

基于数据仓库的行政事业单位合同管理内部控制评价

——以重庆海事局为例

程平^{1,2}(博士生导师), 何昱衡^{1,2}

【摘要】 经济合同管理内部控制评价对行政事业单位规范管理、风险识别、风险应对等方面具有重要意义,而基于数据仓库的合同管理内部控制评价能较好地满足行政事业单位的内部控制评价需求。以重庆海事局为例,基于财政部《行政事业单位内部控制规范(试行)》,依据合同管理内部控制评价指标,设计出行政事业单位业务层面下合同管理内部控制评价数据仓库体系,针对每个合同管理内部控制评价指标构建对应的数据集,并基于此构建数据仓库的概念模型、逻辑模型和物理模型,最后基于Hadoop平台的Hive数据仓库的类SQL语句,以合同管理监控有效性评价为例进行数据分析。

【关键词】 行政事业单位; 合同管理; 数据仓库; 内部控制评价

【中图分类号】 F233; C931

【文献标识码】 A

【文章编号】 1004-0994(2019)19-0059-5

目前的文献很少涉及基于数据仓库的数据分析来进行行政事业单位内部控制评价的。有鉴于此,本文以重庆海事局为例,探讨基于数据仓库的行政事业单位合同管理内部控制评价,期望能为行政事业单位的内部控制评价提供一定的参考。

一、合同管理内部控制评价数据仓库体系结构设计

1. 设计思路。行政事业单位在财政部《行政事业单位内部控制规范(试行)》(简称《内控规范》)的基础上,结合本单位的实际情况对业务层面的合同管理制定出相关管理办法,具体包括合同的职责划分、签订、履行、变更解除、监督和检查等方面。以重庆海事局为例,2018年3月12日,在《内控规范》的基础上,重庆海事局结合本单位的合同管理实际情况发布了《重庆海事局经济合同管理办法》(简称《管理办法》),并以此为标准对重庆海事局的合同管理内部控制建立了一系列评价指标。

重庆海事局目前的合同管理内部控制评价还处

于通过线下调研和考察对合同管理内部控制的效果进行评价打分的阶段。基于数据仓库,从《内控规范》和《管理办法》中关于合同管理内部控制的规定出发,针对合同评价主题建立重庆海事局合同管理内部控制评价数据仓库。首先从不同的系统中采集与合同相关的数据,并通过ETL过程(数据抽取、转换及加载)清洗数据,然后基于数据仓库对合同管理内部控制评价所需的相关信息进行存储,结合合同管理内部控制评价的主题来建立数据集,最后在评价阶段利用联机分析处理等方式进行评价打分,并通过可视化、生成内部控制评价报表等方式来呈现。

2. 设计过程。合同管理内部控制评价数据仓库体系主要由ETL过程、数据仓库存储、数据集、联机分析处理、客户端四部分组成。以重庆海事局为例,其合同管理内部控制评价数据仓库体系结构设计过程图参考《财会月刊》2019年第13期中的文章《基于数据仓库的行政事业单位单位层面内部控制评价——以重庆海事局为例》相关内容^[1]。

【基金项目】 国家社会科学基金项目(项目编号:17BGL194); 重庆海事局基于财务云平台的内部控制信息化系统建设(项目编号:2018Q60)

(1)ETL过程。ETL过程是对存储着合同相关信息的各个业务板块数据库中的信息进行抽取、转换、加载的过程。合同管理涉及相关合同网上公示,合同项目、金额、负责人等基本信息,以及与供应商的往来信息。采用Sqoop技术可对财务云平台内合同的基础信息进行抽取,基于Python的网络爬虫则可对重庆海事局官方网站上的合同公示进行抓取,这样采集的数据会依据数据仓库的特点进行再一次清洗,以方便以后的数据存储和进一步挖掘与分析。

(2)数据仓库存储。重庆海事局合同管理主题数据仓库存储着合同管理相关的业务财务信息数据和元数据。其中的元数据是指数据库的基本数据信息和用来描述信息资源、数据等的数据库。同时,基于Hadoop的Hive数据仓库软件能够将结构化的数据映射为数据库表。

(3)数据集市。它是根据分析需求将数据仓库中的数据按划分的合同管理内部控制评价主题域分类存储。结合重庆海事局现有的系统功能和实务,根据《内控规范》和《管理办法》建立评价指标体系,将合同管理内部控制评价工作分为合同审批流程完整性、合同签订合规性、合同履行监控有效性、合同变更处理合规性、合同归口管理执行性、合同信息准确性、合同档案完整性七个分析主题,从合同主题数据仓库中抽取面向分析需求的信息,建立数据集市。

(4)联机分析处理。联机分析处理(OLAP)是从合同相关的不同业务角度进行分析,对合同分析主题进行多角度、多侧面、多维度的分析。对合同业务、财务相关的数据立方体从金额、责任人、部门等多个角度进行切片和切块操作,从时间、地点等多个层级进行上卷、下钻操作,协助相关责任人和管理人对合同管理流程中的风险进行更准确的把握与判断。

三、基于合同管理内部控制评价数据仓库的数据采集与清洗

1. 数据采集。重庆海事局现有金蝶K3账务处理系统、财务云平台等系统,不同的系统都保存有大量的业务财务数据。合同管理内部控制评价数据仓库的数据源于这些系统。数据源的存放格式包括文件扫描图片、html文件、合同基本信息等非结构化数据、半结构化数据以及结构化数据。

非结构化数据是指合同文本扫描件、廉政合同文本扫描件等通过线下手工完成但需要通过扫描录入合同管理系统的影像文件。这些数据在合同管理主题分析中主要作为手续是否齐全、流程是否完整

的判断依据。半结构化数据是指经济合同审查表、经济合同复核表、合同公示等以html文件格式存储的合同相关信息。结构化数据大多存储于财务云平台数据库中。财务云平台是合同管理的主要系统,其中存储着合同金额、合同收付提醒、合同ID、供应商银行账号等基本信息。同时,金蝶K3账务处理系统存放着来自财务系统的数据,当合同管理内部控制评价的分析主题需要考察某项合同的支付情况时,需要对从金蝶K3账务处理系统中提取到的合同支付时间、金额、款项以及支付科目等信息进行分析。

2. 数据清洗。数据清洗是将采集的数据根据数据仓库的特点和分析需求对部分数据进行删除、补充、转化的过程。对于合同管理中录入的时间字段,需要将其分成年度、季度、月度、日的层次,以方便对合同录入日期进行上卷、下钻操作。对于合同当前所处的履行阶段,在财务云平台合同管理板块中并没有进行有效的反映,但是在合同履行监控有效性的评价中需要依据合同是处于保送审批、复核或是履行、结算的状态对合同进行切块分析。这就需要在数据仓库中添加合同处理状态的字段或新建维度表,专门通过合同状态对合同进行分类。

结合《内控规范》和《管理办法》,针对合同管理内部控制评价的实务需求,分别从1~3个评价点对相应指标进行评价。将各个评价点的相关数据进行选择、抓取、清洗是该步骤的重点,需充分考虑数据与评价指标的相关性、数据间的关联性。每个指标都会涉及合同某一个阶段的业务流程,从业务流程入手,对产生的数据库表中的数据进行选取。对于表中数据,要将与评价指标无关的数据剔除,选择相关字段下的数据。当对两张或多张表进行关联时,要找出其中的中间表,通过其将同一项合同或业务关联起来。

四、基于合同管理内部控制评价数据仓库设计

1. 确定分析主题域。在程平等^[2]研究的基础上,本文结合数据仓库的特点和重庆海事局内部控制评价的需求进行进一步优化,结果见表1。

依据上述评价指标,结合《内控规范》和《管理办法》,将合同管理内部控制评价分为七大主题域,并将合同管理内部控制评价数据仓库的数据综合归集成数据集市再加以分析应用,以服务于每个评价指标的分析。合同管理内部控制评价数据仓库的主题域是在内部控制指标分析要求的基础上进行确定的。在具体设计上,每个主题域对应一个合同管理内部控制指标,从一项或多项评价点对相应的主题域

进行评价。合同管理内部控制评价数据仓库的分析主题依据审计指标分为如表1所示的七个分析主题。

表1 重庆海事局合同管理内部控制评价指标

指标名称	评价点	设计依据
合同审批流程完整性	合同审批流程是否完整	《内控规范》第五十四条
合同签订合规性	用印是否规范	《内控规范》第五十四条、第五十五条及第五十七条
	合同谈判是否合规 重大合同是否经纪检部门审核	
合同履行监控有效性	合同履行提醒是否有效	《内控规范》第五十六条
合同变更处理合规性	针对合同变动是否进行合规处理	《内控规范》第五十七条
	合同的账务处理是否合规	
合同归口管理执行性	评价对合同的归档情况	《内控规范》第五十四条
合同信息准确性	评价合同信息是否准确	《内控规范》第五十八条
合同档案完整性	评价合同档案是否完整	《内控规范》第五十八条

2. 确定粒度。在合同管理内部控制评价中需要对数据按照综合程度和细化程度进行划分,并根据不同深度进行数据分析。而粒度就是进行划分的数据综合和细化程度,细化程度越高,粒度越小,反之粒度越大。将合同按收付类型进行第一层细分,再按收款合同和付款合同两种类型进行进一步细分,上述将合同一层一层细分的过程就是按合同类型确定粒度的过程。例如:在分析合同履行监控有效性时,以时间为维度考察合同提醒是否及时,通过上卷操作考察一个年度合同监控的有效性,通过下钻操作考察一个月内结算的合同监控的有效性;在以部门为维度考察合同监控有效性的时候,可以通过上卷操作考察机关单位合同的监控有效性,也可以通过下钻操作考察机关具体某个部门的合同监控有效性。

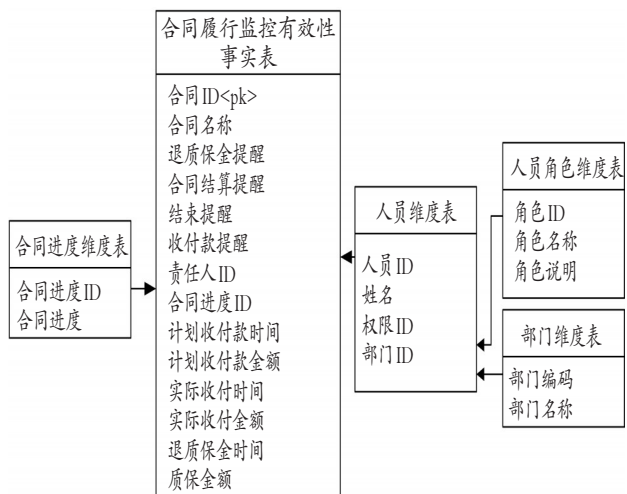
3. 确定维度。在合同管理内部控制评价数据仓库中,维度可以看作是用来分析不同指标的观察广度。在合同管理内部控制审计评价中,每个分析指标对应一个分析主题,每个分析主题都涉及与分析主题相关的各类业务、财务数据,通常以合同类型、合同责任人、合同进度等维度进行分析。不同的指标会根据分析需求设计相应的维度表,例如在重庆海事局合同履行监控有效性分析主题中,设计了合同进度表以便考察在某个进度下的合同项目是否向相关责任人及时推送了进度提醒。在考察合同变更处理

合规性时建立了合同处理维度,将其作为评价合同处理是否合规的依据,当合同实际处理和合同处理维度表中依据《内控规范》和《管理办法》所存储的处理标准数据不符时,可认定合同履行处理不合规。

4. 模型设计。重庆海事局合同管理内部控制评价数据仓库的模型设计主要分为概念模型设计、逻辑模型设计和物理模型设计三个阶段。根据合同管理内部控制评价体系中的七个主题域,对合同管理内部控制评价数据仓库数据集进行模型设计,见表2。

(1)概念模型。概念模型是针对每个分析主题按照业务范围进行划分,考虑到每个合同管理内部控制评价指标所涉及的实体以及各实体之间的关系。如合同履行监控有效性分析主题中按业务范围划分为合同基本信息、合同提醒信息、合同进度等。

(2)逻辑模型。逻辑模型是概念模型的具体化,对实体的具体属性进行定义,将部分实体设计为事实表,部分实体设计为维度表。并根据实体的不同选择使用星型模型、事实星座模型或是雪花模型。若数据在重庆海事局财务云平台的其他板块涉及单位层面的人员部门信息,业务层面的固定资产信息、收支信息等时,则采用事实星座模型将多张事实表关联在一起。例如在合同履行监控有效性分析主题中,责任人信息涉及单位层面的人员信息及部门信息,通过建立事实星座模型将业务层面合同管理板块的合同信息与单位层面人员部门信息关联在一起,以服务于分析需求,具体如图所示。



合同履行监控有效性评价数据仓库逻辑模型图

由上图可知,合同履行监控有效性事实表中显示了合同ID等合同基本信息,退质保金提醒、合同结算提醒等合同监控提醒数据,其中的责任人ID、

表 2

重庆海事局合同管理内部控制评价模型设计

主题域名称	模型	事实表	维度表	主题域描述
合同审批流程完整性	雪花	合同审批流程完整性事实表	合同类型维度表、合同附件维度表	旨在审查合同审批流程是否完整,审查需要进行复核的合同项目经济合同复核表内容是否完整,审查一般合同项目经济合同审查表、复核表、公示记录是否完整
合同签订合规性	事实星座	合同签订合规性事实表、部门事实表	合同类型维度表、人员基本信息维度表	通过合同关联表,追溯到合同申请表中各个部门审批意见,审查各单位审批意见中的疑点数据
合同履行监控有效性	事实星座	合同履行监控有效性事实表、人员信息事实表	合同进度维度表、人员角色维度表、部门维度表	通过合同往来明细表追溯到(收)付款明细当中的收付情况,结合合同信息主表中的计划收付时间,检查合同履行提醒推送疑点数据
合同变更处理合规性	事实星座	合同变更处理合规性事实表、变更合同事实表	合同处理维度表、合同附件查询单维度表、合同类型维度表	旨在审查单位对合同履行情况出现变更时,是否根据规范和合同管理办法完成合同的变更处理
合同归口管理执行性	事实星座	合同归口管理执行性事实表、人员事实表、部门事实表	合同附件维度表	检查规费征稽处对合同的归档情况,包括已经归档合同的经办部门信息是否与合同经办人所属部门一致、合同附件是否完整
合同信息准确性	事实星座	合同信息准确性事实表、供应商事实表、部门事实表	合同附件维度表	旨在考察合同信息是否和实际信息一致,重点检查合同的往来账户是否与合同中供应商的银行账户信息一致
合同档案完整性	雪花	合同档案完整性事实表	合同附件维度表	旨在审查《内控规范》和《管理办法》规定需要保存的合同档案是否完整

合同进度 ID 分别作为人员维度表和合同进度维度表的外键。人员信息从单位层面内部控制评价数据仓库中的人员维度表和人员角色维度表中提取,用以考察该人员的职务。从部门维度表中提取信息,考察该人员的部门信息,方便对相关责任人在业务流程中进行分析信息推送以实现嵌入式风险预警。

(3)物理模型。物理模型又是逻辑模型的具体化,对事实表和维度表的属性进行定义。部分来源于财务云平台内部数据库的数据,可沿用之前的设置。合同管理业务信息大多存储在财务云平台内部数据库中,而合同进度、合同处理等相关附件的表名和属性都需进行定义。具体有:①SS_HT_LXJK【合同履行监控有效性事实表】。包含字段:Contract_id【合同 ID】、Contract_name【合同名称】、AlarYajin【退质保金提醒】、AlarJiesuan【合同结算提醒】、AlarShoufu【收付款提醒】、ZRR_id【责任人 ID】、CtrJinDu_id【合同进度 ID】、PlanPay_Date【计划收付款时间】、PlanPay_Money【计划收付款金额】、RealPay_Date【实际收付金额】、RecAndPayContract【款项进度】、ReYaJin_Date【退质保金时间】和 YaJin_Money【质保金额】。②WD_HT_JD【合同进度维度表】。包含字段:CtrJinDu_id【合同进度 ID】和 JinDu【合同进度】。③NK_YW_SZ_WD_YY【人员维度表】。包含字段:ID【人员编码】、VC_Name【人员名称】、I_NodeID

【权限编码】和 I_Depot【部门编码】。④NK_YW_SZ_WD_BM【部门维度表】。包含字段:ID【部门编码】和 VC_Depot【部门名称】。⑤NK_YW_SZ_WD_QX【权限维度表】。包含字段:ID【权限编码】、VC_NodeName【权限名称】和 VC_NodeEsc【权限描述】。

五、基于数据仓库的合同管理内部控制评价

1. 合同管理内部控制评价方法。合同管理内部控制评价数据仓库可以完成数据挖掘、联机分析处理、HQL 语句查询等数据分析操作。基于决策树、随机森林、K-Means 等算法对合同管理内部控制评价数据仓库中的信息进行数据挖掘,完成合同管理内部控制的疑点和风险识别并向主管人员提供风险应对方法的辅助决策信息。

合同管理内部控制评价的联机分析处理(OLAP)可以将合同内部控制评价的相关信息切片、切块、上卷、下钻等处理以满足分析需求。以合同基本信息数据立方体为例,合同基本信息数据立方体中可分为时间、部门、类型、供应商等多个维度。通过部门对合同信息进行抽取是对合同基本信息数据立方体的切片过程,可以得到某个部门下的合同基本信息。当从部门和时间的角度对合同基本信息数据立方体抽取数据时,可以得到具体某个部门在某年某个季度的合同管理情况,从而完成合同管理内部控制评价工作。HQL 语句是基于 Hive 的类 SQL

语句,能够帮助使用者对合同管理内部控制评价数据仓库进行查询分析。通过合同管理评价每个主体下不同的分析点进行查询分析,能够将数据仓库中存在疑点和风险的合同相关数据检查出来。

2. 基于Hive的合同管理内部控制情况查询分析。HQL语句是基于合同管理内部控制评价模型,通过基于每个合同管理内部控制评价主题模型,从一个或多个分析点进行分析,实现对合同签订、报送审核等流程中风险点的辨识。例如合同监控有效性评价分析主题中,对于合同监控是否有效的评价方法体现在两个方面:一方面,通过合同收付金额和时间以及计划收付金额和时间进行对比来确定是否出现了异常支付;另一方面,从合同两个重要监控提醒的角度对进入相应进度的合同项目进行查询,当合同提醒还未进行时检查合同的相关质保金等是否进行了提前支付。

(1)合同收付情况监控有效性分析。查询进度进入履行阶段的合同,分析该类合同实际收付时间、金额与计划收付时间、金额的匹配情况,当出现实际收付大于计划收付、实际收付时间早于计划收付时间时,可以认为该项合同监控出现了问题。对应分析评价的核心HQL语句如下:

```
Hive>
Select A.Contract_id, A.Contract_name, A.PlanPay_Date, A.PlanPay_Money, A.RealPay_Date, A.RealPay_Money, B.JinDu
From SS_HT_LXJK as A, WD_HT_JD as B
Where A.CtrJinDu_id=B. CtrJinDu_id and B.CtrJinDu_id='004' and (A.RealPay_Money>A.PlanPay_Money or A.RealPay_Date < A.PlanPay_Date)
Order by A Contract_id
```

在该HQL语句中,将合同履行监控有效性事实表和合同进度维度表通过join语句进行关联,检查所有合同进度ID为004即处于合同结算期间的合同,并通过and和or结合的条件查询语句对所有实际收付金额大于计划收付金额或实际收付时间早于计划收付时间的合同项目进行筛选查询,以检查出没有进行有效监控的合同项目。

(2)合同退质保金提醒有效性分析。查询还未对相关责任人进行质保金提醒的合同,分析该类合同质保金是否已经支付。对应分析评价的核心HQL语句如下:

```
Hive>
```

```
Select A.Contract_id, A.Contract_name, B.ID, B.VC_name, C.VC_NodeName, D.VC_Depot
From SS_HT_LXJK as A, NK_YW_SZ_WD_YY as B, NK_YW_SZ_WD_QX as C, NK_YW_SZ_WD_BM as D
where A.ZRR_ID = B.ID and B.I_NodeID = C.ID and B. I_Depot = D.ID and AlarYajin=0 and ReYaJin_date not null
Order by A.Contract_id
```

在该HQL语句中,将合同履行情况监控有效性事实表和人员信息事实表进行联合查询,通过and条件查询语句对退质保金提醒字段值为0的,即还未进入提醒阶段的合同项目和退质保金时间非空的(已经有了质保金退回时间的)合同项目进行查询,并通过人员信息事实表查询其责任人的相关角色、部门信息,以便对其提出风险警报。

(3)合同结算提醒有效性分析。查询还未进入结算阶段的合同,分析该类合同的结算支付是否出现提前支付等疑点数据。对应分析评价的核心HQL语句如下:

```
Hive>
Select A.Contract_id, A.Contract_name, B.ID, B.VC_name, C.VC_NodeName, D.VC_Depot
From SS_HT_LXJK as A, NK_YW_SZ_WD_YY as B, NK_YW_SZ_WD_QX as C, NK_YW_SZ_WD_BM as D
where A.ZRR_ID = B.ID and B.I_NodeID = C.ID and B. I_Depot = D.ID and AlarJieSuan=0 and RealPay_Date not null
Order by A.Contract_id
```

在该HQL语句中,也需进行联合查询,通过and条件查询语句对未进行结算提醒即字段AlarJieSuan的值为0并且实际支付时间不为空的(字段RealPay_Date不为默认值的)合同进行查询,并通过人员信息事实表查询其责任人的相关角色、部门信息,以便对其提出风险警报。

主要参考文献:

- [1]程平,范洵.基于数据仓库的行政事业单位单位层面内部控制评价——以重庆海事局为例[J].财会月刊,2019(13):71~76.
- [2]程平,姜亭杉,冯璐.数据视角下基于财务云平台的合同管理内部控制评价——以重庆海事局为例[J].财会月刊,2017(35):91~96.

作者单位:1.重庆理工大学会计学院,重庆400054; 2.重庆海事局内部控制信息化课题组,重庆400054