



(12)发明专利申请

(10)申请公布号 CN 107945092 A
(43)申请公布日 2018.04.20

(21)申请号 201711330849.5

(22)申请日 2017.12.13

(71)申请人 成都市审计局

地址 610000 四川省成都市高新区锦城大道366号

(72)发明人 曾军 程世斌 孙勇滨

(74)专利代理机构 成都行之专利代理事务所
(普通合伙) 51220

代理人 戴勇灵

(51) Int. Cl.

G06Q 50/26(2012.01)

G06F 17/30(2006.01)

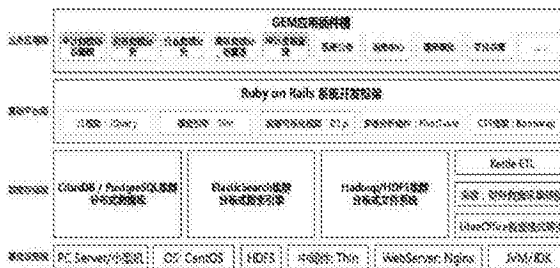
权利要求书1页 说明书6页 附图1页

(54)发明名称

用于审计领域的大数据综合管理方法及系统

(57)摘要

本发明公开了用于审计领域的大数据综合管理方法及系统,包括如下步骤:步骤A)审计人员通过数据库备份恢复的方式获取审计对象的原始数据;步骤B)依据审计署颁发的相应的实务公告的指导,采用兼容市场上90%以上财务账套软件的采集工具对原始数据进行结构分析、数据清洗、数据转换和数据汉化操作,将原始数据转换为大数据平台专用标准数据表;步骤C)将大数据平台专用标准数据表通过ETL工具导入到本系统的分布式数据库中,再将这些数据创建反向数据搜索索引;步骤D)上述步骤完成后,审计人员进行日常的数据分析操作。本发明采用上述原理,自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低。



CN 107945092 A

1. 用于审计领域的大数据综合管理方法,其特征在于,包括如下步骤:

步骤A) 审计人员通过数据库备份恢复的方式获取审计对象的原始数据;

步骤B) 依据审计署颁发的相应的实务公告的指导,采用兼容市场上90%以上财务账套软件的采集工具对原始数据进行结构分析、数据清洗、数据转换和数据汉化操作,将原始数据转换成为大数据平台专用标准数据表;

步骤C) 将大数据平台专用标准数据表通过ETL工具导入到本系统的分布式数据库中,再将这些数据创建反向数据搜索索引;

步骤D) 上述步骤完成后,审计人员进行日常的数据分析操作。

2. 根据权利要求1所述的用于审计领域的大数据综合管理方法,其特征在于,步骤B)中采集工具通过账套、数据库或数据库备份三种方式采集数据。

3. 根据权利要求1所述的用于审计领域的大数据综合管理方法,其特征在于,步骤B)中采集工具内置至少100个采集模板,通过采集模板完成数据的采集、结构匹配、表头汉化和数据清洗的过程,从而将原始数据转换成大数据平台专用标准数据表。

4. 根据权利要求3所述的用于审计领域的大数据综合管理方法,其特征在于,标准数据表能够与A0操作系统的epkg格式相互转换。

5. 根据权利要求1所述的用于审计领域的大数据综合管理方法,其特征在于,步骤D)中审计人员进行日常的数据分析操作步骤如下:

步骤D1) 用户在系统中搜索关键字,系统显示该关键字在全库中的搜索结果;

步骤D2) 用户选择自己感兴趣的数据表,查看搜索到的明细数据;

步骤D3) 用户在明细数据的基础上,运用系统提供的数据筛选功能,自定义筛选条件,进一步缩小数据搜索的范围;

步骤D4) 用户基于上述步骤产生的数据,自助定义不同的维度和汇总的条件,在系统中进行数据多维分析,并观察分析的结果;

步骤D5) 用户根据观察数据分析的结果,若满意则将检索到的数据导出成为电子表格文件,以便在系统外进行进一步的处理,若不满意则返回步骤D1),使用更加精细的关键字进行进一步的数据分析和探索。

6. 用于审计领域的大数据综合管理系统,其特征在于,包括

业务应用层:能够进行审计数据综合管理、财务数据分析、行业数据分析、通过数据分析算法进行数据分析、进行审计思路管理、进行系统公告、设置消息中心、请示审批和文件共享单元;

基础平台层:具备用于开发、部署和维护web应用程序的Ruby on Rails系统开发框架;

数据存储层:包括用于业务应用层调用的CitusDB/PostgreSQL集群分布式数据库、ElasticSearch集群分布式搜索引擎、Hadoop/HDFS集群分布式文件系统、采客和财务数据采集转换单元以及LibreOffice数据格式转换单元;

基础设施层:包括PC Server/小型机、OS操作系统、HDFS分布式文件系统、WebServer服务器和JVM/JDK开发包,这些系统运行所需要的软硬件设备。

7. 根据权利要求6所述的用于审计领域的大数据综合管理系统,其特征在于,Ruby on Rails系统开发框架包括JS框架、模板引擎、数据可视化框架、多维分析组件和CSS框架。

用于审计领域的大数据综合管理方法及系统

技术领域

[0001] 本发明涉及计算机领域中的数据库技术,具体涉及用于审计领域的大数据综合管理方法及系统。

背景技术

[0002] 审计是由国家授权或接受委托的专职机构和人员,依照国家法规、审计准则和会计理论,运用专门的方法,对被审计单位的财政、财务收支、经营管理活动及其相关资料的真实性、正确性、合规性、合法性、效益性进行审查和监督,评价经济责任,鉴证经济业务,用以维护财经法纪、改善经营管理、提高经济效益的一项独立性的经济监督活动。市场上针对审计业务开发的软件不多,专业性不强,不够贴合审计实际工作需要。不能够兼容市场上的大部分财务软件,只能够提供简单的数据查询和SQL语言分析,数据容量小、查询效率低、反应速度慢,缺少关联分析和图形化分析的功能。

发明内容

[0003] 本发明所要解决的技术问题是提高审计领域软件的自动化程度,目的在于提供用于审计领域的大数据综合管理方法及系统,自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低。

[0004] 本发明通过下述技术方案实现:

[0005] 用于审计领域的大数据综合管理方法,包括如下步骤:

[0006] 步骤A) 审计人员通过数据库备份恢复的方式或者其它方式获取审计对象的原始数据;

[0007] 步骤B) 依据审计署颁发的相应的实务公告的指导,采用兼容市场上90%以上财务账套软件的采集工具对原始数据进行结构分析、数据清洗、数据转换和数据汉化操作,将原始数据转换为大数据平台专用标准数据表;

[0008] 步骤C) 将大数据平台专用标准数据表通过ETL工具导入到本系统的分布式数据库中,再将这些数据创建反向数据搜索索引;

[0009] 步骤D) 上述步骤完成后,审计人员进行日常的数据分析操作。

[0010] 现有市面上的财务软件多种多样,且各个软件之间数据格式各不相同,各个财务软件之间并不兼容,而审计局需要对各个公司的财务情况进行审计,而现有市场上针对审计业务开发的软件不多,专业性不强,不够贴合审计实际工作需要。不能够兼容市场上的大部分财务软件,只能够提供简单的数据查询和SQL语言分析,数据容量小、查询效率低、反应速度慢,缺少关联分析和图形化分析的功能。该情况大大增加了审计局的审计工作量。本方法采用兼容市场上90%以上财务账套软件的采集工具对获取审计对象的原始数据进行处理,使得原始数据变为统一格式的数据,方便后期审计对象的数据采集、管理、查询和分析,该方法支持多媒体以及其它格式的数据,又能够达到关系型数据库的数据分析能力,自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低。本方法中还

创建反向数据搜索索引,使数据检索的方式从复杂的SQL查询,变为像Google、百度一样的傻瓜式搜索,审计人员只需要输入几个关键字,就能迅速查找到自己想要的信息。同时,平台还提供自动化、可视化的数据分析功能,大大降低了审计人员进行电子数据分析的门槛,有效解决了审计能力参差不齐的难题,从而更有利大数据审计分析方法的推广和应用。

[0011] 优选的,步骤B)中采集工具通过账套、数据库或数据库备份三种方式采集数据。

[0012] 优选的,步骤B)中采集工具内置至少100个采集模板,通过采集模板完成数据的采集、结构匹配、表头汉化和数据清洗的过程,从而将原始数据转换成大数据平台专用标准数据表。

[0013] 优选的,标准数据表能够与A0操作系统的epkg格式相互转换。本方法能够兼容市场上200多套,400多个不同版本的财务软件,实现在线采集被审计单位财务、业务、报表及其它类型数据;其财务数据包可与审计署《A0审计现场管理系统》相互转换;能够百度式在线查询财务会计凭证、科目余额、账簿、会计报表等数据,并实现财务数据与行业数据的关联查询和图表分析。本方法能够有效加强审计力度,提高监管效率,防止政府资金、资源的违规使用和浪费;利用大数据手段实时、有效监控各行业领域异常业务发生情况,维护社会公平正义;迅速挖掘、打击各类腐败、违纪违法行为,促进政治生态建设健康发展;提高审计信息化建设水平,提升数据审计整体能力。

[0014] 优选的,步骤D)中审计人员进行日常的数据分析操作步骤如下:

[0015] 步骤D1)用户在系统中搜索关键字,系统显示该关键字在全库中的搜索结果;

[0016] 步骤D2)用户选择自己感兴趣的数据表,查看搜索到的明细数据;

[0017] 步骤D3)用户在明细数据的基础上,运用系统提供的数据筛选功能,自定义筛选条件,进一步缩小数据搜索的范围;

[0018] 步骤D4)用户基于上述步骤产生的数据,自助定义不同的维度和汇总的条件,在系统中进行数据多维分析,并观察分析的结果;

[0019] 步骤D5)用户根据观察数据分析的结果,若满意则将检索到的数据导出成为电子表格文件,以便在系统外进行进一步的处理,若不满意则返回步骤D1),使用更加精细的关键字进行进一步的数据分析和探索。

[0020] 用于审计领域的大数据综合管理系统,包括

[0021] 业务应用层:能够进行审计数据综合管理、财务数据分析、行业数据分析、通过数据分析算法进行数据分析、进行审计思路管理、进行系统公告、设置消息中心、请示审批和文件共享单元;

[0022] 基础平台层:具备用于开发、部署和维护web应用程序的Ruby on Rails系统开发框架;

[0023] 数据存储层:包括用于业务应用层调用的CitrusDB/PostgreSQL集群分布式

[0024] 数据库、ElasticSearch集群分布式搜索引擎、Hadoop/HDFS集群分布式文件系统、采客和财务数据采集转换单元以及LibreOffice数据格式转换单元;

[0025] 基础设施层:包括PC Server/小型机、OS操作系统、HDFS分布式文件系统、WebServer服务器和JVM/JDK开发包,这些系统运行所需要的软硬件设备。

[0026] Ruby on Rails系统开发框架包括JS框架、模板引擎、数据可视化框架、多维分析组件和CSS框架。

[0027] 本系统采用分层式的技术架构,包括基础设施层、数据存储层、基础平台层以及业务应用层,通过B/S架构,最终用户仅仅需要使用浏览器即可以使用系统,使数据检索的方式从复杂的SQL查询,变为像Google、百度一样的傻瓜式搜索,审计人员只需要输入几个关键字,就能迅速查找到自己想要的信息,自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低。本系统支持集中式部署方式,部署快捷,升级方便;服务端支持64位CentOS 7.1及以上版本操作系统;客户端支持IE11.0,Chrome56,Firefox52,Safari10,Opera40及以上版本浏览器,兼容性好。

[0028] 系统基于分布式大数据的技术框架进行数据的存储与管理,系统平台层及业务应用层基于Ruby on Rails开发框架进行研发,并使用插件机制来对业务应用层的各个应用APP进行解耦;平台层会将常用的算法和功能进行抽象,形成各种业务引擎,以便上层的业务应用进行复用,开发、部署和维护web应用程序简单方便。同时,平台还提供自动化、可视化的数据分析功能,大大降低了审计人员进行电子数据分析的门槛,有效解决了审计能力参差不齐的难题,从而更有利大数据审计分析方法的推广和应用。

[0029] 本发明与现有技术相比,具有如下的优点和有益效果:

[0030] 1、本发明采用兼容市场上90%以上财务账套软件的采集工具对获取审计对象的原始数据进行处理,使得原始数据变为统一格式的数据,方便后期审计对象的数据采集、管理、查询和分析,该方法支持多媒体以及其它格式的数据,又能够达到关系型数据库的数据分析能力,自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低。

[0031] 2、本发明采用分层式的技术架构,包括基础设施层、数据存储层、基础平台层以及业务应用层,通过B/S架构,最终用户仅仅需要使用浏览器即可以使用系统,使数据检索的方式从复杂的SQL查询,变为像Google、百度一样的傻瓜式搜索,审计人员只需要输入几个关键字,就能迅速查找到自己想要的信息,自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低。

[0032] 3、本发明能够有效加强审计力度,提高监管效率,防止政府资金、资源的违规使用和浪费;利用大数据手段实时、有效监控各行业领域异常业务发生情况,维护社会公平正义;迅速挖掘、打击各类腐败、违纪违法行为,促进政治生态建设健康发展;提高审计信息化建设水平,提升数据审计整体能力。

附图说明

[0033] 此处所说明的附图用来提供对本发明实施例的进一步理解,构成本申请的一部分,并不构成对本发明实施例的限定。在附图中:

[0034] 图1为本发明框架图。

具体实施方式

[0035] 为使本发明的目的、技术方案和优点更加清楚明白,下面结合实施例和附图,对本发明作进一步的详细说明,本发明的示意性实施方式及其说明仅用于解释本发明,并不作为对本发明的限定。

[0036] 实施例1:

[0037] 如图1所示,本发明包括用于审计领域的大数据综合管理方法,包括如下步骤:

[0038] 步骤A) 审计人员通过数据库备份恢复的方式或者其它方式获取审计对象的原始数据;

[0039] 步骤B) 依据审计署颁发的相应的实务公告的指导,采用兼容市场上90%以上财务账套软件的采集工具对原始数据进行结构分析、数据清洗、数据转换和数据汉化操作,将原始数据转换成为大数据平台专用标准数据表;

[0040] 步骤C) 将大数据平台专用标准数据表通过ETL工具导入到本系统的分布式数据库中,再将这些数据创建反向数据搜索索引;

[0041] 步骤D) 上述步骤完成后,审计人员进行日常的数据分析操作。

[0042] 现有市面上的财务软件多种多样,且各个软件之间数据格式各不相同,各个财务软件之间并不兼容,而审计局需要对各个公司的财务情况进行审计,而现有市场上针对审计业务开发的软件不多,专业性不强,不够贴合审计实际工作需要。不能够兼容市场上的大部分财务软件,只能提供简单的数据查询和SQL语言分析,数据容量小、查询效率低、反应速度慢,缺少关联分析和图形化分析的功能。该情况大大增加了审计局的审计工作量。本方法采用兼容市场上90%以上财务账套软件的采集工具对获取审计对象的原始数据进行处理,使得原始数据变为统一格式的数据,方便后期审计对象的数据采集、管理、查询和分析,该方法支持多媒体以及其它格式的数据,又能够达到关系型数据库的数据分析能力,自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低。本方法中还创建反向数据搜索索引,使数据检索的方式从复杂的SQL查询,变为像Google、百度一样的傻瓜式搜索,审计人员只需要输入几个关键字,就能迅速查找到自己想要的信息。同时,平台还提供自动化、可视化的数据分析功能,大大降低了审计人员进行电子数据分析的门槛,有效解决了审计能力参差不齐的难题,从而更有利大数据审计分析方法的推广和应用。

[0043] 实施例2:

[0044] 本实施例在实施例1的基础上优选如下:步骤B) 中采集工具通过账套、数据库或数据库备份三种方式采集数据。

[0045] 步骤B) 中采集工具内置至少100个采集模板,通过采集模板完成数据的采集、结构匹配、表头汉化和数据清洗的过程,从而将原始数据转换成大数据平台专用标准数据表。

[0046] 标准数据表能够与A0操作系统的epkg格式相互转换。本方法能够兼容市场上200多套,400多个不同版本的财务软件,实现在线采集被审计单位财务、业务、报表及其它类型数据;其财务数据包可与审计署《A0审计现场管理系统》相互转换;能够百度式在线查询财务会计凭证、科目余额、账簿、会计报表等数据,并实现财务数据与行业数据的关联查询和图表分析。本方法能够有效加强审计力度,提高监管效率,防止政府资金、资源的违规使用和浪费;利用大数据手段实时、有效监控各行业领域异常业务发生情况,维护社会公平正义;迅速挖掘、打击各类腐败、违纪违法行为,促进政治生态建设健康发展;提高审计信息化建设水平,提升数据审计整体能力。

[0047] 步骤D) 中审计人员进行日常的数据分析操作步骤如下:

[0048] 步骤D1) 用户在系统中搜索关键字,系统显示该关键字在全库中的搜索结果;

[0049] 步骤D2) 用户选择自己感兴趣的数据表,查看搜索到的明细数据;

[0050] 步骤D3) 用户在明细数据的基础上,运用系统提供的数据筛选功能,自定义筛选条

件,进一步缩小数据搜索的范围;

[0051] 步骤D4) 用户基于上述步骤产生的数据,自助定义不同的维度和汇总的条件,在系统中进行数据多维分析,并观察分析的结果;

[0052] 步骤D5) 用户根据观察数据分析的结果,若满意则将检索到的数据导出成为电子表格文件,以便在系统外进行进一步的处理,若不满意则返回步骤D1),使用更加精细的关键词进行进一步的数据分析和探索。

[0053] 该搜索方式自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低,大大降低了审计人员进行电子数据分析的门槛,有效解决了审计能力参差不齐的难题,从而更有利大数据审计分析方法的推广和应用。

[0054] 实施例3:

[0055] 本实施例在上述实施例的基础上优选如下:用于审计领域的大数据综合管理系统,包括

[0056] 业务应用层:能够进行审计数据综合管理、财务数据分析、行业数据分析、通过数据分析算法进行数据分析、进行审计思路管理、进行系统公告、设置消息中心、请示审批和文件共享单元;

[0057] 基础平台层:具备用于开发、部署和维护web应用程序的Ruby on Rails系统开发框架;

[0058] 数据存储层:包括用于业务应用层调用的CitusDB/PostgreSQL集群分布式数据库、ElasticSearch集群分布式搜索引擎、Hadoop/HDFS集群分布式文件系统、采客和财务数据采集转换单元以及LibreOffice数据格式转换单元;

[0059] 基础设施层:包括PC Server/小型机、OS操作系统、HDFS分布式文件系统、WebServer服务器和JVM/JDK开发包,这些系统运行所需要的软硬件设备。

[0060] Ruby on Rails系统开发框架包括JS框架、模板引擎、数据可视化框架、多维分析组件和CSS框架。

[0061] 本系统采用分层式的技术架构,包括基础设施层、数据存储层、基础平台层以及业务应用层,通过B/S架构,最终用户仅仅需要使用浏览器即可以使用系统,使数据检索的方式从复杂的SQL查询,变为像Google、百度一样的傻瓜式搜索,审计人员只需要输入几个关键字,就能迅速查找到自己想要的信息,自动化程度高、操作简单、反应速度快、关联分析好、对审计人员计算机水平要求低。本系统采用NoSQL与关系数据库混合的数据库架构和分布式并行计算技术,在保持大量用户并发下的平台访问速度的同时能够有效提升海量数据的查询分析速度(以社保数据为例,从上亿条数据中检索时间不超过1秒),即支持多媒体以及其它格式的数据,又能够达到关系型数据库的数据分析能力。本系统支持集中式部署方式,部署快捷,升级方便;服务端支持64位CentOS 7.1及以上版本操作系统;客户端支持IE11.0,Chrome56,Firefox52,Safari10,Opera40及以上版本浏览器,兼容性好。

[0062] 系统基于分布式大数据的技术框架进行数据的存储与管理,系统平台层及业务应用层基于Ruby on Rails开发框架进行研发,并使用插件机制来对业务应用层的各个应用APP进行解耦;平台层会将常用的算法和功能进行抽象,形成各种业务引擎,以便上层的业务应用进行复用,开发、部署和维护web应用程序简单方便。同时,平台还提供自动化、可视化的数据分析功能,大大降低了审计人员进行电子数据分析的门槛,有效解决了审计能力

参差不齐的难题,从而更有利大数据审计分析方法的推广和应用。

[0063] 本系统部署于审计的专网或者电子政务网中,需要一台独立的应用服务器和多台数据服务器组成集群;根据实际情况的需要,可以考虑创建分布式数据库系统集群、分布式搜索引擎集群以及分布式文件存储集群。

[0064] 以上所述的具体实施方式,对本发明的目的、技术方案和有益效果进行了进一步详细说明,所应理解的是,以上所述仅为本发明的具体实施方式而已,并不用于限定本发明的保护范围,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

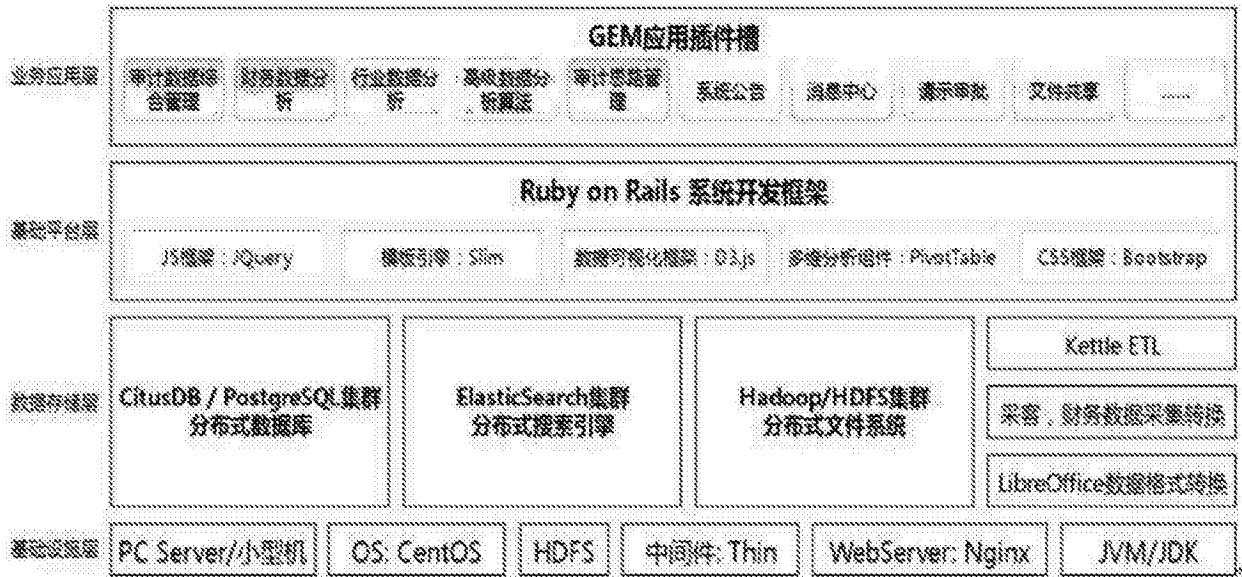


图1