資 料 ３

IoTが生み出すパーソナルデータの利活用と保護

国立情報学研究所　教授／所長補佐

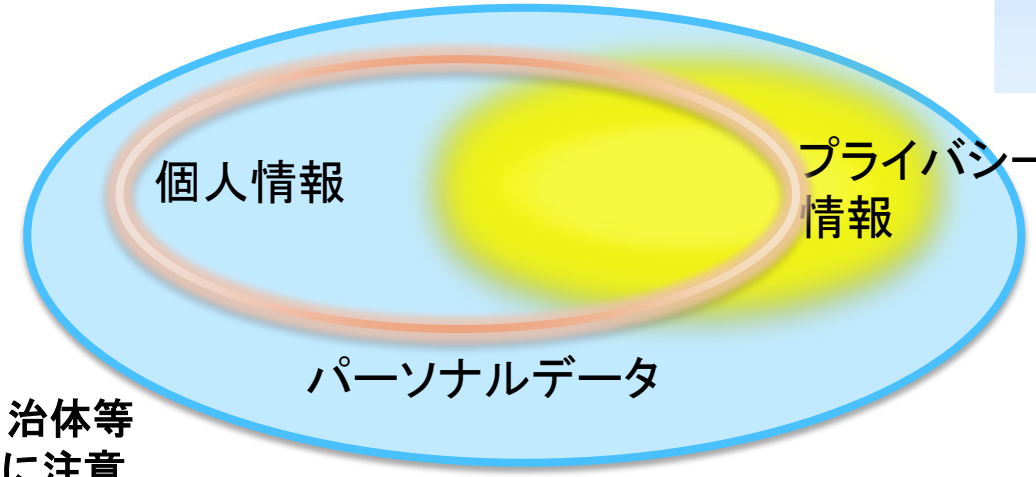
佐藤一郎

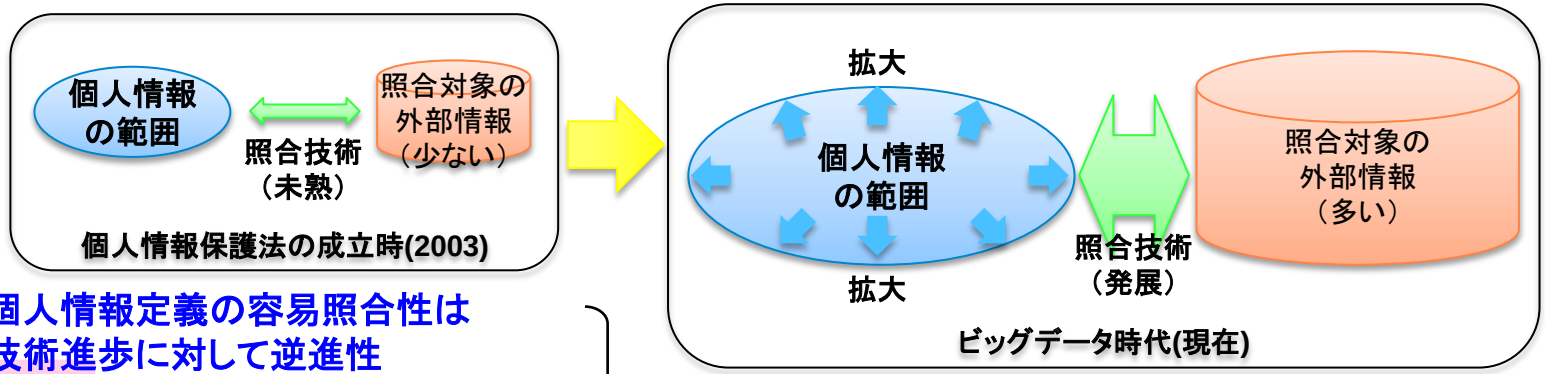
専門はIoTなどの分散システムのミドルウェア研究、プログラミングや実証実験するのが本来のお仕事

Ichiro Satoh

用語整理

- パーソナルデータ
  - 個人に関わる情報全般
- 個人情報
  - 個人情報保護法で定義
  - 特定の個人を識別する情報
    - 民間と行政機関、地方自治体等では定義が相違することに注意
- プライバシー情報
  - 本人がプライバシーだと思う情報（明確な範囲は定義不能）





- 個人情報定義の容易照合性は技術進歩に対して逆進性
（技術進歩により個人情報の範囲拡大）
  - 照合対象の外部情報が増大、照合技術の高度化
  - 新しい外部情報が増えている

個人情報の範囲は時代とともに変わる
（技術進歩への追随性とも言える）

改正個人情報保護法における個人情報の定義は、あくまで明確化
→個人情報範囲は変わっていない

Ichiro Satoh

パーソナルデータの分類

- **Volunteered data**（自発的生成データ）
  - 個人が生成し、明示的に共有されるデータ
    - 例：ユーザ登録、SNSの書き込み

個人情報保護法の範囲
ただし、同法の対象は、特定の個人を識別する情報であり、プライバシーではない。

- **Observed data**（観測データ）
  - 個人の過去の行動に基づくデータ
    - 例：防犯カメラ画像、購買履歴

IoTが生むパーソナルデータ
個人は観測されていることに気づくとは限らないし、利用目的や利用者は明示されていないことが多い。

- **Inferred data**（推定データ）
  - 自発的生成データ及び観測データから、推定・プロファイリングされたデータ
    - 例：SNSのユーザプロファイリング、クレジットカードの与信情報

個人はデータの存在さえ知らない。このためオプトアウト的取り扱いは向かない。
間違った推測による権利侵害が起きうるが、損害賠償による事後救済しかない。

この3分類はThe World Economic Forumが2011年に出した報告書「Personal Data: The Emergence of a New Asset Class」による

IoTによるユーザ行動履歴の収集

- **IoTの対象は現実世界**
  - 現実世界には人間が含まれることは多い
  - IoTが取得する情報には人間に関わるパーソナルデータが含まれる

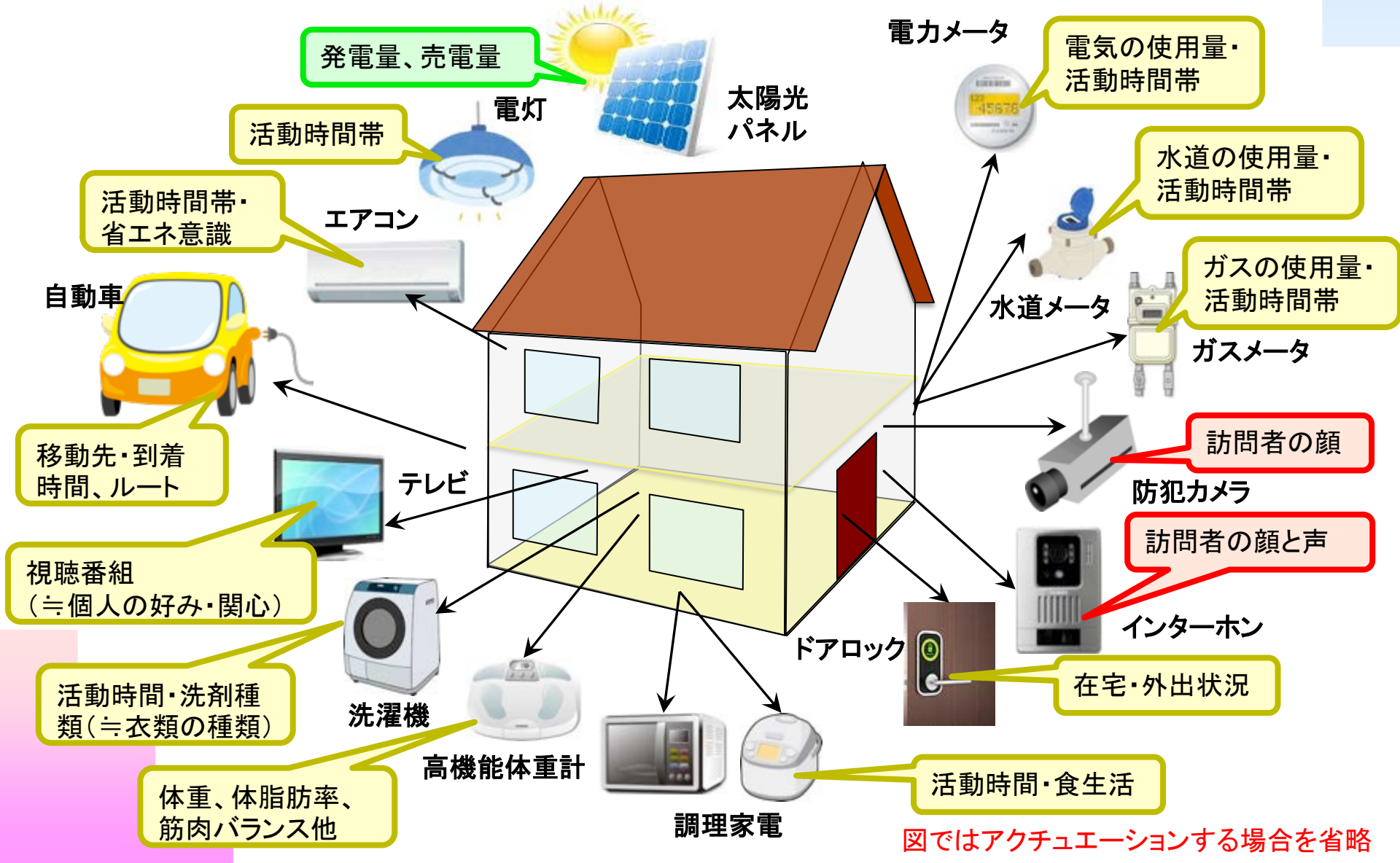
非接触IC カード改札	ポイントカード／ 電子マネー	ウェアラブル デバイス	ソーシャル ネットワークサービス	通話 アプリ
				
乗降駅と 乗降時間	購買品、年齢、 性別、生活圏	心拍数、血圧 ランニング距離・位置	関心事、交友関係、 移住地、学歴、職歴	通話記録、 位置情報、電話帳

大きなビジネスチャンス

- **パーソナルデータ、特に行動履歴から見えてくること**
  - **特定の個人の識別**
  - **個人のプライバシー**

IoTで覗けるプライバシー（住宅）

非個人情報
プライバシー
個人情報



図ではアクチュエーションする場合を省略

IoT利用とプライバシー侵害リスク

- IoTは多様な個人行動の捕捉や個人の識別を可能にする
  - これまで個人情報・プライバシー情報に関わっていなかった事業者もIoTでは関わることになる
  - 知らないうちに個人情報・プライバシー侵害を引き起こすリスク

    → 政府や業界団体による啓蒙活動は重要

- IoTによるデータ利活用でビジネスになりやすいのはパーソナルデータ関連
  - IoTにより個人行動履歴を中心としたパーソナルデータが容易に取得
  - その利活用は大きなビジネスチャンス、一方で個人の権利利益も重要

IoTで取得された情報は、個人特定やプライバシー侵害を高めるのか

現状ではIoTのセキュリティは高いとはいえず、情報漏洩を前提に制度設計すべき

IoTは非力なコンピュータと低速ネットワークで構成 → 高度なセキュリティ・暗号対策は困難

（インターネットとの直接接続ではなく、閉じたネットワークやゲートウェイによる保護が望まれる）

IoTと個人情報

- 断片的な情報でも、外部情報との照合により、個人の特定に至る可能性
  - 同一個人に関する複数情報の突き合わせにより、個人の特定や詳細化
  - 複数情報が同じ個人に関する情報であると見つけるのは簡単ではない

- 現実世界は一人の人間は唯一無二、同じ現実世界を他者と共有
  - IoTにより現実世界の情報の所得は容易化
    - 例：同じ世界（例：同じ空間&同じ時刻）に関する情報が複数存在しうる
  - 物理世界の指標（場所や時刻）により、同じ対象に関する情報と容易にわかる
    - 例：時刻情報が秒単位の鉄道乗降履歴と改札付近の（防犯）カメラ映像

      単体では個人情報にならなくても、同時刻の両情報を突き合わせれば、乗降履歴に関する個人の顔は容易にわかる
      （同じ空間＆同じ時刻に関するIoT情報の突合 → 個人の特定へ）

- IoTにより現実世界は、サイバー空間よりも、個人の特定可能性は高いともいえる
  - 一方で、IoTでは個人の権利利益の侵害につながらない情報も多い
  - IoTが取得したパーソナルデータをすべからく保護するのは賢明ではない
- 課題：個人情報やプライバシー保護しつつ、如何に利活用をすすめるか

購買履歴・ポイントカードから個人を特定できるか

- 日用品の購買履歴は個人の特定にならないが、特定できる場合もありえる
  - 購入品の特殊性
    - 一品物や限定品などの購入して、購入に関する情報が公知の場合
  - 購入形態・組み合わせの特殊性
    - （ありふれた商品でも）購入数や時間に特殊性
      - 例：毎週7個売れるが、一人の顧客が決まった曜日に7個購入（結構、多い）
  - 組み合わせにより、プライバシーは垣間見える
    - 例：カツラの購入と癌に関する書籍の購入→抗癌剤の服用可能性

  なお、購買履歴は店舗名と購入時刻が含まれるので、移動履歴でもある

- ベネッセ顧客情報漏洩事件他、網羅的な個人情報が存在する状況では、従来は個人の特定ができなかった情報でも特定できるようになってしまう
  - 例：スーパーのポイントカードの登録で、匿名ながら生年月日と性別を取得
    - 地域内（例：商圏内）の住民で生年月日と性別の組み合わせが一致する人は一人である可能性は高い（80歳までの生年月日と性別の組み合わせは58440通り）
    - → 外部情報（氏名、住所、生年月日、性別）があれば特定可能性は高い
  - ネットでもISPに接続された端末のIPアドレスと地域性には高い相関がある

鉄道の乗降履歴や位置情報から個人を特定できるか

- 乗降履歴（ICカード番号、乗車駅と時刻、乗降駅と時効）から個人を特定できるか
  - 同様履歴の個人が少数だと、外部情報照合による特定可能性は高まる
    - 例：単体の情報から個人を特定
      - 同一の乗降区間の利用者が少数となる区間の利用
    - 例：長期的な履歴から個人を特定
      - ありふれた履歴でも長期になると該当者は少数化
    - 例：精緻な時刻情報は個人の特定に重要

- 位置情報は状況に応じて個人を特定しうる
  - 人口密度が低い地位域では位置情報は個人を特定されやすい
    - 都会と過疎地域では状況が大きく違う、時間帯による人口変化に注意
      - 深夜の位置情報は自宅住所と考えられる（個人情報と扱うべき）
  - 最近の位置情報ほど、個人を特定されやすい
    - 現時点の位置情報は個人の特定につながりやすい（第三者が知りうる）
  - 移動経路の場合、出発点・到着点に対して、途中経路ではプライバシー的影響少ない傾向
  - 位置情報がプライバシーになるかは人による
    - 居場所を知られたくない人もいる（例：DV被害から特定者から身を隠しているなど）

PD研究会の技術WGの議論を思い出すと

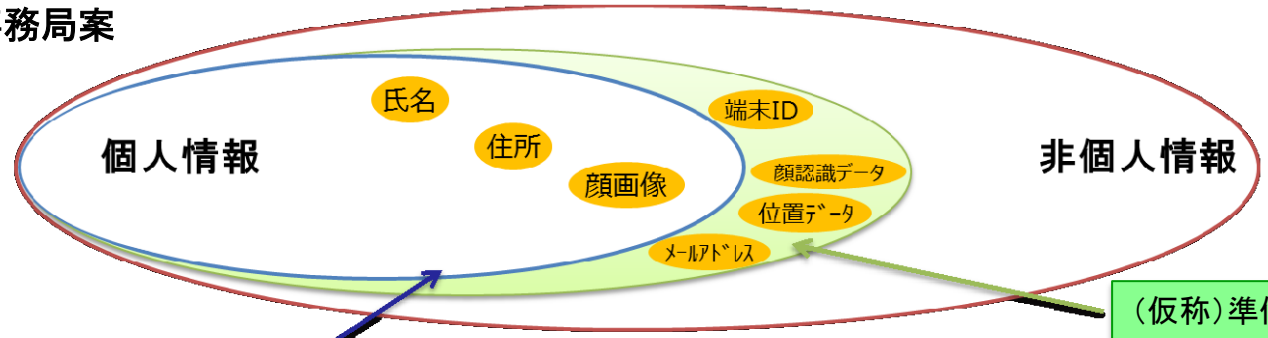
- IoTが生み出す情報そのものは個人情報ではないかもしれないが
  - その情報だけでは、誰とはわからないが、誰か一人に関する情報は、外部情報との照合により、個人の特定にいたる可能性は高い

昔の話をあえて書かせてもらうと

- 2014年、内閣官房PD検討会技術WGにおける議論においては、準個人情報という概念を導入することを前提にしていた
  - （一人ひとりを区別する情報であり）個人を特定しない情報だが
  - 複数事業者で共有される情報
  - 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

三要件

PD検討会事務局案



日本経済団体連合会意見書(2015年2月17日)では「情報の利用形態や用途によっては特定の個人を識別できない場合も存在する#3ことから、画一的な保護対象とすべきではない。」だから、準個人情報を提案したのに反対したのは・・・

IoTにより取得されるパーソナルデータと準個人情報は類似したところがある

Ichiro Satoh

準個人情報からみたIoT情報

- いまはなき「(仮称)準個人情報」の定義に該当する具体的な項目(技術WG報告書から)

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
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
上記ほか特定の個人の識別につながる多量又は多様な情報の収集を可能にするもの	行動履歴(例:Web閲覧履歴、購買履歴、条項履歴)はすべからく準個人情報ではないが、準個人情報として整理すべきものもある

一部は改正個人情報保護法の個人識別符号へ

不明確なまま

技術WGの議論:長期間にわたる行動履歴や該当者が少ない行動履歴は個人の特定に至るリスクが高まることから準個人情報となりうる

- 個人に割り振ったIDなどは比較的短い期間で変更が必要
- 該当者が少ない履歴情報は削除や一般化などの加工が必要
- 行動履歴にはプライバシー性が高い情報とそれ以外が含まれる(前者を保護すべき)
- 行動履歴は多様な情報が含まれることに注意(例:購買履歴は購買店名(場所)と時刻を含む)

- IoTにより取得されるパーソナルデータも同様のはず
  - 複数事業者で共有される情報を含む(ある一人に関する)行動履歴で、長期な行動履歴＆希少な情報は突き合わせによる個人特定リスクが高いことに留意すべき
  - すべからく保護すべきではないが、個人の権利利益侵害可能性がある情報は保護

隠すべき情報だけを隠す

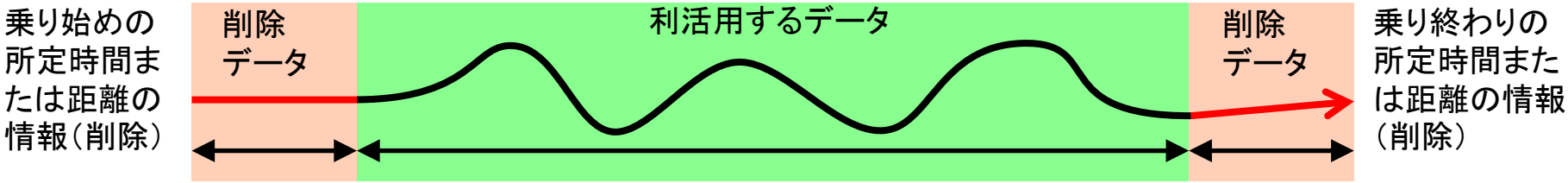
- 守りたい情報は何か、それに応じた対策をたてるべき
  - 情報そのものを隠す必要があるのか
  - その情報に関わる個人を隠したいのか
  - 個人の行動のうち何を隠したいのか

→ 一律に削除・加工するのではなく、守りたい対象を守るべき

例：欧州のカーナビ情報の匿名化（ユーザから利活用の同意取得）

TOMTOM

- 自動車や運転手を特定する情報（自動車ナンバーや所有者他）を削除
- 仮IDはナビゲーション一回ごとに割り当て（継続的トレースはしない）
- 乗り始めと乗り終わりのデータを削除（出発地と目的地の特定を防ぐ）

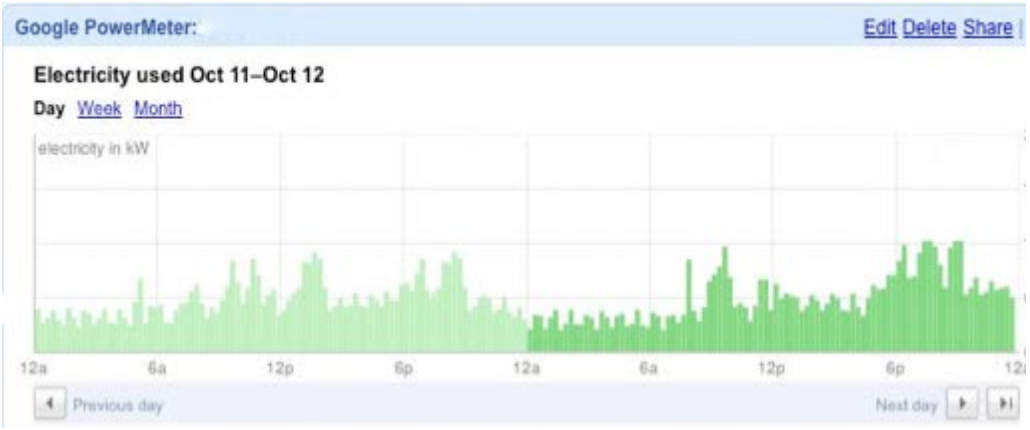


- ここではIoTによる現実世界に関するモニタリングを行う場合を対象としたが、IoTが現実世界にアクチュエーション（変化をもたらす）する場合は保護対象・方法は異なってくる
  - 後者では現実世界そのものが保護対象となりえる
  - 現実世界を守る法制度と、IoTに関わる法制度は結合・一体化する必要

複数データの組み合わせと責任関係

- 自宅（佐藤）の電力監視

  TED 5000 (Energy Inc.)



- 電力監視データをみても本人ですら電力消費と利用家電の相関がわからない
  - 単体のIoTデータの多くは個人特定やプライバシー侵害は少ないが・・・

- 個別データで見えないプライバシーも相違データの組合わせで見えてくる
  - 実例：水道メータとガスメータの遠隔共同検針の実証実験（2002年）
    - プライバシー情報が見えてしまい、実験中止

- データの組み合わせ方は事前に予測できない、組み合わせによる個人特定やプライバシー侵害の責任関係は議論・整理が必要

IoTが取得したパーソナルデータの取り扱い

- IoTで取得したパーソナルデータとその取り扱い
  - 個人本人はデータの取得を知り得えるとは限らない
  - データ取得の回数や種類は膨大となる

} 一般の個人情報の取り扱いは異なるべきかも

		個人情報の取り扱い（参考）	IoTによるパーソナルデータの特殊性
データ取得	利用目的	取得データの利用目的を特定	• 個人は取得に気づいているとは限らない • 目的が空間や状況依存になりやすい
	取得時	利用目的を明示	• 取得回数・手段多く、取得・サービス毎に説明されると煩雑、空間単位や携帯デバイス単位に取得や目的をまとめて明示も一つの方法 • データの保有期間も提示すべきか？（多くのデータは短時間で捨てられる）
取り扱い	目的変更	本人からの同意取得	• データ取得が容易ならば、新しい目的を明示して、取り直しも一つの方法 • 取得時の空間・状況変化時への対応
	正確性の確保	正確かつ最新の内容に保つ	• データ形式はセンサー依存し、人間は読めないことも多い • 測定エラーの正確性は事業者に負担が大きい
第三者提供	（一般の）第三者提供	本人からの同意取得または非個人情報に加工	• IoTが取得する情報はプライバシー性が高いことが多いことに配慮すべき
	匿名加工情報の第三者提供	削除または要加工	• 一般の個人情報よりも加工は困難？

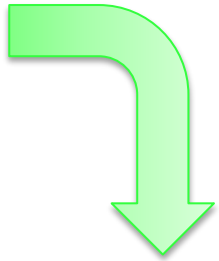
- 個人権利利益の侵害を基準に、情報を分類・保護すべきかもしれない

IoTにより現時世界をアクチュエーションする場合は別議論が必要

ネットサービスのビジネスモデル

ネットサービスのビジネスモデル

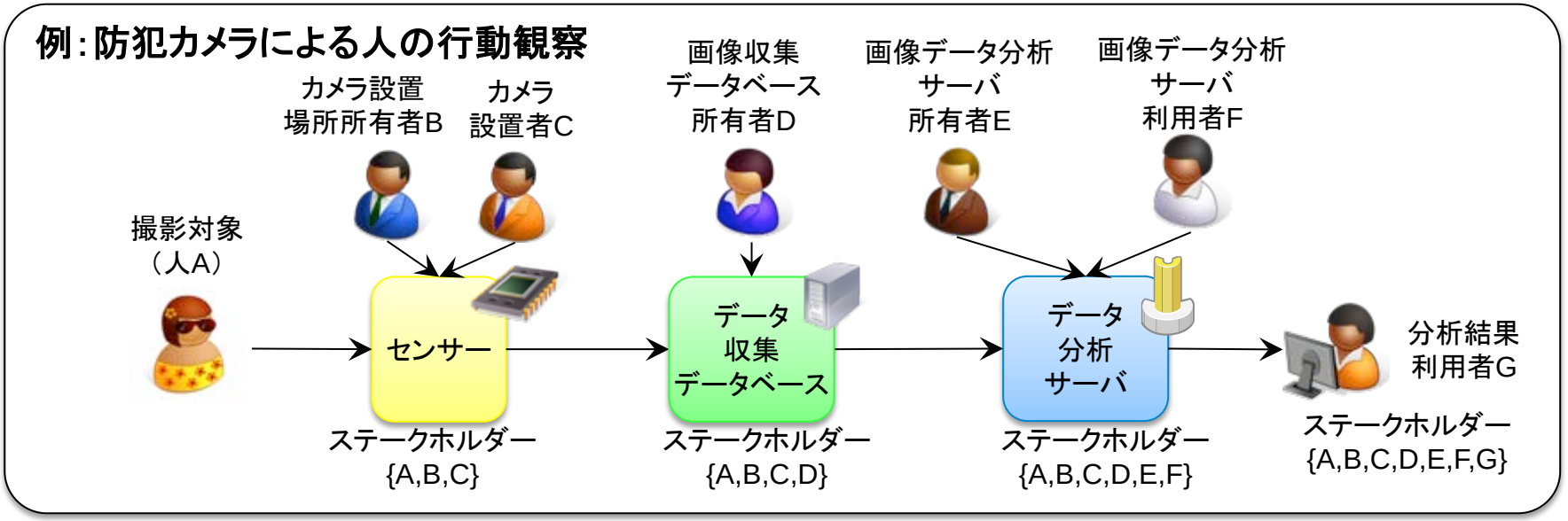
- 広告枠を広告クライアントに販売（ユーザは広告の視聴者）
  - 広告を表示することで、広告主から広告料を稼ぐ
    - 例：既存の多くの無料ネットサービス（Google他）
- サービスをユーザに販売（ユーザが顧客）
  - 所定期間・回数・取得情報に応じてユーザから利用料をもらって稼ぐ
    - 例：ネットゲーム、新聞オンライン版、LINE（スタンプ）他
- ユーザ情報を第三者に販売（ユーザが商品）
  - サービスを通じてユーザに関する情報を収集し、その情報を売って稼ぐ
    - 例：無料ヘルスケアサービス、（Twitter）他



- 現状IoTではネット広告を前提にしたビジネスモデルは成立しない
  - IoTの利用者は人ではなくデバイス、ディスプレーもあるとは限らない
  - 一方で、ユーザはネット広告を前提にした無償サービスに慣れている
    - → ユーザ情報を第三者に販売して収益確保
- 差別や不平等を生み出す可能性
  - 高く売れるパーソナルデータをもつユーザしか、IoTサービスが利用できなくなる可能性もあり、法制度による対策が望まれる

マルチステークホルダー問題

- IoTによる情報の取得、分析、利用にはステークホルダーが多数・複雑
  - データの対象者、所有者、設置者、分析者、分析結果の利用者他多様



- ステークホルダー間利害関係を調整できないとIoTは利用できない
  - IoTにおける個人情報／プライバシー問題は、個人本人とデータ利用者という2者問題となり、マルチステークホルダー問題の一部ともいえる
    - 個人情報・プライバシー問題だけにとらわれるべきではない
  - 解決すべきはマルチステークホルダーを前提にした利害関係の調整
    - 受益者の応分負担と利用者により提供データを削除・加工

まとめ

- IoTは多様化しており、IoTを分類して議論した方が良い
  - 例えば人間が含まれるIoTの系とそれ以外でも分類
  - IoTはマルチステークホルダーとなり、その調整が重要
  - 現実世界へのアクチュエーションがある場合は、保護すべきは現実世界
- IoTによるパーソナルデータの取得を個人本人がわかるとは限らない
  - 改正個人情報保護法は個人が主体的に提供した情報を暗に仮定しているのでは
  - IoTによる情報取得はガイドラインなどで補完することが望まれる
- IoTの特性を考慮して、保護すべき対象だけを保護すべき
  - IoTが集める人間の行動履歴はプライバシーと密接に関わる
  - すべからく保護すべきではないが、個人の権利利益侵害可能性のある情報は保護できる法制度が必要
  - IoTによる複数の行動履歴情報は現実世界が名寄せの起点になりやすい
  - 技術の進歩で保護対象は変わる（例：位置情報の高精度化）
- IoTを開発・運用する側も、起きうる問題を予測し、その問題を解決する技術や法制度を検討すべき
  - 技術と法制度は不可分とし、技術の開発段階から法制度を含めた検討が必要