

パーソナルデータの分散管理による価値の最大化

Decentralized Management of Personal Data for Value Maximilization



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

1986年より2001年まで電子技術総合研究所。その間1988年から1992年まで（財）新世代コンピュータ技術開発機構に外向、2001年から2013年まで産業技術総合研究所。2013年から現職。専門は自然言語処理、認知科学、サービス科学など。日本認知科学会会長、言語処理学会会長等を歴任。

1 感染症とパーソナルデータ

COVID-19 が人々の間を物理的に分断し、リアルな場での交流が制限されている。ほぼ不特定多数の他の学生や教員との交流による学習や発想の機会を提供するという重要な機能が世界中の大学で大きく損なわれているが、その最大の被害者はまだキャンパスでの出会いに恵まれていない新生入生である。大学に限らず企業などの組織や学会などのイベントでも同様の問題が生じている。

この問題をオンラインの手段で解決する必要があるが、それにはパーソナルデータの分散管理（パーソナルデータを管理運用する権限を本人に集約する、つまり個人に分散すること）が有効だろう。出会いはある種のマッチングで実現できるはずである。出会いが有意義であるためには、出会う人々が互いのニーズを満たせる等の条件が必要であり、リアルな場での出会いがその条件を満たす可能性が高いのは同じ大学の同じ学科に属するとか同じ講演会に参加している等の手がかりによる。すると、さらに詳細な手がかりとなるデータを各個人について予め用意しておいてそれを使って個人同士を自動的にマッチングすれば、リアルな場での出会いよりも価値の高い出会いをオンラインで実現できるのではないだろうか。精度が高く安全なマッチングを実現するには各個人のパーソナルデータをすべて本人が活用できることが望ましい。

一方、COVID-19 はいわゆるショックドクトリン、つまり危機対応を理由に（応々にして非民主的な）社会改革が強行されるリスクを生んでいるが、パーソナル

データの分散管理は民主主義の強化によってそれに対抗するにも有効と考えられる。以下では、パーソナルデータの分散管理が大きな価値を生み、事業者の主導によって普及する可能性が高いこと、分散管理が民主的なガバナンスを強化する可能性、分散管理が生む価値と表裏をなすリスク、そのリスクを管理するための民主的なガバナンス等について述べる。

2 集中から分散へ

2018年に施行されたEUの一般データ保護規則（GDPR; General Data Protection Regulation）等を契機として、パーソナルデータ（特定の個人に関するさまざまなデータ）の保護と活用に関する制度の整備が世界中に広がりつつある。それは概ねデータに関する人権を強化するものであり、これによって、パーソナルデータの管理は事業者に集約する（各事業者が多くの個人のデータを集中管理する）これまでの方法から本人に集約する（つまり個人に分散する）方法に移行すると考えられる。

データの集中管理とは多くのデータ主体（個人や法人）のデータを特定他者が管理することであり、分散管理とはデータ主体が自らのデータを管理することであるとする。本稿では主にパーソナルデータについて論ずるが、その議論の多くは企業等のデータにも妥当する。

データの集中管理よりも分散管理の方がさまざまな意味において望ましい。まず、下記の名寄せと利便性は、集中管理（他者による管理）の有無によらず分散管理（本

人または本人に近い代理人による管理) がなされている場合に成り立つ。

名寄せ：従来の集中管理においては、各個人のデータが複数の事業者に断片的に散在していて価値が低い。また、複数の集中管理システムの間でパーソナルデータを名寄せするには、原則として明示的な本人同意が必要であり、またシステムの間で個人IDが統合されていなければ名寄せのためのデータ処理に大きなコストがかかり間違いも生じやすい。これに対し、各集中管理システムからパーソナルデータ(のコピー)を本人に提供すれば、それがすなわち本人の手もとでの正確な名寄せであり、これによってデータの価値が高まる。その際にパーソナルデータへの他者のアクセスは増えないので、集中管理システム間での名寄せの場合に必要なとなるような本人同意は不要である。

利便性：これまではパーソナルデータを他者が管理していたので、データを使うには本人の同意に加えて管理者の同意も必要であり、両方の同意を得るのは非常に難しかった。しかし、パーソナルデータを本人が管理していれば、本人の意思だけでデータが活用できる。また後述のように、活用の際にデータを他者に開示する必要がなく安全なことも多い。

次の安全性と本人利用は、パーソナルデータの管理者が本人だけである(集中管理がなく分散管理のみがなされている)場合に成り立つ。

安全性：たとえば、集中管理されている数千万人分のデータが詐取されたというような事件は枚挙に暇がないが、1千万人分のデータの詐取にかかるコストは1千万円未満だろう。これに対し、分散管理されている1人分のデータを詐取するには1万円以上かかると考えられる。つまり、集中管理をやめて分散管理に移行すれば安全性が桁違いに高まる。また、集中管理を拡大するとデータ漏洩等のリスクも増大するが、分散管理の導入によってリスクを抑制しながらデータの活用を拡大することができる。

本人利用：私的な秘密や黒歴史を含む機微なパーソナルデータは、事業者に開示されることが少なく、従来はあまり活用されなかった。しかし、パーソナルデータを他者が管理せず本人が管理すれば、機微なデータも本人の手もとで活用できる。したがって、パーソナルデータの管理が集中方式から

分散方式に移行することによって、人権が強化されるだけでなく、データの活用が促進され、データが生み出す価値が大幅に増大し、産業や文化の振興につながるだろう。

ただし、いかに分散管理が普及しても集中管理が不要になるわけではない。分散管理の目的はデータ主体自身のメリットを高めることであり、他の目的にデータを用いるには他者による集中管理が必要である。たとえば、納税者はあまり税金を納めたくないの、課税のエビデンスとなるデータは税務当局が集中管理することが望ましい。税だけでなく、社会保障や災害対策など国民に過不足なく提供すべきサービスにも集中管理が必要であり、これらのサービスの基盤が日本の場合にはマイナンバーであり、マイナンバーカードは各国民がそれらを含む行政サービスに電子的にアクセスする手段である。また、企業が顧客や社員のパーソナルデータを管理する必要は10年後にはほとんどなくなるかも知れないが、顧客や社員の連絡先や契約書は集中管理せざるを得ない。

いずれにせよ、今後5~10年でパーソナルデータの分散管理が普及する可能性が高い。それはおそらく、個人が自らの権利を守るために自分のデータを管理しようとするからと言うよりもむしろ、経済的利益のために企業が分散管理を促進するからだろう。後述のようにパーソナルデータの分散管理が集中管理よりも大きな価値を生むが、これは、分散管理の方が個人の行動(特に購買行動など)により直接的に介入できるからである。

3 分散 PDS

パーソナルデータをデータ主体本人が管理運用するための技術的な仕組みをPDS(Personal Data Store)¹⁾と言う。著者は分散型の(つまり本人以外にデータの管理者がいない)PDSの一種であるPLR(Personal Life Repository)^{2) 3) 4)}を開発している。PLRは図1のように個人アプリや事業者アプリに組み込んで使うソフトウェアライブラリである。

各利用者のPLRがPLRクラウドとデータを同期することにより、利用者同士がデータを共有できる。PLRクラウドはGoogleドライブやOneDriveなどのクラウドストレージの集合だが、PLRはデータを暗号化してからクラウドにアップロードし、クラウドからダ

ダウンロードした後に復号するので、クラウド事業者にはデータの意味がわからない。また、クラウド事業者と契約してクラウドを利用するのは PLR や関連アプリの開発・提供者ではなくそれらの利用者なので、この仕組みは利用者が何億人いてもアプリの保守コストだけで安定稼働する。

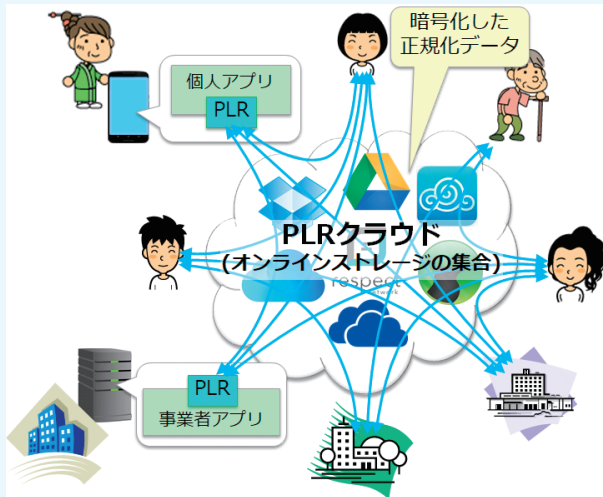


図1 PLRによるデータ共有
PLRはアプリに組み込むライブラリであり、これがPLRクラウドとデータを同期することにより利用者間でデータが共有される。PLRクラウドのデータは暗号化されておりクラウド運営事業者には内容がわからない。この仕組みは利用者が何億人でもアプリの保守コストだけで安定稼働する。

一般に PDS は本人がパーソナルデータを管理することにより前節の名寄せと利便性をもたらすが、さらに安全性と本人利用をも担保できるのは PLR のような分散 PDS だけである。また、大量のデータを頻繁に共有するには PLR クラウドのような共有ストレージが必要である。つまり、PLR 型の分散 PDS のみが、名寄せから本人利用までをすべてサポートしデータ共有の自由度が高い。

4 パーソナルデータの価値

パーソナルデータの活用は下記のようないくつかの場合に分類できる。

0次利用 (マッチング) : パーソナルデータ (が表わす本人のニーズ) に適合する個人向けサービス (商品を含む。以下同様) を選定する。たとえば、個人の身体の採寸データや色の好み等のデータを用いて本人に合う服や靴を選んだり、旅行の予定や食べ物の好みに関するデータと旅先のレストラン等とをマッチングしたりすることが考えられる。

1次利用 (サービス) : 本人へのサービスの価値を高める。たとえば、医療機関が患者の病歴や服薬のデータを参照することによって安全で効果的な治療を施す、保険会社が被保険者の運転のデータから事故のリスクを見積もって保険料を調整することにより安全運転を促す、採寸データを仕立屋に開示して服や靴をテーラードしてもらうなどが考えられる。

2次利用 (統計分析) : 多数の人々のデータを統計分析して一般的な知識を獲得する。多数の患者の診療記録や日常生活行動のデータを収集・分析して生活習慣病や感染症に罹患するリスクを予測するモデルを作る、多くの生徒の教育や生活行動のデータから効果的な教育法を解明する、多くのドライバーの運転と事故のデータから運転の仕方と事故のリスクとの関係を明らかにする、などがその例である。

また、パーソナルデータは活用するだけでなく売買することによって価値 / 収益を生むこともある。これらの関係を図2に示す。パーセンテージは各場面でデータが生み出す価値の対 GDP 比である。これらはきわめて乱暴に推測したかなり不正確な値だが、値の間の大小関係は下記で正しいと考えられる。

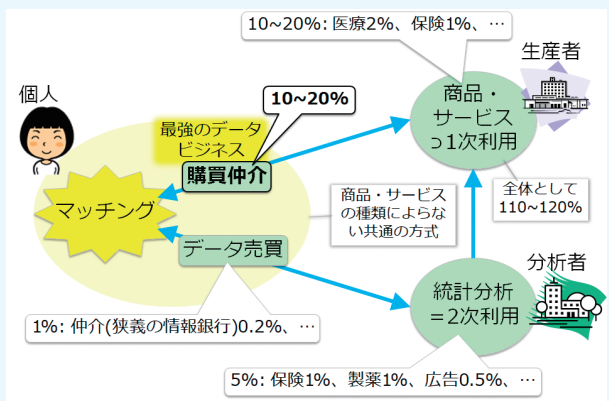


図2 パーソナルデータの価値 (対 GDP%)
PLR に保管されたパーソナルデータとサービスとのマッチング (0次利用) に基づいてサービスの購買を仲介する事業の潜在市場規模は GDP の 10 ~ 20% であり、明らかにこれがパーソナルデータが生み出す最大の価値である。パーソナルデータの1次利用が保険や医療などの各分野で生み出す価値は GDP の 2% 未満と考えられる。2次利用の価値は当然ながら1次利用の価値より小さく、データ売買の価値はさらに小さい。

PLR に付属するパーソナル AI エージェント (PAIA) が PLR に保管されたパーソナルデータとサービスをマッチングすることによって個人に適合するサービスを選定するが、その詳細は後述する。

個人向けサービスは家計消費 (これだけが GDP に含まれる) だけでなく業務の場での個人消費 (個人が勤務

先の負担でサービスを購入すること)と家事や育児などの無償 C2C サービスを含むので、全体としての価値は GDP を越える。そのうちデータの1次利用が生み出す価値は GDP の 10 ~ 20% 程度と思われる。しかし、サービスは多くの分野に分かれており、市場規模の大きい保険や医療の分野でもデータが生む価値はそれぞれ GDP の 2% 未満だろう。たとえば、健康度等に応じて保険料を還付する保険商品の国内市場規模は 8 兆円ほどだが、これが保険分野でパーソナルデータが生み出す価値の主な部分と考えると、保険業全体でパーソナルデータが生む価値は GDP の 2% より少なそうである。また、国民医療費は 45 兆円程度だが、そのうちデータが生む価値はやはり 10 兆円未満だろう。

一方、さまざまなサービス(前述のように商品を含む)とマッチングして購入を仲介する事業は、サービス全体の価値の 10% 超、つまり全体としておそらく GDP の 10 ~ 20% の手数料収入を得ると考えられる。マッチングのアルゴリズムやマッチングに用いるサービスの電子カタログの作り方があらゆるサービスの分野にわたって共通だとすれば、これが最強のデータビジネスであり AI の最も重要な応用と言えよう。

パーソナルデータの2次利用が生み出す価値が大きいのは保険や製薬の分野と考えられるが、いずれにおいてもその価値は GDP の 1% より小さいだろう。たとえば製薬業の国内市場規模は 10 兆円程度だが、そのうちパーソナルデータに由来する価値は半分に満たないと考えられる。また、広告業の市場規模は国内では GDP の 1.2% 程度(世界全体では 0.7% 程度)だが、やはりデー

タが生む価値はその半分より小さいだろう。

パーソナルデータの2次利用全体の価値は1次利用全体の価値(前述のように GDP の 10~20% とと思われる)よりはるかに小さいはずだが、個人がパーソナルデータを2次利用のため分析者に販売する事業の市場規模はさらに小さく、GDP の 1% に満たないだろう。いわゆる(狭義の)情報銀行はこの販売を仲介する事業であり、その市場規模は GDP の 0.2% 未満と考えられる。ちなみに、分散管理の下では、個人が持つデータと分析者のニーズとのマッチングも図2のように PAIA が担当することになるだろう。

以上を要するに、パーソナルデータを用いる事業のうちで0次利用(マッチング)に基づくサービス仲介業が圧倒的に大規模であり、その手数料収入の合計は国家予算に匹敵する GDP の 10 ~ 20% と推定される。サービス仲介業は、このように潜在的な市場規模が大きいだけでなく、後述のように特にマッチングにおいてパーソナルデータを他者に開示する必要がないのでデータの活用が容易であるため実現性も高いという点で、パーソナルデータのきわめて重要なユースケースと考えられる。

5 サービス仲介

このサービス仲介業を営む新種の流通事業者を「メディアータ」と呼ぶ。図3に示すように、メディアータは、生産者(商品の生産者とサービスの提供者)からサービス(商品を含む)の情報を収集することにより、サービスの電子カタログを作成する。個人に属する PAIA は、

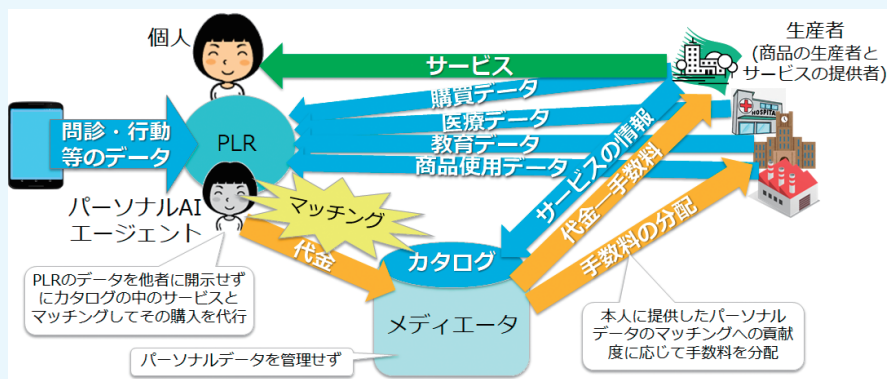


図3 分散マッチングと収益分配

メディアータが生産者からサービスの情報を収集してカタログを作り、生産者から本人に提供され PLR に保管されたパーソナルデータをパーソナル AI エージェント (PAIA) がカタログの中のサービスとマッチングする。マッチングしたサービスが購入されると、メディアータはその代金から手数料を差し引いた金額をサービスの生産者に送金し、マッチングに使われたパーソナルデータを本人に提供した生産者に対してそのデータのマッチングへの貢献度に応じて手数料を分配する。生産者はパーソナルデータを本人に提供することで収益が得られるので、パーソナルデータの本人への集約による活用が拡大する。

生産者から本人に提供され PLR に保管されたパーソナルデータ（購買データや医療データなど）を、このカタログの中のサービスとマッチングし、マッチングした（さらに本人が気に入った）サービスの購入を代行する。

メディアータは、従来の流通事業者と同じく、世の中に多数存在することになるだろう。たとえば医療などの分野に特化したメディアータや地域限定のメディアータも考えられる。また、PAIA にも複数の種類があると想定される。各個人は各時点においてそのうちの 1 種の PAIA を用いて同時に複数のメディアータを利用すると考えられる。各 PLR 利用者は PAIA を通じて随時あらゆるメディアータとつながってカタログを利用することができる。つまり、あらゆる PLR 利用者がいつでもすぐに任意のメディアータの顧客になり得る。

メディアータはサービスの代金から手数料を差し引いた金額をサービスの生産者に送金し、さらに、マッチングに使われたパーソナルデータを本人に提供した生産者に対し、そのデータのマッチングへの貢献度に応じて手数料を分配する。分配の原資は前節で述べたように全体として GDP の 10 ~ 20%（狭義の情報銀行の手数料収入の 50 ~ 100 倍）だから、これは生産者にとって大きな収益になるだろう。さらにメディアータは、パーソナルデータの 1 次利用に対しても生産者から手数料（既述のようにその総額も GDP の 10 ~ 20% 程度と思われる）を徴収し、このデータを本人に提供した生産者にサービスの質向上への貢献度に応じてその手数料を分配することもできるだろう。こうして、生産者はより良質のパーソナルデータをより多く本人に提供することでより大きな収益が得られるので、パーソナルデータがさらに本人に集約されてますます盛んに活用され、ゆえに以上のようなサービス仲介事業が持続的に発展すると考えられる。

GDPR 等の制度により事業者がパーソナルデータを本人に電子的に提供すること（データポータビリティ権の保証）が義務付けられても、それによる事業者のメリットがなければ、データが活用しやすい仕方でも本人に集約されず、あまり価値を生まない。データポータビリティが有効に機能して社会的価値が高まるには、上述のようにデータポータビリティが事業者の利益をもたらすことが必須である。

メディアータは大きな手数料収入を得るが、大量の

パーソナルデータを永続的に管理する必要がないので、必要経費は現在の電子商取引事業者の場合より小さい。生産者と分析者においても大量のパーソナルデータの永続的管理は不要である。

一方、メディアータも生産者も、多数の個人から本人同意や法律に基づいてデータを収集し分析することができる（つまり分析者にもなれる）。その分析結果を、メディアータはカタログの改良に、生産者はサービスの開発や改良に使うと考えられる。同じデータを何度も分析する場合に通信コストを抑制するためデータをしばらく保管することはあり得るが、それを越えて永続的に保管する必要はない。それを含むデータ最小化（data minimization; 所定の目的を越えたデータの保管や利用をしないこと）を DRM や秘密計算によって技術的に強制すれば、事業者（メディアータ、生産者、分析者）にとっても個人にとっても安心・安全が確実に担保される。

6 直接的データによる分散マッチング

上記のマッチングは、パーソナルデータが他者に開示されないという意味において分散的である。この分散マッチングは、機微情報を含む名寄せされた価値の高いデータを（他者に開示せず安全に）活用するので、Amazon 等の流通事業者による従来のリコメンデーション等よりも精度が高い。PAIA は、カタログを個人端末にダウンロードしてパーソナルデータを外部に出さず端末の中でマッチングすることもできるし、必要なパーソナルデータをメディアータのサーバに送ってサーバ内でマッチングすることも考えられるが、後者の場合も、メディアータを含む他者がパーソナルデータを閲覧したりマッチング以外の目的に使ったりすることを技術的に防止できる。

分散マッチングの精度が高いのは、より直接的なデータを用いるからである。ここで直接的であるとは、サービスに対する本人のニーズ（サービスの本人への適合性）を直接的に表現するということである。たとえば身体の採寸データは購買履歴等よりも衣服の適合性に関する直接的なデータである可能性が高い。買い物に出かける前にメモする買い物予定リストも購買履歴より直接的と考えられる。既存のデータを組み合わせることによって直接性が高まることも多いと考えられる。これまでは、コ

コンテンツ閲覧や検索や購買等の履歴が個人適応に使われてきたが、それらは間接的なデータであり、ターゲティングやリコメンデーションに用いた場合の精度が低い。購買履歴に基づく書籍等のリコメンデーションの精度が例外的に高いのは、書籍等の商品の場合には購買データが商品の評価（直接的データ）の良い近似になっているからである。

集中管理よりも分散管理の方が直接的データに適している。つまり、データは直接的であるほど機微性が高いので他者による集中管理に適していない。しかも直接的であれば有用なので活用したいから分散管理が必要である。

直接的データを積極的に生成してデータ主体本人のPLRに蓄積し活用するようなサービスを実現することにより、分散マッチングの精度を高めることが望ましい。直接的データによる分散マッチングのアルゴリズムは協調フィルタリングのような統計的な手法よりも単純な条件判断に近いものになると考えられるので、分散マッチングの精度向上にはリッチな直接的データを活用する方がマッチングエンジンの性能向上よりもはるかに有効であり、それはテクノロジーではなくビジネスモデルの課題である。

7 民主主義のバージョンアップ

パーソナルデータの分散管理は、人権を強化するだけでなく、多様なステークホルダの合議による社会的決定という民主的なガバナンスの基盤にもなる。つまり、個

人が自分のデータを総合的に管理していれば、そのデータを本人の意思によって政府だけでなく大学や研究機関や企業やNPOに開示できるから、それらの機関が多数の個人からデータを収集し分析することにより、客観的根拠に基づいた自由で質の高い言論が可能になる。こうして、政府を含むさまざまな機関や個人が互いを監視しつつ連携するという民主的なガバナンスが成立する。

一方で、インターネット等のテクノロジーは、アラブの春に見られるように民主的な連帯を強化した面もあるものの、逆にフィルターバブルやヘイトスピーチを助長して世界を分断し、また全体主義的な管理も強化しつつある。その状況の概略を図4に示す。著者を含む多くの人々はこの図の上に近い社会を望むだろう。ここでは詳細に立ち入らないが、それは民主的で自由な社会の方が経済的にも精神的にも豊かだとの信念に基づくと考えられる。しかし近年は、上の方でも2つの赤丸の間の分断が深まりつつあり、この分断を止揚できなければ全体主義に陥るリスクがあることは歴史の教訓である。左上の赤丸は自立した市民が科学的根拠に基づいて熟議することにより社会的決定がなされる世界、右側の赤丸はそれと対立する反知性主義やポピュリズムの世界を表わす。ここで反知性主義は必ずしも否定すべきものではない。「経験と勘」や「直観と想像」は合理主義を補完し、科学（仮説検証）における仮説構築に不可欠である。つまり、2つの赤丸の一方に偏るのではなく、両者を連携させる必要がある。それによって民主主義はより強固で効率的な態様に進化するのではないだろうか。

2つの赤丸の間の分断を埋めるということは、ファク

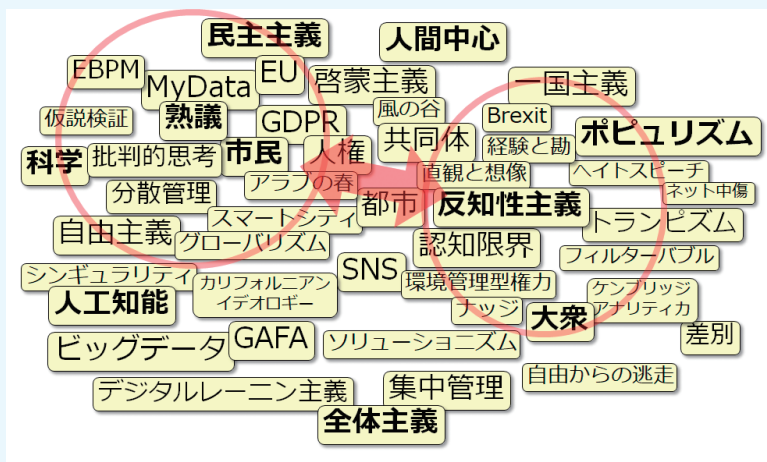


図4 テクノロジーと政治思想と関連事項のタグクラウド

上の方（民主主義や人間中心主義）と下の方（全体主義）が対立し、左上の赤丸（自立した市民による科学的根拠に基づく熟議）と右側の赤丸（反知性主義やポピュリズム）の間の分断が広がっている。この分断を埋めることによって民主主義を強化する必要がある。

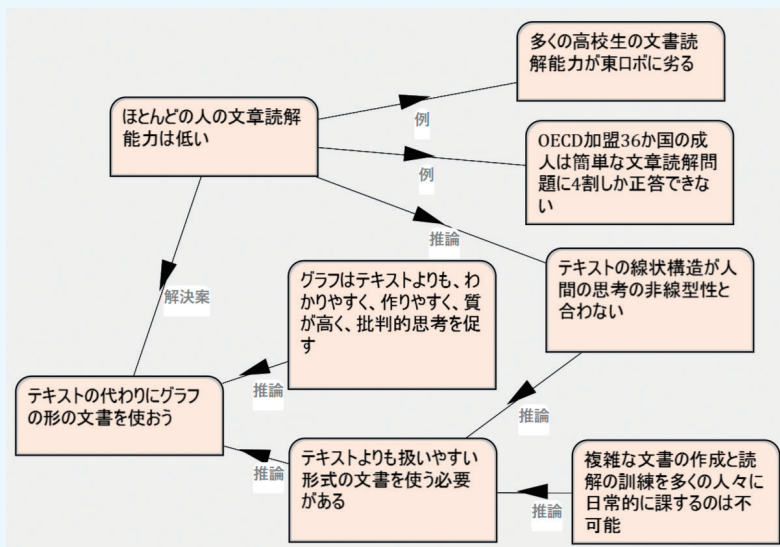


図5 グラフ型文書

図式として意味構造を明示することによって、文書の作成と読解が容易になり、質が向上し、また人間の批判的思考力を高めることが知られている。グラフ文書は、簡単な文や句を含むノードとそれらの間の意味的関係を示すリンクからなり、それによって、任意のテキスト文書の内容をすべて明確に表現できる。これにより、ファクトやロジックと直観や想像との連携を強化する効果が大きいだろう。

トやロジックと直観や想像とを密接に連携させるということである。それにはいわゆる発想支援のようなものが有効だろう。発想支援では何らかの図式によって文書の内容を表現することが多い。グラフ（ネットワーク）や木のような図式として文書を表現することにより、文書の作成と読解が容易になり、質の高い文書を作成でき、また人間の批判的思考力が高まること、先行研究等によって知られている。特に文書の作成はファクトやロジックと直観や想像とを連携させる作業であり、図式による構造の明示によってこの連携が強化されると考えられる。また、図式の作成に際しては構造を明示する必要があるため熟考が促されるだろう。さらに、多くのSNSのような刹那的な反応が抑制され、ヘイトスピーチ等が力を失うのではないだろうか。

図5に示すグラフ型文書はそのような図式の一種である。このグラフにおいては、各ノードが簡単な文や句であり、各リンクは両端のノードの間の意味的な関係を表わす。このようなグラフ文書の重要な特徴は、通常のテキスト文書で表現される内容をすべてテキスト文書以上に明示的に表現できる点にある。つまり、グラフ文書をテキスト文書の代わりに用いることができ、しかもその表現はテキスト文書以上に明瞭である。他の発想支援の方法はこの要件を満たさない。

グラフ文書でテキスト文書を置換することにより、ファクトやロジックと直観や想像との連携を強化する図

式の効果が文書の作成や活用にわたるライフサイクル全体に及ぶので、従来の発想支援よりも連携を強化する効果が大きいだろう。たとえばKJ法では、図解化によってグラフのような図式を作り、それに基づく叙述化によってテキスト文書を作るが、テキスト文書は図式的な構造に乏しいため直観と想像を刺激しにくい。これに対し、図5のように明示的な構造を持つ完全な文書は、その作成・読解を含むあらゆる場面でファクトやロジックと直観や想像を同時に支援することができる。

個人情報や企業秘密を含む非公開文書を作成し限定的に共有して活用するため、グラフ文書も前述のPLRで管理すべきであることは言うまでもない。また、グラフ文書からは詳細な知識を高い精度で抽出できるので、グラフ文書はPLRに付随するパーソナルAIエージェント（PAIA）の最も重要な学習用データになると考えられる。PAIAはその知識に基づいてグラフ文書の作成や活用をきめ細かく支援することができるだろう。

8 パーソナルAIエージェント

分散マッチングの対象はサービス（商品を含む）の購買だけでなく個人のあらゆる行為を含み、PAIAは個人の生活の総合的な最適化を支援することになる。たとえばPAIAは、生活習慣病や感染症を予防したり仕事や学業の成績を伸ばしたりするために個人の行動変容を促す

機能も持つだろう。PAIA はまた、グラフ文書の作成を支援することにより、個人の思考に立ち入って本人の思想の形成に深く関与することになると考えられる。

パーソナルデータの分散管理によって人権が強化されると述べたが、実情はもっと複雑である。分散管理は本人主導のデータ活用を促進するが、それは人間による自己情報コントロールを必ずしも意味しない。パーソナルデータが本人に集約されるとそのデータの使い方に関して判断する機会が増えるが、あらゆる場面で人間がいちいち判断するのは煩雑すぎて不可能だから、PAIA に判断の多くを委ねるしかなかろう。いわゆるシンギュラリティの到来までに 100 年はかかると著者は考えているが、そのはるか前の段階で、たとえばある種の医療診断などの精度についてはすでに AI が人間を凌駕している。したがって、人間が大量の判断を PAIA に委ねた方が全体として本人のメリットが大きいのではないか。現在まだ大きくなくても数年で大きくなるに違いない。

いかに本人のメリットが大きくても、選挙で誰に投票するか、結婚相手に誰を選ぶか、といったことまで AI に決めてもらうのは抵抗があるかも知れない。しかし、そうしたことに関する人間の判断がそもそもどの程度自律的であったかも不明であるから、このような議論の実り多い着地は難しいだろう。ここではむしろ、PAIA の判断が別の意味において社会全体に対しきわめて重大な帰結をもたらす得ることに注意しておきたい。

PAIA は、各個人の行動を少なからず左右して、GDP を上回る価値の交換と国家予算規模の富の分配を司り、さらにデータの 2 次利用も促進することにより、生活や産業や政治や文化にきわめて大きな影響を及ぼす。たとえば、PAIA による分散マッチングが安価なサービスを優先しすぎると市場はデフレに陥り、逆に高価なサービスに偏るとインフレを引き起こす。マッチングが分散的であっても、さまざまなマッチングのアルゴリズムが全体として類似した性質を持てば、それが社会全体を変える力は大きい。すでに SNS 等のフェイクニュースが世論や選挙の結果を左右しているとも言われているが、フェイクニュース等が人間の判断に影響を与えて間接的に行動を変えさせるのに対し、PAIA による選択の支援は人間の行動を直接変えてしまう。グラフ文書の作成支援は、従来の SNS や発想支援システムよりも直接的かつ詳細に個人の思考内容に介入する。これらの支援

サービスは怠惰な人間に適している。そして、われわれ人間が 1 人残らず怠惰であることは言うまでもない。

したがって、PAIA には高度な中立性と倫理性が要求され、AI 関連技術の中でとりわけそのガバナンスが国際的にも重要な課題になることは必定である。既述のように PAIA には複数の種類があり得るが、SNS やメールアプリと同じく、実際にはほとんどの利用者がごく少数の種類のパ IA のいずれかを使うことになるだろう。そのように市場を寡占する PAIA に対してはとりわけ厳格な管理が求められる。たとえば、PAIA のソースコードや学習データは一般公開し、多様なステークホルダーが精査できるような民主的なガバナンスの下に置くべきである。データの管理を分散化してもアプリの機能はある程度標準化する必要があるが、その機能の設計には民主的な分散管理が必須である。

参考文献

- 1) Gordon Bell (2001) A Personal Digital Store. Communications of the ACM, 44: 86-91.
- 2) Kôiti Hasida (2013) Personal Life Repository: Distributed PDS for Data-Driven Improvement of Your Welfare. AAI Spring Symposium 2013 (Data Driven Wellness), Stanford University.
- 3) 橋田浩一 (2015) 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. 社会情報学, 3 (3), 87-98.
- 4) Kôiti Hasida (2019) Decentralized Personal Data Store (PLR) for Convenient, Cost-Minimizing, and Secure Utilization of Your Data. Proceedings of Asia Pacific Society for Computing and Information Technology 2019 Annual Meeting (APSCIT 2019 Annual Meeting)