# 新たな個人情報保護法における匿名加工情報の基準に関する一考察

藤村 明子　間形 文彦　千田 浩司　諸橋 玄武　高橋 克己

NTT セキュアプラットフォーム研究所
180-8585 東京都武蔵野市緑町 3-9-11
fujimura.akiko@lab.ntt.co.jp, magata.fumihiko@lab.ntt.co.jp, chida.koji@lab.ntt.co.jp,
morohashi.gembu@lab.ntt.co.jp, takahashi.katsumi@lab.ntt.co.jp

**あらまし** 個人情報保護法改正により、新たに匿名加工情報という枠組みが設けられる。本稿では、匿名加工情報に関する条文を分析し、プライバシー保護技術やデータの性質論と、法律上の要件との整合性を考察する。

# A study of anonymizing level standards following the new Act on the Protection of Personal Information

Akiko Fujimura　Fumihiko Magata　Koji Chida　Gembu Morohashi
Katsumi Takahashi

NTT Secure Platform Laboratories
3-9-11 Midori-cho, Musashino city, Tokyo 180-8585, JAPAN
fujimura.akiko@lab.ntt.co.jp, magata.fumihiko@lab.ntt.co.jp, chida.koji@lab.ntt.co.jp,
morohashi.gembu@lab.ntt.co.jp, takahashi.katsumi@lab.ntt.co.jp

**Abstract** With the amendment of the Act on the Protection of Personal Information, a new framework named anonymized information has been established. This paper analyses the articles concerning the anonymized information and discusses the consistency between current privacy protection technologies, nature of data and the legal requisites..

## 1　はじめに

### 1.1　背景

個人情報保護法（「個人の保護に関する法律（平成十五年五月法律第五十七号）」）が2005年に全面施行されてから10余年が経過した。

情報通信技術の発展によるビジネスモデルの多様化、グローバル化による国境を越えた執行協力体制の必要性など、現行法では対応困難な問題点を踏まえて、同法の改正法案が国会で審議中である（2015年8月時点）[1]。

改正ポイントは以下である。

（1）個人情報の定義の明確化、（2）適切な規律の下で個人情報等の有用性を確保、（3）個人情報の保護を強化、（4）個人情報保護委員

会の新設及び権限、(5)個人情報保護の取扱いのグローバル化 (6)その他改正事項等である[2]。

上記(2)は第三者提供等に関する規制の緩和と強化を中核に、一定条件の下で本人の同意なく自由な情報の流通を認め、データの利活用を促進するためのルールをプライバシー保護に配慮しつつ整備することを狙いとしている。

2013年9月から内閣官房「パーソナルデータに関する検討会」と、その関連会議「技術検討ワーキンググループ」にて法改正全般に向けた議論が進められた。第三者提供に関する新たな枠組みに関しては「個人特定性低減データ」の整理[3][4]、続いて「匿名加工情報(仮)」[5]という変遷を経て、2015年3月に法案で正式に「匿名加工情報」が位置づけられた。

匿名加工情報の加工方法、安全管理措置、照合の禁止、委員会への公表義務、提供先における再識別行為の禁止といった義務が改正法の条文で規定された[6]。しかし詳細については個人情報保護委員会規則で今後定められる予定となっている。さらに、具体的にどのような加工を行うかは、サービスの特性や取り扱う個人情報、匿名加工情報の内容に応じ、個人情報保護指針等による事業の実態を踏まえた自主的ルールに委ねられている[7]。

### 1.2 本稿の狙い

匿名加工情報の法的位置づけについて立法過程では有識者からの批判もある[8]。しかし、現在の改正法案がいったん成立すれば、個人情報取扱事業者は成立した法律を遵守しなければならない。

そこで本稿は、匿名加工情報の法律上の要件を条文から分析し、既存の匿名化技法でその要件が実現できるかについて考察する。

アプローチとして、まず現行の個人情報保護法の第三者提供の課題を背景に、新たな類型として匿名加工情報が登場した経緯を示す。その上で改正法案の条文を国会質疑応答などから分析して、匿名加工情報の加工方法の法律上の要件を導き、代表的な匿名化技法の機能と匿名性の指標を挙げながらこれらが法律上の要件にあてはまるかを考察していく。

## 2 現行法の第三者提供について

### 2.1 近年の事例

事業者間の第三者提供に関しては、鉄道会社が交通系 IC カードの使用で取得される情報の属性のうち氏名等を除外し、ID を不可逆な識別番号に置き換え処理したデータベースを利用者の同意なくして他社経由で販売しようとした事案があった。確かに ID が加工または削除されていれば元のデータベースと販売しようとしたデータベースを直接照合して個人を特定することは不可能そうである。しかし、このケースでは利用者ごとに一意に付与された ID と数十日分の乗降駅名と利用日時が属性の集合(以下、『レコード』)として記録、形成されるため、レコードの集合が唯一無二のデータとなる可能性が高まることになる。そのため、ID が除外されても、元のデータベースと販売しようとしたデータベースの間で準識別子と非識別子の組み合わせが一致するものを一意に確認することが可能となる。

各属性として、本事案のレコードを識別子、準識別子、非識別子として下記に示す。電話番号については変更可能であるから識別子ではないという見解もあるがここでは識別子とする。

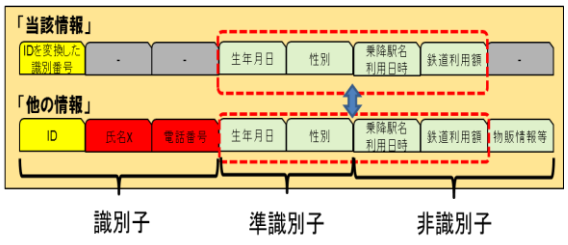


図1: 準識別子と非識別子の組み合わせの一致による特定個人の識別

このような準識別子と非識別子の組み合わせの一致によって特定の個人を識別できるデー

タは、後述する容易照合性があることから個人情報にあたる。

本事案は個人情報であるにもかかわらず第三者提供に求められる法律上の手続きを怠っていたことから現行法上違法といえる。その後同サービスは中断している[9]。

### 2.2 容易照合性と判断主体

容易照合性があるデータとは、特定の個人を識別することができるもののうち、「他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるもの」(現行法2条1項括弧書、改正法案2条1項1号括弧書)をいう。

容易照合性の判断主体は提供元である。これを提供元基準説という。改正法案全体からみると提供元基準説で制度設計されていることが明らかであり、現行法における政府参考人の見解も同説による[10]。

第三者提供のようなデータ移転の場面で匿名加工情報を作成・提供する際には提供元における容易照合性を検討しなければならない。

### 2.3 現行法からの見直し

個人情報を個人情報でなくするアプローチをとるには、何が適切な処理と認められるかが現行法上曖昧であったことから、個人情報取扱事業者には困難が伴った。

容易照合性の理解と処理が不十分な状態にもかかわらず第三者提供に必要な法律上の手続きを省けると誤解したサービス事業者の存在は、現行法上違法となるだけでなく消費者の社会的反発も招き、データの適正な利活用や自由な流通市場形成のむしろ妨げとなってきた。

こうした現行法の課題の中で、改正法案における匿名加工情報は法律の定めを満たしたデータであれば本人の同意なく第三者に提供できるようにする新たな類型として機能することが期待される。匿名加工情報が、法律見直しの趣旨である、“一定条件の下で本人の同意なく自由な情報の流通を認め、データの利活用を促進するためのルールをプライバシー保護に配慮しつつ整備する”ための制度として機能するには、法律の要件を具体的に実現できる個人情報の加工の基準が明らかであることが必要である。

## 3 匿名加工情報について

### 3.1 改正法案の条文

個人情報保護法の改正法案の条文を解釈する上で、現状2つの留意点がある。

1. 匿名加工情報の関連条文文言に関する見解が明確に定まっていない
2. 加工方法の基準は今後個人情報保護委員会規則において定められる予定(新法案36条1項)

上記 1 については、現時点までの政府資料や国会質疑応答を基に見解の変遷と方向性を見出し、それに沿った解釈をした場合の是非を考察する。

上記 2 については、2016年度中をめどに策定見込みであるから、改正法案の文言から要件を導き、詳細を定める委員会規則にどのような匿名化技法や基準を採用すべきかを考察する。

匿名加工情報の加工方法に関する条文は以下の通りである。下線部、強調部、段落替えは筆者による。

「この法律において「匿名加工情報」とは、**次の各号に掲げる<u>個人情報の区分</u>に応じて<u>当該各号に定める措置</u>を講じて**

**特定の個人を識別することができないように個人情報を加工して得られる個人に関する情報であって、当該個人情報を復元することができないようにしたもの**をいう。」(改正法案2条9項柱書)

「個人情報取扱事業者は，匿名加工情報(匿名加工情報データベース等を構成するものに限る。以下同じ。)を作成するときは、

**特定の個人を識別すること及びその作成に**

**用いる個人情報を復元することができないようにするために必要なもの**

として個人情報保護委員会規則で定める基準に従い，当該個人情報を加工しなければならない。」（同36条）

すなわち、法律上の要件としては、

①特定個人の識別性の排除

②個人情報の復元性の排除

という2点の達成が求められている。

同2条1項各号の「個人情報の区分」に応じた加工をして①②が実現できれば、そのデータは匿名加工情報となりえる。

加工の措置については、同2条9項各号に定められた「削除」あるいは「復元することのできる規則性を有しない方法により他の記述等に置き換えること」という措置により実現することになる。

| 区分 | 第一項第一号に該当する個人情報 | 第一項第二号に該当する個人情報 |
|---|---|---|
| 内容 | 【1項1号本文】当該情報に含まれる氏名、生年月日その他の記述等（個人識別符号を除く）により特定の個人を識別することができるもの<br>【1項1号括弧書き】他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるもの | 【1項2号】個人識別符号(※)が含まれるもの<br><br>(※)「個人識別符号」の定義<br>【2項1号】特定の個人の身体の一部の特徴を変換し当該特定の個人を識別することができるもの<br>【2項2号】特定の利用者若しくは購入者又は発行を受ける者を識別することができるもの |
| 措置 | 当該個人情報に含まれる記述等の一部を削除すること<br>（当該一部の記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）。 | 当該個人情報に含まれる個人識別符号の全部を削除すること<br>（当該個人識別符号を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）。 |

表1：2条1項各号、同9項各号の関係

## 3.2 政府解釈の方向性とその考察

①と②の要件を技術的に実現させるにあたり、どこまで求めるかについては国会質疑答弁[10][11]の趣旨を整理する。

①に関する部分は、匿名加工情報は個人情報に該当しないと明言されている。特定個人の識別性については、社会通念上、一般人の判断力や理解力をもって、情報の分析等によって生存する具体的な人物と情報との間に同一性を認めるに至ることが出来るものというこれまでの解釈を踏襲している。

特定個人の識別における容易照合性についても従来同様に特別の調査や費用や手間をかけることなく照合が可能な場合には容易照合性があるとし、技術的な面からは、スーパーコンピューターを用いなければ照合できないものは容易照合性があるとはいわない、通常一般的に今現在誰でも使えるような技術を使っては戻せないものは容易照合性があるとはいえない、という趣旨のことを述べている。加工の程度については、各項目の削除や置きかえの例として、グルーピングや分析対象データの平均から大きく乖離するデータ群をまとめる等の一般的な手法を定めると示しており、特異値の取り扱いについても特異値を消すような関数を使うことを規則で定めることを想定していると述べている。そしてこれらの定めにあたり産業界の意見を聞くことも言及している。

②に関する部分は、「復元することができないようにする」とは、通常の手段を用いて元に戻そうとしても元の個人を識別することができないことをもって復元することができないこととしている。ただし技術的側面から全ての可能性の排除までを求めるものではなく、通常の手段を用いて復元できないということであって、技術的に一〇〇%復元することを不可能にすることまで求めるものではない、としている。

大臣および政府参考人の発言から加工方法の定めの方向性を考察する。

匿名化技法としては①と②が特に厳格に峻別されておらず、一体として機能することを前提に議論がされている様子である。そして、誰でも使えるような技術、通常の手段を用いても戻せないこと、しかし、技術的側面から一〇〇%不可能までは求めないとされている。さらに”一般的な手法”として既存の匿名化技法を例示している点から、既存の匿名化技法として実現可能な機能や性質に基づいた基準が導かれるものと考えられる。

なお、今回の改正法案で匿名加工情報と認められるデータには以下の特色がある。匿名加工情報からは特定の個人の識別性が排除されていて非個人情報の扱いとなっているから、$k=1$ の仮名化はここに含まれなくなった。さらに

現行法ならば非個人情報とされていたものを改正法案の匿名加工情報とするにあたり、各事業者には匿名加工情報としての各種義務が伴うことになるため、むしろ規制強化の側面を有しているのではないかという疑問が残る。

### 3.3　措置を実現する匿名化技法の選択

①と②を実現する措置として改正法案2条9項各号が定める「削除」と「復元することのできる規則性を有しない方法により他の記述等に置き換えること」（以下、“置き換えの措置”と表記）は、どのように実現するか。

「削除」とは対象を完全に消去することである。“置き換えの措置”は条文上特に限定がないため、新たに別の記述に置き換える技法と既存の記述同士をいれかえる技法を含めて考えてよいだろう。

実務上、個人情報に含まれる記述等の一部や個人識別符号の全部の「削除」だけで、①の特定個人の識別性の排除に成功することは、属性がごく少数のシンプルなデータベースの例を除いては、稀である。

多数の属性を有するデータベースにおいては「削除」のみでは容易照合性の問題が残りうる。このことから、実務では加工方法として「削除」とあわせて、“置き換えの措置”を行うことも前提になると考えられる。

匿名加工情報への加工方法としていずれの匿名化技法を選択するかは、容易照合性を失わせることを念頭に、加工対象となるデータベースの規模や属性数、また利用目的や運用面を考慮しながら複数の選択が可能であるべきである。

「削除」や“置き換えの措置”による加工の程度は一般的な手法で定めるとあることから、国会質疑応答の例示を元に、プライバシー保護技術分野や統計分野において既に整理されている代表的な匿名化技法を以下に挙げる。詳細については[12][13]を参照されたい。

・削除
氏名、住所等の属性を削除すること

・トップ（ボトム）・コーディング
属性の値域に上限/下限値を設定してこれを超える値は上限/下限値に置き換えること
（例：100 歳以上の人→「80 歳以上」）

・グルーピング（リコーディング）
属性の値域をより一般化されたものに置き換えたり（例：年齢→年代）、所定の属性値をより一般化された値に置き換えたりすること

・リサンプリング
ミクロデータを全て提供するのではなく、確率的に抽出（サンプリング）したものだけを提供すること

・ソート（配列順の並べ替え）
データの並び順を換えること
（例：データの並びから元データへ推測ができる場合を防ぐ）

・スワッピング
異なる、あるいは類似したレコード間で同一属性の値をいれかえること

・ノイズ付加（誤差の導入）
数値属性にわざとランダムにノイズ（誤差）を混入すること
（例：20 歳を 21 歳に置き換える）。

・PRAM（Post Randomisation Method）
所定の属性値を確率的に別の値に置き換えること

・ミクロアグリゲーション(micro aggregation)
複数のレコードを属性値でグループ化し、属性地を所定の代表値（平均等）に置き換えること
（例：年齢を 2 歳刻みにする、4 捨 5 入する。区町村あるいは都道府県レベルにする）

### 3.4　匿名性の指標の例

匿名性の尺度として、たとえばk匿名性[14]、Pk匿名性[15]がある。

k匿名性とは、単一のミクロデータに対して同じ準識別子の組を持つレコードがk個以上存在するかどうかを測る指標である。同じような人がk人未満に絞り込まれないよう変形することからkの値が大きいと特定個人の識別に結びつきにくくなる利点がある。変形する手法としてトップ（ボトム）・コーディング、グルーピング等が用いられる。kの値を数値で設定できることから、基準の目安として用いやすいといえる。ただしkの値を上げすぎるとデータ分析に必要な情報も多く失われる。また加工なくしてそのまま用いる部分の内容から特定個人が推測できる場合もある。匿名加工情報にこの指標を用いる場合は、同じ準識別子の組をk個以上ではなく、全く同じレコードがk個以上存在するよう加工することで、容易照合性を失わせる方向にはたらかせることができる。

Pk匿名性とは、任意の個人のデータを1/k以下の確率でしか推定できないことを保証する指標である。Pk匿名性をみたす変形の手法としては、ノイズ付加やPRAM等が用いられる。

### 3.5　確率的場面と容易照合性の問題

確率的な手法を用いた匿名化技法で加工したデータの中には、加工前の元データと一致するものが出てくる場合がある。

加工前と加工後のデータを対照させたとき、外観上両者が一致して見える場合がある。

外観上一致して見える場合には、偶然に一致したケースと、本当に一致して特定個人の識別をし得るケースが含まれるが、いずれのケースにあたるかは外観で判別できない。このような場面を確率的場面と呼ぶことにする。

確率的場面におけるデータの容易照合性を法的にどのように評価するかが課題となる。

一律に容易照合性を認めると、偶然一致ケースは特定個人の識別性を失う加工がされていても匿名加工情報から外れる不合理が生じるし、一律に容易照合性を認めないとすれば特定個人の識別性リスクが高いデータが匿名加工情報の中に混ざるおそれがある。

リサンプリングを例にして確率的場面における容易照合性の問題を示す。

| 名前<br>（識別子） | 年齢<br>（準識別子） | 学歴<br>（準識別子） | 職業 | 年収 |
|---|---|---|---|---|
| Alice | 24 | Doctorate | 医者 | $40K |
| Bob | 25 | Doctorate | 弁護士 | $50K |
| Chris | 30 | Bachelor | 看護師 | $30K |
| Dan | 30 | Bachelor | 看護師 | $30K |
| Eve | 25 | Doctorate | 弁護士 | $50K |
| Flora | 32 | Master | 幼稚園教諭 | $20K |

表2：元データ

| 名前<br>（識別子） | 年齢<br>（準識別子） | 学歴<br>（準識別子） | 職業 | 年収 |
|---|---|---|---|---|
| Alice | 24 | Doctorate | 医者 | $40K |
| Bob | 25 | Doctorate | 弁護士 | $50K |
| Chris | 30 | Bachelor | 看護師 | $30K |
| Dan | 30 | Bachelor | 看護師 | $30K |
| Eve | 25 | Doctorate | 弁護士 | $50K |
| Flora | 32 | Master | 幼稚園教諭 | $20K |

表3：リサンプリングしたデータ

元データとリサンプリングしたデータを対照させたとき、25歳で年収が＄50KのDoctorateの弁護士であるBobあるいはEveいずれとも外観上一致して見える。こうした場合Eveの情報との外観の一致をもってEveの情報が特定できているから容易照合性があると主張できてしまいそうである。

しかし、リサンプリングではミクロデータ確率的に抽出（サンプリング）したものだけを提供しているので抽出されたのがBobかEveかは明らかではない。よって、準識別子とその他の非識別子が一致していたとしても、Eveの情報が特定できているとして容易照合性を認めてしまうことは不合理である。

このようなリサンプリングにより生じる確率的場面について統計分野では「特定できたとの主張に対し、特定できたと考えることが適当ではないと主張する方法でもある」と評価している[16]。

リサンプリングに限らず、Pk匿名性をみたすノイズ付加やPRAMなど確率的な手法を用い

た匿名化技法についても、同様の確率的場面が生じることがある。法律施行後の混乱と不合理を防止するためには、委員会規則などにおいて匿名化技法で処理した場合に発生する確率的場面における容易照合性の判断基準作りが必須となる。

### 3.6　今後に向けて

匿名化技法による加工前と加工後のデータ間の容易照合性についての議論は、改正法案に至る過程から現在まで未だ十分に尽くされていない。先述の国会質疑応答においても、スーパーコンピューター利用という手段の側面から言及するにとどまっている。また、容易照合性の概念の意味を誤解し、アクセス制御の問題やデータベース検索の機能の議論と混同されることもある。

現状を踏まえた上で、今後も匿名加工情報に関する委員会規則の制定や個人情報保護指針に基づく自主的ルール化に向け、引き続き検討を続けたい。

## 4　まとめ

現行法の第三者提供における課題を示し、改正法案で新たな類型として設けられた匿名加工情報の要件を条文から分析した。既存の匿名化技法の機能と匿名性の指標の例を示し、確率的場面における容易照合性の基準の問題を考察した。

## 参考文献


[1] 内閣官房 IT 総合戦略室,“第189回通常国会個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律案 新旧対照表,” 2015.3.10
http://www.cas.go.jp/jp/houan/150310/siryou4.pdf

[2] 同 “概要,” 2015.3.10
http://www.cas.go.jp/jp/houan/150310/siryou1.pdf

[3] パーソナルデータに関する検討会技術検討ワーキンググループ,“（仮称）準個人情報」及び「（仮称）個人特定性低減データ」に関する技術的観点からの考察について（中間報告）,” 2014.5
https://www.kantei.go.jp/jp/singi/it2/pd/dai9/siryou2-2.pdf

[4] 高度情報通信ネットワーク社会推進戦略本部,“パーソナルデータの利活用に関する制度改正大綱,” 2014.6
http://www.kantei.go.jp/jp/singi/it2/info/h260625_siryou2.pdf

[5] 内閣官房 IT 総合戦略室,“パーソナルデータの利活用に関する制度改正に係る法律案の骨子,”2014.12.19
http://www.kantei.go.jp/jp/singi/it2/pd/dai13/siryou1.pdf

[6] 内閣官房 IT 総合戦略室,“第189回通常国会個人情報の保護に関する法律及び行政手続における特定の個人を識別するための番号の利用等に関する法律の一部を改正する法律案 要綱,” 2015.3.10
http://www.cas.go.jp/jp/houan/150310/siryou2.pdf

[7] 衆議院,“第 189 回国会 内閣委員会第6号,” 2015.5.15
http://www.shugiin.go.jp/internet/itdb_kaigiroku.nsf/html/kaigiroku/000218920150515006.htm

[8] 高木浩光,“改正個人情報保護法に残された課題と今後の展望, ” 情報法制研究会第2回シンポジウム, 2015.6.28
http://www.dekyo.or.jp/kenkyukai/data/2nd/20150628_doc2.pdf

[9] Suica に関するデータの社外への提供についての有識者会議,“Suica に関するデータの社外への提供について中間とりまとめ, ” 2014.2
http://www.jreast.co.jp/chukantorimatome/20140320.pdf

[10] ]参議院,“第 189 回国会 内閣委員会第 2 号,”2015.3.25
http://www.shugiin.go.jp/internet/itdb_kaigiroku.nsf/html/kaigiroku/000218920150325002.htm

[11]参議院,“第 189 回国会 内閣委員会第 10 号,”2015.5.28
http://kokkai.ndl.go.jp/SENTAKU/sangiin/189/0058/18905280058010c.html
[12] 総務省,“匿名データの作成・提供に係るガイドライン,”2012.8.31 改正版
http://www.stat.go.jp/index/seido/pdf/35glv4.ppd
[13]Khaled El, Luk Arbuckle 著, 木村 映善, 魔 狸監訳,“データ匿名化手法,”オライリージャパン, 2015
[14] L. Sweeney, “k-Anonymity: A Model for Protecting Privacy”,International Journal of Uncertainty,Fuzziness and
Knowledge-Based Systems 10(5), 2002.
[15] 五十嵐大 , 千田浩司 , 高橋克巳, “k- 匿名性の確率的指標への拡張とその適用例,” コンピュータセキュリティシンポジウム 2009,情報処理学会,2009
[16] [12]の“別紙2,”