。通常反映在报表标题上,它体现了表单的基本信息和数据通用类型。

2. 明确表单交叉项目内容,即行字段和列字段共同定义的单元格内容。

3. 分析行字段和列字段各自内部的等级关系(包括内部隐含关系)。

4. 其他外部支持文本或表单,例如其他相关指引和参考说明。

5. 表单内部隐含的相关信息和支撑内容。

(二) 拆分、合并、注解原表单

1. 先按照行字段所具有的层次类别对原表格进行拆分、合并。

2. 再按照列字段所具有的层次类别对原表格进行拆分、合并。

3. 用DPM概念,即“域”、“维度”和“成员”,对每个拆分表格的行和列进行分析,并特别注意字段名中所包含的隐含信息。

(三) 数据点模型原型内容

数据点模型原型一般以EXCEL形式与业务人员沟通,通常包括如下内容:

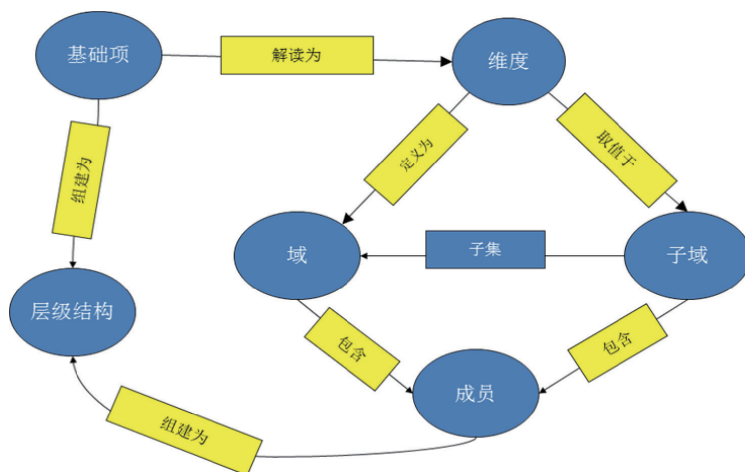


图2 DPM数据结构：概念间的关系

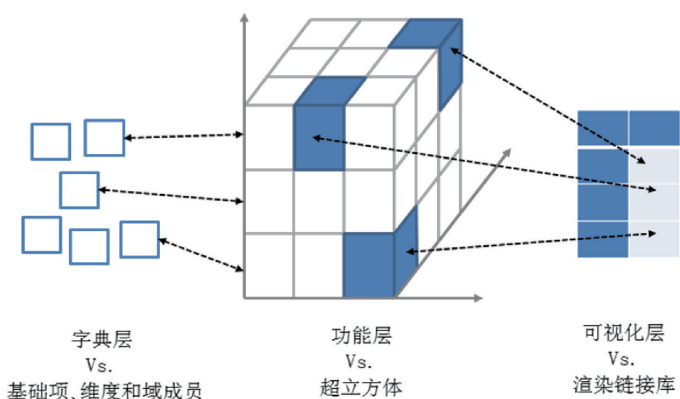


图3 DPM与XBRL分类标准的映射关系

1. 基础项，一般包括：表名称、标签、数据类型、时间类型等。
2. 维度与成员列表，格式为：名称、标签和类型。
3. 域、层次结构表，格式为：层次结构序号、层级、成员名称。
4. 数据点表：每个可用单元格所对应的维度和成员值。

(四) 映射为XBRL分类标准

业务专家建立数据点模型原型后，在XBRL技术专家的协助下进一步添加XBRL相关的技术信息后，可以方便地映射为XBRL分类标准。DPM的三个层面(字典层、功能层和可视化层)同XBRL分类标准(基础项、维度和域成员，超立方体，渲染链接库)相对应，如图3所示。

四、DPM的应用和展望

传统的财务报告XBRL分类标准，以表单为中心进行数据建模，主要遵循直观的层次结构分类。而现代的商业报告则包括财务、风险、环境等诸多内容，以表单为中心将力不从心，甚至无法胜任。以金融行业为例，其所需信息往往具有以下特点：(1) 指标数量众多，如Basel III在欧洲银行业的实施涉及到的财务、风险监管指标就有3.5万个以上；(2) 财务指标和非财务指标往往在同一报告框架中共存，彼此之间有很复杂的逻辑关系；(3) 指标通常涉及多个维

度，如交易对手、金融产品、行业、地区、渠道、期限等多维信息；(4) 数据聚合和报告的自动化程度较高，金融业通常已经采用了数据仓库，可自动生成用于联机分析处理(On-Line Analytical Processing, OLAP)的多维数据集。正是基于这些特点，欧盟开发了以数据为中心的DPM，旨在：(1) 为数据点(指标)提供详细、标准的技术性描述，以确保对报告指标的共同解读；(2) 通过关注指标信息本身的逻辑结构，而非具体模板的结构和上下文，使数据点定义独立于任何特定视图，为分析探索和重新设计未来模板奠定基础；(3) 在使用俗称概念描述数据时应用逻辑和正式的方法，简化业务专家和信息技术专家之间的沟通；(4) 在不依赖于业务专家的情况下，IT数据交换格式，尤其是XBRL分类标准可基于DPM产生。由此可见，DPM这种建模方式将极大地促进业务和IT人员之间的沟通，从而解决数据治理和数据管理中最基础的问题。

欧洲央行和欧洲银行管理局是选择DPM方法的先锋。EBA(欧洲银行管理局)通过对COREP/FINREP报告框架的分解，得到了由107个维度和2174个成员以及域、度量、层次结构、表达式组成的DPM模型，反映了分布于161份表格中的35259个数据点。虽然EBA建议的XBRL分类标准主要用于27个国家监管

部门(National Supervisory Authorities, NSA)和EBA之间的数据传输，尚未强制8500多家信贷机构和投资公司采用，但许多NSA都打算利用该XBRL分类标准收集所监管的信贷机构和投资公司的监管报告。一旦NSA实现上述目标，这个统一的XBRL报告格式将有望提高欧洲信贷机构和投资公司监管合规工作的效率并降低成本。与此同时，欧洲保险和职业养老金管理局(European Insurance and Occupational Pensions Authority, EIOPA)也采用DPM方法开展了Solvency II的实施，同合作运营研究与发展协会(Association for Cooperative Operations Research and Development, ACORD)开发的保险数据标准进行了对接，并针对保险及再保险行业制订了XBRL分类标准。欧洲证券与市场管理局(European Securities and Markets Authority, ESMA)也正在筹划对ESMA标准进行DPM建模并开发相应的XBRL分类标准。可见，DPM这种建模方式因其数据的通用性，将大大促进跨界监管，对我国现阶段的政府监管和信息披露有很强的借鉴意义。

除了在报表层面的应用外，DPM建模还可进一步深入到业务的交易明细层面。例如，EIOPA通过DPM建模深入理解了XBRL Solvency II和ACORD GRLC(Global Reinsurance and Large Commercial)中的数据项，并最终实现了二者之间的映射，从而不仅覆盖了数据仓库及上层分析应用的指标数据和XBRL报送数据，还覆盖了业务系统中的操作数据，切实保障了不同系统间以及系统内外数据含义的一致性、可比性和互操作性。

综上所述，DPM最重要的意义在于：用更少的维度、成员要素来表达更多数目的报告数据项；以无冗余的、一致的和明确的方式表达业务术语；支持IT和业务人员之间的沟通。2012年6月，XBRL Abstract 2.0提出了依据

财务业务一体化在中国联通的应用与实践

唐志勇

目前,中国联合网络通信集团有限公司(以下简称中国联通)在财务业务一体化进程中快速推进,已基于ERP核心系统搭建了大ERP系统体系,初步实现了财务业务一体化架构,建立了一本账一套表的业务透视体系,财务核算功能得到了更好的发挥,财务数据与业务数据同步的程度得到提高。

一、结合自身特点,构建中国联通大ERP系统

1. 中国联通大ERP系统分为ERP核心系统和外围系统

ERP核心系统是指基于Oracle系统实施的与财务核算紧密关联的部分,包括总账模块(GL)、应收模块(AR)、应付模块(AP)、固定资产模块(FA)、工程项目模块(PA)、采购模块(PO)、库存模块(INV)及现金管理模块(CE)。

ERP外围系统是指直接或间接与ERP核心系统有数据交互关系的业务系统,包括报账系统、集中收入管理系统、集中渠道管理系统、合同管理系统、营销资源管理系统、电子渠道销售系统(简称ESS系统)、集中PRM系统、集中结算系统、资金管理系统、人力资源管理系统(简称HR系统)、员工自助管理系统(简称e-HR系统)、税务管理系统、采购管理系统、项目管理系统(简称PMS系统)、主数据管理系统、合作方自服务门户系统等。

2. ERP核心系统和外围系统业务范围

ERP核心系统业务范围:(1)应收模块业务范围涉及接收集中收入管理系统传送的营业日报,营业收款、杂项收款、营业款核销;(2)应付模块业务范围涉及接收报账系统传送的发票及其

付款、预付款核销信息;(3)采购模块业务范围涉及采购管理系统传送的采购订单,接收采购管理系统导入的工程合同;(4)库存模块业务范围涉及采购订单的到货接收,相关物资的入库、出库、调拨等库存管理;(5)项目模块业务范围涉及在建工程核算,工程的预转固、工程的竣工决算,项目信息归集、项目成本归集(含杂项,如基站前六个月租金资本化处理)、利息资本化、人工成本资本化、在建工程经过初验以后转为固定资产,项目、非项目业务进项税转出视同销售业务增值税调整;(6)资产模块业务范围涉及固定资产、无形资产、非合同长期待摊费用的增加、调整、折旧/摊销计提、减值、报废、调拨;(7)总账模块业务范围涉及与外围系统、子系统的集成,接收外围系统、子模块相关数据,内部往来(AGIS)、外围系统及

DPM来构建XBRL分类标准的元模型。随着商业报告的颗粒度越来越细、维度越来越多,为了简化报表生成负担,基于DPM构建XBRL分类标准将成为必然。这种以数据为中心的包含维度信息的XBRL分类标准将有助于监管机构采用多维数据集进行低成本数据交换、存储及使用,并极大地促进信息使用者从多个维度开展切片、切块式的交互式分析挖掘。■

[本文得到教育部人文社会科学研究规

划基金项目“XBRL统一报告研究:基于本体论视角”(11YJA630012)、中央高校基本科研业务专项资金北京语言大学科研项目“XBRL技术规范的语义形式化”(13YBG49)和“可扩展商业报告语言(XBRL)分类标准研究”(09JBT013)的资助]

(作者单位:北京语言大学国际商学院 国际数据管理协会中国分会 <DAMA China> 北京语言大学信息科学学院)

责任编辑 李斐然

参考文献:

1. Ignacio Boixo.2012.The Data Point Model methodology in the European Supervision: COREP/FINREP, 24th World Continuous Auditing and Reporting Symposium “Continuous Auditing and XBRL Coming of Age”, May 3-4, Malatya / Turkey
- 2.EBA.Abstract Description of the Model Represented in Taxonomies Following the DPM Approach.http://www.eurofiling.info/dpm/index.shtml