



(12) 发明专利申请

(10) 申请公布号 CN 105069025 A

(43) 申请公布日 2015. 11. 18

(21) 申请号 201510421285. 0

(22) 申请日 2015. 07. 17

(71) 申请人 浪潮通信信息系统有限公司

地址 250101 山东省济南市高新区浪潮路
1036 号

(72) 发明人 肖雪 龚松柏 沈林江

(74) 专利代理机构 济南信达专利事务所有限公
司 37100

代理人 姜明

(51) Int. Cl.

G06F 17/30(2006. 01)

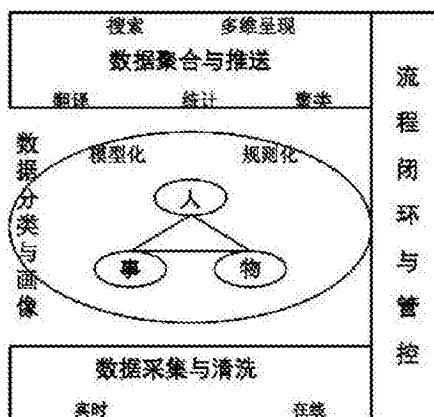
权利要求书1页 说明书4页 附图1页

(54) 发明名称

一种大数据的智能聚合可视化与管控系统

(57) 摘要

本发明公开了一种大数据的智能聚合可视化与管控系统,包括数据采集与清洗、数据分类与画像、数据聚合与推送、流程闭环与管控四个层面的功能群组。该大数据的智能聚合可视化与管控系统与现有技术相比,可以通过搜索方式,自动化、智能化支撑不断爆发的新的数字化服务需求及支撑要求,可扩展性强、可维护性高、二次开发成本低,实用性强,易于推广。



1. 一种大数据的智能聚合可视化与管控系统,其特征在於,包括数据采集与清洗、数据分类与画像、数据聚合与推送、流程闭环与管控四个层面的功能群组,其中:

数据采集与清洗负责从其他系统或来源实时在线获取数据,根据数据清洗规则进行数据清洗及格式化,对清洗出的异常数据进行及时发现及关联分析,回答数据从哪里来;

数据分类与画像负责对采集清洗后的数据进行解剖,包括显性数据的模型化、隐性关系的规则化,对数据进行标签,生成画像,回答数据是什么;

数据聚合与推送负责面向角色对数据画像进行包括翻译、统计、聚类的算法聚合,通过搜索,多维呈现交付数据,回答数据到哪里去;

流程闭环与管控负责管理数据,包括面向角色、安全、质量、闭环,提供系统内数据的自流转。

2. 根据权利要求 1 所述的一种大数据的智能聚合可视化与管控系统,其特征在於,所述数据采集与清洗群组中,数据采集包括通信设备数据采集能力、离线数据批量采集能力、在线数据交互采集能力、互联网数据采集能力、流式数据实时采集能力项,数据清洗包括清洗规则管理、异常数据管理、数据格式化管理、调度管理能力项。

3. 根据权利要求 2 所述的一种大数据的智能聚合可视化与管控系统,其特征在於,所述数据分类与画像群组中,包括已知数据的模型化能力、未知数据的规则化能力,其中已知数据的模型化能力负责提供多维数据模型管理、数据脱敏管理、实时计算能力项,未知数据的规则化负责提供数据的切片、数据关系的挖掘能力项。

4. 根据权利要求 3 所述的一种大数据的智能聚合可视化与管控系统,其特征在於,所述显性数据模型化是指面向用户、网元、终端、业务、网络、区域,建立多维度多视角多口径的模型基础;隐性关系规则化是指基于时间、位置的数据切片、分词统计,挖掘数据关系规则。

5. 根据权利要求 1 所述的一种大数据的智能聚合可视化与管控系统,其特征在於,所述数据分类与画像群组对生产数据、过程数据、组织数据、非结构化数据、互联网数据进行大类、中类、小类三级标签,针对模型化及规则化后的数据,进行数据标签分析、数据定义,生成数据画像,提供前台模糊搜索呈现,同时提供针对大类/中类标签特征数据的实时过滤、实时计算,提供定制化数据画像能力,保证数据画像的时效性及可扩展性。

6. 根据权利要求 1 所述的一种大数据的智能聚合可视化与管控系统,其特征在於,所述数据聚合与推送群组中,数据聚合包括聚合算法管理、搜索能力、语义词义管理,数据推送包括数据服务能力、多维呈现能力;该群组通过对包括用户、网元、终端、业务、网络、区域的数据画像数据的大类、中类、小类三级标签进行包括翻译、统计、聚类的算法聚合,利用搜索,结合聚合数据、聚合工单、聚合工具、聚合应用、聚合业务多维呈现,保证数据聚合的相关性、工具及应用的协同。

7. 根据权利要求 1 所述的一种大数据的智能聚合可视化与管控系统,其特征在於,所述流程闭环与管控组群提供质量规则管理、角色管理、流程管理的基础配置管理能力,通过端到端环节探针收集、质量分析及派单、后评估的服务能力,发现数据质量问题及时告警派单,并能通过数据聚合结果提供派单处理能力,提供系统内数据的自流转及质量保障。

一种大数据的智能聚合可视化与管控系统

技术领域

[0001] 本发明涉及大数据领域,具体地说是一种实用性强、大数据的智能聚合可视化与管控系统。

背景技术

[0002] 移动互联网时代已经到来,当前正处于一个新旧交替、五彩缤纷、未知大于已知的时代。对于移动互联网用户来说,用户个性化、差异化、定制化、端到端的服务需求越来越凸显;而对于传统运营商来说,移动互联网业务对语音、短信等传统业务的替代是颠覆性的,因此,传统运营商需要加速转型,除了做好自己的网络保障及智能管道外,还应该要做好端到端服务支撑及数字化服务,提升客户感知,探索新途径,寻找新增长点。

[0003] 当前移动通信网存在 B、O、M 三域数据,一方面目前三域数据割裂,即使在具体一域内仍然存在烟囱结构,数据孤岛,未整合;另一方面大数据分析 & 处理技术手段难以支撑实时性要求;并且,对于数据的呈现分析手段仍然停留在传统报表及指标,甚至部分系统中已经实现数万张报表,可用性差,响应时间长。这种情况下,数据隐藏在具体系统数据库中,数据一致性准确性差,传统方式下,需要根据具体应用需求及场景,接入各系统数据,生成分析场景数据模型,开发报表及分析功能,不但开发周期长,而且再次生成新的数据孤岛,无法响应实时的个性化、差异化、定制化分析要求,无法提供面向客户的数字化服务需求及端到端服务支撑。

[0004] 基于此,现提供一种大数据的智能聚合可视化与管控系统。在本方法中,首先解决数据孤岛问题,将数据进行标准化共享;其次,面向个性化、差异化、定制化的分析要求,迫切需要建立一套面向用户、网元、业务、终端、网络、区域的一人一档一物一档一事一档的数据画像体系,所需即所在;最后,面向客户的数字化服务需求及端到端服务支撑,则需要能智能聚合相关分析数据,自动形成分析场景,快速提供新的数字化服务能力,所见即所得。

发明内容

[0005] 本发明的技术任务是针对以上不足之处,提供一种实用性强、大数据的智能聚合可视化与管控系统。

[0006] 一种大数据的智能聚合可视化与管控系统,其结构包括数据采集与清洗、数据分类与画像、数据聚合与推送、流程闭环与管控四个层面的功能群组,其中

数据采集与清洗负责从其他系统或来源实时在线获取数据,根据数据清洗规则进行数据清洗及格式化,对清洗出的异常数据进行及时发现及关联分析,回答数据从哪里来;

数据分类与画像负责对采集清洗后的数据进行解剖,包括显性数据的模型化、隐性关系的规则化,对数据进行标签,生成画像,回答数据是什么;

数据聚合与推送负责面向角色对数据画像进行翻译、统计、聚类等算法聚合,通过搜索,多维呈现交付数据,回答数据到哪里去;

流程闭环与管控负责管理数据,包括面向角色、安全、质量、闭环,提供系统内数据的自

流转。

[0007] 所述数据采集与清洗群组中,数据采集包括通信设备数据采集能力、离线数据批量采集能力、在线数据交互采集能力、互联网数据采集能力、流式数据实时采集能力等能力项,数据清洗包括清洗规则管理、异常数据管理、数据格式化管理、调度管理能力项。

[0008] 所述数据分类与画像群组中,包括已知数据的模型化能力、未知数据的规则化能力,其中已知数据的模型化能力负责提供多维数据模型管理、数据脱敏管理、实时计算能力项,未知数据的规则化负责提供数据的切片、数据关系的挖掘能力项。

[0009] 所述显性数据模型化是指面向用户、网元、终端、业务、网络、区域,建立多维度多视角多口径的模型基础;隐性关系规则化是指基于时间、位置的数据切片、分词统计,挖掘数据关系规则。

[0010] 所述数据分类与画像群组对生产数据、过程数据、组织数据、非结构化数据、互联网数据进行大类、中类、小类三级标签,针对模型化及规则化后的数据,进行数据标签分析、数据定义,生成数据画像,提供前台模糊搜索呈现,同时提供针对大类/中类标签特征数据的实时过滤、实时计算,提供定制化数据画像能力,保证数据画像的时效性及可扩展性。

[0011] 所述数据聚合与推送群组中,数据聚合包括聚合算法管理、搜索能力、语义词义管理,数据推送包括数据服务能力、多维呈现能力;该群组通过对包括用户、网元、终端、业务、网络、区域的数据画像数据的大类、中类、小类三级标签进行包括翻译、统计、聚类的算法聚合,利用搜索,结合聚合数据、聚合工单、聚合工具、聚合应用、聚合业务多维呈现,保证数据聚合的相关性、工具及应用的协同。

[0012] 所述流程闭环与管控群组提供质量规则管理、角色管理、流程管理的基础配置管理能力,通过端到端环节探针收集、质量分析及派单、后评估的服务能力,发现数据质量问题及时告警派单,并能通过数据聚合结果提供派单处理能力,提供系统内数据的自流转及质量保障。

[0013] 本发明的一种大数据的智能聚合可视化与管控系统,与现有技术相比,具有以下优点:

本发明提出的一种大数据的智能聚合可视化与管控系统,可以通过搜索方式,自动化、智能化支撑不断爆发的新的数字化服务需求及支撑要求,可扩展性强、可维护性高、二次开发成本低。相比现有技术方式下,动辄需要新开发上万张报表,大大降低了支撑成本,另外一方面,通过快速实时的支撑方式,为数字化服务运营及市场营销提速,实用性强,易于推广。

附图说明

[0014] 附图 1 是本发明的实现示意图。

具体实施方式

[0015] 下面结合附图和具体实施例对本发明作进一步说明。

[0016] 现有技术方式下,应对新的服务需求及支撑要求,需要根据具体应用需求及场景,接入各系统数据,生成分析场景数据模型,开发报表及分析功能,不但开发周期长,而且再次生成新的数据孤岛。

[0017] 本发明提供一种大数据的智能聚合可视化与管控系统,通过对大数据的采集、清洗、标签、分析,采用自动化数据画像技术、智能化数据聚合技术,提供主动推送和闭环管控,建立大数据的聚合可视化与管控体系。应用本发明,可解决数据难以自动模型化、智能可视化,提供面向角色的大数据聚合可视化与管控。

[0018] 如附图 1 所示,其具体结构包括数据采集与清洗、数据分类与画像、数据聚合与推送、流程闭环与管控四个层面的功能群组,其中,

数据采集包括通信设备数据采集能力、离线数据批量采集能力、在线数据交互采集能力、互联网数据采集能力、流式数据实时采集能力等能力项,数据清洗包括清洗规则管理、异常数据管理、数据格式化管理、调度管理等能力项。

[0019] 数据分类与画像包括已知数据的模型化能力、未知数据的规则化能力。已知数据的模型化能力负责提供多维数据模型管理、数据脱敏管理、实时计算等能力项,未知数据的规则化负责提供数据的切片、数据关系的挖掘等能力项。

[0020] 数据聚合包括聚合算法管理、搜索能力、语义词义管理等,数据推送包括数据服务能力、多维呈现能力等。

[0021] 流程闭环与管控提供质量规则管理、角色管理、流程管理等基础配置管理能力,通过端到端环节探针收集、质量分析及派单、后评估等服务能力,发现数据质量问题及时告警派单,并能通过数据聚合结果提供派单处理能力,提供系统内数据的自流转及质量保障。

[0022] 更为具体的,现对四个功能群组进行详细描述:

一、数据采集与清洗。

[0023] 数据采集包括通信设备数据采集能力、离线数据批量采集能力、在线数据交互采集能力、互联网数据采集能力、流式数据实时采集能力等能力项,数据清洗包括清洗规则管理、异常数据管理、数据格式化管理、调度管理等能力项。

[0024] 数据采集与清洗,作为系统的输入源,解决数据从哪里来。通过实时在线获取数据,利用数据清洗规则,对数据进行清洗及格式化,对清洗出的异常数据进行及时发现及关联分析,并作为异常告警直接推送前台可视及分析。

[0025] 二、数据分类与画像。

[0026] 包括已知数据的模型化能力、未知数据的规则化能力。已知数据的模型化能力负责提供多维数据模型管理、数据脱敏管理、实时计算等能力项,未知数据的规则化负责提供数据的切片、数据关系的挖掘等能力项。

[0027] 数据分类与画像,作为系统的基础,体现所需及所在(有),解决数据是什么。通过对采集清洗后的数据进行解剖,显性数据模型化,包括面向用户、网元、终端、业务、网络、区域等,建立多维度多视角多口径的模型基础;隐性关系规则化,包括基于时间、位置的数据切片、分词统计等,挖掘数据关系规则。对生产数据、过程数据、组织数据、非结构化数据、互联网数据等进行大类、中类、小类三级标签,生成数据全貌画像,提供前台模糊搜索呈现。所述的功能包括:

a) 提供对用户、网元、终端、业务、网络、区域等多维度多视角多口径的模型管理,显性数据的模型化;

b) 提供对数据基于时间、位置的切片,对数据的分词及统计能力,隐性关系的规则挖掘;

c) 提供对标签分析的管理,包括大类、中类、小类三级标签,针对模型化及规则化后的数据,进行数据标签分析、数据定义,生成数据画像;

d) 提供针对大类/中类标签特征数据的实时过滤、实时计算,提供定制化数据画像能力,保证数据画像的时效性及可扩展性。

[0028] 三、数据聚合与推送。

[0029] 数据聚合包括聚合算法管理、搜索能力、语义词义管理等,数据推送包括数据服务能力、多维呈现能力等。

[0030] 数据聚合与推送,作为系统的核心,体现所见即所得(给),解决数据到哪里去。通过对用户、网元、终端、业务、网络、区域等数据画像进行翻译、统计、聚类算法智能聚合,利用搜索等多维呈现手段,面向不变的运维过程中的角色,聚合数据、聚合工单、聚合工具、聚合应用、聚合业务,保证数据聚合的相关性、工具及应用的协同。所述的功能包括:

a) 基于标签的数据聚合算法,通过对用户、网元、终端、业务、网络、区域等画像数据的大类、中类、小类三级标签进行翻译、统计、聚类算法聚合;

b) 基于统计的智能搜索关联,提供语义词义管理,基于统计算法进行智能搜索,关联聚合数据;

c) 基于数据及角色的工单、工具、应用、业务聚合,提供所链所用的能力,提供面向个人的运维运营门户;

d) 多维呈现能力,根据包括报表平台、拓扑、GIS、指标平台等多种基础平台能力,结合聚合数据、应用、业务等多维呈现。

[0031] 四、流程闭环与管控。

[0032] 提供质量规则管理、角色管理、流程管理等基础配置管理能力,通过端到端环节探针收集、质量分析及派单、后评估等服务能力,发现数据质量问题及时告警派单,并能通过数据聚合结果提供派单处理能力,提供系统内数据的自流转及质量保障。

[0033] 流程闭环与管控,作为系统的管理,体现可控可管、自流转,解决数据的可信。提供各环节探针收集过程数据、质量数据,根据系统负荷及服务能力,决定相应的处理策略,统一调控采集、画像、聚合三大能力,完成协同工作。

[0034] 本发明的大数据可视化与管控体系,包括数据采集与清洗、数据分类与画像、数据聚合与推送、流程闭环与管控四个层面的功能群组。

[0035] 上述具体实施方式仅是本发明的具体个案,本发明的专利保护范围包括但不限于上述具体实施方式,任何符合本发明的一种大数据的智能聚合可视化与管控系统的权利要求书的且任何所述技术领域的普通技术人员对其所做的适当变化或替换,皆应落入本发明的专利保护范围。

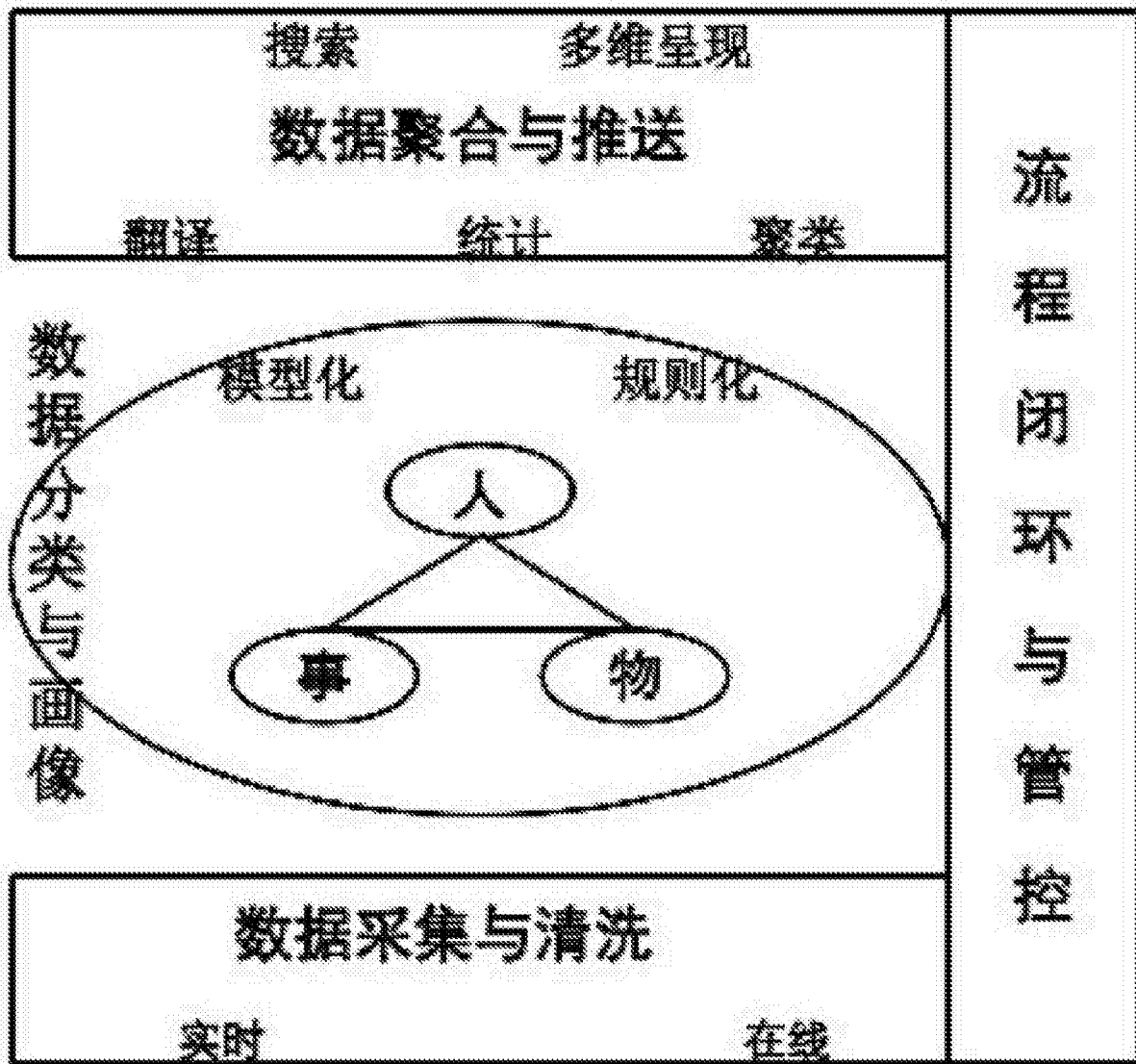


图 1