

匿名加工と プライバシー保護

編集にあたって

—匿名加工による個人に関する情報の利活用—

菊池浩明 | 明治大学／理化学研究所

須川賢洋 | 新潟大学

ビッグデータ活用の時代到来？

2017年5月に全面施行された改正個人情報保護法には「匿名加工情報」という概念が新たに追加された。改正により個人情報を含むデータの取り扱いにより厳密になる一方で、匿名加工情報については本人同意を必要とせず、さまざまな目的を持つ第三者に提供してビッグデータとして活用することが可能となる。従来から統計データへ加工すれば本人同意は不要であったが、これからは誰のデータか分からないように加工すれば、本人に許可を得なくても販売してもよい道が用意された。これらの関係を、表-1に整理しよう。すでに、国や自治体を持つデータについては、2016年12月に官民データ活用推進基本法が施行されて、オープンデータとしての活用が推奨されている。診療報酬明細などの医療情報に至っては、匿名加工医療情報と呼ぶデータに加工して活用する次世代医療基盤法が2017年5月に公布され、2018年には施行予定である。最先端IT国家といえるかどうかはさておき、ビッグデータを官民で活用する基盤が整備されてきたことは確かである。

匿名加工の現実

国内ではすでにいくつもの匿名加工の応用事例が始まっている。認定個人情報保護団体の1つとして会員の情報保護指針などを先導する日本情報経済社会推進協会(JIPDEC)は、匿名加工情報を検討する事業者の参考になるようにと、匿名加工情報の事例集¹⁾を2017年7月に公開している。表-2にそのいくつかの例を挙げる。匿名加工された個人情報がどのように活用されていくのか、具体的な加工方法も含めて紹介されており有益である。専門家の助言を基に安全に加工が行われており、加工済みのデータから特定の個人が再識別されるリスクは低い。

再識別されないように加工することは、実はそんな

表-1 第三者提供の種類

情報の種類	同意	加工方法
非個人情報	不要	無加工
	本人同意あり	
個人情報	本人同意なし	統計データ
		匿名加工
		その他

なに難しくない。個人情報保護委員会は、個人情報の保護に関する法律施行規則（平成 28 年個人情報保護委員会規則第 3 号）や通則ガイドライン、匿名加工情報編²⁾などを整備して、加工の基準を示している。JIPDEC の加工事例はこれらの基準に従っていて、名前や会員番号などはすべて削除されており、安全性は高い。しかしながら、難しいのは、有用性を保ちながら必要十分な加工を行うことである。たとえば、文献 2) の事例 3（購買履歴情報の匿名加工）では、図-1 のような加工がなされている。氏名や会員番号は削除され、生年月日は年代に丸められ、購買日は曜日に置き換えられている。これだけ加工しても、商店街に新規出店を検討している事業者には十分な情報である。しかしながら、購買金額は 200 円ごとの区分に置き換えられているので、平均値などには誤

差が生じている。購買日が曜日に置き換えられているので、月ごとの売り上げの変化を見ることができない。会員番号が削除されているので、1 人あたりの平均購買額や累積顧客数などの統計が得られない。すなわち、それらの有用性を保証して、かつ、誰も再識別されないように加工することは容易ではない。データサイエンスの非常に高度な技術と知識が必要であろう。

匿名加工・再識別コンテスト

そこで、本会 CSEC 研究会では、プライバシーワークショップ（PWS）実行委員会を組織して 2015 年より PWS Cup『匿名加工・再識別コンテスト』を開催している。2017 年は台湾・カナダを含む企業や大学など 21 チーム合計 73 名の参加申し

表-2 JIPDEC 匿名加工情報の事例集

事例	事業者	データ例	備考
1. 所有車データ提供	整備工場が、自動車販売店に対して提供	顧客（数万） 車両（数万） 整備（数百万）	
2. 顧客データ提供	質屋が調査会社に提供	顧客（数千） 取引（数十万）	
3. 購買履歴の提供	商店街が、新規出店事業者提供	顧客（数千） 購買（数十万）	13 カ月
4. 移動履歴	経路サービス事業者が、自治体の委託により、駐輪場事業者提供	顧客（数千） レコード（数千万）	5 人以上の通行者がいる部分を可視化

氏名	会員番号	誕生日	性別	住所	購買日時	購買店	購買額
甲野太郎	0001	1963/1/1	男	S 区 A 町 4 丁目 19 番 1 号	12 月 5 日	D 魚店	300 円
甲野太郎	0001	1963/1/1	男	S 区 A 町 4 丁目 19 番 1 号	12 月 6 日	E 豆腐店	130 円
乙山次郎	0002	1974/2/2	男	T 区 B 町 1 丁目 3 番 2 号	12 月 5 日	D 魚店	300 円
乙山次郎	0002	1974/2/2	男	T 区 B 町 1 丁目 3 番 2 号	12 月 6 日	F 傘店	840 円



氏名	会員番号	誕生日	性別	住所	購買日時	購買店	購買額
(削除)	(削除)	50 代	男	S 区 A 町 4 丁目	月曜	魚店	200-400
		50 代	男	S 区 A 町 4 丁目	火曜	豆腐店	200 未満
		40 代	男	T 区 B 町 1 丁目	月曜	魚店	200-400
		40 代	男	T 区 B 町 1 丁目	火曜	傘店	800 以上

図-1 事例 3：加工方法（文献 1）を基に修正）

込みがあり、このうち 14 チーム約 40 名が 10 月 23 日開催の本戦 PWS Cup にて技術を競った。表-3 に 3 年間のコンテストの概要を示す。

また「セキュアなデータ結合に関する技術・制度の方向性」というタイトルで個人情報保護法の匿名加工情報との関連のパネルディスカッションも行った。本小特集では、これらの 3 年間の活動によって蓄積された匿名加工に関する研究成果について報告する。

「匿名加工とは何か—基本技術とリスク—」(菊池)は、匿名加工に用いるいくつかの基本技術の定義とリスクの種類を整理する。多くの技術がある中で、各国の法制度は多様である。「匿名化に関する制度の国際的な動向」(美馬)では、これまでに本分野を先導してきた英国、米国の法制度の歴史を整理し、韓国、中国、欧州の最新動向を加えている。世界的な法制度整備が進む中、我が国の匿名加工にかかわる制度がどうなっているのか、「我が国における匿名加工の法制度—法律からガイドライン、事務局レ

ポートまで—」(須川)において解説する。

前述の PWS Cup については、コンテスト実行委員会を代表して、「匿名加工・再識別コンテスト—世界唯一の対戦型データ匿名加工コンテスト PWS Cup—」(小栗)が詳しく報告する。「匿名加工再識別コンテストの加工アルゴリズム—PWS Cup 2017 優勝チームより—」(濱田)は、2017 年度の優勝チームのリーダーが加工のコツを伝授する。限られた紙面の中、原理のみを分かりやすく説明しており、これから加工アルゴリズムを研究しようとする人は必読である。最後に、「再識別リスク—匿名化の再識別リスクの考え方—」(野島)では、再識別する側の技術、すなわち、加工データの安全性を正しく評価する方法を解説する。

参考文献

- 1) 匿名加工情報の事例集, JIPDEC (2017), https://www.jipdec.or.jp/protection_org/u71kba0000001hh-att/AOP_006.pdf
- 2) 個人情報の保護に関する法律についてのガイドライン (匿名加工情報編, 平成 28 年個人情報保護委員会告示第 9 号), 個人情報保護委員会 (2016)。

(2018 年 2 月 28 日)

表-3 PWS Cup 匿名加工コンテスト

開催年	2015	2016	2017
開催	10/21 ~ 22 長崎ブリックホール	10/11 ~ 12 秋田キャッスルホテル	10/23 ~ 24 山形国際ホテル
参加者数	13 チーム (20 名)	15 チーム (42 名)	14 チーム (43 名)
データセット	疑似マイクロデータ (世帯消費額)	UCI Dataset "Online Retail" (購買履歴)	
属性数	25	11 (顧客 4 属性 + 履歴 7 属性)	
顧客数	8,333	400	500
履歴数	なし	18,524	44,917
履歴期間	1 年間	1 年間	1 年間 (12 カ月)
参加募集ポスター			