パーソナルデータの安心・安全な利活用を支えるプライバシー保護技術

Privacy Protection Technology to Support Safe and Secure Utilization of Personal Data

● 山岡裕司　● 伊藤孝一　● 伊豆哲也　● 津田　宏

あらまし

スマートフォンやIoT(Internet of Things)の普及により，あらゆる人，モノのデータを収集し，つなげることが可能となった結果，これらのデータを利活用するビジネスの重要度が更に高まると期待されている。特にパーソナルデータは，デジタルマーケティングの実現に必須となる要素であるが，個人のプライバシーに関わるものであるため，取り扱いには十分に注意する必要がある。社会保障・税番号制度(マイナンバー制度)の施行に合わせて，プライバシー保護とパーソナルデータ利活用の両立を図るために個人情報保護法の改正が行われた。この改正によって，個人情報取り扱いの規制が強化される一方，本人を特定できない「匿名加工情報」に変換することにより，本人の同意なしで第三者に提供できるようになった。

本稿では，富士通研究所が開発したパーソナルデータの安心・安全な利活用を行うためのプライバシー保護技術である仮名化(かめいか)やk-匿名化などの匿名化技術，匿名化技術を適切に運用するためのプライバシーリスク評価技術，およびこれらの技術の応用例や利用シーンについて紹介する。

Abstract

The diffusion of smartphones and the Internet of Things (IoT) has made it possible to gather and link together the data of all people and things and, as a result, the importance of business that utilizes these data is expected to grow further. Personal data, in particular, are an element essential to the realization of digital marketing but their handling requires careful attention because they may have a serious effect on the privacy of individuals. In Japan, as the Social Security and Tax Number System (Individual Number System) was enforced, the Act on the Protection of Personal Information was amended in order to achieve both protection of privacy and utilization of personal data. This amendment has tightened the regulations for handling personal information and, at the same time, allowed the information to be provided to third parties without obtaining the owner's consent through conversion to anonymously processed information that cannot be used to identify the owner. This paper presents anonymization technologies such as pseudonymization and k-anonymization that are privacy-protection technologies developed by Fujitsu Laboratories for safe and secure utilization of personal data, and the privacy risk assessment technology for appropriate operation of anonymization technology. It also provides application examples and scenes of use of these technologies.

ま え が き

スマートフォンやIoT（Internet of Things）の普及により，あらゆる人，モノのデータを収集し，つなげることが可能となった結果，これらのデータを利活用するビジネスの重要度が更に高まると期待されている。特に，個人に関連したパーソナルデータの利活用は，デジタルマーケティングの実現に必須となる重要な要素である。その一方で，パーソナルデータには，他人に知られたくない情報が含まれるため，取り扱いには十分な注意が必要である。2013年に，JR東日本がSuicaの乗車履歴データを匿名化して販売するビジネスを発表したところ，多くの利用者から批判を浴びてサービス停止に追い込まれたことを記憶している人も多いだろう。(1)

近年，パーソナルデータの取り扱いルールを明確化する法制度の整備が，世界中で進んでいる。欧州連合（EU）では，各国の法規制整備を求めていた従来のEUデータ保護指令に代わり，全体の統一的なパーソナルデータ保護規制である「EUデータ保護規則」が2016年4月に採択され，2018年5月の施行が予定されている。日本では，2005年に施行された「個人情報保護法」が，社会保障・税番号制度（マイナンバー制度）導入に伴うパーソナルデータの種類や，パーソナルデータの利活用ルールをより明確化するために2015年に改正され，2017年5月に全面施行となった。(2)

この「改正個人情報保護法」では，パーソナルデータを匿名加工することにより，一定の制約のもとで，本人の同意がない場合でも第三者提供が認められるようになった。この「匿名加工情報」は，特定の個人を識別できず，かつ復元が不可能となるように加工された情報である。データ提供元は，匿名加工情報から個人が特定されないよう適正な加工を行うことが義務づけられているが，個人が特定されるリスクを評価し対策するのは容易ではない。海外事例なども考慮すると，専門家による評価を必須とした場合，データを提供するまでに多くの時間がかかる懸念があることが課題となっていた。

そこで富士通研究所では，個別の匿名化手法を実現するための匿名化技術と，匿名化手法に応じて変化するプライバシーリスクとデータ情報量を自動評価するプライバシーリスク評価技術を開発した。(3) この技術を用いることにより，個人特定の可能性に関するリスクを従来より短時間で評価できるようになった。

本稿では，富士通研究所の匿名化技術とリスク評価技術について紹介する。

個人情報保護制度の動向

日本では2005年4月に個人情報保護法が施行されたが，個人情報の定義や取り扱い規則が曖昧で，個人情報を全体として監督する組織もなかった。2013年7月には，前述のSuicaの事例が発生している。この影響もあり，マイナンバー制度の施行に合わせた2015年に個人情報保護法が改正され，2017年5月に全面施行となった。この改正により，個人情報の定義が明確化され，これまで対象外であった5,000人分以下の個人情報を扱う小規模事業者も本法の対象となった。監督機関として個人情報保護委員会が設置され，データベース提供罪などの新たな罰則も追加された。

規制を強化する一方，第三者提供することの公表義務や提供先での個人再識別禁止義務のもと，本人の同意なしで匿名加工情報を第三者提供することを認めている。これによりパーソナルデータの利活用が促進され，新産業・新サービスの創出が期待される。

匿名化技術

パーソナルデータを第三者提供が可能な匿名加工情報にするには，データの一部を変換して，個人を特定できないデータにする必要がある。(4) このような加工を実現する匿名化手法を総称して「匿名化技術」という。本章では，代表的な匿名化手法である「仮名化（かめいか）」と「k-匿名化」，および匿名化技術の利用シーンを紹介する。

● 仮名化とk-匿名化

仮名化は，実名を仮名に置き換えることで個人特定を防ぐ手法である（**図-1**）。この手法は，「田中」という氏名が「ID23052」に置き換わるため，仮名化後も同一人物を追跡できるところが利点であるが，性別，年齢などの属性の組み合わせから，個人が特定されてしまう危険がある。例えば図-1

氏名	性別	年齢	疾患
田中 平助	男	103	糖尿病
吉田 貴子	女	83	がん
山本 洋介	男	23	肺炎
鈴木 三郎	男	26	肺炎

仮名化

削除	仮ID	性別	年齢	疾患
田中 平助	ID23052	男	103	糖尿病
吉田 貴子	ID47353	女	83	がん
山本 洋介	ID98124	男	23	肺炎
鈴木 三郎	ID83041	男	26	肺炎

図-1　仮名化の例

氏名	性別	年齢	疾患
田中 平助	男	103	糖尿病
吉田 貴子	女	83	がん
山本 洋介	男	23	肺炎
鈴木 三郎	男	26	肺炎

k-匿名化

準識別子

削除	性別	年齢	疾患
田中 平助	―	80歳以上	糖尿病
吉田 貴子	―	80歳以上	がん
山本 洋介	男	20歳代	肺炎
鈴木 三郎	男	20歳代	肺炎

k＝2の場合　準識別子（性別，年齢）が同じデータが2件以上あるように加工

図-2　k-匿名化の例

の場合，「男性，103歳」に該当する個人が「田中 平助」さんであることを知っている人には，どのレコードが田中さんに対応するか特定できてしまう。

この仮名化の欠点を回避する手法がk-匿名化である（**図-2**）。k-匿名化は，組み合わせによって個人特定が可能となる属性を準識別子として定義し，同一の準識別子の組み合わせが必ずk人以上存在するようにデータを加工する手法である。図-2は，k＝2，準識別子を性別，年齢と定義した例で，性別，年齢が同じ組み合わせの人が二人以上となり，どのレコードが田中さんであるかを特定できないようにしている。

富士通研究所では，ゲートウェイを通すことで既存のプログラムを変更することなく透過的に実行可能な仮名化処理と，省メモリのサーバでも高速なk-匿名化処理を実現するための技術を開発した。[(5)] 富士通は，この技術を「FUJITSUビジネスアプリケーションNESTGate」として製品化している。[(6), (7)]

●　匿名化技術の利用シーン

匿名化技術の利用シーン例を**図-3**に示す。これ

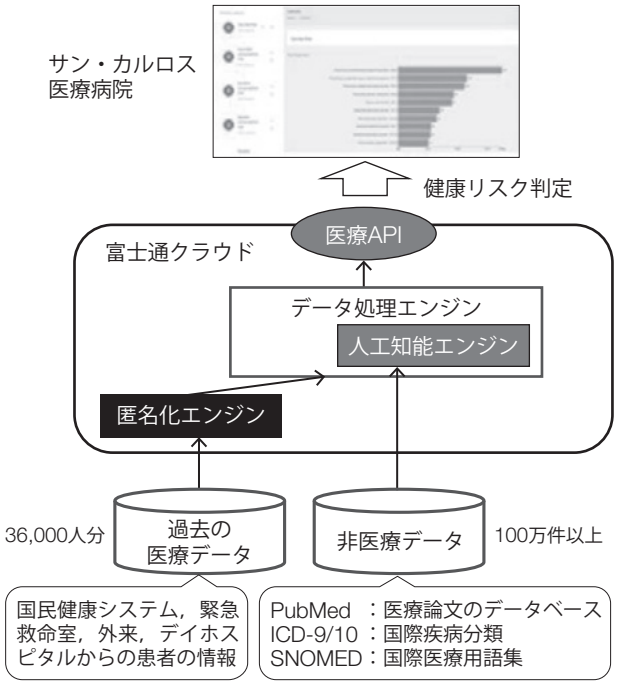

図-3　匿名化技術の利用シーン

は，スペイン・マドリッドのサン・カルロス医療病院，欧州富士通研究所，富士通スペインの3者で行った，医師の迅速な意思決定プロセスを支援す

るための実証実験である。[8]

36,000人以上の精神病患者における過去の医療データと，医療関係の学術論文などのオープンデータを100万件以上取り込んだデータベースを人工知能エンジンで解析することで，想定される患者の健康リスク（自殺，アルコール依存，薬物依存など）を数秒で表示するシステムを開発した。患者の過去の医療データは，富士通研究所の匿名化技術を適用して個人を特定できないようにし，プライバシーを保護している。

プライバシーリスク評価技術

第三者に匿名加工情報を提供するに当たっては，匿名加工基準を満たした加工を行う必要がある。匿名加工基準は，個人情報保護法の施行規則[9]の第19条に定義されている。この規則に対するガイドラインが個人情報保護委員会から公開されており，[4]以下の条件を満たす必要がある。

(1) 氏名や，住所・生年月日の組み合わせなど，特定の個人を識別できる情報を削除もしくは規則性のない方法で変換。
(2) 免許証・被保険者番号などの個人識別符号を削除もしくは規則性のない方法で変換。
(3) そのほか個人を相互連結可能なIDを削除もしくは規則性のない方法で変換。
(4) 特異な値を削除もしくは規則性のない方法で変換。
(5) (1) ～ (4) とは別に，加工対象のデータベースの性質を鑑み，適切な処置を行う。例えば，購買情報や位置情報など，単一では個人識別につながらないが，情報の蓄積により個人識別につながる場合も鑑みて適切な加工を行う。

上記の条件を満たすためには，匿名加工情報から個人が特定されるリスクを評価する必要がある。氏名は仮名化しても，そのほかの属性の組み合わせで個人が特定されてしまう可能性があるが，この評価は容易ではない。例えば，医療機関が持つデータを集め，集めたデータを匿名化して医療研究に利用するケースにおいて，国外の医療機関ではデータが提供できるまで半年以上かかるケースも報告されている。[10]このような事例を踏まえると，全ての属性の組み合わせについてリスクを評価しようとすると，膨大な時間がかかることが予想される。

そこで，富士通研究所はデータの分布に基づき，最も個人を特定しやすい属性の組み合わせを優先的に抽出して探索する技術と，データ中で最も個人を特定しやすい属性の組み合わせを分析し，その特定しやすさ（容易度）を定量化して比較する技術を開発した。その結果，14属性，1万人のデータに対して，3分以内という現実的な時間内での自動リスク評価が可能となった。開発したプライバシー評価技術の特徴を**図-4**に示す。

これらの技術に基づくリスク評価の自動化によって，優先的に匿名化すべき属性が短時間で分かるようになった。加えて，データが漏えいした場合の想定損害賠償額の算出や，匿名加工情報に関するガイドラインへの適合性を判定する技術を開発した。損害賠償額の算出は，日本ネットワークセキュリティ協会が作成したJOモデルに従っている。[11]

富士通研究所では，これらの技術を利用したプ

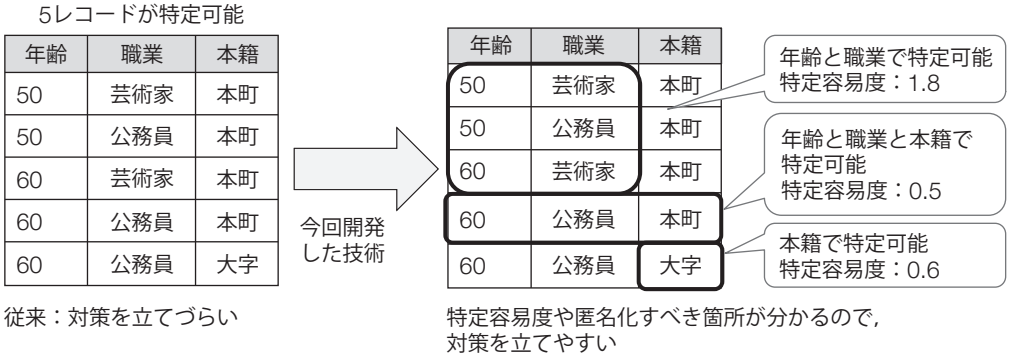

図-4　開発したプライバシーリスク評価技術の特徴

ライバシーリスク評価ツールを2017年度を目処に実用化する予定である。

プライバシーリスク評価技術利用シーン

本章では，プライバシーリスク評価技術の想定利用シーン例を3件述べる。

(1) 匿名化データによる異業種連携

例えば，移動履歴データを匿名加工し，マーケティングに利用する目的で提供することが考えられる（**図-5**）。

本プライバシーリスク評価技術を用いることで，匿名加工されたデータがガイドラインを満たしているか，個人特定される可能性が十分低いデータであるか，情報漏えい時の損害賠償額はどれくらいか，などのリスクを評価できる。リスクが問題ないレベルであると判断された場合にのみ，データ提供を行う。

(2) 医療情報の利活用

2018年の導入を目標に内閣官房が進めており，複数の医療機関のデータを医療情報匿名加工・提供機関（仮称）に集約して匿名加工し，製薬企業などに提供する（**図-6**）。この機関には，匿名加工基準を十分に満たした匿名加工技術と，医療情報の円滑な利活用のための標準や品質水準への対応が求められる。本プライバシーリスク評価技術を用いることで，その要件を満たして，プライバシー保護とデータ利活用の両立が期待できる。

(3) 匿名加工を施したパーソナルデータの目的外利用

改正個人情報保護法で認められている複数の目的で取得した自社データに対し匿名加工を行うことで，別の目的に利用できるようになる（**図-7**）。

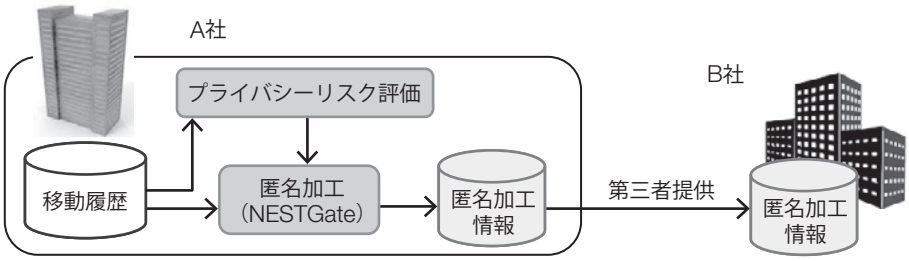

図-5　匿名化データ提供による異業種間連携

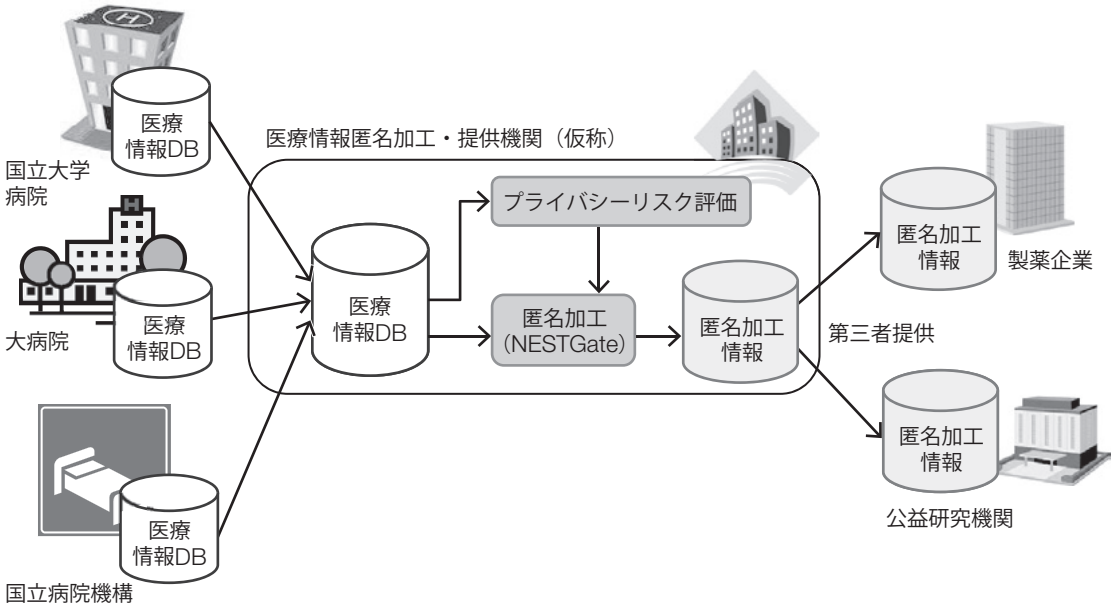

図-6　医療情報匿名加工・提供機関(仮称)による匿名化データ提供

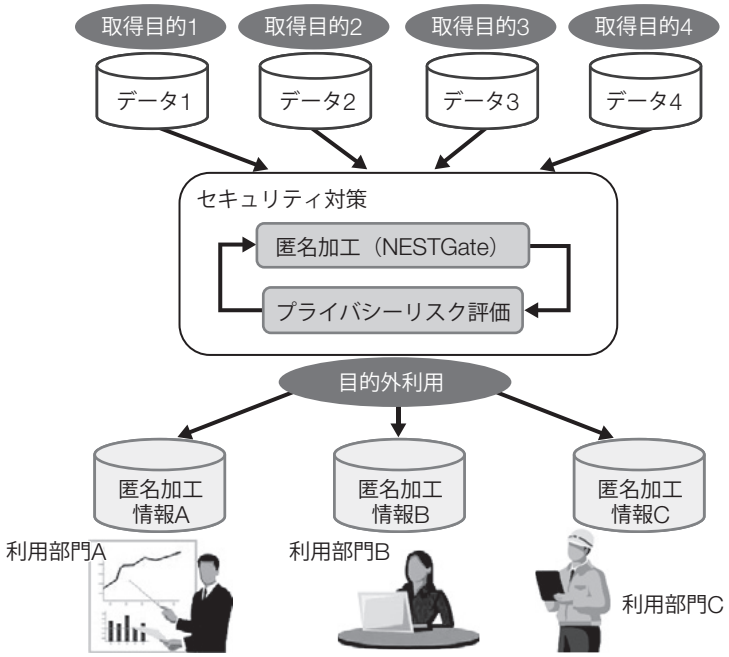

図-7　自社での目的外利用のための匿名加工

本プライバシーリスク評価技術を用いることにより，匿名加工されたデータが分野ごとのガイドラインを満たしているか，情報漏えい時の損害賠償額はどれくらいか，個人特定の可能性が十分低いデータになっているか，などのリスク評価を行う。これとともに，データの情報量を評価することで，提供先となる利用部門の利活用目的に対し十分な品質であるかどうかを確認する。これらが問題ないレベルであると判断された場合にのみ，利用部門にデータを提供する。

む　す　び

本稿では，改正個人情報保護法における匿名加工情報に対応するための匿名化技術と，匿名化技術を適切に運用するためのプライバシーリスク評価技術について，富士通および富士通研究所の取り組みを紹介した。

今後も，安心・安全なパーソナルデータ利活用を実現するため，個人情報保護制度に則した匿名化が効果的・効率的に実行できるよう，プライバシー保護技術の実用化を進めていく。

参考文献

(1) 日経ビジネス：Suica販売履歴の失策.
http://business.nikkeibp.co.jp/article/opinion/20140718/268916/?rt=nocn
(2) 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9月3日成立・同月9日公布）.
http://www.kantei.go.jp/jp/singi/it2/pd/info_h270909.html
(3) 富士通研究所：業界初！パーソナルデータのプライバシーリスクを自動評価する技術を開発.
http://pr.fujitsu.com/jp/news/2016/07/19.html
(4) 個人情報保護委員会：個人情報保護法ガイドライン（匿名加工情報編）.
https://www.ppc.go.jp/files/pdf/guidelines04.pdf
(5) 伊藤孝一ほか：パーソナルデータのプライバシー保護を実現する匿名化・暗号化技術．FUJITSU，Vol.67，No.1，p.26-33（2016）.
http://www.fujitsu.com/jp/documents/about/resources/publications/magazine/backnumber/vol67-1/paper05.pdf
(6) 森沢宜広ほか：k-匿名化技術によりパーソナルデータ保護を実現するNESTGate．FUJITSU，Vol.67，No.1，p.34-40（2016）.
http://www.fujitsu.com/jp/documents/about/

resources/publications/magazine/backnumber/vol67-1/paper06.pdf
(7) 富士通：FUJITSUビジネスアプリケーション NESTGate.
http://www.fujitsu.com/jp/group/fwest/products/software/applications/nestgate/
(8) 富士通：AIで精神病患者の命を救う！サン・カルロス医療病院様との実証実験で医師の診断時間半減に成功. FUJITSU JOURNAL, (2016年12月).
http://journal.jp.fujitsu.com/2016/12/09/01/
(9) 個人情報保護委員会：個人情報の保護に関する法律施行規則.
https://www.ppc.go.jp/files/pdf/290530_personal_commissionrules.pdf
(10) 財団法人日本情報処理開発協会：パーソナル情報の利用のための調査研究報告書.
https://www.jipdec.or.jp/archives/publications/J0004479.pdf
(11) 特定非営利活動法人日本ネットワークセキュリティ協会セキュリティ被害調査ワーキンググループ：2003年度情報セキュリティインシデントに関する調査報告書.
http://www.jnsa.org/houkoku2003/incident_survey2.pdf

著者紹介

山岡裕司（やまおか　ゆうじ）
セキュリティ研究所
データ＆IoTセキュリティプロジェクト
データの匿名化およびプライバシーリスク評価技術に従事。

伊藤孝一（いとう　こういち）
セキュリティ研究所
データ＆IoTセキュリティプロジェクト
日本やEUのプライバシー保護技術の研究開発に従事。

伊豆哲也（いず　てつや）
セキュリティ研究所
サイバーセキュリティプロジェクト
サイバーセキュリティの研究開発に従事。

津田　宏（つだ　ひろし）
セキュリティ研究所
データ＆IoTセキュリティプロジェクト
ブロックチェーン，プライバシー保護，IoTセキュリティの研究開発に従事。