Vol.2020-CSEC-91 No.25
Vol.2020-SPT-40 No.25
Vol.2020-EIP-90 No.25
2020/11/26

パーソナル AI エージェントの社会制度的位置づけ

加藤綾子 [1]　中川裕志 [2]

概要：複雑化を辿る一方の社会情報環境に対処するためには，自分自身に関する情報がどのように使われるかについて，個人が把握し，さらに個人の意思を反映させる仕組みが必要である．本研究は，このような個人データに関する本人の意図する利活用条件が記述されたデータベースを備えたパーソナル AI エージェントを提案する．まず，パーソナル AI エージェントの概念を明らかにし，次に，必要な機能を列挙し，個人データの利活用条件のデータベースの生成の仕組みについて検討する．最後に，外部の事業者，パーソナル AI エージェント，データ主体である本人という 3 者の間にトラストを構築する枠組みについて述べる．

キーワード：パーソナル AI エージェント，プライバシー，個人情報

Social institutional position of personal AI agents

AYAKO　KATO[†1]　HIROSHI NAKAGAWA[†2]

Abstract: In order to deal with the ever-increasing complexity of the social information environment, it is necessary for individuals to understand how information about themselves is used and to reflect their intentions. This research proposes a personal AI agent equipped with a database that describes the terms and conditions of utilization intended by the individual regarding such personal data. First, we will clarify the concept of personal AI agents, then enumerate the necessary functions and examine the mechanism for generating a database of personal data utilization terms and conditions. Finally, we describe the framework for building a trust between the business operator, the personal AI agent, and the person who is the data subject.

Keywords: Personal AI agents,　Privacy,　Personal Information

1. はじめに

現代の複雑な情報環境で自らの意思の通りに，かつ自らの不利益を被らないような情報行動をすることは難しくなってきている．この難しさは幼少者や高齢者では際立つであろうが，社会的にアクティブな活動をしている人であっても無視できない状況になっている．この状況に対応するためには個人が自身の情報を扱う能力を向上させることがまず考えられるものの，外部の情報環境の複雑さ，ときには悪辣さを鑑みると，生身の人間では限界があるのではないだろうか．

この問題状況はつまるところ，個人のデータを外界に対してどのように使うかという指針と言い換えられる．個人データとしては，例えば，生年月日のような個人情報，家族や友人関係，銀行の口座情報，あるいは物品の購入に関する趣味趣向などがある．個人がこれらの情報を使っても良いと思えるかどうかは,その時々の文脈によって異なる．

個人データの使い方を支援ないし管理，監督してくれる AI システムがひとつの解決策となることが期待できる．このような AI システムを以下では「パーソナル AI エージェント」と呼ぶ．

パーソナル AI エージェントは個人データの使い方に係わるシステムだから，そこでの中心的役割は「個人のデータの利用法のデータベース」が担うことになる．具体的に言えば，これは，データの利用法に関する個人の同意条件を書き表したものである．

個人が自らのデータに関して他者によるいかなる使用を認めているかに関する記述は，その個人のアイデンティティであると言え[1]，また，その個人の生存時における，自己の情報の扱いに関する意思そのものであると言える．

AI ガバナンスという観点からみると，データ主体である個人がパーソナル AI エージェントを利用して，外界にいる相手方をガバナンスするという建付けであろう[a].

以下，2 節ではパーソナル AI エージェントに関連するこれまでの議論を概観し，3 節と 4 節では個人の情報の取り扱いに関して AI がガバナンスするという観点で，パーソナル AI エージェントの必須機能や懸念事項などについて検討する．

1 文教大学
Bunkyo University
2 理化学研究所
RIKEN

[a] ちなみに，AI ガバナンスのもう一つの意味は，AI 自体を人間が統御する，ないしは制御可能なものになるように設計するということであろう．

Vol.2020-CSEC-91 No.25
Vol.2020-SPT-40 No.25
Vol.2020-EIP-90 No.25
2020/11/26

2. 背景

国内外で策定されている複数の AI 倫理指針の中で，個人のデータの使い方を支援ないし管理するというパーソナル AI エージェントに近い概念について述べているのは，IEEE の Ethically Aligned Design (EAD)の Version 2 (v2)（以下では IEEE EAD v2 と記す）[2]と First Edition (1e)（以下では IEEE EAD 1e と記す）[3]である．これらに基づいたパーソナル AI エージェントの簡単な概念化が[4]でなされている．他方，橋田浩一の研究では，パーソナル AI エージェントを導入したパーソナルデータエコシステムが検討されている[5][6].

そこで，本節では，IEEE EAD v2，EAD 1e，そして，橋田案のパーソナルデータエコシステムについて概観する．

2.1 IEEE EAD の想定

IEEE EAD v2 [2]では，「Personal Data and Individual Access Control」という章において，既に次のことが指摘されている．

パーソナルデータの権利と個人によるアクセス制御について，個人のデジタルデータの使用に関するインフォームドコンセントと，(データへの) アクセスを定義する権利を個人が持つことが基本的に必要である．個人情報を束ねたり再販したりすることから生じる結果を個人に明示的に認識させるようなポリシーおよび実践と併せて，個人が自己のアイデンティティとパーソナルデータをキュレート（curate）することを支援するメカニズムが必要である．データ収集時点ではなく，データが使用される時点で，このメカニズムは，データ主体によるコントロールがなされることを求める．

(1) データの対象範囲

自己に関するデータの発生は，生前から始まる．母体の中で受精卵の存在が確認された時から，「私」に関する生物学上，遺伝学上のデータが生じることになる．出生後，これらの情報は医療健康関連情報として扱われる．

出生後，人は市民としての権利を獲得し，行政上のデータが生じることになる.その後,人として人生を歩む中で,人は教育を受けたり，移動をしたり，海外へ行ったり，物を購買したり,通信をしたりする訳だが,それらに伴って,教育データ，移動データ，入出国データ，購買データ，通信データ，SNS や IT プラットフォームの利用ログなどが生成される．このほかにも，視聴データ，就業や納税に関するデータ，政治や宗教や何らかのコミュニティなどへの社会参加に関するデータ,資産状況や信用に関するデータ,金融データなどが生じる.そして,これらのデータは死後,ディジタル遺産となり得る．

人が，自己に関するこれらの膨大なデータ（パーソナルデータ）を管理・運用するために，IEEE の Ethically Aligned Design（EAD）では Individual Agency の導入が検討されている．IEEE EAD v2 では，アルゴリズムが支配的である時代におけるパーソナルデータと個人アイデンティティをどのように再定義するかという問題提起がなされている．

(2) 個人に提供されるべき環境

個人にとって信頼できる ID 認証サービス（trusted identity verification services）が必要であり，特に銀行や政府，通信などの規制のある業種は，市民や消費者にデータ認証サービス（data-verification services）を提供すべきである．個人識別可能情報：Personally Identifiable Information (PII)は，法的に保護された個人の資産（the sovereign asset）として，いかなる場合も優先されるべきであり，データが使用される時には，個人がパーソナルデータの使用に関してコントロールするための手段を常時，使えるようにすべきである．

個人によって，パーソナルデータのアクセス・コントロールがなされるべきではあるものの，個人がそれらを行うことは実際には難しい場合がある．そこで，IEEE EAD v2 は，見つけやすくて使いやすいパーソナルデータ・マネジメントツールを，サービスプロバイダーが彼ら自身のサービスのインタフェース内に設けるべきであると指摘している．そのデータマネジメントツールは，誰が何の目的でデータにアクセスするのかを明らかにし，かつ，ユーザーがアクセスの許可をマネージするようなものであるべきであり，EU 一般データ保護規則：General Data Protection Regulation (GDPR)に則ってユーザーが欲すれば，そのサービスから自分自身のデータを簡単に取り除くこともできるようなものであるべきであるとされる．

また，個人本人が，何に同意したかを遡って確認することができることを支援する一方で，個人の環境をスキャンすることで，それに応じたパーソナルなプライバシー設定を行うような，個人情報のコントロールを行うための「アルゴリズミック・ガーディアン」が開発されるべきであるとされる．

(3) A/IS の導入によって懸念されるべき事項

Autonomous and Intelligent system (A/IS)を通じて簡単かつ偶然に共有されるデータが，個人が共有を希望しないかもしれないインタフェースに使われるという問題がある．IEEE EAD v2 は，規約は分かりやすく，同意は条件付きで動的であるべきであると指摘している．そして，データ管理者（Data controllers），プラットフォーム運営者，システム設計者らは,ユーザーが A/IS に直接コンタクトする時にはその成り行きをモニターすべきであるとされている．

今後，多くの A/IS が個人からデータを収集するだろうが，彼らは互いに直接関係をもっていなかったり，システムが個人とは直接に接点を持たなかったりするであろう．そのような状況下で意味のある同意とはどのように提供され得るのだろうか．データ主体がシステムと直接関係をもたないところでは，サービス提供当初の規約に拠らず，同意は動的であるべきである．A/IS は，収集と使用に関して

Vol.2020-CSEC-91 No.25
Vol.2020-SPT-40 No.25
Vol.2020-EIP-90 No.25
2020/11/26

制限を表明している全てのユーザーのデータ選好を解釈するように作られるべきである．

消費者と事業者の契約交渉という観点においては，契約管理プラットフォーム（contract management platform）を作る必要があるだろう．

サードバーティのベンダーがパーソナルデータにアクセスするのを制限するために，あるいは，トランザクションデータをベンダーの内部使用のみに制限するために，無料のデータ交換モデルに代わる有料モデルが必要かもしれない．その場合，事業者はコスト計算に基づく課金モデルや，広告なしの定額料金制を導入しなければならないだろう．

事業者にとってプライバシー影響評価：Privacy Impact Assessment (PIA)は不可欠である．事業者内部においては，雇用主と従業員との間でデータ使用に関する同意を明確にする必要があるとともに，従業員データに関しても PIA が必要である．

個人のプライバシーに係る A/IS には，不正な侵入によって政府当局に情報を送るようなバックドアがあってはならない．

A/IS によってどのような種類の処理がなされているかに関して人々は理解する能力を喪失しつつある．とりわけ，高齢者や成人の精神障碍者による同意には信頼性に欠ける部分がある．

IEEE EAD v2 では，以上のような指摘がなされている．IEEE EAD v2 の想定では，管理されるべきパーソナルデータが，プラットフォーム事業者を含むサービス提供事業者（サービスプロバイダー）のデータベース内にあって，そのデータがさまざまな場面で共有や使用されることに対して個人による管理が及ぶようにするため，サービスプロバイダーが彼ら自身のサービスのインタフェース内に個人データの管理ツールを個人向けに設けるべきであるということが考えられている．

最近の傾向として，事業者がダッシュボードなどと呼ばれる管理画面を設けて，事業者保有の個人データの利用の可否などを個人が設定することができるような手段を個人顧客向けに提供する例が見られる．IEEE EAD v2 の想定はこれと同様に，事業者の責任としてデータの管理ツールが個人に提供されるべきだというものである．これを図示すると図 1 のようになる．

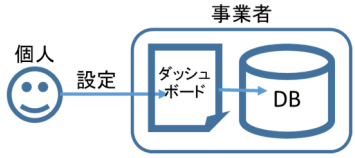


図 1．IEEE EAD Version 2 の
Individual Access Control の概念図
注：ダッシュボードという語は筆者加筆

一方，IEEE EAD v2 よりも後の時期にまとめられた IEEE EAD 1e[3]では，「Personal Data and Individual Agency」という章において，個人が機械可読式の契約条件（machine-readable terms and conditions）を生成することができるようなツールが各個人に提供されるべきであると指摘されている．そして，いかなる環境においてもその個人の契約条件をキュレートするようなアルゴリズミック・エージェントないしパーソナルデータが，各個人に提供されるべきであると指摘されている．さらに，安全かつ明確で限定的なデータ交換をコントロールするために，信頼し得る ID を生成することによってサービスへのアクセスが各個人に提供されるべきであると指摘されている．そこでは，特定の文脈における ID の使用を検証し，証明し，支援するような，信頼し得る ID 認証サービス（trusted identity verification services）を個人が利用できるべきであるとされている．

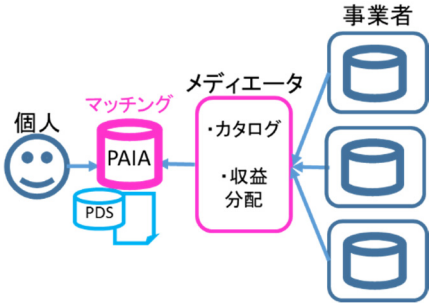
2.2 個人側で動作するパーソナル AI エージェント

橋田浩一の考案するパーソナルデータエコシステム[5][6]においては，パーソナルデータの管理権限が個人に帰着するという大前提を置く．そして，パーソナルデータを閲読して使用することができるのは，基本的には個人本人だけであるという仕組みが検討されている．

この仕組みを実現するためにデータポータビリティが必要であり，事業者から個人本人に還元されたパーソナルデータは各個人の Personal Data Store (PDS)において管理される．そして，個人に代わってパーソナル AI エージェントが，PDS にある膨大なデータを閲読して，個人の判断を補助するための情報の検索や選択肢を絞り込む作業などを行うということが想定されている．

橋田の考案するエコシステムにおいては，データを用いたマッチングを，基本的には個人の手元で，パーソナル AI エージェントが行う．このエコシステムではデータ活用における仲介を「メディエータ」が担うのだが，仲介者であるメディエータは基本的にはデータを保持せず，①事業者の提供する財・サービスの情報が記述された機械可読式のカタログと，②収益分配のみを行う．そして，パーソナル AI エージェントがその機械可読式カタログと PDS のデータ（場合によっては個人が追加入力した機微情報など）とを個人の手元でマッチングする（図 2）．この方法によって個人は財・サービスの購入の可否を判断する．

Vol.2020-CSEC-91 No.25
Vol.2020-SPT-40 No.25
Vol.2020-EIP-90 No.25
2020/11/26



PAIA：パーソナル AI エージェント

図 2．橋田の考案するパーソナルデータエコシステムにおけるパーソナル AI エージェントの概念図

なお，メディエータは多くの個人と事業者を接続する仲介者であるため，既に多くの個人のアカウントを有する Google，Facebook，Amazon，Apple (GAFA)などのプラットフォーム事業者はメディエータになることが容易である．同様の理由で，電力会社や携帯電話の通信キャリアなどのインフラ事業者もメディエータになれると考えられる．

このエコシステムにおいて，パーソナル AI エージェントの開発や提供を行う事業者はメディエータとは独立の事業者であるケースも考えられる．しかし，上記の GAFA はパーソナル AI エージェントの開発元や提供元にもなることができる．あるいは GAFA の利用者インタフェースを個人適応する仕組みによって，プラットフォームの提供するサービスの一環としてパーソナル AI エージェントの機能を提供することもできるであろう．

3. パーソナル AI エージェント

3.1 概念設計

パーソナル AI エージェントにはどのような機能が必要であり，どのような形で提供ないし実装されるべきであるのだろうか．個人の権限や，事業者の責任については，EU の一般データ保護規則：General Data Protection Regulation (GDPR)や IEEE EAD v2, 1e によって洗い出されているが，個人のための代理者の仕組みについてはまだ必ずしも十分に整理されていない．

そこで，本研究は，前節で登場したパーソナルガーディアン, パーソナルエージェントに基礎をおく概念として「パーソナル AI エージェント」を提案し定義する．

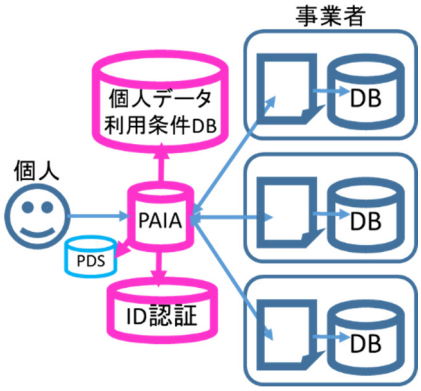
パーソナル AI エージェントの実装方法の一つの候補は，橋田案のように完全に個人側に存在して個人のために働くエージェントである．IEEE EAD v2 で述べられているように，サービス提供事業者が提供するサービスや機能に付属ないし内蔵されて存在して，サービス提供時ないしサービス利用時に動作するという選択肢もある．あるいは，IEEE EAD 1e が指摘するような個人の機械可読式の契約条件が，さまざまなサービスの利用時に適用されて，なおかつ，その時々の個人の意思や文脈に応じてこれが動的に変化して適用されるという方法も候補である．これらの実装方法の候補を念頭におきつつ議論を進める．

今後，A/IS のさらなる導入によって，取引およびデータ利用がますます自動化されることを見据えると，個人が何に同意しているかという情報が「個人のデータの使い方に関するレジストリ」として存在して，取引やデータ利用の都度，そのレジストリが参照されると便利である．この機能は，必ずしも特定の事業者のサービスに付随したものであるとは限らず，個人に付随し，種々の事業者に対応する存在として考えることができる．このことを図 3 に概念化した．

IEEE EAD 1e が指摘するような個人の機械可読式の契約条件は，個人のデータの利用法の条件を記したデータベースとしてパーソナル AI エージェントのシステムに存在する方法がある．その仕組みにおいては，個人がさまざまな事業者のサービスを利用する際などにそのデータベースが参照されて，各個人に特化した契約が事業者と個人との間で結ばれるという案が考えられる．

パーソナルデータそのものが個人の PDS の中にあってもよいが，データがどこに保管されているかということよりも重要なのは，データの使われ方が各個人の希望通りであるかということである．

次項では，このようなパーソナル AI エージェントに必須の機能について検討する．



PAIA：パーソナル AI エージェント

図 3．個人データ利用条件データベースを備えたパーソナル AI エージェントの概念図

3.2 パーソナル AI エージェントの必須機能

前節において概観した IEEE EAD v2 と 1e から，パーソナル AI エージェントには，信頼できる ID 認証サービス（trusted identity verification services），個人が何に同意したかを遡って確認することができる機能，パーソナルなプライバシー設定を行うために各個人の状態や環境をスキャンする機能，機械可読式の契約条件（machine-readable terms and conditions）を生成する機能が少なくとも必要であると

Vol.2020-CSEC-91 No.25
Vol.2020-SPT-40 No.25
Vol.2020-EIP-90 No.25
2020/11/26

いうことが読み取れる．

いかなるオンラインサービスにおいても個人認証の機能は必須である．詳細は次節で述べるが，個人からのアクセスを受ける事業者にとって，その個人が信頼し得る個人であるかを確認したり，既に登録のあるユーザーであるかを確認したりする必要がある．個人およびパーソナル AI エージェント側にとっては，接続先の事業者が信頼し得る事業者であるかを確認できること，および間違いなくその事業者であるかを確認できることが必要である．

このほか，パーソナル AI エージェントは個人データを扱うため，基本的なセキュリティと，各国の個人情報保護法に則ったデータ保護，プライバシー保護が不可欠である．これらの機能を提供する事業者においては，Privacy Impact Assessment (PIA)を行う必要がある．

さらに，パーソナル AI エージェントは，自動化が進む中で，通常時は個人に代わり自律的に動作するが，判断に迷う時，あるいは重大な契約を行う時には，きちんと個人にアラートを示して，個人の判断を仰ぐシステムでなければならない．

パーソナル AI エージェントは，自身が代理しているデータ主体個人の個人データの使い方を定義している．上記のような判断を仰がれた場合にデータ主体個人の判断を支援するために，個人が現在いったい何に同意しようとしているのか，これまで類似ケースでどのような同意をしてきたかを，データ主体個人がいつでも確認することができるような機能が必要である．

上記のことも含め，パーソナル AI エージェントが持つべき必須機能をいかに列挙する．

- 個人認証（identity verification）
- 基本的なセキュリティ
- データ保護，プライバシー保護，および PIA
- 機械可読式の契約条件（machine-readable terms and conditions）の生成
- パーソナルなプライバシー設定を行うために各個人の状態や環境をスキャンする機能
- 自動化と本人関与の両立
- 自分自身が今，何に同意しようとしているのか，何に同意してきたかという履歴を確認できるような機能

3.3 個人データ利用条件データベースの雛型

個人のデータの使い方を定義した個人データ利用条件データベースはどのように作られるのだろうか．パーソナル AI エージェントの中心になる個人データの利用条件データベースを個人ごとに from scratch で形成していくのは時間がかかり，形成するための行動の中で失敗したり損害を被ったりする可能性がある．この状況を避けるために，個人データの利用条件データベースの雛型があると役立つ．パーソナル AI エージェントのソフトウェア開発事業者が，このような雛型を準備しておいてくれることを期待したい．ただし，開発事業者にとっても from scratch で構築することの負担は大きい．むしろ，開発業者は，販売したパーソナル AI エージェントが，利用者個人の利用によって形成される個人データの利用条件データベースの情報を利用者個人の同意の上で，匿名化あるいは統計情報に近い形で提供してもらい，そうして集めた情報から個人データの利用条件データベースの雛型を形成する方法が考えられる．更新された雛型は利用者のパーソナル AI エージェントも**必要な部分**をダウンロードして使うことができれば，パーソナル AI エージェントの開発事業者と利用者の双方にとって win-win な状況となる．

ここで，上記の**必要な部分**をどのように特定するかが問題になる．利用者が今まで買ったことがない商品，例えば高級車，海外旅行，不動産などの購入をしたいときには，購入商品を他の人々が使っているパーソナル AI エージェントがどのような個人データを提供して購入しているかという情報が**必要な部分**ということになる．この仕組みの概要を図 4 に示す．パーソナル AI エージェント開発業者が個人 A, B, C から同意のうえで収集した個人データ利用条件が個人データ利用条件データベースに登録されている．ただし，個人情報保護の観点からデータベース中では A, B, C は 11, 23, 543 という乱数（ハッシュ値）を用いて記述されている．

図中，一番右側にいる別の利用者がある商品（図 4 では車）を購入する場面を想定する．個人ごとに価値観や経済状況も異なるので，自分に近い価値観や経済状況の人々の個人データの利用条件データベースを利用したい．しかし，価値観や経済状況はセンシティブな個人情報なので，直接は開発事業者も収集できないと予想される．そこで，情報技術的に使える方法として協調フィルタリング[7]という推薦システムで使うアリゴリズムが候補として考えられる．協調フィルタリングでは，目下購入しようとしている商品以外の商品購入パタンが自分と似ている人を探す．そうして見つかった人が，今，購入しようとしている商品をどのような条件，すなわち商品提供者にどのような個人データ提供して購入しているかをダウンロードして参考にするという方法である．もちろん，無条件でダウンロードした購入法を真似するのは危険である．したがって，パーソナル AI エージェントの利用者であるデータ主体個人と，インタラクティブなやり取りをして，データ主体個人の意思を決めていくことになる．この作業は知的なユーザインタフェースを必要とするので，AI 技術を利用して設計することになり，今後の課題となる．

さて，この図 4 では仮名化はされていたが，同一の個人のデータが同じレコードに入るという，いわゆる横串を通した状況である．この場合は仮名化されていても実際の個人を特定できる可能性が高い．この特定を避けたいなら，このデータベースの各セルから仮名を除去していわゆる無

Vol.2020-CSEC-91 No.25
Vol.2020-SPT-40 No.25
Vol.2020-EIP-90 No.25
2020/11/26

名化をしてしまう方法が考えられる．この方法を採ると，個人特定は困難であるが，上記の協調フィルタリングの手法は使えないという欠点もある．結局，仮名化で適切な必要な部分を高精度で検索できるか，無名化してあまり高精度で検索できなくするかは，パーソナル AI エージェント開発業者のガバナンスがきちんとできているかどうかに依存する問題である．ガバナンスがきちんとしており，利用者がこの開発事業者をトラストできるなら，高精度のシステムを利用できることになる．

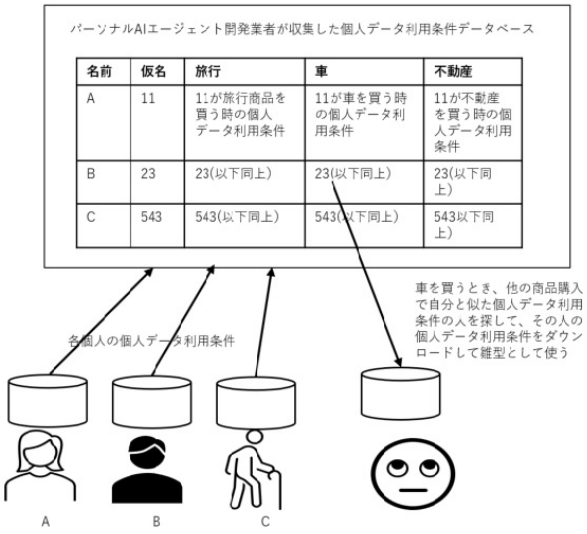


図 4. パーソナル AI エージェント開発事業者が収集している個人データ利用条件データベースから協調フィルタリング手法で，現在必要な部分を探してダウンロード

4. ガバナンスとトラスト

4.1 ガバナンス

デジタルデータの扱いに関連する議論において，ガバナンスという語はやや多義的に用いられている．そこで，まず，誰が何を使って何をガバナンスするのか，ガバナンスとは何か，という問題をここでの議論に当てはめてみよう．

ここまで述べてきた議論では，データ主体の個人が複雑な外界，すなわち時には詐欺師もいる情報環境をパーソナル AI エージェントを使ってガバナンスするという建付けであった．パーソナル AI エージェントが自分の手元にある孤立したソフトウェアであると，いかに優秀なソフトウェアであっても，複雑な外界をガバナンスするには経験が少なすぎると考えられる．そのための教師役を果たすのが前節で述べたパーソナル AI エージェント開発事業者の保持する個人データ利用条件データベースである．このデータベースは既に述べたように，開発業者が自社の販売した多数のパーソナル AI エージェントが，実利用経験の集積結果を収集して構築したものである．したがって，これを適宜ダウンロードして用いれば，多くの人々の知恵によって外界をガバナンスできると期待できる．もっとも，自分にフィットする個人データ利用条件を使うことが大切なので，そのための選択手法として前節では協調フィルタリングの応用を提案したのであった．

4.2 個人とパーソナル AI エージェント間のトラスト

パーソナル AI エージェントの利用者であるデータ主体の個人から見れば，果たして自分はパーソナル AI エージェントを十分にガバナンスできているのかという不安もある．以下でこの不安を払拭するために必要なデータ主体個人とパーソナル AI エージェントの間のトラスト関係について検討してみる．

データ主体個人がパーソナル AI エージェントをトラストできるために必要なことは以下の通りである．

(1)信頼できる開発事業者が開発提供したパーソナル AI エージェントであること．

(2)パーソナル AI エージェントの判断がデータ主体個人にとっていかなるときにも不利益にならないこと．

(3)パーソナル AI エージェントは，過去の個人データ利用法に当てはまらない新規案件がある場合は，独断せずにデータ主体個人と相談すること．

ただし，(3)では，データ主体個人に容易に理解できるように状況を説明する必要がある．また，データ主体個人から質問があれば納得できるように答えることも重要である．このようなやりとりを通じて，データ主体個人はパーソナル AI エージェントをトラストできるようになるだろう[b].

(3)の場合に必要な知的なやり取りを行うためには，AI 技術が必須である．自然言語でやり取りするなら，自然言語処理が必要だが，自然言語でのやり取りはえてして冗長になる．むしろ，選択肢列挙と選択のようなパタン化されたやり取りのほうがデータ主体個人にとっては楽かもしれない．このあたりのインタフェースの設計は今後の課題である．

まったく別のトラストの作り方として，データ主体個人は複数のパーソナル AI エージェントを駆使し，多数派パーソナル AI エージェントの判断を採用するという方法もあるが，相当に面倒である．

いずれにせよ，データ主体個人が最終判断は自分で行ったという感覚を持つことが大切である．

見過ごしがちなのは，逆方向のトラスト，すなわちパーソナル AI エージェントが利用者であるデータ主体個人をトラストできるかという問題である．データ主体個人がパーソナル AI エージェントに間違った情報を与えたり，以前の判断方針すなわち個人データ利用条件を頻繁に変更し

[b] ここでは，主人が執事をトラストすることと似たような状況を考えている．

Vol.2020-CSEC-91 No.25
Vol.2020-SPT-40 No.25
Vol.2020-EIP-90 No.25
2020/11/26

たりすると，パーソナル AI エージェントのアルゴリズムが安定した判断をすることが困難になる．非常に賢いパーソナル AI エージェントができれば，データ主体個人の意図を忖度した上で伺いを立てるようなインタフェースが望ましいが，これは相当に難しい問題であろう．

4.3 パーソナル AI エージェントと事業者間のトラスト

データ主体個人が，自身の情報の使われ方に関して，事業者や政府など他者に対して許諾を与えたり同意をしたりすることを意識的に行えれば，個人としては事業者をトラストするベースラインが得られる．このような生身の個人と事業者間の許諾や同意は従来から行われてきた. しかし，ソフトウェアの利用契約において同意のクリックを，契約文書を読まずに行ってしまうことが多いだろう. あるいは，そのソフトウェアを使いたければ書かれた通りの契約をするしかないような，いわゆる付合契約では，個人側は付き合いでしかたなく行っている状況であり，トラストには至っていないと考えるほうが自然である．このような状況をパーソナル AI エージェントを用いて改善する枠組みを以下の例に基づいて説明する．

従来は，事業者側の都合で利用規約が作成されて個人データの使い方が規定されていた．いわば「事業者のデータの使い方のよくあるパタン」が規約に表現されていたと言える．

個人が自らの情報に関するコントローラビリティないし自己決定権を持つということが重視されるのであれば，個人が「私が許可するデータの使い方のよくあるパタン」を表現して事業者側の規約に対抗する個人データ利用条件として，それを事業者に守らせたり，その「よくあるパタン」から外れた利用がなされた際に異常が検知されてアラートを出してくれたりするような機能があると便利である.

このことを以下の例によって説明する．例えば，ある個人 X が「私が金融機関のオンライン口座の自動引落し契約を新たに結ぶことは基本的には無い」という個人データ利用条件を生成して，この条件を契約中の金融機関に守らせることができるとしよう．他者 Y がその個人のオンライン口座について不正に自動引落し契約を新たに結ぼうと試みた際に，X の個人データ利用条件に照合して異常を検知してアラートが X と金融機関の双方に表示されるような仕組みがあれば，個人 X が不利益を被るような不正な使用を防止することができる．この方法で，2020 年 9 月に報じられたドコモ口座不正利用のようなケースはかなり防ぐことができると予想される．

個人の保有する口座がごく少数であり，条件設定も数える程度であれば，この機能は各金融機関のユーザインタフェースの一つとして用意されれば足りる．ところが，個人の保有する口座数が多く, また, 設定条件も多い場合には，個人側の利用条件を「個人データ利用条件データベース」に書き込んでおき，それがその個人と契約のある複数の口座や金融機関で守られているかをパーソナル AI エージェントが監視する仕組みが有効である．

この場合,「個人データ利用条件データベース」と口座のある金融機関は常に通信することになるだろう．そして，異常を検知したら個人とその金融機関の双方にそれを知らせる仕組みになるであろう．

ある個人が，その口座をどのように使いたいかは，口座保有者本人の意思が反映されるべきである．金融機関側では，いま新規に引落し契約を結ぼうとしている者が，ID を不正に入手した別人であるかどうかは分からない. ID とパスワードが機械的にマッチしてしまった場合，不正な口座の操作が行えてしまう．

これまで金融機関側にとっては，膨大な契約者一人ひとりの口座利用に係る意思など把握しえなかったが，このような「個人データ利用条件データベース」があると，本人の意思に反する利用がなされた時に，すぐさまエラーを検知することができる．このような仕組みは，ある個人が日頃はあまり使用していない金融機関の口座を引き続き向こう何年にも渡り保有し続けようとする時に役立つ．使用頻度の低い口座については頻繁に記帳したりチェックしたりする機会も自ずと乏しくなるので, 上記のパーソナル AI エージェントによるチェック機能は有用である．

さて，上記の例は，個人を代理するパーソナル AI エージェントの視点から事業者とのトラスト構築という視点で述べた．かりにドコモ口座の不正取引のような事案が発生すれば，金融機関側にとっての損失も莫大である．パーソナル AI エージェントによる上記の仕掛けは，このような事案の発生を未然に防ぐ作用が期待される．したがって，事業者側としてはパーソナル AI エージェントによる個人データ利用条件と事業者の提示する条件の不一致をむしろ積極的に利用して，自らの事業の危うい点を予防的に発見するような方向で考えるほうが，長期的には利益が大きい．つまり，事業者と利用者個人の間をパーソナル AI エージェントが仲介することは，事業者と利用者個人の双方にとって win-win な関係を構築できるということになる．

5. おわりに

本稿では，パーソナル AI エージェントを用いて，個人が複雑な外界をガバナンスして，個人データの使用方法を定義したり事業者に守らせたりするという方向性を検討した．パーソナル AI エージェントがいかなるものであるかが見えてくると，ガバナンスのもう一方の側面，すなわち，パーソナル AI エージェント開発事業者をいかにガバナンスするかという議論のポイントも見えてくる．例えば，本研究が示した個人データ利用条件データベースには必ず利用者の情報が蓄積される訳であるから，パーソナル AI エージェント開発事業者は，多くの個人についての「自己の情

Vol.2020-CSEC-91 No.25
Vol.2020-SPT-40 No.25
Vol.2020-EIP-90 No.25
2020/11/26

報に関する意思そのもの」を把握し得ることになる．この場合，現行法による規制や，事業者の自主性に任されるようなガバナンスフレームワークで十分であるのか，ということがより明確に議論することができるようになるだろう．

また，個人データ利用条件データベースは，雛形をダウンロードすることが想定されたが，そのチューンナップの精度が低いと個人の意思が的確に表明されず意図せぬ意思表明をしてしまうことにもなりかねない．このように，個人がパーソナル AI エージェントを用いて外界におけるデータ利用法をガバナンスするということを考えることは，結局，AI をいかにガバナンスするかということや人間社会における技術と責任の問題を考えることにもつながる．

そうした問題を検討していくためにも，まずは，パーソナル AI エージェントの建付けを今後さらに検討して明確にして行く必要がある．

謝辞 本研究は JST-RISTEX-HITE「PATH－AI:人間-AI エコシステムにおけるプライバシー，エージェンシー，トラストの文化を超えた実現方法」の支援を受けた．

参考文献

[1] 崎村夏彦．プライバシーとアイデンティティ～炎上レスでパーソナルデータ連携～．JAPAN IDENTITY & CLOUD SUMMIT (JICS 2014)．2014．http://id.nii.ac.jp/1125/00000097/，https://jics.nii.ac.jp/2014/，(参照 2020-10-28)．

[2] IEEE．Ethically Aligned Design Version 2．2017．https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v2.pdf，(参照 2020-10-28)．

[3] IEEE．Ethically Aligned Design 1e．2019．https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf，(参照 2020-10-28)．

[4] 中川裕志．AI 倫理指針の動向とパーソナル AI エージェント．情報通信政策研究，2020，3 巻 2 号，p.1-24．https://doi.org/10.24798/jicp.3.2_1，(参照 2020-10-28)．

[5] 橋田浩一．分散 PDS と情報銀行-集めないビッグデータによる生活と産業の全体最適化-．情報管理，2017，vol．60，no．4，p.251-260．http://doi.org/10.1241/johokanri.60.25，(参照 2020-10-28)．

[6] 橋田浩一．MyData と PLR．2020．https://www.assemblogue.com/apps/PLRintro.pdf，(参照 2020-10-28)．

[7] 奥 健太．情報推薦システム（Recommender Systems）．人工知能学会．私のブックマーク，No.6 (2013/11)．https://www.ai-gakkai.or.jp/my-bookmark_vol28-no6/，(参照 2020-10-28)．