



(12)发明专利

(10)授权公告号 CN 104737154 B

(45)授权公告日 2017.12.05

(21)申请号 201380054267.2

(22)申请日 2013.08.02

(65)同一申请的已公布的文献号
申请公布号 CN 104737154 A

(43)申请公布日 2015.06.24

(30)优先权数据
61/715,428 2012.10.18 US
13/804,049 2013.03.14 US

(85)PCT国际申请进入国家阶段日
2015.04.17

(86)PCT国际申请的申请数据
PCT/US2013/053373 2013.08.02

(87)PCT国际申请的公布数据
W02014/062278 EN 2014.04.24

(73)专利权人 甲骨文国际公司
地址 美国加利福尼亚

(72)发明人 L·考泽娜 J·K·瑞思

(74)专利代理机构 中国国际贸易促进委员会专
利商标事务所 11038
代理人 罗亚男

(51)Int.Cl.
G06F 17/00(2006.01)

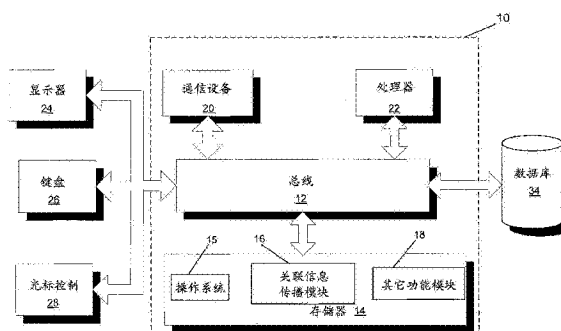
(56)对比文件
CN 1952936 A, 2007.04.25,
CN 102446187 A, 2012.05.09,
CN 102239458 A, 2011.11.09,
US 7707183 B2, 2010.04.27,
US 2007/0255741 A1, 2007.11.01,
CN 1952824 A, 2007.04.25,
审查员 陈静

权利要求书4页 说明书27页 附图19页

(54)发明名称
关联信息传播系统

(57)摘要

提供了一种用于传播关联信息的可见性的关联信息系统。该关联信息系统跟踪数据仓库的数据世系。该关联信息系统还识别主数据元素与关联对象之间的关联,其中主数据元素与关联对象之间的关联被识别为主链接。该关联信息系统还在关联对象映射系统表中存储代表主链接的第一关联对象映射系统记录。该关联信息系统还使用数据世系来识别次要数据元素与关联对象之间的关联,其中次要数据元素与关联对象之间的关联被识别为次要链接,并且其中,次要数据元素与数据世系中的主数据元素有关。该关联信息系统还在关联对象映射系统表中存储代表次要链接的第二关联对象映射记录。



1. 一种其上存储指令的非暂态计算机可读介质,当指令被处理器执行时,使该处理器传播关联信息的可见性,所述传播包括:

跟踪包括一个或多个数据表的数据仓库的数据世系;

在关联对象映射系统表中存储代表所述数据世系的主数据元素与关联对象之间的关联的第一关联对象映射系统记录,其中,所述主数据元素与所述关联对象之间的关联被识别为主链接;

使用所述数据世系来识别次要数据元素与所述关联对象之间的关联,其中,所述次要数据元素与所述数据世系中的所述主数据元素有关;以及

在所述关联对象映射系统表中存储第二关联对象映射记录,第二关联对象映射记录代表所述数据世系的所述次要数据元素与所述关联对象之间的关联并且代表所述主数据元素与所述次要数据元素之间的关联,其中,所述次要数据元素与所述关联对象之间的关联被识别为次要链接;

其中,所述关联对象包括数据世系的信息而不是数据世系的数据元素。

2. 如权利要求1所述的非暂态计算机可读介质,

其中,每个数据表包括一个或多个数据记录;

其中,所述主数据元素包括主数据记录的列;

其中,所述关联对象包括关联信息;

其中,第一关联对象映射系统记录包括所述主数据元素的标识、所述关联对象的标识以及主链接指示符;

其中,所述次要数据元素包括次要数据记录的列;以及

其中,所述第二关联对象映射记录包括所述次要数据元素的标识、所述关联对象的标识以及次要链接指示符。

3. 如权利要求1所述的非暂态计算机可读介质,所述传播还包括:

在用户接口中显示所述主数据元素或所述次要数据元素;

选择所述主数据元素或所述次要数据元素;以及

利用第一关联对象映射系统记录或第二关联对象映射系统记录在所述用户接口中显示所述关联对象。

4. 如权利要求3所述的非暂态计算机可读介质,其中,所述关联对象的所述显示用关联对象类型过滤。

5. 如权利要求1所述的非暂态计算机可读介质,

其中,第一关联对象映射系统记录包括识别所述主数据元素的数据模型的数据元素模型标识、识别所述主数据元素的数据表的数据元素表标识以及识别所述主数据元素的列的数据元素列标识;以及

其中,第二关联对象映射系统记录包括识别所述次要数据元素的数据模型的数据元素模型标识、识别所述次要数据元素的数据表的数据元素表标识和识别所述次要数据元素的列的数据元素列标识。

6. 如权利要求1所述的非暂态计算机可读介质,其中,所述关联对象是以下之一:差异、注释、状况标志、状态指示符、标签、图像、记录、文件、链接或文档。

7. 如权利要求1所述的非暂态计算机可读介质,其中,所述次要数据元素是所述主数据

元素的源数据元素。

8. 如权利要求7所述的非暂态计算机可读介质,其中,第二关联对象映射记录包括所述次要数据元素是优选源数据元素的指示。

9. 如权利要求1所述的非暂态计算机可读介质,其中,所述主数据元素的标识和所述次要数据元素的标识每一个都是代理键。

10. 如权利要求2所述的非暂态计算机可读介质,其中,所述主数据元素包括所述主数据记录;以及

其中,所述次要数据元素包括所述次要数据记录。

11. 如权利要求1所述的非暂态计算机可读介质,所述传播还包括:

删除所述主数据元素;

删除第一关联对象映射记录;

删除第二关联对象映射记录;以及

将所述关联对象的状态修改为关闭状态。

12. 一种用于传播关联信息的可见性的计算机实现的方法,该计算机实现的方法包括:

跟踪包括一个或多个数据表的数据仓库的数据世系;

在关联对象映射系统表中存储代表所述数据世系的主数据元素与关联对象之间的关联的第一关联对象映射系统记录,其中,所述主数据元素与所述关联对象之间的关联被识别为主链接;

使用所述数据世系来识别次要数据元素与所述关联对象之间的关联,其中,所述次要数据元素与所述数据世系中的所述主数据元素有关;以及

在所述关联对象映射系统表中存储第二关联对象映射记录,第二关联对象映射记录代表所述数据世系的所述次要数据元素与所述关联对象之间的关联并且代表所述主数据元素与所述次要数据元素之间的关联,其中,所述次要数据元素与所述关联对象之间的关联被识别为次要链接;

其中,所述关联对象包括数据世系的信息而不是数据世系的数据元素。

13. 如权利要求12所述的计算机实现的方法,

其中,每个数据表包括一个或多个数据记录;

其中,所述主数据元素包括主数据记录的列;

其中,所述关联对象包括关联信息;

其中,第一关联对象映射系统记录包括所述主数据元素的标识、所述关联对象的标识以及主链接指示符;

其中,所述次要数据元素包括次要数据记录的列;

其中,所述次要数据元素与所述数据世系中的所述主数据元素有关;以及

其中,所述第二关联对象映射记录包括所述次要数据元素的标识、所述关联对象的标识以及次要链接指示符。

14. 如权利要求12所述的计算机实现的方法,还包括:

在用户接口中显示所述主数据元素或所述次要数据元素;

选择所述主数据元素或所述次要数据元素;以及

利用第一关联对象映射系统记录或第二关联对象映射系统记录在所述用户接口中显

示所述关联对象。

15. 如权利要求14所述的计算机实现的方法,其中,所述关联对象的所述显示用关联对象类型过滤。

16. 如权利要求12所述的计算机实现的方法,

其中,第一关联对象映射系统记录包括识别所述主数据元素的数据模型的数据元素模型标识、识别所述主数据元素的数据表的数据元素表标识以及识别所述主数据元素的列的数据元素列标识;以及

其中,第二关联对象映射系统记录包括识别所述次要数据元素的数据模型的数据元素模型标识、识别所述次要数据元素的数据表的数据元素表标识和识别所述次要数据元素的列的数据元素列标识。

17. 如权利要求12所述的计算机实现的方法,其中,所述关联对象是以下之一:差异、注释、状况标志、状态指示符、标签、图像、记录、文件、链接或文档。

18. 一种用于传播关联信息的可见性的系统,该系统包括:

处理器;

配置为存储一条或多条指令的存储器;

数据世系模块,配置为跟踪包括一个或多个数据表的数据仓库的数据世系;

关联信息存储模块,配置为在关联对象映射系统表中存储代表所述数据世系的主数据元素与关联对象之间的关联的第一关联对象映射系统记录,其中,所述主数据元素与所述关联对象之间的关联被识别为主链接;

关联信息识别模块,配置为使用所述数据世系来识别次要数据元素与所述关联对象之间的关联,其中,所述次要数据元素与所述数据世系中的所述主数据元素有关;并且

其中,所述关联信息存储模块还配置为在所述关联对象映射系统表中存储第二关联对象映射记录,第二关联对象映射记录代表所述数据世系的所述次要数据元素与所述关联对象之间的关联并且代表所述主数据元素与所述次要数据元素之间的关联,其中,所述次要数据元素与所述关联对象之间的关联被识别为次要链接;

其中,所述关联对象包括数据世系的信息而不是数据世系的数据元素。

19. 如权利要求18所述的系统,

其中,每个数据表包括一个或多个数据记录;

其中,所述主数据元素包括主数据记录的列;

其中,所述关联对象包括关联信息;

其中,第一关联对象映射系统记录包括所述主数据元素的标识、所述关联对象的标识以及主链接指示符;

其中,所述次要数据元素包括次要数据记录的列;

其中,所述次要数据元素与所述数据世系中的所述主数据元素有关;以及

其中,所述第二关联对象映射记录包括所述次要数据元素的标识、所述关联对象的标识以及次要链接指示符。

20. 如权利要求18所述的系统,还包括:

数据元素显示模块,配置为在用户接口中显示所述主数据元素或所述次要数据元素;

数据元素选择模块,配置为选择所述主数据元素或所述次要数据元素;以及

关联对象显示模块,配置为利用第一关联对象映射记录或第二关联对象映射记录在所述用户接口中显示所述关联对象。

21. 如权利要求20所述的系统,其中,所述关联对象的所述显示用关联对象类型过滤。

22. 如权利要求18所述的系统,

其中,第一关联对象映射系统记录包括识别所述主数据元素的数据模型的数据元素模型标识、识别所述主数据元素的数据表的数据元素表标识以及识别所述主数据元素的列的数据元素列标识;以及

其中,第二关联对象映射系统记录包括识别所述次要数据元素的数据模型的数据元素模型标识、识别所述次要数据元素的数据表的数据元素表标识和识别所述次要数据元素的列的数据元素列标识。

23. 如权利要求18所述的系统,其中,所述关联对象是以下之一:差异、注释、状况标志、状态指示符、标签、图像、记录、文件、链接或文档。

24. 一种其上存储指令的非暂态计算机可读介质,当指令被处理器执行时,使该处理器传播关联信息的可见性,所述传播包括:

跟踪包括一个或多个数据表的数据仓库的数据世系;

在关联对象映射系统表中存储代表所述数据世系的主数据元素与关联对象之间的关联的第一关联对象映射系统记录,其中,所述主数据元素与所述关联对象之间的关联被识别为主链接;

使用所述数据世系来识别次要数据元素与所述关联对象之间的关联,其中,所述次要数据元素与所述数据世系中的所述主数据元素有关;以及

在所述关联对象映射系统表中存储第二关联对象映射记录,第二关联对象映射记录代表所述数据世系的所述次要数据元素与所述关联对象之间的关联并且代表所述主数据元素与所述次要数据元素之间的关联,其中,所述次要数据元素与所述关联对象之间的关联被识别为次要链接;删除所述主数据元素;

删除所述第一关联对象映射记录;

删除所述第二关联对象映射记录;和

将所述关联对象的状态修改为关闭状态。

关联信息传播系统

[0001] 对相关申请的交叉引用

[0002] 本申请要求于2012年10月18日提交的美国临时专利申请序列No.61/715,428的优先权,其主题通过引用被结合于此。

技术领域

[0003] 一种实施例针对计算机系统,并且更具体而言,针对管理数据的计算机系统。

背景技术

[0004] 数据仓库可以被用来存储用于报告和分析的数据。计算机系统可以从多个分布式源收集数据并且在数据仓库中的数据模型中存储整合的数据。用户可以对数据应用变换,为基于所存储的数据进行数据审查、数据分析和数据挖掘作准备。另外,数据常常可以与关联的信息,诸如注释、错误报告、标签和标志/状态信息,相关联。

发明内容

[0005] 一种实施例是一种传播关联信息的可见性的系统。该系统追踪包括一个或多个数据表的数据仓库的数据世系。进一步地,该系统在关联对象映射系统表中存储代表数据世系的主数据元素与关联对象之间的关联的第一关联对象映射系统记录,其中,主数据元素与关联对象之间的关联被识别为主链接。该系统还包括使用数据世系来识别次要数据元素与关联对象之间的关联,其中,在数据世系中,次要数据元素与主数据元素相关。该系统还在关联对象映射系统表中存储第二关联对象映射系统记录,该第二关联对象映射系统记录代表数据世系的次要数据元素与关联对象之间的关联并且代表主要数据元素与次要数据元素之间的关联,并且其中,次要数据元素与关联对象之间的关联被识别为次要链接。

附图说明

[0006] 更多的实施例、细节、优点和修改将从以下优选实施例的具体描述变得显然,实施例要结合附图来理解。

[0007] 图1说明了可以实现本发明实施例的系统的框图。

[0008] 图2根据本发明实施例说明了用于系统的逻辑数据模型的框图。

[0009] 图3根据本发明实施例说明了示例数据世系(lineage)的框图。

[0010] 图4根据本发明实施例说明了显示数据世系的示例用户接口。

[0011] 图5根据本发明实施例说明了用于系统的物理模型的框图。

[0012] 图6根据本发明实施例说明了数据世系映射表的框图。

[0013] 图7根据本发明实施例说明了影子表(shadow table)的框图。

[0014] 图8根据本发明实施例说明了数据世系子系统的物理模型的框图。

[0015] 图9根据本发明实施例说明了关联对象映射表的框图。

[0016] 图10根据本发明实施例说明了关联信息子系统的物理模型的框图。

[0017] 图11根据本发明实施例说明了联合变换的示例数据世系追踪。

[0018] 图12根据本发明实施例说明了旋转变换的示例源数据世系追踪。

[0019] 图13根据本发明实施例说明了旋转变换(pivot transformation)的示例目标数据世系追踪。

[0020] 图14根据本发明实施例说明了逆旋转变换(unpivot transformation)的示例源数据世系追踪。

[0021] 图15根据本发明实施例说明了逆旋转变换的示例目标数据世系追踪。

[0022] 图16根据本发明实施例说明了连接变换(join transformation)的示例数据世系追踪。

[0023] 图17根据本发明实施例说明了判例变换(case transformation)的示例数据世系追踪。

[0024] 图18根据本发明实施例说明了数据世系模块的功能的流程图。

[0025] 图19根据本发明实施例说明了关联信息传播模块的功能的流程图。

具体实施方式

[0026] 在一种实施例中,提供了可以通过对可以存储在数据仓库中的一组数据执行的一个或多个变换的途径提供数据世系追踪的系统,其中这组数据存储在一个或多个表中,其中每个表包括一个或多个数据记录。“变换”或“数据变换”是第一格式的一个或多个数据记录到第二格式的一个或多个数据记录的转换,其中数据记录可以包括一个或多个数据值。对于给定的目标数据元素,该系统可以识别产生目标数据元素的一组源数据元素以及从这组源数据元素到该目标数据元素的路径,其中“数据元素”包括存储在数据记录的列中的数据值。类似地,对于给定的源数据元素,系统可以识别从该源数据元素得出的确切的一组目标数据元素以及从该源数据元素到每个目标数据元素的路径。因此,除了数据仓库中所存储的数据之外,该系统还可以提供数据世系追踪信息,以进一步方便所存储的数据的深入分析。在某些实施例中,数据仓库中所存储的数据可以存储在一个或多个“数据记录”中,并且数据世系追踪信息可以存储在一个或多个“系统记录”中。如在这里所定义的,“数据记录”可以是存储数据仓库的数据的任何类型的记录,而“系统记录”可以是存储数据世系追踪信息的任何类型的记录。

[0027] 根据该实施例,为了追踪从源数据元素到目标数据元素以及反过来的数据世系,系统可以为每种变换存储以下数据世系追踪信息:(a)从一个或多个源表定义中的一个或多个列到变换的目标表中的每一列的一个或多个元数据映射,称为“列映射数据”,这是关于由变换执行的数据重新组织的元数据信息;(b)从一个或多个源表中的一个或多个实际数据记录到通过变换的执行所填充的目标表中的每个实际数据记录的一个或多个映射,称为“记录映射数据”;以及(c)特定于变换的附加数据世系追踪信息,诸如变换的类型。列映射数据可以在变换被定义时存储,或者它可以被存储为描述现有变换程序的行为。这种变换列映射数据可以包括:(a)源表和源列;(b)目标表和目標列;(c)变换类型;及(d)特定变换所需的附加信息。当变换被执行时,与通过变换处理的数据记录相关联的记录映射数据可以被存储。该记录映射数据可以包括:a)源数据记录的标识;b)目标数据记录的标识;及c)用于特定变换的附加信息。数据记录的标识可以是单个列,或者可以包括多个列,其值唯

一地识别数据表中的数据记录,也被称为主键或唯一键。根据该实施例,变换的目标表中的每个数据记录可以存储用于为那个目标数据记录在变换中为一个或多个列提供数据的源数据记录的标识,其中这种标识被称为“源键”。在某些实施例中,这些识别值可以组合成可被用来识别数据记录的单个值或者“代理源键”。在某些实施例中,代理键可以包括关于用来把源数据记录变换成目标数据记录的变换的附加信息。在目标数据记录包含来自多个源数据记录的数据的情形下,目标表可以利用适当数目的列来扩展,使得目标数据记录可以存储适当数目的源代理键。然后,系统可以存储目标表中保持记录映射数据的一个或多个源代理键列与和存储为列映射数据的对应源表相关联的目标列之间的关系。这种组合的数据可以被系统用来为数据元素提供源或目标数据世系追踪。

[0028] 在另一种实施例中,提供了一种系统,该系统可以充分利用数据世系追踪信息来支持链接到数据元素的“关联信息”或“关联对象”的可见性并且把关联信息(即,关联对象)传播到相关的源数据元素和目标数据元素供显示和访问。该系统还可以支持关联信息(即,关联对象)通过数据世系中到数据外部源的变换路径而导出到外部系统。

[0029] 根据该实施例,“关联对象”可以是与存储在数据仓库中的数据元素关联的任何关联信息(即,除关联对象与其关联的数据元素之外的任何信息,或者数据仓库中已经通过数据世系关联的任何其它数据元素)。一种类型的关联对象是操作对象。一个或多个关联对象还可以被称为“关联信息”。关联对象的例子是“差异”。“差异”是数据元素的一个或多个特性偏离预期特性的指示。这种偏离可以是基于一个或多个数据驱动的规则由人检测的或者由机器检测的。另外,差异可以由系统的用户手动创建,或者可以由系统自动创建,例如,作为数据验证的结果。差异的例子是错误报告。关联对象的其它例子可以包括,但不限于,“注释”、“状况标志”(status flag)、“状态指示符”(state indicator)、“标签”、“图像”、“记录”、“文件”、“链接”和“文档”。“注释”是由用户输入的关于数据元素的文本集合。“状况标志”是数据元素的状况的指示。“状态指示符”是数据元素的状态的指示。“标签”是关于数据元素的词或短语。“图像”是关于数据元素的图像。“记录”是关于数据元素的任何音频或视频的记录。“文件”是关于数据元素的计算机文件。“链接”是关于数据元素的超本文标记语言(“HTML”)链接。“文档”是关于数据元素的外部文档。关联对象可以链接到数据元素或链接到整个数据记录。

[0030] 根据该实施例,系统可以存储关联对象与和该关联对象相关联的数据元素或数据记录之间的链接。当关联对象最初被创建或者最初被关联(因为关联对象可以是预先存在的关联对象)时,数据元素或数据记录可以被标记为主链接。由系统为了提供数据世系中关联对象的可见性而生成的链接,包括数据元素或数据记录的上游以及数据元素或数据记录的下游,可以被标记为辅助链接。因此,根据该实施例,关联对象,诸如差异,可以在数据元素或数据记录的上游和数据元素或数据记录的下游都可见。为了提供这种可见性,系统可以:(a) 存储关联对象与数据元素(或数据记录)之间的链接;(b) 追踪数据元素(或数据记录)的源和目标世系;(c) 存储关联对象与来自主数据元素的源和目标追踪中的每个数据元素之间的一个或多个次要链接;及(d) 向用户接口提供主链接信息和次链接信息,用于显示关联对象的可见性。另外,相对于遍历数据世系来检测关联对象的存在,由该系统生成的链接可以被用来获得从上游或下游数据元素对关联对象的直接访问。

[0031] 图1说明了可以实现本发明一种实施例的系统10的框图。系统10包括总线12或者

用于在系统10的组件之间传送信息的其它通信机制。系统10还包括操作耦合到总线12的处理器22,用于处理信息并执行指令或操作。处理器22可以是任何类型的通用或专用处理器。系统10还包括存储器14,用于存储信息和要由处理器22执行的指令。存储器14可以是随机存取存储器(“RAM”)、只读存储器(“ROM”)、诸如磁盘或光盘的静态存储器或者任何其它类型的机器或计算机可读介质的任意组合。系统10还包括通信设备20,诸如网络接口卡或其它通信接口,以提供对网络的访问。因此,用户可以直接地,或者通过网络或任何其它方法远程地,与系统10接口。

[0032] 计算机可读介质可以是可被处理器22访问的任何可用介质。计算机可读介质可以包括易失性和非易失性介质、可移除和不可移除介质、通信介质以及存储介质。通信介质可以包括计算机可读指令、数据结构、程序模块或者诸如载波或其它传输机制的调制数据信号中的其它数据,并且可以包括本领域中已知的任何其它形式的信息交付介质。存储介质可以包括RAM、闪存存储器、ROM、可擦可编程只读存储器(“EPROM”)、电可擦除可编程只读存储器(“EEPROM”)、寄存器、硬盘、可移除的盘、压缩光盘只读存储器(“CD-ROM”),或者本领域中已知的任何其它形式的存储介质。

[0033] 处理器22还可以经总线12操作耦合到显示器24,诸如液晶显示器(“LCD”)。显示器24可以向用户显示信息。键盘26和光标控制设备28,诸如计算机鼠标,也可以操作耦合到总线12,以启用用户与系统10的接口。

[0034] 根据一种实施例,存储器14可以存储在被处理器22执行时可提供功能的软件模块。这些模块可以包括操作系统15、关联信息传播模块16以及其它功能模块18。操作系统15可以为系统10提供操作系统功能。关联信息传播模块16可以提供用于传播关联信息的可见性的功能,如以下将更详细描述。在某些实施例中,关联信息传播模块16可以包括多个模块,其中每个模块提供用于传播关联信息的可见性的特定个别功能。系统10也可以是更大系统的一部分。因此,系统10可以包括一个或多个附加的功能模块18,以包括附加的功能。例如,功能模块18可以包括提供附加功能的模块,诸如来自Oracle公司的“Oracle Life Sciences Data Hub”产品的模块。在一种实施例中,功能模块18可以包括数据世系模块,该模块可以提供用于追踪数据世系的功能,如以下将更详细描述。

[0035] 处理器22还可以经总线12操作耦合到数据库34。数据库34可以在逻辑上相关的数据记录、系统记录或数据文件的集成集合中存储数据。数据库34可以是操作数据库、分析数据库、数据仓库、分布式数据库、终端用户数据库、外部数据库、导航数据库、内存中的数据库、面向文档的数据库、实时数据库、关系数据库、面向对象的数据库,数据管理工作台,或者本领域中已知的任何其它数据库。

[0036] 图2根据本发明的实施例说明了用于系统的逻辑数据模型的框图。该逻辑数据模型可以包括数据世系传播模块200。数据世系传播模块200是可以提供用于追踪数据世系的功能以及用于传播关联信息可见性的功能的模块,如以下将更详细描述。在某些实施例中,数据世系传播模块200可以包括提供用于追踪数据世系的功能的单独关联信息传播模块,以及提供用于传播关联信息可见性的功能的单独关联信息传播模块。

[0037] 根据该实施例,数据世系传播模块200可以产生记录映射数据205、列映射数据210以及关联对象映射数据215。如前所述,记录映射数据205包括存储在变换的目标表中的数据记录中的源代理键(即,源数据记录的标识)。而且,如前所述,列映射数据210包括变换的

源表和列、变换的目标表和列、变换类型以及变换所需的任何附加信息。另外，列映射数据210可以从变换定义得出。最后，还如前所述，关联对象映射数据215包括关联对象与和该关联对象相关联的数据元素或数据记录之间的链接（称为主链接）。关联对象映射数据215还包括为了提供关联对象在上游和下游数据模型中的可见性而由数据世系传播模块200生成的链接（称为次要链接）。这种可见性可以包括关联对象的高效访问，尤其是对于在数据显示中可见地识别与关联对象相关联的数据元素或数据记录。数据世系传播模块200还可以包括公共应用编程接口（“API”）220。公共API 220是可以被提供为使得系统的用户可以把列映射数据输入数据世系传播模块200的API。这可以允许用户利用定制的计算机程序，诸如定制的过程语言/结构化查询语言（“PL/SQL”）程序或定制的统计分析系统（“SAS”）程序，存储用于数据的映射。

[0038] 根据该实施例，逻辑数据模型还可以包括变换模块225、验证模块230以及查询建立器235。变换模块225可以识别被映射数据的唯一源并且把源代理键添加到一个或多个目标数据表的定义。另外，基于一个或多个变换定义，变换模块225可以存储用于一个或多个变换的列映射数据210。变换模块225还可以允许用户为变换模块225识别出的定制变换提供定制列映射数据。变换模块225可以实现验证模块230和查询建立器235，其中验证模块230可以读取一个或多个数据记录、测试一个或多个数据记录的数据值，并且为数据值创建一个或多个差异，并且其中查询建立器235可以建立来自一个或多个数据表的数据的查询或数据库视图的查询，同时维护作为变换的从相同的元数据源列到目标列的映射。在某些实施例中，变换映射数据也可以包括“优选源路径”。当变换的目标列具有多个立即的源列时，一个列可以被识别为优选源列。这种信息可以作为列映射数据210的一部分被追踪。因而，“优选源路径”是通过优选源列的变换的路径。换句话说，优选源路径是从多条可能的源路径中被识别为比其它可能的源路径更优选的单条源路径。可以这么做，以便向单个源数据元素提供用于附连可以发送到提供源数据的源系统的查询或其它关联对象的唯一路径。

[0039] 在变换执行期间（图2中说明为变换执行处理240），记录映射数据205可以存储在一个或多个源代理键中，其中一个或多个源代理键可以存储在变换的一个或多个目标表（图2中说明为数据表245）的一个或多个辅助列中。根据该实施例，记录映射数据205和列映射数据215以及其它数据世系追踪信息的组合可以被用来显示源或目标数据元素的数据世系的追踪（图2中说明为世系追踪250），如以下更详细描述。

[0040] 根据该实施例，逻辑数据模型还可以包括关联对象模块255和适配器260。关联对象模块255可以为每个关联对象输入关联对象映射数据215中的数据。来自关联对象映射数据215的数据可以被用来可见地加亮具有与其关联的关联对象的一个或多个数据元素（在图2中说明为加亮关联对象250）。根据某些实施例，关联对象模块255可以为与关联对象相关联的每个数据元素请求优选源路径。另外，一个或多个关联对象可以利用适配器260在系统中创建。适配器260可以使用关联对象模块255在系统中创建一个或多个关联对象，然后关联对象模块255可以使用数据世系传播模块200在源或目标数据元素的数据世系追踪中传播关联对象的可见性。在某些实施例中，逻辑数据模型还可以包括显示来自一个或多个数据表的一个或多个数据元素并且加亮与这—个或多个数据元素关联的一个或多个关联对象的存在的附加模块（图2中未说明）。

[0041] 图3根据本发明实施例说明了示例数据世系的框图。根据实施例，系统可以生成存

储在数据仓库中的数据的数据世系的追踪。在这种数据世系追踪中,系统的用户可以选择目标表中的数据元素并且请求向上游通过一个或多个变换步骤到达得出被选数据元素的一个或多个源数据元素的追踪。另外,用户可以选择源表中的数据元素并且请求向下游通过一个或多个变换步骤到达从被选数据元素得出的一个或多个目标数据元素请求的追踪。关于变换,根据该实施例,存在不同类型的变换。此类变换类型可以包括:(a)直接映射变换(有可能包括过滤和/或导出);(b)联合变换;(c)连接变换(有可能包括单个和/或重复的数据);(d)水平到垂直逆旋转变换;及(e)垂直到水平旋转变换。不同的变换类型在以下更详细地描述。在某些实施例中,可以生成数据世系追踪,以包括通过被创建用于为多步变换存储数据的一个或多个分级表的数据世系。根据该实施例,分级表是可以根据需要用于一个或多个变换的数据仓库的表,并且可以起到对多步变换场景充当中间存储结构的目的。

[0042] 直接映射变换是源表中的源数据记录的一个或多个列在目标表的目标数据记录中被重命名和/或重新排序的变换。在某些实施例中,在直接映射变换中,源表中的源数据记录的一个或多个列可以不变地传播。联合变换是来自两个或更多个源表的一个或多个源数据记录组合到单个目标表的两个或更多个目标数据记录中的变换,使得目标表包括来自两个或更多个源表的一个或多个源数据记录当中的一些或全部。连接变换是来自两个或更多个源表的一个或多个源数据记录组合到单个目标表的单个目标数据记录中的变换,使得目标表的该目标数据记录包括来自两个或更多个源表的一个或多个源数据记录当中的全部列。

[0043] 另外,水平到垂直逆旋转变换是来自源表的一个或多个源数据记录被转换成目标表的一个或多个目标数据记录的变换,其中每个源数据记录的每组一个或多个列被转换成该目标表的单独目标数据记录。例如,如果源表具有包括列A、B、C的单个源数据记录,则逆旋转变换把源数据记录变换成三个目标记录:包括列A的第一目标数据;包括列B的第二目标数据;和包括列C的第三目标数据。垂直到水平旋转变换是来自源表的一个或多个源数据记录被转换成目标表的目标数据记录的变换,其中这一个或多个源数据记录被组合到单个目标数据记录中。例如,如果源表包括三个源数据记录,数据记录1、数据记录2和数据记录3,则旋转变换把这三个源数据记录变换成包括数据记录1、2和3的一个或多个列的单个目标数据记录。判例变换是用于目标数据记录的数据值基于条件从可以来自不同源表的一个或多个源列填充的变换。

[0044] 根据实施例,可以使数据世系追踪信息可用于在系统的用户接口中显示。现在根据示例实施例更详细地描述可以提供给用户接口的信息。对于被选的数据元素的源世系,可以提供产生该被选数据元素的数据元素的层次。这是通过所有关联的变换遍历返回到数据仓库中原始源数据的数据元素路径。该路径可以包括通过数据的若干路径分支。类似地,对于被选数据元素的目标世系,可以提供从该被选数据元素得出的数据元素的层次。

[0045] 根据该实施例,可以为路径中的每个数据元素提供一个或多个以下属性:(a)数据元素在层次中的放置(例如,层次中该数据元素的标识以及其父数据元素的标识);(b)数据模型名称;(c)表名;(d)列名;(e)数据元素的数据类型;(f)数据元素值;(g)优选源路线(是或者否);(h)主键;(i)数据是否是(或者可能是)被屏蔽数据的指示符;(j)数据是否来自内部分级表的指示符;(k)一个或多个关联对象(诸如差异)主要链接到该数据元素的指示符;或者(l)底层数据(例如,表对象标识、列对象标识、代理键值以及任何其它相关的数据)。

“数据模型”是与数据仓库的一个或多个数据表关联的数据方案。

[0046] 图3根据本发明实施例说明了与一系列医学研究相关的病人数据的数据世系。但是,在其它实施例中,数据可以是任何类型的生命科学或医学数据。另外,在其它实施例中,数据可以是任何类型的数据。根据图3所说明的实施例,数据仓库可以包括四个数据模型:源数据模型310、所有站点数据模型320、审查数据模型330以及分析数据模型340。源数据模型310包括三个表:人口统计表311、站点A访问表312和站点B访问表313。人口统计表311包括用于研究中的所有主体的人口统计数据,其中人口统计数据包括每个主体的高度和高度单位。站点A访问表312对于在研究站点A的每个主体的每次访问包括一个数据记录(即,访问数据),其中访问数据包括主体的重量和重量单位。站点B访问表313对于在研究站点B的每个主体的每次访问包括一个数据记录(即,访问数据),其中访问数据包括主体的重量和重量单位。所有站点数据模型320包括一个表:访问表321。访问表321是通过包括在站点A访问表312中的访问数据和包括在站点B访问表313中的访问数据的联合变换创建的站点A访问表312和站点B访问表313的联合。审查数据模型330包括一个表:里程碑访问表331。里程碑访问表331是利用源过滤变换从访问表321中所包括的访问数据得出的每个病人的重要访问的子集(例如,对于每个主体而言,访问1、4、7和10)。分析数据模型340包括一个表:BMI表341。BMI表341是通过人口统计表311中所包括的人口统计数据和里程碑访问表331中所包括的访问数据的连接变换创建的人口统计表311和里程碑访问表331的连接,其中身体质量指数(“BMI”)值是利用主体的重量、重量单位、高度和高度单位得出的。

[0047] 图4根据本发明实施例说明了显示数据世系400的示例用户接口。根据该示例实施例,图3的BMI表341可以包括以下数据:

[0048]

研究	主体	访问	站点	性别	年龄	年龄单位	BMI	BMI 单位
BMI	1001	4	MGH	女	34	岁	22.7	KG/M**2
BMI	1002	4	MGH	女	24	岁	24.5	KG/M**2
BMI	1003	4	MGH	男	42	岁	21.9	KG/M**2
BMI	1004	4	MGH	男	28	岁	26.3	KG/M**2
BMI	1005	4	MGH	女	34	岁	24.6	KG/M**2
BMI	1006	4	MGH	女	29	岁	28.8	KG/M**2
BMI	1007	4	McLean	女	42	岁	33.3	KG/M**2
BMI	1008	4	McLean	女	35	岁	22.5	KG/M**2
BMI	1009	4	McLean	男	34	岁	33.6	KG/M**2
BMI	1010	4	McLean	女	54	岁	30.6	KG/M**2

[0049] 根据该实施例,系统可以在系统的用户接口中显示以上数据。系统的用户可能想调查主体1007、女性、年龄42且BMI值为33.3的源数据。用户可以在用户接口中选择用于主体1007的数据值33.3并且指示用户希望审查该数据值的源数据世系。作为响应,系统可以

在用户接口中显示图4的数据世系400,这在以下更详细地进一步描述。

[0050] 根据所说明的实施例,数据世系400包括多行(即,行401、402、403、404、405、406、407、408和409),其中每一行包括关于一个源数据元素的信息。这些行遍历通过图3所示的这组变换的路径。因此,直接祖先数据元素处于数据世系400中被选数据元素下面的层次级别,并且它们的祖先遵循相似的方式。因此,层次路径向上游从最靠近被选数据元素的数据元素遍历到原始源数据元素。因此,在图4所说明的实施例中,行401代表被选的数据元素(即,用于主体1007的数据值33.3)。行402、403、404和407代表由行401代表的数据元素的直接祖先数据元素。因此,由行401代表的被选数据元素的直接祖先表是图3的人口统计表311和里程碑访问表331。行405代表由行404代表的数据元素的直接祖先数据元素。行406代表由行405代表的数据元素的直接祖先数据元素。行408代表由行407代表的数据元素的直接祖先数据元素。行409代表由行408代表的数据元素的直接祖先数据元素。因此,由行404代表的数据元素以及由行407代表的数据元素的直接祖先表是访问表321。另外,由行405代表的数据元素和由行408代表的数据元素的直接祖先表是站点B访问表313。虽然数据世系400是源数据世系的例子,但是,在备选实施例中,目标数据世系可以在用户接口中显示,其中用于被选数据元素的一个或多个后代数据元素可以显示,其中目标数据世系包括多个行,并且其中每行包括关于一个目标数据元素的信息。

[0051] 数据世系400还包括多个列(即,列410、420、430、440和450)。列410代表数据元素的数据模型名、表名和列名。列420代表数据元素的数据值。列430代表数据元素的数据类型和尺寸。列440代表数据元素是否是优选源路径的一部分的指示(即,“是”或者“否”)。列450代表数据元素的主键值。

[0052] 因此,数据世系400可以是显示数据元素的层次的树表,其中列的标题是静态的,并且用户可以利用树表的成分打开或关闭层次的部分。另外,在一种实施例中,链接到关联对象的数据元素(在图4中说明为数据元素460)可以利用背景色加亮。加亮的背景色可以代表数据元素与关联对象之间的主链接。在主链接的情况下,用户可以选择加亮的数据元素并且请求显示链接到该数据元素的关联对象的细节。

[0053] 在一种实施例中,如果数据世系400包括通过内部分级表的路径,则对应的数据值可以用符号,诸如星号(图4中未说明),做标记。另外,脚注可以在数据世系400下面显示,以解释符号的意思。此外,在一种实施例中,如果数据元素是来自安全性已经应用到其的表或列,并且被屏蔽的数据代替实际数据显示,则数据元素可以利用符号,诸如符号“S”(图4中说明为符号470),做标记。另外,脚注可以在数据世系400下面显示,以解释符号的意思。此外,在一种实施例中,完全的数据世系追踪可能不能完成,或者因为由于应用到数据的安全性而使得用户不能访问追踪路径中的一些数据(上游或下游),或者因为数据记录已经从表中删除并且删除还没有传播到下游表。在任一种情况下,根据该实施例,用户接口都可以显示指示数据追踪已经终止的消息(图4中未说明)。

[0054] 因此,在一种实施例中,系统的用户可以通过数据浏览器导航到数据列表页面并且在用户接口中显示的数据网格上选择数据元素。“数据网格”是给出数据表的行和列或者数据视图的显示。在选择数据元素时,用户可以指示用户期望看到被选数据元素从其得出的源数据。作为响应,系统可以获得源数据元素的集合,并且连同关于用于每个源数据元素的数据模型、表和列的对应信息一起显示源数据元素值的集合。该集合可以被组织并且包

括信息,使得该集合可以在用户接口中显示为组成用于被选数据元素的源数据的路径树。信息也可以被包括,使得优选源路径可以被显示。

[0055] 在另一种实施例中,用户可以通过数据浏览器导航到数据列表页面并且在用户接口中显示的数据网格上选择数据元素。在选择数据元素时,用户可以指示用户期望看到得自被选数据元素的目标数据。作为响应,系统可以获得目标数据元素的集合,并且连同关于用于每个目标数据元素的数据模型、表和列的对应信息一起显示目标数据元素值的集合。该集合可以被组织并且包括信息,使得该集合可以在用户接口中显示为组成用于被选数据元素的目标数据的路径树。

[0056] 另外,关于关联对象,在一种实施例中,在数据仓库中创建或者以别的方式与数据仓库的数据元素相关联的关联对象可以在上游或下游数据模型中被查看,就好像这些对象已经传播到不同的关联对象一样。系统可以自动地把在一个数据模型中创建的关联对象与其在其它数据模型中的对应源和目标数据元素相关联。关联对象与源和目标数据模型中的数据元素之间的关联可以在系统的用户接口中显示,其中关联对象的存在可以连同关联的数据元素以及关联对象分类一起显示。另外,当多个源数据元素被用来得出目标数据元素时,在变换定义期间,可以选择一个源数据元素作为优选源数据元素。优选源数据元素可以构成向上游通过源数据的数据世系路径。最上游的源数据元素可以被用来把关联对象导出到外部系统。另外,当与关联对象相关联的数据被更新时,关联对象的一个或多个标志可以设为指示数据已更新。

[0057] 如前所述,可以使数据元素与关联对象之间的链接可用于在显示数据世系400的用户接口中显示。现在更详细地描述由系统提供给用户接口用于识别具有与其相关联的关联对象的数据元素或数据记录的信息。系统可以向用户接口提供信息,使得具有与其相关联的关联对象的每个数据元素在显示中被加亮。这可以包括关联对象最初在其中创建或者以别方式关联(即,主链接)的表中的数据元素。

[0058] 在一种实施例中,系统可以提供:(a) 关联对象的标识(这可以被用来访问关联对象及其属性);(b) 数据元素参考(诸如表对象标识、列对象标识以及代理键);(c) 与数据元素相关联的关联对象是主链接还是次要链接的指示;(d) (对于次要链接)该链接是主数据元素的源路径还是目标路径的指示;及(e) 关联对象的状态。根据该实施例,如图4中由数据元素460所说明的,系统可以利用背景色加亮数据元素,以指示数据元素链接到关联对象。系统可以利用第一背景色加亮数据元素,以指示主链接。作为替代,系统可以利用第二背景色加亮数据元素,以指示次要链接。如前所述,用户能够请求关联对象链接到的数据元素的源或目标世系。当用户到达包括到关联对象的主链接的数据元素时,用户可以查看关联对象,包括关联对象的一个或多个特性。用户还可以把关联对象导出到外部源。在备选实施例中,不是利用颜色加亮数据元素来指示该数据元素与关联对象相关联,而是系统可以使用备选的指示。备选指示的一些例子包括,但不限于,图标、带色的边界或者音频指示。

[0059] 因此,在一种实施例中,系统的用户可以导航到数据列表、在用户接口中显示的数据网格上选择数据元素、创建关联对象并且把关联对象与被选数据元素相关联。然后,系统可以把关联对象与关于被选数据元素的数据元素相关联。系统可以把新创建的关联对象链接到被选数据元素,并且把该链接指定为主链接。通过使用与数据仓库中现有变换相关联的列映射数据和记录映射数据,系统还可以递归地找出用于被选数据元素的所有源和目标

数据元素。系统随后可以把相同的关联对象链接到相关的源和目标数据元素,并且把这些链接指定为次要链接。因此,在每种相关的源或目标模型中,具有到关联对象的次要链接的数据元素都可以在系统的用户接口中被加亮。

[0060] 在另一种实施例中,系统的用户可以导航到数据列表,在那里,系统可以提供关联对象与主源数据之间的主链接和次要链接。系统还可以从关联对象获得属性,诸如每个关联对象的状态,以过滤显示。在还有另一种实施例中,用户可以删除包括到一个或多个关联对象的链接的数据元素。系统可以除去到与被删除数据相关联的关联对象的主和/或次要链接。当用于主链接的数据被删除时,系统可以除去到与数据元素相关联的关联对象的主和次要链接,并且系统可以删除关联对象。当用于次要链接的数据被删除时,系统可以除去该次要链接并且确定是否应当除去更多链接(并且如果必要的话就除去任何更多链接)。在另一种实施例中,系统可以允许用户创建查询,该查询可以访问列映射数据以及链接的关联对象,以支持关联对象的可见性。

[0061] 在用于数据世系传播模块(诸如图2的数据世系传播模块200)的控制流中,根据一种实施例,一个或多个数据模型可以在数据仓库中定义。用于每个数据模型的元数据被创建或加载。利用变换模块(诸如图2的变换模块225)为一个或多个目标表定义一个或多个变换。对于特定的变换,在需要的时候,一个或多个辅助列添加到每个目标表。这些辅助列在以下更详细地进一步描述。随后,数据世系映射表利用关于列映射的数据世系追踪信息(即,元数据信息)填充,该信息与存储在数据仓库的一个或多个数据表中的数据不同,并且其中数据世系映射表在以下更详细地进一步描述。辅助列标识是列映射元数据的一部分,使得被映射的目标列可以与为被映射的源列保持源表记录的标识的辅助列链接。一个或多个数据模型安装有一个或多个表。一个或多个源表加载数据。变换被执行并且填充一个或多个目标表。在一个或多个目标表加载数据之后,影子表利用数据世系追踪信息(即,记录映射信息)被填充或更新,该信息也与存储在数据仓库的一个或多个数据表中的数据不同,其中影子表在以下更详细地进一步描述。最后,用于关联对象的一个或多个次要链接在数据中传播。次要链接可以被用来获得上游或下游数据元素与关联对象的关联,而无需为了检测关联对象的存在而遍历数据世系映射表和影子表的数据世系追踪信息。

[0062] 因此,根据实施例,数据世系传播模块(诸如图2的数据世系传播模块200)可以维护一个或多个关联,并且可以在两种类型的对象之间追踪:(a)表中数据记录的数据元素;及(b)表中的完整数据记录。关于(a),以下可以被用来唯一地识别数据元素,并且还在为了获得数据元素需要的时候追踪所安装的对象的名字:(1)识别数据仓库的数据模型的模型对象标识;(2)识别数据仓库的表的表实例对象标识;(3)识别代理键的代理键值;及(4)识别表的列的列对象标识。关于(b),一些关联对象可以链接到数据记录而不是该数据记录中特定的数据元素。在这种场景下,由于数据记录中没有特定的列,因此对代理键使用列对象标识,而不是识别表的列的列对象标识。因此,以下可以被用来唯一地识别数据记录:(1)识别数据仓库的数据模型的模型对象标识;(2)识别数据仓库的表的表实例对象标识;(3)识别代理键的代理键值;及(4)用于代理键的列对象标识。在备选实施例中,唯一识别数据仓库的表的唯一表标识符(跨数据仓库的所有数据模型)可以被用来唯一地识别数据记录。

[0063] 另外,根据该实施例,数据世系传播模块维护并传播与关联对象的关联。数据世系传播模块还提供两种类型的关系映射:(a)特定数据元素到一个或多个源或目标数据元素

的映射(即,数据世系可追踪性);及(b)关联对象与数据元素或数据记录之间的映射(即,关联对象的可见性)。

[0064] 根据该实施例,在关于数据仓库定义一个或多个表之后,为了提供数据世系可追踪性,数据世系传播模块利用数据世系追踪信息填充数据世系映射表。数据世系映射表在以下更详细地进一步描述。由于一个或多个变换在数据仓库中被执行,因此一个或多个辅助列被填充,并且影子表加载用于追踪一个或多个数据元素的数据世系的数据世系追踪信息。辅助列和影子表在以下更详细地进一步描述。由于关联对象被创建和/或与数据元素关联,因此数据世系传播模块可以给与源和目标数据元素关联的关联对象加索引以被其它过程高效可视化。因此,上游或下游数据元素与关联对象的关联可以直接获得,而无需遍历数据世系映射表和影子表的数据世系追踪信息。

[0065] 在数据模型和变换定义阶段期间,变换的目标表可以被识别,并且一个或多个辅助列可以添加到识别出的表,作为一个或多个元数据定义。辅助列的数目可以依赖于变换的类型。一个或多个辅助列的主要用途是存储来自变换的源表的源代理键值。如前所述,代理键是数据记录的标识,其中源代理键是源数据记录的标识。一个或多个辅助列可以利用一个或多个源代理键值,或者适于变换的其它数据值,填充。数据世系传播模块可以使用存储在辅助列中的代理键值(或者其它适当的数据值)来回溯通过对目标数据做贡献的源数据元素的路径。

[0066] 如前所述,在对特定变换需要的时候,一个或多个辅助列可以添加到一个或多个目标表。可以为不同类型的变换添加的辅助列的例子如下:

[0067]

变换	辅助列
联合	1 个源代理键列(源代理键可以包含源表标识作为其第一成分)
连接	源代理键列: 对该连接做贡献的每个表一列
水平到垂直逆旋转	1 个源代理键列
垂直到水平旋转	1 个 源代理键列(值可以被修改,以保持代理键根,即,具有用于保持旋转列值的位置的占位符)
直接映射	1 个源代理键列(这可以包括直接映射类型的导出)
判例选择	源代理键列: 连接到源表和列的过滤值(判例标识符)一列, 并且用于源代理键的一列

[0068] 根据该实施例,追踪数据元素的数据世系依赖于一个或多个代理键。每个表可以包括代理键列,该列可以充当表的加索引的唯一键。为了从变换的目标表追踪回源表,来自源表的代理键存储在目标表中。在简单变换的例子中,为了在目标表中追踪被选数据元素的源数据元素,在数据世系映射表中查找目标表(以下更详细地描述),以获得源表、源列和

存储源代理键的辅助列。利用辅助列,获得源代理键的值并将其存储在包含被选数据元素的目标表记录中。源表、源列和源代理键值被用来找出被选目标数据元素的源数据值。用于更复杂的变换的数据世系追踪在以下更详细地描述。

[0069] 为了追踪被选源数据元素的目标数据元素,维护数据世系追踪信息的单独表,被称为影子表(以下更详细地描述)。影子表包括源数据记录到目标数据记录的一个或多个映射。更具体而言,影子表可以包括利用特定的源键连同识别列与源键之间的元数据关系的上下文映射而去规格化的源记录-目标记录映射。因此,根据该实施例,在变换执行结束时,影子表可以包括提供映射数据的加索引版本的一个或多个系统记录,这允许利用数据世系映射表从源数据元素追踪目标数据元素。

[0070] 图5根据本发明实施例说明了用于系统的物理模型的框图。图5说明了用于系统的代表性物理模型。在某些实施例中,系统可以不包括图5中所说明的物理模型的全部组件。在其它实施例中,系统可以包括未在图5中说明的附加组件。数据世系传播模块(诸如图2的数据世系传播模块)可以包括包含一个或多个组件的物理模型,其中每个组件包括一个或多个表。在图5所说明的实施例中,物理模型包括数据世系映射组件510,该组件包括用于从目标数据元素追踪源数据元素的数据世系追踪信息(即,元数据信息)。数据世系映射组件510包括数据世系映射表,这在以下联系图6更详细地进一步描述。另外,物理模型包括影子表组件520,该组件包括用于从源数据元素追踪目标数据元素的加索引的数据世系追踪信息。影子表组件520包括影子表,这在以下联系图7更详细地进一步描述。此外,物理模型包括变换组件530,该组件包括用于一个或多个变换的数据。另外,物理模型包括关联对象映射组件540,该组件包括关于关联对象与数据元素或数据记录之间的链接的数据。关联对象映射组件540包括关联对象映射表,这在以下联系图9更详细地进一步描述。此外,物理模型包括表定义组件550,该组件包括用于一个或多个源表和一个或多个目标表的应用表定义。在一种实施例中,表定义组件550可以包括处于数据模型级的一个或多个应用表定义(例如,图5中的表“CDR_MODEL”)、处于表级的一个或多个应用表定义(例如,图5中的表“CDR_TABLE”、“CDR_TABLE_CONS”和“CDR_TABLES_CONS_COLS”)以及处于列级的一个或多个应用表定义(例如,图5中的表“CDR_COLUMN”和“CDR_TABLES_CONS_COLS”)。另外,物理模型包括关联对象组件560,该组件包括用于一个或多个关联对象的表定义。在一种实施例中,关联对象组件560可以包括用于一个或多个差异的一个或多个表定义(例如,图5中的表“DME_DISCREPANCY”)和用于一个或多个标志的一个或多个表定义(例如,图5中的表“DME_FLAG”)。此外,物理模型包括研究组件570,该组件是把物理模型中的所有其它对象分组到一起的父对象。

[0071] 数据世系映射组件510、影子表组件520和变换组件530连同表定义组件550一起提供数据世系子系统,这在以下联系图8更详细地进一步描述。另外,关联对象映射组件540连同表定义组件550和关联对象组件560一起提供关联信息子系统,这在以下联系图10更详细地进一步描述。

[0072] 图6根据本发明实施例说明了数据世系映射表600的框图。数据世系映射表600包括用于从目标数据元素追踪源数据元素的信息。用于从目标数据元素追踪源数据元素的信息是数据世系追踪信息(即,元数据信息),这与数据仓库的一个或多个数据表中所存储的数据不同。数据世系映射表600可以在数据模型发布阶段填充,并且数据世系映射表600包

括对每种数据变换用于在源数据元素和目标数据元素之间追踪的信息。数据世系映射表600通过对象标识存储对变换列级映射对象、目标表和列、源表和列、保持源代理键的辅助列以及特定变换所需的任何附加信息的引用。

[0073] 根据实施例,存储在数据世系映射表600中的信息组成用来追踪数据世系的数据世系追踪信息的第一主要部分。用来追踪数据世系的数据世系追踪信息的第二主要部分是为了追踪回源数据记录而添加到变换的目标表的表定义的一个或多个辅助列。辅助列可以包含来自对目标表中的数据记录做贡献的每个源表的代理键。当必要时,辅助列可以保持源表的表标识和/或识别用于特定变换的“过滤”值。数据世系映射表600和一个或多个辅助列包括由数据世系传播模块(诸如图2的数据世系传播模块200)用来从目标数据元素追踪回源数据元素的数据。用来追踪数据世系的数据世系追踪信息的第三主要部分是影子表,这在以下联系图7更详细地进一步描述。

[0074] 在图6所说明的实施例中,数据世系映射表600(包括列映射数据)包括对应于变换表级映射标识的“TMAP_ID”列(即,识别变换)。数据世系映射表600还包括定义源列如何与目标列相关的“TMAP_TYPE_RC”列(即,识别变换类型)。数据世系映射表600还包括对应于源数据模型标识的“SRC_DM_OBJ_ID”列、对应于源表标识的“SRC_TAB_OBJ_ID”列以及对应于源列标识的“SRC_COL_OBJ_ID”列。数据世系映射表600还包括对应于目标数据模型标识的“TRG_DM_OBJ_ID”列、对应于目标表标识的“TRG_TAB_OBJ_ID”列以及对应于目标列标识的“TRG_COL_OBJ_ID”列。数据世系映射表600还包括对应于用于目标表中辅助列的对象标识的“SRC_SKEY_COL_OBJ_ID”列,该列存储用于由存储在数据记录的“SRC_TAB_OBJ_ID”列中的源表标识识别出的源表的源表数据记录代理键。根据所说明的实施例,数据世系映射表600还包括可选地对某些变换类型在源表或目标表中提供关于过滤列标识的信息的“FILTER_COL_OBJ_ID”列以及可选地对某些变换类型提供关于表过滤列中过滤值的信息的“FILTER_VALUE”列。

[0075] 图7根据本发明实施例说明了影子表700的框图。影子表700是优化的中心索引(central index),它可以在从源数据元素向目标数据元素追踪时提供数据世系追踪信息的高效查找。影子表700可以允许数据世系处理在单个表中直接解析源记录-目标记录关系,而不必为数据仓库中的每个数据表创建单独的查询。这种数据世系处理可以更高效几个数量级。另外,影子表700可以从数据仓库的一个或多个数据表提取源记录-目标记录关系,并且可以充当用于遍历一个或多个数据表的世系的索引,而无需查询数据表本身。

[0076] 在图7所说明的实施例中,影子表700包括对应于目标表实例标识的“TRG_TAB_OBJ_ID”列。影子表700还包括对应于源表实例标识的“SRC_TAB_OBJ_ID”列。影子表700还包括对应于用于目标表中辅助列的对象标识的“SRC_SKEY_COL_OBJ_ID”列,该列存储源表数据记录代理键。影子表700还包括对应于目标表数据记录的代理键值的“TRG_SKEY_VALUE”和对应于源表数据记录的代理键值的“SRC_SKEY_VALUE”,后者也是用于前面提到的目标表数据记录的辅助列的值。根据所说明的实施例,影子表700还包括可选地对某些变换类型提供关于表过滤列中过滤值的信息的“FILTER_VALUE”列。

[0077] 图8根据本发明实施例说明了数据世系子系统的物理模型的框图。在图8所说明的实施例中,物理模型包括图6的数据世系映射表600和图7的影子表700。该物理模型还包括变换类型表800,该表800包括代表一种或多种变换类型的数据。此外,物理模型包括表定义

组件810,该组件包括用于一个或多个源表和一个或多个目标表的表定义,其中源表和目标表是数据表的例子。在所说明的实施例中,表定义组件810可以包括处于数据模型级的一个或多个表定义(例如,图8中的“CDR_MODEL”表)、处于表级的一个或多个表定义(例如,图8中的“CDR_TABLE”、“CDR_TABLE_CONS”和“CDR_TABLES_CONS_COLS”表)以及处于列级的一个或多个表定义(例如,图8中的“CDR_COLUMN”和“CDR_TABLES_CONS_COLS”表)。

[0078] 根据该实施例,当数据在数据仓库中生成或加载到其中时,以及当一个或多个变换已经为数据仓库定义时,数据世系传播模块(诸如图2的数据世系传播模块200)可以在数据世系子系统的物理模型中生成数据世系追踪信息。更具体而言,对于一种或多种变换当中的每种变换,数据世系传播模块可以在数据世系映射表600中存储每个目标到源映射。数据世系传播模块还可以把一个或多个目标表的一个或多个目标数据值映射到一个或多个源表的一个或多个源数据值。数据世系传播模块还可以通过修改表定义组件810的一个或多个表定义把一个或多个辅助列添加到一个或多个目标表,其中每个辅助列可以存储识别源表的源代理键。另外,对于一种或多种变换当中的每种变换,数据世系传播模块可以在影子表700中存储每个源到目标映射。数据世系传播模块还可以把一个或多个源表的一个或多个源数据值映射到一个或多个目标表的一个或多个目标数据值。

[0079] 关于存储在数据世系映射表600中的一个或多个目标到源映射,数据世系传播模块可以在数据世系映射表600中存储用于每个目标列数据值的以下信息:(a)源表实例对象标识和源列对象标识;(b)目标表实例对象标识和目标列对象标识;(c)已经添加到目标表的每个辅助列的对象标识,其中每个辅助列存储用于被用来产生目标数据值的源数据记录的源代理键;及(d)映射所需的任何附加数据,诸如映射是优选源路线的指示符(例如,标志)。如果相同的源数据记录被用来填充多于一个目标列中的值,则存储源代理键列标识的相同辅助列可以在数据世系映射表600中的多个映射系统记录中重复。

[0080] 数据世系追踪信息可以提供对数据仓库变换的支持,其中示例变换类型包括直接映射变换、转换变换、导出变换、联合变换、垂直到水平旋转变换、水平至垂直旋转变换、判例变换,以及定制类型的变换。在某些实施例中,数据世系追踪信息可以利用数据世系映射技术产生。数据世系映射技术涉及把一个或多个源列映射到目标列,以及添加一个或多个辅助列以保持一个或多个源代理键,其中每个源代理键对应于对目标列的值做贡献的源表。

[0081] 直接映射变换、转换变换和导出变换可以利用数据世系映射技术在数据世系映射表600中产生直接的条目:一个或多个源列映射到每个目标列。这些变换常常从单个源表映射到单个目标表,因此常常只有单个存储源代理键值的辅助列。

[0082] 数据世系追踪信息还可以被用来利用数据世系映射技术支持联合变换和判例变换。但是,数据世系映射技术会对每种变换需要多于一个辅助列,以确保正确地追踪到源数据元素。例如,在五个源表联合成一个目标表时,目标表中的任何数据记录可以来自五个源表当中任何一个。为了避免识别特定数据记录起点的歧义,目标表可以扩展成包括五个辅助列(每个源表一列)。当执行联合变换时,源数据记录代理键插入映射到正确源表的辅助列。因此,对于目标表中的每个数据记录,五个辅助列中只有一个将包含源代理键值,而其它四个辅助列将保持空。

[0083] 相同的场景可以应用到判例变换。在判例变换中,来自目标表的一列中的不同数

据记录的值可以从可来自不同源表的不同源列填充。例如,在第一个数据记录中,目标判例列可以从源Tab1.ColX填充。在第二个数据记录中,目标判例列可以从源Tab1.ColY填充。在第三个数据记录中,目标判例列可以从源Tab2.ColX填充。每个指定由判例变换的条件控制。因此,可能被用来填充目标列的每个源表与列的组合必须被识别,以便在源和目标数据元素之间追踪。将为每种可能的源表和列组合添加一个辅助列。就像关于联合变换,依赖于对那个数据记录哪个判例条件为真,只有一个辅助列将利用每个数据记录中的值填充。另外,将存在到目标列的多个列数据映射,其中每个映射源记录包括不同的源键列标识。

[0084] 数据世系映射技术可以被修改,以支持垂直到水平旋转变换和水平至垂直旋转变换。修改可以涉及在数据世系映射表600中存储附加信息,例如,关于用于变换的旋转列的信息以及过滤值的列表。例如,垂直到水平旋转变换可以取多个数据记录作为输入,每个记录在“旋转”列中具有不同的“过滤”值,并且可以在目标表中创建多个列,每个“过滤”值有一个或多个列。结果可以是目标表,其中多个源数据记录对单个数据记录做贡献。因此,多个辅助列可以添加到目标表,对目标数据记录做贡献的每个源数据记录有一个或多个列。

[0085] 在某些实施例中,先进的技术可以被用来存储数据世系映射数据。在一种先进的技术中,对于联合变换,用于存储数据世系映射数据的先进技术可以涉及在辅助列中所存储的源代理键中包括源表实例对象标识。从目标数据记录追踪回源数据记录的算法可以使用存储在目标数据记录的辅助列中的源代理键值来查找指定的源表中的源数据记录,其中源表可以从源代理键值提取。对于联合变换和判例变换二者,用于存储数据世系追踪数据的另一种先进技术可以涉及包括存储“过滤值”的“过滤”辅助列和存储源代理键的另一辅助列。对于联合变换,存储“过滤值”的“过滤”辅助列存储源表实例对象标识。因此,从目标数据记录追踪回源数据记录的算法可以使用存储在目标数据记录的辅助列中的源代理键值来查找源表中的源数据记录,其中源表在“过滤”辅助列中识别。对于判例变换,用户可以提供作为文本串的“过滤”值,以识别可能填充判例目标列的每种源表和列组合。这些“过滤”值连同适当的源表/列组合以及判例变换目标表/列组合一起存储在数据世系映射表600中。在目标表的每个数据记录中,对应的“过滤值”可以输入到“过滤”辅助列。因此,从目标数据记录追踪回源数据记录的算法可以使用来自目标数据记录的源代理键值,连同“过滤值”和存储在数据世系映射表600中的信息。

[0086] 为了避免对多个辅助列的需求,其它先进技术可以被用来存储数据世系追踪信息。一种先进技术涉及“智能代理键”,其中“智能代理键”是一个或多个源代理键值的串联,以源表标识为前缀。因此,智能代理键可以被用来获得源表标识,并且因此,不需要存储源表标识的单独辅助列。另外,将只需要单个辅助列来存储智能代理键,由此在联合变换、垂直到水平旋转变换和水平至垂直旋转变换的情况下除去了对多个辅助列的需求。另一种先进技术涉及对于垂直到水平旋转变换和水平至垂直旋转变换在智能代理键中包括旋转列值作为“过滤值”。

[0087] 根据实施例,以上提到的数据世系追踪信息,或者其它数据世系追踪信息,可以被充分利用,以支持链接到数据元素或数据记录的关联对象的可见性,并且把关联对象传播到相关的源数据元素和目标数据元素。另外,以上提到的数据世系追踪信息,或者其它数据世系追踪信息,还可以被充分利用,以支持关联对象通过数据世系中的变换路径导出到数据的外部源。根据实施例,关联对象映射表可以存储关于关联对象与数据元素或数据记录

之间的一个或多个链接的数据。数据世系传播模块(诸如图2的数据世系传播模块)可以识别与关联对象的关联以及关联对象作为主链接被创建的数据元素(或数据记录),并且还可以在关联对象映射表中存储该主链接。数据世系传播模块还可以创建并存储到主数据元素(或数据记录)的源和目标数据元素/数据记录的次要链接。当系统的用户查看目标数据的源时以及当用户查看主数据元素(或数据记录)时,这可以提供关联对象的高效可视化。因此,根据该实施例,每个关联对象可以具有一个主链接,并且可以具有一个或多个次要链接。

[0088] 图9根据本发明实施例说明了关联对象映射表900的框图。关联对象映射表900可以存储关于关联对象与数据元素或数据记录之间的一个或多个链接的数据。关联对象映射表900可以存储关于关联对象与数据元素(或数据记录)之间的链接的数据,其中该主链接识别关联对象是为所述数据元素(或数据记录)创建的(即,最初与之关联)。另外,关联对象映射表900可以存储关于关联对象与数据元素(或数据记录)之间的次要链接的数据,其中次要链接识别数据元素(或数据记录)是与关联对象相关联的主数据元素(或数据记录)的源或者目标数据元素(或数据记录)。关于关联对象与数据元素之间的链接的数据,以及关于关联对象与一个或多个数据元素之间的一个或多个次要链接的数据,可以被用来获得上游或下游数据元素与该关联对象的关联。另外,关联不必遍历数据世系以检测关联对象的存在就可以获得。

[0089] 在图9所说明的实施例中,关联对象映射表900包括对应于关联对象标识的“OPOBJ_ID”列(即,识别关联对象)。关联对象映射表900还包括对应于关联对象类型标识的“OPOBJ_TYPE_ID”列(即,识别关联对象的类型)。在备选实施例中,关联对象的标识可以是包括标识值以及至少一个值的复杂标识。

[0090] 根据所说明的实施例,关联对象映射表900还包括对应于数据元素模型标识的“DM_OBJ_ID”列(即,识别关联对象与其关联的数据元素的数据表的数据模型)。关联对象映射表900还包括对应于数据元素表标识的“TAB_OBJ_ID”列(即,识别关联对象与其关联的数据元素的数据表)。关联对象映射表900还包括对应于数据元素列标识的“COL_OBJ_ID”列(即,识别关联对象与其关联的数据元素的列)。关联对象映射表900还包括对应于数据元素数据记录标识的“SKEY_VALUE”列(即,识别关联对象与其关联的数据元素的数据记录)。数据元素数据记录标识也是代理键。

[0091] 根据所说明的实施例,关联对象映射表900还包括识别关联对象与数据元素之间的关联是主链接还是次要链接的“IS_PRIMARY_RC”列。关联对象映射表900还包括识别关联对象是与数据元素还是与数据记录关联的“IS_RECORD_LEVEL_RC”列。关联对象映射表900还包括识别数据元素是否是用于关联对象的优选源数据元素的“PREFERRED_ROUTE_RC”列,并且其值是从数据世系映射表600继承的。关联对象映射表900还包括识别关联对象与数据元素之间的关联是主链接、上游次要链接还是下游主链接的“TRACE_TO_RC”列。

[0092] 图10根据本发明实施例说明了关联信息子系统的物理模型的框图。在图10所说明的实施例中,物理模型包括图9关联对象映射表900。该物理模型还包括关联对象组件1000,该组件包括代表一个或多个关联对象的数据。在一种实施例中,关联对象组件1000可以包括表,该表包括代表一个或多个关联对象的类型的数据(例如,图10中的“DME_OPERATIONAL_OBJECT_TYPE”表)。根据该实施例,关联对象组件1000还可以包括一个或多个

表,这些表包括代表指定的关联对象类型的一个或多个关联对象的数据。在所说明的实施例中,“CDR_DISCREPANCY”表包括代表是差异的一个或多个关联对象的数据,“CDR_FLAG”表包括代表是状况标志的一个或多个关联对象的数据,并且“CDR_COMMENT”表包括代表是注释的一个或多个关联对象的数据。此外,物理模型包括表定义组件1010,该组件包括用于一个或多个源表和一个或多个目标表的表定义。在所说明的实施例中,表定义组件1010可以包括处于数据模型级的一个或多个表定义(例如,图10中的“CDR_MODEL”表)、处于表级的一个或多个表定义(例如,图10中的“CDR_TABLE”表)以及处于列级的一个或多个表定义(例如,图10中的“CDR_COLUMN”表)。

[0093] 根据实施例,系统的用户,或者在系统中被执行的自动程序,可以选择数据元素,并且创建与被选数据元素相关的关联对象,诸如差异。该系统可以调用数据世系传播模块(诸如图2的数据世系传播模块),并且数据世系传播模块可以创建新的关联对象与被选数据元素之间的主关联(称为主链接)。数据世系传播模块还可以在关联对象映射表900中存储主关联(即,主链接)。通过追踪从被选数据元素到一个或多个源数据元素和/或目标数据元素的完全数据世系路径,数据世系传播模块还可以在新的关联对象与一个或多个源数据元素和/或目标数据元素之间创建一个或多个次要关联。数据世系传播模块还可以在关联对象映射表900中存储一个或多个次要关联(即,一个或多个次要链接)。在某些场景下,当向上游追踪到源数据元素时,多个源数据元素可以被映射到单个目标数据元素。作为用于关联对象的优选源数据元素的源数据元素可以这样在关联对象映射表900中被识别。优选源数据元素可以被识别为优选路径,并且优选路径数据世系追踪信息可以存储在数据世系映射表(诸如图6和8的数据世系映射表600)中。

[0094] 根据实施例,该系统可以利用关联对象映射表900中所存储的主链接在数据表的数据模型的显示中显示关联对象的存在,所述关联对象在所述数据表中创建(或者以别方式关联)。该系统还可以利用关联对象映射表900中所存储的一个或多个次要链接在上游和/或下游数据模型中显示关联对象。当对于特定的目标数据元素存在多个源数据元素时,系统可以将一个或多个优选源数据元素标记为优选路径。当系统显示数据模型时,系统还可以显示与该数据模型的数据表的每个数据元素关联的一个或多个关联对象,不管关联对象与数据元素是通过主链接还是次要链接关联。

[0095] 在一种实施例中,当检测到包含具有到关联对象的主链接的数据元素的数据记录时,关联对象的状态可以被修改为关闭状态。用于该关联对象的数据可以从关联对象映射表900中去除。作为替代,用于关联对象的数据可以留在关联对象映射表900中,但是被修改成指示数据已经“实质上被删除”(诸如把“删除”标志或时间戳与数据关联)。为了审计目的,关联对象可以被保留,而不是删除,并且保存。如果经主链接与关联对象关联的数据元素被删除,则到该关联对象的主链接和所有次要链接都可以被删除,并且该关联对象的状态可以被修改成关闭状态。如果经次要链接与关联对象关联的数据元素被删除,则该次要链接可以被删除。

[0096] 根据某些实施例,现在更详细地描述追踪数据世系以及传播关联信息(即,关联对象)的可见性的附加方面。在某些实施例中,系统可以支持用户或系统通过源侧数据世系中的源数据字段/记录从任何数据记录上的任何数据字段向后导航,直到数据世系到达在外部系统(即,源系统)中起源的源数据记录/字段。然后,系统可以允许用户选择对其动作的

源数据记录(或数据字段)。动作的例子是向源系统发送关于数据记录(或数据字段)值的查询或消息、发送文档、调用可以直接或间接对源系统采取动作的API,以及利用上下文中在源系统中源数据记录/字段或者其表现打开源系统的应用窗口。

[0097] 对以上功能的扩展涉及定义从每个表(或列)到紧上的源表(或列)的“优选源路径”的变换定义元数据。这个优选路径使系统能够在源自外部系统的数据记录(或数据字段)选择源世系的单个分支,而无需用户明确地导航到源数据世系。然后,该系统能够对外部系统中的那个数据记录(或数据字段)采取上述动作。

[0098] 在其它某些实施例中,当在用户接口中显示数据记录及其数据字段时,诸如包括所显示字段的表格显示或者格式化形式,关联信息(即,一个或多个关联对象)的优化索引,诸如关联对象映射表,可以被用来为与该关联信息相关联的数据记录和/或数据字段提供关联信息存在的可视指示,诸如加亮或图标。

[0099] 该系统还可以允许用户通过所显示的信息的某个属性把那种信息限定到关联信息的子集(即,一个或多个关联对象)。例子包括关联对象的类型(诸如差异、错误报告、文档、注释)和关联对象的特定类型的子类型(诸如文档的类型、错误报告的状况,或者标签或状况标志的分类)。这个功能可以通过在把数据记录(或数据字段)与到所有其它相关数据记录(或字段)的关联对象关联的关联对象表中包括关于关联对象的类型或子类型的信息来优化,使得系统不必为了执行限定而查看超过该表本身。

[0100] 该系统还可以允许用户基于与关联对象关联的数据记录(或数据字段)到当前显示的关联对象的关系来限定所显示的关联的信息(即,关联对象)。例子包括只显示在数据世系中与所显示数据元素上游的数据记录(或数据字段)关联的关联对象、只显示与所显示数据元素下游的数据记录(或数据字段)关联的关联对象、只显示(经主链接)与所显示数据元素直接关联的关联对象,或者只显示来自特定源或目标表或者特定源表或目标表集合的关联对象。这可以通过在把与该关联对象关联的数据记录(或数据字段)关联到所有其它相关数据记录(或数据字段)的关联对象表中包括关于每个关联对象的相对位置的信息来优化,使得系统不必为了执行限定而查看超过该表本身。

[0101] 根据某些实施例,为了显示关联信息(即,关联对象)的可视指示,会有必要把每个关联对象高效地关联到特定显示的数据字段(或数据记录)。在一种实施例中,系统可以通过使用结构化查询语言(“SQL”)的能力首先对用于关联的数据记录的标识旋转使得每个关联对象处于表格式而执行关联,其中每列的名字或标识具有与所显示的信息相同的表结构,并且其中被旋转的表中的数据元素只有在存在关联对象时才具有数据值。然后,系统可以对所显示的实际数据与被旋转的表之间的给定数据记录标识使用名字或列标识匹配,以便高效地开启对关联对象的存在指示。在另一种实施例中,系统可以通过在具有高效访问方法的本地数据结构中单独检索关联对象来执行关联,以便检索关于特定数据记录标识或特定列标识的信息。然后,该系统可以通过利用数据记录标识或列标识访问这个本地数据结构来高效地识别所显示的数据世系中数据记录和数据字段。在一种实施例中,对于数据元素的数据世系,系统还可以为数据世系中的每个数据记录(或数据字段)显示一个或多个关联对象的存在。

[0102] 在某些实施例中,当系统的用户接口显示一个或多个关联对象存在的指示时,系统允许用户调用对特定数据记录(或数据字段)的动作,利用例如鼠标点击、鼠标右击和菜

单选择或者光标导航到目标数据记录(或数据字段)以及键盘动作。被调用的动作可以包括显示关联对象的列表和到关于关联对象的细节的导航。例如,系统可以显示所关联的注释的弹出表,具有该注释附连到的每个数据表(或数据字段)的指示。另一个例子可以是显示所关联的差异的列表,该列表可以被用来选择和显示关于差异的详细信息。另外,代替地,系统可以把用户放在接口中,而不管用户是否可以对关联对象采取动作,诸如对注释作出响应或者改变差异的状况。

[0103] 在某些实施例中,系统可以使用关于访问不同类型关联对象所需的用户权利和特权或者访问附连到特定上游或下游数据表的关联对象所需的特权的安全信息,以限定用户接口中关联对象的可见性、关于用户可用的关联对象的细节程度或者用户对关联对象采取动作的能力。

[0104] 根据某些实施例,系统可以使用与上游或下游数据记录及与它们关联的关联对象(诸如差异或注释)相关联的追踪时间戳信息,来使得用户能够理解在用户当前查看的数据之前或之后的数据是否与上游或下游数据一致。例如,如果用户在查看数据元素并调用示出在当前数据元素的数据世系中附连到数据的关联的注释的功能,则系统可以指示源数据世系中的源数据和关联的注释是否比该数据元素更近并且因此可以反映还没有传播到当前数据表的变化。类似地,在目标数据世系中,用户可以看到数据和关联的注释是否比当前数据更旧并且因此可以反映还没有被更新以反映已经对当前数据元素所作的变化的数据。相同能力的另一种表现是当用户启动了对与当前数据关联的优选源数据元素提出查询的动作并且系统检测到并提醒用户源数据元素已经比当前数据元素更近地被修改时。这种功能可以防止用户通过世系设施(facility)对可能已经被修改的数据采取动作。

[0105] 如前所述,系统可以通过代理键的使用最小化数据存储并且可以通过影子表的使用优化查找性能,其中,为了优化的查找,影子表把源代理键提取到单个表中。根据一种实施例,对这种设计的备选方案不需要生成和使用代理键。在这种实施例中,源表的完全主键存储在一个或多个目标表中,作为源键,并且在目标表上添加附加的数据字段,以保持该代理键中所捕捉的任何辅助信息,诸如与源主键字段关联的旋转值或表选择符。在这种实施例中,系统元数据描述每个这种源键并且利用代理键的任何功能利用组成该源键的数据字段的全集操作。另外,在不必查询每个个别表的情况下把代理键提取到影子表中用于优化的查询性能的优化可以通过使用单个影子表来实现,这单个影子表具有为每个源或目标键保留的多个列,这些列在需要的时候用来保持用于每个源键的成分的值。然后,通过对由特定表使用的列的适当连接以保持全部源键信息,关于用于每个表的键结构的系统元数据将使得系统对照这单个影子表构造查询。关于源键集合的适当串联索引和关于目标键集合的另一索引将利用单个影子表优化性能。

[0106] 在某些实施例中,系统可以充当使用一个或多个应用的元数据描述来生成数据世系的运行时元素的应用发生器。例如,可以存在用户定义一个或多个源表和目标表之间的映射的用户接口。该映射可以依赖于变换的类型(例如,连接、旋转或者逆旋转)而不同。定义过程可以记录一个或多个源数据元素和一个或多个目标数据元素之间列级的映射。然后,作为部署变换的一部分,系统可以:(a)在目标表安装在数据库中之前把一个或多个辅助列添加到目标表,其中这一个或多个辅助列存储一个或多个源键;(b)记录关于一个或多个源数据元素和目标数据元素之间的关系的元数据映射信息;及(c)生成为了执行变换而

可以被执行的计算机代码。计算机代码可以包括填充目标表中的一个或多个源键的逻辑,以便记录一个或多个源数据元素和目标数据元素之间的关系。根据某些实施例,无需用户干预,所生成的计算机代码可以在完全相同,或者相似,的结构的新应用部署中重用。

[0107] 图11根据本发明实施例说明了联合变换的示例数据世系追踪。在示例实施例中,源表“LAB1”、“LAB2”和“LAB3”(即,源表1110、1120和1130)经联合变换组合到目标表“LAB_SUM”(即,目标表1140)中。另外,在该示例实施例中,数据世系传播模块(诸如图2的数据世系传播模块200)可以为目标表1140中的以下目标数据元素找出源数据元素:列“RBC”和数据记录标识(即,代理键)“[LAB_SUM]~103~1”。数据世系传播模块可以查找数据世系映射表1100并且在三个系统记录:系统记录9、25和35中找到目标表“LAB_SUM”(即,目标表1140)作为目标表,并且列“RBC”作为目标列。这三个系统记录具有“UNION”的变换映射类型(即,列“TMAP_TYPE”),并且它们在列“LAB_SUM.UNION_SRC_KEY”中存储它们的源代理键,在数据世系映射表1100的“SRC_SKEY_COLNAME”列中示出。在目标表1140中,数据世系传播模块可以在列“UNION_SRC_KEY”中查找目标数据记录。这个列包含值“[LAB1]~103~1~LAB1”,并且因为这是智能代理键,所以这个值中的第一个成分(即,“[LAB1]”)是源表标识。因而,数据世系传播模块可以确定源表是源表“LAB1”(即,源表1110)。然后,数据世系传播模块可以返回到在数据世系映射表1100中找到的三个系统记录并且,通过知道源表是源表1110,数据世系传播模块可以在数据世系映射表1100中识别出正确的系统记录:系统记录9。因此,数据世系传播模块可以识别出源数据元素是在源表1110中在具有代理键值“[LAB1]~103~1~LAB1”的列“RBC”中。数据世系传播模块还可以识别出源数据值是“5.3”。

[0108] 另外,在该示例实施例中,数据世系传播模块可以为源表1110中的以下源数据元素找出目标数据元素:列“RBC”和数据记录标识(即,代理键)“[LAB1]~103~1~LAB1”。数据世系传播模块可以在数据世系映射表1100中查看并且在系统记录9中找到源表“LAB1”(即,源表1110)作为源表,并且找出列“RBC”作为源列。根据这个信息,数据世系传播模块可以确定它需要查看目标表1140的“RBC”列中的目标数据元素。然后,利用目标表“LAB_SUM”、源表“LAB1”、SRC_SKEY_COLNAME“UNION_SRC_SKEY”和源表代理键“[LAB1]~103~1~LAB1”的组合,数据世系传播模块可以搜索影子表1150,以确定目标表1140中期望数据记录的代理键是“[LAB_SUM]~103~1”。因此,数据世系传播模块可以识别出目标数据元素是在目标表1140中具有代理键“[LAB_SUM]~103~1”的数据记录的列“RBC”中。数据世系传播模块还可以识别出目标数据值是“5.3”。

[0109] 图12根据本发明实施例说明了旋转变换的示例源数据世系追踪。这是从垂直源表“VS”(即,源表1210)到水平目标表“VITALS”(即,目标表1220)的变换。源表1210具有识别重要符号的旋转列,列“VSTSTCD”,并且重要符号值存储在列“VSORES”中。在“VSTSTCD”列中存在五个不同的重要符号值:“SYSBP”、“DIASBP”、“PULSE”、“HEIGHT”和“WEIGHT”。它们作为过滤值(即,列“FILTER_VALUE”中的值)存储在数据世系映射表1200中,每个都在目标表1220中具有其对应的列。另外,在源表1210中,每个智能代理键值(即,列“SK”中的值)包含来自其“VSTSTCD”列的过滤值作为该智能代理键值的最后一个成分。目标表1220对于所收集的五个重要符号当中每一个包括一列:“SYSBP”、“DIASBP”、“PULSE”、“HT”和“WT”。重要符号值在用于每个主体的适当列中存储。目标表1220中的每个数据记录消费源表1210中的五个数据记录;来自源表1210中的源列“VSTSTCD”的过滤值(即,“SYSBP”、“DIASBP”、“PULSE”、

“HEIGHT”和“WEIGHT”)变换成目标表1220中的列(即,列“SYSBP”、“DIASBP”、“PULSE”、“HT”和“WT”)。

[0110] 在该示例实施例中,数据世系传播模块(诸如图2的数据世系传播模块200)可以为目标表1220中的以下目标数据元素找出源数据元素:列“PULSE”和数据记录标识(即,代理键)“[VITALS]~101~2”。数据世系传播模块可以查看数据世系映射表1200并且为目标表“VITALS”(即,目标表1220)和目标列“PULSE”只找出一个系统记录,系统记录3(在图12中称为过滤后的数据世系映射表1230)。这个系统记录对于具有旋转列“VSTSTCD”和源数据项列“VSORES”的TALL_SKINNYSOURCE表“VS”具有变换映射类型“PIVOT”。相同的系统记录,系统记录3(在图12中称为过滤后的数据世系映射表1230),定义目标记录辅助列,以得到源记录代理键值,作为“VS_SK1”。根据该算法,来自这个列的值(“[VS]~101~2”)可以由数据世系传播模块(诸如图2的数据世系传播模块200)与过滤值“PULSE”组合(即,“[VS]~101~2~PULSE”),以找出源表中的正确记录。因此,数据世系传播模块还可以识别出源数据值为“68”。

[0111] 图13根据本发明实施例说明了旋转变换的示例目标数据世系追踪。如前所述,这是从垂直源表“VS”(即,表12的源表1210)到水平目标表“VITALS”(即,图12的目标表1220)的变换。在该示例实施例中,数据世系传播模块(诸如图2的数据世系传播模块200)可以为源表1210中的以下源数据元素找出目标数据元素:列“VSORES”和数据记录标识(即,代理键)“[VS]~101~2~PULSE”。数据世系传播模块可以查看图12的数据世系映射表1200并且在五个系统记录:系统记录1、2、3、4和5中找出作为源列的列“VSORES”。然后,数据世系传播模块可以利用列“VSTSTCD”中的源数据记录数据值(即,“PULSE”)过滤数据世系映射表1200中的系统记录,以确定数据世系映射表1200中的正确系统记录。因此,数据世系传播模块可以把系统记录3识别为数据世系映射表1200中的正确系统记录(在图13中称为过滤后的数据世系映射表1300)。根据这个信息,数据世系传播模块可以确定它需要在目标表1220的“PULSE”列中查找目标数据元素。根据用于旋转变换的算法,数据世系传播模块可以从源代理键值除去过滤值,然后可以利用代理键“[VS~101~2]”搜索影子表1310,并且确定目标表1220中期望数据记录的代理键是“[VITALS]~101~2”。因此,数据世系传播模块可以识别出目标数据元素在目标表1220中具有代理键值“[VITALS]~101~2”的列“PULSE”中。数据世系传播模块还可以识别出目标数据值为“68”。

[0112] 图14根据本发明实施例说明了逆旋转变换的示例源数据世系追踪。这是从水平源表“VITALS”(即,源表1410)到垂直目标表“VS”(即,目标表1420)的变换。源表1410对所收集的五个重要信号当中每一个具有一列:“SYSBP”、“DIASBP”、“PULSE”、“HT”和“WT”。重要符号值存储在用于每个主体的适当列中。目标表1420具有识别重要符号的旋转列,列“VSTSTCD”,并且重要符号值存储在列“VSORES”中。在“VSTSTCD”列中存在五个不同的重要符号值:“SYSBP”、“DIASBP”、“PULSE”、“HEIGHT”和“WEIGHT”。这些作为过滤值(即,列“FILTER_VALUE”中的值)存储在数据世系映射表1400中,每个都在源表1410中具有其对应的列。此外,在数据世系映射表1400中,目标表“VS”标记为“tall skinny”表。在目标表1420中,每个智能代理键值(即,列“SK”的值)包含来自其“VSTSTCD”列的过滤值作为该智能代理键值的最后一个成分。源表1410中的每个数据记录被变换成目标表1420中的五个数据记录。源表1410中的五个重要符号列当中每一个(即,列“SYSBP”、“DIASBP”、“PULSE”、“HT”和

“WT”)变成目标表1420的列“VSTSTCD”中的过滤值(即,“SYSBP”、“DIASBP”、“PULSE”、“HEIGHT”和“WEIGHT”)。

[0113] 在该示例实施例中,数据世系传播模块(诸如图2的数据世系传播模块200)可以为目标表1420中的以下目标数据元素找出源数据元素:列“VSORES”和数据记录标识(即,代理键)“[VS]~101~2~PULSE”。数据世系传播模块可以查看数据世系映射表1400并且在五个系统记录:系统记录1、2、3、4和5中找到目标表“VS”(即,目标表1420)作为目标表,并且找出列“VSORES”作为目标列。目标列“VSTSTCD”被标记为目标表“VS”的旋转列。然后,数据世系传播模块可以利用列“VSTSTCD”中的目标数据记录数据值(即,“PULSE”)过滤数据世系映射表1400中的系统记录,以确定数据世系映射表1400中正确的系统记录。因此,数据世系传播模块可以识别出系统记录3作为数据世系映射表1400中正确的系统记录(在图14中称为过滤后的数据世系映射表1430)。这个系统记录具有变换映射类型(即,列“TMAP_TYPE”)“UNPIVOT”、旋转列(即,列“PIVOT_COL”)“VSTSTCD”、源表(即,列“SRC_TAB”)“VITALS”以及源列(即,列“SRC_COL”)“PULSE”。根据目标表1420的代理键值(即,列“VITALS_SK1”),数据世系传播模块可以确定源数据记录的代理键是“[VITALS]~101~2”。因此,数据世系传播模块可以识别出源数据元素是在源表1410中具有代理键值“[VITALS]~101~2”的列“PULSE”中。数据世系传播模块还可以识别出源数据值为“68”。

[0114] 图15根据本发明实施例说明了逆旋转变换的示例目标数据世系追踪。如前所述,这是从水平源表“VITALS”(即,图14的源表1410)到垂直目标表“VS”(即,图14的目标表1420)的变换。在该示例实施例中,数据世系传播模块(诸如图2的数据世系传播模块200)可以为源表1410中的以下源数据元素找出目标数据元素:列“PULSE”和数据记录标识(即,代理键)“[VITALS]~101~2”。数据世系传播模块可以查看图14的数据世系映射表1400并且在系统记录3中找到源表“VITALS”(即,源表1410)作为源表,并且找到列“PULSE”作为源列(在图15中称为过滤后的数据世系映射表1500)。这个系统记录具有变换映射类型(即,列“TMAP_TYPE”)“UNPIVOT”、旋转列(即,列“PIVOT_COL”)“VSTSTCD”、目标表(即,列“TRG_TAB”)“VS”以及目标列(即,列“TRG_COL”)“VSORES”。根据这个信息,数据世系传播模块可以确定它需要在目标表1420的“VSORES”中查找目标数据元素。然后,数据世系传播模块可以利用源表1410的代理键“[VITALS]~101~2”搜索影子表1510,并且确定目标表1420中期望的数据记录的代理键是“[VS]~101~2~PULSE”。为了减少影子表中记录的数目,数据世系传播模块(诸如图2的数据世系传播模块200)为逆旋转变换在列“TRG_SKEY_VALUE”中存储缩略值,即,没有过滤值。然后,在利用其作为目标记录标识符之前,数据世系传播模块把过滤值附加到从影子表检索出的列“TRG_SKEY_VALUE”的值。因此,数据世系传播模块可以识别出目标数据元素是在目标表1420中代理键值为“[VS]~101~2~PULSE”的列“VSORES”中。数据世系传播模块还可以识别出目标数据值为“68”。

[0115] 图16根据本发明的实施例说明了连接变换的示例数据世系追踪。在该示例实施例中,源表“DM”和“LAB1”(即,源表1610和1620)连接,以形成目标表“LAB1_SUM”(即,目标表1630)。数据世系映射表1600为目标表1630的每一列包括一个系统记录。每个系统记录包括源表标识(即,列“SRC_TAB”)、源列标识(即,列“SRC_COL”)和源代理键列标识(即,列“SRC_SKEY_COLNAME”)。源代理键列是添加到目标表1630的辅助列,以存储用于对目标表1630中的数据记录做贡献的源表1610(即,列“DM_SKEY”)或源表1620(即,列“LAB1_SKEY”)中的数

据记录的源代理键。对于连接变换,存在用于对连接做贡献的每个表的源代理键辅助列,在这个例子中是两个表(即,源表1610和1620)。源表1610在以下列中包括代表用于每个研究主体的人口统计数据的数据记录:列“AGE”、“AGE_U”和“GENDER”。源表1620在以下列中包括代表对不同主体用于实验室测试的结果的数据记录:列“RBC”、“HEMOGLOBIN”和“HEMATOCRIT”。这两个表连接,以创建目标表1630。目标表1630的每个数据记录包括来自源表1610和1620当中每个源表的值。因此,目标表1630的每个数据记录存储两个源代理键值(即,来自每个做贡献的源表的一个源代理键值)。这些源代理键值存储在列“DM_SKEY”和列“LAB1_SKEY”中。

[0116] 在该示例实施例中,数据世系传播模块(诸如图2的数据世系传播模块200)可以为目标表1630中的以下目标数据元素找出源数据元素:列“HEMOGLOBIN”和数据记录标识(即,代理键)“[LAB1_SUM]~105~1~LAB1”。数据世系传播模块可以查看数据世系映射表1600并在系统记录10中找出目标表“LAB1_SUM”(即,目标表1630)为目标表,并且找出列“HEMOGLOBIN”为目标列。根据这个信息,数据世系传播模块可以知道源数据元素是在源表“LAB1”(即,源表1620)和源列“HEMOGLOBIN”中。数据世系传播模块还知道源代理键值在目标表1630的列“LAB1_SKEY”中找到。在目标表1630中,数据世系传播模块可以在列“LAB1_SKEY”中查找目标数据记录。这个列包含代理键值“[LAB1]~105~1~LAB1”,并且数据世系传播模块知道源表、源列和代理键值。因此,数据世系传播模块可以识别出源数据元素是在源表1620中代理键值为“[LAB1]~105~1~LAB1”的列“HEMOGLOBIN”中。数据世系传播模块还可以识别出源数据值为“111”。

[0117] 另外,在该示例实施例中,数据世系传播模块可以为源表1620中的以下源数据元素找出目标数据元素:列“HEMOGLOBIN”和数据记录标识(即,代理键)“[LAB1]~105~1~LAB1”。数据世系传播模块可以查找数据世系映射表1600并且在系统记录10中找出源表“LAB1”(即,源表1620)为源表,并且找出列“HEMOGLOBIN”为源列。根据这个信息,数据世系传播模块可以确定它需要在目标表“LAB1_SUM”(即,目标表1630)的“HEMOGLOBIN”列查找目标数据元素。然后,数据世系传播模块可以利用源表1620的代理键“[LAB1]~105~1~LAB1”搜索影子表1640,并且确定目标表1630中期望数据记录的代理键是“[LAB1_SUM]~105~1~LAB1”。因此,数据世系传播模块可以识别出目标数据元素是在目标表1630中代理键值为“[LAB1_SUM]~105~1~LAB1”的列“HEMOGLOBIN”中。数据世系传播模块还可以识别出目标数据值为“111”。

[0118] 图17根据本发明实施例说明了判例变换的示例数据世系追踪。在该示例实施例中,源表“NRANGE”(即,源表1710)包括用于每个人口统计组的不同血液测试的正常值范围。源表“LAB”(即,源表1720)包括一组主体的实验室测试结果。目标表“LAB_SUM”(即,目标表1730)包括来自源表1720的列,加上包含来自源表1710的对应正常范围和正常范围单位的列。目标表1730还包括辅助列:一个过滤列(即,列“CASE_FVAL”)和两个源代理键列(即,列“LAB_SKEY”和“NRANGE_SKEY”)。

[0119] 在该示例实施例中,数据世系传播模块(诸如图2的数据世系传播模块200)可以为目标表1730中的以下目标数据元素找出源数据元素:列“RBC_NR”和数据记录标识(即,代理键)“[LAB_SUM]~103~1”。数据世系传播模块可以查看数据世系映射表1700并且在四个系统记录:系统记录11、12、13和14中找出目标表“LAB_SUM”(即目标表1730)作为目标列,并且

找出列“RBC_NR”作为目标列。这四个系统记录具有变换映射类型(即,列“TMAP_TYPE”)“CASE”。数据世系传播模块使用目标表1730中的两个辅助列来确定数据世系映射表1700中正确的系统记录:列“NRANGE_SKEY”和“CASE_FVAL”。列“NRANGE_SKEY”存储源表代理键,并且在数据世系映射表1700中具有对应的标识(即,列“SRC_SKEY_COLNAME”)。列“CASE_FVAL”存储判例条件子句标签,并且在数据世系映射表1700中具有对应的标识(即,列“FILTER_COL”)。数据世系传播模块可以确定目标表1730中的数据记录包括过滤列(即,列“CASE_FVAL”)中的值“MALE”。数据世系传播模块还可以确定数据世系映射表1700中正确的系统记录是在“FILTER_VAL”列中具有匹配标识的系统记录。因此,根据这四个系统记录,数据世系传播模块可以确定在“FILTER_VAL”列中包括“MALE”的系统记录是系统记录11。因此,源数据元素来自源表“NRANGE”(即,源表1710)并且具有列“MALE”。数据世系传播模块还可以确定源代理键值位于目标表1730的列“NRANGE_SKEY”中,并且源代理键值是“[NRANGE]~LAB1~RBC”。数据世系传播模块还可以确定源数据值为“4.7-6.1”。

[0120] 另外,在示例实施例中,数据世系传播模块可以为源表1710中的以下源数据元素找出目标数据元素:列“MALE”和数据记录标识(即,代理键)“[NRANGE]~LAB1~RBC”。数据世系传播模块可以查找数据世系映射表1700并且在系统记录11中找出源表“NRANGE”(即,源表1710)作为源表,并找出列“MALE”作为源列。根据这个信息,数据世系传播模块可以确定它需要在目标表“LAB_SUM”(即,目标表1730)的“RBC_NR”列中查找目标数据元素。然后,数据世系传播模块可以利用源表1710的代理键“[NRANGE]~LAB1~RBC”搜索影子表1740并且确定目标表1730中的期望数据记录的代理键是“[LAB_SUM]~103~1”和“[LAB_SUM]~104~1”。因此,数据世系传播模块可以识别出两个目标数据元素在目标表1730中代理键值为“[LAB_SUM]~103~1”和“[LAB_SUM]~104~1”的列“RBC_NR”中。数据世系传播模块还可以识别出用于这两个目标数据元素的目标数据值都是“4.7-6.1”。

[0121] 图18根据本发明实施例说明了数据世系模块(诸如作为图1的功能模块18的一部分的数据世系模块)的功能的流程图。在一种实施例中,图18的流程图的功能(下述)以及图19的流程图的功能(也在下述)每个都是由存储在存储器或某种其它计算机可读或有形介质中并被处理器执行的软件实现的。在其它实施例中,每个功能可以由硬件(例如,通过专用集成电路(“ASIC”)、可编程门阵列(“PGA”)、现场可编程门阵列(“FPGA”)等等的使用)或者硬件和软件的组合来执行。

[0122] 流程开始并前进到1805。在1805,目标数据元素映射到一个或多个源数据元素,作为包括一个或多个数据表的数据仓库的数据世系的部分,其中每个数据表包括一个或多个数据记录。目标数据元素包括目标数据记录的列。每个源数据元素包括源数据记录的列。目标数据元素是利用数据变换从一个或多个源数据元素得出的。在一种实施例中,数据仓库包括一个或多个有序的数据模型。每个有序的数据模型包括一个或多个数据表。另外,后续数据模型的一个或多个数据表是利用数据变换从前一数据模型的至少一个表构成的。在一种实施例中,数据变换是以下之一:直接映射变换、联合变换、连接变换、水平到垂直逆旋转变换、垂直到水平旋转变换,或者判例变换。流程前进到1810。

[0123] 在1810,一个或多个源代理键存储在目标数据记录的一个或多个辅助列中。每个源代理键包括包含对应源数据元素的源数据记录的标识。在一种实施例中,一个或多个代理键当中的每个代理键包括包含源数据记录的源数据表的标识。在一种实施例中,基于目

标数据元素到一个或多个源数据元素的映射,一个或多个辅助列自动添加到目标数据记录。流程前进到1815。

[0124] 在1815,对于每个源数据元素,数据世系映射系统记录存储在代表目标数据元素与对应源数据元素的映射的数据世系映射系统表中。数据世系映射系统记录包括对应的源数据元素、目标数据元素以及一个或多个辅助列。根据该实施例,数据世系映射系统记录还包括数据变换的类型。流程前进到1820。

[0125] 在1820,源数据元素映射到一个或多个目标数据元素,作为数据世系的一部分。这一个或多个目标数据元素是利用数据变换从源数据元素得出的。流程前进到1825。

[0126] 在1825,对于每个目标数据元素,影子系统记录存储在代表源数据元素与对应目标数据元素的映射的影子系统表中。影子系统记录包括识别源数据元素的源代理键、识别对应目标数据元素的目标代理键,以及一个或多个辅助列。因此,根据该实施例,数据世系包括一个或多个数据世系映射系统记录、一个或多个影子系统记录以及一个或多个源代理键。流程前进到1830。

[0127] 在1830,从利用数据世系映射的一个或多个数据元素选择数据元素。流程前进到1835。

[0128] 在1835,映射到被选数据元素的一个或多个数据元素利用数据世系在用户接口中显示。在一种实施例中,映射到被选数据元素的一个或多个数据元素还包括为每个数据元素显示数据世系内该数据元素的位置。流程前进到1840。

[0129] 在1840,对每个目标数据元素,一个或多个源数据元素其中的一个源数据元素被选作优选源数据元素。流程前进到1845。

[0130] 在1845,基于选定的一个或多个优选源数据元素创建数据世系的优选源路径。流程前进到1850。

[0131] 在1850,旋转列标识存储在目标数据记录的辅助列中。旋转列标识识别源数据记录的旋转列。流程前进到1855。

[0132] 在1855,过滤列标识存储在目标数据记录的第一辅助列中。过滤列标识识别源数据记录的过滤列。流程前进到1860。

[0133] 在1860,过滤值存储在目标数据记录的第二辅助列中,其中过滤值是源数据记录的过滤值。然后,流程结束。

[0134] 图19根据本发明实施例说明了关联信息传播模块(诸如作为图1功能模块18的一部分的关联信息传播模块)的功能的流程图。流程在1905开始。在1905,追踪数据仓库的数据世系被追踪。数据仓库包括一个或多个数据表。每个数据表包括一个或多个数据记录。在一种实施例中,数据仓库包括一个或多个有序的数据模型。每个有序的数据模型包括一个或多个数据表。另外,后续数据模型的一个或多个数据表是利用数据变换从前一数据模型的至少一个表构成的。在一种实施例中,数据变换是以下之一:直接映射变换、联合变换、连接变换、水平到垂直逆旋转变换、垂直到水平旋转变换,或者判例变换。流程前进到1910。

[0135] 在1910,识别主数据元素与关联对象之间的关联。主数据元素包括主数据记录的列。关联对象包括关联信息。主数据元素与关联对象之间的关联被称为主链接。在一种实施例中,关联对象是以下之一:差异、注释、状况标志、状态指示符、标签、图像、记录、文件、链接或文档。而且,在一种实施例中,代替主数据元素包括主数据记录的列,主数据元素包括

主数据记录。

[0136] 在1915,代表主链接的第一关联对象映射系统记录存储在关联对象映射系统表中。第一关联对象映射系统记录包括主数据元素的标识、关联对象的标识以及主链接指示符。在一种实施例中,第一关联对象映射系统记录包括识别主数据元素的数据模型的数据元素模型标识、识别主数据元素的数据表的数据元素表标识,以及识别主数据元素的列的数据元素列标识。流程前进到1920。

[0137] 在1920,利用数据世系识别次要数据元素与关联对象之间的关联。次要数据元素包括次要数据记录的列。次要数据元素与数据世系中的主数据元素相关。次要数据元素与关联对象之间的关联被称为次要链接。在一种实施例中,次要数据元素是主数据元素的源数据元素。而且,在一种实施例中,主数据元素的标识和次要数据元素的标识每个都是代理键。另外,在一种实施例中,代替次要数据元素包括次要数据记录的列,次要数据元素包括次要数据记录。在一种实施例中,次要数据元素与关联对象之间的关联是通过利用数据世系基于主数据元素识别次要数据元素来识别的。在实施例中,附加次要数据元素与关联对象之间的附加关联也利用数据世系来识别。流程前进到1925。

[0138] 在1925,代表次要链接并且代表主数据元素与次要数据元素之间的关联的第二关联对象映射系统记录存储在关联对象映射系统表中。第二关联对象映射系统记录包括次要数据元素的标识、关联对象的标识以及次要链接指示符。在一种实施例中,第二关联对象映射系统记录包括识别次要数据元素的数据模型的数据元素模型标识、识别次要数据元素的数据表的数据元素表标识以及识别次要数据元素的列的数据元素列标识。另外,在一种实施例中,第二关联对象映射系统记录包括辅次要数据元素是优选源数据元素的指示。在其中附加次要数据元素与关联对象之间的附加关联也利用数据世系识别的实施例中,附加的关联对象映射系统记录存储在关联对象映射系统表中。流程前进到1930。

[0139] 在1930,主数据元素或次要数据元素在用户接口中显示。在一种实施例中,主数据元素和次要数据元素都在用户接口中显示。然后,流程前进到1935。

[0140] 在1935,主数据元素或次要数据元素被选择。然后,流程前进到1940。

[0141] 在1940,关联对象利用第一关联对象映射系统记录或第二关联对象映射系统记录在用户接口中显示。更具体而言,如果主数据元素被选择,则关联对象利用第一关联对象映射系统记录在用户接口中显示。如果次要数据元素被选择,则关联对象利用第二关联对象映射系统记录在用户接口中显示。在一种实施例中,关联对象的显示是通过关联对象的类型过滤的。流程前进到1945。

[0142] 在1945,主数据元素被删除。流程前进到1950。

[0143] 在1950,第一关联对象映射系统记录被删除。流程前进到1955。

[0144] 在1955,第二关联对象映射系统记录被删除。流程前进到1960。

[0145] 在1960,关联对象的状态被修改为关闭状态。流程结束。

[0146] 因此,在一种实施例中,可以提供一种系统,该系统可以提供可以被存储在数据仓库中的一组数据的数据世系追踪,其中这组数据存储在一个或多个表中,并且其中每个表包括一个或多个数据记录。对于给定的目标数据元素,系统可以识别产生该目标数据元素的一组源数据元素,其中“数据元素”包括数据记录的列。类似地,对于给定的源数据元素,系统可以识别从该源数据元素得出的一组目标数据元素以及从该源数据元素到这组目标

数据元素的路径。由该系统提供的数据世系追踪还可以方便对所存储的数据的深入分析。

[0147] 另外,在另一种实施例中,可以提供一种系统,该系统可以充分利用数据世系追踪来支持链接到数据元素的“关联信息”或“关联对象”的可见性,并且把关联信息(即,关联对象)传播到相关的源数据元素和目标数据元素。根据该实施例,系统可以存储关联对象与和该关联对象关联的数据元素或数据记录之间的主链接和次要链接。因此,根据该实施例,关联对象,诸如差异,可以在数据世系中数据元素的上游和数据元素的下游都可见。这种关联对象的可见性也可以方便对所存储的数据的深入分析。

[0148] 贯穿本说明书描述的本发明的特征、结构或特性可以在一种或多种实施例中以任何合适的方式组合。例如,“一种实施例”、“一些实施例”、“某个实施例”、“某些实施例”或者其它类似语言的使用贯穿本说明书都指联系该实施例所描述的特定特征、结构或特性可以包括在本发明至少一种实施例中的事实。因此,短语“一种实施例”、“一些实施例”、“某个实施例”、“某些实施例”或者其它类似语言的出现贯穿本说明书不一定全部指相同的实施例组,并且所描述的特征、结构或特性可以在一种或多种实施例中以任何合适的方式组合。

[0149] 本领域普通技术人员将很容易理解,如上讨论的本发明可以以不同次序的步骤实践,和/或以配置中与所公开那些不同的元件实践。因此,虽然本发明已经基于这些优选实施例进行了描述,但是对本领域技术人员来说很显然,在仍然属于本发明精神和范围的情况下,某些修改、变化和备选构造将是显然的。因此,为了确定本发明的界限和范围,应当参考所附权利要求。

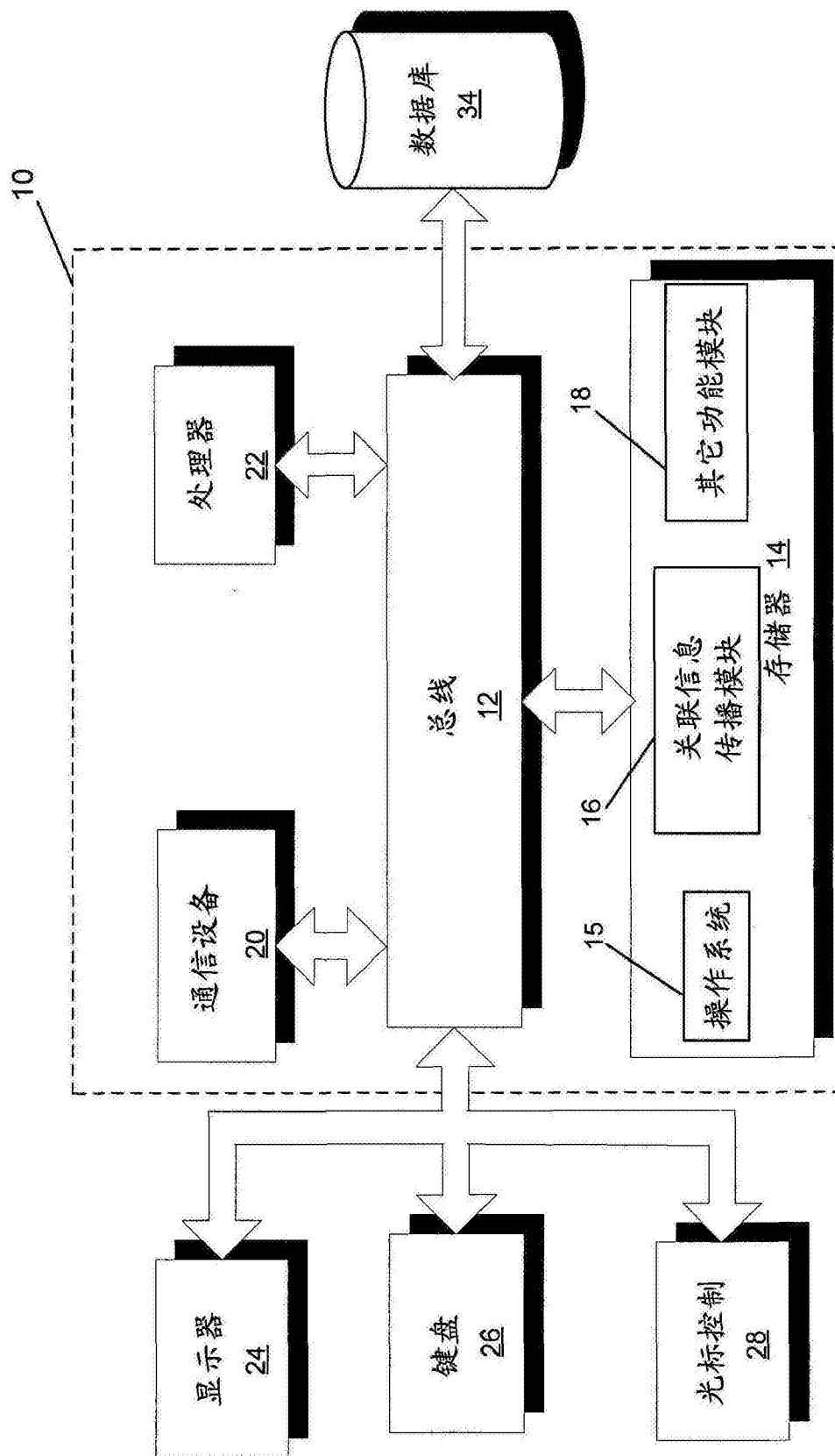


图1

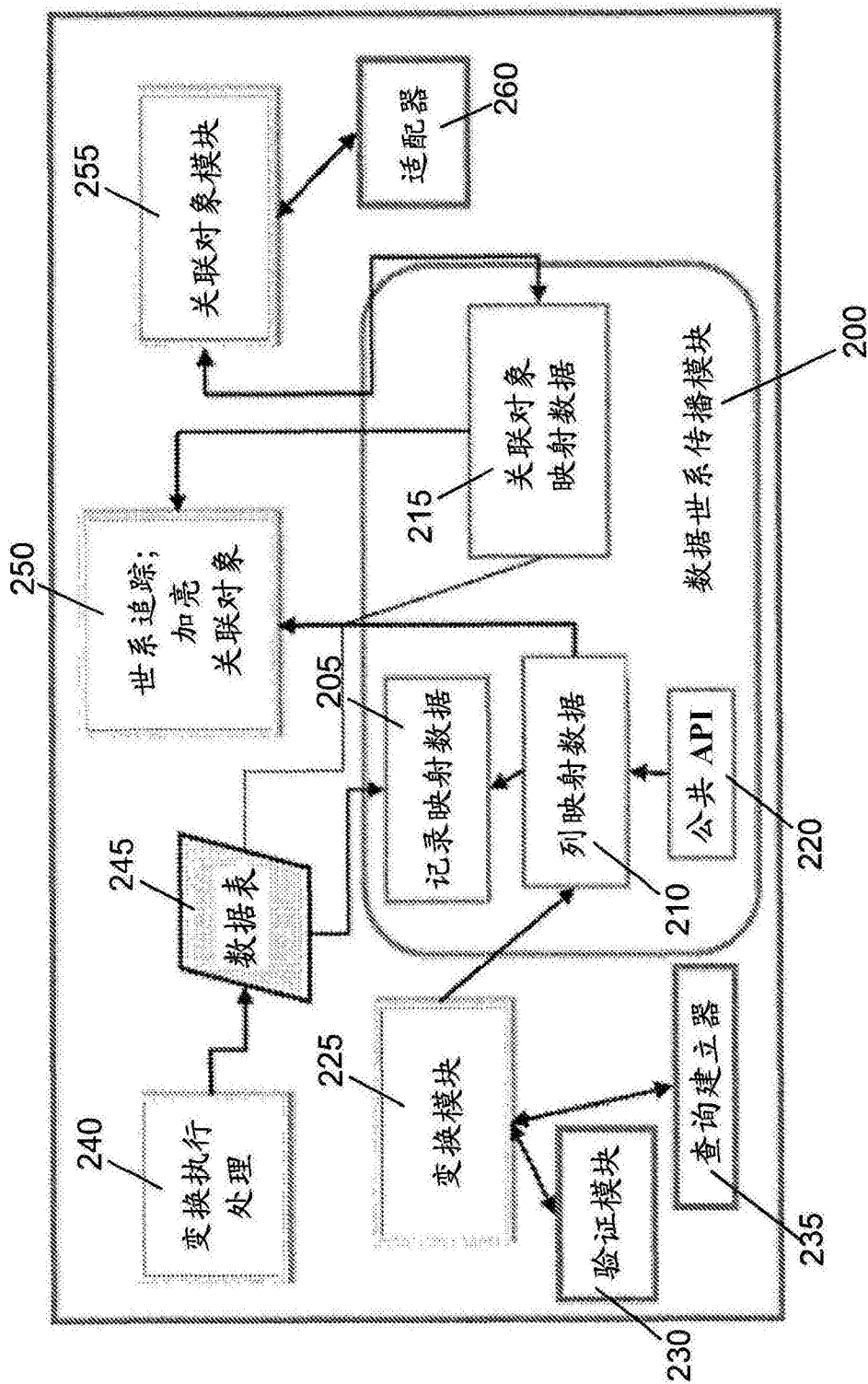


图2

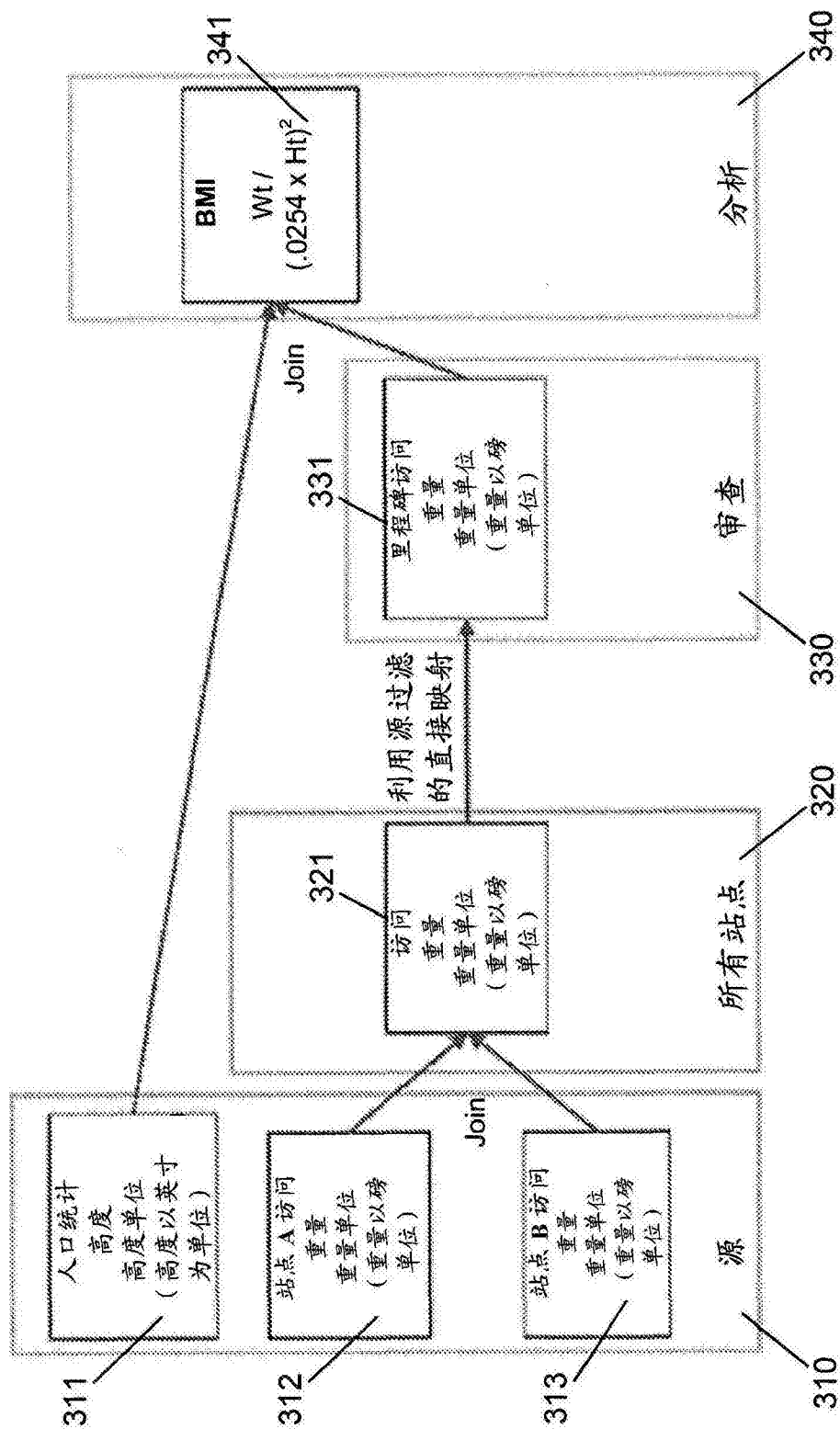


图3

400	410	420	460	430	440	450
401	数据模型.表.列					
402	☒ Analysis.BMI.BMI		33.3	NUMBER(4,1)		Subject:1007, Visit:4, Site:McLean
403	☒ Source.DEMOGRAPHY.HEIGHT		62.0	NUMBER	NO	Subject:1007
404	☒ Source.DEMOGRAPHY.HEIGHTU		IN	VARCHAR2(20)	NO	Subject:1007
405	☒ Review.MILESTONE_VISITS.WEIGHT		82.1	NUMBER	YES	Subject:1007, Visit:4, Site:McLean
406	☒ All_Sites.VISITS.WT		181.0	NUMBER	YES	Subject:1007, Visit:4, Site:McLean
407	☒ Source.SITE_B_VISITS.WT		181.0	NUMBER	YES	Subject:1007, Visit:4, Site:McLean
408	☒ Review.MILESTONE_VISITS.WEIGHTU		KG	VARCHAR2(20)	NO	Subject:1007, Visit:4, Site:McLean
409	☒ All_Sites.VISITS.WTU		LBS	VARCHAR2(20)	NO	Subject:1007, Visit:4, Site:McLean
	☒ Source.SITE_B_VISITS.WTU		LBS	VARCHAR2(20)	NO	Subject:1007, Visit:4, Site:McLean
关闭						
470						

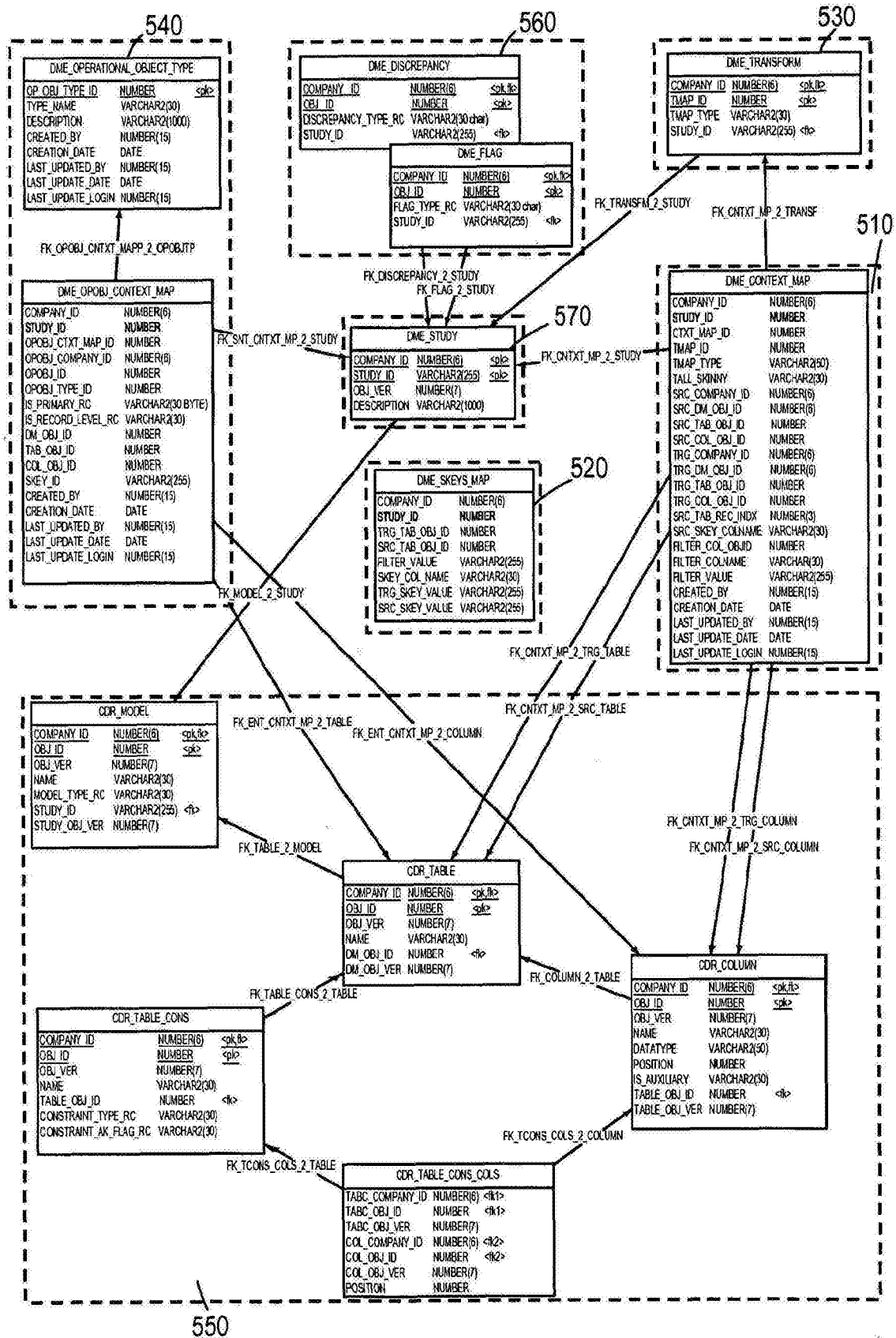


图5

600

DME_CONTEXT_MAP	
COMPANY_ID	NUMBER(6)
STUDY_ID	NUMBER
LIFECYCLE_RC	VARCHAR2(30)
TMAP_ID	NUMBER
TMAP_TYPE_RC	VARCHAR2(50)
TALL_SKINNY_RC	VARCHAR2(30)
SRC_DM_OBJ_ID	NUMBER
SRC_TAB_OBJ_ID	NUMBER
SRC_COL_OBJ_ID	NUMBER
TRG_DM_OBJ_ID	NUMBER
TRG_TAB_OBJ_ID	NUMBER
TRG_COL_OBJ_ID	NUMBER
TRG_PERSISTENT_RC	VARCHAR2(30)
SRC_SKEY_COL_OBJ_ID	NUMBER
PREFERRED_ROUTE_RC	VARCHAR2(30)
FILTER_COL_OBJ_ID	NUMBER
FILTER_VALUE	VARCHAR2(255)
CREATED_BY	NUMBER(15)
CREATION_DATE	DATE
LAST_UPDATED_BY	NUMBER(15)
LAST_UPDATE_DATE	DATE
LAST_UPDATE_LOGIN	NUMBER(15)

图6

700

DME_CTXT_SKEYS_MAP	
COMPANY_ID	NUMBER(6)
STUDY_ID	NUMBER
LIFECYCLE_RC	VARCHAR2(30)
TRG_TAB_OBJ_ID	NUMBER
SRC_TAB_OBJ_ID	NUMBER
FILTER_VALUE	VARCHAR2(255)
SRC_SKEY_COL_OBJ_ID	NUMBER
TRG_SKEY_VALUE	VARCHAR2(4000)
SRC_SKEY_VALUE	VARCHAR2(4000)

图7

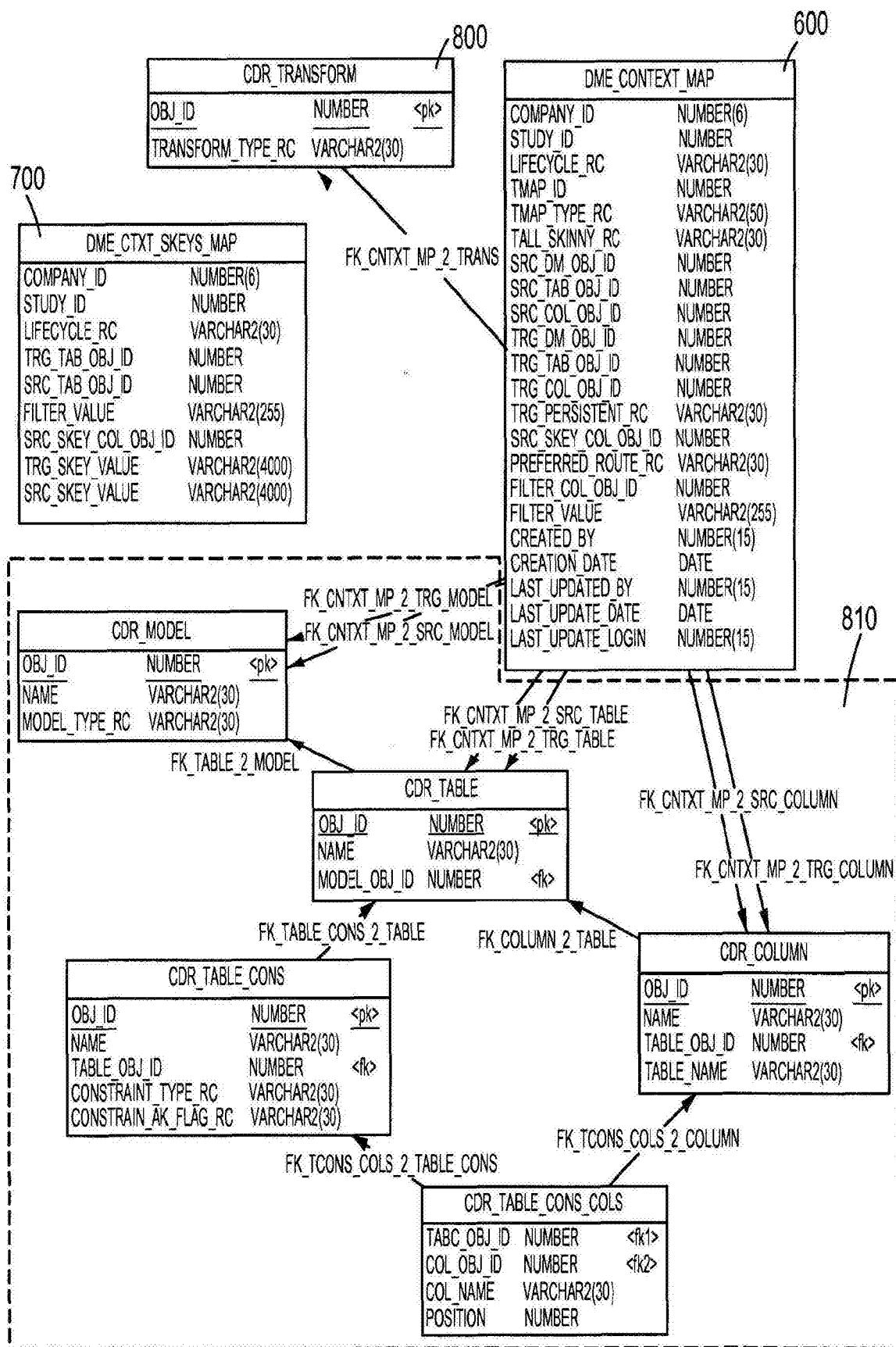


图8

900

DME_OPOBJ_CONTEXT_MAP	
COMPANY_ID	NUMBER(6)
STUDY_ID	NUMBER
LIFECYCLE_RC	VARCHAR2(30)
OPOBJ_ID	NUMBER
OPOBJ_TYPE_ID	NUMBER
IS_PRIMARY_RC	VARCHAR2(30)
IS_RECORD_LEVEL_RC	VARCHAR2(30)
PREFERRED_ROUTE_RC	VARCHAR2(30)
DM_OBJ_ID	NUMBER
TAB_OBJ_ID	NUMBER
COL_OBJ_ID	NUMBER
SKEY_VALUE	VARCHAR2(4000)
CREATED BY	NUMBER(15)
CREATION_DATE	DATE
LAST_UPDATED_BY	NUMBER(15)
LAST_UPDATE_DATE	DATE
LAST_UPDATE_LOGIN	NUMBER(15)

图9

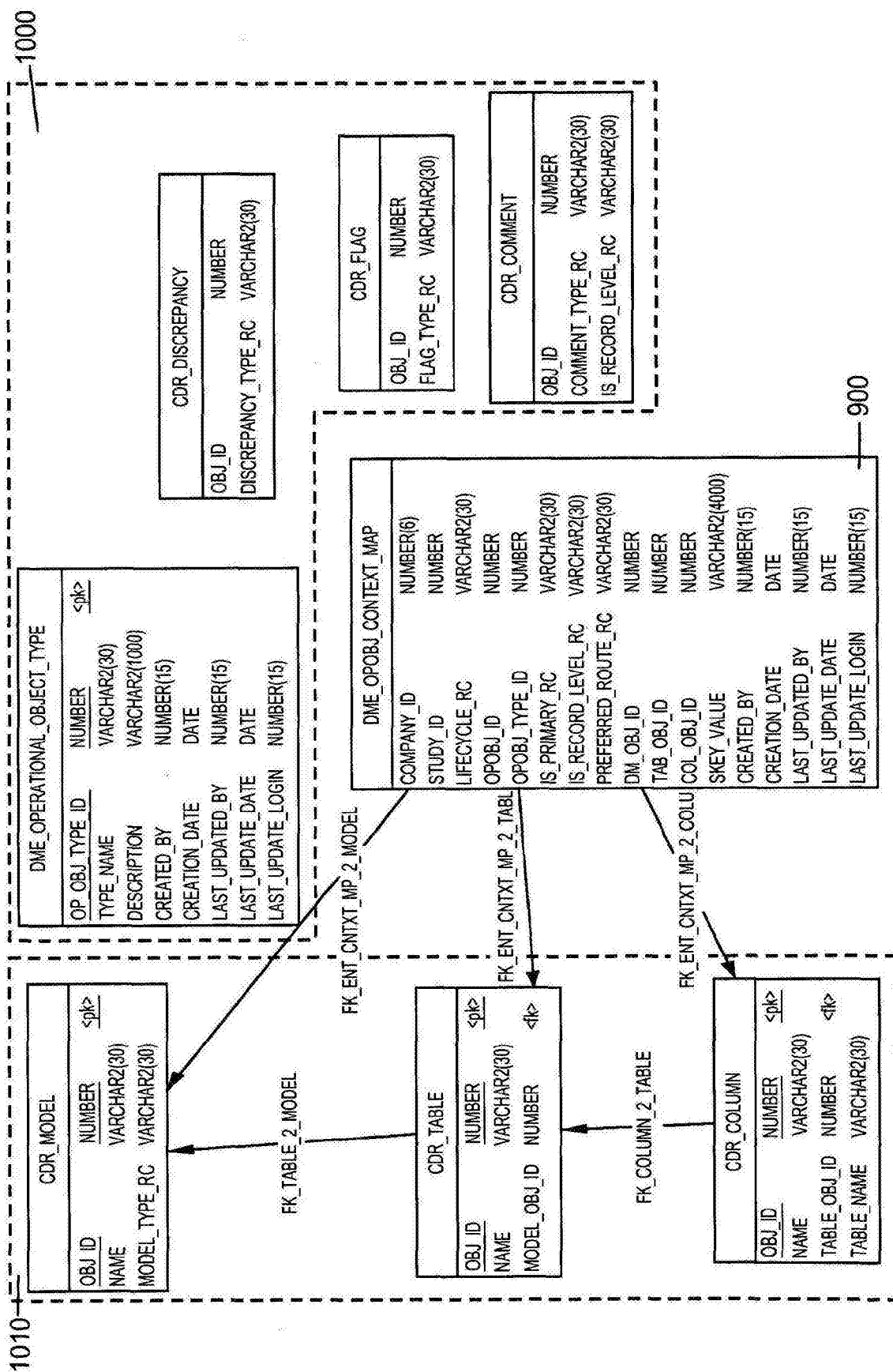


图10

DME_CONTEXT_MAP													
CTX_MAP_ID	STUDY	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	SRC_TAB_REC_IDX	FILTER_COL	FILTER_VAL	SRC_SKEY	COLNAME	
2	ST123	1	UNION	LAB1	STUDYID	LAB_SUM	STUDYID	1			UNION_SRC_SKEY		
3	ST123	1	UNION	LAB1	SUBJID	LAB_SUM	SUBJID	1			UNION_SRC_SKEY		
4	ST123	1	UNION	LAB1	VISITNUM	LAB_SUM	VISITNUM	1			UNION_SRC_SKEY		
5	ST123	1	UNION	LAB1	SITE	LAB_SUM	SITE	1			UNION_SRC_SKEY		
6	ST123	1	UNION	LAB1	AGE	LAB_SUM	AGE	1			UNION_SRC_SKEY		
7	ST123	1	UNION	LAB1	AGE_U	LAB_SUM	AGE_U	1			UNION_SRC_SKEY		
8	ST123	1	UNION	LAB1	GENDER	LAB_SUM	GENDER	1			UNION_SRC_SKEY		
9	ST123	1	UNION	LAB1	RBC	LAB_SUM	RBC	1			UNION_SRC_SKEY		
10	ST123	1	UNION	LAB1	HEMOGLOBIN	LAB_SUM	HEMOGLOBIN	1			UNION_SRC_SKEY		
11	ST123	1	UNION	LAB1	HEMATOCRIT	LAB_SUM	HEMATOCRIT	1			UNION_SRC_SKEY		
12	ST123	1	UNION	LAB2	STUDYID	LAB_SUM	STUDYID	1			UNION_SRC_SKEY		
.....													
25	ST123	1	UNION	LAB2	RBC	LAB_SUM	RBC	1			UNION_SRC_SKEY		
28	ST123	1	UNION	LAB3	STUDYID	LAB_SUM	STUDYID	1			UNION_SRC_SKEY		
.....													
35	ST123	1	UNION	LAB3	RBC	LAB_SUM	RBC	1			UNION_SRC_SKEY		
.....													

LAB1	DMESSKEY	STUDYID	SUBJID	VISITNUM	SITE	AGE	AGE_U	GENDER	RBC	HEMOGLOBIN	HEMATOCRIT
[LAB1]-101-1-LAB1	ST123	101	1	LAB1	35	YEAR	3.8	9	30		
[LAB1]-103-1-LAB1	ST123	103	1	LAB1	21	YEAR	5.3	15	45		
[LAB1]-104-1-LAB1	ST123	104	1	LAB1	28	YEAR	3.2	16	17		

LAB2	DMESSKEY	STUDYID	SUBJID	VISITNUM	SITE	AGE	AGE_U	GENDER	RBC		
[LAB2]-102-1-LAB2	ST123	102	1	LAB2	9	YEAR	4.8				
[LAB2]-105-1-LAB2	ST123	105	1	LAB2	7	YEAR	4.5				

LAB3	DMESSKEY	STUDYID	SUBJID	VISITNUM	SITE	AGE	AGE_U	GENDER	RBC		
[LAB3]-106-1-LAB3	ST123	106	1	LAB3	2	MONTH	4.6				

LAB_SUM	DMESSKEY	STUDYID	SUBJID	VISITNUM	SITE	AGE	AGE_U	GENDER	RBC	HEMOGLOBIN	HEMATOCRIT	UNION_SRC_SKEY
[LAB_SUM]-101-1	ST123	101	1	LAB1	35	YEAR	3.8	9	30			[LAB1]-101-1-LAB1
[LAB_SUM]-102-1	ST123	102	1	LAB2	9	YEAR	4.8					[LAB1]-102-1-LAB2
[LAB_SUM]-103-1	ST123	103	1	LAB1	21	YEAR	5.3	15	45			[LAB1]-103-1-LAB1
[LAB_SUM]-104-1	ST123	104	1	LAB1	28	YEAR	3.2	16	17			[LAB1]-104-1-LAB1
[LAB_SUM]-105-1	ST123	105	1	LAB2	7	YEAR	4.5					[LAB2]-105-1-LAB2
[LAB_SUM]-106-1	ST123	106	1	LAB3	2	MONTH	4.6					[LAB3]-106-1-LAB3

DME_SKEYS_MAP													
TRG_TAB	SRC_TAB	SKEY_COLNAME	TRG_SKEY_VALUE	SRC_SKEY_VALUE									
LAB_SUM	LAB1	UNION_SRC_SKEY	[LAB_SUM]-101-1	[LAB1]-101-1-LAB1									
LAB_SUM	LAB2	UNION_SRC_SKEY	[LAB_SUM]-102-1	[LAB2]-102-1-LAB2									
LAB_SUM	LAB1	UNION_SRC_SKEY	[LAB_SUM]-103-1	[LAB1]-103-1-LAB1									
LAB_SUM	LAB1	UNION_SRC_SKEY	[LAB_SUM]-104-1	[LAB1]-104-1-LAB1									
LAB_SUM	LAB2	UNION_SRC_SKEY	[LAB_SUM]-105-1	[LAB2]-105-1-LAB2									
LAB_SUM	LAB3	UNION_SRC_SKEY	[LAB_SUM]-106-1	[LAB3]-106-1-LAB3									

图11

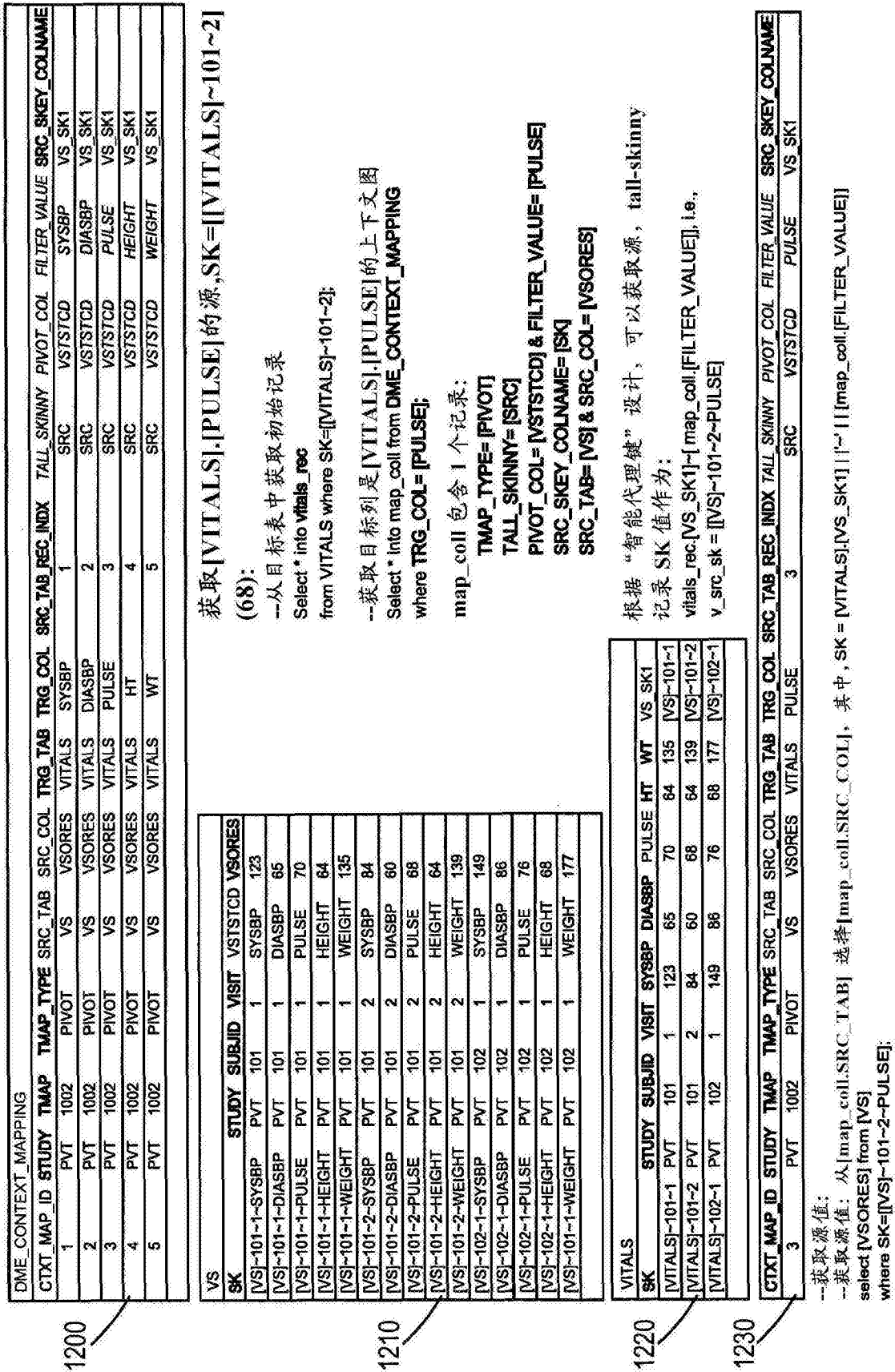


图12

VS	SK	STUDY	SUBJID	VISIT	VS1STD	VSORES
[VS]	-101-1-SYSBP	PVT	101	1	SYSBP	123
[VS]	-101-1-DIASBP	PVT	101	1	DIASBP	65
[VS]	-101-1-PULSE	PVT	101	1	PULSE	70
[VS]	-101-1-HEIGHT	PVT	101	1	HEIGHT	64
[VS]	-101-1-WEIGHT	PVT	101	1	WEIGHT	135
[VS]	-101-2-SYSBP	PVT	101	2	SYSBP	84
[VS]	-101-2-DIASBP	PVT	101	2	DIASBP	60
[VS]	-101-2-PULSE	PVT	101	2	PULSE	68
[VS]	-101-2-HEIGHT	PVT	101	2	HEIGHT	64
[VS]	-101-2-WEIGHT	PVT	101	2	WEIGHT	139
[VS]	-102-1-SYSBP	PVT	102	1	SYSBP	149
[VS]	-102-1-DIASBP	PVT	102	1	DIASBP	86
[VS]	-102-1-PULSE	PVT	102	1	PULSE	76
[VS]	-102-1-HEIGHT	PVT	102	1	HEIGHT	68
[VS]	-102-1-WEIGHT	PVT	102	1	WEIGHT	177

VITALS									
SK	STUDY	SUBJID	VISIT	SV38BP	DIAB3P	PULSE	HT	WT	VS SK1
[NITALS]-101-1	PVT	101	1	123	65	70	64	135	[VS]-101-1
[NITALS]-101-2	PVT	101	2	84	60	68	64	139	[VS]-101-2
[NITALS]-102-1	PVT	102	1	149	86	76	68	177	[VS]-102-1

CXTX_MAP_ID	STUDY	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	SRC_TAB	REC_IDX	TALL_SKINNY	PIVOT_COL	FILTER	VALUE	SRC_SKY	COLNAME
3	PVT	1002	PIVOT	VS	VSCORES	VITALS	PULSE	3		SRC	VSTSTCD	PULSE		VS SK1	

DME_SKESYS_MAP				
TRG TAB	SRC TAB	SKKEY COLNAME	TRG SKKEY VALUE	SRC SKKEY VALUE
VITALS	VS	VS_SK1	VITALSJ-101-1	VSP-101-1
VITALS	VS	VS_SK1	VITALSJ-101-2	VSP-101-2

获取[VS].[VSORES] 的目标, SK=[VS]~101~2~PULSE]
(68):

---从源表中获取初始记录

Select * into vs rec

from VS where SK=[VS]~101~2~PULSE];

```
--获取源列[V$.V$ORES]的上下文图
Select * into map_coll from DME_CONTEXT_MAPPING
where SRC_COL=[V$ORES];
```

map coll 包含 5 个记录:

TMAP_TYPE = [PIVOT]

TALL_SKINNY = [SRC]

PIVOT_COL=[VSTSTCD]

FILTER_VALUE=[SYSBP | DIASBP | PULSE | HEIGHT | WEIGHT]

SRC_SKEY_COLNAME

TRG_TAB=[VITALS]

TRG_COL=[SYSBP | DIASBP | PULSE | HT | WT]

```
--因为 tmap_type==PIVOT, 所以使用列[VSTSTCD]中的源记录值
--作为 FILTER_VALUE 来获取相关上下文图记录
Select * into map_coll from DME_CONTEXT_MAPPING
where SRC_COL=[VSORES] and FILTER_VALUE=[vs_rec.VSTSTCD];
map_coll 包含 1 个记录;
```

TRG TAB= [VITALS] & TRG COL= [PULSE]

---目标表列IVS SK1不具有索引，

因此使用编索引的 DME SKEYS MAP 表来获取目标记录 SK

Select TRG_SKEY_VALUE into v_trg_sk

From DME SKEYS MAP

Where TRG TAB=[VITALS] and SRC TAB=[VS]

And SKEY COLNAME=[VS SK1]

```
And SRC_SKEY_VALUE= Replace(vs_rec.[SK], '~' | map_coll. [FILTER_VALUE]);
```

---获取目标值:从[map coll.TRG TAB] 选择[map coll.TRG COL],其中SK=V TRG SK

select [PULSE] from [VITALS]
where SK=[[VITALS]~101~2];

图 13

DME_CONTEXT_MAPPING															
CTXT_MAP_ID	STUDY	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	SRC_TAB	REC_IDX	TALL	SKINNY	PIVOT_COL	FILTER_VALUE	SRC_SKEY_COLNAME	
1	UnPVT	1002	UNPIVOT	VITALS	SYSBP	VS	VSORES	1	TRG	VSTSTCD	SYSBP	VITALS_SK1			
2	UnPVT	1002	UNPIVOT	VITALS	DIASBP	VS	VSORES	1	TRG	VSTSTCD	DIASBP	VITALS_SK1			
3	UnPVT	1002	UNPIVOT	VITALS	PULSE	VS	VSORES	1	TRG	VSTSTCD	PULSE	VITALS_SK1			
4	UnPVT	1002	UNPIVOT	VITALS	HT	VS	VSORES	1	TRG	VSTSTCD	HEIGHT	VITALS_SK1			
5	UnPVT	1002	UNPIVOT	VITALS	WT	VS	VSORES	1	TRG	VSTSTCD	WEIGHT	VITALS_SK1			
VITALS															
SK	STUDY	SUBJID	VISIT	SYSBP	DIASBP	PULSE	HT	WT							
[VITALS]-101-1	UnPVT	101	1	123	65	70	64	135							
[VITALS]-101-2	UnPVT	101	2	84	60	68	64	139							
[VITALS]-102-1	UnPVT	102	1	149	86	76	68	177							

1400

VITALS															
SK	STUDY	SUBJID	VISIT	SYSBP	DIASBP	PULSE	HT	WT							
[VITALS]-101-1	UnPVT	101	1	123	65	70	64	135							
[VITALS]-101-2	UnPVT	101	2	84	60	68	64	139							
[VITALS]-102-1	UnPVT	102	1	149	86	76	68	177							

1410

VS														
SK	STUDY	SUBJID	VISIT	VSTSTCD	VSORES	VITALS_SK1								
[VS]-101-1-SYSBP	UnPVT	101	1	SYSBP	123	[VITALS]-101-1								
[VS]-101-1-DIASBP	UnPVT	101	1	DIASBP	65	[VITALS]-101-1								
[VS]-101-1-PULSE	UnPVT	101	1	PULSE	70	[VITALS]-101-1								
[VS]-101-1-HEIGHT	UnPVT	101	1	HEIGHT	64	[VITALS]-101-1								
[VS]-101-1-WEIGHT	UnPVT	101	1	WEIGHT	135	[VITALS]-101-1								
[VS]-101-2-SYSBP	UnPVT	101	2	SYSBP	84	[VITALS]-101-2								
[VS]-101-2-DIASBP	UnPVT	101	2	DIASBP	60	[VITALS]-101-2								
[VS]-101-2-PULSE	UnPVT	101	2	PULSE	68	[VITALS]-101-2								
[VS]-101-2-HEIGHT	UnPVT	101	2	HEIGHT	64	[VITALS]-101-2								
[VS]-101-2-WEIGHT	UnPVT	101	2	WEIGHT	139	[VITALS]-101-2								
[VS]-102-1-SYSBP	UnPVT	102	1	SYSBP	149	[VITALS]-102-1								
[VS]-102-1-DIASBP	UnPVT	102	1	DIASBP	86	[VITALS]-102-1								
[VS]-102-1-PULSE	UnPVT	102	1	PULSE	76	[VITALS]-102-1								
[VS]-102-1-HEIGHT	UnPVT	102	1	HEIGHT	68	[VITALS]-102-1								
[VS]-102-1-WEIGHT	UnPVT	102	1	WEIGHT	177	[VITALS]-102-1								

1420

CTXT_MAP_ID															
3	UnPVT	1002	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	TRG_TAB	REC_IDX	TALL	SKINNY	PIVOT_COL	FILTER_VALUE	SRC_SKEY_COLNAME

1430

--如下获取源值:

--从[map_coll.SRC_TAB] 选择 map_coll.SRC_COL, 其中, SK = [vs_rec.[SRC_SKEY_COLNAME]]

select [PULSE] from [VITALS]

where SK=[VITALS]-101-2;

获取[VITALS].[PULSE] 的目标, SK=[VITALS]~101~2 (68):

--从源表中获取初始记录

Select * into vitals_rec
from VITALS where SK=[VITALS]~101~2;

--获取目标列[VSI],[VSORES]的上下文图

Select * into map_coll from DME_CONTEXT_MAPPING
where SRC_COL= [PULSE];

map_coll 包含 1 个记录:

TMAP_TYPE= [UNPIVOT]

TALL_SKINNY= [TRG]

PIVOT_COL= [VSTSTD] & FILTER_VALUE= [PULSE]

SRC_KEY_COLNAME= [SK]

TRG_TAB= [VS] & TRG_COL= [VSORES]

1410

VITALS						
SK	STUDY	SUBJID	VISIT	SYSBP	DIASBP	PULSE HT WT
[VITALS]~101~1	UnPVT	101	1	123	65	70 64 135
[VITALS]~101~2	UnPVT	101	2	84	60	68 64 139
[VITALS]~102~1	UnPVT	102	1	149	86	76 68 177

1420

VS	STUDY	SUBJID	VISIT	VSTSTD	VSORES	VITALS_SK1
[VS]~101~1~SYSBP	UnPVT	101	1	SYSBP	123	[VITALS]~101~1
[VS]~101~1~DIASBP	UnPVT	101	1	DIASBP	65	[VITALS]~101~1
[VS]~101~1~PULSE	UnPVT	101	1	PULSE	70	[VITALS]~101~1
[VS]~101~1~HEIGHT	UnPVT	101	1	HEIGHT	64	[VITALS]~101~1
[VS]~101~1~WEIGHT	UnPVT	101	1	WEIGHT	135	[VITALS]~101~1
[VS]~101~2~SYSBP	UnPVT	101	2	SYSBP	84	[VITALS]~101~2
[VS]~101~2~DIASBP	UnPVT	101	2	DIASBP	60	[VITALS]~101~2
[VS]~101~2~PULSE	UnPVT	101	2	PULSE	68	[VITALS]~101~2
[VS]~101~2~HEIGHT	UnPVT	101	2	HEIGHT	64	[VITALS]~101~2
[VS]~101~2~WEIGHT	UnPVT	101	2	WEIGHT	139	[VITALS]~101~2
[VS]~102~1~SYSBP	UnPVT	102	1	SYSBP	149	[VITALS]~102~1
[VS]~102~1~DIASBP	UnPVT	102	1	DIASBP	86	[VITALS]~102~1
[VS]~102~1~PULSE	UnPVT	102	1	PULSE	76	[VITALS]~102~1
[VS]~102~1~HEIGHT	UnPVT	102	1	HEIGHT	68	[VITALS]~102~1
[VS]~102~1~WEIGHT	UnPVT	102	1	WEIGHT	177	[VITALS]~102~1

1500

CDXT_MAP_ID	STUDY	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	TRG_TAB	TRG_COL	SRC_TAB	REC_INDX	TALL_SKINNY	PIVOT_COL	FILTER_VALUE	SRC_KEY_COLNAME
3	UnPVT	1002	UNPIVOT	VITALS	PULSE	VS	VSORES	VS	VSORES	VSORES	1	TRG	VSTSTD	PULSE	VITALS_SK1

1510

DME_SKEYS_MAP					
TRG_TAB	SRC_TAB	SKEY_COLNAME	TRG_SKEY_VALUE	SRC_SKEY_VALUE	
S	VITALS	VITALS_SK1	[VS]~101~1~HEIGHT	[VITALS]~101~1	
VS	VITALS	VITALS_SK1	[VS]~101~1~WEIGHT	[VITALS]~101~1	
VS	VITALS	VITALS_SK1	[VS]~101~2~SYSBP	[VITALS]~101~2	
VS	VITALS	VITALS_SK1	[VS]~101~2~DIASBP	[VITALS]~101~2	
VS	VITALS	VITALS_SK1	[VS]~101~2~PULSE	[VITALS]~101~2	
VS	VITALS	VITALS_SK1	[VS]~101~2~HEIGHT	[VITALS]~101~2	

--获取目标值:

--从[map_coll.TRG_TAB] 选择[map_coll.TRG_COL], 其中 SK=v_trg_sk

Select [VSORES] from [VS]

where SK=[VS]~101~2~PULSE;

--目标表列[VITALS_SK1]不具有索引

--因此使用编索引的 DME_SKEYS_MAP 表

--来获取目标记录 SK

Select TRG_SKEY_VALUE into v_trg_sk

From DME_SKEYS_MAP

Where TRG_TAB= [VS] and SRC_TAB= [VITALS]

And SKEY_COLNAME= [VITALS_SK1]

And SRC_SKEY_VALUE= [[VITALS]~101~2]

And [PULSE]= [substr([TRG_SKEY_VALUE], instr([TRG_SKEY_VALUE], '~', 1, 1) + 1)];

图15

DME CONTEXT MAP											
CTXT_MAP_ID	STUDY	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	SRC_TAB_REC_IDX	FILTER	COL_FILTER	SRC_SKEY_COLNAME
2	ST123	1	MAP	DM	STUDYID	LAB1_SUM	STUDYID	1			DM_SKEY
3	ST123	1	MAP	DM	SUBJID	LAB1_SUM	SUBJID	1			DM_SKEY
4	ST123	1	MAP	DM	AGE	LAB1_SUM	AGE	1			DM_SKEY
5	ST123	1	MAP	DM	AGE_U	LAB1_SUM	AGE_U	1			DM_SKEY
6	ST123	1	MAP	DM	GENDER	LAB1_SUM	GENDER	1			DM_SKEY
7	ST123	2	MAP	LAB1	VISITNUM	LAB1_SUM	VISITNUM	1			LAB1_SKEY
8	ST123	2	MAP	LAB1	SITE	LAB1_SUM	SITE	1			LAB1_SKEY
9	ST123	2	MAP	LAB1	RBC	LAB1_SUM	RBC	1			LAB1_SKEY
10	ST123	2	MAP	LAB1	HEMOGLOBIN	LAB1_SUM	HEMOGLOBIN	1			LAB1_SKEY
11	ST123	2	MAP	LAB1	HEMATOCRIT	LAB1_SUM	HEMATOCRIT	1			LAB1_SKEY
...											
DM											
DMESSKEY	STUDYID	SUBJID	AGE	AGE_U	GENDER	SITE	VISITNUM	HEMOGLOBIN	HEMATOCRIT	RBC	HEMOGLOBIN
[DM]-101	ST123	101	35	YEAR	FEMALE	LAB1	1	ST123	101	LAB1	3.8
[DM]-102	ST123	102	9	YEAR	FEMALE	LAB1	1	ST123	102	LAB1	4.8
[DM]-103	ST123	103	21	YEAR	MALE	LAB1	1	ST123	103	LAB1	5.3
[DM]-104	ST123	104	28	YEAR	MALE	LAB1	1	ST123	104	LAB1	3.2
[DM]-105	ST123	105	7	YEAR	MALE	LAB1	1	ST123	105	LAB1	4.5
[DM]-106	ST123	106	2	MONTH	FEMALE	LAB1	1	ST123	106	LAB1	4.6
LAB1_SUM											
DMESSKEY	STUDYID	SUBJID	AGE	AGE_U	GENDER	SITE	VISITNUM	HEMOGLOBIN	HEMATOCRIT	DM_SKEY	LAB1_SKEY
[LAB1_SUM]-101-1-LAB1	ST123	101	35	YEAR	FEMALE	LAB1	1	ST123	101	[DM]-101	[LAB1]-101-1-LAB1
[LAB1_SUM]-102-1-LAB1	ST123	102	9	YEAR	FEMALE	LAB1	1	ST123	102	[DM]-102	[LAB1]-102-1-LAB1
[LAB1_SUM]-103-1-LAB1	ST123	103	21	YEAR	MALE	LAB1	1	ST123	103	[DM]-103	[LAB1]-103-1-LAB1
[LAB1_SUM]-104-1-LAB1	ST123	104	28	YEAR	MALE	LAB1	1	ST123	104	[DM]-104	[LAB1]-104-1-LAB1
[LAB1_SUM]-105-1-LAB1	ST123	105	7	YEAR	MALE	LAB1	1	ST123	105	[DM]-105	[LAB1]-105-1-LAB1
[LAB1_SUM]-106-1-LAB1	ST123	106	2	MONTH	FEMALE	LAB1	1	ST123	106	[DM]-106	[LAB1]-106-1-LAB1
DME SKEYS MAP											
TRG_TAB	SRC_TAB	SKEY_COLNAME	TRG_SKEY_VALUE	SRC_SKEY_VALUE							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-101-1-LAB1	[DM]-101							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-101-1-LAB1	[LAB1]-101-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-102-1-LAB1	[DM]-102							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-102-1-LAB1	[LAB1]-102-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-103-1-LAB1	[DM]-103							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-103-1-LAB1	[LAB1]-103-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-104-1-LAB1	[DM]-104							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-104-1-LAB1	[LAB1]-104-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-105-1-LAB1	[DM]-105							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-105-1-LAB1	[LAB1]-105-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-106-1-LAB1	[DM]-106							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-106-1-LAB1	[LAB1]-106-1-LAB1							
...											

DME CONTEXT MAP											
CTXT_MAP_ID	STUDY	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	SRC_TAB_REC_IDX	FILTER	COL_FILTER	SRC_SKEY_COLNAME
2	ST123	1	MAP	DM	STUDYID	LAB1_SUM	STUDYID	1			DM_SKEY
3	ST123	1	MAP	DM	SUBJID	LAB1_SUM	SUBJID	1			DM_SKEY
4	ST123	1	MAP	DM	AGE	LAB1_SUM	AGE	1			DM_SKEY
5	ST123	1	MAP	DM	AGE_U	LAB1_SUM	AGE_U	1			DM_SKEY
6	ST123	1	MAP	DM	GENDER	LAB1_SUM	GENDER	1			DM_SKEY
7	ST123	2	MAP	LAB1	VISITNUM	LAB1_SUM	VISITNUM	1			LAB1_SKEY
8	ST123	2	MAP	LAB1	SITE	LAB1_SUM	SITE	1			LAB1_SKEY
9	ST123	2	MAP	LAB1	RBC	LAB1_SUM	RBC	1			LAB1_SKEY
10	ST123	2	MAP	LAB1	HEMOGLOBIN	LAB1_SUM	HEMOGLOBIN	1			LAB1_SKEY
11	ST123	2	MAP	LAB1	HEMATOCRIT	LAB1_SUM	HEMATOCRIT	1			LAB1_SKEY
...											
DM											
DMESSKEY	STUDYID	SUBJID	AGE	AGE_U	GENDER	SITE	VISITNUM	HEMOGLOBIN	HEMATOCRIT	RBC	HEMOGLOBIN
[DM]-101	ST123	101	35	YEAR	FEMALE	LAB1	1	ST123	101	LAB1	3.8
[DM]-102	ST123	102	9	YEAR	FEMALE	LAB1	1	ST123	102	LAB1	4.8
[DM]-103	ST123	103	21	YEAR	MALE	LAB1	1	ST123	103	LAB1	5.3
[DM]-104	ST123	104	28	YEAR	MALE	LAB1	1	ST123	104	LAB1	3.2
[DM]-105	ST123	105	7	YEAR	MALE	LAB1	1	ST123	105	LAB1	4.5
[DM]-106	ST123	106	2	MONTH	FEMALE	LAB1	1	ST123	106	LAB1	4.6
LAB1_SUM											
DMESSKEY	STUDYID	SUBJID	AGE	AGE_U	GENDER	SITE	VISITNUM	HEMOGLOBIN	HEMATOCRIT	DM_SKEY	LAB1_SKEY
[LAB1_SUM]-101-1-LAB1	ST123	101	35	YEAR	FEMALE	LAB1	1	ST123	101	[DM]-101	[LAB1]-101-1-LAB1
[LAB1_SUM]-102-1-LAB1	ST123	102	9	YEAR	FEMALE	LAB1	1	ST123	102	[DM]-102	[LAB1]-102-1-LAB1
[LAB1_SUM]-103-1-LAB1	ST123	103	21	YEAR	MALE	LAB1	1	ST123	103	[DM]-103	[LAB1]-103-1-LAB1
[LAB1_SUM]-104-1-LAB1	ST123	104	28	YEAR	MALE	LAB1	1	ST123	104	[DM]-104	[LAB1]-104-1-LAB1
[LAB1_SUM]-105-1-LAB1	ST123	105	7	YEAR	MALE	LAB1	1	ST123	105	[DM]-105	[LAB1]-105-1-LAB1
[LAB1_SUM]-106-1-LAB1	ST123	106	2	MONTH	FEMALE	LAB1	1	ST123	106	[DM]-106	[LAB1]-106-1-LAB1
DME SKEYS MAP											
TRG_TAB	SRC_TAB	SKEY_COLNAME	TRG_SKEY_VALUE	SRC_SKEY_VALUE							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-101-1-LAB1	[DM]-101							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-101-1-LAB1	[LAB1]-101-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-102-1-LAB1	[DM]-102							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-102-1-LAB1	[LAB1]-102-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-103-1-LAB1	[DM]-103							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-103-1-LAB1	[LAB1]-103-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-104-1-LAB1	[DM]-104							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-104-1-LAB1	[LAB1]-104-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-105-1-LAB1	[DM]-105							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-105-1-LAB1	[LAB1]-105-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-106-1-LAB1	[DM]-106							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-106-1-LAB1	[LAB1]-106-1-LAB1							
...											

DME CONTEXT MAP											
CTXT_MAP_ID	STUDY	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	SRC_TAB_REC_IDX	FILTER	COL_FILTER	SRC_SKEY_COLNAME
2	ST123	1	MAP	DM	STUDYID	LAB1_SUM	STUDYID	1			DM_SKEY
3	ST123	1	MAP	DM	SUBJID	LAB1_SUM	SUBJID	1			DM_SKEY
4	ST123	1	MAP	DM	AGE	LAB1_SUM	AGE	1			DM_SKEY
5	ST123	1	MAP	DM	AGE_U	LAB1_SUM	AGE_U	1			DM_SKEY
6	ST123	1	MAP	DM	GENDER	LAB1_SUM	GENDER	1			DM_SKEY
7	ST123	2	MAP	LAB1	VISITNUM	LAB1_SUM	VISITNUM	1			LAB1_SKEY
8	ST123	2	MAP	LAB1	SITE	LAB1_SUM	SITE	1			LAB1_SKEY
9	ST123	2	MAP	LAB1	RBC	LAB1_SUM	RBC	1			LAB1_SKEY
10	ST123	2	MAP	LAB1	HEMOGLOBIN	LAB1_SUM	HEMOGLOBIN	1			LAB1_SKEY
11	ST123	2	MAP	LAB1	HEMATOCRIT	LAB1_SUM	HEMATOCRIT	1			LAB1_SKEY
...											
DM											
DMESSKEY	STUDYID	SUBJID	AGE	AGE_U	GENDER	SITE	VISITNUM	HEMOGLOBIN	HEMATOCRIT	RBC	HEMOGLOBIN
[DM]-101	ST123	101	35	YEAR	FEMALE	LAB1	1	ST123	101	LAB1	3.8
[DM]-102	ST123	102	9	YEAR	FEMALE	LAB1	1	ST123	102	LAB1	4.8
[DM]-103	ST123	103	21	YEAR	MALE	LAB1	1	ST123	103	LAB1	5.3
[DM]-104	ST123	104	28	YEAR	MALE	LAB1	1	ST123	104	LAB1	3.2
[DM]-105	ST123	105	7	YEAR	MALE	LAB1	1	ST123	105	LAB1	4.5
[DM]-106	ST123	106	2	MONTH	FEMALE	LAB1	1	ST123	106	LAB1	4.6
LAB1_SUM											
DMESSKEY	STUDYID	SUBJID	AGE	AGE_U	GENDER	SITE	VISITNUM	HEMOGLOBIN	HEMATOCRIT	DM_SKEY	LAB1_SKEY
[LAB1_SUM]-101-1-LAB1	ST123	101	35	YEAR	FEMALE	LAB1	1	ST123	101	[DM]-101	[LAB1]-101-1-LAB1
[LAB1_SUM]-102-1-LAB1	ST123	102	9	YEAR	FEMALE	LAB1	1	ST123	102	[DM]-102	[LAB1]-102-1-LAB1
[LAB1_SUM]-103-1-LAB1	ST123	103	21	YEAR	MALE	LAB1	1	ST123	103	[DM]-103	[LAB1]-103-1-LAB1
[LAB1_SUM]-104-1-LAB1	ST123	104	28	YEAR	MALE	LAB1	1	ST123	104	[DM]-104	[LAB1]-104-1-LAB1
[LAB1_SUM]-105-1-LAB1	ST123	105	7	YEAR	MALE	LAB1	1	ST123	105	[DM]-105	[LAB1]-105-1-LAB1
[LAB1_SUM]-106-1-LAB1	ST123	106	2	MONTH	FEMALE	LAB1	1	ST123	106	[DM]-106	[LAB1]-106-1-LAB1
DME SKEYS MAP											
TRG_TAB	SRC_TAB	SKEY_COLNAME	TRG_SKEY_VALUE	SRC_SKEY_VALUE							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-101-1-LAB1	[DM]-101							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-101-1-LAB1	[LAB1]-101-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-102-1-LAB1	[DM]-102							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-102-1-LAB1	[LAB1]-102-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-103-1-LAB1	[DM]-103							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-103-1-LAB1	[LAB1]-103-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-104-1-LAB1	[DM]-104							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-104-1-LAB1	[LAB1]-104-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-105-1-LAB1	[DM]-105							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-105-1-LAB1	[LAB1]-105-1-LAB1							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-106-1-LAB1	[DM]-106							
LAB_SUM	LAB1	LAB1_SKEY	[LAB1_SUM]-106-1-LAB1	[LAB1]-106-1-LAB1							
...											

DME CONTEXT MAP											
CTXT_MAP_ID	STUDY	TMAP	TMAP_TYPE	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	SRC_TAB_REC_IDX	FILTER	COL_FILTER	SRC_SKEY_COLNAME
2	ST123	1	MAP	DM	STUDYID	LAB1_SUM	STUDYID	1			DM_SKEY
3	ST123	1	MAP	DM	SUBJID	LAB1_SUM	SUBJID	1			DM_SKEY
4	ST123	1	MAP	DM	AGE	LAB1_SUM	AGE	1			DM_SKEY
5	ST123	1	MAP	DM	AGE_U	LAB1_SUM	AGE_U	1			DM_SKEY
6	ST123	1	MAP	DM	GENDER	LAB1_SUM	GENDER	1			DM_SKEY
7	ST123	2	MAP	LAB1	VISITNUM	LAB1_SUM	VISITNUM	1			LAB1_SKEY
8	ST123	2	MAP	LAB1	SITE	LAB1_SUM	SITE	1			LAB1_SKEY
9	ST123	2	MAP	LAB1	RBC	LAB1_SUM	RBC	1			LAB1_SKEY
10	ST123	2	MAP	LAB1	HEMOGLOBIN	LAB1_SUM	HEMOGLOBIN	1			LAB1_SKEY
11	ST123	2	MAP	LAB1	HEMATOCRIT	LAB1_SUM	HEMATOCRIT	1			LAB1_SKEY
...											
DM											
DMESSKEY	STUDYID	SUBJID	AGE	AGE_U	GENDER	SITE	VISITNUM	HEMOGLOBIN	HEMATOCRIT	RBC	HEMOGLOBIN
[DM]-101	ST123	101	35	YEAR	FEMALE	LAB1	1	ST123	101	LAB1	3.8
[DM]-102	ST123	102	9	YEAR	FEMALE	LAB1	1	ST123	102	LAB1	4.8
[DM]-103	ST123	103	21	YEAR	MALE	LAB1	1	ST123	103	LAB1	5.3
[DM]-104	ST123	104	28	YEAR	MALE	LAB1	1	ST123	104	LAB1	3.2
[DM]-105	ST123	105	7	YEAR	MALE	LAB1	1	ST123	105	LAB1	4.5
[DM]-106	ST123	106	2	MONTH	FEMALE	LAB1	1	ST123	106	LAB1	4.6
LAB1_SUM											
DMESSKEY	STUDYID	SUBJID	AGE	AGE_U	GENDER	SITE	VISITNUM	HEMOGLOBIN	HEMATOCRIT	DM_SKEY	LAB1_SKEY
[LAB1_SUM]-101-1-LAB1	ST123	101	35	YEAR	FEMALE	LAB1	1	ST123	101	[DM]-101	[LAB1]-101-1-LAB1
[LAB1_SUM]-102-1-LAB1	ST123	102	9	YEAR	FEMALE	LAB1	1	ST123	102	[DM]-102	[LAB1]-102-1-LAB1
[LAB1_SUM]-103-1-LAB1	ST123	103	21	YEAR	MALE	LAB1	1	ST123	103	[DM]-103	[LAB1]-103-1-LAB1
[LAB1_SUM]-104-1-LAB1	ST123	104	28	YEAR	MALE	LAB1	1	ST123	104	[DM]-104	[LAB1]-104-1-LAB1
[LAB1_SUM]-105-1-LAB1	ST123	105	7	YEAR	MALE	LAB1	1	ST123	105	[DM]-105	[LAB1]-105-1-LAB1
[LAB1_SUM]-106-1-LAB1	ST123	106	2	MONTH	FEMALE	LAB1	1	ST123	106	[DM]-106	[LAB1]-106-1-LAB1
DME SKEYS MAP											
TRG_TAB	SRC_TAB	SKEY_COLNAME	TRG_SKEY_VALUE	SRC_SKEY_VALUE							
LAB_SUM	DM	DM_SKEY	[LAB1_SUM]-101-1-LAB1	[DM]-101							

图16

DME CONTEXT MAP												
CTXT_MAP_ID	STUDY	TMAP	TMAP	SRC_TAB	SRC_COL	TRG_TAB	TRG_COL	SRC_TAB	REC_NDX	FILTER_COOL	FILTER_VAL	SRC_SKEY_COLUMN
2	ST123	1	MAP	LAB	STUDYID	LAB_SUM	STUDYID	1				LAB_SKEY
3	ST123	1	MAP	LAB	SUBJID	LAB_SUM	SUBJID	1				LAB_SKEY
4	ST123	1	MAP	LAB	VISITNUM	LAB_SUM	VISITNUM	1				LAB_SKEY
5	ST123	1	MAP	LAB	SITE	LAB_SUM	SITE	1				LAB_SKEY
6	ST123	1	MAP	LAB	AGE	LAB_SUM	AGE	1				LAB_SKEY
7	ST123	1	MAP	LAB	AGE_U	LAB_SUM	AGE_U	1				LAB_SKEY
8	ST123	1	MAP	LAB	GENDER	LAB_SUM	GENDER	1				LAB_SKEY
9	ST123	1	MAP	LAB	RBC	LAB_SUM	RBC	1				LAB_SKEY
10	ST123	2	MAP	NRANGE	TEST_U	LAB_SUM	RBC_U	1				NRANGE_SKEY
11	ST123	3	CASE	NRANGE	MALE	LAB_SUM	RBC_NR	1		CASE_FVAL	MALE	NRANGE_SKEY
12	ST123	3	CASE	NRANGE	FEMALE	LAB_SUM	RBC_NR	1		CASE_FVAL	FEMALE	NRANGE_SKEY
13	ST123	3	CASE	NRANGE	CHILD	LAB_SUM	RBC_NR	1		CASE_FVAL	CHILD	NRANGE_SKEY
14	ST123	3	CASE	NRANGE	NEWBORN	LAB_SUM	RBC_NR	1		CASE_FVAL	NEWBORN	NRANGE_SKEY

NRANGE									
DMESSKEY	STUDYID	SITE	TEST	TEST_U	MALE	FEMALE	CHILD	NEWBORN	
[NRANGE]-LAB1-RBC	ST123	LAB1	RBC	Meas/mcL	4.7-6.1	4.2-5.4	4.8-4.8	4.8-4.8	
[NRANGE]-LAB1-HEMOGLOBIN	ST123	LAB1	HEMOGLOBIN	g/dL	13-18	12-16	11.5-13.5	17-22	
[NRANGE]-LAB1-HEMATOCRIT	ST123	LAB1	HEMATOCRIT	%	38-51	34-46	33-40	50-82	

LAB									
DMESSKEY	STUDYID	SUBJID	VISITNUM	SITE	AGE	AGE_U	GENDER	RBC	
[LAB]-101-1	ST123	101	1	LAB1	35	YEAR	FEMALE	3.8	
[LAB]-102-1	ST123	102	1	LAB1	9	YEAR	FEMALE	4.8	
[LAB]-103-1	ST123	103	1	LAB1	21	YEAR	MALE	5.3	
[LAB]-104-1	ST123	104	1	LAB1	28	YEAR	MALE	3.2	
[LAB]-105-1	ST123	105	1	LAB1	7	YEAR	MALE	45	
[LAB]-106-1	ST123	106	1	LAB1	2	MONTH	FEMALE	4.8	

DME_SKEYS MAP									
TRG_TAB	SRC_TAB	FILTER	VALUE	SKEY_COLUMN	TRG_SKEY	VALUE	SRC_SKEY	VALUE	
LAB_SUM	NRANGE	FEMALE		NRANGE_SKEY	[LAB_SUM]-101-1		[NRANGE]-LAB1-RBC		
LAB_SUM	NRANGE	CHILD		NRANGE_SKEY	[LAB_SUM]-102-1		[NRANGE]-LAB1-RBC		
LAB_SUM	NRANGE	MALE		NRANGE_SKEY	[LAB_SUM]-103-1		[NRANGE]-LAB1-RBC		
LAB_SUM	NRANGE	MALE		NRANGE_SKEY	[LAB_SUM]-104-1		[NRANGE]-LAB1-RBC		
LAB_SUM	NRANGE	CHILD		NRANGE_SKEY	[LAB_SUM]-105-1		[NRANGE]-LAB1-RBC		
LAB_SUM	NRANGE	NEWBORN		NRANGE_SKEY	[LAB_SUM]-106-1		[NRANGE]-LAB1-RBC		

LAB1_SUM									
DMESSKEY	STUDYID	SUBJID	VISITNUM	SITE	AGE	AGE_U	GENDER	RBC	
[LAB_SUM]-101-1	ST123	101	1	LAB1	35	YEAR	FEMALE	3.8	4.2-5.4
[LAB_SUM]-102-1	ST123	102	1	LAB1	9	YEAR	FEMALE	4.8	4.8-4.8
[LAB_SUM]-103-1	ST123	103	1	LAB1	21	YEAR	MALE	5.3	4.7-6.1
[LAB_SUM]-104-1	ST123	104	1	LAB1	28	YEAR	MALE	3.2	4.7-6.1
[LAB_SUM]-105-1	ST123	105	1	LAB1	7	YEAR	MALE	45	4.6-4.8
[LAB_SUM]-106-1	ST123	106	1	LAB1	2	MONTH	FEMALE	4.8	4.5-4.8

图17

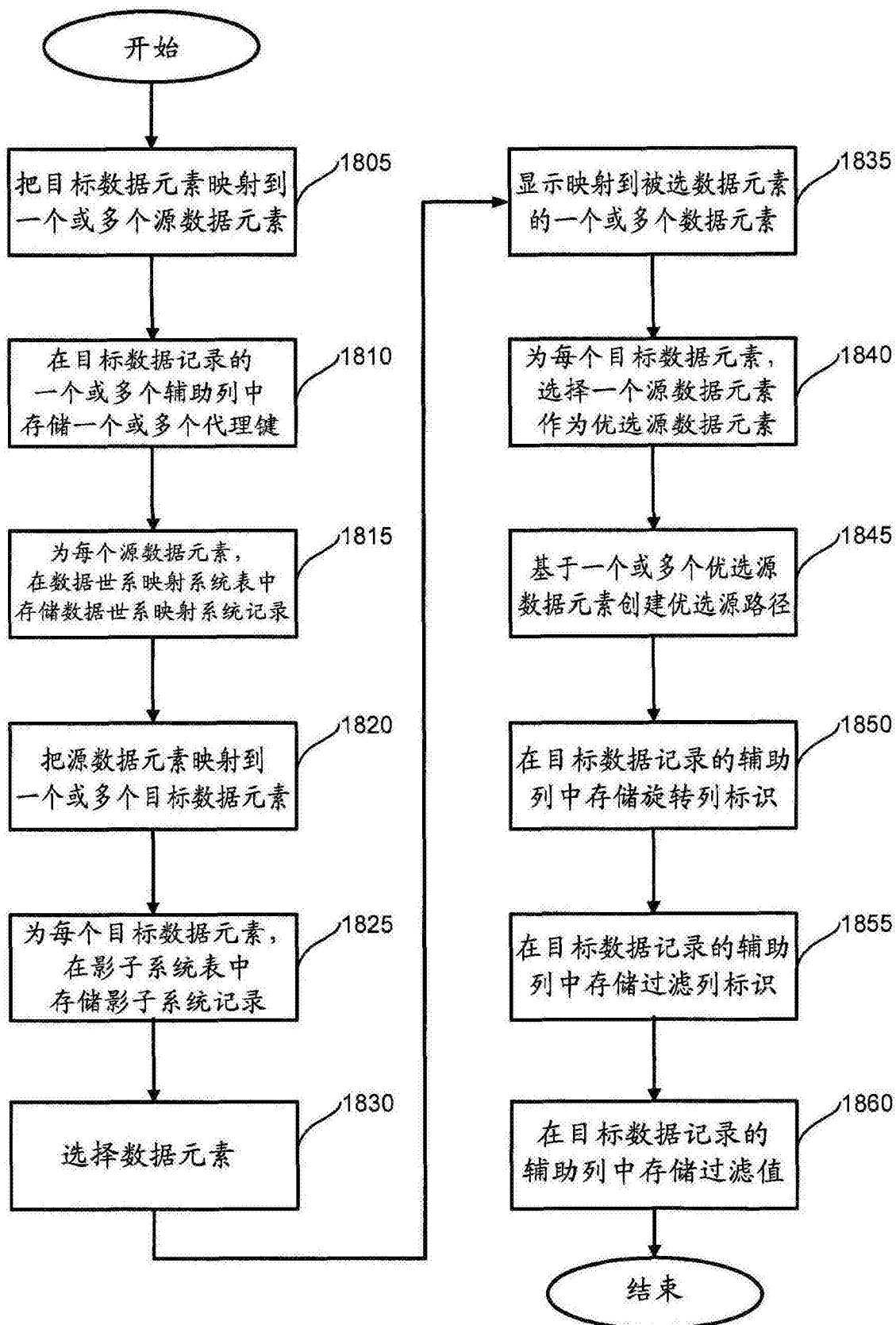


图18

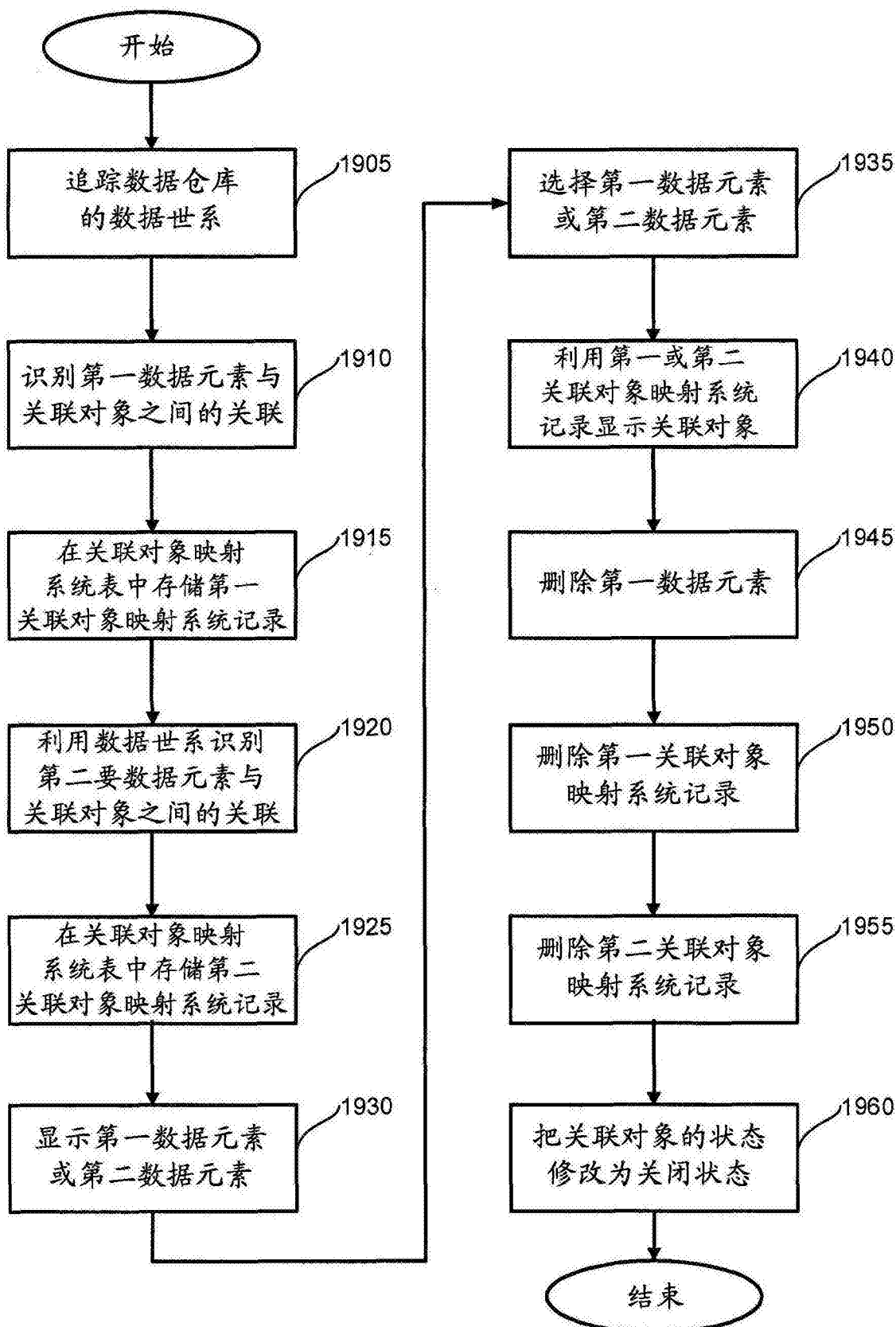


图19