

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5401037号
(P5401037)

(45) 発行日 平成26年1月29日 (2014. 1. 29)

(24) 登録日 平成25年11月1日 (2013. 11. 1)

(51) Int. Cl.

F I

G 0 6 Q 50/22 (2012. 01)

G 0 6 Q 50/22 1 0 0

G 0 6 F 17/30 (2006. 01)

G 0 6 F 17/30 1 7 0 Z

G 0 6 F 17/30 4 1 9 B

請求項の数 12 (全 34 頁)

(21) 出願番号 特願2007-523799 (P2007-523799)
 (86) (22) 出願日 平成17年7月28日 (2005. 7. 28)
 (65) 公表番号 特表2008-508612 (P2008-508612A)
 (43) 公表日 平成20年3月21日 (2008. 3. 21)
 (86) 国際出願番号 PCT/US2005/026750
 (87) 国際公開番号 W02006/015100
 (87) 国際公開日 平成18年2月9日 (2006. 2. 9)
 審査請求日 平成20年5月29日 (2008. 5. 29)
 審判番号 不服2011-27713 (P2011-27713/J1)
 審判請求日 平成23年12月22日 (2011. 12. 22)
 (31) 優先権主張番号 60/591, 734
 (32) 優先日 平成16年7月28日 (2004. 7. 28)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 503324302
 アイエムエス ヘルス インコーポレイテッド
 アメリカ合衆国 ペンシルヴェニア州 1
 9 4 6 2 プリマス ミーティング ウェ
 スト ジャーマンタウン パイク 6 6 0
 (74) 代理人 100147485
 弁理士 杉村 憲司
 (72) 発明者 ヒーサー ツレバ
 アメリカ合衆国 ペンシルヴェニア州 1
 9 4 4 4 ラファイエット ヒル アンド
 ラ グレン コート 5 1 3

最終頁に続く

(54) 【発明の名称】 複数のデータ・ソースから、暗号化及び非暗号化したデモグラフィック情報及びヘルスケア情報を用いて、非識別化した患者記録をリンクする方法。

(57) 【特許請求の範囲】

【請求項 1】

コンピュータを用いて、長期リンクのタグを、対象のデータ記録に割り当てる方法であって、前記対象のデータ記録のデータ属性は、英数字の識別コード属性、第三者属性及び／又は患者ヘルスケア・トランザクションに関するデモグラフィック属性のうち少なくとも幾つかを含む方法において、

(a) 前記コンピュータのプロセッサを用いて、マッチングレベルの階層構造を規定するステップであって、前記マッチングレベルの各々は、既知の長期リンクのタグと関連付けされた参照データ記録との比較のための指定組のデータ属性によって規定されており、前記マッチングレベルの階層構造は、各々の前記マッチングレベルにおける前記データ属性のマッチングの有効性に依拠して編成され、前記マッチングレベルの階層構造は、指定組の英数字の識別コード属性によって各々のマッチングレベルが規定される第1の組のマッチングレベル、及び、デモグラフィック属性を含む指定組の属性によって各々のマッチングレベルが規定される第2の組のマッチングレベルを含むステップと、

(b) 前記コンピュータのプロセッサを用いて、マッチングする参照データ記録を見つける試みにおいて、当該プロセッサを用いて、前記第1の組のマッチングレベルで、前記対象のデータ記録及び前記参照データ記録における前記指定組のデータ属性の値をレベル毎に比較するステップと、

(c) マッチングする参照データ記録が、前記ステップ(b)で見つけられない場合に、前記コンピュータのプロセッサを用いて、少なくとも関連するデモグラフィック属性が、前

記対象のデータ記録に与えられていることを確かめるステップと、

(d) 少なくとも前記関連するデモグラフィック属性が、前記対象のデータ記録に与えられている場合に、前記コンピュータのプロセッサを用いて、マッチングする参照データ記録を見つける試みにおいて、当該プロセッサを用いて、前記対象のデータ記録及び前記参照データ記録における前記指定組のデータ属性の値を、前記第2の組のマッチングレベル毎に比較するステップと、

(e) マッチングする参照データ記録が見つけれられる場合に、前記コンピュータのプロセッサを用いて、前記参照データ記録と関連付けられている長期リンクのタグを、前記対象のデータ記録に割り当てるステップと、
を含む方法。

10

【請求項2】

(f) 前記ステップ(b)～(e)で、前記参照データ記録が、前記対象のデータ記録と首尾よくマッチングしない場合に、前記コンピュータのプロセッサを用いて、新規に長期リンクのタグを生成し、前記タグを前記対象のデータ記録に割り当てるステップを更に含む、請求項1に記載の方法。

【請求項3】

前記第1の組のマッチングレベルにおける少なくとも1つのマッチングレベルが、供給者の患者ID、供給者ID、小売店ID、Rx番号、NCPPD資格付与者、NCPPDIDの属性及びそれらのいずれかの組み合わせからなるグループから選択した一組の属性によって規定されている、請求項1に記載の方法。

20

【請求項4】

前記ステップ(c)において、前記関連する属性が、DOB属性である、請求項1に記載の方法。

【請求項5】

前記第2の組のマッチングレベルにおける少なくとも1つのマッチングレベルが、保険カード番号、DOB、性別、患者のジップコード、姓、名、アドレス番号及びそれらのいずれかの組み合わせからなるグループから選択した一組のデモグラフィック属性によって規定されている、請求項1に記載の方法。

【請求項6】

(g) 前記コンピュータのプロセッサを用いて、前記対象のデータ記録、及びステップ(b)で見つけたマッチングする参照データ記録におけるデモグラフィック属性の値を比較して、前記2つの記録についてのマッチングを確認するステップを更に含む、

30

前記マッチングレベルは、当該マッチングレベルの各々を規定する前記指定組のデータ属性について経験的に定められたマッチングの有効性のランクに応じて、階層的な順序に配列されている、請求項1に記載の方法。

【請求項7】

前記ステップ(g)が、前記対象のデータ記録及び前記マッチングする参照データ記録におけるDOB及び性別の属性の値を比較して、前記2つの記録についてのマッチングを確認するステップを含む、請求項6に記載の方法。

【請求項8】

40

前記ステップ(g)が、前記対象のデータ記録及び前記マッチングする参照データ記録における姓及びジップコードの属性の値を比較して、前記2つの記録についてのマッチングを確認するステップを含む、請求項6に記載の方法。

【請求項9】

前記マッチングレベルが、当該マッチングレベルの各々を規定する前記指定組のデータ属性について経験的に定められたマッチングの有効性のランクに応じて、階層的な順序に配列されている、請求項1に記載の方法。

【請求項10】

コンピュータによって実行される命令を含むコンピュータ読取り可能なメディアであって、前記コンピュータは、前記命令を実行することによって、

50

前記コンピュータのプロセッサを用いて、長期リンクのタグを、対象のデータ記録に割り当てる方法を実行し、前記対象のデータ記録のデータ属性は、英数字の識別コード属性、第三者属性及び／又は患者ヘルスケア・トランザクションに関するデモグラフィック属性のうち少なくとも幾つかを含み、前記方法は、

(a) 前記コンピュータのプロセッサを用いて、マッチングレベルの階層構造を規定するステップであって、前記マッチングレベルの各々は、既知の長期リンクのタグと関連付けされた参照データ記録との比較のための指定組のデータ属性によって規定されており、前記マッチングレベルの階層構造は、各々の前記マッチングレベル内の前記データ属性のマッチングの有効性に依拠して編成され、前記マッチングレベルの階層構造は、指定組の英数字の識別コード属性によって各々のマッチングレベルが規定される第1の組のマッチングレベル、及び、デモグラフィック属性を含む指定組の属性によって各々のマッチングレベルが規定される第2の組のマッチングレベルを含むステップと、

10

(b) 前記コンピュータのプロセッサを用いて、マッチングする参照データ記録を見つける試みにおいて、当該プロセッサを用いて、前記第1の組のマッチングレベルで、前記対象のデータ記録及び前記参照データ記録における前記指定組のデータ属性の値をレベル毎に比較するステップと、

(c) マッチングする参照データ記録が、前記ステップ(b)で見つけられない場合に、前記コンピュータのプロセッサを用いて、少なくとも関連するデモグラフィック属性が、前記対象のデータ記録に与えられていることを確かめるステップと、

(d) 少なくとも前記関連するデモグラフィック属性が、前記対象のデータ記録に与えられている場合に、前記コンピュータのプロセッサを用いて、マッチングする参照データ記録を見つける試みにおいて、前記対象のデータ記録及び前記参照データ記録における前記指定組のデータ属性の値を、前記第2の組のマッチングレベル毎に比較するステップと、

20

(e) マッチングする参照データ記録が見つけれられる場合に、前記コンピュータのプロセッサを用いて、前記参照データ記録と関連付けられている長期リンクのタグを、前記対象のデータ記録に割り当てるステップと、
を含む、コンピュータ読取り可能なメディア。

【請求項11】

コンピュータ読取り可能なメディア上に符号化された長期リンクのタグを割り当てられた、少なくとも1つの対象のデータ記録を含む長期のデータベースであって、

30

前記少なくとも1つの対象のデータ記録は、英数字の識別コード属性、第三者属性及び／又はデモグラフィック属性のうち少なくとも幾つかを含み、

マッチングレベルの階層構造が規定され、前記マッチングレベルの各々は、既知の長期リンクのタグと関連付けされた参照データ記録との比較のための指定組のデータ属性によって規定されており、前記マッチングレベルの階層構造は、各々の前記マッチングレベルにおける前記データ属性のマッチングの有効性に依拠して編成され、前記マッチングレベルの階層構造は、指定組の英数字の識別コード属性によって各々のマッチングレベルが規定される第1の組のマッチングレベル、及び、デモグラフィック属性を含む指定組の属性によって各々のマッチングレベルが規定される第2の組のマッチングレベルを含み、

前記長期のデータベースにリンクされる新規の対象のデータ記録の各々は、まず、前記第1の組のマッチングレベルにおけるレベル毎に、既存の参照データ記録と比較され、マッチングする参照データ記録が見つけれられる場合に、前記参照データ記録と関連付けられている長期リンクのタグを割り当てられ、

40

前記第1の組のマッチングレベルにおいて、マッチングする参照データ記録が見つけれられず、かつ、少なくとも関連するデモグラフィック属性が、前記新規の対象のデータ記録に与えられている場合に、前記新規の対象のデータ記録の各々は、前記第2の組のマッチングレベルにおけるレベル毎に、既存の参照データ記録と比較され、マッチングする参照データ記録が見つけれられる場合に、前記参照データ記録と関連付けられている長期リンクのタグを割り当てられ、

前記第2の組のマッチングレベルにおいて、マッチングする参照データ記録が見つけれ

50

れないか、関連するデモグラフィック属性が、前記新規の対象のデータ記録に与えられていない場合に、前記新規の対象のデータ記録の各々は、新規の長期リンクのタグを割り当てられる、長期のデータベース。

【請求項 12】

前記新規の対象のデータ記録が、指定組のデモグラフィック属性の比較において前記第2の組のマッチングレベルのみで前記参照データ記録と首尾よくマッチングし、前記首尾よくマッチングした前記参照データ記録と関連付けられている前記長期リンクのタグを割り当てられる、請求項 11 に記載の長期のデータベース。

【発明の詳細な説明】

【技術分野】

10

【0001】

この出願は、2004年7月28日に出願された米国特許仮出願第60/591,734号明細書、及び関連する米国特許出願第11/122,589号明細書、米国特許出願第11/122,581号明細書、米国特許出願第11/122,564号明細書、及び、米国特許出願第11/122,565号明細書（全て2005年5月5日に提出された出願）の利益を主張するものであり、これら全ての出願の内容は、本出願の明細書の内容に取り込まれる。

【0002】

本発明は、個々において、個人の健康情報又はデータについての管理に関する。特に、本発明は、個人のプライバシーを保持しながら、長期のデータベースにおける、そのようなデータのアセンブリ及び使用に関する。

20

【背景技術】

【0003】

患者の健康記録の電子データベース化は、商業利用及び非商業的利用の双方に役立つものである。例えば、長期（寿命期間）の患者記録データベースは、疫学的に、又は、時間傾向、因果関係、又は、ある母集団における健康事象の発生率の分析として他の母集団ベースの研究に用いられる。長期のデータベースでアセンブリされた患者記録は、複数のソースから、様々なフォーマットで収集されることが考えられる。患者健康記録の明白なソースは、現代においては健康保険業界であり、その業界は、医療サービスプロバイダに対して保険金を与えるために、電子通信化した患者のトランザクション記録を広く頼りにしている。医療サービスプロバイダ（例えば、薬局、病院または診療所）、又は、医療サービスプロバイダのエージェント（例えば、データクリアリングハウス、プロセッサ又はベンダー）は、個別に特定された患者トランザクション記録を、補償金のために保険業界に供給する。患者トランザクション記録は、個人情報データのフィールド、即ち属性に加えて、例えば、診断、処方箋、治療内容又は治療結果などの他の情報を含むことができる。複数のソースから取得したそのような情報は、長期の研究に対して貴重となる場合がある。しかしながら、個人のプライバシーを保護するために、長期のデータベース装置に統合される患者記録を、“匿名化”、即ち“非識別化”することが重要である。

30

【0004】

データ供給者、即ちデータ・ソースは、転送前に患者プライバシーを保護するために、患者トランザクション記録の個人情報データのフィールド、即ち属性（例えば、名前、社会保障番号、住所、ジップコードなど）を除去又は暗号化できる。患者プライバシーを保護するために、ある個人情報データのフィールドの暗号化又は標準化は、現在、法令と政令によって規制されている。個人の民事的権利に対する関心事は、電子取引における、個人情報の収集についての政府規制と個人の健康データの使用に向けられている。例えば、「医療保険の相互運用性と説明責任に関する法律（1996年）（HIPAA法）」の下で発行された規制は、個人的な健康情報のセキュリティ及び秘匿性をセーフガードするように、緻密な規則を含む。HIPAA規制は、健康保険プラン、ヘルスケア・クリアリングハウス、及び、ある財政用及び管理用のトランザクション（例えば、登録、支払い、及び適正検査）を電子的に行うヘルスケア・プロバイダなどのエンティティをカバーしている。

40

50

(例えば、<http://www.hhs.gov/ocr/hipaa> 参照)。2004年7月15日に出願された、米国特許出願第10/892,021号明細書は、本出願の明細書の内容に取り込まれるものであるが、政令のHIPAA規制又は他の組のプライバシー規則に従って標準化されたフォーマットで個人的な健康情報を収集及び使用するシステム及び方法について、説明されている。

【0005】

患者のプライバシーの侵害についてのリスクを更に最小化するために、長期のデータベースを構成するように用いられる患者記録から、全ての患者識別情報を分離又は除去することが望ましいとも考えられる。しかしながら、患者識別情報をデータ記録から分離し、患者識別情報を完全に“匿名化”することは、格納されたデータ記録又はフィールドが、患者毎に必ず更新される長期のデータベースの構成に対して、互換性がなくなる場合がある。

10

【0006】

2005年5月5日に出願された、共通して所有される米国特許出願第11/122,589号明細書の内容は、本明細書の内容に取り込まれ、複数のソースから取得した患者データ記録が、患者プライバシーの侵害の如何なるリスクも生じさせることなく、個々の患者毎に長期のデータベースへと統合されることを可能とする解決策について開示されている。

【0007】

米国特許出願第11/122,589号明細書に開示される解決策は、複数の暗号鍵を用いて2ステップの暗号化手順を用いる。その手順は、データ・ソース、即ちデータ供給者(“DS”)、長期のデータベース装置(“LDF”)、及び、第三者実装パートナー(“IP”)を包含する。IPは、暗号鍵管理人としても機能し、装置コンポーネントサイトで動作することができる。第1のステップで、各DSは、患者記録において選択したデータ・フィールド(例えば、患者識別属性及び/又は他の標準的な属性データ・フィールド)を暗号化し、患者記録を第1の“匿名化した”フォーマットに変換する。各DSは、2つの鍵(即ち、ベンダー特定鍵と、特定のLDFに関連付けられた共通の長期鍵)を用いて、選択したデータ・フィールドを二重暗号化する。二重暗号化したデータ記録は、装置コンポーネントサイトに送信される。ここで、IPは、更に、二重暗号化したデータ記録を処理できる。データ記録は、第2の匿名化したフォーマットへと処理され、第2の匿名化したフォーマットは、データ記録が元の暗号化していない患者の識別情報を復元することなく、個々の患者毎にリンクされることを可能とするように構成される。このため、DSから受信した患者記録の二重暗号化したデータ・フィールドは、特定のベンダー鍵を用いて部分的に解読される(この際に、二重暗号化したデータ・フィールドは、共通の長期鍵暗号化をまだ保持している)。第3の鍵(例えば、トークン・ベースの鍵)は、長期のデータベースにおいて用いられる、新たな単一の(共通の長期鍵で)暗号化したデータ・フィールド、即ち属性を更に作成するように用いることができる。長期のデータベースに対して内部的となる長期の識別子(ID)又はダミーラベルは、データ記録をタグ付けするように用いることができ、これにより、データ記録は、元の暗号化していない患者の識別情報の知識なく、長期のデータベースにおいて個々のID毎にマッチングさせ、リンクさせることができる。(図1参照)

20

30

40

【0008】

マッチングアルゴリズムについて開示されている2005年5月5日に出願された、共通に所有される米国特許出願第11/122,564号明細書は、本明細書の内容に取り込まれ、本マッチングアルゴリズムは、新規又は前に規定したIDが一組の暗号化したデータ属性に割り当てられるべきであるか否かを決定するように用いることができる。一旦、新規又は前に規定したIDが割り当てられると、そのIDは、詳細なトランザクション情報を含んでいる完全なデータ記録にタグ付けし直してリンクさせるように用いることができる。本マッチングアルゴリズムは、患者データ記録を参照データ記録とマッチングさせることにより、長期リンクのタグを非識別化した患者データ記録に割り当てる。非識別

50

化した患者データ記録は、暗号化及び非暗号化したデータ属性の双方を含むことができる。データ属性の種々に組み合わせ可能なサブセットは、階層構造のレベルでカテゴリ化することができる。データ・フィールド値のサブセットは、1回当たり1レベルずつ参照データ記録と比較される。データ・フィールド値のサブセットについての首尾よい比較、即ちマッチングに基づいて、マッチングした参照データ記録と関連付けられる、長期リンクのタグは、割り当てられる非識別化したデータ記録に割り当てられる。マッチングが見つからない場合、新規の長期リンクのタグを作成し、非識別化したデータ記録に割り当てる。新規のタグ、及び対応するデータ記録の各属性は、更なるマッチング動作のために参照データ記録に加えられる。

【発明の開示】

10

【発明が解決しようとする課題】

【0009】

様々なデータ・ソースから生成又は受信した患者データをマッチングする処理に対して、更なる考慮がここに与えられる。例えば、患者のデータ記録のリンク化が、異なる薬局に渡って同一の患者にサービスすることを可能とするマッチング処理及びロジックに対して注意を向けている。

【課題を解決するための手段】

【0010】

本発明は、マルチソースの患者ヘルスケア・トランザクション記録から長期のデータベースをアセンブリするマッチング処理及びマッチングアルゴリズムを提供する。患者ヘルスケア・トランザクション記録は、患者の識別性又は個人情報について、無許可又は不注意の開示を防ぐために、非識別化される。非識別化した患者ヘルスケア・トランザクションデータ記録は、1つ以上のデータ属性を含むことができ、例えば、英数字の識別コード属性、第三者属性、及び／又は、患者ヘルスケア・トランザクションに関するデモグラフィック属性などを含むことができる。

20

【0011】

マッチング処理及びマッチングアルゴリズムは、入ってくるデータ記録に与えられている選択組のデータ属性（デモグラフィック属性の組を含むことができる。）を、既存の参照データ記録における対応する組のデータ記録と、逐次的に比較するように構成している。マッチングする参照データ記録が見つけれられる場合には、マッチングする参照データ記録と関連付けられている長期リンクのタグが、入ってくるデータ記録に割り当てられる。

30

【0012】

マッチング処理では、入ってくる非識別化した患者記録（対象のデータ記録）を、階層構造のマッチングレベルで参照データ記録に対してマッチングさせる試みが為される。各マッチングレベルでは、指定組のデータ属性が評価される。各マッチングレベルは、対象のマッチング記録に含めることができる英数字の識別コード（例えば、供給者の患者ID、供給者ID、小売店ID及びRx番号）の比較を含む。対象のデータ記録が、早めのマッチングレベルで、参照データ記録に対してマッチングしない場合には、後のマッチングレベルは、対象のマッチング記録に含めることができる1つ以上のデモグラフィック属性（例えば、保険カード番号、DOB、性別、患者ジップ（Zip）、姓、名、アドレス番号など）の比較を含むことができる。

40

【0013】

階層構造のマッチングレベル及び比較のために指定される関連組の属性は、特定のデータ属性又はデータ属性の組について、マッチングの有効性を考慮して選択される。特定のデータ属性又はデータ属性の組についてのマッチングの有効性は、テスト的なデータ記録を用いた、試し動作（trial run）で経験的に定めることもできる。

【0014】

好適なコンピュータプログラムで実現可能なマッチング処理例は、階層構造のレベル（例えば1～13）で行なわれる。これらのうち11個のレベル（例えば、レベル1～4及び6～12）で、本処理は、データ属性値を比較することにより、非識別化した患者記録

50

を参照データ記録に対してマッチングを試みる。マッチングが見つければ直ちに、マッチングした記録の長期の患者IDを入れてくる記録に割り当て、マッチング処理を継続させない。中間レベル（例えば、レベル5）では、マッチング処理は、入ってくるデータ記録を評価して、ほぼマッチングするレベルでマッチング成功の見込み度を示すことができる、関連するデモグラフィックデータ属性（例えば、DOB）の存在を確かめることができる。そのような関連するデモグラフィック属性が存在する場合、残りのマッチングレベルで、マッチング処理を継続させることができる。入ってくるデータ記録が、如何なる既存の参照データ記録ともマッチングしない場合には、入ってくるデータ記録に新たに生成した内部的な識別子を割り当てることができる（例えば、レベル13で）。

【0015】

本発明のマッチング処理は、ヘルスケア又は製薬の業界で用いられる長期の患者データベース（例えば、LRxデータベース）の情報量を豊かにできる。例えば、2つの薬局間で流通する、非識別化した患者についての非識別化したRx記録は、同一のIDに対して認識又はリンクさせることができる。以前、これらのRx記録は、二人の異なる者に対してリンクされていたはずである。ここで、LRxデータベースは、ビジネスのソース（例えば、新スタート、再スタート、継続患者など）、治療パターン（例えば、単剤療法（mono therapy）、併用療法（combination therapy）、及び追加併用療法（add-on therapy））、及びセラピー・進行状況（例えば、第1、第2及び第3ライン、療法の長さ、過去の療法（lapsed therapy）／中止した療法の使用、残留性（persistency）及びコンプライアンス（compliance）、タイトレーション（titration）及び薬剤投与（dosing）、及び一般的な患者の特徴）など患者中心の基準を含むように好適にアップグレードできる。

【0016】

本発明の更なる特徴、その本質、及び様々な利点は、以下の詳細な説明及び図面が明らかとなり、ここに、同様な参照番号は、同様の要素を示している。

【発明を実施するための最良の形態】

【0017】

内部的な長期リンクの識別子又はタグを、非識別化した患者トランザクションデータ記録に割り当てるマッチングアルゴリズムを提供する。割り当てられた長期リンクの識別子でタグ付けしたデータ記録は、識別子毎に容易にリンクさせることができ、個々の患者を識別し得る個人情報にアクセスすることなく、長期のデータベースをアセンブリすることができる。マッチングアルゴリズムは、新規又は前に規定したIDを一組の暗号化したデータ属性に割り当てべきか否かを決定するように用いることができる。一旦、新規又は前に規定したIDが割り当てられると、そのIDは、詳細なトランザクション情報を含む完全なデータ記録にタグ付けし直してリンクさせるように用いることができる。

【0018】

長期のデータベースのアセンブリに関して、まず、患者トランザクションデータ記録は、データ記録のデータ・フィールドが標準化した共通のフォーマットとなるように処理され、次に暗号化される。データ記録は、選択組のデータ属性に対応する少なくとも1つ以上のデータ・フィールドを含む。選択組のデータ属性は、暗号化されていないときに患者識別となるトランザクション属性、更に患者識別とならない他のトランザクション属性を含むことができる。本発明のマッチングアルゴリズムは、データ記録内の暗号化した属性の値を評価し、それに応じて内部的な長期リンクの識別子をデータ記録に割り当てる。そのような評価には、好適な長期リンクの識別子を割り当てるために、繰り返し評価、参照比較、確率論的又は他の統計学的な技法を含めることができる。評価される選択組のデータ属性は、適切な長期リンクの識別子をデータ記録に割り当てる際に生じる誤差を低減する観点から選択される。

【0019】

本発明のマッチングアルゴリズムを（本明細書の内容に取り込まれる米国特許出願第11／122,589号明細書に説明される）有効な解決策の内容に関して、以下説明する。本発明のマッチングアルゴリズムは、マルチソース化した患者データ記録を、個々の患

者ごとに、患者のプライバシーの侵害のリスクを生じることなく、長期のデータベースに統合する。尚、その解決策は、図示のためにのみに参照されるべきものであり、本発明のマッチングアルゴリズムは、非識別化したデータ記録を長期のデータベースに統合する他の解決策のアプリケーションを容易に見出すことができるものであることは明らかである。

【0020】

以下に説明する本発明について、完全に理解するために、米国特許出願第11/122,589号明細書に説明されている解決策の詳細な説明を、以下に説明する。米国特許出願第11/122,589号明細書で表される図1は、マルチソース化した患者データ記録から長期のデータベースをアセンブリする一例の解決策100のシステムコンポーネント及び処理を示す。複数の暗号鍵を用いる2ステップの暗号化手順が、患者データ記録を非識別化するために用いられる。解決策500は、データ・ソース、即ちデータ供給者（“DS”）、長期のデータベース装置（“LDF”）、及び第三者実装パートナー（“IP”）及び／又は鍵管理人を含む。第1のステップで、各DSは、患者記録の選択したデータ・フィールド（例えば、患者を識別する属性及び／又は他の標準的な属性のデータ・フィールド）を暗号化し、その患者記録を第1の“匿名化した”フォーマットへと変換する。各DSは、2つの鍵（即ち、ベンダー特定鍵、及び特定のLDFと関連付けられる共通の長期鍵）を用い、その選択したデータ・フィールドを二重暗号化する。二重暗号化したデータ記録は、装置コンポーネントサイトに送信される。ここで、二重暗号化したデータ記録は更に処理される。データ記録は、第2の匿名化したフォーマットへと処理され、第2の匿名化したフォーマットは、データ記録を、元の暗号化していない患者識別情報を復元することなく、個々の患者毎に効果的にリンクさせることができるように構成される。

【0021】

このため、DSから受信した患者記録の二重暗号化したデータ・フィールドは、特定のベンダー鍵を用いて、部分的に解読する（二重暗号化したデータ・フィールドが、まだ共通の長期鍵の暗号化を保持するように）。更に、長期のデータベースにおいて使用するための、新たな単一の（共通の長期鍵で）暗号化したデータ・フィールド又は属性を作成するように、第3の鍵（例えば、トークン・ベース鍵）を用いることができる。LDFに対して内部的となる、長期の識別子（ID）又はダミーラベルは、データ記録をタグ付けするように用いることができ、タグ付けしたデータ記録は、元の暗号化されていない患者識別情報の知識無く、長期のデータベースにおいて個々の患者毎にマッチングさせ、リンクさせることができる。

【0022】

好適なマッチングアルゴリズムは、前に規定したID又は新規のIDを、一組の暗号化した属性に対して割り当てべきか否か決定するように用いることができる。一旦、IDが決定されると、そのIDは、暗号化した属性を伴う処理を通過した一組の認証されたマッチングする属性を用いて、データ供給者からの詳細なトランザクションデータに対してリンクし戻される。次に、暗号化したデータ属性及び割り当てたIDは、更なるマッチング処理における使用のために参照データベース内に格納される。

【0023】

本発明によれば、データ記録における、ある選択組の属性／フィールドは、入ってくるデータ記録に対してIDを割り当てるために、その入ってくるデータ記録を、長期の参照データベースに格納される参照データ記録とマッチングさせるのに用いられる。マッチング・ルール及びマッチング処理は、入ってくるデータ記録を、増加的にレベル毎に（即ち、属性毎に）マッチングするように設けられる。マッチング・ルール及びマッチング処理は、無関係の薬局又はベンダーから受信した患者の記録をリンクさせることを可能とする。

【0024】

暗号化した患者の変数のわずか（例えば、小売店（例えば、開業されている薬局）、患者の性別及び患者の生まれた年、から非識別化した患者ID）を用いる従来のマッチング

処理は、非識別化した患者IDのみのマッチングによって、患者記録を小売店内でリンクするように用いることができる。本発明のマッチング処理は、追加数となる、現在利用可能な暗号化した非識別化した患者データ属性（例えば、完全な生年月日、保険カードID、姓、名、ジップコードなど）を用いる。これらの暗号化した変数は、小売店、供給者、及び他のデータ・ソースを渡って、非識別化した患者記録のリンクを可能とする。本発明のマッチング処理は、患者が引越し又は患者の姓が変更になった場合でさえ、患者のデータ記録を識別し、リンクすることができる。

【0025】

マッチング処理の一例において、入ってくる非識別化した患者記録を、参照データベース内の既存のデータ記録とマッチングさせる試みが為される。その試みは、階層構造のマッチングレベル（例えば、13レベル）で逐次的に為される。階層構造の一例の13レベルの各々で、入ってくる非識別化したデータ記録を、参照データベース内の記録に対してマッチングさせる試みが為される（即ち、入ってくる非識別化したデータ記録を、既存の長期IDと関連付けるためである）。首尾よくマッチングするものがあれば、非識別化した患者記録は、マッチングした参照データ記録と関連付けられるIDが割り当てられる。首尾よくマッチングするものが無ければ、非識別化した患者記録は、新たな患者ID（“長期のID”）を割り当てることができる。マッチングレベルの各々は、長期のIDに対する非識別化した患者記録をリンク又はマッチングさせるために、種々の組み合わせのデータ属性を用いる。表1は、IDを割り当て、本処理によってリンクすることが可能な患者記録・データ記録に含めることができる製薬上のデータ属性の一例の組をリストしたものである。

【0026】

【表1】

表1 マッチングロジックで用いられる一例の属性の組		
属性	暗号化	説明
供給者の患者ID	No	供給者によって作成した患者IDであり、小売店レベルで、又は、供給者レベルで固有のIDとできる。
供給者ID	No	IMSによって作成した供給者ID
小売店	No	IMSが作成した小売店（販売店）番号
Rx番号	No	小売店で割り当てた処方箋番号
NCPDP資格付与者 ¹	No	NCPDP IDが社会保障番号、軍事ID番号、運転ライセンス番号又はその他であるか否かを、資格付与者は識別する。
NCPDP ID	Yes	個人に対して固有のID（資格付与者に基づく）
患者の完全な生年月日(DOB)	Yes	デモグラフィック変数
患者の保険カードID	Yes	デモグラフィック変数
患者の性別	No	デモグラフィック変数
患者のジップ（5桁）	Yes	デモグラフィック変数
患者の姓	Yes	デモグラフィック変数
患者の名	Yes	デモグラフィック変数
患者のアドレス番号	Yes	一組の番号のうち、最小に現れる見込みのアドレスを検索することにより、作成されるデモグラフィック変数。（“1245メインストリート”は、“1245”になる。）

¹ NCPDP = National Council for Prescription Drug Programs. 薬局で用いられる標準の変数としての定義付けられる。

【0027】

表1にリストする13個のデータ属性の幾つか（例えば、供給者の患者ID、供給者ID、小売店及びRx#）は、患者識別情報について、明らかとならないものであるか、或いは少なくとも直接的に明らかとならない属性とできる。これらの属性の値は、データ記録において暗号化される必要はない。他の属性は、機密の患者識別情報（例えば、NCP

D P I D、患者の完全な生年月日（D O B）、患者の保険カード I D、患者の性別、患者のジップ（5桁）、患者の姓、患者の名、及び患者のアドレス番号）に関連するものとする。そのような属性の値は、データ記録において暗号化される。全ての暗号化したデータ・フィールド／属性は、暗号化される前では、好適な標準のフォーマットとすることができる。例えば、欠落及び無効となる値を識別でき、それぞれキャラクタ“B”と“I”でコード化することができる。“&”、“%”、“{”、及び“^”などの特別なキャラクタは、データ記録から除去できる。先行する空白又はゼロを除去でき、データ・フィールドの左右揃えをすることができる。更に、全てのキャラクタを大文字（又は小文字）に置換できる。

【0028】

図2a及び2bは、“供給者内”のマッチング処理200及び200'の一例のステップを示し、各処理は、I Dをデータ供給者から取得した患者トランザクションデータ記録に割り当てるように用いることができる。供給者I D番号（例えば、供給者I D 210）は、特定のデータ供給者を識別するように用いることができる。マッチング処理200及び200'は、長期のデータベース（例えば、解決策100、図1）をアセンブリするのに好適ないずれかの解決策の構成内で実現することができる。マッチング処理200及び200'は、データ供給者の暗号化した属性及び他のデータ供給者の標準化した属性を含めることができる、入ってくる患者トランザクションデータ記録を処理するように構成される。その入ってくる患者トランザクションデータ記録は、表1に示す13の個の属性の幾つか又は全てを含めることができ、更に、表1に示されていない他の属性を含めることができる。データ記録に含められる特定の属性は、データ供給者の採択又は能力に依存するものとする。マッチング処理200及び200'は、入ってくるデータ記録を参照データベースに格納されている参照データ記録と比較する。入ってくるデータ記録における統計的にかなりの数のデータ属性／フィールドが、特定の参照記録における対応するデータ属性／フィールドとマッチングする場合には、入ってくるデータ記録には、特定の参照記録と関連付けられたI Dを割り当てる。

【0029】

マッチング処理200及び200'は、表2と関連して図2aから理解することができる。表2は、参照データベース内の参照データ記録の対応する組み合わせと比較可能な、入ってくるデータ記録の属性の13個の組み合わせの一例（マッチングレベル1～13）をリストしている。各レベルは、データ記録が統計的に有効な方法でマッチングさせることができるように構成される。レベル1～4は、データ記録において、個々の患者、データ供給者、販売店及び他の当事者を指定するように用いられる固有の数字又は英数字の識別コードなどの属性値でマッチングを試みる。レベル1～3は、個々のデータ供給者の患者内でマッチングを試みることができる。レベル4は、データ供給者内の場合、及び、横断するデータ供給者の場合の双方でマッチングを試みることができる。レベル5は、マッチング処理を続けるために、入ってくる記録が、十分なデモグラフィックデータ又は情報を含むか否かを確かめることができる。入ってくる記録が、十分なデモグラフィックデータ又は情報を含んでいなければ、マッチングの試みを中止して、新規な長期のI Dを入ってくる記録に割り当てることができる。入ってくる記録が、十分なデモグラフィックデータを含んでいれば、より高いマッチングレベルへと進む。レベル6～12は、様々な組み合わせのデモグラフィック変数に基づいて、横断するデータ供給者に対するマッチングの試みを含んでいる。これらのレベルのほとんどに対して、入ってくる記録は、少なくとも5～7のデモグラフィック変数でマッチングしなければならない。唯一の例外はレベル7であり、レベル7では、入ってくるデータ記録が、3つの変数のみでマッチングさせることができる。なぜなら、レベル7は、他のデモグラフィック変数よりも、より固有なものとなる保険カードを利用するためである。入ってくる非識別化した患者記録が、レベル12で首尾よくマッチングしない場合には、その非識別化した患者記録に、新規の長期／患者I Dを割り当て、又は関連付けることができる。

【0030】

【表 2 - 1】

表 2 階層構造のマッチングレベルの説明		
マッチング レベル	入ってくる記録は、以下の 変数で、データベースと マッチングさせる。	説明
1	- 供給者の患者 ID - 供給者 ID	このレベルは、固有の患者 ID を有する供給者内・横断する供給者のマッチングを識別する。(例えば、ライトエイド)
2	- 供給者の患者 ID - 供給者 ID - 小売店	このレベルは、小売店での固有の患者 ID を有する供給者内のマッチングを識別する。
3	- 供給者 ID - 小売店 - R x 番号	このレベルは、固有の患者 ID システムを用いない供給者内のマッチングを識別する。
4	- NCPDP 資格付与者 - NCPDP ID	このレベルは、患者が、自己の社会保障番号、軍事 ID 番号、又は運転ライセンス番号 (NCPDP 資格付与者は、“01”、“02”又は“03”と等しい必要がある) を提供していた場合に、全ての横断する供給者のマッチングを識別する。 現時点では、これらの変数は、通常欠落している。
5	患者の完全な生年月日	レベル 5 は、他のレベルとは異なる。ここで、入ってくる記録がレベル 6 ~ 13 に進む前に、完全な生年月日が設定されていることを確実にするためにチェックする。生年月日が欠落又は無効である場合、その記録には新規の IMS 患者 ID が与えられる。そうでなければ、その記録はレベル 6 に進む。
6	- 患者の完全な生年月日 - 患者の保険カード ID - 患者の性別 - 患者のジップコード (5 桁) - 患者の姓 - 患者の名 - 患者のアドレス番号	このレベルは、7 つのデモグラフィック変数が等しい場合に、全ての横断する供給者のマッチングを識別する。
7	- 患者の完全な生年月日 - 患者の保険カード ID - 患者の性別	このレベルは、3 つのデモグラフィック変数のみを用いてマッチングを識別する (それらのうちの 1 つが数字であれば高い固有の変数であることを意味する。) 姓及びアドレスについての変更を取り込むことができる。
8	- 患者の完全な生年月日 - 患者の性別 - 患者のジップコード (5 桁) - 患者の姓 - 患者の名	このレベルは、5 つのデモグラフィック変数を用いてマッチングを識別する。保険カード ID についての変更を取り込むことができる。
9	- 患者の完全な生年月日 - 患者の性別 - 患者のジップコード (5 桁) - 患者の姓 - 患者のアドレス番号	このレベルは、5 つのデモグラフィック変数を用いてマッチングを識別する。保険カード ID についての変更を取り込むことができる。

【表 2 - 2】

表 2 階層構造のマッチングレベルの説明		
マッチングレベル	入ってくる記録は、以下の変数で、データベースとマッチングさせる。	説明
10	- 患者の完全な生年月日 - 患者の性別 - 患者のジップコード (5桁) - 患者の名 - 患者のアドレス番号	このレベルは、5つのデモグラフィック変数を用いてマッチングを識別する。保険カードID 及び姓 についての変更を取り込むことができる。
11	- 患者の完全な生年月日 - 患者の性別 - 患者の姓 - 患者の名 - 患者のアドレス番号	このレベルは、5つのデモグラフィック変数を用いてマッチングを識別する。保険カードID 及びジップコード についての変更を取り込むことができる。
12	- 患者の完全な生年月日 - 患者のジップコード (5桁) - 患者の姓 - 患者の名 - 患者のアドレス番号	このレベルは、5つのデモグラフィック変数を用いてマッチングを識別する。保険カードID についての変更を取り込むことができる。
13	無し	このレベルは、患者データベースにマッチングしない記録を識別する。新規IMS患者IDが割り当てられる。

10

20

【0032】

データ属性は、マッチングが見つかるまでの階層的な系列で組み合わせ毎に（即ち、レベル毎に）、評価した組み合わせとできる。表1にリストされる特定の組み合わせの属性は、一例を示すものであることは明らかである。データ記録のマッチング処理について様々な実現例では、これらの組み合わせの幾つか又は全て、及び／又は他の所望される又は適切な組み合わせを用いることができる。

30

【0033】

図2aを参照して、マッチング処理200の予備のステップで、入ってくる“対象の”データ記録内の符号化した供給者IDの属性は、特定のデータ供給者IDと関連付けられる供給者ID 210と比較される。対象のデータ記録が、供給者ID 210で符号化されていない場合、対象のデータ記録は、マッチング処理300（図3a）に転送される。マッチング処理300は、異なるデータ供給者に渡って患者記録をマッチングさせることができる。

【0034】

40

対象のデータ記録が、供給者ID 210で符号化される場合、ステップ220で、処理200は、データ供給者（即ち、供給者ID 210）によって対象のデータ記録に含めることができる、何らかの患者識別番号を評価する。データ供給者は、患者データ記録を識別するために、普遍的な、即ち“横断する供給者”の患者識別番号、又は他の固有の識別番号を含めることができる。普遍的な患者識別番号が、対象のデータ記録内に与えられている場合に、参照データ記録内の患者識別番号と比較される。普遍的な患者識別番号が、参照データ記録内の患者識別番号とマッチングする場合、レベル1のマッチングが宣言される。データ供給者が、普遍的な患者識別番号の代わりに、小売店の特定の患者識別番号を含める場合、ステップ220で、対象のデータ記録内のその小売店の特定の患者識別番号及び小売店の属性が、参照データ記録内のそれらと比較される。これら2つのデー

50

タ属性について首尾よくマッチングした場合には、レベル2のマッチングが宣言される。或いは又、ステップ220で、対象のデータ記録にて普遍的な患者識別番号又は小売店の特定の患者識別番号の、いずれも使用しない供給者からのデータに対して、本処理200は、対象のデータ記録内の小売店及び処方箋番号の属性と、参照データ記録内のそれらとの比較による首尾よいマッチングに基づいて、ステップ220で、レベル3のマッチングを宣言するように構成させることができる。首尾よいレベル1、2及び3のマッチングで、対象のデータ記録には、マッチングした参照データ記録と関連付けられる長期のIDが割り当てられる。参照データ記録に対するマッチングが、ステップ220で全く得られない場合には、処理200は、より高いレベルのマッチングを試みる。

【0035】

ステップ230で、処理200は、対象のデータ記録に与えることができる、何らかのNCPPDの患者IDについてのNCPPD資格付与者を調べる。NCPPD資格付与者が、標準の保険タイプ01、02又は03（例えば、社会保障番号、軍事ID番号、又は、運転ライセンス番号のそれぞれに対応する）であれば、ステップ240で、処理200は、対象のデータ記録内に与えられるNCPPD患者IDと、参照データ記録内のそれらとを比較する。処理200は、対象のデータ記録内のNCPPD患者IDと参照データ記録内のそれとを比較して首尾よいマッチングに基づいて、レベル4のマッチングを宣言するように構成させることができる。ステップ230で、処理200が、NCPPD患者IDが与えられていない、又は、標準の保険タイプ01、02又は03のものではないと決定する場合、処理200は、ステップ250へと進み、対象のデータ記録に与えることができる追加のデータ属性（例えば、ドモグラフィックデータ属性）を評価することができる。例えば、ステップ250で、対象のデータ記録に与えられている3つの属性（例えば、保険カード番号、生年月日（DOB）、及び性別の属性）からなる組を、参照データ記録内のそれらと比較することができる。処理200は、対象のデータ記録に与えられているこれら3つの属性と参照データ記録内のそれらと比較して首尾よくマッチングに基づいて、レベル7のマッチングを宣言するように構成させることができる。処理200は、予備のステップ（図示せず）で、まだマッチングしていない対象のデータ記録が、関連する又は重要な追加のデータ属性（例えば、DOB属性）を含んでいることを確認した後のみ、ステップ250へと進み、追加のデータ属性を評価することができる。

【0036】

ステップ250で用いられる3つの属性の組は、保険カード番号及びDOBなどの、数値を有する属性を含む。これらの属性のうち1つを含んでいない対象のデータ記録（即ち、保険カード番号が欠落）を処理するために、処理200は、データ供給者から取得した患者トランザクションデータ記録にこの時点で共通して含まれている追加のデータ属性に基づいて、対象のデータ記録のマッチングを試みる。ステップ260で、処理200は、対象のデータ記録に与えることが可能な5つのデモグラフィック属性の組の値を、参照データ記録内のそれらと逐次的に比較することにより、対象のデータ記録のマッチングを試みる。5つのデモグラフィック変数の組の例は、共通して3つの属性（即ち、DOB、性別、姓の属性）、及び、4つの属性（名、ジップコード、処方者名、アドレス番号）のグループから選択した2つの追加の属性を含む。図2aは、対象のデータ記録のマッチングに逐次的に用いられる5つのデータ属性についての6つのそのような組を示す。本処理200は、対象のデータ記録に与えられている5つのデータ属性のそのようないずれかの組と、参照データ記録内のそれらと比較して首尾よいマッチングに基づいて、マッチング（例えば、レベル8～12又はそれ以上の高いレベル）を宣言するように構成させることができる。

【0037】

或いは又、処理200は、変更したステップ260'で、対象のデータ記録に与えることが可能な4つのデモグラフィック属性の値と参照データ記録内のそれらとの逐次的な比較によって、対象のデータ記録をマッチングするように構成させることができる。4つのデモグラフィック変数の組の例は、共通して、3つの属性（DOB、性別、及び姓）、及

10

20

30

40

50

び、4つの属性（名、ジップコード、処方者名、及びアドレス番号）のグループから選択した1つの追加の属性を含んでいる。図2 aは、対象のデータ記録をマッチングするのに逐次的に用いる、4つのデータ属性の4つのそのような組を示す。処理200は、対象のデータ記録に与えられる4つのデータ属性のそのようないずれかの組と、参照データ記録内のそれらと比較して首尾よいマッチングに基づいて、マッチングを宣言するように構成させることができる。

【0038】

ステップ210～260又は260'で、首尾よいマッチングが見つけれない場合には、処理200は、新規な患者に属するとして、対象のデータ記録を認識するように構成され、それに従い、新規なIDタグを対象のデータ記録に割り当てるように構成される。新規なIDタグは、対象のデータ記録を、長期のデータベースの逐次的にマッチングしたデータ記録にリンクさせるように用いることができる。

10

【0039】

本処理200にて、数字又は英数字の識別コードの属性に基づくマッチングを含むステップ210～230は、低リスクの誤検知の結果を与える。第三者データ（例えば、保険カード番号）に基づくマッチングを含むステップ250は、適度のリスクの誤検知の結果を与える。デモグラフィックデータに基づくマッチングを含むステップ260は、増大したリスクの誤検知の結果を与える。マッチング処理200は、データ処理のステージなしで、LDF（又は任意の他の無許可のエンティティ）が患者識別可能なデータを見ることができるような、プライバシー要件と互換性がある。例えば、LDFは、ある患者（例えば、Mr. XXY&^%FR21gywt30Jjokiau76gfs545aなどの長期IDによってのみ識別される患者）が複数の薬局又は販売店に通っていたことを知ることができる。しかしながら、LDFは、複数の販売店に通っていた特定の“Mr. John Smith”を知ることではなく、識別することはできない。

20

【0040】

図2 bは、長期のデータベースに包含させる対象のデータ記録を処理するように用いることができる別のマッチング処理200'を示す。本処理200'は、IDを対象のデータ記録に割り当てることについて、処理200とは、デモグラフィックデータ属性の異なる組み合わせを用い、異なる論理的フローのマッチングステップを用いる。処理200のように、本処理200'は、特定のデータ供給者（例えば供給者ID 210）からデータ記録を処理するように構成されるが、首尾よいマッチングを確かめるために、追加のステップを含む。処理200'のステップ210～240にて、処理200と同様なステップは、値として英数字の識別コードを有するデータ属性の検討を含むが、確認するステップ270が後に続く。ステップ210～240で、参照データ記録に対する対象のデータ記録の如何なる首尾よいマッチングも、対象データ記録のDOB及び性別の属性が、マッチングした参照データ記録内のそれらと同一であることをチェックすることにより、ステップ270で確かめられる。

30

【0041】

更に、本処理200'のステップ230は、NCPDP患者IDが与えられておらず、標準の保険タイプ01、02又は03のものではないと決定する場合、本処理200'は、ステップ245～295、260及び260'に進み、対象のデータ記録に与えられるデータ属性の追加の組を評価する。ステップ245～295、260及び260'で検討したデータ属性の組は、処理200のステップ250～260'で検討したそれらとは異なるものとできる。ステップ245で、本処理200'は、対象のデータ記録に与えられる小売店識別の属性及び販売店特定の患者識別の属性についてマッチングを試みる。ステップ255で、本処理200'は、対象のデータ記録の患者に与えられる小売店識別の属性及び処方箋番号の属性についてのマッチングを試みる。双方のステップでの首尾よいマッチングは、更に、首尾よいマッチングが宣言される前に、確認するステップ270で確かめられる。

40

【0042】

50

更に、小売店識別の属性及び処方箋番号の属性について少なくとも1つを含まない対象のデータ記録を処理するために、本処理200'のステップ280は、3つの属性（保険カード番号、DOB、及び性別）の組を評価することによって、対象のデータ記録のマッチングを試みる。参照データ記録で如何なるマッチングも見つけれない場合には、処理200'のステップ290は、首尾よいマッチングを宣言する前に、マッチングする参照データ及び対象のデータ記録が同一の姓及びジップコードの属性を有することを確認する。

【0043】

対象のデータ記録が、保険カード番号の属性を含んでいない場合では、本処理200'は、対象のデータ記録に与えられうる追加のデモグラフィック属性に基づいて対象のデータ記録のマッチングを試みることができる。ステップ295で、本処理200'は、6つのデータ属性（例えば、ジップコード、DOB、名、姓、アドレス番号及び性別）の組の値を比較することにより、参照データ記録に対して対象のデータ記録のマッチングを試みる。ステップ260及び260'で、本処理200'は、処理200のように、5つ又は4つのデータ属性の組の値のそれぞれについて比較することにより、参照データ記録に対して、対象のデータ記録のマッチングを試みる。ステップ260での5つのデータ属性の組は、共通して、3つのデータ属性、即ちジップコード、DOB及び性別の属性を含んでいる。5つのデータ属性の組のうち、残りの2つの属性は、名、姓及びアドレス番号の属性からの一対の属性として選択される。ステップ260'での4つの属性の組は、共通して、2つのデータ属性、即ちDOB及び性別を含んでいる。4つのデータ属性のうち、他の対となる属性は、（ジップコード、姓）、（ジップコード、アドレス番号）又は（名、姓）のうちいずれかの対の1つとして選択される。

【0044】

マッチング処理2000及び200'は、特定のデータ供給者から受信した対象のデータ記録を処理するように構成される。また、マッチング処理は、如何なる数のデータ供給者からも受信できる複数のソースの対象のデータ記録（に対して処理し、好適な長期IDを割り当てるように構成させることができる。図3a及び3bは、“横断する供給者”の場合のマッチング処理300及び300'のステップ例をそれぞれ示す。マッチング処理300及び300'は、処理200及び200'のそれぞれと連結して、又は拡張して動作させることができる。

【0045】

処理300及び300'は、処理200及び処理200'の予備ステップ210において、対象のデータ記録が特定の供給者の供給者ID 210で符号化されていないことを示すときに、開始することができる。処理200及び200'とは異なり、“横断する供給者”の場合のマッチング処理300及び300'は、供給者の特定の患者識別の属性を比較することによって（例えば、処理200及び200'参照）、参照データ記録に対して対象のデータ記録のマッチングを試みることはしない。この相違点とは別にして、処理300は、参照データ記録に対する対象のデータ記録のマッチングの試みにおいて、処理200（例えば、ステップ230～260及び260'）と同一又は同様の属性の逐次的な組み合わせ又はレベルを用いることができる。同様に、処理300'は、参照データ記録に対する対象のデータ記録のマッチングの試みにおいて、処理200'（例えば、ステップ230～295、260及び260'）と同一又は同様の属性の逐次的な組み合わせ又はレベルを用いることができる。

【0046】

デモグラフィック属性を含む複数組の属性の組み合わせに基づいている、階層構造のマッチング処理（例えば、処理200、200'、300及び300'）は、長期の処方箋データベース（例えば、LRxデータベース）の情報コンテンツを拡張するために有利に用いることができる。例えば、2つの薬局間を通う非識別の患者についての非識別化したRx記録は、現時点で、同一のIDに対して認識又はリンクさせることができるものとなっている。以前、これらのRx記録は、2人の異なる者に対してリンクされていたはずの

ものである。ここで、L R x データベースは、ビジネスのソース（例えば、新スタート、再スタート、継続患者など）、治療パターン（例えば、単剤療法、併用療法、及び追加併用療法、及びセラピー・進行状況（例えば、第1、第2及び第3ライン、療法の長さ、過去の療法／中止したセラピー使用、残留性及びコンプライアンス、タイトレーション及び薬剤投与、及び一般的な患者の特徴）など患者中心の基準を含むように好適にアップグレードできる。階層的なマッチング処理に用いられる、マッチングロジック及び追加の暗号化した属性は、複数の薬局及びヘルスケア・ファシリティ間を通う非識別の患者をリンクすることにより、患者中心の基準の精度を改善させる。本発明のマッチング処理は、プライバシー要件と互換性がある。本マッチング処理は、データ処理のステージなしで、L D F（又は任意の他の無許可のエンティティ）が患者識別可能なデータを見ることができるように構成される。例えば、L D Fは、ある患者（例えば、Mr. XXY&^%FR21gywt30Jjokiau76gfs545aなどの長期 I D又はコードによってのみ識別される患者）が複数の販売店に通っていたことを知ることができる。しかしながら、L D Fは、複数の販売店に通っている特定の“Mr. John Smith”を知ることはなく、識別することはできない。

10

【0047】

本発明の原理に従って、属性の組み合わせ、階層構造のマッチングレベル、及びマッチング処理に用いられるマッチングロジックは、特定のデータ処理要件又は事情に適切に適合させながら、好適に変更又は変化させることができる。また、マッチングレベルの階層構造のレベルの数は、変更可能である。あるレベルが非効率的であると決定される場合（即ち、多数の非識別の患者をリンクしない）又は異なる順序でマッチングレベルを処理することがより効果的な処理をもたらすことを決定される場合、マッチングロジック及びマッチングレベルの数を変更（例えば、減少）できる。

20

【0048】

特定のデータ処理要件又は事情に適合する好適なマッチング処理又はマッチングアルゴリズムの開発は、最善のマッチングアルゴリズム、又はデータ記録における選択組のデータ属性を与えられるマルチソース化したデータ記録から、個々の患者を固有に識別する最善の方法を決定する問題として定性化させることができる。例えば、マッチング処理 200、200'、300及び300'の開発において、12個のデータ属性の開始組を、参照データ記録に対してマッチングさせる対象のデータ記録において、ユーティリティ用に、及び効率化のために選択し、テストした。12個のデータ属性の開始組は、標準化したデータ属性、例えば、カードホルダ I D、完全な生年月日（D O B）、性別、名、姓、アドレス（クリーンにし、標準化したもの）、ジップコード（5桁）、N C P D P I D及びN C P D P資格付与者、R x番号、小売店、供給者 I D、患者 I D（薬局からのもの）を含ませた。幾つかのデータ属性（例えば、名、姓、アドレス及び患者 I D）は、ヘルスケア・データ供給者から取得されるデータ記録において、ルーチン的(routinely)に暗号化される。当該分野の展開させるためのマッチング処理の評価及び開発は、供給者及び他のデータ暗号化処理の有効性の決定、並びに、アドレス番号の属性を、肯定的な結果を有する完全なアドレス属性の代わりに用いることができるか否かを見るための検討を含むものであった。更に、マッチング処理は、二人の個人としてではなく、一人の個人として、供給者内又は横断する供給者について、2つの異なる薬局又は小売店からの処方薬を購入する患者を識別するように構成される。

30

40

【0049】

マッチング処理のフィールド実現例について、属性又はマッチングレベルの特定の組み合わせ、階層構造のマッチングレベル、及び展開されるマッチングロジックは、論理的又は経験的な考慮に基づいて選択することができる。マッチングで利用される属性の組み合わせ及びタイプは、例えば、参照データ記録とマッチングする対象のデータ記録におけるそれらのポテンシャルに従って選択することができる。処理 200、200'、300及び300'で利用されるマッチングレベル又は属性の組み合わせは、そのような考慮に基づいて選択した。表3及び表4は、長期の処方箋データベース（L R x）の場合に、参照データ記録に対する対象のデータ記録のマッチングにおいて、経験的に決定したマッチン

50

グの有用性又は有効性を有する、処理200、200'、300及び300'で用いられる属性の組についてリストしている。表3は、“供給者内”の場合のデータ記録のマッチングに関するものであり（例えば、マッチング処理200及び200'）、表4は、“横断する供給者”の場合のデータ記録のマッチングに関するものである（例えば、マッチング処理300及び300'）。双方の表において、マッチング処理において用いられるデータ属性及び属性の組み合わせは、マッチング・ポテンシャルに従ってランク付けさせている（例えば、マッチングランクのベスト・マッチング=1、ワースト・マッチング=9）。

【0050】

【表3】

表3 “供給者内”でのLRx患者IDの最小のマッチング													
供給者内	マッチング ランク	供給者ID (2桁)	供給者 特定の 患者ID番号	NCPDP ID番号	NCPDP ID 資格付与者	保険カード 番号	患者 DOB	患者 性別	姓	名	患者 ジップ コード	処方者	アドレ ス 番号
ID番号 第三者 患者 デモグラ フィックス	ベスト マッチ	1	X		X								
		2	X										
		3	X			X	X	X					
		4	X				X	X	X	X	X		
		5	X				X	X	X	X		X	
		6	X				X	X	X	X			X
ワースト マッチ		7	X				X	X	X		X	X	X
		8	X				X	X	X		X		X
		9	X				X	X	X			X	X

供給者内	マッチング ランク	供給者ID (2桁)	供給者 特定の 患者ID番号	NCPDP ID番号	NCPDP ID 資格付与者	保険カード 番号	患者 DOB	患者 性別	姓	名	患者 ジップ コード	処方者	アドレ ス 番号
ID番号 第三者 患者 デモグラ フィックス	ベスト マッチ	1	X										
		2	X		X								
		3	X			X	X	X					
		4	X				X	X	X	X			
		5	X				X	X	X		X		
		6	X				X	X	X			X	
ワースト マッチ		7	X				X	X	X				X

10

20

30

40

50

【0051】

【表4】

表4 “候補する供給者” の場合のLLRx患者IDの最小のマッチング													
候補する供給者	マッチング ランク	供給者ID (2桁)	供給者 特定の 患者ID番号	NCPDP ID番号	NCPDP ID 資格付与者	保険カード 番号	患者 DOB	患者 性別	姓	名	患者 ジップ コード	処方者	アドレス 番号
ID番号	ベスト マッチ			X	X								
第三者 患者 デモグラ フィックス	1					X	X	X	X	X			
	2						X	X	X	X	X		
	3						X	X	X	X			
	4						X	X	X	X		X	
	5						X	X	X	X			X
	6						X	X	X	X			X
ワースト マッチ	7						X	X	X	X	X		X
	8						X	X	X	X		X	X

候補する供給者	マッチング ランク	供給者ID (2桁)	供給者 特定の 患者ID番号	NCPDP ID番号	NCPDP ID 資格付与者	保険カード 番号	患者 DOB	患者 性別	姓	名	患者 ジップ コード	処方者	アドレス 番号
ID番号	ベスト マッチ			X	X								
第三者	2					X	X	X	X	X			
患者 デモグラ フィックス	3						X	X	X	X			
	4						X	X	X	X	X		
	5						X	X	X	X		X	
ワースト マッチ	6						X	X	X	X			X

【0052】

決定論的又は確率論的なマッチング基準又は論理のいずれかをマッチング処理に用いることができ、2つのデータ記録間で比較し、首尾よくマッチングしたことを宣言させる。処理に利用可能なヘルスケア・データ記録では、ほとんどの患者デモグラフィック属性は暗号化されている。更に、ある属性（例えば、姓）が、同一であるべき別の属性（例えば、Smith とSmyth）に対する値において十分に“近い”という状況を判断することは難しい。そのような状況では、対象のデータ記録を参照データ記録と比較する確率論的マッチングロジックを用いることは、誤検知に通じうる。それ故、マッチング処理（例えば、処

10

20

30

40

50

理200、200'、300及び300')は、データ属性を比較するために、正確な、即ち決定論的なマッチングロジックを用いるように構成させることができる。そのようなロジックの下で、2つの患者データ記録を、2値形式でマッチングするか否かを表す。決定論的ロジックを用いて得られる、明確な2値による正又は否の結果は、マッチング処理の実用上の実現に対して有利である。更に、データ記録の比較に関して、完全なアドレス属性の代わりにアドレス番号の属性を用いることは、データ記録のマッチングに有利であることを検討により示された。アドレス番号の使用は、データ記録に完全なアドレスを入力する際に生じうるデータエントリ・エラーを低減させることもできる。

【0053】

マッチング処理（例えば、処理200、200'、300及び300'）におけるマッチングレベルの階層構造は、特定のデータ属性（又は属性の組み合わせ）を選択することにより構成させることができ、一致する際に最善の患者マッチングをもたらす。より優れた患者マッチングは、誤検知のより低い見込み度に対応する。階層構造のマッチングレベルは、最小数のデータ属性を要求し、各レベルでテストするように構成させることができる。対象の患者データ記録が、或るレベルで参照データ記録とのマッチングに失敗する場合、マッチング処理は、次のレベルへと移行し、マッチングを検索する。これは、より低いマッチングレベル番号の首尾よいマッチングが、より良いマッチングとなることを意味している。表3及び4は、処方箋データ記録（LRxデータベース）の場合に“供給者内”及び“横断する供給者”の場合のそれぞれについての首尾よいデータ記録マッチングを宣言する最小のマッチング要件となるように見つけられた、特定の組み合わせの属性を区別して表している。追加の、又は、代替りの属性の組み合わせは、他の状況又は事情で好適に見つけることができる。例えば、処方箋データ記録をリンクするマッチング処理では、マッチングする属性として、“医師”を考慮することが役立つ。

【0054】

マッチング処理は、マッチングしない状況とは対照的に、データ属性／フィールドが対象データ記録から無効又は欠落しているデータ処理状況をアカウントするように好適に構成させることができる。テーブル5は、2つのデータ記録が同一又は一致する患者に属するときでさえ、2つのデータ記録のデータ属性の値が同一である必要はない（即ち、マッチングしない状況）場合のデータ属性のリスト例である。また、テーブル5には、同一又は一致する患者に属する、2つの対象のデータ記録におけるデータ属性が異なることについて、可能性ある理由を示している。

【0055】

テーブル5

どの属性が変更可能でありながら、まだ同一の患者に対するものであるか？	
属性	変更理由
保険カード番号	新保険会社
患者ジップコード	引越し
姓	結婚
名	不正確の入力???
アドレス番号	引越し

【0056】

マッチング処理を、コンピュータプログラムとして実現させることができる。本発明によれば、前述したマッチング処理を実現するソフトウェア（即ち、コンピュータプログラム命令）をコンピュータ読取り可能なメディアで提供することができる。各ステップ（本発明に従って前述した）については、これらステップの如何なる組み合わせも、コンピュータプログラム命令によって実現させることができることは明らかである。如何なる好適なコンピュータプログラム言語も本目的のために用いることができる。図4は、患者データ記録を処理するコンピュータ・サブルーチン例400としてのマッチングアルゴリズムの実現例を示す。サブルーチン400にて、マッチング・ルールは、一連のネストされた

IF-ELSE IF-THENの条件文として、選択組のデータ属性に適用されており、その各々は、あるレベル、即ちテストしたデータ記録（例えば、表1参照）におけるデータ属性の組み合わせに対応している。マッチングアルゴリズムは、9個のレベルの階層構造において13個の変数を用いる（13個の変数の組み合わせ）。

【0057】

コンピュータプログラム命令を、コンピュータ又は他のプログラム可能な装置にロードさせ、マシンを構成させることができ、これにより、コンピュータ又は他のプログラム可能な装置で実行させるコンピュータプログラム命令は、前述したマッチング処理の機能を実現するための手段を作り出す。また、これらのコンピュータプログラム命令をコンピュータ読取り可能なメモリに格納することができ、コンピュータ又は他のプログラム可能な装置を特定の態様で機能するように方向付けることができ、これにより、コンピュータ読取り可能なメモリに格納されるコンピュータプログラム命令は、前述した分布型の推測統計学的なコントローラ及びシステムの機能を実現する命令手段を有する製造物を生成させる。また、コンピュータプログラム命令をコンピュータ又は他のプログラム可能な装置にロードさせ、コンピュータ又は他のプログラム可能な装置で実現させるべき一連の動作ステップを生じさせ、コンピュータ又は他のプログラム可能な装置で実行するコンピュータプログラム命令が、前述したマッチングアルゴリズム及びマッチング処理の機能を実現するステップを与えるように、コンピュータで実現される処理を可能とする。また、前述したマッチングアルゴリズム及びマッチング処理を実現する命令が与えられるコンピュータ読取り可能なメディアは、そのように限定するものではなく、ファームウェア、マイクロコントローラ、マイクロプロセッサ、集積回路、ASIC、及び他の利用可能なメディアを含むことは明らかである。

【0058】

サブルーチン例400のプログラムロジックは、対象の入ってくる患者記録が、参照データベースに格納できる参照データ記録と比較されることが想定されている。各マッチングレベル（例えば、比較のために指定されるデータ属性の組）で、サブルーチン400は、参照データベースをスキャンし、マッチングする参照データ記録を検索する。指定したマッチングレベルのいずれにおいてもマッチング参照データ記録が見つからなければ、サブルーチン400は、対象の入ってくる患者記録が新規の患者又は新規IDと関連付ける必要があることを決定できる。

【0059】

本発明のマッチング処理における他のプログラム実現例は、図4に示されているものより、少ないレベル又は多いレベルを含むことができる。例えば、レベル12は、サブルーチン400の幾つかのフィールド実現例から除去した。更に、本プログラム実現例は、入ってくるデータ記録の整数をチェックするように構成させた好適なルーチンを含めることができる。例えば、サブルーチン400は、入ってくるデータ記録におけるデータ属性の無効又は欠落を識別するステップを含むように、変更又は拡張できる。更に、本プログラム実現例は、種々のマッチングロジックで共通のデータ処理する事項を処理するように構成させることができる。例えば、本プログラム実現例のマッチングロジックは、女性患者の姓の変更を許容でき、女性患者の変更した姓を有するデータ記録を、元の姓を有するデータ記録で肯定的にマッチングさせることができる。

【0060】

本発明のマッチング処理における本プログラム実現例について、データ供給者又はベンダーからのテストデータを用いて、テストし、評価し、最適化することができる。単一のデータ供給者／ベンダー（例えば、McKesson）からのテストデータを用いて、あるプログラム実現例のマッチングロジックについて、以下のようにテストすることができる。

最もマッチングするレベルの判断、及び、マッチング処理のスピードアップのために必要とされるマッチングレベルの階層構造の変更。

各データ属性が、如何に多くB、S又はIと等しくなるかの判断。

人コード／関係コードがマッチング処理の助けとなるか否かの判断。

各要素から作成されたある長期鍵と対照したものとして、個々の要素でのマッチングにより得られるマッチングがどれほど多く得られるかの判断。

入ってくる記録が、如何に多く参照患者マッチングデータベースにおける1つ以上の記録にマッチングするかの判断。これは、如何に多くL D Fが“クリーンアップ”を実行するべきかについて明確な助けとなる。

姓、カードID、アドレス、及びジップの属性が、或る期間内で単一の患者に対し、如何に多く変更するかの判断。

【0061】

第2のデータ供給者／ベンダー（例えば、Pharmacy Benefit Management (PBM)データ供給者）からの追加のテストデータを用いて、プログラム実現例のマッチングロジックを、更に以下のようにテストすることができる。

10

患者が、ある地理的範囲内で、如何に多く第1及び第2のデータ供給者／ベンダー販売店の間を通うかの判断。

この数が、現実的であるか？

他の患者による通いの学習結果から得られたデータと同様であるか？

【0062】

PBMから受信したスクリプトデータ（Rxスクリプト）を、プログラム実現例の評価のテスト前に、処理し、間引き、クリーンにできる。スクリプトデータを区別するのに必要なデータは、供給者ID、小売店番号、保険会社、保険カード番号、DOB、姓（オプション）、名（オプション）、性別、及び（Rx番号、完了日、医師）又は（薬番号、完了日、医師）などの属性を含むことができる。PBMデータの利用（例えば、薬局を横断する患者のマッチング用）は、PBMで供給されたデータ属性に依存するものとできる。例えば、PBMは、小売店の特定のRx番号、完了日、医師、従属のDOB、性別のデータ属性を供給することができる。従って、プログラム実現例で用いられるマッチングロジックは、そのようなデータ属性に基づくものとできる。実用上、データ供給者に渡って患者マッチングの有用性を判断するために、特定組の属性（例えば、州（State）、国、MSA、ジップコード（5桁）及びジップコード（3桁））に注意を向けることができる。また、薬局を横断する患者のマッチング結果は、小売店が、1日に一回以上、同一のRx番号を発行できるか否かに依存するものとできる。

20

【0063】

試し動作では、本発明のマッチングアルゴリズム（例えば、サブルーチン400）は、テストデータセット（即ち、データ供給者PD X及びMcKessonから得られる暗号化したデータ記録）に対して評価してみた。表5～7及び図8A～8Eは、試し動作の結果を示す。

30

【0064】

表5は、LRx患者IDを割り当てる最小のマッチング要件として、試し動作で用いたデータ属性の様々な組み合わせに対する効率的なマッチングランクを示す。表6及び7は、共通のデータ処理状態又はシナリオの下で、2つのデータ記録間でのマッチングを見つけるためのポテンシャルベースとして、データ属性の様々な組み合わせの検討及びテストの結果を示す。

【0065】

40

【表 5】

表 5 L R x 患者 ID の最小のマッチング																			
供給者内 タイプ	レベル	マッチ ング ランク	供給者 ID (2桁)	供給者 特定 ID 番号	NCPDP ID 番号	NCPDP ID 資格付与者		小売店		販売店 特定 ID 番号		R x 番号	保険カード 番号	患者 コード	患者 DOB	名	姓	アドレ ス 番号	患者 性別
						ST	ST	ST	ST	ST	ST								
タイプ ID 番号	レベル					ST	ST	ST	ST	ST	ST	ST	U	U	U	U	U	U	U
	ベスト マッチ	1	X	X	X	X													
		2	X	X	X														
		3	X	X				X	X										
第三者 患者		4	X	X				X	X			X							
		5	X	X									X	X					
		6	X	X								X		X	X	X	X	X	X
		7	X	X										X	X	X	X	X	X
デモグラ フィックス		8	X	X										X	X	X	X	X	X
		9	X	X										X	X	X	X	X	X
		10	X	X										X	X	X	X	X	X
	フースト マッチ	11	X	X										X	X	X	X	X	X

【 0 0 6 6 】

植断する 供給者	タイプ	マッチング ランク	供給者 ID (2桁)	供給者 特定 ID 番号	NCPDP ID 番号	NCPDP ID 資格付与者	小売店	販売店 特定 ID 番号	R x 番号	保険カード 番号	患者 ジップ コード	患者 DOB	姓 名	姓 名	アドレ ス 番号	患者 性別
植断する 供給者	レベル	12	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	ベスト マッチ	13	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		14	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		15	X	X	X	X	X	X	X	X	X	X	X	X	X	X
デモグラ フィックス		16	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		17	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		18	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	フースト マッチ	19	X	X	X	X	X	X	X	X	X	X	X	X	X	X

10

20

30

40

【表 6】

表 6

マッチングログブック

レベル	テスト	マッチ/非マッチの記録	記
1	ライト-エイドタイプのIDでのマッチング	マッチングした	これらのIDが正しいと、現在想定されるか？
2	供給者+小売店レベルの患者IDでのマッチング	マッチングした	
3	供給者+小売店+R x 番号でのマッチング	マッチングした	
4	NCPDP ID+NCPDP 資格付与者でのマッチング	マッチングした	ヒットしたフィールドの正確さは、どれほどの信頼性？
5	DOBが、非マッチング又は欠落	マッチングしなかった	DOBのマッチングは、必須！
6	性別が非マッチング	マッチングしなかった	性別が欠落しており、設定できるのであれば、マッチング必須
7	DOB+ (保険カード+ 5つのグループA変数の全て)	マッチングした	全てのデモグラフィック変数でのマッチング
8	DOB+ (保険カード ID+ 性別)	マッチングした	第三者の支払者用のアカウントであり、アドレス、名及び姓の変更の識別可。
9	性別=女性、且つ姓がマッチングしない場合	マッチングしなかった	男性の姓は、変更できないと想定している。性別=男性、且つ姓が欠落している場合、次のレベルに進む。
10	グループAの3～5の変数がマッチングしない場合	マッチングしなかった	どの変数がマッチングしないかは、問題にしない。
11	DOB+グループAの4～5の変数のマッチングの場合	マッチングした	現金での支払者用のアカウント、及び保険番号の変更。
12	DOB+グループAの3～5の変数のマッチングの場合	できる限りマッチングした	
13	他	マッチングしなかった	

デモグラフィック変数

保険カード

DOB

性別

患者のジップ

姓

名

アドレス番号

グループA

性別

患者のジップ

姓

名

アドレス番号

変更可能な変数

新プランの保険カード

引越した患者のジップ

結婚した姓

ニックネーム、不正確な入力

引越したアドレス番号

【 0 0 6 7 】

10

20

30

40

【表 7】

表 7

マッチングロジック

5C3 = 常にDOB
5セット= (性別、ジップコード、姓、名、アドレス)

10個の組み合わせ			例	コンフリクト	コンフリクトの 見込み度	レフトアウト		レフトアウト変数に関して、非マッ チに対する欠落は問題となるか?	マッチ/ 非マッチ
DOB	性別	ジップ	姓	DOBを有する複数の男性 Smith	6- 中間	名	アドレス	No	?
DOB	性別	ジップ	名	同一のジップコード内で同一の DOBを有する複数の John	7- 中間 高	姓	アドレス	非マッチの姓は、 レベル9で解析される。	非マッチ
DOB	性別	ジップ	アドレス	同一のジップコード内で同一の DOBを有する複数の John	9- 高	姓	名	非マッチの姓は、 レベル9で解析される。	非マッチ
DOB	性別	姓	名	同一のDOBを有する John Smith	6- 中間	ジップ	アドレス	No	?
DOB	性別	姓	アドレス	同一のDOB及び同一のアドレ ス番号を有する複数の Smith	5- 中間	ジップ	名	No	?
DOB	性別	名	アドレス	同一のアドレス番号内の同一の DOBを有する複数の John	6- 中間	ジップ	姓	非マッチの姓は、 レベル9で解析される。	?
DOB	ジップ	姓	名	同一のDOB及び同一のジップ コードを有する複数の John Smith	3- 低	性別	アドレス	非マッチの性別は、 レベル6で解析される。	マッチ
DOB	ジップ	姓	アドレス	同一のDOB及び同一のアドレ ス番号を有する複数の Smith	5- 中間	性別	名	非マッチの性別は、 レベル6で解析される。	?
DOB	ジップ	名	アドレス	同一のDOB及び同一のアドレ ス番号を有する複数の John	6- 中間	性別	姓	非マッチの性別及び姓は、 レベル6及び9で解析される。	?
DOB	姓	名	アドレス	同一のDOB及び同一のアドレ ス番号を有する複数の John Smith	3- 低	性別	ジップ	非マッチの性別は、 レベル6で解析される。	マッチ

* 以下、定性的に割り当てられる。

1=コンフリクトの低い見込み度、又は非マッチ; 5=コンフリクトの中間の見込み度; 10=コンフリクトの高い見込み度

【0068】

表 8A～8Eは、テストデータ記録が特定のタイプのデータ属性を含んでいた又は除外していた種々のシナリオの下で、データ属性の様々な組み合わせに対するマッチング結果を示す。表 8Aは、保険カード、DOB及び性別の情報を含んでいたテストデータ記録に対するテスト結果を示す。表 8Bは、保険カード、DOBの情報を含むが、性別情報は欠いていたテストデータ記録のテスト結果を示す。表 8Cは、保険カード情報は含むが、DOB情報は欠い

10

20

30

40

50

ていたテストデータ記録のテスト結果を示す。更に、表 8 Dは、DOB情報は含むが、保険カード情報は欠いていたテストデータ記録のテスト結果を示す。表 8 Eは、DOB情報及び保険カード情報の双方を欠いていたテストデータ記録のテスト結果を示す。

【0069】

【表 8 A】

表 8 A

供給者内 カウンント	供給者 iD	保険 iD番号	患者 DOB	患者 性別	患者 ジップ コード	姓	名	アドレス 番号	マッチした 数	許可可?	再確認
1	X	X	X	X	X	X	X	X	8	Y	
2	X	X	X	X	X	X	X	O	7	Y	
3	X	X	X	X	X	X	O	X	7	Y	
4	X	X	X	X	X	O	X	X	7	Y	
5	X	X	X	X	O	X	X	X	7	Y	
6	X	X	X	X	X	X	O	O	6	Y	
7	X	X	X	X	X	O	X	O	6	Y	
8	X	X	X	X	O	O	X	X	6	Y	
9	X	X	X	X	X	O	O	X	6	Y	
10	X	X	X	X	O	X	O	X	6	Y	
11	X	X	X	X	O	O	X	X	6	Y	
12	X	X	X	X	X	O	O	O	5	Y	
13	X	X	X	X	O	X	O	O	5	Y	
14	X	X	X	X	O	O	X	O	5	Y	
15	X	X	X	X	O	O	O	X	5	Y	
16	X	X	X	X	O	O	O	O	4	Y	

【0070】

10

20

30

40

【表 8 B】

表 8 B

供給者内 カウ ント	供給者 ID	保険 ID番号	患者 DOB	患者 性別	患者 ジッブ コード	姓	名	アドレス 番号	マッチした 数	許可可?	再確認
保険カード及びDOBを母とする	17	X	X	O	X	X	X	X	7	?	Y
	18	X	X	O	X	X	X	O	6	?	Y
	19	X	X	O	X	X	O	X	6	?	N
	20	X	X	O	X	O	X	X	6	?	Y
	21	X	X	O	O	X	X	X	6	?	Y
	22	X	X	O	X	X	O	O	5	?	N
	23	X	X	O	X	O	X	O	5	?	Y
	24	X	X	O	O	X	X	O	5	?	Y
	25	X	X	O	X	O	O	X	5	?	N
	26	X	X	O	O	O	X	X	5	?	N
	27	X	X	O	O	O	X	X	5	?	Y
	28	X	X	O	X	O	O	O	4	?	N
	29	X	X	O	O	X	O	O	4	?	N
	30	X	X	O	O	O	X	O	4	?	Y
	31	X	X	O	O	O	O	O	3	?	N
	32	X	X	O	O	O	O	O	3	?	N

【0071】

10

20

30

40

【表 8 C】

表 8 C

供給者内 カウント	供給者 ID	保険 ID 番号	患者 DOB	患者 性別	患者 ジューブ コード	姓	名	アドレス 番号	マッチした 数	許可可?	再確認
保険カードを有するが、DOBが欠落	33	X	X	O	X	X	X	X	7	?	N
	34	X	X	O	X	X	X	O	6	?	N
	35	X	X	O	X	X	O	X	6	?	N
	36	X	X	O	X	X	O	O	5	?	N
	37	X	X	O	X	O	X	X	6	?	N
	38	X	X	O	X	O	O	O	5	?	N
	39	X	X	O	X	O	O	O	5	?	N
	40	X	X	O	X	O	O	O	4	?	N
	41	X	X	O	X	X	X	X	6	?	N
	42	X	X	O	X	X	X	O	5	?	N
	43	X	X	O	X	X	O	X	5	?	N
	44	X	X	X	X	O	O	O	5	?	N
	45	X	X	O	X	O	X	X	5	?	N
	46	X	X	O	X	O	X	O	4	?	N
	47	X	X	O	X	O	O	X	4	?	N
	48	X	X	O	X	O	O	O	3	?	N
	49	X	X	O	X	X	X	X	6	?	N
	50	X	X	O	X	X	X	O	5	?	Y
	51	X	X	O	X	X	O	X	5	?	N
	52	X	X	O	X	X	O	O	4	?	N
	53	X	X	O	X	O	X	X	5	?	N
	54	X	X	O	X	O	X	O	4	?	N
	55	X	X	O	X	O	O	X	4	?	N
	56	X	X	O	X	O	O	O	3	?	N
	57	X	X	O	O	X	X	X	5	?	N
	58	X	X	O	O	X	X	O	4	?	N
	59	X	X	O	O	X	O	X	4	?	Y
	60	X	X	O	O	X	O	O	3	?	N
	61	X	X	O	O	O	X	X	4	?	N
	62	X	X	O	O	O	X	O	3	?	N
	63	X	X	O	O	O	O	X	3	?	N
	64	X	X	O	O	O	O	O	2	?	N

【 0 0 7 2 】

【表 8 D】

表 8 D

供給者内 カウント	供給者 ID	保険 ID番号	患者 DOB	患者 性別	患者 ジップ コード	姓	名	アドレス 番号	マッチした 数	許可可?	再確認
保険サービスが提供される、DOBが与えらる	65	X	O	X	X	X	X	X	7	?	
	66	X	O	X	X	X	X	O	6	Y	
	67	X	O	X	X	X	O	X	6	Y	
	68	X	O	X	X	X	O	O	5	Y	
	69	X	O	X	X	O	X	X	6	?	N
	70	X	O	X	X	O	X	O	5	?	N
	71	X	O	X	X	O	O	X	5	?	N
	72	X	O	X	X	O	O	O	5	?	N
	73	X	O	X	X	O	O	O	4	?	N
	74	X	O	X	X	X	X	X	6	?	N
	75	X	O	X	X	X	O	O	5	?	N
	76	X	O	X	X	X	O	O	4	?	N
	77	X	O	X	X	O	X	X	5	?	N
	78	X	O	X	X	O	X	O	4	?	X
	79	X	O	X	X	O	O	O	4	N	
	80	X	O	X	X	O	O	O	3	N	
	81	X	O	X	X	X	X	X	6	Y	
	82	X	O	X	X	X	X	O	5	?	Y
	83	X	O	X	X	X	O	X	5	?	Y
	84	X	O	X	X	X	O	O	4	?	N
	85	X	O	X	X	O	X	X	5	?	N
	86	X	O	X	X	O	X	O	4	?	N
	87	X	O	X	X	O	O	X	4	?	N
	88	X	O	X	X	O	O	O	3	N	
	89	X	O	X	X	O	X	X	5	?	Y
	90	X	O	X	X	O	X	X	O	4	?
	91	X	O	X	X	O	O	O	3	N	
	92	X	O	X	X	O	O	O	4	?	N
	93	X	O	X	X	O	X	X	4	?	N
	94	X	O	X	X	O	O	O	3	N	
	95	X	O	X	X	O	O	X	3	N	
	96	X	O	X	X	O	O	O	2	N	

【0073】

10

20

30

40

【表 8 E - 1】

表 8 E

供給者内 カウント	供給者 ID	保険 ID番号	患者 DOB	患者 性別	患者 ジップ コード	姓	名	アドレス 番号	マッチした 数	許容可?	再確認
97	X	O	O	X	X	X	X	X	6	?	N
98	X	O	O	X	X	X	X	O	5	?	N
99	X	O	O	X	X	X	O	X	5	N	
100	X	O	O	X	X	X	O	O	4	N	
101	X	O	O	X	X	O	X	X	5	N	
102	X	O	O	X	X	O	X	O	4	N	
103	X	O	O	X	X	O	O	X	4	N	
104	X	O	O	X	X	O	O	O	O	3	
105	X	O	O	X	O	X	X	X	5	?	
106	X	O	O	X	O	X	X	O	4	N	
107	X	O	O	X	O	X	O	O	3	N	
108	X	O	O	X	O	O	X	X	4	N	
109	X	O	O	X	O	O	O	X	3	N	
110	X	O	O	X	O	O	X	O	3	N	
111	X	O	O	X	O	O	O	X	3	N	
112	X	O	O	X	O	O	O	O	2	N	
113	X	O	O	O	X	X	X	X	5	N	
114	X	O	O	O	X	X	X	O	4	N	
115	X	O	O	O	X	X	O	O	4	N	
116	X	O	O	O	X	X	O	O	3	N	
117	X	O	O	O	X	O	X	X	4	N	
118	X	O	O	O	X	O	X	O	3	N	
119	X	O	O	O	X	O	O	X	3	N	
120	X	O	O	O	X	O	O	O	2	N	
121	X	O	O	O	O	X	X	X	4	N	
122	X	O	O	O	O	X	X	O	3	N	
123	X	O	O	O	O	X	O	X	3	N	
124	X	O	O	O	O	X	O	O	2	N	
125	X	O	O	O	O	O	X	X	3	N	
126	X	O	O	O	O	O	X	O	2	N	
127	X	O	O	O	O	O	O	X	2	N	
128	X	O	O	O	O	O	O	O	1	N	

保険カード及びDOBを欠落

【表 8 E - 2】

再確認	
許容可?	
マッチした数	
アドレス番号	
名	
姓	
患者 ジップ コード	
患者 性別	
患者 DOB	
保険 ID番号	
供給者 ID	
カウント	129
供給者内	

10

20

30

【0075】

40

表 8A～8Eに示すようなテスト結果の分析は、効率的なマッチング処理実現のために、特定組のデータ属性及びレベルを推奨又は選択する経験的な基礎をもたらす。

【0076】

更に、他のデータ処理状況又は事情に対するマッチングロジックを洗練させるために、分析又はテストを実行することができる。例えば、“横断する供給者”の場合のマッチング処理に関して、“販売店の普遍性”データファイルを分析して、2つのデータ供給者（例えば、McKesson及びPDX）が如何に多く同一の地理的範囲内に販売店を有するかについて判断することができる。この分析は、例えばMcKesson及びPDXの患者間で、通い患者のマッチング期待度の判断を助ける。分析で考慮されうるデータ属性は、例えば、州、国、MSA、ジップコード（5桁）及びジップコード（3桁）を含むことができる。

50

【0077】

前述の例は、本発明の概略的な原理を示したものに過ぎず、本発明の範囲及び趣旨から逸脱することなく、当業者であれば様々な変更が可能であることは明らかであり、本発明は、特許請求の範囲によってのみ制限される。

【図面の簡単な説明】

【0078】

【図1】本発明の原理に従って、米国特許出願第11／122,589号明細書から再現される、マルチソースの患者データ記録から長期のデータベースをアセンブリするシステム例のブロック図であり、本システムで実現できるマッチングアルゴリズムは、図2a、2b、3a及び3bに示される。

10

【図2a】本発明の原理に従って、データ記録の属性をレベル毎にマッチングし、特定のデータ供給者によって供給されるデータ記録に、長期リンクの識別子を割り当てるステップ例を示す概略的な処理フロー図である。

【図2b】本発明の原理に従って、データ記録の属性をレベル毎にマッチングし、特定のデータ供給者によって供給されるデータ記録に、長期リンクの識別子を割り当てるステップ例を示す概略的な処理フロー図である。

【図3a】本発明の原理に従って、データ記録の属性をレベル毎にマッチングし、複数のデータ供給者によって供給されるデータ記録に、長期リンクの識別子を割り当てるステップ例を示す概略的な処理フロー図である。

【図3b】本発明の原理に従って、データ記録の属性をレベル毎にマッチングし、複数のデータ供給者によって供給されるデータ記録に、長期リンクの識別子を割り当てるステップ例を示す概略的な処理フロー図である。

20

【図4-1】本発明の原理に従って、米国特許出願第11／122,589号明細書から再現される、図2a、2b、3a及び3bに示されるマッチング処理の実現に関するソフトウェア・サブルーチンのロジックを示す図である。

【図4-2】本発明の原理に従って、米国特許出願第11／122,589号明細書から再現される、図2a、2b、3a及び3bに示されるマッチング処理の実現に関するソフトウェア・サブルーチンのロジックを示す図である。

【図 1】

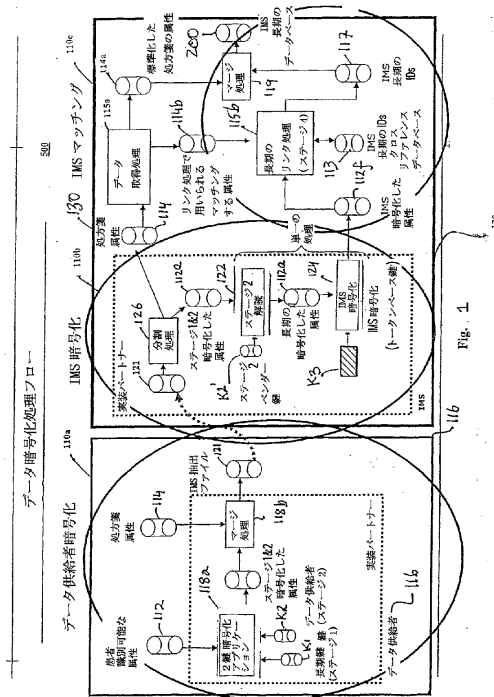


Fig. 1

【図 2 a】

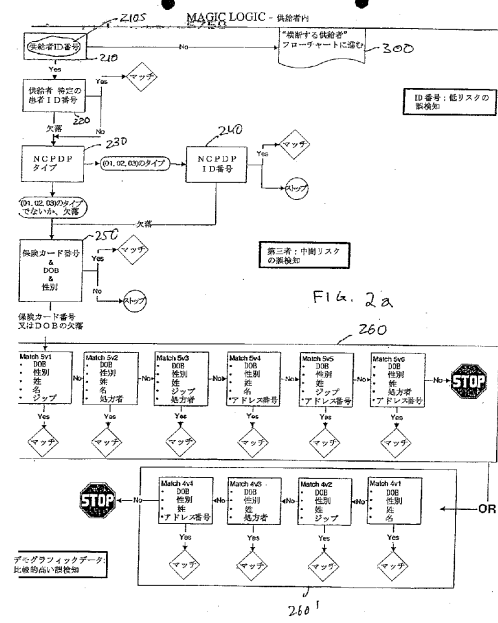


Fig. 2a

【図 2 b】

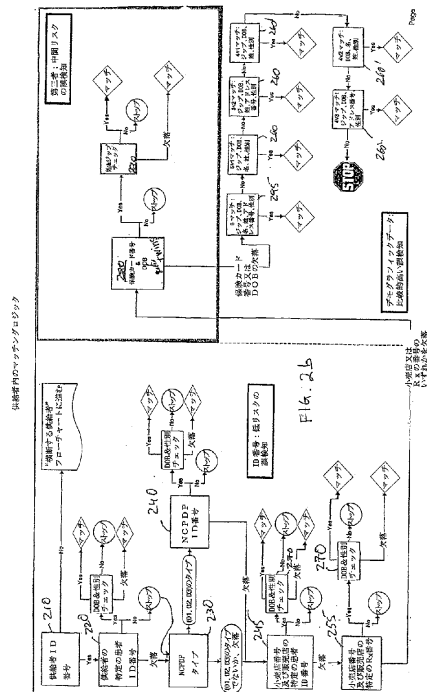


Fig. 2b

【図 3 a】

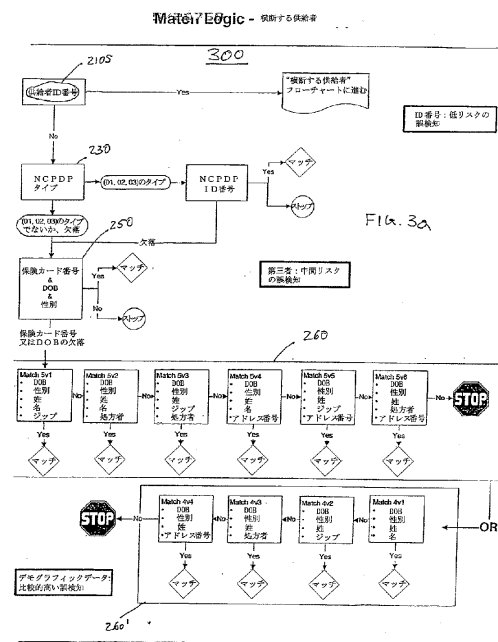


Fig. 3a

フロントページの続き

合議体

審判長 清田 健一

審判官 須田 勝巳

審判官 井上 信一

(56)参考文献 米国特許第6 3 9 7 2 2 4 (U S, B 1)
特開2 0 0 1－3 2 5 4 2 5 (J P, A)

(58)調査した分野(Int.Cl., D B名)
G06Q10/00-50/00