

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第5847585号  
(P5847585)

(45) 発行日 平成28年1月27日(2016.1.27)

(24) 登録日 平成27年12月4日(2015.12.4)

(51) Int.Cl.

F 1

**G06T 1/00 (2006.01)**  
**G06F 3/0481 (2013.01)**  
**G06F 12/00 (2006.01)**  
**G06T 11/00 (2006.01)**

GO6T 1/00  
 GO6F 3/048 657A  
 GO6F 12/00 520J  
 GO6T 11/00

請求項の数 20 (全 17 頁)

(21) 出願番号 特願2011-539651 (P2011-539651)  
 (86) (22) 出願日 平成21年12月2日 (2009.12.2)  
 (65) 公表番号 特表2012-510688 (P2012-510688A)  
 (43) 公表日 平成24年5月10日 (2012.5.10)  
 (86) 國際出願番号 PCT/US2009/066390  
 (87) 國際公開番号 WO2010/065623  
 (87) 國際公開日 平成22年6月10日 (2010.6.10)  
 審査請求日 平成24年11月29日 (2012.11.29)  
 審判番号 不服2014-23404 (P2014-23404/J1)  
 審判請求日 平成26年11月17日 (2014.11.17)  
 (31) 優先権主張番号 61/119,201  
 (32) 優先日 平成20年12月2日 (2008.12.2)  
 (33) 優先権主張國 米国(US)

(73) 特許権者 509123208  
 アビニシオ テクノロジー エルエルシー  
 アメリカ合衆国 O 2 4 2 1 マサチュー  
 セツツ州 レキシントン スプリング ス  
 トリート 201  
 (74) 代理人 100079108  
 弁理士 稲葉 良幸  
 (74) 代理人 100109346  
 弁理士 大貫 敏史  
 (74) 代理人 100117189  
 弁理士 江口 昭彦  
 (74) 代理人 100134120  
 弁理士 内藤 和彦

最終頁に続く

(54) 【発明の名称】データ要素とデータ要素の属性のグラフ表現との関係の視覚化

## (57) 【特許請求の範囲】

## 【請求項 1】

メタデータをデータ格納システムに格納すること、

データ及びメタデータオブジェクトにつながるデータ系統の情報を利用する処理を表す  
前記メタデータオブジェクトを前記データ格納システムに格納すること、

前記データ格納システムに格納された複数のメタデータオブジェクトの各々の1つ以上の  
 特徴を特定する要約データを計算することであって、無効な値を有するデータの比率、  
 及び最近前記データが更新されたか又は前記メタデータに対して修正されたかの指示子か  
 らなるグループから選択された少なくとも1つの特徴を計算することを含み、

所与のメタデータオブジェクトに関連付けて同所与のメタデータオブジェクトを特徴付  
 ける前記要約データを格納すること、及び、

前記所与のメタデータオブジェクトの少なくとも3つの性状を説明する線図の視覚的表  
 現を発生すること、を含む方法であって、

前記少なくとも3つの性状は、

前記所与のメタデータオブジェクトと前記データ系統に従う他のメタデータオブジェクト  
 との間の関係であって前記他のメタデータオブジェクトを表す複数のノードと前記所与  
 のメタデータオブジェクトを表すノードとの間に表される関係、

前記所与のメタデータオブジェクトのコンテクストであって同コンテクストに基くグル  
 ープ内に示される前記所与のメタデータオブジェクトを伴うコンテクスト、及び、

前記所与のメタデータオブジェクトを特徴付ける前記要約データにより特定される特徴

10

20

であって品質、新しさ、又はレコード情報のソースを含む特徴、を含み、  
前記視覚的表現を発生することは、

前記特徴の表現を前記所与のメタデータオブジェクトを表すノードの近傍に重ねること  
を含む方法。

**【請求項 2】**

前記表現は、同表現を分類する凡例と関連付けられる請求項 1 に記載の方法。

**【請求項 3】**

カーソルを前記視覚的表現の上でうろつかせることにより前記表現に関係した情報を含  
むウィンドウを発生する請求項 1 に記載の方法。

**【請求項 4】**

前記表現は、ユーザによって選択可能な特徴を表す請求項 1 に記載の方法。 10

**【請求項 5】**

メタデータをデータ格納システムに格納する手段と、

データ及びメタデータオブジェクトにつながるデータ系統の情報を利用する処理を表す  
前記メタデータオブジェクトを前記データ格納システムに格納する手段と、

前記データ格納システムに格納された複数のメタデータオブジェクトの各々の 1 つ以上の  
特徴を特定する要約データを計算する手段であって、無効な値を有するデータの比率、  
及び最近前記データが更新されたか又は前記メタデータに対して修正されたかの指示子か  
らなるグループから選択された少なくとも 1 つの特徴を計算することを含む、計算する手  
段と、 20

所与のメタデータオブジェクトに関連付けて同所与のメタデータオブジェクトを特徴付  
ける前記要約データを格納する手段と、

前記所与のメタデータオブジェクトの少なくとも 3 つの性状を説明する線図の視覚的表  
現を発生すること、を含むシステムであって、

前記少なくとも 3 つの性状は、

前記所与のメタデータオブジェクトと前記データ系統に従う他のメタデータオブジェクト  
との間の関係であって前記他のメタデータオブジェクトを表す複数のノードと前記所与  
のメタデータオブジェクトを表すノードとの間に表される関係、

前記所与のメタデータオブジェクトのコンテクストであって同コンテクストに基くグル  
ープ内に示される前記所与のメタデータオブジェクトを伴うコンテクスト、及び、 30

前記所与のメタデータオブジェクトを特徴付ける前記要約データにより特定される特徴  
であって品質、新しさ、又はレコード情報のソースを含む特徴、を含み、

前記視覚的表現を発生することは、

前記特徴の表現を前記所与のメタデータオブジェクトを表すノードの近傍に重ねること  
を含むシステム。

**【請求項 6】**

メタデータをデータ格納システムに格納し、

データ及びメタデータオブジェクトにつながるデータ系統の情報を利用する処理を表す  
前記メタデータオブジェクトを前記データ格納システムに格納し、

前記データ格納システムに格納された複数のメタデータオブジェクトの各々の 1 つ以上の  
特徴を特定する要約データを計算することであって、無効な値を有するデータの比率、  
及び最近前記データが更新されたか又は前記メタデータに対して修正されたかの指示子か  
らなるグループから選択された少なくとも 1 つの特徴を計算することを含み、 40

所与のメタデータオブジェクトに関連付けて同所与のメタデータオブジェクトを特徴付  
ける前記要約データを格納し、

前記所与のメタデータオブジェクトの少なくとも 3 つの性状を説明する線図の視覚的表  
現を発生するように構成されたプロセッサとメモリとを含むコンピュータシステムであっ  
て、

前記少なくとも 3 つの性状は、

前記所与のメタデータオブジェクトと前記データ系統に従う他のメタデータオブジェク 50

トとの間の関係であって前記他のメタデータオブジェクトを表す複数のノードと前記所与のメタデータオブジェクトを表すノードとの間に表される関係、

前記所与のメタデータオブジェクトのコンテクストであって同コンテクストに基くグループ内に示される前記所与のメタデータオブジェクトを伴うコンテクスト、及び、

前記所与のメタデータオブジェクトを特徴付ける前記要約データにより特定される特徴であって品質、新しさ、又はレコード情報のソースを含む特徴、を含み、

前記視覚的表現を発生することは、

前記特徴の表現を前記所与のメタデータオブジェクトを表すノードの近傍に重ねることを含むコンピュータシステム。

#### 【請求項 7】

10

コンピュータプログラムを格納するコンピュータ読み出し可能媒体であって、実行可能な命令を含む前記コンピュータプログラムは、コンピュータに、

メタデータをデータ格納システムに格納させ、

データ及びメタデータオブジェクトにつながるデータ系統の情報を利用する処理を表す前記メタデータオブジェクトを前記データ格納システムに格納させ、

前記データ格納システムに格納された複数のメタデータオブジェクトの各々の 1 つ以上の特徴を特定する要約データを計算させることであって、無効な値を有するデータの比率、及び最近前記データが更新されたか又は前記メタデータに対して修正されたかの指示子からなるグループから選択された少なくとも 1 つの特徴を計算させることを含み、

所与のメタデータオブジェクトに関連付けて同所与のメタデータオブジェクトを特徴付ける前記要約データを格納させ、

20

前記所与のメタデータオブジェクトの少なくとも 3 つの性状を説明する線図の視覚的表現を発生させる命令を含み、

前記少なくとも 3 つの性状は、

前記所与のメタデータオブジェクトと前記データ系統に従う他のメタデータオブジェクトとの間の関係であって前記他のメタデータオブジェクトを表す複数のノードと前記所与のメタデータオブジェクトを表すノードとの間に表される関係、

前記所与のメタデータオブジェクトのコンテクストであって同コンテクストに基くグループ内に示される前記所与のメタデータオブジェクトを伴うコンテクスト、及び、

前記所与のメタデータオブジェクトを特徴付ける前記要約データにより特定される特徴であって品質、新しさ、又はレコード情報のソースを含む特徴、を含み、

30

前記視覚的表現を発生させることは、

前記特徴の表現を前記所与のメタデータオブジェクトを表すノードの近傍に重ねることを含む、コンピュータ読み出し可能媒体。

#### 【請求項 8】

請求項 1 に記載の方法であって、さらに、ユーザからの選択を受けて前記要約データによって特定された複数の特徴のうちのいずれが、その表現が前記所与のメタデータオブジェクトを表すノードの近傍に重ねられる特徴であるかを決定することを含む方法。

#### 【請求項 9】

前記表現は、同表現を分類する凡例と関連付けられる請求項 5 に記載のシステム。

40

#### 【請求項 10】

カーソルを前記視覚的表現の上でうろつかせることにより前記表現に関係した情報を含むウィンドウを発生する請求項 5 に記載のシステム。

#### 【請求項 11】

前記表現は、ユーザによって選択可能な特徴を表す請求項 5 に記載のシステム。

#### 【請求項 12】

前記要約データによって特定された複数の特徴のうちのいずれが、該表現が前記所与のメタデータオブジェクトを表すノードの近傍に重ねられる特徴であるかを決定するために、ユーザからの選択を受ける手段をさらに含む、請求項 5 に記載のシステム。

#### 【請求項 13】

50

前記表現は、同表現を分類する凡例と関連付けられる請求項6に記載のコンピュータシステム。

【請求項 14】

カーソルを前記視覚的表現の上でうろつかせることにより前記表現に関係した情報を含むウィンドウを発生する請求項6に記載のコンピュータシステム。

【請求項 15】

前記表現は、ユーザによって選択可能な特徴を表す請求項6に記載のコンピュータシステム。

【請求項 16】

前記要約データによって特定された複数の特徴のうちのいずれが、該表現が前記所与のメタデータオブジェクトを表すノードの近傍に重ねられる特徴であるかを決定するために、ユーザからの選択を受けるように、さらに構成される、請求項6に記載のコンピュータシステム。10

【請求項 17】

前記表現は、同表現を分類する凡例と関連付けられる請求項7に記載のコンピュータ読み出し可能媒体。

【請求項 18】

カーソルを前記視覚的表現の上でうろつかせることにより前記表現に関係した情報を含むウィンドウを発生する請求項7に記載のコンピュータ読み出し可能媒体。

【請求項 19】

前記表現は、ユーザによって選択可能な特徴を表す請求項7に記載のコンピュータ読み出し可能媒体。

【請求項 20】

コンピュータに、前記要約データによって特定された複数の特徴のうちのいずれが、該表現が前記所与のメタデータオブジェクトを表すノードの近傍に重ねられる特徴であるかを決定するために、ユーザからの選択を受けさせる実行可能命令をさらに含む、請求項7に記載のコンピュータ読み出し可能媒体。20

【発明の詳細な説明】

【技術分野】

【0001】

本発明はデータ要素とデータ要素の属性のグラフ表現との関係を視覚化することに関する。

【背景技術】

【0002】

企業は複雑なデータ処理システム、たとえば、データ倉庫保管システム、顧客関係管理システム及びデータマイニング（採掘、mining）等を用いてデータを扱う。多くのデータ処理システムにおいて、データは、多くの異なるデータソース、たとえばデータベースファイル、運用（オペレーションナル）システム、フラットファイル、インターネット等から中央保管所（セントラル・レポジトリ、central repository）に引き出される。しばしば、データはデータシステムにロードされる前に変換される。変換には、クリーンシング（cleansing）、統合及び抽出を含むことができる。データ、そのデータのソース及びデータシステムに格納されたデータに対して発生した変換を追跡するには、メタデータを用いることができる。メタデータ（ときおり、「データについてのデータ」と呼ばれる。）は、データの属性、フォーマット、起源（origin）、履歴、相互関係等を記述するデータである。メタデータ管理は複雑なデータ処理システムにおいて中心的役割を果たすことができる。40

【0003】

場合により、データベースユーザは、データが異なるデータソースからどのように引き出されたかについての調査を希望する。たとえば、データベースユーザは、データセット又はデータオブジェクトがどのように発生されたのか、或いは、データセット又はデータ

10

20

30

40

50

オブジェクトがどのソースからインポートされたのか、を知ることを望む。データセットが引き出されたソースまで遡ってそのデータセットを追跡することは、データ系統 (lineage) 追跡 (又は「上流データ系統追跡」) と称呼される。場合により、データベースユーザは、あるデータセットがどのように使用されたか (「下流データ系統追跡」又は「インパクト解析」と称呼される。)、たとえば、どのアプリケーションが所与のデータセットを読んだか、についての調査を希望する。また、データベースユーザは、データセットが他のデータセットにどのように関係しているかを知ることに興味があるであろう。たとえば、ユーザは、データセットが修正されたか否か、又は、どのテーブルが影響を受けたか、を知ることを欲する場合がある。

#### 【発明の概要】

10

#### 【0004】

##### <関連出願の相互参照>

本願は、2008年12月2日出願の米国特許出願第61/119,201の優先権を主張するものであり、この出願はここに引用されて組み込まれる。

概略において、方法は、メタデータをデータ格納システムに格納することを含む。そのデータ格納システムに格納された複数のメタデータオブジェクトの各々についての1つ以上の特徴を特定 (識別) する要約データ (summary data) が計算され、所与のメタデータオブジェクトに関連付けてこの所与のメタデータオブジェクトを特徴付ける要約データが格納される。メタデータオブジェクトの各々を表すノード及びそのノード間の関係を含む線図 (ダイアグラム、diagram) の視覚的表現が発生される。その視覚的表現を発生することは、所与のメタデータオブジェクトを特徴付ける前記要約データによって特定される特徴の表現を、その所与のメタデータオブジェクトを表すノードの近傍に重ねる (superimpose) ことを含む。

20

#### 【0005】

複数の態様は、以下の1つ以上の特徴を含むことができる。前記表現はメタデータオブジェクトの質 (属性) を表す。前記表現はメタデータオブジェクトが最近更新されたか否かを表す。前記表現はメタデータオブジェクトが最後に更新されたソースを表す。前記表現はこの表現を分類する凡例 (legend) に関連付けられる。カーソルを視覚的な表現上でうろつかせることにより前記表現に関係した情報を含むウィンドウを発生する。前記表現はユーザによって選択可能な特徴を表す。

30

#### 【0006】

全体的な態様において、システムは、メタデータをデータ格納システムに格納する手段と、前記データ格納システムに格納された複数のメタデータオブジェクトの各々の1つ以上の特徴を特定する要約データを計算する手段と、を含む。また、システムは、所与のメタデータオブジェクトに関連付けてこの所与のメタデータオブジェクトを特徴付ける要約データを格納する手段と、対応するメタデータオブジェクトを表す (複数の) ノード及びそれらのノード間の関係を含む線図 (ダイアグラム) の視覚的表現を発生する手段と、を含む。その視覚的表現を発生することは、所与のメタデータオブジェクトを特徴付ける前記要約データによって特定される特徴の表現を、その所与のメタデータオブジェクトを表すノードの近傍に重ねる (superimpose) ことを含む。

40

#### 【0007】

全体的な態様において、コンピュータシステムは、メタデータをデータ格納システムに格納し、且つ、前記データ格納システムに格納された複数のメタデータオブジェクトの各々の1つ以上の特徴を特定する要約データを計算するように構成される。所与のメタデータオブジェクトに関連付けてこの所与のメタデータオブジェクトを特徴付ける前記要約データが格納され、メタデータオブジェクトの各々を表す (複数の) ノード及びそれらのノード間の関係を含む線図の視覚的表現が発生される。前記視覚的表現を発生することは、所与のメタデータオブジェクトを特徴付ける前記要約データによって特定される特徴の表現を、その所与のメタデータオブジェクトを表すノードの近傍に重ねることを含む。

#### 【0008】

50

全体的な態様において、コンピュータ読み出し可能媒体は、コンピュータプログラムを格納し、そのコンピュータプログラムは、メタデータをデータ格納システムに格納せんようにコンピュータシステムを動作させる命令を含む。前記データ格納システムに格納された複数のメタデータオブジェクトの各々の1つ以上の特徴を特定する要約データが計算され、所与のメタデータオブジェクトに関連付けてこの所与のメタデータオブジェクトを特徴付ける前記要約データが格納される。メタデータオブジェクトの各々を表すノード及びそのノード間の関係を含む線図の視覚的表現が発生される。その視覚的表現を発生することは、所与のメタデータオブジェクトを特徴付ける前記要約データによって特定される特徴の表現を、その所与のメタデータオブジェクトを表すノードの近傍に重ねることを含む。

10

## 【0009】

複数の態様は以下の1つ以上の利点を有する。

## 【0010】

本システムはユーザが(複数の)オブジェクト間の関係を視覚化することを可能にし、更に、ユーザがコンテクストの設定(contextual setting)においてオブジェクトのある属性を見ることを可能にする。メタデータと連携することにより、ユーザは、ある動作を行う前にオブジェクトの起源(発生元)を理解することができる。ユーザは、任意の特定のオブジェクトの操作によってどのオブジェクトが影響されるかを知ることができる。また、ユーザは、所与の複数のオブジェクトの間の関係が明白に示される環境において、そのオブジェクトの属性を見ることができます。

20

## 【図面の簡単な説明】

## 【0011】

【図1】計算システムのブロック線図である。

【図2A】データのノード間の関係を示す線図である。

【図2B】データのノード間の関係を示す線図である。

【図2C】データのノード間の関係を示す線図である。

【図2D】データのノード間の関係を示す線図である。

【図2E】データのノード間の関係を示す線図である。

【図3】データのノード上に重ねられた図形のオーバーレイ(overlay)を示す線図である。

30

【図4】データのノード上に重ね合わされた図形のオーバーレイを示す線図である。

【図5】データのノード上に重ね合わされた図形のオーバーレイを示す線図である。

## 【発明を実施するための形態】

## 【0012】

本システムは、ユーザが、幅広い種類のデータ格納システムに格納された(複数の)オブジェクトの間の関係を視覚化することを可能にする。その(複数の)オブジェクトの間の関係は、それらのデータオブジェクトが使用されるアプリケーションに適切な種々の従属性(dependencies)及び/または関連(associations)を表すことができる。このような技術が使用され得るこの種のシステムの一例として、そのオブジェクトがグラフに基づく(グラフベースの)計算環境の要素を表すシステムを説明する。

40

## 【0013】

図1はグラフに基づく計算の開発、実行及び管理のための計算システム100のパートの相互関係を示すブロック線図である。グラフに基づく計算は、構成要素(データファイル又は処理)を表すグラフにおける頂点を伴う有向グラフによって表現される「データフローグラフ」と、構成要素の間のデータの流れを表すグラフにおける有向リンクつまり「辺(edge)」と、を用いて実現される。グラフ開発環境(GDE)102は、実行可能なグラフを特定しあつグラフ構成要素用のパラメータを定義するユーザインターフェースを提供する。GDEは、たとえばAb Initioから入手可能なCO>OPERATING SYSTEM(登録商標)GDEとすることができます。GDE102は、保管所(repository)104及び並列オペレーティング環境106と通信する。また、保管所10

50

4 及び並列オペレーティング環境 106 には、ユーザインターフェースモジュール 108 及び実行部 (executive) 110 が接続されている。

#### 【0014】

幾つかの例において、保管所 104 は、ベースデータ格納部 (store) 105A 及びインターフェースデータ格納部 (store) 105B の両方を含む。ベースデータ格納部は、技術上のメタデータを記憶し、また、たとえばグラフ及び変換などのアプリケーションを、それらのアプリケーションに関連付けられるメタデータとともに含むことができる。技術上のメタデータを記憶することに加え、ベースデータ格納部は、従属性解析（たとえば、後に詳述するように、データ系統を計算すること）を含む種々の解析を実行することができ、或いは、そのような解析の結果を受け取って格納することができる。幾つかの例において、ベースデータ格納部 105A 及びインターフェースデータ格納部 105B は結合されて単一のデータ格納部として実現することができる。10

#### 【0015】

技術上のメタデータは開発者にとって種々の機能において役立つが、よりハイレベルのメタデータが解析されかつ操作される必要があるという多くの場合がある。この、よりハイレベルのメタデータは、時々「企業」メタデータまたは「ビジネス」メタデータと称呼されることがあり、データ解析においてしばしば有用である。ビジネスメタデータの幾つかの例は、どの従業員がデータに対して責任を有しているかを示すデータ管理職務 (stewardship) と、ファイル及びファイル内のフィールドに対するビジネス上の定義であるデータ辞書と、を含む。ビジネスメタデータはデータの技術上の記述を超えて、ベースデータ格納部 105A とは別の、たとえばインターフェースデータ格納部 105B のようなプラットフォーム上に格納することができる。20

#### 【0016】

インターフェースデータ格納部 105B は、ビジネスメタデータを格納する役目を主として果たす関係型データベース（リレーションナルデータベース）とすることができる。インターフェースデータ格納部は、ベースデータ格納部と通信してそのメタデータを引き出すことができ、更に、様々な他のソース（たとえばグラフ、スプレッドシート（マトリックス計算表）、論理モデル、データベースのテーブル、又は、増設の第3者のデータソース等）から情報を引き出すことができる。

#### 【0017】

幾つかの例において、ベースデータ格納部 105A は、拡張性のある（スケーラブルな）オブジェクト指向型データベースシステムであり、グラフに基づくアプリケーションの開発及び実行を支援し、且つ、グラフに基づくアプリケーションと他のシステム（たとえば、他のオペレーティングシステム）との間のメタデータの交換を支援するように設計されている。保管所 104 は、ドキュメンテーション、レコードフォーマット（たとえば、テーブル内のレコードのフィールド及びデータのタイプ）、変換関数、グラフ、ジョブ及び監視情報等を含む総ての種類のメタデータ用の格納システムである。また、保管所 104 は、計算システム 100 によって処理されるべき実際のデータ（外部のデータ格納部 112 に格納されたデータを含む。）を表すメタデータオブジェクトを格納する。種々のソースからメタデータをインポートして管理するための特徴を含む保管所の一例は、2008 年 12 月 2 日出願の同時係属米国仮特許出願第 61/119,148 号、発明の名称「データ保守システム」に記載されており、本明細書に引用され組み込まれる。類似の特徴は保管所 104 に組み込むことができる。3040

#### 【0018】

並列オペレーティング環境 106 は、GDE 102 において発生されたデータフローグラフの仕様を受理し、そのグラフによって規定された処理ロジック及びリソースに対応するコンピュータ命令を発生する。並列オペレーティング環境 106 は、次いで、一般に、複数のプロセッサ（同種である必要ない）上にてそれらの命令を実行する。適切な並列オペレーティング環境の一例は CO>OPERATING SYSTEM（登録商標）である。50

## 【0019】

ユーザインターフェースモジュール108は、保管所104の内容のウェブブラウザベースの表示を提供する。ユーザ103は、ユーザインターフェースモジュール108を用いて、オブジェクトを閲覧でき、新しいオブジェクトを作り出すことができ、現存するオブジェクトを変更することができ、アプリケーションのパラメータを特定することができ、更に、ジョブのスケジューリングをすることができる。ユーザインターフェース108は、ユーザが、保管所104に格納された「オブジェクト及びオブジェクトについての情報」を検索し且つ見るための書式ベース(forms-based)の探索画面を発生する。

## 【0020】

保管所104は、グラフの構成要素を含むグラフに基づくアプリケーションのためのメタデータオブジェクト、及び、計算グラフを構築するための他の関数(functional)オブジェクト、を含むメタデータを格納する。上述したように、保管所104のベースデータ格納部105Aに格納されたメタデータは、たとえば、「技術上の」メタデータ(たとえば、アプリケーションに関係するビジネス規則、レコードフォーマット及び実行統計)を含み、他方、インターフェースデータ格納部105Bはビジネスメタデータ、たとえば職務権限、役割及び責任のユーザが定義するドキュメンテーション、を含むことができる。

10

## 【0021】

メタデータオブジェクトの形式にて保管所104に格納された情報は、アプリケーション及びこれらのアプリケーションによって処理されるデータに関する種々の解析を可能にする。この情報のサブセットはインターフェースデータ格納部105Bに格納され得る。たとえば、さらに後に詳述するように、ユーザは、データ系統についての質問に対する回答を得ることができる(たとえば、所与の値はどこから来たのか?その出力値はどのように計算されたのか?どのアプリケーションがこのデータを作成し、このデータに依存するか?)。開発者は提案された変更の論理的結論を理解することができる(たとえば、この部分が変化したならば、他のどの部分が影響を受けるのか?このソースフォーマットが変化したならば、どのアプリケーションが影響を受けるのか?)。また、ユーザ/開発者は、技術上のメタデータ及びビジネスメタデータの両方に関係する質問(たとえば、このデータを発生し用いることに対してどのグループが責任を有するのか?このアプリケーションを最後に変更したのは誰か?彼らはどのような変更をしたのか?)に対する回答を得ることができる。

20

## 【0022】

また、保管所104は、格納されたメタデータオブジェクトの状態を追跡できる。保管所104に格納されたオブジェクトはバージョン化され、それにより、先週、先月又は昨年の時点における状態を調べることができ、この状態を今日の状態と比較することができる。保管所104は、ジョブ追跡、又は、実行情報を収集する。ジョブ追跡又は実行情報は、動向分析(たとえば、我々のデータがどの速度で成長しているか?)及び能力の計画(たとえば、そのアプリケーションが実行するのにどのくらいの時間がかかったか?アプリケーションはどのくらいの量のデータを処理したか、また、処理速度は?そのアプリケーションはどのリソースを消費したか?何時、別のサーバの追加を必要とするか?)を可能にする。

30

## 【0023】

ユーザは、格納されたメタデータに含まれる情報及び/または格納されたメタデータに関連付けられた情報をユーザインターフェースモジュール108を介して見ることができ(更に、オプションとしてその情報を編集できる)。メタデータビューア環境(メタデータを見る環境)は、ディスプレ上にユーザインターフェースモジュール108によって表されたアイコン及びアイコンのグルーピング(grouping)を含む種々の図形的表現を用いて種々のメタデータオブジェクトを示すことができる。メタデータオブジェクトは、異なるタイプのデータ要素(たとえば実行可能なプログラムの入力又は出力として使用されるデータ)及び/または変換(たとえば、データを処理する構成要素であってたとえばデータを処理するか又はデータを発生するデータフローグラフ等に関連付けられた任意の

40

50

タイプのデータ操作)を表すことができる。ビューリング環境は、後に詳述するように、複数のメタデータオブジェクト又は複数のメタデータオブジェクトの複数グループを表す図形上のノード同士を接続する線として関係を示すことができる。幾つかの場合において、インターフェースデータ格納部105Bは、ベースデータ格納部105又は他のデータソースから関係(たとえば系統情報等)を引き出すことができる。インターフェースデータ格納部105Bはデータ系統のハイレベルの概要(summary)を保持することができる。その系統情報(又は他のデータ従属性解析)は、システム100内において自動的に計算されることができ、或いは、外部システムから又は手動入力から受け取られることができる。たとえば、システム100は、今まで集められ且つ人のコードを解析することによって準備された系統情報を受け取ることができる。系統情報は、種々の予め定められたフォーマット形式(たとえば、スプレッドシート形式)にて、ファイルから保管所104へインポートされ得る。

#### 【0024】

図2Aはメタデータビューリング環境の例を示す。幾つかの例において、メタデータビューリング環境は、ブラウザの上部にて実行するインターフェースである。図2Aの例では、メタデータビューリング環境は、データ系統図200Aに関係する情報を表示する。メタデータビューリング環境の一例は、ユーザがメタデータを見ることができ且つ編集することを可能とするウェブベースのアプリケーションである。メタデータビューリング環境を用いると、ユーザは、企業内のどこからでも標準的なウェブブラウザを用いて、メタデータを探索、解析及び管理することができる。メタデータオブジェクトの各タイプは、1つ以上のビューフォーマット視覚的表現を有する。図2Aのメタデータビューリング環境は、目標要素206Aへの系統図を図示する。

#### 【0025】

たとえば、前記系統図は、保管所104に格納されたメタデータオブジェクトを表すデータ及び/または処理ノードに対する先端から末端までの系統を表示する。即ち、前記系統図は、所与の開始オブジェクトが依存するオブジェクト(ソース)と、所与の開始オブジェクトが影響を及ぼすオブジェクト(ターゲット)と、を表示する。この例においては、データ要素202Aと変換204Aとの間、つまり、2つのメタデータオブジェクト例の間、の接続が示されている。メタデータオブジェクトは線図におけるノードによって表現されている。データ要素202Aは、例えば、データセット、データセット内のテーブル、テーブル内の列、ファイル内のフィールド、メッセージ及びレポート等を表す。変換204Aの一例は、どのようにしてデータ要素の単一の出力が生成されるかを記述する実行ファイル(executables)の要素である。ノード間の接続はメタデータオブジェクト間の関係に基づく。

#### 【0026】

図2Bは、各要素202Bがグループ化され且つコンテキスト(前後関係)に基づくグループ内に示されている点を除き、図2Aに示されたものと同一のターゲット要素206Aへの対応する系統図200Bを図示する。たとえば、データ要素202Bは、データセット208B(たとえば、テーブル、ファイル、メッセージ及びレポート)、アプリケーション210B(たとえばグラフ、計画及びプログラムのような実行ファイル、さらに、それらが影響するデータセットを含む)、及びシステム212B、にグループ化されている。システム212Bは、データ及びこのデータを処理するアプリケーションの機能グループ、つまり、アプリケーション及びデータグループ(たとえば、データベース、ファイルグループ、メッセージングシステム、及びデータセットのグループ)よりなるシステムである。変換204Bは、実行ファイル214B、アプリケーション210B及びシステム212Bにグループ化される。たとえば、グラフ、計画及びプログラム等の実行ファイルは、データセットを読み出しつ書込む。パラメータは、デフォルト設定により、どのグループが展開されるか、及び、どのグループが置き込まれるか、を設定できる。これにより、ユーザは、不要なレベルの詳細を取り除くことによって、ユーザにとって重要なグループのみの詳細を見ることができる。

10

20

30

40

50

## 【0027】

メタデータビューア環境を用いてデータ系統計算を実行することは、多くの点で有用である。たとえば、データ要素と変換との間の関係を計算して図示することは、ユーザが、報告値が所与のフィールドレポートのためにどのように計算されたかを決定することを助けることができる。また、ユーザは、どのデータセットが特定のタイプのデータを格納しているか、及び、どの実行ファイルがそのデータセットを読み出し且つ書込むか、を見ることができる。ビジネス用語の場合においては、データ系統図はどのデータ要素（たとえば、列及びフィールド）があるビジネス用語（企業における定義）に関連するかを図示ことができる。

## 【0028】

また、メタデータビューア環境内に示されたデータ系統図は、ユーザがインパクト解析を行うことを支援することができる。具体的には、ユーザは、列またはフィールドがデータセットに追加された場合に下流側の実行ファイルのどれが影響を受けるのか、及び、誰に通知する必要があるのか、について知ることを希望することができる。インパクト解析は、所与のデータ要素がどこで使用されているかを決定でき、また、そのデータ要素の変更の悪影響（ramifications）を決定することもできる。同様に、ユーザは、実行ファイルの変更によってどのデータセットが影響を受けるか、或いは、生成から「あるデータベーステーブル」を削除することは安全か否か、を視覚的に見ることができる。

## 【0029】

メタデータビューア環境を用いてデータ系統図発生用のデータ系統計算を実行することは、ビジネス用語を管理する上で有用である。たとえば、企業内の従業者が、その企業全体に渡るビジネス用語の意味、これらの用語間の関係、及び、これらの用語が参照するデータ、について合意することがしばしば望まれる。ビジネス用語の一貫性のある使用は、企業データの透明性を高め、且つ、ビジネス上の要求の伝達を促進する。従って、ビジネス用語の基礎となる物理的なデータがどこで発見できるか、及び、どのビジネスロジックが計算に用いられるか、を知ることは重要である。

## 【0030】

また、データノード間の関係を見ることは、メタデータを管理し保守する上で有益である。たとえば、ユーザは、だれがメタデータの一部を変更したのか、どのソース（即ち、「レコードのソース」）がメタデータの一部に対応しているのか、或いは、外部ソースからメタデータをロードするとき又は再ロードするときにどんな変更がなされたのか、を知ることを希望することがある。メタデータを保守する場合、指定されたユーザが、メタデータオブジェクト（たとえばビジネス用語）を創出することができ、メタデータオブジェクトの特徴（たとえばオブジェクトの他のオブジェクトに対する記述及び関係）を編集することができ、或いは、もはやサポートされていない（旧式な）メタデータオブジェクトを削除することができる。これにより、ユーザはそれらのオブジェクトに簡単に戻ることができる。

## 【0031】

メタデータビューア環境は、オブジェクトの多数の図形的ビューを提供し、それにより、ユーザはメタデータを探索し且つ解析することができる。たとえば、ユーザは、システム及びアプリケーションの内容を視覚的に見て、任意のオブジェクトの詳細を探索し、更に、データ系統ビューを用いてオブジェクト間の関係を見ることができます。これにより、ユーザは、上述したデータ系統解析及びインパクト解析のような種々のタイプの従属性の解析を容易に実行することができる。また、オブジェクトの階層も見られることができ、また、その階層は特定のオブジェクトを見つけるために探索されるができる。オブジェクトが一旦見つけられたならば、そのオブジェクトに対するしおりが生成され得る。これにより、ユーザはそれらのオブジェクトに簡単に戻ることができる。

## 【0032】

正式な許可を有するユーザは、メタデータビューア環境においてメタデータを編集することができる。たとえば、ユーザはオブジェクトの記述を更新することができ、ビジネス用語を創出することができ、（ビジネス用語をレポート内のフィールド又はテーブル

10

20

30

40

50

内の列にリンクさせるような) オブジェクト間の関係を定義することができ、(たとえば、データセットを1つのアプリケーションから他のアプリケーションへ移動するような) オブジェクトの移動を実行することができ、或いは、オブジェクトを削除することができる。

#### 【0033】

図2Cにおいては目標要素206Aについての対応する系統図200Cが図示されているが、解像度のレベルは目標データ要素206Aについての計算に参加しているアプリケーション(のレベル)に設定されている。具体的に述べると、アプリケーション202C、204C、206C、208C及び210Cのみが目標データ要素206Aに対する計算に直接参加しているので、これらのアプリケーションが図示されている。ユーザが、異なる解像度レベルにてその系統図のある部分を見ることを望めば(たとえば、線図において、より詳細に又はより簡素に表示することを望む場合)、そのユーザは対応する展開/折り畳みボタン212Cをアクティブにする(activate)ことができる。

#### 【0034】

図2Dは解像度レベルが異なる対応する系統図200Dを示す。この例においては、展開/折り畳みボタン212Cがユーザによってアクティブにされており、メタデータビューアイニング環境は現時点にて同一の系統図を表示しているが、アプリケーション202Cはアプリケーション202C内においてデータセット214D及び実行ファイル216Dを示すように展開されている。

#### 【0035】

図2Eは、解像度レベルが異なる対応する系統図200Eを示す。この例において、ユーザはカスタム展開(あつらえの展開、custom expansion)により展開された総てを示すことを選択している。データの最終的なソースであるフィールド又は列(たとえば、それは上流のシステムを有しない。)が展開されている。加えて、特定のフラグのセット(集合)を有するフィールドも展開されている。この例では、系統における重要な中間点にて特定のフラグがデータセット及びフィールド上に設定されており、また、一つの列はその系統が示されている列である。ユーザインターフェースモジュール108は、どのノードが折り畳まれる必要があるか、どのノードが線図から完全に除外される必要があるか、を決定する。

#### 【0036】

また、ユーザは自身の線図を構築することができる。たとえば、線図がメタデータにおける主キー/外部キーの関係を辿ることができるように構築されることができる。また、フィルタを従属性の解析に適用して系統図から情報を排除することもできる。たとえば、ユーザが系統図から拒否ファイルであるデータセットを除外することを望む場合、そのユーザはその系統図内の拒否ファイルの表示をオン・オフに切り替えることができる。

#### 【0037】

メタデータビューアイニング環境において要素及び関係を見る場合、各ノードを表すノードに関係する情報を付加することによってさらに有益になる。ノードに関係情報を付加する方法の一例は、所定のノードの上部に情報を图形的に重複(overlay)させることである。これらの图形は、ノードによって表わされたデータの値又は特徴、更にはメタデータ内のかなる特徴をも、示すことができる。このやり方は、通常は完全に別の2つ以上の情報(データのノード間の関係及びそのノードによって表わされたデータの特徴)を結びつける点で有利であり、「コンテキスト(前後関係)」内に有益な情報を置くように努める。たとえば、メタデータの質(属性)、メタデータの新しさ、又は、レコード情報のソースのような特徴は、データノードの間の関係の視覚的な描写とともに表示され得る。この情報のいくつかは表形式にてアクセスすることができるが、データの異なるノード間の関係とともにそのデータの特徴を見ることはユーザにとって一層有益であろう。ユーザは、メタデータビューアイニング環境において、データのどの特徴がデータ要素及び/または変換ノードの上部に示されるのかを選択することができる。また、どの特徴を示すのかは初期設定システム(デフォルトシステム)の設定に従って設定することもできる

10

20

30

40

50

。

**【0038】**

図3の例において、ノード300は、そのノードによって表わされたメタデータの新しさに係る情報を含む図形オーバレイ302を表示する。「メタデータの新しさ」は、そのメタデータの更新又は外部からの変更がどのくらい最近に行われたのかを意味する。カーソルを図形オーバレイ302上にてうろつかせる(hover)ことにより、図形オーバレイ302により現時点にて表示されている特徴についての更なる詳細を含むウィンドウ304が画面に呼び出される。図形オーバレイは、凡例(legend)306を介して異なる意味に対応付けられた図形の異なる色をもって色コード化され得る。

**【0039】**

図4の例において、メタデータの質のレベルを表示する図形オーバレイが、ノード400上のオーバレイ402を含むデータ要素ノードの上部に重ねられている。メタデータの質の測定は、ビジネスによって用いられ、たとえばデータをインポート又は処理する前にビジネスパートナーから送られてきた周期的(たとえば月毎)データ入力の輪郭を描く(profile)ことができる。これにより、ビジネスは、「悪い」データ(たとえば閾値より高い無効な値の比率を有するデータ)を発見し、この結果、悪いデータが、操作によって元にもどすことが困難な現存のデータ格納部を汚染しない。先に述べた例と同様、カーソルを図形オーバレイ402上にてうろつかせる(hover)ことにより、図形オーバレイ402によって現時点において表示されている特徴についての更なる詳細を含むウィンドウ404が画面上に呼び出される。

**【0040】**

図5の例においては、レコードのソースのタイプを表す図形オーバレイがデータ要素のノード及び変換ノードの上部に重ねられている。ノード500は、レコードのソースが「アクティブなインポートソース」であることを示すオーバレイ502を有する。これは、メタデータがソース(たとえばスプレッドシート)から自動的にインポートされたことを意味する。カーソルを図形オーバレイ502の上でうろつかせることにより、そのインポートに使用されたファイルのタイプ(この例では、エクセルのスプレッドシート)、ファイル名、ファイル所有者及びインポート日付のような詳細を含むウィンドウ504が呼び出される。ノード506は、レコードのソースが「手動によって保守される」ことを示すオーバレイ508を有する。これは、メタデータが(たとえば、ユーザインターフェースモジュール108を用いて)ユーザによって手動にて修正されたことを意味する。カーソルを図形オーバレイ508の上でうろつかせることにより、メタデータを修正したユーザ名及びその修正の日付のような詳細を含むウィンドウ510が呼び出される。

**【0041】**

システム100のモジュール及びシステム100によって実行される手続を含む上述したレコード記憶及び検索のアプローチは、コンピュータ上で実行されるソフトウェアを用いて実現できる。たとえば、このソフトウェアは、1つ以上のプログラムされたコンピュータシステム又はプログラム可能なコンピュータシステム(分散型、クライアント/サーバ型又はグリッド型等の種々アーキテクチャ型である)上で実行される1つ以上のコンピュータプログラムにおける手続を形成する。この各コンピュータシステムは、少なくとも1つのプロセッサ、(揮発性メモリ及び不揮発性メモリ、及び/または、記憶素子を含む)少なくとも1つの記憶システム、少なくとも1つの入力装置またはポート、並びに、少なくとも1つの出力装置またはポートを含む。そのソフトウェアは、たとえば、計算グラフの設計及び構成に関係した他のサービスを提供するより大きなプログラムの1つ以上のモジュールを形成できる。グラフのノード及び要素は、コンピュータ読み出し可能媒体に格納されたデータ構造、または、データ保管所に格納されたデータモデルに合致する他の系統的なデータでもよい。

**【0042】**

上述したアプローチは、コンピュータ上で実行されるソフトウェアを用いて実現できる。たとえば、このソフトウェアは、1つ以上のプログラムされたコンピュータシステム又

10

20

30

40

50

はプログラム可能なコンピュータシステム（分散型、クライアント/サーバ型又はグリッド型等の種々アーキテクチャ型である）上で実行される1つ以上のコンピュータプログラムにおける手続を形成する。この各コンピュータシステムは、少なくとも1つのプロセッサ、（揮発性メモリ及び不揮発性メモリ、及び／または、記憶素子を含む）少なくとも1つの記憶システム、少なくとも1つの入力装置またはポート、並びに、少なくとも1つの出力装置またはポートを含む。そのソフトウェアは、たとえば、計算グラフの設計及び構成に関係した他のサービスを提供するより大きなプログラムの1つ以上のモジュールを形成できる。グラフのノード及び要素は、コンピュータ読み出し可能媒体に格納されたデータ構造、または、データ保管所に格納されたデータモデルに合致する他の系統的なデータでもよい。

10

#### 【0043】

ソフトウェアは、汎用または専用のプログラマブルコンピュータによって読み出し可能なCD-ROM等の記憶媒体上に提供され、或いは、そのソフトウェアが実行されるコンピュータにネットワークの通信媒体を用いて配信される。すべての機能は、専用のコンピュータ、または、コプロセッサのような専用ハードウェアを用いて実行され得る。ソフトウェアは、ソフトウェアによって特定された異なる計算の部分が異なるコンピュータによって実行されるという分散方式によって実行され得る。好ましくは、上述の各コンピュータプログラムは、汎用または専用プログラマブルコンピュータによって読み出し可能な記憶媒体またはデバイス（たとえば、固体メモリまたは媒体、もしくは磁気的または光学的媒体）に格納され或いはダウンロードされ、記憶媒体またはデバイスがコンピュータシステムによって読み出されて上述した手順を実行するときに、コンピュータを構築及び動作させる。また、本発明のシステムは、コンピュータプログラムとともに構築されたコンピュータ読み出し可能記憶媒体として実現されることが考えられる。ここで、そのように構築された記憶媒体は、上述の機能を実行するために特別且つ予め定義された様式でコンピュータシステムを作動させる。

20

#### 【0044】

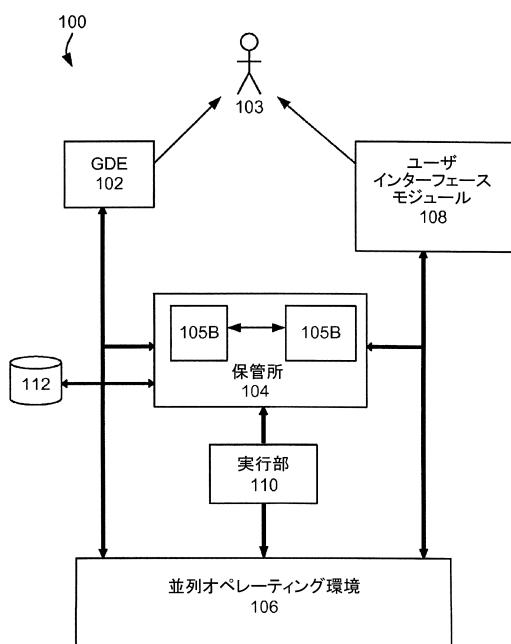
多くの本発明の実施例を記載した。しかしながら、本発明の精神及び範囲から逸脱することなく多種の変更がなされ得ることが理解されよう。たとえば上述のいくつかのステップは、順序が独立であり、従って上述の順序と異なる順序で実行され得る。

30

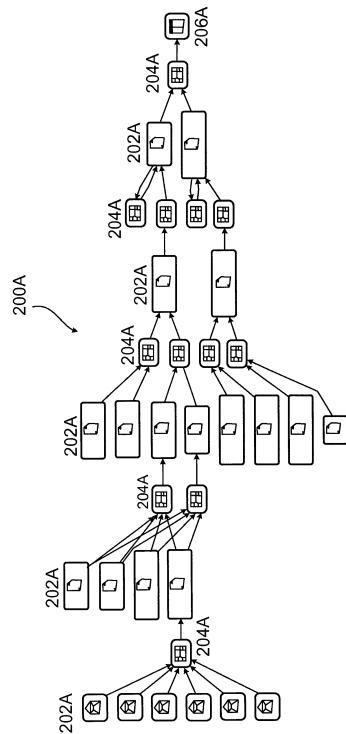
#### 【0045】

上述の記述は説明のためにあり、添付された本発明の特許請求の範囲によって定義される発明の範囲を限定する意図はないことが理解されるべきである。たとえば、上述の機能ステップの多くは、全体の処理に実質的に影響を及ぼすことなく異なる順序で実行され得る。他の実施例は特許請求の範囲の請求項の範囲内にある。

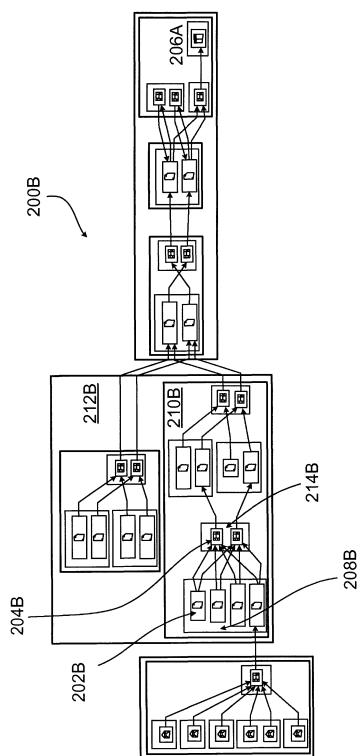
【図1】



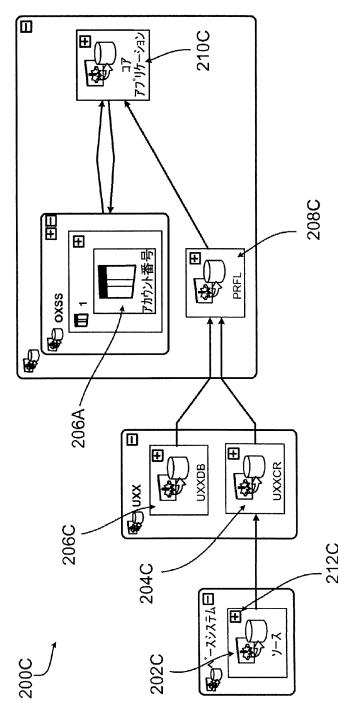
【 図 2 A 】



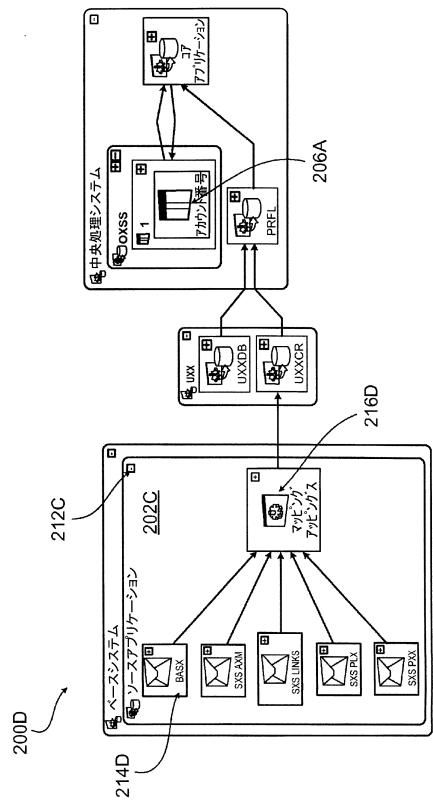
【図2B】



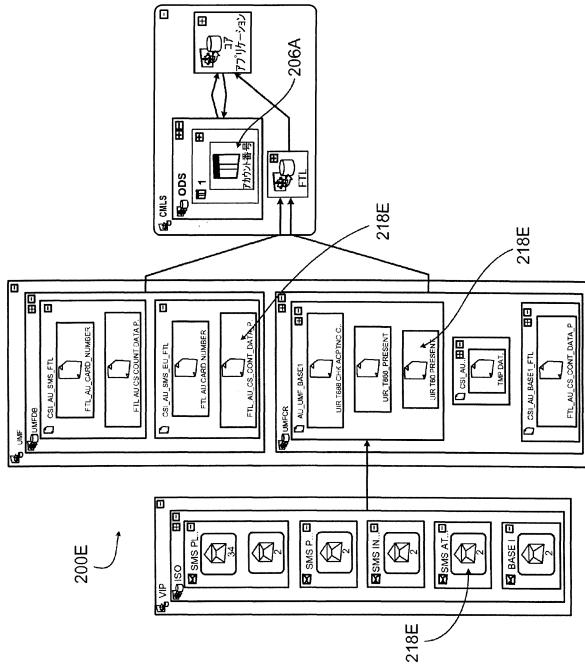
【図2C】



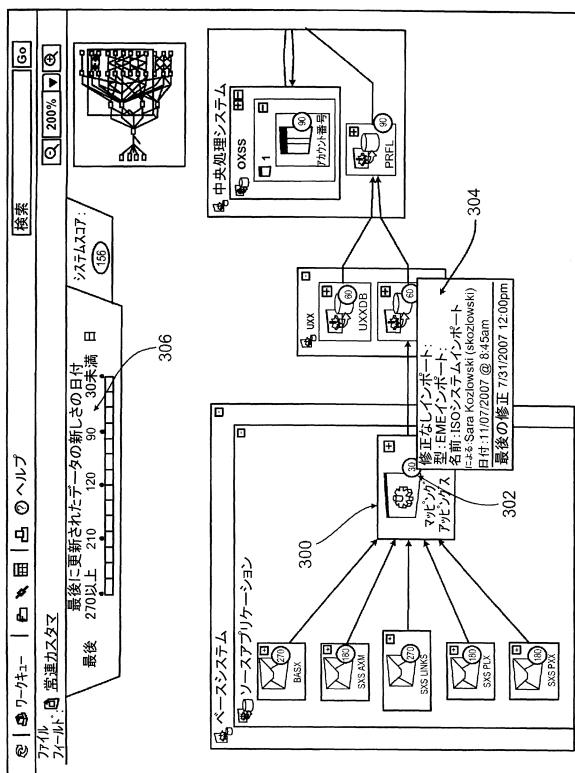
【図2D】



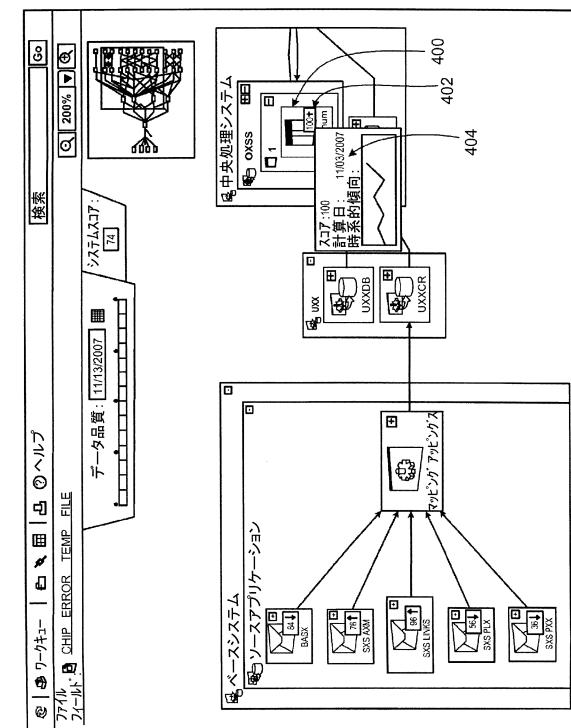
【図2E】



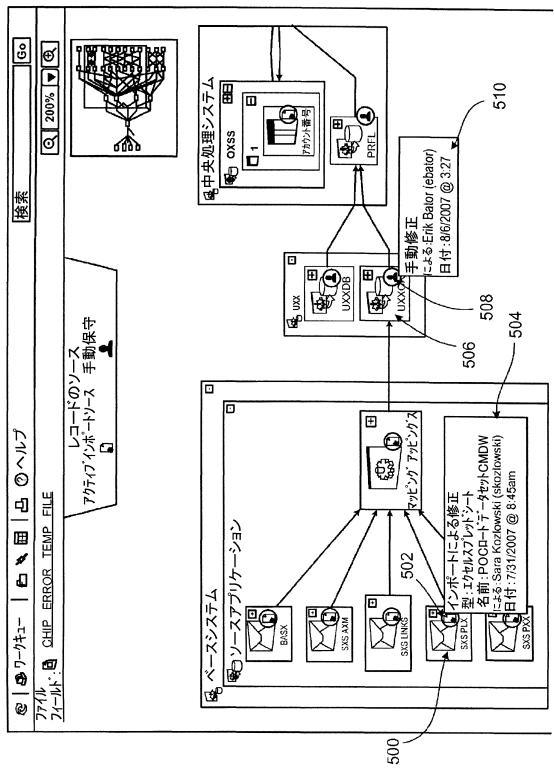
【図3】



【図4】



【図5】



---

フロントページの続き

(72)発明者 バートル エリク  
アメリカ合衆国 02494 マサチューセッツ州 ニーダム セントラル アベニュー 314  
(72)発明者 グールド ジョエル  
アメリカ合衆国 01890 マサチューセッツ州 ウィンチェスター ソーンベリー ロード  
50  
(72)発明者 ラディヴィオジェヴィク ドゥシャン  
アメリカ合衆国 01845 マサチューセッツ州 ノース アンドーヴァー ウェイランド サ  
ークル 140

合議体

審判長 渡邊 聰  
審判官 藤井 浩  
審判官 小池 正彦

(56)参考文献 特表2008-524671(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 1/00  
G06T 11/00  
G06F 12/00  
G06F 3/048