

平成 28 年度総務省委託事業

位置情報に関するプライバシーの適切な保護と  
社会的活用の両立に向けた調査研究 報告書

平成 29 年 3 月

株式会社野村総合研究所



## はじめに（本調査研究の趣旨）

昨今、スマートフォン等の高機能な移動体端末は、その急激な普及により、我が国の国民及びその経済活動にとって、欠かすことのできない重要なインフラとなっている。そうした移動体端末からは、基地局やGPS、Wi-Fiといった手段により電気通信事業者においてその位置情報を取得することが可能になっている。取得された位置情報はビッグデータとして、防災・減災や街づくり、観光地・商店街の活性化、利用者に向けた有用なサービスの展開等様々な社会的効果が期待されている。しかしながらその一方で、個人にかかわる情報を含むビッグデータの利活用については、プライバシー面での懸念が指摘されている。

このような状況を踏まえ、平成25年11月から平成26年5月にかけて、「緊急時等における位置情報の収集等に関する検討会」（座長：長谷部恭男 早稲田大学大学院法務研究科教授）を開催し、パーソナルデータとして利活用が期待されている電気通信事業者が取得する位置情報について、通信の秘密や個人情報、プライバシーを適切に保護しつつ、ビジネス利用も含めたその社会的利活用を促進するための所要の整理を行い、平成26年7月17日に、「位置情報プライバシーレポート」をとりまとめた。

平成27年度の調査研究では、「位置情報プライバシーレポート」を踏まえて、ユースケースに基づく実証研究を行い、位置情報の適切な取扱いに関するより具体的な方法を示した。本調査研究（平成28年度）は、平成27年度の調査研究の結果を踏まえ、電気通信事業者が取り扱う位置情報について、利活用に係るユースケースを設定した上で、「十分な匿名化」に関する検証を継続的に実施するほか、平成27年9月に成立した改正個人情報保護法（平成27年法律第65号）で新たに規定された匿名加工情報について検証等を実施した。

本調査研究の検証結果は、電気通信事業者が扱う位置情報に関するルールに盛り込まれるべき内容（案）として、「十分な匿名化」の取扱い、匿名加工情報の作成方法等、それぞれについてとりまとめた。これらのルール（案）は、電気通信事業分野における業界団体や認定個人情報保護団体に引き継がれ、位置情報を取り扱う際の規範として活用されていくことを想定している。



## 目 次

はじめに（本調査研究の趣旨）

第1 協議会の設置・運営等 .....	1
1. 委員構成.....	1
2. 協議会の開催概要 .....	2
第2 調査研究.....	3
1. ユースケースを通じた「十分な匿名化」及び匿名加工情報の作成方法に関する検証 .....	3
1) ユースケースの設定(昨年度の検討経緯と今年度の射程) .....	3
2) 商用ケースに係る検証 .....	5
3) 観光ケースに係る検証 .....	27
4) 検証結果の整理 .....	41
2. 加工方法等情報に係る安全管理措置の基準の検証 .....	42
1) 安全管理措置に関するルールの整理 .....	42
2) 匿名加工情報の適正な取扱いを確保するために必要な措置の検証 .....	43
3) 今後の展開に向けて.....	45
3. 「十分な匿名化」と匿名加工情報の作成の方法等についての関係等の整理 .....	46
1) 論点の設定 .....	46
2) 論点の検証結果 .....	47
4. 利用者の受容性調査.....	53
1) 調査概要.....	53
2) 調査結果.....	56
5. 電気通信事業者が取り扱う位置情報に関するルールの整理.....	93
1) 「十分な匿名化」に関するルールに盛り込まれるべき内容 .....	94
2) 電気通信事業者が扱う位置情報を用いた匿名加工情報の作成等に関するルールに盛り込まれるべき内容 .....	132
6. 今後の展開に向けて .....	158
1) 業界団体ルール、個人情報保護指針への取込み.....	158
2) 「十分な匿名化」に係る安全管理措置のフォローアップ .....	158
3) 匿名加工情報に係る他分野の取り組みとの整合 .....	158
4) IoT の進展を踏まえたルールづくり.....	159

付属資料 協議会議事概要



## 第1 協議会の設置・運営等

本調査研究では、個人情報保護・通信の秘密を専門分野とする有識者及び情報セキュリティを専門分野とする有識者ら複数名を委員に、個人情報保護委員会事務局、電気通信事業者、関連団体をオブザーバとする「位置情報に関するプライバシーの適切な保護と社会的活用」の両立に向けた調査研究「協議会」（以下「協議会」という。）を設置して検討を行った。

### 1. 委員構成

協議会の委員、オブザーバ、事務局の構成は次の通りである。

#### <委員>

- ◎森 亮二 英知法律事務所 弁護士
- 伊藤 伸介 中央大学経済学部 准教授
- 佐藤 一郎 国立情報学研究所 教授
- 高橋 克巳 NTT セキュアプラットフォーム研究所 主席研究員
- 疋田 敏朗 トヨタ IT 開発センター シニアリサーチャー

◎は主査

#### <オブザーバ>

- 個人情報保護委員会事務局
- 一般社団法人電気通信事業者協会
- 一般財団法人日本データ通信協会
- (株) NTT ドコモ
- KDDI (株)
- ソフトバンク (株)
- エヌ・ティ・ティ・ブロードバンドプラットフォーム (株)

#### <事務局>

- 総務省総合通信基盤局電気通信事業部消費者行政第二課
- (株) 野村総合研究所

## 2. 協議会の開催概要

協議会は全4回実施し、各回の開催日と主な議題は次の通りである。

回	開催日と主な議題
第1回	開催日：平成28年11月8日（火曜） 1. 協議会の趣旨等について 2. 検討項目と主な論点 3. ユースケースの設定と利用者の受容性の評価
第2回	開催日：平成28年12月13日（火曜） 1. 電気通信事業者における匿名化した位置情報の活用状況と活用ニーズ 2. ユースケースの設定 3. 利用者の受容性調査
第3回	開催日：平成29年2月7日（火曜） 1. 「十分な匿名化」と匿名加工情報の作成方法に関する基準の検証 2. 安全管理措置に関するルールの整理 3. 利用者の受容性調査 4. 「十分な匿名化」と匿名加工情報の作成の方法等についての関係等の整理
第4回	開催日：平成29年3月14日（火曜） 1. 安全管理措置に関するルールの整理 2. 「十分な匿名化」と匿名加工情報の作成の方法等についての関係等の整理 3. 電気通信事業者が取り扱う位置情報に関するルールの整理 4. 利用者の受容性調査の結果



## 第2 調査研究

### 1. ユースケースを通じた「十分な匿名化」及び匿名加工情報の作成方法に関する検証

#### 1) ユースケースの設定（昨年度の検討経緯と今年度の射程）

昨年度の調査研究では、「位置情報プライバシーレポート」を受け、電気通信事業者が保有する「通信の秘密」に該当する位置情報について、「十分な匿名化」の枠組みを用いて利活用することに関する実証実験がなされた。

具体的には「防災」、「交通」、「観光」、「商用」分野で、利活用が想定されるユースケースを設定し、それぞれのユースケースにおいて適切と考えられる匿名化の加工を行い、「十分な匿名化」に係る要件が整理された。

図表 1 昨年度検討されたユースケース

名称	概要
震災等災害時の避難行動分析(防災)	GPS測位した位置情報を用いて平常時の「流動人口」の集計を行い、そこから災害が発生した場合の被害を算出した後、内閣府公表の被害想定との比較を行う。
都市交通整備への活用を想定した空港利用者の動態分析(交通)	利用者の契約情報及び位置情報を用いて、首都高速中央環状線(大井JCT～初台IC)の開通前後の羽田・成田空港利用者の動態分析を行う。
スポーツ観戦時におけるユーザーの流動性調査(商用)	九州地区での野球観戦キャンペーンへの応募者から、顧客情報と一定期間における位置情報を取得し、イベント施設周辺エリア及び商業エリア(博多駅等)に滞在した人の①属性、②施設来場頻度・滞在時間、③施設来場前後の立寄り等の分析を行った。
観光客の滞留時間分析、導線分析(観光)	交通要所及び観光スポットに設置しているWi-Fiエリアを活用して位置情報を取得し、訪日外国人等の各スポットでの接続時間を把握・分析する。

出所) 改正個人情報保護法等を踏まえたプライバシー保護検討タスクフォース「議論の取りまとめの方向性」(2016年7月12日)

加工要件に関しては、以下に挙げた措置を行うことが整理されていた。

#### ① 加工を始める際の入口要件

- ・ 直接識別子は削除する
- ・ 仮IDの生成に当たっては、可能な限り短い有効期限を設ける
- ・ ハッシュ化等の一方向の置き換えを行う

## ② 加工を終える際の出口要件

- ・ 原則として、同一時刻のすべての間接識別子について非識別、かつ、すべての時間を通じた履歴において非識別
- ・ 特定の個人が識別されるリスクを定性的に評価するに当たり、評価指標に以下を含める：

1) データの項目	5) 位置データの精度、
2) 場所の特性	6) 追跡可能な移動履歴の長さ・仮 ID の有効期間
3) 利用者の特性	7) データを取得する際の時間間隔
4) 取得期間・時期	8) 標本数

本年度の調査研究においては、業界のルールとしてとりまとめることを想定して、昨年度の「十分な匿名化」に関する検討を踏まえつつ、拡充・詳細化すべき視点を検討するため、より実用に近しいと考えられるケースを設定し、検証を行った。

例えば、以下に挙げる視点は、本調査事業で掘り下げて検証した。

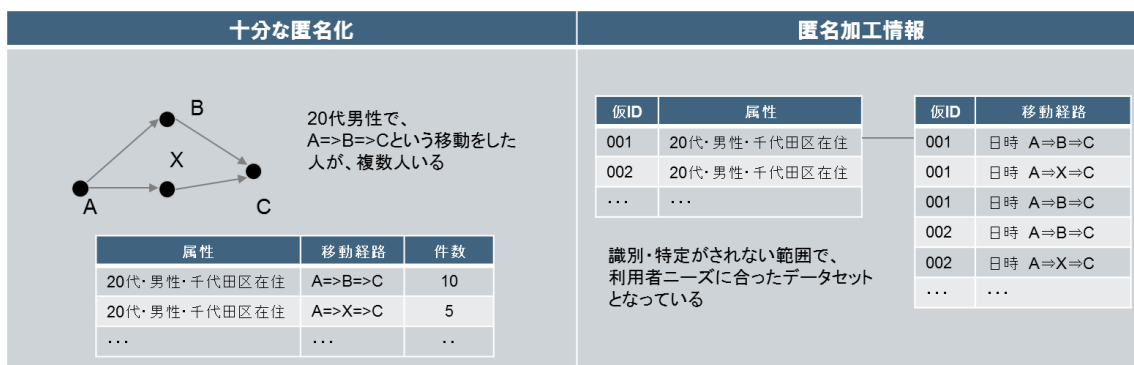
- 消費者の属性情報（以下「位置情報」との対比のため、「付帯情報」と言う。）を追加した場合、「十分な匿名化」の加工基準にどのような影響が生じるか（ただし、「十分な匿名化」では、用いることのできる付帯情報は限定されているため、その範囲での検証とする）
- 位置情報に関して、移動履歴として滞留地点を複数つなげた場合に、「十分な匿名化」の加工基準にどのような影響が生じるか
- 評価指標に、さらに「提供までの期間」を加えて評価してはどうか<sup>1</sup>

なお、「十分な匿名化」における加工方法の意図するところは、昨年度と同様であり、位置情報の履歴、属性を含むデータセットに対して、非識別性を担保し、利用されることを想定して集計したものである（図表 2）。

また、改正個人情報保護法において新たに制度化された匿名加工情報に係る検討も、「十分な匿名化」に係る検討とあわせて実施した。匿名加工情報の加工方法は、個人情報保護委員会によってガイドラインが示されている。本調査研究では、同ガイドラインを踏まえて「位置情報の履歴、属性を含むデータセットを、特定の個人が識別することができないように加工し、かつ元の個人情報を復元することができないようにしたもの」とた（図表 2）。

<sup>1</sup> 「改正個人情報保護法等を踏まえたプライバシー保護検討タスクフォース」における指摘を受けた対応

図表 2 「十分な匿名化」と匿名加工情報のデータイメージ



本年度は、昨年度の経緯を踏まえ、事業者の活用ニーズが高く、かつ2020年を見据えた社会的なニーズがあると考えられる「商用」と「観光」の2つのユースケースを設定した。

なお、データセットの準備等にあって、本調査研究では、いずれのユースケースにおいてもダミーデータを作り、それを分析することとした。各種条件は、現実に即したものになるよう配慮を行っているが、以降で示す「結果」自体は、ダミーデータの分析によって得られたもので、電気通信事業者が実際の運用を通じて取得したデータに基づくものではない点につき、予め留意されたい。

## 2) 商用ケースに係る検証

### (1) ユースケースの設定

昨年度は、商用のユースケースとして、九州地区での野球観戦におけるユーザーの移動履歴を分析した。このケースではユーザーの行動パターンは限定され、かつ位置情報の精度が粗い条件下での分析であった。

本年度は、昨年度実施していなかったポイントとして、

- 複数のエリア（メッシュ）間を、どのように移動するのかを、プライバシーに配慮しながら分析する
- 位置情報メッシュを細かく見る、または、人の趣味嗜好情報と複数のホップ数を重ね合わせる、発地を加味する等の検討をする

を実施し、商業地区・施設におけるマーケティング活動等への適用を目的とし、商業地区への訪問者の特徴・傾向を捉えた情報の作成に関する検討を行った。

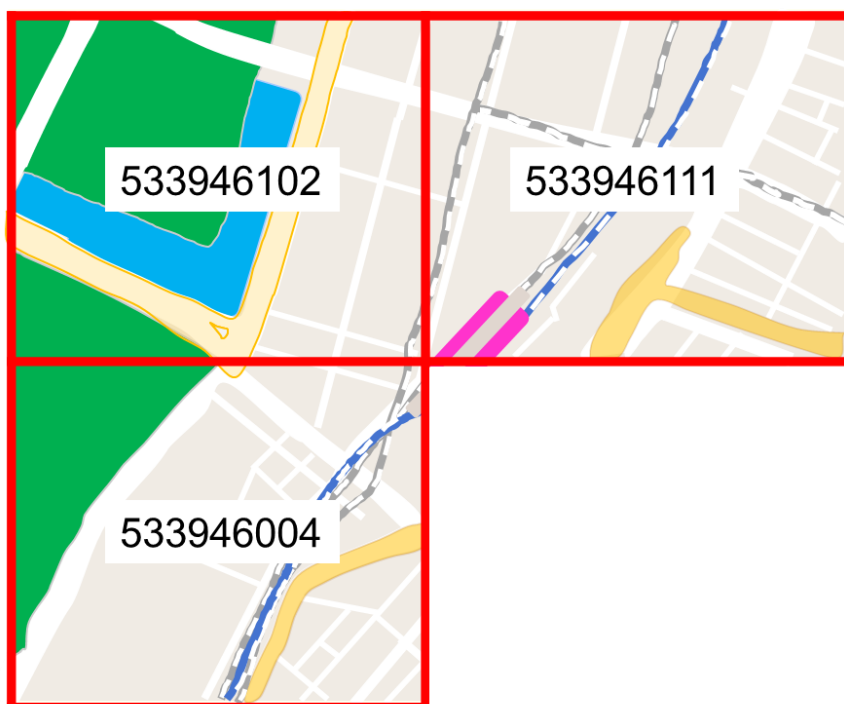
具体的には、通信端末情報等を通して得られた利用者の位置情報を用いることを想定し、有楽町駅西側・日比谷駅周辺において、以下の分析を実施する。

- 付帯情報（本調査事業では趣味嗜好情報を使用）を踏まえた来場者の分析

- 一定期間における来場者の来場頻度の分析
- 来場者の対象エリア内の移動履歴の分析

以下に、ユースケースで利用した実際のエリア情報を示す（図表 3）。なお、地図の赤枠の囲み（メッシュ）については、7 後段で説明する。本エリア内には、オフィス、皇居、日比谷公園、ホテル・イベント施設、映画館・劇場・美術館、ショッピングセンター、飲食街、有楽町駅・日比谷駅、小学校といった施設が含まれている。

図表 3 分析対象としたエリア



## (2) 準備したデータセットの説明

本ユースケースでは、大きく 3 つのデータベース等を利用することで、以下のデータを収集し、活用することを想定した。

- 契約者情報：ユーザーID、性別、年齢、住所コード
- 位置情報：ユーザーID、緯度、経度、日時
- 趣味嗜好情報

対象とした期間は、2016 年 10 月 31 日～11 月 13 日の 2 週間である。この 2 週間を想定した理由は、平日と休日、さらに祝日が含まれる期間であったこと、劇場のイベントがこの期間内に含まれると想定されること等が挙げられる。

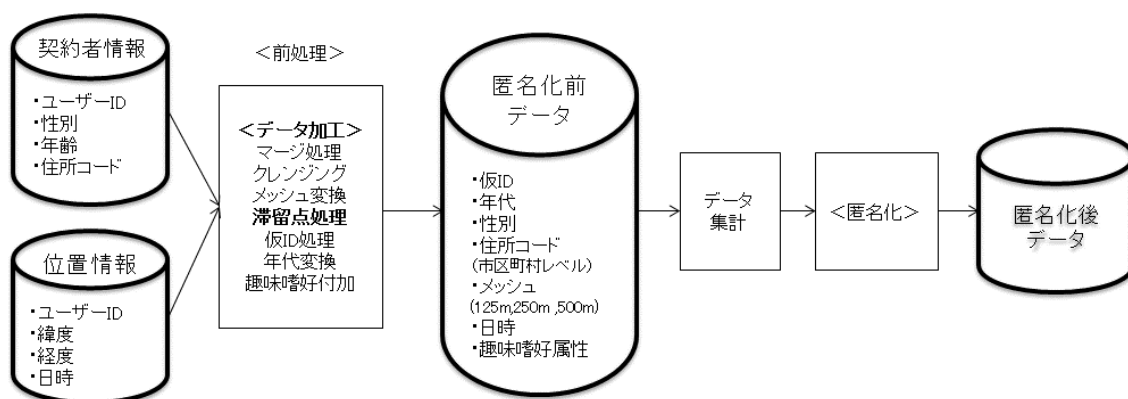
## (3) 「十分な匿名化」に係る検証

### ① 「十分な匿名化」のデータセットの作成

「十分な匿名化」のデータセットの作成においては、入口要件としての前処理と加工を終える際の出口要件の2つの加工が必要となる。

検証の前に、「十分な匿名化」のプロセスの図表に則り、「十分な匿名化」のデータセットを作成するまでの流れを以下で説明する。

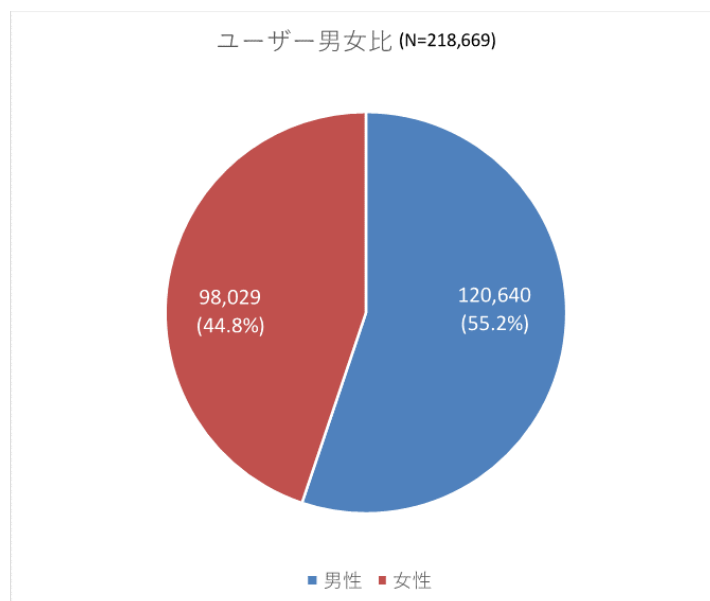
図表 4 「十分な匿名化」のプロセス



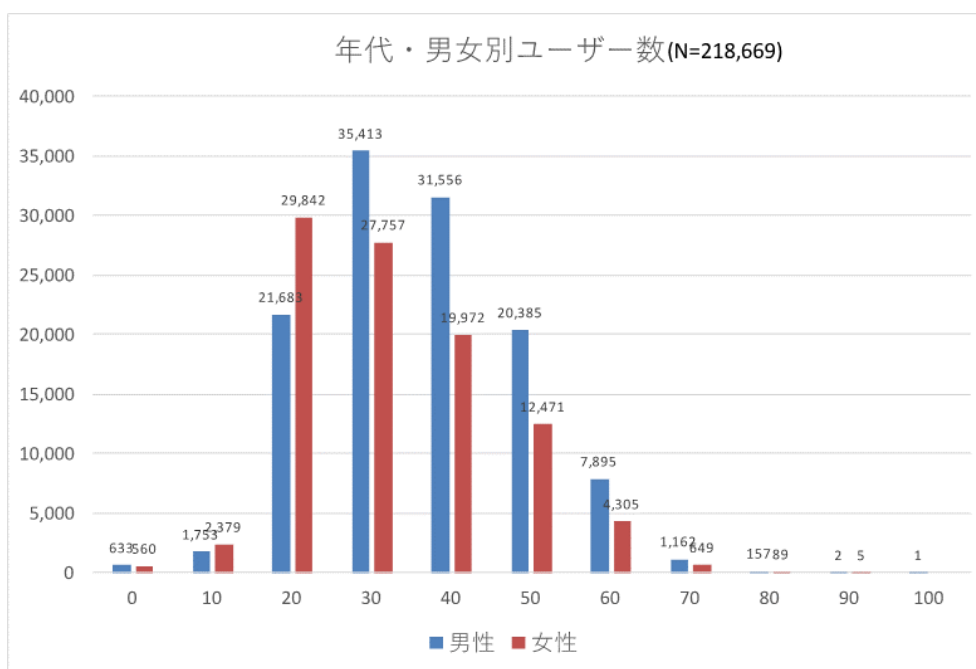
#### 【元となるデータの概要】

- ユーザーID数・レコード数：  
本調査研究で用意したユーザー数は 218,669 人、対象期間内（2 週間）の総レコード数は 2,319,739 件である。
- 性別：  
性別は男女の 2 区分である。
- 年齢：  
本調査研究では、年齢情報は、予め 10 歳区分のものを用意した。
- 住所コード：  
本調査研究では、住所コードは、予め市区町村まで丸めたものを用意した。
- 緯度・経度：  
緯度経度は通信の秘密に係る Wi-Fi 情報から収集した。
- 日時：  
日時は年月日時分秒単位で設定をした。
- 趣味嗜好：  
趣味嗜好情報は、スポーツ観戦、映画鑑賞、観劇、グルメ、ショッピング、読書、ゲーム・アニメ、音楽、健康、マジックの 10 種類を用意した。  
元となるユーザーのデモグラフィック特性に関するデータを以下に示す。

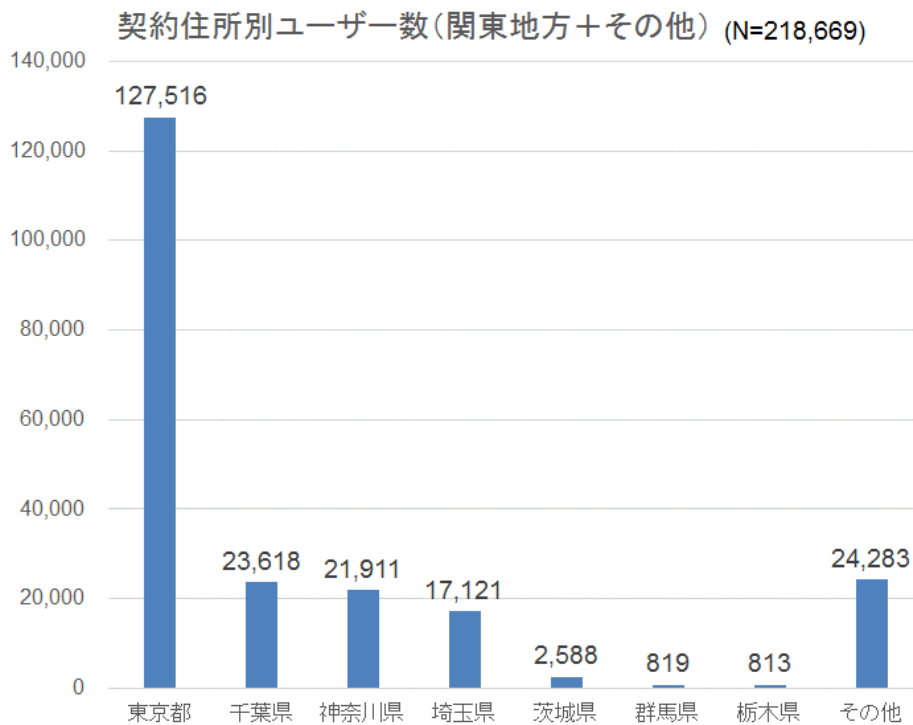
図表 5 ユーザーのデモグラフィック特性① 男女比



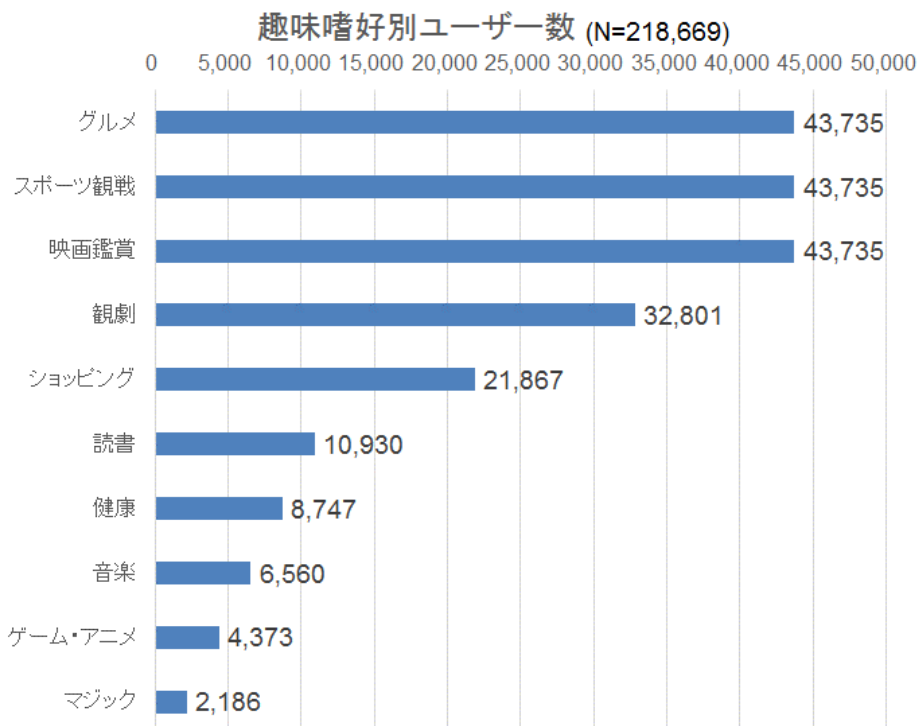
図表 6 ユーザーのデモグラフィック特性② 男女×年齢構成



図表 7 ユーザーのデモグラフィック特性③ ユーザーの居住地



図表 8 ユーザーのデモグラフィック特性④ ユーザーの趣味嗜好情報



【前処理】

前処理では、上述のデータを、ユーザーID を用いて紐付けし、データクレンジングや時間の変換、位置情報の変換、ユーザーID の仮 ID 化等を実施した。以下に、前処理における特徴的な事項を記載する。

- データクレンジング：

Wi-Fi の起動状況等に鑑み、ユーザーの位置情報の履歴が収集できないケースが想定される。このような「不完全」なデータを削除している。

- 日時：

日時情報のうち、深夜は、再識別の可能性を高めることに鑑み、データ収集時間を 6 時～24 時に限定した。また、時間に幅を持たせるために、時分秒の単位を 15 分単位に丸めた。

このような処理の結果、ユーザー数は、218,699 人から 88,851 人まで減少する。また、レコード数も、2,319,739 件から 764,785 件に減少した。

- 位置情報：

本ユースケースの活用の目的は、商圈分析にあることに鑑み、ユーザーが特定の地点に立ち寄っていることを分析したいニーズがあると考えた。そのため、位置の認識を、エリアをメッシュ化し、「滞留点」によって把握することとした。

○エリアのメッシュ化とは：

商用ケース冒頭で示したエリアを、幾つかの区分に分けて把握するものであり、本調査研究では、

- ・ 500m 四方のブロック 3 つ
- ・ 250m 四方のブロック 12 個
- ・ 125m 四方のブロック 48 個

によって、エリアを認識する。すなわち、同一メッシュの中に入れば、同じ場所にいるものと認識する。

○滞留点とは：

同一メッシュ又はエリア内、または回遊対象の建物・施設周辺において、最初に位置情報が測位された時刻から

- ・ 一定時間、連続的に位置情報が存在する場合、及び
- ・ 断続的に存在する場合は、一定時間以上滞在したと想定される場合、に滞留したとみなすことを意図し、単純な通過は、滞留にカウントしない。



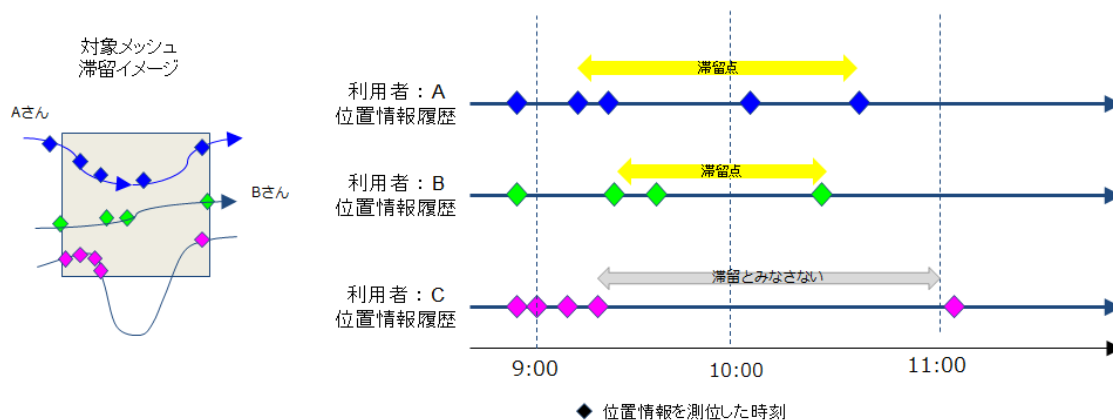
図表 9 に滞留点のイメージを示す。

本調査研究では、滞留の定義を、最低限 125m 四方のメッシュエリアに 15 分以上滞留していることとした。その結果、利用可能なデータは、レコード数で見て 68,564 件にまで減少した。

この滞留点の考え方をを用いて、以降で移動情報の分析を実施する。例えば、複数ホップを移動する際には、利用者の移動履歴情報から、予め定義した滞留点を抽出し、複数の滞留点を回遊した履歴としている。

図表 9 滞留点のイメージ

〈一定時間=1時間の場合、滞留メッシュの例〉



なお、上述の処理によって得られた前処理後のデータサンプルは以下の通りである。

図表 10 前処理後のデータサンプル

仮ID	年代	性別	住所コード	日時	区分	緯度	経度	1/8メッシュコード (125m)	1/4メッシュコード (250m)	1/2メッシュコード (500m)	趣味嗜好
5995	40	1	12227	20161102 123000	平日	35.673	139.762	53394600442	5339460044	533946004	映画鑑賞
5995	40	1	12227	20161102 124500	平日	35.672	139.761	53394600424	5339460042	533946004	映画鑑賞
5995	40	1	12227	20161102 153000	平日	35.672	139.761	53394600424	5339460042	533946004	映画鑑賞
5995	40	1	12227	20161102 160000	平日	35.673	139.762	53394600442	5339460044	533946004	映画鑑賞
5995	40	1	12227	20161110 183000	平日	35.677	139.763	53394611113	5339461111	533946111	映画鑑賞
5995	40	1	12227	20161112 083000	休日	35.677	139.763	53394611113	5339461111	533946111	映画鑑賞

### 【データ集計】

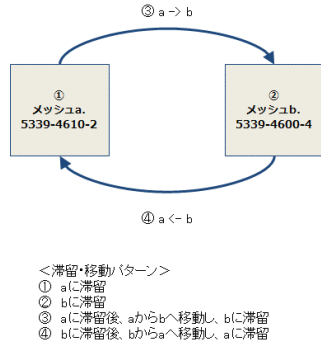
上述の通り、「十分な匿名化」の条件として、データ集計は、各集計結果に該当する人が複数人いることを担保しながら、行われる。担保のレベルは後段の検証で実施するが、個々で先んじて、移動履歴の集計方法に関して整理をする。

下のイメージの通り、移動履歴のレコード数は、日時を無視して集計する場合、必ずしも人数とは一致しないため、定義を誤ると、思わぬところで、秘匿を担保できていないリスク

が生じうる。

本ユースケースの分析では、基本的に、同一人物のレコードは集約してカウントする。

図表 11 移動履歴の集計方法のイメージ (一例)



	Aさん	Bさん
Day 1	a → b	
Day 2		b
Day 3	a	
Day 4	a	a → b
Day 5	a → b	
...		

上記の集計結果		
	人の重複有	重複無し
a	2件	1件
b	1件	1件
a → b	3件	2件
b → a	0件	0件

② 「十分な匿名化」の検証

以上までの前処理後のデータを用いて作成された「十分な匿名化」データのサンプルは、以下ようになる。

図表 12 「十分な匿名化」による評価データのサンプル

ホップ数	ホップ	20161102 (平日)	20161106 (休日)	20161110 (平日)	重複あり	重複なし
0	53394611113	1537	1385	2194	24373	20224
0	53394611142	1180	718	1198	13699	11105
0	53394600444	707	605	816	9740	8557
0	53394600442	589	279	457	5802	5157
0	53394611112	251	301	230	3660	3431
0	53394600441	289	167	264	3190	2687
0	53394600422	243	144	232	2758	2253
0	53394600424	259	141	236	2589	2165
0	53394611122	222	60	221	2597	1930
0	53394611121	201	101	190	2191	1706
0	53394610224	199	91	221	2122	1648
0	53394600443	138	100	135	1817	1633
0	53394611111	130	160	129	1783	1548
0	53394611141	161	55	200	1836	1359
0	53394611133	97	84	129	1344	1095
0	53394600414	150	30	128	1472	877
0	53394600421	99	29	113	979	820
0	53394611132	66	17	77	575	561
0	53394600423	66	34	54	600	508
0	53394611124	73	4	58	636	477
0	53394600432	23	20	31	377	335
0	53394611143	32	12	36	355	326
0	53394611114	38	18	30	360	302
0	53394600434	10	12	9	114	95
0	53394600412	8		5	101	66
0	53394610222	7		7	54	48
0	53394610244	10		5	84	32
1	53394611113→53394600444	13	12	12	161	161
1	53394600444→53394611113	6	4	13	129	129
1	53394611113→53394611112	7	10	8	125	125
1	53394611113→53394600442	4	6	8	104	104
1	53394600444→53394600442	10	5	6	102	100
1	53394611112→53394611113	6	8	10	82	82
1	53394611141→53394611142	9	2	8	84	78
1	53394600442→53394600444	5	6	2	77	77
1	53394611113→53394610224	8	4	9	75	73

※) 表内の 11 桁の数字は、メッシュコードを表している。

本表では、125m 四方のメッシュを取り、どの日に、どこのメッシュにユーザーが何人いるのか、移動したユーザーは、どこに何人いったのかを示している。また、前述の【データ集計】で示したように、移動レコード数と人数が一致しない（同一人物のレコードが複数ある）場合も診ることができるよう、「重複あり」という区分で、その結果を見えるようにしている。

この結果を踏まえ、以降では、エリアの粒度やホップ数等を変えた場合に、どの程度十分な匿名化を満たす情報が作成できるのか、を評価・検証する。

まず、測定対象とした 2 週間全体での、位置情報と人数、非識別性（本報告書では、k-匿名性と同等の意）<sup>2</sup>の評価値としての k 値を整理すると以下のように表される。

図表 13 全期間・全利用者数で見た「十分な匿名化」の評価データの状況

500m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	70,751	7	0.01%	11	0.02%	22	0.03%
0ホップ	69,505	0	0.00%	0	0.00%	0	0.00%
1ホップ	1,189	0	0.00%	0	0.00%	0	0.00%
2ホップ	53	3	5.66%	7	13.21%	18	33.96%
3ホップ	4	4	100.00%	4	100.00%	4	100.00%

250m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	72,589	52	0.07%	82	0.11%	218	0.30%
0ホップ	70,415	0	0.00%	0	0.00%	0	0.00%
1ホップ	2,026	10	0.49%	26	1.28%	88	4.34%
2ホップ	139	35	25.18%	47	33.81%	121	87.05%
3ホップ	5	5	100.00%	5	100.00%	5	100.00%
4ホップ	4	2	50.00%	4	100.00%	4	100.00%

125m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	74,529	257	0.34%	369	0.50%	929	1.25%
0ホップ	70,945	0	0.00%	0	0.00%	0	0.00%
1ホップ	3,247	98	3.02%	180	5.54%	635	19.56%
2ホップ	306	128	41.83%	158	51.63%	263	85.95%
3ホップ	20	20	100.00%	20	100.00%	20	100.00%
4ホップ	8	8	100.00%	8	100.00%	8	100.00%
5ホップ	3	3	100.00%	3	100.00%	3	100.00%

本表は 3 つの表より構成される。一番上の表はメッシュを 500m 四方で区切った際に、そのメッシュに滞留し続けるユーザーが 69,505 人（0 ホップ）、異なるメッシュに一度移動

<sup>2</sup> 対象となるデータセットにおいて、レコード単位で見て、全てのデータ項目について、同じ値を持つレコードが複数ある状態にすること。

したユーザーが 1,189 人 (1 ホップ)、異なるメッシュに移動し、さらに異なるメッシュに移動したユーザーが 53 人 (2 ホップ) いることを意図している。

2 ホップ、3 ホップ移動するユーザーの中には、他と比べて異なる移動をする少数派のユーザーも存在する。2 ホップの場合、 $k=2$  (すなわち、2 人以上同じ移動をしているユーザーがない、つまり、そのユーザー独特の移動をしている) に該当するユーザーは 3 名いる。 $k=3$  と設定すると 7 名、 $k=10$  と設定すると 18 名が、少数派に該当する。

このような少数の挙動は、再識別のリスクを高める危険性を持つため、「十分な匿名化」の出口要件の中で、削除されることになる。この削除の比率を見ていくと、以下の傾向を見ることができる。参考までに下にデータセットで見た秘匿処理の実施状況を示す。

- メッシュの粒度が細かい程、メッシュにいるユーザー数の密度は少なくなり易いため、削除比率は高くなる
- ホップ数を増やすほど、全件数が少なくなるため、削除比率は高くなる
- 少数派の閾値 ( $k$  - 匿名化にたとえば、 $k$  を何に設定するか) を増やすほど、少数と認識されるユーザーの数が増えるため、削除比率は増加する

図表 14 少数派の閾値の変化による評価データの削除領域の広がり  
【元データ】

ホップ数	ホップ	20161102 (平日)	20161106 (休日)	20161110 (平日)	重複あり	重複なし
0	533946004	2,673	1,622	2,549	30,543	24,618
0	533946102	216	91	234	2,262	1,726
0	533946111	4,096	2,998	4,823	54,786	43,161
1	533946004→533946102	1	1		22	21
1	533946004→533946111	21	31	28	412	412
1	533946102→533946004	5	2	6	34	34
1	533946102→533946111	9	7	10	95	91
1	533946111→533946004	44	36	50	541	537
1	533946111→533946102	8	6	12	96	94
2	533946004→533946102→533946004				1	1
2	533946004→533946111→533946004				14	12
2	533946102→533946004→533946102				1	1
2	533946102→533946111→533946004		1		2	2
2	533946102→533946111→533946102				5	5
2	533946111→533946004→533946102				1	1
2	533946111→533946004→533946111	2	3	1	23	23
2	533946111→533946102→533946004	2			2	2
2	533946111→533946102→533946111			2	7	6
3	533946004→533946111→533946004→533946111		1		1	1
3	533946111→533946004→533946111→533946004				1	1
3	533946111→533946004→533946111→533946102		1		1	1
3	533946111→533946102→533946111→533946004		1		1	1

【 k = 2】

ホップ数	ホップ	20161102 (平日)	20161106 (休日)	20161110 (平日)	重複あり	重複なし
0	533946004	2,673	1,622	2,549	30,543	24,618
0	533946102	216	91	234	2,262	1,726
0	533946111	4,096	2,998	4,823	54,786	43,161
1	533946004→533946102				22	21
1	533946004→533946111	21	31	28	412	412
1	533946102→533946004	5	2	6	34	34
1	533946102→533946111	9	7	10	95	91
1	533946111→533946004	44	36	50	541	537
1	533946111→533946102	8	6	12	96	94
2	533946004→533946102→533946004					
2	533946004→533946111→533946004				14	12
2	533946102→533946004→533946102					
2	533946102→533946111→533946004				2	2
2	533946102→533946111→533946102				5	5
2	533946111→533946004→533946102					
2	533946111→533946004→533946111	2	3		23	23
2	533946111→533946102→533946004	2			2	2
2	533946111→533946102→533946111			2	7	6
3	533946004→533946111→533946004→533946111					
3	533946111→533946004→533946111→533946004					
3	533946111→533946004→533946111→533946102					
3	533946111→533946102→533946111→533946004					

【 k = 3】

ホップ数	ホップ	20161102 (平日)	20161106 (休日)	20161110 (平日)	重複あり	重複なし
0	533946004	2,673	1,622	2,549	30,543	24,618
0	533946102	216	91	234	2,262	1,726
0	533946111	4,096	2,998	4,823	54,786	43,161
1	533946004→533946102				22	21
1	533946004→533946111	21	31	28	412	412
1	533946102→533946004	5		6	34	34
1	533946102→533946111	9	7	10	95	91
1	533946111→533946004	44	36	50	541	537
1	533946111→533946102	8	6	12	96	94
2	533946004→533946102→533946004					
2	533946004→533946111→533946004				14	12
2	533946102→533946004→533946102					
2	533946102→533946111→533946004					
2	533946102→533946111→533946102				5	5
2	533946111→533946004→533946102					
2	533946111→533946004→533946111		3		23	23
2	533946111→533946102→533946004					
2	533946111→533946102→533946111				7	6
3	533946004→533946111→533946004→533946111					
3	533946111→533946004→533946111→533946004					
3	533946111→533946004→533946111→533946102					
3	533946111→533946102→533946111→533946004					

【 k = 10】

ホップ数	ホップ	20161102 (平日)	20161106 (休日)	20161110 (平日)	重複あり	重複なし
0	533946004	2,673	1,622	2,549	30,543	24,618
0	533946102	216	91	234	2,262	1,726
0	533946111	4,096	2,998	4,823	54,786	43,161
1	533946004→533946102				22	21
1	533946004→533946111	21	31	28	412	412
1	533946102→533946004				34	34
1	533946102→533946111			10	95	91
1	533946111→533946004	44	36	50	541	537
1	533946111→533946102			12	96	94
2	533946004→533946102→533946004					
2	533946004→533946111→533946004				14	12
2	533946102→533946004→533946102					
2	533946102→533946111→533946004					
2	533946102→533946111→533946102					
2	533946111→533946004→533946102					
2	533946111→533946004→533946111				23	23
2	533946111→533946102→533946004					
2	533946111→533946102→533946111					
3	533946004→533946111→533946004→533946111					
3	533946111→533946004→533946111→533946004					
3	533946111→533946004→533946111→533946102					
3	533946111→533946102→533946111→533946004					

なお、本表と同じ条件のもと、集計方法における人の重複を許した場合は、図表 15 の通りである。

おおよその傾向は、人の重複を廃した場合に類似するが、削除比率や削除件数を個別に見ると、値は異なるものになっている。特に 250 メッシュで 2 ホップの移動を許した場合、 $k$  の値を 10 に設定した際の削除比率は、重複なしで 87.05%、重複ありで 73.79%となり、10 ポイント以上乖離することになる。従って、どちらの考え方に基づいて移動履歴をカウントするのかは、目的とケースに応じて検討すべきである。

続いて、特定期間を切り出して、「十分な匿名化」の情報を作成した場合の状況は、

図表 16 の通りである。期間を特定化することで、対象者数が少なくなり、全期間と比較して総じて削除比率は高くなる傾向にある。

図表 15 人の重複がある場合の全期間・全利用者数で見た「十分な匿名化」の評価データの状況

500m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	88,851	7	0.01%	11	0.01%	23	0.03%
0ホップ	87,591	0	0.00%	0	0.00%	0	0.00%
1ホップ	1,200	0	0.00%	0	0.00%	0	0.00%
2ホップ	56	3	5.36%	7	12.50%	19	33.93%
3ホップ	4	4	100.00%	4	100.00%	4	100.00%

250m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	88,851	52	0.06%	68	0.08%	207	0.23%
0ホップ	86,640	0	0.00%	0	0.00%	0	0.00%
1ホップ	2,055	10	0.49%	26	1.27%	89	4.33%
2ホップ	145	35	24.14%	35	24.14%	107	73.79%
3ホップ	5	5	100.00%	5	100.00%	5	100.00%
4ホップ	6	2	33.33%	2	33.33%	6	100.00%

125m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	88,851	255	0.29%	367	0.41%	938	1.06%
0ホップ	85,208	0	0.00%	0	0.00%	0	0.00%
1ホップ	3,290	98	2.98%	180	5.47%	640	19.45%
2ホップ	321	127	39.56%	155	48.29%	266	82.87%
3ホップ	20	20	100.00%	20	100.00%	20	100.00%
4ホップ	9	7	77.78%	9	100.00%	9	100.00%
5ホップ	3	3	100.00%	3	100.00%	3	100.00%

図表 16 日次で見た「十分な匿名化」の評価データの状況 (1/2)

11月6日(休日)

500m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	4,801	5	0.10%	7	0.15%	23	0.48%
0ホップ	4,711	0	0.00%	0	0.00%	0	0.00%
1ホップ	83	1	1.20%	3	3.61%	16	19.28%
2ホップ	4	1	25.00%	1	25.00%	4	100.00%
3ホップ	3	3	100.00%	3	100.00%	3	100.00%

250m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	4,801	15	0.31%	33	0.69%	78	1.62%
0ホップ	4,667	0	0.00%	0	0.00%	0	0.00%
1ホップ	124	10	8.06%	26	20.97%	68	54.84%
2ホップ	7	2	28.57%	4	57.14%	7	100.00%
3ホップ	2	2	100.00%	2	100.00%	2	100.00%
4ホップ	1	1	100.00%	1	100.00%	1	100.00%

125m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	4,801	82	1.71%	124	2.58%	205	4.27%
0ホップ	4,567	0	0.00%	0	0.00%	4	0.09%
1ホップ	211	62	29.38%	104	49.29%	178	84.36%
2ホップ	16	13	81.25%	13	81.25%	16	100.00%
3ホップ	4	4	100.00%	4	100.00%	4	100.00%
4ホップ	3	3	100.00%	3	100.00%	3	100.00%

図表 17 日次で見た「十分な匿名化」の評価データの状況 (2/2)

11月10日(平日)

500m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	7,715	1	0.01%	3	0.04%	9	0.12%
0ホップ	7,606	0	0.00%	0	0.00%	0	0.00%
1ホップ	106	0	0.00%	0	0.00%	6	5.66%
2ホップ	3	1	33.33%	3	100.00%	3	100.00%

250m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	7,715	17	0.22%	31	0.40%	120	1.56%
0ホップ	7,526	0	0.00%	0	0.00%	5	0.07%
1ホップ	175	7	4.00%	17	9.71%	101	57.71%
2ホップ	14	10	71.43%	14	100.00%	14	100.00%

125m	全件数	k=2		k=3		k=10	
		削除件数	削除比率	削除件数	削除比率	削除件数	削除比率
トータル	7,715	84	1.09%	142	1.84%	301	3.90%
0ホップ	7,405	0	0.00%	0	0.00%	26	0.35%
1ホップ	283	61	21.55%	115	40.64%	248	87.63%
2ホップ	25	21	84.00%	25	100.00%	25	100.00%
3ホップ	1	1	100.00%	1	100.00%	1	100.00%
4ホップ	1	1	100.00%	1	100.00%	1	100.00%



③ 「十分な匿名化」の加工基準に関するまとめ

図表 18 「十分な匿名化」に係る入口要件・出口要件への事業者の対応・評価

指標	指標の特徴及び想定リスク	リスク回避手段	事業者の対応・評価
入口要件			
入口要件	データから個別の通信や特定の個人を識別することができる可能性がある。	氏名は個人識別性が高いため削除する。 端末識別符号等はハッシュ化等により、一方向に仮 ID 化する。	本ケースではユーザーID が端末識別符号等に該当するため、仮 ID 化を行う。
出口要件			
場所の特性	場所自体から、個人の特定につながる可能性がある場所が存在する。特に要配慮情報にかかわる場所や住所等が含まれている場合、リスクは大きい。	特に要配慮情報に関連のある施設・建物付近に滞留していることがあきらかな場合、対象データを削除する、若しくは対象データに限定されないよう分析エリアを広くとる。	本ケースでは、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアを対象として設定。 本ケースに、個人特定のリスクが高い要配慮情報に関連のある建物・施設は含まれていないと認識。 その他、十分なサンプル数を確保 (218,669 ユーザー数)
集団の規模	特定の職業や珍しい趣味嗜好を持つ集団を対象とした場合、集団そのものの規模によっては、対象の限定がしやすくなる。	趣味嗜好をデータ項目として取り扱う場合、出現頻度の少ないカテゴリは、より一般的なカテゴリにまとめサンプル数を確保する。	上述より、不特定多数の人を母集団として、サンプル数も確保 (218,669 ユーザー数) あまり一般的でない趣味については、母集団の数を大きくするために、属性の一般化レベルを上げた。
取得時期の特性	イベントや事件等が特定される場合、他の情報(受領者が独自に入手した情報)との照合により、個人特定のリスクは高まる。	特定のイベントや事件があったことが明らかな場合(特に要配慮情報に関わる場合等)は、対象データを削除する。その他母集団が偏らないようエリアを広くとり、サンプル数を確保する等。	データ取得期間は 2016 年 10 月 31 日～11 月 13 日で、この間、分析対象エリアにて、配慮が必要な特定のイベントや事件等の存在を確認していない。 尚、行政機関による告知、新聞・TV 等で報道されており、大多数の人に周知されているイベントや行事であれば認知可能だが、様々な企業や団体が実施するイベントや行事を把握・認知する事は難しいため、対象データを特定し削除することは難しいと考えられる。
位置の精度	データの精度が細かいほどリスクが大きい。 (例:数メートルの場合、立ち寄った店舗等が判明可能)	緯度・経度(位置情報)は、125m, 250m, 500m メッシュへ変換することにより精度を粗くする(一般化)。	125m、250m、500m メッシュに変換して分析。 単位時間あたりにメッシュに存在するサンプル数が十分確保できる場合、他の掛け合わせる項目の一般化レベルを上げる等の加工を施す場合等は、より小さなサイズを指定する事も可能と考える。

指標	指標の特徴及び想定リスク	リスク回避手段	事業者の対応・評価
移動履歴の期間	<p>移動履歴</p> <ul style="list-style-type: none"> <li>・トレースする期間が長いほど、起点から終点の履歴が長いほど、リスクが大きい。</li> <li>また、ホップ数が多ければ多い程リスクが大きい</li> </ul> <p>・特徴的な場所の推定</p> <p>移動履歴に含まれる位置に個人を特定する為の鍵となる特徴的な場所と解釈可能な場所が含まれる場合（例、自宅住所）、特徴的な場所を使って他の情報（受領者が独自に入手した情報）と照合することにより個人特定がなされる（例、移動履歴から自宅住所→名簿と照合）</p> <p>・パターン性のある移動履歴</p> <p>パターン性のある移動履歴があり、既知である（他手段でも収集されている）場合、両者を照合することで、個人特定のリスクがある。</p> <p>（例、移動履歴にユニークな通勤経路を含む場合、他のサービス等で通勤経路と氏名を把握したデータがある場合、個人特定がなされる）</p> <p>・移動履歴の中に、特定の場所への立ち寄り等、受領者が知り得た情報がある場合、個人特定のリスクがある。</p> <p>（例、目撃者（日付、時間、場所）又は SNS 等での公開情報）</p>	<p>移動履歴</p> <ul style="list-style-type: none"> <li>・適切な K-匿名性を確保するように加工する</li> <li>・特徴的な場所、生活圏（自宅周辺・勤務先通学先周辺）と想定される位置情報は削除する</li> <li>・パターン性のある移動履歴がある場合は削除するか部分削除する等の加工を行う。</li> </ul> <p>仮 ID について適切な有効期間をとる。</p> <ul style="list-style-type: none"> <li>・集計・加工・分析途中の工程では、仮 ID (Hash 化等) へ置換</li> </ul> <p>した上でを行い、加工後は削除する。</p>	<p>本ケースにおいては、125m、250m、500m の各々のメッシュサイズにおいて、Hop 数を 0 から 5 へ変化させて分析。</p> <p>Hop 数の増加に伴い、データ件数が大きく減少する傾向が見られ、本ケースでは、2Hop ないし 3 Hop レベルが限界ではないかと思われる。</p> <p>（Hop 数を増やす為には、データの減少傾向より、元のデータ件数を約 30～50 倍以上増やす必要があると想定）</p>
時間の精度・間隔	<p>データを取得する間隔が短いほど、より詳細な移動履歴情報が蓄積される為、より詳細なトレースが可能となり、個人の特定につながる可能性がある。</p>	<p>データの取得間隔を長く取る。</p> <p>（15分、30分、60分等）</p>	<p>本ケースにおいては、15 分刻みのデータへ加工した。</p> <p>分析結果の利用目的にも依存するが、利用者が徒歩で移動しているケースを想定した場合、メッシュサイズに適したデータの取得間隔が考えられる。</p> <p>例) 例えば一般的な成人の歩行速度 80m/分で移動すると想定した場合、125m メッシュを横切る移動時間は、約 1 分半強となる。</p>

指標	指標の特徴及び想定リスク	リスク回避手段	事業者の対応・評価
			<p>メッシュ内での一時的な滞留が無い場合、</p> <p>125m の場合、1 分半から 2 分程度</p> <p>250m の場合、3 分</p> <p>500m の場合、6 分程度の組み合わせ例が考えられる。</p> <p>ただし、より詳細になるほど、個人の特定につながる可能性があるため、他の付帯情報の一般化レベルを上げる等の配慮が望ましい</p>
対象者数	対象者数が少ないほど個人を特定するリスクが大きい	対象者数を多くする。	<p>本ケースで用いた、2 週間の期間で取得したデータ件数</p> <ul style="list-style-type: none"> <li>位置情報レコード： 2,319,739 レコード</li> <li>ユーザー数： 218,669 ユーザー</li> </ul> <p>前処理後のデータ件数 時間帯：6:00-24:00 で 15 分刻み</p> <ul style="list-style-type: none"> <li>位置情報レコード数： 764,785 レコード</li> <li>ユーザー数： 88,851 ユーザー</li> </ul>
提供までの期間	データ取得から提供までの期間が短いほどリスクが大きい 例えば、よりリアルタイムに近い場合、現在も対象者が付近に存在している可能性がある。 パターン性のある移動履歴がある場合、同一の時間帯に同一の場所に再度来訪している可能性がある。	個人の特定につながる可能性を軽減する為、適切な期間を確保することが望ましい。 または、一般化レベルをあげることが望ましい。	本ケースにおいて具体的な提供までの期間を想定していないため、特段の対応は実施していない。

#### (4) 匿名加工情報に係る検証

##### ① 匿名加工情報のデータセットの作成

匿名加工情報の作成においては、個人情報保護法 36 条 1 項に基づく個人情報保護法施行規則 19 条各号とそれらに対応する個人情報保護法ガイドライン(匿名加工情報編)の記述内容を遵守することが求められる。

##### 【元となるデータの概要】

元となるデータ自体は、「十分な匿名化」において用いたデータと同じである。

匿名加工情報は、通信の秘密に係る情報には適用できないことから、位置情報については、通信の秘密に当たらない Wi-Fi 情報を用いていると想定した。

また、匿名加工情報においては、多種多様な属性情報を付加することが認められるが、本調査研究においては、データセットの準備工数等に鑑み、上述したデータと同様のデータで分析することとした。

なお匿名加工情報の場合、レコード単位での提供となることを想定し、位置情報をつなげた移動履歴としての評価は行わない。

##### 【匿名加工情報作成のための処理】

上述の通り、規則 19 条 1 号～5 号に鑑み、処理を実施する。

##### 【規則 19 条 1 号】特定の個人を識別することができる記述等の削除

- ・ 個人情報に含まれる特定の個人を識別することができる記述等の全部又は一部を削除すること(当該全部又は一部の記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。)

1 号によれば、氏名、住所、生年月日が含まれる個人情報を加工する場合、これらの情報の削除や置き換えが必要となる。

また、会員 ID、氏名等の情報も、同様の措置が必要となる。

本ユースケースでは、ユーザー ID が会員 ID としての性格に近いものの、このデータに関しては、3 号で対処することとする。住所情報は既に市区町村に丸めていることから、会員 ID の他に 1 号に該当し得るデータはない。

##### 【規則 19 条 2 号】個人識別符号の削除

- ・ 個人情報に含まれる個人識別符号の全部を削除すること(当該個人識別符号を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。)

1 号同様、2 号に関しても、本ユースケースでは該当する情報はない。

**【規則 19 条 3 号】情報を相互に連結する符号の削除**

個人情報と当該個人情報に措置を講じて得られる情報を連結する符号(現に個人情報取扱事業者において取り扱う情報を相互に連結する符号に限る。)を削除すること(当該符号を復元することのできる規則性を有しない方法により当該個人情報と当該個人情報に措置を講じて得られる情報を連結することができない符号に置き換えることを含む)。

本ユースケースでは、ユーザーID が該当する。

よって、ユーザーID はハッシュ化等により仮 ID 化している。なお、ユーザーID 以外に個人情報に関わる情報は存在しない。

**【規則 19 条 4 号】特異な記述等の削除**

- ・ 特異な記述等を削除すること(当該特異な記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。)

本ユースケースでは、趣味嗜好情報が該当しうるが、本情報の区分は 10 区分と粗く、特異性も大きくは認められない。

**【規則 19 条 5 号関連】個人情報データベース等の性質を踏まえたその他の措置**

- ・ 前各号に掲げる措置のほか、個人情報に含まれる記述等と当該個人情報を含む個人情報データベース等を構成する他の個人情報に含まれる記述等との差異その他の当該個人情報データベース等の性質を勘案し、その結果を踏まえて適切な措置を講ずること。

本ユースケースでは、日時情報、位置情報、趣味嗜好情報が、該当し得る。

日時については、深夜時間帯(0時～6時)は除いているものの、詳細な時間が明確になるほど、再識別のリスクは高まるおそれがある。よって、本ユースケースでは、上述したデータセット同様、時間の区分を 15 分単位に丸めたものを使用する。

位置情報も、精度が高まるほど、再識別のリスクが高まるおそれがある。また、本ユースケースでは Wi-Fi による測位情報を利用しているが、Wi-Fi では数 m～数十 m の誤差が生じる。よって、上述の通り、ある程度の広さを持ったメッシュ(500m 四方、250m 四方、125m 四方)で位置を認識することとする。

趣味嗜好情報は、区分が細くなるほど、再識別のリスクが高くなる。ただし、本ユースケースでは、10 種類の情報しか用意していないことに鑑み、特段の処理は不要と考える。

以上の規則に則り、次ページ以降にある形で匿名加工情報を作成した。

図表 19 匿名加工情報 評価用データサンプル (1/2)

仮ID	年代	性別	趣味嗜好	市区町村	日付	区分	時間	125mメッシュ	250mメッシュ	500mメッシュ
448929	20	女性	観劇	東京都目黒区	20161031	平日	120000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	131500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	134500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	140000	53394611113	5339461111	533946111
448929	20	女性	観劇	東京都目黒区	20161031	平日	141500	53394611113	5339461111	533946111
448929	20	女性	観劇	東京都目黒区	20161031	平日	144500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	150000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	160000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	164500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	171500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	183000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161031	平日	190000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	110000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	111500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	123000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	131500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	133000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	134500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	140000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	141500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	143000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	150000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	153000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	160000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	161500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	163000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	164500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	180000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	184500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161104	平日	190000	53394611112	5339461111	533946111
448929	20	女性	観劇	東京都目黒区	20161104	平日	191500	53394611112	5339461111	533946111
448929	20	女性	観劇	東京都目黒区	20161110	平日	131500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161110	平日	143000	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161110	平日	144500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161110	平日	151500	53394610224	5339461022	533946102
448929	20	女性	観劇	東京都目黒区	20161110	平日	154500	53394610224	5339461022	533946102



② 匿名加工情報の加工基準に関するまとめ

図表 21 匿名加工情報の加工基準に関するまとめ

指標	想定リスク	リスク回避手段	事業者の評価
規則 19 条 1 号	個人情報に含まれる特定の個人を識別することができる記述等は、個人の特定や識別につながる。	<ul style="list-style-type: none"> <li>・氏名の削除</li> <li>・住所の削除、または、市区町村レベルへの置き換え</li> <li>・生年月日の削除、または、正年月レベルへの置き換え</li> </ul>	<ul style="list-style-type: none"> <li>・氏名は使用しない</li> <li>・住所は市区町村レベルへ置き換え</li> <li>・生年月日は 10 歳刻みの年代レベルへ置き換え</li> </ul> 以上の処理により、リスクを回避
規則 19 条 2 号	個人情報に含まれる個人識別符号は、個人の特定や識別につながる。	<ul style="list-style-type: none"> <li>・生体情報の削除</li> <li>・旅券番号、免許証番号などの削除</li> </ul>	該当する情報は使用していない
規則 19 条 3 号	個人情報と当該個人情報に措置を講じて得られる情報を連結する符号を残しておくことで、個人の特定や識別のリスクがあがる可能性がある。	<ul style="list-style-type: none"> <li>・サービス会員情報に管理用 ID を付与し、連結をしている場合、その管理用 ID を削除</li> </ul>	<ul style="list-style-type: none"> <li>・加工途中に使用した 管理用 ID は削除</li> <li>・管理用 ID と別途新たに生成した ID を付与</li> <li>・ログ等も削除し、情報を連結する可能性を排除</li> </ul> 以上の処理により、リスクを回避
規則 19 条 4 号	特異な記述等があることで、該当する特定の個人を識別するリスクが増加する可能性がある。	<ul style="list-style-type: none"> <li>・症例の極めて少ない病歴を削除</li> <li>・年齢が 116 歳という情報を 90 歳以上に置き換え</li> </ul>	<ul style="list-style-type: none"> <li>・特異な記述を含むデータは使用していない</li> </ul>
規則 19 条 5 号	前各号に掲げる措置のほか、個人情報に含まれる記述等と当該個人情報を含む個人情報データベース等を構成する他の個人情報に含まれる記述等との差異その他の当該個人情報データベース等の性質によっては、特定個人の識別につながるリスクが存在する。	<ul style="list-style-type: none"> <li>・移動履歴を含む個人情報データベース等を加工する際、自宅や職場の推定に繋がる情報を削除</li> <li>・身長 170cm の小学生の身長を 150cm 以上という情報に置き換え</li> </ul>	<ul style="list-style-type: none"> <li>・自宅の推定に繋がるデータ（午前 0 時から 6 時までの間の位置情報）は除外することでリスクを回避</li> <li>・趣味嗜好についてはより一般化レベルの高いカテゴリを使用</li> </ul>



### 3) 観光ケースに係る検証

#### (1) ユースケースの設定

昨年度は、観光のユースケースとして、観光地への外国人観光客の来訪における移動情報の分析を行った。

当ケースでは、外国人の識別は、言語情報によって実施し、情報の収集期間は1ヶ月程度と短い期間に限定しての分析であった。

本年度は、昨年に引き続き、Wi-Fi アクセスポイントへの接続状況から見た外国人を含む観光客の移動状況を、プライバシーに配慮しながら分析し、観光活用および官民連携モデルの具体化に関する検討を実施する。なお、新たな論点として、

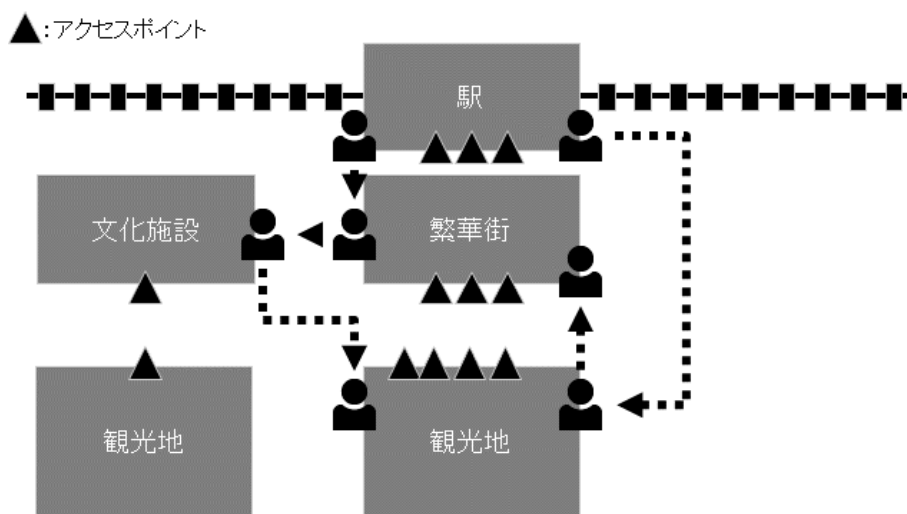
- 長期間に見たデータ・ユーザーを対象とする
- ユーザーのホップ数（移動経路数）をより増やした場合の検討を行う

の2点を取り入れ、検討を行った。

具体的には、Wi-Fi のアクセスポイントと通信端末等を通して得られた利用者の位置情報を用い、外国人観光客の訪れる著名な観光地を対象エリアに設定し、利用者の動線分析を実施する。

なお、本ユースケースの対象とするエリアでは、主要な駅、空港、港、観光地といった公共性の高い場所（特異な場所、例えば専門病院などは含まれていない）において、観光客がどのように移動するのかを検討することを想定している（図表 22 参照）が、以降の分析では地点名を地点 1、地点 2 のように示し、それぞれが、どのような施設なのか、どのような位置関係になっているのか等に関する設定は行っていない。

図表 22 観光地におけるユーザーの移動イメージ

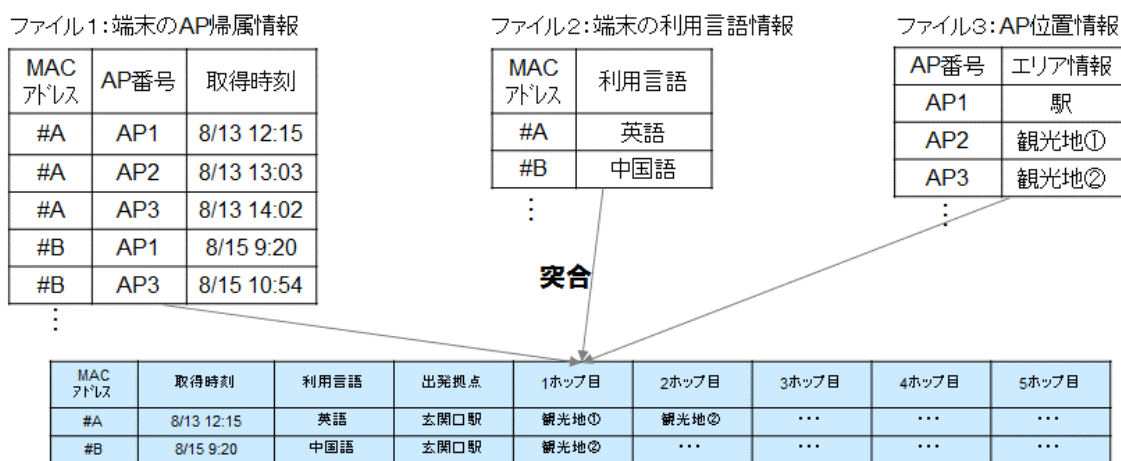


※ 通常、アクセスポイントは施設に複数存在しうるが、今回の分析では、1地点1アクセスポイントと仮定し、データを準備した。

(2) 準備したデータセットの説明

本ユースケースでは、大きく 3 つのデータベース等を利用することで、以下のデータを収集し、活用することを想定した。データセットのイメージは、図表 23 の通りである。

図表 23 準備したデータセット



※AP 番号=アクセスポイント番号

- アクセスポイント位置情報 :  
アクセスポイント、アクセスポイントのエリア情報  
※本ユースケースでは、エリア情報をより端的に、地域 1、地域 2 等と設定
- 端末のアクセスポイント帰属情報 :  
MAC アドレス、アクセスポイント番号、取得時刻
- 端末の利用言語情報  
MAC アドレス、利用言語  
※利用言語は、Wi-Fi の利用の同意を取得する際に選択された言語とみなしている

対象とした期間は、6 ヶ月間であり、これも具体的な開始時期と終了時期については、定めておらず、全体としてダミーなデータとしている。

### (3) 「十分な匿名化」に係る検証

#### ① 「十分な匿名化」のデータセットの作成

「十分な匿名化」のデータセットの作成においては、入口要件としての前処理と加工を終える際の出口要件の2つの加工が必要となる。

検証の前に、「十分な匿名化」のプロセスの図表に則り、「十分な匿名化」のデータセットを作成するまでの流れを以下で説明する。

#### 【元となるデータの概要】

##### ● MAC アドレス数・レコード数：

本調査研究で用意したユーザー数は6.0万人、総レコード数は11.4万件である。

※本レコードは、特定の地点から出発したもののみを対象としている。

※複数ホップ分析が可能なレコードのみを抽出しているため、出発点のみ記録があるようなレコードは対象としないことを前提としている。

##### ● AP 番号・エリア情報：

AP 番号およびエリア番号は、60種類を想定する。

※本来的には、AP 番号は、エリア情報と1:1に紐付かないが、上述の通り、本ユーザースペースでダミーデータをセットするにおいて、AP 番号とエリア番号を1:1に紐付けられるように処理をしている。

※いずれのエリア番号も、著名な観光地等を想定しているため、場所の特性があることは特には想定していない。

※上述により、ある地点から別の地点への移動ルートは59種類あることになる。

##### ● 日時：

日時は年月日時分秒単位で設定をした。

##### ● 利用言語：

利用言語は、以下の7種類の言語を想定する。

※日本語、英語、中国語（繁体字）、中国語（簡体字）、韓国語、フランス語、タイ語である。

元となるデータに係るユーザーのデモグラフィック特性を、図表 24 に示す。

図表 24 ユーザーのデモグラフィック特性（言語別構成比）※全体で6.0万人

日本語	英語	中国語 (繁体字)	中国語 (簡体字)	韓国語	フランス 語	タイ語
60.3%	25.3%	4.6%	3.4%	2.7%	2.3%	1.3%

## 【前処理】

前処理では、上述のデータを、MAC アドレスおよび、アクセスポイント番号を用いて紐付けし、データクレンジングや時間の変換、MAC アドレスの仮 ID 化等を実施した。以下に、前処理における特徴的な事項を記載する。

- データクレンジング：  
Wi-Fi の起動状況等に鑑み、ユーザーの位置情報の履歴が収集できないケースが想定される。このような「不完全」なデータを削除している。  
※本処理は、データセットの作成の段階で、同時に実施している。
- 日時：  
詳細な日時情報は、再識別の可能性を高めることに鑑み、時分秒の単位を 15 分単位に丸めた。
- 位置情報：  
位置情報は上述の通り、著名な観光地等をアクセスポイント情報で認識するとした。なお、先例の商用のユースケースとは異なり、観光においては、滞留の有無ではなく、どのようなルートをユーザーが通ったのかを把握することにニーズがあると考えられるため、特に、「滞留」という考え方は採用していない。すなわち、ある時刻で、ユーザーが Wi-Fi を起動し、アクセスポイントと接続した点を持って、位置情報を認識することとする。  
また、観光地への入り込み客数を延べ人数で把握するため、個人単位で集約して集計することはしない。ただし、個人の識別性への配慮のため、加工時に非識別性の評価は行う。
- MAC アドレス  
MAC アドレスは、仮 ID 化を行うが、本処理は、ダミーデータの作成時に併せて実施している。

これらの処理によって得られた前処理後のデータサンプルは図表 25 の通りである。

図表 25 前処理後のデータサンプル

仮 ID	出発時刻	利用言語	出発拠点	1ホップ目	2ホップ目	3ホップ目	4ホップ目	5ホップ目
dammy000001	2016/1/1 0:00	日本語	玄関口駅	観光地49				
dammy000002	2016/1/1 0:00	日本語	玄関口駅	観光地49	観光地15			
dammy000003	2016/1/1 0:00	日本語	玄関口駅	観光地5	観光地15	観光地2	観光地48	観光地50
dammy000004	2016/1/1 0:00	日本語	玄関口駅	観光地30	観光地16	観光地2	観光地48	
dammy000005	2016/1/1 0:15	英語	玄関口駅	観光地2	観光地48			
dammy000006	2016/1/1 0:15	英語	玄関口駅	観光地13	観光地33	観光地12	観光地13	観光地14
dammy000007	2016/1/1 0:15	日本語	玄関口駅	観光地31	観光地14			
dammy000008	2016/1/1 0:15	日本語	玄関口駅	観光地48	観光地50	観光地54		
dammy000009	2016/1/1 0:30	日本語	玄関口駅	観光地15	観光地54	観光地16	観光地54	観光地19
dammy000010	2016/1/1 0:30	日本語	玄関口駅	観光地4	観光地7	観光地30	観光地31	
dammy000011	2016/1/1 0:30	英語	玄関口駅	観光地3	観光地15			
dammy000012	2016/1/1 0:30	フランス語	玄関口駅	観光地15	観光地51	観光地34	観光地31	観光地29
dammy000013	2016/1/1 0:30	日本語	玄関口駅	観光地49				
dammy000014	2016/1/1 0:45	日本語	玄関口駅	観光地4	観光地12	観光地13	観光地11	
dammy000015	2016/1/1 0:45	英語	玄関口駅	観光地58				
dammy000016	2016/1/1 0:45	タイ語	玄関口駅	観光地3	観光地4	観光地3	観光地4	
dammy000017	2016/1/1 1:00	簡体字	玄関口駅	観光地9	観光地6	観光地31	観光地29	観光地9
dammy000018	2016/1/1 1:00	繁体字	玄関口駅	観光地3	観光地5	観光地8		
dammy000019	2016/1/1 1:00	日本語	玄関口駅	観光地48	観光地50	観光地11	観光地54	
dammy000020	2016/1/1 1:00	英語	玄関口駅	観光地10	観光地49	観光地10	観光地60	観光地5

#### 【データ集計】

上述の通り、「十分な匿名化」の条件として、データ集計は、各集計結果に該当する人が複数人いることを担保しながら、行われる。担保のレベルは後段の検証で実施するが、移動情報のカウントに関して、予め以下を定義する。

レコード数が多数存在するものの、IDのパターンが少ない場合、それは、特定個人が特定のルートを非常に多く利用していることと同義である。このようなレコードは、個人の再識別や推定の可能性を高めるおそれがあることから、同一人物による移動は、「十分な匿名化」においては、1パターンとして集約する。

② 「十分な匿名化」の検証

以上までの前処理後のデータを用いて作成された「十分な匿名化」の評価データのサンプルは、図表 26 のようになる。

図表 27 「十分な匿名化」の評価データのサンプル

#	出発点	第1到着地	第2到着地	第3到着地	レコード数
1	拠点1	拠点2	拠点48	拠点50	2,073
2	拠点1	拠点48	拠点50	拠点54	1,306
3	拠点1	拠点5	拠点4	拠点3	885
4	拠点1	拠点2	拠点1	拠点2	677
5	拠点1	拠点5	拠点3	拠点4	675
6	拠点1	拠点5	拠点49	拠点15	609
7	拠点1	拠点2	拠点48	拠点47	592
8	拠点1	拠点5	拠点4	拠点5	465
9	拠点1	拠点5	拠点3	拠点5	402
10	拠点1	拠点49	拠点50	拠点54	393

11.4 万のレコードのすべてを示すことは困難であるため、本サンプルでは、ホップ数を 3 とした場合に、レコード数が多数存在したルート 10 種類を抽出している。

この結果を踏まえ、以降では、言語を分けずに分析した際のホップ数ごとに見たルートパターンの構成、および言語別に分析した際のホップ数ごとに見たルートパターンの構成を分析する。

図表 28 は、7つの言語を併せた結果を示している。

これを見ると、ホップ数が 1（出発点から別の地点への移動）のルートは 58 ルート存在しており、その中で、10 人以上をカウントできるルートは 98.3%存在することが分かる。ただし、ホップ数が 1 であっても、ユーザー数が 10 人未満のルートは 1.7%（つまり 58 ルートのうち、1 つのルートは明確に、ユーザー数が 10 人未満）となった。

ホップ数が増えるほど、とりうるルートのパターンは多様化し、かつ、多数のユーザー数が存在するルートのパターンも減ってくる。

5 ホップのパターンを容認すると、10 人以上のグルーピングができるルートは 4.2%のみとなり、過半数のルートが 10 人未満のルートとなることが分かる。

図表 28 全言語で見たホップ数とルートとの関係

	1ホップ目	2ホップ目	3ホップ目	4ホップ目	5ホップ目
論理的に発生しうるルート数	59ルート	59の2乗 (3,481)	59の3乗 (20.5万)	59の4乗 (1,200万)	59の5乗 (7億)
実際に発生したルート数	58ルート	1,126ルート	4,239ルート	7,513ルート	9,514ルート
ユーザ数	6.0万	5.6万	5.1万	4.8万	4.4万

K=500~	24.1%	1.9%	0.2%	0.03%	0.02%
K=100~	44.8%	5.2%	1.2%	0.3%	0.1%
K=10~	29.3%	41.6%	19.8%	9.5%	4.1%
K=4~	1.7%	33.8%	48.5%	51.1%	50.9%
K=3	0%	3.9%	7.1%	8.8%	9.9%
K=2	0%	6.0%	9.2%	11.3%	12.9%
K=1	0%	7.7%	14.2%	18.9%	22.1%

図表 29 は、言語別に見た 2 ホップにおけるルートの構成の関係を示している。

図表 29 言語別に見た 2 ホップ数時のルートの構成

	日本語	英語	中国語 (繁体字)	中国語 (簡体字)	韓国語	フランス語	タイ語
K=1以上となるルート数	984	975	498	414	371	329	247
ユーザ数	3.34万	1.42万	0.26万	0.18万	0.15万	0.13万	0.07万

K=500~	1.2%	0.5%	-	-	-	-	-
K=100~	4.0%	2.3%	1.0%	0.2%	0.3%	-	-
K=10~	34.1%	17.1%	7.8%	7.7%	7.3%	7.3%	6.1%
K=4~	39.5%	23.6%	14.9%	10.9%	12.9%	11.6%	9.7%
K=3	9.5%	11.7%	11.6%	9.2%	6.2%	5.8%	6.9%
K=2	6.7%	18.6%	19.3%	16.9%	16.7%	16.4%	15.0%
K=1	5.0%	26.3%	45.4%	55.1%	56.6%	59.0%	62.3%

これを見ると、構成比の多い日本語ユーザーは、10 人以上のグルーピングが可能なルートが 40%近くある一報で、中国語、韓国語、フランス語、タイ語の 4 言語は、そもそもの母集団が少なく、結果として、半数程度ルートがユニーク（1 ユーザーのみが使っている）ルートとなることが分かる。

構成人数が減るほど、再識別や推定のリスクは高まるため、これらの情報は公開されない



対象になりうる。

③ 「十分な匿名化」の加工基準に関するまとめ

以上の議論をもとに、本ユースケースにおける十分な匿名化に係る加工基準は、以下の通りとなる。

図表 30 「十分な匿名化」に係る入口要件・出口要件への事業者の対応・評価

指標	指標の特徴	想定リスク	リスク回避手段	事業者の対応・評価
入口要件				
入口要件	氏名などの個人情報があれば、データは再特定可能となる。	データから個別の通信や特定の個人を識別することができる可能性がある。	氏名は個人識別性が高いため削除する。 端末識別符号等はハッシュ化等により、一方向に仮 ID 化する。	本ケースでは MAC アドレスが端末識別符号等に該当するため、仮 ID 化を、データセット作成時に実施したとみなしている。
出口要件				
場所の特性	特定の職業や同じ趣味・嗜好を持つ母集団が定期的集まる場所や施設の場合、母集団そのものの数によっては、対象の限定がしやすくなる。  機微情報にかかわる場所や住所等が含まれている場合、リスクは大きい。	蓄積された位置情報（移動履歴）と特定の職業や同じ趣味嗜好から、特徴的な行動習慣や移動パターンが読み取れる場合、個人の特定につながる可能性がある。	特定の職業や同じ趣味嗜好を持つ母集団が定期的集まる場所や施設が含まれるのを避ける。  機微情報に関連のある施設・建物付近に滞留していることがあきらかな場合、対象データは削除することが望ましい。（削除）	本ケースでは、アクセスポイントを著名な観光地レベルで集約し、観光地等を対象としているため、場所の特性は少ないものと評価される。  ただし、住宅地や地域の繁華街等にまで範囲を広げた場合は、場所の区分を粗くする（丸め）などの処理が必要になる可能性もあり、このような処理をすることが望ましい。
母集団の特性	特定の職業や同じ趣味嗜好を持つ母集団を対象とした場合、母集団そのものの数によっては、対象の限定がしやすくなる。	蓄積された位置情報（移動履歴）と特定の職業や同じ趣味嗜好から、特徴的な行動習慣や移動パターンが読み取れる場合、個人の特定につながる可	不特定多数の人が来訪する場所を対象とする。	上述に加え、サンプル数が少なく、あまり一般的でない言語情報を分析する場合は、個人の移動パターンを読みとりやすくなる可能性があるため、母集団の

指標	指標の特徴	想定リスク	リスク回避手段	事業者の対応・評価
		能性がある。		数を大きくするべく、属性の一般化レベルを上げる等の加工が望ましい。
取得時期の特性	イベントや事件等が特定される場合、リスクは大きい	特定のイベントや事件のあった日や期間と一致する可能性がある場合、他の情報（受領者が独自に入手した情報）との照合により、個人特定のリスクがある。	特定のイベントや事件があったことが明らかな場合は、対象データを削除する等の処理が望ましい。 (削除)	本ケースは、全てのデータをダミーとし、時期に係る検討も実施していない。  ただし、実際のデータを用いる際には、当該エリアにて、特定のイベントや事件等は確認していないが、配慮することが望ましい。  ただし、行政機関による告知、新聞・TV等で報道されており、大多数の人に周知されているイベントや行事であれば認知可能だが、様々な企業や団体が実施するイベントや行事を把握・認知する事は難しいと考えられる。
位置の精度	精度が細かいほどリスクが大きい。	データの精度により、位置の特定の可能なレベルが異なる。(例、数センチ/数十センチ/数メートル/数十メートル/市区町村/地域)数メートルの場合、立ち寄った店舗等が判明可能。	・緯度・経度(位置情報)は、125m, 250m, 500mメッシュへ変換する、特定の施設のみと紐付けることにより精度を粗くする(一般化)。	本ユースケースでは、特定施設と位置情報を紐付け、特定施設は著名な観光地等と設定しているため、個人の推定、再識別リスクは低い。 ただし、対象施設を広げる場合には、個人の行動の推定に繋がるおそれも高まることから、それらの情報は削除する、または丸めることが望ましい。
移動履歴の期間	移動履歴の蓄積期間が長いほどリスクが大きい。  有効期間が長いほどリスクが大きい	移動履歴 ・トレースする期間が長いほど、起点から終点の履歴が長いほど、リスクが大きい。また、ホップ数が多ければ多い程リスクが	移動履歴 ・匿名性を適切に確保するように加工する  ・特徴的な場所、生活圈(自宅周辺・勤	本ケースにおいては、ホップ数を1~5として、分析を試みた。  ホップ数が増えれば、匿名性を担保し難いルートが増える傾向があ

指標	指標の特徴	想定リスク	リスク回避手段	事業者の対応・評価
		<p>大きい</p> <ul style="list-style-type: none"> <li>特徴的な場所の推定 移動履歴に含まれる位置に個人を特定する為の鍵となる特徴的な場所と解釈可能な場所が含まれる場合(例、自宅住所)、特徴的な場所を使って他の情報(受領者が独自に入手した情報)と照合することにより個人特定がなされる (例、移動履歴から自宅住所→名簿と照合)</li> <li>パターン性のある移動履歴 パターン性のある移動履歴があり、既知である(他手段でも収集されている)場合、両者を照合することで、個人特定のリスクがある。 (例、移動履歴にユニークな通勤経路を含む場合、他のサービス等で通勤経路と氏名を把握したデータがある場合、個人特定がなされる)</li> <li>移動履歴の中に、特定の場所への立ち寄り等、受領者が知り得た情報がある場合、個人特定のリスクがある。 (例、目撃者(日付、時間、場所)又はSNS等での公開情報)</li> </ul>	<p>務先通学先周辺)と想定される位置情報は削除する</p> <ul style="list-style-type: none"> <li>パターン性のある移動履歴がある場合は削除するか部分削除する等の加工を行う。</li> </ul> <p>仮IDの有効期間</p> <ul style="list-style-type: none"> <li>集計・加工・分析途中の工程では、仮ID(Hash化等)へ置換した上でを行い、加工後は削除する。 (削除)</li> </ul>	<p>るため、ホップ数は一定程度に抑えるか、母集団の特性を出さないような加工を行う必要があるものと考えられる。</p>
時間の精度・	間隔が狭いほどリスクが大きい	<ul style="list-style-type: none"> <li>データを取得する間隔が短いほど、より詳細な移動履歴情報が蓄積される為、より詳細なトレースが可能となり、個人の特定</li> </ul>	<p>データの取得間隔を長く取る。 (15分、30分、60分等)</p>	<p>本ケースにおいては、15分刻みのデータへ加工した。</p> <p>また、「十分な匿名化」の中では、時間の概念</p>

指標	指標の特徴	想定リスク	リスク回避手段	事業者の対応・評価
間隔		につながる可能性がある。		を排除した集計を実施している。  目的に応じて、例えば特定の日にち、期間での分析をするにおいては、対象となるユーザー数も減るため、データを見つつ、時間の間隔等を調整していくことが望ましい。
対象者数	少ないほどリスクが大きい	対象者数 対象数が多ければ多い程、多くの人数から個人を特定する確率が低くなる為、多い程リスクは低い。	対象者数を多くする。	本ケースで用いた対象者数は6万人
提供までの期間	データ取得から提供までの期間が短いほどリスクが大きい	・データを取得してから提供するまでの間隔が短いほど、個人の特定につながる可能性がある。 例えば、よりリアルタイムに近い場合、現在も対象者が付近に存在している可能性がある。 パターン性のある移動履歴がある場合、同一の時間帯に同一の場所に再度来訪している可能性がある。	個人の特定につながる可能性を軽減する為、適切な期間を確保することが望ましい。 または、一般化レベルをあげることが望ましい。	本ケースにおいて具体的な提供までの期間を想定していないが、情報を利用する側のニーズとしては、より短い間隔での提供が要望されるケースもあり得ると思われる。 データ取得から提供までの期間が短ければ短いほど、位置情報及び付帯情報の一般化レベルをあげることにより、リスクを軽減・回避する加工を施すことが望ましいと思われる。

#### 4) 匿名加工情報に係る検証

##### (1) 匿名加工情報のデータセットの作成

匿名加工情報の作成においては、個人情報保護法 36 条 1 項に基づく個人情報保護法施行規則 19 条各号とそれらに対応する個人情報保護法ガイドライン（匿名加工情報編）の記述内容を遵守することが求められる。

以降で、どのようなデータを、どのように加工するかに関して述べる。

##### 【元となるデータの概要】

元となるデータ自体は、「十分な匿名化」において用いたデータと同じである。

匿名加工情報は、通信の秘密に係る情報には適用できないことから、位置情報については、通信の秘密に当たらない Wi-Fi 情報を用いていると想定した。

なお、匿名加工情報の作成においては、本来的には、データセットに個人情報が含まれる必要があるが、本ユースケースでは、個人の氏名など、陽として個人情報が含まれていないデータセットで、試験的に匿名加工情報を作成することとした。

また、匿名加工情報においては、多種多様な属性情報を付加することが認められるが、本調査研究においては、データセットの準備工数等に鑑み、上述したデータと同様のデータで分析することとした。

なお匿名加工情報の場合、レコード単位での提供となることを想定し、位置情報をつなげた移動履歴としての評価は行わない。

##### 【匿名加工情報作成のための処理】

上述の通り、規則 19 条 1 号～5 号に鑑み、処理を実施する。

##### 【規則 19 条 1 号】特定の個人を識別することができる記述等の削除

- ・ 個人情報に含まれる特定の個人を識別することができる記述等の全部又は一部を削除すること(当該全部又は一部の記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。)

1 号によれば、氏名、住所、生年月日が含まれる個人情報を加工する場合、これらの情報の削除や置き換えが必要となる。

また、会員 ID、氏名等の情報も、同様の措置が必要となる。

本ユースケースでは、MAC アドレスが会員 ID としての性格に近いものの、このデータに関しては、3 号で対処することとする。

##### 【規則 19 条 2 号】個人識別符号の削除

- ・ 個人情報に含まれる個人識別符号の全部を削除すること(当該個人識別符号を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。)

1号同様、2号に関しても、本ユースケースでは該当する情報はない。

**【規則 19 条 3 号】**情報を相互に連結する符号の削除

個人情報と当該個人情報に措置を講じて得られる情報を連結する符号(現に個人情報取扱事業者において取り扱う情報を相互に連結する符号に限る。)を削除すること(当該符号を復元することのできる規則性を有しない方法により当該個人情報と当該個人情報に措置を講じて得られる情報を連結することができない符号に置き換えることを含む)。

本ユースケースでは、MAC アドレスおよびアクセスポイント番号が該当する。

よって、MAC アドレスはハッシュ化等により仮 ID 化している。

アクセスポイントは、施設・地点名に返還をし、その後はデータを削除する。

**【規則 19 条 4 号】**特異な記述等の削除

- ・ 特異な記述等を削除すること(当該特異な記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。)

本ユースケースでは、言語情報が該当しうるが、本情報の区分は7区分と粗く、特異性も大きくは認められない。

**【規則 19 条 5 号関連】**個人情報データベース等の性質を踏まえたその他の措置

- ・ 前各号に掲げる措置のほか、個人情報に含まれる記述等と当該個人情報を含む個人情報データベース等を構成する他の個人情報に含まれる記述等との差異その他の当該個人情報データベース等の性質を勘案し、その結果を踏まえて適切な措置を講ずること。

本ユースケースでは、時間情報と言語情報が、該当し得る。

時間については、精度を上げれば上げるほど、識別や推定の可能性が高くなるため、15分刻みに丸める。

特に韓国語、フランス語、タイ語のユーザー数が少ない中で、移動パターンを多様化させると、再識別のリスクは高まるおそれがある。よって、本ユースケースでは、上述したデータセットは、必要に応じて削除等が必要になる。以上の規則を踏まえて作成される匿名加工情報は、結果として、以上で示した、「十分な匿名化」における前処理後

のデータと同等で十分なものとする。

図表 31 匿名加工情報サンプル

仮 ID	出発時刻	利用言語	出発地点	1ホップ目	2ホップ目	3ホップ目	4ホップ目	5ホップ目
dammy000001	2016/1/1 0:00	日本語	玄関口駅	観光地49				
dammy000002	2016/1/1 0:00	日本語	玄関口駅	観光地49	観光地15			
dammy000003	2016/1/1 0:00	日本語	玄関口駅	観光地5	観光地15	観光地2	観光地48	観光地50
dammy000004	2016/1/1 0:00	日本語	玄関口駅	観光地30	観光地16	観光地2	観光地48	
dammy000005	2016/1/1 0:15	英語	玄関口駅	観光地2	観光地48			
dammy000006	2016/1/1 0:15	英語	玄関口駅	観光地13	観光地33	観光地12	観光地13	観光地14
dammy000007	2016/1/1 0:15	日本語	玄関口駅	観光地31	観光地14			
dammy000008	2016/1/1 0:15	日本語	玄関口駅	観光地48	観光地50	観光地54		
dammy000009	2016/1/1 0:30	日本語	玄関口駅	観光地15	観光地54	観光地16	観光地54	観光地19
dammy000010	2016/1/1 0:30	日本語	玄関口駅	観光地4	観光地7	観光地30	観光地31	
dammy000011	2016/1/1 0:30	英語	玄関口駅	観光地3	観光地15			
dammy000012	2016/1/1 0:30	フランス語	玄関口駅	観光地15	観光地51	観光地34	観光地31	観光地29
dammy000013	2016/1/1 0:30	日本語	玄関口駅	観光地49				
dammy000014	2016/1/1 0:45	日本語	玄関口駅	観光地4	観光地12	観光地13	観光地11	
dammy000015	2016/1/1 0:45	英語	玄関口駅	観光地58				
dammy000016	2016/1/1 0:45	タイ語	玄関口駅	観光地3	観光地4	観光地3	観光地4	
dammy000017	2016/1/1 1:00	簡体字	玄関口駅	観光地9	観光地6	観光地31	観光地29	観光地9
dammy000018	2016/1/1 1:00	繁体字	玄関口駅	観光地3	観光地5	観光地8		
dammy000019	2016/1/1 1:00	日本語	玄関口駅	観光地48	観光地50	観光地11	観光地54	
dammy000020	2016/1/1 1:00	英語	玄関口駅	観光地10	観光地49	観光地10	観光地60	観光地5

② 匿名加工情報の加工基準に関するまとめ

以上の議論をもとに、本ユースケースにおける十分な匿名化に係る加工基準は、以下の通りとなる。

図表 32 匿名加工情報の加工基準に関するまとめ

指標	想定リスク	リスク回避手段	事業者の評価
規則 19 条 1 号	個人情報に含まれる特定の個人を識別することができる記述等は、個人の特定や識別につながる。	<ul style="list-style-type: none"> <li>氏名の削除</li> <li>住所の削除、または、市区町村レベルへの置き換え</li> <li>生年月日の削除、または、正年月レベルへの置き換え</li> </ul>	本ユースケースでは該当する情報はない。 MAC アドレスについては 3 号で処理をする。
規則 19 条 2 号	個人情報に含まれる個人識別符号は、個人の特定や識別につながる。	<ul style="list-style-type: none"> <li>生体情報の削除</li> <li>旅券番号、免許証番号などの削除</li> </ul>	本ユースケースでは該当する情報はない。
規則 19 条 3 号	個人情報と当該個人情報に措置を講じて得られる情報を連結する符号を残しておくことで、個人の特定や識別のリスクがあがる可能性がある。	<ul style="list-style-type: none"> <li>サービス会員情報に管理用 ID を付与し、連結をしている場合、その管理用 ID を削除</li> </ul>	MAC アドレスは、一方向にハッシュ化する。

指標	想定リスク	リスク回避手段	事業者の評価
規則 19 条 4 号	特異な記述等があることで、該当する特定の個人を識別するリスクが増加する可能性がある。	<ul style="list-style-type: none"> <li>・症例の極めて少ない病歴を削除</li> <li>・年齢が 116 歳という情報を 90 歳以上に置き換え</li> </ul>	本ユースケースでは該当する情報はない。
規則 19 条 5 号	前各号に掲げる措置のほか、個人情報に含まれる記述等と当該個人情報を含む個人情報データベース等を構成する他の個人情報に含まれる記述等との差異その他の当該個人情報データベース等の性質によっては、特定個人の識別につながるリスクが存在する。	<ul style="list-style-type: none"> <li>・移動履歴を含む個人情報データベース等を加工する際、自宅や職場の推定に繋がる情報を削除</li> <li>・慎重 170cm の小学生の身長を 150cm 以上という情報に置き換え</li> </ul>	言語情報は、絶対的なユーザーが少ないものについては、再識別の可能性がある場合は削除することが望ましい。

#### 4) 検証結果の整理

以上までで、商用、観光の 2 つのユースケースを用い、プライバシーに配慮をした「十分な匿名化」と匿名加工情報の作成に係る検討を実施した。これらの情報の作成方法に関する検証結果として、以下が示唆として挙げられる。

##### 【「十分な匿名化」・匿名加工情報に共通して言えること】

- 位置情報に関する集計の手法は、矛盾などを生じさせないように定義をしておくことが必要である。

##### 【「十分な匿名化」の検討】

- 「十分な匿名化」のための加工過程では、作成されるデータは非常に多様にあり、本データの利用目的に合わせて、付帯情報を制限する、位置情報や時間情報の精度を緩める等の措置をすることが必要となる。
- 「十分な匿名化」のための加工過程では、場所の特性を考慮する必要がある。住宅地や人口密度の低いエリアを対象とする場合は、その精度を粗くする等の措置が必要となる。

##### 【匿名加工情報の検討】

- 匿名加工情報は、利用目的や利用シーンによらず、汎用的な情報として加工できる可能性がある。ただし、住宅地を対象とした匿名加工情報や、深夜時間帯を含む匿名加工情報は、個人の推定や再識別のリスクを高めるおそれがあるため、十分に配慮して加工をする必要がある。



## 2. 加工方法等情報に係る安全管理措置の基準の検証

### 1) 安全管理措置に関するルールの整理

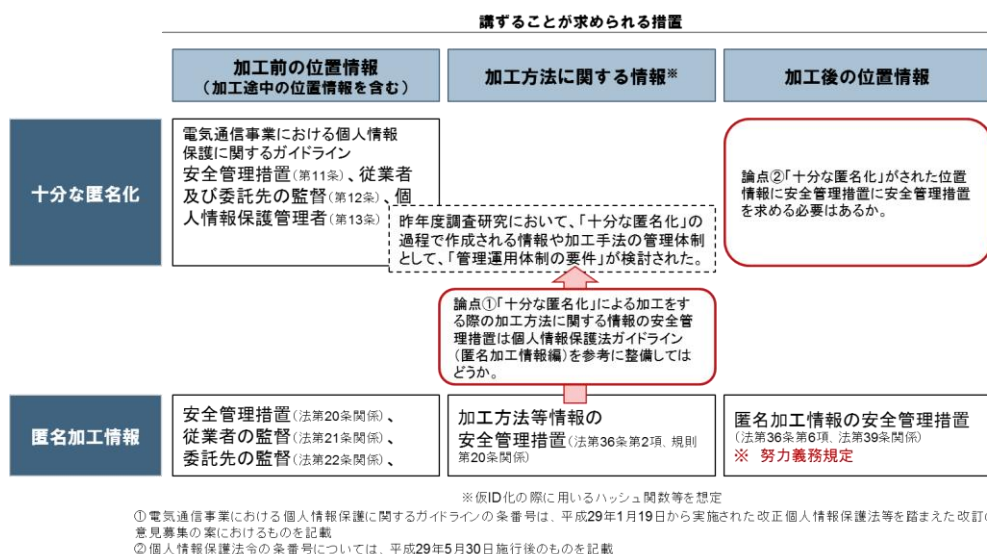
#### (1) 論点の設定

「十分な匿名化」による加工を行う際には、通信の秘密の保護及び個人情報の安全管理のために必要かつ適切な措置（以下、安全管理措置）を講ずることが求められるが、そのルールの整備にあたっては、改正個人情報保護法において整備された匿名加工情報の法制度を参照しながら、以下の通り整理を行い、論点を設定した。

#### < 論点 >

- ① 「十分な匿名化」による加工を行う際の加工方法に関する情報<sup>3</sup>の安全管理措置は個人情報保護法ガイドライン（匿名加工情報編）を参考に整備してはどうか。
- ② 「十分な匿名化」がされた位置情報に安全管理措置を求める必要はあるか。

図表 33 「十分な匿名化」及び匿名加工情報の安全管理措置に関するルールの整理



#### (2) 論点の検証結果

##### < 論点① >

「十分な匿名化」による加工を行う際の加工方法に関する情報の安全管理措置として講ずることが求められる事項は個人情報保護法ガイドライン（匿名加工情報編）の加工

<sup>3</sup> 仮IDに置き換える際の置き換えアルゴリズムに用いられる入力情報に関する情報等を想定

方法等情報の安全管理措置（法第 36 条第 2 項関係、規則第 20 条関係）を参考に整備する。

なお、加工方法に関する情報を用いることで、加工後の位置情報から特定の個人や個別の通信が識別されるおそれがあるため、安全管理措置の基準は、加工前の位置情報に対する措置と同等の水準を講ずることが望ましい。

また、「十分な匿名化」の加工途中の位置情報は、通信の秘密に該当するものとして、通信の秘密と同等の安全管理措置を講じなければならない。

#### < 論点② >

「十分な匿名化」がされた位置情報は技術的に相当程度、特定の個人や個別の通信を識別されるリスクが低減されていることから、電気通信事業者に対して特段の安全管理措置を講ずることは求めない。

## 2) 匿名加工情報の適正な取扱いを確保するために必要な措置の検証

元の個人情報に含まれる個人を特定する記述等に対して、個人情報保護の観点から「電気通信事業における個人情報の保護に関するガイドライン」における「匿名加工情報の作成等（第 28 条 2 項）」に掲げる加工方法に係る安全管理措置に対応する措置が必要である。

また、加工途中の位置情報に対して、通信の秘密を保護するため、「電気通信事業における個人情報の保護に関するガイドライン」における「安全管理措置（第 11 条）」、「従業者及び委託先の監督（第 12 条）」、「個人情報保護管理者（第 13 条）」に対応する措置が必要である。

これらの措置を技術的に満たすものとして暗号技術の導入が期待されている。暗号により、保存される個人情報や位置情報を秘匿することができ、情報漏えい時の対応に役立てることができる。

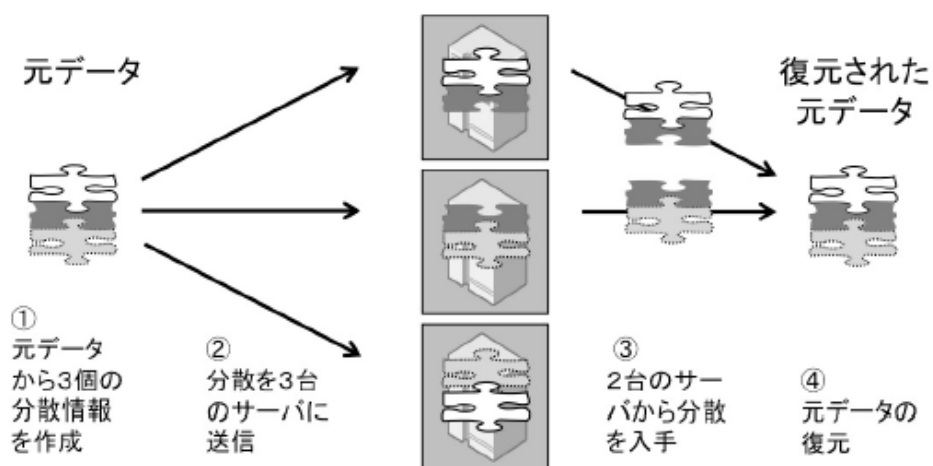
位置情報の十分な匿名化はこれらの情報を（異なる情報源から集めて）加工するものであり、その処理全体を通しての安全管理をより効果的に行うことができる技術が求められている。

秘密計算技術とは暗号の応用技術で、データを暗号化したまま加工をすることができる。秘密計算を十分な匿名化に適用した場合は、たとえば、元の個人情報と通信の位置情報を暗号化し、そのまま両者を突き合わせ、十分な匿名化への加工の結果を出力することが期待できる。この利用方法における従来の暗号化との差分は、単純な暗号化では、個人情報と移動履歴の突き合わせや加工する際は暗号データを復号しなければならないが、秘密計算ではその処理も暗号化したまま行うことができる。

以降ではこの秘密計算技術の概略を述べることにする<sup>4</sup>。実装形態として、暗号技術に秘密分散を用いる秘密分散・秘密計算技術を例に説明する。

秘密分散のイメージは、図表 34 のように描かれる。

図表 34 秘密分散のイメージ

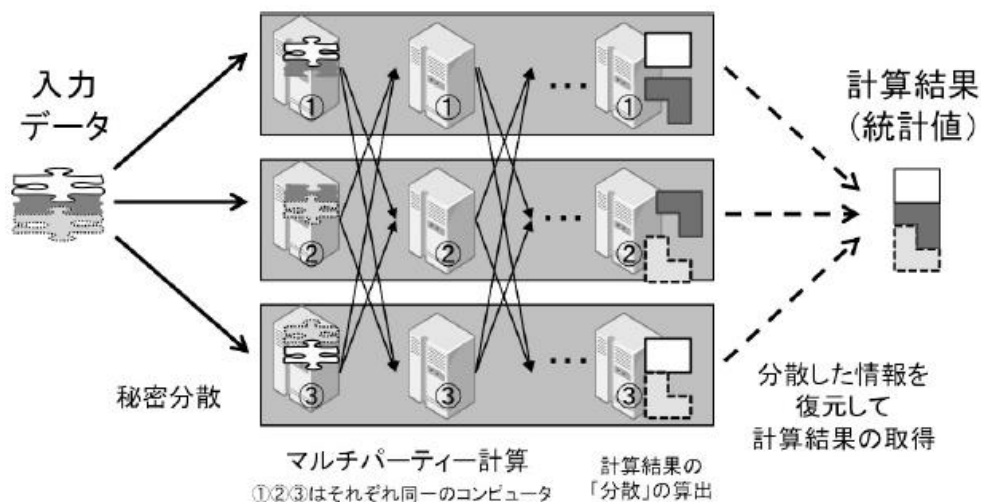


秘密分散は、秘密にする情報を、複数に分けることによって保護するアルゴリズムであり、分散した情報を一定数集めることによって、情報を復元することができ、逆に一定数集めない限り、秘密が漏れることはない技術である。

図表 35 は、情報を 3 つに分散させ、そのうち、2 つを集めることによって復元できる秘密分散を示している。この秘密分散による安全性をベースにし、秘密計算は、データを秘匿したまま、各種計算を行うことと定義される。

図表 35 秘密計算のイメージ

<sup>4</sup> 「公的統計の高度な二次的利用のための秘密計算技術の応用の研究」総務省統計研修所リサーチペーパー34号 P.18~P.19



入力データは秘密分散がされ、複数のコンピュータに配置されている。複数のコンピュータを横串で見ることができず、分析をしたいものは、分析に係る処理を最終的な統計量としての結果のみを見ることができる。

この技術は、民間事業者が企業秘密として厳重に管理している創薬等の化合物データベースの検索技術等の開発に用いられている<sup>5</sup>ほか、世界各国で研究開発が行われているとされている。

我が国では、NTT や日本電気が高速な秘密計算手法を開発している。さらに、バイオデータの分析のための秘密計算の研究が、筑波大学や産業技術総合研究所で進められている。

諸外国においては、エストニアでは Sharemind (企業財務データ分析) や、デンマークでは VIFF (オークション等) といった研究開発が行われている。

また、運用面から、匿名加工情報の適正な取扱いを確保するための措置として、プライバシー影響評価 (PIA) の実施が有効である。PIA の実施方法は、平成 27 年度の調査研究において「十分な匿名化」において講じるべき措置の一つとして整理しており、これを匿名加工情報においても適用することはプライバシー保護の観点からは有効である。

### 3) 今後の展開に向けて

本調査研究においては、「十分な匿名化」がされた位置情報は技術的に相当程度、特定の個人や個別の通信の識別リスクが低減されていることから、電気通信事業者に対し特段の安全管理措置を講ずることは求めないとした。しかし、今後の制度改正や技術の進展を踏まえ、定期的に安全管理措置の見直しを行って、所要の措置を講ずることが望まれる。

<sup>5</sup> 国立研究開発法人産業技術総合研究所 (産総研) ウェブサイト「秘密計算による化合物データベースの検索技術」(2011年11月1日)

### 3. 「十分な匿名化」と匿名加工情報の作成の方法等についての関係等の整理

「十分な匿名化」と匿名加工情報の作成の方法等について、関係等を整理した。特に、電気通信事業者が取り扱う位置情報について、通信の秘密に該当する場合と、通信の秘密に該当しない場合とで、匿名加工の程度にどのような違いがあるのかを検討した。

#### 1) 論点の設定

電気通信事業者が扱う位置情報が通信の秘密への該当するかどうか（通秘／非通秘）の区分と「十分な匿名化」をした情報（対象は個人情報に限らない）と匿名加工情報（対象は個人情報に該当するもの）の区分との2×2のマトリックスを用いて、二つの論点を設定した（図表 35）。

<論点>

- ① そもそも通秘に該当する位置情報を匿名加工 情報として利用することは可能か
- ② 非通秘の位置情報に対して、「十分な匿名化」の水準を求めることの当否

図表 36 「十分な匿名化」と匿名加工情報の作成の方法等についての関係等に係る論点

		作成方法、通知・同意、安全管理措置等	
		「十分な匿名化」をした情報 (対象は個人情報に限らない)	匿名加工情報 (対象は個人情報に該当するもの)
(電気通信事業者が扱うもの) 位置情報	通秘	昨年度の調査事業の対象	<div style="border: 1px solid black; padding: 5px;"> <p style="text-align: center;"><b>論点①</b></p> <p>そもそも通秘に該当する位置情報を匿名加工情報として利用することは可能か？</p> </div>
	非通秘	<p style="text-align: center;"><b>論点② 「十分な匿名化」の水準を求めることの当否</b></p>	

## 2) 論点の検証結果

「論点① そもそも通秘に該当する位置情報を匿名加工 情報として利用することは可能か」について、制度面、技術面、規律面で整理を行って検証した結果を次に示す。

### ○制度面における整理

「十分な匿名化」をした情報と匿名加工情報について、法的根拠、関連文書、保護の趣旨を整理したところ、両者には大きな差異のあることが認められる（図表 36）。

なお関連文書で掲げた「位置情報プライバシーレポート」では、通秘に該当する位置情報の利用に当たっては有効な同意が必要であり、同意無しの利用は認められてきていない。この考え方を今後も維持するならば、匿名加工情報としての利用は認められないこととなる。

図表 37 論点①に係る制度面における整理

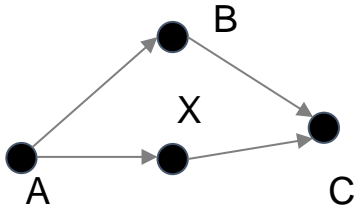
	「十分な匿名化」をした情報	匿名加工情報
対象とする情報	<ul style="list-style-type: none"> <li>電気通信事業者が取り扱う位置情報（通信の秘密に該当するもの）</li> </ul>	<ul style="list-style-type: none"> <li>個人情報にひもづく位置情報</li> </ul>
法的根拠	<ul style="list-style-type: none"> <li>電気通信事業法（通信の秘密の保護）</li> <li>プライバシーの保護</li> </ul>	<ul style="list-style-type: none"> <li>改正個人情報保護法（※2017年5月30日施行予定）</li> </ul>
関連文書	<ul style="list-style-type: none"> <li>位置情報プライバシーレポート</li> <li>電気通信事業における個人情報保護に関するガイドライン（※2017年5月30日施行予定）</li> </ul>	<ul style="list-style-type: none"> <li>個人情報保護法ガイドライン（匿名加工情報編）</li> <li>電気通信事業における個人情報保護に関するガイドライン（※2017年5月30日施行予定）</li> </ul>
保護の趣旨	<ul style="list-style-type: none"> <li>通信の秘密及びプライバシーの保護</li> <li>個人だけでなく、個別の通信を特定されない水準までの保護をする。</li> </ul>	<ul style="list-style-type: none"> <li>個人情報の適正な取扱いに関する、個人の権利利益の保護</li> <li>特定個人の識別性が排除される水準までの保護をする。</li> </ul>

### ○技術面における整理

「十分な匿名化」をした情報と匿名加工情報について、技術面のポイントと具体的なアウトプットのイメージを整理した（図表 37、38）。

「十分な匿名化」では、通信の秘密保護の趣旨から、個人だけでなく個別の通信を特定されないレベルでの加工が求められるため、事実上個人情報保護を目的とする匿名加工情報よりも厳しい匿名化を必要としている。

図表 38 論点①に係る技術面における整理

「十分な匿名化」をした情報														
<p>【ポイント】</p> <ul style="list-style-type: none"> <li>位置情報の履歴、属性を含むデータセットに対して、非識別性を担保できるように集計したもの</li> <li>k-匿名性で評価した場合、いかなるケースにおいても、kは非識別性を確保できる水準であることが求められる</li> </ul>														
<p>【イメージ】</p> <div style="display: flex; align-items: center;">  <table border="1" style="margin-left: 20px;"> <thead> <tr> <th>属性</th> <th>移動経路</th> <th>件数</th> </tr> </thead> <tbody> <tr> <td>20代・男性・千代田区在住</td> <td>A=&gt;B=&gt;C</td> <td>10</td> </tr> <tr> <td>20代・男性・千代田区在住</td> <td>A=&gt;X=&gt;C</td> <td>5</td> </tr> <tr> <td>...</td> <td>...</td> <td>..</td> </tr> </tbody> </table> </div>			属性	移動経路	件数	20代・男性・千代田区在住	A=>B=>C	10	20代・男性・千代田区在住	A=>X=>C	5	...	...	..
属性	移動経路	件数												
20代・男性・千代田区在住	A=>B=>C	10												
20代・男性・千代田区在住	A=>X=>C	5												
...	...	..												
<p>20代男性で、 A=&gt;B=&gt;C という移動をした人が、 12月の休日に、k人いる</p>														

図表 39 論点①に係る技術面における整理 (つづき)

匿名加工情報																															
<p>【ポイント】</p> <ul style="list-style-type: none"> <li>位置情報の履歴、属性を含むデータセットを、特定の個人が識別することができないように加工し、かつ元の個人情報を復元することができないようにしたもの</li> <li>利用ニーズに応じて (Custom made)、位置情報の期間や位置精度、属性の粒度を調整して作成する</li> <li>k-匿名性で評価した場合、ケースによっては、<math>k \geq 2</math> となるように加工することは必ずしも求められない。</li> </ul>																															
<p>【イメージ】</p> <table border="1" style="width: 100%;"> <thead> <tr> <th>仮ID</th> <th>属性</th> <th>仮ID</th> <th>移動履歴</th> </tr> </thead> <tbody> <tr> <td>001</td> <td>20代・男性・千代田区在住</td> <td>001</td> <td>日時 A⇒B⇒C</td> </tr> <tr> <td>002</td> <td>20代・男性・千代田区在住</td> <td>001</td> <td>日時 A⇒X⇒C</td> </tr> <tr> <td>...</td> <td>...</td> <td>001</td> <td>日時 A⇒B⇒C</td> </tr> <tr> <td></td> <td></td> <td>002</td> <td>日時 A⇒B⇒C</td> </tr> <tr> <td></td> <td></td> <td>002</td> <td>日時 A⇒X⇒C</td> </tr> <tr> <td></td> <td></td> <td>...</td> <td>...</td> </tr> </tbody> </table>				仮ID	属性	仮ID	移動履歴	001	20代・男性・千代田区在住	001	日時 A⇒B⇒C	002	20代・男性・千代田区在住	001	日時 A⇒X⇒C	...	...	001	日時 A⇒B⇒C			002	日時 A⇒B⇒C			002	日時 A⇒X⇒C			...	...
仮ID	属性	仮ID	移動履歴																												
001	20代・男性・千代田区在住	001	日時 A⇒B⇒C																												
002	20代・男性・千代田区在住	001	日時 A⇒X⇒C																												
...	...	001	日時 A⇒B⇒C																												
		002	日時 A⇒B⇒C																												
		002	日時 A⇒X⇒C																												
		...	...																												
<p>識別・特定がされない範囲で、 利用者ニーズに合った データセットとなっている</p>																															

○論点①に係る取扱い上の規律に関する整理

制度面、技術面（加工時）について、取扱い上の規律を整理した（図表 39）。規律においても、「十分な匿名化」をした情報と匿名加工情報には大きな差異が認められる。

図表 40 取扱い上の規律

	「十分な匿名化」をした情報	匿名加工情報匿名加工情報
制度面	<ul style="list-style-type: none"> <li>○ 有効な同意が必要である。</li> <li>○ 以下の全ての要件を満たせば、事前の包括同意で活用できる。</li> <li>・ 加工の手法・管理運用体制の適切性</li> <li>・ オプトアウトの提供</li> <li>・ PIA の実施及び利用者への相応の周知</li> </ul>	<ul style="list-style-type: none"> <li>○ 以下の要件などを満たせば、同意なしで活用できる。</li> <li>・ 作成時の公表</li> <li>・ 第三者提供時の明示</li> <li>・ 識別行為の禁止</li> <li>・ 安全管理措置</li> </ul>
技術面 （加工時）	<ul style="list-style-type: none"> <li>・ 再特定化・再識別化が不可能又は極めて困難と言える程度の加工</li> <li>・ 加工対象データの制限（入口要件、出口要件）</li> </ul>	<ul style="list-style-type: none"> <li>・ 加工基準に従った加工</li> </ul>

○論点①に係る整理結果

制度面における整理、技術面における整理、取扱い上の規律に関する整理を踏まえ、通信の秘密に該当する位置情報を、有効な同意なく匿名加工情報に加工して取り扱うことは、通常認められない、として整理した。

図表 41 論点①に係る整理結果

論点	<p><b>論点①（通秘×匿名加工情報）</b></p> <ul style="list-style-type: none"> <li>・ そもそも通秘に該当する位置情報を匿名加工情報として利用することは可能か？</li> </ul>
整理結果	<ul style="list-style-type: none"> <li>・ 通秘に該当する位置情報を、有効な同意なく匿名加工情報に加工して取り扱うことは、通常認められない。</li> </ul>
理由	<ul style="list-style-type: none"> <li>・ 制度面：通秘の利用に当たり、本人関与の仕組みとして、原則として有効な同意を必要としている。</li> <li>・ 技術面：通信の秘密保護の趣旨から、個人だけで無く個別の通信を特定されないレベルでの加工が求められるため、事実上、個人情報保護を目的とする匿名加工情報よりも厳しい匿名化を必要としている。</li> </ul>



「② 非通秘の位置情報に対して、「十分な匿名化」の水準を求めることの当否」について、制度面、技術面で整理を行って検証した結果を次に示す。

○制度面における整理

改正個人情報保護法では、個人情報に紐づく情報の場合、匿名加工情報についての加工基準に従えばよいとされている。

一方で、電気通信事業における個人情報保護に関するガイドライン（※改正後の2017年5月30日施行予定）では、通信の秘密に該当しない位置情報についても、「他人への提供その他の利用においては、利用者の同意を得る場合又は違法性阻却事由がある場合に限定することが強く求められる」とされている。背景として、電気通信事業者は、契約情報等を保有していることが通常であり（図表42）、プライバシー保護の必要性が高いことが理由として挙げられる。

図表 42 論点②に係る制度面における整理

	「十分な匿名化」をした情報	匿名加工情報匿名加工情報
対象とする情報	<ul style="list-style-type: none"> <li>電気通信事業者が取り扱う位置情報（通信の秘密に該当しないもの）</li> </ul>	<ul style="list-style-type: none"> <li>個人情報にひもづく位置情報</li> </ul>
法的根拠	<ul style="list-style-type: none"> <li>プライバシーの保護</li> </ul>	<ul style="list-style-type: none"> <li>改正個人情報保護法（2017年5月30日施行予定）</li> </ul>
関連文書	<ul style="list-style-type: none"> <li>位置情報プライバシーレポート</li> <li>電気通信事業における個人情報保護に関するガイドライン（※2017年5月30日施行予定）</li> </ul>	<ul style="list-style-type: none"> <li>個人情報保護法ガイドライン（匿名加工情報編）</li> <li>電気通信事業における個人情報保護に関するガイドライン（※2017年5月30日施行予定）</li> </ul>
保護の趣旨	<ul style="list-style-type: none"> <li>電気通信事業者は、契約情報等を保有していることが通常であり、プライバシー保護の必要性が高いこと</li> </ul>	<ul style="list-style-type: none"> <li>個人情報の適正な取扱いに関する、個人の権利利益の保護</li> <li>特定個人の識別性が排除される水準までの保護をする。</li> </ul>

図表 43 電気通信事業者における位置情報と契約者情報等との照合可能性

位置情報 を取り扱う電気通 信事業者 の類型	基地局に係る 位置情報		GPS 位置情報	Wi-Fi 位置情報	
	個々の通信の際 に利用される基 地局の位置情報	位置登録情報		端末利用者と アクセスポイ ント設置者間 の通信に基づ く位置情報	端末利用者がアク セスポイントから 外部と通信を行う ことで把握される 位置情報
携帯 キャリア	○ 契約者情報と照 合可能	○ 契約者情報 と照合可能	○ 契約者情報 と照合可能	○ 契約者情報と 照合可能	○ 契約者情報と照合 可能
Wi-Fi 事業者	—	—	—	○ メアド+属性 レベルの情報 の追加入力を 求めることが 多い	○ メアド+属性レベ ルの情報の追加入 力を求めることが 多い
MVNO 事業者 かつ 位置情報 アプリ 事業者	—	—	○ 契約者情報 と照合可能	—	—
位置情報 アプリ 事業者	—	—	○ メアド+属 性レベルの 情報の追加 入力を求め ることが多 い	—	—

### ○技術面における整理

前述したとおり、「十分な匿名化」では、通信の秘密保護の趣旨から、個人だけで無く個別の通信を特定されないレベルでの加工が求められるため、事実上個人情報保護を目的とする匿名加工情報よりも厳しい匿名化を必要としている。

他方、通信の秘密に該当しないGPS情報は、基地局情報より精度の高い位置情報を計測することができ、位置精度の観点ではプライバシー性が高いといえる。

### ○論点②に係る整理結果

制度面における整理、技術面における整理を踏まえ、次の通り整理した。

- 通信の秘密に該当しない（非通秘の）位置情報における匿名化に対して、「十分な匿名化」の水準を求める必然性はない。
- 一方、契約者情報又は高精度の位置情報を利用する場合は、「十分な匿名化」に準じた水準まで加工することは、プライバシー保護の観点からは有効であり、望ましい。

図表 44 論点②に係る整理結果

論点	<b>論点②（非通秘）</b> <ul style="list-style-type: none"><li>• 「十分な匿名化」の水準を求めることの是非<ul style="list-style-type: none"><li>◇ 今後の技術進展によって一層高いプライバシー性を有することになると想定される</li><li>◇ GPS情報は基地局情報より精度が高くプライバシー性が高い</li></ul></li></ul>
整理結果	<ul style="list-style-type: none"><li>• 非通秘の位置情報における匿名化に対して、「十分な匿名化」の水準を求める必然性はない。</li><li>• 一方、契約者情報又は高精度の位置情報を利用する場合は、「十分な匿名化」に準じた水準まで加工することは、プライバシー保護の観点からは有効であり、望ましい。</li></ul>
理由	<ul style="list-style-type: none"><li>• 制度面： 「十分な匿名化」をした情報と匿名加工情報とでは、保護の趣旨、取扱い上の規律が異なっており、別の枠組みとして、対象とする情報の通秘及び個人情報への該当性に応じて使い分けるべきで、一律に「十分な匿名化」の水準を求める必然性はない。</li><li>• 技術面： 「十分な匿名化」の水準は、事実上匿名加工情報における加工基準よりも全般的に厳しいものである。</li></ul>

## 4. 利用者の受容性調査

### 1) 調査概要

#### (1) 調査目的

本調査研究において検証したユースケースに基づき、「十分な匿名化」と匿名加工情報それぞれの加工方法について具体例を示した上で、以下の事項を明らかにする。

- ① 「十分な匿名化」と匿名加工情報とで、一般利用者の受容性に差異が生じるか
- ② 一般利用者の受容性向上、不安軽減に資する対処の方策は何か

#### (2) 調査方法

会場調査（CLT：Central Location Test）とデプスインタビューを併用した。

匿名化の加工方法は一般利用者には理解が容易ではなく、適切な回答を得るためには、内容を十分に理解してもらう必要がある。そのため調査対象者を設定した会場に集め、対面で説明を行うことで、背景情報や目的を詳細に伝える調査手法である CLT を採用した。デプスインタビューでは CLT における各設問に対する回答理由や利用条件を変更した場合の意向の変化について、深堀を行った。

なお、一般利用者の受容性評価には影響が薄いと考えられることから、質問にあたって、通信の秘密に該当するか否かは特に明示していない。

#### (3) 調査対象

アンケート調査会社のモニター登録者（全国約 130 万人）から、一都三県<sup>6</sup>に居住する 20 歳～69 歳のスマートフォン利用者<sup>7</sup>を CLT 対象者 100 人、デプスインタビュー対象者 10 人ずつそれぞれ抽出した。

なお、抽出にあたっては凡そ以下の割付になるよう調整を行った。

#### <CLT>

- ・ 性年代（10 歳刻み）均等割付
- ・ 位置情報の利用に関して許容度を測る事前質問を行い、CLT 対象者全体の回答結果が、過去総務省が実施した一般調査<sup>8</sup>における同様に質問に対する回答結果に近い（図表 1、図表 2 を参照）

---

<sup>6</sup> 東京都、埼玉県、千葉県、神奈川県

<sup>7</sup> 本人及び同居の家族が以下の関連業種勤務者に該当する者を除く  
インターネットサービス関連、マスコミ（テレビ・ラジオ局、新聞・雑誌社、広告代理店）、コンサルティング・市場調査関連

<sup>8</sup> 位置情報プライバシーレポート（2014 年 7 月）における調査

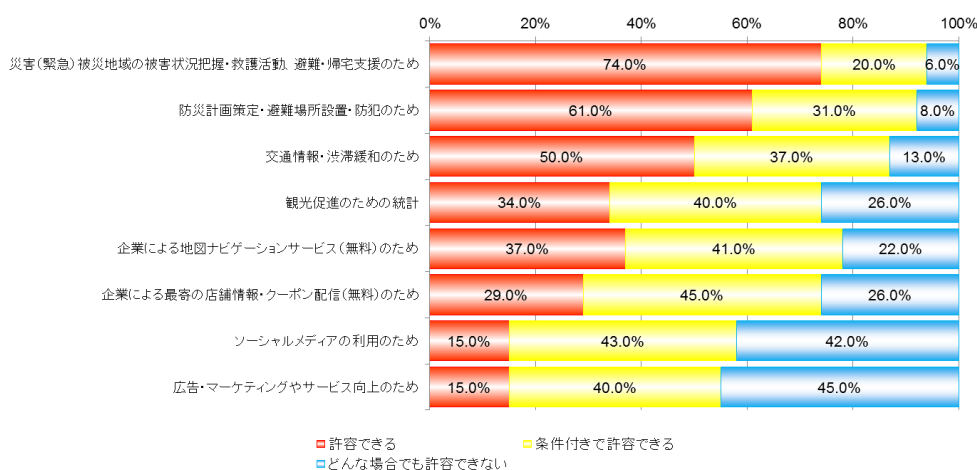
<デプスインタビュー>

- ・ CLT における「十分な匿名化」と匿名加工情報について加工方法の受容性を確認する設問の回答結果が、デプスインタビュー対象者全体で分散する

○位置情報の利用に関して許容度を測る事前質問

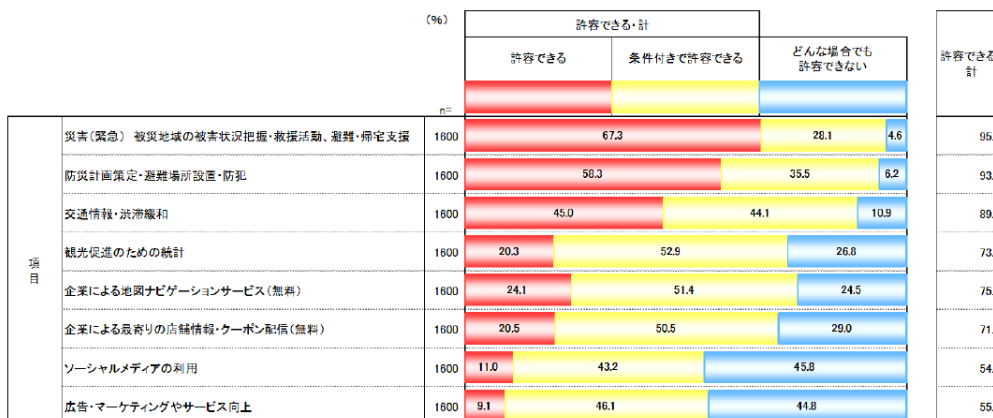
あなたのスマートフォンなど携帯端末の位置情報を、企業や団体が次のような目的に利用することについて、あなたの考えに最も近い（許容できる/条件付きで許容できる/どんな場合でも許容できない）のはどれですか

図表 45 CLT 対象者 (n=100) の回答分布



図表 46 総務省位置情報プライバシーレポート (2014年7月) における同様の質問に対する回答分布

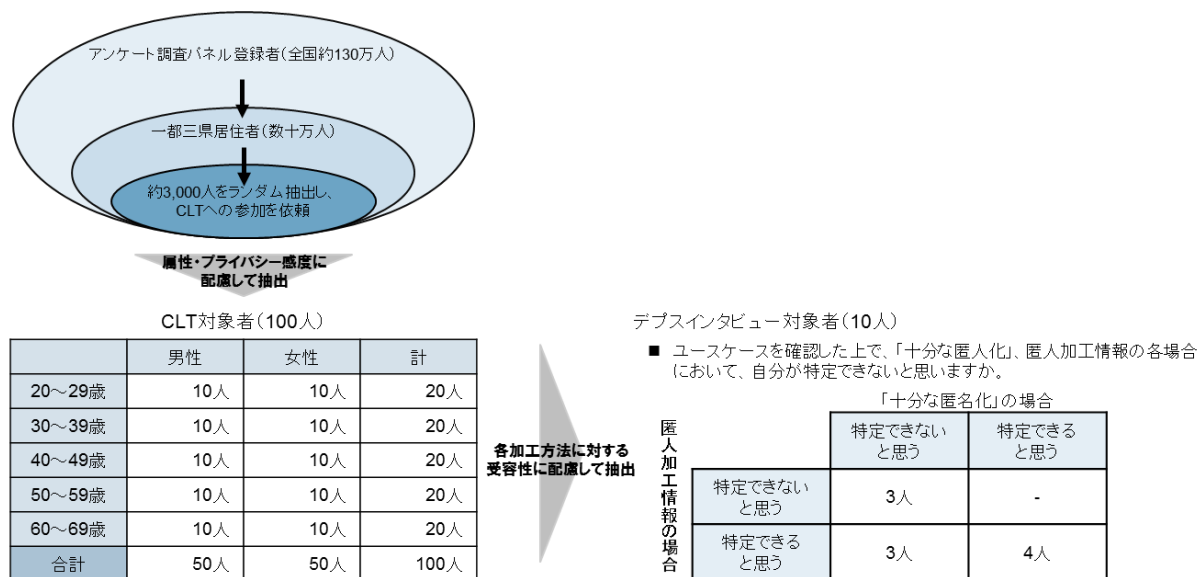
Q あなたの位置情報を次のような目的に利用することについて、あなたの考えに最も近いのはどれですか。



※n=30未満は参考値のため灰色。

○調査対象抽出の流れ

図表 47 調査対象抽出の流れ



(4) 調査期間

2017年1月～2月

(5) 調査項目

1. 位置情報の提供全般

1. 位置情報の取得に対する認知度
2. 位置情報の設定状況
3. 位置情報の提供に対する認知度
4. 位置情報の提供に対する意向
5. 位置情報とともに取得される情報種類の受容性

2. 加工方法

1. 匿名加工情報にした場合の受容性
  2. 「十分な匿名化」を行った場合の受容性
3. 受容性向上、不安軽減に資する対処の方策
1. 安全管理措置に対する要望
  2. オプトアウトを設定した場合の受容性の変化
  3. 提供範囲を限定した場合の受容性の変化
  4. 利用目的を限定した場合の受容性の変化

## 2) 調査結果

### (1) サマリー

#### 調査目的① 「十分な匿名化」と匿名加工情報とで、一般利用者の受容性に差異が生じるか

「十分な匿名化」における集計処理は、一般利用者が匿名加工情報と比べて個人の特定につながりにくくなると感じる要因になっている。契約者情報と紐付けることで加工する項目が多くなるケース2の事例において、特に違いが生じた（図表3を参照）。

「十分な匿名化」、匿名加工情報いずれの加工においても、加工後のデータセットから個人を特定できると思うと回答した一般利用者の6割が居住地（市区町村単位）を最も個人の特定につながりやすい項目と考えており、居住地同様、自宅付近の位置情報は慎重な取り扱いが求められる。

「十分な匿名化」におけるk値に対する感じ方は様々で、k=2 あれば心配に思わないとする意見がある一方で、k=10以上を求める声も聞かれた。

#### 調査目的②一般利用者の受容性向上、不安軽減に資する対処の方策は何か

オプトアウトは利用の有無に関わらず9割、提供先の限定は7割、第三者提供を不安に感じる一般利用者の不安軽減に資する。

図表 48 各ユースケースにおいて「十分な匿名化」、匿名加工情報それぞれの場合で、自分を特定できると思うか否か問う設問の回答結果（n=100）

ケース1 観光		十分な匿名化	
		特定できない	特定できる
匿名加工	特定できない	43%	10%
	特定できる	12%	35%

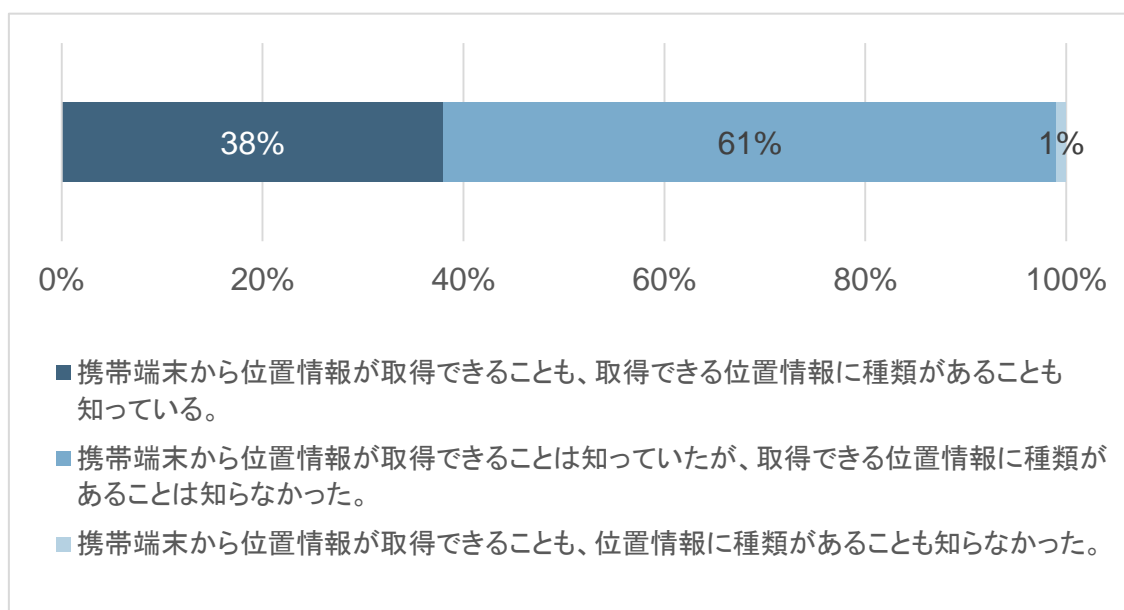
ケース2 商用		十分な匿名化	
		特定できない	特定できる
匿名加工	特定できない	38%	6%
	特定できる	20%	36%

## (2) 個別設問の回答結果

### 1-1. 位置情報の取得に対する認知度

一般に携帯電話、スマートフォン、タブレット端末（以下、携帯端末といいます）からは設定により、その端末がどこにあるかを示す位置情報を取得することができます。位置情報には精度等の違いから、以下①～③の種類があります。

あなたは携帯端末から位置情報が取得できることや、取得できる位置情報に種類があることについて知っていますか。(SA)

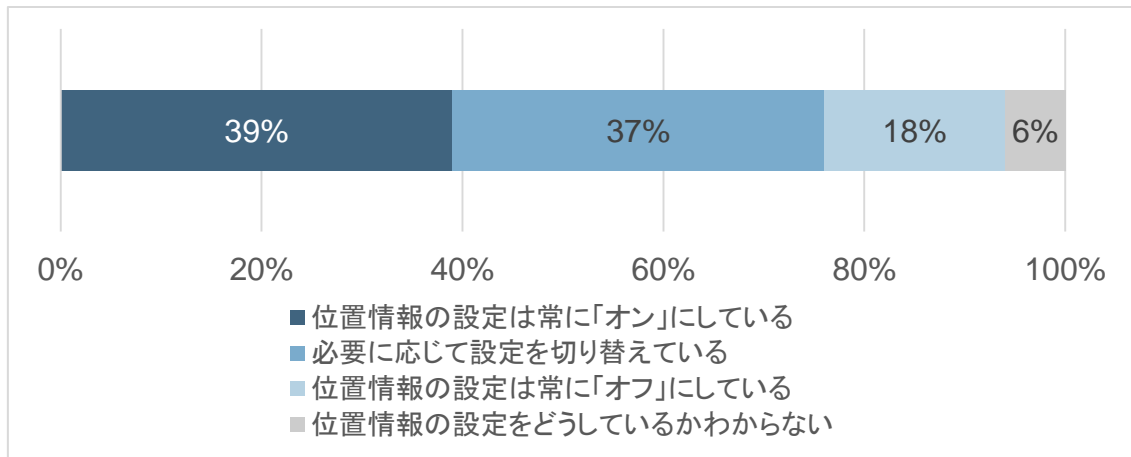




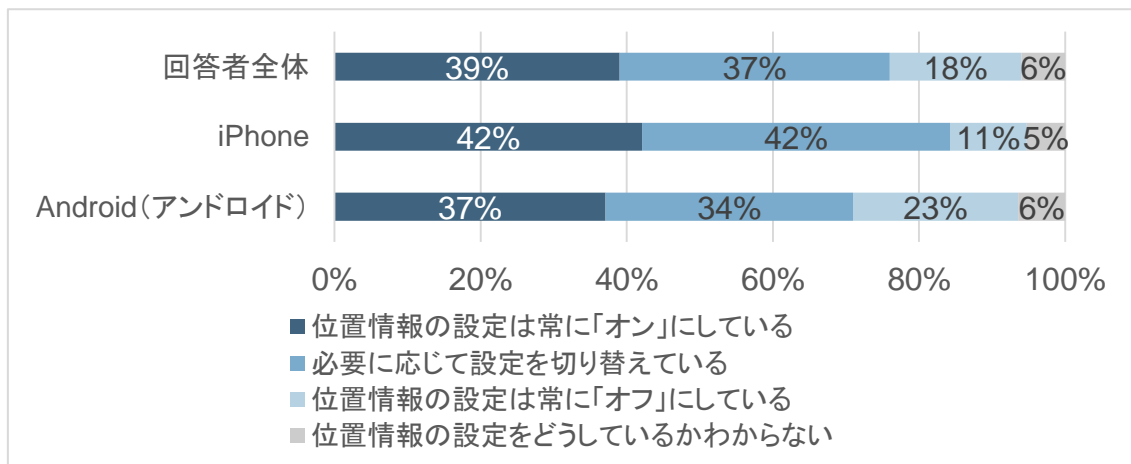
## 1-2. 位置情報の設定状況

携帯端末では、以下のような画面で端末サービスやアプリの利用にあたって位置情報の設定をすることができます。

あなたは普段、ご自分のスマートフォンの位置情報の設定を「オン」にしていますか。(SA)



(参考) 利用するスマートフォンの種類別位置情報の設定状況



### 《デブスインタビューでの発言》

**Q: 位置情報の設定理由。サービス・アプリごとに設定を変えている場合は、設定を「オン」にしているアプリの種類と切り替えの基準。**

➤ **位置情報の設定は常に「オン」にしている**

- Google マップが正確になるという意識があるし、位置サービスのゲームをやっているのでオンにしておいた方が良いと思っている。あとカーナビを使っているのが常にオン。
- 写真を撮る時にアプリでその時の位置情報が入るので便利。地図もよく使うので位置情報はオンにしている。

➤ **必要に応じて設定を切り替えている**

- ナビと地図のアプリを使う時だけオン。基本はオフ。ホットペッパーとかのアプリで「今この近くにおすすぬめがあります。出しますか」とお知らせが出ることもあるが、必要としていないので、そういうものはオフにしている。地図を見るとか特にメリットがあつてその情報を公開するのは良いが、**自分にあまりメリットがないと思うものに対してわざわざ提供する必要がない**と思っている。友人から「怖いよね」という話をすごく聞いたので、どこまでわかっているかわからない時は常にオンはしないでおこうと思った。
- 自分の位置と目的の場所までどのくらいの距離があるかを把握したい時にオンにする。Instagram など写真を投稿する時に自分の位置情報を聞かれるが、Instagram は不特定多数の人が見るので断る。位置情報を知られているのが、ちょっと怖いとか気持ちが悪い。今は画像を撮るだけで位置情報がわかつたりするので、何もしなくてもオンにしているだけで知られているのではないかと思つてしまう。

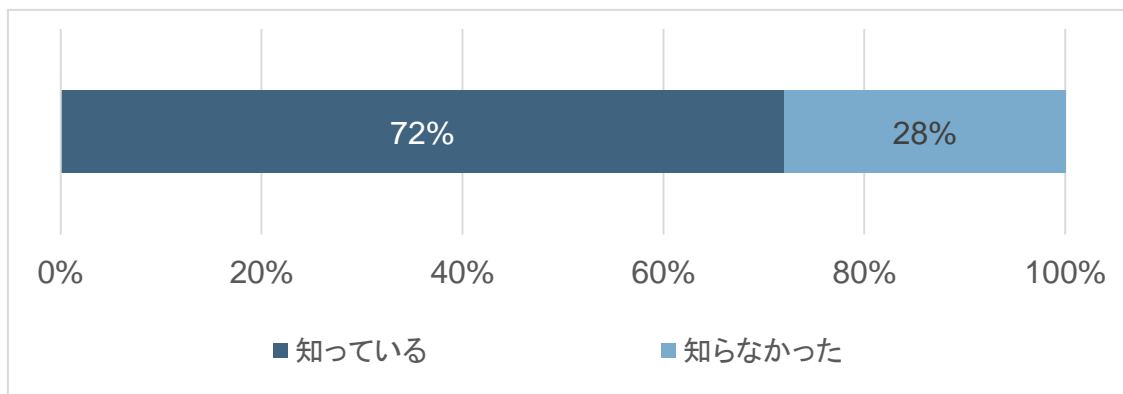
➤ **位置情報の設定は常に「オフ」にしている**

- 機械とか情報とか得意じゃないし、いくら保護されていると言われても嫌。自分の行動を掌握されるのがあまり好きではない。

### 1-3. 位置情報の提供に対する認知度

携帯端末のサービス・アプリを利用するとき位置情報の設定を「オン」にすると、位置情報の提供に同意したとして、あなたの位置情報が端末サービス・アプリの提供者に取得され利用されることがあります。

あなたはこのことを知っていますか。(SA)



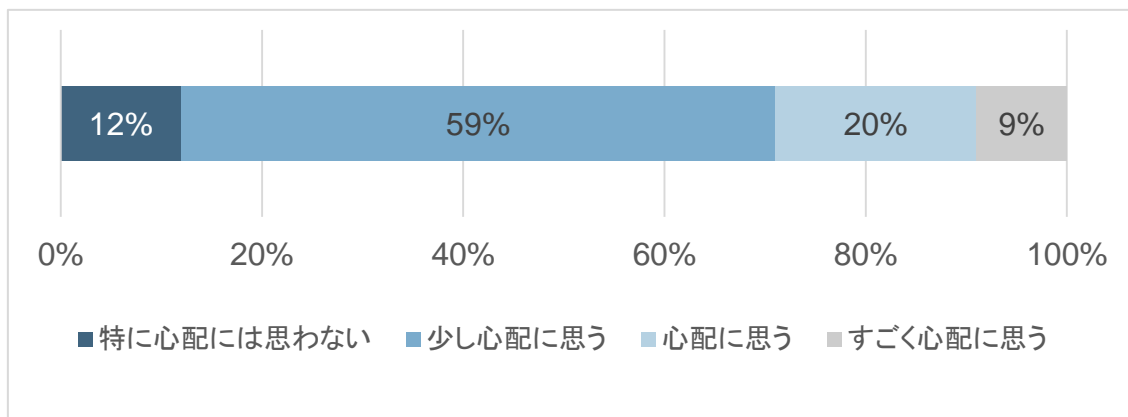
#### 《デブスインタビューでの発言》

##### Q: 位置情報を設定する際に、同意をした認識があるか

- 自分が利用するなら、それなりに相手にも提供しなければいけないというやり取りなんだろうなと思った。
- 許可したという意識はある。
- そのように表示されるから仕方なくした。

#### 1-4. 位置情報の提供に対する意向

あなたは携帯端末を通じて、自分の位置情報が端末サービス・アプリ提供者へ提供されることについて心配に思いますか。(SA)



#### 《デブスインタビューでの発言》

##### Q: 心配に思う理由

###### ➤ 少し心配に思う

- 何がどれくらい提供されるのかわからないので、不安に思う。
- アプリの提供会社でも大手のアプリなら位置情報を提供してもしっかり管理されていると思うが、たまに無料のアプリで大丈夫かなという時にちょっと心配。

###### • 心配に思う

- 特に変なところに行っていないのでバレたらバレてもいいが、自分の生活範囲が全部バレているのかな、と思うと何となく薄気味悪い。家族でも携帯会社に全然バレてもいいが、何となくどこかにバレているのかなと思うと心配になる。
- 情報が漏れたりしたら嫌。会社で個人情報が漏れて謝罪していることがあるが、そういうのを見ると、位置情報もどうしているところに行っているのかわからない。
- 携帯はつながっているものなので、自分の行動が携帯から流れるのではないかと感じてしまう。いくら性別とかわからないようにしても、通信会社には自分のデータが登録されている。そのデータがなければ安心だが、結局つながっているの、自分を特定されて何かに悪いことに使われるのではないかと心配になる。

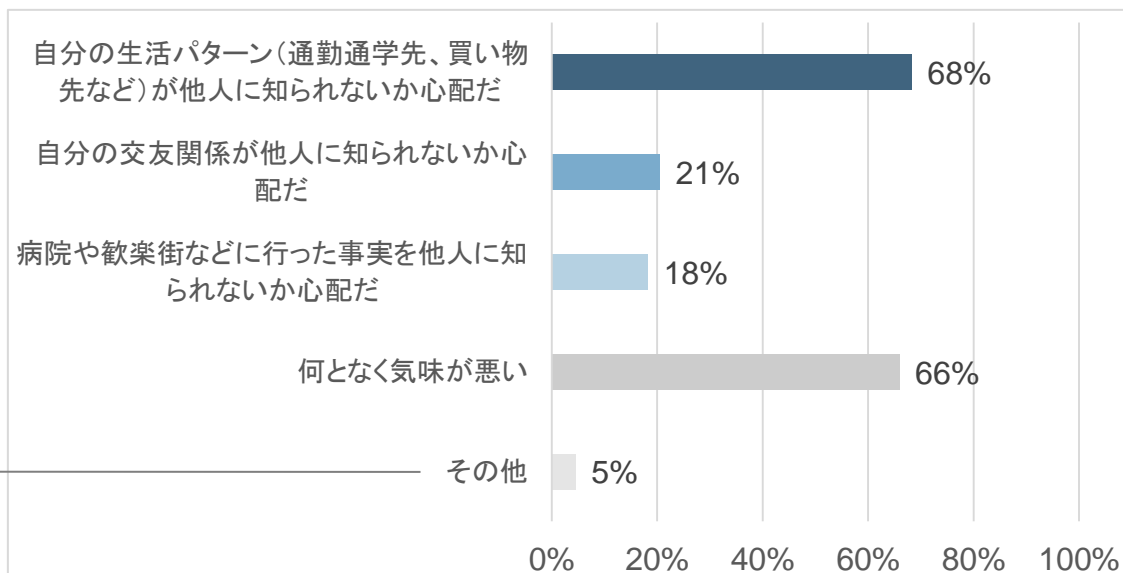
###### • すごく心配に思う

- 「気味が悪い」というのが一番しっくりくる。自分がどこにいるのかを、知らない人に知られるのが嫌。「知らない人」には通信キャリアも含まれる。

#### 1-4. 位置情報の提供に対する意向

前問で 2,3,4 と回答した方にお伺いします。

自分の位置情報が端末サービスやアプリ提供者へ提供されることを心配に思う理由としてあてはまるものをお知らせください。(MA)



#### その他(自由記入内容)

- アプリ提供者に携帯に入っている個人情報が悪用されるかが心配になるから。
- 知らせる必要性が理解できない。

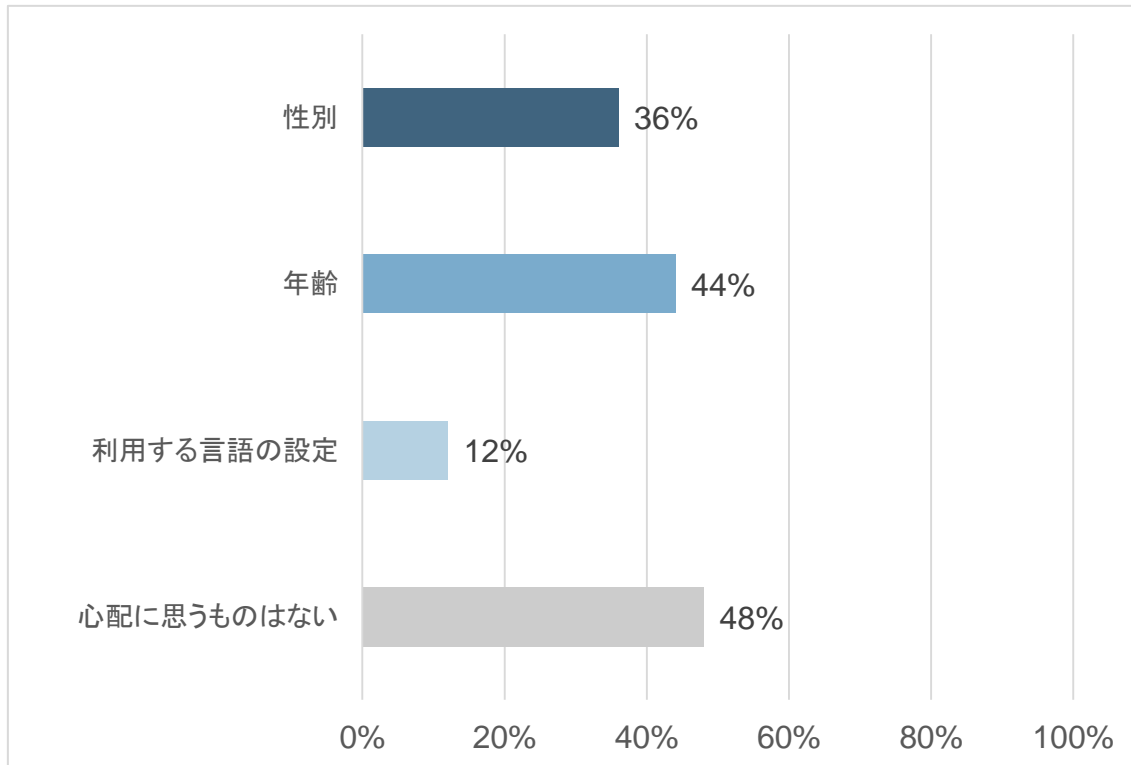
#### 《デブスインタビューでの発言》

##### Q: 心配に思う理由として提示した選択肢(「何となく気味が悪い」以外)の選択理由

- **自分の生活パターン(通勤通学先、買い物先など)が他人に知られないか心配だ**
  - 例えば、自分がいつ自宅にいていつ外に出ているかという情報が第三者に漏れた時は、この時間帯は自宅にいないから空き巣に入られるというような心配があるのではないか。
  - 位置情報を使ったゲームで完全に最寄り駅とか職場の駅とか何があったとか記録しているので、それが漏れると全部が明るみに出てしまい、この人とわかるのが少し怖い。
  - 知らない人がそれを見て、ストーカーのように自分の行動を追われたら嫌
- **自分の交友関係が他人に知られないか心配だ**
  - 位置情報で、頻繁にそこ(友人宅など)に行っているというのがわかると思う。
- **病院や歓楽街などに行った事実を他人に知られないか心配だ**
  - どこかに行ったということが後でどこからか回ってきて、いらぬ誤解をされるのが嫌。

### 1.-5. 位置情報とともに取得される情報種類の受容性

自分の位置情報を端末サービスやアプリ提供者へ提供する際に、以下の情報も付随的に提供することを求められた場合、心配に思うものをお知らせください。(MA)



#### 《デプスインタビューでの発言》

##### Q:項目ごとの心配に思う／思わない理由

###### ➤ 性別、年齢

- 性別は別にいいと思うが、年齢だと人物像が絞り込まれるような気がした。性別より年齢の方が、どれくらいの人かということで自分が特定されるような気がした。
- 性別は男女しかいないので知られてもそれほどでもないと思うが、年齢だとある程度特定されるのではないかと思う。
- 年齢は何十代でも割合細かいと思うので、3段階くらいなら大丈夫。

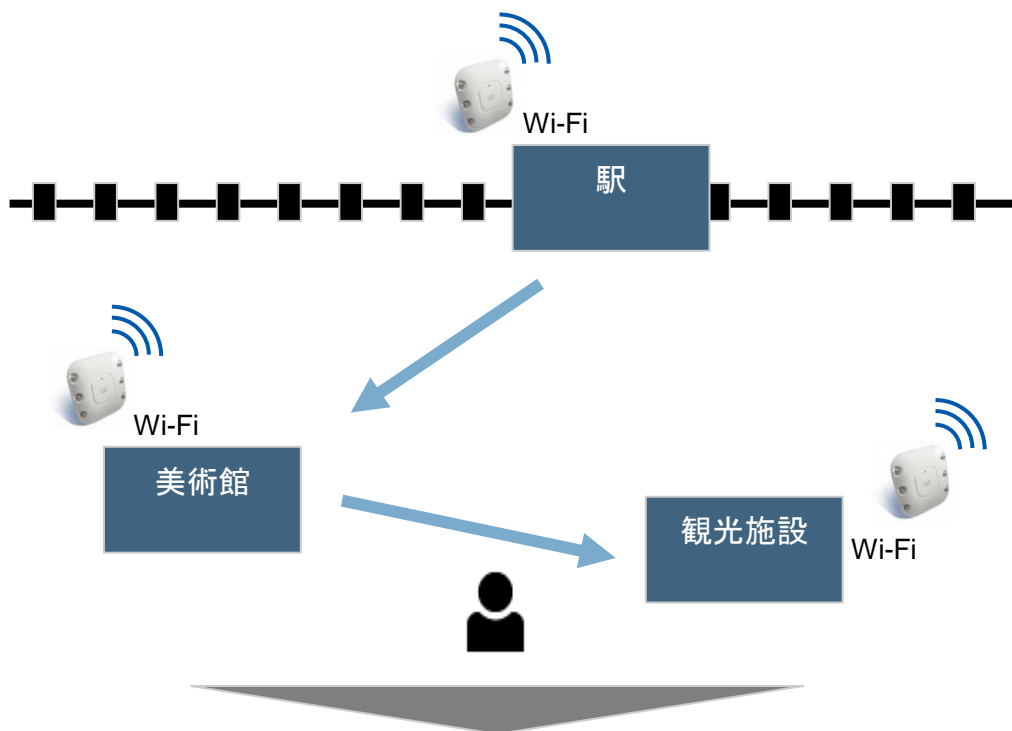
###### ➤ 利用する言語の設定

- 言語はさすがに日本人なので日本にいたら気にしない。
- 言語は、ほとんど関係ないと思う。
- 言語は国内ならあまり心配はない。
- 自分は田舎出身なので、渋谷みたいに外国人がたくさんいる中では、より特定されるような気がしてしまう。

2. 加工方法（ケース1の説明）

A市では観光施設等に設置している無料Wi-Fiサービスを利用した観光客の位置情報を通信会社が取得します。

【図】観光客の移動イメージ



【表】通信会社が取得する位置情報

携帯端末のID	日	時間	Wi-Fi施設	サービス利用言語
aaaaaaaaa	2017/2/11	11時24分38秒	駅	英語
aaaaaaaaa	2017/2/11	12時40分01秒	美術館	英語
aaaaaaaaa	2017/2/11	14時09分10秒	観光施設	英語
bbbbbbbbb	2017/2/12	9時31分50秒	美術館	中国語
cccccccc	2017/2/13	11時9分07秒	美術館	中国語
cccccccc	2017/2/13	12時30分36秒	駅	中国語
dddddddd	2017/2/15	16時37分51秒	観光施設	日本語
dddddddd	2017/2/15	18時22分22秒	美術館	日本語
dddddddd	2017/2/15	19時31分18秒	駅	日本語
eeeeeeee	2017/2/18	20時40分01秒	駅	中国語
eeeeeeee	2017/2/18	21時00分22秒	観光施設	中国語
...	...	...	...	...

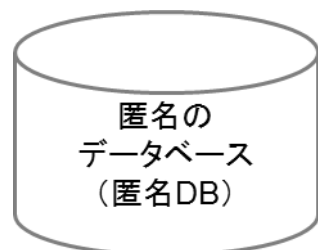
中国人観光客向けのおすすめ観光コースを作成したいと、地元の旅行代理店から要望を受けた通信会社は、取得した位置情報を個人が特定されないよう加工して、提供します。

- 以降の設問は、ご自身を A 市と訪れた観光客に置き換えたと仮定してお答えください。

**【表】通信会社が取得する位置情報**

携帯端末のID	日	時間	Wi-Fi施設	サービス利用言語
aaaaaaaaa	2017/2/11	11時24分38秒	駅	英語
aaaaaaaaa	2017/2/11	12時40分01秒	美術館	英語
aaaaaaaaa	2017/2/11	14時09分10秒	観光施設	英語
bbbbbbbbb	2017/2/12	9時31分50秒	美術館	中国語
cccccccc	2017/2/13	11時9分07秒	美術館	中国語
cccccccc	2017/2/13	12時30分36秒	駅	中国語
dddddddd	2017/2/15	16時37分51秒	観光施設	日本語
dddddddd	2017/2/15	18時22分22秒	美術館	日本語
dddddddd	2017/2/15	19時31分18秒	駅	日本語
eeeeeeee	2017/2/18	20時40分01秒	駅	中国語
eeeeeeee	2017/2/18	21時00分22秒	観光施設	中国語
...	...	...	...	...

個人が特定されないよう加工



このとき、加工の方法には、  
 ・加工パターン①  
 ・加工パターン②  
 のいずれかを用います。



加工パターン①（匿名加工情報）の詳細

《加工手順》

1. 個人を直接特定する情報（携帯端末の ID など）は削除又は仮 ID に変換する。
2. 移動した日付は「平日・休日」に集約する
3. 移動した時間は 15 分単位に集約する
4. 週 5 日以上移動情報が取得される人（A 市に住んでいる人、通勤・通学先が A 市にある人を想定）は除く

加工前

携帯端末のID	日	時間	Wi-Fi施設	サービス利用言語
aaaaaaaaa	2017/2/11	11時24分38秒	駅	英語
aaaaaaaaa	2017/2/11	12時40分01秒	美術館	英語
aaaaaaaaa	2017/2/11	14時09分10秒	観光施設	英語
bbbbbbbbb	2017/2/12	9時31分50秒	美術館	中国語
cccccccc	2017/2/13	11時9分07秒	美術館	中国語
cccccccc	2017/2/13	12時30分36秒	駅	中国語
dddddddd	2017/2/15	16時37分51秒	観光施設	日本語
dddddddd	2017/2/15	18時22分22秒	美術館	日本語
dddddddd	2017/2/15	19時31分18秒	駅	日本語
eeeeeeee	2017/2/18	20時40分01秒	駅	中国語
eeeeeeee	2017/2/18	21時00分22秒	観光施設	中国語
...	...	...	...	...

仮IDに変換

集約

集約

そのまま

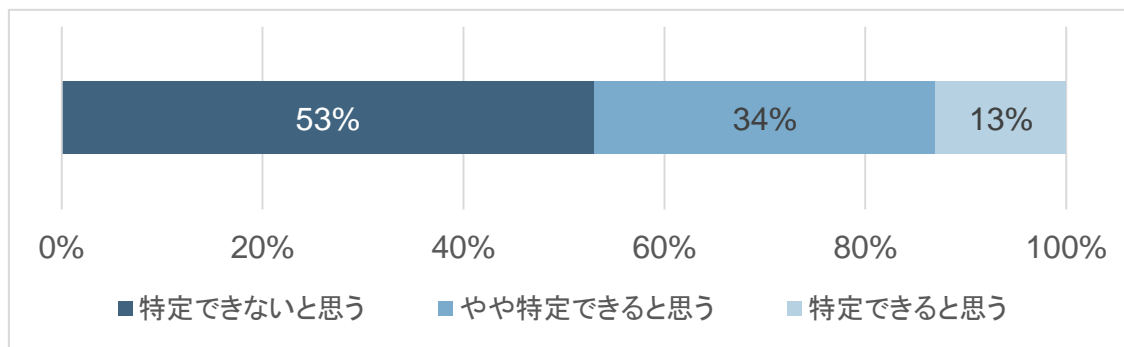
そのまま

加工後  
（匿名DB）

仮ID	日	時間	Wi-Fi施設	サービス利用言語
XXX11XX1	休日	11時30分	駅	英語
XXX11XX1	休日	12時45分	美術館	英語
XXX11XX1	休日	14時15分	観光施設	英語
1X1111XX	休日	9時45分	美術館	中国語
1XXXXXX11	平日	11時15分	美術館	中国語
1XXXXXX11	平日	12時45分	駅	中国語
1X1X1X1X	平日	16時45分	観光施設	日本語
1X1X1X1X	平日	18時30分	美術館	日本語
1X1X1X1X	平日	19時45分	駅	日本語
XXXX111XX	休日	20時45分	駅	中国語
XXXX111XX	休日	21時15分	観光施設	中国語
...	...	...	...	...

## 2-1. 加工パターン①（匿名加工情報）の受容性

加工パターン①のように加工されれば、自分を特定できないと思いますか。（SA）



## 《デブスインタビューでの発言》

### Q: 自分を特定できる／できないと思う理由

#### ➤ 特定できない

##### (十分な匿名化の場合も特定できない)

- 端末 ID が変えられている。日にちが平日か休日の 2 つで、月のどの日かわからない。時間も細かくなくて、15 分単位なのでわからない。誰も行かないところなら特定できるかもしれないが、普通は、これでは個人はわからないと思った。
- 端末 ID。これさえわからなければ、ほぼ特定はできないと思う。時間は結構明確なのでちょっと気になるが、端末 ID が変換されていれば問題ないと思う。

#### ➤ 特定できる

##### (十分な匿名化の場合も特定できない)

- どの時間帯に出かけてどの施設に行くかを、仮 ID になっていても「ある人はこのパターンだ」というまとまりで作られているので、特定できると思った。
- 時間も 15 分単位だとある程度小刻みだと思うので、そのくらいじゃ特定される気がする。午前中とか午後なら心配がないように思う。日にちも平日・休日だと、働いているとか働いていないとかで特定できると思う。

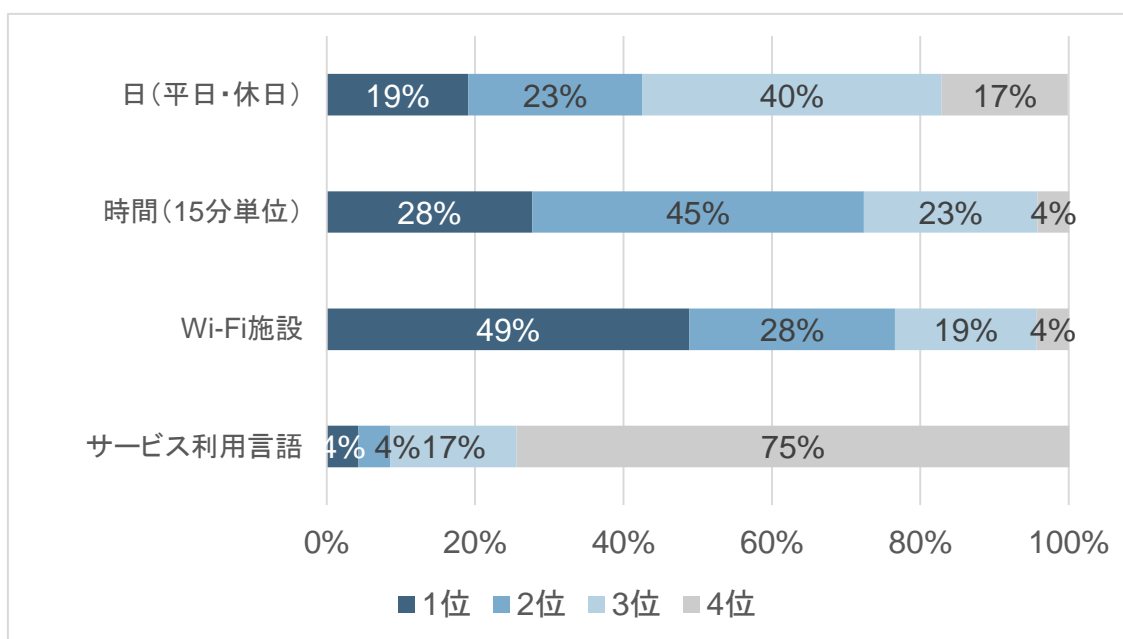
##### (十分な匿名化の場合も特定できる)

- 観光地といっても、渋谷のど真ん中なら人が何千人といるので大丈夫だが、秘湯めぐりなどはどうかな。ちょっとひなびているところだとあのバス停で待っていた誰かというのはわかってしまうだろう。Wi-Fi 施設があつてそこにいたりそこを通ったりして位置情報がつながって、この順で回ったことがわかれば、あそこでお土産を買っていたあの人だとかわかってしまうのかなと思った。そうすると、そこのリピーターになりたくないなと思ってしまう。
- 通信会社には契約者の個人情報が行っていると思うので、探っていけば加工されていても、パソコンに詳しい人なら特定できるのではないかな。
- 加工手順 1 の「加工する」というのが信用できない。それ以降はあまり心配していないが、次の例を含めて加工手順 1 が信用できない
- いずれ特定されると思う。時間がかかると思うが、会社の情報を調べようと思ってそこに行き、そこから調べようと思ったら、一旦加工してあっても元の情報にたどり着くことができると思う。

## 2-1. 加工パターン①（匿名加工情報）の受容性

前問で2,3と回答した方にお伺いします。

どの情報が自分を特定することにつながると感じますか。自分の特定につながりやすいと感じる項目に順位を記載してください。（SA）



## 《デブスインタビューでの発言》

### Q:自分を特定することにつながりやすいと感じた項目の理由

- 時間→Wi-Fi 施設→日→サービス利用言語：特に自分の特定につながりやすいのは時間と Wi-Fi 施設。この時間にここにいたというのは、その人の趣味や行動パターンが特定されるような気がした。

#### ➤ 1位～4位のレベル感

- 時間→日→Wi-Fi 施設→サービス利用言語：4位は離れるが、1、2、3位は僅差。「このくらいの時間にここにいた人」ということで特定されるかなと思った。
- Wi-Fi 施設→時間→日→サービス利用言語：ダントツ Wi-Fi 施設が1位。2位3位がちょっと下で、言語が相当下。
- Wi-Fi 施設→日→時間→サービス利用言語：1位、2位、3位は同じで、4位が離れている。

### Q:外部の情報と組合せた場合、どのような情報だと自分の特定につながると感じるか

#### ➤ コンビニ等での購買記録

- サービスなり商品の提供を自分が受けた時点で、行動も取られているし年齢も住所も分かってしまうのはある程度しょうがないと思うが、それを加工されているなら特定されるとは思わない。
- 例えば美術館でお茶を買ったなどの情報が組み合わせさっても、恐らくわからないと思う。

#### ➤ 防犯カメラの映像記録

- そこまでいくと、ほとんどの情報があるので特定は固まると思う。

加工パターン②（十分な匿名化）の詳細

《加工手順》

1. 個人を直接特定する情報（携帯端末のIDなど）は削除又は仮IDに変換する。
2. 移動した日付は「平日・休日」に集約する
3. ID、日、時間を基準にWi-Fi施設をつなげて移動経路の情報を作成する
4. 週5日以上移動情報が取得される人（A市に住んでいる人、通勤・通学先がA市にある人を想定）は除く
5. 日、移動経路、サービス利用言語の観点から該当する仮IDの数を集計する
6. 集計した結果、1名しか存在しないデータを削除する

加工前

携帯端末のID	日	時間	Wi-Fi施設	サービス利用言語
aaaaaaaa	2017/2/11	11時24分38秒	駅	英語
aaaaaaaa	2017/2/11	12時40分01秒	美術館	英語
aaaaaaaa	2017/2/11	14時09分10秒	観光施設	英語
bbbbbbbb	2017/2/12	9時31分50秒	美術館	中国語
cccccccc	2017/2/13	11時9分07秒	美術館	中国語
cccccccc	2017/2/13	12時30分36秒	駅	中国語
dddddddd	2017/2/15	16時37分51秒	観光施設	日本語
dddddddd	2017/2/15	18時22分22秒	美術館	日本語
dddddddd	2017/2/15	19時31分18秒	駅	日本語
eeeeeeee	2017/2/18	20時40分01秒	駅	中国語
eeeeeeee	2017/2/18	21時00分22秒	観光施設	中国語
...	...	...	...	...

加工手順  
1.~4.

仮IDに変換
集約
ID、日、時間を基準に移動経路化
そのまま

仮ID	日	移動経路	サービス利用言語
XXX11XXX1	休日	駅→美術館→観光施設	英語
1X1111XXX	休日	美術館	英語
1XXXXXX11	平日	美術館→駅	中国語
1X1X1X1XX	平日	観光施設→美術館→駅	日本語
XXXX111XX	休日	駅→観光施設	中国語
XX11X1XX1	平日	美術館→駅	中国語
1XXXXXX111	休日	駅→観光施設	日本語
XX1X1X1XX	平日	観光施設→美術館→駅	日本語
...	...	...	...

加工手順  
5.~6.

集計

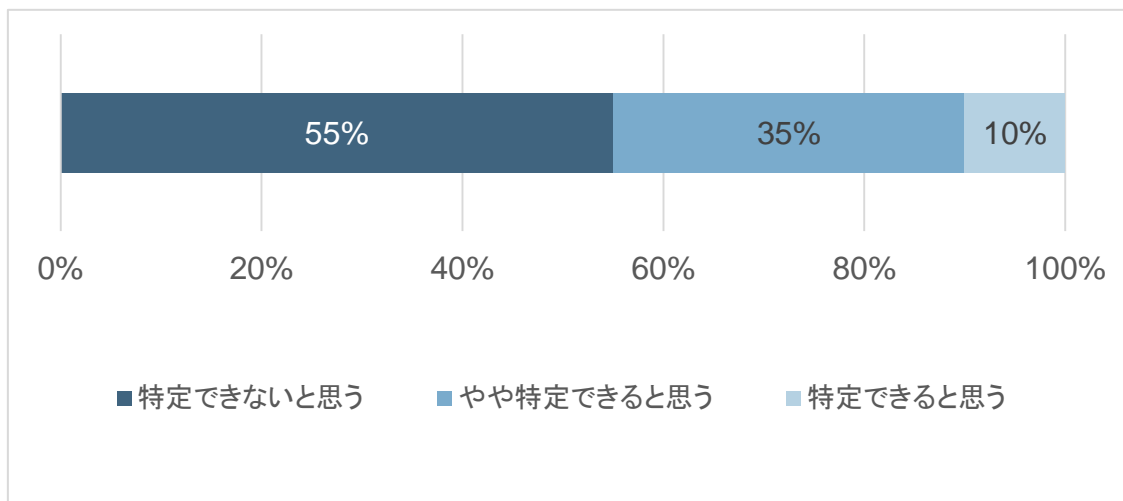
集計する際の観点として利用

該当数	日	移動経路	サービス利用言語
1名	休日	駅→美術館→観光施設	英語
1名	休日	美術館	英語
2名	平日	美術館→駅	中国語
2名	平日	観光施設→美術館→駅	日本語
1名	休日	駅→観光施設	中国語
1名	休日	駅→観光施設	日本語
...	...	...	...

加工後  
(匿名DB)

## 2.2. 加工パターン②（十分な匿名化）の受容性

加工パターン②のように加工されれば、自分が特定できないと思いますか。（SA）



## 《デブスインタビューでの発言》

### Q: 自分を特定できる／できないと思う理由

#### ➤ 特定できない

##### (匿名加工情報の場合も特定できない)

- 加工パターン①も加工パターン②も同じ。ID さえわからなければ、後はどうぞご自由にといい感じ。問題なのはID と個人名と住所。
- 集計前のデータでも、ID も変わって日にちも特定されないの、個人の情報はわからないと思った。更に行動のパターンが1つしかない削除されるということだったので、更に全体のデータのように自分が埋もれると思った。

##### (匿名加工情報の場合は特定できる)

- 1名だと完全にデータ削除とあるので、複数なら特定しないと思った。加工手順の5、6(集計)があるから特定できないと思った。第一段階の手順ではまだ特定できると思う。

#### ➤ 特定できる

##### (十分な匿名化の場合も特定できる)

- 例えば渋谷であれば1,000人、2,000人が闊歩しているので、そうしたらもう特定できないと思うが、2名は何人中の2名かによる。(特定できなくなると思う最低人数は何人か) 10人。
- 109にいたとか大きな施設にいて探すのは無理かもしれないが、2名3名しか人がいないところなら頑張れば特定できそう。(特定できないと思う最低人数は何人か) 1,000人。
- 機械が詳しい人が調べれば、いずれは元のデータにたどり着くと思う
- 加工手順1が信用できないが、手順1以降についてはどれも気になることはない。

### Q: 匿名加工情報の場合と十分な匿名化の場合とで受容性が異なる理由

#### ➤ 匿名加工情報(特定できる)×十分な匿名化(特定できない)

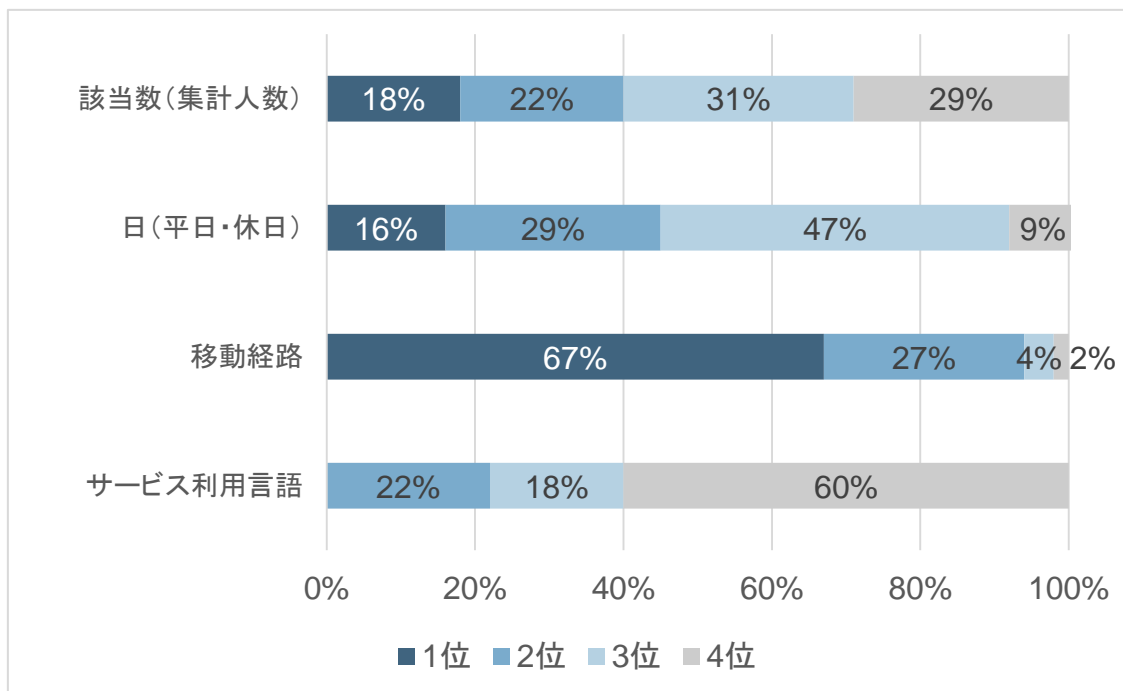
- 2つは大きく違う。2人だと集計されたデータのような気がするので安心。(自分の特定につながる最低人数は何人か) 2名以上なら大丈夫。
- 2つには大きな差がある。加工パターン②の方が特定できなくなったと思う。決め手は、時間がなくなって経路になったこと。最後の集計で人数が多いところだけとなると、より特定できなくなると思うので、断然加工パターン②。(自分の特定につながる最低人数は何人か) 集計された段階で何人でも全然心配はないと思う。
- 集計してしまったら、あまり関係ないと思う。自分の携帯とはつながらない感じがするので、あまり匿名化されないと思う。2名以上であれば大丈夫だと思う。



## 2-2. 加工パターン②（十分な匿名化）の受容性

前問で2,3と回答した方にお伺いします。

どの情報が自分を特定することにつながるとお思いますか。自分の特定につながりやすいと感じる項目に順位を記載してください。（SA）



### 《デブスインタビューでの発言》

#### Q: 自分を特定することにつながりやすいと感じた項目の理由

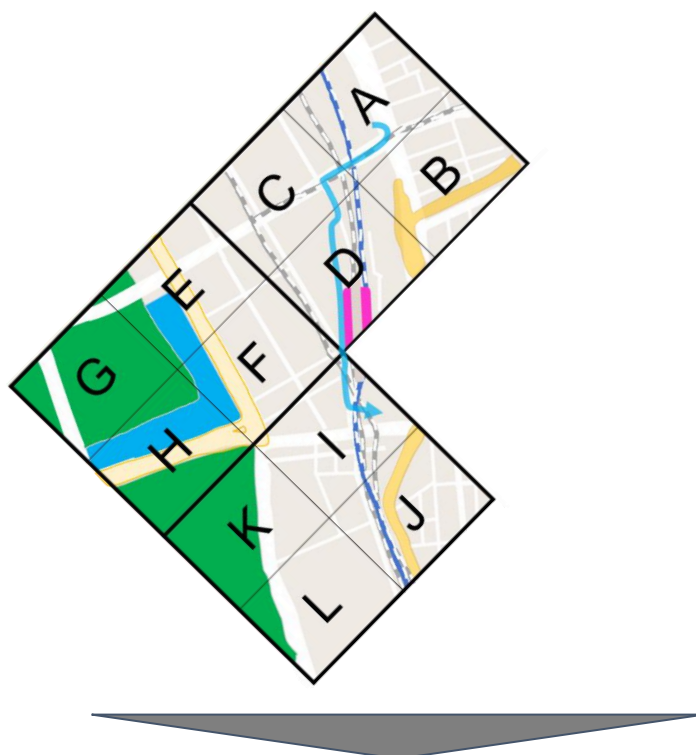
- 移動経路→日→サービス利用言語→該当数：経路は自分が行動した場所なので、ほぼわかってしまう感じがする。何時にどこにいたという情報よりも、こういう風に回っているというのがわかると、自分の出発点とか、目的地とかがわかってしまう気がする。
- 日→移動経路→該当数→サービス利用言語：1位と2位が同じで、3位と4位がちょっと離れている。

2. 加工方法（ケース2の説明）

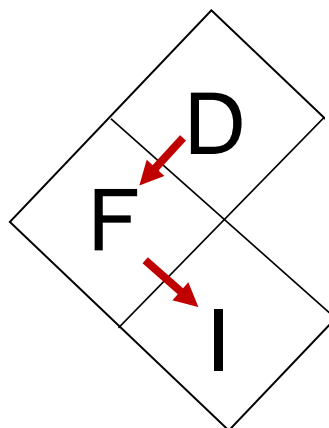
Bさんは2017年2月11日に有楽町マルイで買い物をした後、ジョナサン有楽町店で食事をして、ソニービルへ移動しました。

- 地域をメッシュで区切ったとき、通信会社は携帯端末の位置情報から、Bさんの位置情報をメッシュ単位で取得します。

【図】地図上のBさんの移動経路



【表1】通信会社が取得したBさんの位置情報



携帯端末のID	日	時間	メッシュ番号
xxxxxxxxxx	2017/2/11	11時24分38秒	D
xxxxxxxxxx	2017/2/11	12時40分01秒	F
xxxxxxxxxx	2017/2/11	14時09分10秒	I

通信会社は B さんを含む有楽町近辺を移動した人から取得した位置情報を、携帯端末の ID を元に保有する契約者情報を突合して、【表 4】データベースを作成します。

- 通信会社は携帯端末の ID を用いて、取得した移動情報と既に保有している契約者情報をつなぎ合わせて【表 4】データベースを作成します。

**【表 2】通信会社が取得した B さんを含む有楽町近辺を移動した人の位置情報**

携帯端末のID	日	時間	メッシュ番号
xxxxxxxxxx	2017/2/11	11時24分38秒	D
xxxxxxxxxx	2017/2/11	12時40分01秒	F
xxxxxxxxxx	2017/2/11	14時09分10秒	I
yyyyyyyyyy	2017/2/12	9時31分50秒	H
zzzzzzzzzz	2017/2/13	20時40分01秒	B
zzzzzzzzzz	2017/2/13	21時00分22秒	D
zzzzzzzzzz	2017/2/13	21時29分54秒	F
zzzzzzzzzz	2017/2/13	21時48分18秒	H
...	...	...	...

突合

**【表 3】通信会社が保有している契約者情報**

携帯端末のID	性別	年齢	住所	趣味
xxxxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞
yyyyyyyyyy	男性	42歳	神奈川県川崎市中原区1-2	ゴルフ
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇
...	...	...	...	...

**【表 4】データベース**

携帯端末のID	性別	年齢	住所	趣味	日	時間	メッシュ番号
xxxxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	11時24分38秒	D
xxxxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	12時40分01秒	F
xxxxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	14時09分10秒	I
yyyyyyyyyy	男性	42歳	神奈川県川崎市中原区1-2	ゴルフ	2017/2/12	9時31分50秒	H
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	20時40分01秒	B
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	21時00分22秒	D
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	21時29分54秒	F
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	21時48分18秒	H
...	...	...	...	...	...	...	...

通信会社は【表 4】データベースを個人が特定されないように加工して、匿名のデータベースを作成します。この匿名のデータベースは外部の事業者へ提供されます。

- 「映画のクーポンチケットをどの地域で配ったらより効果的か」という相談を有楽町の映画館から受けた通信会社は、【表 4】データベースを元に個人が特定されないように加工して匿名のデータベース（以下、匿名 DB といいます）を作成し、これを映画館に提供します。

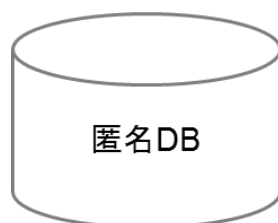
【表 4】データベース

携帯端末のID	性別	年齢	住所	趣味	日	時間	メッシュ番号
xxxxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	11時24分38秒	D
xxxxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	12時40分01秒	F
xxxxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	14時09分10秒	I
yyyyyyyyyy	男性	42歳	神奈川県川崎市中原区1-2	ゴルフ	2017/2/12	9時31分50秒	H
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	20時40分01秒	B
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	21時00分22秒	D
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	21時29分54秒	F
zzzzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	21時48分18秒	H
...	...	...	...	...	...	...	...

個人が特定されないように加工



このとき、加工の方法には、  
 ・加工パターン③  
 ・加工パターン④  
 のいずれかを用います。



匿名 DB の提供を受けた外部の事業者は、データを活用してサービスを検討します。

- 匿名 DB の提供を受けた有楽町の映画館は、映画鑑賞が趣味の男性が千代田区に多いと考え、映画のクーポンチケットを千代田区限定で配ることにしました。
- 以降の設問は、ご自身を B さんに置き換えたと仮定してお答えください。

### 加工パターン③（匿名加工情報）の詳細

《加工手順》

1. 個人を直接特定する情報（氏名、携帯端末の ID など）は削除又は仮 ID に変換する
2. 年齢は 10 歳区切りに集約する
3. 住所は市区町村単位に集約する
4. 移動した日付は「平日・休日」に集約する
5. 移動した時間は 15 分単位に集約する
6. 週 5 日以上移動情報が取得される人（有楽町に住んでいる人、通勤・通学先が有楽町にある人を想定）は除く

### 加工前(データベース)

携帯端末のID	性別	年齢	住所	趣味	日	時間	メッシュ番号
xxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	11時24分38秒	D
xxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	12時40分01秒	F
xxxxxxxx	男性	25歳	東京都千代田区霞ヶ関2-1-2	映画鑑賞	2017/2/11	14時09分10秒	I
yyyyyyyy	男性	42歳	神奈川県川崎市中原区1-2	ゴルフ	2017/2/12	9時31分50秒	H
zzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	20時40分01秒	B
zzzzzzzz	女性	37歳	東京都港区芝公園4-2-8	観劇	2017/2/13	21時00分22秒	D
...	...	...	...	...	...	...	...

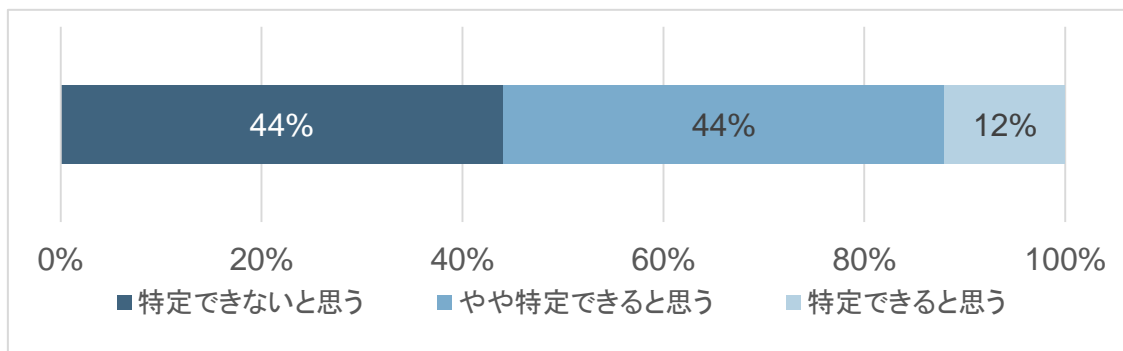
仮 ID に変換    そのまま    集約    集約    そのまま    集約    集約    そのまま

### 加工後(匿名 DB)

仮ID	性別	年齢	住所	趣味	日	時間	メッシュ番号
xxx11xxx1	男性	20代	東京都千代田区	映画鑑賞	休日	11時30分	D
xxx11xxx1	男性	20代	東京都千代田区	映画鑑賞	休日	12時45分	F
xxx11xxx1	男性	20代	東京都千代田区	映画鑑賞	休日	14時15分	I
x1x1x1111	男性	40代	神奈川県川崎市	ゴルフ	休日	9時45分	H
1xxxxxx11	女性	30代	東京都港区	観劇	平日	20時44分	B
1xxxxxx11	女性	30代	東京都港区	観劇	平日	21時00分	D
...	...	...	...	...	...	...	...

### 2-3. 加工パターン③（匿名加工情報）の受容性

加工パターン③のように加工されれば、自分が特定できないと思いますか。（SA）



## 《デブスインタビューでの発言》

### Q:自分を特定できる／できないと思う理由

#### ➤ 特定できない

##### (十分な匿名化の場合も特定できない)

- 携帯の ID が変わっているし、年齢も細かくなく何十代になっている。住所も最後の細かいところが消えている。日にちも平日・休日、時間も 15 分単位になっているので大丈夫だと思った。
- 住所が番地まで記載されていると特定されるかもしれないが、区内・市内までなら特定されないと思った。町名だとちょっと微妙。

#### ➤ 特定できる

##### (十分な匿名化の場合も特定できない)

- 居住地と年代で、特に居住地に市町村まで入っていると特定されると思った。都道府県までなら該当する人がたくさんいると思うが、市町村の何とか区とか何とか市プラス年代が入ると、かなり絞り込まれると思った。
- 居住地とリンクさせるところが一番ドキッとした。通信会社の情報と照らし合わせると、いろいろぼかしてもバレるんじゃないかなと思った。

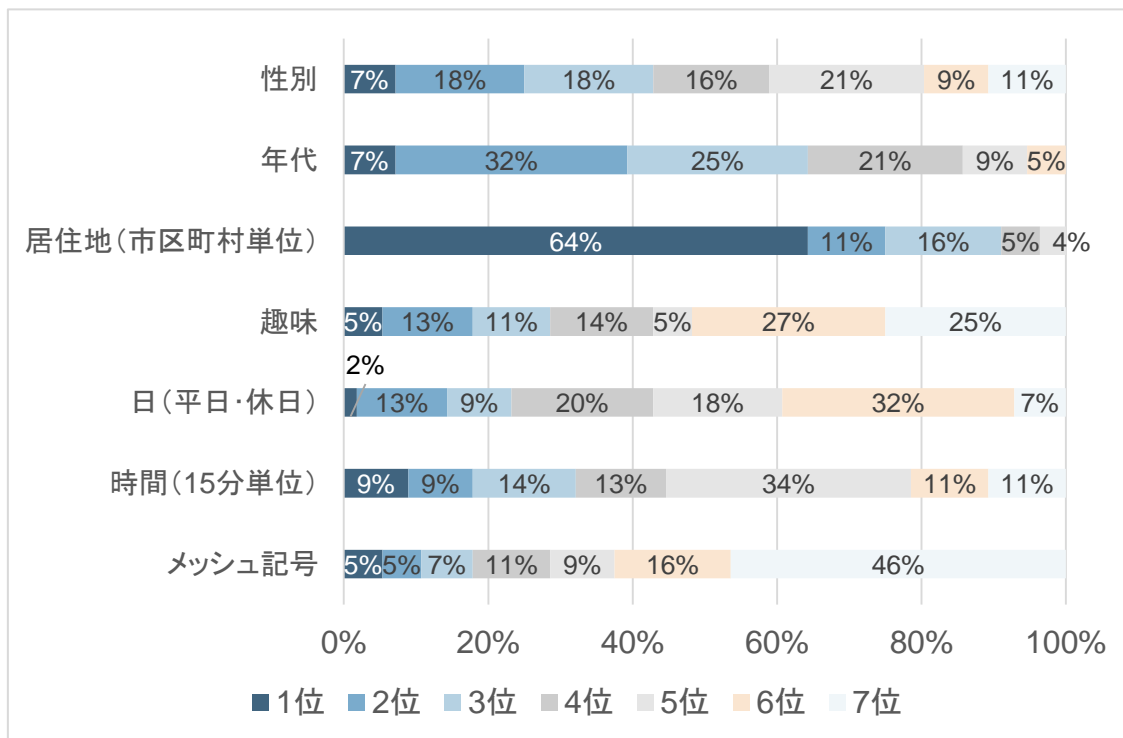
##### (十分な匿名化の場合も特定できる)

- まず居住地はあまり知られたくない。仮 ID も頑張れば加工前の ID にできそう。(居住地が県レベルなら問題ないか) できれば何も知られたくないが、できれば県。仮 ID もあまり表示してほしくない。
- 携帯会社の情報として完全に自分の住所などが行っているの、そこをつなげれば特定できると思った。
- この携帯会社に詳しいデータベースがあり、それが外部に漏れている時点で怖いし、それで特定されてしまうのではないかと思った。どこかで位置情報を取って加工パターン③に加工しても、年代などが詳しくわかったり趣味も入れているので特定できるような気がした。加工パターン②のように地点ではなくメッシュ記号になったので加工パターン③の方が特定できないような気がするが、突き詰めれば皆一緒という気がする。何かうまいことをすれば、特定できてしまうのではないかなという感覚があった

### 2-3. 加工パターン③（匿名加工情報）の受容性

前問で2,3と回答した方にお伺いします。

どの情報が自分を特定することにつながると感じますか。自分の特定につながりやすいと感じる項目に順位を記載してください。（SA）



#### 《デブスインタビューでの発言》

##### Q: 自分を特定することにつながりやすいと感じた項目の理由

- 自分の特定に一番つながりやすいと思う項目は住所。市区町村単位でもそう思う。今ここにいる人で、同じ市区町村の間はそんなにいないだろう。
- 居住地と年代で、特に居住地で市町村まで入っていると特定されると思った。都道府県までなら該当する人がたくさんいると思うが、市町村の何とか区とか何とか町まで入って年代が入ると、かなり絞り込まれると思った。ただし、どちらか一方ならそんなに絞り込まれないと思う。
- （居住地が区市町村より粗ければ特定にはつながりにくいと思うか）そこまでしっかり加工がしてあれば、千葉県の北部とか中部、東京でも23区内とかなら大丈夫。
- 趣味を聞かれてそんなに真剣に答える人もいないと思うので、それで特定にはあまりつながらないと思う。
- （趣味が「恋愛映画」、「邦画鑑賞」、「宝塚鑑賞」など更に細くなると特定につながると思うか）そうは思わない。
- 多趣味の人より限られた趣味の人だと特定されるのではないか。映画鑑賞が趣味の人は多いと思うが、変わったマニアックな趣味の人だと限られてくると思う。



#### 加工パターン④（十分な匿名化）の詳細

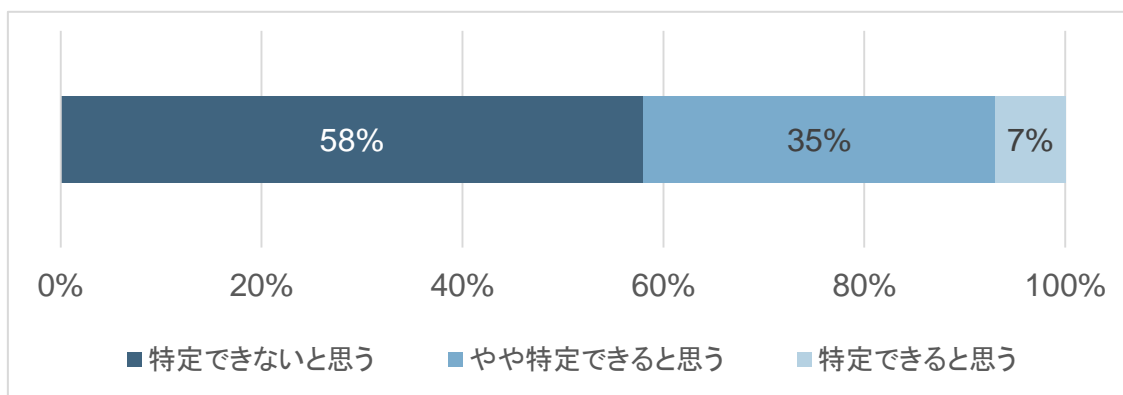
##### 《加工手順》

1. 個人を直接特定する情報（氏名、携帯端末の ID など）は削除又は仮 ID に変換する
2. ID、日、時間を基準にメッシュ番号をつなげて移動経路の情報を作成する
3. 年齢は 10 歳区切りに集約する
4. 住所は市区町村単位に集約する
5. 移動した日付は「平日・休日」に集約する
6. 週 5 日以上移動情報が取得される人（有楽町に住んでいる人、通勤・通学先が有楽町にある人を想定）は除く
7. 属性（性別・年齢・住所）、趣味、日、移動経路の観点から該当する仮 ID の数を集計する
8. 集計した結果、1 名しか存在しないデータを削除する



## 2-4. 加工パターン④（十分な匿名化）の受容性

加工パターン④のように加工されれば、自分が特定できないと思いますか。（SA）



### 《デブスインタビューでの発言》

#### Q: 自分を特定できる／できないと思う理由

##### ➤ 特定できない

###### （十分な匿名化の場合も特定できない）

- ・ 加工パターン③より更に仮 ID が消えて、1 名だったものを全て削除するなど集計をするため、特定できないと思う。
- ・ 2 名以上だったら大丈夫。偶然同じ人がいたら、と思うので、もう少し増やして 10 人いれば安心。
- ・ 本人にはたどり着けないだろう。有楽町にいた人も、どこから来た人なのかはわからない。（その人の住所情報があってもそう思うか）市・区のレベルだとわからない。更に細かい町レベルだとわかってしまうかもしれない。

##### ➤ 特定できる

###### （十分な匿名化の場合も特定できない）

- ・ 加工手順 7、8（集計と K=1 の削除）で複数名になった時点で、特定できなくなるんじゃないかと思う。（自分を特定できないと思う最低人数は何人か）2 名。
- ・ これだとずいぶん加工されているので大丈夫そう。集計していることが大きい。

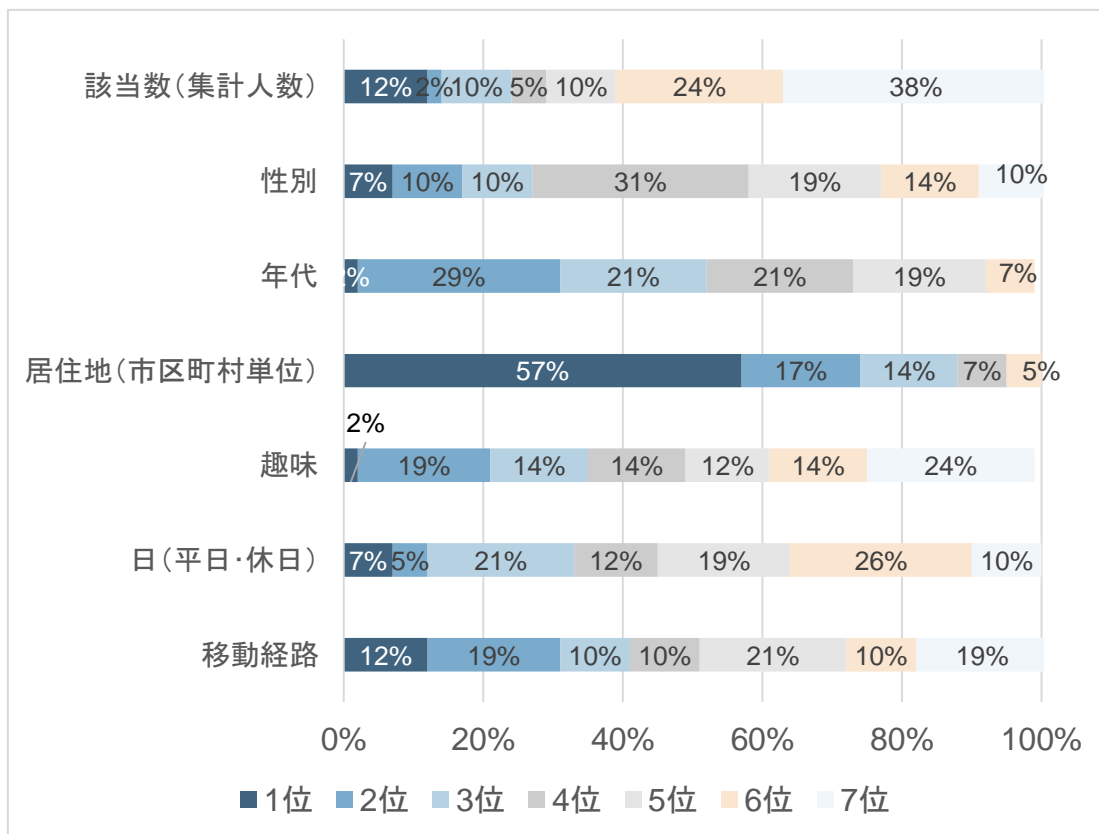
###### （十分な匿名化の場合も特定できる）

- ・ 人数である程度該当者数が絞られて、2 名であればわかりやすくなってしまおう。位置と時間とかある程度の個人情報が結びつけられると見つけられる可能性が高くなると思う。加工パターン③と④とで大差はない。
- ・ 居住地があることで、ほぼ特定できるんじゃないかと思う。それ以外は同じ。加工パターン③も加工パターン④もほとんど変わらないと思う。（集計をしてもあまり変わらないと思うのか）加工パターン④は県単位なら良いかもしれないが、市町村になるとかなり特定されそう。

## 2-4. 加工パターン④（十分な匿名化）の受容性

前問で2,3と回答した方にお伺いします。

どの情報が自分を特定することにつながると感じますか。自分の特定につながりやすいと感じる項目に順位を記載してください。（SA）



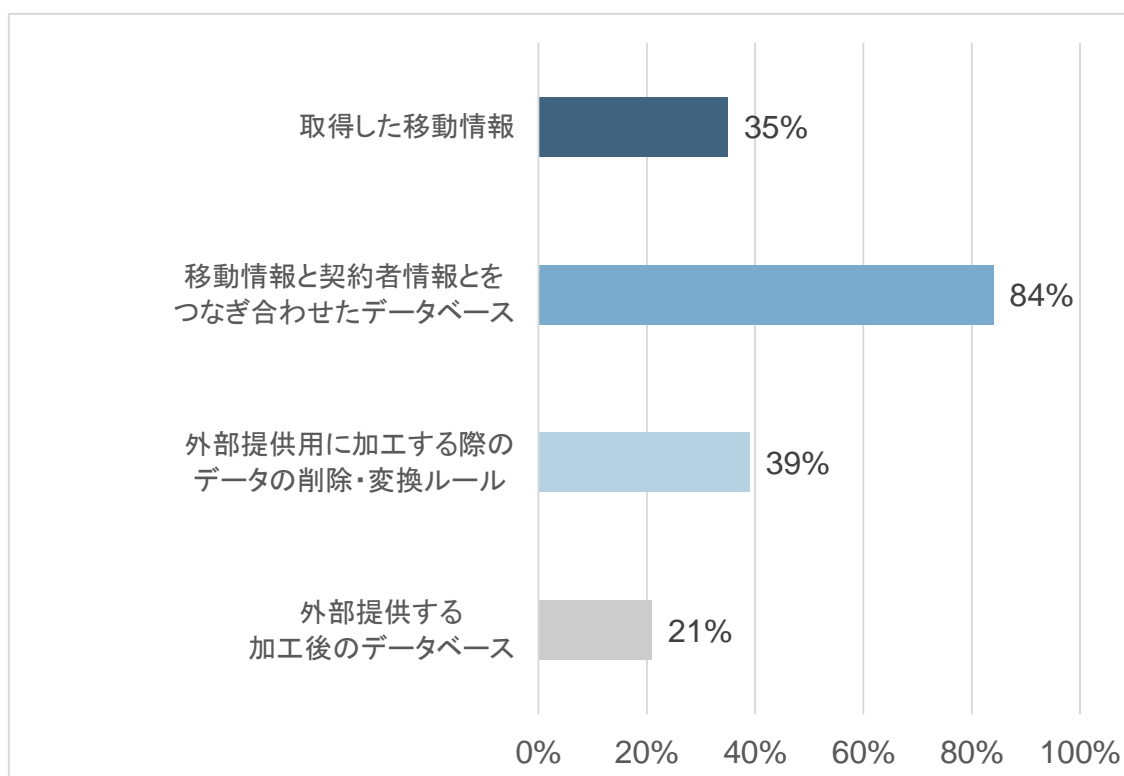
### 《デブスインタビューでの発言》

#### Q: 自分を特定することにつながりやすいと感じた項目の理由

- 居住地→性別→年代→日→移動距離→該当数→趣味：居住地がダントツで、4位5位が同じ。6位7位はあまり気にしない。有楽町は自宅に近いので、ケース2の方がケース1より特定されそう。

### 3-1. 安全管理措置に対する要望

通信会社が位置情報を取得して加工し、外部の事業者に提供するにあたり、通信会社の情報管理体制やITセキュリティを強化して特に保護してほしいと感じるものを次のうちからお答えください。(MA)



#### 《デブスインタビューでの発言》

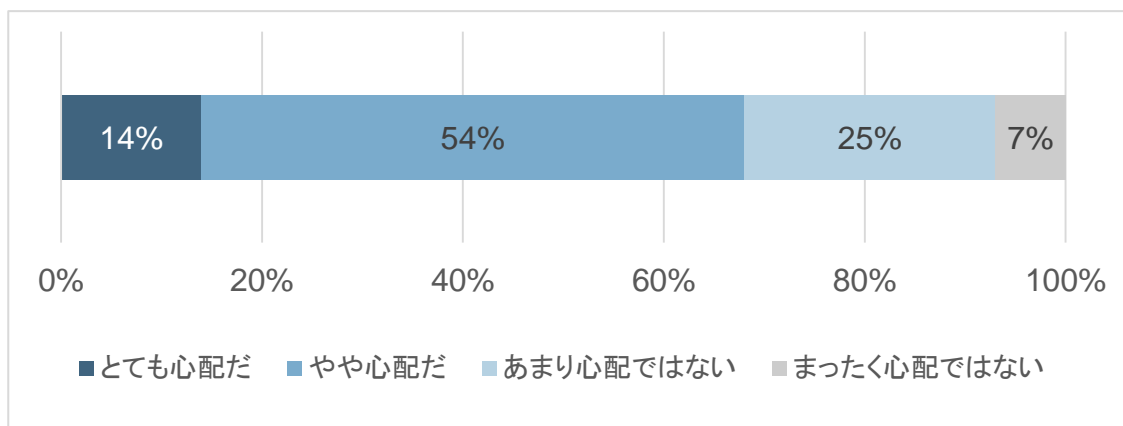
##### Q: ケースの加工手順の中で、特に保護してほしいと感じた部分

- **移動情報と契約者情報とをつなぎ合わせたデータベース**
  - やはり契約者情報が一番。契約者情報と移動がわかると全部わかってしまうという感じがする。
  - つなぎあわせたものは細かい携帯のIDが変わっていないし、年齢も何十代でなく細かくなっていたし、住所は丸々出しているの、保護してほしいと思った。
  - 契約者情報の詳しい住所と年齢にプラス移動情報が組み合わせられると絶対に特定されると思うので、どこかに漏れたりすることはなくしてほしい。
- **外部提供する際のデータの削除・変換ルール**
  - 加工する際のデータの削除・変換ルール。そこが本当にきちっと変換されているかと、秘密にされているか気になるし、保護してほしい。

### 3.2. オプトアウトを設定した場合の受容性の変化

今後、取得した位置情報について個人を特定できないように加工すれば、第三者にデータを提供し、自由に商品開発やマーケティングに活用できるようになります。

そうした加工を行ったデータベースに自分の情報が含まれることについて、心配に思いますか。(SA)



#### 《デブスインタビューでの発言》

##### Q: 心配に思う理由

###### ➤ とても心配だ

- データベースの情報も、加工されたものだけなら大丈夫だと思うが、どれだけ加工されているのかわからないのが心配。
- 知られないで済むならその方が良く、使われたくない。パソコンの技術がすごい人がいて、そういう人がストーカーみたいになったら嫌。
- 悪用しようと思っているすごい人はハッキングで来てしまうと思うし、内部で1人でもすごい技術の人がいて情報を使われ、変なDMとかがたくさん来たら嫌。

###### ➤ やや心配だ

- 大半の場合は大丈夫だと思うが、取得される業者がどんな業者かが心配。データベースを保護している体制があって整備されているような業者に取得されるなら特に心配はないが、新興企業とかでセキュリティ体制があまりしっかりしていないだろうという会社に取得されると心配。
- 第三者というところが嫌。通信会社が取得する分にはまあしょうがないと思っているが、その他には行ってほしくない。
- サービスを含めて、公共的なことに活用できるなら良いと思う。それ以外の特定の会社だといったい何に使われるのかわからない。大きな会社でも心配。

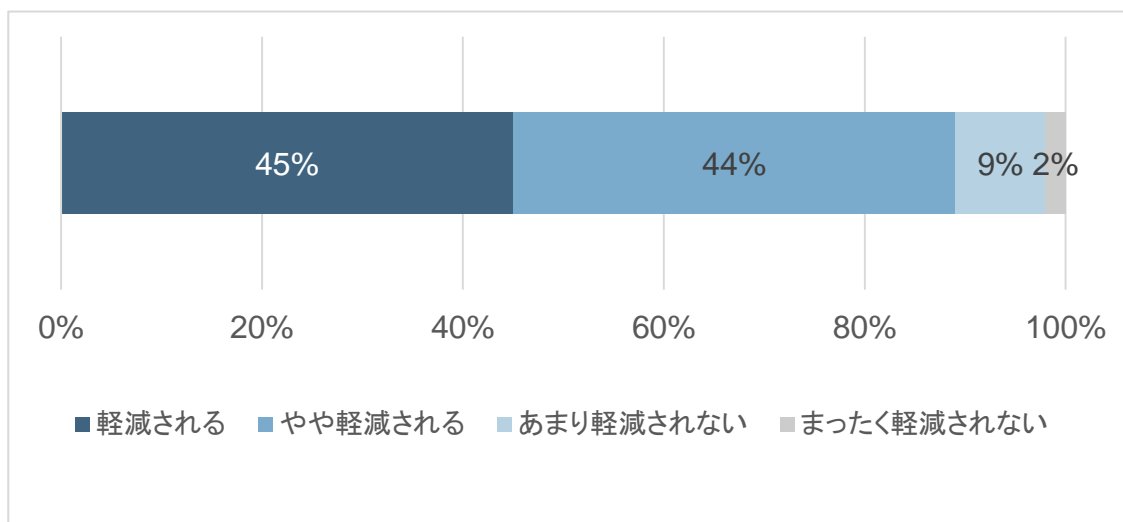
###### ➤ まったく心配ではない

- 企業発展や自分の利益になることがあると思うので、お互いが良くなれば良い。

### 3.-2. オプトアウトを設定した場合の受容性の変化

前問で 1,2,3 と回答した方にお伺いします。

通信会社が位置情報を加工して外部の事業者に提供する場合、自分の情報が含まれないように申請できる仕組みがあるとしたら、心配する気持ちは軽減されますか。 (SA)



#### 《デブスインタビューでの発言》

##### Q: オプトアウトを設定した場合の受容性変化の理由

- **軽減される**
  - ・そもそも、こちらに選べる権利があると思っていなかったし、サービスを利用したら漏れてしまうものかなと思っていた。更にこういう仕組みがあって、こちらに選ぶ権利があれば軽減されるなど思った。
- **まったく軽減されない**
  - ・それでも、元の情報があれば、機械に詳しい人が暴こうとしたら暴けると思う。

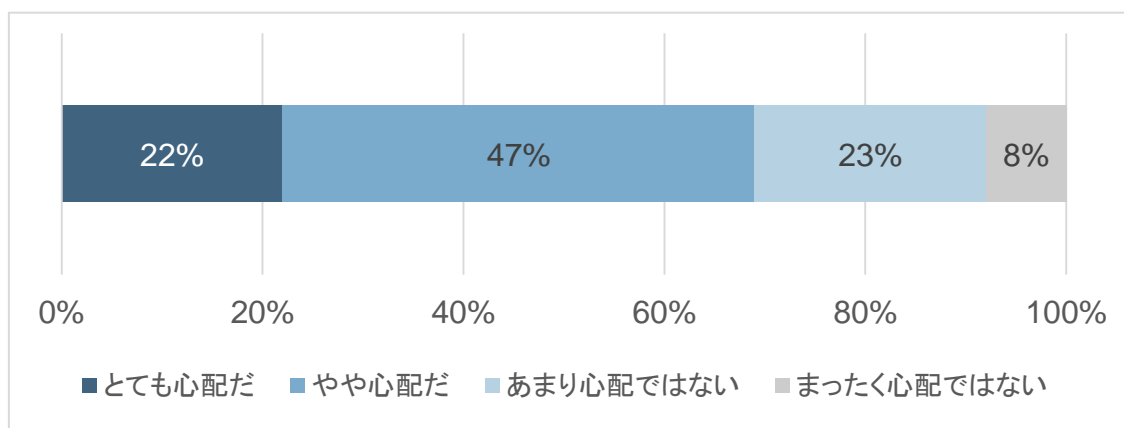
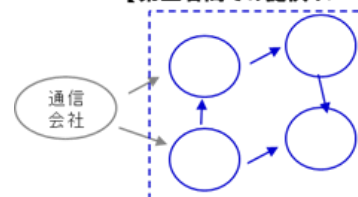
**デブスインタビュー対象者でオプトアウトの利用経験者はいなかった**

### 3-3. 提供範囲を限定した場合の受容性の変化

個人を特定できないように加工して第三者提供された位置情報は、第三者間で提供しあうことが可能になります。

そうした流通するデータベースに自分の情報が含まれることについて、心配に思いますか。(SA)

【第三者間での提供イメージ】



#### 《デプスインタビューでの発言》

##### Q: 心配に思う理由

- **とても心配だ**
  - ・ 通信会社と次の会社までは、その間でちゃんと規定があればしっかり情報が保護されると思うが、そこから次に行くと心配。
  - ・ 自分の知らないところで共有されるのはちょっと怖い。勝手にやりとりしないでほしい。
- **やや心配だ**
  - ・ 第三者だと手を離れるような気がする。通信会社に留まっていればいいが、次から次へというのがちょっと心配。
  - ・ 第三者に行くという時点で嫌だが、そこでその情報を共有する人たちがいると思うと怖い。今の時代だから余計に怖い。
- **あまり心配ではない**
  - ・ ID・年齢なども変わって、2名という集計したレベルで第三者間でその情報を提供しあったとしても、心配ではないと思った。
  - ・ 「4. まったく心配ではない」でもいいが、過去にデータが漏れたと思われる経験しているので。
- **まったく心配ではない**
  - ・ ヤミ金融などに情報が流れるのは心配。食品会社・市とか県は、情報が役に立つなら良い。

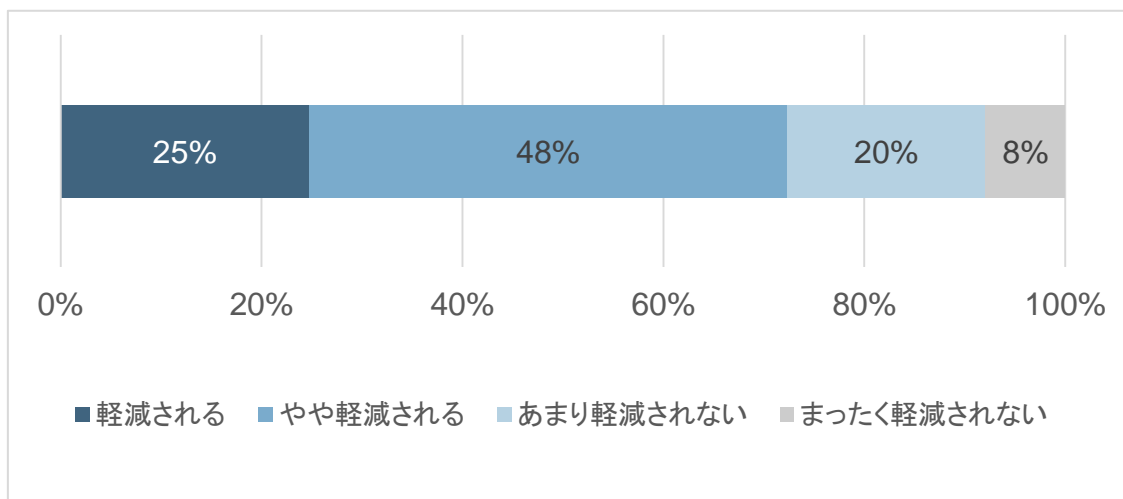
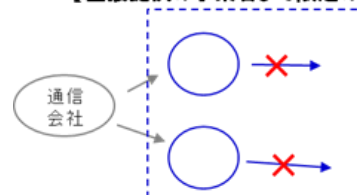


### 3-3. 提供範囲を限定した場合の受容性の変化

前問で1,2,3と回答した方にお伺いします。

通信会社がデータベースを加工して外部の事業者に提供する際に、提供先を通信会社が直接提供する事業者までに限定したら、心配する気持ちは軽減されますか。あてはまるものにお答えください。(SA)

【直接提供の事業者まで限定のイメージ】



#### 《デプスインタビューでの発言》

##### Q: 提供先を限定した場合の受容性変化の理由

###### ➤ 軽減される

- どこまで自分の情報が加工されてどこまで情報がいつているのかわからないより、「ここまでです」とはっきり言われた方が信頼できると思ったので、軽減できるかなと思った。
- 通信会社の傘下までがいい。何か被害があった時に通信会社に言ったら解決してもらえるところまでなら良い。

###### ➤ やや軽減される

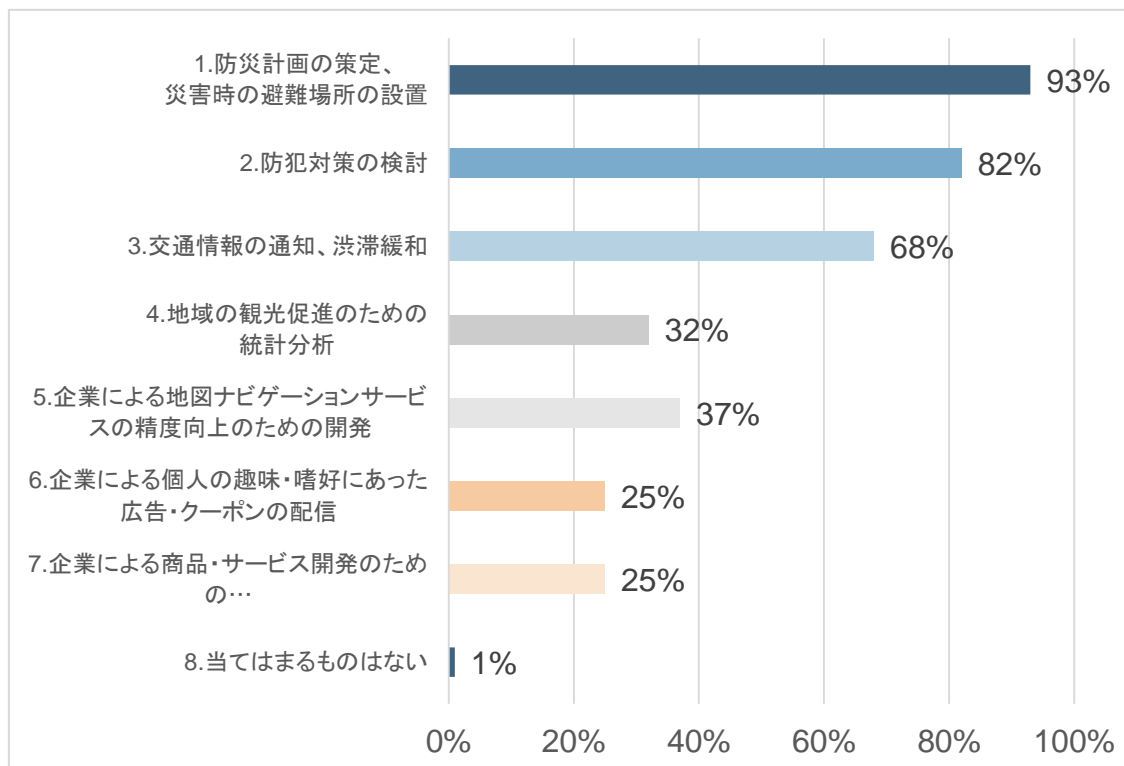
- 通信会社を信頼して契約しているので、そこから第三者へ提供されるのは嫌。できれば、通信会社までで遮断したい。

###### ➤ まったく軽減されない

- グループ企業同士だと、結局グループ内では広がってしまうと思う。

### 3-4. 利用目的を限定した場合の受容性の変化

事業者が個人を特定できないように加工した位置情報を活用するにあたって、どのような目的であれば、自分の位置情報を提供してもよいと思いますか。(MA)



#### 《デプスインタビューでの発言》

##### Q: 位置情報を提供してもよい目的とよくない目的との違い

- 1、2、3は公共性が高いのでOK。4も観光促進で皆のためになるならOK。6は自分にメリットがあるので提供しても良い。5、7は自分にメリットがあまりないような気がするので、すすんで提供したくはない。
- 1、2は自分の命にかかわることなので、それなら別に良いと思う。3～7については自分の足を使って調べてよ、という感じ。自分に還元されることがたくさんあると思うが、企業が楽しんで情報を経て自分の会社の利益に使うというのは納得がいかない。
- 2と3は地域の情報に役立つし、今後の防災・防犯は見直していくべきなので、そういうものにちゃんと情報発信して、何か起きた時に端末で確認できたらいい。6は、個人にとってうれしい情報なら全然問題ない。そういうことでどんどん企業が活性化されていけば、消費者も喜んで情報をもらいたいと思う。それ以外の項目も絶対に嫌というものはない。7の企業による商品・サービス開発のための顧客分析は、変換の内容によっては嫌。どういう使われ方をするのがちょっと心配。
- 加工してあれば使ってもいいし、こういうことに使わないと加工する意味がない。

その他デプスインタビューで確認された特徴的な発言

#### 「十分な匿名化」の加工結果を例示する上での注意点

- 「十分な匿名化」における該当人数の集計結果を見て) この2名とか3名という数字ははじめ、カップルやグループで移動している場合の情報と勘違いしてしまった。集計結果ということであれば、10名以上の数値を例示してもらったほうがわかりやすいと思う。

#### 時間の加工単位に対する意見

- 15分単位というのは細かすぎる。1時間でもちょっと短いので、午前・午後とか2時間くらいが良い。
- 15分単位でそこに該当する人が少ないような気がする。2~3時間単位で、最低人数が10人以上ならまあ特定されないと思う。

## 5. 電気通信事業者が取り扱う位置情報に関するルールの整理

以上述べた1～4の検討結果を踏まえ、電気通信事業者が取り扱う位置情報に関するルールとして、次の二つを整理した。

- 1) 「十分な匿名化」に関するルールに盛り込まれるべき内容
- 2) 電気通信事業者が扱う位置情報を用いた匿名加工情報の作成等に関するルールに盛り込まれるべき内容

1) は、業界団体等に引き継がれて、業界自主ルールとして定められることを、2) は、認定個人情報保護団体に引き継がれて、個人情報保護指針として定められることを想定している。それぞれ、次に掲載する。

## 1) 「十分な匿名化」に関するルールに盛り込まれるべき内容

位置情報に関するプライバシーの適切な保護と社会的活用の両立に向けた調査研究

### 「十分な匿名化」に関するルールに盛り込まれるべき内容

#### 1. 趣旨

本資料は、電気通信事業法における通信の秘密の保護及びプライバシーの保護の趣旨を踏まえ、電気通信事業者が扱う位置情報を、加工して社会的に活用するための方法に関するルールをとりまとめたものである。本資料は、業界団体等に引き継がれて、業界自主ルールとして定められることを想定している。

加工方法の要求レベルは、電気通信事業者が扱う位置情報の精度やユースケースによって判断されるべきもので、普遍的な基準を定めることは困難である。このため、本資料では、ケースバイケースで「十分な匿名化」の要求レベルについて判断することを前提に、最低限遵守すべき基本ルールを定めるとともに、付属資料として具体的なユースケースとその対応方法等を取りまとめている。

また加工方法の他、消費者の安心・安全に資するよう、安全管理措置、消費者への通知及び消費者の同意・選択、プライバシー影響評価（管理運用体制の適切性の評価・検証）に関して、求められる取組みを定めている。

なお、個人情報に該当する位置情報については、個人情報保護法上の規律も参照する必要がある。

#### 2. 用語の定義

##### 2.1. 十分な匿名化

- 加工方法の組合せにより、その時点での技術水準では再特定化・再識別化が不可能又は極めて困難といえる程度に加工することをいう。電気通信事業者が扱う位置情報について、通信の秘密及びプライバシーの保護と社会的活用とを両立するため、位置情報と個別の通信とを紐付けることができないように加工する方法として考案された仕組み。レコード単位で見ると、位置情報（移動履歴を含む）を含む全てのデータ項目について、同じ値を持つレコードが複数あることを作り出し、通信や個人の識別リスクを十分に低減させることを前提に、制度面、技術面、運用面での対処を規定している。

##### 2.2. 位置情報

- 電気通信事業分野ガイドライン第35条第1項に規定する位置情報をいう。携帯電話の基地局に係る位置情報、GPSによる位置情報、Wi-Fiアクセスポイントに係る位置情報等がある。

### 2.3. 付帯情報

- ・ 電気通信事業者が扱う位置情報に付帯する情報。電気通信事業者が保有する契約者に係る情報、電気通信サービスを利用する際に本人が登録する情報等がある。本書では、そのうち電気通信事業者が位置情報と結合して利用することができるものを「付帯情報」という。

### 2.4. 非識別化

- ・ データの匿名化の程度をあらわす概念の一つ。対象となるデータセットにおいて、レコード単位で見て、全てのデータ項目について、同じ値を持つレコードが複数ある状態にすること。個別の通信や特定の個人を識別されないようにする。非識別化の具体的な手法として、「k-匿名化」がある。

### 2.5. オプトアウト

- ・ 本人の明示的な同意を事前に取得せずに、本人の情報を利用し、本人からの求めに応じてその情報の利用を停止する方式。ただし、オプトアウトを採用する場合は、情報の利用目的を事前に通知・公表するとともに、オプトアウトの実質的な機会を確保することが前提となる。

### 2.6. プライバシー影響評価

- ・ ある個人情報の取扱いがプライバシーに対して与える影響を、事前に、識別、分析、評価、診断して、対処する一連の活動をいう。

## 3. 適用範囲

- ・ 電気通信事業者が扱う位置情報を匿名化して利用する場面。当該位置情報のうち、通信の秘密に該当するものを第一義に想定するが、通信の秘密に該当しないものにおいても、本ルールを適用することは、プライバシー保護の観点からは有効であり、望ましい。

## 4. 「十分な匿名化」に係る取扱い

### 4.1. 「十分な匿名化」による加工

#### 4.1.1. 位置情報と付帯情報との結合

- ・ 「十分な匿名化」には、位置情報と付帯情報とを結合して作成したデータを用いることができる。ただし、結合することのできる付帯情報は、次の条件のいずれかに該当しなければならない。
  - 性別、年齢、市区町村までの住所。

- 経時的にデータが積み重ねられることのない情報（性別、年齢、市区町村までの住所を除く）で、以下の①ないし③をすべて満たすもの。
  - ① 単体では個別の通信や個人を特定することができないものであって、かつ他の情報と照合してもなお、個別の通信や個人を特定する可能性が一般に想定されないもの
  - ② 利用者が入力した情報、サービスの提供により電気通信事業者に提供されることが利用者にとって明らかな情報等、電気通信事業者による利用が利用者の想定範囲内にあるもの
  - ③ 「十分な匿名化」をして利用することを公にしているもの。  
現時点で考えられる例としては、利用者が入力した趣味嗜好や言語情報（ただし、十分な対象者数が確保できる場合に限る。）がある。

#### 4.1.2. 入口要件

- ・ 「十分な匿名化」に用いるデータには、「入口要件」として、以下の条件を満たすものに限る。
  - 加工の対象は、位置情報と付帯情報に限る
  - 位置情報と付帯情報とを連結する符号（現に電気通信事業者において扱う情報を相互に連結する符号に限る。）がある場合は当該符号は削除する（当該符号を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）

#### 4.1.3. 出口要件

- ・ 「十分な匿名化」により加工した後のデータは、「出口要件」として、次の条件を満たすことで、当該データと個別の通信や特定の個人とを紐付けることができないようにしなければならない。
  - 全てのデータ項目において、適切な非識別性が確保されていること。
  - 次の1)～9)に掲げる評価要素によって、個別の通信や特定の個人が識別されるリスクを評価し、総合的に判断して、同リスクが十分に低減されていること。  
なお必ずしも1)～9)の全ての評価要素を満たす加工を求めるものではない。

##### 1) 付帯情報

- ・ 付帯情報によっては、個別の通信や特定の個人を識別する可能性が高まることに配慮して選定・加工することが望ましい。  
(対処の例)
  - 加工時に付帯情報として位置情報と結合して用いることが認められる性別であっても、対象とする集団に男女の大きな偏りのあることが想定される場合は、付帯情

報として用いない又は配慮して加工する。

## 2) 場所の特性

- ・ 対象とする位置情報に、自宅、通勤・通学地が含まれる場合は、配慮して加工することが望ましい。
- ・ 対象とする位置情報に、要配慮個人情報に関わる場所が含まれている場合は、配慮して加工することが望ましい。

(対処の例)

- 明らかに自宅、通勤・通学先がわかる場合は、これらを除くことが望ましい。
- 明らかに特定の疾患を対象とする病院に滞留していることが明らかなレコードを、加工対象から除外する。

## 3) 集団の規模

- ・ 特定の学校・職場や稀少な趣味嗜好等を持つ集団を対象とした場合、集団の規模によっては、個別の通信や特定の個人を識別する可能性が高まるため、集団の規模に配慮して加工することが望ましい。

(対処の例)

- 特定の趣味嗜好の集団を取り扱う場合、十分な対象者数が得られることを確認する。

## 4) 取得時期の特性

- ・ 特定のイベントや事件のあった日、時期と一致する可能性がある場合、他の情報を参照することによって、個別の通信や特定の個人を識別する可能性が高まるため、取得時期の特性に配慮して加工することが望ましい。

(対処の例)

- 大規模集客施設において、特定の宗教のイベントが開催されていることが明らかな場合は、当該期間及び当該施設に該当するレコードを、加工対象から除外する。

## 5) 位置の精度

- ・ 高い精度の位置情報は、個別の通信や特定の個人を識別する可能性が高いため、適切に精度を低減することが望ましい。人口密度の低いエリアを対象とする場合は、特に配慮することが望ましい。

(対処の例)

- 位置精度数メートルの緯度・経度情報を、適切な大きさのメッシュ単位の位置情報に変換する。



## 6) 移動履歴の期間・範囲

- ・ 移動履歴の期間は長くなったり、特定の時間帯を対象としたりする場合は、次の a)～c)に係るリスクが高くなるため、これらに配慮して加工することが望ましい。

### a) パターン性

- 定期的に通っている場所、滞留している場所が分かることにより、自宅、通勤・通学地などが推測されて、個別の通信や特定個人の識別性が高まる。

### b) 場所の特性

- 「2) 場所の特性」を参照。

### c) 識別性

- 履歴の一意性が高まる。その一意性をもって、直ちに個別の通信や特定の個人を識別することができないとしても、一定の配慮をすることが望ましい。

(対処の例)

- a)～c)を踏まえ、自宅や通勤・通学地に係るレコードを除外する、移動履歴の期間を短くして提供する、同一の事業者に提供する場合は、履歴の期間が重ならないように提供する等の配慮をする。

## 7) 時間の精度・間隔

- ・ 時間の精度が高まったり、データを取得する際の時間間隔が短くなったりすると、個別の通信や特定の個人を識別する可能性が高まる。また、詳細な時刻情報は位置情報とセットになることで、異なるデータセット間における共通の識別子として機能し得る。このため、適切に時間の精度を低減したり、間隔を開けたりすることが望ましい。

(対処の例)

- 秒単位で取得された時間の精度を、15分単位にまるめる。

## 8) 対象者数

- ・ 加工対象とするデータセットに含まれる対象者数が少ないと、個別の通信や特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。なお対象者数の数え方は、付属資料4を踏まえること。
- ・ 同一の個人が複数台の携帯端末を所持している場合のあることを想定して、携帯端末の台数よりも対象者数が小さくなる可能性のあることに留意することが望ましい。

(対処の例)

- データを対象者数でカウントして、適切な規模の対象者数を確保する。

## 9) データ提供までの期間

- ・ データを取得してから、「十分な匿名化」により加工した位置情報として提供するまでの期間が短い場合は、他の情報を参照することによって、個別の通信や特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。

(対処の例)

- 位置情報を取得してから「十分な匿名化」により加工した情報として提供するまでの期間を、三ヶ月以上確保する。

## 4.2. 安全管理措置

### 4.2.1. 加工方法に関する情報

- ・ 「十分な匿名化」により加工する際、①元の個人情報、②元の個人情報から削除した個人を特定する記述等および加工方法に関する情報の漏えいを防止するため、次に掲げる措置と同等の安全管理措置を講じなければならない。
  - 「電気通信事業における個人情報の保護に関するガイドライン」における「匿名加工情報の作成等（第28条2項）」に掲げる加工方法に係る安全管理措置に対応する措置。

### 4.2.2. 「十分な匿名化」による加工途中の位置情報

- ・ 「十分な匿名化」による加工作業が完了していない位置情報は、通信の秘密を保護するため、次に掲げる措置と同等の安全管理措置を講じなければならない。
  - 「電気通信事業における個人情報の保護に関するガイドライン」における「安全管理措置（第11条）」、「従業者及び委託先の監督（第12条）」、「個人情報保護管理者（第13条）」に対応する措置。
  - 元の個人情報および元の個人情報から削除した個人を特定する記述等が加工後の情報と照合されることを防止する措置<sup>9</sup>

### 4.2.3. 「十分な匿名化」がされた位置情報

- ・ 「十分な匿名化」により加工した位置情報は、個別の通信と紐付けられる可能性が十分に低減されたものであるため、事業者は、通信の秘密保護の観点からは、当該情報に対して特段の安全管理措置を講じなくてよい。

## 4.3. 通知及び同意・選択

### 4.3.1. 通知及び同意の取得

---

<sup>9</sup> 元の個人情報との照合により、十分な匿名化データの安全性が損なわれることがあるので、元の個人情報を第三者提供等する場合は、過去に提供等した十分な匿名化データの安全性が損なわれないように配慮することが望ましい。

- ・ 通信の秘密を含む位置情報を利用するためには、事前に、その旨を本人に通知し、有効な同意を取得しなければならず、その同意は原則として個別かつ明確な同意でなければならない。ただし、「十分な匿名化」により加工した位置情報について、次のすべての条件を満たせば、事前の包括的同意に基づいて活用をすることができる。
  - オプトアウトが適切に提供されている。
  - 加工及び運用管理体制（安全管理措置等）が適切である。
  - プライバシー影響評価（PIA）が適切に運用されている。
- ・ 事前の包括的同意を取得する場合は、付属資料2を参考に、本人にわかりやすく伝えなければならない。

#### 4.3.2. 苦情・相談窓口

- ・ 「十分な匿名化」により加工した位置情報の活用をするためには、苦情・相談窓口を設置し、本人からの苦情や相談へ対応しなければならない。

#### 4.4. オプトアウト

- ・ 事前の包括的同意で「十分な匿名化」により加工した位置情報を活用する場合、本人に、オプトアウトの手段を提供しなければならない。
- ・ 本人が、ウェブサイト、電話等により容易にオプトアウトすることができるように努めなければならない。

#### 4.5. プライバシー影響評価（PIA）－管理運用体制の適切性の評価・検証－

- ・ 「十分な匿名化」により加工した位置情報を活用する場合、プライバシー影響評価（PIA）を実施して、通知及び同意・選択の方法、加工方法、安全管理措置の管理運用に係る一連の取組みにおける適切性を評価・検証し、その評価結果を公表する。
- ・ PIAを実施し、適切な管理運用体制を確保することにより、通信の秘密及びプライバシーの保護を確実にするよう努めなければならない。
- ・ PIAを実施する際は、付属資料3を参考に、適切な評価項目を設定するとともに適切に評価結果を公表すること。

#### 5. ルールの見直し

- ・ 「十分な匿名化」に関するルールは、制度改正や技術の進展を踏まえ、定期的に見直しを行って、所要の措置を講じなければならない。

## 付属資料

- ・付属1：電気通信事業者が扱う位置情報を用いた匿名加工情報の作成等に関するルールに盛り込まれるべき内容<sup>10</sup>
- ・付属2：通知・同意取得に関する文案
  - Web等での説明事項（案）
  - 包括同意約款（案）
- ・付属3：プライバシー影響評価（PIA）の参考資料
  - 評価項目
  - 評価結果の公表方法
- ・付属4：移動履歴のカウント方法
- ・付属5：「十分な匿名化」による加工事例
  - 商用の事例
  - 観光の事例
  - 交通の事例

---

<sup>10</sup> 認定個人情報保護団体による「個人情報保護指針」に盛り込まれることを想定した内容を、ここで参照する。

## 付属 1 : 匿名加工情報の取扱いに関するルール

(略)

付属 2 : 通知・同意取得に関する文案

Web 等での説明事項(案)

No	項目	モデル案(実際に適用する際には必要に応じてカスタマイズする)
ア	取得者(位置情報の利用事業者)	****株式会社(以下、当社)
イ	位置情報の種類 ※基地局情報、GPS 位置情報、Wi-Fi 位置情報等	<p>当社は、お客様が利用する端末の位置情報として、基地局位置情報、GPS 位置情報、Wi-Fi 位置情報を取得します。なお、これらには通信の秘密に該当する位置情報が含まれる場合があります。このサイトでは、これらの位置情報のうち、通信の秘密に該当する位置情報を、〇〇約款第〇条に定める匿名化利用(以下、匿名化利用)する場合についてご説明します。</p> <p>(1)基地局位置情報 お客様の端末が接続した基地局の緯度・経度等の情報およびその時刻</p> <p>(2)GPS 位置情報 〇〇〇</p> <p>(3)Wi-Fi 位置情報 〇〇〇</p>
ウ	位置情報の精度、取得頻度、ID 有効期間	<p>当社は、お客様が利用する端末の位置情報を以下のように取得します。</p> <p>(1)精度 GPS・Wi-Fi 位置情報/数 m～数十 m、基地局位置情報/数百 m</p> <p>(2)取得頻度 通信・通話ごとまたは〇秒ごと</p> <p>(3)ID の有効期間 〇日間以下</p>
エ	加工の手法	<p>当社は、匿名化利用を行う場合、以下の加工手法を組合せて実施することで、お客様から取得した情報(上記位置情報および下記お客様情報)につき、十分な匿名化を行います。</p> <p>利用する加工手法:</p> <p>(1) 項目削除 加工対象となる個人情報データベース等に含まれる個人情報の項目を削除するもの。例えば、年齢のデータを全ての個人情報から削除すること。</p> <p>(2)一般化 加工対象となる情報に含まれる記述等について、上位概念若しくは数値に置き換えること。例えば、「年齢 70 歳」を「高齢者」に置き換えること。</p> <p>(3)丸め(ラウンディング) 加工対象となるデータベースに含まれる数値に対して、四捨五入等して得られた数値に置き換えることとするもの。</p> <p>《上記(1)～(3)は例示であり、特に各社にて状況に応じ変更する箇所になります。》</p>

No	項目	モデル案(実際に適用する際には必要に応じてカスタマイズする)
オ	利用目的	<p>当社は、匿名化利用において、位置情報を以下の目的で利用します。</p> <p>(1)お客さまに有益と認める情報についての、当社による表示・配信  (2)当社による〇〇サービスに関するマーケティング調査および分析  (3)当社〇〇サービスの品質向上や、新商品・新サービスの企画・開発・提供  (4)当社〇〇サービスについての当社によるご利用状況分析  (5)官公庁、公共団体、一般企業等への人口動態分析、マーケティング分析等の各種分析結果の提供</p> <p>《上記(1)～(5)は例示であり、特に各社にて状況に応じ変更する箇所になります。》</p>
カ	第三者提供の有無及びその提供先	<p>当社は、上記の利用目的で、個人を容易に特定できないよう十分に匿名化した位置情報等を、反社会的勢力でない第三者に提供することがあります。また、そのような提供先に対し、特定の個人を識別する目的で他の情報と照合しないことを契約上義務付けます。さらに、提供先が別の会社等に提供を行う場合は、それ以降の提供先にも同様の義務を課すことを求めます。</p>
キ	保存期間	<p>加工対象情報(加工前・加工中の位置情報)は、正当業務行為の範囲を超える〇ヶ月のみ保存することとします。十分な匿名化後の情報については、特に保存期間の制限を設けません。</p>
ク	位置情報に紐付けて利用される情報	<p>当社は、匿名化利用において、上記の利用目的を達成するために、以下のお客様情報を位置情報へ紐付けて利用します。</p> <p>(1)住所(市区町村名までの住所情報とします。)  (2)年齢  (3)性別</p> <p>《お客様情報は、住所、年齢および性別に限られます。》</p>
ケ	本人関与の仕組み	<p>お客様から、位置情報の匿名化利用を停止するようお申し出があった場合、そのお申し出以降、そのお客様の位置情報の匿名化利用を停止いたします(「十分な匿名化」により加工をした場合、加工後の情報は除きます)。  位置情報の匿名化利用停止のご連絡先は以下となります。</p> <p>(1)Web ページから  <a href="http://www.***.com/*****/****">http://www.***.com/*****/****</a></p> <p>(2)電話から</p> <ul style="list-style-type: none"> <li>・ * * 携帯電話から: 局番なし * * * (無料)</li> <li>・ 一般電話から: ****-**-**** (無料)</li> <li>・ 受付時間: 9:00～20:00 (土・日・祝日も受付)</li> </ul>
コ	情報管理体制	<p>通信の秘密に該当する位置情報を匿名化利用する場合、位置情報の匿名化の加工途中の情報ならびに元の個人情報から削除した情報および加工方法に関する情報についての管理運用体制を適正に構築し運用します。また、加工方法および管理運用体制について評価を行い、その結果を別途公表します。</p>

## 包括同意約款(案)

### モデル案(実際に適用する際には必要に応じてカスタマイズする)

#### 【利用する情報】

当社は、通信の秘密に該当する位置情報(通信の場所、日時、端末識別符号に限ります。)を、電気通信業務の提供を目的に取得・利用するほか、お客様情報(市区町村名までの住所情報、年齢、性別に限ります)とともに、次の利用目的のために、十分な匿名化を行ったうえで利用します。(以下、匿名化利用)

匿名化利用の内容は、以下のとおりです。なお、詳細については、「Web 等での説明事項案」をご参照ください。

#### 【匿名化手法】

個人の再特定化・再識別化が極めて困難といえる程度に加工します。

具体的な匿名化手法は「Web 等での説明事項案」に記載します。

#### 【利用目的】

(1) 当社サービスの表示・配信

(2) 当社による〇〇サービスに関するマーケティング調査および分析

(3) 当社〇〇サービスの品質向上や、新商品・新サービスの企画・開発・提供

(4) 当社〇〇サービスについての当社によるご利用状況分析

(5) 官公庁、公共団体、一般企業等への人口動態分析、マーケティング分析等の各種分析結果の提供

《上記(1)~(5)は例示であり、特に各社にて状況に応じ変更する箇所になります。別途「プライバシーポリシー」等で記載されたものを参照することも可としますが、その場合、「プライバシーポリシー」において、本包括同意約款に基づいて通信の秘密に該当する位置情報を利用する場合の利用目的の項目が存在し、かつお客様に分かりやすい場所に明記されている必要があります。》

#### 【第三者提供】

上記位置情報およびお客様情報、端末識別符号は、特定の個人を容易に識別できないよう十分な匿名化を行ったうえで、上記利用目的の範囲内で官公庁、公共団体、一般企業等の第三者に提供することがあります。

#### 【情報の利用・第三者提供の停止】

お客様は、当社が取得した位置情報、お客様情報、その他端末識別符号の当社における匿名化利用(「十分な匿名化」により加工した場合、加工後の情報は除きます。)について、以下の URL から停止を申出ることができます。 <https://...>

#### 【個人情報保護方針】

当社の個人情報保護方針は、別途当社の「プライバシーポリシー」に記載いたします。

#### 【約款の変更】

利用目的の変更や取得するデータ項目の変更に伴い、本約款を変更する場合があります。



## 付属3：プライバシー影響評価（PIA）

### 1. 評価項目

#### (1) 全般的事項

- (ア) 位置情報や付帯情報（以下、「位置情報等」）の取扱いに関して、責任者や担当者が明確になっていること。また、情報漏えい等、インシデント発生時の報告体制が決められていること。
- (イ) 位置情報等を対象に含めたプライバシーポリシーを策定していること。
- (ウ) 個人情報保護(安全管理措置を含む)に関する規程、マニュアルが作られており、位置情報等が保護の対象になっていること。
- (エ) 上記規程、マニュアルの内容が関係者に周知されていること。上記規程、マニュアルの内容が適時見直されていること。
- (オ) 位置情報等を格納する場所が特定されていること。
- (カ) 位置情報等へのアクセス権限は、必要な要員にのみ付与されること。原則として1人1IDとし、共用IDは禁止する。また、不正利用を防止するため、以下のように認証管理を強化すること。また、アクセスが不要になった場合は速やかに削除されること(論理的アクセス/物理的アクセス)。
  - ① パスワードの強度
    - \* 一定以上の文字数
    - \* 英数記号混合
  - ② 以前に使用したパスワードは使用禁止
  - ③ パスワードは一定期間毎に変更
- (キ) 位置情報等へのアクセスについて、以下の記録要件を遵守すること。
  - ① アクセスした者、正確な日時、接続元コンピュータ名、成功・失敗を示す情報等を記録すること
  - ② 上記の記録へのアクセス・初期化・停止(一時停止含む)を記録すること
  - ③ 上記の記録は不正アクセスから保護すること
- (ク) 位置情報等の取扱いが適切に行なわれているかについて適時状況が確認されていること(監査、自己チェック等)。

#### (2) データの取得

- (ア) 位置情報等の取得に先立って、利用者から同意を得ていること。その際、取得するデータの項目や利用目的等、利用者が同意する内容が明確になっていること。
- (イ) 利用者が同意した日時を記録し、いつでも提示できるようにすること。
- (ウ) 不必要な情報を取得しないこと。
- (エ) 利用等停止の申し入れ等、利用者関与の機会が確保されていること。

- (オ) 利用目的の変更を行う場合は、変更前の利用目的と関連性を有すると合理的に認められる範囲を超えて行ってはならない。
- (カ) 合併や営業譲渡等により事業の承継があった場合、承継前の利用目的の達成に必要な範囲を超えて承継前に取得した位置情報等を取扱わないこと。
- (キ) 取得の際に利用者に対して、位置情報等の取扱いについて分かりやすく説明・表示を行っていること。
  - ① 説明・表示の方法
    - 1. 位置情報取得時における同意取得の場面での説明・表示を行うこと
  - ② 説明事項
    - 1. 取得者（位置情報の利用者）
    - 2. 位置情報の種類（基地局情報、GPS 位置情報、Wi-Fi 位置情報等）
    - 3. 精度、取得頻度、追跡期間
    - 4. 加工方法
    - 5. 利用目的
    - 6. 第三者提供の有無その他第三者提供に関する事項
    - 7. 保存期間
    - 8. 付帯情報の種類（位置情報に紐付けて利用される他の利用者情報）
    - 9. 利用者関与の仕組み
    - 10. 情報管理体制 等

### (3) 加工前および加工中のデータ等の保管

- (ア) 加工前・加工中のデータ、加工の際に元の位置情報等から削除した情報および加工の方法に関する情報等（以下、「加工前データ等」）は、所定の場所で安全に管理されること。
- (イ) 加工前データ等について、機密性を高める以下のような施策を講じること。
  - 【機密性を高める施策】**
  - ① アクセスに対するモニタリング
  - ② 保管時の暗号化
  - ③ 秘密分散 等
- (ウ) システムの設置場所等、加工前データ等の保管場所において、媒体や機器の持ち込み・持ち出し制限が行なわれていること。
- (エ) システムの設置場所等、加工前データ等の保管場所において、カメラ等の監視装置が設置されていること。また一定期間その監視記録が参照できるようになっていること。
- (オ) 利用目的に応じた保管期間を定めること。
- (カ) 保管期間経過後または利用目的達成後でも加工前データ等を保管する場合は、

相応の理由を定めること。

- (キ) 安全管理に関する教育を従業者に対して実施すること。
- (ク) 外部からの不正アクセスを防止するための措置(ファイアウォールの設置等)を講ずること。
- (ケ) 加工前データ等を扱う端末等と媒体や機器との接続を制限すること。
- (コ) 加工前データ等については、加工後のデータと照合することで特定の個人を再特定しないことを担保するような管理策・方法が講じられていること。
- (サ) その他加工前データ等の保管において想定されるリスクに対し、必要な対策を講じておくこと。

#### (4) 十分な匿名化加工

- (ア) 加工作業を担当する要員が、業務外で位置情報等を使用、複製等することに対し予防措置を設けること。
- (イ) 加工について作業記録を残すこと、また、加工を行うプログラムの作成・実行・変更・削除の記録を残すこと。作業記録を適時に参照し、意図しない加工が行なわれていないか確認すること。作業記録には、アクセスした者、正確な日時、接続元コンピュータ名等を記録すること。作業記録は不正アクセスから保護すること。
- (ウ) 加工方法は、所定のものを用いること。
- (エ) 加工作業を担当する要員を限定すること ((1)全般的事項(カ)を参照)。
- (オ) 付帯情報は、次の条件のいずれかに該当すること。
  - ① 性別、年齢、市区町村までの住所。
  - ② ①以外の情報で、以下の(i)ないし(iii)を満たすもの。
    - (i) 経時的にデータが積み重ねられることのない情報で、単体では個人を特定することができないものであって、他の情報と照合してもなお、個人を特定する可能性が一般に想定されないものであること、(ii) 利用者が入力した情報、サービスの提供により電気通信事業者に提供されることが利用者にとって明らかな情報等、電気通信事業者による利用が利用者の想定範囲内にあるものであること、(iii) 「十分な匿名化」をして利用することを公にしているものであること。
- (カ) 加工を行うことのできるデータは、位置情報と付帯情報に限定され、「入口要件」として、さらに位置情報と付帯情報とを連結する符号(現に電気通信事業者において扱う情報を相互に連結する符号に限る。)を削除すること。
- (キ) 加工をした後のデータは、「出口要件」として、次の条件を満たすこと。
  - ① 全てのデータ項目において、適切なk-匿名性が確保されていること。
  - ② 次の1)～9)に掲げる評価要素によって、特定の個人が識別されるリスクを定性的に評価し、同リスクが十分に低減されていること。

- 1) 付帯情報
- 2) 場所の特性
- 3) 母集団の特性
- 4) 取得時期の特性
- 5) 位置の精度
- 6) 移動履歴の期間・範囲
- 7) 時間の精度・間隔
- 8) 対象者数
- 9) データ提供までの期間

(ク) その他データの加工において想定されるリスクに対し、必要な対策を講じておくこと。

#### (5) 加工後のデータの提供その他の利用

- (ア) 加工後のデータの利用は、位置情報等の取得時に同意を取った利用目的に限定して行われること。
- (イ) 第三者(委託・再委託も含む)への提供は、位置情報等の取得時に同意を取っている場合に限り行われること
- (ウ) 第三者(委託・再委託も含む)への提供にあたって手続きが定められており、決められた目的以外での提供ができないルールになっていること。
- (エ) 第三者(委託・再委託も含む)への提供の記録が残されること(日付、目的、責任者名)。
- (オ) 提供される加工後のデータの項目や提供方法が、提供に先立って公表されていること(委託・再委託の場合必須ではない)。
- (カ) 加工後のデータを第三者(委託・再委託も含む)に提供する場合、利用者から要求があった場合、または利用者のプライバシーが不当に侵害されていると判断される場合は、加工後のデータの提供を停止できるようにしておくこと。
- (キ) 加工後のデータを第三者(委託・再委託も含む)に提供する場合、本人を識別するために、加工後のデータを他の情報と照合してはならず、加工前データ等を取得しないことを、契約書等で提供先に義務づけること。
- (ク) その他加工後のデータの提供その他の利用において想定されるリスクに対し、必要な対策を講じておくこと。

#### (6) データの消去

- (ア) 加工前および加工中のそれぞれのデータについて、不必要なものは消去すること。加工前データ等も速やかに消去すること。
- (イ) 利用目的を終了した加工後のデータを速やかに消去すること。

- (ウ) データ消去の際には、ダミーデータを上書きする等、容易にデータの復元ができないようにすること。
- (エ) サーバやストレージ、外部記憶媒体等を廃棄する際には、データが復元できないように、消磁等の論理的対応か物理的な破壊を行うこと。
- (オ) その他データの消去において想定されるリスクに対し、必要な対策を講じておくこと。

## 2. 公表方法

電気通信事業者のホームページ等で、有識者の確認結果を踏まえた PIA 評価結果を公表する。

### 【推奨版】

評価項目ごとに、評価結果を公表する。PIA を初めて実施する事業者は、利用者への情報提供において透明性を高めるため、上記案を推奨する。

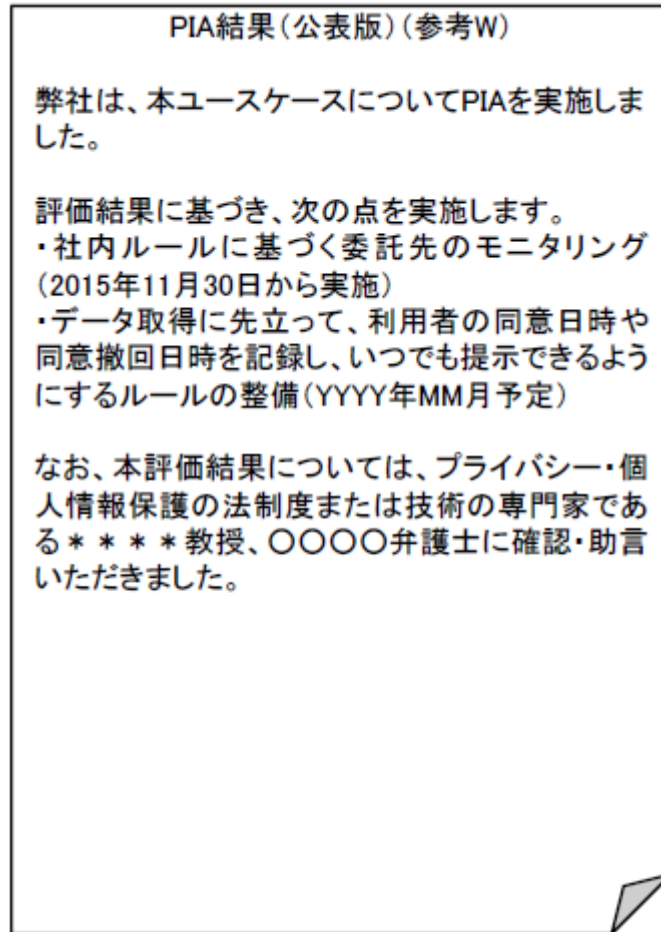
図表 PIA 結果（公表版）①推奨版

PIA 結果(公表版)			
項番	リスク対策	関連する基準等	評価結果
1. 全般的事項			
1.1	位置情報や付帯情報の取扱い・・・	電気通信事業における・・・	YYYY 年 MM 月までに、位置情報の取り扱いの見直しサイクルを社内規程等に明記する予定です。 上記以外について、特段齟齬は見られませんでした。
∫	∫	∫	
1.9	位置情報や付帯情報の取扱いが適切に・・・	電気通信事業における・・・	
2. データの取得（抽出）			
2.1	データの取得（抽出）に先立って・・・		左記リスク対策について、資料閲覧とヒアリングを行った範囲では、特段齟齬は見られませんでした。
∫			
2.11	十分な匿名化を施した場合でも・・・		
3. データの匿名化加工			
4. データの消去			

【簡略版】

PIA 評価項目・評価観点と齟齬が少ない等、プライバシーへの影響が特に小さいと評価される場合は、簡略な方法で公開することも考えられる。

図表 PIA 結果（公表版）②簡略版



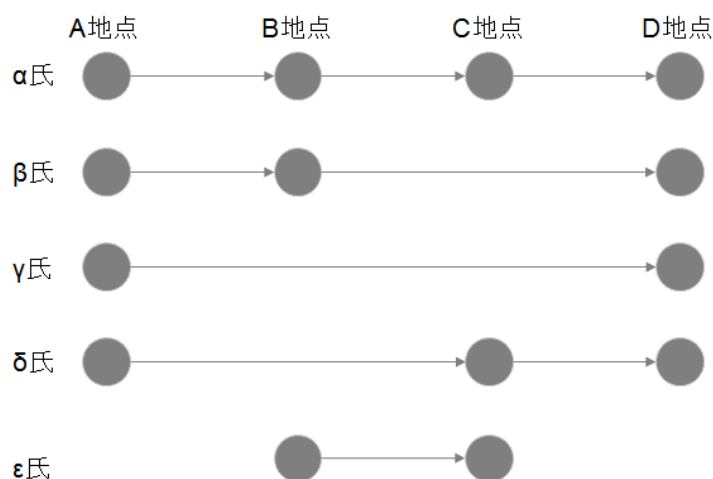
## 付属 4 : 移動履歴のカウント方法

ある人物の位置情報を複数つなげると、その人の移動履歴が構成される。このとき、移動履歴の数を明確にしておかなければ、分析結果を利用する側にとって混乱が生じたり、個別の通信や特定個人の識別リスクが高まったりするおそれがある。このため、ここでは、移動履歴のカウント方法とその適用上の留意事項について整理する。

### ■移動履歴のカウント方法

移動履歴のカウント方法は、「図 1 移動履歴のカウントに係るサンプルモデル」を用いて、「ホップ」と「トリップ」という用語を定義して説明する。本モデルにおいて、●を通過地点（あるいは滞留地点）、時間軸を左から右に経過するものとし、同一時間軸上の地点は時刻が同一もしくは同一とみなすものとする。

図 1 移動履歴のカウントに係るサンプルモデル



#### ホップ

移動履歴の最小単位。2 地点間の移動を 1 ホップ、3 地点間の移動を 2 ホップと数える。1 地点に滞留しているだけの場合は 0 ホップとなる。

図 1 のモデルで考えると、α 氏は、A 地点から D 地点までの移動にあたって、A→B、B→C、C→D の 3 ホップをしていると数える。一方、β 氏は、A 地点から D 地点まで移動にあたって、A→B、B→D の 2 ホップをしていると数える。

#### トリップ

移動履歴のカウント単位。連続する複数のホップをつなげた”一連のホップ”を分析対象とする場合は、これを 1 トリップとしてカウントする。1 ホップのみであっても、これを移動



履歴として分析対象とする場合は、1トリップとしてカウントする。

図1のモデルで考えると、 $\alpha$ 氏において $A \rightarrow B$ 、 $B \rightarrow C$ 、 $C \rightarrow D$ の一連のホップを分析対象とする場合は、 $A \rightarrow B \rightarrow C \rightarrow D$ という一つのトリップとしてカウントする。このとき、図1では、 $A \rightarrow B \rightarrow C \rightarrow D$ のトリップに該当するのは、 $\alpha$ 氏のみである。

また $\alpha$ 氏において、 $A \rightarrow B$ 、 $B \rightarrow C$ 、 $C \rightarrow D$ のホップそれぞれを分析対象とする場合は、各ホップを独立のトリップとしてカウントする。このとき、図1では、 $A \rightarrow B$ のトリップに該当するのは $\alpha$ 氏と $\beta$ 氏の2名、 $B \rightarrow C$ のトリップに該当するのは $\alpha$ 氏と $\varepsilon$ 氏の2名、 $C \rightarrow D$ のトリップに該当するのは $\alpha$ 氏と $\delta$ 氏の2名となる。

### ■移動履歴をカウントする際の留意事項

次の1～3に留意して、移動履歴をカウントすることが有効である。

#### 1. 対象とするトリップの分け方を明確にする。

移動履歴として、どのホップまたは一連のホップをトリップの対象とするのかを特定する。なお、同一データセットに異なるトリップの分け方を適用して異なる匿名化データを作成することには注意が必要である。たとえば、同一のデータセットから作成した2種類の匿名化データにおいて、一方の匿名化データで $A \rightarrow B \rightarrow C$ が5件、もう一方の匿名化データで $B \rightarrow C \rightarrow D$ が6件あった場合、 $A \rightarrow B \rightarrow C \rightarrow D$ と移動した履歴が1件であることがわかってしまう場合があるからである。

#### 2. データセットにおいて、同時に、ホップが重なり合う異なるトリップを集計しない。

図1のモデルで考えると、 $A \rightarrow B \rightarrow C \rightarrow D$ のトリップを集計する場合は、同時に、 $B \rightarrow C$ の集計を行ってはならない。同一人物の移動を2重にカウントすることになるからである。

#### 3. リピータ（同一人物の同一トリップ）へ配慮する。

同一人物が同一のトリップを繰り返す場合がある。通勤、通学、通院のような日常的な移動に加え、観光やレジャーなどの不定期の移動も、同一人物による同一トリップに該当する可能性がある。

リピータのトリップを分析対象とする場合は、個人単位でトリップを集約するなどの配慮が必要である。

一方、観光地への入込客を延べ人数で集計したい場合は、個人単位で集約して集計することはしないが、個人の識別性への配慮は必要である。たとえば、対象とするトリップが100件あったとしても、すべて同一人物によるものである場合は、識別リスクが高まるため、個人単位で集計して非識別性を評価する必要がある。

## 付属 5 : 「十分な匿名化」の加工事例

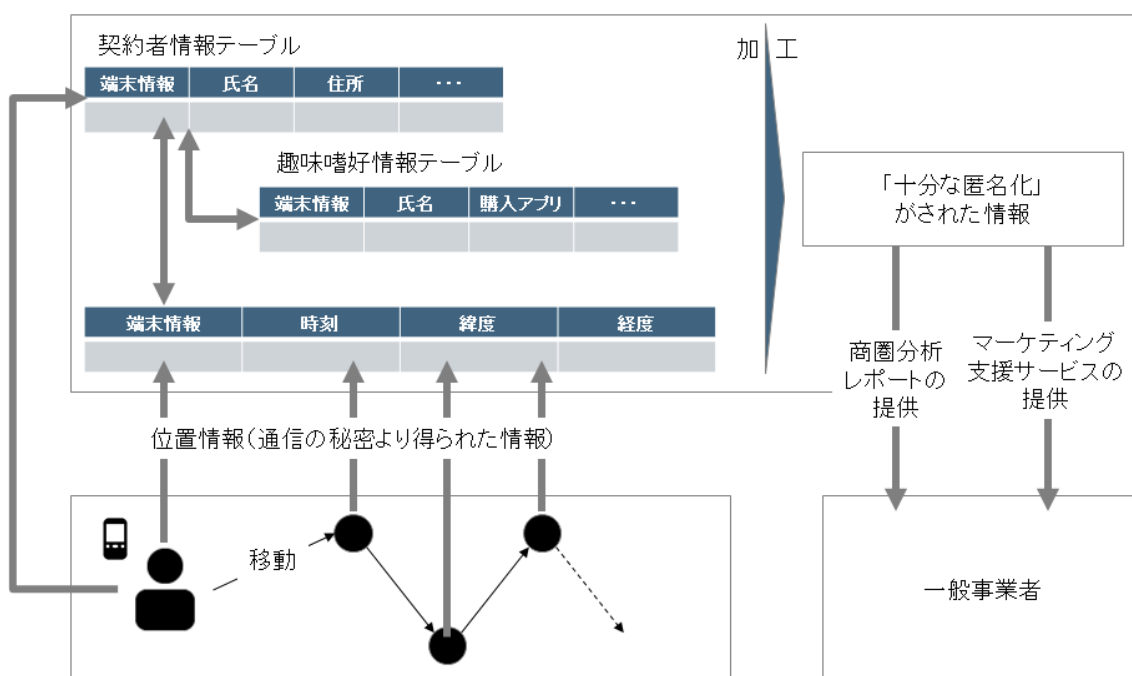
### 1. 商用の事例<sup>11</sup>

#### 1) ユースケース

本ユースケースは、電気通信事業者が保有する通信の秘密に該当する位置情報や契約者情報等を、「十分な匿名化」の枠組みを活用して、一般事業者等へ提供するものである。

電気通信事業者においては、店舗の商圈の分析<sup>12</sup>や、マーケティング支援サービス等の販売促進活動に活用することが想定される。

図表 1-1 電気通信事業者が保有する位置情報を利用するユースケースのイメージ

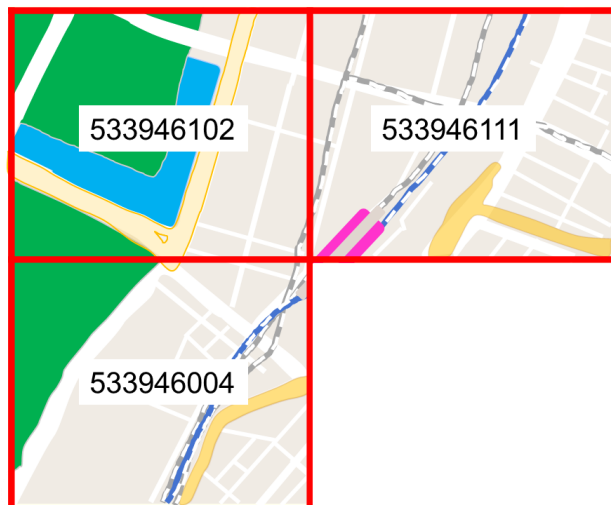


本ユースケースで取扱う位置情報は、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアを対象とし、スマートフォン等の携帯端末の基地局との通信で得られる位置情報を、匿名加工の前処理段階で、1/2 メッシュコード (500m 四方のメッシュコード) に変換して用いる。

<sup>11</sup> 本事例の作成に用いたデータはすべて、加工方法の検討用に作成したダミーデータを用いている。

<sup>12</sup> 例えば、新店舗の立地条件を考える際に、ターゲットとしたい顧客層の移動状況を分析することが考えられる。

図表 1-2 対象エリアにおけるメッシュのイメージ (数値はメッシュコードを表す)



付帯情報については、端末契約時の契約者情報、及び電気通信事業者が提供するアプリ購入時に、契約者から同意を得て取得する趣味嗜好情報から構成される。

全てのテーブルは、端末識別番号によって紐付けが可能である。

図表 1-3 位置情報に関するデータのレイアウトイメージ

契約者情報

端末識別番号	氏名	性別	生年月日	住所
008542314-5	田中 一郎	男	1972 年 4 月 4 日	東京都千代田区丸の内 6-5-3
088539881-2	佐藤 幸子	女	1993 年 12 月 9 日	東京都西東京市保谷町 2-8-18
063381312-1	鈴木 博	男	1938 年 8 月 23 日	神奈川県横浜市旭区二俣川 236

趣味嗜好情報

端末識別番号	趣味嗜好
008542314-5	オペラ観劇
028809315-1	野球観戦
088539881-2	マジックショー

位置情報

端末識別番号	日時	メッシュコード
008542314-5	20161128132415	533946004
008542314-5	20161128150212	553946004
028809315-1	20161130020249	533946111

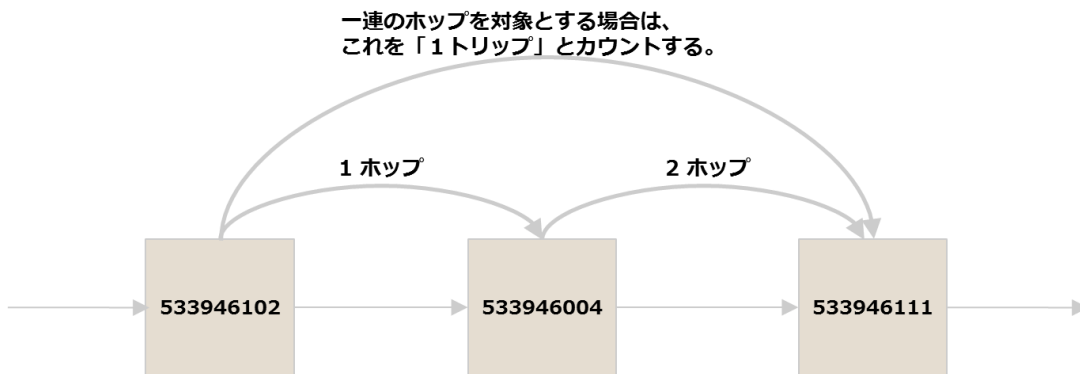
#### <移動履歴のカウント方法>

本ユースケースでは、利用者の移動履歴から予め定義した滞留点を抽出し、二つの滞留点間の移動を1 ホップ、三つの滞留点間の移動を2 ホップとする。一つの滞留点のみの位置情報の場合は、0 ホップとする。同一メッシュで15分以上連続した位置情報がある場合は、利用者が、同メッシュ内のエリアに滞留しているものとみなし、当該メッシュを滞留点とする。

連続する複数のホップをつなげた”一連のホップ”を分析対象とする場合は、これを1つの移動履歴（トリップ）としてカウントする。なお2重カウントを避けるため、トリップを構成するホップを、同時には移動履歴の単位としてカウントしない。

また、個人単位でトリップを集約して、リピータの移動も把握できるようにする。

図表 1-4 移動履歴のカウント方法



## 2) 考慮すべき事項とリスクに対応した具体的な加工方法の検討

#### <位置情報と付帯情報との結合>

付帯情報の条件を踏まえ、契約者情報、趣味嗜好情報を次の通りに加工した上で、端末識別符号をキーにて、位置情報と結合する。

- 氏名は、削除する。
- 性別は、加工せず、そのまま用いる。
- 生年月日は、20歳未満、20代、30代、40代、50代、60代、70歳以上の7つの年代に置き換える。(丸め)
- 趣味嗜好は、一般的なカテゴリとする。マイナーなカテゴリは、他のカテゴリと統合して、一般化する。

#### <入口要件>

端末識別符号が、位置情報と付帯情報とを連結する符号に該当する。位置情報と付帯情報

とを結合した後、これを削除する、又はハッシュ化等により一方向に仮 ID 化する。

#### < 出口要件 >

全てのデータ項目において適切な非識別性を確保し、あわせて、9つの評価要素によって、個別の通信や特定の個人が識別されるリスクを評価する。

本ユースケースにおける 1) ~ 9) の評価指標に基づいた処理方法および、1) ~ 9) を踏まえた総合評価に基づく処理方法を示す。

#### 1) 付帯情報

- 付帯情報によっては、個別の通信や特定の個人を識別する可能性が高まることに配慮して選定・加工することが望ましい。
- 本ユースケースでは、配慮すべき対象として、趣味嗜好情報が該当する。例えば、この情報に「マジックショー」のような出現頻度の低いカテゴリがある場合、それらは「観劇」のカテゴリに統合して、一般化する。

#### 2) 場所の特性

- 対象とする位置情報に、自宅、通勤・通学地が含まれる場合は、配慮して加工することが望ましい。また、対象とする位置情報に、要配慮個人情報に関わる場所が含まれている場合は、配慮して加工することが望ましい。
- 本ユースケースでは、位置情報の精度を 500 メートル四方のメッシュに変換していることから、自宅、通勤・通学先が特定される可能性は低いが、万全を期す場合、契約者情報を用いて、対象とするエリアに自宅のある個人のレコードを削除する。要配慮個人情報に関わる場所については、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアであることから、特に配慮はしない。

#### 3) 集団の規模

- 特定の学校・職場や稀少な趣味嗜好等を持つ集団を対象とした場合、集団の規模によっては、個別の通信や特定の個人を識別する可能性が高まるため、集団の規模に配慮して加工することが望ましい。
- 本ユースケースでは、個別の通信や特定の個人を識別する可能性のある集団を対象としていないため、特に配慮しない。

#### 4) 取得時期の特性

- 特定のイベントや事件のあった日、時期と一致する可能性がある場合、他の情報を参照することによって、個別の通信や特定の個人を識別する可能性が高まるため、取得時期

の特性に配慮して加工することが望ましい。

- 本ユースケースでは、対象エリアにて、配慮が必要な特定のイベントや事件等が期間内になかったなかつたとみなし、特に配慮しない。

#### 5) 位置の精度

- 高い精度の位置情報は、個別の通信や特定の個人を識別する可能性が高いため、適切に精度を低減することが望ましい。人口密度の低いエリアを対象とする場合は、特に配慮することが望ましい。
- 本ユースケースでは、位置情報を 500m 四方のメッシュとしていること、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアであることから、特に配慮はしない。

#### 6) 移動履歴の期間・範囲

- 移動履歴の期間は長くなったり、特定の時間帯を対象としたりする場合は、次の a)~c) に係るリスクが高くなるため、これらに配慮して加工することが望ましい。
  - a) パターン性
    - 定期的に通っている場所、滞留している場所が分かることにより、自宅、通勤・通学地などが推測されて、個別の通信や特定個人の識別性が高まる。
  - b) 場所の特性
    - 「2) 場所の特性」を参照。
  - c) 識別性
    - 履歴の一意性が高まる。その一意性をもって、直ちに個別の通信や特定の個人を識別することができないとしても、一定の配慮をすることが望ましい。
- 本ユースケースでは、位置情報を 500m 四方のメッシュとしていること、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアであることから、更なる加工はしない。ただし期間を長くすることで、差異の大きな個人が識別される場合は、トップコーディングを行う。また同一の事業者に提供する場合は、履歴の期間が重ならないように提供する等の配慮をし、各期間の履歴が結びつかないようにする。

#### 7) 時間の精度・間隔

- 時間の精度が高まったり、データを取得する際の時間間隔が短くなったり、個別の通信や特定の個人を識別する可能性が高まる。また、詳細な時刻情報は位置情報とセットになることで、異なるデータセット間における共通の識別子として機能し得る。このため、適切に時間の精度を低減したり、間隔を開けたりすることが望ましい。
- 本ユースケースでは、秒単位で取得された時間の精度を、15分単位に丸める。

8) 対象者数

- 加工対象とするデータセットに含まれる対象者数が少ないと、個別の通信や特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。また、同一の個人が複数台の携帯端末を所持している場合のあることを想定して、携帯端末の台数よりも対象者数が小さくなる可能性のあることに留意することが望ましい。
- 本ユースケースでは、本ユースケースでは、対象者数でカウントして、一定規模の対象者数が確保されることを確認する。

9) データ提供までの期間

- データを取得してから、「十分な匿名化」により加工した情報として提供するまでの期間が短い場合は、他の情報を参照することによって、個別の通信や特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。
- 本ユースケースでは、位置情報を 500m 四方のメッシュとしていること、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアであることから、特に配慮はしない。

<加工後のデータのイメージ>

上記の考え方に基づいて加工されたデータは、図表 1-5 のようになる。

図表 1-5 商用のユースケースにおける加工後のデータのイメージ

時期	年代	性別	住所	趣味嗜好	移動履歴	人数
2016 年 10 月 12 日 13 時～15 時	40 代	女性	埼玉県 さいたま市	観劇	5339460004 → 553946004	15 人
	60 代	女性	東京都 練馬区	観劇	5339460004 → 553946004	35 人
	50 代	男性	群馬県	スポーツ観 戦	5339460004 → 553946004	13 人
	60 代	男性	東京都 練馬区	観劇	5339460004 → 533946111	8 人
	40 代	女性	千葉県 千葉市	観劇	5339460004	75 人
...	...	...	...	...	...	...

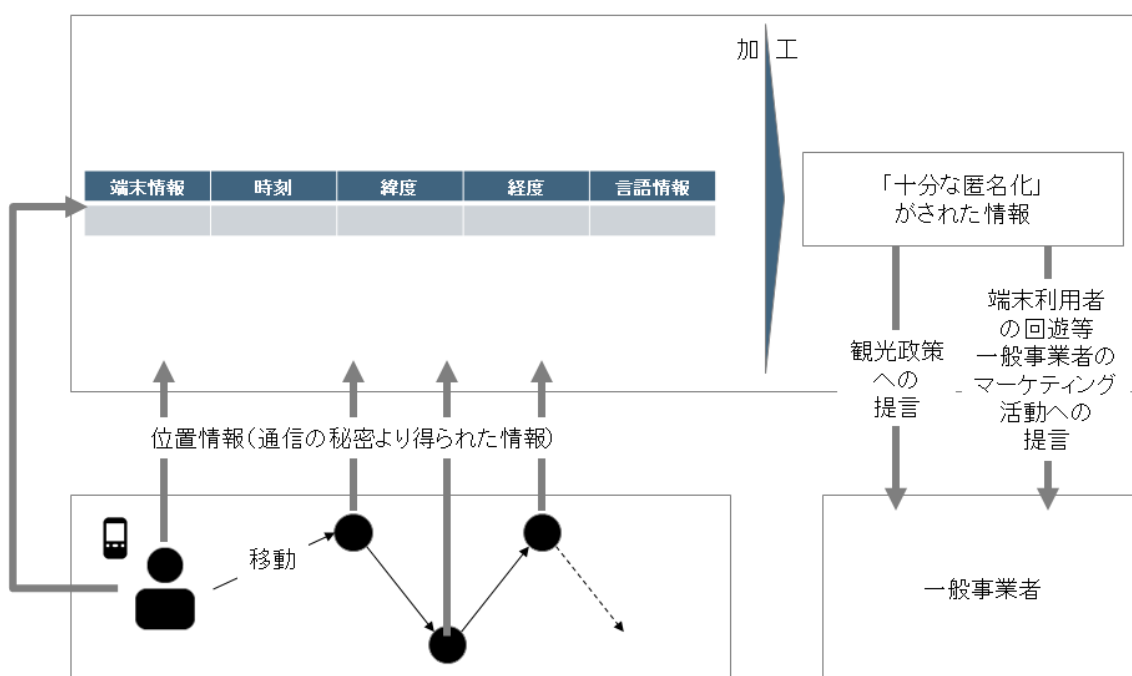
## 2. 観光の事例<sup>13</sup>

### 1) ユースケース

本ユースケースは、通信事業者が保有する通信の秘密に該当する Wi-Fi 位置情報（端末利用者がアクセスポイントから外部と通信を行うことで把握される位置情報）やその他の情報を、「十分な匿名化」の枠組みを活用して、一般事業者等へ提供するものである。

電気通信事業者においては、利用者（海外からの観光客等も含む）の動態や利用ルート等の把握をし、それを観光政策や立地戦略に資するサービス<sup>14</sup>として、第三者に提供することが想定される。

図表 2-1 通信事業者が保有する位置情報を利用するユースケースのイメージ



本ユースケースで取扱う位置情報は、外国人観光客がよく訪れる観光地を対象とし、スマートフォン等の携帯端末による、鉄道の駅や文化施設、観光スポット等に設置されたアクセスポイントとの通信で得られるものを対象とする。複数のアクセスポイントが一つの駅や施設等に設置されている場合は、匿名加工の前処理段階で、駅や施設等の単位に集約する処理を行う。

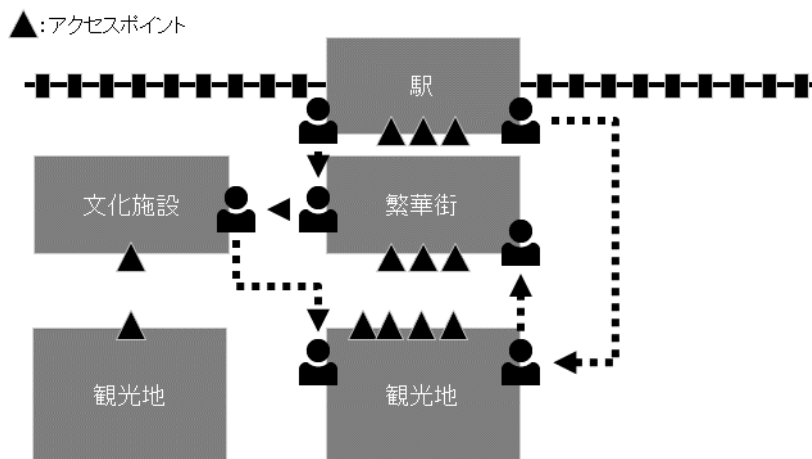
<sup>13</sup> 本事例の作成に用いたデータはすべて、加工方法の検討用に作成したダミーデータを用いている。

<sup>14</sup> 例えば、観光客が回遊するルートとそうではないルートの可視化、特定の利用者が主回遊するルートの可視化を通して、観光政策の検討や観光客取り込みに向けたマーケティング戦略の検討に資する提言をサービスとして実施することが考えられる。



付帯情報については、アクセスポイント利用時に選択する言語情報、アクセスポイントの具体的な地名・位置とを紐付ける情報から構成される。テーブルは、MAC アドレスおよびアクセスポイントによって紐付けが可能である。

図表 2-2 観光地におけるユーザーの移動イメージ



図表 2-3 位置情報に関するデータのレイアウトイメージ

利用者情報

MAC アドレス	言語
19-88-76-55-18-99	英語
26-99-88-25-55-66	中国語（簡体字）
18-66-70-09-81-33	日本語

※言語情報は、利用者がアクセスポイント接続時に選んだ言語情報を使用することを想定

位置情報

MAC アドレス	取得時刻	アクセスポイント番号
19-88-76-55-18-99	20170223145731	XXX-XXXX-XX1
19-88-76-55-18-99	20170223192651	XXX-XXXX-XX5
26-99-88-25-55-66	20170228124539	XXX-XXXX-YZ3

アクセスポイント位置情報変換テーブル

アクセスポイント番号	地点名
XXX-XXXX-XX1	東京都 東京駅
XXX-XXXX-XX5	東京都 東京駅
XXX-XXXX-YZ3	東京都 東京駅丸の内ビル

※アクセスポイントは、特定の施設等に立地されることが多いため、位置情報もそれに紐付けて整理する

※アクセスポイントは、1つの施設に複数存在することが多いが、本ユースケースでは、それらは1つの施設に紐付けて整理する

#### <移動履歴のカウント方法>

本ユースケースでは、利用者の移動履歴を、連続する複数のホップをつなげた一連のホップとしてカウントする。2重カウントを避けるため、トリップを構成するホップを、同時には移動履歴の単位としてカウントしない。

観光地への入り込み客数を延べ人数で把握するため、個人単位で集約して集計することはしない。ただし、個人の識別性への配慮のため、加工時に非識別性の評価は行う。

## 2) 考慮すべき事項とリスクに対応した具体的な加工方法の検討

#### <位置情報と付帯情報との結合>

付帯情報の条件を踏まえ、利用者情報を次の通りに加工した上で、端末識別符号をキーにて、位置情報と結合する。

- 言語情報は、一般的なカテゴリとする。マイナーなカテゴリは、他のカテゴリと統合して、一般化する。

#### <入口要件>

MACアドレスが、位置情報と付帯情報とを連結する符号に該当する。位置情報と付帯情報とを結合した後、これを削除する、又はハッシュ化等により一方向に仮ID化する。

#### <出口要件>

全てのデータ項目において適切な非識別性を確保し、あわせて、9つの評価要素によって、個別の通信や特定の個人が識別されるリスクを評価する。

本ユースケースにおける1)～9)の評価指標に基づいた処理および、1)～9)を踏まえた総合評価に基づく処理を示す。

#### 1) 付帯情報

- 付帯情報によっては、個別の通信や特定の個人を識別する可能性が高まることに配慮して選定・加工することが望ましい。
- 本ユースケースでは、配慮すべき対象として、言語情報が該当する。相対的に少ないと

想定される言語情報については、「その他の言語」のカテゴリに統合して、一般化する。  
なお本ユースケースでは、このような出現頻度の低い言語は想定しておらず、よって、  
特段処理は実施しない。

## 2) 場所の特性

- 対象とする位置情報に、自宅、通勤・通学地が含まれる場合は、配慮して加工することが望ましい。また、対象とする位置情報に、要配慮個人情報に関わる場所が含まれている場合は、配慮して加工することが望ましい。
- 本ユースケースでは、場所を地名に置き換える処理をして取り扱っていることから、本項目に係るさらなる処理は実施しない。

## 3) 集団の規模

- 特定の学校・職場や稀少な趣味嗜好等を持つ集団を対象とした場合、集団の規模によっては、個別の通信や特定の個人を識別する可能性が高まるため、集団の規模に配慮して加工することが望ましい。
- 本ユースケースでは、個別の通信や特定の個人を識別する可能性のある集団を対象としていないため、特に配慮しない。

## 4) 取得時期の特性

- 特定のイベントや事件のあった日、時期と一致する可能性がある場合、他の情報を参照することによって、個別の通信や特定の個人を識別する可能性が高まるため、取得時期の特性に配慮して加工することが望ましい。
- 本ユースケースでは、対象エリアにて、配慮が必要な特定のイベントや事件等が期間内になかったとみなし、特に配慮しない。

## 5) 位置の精度

- 高い精度の位置情報は、個別の通信や特定の個人を識別する可能性が高いため、適切に精度を低減することが望ましい。人口密度の低いエリアを対象とする場合は、特に配慮することが望ましい。
- 本ユースケースでは、アクセスポイントを地名に置き換えること、外国人観光客がよく訪れる観光地を対象としていることから、特に配慮はしない。

## 6) 移動履歴の期間・範囲

- 移動履歴の期間は長くなったり、特定の時間帯を対象としたりする場合は、次の a)~c) に係るリスクが高くなるため、これらに配慮して加工することが望ましい。
  - a) パターン性

➤ 定期的に通っている場所、滞留している場所が分かることにより、自宅、通勤・通学地などが推測されて、個別の通信や特定個人の識別性が高まる。

b) 場所の特性

➤ 「2) 場所の特性」を参照。

c) 識別性

➤ 履歴の一意性が高まる。その一意性をもって、直ちに個別の通信や特定の個人を識別することができないとしても、一定の配慮をすることが望ましい。

- 本ユースケースでは、アクセスポイントを地名に置き換えること、外国人観光客がよく訪れる観光地を対象としていることから、更なる加工はしない。ただし、期間を長くすることで、差異の大きな個人が識別される場合は、トップコーディングを行う。また同一の事業者を提供する場合は、履歴の期間が重ならないように提供する等の配慮をし、各期間の履歴が結びつかないようにする。

7) 時間の精度・間隔

- 時間の精度が高まったり、データを取得する際の時間間隔が短くなったりすると、個別の通信や特定の個人を識別する可能性が高まる。また、詳細な時刻情報は位置情報とセットになることで、異なるデータセット間における共通の識別子として機能し得る。このため、適切に時間の精度を低減したり、間隔を開けたりすることが望ましい。
- 本ユースケースでは、一定期間内における利用者による地点間の移動を把握することを目的とするため、時間の精度・間隔に係るデータ項目は、加工後のデータセットには含まれない。このため、特に配慮はしない。

8) 対象者数

- 加工対象とするデータセットに含まれる対象者数が少ないと、個別の通信や特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。また、同一の個人が複数台の携帯端末を所持している場合のあることを想定して、携帯端末の台数よりも対象者数が小さくなる可能性のあることに留意することが望ましい。
- 本ユースケースでは、本ユースケースでは、対象者数でカウントして、一定規模の対象者数が確保されることを確認する。

9) データ提供までの期間

- データを取得してから、「十分な匿名化」により加工した情報として提供するまでの期間が短い場合は、他の情報を参照することによって、個別の通信や特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。
- 本ユースケースでは、アクセスポイントを地名に置き換えること、外国人観光客がよく訪れる観光地を対象としていること、時間の精度・間隔は集計の過程で削除しているこ

とから、特に配慮はしない。

<加工後のデータのイメージ>

上記の考え方に基づいて加工されたデータは、図表 2-4 のようになる。

図表 2-4 観光のユースケースにおける加工後のデータのイメージ

時期	言語	移動履歴	人数
2016 年 10 月	英語	東京駅→新宿駅→横浜駅→箱根駅	35 人
	英語	有楽町駅→皇居→歌舞伎座→有楽町駅	13 人
	英語	東京駅→浅草駅→雷門→両国国技館→東京駅	10 人
	中国語	東京駅→銀座→舞浜駅→ディズニーランド→・・・	120 人
・・・	・・・	・・・	・・・

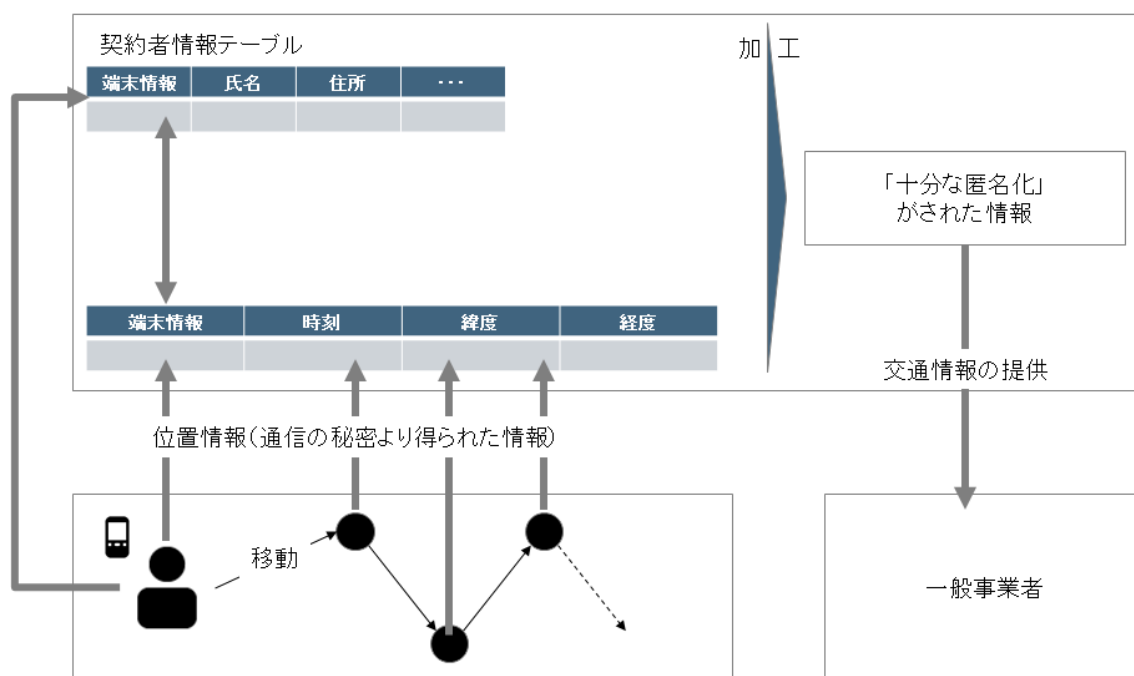
### 3. 交通の事例<sup>15</sup>

#### 1) ユースケース

本ユースケースは、電気通信事業者が保有する通信の秘密に関する位置情報（個々の通信の際に把握される基地局の位置情報）や契約者情報等を、「十分な匿名化」の枠組みを活用して、一般事業者へ提供するものである。

電気通信事業者においては、交通渋滞情報の提供や、路線の交通量の逼迫情報を分析したりするサービスに活用することが想定される。

図表 3-1 電気通信事業者が保有する位置情報を利用するユースケースのイメージ



本ユースケースは、出発地点または到着地点を羽田空港あるいは成田空港に設定して、両空港間の移動動態に係る分析データの提供を対象とする。

スマートフォン等携帯端末の基地局との通信で得られる位置情報と、端末契約時の契約者情報から構成される、図表 3-2 のようなデータ構造を前提として検討する。全てのテーブルは、端末識別符号によって紐付けが可能である。

<sup>15</sup> 本事例は、平成 27 年度に実施した調査研究におけるユースケースを参考に作成している。

図表 3-2 位置情報に関するデータのレイアウトイメージ

契約者情報

端末識別符号	氏名	性別	生年月日	住所
008542314-5	田中 一郎	男	1972 年 4 月 4 日	東京都千代田区丸の内 6-5-3
088539881-2	佐藤 幸子	女	1993 年 12 月 9 日	東京都西東京市保谷町 2-8-18
063381312-1	鈴木 博	男	1938 年 8 月 23 日	神奈川県横浜市旭区二俣川 236

位置情報

端末識別符号	時刻	緯度	経度
008542314-5	20161128132415	35.658593	139.745441
008542314-5	20161128150212	35.658593	139.745992
028809315-1	20161130020249	38.688721	140.663542

※時刻は年月日時刻（秒単位）まで収集可能と想定

<移動履歴のカウント方法>

本ユースケースでは、利用者の移動履歴を、出発駅、経由地、到着地の3地点間のホップとトリップでカウントする。

交通量を把握するため、個人単位で集約して集計することはない。ただし、個人の識別性への配慮のため、加工時に非識別性の評価は行う。

2) 考慮すべき事項とリスクに対応した具体的な加工方法の検討

<位置情報と付帯情報との結合>

付帯情報の条件を踏まえ、契約者情報、趣味嗜好情報を次の通りに加工した上で、端末識別符号をキーにて、位置情報と結合する。

氏名は、削除する。

- 性別は、加工せず、そのまま用いる。
- 生年月日は、20歳未満、20代、30代、40代、50代、60代、70歳以上の7つの年代に置き換える。(丸め)
- 住所は県と市区町村のレベルまでに限定

<入口要件>

端末識別符号が、位置情報と付帯情報とを連結する符号に該当する。位置情報と付帯情報とを結合した後、これを削除する、又はハッシュ化等により一方向に仮ID化をする。

#### <出口要件>

全てのデータ項目において適切な非識別性を確保し、あわせて、9つの評価要素によって、個別の通信や特定の個人が識別されるリスクを評価する。

本ユースケースにおける1)～9)の評価指標に基づいた処理方法および、1)～9)を踏まえた総合評価に基づく処理方法を示す。

##### 1) 付帯情報

- 付帯情報によっては、個別の通信や個人を特定する可能性が高まることに配慮して選定・加工することが望ましい。
- 本ユースケースでは、特にこれに該当する情報は見受けられないため、本項目に係る処理は実施しない。

##### 2) 場所の特性

- 対象とする位置情報に、自宅、通勤・通学地が含まれる場合は、配慮して加工することが望ましい。また、対象とする位置情報に、要配慮個人情報に関わる場所が含まれている場合は、配慮して加工することが望ましい。
- 本ユースケースでは、契約者情報を用いて、対象とするエリアに自宅のある個人のレコードを削除する。要配慮個人情報に関わる場所については、空港、駅および高速道路のジャンクションといった公共性の高い場所を対象としていることから、特に配慮しない。

##### 3) 集団の規模

- 特定の学校・職場や稀少な趣味嗜好等を持つ集団を対象とした場合、集団の規模によっては、個別の通信や特定の個人を識別する可能性が高まるため、集団の規模に配慮して加工することが望ましい。
- 本ユースケースでは、個別の通信や特定の個人を識別する可能性のある集団を対象としていないため、特に配慮しない。

##### 4) 取得時期の特性

- 特定のイベントや事件のあった日、時期と一致する可能性がある場合、他の情報を参照することによって、個別の通信や特定の個人を識別する可能性が高まるため、取得時期の特性に配慮して加工することが望ましい。
- 本ユースケースでは、対象エリアにて、配慮が必要な特定のイベントや事件等が期間内になかったとみなし、特に配慮しない。



#### 5) 位置の精度

- 高い精度の位置情報は、個別の通信や特定の個人を識別する可能性が高いため、適切に精度を低減することが望ましい。人口密度の低いエリアを対象とする場合は、特に配慮することが望ましい。
- 本ユースケースでは、緯度・経度情報を、空港、駅および高速道路のジャンクション等の単位に変換する。

#### 6) 移動履歴の期間・範囲

- 移動履歴の期間は長くなったり、特定の時間帯を対象としたりする場合は、次の a)~c) に係るリスクが高くなるため、これらに配慮して加工することが望ましい。
  - a) パターン性
    - 定期的に通っている場所、滞留している場所が分かることにより、自宅、通勤・通学地などが推測されて、個別の通信や特定個人の識別性が高まる。
  - b) 場所の特性
    - 「2) 場所の特性」を参照。
  - c) 識別性
    - 履歴の一意性が高まる。その一意性をもって、直ちに個別の通信や特定の個人を識別することができないとしても、一定の配慮をすることが望ましい。
- a)~c)を踏まえ、本ユースケースでは、
  - ・ 自宅や通勤・通学等の生活圏に係るレコードを除外する。
  - ・ 仮 ID の振り方を 1 日単位で変えることで、識別性を低減する。といった処理をする。また同一の事業者を提供する場合は、履歴の期間が重ならないように提供する等の配慮をし、各期間の履歴が結びつかないようにする。

#### 7) 時間の精度・間隔

- 時間の精度が高まったり、データを取得する際の時間間隔が短くなったりすると、個別の通信や特定の個人を識別する可能性が高まる。また、詳細な時刻情報は位置情報とセットになることで、異なるデータセット間における共通の識別子として機能し得る。このため、適切に時間の精度を低減したり、間隔を開けたりすることが望ましい。
- 本ユースケースでは、一定期間内における利用者による地点間の移動を把握することを目的とするため、時間の精度・間隔に係るデータ項目は、加工後のデータセットには含まれない。このため、特に配慮はしない。

#### 8) 対象者数

- 加工対象とするデータセットに含まれる対象者数が少ないと、個別の通信や特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。また、同一の個人が複数台の携帯端末を所持している場合のあることを想定して、携帯端末の台数よりも対象者数が小さくなる可能性のあることに留意することが望ましい。
- 本ユースケースでは、本ユースケースでは、対象者数でカウントして、一定規模の対象者数が確保されることを確認する。

9) データ提供までの期間

- データを取得してから、「十分な匿名化」により加工した情報として提供するまでの期間が短い場合は、他の情報を参照することによって、個別の通信や特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。
- 本ユースケースでは、場所が、空港、駅および高速道路のジャンクションといった公共性の高い場所を対象としていること、時間の精度・間隔は集計の過程で削除していることから、特に配慮はしない。

<加工後のデータのイメージ>

上記の考え方に基づいて加工されたデータは、図表 3-4 のようになる。

図表 3-4 交通のユースケースにおける加工後のデータのイメージ

時期	年代	性別	出発駅	到着地	経由地	人数
2016 年 10 月平日	40 代	女性	新宿駅	羽田空港	浜松町駅	50 人
	60 代	女性	新宿駅	羽田空港	大井 JCT	75 人
	40 代	男性	浜松町駅	羽田空港	-	25 人
	60 代	男性	高田馬場駅	成田空港	大井 JCT	8 人
2016 年 10 月 休日祝日	40 代	女性	新宿駅	羽田空港	大井 JCT	100 人
	40 代	女性	池袋駅	羽田空港	浜松町駅	100 人
	20 代	女性	新宿駅	成田空港	大井 JCT	40 人
	50 代	男性	浜松町駅	羽田空港	-	14 人
...	...	...	...	...	...	...

## 2) 電気通信事業者が扱う位置情報を用いた匿名加工情報の作成等に関するルールに盛り込まれるべき内容

### 電気通信事業者が扱う位置情報を用いた匿名加工情報の作成等 に関するルールに盛り込まれるべき内容

#### 1. 趣旨

本資料は、電気通信事業における個人情報の保護を踏まえ、電気通信事業者が扱う位置情報を、個人情報保護法で定める匿名加工情報として活用するための方法に関するルールをとりまとめたものである。本資料は、認定個人情報保護団体に引き継がれて、個人情報保護指針として定められることを想定している。

加工の要求レベルは、電気通信事業者が扱う位置情報の精度やユースケースによって判断されるべきもので、普遍的な基準を定めることは困難である。このため、本資料では、ケースバイケースで匿名加工情報の要求レベルについて判断することを前提に、最低限遵守すべき基本ルールを定めるとともに、別添として具体的なユースケースとその対応方法をとりまとめている。

作成方法以外の匿名加工情報の取扱いに係る規律は、基本的に、個人情報保護法ガイドライン(匿名加工情報編)及び個人情報保護委員会事務局レポート<sup>16</sup>を参照することを想定し、本資料では対象としない。

なお、通信の秘密に該当する位置情報を利用する場合は、電気通信事業における個人情報保護に関するガイドライン(以下「電気通信事業分野ガイドライン」という。)第35条第2項及び同項の解説も参照する必要がある。

#### 2. 用語の定義

##### 2.1. 位置情報

- ・ 電気通信事業分野ガイドライン第35条1項に規定する位置情報をいう。携帯電話の基地局に係る位置情報、GPSによる位置情報、Wi-Fiアクセスポイントに係る位置情報等がある。

##### 2.2. 付帯情報

- ・ 電気通信事業者が扱う位置情報に付帯する情報。電気通信事業者が保有する契約者に係る情報、電気通信サービスを利用する際に本人が登録する情報等がある。

---

<sup>16</sup> 個人情報保護委員会事務局「個人情報保護委員会事務局レポート：匿名加工情報 パーソナルデータの利活用促進と消費者の信頼性確保の両立に向けて」(2017年2月)

## 2.3. オプトアウト

- ・ 本人の明示的な同意を事前に取得せずに、本人の情報を利用し、本人からの求めに応じてその情報の利用を停止する方式。ただし、オプトアウトを採用する場合は、情報の利用目的を事前に通知・公表するとともに、オプトアウトの実質的な機会を確保することが前提となる。

## 3. 適用範囲

- ・ 電気通信事業者が扱う位置情報を用いた匿名加工情報の作成方法に関する基準

## 4. 匿名加工情報の作成方法の基準

### 4.1. 個人情報保護法 36 条 1 項関連に係る規定の遵守

- ・ 個人情報保護法 36 条 1 項に基づく個人情報保護法施行規則 19 条各号とそれらに対応する個人情報保護法ガイドライン（匿名加工情報編）の記述内容を遵守すること。

#### 【規則 19 条 1 号】特定の個人を識別することができる記述等の削除

- ・ 個人情報に含まれる特定の個人を識別することができる記述等の全部又は一部を削除すること（当該全部又は一部の記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）。

事例 1) 氏名、住所、生年月日が含まれる個人情報を加工する場合に次の 1 から 3 までの措置を講ずる。

- 1) 氏名を削除する。
- 2) 住所を削除する。又は、〇〇県△△市に置き換える。
- 3) 生年月日を削除する。又は、日を削除し、生年月に置き換える。

事例 2) 会員 ID、氏名、住所、電話番号が含まれる個人情報を加工する場合に次の 1、2 の措置を講ずる。

- 1) 会員 ID、氏名、電話番号を削除する。
- 2) 住所を削除する。又は、〇〇県△△市に置き換える。

#### 【規則 19 条 2 号】個人識別符号の削除

- ・ 個人情報に含まれる個人識別符号の全部を削除すること（当該個人識別符号を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）。

#### 【規則 19 条 3 号】情報を相互に連結する符号の削除

- ・ 個人情報と当該個人情報に措置を講じて得られる情報を連結する符号（現に個人情報取扱事業者において取り扱う情報を相互に連結する符号に限る。）を削除すること

(当該符号を復元することのできる規則性を有しない方法により当該個人情報と当該個人情報に措置を講じて得られる情報を連結することができない符号に置き換えることを含む。)

事例 1) サービス会員の情報について、氏名等の基本的な情報と購買履歴を分散管理し、それらを管理用 ID を付すことにより連結している場合、その管理用 ID を削除する。

事例 2) 委託先へ個人情報の一部を提供する際に利用するために、管理用 ID を付すことにより元の個人情報と提供用に作成した情報を連結している場合、当該管理用 ID を仮 ID に置き換える。

**【規則 19 条 4 号】 特異な記述等の削除**

- ・ 特異な記述等を削除すること（当該特異な記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）

事例 1) 症例数の極めて少ない病歴を削除する。

事例 2) 年齢が「116 歳」という情報を「90 歳以上」に置き換える。

**【規則 19 条 5 号関連】 個人情報データベース等の性質を踏まえたその他の措置**

- ・ 前各号に掲げる措置のほか、個人情報に含まれる記述等と当該個人情報を含む個人情報データベース等を構成する他の個人情報に含まれる記述等との差異その他の当該個人情報データベース等の性質を勘案し、その結果を踏まえて適切な措置を講ずること。

事例 1) 移動履歴を含む個人情報データベース等を加工の対象とする場合において、自宅や職場などの所在が推定できる位置情報（経度・緯度情報）が含まれており、特定の個人の識別又は元の個人情報の復元につながるおそれがある場合に、推定につながり得る所定範囲の位置情報を削除する。**(項目削除／レコード削除／セル削除)**

事例 2) ある小売店の購買履歴を含む個人情報データベース等を加工の対象とする場合において、当該小売店で購入者が極めて限定されている商品の購買履歴が含まれており、特定の個人の識別又は元の個人情報の復元につながるおそれがある場合に、具体的な商品情報（品番・色）を一般的な商品カテゴリに置き換える。**(一般化)**

事例 3) 小学校の身体検査の情報を含む個人情報データベース等を加工の対象とする場合において、ある児童の身長が 170 cm という他の児童と比べて差異が大きい情報があり、特定の個人の識別又は元の個人情報の復元につながるおそれがある場合に、当該身長情報を一般的な身長範囲に置き換える。**(範囲化)**

る場合に、身長が 150cm 以上の情報について「150 cm以上」という情報に置き換える。(トップコーディング)

#### 4.2. 位置情報に特有の個人情報データベース等の性質を踏まえたその他の措置

- ・ 個人情報保護法施行規則 19 条 5 号に関して、位置情報特有の性質へ対処するため、次の 1) ～ 9) に掲げる評価要素による特定の個人が識別されるリスクを定性的に評価し、総合的に判断して、同リスクが十分に低減するよう努めなければならない。なお必ずしも 1) ～ 9) の全ての評価要素を満たす加工を求めるものではない。

##### 10) 付帯情報

- ・ 付帯情報によっては、個人を特定する可能性が高まることに配慮して選定・加工することが望ましい。

(対処の例)

- 性別であっても、対象とする集団に男女の大きな偏りのあることが想定される場合は、付帯情報として用いない又は配慮して加工する。

##### 11) 場所の特性

- ・ 対象とする位置情報に、自宅、通勤・通学地が含まれる場合は、配慮して加工することが望ましい。
- ・ 対象とする位置情報に、要配慮個人情報に関わる場所が含まれている場合は、配慮して加工することが望ましい。

(対処の例)

- 明らかに自宅、通勤・通学先がわかる場合は、これらを除くことが望ましい。
- 特定の疾患を対象とする病院に滞留していることが明らかなレコードを、加工対象から除外する。

##### 12) 集団の規模

- ・ 特定の学校・職場や稀少な趣味嗜好等を持つ集団を対象とした場合、集団の規模によっては、特定の個人を識別する可能性が高まるため、集団の規模に配慮して加工することが望ましい。

(対処の例)

- 特定の趣味嗜好の集団を取り扱う場合、十分な対象者数が得られることを確認する。

### 1 3) 取得時期の特性

- ・ 特定のイベントや事件のあった日、時期と一致する可能性がある場合、他の情報を参照することによって、特定の個人を識別する可能性が高まるため、取得時期の特性に配慮して加工することが望ましい。

(対処の例)

- 大規模集客施設において、特定の宗教のイベントが開催されていることが明らかでない場合は、当該期間及び当該施設に該当するレコードを、加工対象から除外する。

### 1 4) 位置の精度

- ・ 高い位置精度の情報は、特定の個人を識別する可能性が高いため、適切に精度を低減することが望ましい。人口密度の低いエリアを対象とする場合は、特に配慮することが望ましい。

(対処の例)

- 位置精度数メートルの緯度・経度情報を、適切な大きさのメッシュ単位の位置情報に変換する。

### 1 5) 移動履歴の期間・範囲

- ・ 移動履歴の期間は長くなったり、特定の時間帯を対象としたりする場合は、次の a)～c)に係るリスクが高くなるため、これらに配慮して加工することが望ましい。

#### d) パターン性

- 定期的に通っている場所、滞留している場所が分かることにより、自宅、通勤・通学地などが推測されて、特定の個人の識別性が高まる。

#### e) 場所の特性

- 「2) 場所の特性」を参照。

#### f) 識別性

- 履歴の一意性が高まる。その一意性をもって、直ちに個人を特定することができないとしても、一定の配慮をすることが望ましい。

(対処の例)

- a)～c)を踏まえ、移動履歴の期間を短くして提供する。同一の事業者に提供する場合は、履歴の期間が重ならないように提供する等の配慮をし、各期間の履歴が結びつかないようにする。

### 1 6) 時間の精度・間隔

- ・ 時間の精度が高まったり、データを取得する際の時間間隔が短くなったりすると、特定の個人を識別する可能性が高まる。また、詳細な時刻情報は位置情報とセットにな

ることで、異なるデータセット間における共通の識別子として機能し得る。このため、適切に時間の精度を低減したり、間隔を開けたりすることが望ましい。

(対処の例)

- 秒単位で取得された時間の精度を、15分単位にまるめる。

#### 17) 対象者数

- ・ 加工対象とするデータセットに含まれる対象者数が少ないと、特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。なお対象者数の数え方は、付属資料2を踏まえること。
- ・ 同一の個人が複数台の携帯端末を所持している場合のあることを想定して、携帯端末台数よりも対象者数が小さくなる可能性のあることに留意することが望ましい。

(対処の例)

- データを対象者数でカウントして、適切な規模の対象者数を確保する。

#### 18) データ提供までの期間

- ・ データを取得してから、匿名加工情報として提供するまでの期間が短い場合は、他の情報を参照することによって、特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。

(対処の例)

- 位置情報を取得してから匿名加工情報として提供するまでの期間を、三ヶ月以上確保する。

### 5. オプトアウト

- ・ オプトアウトの手段を提供し、本人からの申し出に応じて、匿名加工情報へのデータ利用を停止できるようにすることが望ましい。
- ・ オプトアウトの手段は、本人が、ウェブサイト、電話等により容易にオプトアウトすることができるように努めなければならない。

### 6. ルールの見直し

- ・ 電気通信事業者が扱う位置情報を用いた匿名加工情報の作成等に関するルールは、制度改正や技術の進展を踏まえ、定期的に見直しを行って、所要の措置を講じなければならない。

### 付属資料

- ・ 付属1：「十分な匿名化」に関するルールに盛り込まれるべき内容
- ・ 付属2：移動履歴のカウント方法



- ・ 付属 3 : 電気通信事業者が扱う位置情報を用いた匿名加工情報の加工の事例
  - 商用の事例
  - 観光の事例

## 付属 1 : 「十分な匿名化」に関するルール

(略)

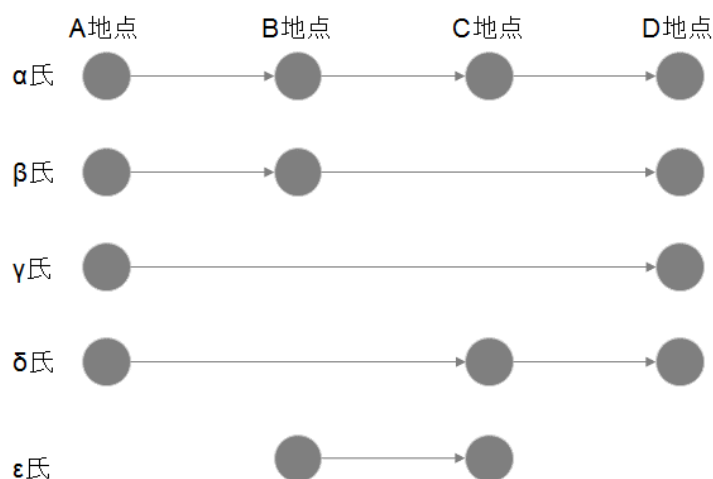
## 付属 2 : 移動履歴のカウント方法

ある人物の位置情報を複数つなげると、その人の移動履歴が構成される。このとき、移動履歴の数を明確にしておかなければ、分析結果を利用する側にとって混乱が生じたり、特定の個人の識別リスクが高まったりするおそれがある。このため、ここでは、移動履歴のカウント方法とその適用上の留意事項について整理する。

### ■移動履歴のカウント方法

移動履歴のカウント方法は、「図 1 移動履歴のカウントに係るサンプルモデル」を用いて、「ホップ」と「トリップ」という用語を定義して説明する。本モデルにおいて、●を通過地点（あるいは滞留地点）、時間軸を左から右に経過するものとし、同一時間軸上の地点は時刻が同一もしくは同一とみなすものとする。

図 1 移動履歴のカウントに係るサンプルモデル



#### ホップ

移動履歴の最小単位。2 地点間の移動を 1 ホップ、3 地点間の移動を 2 ホップと数える。1 地点に滞留しているだけの場合は 0 ホップとなる。

図 1 のモデルで考えると、 $\alpha$  氏は、A 地点から D 地点までの移動にあたって、A→B、B→C、C→D の 3 ホップをしていると数える。一方、 $\beta$  氏は、A 地点から D 地点まで移動にあたって、A→B、B→D の 2 ホップをしていると数える。

#### トリップ

移動履歴のカウント単位。連続する複数のホップをつなげた”一連のホップ”を分析対象とする場合は、これを 1 トリップとしてカウントする。1 ホップのみであっても、これを移動

履歴として分析対象とする場合は、1トリップとしてカウントする。

図1のモデルで考えると、 $\alpha$ 氏において $A \rightarrow B$ 、 $B \rightarrow C$ 、 $C \rightarrow D$ の一連のホップを分析対象とする場合は、 $A \rightarrow B \rightarrow C \rightarrow D$ という一つのトリップとしてカウントする。このとき、図1では、 $A \rightarrow B \rightarrow C \rightarrow D$ のトリップに該当するのは、 $\alpha$ 氏のみである。

また $\alpha$ 氏において、 $A \rightarrow B$ 、 $B \rightarrow C$ 、 $C \rightarrow D$ のホップそれぞれを分析対象とする場合は、各ホップを独立のトリップとしてカウントする。このとき、図1では、 $A \rightarrow B$ のトリップに該当するのは $\alpha$ 氏と $\beta$ 氏の2名、 $B \rightarrow C$ のトリップに該当するのは $\alpha$ 氏と $\varepsilon$ 氏の2名、 $C \rightarrow D$ のトリップに該当するのは $\alpha$ 氏と $\delta$ 氏の2名となる。

### ■移動履歴をカウントする際の留意事項

次の1～3に留意して、移動履歴をカウントすることが有効である。

#### 1. 対象とするトリップの分け方を明確にする。

移動履歴として、どのホップまたは一連のホップをトリップの対象とするのかを特定する。なお、同一データセットに異なるトリップの分け方を適用して異なる匿名化データを作成することには注意が必要である。たとえば、同一のデータセットから作成した2種類の匿名化データにおいて、一方の匿名化データで $A \rightarrow B \rightarrow C$ が5件、もう一方の匿名化データで $B \rightarrow C \rightarrow D$ が6件あった場合、 $A \rightarrow B \rightarrow C \rightarrow D$ と移動した履歴が1件であることがわかってしまう場合があるからである。

#### 2. データセットにおいて、同時に、ホップが重なり合う異なるトリップを集計しない。

図1のモデルで考えると、 $A \rightarrow B \rightarrow C \rightarrow D$ のトリップを集計する場合は、同時に、 $B \rightarrow C$ の集計を行ってはならない。同一人物の移動を2重にカウントすることになるからである。

#### 3. リピータ（同一人物の同一トリップ）へ配慮する。

同一人物が同一のトリップを繰り返す場合がある。通勤、通学、通院のような日常的な移動に加え、観光やレジャーなどの不定期の移動も、同一人物による同一トリップに該当する可能性がある。

リピータのトリップを分析対象とする場合は、個人単位でトリップを集約するなどの配慮が必要である。

一方、観光地への入込客を延べ人数で集計したい場合は、個人単位で集約して集計することはしないが、個人の識別性への配慮は必要である。たとえば、対象とするトリップが100件あったとしても、すべて同一人物によるものである場合は、識別リスクが高まるため、個人単位で集計して非識別性を評価する必要がある。

### 付属 3：電気通信事業者が扱う位置情報を用いた匿名加工情報の加工の事例

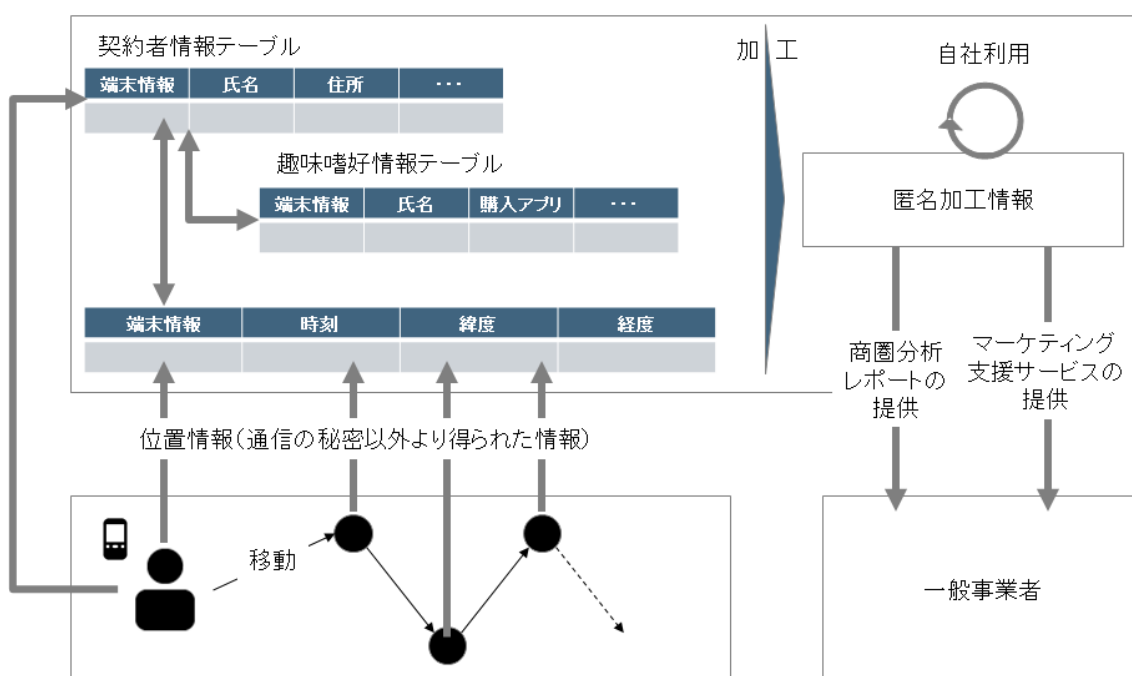
#### 1. 商用の事例<sup>17</sup>

##### 1) ユースケース

本ユースケースは、電気通信事業者が保有する通信の秘密に該当しない位置情報や契約者情報等を匿名加工したうえで、匿名加工情報の枠組みを活用して、自社で利用したり、一般事業者へ提供したりするものである。

電気通信事業者や一般事業者においては、店舗の商圈分析<sup>18</sup>や、マーケティング支援サービス等の販売促進活動に活用することが想定される。

図表 1-1 電気通信事業者が保有する位置情報を利用するユースケースのイメージ

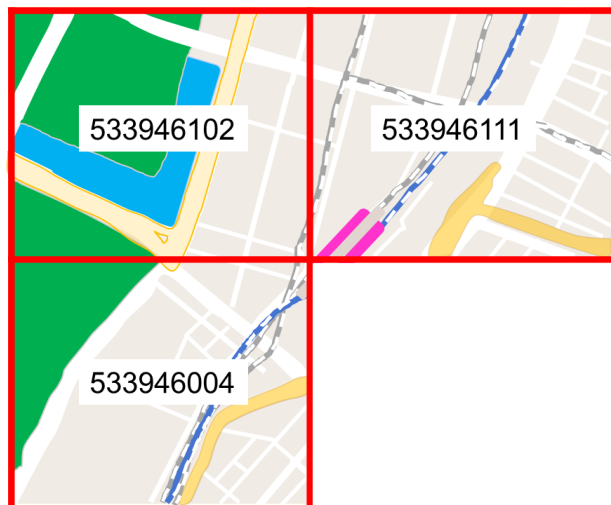


本ユースケースで取扱う位置情報は、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアを対象とし、スマートフォン等の携帯端末の基地局との通信で得られる位置情報を、匿名加工の前処理段階で、1/2 メッシュコード (500m 四方のメッシュコード) に変換して用いる。

<sup>17</sup> 本事例の作成に用いたデータはすべて、加工方法の検討用に作成したダミーデータを用いている。

<sup>18</sup> 例えば、新店舗の立地条件を考える際に、ターゲットとしたい顧客層の移動状況を分析することが考えられる。

図表 1-2 対象エリアにおけるメッシュのイメージ (数値はメッシュコードを表す)



付帯情報については、端末契約時の契約者情報、及び電気通信事業者が提供するアプリ購入時に、契約者から同意を得て取得するアプリ等に基づく趣味嗜好情報から構成される。全てのテーブルは、端末識別番号によって紐付けが可能である。

図表 1-3 位置情報に関するデータのレイアウトイメージ

契約者情報

端末識別番号	氏名	性別	生年月日	住所
008542314-5	田中 一郎	男	1972 年 4 月 4 日	東京都千代田区丸の内 6-5-3
088539881-2	佐藤 幸子	女	1993 年 12 月 9 日	東京都西東京市保谷町 2-8-18
063381312-1	鈴木 博	男	1938 年 8 月 23 日	神奈川県横浜市旭区二俣 236

趣味嗜好情報

端末識別番号	趣味嗜好
008542314-5	オペラ観劇
028809315-1	野球観戦
088539881-2	マジックショー

位置情報

端末識別番号	日時	メッシュコード
008542314-5	20161128132415	533946004
008542314-5	20161128150212	553946004
028809315-1	20161130020249	533946111

## 2) 考慮すべき事項とリスクに対応した具体的な加工方法の検討

電気通信事業者は、位置情報と付帯情報とを結合した情報を、個人情報保護法 36 条 1 項に基づく個人情報保護法施行規則 19 条各号とそれらに対応する個人情報保護法ガイドライン（匿名加工情報編）を遵守し、匿名加工情報を作成する。この際、位置情報に特有の個人情報データベース等の性質を踏まえたその他の措置を実施することに努めなければならない。

以下に、具体的な加工方法を示す。

<個人情報保護法 36 条 1 項関連に係る規定の遵守に係る加工方法>

### ①. 【規則 19 条 1 号】 特定の個人を識別することができる記述等の削除

- 1 号では、「個人情報に含まれる特定の個人を識別することができる記述等の全部又は一部を削除すること（当該全部又は一部の記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）」が求められている。
  - 本ユースケースでは、個人を識別することのできる情報として、氏名、住所、生年月日が挙げられる。それぞれの情報については、
    - ・ 氏名は削除
    - ・ 住所は県と市区郡のレベルまでに限定
    - ・ 生年月日は 20 歳未満、20 代、30 代、40 代、50 代、60 代、70 歳以上の 7 つの年代に置き換える。（丸め）といった処理をすることが望ましい。
- なお、商圈分析では、電車移動における乗車駅等を考えるニーズもあることから、契約者の住所を最寄り駅単位で一般化する処理方法も考えられる。

### ②. 【規則 19 条 2 号】 個人識別符号の削除

- 2 号では、「個人情報に含まれる個人識別符号の全部を削除すること（当該個人識別符号を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）」が求められている。
- 本ユースケースでは、当該識別符号はないため、本規則に則った加工は実施しない。

### ③. 【規則 19 条 3 号】 情報を相互に連結する符号の削除

- 3 号では、「個人情報と当該個人情報に措置を講じて得られる情報を連結する符号（現に個人情報取扱事業者において取り扱う情報を相互に連結する符号に限る。）を削除すること（当該符号を復元することのできる規則性を有しない方法により当該個人情報と当該個人情報に措置を講じて得られる情報を連結することができない符号に置き換えることを含む。）」が求められている。
- 本ユースケースでは、端末識別符号が該当する。端末識別符号を用いて、位置情報や趣

味嗜好情報を連結している。よって、端末識別符号は、ハッシュ化等により一方向に仮ID化する。

#### ④. 【規則 19 条 4 号】 特異な記述等の削除

- 4号では、「特異な記述等を削除すること（当該特異な記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）」が求められている。
- 本ユースケースでは、当該記述はないため、本規則に則った加工は実施しない。

#### ⑤. 【規則 19 条 5 号関連】 個人情報データベース等の性質を踏まえたその他の措置

- 5号では、「前各号に掲げる措置のほか、個人情報に含まれる記述等と当該個人情報を含む個人情報データベース等を構成する他の個人情報に含まれる記述等との差異その他の当該個人情報データベース等の性質を勘案し、その結果を踏まえて適切な措置を講ずること。」が求められている。
- 本ユースケースでは、趣味嗜好情報が該当する。趣味嗜好は多種多様に存在し、スポーツ観戦のようなメジャーなものからマイナーなものまで多岐に渡る。例えば「マジックショー」を趣味とする人が、特定の日時に特定の場所に出現する場合は、規則 19 条 5 号の個人情報に含まれる記述等と当該個人情報を含む個人情報データベース等を構成する他の個人情報に含まれる記述等との差異に該当すると考えられる。このような出現頻度の低いカテゴリは、例えば「観劇」のカテゴリに統合して一般化する。
- また、位置情報特有の性質へ対処するため、次の 1)～9) に掲げる評価要素による特定の個人が識別されるリスクを定性的に評価し、総合的に判断して、同リスクが十分に低減するための措置を実施することに努めなければならない。なお必ずしも 1)～9) の全ての評価要素を満たす加工を求めるものではない。

<位置情報に特有の個人情報データベース等の性質を踏まえたその他の措置>

##### 1) 付帯情報

- 付帯情報によっては、個人を特定する可能性が高まることに配慮して選定・加工することが望ましい。
- 本ユースケースでは、特にこれに該当する情報は見受けられないため、本項目に係る処理は実施しない。

##### 2) 場所の特性

- 対象とする位置情報に、自宅、通勤・通学地が含まれる場合は、配慮して加工することが望ましい。また、対象とする位置情報に、要配慮個人情報に関わる場所が含まれている場合は、配慮して加工することが望ましい。



- 本ユースケースでは、位置情報の精度を 500 メートル四方のメッシュに変換していることから、自宅、通勤・通学先が特定される可能性は低いですが、万全を期す場合、契約者情報を用いて、対象とするエリアに自宅のある個人のレコードを削除する。要配慮個人情報に関わる場所については、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアであることから、特に配慮はしない。

### 3) 集団の規模

- 特定の学校・職場や稀少な趣味嗜好等を持つ集団を対象とした場合、集団の規模によっては、特定の個人を識別する可能性が高まるため、集団の規模に配慮して加工することが望ましい。
- 本ユースケースでは、特定の個人を識別する可能性のある集団を対象としていないため、特に配慮しない。

### 4) 取得時期の特性

- 特定のイベントや事件のあった日、時期と一致する可能性がある場合、他の情報を参照することによって、特定の個人を識別する可能性が高まるため、取得時期の特性に配慮して加工することが望ましい。
- 本ユースケースでは、対象エリアにて、配慮が必要な特定のイベントや事件等が期間内になかったなかつたとみなし、特に配慮しない。

### 5) 位置の精度

- 高い位置精度の情報は、特定の個人を識別する可能性が高いため、適切に精度を低減することが望ましい。人口密度の低いエリアを対象とする場合は、特に配慮することが望ましい。
- 本ユースケースでは、位置情報を 500m 四方のメッシュとしていること、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアであることから、特に配慮はしない。

### 6) 移動履歴の期間・範囲

- 移動履歴の期間は長くなったり、特定の時間帯を対象としたりする場合は、次の a)~c) に係るリスクが高くなるため、これらに配慮して加工することが望ましい。
  - a) パターン性
    - 定期的に通っている場所、滞留している場所が分かることにより、自宅、通勤・通学地などが推測されて、特定の個人の識別性が高まる。
  - b) 場所の特性
    - 「2) 場所の特性」を参照。

#### c) 識別性

➤ 履歴の一意性が高まる。その一意性をもって、直ちに個人を特定することができないとしても、一定の配慮をすることが望ましい。

- 本ユースケースでは、位置情報の精度を 500 メートル四方のメッシュに変換していることから、自宅、通勤・通学先が特定される可能性は低いですが、万全を期す場合、契約者情報を用いて、対象とするエリアに自宅のある個人のレコードを削除する（「2) 場所の特性」での記載の再掲）。また期間を長くすることで、差異の大きな個人が識別される場合は、トップコーディングを行う。さらに同一の事業者に提供する場合は、履歴の期間が重ならないように提供する等の配慮をし、各期間の履歴が結びつかないようにする。

#### 7) 時間の精度・間隔

- 時間の精度が高まったり、データを取得する際の時間間隔が短くなったりすると、特定の個人を識別する可能性が高まる。また、詳細な時刻情報は位置情報とセットになることで、異なるデータセット間における共通の識別子として機能し得る。このため、適切に時間の精度を低減したり、間隔を開けたりすることが望ましい。
- 本ユースケースでは、秒単位で取得された時間の精度を、15分単位に丸める。

#### 8) 対象者数

- 加工対象とするデータセットに含まれる対象者数が少ないと、特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。また、同一の個人が複数台の携帯端末を所持している場合のあることを想定して、携帯端末の台数よりも対象者数が小さくなる可能性のあることに留意することが望ましい。
- 本ユースケースでは、対象者数でカウントして、一定規模の対象者数が確保されることを確認する。

#### 9) データ提供までの期間

- データを取得してから、匿名加工情報として提供するまでの期間が短い場合は、他の情報を参照することによって、特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。
- 本ユースケースでは、位置情報を 500m 四方のメッシュとしていること、都心で不特定多数の人が来訪する一般的なオフィス・商業エリアであることから、特に配慮はしない。

以上の本ユースケースにおける各情報についての加工の方向性をまとめると、次のようになる。

図表 1-4 商用のユースケースにおける加工の方向性

契約者情報、趣味嗜好情報

データ項目	想定されるリスク	加工の方向性
端末識別符号	契約者を特定されるリスクがある。	仮 ID に置き換える。
氏名	単体で個人を特定できる。	全部削除する。
性別	他の情報との組合せによって、個人が特定されるリスクがある。	加工しない。他のデータ項目を加工することで個人が特定されるリスクを下げることで対応する。
生年月日	他の情報との組合せによって、個人が特定されるリスクがある。	20 歳未満、20 代、30 代、40 代、50 代、60 代、70 歳以上の 7 つの年代に置き換える。
住所	他の情報との組合せによって、個人が特定されるリスクがある。 他の事業者でも収集している可能性があり、それと照合して個人が特定されるリスクがある。 また、本人にアクセスされるリスクがある。	市区郡単位より細かい情報を削除する。
趣味嗜好	他の情報との組合せによって、個人が特定されるリスクがある。	出現頻度の低いカテゴリがある場合は、上位のカテゴリに統合して一般化する。

位置情報

データ項目	想定されるリスク	加工の方向性
端末識別符号	契約者を特定されるリスクがある。	仮 ID に置き換える。
日時	位置情報に特有の性質を踏まえた	15 分単位にまるめる。
位置情報（メッシュコード）	9 つの評価要素に係るリスクがある。 1) 付帯情報 2) 場所の特性	500m 四方のメッシュコードに変換する（前処理段階で実施）。  契約者情報を用いて、対象とするエ

データ項目	想定されるリスク	加工の方向性
	3) 集団の規模 4) 取得時期の特性 5) 位置の精度 6) 移動履歴の期間・範囲 7) 時間の精度・間隔 8) 対象者数 9) データ提供までの期間	リアに自宅のある個人のレコードを削除する。  対象者数でカウントして、一定規模の対象者数が確保されることを確認する。  期間を長くすることで、差異の大きな個人が識別される場合は、トップコーディングを行う。また同一の事業者を提供する場合は、履歴の期間が重ならないように提供する等の配慮をし、各期間の履歴が結びつかないようにする。

<加工後のデータのイメージ>

上記の考え方に基づいて加工されたデータは、図表 1-5 のようになる。

図表 1-5 商用のユースケースにおける加工後のデータのイメージ

契約者情報、趣味嗜好情報

仮 ID	性別	年代	住所	趣味嗜好
12fa7d1	男	40 代	東京都千代田区	観劇
B66747ff9	女	20 代	東京都西東京市	スポーツ観戦
6c7de4b	男	60 代	神奈川県横浜市旭区	観劇

位置情報

仮 ID	日時	メッシュコード
12fa7d1	201611281330	533946004
12fa7d1	201611281500	553946004
12fa7d1	201611300200	533946111

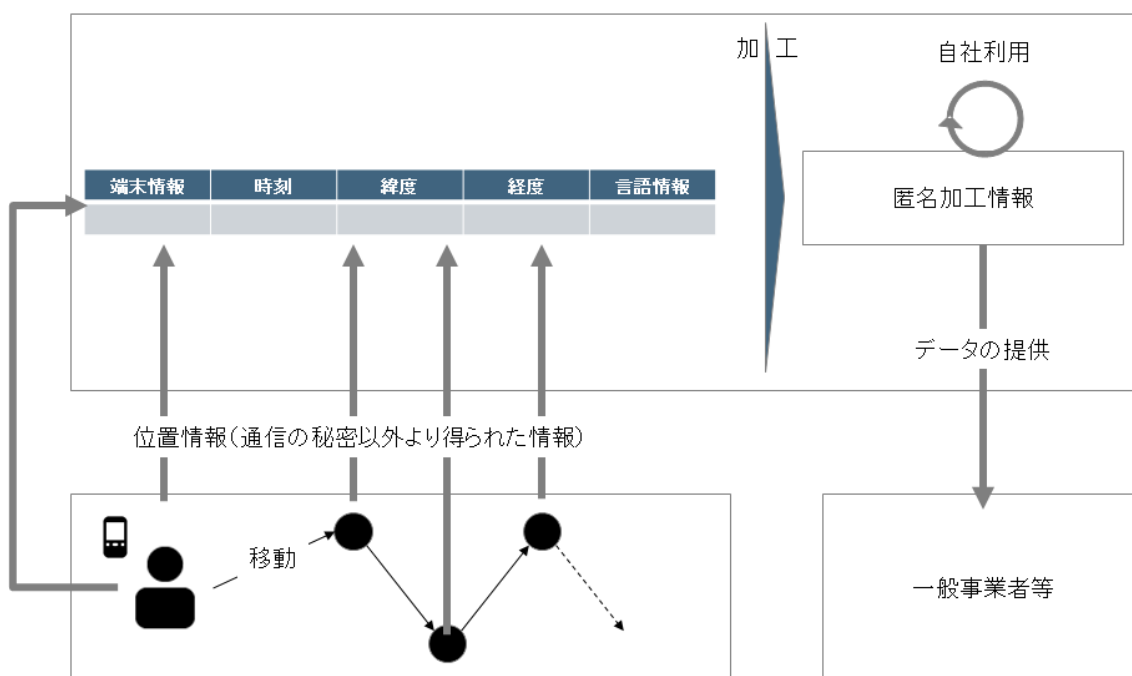
## 2. 観光の事例<sup>19</sup>

### 1) ユースケース

本ユースケースは、電気通信事業者が保有する通信の秘密に属さない Wi-Fi 位置情報（利用者とアクセスポイント設置者間の通信に基づく位置情報）や他の情報について、匿名加工した上で、匿名加工情報の枠組みを活用して、自社で活用したり、一般事業者へ提供したりするものである。

電気通信事業者や一般事業者においては、利用者（海外からの観光客等も含む）の動態や利用ルート等を把握することに活用することが想定される。

図表 2-1 電気通信事業者が保有する位置情報を利用するユースケースのイメージ



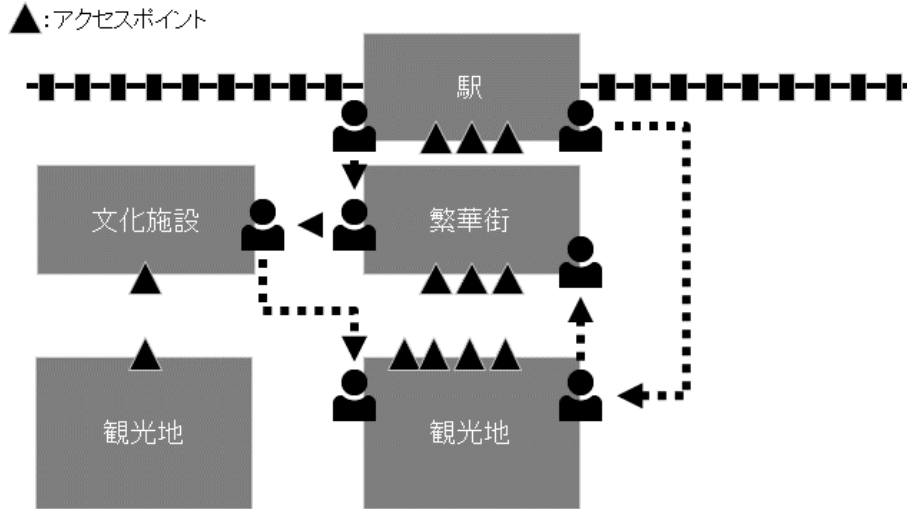
本ユースケースで取扱う位置情報は、外国人観光客がよく訪れる観光地を対象とし、スマートフォン等の携帯端末による、鉄道の駅や文化施設、観光スポット等に設置されたアクセスポイントとの通信で得られるものを対象とする。複数のアクセスポイントが一つの駅や施設等に設置されている場合は、匿名加工の前処理段階で、駅や施設等の単位に集約する処理を行う。

付帯情報については、アクセスポイント利用時に選択する言語情報を含む利用者情報（言語選択の際、氏名も入力することを想定している）を利用する。テーブルは、MAC アドレ

<sup>19</sup> 本事例の作成に用いたデータはすべて、加工方法の検討用に作成したダミーデータを用いている。

スおよびアクセスポイントによって紐付けが可能である。

図表 2-2 観光地におけるユーザーの移動イメージ



図表 2-3 位置情報に関するデータのレイアウトイメージ

利用者情報

MAC アドレス	言語	氏名
19-88-76-55-18-99	英語	XXXX XXXX
26-99-88-25-55-66	中国語（簡体字）	YYYY YYYY
18-66-70-09-81-33	タイ語	ZZZZ ZZZZ

※言語情報は、アクセスポイント接続時に選ばれた言語情報を使用することを想定

位置情報

MAC アドレス	取得時刻	アクセスポイント番号
19-88-76-55-18-99	20170223145731	XXX-XXXX-XX1
19-88-76-55-18-99	20170223192651	XXX-XXXX-XX5
26-99-88-25-55-66	20170228124539	XXX-XXXX-YZ3

アクセスポイント位置情報変換テーブル

アクセスポイント番号	地点名
XXX-XXXX-XX1	東京都 東京駅
XXX-XXXX-XX5	東京都 東京駅
XXX-XXXX-YZ3	東京都 東京駅丸の内ビル

※アクセスポイントは、特定の施設等に敷設されることが多いため、地点名もそれに紐づけ

て整理する

※アクセスポイントは、1つの施設に複数存在することが多いが、本ユースケースでは、それらは1つの施設に紐つけて整理する

## 2) 考慮すべき事項とリスクに対応した具体的な加工方法の検討

電気通信事業者は、位置情報と付帯情報とを結合した情報を、個人情報保護法 36 条 1 項に基づく個人情報保護法施行規則 19 条各号とそれらに対応する個人情報保護法ガイドライン（匿名加工情報編）を遵守し、匿名加工情報を作成する。この際、位置情報に特有の個人情報データベース等の性質を踏まえたその他の措置を実施することに努めなければならない。

以下に具体的な加工方法を示す。

<個々の加工方法>

### ①. 【規則 19 条 1 号】特定の個人を識別することができる記述等の削除

- 1 号では、「個人情報に含まれる特定の個人を識別することができる記述等の全部又は一部を削除すること（当該全部又は一部の記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）」が求められている。
- 本ユースケースでは、言語選択の際に登録した氏名が該当する。氏名は個人情報であり、かつ、データの紐付けには用いないことから、削除する。

### ②. 【規則 19 条 2 号】個人識別符号の削除

- 2 号では、「個人情報に含まれる個人識別符号の全部を削除すること（当該個人識別符号を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）」が求められている。
- 本ユースケースでは、当該識別符号はないため、本規則に則った加工は実施しない。

### ③. 【規則 19 条 3 号】情報を相互に連結する符号の削除

- 3 号では、「個人情報と当該個人情報に措置を講じて得られる情報を連結する符号（現に個人情報取扱事業者において取り扱う情報を相互に連結する符号に限る。）を削除すること（当該符号を復元することのできる規則性を有しない方法により当該個人情報と当該個人情報に措置を講じて得られる情報を連結することができない符号に置き換えることを含む。）」が求められている。
- 本ユースケースでは、MAC アドレスが該当する。MAC アドレスを用いて、位置情報と利用言語を連結する。また、アクセスポイントも具体的な地名を照合する際に、連結して使用する。

MAC アドレスは、ハッシュ化等により一方向に仮 ID 化をする。

アクセスポイントについては、地名照合後は、本情報は地名に代替されるため、削除することが望ましい。

#### ④.【規則 19 条 4 号】特異な記述等の削除

- 4 号では、「特異な記述等を削除すること（当該特異な記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）」が求められている。
- 本ユースケースでは、アクセスポイントおよび地名が該当しうるが、観光地を中心とした情報の収集においては、「特異な記述」にあたるものは、本ユースケースの限りではない。よって、本ユースケースにおいては、本規則に則った加工は実施しない。

#### ⑤.【規則 19 条 5 号関連】個人情報データベース等の性質を踏まえたその他の措置

- 5 号では、「前各号に掲げる措置のほか、個人情報に含まれる記述等と当該個人情報を含む個人情報データベース等を構成する他の個人情報に含まれる記述等との差異その他の当該個人情報データベース等の性質を勘案し、その結果を踏まえて適切な措置を講ずること。」が求められている。
- 本ユースケースでは、言語情報が該当しうる。図表 2-3 に挙げた言語情報は、英語、中国語（簡体字）、タイ語の 3 言語であり、我が国へのインバウンド件数は多く見込まれる。ただし、言語情報を増やした際には、インバウンド数が少なく、母集団数が少ない言語が存在する可能性があり、その場合、特定の個人を識別する可能性が生じる。一定程度の規模のレコード数に満たない言語に関しては、「その他」言語に区分をする（一般化）または、その言語が含まれるレコードを削除する等の処理をする。
- また、位置情報特有の性質へ対処するため、次の 1)～9) に掲げる評価要素による特定の個人が識別されるリスクを定性的に評価し、総合的に判断して、同リスクが十分に低減するための措置を実施することに努めなければならない。なお必ずしも 1)～9) の全ての評価要素を満たす加工を求めるものではない。

<位置情報に特有の個人情報データベース等の性質を踏まえたその他の措置>

##### 1) 付帯情報

- 付帯情報によっては、個人を特定する可能性が高まることに配慮して選定・加工することが望ましい。
- 本ユースケースでは、特にこれに該当する情報は見受けられないため、本項目に係る処理は実施しない。

##### 2) 場所の特性

- 対象とする位置情報に、自宅、通勤・通学地が含まれる場合は、配慮して加工すること



が望ましい。また、対象とする位置情報に、要配慮個人情報に関わる場所が含まれている場合は、配慮して加工することが望ましい。

- 本ユースケースでは、場所を地名に置き換える処理をして取り扱っていることから、本項目に係るさらなる処理は実施しない。

### 3) 集団の規模

- 特定の学校・職場や稀少な趣味嗜好等を持つ集団を対象とした場合、集団の規模によっては、特定の個人を識別する可能性が高まるため、集団の規模に配慮して加工することが望ましい。
- 本ユースケースでは、特定の個人を識別する可能性のある集団を対象としていないため、特に配慮しない。

### 4) 取得時期の特性

- 特定のイベントや事件のあった日、時期と一致する可能性がある場合、他の情報を参照することによって、特定の個人を識別する可能性が高まるため、取得時期の特性に配慮して加工することが望ましい。
- 本ユースケースでは、対象エリアにて、配慮が必要な特定のイベントや事件等が期間内になかったとみなし、特に配慮しない。

### 5) 位置の精度

- 高い位置精度の情報は、特定の個人を識別する可能性が高いため、適切に精度を低減することが望ましい。人口密度の低いエリアを対象とする場合は、特に配慮することが望ましい。
- 本ユースケースでは、外国人観光客がよく訪れる観光地を対象としていること、場所を地名に置き換える処理をして取り扱っていることから、本項目に係るさらなる処理は実施しない。

### 6) 移動履歴の期間・範囲

- 移動履歴の期間は長くなったり、特定の時間帯を対象としたりする場合は、次の a)~c) に係るリスクが高くなるため、これらに配慮して加工することが望ましい。
  - a) パターン性
    - 定期的に通っている場所、滞留している場所が分かることにより、自宅、通勤・通学地などが推測されて、特定の個人の識別性が高まる。
  - b) 場所の特性
    - 「2) 場所の特性」を参照。
  - c) 識別性

➤ 履歴の一意性が高まる。その一意性をもって、直ちに個人を特定することができな  
いとしても、一定の配慮をすることが望ましい。

- 本ユースケースでは、アクセスポイントを地名に置き換えること、外国人観光客がよく訪れる観光地を対象としていることから、更なる加工はしない。ただし、期間を長くすることで、差異の大きな個人が識別される場合は、トップコーディングを行う。また同一の事業者を提供する場合は、履歴の期間が重ならないように提供する等の配慮をし、各期間の履歴が結びつかないようにする。

#### 7) 時間の精度・間隔

- 時間の精度が高まったり、データを取得する際の時間間隔が短くなったりすると、特定の個人を識別する可能性が高まる。また、詳細な時刻情報は位置情報とセットになることで、異なるデータセット間における共通の識別子として機能し得る。このため、適切に時間の精度を低減したり、間隔を開けたりすることが望ましい。
- 本ユースケースでは、秒単位で取得された時間の精度を、15分単位に丸める。

#### 8) 対象者数

- 加工対象とするデータセットに含まれる対象者数が少ないと、特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。また、同一の個人が複数台の携帯端末を所持している場合のあることを想定して、携帯端末の台数よりも対象者数が小さくなる可能性のあることに留意することが望ましい。
- 本ユースケースでは、対象者数でカウントして、一定規模の対象者数が確保されることを確認する。

#### 9) データ提供までの期間

- データを取得してから、匿名加工情報として提供するまでの期間が短い場合は、他の情報を参照することによって、特定の個人を識別する可能性が高まることに配慮して加工することが望ましい。
- 本ユースケースでは、アクセスポイントを地名に置き換えること、外国人観光客がよく訪れる観光地を対象としていることから、特に配慮はしない。

以上の本ユースケースにおける各情報についての加工の方向性をまとめると、次のようになる。

図表 2-4 観光のユースケースにおける加工の方向性

利用者情報

データ項目	想定されるリスク	加工の方向性
MAC アドレス	利用者を特定されるリスクがある。	仮 ID に置き換える。
氏名	単体で個人を特定できる。	全部削除する。
言語	他の情報との組合せによって、個人が特定されるリスクがある。	一定程度の規模のレコード数に満たない言語に関しては、「その他」言語に区分をする（一般化）または、その言語が含まれるレコードを削除する等の処理をする。

位置情報

データ項目	想定されるリスク	加工の方向性
MAC アドレス	契約者を特定されるリスクがある。	仮 ID に置き換える。
日時	位置情報に特有の性質を踏まえた	15分単位にまるめる。
アクセスポイント番号	9つの評価要素に係るリスクがある。 1) 付帯情報 2) 場所の特性 3) 集団の規模 4) 取得時期の特性 5) 位置の精度 6) 移動履歴の期間・範囲 7) 時間の精度・間隔 8) 対象者数 9) データ提供までの期間	地点名に置き換える。1地点に複数のアクセスポイントが設置されている場合は、同じ地点名とする。  対象者数でカウントして、一定規模の対象者数が確保されることを確認する。  期間を長くすることで、差異の大きな個人が識別される場合は、トップコーディングを行う。また同一の事業者を提供する場合は、履歴の期間が重ならないように提供する等の配慮をし、各期間の履歴が結びつかないようにする。

<加工後のデータのイメージ>

上記の考え方に基づいて加工されたデータは、図表 2-5 のようになる。

図表 2-5 商用のユースケースにおける加工後のデータのイメージ

利用者情報

仮 ID	言語
12fa7d1	英語
B66747ff9	中国語（簡体字）
6c7de4b	タイ語

位置情報

仮 ID	日時	地点名
12fa7d1	201611281330	東京駅
12fa7d1	201611281500	両国駅
12fa7d1	201611300200	両国国技館

## 6. 今後の展開に向けて

### 1) 業界団体ルール、個人情報保護指針への取込み

本調査研究では、「十分な匿名化」に関するルール、及び電気通信事業者が扱う位置情報を用いた匿名加工情報の作成等に関するルールにそれぞれ盛り込まれるべき内容を取りまとめた。これらは、電気通信事業分野における業界団体や認定個人情報保護団体に引き継がれ、位置情報を取り扱う際の規範として活用されていくことを想定している。

改正個人情報保護法の全面施行は、平成 29 年 5 月 30 日に予定されており、これと同時期に、「電気通信事業における個人情報保護に関するガイドライン」も施行されることになっている。電気通信事業分野における業界団体や認定個人情報保護団体は、本調査研究でとりまとめた内容を踏まえて、自主ルールや個人情報保護指針を策定することが期待されており、総務省においてもこれらの活動を支援する取り組みが求められる。

### 2) 「十分な匿名化」に係る安全管理措置のフォローアップ

匿名加工情報は、加工後の情報であっても、安全管理措置の努力義務が課せられている。一方、「十分な匿名化」による情報は、文字通り、匿名化が十分なされているとして、加工後の情報に対しては、安全管理措置を求めなくてもよいと本調査研究では整理した。しかし、本調査研究で設置した協議会で指摘されたとおり、技術の進展によっては、加工後の情報から個別の通信や個人を識別される蓋然性を完全に排除することは困難である。

このため、総務省は、事業者のルールの運用状況をモニタリングしつつ、制度改正や技術の進展を踏まえ、定期的に見直しを行って、所用の措置を講じる必要がある。

### 3) 匿名加工情報に係る他分野の取り組みとの整合

本調査研究でとりまとめた匿名加工情報の作成に係るルール案では、「場所の特性」のように、要配慮個人情報に関わる場所が含まれている場合は、配慮して加工することが望ましいと整理している。一方、個人情報保護委員会が定める委員会規則やガイドラインにおいては、特定の個人が識別されないことを匿名加工情報の加工基準としており、要配慮個人情報への配慮までは求めていない。すなわち、本調査研究でとりまとめたルールは、一般的な分野よりも厳しい加工基準を求めている。

匿名加工情報は、制度ができたばかりで、まだ運用は開始されておらず、行政も事業者も試行錯誤しながら取り組むことになると想定される。本調査研究でとりまとめたルールも、他の分野の取り組みを踏まえて、試行錯誤しながら整合を図っていることが望まれる。

#### 4) IoTの進展を踏まえたルールづくり

本調査研究で対象とした位置情報は、基地局に係る位置情報、GPS位置情報、Wi-Fi位置情報の3種類であり、これらには通信の秘密に該当するもの、該当しないものがある。これら以外に、Bluetooth機能を備えたスマートフォン等の高機能な携帯端末では、ビーコンによっても位置情報を把握することができる。IoTの進展とともに、日常空間に様々なセンサーが設置され、様々な手段で、様々な主体が位置情報を収集し、活用されていくものと想定される。

こうしたデータの取得・分析・利用においては、本人とデータ利用者との二者の関係に留まらず、仲介者や分析者などの多くのステークホルダーが存在し、本人の意思確認や事業者間における契約関係等の調整が不可欠となる。このため、IoTの進展を踏まえて、位置情報をはじめとするプライバシーに関わりの深い情報を適切に活用するためのルールづくりを進めていくことが求められる。

## 付属資料

位置情報に関するプライバシーの適切な保護と社会的活用の両立に向けた調査研究  
第1回協議会 議事概要

日時：2016年11月8日 15:00-17:00

場所：野村総合研究所丸の内北口ビル 9F 大会議室 2

出席者：

委員：森主査、伊藤委員、佐藤委員、高橋委員、疋田委員

オブザーバ：個人情報保護委員会事務局、電気通信事業者協議会、日本データ通信協会、

NTT ドコモ、KDDI、ソフトバンク、NTT BP

総務省：湯本課長、徳光課長

野村総合研究所（事務局）：小林、小島、南島

## 開会

- ・委員・オブザーバご挨拶
- ・主査の選出（満場一致で森亮二氏を主査に選任）
- ・森主査ご挨拶
- ・湯本課長ご挨拶

## 1. 協議会の趣旨等について

資料1に基づき事務局より説明。委員・オブザーバからの質疑はなし。

## 2. 検討項目と主な論点

資料2、参考資料1、参考資料2に基づき事務局より説明。

### 【質疑応答】

疋田委員：

検討項目1「十分な匿名化」に関する検証において、本年度の検証対象としてサンプリングの無作為性を挙げているが、標本の抽出方法と同様に、抽出する標本のボリュームそのものにも留意すべきと考える。標本数を2倍に増やしても、同じだけ匿名性が上がるものではなく、一般に標本数と匿名性との関係は比例しない。

伊藤委員：

標本数と有用性との関係は重要で、誤差をどれ位認めるかによってサンプリングの方法が変わってくるので、必要な標本数とサンプリング方法は同時に議論すべきである。

サンプリングの無作為性といった際、よくランダムサンプリングが挙がるが、一口にランダムサンプリングといっても、複数の手法が存在する。

森主査：

検証を行う際の視点の一つとして受け止める。



佐藤委員：

本年度の論点としてリスク管理指標の一般化を挙げているが、一般化は事業者の利便性を考えて、すべきと思う反面、一般化の過程で基準が厳しめになる可能性があるため、利活用を目的とした場合、安易な一般化は避けるべきではないか。

また、本協議会における検討対象の範囲を確認したい。位置情報を取得する方法は端末に限られないが、本検討会では通信にかかる位置情報を検討の範囲とするということか。

森主査：

佐藤委員の指摘について、検討項目 5 を含めて整理したい。本協議会では、電気通信事業者が扱う位置情報について議論することを想定している。通秘・非通秘と分かれるが、いずれも端末を通じて取得する情報である。昨年度は通秘の位置情報について十分な匿名化が必要という議論をしたが、位置情報だから何が必要、通秘だから何が必要という具合に分けて議論してこなかったため、今年度はそこにメスを入れたい。

佐藤委員：

端末にはどこまで入るか。自動車も入るか。

疋田委員：

同じ GPS 情報でも取得する主体が電気通信事業者か否かによって、取り扱いが変わることには違和感がある。電気通信事業者以外についても考慮した位置情報について議論すべきでないか。

湯本課長：

電気通信事業者は契約者情報等の個人情報を保有しており、それが位置情報と結びつけられる点で電気通信事業者以外の事業者と大きく異なると考える。検討の範囲については、H26 年に公表した位置情報プライバシーレポートを具体化していくことが本調査自体の目的であり、P3 で示される通信キャリアの取得する位置情報が議論のメインであると理解いただきたい。はみ出る部分が相当程度あることは確かにそうだが、次年度に引き続き議論できる場を用意したいと考えている。

疋田委員：

小売の位置情報を通信事業者が取り扱う場合は本協議会の議論に含まれるか。

湯本課長：

別の会議体で電気通信サービスに付随するサービスの範囲を議論しているところ。そこでの議論では、ポイントの付与サービスは含まれると解しているが、本協議会の議論対象には該当しない。

高橋委員：

昨年度は十分な匿名化について議論を重ねた。今年度は昨年度の成果を引き継ぎ、使いやすくするという意味で、検討項目 1 について特に注力したい。具体的にはユースケースで

取り扱うデータ数を増やしたい。昨年度は限定的な数で行ったが、数が増えれば、使えることが増えるはず。概念的な目標として、ここでの議論をベースに事業者が自主的に判断できるようにすればよいと考える。そのために一般化にあたっては工夫をしたい。

森主査：

安易な一般化を避ける、又は一般化にあたり工夫するとは具体的にどういうことか。

佐藤委員：

ガイドラインやチェックリストのような統一的な基準は、事業者にとって従えば安全という点でメリットがあるが、反面自主的な利活用の検討を阻害する恐れがある。

高橋委員：

事業者の判断に資する良い事例を示すという方法も考えられる。メジャーな観光地・空港に人が束になって動くケースについて良い事例を示せば、事業者も使いやすいのではないか。今回は観光地等への人の移動についての良い事例を示して、他のケースは置いておくような漸近的な進め方でもよいのではないか。

森主査：

一般的な基準が示されず、ケースごとに有識者判断になるのは良くないと思うが、そういうことであれば、同意する。

高橋委員：

検討項目 2 について懸念がある。昨年度の調査を通じて良い事例を引き出す作業には相当の労力、時間を要すると感じた。今年度の制約の中で匿名加工情報についてこれを行うことは難しいと考える。検討項目 4 で挙げている関係性の整理は有用だが、検討項目 1 を侵さない前提で取捨選択すべきと考える。

佐藤委員：

匿名加工情報は第三者提供を前提とするため、売り手と買い手が存在するが、本協議会は売り手となることが想定される主体を中心に構成しているため、位置情報にかかる匿名加工情報にどんな需要があるか不明。位置情報の匿名加工を買う人が何を求めているか、考慮しながら検討すべきでないか。

事務局：

ユースケースについては次の議題の中でご議論いただきたい。

森主査：

検討項目は次回以降、度々立ち戻ることとし、本日の議論はこれまでとする。

### 3. ユースケースの設定と利用者の受容性の評価

資料 3 に基づき事務局より説明。

#### 【質疑応答】

ソフトバンク（東海オブザーバ）：

昨年度のヤフオクドームのケースは個別に同意を取得して使用したものであり、今年度用いることはできない。時間と工数の制約を考え、都内の繁華街・商業エリアの Wi-Fi データを利用できないか調整している。昨年度との差分として、属性のデータ項目に興味嗜好というカテゴリを追加したい。但し、そこはダミーデータになる。性別・年代に紐付く趣味嗜好のデータを外部から入手する予定。どれくらいのサンプル数が必要かについては委員の方々からご教示いただきたい。

NTT BP（北條オブザーバ）：

昨年度は 6 万サンプルデータを利用。想定しているケースではアプリ経由で端末やブラウザの言語情報を取得する。アプリ登録時に言語情報の提供に同意してもらえれば、取得できる整理。オリンピックが決定してから訪日外国人が増加しており、彼らの動きに対しては自治体の興味も大きくなっている。アジアからの観光客といっても、国ごとに観光経路は違う。彼らの移動情報から、パスされてしまう観光地があれば、その魅力を発信することもできる。

NTT BP は Wi-Fi 事業者であるため、補足できるのはその観光地に訪れた Wi-Fi 利用者に限られ、観光地に訪れている人全員の動きを掴めるわけではないことは留意が必要。

伊藤委員：

ユースケースの有用性について資料には「○」「×」「△」と記載しているがどう評価したのか。

ソフトバンク（東海オブザーバ）：

社内のマーケティング部署の人間が確認した結果。厳密な軸をもって評したものではない。

高橋委員：

補足すると「×」と評している箇所については K=10 で匿名化すると性別と年代に関する情報が失われてしまったといった、客観的な事実に基づき、最終的な判断をソフトバンクで実施した。

疋田委員：

「×」としている箇所は年代等一部のセグメントにおいて使えないから×という判断だったのか。

ソフトバンク（東海オブザーバ）：

年代に絞らず、幅広いターゲティングをしたときに使えるかどうかを判断した。

疋田委員：

「×」がついても年代層を限れば、使用に耐えるデータがあるということか。

ソフトバンク（東海オブザーバ）：

ご認識の通り。

高橋委員：

有用性という表現が適切でなかった可能性がある。

疋田委員：

「×」と評している箇所も 20・30 代向けのマーケティングなら問題なかった可能性がある。徒な一般化による弊害になりうる。

高橋委員：

いずれにせよ示された 2 つのユースケースを広げていくという理解でよいか。

事務局：

別途委員・オブザーバから別のケースが提供いただけるのであれば、それを利用したいが、現時点では事務局としてはその想定。

高橋委員：

昨年度の発展系とするなら妥当なケースと考える。但し、両ケースとも Wi-Fi 情報を想定している点は気になる。

事務局：

今年入手できるのが、Wi-Fi データのみという理由による。

高橋委員：

了。

疋田委員：

受容性評価にあたり、5 つの年代で男女各 20 名ずつサンプル取得しようとする、全体で 200 サンプル集めることになるが、資料上は計 100 サンプルとある。これは誤植か。その上で、各年代、男女それぞれ 10 名ずつというサンプル数でクロス集計を行うことは統計的にやや乱暴ではないか。

事務局：

誤植である。各 10 名ずつに訂正させていただきたい。

伊藤委員：

恐らく誤差が大きくて評価に苦しむことが想定される。

疋田委員：

年代を 5 つではなく、20・30 代に絞る等、各セグメントのサンプル数を増やすべきではないか。

高橋委員：

サンプル数もそうだが、何を確認するかも重要になる。一般の方に加工手順について尋ねても評価のしようがなく、こうした加工手順であなたのプライバシーが守られていると思いますかといった確認の仕方が適切と考える。

事務局：

過去の類似する実証においては、加工の方法を説明すると、一向に構わないという回答者

と気になるという回答者に分かれる傾向があった。今回は同意やオプトアウトを組み合わせたときの違いを見たい。質問の仕方については別途パッケージを示す。

疋田委員：

評価対象のサンプルをどう選ぶかによる。事前に Web で選出して、十分な説明をする場合と、しない場合とで評価がどう違うか確認できると良いかもしれない。

事務局：

元々のプライバシーに対する考え方は考慮して意図を確認することを想定している。

高橋委員：

説明にあたり位置情報の取得方法は明示すべき。位置情報以外にも言語情報が取得されることが自明であることが一般的かどうか、疑問がある。説明のパッケージに加えるべき。

森主査：

本評価調査における説明パッケージは今後、電気通信事業者が説明を行う際、活かせるはずである。

佐藤委員：

位置情報を匿名加工情報にした場合のニーズについてはどう考えるか。どういう位置情報を匿名加工情報にしてどこに売るかについて、イメージを共有できれば、議論が発散しないのではないか。

事務局：

地域に詳しいコンサルタントが自治体向けのコンサルティングにおいてデータをほしいとすることもニーズの一つと考えている。

森主査：

制度的に通秘を多分に含む位置情報を匿名加工情報にできるのかという議論もある。

湯本課長：

そもそも整理が必要かという認識でいる。事業者として位置情報の匿名加工情報化に対しどこまでウィルがあるかについては次回検討会までに確認したい。

疋田委員：

具体的なニーズの中身を確認しようとしても事業者も発言しづらいと考える。ニーズの確認方法として、JR 東日本が日立に提供しようとした事例など、過去に報道があった案件に対し、ニーズがあるかないかを確認することが適切ではないか。

佐藤委員：

日立の事例は不動産開発にあたり日常的な人の流出入を確認したいというニーズに基づくものなので、今回扱おうとしているユースケースとマッチしないと考える。匿名加工するならどんなユースケースがいいかは考えるべき。

湯本課長：

ニーズについては経済産業省の別会議体での議論の内容を共有することが可能か事務局

と相談する。

疋田委員：

位置情報は、今の位置自体をスナップショットとして利用するか、履歴として利用したいかは一度整理したほうがよい。位置情報は一般的には今、自分がどこにいるかがイメージされる。履歴が長くなればなるほど、個人の特定性が高まる。

森主査：

ユースケースの検討において議論する。

以上

位置情報に関するプライバシーの適切な保護と社会的活用の両立に向けた調査研究  
第2回協議会 議事概要

日時：2016年12月13日 15:00-17:00

場所：野村総合研究所大手町グランキューブ 29F 大会議室 14

出席者：

委員：森主査、伊藤委員、高橋委員、疋田委員

オブザーバ：個人情報保護委員会事務局、電気通信事業者協議会、日本データ通信協会、  
NTT ドコモ、KDDI、ソフトバンク、NTT BP

総務省：湯本課長、徳光課長

野村総合研究所（事務局）：小林、小島、南島

## 1. 電気通信事業者における位置情報の取扱い状況と活用ニーズ

資料1に基づき事務局より説明。

### 【質疑応答】

森主査：

資料1 P.2に記載された、「「基地局に係る位置情報」と「GPS位置情報」とを組み合わせた利活用は行われていない」とはどういう意味か。

事務局：

基地局情報は通信端末に係る契約者情報であり、GPSは携帯アプリ等と紐づく情報である。これらの情報を、ID等を用いて紐付けてはならず、そのため、組み合わせた利活用は行われていないと記載している。

疋田委員：

資料1 P.1に関して、Wi-Fiの利活用は、無料と有料とで、保有しているデータが異なるものと考えられる。P.1で意図されたユースケースは、無料利用を想定しているのか。

NTT BP 北條オブザーバ：

Wi-Fi事業者は、プラットフォームを提供するも、携帯電話キャリア各社の保有する契約者情報は見られないという意味である。

森主査：

資料1 P.2の図表に関し、A社は通秘の基地局情報を利用、B社は通秘のWi-Fi情報を利用としているが、それぞれどのように匿名化をしているのか。

事務局：

A社、B社ともに、個別同意を取っている。なお、「利活用」と言うと活用が進んでいるように見えるが、現状は、個別同意というスキーム上、活用可能なデータ量は膨大にある訳

ではなく、大いに活用されている状況には至っていない。

疋田委員：

これらの利活用は現状、個別同意の中で実施され、匿名化した位置情報は、各社社内で活用し、外に提供している訳ではないという認識である。

伊藤委員：

①. 資料1 P.2 で観光動態把握に位置情報を使用しているが、どの様な目的で利用しているのか、②. 基地局と Wi-Fi のそれぞれの位置情報の利用に関して、そもそも基地局や Wi-Fi のアクセスポイントの普及状況やカバー範囲はどのようになっているのか。

事務局：

①については、観光地への訪問状況の確認や、訪問されない観光地の特定のために、利用している。②については、事業者各社によって、かつ地域によって状況が異なっている。

伊藤委員：

①に関して、観光に関するデータ利活用の中で、オープンデータがキーワードの 1 つになっていると考えている。紹介された取組みでは、統計や行政記録だけではなく、民間のデータもオープンデータ化されうると理解してよいのか。

事務局：

将来的にその可能性は考えうるが、現状事業者は、自社でデータ取得、データ分析をした結果で事業を成り立たせており、民間データのオープンデータ化という傾向にある訳ではない。

森主査：

資料1 P.3 に関して、実務では Beacon と合わせた位置情報の利活用がされていると記載されているが、具体的にはどういうことか。

事務局：

Bluetooth 通信を活用した、安価で電池寿命の長い Beacon 端末を活用し、その場その場で広告や販促を行う活用がされている。

## 2. ユースケースの設定

資料2に基づき事務局より説明。加えて事業者、高橋委員より下記の補足とコメントを頂く。

### 【補足とコメント】

NTT BP 大島オブザーバ：

昨年度の 1 ヶ月間相当の分析期間を半年に拡大し、委員の皆様にはアドバイスを頂きながら、分析を進めていきたい。



SB 東海オブザーバ :

昨年の分析との大きな違いは、場所の違いとデータがダミーである点が挙げられる。実態と異なることを前提とした上で、ダミーの有用性という観点から、今後の分析を進めていきたい。

高橋委員 :

十分な匿名化と匿名加工情報のそれぞれの定義と違い、検証ポイントに関して確認・補足をしたい。

十分な匿名化は、通常のk-匿名化のような、ターゲットの属性を残しつつ、住所、年齢などを匿名化する形態ではなく、資料2 P.3にあるような、位置等の履歴を組み合わせた一式に対して、非特定性を担保しようとしている点が、認識すべきポイントである。

十分な匿名化の論点は、資料2 P.7で記載されているが、各項目の組み合わせ(例えば、時間の細かさ、場所の精度、ホップ数等)によって、全体としてk-匿名性を担保し、それを一般化していく観点で、検討を進めていくことが望ましい。

一方、匿名加工情報は、十分な匿名化のような限定されたデータセットではなく、例えば、住所、性別、年齢、移動履歴、その他各種属性情報といった形で、項目を多数追加しうる点が、重要なポイントである。

データ項目が増えると、有用性は高まるが、特定性も高まるというトレードオフが生じるため、どの項目は重視する/諦めるといった整理をする必要があることを前提に、ケースの定量分析だけでなく広い意味で、検討を進める必要がある。即ち、匿名加工情報では用途ごとに様々なバリエーションが考えられるため、細則の論点だしよりも、基本スタンスを決定することが重要である。

#### 【質疑応答】

伊藤委員 :

手法としてサンプリングが挙げられているため、下記4点を考慮していく必要があると考える。

①. データ加工のタイミングとして、サンプリング前にデータを処理するか、サンプリング後にデータを処理するかで、結果が変わってくるため、手順を予め決める必要がある。

②. サンプリング率の設定(最大80%とするのか、50%とするのか、どのようなパターンでサンプリング率を設定するか等)は、作業ボリュームに大きな影響を与えるため、予め決める必要がある。

③. 適切なサンプリングの評価手法を考える必要がある。統計では例えば、ユニークな標本を数え上げる(サンプルユニーク)手法が採用されるが、この手法が位置情報において適切か、考える必要がある。さらに、評価結果の判断手法も検討することが重要である。

④. サンプリングとk-匿名化との関連の検討が必要である。資料2では、十分な匿名化、

匿名加工情報とともにk-匿名化の議論がされているが、例えば、十分な匿名化においても、サンプリングを用いた場合、k=1が許容されるのではと考える。

また、十分な匿名化と匿名加工情報とでは、趣味嗜好の分析手法が異なってくる可能性がある。十分な匿名化においては、趣味嗜好情報は、入り口要件で大分類に加工する必要があるが、匿名加工情報は、はかかなり緻密な趣味嗜好情報を用いて分析することが可能ではないかと考えており、考え方を整理する必要がある。

高橋委員：

サンプリングは、十分な匿名化においては検討せず、匿名加工情報においては検討してもよいという位置づけの技法と考えている。理由としては、十分な匿名化は、組み合わせをN数で稼ぐ発想をしており、サンプリングの考え方と異なることが挙げられる。

十分な匿名化でサンプリングを容認するのはk=1の場合のみと考えるが、この場合はサンプリング前の結果で $k \geq 2$ を担保していることと、ほぼ同義であると考えている。

一方、匿名加工情報においては、k=1においても個人情報を守るという観点で、サンプリングを検討する意義があると考えている。ただし、このようなサンプリングの利用に関する実例は、国内外でも見当たらず、相当にチャレンジングな検討になると感じている。

森主査：

匿名加工情報において、これまでサンプリングがあまり議論をされていなかったのは、幅広い業種・業態を意識した際に、しっかりとサンプリングしきることは困難と整理してきたためだと考えている。ただし、本事業で対象となっている電気通信事業者の実力に鑑み、本事業で、サンプリングを検討することには意義があると考えている。

高橋委員：

取り組むことにメリットが少ない検討は、避けるべきだと考える。学術的には興味深いが、このようなニーズが本当にあるのかは、分からない。特に、十分な匿名化の検討においては、パターンが多々生じて、相当に難しい検討になると感じている。

伊藤委員：

統計では、標本調査をする際、調査結果が母集団をどの程度代表しているか、ユニークな標本が検出された際、それが母集団のどの程度を占めているか、という議論がある。サンプリングによって、安全性を確保することは、考え方としてはあると考える。

事務局：

十分な匿名化でサンプリングの検討をする意義に質問を頂いたが、そもそも十分な匿名化におけるkの議論も未だ確定したものではない。k-匿名化とサンプリングを組み合わせることで、より匿名性を確保できるという考え方はあるのではないかと。

高橋委員：

匿名加工情報において、サンプリングの議論が少なかった技術的な経緯は、ビッグデータに係る検討において、母集団の「安定さ」が確保されるとは限らない点が挙げられる。

サンプリングの正当性は、サンプリングをすることで、「特定個人の情報が含まれていないかもしれない」ことに裏づけされることが多いが、その重要な前提として、サンプリングが、「正しく」抜き出しをしていたかという点が挙げられる。

ある地域で 10 代の人口が安定的に存在するのであれば、サンプリングの効果はあるが、六本木ヒルズの最上階に住む年収 5 億円の人をサンプリングする行為は意味がない（サンプリングしようがしまいが、特定されてしまう）。

ビッグデータは、安定的な母集団であるとは言いがたく、この前提が確立されていない中でサンプリングを活用するのはあまり適切ではないというのが、技術から見た、上述の根拠だと考えている。

その中で、移動履歴は、ビッグデータ同様、安定した母集団と言えとは限らない。

さらに、今後の利活用のことを考えた場合、データの利用に係る技術や説明は、シンプルで分かりやすいものであるべきと考え、サンプリングのような技術に傾倒した手法は、優先的な検討項目とは考えていない。

見なければならぬパラメータは、他にも多数存在するため、それを先にやるべきではないか。サンプリングは、できれば今回は、検討項目から外したい。

森主査：

「安定さ」というのは、データを構成する項目全てに対して求められるのか。

伊藤委員：

クラスターをどう考えるかによって、変わってくる。データの特性に依存するため、一概には質問のように整理できない。

疋田委員：

しっかりとクラスターが出来るのであれば、安定していると評価されるが、移動履歴は、指紋とのアナロジーで考えてみると、サンプリングしても特定可能になりやすいと考える。前処理や場所の特性を踏まえれば、安定したデータとなる（例えば山手線内）し、そうでなければ、クラスターの形成が不可能となる可能性がある。

疋田委員：

資料 2 P.7 について、半年分の観光データを収集するとしているが、外国人観光客を想定しているため、同一人物の動きをトレースする訳ではないと理解してよいか。

NTT BP 大島オブザーバ：

目的がそもそも観光分析であるため、現時点では、地元の人などは対象外とすることを想定している。

疋田委員：

半年の分析をする際、時間の概念を無視して分析しても良いか。夏と冬では観光動線が異なるということも想定される。

事務局：

本事業での検証ポイントは、前年度は1ヶ月のスパンでの検討であったことを、6ヶ月に拡張したというところにある。

正田委員：

季節属性がない観光地域など、選ぶ場所に配慮があれば、問題なく分析できるだろう。

NTT ドコモ 松岡オブザーバ：

サンプリングに関して、自社のシェアという概念に鑑み、自社のみがもつデータは、サンプリングされた結果だと考えることは可能か。

正田委員：

期間が短ければそれも可能だが、長期的に／広域にデータを見ると、移動に典型的な傾向が見えてくることも多く、その場合は、サンプリングされた結果とみなすことは不可能となる。例えば、新宿～渋谷での移動を見るだけなら、全く問題はないだろうが、長距離化すると、特徴点が露見する可能性は高い。

高橋委員：

仮に、本当に十分に匿名化されているのであれば、松岡オブザーバの考え方は成立する。もっとも、ここで言う「本当に十分に」とは、どこを見ても当たり前のようにデータが存在するような状況のことであり、現実、そうはいかないことが多い。

伊藤委員：

サンプリングとは、母集団を標本集団が代表することであり、サンプリングをするということは、ウェイトバックをすることとセットで考える必要がある。この考え方に基づくと、そもそも今回のケースで、ウェイトの議論が生じるのか、疑問である。この疑問は、ビッグデータに関するサンプリング適用の際にも生じる。

正田委員：

商用の事例で、有楽町駅ははじめから分析対象としないという理解でよいか。

SB 東海オブザーバ：

現状は対象外にしているが、分析をしながら、対象にするかどうかの検討は別途行う可能性がある。

森主査：

本事業に関して、個人情報保護法に係る事務局レポートが来年2月に公表されると聞いており、それらの情報は、本事業でも確認していく方がよいと考える。

高橋委員：

分析を担当する事業者にお願いがある。本事業では、一般化に向けた検討がゴールである。例えばホップ数についても、2ホップを前提としてしまえば、3ホップ以降の議論が出来な

くなる。即ち、一般化に向けた検討が出来なくなる。

出来るのであれば、実施事項を含め、今後も相談やお願いをさせて欲しい。

**SB 東海オブザーバ：**

分析の工数を昨年度との比較で考えた場合、全ての検証ポイント、パターンをやりきるの  
は不可能であるとする。どれを優先的にやるか、どれはやらないか等の整理をしていく必  
要がある。

### 3. 利用者の受容性の評価

資料3に基づき事務局より説明。

#### 【質疑応答】

正田委員：

事務局提案の手法を実行した際、矛盾した回答者が生じる点を懸念している。保守的な回  
答をする回答者が、実は Google マップを使っていた、というケースは容易に想定され、回  
答特性と行動特性が一致していない可能性が考えられる。スクリーニングをする際には、例  
えば、「地図アプリを使っているか？」等の質問を入れておいた方がよいだろう。精度を上  
げるための質問設計を考えたい。

事務局：

基本的には、位置情報プライバシーレポートを踏まえてスクリーニングすること原則と  
し、問題提起頂いた回答特性と行動特性との差異は、別途検証項目に入れたい。

正田委員：

商業利用に関して、「あなたに位置情報を返すのは許容できるか」と聞くのと、「第三者に  
提供するのを許容できるか」と聞くのとでは、回答者の心象が大きく異なってくるため、質  
問の文言にも配慮をしたい。

高橋委員：

十分な匿名化と匿名加工情報の受容性の差異は、最も関心を持っている。おそらく、十分  
な匿名化の受容性は高く、匿名加工情報の受容性が低いと予測されるが、それをどう評価す  
るかに関して踏み込める質問を用意しておきたい。

事務局：

この差異は、アンケートによって定量的に評価するだけでなく、デプスインタビューで検  
証していくことを想定している。

以上

位置情報に関するプライバシーの適切な保護と社会的活用の両立に向けた調査研究  
第3回協議会 議事概要

日時：2017年2月7日 10:00-12:20

場所：総務省 10F 1001 会議室

出席者：

委員：森主査、伊藤委員、佐藤委員、高橋委員、疋田委員

オブザーバ：個人情報保護委員会事務局、電気通信事業者協議会、日本データ通信協会、  
NTT ドコモ、KDDI、ソフトバンク、NTT BP

総務省：湯本課長、徳光課長、景山企画官

野村総合研究所（事務局）：小林、小島、南島

### 1. 「十分な匿名化」と匿名加工情報の作成方法に関する基準の検証

資料1-1、資料1-2、資料1-3に基づき事務局、ソフトバンク、NTT BP より説明。

#### 【質疑応答】

高橋委員：

各データ項目について、ソフトバンク（以下、SB）とNTT BP（以下、BP）の扱い方を確認したい。先ず位置について、SBは緯度・経度を3種類のメッシュに変換して表現、BPは拠点と表現している。BPの拠点とは意味的にひとまとまりなものと理解してよいか。つまり日比谷公園に複数のWi-Fi機器が設置してあったとしても、それらから取得される拠点としての情報は「日比谷公園」と表現されるという認識でいる。

NTT BP 大島オブザーバ：

認識に相違ない。

高橋委員：

次に情報取得の条件として、SBは建物に立ち寄ったという、分析にあたっての有用性に基づき15分間の同一メッシュでの滞留を条件にしているが、BPはそうした条件を定めておらず、通り過ぎた場合も含んでいる。時間については両者とも一番細かい粒度でも15分単位に集約している。最後に紐付く属性情報について、SBは年齢、性別、住所コード、趣味嗜好を、BPは言語情報のみを紐付けている。

SB 東海オブザーバ：

相違ない。

NTT BP 大島オブザーバ：

相違ない。

森主査：

SBのケースにおいて、500mの3メッシュについてはホップではなく、単一のメッシュ

内で分析しているため、250m メッシュ間、125m メッシュ間でのホップが資料上綺麗に整理されているという理解でよいか。

SB 東海オブザーバ：

相違ない。

佐藤委員：

ケースを拝見したところ予想以上に多様な加工結果となっており、用途をある程度特定しない限り、加工方法等のルール化は難しいと感じた。今後の議論の進め方として、用途を限定するか否かによって本調査研究のアウトプットとして示す加工方法・加工基準が異なると思うので、これについてどこかのタイミングで議論したい。

疋田委員：

資料1-2 ④<補足説明>「移動履歴の集計について」に関して、IDの重複について議論の際は、位置情報のカウントの仕方も同時に議論すべきと考える。SBの例をとると、Day4において、Bさんはa→bの移動をしたことになっているが、同時にa、bに滞在したということでもある。そのため、このケースにおいて、a、b、a→bと件数を数えるか、a→bと数えるかということはケースに応じて有用性が変わる。

また、事務局資料（資料1-1）において「商用のケースを見る限り、学校や病院などのプライバシー性が高いとされる場所があったとしても、十分に匿名化がされていれば、個人特定に繋がるとは考えにくい」とあるが、今回のSBの有楽町のケースでは確かにその通りだが、機微情報に該当する可能性が高いエリアは日本にいくつか存在するので、このようにストレートに言い切ってしまうとよいか個人的には難しいと感じる。

森主査：

疋田委員の2つ目の指摘は何を持って匿名化と見なすかという話と関係するが、個人の特定に繋がらなければ匿名化であると事務局は考えているため、こうした表現を行っていると考え。

伊藤委員：

事務局資料（資料1-1）ではK=2以上の場合であっても十分な匿名化がされていると示されているが、これはデータセットに含まれる変数のカテゴリに左右されると考える。例えば趣味嗜好についてかなり特殊な趣味がカテゴリに含まれる場合は、K=2以上であったとしても特定に繋がりがやすい。変数のデータカテゴリの特性に応じて加工した情報の安全性は変わってくる。

また、匿名加工情報の加工方法について今後議論していく上で、十分な匿名化の加工粒度を上げることが匿名加工情報に移行するものではないと考えている。そのためデータカテゴリの考え方も、十分な匿名化で良しとされたものがそのまま匿名加工情報の場合に適用されるのではなく、両者は分けて考えるべきと思う。

SB 東海オブザーバ：

ご指摘いただいた点の1つ目に関連して補足させていただくと、今回のケースにおいて

後から検証できるかもしれないということで、趣味嗜好のカテゴリに「マジック」という項目を入れている。これは有楽町の分析対象エリアにマジックの専門店があり、該当する趣味嗜好の人の移動履歴から専門店への滞留を検出できるか見てみたいという狙いに基づくものである。また、分析の便宜上、趣味嗜好を 10 種類にまとめたが、より細分化した場合の有用性は確かめてみたいと事業者としても考えている。

事務局：

事務局としてはアウトプットの方向性として、十分な匿名化は汎用的なもの、匿名加工情報は個別ケースによるものと打ち出したいと考えている。そうした枠組みの中で、正田委員にご指摘いただいたようなセンシティブな場所が多分に含まれる場合は、個別ケースによる判断が必要といった整理をするなど、ルール化をしていくにあたり、ケース依存性をどの程度認め、整理するか委員の皆さまからお知恵をお借りしたいと考えている。

正田委員：

資料 1-1 P16 の K=1 の場合の匿名加工情報の取り扱いに関して、商業地域においては問題ないと思うが、住宅地など滞留する人口数が一定数に満たない地域の場合は配慮が必要と考える。配慮の仕方としては人口が一定数確保されるまでメッシュ粒度を上げるといった調整が考えられる。

森主査：

センシティブな場所を含む場合の注意事項を考慮したとしても、十分な匿名化したデータセットが技術的に再特定・再識別のリスクが低減されており、汎用的なものであるという主張は受け入れられると考える。

高橋委員：

十分な匿名化についてエリアや時間の刻みを粗くすれば汎用的なものを作ることはできそうなものの、ホップ数を増やしたり、様々な属性データを紐付けようとしたら、その汎用性の維持には限界があることが分かった。ユースケースでの分析結果を踏まえ、ケースを考えないとルール化の議論はできないという領域まで検討が進んできている認識でいる。

森主査：

十分な匿名化は属性を紐付けた上で加工することを想定しているため、当然のことと思うがどうか。

高橋委員：

十分な匿名化が年齢などの基本 4 情報以外も紐付けてデータキューブにすることを指すものである認識に変わりないが、色々な属性を紐付けるほどにキューブは小さくなることから、ケースに応じて属性を取捨選択する必要があるという認識である。

森主査：

了。

佐藤委員：

個別ケースによる形でルールをとりまとめる場合、総務省の別の検討会において、独法や



行政機関の場合は民間事業者に加工方法も提案してもらおうという検討をした。そういう究極的な解もある。

正田委員：

汎用性を求めるのであれば、むしろ匿名加工情報を利用することが現状としては正解ではないか。十分な匿名化で、汎用的なものを作ろうとするとある程度の属性の紐付けが必要になるが、キューブが小さくなるため技術的に厳しい。また  $K=2$ ギリギリで加工した結果、紐付けた色々な属性から個人が特定されうるということにもなりかねない。加工の方法や程度は目的と有用性を考えながら議論すべきと考える。

伊藤委員：

資料1-1 P4にデータキューブという言葉があるが、これは統計用語で、利用者が集計項目を選択して利用する際の元となる保管データの体系を指す。なお、正田委員から指摘があったように十分な匿名化された情報であっても、マッチングされる危険性があるならば、何らかの追加的な措置を講じる検討はした方がよいと考える。その際の手法として、本協議会では検討の外に置かれているスワッピングやノイズをいれるといった措置の可能性を将来的にでも考えた方がよいのではないかと思う。

高橋委員：

攪乱的手法やサンプリングについて、今回は扱わないがその有用性を否定するものではないという認識でいる。また、 $K=2$ であれば問題ないという表現は、今回のユースケースを見る限りその通りと思うが、今後、言葉が独り歩きして誤解されることを防ぐため、戦略的な意味で、用いない方が適切だと考える。

## 2. 安全管理措置に関するルールの整理

資料2に基づき事務局より説明。

### 【質疑応答】

高橋委員：

十分な匿名化をした情報は安全管理措置が必要ない程度に加工したものであり、論点②については事務局案に賛成する。論点①については、森主査のご意見も伺いたいが、電気通信事業者に対し通秘部分への安全管理措置は既に求められており、それ以外の部分についても匿名加工情報と同じ水準を求めることは電気通信事業者だけの上乗せ規制として合理性がないのではないかと考える。

正田委員：

十分な匿名化の加工過程は個人情報保護法第20条とそれに関わる電気通信分野の事業者ガイドラインが直接的に適用される部分と認識しており、当該部分で安全管理措置に関するものは既にカバーされていると考える。仮に個人情報保護法第20条に基づく安全管理措置では足りないとして、匿名加工情報の基準を適用した場合、事業者が内部で個人情報を加工する際の安全管理措置はどうかといった形で話が広がっていく可能性が危惧される。

また、論点②については十分な匿名化した情報が再識別される技術的リスクが否定できない以上、何らかのモニタリングを実施することを推奨すべきと考える。

森主査：

十分な匿名化の場合、通秘を扱うので個人情報保護法第 20 条だけでなく、何かしら上乘せされたものがあるべきと考える。ここに関しては意見が分かれると考えており、一度引き取らせて頂き、事務局と相談させて頂きたい。なお、昨年度の調査研究において、十分な匿名化における管理運用体制の要件は加工のプロセスに対して行われるものと整理した。一方で、匿名加工情報における安全管理措置は加工のルールそのものに対する措置を想定しており、両者の保護対象は必ずしも一致しない。

### 3. 利用者の受容性調査

資料 3 に基づき事務局より説明。

【質疑応答】

伊藤委員：

CLT 調査についてサンプル数 100 名程度だと詳細なクロス分析は厳しいと考えるがどうか。

事務局：

本調査においては基本的にクロス集計による分析は想定していない。クロスしても 1 変数での集計である。

伊藤委員：

了。

正田委員：

資料 3 P34、P35 では意図的にリスクを提示した上で受容性を質問しているかが、そうした誘導的な質問の仕方は統計的に問題ないか。

伊藤委員：

この調査票の狙いは対象者を絞り込んで、より詳細な質問を行うところにあると認識しているが、そうした調査票の設計は一般的なものとする。

佐藤委員：

SB のケースにおいてはメッシュの大きさがポイントになっているが、この点を受容性調査においても明確に調べなくてよいか。調査研究の報告書作成にあたり、基準を揃えたほうが意義深いものになると考える。

森主査：

メッシュの違いについては一般の利用者には理解が難しいと考える。また同じ理由から匿名加工情報という示しはしないほうが良いと思う。

事務局：表現については今後調整する。

#### 4. 「十分な匿名化」と匿名加工情報の作成の方法等についての関係等の整理

資料4に基づき事務局より説明。

##### 【質疑応答】

森主査：

論点①は通秘に係るものを匿名加工情報とすることが可能かどうか、論点②は通秘に該当しないものを十分な匿名化する必要があるかどうか、である。事務局の整理では論点①については匿名加工情報にすることは認められず、十分な匿名化すべきとしており、論点②については十分な匿名化しなければならないものではないが、十分な匿名化することが望ましいとしている。

高橋委員：

論点①について内容は正しいが自明のことなので、改めて明言する必要があるかと感じる。元々は通秘に係る情報について、十分な匿名化と包括同意を行うことで、利活用の道が開けた。それ以外の場合については十分な議論が重ねられているわけではないので、断定的にできないと示すこと自体どうかと思う。また、論点②について、望ましいという表現は強すぎると思う。十分な匿名化が匿名加工情報を技術的な性質でカバーしていることは認識に相違ないが、両者の出自は異なるものであり、一方の手法を望ましいと表現する関係性にはないと思う。

佐藤委員：

そもそも十分な匿名化は通秘に対応するために生み出された概念であり、それを通秘以外にも適用するかという議論はナンセンスだと思う。十分な匿名化の対象や意図を明確化することで、対象外のものについても自ずから明らかになると思う。

森主査：

論点①について、通秘の場合、十分な匿名化をすれば利活用してもよいという議論をしただけであり、匿名加工情報の場合は駄目という議論はしていないということは尤もだが、当初、位置情報プライバシーレポートでは特定性低減データとの対比で十分な匿名化を議論してきた。結果、特定性低減データでは安全性が足りないと位置情報プライバシーレポートでは結論付けられたが、新たに登場した匿名加工情報についても、その安全性を議論すること自体は自然な流れと考える。その上で両者を比べると、匿名工情報は十分な匿名化ほど安全性が確保されていないため、事務局案が出てきていると考える。また、論点②については高橋委員、佐藤委員の指摘が尤もとして、再度事務局には検討頂きたい。

総務省：

森主査、両委員のご指摘はその通りだが、匿名加工情報という新しい概念が生まれたことで、それとの対応を行政として説明が求められていることが、論点設定の背景にあることをご理解いただきたい。なお、論点②については、非通秘であってもGPS等通秘と同様に精度の高い位置情報が取得できてしまう問題や同じ電気通信事業者でも強力な契約者情報を有する事業者とそうでない事業者が混在している問題を考慮した上で、次回引き続き議論

させていただきたい。

正田委員：

論点②に関し安全管理措置も考慮すると、提供先の管理が求められる匿名加工情報の方が望ましいのではないかと考える。十分な匿名化の場合の方が望ましいと言い切るのではなく、ユースケースごとに事業者が選択すべきとした方が望ましい表現ではないか。

位置情報に関するプライバシーの適切な保護と社会的活用の両立に向けた調査研究  
第4回協議会 議事概要

日時：2017年3月14日 13:00-15:10

場所：総務省 8F 第4特別会議室

出席者：

委員：森主査、伊藤委員、佐藤委員、高橋委員、疋田委員  
オブザーバ：個人情報保護委員会事務局、日本データ通信協会、  
NTT ドコモ、KDDI、ソフトバンク、NTT BP  
総務省：湯本課長、徳光課長、渡邊専門職  
野村総合研究所（事務局）：小林、小島、南島

## 1. 安全管理措置に関するルールの整理

資料1に基づき事務局より説明。

### 【質疑応答】

森主査：

個人情報保護法第21条関係に関しては、「委託先の監督」も合わせて資料に記載したい。

佐藤委員：

加工後の情報に関して、十分な匿名化で特段の安全管理措置を講じることを求めない点は、「十分な匿名化」がどの程度十分に匿名化をするのかの観点で、懸念がある。「十分な匿名化」の定義を明確にしなければ、「求めない」という判断はできない。

森主査：

「十分な匿名化」は、一意に分かるデータを残さない特徴がある。この特徴に鑑み、個人の特定の観点で、安全管理措置をしようという話にはならないと考えている。

疋田委員：

安全管理措置に関しては、「十分な匿名化」と匿名加工情報のみを比較するのではなく、事業者等が保有する一般的な個人情報等も合わせて比較・評価をしておくべきではないか。

この2つの情報のみを考慮すると、全体としての安全管理措置に係る状況の把握（例えば、「十分な匿名化」や匿名加工情報でなくとも、事業者は不正アクセス防止や持出し制限を実施している）ができなくなる懸念がある。

森主査：

一般的な情報の匿名化に係る法制度は存在しない。また、改正個人情報保護法によって匿名加工情報のスキームは確立されている。これらを踏まえて、本資料では「十分な匿名化」と匿名加工情報の比較を行っている。

なお、昨年度はPIA等によってデータ作成における「動的」な安全管理措置を検討したが、本年度はその管理に係る「静的」な措置に対する検討を入れることを目的としている。

伊藤委員：

加工中のデータの安全管理、十分に匿名化されたデータの管理共に、k-匿名性がどの程度担保されているのかによって決められるものとする。kの値を大きくすれば加工後の安全管理措置は不要等の整理もし得る。

事務局：

事務局としては、本ルールでは、k-匿名性と9つの定性評価指標による総合評価で「十分な匿名化」をすることとしており、k-匿名性の程度で、安全管理措置を判断するものではないものと考えている。

森主査：

以上の議論をまとめ、加工中のデータの安全管理については、匿名加工情報の法制度を参照した安全管理措置を整備し、加工後のデータについては、安全管理措置を特段講じることは求めないレベルまで十分に匿名化をする、という整理をしたい。

## 2. 「十分な匿名化」と匿名加工情報の作成の方法等についての関係等の整理

資料2に基づき事務局より説明。

### 【質疑応答】

高橋委員：

通信の秘密に該当する位置情報を匿名加工情報として利用することは通常認めないとする点は、賛同する。

通信の秘密に該当しない位置情報を利用し、かつ、契約者情報や高精度の位置情報を利用する場合に「十分な匿名化」に準じた加工をすることは、必然性はないが、加工する方が望ましいと考えている。具体的なルールを適用した際に、リスクと手間とのバランスがしっかりと取られていれば良いと考える。

正田委員：

契約者情報や高精度の位置情報の利用において、「十分な匿名化」に準じた加工をすることは、プライバシーの観点では望ましいが、そもそも位置情報が個人情報に該当しない場合まで適用するのは困難と考える。例えばオープンストリートマップの作成に利用されているデータは、ある地点の高精度な位置情報を与えているが、これ自体は個人情報ではない。

また、資料2 P.11の「十分な匿名化」をした情報の個人情報保護法における取り扱いにおいて、全ての位置情報が個人情報になる訳ではない点に鑑み、まずは当該情報が個人情報に該当するか否かから整理をすることが望ましいと考える。

森主査：

個人情報になる位置情報を考えた場合、これらを加工した結果は、匿名加工情報、統計情報の二択となるのかを考える必要がある。

そうではないケースがあるのであれば、資料2 P.11のロジック分岐は2つで済むと整理することは困難である。

個人情報保護委員会事務局：

位置情報に関してそういう場合があるのかすぐにイメージできないが、他の個人情報においては、加工の結果、統計情報でも匿名加工情報でもないものになるものもあると理解している。

高橋委員：

位置情報においても、個人情報委員会事務局指摘の事象はあると考えている。

伊藤委員：

「統計」という言葉は非常に広い概念を含むため、整理をするにおいても、定義を明確にしておく必要がある。個人情報保護委員会の事務局レポートにおいても、統計の定義の脚注の中で、統計は、「統計作成者の責任のもとで作成された」という言葉がある。このような定義を抜いて整理をすることは、困難であると考えている。

佐藤委員：

伊藤委員の「統計」の定義に加え、匿名加工情報の加工基準にも定義が必要である。加工基準を個人情報保護委員会の出した基準に限定するのであれば定義は明確だが、実際の加工基準は、業界団体等を経て業界ごとにカスタマイズされることになる。この場合、どの加工基準が妥当なのかは曖昧となる。フローチャートでの整理は利用者の理解を容易にし得るが、上述の定義を明確にしておかないと、かえって混乱を来すと考える。

森主査：

通信の秘密に該当する位置情報を匿名加工情報として利用することは通常認めない点、通信の秘密に該当しない位置情報を利用し、かつ、契約者情報や高精度の位置情報を利用する場合に「十分な匿名化」に準じた加工をする点は、合意が取れたと考える。

一方で、「十分な匿名化」をした情報に関する個人情報保護法上の取り扱いは、上述の通り、言葉の定義を明確にした上で、フローチャート等で整理をするのであれば、更なる精査が必要と考える。

### 3. 電気通信事業者が取り扱う位置情報に関するルールの整理

資料3-1、資料3-2に基づき事務局より説明。

#### 【質疑応答】

高橋委員：

「十分な匿名化」におけるk-匿名性の議論について、kの値が十分に大きい場合には、リスク評価の工数等を軽減する議論が昨年度の実証事業ではされていたが、本年度の整理案からは除かれている。直近の事業者の想定ユースケースを考えたルールとしては、これでよいのかも知れない。ただし、将来的に多様に「十分な匿名化」が扱われることを想定した際、この整理で本当に良いのかを考えたい。

kの値を何に設定しても、データセットに依存する形で懸念は生じ得る。ただし、その結果、一概に9つの出口要件の指標を適用するとなれば、例えば、場所の特性の観点で見る

と、全ての病院をデータセットから除外する等の処理をしなければならなくなる。運用面も踏まえて、どの程度の加工や処理を、どの場合に実施するか、を考えておくことが望ましい。

正田委員：

「十分な匿名化」が、 $k$ -匿名性と同義であるような定義になっていることに違和感がある。 $k$ -匿名化は「十分な匿名化」された情報を作成する1つの手法ではあるが、「十分な匿名化」された情報を作成する手段は、 $k$ -匿名化以外にもあり得る。なお、現時点での有効な手法として取り上げる分には問題ないと考える。

場所の特性の例として病院を挙げるのは、範囲が広すぎるように感じる。クリニック等の一般的な病院の位置情報自体にはリスクはなく、特定疾患の処方がある病院のみ（例えば国立がん研究センター等）に、ターゲットを絞れば良いと考えている。

本調査事業は、電気通信事業者に係るルールの検討であることに鑑み、電気通信事業ならではの特性をルールに入れ込むことが望ましい。例えば対象者数に関し、この数が、契約者数を意図しているのか、端末数を意図しているのかは十分に配慮する必要がある。今日、携帯電話等の端末は、1人で2台以上保有している利用者も多く、仮に端末数ベースで集計し、 $k$ が2以上であったとしても、利用者数ベースでは $k$ が1となる可能性もある。こうした電気通信事業の特性を踏まえたルールが、ガイドラインに記載されているべきと考える。

佐藤委員：

$k$ -匿名性と $k$ -匿名化の言葉の定義と使い分けは、ルールの中で徹底されたい。 $k$ -匿名化は一般用語、 $k$ -匿名性は評価指標としての言葉である。

その上で、加工基準を考えると、上述の病院も然り、個人情報だけでなく、プライバシーの観点で情報の削除をする点が議論になっていると認識している。情報の利用に多様性を持たせなければ、事業者が運用しづらいルールになってしまうことを懸念している。例えば、欧州のカーナビ事業者は、出発地周辺と到着地周辺の移動情報を集計しており、走行のパターンを幅広く収集したサービスを提供している。

伊藤委員：

出口要件にある「母集団の特性」に関して、「母集団」は統計では標本の元になる集団を指す言葉である。出口要件で処理する以前に、一部抽出等がなされているデータセットに対して、「母集団」という言葉を利用すると、混乱を来す可能性がある。

その上で、出口要件に対する議論がされているが、入口要件で実施している一部抽出は、最終的なデータの特性に大きな影響を与えることを考えると、入口要件や一部抽出の手法に関するルール等も検討する必要があると考えている。特に $k$ -匿名化における $k$ の値を一義に決められないのであれば、一部抽出の重要性が増すのではないかと考える。

例えば、母集団のデータセットに含まれた趣味嗜好情報の分布や特性を踏まえて、一部抽出をしてデータセットを作成し、その上で出口要件にかけるような処理も検討する必要があると考えている。



ソフトバンク 東海オブザーバ：

ソフトバンクとしては、法人ユーザーやオプトアウト希望者等を対象外とした上で、全てのデータを利用しており、分析目的に応じて一部抽出の仕方を変えることは、現時点では想定していない（ただしサービス別等での抽出は存在）。

佐藤委員：

ルールの方策という観点で見て、GPSに基づく位置情報、基地局を介して得られた位置情報、その他手法に基づく位置情報を一まとめにして整理することが可能かは、検討が必要と考える。本調査事業に参加していない、前提知識のない人が、ルールを見て、要らぬ拡大解釈等をしないように定義をする必要があると考えている。

位置情報は特性として、誤差を含んでおり、また、この誤差はランダムに生じる。この誤差を同見するのに関しても、ルールの中で言及しておくことも必要ではないかと考えている。

森主査

佐藤委員の指摘は重要ではあるが、本事業の「十分な匿名化」の論点は、あくまで通信の秘密に係る位置情報であるため、本事業の中での優先順位はそれ程高いとは考えていない。

先般の高橋委員の指摘に関し、 $k$ が大きい場合は出口要件の多くを簡素化／省略できるとした場合の $k$ の値を具体化できないか。

高橋委員：

データセットの特性等が影響するため、ケースバイケースの判断にならざるを得ない。なお、9つの評価指標を省略することは考えておらず、 $k$ の値に応じて、全てを厳密に評価しなくても済む状況を作ることを考えている。

正田委員：

この $k$ の値であれば、評価指標のこの部分は、ある程度見るだけで良いとし、その代わり、別の指標はしっかり見ないといけない、というバランスを取ることが重要であろう。ただし、全ての評価指標は何らかしらの形で見ることが必要である。

その意味で、 $k$ の値が大きい場合と小さい場合で、加工や処理方法を分けるのではなく、 $k$ の値に鑑み、どうバランスを取るか、という処理を実施していくべきと考える。

高橋委員：

例えば、位置の精度に関して、メッシュの粒度を粗くすれば、そのメッシュの中に自宅が残っていても、それが自宅とは判断が付かないため、問題ないと整理できる。ただし、自宅があること自体を考慮しなくても良いわけではない。

事務局：

ルールに記載として、出口要件として、総合的な判断を行うことを意図していたが、現状のルールでは、9つの出口要件を実施した上で、総合的な判断があるように見えてしまっているので書きぶりを見直す。

高橋委員：

ルールを個別に見た際に、場所の特性として、病院が挙げられた点に対し、実務上の処理が困難であること、また全体として見た際に、厳密に 9 つの出口要件を処理しなければならないような書き方は、運用上困難であることを懸念していた。k の値も踏まえて、9 つの評価指標を、選択的かつ総合的に用いるのであれば、運用可能と考える。

なお、これまでも k-匿名性を用いて議論を進めてきたが、k-匿名性という言葉を使うことは、今後は控えた方が良く考えている。k-匿名化は、「準識別子が k 個存在する」状態を指しているが、本検討においては準識別子かどうかという議論は行っていない。k を用いた議論や説明は分かり易いが、他の識者が見た際にあらぬ誤解をされる可能性もあると感じている。

高橋委員：

匿名加工情報について、「十分な匿名化」との差異を事務局はどのように考えているか。

事務局：

① k-匿名性を前提として処理するのが「十分な匿名化」であり、匿名加工情報はその限りではない点、②「十分な匿名化」では付帯情報は限定列挙されている一方、匿名加工情報は柔軟に属性情報を付与できる点の 2 点が、大きな違いと認識している。

高橋委員：

上記認識の違いを前提に、匿名加工情報のルール案を考えると、このルールを適用した結果は、かなり「表面的」なイメージがあり、「十分な匿名化」と同等のレベルにあるものと感じられる。

正田委員：

個人情報保護委員会事務局レポートの匿名加工情報に係るユースケースはかなり詳しいデータセットの開示を認めている。一方、本事業で提示されたルールは、厳格に適用すると、「十分な匿名化」の加工レベルに近い生成物が作られるように感じている。

森主査：

例えば匿名加工情報のルールのユースケースについて、日時に係る情報を丸め、その代わり、他の属性情報を付加する等を検討しても良いものと考えられる。

なお、匿名加工情報のルールの資料 3-2 P.15 において、趣味嗜好情報に関し、5 号と 4 号の記載が混同している。

事務局：

指摘頂いた事項は、記載ミスにつき、訂正する。

高橋委員：

匿名加工情報で、どこまでプライバシーに配慮するのは論点であると考えている。例えば、個人情報保護法上、病院など要配慮個人情報問題は問題視されていないが、現状のルール案ではプライバシーの要素が入っている。総務省として、この点をどのように考えているのかを知りたい。

総務省 湯本課長：

議論の整理として、「十分な匿名化」、匿名加工情報ともに、総合的な判断が必要であり、定量的に一義に決めることは困難であることを理解した。その上で、未だ想定されるケースも明らかにはなっていないタイミングで、最適なルールを検討することは困難であると考えている。

よって、電気通信事業者向けの匿名加工情報の第一弾のルールとしては、十分な匿名化の基準に基づきルールを策定し、半年、1年後の然るべきタイミングで、再度ルールを検討していく必要があると考えている。

現状の匿名加工情報は、個人情報保護委員会の事務局レポートにあるレベルよりも厳しい基準が採択されているという指摘は、印象としては、私もそのような感じと理解している。  
高橋委員：

まずは安全を担保したルールから始める、という理解をした。

森主査：

場所の特性については、要配慮個人情報を含んでいる点が論点になると理解した。

総務省 湯本課長：

情報通信事業の分野で、要配慮個人情報を、どの程度留意するかは、やりながら考えていく形にせざるを得ないと考えている。ただし、現時点では配慮は必須であると感じている。

正田委員

要配慮個人情報に関わる論点として、欧州では、例えば宗教上の施設の位置情報への配慮が求められている。その意味では、要配慮個人情報は匿名加工情報の中で検討すべき論点になりうると考える。

森主査：

本協議会で議論しきれなかった部分については、別途メール等で意見を頂く形としたい。

#### 4. 利用者の受容性調査の結果

資料4に基づき事務局より説明。

##### 【質疑応答】

伊藤委員：

スクリーニング前の母集団とスクリーニング後の標本集団の傾向は、一概に同じとは言えず、調査項目によっては傾向が逆に出ているものも存在する。よって、報告書として取り纏める際には、プライバシーリスクに対して緩和された標本集団を用いて分析をしていることを明確にしておくことが望ましい。

正田委員：

消費者の声を分析した意義は大きく、参考になる。例えば、居住地に対するリスクに係る分析的な判断と利用者の感覚とはずれるなどが明確になっている。どのようなことを実施した際に、消費者心情と離れる可能性があるのか、等は注意喚起として、ガイドライン等に何らかしらの形で導入するのが望ましい。