(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6563812号
(P6563812)

(45)発行日　令和1年8月21日(2019.8.21)　(24)登録日　令和1年8月2日(2019.8.2)

(51)Int.Cl.		F I		
G06F 21/62	(2013.01)	G06F	21/62	354
G06F 21/31	(2013.01)	G06F	21/31	
G06F 21/60	(2013.01)	G06F	21/60	360
G16H 15/00	(2018.01)	G16H	15/00	

請求項の数 13　(全 16 頁)

(21)出願番号　特願2015-546142(P2015-546142)
(86)(22)出願日　平成25年12月9日(2013.12.9)
(65)公表番号　特表2016-505948(P2016-505948A)
(43)公表日　平成28年2月25日(2016.2.25)
(86)国際出願番号　PCT/IB2013/060736
(87)国際公開番号　WO2014/091385
(87)国際公開日　平成26年6月19日(2014.6.19)
審査請求日　平成28年11月29日(2016.11.29)
(31)優先権主張番号　61/735,245
(32)優先日　平成24年12月10日(2012.12.10)
(33)優先権主張国・地域又は機関
米国(US)

前置審査

(73)特許権者　590000248
コーニンクレッカ　フィリップス　エヌヴェ
KONINKLIJKE PHILIPS N.V.
オランダ国　5656　アーエー　アインドーフェン　ハイテック　キャンパス　5
High Tech Campus 5, NL-5656 AE Eindhoven
(74)代理人　100122769
弁理士　笛田　秀仙
(74)代理人　100163809
弁理士　五十嵐　貴裕

最終頁に続く

(54)【発明の名称】マルチサイトパフォーマンス測定を匿名にし、匿名データの処理及び再識別を制御する方法及びシステム

(57)【特許請求の範囲】
【請求項１】
データソースを匿名にしたデータを送信する複数のデータソースを有するシステムであって、
各データソースが、
臨床データを作成するデータ作成エンジンであって、前記臨床データが、前記データソースの識別が推定されることを可能にする１つ又は複数の属性を含む、データ作成エンジンと、
前記臨床データを非正規化されたデータへと変換し、前記非正規化されたデータを遠隔計算リソースに送信する変換エンジンとを含み、
前記変換が、各グループに関する前記属性の値が目標数を超えないよう、前記臨床データを複数のグループに分割し、前記複数のグループの各グループに個別の識別子を割り当てることにより行われる、システム。
【請求項２】
前記変換エンジンが、前記複数のグループの各グループに割り当てられる個別の識別子を対応する前記１つ又は複数の属性へ戻すためのマッピングを規定する表を生成する、請求項１に記載のシステム。
【請求項３】
前記遠隔計算リソースが、前記１つ又は複数のデータソースからの前記非正規化されたデータを利用し、パフォーマンス分析及び／又はレポートを提供する、請求項１又は２に

記載のシステム。
【請求項４】
前記データソースが、健康プロバイダである、請求項１乃至３の任意の一項に記載のシステム。
【請求項５】
データソースを匿名にしたデータを送信する方法において、
複数のデータソースにより、臨床データを作成するステップであって、前記臨床データが、前記データソースの識別が推定されることを可能にする１つ又は複数の属性を含む、ステップと、
変換エンジンにより、前記臨床データを非正規化されたデータへと変換するステップと、
前記非正規化されたデータを遠隔計算リソースに送信するステップとを有し、
前記変換するステップが、各グループに関する前記属性の値が目標数を超えないよう、前記臨床データを複数のグループに分割し、前記複数のグループの各グループに個別の識別子を割り当てることにより行われる、方法。
【請求項６】
データソースを匿名にしたデータを送信する複数のデータソースを有するシステムであって、
各データソースが、臨床データを作成し、前記臨床データを遠隔計算リソースに送信するデータ作成エンジンであって、前記臨床データが、前記データソースの識別が推定されることを可能にする１つ又は複数の属性を含む、データ作成エンジンを含み、
前記遠隔計算リソースは、前記複数のデータソースから前記臨床データを受信してこれを格納し、及び
前記遠隔計算リソースは、前記臨床データを非正規化されたデータへと変換し、前記非正規化されたデータを格納する変換エンジンを含み、
前記変換が、各グループに関する前記属性の値が目標数を超えないよう、前記臨床データを複数のグループに分割し、前記複数のグループの各グループに個別の識別子を割り当てることにより行われる、システム。
【請求項７】
前記遠隔計算リソースが、前記非正規化されたデータへのアクセス及び再識別を制御する、請求項６に記載のシステム。
【請求項８】
前記遠隔計算リソースが、前記データプロバイダにログインしたユーザの役割に基づき、前記非正規化されたデータに関して呼び出される処理を制御する、請求項７に記載のシステム。
【請求項９】
各データソースが、再識別処理を呼び出すために前記遠隔計算リソースに証明書を送信する認証エンジンを有する、請求項８に記載のシステム。
【請求項１０】
前記遠隔計算リソースが、前記再識別処理を認証するため、前記データソースから送信される証明書と格納された証明書とを比較する、請求項９に記載のシステム。
【請求項１１】
前記変換エンジンが、前記複数のグループの各グループに割り当てられる個別の識別子を対応する前記１つ又は複数の属性へ戻すためのマッピングを規定する表を生成する、請求項６乃至１０の任意の一項に記載のシステム。
【請求項１２】
データソースを匿名にする変換エンジンデバイスであって、
メモリとプロセッサとを有し、前記プロセッサが、
臨床データを非正規化されたデータへと変換し、前記非正規化されたデータを遠隔計算リソースに送信しており、

前記臨床データが、前記データソースの識別が推定されることを可能にする１つ又は複数の属性を含み、
前記変換が、各グループに関する前記属性の値が目標数を超えないよう、前記臨床データを複数のグループに分割し、前記複数のグループの各グループに個別の識別子を割り当てることにより行われる、変換エンジンデバイス。
【請求項１３】
データソースを匿名にする変換エンジンデバイスであって、
メモリとプロセッサとを有し、前記プロセッサが、
複数のデータソースから臨床データを受信し、及び
前記臨床データを非正規化されたデータへと変換し、前記非正規化されたデータを前記メモリに格納しており、
前記臨床データが、前記データソースの識別が推定されることを可能にする１つ又は複数の属性を含み、及び
前記変換が、各グループに関する前記属性の値が目標数を超えないよう、前記臨床データを複数のグループに分割し、前記複数のグループの各グループに個別の識別子を割り当てることにより行われる、変換エンジンデバイス。
【発明の詳細な説明】
【技術分野】
【０００１】
本願は、遠隔計算リソースにより臨床データを分析することに関する。それは、マルチサイトパフォーマンス測定を匿名にするシステム及び方法と連動して特定の用途を見いだす。それは更に、匿名データの処理及び再識別を制御するシステム及び方法と連動して特定の用途を見いだす。しかしながら、それは、他の使用シナリオにおける用途も見いだし、必ずしも上述した用途に限定されるわけではない点を理解されたい。
【背景技術】
【０００２】
複数の健康サイトのパフォーマンスを測定することは、任意の品質改良イニシアティブの部分であるだけでなく、多くの場合、州、連邦政府、プライベートな投資家等によっても必要とされる。概して、健康サイトはレポーティングを共有するよう求められるとき、データに貢献する健康サイトによる抵抗又はデータ信頼性の低下が存在する。これは概して、パフォーマンスが欠如している領域を他者に見られることを恐れるが故である。これは、動作改良プロジェクトを識別する際、さらに重要なことに、他者を明らかに好適に上回る看護プロセスを実施する破壊的な革新者を識別する際に遅延を生じさせる。
【発明の概要】
【発明が解決しようとする課題】
【０００３】
一般的なＩＤにサイトをマッピングすることは、パフォーマンスデータを部分的に非識別する（de-identify）態様として使用される。しかし、どれくらい多くのサイト及びユニットがシステムの一部であるかを知ることにより、複数のサイトを復号化することがまだ可能である。例えば、１２ベッドの集中治療室（ＩＣＵ）及び１２ベッドの冠疾患病室（ＣＣＵ）を持つ単一の病院と、それぞれが６つのＩＣＵ（新生児集中治療室（ＮＩＣＵ）を含む）を持つ６つの病院を持つ企業とが符号化される場合、３６のユニット（室）が大きなサイトに帰属し、２のユニットが小さな施設に帰属することが明らかである。同様に、あるサイトでは年間３６００人の訪問者があり、別のサイトでは年間７８０００人の訪問者がある場合、又はあるサイトのデータにＮＩＣＵデータがある場合、どのサイトがどちらに対応するかは明らかである。
【０００４】
群における他者がパフォーマンスを推定することができるというリスクなしに、任意のサイトがかれらの結果を群と比較することができるという態様で、本願は多くのパフォーマンスデータベースの連合を支持する。これは、共有されたパフォーマンス結果に対する

各サイトの匿名化を確実にしつつ、一方で、データセットにおける特定の非特異属性が維持される。

【０００５】

更に、非識別されたデータセットにおけるリアルタイムのデータ解析は、「ビッグデータ」臨床調査に対する基礎である。１つのチャレンジは、いったん「正しい」患者セットが得られると、異なるデータを集めるか、又は解析リポジトリにおけるデータを他の非識別されていない（non-de-identified）データとマージする必要がある点にある。これは、非正規化されたデータセットの再識別を必要とする。

【０００６】

データ処理プロセスを駆動するためクラウドにおいて固定された役割を使用し、それらのパーミッションに対してオンプレミス（on-premise）ユーザ認証を呼び出すことにより、データを所有する（非識別された解析システムに貢献する）サイトは、そのデータに関して許される処理を制御する（表示する、レポートを実行する、再識別する、エクスポートする...）。更に、サイト発行トークンの形で、サイト提供される処理サービスに対して、認可トークンがアプリケーションにより発行される。これは、サイトがユーザプールを管理し、クラウド認証システムにユーザを追加することなしに、彼らが処理許可を持たせたいと望む者を許可することを可能にする。さらに重要なことに、トークンは、クラウドにおいてユーザ役割又はパスワードを学習した未許可のユーザ又は不正なユーザが、まず所有サイトにより認証されることなしに、再識別サービスを起動することを防止する。

【０００７】

本願は更に、そのデータに対する特定の処理を実行するために誰が権限を持つかをサイトが管理することができるという態様で、多くの非識別された臨床データベースの連合を支持する。更に、このシステムは、所有サイトのユーザ認証なしに、処理が呼び出されることができないことを確実にする。更に、このシステムは、データ再識別プロセスが、オンプレミスサーバで実行され、従って、クラウドに対して被保護健康情報（ＰＨＩ）が送信されないことを確実にする。

【０００８】

本願は、上述した課題及びその他を解決する、新規かつ改良された装置及び方法を提供する。

【課題を解決するための手段】

【０００９】

１つの側面によれば、データソースを匿名にするシステムが提供される。このシステムは、複数のデータソースを有し、各データソースが、共通スキーマに基づき、正規化されたデータを作るデータ作成エンジンと、上記データソースが他のデータソースにより推定されることができないよう、上記正規化されたデータを非正規化し、上記非正規化されたデータを遠隔計算リソースに送信する変換エンジンとを含む。遠隔計算リソースは、上記複数のデータソースから上記非正規化されたデータを受信し及びこれを格納する。

【００１０】

別の側面によれば、データソースを匿名にする方法が提供される。この方法は、複数のデータソースにより共通スキーマに基づき、正規化されたデータを作るステップと、上記データソースが他のデータソースにより推定されることができないよう、変換エンジンにより、上記正規化されたデータを非正規化するステップと、上記非正規化されたデータを遠隔計算リソースに送信するステップと、上記遠隔計算リソースにおいて上記複数のデータソースからの上記非正規化されたデータを格納するステップとを有する。

【００１１】

別の側面によれば、データソースを匿名にするシステムが提供される。このシステムは、複数のデータソースを有し、各データソースが、共通スキーマに基づき、正規化されたデータを作り、上記正規化されたデータを遠隔計算リソースに送信するデータ作成エンジンを含む。遠隔計算リソースは、上記複数のデータソースから上記正規化されたデータを受信してこれを格納し、及び上記データソースが他のデータソースにより推定されること

ができないよう、上記正規化されたデータを非正規化し、上記非正規化されたデータを格納する変換エンジンを含む。

【００１２】

１つの利点は、健康臨床データを匿名化する点にある。

【００１３】

別の利点は、ホスティングドメインの外側で認証情報を送信することなく、ユーザ認証に基づき、非正規化されたデータを再識別する点にある。

【００１４】

別の利点は、実際の認証されたユーザをクラウド環境に露出させることなしに、ユーザの役割に基づき、処理を制御する点にある。

【００１５】

別の利点は、送信の前にどの送信データをどのように縮小させるかを規定する連合化されたホストにある。こうして、ソースデータが傍受から更に保護される。

【００１６】

別の利点は、サイト、ユニット、ユーザ又は患者情報の無許可の再識別のリスクが低減されるため、ベンチマーキングデータへのアクセスが改善される点にある。

【００１７】

別の利点は、本発明により可能にされるベンチマーキングから生じる改善された臨床ケア及び効率にある。

【図面の簡単な説明】

【００１８】

【図１】本願によるＩＴ基盤のブロック図である。

【図２】マルチサイトパフォーマンス測定を本願に基づき匿名にする方法のフローチャート図である。

【図３】本願による健康プロバイダ構成の図である。

【図４】本願による健康プロバイダのアクセスシナリオの図である。

【図５】本願による再識別ロジックの図である。

【図６】本願による直接的なクラウドのアクセスシナリオの図である。

【図７】本願による匿名データの処理及び再識別を制御する方法のフローチャート図である。

【発明を実施するための形態】

【００１９】

本発明の更に追加的な利点は、以下の詳細な説明を読み及び理解することにより当業者に理解されるだろう。

【００２０】

本発明は、様々な要素及び要素の配列の形式並びに様々なステップ及びステップの配列の形式を取ることができる。図面は、好ましい実施形態を説明するためだけにあり、本発明を限定するものとして解釈されるべきものではない。

【００２１】

図１を参照すると、ブロック図が、遠隔計算リソースにより分析される臨床データに対するアクセスを制御するシステムの情報技術（ＩＴ）基盤１０の一実施形態を示す。ＩＴ基盤１０は、通信ネットワーク１６を介して相互接続される１つ又は複数の健康システム及びプロバイダ１２、遠隔計算リソース又は連合化されたデータストア１４等を適切に含む。通信ネットワーク１６は、イントラネット、局所エリアネットワーク、ワイドエリアネットワーク、無線ネットワーク、有線ネットワーク、セルラー電話網、データバス、個人エリアネットワーク等の１つ又は複数を含むと考えられる。健康プロバイダ１２は、健康システム又は医学機関によりケアされる１人又は複数の患者に関連付けられる臨床データを集める。このデータは、遠隔計算リソース１４により分析及び格納される。このリソースは、ある実施形態ではクラウド基盤に配置される。他の実施形態において、このシステムは、単一の位置に配置されることができる。更に別の実施形態では、このシステムは

、セキュア環境に存在することができる。しかし、データ通信は、公衆メディア又は共有基盤にわたり行われる。

【００２２】

本書に使用される「クラウド」は、オフサイト又はオフプレミスの通信相手（例えば、サードパーティ）により提供及び維持される資源の集合（例えば、ハードウェア、データ及び／又はソフトウェア）を指すことができる。この場合、データ及び資源の集合は、識別されたユーザによりネットワークを介してアクセスされることができる。資源は、データストレージサービス、データ処理サービス（例えば、アプリケーション）及び従来はパーソナルコンピュータ、ローカル又は「オンプレミス」サーバに関連付けられ、及びこの中に存在する他の多くのサービスを含むことができる。こうしたコンピュータ又はサーバは、例えばマイクロプロセッサ、グラフィックスプロセッサ及び関連付けられる要素といった少なくとも１つの処理デバイスを持つ。一般に、クラウドコンピューティングは、分離的な態様でサービスを実行するために使用されることができる。即ち、クライアントは、予想されるクオリティオブサービスでサービスが実行される限り、サービスがどこで実行されるかについては知ることができない。

【００２３】

本書に使用される「臨床データ」は、患者又は医療機関から、任意の数の従来の態様において集められるデータを指すことができる。例えば、臨床データは、医師又は臨床医といった健康プロバイダにより、フィールドにおいて集められることができる。別の実施形態では、臨床データは、所与の患者又はサブ集団の健康に関連するデータを含む。別の実施形態では、臨床データは、健康プロバイダの構造及びローカルパフォーマンスに関連するデータを含む。代替的に、患者は、病院又は緊急クリニックといった健康プロバイダに入院し、関連付けられる臨床データが例えば、健康プロバイダでの入院又は管理により集められることができる。臨床データは、他の医療デバイスにより集められることができる。それは例えば、Ｓｐ０２、温度、血圧、心拍等の各バイタルサインに関するさまざまなサブシステムを含む患者モニタ、さまざまな撮像装置、ペースメーカモニタ、探査デバイス、研究室装置及び他の臨床データ収集システムである。臨床データは、患者のホーム監視システムにより集められることもできる。このシステムは、物理的、化学的、電気的、又は他の患者の臨床パラメータを報告することができる。本書において用いられるデータ収集は、所定のイベント又は確率過程に基づかれ偶発的、例えば４時間毎といった周期的、又は、連続的とすることができる。データ収集は、リアルタイム、ほぼリアルタイムに実行されるか、又は以前に取得され、後でアップロードされることができる。

【００２４】

健康システム、アプリケーション、処理又はプロバイダ（これより前において健康プロバイダと呼ばれる）１２は、データのオーナーの代わりにデータを作成するか、又は収集された臨床データを処理し、分析及び／又はレポーティングのために遠隔計算リソース１４に対して安全に臨床データを送信する。臨床データを受信した後、遠隔計算リソース１４は、臨床データを処理し、この分析から１つ又は複数の結果及び／又はレポートを生成する。上述したように、健康プロバイダがレポーティングを共有するよう要求されるとき、臨床データに貢献する健康サイトによる抵抗又は臨床データ信頼性の低下が存在する。これは概して、パフォーマンスが欠如している領域を他の健康プロバイダに見られることを恐れることが原因である。一般的なＩＤにサイトをマッピングすることは、臨床データを部分的に非識別する態様として以前は使用されていた。しかしながら、臨床データのコンテンツが原因で、特定の健康プロバイダに関する基本的な知識を知ることにより、複数のサイトを復号化することがまだ可能である。そのようなものとして、本願は、他のプロバイダがフォーマンスを推定することができるというリスクなしに、任意の健康プロバイダが、それらの結果及び／又はレポートを一団の健康プロバイダと比較することができるという態様で、臨床データの収集をサポートする。これは、共有された結果及び／又はレポートに対して各健康プロバイダの匿名化を行い、一方、臨床データセットにおける特定の非特異属性が維持されることを確実にする。

【００２５】
特に、健康プロバイダ１２は、例えば患者、健康プロバイダ自体等といったデータのオーナーの代りに、正規化された形において臨床データを作成するデータ作成エンジン１８を含む。このデータはその後、正規化されたデータセットデータベース２０に格納される。ある実施形態において、データ作成エンジン１８により作成される臨床データは、任意の数の従来の態様において患者又は医療機関から集められるデータを含む。別の実施形態では、データ作成エンジン１８は、共通スキーマに基づき臨床データを作成する。その結果、データが、特定されることができ、他の健康プロバイダ及び遠隔計算リソース１４における一般の項へとマッピング可能でありえる。データ作成エンジン１８により作成される臨床データは、健康プロバイダに関する一意な属性を含む点も理解されたい。これは、臨床データのソースとして健康プロバイダを推定する通常の態様である。属性は例えば、ユニットの数、ユニットタイプ、手順の数及びタイプ、試験の数及びタイプ、ベッドの数、患者遭遇者の数、施設の数等を含むことができる。

【００２６】
他の健康プロバイダが臨床データのソースを推定することができないよう、臨床データを非正規化するため、健康プロバイダ１２は、健康プロバイダ１２により作成された臨床データを非正規化する変換エンジン２２を含む。このデータは、非正規化されたデータセットデータベース２４に格納される。特に、他の臨床データセットに対して見られるとき、各一意な属性が逆正規化され及び匿名であるよう、変換エンジン２２は臨床データを変換する。これを実現するため、遠隔計算リソース１４は、変換エンジン２２に送られ又はこれにより参照される臨床データの各一意な属性に関して、エントリの目標数を決定する。例えば、遠隔計算リソース１４は、各一意な属性に関するエントリの目標数を含む臨床データをどのように縮小させるかを変換エンジン２２に指示する変換フォーマットを、各健康プロバイダ１２に伝達する。臨床データにおいて各一意な属性に関する目標数がセットされる。ここで、健康プロバイダの匿名性が必要とされる。例えば、第１の健康プロバイダデータセットが６つの看護ユニットを含み、第２の健康プロバイダデータセットが２つの看護ユニットを含む場合、遠隔計算リソース１４は、看護ユニット属性に関するエントリの目標数をデータセットのエントリの最低の名数であるよう決定する。この場合、第１の健康プロバイダデータセットは、６つの看護ユニットエントリを３つの異なる２つの看護ユニットエントリに変換する。斯かる変換は、第１の健康プロバイダデータセットを第２の健康プロバイダデータセットから匿名にする。臨床データのソースが匿名であるよう、変換エンジン２２は、各一意な属性に関する目標数に基づき、臨床データを変換する。特に、変換エンジン２２は、匿名化される必要のある一意な属性、各一意な属性の目標数、及び遠隔計算リソース１４から受信されるデータキューブ規定に基づき、１つ又は複数のデータキューブへと臨床データを非正規化する。非正規化された臨床データが他の健康プロバイダ臨床データに対して効率的に比較されることができるよう、データキューブ規定は、遠隔計算リソース１４により利用される標準的なデータフォーマットである。別の実施形態では、変換エンジン２２は、非正規化されたデータにおける各属性を正規化されたデータフォーマットへマッピングするグローバルユニーク識別子（ＧＵＩＤ）の表を生成する。非正規化されたデータを再識別するため、再識別エンジン２６は、ＧＵＩＤの表を利用し、非正規化されたデータをそのオリジナルの正規化されたフォーマットにマッピングしなおす。健康プロバイダ１２は、臨床データを入力する、又は各一意な属性の目標数及び／若しくはデータキューブ規定を調整するためのユーザ入力デバイスも含む。いくつかの実施形態において、健康プロバイダ１２は、臨床データを手動で入力する、並びに／又は臨床データの生成されたレポート及び／若しくは分析を表示するユーザインタフェースをユーザに提供するディスプレイデバイスを含む。

【００２７】
例えば、表Ａに表されるシナリオにおいて、健康プロバイダＡは、さまざまなユニット（ＭＩＣＵ、ＳＩＣＵ、ＣＣＵ、ＮＩＣＯ、ＩＣＵ、ＣＶＩＣＵ、ＲＩＣＵ及びＰＩＣＵ）及びユニット当たり個別の数のベッドを持つ３つの病院（病院１、２及び３）を含む。

健康プロバイダＢは、３つのユニット（ＣＣＵ、ＩＣＵ及びＮＩＣＵ）及びユニット当たり個別の数のベッドを持つ単一の病院を含む。

【表１】

システム	病院	ユニット	ベッド
A	1	MICU	8
		SICU	8
		CCU	8
		NICU	22
	2	CCU	6
		ICU	6
	3	CVICU	16
		CCU	24
		ICU	16
		MICU	8
		RICU	8
		PICU	16
		NICU	60
B	1	CCU	12
		ICU	12
		NICU	24

表A

【００２８】

遠隔計算システム１４には、健康プロバイダ１２により読み出され、変換キューブ規定に加えて、ＧＵＩＤ置換に関する属性パラメータの目標数を含む構成が存在する。この例では、ユニットサイズは６であり、病院当たりのユニットのユニット数は、２である。この例では、ソースベッドを匿名化する必要がある。データがベッドに関連付けられずに、代わりに遭遇者に関連付けられる場合、遭遇者データは、新規な連合化されたユニット表現にランダムに割り当てられることができる。

【００２９】

変換エンジン２２は、１つの看護ユニットからのベッドを、６つのベッドのユニット及び２つの看護ユニットを持つ病院のシステムへとランダムに分けることにより、臨床データを非正規化する。前述のように、遠隔計算リソース１４は、各一意な属性に関するエントリの目標数を含む臨床データをどのように縮小させるかを変換エンジン２２に指示する変換フォーマットを、各健康プロバイダ１２に伝達する。これらのベッドに入れられる患者からの遭遇者データは、他の施設がソースを決定する方法を与えることなしに、臨床データとして表現されることができる。下記の表Ｂは、変換を理解することを簡単にするための変換識別の連続的な態様を表す。しかし、好ましい実施形態において、変換識別は、任意の識別可能な順にはなく、むしろ長い英数字ランダム性のＧＵＩＤを含む。以下は、結果として生じる変換を表す。

【表２】

システム	病院	ユニット	ベッド
A	1a	MICU1	6
	1a	MICU2	2
	1b	SICU1	6
	1b	SICU2	2
	1c	CCU1	6
	1c	CCU2	2
	1d	NICU1	6
	1d	NICU2	6
	1e	NICU3	6
	1e	NICU4	4
	2a	CCU3	6
	2a	ICU1	6
	3a	CVICU1	6
	3a	CVICU2	6
	3b	CVICU3	4
	3b	CCU4	6
	3c	CCU5	6
	3c	CCU6	6
	3d	CCU7	6
	3d	ICU2	6
	3e	ICU3	6
	3e	ICU4	4
	3f	MICU3	6
	3f	MICU4	2
	3g	RICU1	6
	3g	RICU2	2
	3h	PICU1	6
	3h	PICU2	6
	3i	PICU3	4
	3i	NICU1	6
	3j	NICU2	6
	3j	NICU3	6
	3k	NICU4	6
	3k	NICU5	6
	3l	NICU6	6
	3l	NICU7	6
	3m	NICU8	6
	3m	NICU9	6
	3n	NICU10	6
B	1aa	CCU8	6
	1aa	CCU9	6
	1ab	ICU5	6
	1ab	ICU6	6
	1ac	NICU11	6
	1ac	NICU12	6
	1ad	NICU13	6
	1ad	NICU14	6

表B

変換エンジン２２は、（この場合病院及びユニットにおいて）変更された記述を表すＧＵＩＤの表を作成し、ソース健康プロバイダにおいてデータを再識別するための基準テーブルを格納する。病院健康プロバイダＢに関して、表Ｃは、この例におけるＧＵＩＤの表を表す。

【表３】

システム	病院	ケアユニット	連合化された病院	連合化されたユニット	含まれる連合化されたベッド
B	1	CCU	1aa	CCU8	1, 3, 4, 7, 8, 9
			1aa	CCU9	2, 5, 6, 10, 11, 12
		ICU	1ab	ICU5	2, 3, 6, 9, 10, 11
			1ab	ICU6	1, 4, 5, 7, 8, 12
		NICU	1ac	NICU11	1, 6, 9, 11, 16, 18
			1ac	NICU12	3, 7, 8, 14, 15, 21
			1ad	NICU13	2, 4, 5, 17, 22, 24
			1ad	NICU14	10, 12, 13, 19, 20, 23

表C

【００３０】

この例では、健康プロバイダＢは、病院ＩＤ１ａａ～１ａｄが、それらの看護ユニットを表し、ＣＣＵ８及びＣＣＵ９が、特に健康プロバイダＢの１６のベッド看護ユニットであることを知る。しかし、データを見ても他の健康プロバイダは、ベッドＣＣＵのどれがどの病院に関連付けられるかを識別することができない。健康プロバイダ１２が、格納さ

れた連合化されたデータストアに基づきレポート及び／又は分析を要求するとき、健康プロバイダ１２は、どのソース機能が一団の結果に属するかを知ることなしに、他の類似する看護ユニット及び施設とそれらのパフォーマンスを比較することができる。健康プロバイダが、互いにベンチマークすることを選択をする場合、彼らは、パフォーマンス及び任意のベンチマークを見るために、識別性を共有することに同意する健康プロバイダと単にＧＵＩＤの表を共有することを必要とする。上記の例は、１つの病院を含む健康プロバイダ１２で臨床データの変換をもたらすよう容易に拡張されることができる。その結果、上の表は、４３の企業からのものであるあるように見える。各１つの病院は、６を超えないベッドを含む２つの看護ユニットを持つ。オペレーショナルリサーチを支援するため、追加的な記述がキューブ規定に加えられることができる。例えば、新規パラメータが、臨床データ表現に加えられることができる。このパラメータは、連合化されたユニットが、大きいユニット（即ち１６以上のベッド）又は小さいユニットかを示す。

【００３１】

健康プロバイダ１２はその後、分析及びレポーティングエンジン２８を介した追加的な分析及び／又はレポーティングのため、遠隔計算リソース１４に対して非正規化された臨床データを送信する。例えば、分析及びレポーティングエンジン２８は、例えばローカルパフォーマンスデータといった健康プロバイダの臨床データをベースライン臨床データと比較することができる。これらのすべては、遠隔計算リソース１４に格納される、又はこれによりアクセス可能とすることができる。分析は、分析及びレポーティングエンジン２８による１つ又は複数のレポートの生成を含むこともできる。このレポートは、パフォーマンスレポート、臨床勧告若しくはアドバイザリ、又は時系列的なグラフィックレポートを含むことができる。グラフィックレポートは例えば、明確で読みやすい表示フォーマットにおいて、健康及び不健康な結果を示すチャートを含む。ある例示的な実施形態では、結果及び／又はこの結果に対応するレポートデータが、追加的な処理のため健康システム、プロバイダ又は実際のデータのオーナー１２へ送信される。

【００３２】

別の実施形態では、１つ又は複数の健康プロバイダ１２は、正規化された臨床データを遠隔計算リソース１４に送信する。このリソースは、遠隔の正規化されたデータセットデータベース３０に正規化された臨床データを格納する。他の健康プロバイダが臨床データのソースを推定することができないよう、臨床データを非正規化するため、遠隔計算リソース１４は、遠隔の非正規化されたデータセットデータベース３２に格納される健康プロバイダ１２により作成される臨床データを非正規化する変換エンジン２２を含む。変換エンジン２２は、上述したように各一意な属性に関する目標数に基づき、臨床データを変換する。特に、変換エンジン２２は、匿名化されることを必要とする一意な属性、各一意な属性の目標数及び遠隔計算リソース１４のデータキューブ規定に基づき、１つ又は複数のデータキューブへと臨床データを非正規化する。変換エンジン２２は、正規化されたデータフォーマットへと非正規化されたデータにおける各属性をマッピングしなおすグローバルユニーク識別子（ＧＵＴＤ）の表も生成する。非正規化されたデータを再識別するため、再識別エンジン２６は、ＧＵＩＤの表を利用して、そのオリジナルの正規化されたフォーマットへと非正規化されたデータをマッピングしなおす。遠隔計算リソース１４は、臨床データを入力する、又は各一意な属性の目標数及び／若しくはデータキューブ規定を調整するためのユーザ入力デバイスも含む。いくつかの実施形態において、遠隔計算リソース１４は、臨床データを手動で入力する、並びに／又は臨床データの生成されたレポート及び／若しくは分析を表示するユーザインタフェースをユーザに提供するディスプレイデバイスを含む。

【００３３】

図２を参照すると、マルチサイトパフォーマンス測定を匿名にする方法のフローチャート図２００が示される。ステップ２０２において、１つ又は複数の健康プロバイダが遠隔計算リソースに接続される。ステップ２０４において、１つ又は複数の健康プロバイダは、遠隔計算リソースからデータキューブ規定を読み出す。ステップ２０６において、１つ

又は複数の健康プロバイダは、属性非正規化リスト及び各属性に関するパラメータの目標数を遠隔計算リソースから読み出す。ステップ２０８において、１つ又は複数の健康プロバイダは、データキューブ規定に基づきデータキューブを作成する。ステップ２１０において、１つ又は複数の健康プロバイダは、キューブ規定及び属性ランダム化リストに従って、ランダム化されたＧＵＩＤを作成することにより、データを変換する。ステップ２１２において、データの再識別に関するＧＵＩＤの表が、健康プロバイダ上で作成される。ステップ２１４において、変換されたデータが、遠隔計算リソースに送られる。ステップ２１６において、１つ又は複数の健康プロバイダは、変換スケジュールに従って、新たなキューブ規定をチェックする。

【００３４】

引き続き図１を参照し、臨床データを非正規化する遠隔計算リソース１４に対して１つ又は複数の健康プロバイダ１２が正規化された臨床データを送信するというシナリオにおいて、遠隔計算リソース１４は、非正規化されたデータに対するアクセス及び再識別を制御する。特に、遠隔計算リソースは、非正規化された臨床データにアクセスしてこれを再識別することを試みるとき、健康プロバイダ１２のユーザによりどんな処理が呼び出されることができるかに関する所定の役割を含む。これらの役割は、認証のため健康プロバイダのユーザ役割と同期化される。特に、健康プロバイダ１２ユーザは、健康プロバイダアプリケーションにログインする。それらの承認された役割に基づき、ユーザは、レポートリスト、ユニットリスト、患者レポート等を作成するといった、非正規化された臨床データに関する提供された処理を呼び出す。一旦ユーザが非正規化されたデータに基づき関心レポートを確立すると、ユーザは、非正規化された臨床データを再識別するため遠隔計算リソース１４を呼び出すことができる。この処理は、健康プロバイダ１２から遠隔計算リソース１４の再識別エンジン２６へと証明書が送信されることをもたらす。非正規化された臨床データは、ＧＵＩＤの表に基づき、その後再識別され、遠隔計算リソース１４は、遠隔計算リソース１４に格納される証明書と健康プロバイダ１２から受信した証明書とを整合させる。

【００３５】

特に、１つ又は複数の健康プロバイダ１２は、認証エンジン３４を含む。このエンジンは、ユーザに関してセットされる役割に基づき、遠隔計算リソース１４の分析及びレポーティングエンジン２８に呼び出される１つ又は複数の処理を選択するため、その特定のユーザが健康プロバイダアプリケーションにログインすることを可能にする。例えば、ユーザの承認された役割が医師である場合、分析及びレポーティングエンジン２８は患者レポートを作成する能力を彼らに提供する。同様に、ユーザの役割が管理者である場合、分析及びレポーティングエンジン２８は、このユーザが１つ又は複数の健康プロバイダのパフォーマンスに関するレポートを作成することを可能にする。非正規化されたデータに基づきユーザが関心処理を選択した後、ユーザは、それらの健康プロバイダ１２に関連付けられる非正規化された臨床データを再識別するためのオプションを与えられる。非正規化された臨床データを再識別することをユーザが選択する場合、認証エンジン３４は、遠隔計算リソース１４の再識別エンジン２２に証明書を送信する。遠隔計算リソース１４の再識別エンジン２２が、認証エンジン３４から受信される証明書が遠隔計算リソース１４に格納される証明書と整合すると決定する場合、再識別エンジン２２は、ＧＵＩＤの表を利用して、健康プロバイダに関連付けられる非正規化された臨床データをそのオリジナルの正規化されたフォーマットへとマッピングしなおす。上記の認証プロセスが、類似する態様で非正規化された臨床データに対するアクセスを制御するのに利用されることもできる点を理解されたい。

【００３６】

ＩＴインフラストラクチャ１０の要素は、前述の機能を実現するコンピュータ実行可能な命令を実行するプロセッサ４０を適切に含む。ここで、コンピュータ実行可能な命令は、プロセッサ４０に関連付けられるメモリ４２に格納される。しかしながら、前述の機能の少なくとも一部は、プロセッサを用いることなしに、ハードウェアにおいて実現される

ことができることも想定される。例えば、アナログ回路が使用されることができる。更に、ＩＴインフラストラクチャ１０の要素は、通信ネットワーク２０にわたり通信するためのインタフェースをプロセッサ４０に提供する通信ユニット４４を含む。更に、ＩＴインフラストラクチャ１０の前述の要素は別々に表されているが、これらの要素は組み合わせられることができる点を理解されたい。

【００３７】

図３を参照すると、健康プロバイダ構成の図３００が示される。ユーザは、健康プロバイダ表示アプリケーション３０４にログインする（ステップ３０２）。ユーザは、健康プロバイダのアクティブディレクトリ３０８を介して認証される（ステップ３０６）。健康プロバイダ表示アプリケーション３０４は、ユーザの役割を報告する（ステップ３１０）。非正規化されたデータに関して呼び出す処理が、ユーザにより選択され（ステップ３１２）、遠隔計算リソースに通信される（ステップ３１４）。ユーザの役割及び選択された処理が、追加的な処理のためローカルに格納される（ステップ３１６）。

【００３８】

図４を参照すると、健康プロバイダのアクセスシナリオの図４００が示される。ユーザは、健康プロバイダ表示アプリケーション４０４にログインする（ステップ４０２）。ユーザは、健康プロバイダのアクティブディレクトリ４０８を介して認証される（ステップ４０６）。健康プロバイダ表示アプリケーション４０４は、非正規化されたデータに関して呼び出す、ユーザにより選択される処理を探す（ステップ４１０）。ユーザに対して利用可能な非正規化されたデータに関して呼び出す処理のリストが、遠隔計算リソースから提供され（ステップ４１２）、健康プロバイダ表示アプリケーション４０４でユーザに表示される（ステップ４１４）。ユーザは、健康プロバイダ表示アプリケーション４０４上で処理をクリックする（ステップ４１６）。処理をクリックすることに基づき、セッショントークンが、再識別サービス４２０に送信される（ステップ４１８）。ユーザにより呼び出される処理４２２が、トークンと共に開始される（ステップ４２４）。呼び出された処理を開始することに基づき、トークンが再識別サービス４２０を介して有効にされる（ステップ４２６）。有効化が失敗する場合（ステップ４２８）、エラーページ４３０が表示される。有効化が成功する場合（ステップ４３２）、処理が呼び出され、データが正規化され、ユーザに表示される。

【００３９】

図５を参照すると、再識別ロジックの図５００が示される。呼び出す処理５０２を選択するとき、再識別命令５０４が、再識別サービスに送信される。トークンの有効化が失敗する場合（ステップ５０６）、エラーページ５０８が表示される。トークンの有効化が成功する場合（ステップ５１０）、処理が呼び出され、データが正規化され（ステップ５１２）、ユーザに表示される。

【００４０】

図６を参照すると、直接的なクラウドアクセスシナリオの図６００が示される。ユーザは、健康プロバイダマネージャ６０４にログインする（ステップ６０２）。ユーザは、クラウドアクティブディレクトリ６０８を介して認証される（ステップ６０６）。非正規化されたデータに関して呼び出す処理へのアクセスが、ユーザ６１０の役割に基づき制御される。この役割は、クラウドアクティブディレクトリ６０８から受信される（ステップ６１２）。ユーザの役割に基づき、非正規化されたデータの表示及びこれへのアクセス機能が制御される（ステップ６１４）。

【００４１】

図７を参照すると、匿名データの処理及び再識別を制御する方法のフローチャート図７００が示される。ステップ７０２において、非正規化されたデータが遠隔計算リソースに格納される。ステップ７０４において、ユーザは、健康プロバイダアプリケーションにログインする。ステップ７０６において、遠隔計算リソースは、ユーザの役割に基づき、非正規化されたデータに関して呼び出すユーザ処理を提供する。ステップ７０８において、健康プロバイダは、正規化されたデータを再識別するため、ユーザが処理を呼び出すこと

に基づき、遠隔計算リソースに証明書を送信する。ステップ７１０において、遠隔計算リソースは、遠隔計算リソースに格納される証明書と健康プロバイダから受信される証明書とが整合することに基づき、非正規化されたデータを再識別する。

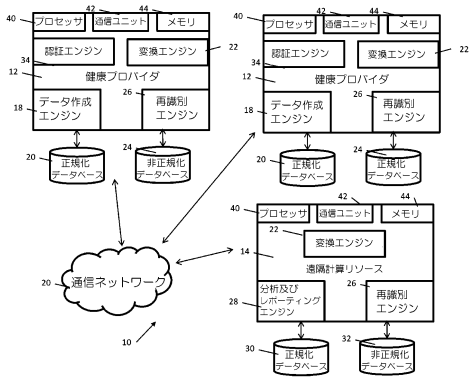
【００４２】

本書で使用されるメモリは、非一時的なコンピュータ可読媒体；磁気ディスク又は他の磁気ストレージ媒体；光学ディスク又は他の光学ストレージ媒体；ランダムアクセスメモリ（ＲＡＭ）、リードオンリーメモリ（ＲＯＭ）又は他の電子メモリデバイス又はチップ又は動作可能に相互接続されるチップのセット；格納された命令がインターネット／イントラネット又はローカルエリアネットワークを介して取得されることができるインターネット／イントラネットサーバ；等の１つ又は複数を含む。更に、本書で使用されるプロセッサは、マイクロプロセッサ、マイクロコントローラ、グラフィック処理ユニット（ＧＰＵ）、特定用途向け集積回路（ＡＳＩＣ）、フィールドプログラム可能ゲートアレイ（ＦＰＧＡ）等の１つ又は複数を含み；ユーザ入力デバイスは、マウス、キーボード、タッチスクリーンディスプレイ、１つ又は複数のボタン、１つ又は複数のスイッチ、１つ又は複数のトグル等の１つ又は複数を含み；ディスプレイデバイスは、液晶ディスプレイ、ＬＥＤディスプレイ、プラズマディスプレイ、プロジェクションディスプレイ、タッチスクリーンディスプレイ等の１つ又は複数を含む。

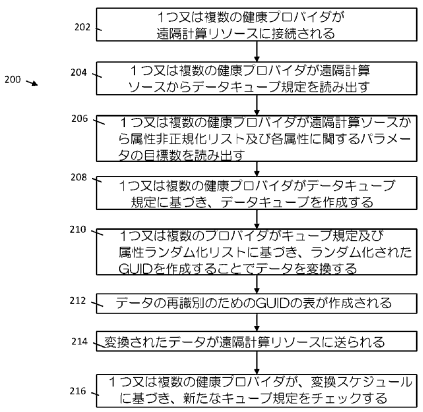
【００４３】

本発明が、好ましい実施形態を参照して説明されてきた。上記の詳細な説明を読み及び理解すると、第三者は、修正及び変更を思いつくことができる。それらの修正及び変更が添付の特許請求の範囲又はその均等物の範囲内にある限り、本発明は、すべての斯かる修正及び変更を含むものとして構築されることが意図される。

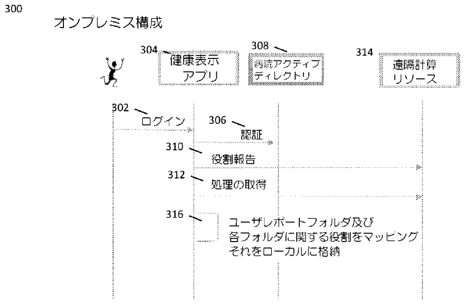
【図１】



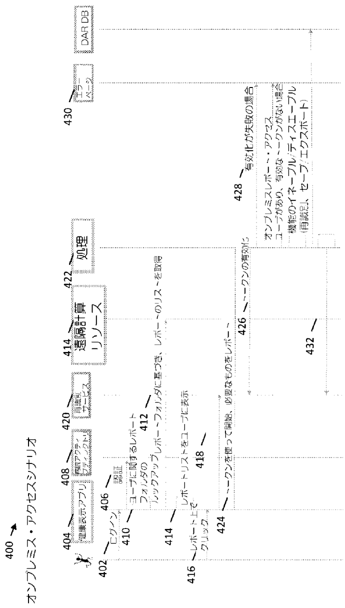
【図２】



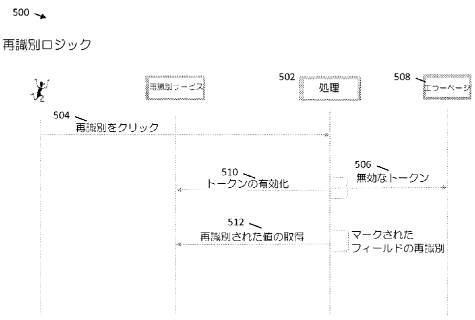
【図３】


300
オンプレミス構成
304 健康表示アプリ
308 持続アクティブディレクトリ
314 遠隔計算リソース
302 ログイン
306 認証
310 役割報告
312 処理の取得
316 ユーザレポートフォルダ及び各フォルダに関する役割をマッピングそれをローカルに格納

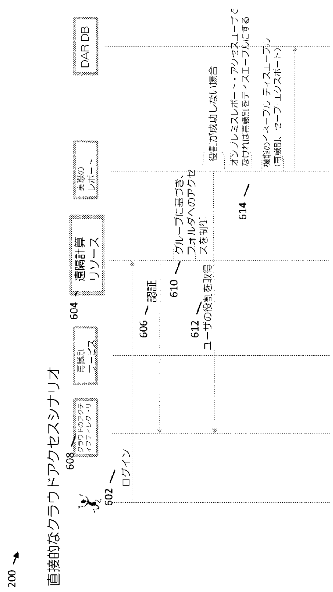
【図４】



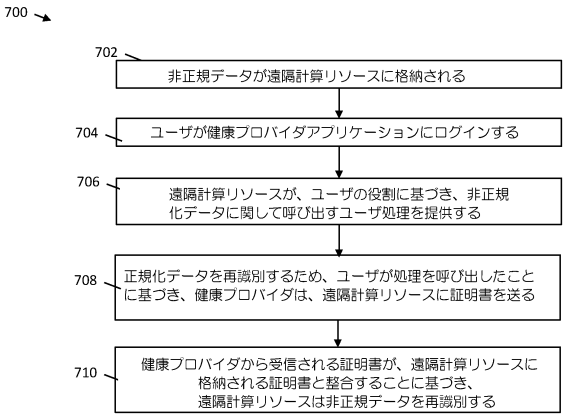
【図５】


500
再識別ロジック
502 処理
508 エラーページ
504 再識別をクリック
510 トークンの有効化
506 無効なトークン
512 再識別された値の取得
マークされたフィールドの再識別

【図６】



【図７】


700
702
非正規データが遠隔計算リソースに格納される
704
ユーザが健康プロバイダアプリケーションにログインする
706
遠隔計算リソースが、ユーザの役割に基づき、非正規化データに関して呼び出すユーザ処理を提供する
708
正規化データを再識別するため、ユーザが処理を呼び出したことに基づき、健康プロバイダは、遠隔計算リソースに証明書を送る
710
健康プロバイダから受信される証明書が、遠隔計算リソースに格納される証明書と整合することに基づき、遠隔計算リソースは非正規データを再識別する

フロントページの続き

(72)発明者　フロス　ブリアン　ダフィト
オランダ国　５６５６　アーエー　アインドーフェン　ハイ　テック　キャンパス　ビルディング　５
(72)発明者　エルド　イサック
オランダ国　５６５６　アーエー　アインドーフェン　ハイ　テック　キャンパス　ビルディング　５

審査官　青木　重徳

(56)参考文献　特表２０１１－５０１８３４（ＪＰ，Ａ）
国際公開第２０１２／１６５５１８（WO，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ　２１／６２
Ｇ０６Ｆ　２１／３１
Ｇ０６Ｆ　２１／６０
Ｇ０６Ｑ　５０／２２