

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7237905号
(P7237905)

(45)発行日 令和5年3月13日(2023.3.13)

(24)登録日 令和5年3月3日(2023.3.3)

(51)国際特許分類	F I
G 0 6 F 16/22 (2019.01)	G 0 6 F 16/22
G 0 6 F 16/27 (2019.01)	G 0 6 F 16/27
G 0 6 F 16/21 (2019.01)	G 0 6 F 16/21

請求項の数 10 外国語出願 (全31頁)

(21)出願番号	特願2020-177565(P2020-177565)	(73)特許権者	500575824
(22)出願日	令和2年10月22日(2020.10.22)		ハネウェル・インターナショナル・イン
(65)公開番号	特開2021-68448(P2021-68448A)		コーポレーテッド
(43)公開日	令和3年4月30日(2021.4.30)		Honeywell International Inc.
審査請求日	令和2年12月9日(2020.12.9)		アメリカ合衆国、28202 ノース・
(31)優先権主張番号	16/660,122		カロライナ州、シャーロット、サウス・
(32)優先日	令和1年10月22日(2019.10.22)		ミント・ストリート、855
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100118902
			弁理士 山本 修
		(74)代理人	100106208
			弁理士 宮前 徹
		(74)代理人	100196508
			弁理士 松尾 淳一
		(74)代理人	100119781

最終頁に続く

(54)【発明の名称】 データマッピングのための方法、装置、及びシステム

(57)【特許請求の範囲】

【請求項1】

コンピュータ実施方法であって、
第1のデータベーススキーマに対してデータマッピングプロセスを実行するための要求を受信することと、

前記要求に応答して、

データベースから、前記第1のデータベーススキーマに関連付けられた第1のデータモデルに関する第1の複数のデータオブジェクトを取得することと、

前記第1の複数のデータオブジェクトに少なくとも基づいて、第1の機械学習モデルを用いて、前記第1のデータベーススキーマに対応する第1のデータ分類子を決定することと、

10

前記第1のデータ分類子に関連付けられた信頼スコアを計算することと、

前記信頼スコアが、所定の閾値を満たすか否かを決定することと、

前記第1のデータ分類子が、前記第1の複数のデータオブジェクトに対して正確であるか否かを示す、ユーザからのフィードバックデータを生成することと、

少なくとも部分的に、前記フィードバックデータに基づいて、前記第1の機械学習モデルを訓練することと、

前記信頼スコアが前記所定の閾値を満たすという決定に応答して、

前記第1のデータ分類子及び前記第1の複数のデータオブジェクトに関連付けられたメタデータに少なくとも部分的に基づいて、マッピング仕様に関連付けられたデータ動

20

作に関するルールのセットを決定することと、

前記マッピング仕様に関連付けられた前記ルールのセットに基づいて、前記第 1 のデータベーススキーマに関連付けられた前記第 1 の複数のデータオブジェクトを、第 2 のデータベーススキーマに関連付けられた第 2 のデータモデルに関する第 2 の複数のデータオブジェクトに変換することと、

前記第 2 の複数のデータオブジェクトに対して 1 つ以上のデータの本質的な特徴の理解を提供することと、を含む、コンピュータ実施方法。

【請求項 2】

前記第 1 の複数のデータオブジェクトが第 1 のデータテーブルを含み、前記第 1 のデータテーブルは、少なくとも 1 つのデータフィールドを含む、請求項 1 に記載のコンピュータ実施方法。

10

【請求項 3】

前記第 1 のデータテーブルが、名前メタデータ、列メタデータ、又は行メタデータのうちの少なくとも 1 つを含む、請求項 2 に記載のコンピュータ実施方法。

【請求項 4】

前記第 1 のデータベーススキーマに対応する前記第 1 のデータ分類子を決定することが、前記第 1 の複数のデータオブジェクトに関連付けられた前記名前メタデータ、前記列メタデータ、又は前記行メタデータのうちの少なくとも 1 つを取得することを更に含む、請求項 3 に記載のコンピュータ実施方法。

【請求項 5】

前記第 1 のデータベーススキーマに対応する前記第 1 のデータ分類子を決定することが、前記名前メタデータ、前記列メタデータ、又は前記行メタデータのうちの少なくとも 1 つに更に基づいて、前記第 1 のデータ分類子を決定することを更に含む、請求項 4 に記載のコンピュータ実施方法。

20

【請求項 6】

前記第 1 の複数のデータオブジェクトが第 2 のデータテーブルを含み、前記コンピュータ実施方法は、

前記第 1 のデータテーブル及び前記第 2 のデータテーブルに関連付けられた関連メタデータを決定することを更に含む、請求項 2 に記載のコンピュータ実施方法。

【請求項 7】

前記第 1 の複数のデータオブジェクトが第 2 のデータテーブルを含み、前記コンピュータ実施方法は、

前記関連メタデータに更に基づいて、前記第 1 のデータ分類子を決定することを更に含む、請求項 6 に記載のコンピュータ実施方法。

30

【請求項 8】

前記ルールのセットを決定することが、第 2 の機械学習モデルを用いて、自動的に、前記マッピング仕様を生成することを含む、
請求項 1 に記載のコンピュータ実施方法。

【請求項 9】

前記第 1 のデータベーススキーマが、前記第 1 の複数のデータオブジェクトの 1 つ又は複数のデータオブジェクトの間のデータ編成を規定する構造を含み、

40

前記第 2 のデータベーススキーマが、前記第 2 の複数のデータオブジェクトの 1 つ又は複数のオブジェクトの間のデータ編成を規定する構造を含む、

請求項 1 に記載のコンピュータ実施方法。

【請求項 10】

前記信頼スコアを計算することが、前記第 1 の複数のデータオブジェクトのデータフィールドと、教師あり学習モデルに関連付けられた訓練データとの間の類似性の度合いに基づいて前記信頼スコアを計算することを含む、請求項 1 に記載のコンピュータ実施方法。

【発明の詳細な説明】

【技術分野】

50

【0001】

本開示は、概して、データマッピングのための方法、装置、及びシステムに関し、より具体的には、データマッピングのためのデータ分類子を生成するための方法、装置、及びシステムに関する。

【背景技術】

【0002】

データベースは、コンピューティングシステムによって電子的に記憶、アクセス、及び/又は管理され得る情報又はデータ(データオブジェクトなど)の集合である。データベーススキーマとは、データベース内でデータがどのように編成され、かつ/又は互いに関連付けられるかを定義し得る構造を指す。データベースは、限定するものではないが、ヘルスケア産業、小売業界、及び財務サービスを含む、様々な分野で実現され得る。

10

【0003】

多くのシステム及び方法は、データベースに関連する技術的課題及び困難を克服しない。例えば、データは、共通データモデルに準拠しない場合があるレガシーデータベースに記憶される可能性があり、多くのシステム及び方法は、コンピューティングシステムによって分析され得るようにこれらのデータを変換する能力を提供しない。これらの課題及び困難は、レガシーデータベースが分析を必要とする大量のデータを記憶するときに更に増幅され得る。

【発明の概要】

【0004】

様々な実施例によれば、装置が提供され得る。この装置は、少なくとも1つのプロセッサと、プログラムコードを含む少なくとも1つの非一時的メモリと、を備え得る。少なくとも1つの非一時的メモリ及びプログラムコードは、少なくとも1つのプロセッサを用いて、装置に少なくとも、第1のデータベーススキーマに関連付けられた第1の複数のデータオブジェクトを取得させ、第1の複数のデータオブジェクトに少なくとも基づいて第1のデータベーススキーマに対応する第1のデータ分類子を決定させ、第1のデータ分類子及び第1の複数のデータオブジェクトに少なくとも部分的に基づいてマッピング仕様を生成させ、第1の複数のデータオブジェクト及びマッピング仕様に少なくとも部分的に基づいて第2の複数のデータオブジェクトを生成させるように構成されてもよい。いくつかの実施例では、マッピング仕様は、第1のデータベーススキーマに関連付けられた第1の複数のデータオブジェクトを、第2のデータベーススキーマに関連付けられた第2の複数のデータオブジェクトに変換するように構成されてもよい。

20

30

【0005】

いくつかの実施例では、第1の複数のデータオブジェクトは、第1のデータテーブルを含んでもよい。いくつかの実施例では、第1のデータテーブルは、少なくとも1つのデータフィールドを含んでもよい。いくつかの実施例では、第1のデータテーブルは、名前メタデータ、列メタデータ、又は行メタデータのうちの少なくとも1つを含んでもよい。

【0006】

いくつかの実施例では、第1のデータベーススキーマに対応する第1のデータ分類子を決定するとき、少なくとも1つの非一時的メモリ及びプログラムコードは、少なくとも1つのプロセッサを用いて、装置に、第1の複数のデータオブジェクトに関連付けられた名前メタデータ、列メタデータ、又は行メタデータのうちの少なくとも1つを取得させ、名前メタデータ、列メタデータ、又は行メタデータのうちの少なくとも1つに更に基づいて第1のデータ分類子を決定させるように構成されてもよい。

40

【0007】

いくつかの実施例では、第1の複数のデータオブジェクトは、第2のデータテーブルを含んでもよい。いくつかの実施例では、少なくとも1つの非一時的メモリ及びプログラムコードは、少なくとも1つのプロセッサを用いて、装置に、第1のデータテーブル及び第2のデータテーブルに関連付けられた相関メタデータを決定させ、相関メタデータに更に基づいて第1のデータ分類子を決定させるように構成されてもよい。

50

【 0 0 0 8 】

いくつかの実施例では、第 1 のデータベーススキーマに対応する第 1 のデータ分類子を決定するとき、少なくとも 1 つの非一時的メモリ及びプログラムコードは、少なくとも 1 つのプロセッサを用いて、装置に、第 1 のデータテーブルに関連付けられたドメインメタデータを決定させ、ドメインメタデータに更に基づいて第 1 のデータ分類子を決定させるように構成されてもよい。

【 0 0 0 9 】

いくつかの実施例では、マッピング仕様を生成する前に、少なくとも 1 つの非一時的メモリ及びプログラムコードは、少なくとも 1 つのプロセッサを用いて、装置に更に、第 1 のデータ分類子に関連付けられた信頼スコアを計算させ、信頼スコアが所定の閾値を満たすかどうかを決定させるように構成されてもよい。

10

【 0 0 1 0 】

いくつかの実施例では、マッピング仕様を生成することは、信頼スコアが所定の閾値を満たしているという判定に対する応答であってもよい。

【 0 0 1 1 】

いくつかの実施例では、少なくとも 1 つの非一時的メモリ及びプログラムコードは、少なくとも 1 つのプロセッサを用いて、装置に更に、信頼スコアが所定の閾値を満たしていないと判定させ、第 1 のデータ分類子に関連付けられたユーザ入力要求を生成させ、ユーザ入力要求に応答してユーザ入力を受信させるように構成されてもよい。いくつかの実施例では、ユーザ入力要求は、第 1 のデータ分類子を確認するための電子要求を含んでもよい。

20

【 0 0 1 2 】

いくつかの実施例では、ユーザ入力は、第 1 のデータ分類子の確認を含んでもよい。いくつかの実施例では、マッピング仕様を生成することは、第 1 のデータ分類子の確認に対する応答であってもよい。

【 0 0 1 3 】

いくつかの実施例では、ユーザ入力は、第 1 のデータ分類子の修正を含んでもよく、少なくとも 1 つの非一時的メモリ及びプログラムコードは、少なくとも 1 つのプロセッサを用いて、装置に更に、ユーザ入力に基づいて第 1 のデータ分類子を修正させるように構成される。

30

【 0 0 1 4 】

いくつかの実施例では、少なくとも 1 つの非一時的メモリ及びプログラムコードは、少なくとも 1 つのプロセッサを用いて、装置に更に、ユーザ入力に基づいてフィードバックデータを生成させ、第 3 のデータベーススキーマに関連付けられた第 3 の複数のデータオブジェクトを取得させ、第 3 の複数のデータオブジェクト及びフィードバックデータに少なくとも基づいて第 3 のデータベーススキーマに対応する第 2 のデータ分類子を決定させるように構成されてもよい。

【 0 0 1 5 】

様々な実施例によれば、コンピュータ実施方法が提供され得る。コンピュータ実施方法は、データベースから第 1 のデータベーススキーマに関連付けられた第 1 の複数のデータオブジェクトを取得することと、第 1 の複数のデータオブジェクトに少なくとも基づいて第 1 のデータベーススキーマに対応する第 1 のデータ分類子を決定することと、第 1 のデータ分類子及び第 1 の複数のデータオブジェクトに少なくとも部分的に基づいてマッピング仕様を生成することと、第 1 の複数のデータオブジェクト及びマッピング仕様に少なくとも部分的に基づいて第 2 の複数のデータオブジェクトを生成することと、を含んでもよい。いくつかの実施例では、マッピング仕様は、第 1 のデータベーススキーマに関連付けられた第 1 の複数のデータオブジェクトを、第 2 のデータベーススキーマに関連付けられた第 2 の複数のデータオブジェクトに変換するように構成されてもよい。

40

【 0 0 1 6 】

様々な実施例によれば、コンピュータプログラム製品が提供され得る。コンピュータ

50

プログラム製品は、その中に記憶されたコンピュータ可読プログラムコード部分を有する少なくとも1つの非一時的コンピュータ可読記憶媒体を含んでもよい。コンピュータ可読プログラムコード部分は、データベースから第1のデータベーススキーマに関連付けられた第1の複数のデータオブジェクトを取得し、第1の複数のデータオブジェクトに少なくとも基づいて第1のデータベーススキーマに対応する第1のデータ分類子を決定し、第1のデータ分類子及び第1の複数のデータオブジェクトに少なくとも部分的に基づいてマッピング仕様を生成し、第1の複数のデータオブジェクト及びマッピング仕様に少なくとも部分的に基づいて第2の複数のデータオブジェクトを生成するように構成された実行可能部分を含んでもよい。いくつかの実施例では、マッピング仕様は、第1のデータベーススキーマに関連付けられた第1の複数のデータオブジェクトを、第2のデータベーススキーマに関連付けられた第2の複数のデータオブジェクトに変換するように構成されてもよい。

10

【0017】

前述の実例となる概要、並びに本開示の他の例示的な目的及び/又は利点、並びにそれが達成される方法は、以下の発明を実施するための形態及びその添付図面において更に説明される。

【図面の簡単な説明】

【0018】

図示の実施形態の説明は、添付の図面と併せて読むことができる。特に記載がない限り、図の簡略化及び明確化のために、図面に示される要素は必ずしも縮尺どおりに描かれていないことが理解されよう。例えば、特に記載のない限り、要素のうちのいくつかの寸法は、他の要素に対して誇張されている場合がある。本開示の教示を組み込む実施形態は、本明細書に提示される図に関連して示され、説明される。

20

【0019】

【図1】本開示の様々な実施形態による例示的なシステムの例示的な概略図を示す。

【0020】

【図2】本開示の様々な実施形態による例示的な装置の例示的なブロック図を示す。

【0021】

【図3】本開示の様々な実施形態による例示的なフロー図を示す。

【0022】

【図4】本開示の様々な実施形態による例示的なフロー図を示す。

30

【0023】

【図5】本開示の様々な実施形態による例示的なフロー図を示す。

【0024】

【図6】本開示の様々な実施形態による例示的なフロー図を示す。

【0025】

【図7】本開示の様々な実施形態による例示的なフロー図を示す。

【発明を実施するための形態】

【0026】

次に、本開示のいくつかの実施形態について添付図面を参照しながら以下により詳細に説明するが、本開示の全てではなくいくつかの実施形態を示すものである。実際に、本開示は、多くの異なる形態で具現化されてもよく、本明細書に記載される実施形態に限定されるものとして解釈されるべきではなく、むしろ、これらの実施形態は、本開示が適用可能な法的要件を満たすように提供される。同様の数字は、全体をとおして同様の要素を指す。

40

【0027】

語句「一実施形態では」、「一実施形態によれば」、「例えば」、「いくつかの実施例では」、「一例として」、及び同様の語句は、その語句に続く特定の特徴、構造又は特性が、本開示の少なくとも1つの実施形態に含まれてもよいこと、及び本開示の2つ以上の実施形態に含まれてもよい（そのような語句は必ずしも同じ実施形態に言及しない）ことを一般に意味する。

50

【 0 0 2 8 】

本明細書で使用するとき、用語「例」又は「例示的な」は、「一例、事例、又は実例としての役割を果たすこと」を意味する。「例」又は「例示的な」として本明細書に記載される任意の実施態様は、必ずしも他の実施態様よりも好ましい又は有利であると解釈されなくてよい。

【 0 0 2 9 】

本明細書が、ある構成要素又は特徴が含まれる又は特性を有することに関して「～してもよい (may)」、「～し得る (can、could)」、「～べきである (should)」、「～であろう (would)」、「好ましくは (preferably)」、「場合により (possibly)」、「典型的には (typically)」、「任意選択的に (optionally)」、「例えば (for example)」、「いくつかの実施例では (in some examples)」、多くの場合 (often)」又は「～かもしれない (might)」(又は他のそのような言語)と提示している場合、その特定の構成要素又は特徴は、含まれること又は特性を有することを必要としない。このような構成要素又は特徴は、いくつかの実施形態に任意選択的に含まれてもよく、又は除外されてもよい。

10

【 0 0 3 0 】

用語「回路」は、ハードウェアを含み、いくつかの実施形態では、ハードウェアを構成するためのソフトウェアを含むことが広く理解されるべきである。したがって、装置の構成要素に関して、本明細書で使用するとき、用語「回路」は、本明細書で説明される特定の回路に関連付けられた機能を実行するように構成された特定のハードウェアを含むことが理解されるべきである。例えば、いくつかの実施形態では、「回路」は、処理回路、記憶媒体、ネットワークインターフェース、入力/出力装置などを含んでもよい。

20

【 0 0 3 1 】

本開示において、用語「電子的に結合された」、「電子的に結合している」、「電子的に結合する」、「～と通信する」、「～と電子通信する」、又は「接続された」という用語は、2つ以上の構成要素(例えば、限定するものではないが、クライアントデバイス(単数又は複数)、データ分析システム(単数又は複数)、データベース(単数又は複数))が有線手段(例えば、限定されるものではないが、有線イーサネット)及び/又は無線手段(例えば、限定するものではないが、Wi-Fi、Bluetooth、ZigBee)を介して接続され、それにより、データ及び/又は情報は、これらの構成要素に送信され、かつ/又はこれらから受信され得ることを指す。

30

【 0 0 3 2 】

用語「データ分析システム」は、データベースに記憶されたデータの意味を示し得るデータ分類子を生成し、かつ/又はデータ分類子に基づいてマッピング仕様を生成するように構成され得るシステム又は仮想環境を指し得る。データ分析システムは、ソフトウェアアプリケーションを実行し、デジタルコンテンツアイテム、アプリケーション関連データ、及び/又は同様のものを記憶する1つ以上のデータベースへのアクセスを有する、1つ以上の追加のサーバーと通信するように配設された、1つ以上の中央サーバーの形態をとり得る。例示的なデータ分析システムは、本明細書の図1に少なくとも関連して記載される。

40

【 0 0 3 3 】

用語「ユーザ」は、個人、個人のグループ、ビジネス、組織などを指すものと理解されるべきである。本明細書で言及されるユーザは、クライアントデバイスを使用してデータ分析システムにアクセスすることができる。用語「クライアントデバイス」は、データ分析システムにアクセスするように構成されたコンピュータハードウェア及び/又はソフトウェアを指す。クライアントデバイスとしては、限定するものではないが、スマートフォン、タブレットコンピュータ、ラップトップコンピュータ、ウェアラブル、パーソナルコンピュータ、企業コンピュータなどが挙げられ得る。

【 0 0 3 4 】

いくつかの実施例では、データ及び情報(電子要求など)は、データ分析システムに

50

送信され、かつ／又はそこから受信され得る。例えば、「データ分類要求」は、データベースに記憶されたデータの意味を示し得るデータ分類子を生成するための電子要求を示し得る。いくつかの実施例では、データ分類要求は、クライアントデバイスからデータ分析システムに送信されてもよい。別の例として、「ユーザ入力要求」は、ユーザ入力を提供するための電子要求を示し得る。いくつかの実施例では、ユーザ入力要求は、データ分析システムからクライアントデバイスに送信されてもよく、データ分類子を確認するための電子要求を含んでもよく、その詳細は本明細書に記載される。

【0035】

用語「データオブジェクト」は、データに関連付けられた1つ以上の値を表し得るデータ構造を指す。データオブジェクトはデータベースに記憶されてもよく、1つ以上の「データフィールド」を含んでもよい。いくつかの実施例では、データフィールドは、情報交換用米国標準コード(American Standard Code for Information Interchange: ASCII)テキスト、ポインタ、メモリアドレスなどの形態であってもよく、データオブジェクトに関連付けられた少なくとも1つの値を含んでもよい。

10

【0036】

例えば、データオブジェクトは、行(単数又は複数)及び／又は列(単数又は複数)を含み得る表形式又は半表形式で値を表し得る「データテーブル」を含んでもよい。いくつかの実施例では、データテーブルは、少なくとも1つのデータフィールドを含んでもよい。例示的なデータテーブル内の例示的なデータフィールドを以下に示す。

【表1】

20

(名前)	(列1)	(列2)
(行1)	データフィールド1	データフィールド2
(行2)	データフィールド3	データフィールド4

上記の例では、データテーブルは、2つの行、2つの列、及び4つのデータフィールド(データフィールド1、データフィールド2、データフィールド3、及びデータフィールド4など)を含んでもよい。

【0037】

いくつかの実施例では、データテーブルの1つ以上のデータフィールドは、行識別子及び／又は列識別子に関連付けられてもよい。行識別子は、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよく、データフィールドに関連付けられた行を一意に識別することができる。列識別子は、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよく、データフィールドに関連付けられた列を一意に識別することができる。

30

【0038】

上記の例から続けると、データフィールド1は、行識別子「行1」及び列識別子「列1」に関連付けられてもよい。データフィールド2は、行識別子「行1」及び列識別子「列2」に関連付けられてもよい。データフィールド3は、行識別子「行2」及び列識別子「列1」に関連付けられてもよい。データフィールド4は、行識別子「行2」及び列識別子「列2」に関連付けられてもよい。

40

【0039】

用語「メタデータ」は、他のデータ(データオブジェクトのデータフィールドなど)を記述し得るデータを指す。いくつかの実施例では、データオブジェクトは、1つ以上のメタデータを含んでもよく、かつ／又は関連付けられてもよい。いくつかの実施例では、メタデータは、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよく、データフィールド内の値の意味に関連付けられた情報を含んでもよい。

【0040】

いくつかの実施例では、データテーブルは、以下の例に示されるように、名前メタデータ、列メタデータ、及び／又は行メタデータのうちの少なくとも1つを含んでもよい。

50

【表 2】

名前メタデータ1	列メタデータ1	列メタデータ2
行メタデータ1	データフィールド1	データフィールド2
行メタデータ2	データフィールド3	データフィールド4

上記の例では、名前メタデータ（名前メタデータ1など）、列メタデータ（列メタデータ1及び列メタデータ2など）、及び行メタデータ（行メタデータ1及び行メタデータ2など）は、データテーブル内のデータフィールド（データフィールド1、データフィールド2、データフィールド3、及びデータフィールド4など）の意味に関連付けられてもよい。

10

【0041】

いくつかの実施例では、名前メタデータは、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよく、データテーブルの名前を含んでもよい。いくつかの実施例では、列メタデータは、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよく、列識別子及び/又は列の名前を含んでもよい。いくつかの実施例では、行メタデータは、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよく、行識別子及び/又は行の名前を含んでもよい。いくつかの実施例では、データ分析システムは、名前メタデータ、列メタデータ、及び/又は行メタデータのうちの少なくとも1つに基づいてデータフィールドの意味を決定するために、機械学習モデルを実施してもよく、その詳細は本明細書に記載される。

20

【0042】

上記の例示的なデータテーブルは、例示的な名前メタデータ、例示的な列メタデータ、及び例示的な行メタデータを示しているが、本開示の範囲は、これらのメタデータに限定されないことに留意されたい。例えば、データ分析システムは、2つ以上のデータオブジェクト間の関係を示し得る「相関メタデータ」を決定することができる。いくつかの実施例では、データ分析システムは、データオブジェクトに関連付けられたドメインを示し得る「ドメインメタデータ」を決定することができる。相関メタデータ及びドメインメタデータの例は、本明細書に記載される。

【0043】

用語「データオブジェクト識別子」は、複数のデータオブジェクトから、かつ/又は1つ以上のデータベースからデータオブジェクトを一意に識別及び/又は検索し得る識別子を指す。いくつかの実施例では、データオブジェクト識別子は、ASCIIテキスト、メモリアドレス、ネットワークアドレスなどの形態であってもよい。

30

【0044】

いくつかの実施例では、データベース内に記憶されたデータオブジェクトは、データベースのデータベーススキーマに関連付けられてもよい。上述のように、「データベーススキーマ」は、データベース内でデータがどのように編成され、かつ/又は互いに関連付けられるかを定義し得るデータベースの構造を指し得る。いくつかの実施例では、データベーススキーマは、データモデルの物理的な実現であってもよい。用語「データモデル」は、データオブジェクトを編成し、それらの関係を標準化し得る抽象モデルを指し得る。例示的なデータモデルとしては、（限定するものではないが）アプリケーション及び/又はデータソースの間で共有されるデータオブジェクトの標準化された編成を提供し得る共有データモデルである、共通データモデル（CDM）が挙げられ得る。

40

【0045】

用語「データベース識別子」は、データベース及び/又はデータベーススキーマを一意に識別し、かつ/又は検索し得る識別子を指す。いくつかの実施例では、データオブジェクト識別子は、ASCIIテキスト、メモリアドレス、ネットワークアドレスなどの形態であってもよい。

【0046】

用語「データ分類子」は、分類を示し、説明を含み、かつ/又は1つ以上のデータオ

50

プロジェクトに関連付けられた情報の意味を提供し得るデータを指し得る。データ分類子は、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよい。例えば、データ分類子は、1つ以上のデータオブジェクト内の1つ以上のデータフィールドに関連付けられたデータカテゴリを含み得る、テキスト文字列の形態であってもよい。例示的なデータカテゴリとしては、限定するものではないが、ユーザ名、電子メールアドレス、バッテリーレベル値、測定値が挙げられ得る。いくつかの実施例では、データ分類子は、データ分析システムによって生成されてもよく、その例示的な詳細は本明細書に記載される。

【0047】

用語「マッピング仕様」は、あるデータモデルに関連付けられたデータが別のデータモデルに基づいてどのように表され得るかを決定し得る、動作及び/又は変換を記述及び/又は指定し得るデータオブジェクトを指し得る。例えば、マッピング仕様は、第1のデータベーススキーマに関連付けられた第1の複数のデータオブジェクトを、第2のデータベーススキーマに関連付けられた第2の複数のデータオブジェクトに変換するように構成されてもよい。いくつかの実施例では、マッピング仕様は、データ分析システムによって生成されてもよく、その例示的な詳細は本明細書に記載される。

10

【0048】

上述のように、多くのシステム及び方法は、データベースに関連する技術的課題及び困難を克服しない。例えば、多くのシステムは、最新のデータベースモデルに準拠しない場合があるレガシーデータベースに記憶されたデータを分析する能力を欠いている。いくつかの実施例では、これらのデータを用いてどのようなデータの本質的な特徴の理解が生成され得るか不明である場合に、データ分析の目的のためにこれらのデータを消費するのはコストがかかる可能性がある。例えば、レガシーデータベースに記憶されたデータは、数字の文字列を含む可能性があり、これらの数字の文字列が何を表し得るかは不明である場合がある。したがって、分析のために使用され得る前に、データを理解するのにコスト及び時間がかかる場合がある。いくつかの実施例では、データを理解した結果として有用なデータの本質的な特徴の理解が生成され得ない場合、コンピューティングリソースが無駄になり得る。

20

【0049】

対照的に、本開示による様々な実施例は、これらの課題及び困難を克服し得る。いくつかの実施例では、機械知能ソリューション（機械学習）及び他の技術を使用するなど）は、データの名前、データの種類、及びデータ間の関係を検査することによって、データベース内のデータの意味を推測し得る。

30

いくつかの実施例では、これらのデータは、データの意味において十分な信頼性がある場合、分析目的のためにデータモデルに自動的にマッピングされてもよい。

いくつかの実施例では、データの意味における信頼性が十分でない場合、ユーザは、データの意味に関するユーザ入力を提供することができ、それを使用して、機械知能ソリューションが改善され得る。したがって、本開示の様々な実施例は、データマッピングに必要とされるコスト及びコンピューティングリソースを低減させることができ、データからの有用なデータの本質的な特徴の理解の質を高めることができる。

【0050】

40

本開示の方法、装置、及びコンピュータプログラム製品は、様々なデバイスのうちのいずれかによって具現化され得る。例えば、例示的な実施形態の方法、装置、及びコンピュータプログラム製品は、1つ以上のクライアントデバイスなどの1つ以上のデバイスと通信するように構成されたサーバー又は他のネットワークエンティティなどのネットワークデバイス（例えば、データ分析システム）によって具現化されてもよい。加えて、又はあるいは、コンピューティングデバイスは、パーソナルコンピュータ又はコンピュータワークステーションなどの固定コンピューティングデバイスを含んでもよい。加えて、又はあるいは、例示的な実施形態は、ポータブルデジタルアシスタント（PDA）、携帯電話、スマートフォン、ラップトップコンピュータ、タブレットコンピュータ、ウェアラブル、又は前述のデバイスの任意の組み合わせなどの様々なモバイル機器のいずれかによって

50

具現化されてもよい。

【0051】

図1は、本開示の実施形態がその中で動作し得る、例示的なシステムアーキテクチャ100を示す。ユーザは、クライアントデバイス101A、101B、101C、...、101Nを使用して、通信ネットワーク103を介してデータ分析システム105にアクセスすることができる。

【0052】

クライアントデバイス101A~101Nは、上述のような任意のコンピューティングデバイスであってもよい。クライアントデバイス101A~101Nからデータ分析システム105によって受信される電子データは、様々な形態で、様々な方法を介して提供されてもよい。例えば、クライアントデバイス101A~101Nとしては、デスクトップコンピュータ、ラップトップコンピュータ、スマートフォン、ネットブック、タブレットコンピュータ、ウェアラブルなどが挙げられ得る。いくつかの実施例では、クライアントデバイス101A~101Nのうちの1つ以上はそれぞれ、クライアントデバイスを一意に識別し得るクライアントデバイス識別子に割り当てられてもよい。いくつかの実施例では、クライアントデバイス識別子は、ASCIIテキスト、ポインタ、メモリアドレスなどを含んでもよい。

【0053】

クライアントデバイス101A~101Nのうちの1つのクライアントデバイスが、スマートフォン又はタブレットなどのモバイルデバイスである実施形態では、クライアントデバイスは、データ分析システム105と相互作用するための「アプリ」を実行してもよい。そのようなアプリは、典型的には、タブレット又はスマートフォンなどのモバイルデバイス上で実行するように設計されている。例えば、iOS(登録商標)、Android(登録商標)、又はWindows(登録商標)などのモバイルデバイスオペレーティングシステム上で実行するアプリが提供され得る。これらのプラットフォームは、典型的には、アプリが互いにモバイルデバイスの特定のハードウェア及びソフトウェア構成要素と通信することを可能にするフレームワークを提供する。例えば、上記のモバイルオペレーティングシステムは各々、場所サービス回路、有線及び無線ネットワークインターフェース、ユーザ連絡先、並びに他のアプリケーションと相互作用するためのフレームワークを提供する。アプリの外部で実行するハードウェア及びソフトウェアモジュールとの通信は、典型的には、モバイルデバイスオペレーティングシステムによって提供されるアプリケーションプログラミングインターフェース(Application Programming Interface: API)を介して提供される。いくつかの実施例では、アプリは、ユーザがデータ分析システム105と相互作用することを可能にし得るユーザインターフェースを提供してもよい。

【0054】

加えて、又はあるいは、クライアントデバイス101A~101Nは、ウェブブラウザを介してデータ分析システム105と相互作用してもよい。加えて、又はあるいは、クライアントデバイス101A~101Nは、データ分析システム105とインターフェースするように設計された様々なハードウェア又はファームウェアを含んでもよい。

【0055】

通信ネットワーク103としては、例えば、有線又は無線ローカルエリアネットワーク(LAN)、パーソナルエリアネットワーク(PAN)、メトロポリタンエリアネットワーク(MAN)、広域ネットワーク(WAN)などを含む、1つ以上の有線又は無線通信ネットワーク、並びに、1つ以上のネットワークを実装するのに必要な任意のハードウェア、ソフトウェア、及び/又はファームウェア(例えば、ネットワークルーターなど)が挙げられ得る。例えば、通信ネットワーク103としては、汎用パケット無線サービス(GPRS)ネットワーク、符号分割多元接続2000(CDMA2000)ネットワーク、広帯域符号分割多元接続(WCDMA)ネットワーク、Global System for Mobile Communications(GSM)ネットワーク、GSM進

10

20

30

40

50

化型高速データレート(EDGE)ネットワーク、Time Division - Synchronous Code Division Multiple Access(TD-SCDMA)ネットワーク、Long Term Evolution(LTE)ネットワーク、High Speed Packet Access(HSPA)ネットワーク、高速ダウンリンクパケットアクセス(HSDPA)ネットワーク、IEEE 802.11(Wi-Fi)、Wi-Fi Direct、IEEE 802.16(WiMAX)、及び/又は同様のものが挙げられ得る。加えて、又はあるいは、通信ネットワーク103としては、パブリックネットワーク(インターネットなど)、プライベートネットワーク(イントラネットなど)、又はこれらの組み合わせが挙げられ得る。

【0056】

いくつかの実施例では、通信ネットワーク103は、限定するものではないが、ハイパーテキスト転送プロトコル(HTTP)プロトコル、1つ以上の伝送制御プロトコル/インターネットプロトコル(TCP/IP)ベースのネットワークングプロトコル、近距離無線通信(NFC)プロトコル、Bluetoothプロトコル、及び/又はZigBeeプロトコルを含む、ネットワークングプロトコルを利用してもよい。例えば、ネットワークングプロトコルは、データ分析システム105の必要性に適合するようにカスタマイズされてもよい。いくつかの実施形態では、プロトコルは、Websocketチャンネルを介して送信されるJSONオブジェクトのカスタムプロトコルであってもよい。いくつかの実施形態では、プロトコルは、JSON-RPC、JSON-REST/HTTPなどであってもよい。

【0057】

図1を再び参照すると、データ分析システム105は、上述のコンピューティングデバイスとして具現化され得る。例えば、データ分析システム105は、少なくとも1つのプロセッサと、コンピュータプログラム命令を記憶する少なくとも1つの非一時的メモリと、を含んでもよい。これらのコンピュータプログラム命令は、少なくとも1つの非一時的メモリに記憶された命令が製品を製造し得るように、特定の方法で機能するようにデータ分析システム105に指示することができ、その実行によって本開示の実施形態が実施され得る。したがって、本開示のいくつかの実施例では、データ分析システム105は、データベースコネクタ107、データインタプリタ109、及び/又はデータマッパー111を含んでもよい。

【0058】

データベースコネクタ107、データインタプリタ109、及び/又はデータマッパー111は、ハードウェア手段(回路(単数又は複数)など)、ソフトウェア手段(コンピュータプログラムコードなど)、又はハードウェア手段とソフトウェア手段との組み合わせで具現化されてもよい。いくつかの実施例では、データベースコネクタ107は、1つ以上のデータベース(限定するものではないが、図1に示すようにデータベース113A~113Nを含む)から1つ以上のデータオブジェクトを取得するように構成されてもよい。いくつかの実施例では、データインタプリタ109は、1つ以上のデータ分類子を生成するように構成されてもよい。いくつかの実施例では、データマッパー111は、1つ以上のマッピング仕様を生成するように構成されてもよい。データ分析システム105の様々な例示的構成要素を示す例示的な図は、少なくとも図2に関連して示され、説明される。

【0059】

データ分析システム105内の様々な構成要素は、本開示の実施例に従って動作を実施するために、同じコンピュータ又はコンピューティング装置を活用してもよいことに留意されたい。例えば、データベースコネクタ107、データインタプリタ109、及び/又はデータマッパー111は、これらの機能を実施するために、同じプロセッサ又はメモリを活用してもよい。いくつかの実施例では、データベースコネクタ107、データインタプリタ109、及び/又はデータマッパー111は、分離された回路を活用してもよい。

【0060】

10

20

30

40

50

本開示の様々な実施形態では、限定するものではないが、データ分類要求及び/又はデータマッピング要求を含む1つ以上の電子要求が、データ分析システム105に送信されてもよい。いくつかの実施例では、これらの電子要求は、HTTP要求の形態であってもよい。いくつかの実施例では、これらの電子要求は、クライアントデバイス101A~101Nのうちの1つのクライアントデバイスによって、通信ネットワーク103を介してデータ分析システム105に直接送信されてもよい。加えて、又はあるいは、電子要求は、中間を介してデータ分析システム105に送信されてもよい。

【0061】

いくつかの実施例では、データ分類要求の受信時に、データ分析システム105は、1つ以上のデータ分類子を生成してもよい。いくつかの実施例では、1つ以上のデータ分類子に基づいて、データ分析システム105は、1つ以上のマッピング仕様を生成してもよい。

10

【0062】

図1を再び参照すると、データベース113A~113Nは、ネットワーク接続ストレージ(NAS)デバイス(単数又は複数)などの1つ以上のデータ記憶デバイスとして、又は別個のサーバー(単数又は複数)として具現化されてもよい。データベース113A~113Nは、データ分析システム105及び/又はクライアント装置101A~101Nにアクセス可能なデータ及び/又は情報を含んでもよい。

【0063】

いくつかの実施例では、データベース113A~113Nは、限定するものではないが、1つ以上のデータオブジェクトなどのデータを記憶してもよい。いくつかの実施例では、データ分類要求の受信時に、データ分析システム105は、データベース113A~113Nに電子要求を送信して、データベース113A~113Nから1つ以上のデータオブジェクトを取得又はフェッチしてもよい。いくつかの実施例では、データ分析システム105は、データベース113A~113Nに1つ以上のデータオブジェクトを記憶してもよい。

20

【0064】

データベース113A~113Nは、上述の動作を実施するために、同じコンピュータ又はコンピューティング装置を活用してもよいことに留意されたい。例えば、データベース113A~113Nは、データベース113A~113Nがデータ分析システム105の一部であり得るように、データ分析システム105内に統合されてもよい。いくつかの実施例では、データベース113A~113N及びデータ分析システム105は、分離された回路を利用してもよい。

30

【0065】

図1のデータ分析システム105は、図2に示すような装置200などの1つ以上のコンピューティングシステムによって具現化されてもよい。装置200は、プロセッサ202、メモリ204、入出力回路206、及び/又は通信回路208を含んでもよい。装置200は、図1に関して上述され、図3~7に関して後述される動作を実行するように構成されてもよい。

【0066】

これらの構成要素は、機能的制限に関して記載されているが、特定の実装は、特定のハードウェアの使用を必然的に含むことを理解されたい。これらの構成要素の特定のものは、同様の又は共通のハードウェアを含んでもよいことも理解されたい。例えば、回路の2つのセットは両方とも、同じプロセッサ、ネットワークインターフェース、記憶媒体などの使用を活用して、それらの関連する機能を実施するため、回路のそれぞれのセットに対して重複するハードウェアが不要である。

40

【0067】

いくつかの実施形態では、プロセッサ202(及び/又はプロセッサを補助するか、ないしは別の方法でプロセッサに関連付けられた任意の他の処理回路)は、装置の構成要素間で情報を渡すためのバスを介してメモリ204と通信してもよい。メモリ204は、

50

非一時的であり、例えば、1つ以上の揮発性及び/又は不揮発性メモリを含んでもよい。換言すれば、例えば、メモリ204は、電子記憶装置(例えば、コンピュータ可読記憶媒体)であってもよい。メモリ204は、装置が、本開示の例示的な実施形態による様々な機能を実行することを可能にするために、情報、データ、コンテンツ、アプリケーション、命令などを記憶するように構成されてもよい。

【0068】

図2に示す例では、メモリ204は、データベースコネクタモジュール210、データインタプリタモジュール212、及び/又はデータマッパーモジュール214を含み得るコンピュータプログラム命令を記憶してもよい。データベースコネクタモジュール210がプロセッサ202によって実行されるとき、装置200は、1つ以上のデータベース(例えば、限定するものではないが、図1に関連して上述したデータベース113A~113Nなど)から1つ以上のデータオブジェクトを取得するように構成されてもよい。データインタプリタモジュール212がプロセッサ202によって実行されるとき、装置200は、1つ以上のデータ分類子を生成するように構成されてもよい。データマッパーモジュール214がプロセッサ202によって実行されるとき、装置200は、1つ以上のマッピング仕様を生成するように構成されてもよい。

10

【0069】

加えて、又はあるいは、装置200は、データベースコネクタモジュール210、データインタプリタモジュール212、及び/又はデータマッパーモジュール214用に構成された1つ以上の指定されたハードウェアを含んでもよい。例えば、装置200は、データベースコネクタモジュール210、データインタプリタモジュール212、及び/又はデータマッパーモジュール214の機能を実行するように構成された、別個のプロセッサ、特別に構成されたフィールドプログラマブルゲートアレイ(FPGA)、又は特定用途向け集積回路(ASIC)を含んでもよい。

20

【0070】

図2を再び参照すると、プロセッサ202は、多くの異なる方法で具現化されてもよく、例えば、独立して実施するように構成された1つ以上の処理装置を含んでもよい。いくつかの実施例では、プロセッサ202は、命令、パイプライン、及び/又はマルチスレッドの独立した実行を可能にするためにバスを介してタンデム型に構成された1つ以上のプロセッサを含んでもよい。用語「プロセッサ」又は「処理回路」の使用は、シングルコアプロセッサ、マルチコアプロセッサ、装置内部の複数のプロセッサ、及び/又はリモートプロセッサ若しくは「クラウド」プロセッサを含むと理解されてもよい。

30

【0071】

上述のように、プロセッサ202は、メモリ204に記憶された命令を実行するか、ないしは別の方法でプロセッサ202にアクセス可能な命令を実行するように構成されてもよい。いくつかの好ましい、かつ非限定的な実施形態では、プロセッサ202は、ハードコードされた機能を実行するように構成されてもよい。したがって、ハードウェア方法又はソフトウェア方法によって構成されるか、又はそれらの組み合わせによって構成されるかにかかわらず、プロセッサ202は、それに応じて構成されている間、本開示の実施形態による動作を実施することができる(例えば、回路内で物理的に具現化された)エンティティを表してもよい。あるいは、別の例として、プロセッサ202がソフトウェア命令の実行体として具現化される場合、命令は、命令が実行されたときに本明細書に記載されるアルゴリズム及び/又は動作を実施するようにプロセッサ202を具体的に構成してもよい。

40

【0072】

通信回路208は、データを装置200と通信するネットワーク及び/又は任意の他の装置、回路、若しくはモジュールから受信する、かつ/又はそれらに送信するように構成されたハードウェア又はハードウェアとソフトウェアとの組み合わせのいずれかで具現化されたデバイス又は回路などの任意の手段であってもよい(図1に関連して上述されたクライアントデバイス101A~101N及び/又はデータベース113A~113Nな

50

ど)。この点に関して、通信回路 208 は、例えば、有線又は無線通信ネットワーク（図 1 に関連して上述した通信ネットワーク 103 など）との通信を可能にするためのネットワークインターフェースを含んでもよい。例えば、通信回路 208 は、1 つ以上のネットワークインターフェースカード、アンテナ、バス、スイッチ、ルーター、モデム、並びにサポートハードウェア及び/若しくはソフトウェア、又はネットワークを介した通信を可能にするのに好適な任意の他のデバイスを含んでもよい。加えて、又はあるいは、通信回路 208 は、アンテナ（単数又は複数）と相互作用して、アンテナを介した信号の伝送を引き起こすため、又はアンテナを介して受信した信号の受信を処理するための回路を含んでもよい。

【0073】

いくつかの実施形態では、装置 200 は、入出力回路 206 を任意選択的に含んでもよく、入出力回路 206 は、プロセッサ 202 と通信して、ユーザに出力を提供し、いくつかの実施形態では、ユーザ入力 of 指示を受信することができる。入出力回路 206 は、ユーザインターフェース回路を含んでもよく、ウェブユーザインターフェース、モバイルアプリケーション、クライアントデバイス、キオスクなどを含み得るディスプレイを含んでもよい。いくつかの実施形態では、入出力回路 206 はまた、キーボード、マウス、ジョイスティック、タッチスクリーン、タッチエリア、ソフトキー、マイクロフォン、スピーカ、又は他の入出力機構を含んでもよい。プロセッサ及び/又はプロセッサを含むユーザインターフェース回路は、プロセッサ（例えば、メモリ 204、及び/又は同様のもの）にアクセス可能なメモリ上に記憶されたコンピュータプログラム命令（例えば、ソフトウェア及び/又はファームウェア）を介して 1 つ以上のユーザインターフェース要素の 1 つ以上の機能を制御するように構成されてもよい。

【0074】

本明細書で論じる情報の全て又は一部は、装置 200 の 1 つ以上の構成要素によって受信、生成、及び/又は維持されるデータに基づき得ることに留意されたい。いくつかの実施形態では、1 つ以上の外部システム（リモートクラウドコンピューティング及び/又はデータストレージシステムなど）はまた、本明細書で論じられる機能性の少なくともいくつかを提供するために活用されてもよい。

【0075】

いくつかの実施形態では、装置 200 の他の要素は、特定の回路の機能性を提供又は補完してもよい。例えば、プロセッサ 202 は、処理機能を提供してもよく、メモリ 204 は、記憶機能を提供してもよく、通信回路 208 は、ネットワークインターフェース機能を提供してもよい。理解されるように、コンピュータ、プロセッサ、又はマシン上でコードを実行する他のプログラム可能回路が、本明細書に記載される機能を含む様々な機能を実施するための手段を生むことができるように、任意のかかるコンピュータプログラム命令及び/又は他の種類のコードが、コンピュータ、プロセッサ、又は他のプログラム可能な装置の回路上にロードされて、機械が製造されてもよい。

【0076】

上述したように、かつ本開示に基づいて理解されるように、本開示の実施形態は、方法、モバイルデバイス、バックエンドネットワークデバイスなどとして構成されてもよい。したがって、実施形態は、ハードウェア全体、又はソフトウェアとハードウェアとの任意の組み合わせを含む様々な手段を含んでもよい。

【0077】

ここで図 3 ~ 図 7 を参照すると、本開示の様々な実施形態による例示的な方法が示されている。いくつかの実施例では、フローチャートのそれぞれのブロック又は工程、並びにフローチャート内のブロック及び/又は工程の組み合わせは、ハードウェア、回路、及び/又は 1 つ以上のコンピュータプログラム命令を含むソフトウェアの実行に関連付けられた他のデバイスなどの様々な手段によって実施されてもよい。

【0078】

いくつかの実施例では、図に記載される手順のうちの 1 つ以上は、本開示の実施形態

10

20

30

40

50

を用いる装置のメモリ回路（非一時的メモリなど）によって記憶され、装置の処理回路（プロセッサなど）によって実行され得るコンピュータプログラム命令によって具現化されてもよい。これらのコンピュータプログラム命令は、メモリ回路に記憶された命令によって製品が製造され得るように、装置に対して特定の方法で機能するように指示してもよく、その実行によって、フロー図のブロック（単数又は複数）内に指定された機能が実装され得る。更に、装置は、例えば、通信回路及び/又は入出力回路などの1つ以上の他の構成要素を含んでもよい。装置の様々な構成要素は、互いにデータを送信し、及び/又は互いからデータを受信するために、互いの間及び/又は互いの中で電子通信してもよい。

【0079】

いくつかの実施例では、実施形態は、コンピュータ可読プログラム命令（例えば、コンピュータソフトウェア）を記憶する非一時的コンピュータ可読記憶媒体上のコンピュータプログラム製品の形態をとってもよい。非一時的ハードディスク、CD-ROM、フラッシュメモリ、光学記憶デバイス、及び/又は磁気記憶デバイスを含む、任意の好適なコンピュータ可読記憶媒体が利用されてもよい。

10

【0080】

ここで図3を参照すると、本開示のいくつかの実施形態による例示的な方法300が示されている。具体的には、例示的な方法300は、データオブジェクトを分類し、データオブジェクトを変換するためのマッピング仕様を生成する例示的な実施形態を示し得る。いくつかの実施例では、方法300は、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202）によって実施されてもよい。

20

【0081】

方法300は、ブロック301で開始する。

【0082】

ブロック303では、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202）は、第1の複数のデータオブジェクトを取得し得る。いくつかの実施例では、第1の複数のデータオブジェクトは、第1のデータベーススキーマに関連付けられてもよい。

【0083】

いくつかの実施例では、処理回路は、クライアントデバイス（限定するものではないが、図1に示すように、クライアントデバイス101A~101Nのうちの1つなど）からデータ分類要求を受信してもよい。データ分類要求は、1つ以上のデータオブジェクト識別子及び/又はデータ分類子を生成する要求を含んでもよい。

30

【0084】

上述のように、データオブジェクト識別子は、1つ以上のデータベース（限定するものではないが、図1に示されるようにデータベース113A~113Nなど）からデータオブジェクトを識別及び/又は検索し得る。データ分類要求の受信時に、処理回路は、データ検索要求（例えば、1つ以上のデータオブジェクト識別子を含んでもよい）を1つ以上のデータベース（限定するものではないが、図1に示されるようにデータベース113A~113Nなど）に送信してもよい。処理回路は、データ取得要求に回答して、1つ以上のデータベースからデータオブジェクト識別子に関連付けられた1つ以上のデータオブジェクトを受信してもよい。

40

【0085】

いくつかの実施例では、処理回路は、データベースコネクタ構成要素（例えば、限定するものではないが、図1に示されるようにデータベースコネクタ107）を含んでもよい。データベースコネクタ構成要素は、1つ以上のデータベースからデータオブジェクト及び/又はデータベーススキーマ情報をフェッチしてもよく、データオブジェクト及び/又はデータベーススキーマ情報をデータインタプリタ構成要素（例えば、限定するものではないが、図1に示すようにデータインタプリタ109）に送信してもよい。

【0086】

50

ブロック 305 では、処理回路（例えば、図 1 に関連して説明されるデータ分析システム 105 の処理回路、及び / 又は図 2 に関連して説明される装置 200 のプロセッサ 202）は、第 1 のデータベーススキーマに対応する第 1 のデータ分類子を決定し得る。いくつかの実施例では、処理回路は、ブロック 303 で取得された第 1 の複数のデータオブジェクトに少なくとも基づいて、第 1 のデータ分類子を決定してもよい。

【0087】

いくつかの実施例では、処理回路は、データインタプリタ構成要素（例えば、限定するものではないが、図 1 に示されるようにデータインタプリタ 109）を含んでもよい。データインタプリタ構成要素は、データベースコネクタ構成要素（例えば、限定するものではないが、図 1 に示されるようにデータベースコネクタ 107）からデータを受信してもよい。

10

【0088】

上述したように、データ分類子は、分類を示し、記述を含み、かつ / 又は 1 つ以上のデータオブジェクトに関連付けられた情報の意味を提供し得る ASCII テキスト、ポインタ、メモリアドレスなどの形態であってもよい。いくつかの実施例では、処理回路は、機械学習モデル、知的エージェント（IA）、及び / 又は人工知能（AI）ツールを適用して、データ分類子を決定してもよい。

【0089】

いくつかの実施例では、処理回路は、人工ニューラルネットワークを実装して、第 1 のデータ分類子を決定してもよい。例示的な人工ニューラルネットワークは、複数の相互接続されたノードを含んでもよく、それぞれのノードは、（ノードから受信した）入力に基づいて（ノードへの）出力を生成し得る数学関数を表し得る。複数のノードは、入力層、1 つ以上の中間層、及び出力層などの層に分割されてもよい。

20

【0090】

一例として、以下のデータオブジェクト（データテーブルの形態）が、処理回路によって例示的な人工ニューラルネットワークへの入力として提供され得る。

【表 3】

John	Doe	Jr.	John.doe@email.com
Adam	Davis		Adam.davis@email.com
Richard	Roe	II	Richard.roe@email.com

30

この例では、人工ニューラルネットワークは、データテーブル内のそれぞれのデータフィールド（例えば、「John」、「Doe」、「Jr.」、「John.doe@email.com」）に対してノードを生成し得る。相互接続されたノード及びそれらの関連する数学関数を通じて、人工ニューラルネットワークは、データテーブルのそれぞれの行及び / 又は列の分類を示す 1 つ以上のデータ分類子を出力し得る。例えば、人工ニューラルネットワークは、データテーブル用のデータ分類子（例えば、テキスト文字列の形態であってもよい）を生成し得る。いくつかの実施例では、データ分類子は、第 1 の列に関連付けられたデータフィールドがファーストネームであり、第 2 の列に関連付けられたデータフィールドがラストネームであり、第 3 の列に関連付けられたデータフィールドが名前の接尾辞であり、及び / 又は、第 4 の列に関連付けられたデータフィールドが電子メールアドレスであることを示してもよい。

40

【0091】

いくつかの実施例では、処理回路は、第 1 のデータ分類子を決定するために決定木アルゴリズムを実施してもよい。例示的な決定木は、1 つ以上の葉を含んでもよく、それぞれの葉は、例えば、データの可能な分類を表し得る。加えて、又はあるいは、決定木は、例えば、分類の可能な組み合わせ（すなわち、決定木の葉）を表し得る 1 つ以上の枝を含んでもよい。

【0092】

例えば、決定木アルゴリズムは、データテーブルのそれぞれの列の 1 つ以上の分類を

50

決定し得る。上記の例示的なデータテーブルから継続して、処理回路は、第 1 の列に関連付けられたデータフィールドを、例示的な決定木アルゴリズムへの入力として提供してもよい。例示的な決定木アルゴリズムは、これらのデータフィールドが名前（例えば、決定木内の上位ノードとしての「名前」分類）を表すことを決定してもよく、それらがファーストネームを表すか又はラストネームを表すか（例えば、決定木内の最上位ノードの下位ノードとしての「ファーストネーム」下位分類及び「ラストネーム」下位分類）を更に決定してもよい。第 1 の列に関連付けられたデータフィールドに基づいて、決定木アルゴリズムは、これらのデータフィールドがファーストネームを表す第 1 の確率、及びこれらのデータフィールドがラストネームを表す第 2 の確率を計算し得る。決定木アルゴリズムは、第 1 の確率を第 2 の確率と比較してもよく、第 1 の確率が第 2 の確率よりも高いことを決定し得る。この例では、決定木アルゴリズムは、出力（第 1 の列がファーストネームを表すことを示し得る）を生成してもよく、処理回路は、決定木アルゴリズムからの出力に基づいてデータ分類子を決定してもよい。

10

【 0 0 9 3 】

いくつかの実施例では、処理回路は、第 1 のデータ分類子を決定するために教師あり学習モデルを実施してもよい。例示的な教師あり学習モデルでは、モデルによって受信された入力は、例示的な入出力ペア（例えば、トレーニングデータ）に基づいて出力にマッピングされてもよい。例示的な教師あり学習モデルは、トレーニングデータを解析してもよく、例示的な入出力ペアから 1 つ以上の関数を推測してもよい。例示的な教師あり学習モデルは、推測された関数を利用して 1 つ以上の出力を生成してもよい。例示的な教師あり学習モデルとしては、限定するものではないが、サポートベクターマシンが挙げられ得る。

20

【 0 0 9 4 】

上記の例示的なデータテーブルから続けると、例示的な名前の接尾辞は、トレーニングデータとして例示的な教師あり学習モデルに提供されてもよい。例えば、処理回路は、例示的な教師あり学習モデルに例示的な名前の接尾辞 I I、I I I、I V、J r .、S r .、M D、P h D を提供してもよい。教師あり学習モデルは、それぞれの例示的な名前の接尾辞を名前の接尾辞分類と接続してもよい。教師あり学習モデルが、上記の例示的なデータテーブルからデータフィールド「J r .」を受信したとき、教師あり学習モデルは、データフィールドが名前の接尾辞に関連付けられていることを示す出力を生成してもよく、処理回路は、出力に基づいてデータ分類子を生成してもよい。

30

【 0 0 9 5 】

上記の例は、人工ニューラルネットワーク、決定木アルゴリズム、及び/又は教師あり学習モデルに基づいて例示的なデータ分類子を生成することを示しているが、本開示の範囲は、これらの機構のみに限定されないことに留意されたい。いくつかの実施例では、限定するものではないが、ベネジアンネットワーク、遺伝的アルゴリズム、回帰モデル、及び/又はランダムフォレストを含む、他の機械学習モデル、I A、及び/又は A I ツールが、データ分類子を決定するために追加的又は代替的に利用されてもよい。

【 0 0 9 6 】

いくつかの実施例では、処理回路は、データオブジェクトのデータフィールドに加えて、又はその代わりに、データオブジェクトに関連付けられたメタデータに基づいて第 1 のデータ分類子を決定してもよい。例えば、処理回路は、第 1 の複数のデータオブジェクトに関連付けられた名前メタデータ、列メタデータ、行メタデータ、相関メタデータ、及び/又はドメインメタデータに少なくとも部分的に基づいて、第 1 のデータ分類子を決定することができ、その例の詳細は、少なくとも図 4、図 5、及び図 6 に関連して説明される。

40

【 0 0 9 7 】

ブロック 3 0 7 では、処理回路（例えば、図 1 に関連して説明されるデータ分析システム 1 0 5 の処理回路、及び/又は図 2 に関連して説明される装置 2 0 0 のプロセッサ 2 0 2）は、マッピング仕様を生成し得る。いくつかの実施例では、マッピング仕様は、第

50

1のデータベーススキーマに関連付けられた第1の複数のデータオブジェクトを、第2のデータベーススキーマに関連付けられた第2の複数のデータオブジェクトに変換するように構成されてもよい。

【0098】

いくつかの実施例では、処理回路は、クライアントデバイス（限定するものではないが、図1に示すように、クライアントデバイス101A~101Nのうちの1つなど）からデータマッピング要求を受信してもよい。データマッピング要求は、第1のデータベーススキーマに関連付けられた第1のデータベース識別子と、第2のデータベーススキーマに関連付けられた第2のデータベース識別子と、第1のデータベーススキーマから第2のデータベーススキーマにデータオブジェクトを変換するための要求と、を含んでもよい。

10

【0099】

上述のように、データベース識別子は、データベースに関連付けられたデータベーススキーマ（限定するものではないが、図1に示されるようなデータベース113A~113Nなど）を識別及び/又は検索し得る。データマッピング要求の受信時に、処理回路は、例えば、ブロック303及びブロック305に関連して上述したものと同様に、第1のデータベーススキーマに関連付けられたデータオブジェクトを取得し、データ分類子を決定してもよい。いくつかの実施例では、データマッピング要求は、データ分類要求と組み合わせられてもよく、クライアントデバイスは、処理回路をトリガしてデータ分類子及びマッピング仕様を生成するために、1つの電子要求を送信してもよい。

【0100】

20

いくつかの実施例では、処理回路は、第1の複数のデータオブジェクト（ブロック303で取得された）及び第1のデータ分類子（ブロック305で決定された）に少なくとも部分的に基づいてマッピング仕様を生成してもよい。上述のように、用語「マッピング仕様」は、1つのデータモデルに関連付けられたデータは、別のデータモデルに基づいてどのように表され得るかを決定し得る、動作及び/又は変換を記述する、かつ/又は指定することができるデータオブジェクトを指し得る。

【0101】

以下のデータテーブルに関連する実施例から続ける。

【表4】

	(列1)	(列2)	(列3)	(列4)
(行1)	John	Doe	Jr.	John.doe@email.com
(行2)	Adam	Davis		Adam.davis@email.com
(行3)	Richard	Roe	II	Richard.roe@email.com

30

ブロック305で生成された例示的なデータ分類子は、列1に関連付けられたデータフィールドがファーストネームであり、列2に関連付けられたデータフィールドがラストネームであり、列3に関連付けられたデータフィールドが名前の接尾辞であり、かつ/又は、列4に関連付けられたデータフィールドが電子メールアドレスであることを示してもよい。一例として、処理回路によって受信されたデータマッピング要求は、上記のソースデータテーブルからのデータを、異なるタイプのデータテーブル及び/又は異なるデータベーススキーマに基づいたデータオブジェクトに変換するための要求を示し得る。例えば、データマッピング要求は、上記のソースデータテーブルを、共通データモデルに準拠するターゲットデータオブジェクトに変換するための要求を含んでもよい。

40

【0102】

上述のように、処理回路は、データ分類子（例えば、ブロック305で決定されたデータ分類子）に少なくとも部分的に基づいてマッピング仕様を生成してもよい。上記の例から続けると、データ分類子は、ソースデータテーブルの列4に関連付けられたデータフィールドが電子メールアドレスであることを示し得る。処理回路は、ターゲットデータオブジェクト（例えば、共通データモデルに基づくデータテーブル）が、電子メールアドレスのデータフィールドを含み得ることを決定し得る。処理回路は、ソースデータテーブル

50

の列 1 をターゲットデータオブジェクト内の電子メールアドレスの対応するデータフィールドに変換するように、マッピング仕様内にルールステートメントを生成してもよい。

【 0 1 0 3 】

いくつかの実施例では、処理回路は、データ分類子に基づいて 1 つ以上のデータフィールドを組み合わせるように、マッピング仕様内に 1 つ以上のルールステートメントを生成してもよい。上記の例では、データ分類子は、列 1、列 2、及び列 3 が名前に関連付けられていることを示し得る。データ分類子に基づいて、処理回路は、それぞれの行の列 1、列 2、及び列 3 のデータフィールドを組み合わせてもよく、それらをターゲットデータテーブル内の名前の対応するデータフィールドに提供してもよい。

【 0 1 0 4 】

いくつかの実施例では、処理回路は、データフィールドをデータ分類子に基づいて複数のデータフィールドに分割又は複製するように、マッピング仕様内に 1 つ以上のルールステートメントを生成してもよい。例えば、データ分類子が、データフィールドは Y Y Y Y - M M - D D 形式の日付を含むことを示し、処理回路が、ターゲットデータテーブルは年、月、及び日の分離された列を含むことを決定した場合、処理回路は、データフィールドを分割して年値、月値、及び日値を単離するように、1 つ以上のルールステートメントを生成してもよい。

【 0 1 0 5 】

上記の例は、データテーブル間のデータフィールドの変換 (converting) 及び変換 (transforming) に関連付けられた例示的なマッピング仕様を示し得るが、本開示の範囲はデータテーブルに限定されないことに留意されたい。いくつかの実施例では、処理回路は、他の種類 (単数又は複数) のデータオブジェクトに基づいてマッピング仕様を生成してもよい。

【 0 1 0 6 】

例えば、複数のデータオブジェクトは、1 つ以上の A S C I I 文字を含み得るテキスト文書を含んでもよい。ブロック 3 0 5 に関連して上述したように、処理回路は、テキスト文書に基づいて第 1 のデータ分類子を生成してもよい。一例として、処理回路は、テキスト文書上の自然言語処理を行うために機械学習モデルを実施してもよい。処理回路は、自然言語処理の結果に基づいて、テキスト文書の意味を示し得るデータ分類子を生成してもよい。

【 0 1 0 7 】

非限定的な例として、データ分類子は、テキスト文書が 1 つ以上のデバイスのバッテリーレベル値を記述し得ることを示してもよい。処理回路は、データ分類子に少なくとも部分的に基づいて、ソーステキスト文書内のバッテリーレベル値を、ターゲットデータオブジェクト内のデータフィールドに変換するように構成され得るマッピング仕様を生成してもよい。

【 0 1 0 8 】

いくつかの実施例では、処理回路は、マッピング仕様を生成するために、機械学習モデル、I A、及び / 又は A I ツールを実施してもよい。例えば、処理回路は、最も効率的なマッピング仕様を決定するために、人工ニューラルネットワークを実施してもよい。この実施例では、処理回路は、ルールステートメントで指定されるデータ動作に基づいて、例示的な人工ニューラルネットワーク内のノードを生成してもよい。次いで、処理回路は、最も効率的なマッピング仕様を決定するために、最も少ない動作を必要とし得る例示的な人工ニューラルネットワーク内の経路を計算してもよい。いくつかの実施例では、他の機械学習モデル、I A、及び / 又は A I ツールは、マッピング仕様を生成するための処理回路によって利用されてもよい。

【 0 1 0 9 】

ブロック 3 0 9 では、処理回路 (例えば、図 1 に関連して説明されるデータ分析システム 1 0 5 の処理回路、及び / 又は図 2 に関連して説明される装置 2 0 0 のプロセッサ 2 0 2) は、第 2 の複数のデータオブジェクトを生成し得る。

10

20

30

40

50

【 0 1 1 0 】

いくつかの実施例では、処理回路は、第 1 の複数のデータオブジェクト及びマッピング仕様に少なくとも部分的に基づいて、第 2 の複数のデータオブジェクトを生成してもよい。例えば、処理回路は、マッピング仕様のルールステートメントに基づいて、ソースデータオブジェクト内のデータフィールドをターゲットデータオブジェクトのデータフィールドに変換してもよい。

【 0 1 1 1 】

名前及び電子メールアドレスに関連する上記の例示的なソースデータテーブルから続けると、処理回路は、マッピング仕様におけるルールステートメントに基づいて、ソースデータテーブルをターゲットデータテーブルに変換してもよい。上述のように、ルールステートメントは、データ分類子に基づいて生成されてもよい。この実施例では、処理回路は、ルールステートメントに基づいて、ソースデータテーブルからそれぞれの行の列 1、列 2、及び列 3 に関連付けられた複合データフィールド (combine data fields) を (例えば、行識別子に基づいて) 生成し、ターゲットデータオブジェクトの複数の第 1 のデータフィールドに対して複合データフィールドを入力し得る。加えて、又はあるいは、処理回路は、ソースデータテーブルの列 4 から、ターゲットデータオブジェクトの複数の第 2 データフィールドにデータフィールドを変換してもよい。

10

【 0 1 1 2 】

上記の実施例は、例示的なデータオブジェクトとして例示的なデータテーブルを示しているが、本開示の範囲はデータテーブルに限定されないことに留意されたい。いくつかの実施例では、複数のデータオブジェクトは、ログ、数字文字列などを含んでもよい。

20

【 0 1 1 3 】

方法 3 0 0 は、ブロック 3 1 1 で終了する。

【 0 1 1 4 】

ここで図 4、図 5、及び図 6 を参照すると、本開示のいくつかの実施形態による例示的な方法が示されている。具体的には、これらの例示的な方法は、データ分類子 (例えば、上述した図 3 のブロック 3 0 5 に関連し得る) を決定する例示的な実施形態を示し得る。

【 0 1 1 5 】

ここで図 4 を参照すると、本開示のいくつかの実施形態による例示的な方法 4 0 0 が示されている。具体的には、例示的な方法 4 0 0 は、データベース内のデータオブジェクトに関連付けられたメタデータに少なくとも基づいてデータ分類子を決定する例示的な実施形態を示し得る。

30

【 0 1 1 6 】

いくつかの実施例では、方法 4 0 0 は、処理回路 (例えば、図 1 に関連して説明されるデータ分析システム 1 0 5 の処理回路、及び / 又は図 2 に関連して説明される装置 2 0 0 のプロセッサ 2 0 2) によって実施されてもよい。

【 0 1 1 7 】

方法 4 0 0 は、ブロック A から開始してもよい。再び図 3 を参照すると、ブロック A は、第 1 の複数のデータオブジェクトを取得した (ブロック 3 0 3) 後であってもよい。

【 0 1 1 8 】

ブロック 4 0 2 では、処理回路 (例えば、図 1 に関連して説明されるデータ分析システム 1 0 5 の処理回路、及び / 又は図 2 に関連して説明される装置 2 0 0 のプロセッサ 2 0 2) は、複数のデータオブジェクト (例えば、図 3 に関連して説明される第 1 の複数のデータオブジェクト) に関連付けられたメタデータを取得し得る。

40

【 0 1 1 9 】

図 4 に示される実施例では、処理回路によって取得されたメタデータは、第 1 の複数のデータオブジェクトに関連付けられた名前メタデータ 4 0 4、列メタデータ 4 0 6、又は行メタデータ 4 0 8 のうちの少なくとも 1 つを含んでもよい。例えば、名前メタデータ、列メタデータ、及び / 又は行メタデータは、データテーブルに関連付けられてもよい。

【 0 1 2 0 】

50

いくつかの実施例では、名前メタデータ 404 は、ASCII テキスト、ポインタ、メモリアドレスなどの形態であってもよく、データオブジェクト（データテーブルなど）の名前を含んでもよい。例えば、名前メタデータ 404 は、データテーブルが消費者情報に関連付けられていることを示してもよい。

【0121】

いくつかの実施例では、列メタデータ 406 は、ASCII テキスト、ポインタ、メモリアドレスなどの形態であってもよく、列識別子及び/又は列の名前を含んでもよい。例えば、列メタデータ 406 は、列に関連付けられたデータフィールドが、消費者の名前に関連することを示してもよい。

【0122】

いくつかの実施例では、行メタデータ 408 は、ASCII テキスト、ポインタ、メモリアドレスなどの形態であってもよく、行識別子及び/又は行の名前を含んでもよい。例えば、行メタデータ 408 は、行に関連付けられたデータフィールドが、消費者識別子に関連することを示してもよい。

【0123】

ブロック 410 では、処理回路（例えば、図 1 に関連して説明されるデータ分析システム 105 の処理回路、及び/又は図 2 に関連して説明される装置 200 のプロセッサ 202）は、メタデータに基づいて第 1 のデータ分類子を決定してもよい。例えば、処理回路は、少なくとも、名前メタデータ 404、列メタデータ 406、及び/又は行メタデータ 408 のうちの少なくとも 1 つに基づいて、第 1 のデータ分類子を決定してもよい。

【0124】

上述のように、処理回路は、データ分類子を決定するために、機械学習モデル、IA、及び/又は AI ツールを適用してもよい。いくつかの実施例では、ブロック 402 で取得されたメタデータは、機械学習モデル、IA、及び/又は AI ツールへの入力データとして提供されてもよい。例えば、名前メタデータ 404 は、データテーブルが消費者情報に関連付けられていることを示してもよく、列メタデータ 406 は、列に関連付けられたデータフィールドが消費者の名前に関連することを示してもよく、かつ/又は行メタデータ 408 は、行に関連付けられたデータフィールドが消費者識別子に関連することを示してもよい。処理回路は、図 4 に関連して上述したものと同様の、名前メタデータ 404、列メタデータ 406、及び/又は行メタデータ 408 のうちの少なくとも 1 つに少なくとも

【0125】

ブロック 410 に続いて、方法 400 はブロック B に戻ってもよい。図 3 に示すように、ブロック B は、ブロック 307 においてマッピング仕様を生成する前であってもよい。いくつかの実施例では、処理回路は、図 4 に関連して説明されるように、データ分類子に少なくとも部分的に基づいてマッピング仕様を生成してもよい。

【0126】

ここで図 5 を参照すると、本開示のいくつかの実施形態による例示的な方法 500 が示されている。具体的には、例示的な方法 500 は、データベース内のデータオブジェクトに関連付けられた関連メタデータに少なくとも基づいてデータ分類子を決定する例示的な実施形態を示し得る。

【0127】

いくつかの実施例では、方法 500 は、処理回路（例えば、図 1 に関連して説明されるデータ分析システム 105 の処理回路、及び/又は図 2 に関連して説明される装置 200 のプロセッサ 202）によって実施されてもよい。

【0128】

方法 500 は、ブロック A から開始してもよい。再び図 3 を参照すると、ブロック A は、第 1 の複数のデータオブジェクトを取得した（ブロック 303）後であってもよい。

【0129】

10

20

30

40

50

ブロック501では、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202）は、相関メタデータを決定し得る。いくつかの実施例では、相関メタデータは、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよく、2つ以上のデータオブジェクト間の関係を示してもよい。

【0130】

いくつかの実施例では、相関メタデータは、第1のデータテーブルと第2のデータテーブルとの間の関係を示してもよい。例えば、相関メタデータは、第1のデータテーブル及び第2のデータテーブルの両方が消費者情報に関連付けられていることを示してもよい。加えて、又はあるいは、相関メタデータは、第1のデータテーブルが消費者名に関連付けられたデータフィールドを含み得、第2のデータテーブルが消費者電子メールアドレスに関連付けられたデータフィールドを含み得ることを示してもよい。

10

【0131】

いくつかの実施例では、処理回路は、相関データを決定するために、機械学習モデル、IA、及び/又はAIツールを実施してもよい。例えば、処理回路は、第1のデータテーブルと第2のデータテーブルとの間の関係を決定するために教師あり学習モデルを実施してもよい。

【0132】

ブロック503では、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202）は、相関メタデータに基づいて第1のデータ分類子を決定し得る。

20

【0133】

上述のように、処理回路は、データ分類子を決定するために、機械学習モデル、IA、及び/又はAIツールを適用してもよい。いくつかの実施例では、ブロック501で決定された相関メタデータは、機械学習モデル、IA、及び/又はAIツールへの入力データとして処理回路によって提供されてもよい。処理回路は、図4に関連して上述したものと同様の、相関メタデータに少なくとも基づいて、第1のデータ分類子を決定するために、機械学習モデル、IA、及び/又はAIツールを実施してもよい。

【0134】

ブロック503に続いて、方法500はブロックBに戻ってもよい。図3に示すように、ブロックBは、ブロック307においてマッピング仕様を生成する前であってもよい。いくつかの実施例では、処理回路は、図5に関連して説明されるように、データ分類子に少なくとも部分的に基づいてマッピング仕様を生成してもよい。

30

【0135】

ここで図6を参照すると、本開示のいくつかの実施形態による例示的な方法600が示されている。具体的には、例示的な方法600は、ドメインメタデータに少なくとも基づいてデータ分類子を決定する例示的な実施形態を示し得る。

【0136】

いくつかの実施例では、方法600は、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202）によって実施されてもよい。

40

【0137】

方法600は、ブロックAから開始してもよい。再び図3を参照すると、ブロックAは、第1の複数のデータオブジェクトを取得した（ブロック303）後であってもよい。

【0138】

ブロック602では、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202）は、ドメインメタデータを決定し得る。いくつかの実施例では、ドメインメタデータは、ASCIIテキスト、ポインタ、メモリアドレスなどの形態であってもよく、データオブジェクト（限定するものではないが、第1のデータテーブルなど）に関連付けられ

50

た主題ドメインを示してもよい。

【0139】

用語「ドメイン」又は「主題ドメイン」は、複数のデータオブジェクト間の共通の属性及び／又は機能のセットを指し得る。いくつかの実施例では、同じドメインに関連付けられたデータオブジェクトは、共通のエンティティに関連付けられた知識又は活動の領域を表し得る。例えば、名前、電子メールアドレス、及び電話番号を表すデータオブジェクトは、消費者情報のドメインにグループ化されてもよい。

【0140】

いくつかの実施例では、処理回路は、ドメインデータを決定するために、機械学習モデル、IA、及び／又はAIツールを実施してもよい。例えば、処理回路は、第1のデータテーブルに関連付けられた対応するドメインを決定するために、人工ニューラルネットワークを実施してもよい。

10

【0141】

ブロック604では、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び／又は図2に関連して説明される装置200のプロセッサ202）は、ドメインメタデータに基づいて第1のデータ分類子を決定し得る。

【0142】

上述のように、処理回路は、データ分類子を決定するために、機械学習モデル、IA、及び／又はAIツールを適用してもよい。いくつかの実施例では、ブロック602で決定されたドメインメタデータは、機械学習モデル、IA、及び／又はAIツールへの入力データとして処理回路によって提供されてもよい。処理回路は、図4に関連して上述したものと同様の、ドメインメタデータに少なくとも基づいて、第1のデータ分類子を決定するために、機械学習モデル、IA、及び／又はAIツールを実施してもよい。

20

【0143】

ブロック604に続いて、方法600はブロックBに戻ってもよい。図3に示すように、ブロックBは、ブロック307においてマッピング仕様を生成する前であってもよい。いくつかの実施例では、処理回路は、図6に関連して説明されるように、データ分類子に少なくとも部分的に基づいてマッピング仕様を生成してもよい。

【0144】

図4、図5、及び図6は、名前メタデータ、列メタデータ、行メタデータ、相関メタデータ、及び／又はドメインメタデータに基づいてデータ分類子を決定する例示的な方法を示すが、本開示の範囲は、これらのメタデータに限定されないことに留意されたい。いくつかの実施例では、上記のこれらのメタデータに加えて、又はその代わりに、他のメタデータがデータ分類子を決定するために使用されてもよい。

30

【0145】

ここで図7を参照すると、本開示のいくつかの実施形態による例示的な方法700が示されている。具体的には、例示的な方法700は、データオブジェクトを分類し、データオブジェクトを変換するためのマッピング仕様を生成する例示的な実施形態を示し得る。いくつかの実施例では、方法700は、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び／又は図2に関連して説明される装置200のプロセッサ202）によって実施されてもよい。

40

【0146】

方法700は、工程701で開始する。

【0147】

ブロック703では、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び／又は図2に関連して説明される装置200のプロセッサ202）は、第1のデータ分類子を決定し得る。

【0148】

いくつかの実施例では、第1のデータ分類子は、第1の複数のデータオブジェクトに関連付けられてもよい。いくつかの実施例では、処理回路は、図3、図4、図5、及び／

50

又は図 6 に関連して説明されるものと同様の方法に基づいて第 1 のデータ分類子を生成してもよい。

【 0 1 4 9 】

ブロック 7 0 5 では、処理回路（例えば、図 1 に関連して説明されるデータ分析システム 1 0 5 の処理回路、及び / 又は図 2 に関連して説明される装置 2 0 0 のプロセッサ 2 0 2 ）は、第 1 のデータ分類子に関連付けられた信頼スコアを計算し得る。

【 0 1 5 0 】

「信頼スコア」は、対応するデータが正しい及び / 又は真の状態を表す可能性を示し得る数学的な値を指し得る。例えば、データ分類子に関連付けられた信頼スコアは、データ分類子がデータオブジェクト（単数又は複数）の正しい分類及び / 又は正しい記述を含む可能性を示し得る。

10

【 0 1 5 1 】

図 3、図 4、図 5、及び図 6 に関連して上述したように、処理回路は、例えば、機械学習モデル、I A、及び / 又は A I ツールに基づいて第 1 の分類子を決定してもよい。いくつかの実施例では、処理回路は、データ分類子を決定するために使用されるもの（単数又は複数）と同じ又は異なる機械学習モデル、I A、及び / 又は A I ツールを使用して信頼スコアを生成してもよい。

【 0 1 5 2 】

一例として、処理回路が教師あり学習モデルを利用して第 1 のデータ分類子を決定するとき、処理回路は、データオブジェクトのデータフィールドをトレーニングデータ内の入出力ペアと比較してもよい。処理回路は、データフィールドとトレーニングデータとの間の類似性レベルを示し得る近接値を計算してもよい。類似性が高いほど、教師あり学習モデルがデータオブジェクトの正しい分類を生成する可能性は高くなる。したがって、処理回路は、近接値に基づいて信頼スコアを生成してもよい。

20

【 0 1 5 3 】

加えて、又はあるいは、処理回路は、信頼スコアを計算するために他の機械学習モデル、I A、及び / 又は A I ツールを利用してもよい。

【 0 1 5 4 】

ブロック 7 0 7 では、処理回路（例えば、図 1 に関連して説明されるデータ分析システム 1 0 5 の処理回路、及び / 又は図 2 に関連して説明される装置 2 0 0 のプロセッサ 2 0 2 ）は、信頼スコアが所定の閾値を満たしているかどうかを決定し得る。

30

【 0 1 5 5 】

いくつかの実施例では、処理回路は、例えばシステム要件に基づいて所定の閾値を決定してもよい。例えば、例示的なシステム要件は、データマッピングに必要とされる精度のレベルを示し得る。この例では、必要とされる精度のレベルが高いほど、所定の閾値の値は高くなる。

【 0 1 5 6 】

いくつかの実施例では、所定の閾値は、ユーザによって設定されてもよい。例えば、ユーザは（クライアント装置を動作させている間）、処理回路にユーザ入力を提供することができる。ユーザ入力は、所定の閾値に対応し得る値を含んでもよい。

40

【 0 1 5 7 】

再び図 7 を参照すると、ブロック 7 0 7 において、処理回路が、信頼スコアが所定の閾値を満たしていると決定した場合、方法 7 0 0 はブロック 7 0 9 に進み得る。

【 0 1 5 8 】

一例として、ブロック 7 0 5 で計算された信頼スコアが 0 . 8 であり、所定の閾値が 0 . 6 である場合、処理回路は、信頼スコアが閾値よりも高く、したがって、信頼スコアが所定の閾値を満たしていると決定し得る。

【 0 1 5 9 】

ブロック 7 0 9 では、処理回路（例えば、図 1 に関連して説明されるデータ分析システム 1 0 5 の処理回路、及び / 又は図 2 に関連して説明される装置 2 0 0 のプロセッサ 2

50

02)は、信頼スコアが所定の閾値を満たしているという決定にตอบสนองしてマッピング仕様を生成し得る。

【0160】

いくつかの実施例では、処理回路は、図3に関連して上述したものと同様に、第1のデータ分類子に少なくとも部分的に基づいてマッピング仕様を生成してもよい。

【0161】

再び図7を参照すると、処理回路が、ブロック707において、信頼スコアが所定の閾値を満たしていないと判定した場合、方法700はブロック711に進み得る。

【0162】

一例として、ブロック705で計算された信頼スコアが0.7であり、所定の閾値が0.9である場合、処理回路は、信頼スコアが閾値よりも低く、したがって、信頼スコアは所定の閾値を満たしていないと決定し得る。

10

【0163】

ブロック711では、処理回路(例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202)は、ユーザ入力要求を生成し得、ユーザ入力要求をクライアントデバイス(図1に関連して上述されたクライアントデバイス101A~101Nなど)に送信し得る。

【0164】

いくつかの実施例では、ユーザ入力要求は、第1のデータ分類子に関連付けられてもよい。例えば、ユーザ入力要求は、第1のデータ分類子が正しいかどうかを確認するための電子要求を含んでもよい。

20

【0165】

いくつかの実施例では、ユーザ入力要求は、エキスパートに関連付けられたクライアントデバイスに送信されてもよい。クライアントデバイスは、ブロック703で決定されたデータオブジェクト及びデータ分類子のサンプルを含み得るユーザ入力要求を表示用にレンダリングしてもよい。一例として、クライアントデバイスは、データテーブルと、データテーブルが消費者情報に関連付けられていると決定されたことを示すデータ分類子と、を表示してもよい。エキスパートは、分類が正しいか又は正しくないかを選択することができる。

【0166】

ブロック713では、処理回路(例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202)は、ユーザ入力を受信し得る。ユーザ入力は、ブロック711のユーザ入力要求への応答であってもよい。

30

【0167】

いくつかの実施例では、ユーザ入力は、ブロック703で決定されたデータ分類子が正しいか否かに関するユーザからの指示を含んでもよい。一例として、ユーザ入力は、データ分類子の確認を含んでもよく、これは例えば、エキスパートが、データ分類子がデータオブジェクトの適切な分類を表すことを確認することを示し得る。別の例として、ユーザ入力は、データ分類子の修正を含んでもよく、これは例えば、エキスパートが、ブロック703で決定されたデータ分類子が正しくないか又は正確でないことを示し得る。

40

【0168】

ブロック715では、処理回路(例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202)は、ブロック713で受信されたユーザ入力が、ブロック703で決定された第1のデータ分類子が正しいことを示しているか否かを決定し得る。

【0169】

ブロック715において、処理回路が、ユーザ入力はデータ分類子が正しいことを確認するものと決定した場合、方法700はブロック709に進み得る。例えば、ユーザ入

50

力が第1のデータ分類子の確認を含む場合、処理回路は、ブロック709における第1のデータ分類子の確認に応答してマッピング仕様を生成してもよい。

【0170】

ブロック715において、処理回路が、ユーザ入力はデータ分類子が適切であることを確認しないものと決定した場合、方法700はブロック719に進み得る。ブロック719では、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202）は、ユーザ入力に基づいて第1のデータ分類子を修正し得る。

【0171】

いくつかの実施例では、ユーザ入力は、第1のデータ分類子の修正を含んでもよい。一例として、ブロック703で決定された第1のデータ分類子が、データオブジェクトは「ファーストネーム」に関連付けられていると決定されたことを示している場合、ユーザ入力は、データオブジェクトが「ラストネーム」に関連付けられていることを示してもよい。

10

【0172】

いくつかの実施例では、処理回路は、ユーザ入力に基づいて第1のデータ分類子を修正してもよい。上記の例から続けると、処理回路は、ユーザ入力に基づいて「ファーストネーム」を示すものから「ラストネーム」を示すものにデータ分類子を変更してもよい。

【0173】

ブロック719で第1のデータ分類子を修正することに続いて、処理回路はブロック709に進み得る。ブロック709において、処理回路は、修正された第1のデータ分類子に基づいてマッピング仕様を生成してもよい。

20

【0174】

図7を再び参照すると、ブロック717では、処理回路（例えば、図1に関連して説明されるデータ分析システム105の処理回路、及び/又は図2に関連して説明される装置200のプロセッサ202）は、ユーザ入力に基づいてフィードバックデータを生成し得る。

【0175】

いくつかの実施例では、フィードバックデータは、データ分類子を生成する際の精度を向上させるために、機械学習モデル、IA、及び/又はAIツールに提供されてもよい。例えば、処理回路は、（第1の複数のデータオブジェクトのものと比較して）同じ又は異なるデータベーススキーマに関連付けられた別の複数のデータオブジェクトを取得してもよく、データオブジェクト及びフィードバックデータに基づいてデータ分類子を生成してもよい。

30

【0176】

一例として、処理回路が教師あり学習モデルを利用して第1のデータ分類子を決定するとき、処理回路は、ユーザ入力に基づいて入出力ペア（すなわち、教師あり学習モデル用のトレーニングデータ）の形態でフィードバックデータを生成してもよい。例えば、ユーザ入力が、データオブジェクトがラストネームに関連付けられていることを示しているとき、処理回路は、入力（データオブジェクト）及び出力（ラストネーム分類）のペアを生成してもよい。処理回路は、教師あり学習モデルをトレーニングするための入力及び出力のペアを提供してもよく、それにより、教師あり学習モデルは、入出力ペア内のものと同様の他のデータオブジェクトについてデータ分類子を決定する際の精度を向上させ得る。

40

【0177】

方法700は、ブロック721で終了する。

【0178】

本開示は、開示される特定の実施形態に限定されるものではないこと、並びに修正及び他の実施形態は、添付の特許請求の範囲内に含まれることが意図されることを理解されたい。特定の用語が本明細書で用いられているが、特に記載のない限り、これらは一般的か

50

つ記述的な意味でのみ使用され、限定の目的では使用されない。

本願の実施例に係る発明には、例えば、以下の各発明が含まれる：

[発明 1]

コンピュータ実施方法であって、

第 1 のデータベーススキーマに対してデータマッピングプロセスを実行するための要求を受信することと、

前記要求に応答して、

データベースから、前記第 1 のデータベーススキーマに関連付けられた第 1 のデータモデルに関する第 1 の複数のデータオブジェクトを取得することと、

前記第 1 の複数のデータオブジェクトに少なくとも基づいて、前記第 1 のデータベーススキーマに対応する第 1 のデータ分類子を決定することと、

前記第 1 のデータ分類子に関連付けられた信頼スコアを計算することと、

前記信頼スコアが所定の閾値を満たすという決定に応答して、

前記第 1 のデータ分類子及び前記第 1 の複数のデータオブジェクトに関連付けられたメタデータに少なくとも部分的に基づいて、マッピング仕様に関連付けられたデータ動作に関するルールのセットを決定することと、

前記マッピング仕様に関連付けられた前記ルールのセットに基づいて、前記第 1 のデータベーススキーマに関連付けられた前記第 1 の複数のデータオブジェクトを、第 2 のデータベーススキーマに関連付けられた第 2 のデータモデルに関する第 2 の複数のデータオブジェクトに変換することと、

前記第 2 の複数のデータオブジェクトに対して 1 つ以上のデータの本質的な特徴の理解を提供することと、を含む、コンピュータ実施方法。

[発明 2]

前記第 1 の複数のデータオブジェクトが第 1 のデータテーブルを含み、前記第 1 のデータテーブルは、少なくとも 1 つのデータフィールドを含む、発明 1 に記載のコンピュータ実施方法。

[発明 3]

前記第 1 のデータテーブルが、名前メタデータ、列メタデータ、又は行メタデータのうちの少なくとも 1 つを含む、発明 2 に記載のコンピュータ実施方法。

[発明 4]

前記第 1 のデータベーススキーマに対応する前記第 1 のデータ分類子を決定することが、前記第 1 の複数のデータオブジェクトに関連付けられた前記名前メタデータ、前記列メタデータ、又は前記行メタデータのうちの少なくとも 1 つを取得することを更に含む、発明 3 に記載のコンピュータ実施方法。

[発明 5]

前記第 1 のデータベーススキーマに対応する前記第 1 のデータ分類子を決定することが、前記名前メタデータ、前記列メタデータ、又は前記行メタデータのうちの少なくとも 1 つに更に基づいて、前記第 1 のデータ分類子を決定することを更に含む、発明 4 に記載のコンピュータ実施方法。

[発明 6]

前記第 1 の複数のデータオブジェクトが第 2 のデータテーブルを含み、前記コンピュータ実施方法は、

前記第 1 のデータテーブル及び前記第 2 のデータテーブルに関連付けられた相関メタデータを決定することを更に含む、発明 2 に記載のコンピュータ実施方法。

[発明 7]

前記第 1 の複数のデータオブジェクトが第 2 のデータテーブルを含み、前記コンピュータ実施方法は、

前記相関メタデータに更に基づいて、前記第 1 のデータ分類子を決定することを更に含む、発明 6 に記載のコンピュータ実施方法。

[発明 8]

10

20

30

40

50

前記第 1 のデータテーブルに関連付けられたドメインメタデータを決定することを更に含む、発明 2 に記載のコンピュータ実施方法。

[発明 9]

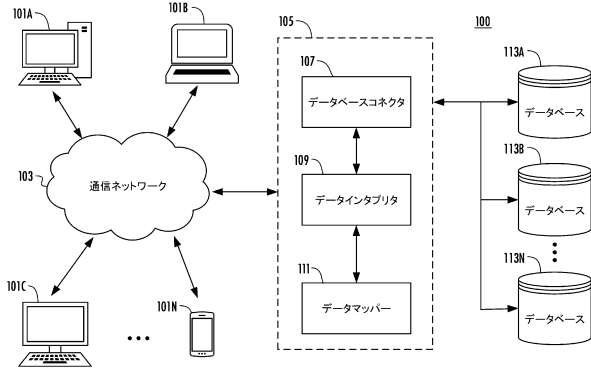
前記ドメインメタデータに更に基づいて、前記第 1 のデータ分類子を決定することを更に含む、発明 8 に記載のコンピュータ実施方法。

[発明 10]

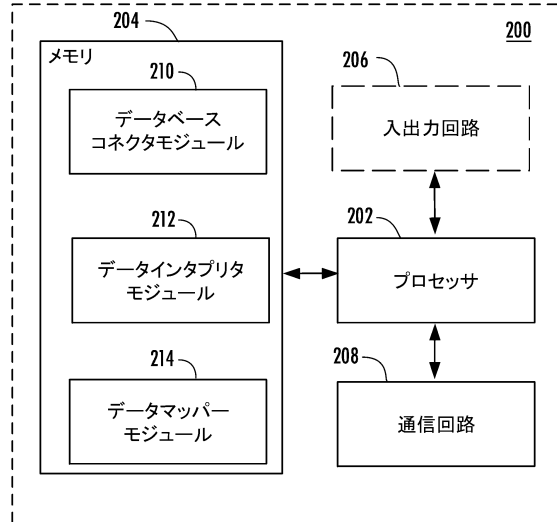
前記信頼を計算することが、前記第 1 の複数のデータオブジェクトと、教師あり学習モデルに関連付けられたデータとの間の類似性の度合いに基づいて前記信頼スコアを計算することを含む、発明 1 に記載のコンピュータ実施方法。

【図面】

【図 1】



【図 2】



10

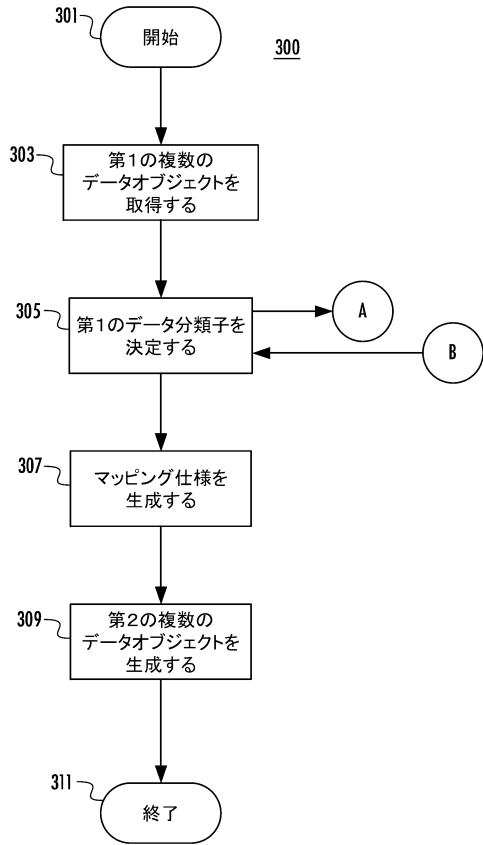
20

30

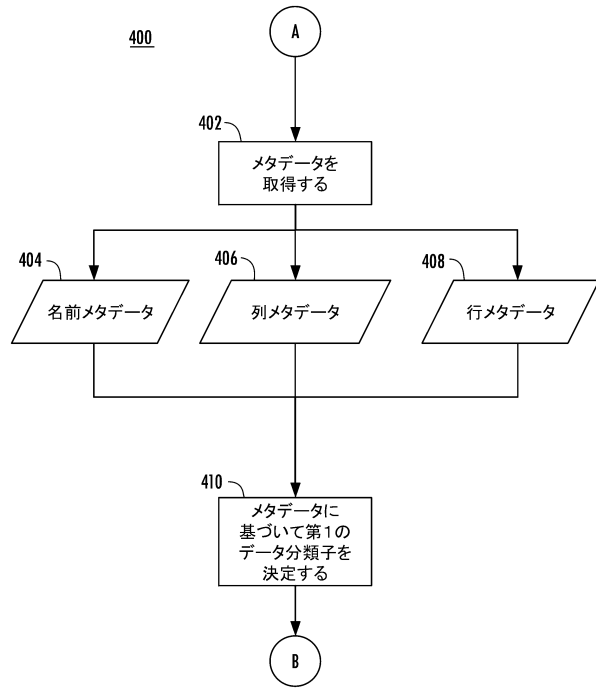
40

50

【図 3】



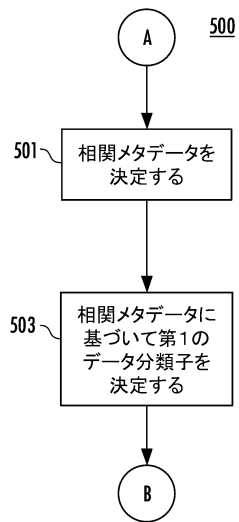
【図 4】



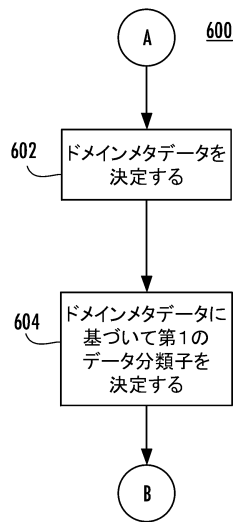
10

20

【図 5】



【図 6】

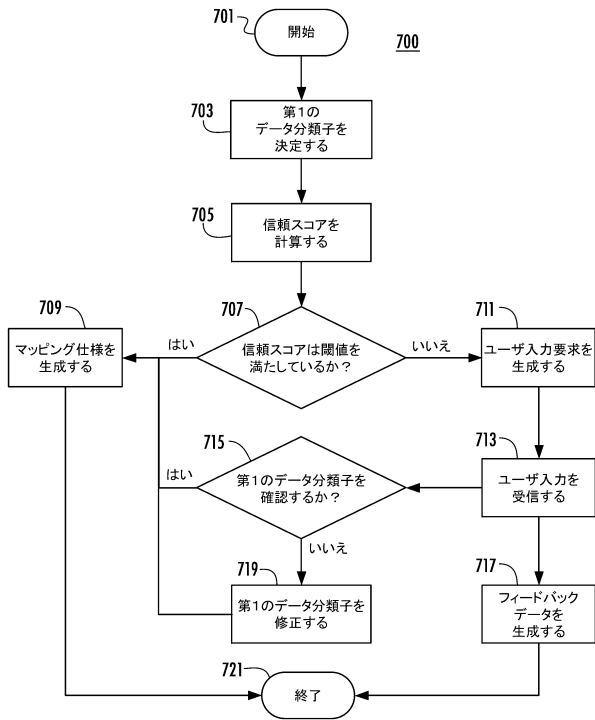


30

40

50

【 図 7 】



10

20

30

40

50

フロントページの続き

- 弁理士 中村 彰吾
- (72)発明者 シャウン・ザベル
アメリカ合衆国 サウスカロライナ州 29707-7539 フォートミル, オールドベイルスロード 9680
- (72)発明者 シヴァ・ラボル
インド 500019 トリシュール県 ハイデラバード, ナナックラムギューダ ヴィレッジ (ピー), ロット 115
- (72)発明者 ニーラジャ・サンジェーヴァルコット
インド 500019 トリシュール県 ハイデラバード, ナナックラムギューダ ヴィレッジ (ピー), ロット 115
- (72)発明者 スウェサ・シダリンガッパ
インド 560103 カルナータカ州 バンガロール, ケーアール バーサーホプリ, デバラビサナハリ ヴィレッジ
- (72)発明者 マハブベラバンニ・シャイク
インド 500019 トリシュール県 ハイデラバード, ナナックラムギューダ ヴィレッジ (ピー), ロット 115
- (72)発明者 シャルロtte・ルーミス
アメリカ合衆国 ノースカロライナ州 28202 シャルロtte, 300サウス テイロンストリート, スイート600 ハネウエル・インターナショナル・インコーポレーテッド インテレクチュアルプロパティサービスグループ
- (72)発明者 ジェッセ・ゲイトリー
アメリカ合衆国 サウスカロライナ州 29707-7539 フォートミル, オールドベイルスロード 9680
- (72)発明者 ロバート・гент
アメリカ合衆国 カリフォルニア州 92154, サンディエゴ, ウォーターヴィルロード 7828
- (72)発明者 ヴィナサ・バビプラカシュ
イギリス国 RG12 1EB ブラックネル パークス, ダウンシェアウェイ, スキンブドヒルレーン
- (72)発明者 ヴォジュテック・ソジュカ
チェコ共和国 62700 ブルノBM, トランカ 1387/100
- 審査官 齊藤 貴孝
- (56)参考文献 米国特許出願公開第2019/0258942 (US, A1)
米国特許出願公開第2018/0232528 (US, A1)
特表2019-528491 (JP, A)
米国特許出願公開第2019/0220449 (US, A1)
- (58)調査した分野 (Int.Cl., DB名)
G06F 16/00 - 16/958