

分散PDSによる個人データの自己管理

— ビッグパーソナルデータの地平 —

Self-Management of Personal Data by Distributed PDS

東京大学大学院情報理工学系研究科ソーシャルICT 研究センター教授 **橋田 浩一**

PROFILE: 1986～2001年電子技術総合研究所、2001～2013年産業技術総合研究所、2013年7月より現職

1 声の網

星新一（1927～1997）は「ショートショート的神様」と呼ばれているが、長編小説もいくつか手がけており、その中に1970年に発表した「声の網」という作品がある。ネットワークで相互接続された多数のコンピュータが電話で人間からデータを収集し、そのデータを駆使して人間に気付かれないように社会を支配する、というディストピア小説である。

その中に登場する「情報銀行」と呼ばれるサービス事業者は、顧客の個人データをコンピュータの中に保管し、顧客はそのデータを必要なときに引き出して利用することができる。情報銀行は、息子に誕生日のプレゼントを買うべきことをリマインドしてくれたり、個人情報を分析して毎日の食事のメニューを顧客の好みや体調に合わせて選んでくれたりもする。

このような個人情報の管理がコンピュータによる支配の要である。電話によるデータ収集は、通話の内容だけでなく、脳波などのバイタルデータにも及ぶ。コンピュータが人間の声色をまねて電話をかけ、人の弱みに漬け込んで電話のシステムを秘かに改造させることにより、電話器で盗聴したり、受話器から人間に薬剤を投与したりすることも可能である。警察のコンピュータと電話システムが連携して反乱分子の身元を洗い出し、社会的に抹殺したり記憶を奪ったりすることもできるので、支配は盤石である。

ここで描かれている情報銀行は、上記の通り、元来は個人データを蓄積・保管することによって当該個人の

生活や業務を支援してくれるはずのものである。しかし、個人データの共有に関する設計に不備があり、かつコンピュータが人間の予想を越える能力を持っていたために、コンピュータネットワークが影のビッグブラザーになってしまう。情報銀行はもちろん顧客の個人情報を無闇に漏らしてはいけないことになっているのだが、おそらく情報システムの設計にバグがあり、人間の想定を越えて優秀なコンピュータたちがその不備を突いて個人データを人間に断りなく共有し悪用してしまったわけである。

しかし幸いにと言うべきか、あと数十年ほどの間、コンピュータがそれほど有能になる心配はなからう。そうすると、情報システムおよび社会制度を入念に設計することにより、「声の網」のような問題が生じないような情報銀行が実現できるのではないか。実際、最近日本で提唱されている情報銀行[iBank]は、名称は「声の網」のそれに因むものではないとのことだが、コンセプトはよく似ており、情報システムと制度の適正な設計によって安全かつ公正な個人データの活用を目指している。

2 PDS

そのコンセプトは2000年ごろから提唱されているらしく[Mitchell 01]、一般にはPDS (personal data store) と呼ばれる (storeの代わりにstorage、service、bank、vault、locker等が使われることもある)。PDSとは、個人が本人のデータを蓄積・管理し、他者と限定的に共有して活用することを可能にする仕組み

みである。個人のデータは本人が管理して自由に活用すべきだというのは、個人の生活における多様な気付きを促すと考えられる、自然かつ健全な常識であり、PDSはそれを実現する手段である。「声の網」では、情報銀行が顧客のデータを他の事業者の開示することが制度設計において想定されていないように見えるが、現代の情報銀行を含むPDSにおいては、個人データを本人が蓄積して自分のために使うだけでなく、そのデータを社会的に共有することによりさらに広く活用することを重視している。複数の種類のPDSがあってももちろん構わないが、それら間でのデータ連携が自動化されること、および共通のアプリが使えることが重要である。

事業者による個人データの囲い込みを防ぎ、プライバシーを担保しつつ個人データを社会的に流通させビッグデータとして活用できるようにするには、PDSが必要と考えられる。以下に述べるように、それによって個人データにまつわる諸問題が解消し、いわゆるビッグブラザーが無効になり、個人にさまざまな気付きがもたらされ、サービスの内容や評価に関する情報が社会的に共有されて、社会全体にわたり価値が増大する。

最近では、購買や移動などの行動の履歴や嗜好や家族構成や職業や性別や年齢などの個人データを事業者が分析してマーケティング等に利用することが広く行なわれるようになってきているが、図1に示すように、この状況は多くの問題を孕む。

そもそも第一の問題として、自分のどのようなデータが収集され何に使われるのかを個人が明確に認識し同意していない場合がほとんどである。さらに、個人は自分のデータが手もとにないのでそれを活用できず、気付き

の機会も失っている。たとえば、ある病院での医療記録が患者の手もとにないので、そのデータを別の病院で開示して継続的・体系的な治療を受けることができない。あるいは、Amazonのサーバに蓄積されている自分の購買履歴をショッピングモールで開示して適切なサービスや商品を推薦してもらうこともできない。また、リスクの高い疾病等についてはコストパフォーマンスの比較評価に基づいて病院を選びたいのが人情であり、どれほど重度の何という病気での程度の時間をかけてどんな治療を受けていくら払って結果がどうなったかというデータを患者が所有していれば、そのデータを持ち寄って客観的な比較評価ができるはずだが、患者がデータを持っていない現状ではそんなことはできない。

一方、事業者にとっても、自らが提供するサービス等のデータは収集できるが、それ以外のデータを取得するのは困難である。前記の繰返しになるが、各病院は自院内での治療のデータを蓄積・活用できるが、他院のデータや患者の日常生活のデータは取得できない。Amazonは顧客が近所の商店でどんな買い物をしているのか知る由もない。Googleは利用者の情報検索や位置の履歴を取得できるが電車やバスの乗降履歴を知り得ず、JRはその逆である。

事業者による個人データの管理によって生ずる最大の問題は、個人およびそのデータが特定の事業者にロックインされ、また市場の参入障壁が高くなり、ゆえに市場が機能せず、社会全体にわたる価値の向上が阻害されることである。たとえば、これまで治療を受けてきた病院から別の病院に移るのは難しいので、病院の間での競争が生じにくい。また、顧客の個人データを収集している

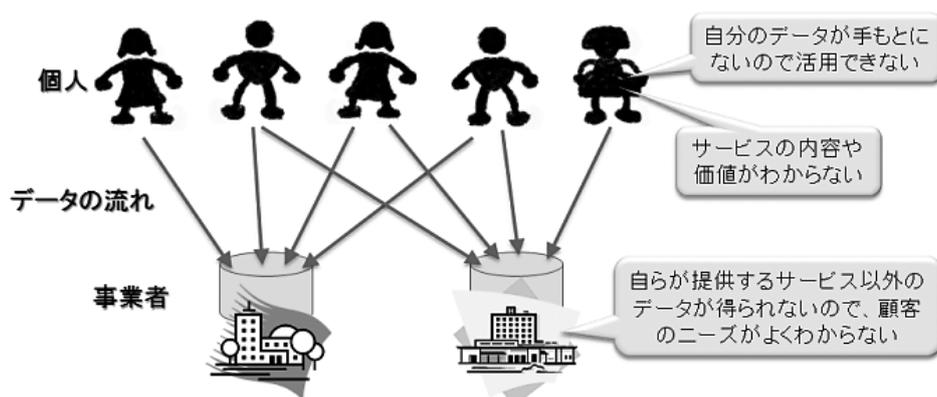


図1: 事業者による個人データの管理

ことが事業に必要なだとすると、そのようなデータを持たない新興事業者は既存の事業者からデータを入手するしかならないから、市場への参入が難しい。

ここで市場を十全に機能させ社会的価値を最大限に向上させるには、PDSが必要である。すなわち、図2のように個人データを本人が蓄積・管理し、相手を自由に選んで共有できるようにすべきである。他人の個人データを第三者に開示しようとするとは法的・倫理的な問題がしばしば発生するが、猥褻物陳列罪等に該当しない限り個人が自分のデータをどうしようと勝手である。

個人が本人のデータを管理し自らの意志で流通させることにより、個人は生活に関する多様な気付きを得るとともに、さまざまなサービスや商品の自分にとっての価値を正確に評価でき、その評価に基づいてサービスや商品を自由に選べる。一方、事業者は自らの事業に直接由来しないデータも取得して顧客のニーズをより正確に把握することにより、サービスや商品の価値を高めることができる。たとえば、ある病院での診療記録を別の病院で開示したり日常生活のデータを病院で開示したりすることによって、無駄な検査を防いだり効果の高い医療を受けたりすることが可能だろう。

問診票の記入などのデータ作成は当然ながら人間がしなければならぬが、センサデータの生成には人手がかからないし、また得られたデータの蓄積・管理は自動化可能である。たとえば診療明細書や検査データや処方箋は、患者が病院から紙でもらうのではなく、病院の情報システムが患者のPDSに自動的に送信すれば良い。

個人データの管理において本質的に難しいのは、サー

ビス事業者等の他者からデータ開示の要請を受けたときに、その目的（データの取り扱いやそれに基づくサービスの内容）と開示するデータや対価が釣り合っているかどうかを判断することである。この問題を解決するには、多くの人々のPDSのデータを集約することによって事業者や個人の信用度を評価しそれを社会的に共有することにより、一種の相場感を醸成するのが有効だろう。それでも高齢者や幼児が正しく判断するのは難しいだろうが、そのような場合に対処するには、家族や後見人が判断を代行する機能がPDSにあれば良い。

3 VRM

個人の生活や業務を支援する産業が機能するには、何らかの仕方で個人と事業者をマッチングする必要がある、従来は広告や推薦がその役割を果たしていた。PDSに蓄積される個人データを用いれば、そのマッチングの精度は、従来の方法よりも高くなるはずである。個人を主体とするそのようなマッチング、およびそれによってサービスや商品の組み合わせを個人にとって最適化することをVRM (vendor relationship management; 業者関係管理) と言う [Searls 12] [Siegel 09][VRM]。これはCRM (customer relationship management; 顧客関係管理) の対概念であり、CRMよりも社会全体としてのコストが安く、かつマッチングの精度が高い。「声の網」の情報銀行は、食事のメニューの最適な選択など、VRMの機能を備え

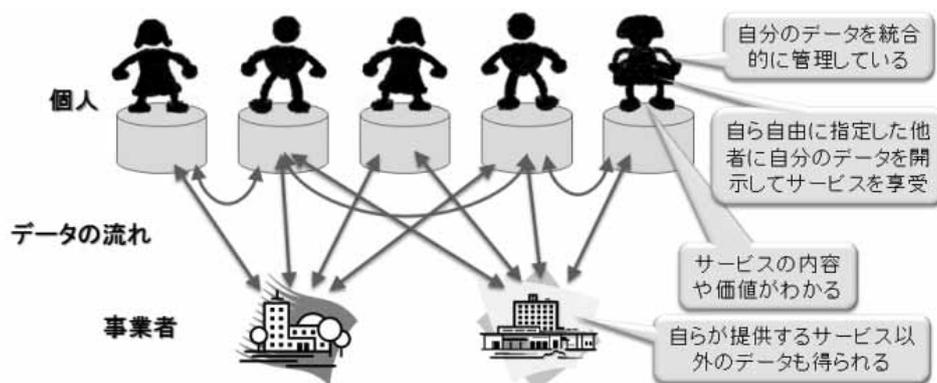


図2: 本人による個人データの管理

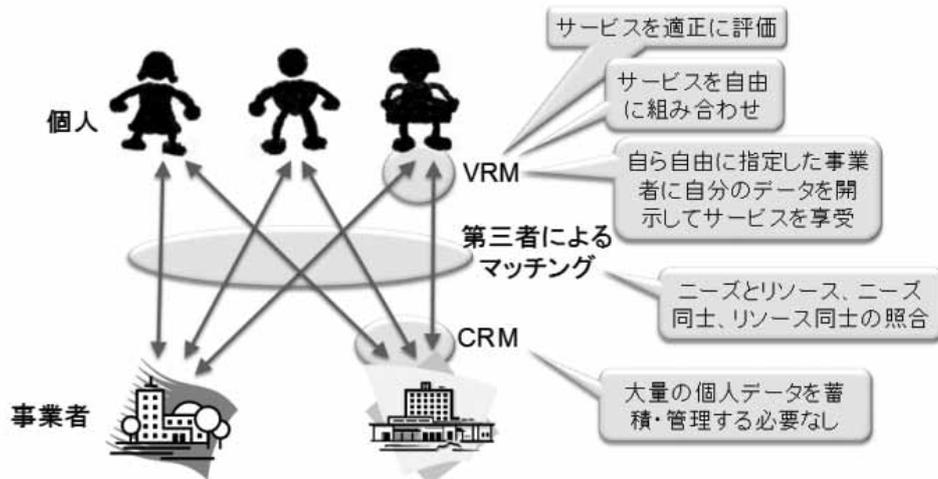


図 3: VRM とマッチング

ているように見える。

図 3 に示すように、VRM は顧客である個人が主体、CRM は事業者が主体となって、顧客と事業者とのマッチングを行なう。これらとは別に、顧客と事業者以外の第三者がマッチングを行なう形態も考えられるが、顧客の詳細なデータを用いれば、第三者によるマッチングも VRM と同程度に精度が高いだろう。現在の広告代理店や結婚相談所や就職・転職の仲介業者やレストランの紹介サイトや宿泊予約サイトがこの第三者に相当するのかもしれないが、これらは顧客の詳細なデータを使うことによってマッチングの精度を高められる余地が大きい。逆に言えば、PDS で個人が本人のデータを蓄積・管理するようになると、VRM のみならず第三者によるマッチングも精度の大幅な向上が期待できる。むしろ、資源の最適配分のためには、VRM よりも第三者によるマッチングの方が社会全体の価値の向上への貢献が大きい場合が多いかもしれない。たとえば、部分的な相乗りを許容するカーシェアリングやデマンドバスの運用のためのマッチングは、各乗客が個別に VRM を実行したのでは全体最適から大きくはずれてしまうことがしばしばあるだろう。

VRM の概念を敷衍すると、多数の個人から PDS のデータを集めて分析することによりサービスや商品进行评估し、さらにその評価を社会的に共有することにより市場の機能を高めることができるだろう。特に、いわゆる情報の非対称性の高いサービスにおいてはサービスの価値がステークホルダの間で十分に共有されず、ゆえにス

テークホルダの間での協調による価値の共創がうまく行かないことが多いが、この問題はサービスの評価を社会的に共有することによって解決できると考えられる。

たとえば医療サービスでもステークホルダの目的がばらばらである。患者の主な目的は健康度の向上にある。病院の目的は経営の観点からは利益である。保険者の本来の目的は被保険者の健康度の向上であるはずだが、現実には保険の償還金額の抑制が主目的になりがちである。このようにステークホルダの間で目的が共有されていないと価値がうまく共創されない。たとえば、病院の利益は主として診療報酬に基づき、診療報酬は保険点数に比例し、検査や注射や手術をすればするほど点数が稼げるから、過剰診療や薬漬けが起こり、患者にとっての価値が損なわれる。

本来追求すべき価値は患者の健康度だから、患者から見た病院や保険者のコストパフォーマンスの評価を社会的に共有することによって、その価値を高めるような協調と市場での競争とを促すことができるだろう。医療費が法律で決まっているために市場による価格調整がなされないという問題を解決するには法制度の変更が必要だが、各患者にとっての各医療サービスの価値が定量的に評価できれば、韓国で試みられているように、当該医療サービスの価格をその価値に応じて決めることも考えられる。

患者にとっての医療サービスの価値を客観的に評価するには、診療明細書、処方箋、および検査結果のデータ（カルテのデータはおそらく不要であろう）を各患者が

自分の PDS で蓄積・管理し、多数の患者からそのデータを収集して分析すれば良い。診療明細書等は医療機関が患者に紙を手渡す代わりに患者の PDS に格納すれば、データの取得・蓄積に患者の手間がかからない。電子カルテは特に中小の病院にはまだあまり普及していないが、診療明細書は歯科医院を除けばほぼすべての医療機関で電子化されており、それを患者の PDS に送るのは技術的には容易である。ほとんどの医療機関には診療明細書等を患者に電子的に提供するインセンティブがさしあたりはないが、多くの患者が PDS を使うようになれば、医療機関は PDS と電子的に連携せざるを得なくなるだろう。医療機関の間でのデータ連携のメリットを直接享受するのは医療機関ではなく患者であり、医療機関の経営の観点からはデータ連携のインセンティブはあまり生じない。ゆえにデータ連携は患者主導で進めるしかなく、PDS はそのためのツールとなる。

国民の 5% ぐらいが PDS に医療データを蓄積していて、それを集約し分析できれば、主要な医療機関のコストパフォーマンスの客観的評価（これぐらいの重症度のこんな病気でこの病院にかかると 5 年生存率は 60% だ、など）ができるだろう。しかし、PDS の医療応用がいきなりそこまで普及するわけではない。そこで、たとえば「あなたみたいな病状の人は世の中に 4% ぐらいいます」とか「あなたに似た人にはこの薬が効くみたいです」のようなアドバイスなど、データが少なくとも実現可能なサービスを提供することによって医療 PDS の利用者を増やす必要がある。

4 PDS の動向

PDS は、「声の網」があったにもかかわらず日本ではそもそもあまり知られていないが、欧米ではすでに多くの民間企業によって事業化されている。しかしそのほとんどは、個人が自分のデータを自由に活用できるようなものではなく、その意味で PDS としてはきわめて不完全である。それらのいわゆるデータブローカは、多数の個人から集めたデータをマーケティング等のため他の企業や政府に売ったりデータの分析を代行したりする事業者である。たとえば Acxiom [Acxiom] は購買

履歴や資産や健康に関する 2 億人以上の個人データを保有し販売している。Bluekai [Bluekai] は個人による Web の閲覧の記録を販売するとともにそれらのデータを解析するサービスも提供している。米国の連邦取引委員会 (FTC; Federal Trade Commission) は、一部のデータブローカが個人のプライバシーを侵害していることを指摘し、データブローカがどのようなデータを保有してそれをどのように使っているかを個人に開示することを義務付けるような規制を検討している。

一方、情報システムとサービスの設計にプライバシー対策を埋め込むための 7 原則からなる PbD (privacy by design) の考え方 [PbD] がデファクト標準になりつつあり、それに基づいて個人データの利活用を推進するための Personal Data Ecosystem Consortium という民間企業のコンソーシアムが 2010 年に発足している [PDEC]。また、英国の midata プロジェクト [BIS 11]、EU の個人データ保護規則案 [EU 12]、米国の消費者プライバシー権利章典 [WH 12] など、欧米諸国では、個人が本人のデータの使われ方をコントロールできるようにすべきだという、PDS に通じる考え方が具体的な政策にも反映されつつあるかに見える。

しかし、2013 年 6 月に CIA の元職員が暴露した PRISM システム等から推察するに、米国政府にとっては国家の安全保障が国民のプライバシーに優先するらしい。EU も表向きは人権を掲げながら IT 関連ビジネスにおいて米国を牽制したい本音が見え隠れする。このように考えると、PDS の理想を政治主導でトップダウンに実現するのは難しいと思われる。

5 分散 PDS

ボトムアップの普及も簡単ではなさそうである。個人データの二次利用に特化したデータブローカ事業でなく、データの一次利用と二次利用の両面にわたって個人の生活や業務を支援するような真の PDS は、収益事業として確立していない。たとえば、PHR (personal health record; 個人健康録) は医療に関する PDS であり、ヨーロッパの比較的小さな国々では政府によるトップダウンの施策の下で普及しているところもある

が、米国や日本では民間企業等による事業化の試みがことごとく失敗している。米国では、Google Health は閉鎖してしまったし、Microsoft HealthVault が普及しているという話も聞こえて来ない。日本でも、民間企業等による PHR 事業の立ち上げを目指した国家プロジェクトがいくつもあったが、持続可能な事業は実現されていない。

このように、単なるデータブローカでない PDS が収益事業として成立していないのは、コストがかかり過ぎ、その割にはサービスの価値が低く、利用者が増えないからだろう。特に Google Health や MS HealthVault では、個人データを囲い込まれる不安も、利用者が増えにくい大きな原因のひとつだと考えられる。また、期間限定のプロジェクトでは、事業者がデータを管理するのにかかるコストがプロジェクトの終了とともに賄えなくなるため、金の切れ目が縁の切れ目で、サービスが持続不可能になる。このようにサービスの利用者が増えず長続きもしなければ、価値の高いサービスを創造するための仮説検証がほとんど起こらず、ゆえに良いサービスが生まれず、という悪循環が生ずる。

この問題の本質的な原因は、PDS の運用事業者が多くの個人のデータをサーバに集めて集中管理している点にある。集中管理用のサーバの運用にコストがかかるのは言うまでもないが、運用事業者はデータ漏洩のリスクも負わねばならない。実際、多数の顧客の個人データが漏洩したという事件は枚挙に暇がない。大量の個人データの保管に伴うコストとリスクに耐えられるのは一部の大企業だけであり、そのような大企業も PDS 事業を持続可能にする価値の高いサービスを作り出せずにいる。ましてや、中小零細企業が大量の個人データを用いていきなり価値の高いサービスを提供するのは難しいだろう。日本の病院は治療の記録を 5 年間保存することが義務付けられているが、各患者の生涯にわたる記録を保管するコストとリスクを負うのは不可能と思われる。

PDS を持続可能な収益事業にするには、プライバシーに関する利用者の懸念を払拭し、運用コストを下げる必要がある。そのためには個人データの集中管理をやめるしかない。

プライバシーに関する懸念を解消するには、個人データを実質的に本人が管理する必要がある。つまり、データ

を格納するサーバを運営する（当該個人以外の）事業者にはデータの内容がわからず、またそのデータを本人がいつでも自由に利用できるような仕方で PDS を運用する必要がある。そのような PDS を分散 PDS (distributed PDS) と呼ぼう。

分散 PDS は、サーバに格納する個人データをサーバ運営事業者には内容がわからないように暗号化または秘密分散することによって実現できる。そのサーバとして Dropbox や Google Drive 等の基本料金無料のクラウドストレージを使えば、分散 PDS の基本運用コストはゼロである。さらに、いつでも自由に自分のデータを活用できて、他のクラウドストレージやホームサーバにデータを移すことも容易である、という意味で、データの可用性と可搬性が担保される。以下ではこのような分散 PDS のみを考えることにする。われわれが開発している PLR (personal life repository; 個人生活録) という分散 PDS のクラウドは、MyCloud [堀内 12] の方法を用いて、図 4 のように複数のクラウドストレージを組み合わせる構成される。

分散 PDS の基本運用コストがゼロなのは、Dropbox 等におけるフリーミアムのビジネスモデルのゆえであり、実際に生ずるコストはこのモデルによって回収される。つまり、分散 PDS に基づくサービスの価値があまり高くなくて分散 PDS の利用者が少ないうちはクラウドストレージ事業に対するインパクトはほとんどないが、サービスの価値が高まるにつれて分散 PDS の利用者が増え、クラウドストレージに料金を支払う利用者の割合も増し、クラウドストレージ事業者の収益も拡大する。

分散 PDS は、個人データを収集・分析する事業者のコストとリスクを激減させる。まず、個人データはそれぞれ本人が管理しているので、事業者は個人データを保管するための大規模なサーバを恒常的に運用する必要がない。本人のデータを蓄積・管理している多数の個人の PDS にアクセスできれば、各個人の同意の下でそのデータを使うことができる。

何のためにどんなデータをどう分析するかもわからないうちに莫大なデータを集めようとして資金を浪費しているケースが多いが、分析の目的と方法が明確になってから種類を絞ってデータを集めた方が明らかに効率的で

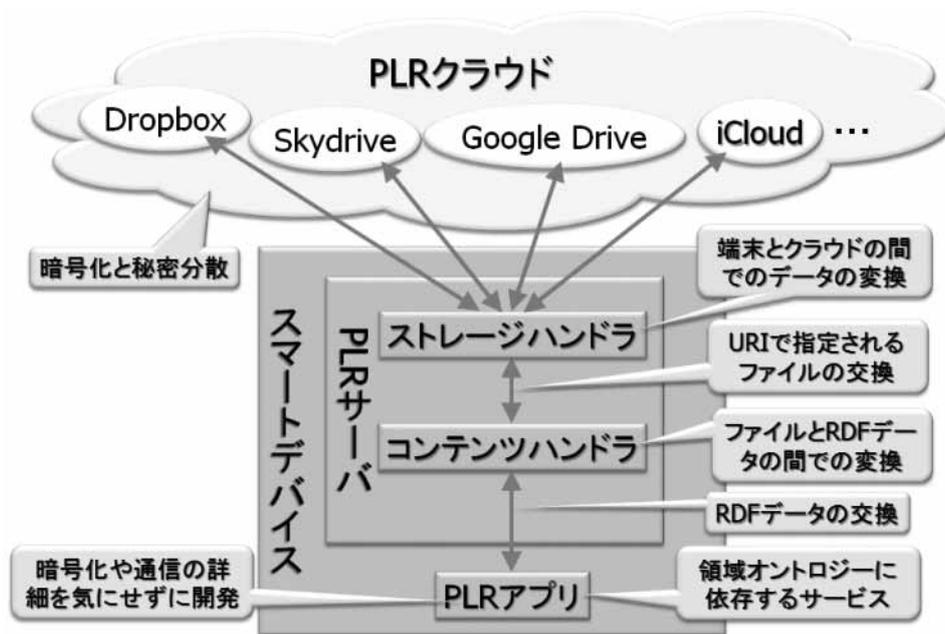


図 4: PLR クラウドの構成

あり、それは分散 PDS によって可能になる。データを集めて分析が終わったら分析結果を残して元のデータをすぐに破棄すれば良いから、個人データが漏洩するリスクもほとんどない。

多数の個人のデータを集中管理するよりも個人が本人のデータを分散管理した方が、社会全体でのセキュリティが高いことは言うまでもない。N 人分の個人データを集中管理していると、その管理者に対する 1 回の不正なアクセスによって N 人分のデータが漏洩する。これに対し、個人データを各個人が分散的に管理していれば、1 回の不正アクセスによって漏洩する個人データは 1 人分だけである。N 人のデータを取得するためのコストは、分散管理の場合には集中管理の場合のほぼ N 倍であるから、多人数の個人データへの不正アクセスには圧倒的に大きなコストがかかる。また、集中管理の場合と分散管理の場合とで暗号や秘密分散等のセキュリティ技術のレベルは同等だから、総じて分散管理の方がはるかに安全である。

6 分散 SNS

Facebook や Twitter や mixi などの従来の SNS では、利用者のコンテンツが Facebook 社や Twitter 社などの事業者のサーバに格納される。Gmail や Yahoo! メール等の Web メールも同様だが、事業者は SNS やメールのコンテンツをプログラムで分析することにより、各々の利用者に適合すると思われる広告を配信する。これは明らかに通信の秘密の侵害であり、不快・不安を抱く利用者も多いと考えられるが、いずれの国でも産業振興等のためか不問に附されている。Web メールにおける通信の秘密は文面の暗号化等によって担保できるが、それには発信者と受信者が暗号化に対応する必要があるが、それには煩雑だし、また暗号化を SNS に適用するのはさらに難しい。

コンテンツの閲覧・編集のための利用者インタフェースを分散 PDS に付加すると分散 SNS (distributed SNS) ができる。ここで言う分散 SNS は、コンテンツを事業者が大がかりなサーバで管理するのではなく各利用者が本人の分散 PDS で管理するような SNS である。その運用コストはほぼアプリの保守に係る費用のみ

なので、スケーラビリティが非常に高く、利用者が数十億人になっても構わない。さらに、非公開のコンテンツは本人が指定した他者にのみ開示されるという意味において、分散 SNS は、通信の秘密を守り言論の自由を担保する。

われわれはそのような分散 SNS として assemblogue を開発している。図 5 に示すように assemblogue のコンテンツは RDF グラフである。その中のノードはハイパーノード（グラフを含むノード）でも良い。各利用者の閲覧・編集権限はノードごとに設定できる。分散 SNS はこれまでもいくつか提案されてきたが、PLR のような手軽な運用基盤を欠いていたために普及しなかった。

assemblogue では各利用者のコンテンツを本人が PLR で管理するので、コンテンツを検閲したり流通を制限したりするには、各利用者個別にコンテンツを開示させたり開示を制約したりする必要がある。しかし、利用者が非常に多ければそれは事実上不可能であろう。また、PLR と assemblogue 自体がアプリを配布する手段であれば、各利用者による PLR と assemblogue の利用や更新を制限することもできない。クラウドストレージサービスは Dropbox や Google Drive の他にいくらでもあるので、その利用を制限するのも不可能だ

ろう。一方、assemblogue をアプリ配布のプラットフォームとすることにより、たとえば個人データを不正使用する恐れのあるアプリの配布を防ぐこともできる。

米国では、2011 年に成立したいわゆるパトリオット法により、通信の秘密を完全に担保するようなサービスやアプリを特定の企業が提供することができない。すなわち、Google や Amazon が米国で分散 SNS や分散 PDS を配布するのは不可能である。したがって、米国を含めて広く分散 SNS/PDS を普及させるには、それをオープンソースのフリーソフトウェアとし、不特定多数の人々がメンテナンスし続ける体制を作る必要があるだろう。その意味で分散 SNS/PDS は独占不可能である。

7 集めないビッグデータ

運用コストがかからない分散 PDS は、ほどほどのサービスをバンドルすることにより、利用者のメリットに基づいて普及する可能性が高い。また前述の通り、個人データの管理に伴うコストとリスクがほぼゼロなので、中小企業にも個人データの活用が可能であり、個人向けサービス等の市場への新規参入が容易である。こう

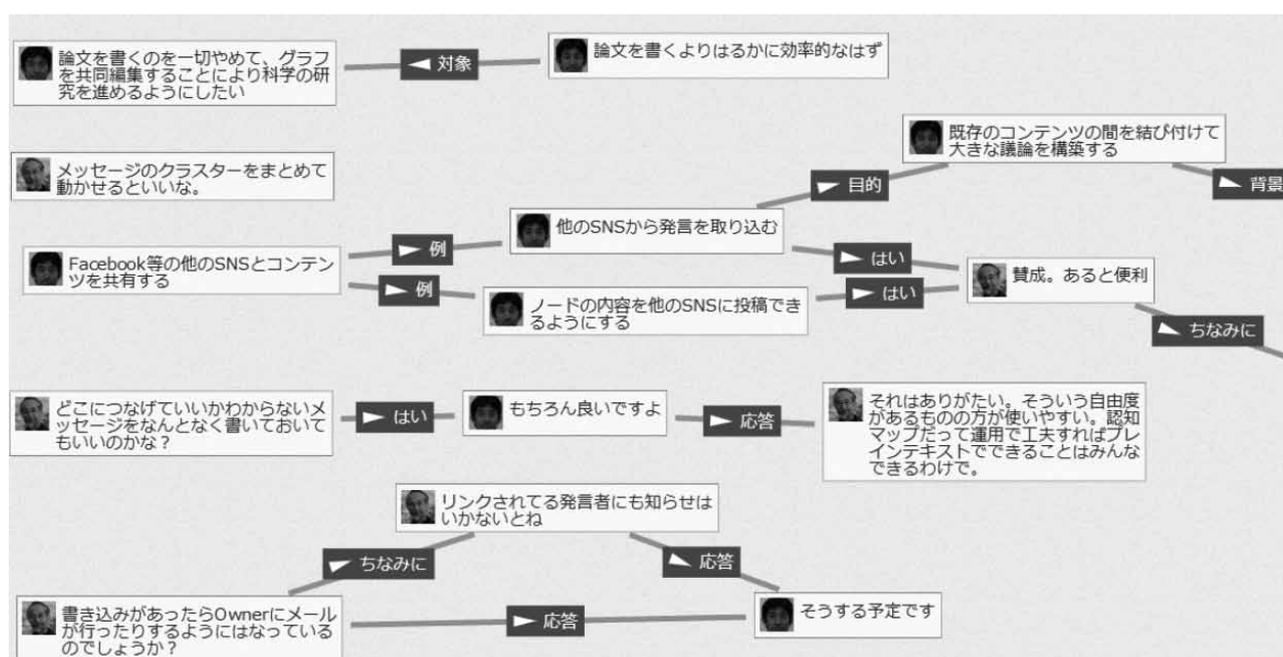


図 5: assemblogue の利用者インターフェース



して、既得権益や交渉力による作為ではなく、サービスの価値に関する競争が促され、市場が健全に機能する。

このように分散 PDS は個人や中小企業にとってのメリットが大きいと考えられる。しかし実は、分散 PDS の最大の受益者は、多数の顧客に個別に連絡する手段を持つ大企業や公的機関であろう。つまり、Google や Amazon やクラウドストレージ事業者や通信事業者や政府や自治体は、顧客（政府・自治体の場合には国民・住民）に分散 PDS を普及させるだけで、ほぼ労せずして巨大な個人データにアクセスできるようになる。

Google はすでに大量の個人データを収集しているが、その種類は個人の生活のごく一部の側面に関するものに過ぎず、医療データの収集にも失敗している。政府でも Google のような大企業でも、各個人の生活のあらゆる面に関するデータの収集は明らかに不可能である。しかし Google は、Google Drive の利用者のうち分散 PDS の利用者を簡単に判別でき、本人の同意の下に彼らから個人データを容易に収集できるだろう。そのために必要な投資は非常に小さく、またデータの収集は Google Drive 内での処理だから効率が良い。Dropbox や Amazon についても同様である。日本の通信事業者は顧客の電話番号を自由に使えないことになっているが、自社が発行する別の ID を使う設定の分散 PDS をバンドルしたスマートフォンを普及させれば、Google 等と同様にビッグパーソナルデータが簡単に得られる。政府が国民から個人データを安価かつ大量に収集できるようになれば、政策の立案や評価の精度を大幅に向上させられるだろうが、日本政府がそれをするには、マイナンバーを分散 PDS の ID として使えるようにすれば良い（PLR では、各個人は複数の ID を持ち、それらを名寄せできるのは本人だけである）。

しかしながら、Google は個人の全データを集められるという幻想に囚われてデータを囲い込もうとしているかに見える。だとすれば、上記のような簡単な理屈をすぐには理解しないだろう。今のところ、通信事業者等も個人データを囲い込もうとして資金を無駄使いしている。一方、Amazon は自社こそ VRM に最適の事業者であることを理解していると思われる。従来の推薦に代わって顧客の分散 PDS によるはるかに高精度のマッチングを使えば、Amazon はコストを大幅に削減しつつ

売り上げを増大させることができるだろう。これは分散 PDS によるイノベーションのほんの一例に過ぎないが、そのような大変動が数年の内に起こる可能性が高い。

参考文献

- [Acxiom] <http://acxiom.com/>
- [BIS 11] The Midata Vision of Consumer Empowerment. <http://www.bis.gov.uk/news/topstories/2011/nov/midata> (2011)
- [Bluekai] <http://bluekai.com/>
- [EU 12] Commission proposes a comprehensive reform of data protection rules to increase users' control of their data and to cut costs for businesses. <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/12/46&format=HTML&aged=0&language=EN&guiLanguage=en> (2012)
- [堀内 12] 堀内 公平：MyCloud: 複数ベンダのクラウドを用いて構成する高速で高信頼な分散ストレージ. コンピュータソフトウェア, 29 (2), 158-167 (2012)
- [iBank] 東京大学 空間情報科学研究センター / 地球観測データ統融合連携研究機構：情報銀行 <https://ibank.csis.u-tokyo.ac.jp/ibank/index>
- [Mitchell 01] Alan Mitchell: Right Side Up: Building Brands in the Age of the Organized Consumer. Harper Collins Business (2001)
- [PbD] <http://www.privacybydesign.ca/>
- [Searls 12] Doc Searls: The Intention Economy: When Customers Take Charge. Harvard Business Review Press (2012)
- [PDEC] Personal Data Ecosystem Consortium. <http://pde.cc/>
- [Siegel 09] David Siegel: Pull: The Power of the Semantic Web to Transform Your Business. Penguin (2009)
- [VRM] Project VRM: http://cyber.law.harvard.edu/projectvrm/Main_Page
- [WH 12] The White House: Consumer Data Privacy in a Networked World: A Framework

for Protecting Privacy and Promoting
Innovation in the Global Digital Economy.
[http://www.whitehouse.gov/sites/default/
files/privacy-final.pdf](http://www.whitehouse.gov/sites/default/files/privacy-final.pdf) (2012)