

(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2013-105207

(P2013-105207A)

(43) 公開日 平成25年5月30日(2013.5.30)

(51) Int.Cl.		F I		テーマコード (参考)
G06F 17/30	(2006.01)	G06F 17/30	120A	
G06F 21/60	(2013.01)	G06F 21/24	160C	
		G06F 17/30	350C	

審査請求 未請求 請求項の数 13 O L (全 28 頁)

(21) 出願番号	特願2011-246817 (P2011-246817)	(71) 出願人	000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(22) 出願日	平成23年11月10日 (2011.11.10)	(74) 代理人	100103528 弁理士 原田 一男
		(72) 発明者	高 杰 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		(72) 発明者	片山 佳則 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		(72) 発明者	森川 郁也 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

最終頁に続く

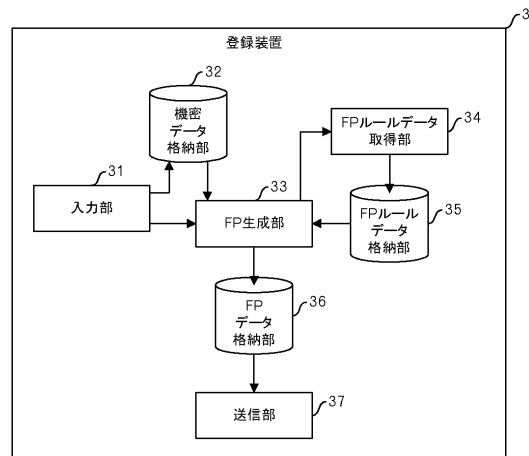
(54) 【発明の名称】 秘匿化データの検索のための情報処理方法及び装置

(57) 【要約】

【課題】 秘匿化したまま類似するデータを抽出できるようにする。

【解決手段】 本方法は、データ格納部に格納されており且つ第1の数値を含むテキストデータから、第1の数値及び当該第1の数値の周辺に存在する複数個の特徴語を抽出する工程と、抽出された第1の数値から、当該第1の数値と近似するか否かを判断する上で基準となる1又は複数の第2の数値を生成する工程と、1又は複数の第2の数値と複数個の特徴語との各々について秘匿化処理を行って秘匿化データを生成し、データ格納部に格納する工程とを含む。

【選択図】 図2



【特許請求の範囲】

【請求項 1】

データ格納部に格納されており且つ第 1 の数値を含むテキストデータから、第 1 の数値及び当該第 1 の数値の周辺に存在する複数個の特徴語を抽出する処理と、

抽出された前記第 1 の数値から、当該第 1 の数値と近似するか否かを判断する上で基準となる 1 又は複数の第 2 の数値を生成する生成処理と、

前記 1 又は複数の第 2 の数値と前記複数個の特徴語との各々について秘匿化処理を行って秘匿化データを生成し、前記データ格納部に格納する処理と、

を含む処理を、コンピュータが実行する情報処理方法。

【請求項 2】

前記生成処理が、

抽出された前記第 1 の数値を異なる有効桁数で表した複数の第 2 の数値を生成する処理である請求項 1 記載の情報処理方法。

【請求項 3】

前記生成処理が、

抽出された前記第 1 の数値を含む所定の数値範囲の上限値及び下限値である複数の第 2 の数値を特定する処理と、

前記第 1 の数値と前記下限値との差と、前記第 1 の数値と前記上限値との差とを算出し、前記データ格納部に格納する処理と、

を含む請求項 1 記載の情報処理方法。

【請求項 4】

前記生成処理が、

抽出された前記第 1 の数値と、数値を分類するための数値範囲の設定とから、前記第 1 の数値を代表する 1 又は複数の第 2 の数値を特定する処理と、

前記 1 又は複数の第 2 の数値と、前記第 1 の数値との差を算出し、前記データ格納部に格納する処理と、

を含む請求項 1 記載の情報処理方法。

【請求項 5】

前記生成処理が、

抽出された前記第 1 の数値を、指示された有効桁数で表した第 2 の数値を 1 つ生成する処理

である請求項 1 記載の情報処理方法。

【請求項 6】

前記複数個の特徴語の同義語を抽出する処理と、

前記同義語の秘匿化を行って秘匿化データを生成し、前記データ格納部に格納する処理と、

を前記処理がさらに含む請求項 1 乃至 5 のいずれか 1 つ記載の情報処理方法。

【請求項 7】

第 1 の数値の第 1 の秘匿化データ値と複数個の第 1 の特徴語の第 2 の秘匿化データ値とを含む 1 又は複数の検索データブロックを含む検索要求を受信する処理と、

複数の第 2 の数値の第 3 の秘匿化データ値と複数個の第 2 の特徴語の第 4 の秘匿化データ値とを含む 1 又は複数のデータブロックと識別情報とを含む案件データブロックを複数格納するデータ格納部に格納されている前記案件データブロックの各々について、前記第 1 の秘匿化データ値と前記第 3 の秘匿化データ値とから算出される、数値についての第 1 の類似度と、前記第 2 の秘匿化データ値と一致する前記第 4 の秘匿化データ値の個数とから、処理対象の案件データブロックに含まれる前記データブロックと前記検索データブロックとの各組み合わせについての第 2 の類似度の合計値である第 3 の類似度を算出する算出処理と、

前記第 3 の類似度が閾値を超えた案件データブロックの識別情報又は前記第 3 の類似度が上位所定数の案件データブロックの識別情報を、前記検索要求の送信元に送信する処理

10

20

30

40

50

と、

を含む処理を、コンピュータが実行する情報処理方法。

【請求項 8】

前記算出処理が、

前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値が存在する場合には前記第 1 の類似度を 1 に設定し、前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値が存在しない場合には前記第 1 の類似度を 0 に設定する処理

を含む請求項 7 記載の情報処理方法。

【請求項 9】

前記第 1 の秘匿化データ値が、前記第 1 の数値の元の数値についての代表値の秘匿化データ値であり、

前記検索要求には、前記第 1 の数値の元の数値についての代表値との差である第 1 の補助数値と、近似判定のための範囲のデータとをさらに含み、

前記複数の第 2 の数値が、元の数値が属する値域の下限値及び上限値であり、

前記データブロックには、前記第 2 の数値の元の数値が属する値域の下限値と当該元の数値との差である第 2 の補助数値と当該元の数値と前記上限値との差である第 3 の補助数値とをさらに含み、

前記算出処理が、

前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値が存在する場合には、前記第 1 の秘匿化データ値についての前記第 1 の補助数値と、前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値についての前記第 2 の補助数値又は前記第 3 の補助数値との差を算出する処理と、

前記第 1 の補助数値と前記第 2 の補助数値又は前記第 3 の補助数値との差が、前記近似判定のための範囲内であれば、前記第 1 の類似度を 1 に設定し、前記第 1 の補助数値と前記第 2 の補助数値又は前記第 3 の補助数値との差が、前記近似判定のための範囲内でない場合には前記第 1 の類似度を 0 に設定する処理と、

を含む請求項 7 記載の情報処理方法。

【請求項 10】

前記第 1 の数値の第 1 の秘匿化データ値が複数前記データブロックに含まれ、

前記算出処理が、

前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値の個数に応じた類似度を前記第 1 の類似度に設定する処理

を含む請求項 7 記載の情報処理方法。

【請求項 11】

前記第 1 の秘匿化データ値が、前記第 1 の数値の元の数値についての代表値の秘匿化データ値であり、

前記検索要求には、前記第 1 の数値の元の数値についての代表値との差である第 1 の補助数値と、近似判定のための範囲のデータとをさらに含み、

前記複数の第 2 の数値が、元の数値が属する値域の下限値及び上限値であり、

前記データブロックには、前記第 2 の数値の元の数値が属する値域の下限値と当該元の数値との差である第 2 の補助数値と当該元の数値と前記上限値との差である第 3 の補助数値とをさらに含み、

前記算出処理が、

前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値が存在する場合には、前記第 1 の秘匿化データ値についての前記第 1 の補助数値と、前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値についての前記第 2 の補助数値又は前記第 3 の補助数値との差を算出する処理と、

前記第 1 の補助数値と前記第 2 の補助数値又は前記第 3 の補助数値と、前記近似判定のための範囲を表す数値と、の差に応じた類似度を前記第 1 の類似度に設定する処理と、

を含む請求項 7 記載の情報処理方法。

10

20

30

40

50

【請求項 1 2】

データ格納部と、

前記データ格納部に格納されており且つ第 1 の数値を含むテキストデータから、第 1 の数値及び当該第 1 の数値の周辺に存在する複数個の特徴語を抽出し、抽出された前記第 1 の数値から、当該第 1 の数値と近似するか否かを判断する上で基準となる 1 又は複数の第 2 の数値を生成し、前記 1 又は複数の第 2 の数値と前記複数個の特徴語との各々について秘匿化処理を行って秘匿化データを生成し、第 2 のデータ格納部に格納する生成部と、
を有する情報処理装置。

【請求項 1 3】

第 1 の数値の第 1 の秘匿化データ値と複数個の第 1 の特徴語の第 2 の秘匿化データ値とを含む 1 又は複数の検索データブロックを含む検索要求を受信する受信部と、

複数の第 2 の数値の第 3 の秘匿化データ値と複数個の第 2 の特徴語の第 4 の秘匿化データ値とを含む 1 又は複数のデータブロックと識別情報とを含む案件データブロックを複数格納するデータ格納部に格納されている前記案件データブロックの各々について、前記第 1 の秘匿化データ値と前記第 3 の秘匿化データ値とから算出される、数値についての第 1 の類似度と、前記第 2 の秘匿化データ値と一致する前記第 4 の秘匿化データ値の個数とから、処理対象の案件データブロックに含まれる前記データブロックと前記検索データブロックとの各組み合わせについての第 2 の類似度の合計値である第 3 の類似度を算出する検索処理部と、

前記第 3 の類似度が閾値を超えた案件データブロックの識別情報又は前記第 3 の類似度が上位所定数の案件データブロックの識別情報を、前記検索要求の送信元に送信する送信部と、

を有する情報処理装置。

【発明の詳細な説明】

【技術分野】

【0001】

本技術は、秘匿化データの検索技術に関する。

【背景技術】

【0002】

クラウドの広がりと共に、情報をクラウドに預けてクラウド本来の特徴を生かした情報共有及び活用が進んでいる。その中で、クラウドでの協業や分業における機密データの活用が期待されている。例えば、個人が健康に関する情報をクラウドに預け、これを信頼できる公的機関などに分析及び整理してもらうというような利用方法が考えられる。

【0003】

こういった場面では、数値を含むテキストデータが共有される。例えば、医療関係では体温や血圧など患者の検査データに数値が含まれることになる。このようなデータを共有することは、関係者には有用である。

【0004】

一方で、セキュリティとプライバシー保護のため、このようなデータは秘匿化してからクラウドに預けるのは一般的である。そうすると、セキュリティとプライバシーが守られるが、データの活用という面では制限が生ずる。すなわち、秘匿化データは、従来の分析アプリケーションや検索サービスでは適切に処理できない。例えば、患者の症状と類似する診療例を検索したい場合であっても、診療データが秘匿化されていると、単純な検索では適切な診療例を見つけることが難しい。

【0005】

なお、文書を検索キーワードで検索して、検索キーワードが出現すると当該文書内で検索キーワードに関連する数値と、検索キーワードと共に指定された数値とを比較するような技術が存在している。しかしながら、データを秘匿化することは考慮されていないので、秘匿化すると適切なデータを検索で抽出することは難しい。

【0006】

10

20

30

40

50

また、検索対象のデータを秘匿化してサーバに保持しておき、検索時にも検索条件を同じように秘匿化して検索を行う技術も存在している。しかしながら、暗号化やハッシュ値算出を行うと、完全一致するデータのみしか抽出できないという問題がある。

【0007】

さらに、検索条件入力データとして、数値範囲を指定することができ、数値範囲に少なくとも一致するデータを抽出する技術も存在している。しかしながら、暗号化やハッシュ値算出を行うことは考慮されていない。

【0008】

また、秘匿化したデータを一旦安全な場所で復元し、検索条件とマッチング処理を行う技術も存在している。マッチング処理は平文の状態で行われるので、類似するデータをも抽出できるが、検索条件についても平文で入力するので、マッチング処理を行うサーバには検索条件は知られてしまう。

【先行技術文献】

【特許文献】

【0009】

【特許文献1】特開2000-11001号公報

【特許文献2】特開2007-52698号公報

【特許文献3】特開2005-242740号公報

【特許文献4】特開2002-108911号公報

【特許文献5】特開2004-213649号公報

【特許文献6】特開平1-58019号公報

【発明の概要】

【発明が解決しようとする課題】

【0010】

従って、本技術の目的は、一側面としては、秘匿化したまま類似するデータを抽出できるようにするための技術を提供することである。

【課題を解決するための手段】

【0011】

本技術の第1の形態に係る情報処理方法は、(A)データ格納部に格納されており且つ第1の数値を含むテキストデータから、第1の数値及び当該第1の数値の周辺に存在する複数の特徴語を抽出するステップと、(B)抽出された第1の数値から、当該第1の数値と近似するか否かを判断する上で基準となる1又は複数の第2の数値を生成する生成ステップと、(C)1又は複数の第2の数値と複数の特徴語との各々について秘匿化処理を行って秘匿化データを生成し、データ格納部に格納するステップとを含む。

【0012】

本技術の第2の技術に係る情報処理方法は、(A)第1の数値の第1の秘匿化データ値と複数の第1の特徴語の第2の秘匿化データ値とを含む1又は複数の検索データブロックを含む検索要求を受信するステップと、(B)複数の第2の数値の第3の秘匿化データ値と複数の第2の特徴語の第4の秘匿化データ値とを含む1又は複数のデータブロックと識別情報とを含む案件データブロックを複数格納するデータ格納部に格納されている案件データブロックの各々について、第1の秘匿化データ値と第3の秘匿化データ値とから算出される、数値についての第1の類似度と、第2の秘匿化データ値と一致する第4の秘匿化データ値の個数とから、処理対象の案件データブロックに含まれるデータブロックと検索データブロックとの各組み合わせについての第2の類似度の合計値である第3の類似度を算出する算出ステップと、(C)第3の類似度が閾値を超えた案件データブロックの識別情報又は第3の類似度が上位所定数の案件データブロックの識別情報を、検索要求の送信元に送信するステップとを含む。

【発明の効果】

【0013】

本技術の一側面によれば、秘匿化したまま類似するデータを抽出できるようになる。

【図面の簡単な説明】

【0014】

【図1】図1は、実施の形態のシステム構成図である。

【図2】図2は、登録装置の機能ブロック図である。

【図3】図3は、管理装置の機能ブロック図である。

【図4】図4は、検索装置の機能ブロック図である。

【図5】図5は、登録時の処理フローを示す図である。

【図6】図6は、FP生成処理の処理フローを示す図である。

【図7】図7は、機密データの一例を示す図である。

【図8A】図8Aは、データブロック（第1の方式）の一例を示す図である。

10

【図8B】図8Bは、データブロック（第2の方式）の一例を示す図である。

【図9A】図9Aは、データブロック（第1の方式）の他の例を示す図である。

【図9B】図9Bは、データブロック（第2の方式）の他の例を示す図である。

【図10】図10は、管理装置のDBに蓄積されるデータの一例を示す図である。

【図11】図11は、検索時の処理フローを示す図である。

【図12】図12は、第2FP生成処理の処理フローを示す図である。

【図13】図13は、検索条件となる機密データの一例を示す図である。

【図14A】図14Aは、検索FPデータ（第1の方式の第1の例）の一例を示す図である。

【図14B】図14Bは、検索FPデータ（第1の方式の第2の例）の一例を示す図である。

20

【図15A】図15Aは、検索FPデータ（第2の方式の第1の例）の一例を示す図である。

【図15B】図15Bは、検索FPデータ（第2の方式の第2の例）の一例を示す図である。

【図16】図16は、検索処理の処理フローを示す図である。

【図17】図17は、類似度算出処理の処理フローを示す図である。

【図18】図18は、第1の方式を採用した場合における数値のハッシュ値の比較について説明するための図である。

【図19】図19は、第2の方式を採用した場合における数値のハッシュ値の比較について説明するための図である。

30

【図20】図20は、類似度算出処理の処理フローを示す図である。

【図21】図21は、類似度算出処理の処理フローを示す図である。

【図22】図22は、出力例を示す図である。

【図23】図23は、FP生成処理の他の例を示す図である。

【図24】図24は、コンピュータの機能ブロック図である。

【発明を実施するための形態】

【0015】

本技術の実施の形態に係るシステムの構成例を図1に示す。図1に示すように、例えばインターネットであるネットワーク1には、登録装置3と、管理装置5と、検索装置7とが接続されている。登録装置3は、以下で述べる処理を行って機密データを秘匿化して、管理装置5に登録する装置であり、登録装置3の数に制限はない。また、検索装置7は、以下で述べる処理を行って検索条件に係る機密データを秘匿化して、秘匿化データと他の検索条件とを含む検索要求を管理装置5に送信し、管理装置5から検索結果を受信する装置であり、検索装置7の数に制限はない。登録装置3と検索装置7は、専用の装置であっても良いし、秘匿化データを登録する際には登録装置3として機能し、検索を行う際には検索装置7として機能する装置であっても良い。

40

【0016】

図2に、登録装置3の機能ブロック図を示す。登録装置3は、入力部31と、機密データ格納部32と、FP（Finger Print）生成部33と、FPルールデータ取得部34と、

50

F P ルールデータ格納部 3 5 と、F P データ格納部 3 6 と、送信部 3 7 とを有する。入力部 3 1 は、ユーザからの指示に応じて、機密データ格納部 3 2 に、管理装置 5 に格納すべきデータを格納したり、ユーザから機密データの選択指示を受け付け、当該選択指示を F P 生成部 3 3 に出力する。F P 生成部 3 3 は、F P ルールデータ格納部 3 5 に格納されている F P ルールデータに従って F P データを生成して、F P データ格納部 3 6 に格納する。なお、F P ルールデータ格納部 3 5 に F P ルールデータが格納されていない場合には、F P 生成部 3 3 は、F P ルールデータ取得部 3 4 に対して管理装置 5 から F P ルールデータを取得するように指示する。F P ルールデータ取得部 3 4 は、F P 生成部 3 3 からの指示に応じて、管理装置 5 から F P ルールデータを取得して、F P ルールデータ格納部 3 5 に格納する。送信部 3 7 は、F P データ格納部 3 6 に格納されている F P データを、管理装置 5 に送信する。

10

【 0 0 1 7 】

図 3 に、管理装置 5 の機能ブロック図を示す。管理装置 5 は、F P ルールデータ格納部 5 1 と、F P ルールデータ配布部 5 2 と、F P 登録部 5 3 と、データベース (D B) 5 4 と、検索処理部 5 5 と、検索要求受信部 5 6 と、検索結果送信部 5 7 とを有する。F P ルールデータ配布部 5 2 は、F P ルールデータ格納部 5 1 に格納されている F P ルールデータを、要求に応じて配信する。F P 登録部 5 3 は、登録装置 3 から F P データを受信し、D B 5 4 に格納する。検索要求受信部 5 6 は、検索装置 7 から、検索要求を受信し、受信した検索要求のデータを検索処理部 5 5 に出力する。検索結果送信部 5 7 は、検索処理部 5 5 から検索結果を受信すると、検索要求の送信元の検索装置 7 へ検索結果を送信する。検索処理部 5 5 は、F P ルールデータに従って、検索要求受信部 5 6 から受け取った検索要求に含まれる秘匿化データ及び検索条件などを用いた検索処理を実施して、検索結果を検索結果送信部 5 7 に出力する。

20

【 0 0 1 8 】

図 4 に、検索装置 7 の機能ブロック図を示す。検索装置 7 は、入力部 7 1 と、機密データ格納部 7 2 と、F P 生成部 7 3 と、F P ルールデータ取得部 7 4 と、F P ルールデータ格納部 7 5 と、検索条件データ格納部 7 6 と、F P データ格納部 7 7 と、検索要求部 7 8 と、出力部 7 9 とを有する。入力部 7 1 は、ユーザからの指示に応じて、機密データ格納部 7 2 に、検索のための機密データを格納したり、ユーザから機密データの選択指示を受け付け、当該選択指示を F P 生成部 7 3 に出力する。また、入力部 7 1 は、ユーザから検索条件のデータを受け付け、検索条件データ格納部 7 6 に格納する。

30

【 0 0 1 9 】

F P 生成部 7 3 は、F P ルールデータ格納部 7 5 に格納されている F P ルールデータに従って F P データ等を生成して、F P データ格納部 7 7 に格納する。なお、F P ルールデータ格納部 7 5 に F P ルールデータが格納されていない場合には、F P 生成部 7 3 は、F P ルールデータ取得部 7 4 に対して管理装置 5 から F P ルールデータを取得するように指示する。F P ルールデータ取得部 7 4 は、F P 生成部 7 3 からの指示に応じて、管理装置 5 から F P ルールデータを取得して、F P ルールデータ格納部 7 5 に格納する。検索要求部 7 8 は、F P データ格納部 7 7 に格納されている F P データ等と、検索条件データ格納部 7 6 に格納されている検索条件データとを読み出して検索要求を生成して、管理装置 5 に送信する。また、検索要求部 7 8 は、管理装置 5 から検索結果を受信すると、出力部 7 9 に出力して、例えば表示装置などに検索結果を表示する。

40

【 0 0 2 0 】

次に、図 1 乃至図 4 に示した装置の処理内容について説明する。まず、図 5 乃至図 1 0 を用いて、F P データの登録処理について説明する。まず、入力部 3 1 は、F P 生成対象の機密データの指定を受け付ける (図 5 : ステップ S 1) 。例えば、機密データ格納部 3 2 に格納されている機密データを列挙して選択させるようにしても良いし、指定された機密データを他のコンピュータなどから取得して機密データ格納部 3 2 に格納するようにしても良い。そして、入力部 3 1 は、指定された機密データを F P 生成部 3 3 に通知する。

【 0 0 2 1 】

50

F P生成部 33 は、F Pルールデータ格納部 35 にF Pルールデータが格納されているか確認する(ステップ S3)。F Pルールデータが格納されていない場合には(ステップ S5: No ルート)、F P生成部 33 は、F Pデータ取得部 34 に、F Pルールデータを取得させ、F Pルールデータ格納部 35 に格納させる(ステップ S7)。

【0022】

一方、F PルールデータがF Pルールデータ格納部 35 に格納されている場合(ステップ S5: Yes ルート)、又はステップ S7 の後に、F P生成部 33 は、F Pルールデータに従って、ユーザにより指定された機密データのF P生成処理を実施する(ステップ S9)。F P生成処理については、後に詳しく述べる。これによって、生成されたF Pデータは、F Pデータ格納部 36 に格納される。

10

【0023】

そして、送信部 37 は、F Pデータ格納部 36 に格納されているF Pデータを、管理装置 5 に送信する(ステップ S11)。これに対して、管理装置 5 のF P登録部 53 は、登録装置 3 からF Pデータを受信すると、当該受信したF Pデータ及び識別情報などを、DB 54 に格納する(ステップ S13)。識別情報は、例えば登録装置 3 の登録者IDと、登録日とを含み、F P登録部 53 が発行したF PIDをも含む。

【0024】

このような処理を繰り返すことで、DB 54 にF Pデータが蓄積されてゆく。

【0025】

次に、図 6 乃至図 9 を用いて、F P生成処理について説明する。F P生成部 33 は、指定された機密データに対して正規化処理を実施する(図 6: ステップ S21)。本実施の形態における機密データは、数値を含むテキストデータである。しかしながら、数値は、半角数字、全角数字、漢数字、アラビア数字などで表されている場合があり、さらに単位の違いも含まれる可能性がある。本実施の形態における正規化処理では、このような異なる表現を統一させる処理である。例えば、全角で「7000」を、半角の「7000」へ、「1万円」を半角数字の「10000」に変換する。この正規化処理についてはよく知られているので、これ以上述べない。

20

【0026】

その後、F P生成部 33 は、指定された機密データ中の数値及び特徴語を抽出し、例えばメインメモリなどの記憶装置に格納する(ステップ S23)。例えば、機密データのテキストを形態素解析により形態素に分解し、さらにその中から数値及び特徴語(例えば一般名詞、固有名詞など)を抽出する。

30

【0027】

例えば、図 7 に示すようなテキストを処理する場合を考える。この例では、「患者」「基本」「情報」「主訴」「朝」「体温」「38.5」「発熱」「症状」「検査」「心拍数」「測定」「結果」「85」「以上」「値」「血液検査」...「治療」「方針」などが抽出される。

【0028】

次に、F P生成部 33 は、抽出された数値のうち未処理の数値を1つ特定する(ステップ S25)。そして、F P生成部 33 は、F Pルールデータ格納部 35 に格納されているF Pルールデータに従って、特定された数値から、F Pのための数値を生成し、メインメモリなどの記憶装置に格納する(ステップ S27)。本実施の形態では、数値の近似を判断できるようにするために、単純に数値を秘匿化するのではなく、例えば2つの方式のいずれかで、特定された数値を展開する。

40

【0029】

第1の方式では、複数の有効桁数で、特定された数値を表すようにする。例えば、「38.5」であれば、有効桁数が1であれば「 3×10^1 」、有効桁数が2であれば「 3.8×10^1 」、有効桁数が3であれば「 3.82×10^1 」というように表現を変更する。使用すべき有効桁数についてのデータは、F Pルールデータに含まれている。これによって、近似判断の幅を表す数値を生成している。

50

【 0 0 3 0 】

第2の方式では、予め定められた数値の範囲のいずれに、特定された数値が属するかを判断し、特定された数値が属する範囲の上限値及び下限値を特定する。なお、補助データとして、下限値からの差及び上限値からの差をさらに算出する。例えば、10刻みで範囲が規定されている場合には、「38.2」の場合、30乃至40という範囲に属するので、上限値「40」及び下限値「30」が特定される。補助データは、下限値からの差「8.2」と上限値からの差「-1.2」が算出される。FPルールデータには、数値の範囲についての定義が含まれる。このようにして、近似判断の幅を表す数値とその補助数値とが生成される。

【 0 0 3 1 】

そして、FP生成部33は、生成されたFPのための数値における秘匿部分に対するハッシュ値を生成し、メインメモリなどの記憶装置に格納する(ステップS29)。ハッシュ値ではなく、暗号化であっても良い。鍵を用いる場合には、登録装置3及び検索装置7で共通の鍵を用いる。第1の方式の場合には、有効桁数が1乃至3であれば、「 3×10^1 」、「 3.8×10^1 」及び「 3.82×10^1 」のそれぞれについてハッシュ値を算出する。第2の方式の場合には、特定された数値が属する範囲の上限値及び下限値のそれぞれについてハッシュ値を算出する。補助データについてはハッシュ値を算出しない。

【 0 0 3 2 】

このように、FPデータの登録処理の場合には、複数の数値に対して複数のハッシュ値が算出される。単純に数値のハッシュ値を1つだけ算出するだけでは、数値が完全一致するか否かしか判断できない。しかしながら、第1の方式によれば、1桁一致、2桁一致、3桁一致といったように、有効桁数の範囲で一致不一致を判断できるため、近似する数値の有無を判断できる。第2の方式によれば、特定された数値が属する範囲の上限値又は下限値が一致する場合を特定でき、以下で述べるように実際に数値の差をも判断できるため、近似する数値の有無も判断できる。

【 0 0 3 3 】

FP生成部33は、特定された数値の周辺における特徴語を所定個数特定する(ステップS31)。所定個数は、例えばFPルールデータに規定されている。そして、FP生成部33は、特定された各特徴語についてハッシュ値を算出し、メインメモリなどの記憶装置に格納する(ステップS33)。

【 0 0 3 4 】

そして、FP生成部33は、数値のハッシュ値等(補助データがある場合には当該補助データ)と特徴語のハッシュ値とを含むデータブロックを、FPデータ格納部36に格納する(ステップS35)。

【 0 0 3 5 】

ここまで処理すると図8Aに示すようなデータブロックが、FPデータ格納部36に格納される。図8Aの例では、数値の周辺4個の特徴語についてハッシュ値を算出するようになっている。なお、図8Aは、第1の方式を採用した場合の例を示している。また、Hash(X)は、Xのハッシュ値を表す。一方、第2の方式を採用した場合には、図8Bに示すようなデータブロックが生成される。

【 0 0 3 6 】

その後、FP生成部33は、機密データから抽出された数値の中で未処理の数値が存在するか判断する(ステップS37)。未処理の数値が存在している場合には処理はステップS25に戻る。一方、未処理の数値が存在していない場合には呼び出し元の処理に戻る。図7の機密データを第1の方式で処理すると、図9Aに示すようなもう一つデータブロックが生成される。一方、第2の方式で処理すると、図9Bに示すようなもう一つのデータブロックが生成される。このように、FPデータは、1又は複数のデータブロックを含む。

【 0 0 3 7 】

一般的には、管理装置5のDB54は、例えば図10に示すようなデータが蓄積される

10

20

30

40

50

。図10の例では、FPIDと、登録者IDと、登録日と、FPデータとが登録されるようになっている。FPデータは、データブロックのIDであるブロック番号と、数値部分と、特徴語部分とを含む。各データブロックの数値部分には、複数の数値のハッシュ値（NUM（1，1），NUM（1，2）など）と、補助データがある場合には補助データ（AUX1など）とを含む。さらに特徴語部分には、複数の特徴語のハッシュ値（KW（1，1），KW（1，2）などM個の特徴語のハッシュ値）を含む。図10の例では、N個のデータブロックが含まれ、各データブロックの特徴語はM個である例を示している。

【0038】

次に、検索時に行われる処理について図11乃至図21を用いて説明する。まず、検索装置7の入力部71は、ユーザから検索に係る機密データの指定を受け付け、FP生成部73に機密データの指定を出力する（図11：ステップS41）。機密データ格納部72に格納されていない場合には、例えば他のコンピュータから、指定された機密データを取得して、FP生成部73に出力するようにしても良い。

10

【0039】

また、入力部71は、ユーザから検索条件の入力を受け付け、検索条件データ格納部76に格納する（ステップS43）。以下で具体的な検索処理において用いられるパラメータを、ユーザが指定する。例えば、類似度の閾値や結果の出力数などが指定される。どのパラメータを指定すべきかは、例えばFPルールデータに含まれる場合もある。また、FPデータを生成する上で用いられるパラメータについては、FP生成部73に出力される。

20

【0040】

そして、FP生成部73は、FPルールデータ格納部75に、FPルールデータが格納されているか判断する（ステップS45）。FPルールデータがFPルールデータ格納部75に格納されていない場合には（ステップS47：Noルート）、FP生成部73は、FPデータ取得部74に、FPルールデータを取得させ、FPルールデータ格納部75に格納させる（ステップS49）。

【0041】

一方、FPルールデータがFPルールデータ格納部75に格納されている場合（ステップS47：Yesルート）、又はステップS49の後に、FP生成部73は、FPルールデータに従って、ユーザにより指定された機密データの第2FP生成処理を実施する（ステップS51）。第2FP生成処理については、図12乃至図15Bを用いて説明する。

30

【0042】

FP生成部73は、指定された機密データに対して正規化処理を実施する（図12：ステップS71）。ステップS21と同様である。

【0043】

その後、FP生成部73は、指定された機密データ中の数値及び特徴語を抽出し、例えばメインメモリなどの記憶装置に格納する（ステップS73）。ステップS23と同様である。

【0044】

例えば、図13に示すようなテキストを処理する場合を考える。この例では、「患者」「基本」「情報」「主訴」「体温」「38」「測定」「発熱」「判定」「検査」「心拍数」「測定」「値」「80」「正常値」...「治療」「方針」などが抽出される。

40

【0045】

次に、FP生成部73は、抽出された数値のうち未処理の数値を1つ特定する（ステップS75）。そして、FP生成部73は、FPルールデータ格納部75に格納されているFPルールデータに従って、特定された数値から、FPのための1又は複数の数値を生成し、メインメモリなどの記憶装置に格納する（ステップS77）。本ステップについてもステップS27と基本的には同様である。

【0046】

但し、本実施の形態では、第1の方式を採用する場合には、有効桁数の指定が検索条件

50

に含まれる場合がある。その場合には、FPのための数値について、複数の数値を生成するのではなく、指定された有効桁数の数値を生成する。

【0047】

図13の例の場合、「38」については、指定された有効桁数が「2」であれば、「 3.8×10^1 」というような表現の数値が生成される。

【0048】

また、第2の方式を採用する場合には、上で述べた方法と同様の方法を採用しても良い。例えば、10刻みで数値の範囲が規定されている場合には、「38」の場合、30乃至40という範囲に属するので、上限値「40」及び下限値「30」が特定される。補助データは、下限値からの差「8」と上限値からの差「-2」が算出される。この場合、下限値及び上限値が代表値としてハッシュ値の算出対象数値として取り扱われる。

10

【0049】

但し、第2の方式の場合、検索条件として近似と判断する範囲を指定するため、この近似と判断する範囲に基づき、特定された数値を展開してもよい。例えば、プラスマイナス1の範囲が近似と判断する範囲として指定された場合、特定された数値が「38」であれば「37」から「39」までであれば近似していると判断される。従って、特定された数値が属する数値の範囲を超えて近似と判断されることがないので、所属範囲の下限値「30」を代表値として特定し、当該代表値からの差「8」が補助データとして特定される。一方、特定された数値が「41」である場合に、プラスマイナス3の範囲が近似と判断する範囲として指定されると、「38」から「44」までであれば近似していると判断される。従って、代表値としては、所属する範囲の1つ下の範囲の下限値「30」と、所属する範囲の下限値「40」を代表値として特定し、補助データとして、第1の下限値からの差「11」と第2の下限値からの差「1」が算出される。

20

【0050】

上で述べた例では、数値の範囲の刻みを超えて近似と判断される範囲が規定されないという前提があるが、このような前提が成り立たない場合には、以下のようにする。例えば、10刻みで範囲が規定されているが、特定された数値が「123」で、プラスマイナス15が近似と判断する範囲と指定された場合、「108」乃至「138」が近似と判断される。従って、「100」「110」「120」「130」を代表値として特定し、それぞれとの差を補助データとして生成する。

30

【0051】

そして、FP生成部73は、生成された1又は複数の数値における秘匿部分に対するハッシュ値を生成し、メインメモリなどの記憶装置に格納する(ステップS79)。ステップS29と同様である。第2の方式の場合には、補助データはハッシュ値を算出しない。

【0052】

このように、検索を行う場合には、特定された数値に対して、FPのための数値として1又は複数の数値が生成される。但し、FP登録時と同様にFPデータを生成しても良い。上で述べたようなオプションについては、FPルールデータに規定されているものとする。

【0053】

FP生成部73は、特定された数値の周辺における特徴語を所定個数特定する(ステップS81)。所定個数は、例えばFPルールデータに規定されている。そして、FP生成部73は、特定された各特徴語についてハッシュ値を算出し、メインメモリなどの記憶装置に格納する(ステップS83)。

40

【0054】

そして、FP生成部73は、数値のハッシュ値等(補助データがある場合には当該補助データ)と特徴語のハッシュ値とを含むデータブロックを、FPデータ格納部77に格納する(ステップS85)。ステップS35と同様である。

【0055】

その後、FP生成部73は、機密データから抽出された数値の中で未処理の数値が存在

50

するか判断する（ステップS 87）。未処理の数値が存在している場合には処理はステップS 75に戻る。一方、未処理の数値が存在していない場合には呼び出し元の処理に戻る。

【0056】

例えば、図13に示した機密データについて、単純な第1の方式を採用した場合には、図14Aに示すようなFPデータが生成される。また、第1の方式で有効桁数が「2」である場合には、例えば図14Bに示したようなFPデータが生成される。一方、単純な第2の方式によれば、図15Aに示したようなFPデータが生成される。さらに、検索条件で近似と判断される範囲がプラスマイナス3であれば、図15Bに示したようなFPデータが生成される。

10

【0057】

このようにすれば、完全一致だけではなく数値が近似しているか否かを判定できるようになる。

【0058】

図11の処理の説明に戻って、検索装置7の検索要求部78は、FPデータ格納部77に格納されているFPデータ（以下、区別するため検索FPデータと呼ぶ）と検索条件データ格納部76に格納されているデータとを含む検索要求を、管理装置5に送信する（ステップS 53）。

【0059】

管理装置5の検索要求受信部56は、検索装置7から、検索FPデータ及び検索条件を含む検索要求を受信すると（ステップS 55）、検索要求のデータを検索処理部55に出力する。検索処理部55は、検索要求のデータを受け取ると、検索処理を実施する（ステップS 57）。この検索処理については、図16乃至図21を用いて説明する。

20

【0060】

検索処理部55は、FPルールデータ格納部51からFPルールデータを読み出す（図16：ステップS 91）。そして、検索処理部55は、類似すると判定されたFPについての識別情報を格納する類似FP配列を初期化する（ステップS 93）。さらに、検索処理部55は、類似判定のための閾値Tを、FPルールデータ又は検索条件から設定する（ステップS 95）。閾値は固定のこともあり、その場合にはFPルールデータに含まれる。

30

【0061】

その後、検索処理部55は、DB54内の未処理のFPデータを特定する（ステップS 97）。そして、検索処理部55は、特定されたFPデータと検索FPデータとについて類似度算出処理を実施する（ステップS 99）。類似度算出処理については、図17乃至図21を用いて説明する。

【0062】

まず、検索処理部55は、数値の類似度に応じた特徴語の共通度合いの累計値を算出するための変数c1及び特定されたFPデータに含まれるデータブロック数をカウントするための変数c2を0に初期化する（図17：ステップS 111）。また、検索処理部55は、検索FPデータに含まれるデータブロックのうち未処理のデータブロックの数値データN1を特定する（ステップS 113）。ハッシュ値が複数ある場合、補助データがある場合も、それらを含めてN1として特定する。

40

【0063】

さらに、検索処理部55は、特定されたFPデータに含まれるデータブロックのうち未処理のデータブロックの数値データN2を特定する（ステップS 115）。ここでも、ハッシュ値が複数ある場合、補助データがある場合も、それらも含めてN2として特定する。

【0064】

そして、検索処理部55は、数値データN1と数値データN2とを比較して、数値類似度Simを設定する（ステップS 119）。本実施の形態では、上で述べた2つの方式が

50

存在する。最初に、単純な比較方式について説明する。

【0065】

第1の方式の場合、数値データN1には1又は複数のハッシュ値が含まれ、数値データN2には複数のハッシュ値が含まれる。例えば、数値データN1に複数のハッシュ値が含まれる例を図18に示す。図18の例では、数値データN1の元の数値は38.2で、有効桁数1乃至3の場合のハッシュ値が数値データN1に含まれる。一方、図18には、元の数値が38.2であるデータブロック(A)の数値データN2と、元の数値が38であるデータブロック(B)の数値データN2と、元の数値が39であるデータブロック(C)の数値データN2とが比較対象として示されている。このように複数のハッシュ値が数値データN1に含まれる場合には、いずれかのハッシュ値が、比較対象の数値データN2に含まれるハッシュ値と一致すれば、Simに1を設定し、いずれのハッシュ値も、比較対象の数値データN2に含まれるハッシュ値と一致しなければ、Simに0を設定する。図18の例では、データブロック(A)乃至(C)のいずれも有効桁数「1」について一致するので、Sim = 1と設定される。

10

【0066】

一方、数値データN1に、指定された有効桁数のハッシュ値が1つだけ含まれる場合には、その1つのハッシュ値に一致するか否かを判断する。例えば、数値データN1について有効桁数2である「 3.8×10^1 」のみが含まれる場合には、データブロック(A)及び(B)については有効桁数2についてのハッシュ値が一致するが、データブロック(C)については一致するハッシュ値がないと判断される。

20

【0067】

第2の方式の場合、数値データN1には1又は複数のハッシュ値及び対応する補助データとが含まれ、数値データN2には複数のハッシュ値及び対応する補助データが含まれる。図19に、数値データN1と数値データN2との比較例を模式的に示す。例えば、元の数値「38.2」の数値データN2には、Hash(30)及び補助データ「8.2」とHash(40)及び補助データ「-1.8」とが含まれている。これに対して、元の数値「39.1」の数値データN1には、Hash(30)及び補助データ「9.1」とHash(40)及び補助データ「-0.9」とが含まれる。なお、検索条件として近似と判断する範囲のデータが指定され、ここではプラスマイナス1が指定されているものとする。

30

【0068】

この場合、数値データN1に含まれるハッシュ値と、数値データN2に含まれるハッシュ値とを比較して一致するものがあるか判断する。図19の例ではhash(30)及びhash(40)のいずれも一致すると判断される。そして、hash(30)の場合には、数値データN2の補助データ「8.2」と数値データN1の補助データ「9.1」との差が、指定された範囲内であるか否かを判断する。この場合、 $|9.1 - 8.2| = 0.9$ であるから、指定された範囲内であるので、本実施の形態では、数値類似度Sim = 1に設定する。もし、補助データの差が、指定された範囲を超えている場合には、数値類似度Sim = 0に設定する。hash(40)については同じ値が得られるので、処理しなくとも良い。

40

【0069】

次に、数値の類似度合いに応じて数値類似度Simを0から1までの実数を設定する方式について説明する。第1の方式の場合には、図18に示すように、数値データN1に複数のハッシュ値が含まれ、数値データN2にも複数のハッシュ値が含まれる。従って、同一の有効桁数のハッシュ値同士を比較して、一致する回数をカウントする。例えばデータブロック(A)の場合、元の数値が一致するので、3回一致する。データブロック(B)の場合、有効桁数2まで一致するので、2回一致する。データブロック(C)については、有効桁数1まで一致するので、1回一致する。従って、データブロック(A)については、 $Sim = 3回 / 3 (= 有効桁数の種類数) = 1$ を設定し、データブロック(B)については、 $Sim = 2回 / 3 = 0.67$ を設定し、データブロック(C)については、Si

50

$m = 1 \text{ 回} / 3 = 0.33$ を設定する。

【0070】

一方、第2の方式の場合、上で述べたように補助データの差が算出されるので、(指定された範囲 - 補助データの差の絶対値) / (指定された範囲)で算出する。上で述べた例では、 $Sim = |1 - 0.9| / 1 = 0.1$ と算出される。

【0071】

その後、検索処理部55は、数値類似度 Sim が0を超えているか判断する(ステップS121)。数値類似度 Sim が0である場合には、端子Bを介して図21のステップS137に移行する。これは、数値類似度 Sim との乗算によってそのデータブロックについての類似度が決定されるので、数値類似度 $Sim = 0$ であれば、当該データブロックについて比較を行っても全体で0となってしまうためである。一方、数値類似度 $Sim > 0$ であれば、端子Aを介して図20のステップS123の処理に移行する。

10

【0072】

図20の処理の説明に移行して、検索処理部55は、数値データ $N1$ に対応するデータブロックに含まれる特徴語のうち未処理の特徴語のハッシュ値 $KW1$ を特定する(ステップS123)。また、検索処理部55は、数値データ $N2$ に対応するデータブロックに含まれる特徴語のうち未処理の特徴語のハッシュ値 $KW2$ を特定する(ステップS125)。そして、検索処理部55は、ハッシュ値 $KW1$ とハッシュ値 $KW2$ とを比較する(ステップS127)。

20

【0073】

なお、本実施の形態では同一のFPデータについて類似度を算出した場合には、1になることを前提としている。しかし、一般的には、数値について同一のハッシュ値が異なるデータブロックで出現する場合がある。この場合、異なる特徴語のハッシュ値が対応付けられている場合には特に問題ないが、同一の特徴語のハッシュ値が対応付けられている場合には同一のFPデータについて類似度を算出すると全体として類似度が1を超えてしまう。そこで、数値についてのハッシュ値と特徴語についてのハッシュ値との組み合わせが既に出現していたことが判明した場合には、その比較結果を類似度に反映しないようにする。

【0074】

従って、検索処理部55は、 $KW1 = KW2$ であって且つ数値データ $N1$ と $KW1$ の組み合わせが初出であるか判断する(ステップS129)。 $KW1$ と $KW2$ とが一致しない場合、又は数値データ $N1$ と $KW1$ の組み合わせが既出である場合には、ステップS133に移行する。

30

【0075】

一方、 $KW1 = KW1$ であって且つ数値データ $N1$ と $KW1$ の組み合わせが初出である場合、検索処理部55は、変数 $c1$ に数値類似度 Sim を加算して新たな変数 $c1$ の値として設定する(ステップS131)。数値類似度 Sim が0又は1の場合には、変数 $c1$ には、共通する特徴語の数が設定される。一方、数値類似度 Sim が0から1までの値で変化する場合には、変数 $c1$ には、データブロック毎に数値類似度 Sim で重み付けされた共通特徴語の数が累積される。

40

【0076】

そして、検索処理部55は、数値データ $N2$ に対応するデータブロックに未処理の特徴語のハッシュ値 $KW2$ が存在するか判断する(ステップS133)。未処理の特徴語のハッシュ値が存在する場合にはステップS125に戻る。一方、未処理の特徴語のハッシュ値が存在しない場合には、検索処理部55は、数値データ $N1$ に対応するデータブロックに未処理の特徴語が存在するか判断する(ステップS135)。未処理の特徴語のハッシュ値が存在する場合にはステップS123に戻る。一方、未処理の特徴語のハッシュ値が存在しない場合には、端子Bを介して図21のステップS137に移行する。

【0077】

図21の処理の説明に移行して、検索処理部55は、変数 $c2$ を1インクリメントする

50

(ステップ S 1 3 7)。検索処理部 5 5 は、特定された F P データに含まれるデータブロックに未処理のデータブロックがあるか判断する(ステップ S 1 3 9)。特定された F P データに未処理のデータブロックが存在している場合には、処理は端子 C を介して図 1 7 のステップ S 1 1 5 に戻る。一方、特定された F P データに未処理のデータブロックが存在しない場合には、検索処理部 5 5 は、検索 F P データに含まれるデータブロックに未処理のデータブロックが存在するか判断する(ステップ S 1 4 1)。検索 F P データに未処理のデータブロックが存在する場合には、処理は端子 D を介して図 1 7 のステップ S 1 1 3 に戻る。一方、検索 F P データに未処理のデータブロックが存在しない場合には、検索処理部 5 5 は、 $c 1 / (c 2 \times \text{ブロックサイズ})$ により類似度を算出し、F P データの識別情報に対応付けて例えばメインメモリなどの記憶装置に格納する(ステップ S 1 4 3)。ブロックサイズは、1 データブロックに含まれる特徴語の数である。そして呼び出し元の処理に戻る。

【0078】

このような処理を実施することで、数値をベースに特徴語も類似する機密データを秘匿性を保持しつつ検索することができる。数値についても近似しているか否かを秘匿化したままで判断できる。さらに検索 F P データについても秘匿化されており、管理装置 5 に対しても、どのような検索を行っているのかについて秘密が保持されている。

【0079】

なお、ステップ S 1 4 3 で計算される類似度については、特定された F P データにフォーカスし、そのデータに含まれるブロック数 $c 2$ を類似度計算式に入れた。そのほかに、検索 F P データに含まれるブロック数 N_Q を $c 2$ の代わりに使い、特徴語の共通度合い $c 1$ が検索 F P データのサイズ(ブロック数 $N_Q \times \text{ブロックサイズ}$)のどの程度の割合を占めるかを表す類似度も考えられる。その計算式は以下の式で表される。同様に、利用場面によっては、 $c 2$ と N_Q の大きい方 $\max(c 2, N_Q)$ 、または小さい方 $\min(c 2, N_Q)$ を $c 2$ の代わりに使うことも考えられる。

【0080】

【数 1】

$$\text{Similarity}(Q, D) = \sum_{i, j} \frac{\text{Sim}(\text{Num}_{Qi}, \text{Num}_{Dj}) \times (B_{Qi} \cap B_{Dj})}{\text{Block_size} \times N_Q}$$

【0081】

ここで Q が検索 F P データを表し、 D が比較対象の F P データを表す。そして、Block_size は、上で述べたブロックサイズであり、 N_Q は、検索 F P データのデータブロック数を表す。 Num_{Qi} は、検索 F P データにおける i 番目のデータブロックの数値データを表し、 Num_{Dj} は、比較対象の F P データにおける j 番目のデータブロックの数値データを表す。 $\text{Sim}(\text{Num}_{Qi}, \text{Num}_{Dj})$ は、検索 F P データにおける i 番目のデータブロックの数値データと、比較対象の F P データにおける j 番目のデータブロックの数値データとの類似度 Sim を表す。 B_{Qi} 、 B_{Dj} は、検索 F P データにおける i 番目のデータブロックに含まれる特徴語のハッシュ値と、比較対象の F P データにおける j 番目のデータブロックに含まれる特徴語のハッシュ値とで共通するハッシュ値の個数を表す。

【0082】

図 1 6 の処理の説明に戻って、検索処理部 5 5 は、算出した類似度が、検索条件で指定された閾値 T を超えているか判断する(ステップ S 1 0 1)。類似度が閾値 T を超えている場合には、検索処理部 5 5 は、特定された F P データの識別情報(図 1 0 における F P ID、登録者 ID 及び登録日など)を含む書誌データを、類似 F P 配列に追加する(ステップ S 1 0 3)。検索者の参照のため、類似度の数値自体を書誌データに含めるようにしても良い。一方、類似度が閾値 T 以下である場合には、ステップ S 1 0 5 に移行する。

【0083】

ステップ S 1 0 1 で類似度が閾値 T 以下であると判断された場合又はステップ S 1 0 3

10

20

30

40

50

の後に、検索処理部 55 は、DB 54 内に未処理の F P データが存在しているか判断する（ステップ S 105）。未処理の F P データが存在している場合には処理はステップ S 97 に戻る。一方、未処理の F P データが存在していない場合には、検索処理部 55 は、類似 F P 配列のデータを検索結果送信部 57 に出力する（ステップ S 107）。そして呼び出し元の処理に戻る。なお、登録者の詳細データについて追加した形で、検索結果送信部 57 に出力するようにしても良い。

【0084】

このようにして完全一致だけではなく類似する F P データを特定して、当該 F P データに関連するデータが抽出される。

【0085】

図 11 の処理の説明に戻って、検索結果送信部 57 は、検索処理部 55 から受け取った検索結果のデータを、検索要求の送信元である検索装置 7 に送信する（ステップ S 59）。検索装置 7 の検索要求部 78 は、検索結果を管理装置 5 から受信し、出力部 79 に出力する（ステップ S 61）。そして、出力部 79 は、検索結果を表示装置などに出力する。例えば、図 22 に示すようなデータが表示装置に表示される。図 22 の例では、F P I D と、登録者と、登録日と、類似度とが表示される。このように類似度が高い順にソートされた結果が提示されるようにしても良い。

【0086】

これによって、検索者は、類似する F P データの登録者を特定できるので、当該登録者に具体的な情報提供を依頼することができるようになる。

【0087】

例えば、診療データについて F P データを登録する場合には、診療データそのものを開示することがないので、プライバシー保護やセキュリティ保護の観点で問題が生じず、管理装置 5 への F P データ登録が促進される。一方、検索側でも患者のデータは秘匿化されたままであり、プライバシー保護やセキュリティ保護の観点で問題は無いので、利用の促進も図られる。そして、具体的に類似する症例の存在が確認できれば、別途問い合わせを行うことで、治療法などの情報を早期に取得でき、患者にも有効である。

【0088】

以上本技術の実施の形態を説明したが、本技術はこれに限定されるものではない。例えば、機密データから特徴語を特定するような処理を実施していたが、これに加えて、図 23 に示すような処理フローを実施しても良い。ステップ S 201 乃至 S 205 以外は、図 6 と同様である。F P 生成部 33 は、特定された特徴語の同義語を辞書から抽出し（ステップ S 201）、各特徴語及び各同義語についてハッシュ値を算出する（ステップ S 203）。又、F P 生成部 33 は、特徴語に加えて同義語のハッシュ値をも含むデータブロックを F P データ格納部 36 に格納する（ステップ S 205）。このようにして、同義語についてのハッシュ値をも F P データに含めるようにしても良い。

【0089】

さらに、上では閾値 T を検索条件に含める例を示したが、例えば類似度が高い順で上位指定個数の F P データを抽出するようにしても良い。

【0090】

さらに、上で示した機能ブロック図は一例であって、必ずしも実際のプログラムモジュール構成と一致しない場合もある。さらに、処理フローについても、処理結果が変わらない限り処理ステップの順番を入れ替えたり、並列実行するようにしても良い。

【0091】

また、F P ルールデータは、管理装置 5 以外で管理しても良い。

【0092】

なお、上で述べた登録装置 3、管理装置 5 及び検索装置 7 は、コンピュータ装置であって、図 24 に示すように、メモリ 2501 と CPU (Central Processing Unit) 2503 とハードディスク・ドライブ (HDD: Hard Disk Drive) 2505 と表示装置 2509 に接続される表示制御部 2507 とリムーバブル・ディスク 2511 用のドライブ装置

10

20

30

40

50

2513と入力装置2515とネットワークに接続するための通信制御部2517とがバス2519で接続されている。オペレーティング・システム（OS：Operating System）及び本実施例における処理を実施するためのアプリケーション・プログラムは、HDD2505に格納されており、CPU2503により実行される際にはHDD2505からメモリ2501に読み出される。CPU2503は、アプリケーション・プログラムの処理内容に応じて表示制御部2507、通信制御部2517、ドライブ装置2513を制御して、所定の動作を行わせる。また、処理途中のデータについては、主としてメモリ2501に格納されるが、HDD2505に格納されるようにしてもよい。本技術の実施例では、上で述べた処理を実施するためのアプリケーション・プログラムはコンピュータ読み取り可能なリムーバブル・ディスク2511に格納されて頒布され、ドライブ装置2513からHDD2505にインストールされる。インターネットなどのネットワーク及び通信制御部2517を経由して、HDD2505にインストールされる場合もある。このようなコンピュータ装置は、上で述べたCPU2503、メモリ2501などのハードウェアとOS及びアプリケーション・プログラムなどのプログラムとが有機的に協働することにより、上で述べたような各種機能を実現する。

10

【0093】

以上述べた本実施の形態をまとめると、以下のようになる。

【0094】

本実施の形態の第1の形態に係る情報処理方法は、（A）データ格納部に格納されており且つ第1の数値を含むテキストデータから、第1の数値及び当該第1の数値の周辺に存在する複数個の特徴語を抽出する処理と、（B）抽出された第1の数値から、当該第1の数値と近似するか否かを判断する上で基準となる1又は複数の第2の数値を生成する生成処理と、（C）1又は複数の第2の数値と複数個の特徴語との各々について秘匿化処理を行って秘匿化データを生成し、データ格納部に格納する処理とを含む。

20

【0095】

このように第2の数値を生成して当該第2の数値の秘匿化データを生成すれば、完全一致だけではなく近似する数値についても検出できるようになる。なお、第2の数値を1つだけ生成するのではなく複数生成すれば、より近似する数値を検出し易くなる。これはデータ登録時でもデータ検索時でも同様である。なお、第1の数値と類似するか否かを判断する上で基準となる数値は、近似判断の幅を表す数値とも言える。

30

【0096】

また、上で述べた生成処理が、抽出された第1の数値を異なる有効桁数で表した複数の第2の数値を生成する処理である場合もある。このようにすれば、有効桁数によって近似の精度を調整できる。

【0097】

さらに、上で述べた生成処理が、抽出された第1の数値を含む所定の数値範囲の上限値及び下限値である複数の第2の数値を特定する処理と、第1の数値と下限値との差と、第1の数値と上限値との差とを算出し、データ格納部に格納する処理とを含むようにしても良い。このようにすれば、検索の際に、元の数値との差を計算しやすくなる。

【0098】

さらに、上で述べた生成処理が、抽出された第1の数値と、数値を分類するための数値範囲の設定とから、第1の数値を代表する1又は複数の第2の数値を特定する処理と、1又は複数の第2の数値と、第1の数値との差を算出し、データ格納部に格納する処理とを含むようにしても良い。例えば、検索のためのデータを生成する際には、近似と判断する範囲なども加味して第2の数値を生成すれば、検索時に近似する数値についての秘匿化データを正確に特定できるようになる。

40

【0099】

また、上で述べた生成処理が、抽出された第1の数値を、指示された有効桁数で表した第2の数値を1つ生成する処理である場合もある。検索時にはこのように有効桁数を指定することで、所望の精度で近似を判断できるようになる。

50

【 0 1 0 0 】

さらに、本実施の形態の第 1 の形態に係る情報処理方法は、複数個の特徴語の同義語を抽出する処理と、同義語の秘匿化を行って秘匿化データを生成し、データ格納部に格納する処理とをさらに含むようにしても良い。これによれば、類似する秘匿化データを抽出し易くなる。

【 0 1 0 1 】

本実施の形態の第 2 の態様に係る情報処理方法は、(A) 第 1 の数値の第 1 の秘匿化データ値と複数個の第 1 の特徴語の第 2 の秘匿化データ値とを含む 1 又は複数の検索データブロックを含む検索要求を受信する処理と、(B) 複数の第 2 の数値の第 3 の秘匿化データ値と複数個の第 2 の特徴語の第 4 の秘匿化データ値とを含む 1 又は複数のデータブロックと識別情報とを含む案件データブロックを複数格納するデータ格納部に格納されている案件データブロックの各々について、第 1 の秘匿化データ値と第 3 の秘匿化データ値とから算出される、数値についての第 1 の類似度と、第 2 の秘匿化データ値と一致する第 4 の秘匿化データ値の個数とから、処理対象の案件データブロックに含まれるデータブロックと検索データブロックとの各組み合わせについての第 2 の類似度の合計値である第 3 の類似度を算出する算出処理と、(C) 第 3 の類似度が閾値を超えた案件データブロックの識別情報又は第 3 の類似度が上位所定数の案件データブロックの識別情報を、検索要求の送信元に送信する処理とを含む。

【 0 1 0 2 】

このようにすれば、データ格納部に格納されている案件データブロックも、検索要求に含まれるデータブロックについても秘匿された状態で、数値についての類似度も特定でき、全体としての類似度も算出できる。従って、より類似度の高い案件データブロックを特定できるようになる。

【 0 1 0 3 】

なお、上で述べた算出処理が、第 1 の秘匿化データ値に一致する第 3 の秘匿化データ値が存在する場合には第 1 の類似度を 1 に設定し、第 1 の秘匿化データ値に一致する第 3 の秘匿化データ値が存在しない場合には第 1 の類似度を 0 に設定する処理を含むようにしても良い。例えば元の値の近似判断を表す複数の第 2 の数値について第 3 の秘匿化データ値を用意しておけば、近似する数値の存在を検出しやすくなる。

【 0 1 0 4 】

また、上で述べた第 1 の秘匿化データ値が、第 1 の数値の元の数値についての代表値の秘匿化データ値である場合もある。そして、上で述べた検索要求には、第 1 の数値の元の数値についての代表値との差である第 1 の補助数値と、近似判定のための範囲のデータとをさらに含む場合もある。そして、複数の第 2 の数値が、元の数値が属する値域の下限値及び上限値であり、上で述べたデータブロックには、第 2 の数値の元の数値が属する値域の下限値と当該元の数値との差である第 2 の補助数値と当該元の数値と上記上限値との差である第 3 の補助数値とをさらに含むようにしてもよい。このような場合、上で述べた算出処理が、第 1 の秘匿化データ値に一致する第 3 の秘匿化データ値が存在する場合には、第 1 の秘匿化データ値についての第 1 の補助数値と、第 1 の秘匿化データ値に一致する第 3 の秘匿化データ値についての第 2 の補助数値又は第 3 の補助数値との差を算出する処理と、第 1 の補助数値と第 2 の補助数値又は第 3 の補助数値との差が、近似判定のための範囲内であれば、第 1 の類似度を 1 に設定し、第 1 の補助数値と第 2 の補助数値又は第 3 の補助数値との差が、近似判定のための範囲内でない場合には第 1 の類似度を 0 に設定する処理とを含むようにしても良い。

【 0 1 0 5 】

さらに、第 1 の数値の第 1 の秘匿化データ値が複数データブロックに含まれる場合には、上で述べた算出処理が、第 1 の秘匿化データ値に一致する第 3 の秘匿化データ値の個数に応じた類似度を第 1 の類似度に設定する処理を含むようにしても良い。このようにすれば、0 又は 1 だけではない第 1 の類似度を設定できるようになる。

【 0 1 0 6 】

また、第1の秘匿化データ値が、第1の数値の元の数値についての代表値の秘匿化データ値であり、検索要求には、第1の数値の元の数値についての代表値との差である第1の補助数値と、近似判定のための範囲のデータとをさらに含むようにしても良い。さらに、上で述べた複数の第2の数値が、元の数値が属する値域の下限値及び上限値であり、上で述べたデータブロックには、第2の数値の元の数値が属する値域の下限値と当該元の数値との差である第2の補助数値と当該元の数値と上記上限値との差である第3の補助数値とをさらに含むようにしても良い。そして、上で述べた算出処理が、第1の秘匿化データ値に一致する第3の秘匿化データ値が存在する場合には、第1の秘匿化データ値についての第1の補助数値と、第1の秘匿化データ値に一致する第3の秘匿化データ値についての第2の補助数値又は第3の補助数値との差を算出する処理と、第1の補助数値と第2の補助数値又は第3の補助数値と、近似判定のための範囲を表す数値と、の差に応じた類似度を第1の類似度に設定する処理とを含むようにしても良い。このようにすれば0又は1だけではない第1の類似度が設定できるようになる。

10

【0107】

なお、上で述べたような処理をコンピュータに実施させるためのプログラムを作成することができ、当該プログラムは、例えばフレキシブル・ディスク、CD-ROMなどの光ディスク、光磁気ディスク、半導体メモリ（例えばROM）、ハードディスク等のコンピュータ読み取り可能な記憶媒体又は記憶装置に格納される。

【0108】

以上の実施例を含む実施形態に関し、さらに以下の付記を開示する。

20

【0109】

(付記1)

データ格納部に格納されており且つ第1の数値を含むテキストデータから、第1の数値及び当該第1の数値の周辺に存在する複数個の特徴語を抽出する処理と、

抽出された前記第1の数値から、当該第1の数値と近似するか否かを判断する上で基準となる1又は複数の第2の数値を生成する生成処理と、

前記1又は複数の第2の数値と前記複数個の特徴語との各々について秘匿化処理を行って秘匿化データを生成し、前記データ格納部に格納する処理と、

を含む処理を、コンピュータが実行する情報処理方法。

30

【0110】

(付記2)

前記生成処理が、

抽出された前記第1の数値を異なる有効桁数で表した複数の第2の数値を生成する処理である付記1記載の情報処理方法。

【0111】

(付記3)

前記生成処理が、

抽出された前記第1の数値を含む所定の数値範囲の上限値及び下限値である複数の第2の数値を特定する処理と、

前記第1の数値と前記下限値との差と、前記第1の数値と前記上限値との差とを算出し、前記データ格納部に格納する処理と、

40

を含む付記1記載の情報処理方法。

【0112】

(付記4)

前記生成処理が、

抽出された前記第1の数値と、数値を分類するための数値範囲の設定とから、前記第1の数値を代表する1又は複数の第2の数値を特定する処理と、

前記1又は複数の第2の数値と、前記第1の数値との差を算出し、前記データ格納部に格納する処理と、

を含む付記1記載の情報処理方法。

50

【 0 1 1 3 】

(付記 5)

前記生成処理が、
抽出された前記第 1 の数値を、指示された有効桁数で表した第 2 の数値を 1 つ生成する
処理

である付記 1 記載の情報処理方法。

【 0 1 1 4 】

(付記 6)

前記複数個の特徴語の同義語を抽出する処理と、
前記同義語の秘匿化を行って秘匿化データを生成し、前記データ格納部に格納する処理
と、

を前記処理がさらに含む付記 1 乃至 5 のいずれか 1 つ記載の情報処理方法。

【 0 1 1 5 】

(付記 7)

第 1 の数値の第 1 の秘匿化データ値と複数個の第 1 の特徴語の第 2 の秘匿化データ値と
を含む 1 又は複数の検索データブロックを含む検索要求を受信する処理と、

複数の第 2 の数値の第 3 の秘匿化データ値と複数個の第 2 の特徴語の第 4 の秘匿化データ
値とを含む 1 又は複数のデータブロックと識別情報とを含む案件データブロックを複数
格納するデータ格納部に格納されている前記案件データブロックの各々について、前記第
1 の秘匿化データ値と前記第 3 の秘匿化データ値とから算出される、数値についての第 1
の類似度と、前記第 2 の秘匿化データ値と一致する前記第 4 の秘匿化データ値の個数と
から、処理対象の案件データブロックに含まれる前記データブロックと前記検索データブ
ロックとの各組み合わせについての第 2 の類似度の合計値である第 3 の類似度を算出する算
出処理と、

前記第 3 の類似度が閾値を超えた案件データブロックの識別情報又は前記第 3 の類似度
が上位所定数の案件データブロックの識別情報を、前記検索要求の送信元に送信する処理
と、

を含む処理を、コンピュータが実行する情報処理方法。

【 0 1 1 6 】

(付記 8)

前記算出処理が、
前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値が存在する場合には前
記第 1 の類似度を 1 に設定し、前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化デ
ータ値が存在しない場合には前記第 1 の類似度を 0 に設定する処理

を含む付記 7 記載の情報処理方法。

【 0 1 1 7 】

(付記 9)

前記第 1 の秘匿化データ値が、前記第 1 の数値の元の数値についての代表値の秘匿化デ
ータ値であり、

前記検索要求には、前記第 1 の数値の元の数値についての代表値との差である第 1 の補
助数値と、近似判定のための範囲のデータとをさらに含み、

前記複数の第 2 の数値が、元の数値が属する値域の下限値及び上限値であり、

前記データブロックには、前記第 2 の数値の元の数値が属する値域の下限値と当該元の
数値との差である第 2 の補助数値と当該元の数値と前記上限値との差である第 3 の補助
数値とをさらに含み、

前記算出処理が、

前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値が存在する場合には、
前記第 1 の秘匿化データ値についての前記第 1 の補助数値と、前記第 1 の秘匿化データ
値に一致する前記第 3 の秘匿化データ値についての前記第 2 の補助数値又は前記第 3 の補助
数値との差を算出する処理と、

10

20

30

40

50

前記第 1 の補助数値と前記第 2 の補助数値又は前記第 3 の補助数値との差が、前記近似判定のための範囲内であれば、前記第 1 の類似度を 1 に設定し、前記第 1 の補助数値と前記第 2 の補助数値又は前記第 3 の補助数値との差が、前記近似判定のための範囲内でない場合には前記第 1 の類似度を 0 に設定する処理と、

を含む付記 7 記載の情報処理方法。

【 0 1 1 8 】

(付記 1 0)

前記第 1 の数値の第 1 の秘匿化データ値が複数前記データブロックに含まれ、前記算出処理が、

前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値の個数に応じた類似度を前記第 1 の類似度に設定する処理

10

を含む付記 7 記載の情報処理方法。

【 0 1 1 9 】

(付記 1 1)

前記第 1 の秘匿化データ値が、前記第 1 の数値の元の数値についての代表値の秘匿化データ値であり、

前記検索要求には、前記第 1 の数値の元の数値についての代表値との差である第 1 の補助数値と、近似判定のための範囲のデータとをさらに含み、

前記複数の第 2 の数値が、元の数値が属する値域の下限値及び上限値であり、

前記データブロックには、前記第 2 の数値の元の数値が属する値域の下限値と当該元の数値との差である第 2 の補助数値と当該元の数値と前記上限値との差である第 3 の補助数値とをさらに含み、

20

前記算出処理が、

前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値が存在する場合には、前記第 1 の秘匿化データ値についての前記第 1 の補助数値と、前記第 1 の秘匿化データ値に一致する前記第 3 の秘匿化データ値についての前記第 2 の補助数値又は前記第 3 の補助数値との差を算出する処理と、

前記第 1 の補助数値と前記第 2 の補助数値又は前記第 3 の補助数値と、前記近似判定のための範囲を表す数値と、の差に応じた類似度を前記第 1 の類似度に設定する処理と、

を含む付記 7 記載の情報処理方法。

30

【 0 1 2 0 】

(付記 1 2)

データ格納部と、

前記データ格納部に格納されており且つ第 1 の数値を含むテキストデータから、第 1 の数値及び当該第 1 の数値の周辺に存在する複数の特徴語を抽出し、抽出された前記第 1 の数値から、当該第 1 の数値と近似するか否かを判断する上で基準となる 1 又は複数の第 2 の数値を生成し、前記 1 又は複数の第 2 の数値と前記複数の特徴語との各々について秘匿化処理を行って秘匿化データを生成し、第 2 のデータ格納部に格納する生成部と、

を有する情報処理装置。

【 0 1 2 1 】

40

(付記 1 3)

第 1 の数値の第 1 の秘匿化データ値と複数の第 1 の特徴語の第 2 の秘匿化データ値とを含む 1 又は複数の検索データブロックを含む検索要求を受信する受信部と、

複数の第 2 の数値の第 3 の秘匿化データ値と複数の第 2 の特徴語の第 4 の秘匿化データ値とを含む 1 又は複数のデータブロックと識別情報とを含む案件データブロックを複数格納するデータ格納部に格納されている前記案件データブロックの各々について、前記第 1 の秘匿化データ値と前記第 3 の秘匿化データ値とから算出される、数値についての第 1 の類似度と、前記第 2 の秘匿化データ値と一致する前記第 4 の秘匿化データ値の個数とから、処理対象の案件データブロックに含まれる前記データブロックと前記検索データブロックとの各組み合わせについての第 2 の類似度の合計値である第 3 の類似度を算出する検

50

索処理部と、

前記第3の類似度が閾値を超えた案件データブロックの識別情報又は前記第3の類似度が上位所定数の案件データブロックの識別情報を、前記検索要求の送信元に送信する送信部と、

を有する情報処理装置。

【符号の説明】

【0122】

3 登録装置

3 1 入力部

3 2 機密データ格納部

3 3 FP生成部

3 4 FPRルールデータ取得部

3 5 FPRルールデータ格納部

3 6 PFデータ格納部

3 7 送信部

5 管理装置

5 1 FPRルールデータ格納部

5 2 FPRルールデータ配布部

5 3 PF登録部

5 4 DB

5 5 検索処理部

5 6 検索要求受信部

5 7 検索結果送信部

7 検索装置

7 1 入力部

7 2 機密データ格納部

7 3 FP生成部

7 4 FPRルールデータ取得部

7 5 FPRルールデータ格納部

7 6 検索条件データ格納部

7 7 FPデータ格納部

7 8 検索要求部

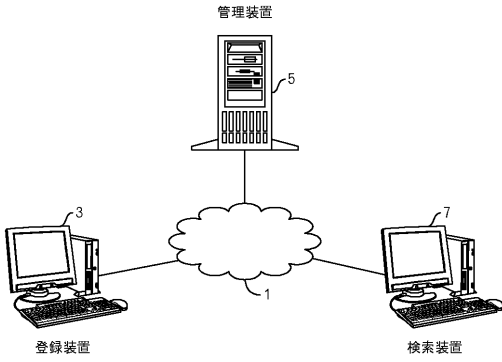
7 9 出力部

10

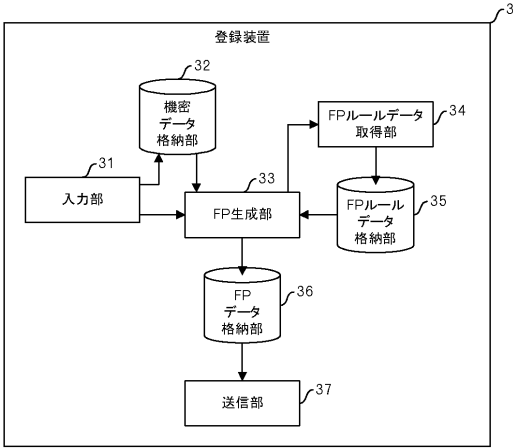
20

30

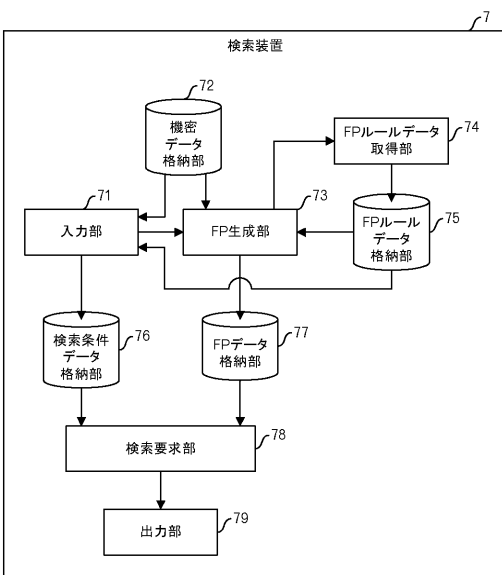
【 図 1 】



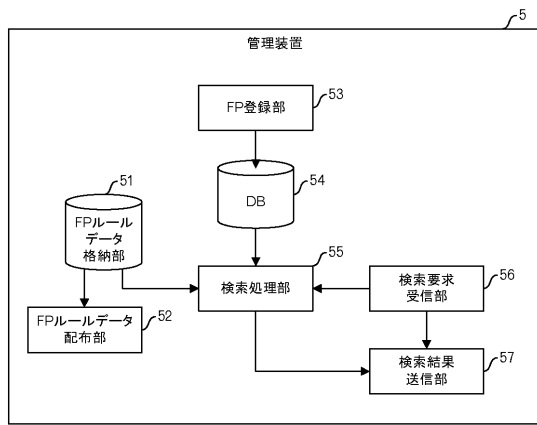
【 図 2 】



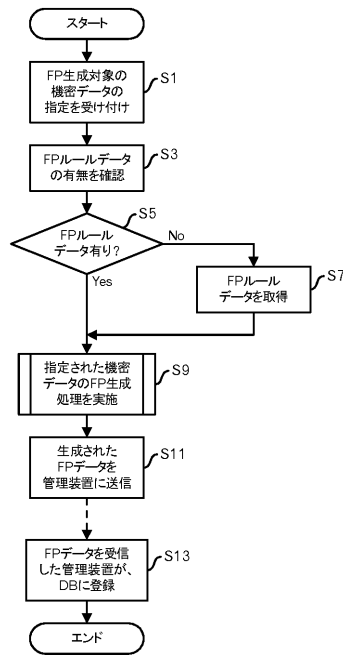
【 図 4 】



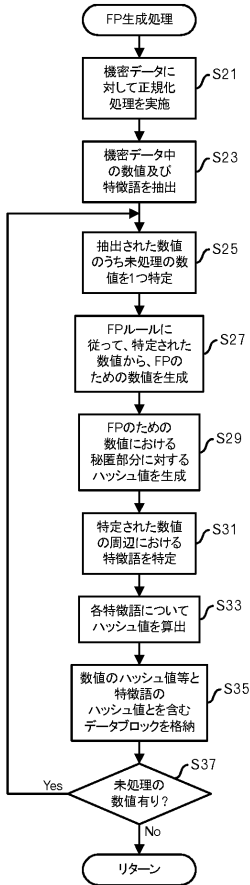
【 図 3 】



【 図 5 】



【 図 6 】



【 図 9 B 】

Hash(80),+5 Hash(90),-5	Hash(心拍数), Hash(測定), Hash(結果),Hash(異常)
----------------------------	---

【 図 7 】

【患者の基本情報】
 【主訴】
 朝起きたら、急に体温が38.5度まで上昇し、発熱したような症状がみられた
 【検査】
 心拍数を測定した結果、85という少し異常な値と考えられる。血液検査…
 【治療方針】
 …

【 図 8 A 】

Hash(3×10^1) Hash(3.8×10^1) Hash(3.85×10^1)	Hash(朝), Hash(体温),Hash(上昇), Hash(発熱)
--	---

【 図 8 B 】

Hash(30),+8.2 Hash(40),-1.8	Hash(朝), Hash(体温),Hash(上昇), Hash(発熱)
--------------------------------	---

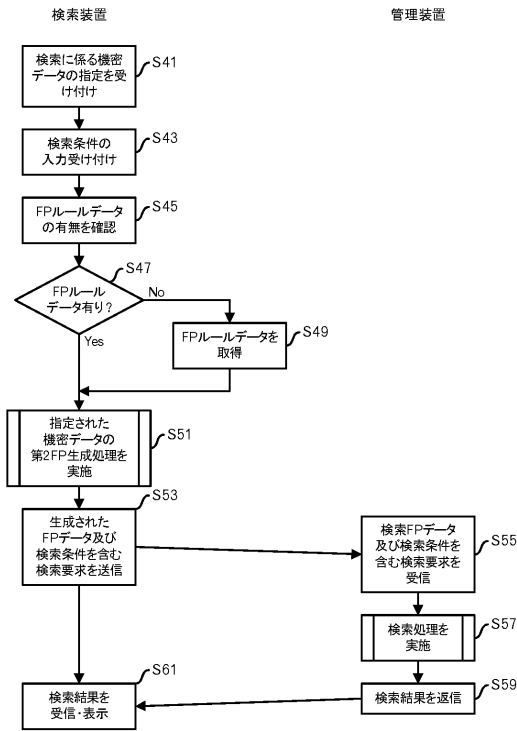
【 図 9 A 】

Hash(8×10^1) Hash(8.5×10^1) Hash(8.50×10^1)	Hash(心拍数), Hash(測定), Hash(結果),Hash(異常)
--	---

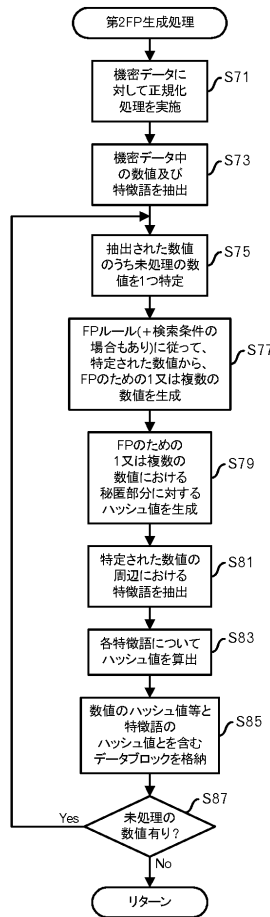
【 図 10 】

FPID	登録者ID	登録日	FPデータ		
			数値部分	特徴語部分	
A0001	A	2011/10/1	1	NUM(1,1), NUM(1,2),... (AUX1,...)	KW(1,1),KW(1,2),... KW(1,M)
			2	NUM(2,1), NUM(2,2),... (AUX1,...)	KW(2,1),KW(2,2),... KW(2,M)
			3	NUM(3,1), NUM(3,2),... (AUX1,...)	KW(3,1),KW(3,2),... KW(3,M)
			⋮	⋮	⋮
			N	NUM(N,1), NUM(N,2),... (AUX1,...)	KW(N,1),KW(N,2),... KW(N,M)

【 図 1 1 】



【 図 1 2 】



【 図 1 3 】

【患者の基本情報】
【主訴】
 体温が38度を測定し、発熱と判定
【検査】
 心拍数の測定値が80だと、正常値よりはやや速い…
【治療方針】
 …

【 図 1 5 A 】

Hash(30),+8 Hash(40),-2	Hash(体温), Hash(測定), Hash(発熱),Hash(判定)
Hash(80),0 Hash(90),-10	Hash(心拍数), Hash(測定), Hash(値),Hash(正常値)

【 図 1 5 B 】

Hash(30),+8 Hash(40),-2	Hash(体温), Hash(測定), Hash(発熱),Hash(判定)
Hash(70),+10 Hash(80),0	Hash(心拍数), Hash(測定), Hash(値),Hash(正常値)

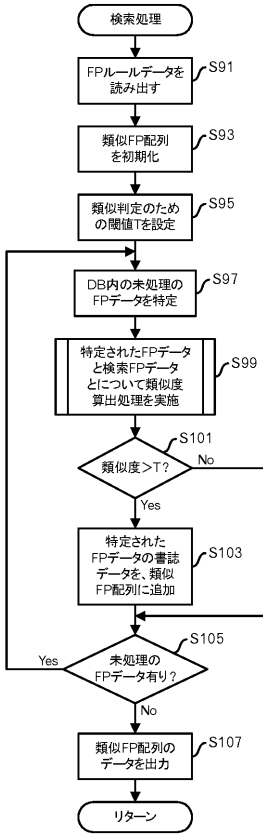
【 図 1 4 A 】

Hash(3 × 10 ¹) Hash(3.8 × 10 ¹) Hash(3.80 × 10 ¹)	Hash(体温), Hash(測定), Hash(発熱),Hash(判定)
Hash(8 × 10 ¹) Hash(8.0 × 10 ¹) Hash(8.00 × 10 ¹)	Hash(心拍数), Hash(測定), Hash(値),Hash(正常値)

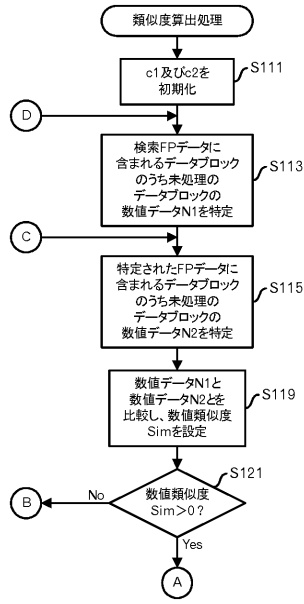
【 図 1 4 B 】

Hash(3.8 × 10 ¹)	Hash(体温), Hash(測定), Hash(発熱),Hash(判定)
Hash(8.0 × 10 ¹)	Hash(心拍数), Hash(測定), Hash(値),Hash(正常値)

【 図 1 6 】



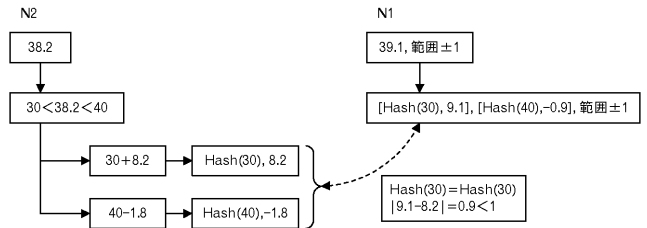
【 図 1 7 】



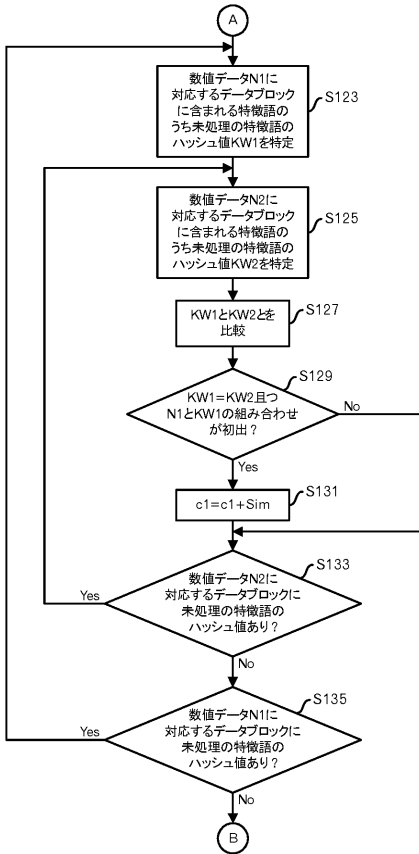
【 図 1 8 】

	有効桁数			比較結果		
	1	2	3	=	≠	≠
(A) N1	38.2	3×10^1	3.8×10^1	=	=	=
(A) N2	38.2	3×10^1	3.82×10^1	=	≠	≠
(B) N1	38	3×10^1	3.8×10^1	=	=	=
(B) N2	38	3×10^1	3.80×10^1	=	≠	≠
(C) N1	39	3×10^1	3.9×10^1	=	≠	≠
(C) N2	39	3×10^1	3.90×10^1	=	≠	≠

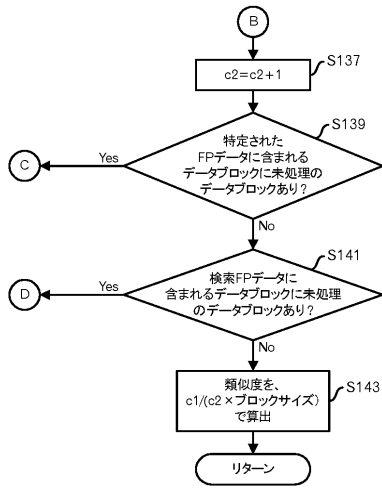
【 図 1 9 】



【図20】



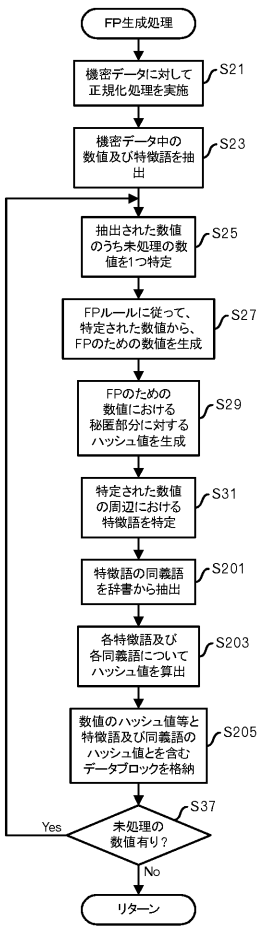
【図21】



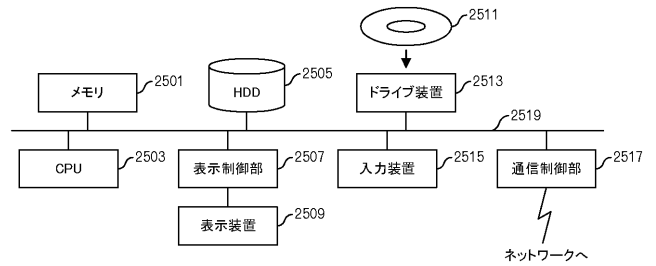
【図22】

FPID	登録者	登録日	類似度
A007	A	2011/1/1	0.8
A009	A	2011/2/1	0.7
B011	B	2011/5/1	0.6

【図23】



【図24】



フロントページの続き

(72)発明者 津田 宏

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内