

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4906292号
(P4906292)

(45) 発行日 平成24年3月28日 (2012.3.28)

(24) 登録日 平成24年1月20日 (2012.1.20)

(51) Int. Cl. F 1
G 0 6 F 12/00 (2006.01) G O 6 F 12/00 5 1 3 A
G 0 6 F 17/21 (2006.01) G O 6 F 17/21 5 9 0 E

請求項の数 6 (全 12 頁)

<p>(21) 出願番号 特願2005-253870 (P2005-253870) (22) 出願日 平成17年9月1日 (2005.9.1) (65) 公開番号 特開2006-107466 (P2006-107466A) (43) 公開日 平成18年4月20日 (2006.4.20) 審査請求日 平成20年8月26日 (2008.8.26) (31) 優先権主張番号 10/955,433 (32) 優先日 平成16年9月30日 (2004.9.30) (33) 優先権主張国 米国 (US)</p>	<p>(73) 特許権者 500046438 マイクロソフト コーポレーション アメリカ合衆国 ワシントン州 9805 2-6399 レッドモンド ワン マイ クロソフト ウェイ (74) 代理人 100077481 弁理士 谷 義一 (74) 代理人 100088915 弁理士 阿部 和夫 (72) 発明者 スミ エヌ, シン アメリカ合衆国 98052 ワシントン 州 レッドモンド ワン マイクロソフト ウェイ マイクロソフト コーポレーシ ョン内</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

最終頁に続く

(54) 【発明の名称】 電子文書で使用するために複数のデータソースのデータをマージする方法、システム、およびコンピュータ読み取り可能媒体

(57) 【特許請求の範囲】

【請求項1】

電子文書で使用するために複数のデータソースのデータをマージする方法であって、コンピュータに記憶されたコンピュータ実行可能命令を前記コンピュータが実行することによって実施される該方法は、

- 第1のデータソースのフィールド名を受け取ること、
- 少なくとも1つの第2のデータソースのフィールド名を受け取り、前記第1のデータソースの前記フィールド名にマッピングし、関連付けること、
- 前記マッピングされたフィールド名に基づいて受取りリストスキーマを構築すること、
- 前記第1のデータソースおよび前記第2のデータソースからデータを受け取るために、前記受取りリストスキーマに従って受取りリストテーブルを作成すること、
- 前記第1のデータソースおよび前記第2のデータソースからデータを取り出して、前記受取りリストテーブルを埋めること、
- 前記受取りリストテーブル内の各データについてハッシュ値を算出すること、
- ユーザ入力に応じて、前記受取りリストテーブル内のデータを修正すること、
- 前記受取りリストテーブル内の修正されたデータに対応する前記ハッシュ値を参照して、前記第1のデータソースおよび前記第2のデータソース中の対応するデータの位置を突き止めて更新すること、
- ユーザ入力に応じて、前記受取りリストテーブル中の前記第1のデータソースと前記第2のデータソースとの間の重複を解決することであって、

10

20

前記受取りリストテーブルから前記ハッシュ値を含むハッシュテーブルを作成すること

前記ハッシュテーブル内の前記ハッシュ値を比較すること、および、
前記ハッシュ値が合致した場合に、前記合致したハッシュ値を有するデータをユーザ
インターフェースウィンドウに表示することを含むこと、

前記受取りリストテーブル内のデータをファイルに保存すること
 を備えることを特徴とする方法。

【請求項 2】

前記受取りリストスキーマを構築することは、

前記第 1 のデータソースの前記フィールド名のいずれにも対応しない前記第 2 のデータ
ソースの前記フィールド名を、前記受取りリストスキーマに追加すること
 を備えることを特徴とする請求項 1 に記載の方法。

【請求項 3】

電子文書で使用するために複数のデータソースのデータをマージするシステムであって、
 前記複数のデータソースのデータを使用するアプリケーションプログラムを実行するた
 めに動作可能なクライアントコンピュータを備え、該クライアントコンピュータが、

第 1 のデータソースのフィールド名を受け取り、

少なくとも 1 つの第 2 のデータソースのフィールド名を受け取り、前記第 1 のデータソ
ースの前記フィールド名にマッピングし、関連付け、

前記マッピングされたフィールド名に基づいて受取りリストスキーマを構築し、

前記第 1 のデータソースおよび前記第 2 のデータソースからデータを受け取るために、
前記受取りリストスキーマに従って受取りリストテーブルを作成し、

前記第 1 のデータソースおよび前記第 2 のデータソースからデータを取り出して、前記
受取りリストテーブルを埋め、

前記受取りリストテーブル内の各データについてハッシュ値を算出し、

ユーザ入力に応じて、前記受取りリストテーブル内のデータを修正し、

前記受取りリストテーブル内の修正されたデータに対応する前記ハッシュ値を参照して、
前記第 1 のデータソースおよび前記第 2 のデータソース中の対応するデータの位置を突き
止めて更新し、

ユーザ入力に応じて、前記受取りリストテーブル中の前記第 1 のデータソースと前記第 2
のデータソースとの間の重複を解決することであって、

前記受取りリストテーブルから前記ハッシュ値を含むハッシュテーブルを作成し、

前記ハッシュテーブル内の前記ハッシュ値を比較し、

前記ハッシュ値が合致した場合に、前記合致したハッシュ値を有するデータをユーザ
インターフェースウィンドウに表示し、

前記受取りリストテーブル内のデータをファイルに保存する
 ために動作可能であることを特徴とするシステム。

【請求項 4】

前記受取りリストスキーマを構築することは、

前記第 1 のデータソースの前記フィールド名のいずれにも対応しない前記第 2 のデータ
ソースの前記フィールド名を、前記受取りリストスキーマに追加すること
 を備えることを特徴とする請求項 3 に記載のシステム。

【請求項 5】

コンピュータ実行可能命令が格納されたコンピュータ読み取り可能媒体であって、前記
 コンピュータ実行可能命令は、コンピュータによって実行されると、電子文書で使用する
ために複数のデータソースのデータをマージする方法を前記コンピュータに実行させ、前
記方法は、

第 1 のデータソースのフィールド名を受け取ること、

少なくとも 1 つの第 2 のデータソースのフィールド名を受け取り、前記第 1 のデータソ
ースの前記フィールド名にマッピングし、関連付けること、

10

20

30

40

50

前記マッピングされたフィールド名に基づいて受取リストスキーマを構築すること、
 前記第1のデータソースおよび前記第2のデータソースからデータを受け取るために、
 前記受取リストスキーマに従って受取リストテーブルを作成すること、
 前記第1のデータソースおよび前記第2のデータソースからデータを取り出して、前記
 受取リストテーブルを埋めること、

前記受取リストテーブル内の各データについてハッシュ値を算出すること、
 ユーザ入力に応じて、前記受取リストテーブル内のデータを修正すること、
 前記受取リストテーブル内の修正されたデータに対応する前記ハッシュ値を参照して、
 前記第1のデータソースおよび前記第2のデータソース中の対応するデータの位置を突き
 止めて更新すること、

ユーザ入力に応じて、前記受取リストテーブル中の前記第1のデータソースと前記第2
 のデータソースとの間の重複を解決することであって、

前記受取リストテーブルから前記ハッシュ値を含むハッシュテーブルを作成すること

前記ハッシュテーブル内の前記ハッシュ値を比較すること、および、
 前記ハッシュ値が合致した場合に、前記合致したハッシュ値を有するデータをユーザ
 インターフェースウィンドウに表示することを含むこと、

前記受取リストテーブル内のデータをファイルに保存すること
 を含むことを特徴とするコンピュータ読み取り可能媒体。

【請求項6】

前記受取リストスキーマを構築することは、
 前記第1のデータソースの前記フィールド名のいずれにも対応しない前記第2のデータ
 ソースの前記フィールド名を、前記受取リストスキーマに追加すること
 を含むことを特徴とする請求項5に記載のコンピュータ読み取り可能媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ソフトウェアアプリケーションの分野、より詳細には、電子文書で使用する
 ための複数のデータソースのデータのマージに関する。

【背景技術】

【0002】

最新のデスクトップパブリッシングアプリケーションは、ユーザが容易に電子文書を作
 成することを可能にする。電子文書を作成するための多くのデスクトップパブリッシング
 アプリケーションで使用可能な一機能は、メールマージと呼ばれる。メールマージは、ユ
 ーザが単一のデータソースに接続し、そこからのデータをマージすることを可能にするこ
 とにより、様々なデータ文書を作成するプロセスを自動化する。メールマージ機能を使用
 する際、データは単一のデータソース（メーリングリストなど）から引き出され、文書の
 内のマークされた位置に挿入される。

【発明の開示】

【発明が解決しようとする課題】

【0003】

最新のデスクトップパブリッシングアプリケーションによって提供されるメールマージ
 機能に関連する一つの欠点は、ユーザがしばしば異なるフォーマットで、コンピュータシ
 ステム内の複数の場所にデータ（メーリングリストなど）を格納することである。例えば
 、ユーザは、ある連絡先のリストを、連絡先マネージャプログラムで読み取り可能な連絡先
 ファイルとして格納することがあり、別の連絡先のリストを、スプレッドシートプログラ
 ムで読み取り可能なスプレッドシートファイルとして格納することがある。したがって、ユ
 ーザはしばしば、デスクトップパブリッシングアプリケーションによって提供されるメー
 ルマージ機能を利用するために、複数のデータソースからの異なるフォーマットのデー
 タを単一のデータソースとして集める必要がある。

10

20

30

40

50

【0004】

こうした考慮すべき点やその他の点に関して、本発明の様々な実施形態が作成された。

【課題を解決するための手段】

【0005】

本発明によれば、上記およびその他の問題は、電子文書で使用するために複数のデータソースのデータをマージする方法、システム、およびコンピュータ読み取り可能媒体によって解決される。本発明の一態様によれば、この方法は、初期データソースからフィールド名およびフィールドデータを受け取ること、その後で追加されたデータソースのフィールド名を初期データソースにマッピングすること、ならびに初期データソースのフィールド名、およびその後で追加されたデータソースのマッピング後フィールド名に基づいて受取りリストスキーマを構築することを含む。受取りリストスキーマは、受取りリスト内のフィールド名と、初期データソースおよびその後で追加されたデータソース内のフィールド名との関係を定義する。受取りリストの構築は、初期データソースのフィールド名に対して、その後で追加されたデータソースのマッピング後フィールド名を比較することを含み、後で追加されたデータソースのマッピング後フィールド名のいずれも初期データソースのフィールド名に対応しない場合、この方法は、マッピング後フィールド名を追加することを含む。

10

【0006】

この方法はさらに、受取りリストスキーマに従って受取りリストを作成すること、および受取りリストを再作成することを可能にするファイルを保存することを含む。受取りリストスキーマによる受取りリストの作成は、マスタデータソースを作成するためにその後で追加されたデータソースからのフィールドデータの行を一時受取りリストに追加すること、およびマスタデータソース中のフィールドデータの各行についてハッシュ値を計算することを含む。保存した更新後受取りリストファイルは、初期データソースに対する参照、その後で追加されたデータソースに対する参照、およびマスタデータソース中のフィールドデータの各行について計算したハッシュ値を含む。この方法はさらに、初期データソース中のフィールドデータ、およびその後で追加されたデータソース中のフィールドデータを修正するために受取りリストを修正すること、および受取りリスト内の初期データソースとその後で追加されたデータソースとの間の重複を解決することを含む。

20

【0007】

本発明は、コンピュータプロセス、コンピューティングシステム、またはコンピュータプログラム製品やコンピュータ読み取り可能媒体などの製品として実装することができる。コンピュータプログラム製品は、コンピュータシステムで読み取り可能であり、コンピュータプロセスを実行する命令のコンピュータプログラムを符号化するコンピュータ記憶媒体でよい。コンピュータプログラム製品はまた、コンピューティングシステムで読み取り可能であり、コンピュータプロセスを実行する命令のコンピュータプログラムを符号化する搬送波上で伝播される信号でもよい。

30

【0008】

本発明を特徴付ける上記およびその他の様々な機能ならびに利点は、以下の詳細な説明を読み、関連する図面を検討することによって明らかとなるであろう。

40

【発明を実施するための最良の形態】

【0009】

次に図面を参照して、本発明の様々な態様を説明する。各図面では、同様の番号が同様の要素を表す。具体的には、図1およびそれに対応する議論は、本発明の実施形態を実施することのできる適切なコンピューティング環境の簡潔で一般的な説明を与えるためのものである。パーソナルコンピュータのオペレーティングシステム上で動作するプログラムモジュールと共に実行されるプログラムモジュールの一般的文脈で本発明を説明するが、他のタイプのコンピュータシステムおよびプログラムモジュールと組み合わせて本発明を実施することもできることを当業者は理解されよう。

【0010】

50

一般に、プログラムモジュールは、特定のタスクを実行し、または特定の抽象データ型を実装する、ルーチン、プログラム、コンポーネント、データ構造、およびその他のタイプの構造を含む。さらに本発明は、ハンドヘルド装置、マルチプロセッサシステム、マイクロプロセッサベースの、またはプログラム可能な消費者向け電子機器、ミニコンピュータ、メインフレームコンピュータなどを含むその他のコンピュータシステム構成で実施できることを当業者は理解されよう。本発明は、通信ネットワークを介してリンクされたりリモート処理装置によってタスクが実施される分散コンピューティング環境でも実施することができる。分散コンピューティング環境では、プログラムモジュールがローカルメモリ記憶装置とリモートメモリ記憶装置のどちらにも位置することができる。

【0011】

10

次に図1を参照して、本発明の様々な実施形態で使用するコンピュータ2用の例示的コンピュータアーキテクチャを説明する。図1に示すコンピュータアーキテクチャは、中央演算処理装置5(「CPU」)と、ランダムアクセスメモリ9(「RAM」)および読取り専用メモリ(「ROM」)11を含むシステムメモリ7と、メモリをCPU5に結合するシステムバス12とを含む従来型デスクトップコンピュータまたはラップトップコンピュータを示す。起動時などにコンピュータ内の要素間で情報を転送する助けになる基本ルーチンを含む基本入出力システムが、ROM11に格納される。コンピュータ2はさらに、以下でより詳細に説明するオペレーティングシステム16、アプリケーションプログラム、およびその他のプログラムモジュールを格納する大容量記憶装置14を含む。

【0012】

20

大容量記憶装置14は、バス12に接続された大容量記憶コントローラ(図示せず)を介してCPU5に接続される。大容量記憶装置14およびその関連するコンピュータ読み取り可能媒体は、コンピュータ2における不揮発性記憶領域を提供する。本明細書に含まれるコンピュータ読み取り可能媒体の説明は、ハードディスクやCD-ROMドライブなどの大容量記憶装置を指すが、コンピュータ読み取り可能媒体はコンピュータ2によってアクセスすることのできる入手可能などんな媒体でもよいことを理解されたい。

【0013】

例えば、それだけに限らないが、コンピュータ読み取り可能媒体は、コンピュータ記憶媒体および通信媒体を含む。コンピュータ記憶媒体は、コンピュータ読み取り可能命令、データ構造、プログラムモジュール、またはその他のデータなどの情報を記憶するために何らかの方法または技術で実装された、揮発性および不揮発性の、取外し可能および取外し不能な媒体を含む。コンピュータ記憶媒体には、それだけに限らないが、RAM、ROM、EPROM、EEPROM、フラッシュメモリ、またはその他の固体メモリ技術、CD-ROM、デジタル多用途ディスク(「DVD」)、またはその他の光記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置、またはその他の磁気記憶装置、あるいは、所望の情報を格納するために使用することができ、コンピュータ2によってアクセスすることができる他のどんな媒体も含まれる。

30

【0014】

本発明の様々な実施形態によれば、コンピュータ2は、インターネットなどのネットワーク18を介したりリモートコンピュータに対する論理接続を使用するネットワーク環境で動作することができる。コンピュータ2は、バス12に接続されたネットワークインターフェースユニット20を介してネットワーク18に接続する。ネットワークインターフェースユニット20は、他のタイプのネットワークおよびリモートコンピュータシステムに接続するために利用することもできることを理解されたい。コンピュータ2はまた、キーボード、マウス、または電子スタイラス(図1には図示せず)を含むいくつかの他の装置から入力を受け取って処理する入出力コントローラ22も含む。同様に、入出力コントローラ22は、ディスプレイ画面、プリンタ、その他のタイプの出力装置に出力を提供する。

40

【0015】

上記で簡潔に述べたように、いくつかのプログラムモジュールおよびデータファイルは

50

、コンピュータ2の大容量記憶装置14およびRAM9に記憶される。それらのプログラムモジュールおよびデータファイルは、ワシントン州レッドモンドのMICROSOFT CORPORATIONによるWINDOWS（登録商標）XPオペレーティングシステムなどの、ネットワークパーソナルコンピュータの動作を制御するのに適したオペレーティングシステム16を含む。大容量記憶装置14およびRAM9は、1つまたは複数のプログラムモジュールも格納する。具体的には、大容量記憶装置14およびRAM9は、デスクトップパブリッシングアプリケーション10を格納する。当業者には周知のように、デスクトップパブリッシングアプリケーション10は、電子文書を作成および編集する機能を提供するために動作する。本発明の一実施形態によれば、デスクトップパブリッシングアプリケーション10は、MICROSOFT CORPORATIONによるPUBLISHERというワードプロセッシングアプリケーションプログラムを含む。しかし、本発明の様々な態様を具体化するために、他の製造業者によるデスクトップパブリッシングアプリケーションも使用できることを理解されたい。本発明の様々な態様はデスクトップパブリッシングアプリケーションに限定されず、MICROSOFT CORPORATIONによるWORDプログラムならびにスプレッドシートプログラムおよびデータベースプログラムなどの、テキストを処理することができるその他のプログラムも使用できることをさらに理解されたい。

【0016】

受取リストの作成に関連して、デスクトップパブリッシングアプリケーション10は、ユーザがデータソース26（データソース1）、28（データソース2）、および30（データソースN）を電子文書の様々な位置にマージすることを可能にする機能を提供する。各データソース26、28、および30は、1つまたは複数のフィールドに分割されたデータのリストまたはテーブルであることが理解されよう。例えば、データソースは、名前、会社、および住所をデータフィールド「Last Name」、「First Name」、「Title」、「Company Name」、および「Address」内に含む連絡先情報を格納することができる。複数のデータソースをマージする際に、デスクトップパブリッシングアプリケーション10は、マージされたデータを格納する受取リストファイル24を作成することを理解されたい。受取リストファイル24を作成するためにデータソース26（Data Source 1）、28（Data Source 2）、および30（Data Source N）をマージするデスクトップパブリッシングアプリケーション10によって使用される例示的方法を、図2に関して以下でより詳細に説明する。

【0017】

ここで図2を参照して、複数のデータソースのデータをマージするためにデスクトップパブリッシングアプリケーション10によって実施されるプロセスを示す例示的ルーチン200を説明する。本明細書で説明する本発明の実施形態をデスクトップパブリッシングアプリケーション10の状況で提示するが、本発明は、ワードプロセッシング、スプレッドシート、データベースプログラムなどのテキスト処理をサポートする他のタイプのアプリケーションプログラムでも使用できることを理解されたい。

【0018】

本明細書で提示されるルーチンの説明を読む際に、本発明の様々な実施形態の論理演算が、（1）コンピューティングシステム上で動作する、一連のコンピュータで実施される動作またはプログラムモジュールとして、および/または（2）コンピューティングシステム内の相互接続された機械論理回路または回路モジュールとして実装されることを理解されたい。実装は、本発明を実装するコンピューティングシステムの性能要件に依存して選択される。したがって、図2に示し、本明細書で説明する本発明の実施形態を構成する論理演算は、オペレーション、構造装置、動作、またはモジュールと様々な形で呼ばれる。本明細書に記載の特許請求の範囲内で列挙される本発明の精神および範囲から逸脱することなく、こうしたオペレーション、構造装置、動作、およびモジュールはソフトウェア、ファームウェア、特殊目的デジタル論理、およびそれらの任意の組合せとして実装でき

10

20

30

40

50

ることを当業者は理解されよう。

【0019】

ここで図2を参照すると、ルーチン200はオペレーション210で始まり、デスクトップパブリッシングアプリケーション10がデータソース26(Data Source 1)、28(Data Source 2)、および30(Data Source N)を受け取り、各データソース中のフィールドを「マッピング」する。具体的には、各データソース中のフィールドをマッピングする際に、デスクトップパブリッシングアプリケーション10は、最初のデータソースを初期データソースとして取り出し、取り出した後続の各データソースのフィールドを初期データソースのフィールドと関連付け、初期データソース中に出現しない後続のデータソース中のフィールドを追加する。例えば、データソース26(Data Source 1)は、データフィールド「First Name」および「Last Name」を含み、一方、データソース28(Data Source 2)は、「First Name」、「Last Name」、および「Title」を含む。データフィールドをマッピングする際に、ワードプロセッシングアプリケーションプログラムは、データソース26を初期データソースとして指定し、各データソースによって共有される共通フィールド(すなわち、「First name」および「Last Name」)をマッピングし、データソース28の固有のフィールド「Title」を追加する。本発明の様々な実施形態では、デスクトップパブリッシングアプリケーション10のユーザは、初期データソースにマッピングまたは追加すべき後続のデータソースのフィールドを指定する。さらに、マッピング済みフィールドをマッピング解除することもできることを理解されよう。

10

20

【0020】

ルーチン200は、オペレーション210からオペレーション220に進み、デスクトップパブリッシングアプリケーション10は、データソース26、28、および30のマッピング済みフィールドから、受取リスト(すなわちマスタデータソース)スキーマを構築する。本明細書では、「スキーマ」は、各入力データソースからマッピングされたフィールド名を定義する。例えば、データソース26およびデータソース28についての受取リストスキーマは、(データソース26および28の)フィールド「First Name」および「Last Name」、ならびに(データソース28のみの)フィールド「Title」を含む。

30

【0021】

ルーチン200は、オペレーション220からオペレーション230に進み、デスクトップパブリッシングアプリケーション10は、受取リストスキーマから一時受取リストを作成する。具体的には、デスクトップパブリッシングアプリケーション10は、各入力データソースのフィールドに関連するデータを受け取るために、受取リストスキーマからフィールドのテーブルを作成する。ルーチン200は、オペレーション230からオペレーション240に進み、デスクトップパブリッシングアプリケーション10は、受取リストを作成するために各入力データソースからデータの1つまたは複数の行を取り出す。具体的には、デスクトップパブリッシングアプリケーション10は、各データソースからデータを取り出し、受取リストスキーマに従って一時受取リストを埋める。

40

【0022】

ルーチン200は、オペレーション240からオペレーション250に進み、デスクトップパブリッシングアプリケーション10は、受取リスト中の各行についてハッシュを作成する。当業者には周知の通り、ハッシュは、テキスト文字列から生成される数値である。それは、データメソッドにアクセスするために使用される。ハッシュ(またはハッシュ値)は、他の何らかのテキストが同一のハッシュ値を生成する可能性が極めて低くなるような式で生成される。ハッシュ値を生成する様々な方法は当業者に周知であり、したがって、本明細書ではさらに詳細には論じない。入力データソース中のデータに対するリンクとして働くハッシュを使用して、受取リスト中のマージデータを区別することができることを理解されよう。

50

【 0 0 2 3 】

ルーチン 2 0 0 は、オペレーション 2 5 0 からオペレーション 2 6 0 に進み、デスクトップパブリッシングアプリケーション 1 0 は、受取リストをファイル（受取リストファイル 2 4 など）に保存する。受取リストを保存する際、受取リストを再作成することを可能にするファイルが作成されることを理解されよう。さらに受取リストファイル 2 4 は、初期データソースに対する参照、後続のまたは追加された各データソースに対する参照、および受取リスト中のフィールドデータの各行について計算されたハッシュ値を含む。ルーチン 2 0 0 は、オペレーション 2 6 0 からオペレーション 2 7 0 に進み、デスクトップパブリッシングアプリケーション 1 0 は、ユーザからの入力に回答して、受取リストを修正する。具体的には、ユーザは、データを変更または削除することによって受取リスト中のデータを更新することができる。受取リスト中の修正により、供給側データソース中のデータも修正されることを理解されよう。例えば、データソース 2 6 から取り出された受取リスト中のラストネームデータをユーザが修正する結果として、データソース 2 6 中の同一のラストネームデータが修正される。受取リスト中のデータを修正する際に、デスクトップパブリッシングアプリケーション 1 0 は、影響を受けるデータソース中の対応するデータの位置を突き止めて更新するために、受取リスト中のデータの修正された各行について計算されたハッシュ値を参照することを理解されよう。

10

【 0 0 2 4 】

ルーチン 2 0 0 は、オペレーション 2 7 0 からオペレーション 2 8 0 に進み、デスクトップパブリッシングアプリケーション 1 0 は、受取リスト中の重複データを解決する。具体的には、データソース 2 6、2 8、および 3 0 は、受取リスト中の重複エントリとして現れる可能性のある同一のデータを含むことがあることを理解されよう。例えば、ユーザは、2 つの異なるデータソース中にある同一人物についての名前および住所の情報を有する可能性がある。

20

【 0 0 2 5 】

一実施形態では、デスクトップパブリッシングアプリケーション 1 0 は、受取リストからハッシュ値のハッシュテーブルを作成し、2 つのハッシュ値が指すデータエントリを比較することによって、重複データエントリを突き止めるように構成されることを理解されよう。本発明の様々な例示的实施形態では、あらゆる行をあらゆる他の行と比較することによって重複を見つける。各行について計算されたハッシュ値はメモリに記憶される。あらゆる行中のあらゆる列について、1 つのハッシュ値が記憶されることを理解されたい。次いでハッシュ値が比較される。2 つのデータエントリが合致するとき、そのハッシュ値も合致する。2 つの行が所定の重複規則（例えば 2 つの行中のデータの 7 5 % が合致しなければならない）を満足するとき、それらはマスタデータソースからロードされ、実際のデータエントリがユーザによって比較され、それらが実際に互いに重複するかどうか判定される。その後ルーチン 2 0 0 は終了する。

30

【 0 0 2 6 】

次に図 3 を参照して、ユーザがデータソースを受取リストスキーマにマッピングすることを可能にする例示的ユーザインターフェースを説明する。図 3 に、データソースのフィールドをマッピングするいくつかのユーザインターフェースコンポーネントを含むユーザインターフェースウィンドウ 3 0 0 を示す。具体的には、ウィンドウ 4 0 はデータソースのフィールドを表示し、それをウィンドウ 4 0 からドラッグして、フィールド 4 2 またはフィールド 4 4 にドロップすることによって、受取リストスキーマにマッピングする。ウィンドウ 4 4 に追加されるフィールドは、受取リストスキーマ中の新しい列として追加される。マッピングは、ウィンドウ 4 2 または 4 4 のフィールドをウィンドウ 4 0 にドラッグすることによってやり直すことができる。データソースのフィールドのデフォルトマッピングを選択するためのデフォルトマップボタン 4 6 が提供される。デフォルトマッピングは、デスクトップパブリッシングアプリケーション 1 0 によって自動的に決定される。

40

【 0 0 2 7 】

次に図 4 を参照して、2 つのデータソースから作成された受取リストを表示する例示的

50

ユーザインターフェースを説明する。図4は、受取リスト、ならびにマッピングされたフィールドの行、および関連するデータを表示するユーザインターフェースウィンドウ400を示す。ウィンドウ52は、受取リストを作成するために使用されたデータソースを特定する。ファイル、データベース、または連絡先リストから受取リストにデータソースを追加するためのリンク54および56が提供される。新しいリストをタイプすることによってデータソースを作成するためのリンク58が提供される。

【0028】

次に図5を参照して、ユーザがデータソースを編集することを可能にする例示的ユーザインターフェースを説明する。図5は、データソーステーブル60およびフィールドデータの行を表示するユーザインターフェースウィンドウ500を示す。新規エントリをデータソーステーブル60に追加する新規エントリ追加ボタン62、および、データソーステーブル60中に表示されるエントリを削除するエントリ削除ボタン64が提供される。ユーザインターフェースウィンドウ500は、別の場所に格納された元のデータソースにリンクする受取リストのマッピングされたデータまたはコピーされたデータを含み、それによって変更が元のデータソースに反映されることが理解されよう。

10

【0029】

次に図6を参照して、図2に関連して上記で論じたデスクトップパブリッシングアプリケーション10によって、ユーザが、受取リスト中に見つかった重複を除去することを可能にする例示的ユーザインターフェースを説明する。図6は、重複データエントリを列挙するデータソーステーブル70を表示するユーザインターフェース600を示す。ユーザが受取リストから除去すべきエントリを選択することを可能にするチェックボックス72の列が提供される。

20

【0030】

上記に基づいて、本発明の様々な実施形態は、電子文書で使用するために複数のデータソースのデータをマージする方法、システム、およびコンピュータ読み取り可能媒体を含むことを理解されたい。上記の仕様、例、およびデータは、本発明の構成の製造および使用について十分な説明を与える。本発明の精神および範囲から逸脱することなく、本発明の多数の実施形態を作成できるので、本発明は、添付の特許請求の範囲に属する。

【図面の簡単な説明】

【0031】

30

【図1】本発明の様々な実施形態で使用され、提供されるコンピュータシステムを示すコンピュータシステムアーキテクチャ図である。

【図2】本発明の例示的实施形態による、複数のデータソースからのデータをマージするために図1のコンピュータシステム内のデスクトップパブリッシングアプリケーションで実施される例示的ルーチンである。

【図3】本発明の様々な実施形態による、ユーザが複数のデータソースからのデータをマージして、マージ後データを管理する機能を提供する本発明の一態様を示す画面図である。

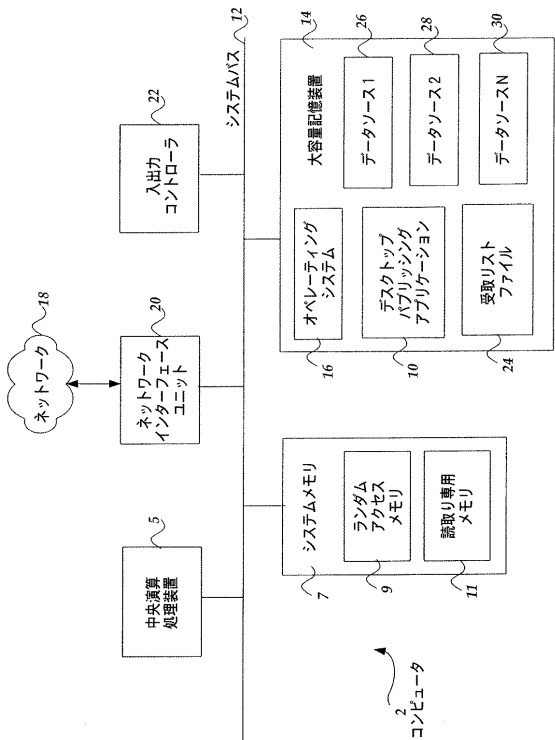
【図4】本発明の様々な実施形態による、ユーザが複数のデータソースからのデータをマージして、マージ後データを管理する機能を提供する本発明の一態様を示す画面図である。

40

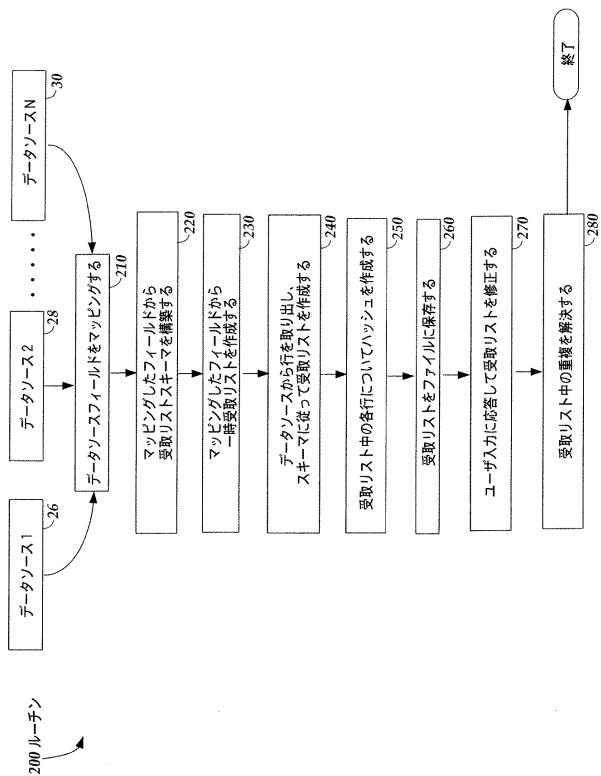
【図5】本発明の様々な実施形態による、ユーザが複数のデータソースからのデータをマージして、マージ後データを管理する機能を提供する本発明の一態様を示す画面図である。

【図6】本発明の様々な実施形態による、ユーザが複数のデータソースからのデータをマージして、マージ後データを管理する機能を提供する本発明の一態様を示す画面図である。

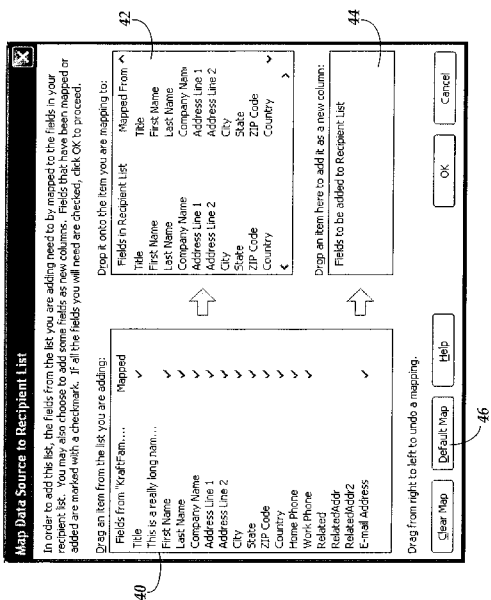
【図 1】



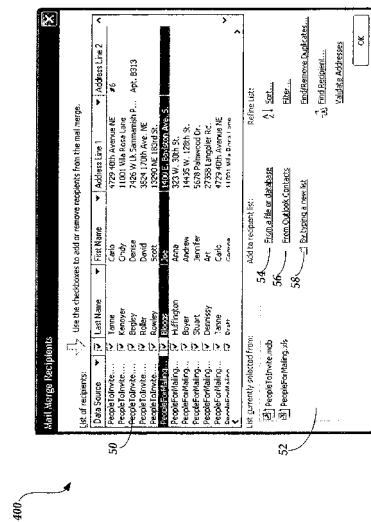
【図 2】



【図 3】



【図 4】



5

500

Edit Data Source

You can edit your data source by clicking directly in the grid below. Column headings display the column name in your data source and the column it is currently mapped to (in parentheses).

Data source being edited: PeopleForMailing.xls

First Name (First Name)	Last Name (Last Name)	Address (Address Line...)	City (City)	State (State)
Anna	Huffington	323 W. 30th St.	Los Angeles	CA
Andrew	Boyer	14435 W. 128th...	Burbank	CA
Jennifer	Stuart	5678 Palmwood ...	Pasadena	CA
Art	Dennisy	27353 Langoler ...	Bellevue	WA
Carlo	Tanne	4729 40th Aven...	Seattle	WA
George	Pratt	11001 Villa Rosa...	Mukilteo	WA
Theresa	Stallamp	7426 W Lik Sam...	Redmond	WA

60

62

64

Buttons: Add New Entry, Find..., Change Entry, Cancel

6

600

Find duplicates

The following message will be displayed. Use the checkboxes to specify the entries you would like to use in your mail merge.

72	73	74	75	76	77	78	79	80
Check Source	Last Name	First Name	Address Line 1	Address Line 2	City	State	Zip	Country
<input type="checkbox"/>	David	David	4729 40th Aven...		Seattle	WA	98148	USA
<input type="checkbox"/>	David	David	323 W. 30th St.		Los Angeles	CA	90007	USA
<input type="checkbox"/>	David	David	323 W. 30th St.		Los Angeles	CA	90007	USA

70

Buttons: OK, Cancel

フロントページの続き

- (72)発明者 ジュラジ ゴットワイス
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 ジョン イー・ディミック ジュニア
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 タラ エム・クラフト
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内

審査官 田川 泰宏

- (56)参考文献 特開2001-109758(JP,A)
橋 直彦, 電子メールの導入・相互接続とグループウェア, OPEN DESIGN, 日本, C
Q出版株式会社, 1997年 5月 1日, No. 9 第2版, p.30~p.47
内藤 一彦, B2Bに威力を発揮! 主要DBのXML対応動向と最新ソリューション, DB M
a g a z i n e, 日本, 株式会社翔泳社 SHOEISHA, 2000年 8月 1日, 第10巻第5号, p
.097~p.103

- (58)調査した分野(Int.Cl., DB名)
G06F 12/00
G06F 17/21