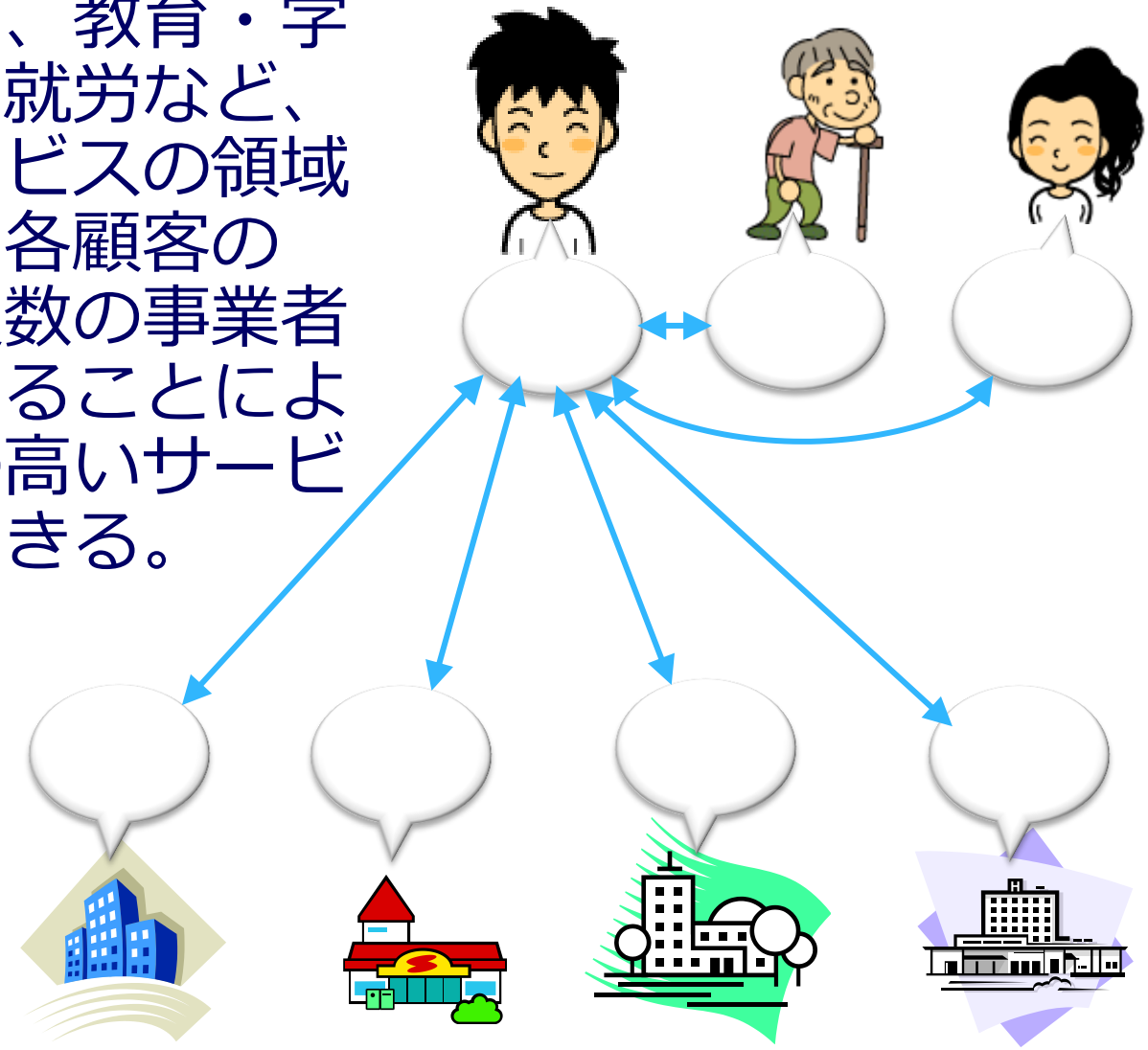
自己情報コントロールとPDS

2017-01-25 橋田浩一

東京大学大学院情報理工学系研究科
ソーシャルICT研究センター

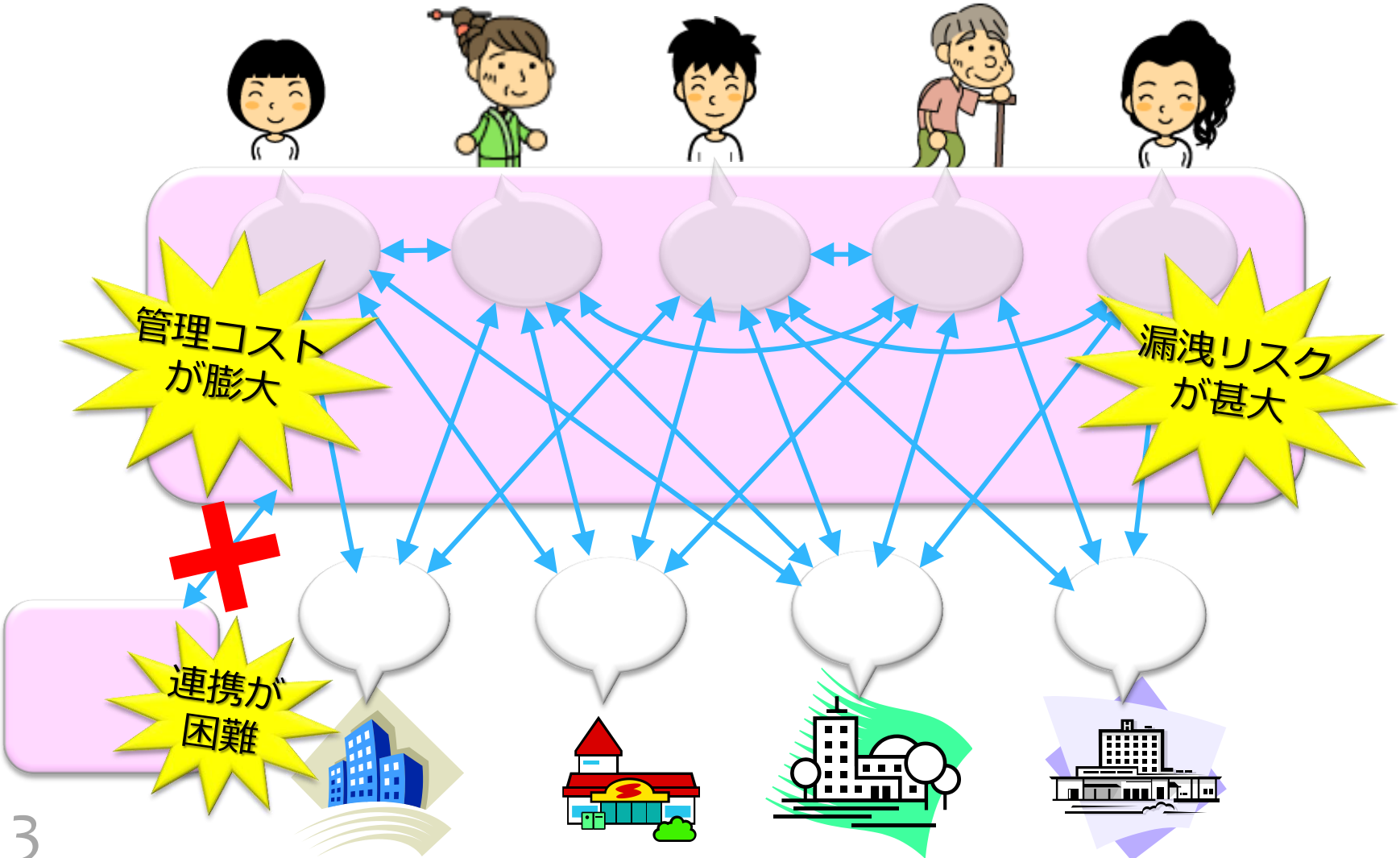
個人のデータを他者が共有するニーズ

ヘルスケア、教育・学習、観光、就労など、多様なサービスの領域において、各顧客のデータを複数の事業者等が共有することにより、価値の高いサービスが提供できる。



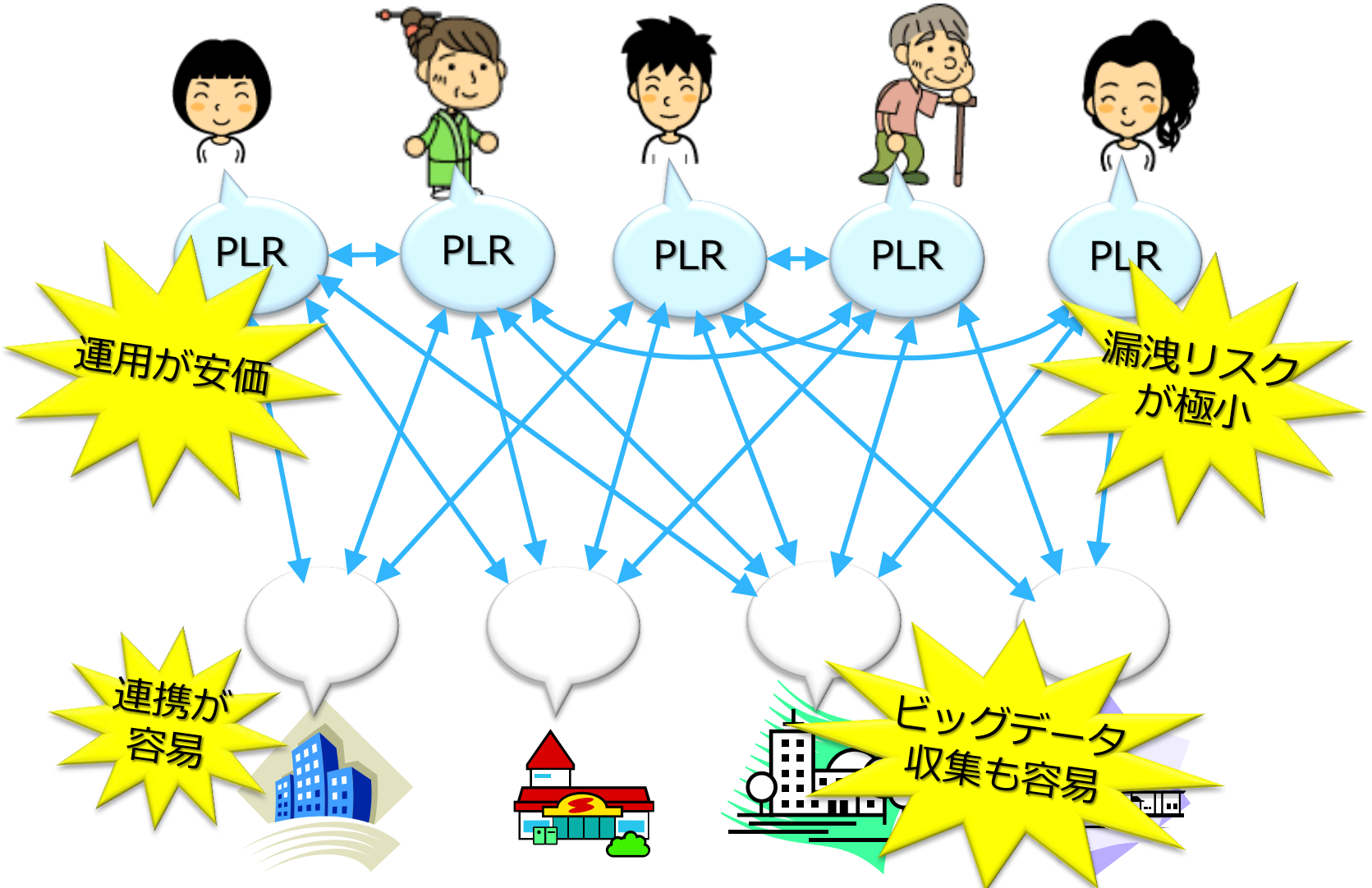
集中管理によるデータ共有

- 多数の個人のデータの集中管理は、各個人のデータの共有には不要であり、余分なコストとリスクを生む。
- 集中管理システムの間の直接的連携が困難なので、データ共有の一般解になり得ない。



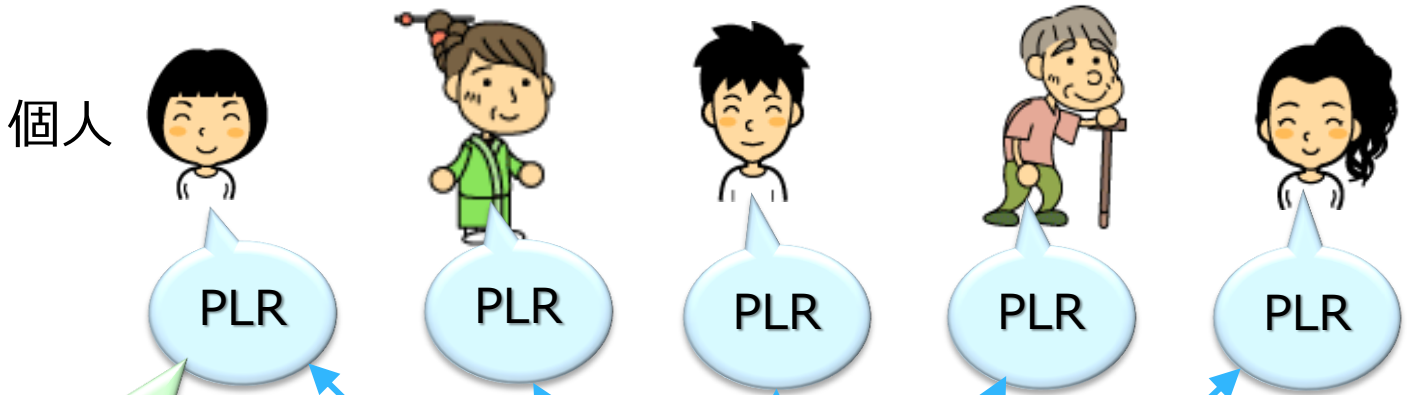
分散管理によるデータ共有

各個人が本人のデータを管理して必要十分な範囲で共有すれば、安価かつ安全に一般的なデータ共有が実現できる。



分散管理に基づくビッグデータ活用

- 個人が本人のデータを管理していれば本人同意に基づくデータ収集が簡単。
- ビッグデータ利用者は多人数のデータを永続的に管理する必要がない。



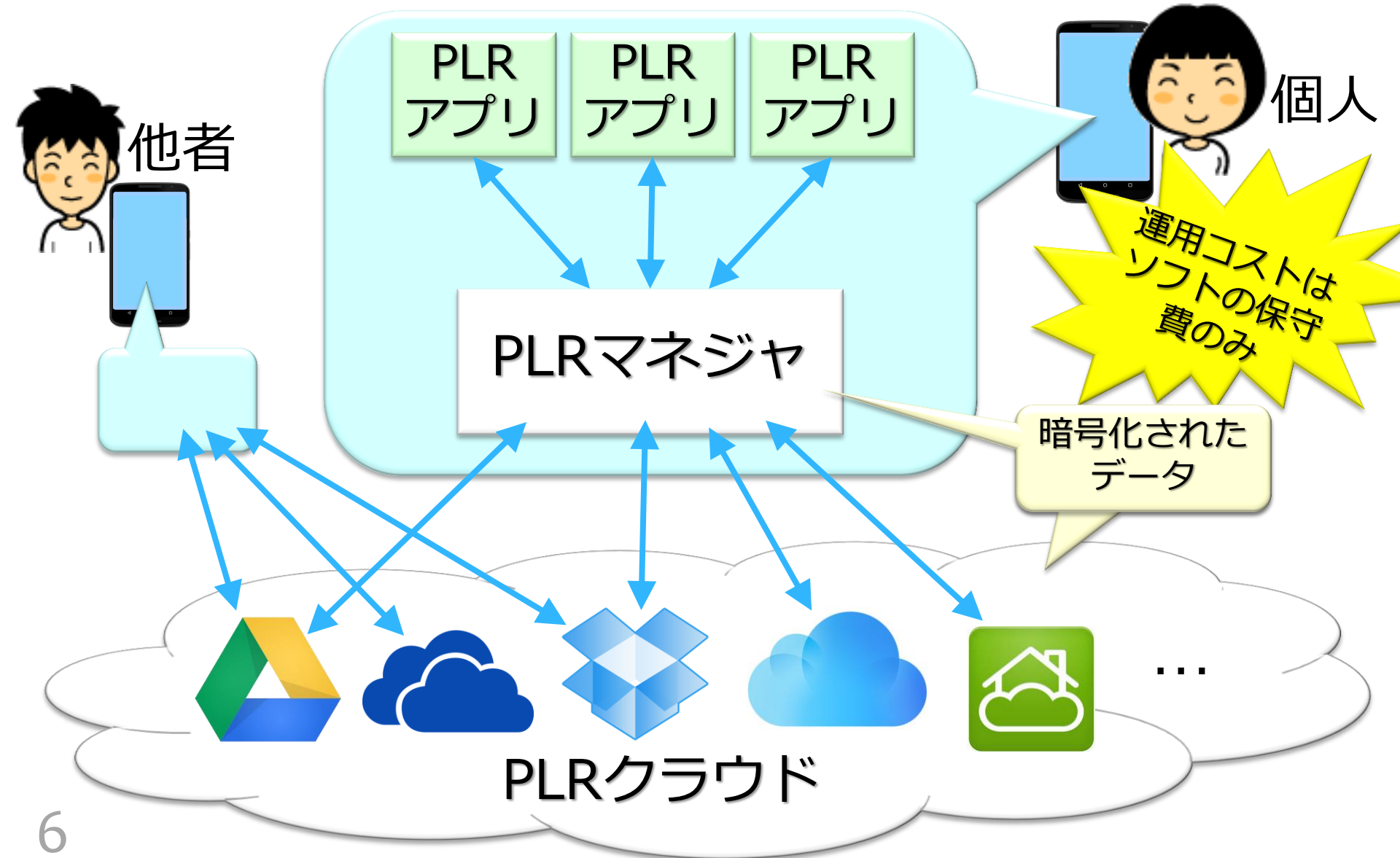
自分のデータがデータ提供募集の条件に合えば本人同意で提供(自動化可能)

ビッグデータ利用者

- ~~100万人のデータを集めて永続的に保持~~
- 必要に応じて1万人のデータを収集
- 分析が終わったら手元の生データを消去

PLR (personal life repository)

個人が必ずしも事業者に依存せずに本人のデータを他者と共有して活用



PLRは安全

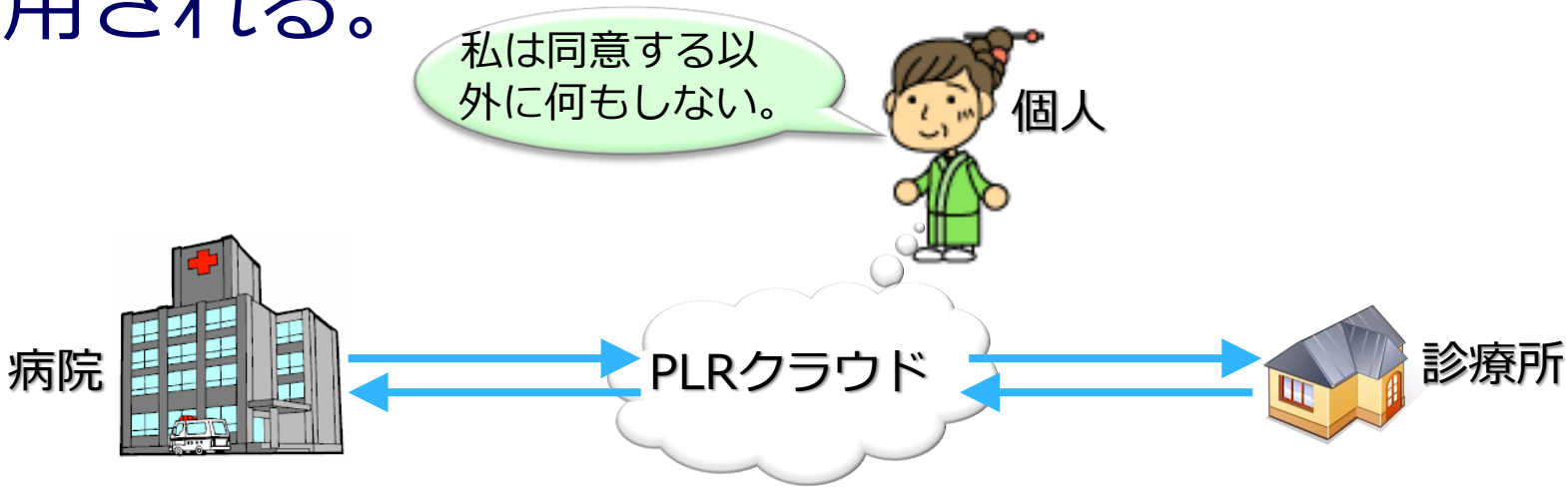
- データ管理を個人に分散
  - データを盗むコスト > メリット
- 多要素統合認証
  - 多要素認証
    - クラウドのアカウント(またはAPIトークンの入った端末)
    - PLRのパスワード
    - 公的個人認証(予定): マイナンバーカード + パスワード
  - アカウントアグリゲーション
    - 多数のアカウントを少数(5程度以下)にまとめることによって、人手による管理を可能にする
- DRM (デジタル権利管理): 暗号化 + アプリの機能制限
  - 平文データを書き出したり送信したりできない
    - 本人が間違ったり騙されたりしても、他人に認証を破られても、多量の情報が一挙に洩れることがない
  - 不適切なデータ共有を防止(予定)
    - 血液型占いのために住所を開示したりしない。

PLRは安心

- PLRのソフトウェアは企業(アセンブローグ社)が保守管理
  - Googleドライブがサービスを閉鎖するとわかったらOneDriveに乗り換える
  - 量子コンピュータが現われたら暗号の代わりに秘密分散を使う
- データは本人または代理人が運用
  - 運用者が開示しない限り、データの内容は上記企業にもわからない
- 個人が運用する場合、データが洩れるよりもパスワードを忘れたりIDカードを失くしたりするリスクの方が大きいが、それはアカウント管理支援サービスで対処可能

PLRは簡単

- ITリテラシは不要: データ共有を設定(委託可能)した後は、本人が端末を操作しなくても、指定された者の間でパーソナルデータが共有・活用される。



- 専門知識も不要: データのさまざまな部分の運用をPLRで他者に委託(信託)できる。

PLRの用途

- 多人数のデータを集める必要のないサービスは個人端末でできる
  - 業務システム(電カル等を含む)、EHR、PHR、他
  - 本人に中味がわからないパーソナルデータを本人が管理して専門家等に開示するEHR的な運用も可
- 多人数のデータを集める必要のあるサービス(検索や分析など)はサーバマシンでできる
  - たとえば、平文データをファイルに書き出したり外部に送したりする可能性のある不正なアプリをOSが排除すれば、不正なOSのインストールを防ぐ管理だけでセキュリティを安価に担保できる

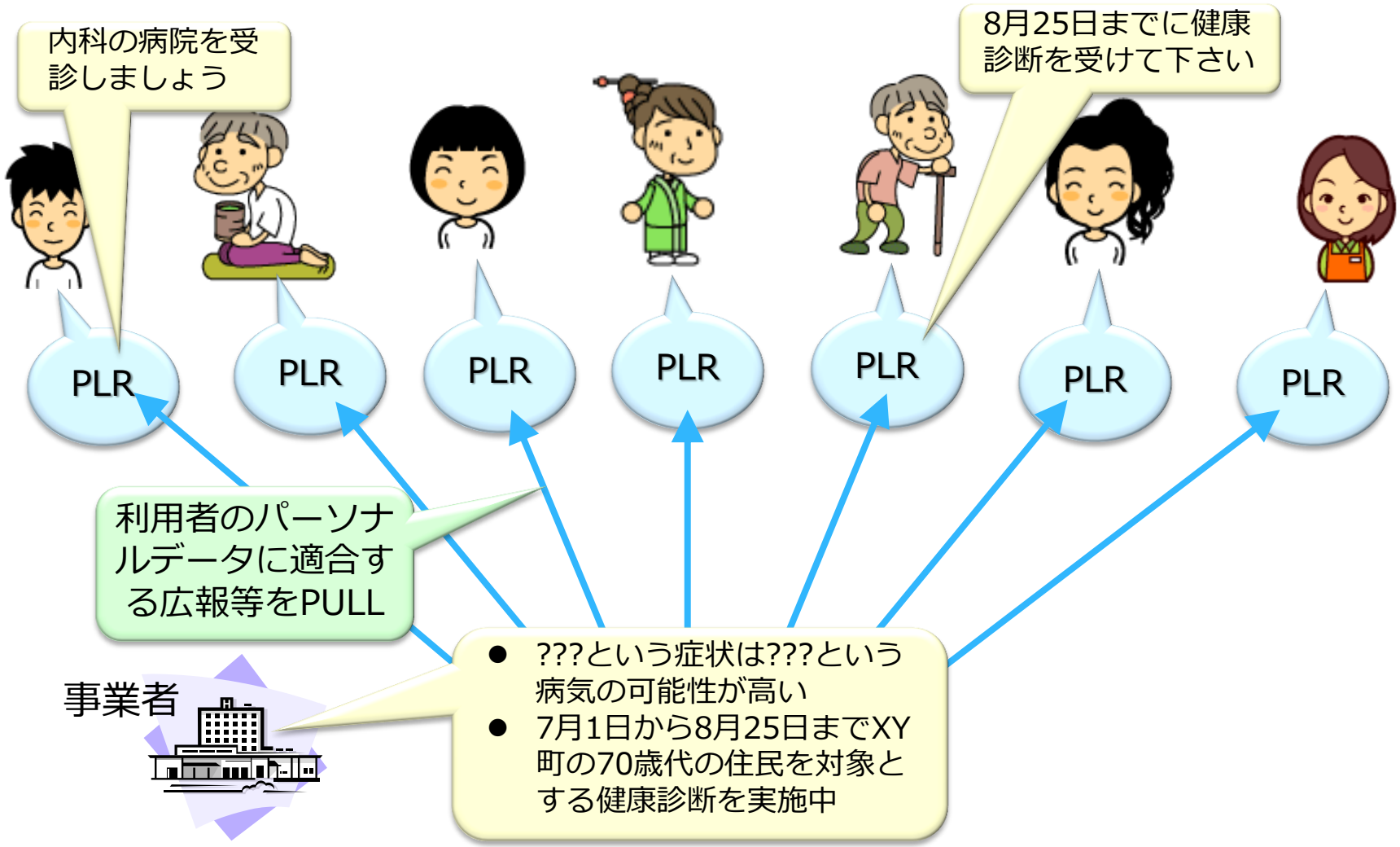
VRM: 業者関係管理

Vendor Relationship Management

- CRM (顧客関係管理; customer relationship management)の逆
- 顧客が自らの意思とデータに基づいて業者からのサービスや商品の買い方を最適化
  - 顧客のソフトウェアエージェントがパーソナルデータに適合するサービスや商品を選択
- 広告や推薦よりはるかに高精度で安価
- Berkman Center for Internet and Society, Harvard Univ.の研究プロジェクト

VRMの基本形

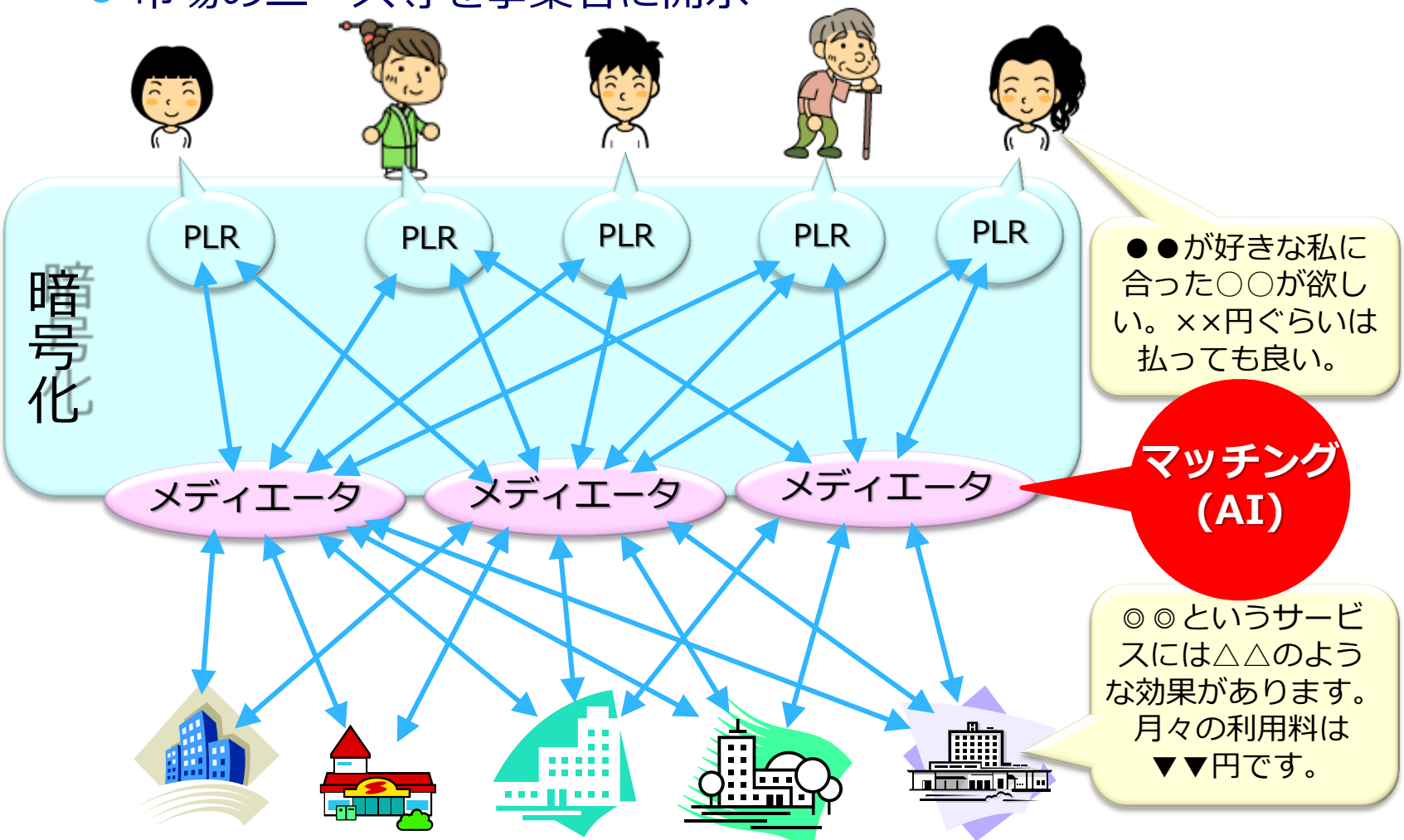
- 利用者のパーソナルデータに適合する広報等を分散PDSがPULL
- 事業者は個人情報を見ずに行動ターゲティング以上のことができる



メディエータ: 大規模個別マーケティング

プライバシを担保しながら需要と供給の間の相互作用を運営

- 個人のニーズとサービス・商品の情報を集約してマッチング
- マッチングの結果を個人に開示
- 市場のニーズ等を事業者に開示



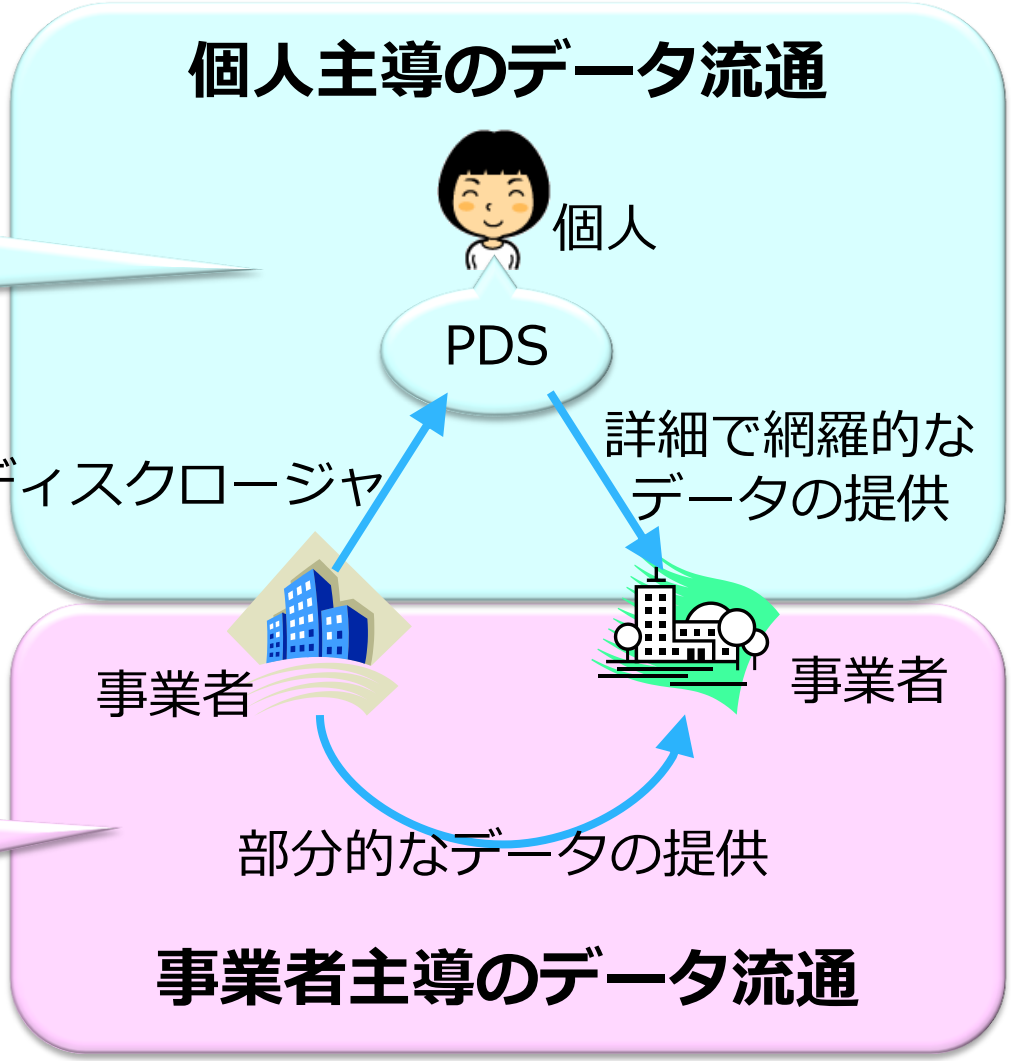
基本概念

- 自己情報コントロール
  - 自分に関するデータを自らの意思で活用すること
  - 活用: 本人による利用と他者への開示
- データポータビリティ
  - (本人の意思に応じて個人の)データを移転できること
  - 扱いやすい(標準形式であるか、少なくとも仕様が開示されている)電子データの移転が前提
  - 自己情報コントロールに必要
- スマートディスクロージャ
  - 扱いやすい電子データによる情報開示
  - 特に、パーソナルデータを保有する事業者等がそれを扱いやすい電子データとして本人に開示すること

パーソナルデータの流通

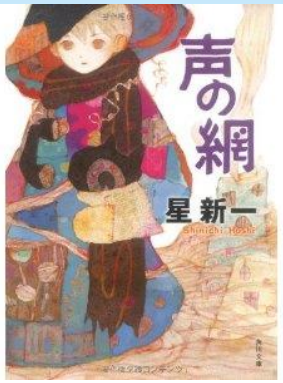
従来の事業者主導のデータ流通に加えて、個人主導のデータ流通の普及が必要

個人は他者よりも詳細で網羅的な本人のデータを集めることができ、それを自らの意思で他者に提供することもできる。

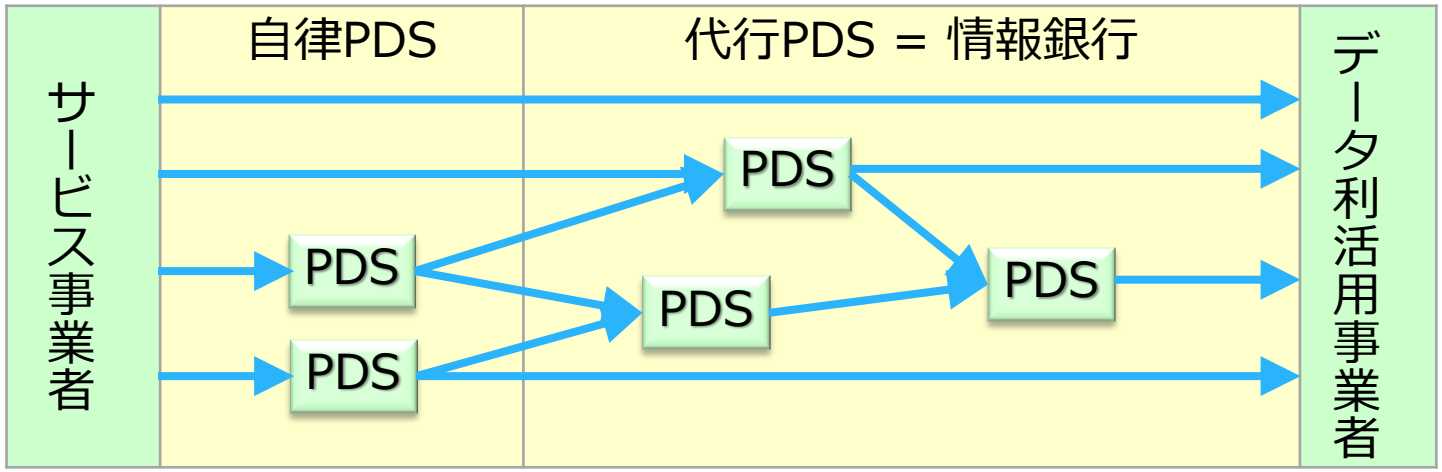


事業者が保有する個人のデータはきわめて部分的であり、匿名化するとさらに有用性が下がる。

PDS



- Personal Data Store
  - パーソナルデータを本人の意思に基づいて運用する仕組み
    - 運用 = データ開示の実効的判断
    - 必然的にデータポータビリティを満たす
  - 概念自体は新しくない: 星新一(1970) 声の網. (情報銀行)
  - Gordon Bell (2001) A Personal Digital Store. *CACM*, 44: 86–91.
- PDSシステム(PLR、Personium、OpenPDS、MesInfos、etc.)の用法
  - 自律(autonomous): 本人が自ら運用
    - より分散的(decentralized): 少人数のデータを運用
  - 代行(surrogate): 他者が運用を代行
    - より集中的(centralized): 多人数のデータを運用
  - 各PDSシステムはいずれの用法も可能



データの運用 = マッチング+開示の判断

- パーソナルデータとのマッチングの対象
  - 商品、サービス、求人、求職、結婚相手、他
- マッチングに応じたデータ開示
  - マッチングしたサービスの享受に必要なデータをサービス提供者に開示
  - マッチングした研究グループにデータを開示
- 運用代行が必要な場合
  - マッチングの対象に関するデータが大きすぎて個人端末に入らない
  - 本人も個人端末の中のAIもマッチングができない
- メディエータは大規模な代行PDS(情報銀行)

制度の動向

- 日本
  - 2015年に改正個人情報保護法改正と改正マイナンバー法が可決
  - 2016年からマイナンバーの施行
  - 2016年に官民データ活用推進基本法が可決
    - 第十二条 本人関与の下でのパーソナルデータの流通
  - 2017年7月からマイナポータルの運用
- EU一般データ保護規則(GDPR)が2016年4月に可決
  - 消去の権利とデータポータビリティの権利
  - 「第三国条項」…パーソナルデータに関する十分なレベルの保護が行われていない第三国へのデータ移動の禁止。
  - 個人の仕事の業績や経済状況、健康状態、嗜好等の自動分析(プロファイリング)に制限
- EU (法規制) vs. 米国(自主規制)
  - 2015年10月セーフハーバー協定無効判決
  - 2016年2月EU-USプライバシーシールドに合意
- 日本企業が海外(特にEU)でパーソナルデータを扱う事業をするにも、本人同意によるデータ活用の仕組みが必須。
- EUのPSD2 (Revised Payment Service Directive)が2018年に発効
  - 銀行はAPIを開示せねばならない → 客のスマホがPOSレジ?

医療制度改革

医療機関等の間でのデータ共有が必須に

- 医療機関(病床)の機能分類を2018年から運用
  - 高度急性期、急性期、回復期、療養期、診療所
- 異種医療機関同士のデータ共有
  - 急性期病院は、退院患者の再入院を防ぐため、受入先の回復期病院や診療所に患者のデータを渡さねばならない。
  - 回復期病院や診療所は、急性期病院等から退院患者を多く受け入れるため、患者のデータを受け取って治療の成績を高める必要がある。

医療制度改革

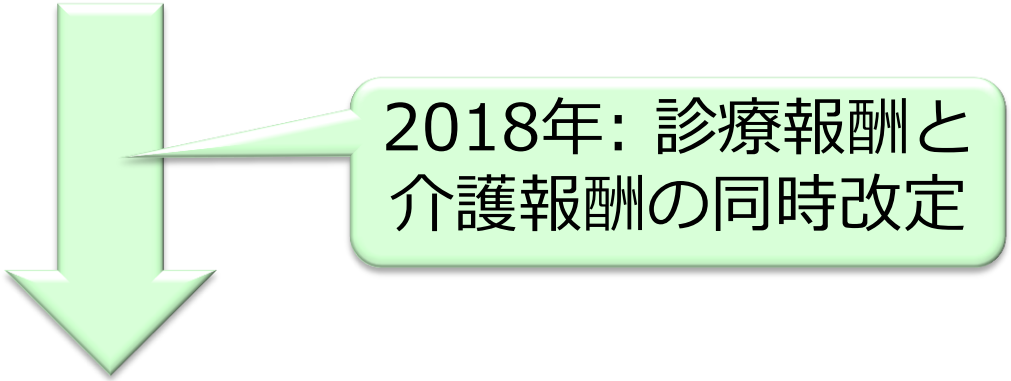
医療機関等の間でのデータ共有が必須に

- 診療所同士のデータ共有
  - 各患者に24時間365日の在宅医療を提供するため、複数の診療所(各々はほとんどが医師1人)がグループを組んで患者のデータを共有せねばならない。

医療制度改革

医療機関等の間でのデータ共有が必須に

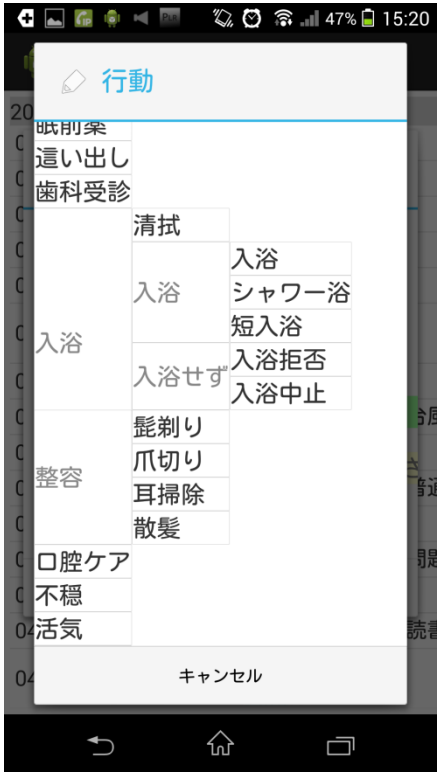
- データを共有しても儲からない

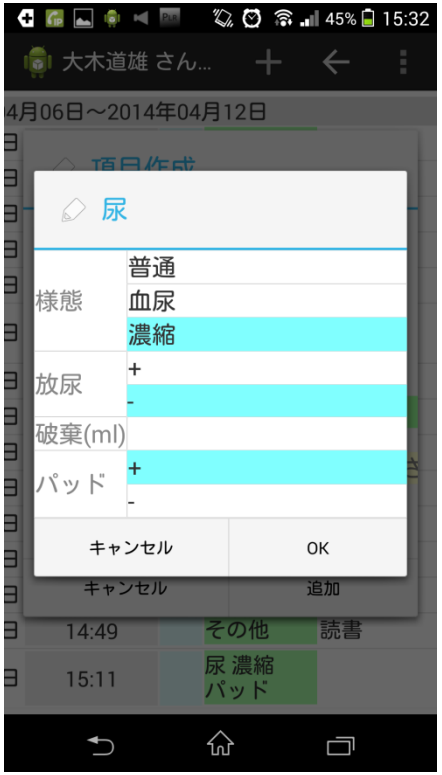


- データを共有しないと経営が成り立たない

PLR介護記録アプリ

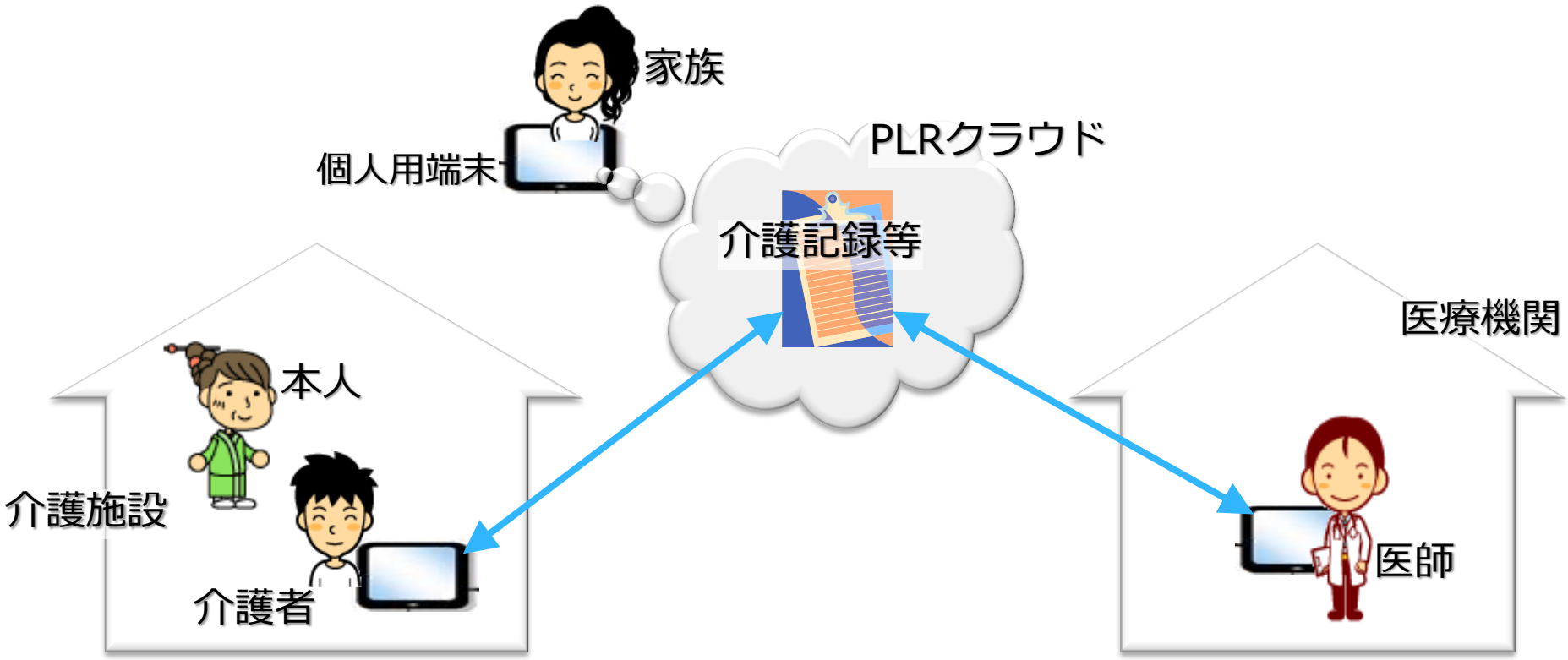
- 山梨と鳥取で被介護者70名を対象に実運用中
- サーバレス
  - 利用サポート不要
  - 無料で配布・運用可能
- オントロジー(データのスキーマ)の変更が容易
  - 訪問医療や訪問看護用のカスタマイズ
  - がん連携手帳やお薬手帳の実装
  - 電子カルテシステムの簡易版も





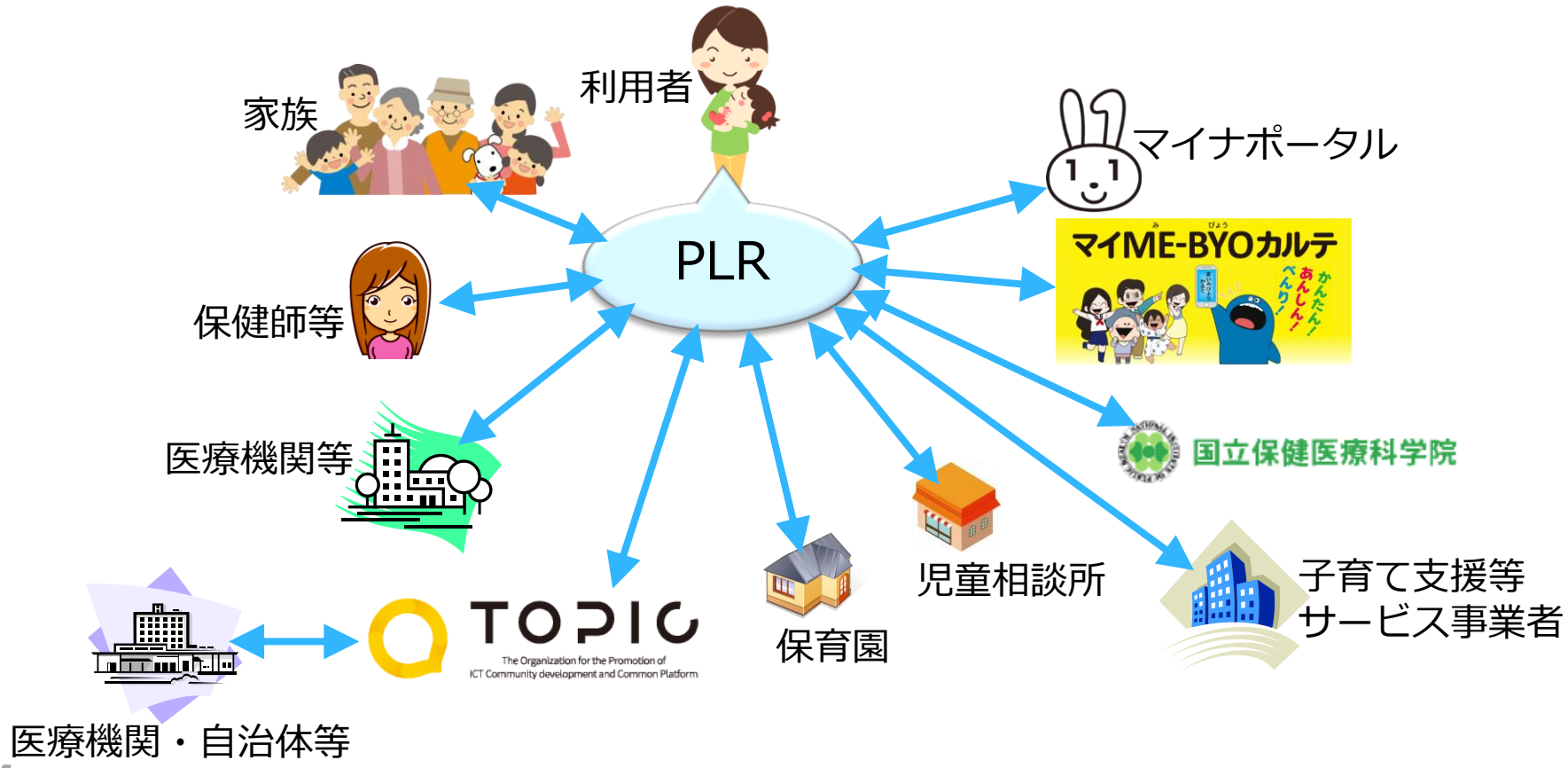
ヘルスケアデータの個人管理

- 介護記録のデータを本人(の家族)が管理して関係者と共有可能に
- 山梨と鳥取で約70人のうち2人の高齢者について運用中
- 分散システムなのでそのまま何億人にでも拡張可能



母子保健での活用
(AMEDのH28年度PHR利活用研究事業1次公募の(1))

- 母親が医療機関等の事業者からデータを取得
- 保健師等の業務を電子化することにより、負担を軽減するとともに、母親とリアルタイム・双方向に情報共有

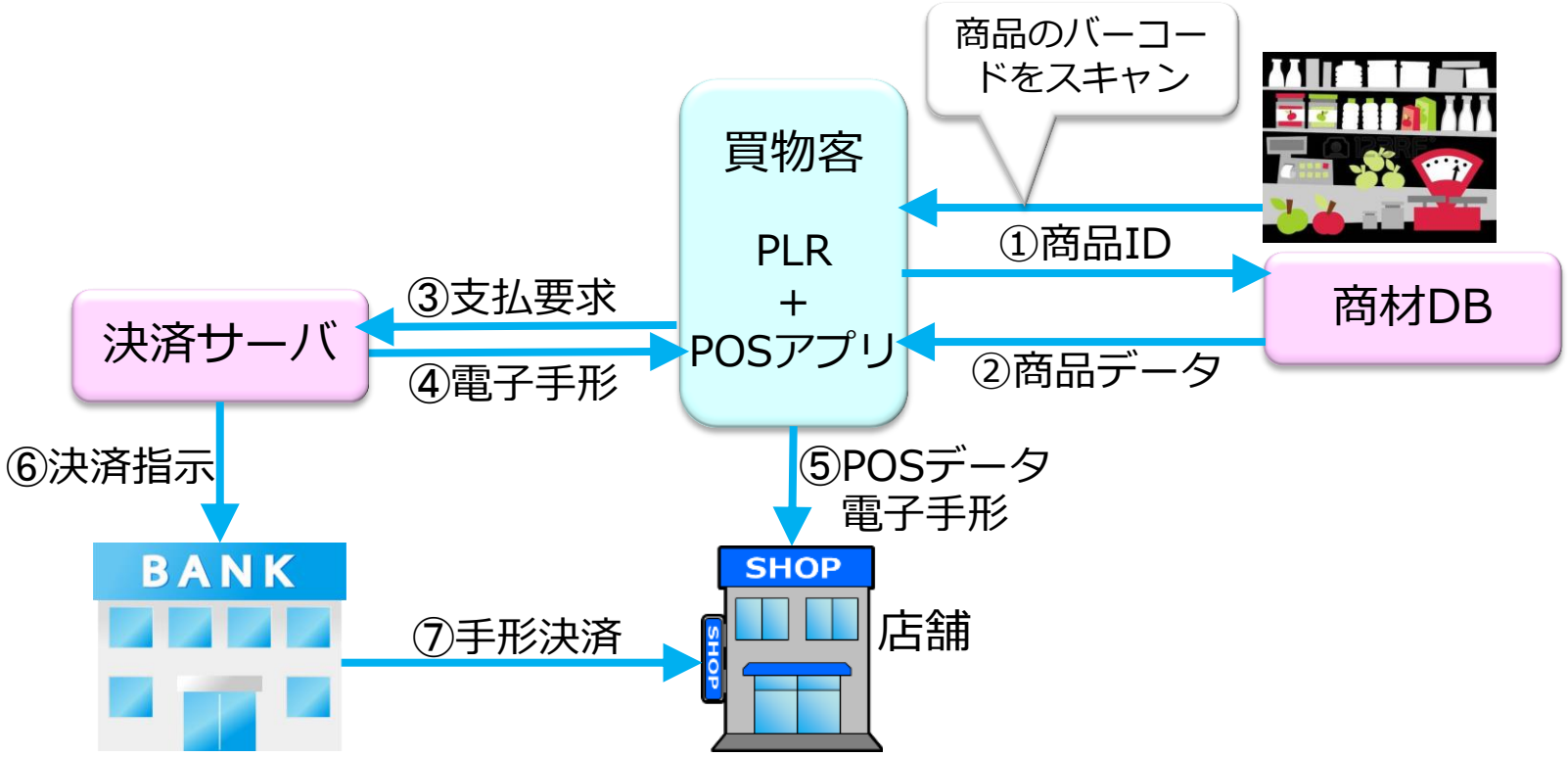


臨床での活用
(AMEDのH28年度PHR利活用研究事業2次公募の(2))

- 各実証フィールドにおいて既存のEHR等とPLRとを連携

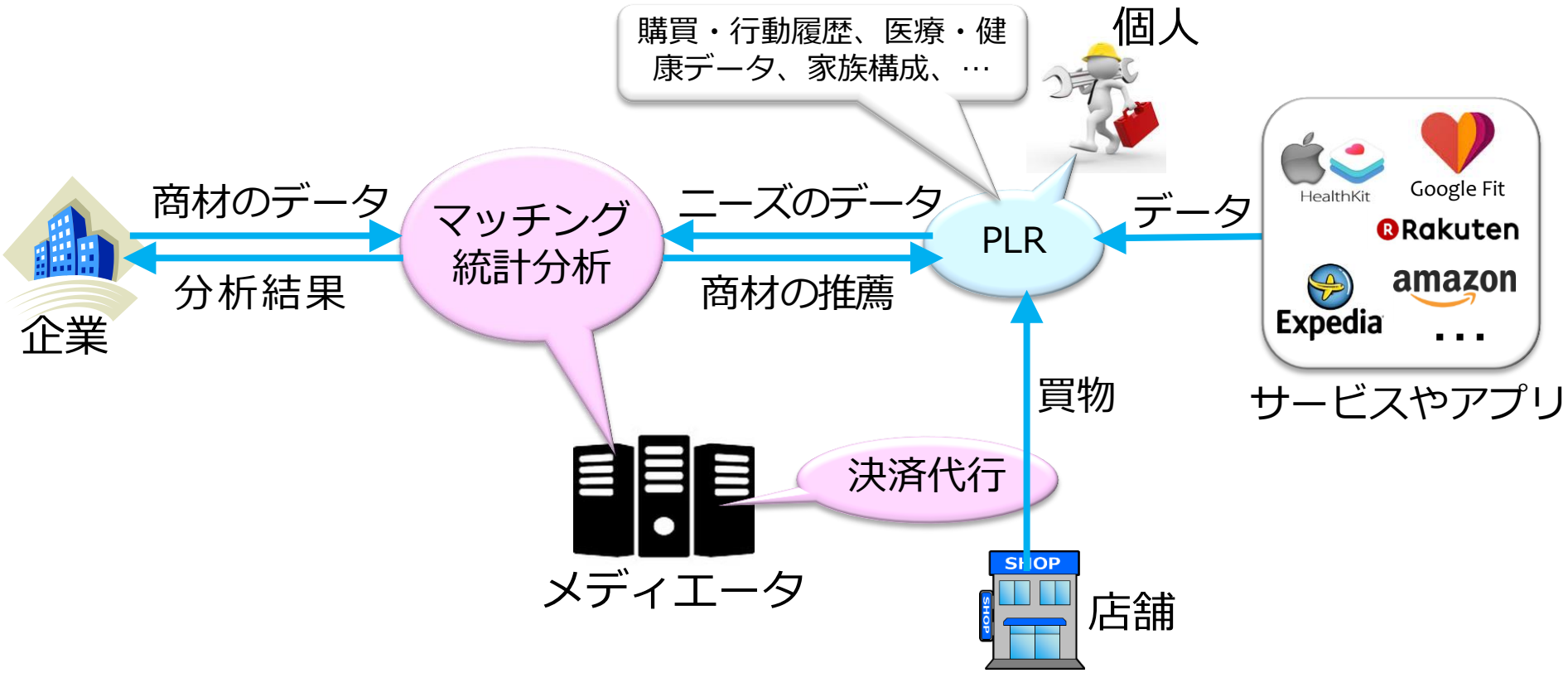
客がレジに並ばず自分のスマホで決済

- PSD2等で銀行がAPIを開示すれば実現は容易
- 買物客のPLR端末
  - POSデータを生成して店舗に渡す
  - 電子手形を決済サーバから取得して店舗に渡す
- 店舗は電子手形を銀行で換金



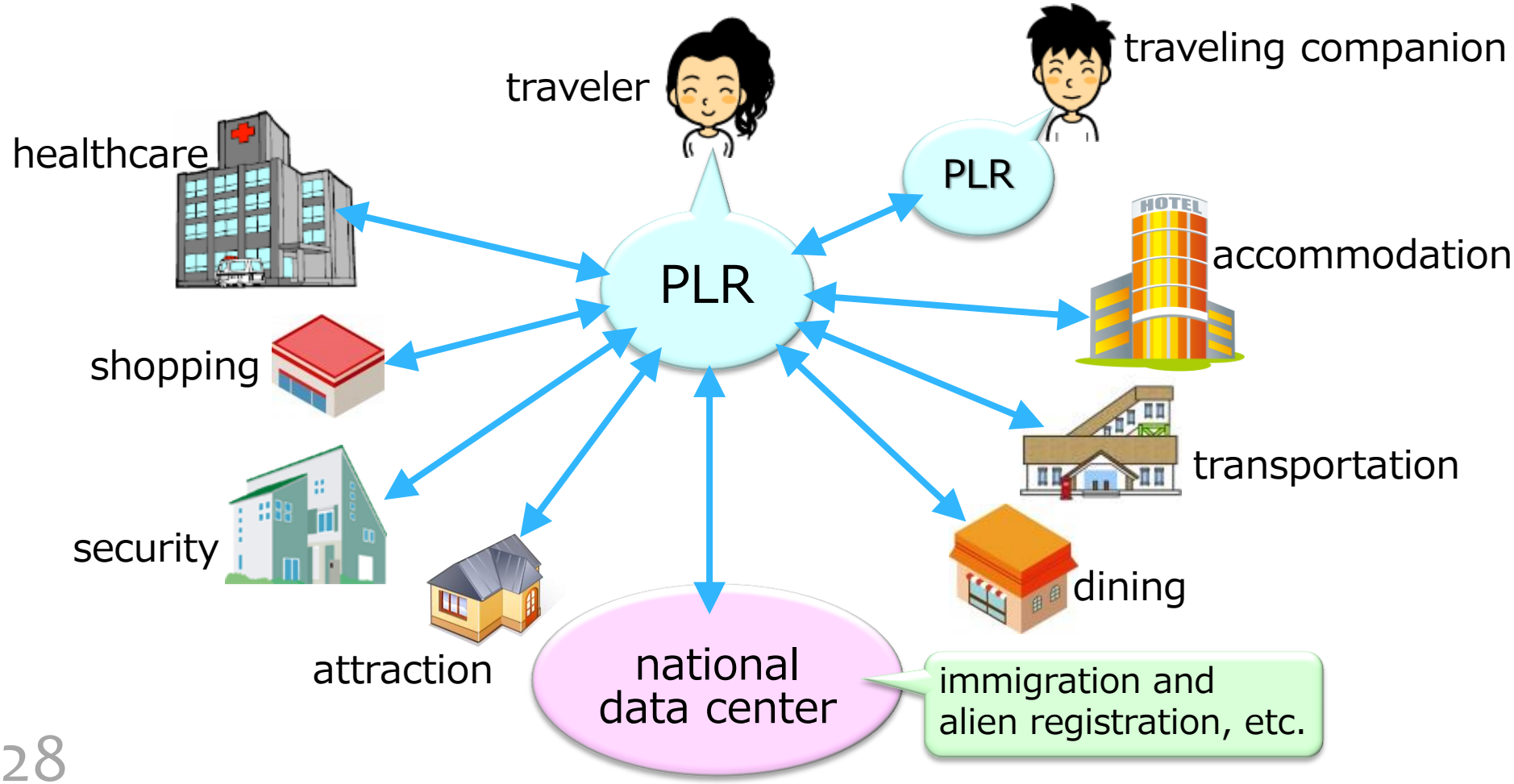
購買支援メディエータ

- 前頁の決済サービスを個人と店舗に無料で提供
- 企業の商材と個人のニーズをマッチング
  - 成約時に仲介手数料を企業から徴収
- ビッグデータの分析結果を企業や自治体に販売



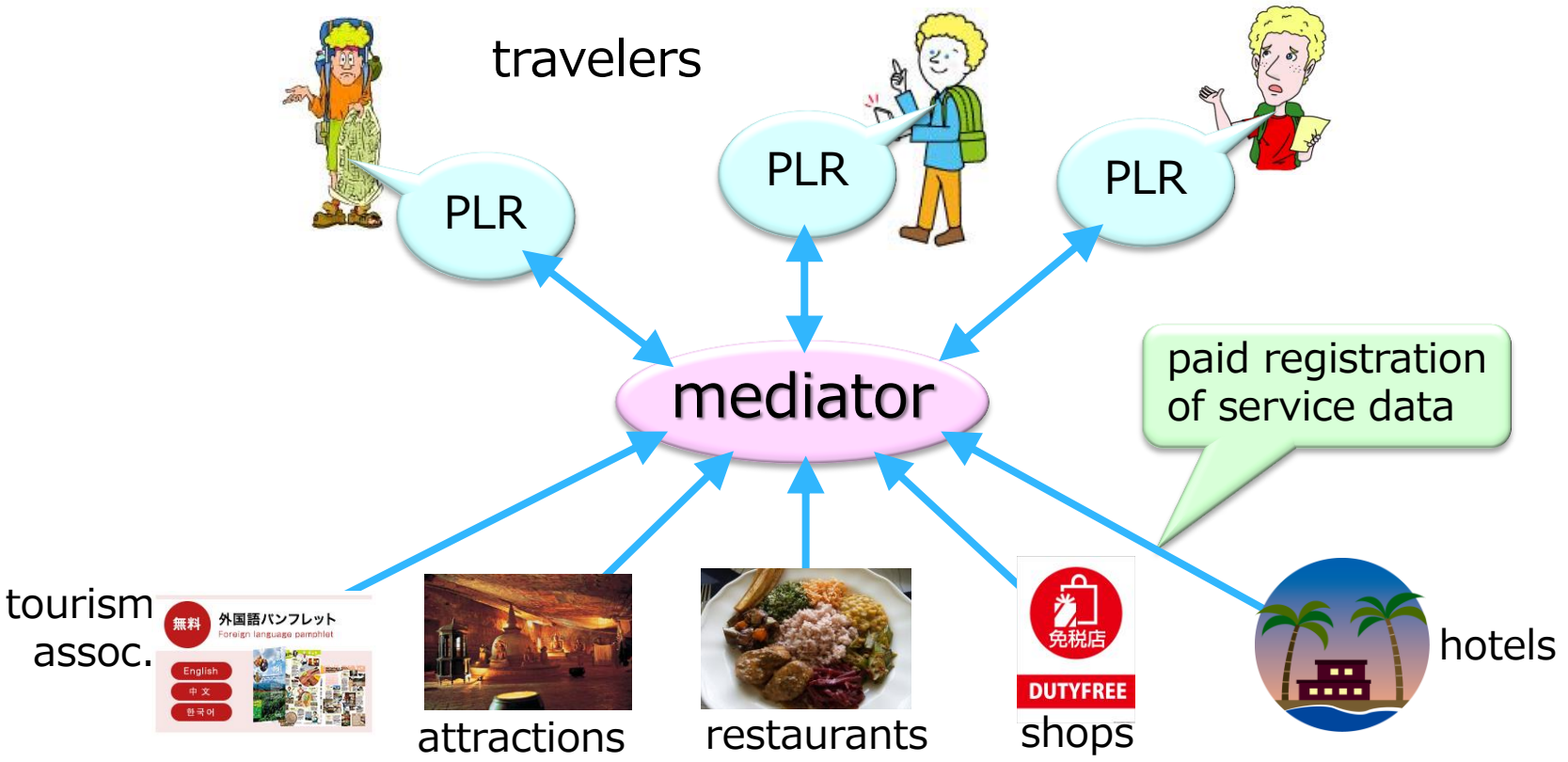
Tourism

Tourist services are federated and optimized per traveler by traveler-mediated real-time data sharing across travelers and tourism vendors.



Tourism Mediation

- Mediators provide info. to travelers.
  - based on
    - personal data from travelers
    - service data from vendors
  - (possibly paid) concierge service
    - human/AI assistance through secure PLR messaging



AIの基盤としてのデータ整備

- AIによる社会の自動化にはデータ整備が必須
  - 研究開発と実用の両方で、良質のデータが容易に取得できる必要あり
    - 研究開発コストのほとんどがデータ整備にかかる
    - サービスの受容者に関する詳しいデータが必要
- データ整備 ＝ BPR (業務改革)
  - 意味構造化データが潤沢に流通する社会の構築
    - 意味構造化 ＝ 機械にも人間にも意味がわかる
- 業務改革には意味の理解が必要だが、AIは意味を理解しないので、AIによるデータ整備は無理
- PDSは意味構造化されたパーソナルデータを本人同意で流通させる
  - それによりマッチングをAIで自動化

AIは意味を理解しない