



(12) 发明专利申请

(10) 申请公布号 CN 116097244 A

(43) 申请公布日 2023. 05. 09

(21) 申请号 202180057053.5

(22) 申请日 2021.07.30

(30) 优先权数据

202041033832 2020.08.07 IN

17/333,738 2021.05.28 US

(85) PCT国际申请进入国家阶段日

2023.02.07

(86) PCT国际申请的申请数据

PCT/IN2021/050739 2021.07.30

(87) PCT国际申请的公布数据

W02022/029797 EN 2022.02.10

(71) 申请人 甲骨文金融服务软件有限公司

地址 印度孟买

(72) 发明人 R·N·瓦达潘德施瓦拉

T·N·坎特

(74) 专利代理机构 中国贸促会专利商标事务所

有限公司 11038

专利代理师 冯薇

(51) Int.Cl.

G06F 16/24 (2006.01)

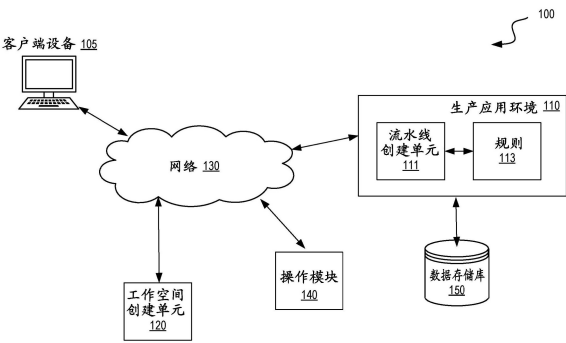
权利要求书4页 说明书21页 附图8页

(54) 发明名称

分析工作空间的物化

(57) 摘要

公开了用于创建工作空间的技术。数据处理系统接收创建工作空间以实现部署在生产环境中的模型的至少一部分的请求。获得与模型的该部分相关的一个或多个数据对象及其相关联的元数据,以及用于工作空间的执行场所。生成指令集,用于在工作空间中执行所述一个或多个数据对象和相关联的元数据。通过在工作空间中实例化模型的所述部分、所述一个或多个数据对象以及相关联的元数据,在执行场所内创建工作空间。根据指令集使用所述一个或多个数据对象和相关联的元数据在工作空间中处理模型的所述部分,并且由数据处理系统基于该处理来更新生产环境。



1. 一种方法,包括:

由数据处理系统接收创建工作空间以实现部署在生产环境中的模型的至少一部分的请求;

由数据处理系统获得与模型的所述部分相关的一个或多个数据对象及其相关联的元数据;

由数据处理系统获得用于工作空间的执行场所;

由数据处理系统生成指令集,用于在工作空间中执行所述一个或多个数据对象和相关联的元数据;

由数据处理系统在执行场所内创建工作空间,其中所述创建包括在工作空间中实例化模型的所述部分、所述一个或多个数据对象以及相关联的元数据;

根据所述指令集使用所述一个或多个数据对象和相关联的元数据在工作空间中处理模型的所述部分;以及

由数据处理系统基于所述处理来更新生产环境。

2. 如权利要求1所述的方法,还包括:

从与生产环境相关联的一个或多个数据库获得多个数据对象和多个元数据,所述一个或多个数据库维护部署在生产环境中的模型与相关联的数据对象的映射;

过滤所述多个数据对象和所述多个元数据以获得与模型的所述部分相关的所述一个或多个数据对象和相关联的元数据;以及

迭代地执行关于与生产环境相关联的所述一个或多个数据库的查找操作的集合以获得与所述一个或多个数据对象相关的至少一个数据对象。

3. 如权利要求1或2所述的方法,其中执行场所是生产环境内的应用实例或远程应用实例中的区域。

4. 如权利要求1、2或3所述的方法,还包括:

从用户接收在工作空间中执行模型的所述部分的请求;

核实用户是否被允许访问与模型的所述部分相关的所述一个或多个数据对象及其相关联的元数据;以及

响应于成功的核实而生成在执行场所内创建工作空间的供应可执行文件。

5. 如前述权利要求中的任一项所述的方法,其中工作空间还被配置用于

从数据处理系统外部的数据源摄取数据;以及

基于模型的所述部分创建至少一个沙箱应用或分析流水线,并且其中使用所述至少一个沙箱应用或分析流水线在工作空间中处理模型的所述部分。

6. 如前述权利要求中的任一项所述的方法,其中指令集以可变换成能够在托管工作空间的执行场所中执行的等效指令集的格式生成。

7. 如前述权利要求中的任一项所述的方法,其中更新生产环境还包括:

获得在执行场所内创建工作空间之前生产环境的初始状态,该初始状态包括部署在生产环境中的模型的第一输出;

将部署在生产环境中的模型的所述部分替换为在工作空间中被处理的模型的所述部分;以及

基于所述替换生成生产环境的更新后的状态,该更新后的状态包括部署在生产环境中

的模型的第二输出。

8. 一种计算设备,包括:

处理器;以及

存储器,包括指令,所述指令在用处理器执行时使得计算设备至少执行以下操作:

接收创建工作空间以实现部署在生产环境中的模型的至少一部分的请求;

获得与模型的所述部分相关的一个或多个数据对象及其相关联的元数据;

获得用于工作空间的执行场所;

生成指令集,用于在工作空间中执行所述一个或多个数据对象和相关联的元数据;

通过在工作空间中实例化模型的所述部分、所述一个或多个数据对象以及相关联的元数据而在执行场所内创建工作空间;

根据所述指令集使用所述一个或多个数据对象和相关联的元数据在工作空间中处理模型的所述部分;以及

基于所述处理来更新生产环境。

9. 如权利要求8所述的计算设备,其中处理器还被配置为:

从与生产环境相关联的一个或多个数据库获得多个数据对象和多个元数据,所述一个或多个数据库维护部署在生产环境中的模型与相关联的数据对象的映射;

过滤所述多个数据对象和所述多个元数据以获得与模型的所述部分相关的所述一个或多个数据对象和相关联的元数据;以及

迭代地执行关于与生产环境相关联的所述一个或多个数据库的查找操作的集合以获得与所述一个或多个数据对象相关的至少一个数据对象。

10. 如权利要求8或9所述的计算设备,其中执行场所是生产环境内的应用实例或远程应用实例中的区域。

11. 如权利要求8、9或10所述的计算设备,其中处理器还被配置为:

从用户接收在工作空间中执行模型的所述部分的请求;

核实用户是否被允许访问与模型的所述部分相关的所述一个或多个数据对象及其相关联的元数据;以及

响应于成功的核实而生成在执行场所内创建工作空间的供应可执行文件。

12. 如权利要求8至11中的任一项所述的计算设备,其中工作空间还被配置为

从外部数据源摄取数据;以及

基于模型的所述部分创建至少一个沙箱应用或分析流水线,并且其中使用所述至少一个沙箱应用或分析流水线在工作空间中处理模型的所述部分。

13. 如权利要求8至12中的任一项所述的计算设备,其中指令集以可变换成能够在托管工作空间的执行场所中执行的等效指令集的格式生成。

14. 如权利要求8至13中的任一项所述的计算设备,其中处理器还被配置为:

获得在执行场所内创建工作空间之前生产环境的初始状态,该初始状态包括部署在生产环境中的模型的第一输出;

将部署在生产环境中的模型的所述部分替换为在工作空间中被处理的模型的所述部分;以及

基于模型的所述部分被替换,生成生产环境的更新后的状态,该更新后的状态包括部

署在生产环境中的模型的第二输出。

15. 一种存储特定计算机可执行指令的非暂态计算机可读介质, 该指令在由处理器执行时使得计算机系统至少执行以下操作:

接收创建工作空间以实现部署在生产环境中的模型的至少一部分的请求;

获得与模型的所述部分相关的一个或多个数据对象及其相关联的元数据;

获得用于工作空间的执行场所;

生成指令集, 用于在工作空间中执行所述一个或多个数据对象和相关联的元数据;

通过在工作空间中实例化模型的所述部分、所述一个或多个数据对象以及相关联的元数据而在执行场所内创建工作空间;

根据所述指令集使用所述一个或多个数据对象和相关联的元数据在工作空间中处理模型的所述部分;

以及基于所述处理来更新生产环境。

16. 如权利要求15所述的非暂态计算机可读介质, 其中计算机系统还被配置为:

从与生产环境相关联的一个或多个数据库获得多个数据对象和多个元数据, 所述一个或多个数据库维护部署在生产环境中的模型与相关联的数据对象的映射;

过滤所述多个数据对象和所述多个元数据以获得与模型的所述部分相关的所述一个或多个数据对象和相关联的元数据; 以及

迭代地执行关于与生产环境相关联的所述一个或多个数据库的查找操作的集合以获得与所述一个或多个数据对象相关的至少一个数据对象。

17. 如权利要求15或16所述的非暂态计算机可读介质, 其中执行场所是生产环境内的应用实例或远程应用实例中的区域。

18. 如权利要求15、16或17所述的非暂态计算机可读介质, 其中计算机系统还被配置为:

从用户接收在工作空间中执行模型的所述部分的请求;

核实用户是否被允许访问与模型的所述部分相关的所述一个或多个数据对象及其相关联的元数据; 以及

响应于成功的核实而生成在执行场所内创建工作空间的供应可执行文件。

19. 如权利要求15至18中的任一项所述的非暂态计算机可读介质, 其中计算机系统还被配置为:

从外部数据源摄取数据; 以及

基于模型的所述部分创建至少一个沙箱应用或分析流水线, 并且其中使用所述至少一个沙箱应用或分析流水线在工作空间中处理模型的所述部分。

20. 如权利要求15至19中的任一项所述的非暂态计算机可读介质, 其中计算机系统还被配置为:

获得在执行场所内创建工作空间之前生产环境的初始状态, 该初始状态包括部署在生产环境中的模型的第一输出;

将部署在生产环境中的模型的所述部分替换为在工作空间中被处理的模型的所述部分; 以及

基于模型的所述部分被替换, 生成生产环境的更新后的状态, 该更新后的状态包括部

署在生产环境中的模型的第二输出。

分析工作空间的物化

[0001] 相关申请的交叉引用

[0002] 本申请要求于2020年8月7日提交的印度临时申请No.202041033832和于2021年5月28日提交的美国非临时申请No.17/333,738的权益和优先权,出于所有目的其全部内容通过引用并入本文。

技术领域

[0003] 本公开一般而言涉及数据处理系统,并且更具体地涉及以无缝方式创建工作空间(workspace)以实现部署在生产环境中的模型的至少一部分的技术。

背景技术

[0004] 随着互联网的普及,各行各业的公司每时每刻都在产生和处理海量的数据。为了利用数据,不同的公司根据公司的业务偏好从数据中提取大量不同的数据指标(index),并且数据指标被用于测量公司的业务发展状况。例如,在电子商务网站应用中,测量公司服务的数据指标可以包括诸如每天活跃用户访问网页的次数、用户的点击率、每个网页的转换速率之类的因素。

[0005] 通常,组织所采用的过程被建模为工作流程,即,流水线(pipeline),其中不同人员(例如数据科学家、建模师、分析师等)可以在其中监视工作流程的性能。在此类工作流程中,构建复杂的数据变换集(包括确定性的和统计性的)以递送支持推导预测结果的用例通常要求创建私有工作空间。此类私有工作空间通常经由桌面应用或大量手动过程来解决,以将结果关联回源数据并将开发的逻辑/计算变换与生产用例联系起来。这种创建工作空间的方法效率低下,因为它往往会分离成孤岛并失去整个工作流程的沿袭性。

[0006] 本文描述的实施例单独地和共同地解决这些和其它问题。

发明内容

[0007] 提供了用于创建工作空间以实现在生产环境中部署的模型的至少一部分的技术(例如,方法、系统、存储可由一个或多个处理器执行的代码或指令的非暂态计算机可读介质)。本文描述了各种实施例,包括方法、系统、存储可由一个或多个处理器执行的程序、代码或指令的非暂态计算机可读存储介质等。

[0008] 在各种实施例中,提供了一种方法,该方法包括:由数据处理系统接收创建工作空间以实现部署在生产环境中的模型的至少一部分的请求;由数据处理系统获得与模型的所述部分相关的一个或多个数据对象及其相关联的元数据;由数据处理系统获得用于工作空间的执行场所;由数据处理系统生成指令集,用于在工作空间中执行所述一个或多个数据对象和相关联的元数据;由数据处理系统在执行场所内创建工作空间,其中所述创建包括在工作空间中实例化模型的所述部分、所述一个或多个数据对象以及相关联的元数据;根据指令集使用所述一个或多个数据对象和相关联的元数据在工作空间中处理模型的所述部分;以及由数据处理系统基于所述处理来更新生产环境。

[0009] 在一些实施例中,该方法还包括:从与生产环境相关联的一个或多个数据库获得多个数据对象和多个元数据,所述一个或多个数据库维护部署在生产环境中的模型与相关联的数据对象的映射;过滤所述多个数据对象和所述多个元数据以获得与模型的所述部分相关的所述一个或多个数据对象和相关联的元数据;以及迭代地执行关于与生产环境相关联的所述一个或多个数据库的查找操作的集合以获得与所述一个或多个数据对象相关的至少一个数据对象。

[0010] 在一些实施例中,执行场所是生产环境内的应用实例或远程应用实例中的区域。

[0011] 在一些实施例中,该方法还包括:从用户接收在工作空间中执行模型的所述部分的请求;核实用户是否被允许访问与模型的所述部分相关的所述一个或多个数据对象及其相关联的元数据;以及响应于成功的核实而生成在执行场所内创建工作空间的供应可执行文件(executable)。

[0012] 在一些实施例中,工作空间还被配置用于:从数据处理系统外部的数据源摄取(ingest)数据;以及基于模型的所述部分创建至少一个沙箱应用或分析流水线,并且其中使用所述至少一个沙箱应用或分析流水线在工作空间中处理模型的所述部分。

[0013] 在一些实施例中,指令集以可变换成能够在托管工作空间的执行场所中执行的等效指令集的格式生成。

[0014] 在一些实施例中,更新生产环境还包括:获得在执行场所内创建工作空间之前生产环境的初始状态,该初始状态包括部署在生产环境中的模型的第一输出;将部署在生产环境中的模型的所述部分替换为在工作空间中被处理的模型的所述部分;以及基于所述替换生成生产环境的更新后的状态,该更新后的状态包括部署在生产环境中的模型的第二输出。

[0015] 在一些实施例中,部署在生产环境中的模型包括多个子模型,并且模型的所述部分与所述多个子模型中的子模型对应。

[0016] 在一些实施例中,在工作空间中处理模型的所述部分还包括:关于训练数据集重新训练模型的所述部分;以及基于所述重新训练来修改模型的所述部分的一个或多个参数。

[0017] 在一些实施例中,在执行场所内创建工作空间还包括:由数据处理系统生成能够在执行场所内创建工作空间的供应可执行文件;以及执行供应可执行文件以在执行场所内创建工作空间。

[0018] 在一些实施例中,该方法还包括:在数据处理系统的数据库中存储工作空间的配置信息,该配置信息至少包括与模型的该部分相关的所述一个或多个数据对象及其相关联的元数据以及与执行场所相关信息;由数据处理系统删除在执行场所内创建的工作空间;以及由数据处理系统基于配置信息在另一个执行场所中生成新的工作空间。

[0019] 在一些实施例中,所述多个子模型包括神经网络模型、线性回归模型、支持向量机模型或聚类模型。

[0020] 在一些实施例中,提供了一种计算设备,该计算设备包括处理器和包括指令的存储器,该指令在用处理器执行时使得计算设备至少执行以下操作:接收创建工作空间以实现部署在生产环境中的模型的至少一部分的请求;获得与模型的所述部分相关的一个或多个数据对象及其相关联的元数据;获得用于工作空间的执行场所;生成指令集,用于在工作

空间中执行所述一个或多个数据对象和相关联的元数据;通过在工作空间中实例化模型的所述部分、所述一个或多个数据对象以及相关联的元数据而在执行场所内创建工作空间;根据所述指令集使用所述一个或多个数据对象和相关联的元数据在工作空间中处理模型的所述部分;以及基于所述处理来更新生产环境。

[0021] 在一些实施例中,提供了一种存储特定计算机可执行指令的非暂态计算机可读介质,该指令在由处理器执行时使得计算机系统至少执行以下操作:接收创建工作空间以实现部署在生产环境中的模型的至少一部分的请求;获得与模型的所述部分相关的一个或多个数据对象及其相关联的元数据;获得用于工作空间的执行场所;生成指令集,用于在工作空间中执行所述一个或多个数据对象和相关联的元数据;通过在工作空间中实例化模型的所述部分、所述一个或多个数据对象以及相关联的元数据而在执行场所内创建工作空间;根据所述指令集使用所述一个或多个数据对象和相关联的元数据在工作空间中处理模型的所述部分;以及基于所述处理来更新生产环境。

[0022] 上面和下面描述的技术可以通过多种方式和在多种情境中实现。参考以下附图设置若干示例实施方式和情境,如下面更详细描述。但是,以下实施方式和情境只是其中的一小部分。

附图说明

[0023] 图1是根据各种实施例的数据处理系统的图示。

[0024] 图2是根据各种实施例供应工作空间的图示。

[0025] 图3图示了根据各种实施例的供应工作空间的细节。

[0026] 图4图示了根据各种实施例的数据接口单元。

[0027] 图5描绘了说明根据各种实施例的工作空间供应的过程的流程图。

[0028] 图6描绘了用于实现各种实施例的分布式系统的简化图。

[0029] 图7是根据各种实施例的系统环境的一个或多个组件的简化框图,通过该系统环境可以将实施例系统的一个或多个组件提供的服务提供为云服务。

[0030] 图8图示了可以用于实现各种实施例的示例计算机系统。

具体实施方式

[0031] 在以下描述中,出于解释的目的,阐述了具体细节以便提供对某些发明性实施例的透彻理解。但是,显然可以在没有这些具体细节的情况下实践各种实施例。附图和描述并非旨在限制。“示例性”一词在本文用于表示“用作示例、实例或说明”。本文描述为“示例性”的任何实施例或设计不一定被解释为优于或有利于其它实施例或设计。

[0032] 数据处理系统的概述

[0033] 企业通常要求持续监视工作流程以测量风险,以便符合普遍的监管和监督标准。在此类工作流程中,测量常常需要对企业数据进行大量计算和验证。数据被变换以支持此类测量和计算。数据的变换所涉及的处理包括大量的时间段数据,并且其处理方法被描述为执行数据变换和动作(包括计算出的值的插入/更新)的一系列任务,这些任务可以通过定义的规则的集合来实现。

[0034] 复杂的分析工作流程被建模为整体模型(例如,机器学习模型),其由若干子模型

构成。具体而言,将若干个子模型以预定方式拼接在一起(即,组合)以生成整体模型。子模型以预定次序被执行以确定整体模型的一个或多个输出变量。每个子模型包括一系列任务,这一系列任务对一个或多个数据集进行操作并执行数据变换以计算整个模型的一个或多个输出变量。因此,整体模型的执行包括:(i) 确定多个输入变量(即,一个或多个子模型的输入参数),(ii) 根据预定的次序执行一个或多个子模型,以及(iii) 生成整体模型的一个或多个输出变量。这种整体模型部署在企业的生产环境中。来自企业不同部门的不同人员(例如,开发人员、数据科学家等)可以监视生产环境中的整体模型。子模型可以包括机器学习模型,诸如神经网络模型、线性回归模型、支持向量机模型或聚类模型。

[0035] 这种复杂的分析工作流程中的要求是,不同的实体可能期望提取整体模型的一部分(例如,子模型)以用于进行进一步处理和开发。例如,构建整体模型的特定子模型的企业部门可能期望提取子模型(以及用于构建模型的相关联的数据对象和元数据)以进一步处理和增强该子模型,例如,关于新的训练数据重新训练该子模型、微调子模型的一个或多个参数以增强性能等。可替代地,也可能期望获得生产环境数据(例如,数据对象、数据记录等)的切片,用于构建新的预测模型,或评估场景以进行异常检测,研究数据模式等。本文描述的是使得能够创建工作空间(例如,现有应用实例内、远程应用实例内或混合环境中的区域),可以在其中提取期望的子模型、数据对象、元数据等以进行进一步处理而不会影响整体模型在生产环境中的运作。通过一些实施例,流水线创建单元111包括模型聚合器。模型聚合器被编程为获得一个或多个子模型(数据对象及其相关联的元数据)并以用户定义的方式组合一个或多个子模型以生成整体模型。模型聚合器可以被实现为获得要被拼接(或组合)的一个或多个子模型以形成整体模型的脚本,例如,Python脚本。

[0036] 本文描述的各种实施例可以用在各种不同的领域和情境中。某些实施例特别适用于以分布式计算技术提供的企业应用软件。但是,本文描述的系统和方法可以被用于为期望创建工作空间的任何系统或应用框架提供数据变换功能性。公开了使数据处理系统能够实例化工作空间并(从生产环境中)将整体模型的相关部分(即,子模型、特定数据对象(及其相关联的元数据))提取到工作空间以供进一步处理的技术(例如,系统、方法和机器可读介质)。

[0037] 图1描绘了根据各种实施例的数据处理系统100,其包括客户端设备105、生产应用环境110、工作空间创建单元120、网络130、操作模块140和数据存储装置150(例如,存储器存储库)。生产应用环境110包括流水线创建单元111和规则单元113。

[0038] 可以向操作客户端设备105的用户呈现一个或多个应用界面(例如,图形用户界面),该一个或多个应用界面接受输入以使用户能够与生产应用环境110和/或工作空间创建单元120进行交互。客户端设备105的示例包括但不限于工作站、个人计算机(PC)、膝上型计算机、诸如智能电话之类的移动设备、可穿戴计算机或其它联网的电子设备。可使用在客户端设备105上执行的应用来访问在客户端设备105处呈现的界面。客户端设备105、生产应用环境110、工作空间创建单元120和操作模块140可以经由一个或多个通信网络130来通信。通信网络130的示例可以包括移动网络、无线网络、蜂窝网络、局域网(LAN)、广域网(WAN)、其它无线通信网络或其组合。

[0039] 通过一些实施例,包括在生产应用环境110中的流水线创建单元111规定(provision)用户将整体模型定义为流水线(本文也称为工作流程),即,被拼接/组合以形

成整体模型的一个或多个子模型的组合。流水线创建单元111还规定用户定义和执行在数据处理系统100中变换数据所需的规则集、报告对象和过程。流水线创建单元111提供允许用户经由例如客户端设备105定义和执行规则、过程并且管理用于形成整体模型的每个子模型的定义的接口。例如,许多企业要求持续监视和测量风险,以便符合普遍的监管和监督标准。这些测量常常需要对企业数据进行大量计算和验证。数据经常被变换以支持此类测量和计算。这些数据变换可以通过定义的规则的集合来实现。

[0040] 流水线创建单元111规定用户定义用于执行整体模型的规则113,例如,定义形成整体模型的一个或多个子模型的执行次序。例如,规则可以包括并发执行的一个或多个子模型的第一子集,而一个或多个子模型的第二子集被顺序执行以便确定整体模型的一个或多个输出变量。此外,流水线创建单元111还规定用户经由应用接口定义例如要从外部数据源获得的一个或多个数据集,将关于这些数据集评估整体模型。此类数据集可以经由操作模块140从外部数据源获得。

[0041] 规则单元113促进用户定义(例如,关于子模型)被实现的规则集、报告对象和过程以变换存储设备(诸如数据存储装置150)中的数据。例如,规则单元113提供促进变换的定义和维护的框架。元数据抽象层可以用在规则的定义中,其中允许用户对数据存储库150中的属性重新分类,从而变换数据。底层元数据对象(诸如非大型或非列表的层次结构、数据集和处理器)驱动规则功能性。规则可以包括各种类型。例如,类型1规则可以涉及基于某些过滤器从数据模型中的给定记录集创建记录的子集。这个过程可以涉及变换、聚合或其组合。类型1规则定义可以通过表到表(T2T)提取来实现。类型2规则可以涉及基于准则对数据模型中的表中的记录进行重新分类,这些准则包括表中的复杂group-by子句和子查询。类型3规则可以涉及基于简单测量来计算新值或度量,并用计算出的值更新数据模型内已识别出的记录集。流水线创建单元111包括工具的集合,其帮助用户操作各种规则定义。

[0042] 在各种实施例中,规则被定义为包括一个或多个组件。这些组件可以包括数据集、源、目标、映射和/或节点标识符。数据集是通过键(key)接合在一起的表的集合。数据集可以至少有一个事实表。事实表存储过程的测量、度量或事实。事实表可以位于星型模式或雪花模式的中心,被维度表包围。维度表存储参考数据(例如,实体处理的产品的列表、分支的列表、地理的列表、货币的列表等,这些是相对静态的)。在使用多个事实表的情况下,这些可以被布置为事实星座模式。事实表通常有两种类型的列:包含事实的列和作为维表的外键的列。事实表的主键通常是由其所有外键组成的组合键。事实表包含数据存储库150的内容并存储不同类型的测量,如加性、非加性和半加性度量。类型3规则定义可以基于包含多于一个事实表的数据集。类型2规则定义可以基于包含单个事实表的数据集。数据集内事实表的一列或多列中的值可以用新值进行变换。

[0043] 根据一些实施例,规则的集合共同形成过程。过程定义可以被表示为过程树。规则单元113提供促进过程的定义和维护的框架。通过定义过程,用户可以对属于功能过程的规则的集合进行逻辑分组。例如,用户可以使用层次结构用现有的元数据对象定义过程,这促进过程树的构建。过程树可以有很多级别并且每个级别内有一个或多个节点。以这种方式,生产应用环境110可以利用规则单元113来定义用于复杂分析工作流程的子模型的过程。可以认识到的是,子模型可以包括一个或多个任务,每个任务都将被执行以便完成子模型的执行。(例如,子模型的任务的)子过程可以在级别成员上定义并且过程层次结构成员形成

树的叶成员。因此,过程可以包括两种类型:(i)端到端过程,其表示功能完整性并准备好执行,以及(ii)非端到端过程,这是规则的逻辑集合并且本身不能被执行。非端到端过程可以被定义为端到端过程中的子过程来执行。规则单元113可以被用于通过添加一个或多个称为过程节点的成员来构建过程树。如果存在与任何成员相关联的前置(predecessor)任务,那么被定义为前置任务的任务先于该成员的执行。

[0044] 由规则单元113创建的定义产生唯一的批次组。然后可以经由由生产应用环境110提供的接口来调度这些批次以供执行。针对过程的执行(例如,子模型的执行)的每个请求都可以成为批次组中的批次。生产应用环境110使用户能够具有用于执行过程的工作过程。通过一些实施例,然后可以从操作模块140执行批次。例如,用户可以经由应用编程接口(API)来选择要被标记为执行的过程定义,其中所选择的过程定义的名称出现在API的运行描述文本框中。API显示其它信息,例如,诸如请求执行ID、请求执行描述、请求状态、执行状态和编辑之类的选项。

[0045] API可以有选项来选择为执行定义的批次。请求状态“打开”指示所选择的过程ID已准备好执行。执行状态“未启动”指示特定过程定义/批次尚未被执行。用户可以选择编辑选项来编辑针对执行的请求。用户还可以使用搜索和过滤选项以基于过程名称、过程执行描述、过程执行ID、过程类型、请求状态等来搜索特定过程。还可以提供分页选项来帮助用户管理数据处理系统内现有过程的协作视图。

[0046] 通过一些实施例,操作模块140促进用户管理和处理数据(例如,企业数据)以在数据处理系统100内创造最高级别的效率并基于一个或多个指定的规则得出结果。操作模块140可以包括诸如批次维护、批次执行、批次调度器、批次监视器、批次处理报告、批次取消、查看日志和报告生成之类的单元。

[0047] 操作模块140内的批次维护框架促进用户创建和维护批次定义。用户可以从批次维护以及还从其它模块(诸如高级分析基础设施模块(未示出)和规则单元113)处理被调度执行的批次。操作模块140的批次执行单元被配置为发起批次以供处理。通过一些实施例,当批次被提交以供执行时,关于定义的组件和参数的一系列命令被传输到数据存储库150。当命令被成功执行时,这进而返回更新计数(例如,值定义)的阵列。操作模块140中的批次调度器单元促进用户调度批次以供随后处理。用户可以定义新的批次调度或更新先前定义的批次调度以供处理。

[0048] 操作模块140的批次监视单元促进用户查看执行的批次定义的状态以及任务细节。用户可以定期跟踪问题(如果有的话)并确保顺利的批次执行。事件日志为用户提供已执行的批次的实时状态。操作模块140中的批次处理报告促进用户查看批次中定义的每个任务组件的执行情况。操作模块140的批次取消单元促进用户取消或中止被调度或者正在执行处理中的批次或特定任务。查看日志单元促进用户查看批次中定义的每个任务组件的执行状态。操作模块140的报告生成单元被配置为生成要提供给客户端设备105的整体模型的性能报告。

[0049] 工作空间创建单元120允许例如建模者、数据和领域分析师创建一个或多个工作空间。在此类工作空间中,数据对象的切片、逻辑抽象、元数据和/或来自生产环境的整体模型的部分(即,子模型)可以被导出,使得导出的信息可以以独立的方式被处理(即,与生产环境分开)用于不同的目的。例如,导出的数据对象、元数据和逻辑抽象可以被用于构建预

测模型或场景以进行异常检测或研究数据模式。还可以训练和增强导出的子模型(例如,优化子模型的参数)用于增强的性能。工作空间创建单元120允许将数据和相关联的对象(逻辑/计算)两者拉入工作空间,即,现有应用实例内的区域或者远程实例或混合环境中,同时通过交互式 and 离线生成的指令集应用通用/一致的元数据。这种方法允许复制、虚拟化和部署经训练的逻辑/计算/模型,并自动绑定到由安全性权利、自我增强和控制管理的目标。

[0050] 根据一些实施例,工作空间创建单元120可以被实现为提取子模型(即,子模型的代码)的副本(即,获得其快照)以及与子模型关联的元数据的副本的可执行脚本,例如,Java脚本、SQL脚本等。元数据可以包括例如数据对象、外部软件包和来自与子模型相关联的外部数据源的数据。子模型的提取出的副本和与其相关联的元数据被打包在档案(例如,可执行的zip档案)中。被实现为可执行脚本的工作空间创建单元120还可以定义与子模型相关联的元数据中的依赖性(例如,数据依赖性)以便确保(执行子模型所需的)所有相关数据都被包括在zip档案中。因此,可以将档案导出到为特定用户创建的工作空间,以进一步开发或改进子模型,而不会影响整体模型在生产环境中的运作。

[0051] 通过一些实施例,一个或多个用户(例如,建模者、数据/领域分析师等)可以创建他们的被导出到用户的相应工作空间的个人档案,即,打包的子模型。可以认识到的是,多个实体可以创建与同一子模型对应的zip档案。用户以独立的方式在其相应的工作空间内处理子模型。例如,不同的用户可以利用不同的数据集、不同的参数等重新训练子模型。在用户的不同工作空间内完成子模型的处理后,通过一些实施例,选择来自其中一个工作空间的特定的子模型以替换生产环境中的原始子模型。例如,将在用户的工作空间中执行的每个子模型的性能与其它类似子模型的性能进行比较。以这种方式,选择其性能优于其它子模型(例如,实现子模型更高吞吐量)的子模型之一来替换生产环境中的原始子模型。此后,关于被替换的子模型执行各种用户对子模型的任何进一步处理/精炼。可以认识到的是,子模型的较旧版本可以存储在与生产环境相关联的数据存储库150中以供以后使用或用于审计目的。

[0052] 根据一些实施例,工作空间创建单元120向客户端设备105提供使客户端设备105能够创建工作空间的用户接口向导(例如,操作模块115的一部分)。工作空间的创建可以包括:(i)具有受控的用户权限的数据实体选择,(ii)自动列出与所选择的逻辑数据元素(即,数据集或数据对象)相关联的元数据(逻辑/计算),(iii)根据需要允许对相关元数据的选择,(iv)扫描可以托管工作空间的本地、远程环境(即,用于工作空间的执行场所);以及(v)列出用于工作空间部署的选择(例如,规定数据物化(materialization)是混合的)。如本文将参考图3描述的,为上面列出的每个步骤生成指令集以创建工作空间并将相关数据导出到工作空间。可以认识到的是,指令集可以以特定于领域的语言或脚本语言(如Shell、Perl、Python)生成,包括与逻辑模型、元数据对象、数据文件的关联。另外,指令集以供指令集在工作空间的期望执行场所中被执行的方式生成。下面参考图2-5描述了关于创建工作空间的详细信息。

[0053] 图2是根据各种实施例供应工作空间的图示。如图2的图示200中所示,数据模型205(例如,整体模型)被部署在生产环境201中。可以基于从外部数据源210获得的数据来构造数据模型205。工作空间创建单元120从客户端设备(例如,图1的客户端设备105)接收创建工作空间以实现部署在生产环境201中的数据模型205的至少一部分的请求。

[0054] 根据一些实施例,为了进一步处理和增强子模型,例如,为了关于新的训练数据重新训练子模型、微调子模型的一个或多个参数以增强性能等,期望在工作空间中实现的数据模型205的上述部分可以与数据模型的特定子模型(以及用于构建子模型的相关联的数据对象和元数据)对应。此外,为了构建新的预测模型或评估场景以进行异常检测、研究数据模式等,期望在工作空间中实现的数据模型205的上述部分可以与生产环境数据的切片(例如,数据对象、数据记录等)对应。期望在工作空间中实现的数据模型205的该部分可以被包括在客户端设备向工作空间创建单元120发出的请求中。

[0055] 可以认识到的是,可以从生产环境201中提取数据对象、元数据或数据模板(例如,报告模板、分析模板、产品层次模板、组织层次模板等),以包括在工作空间中以供进一步处理。换句话说,工作空间创建单元120创建被包括在生产环境201中的数据对象、元数据等的副本,并在工作空间中实例化(即,物理部署)副本。实例化的数据对象、元数据等可以基于在工作空间中执行的处理(例如,测试子模型、更新子模型等)而被更改。通过一些实施例,在工作空间中执行的处理完成时,在工作空间中被更改的数据对象、元数据、模型的一部分等可以经由工作空间创建单元120被推回到生产环境。例如,在数据模型的子模型在工作空间中被实例化(例如,为了微调或重新训练子模型的目的)的场景中,经重新训练/微调的子模型可以被工作空间创建单元120推回(在完成时)到生产环境201。

[0056] 以这种方式,部署在生产环境201中的较旧版本的子模型可以被来自工作空间的经重新训练/微调的子模型替换。此外,通过一些实施例,图1的数据处理系统可以提供在生产环境201中维护子模型的旧版本以及子模型的更新后的版本两者的选项。根据一些实施例,在创建工作空间之前,可以将生产环境201的初始状态存储在(与生产环境201相关联的)数据存储装置中,其中初始状态包括部署在生产环境201中的数据模型205的性能参数(例如,第一输出)。例如,在用(从工作空间获得的)子模型的经保留/微调的版本替换较旧版本的数据模型的子模型后,数据处理系统可以生成生产环境的更新后的状态,其中更新后的状态包括部署在生产环境中的模型的新输出(即,与第一输出不同的第二输出)。部署在生产环境中的数据模型的更新后的性能可以经由应用接口被呈现给客户端设备。

[0057] 由客户端设备向工作空间创建单元120发出的请求包括关于要创建的工作空间的执行场所的信息。具体而言,执行场所与期望要创建工作空间的目标环境对应。根据一些实施例,执行场所与生产环境(即,存储器中工作空间实施方式)或远程应用实例或混合环境中的应用实例内的区域对应。混合环境在本文中被定义为托管工作空间的环境,(除了从生产环境201提取的数据之外)该工作空间被配置为还从其它工作空间、外部数据源210等提取数据。

[0058] 不同的客户端设备(例如,来自组织的不同部门)可以向工作空间创建单元120发出工作空间创建请求。注意的是,每个请求包括关于工作空间的执行场所的信息以及要在工作空间中实例化的数据模型205的一部分。工作空间创建单元120被配置为为每个请求创建工作空间,例如,工作空间A215、工作空间B220和工作空间C225。在每个工作空间中,相关数据对象、元数据和/或数据模型的部分被复制并存储在本地以供进一步处理。例如,如图2中所示,工作空间A包括存储由客户端设备请求的对应数据对象和元数据的副本的数据库216和217,而工作空间B包括存储所请求的数据对象和元数据的相应副本的数据库221和222,而工作空间C(部署在混合环境中)包括存储由客户端设备请求的对应数据对象和元数

据的副本的数据库226和227。如将参考图3详细描述,过滤机制可以被工作空间创建单元120利用以获得要在工作空间中被实例化的相关数据。

[0059] 另外,如图2中所示,可以认识到的是,工作空间C部署在混合环境中,因为它被配置为从外部数据源210提取数据以及从其它工作空间(例如,工作空间B 220)提取数据。通过一些实施例,所创建的工作空间215、220和225中的每一个还可以经由应用编程接口230(例如,REST API)创建有助于在工作空间中执行的(例如,子模型的)处理/评估的至少一个沙箱应用(241、242)或分析流水线(243)。例如,在软件开发的情境中,沙箱应用是隔离的环境,可以在其中评估和测试与子模型有关的代码。以类似的方式,特定工作空间(例如,工作空间A215)可以创建分析流水线243,其可以被用于模拟/测试包括在正在工作空间A 215中被处理的子模型中的不同部分(例如,过程)。

[0060] 图3图示了根据各种实施例的供应工作空间的细节。具体而言,图3中描绘的示意图300图示了负责在执行场所内创建工作空间的工作空间创建单元的组件。如图3中所示,工作空间创建单元(例如,图1的工作空间创建单元120)包括供应应用330,其从远程数据存储器310和/或生产环境320获得数据对象和相关联的元数据并且在一个或多个工作空间(例如,工作空间A 341、工作空间B 343和工作空间C 345)中实例化(即,物理地物化)所获得的信息。

[0061] 通过一些实施例,供应应用330包括数据接口单元333、对象接口单元335、指令单元336、配置数据库337和物化器单元338。数据接口单元333被配置为从远程数据存储器310以及被包括在生产环境320中的数据存储器322提取数据(例如,数据对象)。对象接口单元335被配置为从包括在生产环境320中的元存储器323提取元数据(例如,与数据对象相关联的元数据)。提取出的元数据连同提取出的数据对象由工作空间(例如,工作空间A341、工作空间B343和工作空间C345)中的供应应用进行实例化。如图3中所示,工作空间创建单元的供应应用330可以物化(即,物理部署)不同类型的工作空间。例如,工作空间A 341和工作空间B 343部署在作为远程应用实例或混合环境中的执行场所中,而工作空间C部署在生产环境320内的应用实例内的区域中(即,存储器内工作空间实施方式)。

[0062] 通过一些实施例,数据接口单元333以及对象接口单元335可以利用过滤机制从远程数据存储器310和生产环境320获得相关数据(和元数据)。关于数据接口单元333的操作的细节在本文参考图4来描述。经由数据接口单元333获得的数据对象和经由对象接口单元335获得的元数据被输入到指令单元337。指令单元337生成用于执行将在执行场所内创建的工作空间中的一个或多个数据对象和相关联的元数据的指令集。具体而言,指令集由指令单元337以可变换成可在托管工作空间的执行场所中执行的等效指令集的格式生成。因此,指令单元337将指令集生成为可以在不同工作空间环境中执行的可执行脚本(例如,工件(artifact))。换句话说,由指令单元337生成的指令集是目标不可知的(即,工作空间不可知的),其中数据对象和元数据可以经由可执行脚本的执行在不同的工作空间环境中被物化。例如,工作空间创建单元可以从不同类型的数据存储器(例如,比特数据存储器、关系数据存储器等)提取数据,并将提取出的数据在任何环境(例如大数据环境)中物化。

[0063] 通过一些实施例,对于由客户端设备发出的工作空间创建请求,与生成的指令集(以及相关的数据对象和元数据)、工作空间的执行场所等有关的信息由指令单元337存储为配置数据库336中的工作空间的配置信息。将不同工作空间的这种配置信息存储在配

置数据库336中使得工作空间创建单元120能够在未来无缝地创建工作空间。例如,考虑特定工作空间(例如,工作空间A 341)的配置信息被存储在配置数据库336中。假设在某个时间点,工作空间A经由客户端设备发出的请求被删除。此后,如果客户端设备想要重新创建工作空间A,那么工作空间创建单元可以简单地从配置数据库336中检索配置信息以物化工作空间A。在这样做时,工作空间创建单元避免了(再次)从远程数据存储库310和/或生产环境320提取所有数据对象和相关联的元数据的需要,从而节省数据处理系统的计算资源。此外,工作空间创建单元可以向客户端设备(经由API)提供先前与工作空间相关联的数据对象和元数据的列表。客户端设备可以利用数据对象和元数据的列表,以及指示工作空间创建单元在重新创建工作空间的新实例时检索附加的数据对象(或元数据)。

[0064] 根据一些实施例,物化器单元339被配置为处理提取出的数据对象和元数据,并生成能够(在执行时)在执行场所中物理地创建工作空间的供应可执行文件(例如,脚本)。在创建工作空间后,工作空间创建单元可以实例化工作空间内提取出的数据对象、元数据和/或数据模型的一部分(即,子模型)以供进一步处理。

[0065] 图4图示了根据各种实施例的数据接口单元。具体而言,图4描绘了图示被包括在图3的供应应用330内的数据接口单元333的组件的示意图400。数据接口单元333包括请求处理单元410、过滤单元411、数据检索单元413、影响分析单元415、数据处理单元417、权利419和中性语言转化器421。

[0066] 根据一些实施例,由客户端设备发出的用于创建工作空间的请求(例如,经由图1的客户端设备105发出的用户请求)被指向工作空间创建单元。具体而言,该请求被指向数据接口单元333的请求处理单元410。请求处理单元410被配置为对请求进行处理,例如,提取与工作空间的执行场所有关的信息,获得与将在工作空间中实现的模型(例如,部署在生产环境201中的数据模型205)的一部分有关的信息等。在处理针对工作空间创建的请求后,请求处理单元410触发数据检索单元413。

[0067] 数据检索单元413与被包括在生产环境410中的驱动程序403进行通信。驱动程序403关于(生产环境410的)数据存储库401执行查找操作以检索数据对象、函数等。从数据存储库401中提取的信息被传递到数据接口单元333的数据检索单元413。根据一些实施例,当用户请求与在工作空间内子模型的创建有关时,数据检索单元413向驱动程序403传输与子模型相关联的标识符。驱动程序403利用该标识符对数据存储库401执行查找操作,该数据存储库401维护部署在生产环境中的模型和相关联的数据对象的映射。以这种方式,驱动程序可以提取与子模型相关联的相关数据对象,并将提取出的数据对象传输到数据检索单元413。

[0068] 通过一些实施例,数据检索单元413与影响分析单元415通信以执行对检索出的数据对象的分析。例如,影响分析单元415检查提取出的数据对象的依赖性,即,确定依赖于检索出的数据对象的一个或多个附加数据对象,并通知数据检索单元413检索附加数据对象。通过一些实施例,数据接口单元333利用过滤单元411对检索出的数据对象执行(一个或多个)过滤操作以获得关于请求的相关数据对象的集合。例如,创建工作空间的请求可以包括与用户对于数据对象的偏好有关的信息。此类用户偏好可以被用作过滤机制以获得将在工作空间中被物化的最相关的数据对象集合。

[0069] 另外,检索出的和/或经过滤的数据对象被传递到数据处理单元417,该数据处理

单元417利用权利(entitlement) 419(即,一个或多个访问权限的集合)来确定用户(例如,发起工作空间创建请求的用户)是否被授权访问每个检索出的和/或经过滤的数据对象。响应于数据处理单元417确定用户未被授权访问检索出的数据对象中的一个或多个,数据处理单元417可以触发请求处理单元410向客户端设备传输电子消息以通知用户这样的确定。此外,电子消息可以包括向用户说明工作空间创建请求由于访问问题而不能完成的通知。但是,响应于成功地确定用户被授权访问一个或多个数据对象中的每一个,数据处理单元417激活中性语言转化器421。

[0070] 通过一些实施例,中性语言转化器421被配置为将数据对象、元数据等的格式转换成工作空间的执行场所的环境可接受的格式。在这样做时,数据接口单元333确保提取出的一个或多个数据对象可以被用于进一步处理工作空间。通过一些实施例,数据处理单元417在成功确定用户被授权访问一个或多个数据对象中的每一个时触发物化器(图3的单元338)以生成在执行场所内创建工作空间的供应可执行文件。此外,数据处理单元417还可以触发请求处理单元410向客户端设备传输指示用户关于一个或多个数据对象的成功授权的电子消息。由中性语言转化器421处理的一个或多个数据对象被传递到指令单元(即,图3的单元336)以开始在执行场所中创建工作空间的过程。

[0071] 图5图示了描绘根据各种实施例的由数据处理系统的工作空间创建单元执行的工作空间供应的过程的流程图500。图5中描绘的处理可以在由相应系统、硬件或其组合的一个或多个处理单元(例如,处理器、核(core))执行的软件(例如,代码、指令、程序)中实现。软件可以存储在非暂态存储介质上(例如,存储器设备上)。图5中呈现并在下文描述的方法旨在是说明性而非限制性的。虽然图5描绘了以特定顺序或次序发生的各种处理步骤,但这并不是限制性的。在某些替代实施例中,可以以某种不同的次序执行步骤,或者也可以并行地执行一些步骤。

[0072] 该过程开始于步骤510,其中工作空间创建单元(例如,从客户端设备)接收创建工作空间以实现部署在生产环境中的模型的至少一部分的请求。为了进一步处理和增强子模型以例如关于新的训练数据重新训练子模型、微调子模型的一个或多个参数以增强性能等,期望在工作空间中实现的模型(例如,图5的数据模型205)的上述部分可以与数据模型的特定子模型(以及用于构建子模型的相关联的数据对象和元数据)对应。通过一些实施例,为了构建新的预测模型或评估场景以进行异常检测、研究数据模式,期望在工作空间中实现的模型的上述部分可以与生产环境数据的切片(例如,数据对象、数据记录等)对应。

[0073] 过程然后前进到步骤520,其中工作空间创建单元从生产环境(例如,生产环境中包括的数据存储库和元存储库)获得与模型的上述部分相关的数据对象和相关联的元数据。工作空间创建单元通过对数据存储库和元存储库执行查找操作来获得相关数据对象和相关联的元数据,这维护了部署在生产环境中的模型与由模型使用的相关联的数据对象和元数据的映射。在步骤530中,工作空间创建单元获得执行场所,即,要创建工作空间的目的地。工作空间创建单元提取由客户端设备发出的工作空间创建请求中嵌入的执行场所信息。

[0074] 此后过程前进到步骤540,其中工作空间创建单元生成用于在工作空间中执行一个或多个数据对象和相关联的元数据的指令集。可以认识到的是,该指令集以可变换成等效指令集的格式生成,该等效指令集可以在托管工作空间的执行场所中执行。在步骤550

中,工作空间创建单元生成能够在执行(步骤560)时在执行场所中物理地创建工作空间的供应可执行文件(例如,脚本)。可以认识到的是,通过一些实施例,工作空间创建单元响应于确定发出创建工作空间的请求的(与客户端设备相关联的)用户被授权访问在步骤520中获得的相关数据对象和相关联的元数据而生成供应可执行文件。

[0075] 在步骤570中,工作空间创建单元在创建的工作空间中执行指令集。具体而言,工作空间创建单元实例化工作空间内提取出的数据对象、元数据和/或数据模型的一部分(即,子模型)以供进一步处理。在完成工作空间中的子模型的处理后,步骤580中的过程更新生产环境。例如,部署在生产环境201中的子模型的较旧版本可以被替换为来自工作空间的经处理的(即,经重新训练/经微调的)子模型。

[0076] 此外,通过一些实施例,数据处理系统可以提供在生产环境中维护子模型的较旧版本以及子模型的更新后的版本两者的选项。如前所述,在创建工作空间之前,可以将生产环境的初始状态存储在与生产相关联的数据存储装置中,其中初始状态包括部署在生产环境201中的数据模型的性能参数(例如,第一输出)。例如,在用(从工作空间获得的)子模型的经重新训练/经微调的版本替换数据模型的子模型的较旧版本后,数据处理系统可以生成生产环境的更新后的状态,其中更新后的状态包括部署在生产环境中的模型的新输出(即,与第一输出不同的第二输出)。

[0077] 说明性系统

[0078] 图6描绘了分布式系统600的简化图。在所示示例中,分布式系统600包括经由一个或多个通信网络610耦合到服务器612的一个或多个客户端计算设备602、604、606和608。客户端计算设备602、604、606和608可以被配置为执行一个或多个应用。

[0079] 在各种示例中,服务器612可以适于运行一个或多个服务或软件应用,该一个或多个服务或软件应用实现本公开中描述的一个或多个实施例。在某些示例中,服务器612还可以提供可以包括非虚拟和虚拟环境的其它服务或软件应用。在一些示例中,这些服务可以作为基于web的服务或云服务(诸如在软件即服务(SaaS)模型下)被提供给客户端计算设备602、604、606和/或608的用户。操作客户端计算设备602、604、606和/或608的用户进而可以利用一个或多个客户端应用与服务器612进行交互以利用由这些组件提供的服务。

[0080] 在图6所描绘的配置中,服务器612可以包括实现由服务器612执行的功能的一个或多个组件618、620和622。这些组件可以包括可以由一个或多个处理器执行的软件组件、硬件组件或其组合。应该认识到的是,各种不同的系统配置是可能的,其可以与分布式系统600不同。因此,图6中所示的示例是用于实现示例系统的分布式系统的一个示例并且不旨在进行限制。

[0081] 用户可以使用客户端计算设备602、604、606和/或608来执行一个或多个应用、模型或聊天机器人,该一个或多个应用、模型或聊天机器人可以生成一个或多个事件或模型,然后可以根据本公开的教导来实现或服务该一个或多个事件或模型。客户端设备可以提供使客户端设备的用户能够与客户端设备进行交互的接口。客户端设备也可以经由该接口向用户输出信息。虽然图6仅描绘了四个客户端计算设备,但可以支持任何数量的客户端计算设备。

[0082] 客户端设备可以包括各种类型的计算系统,诸如便携式手持式设备、通用计算机(诸如个人计算机和膝上型计算机)、工作站计算机、可穿戴设备、游戏系统、瘦客户端、各种

消息传送设备、传感器和其它感测设备等。这些计算设备可以运行各种类型和版本的软件应用和操作系统(例如,Microsoft **Windows**[®]、Apple **Macintosh**[®]、**UNIX**[®]或类UNIX操作系统、Linux或类Linux操作系统,诸如Google Chrome[™]OS),包括各种移动操作系统(例如,Microsoft Windows **Mobile**[®]、**iOS**[®]、Windows **Phone**[®]、Android[™]、**BlackBerry**[®]、Palm **OS**[®])。便携式手持式设备可以包括蜂窝电话、智能电话(例如,**iPhone**[®])、平板电脑(例如,**iPad**[®])、个人数字助理(PDA)等。可穿戴设备可以包括Google **Glass**[®]头戴式显示器和其它设备。游戏系统可以包括各种手持式游戏设备、具有互联网功能的游戏设备(例如,具有或不具有**Kinect**[®]手势输入设备的

[0083] Microsoft **Xbox**[®]游戏机、Sony **PlayStation**[®]系统、由**Nintendo**[®]提供的各种游戏系统等)等。客户端设备可以能够执行各种不同的应用,诸如各种与互联网相关的应用、通信应用(例如,电子邮件应用、短消息服务(SMS)应用),并且可以使用各种通信协议。

[0084] (一个或多个)网络610可以是本领域技术人员熟悉的任何类型的网络,其可以使用多种可用协议中的任何一种来支持数据通信,包括但不限于TCP/IP(传输控制协议/互联网协议)、SNA(系统网络体系架构)、IPX(互联网分组交换)、**AppleTalk**[®]等。仅仅作为示例,(一个或多个)网络610可以是局域网(LAN)、基于以太网的网络、令牌环、广域网(WAN)、互联网、虚拟网络、虚拟专用网(VPN)、内联网、外联网、公共电话交换网(PSTN)、红外网络、无线网络(例如,在任何电气和电子协会(IEEE) 1002.11协议套件、**蓝牙**[®]和/或任何其它无线协议下操作的网络)和/或这些网络和/或其它网络的任何组合。

[0085] 服务器612可以由以下组成:一个或多个通用计算机、专用服务器计算机(作为示例,包括PC(个人计算机)服务器、**UNIX**[®]服务器、中档服务器、大型计算机、机架安装的服务器等)、服务器场、服务器集群、或任何其它适当的布置和/或组合。服务器612可以包括运行虚拟操作系统的一个或多个虚拟机,或者涉及虚拟化的其它计算体系架构,诸如可以被虚拟化以维护服务器的虚拟存储设备的逻辑存储设备的一个或多个灵活的池。在各种示例中,服务器612可以适于运行提供前述公开中描述的功能的一个或多个服务或软件应用。

[0086] 服务器612中的计算系统可以运行一个或多个操作系统,包括以上讨论的任何操作系统以及任何商用的服务器操作系统。服务器612还可以运行各种附加服务器应用和/或中间层应用中的任何一种,包括HTTP(超文本传输协议)服务器、FTP(文件传输协议)服务器、CGI(通用网关接口)服务器、**JAVA**[®]服务器、数据库服务器等。示例性数据库服务器包括但不限于可从**Oracle**[®]、**Microsoft**[®]、**Sybase**[®]、**IBM**[®](国际商业机器)等商购获得的数据库服务器。

[0087] 在一些实施方式中,服务器612可以包括一个或多个应用以分析和整合从客户端计算设备602、604、606和608的用户接收到的数据馈送和/或事件更新。作为示例,数据馈送和/或事件更新可以包括但不限于从一个或多个第三方信息源和连续数据流接收到的**Twitter**[®]馈送、**Facebook**[®]更新或实时更新,其可以包括与传感器数据应用、金融报价机、网络性能测量工具(例如,网络监视和流量管理应用)、点击流分析工具、汽车流量监视

等相关的实时事件。服务器612还可以包括经由客户端计算设备602、604、606和608的一个或多个显示设备显示数据馈送和/或实时事件的一个或多个应用。

[0088] 分布式系统600还可以包括一个或多个数据储存库614、616。在某些示例中,这些数据储存库可以被用于存储数据和其它信息。例如,数据储存库614、616中的一个或多个可以被用于存储信息,诸如与聊天机器人性能或生成的模型相关的信息,以供服务器612在根据各种实施例执行各种功能时使用的聊天机器人使用。数据储存库614、616可以驻留在各种位置。例如,服务器612使用的数据储存库可以在服务器612本地或可以远离服务器612并且经由基于网络的连接或专用连接与服务器612进行通信。数据储存库614、616可以是不同类型的。在某些示例中,服务器612使用的数据储存库可以是数据库,例如关系数据库,诸如由Oracle公司®和其它供应商提供的数据库。这些数据库中的一个或多个可以适于响应于SQL格式的命令来实现去往和来自数据库的数据的存储、更新和检索。

[0089] 在某些示例中,应用还可以使用数据储存库614、616中的一个或多个来存储应用数据。应用使用的数据储存库可以是不同类型的,诸如,例如键-值存储库、对象存储库或文件系统支持的通用存储库。

[0090] 在某些示例中,本公开中描述的功能可以经由云环境作为服务来提供。图7是根据某些示例的其中各种服务可以被提供为云服务的基于云的系统环境的简化框图。在图7中描绘的示例中,云基础设施系统702可以提供一个或多个云服务,该一个或多个云服务可以由使用一个或多个客户端计算设备704、706和708的用户来请求。云基础设施系统702可以包括一个或多个计算机和/或服务器,其可以包括上面针对服务器612描述的那些计算机和/或服务器。云基础设施系统702中的计算机可以被组织为通用计算机、专用服务器计算机、服务器场、服务器集群或任何其它适当的布置和/或组合。

[0091] (一个或多个)网络710可以促进客户端704、706和708与云基础设施系统702之间的通信和数据交换。(一个或多个)网络710可以包括一个或多个网络。网络可以是相同的或不同的类型。(一个或多个)网络710可以支持一种或多种通信协议,包括有线和/或无线协议,以促进通信。

[0092] 图7中描绘的示例仅仅是云基础设施系统的一个示例,并且不旨在进行限制。应该认识到的是,在一些其它示例中,云基础设施系统702可以具有比图7所示的组件更多或更少的组件、可以组合两个或更多个组件、或者可以具有不同的组件配置或布置。例如,虽然图7描绘了三个客户端计算设备,但是在替代示例中可以支持任何数量的客户端计算设备。

[0093] 术语“云服务”通常用于指由服务提供商的系统(例如,云基础设施系统702)根据需要并且经由诸如互联网之类的通信网络使用户可用的服务。典型地,在公共云环境中,组成云服务提供商系统的服务器和系统与客户自己的本机服务器和系统不同。云服务提供商的系统由云服务提供商管理。客户因此可以自己利用由云服务提供商提供的云服务,而不必为服务购买单独的许可证、支持或硬件和软件资源。例如,云服务提供商的系统可以托管应用,并且用户可以经由互联网按需订购和使用应用,而用户不必购买用于执行应用的基础设施资源。云服务旨在提供对应用、资源和服务的轻松、可扩展的访问。几个提供商提供云服务。例如,由加利福尼亚州Redwood Shores的Oracle®公司提供了若干种云服务,诸如中间件服务、数据库服务、Java云服务等。

[0094] 在某些示例中,云基础设施系统702可以使用诸如软件即服务(SaaS)模型、平台即

服务 (PaaS) 模型、基础设施即服务 (IaaS) 模型以及包括混合服务模型的其它模型之类的不同模型提供一个或多个云服务。云基础设施系统702可以包括一套应用、中间件、数据库以及使得能够供给各种云服务的其它资源。

[0095] SaaS模型使得应用或软件能够作为服务通过如互联网的通信网络交付给客户,而客户不必为底层应用购买硬件或软件。例如,SaaS模型可以用于为客户提供对由云基础设施系统702托管的按需应用的访问。由 **Oracle®**公司提供的SaaS服务的示例包括但不限于用于人力资源/资本管理、客户关系管理 (CRM)、企业资源规划 (ERP)、供应链管理 (SCM)、企业绩效管理 (EPM)、分析服务、社交应用及其它的各种服务。

[0096] IaaS模型通常用于向客户提供基础设施资源 (例如,服务器、存储装置、硬件和联网资源) 作为云服务,以提供弹性计算和存储能力。**Oracle®**公司提供了各种IaaS服务。

[0097] PaaS模型通常用于提供平台和环境资源作为服务,其使得客户能够开发、运行和管理应用和服务,而客户不必采购、构建或维护此类资源。由 **Oracle®**公司提供的PaaS服务的示例包括但不限于Oracle Java云服务 (JCS)、Oracle数据库云服务 (DBCS)、数据管理云服务、各种应用开发解决方案服务等。

[0098] 云服务通常基于按需自助服务、基于订阅、弹性可缩放、可靠、高度可用和安全的方式提供。例如,客户可以经由订阅订单订购由云基础设施系统702提供的一个或多个服务。然后,云基础设施系统702执行处理,以提供客户的订阅订单中所请求的服务。例如,用户可以使用话语来请求云基础设施系统采取某个动作 (例如,意图),如上所述,和/或为本文描述的聊天机器人系统提供服务。云基础设施系统702可以被配置为提供一种或甚至多种云服务。

[0099] 云基础设施系统702可以经由不同的部署模型来提供云服务。在公共云模型中,云基础设施系统702可以由第三方云服务提供商拥有,并且云服务被提供给任何普通公众客户,其中客户可以是个人或企业。在某些其它示例中,在私有云模型下,可以在组织内 (例如,在企业组织内) 以及在向客户提供的组织内的服务操作云基础设施系统702。例如,客户可以是企业的各个部门,诸如人力资源部门、工资部门等,甚至是企业内的个人。在某些其它示例中,在社区云模型下,云基础设施系统702和所提供的服务可以由相关社区中的若干组织共享。也可以使用各种其它模型,诸如上面提到的模型的混合。

[0100] 客户端计算设备704、706和708可以是不同类型的 (诸如图6中描绘的客户端计算设备604、606和608), 并且可以能够操作一个或多个客户端应用。用户可以使用客户端设备与云基础设施系统702进行交互,诸如请求由云基础设施系统702提供的服务。例如,用户可以使用客户端设备来从聊天机器人请求信息或动作,如本公开中描述的。

[0101] 在一些示例中,由云基础设施系统702执行的用于提供服务的处理可以涉及模型训练和部署。这种分析可以涉及使用、分析和操纵数据集来训练和部署一个或多个模型。这种分析可以由一个或多个处理器执行,可能并行地处理数据,使用数据来执行模拟等。例如,大数据分析可以由云基础设施系统702执行,用于生成和训练用于聊天机器人系统的一个或多个模型。用于这种分析的数据可以包括结构化数据 (例如,存储在数据库中或根据结构化模型结构化的数据) 和/或非结构化数据 (例如,数据blob (二进制大对象))。

[0102] 如在图7的示例中所描绘的,云基础设施系统702可以包括基础设施资源730,其用于促进由云基础设施系统702提供的各种云服务的供应。基础设施资源730可以包括例如处

理资源、存储或存储器资源、联网资源等。在某些示例中,可用于服务从应用请求的存储的存储虚拟机可以是云基础设施系统702的一部分。在其它示例中,存储虚拟机可以是不同系统的一部分。

[0103] 在某些示例中,为了促进这些资源的高效供应以支持由云基础设施系统702为不同客户提供的各种云服务,可以将资源捆绑成资源模块或资源的集合(也称为“群聚(pod)”)。每个资源模块或群聚可以包括一种或多种类型的资源的预先集成和优化的组合。在某些示例中,可以为不同类型的云服务预先供应不同的群聚。例如,可以为数据库服务供应第一群聚集合,可以为Java服务供应第二群聚集合,该第二群聚集合可以包括与第一群聚集合中的群聚不同的资源组合。对于一些服务,可以在服务之间共享为供应服务而分配的资源。

[0104] 云基础设施系统702本身可以内部使用服务732,该服务732由云基础设施系统702的不同组件共享并且促进云基础设施系统702的服务供应。这些内部共享的服务可以包括但不限于安全和身份服务、集成服务、企业储存库服务、企业管理器服务、病毒扫描和白名单服务、高可用性、备份和恢复服务、用于启用云支持的服务、电子邮件服务、通知服务、文件传输服务等。

[0105] 云基础设施系统702可以包括多个子系统。这些子系统可以用软件或硬件或其组合来实现。如图7所示,子系统可以包括用户接口子系统712,该用户接口子系统712使得云基础设施系统702的用户或客户能够与云基础设施系统702进行交互。用户接口子系统712可以包括各种不同的界面,诸如web界面714、在线商店界面716(其中由云基础设施系统702提供的云服务做广告并且可由消费者购买)和其它界面718。例如,客户可以使用客户端设备,使用接口714、716和718中的一个或多个来请求(服务请求734)由云基础设施系统702提供的一个或多个服务。例如,客户可以访问在线商店、浏览由云基础设施系统702提供的云服务、并且为客户希望订阅的由云基础设施系统702提供的一个或多个服务下订阅订单。服务请求可以包括识别客户和客户期望订阅的一个或多个服务的信息。例如,客户可以为云基础设施系统702提供的服务下订阅订单。作为订单的一部分,客户可以提供识别要为其提供服务的聊天机器人系统的信息以及可选的用于聊天机器人系统的一个或多个凭证。

[0106] 在某些示例中,诸如图7中描绘的示例,云基础设施系统702可以包括被配置为处理新订单的订单管理子系统(OMS) 720。作为该处理的一部分,OMS 720可以被配置为:为客户创建账户(如果尚未创建);接收来自客户的账单和/或会计信息,该账单和/或会计信息将用于针对向客户提供所请求的服务向客户计费;核实客户信息;在核实后,为客户预订订单;以及编排各种工作流程来准备用于供应的订单。

[0107] 一旦被正确地验证,OMS 720就可以调用订单供应子系统(OPS) 724,其被配置为针对订单供应资源,包括处理资源、存储器资源和联网资源。供应可以包括为订单分配资源,以及配置资源以促进由客户订单所请求的服务。为订单供应资源的方式和供应资源的类型可以取决于客户已订购的云服务的类型。例如,根据一个工作流程,OPS724可以被配置为确定所请求的特定云服务,并且识别可能已经针对该特定云服务而被预先配置的多个群聚。为订单分配的群聚的数量可以取决于所请求的服务的大小/数量/级别/范围。例如,可以基于服务所支持的用户数量、正在请求的服务的持续时间等来确定要分配的群聚的数量。然后,可以针对特定的请求客户定制所分配的群聚,用于提供所请求的服务。

[0108] 在某些示例中,如上所述的设置阶段处理可以由云基础设施系统702作为供应处理的一部分来执行。云基础设施系统702可以生成应用ID并从云基础设施系统702自身提供的存储虚拟机中或从除云基础设施系统702以外的其它系统提供的存储虚拟机中为应用选择存储虚拟机。

[0109] 云基础设施系统702可以向发出请求的客户发送响应或通知744以在所请求的服务现在准备好使用时进行指示。在一些情况下,可以将信息(例如,链接)发送给客户,使客户能够开始使用和利用所请求服务的益处。在某些示例中,对于请求服务的客户,响应可以包括由云基础设施系统702生成的聊天机器人系统ID和识别由云基础设施系统702为与聊天机器人系统ID对应的聊天机器人系统选择的聊天机器人系统的信息。

[0110] 云基础设施系统702可以向多个客户提供服务。对于每个客户,云基础设施系统702负责管理与从客户接收到的一个或多个订阅订单相关的信息、维护与订单相关的客户数据、以及向客户提供所请求的服务。云基础设施系统702还可以收集关于客户对已订阅的服务的使用的使用统计信息。例如,可以针对使用的存储量、传输的数据量、用户的数量以及系统正常运行时间和系统停机时间量等收集统计信息。该使用信息可以用于向客户计费。计费可以例如按月周期进行。

[0111] 云基础设施系统702可以并行地向多个客户提供服务。云基础设施系统702可以存储这些客户的信息,包括可能的专有信息。在某些示例中,云基础设施系统702包括身份管理子系统(IMS) 728,其被配置为管理客户信息并提供所管理的信息的分离,使得与一个客户相关的信息不可被另一个客户访问。IMS 728可以被配置为提供各种与安全相关的服务,诸如身份服务,诸如信息访问管理、认证和授权服务、用于管理客户身份和角色及相关能力的服务,等等。

[0112] 图8图示了计算机系统800的示例。在一些示例中,计算机系统800可以用于在分布式环境以及上述各种服务器和计算机系统内实现任何数字助理或聊天机器人系统。如图8所示,计算机系统800包括各种子系统,包括经由总线子系统802与多个其它子系统通信的处理子系统804。这些其它子系统可以包括处理加速单元806、I/O子系统808、存储子系统818和通信子系统824。存储子系统818可以包括非暂态计算机可读存储介质,其包括存储介质822和系统存储器810。

[0113] 总线子系统802提供用于使计算机系统800的各种组件和子系统按照期望彼此通信的机制。虽然总线子系统802被示意性地示为单条总线,但是总线子系统的替代示例可以利用多条总线。总线子系统802可以是若干种类型的总线结构中的任何一种,包括存储器总线或存储器控制器、外围总线、使用任何各种总线体系架构的局部总线等。例如,此类体系架构可以包括工业标准体系架构(ISA)总线、微通道体系架构(MCA)总线、增强型ISA(EISA)总线、视频电子标准协会(VESA)局部总线和外围组件互连(PCI)总线,其可以实现为根据IEEE P1386.1标准制造的夹层(Mezzanine)总线,等等。

[0114] 处理子系统804控制计算机系统800的操作,并且可以包括一个或多个处理器、专用集成电路(ASIC)或现场可编程门阵列(FPGA)。处理器可以包括单核或多核处理器。可以将计算机系统800的处理资源组织成一个或多个处理单元832、834等。处理单元可以包括一个或多个处理器、来自相同或不同处理器的一个或多个核、核和处理器的组合、或者核和处理器的其它组合。在一些示例中,处理子系统804可以包括一个或多个专用协处理器,诸如

图形处理器、数字信号处理器 (DSP) 等。在一些示例中,处理子系统的一些或全部处理单元可以使用定制电路(诸如专用集成电路 (ASIC) 或现场可编程门阵列 (FPGA)) 来实现。

[0115] 在一些示例中,处理子系统804中的处理单元可以执行存储在系统存储器810中或计算机可读存储介质822上的指令。在各个示例中,处理单元可以执行各种程序或代码指令,并且可以维护多个并发执行的程序或进程。在任何给定的时间,要执行的程序代码中的一些或全部可以驻留在系统存储器810中和/或计算机可读存储介质822上,包括可能在一个或多个存储设备上。通过适当的编程,处理子系统804可以提供上述各种功能。在计算机系统800正在执行一个或多个虚拟机的情况下,可以将一个或多个处理单元分配给每个虚拟机。

[0116] 在某些示例中,可以可选地提供处理加速单元806,以用于执行定制的处理或用于卸载由处理子系统804执行的一些处理,从而加速由计算机系统800执行的整体处理。

[0117] I/O子系统808可以包括用于向计算机系统800输入信息和/或用于从或经由计算机系统800输出信息的设备和机制。一般而言,术语“输入设备”的使用旨在包括用于向计算机系统800输入信息的所有可能类型的设备和机制。用户接口输入设备可以包括,例如,键盘、诸如鼠标或轨迹球之类的指向设备、并入到显示器中的触摸板或触摸屏、滚轮、点击轮、拨盘、按钮、开关、小键盘、带有语音命令识别系统的音频输入设备、麦克风以及其它类型的输入设备。用户接口输入设备还可以包括使用户能够控制输入设备并与之交互的诸如 Microsoft **Kinect®** 运动传感器的运动感测和/或姿势识别设备、Microsoft **Xbox®** 360游戏控制器、提供用于接收使用姿势和口语命令的输入的接口的设备。用户接口输入设备还可以包括眼睛姿势识别设备,诸如从用户检测眼睛活动(例如,当拍摄图片和/或进行菜单选择时的“眨眼”)并将眼睛姿势转换为到输入设备(例如,Google **Glass®**)的输入的 **GoogleGlass®** 眨眼检测器。此外,用户接口输入设备可以包括使得用户能够通过语音命令与语音识别系统(例如,**Siri®** 导航器)进行交互的语音识别感测设备。

[0118] 用户接口输入设备的其它示例包括但不限于:三维 (3D) 鼠标、操纵杆或指示杆、游戏板和图形平板、以及音频/视频设备,诸如扬声器、数字相机、数字摄像机、便携式媒体播放器、网络摄像机、图像扫描仪、指纹扫描仪、条形码读取器3D扫描仪、3D打印机、激光测距仪、以及眼睛注视跟踪设备。此外,用户接口输入设备可以包括例如医疗成像输入设备,诸如计算机断层摄影、磁共振成像、位置发射断层摄影、以及医疗超声检查设备。用户接口输入设备也可以包括例如音频输入设备,诸如MIDI键盘、数字乐器等。

[0119] 一般而言,术语“输出设备”的使用旨在包括所有可能类型的设备和用于从计算机系统800向用户或其它计算机输出信息的机制。用户接口输出设备可以包括显示子系统、指示器灯或诸如音频输出设备的非可视显示器等。显示子系统可以是阴极射线管 (CRT)、诸如利用液晶显示器 (LCD) 或等离子体显示器的平板设备、投影设备、触摸屏等。例如,用户接口输出设备可以包括但不限于,可视地传达文本、图形和音频/视频信息的各种显示设备,诸如监视器、打印机、扬声器、耳机、汽车导航系统、绘图仪、语音输出设备和调制解调器。

[0120] 存储子系统818提供用于存储由计算机系统800使用的信息和数据的储存库或数据存储库。存储子系统818提供了用于存储提供一些示例的功能的基本编程和数据构造的有形非暂态计算机可读存储介质。存储子系统818可以存储软件(例如,程序、代码模块、指

令), 该软件在由处理子系统804执行时提供上述功能。软件可以由处理子系统804的一个或多个处理单元执行。存储子系统818还可以提供根据本公开的教导的认证。

[0121] 存储子系统818可以包括一个或多个非暂态存储器设备, 包括易失性和非易失性存储器设备。如图8所示, 存储子系统818包括系统存储器810和计算机可读存储介质822。系统存储器810可以包括多个存储器, 包括用于在程序执行期间存储指令和数据的易失性主随机存取存储器 (RAM) 以及其中存储有固定指令的非易失性只读存储器 (ROM) 或闪存。在一些实施方式中, 基本输入/输出系统 (BIOS) 可以典型地存储在ROM中, 该基本输入/输出系统 (BIOS) 包含有助于例如在启动期间在计算机系统800内的元件之间传递信息的基本例程。RAM通常包含当前由处理子系统804操作和执行的数据和/或程序模块。在一些实施方式中, 系统存储器810可以包括多种不同类型的存储器, 诸如静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM) 等。

[0122] 作为示例而非限制, 如图8所示, 系统存储器810可以加载正在被执行的可以包括各种应用 (诸如Web浏览器、中间层应用、关系型数据库管理系统 (RDBMS) 等) 的应用程序812、程序数据814和操作系统816。作为示例, 操作系统816可以包括各种版本的Microsoft **Windows®**、Apple **Macintosh®** 和/或Linux操作系统、各种商用 **UNIX®** 或类UNIX操作系统 (包括但不限于各种GNU/Linux操作系统、Google **Chrome®** OS等) 和/或移动操作系统, 诸如iOS、**Windows®** Phone、**Android®** OS、**BlackBerry®** OS、**Palm®** OS操作系统, 以及其它操作系统。

[0123] 计算机可读存储介质822可以存储提供一些示例的功能的编程和数据构造。计算机可读介质822可以为计算机系统800提供计算机可读指令、数据结构、程序模块和其它数据的存储。当由处理子系统804执行时, 提供上述功能的软件 (程序、代码模块、指令) 可以存储在存储子系统818中。作为示例, 计算机可读存储介质822可以包括非易失性存储器, 诸如硬盘驱动器、磁盘驱动器、诸如CD ROM、DVD、Blu-**Ray®** (蓝光) 盘或其它光学介质的光盘驱动器。计算机可读存储介质822可以包括但不限于: **Zip®** 驱动器、闪存存储器卡、通用串行总线 (USB) 闪存驱动器、安全数字 (SD) 卡、DVD盘、数字视频带等。计算机可读存储介质822也可以包括基于非易失性存储器的固态驱动器 (SSD) (诸如基于闪存存储器的SSD、企业闪存驱动器、固态ROM等)、基于易失性存储器的SSD (诸如基于固态RAM、动态RAM、静态RAM、DRAM的SSD、磁阻RAM (MRAM) SSD), 以及使用基于DRAM和基于闪存存储器的SSD的混合SSD。

[0124] 在某些示例中, 存储子系统818还可以包括计算机可读存储介质读取器820, 其还可以连接到计算机可读存储介质822。读取器820可以接收并且被配置为从诸如盘、闪存驱动器等的存储器设备读取数据。

[0125] 在某些示例中, 计算机系统800可以支持虚拟化技术, 包括但不限于处理和存储器资源的虚拟化。例如, 计算机系统800可以提供用于执行一个或多个虚拟机的支持。在某些示例中, 计算机系统800可以执行诸如促进虚拟机的配置和管理的管理程序之类的程序。可以为每个虚拟机分配存储器、计算 (例如, 处理器、核)、I/O和联网资源。每个虚拟机通常独立于其它虚拟机运行。虚拟机通常运行其自己的操作系统, 该操作系统可以与由计算机系统800执行的其它虚拟机执行的操作系统相同或不同。因而, 计算机系统800可以潜在地同

时运行多个操作系统。

[0126] 通信子系统824提供到其它计算机系统和网络的接口。通信子系统824用作用于从计算机系统800接收数据以及向其它系统传输数据的接口。例如,通信子系统824可以使得计算机系统800能够经由互联网建立到一个或多个客户端设备的通信信道,以用于从客户端设备接收信息以及向客户端设备发送信息。例如,当计算机系统800被用于实现图1中描绘的机器人系统120时,通信子系统可以被用于与为应用选择的聊天机器人系统进行通信。

[0127] 通信子系统824可以支持有线通信协议和/或无线通信协议两者。在某些示例中,通信子系统824可以包括用于(例如,使用蜂窝电话技术、高级数据网络技术(诸如3G、4G或EDGE(全球演进的增强数据速率)、WiFi (IEEE 802.XX族标准)、或其它移动通信技术、或其任何组合)接入无线语音和/或数据网络的射频(RF)收发器组件、全球定位系统(GPS)接收器组件和/或其它组件。在一些示例中,作为无线接口的附加或替代,通信子系统824可以提供有线网络连接(例如,以太网)。

[0128] 通信子系统824可以以各种形式接收和传输数据。在一些示例中,除了其它形式之外,通信子系统824还可以以结构化和/或非结构化的数据馈送826、事件流828、事件更新830等形式接收输入通信。例如,通信子系统824可以被配置为实时地从社交媒体网络的用户和/或诸如**Twitter**[®]馈送、**Facebook**[®]更新、诸如丰富站点摘要(RSS) 馈送的web馈送的其它通信服务接收(或发送)数据馈送826,和/或来自一个或多个第三方信息源的实时更新。

[0129] 在某些示例中,通信子系统824可以被配置为以连续数据流的形式接收本质上可能是连续的或无界的没有明确结束的数据,其中连续数据流可以包括实时事件的事件流828和/或事件更新830。生成连续数据的应用的示例可以包括例如传感器数据应用、金融报价机、网络性能测量工具(例如网络监视和流量管理应用)、点击流分析工具、汽车流量监视等。

[0130] 通信子系统824也可以被配置为将数据从计算机系统800传送到其它计算机系统或网络。数据可以以各种不同的形式(诸如结构化和/或非结构化的数据馈送826、事件流828、事件更新830等)传送给一个或多个数据库,该一个或多个数据库可以与耦合到计算机系统800的一个或多个流传输数据源进行通信。

[0131] 计算机系统800可以是各种类型中的一种,包括手持便携式设备(例如,**iPhone**[®]蜂窝电话、**iPad**[®]计算平板、PDA)、可穿戴设备(例如,Google**Glass**[®]头戴式显示器)、个人计算机、工作站、大型机、信息站、服务器机架或任何其它数据处理系统。由于计算机和网络不断变化的性质,对图8中绘出的计算机系统800的描述旨在仅仅作为具体示例。具有比图8中所绘出的系统更多或更少组件的许多其它配置是可能的。基于本文所提供的公开内容和教导,应当认识到的是,存在实现各种示例的其它方式和/或方法。

[0132] 虽然已经描述了具体的示例,但是各种修改、变更、替代构造以及等同物都是可能的。示例不限于在某些特定数据处理环境内的操作,而是可以在多个数据处理环境内自由操作。此外,虽然已经使用一系列特定的事务和步骤描述了某些示例,但是对于本领域技术人员来说显而易见的是,这并不旨在进行限制。虽然一些流程图将操作描述为顺序处理,但是许多操作可以并行或同时执行。此外,操作的次序可以被重新布置。处理可能具有图中未

包括的附加步骤。上述示例的各种特征和方面可以被单独使用或联合使用。

[0133] 另外,虽然已经使用硬件和软件的特定组合描述了某些示例,但是应该认识到的是,硬件和软件的其它组合也是可能的。某些示例可以仅用硬件或仅用软件或其组合来实现。本文描述的各种处理可以以任何组合在相同的处理器或不同的处理器上实现。

[0134] 在将设备、系统、组件或模块描述为被配置为执行某些操作或功能的情况下,这样的配置可以通过以下方式来实现,例如,通过设计电子电路来执行操作、通过对可编程电子电路(诸如微处理器)进行编程来执行操作,诸如通过执行计算机指令或代码,或被编程为执行存储在非暂态存储介质上的代码或指令的处理器或核、或其任何组合来执行操作。进程可以使用各种技术进行通信,这些技术包括但不限于用于进程间通信的常规技术,并且不同对的进程可以使用不同的技术,或者同一对进程可以在不同时间使用不同的技术。

[0135] 在本公开中给出了具体细节以提供对示例的透彻理解。但是,可以在没有这些具体细节的情况下实践示例。例如,已经示出了众所周知的电路、处理、算法、结构和技术,而没有不必要的细节,以避免使示例模糊。本描述仅提供示例性示例,并且不旨在限制其它示例的范围、适用性或配置。相反,示例的先前描述将为本领域技术人员提供用于实现各种示例的使能描述。可以对元件的功能和布置进行各种改变。

[0136] 因而,说明书和附图应被认为是说明性的而不是限制性的。但是,将显而易见的是,在不脱离权利要求书所阐述的更广泛的精神和范围的情况下,可以对其进行添加、减少、删除以及其它修改和改变。因此,虽然已经描述了具体的示例,但是这些示例并不旨在进行限制。各种修改和等同形式均在所附权利要求的范围内。

[0137] 在前述说明书中,参考其具体示例描述了本公开的方面,但是本领域技术人员将认识到本公开不限于此。上述公开的各种特征和方面可以被单独或联合使用。此外,在不脱离本说明书的更广泛的精神和范围的情况下,可以在超出本文描述的环境和应用的任何数量的环境和应用中利用示例。因而,本说明书和附图被认为是说明性的而不是限制性的。

[0138] 在前面的描述中,为了说明的目的,以特定顺序描述了方法。应该认识到的是,在替代示例中,可以以与所描述的顺序不同的顺序来执行这些方法。还应该认识到的是,上述方法可以由硬件组件执行或可以实施在机器可执行指令的序列中,这些指令可以用于使机器(诸如通用或专用处理器或用指令编程的逻辑电路)执行方法。这些机器可执行指令可以存储在一个或多个机器可读介质上,诸如CD-ROM或其它类型的光盘、软盘、ROM、RAM、EPROM、EEPROM、磁卡或光卡,闪存,或适合存储电子指令的其它类型的机器可读介质。替代地,这些方法可以通过硬件和软件的组合来执行。

[0139] 在组件被描述为被配置为执行某些操作的情况下,这种配置可以通过例如设计电子电路或其它硬件来执行操作、通过对可编程电子电路(例如,微处理器或其它合适的电子电路)进行编程来执行操作,或其任何组合来完成。

[0140] 虽然本文已经详细描述了本申请的说明性示例,但是应该理解的是,本发明构思可以以其它方式被不同地实施和采用,并且所附权利要求旨在被解释为包括这样的变型,受现有技术限制的除外。

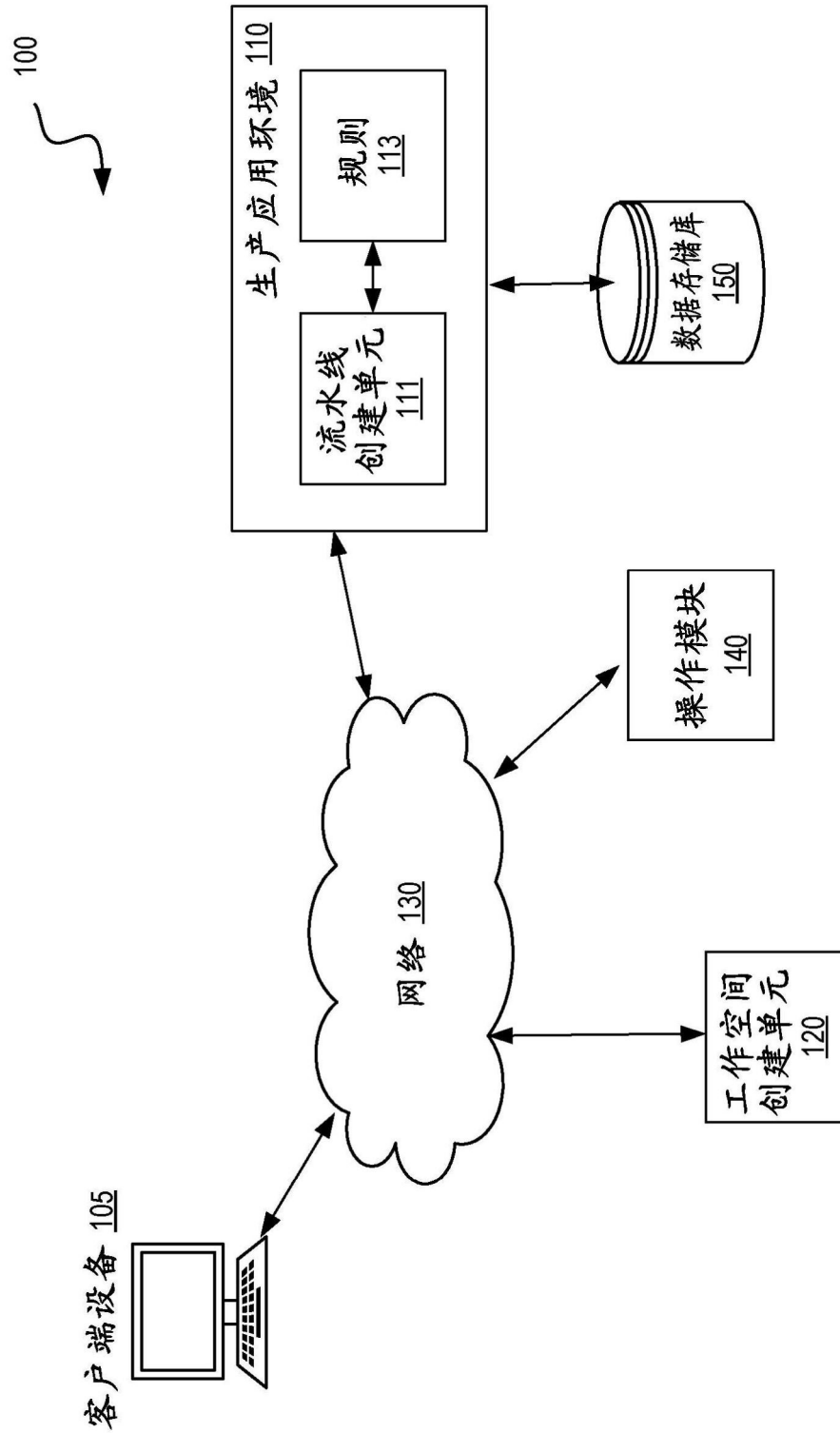


图1

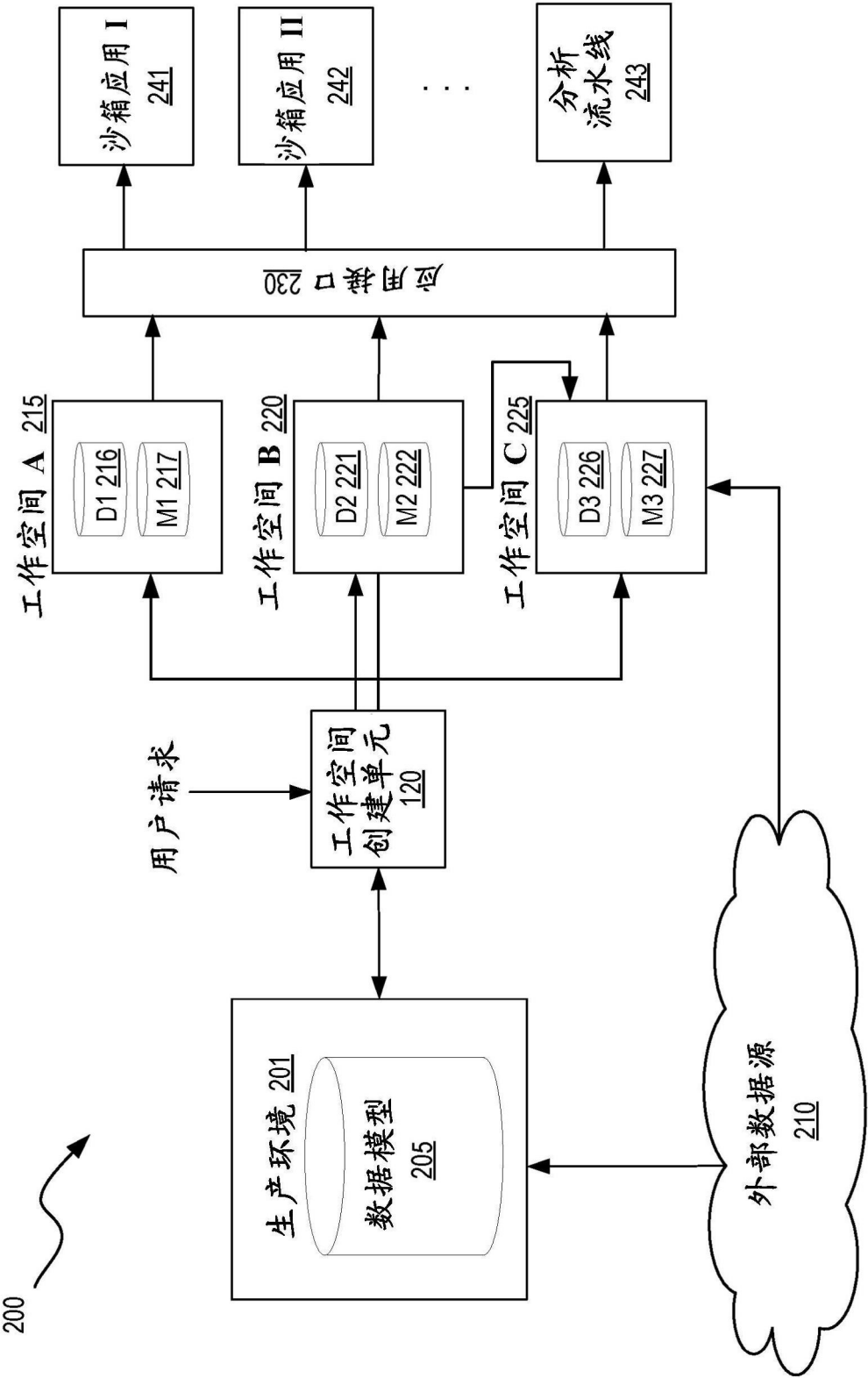


图2

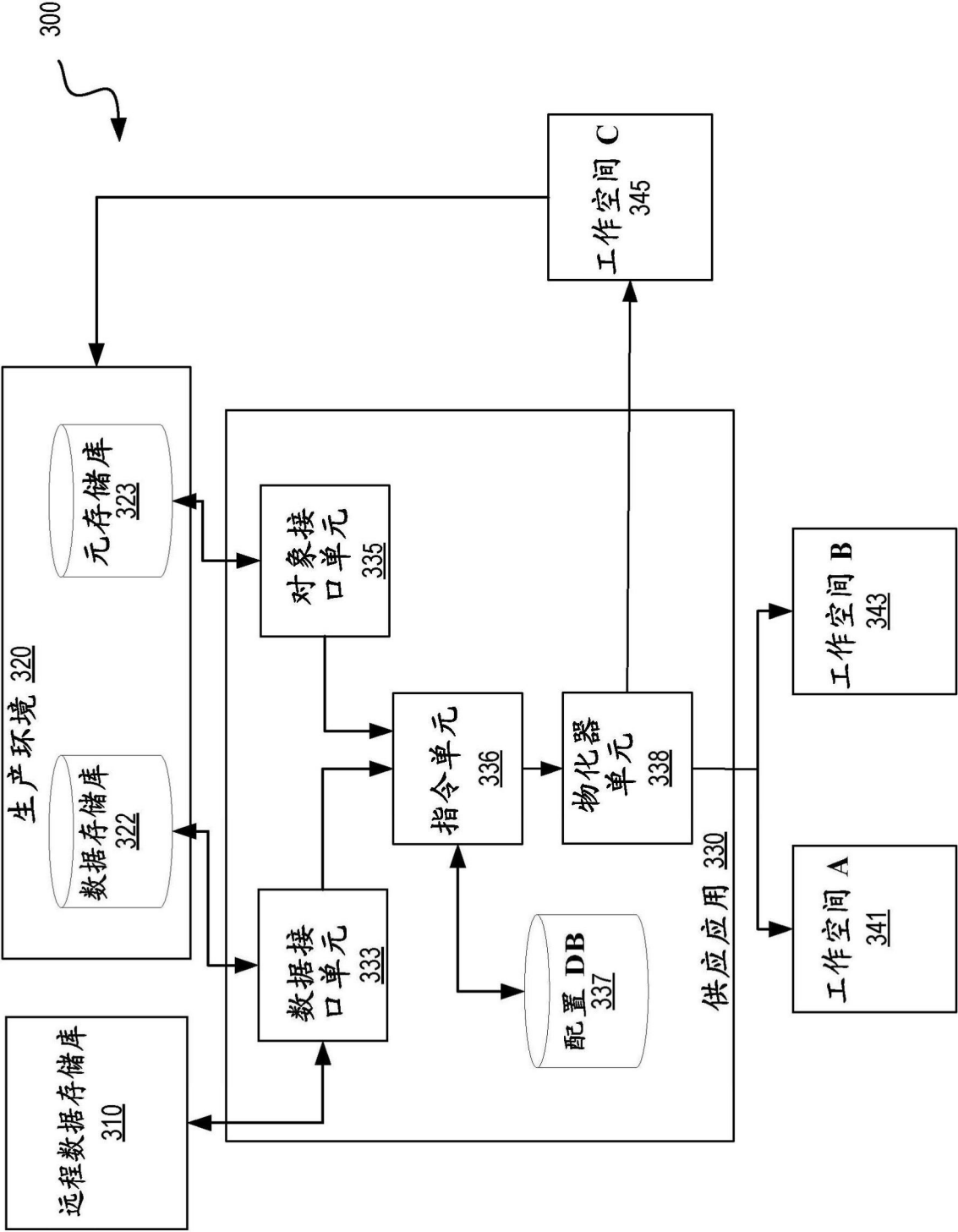


图3

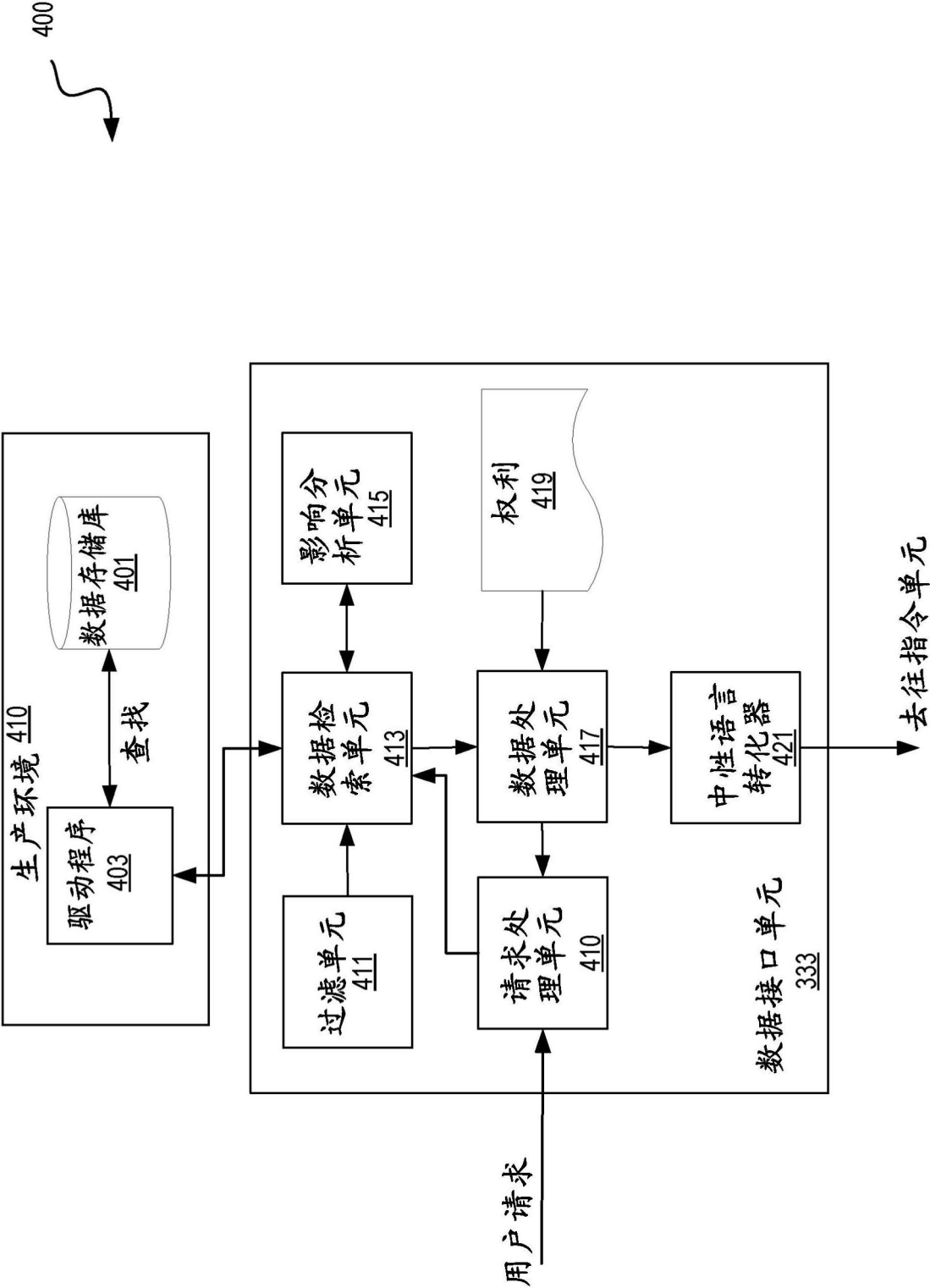


图4

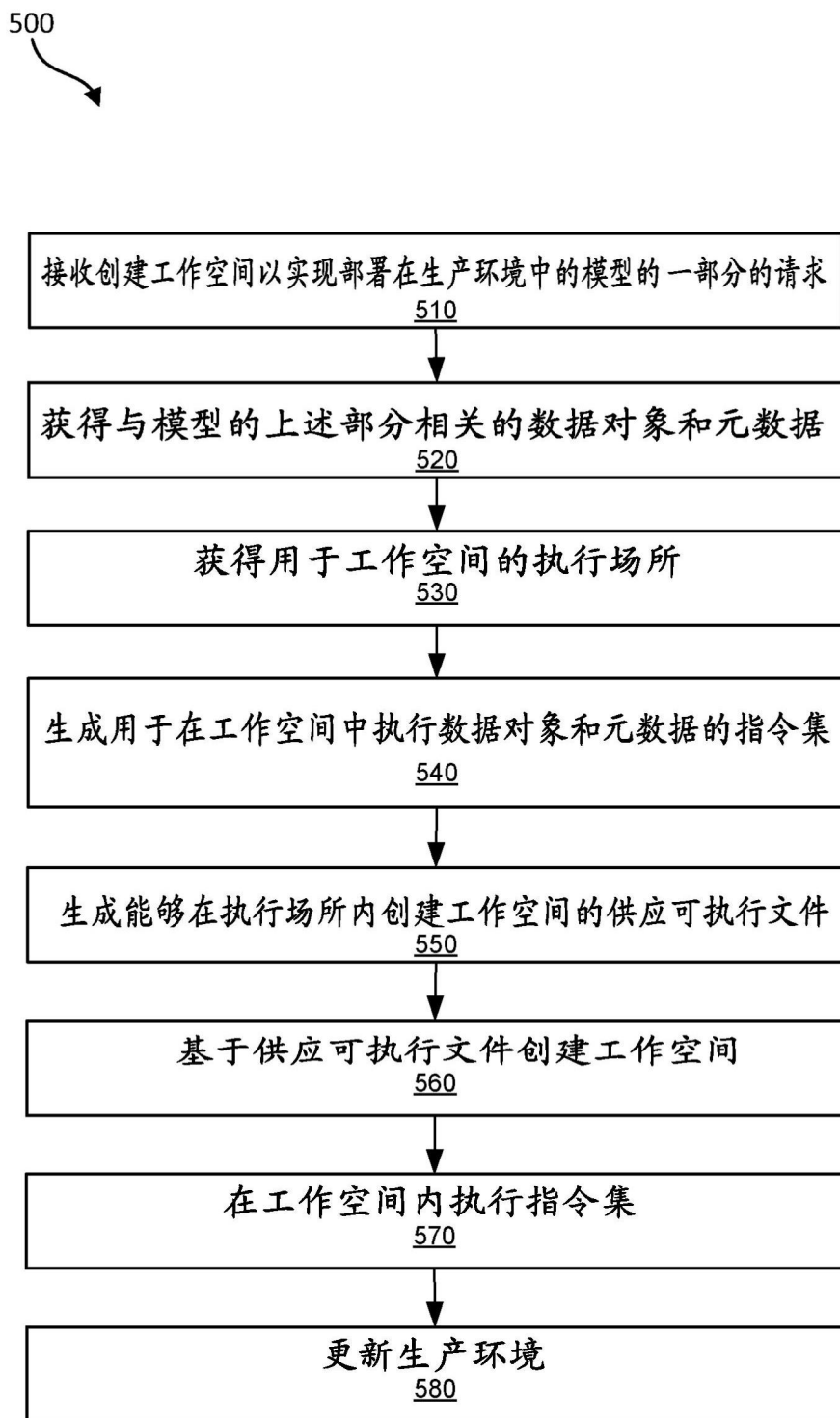


图5

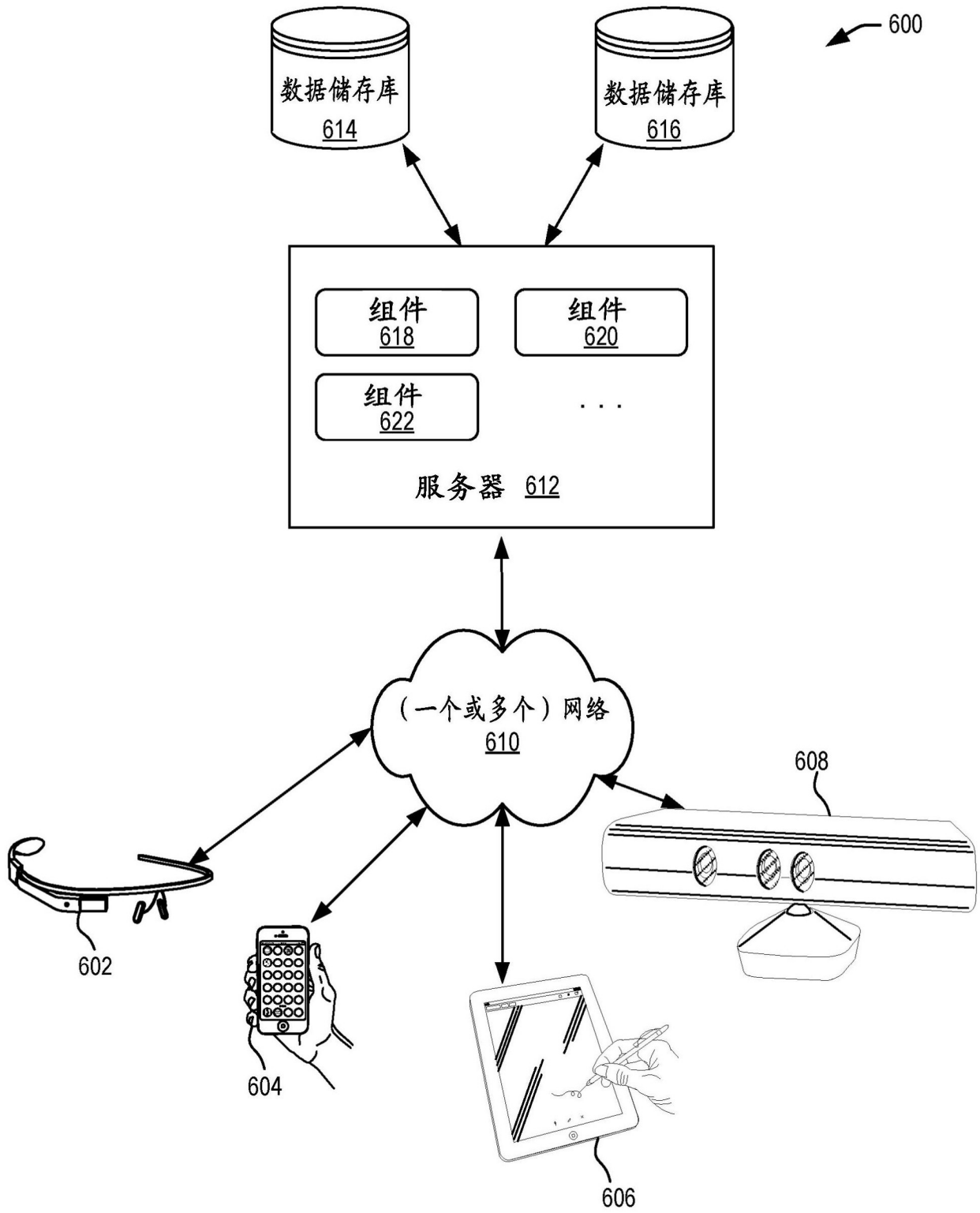


图6

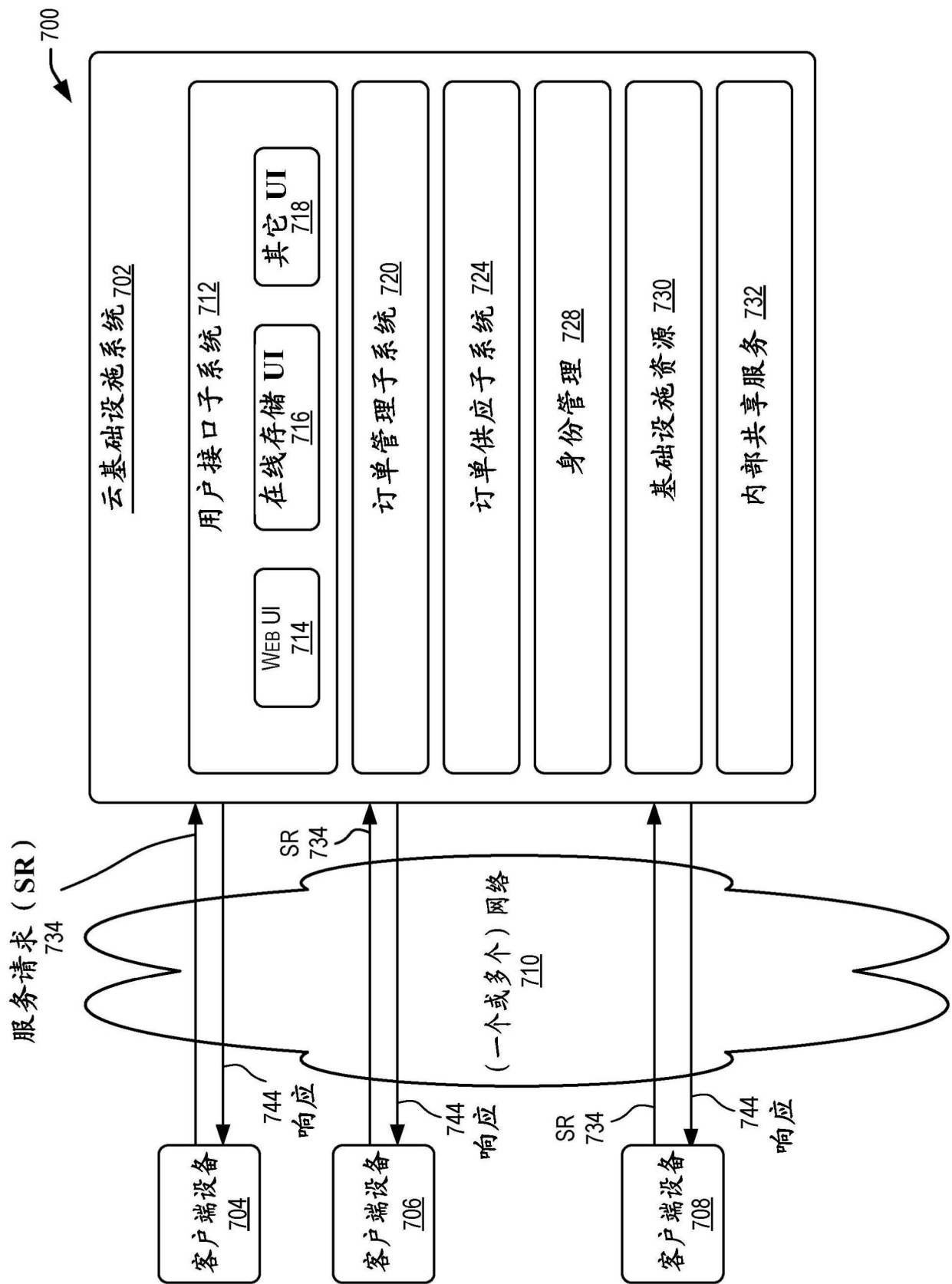


图7

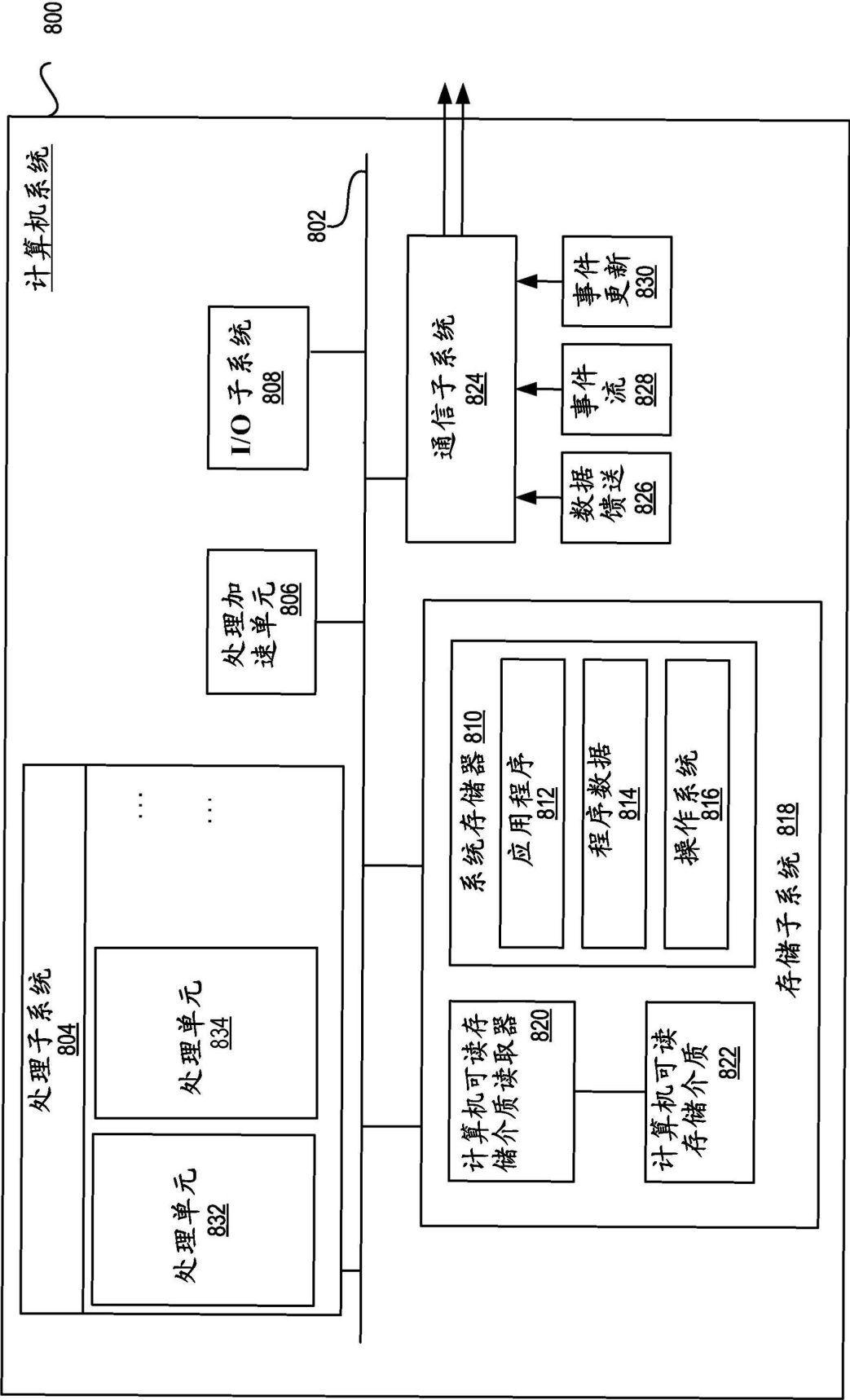


图8