

ビッグデータとパーソナルデータ

国立情報学研究所

佐藤一郎

E-mail: ichiro@nii.ac.jp

Ichiro Satoh

自己紹介: 佐藤一郎

- 国立情報学研究所・アーキテクチャ科学研究系・教授
- 国立大学法人総合研究大学院大学・複合科学研究科・情報学専攻・教授
- 政府IT総合戦略本部「パーソナルデータに関する検討会」委員／
同検討会・技術検討ワーキング主査
- 専門: OSやミドルウェアなどのシステムソフトウェア ニュースゼロ(日テレ)



NHK週間ニュース深読み



国立科学博物館(上野)

やじうまテレビ
(テレ朝)



講演概要

- ビッグデータ
- 個人情報保護法改正に向けた動き
- 法改正の3つのポイント
 - 個人情報の範囲
 - 同意なし新しい第三者提供
 - 第三者機関(プライバシーコミッショナー制度)
- まとめ

注意:

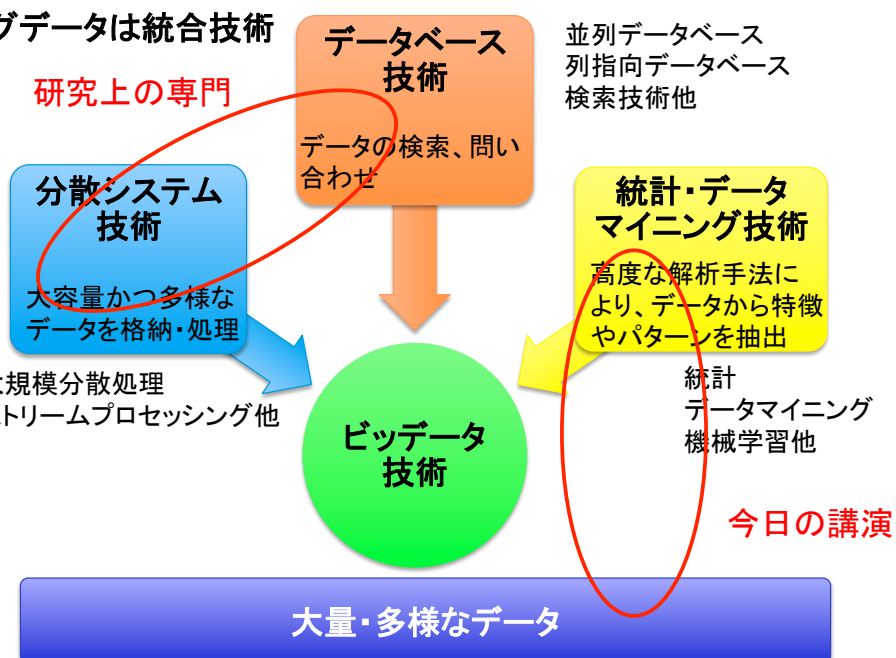
- 政府方針を話しますが、当方個人の意見とってください
(あたりまえですが、政府を代表するわけではないから)

Ichiro Satoh

技術としてのビッグデータ

- ビッグデータは統合技術

研究上の専門

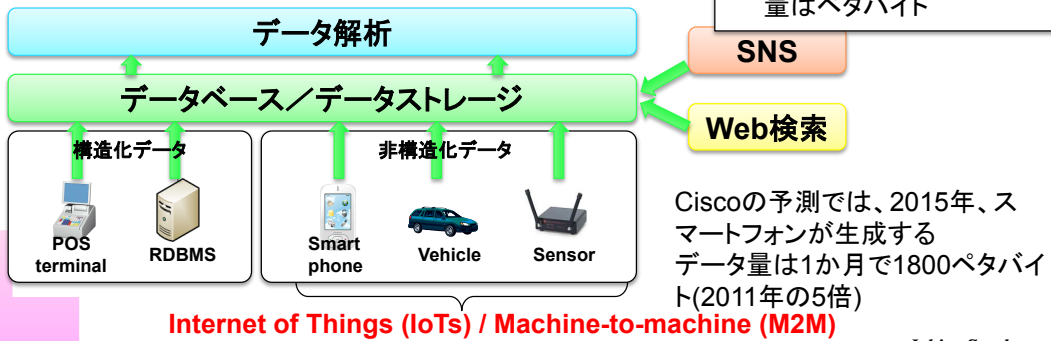


Ichiro Satoh

なぜいまビッグデータなのか

- **大量データ** 既存データの90%は2年以内に生成されている
 - センサー等から大量データが生まれている
- **多様性**
 - 非構造化データへの対応、相違なデータの利用
- **複雑性**
 - 相違なデータの関連性抽出
- **迅速性**
 - 発生頻度が高いデータを扱う

- 旅客機の**ジェットエンジン**に搭載された各種センサーや制御システムは、**1時間に10TBのデータを生成できる**(異常検出目的なので、正常値(大部分)は捨てている)
- 1日に1万機を超える航空機が運航しており、データ量はペタバイト



Ichiro Satoh

マーケティングの変化

- **ビジネス環境の変化**
 - 消費者の影響力増大、グローバル化、ネット化
 - 消費者は、企業よりも他の消費者を信頼
- **マス市場を前提としたマーケティングは通用しない**
 - **個々の消費者行動の分析が必須**

	これまでのマーケティング	いまのマーケティング
製品管理	4P (製品、価格、流通、プロモーション)	協創
顧客管理	STP(セグメンテーション、ターゲティング、ポジショニング)	コミュニティ化
ブランド管理	ブランド構築	キャラクターの構築

一部のネット通販サイトはマウスの動き(≒視線の動き)やリンクを開く時間も収集

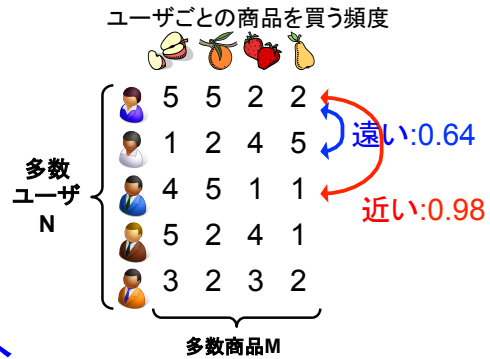
Ichiro Satoh

いまのマーケティング

■ マーケティングの変化

- ネット情報、SNS、消費者の影響力増大
- 消費者は、企業のマス広告よりも他の消費者を信頼

個々の消費者行動の分析が必須



■ ディメンションデータからファクトデータへ

- 従来: コンビニ1店舗のアイテム数: 3,000個 → 3000個の日販数データ
- 現在: 各コンビニ店の一日来店数: 1000人で、平均購入数: 3個
→ 1000 × (アイテム1, アイテム2, アイテム3)

Ichiro Satoh

なぜ個人情報保護法の改正なのか

- 個人情報保護法は技術進歩に密接に関わる分野を扱う
- 現行法の制定は2003年で、その後の技術の進歩に追いついていない
(付帯決議で2008年には見直すはずだった)
- ビッグデータ技術により、大量・多様なデータの処理が可能に
 - ビッグデータの対象はパーソナルデータだけではないけど
 - とはいってもビジネス的な見地からパーソナルデータに関心が集中
 - 顧客データを売りたいとか...

■ 個人情報に関わる問題が顕在化

- 個人に関わる情報を利活用したい
- 個人の情報に関わる不安
- 海外の個人情報保護との整合性

法改正へ

Ichiro Satoh

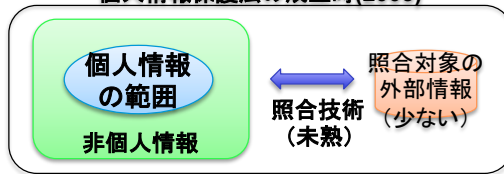
技術進歩と個人情報保護法

「個人情報」とは
(個人情報保護法第2条第1項より)

- ① 生存する個人に関する情報であり、
- ② 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をいう(「識別性」)
- ③ 他の情報と容易に照合することができ、それにより特定個人を識別することができるものも含まれる(「容易照合性」)

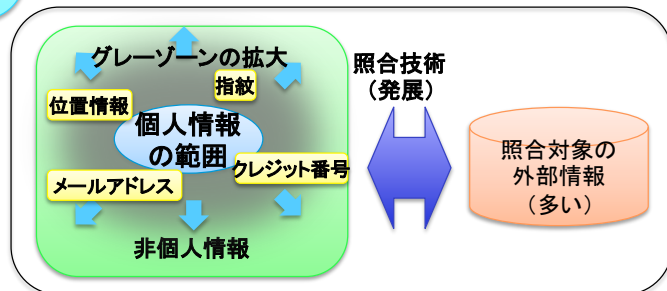
- 容易照合性の問題
 - 法解釈が技術進歩により拡大していいのか(識別範囲の拡大)
 - 技術進歩が利活用に逆インセンティブになっていいのか
- しかし、変更も難しい

個人情報保護法の成立時(2003)



■ グレーゾーンが増大

- 個人に関わる情報が増加
- ビッグデータにより照合技術が向上
- インターネットをはじめとして照合しうる外部情報が増大



ビッグデータ時代(現在)

Ichiro Satoh

個人情報保護法の改正の背景及び趣旨 政府(IT総合戦略本部)方針

■ 個人情報保護法の問題:

- 曖昧さ、技術進歩と合致しない、過度の萎縮、個人の不安……

■ ビッグデータ時代におけるデータ利活用の推進

- データ活用による新しいビジネスの創出
- プライバシー問題へ不安

■ データの国間流通の実現

- グローバルビジネスの推進
- 世界の標準にあったパーソナルデータ保護
- 安全管理措置

個人情報保護法の抜本改正

目標: 個人情報及びプライバシーの保護に配慮した
パーソナルデータの利用・流通を促進

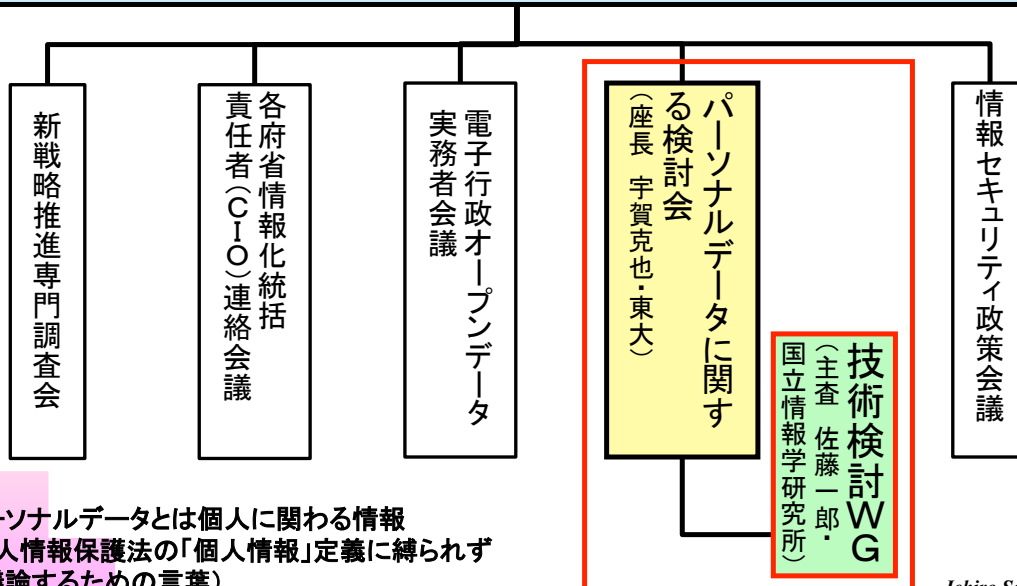
- 透明性の確保と曖昧性の最小化
- (同意不要な) 第三者提供
- 第三者機関の導入
- 最新技術を考慮
- 企業活動の円滑化
- 世界水準の保護

パーソナルデータに関する検討会
と技術検討WGを設置

Ichiro Satoh

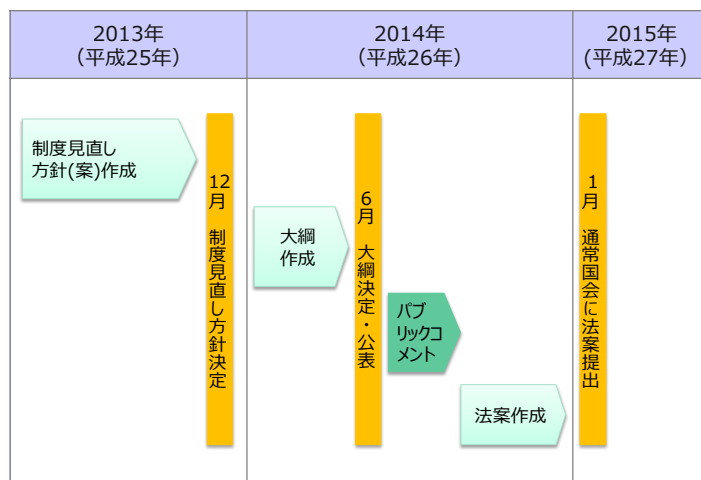
▶ パーソナルデータに関する検討会の体制

高度情報通信ネットワーク社会推進戦略本部
(IT総合戦略本部) (本部長: 内閣総理大臣)



▶ スケジュール

- 検討会を12回開催して、6月中旬に大綱案(検討会案)とりまとめ
- 6月下旬 IT総合戦略本部における大綱決定
- 7月パブリックコメント
- 来年の通常国会提出となると、大綱は限りなく法案に近いレベルにする必要がある



※ 欧米を含めた諸外国の制度についても現在変更に向けた作業が行われているため、これらとの整合性を取るためにある程度の時間が必要となる。

(例: EUデータ保護規則案 2014年4月に欧州議会本会議で採択の見込み)

▶ 法改正の3つのキーポイント

■ グレーゾーンを減らす

- 従来の個人情報に加えて、個人の特定につながるパーソナルデータを整理

技術WGの識別
非特定情報が
ベース

■ 同意なし第三者提供を前提にした新たな種類の導入

- パーソナルデータの保護と利活用を両立

技術WGの(仮称)
法第23条1項適用
除外情報がベース

■ 第三者機関(プライバシーコミッショナー)の設置

- 主務大臣が有している機能・権限に加え、
立入検査等の機能・権限

- この他、パーソナルデータの海外移転、請求権、自主規制他

Ichiro Satoh

▶ グレーゾーンを減らす

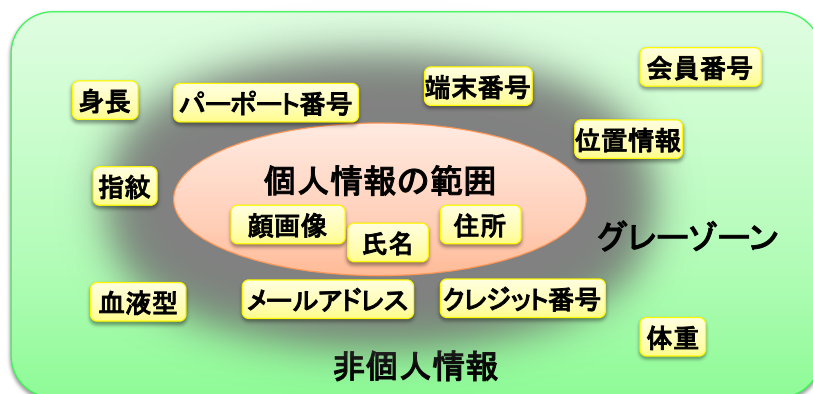
- ただちに特定の個人を識別しないけれど

- 外部データとの突き合わせなどで、特定の個人を識別する情報で
- 取り扱いによっては個人の権利利益侵害の可能性がある情報

保護
対象

(仮称)
準個人
情報

ITが新しい
パーソナル
データ
を作り出した



技術WG: 保護すべき情報を整理・明確化

Ichiro Satoh

▶ (仮称)準個人情報とは何だったのか

日本経済新聞

平成26年5月26日(月)朝刊15面

■ (仮称)準個人情報

「準個人情報」新区分が浮上
メールアドレスや駅の乗降履歴など

個人データの新たな分類イメージ

個人情報	準個人情報(新設)
氏名、住所、顔写真など	端末ID、認証データ、位置データ、メールアドレスなど
目的外利用は本人の同意が必要	分類を新たに設け、一定条件を満たせば利用できるように制度改正
第三者への提供は本人	

政府検討会「データ」

- グレーゾーン上の情報を整理するために、ひとまず範囲を設定
- 範囲設定の3例 (条文に落とし込むときに変わリエる)



▶ 技術検討WGの立場

- パーソナルデータによる本人の権利利益侵害ケース
 - ① 情報から何らかの状況で個人が特定されてしまうことで、権利利益侵害が生じるケース
 - ② 情報から個人が特定されないままで、権利利益侵害が生じるケース
- 技術WGでは、①を前提に検討、②はその情報を利用して、個人に関する情報をマッチングしたケースであり、個人の権利利益に侵害が及ぶケースと及ばないケースがあるが、技術だけでは線が引けない。
- 事前規制でパーソナルデータを守るという立場。海外では事後救済を前提にしていることが多いが、個人情報保護法は事前規制が前提。

▶ 技術WGの項目選定基準

- 個人情報ではないが、取り扱いによっては特定の個人を識別しうる情報項目で、個人の権利・利益侵害につながる情報は保護すべき
- 技術WGが提案した保護されるべき情報の要件(すべてを満足する場合)
 - 本人に付与されるものか、所有物に付与されるものか
 - 本人または本人に密接に関わるモノに割り当てられていること
 - 一意性(重複性)／単射性(一対一)
 - 識別子に重複がない／一つの識別子は一つの対象を参照
 - 共用性
 - 複数事業者間、または(一事業者であっても)相違サービス間で利用される識別子
 - 変更可能性／利用停止可能性／不変性
 - 利用者により、識別子の変更、停止ができる機能がないとき

Ichiro Satoh

▶ (準個人情報として)新たに保護されるべき情報項目(技術WG案)

- 個人情報ではないが、取り扱いによっては特定の個人を識別しうる情報項目で、個人の権利・利益侵害につながる情報は保護すべき

個人又は個人が使用する通信端末機器等にかかるもの	運転免許証番号、パスポート(旅券)番号、健康保険証の記号・番号(健康保険被保険者証記号番号等)、雇用保険被保険者番号、外国人の在留に関する番号(在留カード番号、特別永住者証明書番号、外国人登録証明書番号)、金融機関の口座に関する番号、クレジットカード番号、メールアドレス、ナンバープレート(自動車登録番号標等)番号、固定電話番号、携帯電話番号、情報通信端末シリアルナンバー(携帯電話端末シリアルナンバー等)、MACアドレス、情報通信端末ID、ICカードの固有ID、ソフトウェアシリアル番号、不動産番号、IPアドレス(V6)
個人の身体的特性にかかるもの	声紋、指紋、静脈パターン、虹彩、DNA、顔認識データ、掌形、生体認証で使用されるデータ、歩行パターン、筆跡

- 網羅的に挙げたのではなく例示(技術動向などでかわる)
- 政令などで追加可能にすべき
- 上記に含まれても個人情報になる場合がある(例:メールアドレス)
- **上記は個人情報と組み合わせられていたら、個人情報**

Ichiro Satoh

▶ 準個人情報として保護されるべき 情報項目(技術WG案)の備考

- 免許書番号やパスポート番号等は、番号自体では個人の特定はできなくても、当該データベース等から個人を特定可能
- メールアドレスはユーザ名やドメイン名から個人を特定できる場合は個人情報
- 付番された本来の用途まで制約しない(通信目的の識別子など)
- 変更が困難な端末識別子は保護対象
(ただし、変更・停止可能性により個別判断)
- IPv4は使い回しされるケースが多いので、IPv6を対象にしてある
- 顔識別データは個人を特定できる顔画像になるならば個人情報

Ichiro Satoh

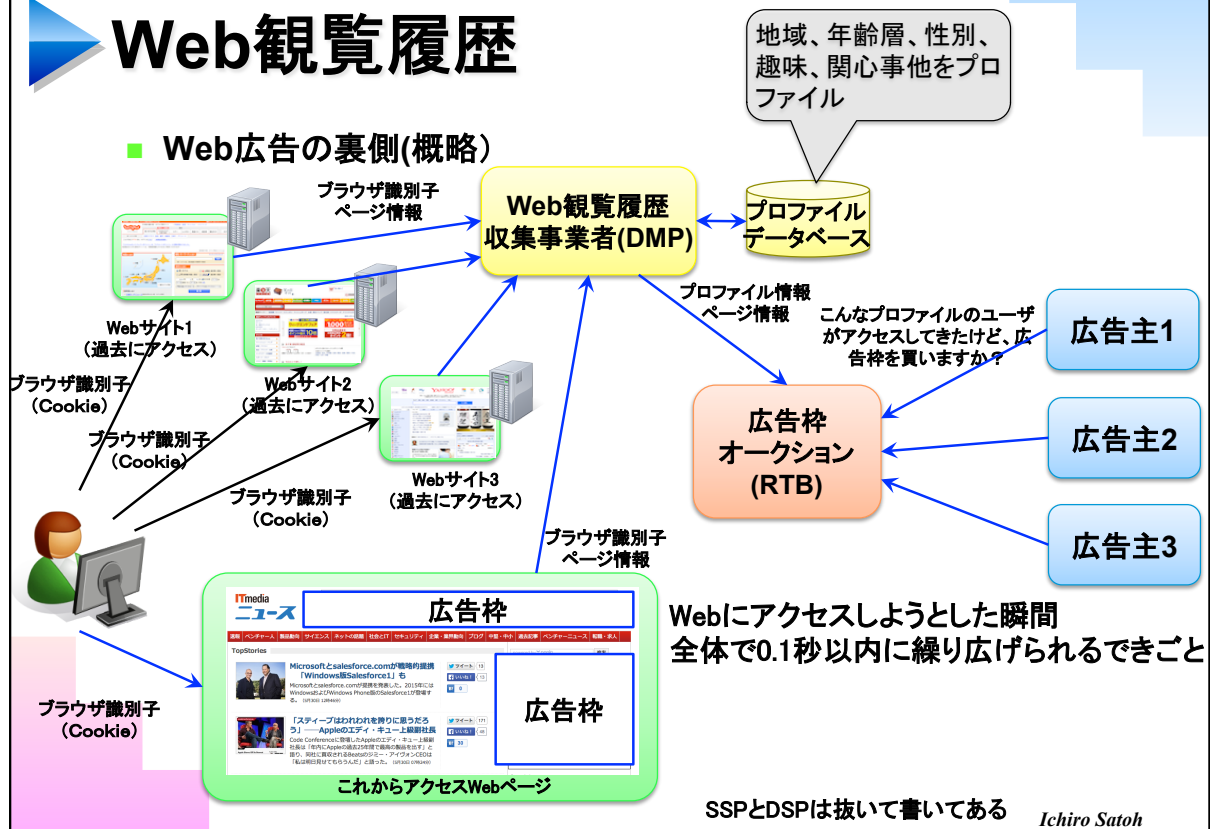
▶ 移動／購買履歴の取り扱い (技術WG報告書)

- **移動履歴を準個人情報(保護対象)にすべきかは個別事例で判断**
 - 情報量、取得期間・取得頻度、位置情報や時刻の精度等の個別の条件によって、特定の個人を識別できる場合、できない場合があるが、基準は技術WGでは引けない
 - 高精度の位置情報+時刻情報だけでも、個人の識別性をもつので、むしろ個人情報かも(例:深夜2時の位置情報≒自宅住所)
- **購買履歴は特徴がある履歴以外は準個人情報ではない**
 - 一品物や購入に特殊性がなければ個人特定性は低い
 - 逆に長期の履歴や、一品物や特異な購入は特徴のある履歴
 - 書籍や医薬品は別途考慮が必要(世論の判断)
 - 販売店は位置情報として扱う
 - 一意な番号(RFIDタグの番号)やシリアル番号が商品毎に付番され、外部からアクセスできる場合は「(仮称)準個人情報」

Ichiro Satoh

Web観覧履歴

■ Web広告の裏側(概略)

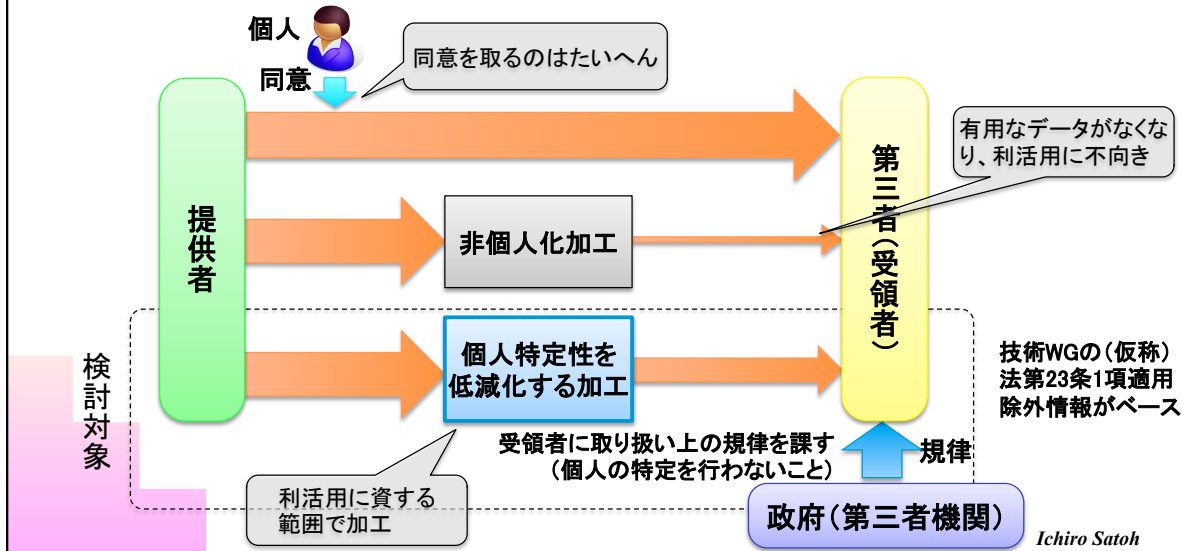


Web観覧履歴(技術WGの報告書から)

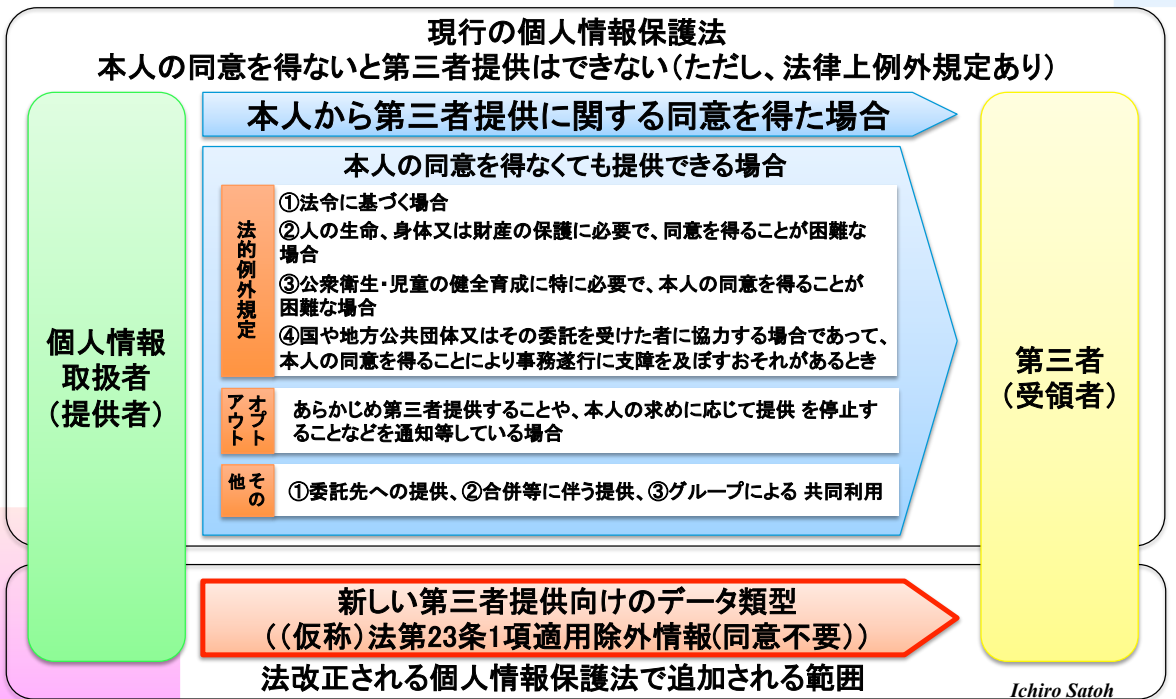
- 個々のケースで判断すべき(技術では整理できない)、ただし
 - **Cookieを含む Web に関わる端末やブラウザを識別する情報**
 - 利用者が識別子としての番号そのものをリセット又は取得・利用停止を選択できる機能(例えば、オプトアウト)を有している場合は準個人情報から外するのが妥当
 - 上記の機能に関わる制度とともに検討すべき
 - **Web観覧履歴**
 - 少なくとも観覧したWebを示すURLなどの情報に、個人情報や準個人情報が含まれる場合は、それぞれの情報とすべき(一般の場合は判断が付きず)

▶ パーソナルデータの第三者提供

- 産業界の要求
 - データをもっている事業者が、そのデータを使うとは限らない
 - 他社のもっているデータを利用したい
- しかし、第三者提供には同意が必須で手間かかる(他の手段を作れ！)

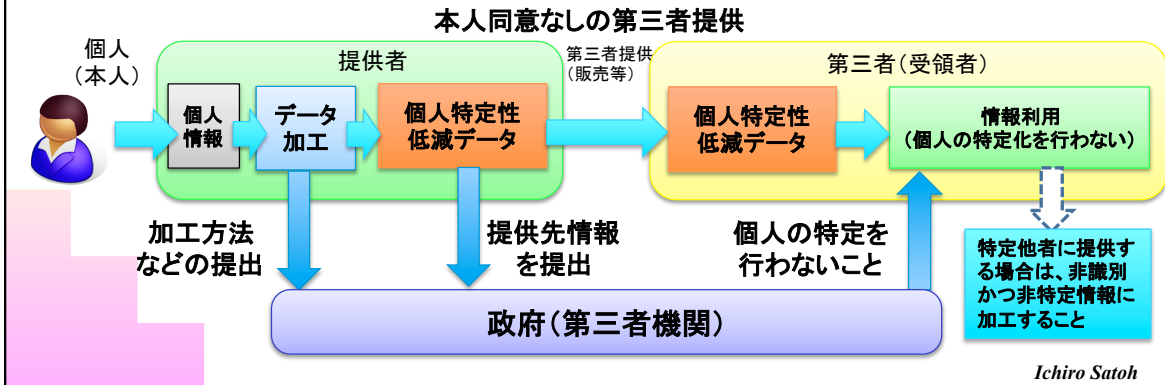


▶ 個人情報保護法における第三者提供



▶ (仮称)個人特定性低減データ

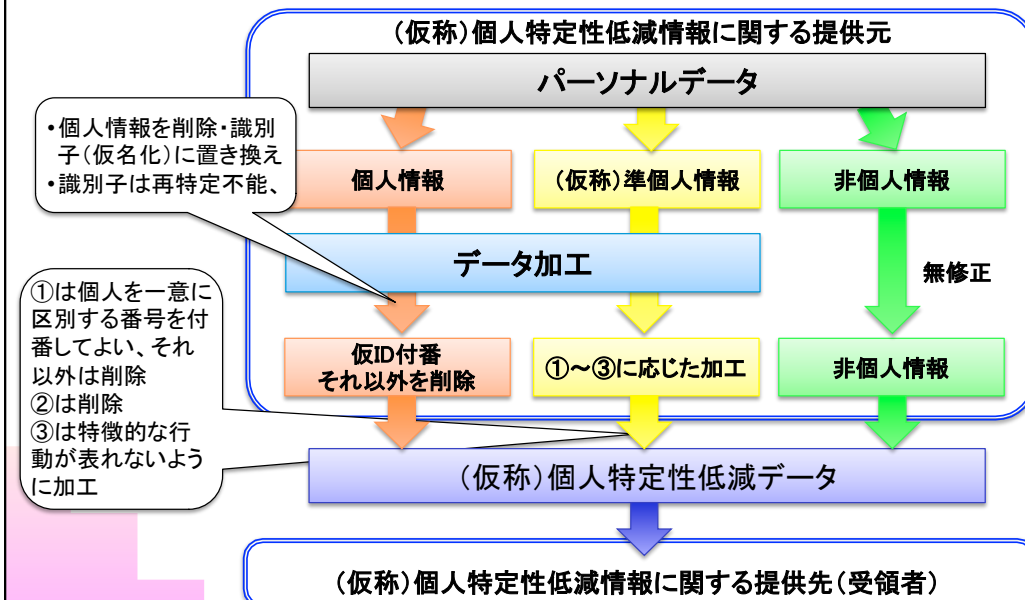
- 提供者はデータ提供の前に、所定の加工を規律で定める
 - 氏名や会員番号など直接個人が特定できる情報を削除
 - 個人が推定しうる情報(生年月日や住所)を曖昧化
 - 加工や提供先を第三者機関に提出
- 受領者は個人の特定を行わないという制約を課す
 - 2次提供はしてよい(不特定他者への提供の場合は、非個人化する)



Ichiro Satoh

▶ (仮称)個人特定性低減データ

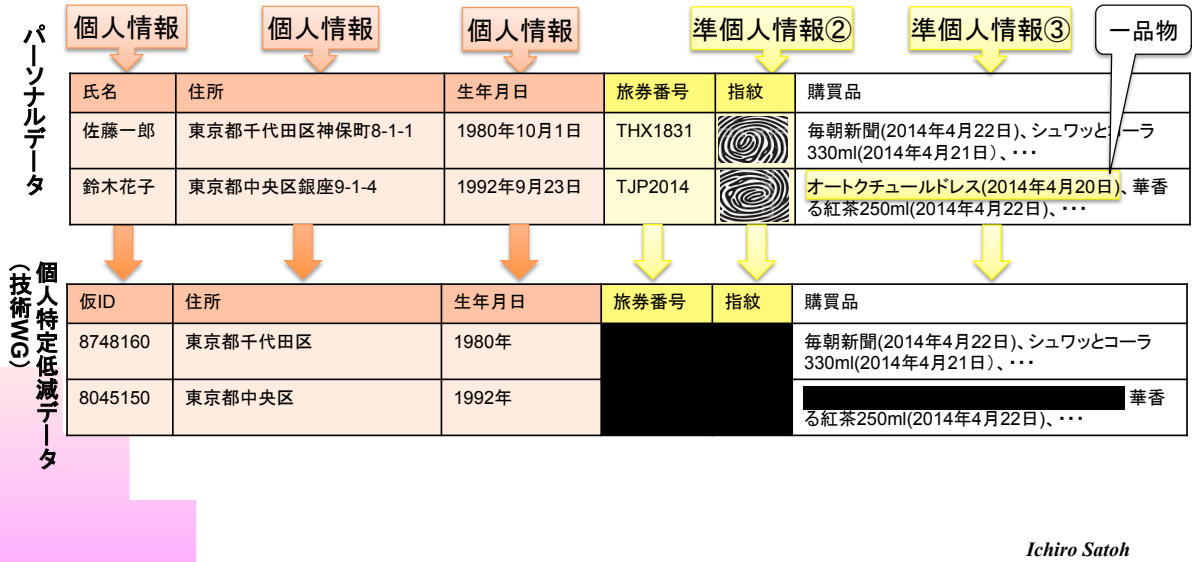
- 任意のデータに対する一律の加工方法は技術的に定義困難



Ichiro Satoh

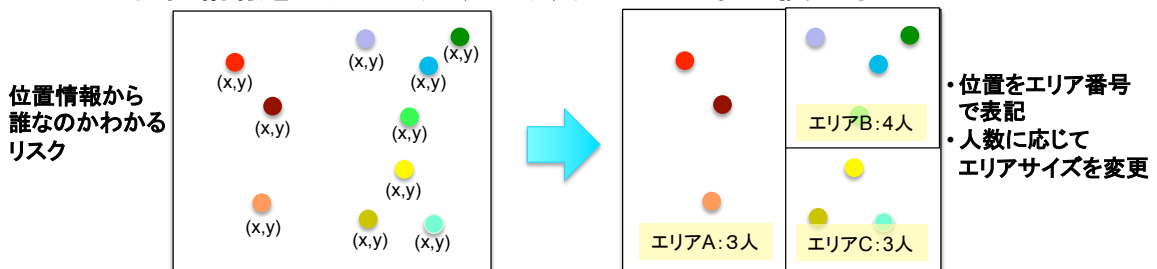
▶ 低減データ: 個人情報の加工

技術WGによる低減データ加工イメージ

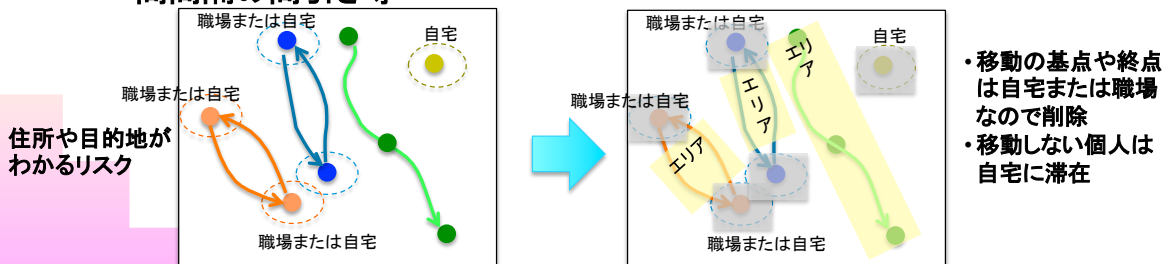


▶ 位置情報に関する加工

- 位置情報をエリア化(一般化)、ランダムな置き換え等

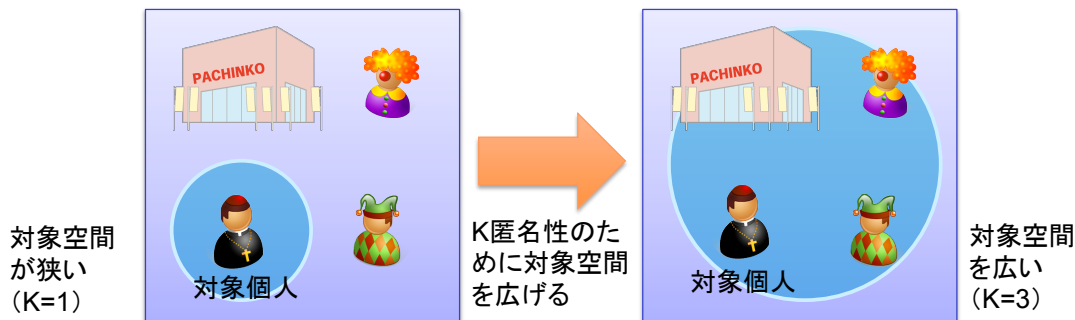


- 移動情報では仮IDの短期再割り当て(長い履歴の削除)、分割や時間間隔の間引き等



匿名化が招く問題

- k-匿名性やエリア化を実現すると、複数人が位置的には区別不能になるが
 - 範囲内の他の個人や施設にいと間違えられるケースも
 - 位置情報以外でも同様の問題がある



- 自衛策:個人は人違いされたくない人たちがいる地域や、いると思われたくない場所のある地域を避けること、ただし結果として
 - 住居や行動範囲に地域的分断が生じる → 深刻な社会問題を引き起こす

Ichiro Satoh

低減では劇薬と行ってもいいほどの強い薬

- 強い薬は効用もあるけど副作用も大きい
 - 効用:同意なしの第三者提供
 - 副作用:パーソナルデータの漏洩や不正利用
- 個人の特性可能性が残った情報を第三者提供する代わりに、規律で縛る
 - 提供者:個人を直接的に特定する情報を加工(削除または仮名化)
 - 受領者:データから個人の特定を禁止
- 副作用を抑えるには提供者と受領者に規律を守らせる仕組みが必要
 - それを緩めるのはたいへん危険(自主基準・規制は向かない)
 - 低減データに関わる脱法は監視・取り締まりが困難(低減データを十分加工しておくことでリスクを最小化するぐらいしか対処策がない)

Ichiro Satoh

▶ 低減データの加工や取り扱いは厳しいのか？

経済界等から、加工程度や取り扱いは緩めるべきとご意見を頂くわけですが

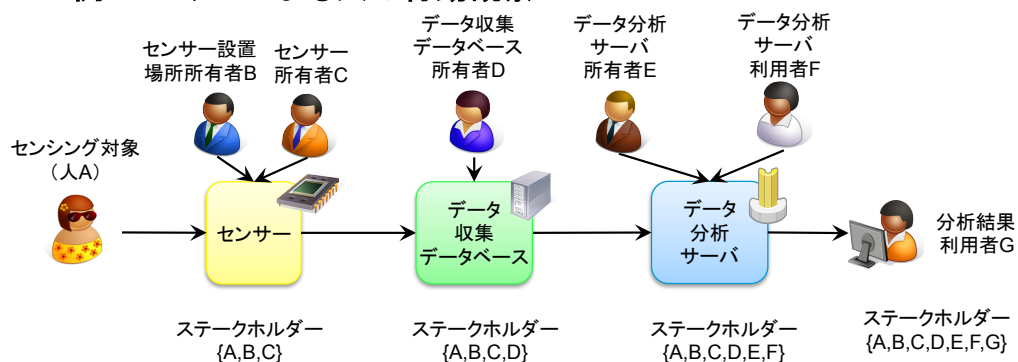
- 第三者提供は同意が基本、低減データは例外的に同意に代える手段。それ相応の加工／取り扱いは必要、海外との整合性も考慮すべき
- 低減データの加工を緩めると、個人は国内事業者は勝手にパーソナルデータを買ってしまうと考え、同意をきっちりとする海外事業者の方を信頼(海外事業者にパーソナルデータが集まる)
- 加工程度が緩いデータは受領側企業の経営リスクとなり、低減データの買い手がつかない
 - 加工程度が緩いデータの情報漏洩時の影響が大きい
 - 仮に提供側が属性情報に、(準)個人情報を残して提供してしまった場合、受領者が気が付かず利活用・提供の可能性(毒饅頭)

Ichiro Satoh

▶ マルチステークホルダー問題

- ビッグデータでは多数のステークホルダーが登場

- 例: センサーによる人の行動観察



- パーソナルデータの問題は、本人とデータ利用者などの2者問題
- それはマルチステークホルダー問題の一部にしか過ぎない

経産省の支援で、マルチステークホルダー問題研究会を設置

Ichiro Satoh

▶ なぜ個人情報保護法改正に関わっているのか

- 技術屋としての思い
 - 技術が生み出した問題は、技術で解決したい
- 現実
 - 技術が生み出した問題を、技術だけで解決できない
 - 制度による補完に頼るしかない
- これからのITと制度の関係
 - **技術と制度は不可分、一体で考えるべき時代**
 - 技術屋だから、制度は無関係といえない
 - 新しい技術を作ったならば、その技術を適切に使うための制度設計も行うべき

Ichiro Satoh

▶ 研究開発と制度設計は不可分



- 某社のウェアラブルカメラから学ぶ教訓
 - ウェアラブルカメラのプライバシー問題は知られていた問題
 - 製品発売と法制度の提案のセットでもよかったかも

Ichiro Satoh

まとめ

- ビッグデータの対象は多様だが、パーソナルデータに関心が集中
 - 保護と利活用の両立が必要
 - パーソナルデータは保護範囲にグレーゾーンが存在
- 個人情報保護法改正では
 - **グレーゾーンを減らすために、各種情報を整理**
 - **規律による保護を前提に同意なし第三者提供を導入**
- **技術と制度は不可分、一体で考えるべき**
 - **技術とともに、それを適切に利用する制度も設計すべき**

詳細は政府IT総合戦略本部「パーソナルデータに関する検討会」及び同技術検討WGの資料をご覧ください

Ichiro Satoh

技術WG 報告書

- 官邸ホームページに掲載

パーソナルデータに関する検討会 議事次第

http://www.kantei.go.jp/jp/singi/it2/pd/dai10/gijisidai.html

首相官邸 Prime Minister of Japan and His Cabinet

政策会議

トップ > 会議等一覧 > 高度情報通信ネットワーク社会推進戦略本部（IT総合戦略本部） > 関係会議等 > パーソナルデータに関する検討会 > 議事次第

第10回 パーソナルデータに関する検討会 議事次第

1. 日時 平成26年5月29日（木）17:30～20:00
2. 場所 内閣府本府 仮設庁舎講堂
3. 議事
 - (1) 開会
 - (2) 山本IT政策担当大臣あいさつ
 - (3) 技術検討WGからの報告
 - (4) オプトアウトに関する論点について
 - (5) 論点の取りまとめについて
 - (6) 閉会

<配付資料>

- 【資料1-1】 [技術検討ワーキンググループ報告書<中間報告からの変更点>（佐藤委員提出資料）](#)
- 【資料1-2】 [技術検討ワーキンググループ報告書（佐藤委員提出資料）](#)
- 【資料2-1】 [第三者提供におけるオプトアウトの適正な執行について](#)
- 【資料2-2】 [「名簿屋」に関する相談事例について（消費者庁提出資料）](#)
- 【資料3】 [パーソナルデータに関する検討会の検討予定（見直し版）](#)
- 【資料4-1】 [パーソナルデータの利活用に関する制度改正の基本的な考え方について](#)