



引用格式:何琰. 大数据技术在审计中的应用[J]. 郑州轻工业学院学报(社会科学版), 2016,17(3):67-71.

中图分类号:F239.1 文献标识码:A

DOI:10.3969/j.issn.1009-3729.2016.03.011

文章编号:1009-3729(2016)03-0067-05

# 大数据技术在审计中的应用

## Application of big data technology in the audits

何琰

HE Yan

河南省审计厅 计算机审计中心, 河南 郑州 450003

**摘要:**大数据技术的发展,对未来审计信息化进程将产生重要影响。大数据技术引发了审计目标、审计内容、分析技术和审计思维模式的转变。应用大数据分析技术,审计方法将从数据验证性分析向数据挖掘性分析转变;应用大数据分析模式,审计方式将从发现问题向风险预警转变;应用大数据审计作业平台,可实现单机审计向云审计的转变;构建专业的审计分析队伍,可实现传统纸质账本审计向大数据审计的转变。面对大数据时代的机遇和挑战,审计工作应创新大数据工作模式,探索部门预算执行审计全覆盖的方法;应用数据分析作业平台,开展跟踪审计;多领域数据融合,开展经济责任审计;深入挖掘数据,开展宏观经济形势预测,不断提升大数据技术在审计中的有效运用。

**关键词:**

大数据技术;

审计;

数据分析;

收稿日期:2016-05-01

作者简介:何琰(1985—),女,河南省邓州市人,河南省审计厅审计师,网络工程师,主要研究方向:计算机审计。

同人力资源、自然资源一样,大数据也是一种资源且是非常重要的战略资源<sup>[1]</sup>,它将世界的本质看成是数据的集合,用数据化思维和先进的处理技术去探索海量数据之间的关系。大数据的特征可以归纳为“4个V”:容量(volume)、多样性(variety)、生成速度(velocity)、价值(value)。目前,大数据已经成为经济社会发展的巨大引擎,其在提升产业竞争力、推动商业模式创新、国情分析和社会管理等方面发挥着越来越重要的作用。大数据技术,是指从各种类型的数据中快速获得有价值信息的一种技术,其水平反映了提取有用信息的能力。鉴于此,本文拟在分析大数据对传统审计带来的挑战的基础上,探索大数据审计的方式和途径,以应对大数据时代给审计工作带来的深刻变革。

## 一、大数据对审计工作带来的影响和挑战

大数据审计是对各类电子数据、账套、文件进行审计,是计算机审计的延伸和发展,是对多行业、多领域、多种形式海量数据的电子审计方式。大数据审计既要处理结构化数据,又要处理大量的非结构化数据,对审计目标、审计内容、分析技术和审计思维模式都提出了新的、更高的要求。

### 1. 审计目标的转变

在信息化发展初期,计算机审计能够通过观察和分析发现疑点,为审计提供线索。随着大数据时代的到来,审计目标逐步转向发现线索、评估风险、关注效益等方面,审计不仅要发现违法违规问题,更多的是要揭示制度方面存在的问题,评估内控风险,通过对经济社会相关大数据的获取和分析,洞察行业整体走向,探索发展规律,对国家、行业、部门的制度出台与发展策略做出前瞻性的思考和战略性的分析。

### 2. 审计内容的转变

传统意义上的“数据”大多是指数字,如主营业务收入金额、营业费用等,由简单数字或用数字编码的简单文本组成,传统的数据分析工具基本上能够满足其需要。而大数据不再单纯指数字,其涵盖的内容更加宽泛;传统意义上的“数据”是关系型的、结构化的二维表数据,而大数据则是指在不同时间生成的复杂、异构、多样化的数据,包括xml、html、文本、音频、视频和“影子数据”(如访问日志和web搜索历史记录)等。

### 3. 分析技术的转变

顾名思义,大数据分析就是指对规模巨大的数据进行分析。大数据分析包括5个基本方面:一是可视化分析,利用图表完整展示数据分析的过程和数据链走向;二是数据挖掘算法,从海量的、随机的、不相关的数据中构建数据挖掘模型,提取有价值的、隐含在数据内部的信息;三是预测性分析能力,根据可视化分析和数据挖掘的结果,做出预测性的判断;四是语义引擎,借助一系列的工具去提取、解析和分析数据,从“文档”数据中自动提取信息;五是数据质量和数据管理,通过标准化的流程和工具对数据进行处理,保证数据分析结果的质量。

### 4. 审计思维模式的转变

大数据时代带来审计思维模式三大转变:一是实现审计全覆盖而非抽样审计,对公共资金、国有资产、国有资源和领导干部履行经济责任情况实行审计全覆盖<sup>[2]</sup>。二是审计数据的混杂性而非精准性。传统审计以财务数据为审计重点,其核心是结构化的报表和凭证信息,而大数据具有高度的数据混杂性,数据类型涵盖了结构化、非结构化和半结构化的数据,保留了最宝贵、最有价值的信息,其内涵具备更深、更广的洞察力。三是大数据主要反映事物的相关关系而非因果关系。大数据没有改变审计事务间

的因果关系,而是注重对相关关系的开发和利用,使得数据分析对因果逻辑关系的依赖性降低了,甚至更多地倾向于基于相关关系的数据分析。

## 二、大数据审计的方式和途径

面对大数据时代对审计工作带来的挑战,审计方式和途径将实现以下4个方面的转变。

### 1. 应用大数据分析技术,实现审计方法从数据验证性分析向数据挖掘性分析转变

传统的计算机审计,是通过电子数据采集转换对数据进行验证,通过构建查询分析、多维分析等方法模型进行数据分析,而应用大数据分析技术,则能够使审计数据分析逐步由传统的验证性分析向挖掘性分析转变。挖掘性分析是指采用大数据处理技术,利用数据仓库、数据挖掘和模型预测工具进行审计分析<sup>[3]</sup>,从大量数据中发现蕴涵的数据模式和规律,包括分类分析、聚类分析、关联分析和序列分析等。其中,分类分析按照数据属性,将数据分派到不同的组中;聚类分析根据数据的相似性原则,将数据划分为若干类别,使同类数据差别极小,不同类别数据差异尽可能大;关联分析可发现隐藏在数据间的相互关系,通过挖掘发现不同数据间的密切度或关系;序列分析可发现数据出现时间、出现序列的规律,分析数据间的关联性。

例如,在金融审计中,运用数据挖掘中决策树的关联规则和挖掘方法,可对不良贷款与正常贷款之间相互错分的规律进行研究,创建对不良贷款和正常贷款具有较强分类能力的条件属性指标,从而增加分类预测模型的可理解性,并在一定程度上揭示商业银行信贷资产风险等级分类的相关规律<sup>[4]</sup>;在税收征管审计中,利用聚类方法,根据聚簇规则,可将纳税户进行分组,使组内纳税户具有最大的相似性,组间纳税户具有尽可能大的相异性,从而产生聚类标准,

并按照标准对纳税户进行分行业、分税种聚类分析,确定纳税群体的总体情况和特征,为审计目标和审计重点的建立提供依据。

### 2. 应用大数据分析模式,实现审计方式从发现问题向风险预警转变

传统审计工作以发现问题为主,对经济形势进行预测分析,因而须等到相关事件发生并且形成一定规模后,再根据搜集到的足够数据进行分析研究,具有滞后性<sup>[7]</sup>。而大数据技术可通过对跨领域的大规模经济、社会行为数据进行分析,对经济社会相关异常动态实现早期关注,利用其对异常数据的敏感性实现早期预警。当前,审计可以运用大数据相关技术,对宏观经济社会风险问题展开初步分析。例如,可以综合利用政府债务审计数据、社保审计数据、宏观经济运行数据、金融市场数据等多种数据库构建大数据集合,运用数据挖掘等数据分析工具,结合多个数据库信息交叉分析与宏观经济运行相关的事件,通过基于个别领域、结构化的数据分析,不断将半结构化、非结构化的跨领域数据加入分析框架中,逐步提高大数据审计的分析能力。

### 3. 应用大数据审计作业平台,实现单机审计向云审计转变

以审计大数据为中心建设“云审计”平台,实现远程存储和移动计算,使审计机关能够通过网络接入“云”实施审计,利用大数据分析、人工智能等信息技术,解决数据采集分析和管理中存在的问题,实现审计成果共享。其次,应完善联网审计系统,逐步建立预算、执行、财政、地税、社会保障、医疗机构、公积金等重要行业和部门的审计实时监督系统。再次,应建设审计数据综合分析平台,运用大数据技术,加大业务数据与财务数据、单位数据与行业数据,以及跨行业、跨领域数据的综合比对和关联分析,提高运用信息化技术查核问题、评价判断、宏观分

析的能力。最后,应推广“总体分析、发现疑点、分散核实、系统研究”的审计模式。

#### 4. 构建专业的审计分析队伍,实现传统纸质账本审计向大数据审计转变

审计工作应实现“六大转变”,即由单点离散审计向多点联动审计转变、由局部审计向全覆盖审计转变、由静态审计向静态与动态审计相结合转变、由事后审计向事后与事中审计相结合转变、由现场审计向现场审计与非现场审计相结合转变、由微观审计向微观与宏观审计相结合转变。为此,需要在组织方式、人员结构、思维方式等方面与之相适应。在组织方式上,应尝试开展无项目审计,依托审计数据中心积累的数据资源,横向关联比对分析,纵向深入挖掘分析,从数据中发现审计疑点和线索。在人员结构上,应不断提升“四种能力”,即大数据分析能力、综合研究能力、创新能力和跨领域知识运用能力,不断加强对大数据先进理念和前沿技术的学习,掌握大数据分析方法,提升审计人员综合素质。在思维方式上,应培养“数据先行”意识,以数据为核心,使数据分析在审计工作开展前先行实施,根据数据分析结果,有重点、有步骤、有深度地在审计实施过程中进行核查验证、追踪线索、发现问题,全面深化大数据技术在审计工作中的应用。

### 三、大数据技术在审计中的应用

大数据技术在审计中的应用广泛,具体来说,主要表现在以下4个方面。

#### 1. 创新大数据工作模式,探索部门预算执行审计全覆盖

传统的以处(科)室为单位的孤立审计,已不适应大数据审计的要求,需要打破部门界限,以审计项目为管理主线成立大审计组,进行扁平化管理。大审计组分领导小组、数据分析组和问题核查组三个层面:领导小组负责审计方

案的制定、领导决策等,由牵头机构组建;数据分析组由审计能手和计算机操作骨干组成,具体负责审计实施,筛查审计疑点和线索;问题核查组由审计一线人员组成,进行实地延伸核查。例如,在“同级审”项目中,采集各部门预算执行审计数据,通过对数据进行集中分析,发现是否存在部分预算项目缺乏科学论证、影响预算执行的问题;结合各预算部门的财务数据,发现是否存在预算项目因连年结转的情况下仍然安排新增预算、造成资金闲置的问题;通过对数据进行宏观整体分析,发现是否存在预算执行效率不高、分配下达预算不及时、拨付转移支付资金超期等情况。

#### 2. 应用数据分析作业平台,开展跟踪审计

通过建设审计数据综合分析平台,搭建关系国计民生的重点行业联网审计系统,用 Hadoop 等专业工具处理半结构化、非结构化数据,规范高效地汇集和处理大规模数据信息。例如,在地税审计中,可利用地税联网审计系统,集中进行全省地税数据整理分析,探索“数据集中采集、集中统一分析、疑点分布落实、资源充分共享”的大数据审计模式,实现全省联动审计;通过构建查询分析模型、多维分析模型,固化审计思路,对地税数据进行全面分析,重点选择税款征收、税款缴纳和税款管理等信息,在财政审计中,运用部门预算执行分析平台,将独立运行、分散存储的各部门预算编制、预算指标、部门决算、部门财务核算等业务数据,进行标准化清洗和整理,同时与组织预算执行数据进行整合,实现各类资金的横向关联,对资金分配结构、资金使用流向、资金管理情况进行总体分析,全面反映预算执行整体情况,实现对预算单位的审计监督全覆盖。

#### 3. 多领域数据融合,开展经济责任审计

运用关联分析,找出数据间的相互联系,分析关联规则,发现异常联系和异常数据,寻找审

计疑点。在经济责任审计中,可利用财政、税务、社保、公安、工商、养老、房管、教育、培训等数据,进行纵向比对、横向关联分析。例如,可将城镇居民医保、城镇职工医保、新农合等信息进行综合分析,检查是否存在同时参保、重复报销、虚假报销、套取资金等情况;将小额担保贷款与人社部门行政事业单位人员信息进行关联分析,检查是否存在违规向行政事业单位人员发放担保贷款和进行财政贴息的问题;将职业技能培训补贴、创业培训补贴与企业法人和行政事业单位人员进行关联分析,检查是否存在向不符合条件人员发放就业补贴的问题;将人社、农业、扶贫、教育、民政管理的各项培训资金进行关联分析,检查是否存在虚报冒领财政资金的问题;将义务教育、中职中专、中学学籍数据与中职助学金发放、免学费补贴等信息进行关联分析,检查是否存在虚报冒领或者套取中职助学金、免学费补贴等专项资金的问题。

#### 4. 深入挖掘数据,开展宏观经济形势预测

利用数据仓库、数据挖掘和模型预测工具进行审计分析,运用回归分析等统计方法,分析趋势变动或数据关系,发现总体情况和规律,进而揭示体制机制性问题。例如,在“三公经费”专项审计中,可通过各预算单位财务核算数据提取“三公经费”预算编制和实际支出情况,结合单位人员花名册,构建统计分析模型,进行结构分析和时间序列分析,横向比对所有预算单位“三公经费”的支出规模、各项支出占比、人均支出金额与年度支出趋势等,分析其压缩控

制比例是否符合要求,各项支出比例和人均支出金额是否存在异常和明显偏离,进而挖掘行业性、趋势性问题,并对离群点进行延伸,发现违规违纪问题线索等。再如,利用城镇或农村居民家庭人均收入指数和各地制定的最低收入保障标准进行分析,查找人均收入指数与最低收入标准间的差异,分析各地低收入标准制定是否科学,为政府制定宏观经济政策提供参考。

综上所述,大数据技术的发展对审计工作提出了新的、更高的要求,也为审计提供了新的工具。大数据审计应从数据、人才和技术等方面逐步积累基础资源,创新工作思路,研究技术方法,培养高素质人才,不断提升大数据技术在审计中的有效运用。

#### 参考文献:

- [1] 怀进鹏. 大数据是国家战略资源[J]. 中国经济和信息化,2013(8):49.
- [2] 秦荣生.《关于完善审计制度若干重大问题的框架意见》学习笔谈摘要[J]. 审计研究,2016(1):3.
- [3] 李玲,刘汝焯. 计算机数据审计[M]. 北京:清华大学出版社,2012:4-36.
- [4] 中华人民共和国审计署. 基于数据挖掘的贷款风险等级分类真实性审计研究[J]. 审计研究报告,2012(11):2.
- [5] 中华人民共和国审计署. 大数据在公共管理中的运用与审计启示[R]. 审计研究报告,2013(7):1.