

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2009-276854
(P2009-276854A)

(43) 公開日 平成21年11月26日(2009.11.26)

(51) Int.Cl.
G06F 17/21 (2006.01)F 1
G06F 17/21 570M
G06F 17/21 538M
G06F 17/21 538A
G06F 17/21 501Tテーマコード(参考)
5B109

審査請求 未請求 請求項の数 8 O L (全 18 頁)

(21) 出願番号 特願2008-125151 (P2008-125151)
(22) 出願日 平成20年5月12日 (2008.5.12)(71) 出願人 000001007
キヤノン株式会社
東京都大田区下丸子3丁目30番2号
(74) 代理人 100090273
弁理士 國分 孝悦
(72) 発明者 齊野 陽子
東京都大田区下丸子3丁目30番2号 キ
ヤノン株式会社内
F ターム(参考) 5B109 ND01 ND03 TB13

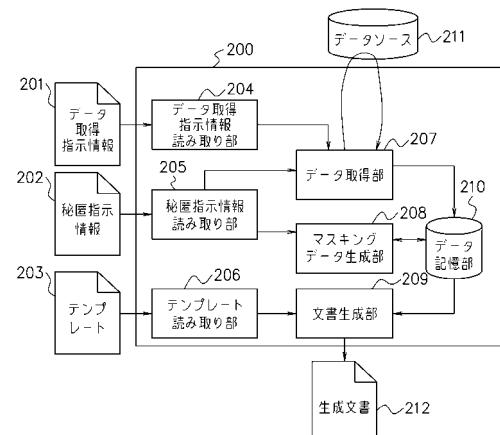
(54) 【発明の名称】情報処理装置、その制御方法及びプログラム

(57) 【要約】

【課題】 テンプレートデータを変更する場合であっても、データを秘匿する方法を改めて指定することなく、データが秘匿された文書等を生成することを可能とする。

【解決手段】 データ取得指示情報読み取り部 204 は、テンプレートデータに合成するデータを取得するためのデータ取得情報を入力する。また、秘匿指示情報読み取り部 205 は、データ取得情報に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿情報を入力する。データ取得部 207 は、記データ取得情報に基づいてデータを取得し、マスキングデータ生成部 208 は、秘匿情報に基づいて、秘匿するデータに対応する秘匿データを生成する。文書生成部 209 は、取得されたデータ及び生成された秘匿データをテンプレートデータに合成する。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

テンプレートデータに合成するデータを取得するためのデータ取得情報と、前記データ取得情報に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿情報とを入力する入力手段と、

前記データ取得情報に基づいてデータを取得する取得手段と、

前記秘匿情報に基づいて、前記秘匿するデータに対応する秘匿データを生成する生成手段と、

前記取得手段により取得されたデータ及び前記生成手段により生成された秘匿データを前記テンプレートデータに合成する合成手段とを有することを特徴とする情報処理装置。 10

【請求項 2】

前記データ取得情報は、前記テンプレートデータに合成するデータの属性情報を指定した情報であり、前記秘匿情報は、前記データ取得情報により指定される前記テンプレートデータに合成するデータの属性情報のうち、前記秘匿するデータに対応する属性情報を指定した情報であることを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】

前記取得手段は、前記データ取得情報により指定される属性情報のうち、前記秘匿情報により指定される属性情報に対応するデータを取得しないことを特徴とする請求項 2 に記載の情報処理装置。 20

【請求項 4】

前記取得手段は、前記秘匿情報により指定される属性情報に対応するデータのうちの一部のデータが秘匿対象のデータである場合、当該属性情報に対応するデータを取得し、前記生成手段は、前記取得手段により取得された当該属性情報に対応するデータを使用して前記秘匿データを生成することを特徴とする請求項 2 に記載の情報処理装置。 30

【請求項 5】

前記属性情報は、データの型を示す情報であることを特徴とする請求項 2 乃至 4 の何れか 1 項に記載の情報処理装置。

【請求項 6】

前記秘匿情報は、前記秘匿データを指定する情報を含み、前記生成手段は、前記秘匿情報に指定される前記秘匿データを生成することを特徴とする請求項 1 乃至 5 の何れか 1 項に記載の情報処理装置。 30

【請求項 7】

テンプレートデータに合成するデータを取得するためのデータ取得情報と、前記データ取得情報に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿情報とを入力する入力ステップと、

前記データ取得情報に基づいてデータを取得する取得ステップと、

前記秘匿情報に基づいて、前記秘匿するデータに対応する秘匿データを生成する生成ステップと、

前記取得ステップにより取得されたデータ及び前記生成ステップにより生成された秘匿データを前記テンプレートデータに合成する合成ステップとを含むことを特徴とする情報処理装置の制御方法。 40

【請求項 8】

テンプレートデータに合成するデータを取得するためのデータ取得情報と、前記データ取得情報に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿情報とを入力する入力ステップと、

前記データ取得情報に基づいてデータを取得する取得ステップと、

前記秘匿情報に基づいて、前記秘匿するデータに対応する秘匿データを生成する生成ステップと、

前記取得ステップにより取得されたデータ及び前記生成ステップにより生成された秘匿データを前記テンプレートデータに合成する合成ステップとをコンピュータに実行させる 50

ためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、テンプレートデータにデータを合成して電子文書等を生成するための技術に関するものである。

【背景技術】

【0002】

従来、帳票の生成処理方法として、帳票のデザインを表すテンプレートに、データが格納されたファイルやデータベース等のデータソースから取得したデータを合成し、帳票を作成することが行われている。

【0003】

一方、伝票、帳簿、契約書等の取引の詳細情報や特定の個人を表す情報等を含むドキュメントを印刷や表示出力する際に、それら情報の一部又は全てを別の文字で表したり、塗りつぶしたり、非表示にする等の伏せ字処理（マスキング処理）が行われている。

【0004】

マスキング処理に関する技術は、例えば特許文献1、2に開示されている。特許文献1は、帳票のレイアウトデザインを示すテンプレートであるフォームファイル上に定義されたフィールドに対し、データファイルからデータを供給し、重ね合わせることで帳票を作成する技術である。各フィールドに重ね合わせるデータのマスキング方法を指定し、指定されたマスキング方法に応じてそのフィールドに入るデータの文字列を出力制御する。例えば、フィールドに「_」が指定されていればその位置に対応するデータの文字を読み飛ばし、「K」が指定されていればその位置の文字を出力する。「あいうえおかきくけこ」がデータとして入るフォームファイルのフィールドに「KKKKKK」が指定してあれば、先頭から6文字「あいうえおか」を出力する。また、フィールドに「KKK_____KKK」と指定してあれば、「_」の位置に該当する「えおかき」を読み飛ばして「あいうくけこ」を出力する。

【0005】

また、特許文献2では、文書論理構造をベースにしてXMLのタグ付けをしたXML文書を作成して、そのXML文書に基づいてマスキングする。具体的にはマスキング対象の文書に対して文書番号、日付、見出し等の文書項目をXMLのタグで囲む。それに加え、ユーザが任意に指定した本文中の特定の範囲もその情報に見合うXMLのタグで囲む。そして、それぞれのタグに対して非公開にするレベルを設定しておく。文書出力する際に指定された非公開レベルに該当するXMLタグを検出し、検出したタグに囲まれた箇所をマスキングした文書を作成し出力する。

【0006】

【特許文献1】特開2002-170072号公報

【特許文献2】特開2004-94542号公報

【発明の開示】

【発明が解決しようとする課題】

【0007】

しかしながら、特許文献1に記載された技術によると、各フィールドに対してマスク処理を指定する必要があるため、そのデータが帳票中の複数の箇所に出現する場合はその箇所ごとにマスク方法を指定する手間がかかった。また、同じデータファイルからデータを供給する場合でもデザインテンプレートになった場合は、新しいデザインテンプレートにマスク方法を改めて指定する手間がかかった。

【0008】

また、特許文献2に記載された技術では、帳票のデザインを表すデザインテンプレートとそこにデータを埋め込む方法で帳票を作成するための技術ではないが、文書論理構造にマスキングする箇所を指定するものである。そのため、文書中に出現するマスキングが

10

20

30

40

50

必要箇所を全て特定のXMLのタグをつける必要がある。従って、同じ情報であってもその情報が文書論理構造中に複数出現している場合はその全てにXMLのタグをつけるという手間がある。

【0009】

そこで、本発明の目的は、テンプレートデータを変更する場合であっても、データを秘匿する方法を改めて指定することなく、データが秘匿された文書等を生成することを可能とすることにある。

【課題を解決するための手段】

【0010】

本発明の情報処理装置は、テンプレートデータに合成するデータを取得するためのデータ取得情報と、前記データ取得情報に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿情報とを入力する入力手段と、前記データ取得情報に基づいてデータを取得する取得手段と、前記秘匿情報に基づいて、前記秘匿するデータに対応する秘匿データを生成する生成手段と、前記取得手段により取得されたデータ及び前記生成手段により生成された秘匿データを前記テンプレートデータに合成する合成手段とを有することを特徴とする。

本発明の情報処理装置の制御方法は、テンプレートデータに合成するデータを取得するためのデータ取得情報と、前記データ取得情報に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿情報とを入力する入力ステップと、前記データ取得情報に基づいてデータを取得する取得ステップと、前記秘匿情報に基づいて、前記秘匿するデータに対応する秘匿データを生成する生成ステップと、前記取得ステップにより取得されたデータ及び前記生成ステップにより生成された秘匿データを前記テンプレートデータに合成する合成ステップとを含むことを特徴とする。

本発明のプログラムは、テンプレートデータに合成するデータを取得するためのデータ取得情報と、前記データ取得情報に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿情報とを入力する入力ステップと、前記データ取得情報に基づいてデータを取得する取得ステップと、前記秘匿情報に基づいて、前記秘匿するデータに対応する秘匿データを生成する生成ステップと、前記取得ステップにより取得されたデータ及び前記生成ステップにより生成された秘匿データを前記テンプレートデータに合成する合成ステップとをコンピュータに実行させることを特徴とする。

【発明の効果】

【0011】

本発明においては、データ取得情報に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿情報に基づいて上記秘匿するデータに対応する秘匿データを生成し、取得されたデータと秘匿データとをテンプレートデータに合成するようにしている。従って、本発明によれば、テンプレートデータを変更する場合であっても、データを秘匿する方法を改めて指定することなく、データが秘匿された文書等を生成することが可能となる。

【発明を実施するための最良の形態】

【0012】

以下、本発明を適用した好適な実施形態を、添付図面を参照しながら詳細に説明する。

【0013】

図1は、本発明の実施形態に係る情報処理装置100の構成を示すブロック図である。図1に示すように、情報処理装置100は、主な構成要素として、情報処理装置100全体の制御を司るコントローラ部101、操作部102、表示部103、ハードディスク104及びネットワークアダプタ108を備えている。

【0014】

コントローラ部101は、CPU105、ROM106、RAM107等を備える。CPU105が実行する各種プログラムや、プログラムが必要とするデータ等は、予めROM106に記憶されている。また、一部のプログラムやデータは、ハードディスク104

10

20

30

40

50

に予め格納しておき、RAM107に読み出して実行することも可能である。本実施形態では、ハードディスク104に文書生成を行うプログラム、生成した文書、文書生成に使用するテンプレート、データ取得指示情報、秘匿指示情報等を記憶する。RAM107にはデータソースから取得したデータ、プログラム実行中に生成する中間データを一時的に記憶する。

【0015】

操作部102はユーザからの入力操作を受け付ける。キーボード、マウスがこれに該当する。表示部103は処理を実行するアプリケーションのGUI(グラフィカルユーザインターフェース)を表示する。ディスプレイにはアプリケーションのパネル、ボタン、テキストフィールド等のGUIが表示され、キーボード、マウス等の操作部102から入力を受けると、それぞれのGUI部品に対応付けられたプログラムが実行される。

10

【0016】

ネットワークアダプタ108は、ネットワーク109を介してネットワーク上のデータベース110やファイルサーバ111等と通信するためのもので例えれば、LANカードである。

【0017】

図2は、情報処理装置100によって実現する文書生成システム200の構成を示す図である。文書生成システム200は、図1のデータベース110に相当するデータソース211からデータを取得し、指定されたテンプレートとデータソース211から取得したデータとを合成して文書を生成するシステムである。生成される文書は、指定された秘匿指示情報に従って、合成するデータの一部がマスキングされて出力されるものである。以下、文書生成システム200に入力される各情報、文書生成システム200を構成する各構成について説明する。なお、文書生成システム200は、本発明の情報処理装置の適用例となる構成である。

20

【0018】

データ取得指示情報201とは、データソース211から取得するデータを指定する情報である。データ取得指示情報201は、本発明のデータ取得情報の適用例となる構成である。

【0019】

テンプレート203とは、データソース211から取得されるデータを埋め込む指示子を含んだ文書である。本実施形態では、文書のレイアウトデザインに指示子を加えた文書をテンプレートとして説明する。また、テンプレート203は、本発明のテンプレートデータの適用例となる構成である。

30

【0020】

秘匿指示情報202とは、データ取得指示情報201に指定されたデータのマスキング方法を指示する情報である。即ち、秘匿指示情報202は、データソース211から取得するどのデータに対してマスキングを行うかをデータの属性情報により指定している。また、本実施形態では、マスキング対象のデータの指定とともに、マスキングに使用する代替データも指定する。例えば、特定の文字列や画像を代替データとして指定する。本実施形態では、黒く塗りつぶしたような出力することとともに、指定されたデータを特定のデータで置き換えて出力することもマスキングと表現する。データ取得指示情報201、テンプレート203及び秘匿指示情報202の各文書の例については後述する。また、秘匿指示情報202は、本発明の秘匿情報の適用例となる構成である。

40

【0021】

データ取得指示情報読み取り部204、秘匿指示情報読み取り部205及びテンプレート読み取り部206はそれぞれ、指定されたデータ取得指示情報201、秘匿指示情報202及びテンプレート203を読み取るための機能構成である。なお、データ取得指示情報読み取り部204及び秘匿指示情報読み取り部205は、本発明の入力手段の適用例となる構成である。

【0022】

50

データ取得部 207 は、データ取得指示情報 201 によって指定されたデータをデータソース 211 から取得する。データ取得指示情報 201 には、データ取得先と取得するデータを特定するための情報とが指定されている。例えば、データソース 211 とは、外部のリレーショナルデータベースであり、データソース 211 を操作するための言語 SQL (Structure Query Language) を使用してデータを取得する。データ取得指示情報 201 には、対象となるデータソース 211 のアクセス先、取得するデータのテーブル名、フィールド名及びデータの型等の情報を指定する。データ取得部 207 は、本発明の取得手段の適用例となる構成である。

【0023】

データ取得部 207 では、テーブル名やフィールド名から SQL の問い合わせ文を作成し、指定されたデータソース 211 にその SQL 文で問い合わせをする。そして、データ取得部 207 は、データソース 211 が返してきたデータを取得する。一方で、データ取得指示情報 201 で指定されているデータの中で、秘匿指示情報 202 にてマスキングすることが指定されているものについては、マスキングするための代替データと置き換えられる。つまり、代替データがマスキングした結果のデータ、マスキングデータとして扱われる。従って、マスキング対象のデータは、データソース 211 から取得する必要がない。そのため、データ取得部 207 は、データ取得指示情報 201 で指示されたデータのうち、秘匿指示情報 202 でマスキング対象に指定されているデータについては SQL 文で取得の指示を行わないようとする。

【0024】

このように、マスキング対象のデータについては取得しないことで、データソース 211 から冗長な情報を取得する時間を省くことができる。データソース 211 から取得したマスキング対象のデータ以外のデータをデータ記憶部 210 に格納する。本実施形態では、データ記憶部 210 は RAM 107 であるとする。

【0025】

但し、秘匿指示情報 202 に記すマスキング方法の指定として、データ全体を置き換えるものではなく、データの一部をマスキングする指定をすることもできる。その場合は、マスキング対象のデータであっても、データソース 211 からデータを取得する必要がある。例えば、データソース 211 から取得したデータが「090-1111-1111」であり、このデータに対して、マスキング方法の指示として「- - - - -」があったとする。そして、この指示が、「-」はその位置に出現するデータをそのまま出力し、それ以外の文字は代替文字としてマスキングに使用することを意味したとする。その場合、「090- - - - -」を、マスキング指示を実行した結果のデータ、マスキングデータとして扱う。このように、実際のデータを使用してマスキングデータを作成する場合を、ここでは「部分マスキング」と呼ぶことにする。

【0026】

もちろん、データ取得指示情報 201 に記述のあるデータを全てデータソース 211 から取得した後、秘匿指示情報 202 で指定されたデータ以外のデータをデータ記憶部 210 に格納してもかまわない。また、データ取得指示情報 201 で指定されている全ての情報をデータソース 211 から取得してデータ記憶部 210 に格納しておき、この後の処理で、マスキングデータ生成部 208 で作成されたマスキングデータで上書きするようにしてもよい。

【0027】

マスキングデータ生成部 208 は、秘匿指示情報読み取り部 205 から秘匿指示情報 202 を得て、マスキングデータを生成する。秘匿指示情報 202 には、データ取得指示情報 201 に記述されたデータのうち、どのデータをマスキングするか記述されている。また、マスキング対象のデータをマスキングする際の代替データが指示されている。なお、マスキングデータ生成部 208 は、本発明の生成手段の適用例となる構成であり、マスキングデータは、本発明の秘匿データの適用例となる構成である。

【0028】

10

20

30

40

50

マスキングデータ生成部 208 は、まず、マスキング対象のデータを認識する。そして、そのデータに対してマスキングに使用する代替データを確定する。代替データとして特定の文字列が指定されている場合はその文字列をデータ記憶部 210 に格納する。また、前述したように秘匿指示情報 202 に部分マスキングを行う指示がある場合は、指示に従って一部をマスキングしたデータをデータ記憶部 210 に格納する。

【0029】

文書生成部 209 は、テンプレート読み取り部 206 から得たテンプレートと、データ記憶部 210 に格納されたデータとから文書を生成する。テンプレートには文書のデザイン、文書の論理構造やテキスト、画像等の文書を構成するコンテンツを含む文書情報である。また、テンプレートにはテンプレートに合成するデータが指定されている。文書生成部 209 は、テンプレートに指定された合成対象データをデータ記憶部 210 から取得して、テンプレートに合成して文書を生成する。文書生成部 209 は、テンプレートに記述された全ての合成指示に対して合成処理を行うことによって生成文書 212 を得る。なお、文書生成部 209 は、本発明の合成手段の適用例となる構成である。

10

【0030】

図 3 に、データソース 211 から取得するデータを指示するためのデータ取得指示情報 201 の一例を示す。本実施形態では、データ取得指示情報 201 を XML (Extensible Markup Language) 文書として記述する。また、ここではデータソース 211 をデータベースとし、データベースに保存されたデータを取得するためのデータ取得指示情報 201 の例を示す。

20

【0031】

図 3 に示すように、データ取得指示情報 201 では、データベースアクセス先と取得するテーブルの種類及びフィールドが指定されている。<server>要素はデータベースへのアクセス先を指示する要素で、<http://database.hogehoge.com>でアクセスできることを示している。また、<table>要素と<field>要素はそれぞれデータベースから取得する対象データのテーブルとそのフィールドを示している。<table>要素のname属性はテーブルの名前を表している。<field>要素のtype属性はフィールドのデータの型を示している。それぞれ “string” は文字列、“email” はデータが電子メール情報を持った文字列、“address” はデータが住所情報を持った文字列、“phone” はデータが電話番号情報を持った文字列であることを示している。<field>要素の内容はフィールドの名前を表している。具体的には、このデータ取得指示情報 201 は、アクセスできるデータベースのCustomerテーブルからName、Email、Prefecture、City、HomePhone、CellPhoneフィールドのデータを取得することを意味している。

30

【0032】

図 4 は、データソース 211 であるデータベースのデータ例を示す図である。データベースのCustomerテーブルにデータが格納されている状態を示している。CustomerテーブルはName、Email、Prefecture、City、HomePhone、CellPhoneフィールドを有する。Customerテーブルには、それぞれのフィールドに値を持つID001からID003までの 3 件のデータが格納されている様子を表している。図 3 に示すデータ取得指示情報が指定されてこのテーブルにアクセスした場合は、このテーブルのID001、ID002、ID003の 3 件のデータが取得される。

40

【0033】

図 5 は、データ取得指示情報 201 で指定されたデータに対するマスキング処理について指示する秘匿指示情報 202 の一例を示す図である。本実施形態では、秘匿指示情報 202 を XML 文書として記述している。

【0034】

<Mask>要素は、秘匿情報のルート要素である。<data>要素は、データ取得指示情報 201 で指定されたデータのうち、マスキング処理を行う対象のデータを指定している。table属性は、データベースのテーブルの名前、field属性はtable属性で指定されたデータベースのテーブルのフィールド名を示している。これらは、ここで指定したテーブルのフ

50

ィールドの値をテンプレートに合成する際にマスキングすることを指示している。そして、type属性はデータの型を示し、field属性の替わりに指定されている場合にこの属性値に示された型を持つデータをマスキングすることを意味する。

【0035】

data要素の内容にはマスキング処理を行うときに使用する代替データが指定される。何も指定されていない場合はデフォルトの処理である、黒い四角でマスキング処理される。文字列が指定されている場合は、指定された文字列を代替データとしてマスキング処理を行う。この例では、CustomerテーブルのEmailフィールドの値にはデフォルトの黒い四角を代替データとしてマスキングに使用することを意味している。また、Cityフィールドの値には「非表示」の文字列を代替データとしてマスキングに使用することを意味している。そして、フィールドの型が“phone”である値には「-----」の文字列を代替データとしてマスキングに使用することを意味している。

10

【0036】

図6は、テンプレートの例を示す図である。本実施形態ではテンプレートをXMLの文書として示している。<table>、<tr>、<td>要素はHTMLの要素と同様の意味を持つものとして例を挙げた。<table>要素は表のデータであることを示し、<tr>要素は行のデータ、<td>要素は列のデータであることを示している。また、<for>要素は子要素を繰り返し出力することを意味する。listName属性値にはデータ取得指示情報201で指定されたデータベースのテーブル名が記述される。<for>要素はlistName属性値で指定されたデータを取得した個数分、子要素を繰り返し出力する処理することを意味している。

20

【0037】

<data>要素は、データ取得指示情報201により取得したデータを<data>要素の替わりに合成することを指定する要素である。name属性には合成するデータを“テーブル名”.“フィールド名”的書式で指定する。つまり、“Customer.Name”と指定した場合は、CustomerテーブルのNameフィールドの値を代入することを意味する。

【0038】

図7に、図6のテンプレートにデータベースから取得したデータを合成した例を示す。図7は、図3に示すデータ取得指示情報201、図4に示すデータベースのデータが用いられ、秘匿指示情報は指定されていない場合の生成文書212の例を示している。<for>要素の指定により、データベースから取得したCustomerテーブルの各フィールドの値が取得した順番毎に、所定の位置に合成されて出力される。ID001からID003の3データが出力される。

30

【0039】

説明をわかりやすくするため、図7を表示できる形態としてレンダリングした様子を図8に示す。図6のテンプレートはHTMLの書式にデータベースからのデータを合成するための記述を加えたものであり、文書生成部209で作成されるデータベースから取得したデータが合成された文書はHTML文書として出力される。この文書をWebブラウザで表示すると図8のような表として出力される。一行目に表の各項目名、2行目、3行目、4行目にデータベースから取得したID001からID003までのデータが表示される。

40

【0040】

図9は、文書生成部209で作成された生成文書212の例を示す図である。この生成文書212は、図3のデータ取得指示情報201、図5の秘匿指示情報202、図6のテンプレート203を使用し、データベースのCustomerテーブルに格納されたデータが図4であった場合に出力される生成文書212である。図5の秘匿指示情報に基づいて、Emailフィールド、Cityフィールド、そして型が“phone”であるHomePhone、CellPhoneフィールドに対してそれぞれに指定された代替データを使用してマスキングされている。

【0041】

図10に、図9の生成文書212をレンダリングした様子を示す。上述したように図7のテンプレート203はHTMLの書式と合成するデータを指示する書式に従って記述されているため、Webブラウザでレンダリング処理して表示した様子である。図5の秘匿

50

指示情報に基づいて指定のデータが指定の代替データでマスキングされて表示される。

【0042】

次に、図11を参照しながら、フローチャートを使用して本実施形態の処理の流れを説明する。図11は、全体の処理の流れを示すフローチャートである。

【0043】

まず、データ取得指示情報201、秘匿指示情報202、テンプレート203を読み取る処理を行う。本フローチャートでは、この順で読み込む処理を行っているが、順不同でよい。

【0044】

ステップS1101では、データ取得指示情報読み取り部204は指定されたデータ取得指示情報201を読み取り、データ取得部207に読み取った情報を渡す。

10

【0045】

ステップS1102では、秘匿指示情報読み取り部205は、指定された秘匿指示情報202を読み取り、読み取った情報をデータ取得部207、マスキングデータ生成部208に渡す。

20

【0046】

ステップS1103では、テンプレート読み取り部206は、テンプレート203を読み取り、文書生成部209に読み取ったデータを渡す。本実施形態では、データ取得指示情報201、秘匿指示情報202、テンプレート203はファイルとしてハードディスク104に保存されているので、指定されたファイルを読み込む処理を行う。もし、外部のファイルサーバ111等に保存されており、そのファイルサーバのURLが指定されている場合は、指定されたURLにアクセスしてファイルを取得して読み取り処理を行う。

20

【0047】

ステップS1104では、データ取得部207は、データ取得指示情報読み取り部204から得たデータ取得指示情報201と、秘匿指示情報読み取り部205から得た秘匿指示情報202とに基づいてデータベースであるデータソース211からデータを取得する。そして、データ取得部207はデータ記憶部210にデータを格納する。秘匿指示情報202で指定されたデータは指定された代替データでマスキングされるため、データソース211からデータを取得する必要がない。そのため、データ取得部207は、秘匿指示情報202を用いてマスキング対象のデータを判別して、データソース211から取得しないようにしている。しかし、冗長ではあるが、もし、マスキング対象のデータもデータソース211から取得してデータ記憶部210に格納するのであれば、データ取得部207は秘匿指示情報202を使用しない。この場合、秘匿指示情報読み取り部205は、ステップS1102で読み取った秘匿指示情報202をデータ取得部207に渡さなくてもよい。

30

【0048】

ステップS1105では、マスキングデータ生成部208は、秘匿指示情報208に従って、マスキングに使用する代替データをマスキング対象のデータに対応させてデータ記憶部210に記憶する。

30

【0049】

ステップS1106では、文書生成部209は、テンプレート読み取り部206から取得したテンプレート203とデータ記憶部210に格納されたデータとを使用して生成文書212を作成する。テンプレート203には、テンプレートに合成するデータが記述されている。文書生成部209は、その記述で指定されたデータをデータ記憶部210から取得してテンプレートに合成して生成文書212を作成する。

40

【0050】

次に、図12を参照しながら、データ取得部207の処理について説明する。図12は、データ取得部207の処理を示すフローチャートである。ここでは、データ取得指示情報読み取り部204から図3に示すデータ取得指示情報を得た後の処理について説明する。

50

【0051】

ステップS1201では、データ取得部207は、データ取得指示情報201から取得するデータを認識する。ここでは、データ取得部207は、データ取得指示情報201を上から順次読み出して各要素の情報を処理する。具体的には、データ取得部207は、<server>要素の内容である "http://database.hogehoge.com" をアクセス先として検出する。続いて、データ取得部207は、<table>要素のname属性からテーブル名、<field>のtype属性と内容からフィールドの型とフィールド名を検出する。図3のデータ取得指示情報201の場合は以下のデータが検出される。CustomerテーブルのNameフィールド(string型)、Emailフィールド(email型)、Prefecture(address型)、Cityフィールド(address型)、HomePhoneフィールド(phone型)、CellPhone(phone型)である。

10

【0052】

ステップS1202では、データ取得部207は、秘匿指示情報を上から順次読み出して、マスキング処理する情報を検出する。図5が秘匿指示情報である場合、マスキング対象データのテーブル名とフィールド名、フィールドの型が、<data>要素のtable属性値、field属性値、type属性値として記述されている。つまり、データ取得部207は、CustomerテーブルのEmailフィールド、Cityフィールド、そして、データ型がphoneである、HomePhoneフィールド、CellPhoneフィールドのデータをマスキング対象データとして検出する。

20

【0053】

ステップS1203では、データ取得部207は、データソース211から取得するデータを決定する。データ取得部207は、ステップS1201で検出したフィールド名からS1202で検出したデータ名を抜いたフィールド名のデータを取得する。つまり、データ取得部207は、Name、Email、Prefecture、City、HomePhone、CellPhoneフィールドからEmail、City、HomePhone、CellPhoneを抜いた、Name、Prefectureフィールドのデータを取得することを決定する。

20

【0054】

ステップS1204では、データ取得部207は、上記決定を基にデータソース211からデータを取得する。ステップS1204では先ず、データ取得部207は、ステップS1203で決定したフィールドに関してデータソース211から取得するためのSQL文を作成する。ステップS1203までにCustomerテーブルのName、Prefectureフィールドの値を取得することが決定している。従って、データ取得部207は、そのデータを取得するためのSQL文「SELECT Name, Prefecture FROM Customer」を作成する。そして、データ取得部207は、データソース211であるデータベースdatabase.hogehoge.comに対して、作成したSQL文で問い合わせてデータを取得する。

30

【0055】

しかしながら、秘匿指示情報で全てのフィールドに対してマスキング処理が指定されている場合もある。その場合は、Customerテーブルに登録されているデータの件数を取得するためのSQL文「SELECT COUNT(*) FROM Customer」を作成し、データソース211に問い合わせる。

40

【0056】

ステップS1205では、データ取得部207は、データソース211から取得したデータをデータ記憶部210に格納する。その際、データ取得部207は、データ記憶部210にCustomerデータを格納するためのデータ管理テーブルを作成して、そこに取得したデータを格納する。

【0057】

データ管理テーブルの例を図13に示す。データ管理テーブル1501はデータ取得指示情報に指定された全てのフィールド名とその型の情報とを有し、データソース211から取得したデータを対応したフィールド名の列に格納する。

【0058】

図13では、ステップS1204で、ID001からID003のNameフィールド、Prefectureフ

50

フィールドの値を取得した場合のデータ格納状態を表している。もし、秘匿指示情報で、データ取得指示情報に記述された全てのデータがマスキング処理されることが指定されている場合、データ管理テーブルは、ステップ S 1 2 0 4 で取得したデータの件数のみを格納したものになる。

【0059】

次に、マスキングデータ生成部 2 0 8 の処理の流れについて説明する。秘匿指示情報読み取り部 2 0 5 から図 5 の秘匿指示情報 2 0 2 を取得した後の処理について説明する。

【0060】

まず、マスキングデータ生成部 2 0 8 は、取得した秘匿指示情報 2 0 2 に記載されたデータフィールドとマスキング文字列を抽出する。即ち、マスキングデータ生成部 2 0 8 は、図 5 の秘匿指示情報 2 0 2 から<data>要素のtable属性値、field属性値、type属性値、そして、内容を読み出す。前述したようにtable属性はデータソース 2 1 1 のテーブル名、field属性はフィールド名、type属性はフィールドのデータ型を示している。属性値でデータソース 2 1 1 上のどのフィールドのデータか特定することができる。また、<data>要素の内容はマスキング処理をするときに使用する代替データを示しており、指定された文字列が代替データとして使用される。内容がない場合はデフォルトの文字列である、黒い四角が使用されることを意味する。この意味に従い、マスキングデータ生成部 2 0 8 は、図 5 の秘匿指示情報 2 0 2 は上から順に、CustomerテーブルのEmailフィールドをデフォルトの黒い四角の文字列を使用してマスキング処理することを認識する。また、マスキングデータ生成部 2 0 8 は、CustomerテーブルのCityフィールドを「非表示」の文字列でマスキング処理することを認識する。また、マスキングデータ生成部 2 0 8 は、Customerテーブルのデータ型が "phone" であるフィールドを「 - - - - - 」でマスキング処理することを認識する。

10

20

20

【0061】

次に、マスキングデータ生成部 2 0 8 は、S 1 3 0 1 で認識した情報に従ってデータ記憶部 2 1 0 のデータ管理テーブルにマスキングに使用する代替データを格納する。代替データを格納した結果のデータ管理テーブルを図 1 4 に示す。

【0062】

図 1 4 に示すように、データ管理テーブルのEmailフィールドに対しては黒い四角を代替データとして用いるので、データ管理テーブルのEmailフィールドの列に黒い四角を格納する。そして、データ管理テーブルのCityフィールドに対しては「非表示」という文字列を用いる。従って、マスキングデータ生成部 2 0 8 は、データ管理テーブルのCityフィールドの列に「非表示」の文字列を格納する。また、データ管理テーブルのphone型のフィールドに対しては、「 - - - - - 」文字列をマスキング処理に用いる。従って、マスキングデータ生成部 2 0 8 は、データ管理テーブルのphone型データを検索して、代替データ「 - - - - - 」を格納する。この例の場合、HomePhoneフィールド、CellPhoneフィールドがphone型と検索されるので、マスキングデータ生成部 2 0 8 は、この 2 つのフィールドの列に対してこの代替データを格納する。以上のようにしてデータ管理テーブルにマスキングに使用する代替データを格納した結果、図 1 4 に示すようなデータ管理テーブル 1 6 0 1 が完成する。

30

40

【0063】

最後に、文書生成部 2 0 9 の処理の流れを図 1 5 のフローチャートを用いて説明する。まず、文書生成部 2 0 9 は、テンプレート読み取り部 2 0 6 から得たテンプレートを読み込んでテンプレート中のデータ合成に関する指示が記載されている記述を検索する。図 6 に示すテンプレートの場合は、子要素の繰り返し出力を指示する<for>要素と指定するデータとの入れ替えを指示する<data>要素がこれに該当する。次に、文書生成部 2 0 9 は、データ合成に関する指示記述に従い、データ管理テーブルから指定されたデータを取得してテンプレートにデータを合成する。

【0064】

図 6 に示すテンプレートに対する文書生成部 2 0 9 の処理の流れを図 1 5 のフローチャ

50

ートを用いて説明する。ここでは、XMLの要素を上から順次処理する。

【0065】

ステップS1801では、文書生成部209は、テンプレート中の全ての要素を処理したか否かを判定し、全て処理した場合に文書生成部209の処理が終了する。それまでは、文書生成部209は読み出したXMLの要素に従った処理を行う。

【0066】

ステップS1802では、文書生成部209は、読み出した要素が`<for>`要素かを判別する。`<for>`要素である場合、文書生成部209は、`<for>`要素の`listName`属性値に記述のあるテーブル名のデータ管理テーブルから格納されているデータの数を取得し、その回数、子要素以下の要素に対する出力処理を繰り返す。`<for>`要素は子要素の繰り返し処理を指定するための要素であるため、生成文書には出力しない。テンプレート601の`for`要素の`listName`属性値は"Customer"なので、文書生成部209は、データ記憶部210に記憶されたデータ管理テーブルからデータの個数を取得する。

10

【0067】

図14に示すデータ管理テーブルの場合、データの件数は3であるため、`<for>`要素の子要素`<tr>`及びそれ以下の要素の出力が3回繰り返されることを意味する。そして、さらに、子要素及びそれ以下の要素の`<data>`要素でデータ管理テーブルのデータを合成する指示がある場合、文書生成部209は、繰り返す毎にデータ管理テーブルから順次Noの小さいデータから大きいデータを読み出し、データ管理テーブルに格納されているデータを全てテンプレートに合成することを意味する。

20

【0068】

ステップS1804では、文書生成部209は、読み出した要素が`<data>`要素であるか否かを判別する。`<data>`要素である場合、文書生成部209は、`<data>`要素の`name`属性値に指定されたデータをデータ管理テーブルから取得し、この`<data>`要素の替わりに生成文書のデータとして出力する(S1805)。例えば、`name`属性値に"Customer.Name"記述されている場合、文書生成部209は、データ管理テーブル1601からNameフィールドの値を読み出す。`<for>`要素による最初の処理であればNo.1に記憶されている「田中花子」を生成文書に出力する。同様に、`name`属性値、"Customer.Email"、"Customer.Prefecture"、"Customer.City"、"Customer.HomePhone"、"Customer.CellPhone"に対してはそれぞれ、「
- - - - -」、「東京都」、「非表示」、「
- - - - -」、「
- - - - -」を出力する。

30

【0069】

テンプレートから読み出した要素がデータ合成に関するものでないと判断した場合、つまり、本実施形態においては、文書生成部209は、`<for>`要素、`<data>`要素以外の要素であると判断した場合、要素をそのまま出力する(S1806)。

【0070】

テンプレート中の全ての要素を処理し、文書生成部209は処理を終える。出力された要素で構成された文書が生成文書212である。

【0071】

図6に示すテンプレートに対して図14のデータ管理テーブルが用いられた場合、図9に示す生成文書が作成される。

40

【0072】

以上、説明した方法によりテンプレートにマスキング処理を指定せずにマスキング処理した文書の生成を行う。

【0073】

なお、本実施形態では、テンプレートとしてXML文書を例に挙げて説明したが、それに限定するものではなく、他の形式の文書であってもよい。テンプレートとは、最終的に生成する文書のフォーマットに、データソース211から取得したデータを合成する指示記述を加えたものである。また、データ合成を指示するための記述とは、それ以外の記述、つまり最終的に生成する文書フォーマットに従った記述と区別可能な記述である。本実

50

施形態における文書生成部 209 は、データ合成を指示する記述以外はそのまま出力するため、そのまま出力するデータのフォーマットは認識しない。それゆえ、生成文書 212 は本実施形態で説明した XML 文書に限るものではない。例えば、PDF (Portable Document Format) やマイクロソフト社の Word、CSV (Comma Separated Values) データ等の他形式の文書であってもよい。

【0074】

本実施形態では 1 つの秘匿指示情報を例に挙げて説明した。もし、文書システムを使用するユーザやユーザが扱える機密情報のレベルによってマスキングされる箇所が異なる文書を生成したい場合は、ユーザや機密情報のレベルに応じて秘匿指示情報読み取り部に渡す秘匿指示情報を変えることにより実現できる。

10

【0075】

上述したように、マスキングデータ生成部 208 は、データ取得指示情報 201 に基づいて取得されるデータのうち、秘匿するデータを指定するための秘匿指示情報 292 に基づいて上記秘匿するデータに対応するマスキングデータを生成する。そして、文書生成部 209 は、取得されたデータとマスキングデータとをテンプレート 203 に合成するよう正在実施している。従って、テンプレート 203 を変更する場合であっても、データを秘匿する方法を改めて指定することなく、マスキングされた文書等を生成することが可能となる。

20

【0076】

このように、本実施形態においては、テンプレート 203 のデザインだけを変更して同じデータを表示する場合、マスキング処理についての指定を変更することなく対応可能である。従って、マスキング箇所を指定することなく、目的であるテンプレート 203 のデザインを変更するだけで、マスキングされた文書を作成することができる。

20

【0077】

また、本実施形態においては、マスキング箇所の指定はデータソース 211 から取得されるデータに対して行う。従って、テンプレート 203 の複数の箇所にデータをレイアウトする場合であっても、テンプレート 203 のどこにデータが出現するのかを検索しながら、その各々の箇所にマスキングを指定する手間が省ける。

30

【0078】

なお、テンプレート 203 とデータソース 211 とは密接な関係がある。そのため、テンプレート 203 とデータソース 211 とは、一対として扱われるこが考えられる。このような場合、テンプレート 203 にマスキング箇所を指定する方法では、マスキングするデータの違いによってテンプレート 203 とデータソース 211 との対を複数用意する必要がある。しかし、本実施形態においては、マスキング箇所の指定をテンプレート 203 に行わないため、テンプレート 203 とデータソース 211 との対を複数用意する必要がない。

30

【0079】

上述した本発明の実施形態を構成する各手段及び各ステップは、コンピュータの RAM や ROM 等に記憶されたプログラムが動作することによって実現できる。このプログラム及び前記プログラムを記録したコンピュータ読み取り可能な記録媒体は本発明に含まれる。

40

【0080】

また、本発明は、例えば、システム、装置、方法、プログラムもしくは記録媒体等としての実施形態も可能であり、具体的には、一つの機器からなる装置に適用してもよい。

【0081】

なお、本発明は、上述した実施形態の機能を実現するソフトウェアのプログラムを、システム又は装置に直接、又は遠隔から供給する。そして、そのシステム又は装置のコンピュータが前記供給されたプログラムコードを読み出して実行することによっても達成される場合を含む。

【0082】

従って、本発明の機能処理をコンピュータで実現するために、前記コンピュータにイン

50

ストールされるプログラムコード自体も本発明を実現するものである。つまり、本発明は、本発明の機能処理を実現するためのコンピュータプログラム自体も含まれる。その場合、プログラムの機能を有していれば、オブジェクトコード、インタプリタにより実行されるプログラム、OSに供給するスクリプトデータ等の形態であってもよい。

【0083】

また、コンピュータが、読み出したプログラムを実行することによって、前述した実施形態の機能が実現される。更に、そのプログラムの指示に基づき、コンピュータ上で稼動しているOS等が、実際の処理の一部又は全部を行い、その処理によっても前述した実施形態の機能が実現され得る。

【0084】

更に、その他の方法として、まず記録媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれる。そして、そのプログラムの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPU等が実際の処理の一部又は全部を行い、その処理によっても前述した実施形態の機能が実現される。

【図面の簡単な説明】

【0085】

【図1】本発明の実施形態に係る情報処理装置の構成を示すブロック図である。

【図2】本発明の実施形態に係る情報処理装置によって実現する文書生成システムの構成を示す図である。

【図3】データソースから取得するデータを指示するためのデータ取得指示情報の一例を示す図である。

【図4】データソースであるデータベースのデータ例を示す図である。

【図5】データ取得指示情報で指定されたデータに対するマスキング処理について指示した秘匿指示情報の一例を示す図である。

【図6】テンプレートの例を示す図である。

【図7】図6のテンプレートにデータベースから取得したデータを合成した例を示す図である。

【図8】図7を表示できる形態としてレンダリングした様子を示す図である。

【図9】文書生成部で作成された生成文書の例を示す図である。

【図10】図9の生成文書をレンダリングした様子を示す図である。

【図11】全体の処理の流れを示すフローチャートである。

【図12】データ取得部の処理を示すフローチャートである。

【図13】データ管理テーブルの例を示す図である。

【図14】代替データを格納した結果のデータ管理テーブルの例を示す図である。

【図15】文書生成部の処理の流れを示すフローチャートである。

【符号の説明】

【0086】

100 情報処理装置

101 コントローラ部

102 操作部

103 表示部

104 ハードディスク

105 CPU

106 ROM

107 RAM

108 ネットワークアダプタ

109 ネットワーク

110 データベース

111 ファイルサーバ

10

20

30

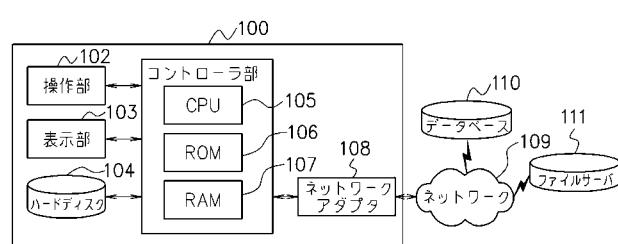
40

50

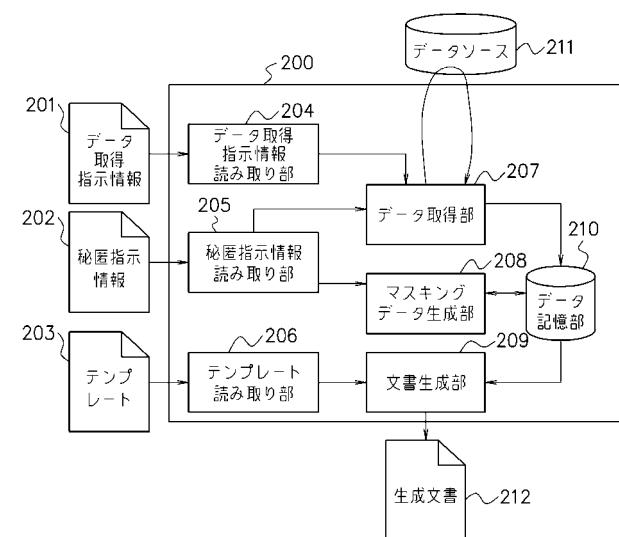
2 0 0 文書生成システム
 2 0 1 データ取得指示情報
 2 0 2 秘匿指示情報
 2 0 3 テンプレート
 2 0 4 データ取得指示情報読み取り部
 2 0 5 秘匿指示情報読み取り部
 2 0 6 テンプレート読み取り部
 2 0 7 データ取得部
 2 0 8 マスキングデータ生成部
 2 0 9 文書生成部
 2 1 0 データ記憶部
 2 1 1 データソース
 2 1 2 生成文書

10

【図1】



【図2】



【図3】

```

<dataRetrieve>
  <server>
    http://database.hogehoge.com
  </server>
  <table name="Customer">
    <field type="string">Name</field>
    <field type="email">Email</field>
    <field type="address">Prefecture</field>
    <field type="address">City</field>
    <field type="phone">HomePhone</field>
    <field type="phone">CellPhone</field>
  </table>
</dataRetrieve>

```

201

【図4】

Customer テーブル					
ID	Name	Email	Prefecture	City	HomePhone
001	田中花子	tanaka@abcd.com	東京都	大田区	03-1234-5678
002	山田太郎	yamada@defg.com	神奈川県	横浜市	045-4621-8765
003	佐藤一郎	sato@hijk.com	北海道	千歳市	0123-45-6789

【図5】

```

<Mask>
<data table="Customer" field="Email"></data>
<data table="Customer" field="City">非表示</data>
<data table="Customer" type="phone">●●●-●●●-●●●●</data>
</Mask>

```

202

【図7】

名前	メール	都道府県	市区町村	自宅電話	携帯電話
田中花子	tanaka@abcd.com	東京都	大田区	03-1234-5678	090-111-1111
山田太郎	yamada@defg.com	神奈川県	横浜市	045-4621-8765	090-222-2222
佐藤一郎	sato@hijk.com	北海道	千歳市	0123-45-6789	090-333-3333

【図6】

```

<table border="1">
<tr>
  <td>名前</td>
  <td>メール</td>
  <td>都道府県</td>
  <td>市区町村</td>
  <td>自宅電話</td>
  <td>携帯電話</td>
</tr>
<for listName="Customer" />
<tr>
  <td><data name="Customer.Name"/></td>
  <td><data name="Customer.Email"/></td>
  <td><data name="Customer.Prefecture"/></td>
  <td><data name="Customer.City"/></td>
  <td><data name="Customer.HomePhone"/></td>
  <td><data name="Customer.CellPhone"/></td>
</tr>
</for>
</table>

```

601

【図 8】

名前	メール	都道府県	市区町村	市町村	自宅電話	携帯電話
田中花子	tanaka@abcd.com	東京都	大田区	03-1234-5678	090-111-1111	
山田太郎	yamada@defg.com	神奈川県	横浜市	045-4321-8765	090-222-2222	
佐藤一郎	sato@hk.com	北海道	千歳市	0123-45-6789	090-333-3333	

【図 9】

```

<table border="1">
<tr>
<td>名前</td>
<td>メール</td>
<td>都道府県</td>
<td>市区町村</td>
<td>自宅電話</td>
<td>携帯電話</td>
</tr>
<tr>
<td>田中花子 </td>
<td>■■■■■ </td>
<td>東京都 </td>
<td>非表示 </td>
<td>●●●-●●●●-●●●●</td>
<td>●●●-●●●●-●●●●</td>
</tr>
<tr>
<td>山田太郎 </td>
<td>■■■■■ </td>
<td>神奈川県 </td>
<td>非表示 </td>
<td>●●●-●●●●-●●●●</td>
<td>●●●-●●●●-●●●●</td>
</tr>
<tr>
<td>佐藤一郎 </td>
<td>■■■■■ </td>
<td>北海道 </td>
<td>非表示 </td>
<td>●●●-●●●●-●●●●</td>
<td>●●●-●●●●-●●●●</td>
</tr>
</table>

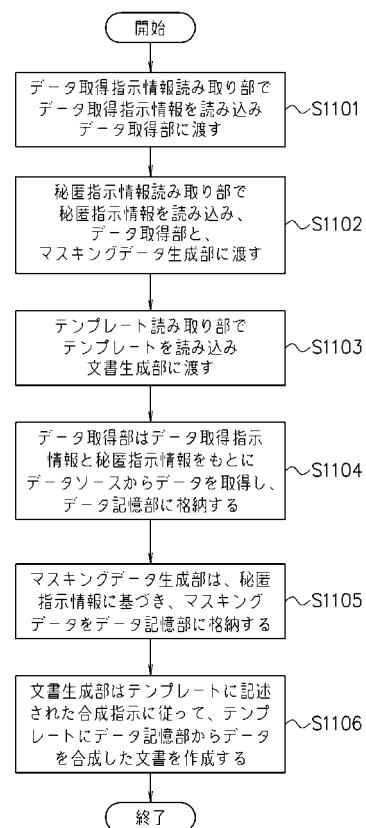
```

212

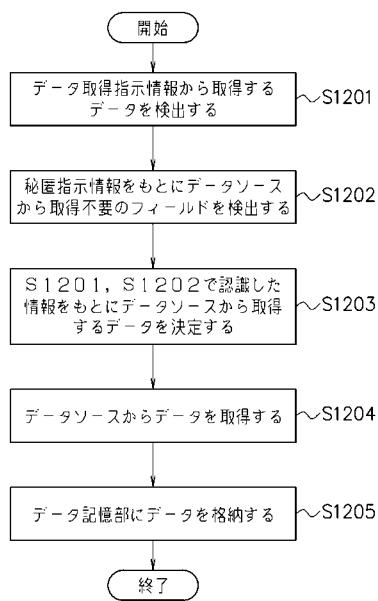
【図 10】

名前	メール	都道府県	市区町村	自宅電話	携帯電話
田中花子	tanaka@abcd.com	東京都	大田区	03-1234-5678	090-111-1111
山田太郎	yamada@defg.com	神奈川県	横浜市	045-4321-8765	090-222-2222
佐藤一郎	sato@hk.com	北海道	千歳市	0123-45-6789	090-333-3333

【図 11】



【図12】



【図13】

データ記憶部 (データ記憶部)

No.	Name (string型)	Email (email型)	Prefecture (address型)	City (address型)	HomePhone (phone型)	CellPhone (phone型)
1	田中花子	■■■■■	東京都	非表示	■■■■■	■■■■■
2	山田太郎	■■■■■	神奈川県	非表示	■■■■■	■■■■■
3	佐藤一郎	■■■■■	北海道	非表示	■■■■■	■■■■■

合計件数: 3

【図14】

データ管理テーブル

No.	Name (string型)	Email (email型)	Prefecture (address型)	City (address型)	HomePhone (phone型)	CellPhone (phone型)
1	田中花子	■■■■■	東京都	非表示	■■■■■	■■■■■
2	山田太郎	■■■■■	神奈川県	非表示	■■■■■	■■■■■
3	佐藤一郎	■■■■■	北海道	非表示	■■■■■	■■■■■

合計件数: 3

【図15】

