

(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2014-199504
(P2014-199504A)

(43) 公開日 平成26年10月23日(2014.10.23)

(51) Int.Cl.		F I		テーマコード (参考)
G06F 12/00	(2006.01)	G06F 12/00	511Z	
G06Q 50/10	(2012.01)	G06Q 50/10	100	
G06F 17/30	(2006.01)	G06F 17/30	220Z	

審査請求 未請求 請求項の数 9 O L (全 34 頁)

(21) 出願番号 特願2013-73930 (P2013-73930)
(22) 出願日 平成25年3月29日 (2013.3.29)

(71) 出願人 000233491
株式会社日立システムズ
東京都品川区大崎一丁目2番1号
(74) 代理人 100091694
弁理士 中村 守
(72) 発明者 橘 正人
東京都品川区大崎1-2-1 株式会社日立システムズ内
(72) 発明者 津嶋 紀宏
東京都品川区大崎1-2-1 株式会社日立システムズ内

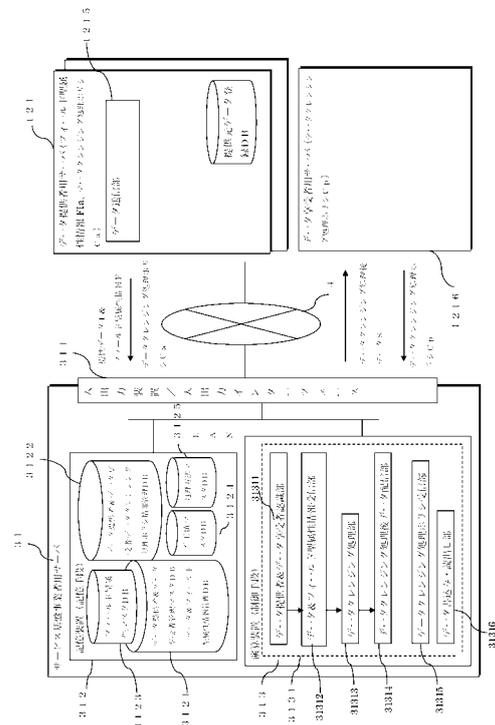
(54) 【発明の名称】 顧客別データクレンジング処理システム及び顧客別データクレンジング処理方法

(57) 【要約】

【課題】データ提供者側の数がN者、データ享受者側の数がM者のように数が増えるシステムにあっても、データクレンジング処理をシステム的に行えるデータクレンジング処理制御装置を備えたサービス提供システムを提供する。

【解決手段】基盤事業者用装置に、データ提供者用装置から提供される元データの不正値の型に対する処理ポリシーを含むデータクレンジング処理ポリシーマスタを元にデータクレンジング処理するデータクレンジング処理制御装置（データクレンジング処理サーバ）を有する。

【選択図】図2



【特許請求の範囲】**【請求項 1】**

少なくとも1つのサービス基盤事業者装置と、複数のデータ提供者装置と、複数のデータ享受者装置とを有したデータ提供システムであって、

前記サービス基盤事業者装置は、
入出力手段、記憶手段、制御手段、を有し、

前記入出力手段は、

前記データ提供者装置及び前記データ享受者装置との間でデータの授受を行い、

前記記憶手段は、前記データ提供者装置から提供されるデータに対するデータクレンジング処理ポリシーマスタを有し、

10

前記制御手段は、前記データクレンジング処理ポリシーマスタのデータクレンジング処理ポリシーをもって、前記データのデータクレンジング処理を行う機能を有し、

前記データ提供者装置から提供されるデータ、及び当該データのフィールド型属性情報を受信する受信部と、

前記データクレンジング処理ポリシーを受信するデータクレンジング処理ポリシー受信部と

、
前記データクレンジング処理ポリシーに従って、前記データ提供者装置から提供されるデータをデータクレンジング処理するデータクレンジング処理手段と、

前記データクレンジング処理手段にてデータクレンジング処理した後のデータクレンジング処理後データを前記データ享受者装置に送信する手段と、

20

を有することを特徴とする顧客別データクレンジング処理システム。

【請求項 2】

請求項 1 に記載された顧客別データクレンジング処理システムにおいて、

前記制御手段は、更に、

前記データ享受者装置のデフォルトデータクレンジング処理ポリシー、前記データ提供者装置から提供されるデータ用のデータ享受者データクレンジング処理ポリシー、及び前記データ提供者装置のデフォルトデータクレンジング処理ポリシー、前記データ提供者装置から提供されるデータ用のデータ提供者データクレンジング処理ポリシーを受信し、前記データクレンジング処理ポリシーマスタに反映する手段、

を有することを特徴とする顧客別データクレンジング処理システム。

30

【請求項 3】

請求項 2 に記載された顧客別データクレンジング処理システムにおいて、
前記制御手段は、更に、

フィールド型属性マスタ、不正值マスタ、処理方法マスタ、を有し、当該フィールド型属性マスタ、不正值マスタ、処理方法マスタ、から前記サービス基盤事業者装置のデータクレンジング処理ポリシーを作成する手段を有し、前記データクレンジング処理ポリシーマスタに反映する手段は、前記サービス基盤事業者装置のデータクレンジング処理ポリシーを前記データクレンジング処理ポリシーマスタに反映する

ことを特徴とする顧客別データクレンジング処理システム。

40

【請求項 4】

請求項 1 に記載された顧客別データクレンジング処理システムにおいて、

前記データクレンジング処理ポリシーマスタは、前記データ提供者装置から提供されるデータの元データフィールド名、当該フィールド名のフィールド型属性に生じうる各々の不正值に対応する処理方法を含み、

前記不正值は、「空白 (N U L L)」、「全角半角不統一」、「不正文字」、「異常値」、「ひらがな文字不統一」、「カナ文字不統一」、「英字不統一」、「数値不統一」、「記号不統一」、「漢字文字不統一」、「電話番号非正規化」、「郵便番号非正規化」、「住所非正規化」、「金額非正規化」、の1つ以上であり、

前記フィールドの型属性は、「主キー」、「名前」、「郵便番号」、「住所」、「数量」、「金額」、「日付」、の1つ以上であり、

50

前記処理方法は、「何もしない」、「レコード全てを削除」、「データ提供者フォーマットに合わせる」、「データ提供者正規化に合わせる」、「空白にする」、「平均値に置き換える」、の1つ以上である

ことを特徴とする顧客別データクレンジング処理システム。

【請求項5】

事業者のデータ提供者用装置から提供される元データを受け、当該元データをデータ享受者用装置に提供する基盤事業者用装置を備え、複数の事業者間でデータを共有するデータをデータクレンジング処理する顧客別データクレンジング処理システムにおいて、

前記基盤事業者装置は、

前記データ提供者用装置からの元データをデータクレンジング処理するデータクレンジング処理制御装置及び前記元データの各フィールドに生じうる不正値を定義する不正値マスタ、前記元データの各フィールドに生じうる不正値に対する処理方法を有するデータクレンジング処理ポリシマスタを有し、

前記データ提供者用装置は、

前記元データに対して、当該データの値が不正値の場合、当該不正値が前記不正値マスタのどの項目に対応するかを指定する手段を有し、

前記データ享受者用装置は、

前記元データに対して網羅的に準備したフィールド型属性型に対して、不正値に対する処理ポリシを設定する手段を有し、

前記データクレンジング処理制御装置は、

前記データ提供者用装置側にて設定した元データのフィールド型属性情報に対して、前記データ享受者用装置側にて設定した処理ポリシを前記データクレンジング処理ポリシマスタから読み取る読取手段と、当該読み取ったデータクレンジング処理ポリシに従って前記元データをデータクレンジング処理する処理する制御手段と、を有する

ことを特徴とする顧客別データクレンジング処理システム。

【請求項6】

少なくとも1つのサービス基盤事業者装置と、複数のデータ提供者装置と、複数のデータ享受者装置とを有したデータ提供システムにおける顧客別データクレンジング処理方法であって、

前記サービス基盤事業者装置は、

入出力手段、記憶手段、制御手段、を有し、

前記入出力手段は、

前記データ提供者装置及び前記データ享受者装置との間でデータの授受を行い、

前記記憶手段は、前記データ提供者装置から提供されるデータに対するデータクレンジング処理ポリシマスタを有し、

前記制御手段は、前記データクレンジング処理ポリシマスタのデータクレンジング処理ポリシをもって、前記データのデータクレンジング処理を行う機能を有し、

前記データ提供者装置から提供されるデータ、及び当該データのフィールド型属性情報を受信するステップと、

前記データクレンジング処理ポリシマスタのデータクレンジング処理ポリシを受信するステップと、

前記データクレンジング処理ポリシに従って、前記データ提供者装置から提供されるデータをデータクレンジング処理するステップと、

前記データクレンジング処理手段にてデータクレンジング処理した後のデータクレンジング処理後データを前記データ享受者装置に送信するステップと、

を有することを特徴とする顧客別データクレンジング処理方法。

【請求項7】

請求項6に記載された顧客別データクレンジング処理方法において、

前記制御手段は、更に、

前記データ享受者装置のデフォルトデータクレンジング処理ポリシ、前記データ提供者

装置から提供されるデータ用のデータ享受者データクレンジング処理ポリシー、及び前記データ提供者装置のデフォルトデータクレンジング処理ポリシー、前記データ提供者装置から提供されるデータ用のデータ提供者用データクレンジング処理ポリシーを受信し、前記データクレンジング処理ポリシーマスタに反映するステップ、

を有することを特徴とする顧客別データクレンジング処理方法。

【請求項 8】

請求項 7 に記載された顧客別データクレンジング処理システムにおいて、前記制御手段は、更に、

フィールド型属性マスタ、不正值マスタ、処理方法マスタ、を有し、当該フィールド型属性マスタ、不正值マスタ、処理方法マスタ、から前記サービス基盤事業者装置のデータクレンジング処理ポリシーを作成するステップを有し、前記データクレンジング処理ポリシーマスタに反映する手段は、前記サービス基盤事業者装置のデータクレンジング処理ポリシーを前記データクレンジング処理ポリシーマスタに反映するステップを有する

10

ことを特徴とする顧客別データクレンジング処理システム。

【請求項 9】

請求項 6 に記載された顧客別データクレンジング処理方法において、

前記データクレンジング処理ポリシーマスタは、前記データ提供者装置から提供されるデータの元データフィールド名、当該フィールド名のフィールド型属性に生じうる各々の不正值に対応するフィールドの型属性、当該不正值に対する処理方法を含み、

前記不正值は、「空白 (NULL)」、「全角半角不統一」、「不正文字」、「異常値」、「ひらがな文字不統一」、「カナ文字不統一」、「英字不統一」、「数値不統一」、「記号不統一」、「漢字文字不統一」、「電話番号非正規化」、「郵便番号非正規化」、「住所非正規化」、「金額非正規化」、の 1 つ以上であり、

20

前記フィールドの型属性は、「主キー」、「名前」、「郵便番号」、「住所」、「数量」、「金額」、「日付」、の 1 つ以上であり、

前記処理方法は、「何もしない」、「レコード全てを削除」、「データ提供者フォーマットに合わせる」、「データ提供者正規化に合わせる」、「空白にする」、「平均値に置き換える」、の 1 つ以上である

ことを特徴とする顧客別データクレンジング処理方法。

【発明の詳細な説明】

30

【技術分野】

【0001】

本発明は、顧客別データクレンジング処理システム及び顧客別データクレンジング処理方法に関する。

【0002】

更に、詳しくは、複数のデータ提供者、例えば複数のサービス提供事業者などから提供される各種データを収集し、例えば、当該データの集計、分析、データマイニングなどを実行した結果をデータ享受者側に提供することが可能なデータ提供システムに適用して好適な顧客別データクレンジング処理システム及び顧客別データクレンジング処理方法に関する。

40

【背景技術】

【0003】

インターネットの発展に伴い、企業の業種や業務内容を切り口にして、それぞれに有効なサービスを提供する各種サービス事業者、例えば飲食業界に対して人気メニューなどの商品情報を提供するサービス事業者や当該サービス事業者食材などを提供したりするサービス事業者や小売サービス事業者、食材などの生産事業者などが増加している。

【0004】

例えば、近年、インターネットや Web などを利用して、飲食店などのサービス享受事業者側に特定の業務サービスを提供するサービス提供事業者側に対して、その特定の業務サービスの基盤となるサービス基盤を提供し、該サービス基盤に基づく業務サービスをも

50

ってサービス享受事業者側のビジネスを支援するサービス提供システムが存在する。

【0005】

このようなサービス提供システムは、サービス基盤事業者装置、サービス提供事業者装置、サービス享受事業者装置、及びユーザ端末装置を有する。

【0006】

サービス基盤事業者装置は、提供するサービス基盤に基づく機能（サービス）、例えば SaaS（Software as a Service）型による業務システムをサービス提供事業者装置に提供するシステム又は装置である。

【0007】

サービス提供事業者装置は、サービス基盤事業者装置から提供されるサービス基盤に基づく業務システム（業務 SaaS）を受けるシステム又は装置である。また、例えば、当該業務システムをサービス享受事業者（或いは直接ユーザから）の要求に応じてカスタマイズしてサービス享受事業者に提供するシステム又は装置である。

【0008】

サービス享受事業者装置は、サービス提供事業者から提供されるサービスを受けるシステム又は装置である。また、例えばユーザから提供されるユーザ情報をサービス提供事業者装置側に提供する装置である。

【0009】

ユーザ端末装置は、例えば、自らユーザ情報を入力し、サービス享受事業者装置側に提供する端末である。また、このユーザ端末装置は、ユーザ自身の携帯端末をかざすことにより、当該携帯端末の情報を自動的に収集することができるようなものであってもよい。

そして、サービス享受事業者やサービス提供事業者が所有するオリジナルユーザ情報は、有益な情報を含んでいる。サービス基盤事業者は、この有益なオリジナルユーザ情報の提供を受け、当該情報の分析を可能とすれば、よりよいビジネス支援を行うことが期待できる。

【0010】

本技術分野の背景技術として、特開 2009 - 93569 号公報（特許文献 1）、特開平 9 - 153094 号公報（特許文献 2）、特開 2012 - 141847 号公報（特許文献 3）がある。

【0011】

特許文献 1 には、SaaS 環境から SI 環境へ変更する場合、SaaS 形態で利用を始めた業務サービスにおいて、埋没コストや機会損失の発生を軽減するものであり、業務サービスを SaaS 事業者環境（A 社）から SI 形態の環境（B 市）に移行する場合、B 市に業務サービスのアドレスを記録するサービスアドレスリストを設け、業務サービスの移行に合わせてサービスアドレスリストのアドレスを更新する。また、A 社に蓄積された業務データも移行する場合、A 社及び B 市に業務データのアドレスを記録するデータアドレスリストを設け、業務データの移行に合わせて双方のデータアドレスリストのアドレスを更新する。それらの更新においては、アドレスが論理的に同一になるようにアクセス方法を変更する技術が開示されている。

【0012】

特許文献 2 には、顧客に関する情報と、顧客が商品の購入を検討したときに参照した商品の情報と、顧客が購入した商品の情報とから商品の販売促進効果を分析する販売促進効果分析方法及びその実施システムが開示されている。

【0013】

特許文献 3 には、後続の各処理で用いる処理条件を定義する定義ファイルを作成し、記憶領域に格納する定義ファイル作成部と、記憶領域から読み出した定義ファイルを用いて、データクレンジング処理前又は後の移行元 DB のデータの妥当性のチェック処理を実施し、チェック処理結果を出力するチェック処理部と、記憶領域から読み出した定義ファイルを用いて、ロード後の移行先 DB に格納されたデータの妥当性を検証し、検証処理結果を出力する検証処理部とを備え、画面処理部は、表示装置にチェック処理結果および検証

10

20

30

40

50

処理結果を表示するデータ移行システムが記載されている。

【先行技術文献】

【特許文献】

【0014】

【特許文献1】特開2009-93569号公報

【特許文献2】特開平9-153094号公報

【特許文献3】特開2012-141847号公報

【発明の開示】

【発明が解決しようとする課題】

【0015】

特許文献1に記載されたようなサービス提供システムによれば、サービス基盤事業者は、サービス提供事業者の業務支援ができる。また、サービス提供事業者は、サービス享受事業者の業務支援ができる。そして、これらの業務支援により、システム全体の業務を効率化でき、業務効率を向上することができる。

【0016】

また、係るサービスを提供する各サービス提供事業者及びサービス基盤事業者側では、それぞれ各ユーザに関するデータ、例えば嗜好や行動履歴などのデータを保持（蓄積）、管理しているのが通常である。

【0017】

従って、このユーザに関するデータを分析し、当該分析結果に基づく有益な情報やオリジナルデータをサービス提供事業者からサービスの提供を受ける側のサービス享受事業者（店舗、一般企業など）に対して、提供できるシステムがあれば、より業務効率を向上することが期待できる。

【0018】

ここで、本発明における「ユーザ」とは、本システムを利用する個人、団体、企業であり、特定ユーザ（一の個人、一の団体、一の企業が特定可能なユーザ）と不特定ユーザ（個人、団体、企業が特定不可能又は特定困難なユーザ）を含む。ユーザが本システムで使用する端末等を「ユーザ側端末」と言う。以下、本発明において、ユーザに関して「個人」とした用語はユーザとしての団体及び企業を含むものである。

【0019】

また、「サービス享受事業者」とは、本システムによる本来事業支援サービス（第1のサービス）の提供を受ける事業体であり、当該サービス享受事業者側に設けられたシステム構成装置を「サービス享受事業者装置」と言う。ここでの「側に設ける」との用語の意義は、サービス享受事業者内に設けることのみを意味するものではない。この「側に設ける」との用語は、以下同様に用いる。

【0020】

また、「サービス提供事業者」とは、本システムによる本来事業支援サービス（第1のサービス）を提供する事業体であり、当該サービス提供事業者側に設けられたシステム構成装置を「サービス提供事業者装置」と言う。

【0021】

また、「サービス基盤事業者」とは、複数のサービス提供事業者を束ねる企業であって、サービスの基盤、例えば業務や予約機能（業務サービスや予約サービス）などをサービス提供事業者及び/又はサービス享受事業者に提供する事業体である。そして、契約に基づきサービス享受事業者装置及び/又はサービス提供事業者装置から、情報を随時収集し、当該情報を例えば分析して、サービス享受事業者側及び/又はサービス提供事業者側に分析情報（第2のサービス）として提供する。ここで、提供する情報は、分析情報に限る必要はなく、例えばデータ提供者からの要求に応じて、オリジナルデータをそのままの非分析情報を提供するものであってもよい。当該サービス基盤事業者側に設けられたシステム構成装置を「サービス基盤事業者装置」と言う。

【0022】

10

20

30

40

50

また、「データ提供事業者」、また「データ享受事業者」とは、サービス提供事業者やサービス享受事業者を指している。そして、例えばオリジナルユーザ情報 I a、I b をそのままの生データ、或いは k - 匿名性を満たす匿名化データ X I a、X I b としてサービス基盤事業者側に提供するデータとして送信し、また当該データをサービス基盤事業者側にて分析した分析データ、或いは生データとして受信する。当該データ提供事業者側に設けられたシステム構成装置を「データ提供事業者装置」と言い、当該データ享受事業者側に設けられたシステム構成装置を「データ享受者装置」と言う。

なお、データ提供者側として、「データ提供事業者」とし、データ享受者側として、「データ享受事業者」としているが、「事業者」に限らず、「ユーザ」個人であってもよい。従って、後述する実施例では、単に「データ提供者」、「データ享受者」と言う場合もある。

10

【0023】

また、サービス基盤事業者は、複数のサービス提供事業者を束ねる企業であることから、当該サービス基盤事業者側において、複数のサービス提供事業者側に蓄積している各種データを融合し、又は統合し、分析することができれば、より高付加価値な分析が可能となり、その結果として、サービスをサービス提供事業者、及び/又はサービス享受事業者を含むユーザに対して、より高付加価値な分析結果を提供することが可能となる。

【0024】

大容量データの分析が可能なIT環境が整いつつある現状にあっては、技術的には、大容量データから、上述した高付加価値な分析を行うことは可能である。

20

【0025】

「高付加価値な分析」とは、パーソナライゼーション（Personalization：得意客、及び見込みユーザごとの好みや属性に応じた情報を電子メールなどで発信すること）や併売リコメンド（recommendation：ユーザの好みを分析し、各ユーザごとに興味のある情報を選択して表示するサービスのこと）などを実現可能にする基礎データを提供するための分析を言う。

【0026】

上述したサービス提供者側及びサービス基盤事業者側において、サービス提供者側から各種データを収集し、例えば、収集したデータを集計・分析するとき、当該データを、集計、分析（解析）などに耐えうるようにするためにデータクレンジング（data cleansing）処理を行う必要がある。その理由は、データ入力者の入力ミスやデータ処理の不具合などによりデータの一部が不正値となることがあるためである。

30

【0027】

このデータクレンジング処理する場合、主に、以下の二通りの方法が一般的である。

第一は、サービス提供者（以下、データ提供者と言う）用装置10側から提供するデータ（元データ）を、当該データ提供者用装置側にて、データクレンジング処理を行い、当該データクレンジング処理済みのデータ（データクレンジング処理後データ）をデータ享受者（以下、データ享受者と言う）用装置20側に提供する方法である。

この方法の場合は、データクレンジング処理ポリシーをデータ提供者側のポリシーに従うか、もしくはデータ享受者側のポリシーに従うかは、2社間で取り決めておくことが要求される。

40

【0028】

第二は、データ提供者用装置10側から提供される元データを受け取るデータ享受者用装置側にて、当該元データのデータクレンジング処理を行う方法である。

この方法の場合、（1）データ享受者用装置20側は、元データが正常値か不正値かの判断が難しい。また、（2）データ提供者用装置10側のデータクレンジング処理ポリシーに従うか、もしくはデータ享受者用装置20側のデータクレンジング処理ポリシーに従うかの判断も必要であり、これら（1）及び（2）は、データ提供者側とデータ享受者側の2社間の取り決めておくことが要求される。

【0029】

50

因みに、元データの値が正常値であるか、または不正値（データクレンジング処理すべき値）であるかを、一番熟知しているのは、元データの所有者であるデータ提供者側である。

【0030】

一方、データ提供者側が、その元データをデータ享受者側に提供する場合、提供データを利用するのはデータ享受者側であり、データクレンジング処理は、データを利用する側であるデータ享受者側のデータクレンジング処理ポリシーに従う場合が好適である。

【0031】

このように、データ提供者側、又はデータ供給者側でデータクレンジング処理を行う方法は、データ提供者側とデータ享受者側が1対1の場合は、2社間の取り決めで容易に実現可能である。

【0032】

しかし、データクレンジング処理ポリシーが相違する複数のデータ提供者及び複数のデータ享受者を対象とするシステム、例えば、データ提供者側の数がN者、データ享受者側の数がM者のように数が増えるシステムへの適用については考慮されていない。

【0033】

つまり、当該システムへの適用の場合、例えば、データ提供者側の数がN者、データ享受者側の数がM者のように数が増えるシステムにあっては、システム構成が複雑となり、以下のような課題がある。

【0034】

N者いるデータ提供者側でデータクレンジング処理すべき値を処理する場合、M者いるデータ享受者側の各々のデータクレンジング処理ポリシーに従ってデータクレンジング処理をすることになり、データ提供者側は、N者に対して、各々M回データクレンジング処理を行わなければならない。

【0035】

逆に、M者いるデータ享受者側でデータクレンジング処理をする場合、N者いるデータ提供者側から元データを取得して、同時にN者いるデータ提供者側からデータクレンジング処理すべき値（不正値）を設定されている場合、M者いるデータ享受者側は、各々N回データクレンジング処理をしなければならない。

【0036】

即ち、データ提供者側がN者、データ享受者側がM者いる場合、効率的なデータクレンジング処理する装置が必要である、という課題があった。

【0037】

また、前記特許文献1には、クライアントからの指示に基づき、移行元DBをデータクレンジング処理するデータクレンジング処理部を有し、ユーザが、移行先システムのデータの妥当性を判断することが可能なデータ移行システムが記載されている。しかし、複数のデータ提供者から複数のデータを受信、収集し、複数のデータ享受者にデータクレンジング処理したデータを提供する場合において、データクレンジング処理をシステム的に行う、ことまでは考慮されていない。特許文献2, 3も同様である。

【0038】

そこで、本発明は、データ提供者側の数がN者、データ享受者側の数がM者のように数が増えるシステムにあって、データクレンジング処理をシステム的に行える顧客別データクレンジング処理システム及び顧客別データクレンジング処理方法を提供する。

【0039】

上記課題を解決するために、本発明は、基盤事業者用装置側にデータクレンジング処理ポリシーマスタ及び当該マスタのデータ提供者用装置から提供される元データの不正値の型に対するデータクレンジング処理ポリシーを元にデータクレンジング処理する顧客別データクレンジング処理手段（データクレンジング処理サーバ）を有し、当該データクレンジング処理ポリシーをデータ享受者用装置から設定可能な構成とする。

【0040】

10

20

30

40

50

例えば、本発明の顧客別データクレンジング処理システムは、
少なくとも1つのサービス基盤事業者装置と、複数のデータ提供者装置と、複数のデータ享受者装置とを有したデータ提供システムであって、

前記サービス基盤事業者装置は、
入出力手段、記憶手段、制御手段、を有し、
前記入出力手段は、
前記データ提供者装置及び前記データ享受者装置との間でデータの授受を行い、
前記記憶手段は、前記データ提供者装置から提供されるデータにデータクレンジング処理するデータクレンジング処理ポリシーマスタを有し、

前記制御手段は、前記データクレンジング処理ポリシーマスタのデータクレンジング処理ポリシーをもって、前記データのデータクレンジング処理を行う機能を有し、

前記データ提供者装置から提供されるデータ、及び当該データのフィールド型属性情報を受信する受信部と、

前記データクレンジング処理ポリシーを受信するデータクレンジング処理ポリシー受信部と

、
前記データクレンジング処理ポリシーに従って、前記データ提供者装置から提供されるデータをデータクレンジング処理するデータクレンジング処理手段と、

前記データクレンジング処理手段にてデータクレンジング処理した後のデータクレンジング処理後データを前記データ享受者装置に送信する手段と、

を有することを特徴とする。

【0041】

前記制御手段は、更に、

前記データ享受者装置のデフォルトデータクレンジング処理ポリシー、前記データ提供者装置から提供されるデータ用のデータ享受者データクレンジング処理ポリシー、及び前記データ提供者装置のデフォルトデータクレンジング処理ポリシー、前記データ提供者装置から提供されるデータ用のデータ提供者データクレンジング処理ポリシーを受信し、前記データクレンジング処理ポリシーマスタに反映する手段、

を有することを特徴とする。

【0042】

前記制御手段は、更に、

フィールド型属性マスタ、不正值マスタ、処理方法マスタ、を有し、当該フィールド型属性マスタ、不正值マスタ、処理方法マスタ、から前記サービス基盤事業者装置のデータクレンジング処理ポリシーを作成する手段を有し、前記データクレンジング処理ポリシーマスタに反映する手段は、前記サービス基盤事業者装置のデータクレンジング処理ポリシーを前記データクレンジング処理ポリシーマスタに反映する

ことを特徴とする。

【0043】

前記データクレンジング処理ポリシーマスタは、前記データ提供者装置から提供されるデータの元データフィールド名、当該フィールド名のフィールド型属性に生じうる各々の不正值に対応する処理方法を含み、

前記不正值は、「空白(NULL)」、「全角半角不統一」、「不正文字」、「異常値」、「ひらがな文字不統一」、「カナ文字不統一」、「英字不統一」、「数値不統一」、「記号不統一」、「漢字文字不統一」、「電話番号非正規化」、「郵便番号非正規化」、「住所非正規化」、「金額非正規化」、の1つ以上であり、

前記フィールドの型属性は、「主キー」、「名前」、「郵便番号」、「住所」、「数量」、「金額」、「日付」、の1つ以上であり、

前記処理方法は、「何もしない」、「レコード全てを削除」、「データ提供者フォーマットに合わせる」、「データ提供者正規化に合わせる」、「空白にする」、「平均値に置き換える」、の1つ以上である

ことを特徴とする。

【 0 0 4 4 】

また、本発明の顧客別データクレンジング処理システムは、
事業者のデータ提供者用装置から提供される元データを受け、当該元データをデータ享受者用装置に提供する基盤事業者用装置を備え、複数の事業者間で共有するデータをデータクレンジング処理する顧客別データクレンジング処理システムにおいて、

前記基盤事業者装置は、

前記データ提供者用装置からの元データをデータクレンジング処理するデータクレンジング処理制御装置及び前記元データの各フィールドに生じうる不正値を定義する不正値マスタ、前記元データの各フィールドに生じうる不正値に対する処理方法を有するデータクレンジング処理ポリシマスタを有し、

前記データ提供者用装置は、

前記元データに対して、当該データの値が不正値の場合、当該不正値が前記不正値マスタのどの項目に対応するかを指定する手段を有し、

前記データ享受者用装置は、

前記元データに対して網羅的に準備したフィールド型属性型に対して、不正値に対する処理ポリシーを設定する手段を有し、

前記データクレンジング処理制御装置は、

前記データ提供者用装置側にて設定した元データのフィールド型属性情報に対して、前記データ享受者用装置側にて設定した処理ポリシーを前記データクレンジング処理ポリシマスタから読み取る読取手段と、当該読み取ったデータクレンジング処理ポリシーに従って前記元データをデータクレンジング処理する制御手段と、を有する

ことを特徴とする。

【 0 0 4 5 】

また、本発明の顧客別データクレンジング処理方法は、

少なくとも1つのサービス基盤事業者装置と、複数のデータ提供者装置と、複数のデータ享受者装置とを有したデータ提供システムにおける顧客別データクレンジング処理方法であって、

前記サービス基盤事業者装置は、

入出力手段、記憶手段、制御手段、を有し、

前記入出力手段は、

前記データ提供者装置及び前記データ享受者装置との間でデータの授受を行い、

前記記憶手段は、前記データ提供者装置から提供されるデータにデータクレンジング処理するデータクレンジング処理ポリシマスタを有し、

前記制御手段は、前記データクレンジング処理ポリシマスタのデータクレンジング処理ポリシーをもって、前記データのデータクレンジング処理を行う機能を有し、

前記データ提供者装置から提供されるデータ、及び当該データのフィールド型属性情報を受信するステップと、

前記データクレンジング処理ポリシマスタのデータクレンジング処理ポリシーを受信するステップと、

前記データクレンジング処理ポリシーに従って、前記データ提供者装置から提供されるデータをデータクレンジング処理するステップと、

前記データクレンジング処理手段にてデータクレンジング処理した後のデータクレンジング処理後データを前記データ享受者装置に送信するステップと、

を有することを特徴とする。

【 0 0 4 6 】

前記制御手段は、更に、

前記データ享受者装置のデフォルトデータクレンジング処理ポリシー、前記データ提供者装置から提供されるデータ用のデータ享受者データクレンジング処理ポリシー、及び前記データ提供者装置のデフォルトデータクレンジング処理ポリシー、前記データ提供者装置から提供されるデータ用のデータ提供者データクレンジング処理ポリシーを受信し、前記データ

10

20

30

40

50

クレンジング処理ポリシマスタに反映するステップ、
を有することを特徴とする。

【0047】

前記制御手段は、更に、

フィールド型属性マスタ、不正值マスタ、処理方法マスタ、を有し、当該フィールド型属性マスタ、不正值マスタ、処理方法マスタ、から前記サービス基盤事業者装置のデータクレンジング処理ポリシを作成するステップを有し、前記データクレンジング処理ポリシマスタに反映する手段は、前記サービス基盤事業者装置のデータクレンジング処理ポリシを前記データクレンジング処理ポリシマスタに反映するステップを有することを特徴とする。

10

【0048】

前記データクレンジング処理ポリシマスタは、前記データ提供者装置から提供されるデータの元データフィールド名、当該フィールド名のフィールド型属性に生じうる各々の不正值に対応するフィールドの型属性、当該不正值に対する処理方法を含み、

前記不正值は、「空白(NULL)」、「全角半角不統一」、「不正文字」、「異常値」、「ひらがな文字不統一」、「カナ文字不統一」、「英字不統一」、「数値不統一」、「記号不統一」、「漢字文字不統一」、「電話番号非正規化」、「郵便番号非正規化」、「住所非正規化」、「金額非正規化」、の1つ以上であり、

前記フィールドの型属性は、「主キー」、「名前」、「郵便番号」、「住所」、「数量」、「金額」、「日付」、の1つ以上であり、

20

前記処理方法は、「何もしない」、「レコード全てを削除」、「データ提供者フォーマットに合わせる」、「データ提供者正規化に合わせる」、「空白にする」、「平均値に置き換える」、の1つ以上である

ことを特徴とする。

【発明の効果】

【0049】

本発明によれば、データ提供者側の数がN者、データ享受者側の数がM者のように数が増えるシステムであっても、データクレンジング処理をシステム的に行える顧客別データクレンジング処理システム及び顧客別データクレンジング処理方法を提供することができる。

30

【0050】

また、データクレンジング処理を基盤事業者側の装置に集約して、「データクレンジング処理ポリシマスタ」を保持することにより、データ提供者側は一度「元データ」の「不正值」を指定し、また、データ享受者側は、一度「不正值」に対する「データクレンジング処理ポリシ」を設定すればよい。また、データ提供者用装置や基盤事業者用装置側にて指定されたデータクレンジング処理ポリシに対しても変更したり、追加することができる。

【0051】

上述した以外の課題、構成及び効果は、以下の実施形態の説明により明らかにされる。

【図面の簡単な説明】

40

【0052】

【図1】本発明の顧客別データクレンジング処理システムを採り入れたサービス提供システムの全体概要を示す図である。

【図2】図1のサービス基盤事業者用サーバの構成及びデータ提供事業者用サーバとデータ享受者用サーバとの関係を示すブロック図である。

【図3】図2のサービス基盤事業者用サーバの演算装置(制御手段)の構成例を示すブロック図である。

【図4】データ提供者用サーバから提供されるデータ(元データ)の一例を示す図である。

【図5】データ提供者用サーバから提供されるデータ(Ia)の、サービス基盤事業者用

50

サーバにて準備してある元データを示すマスタテーブルの例である。

【図 6】サービス基盤事業者用サーバにおけるデータ提供者（A社）の情報示すマスタテーブルの例である。

【図 7】データ提供者用サーバから提供されるデータの、サービス基盤事業者用サーバにて準備してある元データフィールド型属性情報示すマスタテーブルの例である。

【図 8】データ提供者用サーバから提供されるデータのフィールド型属性を示すマスタテーブルの例である。

【図 9】サービス基盤事業者用サーバにて網羅的に準備してある不正値を示すマスタテーブルの例である。

【図 10】サービス基盤事業者用サーバにて網羅的に準備してある処理方法を示すマスタテーブルの例である。

10

【図 11】サービス享受者用サーバにて、データクレンジング処理ポリシーを設定するデータクレンジング処理ポリシー設定画面の例を示す図である。

【図 12】データ提供者（A社）のデータ（Ia）デフォルトデータクレンジング処理ポリシーマスタの例を示す図である。

【図 13】サービス基盤事業者用サーバにおけるデータ享受者（P社）マスタの例を示す図である。

【図 14】サービス基盤事業者用サーバにおけるデータクレンジング処理ポリシーマスタの例を示す図である。

【図 15】データ提供者（A社およびB社）側から提供されるデータ（IaおよびIb）に対して、サービス基盤事業者用サーバ側にて、データクレンジング処理し、当該データクレンジング処理したデータ（データクレンジング処理後データ）をデータ享受者用サーバに送信するフローチャートを示す図である。

20

【図 16】データ提供者（A社）側から基盤事業者側に対して、データ（Ia）、当該データのデフォルトデータクレンジング処理ポリシー（CaIa）、データ提供者デフォルトデータクレンジング処理ポリシー（Ca）を送信する様子を模式的に示す図である。

【図 17】データ享受者（P社）側から基盤事業者側に対して、データ享受者（P社）側のデフォルトデータクレンジング処理ポリシー（Cp）、データ提供者側のデータ（Ia）用データ享受者（P社）データクレンジング処理ポリシー（CpIa）、を送信し、基盤事業者デフォルトデータクレンジング処理ポリシーCgおよび当該データのデータ提供者データクレンジング処理ポリシー（CaIa）およびデータ提供者デフォルトデータクレンジング処理ポリシー（Ca）の少なくとも1つを使用してデータクレンジング処理を行う様子を模式的に示す図である。

30

【図 18】データ享受者用サーバにおけるデータクレンジング処理ポリシー設定画面を表示生成するためのフローチャートを示す図である。

【図 19】図 16における処理フローチャートを示す図である。

【図 20】図 17における処理フローチャートを示す図である。

【図 21】図 20におけるデータ、レコード、フィールドに関する処理フローチャートを示す図である。

【図 22】処理方法に従った一処理フローチャートを示す図である。

40

【図 23】処理方法に従った他の処理フローチャートを示す図である。

【図 24】処理方法に従った他の処理フローチャートを示す図である。

【図 25】処理方法に従った他の処理フローチャートを示す図である。

【図 26】処理方法に従った他の処理フローチャートを示す図である。

【発明を実施するための形態】

【0053】

以下、その実施態様について図面を参照して説明する。まず、サービス享受事業者のビジネスを支援するサービス提供システムの概略について説明する。

サービス提供システムは、アプリケーションサービスやS I、つまり予約機能や業務SaaSなどのサービス（第1のサービスS1）をサービス提供事業者側に提供するサービ

50

ス基盤事業者システム又は装置（以下、サービス基盤事業者装置と称する）と、サービス享受事業者（飲食店など）の要望に応じて第1のサービスをカスタマイズしてサービス享受事業者に提供するサービス提供事業者システム又は装置（以下、サービス提供事業者装置と称する）と、から構成される。

【0054】

サービス享受事業者（飲食店など）は、ユーザ（消費者）に飲食や商材などを提供し、またユーザからユーザ情報（オリジナル情報）の提供を受ける。

【0055】

サービス基盤事業者装置は、SaaS以外、例えばOSを含むプラットフォームを貸し出しするPaaS（Platform as a Service）、ASP（Application Service Provider）サービス等を提供する装置であってもよい。

10

【0056】

以下、飲食業界（飲食店）と、そのサービス提供事業者と、サービス基盤事業者とを主として説明する。

まず、サービス提供システムの概要について説明する。サービス提供システムのサービス基盤事業者装置は、サービス提供事業者装置に対して予約管理サービス（アプリケーションサービス）などを提供するアプリケーションサーバ（以下、サービス基盤事業者用サーバと言う）を有する。

【0057】

サービス提供事業者装置は、この予約管理サービスをサービス享受事業者の要望に沿ってカスタマイズし、サービス享受事業者からユーザ（消費者）の予約情報を受付け管理するアプリケーションサーバ（以下、サービス提供事業者用サーバと言う）を有する。つまり、サービス基盤事業者装置とサービス提供事業者装置との協働をもって、サービス享受事業者のビジネスを支援する仕組みである。

20

【0058】

このような仕組みのシステムは、一般的にサービス享受事業者ごとの個別案件ベースで対応するものである。

従って、サービス基盤事業者装置は、サービス提供事業者装置に対して、第1のサービスS1を提供し、サービス提供事業者装置はサービス享受事業者に対して、カスタマイズした第1のサービス（サービス開始）を提供し、また該提供した第1のサービスに基づくユーザ情報（オリジナル情報）Ia、Ibなどを管理するのみであった。

30

【0059】

このサービス提供方法としては、例えばSaaSが挙げられる。SaaSとは、情報システム（ソフトウェア）が提供する機能（第1のサービス）を、サービス提供事業者やサービス享受事業者が、特に新たなシステムを構築・導入することなく、ネットワーク経由で利用することを可能にするICT活用環境である。

【0060】

一般的には、SaaSでは、必要なときに必要なサービスを、特別なシステム構築の必要もなく素早く利用できるというメリットがある。また、SaaSでは、社内に特別な環境を用意する必要はないことから、その機能（サービス）の運用開始時間を早め、コストも低減できる。また、インターネットやWANを使用することになるが、ネットワークやブラウザの技術革新により、社内システムを利用するのとほぼ同等の操作が可能である。

40

【0061】

また、サービス提供事業者やサービス享受事業者にとっては、サービス基盤事業者から提供されるSaaSやSI及び運用・保守を利用することにより、コスト削減、売上拡大を図ることが期待できる。

【0062】

本発明は、斯様なデータ提供事業者の装置（サービス提供事業者装置A、B、サービス享受事業者装置a1、a2、b1、b2など）からのデータをデータマイニングした結果（データ）を、当該データ享受者ごとにデータクレンジング処理し、データクレンジング

50

処理後データとして当該データ享受者の装置側に提供することを可能とするものである。

【0063】

以下、本発明の実施例について図面に基づき説明する。本実施例は、データクレンジング処理が不揃いのデータであっても、受け入れ、例えば、当該受け入れた複数のデータを融合や統合して効率的に分析することを可能とし、また分析データ、或いはオリジナルデータ(元データ)をデータ享受者装置の設定に応じたデータクレンジング処理し、データ享受者側からの配信要求に応じて送信するものである。

【0064】

本実施例では、

(1) データ提供者の数がN者、データ享受者の数がM者いる場合、基盤事業者を通してデータの享受を行う。 10

(2) 基盤事業者のサーバ上に、データ提供者の元データに対して、データクレンジング処理する「データクレンジング処理手段」を搭載する。つまり、基盤事業者を中心に本発明のシステムを実装する。

(3) また、基盤事業者のサーバ上に「データクレンジング処理ポリシマスタ」(図14)を持たせる。

(4) 「データクレンジング処理ポリシマスタ」(図14)は、デフォルト値として、データ提供者が設定する「デフォルトデータクレンジング処理マスタ」(図12)の値が設定される

(5) データ提供者は、「元データ」(図4)のフィールド毎にフィールドの「属性」を「元データフィールド型属性情報マスタ」(図7)に登録する。「元データ」は「元データソースマスタ」(図2)で管理する 20

(6) データ提供者は、「元データフィールド型属性情報マスタ」(図7)の中の「フィールド型ID」は「フィールド型属性マスタ」(図8)から選択する。

(7) 「フィールド型属性マスタ」(図8)は、網羅的にフィールドの型を準備しておく。もし「フィールド型属性マスタ」に対応する型が無い場合は、データ提供者は「フィールド型属性マスタ」にフィールドの型を追加する。

(8) 「不正値」は、「不正値マスタ(図5)」で管理する。「不正値マスタ(図5)」は、網羅的に不正値を準備しておく。もし「不正値マスタ(図5)」に対応する不正値が無い場合は、データ提供者は「不正値マスタ」に不正値を追加する。 30

(9) 不正値の「処理方法」は、「処理方法マスタ」(図10)で管理する。「処理方法マスタ」は、網羅的に処理方法を準備しておく。もし「処理方法マスタ」に対応する処理方法が無い場合は、データ提供者は「処理方法マスタ」(図10)に処理方法を追加する。

(10) データ提供者は、「フィールド型属性」に対する「処理方法」を「データクレンジング処理ポリシ設定画面」(図11)で指定して「デフォルトデータクレンジング処理マスタ」(図12)の各ID(値)を設定する。データ提供者IDはデフォルト値として「0」とする。

(11) 「データクレンジング処理ポリシ設定画面」(図11)は、同図に示すように、「不正値」(「空白」、「全角半角不統一」、「不正文字」、「非フォーマット」、・・・)に対するフィールドの型属性(「主キー」、「名前」、「郵便番号」、「住所」、「数量」、「金額」、・・・)の値として「処理ID」を設定する。 40

(12) データ享受者は、「データクレンジング処理ポリシ設定画面」(図11)をもって「データクレンジング処理ポリシ」を設定できるようにする。設定された「データクレンジング処理ポリシ」は、「データクレンジング処理ポリシマスタ」(図14)に反映する。ここで、データ享受者は「データ享受者マスタ」(図13)で管理される。

(13) 「データクレンジング処理ポリシマスタ」(図14)の値は、一度データ享受者により、データの「不正値」に対する処理ポリシ、つまり「処理方法」を設定する。このようにデータ提供者の元データの「不正値」に対して処理ポリシを設定することにより、結果として、データ享受者は一回の処理ポリシ設定で、N者のデータ提供者の元データに 50

対するデータクレンジング処理ポリシーを設定することができる。

(14) データ提供者およびデータ享受者は、データ提供者の特定のデータに対するデータクレンジング処理ポリシーを設定できる。特定のデータに対するデータクレンジング処理ポリシーが設定されている場合には、デフォルトデータクレンジング処理ポリシーよりも優先して使用される。

(15) 基盤事業者のデータクレンジング処理サーバは、データ提供者側にて指定した「不正値」(図12の不正値ID)に対して、データ享受者の指定した「処理方法」を「データクレンジング処理ポリシーマスタ」から読み取って、処理プログラムを実行する。

(16) 基盤事業者は、「不正値マスタ」(図9)と「処理方法マスタ」(図10)に対応する処理を実現するプログラムを基盤事業者側のサーバ上に持つ。

以下、その詳細な実施例について説明する。

【実施例1】

【0065】

図1は、本発明の顧客別データクレンジング処理システムをサービス基盤事業者装置に適用したサービス提供システムの全体構成を示す図である。

【0066】

サービス提供システムは、データ提供事業者のデータ提供事業者装置12、サービス基盤事業者のサービス基盤事業者装置3、を有する。

【0067】

データ提供事業者装置12は、複数のサービス提供事業者装置(A、B・・・)を含む複数のサービス提供事業者装置2、複数のサービス享受事業者装置(a、b・・・)を含むサービス享受事業者装置1、を有する。

【0068】

データ提供事業者装置12は、データ提供事業者用サーバ121からなる。

各サービス享受事業者装置1は、例えば複数の店舗(a1、a2)や複数の企業(b1、b2)側の装置であって、サービス享受事業者用サーバ11からなる。そして、各サービス享受事業者用サーバ11は、オリジナルユーザ情報、つまり元データ(IaやIb)をサービス提供事業者装置2に対して出力する。

【0069】

このオリジナルユーザ情報IaやIbは、サービス基盤事業者装置3からネットワーク4を経由して提供する第1のサービスS1(業務/予約機能など)に対応してユーザ側から送られてくるものである。

【0070】

各サービス提供事業者装置2は、サービス提供事業者用サーバ21からなる。サービス提供事業者用サーバ21は、各元データI(Ia、Ib)を受け、当該元データを直接、又はその出所元、つまりユーザが特定できない程度に周知の匿名化技術を利用して匿名化データ(含提供データ=元データI(Ia)、当該データのフィールド型属性情報F(FIa)、及びデータクレンジング処理ポリシーC(Ca))としてネットワーク4を介してサービス基盤事業者装置3側に送信する。

【0071】

サービス基盤事業者装置3は、サービス基盤事業者用サーバ31からなる。サービス基盤事業者用サーバ31は、前記データ提供事業者(A社)用サーバ121から提供されるデータI(Ia)、フィールド型属性情報F(FIa)及びデータクレンジング処理ポリシーC(Ca)を受ける。

【0072】

また、サービス基盤事業者装置3は、データ享受者(P社)用サーバ(図2の1216参照)からデータクレンジング処理ポリシーC(Cp)を受ける。

そして、当該データクレンジング処理ポリシーに応じて、データ提供者事業者用サーバから提供される元データをデータクレンジング処理し、データクレンジング処理後データSとして出力する機能を有する。

10

20

30

40

50

【 0 0 7 3 】

図 2 は、サービス基盤事業者用サーバ 3 1 及びデータ提供事業者用サーバ 1 2 1、データ享受者用サーバ 1 2 1 6 の構成を示すブロック図である。

【 0 0 7 4 】

本例では、図 1 のサービス提供事業者用サーバ 2 1 をデータ提供者用サーバ 1 2 1 とし、サービス享受事業者用サーバ 1 1 をデータ享受者用サーバ 1 2 1 6 とし、当該データ提供者用サーバ 1 2 1 からサービス基盤事業者用サーバ 3 1 に提供するデータをオリジナルデータ（元データ）I（I a、I b、・・・）を前提として説明する。また、データ提供者用サーバ 1 2 1 からのデータをデータクレンジング処理し、当該データクレンジング処理後データ S をデータ享受者用サーバ 1 2 1 6 に送信することを前提として説明する。

10

なお、ここで、元データに対して、所望の分析を必要に応じて行ってもよい。また、この分析は、データクレンジング処理後データを元にデータ享受者側にて行なってもよいが、本発明の本質ではないので、その説明は省略する。

【 0 0 7 5 】

データ提供事業者用サーバ 1 2 1 は、元データ I、当該データのフィールド型属性を示す情報 F 及びデータクレンジング処理ポリシー C、などをサービス基盤事業者用サーバ 3 1 に提供するものであるが、サービス基盤事業者用サーバ 3 1 からデータクレンジング処理後データ S を享受するデータ享受者用サーバとして利用してもよい。またデータ享受者用サーバ 1 2 1 6 は、データクレンジング処理後データ S を享受するに際して、データ享受者側のデータクレンジング処理ポリシー C p を提供するものであるが、サービス基盤事業者用サーバに対して、データを提供しデータ提供事業者用サーバとして利用してもよい。

20

【 0 0 7 6 】

データ享受者用サーバ 1 2 1 6 が、サービス基盤事業者用サーバ 3 1 からデータクレンジング処理後データ S を享受するには、サービス基盤事業者用サーバ 3 1 に対して、その旨（データ享受要求信号）を要求すると共にデータのデータクレンジング処理ポリシー C（C p）をサービス基盤事業者用サーバ 3 1 側に通知する。このデータクレンジング処理ポリシー C（C p）は、予めサービス基盤事業者用サーバ 3 1 側に設定登録しておいてもよい。

【 0 0 7 7 】

サービス基盤事業者用サーバ 3 1 は、入出力装置（入出力インターフェース）3 1 1、記憶装置（記憶手段）3 1 2、演算装置（演算手段）3 1 3、を有する。これらは LAN により接続する。

30

【 0 0 7 8 】

記憶装置 3 1 2 は、データ提供者及びデータ享受者管理マスタ DB / データ及びフィールド型属性情報管理 DB 3 1 2 1、データ提供者・データ享受者・データクレンジング処理ポリシー情報管理 DB 3 1 2 2、フィールド型属性マスタ DB 3 1 2 3、不正値マスタ DB 3 1 2 4、処理方法マスタ DB 3 1 2 5、を有する。

【 0 0 7 9 】

データ提供者及びデータ享受者管理マスタ DB / データ及びフィールド型属性情報管理 DB 3 1 2 1 は、データ提供者及びデータ享受者のデータやフィールド型属性情報を管理するものであり、またデータ提供者用サーバ 1 2 1 から提供されるデータ及び当該データのフィールド型属性情報をそれぞれ対応付けて記憶し、管理する。

40

【 0 0 8 0 】

データ提供者・データ享受者・データクレンジング処理ポリシー情報管理 DB 3 1 2 2 は、データ提供者やデータ享受者側にて設定されるデータクレンジング処理ポリシー（「データ提供者 ID / データ享受者 ID」、「元データソース ID」、「元データフィールド ID」、「元データフィールド名」、「不正値」、「処理方法 ID」、など）を記憶する。

【 0 0 8 1 】

換言すれば、データ提供者用サーバ 1 2 1 からの提供データ I をサービス基盤事業者用サーバ 3 1 にて、データクレンジング処理するデータクレンジング処理ポリシーマスタを管

50

理する。

【0082】

フィールド型属性マスタDB3123は、「フィールド型属性ID」、「属性」、つまり「データ番号」、「購入日」、「店舗名」、「顧客名」、「商品名」、「金額」、「購入数」、「日付」などを記憶する。

【0083】

不正値マスタDB3124は、「不正値ID」、「不正値の型」、つまり「空白」、「全角半角不統一」、「不正文字」、「非フォーマット」、などを記録する。

【0084】

処理方法マスタDB3125は、「処理ID」、「処理」、つまり「何もしない」、「レコード全てを削除」、「空白にする」、「平均値に置き換える」、「データ享受者データクレンジング処理ポリシーに従う」、「データ享受者のデフォルトデータクレンジング処理ポリシーに従う」、「データ提供者データクレンジング処理ポリシーに従う」、「基盤事業者データクレンジング処理ポリシーに従う」、「データ提供者デフォルトデータクレンジング処理ポリシーに従う」、などを記憶する。

10

これらのDBの内容については後述する。

【0085】

演算装置313は、元データIをデータクレンジング処理する制御手段3131を有する。

制御手段3131は、データ提供者及びデータ享受者認識部31311、データ及びフィールド型属性情報受信部31312、データクレンジング処理部31313、データクレンジング処理後データ配信部31314、データクレンジング処理ポリシー受信部31315、データ書込み・読出し部31316、などを有する。

20

【0086】

データ提供者及びデータ享受者認識部31311は、データ提供者及びデータ享受者を認識する。

【0087】

データ及びフィールド型属性情報受信部31312は、データ提供者用サーバ121からのデータI及びフィールド型属性情報Fを受信する。

【0088】

データクレンジング処理ポリシー受信部31315は、データ提供者用サーバ121およびデータ享受者用サーバ1216から提供される元データのデータクレンジング処理ポリシー情報C(Ca、Cp)を受信する。

30

【0089】

データクレンジング処理部31313は、データ提供者用サーバ121から提供されるデータを、データクレンジング処理ポリシー情報管理DB3122のデータクレンジング処理ポリシーに従って、データクレンジング処理する。

【0090】

このデータクレンジング処理は、例えば、サービス基盤事業者用サーバ31にて、データ提供者用サーバ121のデータを分析できるようにするためである。

これにより、データ享受者用サーバ1216は、一度、データクレンジング処理して欲しいデータクレンジング処理ポリシーを設定するだけで、サービス基盤事業者用サーバ31からデータクレンジング処理後データを享受することができる。

40

【0091】

データ書込み・読出し部31316は、上述したデータ提供者用サーバの各データI(Ia、Ib、...)や各データのデータクレンジング処理ポリシーC(Ca、Cp)をDB3121に書き込み、当該DBから読み出しする。

【0092】

データクレンジング処理後データ配信部31315は、データクレンジング処理部31313によりデータクレンジング処理したデータをデータ享受者用サーバ1216に送信

50

する。

【 0 0 9 3 】

図 3 は、サービス基盤事業者用サーバ 3 1 の一例を示す機能ブロック図である。同図において、サーバ 3 1 は、提供データ I 及び当該データのフィールド型属性情報 F をデータ及びフィールド型属性情報受信部 3 1 3 1 2 (図 2 参照) にて受信し、データに対応付けてデータ及びフィールド型属性情報管理 DB 3 1 2 1 に格納する。

【 0 0 9 4 】

提供データ I は、データ提供者用サーバ 1 2 1 から提供されるデータフォーマットのままでデータ及びフィールド型属性情報管理 DB 3 1 2 1 に格納し、その後、データクレンジング処理部 3 1 3 1 3 により、当該データをデータクレンジング処理する。

10

【 0 0 9 5 】

また、データクレンジング処理部 3 1 3 1 3 により、データクレンジング処理したデータクレンジング処理後データ S をデータ及びフィールド型属性情報管理 DB 3 1 2 1 に格納してもよい。

【 0 0 9 6 】

当該データのフィールド型属性情報は、提供データ I に対応付けてデータ及びフィールド型属性情報管理 DB 3 1 2 1 に格納する。

【 0 0 9 7 】

以下、各 DB のテーブル内容について説明する。

図 4 は、データ提供者用サーバから提供されるデータ (元データ I) の一例を示す図であって、図 5 の元データマスタのデータ提供者名「かかか商店 株式会社」の元データ名「売上ジャーナルテーブル」を示すテーブルである。

20

【 0 0 9 8 】

同図において、元データには、元データ主キー (「 1 」 、 「 2 」 、 「 3 」 、 「 4 」 、 ・ ・ ・) 、名前 (「 田 郎 」 、 「 木 郎 」 、 「 立 郎 」 、 「 J フェリー 」 、 ・ ・ ・) 、郵便番号 (「 1 1 1 - 1 1 1 1 」 、 「 1 1 1 1 1 2 」 、 「 1 1 1 1 2 2 」 、 「 1 1 1 - 2 2 2 2 」 、 ・ ・ ・) 、住所 (「 札幌市旭区 」 、 「 八戸市港町 1 - 1 」 、 「 秋田県秋田市 」 、 「 盛岡市川辺 4 丁目 」 、 ・ ・ ・) 、商品名 (「 x x x x 」 、 「 y y y y 」 、 「 z z z z 」 、 「 A b c d e 」 、 ・ ・ ・) 、金額 (「 1 0 0 0 」 、 「 2 0 , 0 0 0 」 、 「 3 0 , 0 0 0 」 、 「 4 0 0 0 0 」 、 ・ ・ ・) 、などの各情報 4 0 1 ~ 4 0 6 を含んでいる。

30

【 0 0 9 9 】

図 5 は、データ提供者用サーバから提供されるデータ (元データ I a) の、サービス基盤事業者用サーバにて準備してある元データマスタを示すテーブルである。

【 0 1 0 0 】

同図において、元データソースマスタは、元データ ID (「 1 」 、 「 2 」 、 「 3 」 、 「 4 」 、 ・ ・ ・) 、元データ名 (「 売上ジャーナルテーブル 」 、 「 予約ジャーナルテーブル 」 、 「 時間帯別売上テーブル 」 、 「 予約ジャーナルテーブル 」 、 ・ ・ ・) 、データ提供者 ID (「 1 」 、 「 2 」 、 「 3 」 、 「 3 」 、 ・ ・ ・) 、データ提供者名 (「 かかか商店 株式会社 」 、 「 きききホテル 株式会社 」 、 「 けけけ美容院 株式会社 」 、 「 けけけ美容院 株式会社 」 、 ・ ・ ・) 、などの各情報 5 0 1 ~ 5 0 4 を含んでいる。

40

【 0 1 0 1 】

図 6 は、サービス基盤事業者用サーバにおけるデータ提供者 (A 社) マスタを示すテーブルである。

【 0 1 0 2 】

同図において、データ提供者マスタは、データ提供者 ID (「 1 」 、 「 2 」 、 「 3 」 、 ・ ・ ・) 、データ提供者名 (「 株式会社 さささ製作所 」 、 「 有限会社 ししし製麺 」 、 「 株式会社 すすす商事 」 、 ・ ・ ・) 、などの各情報 6 0 1 ~ 6 0 2 を含んでいる。

【 0 1 0 3 】

図 7 は、データ提供者用サーバから提供されるデータの、サービス基盤事業者用サーバにて準備してある元データフィールド型属性マスタを示すテーブルである。

50

【0104】

同図において、元データフィールド型属性マスタは、フィールドID（「1」、「2」、「3」、「4」、「5」、「6」、「7」、・・・）、元データID（「1」、「1」、「1」、「1」、「1」、「1」、・・・）フィールド名（「データ番号」、「購入日」、「店舗名」、「顧客名」、「商品名」、「購入数」、「金額」、・・・）、フィールド型属性ID（「1」、「8」、「3」、「3」、「3」、「7」、「6」、・・・）、などの各情報701～704を含んでいる。

【0105】

図8は、データ提供者用サーバから提供されるデータのフィールド型属性マスタを示すテーブルである。

10

【0106】

同図において、フィールド型属性マスタは、フィールド型属性ID（「1」、「2」、「3」、「4」、「5」、「6」、「7」、「8」、・・・）、属性（「データ番号」、「購入日」、「店舗名」、「顧客名」、「商品名」、「金額」、「購入数」、「日付」、・・・）などの各情報801～802を含んでいる。

【0107】

図9サービス基盤事業者用サーバにて網羅的に準備してある不正値マスタを示すテーブルである。

同図において、不正値マスタは、不正値ID（「1」、「2」、「3」、「4」、・・・）、不正値の型（「空白」、「全角半角不統一」、「不正文字」、「非正規化」、・・・）などの各情報901～902を含んでいる。

20

【0108】

図10サービス基盤事業者用サーバにて網羅的に準備してある処理方法マスタを示すテーブルである。

【0109】

同図において、処理方法マスタは、処理方法ID（「1」、「2」、「3」、「4」、「5」、「6」、「7」、「8」、「9」、・・・）、処理（「何もしない」、「レコード全てを削除」、「空白にする」、「平均値に置き換える」、「データ享受者データクレンジング処理ポリシーに従う P社データクレンジング処理ポリシー Cp Ia」、「データ享受者のデフォルトデータクレンジング処理ポリシーに従う P社デフォルトデータクレンジング処理ポリシー Cp」、「データ提供者のデータクレンジング処理ポリシーに従う A社元データ Ia 用 A社データクレンジング処理ポリシー Ca Ia」、「基盤事業者データクレンジング処理ポリシーに従う 基盤事業者デフォルトデータクレンジング処理ポリシー Cq」、「データ提供者のデフォルトデータクレンジング処理ポリシーに従う A社デフォルトデータクレンジング処理ポリシー Ca」、・・・）、などの各情報1001～1002を含んでいる。

30

【0110】

図11はサービス享受者用サーバにて、データクレンジング処理ポリシーを設定するデータクレンジング処理ポリシー設定画面を示す図である。

【0111】

同図において、データクレンジング処理ポリシー設定画面には、横方向にフィールド名703（図7参照）の「データ番号」7031、「購入日」7032、「店舗名」7033、「顧客名」7034、「商品名」7035、「購入日」7036、「金額」7037、及び当該フィールド名のフィールド型属性の「主キー」7041'、「日付」7042'、「名前」7043'、「名前」7044'、「名前」7045'、「数量」7046'、「金額」7047'などが表示される。

40

また、縦方向には、「不正値」（図9参照）の「空白」、「全角半角不統一」、「不正文字」、「非フォーマット」、などが表示される。

そして、「不正値」に対応する「フィールド型属性」は、セレクトボックスをもって、処理方法 図10参照 の処理ID（1～9、・・・）を設定する。

50

【 0 1 1 2 】

図 1 2 は、データ提供者（A 社）のデータ（I a）デフォルトデータクレンジング処理ポリシーマスタの例を示す図である。

【 0 1 1 3 】

同図において、デフォルトデータクレンジング処理ポリシーマスタは、データ提供者 ID（「1」、「1」、「1」、「1」、・・・）、元データソース ID（「1」、「1」、「1」、「1」、・・・）、元データフィールド ID（「1」・・・、「2」・・・、「3」・・・、「4」・・・、・・・）、元データフィールド名（「元データキー」・・・、「名前」・・・、「郵便番号」・・・、「住所」、・・・）、不正値 ID（「1」、「2」、「3」、「4」、「1」、「2」、「3」、「4」、「1」、「2」、「3」、「4」・・・）、処理方法 ID（「2」、「2」、「2」、「2」、「1」、「1」、「3」、「3」、「1」、「5」、「3」、「3」、「1」、「5」、「5」、「3」）、などの各情報 1 2 0 1 ~ 1 2 0 6 を含んでいる。

10

【 0 1 1 4 】

図 1 3 は、サービス基盤事業者用サーバにおけるデータ享受者（P 社）マスタの例を示す図である。

【 0 1 1 5 】

同図において、データ享受者マスタは、データ享受者 ID（「1」、「2」、「3」、・・・）、データ享受者名（「株式会社 さささ製作所」、「有限会社 ししし製麺」、「株式会社 すすす商事」、・・・）、などの各情報 1 3 0 1 ~ 1 3 0 2 を含んでいる。

20

【 0 1 1 6 】

図 1 4 サービス基盤事業者用サーバにおけるデータクレンジング処理ポリシーマスタの例を示す図である。

【 0 1 1 7 】

同図において、データクレンジング処理ポリシーマスタは、データ享受者 ID（「1」、「1」、「1」、「1」、・・・）、元データソース ID（「1」、「1」、「1」、「1」、・・・）、元データフィールド ID（「1」・・・、「2」・・・、「3」・・・、「4」・・・、・・・）、元データフィールド名（「元データキー」・・・、「名前」・・・、「郵便番号」・・・、「住所」、・・・）、不正値 ID（「1」、「2」、「3」、「4」、「1」、「2」、「3」、「4」、「1」、「2」、「3」、「4」・・・）、処理方法 ID（「2」、「2」、「2」、「2」、「1」、「1」、「3」、「3」、「1」、「5」、「3」、「3」、「1」、「5」、「5」、「3」）、などの各情報 1 4 0 1 ~ 1 4 0 6 を含んでいる。

30

【 0 1 1 8 】

次に、顧客別データクレンジング処理方法の処理手順について説明する。

図 1 5 は、本発明の顧客別データクレンジング処理方法の処理フローを説明する図である。本例では、データ提供者用サーバ 1 2 1 が 2 つの場合を示し、データ享受者用サーバ 1 2 1 6 が 1 つの場合を示している。また、各者のデフォルトデータクレンジング処理ポリシーは設定済みとし、かつフィールド型属性マスタ、不正値マスタ、処理方法マスタがデータ提供者用サーバおよびデータ享受者用サーバに共有されているものとする。図において以下の処理を実行する。

40

(1) データ提供者（A 社）は、上述した 3 つのマスタを元にデータ I a 用フィールド型属性情報 F I a、およびデータ I a 用 A 社データクレンジング処理ポリシー C a I a を作成する。

(2) そして、データ提供者（A 社）は、基盤事業者に対して、データ I a & フィールド型属性情報 F I a、およびデータ I a 用 A 社データクレンジング処理ポリシー C a I a を送信する。

(3) データ提供者（B 社）も、上記 (2)、(3) と同様な処理を行う。

(4) データ享受者（P 社）は、基盤事業者に対して、データ I a のフィールド型属性情報 F I a、およびデータ I a 用 A 社データクレンジング処理ポリシー C a I a を要求する。

50

(5) 基盤事業者は、この要求に応答し、データ享受者(P社)に対して、データ提供者A社のデータIaのフィールド型属性情報FIa、およびデータIa用A社データクレンジング処理ポリシーCaIaを送信する。

(6) データ享受者(P社)は、FIa、CaIaを受信し、これらを参照して、データIa用データクレンジング処理ポリシーCpIaを作成し、基盤事業者に送信する。

(7) また、データ享受者(P社)は、基盤事業者に対して、データクレンジング処理後データを要求する。

(8) 基盤事業者は、この要求を元にデータクレンジング処理を実行し、データクレンジング処理後データをサービス享受者(P社)に送信する。

(9) サービス享受事業者(P社)は、このデータクレンジング処理後データを受信する。

10

【0119】

データ提供者(A社)用サーバ121は、ステップS1211にて、データIa及び当該データのフィールド型属性FIaをサービス基盤事業者用サーバ31に送信する。

【0120】

また、別のデータ提供者(B社)用サーバ121は、同様にステップS1212にて、データIb及び当該データのフィールド型属性FIbをサービス基盤事業者用サーバ31に送信する。

【0121】

一方、データ享受者(P社)用サーバ1216は、ステップS12161にて、サービス基盤事業者用サーバ31に対して、データクレンジング処理したデータ(データクレンジング処理後データ)の享受を要求すると共にデータクレンジング処理ポリシーCpを送信する。

20

【0122】

サービス基盤事業者用サーバ31は、ステップS311にて、これらのデータIa、Ib及びフィールド型属性情報FIa、FIbを受信処理し、データ及びフィールド型属性情報管理DB3121に書き込み、格納する。

また、ステップS315にて、データ享受者(P社)用サーバ1216からの要求に応答し、データ享受者(P社)に対して、データ提供者A社のデータIaのフィールド型属性情報FIa、およびデータIa用A社データクレンジング処理ポリシーCaIaを送信する。

30

【0123】

また、ステップS312にて、データ要求及びデータクレンジング処理ポリシーCpを受信し、データ提供者及びデータ享受者データクレンジング処理ポリシー情報管理DB3122に書き込み、格納する。

【0124】

次に、ステップS313にて、データクレンジング処理ポリシー情報管理DB3122のデータクレンジング処理ポリシーCpに従った処理方法を元にデータクレンジング処理を行う。このデータクレンジング処理ポリシーCpに従ったデータクレンジング処理は、元データ毎に行う。

40

【0125】

また、このデータクレンジング処理は、必要に応じて、データ提供者側にて設定したデフォルトデータクレンジング処理ポリシーCa、データ提供者データ用デフォルトデータクレンジング処理ポリシーCaIa、又は基盤事業者側にて設定した基盤事業者データクレンジング処理ポリシーCg、或いはデータ提供者にて設定したデータ提供者用デフォルトデータクレンジング処理ポリシーCaIaに対して、データ享受者側にて一部変更設定した享受者データクレンジング処理ポリシーCpIaにて行ってもよい。

【0126】

しかるのち、ステップS314にて、データ提供者用サーバ121からの提供データ(元データ)Iaをデータクレンジング処理したデータクレンジング処理後データをデータ

50

享受者用サーバ121に対して送信する。

【0127】

このとき、図示していないが、必要に応じて、データ提供者用サーバ121からのデータのデータフォーマットをサービス基盤事業者用サーバ31のデータクレンジング処理ポリシーに従ったデータフォーマットに変換処理し、統一の基盤データフォーマットのデータを所望の分析手法（データマイニング）により、データ分析処理する。あるいは、基盤データフォーマットのまま、あるいはデータ提供者用サーバ121のデータフォーマットのままであってもよい。これらは、データ享受者用サーバ1216からのデータフォーマット形式要求に応じてなされる。

【0128】

データ享受者用サーバ1216は、ステップS12162にて、データクレンジング処理結果データを受信する。

【0129】

これにより、データ享受者側のデータクレンジング処理ポリシーを反映し、当該データ享受者側の希望するデータクレンジング処理データを楽しむことができる。

【0130】

図16は、データ提供者（A社）側から基盤事業者側に対して提供する元データ、フィールド型属性およびデータクレンジング処理ポリシーを送信する流れを模式的に示した図である。

【0131】

同図において、データ提供者（A社）側は、まず、基盤事業者側の記憶装置からフィールド型属性マスタ（図8参照）、不正値マスタ（図9参照）、処理方法マスタ（図10参照）、を取り込む。そして、これらを元にA社デフォルトデータクレンジング処理ポリシーCaを作成し、基盤事業者側に送信する。

【0132】

また、A社から提供するA社元データIaを基盤事業者側に送信する。このとき、A社元データIaと基盤事業者側から取り込んだフィールド型属性マスタ（図8参照）、不正値マスタ（図9参照）、処理方法マスタ（図10参照）を元にA社元データIaのフィールド型属性FIaおよびA社データIa用デフォルトデータクレンジング処理ポリシーCaIa、を作成し、基盤事業者側に送信する。

【0133】

基盤事業者側は、A社デフォルトデータクレンジング処理ポリシーCa、A社元データIa、A社元データIaのフィールド型属性FIa（図8参照）およびA社データIa用デフォルトデータクレンジング処理ポリシーCaIaを受信し、これらを記憶装置のデータクレンジング処理ポリシー情報管理DB31323に格納する（図12参照）。

【0134】

図17は、データ享受者（P社）側から基盤事業者に対してデータクレンジング処理ポリシーを送信し、また基盤事業者側からデータ享受者側に対して、データクレンジング処理されたデータを送信する流れを模式的に示した図である。

【0135】

同図において、データ享受者（P社）側は、まず、基盤事業者側の記憶装置からフィールド型属性マスタ（図8参照）、不正値マスタ（図9参照）、処理方法マスタ（図10参照）を取り込む。そして、これらを元にP社デフォルトデータクレンジング処理ポリシーCpを作成し、基盤事業者側に送信する。

【0136】

また、データ提供者（A社）側で作成したA社元データIa、A社元データIaのフィールド型属性FIa、A社元データIa用A社データクレンジング処理ポリシーCaIa、を基盤事業者側の記憶装置から取り込む。

【0137】

そして、A社元データIaに対して、A社データIa用デフォルトデータクレンジング

10

20

30

40

50

処理ポリシー C a I a に変更を加えたい場合、つまりデータ享受者側の意志による処理ポリシーを設定する場合には、データ享受者（P社）側から所望のデータクレンジング処理ポリシー C p I a を設定し、基盤事業者側に送信する。例えば、基盤事業者側から取り込んだ A 社元データ I a のフィールド型属性情報 F I a 及び A 社元データ I a 用 A 社データクレンジング処理ポリシーを参照し、A 社元データ I a 用 P 社データクレンジング処理ポリシーを作成し、当該作成ポリシーを基盤事業者側のデータクレンジング処理マスタに反映し、当該ポリシーをもってデータクレンジング処理を行えるようにする。

【0138】

このデータクレンジング処理ポリシー設定は、例えば、図10に示すようなデータクレンジング処理設定画面をもって行なう。

10

【0139】

基盤事業者側は、データ享受者（P社）側からのP社デフォルトデータクレンジング処理ポリシー C p、A 社元データ I a 用 P 社データクレンジング処理ポリシー C p I a を受信する。そして、これらの処理ポリシーを元に、又は A 社元 I a 用 A 社データクレンジング処理ポリシー C a I a、基盤事業者デフォルトデータクレンジング処理ポリシー C g、A 社デフォルトデータクレンジング処理ポリシー C a をもって、A 社元データ I a に対して、データクレンジング処理を行う。

【0140】

しかるのち、所望のフォーマット変換処理、所望の分析を行い、その結果データをデータ享受者（P社）側に送信する。

20

【0141】

図18は、データ享受者側と基盤事業者側との間におけるデータクレンジング処理ポリシー設定画面の表示生成するためのフローチャートである。同図において、データ享受者側は、ステップ S 1 1 1 1 にて、基盤事業者側の「不正値マスタ」から不整値一覧を取得し、ステップ S 1 1 1 2 にて、不正値を「データクレンジング処理ポリシー設定画面」（図11参照）の縦軸に表示する。

【0142】

また、ステップ S 1 1 1 3 にて、「フィールド型属性マスタ」（図8参照）からフィールド型属性一覧を取得し、ステップ S 1 1 1 4 にて、フィールド型属性の値を「データクレンジング処理ポリシー設定画面」の横軸に表示する。

30

【0143】

ステップ S 3 1 1 5 にて、「処理方法マスタ」から処理方法一覧を取得し、ステップ S 1 1 1 6 にて、当該処理方法の値を「データクレンジング処理ポリシー設定画面」の値欄で選択可能に表示する。

【0144】

ステップ S 1 1 1 7 にて、データ享受者側では、「データクレンジング処理ポリシー設定画面」で各「フィールド型属性」の各「不正値」に対する「処理方法」を設定し、ステップ S 3 1 1 8 にて、当該設定した「処理方法」を基盤事業者（サーバ）側に送信する。

基盤事業者側は、基盤事業者側からの「処理方法」を受け、当該「処理方法」を元に、ステップ S 3 1 1 9 にて、「データクレンジング処理ポリシーマスタ」を作成し、記憶装置に記憶する。

40

【0145】

図19は、データ提供者側と基盤事業者側との間におけるデータクレンジング処理ポリシー設定画面の表示生成するためのフローチャートである。

【0146】

同図において、データ提供者側（データ提供者 A 社）は、ステップ S 1 1 2 1 にて、基盤事業者側から「フィールド型属性マスタ」、「不正値マスタ」、「処理方法」マスタを取り込む。

【0147】

基盤事業者側は、ステップ S 1 1 2 2 にて、「フィールド型属性情報マスタ」、「不正

50

値マスタ」、「処理方法マスタ」から「データクレンジング処理ポリシー設定画面」を利用して「基盤事業者データクレンジング処理ポリシーCg」を設定する。

【0148】

データ提供者A社は、ステップS1123にて、「フィールド型属性情報マスタ」「不正値マスタ」「処理方法マスタ」から「データクレンジング処理ポリシー設定画面」を利用して「A社デフォルトデータクレンジング処理ポリシーCa」を設定する。

【0149】

データ提供者A社は、ステップS1124にて、基盤事業者に「A社元データIa」を送信する。

【0150】

データ提供者A社は、ステップS1125にて、基盤事業者に「A社元データIa」の「フィールド型属性情報型FIa」を送信する。

【0151】

データ提供者A社は、ステップS1126にて、「A社デフォルトデータクレンジング処理ポリシーCa」を参考にして、「A社データIa用デフォルトデータクレンジング処理ポリシーCaIa」を設定する。

【0152】

データ提供者A社は、ステップS1127にて、基盤事業者に「A社データIa用デフォルトデータクレンジング処理ポリシーCaIa」を送信する。

【0153】

図20～図26は、データクレンジング処理ポリシーの作成からデータクレンジング処理までの一手順例を示すフローチャートである。

【0154】

図20において、データ享受者P社は、ステップS1411にて、基盤事業者から「フィールド型属性情報マスタ」、「不正値マスタ」、「処理方法マスタ」を取り込み、「データクレンジング処理ポリシー設定画面」を利用して「P社デフォルトデータクレンジング処理ポリシーCp」を作成して、基盤事業者に送信する。

【0155】

データ享受者P社は、ステップS1412にて、基盤事業者から「A社データIa用フィールド型属性FIa」及び「A社データIa用データクレンジング処理ポリシーCaIa」を取り込む。

【0156】

データ享受者P社は、ステップS1413にて、「A社データIa用フィールド型属性FIa」及び「A社データIa用データクレンジング処理ポリシーCaIa」を参照して「データクレンジング処理ポリシー設定画面」を利用して「A社元データIa用P社データクレンジング処理ポリシーCpIa」を作成して、基盤事業者に送信する。

【0157】

しかる後、図21の*1へ進む。同図において、ステップS1421にて、データIaのレコード数iは、1からレコード数に達するまで(iレコード数処理)を実行し、ステップS1422にて、データIaのレコードiにおけるフィールドjは、1からフィールド数まで(jフィールド数)処理を実行し、ステップS1423に進み、当該ステップにて、フィールド(i、j)は不正値か否かを判定する。

【0158】

その結果、不正値の場合(Y)は、ステップS1424に進み、当該ステップにて、A社元データIa用P社データクレンジング処理ポリシーCpIaの有無を判定する。

不正値で無い場合(N)は、ステップS1430に進み、後述する処理を行う。

【0159】

当該ステップS1424にて、CpIaが有る場合は、図25の*0へ進み、後述する図10の処理方法に従った処理を実行する。CpIaが無い場合は、ステップS1425へ進み、当該ステップにて、A社データIa用データクレンジング処理ポリシーCaIaの

10

20

30

40

50

有無を判定する。

【0160】

当該ステップS1425にて、CaIaが有る場合は、図23の*7outへ進み、後述する図10の処理方法に従った処理を実行する。CaIaが無い場合は、ステップS1426に進み、当該ステップにて、P社デフォルトデータクレンジング処理ポリシーCpの有無を判定する。

【0161】

当該ステップにて、Cpが有る場合は、図24の6outへ進み、後述する図10の処理方法に従った処理を実行する。Cpが無い場合は、ステップS1427に進み、当該ステップにて、A社デフォルトデータクレンジング処理ポリシーCaの有無を判定する。

10

【0162】

当該ステップにて、Caが有る場合は、図25の*9outへ進み、後述する図10の処理方法に従った処理を実行する。Caが無い場合は、ステップ1428に進み、当該ステップにて、基盤事業者デフォルトデータクレンジング処理ポリシーCgの有無を判定する。

【0163】

当該ステップにて、Cgが有る場合は、図26の*8outへ進み、後述する図10の処理方法に従った処理を実行する。Cgが無い場合は、*100へ戻り、ステップS1430にて、jに1を追加($j = j + 1$)し、ステップS1431にて、iに1を追加($i + 1$)し、処理を終了する。

20

【0164】

次に、図22～図26において、図10の処理方法に従った処理について説明する。

図21のステップS1424にて、CpIaが有る場合は、図22において、ステップS14241にて、図14の処理ポリシーを参照し、処理方法IDを取得する。

【0165】

次に、ステップS14242にて、処理方法IDが1、2、3、4、5の何れで有るか無いかを判定する。ここで、ID=1、ID=2、ID=3、ID=4、ID=5で有る場合(Y)は、図21の*4に戻り、ステップS1429にて、処理方法IDに従った処理を実行する。また、上述したステップS1430、S1431を経て処理を終了する。

【0166】

図22のステップS14243にて、ID=1、ID=2、ID=3、ID=4、ID=5で無い場合(N)は、ステップS14243に進み、当該ステップにて、処理方法IDがID=7で有るか無いかを判定する。

30

【0167】

ここで、処理方法ID=7で有る場合(Y)は、図21の*7inに戻り、上述した処理を実行する。処理方法ID=7で無い場合(N)は、ステップS14244に進み、当該ステップにて、処理方法IDがID=6で有るか無いかを判定する。

【0168】

ここで、処理方法ID=6で有る場合(Y)は、図21の*6inに戻り、上述した処理を実行する。処理方法ID=6で無い場合(N)は、ステップS14245に進み、当該ステップにて、処理方法IDがID=9で有るか無いかを判定する。

40

【0169】

ここで、処理方法ID=9で有る場合(Y)は、図21の*9inに戻り、上述した処理を実行する。処理方法ID=9で無い場合(N)は、ステップS14246に進み、当該ステップにて、処理方法ID=8で有るか無いかを判定する。

【0170】

ここで、処理方法ID=8で有る場合(Y)は、図21の*8inに戻り、上述した処理を実行する。処理方法ID=8で無い場合(N)は、処理を終了する。

【0171】

図21のステップS1425にて、CaIaが有る場合は、図23において、ステップ

50

S 1 4 2 5 1にて、図 1 4 の処理ポリシーを参照し、処理方法 I D を取得する。

【 0 1 7 2 】

次に、ステップ S 1 4 2 5 2にて、処理方法 I D が 1、2、3、4、5 の何れで有るか無いかを判定する。ここで、I D = 1、I D = 2、I D = 3、I D = 4、I D = 5 で有る場合 (Y) は、図 2 1 の * 4 に戻り、ステップ S 1 4 2 9 にて、処理方法 I D に従った処理を実行する。また、上述したステップ S 1 4 3 0、S 1 4 3 1 を経て処理を終了する。

【 0 1 7 3 】

I D = 1、I D = 2、I D = 3、I D = 4、I D = 5 で無い場合 (N) は、ステップ S 1 4 2 5 4 に進み、当該ステップにて、処理方法 I D が I D = 6 で有るか無いかを判定する。

10

【 0 1 7 4 】

ここで、処理方法 I D = 6 で有る場合 (Y) は、図 2 1 の * 6 i n に戻り、上述した処理を実行する。処理方法 I D = 6 で無い場合 (N) は、ステップ S 1 4 2 5 5 に進み、当該ステップにて、処理方法 I D が I D = 9 で有るか無いかを判定する。

【 0 1 7 5 】

ここで、処理方法 I D = 9 で有る場合 (Y) は、図 2 1 の * 9 i n に戻り、上述した処理を実行する。処理方法 I D = 9 で無い場合 (N) は、ステップ S 1 4 2 5 6 に進み、当該ステップにて、処理方法 I D = 8 で有るか無いかを判定する。

【 0 1 7 6 】

ここで、処理方法 I D = 8 で有る場合 (Y) は、図 2 1 の * 8 i n に戻り、上述した処理を実行する。処理方法 I D = 8 で無い場合 (N) は、処理を終了する。

20

【 0 1 7 7 】

図 2 1 のステップ S 1 4 2 6 にて、C p が有る場合は、図 2 4 において、ステップ S 1 4 2 6 1 にて、図 1 4 ポリシを参照し、処理方法 I D を取得する。

【 0 1 7 8 】

次に、ステップ S 1 4 2 6 2 にて、処理方法 I D が 1、2、3、4、5 の何れで有るか無いかを判定する。ここで、I D = 1、I D = 2、I D = 3、I D = 4、I D = 5 で有る場合 (Y) は、図 2 1 の * 4 に戻り、ステップ S 1 4 2 9 にて、処理方法 I D に従った処理を実行する。また、上述したステップ S 1 4 3 0、S 1 4 3 1 を経て処理を終了する。

【 0 1 7 9 】

I D = 1、I D = 2、I D = 3、I D = 4、I D = 5 で無い場合 (N) は、ステップ S 1 4 2 6 5 に進み、当該ステップにて、処理方法 I D が I D = 9 で有るか無いかを判定する。

30

【 0 1 8 0 】

ここで、処理方法 I D = 9 で有る場合 (Y) は、図 2 1 の * 9 i n に戻り、上述した処理を実行する。処理方法 I D = 9 で無い場合 (N) は、ステップ S 1 4 2 6 6 に進み、当該ステップにて、処理方法 I D = 8 で有るか無いかを判定する。

【 0 1 8 1 】

ここで、処理方法 I D = 8 で有る場合 (Y) は、図 2 1 の * 8 i n に戻り、上述した処理を実行する。処理方法 I D = 8 で無い場合 (N) は、処理を終了する。

40

【 0 1 8 2 】

図 2 1 のステップ S 1 4 2 7 にて、C a が有る場合は、図 2 5 において、ステップ S 1 4 2 7 1 にて、図 1 4 ポリシを参照し、処理方法 I D を取得する。

【 0 1 8 3 】

次に、ステップ S 1 4 2 7 2 にて、処理方法 I D が 1、2、3、4、5 の何れで有るか無いかを判定する。ここで、I D = 1、I D = 2、I D = 3、I D = 4、I D = 5 で有る場合 (Y) は、図 2 1 の * 4 に戻り、ステップ S 1 4 2 9 にて、処理方法 I D に従った処理を実行する。また、上述したステップ S 1 4 3 0、S 1 4 3 1 を経て処理を終了する。

【 0 1 8 4 】

I D = 1、I D = 2、I D = 3、I D = 4、I D = 5 で無い場合 (N) は、ステップ S

50

1 4 2 7 5 に進み、当該ステップにて、処理方法 I D が I D = 8 で有るか無いかを判定する。

【 0 1 8 5 】

ここで、処理方法 I D = 8 で有る場合 (Y) は、図 2 1 の * 8 i n に戻り、上述した処理を実行する。処理方法 I D = 8 で無い場合 (N) は、処理を終了する。

【 0 1 8 6 】

図 2 1 のステップ S 1 4 2 8 にて、C g が有る場合は、図 2 9 において、ステップ S 1 4 2 8 1 にて、図 1 4 ポリシを参照し、処理方法 I D を取得する。

【 0 1 8 7 】

次に、ステップ S 1 4 2 8 2 にて、処理方法 I D が 1、2、3、4、5 の何れで有るか無いかを判定する。ここで、I D = 1、I D = 2、I D = 3、I D = 4、I D = 5 で有る場合 (Y) は、図 2 1 の * 4 に戻り、ステップ S 1 4 2 9 にて、処理方法 I D に従った処理を実行する。また、上述したステップ S 1 4 3 0、S 1 4 3 1 を経て処理を終了する。

【 0 1 8 8 】

I D = 1、I D = 2、I D = 3、I D = 4、I D = 5 で無い場合 (N) は、図 2 1 の * 1 0 0 へ戻り、上述した処理を行って終了する。

【 0 1 8 9 】

なお、本発明は上述した実施例に限定されるものではなく、様々な変形例が含まれる。例えば、上述した実施例は本発明を分かりやすく説明するために詳細に説明したものであり、必ずしも説明した全ての構成を備えるものに限定されるものではない。また、ある実施例の構成の一部を他の実施例の構成に置き換えることが可能であり、また、ある実施例の構成に他の実施例の構成を加えることも可能である。また、各実施例の構成の一部について、他の構成の追加・削除・置換をすることが可能である。

【 0 1 9 0 】

また、上記の各構成、機能、処理部、処理手段等は、それらの一部又は全部を、例えば集積回路で設計する等によりハードウェアで実現してもよい。また、上記の各構成、機能等は、プロセッサがそれぞれの機能を実現するプログラムを解釈し、実行することによりソフトウェアで実現してもよい。各機能を実現するプログラム、テーブル、ファイル等の情報は、メモリや、ハードディスク等の記録装置に置くことができる。

【 0 1 9 1 】

また、制御線や情報線は説明上必要と考えられるものを示しており、製品上必ずしも全ての制御線や情報線を示しているとは限らない。実際には殆ど全ての構成が相互に接続されていると考えてもよい。

【 符号の説明 】

【 0 1 9 2 】

- 1 サービス享受事業者装置
- 2 サービス提供事業者装置
- 3 サービス基盤事業者装置
- 3 1 基盤事業者用サーバ
- 3 1 1 入出力装置 (入出力インターフェース手段)
- 3 1 2 記憶装置 (記憶手段)
- 3 1 3 演算装置 (制御手段)
- 1 2 1 サービス提供事業者用サーバ
- 3 1 サービス基盤事業者用サーバ
- 1 2 1 データ提供者用サーバ
- 1 2 1 6 データ享受者用サーバ
- 3 1 2 1 データ提供事業者及びデータ享受者マスタ D B / データ及びフィールド型属性情報管理 D B
- 3 1 2 2 データ提供者 & データ享受者データクレンジング処理ポリシ情報管理マスタ D B

10

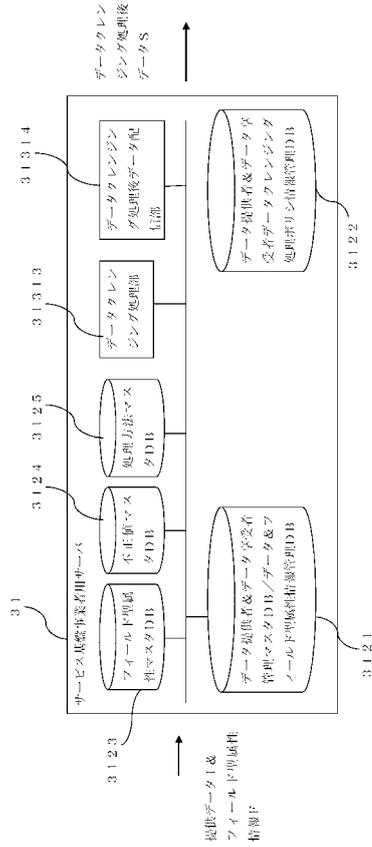
20

30

40

50

【図3】



【図4】

元データ (例:「かいか商店 株式会社」の「売上ジャーナルテーブル」)

401	402	403	404	405	406	407
元データ主キー	購入日	店舗名	顧客名	商品名	購入数	金額
1	2013.3.2	渋谷中央	〇田 〇部	光学式マウス	5	10000
2	2013.3.2	大崎東	〇木 〇部	4色ボールペン	100	20000
3	2013.3.3	大崎東	〇立 〇部	100枚名刺入れ	10	30000
4	2013.3.4	品川本店	J フェリー	筆箱	5	40000
.
.

【図5】

元データソースマスタ

501	502	503	504
元データ ID	元データ名	データ提供者 ID	データ提供者名
1	売上ジャーナルテーブル	1	かいか商店 株式会社
2	予約ジャーナルテーブル	2	きききホテル 株式会社
3	時間別売上テーブル	3	けけけ美容院 株式会社
4	予約ジャーナルテーブル	3	けけけ美容院 株式会社
.	.	.	.
.	.	.	.
.	.	.	.

【図6】

データ提供者マスタ

601	602
データ提供者 ID	データ提供者名
1	かいか商店 株式会社
2	きききホテル 株式会社
3	けけけ美容院 株式会社
.	.
.	.
.	.

【 図 7 】

元データフィールド属性マスタ

フィールド ID	元データ ID	フィールド名	フィールド型属性 ID
1	1	データ番号	1
2	1	購入日	8
3	1	店舗名	3
4	1	顧客名	3
5	1	商品名	3
6	1	購入数	7
7	1	金額	6
・	・	・	・
・	・	・	・
・	・	・	・

【 図 8 】

フィールド型属性マスタ

フィールド型属性 ID	属性
1	データ番号
2	購入日
3	名前
4	住所
5	電話番号
6	金額
7	数量
8	日付
・	・
・	・
・	・

【 図 9 】

不正値マスタ

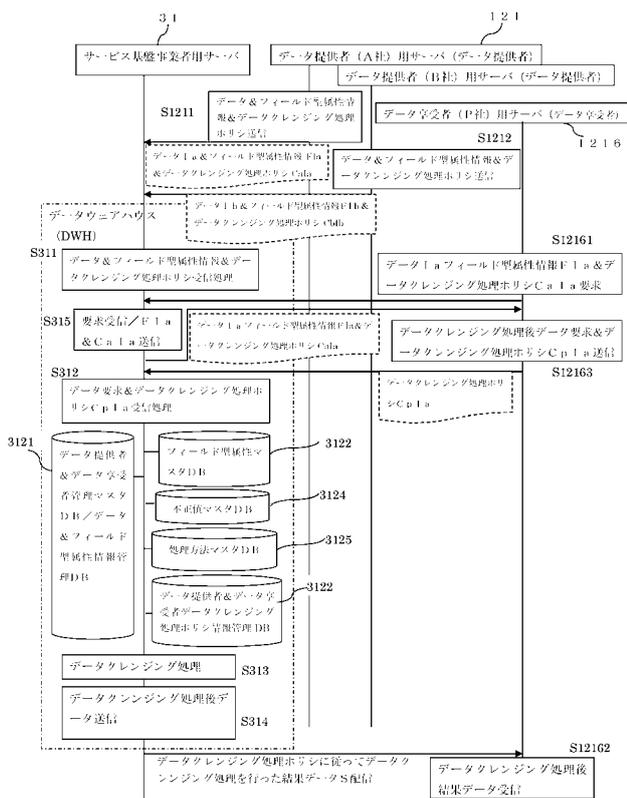
不正値 ID	不正値の型
1	空白
2	全角半角不統一
3	不正文字
4	非フォーマット
・	・
・	・
・	・

【 図 10 】

処理方法マスタ

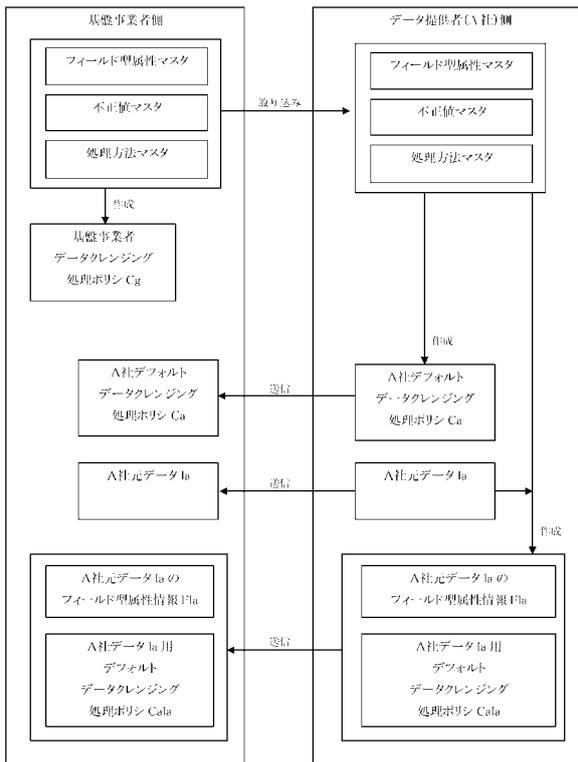
処理方法 ID	処理
1	何もしない
2	レコード全てを削除
3	空白にする
4	平均値に置き換える
5	英数字を全角文字にする
6	データ享受者のデフォルトデータクレンジング処理ポリシーに従う ⇒P 社デフォルトデータクレンジング処理ポリシー Cp
7	当該データ用データ提供者データクレンジング処理ポリシーに従う ⇒△社元データ Ia 用 A 社データクレンジング処理ポリシー CaIa
8	基盤事業者デフォルトデータクレンジング処理ポリシーに従う ⇒基盤事業者デフォルトデータクレンジング処理ポリシー Cg
9	データ提供者のデフォルトデータクレンジング処理ポリシーに従う ⇒△社デフォルトデータクレンジング処理ポリシー Ca
・	・
・	・
・	・

【図15】



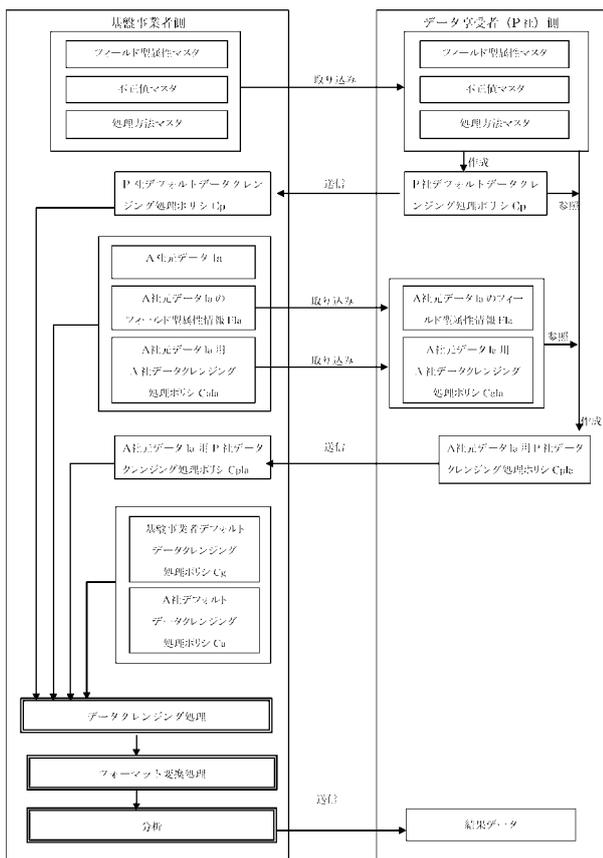
【図16】

データ提供者から基礎事業者への元データ、フィールド型属性およびデータクレンジング処理ホシの送信



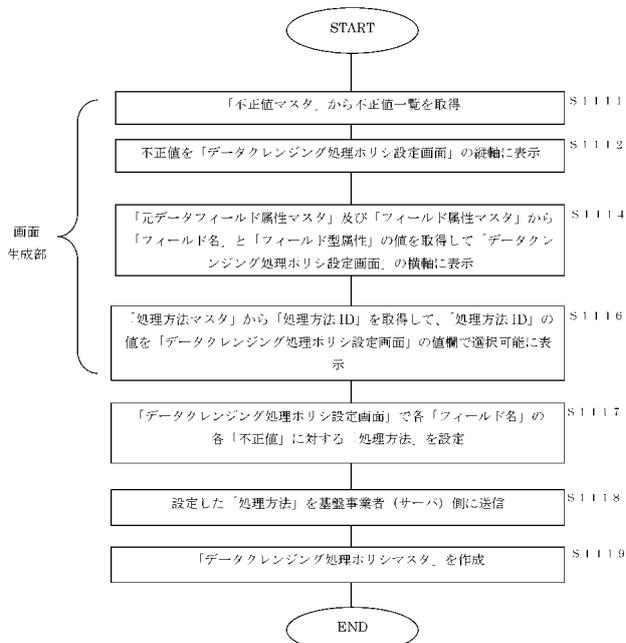
【図17】

データ受取者へのデータクレンジング処理後データの送信



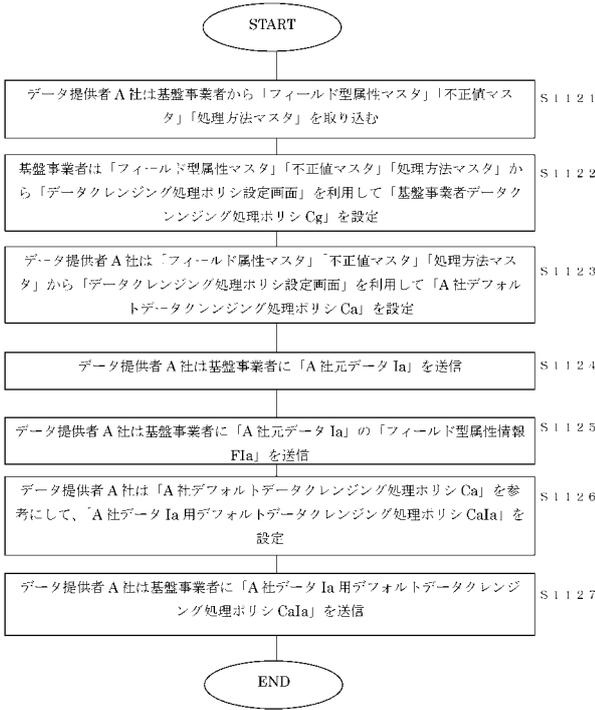
【図18】

「データクレンジング処理ホシ設定画面」の表示生成するためのフローチャート



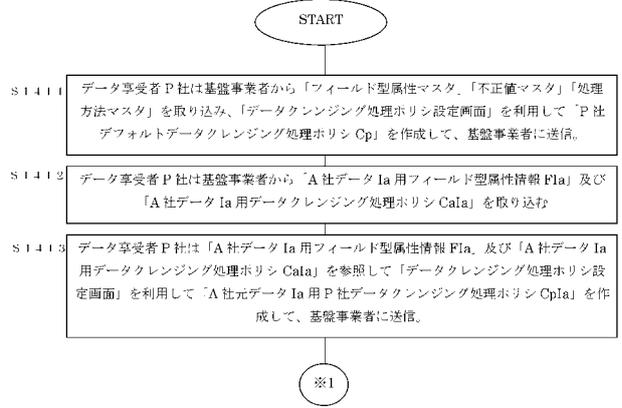
【 図 1 9 】

図 1 6 の処理フローチャート

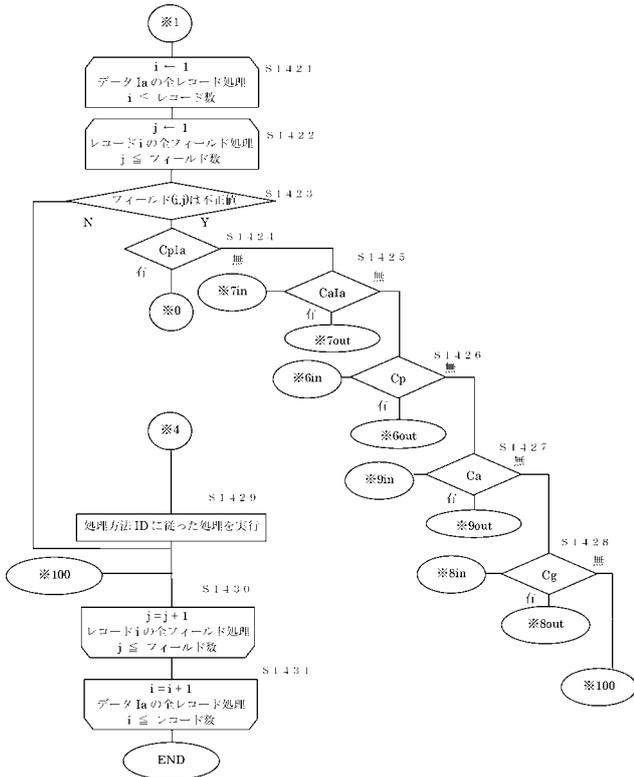


【 図 2 0 】

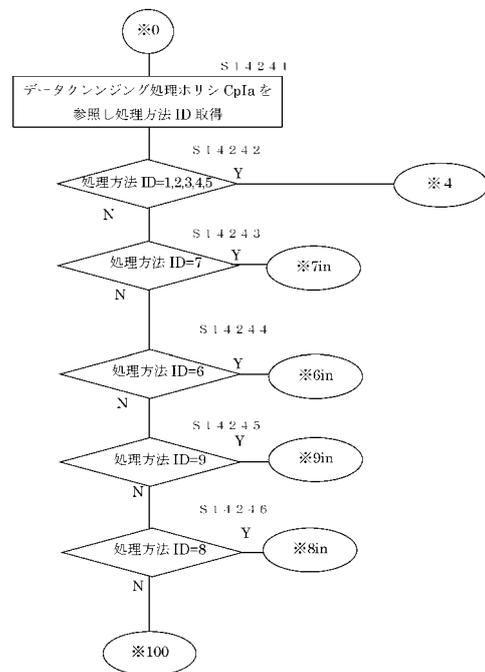
図 1 7 の処理フローチャート



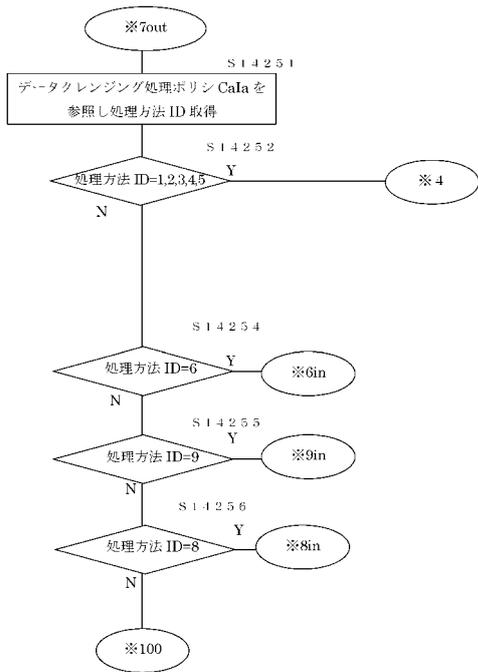
【 図 2 1 】



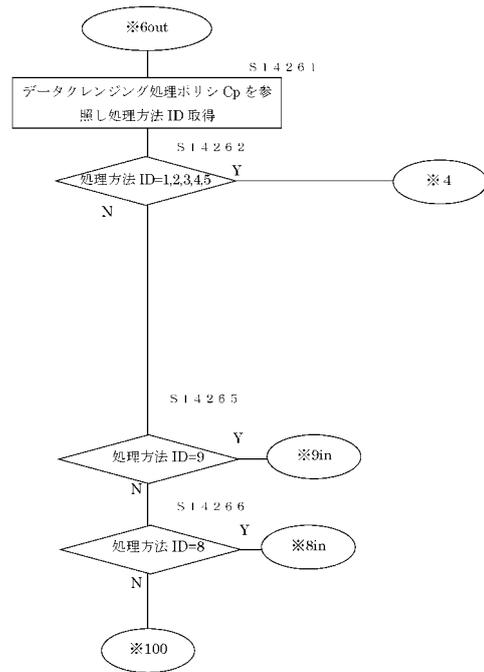
【 図 2 2 】



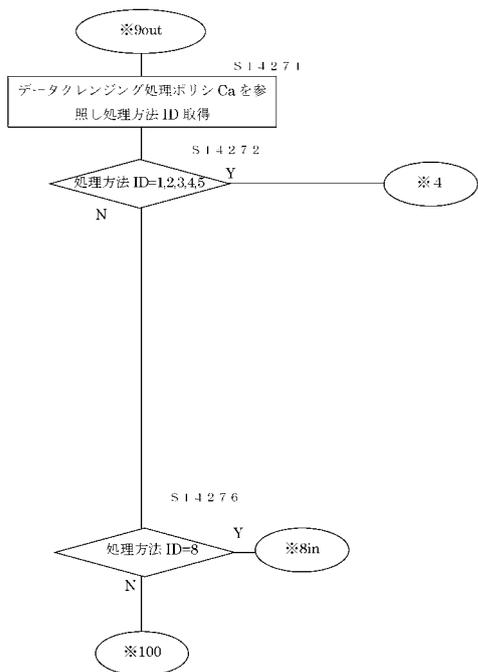
【 図 2 3 】



【 図 2 4 】



【 図 2 5 】



【 図 2 6 】

