Computer Security Symposium 2018
22 - 25 October 2018

自己情報コントロールと権限分散を両立するパーソナルデータ流通方式の提案

山岡 裕司[1,a)]　前田 若菜[1]

概要：パーソナルデータの流通・活用への期待が高まる一方，プライバシー保護の重要性も増す中，日本政府により個人の関与の下でデータ流通・活用を進める仕組みである情報銀行等の社会実装が推進されている．我々はこれまでに，データ流通・活用の活発化を阻害する，権限集中によるリスクと，データの信頼性への不安という 2 大課題を解決する方式を提案した．しかし，流通させるデータ項目の組み合わせを個人が制御できないという課題が生じてしまっていた．本稿では，その課題を解決する拡張方式を提案する．本方式では，データ提供者それぞれが，個人から許可規則を受け，仲介者を経由してきた購入申請のデータ項目組み合わせがその許可規則に合っていた場合のみ購入申請に応じる．暗号技術を利用することで，たとえ仲介者が許可規則を無視した不正な購入申請をしたとしても，それにより取得できるデータを制限できる．本稿ではデータ流通・活用の活発化のためのプライバシー要件を 4 つに整理し，本方式がそれらを満たしていることを示す．

キーワード：プライバシー保護，データ流通，情報銀行，権限分散，仮名化

A Proposal for Personal Data Distribution Coping with Control over Information and Authority Decentralization

Yamaoka Yuji[1,a)]　Maeda Wakana[1]

1. はじめに

パーソナルデータは「インターネットの新たな石油，デジタル世界の新たな通貨」[4] といわれ，近年その流通・活用への期待が高まっている．一方，パーソナルデータの流通・活用に伴い，データ主体である個人のプライバシーについての事案が発生している．そのためプライバシー保護の重要性も増しており，たとえば欧州連合ではパーソナルデータの処理に関する保護を基本的人権とする GDPR (General Data Protection Regulation) が成立・適用された．日本政府は，パーソナルデータを安全に流通させ，データ流通を拡大させ，超少子高齢社会における諸課題の解決につなげるべく，2017 年に IT 戦略「世界最先端 IT 国家創造宣言・官民データ活用推進基本計画」を実施し，2018 年の「世界最先端デジタル国家創造宣言・官民データ活用推進基本計画」でもデータ流通を促進する取り組みを継続している．

その中で，個人の関与の下でデータ流通・活用を進める仕組みである情報銀行等の社会実装が進められている．情報銀行とは，「個人とのデータ活用に関する契約等に基づき，PDS 等のシステムを活用して個人のデータを管理するとともに，個人の指示又は予め指定した条件に基づき個人に代わり妥当性を判断の上，データを第三者（他の事業者）に提供する事業」[9] である．総務省の調査では，消費者は，提供するパーソナルデータが商業目的に活用される場合，「自分への経済的なメリット（割引等）」があるなら提供を最も許容しやすい [8]．そのため，情報銀行の典型的なユースケースとして，個人が自身のデータを第三者に販売するために，情報銀行に仲介を依頼することが考えられ

[1] 株式会社富士通研究所
FUJITSU LABORATORIES LTD.
[a)] yamaoka.yuji@jp.fujitsu.com

る．情報銀行の実現方式のうち，パーソナルデータとして個人が内容を制御できるもののみを取り扱うものを，本稿では**情報銀行方式**と呼ぶ．

情報銀行方式は現行法下でも実現可能ともいわれる一方，データ流通・活用の活発化を阻害する様々な課題が指摘されているが，我々はまず次の 2 つに着目した．

権限集中によるリスク　情報銀行方式は販売可能な価値の高いパーソナルデータを管理しており，それら全データにアクセス可能なため，仲介者（情報銀行事業者）が悪意ある者に制御されると全データが漏洩する恐れがある．価値の高いパーソナルデータは機微性が高い場合も多く，プライバシーへの影響が大きい．仲介者への権限集中をリスクと考える個人もおり，その傍証として，アンケート結果において情報銀行を利用したくない理由の第 2 位が「自分の情報を集約すると漏洩した場合が怖いから」となった調査結果 [6] がある．

データの信頼性への不安　販売データの内容を個人が制御できると，高く売れるデータを偽造する個人の参入が想定され，データの信頼性が低くなる．日本政府の検討会でも，データ流通・活用においてデータの信頼性は必要 [9] とされている．

我々はこれらの課題を解決すべく，すでに個人の信頼を得てパーソナルデータを収集・管理している事業者を信頼点とし，その事業者が個人のデータ販売の代理人となり，さらに仲介者を通して受領者に販売する方式を提案した [7]．仲介者は需給マッチングなどが主な役割で，販売データを直接取り扱わないため，たとえ悪意ある者に制御されてもプライバシー影響は限定的である．また，事業者が管理しているデータは，その事業で実際に使用しているという意味で信頼性が高い．事業者が高く売れるデータを偽造することは可能だが，各個人顧客の代理としてデータ販売している立場であり，偽造を含め不正行為は顧客からの信頼を損なう要因となるため，不正をおこなうインセンティブは低い．事業者は受領者にデータを販売（提供）するため，以降では提供者と呼ぶ．また，本方式を以降では ***SCIS2018* 方式**と呼ぶ．

SCIS2018 方式の特長の一つは，横串データの売買が可能なことである．横串データとは，複数の提供者が管理しているそれぞれのデータ同士を，ある個人同一人物で紐付けたデータのことである．たとえば，店舗 O_1 が管理する購買履歴と，携帯キャリア O_2 が管理する移動履歴について，それぞれに個人 P_1 のデータがあったとして，P_1 の両データを同一人物データとして共通 ID などで紐付けたものが横串データである．この例の場合，この横串データの受領者は，ある人物がどこから店舗 O_1 に来て何を買ってどこへいったか，といったことがわかるようになる．横串データは，単一の提供者では提供できないデータであり，より価値の高いデータであると考えられる．なお，原則的に横串データの共通 ID は仮 ID（仮名）であり，受領者にとって個人 P_1 を特定する必要はないとする．

しかし，SCIS2018 方式では，販売データ項目の組み合わせを個人が制御できないという課題が生じてしまっていた．たとえば，店舗 O_1 は，詳細な情報である購買履歴を販売する代わりにプライバシーに配慮して，より一般化した月次購入額を販売できるようにすることもできる．このとき，個人 P_1 は，携帯キャリア O_2 にある移動履歴とは，店舗 O_1 にある購買履歴を組み合わせ横串データにして販売したいと考えるかもしれない．一方，健康サービス O_3 にある健診情報とは，店舗 O_1 にある購買履歴は横串データにせず，代わりに月次購入額を横串データにして販売したいと考えるかもしれない．情報銀行方式においては，このように，販売するデータの項目を個人が制御できるのは当然のことであった．しかし，SCIS2018 方式では，横串データを作る提供者の数を制御することはできたものの，データ項目については制御できず，上例のような販売意向を実現できなくなってしまっていた．

本稿では，SCIS2018 方式を拡張し，販売データ項目の組み合わせを個人が制御できるようにした方式を提案する．提案方式では，提供者それぞれが，個人から許可規則，具体的には許可 ID 付き提供許可項目集合を受け，仲介者を経由してきた購入申請のデータ項目組み合わせがその許可規則に合っていた場合のみ購入申請に応じる．これにより，本人が許可規則として設定したデータ項目組み合わせのみが販売されるようになる．個人は，許可 ID は提供者間で同一のものを提供するが，提供許可項目は各提供者にそれぞれ自身が管理するデータの項目のみを提供する．それにより，各提供者に，個人が他のどの提供者のサービスを利用しているか知られなくて済む．他の特徴として，暗号技術を利用することで，仲介者が許可規則を無視した不正な購入申請により取得できる横串データを制限できる．

本稿の貢献は，上記方式の提案であり，提案方式は**表 1** に示す要件を全て実現する初めての方式である．各要件の詳細は次節で述べる．

本稿の以降の構成は次の通りである．第 2 節では，提案方式に求められる要件を整理する．第 3 節では，従来方式として，情報銀行方式，UMA，そして SCIS2018 方式を説明する．第 4 節では，提案方式を説明する．第 5 節では，提案方式について考察する．第 6 節では，関連研究を説明する．第 7 節にてまとめる．

2. 要件整理

我々は，個人の関与の下でデータ流通・活用を進める仕組みに必要な要件を表 1 のように考えた．本節では，これら各要件について，必要な理由も含め説明する．

表 1: 要件と各方式の対応
Table 1 Requirements and correspondence of each method

		情報銀行方式	UMA	SCIS2018 方式	提案方式
機能要件	(F1) 横串データを売買可能	Y	Y	Y	Y
セキュリティ要件	(S1) 提供者による代理販売（個人販売でない）	N	Y	Y	Y
プライバシー要件	(P1) 購入せずに全データにアクセス可能な者なし	N	N	Y	Y
	(P2) 販売データは無名化可能	Y	N	Y	Y
	(P3) 販売データ項目組み合わせを個人が制御可能	Y	Y	N	Y
	(P4) 提供者が特定個人顧客の有無を秘匿可能	N	N	N	Y

2.1 (F1) 横串データを売買可能

前述の通り，横串データは複数の提供者を横断するデータで，単一の提供者では提供できないデータであり，より価値の高いデータであると考えられる．そのため，データ流通・活用の活発化のためには，横串データを売買可能なことは必須であろう．

2.2 (S1) 提供者による代理販売（個人販売でない）

前述の通り，データの信頼性確保は一般的に課題とされている．個人販売できる，より正確には販売データの内容を個人が自由に編集できると，データの信頼性が低くなることが想定されるのは前述の通りである．よって，提供者が実際に使用しており個人が自由に編集できないデータを，提供者が個人の代理として販売することが，セキュリティ要件として望まれる．

2.3 (P1) 購入せずに全データにアクセス可能な者なし

前述の通り，権限集中によるリスクは個人へのアンケートからもうかがえる課題である．個人が全データを横串データとして販売することを許可した場合，それを購入した受領者がその個人の全データにアクセス可能になるのは，個人の意向に沿っているため問題としない．しかし，それ以外の場合には，各者が単独で不正をしたとしても，各個人の全データにアクセス可能な者は当該個人以外に存在しないことがプライバシー要件として望まれる．

2.4 (P2) 販売データは無名化可能

無名化とは，氏名やメールアドレス等，個人 ID 相当のデータを取り除くこととする．消費者アンケート結果において「匿名化」されたデータはそうでないデータに比べ提供意向が約 4〜6 倍高いという調査結果がある [5]．そのため，データ流通・活用の活発化のためには，受領者に個人 ID 相当の情報を渡さないで済むようにできることが，プライバシー要件として望まれる．

2.5 (P3) 販売データ項目組み合わせを個人が制御可能

もし販売データ項目組み合わせを個人が制御可能でないならば，提供者は個人顧客のうち最も保守的な思想に合わせたデータ項目しか販売できなくなる可能性が高い．たとえば，店舗 O_1 が管理する購買履歴の販売を考えると，月次購入額の粒度でしか販売したくないという個人がいた場合，それより詳細な粒度での販売が難しくなる．なぜなら，個人が制御可能でない以上，提供者がデータ項目を決めることになるが，詳細な粒度のデータ項目の提供を決めると，プライバシーを理由に販売を希望しない個人が多くなりデータ売買が成立しなくなりがちなためである．そのため，データ流通・活用の活発化のためには，本要件もプライバシー要件として望まれる．

2.6 (P4) 提供者が特定個人顧客の有無を秘匿可能

特に，データを無名化して販売したい個人は，自分がある提供者の顧客であることを，その提供者以外に知られたくない場合があると考えられる．そのような場合，たとえば提供者 O_1 と O_2 に自分が両方の顧客であることを知らせないとそれらの横串データを販売できない場合には，その個人は販売を諦めるかもしれない．一般的に，不必要な情報提供が少なくなるほど，プライバシーリスクは低減するため，個人の販売意向は増大するはずである．そのため，データ流通・活用の活発化のためには，当該個人の開示意向がない限りは提供者が特定個人顧客の有無を秘匿できることが，プライバシー要件として望まれる．

ただし，本要件の制限として，横串データを販売可能とすると必然的に，提供者が受領者になった（あるいは受領者と結託した）場合にはこの要件を達成するのは困難といえる．後に詳述するデータセットでの照合を行えば，提供者が自身の提供するデータを含む横串データを購入することで，横串データに含まれる各個人が自身のどの顧客であるか特定でき，その顧客の他提供者でのデータが得られてしまう．この制限を技術的に緩和するのは困難と考えられ，緩和する運用としてたとえば仲介者が提供者には販売

しないことが考えられる．

3. 従来方式

個人の関与の下でデータ流通・活用を進める仕組みの従来方式として，前述の情報銀行方式が代表的である．また，関連する標準プロトコルとして UMA がある．さらに，表 1 の要件 (S1), (P1) に対応すべく我々が SCIS2018 方式を提案しているのは前述の通りである．本節では，これらについて説明する．

3.1 情報銀行方式

情報銀行方式は，個人による自身のデータの管理や販売を，仲介者が支援・代行する．日本政府からの関心が高いが，その背景に GDPR で導入されたいわゆるデータポータビリティ権がある．つまり，事業者が管理しているデータについて，データポータビリティ権があれば個人は自身のデータを自分の制御下におけるため，横串データの販売を実現しやすくなってきている．

情報銀行方式における仲介者は，個人の全データの制御権を当該個人から委譲される．よって，個人の指示通りにデータ項目を組み合わせた横串データを作れるし，横串データの無名化もできる．つまり，要件 (F1), (P2), (P3) を満たす．

しかし，他の要件は満たさない．まず，個人販売の支援・代行なため，要件 (S1) を満たさない．次に，権限を持つ故リスクも大きく，仲介者は全データにアクセス可能なため要件 (P1) を満たさない．さらに，典型的には，提供者（事業者）が管理していたデータの場合はその提供者の ID も仲介者が管理すると考えられ，要件 (P4) を満たさない．

3.2 UMA

UMA (User-Managed Access)[2] は Kantara Initiative により開発・更新が進められている，アクセス管理プロトコル標準である．データ流通の手順の概要は次のようになる．

(1) 個人（resource owner）は，提供者（resource owner）が管理しているデータの提供についてのポリシーを仲介者（authorization server）に設定する．ポリシーの内容は，どのデータ項目をどの受領者（requesting party）に提供して良い，といったものである．
(2) 受領者は，特定の提供者に特定の個人のデータを要求する．
(3) 仲介者は受領者を認証し，データ要求がポリシーに合うか否かを提供者に知らせる．
(4) ポリシーに合っている場合，提供者は要求されたデータを受領者に提供する．

UMA は次の要件を満たす．受領者は，複数の提供者からのデータを個人 ID で紐付けて横串データを作れるため，要件 (F1) を満たす．データは提供者から受領者に直接提供されるため，要件 (S1) を満たす．個人はデータ項目組み合わせをポリシーとして設定可能なため，要件 (P3) を満たす．

しかし，他の要件は満たさない．仲介者が不正をし，個人のポリシーを無視して任意アクセスを許可することで，仲介者は全データにアクセス可能なため，要件 (P1) を満たさない．また，個人 ID は秘匿せずに使用されるため，要件 (P2), (P4) を満たさない．

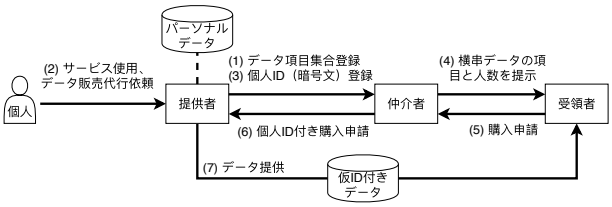

図 1: SCIS2018 方式
Fig. 1 SCIS2018 method

3.3 SCIS2018 方式

SCIS2018 方式 [7] は，上記 2 方式と違い，権限分散により仲介者が悪意ある者に制御されても全データにアクセスされることを防ぐ方式である．すでに個人の信頼を得てデータを収集・管理している事業者を信頼点とし，個人が各事業者に対してそれぞれのデータの提供に関する設定をおこなう．設定には，横串データを作る際のデータ管理元事業者の数の上限を含む．各事業者はその上限を越える横串データの購入申請を拒絶するため，仲介者が個人の全データにまたがる横串データを（購入せずに）入手することはできない．

図 1 は SCIS2018 方式の主な情報の流れを示した図である．各流れに沿った処理内容を次に示す．

(1) 提供者は仲介者に対し，仲介依頼を合意し，提供者 ID o と，販売し得るデータの項目の集合 A_o を仲介者に登録する．A_o は，たとえば A_o = {“購買履歴”, “月次購入額”} といったものである．
(2) 個人は提供者 ID o の提供者に対し，個人 ID p を登録し，サービス等を使用することで，提供者によってパーソナルデータ $T_{o,p}$ が管理されるようにする．さらに，販売代行を依頼し，暗号鍵 k_p と上限数 u_p を登録する．たとえば，提供者が小売店であれば，個人はそこで購買すると，購買履歴が T としてその提供者に管理される．なお，$T_{o,p}$ は個人 ID p を含まない．
(3) 提供者は仲介者に対し，販売代行依頼をしてきた個人 p の，上限数 u_p と，暗号化された個人 ID $p^* = E(k_p, p)$ を登録する．
(4) 仲介者は受領者に対し，全提供者の組み合わせについての販売可能なはずの横串データの情報に相当する，提供者の組み合わせと，各提供者のデータ項目集合と，その組

み合わせでの横串データを販売代行依頼している個人の人数と，の組 $(\{(o_1, A_{o_1}), (o_2, A_{o_2}), \ldots, (o_m, A_{o_m})\}, n)$，ここで n は o_1 から o_m に販売代行依頼しかつ $u_p \leq m$ となる個人の数，を提示する．

(5) 受領者は仲介者に対し，購入申請として提供者の組み合わせと，自身の ID である受領者 ID b との組 $(\{o_1, o_2, \ldots, o_m\}, b)$ を送付する．

(6) 仲介者は購入申請で示された各提供者に対し，受けた購入申請と，その組み合わせでの横串データを販売代行依頼している個人の暗号化された個人 ID と，の組 $(\{o_1, o_2, \ldots, o_m\}, b, \{p_1^*, p_2^*, \ldots\})$ を購入申請として送付する．

(7) 購入申請を受けた提供者 ID o の提供者は，購入申請で示された提供者 ID 集合に自身の ID o が含まれていることと，購入申請で示された各個人 p（p^* を復号して取得）について $u_p \leq m$ であることを確認する．確認できた場合，購入申請で示された受領者 ID b の受領者に対し，購入申請で示された各個人 p について，仮 ID $p^{*+} = E(k_p, p, \{o_1, o_2, \ldots, o_m\})$ とデータとの組 $(p^{*+}, T_{o,p})$ を提供する．

SCIS2018 方式は次の要件を満たす．受領者は，複数の提供者からのデータを仮 ID p^{*+} で紐付けて横串データを作れるため，要件 (F1) を満たす．データは提供者から受領者に直接提供されるため，要件 (S1) を満たす．多数の提供者間にわたる横串データを各提供者から提供を受けようとした場合は購入申請に多数の提供者の ID 集合を含める必要があるが，そうすると上限数 u_p の制限に引っかかる．よって，多数の提供者間にわたる横串データの提供を受けることは単独では誰にもできず，要件 (P1) を満たす．横串データの共通 ID は仮 ID p^{*+} で，暗号鍵 k_p を入手しない限り p^{*+} から個人 ID p を導出できないため，要件 (P2) を満たす．

しかし，他の要件は満たさない．個人は販売データをデータ項目単位で制御できないため，要件 (P3) を満たさない．また，提供者は，受け取った購入申請 $(\{o_1, o_2, \ldots, o_m\}, b, \{p_1^*, p_2^*, \ldots\})$ から，個人 ID を復号し，個人 $\{p_1, p_2, \ldots\}$ が提供者 $\{o_1, o_2, \ldots, o_m\}$ の顧客であることがわかる，つまり他の提供者に特定顧客がいることを秘匿できていないため，要件 (P4) を満たさない．

4. 提案方式

提案方式は，SCIS2018 方式を要件 (P3), (P4) を満たすように拡張した方式である．本節で詳説する．

方式の要点は，提供者それぞれが，個人から許可 ID 付き提供許可項目集合を受け，それに仲介者から受けた購入申請が合っているか確認することで要件 (P3) を達成し，他の提供者には提供許可項目集合の平文を渡さないようにすることで要件 (P4) を達成する点である．

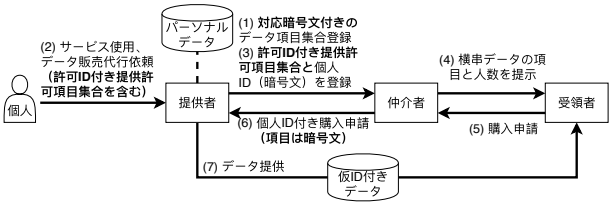

図 2: 提案方式

Fig. 2 Proposed method

図 2 は提案方式の主な情報の流れを示した図である．各流れに沿った処理内容を，SCIS2018 方式からの主な拡張部分を強調して次に示す．

(1) 提供者 ID o の提供者は，**暗号鍵 k_o を用意し**，仲介者に対し，仲介依頼を合意し，提供者 ID o と，販売し得るデータの項目と**その暗号文との組の集合** $A_o^+ = \{(a_{o,1}, a_{o,1}^*), (a_{o,2}, a_{o,2}^*), \ldots\}, a_o^* = E(k_o, a_o)$ を仲介者に登録する．

(2) 個人は提供者 ID o の提供者に対し，個人 ID p を登録し，サービス等を使用することで，提供者によってパーソナルデータ $T_{o,p}$ が管理されるようにする．さらに，販売代行を依頼し，暗号鍵 k_p 一つと**許可 *ID* 付き提供許可項目集合** $(i_p, A'_{o,p}), A'_{o,p} \subseteq A_o$ **を任意の数だけ**登録する．**このとき，個人は横串データを作りたい項目集合については提供者にまたがって同一の許可 *ID* i_p を登録する**．

(3) 提供者は仲介者に対し，販売代行依頼をしてきた個人 p の，**許可 *ID* 付き提供許可項目集合** $(i_p, A'_{o,p})$ と，暗号化された個人 ID $p^* = E(k_p, p)$ を登録する．

(4) 仲介者は受領者に対し，全提供者の組み合わせについての販売可能なはずの横串データの情報に相当する，**提供者をまたがったデータ項目の組み合わせ** C,

$$\begin{aligned} C = \{&(o_1, a_{o_1,1}), (o_1, a_{o_1,2}), \ldots \\ &, (o_2, a_{o_2,1}), (o_2, a_{o_2,2}), \ldots \\ &, (o_m, a_{o_m,1}), (o_m, a_{o_m,2}), \ldots\} \end{aligned}$$

と，その組み合わせでの横串データを販売代行依頼している個人の人数と，の組 (C, n)，ここで n は C **の全データ項目からなる項目集合に同一の許可 *ID* を登録した**個人の数，を提示する．

(5) 受領者は仲介者に対し，購入申請として**データ項目の組み合わせ** C と受領者 ID b との組 (C, b) を送付する．

(6) 仲介者は購入申請で示された各提供者に対し，受けた購入申請を**データ項目について対応する暗号文に置換したもの**と，その組み合わせでの横串データを販売代行依頼している個人の暗号化された個人 ID と**許可 *ID* との組**と，の組 $(C^*, b, \{(p_1, i_{p_1}), (p_2, i_{p_2}), \ldots\}), C^* = \{a_{o_1,1}^*, a_{o_1,2}^*, \ldots, a_{o_2,1}^*, a_{o_2,2}^*, \ldots, a_{o_m,1}^*, a_{o_m,2}^*, \ldots\}$ を購入申請として送付する．

(7) 購入申請を受けた提供者 ID o の提供者は，購入申請

で示されたデータ項目集合の**暗号文のうち復号できるものについては平文の集合 A''_o にし**，購入申請で示された各個人 p（p^* を復号して取得）について，**購入申請で示された許可 ID i_p に紐付けて登録されている提供許可項目集合 $A'_{o,p}$ が $A'_{o,p} \subseteq A''_o$** であることを確認する．確認できた場合，購入申請で示された受領者 ID b の受領者に対し，購入申請で示された各個人 p について，仮 ID $p^{*++} = E(k_p, p, C^*)$ とデータとの組 $(p^{*++}, T_{o,p})$ を提供する．

提案方式は全ての要件を満たす．

- 受領者は，複数の提供者からのデータを仮 ID p^{*++} で紐付けて横串データを作れるため，要件 (F1) を満たす．
- データは提供者から受領者に直接提供されるため，要件 (S1) を満たす．
- 多数の提供者間にわたる横串データを各提供者から提供を受けようとした場合は購入申請に多数の提供者のデータ項目集合を含める必要があるが，提供許可項目集合 $A'_{o,p}$ の制限を受ける．よって，個人の提供許可項目集合を越える横串データの提供を受けることは単独では誰にもできず，要件 (P1) を満たす．
- 横串データの共通 ID は仮 ID p^{*++} で，暗号鍵 k_p を入手しない限り p^{*++} から個人 ID p を導出できないため，要件 (P2) を満たす．
- 個人は，許可 ID 付き提供許可項目集合の登録により，販売データをデータ項目単位で制御できるため，要件 (P3) を満たす．
- 各提供者は，他の提供者の提供者 ID や提供許可項目集合を暗号文でしか受け取れないため，特定個人顧客の有無を知ることができず，要件 (P4) を満たす．

5. 考察

提案方式の制限や，今後の課題について述べる．

制限として，データセットの照合によるプライバシーリスクの増大がある．データセットの照合とは，ID によらないいわゆる名寄せのことで，複数の横串データに対し，同一個人のものと推定されるレコード同士を対応づけてさらに大きな横串データにすることである．履歴データなどのパーソナルデータは個人毎に異なり，各個人のレコードを識別できる状態にあることが多いため，そのような対応づけが実現しやすい．データセットの照合により，個人が販売を意図している横串データの範囲を超えた横串データを作れる可能性がある．しかし，このリスクを技術的に緩和するのは困難と考えられる．なお，緩和する運用としては，個人にそのようなリスクがあることを伝達し，リスクを認識した上で販売データ項目を組み合わせるよう啓蒙することが考えられる．

今後の課題として，入力支援機能の開発がある．セキュリティやプライバシーの機能は多くの個人にとって理解が難しい上，提案方式は暗号鍵を含め複数の情報を個人に設定することを要求するため，個人に利用の敷居が高いと思われてしまう懸念がある．特に，上述のデータセットの照合によるプライバシーへの影響の把握は容易ではないだろう．この課題を緩和するために必要な研究の対象として，プライバシーの影響を理解しやすくするような UI や，AI 技術を利用した設定代行機能などが考えられる．

6. 関連研究

日本政府が整備しているデータ流通・活用を進める他の仕組みとして，匿名加工情報がある．匿名加工情報は 2015 年に改正された個人情報保護法で創設された制度で，提供者は個人情報を加工して匿名加工情報にすることで，一定のルールの下で，個人の同意を得ることなく流通させることができる．日本だけでなく，GDPR 等を含め，多くの国の個人情報保護の法律では，個人情報を匿名化した情報には規制をしていないか規制を緩和している．また，安全管理措置として匿名化やそれに類する加工も有効とされている．そのため，匿名化に関する研究は多く，k-匿名化 [3] や差分プライバシー [1] などがある．匿名化は個人の同意が不要なため，提供意向の個人差等によるデータの偏りが生じない利点がある一方，加工によりデータの有用性が元データに比べ低下しがちという欠点がある．

それに対し提案方式は，個人の同意を得て流通させる方式である．同意さえ取れれば，データの有用性を元データと同じに保てる場合もあること，個人情報保護の法律に抵触するリスクが低いこと，が主な利点である．ただし，先述の通り匿名化されたデータの方が個人の提供意向が高いため，同意を取り易くするために本稿で要件 (P2) とした無名化など，匿名化やそれに類する加工の技術は同意を得て流通させる方式においても有効であろう．

7. まとめ

本稿では，個人の関与の下でパーソナルデータ流通・活用を進める仕組みに必要な機能／セキュリティ／プライバシー要件を提案し，それらを全て実現する初めての方式を提案した．特に，プライバシー要件である，自己情報コントロールと権限分散の両立が特長である．ここで，自己情報コントロールとは，個人が第三者提供したい自身のデータの項目を，その個人が制御可能なことを指す．また，権限分散とは，データを購入することなく，個人の全データにアクセス可能な者がいないことを指す．提案方式は，権限分散を実現している SCIS2018 方式を，自己情報コントロール可能なように拡張したものである．データ提供者それぞれが，個人から許可規則を受け，仲介者を経由してきた購入申請のデータ項目組み合わせがその許可規則に合っていた場合のみ購入申請に応じる．これにより，本人が許

可規則として設定したデータ項目組み合わせのみが販売されるようになる．今後の課題として，入力支援機能の開発がある．

参考文献

[1] Dwork, C.: Differential privacy, *in ICALP*, Springer, pp. 1–12 (2006).

[2] Kantara Initiative: User Managed Access, https://kantarainitiative.org/confluence/display/uma/Home. 2018 年 8 月 15 日参照.

[3] Sweeney, L.: Achieving k-anonymity privacy protection using generalization and suppression, *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, Vol. 10, pp. 571–588 (online), DOI: 10.1142/S021848850200165X (2002).

[4] World Economic Forum: Personal Data: The Emergence of a New Asset Class, https://www.weforum.org/reports/personal-data-emergence-new-asset-class. 2018 年 8 月 15 日参照.

[5] 株式会社 NTT データ経営研究所：パーソナルデータに関する一般消費者の意識調査，http://www.keieiken.co.jp/aboutus/newsrelease/161122/supplementing01.html. 2018 年 8 月 15 日参照.

[6] 株式会社インテージ：PDS/情報銀行の受容性と課題，https://www.kantei.go.jp/jp/singi/it2/senmon_bunka/data_ryutsuseibi/detakatsuyo_wg_dai8/siryou1.pdf. 2018 年 8 月 15 日参照.

[7] 前田若菜，山岡裕司：パーソナルデータ流通市場における本人特定を抑制する仮 ID 発行方式，2018 年暗号と情報セキュリティシンポジウム予稿集 (2018).

[8] 総務省：平成 29 年版情報通信白書, http://www.soumu.go.jp/johotsusintokei/whitepaper/h29.html. 2018 年 8 月 15 日参照.

[9] 内閣官房：AI、IoT 時代におけるデータ活用ワーキンググループ中間とりまとめ，https://www.kantei.go.jp/jp/singi/it2/senmon_bunka/data_ryutsuseibi/dai2/siryou2.pdf. 2018 年 8 月 15 日参照.