

DEVICE AND METHOD FOR
CONTROLLING GRANULARITY OF ANONYMIZED DATA

BACKGROUND OF THE INVENTION

5 Field of the Invention

[0001]

The present invention relates to anonymized-data
granularity control device and method that have a function of
determining and controlling the granularity of anonymized data
10 provided by data providers, such as service providers and
service receivers.

[0002]

More specifically, the invention relates to anonymized-
data granularity control device and method that have functions
15 of receiving anonymized data which is provided by, for example,
multiple service providers or multiple service receivers and
satisfies k-anonymity (hereafter referred to as "k-anonymized
data"), determining whether to accept the k-anonymized data on
the basis of the granularity of the k-anonymized data and the
20 value of k, rejecting data having a high information leakage
risk, that is, data by which individuals may be identified, of
the k-anonymized data, and accepting and controlling data
other than such data.

The anonymized-data granularity control device and method
25 according to the present invention may have functions of

making uniform the data granularities of data items which are not related to anonymization (non-anonymized data) and controlling the acceptance of such data.

5 Description of the Related Art

[0003]

With the development of the Internet, there have been increased the number of various types of service providers that provide services useful for the categories or business types of companies. Examples of such service providers include service providers that provide merchandise information, such as popular menu items, to the restaurant industry, wholesale service providers and retail service providers that provide food material or the like to such service providers, and producers of food materials or the like.

[0004]

For example, there have existed service providing systems in recent years where through the internet or Web, service providers that provide particular business services to service receivers, such as restaurants, are provided with an infrastructure of the particular business services and then support the business of the service receivers by providing business services based on the service infrastructure.

[0005]

Such a service providing system includes a service

infrastructure provider device, service provider devices,
service receiver devices, and user terminal devices.

[0006]

The service infrastructure provider device is a system or
5 device that provides functions (services) based on the service
infrastructure to be provided, for example, a SaaS (Software
as a Service)-type operation system, to the service provider
devices.

[0007]

10 A service provider device is a system or device that
receives an operation system (business SaaS) based on the
service infrastructure provided by the service infrastructure
provider device. For example, the service provider device
customizes the operation system in accordance with a request
15 from a service receiver (or directly from a user) and provides
the customized operation system to the service receiver.

[0008]

A service receiver device is, for example, a device that
provides user information provided by a user to the service
20 provider device.

[0009]

A user terminal device is, for example, a terminal with
which the user himself or herself inputs user information and
provides the user information to the service receiver device.
25 A user terminal device may be a device which when the user

holds his or her mobile terminal over the user terminal device,
it can automatically gather information of the mobile terminal.

Original user information possessed by a service receiver
or service provider includes useful information. If the
5 service infrastructure provider can receive and analyze such
useful original user information, it can perform more suitable
business support.

[0010]

Related technologies in this technical field include
10 Japanese Patent Laid-Open Publication No. 2009-93569 (Patent
Document 1), Japanese Patent Laid-Open Publication No. H09-
153094 (Patent Document 2), and "Personal Information
Anonymization Technology",
http://www.meti.go.jp/policy/it_policy/daikoukai/igvp/contents
15 /pdf/C-1.pdf (Non-Patent Document 1).

[0011]

When changing an SaaS environment to an SI environment,
the technology of Patent Document 1 reduces occurrence of sunk
cost or opportunity loss in a business service whose use has
20 been started in SaaS form. Specifically, when transferring
business services from an SaaS provider environment (company
A) to an SI-form environment (city B), the technology provides
the city B with a service address list for recording the
addresses of the business services and, each time a business
25 service is transferred, updates an address in the service

address list. Further, when transferring business data stored in the company A, the technology provides the company A and the city B with data address lists for recording the addresses of pieces of business data and, each time a piece of business data is transferred, updates the addresses in both data address lists. With respect to the update of the addresses, there is disclosed a technology which changes the access method so that the addresses logically become the same.

[0012]

Patent Document 2 discloses a method for analyzing the effects of merchandise sales promotion from information about customers, merchandise information that the customers have referred to when considering the purchase of merchandise, and information about merchandise purchased by the customers; and a system for performing the method.

[0013]

Non-Patent Document 1 discloses a k-anonymization technology for using gathered personal information as corporate assets.

Citation List

Patent Document

[0014]

Patent Document 1: JP-A-2009-93569

Patent Document 2: JPA-H09-153094

Non-Patent Document

[0015]

Non-Patent Document 1:

http://www.meti.go.jp/policy/it_policy/daikoukai/igvp/contents/pdf/C-1.pdf

5

SUMMARY OF THE INVENTION

[0016]

According to a service providing system as described in Patent Document 1, a service infrastructure provider can
10 support the business of service providers. Each service provider can support the business of service receivers. Such business support can increase the business efficiency of the entire system.

[0017]

15 The service providers and service infrastructure provider, which provide such services, typically hold (store) and manage data about users, for example, data such as preferences or behavior histories of the users.

[0018]

20 Accordingly, if there is a system that can analyze the data about users and provide useful information based on the analysis result to service receivers (stores, general companies, or the like), which receive services from service providers, such a system can increase the business efficiency.

25 [0019]

In the present invention, "a user" refers to an individual, organization, or corporation that uses the present system and includes a specified user (a user which can be identified as a single individual, single organization, or single corporation) and an unspecified user (a user which cannot be identified as any individual, organization, or corporation or is difficult to identify as any individual, organization, or corporation). A terminal or the like used by a user in the present system will be referred to as "a user terminal." In the present invention, the term "individual" described with respect to a user includes an organization and corporation serving as users.

[0020]

"A service receiver" refers to a business entity that receives a core business support service according to the present system (first service). A system component device disposed on the side of the service receiver will be referred to as "a service receiver device". As used herein, "disposed on the side of" does not mean only being disposed inside the service receiver. The term "disposed on the side of" will be used in a similar manner.

[0021]

"A service provider" refers to a business entity that provides the core business support service according to the present system (first service). A system component device

disposed on the side of the service provider will be referred to as "a service provider device".

[0022]

"A service infrastructure provider" refers to a business
5 entity that organizes multiple service providers, and provides
an infrastructure of a service, for example, a business
function or reserve function (business service or reserve
service) or the like, to the service providers and/or service
receivers. The service infrastructure provider gathers
10 information from the service receiver devices and/or service
provider devices on the basis of a contract as necessary,
performs, for example, an analysis on the gathered information,
and provides the analyzed information (second service) to the
service receivers and/or service providers. A system component
15 device disposed on the side of the service infrastructure
provider will be referred to as "a service infrastructure
provider device".

[0023]

"A data provider" refers to a service provider or service
20 receiver. The data provider converts pieces of original user
information, Ia and Ib, into pieces of anonymized data Xia and
XIb satisfying k-anonymity and transmits them to the service
infrastructure provider. A system component device disposed on
the side of the data provider will be referred to as "a data
25 provider device".

[0024]

If the service infrastructure provider, which is a business entity organizing multiple service providers, can fuse and analyze various types of data stored in the service providers, it can perform higher-value added analysis. As a result, the service infrastructure provider can provide, as a service, the result of the higher-value added analysis to users including the service providers and/or service receivers.

[0025]

Under circumstances where an IT environment which allows analysis of large-volume data is being formed, it is technically possible to perform the above high-value added analysis on large-volume data.

[0026]

"A high-value added analysis" refers to an analysis for providing basic data which allows personalization (transmission of information corresponding to preferences or attributes of each client or prospective user through emails or the like), recommendation (a service where preferences of a user are analyzed and information which the user may be interested in is selected and displayed), or the like.

[0027]

"Granularity" refers to the coarseness or fineness of data to be analyzed and indicates to what unit the data is analyzed. In this example, the granularity of k-anonymized

data will be mainly described to simplify the description.

[0028]

When the service provider or service infrastructure provider performs the above high-value added analysis, it
5 faces a major challenge of how it addresses desires as described below of service receivers.

[0029]

For example, service receivers in the same industry (e.g., restaurants, or the like) have a desire "not to provide
10 information to the stores (service receivers) in the same area and the same category."

[0030]

Service receivers using a business service provider have a desire "not to provide information to other service
15 receivers in the same industry."

[0031]

Examples of conceivable, traditional systems which can address these desires include a "simple anonymization"-using service (hereafter referred to as "the simple anonymized
20 service") where service providers and service infrastructure provider fuse various types of data gathered from multiple service receivers, perform higher-value added analysis on the data, and provide the analyzed information to the service receivers as a service which the service receivers consider
25 the information as being useful for their business. The simple

anonymized service is a method where a service provider separates personal information from information about a user, performs "simple anonymization" such as "pseudonymization" on "identification information by which an individual can be
5 identified (information which singly allows an individual to be uniquely identified; for example, a name, full address, email address, or the like)," and then manages the resulting information. These pieces of information are managed in a separated manner, as well as associated with each other using
10 a uniform ID. With the relatedness between the pieces of information maintained, subtle information is managed.

[0032]

Since "information about each individual" is anonymized information in this simple anonymization service, information
15 risk reduction effects can be obtained. However, such a simple anonymization service has a problem of difficulty in completely eliminating the personal information leakage risk in terms of protection of personal information.

[0033]

20 Conceivable methods for correcting this problem (the personal information leakage risk alone) include use of the k-anonymization method described in Non-Patent Document 1.

[0034]

However, when gathering data using the k-anonymization
25 technology, the data granularity in the k-anonymization

process cannot be controlled between multiple data providers.

Specifically, how to provide simple anonymization service is left to service providers having desires, such as a desire "to anonymize part of information and then provide it," a
5 desire "to manage information about each individual in such a manner that the individual cannot be identified," or a desire "to provide services using only anonymous information." This poses a new problem that when fusing and analyzing anonymized data of multiple data providers, the differences in data
10 granularity prevent matching of the analysis axes, resulting in an inability to analyze or a reduction in analysis accuracy.
[0035]

Conceivable analysis methods include a method of analyzing information gathered from service receivers using a
15 statistical analysis technique, as described in Patent Document 2. However, the statistical analysis is based on a statistical process and therefore cannot produce analysis results considering the individual service receivers. Accordingly, the service receivers cannot obtain high-value
20 added analysis information as described above.
[0036]

Service receivers, on their part, have a desire to provide particular information of information possessed thereby to other service providers or other service receivers
25 for a fee. In other words, service receivers desire to

actively "sell" information possessed thereby unless the sale of the information is disadvantageous therefor.

[0037]

In view of the foregoing, an object of the present
5 invention is to provide anonymized-data granularity control device and method which allow, for example, a service infrastructure provider to gather various k-anonymized data tables (hereafter referred to as k-anonymized data.tb) generated by multiple data providers on the basis of
10 granularity definition documents provided by the service infrastructure provider and to determine whether to accept the data on the basis of information included in the k-anonymized data.tb.

[0038]

15 Another object of the invention is to provide anonymized-data granularity control device and method which can add a time stamp, charging information, or the like to accepted data.

Solution to Problem

20 [0039]

To achieve the objects, a service infrastructure provider device according to the present invention includes means configured to transmit, to a data provider device, a granularity definition document file (hereafter referred to as
25 "a granularity definition document.csv") defining the

granularity of data, means configured to receive a k-anonymized data.tb generated by the data provider on the basis of the granularity definition document.csv and including k-anonymized data, and data determination means configured to
5 accept only data that is included in the k-anonymized.tb and meets the value of k and/or a data granularity condition which serve as safety indices of data.

[0040]

The k-anonymized data j.tb may include the k-anonymized
10 data, as well as non-anonymized data. A k-anonymization process is performed on a quasi-identifier (name, address, or the like) of provided data.

[0041]

Specific examples are as follow.

15 [0042]

An anonymized-data granularity control device of the present invention includes: a data provider device; and a service infrastructure provider device. The service infrastructure provider device includes: means configured to
20 distribute, to the data provider device, a granularity definition document.csv defining granularity of data; means configured to receive a k-anonymized data.tb generated by the data provider device based on the granularity definition document.csv, the k-anonymized data.tb including k-anonymized
25 data; and means configured to determine whether to accept the

k-anonymized data, on the basis of granularity of the k-anonymized data and a value of k included in the k-anonymized data.tb. It is possible to prevent a risk that personal information may be leaked from the k-anonymized data and to
5 improve fusion and analysis accuracy of multiple pieces of k-anonymized data.

[0043]

When the k-anonymized data.tb receiving means receives the k-anonymized data which is included in the k-anonymized
10 data.tb and which has been subjected to a k-anonymization process defined in the granularity definition document.csv, the means receives an ID identifying a provider of the data and a granularity definition document No. from the data provider device.

15 [0044]

The anonymized-data granularity control device further includes means configured to, when the data determination means determines that the data is acceptable, add a time stamp and reward information to the data; and means configured to,
20 when the data determination means determines that the data is unacceptable, notify the data provider device of the determination.

[0045]

The anonymized-data granularity control device includes:
25 means configured to, when the data determination means

determines that the data is acceptable, register the data in the storage means; and data analysis means configured to analyze the accepted data. The service infrastructure provider device can fuse, as data to be analyzed, multiple pieces of acceptable data from the data provider device, analyze the fused data using the data analysis means, and distribute the analyzed data to the data provider device.

[0046]

In the anonymized-data granularity control device, the granularity definition document.csv includes one data granularity definition or multiple different data granularity definitions, and the k-anonymized data.tb generated based on the granularity definition document.csv includes one type of data or a group of two or more types of data.

[0047]

A method for controlling granularity of anonymized-data of the present invention is a method by which a service infrastructure provider device controls granularity of anonymized data transmitted from a data provider device. The method includes: distributing, to the data provider device, a granularity definition document.csv defining granularity of data; receiving a k-anonymized data.tb generated by the data provider device based on the granularity definition document.csv, the k-anonymized data.tb including k-anonymized data; and determining whether to accept the k-anonymized data,

on the basis of granularity of the k-anonymized data and a value of k included in the k-anonymized data.tb. It is possible to prevent a risk that personal information may be leaked from the k-anonymized data and to improve fusion and analysis accuracy of multiple pieces of k-anonymized data.

[0048]

The method for controlling granularity of anonymized-data further includes: when the data determination means determines that the data is acceptable, adding a time stamp and reward information to the data; and when the data determination means determines that the data is unacceptable, notifying the data provider device of the determination.

[0049]

The method for controlling granularity of anonymized data further includes: when the data determination means determines that the data is acceptable, registering the data in the storage means; and analyzing the accepted data. The service infrastructure provider device fuses, as data to be analyzed, multiple pieces of acceptable data from the data provider device, analyzes the fused data using the data analysis means, and distributes the analyzed data to the data provider device.

Advantageous Effects of Invention

[0050]

According to the present invention, the service

infrastructure provider defines the granularity of data. Thus,
it is possible to gather k-anonymized data where a quasi-
identifier of k-anonymized data from the data providers, for
example, the place or time is uniform in granularity. Further,
5 when integrating or fusing the data of the data providers and
analyzing the resulting data, it is possible to smoothly
perform analysis work with the analysis axes matched.

[0051]

Further, since the service infrastructure provider
10 specifies the value of k of k-anonymized data, the data
gathered from the data providers are previously k-anonymized.
Thus, advantageously, the service infrastructure provider does
not need any more to perform a personal information leakage
measure (k-anonymization process).

15 [0052]

While simple anonymization, for example, involves a
personal information leakage risk, k-anonymization can
eliminate the personal information risk. Thus, it is possible
to provide safer, more reassuring services. Further, the k-
20 anonymized data of the data providers can also be shared, or
integrated or fused.

[0053]

Problems, configurations, and effects other than those
described above will be clarified in the description of the
25 embodiments below.

BRIEF DESCRIPTION OF THE DRAWINGS

[0054]

Fig. 1 is a diagram showing the overall configuration of
5 a service providing system where an anonymized-data
granularity control device of the present invention is
provided;

Fig. 2 is a block diagram showing the configurations of a
service infrastructure provider device and a data provider
10 device of Fig. 1;

Fig. 3 is a block diagram showing an example
configuration of a computation device (control means) of the
service infrastructure provider device of Fig. 2;

Fig. 4 is a diagram showing an example of a table
15 included in a granularity definition document management
master DB of Fig. 2;

Fig. 5 is a diagram showing an example of a granularity
definition documents.csv included in the granularity
definition document management master DB of Fig. 2;

20 Fig. 6 is a diagram showing an example configuration of
data included in a k-anonymized data.tb generated by a k-
anonymized data generation unit of Fig. 2;

Fig. 7 is a diagram showing one example (k=1, data is
unacceptable) of a k-anonymized data.tb (table name G0001.tb);

25 Fig. 8 is a diagram showing one example (k=2, data is

acceptable) of a k-anonymized data.tb (table name G0001.tb);

Fig. 9 is an example of a table included in a k-anonymized data management master DB 3122; and

Fig. 10 is a flowchart showing the flow of a process
5 performed by the anonymized-data granularity control device of
the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0055]

10 Now, an embodiment of the present invention will be
described with reference to the drawings. First, a service
providing system for supporting the business of service
receivers will be outlined.

[0056]

15 The service providing system includes a service
infrastructure provider system or device configured to provide
application services or SI, that is, services such as a
reserve function or business SaaS (first service S1) to
service providers (hereafter referred to as "the service
20 infrastructure provider device"); and service provider systems
or devices configured to customize the first service in
accordance with desires of service receivers (restaurants,
etc.) and to provide the customized first service to the
service receivers (hereafter referred to as "the service
25 provider devices").

[0057]

A service receiver (restaurant, or the like) provides users (consumers) with food and drink, merchandise, or the like and receives user information (original information) from
5 users.

[0058]

The service infrastructure provider device may be an device configured to provide, for example, PaaS (Platform as a Service) services for lending a platform including an OS, ASP
10 (Application Service Provider) services, or the like rather than SaaS.

[0059]

Hereafter, the restaurant industry (restaurants), service providers therefor, and a service infrastructure provider will
15 be mainly described.

[0060]

First, a service providing system will be outlined. The service infrastructure provider device of the service providing system includes an application server configured to
20 provide a reserve management service (application service) or the like to the service provider devices (hereafter referred to as "the service infrastructure provider server").

[0061]

Each service provider device includes an application
25 server configured to customize this reserve management service

in accordance with desires of the service receivers and to receive and manage user (consumer) reserve information from the service receivers (hereafter referred to as "the service provider server"). That is, the service providing system has a
5 mechanism where the service infrastructure provider device and the service provider devices collaborate with each other to support the business of the service receivers.

[0062]

The system having such a mechanism typically addresses an
10 individual case corresponding to each service receiver.

Accordingly, the service infrastructure provider device only has to provide the first service S1 to each service provider device, while each service provider device only has to provide the customized first service (start of service) to
15 each service receiver and to manage pieces of user information (original information) Ia and Ib based on the provided first service, or the like.

[0063]

Examples of the method for providing this service include
20 SaaS. SaaS is an ICT utilization environment which allows a service provider or service receiver to use a function provided by an information system (software) (this function will be referred to as "the first service") via a network without having to construct or introduce a new system.

25 [0064]

SaaS typically has an advantage that it allows a service provider or service receiver to immediately use a needed service when needed without having to construct a special system. Further, since SaaS does not need to prepare a special environment in the company, it is possible to advance the time when the function (service) of SaaS starts operating and to reduce cost. Owing to the technological innovation of networks or browsers, a service provider or service receiver can operate SaaS, which uses the Internet or WAN, just as it uses an intra-company system.

[0065]

Furthermore, the service provider or service receiver can reduce cost or increase sales by operating or maintaining SaaS or SI provided by the service infrastructure provider.

[0066]

The present invention is intended to determine whether to accept data from the data provider devices described above (service provider devices A and B, service receiver devices a1, a2, b1, and b2, and the like), to share or fuse the data, and to eliminate the personal information leakage risk.

[0067]

Now, an embodiment of the present invention will be described with reference to the drawings. According to the present embodiment, it is possible to reject provided data by which an individual can be indirectly identified even when the

data is K-anonymized data (personal information leakage prevention), to accept only data by which any individual cannot be identified even when the data is disclosed to third parties, and to fuse the accepted pieces of data and analyze
5 the resulting data efficiently.

First Embodiment

[0068]

Fig. 1 is a diagram showing the overall configuration of
10 a service providing system where an anonymized-data granularity control device of the present invention is applied to a service infrastructure provider device.

[0069]

The service providing system includes a data provider
15 device 12 of a data provider and a service infrastructure provider device 3 of a service infrastructure provider.

[0070]

The data provider device 12 includes service provider
devices 2 of service providers A, B, and the like and service
20 receiver devices 1 of service receivers a, b, and the like.

[0071]

A data provider server 121 is used as the data provider
device 12.

The service receiver devices 1 are, for example, devices
25 of stores a1 and a2 or companies b1 and b2, and service

receiver servers 11 are used as the service receiver devices 1. Each service receiver server 11 outputs original user information Ia or Ib to the corresponding service provider device 2.

5 [0072]

The original user information Ia or Ib is sent from a user in response to the first service S1 (business/reserve function, or the like) provided by the service infrastructure provider device 3 through a network 4.

10 [0073]

Each service provider device 2 receives the original user information Ia or Ib, k-anonymizes the source of the original user information, that is, the user to the extent that the user cannot be identified, and transmits, as k-anonymized data, 15 the anonymized user information XIa or XIb to the service infrastructure provider device 3 through the network 4.

[0074]

K-anonymization is performed using, for example, the k-anonymization technique of Non-Patent Document 1 described 20 above. The k-anonymization will be described later.

[0075]

A service infrastructure provider server 31 is used as the service infrastructure provider device 3. The service infrastructure provider server 31 has functions of 25 distributing, to the data provider server 121, a granularity

definition document.csv (comma-separated values, that is,
comma separated file format) defining the granularity of data,
receiving data which is generated based on the granularity
definition document.csv, provided by the data provider server
5 121, and satisfies k-anonymity, and determining whether to
accept the data.

[0076]

Fig. 2 is a block diagram showing the configurations of
the service infrastructure provider server 31 and the data
10 provider server 121.

[0077]

The service infrastructure provider server 31 includes an
input/output device (input/output interface) 311, a storage
device (storage means) 312, and a computation device
15 (computation means) 313.

[0078]

The input/output device 311 have functions of receiving a
k-anonymized data.tb including the k-anonymized data and
transmitting a data rejection notification (to be described
20 later) or analyzed data to the data provider device 12.

[0079]

The storage device 312 includes a data provider
management master DB 3121, a k-anonymized data management DB
3122, a granularity definition document management master DB
25 3123, and an analyzed data registration DB 3124 (see Fig. 3).

[0080]

The data provider management master DB 3121 includes a data provider master table (not shown). The data provider master table forms a table scheme for storing and managing
5 pieces of profile information related to the service provider devices 2, for example, pieces of information such as service provider names (A, B, and the like), email addresses, and postal addresses.

[0081]

10 The k-anonymized data management DB 3122 includes a k-anonymized data management table T2. The k-anonymized data management table T2 (see Fig. 9) will be described later.

[0082]

The granularity definition document management master DB
15 3123 includes multiple granularity definition documents: a granularity definition document 1.csv, a granularity definition document 2.csv, and a granularity definition document 3.csv. The granularity definition documents.csv (see Fig. 5) will be described later.

20 [0083]

The computation device 313 includes control means 3131, such as a CPU. The control means 3131 includes a data provider identification unit 31311, a k-anonymized data determination unit 31312, a data registration unit 31313, and a data
25 analysis unit 31314.

[0084]

The data provider identification unit 31311 identifies a provider which has transmitted k-anonymized data from the data provider server 121, using a provider ID.

5 [0085]

The k-anonymized data determination unit 31312 determines whether to accept the k-anonymized data from the data provider server 121. This determination is made based on the granularity of the data and the value of k.

10 [0086]

If the k-anonymized data determination unit 31312 determines that the data is acceptable, the data registration unit 31313 registers the data in the k-anonymized data fusion management master DB 3122.

15 [0087]

The data analysis unit 31314 analyzes the accepted pieces of data and provides the analyzed data in accordance with a request from the data provider server.

[0088]

20 As shown in Fig. 3, the control means 3131 also includes a time stamping unit 31315, a reward information addition unit 31316, a data analysis unit 31314, an analyzed data distribution unit 31317, and a data acceptance/rejection result notification unit/data discarding unit 31318 having
25 functions of notifying of whether to accept data and

discarding data.

[0089]

The time stamping unit 31315 adds a time stamp to the accepted data. The reward information addition unit 31316 adds,
5 to the accepted data, information indicating that the data is the subject of reward.

[0090]

The analyzed data distribution unit 31317 distributes the analysis result made by the data analysis unit 31314 to the
10 data provider server 121.

[0091]

The data acceptance/rejection result notification unit/data discarding unit 31318 notifies the data provider server 121 of the determination result made by the k-
15 anonymized data determination unit 31312 as to whether to accept the data. If the k-anonymized data is determined to be unacceptable, the data acceptance/rejection result notification unit/data discarding unit 31318 notifies the data provider server 121 of the determination and discards the data,
20 with respect to which the determination has been made.

[0092]

The data provider server 121 includes a granularity definition document.csv request unit 1211, a granularity definition document.csv receiving unit 1212, a source data
25 registration DB 1213, a k-anonymized data generation unit 1214,

a k-anonymized data transmission unit 1215, and a data acceptance/rejection notification and analyzed data receiving unit 1216.

[0093]

5 When the granularity definition document.csv request unit 1211 attempts to provide source data, that is, original user information to third parties, it makes a granularity definition document.csv request to the service infrastructure provider server 31.

10 [0094]

 The granularity definition document.csv receiving unit 1212 receives the granularity definition documents.csv (the granularity definition document 1.csv, the granularity definition document 2.csv, and the granularity definition
15 document 3.csv) from the service infrastructure provider server 31. Alternatively, the data provider may download a granularity definition document.csv suitable for the data provider, of the granularity definition documents.csv which are previously made public on a network.

20 [0095]

 Based on the granularity definition documents.csv, the k-anonymized data generation unit 1214 receives source data (original user information) from the source data registration DB 1213 and generates k-anonymized data.

25 [0096]

The k-anonymized data transmission unit 1215 transmits the k-anonymized data (age, locality, or the like) generated by the k-anonymized data generation unit 1214 to the service infrastructure provider server 31. At this time, the k-anonymized data transmission unit 1215 transmits a provider ID (G, A, Y, or the like) representing the data provider and a definition document No. (1, 2, or 3) in a different file.

[0097]

When the data is rejected, the data acceptance/rejection notification and analyzed data receiving unit 1216 receives a notification from the service infrastructure provider server 31 and receives the analyzed data.

[0098]

Fig. 4 is a diagram showing details of a table T1 of the granularity definition management master DB 3123. In Fig. 4, the table T1 includes information indicating definition document Nos. (1, 2, 3) and definition document names (definition document 1.csv, definition document 2.csv, and definition document 3.csv).

[0099]

Fig. 5 is a diagram showing examples of the definition document 1.csv, the definition document 2.csv, and the definition document 3.csv. In Fig. 5, the definition document 1.csv includes definitions items, such as "k, age (range), age (beginning), residential area, working area, day, week

(beginning), and year (beginning)."

[0100]

As an example of "k, age (range), age (beginning), residential area, working area, day, week (beginning), year (beginning)," the definition document 1.csv shows "2, 3 year
5 old, 0 year old, municipality, prefecture, 5 to 29 o'clock, Monday, Jan. 1st."

As an example of "k, age (range), age (beginning), residential area, working area, day, week (beginning), year (beginning)," the definition document 2.csv shows "2, 5 year
10 old, 0 year old, municipality, prefecture, 5 to 24 o'clock, Monday, Jan. 1st." As an example of "k, age (range), age (beginning), residential area, working area, day, week (beginning), year (beginning)," the definition document 3.csv
15 shows "5, 3 year old, 0 year old, municipality, prefecture, 5 to 29 o'clock, Monday, Jan. 1st."

As an example of "k, age (range), age (beginning), residential area, working area, day, week (beginning), year (beginning)," a definition document n.csv shows "5, 10 year
20 old, 0 year old, municipality, prefecture, 0 to 24 o'clock, Sunday, Sunday of the first week."

[0101]

Fig. 6 is a diagram showing the data configuration of the k-anonymized data.tb. In Fig. 6, the k-anonymized data.tb
25 includes, as data, age range, residential area, and others.

Since the age range and residential area are quasi-identifiers, the data provider k-anonymizes these pieces of data in accordance with the granularity definition documents.csv transmitted from the service infrastructure provider.

5 [0102]

Fig. 7 shows one specific example of Fig. 6 and is a diagram showing an example where the table name of the k-anonymized data.Tb is G0001.tb and k is 2. In Fig. 7, a table T3 includes two items of data, age range and residential area, as quasi-identifiers. The table T3 also includes data
10 indicating the payment method and the amount of money, as others.

[0103]

"Age range: 30-34, residential area: Shinagawa Ward, Tokyo," "age range: 35-39, residential area: Minato Ward, Tokyo," "age range: 20-24, residential area: Minato Ward, Tokyo," and "age range: 25-29, residential area: Shinagawa Ward, Tokyo" each satisfy $k = 2$. Accordingly, these pieces of data are determined to be acceptable. On the other hand, "age
15 range: 30-34, residential area: Minato Ward, Tokyo" and "age range: 60-64, residential area: Shinagawa Ward, Tokyo" are each $k = 1$. Accordingly, these pieces of data are determined to be unacceptable.

[0104]

25 Fig. 8 is a diagram showing an example of a table T3'

including only the acceptable data on the basis of the determination made by the k-anonymized data determination unit 31312 with respect to the table name: G0001.tb, $k = 2$.

[0105]

5 Fig. 9 is a diagram showing details of the table T2 of the k-anonymized data management master DB 3122. In Fig. 9, the table T2 includes information, such as table names (G0001.tb, G0002.tb, A0001.tb, G0003.tb, Y0001.tb, G0004.tb), data provider IDs (G, G, A, G, Y, G), times (time stamps),
10 reward, the number of records, the number of items (the number of k-anonymized quasi-identifiers: age and area), and granularity definition document Nos. (1, 1, 2, 2, 3, 3).

[0106]

 Fig. 10 is a flowchart showing process steps performed by
15 the service infrastructure provider server and the data provider server.

[0107]

 In step S121, the data provider server 121 makes a granularity definition document request to the service
20 infrastructure provider server 31.

[0108]

 In step S311, the service infrastructure provider server 31 receives the granularity definition document request. In step S312, the service infrastructure provider server 31
25 extracts a granularity definition document.csv (for example,

the granularity definition document 1.csv) corresponding to the granularity definition document request from the granularity definition document management master DB 3123 and transmits it to the data provider server 121.

5 [0109]

In step S122, the data provider server 121 receives the granularity definition document 1.csv.

[0110]

In step S123, the data provider server 121 generates k-anonymized data on the basis of the granularity definition document 1.csv and transmits it to the service infrastructure provider server 31.

[0111]

In step S313, the service infrastructure provider server 31 receives the k-anonymized data. At this time, the service infrastructure provider server 31 receives the k-anonymized data, as well as a data provider ID (see Fig. 9) identifying the provider of the data and a definition document No. (see Fig. 4). Note that the data provider server 121 transmits the k-anonymized data, and the data provider ID and the granularity definition document No. in different files.

[0112]

In step S314, the service infrastructure provider server 31 determines whether to accept the k-anonymized data. This data determination is performed based on the granularity of

the k-anonymized data and the value of k.

[0113]

For example, assuming that the service infrastructure provider server 121 defines the value of k as 2, if the value of k is 2 or greater, the service infrastructure provider server 31 determines that the data is acceptable. If the value of k is 2 or less, it determines that personal information is very likely to be indirectly identified, although the data is k-anonymized and then determines that the data is unacceptable. The value of k is set to 2 or greater. Alternatively, the value of k may be set by the data provider.

[0114]

In step S315, the service infrastructure provider server 31 determines whether to accept the data and, in step S322, transmits the determination result to the data provider server 121. If the service infrastructure provider server 31 determines in step S315 that the data is unacceptable, it transmits a notification indicating the determination result to the data provider server 121 and, in step S321, discards the data, with respect to which the determination has been made. In step S126, the data provider server 121 receives the notification and can know whether the data is acceptable, based on this notification.

[0115]

If the service infrastructure provider server 31

determines that the data is acceptable, it adds, to the accepted data, a time stamp indicating the acceptance date of the data in step S316. In step S317, the service infrastructure provider server 31 adds, to the accepted data, reward information indicating the value of the data. These steps are performed optionally.

[0116]

In step S318, the service infrastructure provider server 31 registers the accepted data in the k-anonymized data registration DB 3122.

[0117]

In step S319, for example, based on a request from the data provider, the service infrastructure provider server 31 extracts k-anonymized data (1 or more pieces) suitable for the request from the k-anonymized data DB 3122 and performs a data analysis on the extracted data.

[0118]

In step S320, the service infrastructure provider server 31 transmits the analyzed data to the data provider server 121. For example, the analyzed data is transmitted based on a k-anonymized data request which is made by the data provider server 121 in step S124. In step S125, the data provider server 121 receives the analyzed k-anonymized data.

[0119]

By systematizing the determination and analysis of data

as described in the above embodiment, it is possible to easily achieve a mechanism where if the provided data conforms to the granularity definition document and satisfies the data determination criterion, reward (price) information is added to the data and then a payment is made.

[0120]

Further, by systematizing an analysis, it is possible to easily achieve a mechanism where if the provided data conforms to the granularity definition document, reward (price) information is added to the data and then a payment is made.

[0121]

The above embodiment includes the steps of: receiving k-anonymized data; subsequently determining whether to accept the data; and registering only the accepted data in the k-anonymized data management master DB 3122. Alternatively, the embodiment may include the steps of: temporarily registering k-anonymized data in the k-anonymized data management master DB 3122; subsequently determining whether to accept the data as data to be analyzed; and regarding only acceptable data of the data as data to be analyzed.

[0122]

The present invention is not limited to the above embodiment and includes various modifications. For example, while the embodiment has been described in detail to clarify the present invention, the invention need not always include

all the described components. Further, some components of an embodiment can be replaced with components of another embodiment, and an embodiment can additionally include components of another embodiment. Further, for some components
5 of each embodiment, other components can be added, deleted, or replaced.

[0123]

Some or all of the above components, functions, processing units, processing means, and the like may be
10 realized as hardware, for example, by designing these using integrated circuits. Alternatively, the above components, functions, and the like may be realized as software when a processor interprets and executes a program for realizing the functions. Information for realizing the functions, such as
15 programs, tables, and files can be stored in a storage device, such as a memory, hard disk, or solid-state drive (SSD), or a storage medium, such as an IC card, SD card, or DVD.

[0124]

The drawings show only control lines and information
20 lines considered necessary for description and do not necessarily show all of the control lines and information lines of the product. It can be considered that almost all the components are actually coupled to each other.

25 Reference Sings List

[0125]

1: service receiver device

2: service provider device

21: service provider server

5 3: service infrastructure provider device

31: service infrastructure provider server

121: data provider server

1211: granularity definition document request unit

1212: granularity definition document receiving unit

10 1213: source data registration DB

1214: k-anonymized data generation unit

1215: k-anonymized data transmission unit

311: input/output device

312: storage device

15 313: computation device

3131: control means

3121: data provider management master DB

3122: k-anonymized fusion management master DB

3123: granularity definition document management master DB

20 3124: analyzed data registration DB

CLAIMS:

1. An anonymized-data granularity control device comprising:
a data provider device; and

5 a service infrastructure provider device,
wherein the service infrastructure provider device
includes:

means configured to distribute, to the data provider
device, a granularity definition document.csv defining
10 granularity of data;

means configured to receive a k-anonymized data.tb
generated by the data provider device based on the granularity
definition document.csv, the k-anonymized data.tb including k-
anonymized data; and

15 means configured to determine whether to accept the k-
anonymized data, on the basis of granularity of the k-
anonymized data and a value of k included in the k-anonymized
data.tb, and

wherein it is possible to prevent a risk that personal
20 information may be leaked from the k-anonymized data and to
improve fusion and analysis accuracy of a plurality of pieces
of k-anonymized data.

2. The anonymized-data granularity control device according
25 to claim 1,

wherein when the k-anonymized data.tb receiving means receives the k-anonymized data which is included in the k-anonymized data.tb and which has been subjected to a k-anonymization process defined in the granularity definition document.csv, the means receives an ID identifying a provider of the data and a granularity definition document No. from the data provider device.

3. The anonymized-data granularity control device according to claim 1,

wherein the service infrastructure provider device further includes:

means configured to, when the data determination means determines that the data is acceptable, add a time stamp and reward information to the data; and

means configured to, when the data determination means determines that the data is unacceptable, notify the data provider device of the determination.

4. The anonymized-data granularity control device according to claim 1,

wherein the service infrastructure provider device includes:

means configured to, when the data determination means determines as a result of the determination that the data is

acceptable, register the data in the storage means; and

data analysis means configured to analyze the accepted data, and

wherein the service infrastructure provider device fuses,
5 as data to be analyzed, a plurality of pieces of acceptable data from the data provider device, analyzes the fused data using the data analysis means, and distributes the analyzed data to the data provider device.

10 5. The anonymized-data granularity control device according to any one of claims 1 to 4,

wherein the granularity definition document.csv includes one data granularity definition or a plurality of different data granularity definitions, and

15 the k-anonymized data.tb generated based on the granularity definition document.csv includes one type of data or a group of two or more types of data.

6. A method including a service infrastructure provider
20 device and a data provider device, the service infrastructure provider device controlling granularity of anonymized data transmitted from the data provider device, the method comprising:

distributing, to the data provider device, a granularity
25 definition document.csv defining granularity of data;

receiving a k-anonymized data.tb generated by the data provider device based on the granularity definition document.csv, the k-anonymized data.tb including k-anonymized data; and

5 determining whether to accept the k-anonymized data, on the basis of granularity of the k-anonymized data and a value of k included in the k-anonymized data.tb, and

wherein it prevents a risk that personal information may be leaked from the k-anonymized data and improves fusion and
10 analysis accuracy of a plurality of pieces of k-anonymized data.

7. The method of controlling granularity of anonymized data according to claim 6, further comprising:

when the data determination means determines that the
15 data is acceptable, adding a time stamp and reward information to the data; and

when the data determination means determines that the data is unacceptable, notifying the data provider device of the determination.

20

8. The method of controlling granularity of anonymized data according to claim 6, further comprising:

when the data determination means determines as a result of the determination that the data is acceptable, registering
25 the data in the storage means; and

analyzing the accepted data,
wherein the service infrastructure provider device fuses,
as data to be analyzed, a plurality of pieces of acceptable
data from the data provider device, analyzes the fused data
5 using the data analysis means, and distributes the analyzed
data to the data provider device.