匿名加工・再識別コンテストにおける 有用性と安全性指標の社会実装に向けた検討

小栗 秀暢[†1]　　松井 くにお[†1]　　黒政 敦史[†1]

概要：個人情報保護法の改正により，匿名加工情報という新たな情報の類型が定義された．匿名化処理システムは，処理されたデータの有用性と安全性を計測し，匿名加工情報の認定プロセスに対して，適切に評価値を提供することが求められる．2015 年 10 月に行われた匿名加工・再識別コンテスト（PWSCUP 2015）では，匿名加工されたデータに対し，多様な観点から有用性と安全性の指標を実装し，その評価を試みた．本稿では，それらの有用性と安全性の基準について，データの利用目的に合わせて整理し，有用性・安全性指標に求められる要素を検討する．また，それらの指標群の社会実装に向けた課題の検討を行う．

キーワード：PWS2016，PWSCUP，匿名化処理，再識別処理，匿名加工情報

A Study for the practical implementation of the evaluation system of utility and security, through the data anonymization and re-identification competition.

Hidenobu Oguri[†1]　Kunio Matsui[†1]　Atsushi Kuromasa[†1]

Abstract: Because of revising the Privacy Preserving Law in Japan, Anonymized data —a new type of information was defined. Concerning the certification process, the anonymizing system that can measure the safety and utility of the data is demanded. PWSCUP 2015, the data anonymization and re-identification competition held in Oct 2015, proposed various evaluation index of the security and utility of anonymised data, and attempted to verify the efficiency of the indicators. In this paper, we discussed about the indicators of security and utility of anonymised data for the practical implementation of the evaluation system, and we organized the problems of indicators through the competition. Furthermore, we supposed how to implement those indicators to the society and evaluate the data which provided by participants efficiently.

Keywords: PWS2016，PWSCUP，Anonymization，Reidentification，De-identified data.

1. はじめに

個人情報の保護に関する法律に，2015 年 9 月に成立した同法の改正法[1](以後，改正後の同法を「改正法」という)により，匿名加工情報という新たな情報の類型が定義された．

匿名加工情報とは個人情報保護委員会規則で定める基準に従い，個人情報を加工して特定の個人を識別することができないようにするとともに，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(改正法三十六条)．匿名加工情報は，一定の条件の下で，本人の同意がなくても第三者に提供することが可能となる．それによって情報共有やマーケティング分析，機械学習の教師データ等への活用が期待できる．

匿名加工情報の加工基準については，個人情報保護委員会の規則において定める基準に合わせて，業界ごとの認定個人情報保護団体によって指針が整備される予定である．本稿では，匿名加工情報として認定される際に必要となる安全管理措置としての技術的な加工処理を「匿名化処理」，また，匿名化処理されたデータを「匿名化データ」と定義する．

しかし，各業界で利用されるデータ種類，利用方法，またはそのデータに適用される匿名化処理のアルゴリズム等の議論は進んでいない．

世界的に見ても，EU データ保護規則の承認や，ISO/IEC での匿名加工方式の国際標準化などが進むなか，日本国内における匿名化処理に関する研究の振興が求められている．このような状況下において，コンピュータセキュリティシンポジウム 2015 では，匿名加工技術の開発と再識別に対する公平な安全性評価手法の確立を目的とした世界初の大

†1 ニフティ株式会社
NIFTY Corporation,
Shinjuku, Tokyo 169-8333 Japan.

会『PWSCUP -匿名加工・再識別コンテスト"Ice & Fire"』[2](以後 PWSCUP)が開催された.

PWSCUP ではコンテストとしての公正性を保った上で,匿名加工情報とするに足る技術的な基準を検討するため,コンテスト参加者同士による対戦形式という形で相対的に評価する試みがなされた.しかし,PWSCUP で匿名化処理されたデータ評価のために提案された指標群は,あくまでコンテストの公正な運営のために検討された指標であり,そのまま,現実の匿名化データの流通に適用することはできない.

そこで,本稿では提案された指標の示す範囲を明確化し,社会実装に向けて必要な指標のあり方を検討し,実装可能な指標群を提案する.

2. PWSCUP 概要

PWSCUP は最大知識攻撃者モデル(maximum-knowledge attacker[3])を想定し,匿名加工・再識別の参加者双方が元データセットを共有した状態で行われ,匿名加工と再識別の 2 つのフェイズが存在する.

PWSCUP の運用の流れを図 1 にまとめる.

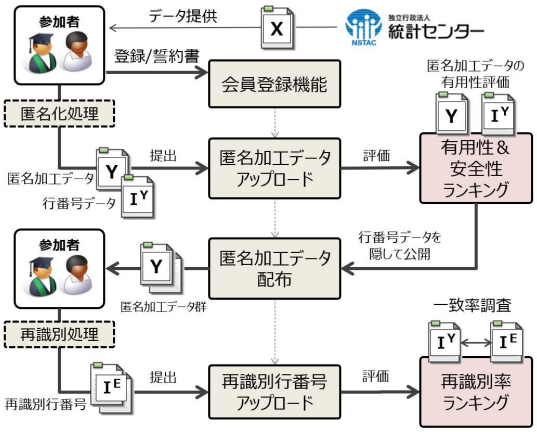

図 1　PWSCUP 運営の流れ

PWSCUP で使用したデータは,教育機関などの演習用として独立行政法人 統計センターが作成した疑似ミクロデータ[4]である.これを個人情報 X とした時,処理された匿名加工データを Y と表す.

1) 匿名加工フェイズ:参加者が擬似ミクロデータ X を匿名化処理した結果 Y と,その行番号データ I^Y を提出する.
2) 実行委員が作成した 13 項目の評価指標によって,Y の有用性と安全性のスコア(Rank)を算出する.
3) 再識別フェイズ:提出された Y を配布し,参加者は再識別行番号 I^E を提出する,I^Y と I^E を対照して再識別率を算出する.

これら PWSCUP で利用した評価指標と評価アルゴリズム,及び順位を定めるためのルール等は菊池らが[2]にて発表している.

3. 従来研究

匿名化データの安全性や有用性を定量化する指標は多く研究されている.

まず,安全性の指標として,k-匿名性[5]や,Pk-匿名性[6],ℓ-多様性[7]等の,個人が再識別される可能性,又は属性を推定される可能性を減少させる指標が多く提案されている.それらの指標群は,パーソナルデータに対する攻撃手法によって区分することができる.Fung らは[8]にてパーソナルデータを匿名化処理し,外部の Untrusted なデータ利用者に提供する PPDP(プライバシー保護データパブリッシング)の脅威モデルを図 6 の形で定義した.

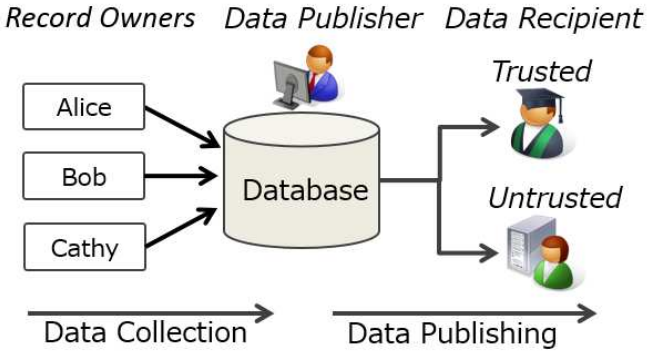

図 6　PPDP への攻撃モデル範囲

その中で,攻撃モデルをレコード結合(Record linkage),属性結合 (Attribute linkage),テーブル結合 (Table linkage),確率的攻撃 (Probabilistic Attack)の 4 種類と定義し,それぞれに有効な安全性指標を整理した.

しかし,匿名化処理によってデータ提供の安全性を高めることによって,そのデータの有用性を損なうことがある.特に,k-匿名性などの安全性を高める処理は,元データの属性値に対して一般化や統合,削除などを行うことから,データの安全性と有用性を両立させることは難しい.また,その際に,有用性を定義するための,データの利用目的も多様に存在する.

有用性を元データと匿名化データの数値属性における差,と定義する場合,情報量(エントロピー)の比較や,InfoLoss[9]などの情報の差を計算する指標が提案されている.一方,数値属性だけでなく,カテゴリ属性を用いた場合の指標も提案されており,Prec[10]や DIS[11]等,一般化階層を用いた際の抽象化レベルを計測する指標が提案されている.

匿名化データの有用性は,データの利用目的に依存するため,求める基準が異なる.また,分析の目的に応じて必要な属性の優先度が異なるなど,個別の対応も必要である.そのため,複数の安全性,有用性指標を組み合わせて評価

を行うことが求められる．

PWSCUP では，コンテスト形式を通じて，安全性指標と有用性指標の総合値による評価を試みた．本指標群を発展させ，社会実装に向けた指標の改良のため，総合的な安全性と有用性評価に求められる要件を検討する．

4. 指標検討の範囲

まず，PWSCUP で利用された共通データが適用されるデータの種類の範囲を定義する．まず，パーソナルデータの種類を「マスターデータ(M)」と「トランザクションデータ(T)」と区分する．

マスターデータは，パーソナルデータに含まれるレコードが，ある 1 ユーザの属性を示し，かつ，それらのレコードが重複しないものと定義する．一方トランザクションデータは，データ中にある 1 ユーザが複数回登場するものである．

PWSCUP（2015）で行われたものはマスターデータである．また，PWSCUP は 2016 年にも行われるが，そちらはマスターデータとトランザクションデータの両方を用いて行われている．今後も多くのデータ形式に対して適用範囲を広げていくことが求められる．表 1 にその内容をまとめる．

表 1　データ形式と採用した属性

データ形式	属性	適用
マスターデータ	数値	PWSCUP
マスターデータ	数値＋カテゴリ	-
トランザクションデータ	数値	-
トランザクションデータ	数値＋カテゴリ	PWSCUP (2016)

PWSCUP で利用された擬似ミクロデータは，全てのレコードが数値で示されている．正確には，カテゴリ属性が多く含まれているが，指標上では全て数値属性として計算可能な属性値として評価されている．

即ち PWSCUP で適用されたのは，全てをマスターデータの数値属性と認識した場合の評価指標群であり，カテゴリ属性を交えた安全性，有用性評価はカバーしていない．これは，パーソナルデータに含まれる属性の一般化，統合等の処理が発生せず，値の変更，削除，かく乱のみを対象とした匿名化データと定義できる．今後の技術の進展によって，この範囲が拡大されることが求められる．

5. PWSCUP のユースケースと採用指標

PWSCUP の想定するデータ提供と攻撃者のユースケースを図 2 に示す．

まず，データ提供者が個人情報 X の匿名加工データ Y を生成し，連結可能キー I^Y を廃棄してデータ利用者に提供する．

それに対して，X を入手可能な最大知識攻撃者が，Y の再識別を依頼されたため，X と Y を対照して再識別を試みるという処理である．

改正個人情報保護法 2 条 9 項の匿名加工情報の定義によると，***「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できないようにしたもの」***との記述がある．

即ち，提供元による情報の有用性，安全性の事前審査を行う形で指標が適用される．

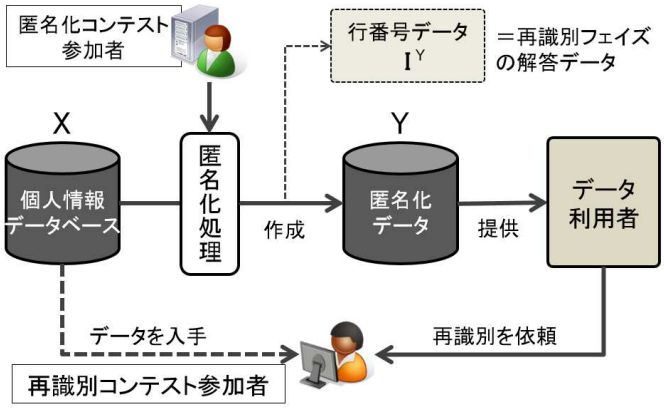

図 2　PWSCUP の想定ユースケース

この匿名化データを評価するための指標として PWSCUP では Ui:有用性評価, Si:安全性評価(k-匿名性), Ei：安全性評価(再識別率)，計 14 指標が設定された．表 2 にそれらの指標を示す．これら 14 の指標について，全ての指標は元情報 X, Y 及び IY を知っているものとして評価を行う．これによって，パーソナルデータを提供する側の企業における自己評価の形で行う．これらの総合的な匿名化データの評価は，各指標の相対的な順位を用いて評価を行った．

表 2　PWSCUP で採用した指標

No.	指標	指標説明
U1	meanMAE	SA 平均絶対誤差
U2	crossMean	クロス集計値の平均絶対誤差
U3	crossCnt	クロス集計数の平均絶対誤差
U4	corMAE	SA の相関係数の平均絶対誤差
U5	IL	データ各値の平均絶対誤差
U6	Nrow	データのレコード数
S1	k-anony	k-匿名性指標の最小値
S2	k-anonymean	k-匿名性指標の平均値
E1	IdRand	QI からランダムな再識別率
E2	IdSA	QI から SA15 列による再識別率
E3	Sort	SA 総和ソートによる再識別率
E4	SA21	SA21 列について再識別率
E5	AYA	山岡攻撃の検知による再識別率
E6	Reiduser	他参加者による再識別数の最大値

Ui～Si の指標は，実行委員が作成したサンプルプログラムによる評価を行っているが，Ei は，サンプルプログラムとコンテスト参加者による再識別結果を利用する．

表 3　各評価指標間の相関係数表

	U1	U2	U3	U4	U5	U6	S1	S2	E1	E2	E3	E4	MaxE1−4	E5	Euser	max E
U1	1.000															
U2	−0.272	1.000														
U3	−0.158	0.775	1.000													
U4	0.008	0.112	−0.147	1.000												
U5	−0.144	0.376	0.277	0.099	1.000											
U6	–	–	–	–	–	1.000										
S1	−0.151	0.713	0.595	0.368	0.399	–	1.000									
S2	−0.115	0.606	0.874	−0.175	0.486	–	0.574	1.000								
E1	0.243	−0.336	−0.234	−0.359	−0.717	–	−0.236	−0.195	1.000							
E2	0.221	−0.219	−0.054	−0.430	−0.831	–	−0.277	−0.219	0.881	1.000						
E3	−0.081	0.267	0.327	−0.265	−0.352	–	−0.139	−0.116	−0.081	0.336	1.000					
E4	−0.058	0.143	0.239	−0.342	−0.585	–	−0.186	−0.143	0.263	0.634	0.842	1.000				
MaxE1−4	0.225	−0.186	−0.016	−0.428	−0.836	–	−0.279	−0.220	0.839	0.996	0.416	0.689	1.000			
E5	−0.002	0.086	0.034	−0.520	0.345	–	0.025	0.037	0.058	0.098	0.087	0.074	0.096	1.000		
Euser	0.595	−0.279	−0.068	−0.335	−0.735	–	−0.311	−0.238	0.720	0.854	0.354	0.575	0.869	0.027	1.000	
max E	0.364	−0.053	0.016	−0.559	−0.047	–	−0.119	−0.076	0.346	0.482	0.246	0.357	0.492	0.840	0.535	1.000

PWSCUP では，ある再識別処理 E によって再識別を試みた結果データの一致率を $Reid^{E}$ と定義し，複数の再識別処理 Ei，及びコンテスト参加者による再識別処理 Euser を交えた最大識別率 $Max(Reid^{Ei})$を安全性指標 E として定義している．

PWSCUP 本戦におけるこれらの指標同士の相関係数を計測したものを表 3 に示す．これによると，有用性指標 U1～U5 の値は，全体的に再識別率との逆相関の数値が出ており，安全性と有用性は緩やかなトレードオフが発生している．しかし，指標の中には，U1 のようにあまりトレードオフが明確に出ていない値もある．これらの値を総括し，どの値が評価として優れていたかを判断することは難しい．

そのため，本稿では，まず，改正法が求めている安全管理措置の範囲を検討し，PWSCUP の指標がカバーしている範囲との比較を行う．そのために，既存の攻撃手法の定義，及びデータの有用性の定義との関係性を明確化した上で，必要な指標セットを提案する．

6. 改正法の範囲と照合の定義

改正法における匿名加工情報は以下のように定義されている．***「個人情報保護委員会規則で定める基準に従い，個人情報を加工して特定の個人を識別することができないようにするとともに，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(改正法三十六条)」***

この定義をユースケースに適用すると，匿名加工情報に対して，最大知識攻撃者（データの提供元）が攻撃を行い，匿名加工情報の各レコードにおける再識別を防止し，かつ，元情報への復元ができないことが求められる．

しかし，0%の再識別率と復元率を達成することは難しい．元情報に対して再識別される可能性は，最低でもそのデータレコード数を n としたとき，1/n の再識別可能性は存在することになる．そこで復元数を評価する妥当な基準が必要となる．

また，改正法には匿名加工情報を照合する対象が明記されていないが，本対象範囲は匿名加工情報を提供する提供元が負うべき責任である，という解釈がされている．

図 3 に，データ提供者（DP）が，匿名加工情報をデータ利用者(DU)に提供した場合におけるデータベース照合の範囲を示す．本図は[12]を基に，論文用に再構築したものである．

個人情報データベース X に対して，匿名化処理を施し，安全性評価を行ったものを匿名化データ Y とする．その Y に対し，個人情報保護委員会が定める規則等の規準を適用したものを匿名加工情報 Y'として DU に提供する．

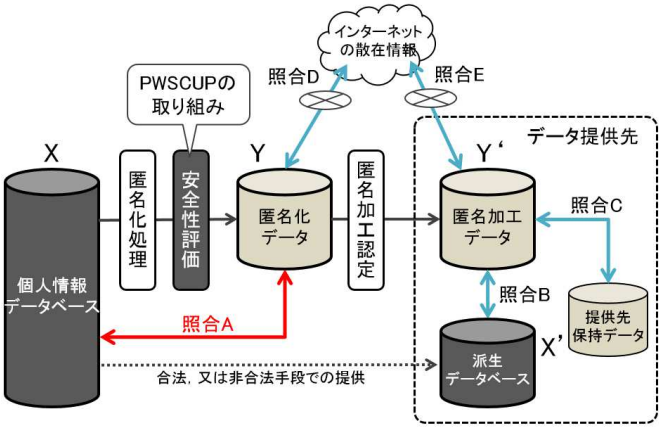

図 3 データベースの照合の種類

このとき，匿名化データと外部データベースの照合として，以下の 5 種類が存在する．

・照合 A： DP が X と Y(Y')を照合する．

・照合 B： DU が X から派生した X'を入手し，X'と Y'を照合する．

・照合 C： DU が入手した他のデータベースと Y'を照合する．

・照合 D： DP が Y（Y'）をインターネットなどの散在情報と照合する．

・照合 E： DU が，Y'をインターネットの散在情報と照合する．

提供元が行う匿名化処理は，この照合 A に関する安全管理措置である，と考えることができる．具体的には X→Y における情報の非識別性，不可逆性が求められている．

また，DP の義務として照合 D も求められると考えることができるが，改正法によってインターネット上に存在する散在情報を利用したプライバシーの侵害は，現状では考慮されていない．

今後，これらの問題が社会的な問題に発展した場合，考慮する必要性が高まることは充分に考えられる．

PWSCUP はこのような状況における X と Y の安全性と有用性を各指標で評価しており，現状の改正法の考え方に沿ってルールを構築されている．しかし，これらの指標とその運用方法が，パーソナルデータにおけるどのような攻撃を防止し，何に対して防止していないのか，明確に定義されていない．

次章では，改正法の求める範囲と，PWSCUP の提案指標が防止できる攻撃範囲について，Fung らが分析した，パーソナルデータ流通における攻撃モデル[8]に対応する範囲と比較し，その適用範囲を明確化する．

7. 攻撃モデルと指標の関係性

前章で述べたとおり，改正法において求められる安全性基準は，X→Y 間に閉じた安全性であると考える場合，その安全性の範囲は 2 章で紹介した Fung による PPDP（プライバシー保護データパブリッシング）の問題と同義であると考えることができる．そこで，第 2 章で述べた Fung らの攻撃モデルに対応した PWSCUP の指標について検討する．

7.1 レコード結合(Record linkage)

まず，レコード結合は，提供したデータベースにおいて，個人が一意に識別されるレコードが存在する場合の攻撃方法である．PWSCUP では，この問題を主に検討を行ってきた．

まず，レコード識別を定義する場合に，準識別子（QID）とセンシティブ属性（SA）を定義する必要がある．QID と SA は，多くの国・地域で定義が異なる．

まず通常の統計分析の考え方では，「QID」は分析対象における「説明変数」であり，「SA」はその「目的変数」と考えることができる．また，他にも QID を「特徴量」と呼称する場合もある．通常，分析対象である SA は残すべき重要な値であるため，SA から個人が識別される問題を考慮しない．

しかし，日本においては，何が QID で何が SA であるかとの判断ができないことから，全ての値を QID としてみなすべき，という議論がある．

PWSCUP では，このような問題に対して検証を行う目的で，QID と SA のどちらを用いても構わない形で再識別を行うルールを採用し，再識別リスクについて検証を行った．

しかし，QID だけの識別可能性を完全に排除することは，国際的な安全性基準の定義として相応しくないため，安全性を Si：安全性評価(k-匿名性)，Ei：安全性評価(再識別率)の2種類設定し，Si として，k-匿名性(指標 S1: k-anony)と k-匿名性平均値(指標 S2:k-anonymean)を採用した．この 2 つの指標は，X に含まれる属性 1～12 までを QID であると仮定したものである．

また，k-匿名性平均 (k-anonymean)を用いた理由として，k-匿名性はその情報の持つ安全性の最小値を示すが，それによって全ての等価クラスサイズを示すものではない．そのため，全ての等価クラスサイズの平均値を使用することで，総合的な安全性を評価するために用いられた．

k-匿名性平均の逆数は，QID の等価クラスに対して一律の再識別攻撃をかけた場合の再識別成功率と等しくなる．

表 4　k-匿名性平均の例

ID.	性別 (QID)	等価クラス	再識別攻撃	再識別結果
1	男性	A	1	○
2	男性		1	×
3	女性	B	3	○
4	女性		3	×
5	女性		3	×
6	女性		3	×

例として，表 4 に k-匿名性平均を安全性の基準とする場合の例を示す．ある個人情報を匿名化処理した結果，QID として性別(男性，女性)に区分し，等価クラス（A,B）が作成された．このとき，最大知識攻撃者ならば，少なくとも等価クラス A に所属しているユーザ ID の 1 名を選択し，等価クラスに対して全て同じ ID を入力する再識別攻撃が有効である．例では，k-匿名性は 2 であり，k-匿名性平均は 3，全体に対して一律値を入力する再識別攻撃を受けた場合の再識別率は 1/3 となる．

本攻撃手法は，アルゴリズム等を用いることなく，元情報を入手できる最大知識攻撃者の場合，容易に達成できるものである．そのため，この値を超える範囲で再識別が可能になった場合，効果的な再識別アルゴリズムであると判断することができるだろう．そのため，k-匿名性と合わせ，本指標が Si として定義されている．

Si の値が，その情報の持つ安全性の理論的な基準であるのに対し，Ei の値は，具体的な攻撃方法によって個人情報が再識別された値を示す．その時に，QID のみならず SA の値から個人を再識別されるリスクが発生することを考慮している．

レコード結合に対応する安全性指標として，PWSCUP では E1: IdRand と E2: IdSA の再識別アルゴリズムを採用した．

本アルゴリズムは，QID を集合化した等価クラスを定義し，その等価クラスの持つ SA 値をランダムに再識別（IdRand）するか，SA15 を基準としてソートした順番に再識別する（IdSA）方式を加えている．これにより，SA に対して無加工な匿名化データは，ほぼ全て一意に再識別されるため，SA の値とその順番を崩さずに，再識別率を下げるための工夫が必要となる．PWSCUP の指標群では，等価クラスにおける SA について，ある1つの SA 値を対象とした再識別のアルゴリズムのみが適用されており，他の SA は利用されない．そのため，社会実装を進める上では，あらゆる属性に着目して QID と SA でソートを行い，再識別を行う手法が必要となる．

しかし，この再識別処理を全ての QID と SA に適用する処理は，その情報に対して行われた加工処理そのものを解明する処理と同義といえる．

PWSCUP のコンテスト中において，再識別者が評価指標の

数値を参照して，摂動化が行われた属性を推定する攻撃が多く行われた．例えば，類似処理を行った結果，評価値が近似したパターンを探索することで再識別の手がかりとする手法などが提案された．他にも．機械学習によって摂動化処理の傾向を学習させて再識別する手法も行われている．

これら評価結果の保持は X→Y の不可逆性を担保する目的であるため，全ての値を残すことでアルゴリズムの可逆的探索の手がかりを残すことは本末転倒であり，安全管理の面から推奨できない．そのため，全属性値を用いた平均値などを用いて安全性の指標として用いるべきと考える．

7.2 属性結合(Attribute linkage)

属性結合は，QID による集合化がされている場合においても，個人と SA を紐付ける攻撃である．例えば，ℓ-多様性，t-近傍性のような，等価クラスと SA の関係性や分散状態を用いて個人の所属する属性値を知る攻撃方法である．

改正法によって求められているのは，個人の再識別と復元であることから，あるレコードに関する属性値を推定される属性結合攻撃は，法的に想定されていないと考えられる．

そのため，本攻撃を対象とした安全性の定義は，追加的な安全管理措置である．しかし，EU 一般データ規則等の規定によると，個人を識別しないが，行動による個人の自動的なプロファイリングは問題視され，規制対象となっている．

PWSCUP では全体の SA の分布から再識別を行う攻撃を想定して，E3: Sort と E4: SA21 を定義した． Sort は SA の総和を用いてソートを行い，個人の再識別を行う．SA21 は 21 番目の SA が分析に重要な値であるとの定義を行った上で利用した．これらの値を用いた再識別プログラムは，SA 全体の分散状況などを利用したものであるため，t-近傍性などの全体分散との関係性から得られる情報から再識別する手法と類似している．しかし，直接的な ℓ-多様性などの指標を計測する指標，再識別アルゴリズムが存在しなかった．

具体的に社会実装を行うためには，1）複数の SA に対応する再識別処理の追加，2）ℓ-多様性等の評価の追加 が必要と考えることができる．

7.3 テーブル結合(Table linkage)

テーブル結合は，公開された匿名化データにおける個人のデータが，何らかの別の方法にて攻撃者に知られていた場合に，個人を再識別できる可能性が高まる攻撃方法である．

例えば，あるサービスプロバイダーが保持するパーソナルデータ X を匿名化処理して，匿名化データ Y を公開したときに，Y⊆X であることが期待できる場合，X の派生 DB である X'を用いて，X'に含まれる要素が Y に含まれている確率を計算することができる． これは図 3 における照合 B のモデルである．

テーブル結合による攻撃方法は，データの提供先におけるテーブルの照合によって発生する攻撃方法を示している．そのため，最大知識攻撃者モデルによる提供元による攻撃に比べて，再識別または復元に関する攻撃強度が弱いと考えることができる．

データ提供先において，最大知識攻撃者以上に強力な比較データを保持している場合，その情報の提供自体が問題であり，匿名化データの安全性として評価せず，定性的な判断によって提供の可否を検討するべきだろう．

今後，SNS などが保持する情報が拡大し，パーソナルデータと結合可能なデータベースとして利用可能になった場合には，脅威が大きくなる．しかし，現状における SNS データなどは散在情報であり，個別のレコードに関する再識別攻撃はレコード結合と属性結合で排除できる．

7.4 確率的攻撃(Probabilistic Attack)

確率的攻撃は，パーソナルデータにおけるレコードや属性値を用いるのではなく，その公開されたデータの集計値や統計情報に対して行う．過去に提供したデータについて，その後，時間を置いた後に再度提供したデータとの統計的差異を検証することで，変化した個人を識別する．所謂，差分プライバシーと呼ばれる問題である．

本攻撃は，ある匿名化データと類似したデータの統計情報を用いて再識別を試みる攻撃である．そのため本コンテストでは，再識別攻撃者に対して，匿名化データの有用性，安全性指標の結果を小数点 8 桁(本戦では 5 桁)まで提供した．コンテスト参加者は，その統計数値の変更量から再識別を試み，いくつかのグループはそれによって多くの再識別成功数を記録した．現状では，明確なアルゴリズムは存在しないが，Euser による指標の確認によって再識別されるレコードが発生したと考えることができる．

本攻撃は，匿名化データのレコードに対する攻撃ではないことから，やはり改正法の範囲には含まれない．また，最大知識攻撃者に対しては，元情報に含まれるタイムスタンプの差異を用いた再識別攻撃も可能であることから，レコード結合攻撃の一種として考えることも可能である．

ただし，本問題は，識別子が重複しないマスターデータにおいては識別可能性を容易に検証できるが，トランザクションデータに対しては考慮していないことを確認しておく．

8. 安全性指標のまとめ

安全性指標の考え方についてまとめる．PPDP における4つの攻撃モデルに対応する安全性指標として，改正法で求めている要件と，PWSCUP で評価した指標について整理する．

4 つの攻撃に対して，PWSCUP が明確に指標として定義したのは，レコード結合攻撃に対してであり，改正法もその範囲の安全性を明確に求めている．しかし，改正法はそれ以外の攻撃モデルに対して考慮しないため，PWSCUP では，追加的な指標として，属性結合，確率的攻撃に対する安全性を確認する指標を定義した．それらの指標のまとめを表 5 にて示す．

表 5　攻撃モデルと対応指標のまとめ

攻撃モデル.	代表的な指標	PWSCUP
レコード結合	k-匿名性	k-anony k-anonymean IdRand IdSA
属性結合	ℓ-多様性 t-近傍性	Sort SA21
テーブル結合	δ-存在性	レコード結合に含まれる
確率的攻撃	ε-差分プライバシー	Euser による再識別処理

PWSCUP の指標群だけでは，属性結合攻撃に対する代表的な指標である ℓ-多様性などに関する評価がカバーされていない．また，コンテストの運営の都合上，QID と SA の一部をピックアップして指標化している箇所があるため，それらを他の属性にも展開し，可逆性を残さないように指標化する必要がある．また，確率的攻撃に対しては統計指標の提供による再識別攻撃をユーザが行う形で求めたが，これらの値を用いた攻撃方法を定式化し，評価アプリケーションとして開発することが求められるだろう．

匿名化データの評価システムの社会実装に向け，本稿にて検討した安全性指標セットを表 6 にて示す．

表 6　安全性指標セットの提案

No.	指標名	説明
1	nrow	元データと比較したサンプリング率.
2	k-anony	ある属性群を QID と設定した場合の k-匿名性.
3	k-anony mean	ある属性群を QID と設定した場合の等価クラスサイズの平均値.
4	IdRand	ある等価クラスが含む SA の値をランダムに再識別.
5	IdSA mean	ある等価クラスが含む SA の値をソートして再識別．結果を SA の平均とする.
6	Sort	SA 総和によってレコード全体をソートし，再識別.
7	SAmean	ある SA によってレコード全体をソートし，再識別．結果を SA の平均とする.
8	l-diversity	ある属性群を QID と設定した場合の SA の多様性.
9	l-diversity mean	ある属性群を QID と設定した場合の SA の多様性の平均値.

PWSCUP の安全性指標に対して，IdSA と SA の再識別プログラムを SA 全体に適用し，その平均値を取得する．また，属性結合攻撃に対応するため，ℓ-多様性と，各等価クラスにおける平均多様性を用いて，指標セットとすることを提案する．

9.　有用性指標について

前章までは安全性指標について検討してきたが，同様に有用性の指標セットも必要である．しかし，有用性評価はデータの利用目的に応じて多様なセットが必要であり，かつ，改正法が求める要件などが存在しない．また，表 1 で定義したように，数値属性とカテゴリ属性の区分についても検討が必要である．

表 7 は，匿名化データが利用可能な目的について検討した表である．これ以外にも多くの利用方法が考えられるが，匿名化データという個人に結びつかないデータであることから，ある程度用途を限定することができる．1）データ分析，2）情報配信，3）データ共有である．

表 7　匿名化データの用途分析

大項目	小項目	主な利用方法
データ分析	調査・研究	・相関/回帰分析等の手法で詳細分析 ・統計/表計算ソフトへの投入，連携
	機械学習	・機械学習用の教師データとして利用 ・次元の変更とデータ量の調整が必要
情報配信	再結合・再投入	・広告配信DSPやABテストへの投入 ・レコメンドエンジン等への活用
	広告・検索	・自社/他社の広告媒体の調査や発注に利用. ・求めるデータを検索してマッチング
データ共有	表・グラフ化	・結果をGoogle Docs等でメンバーに共有 ・参加者の位置や軌跡情報をWEB上で公開.
	報告・連動	・属性サマリをレポート化して共有 ・一定の属性のユーザが蓄積したことを連絡

PWSCUP では，有用性の定義として 6 つの指標を採用した．それらは基本的には上記の用途分析における，データ分析上の目的を達成するために採用された指標である．そのため，その他の利用目的がある場合は，また別途の有用性定義を行う必要がある．

本指標を検討する中で，IL（情報損失）に関する定義から，分析に必要とされるデータに関する要件を検討する．

2 章で検討した InfoLoss[9]の指標は，ある情報に関する総合的な情報損失の計測手段として，表 8 の 5 つの変化の平均を用いている．この定義を用いることで，分析に必要な情報損失の要素を定義することが可能である．

また，本指標群は，使用する目的に応じて指標の重要性を変化させることが可能である．例えば，分析目的が SA の合計値の評価である場合は，値の変化率を重視，また，値の相関関係の調査に利用する場合は相関行列の値を重視することで，多くの分析に適用することが可能である．

表 8 情報損失 ILoss の定義と PWSCUP 指標

No.	指標種類	PWSCUP
1	値の平均変化率	IL
2	属性平均値の平均変化率	meanMAE
3	共分散行列の平均変化率	crossMean
4	分散の平均変化率	存在しない
5	相関行列の平均変化量	corMAE

しかし，この内 4.分散の平均変化率に関しては，PWSCUP では実装されていない．そのため，本指標を加えた有用性指標セットを表 8 として提案する．

表 8　有用性指標セットの提案

No.	指標名	説明
1	nrow	元データと比較したサンプリング率.
2	meanMAE	SA 平均絶対誤差
3	crossMean	クロス集計値の平均絶対誤差
4	crossCnt	クロス集計数の平均絶対誤差
5	corMAE	SA の相関係数の平均絶対誤差
6	IL	データ各値の平均絶対誤差
7	Dist	属性単位の分散の誤差，又は今日分散行列の平均変化率

10. まとめ

PWSCUP の評価指標を決定する過程において，各実行委員が提案した指標とその評価アプリケーションを数多く作成し，共通データセットを用いて評価を行った．それらの結果として，大枠において匿名化データの安全性と有用性を評価する指標として成立する指標群が提供できた．

しかし，社会実装に向けた実データのあり方まで考慮に入れると，指標の持つ位置づけが変化するため，そのままでは利用できない点も多く存在する．

現状では匿名化データに関する多面評価のための仕組みが整備されていないため， PWSCUP を通じて得られた指標の社会実装に向けた展開と，その防止する攻撃範囲を明確化することによって，多様な分野のデータ評価における指標のセットとして利用していきたい．

また，コンテストを通じて得られた指標の結果数値を用いて，データの相対的な評価につなげることも可能となる．現状では仮に優れた指標を提案したとしても，出力結果を比較する基準やシステムが存在せず，絶対的な数値の意味を吟味する必要がある．その際に PWSCUP を通じて得られた結果と比較することによって，匿名化処理を試みた研究者の出力データと比較して，どのレベルの安全性である，という位置づけも明確になるだろう．

また，多種・多様なデータを利用することから，評価システムに求められるシステム的なパフォーマンスの問題や，それらの評価を漏洩せずに保持しておくシステムなども必要となるため，更に実データを用いた検証なども必要とされる．

今後も PWSCUP を通じて得られた指標に関する検討結果を社会に還元できるよう，検討を続けていきたい．

謝辞

本指標の検討にあたり，菊池先生，佐久間先生，山口様，濱田様，山岡様をはじめ PWSCUP 2015 の実行委員の皆様，及び，中川先生より助言，激励を頂き，厚く御礼申し上げます．

参考文献

[1] 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(平成 27 年法律第 65 号)

[2] 菊池 浩明,山口 高康,濱田 浩気,山岡 裕司,小栗 秀暢,佐久間 淳, "匿名加工・再識別コンテスト Ice & Fire の設計",コンピュータセキュリティシンポジウム 2015 論文集, 2015(3), pp.363-370,2015-10-14

[3] J.Domingo-Ferrer,S.Ricci,J.Soria-Comas, "Disclosure Risk Assessment via Record Linkage by a Maximum-Knowledge At-tacker", 2015 Thirteenth Annual Conference on Privacy, Security and Trust (PST), IEEE,2015.

[4] 秋山 裕美,山口 幸三,伊藤 伸介,星野 なおみ,後藤 武彦, "教育用擬似ミクロデータの開発とその利用～平成 16 年全国消費実態調査を例として～", 統計センター製表技術参考資料,16,pp.1-43,2012

[5] L.Sweeney, "k-anonymity: a model for protecting privacy", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, pp.557-570, 2002

[6] 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(CSS2009) 論文集,2009,pp.1-6,2011-10-12

[7] Machanavajjhala, A., Kifer, D., Gehrke, J. and Venkitasubramaniam, M., "l-diversity: Privacy beyond k-anonymity", ACM Transactions on Knowledge Discovery from Data (TKDD), 1(1), pp.3, (2007).

[8] Fung, B., Wang, K., Chen, R. and Yu, P.S., "Privacy-preserving data publishing: A survey of recent developments", ACM Computing Surveys (CSUR), 42(4), pp.14, (2010).

[9] J. Domingo-Ferrer and V. Torra, "A quantitative comparison of disclosure control methods for microdata. Confidentiality", Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies, pp.111-133, (2001).

[10] Sweeney, L., "Guaranteeing anonymity when sharing medical data, the Datafly System.", Proceedings of the AMIA Annual Fall Symposium, pp.51, (1997).

[11] 村本 俊祐, 上土井 陽子, 若林 真一, "データを極小歪曲し k-匿名性を保持したデータに変換するプライバシー保護アルゴリズム", 日本データベース学会 letters, 6(1), pp.97-100, (2007).

[12] 鈴木正朝, 改正個人情報保護法の解説, “URL: http://www.agent.cyber.niigata.jp/pdf/antisocial/forum2015_1.pdf”, (2015).