

パーソナル情報保護と活用のための調査報告

小松文子† 佐藤祥太郎‡ 宮澤泰弘‡ 美馬 正司‡

†(独)情報処理推進機構
113-6591 東京都文京区本駒込 2-28-8

‡(株)日立コンサルティング
102-0083 東京都千代田区麹町二丁目 4 番地 1 麹町大通りビル

あらまし ビッグデータ時代が到来する中、パーソナル情報を活用したビジネスが期待される一方でプライバシー懸念が高まっています。プライバシー懸念を軽減するために、国内外の活用事例、先端技術の動向、法制度、市場規模を調査した。その結果、わが国において健全なパーソナル情報の活用市場が成立するために必要な課題として、二次利用に必要なデータ加工などの条件を前提とした制度とプライバシー懸念を軽減させるための技術の客観的な評価のしくみの必要性、自己情報コントロールの要素と対応技術を明らかにした。市場規模を推計したところ、二次利用の領域は他の領域と比較し規模が小さいことが明らかであった。

A report about personal data protection and practical use

Ayako Komatsu† Shotaro Sato‡ Yasuhiro Miyazawa‡ Shoji Mima‡

†Information-Technology Promotion Agency
2-28-8 Honkomagome, Bunkyo-ku, Tokyo, 113-6591, JAPAN

‡Hitachi Consulting
2-4-1 Koujimachi, Chiyoda-ku, Tokyo 102-0083 JAPAN

Abstract While the big data age comes, and the business which utilized personal information is expected, privacy concern is increasing. In order to reduce privacy concern, a practical use example in and outside the country, the trend of high technology, a legal system, and market size were investigated. As a result, as a issues required since the practical use market of healthy personal information is materialized in our country, The system which enables secondary use on condition of conditions, such as data processing and the structure of objective evaluation of the technology for making privacy concern reduce are required. Moreover, the element and correspondence technology of self-information control were clarified. When market size was estimated, as compared with other domains, it was clear that the domain's of secondary use a scale is small..

1 はじめに

ビッグデータ時代の到来の中、サイバー空間で増大するパーソナル情報(*)を活用した新たな市場の可能性が期待される一方で、利用者の個人情報の悪用等に対する不安や、プライバシーが侵される事への懸念が高まっている。また、先端的なデータマイニング技術により分散した断片的なパーソナル情報から個人が特定されるなど、個人情報保護法における個人情報の定義が IT 技術に十分に対応できない場合もある。さらに、個人が自分の情報の取り扱いについてコントロールできる権利を保護するためには、どのような技術が必要であるかが不明瞭であるなど、法制度と IT 技術のアンマッチも見られる。このような背景から、IPA では、パーソナル情報活用に関する国内外のビジネスの状況調査、技術動向調査、法制の調査および市場を推計した調査を実施したので報告する。

2 ビジネスの動向

パーソナル情報活用の要件を明らかにするために、パーソナル情報を活用した創出例となんらかの阻害要因によりビジネスが成立していない事例を海外事例も含めて調査した。

2.1 ビジネス創出例

パーソナル情報を活用したサービスの事例として表 1 に記す事例を選定し、「ビジネスモデル」、「法制度」、「利用者のメリット」という観点で状況を調査した。その結果、パーソナル情報を活用している事例では、法制度への順守と、利用者への情報取り扱いについての説明が十分になされていることが認められた。例えば、インターネット視聴率調査(ネットレイティングス(株))では、本人へ十分な説明を行い、同意のもと情報収集が行われるだけでなく、個人情報保護方針が Web で示され、パネルから取得する情報、その利用目的、第三者提供を行う旨、利用停止方法等について、詳細な説明がされている。

2.2 ビジネス阻害例

海外では既に実際されているが我が国ではまだサービスインしていない事例や、サービスの構想はあるもののまだサービスインしていない事例を阻害事例として、表 2 に記す事例を調査した。各阻害事例について、「ビジネスモデル」、「法制度」、「利用者の受容性」という観点で状況を調査し、サービスを阻害する要因を抽出した。

その結果、海外では法制度が整備されており、データの取り扱いの基準が明確である事例が複数認められたため、この部分がサービス創生につながっている、すなわち日本ではこれが十分でない事が阻害要因の一つとなっていると考えられる。例えば、米国では、医療データを扱うにあたって HIPAA が制定されており、二次利用に向けて特定の 18 項目の情報を削除するなど、情報の加工方法の条件が規定されている。このように匿名化等の方法や仕組み等を制度的に明らかにすることで、安全性と流通性を両立したパーソナル情報の生成が可能になり、データマーケットプレイス等の発展も期待できる。

2.3 パーソナル情報悪用事例

パーソナル情報が悪用された事例、第三者の事業者に露出した事例、不特定多数の一般人に露出した事例、特定の一般人に露出した事例として、表 3 に記す事例を選定し、「個人情報保護法の遵法性」、「プライバシー侵害の可能性」という観点で状況を調査した。調査では、有識者の助言に基づいてプライバシー侵害の可能性に関する評価項目を設け、評価を行った。

その結果、パーソナル情報が露出・悪用されている事例が複数確認された。また、法制度には順守しているものの、利用者の受容性の観点からプライバシー懸念が解消されない点が問題となり、事業の停止に追い込まれる事例も認められた。

3 関連技術の動向

3.1 個人を特定する技術

パーソナル情報活用ビジネスの実現に関連す

る先端技術について、その研究および実用化の動向について調査を行った。

個人情報保護法では、第二条において、「個人情報」を「生存する個人に関する情報であつて、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの（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を含む）」と定義している。ITが発達したことで、多数のパーソナル情報が創出される反面、意図しないところで個人が識別できる可能性が高まってきている。すなわち、氏名や住所など、個人を特定するためのデータだけでなく、周辺の情報（コンテキストウェアネス情報）である情報が、インターネット、SNS、位置情報、画像情報などから抽出され個人を特定することができるようになってきている。ここでは、パーソナル情報活用ビジネスの実現に関連する技術として、個人を特定するため、または特定しないための技術に関して、研究・実用化動向を述べ、次に法制度への影響についてまとめた。

(1) 個人を識別するための技術

個人を識別するための技術(re-identification techniques)は、あるデータを他のデータと突き合わせて、ユニークなデータを抽出するものが多い。突き合わせるデータは、一般的に公開されているデータが多いが、観測情報や特定の個人が有している情報等が用いられる場合もある。従来、一般に公開されている匿名化されたデータから個人が識別される危険性があることを示すための研究が中心に行われてきており、これに基づいて、後述する個人を特定しないための技術が発展してきた。昨今では、SNS(Social Network Service)、スマートフォン(附随するカメラ)等の普及を背景として、ソーシャルメディア、位置情報、画像・映像等を活用して個人を特定する研究が進められている。個人を識別するための技術については、本人の認証を除けば、その活用はプライバシー侵害になる可能性が高く、治安維持のための犯罪者探索等、実用分野は限定的となる。

(2) 個人の特性を推定する技術

厳密に個人を特定しないが、その行動等から個人の特性を推定する技術の研究が進められている。これはユーザプロファイリング(技術)と呼ばれ、土方嘉徳(2004)によると、明示的手法(explicit method)と暗黙的手法(implicit method)に分けることができる。また、ユーザプロファイリング技術には、収集したデータに基づき作成したプロフィールを固定して活用するものと、プロフィールに基づいたレコメンデーション等を行い、その結果によってプロフィールの適合性を評価し、評価結果に基づきプロフィールを改良する学習型のものが存在する。このようなユーザプロファイリング技術は、Web閲覧履歴等をインプットとした研究が進められてきたが、昨今ではSNSやスマートフォン、購買履歴等も活用されるようになっており、実サービスで活用されているものも少なくない(表4)。

3.2 個人を特定しない技術

個人を特定しないための技術はその手法から、匿名化手法、攪乱手法、暗号化手法に分けることができる。それぞれに概要、研究動向、実用化動向を整理する。多様な研究が進んでいるものの、実用化している事例は少ない状況であり、ビジネスとして展開されているのは匿名化に関する一部の事例のみである(表5)。

4 自己情報コントロール権

4.1 自己情報コントロールに関する法曹界の議論と構成要素

自己情報コントロールという用語は、昭和45年佐藤幸治が「自己情報コントロール権」=プライバシー権利と主張し、「個人が道徳的自律の存在として、自ら善であると判断する目的を追求して、他者とコミュニケーションし、自己の存在にかかわる情報を開示する範囲を選択できる権利」と定義している。ただし、佐藤は自己情報コントロール権の要素として、閲覧請求権、訂正・削除要求権、利用・伝播統制権を示し、かつ収集、保有の制限も必要であると述べている[3]。そのほか、芦部信喜、鈴木正朝、日本弁護士連合会、青柳武彦の議論

を参考とし、自己情報コントロールが、統一的な定義を持たない概念であることを明らかにした。

また、佐藤の説である「自己情報コントロール権」は、現在の日本で有力な説となっているものの、その考えに異議を唱える専門家も少なくない。ただし、青柳のように、自己情報コントロール権の考えに批判的な人物の主張を見ても、自己情報コントロールの必要性(例えば自己情報へのアクセス、訂正、削除等)を根本的に否定しているわけではなく、その内容の強弱や対象とする情報の範囲、実現性、当該内容を権利として認めるかどうか、についての考えの違いを主張しているものと考えられる。なお、自己情報コントロール権は憲法13条(生命・自由・幸福追求権)を根拠として認められると主張されている。

さらに、自己情報コントロールについて、国内の学識有識者等見解やOECD8原則、判例、EUデータ保護指令などを比較検討し、自己情報コントロールの構成要素を、個人の権利としての「アクセス(確認)」、「訂正」、「削除」、「利用停止」、「データポータビリティ」と、事業者の義務としての「収集制限」、「保有条件(データの正確性の維持)」、「利用制限」、「事業者情報の公開」に分類した。

4.2 自己情報コントロールのまとめ

日本の個人情報保護法は、個人の権利を規定していないものの、事業者の義務によって実質的に権利が規定されているものとみなすという意見もあった。一方、EU及び米国などにおける法制度の動向をみると、事業者だけでなく個人の権利を強化する方向であることがうかがえる。これはパーソナル情報を活用したサービスの発展と連動した流れと考えることができる。このような制度環境の変化は企業に対し自己情報コントロール機能の実装を求めることになると想定され、実際に米国では、まだ法案段階であるにもかかわらず、Do Not Trackに対応した機能を各Webブラウザで既に実装している。一方で、自己情報コントロール機能は利用者によるサービスの利用やパーソナル情報の提供に対して一定の効果があると考えられる。したがって、利用者の側からも自己情報

コントロールが求められていることがうかがえる。このことから、今後の制度環境の変化への対応や、利用者におけるサービスの受容性の向上等の観点から、自己情報コントロール機能の実装が有用と考えられる。

ただし、法制度等の調査結果からも明らかなように、自己情報コントロールの構成要素やその機能については現状で明確な定義は存在しないことから、企業が実際に自己情報コントロール機能を整備する意思を持った場合にも、対応すべき内容は不明確である。本調査で整理した自己情報コントロールの構成要素や必要とされる自己情報コントロール機能について、更に具体的な検討を行った上で、共通概念を形成していく必要があると考えられる。また、活用されるパーソナル情報の種類によって求められる自己情報コントロールの範囲や機能、求められる認証の厳格さ等も異なるため、ある程度の共通概念が形成された後には、業界毎に実装すべき自己情報コントロール機能について具体的に検討することが望ましいと考えられる。

5 市場の推計

パーソナル情報を活用した市場全体を図3に記すようにモデル化し、これに基づいて既存の文献等を活用して市場規模の推計を行った。パターンIについては、各産業におけるパーソナル情報に関するIT投資額を市場規模として捉え、パターンII～IVについては、該当するサービスに関わる文献や統計情報を活用し、推計を行った。また、網羅性を担保するため、日本標準産業分類からパーソナル情報を活用した産業のグルーピングを行い、それぞれについて市場規模を算出している。算出の結果、2010年度のパーソナル情報市場は約1兆7,048億円と推計される(表6)。パターンIVは、パーソナル情報の二次利用の対象となる領域であるが、全体の1割強の市場であり、他領域と比べ市場が未成熟であると推察できる。

6 健全なパーソナル情報活用市場 の発展のために

これまで述べたパーソナル情報市場、技術、制度の関係を総括すると図 4 に記すように整理できる。健全なパーソナル市場の発展のためには、関連した個人を特定するための技術や特定しないための技術の活用促進、プライバシー懸念に配慮した企業等の自主的な取り組みの推進やこれを支援する制度・政策面での整備、また、技術の活用に対する推進や不適切な利用防止のための制度面からの整備等が不可欠である。

パーソナル情報市場については、既に約 1 兆 7,048 億円の市場が存在するものの、海外の事例等を参考とすると更なる発展が期待できる。そのためには、既存のデータに加え、携帯電話やソーシャルメディア等によって集積が進むパーソナル情報の安全かつ高度な活用、流通を可能とする環境整備が不可欠である。

技術の観点からは、個人を特定するため、又は特定しないための技術が進展してきており、この適切な活用促進と不適切な利用の防止が非常に重要である。現時点において、このような技術の進展に対して、制度面の環境整備は十分とは言えない。例えば、個人を特定するための技術の進展は、個人情報保護法の個人情報に該当するデータの範囲を広めている。また、米国の HIPAA のように、個人を特定しないための技術によって処理した情報を活用できることを明示した制度は規定されていない。今後、個人の識別性を排除したパーソナル情報を活用するため、あるいは悪意のある個人の識別を防止するため等に向けた制度設計が求められる。

また、個人がパーソナル情報を企業へ提供する際に、自己の情報の流通をコントロールする機能(自己情報コントロール機能)が有効であるとの調査結果があるが、自己情報コントロールについては、法制度上も、技術実装上も、その言葉の意味する統一した概念は規定されていない。本報告書では、試みに個人の権利としての自己情報コン

トロールの構成要素をまとめた。そして、実例の調査を通して自己情報コントロール機能は、技術および運用による実装が進められているが、その機能をすべて充足していないことも明らかになった。個人、企業間の共通的な概念形成が必要であり、ビジネスへの適用には企業、業界団体等の自主的な取り組みが求められる。また、この点で制度面での関与が役立つと考えられる。

一方、パーソナル情報のビジネス活用全般において個人情報保護法だけでなくプライバシーへの配慮がより求められるようになってきている。本報告で述べた露出、悪用事例では、制度に順守しても、消費者に受容されないため、ビジネスが成立しないケースが目立った。これは、企業によるビジネス展開が、制度とともに、個人の受容性にも立脚していることを意味している。特に法律等の制度面の整備においては、このビジネス、制度、個人の受容性のバランスを前提として進める必要がある。また制度整備に相応の時間を要すると考えられ、急速に変化するパーソナル情報を活用したビジネスの中では、「利用者の視点に立った対応が最終的にビジネスの発展につながる」という観点から各企業や業界団体で対応を図ることが求められる。

以上のことから、進展する技術を活用したパーソナル情報市場の発展のためには、制度面での拡充が求められるものの、制度整備に依存するだけではなく、ビジネス側での自主的な取り組みも踏まえたバランスのあるパーソナル情報市場のエコシステムの形成が望まれる。

参考文献

- [1] (独) 情報処理推進機構、パーソナル情報と IT 技術の調査報告書、2012.8
<http://www.ipa.go.jp/about/press/20120823.html>
- [2] 土方嘉徳、情報推薦・情報フィルタリングのためのユーザプロファイリング技術、人工知能学会誌、Vol. 19, No. 3, pp. 365—372, 2004.
- [3] 衆議院憲法調査会、「知る権利・アクセス権とプライバシー権に関する基礎的資料」、2003.5
http://www.shugiin.go.jp/itdb_kenpou.nsf

/html/kenpou/shukenshi028.pdf/\$File/shuk
enshi028.pdf

表 1 パーソナル情報 活用事例

No.	サービスの分類	活用事例	サービス事業者
1	レコメンド／コンテンツ配信	i コンシェル	(株)エヌ・ティ・ティ・ドコモ
2	Web マーケティング	インターネット視聴率調査	ネットレイティングス(株)
3	オンライン広告	Google Adwords	Google
4	販売促進／マーケティング支援	おすすめ商品	Amazon.com
5	販売促進／マーケティング支援	T ポイントサービス	カルチュア・コンビニエンス・クラブ(株)
6	SNS 広告	スポンサー広告	Facebook
7	O2O	Edy au	KDDI(株)、楽天(株)
8	PHR	ポケットカルテ	(特非)日本サスティナブル・コミュニティ・センター

表 2 パーソナル情報阻害事例

No.	サービスの分類	阻害事例	サービス事業者／提案者
1	Personal Data Service	データエクスチェンジ事業	STATZ
2	医療データ活用	医療データ販売事業	MedMining
3	医療データ活用	Heritage Health Prize	Heritage Provider Network
4	Personal Data Service	情報銀行モデル	東京大学 空間情報科学研究センター 柴崎研究室
5	データマーケットプレイス	Windows Azure Marketplace	Microsoft

表 3 パーソナル情報悪用・露出事例

No.	露出、悪用事例の分類	事例	サービス事業者
1	悪用	CarrierIQ 問題	CarrierIQ、携帯電話事業者
2	悪用	iPhone の位置情報収集	Apple
3	不特定多数の一般人に露出	Google Buzz	Google
4	第三者の事業者に露出	AppLog	(株)ミログ
5	特定の一般人に露出	カレログ	(有)マニユスクリプト
6	不特定多数の一般人に露出	Google Location Server	Google
7	不特定多数の一般人に露出	People Search	Intelius

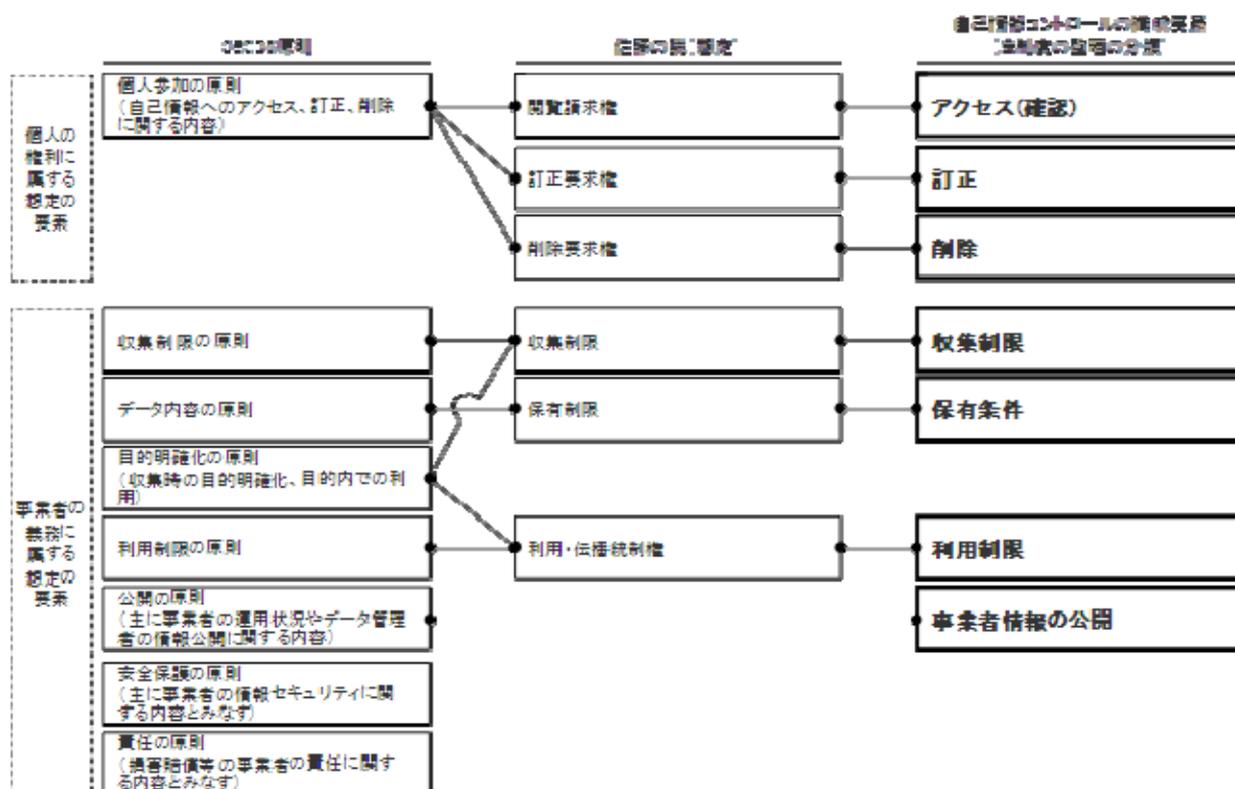
表 4 プロファイル特定手法の分類

手法	概要	インプット
明示的手法 (explicit method)	ユーザーが明示した内容からプロファイルを推定	<ul style="list-style-type: none"> アンケート調査 ブログ、SNSの書き込み 見た目
暗黙的手法 (implicit method)	ユーザーの行動からプロファイルを推定	<ul style="list-style-type: none"> Webの閲覧履歴 SNSのネットワーク構造 購買履歴 位置情報の履歴

表 5 個人を識別しない技術

手法	概要	研究/実用化動向
匿名化手法	<ul style="list-style-type: none"> あるデータに対して、情報の削除やあいまい化等の加工を行うことで、データを参照した人が個人を識別できない、又はセンシティブな情報が分からない状態にする手法 コードを匿名化する際に用いられる指標として k-匿名性、l-多様性、t-近似性等 	<ul style="list-style-type: none"> 近年の研究傾向として、新指標を提起するものと実データへの適用を検討するものが存在 新指標を提起するものは、データの偏在や攻撃者の背景知識等による攻撃を想定し、安全性に関するより強固な指標を提示 実データ適用では、逐次追加・更新されるデータや多次元データ等の実データに近いデータに匿名化手法を適用 実用化している事例は少なく、カナダの PARAT、オランダの μ-ARGUS 等でビジネス展開
攪乱手法	<ul style="list-style-type: none"> あるデータに対して、ノイズ付加や属性値のシャッフル等の加工を行うことで、データを参照した人が個人を識別できない、又はセンシティブな情報が分からない状態にする手法 レコードを攪乱する指標として Pk-匿名性、クエリ応答を攪乱する指標として差分プライバシー等 	<ul style="list-style-type: none"> 近年、ランダム化については安全指標を提起する研究、クエリ推論制御については差分プライバシーへの対応に関する研究が存在 ランダム化には、これまで安全性に関する指標が存在していなかったため指標を提起する研究が展開 差分プライバシーについては、位置情報、二つのパーティーからのクエリ等を想定した研究が展開 実用化事例は少なく、Harvard Medical School 等で実証的に導入
暗号化手法	<ul style="list-style-type: none"> あるデータに対して、暗号化処理を施して、データを参照した人がデータベースの中身を見れない状態のまま、計算結果のみを得る手法 複数のパーティーに分散している情報を他パーティーに秘匿したまま処理を行う手法 (MPC) や、準同型暗号を利用して暗号化したデータベース上で計算を行う手法等 	<ul style="list-style-type: none"> 近年、Craig Gentry (2009) の完全準同型暗号を起点とした研究が存在 MPC は、実用化に向けて、ネックとなっている計算処理の効率化に取り組む研究や、実データへの適用を試みる研究が存在 実用化事例は少なく、産総研の「秘密計算による化合物データベースの検索技術」、MIT の CryptDB 等が実用的な研究としてあり

図1 自己情報コントロールの構成要素



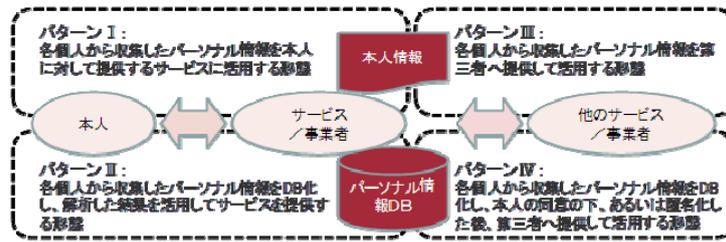


図 2 パーソナル情報のビジネスパターン

表 6 パーソナル情報の市場推計のまとめ

グループ	パターンⅠ	パターンⅡ	パターンⅢ	パターンⅣ	単位 億円	
					合計	重複部分
インフラ関連業界	262	43	—	—	305	
通信業	296	320	3,060	—	3,676	
情報処理・提供サービス業	—	—	—	1,666	1,666	
インターネット・近隣サービス業	34	320	271	2	627	パターンⅠ
鉄道業、航空運輸業	296	320	—	—	616	パターンⅠ
道路貨物運送業、郵便業	62	—	—	—	62	
小売業	267	4,230	—	—	4,497	パターンⅠの一部
金融・保険業	1,293	4,230	213	—	6,036	パターンⅠ
宿泊業、旅行業	198	4,230	—	—	4,428	パターンⅠ
教育、学習支援業	264	—	—	—	264	
医療	1,140	—	26	363	1,529	
保健	304	—	—	—	304	
福祉	171	—	28	—	199	パターンⅢ
地産地消	—	—	1,000	—	1,000	
総計	1,470	—	—	—	1,470	
合計	6,296	4,273	4,669	1,970	17,048	

図 3 パーソナル情報と技術・制度

