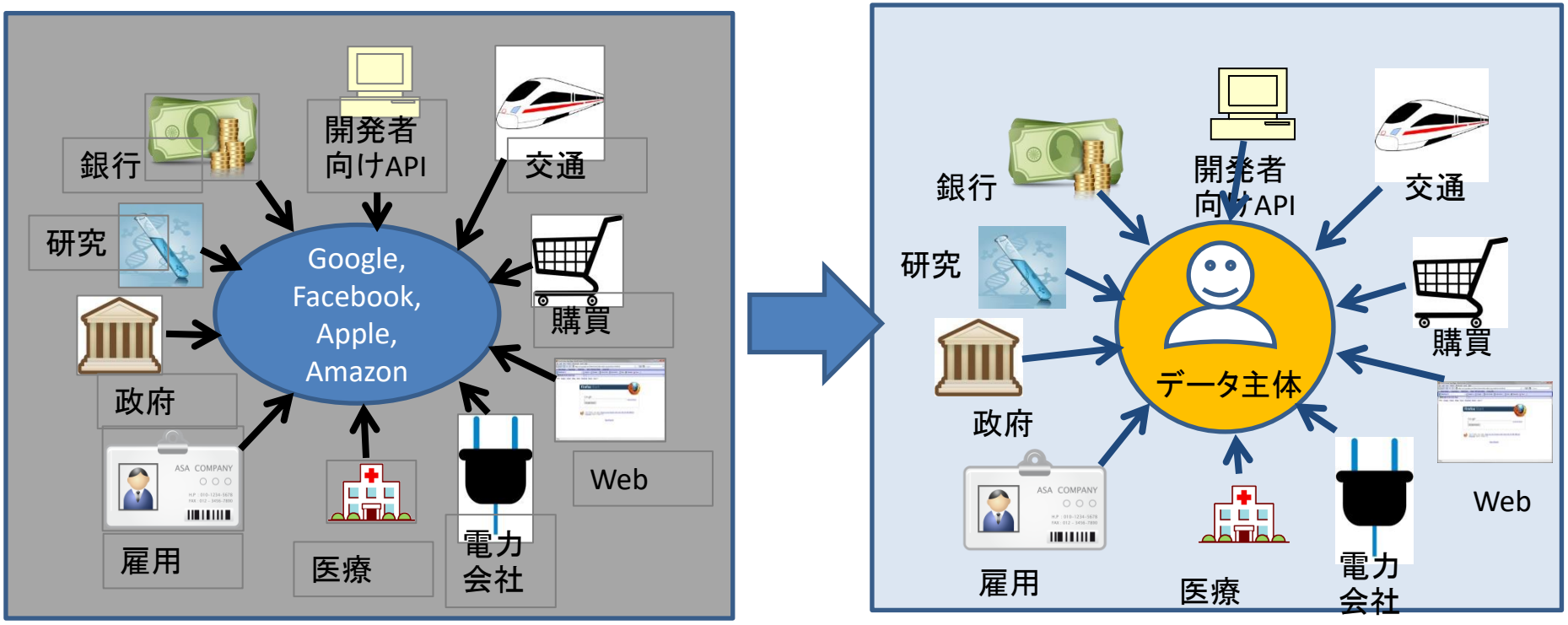
MyDataをめぐる状況

ー MyData 2016の報告ー

中川裕志
東京大学／理研AIP

個人データ管理は
データ主体の個人へ



背景：個人データ保護の流れ

- インターネットの普及によりデータ主体である個人の個人情報、個人データはネットに氾濫している。
- この情況に対してプライバシー保護の動きが強まっている。

- EUの動き
  - OECD8原則(EUだけではない)→ 種々のプライバシー保護法制の基礎
  - 以前のEUデータ保護指令　→ 法律は国毎に違う

- EU全域で統一したプライバシー保護の法律として
- GDPR(General Data Protection Rule)が2016/4/14(欧州議会採択)、2016/5/24施行、2018/5/25適用

背景:個人データ保護の流れ

- EU:GDPR(General Data Protection Rule)
  - 匿名化処理によって個人情報ではなくなり自由流通できると思う人もいるかもしれないが、
    - EUではこのような匿名化手法は基本的に存在しないとしている。
  - つまり、個人データ流通は以下が想定
    1. 事業者の説明責任と
    2. データ主体の同意による
  - 個人データを一業者が囲い込むことを許さず、データ主体個人の意志で個人に還元できるデータポータビリティを保障(GDPR　第20条)　→ MyDataに近い考え方

背景：個人データ保護の流れ

- 米国：プライバシー保護のための連邦法はない
  - 州法が主体。カリフォルニアでは先進的なプライバシー保護の州法がある
- オバマ政権のときに消費者プライバシー権利章典法 2015/5/27　が公開された。ただし、連邦法とはなっていない。
- FTC(Federal Trade Commission)や商務省が個別に管轄。
  - FTC3要件、FTC5条　が有名

背景：個人データ保護の流れ

- 日本：個人情報保護法改正 2015/9
  - 匿名加工情報の導入

- 世界の潮流
  - 忘れられる権利
  - プロファイリングの自動処理で得られた結果に服さなくてよい権利
  - Do Not Track (追跡拒否権)
- 等、個人のデータの収集、利用を制限する

  - 個人データは個人が管理し、同意に基づいて使わせる仕組み

残された課題：プロファイリングの問題

- IT事業者が収集したデータ主体の個人データを用いたデータ主体のプロファイリング
  - プロファイルを用いたターゲット広告：強力なビジネスモデル
- えてして不正確な個人のプロファイル
  - 自分のことを自分よりよく知っている！？
  - とんでもない被害
  - cf：英国では約90%に人がIT業者の個人データプライバシー保護を心配している。

残された課題：プロファイリングの問題

- プロファイルされた情報に基づくデータ主体への判断に服さなくてよい権利(GDPR 第22条)
  - プロファイル情報の開示要求
  - 間違ったプロファイル情報の問題
  - プロファイル拒否→追跡拒否権 (Do Not Track :DNT)
    - 業者への効力がほとんどない
  - 個人が同意して提出した個人データのほうが推定処理をしたプロファイル情報より正確でup to date

背景:IT企業と個人データ

- 米国のIT企業GAFA: Google Amazon Facebook Apple がパーソナルデータをどんどん収集して囲い込み、利益を上げている現状

- 収奪されるEU、収奪されるデータ主体の個人
- GDPRで反撃しているが、それだけではEUの産業は育たない
- EUの個人データのプライバシー(＝人権)の危機。だが、産業は興さないと低落するのみ

- 個人データはデータ発生源であるデータ主体の個人が管理
- その枠組みの標榜と、ビジネス育成がテーマ

- 2016年8月30日から9月1日　Helsinkiにて　MyData2016の会議開催　(今年も同時期に開催)

会場の様子

ORGANIZERS

OPEN KNOWLEDGE FINLAND

Aalto University

Fing

規模

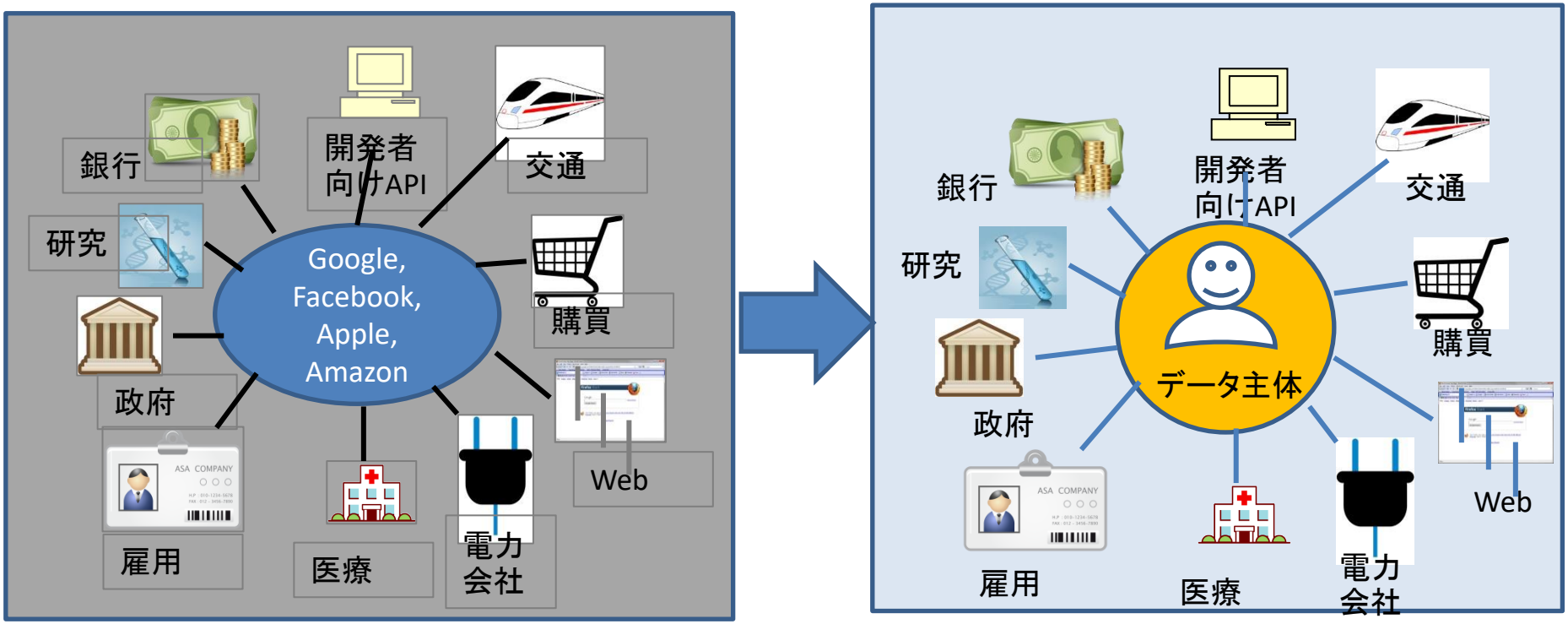
- 参加者　650人
- 発表者　140人
- 7会場の並列セッション構成
  - プログラムはここからアクセスできる

- フィンランド交通通信省が推進するMyDataプロジェクトが運営母体
  - 最終日に交通通信省の大臣が来て講演

- **SNSの活用**
- セッション中に会議サイトにコメントを書き込むと、その会議サイトで即時閲覧可能
  - どんな意見なのかがリアルタイムで分かる。
  - 講演への質問も書き込める

個人データ管理はGAFAから
データ主体の個人へ

- 個人データをディジタル人権に基づき産業に応用する。
- 標語：Make it happen, make it right!



メイン会場のセッション

- Opening - why are we here?
- Challenges for the data-driven society
- Show me the power of individuals
- Empowering people with their data
- Collaborating for a better data future
- Closing - ACTION!

MyDataのポリシー：
よりよいデータの未来像へ向けて
コラボしよう

- データ、アイデンティティ、プライバシー、セキュリティ、同意が合わさって、EUのディジタル経済を根本的に変えていくというメッセージ
  - 所有可能、取引可能、だが独占は許さない
  - → オープンデータ化に進展をめざす
  - API-of-Me による多数の人、企業からの操作は可能か
  - 公共の個人データをこのようにオープン化するには政治的なポリシーが必要と主張

本レポートの以下の部分

- MyData 2016での提案に沿って、以下のテーマを報告します。
  - 人々のMyDataに対する感覚
  - MyDataを実現する技術的要素
  - MyDataを用いたビズネスモデルの提案

人々の感覚の調査

- A trust-based framework for the data-driven economy：
  - Nicolo Zingales, Tilburg Institute for Law

  – 32%のプラットフォームは匿名化、仮名化を拒否

  – EUのユーザの52%は同意なしのデータ削除、88%は同意なしのアカウント削除にOK

  – 80%はユーザは第3者がのぞきにくることOK

  – 62%のユーザは商業利用でデータの共有OK

  – 52%のユーザはサービスをまたがるデータ集約OK

  – 38%のユーザ複数IoTデバイスをまたがるデータ集約OK

- TYPES　Project
  - Miguel Perez Subias : AUI Internet Users Association

- 90%は収集された個人データをcontrolできることは重要と思う

- 71%は代替サービスがないので、しかたなく個人データを晒してサービスを利用する。

- 1/7の人は収集されたデータの目的外利用に関心を持つ

- データポータビリティや透明性があれば、サービス提供者側に競争が生まれるので、良い方向になるが。

- TYPES (Towards transparency and privacy in the online advertising business) というプロジェクト
  - プライバシー保護を確保しつつ、サービス業者側が広告目的でユーザデータを利用できる仕組みを作る
  - つまり、ユーザに相当な自己情報コントロール能力を与える
  - この目的のための　ブラウザ・プラグインを開発

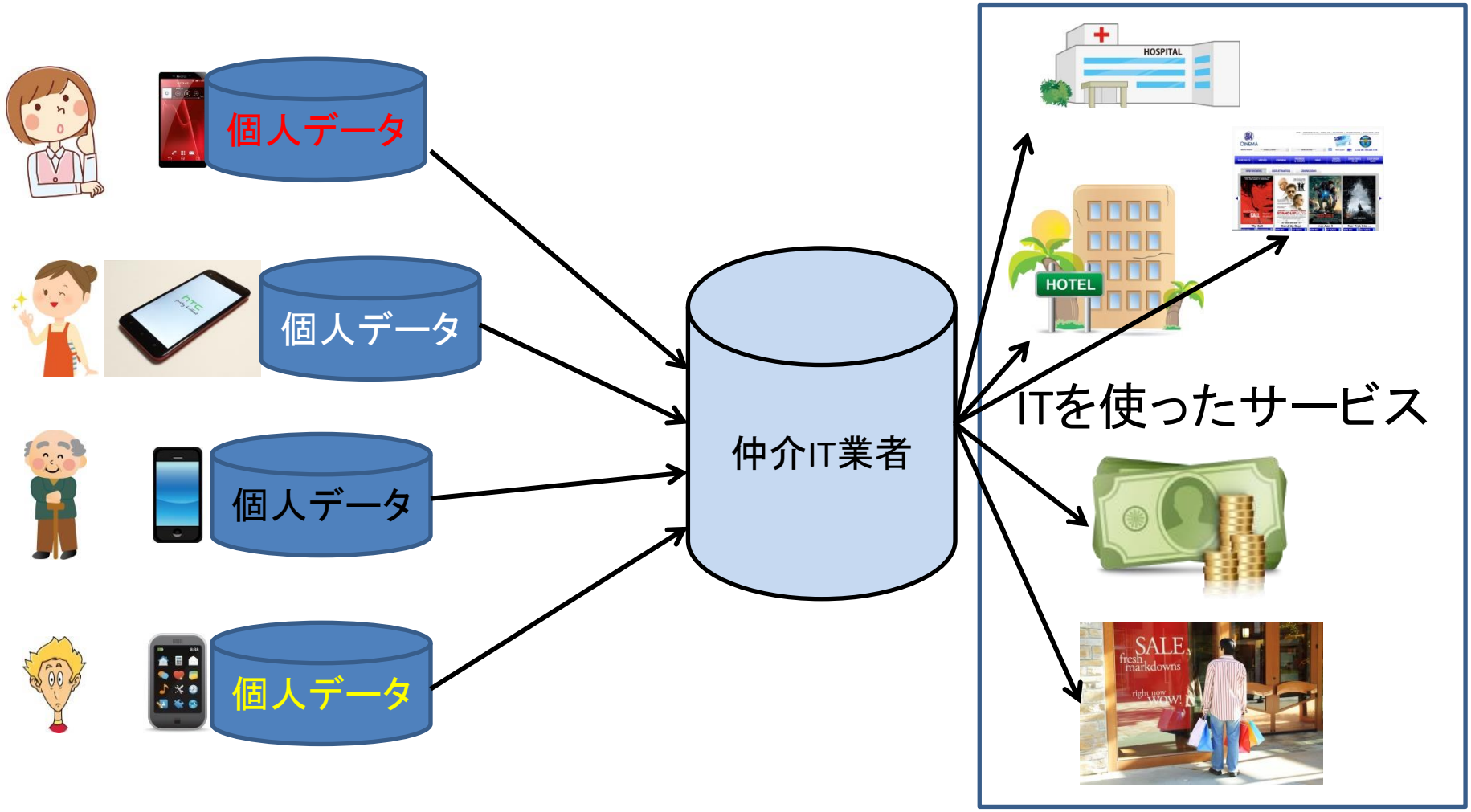
主要な技術的ポイント

- パーソナルクラウド
- インターネットにおける Identity 認証
- 個人データのポータビリティ
- Block Chain による個人の Identity 認証
- プライバシー保護(暗号化,複数当事者による計算 : MPC , etc.)
- 公平性、透明性の確保手段

ポリシーを実現する技術の様相

- インターネットIdentityの認証技術がテーマである。
  - OAuth, Open ID Connectの紹介
    - これらはインターネットIdentityの認証技術
    - OIDCはOAuthの上に構築されたIdentityの認証プロトコル
  - UMA (user managed access):
    - OAuth を用いて、情報資源管理者が利用者からの保護されている情報資源アクセス要求を制御するプロトコル
  - XDI:
    - 参加者間の契約(link contract)に基づくデータ交換プロトコル
  - VRM:
    - Doc Searlsのプロジェクトのツール群として実装

個人データを個人のデバイスで個人が管理



- 個人管理は面倒！（いい加減なセキュリティになりがち）

パーソナル・データ・ストレージ（PDS)

- パーソナル・データ・ストア／ボールト
- あるいは
- パーソナル・データ・クラウド








個人データ

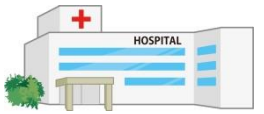
個人データ

パーソナル・データ・ストレージ

個人データ

個人データ

仲介IT業者



ITを使ったサービス


- 自動アップロード
- 個人キーで暗号化
- 個人ID認証
- API-of-Me

- 利用ログ
- 流通経路トレース
- 統一データ形式
- ポータビリティ

パーソナル・データ・ストレージ
Cozy Cloud

- 種々の操作での共通化、データ主体の同意管理は大変

- そこで personal data storage（PDS)からpersonal cloudに移行＝Cozy Cloud → PDSのクラウド化をサポートするシステム
- MesInfos:(Frace Fing) → データ主体が管理するパーソナルクラウド

- 特徴:
  - Single sign-on,global search
  - APPとの円滑な統合、IoTのハブになる
  - プライバシー確保

- 20名の社員で5,200,000ユーロ

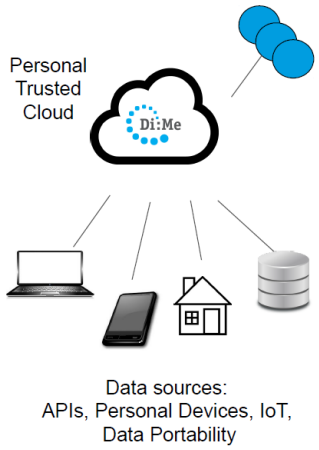
- Together, let's Uberize GAFAs!!

パーソナル・データ・ストレージ
Digital me

- Knowledge Worker（知識産業の従事者）が対象

- Digital Interaction は全てlogされる=Digital Footprint
  - プライバシー保護
  - 個人情報流通のトレース
  - → 個人による個人データのcontrolが必要

- 技術
  - Java, Spring Boot,
  - RESTful API, JSON
  - DB: SQL Hibernate or MongoDB
  - テキスト検索はLucene
  - MacOS X, Linux, Windows でローカルに稼働



パーソナル・データ・ストレージ
My Data Store

- TIM(Telecom Italia)
- ユーザによる管理と透明性によって信頼されたサービス (**Trusted** Services) に基づくエコシステム

  - 多ソースからのデータ収集
  - プライバシー バイ デザイン
  - 個人の完全なcontrol と透明性、

  - トラストできるAPPによるエコシステム、

  - 他との比較可能、詳細データのauditができる

個人データのためのブロックチェイン

- 個人のDigital IdentityをBlock Chainで管理する方法が焦点
  - Digital CATAPULT

  – コピー制限、流通経路の把握、identityの確認
  – プライバシーと個人IDの流出が危ない
  – 悪意ある行為の抑止方法
    ➢ゼロ知識証明で解決する方法の示唆あり

プライバシー保護データマイニング

- データマイニング結果の公平性確保
- データ処理事業者の説明責任
  - 開示要求：使われた個人データと利用法
  - 開示要求の受付と、処理の透明化
    - 機械学習アルゴリズム、データ処理プロセスに関する理解可能な説明
    - センシティブな個人データに関しては、データマイニングなどの結果のみ表示。
    - → 暗号化したまま行える秘密計算：MPC（Multi Party Computation）個人データは暗号化されたままなので、秘匿は完璧（計算プロトコルの工夫が必要ではあるが）

プライバシー保護データマイニング

- データ処理事業者の説明責任

◆実際は説明責任の実効的実装は困難

➢トラストの重視

➢定義：説明を受けなくとも信頼すること

➢過去の事業履歴＋　事業者の評判＋
想定被害額が大きくない

などの条件によりトラストが増加

➢ただし、なぜトラストできるか、あるいはトラストが崩壊するプロセスなどの　トラストの構造分析　はこれからの課題

パーソナル・データの利活用分野

- 健康、医療、福祉
- エネルギー、スマートシティ
- IoT
- 教育

- 移動（自家用自動車、公共交通）
- 保険とファイナンス
- 研究開発
- 雇用(雇用者の健康状態）

個人のエネルギー消費

- ENEDIS(フランスの電力会社)はスマートメータを導入(全家庭の20%)
  - 個人が自分のエネルギー消費量を知ることになより環境保護意識が促進できる
  - 家庭、ないし個人のエネルギー消費データを集めることはできるが、プライバシーが脅かされる。

Internet of Things: IoT

- センサーデータの管理と共有

- プライバシーを保護しつつ、データの共有と個人へのアクセス手段の提供を異なるプラットフォーム上で実現したい(TATAのTCS Crystal Ball))

- IoTから集まる個人データの収集経路などでのプライバシーリスクの明確化
  - 暗号化はデフォールトであろう
  - 行動データはかなり危険に晒されている

- 個人データの真正性はブロックチェインの提案多し

- Identityの確認（既存の　internet identity の仕掛けに言及したものは見当たらず）
- データの転送
  - 居場所、移動情報を含むトランザクション化、およびその内容の整備

- IoT固有というよりは、個人データ管理の問題一般に通ずる話

保険とファイナンス

- AXA フランス
- AXA Groupのデータプライバシーへの取り組み
  - ドライビング・パターンの抽出と保険
  - ビッグデータ分析の可能性と課題，法制度，倫理
  - 第三者機関French Data Protection Authority (CNIL) および
  - EUの15のデータ保護機関から承認を得ている.

医療応用プラットフォーム

- André Golliez, Adrian Wyss, Aline Zaugg “– MIDATA.coop – my data our health (SWISS)”（Platform economys/Engage ８／３１）

- 医療応用の話、様々なユースケースに対して実践的に実践している印象あり
- ビジネスモデルとしての共同組合方式(coop)など、興味深い

- ユースケース１ 手術後のフォローアップ
  - 対象　脂肪過多、胃のバイパス手術を受けた患者、
  - 利用アプリ：　MIMOTI – „Mini Motivation“
- 他にも多くのユースケースあり

- 各ユースケースでの共通項目　→信頼
  - Data protection, Data sharing, Mobile access through Apps, Citizen empowerment, Added value, Open platform with clear governance

信頼できる被雇用者

- 個人の健康状態などまで定量評価された被雇用者のイメージ
- 当然、以下のような問題がある
  - 労働者の権利
  - プライバシー
  - 技術的プラットフォーム

信頼できる被雇用者
Digi Clinic

- Mehiläinen：ヘルシンキの代表的な私立病院の一つ
  - 外務省ホームページで紹介されている3つの病院うちの1つ

- データの種類
  - ① 同社が有する Occupational Data
  - ② KanTa のデータ
  - ③ Self Measurement Data
    - 自己測定データが容易に測定・蓄積可能になってきたし，医師がそれを参照できるようになってきた．(例：血圧，睡眠時間，労働時間など)

- 産業保健において，これらのデータを用いると
  - 医師がデータから労働環境を評価できたり，比較できたりする．
  - 労働能力(Working Ability)を評価できる。これらのデータを集約(gather)しようとしている

Digi Clinic　つづき

- 労働環境と個人の健康・幸福
  - HeiaHeia (プラットフォーム)
    - ウェルネスや，ライフスタイルのデータ
  - HR プロフェッショナル (?) ＝法人顧客の話しを引用：
    - (法人は)被雇用者が自身の健康や労働能力に責任が持てるようにする必要がある.
  - 当初は，被雇用者がそうするためのツールが課題であった.
  - だが，Aikaniなど，ワークフローや時間を測定するためのアプリケーションが登場し，データを連携(combine)して，何にどれだけの時間を費やしているかが把握できるようになってきている.
  - メイン・イシュー
    - 産業保健の観点からは，従業員と医師の対話(Sessions)を支援するようなデータをどのように取得するかが問題である.
- 産業保健プロバイダーとして法人顧客に何をすべきか?
  - 我々は，被雇用者に対してコンサルやサポートができるし，彼らが働くことに関して責任が持てるよう補助することができる.

データ管理の経済性

- プライバシーはビジネスの起爆剤
  - 以下が一例

- スキポール空港を世界最高のデジタル空港のする
  → ポイントは
  - プライバシー保護
  - データセキュリティ
  - Trust＝（親密さ＋信用）／リスク
    - 親密さ＝対話性、透明性、公開性
- GDPRの精神を活かしたい
  - Privacy By Design

まとめ

- 個人データを個人管理というMyData
  - 必要性
  - 技術
  - 応用例

- ビジネスモデルとしての定着へ向けて

  - いろいろな問題が山積。例えば、
  - プライバシー・ポリシーを全部読むと、年間76日かかる。しかし、プライバシーは重要
  - ➢ 解決法は？

➢ビジネスチャンスと考えよう！