2014年総会シンポジウム「ビッグデータの可能性と課題 ――監視・シミュレーション・プライバシー」・論文

集めないビッグデータ: 情報の分散管理による個人の尊厳と公共の福祉

Distributed Big Data: Individuals' Dignity and Public Welfare by Decentralized Information Management

キーワード：
個人データ，分散PDS，PLR，VRM
keyword：
personal data, decentralized PDS, PLR, VRM

東京大学大学院情報理工学系研究科　橋　田　浩　一
Graduate School of Information Science and Technology, the University of Tokyo
Kôiti HASIDA

要　約

個人データの集中管理とは，管理者が多数の個人のデータをまとめて管理することであり，これによって多数の個人のデータが一挙に活用されたり漏洩したりし得る。一方，個人データの分散管理とは，個人または代理人が本人分のデータのみを管理することであり，一挙に利用されたり漏洩したりするデータは１人分に限られるので，セキュリティが高い。個人が本人のデータを電子的に蓄積して他者と安全に共有し活用するための仕組みをPDS (個人データ保管庫)と言う。個人はPDSによって自らの利益を高めるように自分のデータを活用することができ，こうしてB2Cサービス全体の社会的価値も向上する。PDSにも集中型のものと分散型のものがあり，分散PDSの方がセキュリティと利便性が高い。分散PDSにもサーバ主導のものと個人端末主導のものがあるが，PLR (個人生活録)は個人端末主導の分散PDSであり，個人端末もサーバ側の仕組みもすでにコモディティになっているものを使うので，導入・運用コストが非常に小さい。マーケティング，電力小売の自由化，マイナンバー，スーパーハイビジョン放送，医療制度改革等に伴う個人データの安全・安価な管理と活用に対するニーズの高まりがPLRのような分散PDSを普及させる契機になるものと考えられる。

Abstract

Centralized management of personal data means that the manager is in charge of many people's personal data, so that this big data can be utilized or leak at one time. On the other hand, decentralized management of personal data means that each individual or her agent manages her own data only, so that the data utilized or leaking at one time is just part of her own data, entailing much higher security. PDS (personal data store) is a mechanism to allow each individual to electronically accumulate and utilize her own data by sharing the data with others. Individuals can use their own data so as to improve their own benefits, hence improving the whole social value of B2C services. PDS is either centralized or decentralized, and decentralized PDSs are much more secure and convenient. Decentralized PDSs are either server oriented or personal-device oriented. PLR (personal life repository) is a personal-device oriented, decentralized PDS, whose introduction and operation costs are very low because it uses commodities for both personal devices and servers. The spread of such decentralized PDSs will be triggered by the increasing needs for secure and low-cost management and utilization of personal data in line with to marketing, liberalization of power retailing, My Number, super high vision broadcasting, reform of medical system, and so on.

1　個人データの管理

データの管理とは，データを保管しつつ，データの利用の可否を決定するということである。管理者の意思または過失によってデータが利用されたり漏洩したりするわけである。

個人データの集中管理（centralized management）とは，管理者が多数の個人のデータをまとめて管理することである。ゆえに集中管理の下では，多数の個人のデータを一挙に利用することができ，また多数の個人のデータが一挙に漏洩することがあり得る。個人データの利用に際してその都度本人の同意を求めなければ，多人数のデータを活用するのが簡単である。したがって集中管理は，本人たちのメリットがあまり明確でないデータの利用に適する。それは典型的には多数の人々のデータの分析であり，たとえばある疾患の治療法を発見するために多数の患者から集めたビッグデータを分析する場合などが考えられる。

逆に，本人のメリットが明確なデータの利用には集中管理は適さない。これは，管理者の利害と本人の利害が一致しないことが多いからである。たとえば，病院が治療のデータを他の病院や診療所と共有すれば患者のメリットになるはずだが，従来はデータを共有しても病院が儲からなかったので，医療データの共有はほとんど進んでいない。

一方，個人データの分散管理（decentralized management）とは，管理者が１人分の個人データのみを管理することだが，その管理者は典型的には本人または代理人や後見人であろう。分散管理の下では一挙に利用されたり漏洩したりし得るデータは１人分のみである。したがって，たとえば個人データが数千万人分あるとすると，集中管理よりも分散管理の方が数千万倍安全である。また，分散管理においては個人データの利用が本人（代理人）の同意に基づくので，分散管理は本人のメリットが明確なデータの利用に適する。たとえば，ある病院での治療の記録を別の病院で開示することによって安全で効果的な治療を受ける場合などが考えられる。逆に，上述したビッグデータの分析のように本人のメリットが不明確なデータ利用のためには分散管理は不適切かも知れない。

このように，個人データの集中管理と分散管理はそれぞれに有用で相補う関係にあり，いずれも必要だが，問題は分散管理が実際にはほとんどなされていないことである。前記の通り，個人データの分散管理は本人のメリットを高めるために有効だが，現在は個人が自分のデータを体系的に管理できていないことが多く，過剰な集中管理と相俟って，それが個人のプライバシを脅かすのみならず，B2Cサービスの価値の向上や市場の拡大を阻害している。前述の医療の例のように，個人が本人のデータを自らの判断に基づいて自ら指定する事業者と簡単に共有できれば，各B2Cサービスの価値が高まり，市場全体のパイが大きくなるはずである。さらに，商品やサービスの利用等に関するデータを個々の消費者が管理し社会的に共有し分析して事業者を比較評価することで競争環境を整え，その産業の国際的な競争的および他産業との競争力を高めることによっても，当該産業のパイが拡大するだろう。

2　PDS

個人が本人のデータを電子的に蓄積・保管して他者と共有し活用できるようにする仕組みをPDS（personal data store，personal digital store，personal data service，personal data vaultなど）（Bell, 2001）と言う。AcxiomやBluekaiなどのいわゆるデータブローカは，（個人データの売買以外に）個人データを本人のために活用する機能を持たないので，本稿ではPDSから除外して考える。

PDSは，後述のような官製のトップダウンな仕組み以外にはまだほとんど普及していない。その原因は，良い収益モデルがまだ確立しておらず，

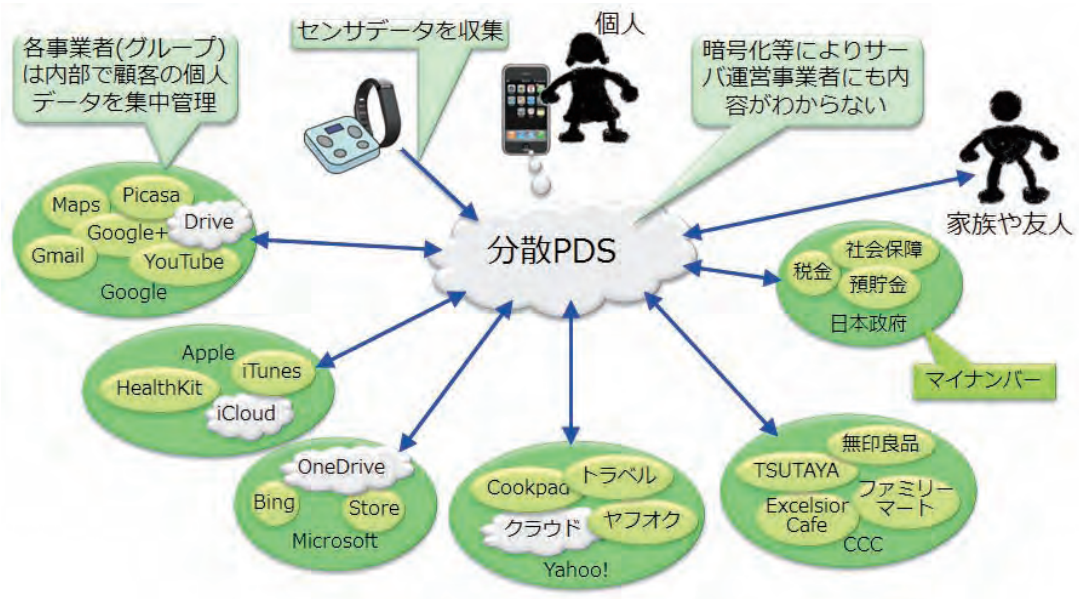
サードパーティとして参画するサービス事業者のメリットが不明確なことなどであろう。以下では，PDSの諸相について本節で論じた後に，サービス事業者にとってのPDSのメリットや日本の政策の動向などPDS（特に後述の分散PDS）を普及させる契機について３節以降で述べる。

2.1　集中PDS

PDSにも個人データを集中管理するものと分散管理するものとがある。集中PDS（centralized PDS）は，個人データを本人の役に立てるだけでなく，多数の個人のデータに匿名化等の処理を施した上で事業者に提供するデータブローカの機能も備える場合が多い。ヘルスケアに関する集中PDSとして，Google HealthやMS HealthVaultやPicnicHealth（PicnicHealth, 2014）や日本の「どこでもMY病院」構想の下で開発されたシステム（厚生労働省, 2013a）など，集中管理に基づくPHR（personal health record; 個人が本人の医療データを管理しヘルスケア事業者と共有して活用する仕組み）が挙げられる。これらはヘルスケアのための民間のPDSだが，デンマークのBorger（Borger, 2014）などは，ヘルスケアに限らない多様な個人データを政府が集中管理して本人の役に立てたり企業等に提供したりすることによって産業や文化の振興を図っている。米国政府が運用するBlue Button（HealthIT.gov, 2014）とGreen Button（Department of Energy, 2014）はそれぞれヘルスケアと電力エネルギーに関する集中PDSと言えるだろう。現在開発が構想されている日本の情報銀行（東京大学, 2013）もデータブローカの機能を備える集中PDSの一種と言えよう。

2.2　分散PDS

分散PDS（decentralized PDSまたはdistributed PDS）は，図－1のように，集中PDSを含む多くの集中管理型サービスを個人が自由に組み合わせて利用することを可能にする。この図は，GoogleやCCCや日本政府がそれぞれID連携等の仕組みを用いたデータの集中管理に基づくサービスを行ない，個人がそうした複数のサービスを利用している様子を表わす。たとえば，YouTube



図－1　多数の集中管理型サービスを個人利用者が分散PDSで相互連携させる

の利用履歴とiTunesの利用履歴を統合して分析したいとか，日本政府がマイナンバーで管理する預貯金のデータとCCCがT-IDで管理する購買のデータを統合して分析したいとか言っても，GoogleとAppleが顧客のデータを共有するとか，日本政府とCCCが個人データを融通し合うとかいうことはあり得ないし，あってはなるまい。YouTubeとiTunesの利用履歴を統合するとか預貯金のデータと購買のデータを統合するとかいうことは，個人が本人の意思に基づいて分散PDSで行なうしかない。さらに，この統合を多数の個人のデータに拡張するにはその人々から同意を取得する必要があるが，それも分散PDSによって容易になる。

だが，各集中型サービスがより多くの個別サービスを相互連携させ，またその顧客が増えると，分散PDSによって統合すべき集中型サービスが少なくなるので，集中型サービスの間での統合は分散PDSの運用にとって好都合である。たとえば，EHR（electronic health record; 病院等が運営するサーバによって複数の医療機関の間で患者のデータを共有する仕組み）によって多くの医療機関がつながっていれば，個人が分散PDSによってつながるべき医療機関（またはEHRシステム）が少ないので，ヘルスケアデータの社会的共有や地域医療連携が進みやすい。

多数のサービスを相互連携させる機能の一環として，分散PDSはそれらのサービスに関するシングルサインオンの機能を備え，それによってパスワードの使い回しを防ぐなど，個人レベルでのセキュリティを向上させる。一般にデータは分散させた方がセキュリティが高いが，１人分のデータをあまり多数に分散させると，管理者である本人や代理人の認知限界を越えてしまい，アカウントを忘れたりパスワードを使い回したりするので却って危ない。１人分のデータはひとまとめにして管理するのが実際には最も簡単で安全と思われる。データの開示は一般には１人分の全データを対象とせず，各々の場合の目的に応じたごく一部のデータに限ることは言うまでもない。

分散PDSには，個人用端末同士のP2P通信によって利用者間でのデータ共有を実現するものと，個人用端末以外にサーバ（ホームサーバや事業者が運営するサーバ）を用いてサーバ同士またはサーバと個人用端末との通信によりデータを共有するものとがある。

前者のP2P型分散PDSとしてはPersonal Server（Want, 2002）などが提案されている。しかし，その後はいまのところ実験的にでも稼働しているものはなさそうである。スマートフォン等の個人用端末によるP2P型の分散PDSの実装は，端末の通信量や電力消費量に関する制約により当面は難しいだろう。

データ共有にサーバを用いる分散PDSは，サーバと個人端末との役割分担の観点から分類することができる。サーバが主要な役割を果たす方式の分散PDSとしては，Persona（Baden, 2009），VIS（Cáceres, 2009），PDV（Mun, et al., 2010），PrPl（Seon, et al., 2010），openPDS（de Montjoye, et al., 2014），Respect Network（RespectNetwork, 2014）などがある。これらのうちPersonaは個人ごとのデータの暗号化によって，VISとPDVとPrPlとopenPDSは各個人専用の仮想計算機等を用いて分散管理を実現する。Respect Networkにおける個人データの分散管理は，利用者との（個人データの利用が本人同意を必要とする旨の）契約に基づく運用によるものと思われる。

一方，個人端末が主要な役割を果たす分散PDSとしてPLR（personal life repository; 個人生活録）（橋田, 2013; Hasida, 2014）がある。PLRでは，すでにコモディティになっているGoogle DriveやDropbox等のクラウドストレージサービスをそのままデータ共有用のサーバとして使い，データ共有以外（データの視覚化や暗号化やクラウド間連携など）の情報処理をすべてスマート

フォン等の個人端末が担う。PLRは，非公開の個人データを暗号化してからクラウドストレージに送信し，クラウドストレージからデータを取得した後に手もとで復号する。その復号に必要な鍵を，クラウドストレージ事業者にもPLRを開発する事業者にも原則として開示せず，利用者が自ら指定した他者にのみ開示することにより，個人データの分散管理を実現する。

まだコモディティになっていない類のクラウドストレージやID連携などのサーバ機能を（P2P型以外の）他の分散PDSが必要とするのに対し，PLR本体はスマートフォン等のアプリに過ぎずサーバとしては既存のコモディティをそのまま活用する。この意味において，PLRはきわめて簡便で運用コストの低い分散PDSと言えよう。Respect Networkのようなサービスがコモディティとして安価に使えるようになれば，個人端末の機能の多くをそれらのサービスに委ねることにより，PLR本体の実装をさらに簡素化し保守を容易にできるだろう。つまり，中長期的にはPLRと他の分散PDSとが融合する可能性がある。もうひとつの可能性は，個人端末の性能向上によってP2P型の分散PDSが普及することであろう。PLRは両方の可能性を視野に入れて設計されている。さらには，PLRと情報銀行とが連携して，PLRが個人データの分散管理，情報銀行がそのデータの収集とビッグデータとしての活用を担う可能性も考えられる。

3　個人情報漏洩リスクの管理

民間のPDSがあまり普及していない主な原因のひとつは，既存のサービス事業者がPDSと連携するメリットを理解していないことである。以下ではまず，個人情報漏洩のリスクを管理するために分散PDSが有効であることを指摘する。

2014年７月９日以来の報道によれば，ベネッセホールディングズが保有する多数の顧客の個人情報が漏洩し，名簿業者に売却され，ジャストシステム等に渡った。そのデータは約4,800万人分に上り，その内容は住所・氏名・電話番号・家族構成を含むらしい。米国のTargetやHome Depotでも数千万〜数億件のクレジットカード等の情報が流出するなど，個人情報漏洩事件は枚挙に暇がない。

ベネッセでは社内システムを24時間監視しており，全社員や委託先の従業員に個人情報の取り扱いに関する教育をし，定期的な外部監査も受け，個人情報を適切に管理しているという旨の「プライバシーマーク」も取得していた。社内で顧客の情報にアクセスできるのは透明のガラスで仕切られた小部屋にある専用端末のみであり，その部屋に入るには，事前の予約とセキュリティーカードによるチェックが義務づけられ，端末から顧客情報を取り出す際には，パスワードの入力が必要だという。

このようにベネッセの管理はかなりしっかりしていたと言えるが，今回の犯人はグループ会社の下請けのSEであり，スマートフォンの持ち込みが禁止されていなかったという盲点を突いてスマートフォンでデータを持ち出したとのことである。ではさらに管理を徹底してスマートフォンの持ち込みを禁止しておけば良かったではないかと言う向きもあろうが，徹底の度合いには際限がなく，徹底すればするほどコストがかかる。スマートフォンの持ち込みを禁止していれば良かったなどと言うのは結果論であり，そのような管理の盲点を予めすべてつぶしておくことなど不可能であろう。

つまり，このような個人情報漏洩のリスクは多くの事業者にとって人ごとではない。以下ではまず，分散PDSによって個人情報の漏洩がいかにして防げるかについて考えてみよう。

3.1　暗号化

PersonaやPLR等の分散PDSは，利用者の個人

データを暗号化した状態で端末やクラウドに保存する。多くのユースケースにおいては，復号されたデータはプログラムの実行時のメモリ空間に一時的に存在するのみであり，他のプログラムから直接アクセスできる形で保存されることはない。

このような暗号化の運用によって管理が行き届きやすくなることが期待できる。つまり，PLRのような分散PDSを業務に用いる場合，ベネッセのように情報システムの管理を部分的に外注する際は，非公開のデータを復号する必要のない作業に外注の範囲を限定し，データの復号が必要なメンテナンス作業を内製化することにより，個人情報漏洩等のリスクを低減できるだろう。

しかし，業態によっては，システム管理作業においてデータを復号せざるを得ないことが多く，それをすべて内製化できない場合もあるだろう。そもそも，個人データを集中管理している限り，そのデータが一挙に漏洩する恐れは常にある。

3.2　アドホックなデータ収集

データを集めてしまうと必然的に集中管理が発生する。しかし，集中管理の常態化を分散管理で防ぐことによってリスクを低減できるだろう。

つまり，いきなり大量のデータを集めてずっと抱え込んでおくのではなく，必要に応じてアドホックにデータを集め，分析が終わったらその結果だけ残して元データを破棄してしまえば，漏洩のリスクが激減する。現状では，何のためにどう使うのかよくわからないような大量の個人データを不用意に集め，さらに用済みのデータまで未練がましく保管して管理コストと漏洩リスクを無闇に高め，リソースを浪費し経営を危険に晒している事業者があまりにも多い。

分散PDSの多くの利用者とつながっていれば，いきなり大量の個人データを集めるのではなく，分析の目的と方法と効能が十分明らかになった後に分散PDS利用者からデータを本格的に集めることができる。その気になればいつでも再びデータを集められるから，分析し終わったデータはすぐに消去して構わない。大量のデータを保管してわざわざ漏洩のリスクに晒す必要はない。もちろん，短期間のうちに繰り返し使うようなデータは破棄せずに保存しておいた方が効率が良いこともあろう。しかし，保存が長期に及ぶと，漏洩のリスクが高まるだけでなく，保存中のデータが多岐にわたり，データ同士の照合禁止等々に関するコンプライアンスの管理が煩雑になり，それもリスクの元になる。

従来，各事業者は自ら提供するサービスに直結するデータは普通に取得できた。たとえば，クレジット機能付の会員カードを顧客に使ってもらえば，顧客別の購買データが取れる。同様のデータを分散PDSで取得・保管することは易しい。クレジット機能付会員カード等によって自分のID付きの購買データを事業者が取得することに同意していた顧客であれば，代わりに分散PDSを使って自分のID付きの購買データを提供することに抵抗はなかろう。

さらには，顧客が他の事業者からの購買のデータや日常生活行動のデータや医療記録やバイタルデータを分散PDSで蓄積していれば，事業者はそれも顧客本人の同意の下で見せてもらうことができる。通常のクレジットカードと違ってオンラインでつながっているので，顧客にポイントを付与するだけでなくクーポンを配ったりオンラインで広告を配信したりキャンペーンを打ったりできることは言うまでもない。しかも，そのために購買記録等の個人データをすべて常に保管しておかなくても済む。個人データは必要に応じてアドホックに顧客から集めれば良い。

だが，データを集めることができるということは，集めたデータが漏れる恐れがあるということである。データ利用の利便性と漏洩のリスクは同じコインの表裏である。したがって，氏名や住所や電話番号や医療記録など機微性の高いデータの収集・利用を面倒くさくすることによってその漏

洩のリスクを低減させるしかない。

そのためには機微なデータを集める際の認証を厳しくすれば良い。分散PDSを使えばそれは簡単である。各個人にとって機微性の高いデータを開示する際には本人の分散PDSが通常より厳しい認証を要求するわけである。個人の分散PDSの設定を変更する権限は原則として本人だけにあるので，事業者がその設定を変更して機微性の高いデータを大量に集めるのはほぼ不可能である。このように，分散PDSによって個人データの主たる管理者を事業者ではなく本人とすることにより，事業者からのデータ漏洩をかなりの程度まで防止できるだろう。個人が本人のデータだけを管理するという意味での分散管理は，一度に漏洩するデータを1人分に限るゆえに集中管理より圧倒的に安全であるが，事業者によるデータ収集を統制しやすいという意味においても安全性が高い。

3.3 VRM

また，個人データを本人が分散PDSで管理していれば，事業者は個人データに直接触れずに結構いろいろなことが今までよりも効果的にできる。

たとえば，イベントや新商品のお知らせを個人に届けるためにダイレクトメールや電子メール等が使われてきたが，その際に各個人に合ったお知らせを送るには当該個人に関する情報が必要である。ベネッセの事件では，漏洩したデータが家族構成等の情報を含んでおり，名簿屋からデータを購入したジャストシステム等はそれを用いてダイレクトメールの送付先を選んだわけである。

しかし，事業者が個人にダイレクトメール等を送り付ける代わりに，個人利用者が用いる何らかのアプリ（これを「VRMアプリ」と呼ぼう）が利用者の分散PDSに蓄積された個人データを参照することにより事業者の広報サイトから利用者に適合した情報を抽出して利用者に通知することも可能である。たとえば，ある事業者がコンサートの開催予定を広報しているとき，そのコンサートの演目やアーティストを好む利用者だけが自分の分散PDS からコンサートに関する通知を受け取る，といった具合である。このように個人が自分に合った商品やサービスの情報を自分のアプリで取ってくれれば，事業者はダイレクトメール等の送り先を選ぶための個人情報を必要としない。ベネッセのような事業者も，顧客の住所や家族構成等の情報を持たずに各顧客に合った情報を提供できる。

CRM（Customer Relationship Management; 顧客関係管理）とは事業者が顧客への売り方を最適化するということだが，VRM（vendor relationship management; 業者関係管理）(Project VRM, 2014; Searls, 2012）は逆に顧客が事業者からの買い方を最適化するということであり，上記のVRMアプリはその基本的な機能を持つ。個人のデータを本人が蓄積･管理していれば，個人はいかなる事業者よりも本人のデータを多く持つから，自分に適した商品やサービスを事業者よりもずっと正確に同定できる。B2C事業者は個人情報管理を含むCRMのコストとリスクから解放される。

当然ながら，VRMが普及すれば事業者が顧客を囲い込むのが難しくなる。しかし事業者は，先行してVRMに対応することにより，CRMのコストとリスクを低減させつつVRMに即した事業を早期に開始し，市場における競争優位性を確保できるだろう。

3.4 さらなる分散

一方，ビッグデータを1か所に集約せず分散させたままリアルタイムで分析や学習を行なう技術も進歩しつつある。上記のVRMに加えてこのような技術を使えば，そもそも個人データを集める必要が実際にはあまりなくなるかも知れない。

しかしながら，管理を個人（だけ）に任せられない個人データもある。たとえば，普通の納税者はなるべく税金を払いたくないので，課税の根拠

となる個人の収入や資産のデータは国や自治体が集中管理すべきだろう。それがマイナンバーを導入する目的のひとつである。また，個人と事業者との間の契約書も，個人にとって不利な内容を含む場合（つまりほとんどの場合）には事業者が保管する必要がある。

だがそのような場合でも，通常はなるべく集中管理されているデータを使わずに個人ごとに安価に分散管理されているデータを使うことにより，セキュリティを格段に高めることができるだろう。個人は課税の根拠や契約書を改竄したくなるかも知れないが，政府や企業が正しいデータを持っているのでそれは無効である。契約書等の場合は事業者の電子署名によって真正性を保証することもできる。

4　自律分散協調ヘルスケア

病院や診療所などの医療機関が医療データを共有することが医療の価値を高めるためには望ましい。各患者に対してどのような治療が安全で効果的であるかは当該患者に対するこれまでの診断や治療の内容に依存するから，複数の医療者が１人の患者の治療に当たるには，その患者に関するデータを共有すべきである。

そのことはほとんどの医療関係者が認識しているはずであるが，これまでデータ共有はあまり進んでいない。データを共有しても病院や診療所の収益に直結しないからである。

4.1　医療制度改革

しかし，現在進行中の医療制度改革（厚生労働省, 2014）によってデータ共有が必須になる。厚生労働省は2025年までに新たな医療提供体制を確立することを目指して着々と制度の改革を進めている。それに伴って医療データの共有が医療機関の経営の観点からも必須になりつつある。

この医療制度改革において特に重要なのは下記の２点だろう。

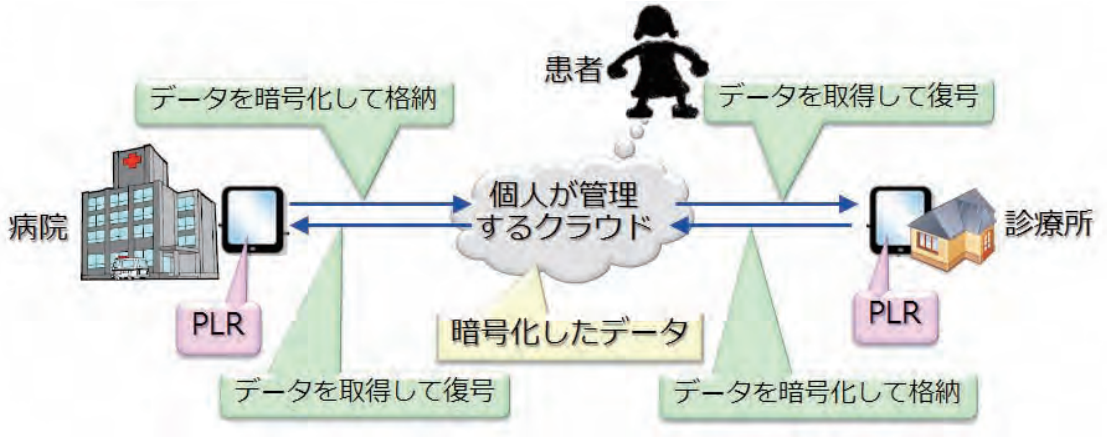
- 病院の間の役割分担（厚生労働省, 2013b）
- 在宅医療の推進（厚生労働省, 2012）

病院は，その機能に関して，急性期，回復期，療養期などに分類され，各々の段階の入院患者のケアに特化しつつあり，2018年にはこの分類が完了する予定である。たとえば急性期病院への保険点数の付与は急性期の入院医療と紹介による外来診療に重点化される。このように，各種の医療機関が特有の機能に専門化することによって医療サービスの質が高まり，異種の医療機関の間での連携が強化されるだろう。

異種の医療機関や介護・看護事業者の間の連携を強化して体系的・継続的なヘルスケアを提供するには，それら関係者の間でのデータの共有が必要だろう。在宅医療に関しても，複数の診療所（訪問医）の間で患者のデータを共有する必要が生ずる。新たな診療報酬制度の下では，多くの患者について24時間365日の対応が訪問医に求められるからである。診療所のほとんどは医師が１人で看護師が２～３人の体制だが，それではとてもそんな対応は無理なので，複数の診療所がグループを組む必要がある。グループが機能するには，各患者のデータをグループ内で共有して医師や看護師が外出中にも参照できるようにせねばなるまい。もちろんそのデータ共有は病院や介護・看護事業者にも及ぶ必要がある。

4.2　集中管理から分散管理へ

医療や介護のデータを各患者や被介護者本人（または代理人）に分散して管理すれば，集中管理方式の場合よりも圧倒的に安全にかつ小さな費用でデータの共有が実現できる。つまり，PLRの仕組みを部分的に用いて，図－２のように，個人が本人のデータをGoogle DriveやDropbox等の基本無料のクラウドストレージに格納して家族やヘルスケア事業者と共有すれば良い。非公開のデータは暗号化してからクラウドに格納しクラウ



図－２　患者個人をハブとして医療機関同士が患者のデータを共有する

ドから取得の後に復号することにより，Google社やDropbox社にも内容がわからないようにデータを運用できる。

図－2においては，各患者は（健康な人も）まずGoogle Drive等にアカウントを作り，そこに所定の形のフォルダを作成し，それを自分がかかる可能性の高い病院や診療所や老人ホーム等のヘルスケア事業者（のPLRまたはAPIを設けた医療情報システム等）と共有するだけである。その共有の際にこのフォルダ用の暗号鍵を事業者に渡す必要があり，それにはPLRが必要だが，各患者が自らスマートフォン等によってPLRを利用する必要はなく，他人のPLRアプリでこの作業を代行することができる。つまり，そのような相互扶助により患者にとっての基本コストはほぼゼロになる。自らPLRを利用しなければ，ヘルスケア事業者の間での自分のデータの共有を仲介するだけだが，スマートフォン等を持っている患者は自らPLRを使うことによって自分のデータを閲覧したり分析したりすることができる。

一方，各ヘルスケア事業者は（病院の場合は診療科ごとでも可）タブレットPC等を導入してPLRとアプリをインストールし，既存の電子カルテシステム等があればそれとつないで（または既存のシステムにAPIを付加することにより）データ連携する必要があるが，それにかかる事業者（病院の場合は診療科）のコストはきわめて小さい。まず，タブレットPCは３万円ほどで十分な性能のものが入手可能である。また，各電子カルテシステム等は数百の医療機関に導入されているのが普通であり，ひとつのシステムをPLRと連携させるための改修にかかる費用はせいぜい200万円程度だろうから，各機関あたり数千円程度で済む。これらに間接経費等を加えても，各事業者が負担する導入コストはEHRの場合の数百万円より圧倒的に安い。運用コストもPLRの場合はタブレットPCの減価償却費程度であろう。電子カルテシステム等にAPIを付加する場合も同様である。

分散PDSによってこのように低コストのデータ共有が実現できるだけでなく，既述の通り，セキュリティも大幅に高まる。さらに，集中管理方式だと医療機関の間でしかデータが共有できないのに対して，個人による分散管理方式は，患者が自分の端末で任意の医療者等にデータを開示できるという意味で利便性も高い。

一方，ヘルスケア事業者が集中管理型のデータ共有システム（EHR）を用いるメリットとして，たとえば病院が診療所や老人ホームや患者を囲い込めると考える向きもあるかも知れない。しかし実際にはそのような囲い込みは不可能である。囲

い込みが可能なのは他の選択肢を選ぶのに伴うコストが大きい場合だが，上記のように分散管理は圧倒的に安価であり，しかも安全性においても利便性において集中管理を凌駕するからである。

5　展望

以上のように，個人情報漏洩のリスク管理や医療制度改革が分散PDSを普及させるきっかけになる可能性が高い。また，分散PDSを普及させるきっかけは他にもある。

総務省がSHV（スーパーハイビジョン）放送の普及を推進しているが，地上波の帯域ではSHVに対応できないので，SHVのコンテンツは光ケーブルで配信することになるだろう。地上波放送の帯域が通信用に割り当てられ，現在の地上波放送がインターネットに移行すれば，独居老人宅なども含む全国のほぼ全世帯のテレビがインターネットにブロードバンドで接続される。それにより，ほとんどどの家にもあるテレビが医療・介護や購買支援など生活サービス用の端末になると考えられる。

テレビを通じた総合生活サービスにおいて個人データを安全かつ安価に管理・活用するためにはPLR等の分散PDSが必要と考えられる。それに関連して，マイナンバーに基づく本人認証を分散PDSのアプリとすることでマイナンバーと分散PDSを連携させ，図－1 のように数多あるサービスIDのひとつとしてマイナンバーを位置付けることにより，マイナンバーの運用に伴うプライバシに関する懸念を払拭し，マイナンバーの普及を促すことにもなるだろう。

さらに，自由化後の電力小売市場において適正な競争を促すには，各需要家が自分のエネルギー消費に関するデータを持ち，それに基づいて電力小売事業者を正しく選択できるようにする必要がある。そのために米国のGreen Buttonのような集中PDSを構築するよりも，分散PDSを用いた方がはるかに社会的コストが安く，かつデータ流通の自由度が高くなるだろう。

個人情報漏洩の防止，ヘルスケア，SHV，マイナンバー，エネルギー管理等に関する分散PDSの普及が相互に促進し合うことは言うまでもない。これらの領域のうちいずれか 1 つにおいて分散PDSが普及すれば，ドミノ倒し的に他の領域でも分散PDSが広まることになるだろう。電力小売の自由化とマイナンバーの利用開始は2016年，SHVの普及は2020年の東京オリンピックまで，医療制度改革は2025年までの予定だから，このドミノ倒しは2020年までに生ずる可能性が高い。東京大学は，2014年10月に「集めないビッグデータ」コンソーシアムを設立し，その可能性の具現化を目指している。

注

本稿は『人工知能学会誌』2014年11月号に掲載した解説を修正したものである。

参考文献

Randy Baden, Adam Bender, Neil Spring, Bobby Bhattacharjee, and Daniel Starin (2009) Persona: an online social network with user-defined privacy. *ACM SIGCOMM Computer Communication Review*. Vol. 39, pp.135-146.

Gordon Bell (2001) A Personal Digital Store. *Communications of the ACM*, 44: 86-91.

Borger (2014) <https://www.borger.dk/> Accessed 2014-09-30.

Ramon Cáceres, Landon Cox, Harold Lim, Amre Shakimov, and Alexander Varshavsky (2009) Virtual individual servers as privacy preserving proxies for mobile devices. *Proceedings of the 1st ACM workshop on Networking, systems, and applications for mobile handhelds*, pp.37-42.

Yves-Alexandre de Montjoye, Erez Shmueli, Samuel S. Wang, and Alex Sandy Pentland (2014) openPDS: Protecting the Privacy of Metadata through SafeAnswers. *PROS ONE*, 9(7): e98790 doi:10.1371/journal.pone.0098790.

Department of Energy (2014) Green Button. <http://energy.gov/data/green-button> Accessed 2014-09-30.

橋田 浩一 (2013)「分散PDSによる個人データの自己管理」.『人工知能学会誌』, 28(6), 872-878.

Kôiti Hasida (2014) Personal Life Repository as a Distributed PDS and Its Dissemination Strategy for Healthcare Services. *Big Data Becomes Personal: Knowledge into Meaning*, 2014 AAAI Spring Symposim Series.

HealthIT.gov (2014) Your Health Record. <http://www.healthit.gov/patients-families/your-health-records> 2014-09-30.

東京大学 (2013) 情報銀行. <https://ibank.csis.u-tokyo.ac.jp/ibank/index> Accessed 2014-09-30.

厚生労働省 (2012) 在宅医療・介護の推進について. <http://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryou/iryou/zaitaku/dl/zaitakuiryou_all.pdf> Accessed 2014-09-30.

厚生労働省 (2013a) シームレスな健康情報活用基盤実証事業（国庫債務負担行為に係るもの）平成24年度事業成果報告書. (2013) <http://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryou/iryou/johoka/johokatsuyou/dl/houkokusho08.pdf> Accessed 2014-09-30.

厚生労働省 (2013b) 病床機能報告制度及び地域医療ビジョンについて. <http://www.mhlw.go.jp/file/05-Shingikai-12601000-Seisakutoukatsukan-Sanjikanshitsu_Shakaihoshoutantou/0000023379.pdf> 2014-09-30.

厚生労働省 (2014) 平成26年度の診療報酬改定の概要. <http://www.mhlw.go.jp/file/06-Seisakujouhou-12400000-Hokenkyoku/0000039891.pdf> Accessed 2014-09-30.

Min Mun, Shuai Hao, Nilesh Mishra, Katie Shilton, Jeff Burke, Deborah Estrin, Mark Hansen, and Ramesh Govindan (2010) Personal data vaults: a locus of control for personal data streams. *Proceedings of the 6th International Conference*. ACM.

PicnicHealth (2014) <https://picnichealth.com/> Accessed 2014-09-29.

Respect Network (2014) <https://www.respectnetwork.com/> Accessed 2014-09-29.

Doc Searls (2012) *The Intention Economy*: *When Customers Take Charge*. Harvard Business Review Press. (邦訳 栗原 潔 (2013)『インテンション・エコノミー』. 翔泳社)

Seok-Won Seong, Jiwon Seo, Matthew Nasielski, Debangsu Sengupta, Sudheendra Hangal, Seng Keat Teh, Ruven Chu, Ben Dodson, and Monica S. Lam (2010) PrPl: A Decentralized Social Networking Infrastructure. *ACM Workshop on Mobile Cloud Computing and Services: Social Networks and Beyond*.

Project VRM (2014) <http://cyber.law.harvard.edu/projectvrm/Main_Page> Accessed 2014-09-29.

Roy Want, Trevor Pering, Gunner Danneels, Muthu Kumar, Murali Sundar, and John Light (2002) The personal server: Changing the way we think about ubiquitous computing. *Ubicomp* 2002, 223-230.