個人情報保護

ネットワークと情報セキュリティ12

菊池　浩明

Contents

- 個人情報保護に関する法制度
- 匿名加工技術
- 秘匿計算技術
  - プライバシー保護データマイニング

個人情報保護に関する法制度

ビッグデータ利活用とリスク

「個人情報」とは

- 個人情報（法2条）
  - 生存する個人に関する情報であって，当該情報に含まれる氏名，生年月日その他によって特定の個人を識別することができるもの．
  - 他の情報と容易に照合することができ，それにより特定の個人を識別することができることになるもの（*容易照合性*），（*提供元基準*）

問題

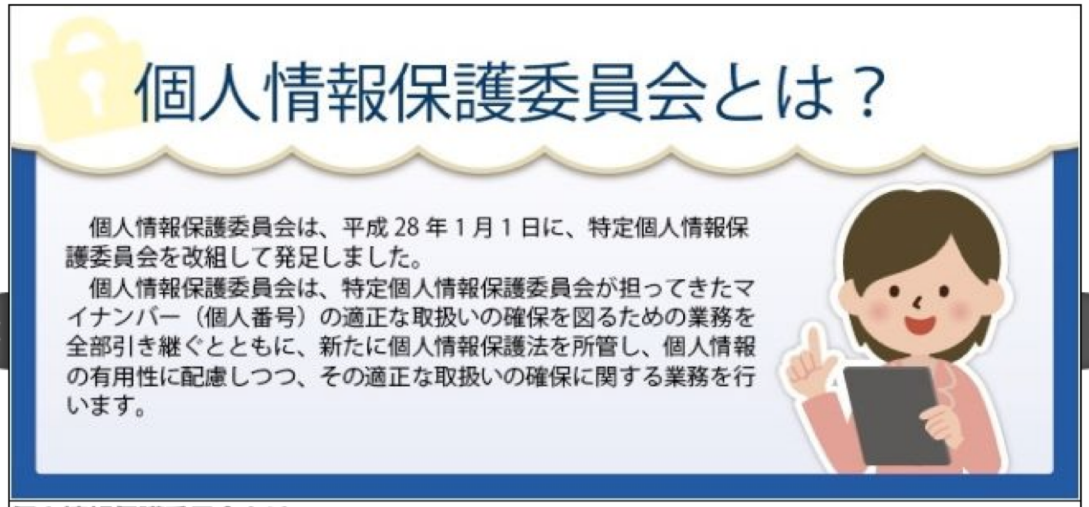
- 個人情報はどれか？
  - 1. 住所
  - 2. 顔画像
  - 3. スマートフォンのMACアドレス
  - 4. 訃報
  - 5. 電車の乗降履歴

個人情報保護委員会

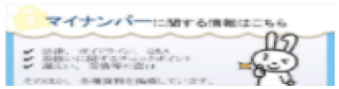
Personal Information Protection Commission
法人番号：4000012010025

文字サイズ変更 標準 大きめ 検索

ホーム | 委員会の概要 | 個人情報保護法について | マイナンバーについて | 委員会の活動 | お知らせ | お問合せ・申請

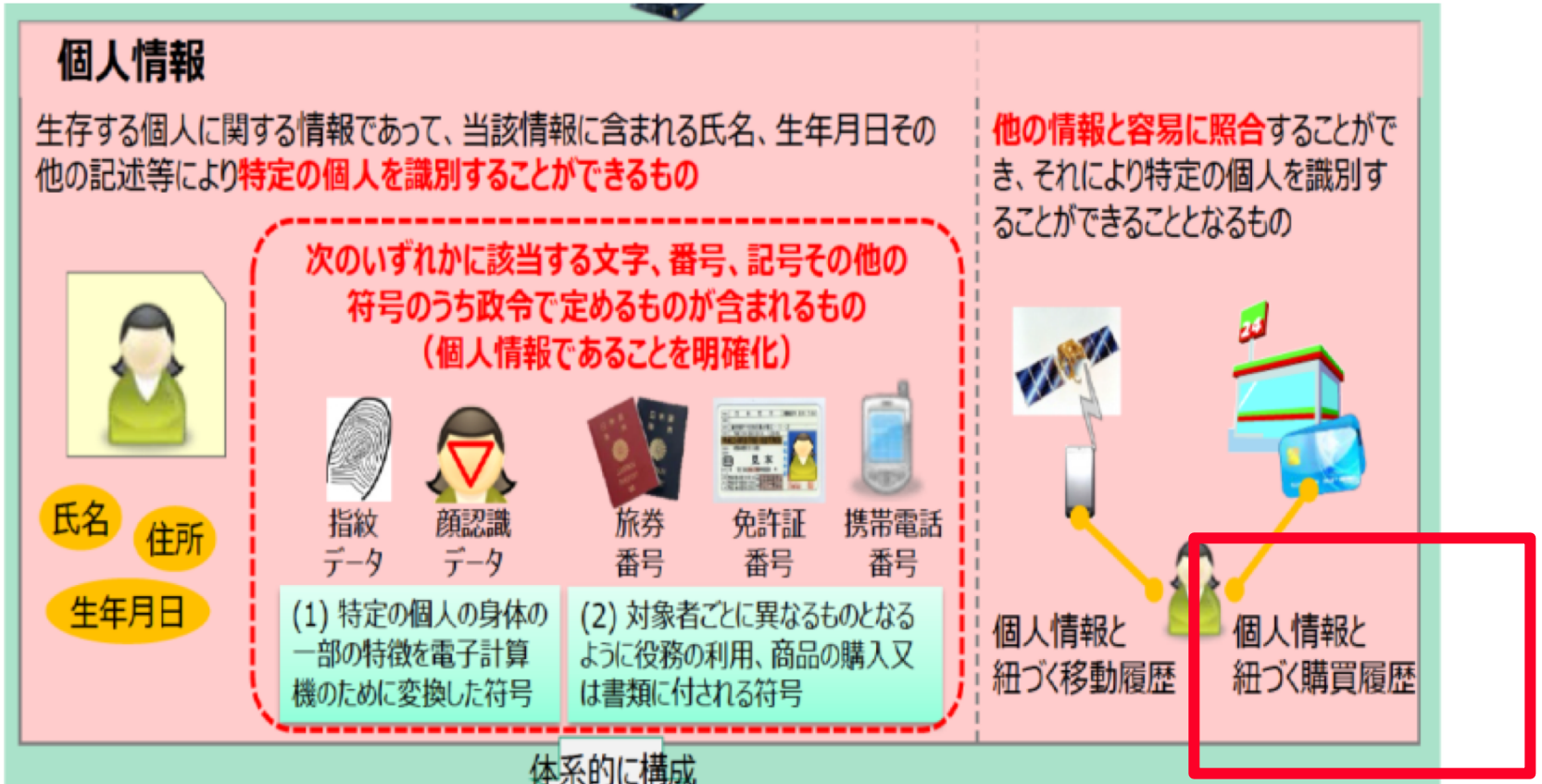


個人情報保護委員会とは



- ホーム
- 委員会の概要
  - 個人情報保護委員会について
  - 委員長・委員紹介
  - 広報
  - キッズページ
- 個人情報保護法について
  - 法令・ガイドライン等
  - 漏えい等の対応（個人情報）
  - 中小企業サポートページ（個人情報保護法）
  - 認定個人情報保護団体
  - 匿名加工情報・非識別加工情報

個人情報の範囲の明確化



	個人情報（旧法2条1項）		（新規）
	従来型の個人情報 （法2条1項1号）	個人識別符号 （法2条1項2号，法2条2項1号2号）	要配慮個人情報 （法2条3項）
従来法（旧）	生存する個人に関する情報であって、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの		
改正法（新）	「その他の記述等」についての定め方が詳細になったものの**実質的**な変更なし	個人識別符号が含まれるもの	本人の人種、信条、社会的身分、病歴、犯罪の経歴、犯罪により害を被った事実その他本人に対する不当な差別、偏見その他の不利益が生じないようにその取扱いに特に配慮を要するものとして政令で定める記述等が含まれる個人情報
例	本人の氏名，生年月日、連絡先（住所・居所・電話番号・メールアドレス）、会社における職位又は所属に関する情報について、それらと本人の氏名を組み	1号個人識別符号：DNA，顔貌，虹彩，声紋，歩容等 2号個人識別符号：パスポート番号，基礎年金番号，免許証番号，住民票コード，マイナンバー等	人種、信条、社会的身分、病歴、犯罪の経歴、犯罪により害を被った事実（法2条3項），健康診断等の結果（施行令2条2号）等。

保護法改正のポイント

- 1. 個人情報の定義の明確化
  - 個人識別符号（第2条2項），要配慮個人情報（第2条3項）
- 2. 個人情報の有用性確保
  - 匿名加工情報（第2条9項）
  - 認定個人情報保護団体による個人情報保護指針（第53条）
- 3. 個人情報の保護を強化
  - 個人情報データベース等提供罪（第83条）
- 4. 個人情報保護委員会の新設と権限
  - 新設（第5章）
- 5. 個人情報の取り扱いのグローバル化
  - 国境を越えた適用（第75条），外国執行局への情報提供（第24条）
- 6. オプトアウトなど
  - 届出厳格化（第23条2項），利用目的の変更禁止（第15条2項）

個人情報保護に関する法律体系
(個人情報保護委員会資料)

個人情報保護に関する法律・ガイドラインの体系イメージ

民間分野

ガイドライン
（通則編・外国第三者提供編・確認記録義務編・匿名加工情報編）
（*2）

個人情報保護法（*1）
（4～7章：個人情報取扱事業者等の義務、罰則等）
（対象：民間事業者）

公的分野

行政機関個人情報保護法（*3）	独立行政法人個人情報保護法（*4）	個人情報保護条例（*5）
（対象：国の行政機関）	（対象：独立行政法人等）	（対象：地方公共団体等）

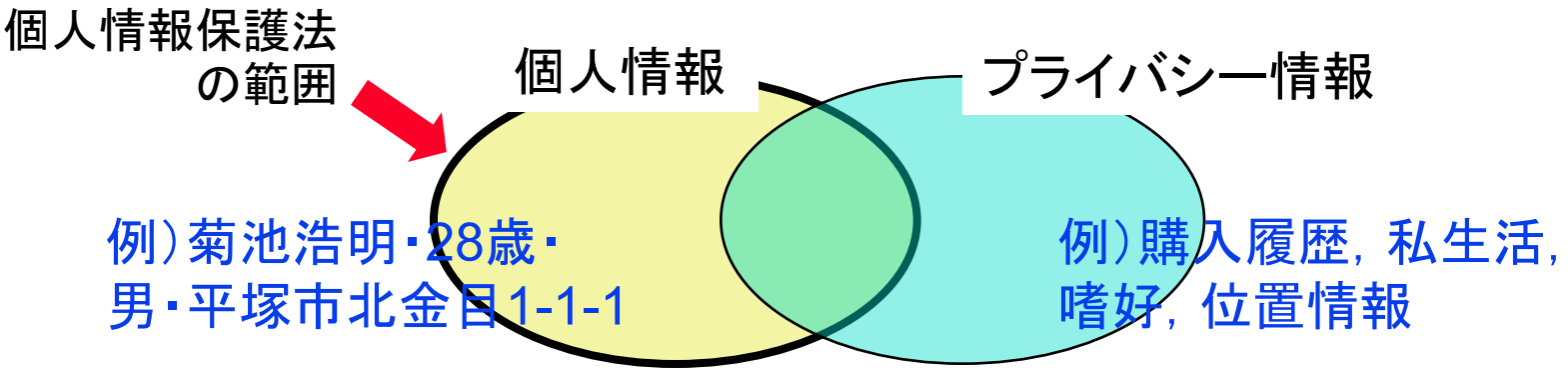
個人情報保護法（*1）
（1～3章：基本理念、国及び地方公共団体の責務・個人情報保護施策等）

個人情報の保護に関する基本方針

（*1）個人情報の保護に関する法律
（*2）金融関連分野・医療関連分野・情報通信関連分野等においては、別途のガイドライン等がある。
（*3）行政機関の保有する個人情報の保護に関する法律
（*4）独立行政法人等の保有する個人情報の保護に関する法律
（*5）個人情報保護条例の中には、公的分野における個人情報の取扱いに関する各種規定に加えて、事業者の一般的責務等に関する規定や、地方公共団体の施策への協力に関する規定等を設けているものもある。

個人情報 ≠ プライバシー情報

- 個人情報
  - 生存する個人に関する情報であって、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの
- プライバシー
  - 私生活上の事項をみだりに公開されない法的な保障と権利.（憲法13条個人の尊重）
  - 個人情報をコントロールする権利
  - 他人に知られたくない（主観的）情報



位置のプライバシー

- 国内
  - 総務省「スマートフォンプライバシーイニシアティブⅢ」
  - （単体では）個人識別性を有しない
  - 同一IDに紐づけて行動履歴を集積する場合は個人情報に準じる
- 海外
  - EUでは個人データとみなされている
  - 例）英国 Renew London社のMACアドレス収集をロンドン市が中止する通達
  - GDPRではクッキーもPII（個人識別情報）

OS が生成する ID（Android ID）、独自端末識別番号（UDID）、加入者識別 ID（IMSI）、IC カード識別番号（ICCID）、端末識別 ID（IMEI）、MAC アドレス等	× 端末交換や契約変更をしない限り変更が困難	・スマートフォンの OS やシステムプログラム、SIM カード、端末そのもの等に割り振られ管理される。利用者は端末交換や契約変更をしない限り変更困難。 ・単体では個人識別性を有しない。他の情報と容易に照合できる場合、個人識別性を獲得する。 ・同一 ID に紐付けて行動履歴や位置情報を集積す

容易照合性（個人と紐づく履歴）

- 顧客データベース M（個人情報）

顧客 ID	性別	誕生日	国籍
12360	M	1876/2/24	Australia
12361	F	1954/2/14	Belgium
12362	F	1963/12/2	Belgium
12364	F	1960/9/16	Belgium

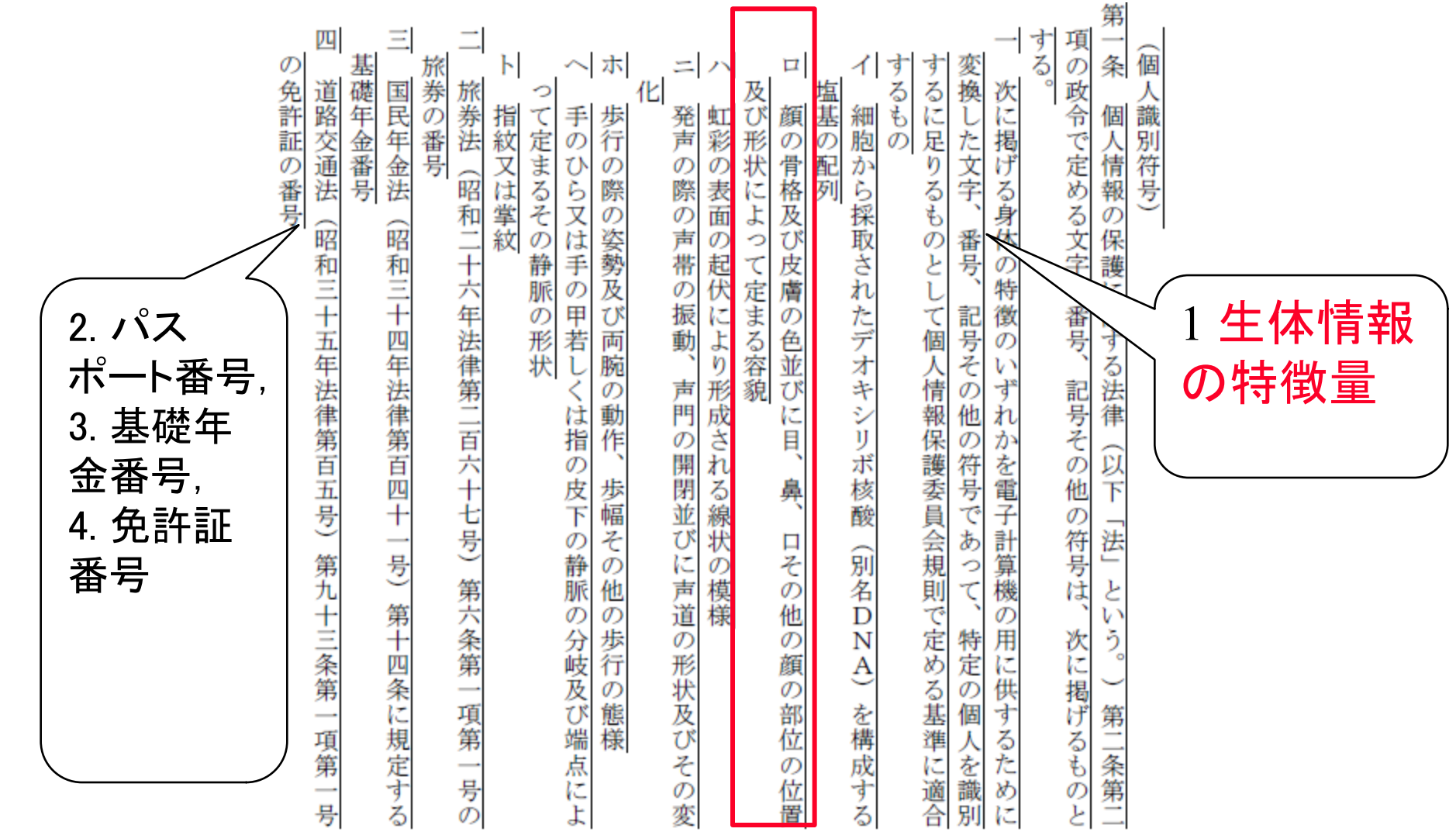
- 購買履歴 T

顧客 ID	伝票ID	日付	時刻	商品ID	単価
12362	544203	2011/2/17	10:30	21913	3.75
12362	544203	2011/2/17	10:30	22431	1.95
12361	545017	2011/2/25	13:51	22630	1.95
12361	545017	2011/2/25	13:51	22326	2.95

個人データの定義

<table>
<tr><td rowspan="4">個人情報
（法2条2項，個人を識別できる情報）</td><td colspan="3">・名刺の束
・アンケート用紙</td></tr>
<tr><td rowspan="3">個人データ
（法2条6項，検索などを可能にしたデータベースの構成要素）</td><td colspan="2">・委託されて入力した個人データ</td></tr>
<tr><td rowspan="2">保有個人データ
（法2条7項，開示，訂正の権限を持つもの）</td><td>・6か月以内
・業務に要していない</td></tr>
<tr><td>個人データベース（法2条4項）</td></tr>
</table>

「個人識別符号」施行令(2016年)



（個人識別符号）

第一条　個人情報の保護に関する法律（以下「法」という。）第二条第二項の政令で定める文字、番号、記号その他の符号は、次に掲げるものとする。

一　次に掲げる身体の特徴のいずれかを電子計算機の用に供するために変換した文字、番号、記号その他の符号であって、特定の個人を識別するに足りるものとして個人情報保護委員会規則で定める基準に適合するもの

イ　細胞から採取されたデオキシリボ核酸（別名ＤＮＡ）を構成する塩基の配列

ロ　顔の骨格及び皮膚の色並びに目、鼻、口その他の顔の部位の位置及び形状によって定まる容貌

ハ　虹彩の表面の起伏により形成される線状の模様

ニ　発声の際の声帯の振動、声門の開閉並びに声道の形状及びその変化

ホ　歩行の際の姿勢及び両腕の動作、歩幅その他の歩行の態様

ヘ　手のひら又は手の甲若しくは指の皮下の静脈の分岐及び端点によって定まるその静脈の形状

ト　指紋又は掌紋

二　旅券法（昭和二十六年法律第二百六十七号）第六条第一項第一号の旅券の番号

三　国民年金法（昭和三十四年法律第百四十一号）第十四条に規定する基礎年金番号

四　道路交通法（昭和三十五年法律第百五号）第九十三条第一項第一号の免許証の番号

1 生体情報の特徴量

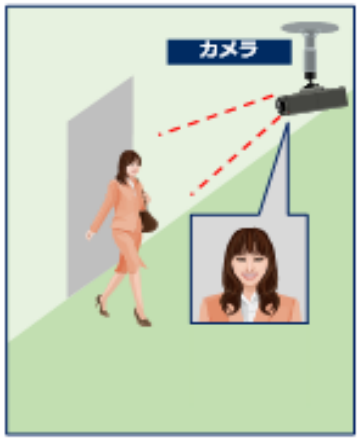
2. パスポート番号,
3. 基礎年金番号,
4. 免許証番号

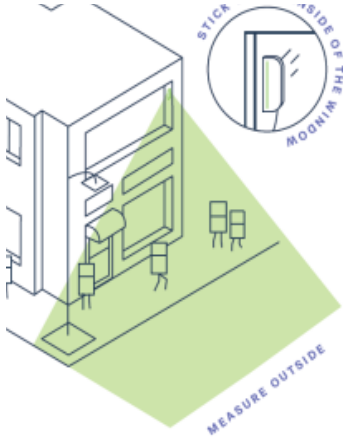
個人情報になりそうな情報

- 個人情報（一号）
  - スマートテレビの視聴履歴
  - スマートホームのパーソナライズ嗜好記録
  - ライドシェアの乗車履歴
  - オンラインショッピングの購買記録
  - スマートメーターの電力消費履歴
- 個人識別符号（二号）
  - Webカメラのストリームデータ
  - ウェアラブルカメラの撮影データ
  - 歩容認証センサーデータ
  - コールセンターの音声データ
  - 虹彩認証，静脈認証のテンプレート

カメラ応用4類型

1. 店舗設置カメラ　2. 野外設置カメラ　3. 公共空間設置カメラ4. リピート分析カメラ





- 来店者の年齢性別などの人物属性
- 属性分析後消去

- 通行者のカウント

- 人物の移動履歴（人流情報）のみ.
アイコンで可視化

- 顧客の来店頻度

問）　どれが個人情報（個人情報データベース）でしょう？

個人情報保護委員会「ガイドラインに関するQ&A」

「個人情報の保護に関する法律についてのガイドライン」及び「個人データの漏えい等の事案が発生した場合等の対応について」に関するQ＆A

Q1－13 カメラ画像から抽出した性別や年齢といった属性情報や、人物を全身のシルエット画像に置き換えて作成した移動軌跡データ（人流データ）は、個人情報に該当 しますか。

A1－13 個人情報とは、特定の個人を識別することができる情報をいいます。性別、年 齢、又は全身のシルエット画像等による移動軌跡データのみであれば、抽出元の本人を判別可能なカメラ画像や個人識別符号等本人を識別することができる情報と容易に照合 することができる場合を除き、個人情報には該当しません。

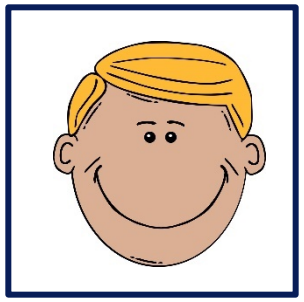
Q1－38 防犯カメラやビデオカメラなどで記録された映像情報は、本人が判別できる 映像であれば、個人情報データベース等に該当しますか。

A1－38 本人が判別できる映像情報であれば、個人情報に該当しますが、特定の個人情 報を検索することができるように「体系的に構成」されたものでない限り、個人情報データベース等には該当しないと解されます。すなわち、記録した日時について検索する ことは可能であっても、特定の個人に係る映像情報について検索することができない場 合には、個人情報データベース等には該当しないと解され ま

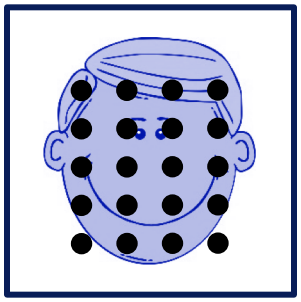
特徴量とは何か？

- 多重変動分析法

「不可逆処理」
元の画像は復元不能



撮影情報



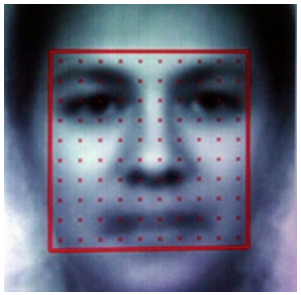
サンプリング

3	12	18	2
5	20	24	8
3	3	5	7
0	4	5	1

特徴量

個人を追跡したり，行動履歴を把握される可能性あり

照合

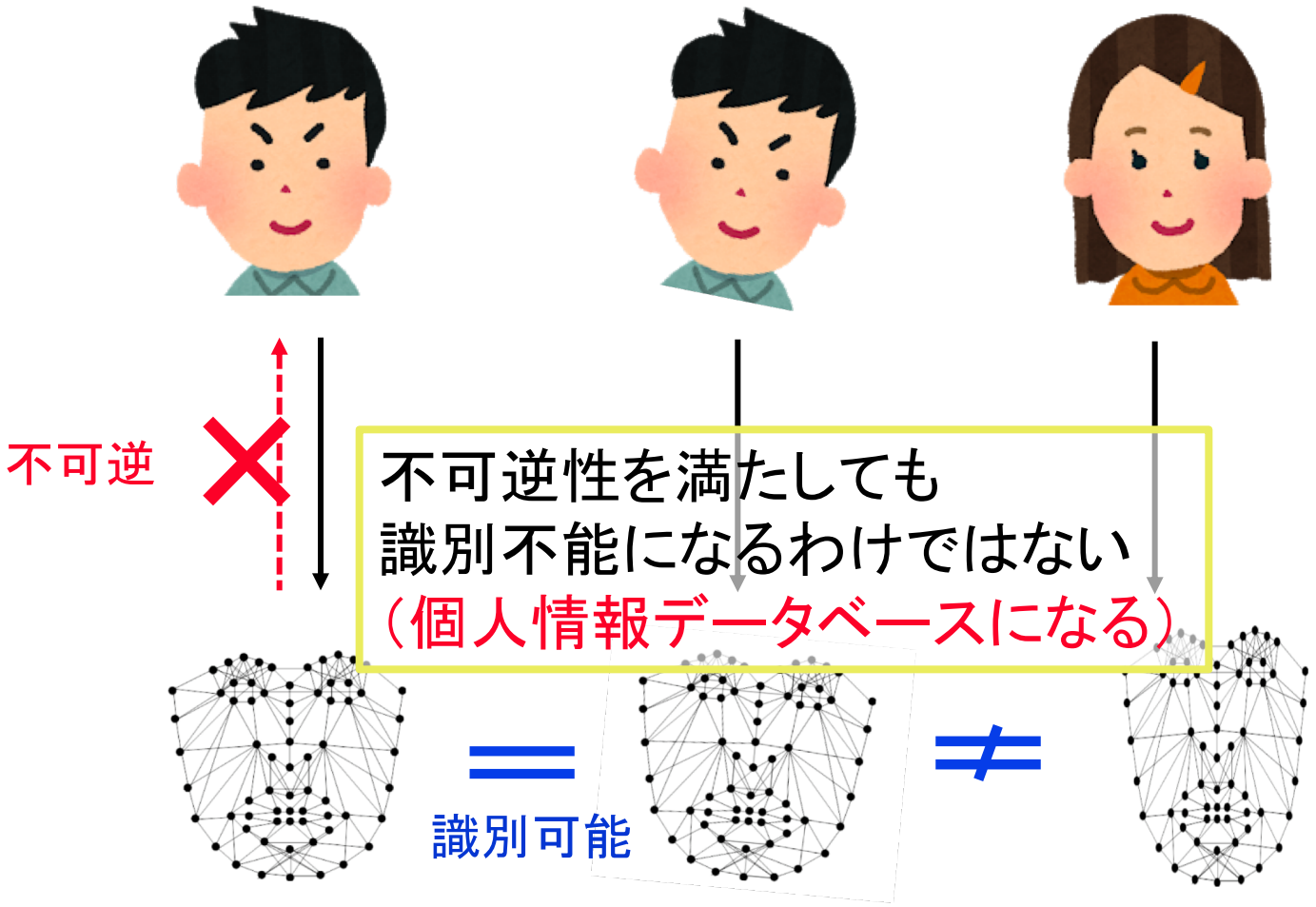


標準画像

5	10	20	0
5	20	25	10
5	5	5	10
0	5	5	0

http://www.itmedia.co.jp/news/0311/06/nj00_hitachi_face.html

不可逆性と識別可能性



利用（識別）を停止するには？

- 1. 個人情報のオプトアウト請求
  - ❑顔特徴量を提出して個人情報削除要求
  - ❑誤認識による道ずれ削除のリスク

- 2. 取得の否定
  - ❑“Don’t Truck” カード
  - ❑マスク
  - ❑プライバシーバイザー（NII越前功研究室）

高崎晴夫,「個人情報保護に係る法制度をめぐるEUの状況」, 情報処理, Vol. 55, 2014.

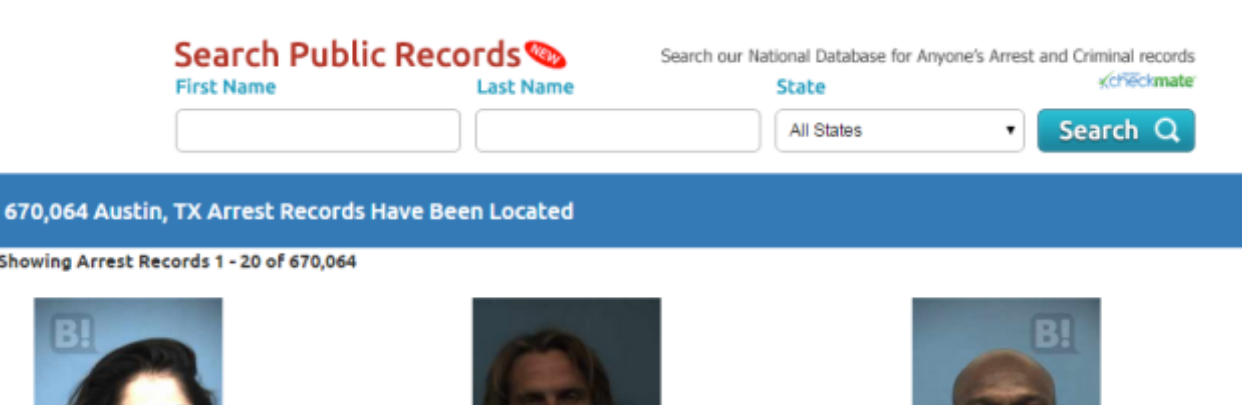
1. 欧州の動向

- EU一般データ保護規制 GDPR
  - 1995年のEUデータ保護指令の改正
  - 2014 LIBE委員会
    - »（Civil Liberties, Justice and home affairs)
  - 法的拘束力の強化(「指令directive」から「規制regulation」へ.
  - 個人情報の拡大
    - » 生体データ, 遺伝的データ, 健康データ
    - » 仮名化データ (pseudonymised), プロファイリング
  - データ主体の権利の強化
    - » 忘れられる権利, 削除権
    - » プロファイリングへの異議申し立て

石井夏生利，「アメリカのプライバシー保護に関する動向」，情報処理，Vol. 55, 2014.

2. 米国の動向

- 2012年消費者プライバシー権利章典
  - FTC法執行権限，国際的相互運用性
- FTCレポート
  - Privacy by Design (PbD), 透明性，教育目的収集の徹底，差別阻止，
- データブローカー
  - Acxiom, busted in など，個人データ収集やプロファイリングが合法的なビジネス．



3. 中国：進む監視カメラ

- 1億7千万台の監視カメラ
- 2018年5月20日
  - 張学友（ジャッキーチュン）さんのコンサート会場で逃亡犯が逮捕されている
  - 入場ゲートに設置された監視カメラの顔認証システムが3年前から詐欺容疑で逃亡を続けていた男性を識別

http://j.people.com.cn/n3/2018/0816/c95952-9491225.html

演習）次の情報を分類せよ

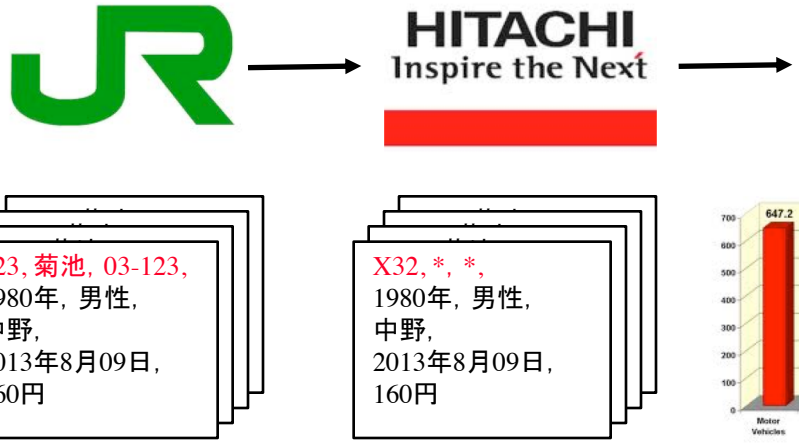
- A. 個人情報（1号），B. 個人識別符号（2号），C. 要配慮情報，D. プライバシー情報
  - ❑基本4情報（氏名，性別，生年月日，住所） A
  - ❑個人番号(マイナンバー） B
  - ❑YouTubeの閲覧経験 _
  - ❑Skypeの通話データ（音声） _
  - ❑クレジットカード情報 _
  - ❑ブラウザのBookmark _
  - ❑肌の色 _

匿名加工情報

同意なき第三者提供の課題

第三者提供例)「Suica案件」

- 2013年7月26日
  - JR東日本は4,300万枚のSuicaの乗降履歴を，7月から販売していた.
  - 名前，連絡先は除外，性別と年月日は含む.
  - 日立製作所が購入．市場調査用統計リポート，10駅分500万円.
  - 7月31日までに8,823件の除外申請.



http://digital.asahi.com/articles/TKY201307260002.html

第三者提供の方法

方法	備考	法	個人情報	要配慮個人情報
1. 同意	オプトアウト（通知，又は本人が容易に知り得る状態に置くとき）	23条2項	○	×
	例外規定（生命，身体，財産の保護，児童など同意が困難な場合）	23条1項	○	○
2. 委託・共同利用	第三者に該当しない.	23条5項	○	○
3. 匿名加工	公表義務	37条		
	加工基準，安全管理措置	36条，保護委員会規則		

匿名加工情報の定義

- 匿名加工情報（法第二条9項）
  - 個人情報の区分に応じて当該各号に定める措置を講じて特定の個人を識別することができないように個人情報を加工して得られる個人に関する情報であって、当該個人情報を復元することができないようにしたものをいう
  - 個人情報に含まれる記述等の一部を削除すること（当該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。）

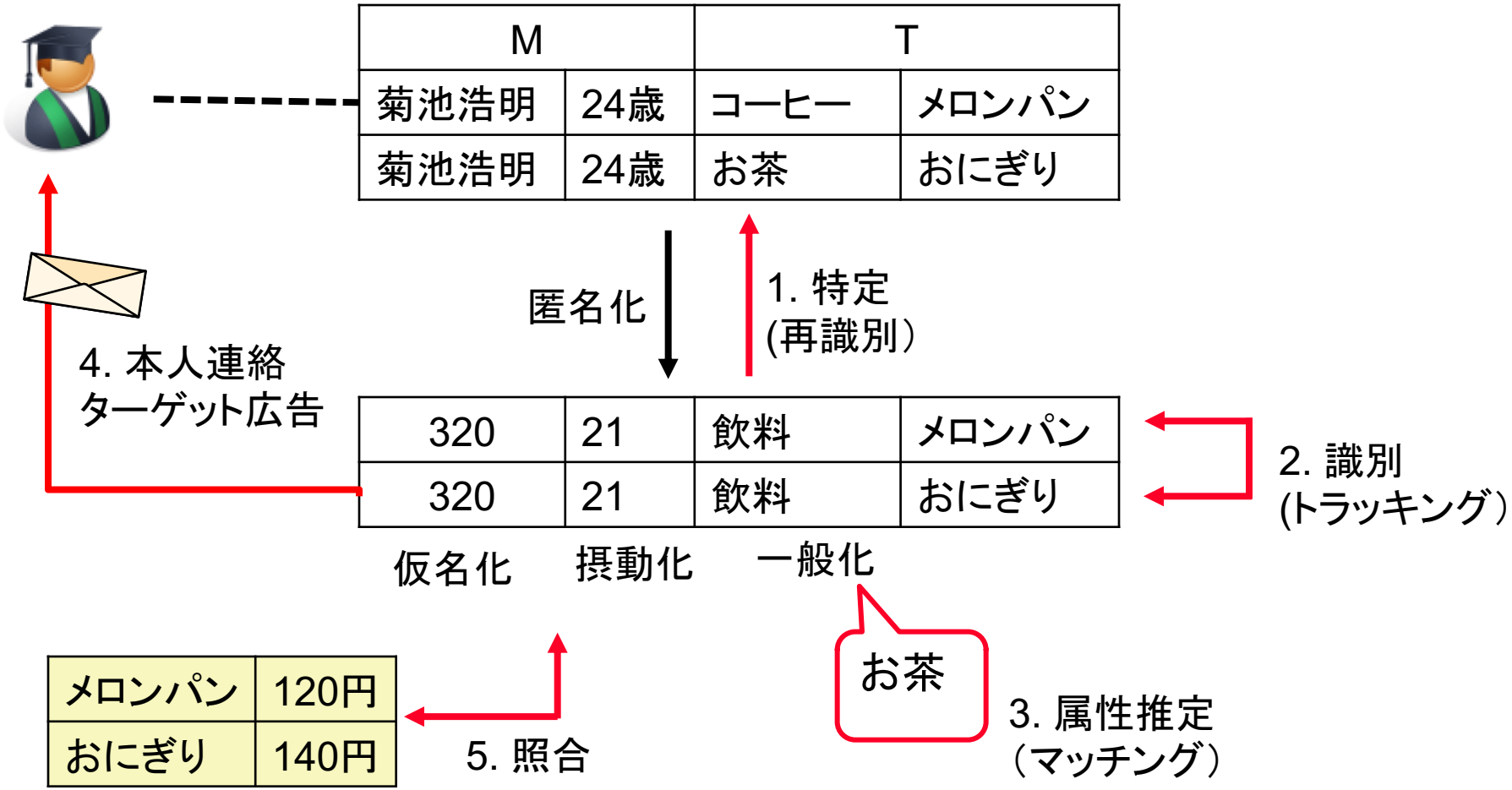
＝仮名化

JIPDEC匿名加工情報事例集

- 2017年7月7日．日本情報経済社会推進協会
- 認定個人情報保護団体として，会員の個々の事例について，整理したもの．

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

匿名加工情報の脅威


M
T
菊池浩明 24歳 コーヒー メロンパン
菊池浩明 24歳 お茶 おにぎり
匿名化
1. 特定
(再識別)
4. 本人連絡
ターゲット広告
320 21 飲料 メロンパン
320 21 飲料 おにぎり
2. 識別
(トラッキング)
仮名化
摂動化
一般化
お茶
3. 属性推定
(マッチング)
メロンパン 120円
おにぎり 140円
5. 照合

加工基準

「個人情報保護ガイドライン（匿名加工情報編）」より抜粋

規則	対象	例
(1)	個人情報の全部又は一部を削除すること	氏名を削除
(2)	個人識別符号の全部を削除すること	マイナンバーを削除
(3)	個人情報と当該個人情報に措置を講じて得られる情報とを連結する符号を削除する	管理用 ID を 削除
(4)	特異な記述等を削除	116歳を90歳以上に置換
(5)	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	・自宅住所が推定できる位置情報削除 ・購入者が限定される特殊商品を一般カテゴリーに置換

匿名加工手法例

手法名	解説
項目削除／レコード削除／セル削除	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記述等を削除するもの。 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（項目削除）、特定の個人の情報を全て削除すること（レコード削除）、又は特定の個人の年齢のデータを削除すること（セル削除）。
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又は数値を四捨五入などして丸めることとするもの。 例えば、購買履歴のデータで「きゅうり」を「野菜」に置き換えること。
トップ（ボトム）コーディング	加工対象となる個人情報データベース等に含まれる数値に対して、特に大きい又は小さい数値をまとめることとするもの。 例えば、年齢に関するデータで、80 歳以上の数値データを「80 歳以上」というデータにまとめること。
ミクロアグリゲーション	加工対象となる個人情報データベース等を構成する個人情報をグループ化した後、グループの代表的な記述等に置き換えることとするもの。
データ交換（スワップ）	加工対象となる個人情報データベース等を構成する個人情報相互に含まれる記述等を（確率的に）入れ替えることとするもの。
ノイズ（誤差）の付加	一定の分布に従った乱数的な数値を付加することにより、他の任意の数値へと置き換えることとするもの。
疑似データ生成	人工的な合成データを作成し、これを加工対象となる個人情報データベース等に含ませることとするもの。

万能な匿名加工はない

- 多用なビッグデータに対して，共通して適用可能な加工技術が未確立

加工＼リスク	特定	識別	属性推定
削除	部分的	部分的	部分的
仮名化	No	部分的	No
一般化	No	部分的	部分的
k-匿名性	Yes	部分的	No
ℓ-多様性	No	No	Yes
差分プライバシー	Yes	Yes	部分的

ISO/IEC DIS 20889:2018，Table A.1から抜粋

k-匿名化

1. オリジナル

M		T
A大学	24歳	コーヒー
A大学	24歳	お茶
B大学	28歳	お茶
C総研	35歳	コーヒー
C総研	35歳	コーヒー

2. 一般化（所属）

M		k	T
大学	24歳	2	コーヒー
大学	24歳		お茶
大学	28歳	1	お茶
企業	35歳	2	コーヒー
企業	35歳		コーヒー

k=1

3. 一般化（年齢）

M		k	T
大学	20代	3	コーヒー
大学	20代		お茶
大学	20代		お茶
企業	30代	2	コーヒー
企業	30代		コーヒー

k=2, ℓ =1

ℓ-多様性

3. 一般化（年齢）

M		k	T
大学	20代	3	コーヒー
大学	20代		お茶
大学	20代		お茶
企業	30代	2	コーヒー
企業	30代		コーヒー

k=2, ℓ =1

4. 摂動化（履歴）

M		k	T	ℓ
大学	20代	3	コーヒー	2
大学	20代		お茶	
大学	20代		お茶	
企業	30代	2	コーヒー	2
企業	30代		お茶	

k=2, ℓ =2

改正法律案

- 第五十三条
  - 認定個人情報保護団体は，
    対象事業者の個人情報等の適正な取り扱いの確保のために，
    (1) 個人情報に関わる利用目的の特定，
    (2) 安全管理のための措置，
    (3) 開示等の請求等に応じる手続き…又は匿名加工情報に係る作成の方法..

    に関し，消費者の意見を代表する者その他の関係者の意見を聴いて，

    この法律の規定の趣旨に沿った指針を作成するように努めなくてはならない

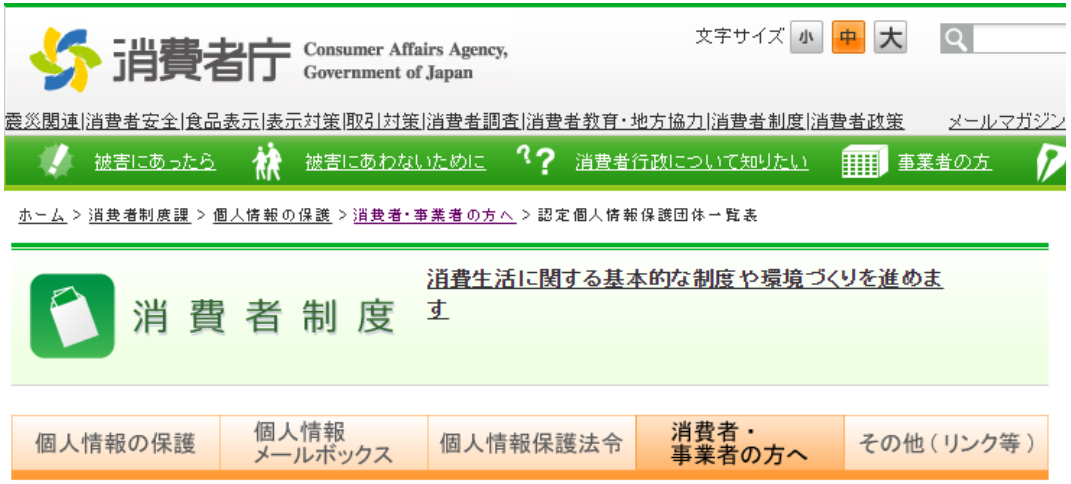
個人情報保護委員会の認定を受けた法人（第47条）

マルチステークホルダ

個人情報保護指針

認定個人情報保護団体

- 業種ごとに認定
  41団体
  - 電気通信事業
    総務省・経産省
    日本データ通信協会
  - クレジット事業
    経産省
    日本クレジット協会



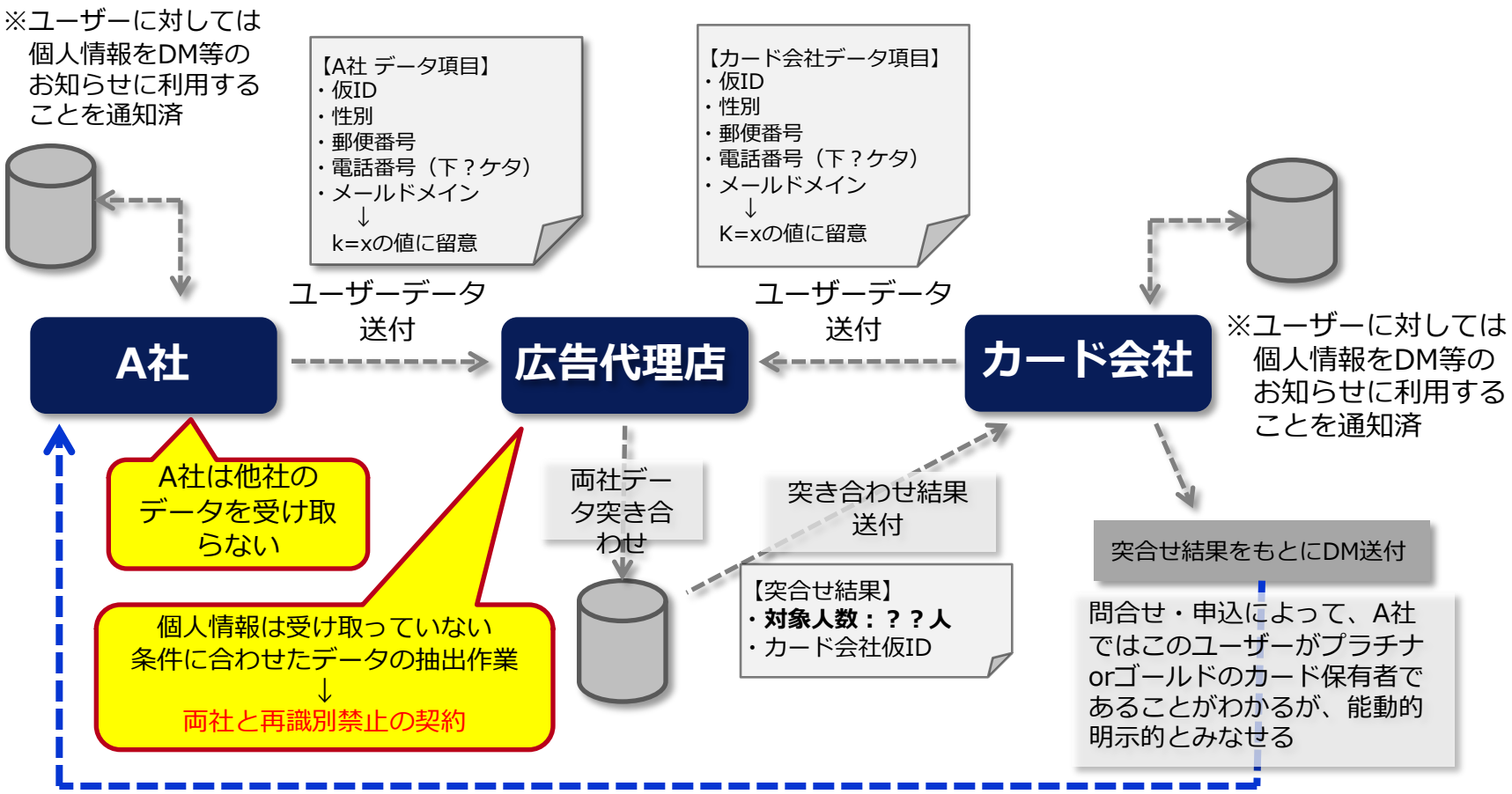
認定個人情報保護団体一覧表

平成26年11月26日現在

対象事業等分野	所管府省	名称	苦情処理相談窓口の電話番号	所在地	認定年月日	ガイドラインの名称
警備業	国家公安委員会	一般社団法人　全国警備業協会	03-3342-5821	東京都新宿区西新宿1-9-18　永和ビル7階	平成20年11月21日	警備業における個人情報の保護に関するガイドライン
指定自動車教習所業	国家公安委員会	一般社団法人　全日本指定自動車教習所協会連合会	03-3556-0070	東京都千代田区九段南2-3-9　サンク九段ビル4階	平成26年10月9日	指定自動車教習所業における個人情報保護に関する指針
証券業	金融庁	日本証券業協会	03-3667-8451	東京都中央区日本橋茅場町1-5-8	平成17年4月1日	個人情報の保護に関する指針

ケース7

異業種2社の会員データを活用したプライベートデータエクスチェンジ



（パーソナルデータ利活用に関するマルチステークホルダープロセスの実施方法等の調査事業）報告書，図表41　イメージ図（抜粋）

識別行為の禁止

- 第三十八条

  匿名加工情報取扱事業者は，匿名加工情報を取り扱うにあたっては，当該匿名加工情報の作成に用いられた個人情報に係る本人を識別するために，当該個人情報から削除された記述等若しくは個人識別符号若しくは第三十六条の規定により行われた加工の方法に関する情報を取得し，又は当該匿名加工情報を他の情報と照合してはならない

ケース7の問題点

A社 =個人情報取扱事業者

菊池浩明	77	24歳	お茶
松本 慶浩	85	42歳	お茶
佐藤 泰	90	32歳	お茶

匿名加工
第三者提供

77
85
90

広告代理店 =匿名加工情報
取扱事業者

識別の為
の照合

38条禁止
行為

カード会社B =個人情報取扱事業者
=匿名加工情報取扱事業者

菊池浩明	77	24歳	おにぎり
木村 愛	40	18歳	おにぎり
佐藤 泰	90	32歳	おにぎり

匿名加工
第三者提供

77
40
90

識別の為の照合

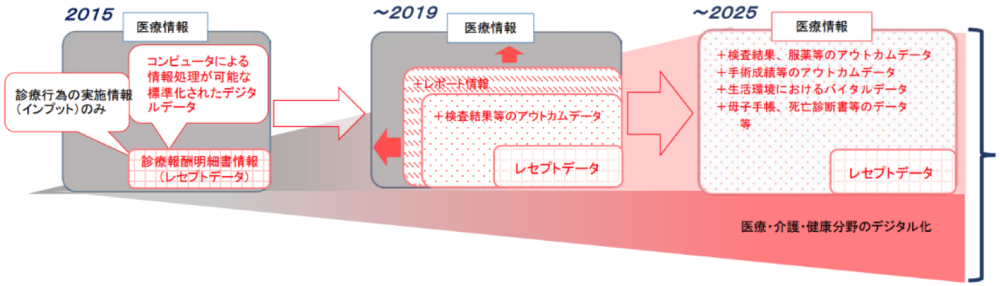
77
90

38条禁止行為

2018年5月11日

- 次世代医療基盤法施行

次世代医療基盤法

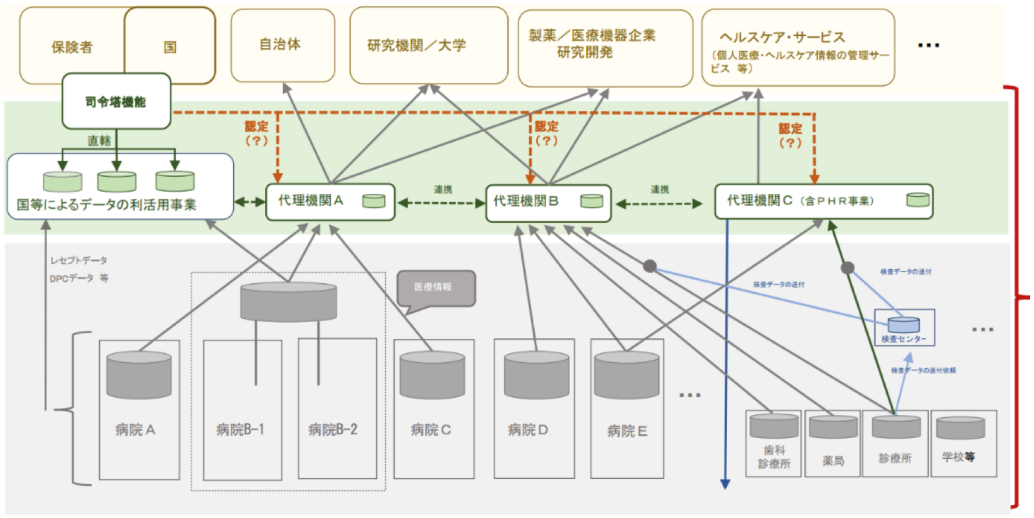


ポイント

1. 医療現場のICTによる高度化、及び医療データの収集・利活用の仕組みを車の両輪として確立。

2. 国民・患者への新しい付加価値の創出（医療行政、医療／医療サービス、研究開発　等）

代理機関（仮）制度を利用したオールジャパンの医療情報利活用イメージ】



医療データの収集・利活用の仕組づくりの留意点

代理機関（仮）制度の適切な実施

①事業者のイメージ
②事業内容、利活用イメージ
③医療における役割イメージ
・・等

第2回　次世代医療ICT基盤協議会，H27.12.25, 事務局資料

二つの匿名加工情報

	匿名加工医療情報	匿名加工情報	備考
法律	「次世代医療基盤法」 2018年5月11日施行	「個人情報保護法」 2017年5月30日施行	
対象	医療情報 医療情報データベース	個人情報 個人情報データベース	
病歴等の第三者提供	○ オプトアウト	× 要配慮個人情報はオプトイン	一般は両方とも○
転々流通	× 契約がGLで推奨	○ 公表義務	
加工事業者	認定 一般医局が行わない	任意 委託事業者等	200社以上実施
識別のための照合	○ 規定なし	× 法38条禁止行為	仮IDの履歴長
加工方法作成基準	施行規則18条 1－5項	施行規則19条 1－5項	等価

プライバシー保護データマイニング

足し算プロトコル Step 1

A = 1 = 2 + 3 – 4 mod p

1 = good

0 = bad

足し算プロトコル Step 2

2		3		-4
+4		-4		+1
+2		+2		-4

8 + 1 + -7 = 2

2

A: 1 = 2 + 3 – 4

B: 1 = 4 – 4 + 1

C: 0 = 2 + 2 – 4

研究目的: 安全な疫学調査

- 研究目的: 組織A, Bが互いのデータセットを秘匿して，相対危険度RRを求める

Aにも誰か不明

氏名	年齢	部位
菊池浩明	25	胃境界
佐久間淳	28	胃底部
三上春雄	35	幽門

Bには秘密

氏名	年齢	検診日
吉田哲郎	45	2001
菊池浩明	25	2001
古川和彦	35	2002

Aには秘密

組織A (千葉がんセンター)　組織B (厚生省・保健所)

だれか1名共通な人がいる.

疫学調査

- 患者-対象調査

要因	がん罹患	対象(無)	罹患率
ピロリ菌	a	b	a/(a+b)
未感染	c	d	c/(c+d)

- 相対危険度 (Relative Risk)

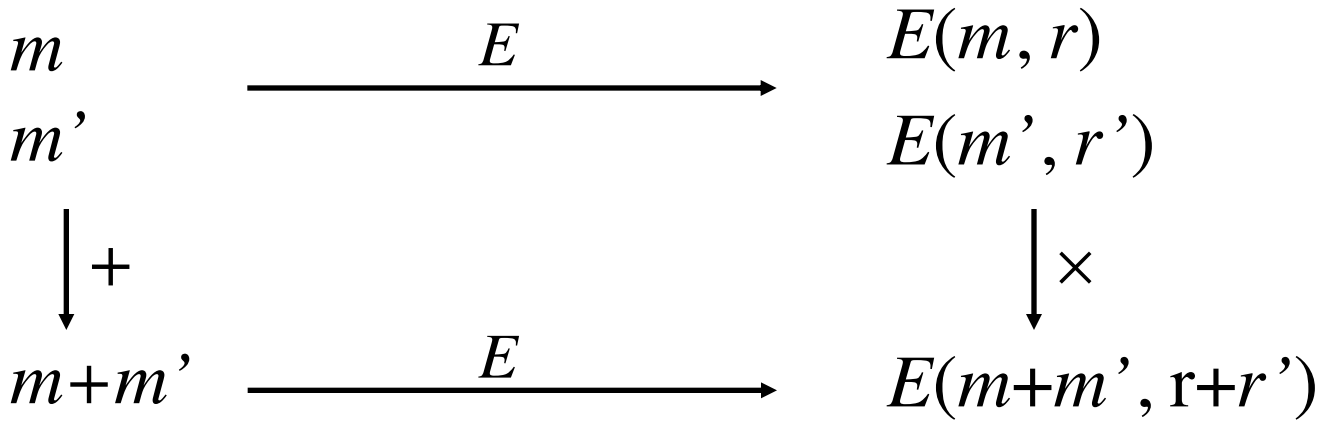
$$RR = \frac{a}{a+b} / \frac{c}{c+d} \approx \frac{ad}{bc}$$

- 統計量

$$\chi = \frac{\sqrt{N-1}((ad-bc) \pm N/2)}{\sqrt{(a+c)(b+d)(a+b)(c+d)}}$$

加法準同型性

- 可換な図



- Privacy-preserving computations

$$E(m,r) \times E(m', r') = E(m+m', r+r')$$
$$E(m,r)^x = E(mx,rx)$$

準同型性暗号(セキュア内積プロトコル)

A	B
x = (上原，菊池，吉田，若林)	y = (上原，菊池，吉田，若林)
= (1, 0, 1, 1)	= (1, 1, 1, 0)

Xの暗号化

$E(1)$

$E(1), E(0), E(1), E(1)$ →

$c = E(1)^1 E(0)^1 E(1)^1 E(1)^0$

$= E(1*1)\ E(0*1)\ E(1*1)\ E(1$

$= E(1 + 0 + 1 + 0)$

$= E(2)$

← c

復号

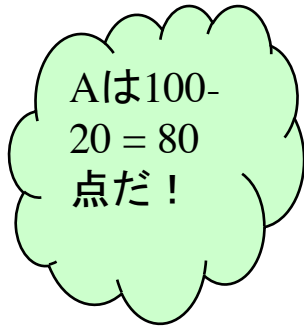
$D(c) = D(E(2)) = 2$

$= |x \cap y|$ 合計のみ分かる

出力プライバシー

- 例)「二人で平均点を秘密計算」

A	B	平均
80点	20点	50点



- 防止策=差分プライバシー
  - R = Δ_f/εとする $f(X)$ + rは差分プライバシーを満足する. ここで, r はラプラス分布Lap(|x|/R)に従う乱数

A	B	平均+Lap(x/R)
80点	20点	75点

まとめ

- 氏名などの属性およびその他の情報を照合することで特定の個人を（　　）できる情報を個人情報という．個人番号や生体情報などを（　　）という．
- 匿名加工には，氏名などをランダムな番号に置き換える（　　），データにノイズを加える（　　），より上位の概念に置き換える一般化などがある．匿名加工することで，本人の（　　）で第三者提供が認められる．
- 準同型性暗号などを用いることで，情報を暗号化したままデータマイニングすることを（　　）という．

演習 問2

- (1) 疑似識別子となる属性の組み合わせはいくつあるか
- (2) {性別，通勤}, {性別，メガネ}, {通勤，メガネ}の疑似識別子について，kを求めよ.

ID	性別	通勤	メガネ	年収
1	M	電車	あり	1000万
2	F	電車	なし	800万
3	F	電車	なし	500万
4	F	自転車	あり	60万
5	M	自転車	あり	500万
6	M	電車	あり	600万